



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

SERIES DE TIEMPO
CON VALORES DISCRETOS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A:

ERIK ARMANDO VÁZQUEZ RIVERA



TUTOR:
DR. RAMSÉS HUMBERTO MENA CHÁVEZ

2007



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

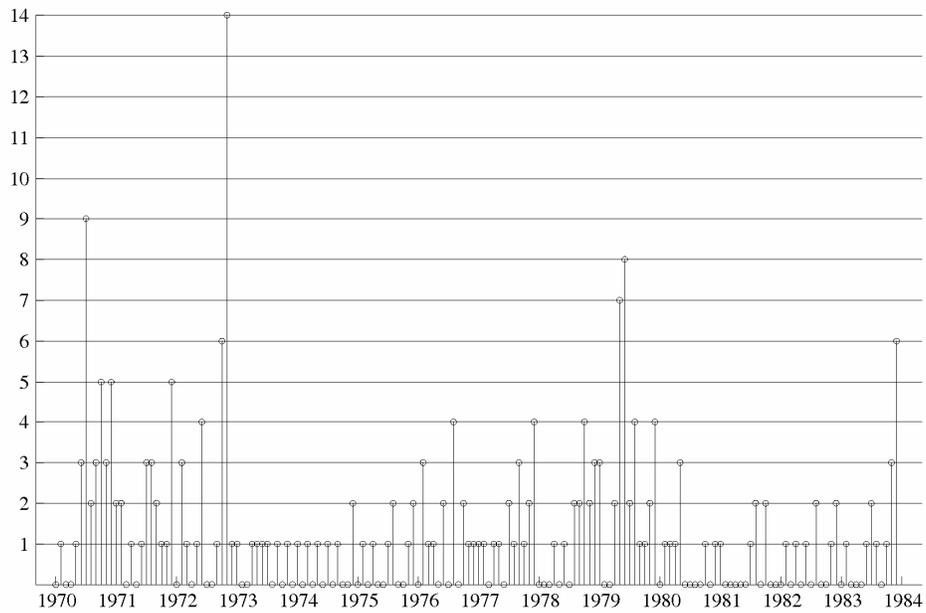
El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Hoja de Datos del Jurado

<p>1. Datos del alumno Vázquez Rivera Erik Armando 57 30 16 54 Universidad Nacional Autonoma de Mexico Facultad de Ciencias Actuaría 099341024</p>
<p>2. Datos del tutor Dr Ramsés Humberto Mena Chávez</p>
<p>3. Datos del sinodal 1 Dr Mogens Bladt Petersen</p>
<p>4. Datos del sinodal 2 Dra Silvia Ruiz Velasco Acosta</p>
<p>5. Datos del sinodal 3 Dr Alberto Contreras Cristán</p>
<p>6. Datos del sinodal 4 Dra Ruth Selene Fuentes García</p>
<p>7. Datos del trabajo escrito Series de tiempo con valores discretos 135 p 2007</p>

*A mis padres:
por su amor y ejemplo.
Como una muestra de gratitud
por apoyarme en los momentos
definitivos de mi vida.*

Series de tiempo con valores discretos



Erik Armando Vázquez Rivera

Agradecimientos

Al Dr. Mena por todo su apoyo y empuje durante la elaboración de este trabajo.

A mis sinodales por sus observaciones y sugerencias para la versión final de esta tesis.

A toda la gente que me motivó a terminar este trabajo, cuyos nombres son más de uno.

A todos aquellos que me muestran su amor y cariño.

A Dios por regalarme la existencia.

*Mil Gracias
de Corazón.*

Índice general

Resumen	v
Agradecimientos	vii
Notación y abreviaturas	ix
1. Introducción	1
1.1. Series de tiempo con valores discretos	1
1.2. Reseña histórica	2
1.3. Sobre esta tesis	3
2. Preliminares	5
2.1. Procesos estocásticos	5
2.2. Series de tiempo	7
2.3. Distribuciones	16
3. Motivación y antecedentes	19
3.1. Modelos con distribución Exponencial	19
3.1.1. EAR(1)	20
3.1.2. EMA(1)	22
3.1.3. EARMA(1,1)	24
3.2. Modelos discretos generados por mezclas	28
3.2.1. DARMA(1, $N + 1$)	28
3.2.2. DAR(p)	34
3.2.3. DARMA(p , $N + 1$)	39
3.2.4. NDARMA(p , N)	41
4. Sobre descomponible-por sí misma	45
4.1. Descomponibilidad-por sí misma	45
4.1.1. GAR(1)	47
4.2. Descomponible-por sí misma discreta	51

4.2.1. Operador de adelgazamiento Binomial	55
5. Modelos basados en adelgazamiento	61
5.1. INAR(1)	61
5.1.1. Poisson	64
5.1.2. Geométrica	65
5.1.3. Binomial Negativa	67
5.1.4. Propiedades generales	70
5.2. Otros AR(1) discretos	73
5.2.1. Binomial Negativa	73
5.2.2. Binomial	80
5.3. MA(1) discretos	83
5.3.1. Poisson	83
5.3.2. Geométrico	85
5.3.3. Binomial	87
5.4. ARMA(1,1) discretos	88
5.4.1. Poisson	88
5.4.2. Geométrico	90
5.4.3. Binomial	93
6. Modelo AR(1) con marginal <i>ED</i>	97
6.1. Modelos de dispersión exponencial	97
6.1.1. Ejemplos	98
6.2. Operador de adelgazamiento de Joe	100
6.2.1. Ejemplos	101
6.2.2. Propiedades generales	106
7. Aplicación	109
7.1. Poliomielitis en U.S. de 1970 a 1983	109
7.1.1. Distribución	110
7.1.2. Modelo	111
7.2. Parámetros	113
7.2.1. Discusión	117
7.3. Estimación	121
7.3.1. Código	122
7.4. Datos	127
8. Conclusiones	129

Resumen

Esta tesis trata de los principales modelos de series de tiempo con valores discretos, primordialmente de orden uno, que aparecieron en la literatura durante los años 1978-1998. Centrándonos de manera exhaustiva en procesos de tipo Autoregresivo como el INteger-valued Autoregressive de orden uno, INAR(1), y Joe (1996); en menor proporción estudiamos procesos con estructuras MA y ARMA, con distintas distribuciones marginales como la Poisson, la Geométrica, la Binomial Negativa y la Binomial. Se ejemplifica la teoría por medio de un caso concreto: usamos la serie de los casos mensuales de poliomielitis reportados por los Centros para el Control y Prevención de Enfermedades -CDC- de los Estados Unidos entre los años 1970-1983 para aplicar el modelo de Joe (1996).

Capítulo 1

Introducción

Una serie de tiempo es una sucesión de observaciones, usualmente, ordenadas en el tiempo. En algunos casos dicho ordenamiento puede ser hecho en otro índice, por ejemplo el espacio. Mientras en muchos problemas estadísticos las observaciones se consideran independientes, en series de tiempo, observaciones sucesivas pueden no serlo.

Existen dos grandes enfoques para el análisis de series de tiempo. El enfoque del dominio de frecuencias y el enfoque del dominio en el tiempo. El primero está basado en la teoría del análisis de Fourier. El segundo tiene su origen en la teoría clásica de correlación, y es el que nosotros adoptamos en este trabajo.

El objetivo principal de la teoría de series de tiempo es el estudio, la descripción y el análisis de la dinámica que gobierna una sucesión de datos aleatorios; se pretende hacer inferencia acerca del mecanismo probabilístico que produce la serie. Este análisis puede tener diversas finalidades, por ejemplo el pronóstico de valores futuros, controlar eventos futuros vía intervención o simplemente una mejor comprensión y explicación de un fenómeno observado.

1.1. Series de tiempo con valores discretos

En las últimas cuatro décadas el estudio de series de tiempo ha recibido mucha atención. La clase más popular de modelos de series de tiempo consiste en los modelos Autoregresivos-Promedios Móviles (ARMA). Éstos son adecuados para modelar sucesiones dependientes estacionarias bajo la suposición de Normalidad, esto es, que la sucesión de innovación es Normal. Sin embargo esta suposición es verdaderamente difícil de cumplir en la práctica. Una situación crítica en la que la suposición de Normalidad no es válida es cuando el fenómeno en cuestión toma valores en un espacio discreto. Regularmente este último tipo de series aparece como el conteo de un evento, objetos o individuos. Por ejemplo el número de pacientes en un hospital, el número de personas en una cola esperando por un servicio, el número de accidentes en una fábrica, el número de artículos defectuosos en muestras sucesivas de una línea de producción, el número de personas reportadas

con cierta enfermedad, durante un lapso de tiempo determinado, un día, un mes, un año, etc.

A pesar de la clara necesidad de modelar y simular este tipo de datos, las cadenas de Markov representaban la única clase general de modelos disponibles para este fin, hasta finales de los setentas. Serios intentos han sido hechos desde entonces para desarrollar modelos para series de tiempo con valores discretos que exhiben una estructura de correlación reconocible.

No tenemos conocimiento de la existencia de un libro propio sobre el tema. Mucho menos en español. La mayor referencia es MacDonald y Zucchini (1997). En la primera parte de este libro los autores presentan una breve reseña de varios de los modelos de los capítulos 2 y 4 de este trabajo, sin embargo no profundizan en éstos, ni tocan los enfoques más generales. Otra referencia de este estilo es el artículo de McKenzie (2003). Se cuenta, sin embargo, con una gran cantidad de trabajos que se han ido publicando a lo largo de los años, varios (no todos) de los cuales nos hemos dado a la tarea de estudiar, reunir e integrar en este trabajo.

1.2. Reseña histórica

A continuación una breve reseña histórica sobre la aparición y rol de los trabajos esenciales que se revisaron para la elaboración de esta tesis:

Los primeros intentos por proporcionar modelos de series de tiempo estacionarios con distribución marginal no-Gaussiana se encuentran en Gaver y Lewis (1980), Lawrance y Lewis (1977), Jacobs y Lewis (1977). Estos trabajos proponen modelos con distribución marginal Exponencial y estructura AR(1), MA(1) y ARMA(1,1) respectivamente. La exposición general de dicho enfoque resultó en un modelo Exponencial Autoregresivo-Promedios Móviles (EARMA(p, q)) presentado por Lawrance y Lewis (1980). Gaver y Lewis (1980) y Lawrance (1982), bajo la misma línea, proponen un proceso AR(1) con distribución marginal Gamma, al que llamarían GAR(1). Mientras que, Lewis (1983) y Lewis *et al.* (1989) abordan igualmente la distribución Gamma pero bajo un enfoque distinto.

Los trabajos de Gaver y Lewis (1980), Lawrance y Lewis (1977), y Jacobs y Lewis (1977) juegan un papel muy importante en el desarrollo de modelos de series de tiempo con valores discretos. Pues, son la pauta que motivaría la aparición de los artículos de Jacobs y Lewis (1978a-c), y además, son el análogo continuo de los modelos con distribución marginal Geométrica presentados por McKenzie (1986).

Uno de los primeros trabajos para series de tiempo con valores en los enteros en la literatura son los modelos Discretos de mezclas Autoregresivos-Promedios Móviles (DARMA), propuestos en Jacobs y Lewis (1978a-c). Éstos modelos son obtenidos por una mezcla probabilística de una sucesión de v.a.'s i.i.d. con distribución discreta. El enfoque Jacobs-Lewis puede ser visto como análogo al enfoque Box-Jenkins en el cual la combinación lineal de v.a.'s continuas es reemplazada por una mezcla probabilística. Como resultado de esto, una realización del proceso, en general, contendrá muchas corridas de un mismo valor. Jacobs y Lewis (1983) conjuntan los tres trabajos anteriores (1978a-c)

en un proceso DARMA($p, N + 1$), e introducen uno nuevo al que llaman NDARMA.

En Gaver y Lewis (1980) se trata ya la cuestión de las distribuciones descomponibles-por sí mismas (S-D, *self-decomposable*) como distribuciones marginales para la solución estacionaria de un proceso AR(1) con marginal positiva. El artículo de Steutel y van Harn (1979) al establecer el concepto de S-D caso discreto, se convierte en la llave que permitiría la creación de un proceso AR(1) con marginal discreta, bastante semejante a su análogo continuo AR(1), no sólo en la forma sino también en la estructura, denominado INAR(1) (INteger-valued Autoregressive). El caso en que la marginal de este proceso es la distribución Geométrica es estudiado por McKenzie (1986). Mientras que, McKenzie (1988b) y Al-Osh y Alzaid (1987), de manera independiente, tratan de la distribución Poisson como la distribución marginal requerida.

Otros modelos como Al-Osh y Alzaid (1991) y Al-Osh y Aly (1992) aparecen años después para procesos con distribución marginal Binomial y Binomial Negativa respectivamente.

Estructuras como la MA(1) y la ARMA(1,1) son tratadas en varios de estos trabajos, con la respectiva distribución marginal. En repetidas ocasiones el apelativo MA y ARMA es más una forma de relacionarlos con los casos continuos, teniendo como referencia la correlación serial de los procesos.

Posteriormente Joe (1996) introduce un proceso que unifica y generaliza los trabajos de Lewis (1983) y Lewis *et al.* (1989), McKenzie (1986), McKenzie (1988b) y Al-Osh y Alzaid (1991) para procesos AR(1) con marginal no-Gaussiana y de manera general con marginal en las familias paramétricas infinitamente-divisibles cerradas bajo convolución. Jørgensen y Song (1998) concretan el trabajo de Joe (1996) en los modelos de dispersión exponencial cerrados bajo convolución (*ED*).

1.3. Sobre esta tesis

Como ya se mencionó esta tesis trata de modelos de series de tiempo con valores discretos. Haciendo un especial énfasis en procesos Autoregresivos de orden uno.

En particular, consiste en la revisión de poco menos de una veintena de artículos sobre el tema y en la aplicación real de al menos uno de los modelos vistos a lo largo del trabajo. Comenzando por los procesos DARMA de Jacobs y Lewis (1978a-c) hasta el proceso AR(1) de Joe (1996) en el contexto de Jørgensen y Song (1998). En la mayoría de los casos los artículos omiten las pruebas y se limitan a la presentación de resultados, por razones de espacio. En este trabajo presentamos y desglosamos las principales características de cada modelo, tales como estacionariedad, correlación serial, distribución conjunta, reversibilidad en el tiempo, esperanza y varianza condicionales, así como las probabilidades de transición para procesos AR(1).

El orden y contenido, de manera general, de este trabajo es conforme al de la aparición de los artículos arriba mencionados.

El segundo capítulo está dedicado a la presentación de una serie de conceptos y resultados elementales de procesos estocásticos y de la teoría de series de tiempo, bajo el enfoque del dominio en el tiempo, que consideramos pueden ayudar a una mejor comprensión del trabajo.

La primera parte del Capítulo 3 trata de los trabajos de Gaver y Lewis (1980), Lawrance y Lewis (1977), Jacobs y Lewis (1977) como el precedente de los procesos DARMA y NDARMA de Jacobs y Lewis (1978a-c; 1983) que son abordados en la segunda parte del mismo.

El Capítulo 4 introduce a la teoría de S-D. Mostramos que las distribuciones S-D son las distribuciones marginales para la solución estacionaria de un proceso AR(1) con valores en \mathbb{R}^+ . En el caso en que el proceso tome valores en \mathbb{N}_0 será necesario reemplazar la multiplicación escalar por el operador de adelgazamiento Binomial, usando el concepto de descomponibilidad por sí misma caso discreto.

El Capítulo 5 trata de modelos de series de tiempo con valores discretos de orden uno basados principalmente en adelgazamiento Binomial. Estudiamos concretamente tres distribuciones: Poisson, Geométrica y Binomial, y las estructuras AR(1), MA(1) y ARMA(1,1) para las tres. Dos modelos son la excepción respecto al uso de este adelgazamiento. Una generalización del adelgazamiento Binomial es usada por Al-Osh y Aly (1992), en tanto que Al-Osh y Alzaid (1991) usan adelgazamiento Hipergeométrico.

El Capítulo 6 introduce los conceptos básicos de la teoría de modelos de dispersión exponencial infinitamente-divisibles (*ED*). Posteriormente integra de la manera más explícita posible los trabajos de Joe (1996) y Jørgensen y Song (1998) en cuanto a procesos AR(1) se refiere, logrando una gran compacidad que se ve reflejada en los Ejemplos de la subsección 6.2.1.

Finalmente el Capítulo 7, presenta la aplicación a datos reales del modelo AR(1) de Joe (1996) con distribución marginal Binomial Negativa, para la obtención de los parámetros de la distribución. Ciertamente varios programas por diversos métodos pueden darnos una estimación de parámetros, pero todos ellos asumen independencia en los datos. En nuestro caso los datos exhiben correlación. Se elabora un código en `OX` que implementa dicha estimación por el método de Máxima Verosimilitud (MV). Una discusión amplia sobre varios posibles ajustes se encuentra en este capítulo. La serie analizada es el número de casos mensuales de poliomielitis en los Estados Unidos de 1970 a 1983.

Confiamos en que el trabajo será una referencia valiosa para aquellos que se están introduciendo al tema por vez primera y del interés de aquellos lectores más familiarizados en el campo. Además, creemos que puede ayudar a fomentar una mayor exploración del tema a nivel de licenciatura.

Septiembre 2006

Capítulo 2

Preliminares

2.1. Procesos estocásticos

Un *espacio medible* (S, ζ) es un conjunto S con una σ -álgebra ζ de subconjuntos de S . Un *espacio de probabilidad* $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio medible (Ω, \mathcal{F}) con una medida de probabilidad \mathbb{P} sobre \mathcal{F} . Sea X una función de Ω a S ($X : \Omega \rightarrow S$), se dice que X es (\mathcal{F}, ζ) -medible si y solo si

$$X^{-1}(A) \in \mathcal{F}, \quad \text{para toda } A \in \zeta.$$

Si S es el conjunto \mathbb{R} de los Reales y, $\zeta = \mathcal{B}(\mathbb{R})$ la clase de conjuntos de Borel en \mathbb{R} , entonces X es llamada *variable aleatoria* (v.a.).

Definición 2.1 (Proceso estocástico) . Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, (S, ζ) un espacio medible arbitrario, y T un conjunto arbitrario. Un *proceso estocástico* sobre $(\Omega, \mathcal{F}, \mathbb{P})$ con *espacio de estados* (S, ζ) y conjunto *índice* T es una función definida sobre $\Omega \times T$ con valores en S , tal que para toda $t \in T$ fija, $X(., t) : \Omega \rightarrow S$ es (\mathcal{F}, ζ) -medible.

Note que si T es el conjunto de los Reales positivos, \mathbb{R}^+ , o el conjunto de los enteros no negativos, \mathbb{N}_0 , y $S = \mathbb{R}$ o $S = \mathbb{Z}$, un proceso estocástico es una familia $\{X(\omega, t); \omega \in \Omega, t \in T\}$ de variables aleatorias.

Por brevedad usualmente un proceso estocástico es denotado por $\{X_t; t \in T\}$, $\{X(t); t \in T\}$, o $\{X_t\}_{t \in T}$.

Frecuentemente el conjunto índice de un proceso estocástico representa el tiempo. Por ello la sucesión $\{X_n; n \in \mathbb{N}_0\}$ es llamada *proceso estocástico a tiempo discreto*, y a la familia $\{X_t; t \geq 0\}$ nos referimos como un *proceso estocástico en tiempo continuo*.

Definición 2.2 (Cadena de Markov) . Una sucesión de variables aleatorias $\{X_n\}_{n=0,1,\dots}$ con valores en S numerable, diremos que es una *cadena de Markov* (a tiempo discreto) si

$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0] = \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n], \quad (2.1)$$

para todo $x_0, x_1, \dots, x_{n+1} \in S$ y n un entero no negativo.

La ecuación (2.1) es llamada la *propiedad de Markov*. Las probabilidades condicionales $p(x_n, x_{n+1}) = \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n]$ son llamadas *probabilidades de transición* de la cadena. Si estas probabilidades no dependen del tiempo, i.e. de n , diremos que la cadena es *homogénea en el tiempo*.

Hablamos de una cadena de Markov de orden r , si dado un $r \geq 1$ fijo, se cumple

$$\begin{aligned} \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0] \\ = \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n, \dots, X_{n+1-r} = x_{n+1-r}], \end{aligned}$$

para todo $x_0, x_1, \dots, x_{n+1} \in S$ y n un entero no negativo.

Esta última es la propiedad de Markov de orden r . La Definición 2.2, es el caso particular en que $r = 1$.

Definición 2.3 Un proceso estocástico $\{X_t; t \in T\}$ se dice que tiene *incrementos independientes* si para cualquier entero $n \geq 1$, y cualquier colección de índices $t_1 < t_2 < \dots < t_n$ se cumple que las variables aleatorias

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}},$$

son independientes.

Definición 2.4 Un proceso estocástico $\{X_t; t \in T\}$ diremos que tiene *incrementos estacionarios* si para cualquier entero $n \geq 1$, $h > 0$ y cualquier colección de índices $t_1 < t_2 < \dots < t_n$, se cumple

$$(X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}) \stackrel{d}{=} (X_{t_2+h} - X_{t_1+h}, \dots, X_{t_n+h} - X_{t_{n-1}+h}).$$

Un ejemplo básico de proceso estocástico a tiempo continuo es el llamado *proceso Poisson*, $N(t)$. $N(t)$ cuenta el número de veces que ocurre un determinado evento, partiendo de un inicio (tiempo cero) al tiempo $t > 0$.

Sea $\{T_n\}_{n=1,2,\dots}$ la sucesión que denota los tiempos en que ocurre el evento y $\{W_n\}_{n=1,2,\dots}$ la sucesión de *tiempos de interocurrencia*, $W_n = T_n - T_{n-1}$, $n = 1, 2, \dots$, con $T_0 = 0$.

Definición 2.5 (Proceso Poisson) . $\{N(t), t \geq 0\}$, es un proceso Poisson (homogéneo) con intensidad λ , si y solo si

1. $N(0) = 0$.
2. Tiene incrementos independientes y estacionarios.
3. $N(t)$ tiene distribución Poisson(λt).

Además, los tiempos de interocurrencia $\{W_n\}_{n=1,2,\dots}$ son variables aleatorias *independientes con distribución exponencial*(λ).

2.2. Series de tiempo

Definición 2.6 (Serie de tiempo) . Una serie de tiempo es un conjunto de observaciones ordenadas en el tiempo.

Si con el conocimiento del pasado, el comportamiento futuro de la serie puede ser determinado se dice que la serie es *determinista*. De otra manera decimos que se trata de una serie *estadística*. Para este último tipo de series, en el mejor de los casos, el conocimiento del pasado puede sólo indicar la estructura probabilística del comportamiento futuro.

Una serie de tiempo estadística puede ser considerada como la realización **única** de un proceso estocástico implícito.

A continuación definimos una serie de funciones propias del enfoque de series de tiempo en el dominio del tiempo.

Definición 2.7 (Función de autocovarianza) . Definimos la función de autocovarianza (acv.f.) $\gamma(t_1, t_2)$ de un proceso $\{X_t; t \in T\}$, como la covarianza entre X_{t_1} y X_{t_2} , esto es

$$\gamma(t_1, t_2) = E\left[(X_{t_1} - \mu_{t_1})(X_{t_2} - \mu_{t_2})\right],$$

con $\mu_t = E[X_t]$ para cualquier t .

La función varianza es un caso especial de la acv.f. con $t_1 = t_2$.

Definición 2.8 (Estacionario fuerte) . Un proceso $\{X_t; t \in T\}$ se dice que es *estrictamente estacionario* o *estacionario fuerte* si para todo entero $n \geq 1$, para cualesquiera t_1, t_2, \dots, t_n elementos de T y para toda h tal que $t_1 + h, t_2 + h, \dots, t_n + h \in T$, la distribución conjunta del vector $(X_{t_1}, X_{t_2}, \dots, X_{t_n})$ es igual a la distribución conjunta del vector $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$.

En particular la definición anterior es válida para $n = 1$, en cuyo caso tenemos que la distribución de X_t es la misma para toda t . En otras palabras, la estructura probabilística de un proceso estrictamente estacionario no cambia al correr el tiempo.

Ésta es una fuerte suposición cuando se modelan datos reales por lo que en la práctica es necesario definir estacionariedad en un sentido menos restrictivo. En lugar de fijar la distribución de todo el proceso, pediremos que sólo los dos primeros momentos de la distribución conjunta sean los mismos ante corrimientos en el tiempo, concretamente.

Definición 2.9 (Estacionario débil) . Un proceso $\{X_t; t \in T\}$ se dice que es *estacionario de segundo orden* o *estacionario débil* si

$$\mu_t = \mu, \quad \text{para todo } t$$

y

$$E[X_{t_1}X_{t_2}] = E[X_{t_1+h}X_{t_2+h}], \quad (2.2)$$

para todo $t_1, t_2, t_1 + h, t_2 + h \in T$.

Algunos autores prefieren enunciar la ecuación (2.2) como $E[X_{t_1}X_{t_2}] = f(|t_1 - t_2|)$, esto es, que $E[X_{t_1}X_{t_2}]$ es una función que depende únicamente de la separación en el tiempo entre las variables X_{t_1}, X_{t_2} .

Sea $\{X_t; t \in T\}$ un proceso estacionario de segundo orden, entonces

$$\gamma(t_1, t_2) = E[X_{t_1}X_{t_2}] - \mu^2 = E[X_{t_1+h}X_{t_2+h}] - \mu^2.$$

Es obvio que en tal caso

$$\gamma(t_1, t_2) = f(|t_1 - t_2|).$$

Para escribir que la acv.f. depende únicamente de la separación en el tiempo entre las variables usamos la notación $\gamma(\tau)$ (función de τ) con $\tau = t_2 - t_1$. En series de tiempo τ comúnmente recibe el nombre de rezago.

En base a esto es más frecuente encontrarse con la siguiente definición.

Definición 2.10 Un proceso $\{X_t; t \in T\}$ se dice que es *estacionario de segundo orden estacionario débil* si su media es constante y su acv.f. depende únicamente del rezago, esto es

$$\mu_t = \mu, \quad \text{para todo } t$$

y

$$\text{Cov}(X_t, X_{t+\tau}) = \gamma(\tau).$$

En particular si $\tau = 0$, se tiene que la varianza del proceso (σ_x^2), al igual que la media, es constante. La definición también implica que ambas, media y varianza, deben ser finitas.

Obsérvese que si $\{X_t\}$ es un proceso estacionario de segundo orden

$$\gamma(\tau) = \text{Cov}(X_t, X_{t+\tau}) = \text{Cov}(X_{t+\tau}, X_t) = \text{Cov}(X_t, X_{t-\tau}) = \gamma(-\tau),$$

por lo tanto, en tal caso, la acv.f. es una función par. Además, la magnitud de la acv.f. depende de las unidades en que el proceso es medido. Por ello, para propósitos de interpretación, es de ayuda estandarizarla.

Definición 2.11 (Función de autocorrelación) . Si $\{X_t; t \in T\}$ es un proceso estacionario de segundo orden definimos la función de autocorrelación (ac.f.) $\rho(\tau)$, como

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)} = \frac{\gamma(\tau)}{\sigma_x^2}.$$

Obviamente $-1 \leq \rho(\tau) \leq 1$, para cualquier valor de τ , y la ac.f. hereda de la acv.f. la simetría par, $\rho(\tau) = \rho(-\tau)$. Por lo que sólo es necesario calcularla para $k \geq 0$ (ó $k \leq 0$).

Por último, ¿cuál es la relación que guardan los dos tipos de estacionariedad vistos arriba?

Estacionariedad fuerte \implies Estacionariedad débil
siempre que los dos primeros momentos existan.

Estacionariedad débil \implies Estacionariedad fuerte
si el proceso es Gaussiano

Definición 2.12 (Proceso Gaussiano) . El comportamiento probabilístico de un proceso estocástico está completamente determinado por todas sus distribuciones finito-dimensionales. Un proceso $\{X_t; t \in T\}$ se dice que es Gaussiano si todas sus distribuciones finito-dimensionales son Gaussianas (Normales).

A continuación definimos una serie de procesos estocásticos usados ampliamente en la modelación de series de tiempo.

Definición 2.13 (MA(q)) . Sea $\{Z_t\}$ una sucesión de v.a.'s i.i.d. con media cero y varianza σ_z^2 . Entonces el proceso $\{X_t\}$ es llamado de *Promedios Móviles* de orden q (MA(q)) si

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \dots + \beta_q Z_{t-q},$$

con $\{\beta_i\}$ constantes. Usualmente -y para nosotros así será- $\beta_0 = 1$.

Proceso	$\gamma(k)$	k	$\rho(k)$
MA(q)	$\sigma_z^2 \sum_{i=0}^{q-k} \beta_i \beta_{i+k}$	$k = 0, 1, \dots, q$	$\sum_{i=0}^{q-k} \beta_i \beta_{i+k} / \sum_{i=0}^q \beta_i^2$
	0	$k > q$	0
	$\gamma(-k)$	$k < 0$	$\rho(-k)$

Cuadro 2.1: Función de autocovarianza y autocorrelación de un proceso MA(q).

Puesto que $\gamma(k)$ no depende de t , y la media es constante, el proceso MA(q) es estacionario de segundo orden. Note además que la ac.f. se corta exactamente en el rezago q . Ésta es una propiedad característica del proceso MA(q).

En particular para el caso $q = 1$, tenemos

MA(1)

$$X_t = Z_t + \beta Z_{t-1}, \quad (2.3)$$

con $\{Z_t\}$ una sucesión de v.a.'s i.i.d. con media cero y varianza σ_z^2 .

Proceso	$\gamma(k)$	k	$\rho(k)$
	$\sigma_z^2 \beta$	$k = \pm 1$	$\beta/(1 + \beta^2)$
MA(1)	$\sigma_z^2(1 + \beta^2)$	$k = 0$	1
	0	e.o.c.	0

Definición 2.14 (AR(p)) . Sea $\{Z_t\}$ una sucesión de v.a.'s i.i.d. con media cero y varianza σ_z^2 . Entonces el proceso $\{X_t\}$ es llamado *Autoregresivo* de orden p (AR(p)) si

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

El Cuadro 2.2 ilustra el sistema de ecuaciones que determina la función de autocovarianza y autocorrelación de un proceso AR(p). Dicho sistema recibe el nombre de *ecuaciones de Yule-Walker*.

Proceso	$\gamma(k)$	k
	σ_x^2	$k = 0$
AR(p)	$\alpha_1 \gamma(k-1) + \alpha_2 \gamma(k-2) + \dots + \alpha_p \gamma(k-p)$	$k > 0$
	$\gamma(-k)$	$k < 0$
Proceso	$\rho(k)$	k
	1	$k = 0$
	$\alpha_1 \rho(k-1) + \alpha_2 \rho(k-2) + \dots + \alpha_p \rho(k-p)$	$k > 0$
	$\rho(-k)$	$k < 0$

Cuadro 2.2: Función de autocovarianza y autocorrelación de un proceso AR(p).

Un resultado benéfico que se sigue de la propiedad de Markov de orden p , es que todas las propiedades de distribución de un proceso estacionario AR(p) pueden ser determinadas a partir de su distribución conjunta de orden $(p+1)$. Conviene saber que el resultado sigue siendo válido para el proceso MA, aún cuando éste no es de Markov (McKenzie 1988a).

AR(1) Un proceso X_t , AR(1), está definido mediante la ecuación

$$X_t = \alpha X_{t-1} + Z_t, \quad t = 0, 1, 2, \dots, \quad (2.4)$$

donde las $\{Z_t\}$ son una sucesión de v.a.'s i.i.d. . El término Z_t en (2.4) suele recibir el apelativo de *error*, *ruido*, *innovación* o *incremento* según el contexto en el que se esté trabajando y las

propiedades que caractericen a éste.

La condición de estacionariedad de segundo orden para el proceso AR(1) es $|\alpha| < 1$. Además en tal caso, X_t puede expresarse como un proceso de Promedios Móviles de orden infinito (MA(∞))

$$X_t = \alpha X_{t-1} + Z_t = \sum_{j=0}^{\infty} \alpha^j Z_{t-j}. \quad (2.5)$$

Proceso	$\gamma(k)$	k	$\rho(k)$
	σ_x^2	$k = 0$	1
AR(1)	$\alpha\gamma(k-1) = \alpha^k \sigma_x^2$	$k \geq 0; \alpha < 1$	α^k
	$\gamma(-k)$	$k < 0$	$\alpha^{ k }$

De maneras muy variadas se ha probado en la literatura que la ac.f. de un proceso estacionario de segundo orden, AR(1), es α^k , $k = 0, 1, \dots$. A continuación presentamos una manera muy general de hacerlo, basada en la estructura lineal de la media condicional. Este resultado será de gran ayuda en el resto de los capítulos.

Proposición 2.15 Sea $\{X_t\}$ un proceso de media constante μ , y tal que $E[X_{t+1}|X_t] = \alpha X_t + (1 - \alpha)\mu$, entonces

$$E[X_{t+k} | X_t] = \alpha^k X_t + (1 - \alpha^k)\mu.$$

Demostración: Por inducción.

Para $k = 1$ se cumple por definición.

Supongamos que se cumple para $k - 1$, por demostrar, que se cumple para k

$$\begin{aligned} E[X_{t+k} | X_t] &= E\left[E[X_{t+k}|X_{t+k-1}] | X_t\right] = E\left[(\alpha X_{t+k-1} + (1 - \alpha)\mu) | X_t\right] \\ &= \alpha E[X_{t+k-1}|X_t] + (1 - \alpha)\mu = \alpha\left(\alpha^{k-1} X_t + (1 - \alpha^{k-1})\mu\right) + (1 - \alpha)\mu \\ &= \alpha^k X_t + (1 - \alpha^k)\mu. \quad \square \end{aligned}$$

Corolario 2.16 Sea $\{X_t\}$ como en la Proposición anterior, entonces

$$E[X_t X_{t+k}] = \alpha^k E[X_t^2] + (1 - \alpha^k)\mu^2.$$

Demostración: Basta observar que

$$E[X_t X_{t+k}] = E\left[X_t E[X_{t+k}|X_t]\right]. \quad \square$$

Finalmente es inmediato el siguiente resultado.

Corolario 2.17 Sea $\{X_t\}$ un proceso con media y varianza constantes, μ y σ_x^2 respectivamente, y esperanza condicional dada por $E[X_{t+1}|X_t] = \alpha X_t + (1 - \alpha)\mu$, entonces

$$\text{Corr}(X_t, X_{t+k}) = \frac{E[X_t X_{t+k}] - E[X_t]E[X_{t+k}]}{\sqrt{\text{Var}(X_t)}\sqrt{\text{Var}(X_{t+k})}} = \alpha^k. \quad \square$$

En particular un proceso estacionario de segundo orden AR(1) cumple con las condiciones del Corolario 2.17, puesto que $E[X_t|X_{t-1}] = \alpha X_{t-1} + E[Z_t]$, y ya que por hipótesis la media del proceso es constante, aplicando esperanza en ambos lados de la ecuación (2.4), se tiene $E[Z_t] = \mu(1 - \alpha)$.

Como hemos mencionado arriba (Cuadro 2.1) la ac.f. de un proceso MA(q) después del rezago q es cero. Proviendo así, la ac.f., información sobre el orden del proceso cuando éste es de Promedios Móviles. Sin embargo si el proceso es un AR(p) la ac.f. por sí sola no es suficiente para determinar el orden de dependencia. Motivados por esto definimos la siguiente función.

Definición 2.18 (Función de autocorrelación parcial) . Definimos la *función de autocorrelación parcial* (pac.f.), Ψ_{kk} , de un proceso $\{X_t; t \in T\}$, como

$$\Psi_{kk} = |\mathbf{P}_k^*| / |\mathbf{P}_k|, \quad k = 1, 2, \dots,$$

donde \mathbf{P}_k es la *matriz de autocorrelación* de $k \times k$,

$$\mathbf{P}_k = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{pmatrix}$$

y \mathbf{P}_k^* es \mathbf{P}_k con la última columna reemplazada por

$$(\rho_1, \rho_2, \dots, \rho_k)'$$

Se puede probar que la pac.f. de un proceso AR(p) se corta en el rezago p . Desarrollando así, ésta última, un papel equivalente al de la ac.f. para procesos MA(q). Véase al respecto el Cuadro 2.3.

Así, por ejemplo, para un AR(1)

$$\Psi_{11} = \rho(1) = \alpha, \quad \text{y} \quad \Psi_{kk} = 0, \quad k > 1.$$

Mientras que para un MA(1)

$$\Psi_{kk} = \frac{(-1)^{k-1} \beta^k (1 - \beta^2)}{1 - \beta^{2(k+1)}}, \quad k \geq 1.$$

Una clase de modelos de gran ayuda para series de tiempo es formada combinando los procesos MA y AR.

Proceso	ac.f.	pac.f.
AR(p)	Decae exponencialmente u oscilando	Cero después del rezago p
MA(q)	Cero después del rezago q	Decae exponencialmente u oscilando

Cuadro 2.3: Comportamiento de la auto y parcial correlación.

Definición 2.19 (ARMA(p, q)) . Un proceso *Autoregresivo-Promedios Móviles*, con p términos AR y q términos MA se dice que es un proceso ARMA(p, q) y está dado por

$$X_t = \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + Z_t + \beta_1 Z_{t-1} + \dots + \beta_q Z_{t-q}, \quad (2.6)$$

con $\{Z_t\}$ una sucesión de v.a.'s i.i.d. de media cero y varianza σ_z^2 .

Su acv.f. viene dada por la ecuación

$$\begin{aligned} \gamma(k) = & \alpha_1 \gamma(k-1) + \alpha_2 \gamma(k-2) + \dots + \alpha_p \gamma(k-p) + \gamma_{zx}(k) + \beta_1 \gamma_{zx}(k-1) + \\ & \dots + \beta_q \gamma_{zx}(k-q), \end{aligned} \quad (2.7)$$

donde

$$\gamma_{zx}(j) = \text{Cov}(Z_t, X_{t-j}) = E[Z_t X_{t-j}]$$

y naturalmente

$$\gamma_{zx}(j) = 0, \quad j > 0,$$

puesto que X_t es independiente de cualquier Z posterior. Entonces para $k > q$, la ecuación (2.7) se reduce a

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

que no involucra los parámetros del proceso MA. Por lo tanto, después del rezago q , la acv.f. se comporta como la de un proceso AR(p).

Una forma más práctica de escribir el modelo (2.6) es

$$b(B)X_t = a(B)Z_t, \quad (2.8)$$

donde B denota el *operador de rezago*, el cual está definido como

$$B^k X_t = X_{t-k}, \quad k = \pm 1, \pm 2 \dots$$

y, $a(\cdot)$ y $b(\cdot)$ son polinomios definidos como

$$b(w) = 1 - b_1 w - \dots - b_p w^p, \quad a(w) = 1 + a_1 w + \dots + a_q w^q.$$

Un proceso ARMA(p, q) dado por (2.8) es estacionario si $b(w) \neq 0$ para todo número complejo w tal que $|w| \leq 1$.

Si $\{Z_t\}$ es una sucesión de v.a.'s i.i.d. con distribución Normal($0, \sigma^2$) y $b(w) \neq 0$ para todo número complejo w tal que $|w| \leq 1$, $\{X_t\}$ definido por (2.8) es un proceso estrictamente estacionario Gaussiano.

Definición 2.20 (Invertibilidad) . Decimos que el proceso ARMA(p, q) es *invertible* si $a(w) \neq 0$ para todo número complejo w con $|w| \leq 1$.

La invertibilidad de un proceso implica que éste puede ser expresado en la forma de un proceso Autoregresivo, posiblemente de orden infinito, cuyos coeficientes forman una suma convergente. Es claro que un proceso AR es invertible por definición, mientras que un proceso MA sólo lo será bajo las condiciones arriba mencionadas.

Lo anterior puede resumirse en el siguiente cuadro

Proceso	Estacionariedad	Invertibilidad
AR(p)	Si $b(w) \neq 0$ para $ w \leq 1$	Por definición
MA(q)	Por definición	Si $a(w) \neq 0$ para $ w \leq 1$

Es de particular interés para nosotros el proceso ARMA con $p = 1 = q$.

ARMA(1,1)

$$X_t = \alpha X_{t-1} + Z_t + \beta Z_{t-1}, \quad (2.9)$$

con $\{Z_t\}$ una sucesión de v.a.'s i.i.d. de media cero y varianza σ_z^2 .

Usando (2.7)

$$\gamma(k) = \alpha\gamma(k-1) + \gamma_{zx}(k) + \beta\gamma_{zx}(k-1),$$

entonces

$$\gamma(0) = \alpha\gamma(1) + \sigma_z^2 + \beta\gamma_{zx}(-1), \quad (2.10)$$

$$\gamma(1) = \alpha\gamma(0) + \beta\sigma_z^2 \quad (2.11)$$

y

$$\gamma(k) = \alpha\gamma(k-1), \quad k > 1. \quad (2.12)$$

Multiplicando por Z_{t-1} , (2.9) y tomando esperanza

$$\gamma_{zx}(-1) = \alpha\sigma_z^2 + \beta\sigma_z^2.$$

Sustituyendo esto en (2.10) y resolviendo el sistema de ecuaciones tenemos

$$\begin{aligned}\gamma(0) &= \frac{(1 + 2\beta\alpha + \beta^2)}{1 - \alpha^2} \sigma_z^2, \\ \gamma(1) &= \frac{(\alpha + \beta)(1 + \beta\alpha)}{1 - \alpha^2} \sigma_z^2\end{aligned}$$

y el resto pueden ser determinadas por medio de la ec. (2.12).

Entonces la ac.f. viene dada por

$$\rho(k) = \alpha^{k-1} \frac{(\alpha + \beta)(1 + \beta\alpha)}{(1 + 2\beta\alpha + \beta^2)}, \quad k = 1, 2, \dots \quad \square$$

Ya que un modelo ARMA invertible tiene una representación AR infinita, la pac.f. no se corta y persiste.

Por último analizamos el concepto de *reversibilidad en el tiempo* para procesos con distribución marginal positiva. Sólo para recordar,

Definición 2.21 (Transformada de Laplace-Stieltjes) . Sea X una v.a.. Su transformada de Laplace-Stieltjes está dada como la función

$\Phi_X: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ con,

$$\Phi_X(t) := E[e^{-tX}] = \int_{\mathbb{R}^+} e^{-tx} dF_X(x).$$

La definición para el caso conjunto es análoga.

Definición 2.22 (Reversible en el tiempo) . Una serie de tiempo estacionaria $\{X_t\}$ es *reversible en el tiempo* si la distribución conjunta del vector (X_1, X_2, \dots, X_k) es la misma que la del vector $(X_k, X_{k-1}, \dots, X_1)$, para toda $k \geq 2$.

Suponga que está tratando con un proceso $\{X_t\}$ con distribución marginal positiva. Una forma muy simple de comprobar la reversibilidad en el tiempo es usando la transformada de Laplace de la distribución conjunta de varias X_i 's consecutivas. La condición de reversibilidad viene dada por

$$\begin{aligned}\Phi_{X_1, X_2, \dots, X_n}(s_1, s_2, \dots, s_n) &= \Phi_{X_n, X_{n-1}, \dots, X_1}(s_1, s_2, \dots, s_n) \\ &= \Phi_{X_1, X_2, \dots, X_n}(s_n, s_{n-1}, \dots, s_1),\end{aligned}$$

para cualquier s_1, s_2, \dots, s_n , no negativos, y $n \geq 2$.

Así, si la transformada de Laplace conjunta es simétrica entonces el proceso es reversible en el tiempo y viceversa.

Más aún, para un proceso AR(p) (MA(q)) basta observar la distribución conjunta de $(p+1)$ ($(q+1)$) observaciones consecutivas para determinar si el proceso es reversible en el tiempo.

2.3. Distribuciones

Las distribuciones que usaremos a lo largo de este trabajo son

Bernoulli(p)

$$\begin{aligned} f_X(x) &= p^x(1-p)^{1-x}, & x &= 0, 1, \\ & & 0 < p < 1. \\ E[X] &= p, & \text{Var}(X) &= p(1-p). \\ \Phi_X(t) &= pe^{-t} + q, & \text{con } q &= 1-p. \end{aligned}$$

Binomial(n, p)

$$\begin{aligned} f_X(x) &= \binom{n}{x} p^x (1-p)^{n-x}, & x &= 0, 1, \dots, n, \\ & & n &= 1, 2, \dots, \\ & & 0 < p < 1. \\ E[X] &= np, & \text{Var}(X) &= np(1-p). \\ \Phi_X(t) &= (pe^{-t} + q)^n, & \text{con } q &= 1-p. \end{aligned}$$

Beta-Binomial(α, β, n)

$$\begin{aligned} f_X(x) &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha + \beta + n)} \binom{n}{x} \Gamma(\alpha + x)\Gamma(\beta + n - x), \\ & & x &= 0, 1, \dots, n, \\ & & n &= 1, 2, \dots, \\ & & \alpha &> 0, \\ & & \beta &> 0. \\ E[X] &= n \frac{\alpha}{\alpha + \beta}, & \text{Var}(X) &= \frac{n\alpha\beta}{(\alpha + \beta)^2} \frac{\alpha + \beta + n}{\alpha + \beta + 1}. \end{aligned}$$

Hipergeométrica(N, M, n)

$$\begin{aligned} f_X(x) &= \frac{\binom{N-M}{n-x} \binom{M}{x}}{\binom{N}{n}} = \frac{\binom{n}{x} \binom{N-n}{M-x}}{\binom{N}{M}}, \\ & & x &= a, a+1, \dots, b, \\ & & a &= \max\{0, M - (N - n)\}, \\ & & b &= \min\{n, M\}, \\ & & N &= 1, 2, \dots, \\ & & M &= 1, 2, \dots, N, \\ & & n &= 1, 2, \dots, N. \\ E[X] &= n \frac{M}{N}, & \text{Var}(X) &= \frac{N-n}{N-1} npq, & \text{con } p &= \frac{M}{N}, \text{ y } q = 1-p. \end{aligned}$$

Binomial Negativa(r, p)

$$f_X(x) = \binom{r+x-1}{r-1} (1-p)^x p^r, \quad \begin{array}{l} x = 0, 1, \dots, \\ r > 0, \\ 0 < p < 1. \end{array}$$

$$E[X] = r \frac{1-p}{p}, \quad \text{Var}(X) = r \frac{1-p}{p^2}.$$

$$\Phi_X(t) = \left(\frac{p}{1 - qe^{-t}} \right)^r, \quad \text{con } q = 1 - p.$$

Geométrica(p)

$$f_X(x) = (1-p)^x p, \quad \begin{array}{l} x = 0, 1, \dots, \\ 0 < p < 1. \end{array}$$

$$E[X] = \frac{1-p}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2}.$$

$$\Phi_X(t) = \frac{p}{1 - qe^{-t}}, \quad \text{con } q = 1 - p.$$

Caso particular de la Binomial Negativa con $r = 1$.

Poisson(λ)

$$f_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad \begin{array}{l} x = 0, 1, 2, \dots, \\ \lambda > 0. \end{array}$$

$$E[X] = \lambda, \quad \text{Var}(X) = \lambda.$$

$$\Phi_X(t) = e^{\lambda(e^{-t}-1)}.$$

Beta(α, β)

$$f_X(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad \begin{array}{l} 0 < x < 1, \\ \alpha > 0, \beta > 0. \end{array}$$

$$E[X] = \frac{\alpha}{\alpha + \beta}, \quad \text{Var}(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

Uniforme(a, b)

$$f_X(x) = \frac{1}{b-a}, \quad \begin{array}{l} a < x < b, \\ b > a. \end{array}$$

$$E[X] = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}.$$

$$\Phi_X(t) = \frac{e^{-tb} - e^{-ta}}{-t(b-a)}.$$

Gamma(k, λ)

$$f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}, \quad \begin{array}{l} x > 0, \\ k > 0, \text{ y } \lambda > 0. \end{array}$$

$$E[X] = \frac{k}{\lambda}, \quad \text{Var}(X) = \frac{k}{\lambda^2}.$$

$$\Phi_X(t) = \left(\frac{\lambda}{\lambda+t} \right)^k.$$

Exponencial(λ)

$$\begin{aligned}
 f_X(x) &= \lambda e^{-\lambda x}, & x > 0, \\
 & & \lambda > 0. \\
 E[X] &= \frac{1}{\lambda}, & \text{Var}(X) &= \frac{1}{\lambda^2}. \\
 \Phi_X(t) &= \frac{\lambda}{\lambda+t}.
 \end{aligned}$$

Caso particular de la Gamma con $k = 1$.

Normal(μ, σ^2)

$$\begin{aligned}
 f_X(x) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, & -\infty < x < +\infty, \\
 & & -\infty < \mu < +\infty, \\
 & & \sigma > 0. \\
 E[X] &= \mu, & \text{Var}(X) &= \sigma^2. \\
 \Phi_X(t) &= e^{-t\mu + \frac{1}{2}t^2\sigma^2}.
 \end{aligned}$$

Degenerada(δ)

$$\begin{aligned}
 f_X(x) &= \begin{cases} 1 & \text{si } x = \delta \\ 0 & \text{e.o.c.} \end{cases} & \text{con } \delta \in \mathbb{R}. \\
 E[X] &= \delta, & \text{Var}(X) &= 0. \\
 \Phi_X(t) &= e^{-t\delta}.
 \end{aligned}$$

Teorema 2.23 (Transformación de una variable aleatoria) . Sea X una v.a. que toma valores en $(a, b) \subseteq \mathbb{R}$, y función de densidad f_X . Sea $\varphi : (a, b) \rightarrow \mathbb{R}$, una función continua estrictamente creciente o decreciente, y con inversa diferenciable φ^{-1} . Entonces la v.a. $Y = \varphi(X)$ toma valores en $\varphi(a, b)$ y tiene función de densidad

$$f_Y(y) = \begin{cases} f_X(\varphi^{-1}(y)) \left| \frac{d}{dy} \varphi^{-1}(y) \right| & \text{si } y \in \varphi(a, b), \\ 0 & \text{e.o.c.} \end{cases}$$

Capítulo 3

Motivación y antecedentes

Este capítulo se encuentra dividido en dos secciones. En la primera estudiamos modelos de series de tiempo con distribución marginal Exponencial. El primer modelo que estudiamos es el modelo EAR(1) (Exponencial Autoregresivo de orden 1) presentado en Gaver y Lewis (1980), aunque como Lawrance y Lewis (1980, pg. 151) nos informan, el trabajo fue desarrollado desde 1974. La innovación de este proceso es una *mezcla* de una componente discreta en cero y una v.a. Exponencial. Sería de este elemento de mezcla del que se valdrían Lawrance y Lewis (1977) y Jacobs y Lewis (1977) para la creación de modelos con distribución marginal Exponencial MA(1) y ARMA(1,1) respectivamente. Dicho concepto también fue utilizado por Jacobs y Lewis (1978a-c; 1983) para proveer el primer intento real de una clase de modelos para series de tiempo con valores discretos. Sus dos clases de modelos, DARMA y NDARMA, son formadas por una mezcla probabilística de v.a.'s discretas i.i.d. con la distribución requerida para el proceso. Como los autores prueban, un proceso de este tipo contendrá, en general, muchas corridas de un mismo valor, sección 2.

Recordatorio: Una v.a. W es una *mezcla* de las v.a.'s Z_1, Z_2, \dots, Z_n , si W es una selección aleatoria de una y sólo una de éstas.

3.1. Modelos con distribución Exponencial

Las motivaciones que dieron origen a estos modelos son:

1. Como alternativa a la teoría de series de tiempo con distribución Gaussiana.
2. Como modelos con variables aleatorias correlacionadas positivamente y distribución marginal Exponencial.
3. Y principalmente como modelos de procesos puntuales con los cuales analizar series de eventos que no son Poisson, i.e., procesos -puntuales- de conteo que tiene a la sucesión $\{X_n\}$

como los tiempos de interocurrencia. A diferencia de la Definición 2.5, estos tiempos **no** son independientes.

3.1.1. EAR(1)

Consideremos el modelo AR(1)

$$X_n = \rho X_{n-1} + \varepsilon_n, \quad n = 0, 1, 2, \dots, \quad (3.1)$$

donde $\{\varepsilon_n\}$ es una sucesión de v.a.'s i.i.d.

¿Dada una distribución para las $\{X_n\}$ la ecuación autoregresiva (3.1) tiene solución?, esto es, ¿existe un proceso $\{\varepsilon_n\}$, tal que la sucesión de v.a.'s $\{X_n\}$ tenga la distribución marginal deseada?

Puesto que X_{n-1} es independiente de ε_n la transformada de Laplace, $\Phi_{X_n}(s)$, de la distribución de X_n es:

$$\begin{aligned} \Phi_{X_n}(s) &= E[\exp(-sX_n)] && s > 0 \\ &= E[\exp\{-s(\rho X_{n-1} + \varepsilon_n)\}] = \Phi_{X_{n-1}}(\rho s)\Phi_{\varepsilon_n}(s). \end{aligned}$$

Por lo tanto

$$\Phi_{\varepsilon_n}(s) = \frac{\Phi_{X_n}(s)}{\Phi_{X_{n-1}}(\rho s)}. \quad (3.2)$$

Supongamos que X_n es un proceso marginalmente estacionario, entonces

$$\Phi_{\varepsilon}(s) = \frac{\Phi_X(s)}{\Phi_X(\rho s)}. \quad (3.3)$$

Hasta ahora la única restricción implícita que hemos hecho sobre ρ , es $|\rho| < 1$, pero si requerimos que X_n sea una v.a. positiva, habremos de pedir que $0 \leq \rho < 1$. La razón es que si ρ es negativa, ρX_{n-1} también lo es, y entonces necesitamos que el término ε_n (que es independiente de X_{n-1}) haga positiva a X_n , sin embargo con tales condiciones ($\rho < 0$) la ec. (3.2) no tiene solución.

En particular si requerimos que las X_n 's tengan distribución Exponencial (λ), $\Phi_X(s) = \frac{\lambda}{\lambda+s}$. Sustituyendo en (3.2)

$$\begin{aligned} \Phi_{\varepsilon}(s) &= \frac{\lambda}{\lambda+s} \frac{\lambda + \rho s}{\lambda} = \frac{\lambda + \rho s}{\lambda + s} \\ &= \rho + (1 - \rho) \frac{\lambda}{\lambda + s}. \end{aligned} \quad (3.4)$$

Esta última ecuación es la transformada de Laplace de una v.a. no negativa que con probabilidad ρ toma el valor cero, y con $(1 - \rho)$ es una v.a. Exponencial(λ).

Si queremos entonces que el modelo (3.1) tenga distribución marginal Exponencial(λ), éste estará dado por

$$\begin{aligned} X_n &= \rho X_{n-1} + \varepsilon_n \\ &= \begin{cases} \rho X_{n-1} & \text{c.p. } \rho, \\ \rho X_{n-1} + G_n & \text{c.p. } (1 - \rho), \end{cases} \end{aligned} \quad (3.5)$$

donde $\{G_n\}$ es una sucesión de v.a.'s i.i.d. Exponenciales(λ).

Este modelo recibe el nombre de EAR(1) (Exponencial Autoregresivo de orden 1).

Varios aspectos de este modelo deben resaltarse:

- ε_n es una mezcla de una componente discreta en cero y una v.a. Exponencial.
- G_n no depende del valor de ρ .
- Puesto que la esperanza de una mezcla es la mezcla de las esperanzas, resulta fácil de obtener

$$E[\varepsilon_n] = (1 - \rho) \frac{1}{\lambda}, \quad \text{Var}(\varepsilon_n) = (1 - \rho^2) \frac{1}{\lambda^2}.$$

Estacionariedad El proceso es estacionario si tomamos $X_0 = G_0$, (i.e. con distribución Exponencial(λ) independiente del resto de las $\{G_n; n \geq 1\}$).

La prueba es obvia

$$\begin{aligned} \Phi_{X_n}(s) &= E[\exp(-sX_n)] \\ &= E[\exp(-s\rho X_{n-1})]\rho + E[\exp(-s\rho X_{n-1} - sG_n)](1 - \rho) \\ &= \frac{\lambda}{\lambda + \rho s} \rho + \frac{\lambda}{\lambda + \rho s} \frac{\lambda}{\lambda + s} (1 - \rho) = \frac{\lambda}{\lambda + s}, \quad n = 0, 1, \dots \end{aligned} \quad (3.6)$$

Correlación serial Obsérvese que $E[X_{n+1}|X_n] = \rho X_n + (1 - \rho)\mu$, con $\mu = \frac{1}{\lambda} = E[G_n] = E[X_n]$, por lo tanto haciendo uso del Corolario 2.17 se tiene

$$\rho(k) = \text{Corr}(X_n, X_{n+k}) = \rho^k, \quad 0 \leq \rho < 1. \quad (3.7)$$

Observemos que ésta es siempre positiva.

Distribución conjunta La transformada de Laplace de la distribución conjunta de X_n y X_{n+1} es

$$\begin{aligned} \Phi_{X_n, X_{n+1}}(s_1, s_2) &= E[\exp(-s_1 X_n - s_2 X_{n+1})] \\ &= E[\exp\{-(s_1 + \rho s_2) X_n - s_2 \varepsilon_{n+1}\}] \\ &= E[\exp\{-(s_1 + \rho s_2) X_n\}] \Phi_{\varepsilon_{n+1}}(s_2) \\ &= \frac{\lambda}{\lambda + s_1 + \rho s_2} \left(\frac{\lambda + \rho s_2}{\lambda + s_2} \right) \end{aligned} \quad (3.8)$$

$$\begin{aligned}
&= \frac{\lambda}{\lambda + s_1 + \rho s_2} \left(\rho + (1 - \rho) \frac{\lambda}{\lambda + s_2} \right) \\
&= \rho \frac{\lambda}{\lambda + s_1 + \rho s_2} + (1 - \rho) \left(\frac{\lambda}{\lambda + s_1 + \rho s_2} \right) \left(\frac{\lambda}{\lambda + s_2} \right). \tag{3.9}
\end{aligned}$$

La distribución bivariada de (X_n, X_{n+1}) tiene una componente singular con probabilidad ρ , el primer término en (3.9), que es la transformada de Laplace entre X_n y ρX_n , y una componente conjunta con probabilidad $(1 - \rho)$, el segundo término en (3.9), que es la transformada de Laplace entre X_n y $\rho X_n + G_n$.

Reversible en el tiempo En nuestro caso, el proceso EAR(1) no es reversible en el tiempo puesto que (3.9) es no simétrica en s_1 y s_2 .

Sólo por recordar, el modelo AR(1) Gaussiano es reversible en el tiempo.

Esperanza y varianza condicionales Consideremos X_n y X_{n+1} . De (3.5) la regresión de X_{n+1} sobre $X_n = x$, es lineal en x :

$$E[X_{n+1}|X_n = x] = \rho E[\rho x] + (1 - \rho) E[\rho x + G_{n+1}] = E[\rho x] + (1 - \rho) E[G_{n+1}],$$

simplificando

$$E[X_{n+1}|X_n = x] = \rho x + (1 - \rho) \frac{1}{\lambda}. \tag{3.10}$$

Además

$$\begin{aligned}
E[X_{n+1}^2|X_n = x] &= \rho E[(\rho x)^2] + (1 - \rho) E[(\rho x + G_{n+1})^2] \\
&= \rho^3 x^2 + \left[\rho^2 x^2 + 2\rho x \frac{1}{\lambda} + \frac{2}{\lambda^2} \right] (1 - \rho).
\end{aligned}$$

Quedándonos que

$$\text{Var}(X_{n+1}|X_n = x) = \frac{(1 - \rho^2)}{\lambda^2}, \tag{3.11}$$

la cual es una constante igual a la varianza de la innovación.

3.1.2. EMA(1)

Sea $\{\varepsilon_n\}$ una sucesión de v.a.'s i.i.d. Exponenciales(λ). El proceso EMA(1) (Exponencial de Promedios Móviles de orden 1), $\{X_n\}$, está definido como

$$X_n = \begin{cases} \beta \varepsilon_n & \text{c.p. } \beta, \\ \beta \varepsilon_n + \varepsilon_{n-1} & \text{c.p. } (1 - \beta), \end{cases} \tag{3.12}$$

$$0 \leq \beta \leq 1, \quad n = 0, \pm 1, \pm 2, \dots$$

Estacionariedad El proceso (3.12) tiene distribución marginal Exponencial.

Demostración:

$$\begin{aligned}\Phi_{X_n}(s) &= E[\exp(-sX_n)] = E[\exp(-s\beta\varepsilon_n)]\beta + E[\exp(-s\beta\varepsilon_n - s\varepsilon_{n-1})](1-\beta) \\ &= \frac{\lambda}{\lambda + \beta s}\beta + \frac{\lambda}{\lambda + \beta s}\frac{\lambda}{\lambda + s}(1-\beta) = \frac{\lambda}{\lambda + s}.\end{aligned}\quad (3.13)$$

Compare esta última ecuación con (3.6).

Correlación Serial La distribución de $X_n X_{n+1}$ es

$$\left\{ \begin{array}{ll} \beta^2 \varepsilon_n \varepsilon_{n+1} & \text{c.p. } \beta^2, \\ \beta^2 \varepsilon_n \varepsilon_{n+1} + \beta \varepsilon_n^2 & \text{c.p. } \beta(1-\beta), \\ \beta^2 \varepsilon_n \varepsilon_{n+1} + \beta \varepsilon_{n+1} \varepsilon_{n-1} & \text{c.p. } \beta(1-\beta), \\ \beta^2 \varepsilon_n \varepsilon_{n+1} + \beta \varepsilon_n^2 + \beta \varepsilon_{n+1} \varepsilon_{n-1} + \varepsilon_n \varepsilon_{n-1} & \text{c.p. } (1-\beta)^2. \end{array} \right.$$

Su esperanza es fácilmente calculada. Agrupando términos semejantes se tiene

$$\begin{aligned}E[X_n X_{n+1}] &= E[\beta^2 \varepsilon_n \varepsilon_{n+1}] + (1-\beta)E[\beta \varepsilon_n^2] + (1-\beta)E[\beta \varepsilon_{n-1} \varepsilon_{n+1}] + (1-\beta)^2 E[\varepsilon_{n-1} \varepsilon_n] \\ &= \beta^2 \frac{1}{\lambda^2} + \beta(1-\beta) \frac{2}{\lambda^2} + \beta(1-\beta) \frac{1}{\lambda^2} + (1-\beta)^2 \frac{1}{\lambda^2}.\end{aligned}$$

Simplificando

$$E[X_n X_{n+1}] = \frac{1}{\lambda^2} + \frac{\beta - \beta^2}{\lambda^2}.$$

Por lo tanto

$$\rho(1) = \text{Corr}(X_n, X_{n+1}) = \beta(1-\beta).\quad (3.14)$$

Se puede observar directamente de la definición del modelo que las correlaciones de orden más alto son cero.

La correlación de orden uno, (3.14), limita en gran medida el modelo pues además de no poder tomar valores negativos está acotada por 0.25.

Distribución conjunta De los cálculos de arriba, la distribución conjunta de X_n y X_{n+1} es la mezcla de las distribuciones conjuntas

$$\begin{aligned}(\beta\varepsilon_n, \beta\varepsilon_{n+1}) & \text{ c.p. } \beta^2, \\ (\beta\varepsilon_n, \beta\varepsilon_{n+1} + \varepsilon_n) & \text{ c.p. } \beta(1-\beta), \\ (\beta\varepsilon_n + \varepsilon_{n-1}, \beta\varepsilon_{n+1}) & \text{ c.p. } \beta(1-\beta), \\ (\beta\varepsilon_n + \varepsilon_{n-1}, \beta\varepsilon_{n+1} + \varepsilon_n) & \text{ c.p. } (1-\beta)^2.\end{aligned}$$

Entonces la transformada de Laplace conjunta de X_n y X_{n+1} se calcula mediante

$$\begin{aligned}\Phi_{X_n, X_{n+1}}(s_1, s_2) &= E[\exp(-s_1 X_n - s_2 X_{n+1})] \\ &= \beta^2 E[\exp(-s_1 \beta \varepsilon_n - s_2 \beta \varepsilon_{n+1})] \\ &\quad + \beta(1 - \beta) E[\exp(-s_1 \beta \varepsilon_n - s_2 \{\beta \varepsilon_{n+1} + \varepsilon_n\})] \\ &\quad + \beta(1 - \beta) E[\exp(-s_1 \{\beta \varepsilon_n + \varepsilon_{n-1}\} - s_2 \beta \varepsilon_{n+1})] \\ &\quad + (1 - \beta)^2 E[\exp(-s_1 \{\beta \varepsilon_n + \varepsilon_{n-1}\} - s_2 \{\beta \varepsilon_{n+1} + \varepsilon_n\})].\end{aligned}$$

Sea $\psi(s)$ la transformada de Laplace de la f.d.p. de las $\{\varepsilon_n\}$, en este caso $\psi(s) = \frac{\lambda}{\lambda + s}$. Haciendo uso de esta notación nos queda

$$\begin{aligned}\Phi_{X_n, X_{n+1}}(s_1, s_2) &= \psi(\beta s_2) \left[\beta^2 \psi(\beta s_1) + \beta(1 - \beta) \psi(s_2 + \beta s_1) + \beta(1 - \beta) \psi(\beta s_1) \psi(s_1) \right. \\ &\quad \left. + (1 - \beta)^2 \psi(s_2 + \beta s_1) \psi(s_1) \right] \\ &= \psi(\beta s_2) \left[\beta \psi(\beta s_1) + (1 - \beta) \psi(s_2 + \beta s_1) \right] \left[\beta + (1 - \beta) \psi(s_1) \right] \\ &= \psi(\beta s_2) \frac{\lambda^2 + \lambda \beta s_1 + \lambda \beta s_2}{(\lambda + \beta s_1)(\lambda + \beta s_1 + s_2)} \frac{\lambda + \beta s_1}{\lambda + s_1} \\ &= \frac{\lambda^2 (\lambda + \beta s_1 + \beta s_2)}{(\lambda + \beta s_2)(\lambda + \beta s_1 + s_2)(\lambda + s_1)}.\end{aligned}\tag{3.15}$$

Reversible en el tiempo Aunque como Jacobs y Lewis (1978a, pg. 100) aclaran, la reversibilidad en el tiempo de este proceso no puede ser determinada a partir de las propiedades de segundo orden, el proceso es no reversible en el tiempo.

3.1.3. EARMA(1,1)

Mediante la combinación de los dos modelos anteriores, se obtiene el proceso EARMA(1,1) (Exponencial Autoregresivo-Promedios Móviles, ambos de orden 1). El cual está definido como

$$\begin{aligned}X_n &= \begin{cases} \beta \varepsilon_n & \text{c.p. } \beta, \\ \beta \varepsilon_n + A_{n-1} & \text{c.p. } (1 - \beta), \end{cases} \\ &0 \leq \beta \leq 1, \quad n = 1, 2, \dots,\end{aligned}\tag{3.16}$$

con

$$\begin{aligned}A_n &= \begin{cases} \rho A_{n-1} & \text{c.p. } \rho, \\ \rho A_{n-1} + \varepsilon_n & \text{c.p. } (1 - \rho), \end{cases} \\ &0 \leq \rho < 1, \quad n = 1, 2, \dots.\end{aligned}\tag{3.17}$$

Análogamente $\{\varepsilon_n\}$ es una sucesión de v.a.'s i.i.d. Exponenciales(λ).

Estacionariedad Comparando (3.17) con (3.5), se observa que $\{A_n\}$ es un proceso EAR(1) y por lo tanto A_n tiene distribución marginal Exponencial(λ). Por lo tanto (3.16) tiene la misma estructura que (3.12) y entonces X_n se distribuye Exponencial(λ) para cada n .

Del estudio hecho a los dos modelos anteriores sabemos que el enunciado anterior se cumple siempre que $A_0 = \varepsilon_0$, i.e. A_0 se distribuya de manera Exponencial(λ) independiente del resto de las $\{\varepsilon_n; n = 1, 2, \dots\}$.

Correlación serial La distribución de $X_n X_{n+k}$ viene dada por

$$\begin{cases} \beta^2 \varepsilon_n \varepsilon_{n+k} & \text{c.p. } \beta^2, \\ \beta^2 \varepsilon_n \varepsilon_{n+k} + \beta \varepsilon_n A_{n+k-1} & \text{c.p. } \beta(1 - \beta), \\ \beta^2 \varepsilon_n \varepsilon_{n+k} + \beta \varepsilon_{n+k} A_{n-1} & \text{c.p. } \beta(1 - \beta), \\ \beta^2 \varepsilon_n \varepsilon_{n+k} + \beta \varepsilon_n A_{n+k-1} + \beta \varepsilon_{n+k} A_{n-1} + A_{n+k-1} A_{n-1} & \text{c.p. } (1 - \beta)^2. \end{cases}$$

Sea $k \neq 0$. En tal caso, agrupando términos, la esperanza viene dada por:

$$\begin{aligned} E[X_n X_{n+k}] &= E[\beta^2 \varepsilon_n \varepsilon_{n+k}] + (1 - \beta) \left(E[\beta \varepsilon_n A_{n+k-1}] + E[\beta \varepsilon_{n+k} A_{n-1}] \right) \\ &\quad + (1 - \beta)^2 E[A_{n+k-1} A_{n-1}] \\ &= \beta(1 - \beta) E[\varepsilon_n A_{n+k-1}] + (1 - \beta)^2 E[A_{n+k-1} A_{n-1}] + \frac{\beta}{\lambda^2}. \end{aligned} \quad (3.18)$$

Ahora bien, para el cálculo del primer sumando se tiene,

$$\varepsilon_n A_{n+k} = \begin{cases} \rho A_{n+k-1} \varepsilon_n & \text{c.p. } \rho, \\ \rho A_{n+k-1} \varepsilon_n + \varepsilon_{n+k} \varepsilon_n & \text{c.p. } (1 - \rho). \end{cases}$$

Comencemos cuando $k = 0$,

$$\begin{aligned} E[\varepsilon_n A_n] &= E[\rho A_{n-1} \varepsilon_n] + (1 - \rho) E[\varepsilon_n^2] \\ &= \rho E[A_{n-1} \varepsilon_n] + (1 - \rho) E[\varepsilon_n^2] \\ &= \rho \frac{1}{\lambda^2} + (1 - \rho) \frac{2}{\lambda^2} = \frac{(2 - \rho)}{\lambda^2}. \end{aligned}$$

Mientras que para $k \geq 1$, puesto que las $\{\varepsilon_n\}$ son v.a.'s i.i.d. Exponenciales(λ), se tiene

$$\begin{aligned} E[\varepsilon_n A_{n+k}] &= \rho E[A_{n+k-1} \varepsilon_n] + (1 - \rho) \frac{1}{\lambda^2} \\ &= \rho \left(\rho E[A_{n+k-2} \varepsilon_n] + \frac{1 - \rho}{\lambda^2} \right) + \frac{1 - \rho}{\lambda^2} \\ &= \rho \left(\rho \left(\rho E[A_{n+k-3} \varepsilon_n] + \frac{1 - \rho}{\lambda^2} \right) + \frac{1 - \rho}{\lambda^2} \right) + \frac{1 - \rho}{\lambda^2} \\ &\quad \vdots \end{aligned}$$

$$= \rho^k E[A_n \varepsilon_n] + \rho^{k-1} \frac{1-\rho}{\lambda^2} + \rho^{k-2} \frac{1-\rho}{\lambda^2} + \dots + \rho \frac{1-\rho}{\lambda^2} + \frac{1-\rho}{\lambda^2}. \quad (3.19)$$

Por lo tanto

$$E[\varepsilon_n A_{n+k}] = \rho^k \frac{(2-\rho)}{\lambda^2} + \frac{(1-\rho^k)}{\lambda^2} = \frac{\rho^k}{\lambda^2} (1-\rho) + \frac{1}{\lambda^2}, \quad k \geq 1. \quad (3.20)$$

Por otra parte, respecto al segundo sumando: puesto que $\{A_n\}$ es un proceso EAR(1), se sabe que su correlación serial es $\rho_A(k) = \rho^k$, y por consiguiente $E[A_n A_{n+k}] = \rho_A(k) \text{Var}(A) + E^2[A] = \rho^k \frac{1}{\lambda^2} + \frac{1}{\lambda^2}$.

Sustituyendo estos dos últimos resultados en (3.18)

$$\begin{aligned} E[X_n X_{n+k}] &= \beta(1-\beta) \left[\frac{\rho^{k-1}}{\lambda^2} (1-\rho) + \frac{1}{\lambda^2} \right] + (1-\beta)^2 \left[\frac{\rho^k + 1}{\lambda^2} \right] + \frac{\beta}{\lambda^2} \\ &= \beta(1-\beta)(1-\rho) \frac{\rho^{k-1}}{\lambda^2} + (1-\beta)^2 \left(\frac{\rho^k}{\lambda^2} \right) + \frac{1}{\lambda^2}. \end{aligned}$$

Entonces

$$\rho(1) = \text{Corr}(X_n, X_{n+1}) = \beta(1-\beta)(1-\rho) + (1-\beta)^2 \rho = g(\beta, \rho) \quad (3.21)$$

y en general

$$\rho(k) = \text{Corr}(X_n, X_{n+k}) = \rho^{k-1} g(\beta, \rho), \quad k = 1, 2, \dots \quad (3.22)$$

Vale la pena distinguir que cuando

$\rho = 0$, tenemos la correlación del proceso EMA(1), (3.14), y si

$\beta = 0$, entonces la correlación es la del proceso EAR(1), (3.7).

Escribimos $\rho(1)$, en la forma

$$\rho(1) = g(\beta, \rho) = \beta(1-\beta) + \rho(2\beta^2 - 3\beta + 1) = \beta(1-\beta) + \rho(1-\beta)(1-2\beta) = g_1(\beta) + \rho g_2(\beta).$$

Tanto g_1 como g_2 son no negativas para $\beta \in [0, 0.5]$. Mientras que para $\beta \in (0.5, 1]$ g_1 se mantiene no negativa y g_2 cambia de signo alcanzando su mínimo absoluto en $\beta = 0.75$ tomando un valor de -0.125 , para este mismo valor de β , g_1 vale 0.1875 .

Así pues cuando $\rho = 1$, $g(\beta, 1) = g_1(\beta) + g_2(\beta) = \beta^2 - 2\beta + 1$, que es un polinomio de segundo grado con mínimo absoluto igual a 0 en $\beta = 1$. De hecho en el intervalo $[0, 1]$ $g(\beta, 1)$ decrece monótonamente de uno a cero. Para $0 \leq \rho \leq 1$ fija, $g(\beta, \rho) = g_1(\beta) + \rho g_2(\beta)$ tiene un comportamiento totalmente análogo, decrece monotonamente de ρ a cero. Teniendo así que $\rho(k)$ es no negativa para toda k .

Distribución conjunta La transformada de Laplace de la distribución conjunta de X_n y X_{n+1} , la obtenemos siguiendo la misma estrategia que hemos usado en los modelos anteriores

$$\begin{aligned} \Phi_{X_n, X_{n+1}}(s_1, s_2) &= E[\exp(-s_1 X_n - s_2 X_{n+1})] \\ &= \beta^2 E[\exp(-s_1 \beta \varepsilon_n - s_2 \beta \varepsilon_{n+1})] \end{aligned}$$

$$\begin{aligned}
& +\beta(1-\beta)E[\exp(-s_1\beta\varepsilon_n - s_2\{\beta\varepsilon_{n+1} + A_n\})] \\
& +\beta(1-\beta)E[\exp(-s_1\{\beta\varepsilon_n + A_{n-1}\} - s_2\beta\varepsilon_{n+1})] \\
& +(1-\beta)^2E[\exp(-s_1\{\beta\varepsilon_n + A_{n-1}\} - s_2\{\beta\varepsilon_{n+1} + A_n\})],
\end{aligned}$$

donde en el segundo y cuarto sumandos debemos de ser cuidadosos con A_n puesto que este término no es independiente de ε_n . Sea $\psi(s)$ la transformada de Laplace de los $\{\varepsilon_n\}$ y $\phi(s)$ la transformada de las $\{A_n\}$ (ambas son iguales a $\lambda/(\lambda + s)$, si hacemos la diferencia es para permitir una mejor comprensión del desarrollo), entonces nos queda

$$\begin{aligned}
\Phi_{X_n, X_{n+1}}(s_1, s_2) &= \beta^2\psi(\beta s_1)\psi(\beta s_2) \\
& +\beta(1-\beta)\rho\psi(\beta s_1)\psi(\beta s_2)\phi(\rho s_2) + \beta(1-\beta)(1-\rho)\psi(\beta s_1 + s_2)\psi(\beta s_2)\phi(\rho s_2) \\
& +\beta(1-\beta)\psi(\beta s_1)\phi(s_1)\psi(\beta s_2) \\
& +(1-\beta)^2\rho\psi(\beta s_1)\psi(\beta s_2)\phi(s_1 + \rho s_2) + (1-\beta)^2(1-\rho)\psi(\beta s_1 + s_2)\psi(\beta s_2)\phi(s_1 + \rho s_2).
\end{aligned}$$

Agrupando el primer y tercer término, y el segundo con el último, nos queda que

$$\begin{aligned}
\Phi_{X_n, X_{n+1}}(s_1, s_2) &= \beta\psi(\beta s_1)\psi(\beta s_2) \left[\beta + (1-\beta)\phi(s_1) \right] \\
& + \left[\rho\psi(\beta s_1) + (1-\rho)\psi(\beta s_1 + s_2) \right] \psi(\beta s_2) \left[\beta(1-\beta)\phi(\rho s_2) + (1-\beta)^2\phi(s_1 + \rho s_2) \right] \\
& = \beta\psi(\beta s_1)\psi(\beta s_2) \left(\frac{\lambda + \beta s_1}{\lambda + s_1} \right) \\
& + \left(\frac{\lambda(\lambda + \beta s_1 + \rho s_2)}{(\lambda + \beta s_1)(\lambda + \beta s_1 + s_2)} \right) \psi(\beta s_2) \left(\frac{(1-\beta)\lambda(\lambda + \beta s_1 + \rho s_2)}{(\lambda + \rho s_2)(\lambda + s_1 + \rho s_2)} \right).
\end{aligned}$$

Finalmente

$$\begin{aligned}
\Phi_{X_n, X_{n+1}}(s_1, s_2) &= \left(\frac{\lambda}{\lambda + \beta s_1} \right) \left(\frac{\lambda}{\lambda + \beta s_2} \right) \\
& \times \left[\frac{\beta(\lambda + \beta s_1)}{(\lambda + s_1)} + \frac{(1-\beta)\lambda(\lambda + \beta s_1 + \rho s_2)^2}{(\lambda + \beta s_1 + s_2)(\lambda + \rho s_2)(\lambda + s_1 + \rho s_2)} \right]. \tag{3.23}
\end{aligned}$$

El lector puede comprobar que si

$\beta = 0$, entonces se tiene la transformada de Laplace del modelo EAR(1), (3.8), y cuando

$\rho = 0$, entonces tenemos la transformada de Laplace del modelo EMA(1), (3.15).

Por último, escribimos el proceso EARMA(1,1) descrito por las ecuaciones (3.16) y (3.17) en la siguiente forma:

Sean $\{U_n\}$ y $\{V_n\}$ sucesiones independientes de v.a.'s i.i.d. tales que, $\mathbb{P}[U_n = 0] = 1 - \mathbb{P}[U_n = 1] = \beta \in [0, 1]$ y, $\mathbb{P}[V_n = 0] = 1 - \mathbb{P}[V_n = 1] = \rho \in [0, 1]$.

Entonces el proceso EARMA(1,1), toma la forma

$$X_n = \beta\varepsilon_n + U_n A_{n-1}, \tag{3.24}$$

con

$$A_n = \rho A_{n-1} + V_n \varepsilon_n, \quad (3.25)$$

$n = 1, 2, \dots$, y $\{\varepsilon_n\}$ una sucesión de v.a.'s independientes Exponenciales(λ).

Esta notación no representa ventaja alguna con respecto a la anterior. Pero nos permite visualizar su estrecha relación con respecto al primer gran intento de definir una clase de modelos de series de tiempo con valores discretos análoga a la familia ARMA, los modelos DARMA. Modelos que a continuación estudiamos.

3.2. Modelos discretos generados por mezclas

Hasta antes de los modelos que serán estudiados en esta sección sólo se contaba con modelos de cadenas de Markov para analizar sucesiones dependientes de v.a.'s discretas.

Sin embargo estos modelos suelen tener dos inconvenientes:

1. Estar sobreparametrizados.
2. Frecuentemente se presentan datos cuya estructura se puede mostrar que no es de Markov, al menos no de primer orden. Cadenas de Markov de orden más alto pueden usarse pero esto sólo empeora el inconveniente anterior.

Los modelos DARMA sin embargo permiten trabajar con sucesiones de v.a.'s discretas que no son de Markov.

3.2.1. DARMA(1, $N + 1$)

Sea $\{Y_n\}$ una sucesión de v.a.'s i.i.d. que toma valores en un subconjunto contable E de la línea real con distribución π , esto es, $\mathbb{P}[Y_n = i] = \pi(i)$ para toda $i \in E$.

Sean $\{U_n\}$ y $\{V_n\}$ sucesiones independientes de v.a.'s i.i.d. que toman el valor 0 y 1 de la siguiente manera,

$$\mathbb{P}[U_n = 1] = \beta \quad y \quad \mathbb{P}[V_n = 1] = \rho,$$

con $0 \leq \beta \leq 1$ y, $0 \leq \rho < 1$, fijas.

Sea $\{D_n\}$ una sucesión de v.a.'s i.i.d., tal que

$$\mathbb{P}[D_n = i] = \delta_i, \quad i = 0, 1, 2, \dots, N,$$

con N un entero no negativo fijo.

La primer clase de modelos para series de tiempo con valores discretos que estudiaremos, es la clase DARMA(1, $N + 1$) (Discreto de mezclas Autoregresivo-Promedios Móviles de orden uno y $N + 1$,

respectivamente). Ésta fue publicada en 1978 por Jacobs y Lewis, en dos diferentes artículos, y está definida como

$$X_n = U_n Y_{n-D_n} + (1 - U_n) Z_{n-(N+1)}, \quad (3.26)$$

$$n = 1, 2, \dots,$$

con componente autoregresiva

$$Z_n = V_n Z_{n-1} + (1 - V_n) Y_n, \quad (3.27)$$

$$n = -N, -N + 1, -N + 2, \dots,$$

Estacionariedad De (3.27) observamos que la v.a. Z_n puede ser expandida hacia atrás, y obtener $Z_n = Y_{n-j}$ con probabilidad $\rho^j(1 - \rho)$ para $0 \leq j \leq N + n$, y $Z_n = Z_{-(N+1)}$ con probabilidad ρ^{N+n+1} . Esto es, Z_n es una mezcla de las v.a.'s $Y_n, Y_{n-1}, \dots, Y_{-N}$, y $Z_{-(N+1)}$.

Afirmación: Si empezamos el proceso con $Z_{-(N+1)}$ teniendo distribución π independiente de $\{Y_n; n \geq -N\}$, $\{U_n\}$, $\{V_n\}$ y $\{D_n\}$, entonces $\{X_n\}$ forma una sucesión estacionaria de v.a.'s dependientes discretas con distribución marginal π .

Demostración: Por la observación anterior, para cualquier $i \in E$

$$\begin{aligned} \mathbb{P}[Z_n = i] &= \sum_{j=0}^{N+n} \mathbb{P}[Y_{n-j} = i] \rho^j (1 - \rho) + \mathbb{P}[Z_{-(N+1)} = i] \rho^{N+n+1} \\ &= \sum_{j=0}^{N+n} \pi(i) \rho^j (1 - \rho) + \pi(i) \rho^{N+n+1} = \pi(i), \end{aligned} \quad (3.28)$$

para $n = -N, -N + 1, -N + 2, \dots$

Para $\{Y_{n-D_n}\}$, tenemos

$$\mathbb{P}[Y_{n-D_n} = i] = \sum_{j=0}^N \mathbb{P}[Y_{n-j} = i, D_n = j] = \pi(i) \sum_{j=0}^N \mathbb{P}[D_n = j] = \pi(i), \quad (3.29)$$

para todo $i \in E$ y $n = 1, 2, \dots$. Entonces

$$\mathbb{P}[X_n = i] = \beta \mathbb{P}[Y_{n-D_n} = i] + (1 - \beta) \mathbb{P}[Z_{n-(N+1)} = i] = \pi(i),$$

para todo $i \in E$ y $n = 1, 2, \dots$ \square

Correlación serial Aunque las X_n tienen distribución estacionaria π , no son independientes como las Y_n , de hecho la covarianza entre X_{n+j} y X_n , es

$\text{Cov}(X_{n+j}, X_n)$

$$= \text{Cov}(U_{n+j} Y_{n+j-D_{n+j}} + (1 - U_{n+j}) Z_{n+j-(N+1)}, U_n Y_{n-D_n} + (1 - U_n) Z_{n-(N+1)})$$

$$\begin{aligned}
&= \beta^2 \text{Cov}(Y_{n+j-D_{n+j}}, Y_{n-D_n}) \\
&\quad + \beta(1-\beta) [\text{Cov}(Y_{n+j-D_{n+j}}, Z_{n-(N+1)}) + \text{Cov}(Y_{n-D_n}, Z_{n-(N+1)+j})] \\
&\quad + (1-\beta)^2 \text{Cov}(Z_{n-(N+1)+j}, Z_{n-(N+1)}). \tag{3.30}
\end{aligned}$$

Examinemos con detenimiento cada uno de los cuatro sumandos.

Para el primero se tiene que, si $j > N$

$$\text{Cov}(Y_{n+j-D_{n+j}}, Y_{n-D_n}) = 0, \tag{3.31}$$

para esto basta ver que $n+j-D_{n+j} \neq n-D_n$, para toda n y D_n .

Mientras que para $1 \leq j \leq N$, si $D_{n+j} = k$ y $D_n = k-j$, obtenemos la varianza de $\{Y_n\}$. Ahora bien, como $D_n \in \{0, 1, \dots, N\}$ se tiene que k debe empezar en j y terminar en N . Por lo tanto

$$\text{Cov}(Y_{n+j-D_{n+j}}, Y_{n-D_n}) = \text{Var}(Y_1) \sum_{k=j}^N \mathbb{P}[D_{n+j} = k] \mathbb{P}[D_n = k-j].$$

Entonces,

$$\text{Corr}(Y_{n+j-D_{n+j}}, Y_{n-D_n}) = \sum_{k=j}^N \delta_k \delta_{k-j} = \sum_{k=0}^{N-j} \delta_k \delta_{k+j} \tag{3.32}$$

Para el segundo sumando es claro que siempre se cumple la relación $n+j-D_{n+j} > n-(N+1)$, por lo tanto

$$\text{Corr}(Y_{n+j-D_{n+j}}, Z_{n-(N+1)}) = 0, \quad \text{para } j \geq 1. \tag{3.33}$$

Para que el tercer sumando tenga valor distinto de cero, necesitamos $n-D_n \leq n-(N+1)+j$, o equivalentemente $(N+1)-j \leq D_n \leq N$, entonces si $1 \leq j \leq N$, D_n tomará valores desde $(N+1)-j$ hasta N , mientras que para $j > N$, D_n va de cero a N . Así entonces, para $1 \leq j \leq N$,

$$\begin{aligned}
&\text{Cov}(Y_{n-D_n}, Z_{n-(N+1)+j}) \\
&= \sum_{k=N+1-j}^N \text{Cov}(Y_{n-k}, Z_{n-(N+1)+j}) \mathbb{P}[D_n = k] \\
&= (1-\rho) \text{Var}(Y_{n-(N+1)+j}) \delta_{N+1-j} + \rho(1-\rho) \text{Var}(Y_{n-(N+2)+j}) \delta_{N+2-j} \\
&\quad + \rho^2(1-\rho) \text{Var}(Y_{n-(N+3)+j}) \delta_{N+3-j} + \dots + \rho^{j-1}(1-\rho) \text{Var}(Y_{n-N}) \delta_N.
\end{aligned}$$

Como las Y_n son idénticamente distribuidas

$$\begin{aligned}
\text{Corr}(Y_{n-D_n}, Z_{n-(N+1)+j}) &= \sum_{k=N+1-j}^N (1-\rho) \rho^{k-(N+1-j)} \delta_k \\
&= (1-\rho) \rho^{-(N+1-j)} \sum_{k=N+1-j}^N \rho^k \delta_k, \quad 1 \leq j \leq N. \tag{3.34}
\end{aligned}$$

Y para $j > N$,

$$\text{Corr}(Y_{n-D_n}, Z_{n-(N+1)+j}) = (1-\rho) \rho^{j-(N+1)} \sum_{k=0}^N \rho^k \delta_k. \tag{3.35}$$

Finalmente para el cuarto tenemos

$$\begin{aligned} & \text{Cov}(Z_{n-(N+1)+j}, Z_{n-(N+1)}) \\ &= \rho \text{Cov}(Z_{n-(N+1)+j-1}, Z_{n-(N+1)}) + (1 - \rho) \text{Cov}(Y_{n-(N+1)+j}, Z_{n-(N+1)}), \end{aligned}$$

donde el segundo sumando es cero para toda $j \geq 1$, y para el primero es fácil ver que

$$\text{Cov}(Z_{n-(N+1)+j}, Z_{n-(N+1)}) = \rho^j \text{Var}(Z_{n-(N+1)}).$$

Por lo tanto

$$\text{Corr}(Z_{n-(N+1)+j}, Z_{n-(N+1)}) = \rho^j. \quad (3.36)$$

Uniendo los cuatro términos se tiene que la correlación serial del proceso es la siguiente:

Si $1 \leq j \leq N$

$$\begin{aligned} \rho(j) &= \text{Corr}(X_{n+j}, X_n) \\ &= \beta^2 \sum_{k=0}^{N-j} \delta_k \delta_{k+j} + \beta(1 - \beta)(1 - \rho)\rho^{-(N+1-j)} \sum_{k=N+1-j}^N \rho^k \delta_k + (1 - \beta)^2 \rho^j \end{aligned} \quad (3.37)$$

y para $j > N$

$$\rho(j) = \text{Corr}(X_{n+j}, X_n) = \beta(1 - \beta)\rho^{j-(N+1)} \sum_{k=0}^N \rho^k (1 - \rho)\delta_k + (1 - \beta)^2 \rho^j. \quad (3.38)$$

Note que la correlación del proceso es independiente de la distribución marginal estacionaria. Además, observe que las correlaciones son todas no negativas.

DAR(1)

El proceso DAR(1) está definido por la ecuación (3.27).

Por razones de presentación abordamos su estudio en la siguiente subsección, DAR(p).

DMA(1)

El proceso DMA(N) está definido como $X_n = Y_{n-D_n}$ el cual no es de Markov en general.

Por (3.29), (3.31) y (3.32) tenemos que el proceso DMA(N) es estacionario con distribución marginal π , y correlación

$$\rho(j) = \text{Corr}(X_n, X_{n+j}) = \sum_{k=0}^{N-j} \delta_k \delta_{k+j}, \quad 1 \leq j \leq N$$

y $\rho(j) = 0$, para $j > N$.

Cuando $N = 1$, el proceso DMA(1) queda como

$$X_n = \begin{cases} Y_n & \text{c.p. } \alpha, \\ Y_{n-1} & \text{c.p. } (1 - \alpha). \end{cases}$$

con $\alpha = \delta_0$.

En este caso

$$\rho(1) = \alpha(1 - \alpha) \quad \text{y} \quad \rho(j) = 0, \quad j \geq 2.$$

Longitud de una corrida Para $i \in E$ fijo, definimos

$$T_i = \inf\{n \geq 1; Z_n \neq i\} - 1,$$

como la longitud de una corrida de i , empezando al tiempo uno, y $T_i \in \{0, 1, \dots\}$

Deseamos calcular la $E[T_i]$.

Sea $a_0 = 1$ y $a_n = \mathbb{P}[T_i \geq n] = \mathbb{P}[X_1 = X_2 = \dots = X_n = i]$ para $n \geq 1$, entonces

$$a_1 = \mathbb{P}[X_1 = i] = \pi(i)$$

$$a_2 = \mathbb{P}[X_1 = X_2 = i] = \alpha\pi(i)a_1 + (1 - \alpha)\pi(i) [\alpha a_0 + (1 - \alpha)\pi(i)]$$

$$= \alpha\pi(i)a_1 + \alpha(1 - \alpha)\pi(i)a_0 + (1 - \alpha)^2\pi(i)^2.$$

Para aclarar un poco esta fórmula, presentamos el siguiente razonamiento

$$X_2 = \begin{cases} Y_2 & \text{c.p. } \alpha, \\ Y_1 & \text{c.p. } (1 - \alpha), \end{cases} \quad X_1 = \begin{cases} Y_1 & \text{c.p. } \alpha, \\ Y_0 & \text{c.p. } (1 - \alpha). \end{cases}$$

Caso 1: Con probabilidad α , X_2 toma el valor Y_2 , luego le pido que tome el valor i , con probabilidad $\pi(i)$, y además necesito que $X_1 = i$, que da el primer sumando de a_2 ($\alpha\pi(i)a_1$).

Caso 2: Con probabilidad $1 - \alpha$, X_2 toma el valor Y_1 , pedimos que tome el valor i , entonces pueden suceder dos cosas,

i) con probabilidad α , $X_1 = Y_1$, y ya acabamos $((1 - \alpha)\pi(i)\alpha)$,

ii) con probabilidad $1 - \alpha$, $X_1 = Y_0$, en tal caso pedimos que tome el valor i ,
 $((1 - \alpha)\pi(i)(1 - \alpha)\pi(i))$.

Para $n = 3$,

$$X_3 = \begin{cases} Y_3 & \text{c.p. } \alpha, \\ Y_2 & \text{c.p. } (1 - \alpha), \end{cases} \quad X_2 = \begin{cases} Y_2 & \text{c.p. } \alpha, \\ Y_1 & \text{c.p. } (1 - \alpha), \end{cases} \quad X_1 = \begin{cases} Y_1 & \text{c.p. } \alpha, \\ Y_0 & \text{c.p. } (1 - \alpha), \end{cases}$$

$$a_3 = \alpha\pi(i)a_2 + (1 - \alpha)\pi(i) [\alpha a_1 + (1 - \alpha)\pi(i) \{ \alpha a_0 + (1 - \alpha)\pi(i) \}]$$

$$= \alpha\pi(i)a_2 + \alpha(1 - \alpha)\pi(i)a_1 + \alpha(1 - \alpha)^2\pi(i)^2 + (1 - \alpha)^3\pi(i)^3.$$

Explicación:

Caso 1: Con probabilidad α , $X_3 = Y_3$, pedimos que tome el valor i , y además necesitamos que $X_2 = X_1 = i$. Todo esto sucede con probabilidad, $\alpha\pi(i)a_2$.

Caso 2: Con probabilidad $1 - \alpha$, $X_2 = Y_2$, le pido que tome el valor i , entonces

- i) si $X_2 = Y_2$, con probabilidad α , sólo nos resta pedir que $X_1 = i$ ($(1 - \alpha)\pi(i)\alpha a_1$),
- ii) si $X_2 = Y_1$, con probabilidad $1 - \alpha$, pedimos que tome el valor i , y entonces
- a) si $X_1 = Y_1$ con probabilidad α , ($(1 - \alpha)\pi(i)(1 - \alpha)\pi(i)\alpha$),
- b) si $X_1 = Y_0$ con probabilidad $1 - \alpha$, pedimos que tome el valor i ,
 $((1 - \alpha)\pi(i)(1 - \alpha)\pi(i)(1 - \alpha)\pi(i))$.

En general se puede probar por inducción que

$$\begin{aligned}
 a_n &= \mathbb{P}[X_1 = X_2 = \dots = X_n = i] \\
 &= \alpha\pi(i)a_{n-1} + \bar{\alpha}\pi(i) \left[\alpha a_{n-2} + \bar{\alpha}\pi(i) \left[\alpha a_{n-3} + \bar{\alpha}\pi(i) \left[\dots \right. \right. \right. \\
 &\qquad\qquad\qquad \left. \left. \left. \dots \dots \dots \{ \alpha a_1 + \bar{\alpha}\pi(i) [\alpha a_0 + \bar{\alpha}\pi(i)] \} \right] \right] \right] \\
 &= \alpha\pi(i)a_{n-1} + \alpha\bar{\alpha}\pi(i)a_{n-2} + \alpha\bar{\alpha}^2\pi(i)^2a_{n-3} + \dots + \alpha\bar{\alpha}^{n-1}\pi(i)^{n-1}a_0 + \bar{\alpha}^n\pi^n,
 \end{aligned}$$

con $\bar{\alpha} = 1 - \alpha$.

Entonces calculamos la esperanza de T_i como

$$\begin{aligned}
 E[T_i] &= \sum_{n=1}^{\infty} \mathbb{P}[T_i \geq n] = \sum_{n=1}^{\infty} a_n \\
 &= a_1 \\
 &\quad + \alpha\pi(i)a_1 \quad + \alpha\bar{\alpha}\pi(i)a_0 \quad + \bar{\alpha}^2\pi(i)^2 \\
 &\quad + \alpha\pi(i)a_2 \quad + \alpha\bar{\alpha}\pi(i)a_1 \quad + \alpha\bar{\alpha}^2\pi(i)^2a_0 \quad + \bar{\alpha}^3\pi(i)^3 \\
 &\quad + \alpha\pi(i)a_3 \quad + \alpha\bar{\alpha}\pi(i)a_2 \quad + \alpha\bar{\alpha}^2\pi(i)^2a_1 \quad + \alpha\bar{\alpha}^3\pi(i)^3a_0 \quad + \bar{\alpha}^4\pi(i)^4 \\
 &\quad + \dots
 \end{aligned}$$

Por lo tanto

$$\begin{aligned}
 E[T_i] &= \sum_{n=1}^{\infty} a_n \\
 &= a_1 + \alpha\pi(i) \sum_{n=1}^{\infty} a_n + \alpha(1 - \alpha)\pi(i) \left[\sum_{n=0}^{\infty} a_n \right] + \alpha(1 - \alpha)^2\pi(i)^2 \left[\sum_{n=0}^{\infty} a_n \right] + \dots \\
 &\qquad\qquad\qquad \dots + \alpha(1 - \alpha)^n\pi(i)^n \left[\sum_{n=0}^{\infty} a_n \right] + \dots + \sum_{n=2}^{\infty} (1 - \alpha)^n\pi(i)^n \\
 &= a_1 + \alpha\pi(i) \sum_{n=1}^{\infty} a_n + \sum_{n=0}^{\infty} a_n \left[\frac{\alpha(1 - \alpha)\pi(i)}{1 - (1 - \alpha)\pi(i)} \right] + \frac{(1 - \alpha)^2\pi(i)^2}{1 - (1 - \alpha)\pi(i)}.
 \end{aligned}$$

Sustituyendo

$$E[T_i] = \pi(i) + \frac{\alpha(1 - \alpha)\pi(i) + (1 - \alpha)^2\pi(i)^2}{1 - (1 - \alpha)\pi(i)} + \left[\alpha\pi(i) + \frac{\alpha(1 - \alpha)\pi(i)}{1 - (1 - \alpha)\pi(i)} \right] E[T_i].$$

Agrupando términos

$$E[T_i] = \frac{\pi(i)[1 + \alpha(1 - \alpha)(1 - \pi(i))]}{\{1 - \alpha(1 - \alpha)\pi(i)\}\{1 - \pi(i)\}}.$$

Si α es cero o uno, $\{X_n\}$ es una sucesión de v.a.'s independientes. En tal caso $E[T_i] = \pi(i)/(1-\pi(i))$, mientras que para $0 \leq \alpha \leq 1$,

$$E[T_i] \geq \frac{\pi(i)}{1-\pi(i)}.$$

Esto significa que la longitud esperada de una corrida de i 's para el proceso DMA(1) es siempre mayor o igual a la longitud esperada de una corrida de i 's para una sucesión de v.a.'s independientes. Además, dada una distribución π , el máximo de $E[T_i]$ ocurre en $\alpha = 1/2$. En este caso

$$E[T_i] = \frac{\pi(i)}{1-\pi(i)} \left(1 + \frac{1}{4-\pi(i)} \right).$$

Por último, es de mencionar que los autores muestran que el proceso DMA(1) es reversible en el tiempo.

DARMA(1,1)

El proceso $\{X_n\}$, DARMA(1,1), está dado por

$$X_n = \begin{cases} Y_n & \text{c.p. } \beta, \\ Z_{n-1} & \text{c.p. } (1-\beta) \end{cases} \quad \text{y} \quad Z_n = \begin{cases} Z_{n-1} & \text{c.p. } \rho, \\ Y_n & \text{c.p. } (1-\rho), \end{cases}$$

el cual es no Markoviano en general.

En el proceso DARMA(1,1), hay una elección aleatoria entre Z_{n-1} y Y_n . Z_{n-1} contiene toda la información disponible del pasado y si Y_n es escogida, entonces la memoria del proceso antes del tiempo n es perdida para siempre.

Haciendo $N = 0$, (3.31)-(3.34) son cero y, de (3.35) y (3.36) se tiene que

$$\rho(j) = \text{Corr}(X_n, X_{n+j}) = \rho^{j-1}[\beta(1-\beta)(1-\rho) + (1-\beta)^2\rho], \quad j \geq 1.$$

3.2.2. DAR(p)

El modelo DAR(p) es la generalización del modelo DAR(1), ec. (3.27). Permite generar una sucesión estacionaria de v.a.'s discretas con dependencia de Markov de orden p y distribución marginal dada.

Sea $\{Y_n\}$ y $\{V_n\}$ como en la subsección 2.2.1., y $\{A_n\}$ una sucesión de v.a.'s independientes que toma valores en $\{1, 2, \dots, p\}$, con p en los naturales y tal que

$$\mathbb{P}[A_n = i] = \alpha_i, \quad i = 1, 2, \dots, p.$$

con $\sum_{i=1}^p \alpha_i = 1$. Definimos el proceso DAR(p) (Discreto Autoregresivo de orden p) como

$$X_n = V_n X_{n-A_n} + (1-V_n)Y_n, \quad n = 1, 2, \dots \quad (3.39)$$

Directamente de la definición se tiene

$$\mathbb{P}[X_{n+1} = j | X_n = i_1, \dots, X_{n-p+1} = i_p] = (1 - \rho)\pi(j) + \sum_{k=1}^p \rho\alpha_k I_j(i_k),$$

donde

$$I_j(i_k) = \begin{cases} 1 & \text{si } i_k = j, \\ 0 & \text{e.o.c.} \end{cases}$$

Estacionariedad La definición del proceso DAR(p) requiere la especificación de la distribución conjunta del vector $(X_{-p+1}, \dots, X_{-1}, X_0)$ que haga a la sucesión $\{X_n\}$ estacionaria.

En Jacobs y Lewis (1978c) se muestra que existe una única distribución límite ν (con distribución marginal π) para la cadena de Markov

$$\{(X_n, X_{n+1}, \dots, X_{n+p-1}), n = 1, 2, \dots\}.$$

Por lo tanto si usamos ν como la distribución inicial para $(X_{-p+1}, \dots, X_{-1}, X_0)$, el proceso $\{X_n\}$, DAR(p), es estacionario con distribución marginal π .

Correlación serial Sea $\{X_n\}$ un proceso estacionario DAR(p) con distribución marginal π .

De (3.39) se tiene que $X_{n+k} = X_{n+k-j}$ con probabilidad $\rho\alpha_j$, para $j = 1, 2, \dots, p$.

Entonces

$$\text{Cov}(X_n, X_{n+1})$$

$$= \rho\{E[X_n X_{n+1-A_n}] - E[X_n]E[X_{n+1-A_n}]\} + (1 - \rho)\{E[X_n Y_{n+1}] - E[X_n]E[Y_{n+1}]\}.$$

El segundo sumando es cero y por el primero se tiene que

$$\text{Cov}(X_n, X_{n+1}) = \rho \sum_{j=1}^p \alpha_j \text{Cov}(X_n, X_{n+1-j}).$$

Por lo tanto

$$r(1) = \rho\alpha_1 r(0) + \rho\alpha_2 r(1) + \dots + \rho\alpha_p r(p-1),$$

con $r(k) = \text{Corr}(X_n, X_{n+k})$.

Haciendo lo mismo con X_{n+2}, X_{n+3}, \dots obtenemos el conjunto completo de las ecuaciones de Yule-Walker,

$$\begin{aligned} r(1) &= \rho\alpha_1 r(0) & + \rho\alpha_2 r(1) & + \dots & + \rho\alpha_p r(p-1) \\ r(2) &= \rho\alpha_1 r(1) & + \rho\alpha_2 r(0) & + \dots & + \rho\alpha_p r(p-2) \\ &\vdots \\ r(p) &= \rho\alpha_1 r(p-1) & + \rho\alpha_2 r(p-2) & + \dots & + \rho\alpha_p r(0) \end{aligned}$$

y para $k \geq 1$,

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

Afirmación: Las correlaciones seriales del proceso DAR(p) son todas no negativas.

Demostración: Sea $q(i), i \geq 1$ la probabilidad de que X_n y X_{n+i} tomen el mismo valor aleatorio Y_k con $k \leq n$. Sea R_n el índice aleatorio de las Y_k $k \leq n$ que X_n toma, i.e. $X_n = Y_{R_n}$,

$$\begin{aligned} E[X_n X_{n+l}] &= E[Y_{R_n} Y_{R_{n+l}}] \\ &= \sum_{k=1}^n \sum_{j=1}^{n+l} E[Y_k Y_j] \mathbb{P}[R_n = k, R_{n+l} = j] \\ &= \sum_{k=1}^n E[Y_k^2] \mathbb{P}[R_n = R_{n+l} = k] + \sum_{k=1}^n \sum_{\substack{j=1 \\ j \neq k}}^{n+l} E[Y_k Y_j] \mathbb{P}[R_n = k, R_{n+l} = j]. \end{aligned}$$

Puesto que las Y_n , son v.a.'s i.i.d.

$$E[X_n X_{n+l}] = E[Y_1^2] \mathbb{P}[R_n = R_{n+l}] + E^2[Y_1] \mathbb{P}[R_n \neq R_{n+l}].$$

Restando $E[X_n]E[X_{n+l}] = E[Y_{R_n}]E[Y_{R_{n+l}}] = E^2[Y_1]$

$$\begin{aligned} \text{Cov}(X_n, X_{n+l}) &= E[Y_1^2] \mathbb{P}[R_n = R_{n+l}] + E^2[Y_1] (\mathbb{P}[R_n \neq R_{n+l}] - 1) \\ &= \text{Var}(Y_1) \mathbb{P}[R_n = R_{n+l}], \end{aligned}$$

entonces

$$r(l) = \text{Corr}(X_n, X_{n+l}) = \mathbb{P}[R_n = R_{n+l}] = q(l), \quad i \geq 1.$$

Por lo tanto las correlaciones del proceso DAR(p) son todas no negativas. \square

DAR(1)

El proceso DAR(1) está dado por

$$X_n = V_n X_{n-1} + (1 - V_n) Y_n, \quad n = 1, 2, \dots$$

y ya ha sido tratado en la subsección 3.2.1, obteniéndose resultados como las condiciones de estacionariedad y su correlación serial (ecs. (3.28) y (3.36)) bajo el enfoque de mezclas. No obstante, presentamos de nuevo estos dos resultados pero bajo el punto de vista de cadenas de Markov y utilizando la linealidad del proceso.

Directamente de la definición el proceso DAR(1) es una cadena de Markov (de orden 1) con matriz de probabilidades de transición $[\mathbb{P}]_{i,j \in E}$, dada por

$$\mathbb{P}[X_{n+1} = j | X_n = i] = p(i, j) = (1 - \rho)\pi(j) + \rho I_j(i),$$

con

$$I_j(i) = \begin{cases} 1 & \text{si } i = j, \\ 0 & \text{e.o.c.} \end{cases}$$

Estacionariedad Por demostrar

$$\sum_{i \in E} \pi(i)p(i, j) = \pi(j), \quad j \in E.$$

Demostración: Sean $i, j \in E$,

$$\begin{aligned} \sum_{i \in E} \pi(i)p(i, j) &= (1 - \rho)\pi(j) \sum_{i \in E} \pi(i) + \rho \sum_{i \in E} I_j(i)\pi(i) \\ &= (1 - \rho)\pi(j) + \rho\pi(j) = \pi(j). \end{aligned}$$

Por lo tanto basta que X_0 tenga distribución π para que el proceso sea estacionario. \square

Cabe mencionar que se ha llevado un dirección opuesta a la que normalmente se lleva en la teoría de cadenas de Markov: primero hemos especificado la distribución estacionaria asociada a la cadena y luego la matriz de transición. Además al cambiar el parámetro ρ no se afecta a la distribución estacionaria, π .

Correlación serial Tomando esperanza condicional al proceso DAR(1)

$$E[X_{n+1} | X_n = x] = \rho x + (1 - \rho)E[Y_n],$$

ya que la esperanza de $\{Y_n\}$ y $\{X_n\}$ es la misma, aplicamos el Corolario 2.17, y corroboramos una vez más que

$$\text{Corr}(X_n, X_{n+k}) = \rho^k, \quad k = 0, 1, \dots$$

Longitud de una corrida Para $i \in E$ fijo, definimos $T_i = \inf\{n \geq 1; Z_n \neq i\} - 1$, como antes.

Afirmación: Para este proceso, al igual que para el caso del modelo DMA(1), la longitud esperada de una corrida de i 's, es mayor o igual que la de una sucesión de v.a.'s independientes.

Demostración: Para $n \geq 1$,

$$\begin{aligned} \mathbb{P}[T_i \geq n] &= \mathbb{P}[Z_1 = Z_2 = \dots = Z_n = i] \\ &= \mathbb{P}[Z_n = i | Z_{n-1} = i] \mathbb{P}[Z_{n-1} = i | Z_{n-2} = i] \dots \mathbb{P}[Z_1 = i] = \pi(i)p(i, i)^{n-1}. \end{aligned}$$

Y, $\mathbb{P}[T_i = 0] = 1 - \pi(i)$.

Entonces calculamos la esperanza de T_i como

$$\begin{aligned} E[T_i] &= \sum_{n=1}^{\infty} \mathbb{P}[T_i \geq n] = \sum_{n=1}^{\infty} \pi(i)p(i, i)^{n-1} \\ &= \pi(i) \frac{1}{1 - p(i, i)} = \frac{\pi(i)}{(1 - \rho)(1 - \pi(i))}. \end{aligned}$$

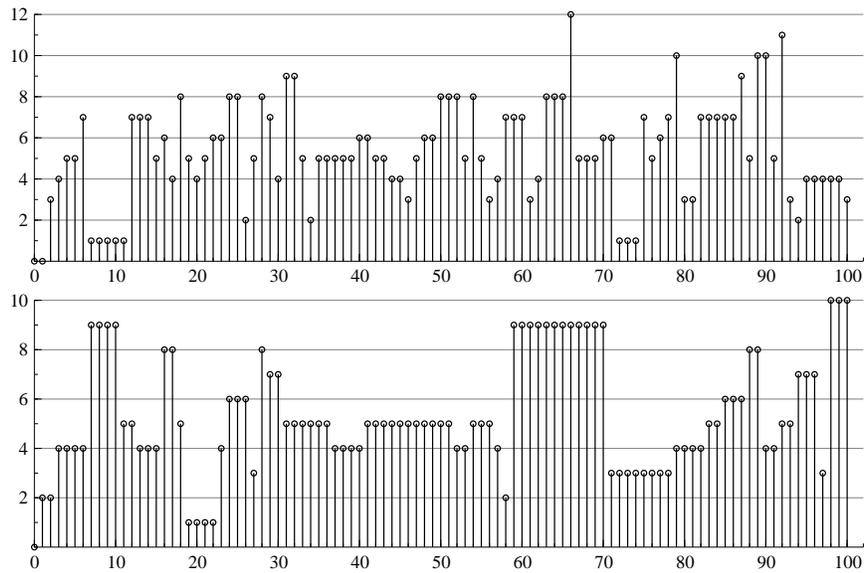


Figura 3.1: Simulación de un proceso Poisson DAR(1), con $\lambda = 5$ y $n = 100$. En la parte superior: $\rho = 0.3$. En la parte inferior: $\rho = 0.7$.

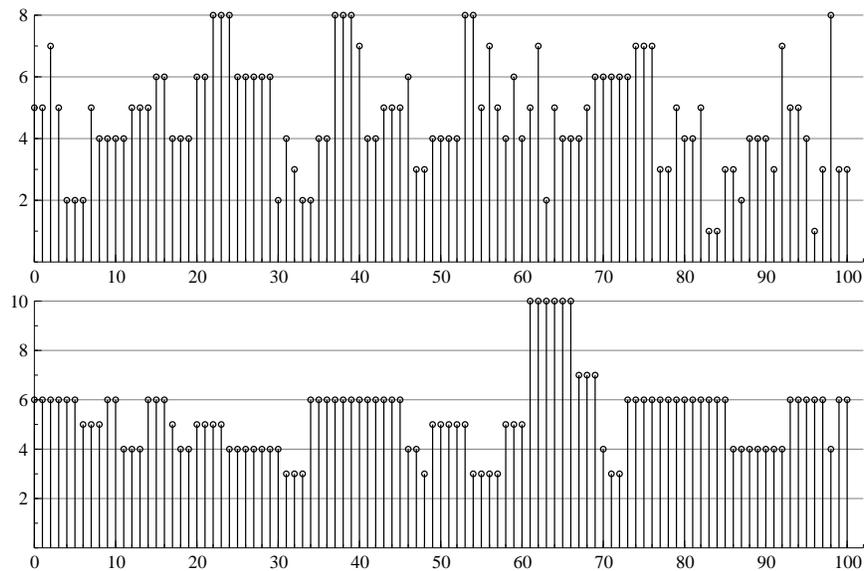


Figura 3.2: Simulación de un proceso Binomial DAR(1), con $N = 10$, $p = 0.5$ y $n = 100$. En la parte superior: $\rho = 0.3$. En la parte inferior: $\rho = 0.7$.

Si $\rho = 0$, entonces $Z_n = Y_n$, i.e. $\{Z_n\}$ es una sucesión de v.a.'s independientes y $E[T_i] = \pi(i)/(1 - \pi(i))$. Mientras que para $0 \leq \rho < 1$

$$E[T_i] \geq \frac{\pi(i)}{1 - \pi(i)},$$

como afirmamos. Además note que este incremento es constante para todos los estados ($1/(1 - \rho)$).

□

Por otra parte resulta un cálculo muy simple la $\mathbb{P}[T_i = n]$.

Para $n \geq 1$,

$$\mathbb{P}[T_i = n] = \mathbb{P}[T_i \geq n] - \mathbb{P}[T_i \geq n + 1] = \pi(i)P(i, i)^{n-1}[1 - P(i, i)]. \quad \square$$

Las figuras 3.1 y 3.2 muestran simulaciones de procesos DAR(1) con distribución marginal Poisson y Binomial respectivamente.

3.2.3. DARMA($p, N + 1$)

Sean $\{Y_n\}$, $\{U_n\}$, $\{V_n\}$, $\{A_n\}$ y $\{D_n\}$ como hasta ahora. El proceso DARMA($p, N + 1$) (Discreto de mezclas Autoregresivo-Promedios Móviles de orden p y $N + 1$, respectivamente) está definido como

$$X_n = U_n Y_{n-D_n} + (1 - U_n) Z_{n-(N+1)}, \quad (3.40)$$

$$n = 1, 2, \dots,$$

con cola autoregresiva dada por

$$Z_n = V_n Z_{n-A_n} + (1 - V_n) Y_n, \quad (3.41)$$

$$n = -N, -N + 1, \dots$$

Como podemos observar este modelo conjunta a los 2 anteriores.

Estacionariedad Si $\{Z_{-N-p}, Z_{-N-(p-1)}, \dots, Z_{-N-1}\}$ tiene distribución conjunta ν como en la subsección 3.2.2, DAR(p), el proceso DARMA($p, N + 1$), $\{X_n; n = 1, 2, \dots\}$ es estacionario con distribución marginal π .

Correlación serial Sea $\{X_n\}$ un proceso estacionario DARMA($p, N + 1$).

Sea $r_A(k) = \text{Corr}(Z_n, Z_{n+k})$, $k \geq 1$, que sabemos se obtiene mediante las ecuaciones de Yule-Walker (subsección anterior).

Por otra parte,

$$\text{Cov}(Z_n, Y_n) = \rho(E[Z_{n-A_n} Y_n] - E[Z_{n-A_n}]E[Y_n]) + (1 - \rho)(E[Y_n^2] - E[Y_n]E[Y_n]),$$

puesto que el primer término es cero se tiene que

$$\text{Corr}(Z_{n-N}, Y_{n-N}) = (1 - \rho). \quad (3.42)$$

Sea $r_B(k) = \text{Corr}(Z_{n+k-(N+1)}, Y_{n-D_n}) = \sum_{j=0}^N \text{Corr}(Z_{n+k-(N+1)}, Y_{n-j})\delta_j$. Para que la correlación sea distinta de cero necesitamos que $n+k-(N+1) \geq n-j$ o equivalentemente $j \geq (N+1)-k$.

Entonces para

$$k = 0,$$

$$r_B(0) = 0.$$

$$k = 1,$$

$$r_B(1) = \delta_N \text{Corr}(Z_{n+1-(N+1)}, Y_{n-N}) = \delta_N(1 - \rho).$$

$$k = 2,$$

$$\begin{aligned} r_B(2) &= \delta_{N-1} \text{Corr}(Z_{n+2-(N+1)}, Y_{n-(N-1)}) + \delta_N \text{Corr}(Z_{n+2-(N+1)}, Y_{n-N}) \\ &= \delta_{N-1}(1 - \rho) + \delta_N \rho \alpha_1 \text{Corr}(Z_{n+1-(N+1)}, Y_{n-N}) = \delta_{N-1}(1 - \rho) + \rho \alpha_1 r_B(1). \end{aligned}$$

$$k = 3,$$

$$\begin{aligned} r_B(3) &= \delta_{N-2} \text{Corr}(Z_{n+3-(N+1)}, Y_{n-(N-2)}) \\ &\quad + \delta_{N-1} \text{Corr}(Z_{n+3-(N+1)}, Y_{n-(N-1)}) + \delta_N \text{Corr}(Z_{n+3-(N+1)}, Y_{n-N}) \\ &= \delta_{N-2}(1 - \rho) + \delta_{N-1} \rho \alpha_1 \text{Corr}(Z_{n+2-(N+1)}, Y_{n-(N-1)}) \\ &\quad + \delta_N [\rho \alpha_1 \text{Corr}(Z_{n+2-(N+1)}, Y_{n-N}) + \rho \alpha_2 \text{Corr}(Z_{n+1-(N+1)}, Y_{n-N})] \\ &= \delta_{N-2}(1 - \rho) + \rho \alpha_1 r_B(2) + \rho \alpha_2 r_B(1). \end{aligned}$$

Y en general se puede comprobar que

$$\begin{aligned} r_B(k) &= \delta_{N-(k-1)}(1 - \rho) + \rho \alpha_1 r_B(k-1) + \rho \alpha_2 r_B(k-2) + \dots + \rho \alpha_{k-1} r_B(1) + \rho \alpha_k r_B(0) \\ &= \delta_{N-(k-1)} + \sum_{i=1}^k \rho \alpha_i r_B(k-i), \quad k \geq 1, \end{aligned}$$

donde $\alpha_i = 0$ si $i > p$, y $\delta_j = 0$ si $j < 0$.

Ahora,

$$\begin{aligned} \text{Cov}(X_n, X_{n+k}) &= \beta^2 (E[Y_{n-D_n} Y_{n+k-D_{n+k}}] - E[Y_{n-D_n}] E[Y_{n+k-D_{n+k}}]) \\ &\quad + \beta(1 - \beta) (E[Y_{n-D_n} Z_{n+k-(N+1)}] - E[Y_{n-D_n}] E[Z_{n+k-(N+1)}]) \\ &\quad + \beta(1 - \beta) (E[Y_{n+k-D_{n+k}} Z_{n-(N+1)}] - E[Y_{n+k-D_{n+k}}] E[Z_{n-(N+1)}]) \\ &\quad + (1 - \beta)^2 (E[Z_{n+k-(N+1)} Z_{n-(N+1)}] - E[Z_{n+k-(N+1)}] E[Z_{n-(N+1)}]). \end{aligned}$$

El tercer término es cero pues $n + k - D_{n+k} > n - (N + 1)$ para toda $k \geq 0$.

Mientras que el primero, si $1 \leq k \leq N$

$$\beta^2(E[Y_{n-D_n}Y_{n+k-D_{n+k}}] - E[Y_{n-D_n}]E[Y_{n+k-D_{n+k}}]) = \beta^2 \sum_{i=0}^{N-k} \delta_i \delta_{i+k} \text{Var}(Y_1).$$

Por lo tanto, $\{r(k)\}$ con $r(k) = \text{Corr}(X_n, X_{n+k})$ puede ser obtenido mediante el sistema de ecuaciones

$$r(1) = \beta^2 \sum_{i=0}^{N-1} \delta_i \delta_{i+1} + \beta(1 - \beta)r_B(1) + (1 - \beta)^2 r_A(1)$$

$$r(2) = \beta^2 \sum_{i=0}^{N-2} \delta_i \delta_{i+2} + \beta(1 - \beta)r_B(2) + (1 - \beta)^2 r_A(2)$$

⋮

$$r(N) = \beta^2 \delta_0 \delta_N + \beta(1 - \beta)r_B(N) + (1 - \beta)^2 r_A(N)$$

$$r(N + k) = \beta(1 - \beta)r_B(N + k) + (1 - \beta)^2 r_A(N + k), \quad k \geq 1.$$

Las correlaciones de este proceso al igual que en el caso del modelo $\text{DAR}(p)$ son todas no negativas de hecho se puede usar exactamente la misma prueba.

3.2.4. NDARMA(p, N)

Una mezcla de un número finito de v.a.'s cada una con función de densidad de probabilidad π , tiene función de densidad de probabilidad π aún si las v.a.'s son dependientes. Basados en esto Jacobs y Lewis (1983) presentaron un Nuevo modelo en el que, a diferencia del proceso $\text{DARMA}(p, N + 1)$, la autoregresión no está definida sobre una cola autoregresiva sino sobre sí misma:

Sea $\{V_n\}, \{A_n\}$ y $\{D_n\}$ como antes. Definimos el proceso $\{X_n\}$ NDARMA($p, N + 1$) como

$$X_n = V_n X_{n-A_n} + (1 - V_n) Y_{n-D_n}, \quad (3.43)$$

$$n = 1, 2, \dots$$

Entonces con probabilidad ρ , X_n es uno de los p valores anteriores $X_{n-1}, X_{n-2}, \dots, X_{n-p}$ y con probabilidad $(1 - \rho)$ es una mezcla de las Y_k 's, $n - N \leq k \leq n$.

Obsérvese que si $\rho = 0$ entonces $\{X_n\}$ es el proceso $\text{DMA}(N)$ y si $\mathbb{P}[D_n = 0] = 1$ entonces $\{X_n\}$ es el proceso $\text{DAR}(p)$.

Estacionariedad Jacobs y Lewis (1983) prueban que el proceso NDARMA(p, N) es estacionario con distribución marginal π .

Correlación serial La correlación serial para un proceso estacionario NDARMA(p, N) satisface las ecuaciones de Yule-Walker para un proceso ARMA(p, N) con restricciones en el rango de los coeficientes.

Sea $r_B(i) = \text{Corr}(X_n, Y_{n-i})$, la cual se puede calcular recursivamente mediante

$$r_B(0) = \text{Corr}(X_n, Y_n) = \rho \text{Corr}(X_{n-A_n}, Y_n) + (1 - \rho) \text{Corr}(Y_{n-D_n}, Y_n) = (1 - \rho) \delta_0.$$

$$\begin{aligned} r_B(1) &= \text{Corr}(X_n, Y_{n-1}) = \rho \text{Corr}(X_{n-A_n}, Y_{n-1}) + (1 - \rho) \text{Corr}(Y_{n-D_n}, Y_{n-1}) \\ &= \rho \alpha_1 \text{Corr}(X_{n-1}, Y_{n-1}) + (1 - \rho) \delta_1 \\ &= \rho \alpha_1 r_B(0) + (1 - \rho) \delta_1. \end{aligned}$$

$$\begin{aligned} r_B(2) &= \text{Corr}(X_n, Y_{n-2}) = \rho \text{Corr}(X_{n-A_n}, Y_{n-2}) + (1 - \rho) \text{Corr}(Y_{n-D_n}, Y_{n-2}) \\ &= \rho \alpha_1 \text{Corr}(X_{n-1}, Y_{n-2}) + \rho \alpha_2 \text{Corr}(X_{n-2}, Y_{n-2}) + (1 - \rho) \delta_2 \\ &= \rho \alpha_1 r_B(1) + \rho \alpha_2 r_B(0) + (1 - \rho) \delta_2. \end{aligned}$$

En general, es sencillo ver que

$$r_B(i) = \rho \alpha_1 r_B(i-1) + \rho \alpha_2 r_B(i-2) + \cdots + \rho \alpha_i r_B(0) + (1 - \rho) \delta_i,$$

donde $\alpha_i = 0$ si $i > p$, y $\delta_j = 0$ si $j > N$.

Sea $r(k) = \text{Corr}(X_n, X_{n+k})$

$$= \rho \sum_{i=1}^p \alpha_i \text{Corr}(X_n, X_{n+k-i}) + (1 - \rho) \sum_{i=k}^N \delta_i \text{Corr}(X_n, Y_{n+k-i}),$$

entonces, para $1 \leq k \leq N$

$$r(k) = \rho \alpha_1 r(k-1) + \rho \alpha_2 r(k-2) + \cdots + \rho \alpha_p r(k-p) + (1 - \rho) \sum_{i=k}^N \delta_i r_B(i-k), \quad (3.44)$$

con $r(0) = 1$.

Cuando $k > N$ el último término desaparece.

Las correlaciones seriales del proceso NDARMA(p, N) son todas no negativas. El argumento es el mismo que se usó para el modelo DAR(p).

NDARMA(1,1)

El proceso $\{X_n\}$ NADRMA(1,1) viene dado por

$$X_n = \begin{cases} X_{n-1} & \text{c.p. } \rho, \\ Y_n & \text{c.p. } (1 - \rho)\delta_0, \\ Y_{n-1} & \text{c.p. } (1 - \rho)(1 - \delta_0), \end{cases}$$

con $0 \leq \delta_0 \leq 1$ y $0 \leq \rho < 1$.

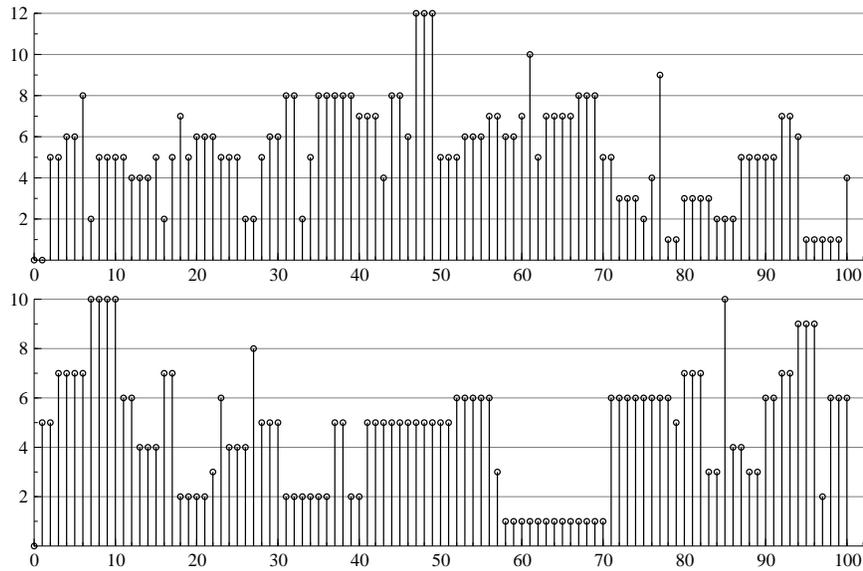


Figura 3.3: Simulación de un proceso Poisson NDARMA(1), con $\lambda = 5$, $n = 100$ y $\delta_0 = 0.5$. En la parte superior: $\rho = 0.3$. En la parte inferior: $\rho = 0.7$.

En este caso $\alpha_1 = 1$, entonces por (3.44)

$$\begin{aligned} r(1) &= \rho + (1 - \rho)\delta_1 r_B(0) \\ &= \rho + (1 - \rho)^2 \delta_0 (1 - \delta_0) \end{aligned}$$

y

$$r(k) = \rho r(k - 1), \quad k > 1.$$

Por lo tanto

$$r(k) = \text{Corr}(X_n, X_{n+k}) = \rho^{k-1} [\rho + (1 - \rho)^2 \delta_0 (1 - \delta_0)], \quad k > 0.$$

Las figuras 3.3 y 3.4 muestran las simulaciones de procesos NDARMA(1,1) con marginal Poisson y Binomial respectivamente.

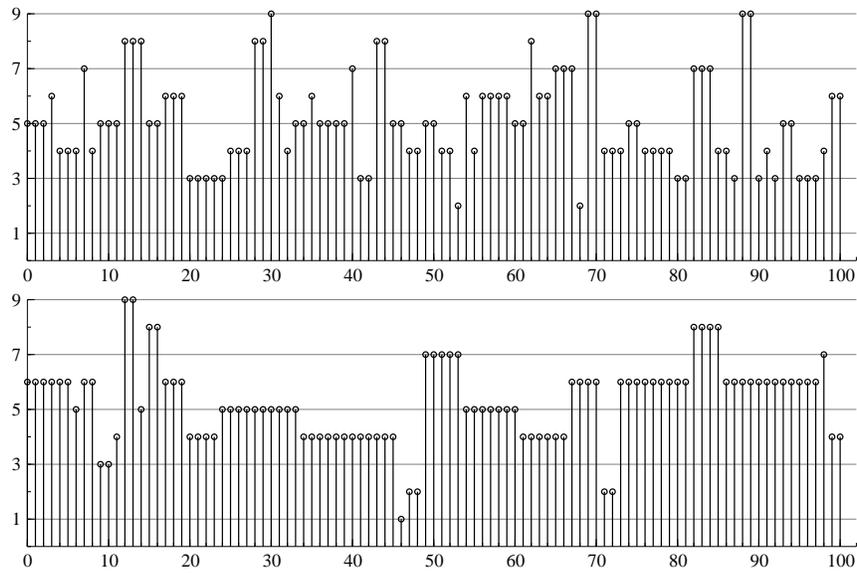


Figura 3.4: Simulación de un proceso Binomial NDARMA(1), con $N = 10$, $p = 0.5$, $n = 100$ y $\delta_0 = 0.5$. En la parte superior: $\rho = 0.3$. En la parte inferior: $\rho = 0.7$.

Capítulo 4

Sobre descomponible-por sí misma

Este capítulo pretende servir como un puente entre nuestro capítulo anterior y el siguiente. Ilustramos el concepto de *descomponible-por sí misma* (S-D) para variables aleatorias en \mathbb{R}^+ . Concepto que de manera implícita fue usado en el capítulo anterior. Es un resultado bien sabido que una distribución (no degenerada) descomponible-por sí misma es absolutamente continua.

Introducimos el *operador de adelgazamiento Binomial* que nos permitirá dar un concepto de descomponibilidad-por sí misma para variables aleatorias con valores en los enteros no negativos. Fue a partir de este operador de donde surgió la posibilidad de crear modelos para series de tiempo con valores discretos con estructura semejante a la de sus análogos continuos AR y MA, sobre todo de orden uno. Modelos que serán abordados en detalle en el siguiente capítulo.

La teoría de descomponibilidad-por sí misma es bastante extensa. Nos limitaremos a presentar algunos conceptos y resultados que son básicos en nuestro estudio sin entrar en demasía en los detalles. Para una visión más profunda del tema exhortamos al lector a que consulte el texto de Steutel y van Harn (2004).

4.1. Descomponibilidad-por sí misma

Definición 4.1 Una variable aleatoria X se dice que es *descomponible-por sí misma* (S-D) si **para todo** $\alpha \in (0, 1)$ existe una v.a. X_α tal que

$$X \stackrel{d}{=} \alpha X + X_\alpha, \quad (4.1)$$

y en el lado derecho X y X_α son independientes.

Claramente descomponibilidad por sí misma es una propiedad de la distribución de X , por lo que ésta también recibe dicho nombre.

Obviamente toda distribución degenerada es S-D. Además se sabe que toda distribución S-D es unimodal.

Si X toma valores en \mathbb{R}^+ entonces sus componentes X_α también. En este caso usaremos transformadas de Laplace para escribir (4.1).

Definición 4.2 Sea X una v.a. con valores en \mathbb{R}^+ y transformada de Laplace Φ_X . X (o su distribución) es descomponible-por sí misma si y solo si para todo $\alpha \in (0, 1)$, existe una transformada de Laplace Φ_α tal que

$$\Phi_X(s) = \Phi_X(\alpha s)\Phi_\alpha(s). \quad (4.2)$$

Ya que la transformada de Laplace (T.L.) es siempre positiva, la relación (4.2) puede ser escrita como

$$\Phi_\alpha(s) = \frac{\Phi_X(s)}{\Phi_X(\alpha s)}. \quad (4.3)$$

Es claro entonces que las distribuciones de las componentes X_α de X en (4.1) están determinadas de manera única por X .

Conviene saber dos cosas sobre distribuciones descomponibles-por sí mismas:

1. La clase de distribuciones descomponibles-por sí mismas es una amplia subclase de las distribuciones *infinitamente-divisibles*.

Estas últimas se encuentran definidas de la siguiente manera

Definición 4.3 Una v.a. X se dice que es *infinito-divisible* (I-D) si **para todo** $n \in \mathbb{N}$ existen $X_{n,1}, X_{n,2}, \dots, X_{n,n}$ v.a.'s independientes tales que

$$X \stackrel{d}{=} X_{n,1} + \dots + X_{n,n}, \quad (4.4)$$

y $X_{n,j} \stackrel{d}{=} X_n$ para toda j y alguna X_n llamada el *factor de orden n -ésimo* de X .

El termino infinito-divisible también aplica a la distribución de X . Se sigue que una función de distribución F es I-D si y solo si para todo $n \in \mathbb{N}$, F es la n -ésima convolución de una función de distribución F_n con ella misma. Y si X posee T.L., Φ_X , entonces X es I-D si y solo si para toda $n \in \mathbb{N}$, Φ_X es la n -ésima potencia de una T.L. Φ_n :

$$F = F_n \otimes^{(n)}, \quad n \in \mathbb{N}; \quad \Phi_X(s) = \{\Phi_n(s)\}^n, \quad n \in \mathbb{N}, \quad (4.5)$$

donde \otimes denota convolución.

Así mismo se tiene de la definición que si X es I-D entonces cX también lo es, para cualquier $c \in \mathbb{R}$; y si X y Y son independientes e I-D ambas, entonces $X + Y$ es I-D, se dice entonces que la clase infinitamente-divisible se preserva bajo convolución.

De (4.5) se obtiene el siguiente resultado:

Proposición 4.4 Una v.a. X con T.L. Φ_X es infinito-divisible si y solo si Φ_X^t es una T.L. para todo $t \in (0, \infty)$.

Ejemplo 4.5 Sea $\mu \in \mathbb{R}$ y $\sigma > 0$. Si X tiene distribución $\text{Normal}(\mu, \sigma^2)$, X es infinito-divisible. Para $n \in \mathbb{N}$ el factor de orden n -ésimo X_n tiene distribución $\text{Normal}(\mu/n, \sigma^2/n)$. \square

Ejemplo 4.6 Sea $k > 0$ y $\lambda > 0$. Si X tiene distribución $\text{Gamma}(k, \lambda)$, con función de densidad

$$f_X(x) = \frac{\lambda^k}{\Gamma(k)} x^{k-1} e^{-\lambda x}, \quad (4.6)$$

X es infinito-divisible. Para $n \in \mathbb{N}$ el factor de orden n -ésimo X_n tiene distribución $\text{Gamma}(k/n, \lambda)$. Obsérvese que si $k = 1$ se tiene que la distribución $\text{Exponencial}(\lambda)$ es infinito-divisible. \square

2. Toda distribución (no degenerada) descomponible-por sí misma es **absolutamente continua** (Steutel y van Harn (2004), pgs. 235 y 278), por lo que una variable aleatoria discreta (con valores en \mathbb{N}_0) no podrá serlo.

¿Habrà alguna manera anàloga a (4.1) para definir descomponibilidad-por sí misma para v.a.'s con valores en los enteros no negativos?, y en caso afirmativo, ¿qué significará ello?

Antes de responder a estas interrogantes presentamos el siguiente ejemplo para ilustra el uso de S-D.

4.1.1. GAR(1)

Consideremos nuestro proceso estacionario AR(1)

$$X_n = \rho X_{n-1} + \varepsilon_n, \quad n = 0, \pm 1, \pm 2, \dots, \quad (4.7)$$

con $|\rho| \leq 1$ y, $\{\varepsilon_n\}$ una sucesión de v.a.'s i.i.d.

En este apartado nos planteamos el siguiente problema: ¿es posible obtener un proceso estacionario con distribución marginal $\text{Gamma}(k, \lambda)$ que cumpla con el modelo (4.7)?

En el capítulo anterior hemos visto que si pretendemos que las $\{X_n\}$ tengan distribución marginal positiva es necesario pedir $\rho \in [0, 1)$. Además, encontramos que si suponíamos estacionariedad y dábamos la distribución para las X_n 's la ecuación que determinaría la distribución de las ε_n , es

$$\Phi_\varepsilon(s) = \frac{\Phi_X(s)}{\Phi_X(\rho s)}. \quad (4.8)$$

Que esta última sea la transformada de Laplace de una distribución (para todo $\rho \in (0, 1)$) no es otra cosa sino saber si X es descomponible-por sí misma. Se tiene entonces que:

Todas las distribuciones que sean descomponibles-por sí mismas pueden ser distribuciones marginales para la solución estacionaria de la ecuación (4.7).

Queda resuelta así, la primera pregunta que se planteó en este trabajo (pg. 20).

Nuestro problema es ahora, saber si la distribución $\text{Gamma}(k, \lambda)$ es S-D y en tal caso deducir cómo son las ε_n .

La transformada de Laplace para una variable aleatoria X , con función de densidad como en (4.6), está dada por $\Phi_X(s) = \left(\frac{\lambda}{\lambda+s}\right)^k$, sustituyendo en (4.8) se tiene,

$$\Phi_\varepsilon(s) = \left(\frac{\lambda + \rho s}{\lambda + s}\right)^k = \left(\rho + (1 - \rho)\frac{\lambda}{\lambda + s}\right)^k. \quad (4.9)$$

Por el Ejemplo III.11.8 de Steutel y van Harn (2004), con $\rho = \lambda/(\lambda+a)$, se tiene que $(\lambda + \rho s)/(\lambda + s)$ es una T.L. infinitamente-divisible. Lo que asegura, Proposición 4.4, que $\left(\frac{\lambda + \rho s}{\lambda + s}\right)^k$ es una transformada de Laplace para cualquier k positiva. Por lo tanto X es descomponible-por sí misma.

Tenemos entonces la mitad del problema, tal vez menos, resuelto. Ahora la interrogante es: ¿cuál es la distribución que está implícita en (4.9)? Uno puede ver que cuando $k \in \{1, 2, 3, \dots\}$, ε es una mezcla finita de distribuciones Gamma. Por ejemplo cuando $k = 2$,

$$\Phi_\varepsilon(s) = \rho^2 + 2\rho(1 - \rho)\frac{\lambda}{\lambda + s} + (1 - \rho)^2 \left(\frac{\lambda}{\lambda + s}\right)^2, \quad (4.10)$$

que es una v.a. que con probabilidad ρ^2 toma el valor 0, con probabilidad $2\rho(1 - \rho)$ es una v.a. Exponencial(λ) y con $(1 - \rho)^2$ una v.a. Gamma($\lambda, 2$).

¿Y cuándo k no es un entero -positivo-? Lawrance (1982) mostró que la variable aleatoria de innovación, ε_n , con transformada de Laplace (4.9), puede ser generada como

$$\varepsilon_n = \sum_{r=1}^N \rho^{U_r} Y_r, \quad \text{para cada } n, \quad (4.11)$$

donde

- $\{Y_r\}$ es una sucesión de v.a.'s i.i.d. con distribución Exponencial(λ),
- $\{U_r\}$ una sucesión de v.a.'s i.i.d. con distribución Uniforme(0,1),
- N una v.a. Poisson con media $k \ln(1/\rho)$,
- $\{Y_r\}$, $\{U_r\}$ y N , son independientes.

Proposición 4.7 La transformada de Laplace de (4.11) es (4.9).

Demostración: Haciendo uso del Teorema 2.23, se tiene que la distribución de $Z = \rho^U$, con $\rho \in (0, 1)$ y U una v.a. Uniforme(0,1) es

$$f_Z(z) = -\frac{1}{z} \frac{1}{\ln(\rho)}, \quad z \in (\rho, 1). \quad (4.12)$$

Entonces si Y es una v.a. Exponencial(λ) independiente de U , y $s > 0$

$$\begin{aligned}
E[e^{-sZY}] &= - \int_{\rho}^1 \int_0^{\infty} e^{-szy} \frac{1}{z} \frac{1}{\ln(\rho)} \lambda e^{-\lambda y} dy dz \\
&= - \frac{1}{\ln(\rho)} \int_{\rho}^1 \frac{\lambda}{z} \int_0^{\infty} e^{-y(\lambda+sz)} dy dz \\
&= - \frac{1}{\ln(\rho)} \int_{\rho}^1 \frac{\lambda}{z} \frac{1}{\lambda+sz} dz = - \frac{1}{\ln(\rho)} \ln \left[\frac{z(\lambda+s)}{\lambda+sz} \right] \Bigg|_{\rho}^1 \\
&= \frac{\ln \left[\frac{\rho(\lambda+s)}{\lambda+\rho s} \right]}{\ln(\rho)} = \frac{\ln(\rho) - \ln(\lambda+\rho s) + \ln(\lambda+s)}{\ln(\rho)} \\
&= 1 - \frac{\ln \left(\frac{\lambda+\rho s}{\lambda+s} \right)}{\ln(\rho)}. \tag{4.13}
\end{aligned}$$

La T.L. de una v.a. Poisson compuesta es un cálculo bastante sencillo y conocido, que se encuentra más abajo. En particular para ε_n como en (4.11),

$$\Phi_{\varepsilon_n}(s) = \Phi_N(-\ln\{\Phi_R(s)\}), \tag{4.14}$$

donde $R = ZY$, con Z y Y como arriba.

Sustituyendo (4.13) y recordando que la T.L. para una v.a. Poisson(h) es $\Phi_X(s) = e^{h(e^{-s}-1)}$. En particular si $h = -k \ln(\rho)$, $\Phi_X(s) = \exp\{-k \ln(\rho)[e^{-s} - 1]\}$. Así entonces,

$$\begin{aligned}
\Phi_{\varepsilon_n}(s) &= \exp\{-k \ln(\rho)(e^{\ln(\Phi_R(s))} - 1)\} \\
&= \exp\{-k \ln(\rho)(\Phi_R(s) - 1)\} = \exp \left[k \ln \left(\frac{\lambda + \rho s}{\lambda + s} \right) \right],
\end{aligned}$$

que es la misma que (4.9). \square

Acabamos de probar que el termino de innovación ε_n en (4.7) tiene distribución Poisson compuesta como en (4.11). Esta forma nos permite generar ε_n y por lo tanto el modelo GAR(1) (Gamma Autoregresivo de orden 1) para cualquier par de valores k y ρ , incluso cuando éstos son muy pequeños.

Por último, la T.L. de una v.a. Poisson compuesta se puede obtener de la siguiente manera.

Proposición 4.8 Sea $M = \sum_{i=1}^B W_i$, con $W_1, W_2 \dots$ v.a.'s i.i.d. y B una v.a. independiente de W_i para toda i , entonces la transformada de Laplace de M , $\Phi_M(t)$, está dada por

$$\Phi_M(t) = E [(\Phi_W(t))^B].$$

Demostración: $\Phi_M(t) = E[e^{-tM}] = E[e^{-t(W_1+W_2+\dots+W_B)}]$

$$\begin{aligned} &= E\left[\prod_{i=1}^B e^{-tW_i}\right] = E\left[E\left[\prod_{i=1}^B e^{-tW_i} \mid B\right]\right] \\ &= E\left[\prod_{i=1}^B E[e^{-tW_i} \mid B]\right] = E\left[\prod_{i=1}^B E[e^{-tW_i}]\right] \\ &= E\left[\prod_{i=1}^B \Phi_{W_i}(t)\right] = E[(\Phi_W(t))^B]. \end{aligned}$$

Que podemos reescribirlo como

$$\Phi_M(t) = E[(\Phi_W(t))^B] = E\left[e^{B \ln\{\Phi_W(t)\}}\right] = \Phi_B(-\ln\{\Phi_W(t)\}),$$

siempre que $0 < \Phi_W(t) < 1$, condición que asegura la positividad del argumento. En nuestro caso, ya que $\rho \in (0, 1)$ y $s > 0$, nos queda que $\Phi_R(s)$ cumple con esta condición. \square

Estacionariedad Por construcción el proceso es estacionario con distribución marginal Gamma(k, λ), siempre que X_0 tenga dicha distribución.

Correlación Serial Sabemos que $\Phi'_\varepsilon(0) = -E[\varepsilon]$. Calculando esta derivada se obtiene que $E[\varepsilon] = (1 - \rho)\frac{k}{\lambda}$.

Además como X_n se distribuye Gamma(k, λ) se tiene que $E[X_n] = \frac{k}{\lambda}$, por lo tanto del Corolario 2.17 se tiene que

$$\rho(k) = \rho^k, \quad k = 0, 1, \dots$$

Distribución Conjunta La transformada de Laplace de la distribución conjunta de X_n y X_{n+1} es

$$\begin{aligned} \Phi_{X_n, X_{n+1}}(s_1, s_2) &= E[\exp(-s_1 X_n - s_2 X_{n+1})] \\ &= \Phi_X(s_1 + \rho s_2) \Phi_\varepsilon(s_2) \\ &= \left(\frac{\lambda}{\lambda + s_1 + \rho s_2}\right)^k \left(\rho + (1 - \rho)\frac{\lambda}{\lambda + s_2}\right)^k \\ &= \left(\rho \frac{\lambda}{\lambda + s_1 + \rho s_2} + (1 - \rho)\frac{\lambda}{\lambda + s_1 + \rho s_2} \frac{\lambda}{\lambda + s_2}\right)^k. \end{aligned}$$

Ya que esta última es no simétrica en s_1 y s_2 , se sigue que el proceso es no reversible en el tiempo.

Esperanza y Varianza Condicionales Consideremos X_n y X_{n+1} .

$$E[X_{n+1}|X_n = x] = \rho x + E[\varepsilon_{n+1}] = \rho x + (1 - \rho)\frac{k}{\lambda},$$

$$E[X_{n+1}^2|X_n = x] = \rho^2 x^2 + 2\rho x E[\varepsilon_{n+1}] + E[\varepsilon_{n+1}^2].$$

Derivando dos veces la transformada de Laplace de ε_n (4.9), y evaluando en cero se obtiene el último sumando, quedando

$$E[X_{n+1}^2|X_n = x] = \rho^2 x^2 + 2\rho x(1 - \rho)\frac{k}{\lambda} + \frac{k}{\lambda^2}(\rho - 1)(k\rho - k - \rho - 1).$$

Así, la varianza condicional es

$$\text{Var}(X_{n+1}|X_n = x) = (1 - \rho^2)\frac{k}{\lambda^2},$$

que es una constante, igual a k veces la varianza condicional del proceso EAR(1).

4.2. Descomponible-por sí misma discreta

Definición 4.9 Sea X una variable aleatoria con valores en $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ y función de densidad $p_k (= P[X = k])$. La *función generadora de probabilidad* (f.g.p.) de X está definida como $P_X : [0, 1] \rightarrow [0, 1]$, con

$$P_X(s) := E[s^X] = \sum_{k=0}^{\infty} p_k s^k, \quad 0 \leq s \leq 1.$$

Es inmediato que si X y Y son v.a.'s independientes entonces $P_{X+Y} = P_X P_Y$. Existe el teorema de unicidad ($P_X = P_Y$ si y solo si $X \stackrel{d}{=} Y$). Además, la densidad p_k de X , puede obtenerse como

$$p_k = \frac{1}{k!} P_X^{(k)}(0), \quad k \in \mathbb{N}_0.$$

En la sección anterior nos planteamos la siguiente pregunta: ¿Podemos hablar de descomponibilidad-por sí misma para v.a.'s con valores en \mathbb{N}_0 ?, ¿Habría una forma análoga a (4.1) para definir descomponibilidad-por sí mismas para estas variables aleatorias? La respuesta es sí.

En 1979 Steutel y van Harn presentaron un trabajo en el que propusieron, y de hecho establecieron, el análogo discreto de S-D.

Sea X una v.a. con valores en \mathbb{N}_0 y $\alpha \in (0, 1)$. Para dar una definición de descomponible-por sí misma para X es necesario reemplazar la multiplicación escalar αX en (4.1) por $\alpha \circ X$, con

$$\alpha \circ X = Y_1 + Y_2 + \dots + Y_X,$$

donde Y_1, Y_2, \dots son v.a.'s i.i.d. Bernoulli de media α , e independientes de X . Si $X = 0$ la suma es cero.

Conviene definir $\alpha \circ X$ también para $\alpha \in \{0, 1\}$. De manera natural hacemos, $0 \circ X := 0$ y $1 \circ X := 1$. Obsérvese que $\alpha \circ X$, también es una v.a. con valores en \mathbb{N}_0 y definida en distribución por su f.g.p. como

$$P_{\alpha \circ X}(s) = P_X(1 - \alpha + \alpha s). \quad (4.15)$$

Algunos resultados como éste sobre el operador “ \circ ” se encuentran al final del capítulo.

Estamos ya listos para dar la definición de descomponibilidad-por sí misma caso discreto.

Definición 4.10 Una v.a. X que toma valores en \mathbb{N}_0 se dice que es *descomponible-por sí misma discreta* si para todo $0 < \alpha < 1$, existe una v.a. X_α tal que

$$X \stackrel{d}{=} \alpha \circ X + X_\alpha, \quad (4.16)$$

donde en el lado derecho $\alpha \circ X$, y X_α son independientes.

Más adelante presentamos la función de densidad de $\alpha \circ X$, ec. (4.21), de ésta (y puesto que $0 < \alpha < 1$) se obtiene que $\mathbb{P}[\alpha \circ X = 0] > 0$. Esto implica que las componentes X_α de X , también toman valores en \mathbb{N}_0 . Por lo tanto podemos usar la f.g.p. para reescribir (4.16).

Definición 4.11 Sea X una v.a. con valores en \mathbb{N}_0 y f.g.p. P_X . X (o su distribución) es S-D discreta si y solo si para todo $0 < \alpha < 1$, existe una f.g.p. P_α tal que

$$P_X(s) = P_X(1 - \alpha + \alpha s)P_\alpha(s). \quad (4.17)$$

Si definimos $G_X(s) := P_X(1 - s)$ y $G_\alpha(s) := P_\alpha(1 - s)$, (4.17) puede ser escrita como

$$G_X(s) = G_X(\alpha s)G_\alpha(s), \quad (4.18)$$

que es más similar a (4.2) para el caso en \mathbb{R}^+ . La demostración de (4.18) se encuentra al final del capítulo.

De la Proposición V.4.1. de Steutel y van Harn (2004) la componente X_α en (4.16) también toma el valor cero con probabilidad positiva. De esto se deduce un resultado interesantísimo: distribuciones degeneradas en \mathbb{N}_0 que son descomponibles-por sí mismas en el sentido clásico (Sección 3.1.) no pueden ser S-D discretas. Más aún, ya que distribuciones (no degeneradas) descomponibles-por sí mismas en \mathbb{R}^+ (y en todo \mathbb{R}), son absolutamente continuas, se tiene que distribuciones con valores en \mathbb{N}_0 no pueden ser ambas, descomponibles-por sí mismas discretas y descomponibles-por sí mismas en el sentido clásico.

Por otra parte en analogía a la sección anterior se sabe que:

Proposición 4.12 Una distribución descomponible-por sí misma discreta es infinito-divisible discreta.

Donde, una variable aleatoria X que toma valores en \mathbb{N}_0 , se dice que es *infinito-divisible discreta*, si cumple con la Definición 4.3 y además $\mathbb{P}[X = 0] > 0$. En términos de su f.g.p. se tiene

Proposición 4.13 Una v.a. X con valores en \mathbb{N}_0 y f.g.p. P_X , con $P_X(0) > 0$, es infinito-divisible discreta si y solo si P_X^t es una f.g.p. para todo $t \in (0, \infty)$.

Ejemplo 4.14 Sea $\lambda > 0$. Sea X una v.a. con distribución $\text{Poisson}(\lambda)$, $p_k = e^{-\lambda} \frac{\lambda^k}{k!}$, $k \in \mathbb{N}_0$. X es I-D discreta con factor de orden n -ésimo X_n con distribución $\text{Poisson}(\lambda/n)$. \square

Ejemplo 4.15 Para $r = 1, 2, \dots$ y $\theta \in (0, 1)$, sea X una v.a. con distribución Binomial Negativa(r, θ)

$$p_k = \binom{r+k-1}{k} \theta^k (1-\theta)^r, \quad k = 0, 1, \dots$$

X es I-D discreta con factor de orden n -ésimo X_n con distribución Binomial Negativa($r/n, \theta$). Si $r = 1$ obtenemos que la distribución Geométrica(θ) es I-D. \square

Retomando nuestro estudio sobre descomponibilidad-por sí misma caso discreto, reescribimos la ec. (4.17) como

$$P_\alpha(s) = \frac{P_X(s)}{P_X(1-\alpha + \alpha s)}, \quad (4.19)$$

y llamamos a esta función, la *función- P_α* de P_X .

Proposición 4.16 Sea X una v.a. con valores en \mathbb{N}_0 y f.g.p. P_X . X es S-D discreta si y solo si para toda $\alpha \in (0, 1)$, la función- P_α de P_X , es absolutamente monótona.

Donde, sea H una función que toma valores en \mathbb{R} y dominio $[0,1)$. Se dice que H es *absolutamente monótona* si es no negativa y posee derivadas no negativas de todos los ordenes, esto es

$$H^{(n)}(z) \geq 0, \quad n \in \mathbb{N}_0; \quad 0 \leq z < 1.$$

Considere el siguiente Teorema:

Teorema 4.17 Sea $H : [0, 1) \rightarrow \mathbb{R}^+$. H es absolutamente monótona si y solo si existe una sucesión $\{h_k\}_{k \in \mathbb{N}_0}$ de números no negativos tales que

$$H(z) = \sum_{k=0}^{\infty} h_k z^k, \quad 0 \leq z < 1.$$

A partir de este Teorema, la Proposición 4.16 es clara: La Definición 4.11 pide que la función- P_α , sea una f.g.p. Si pensamos en el Teorema 4.17 a la sucesión $\{h_k\}$ como probabilidades, éste nos dice que hay una equivalencia entre ser una f.g.p. y ser una función absolutamente monótona.

Ejemplo 4.18 Sea X una v.a. con distribución Poisson(θ). Su f.g.p. es fácilmente calculada como

$$P_X(s) = e^{-\theta(1-s)}.$$

Para $\alpha \in (0, 1)$ la función- P_α de P_X queda como

$$P_\alpha(s) = e^{-\theta(1-\alpha)(1-s)}.$$

Se tiene $P_\alpha^{(n)}(s) = [\theta(1-\alpha)]^n \exp\{-\theta(1-\alpha)(1-s)\} > 0$, para todo $0 < \alpha < 1$. Por lo tanto P_α es absolutamente monótona y X es S-D discreta. \square

Ejemplo 4.19 Sea X una v.a. con distribución Binomial Negativa(r, θ).

$$\begin{aligned} E[s^X] &= \sum_{k=0}^{\infty} s^k \binom{r+k-1}{k} \theta^k (1-\theta)^r \\ &= (1-\theta)^r \sum_{k=0}^{\infty} \binom{r+k-1}{k} (\theta s)^k = (1-\theta)^r \frac{1}{(1-\theta s)^r}, \end{aligned}$$

i.e.

$$P_X(s) = \left(\frac{1-\theta}{1-\theta s} \right)^r.$$

Para $\alpha \in (0, 1)$, la función- P_α de P_X , queda como

$$P_\alpha(s) = \left(\frac{1-\theta(1-\alpha+\alpha s)}{1-\theta s} \right)^r = \left[\alpha + (1-\alpha) \frac{1-\theta}{1-\theta s} \right]^r.$$

Del Ejemplo II.11.15 de Steutel y van Harn (2004), P_α es la r -ésima potencia de una f.g.p. infinitamente-divisible, puesto que $\alpha \in (0, 1)$. Lo que asegura que $\left(\frac{1-\theta(1-\alpha+\alpha s)}{1-\theta s} \right)^r$ es una f.g.p. para toda r positiva, Proposición 4.13. Por lo tanto X es S-D discreta. \square

Cabe mencionar que una distribución (no degenerada) acotada no podrá ser descomponible-por sí misma ni en el sentido clásico ni en el discreto. Esto se deriva del siguiente resultado.

Proposición 4.20 Una variable aleatoria no degenerada acotada no puede ser infinito-divisible (discreta).

Demostración: Sea X una v.a. tal que $|X| \leq a$ para alguna $a \in \mathbb{R}$ y supongamos que X es infinito-divisible. Entonces por (4.4), para $n \in \mathbb{N}$ el factor de orden n -ésimo X_n de X satisface $|X_n| \leq a/n$ y la Varianza de X puede ser estimada como

$$\text{Var}(X) = n \text{Var}(X_n) \leq n E[X_n^2] \leq a^2/n,$$

haciendo tender n a infinito se tiene que $\text{Var}(X)=0$, por lo tanto X es degenerada. \square

4.2.1. Operador de adelgazamiento Binomial

Definición 4.21 Sea X una v.a. con valores en los enteros no negativos. El *operador de adelgazamiento Binomial* “ \circ ” está definido como

$$\alpha \circ X = \begin{cases} \sum_{j=1}^X Y_j & \text{si } X > 0, \\ 0 & \text{si } X = 0, \end{cases}$$

con $\{Y_j\}$ una sucesión de v.a.'s i.i.d. Bernoulli con $P[Y_j = 1] = 1 - P[Y_j = 0] = \alpha$, e independientes de X .

La sucesión Y_1, Y_2, \dots es llamada la *serie de conteo* de $\alpha \circ X$.

Además $\alpha \circ X$, está definida en distribución por su f.g.p.

Proposición 4.22 La f.g.p. de $\alpha \circ X$, es $P_{\alpha \circ X}(s) = P_X(1 - \alpha + \alpha s)$.

Demostración:

$$\begin{aligned} E[s^{Y_1 + \dots + Y_X}] &= E\left[\prod_{j=1}^X s^{Y_j}\right] = E\left[E\left[\prod_{j=1}^X s^{Y_j} \mid X\right]\right] \\ &= E\left[\prod_{j=1}^X E[s^{Y_j} \mid X]\right] = E\left[\prod_{j=1}^X E[s^{Y_j}]\right] = E\left[\left(E[s^Y]\right)^X\right] \\ &= E\left[\left(P_Y(s)\right)^X\right] = P_X(P_Y(s)) \end{aligned}$$

y

$$P_Y(s) = E[s^Y] = \sum_{y=0,1} s^y \alpha^y (1 - \alpha)^{1-y} = 1 - \alpha + \alpha s. \quad \square \quad (4.20)$$

Proposición 4.23 Dada X , $\alpha \circ X$ se distribuye Binomial (X, α) .

Demostración: Intuitivamente el resultado es claro. Basta recordar que la suma de v.a.'s Bernoulli es Binomial. Una prueba es la siguiente.

Si X es fija, digamos x , entonces $\mathbb{P}[X = x] = 1$, y por lo tanto $P_X(s) = E[s^X] = s^x$, y

$$P_{\alpha \circ X}(s) = P_X(1 - \alpha + \alpha s) = (1 - \alpha + \alpha s)^x.$$

De (4.20) se sigue que esta es la f.g.p. de una Binomial (x, α) . \square

Debido al resultado anterior es que el operador “ \circ ”, suele recibir el nombre de *adelgazamiento Binomial*.

La v.a. $\alpha \circ X$ también toma valores en \mathbb{N}_0 y su distribución puede obtenerse como

$$\mathbb{P}[\alpha \circ X = k] = \sum_{n=k}^{\infty} \mathbb{P}[\alpha \circ X = k \mid X = n] \mathbb{P}[X = n]$$

$$= \sum_{n=k}^{\infty} \binom{n}{k} \alpha^k (1-\alpha)^{n-k} p_n. \quad (4.21)$$

En particular, $\mathbb{P}[\alpha \circ X = 0] = p_0 + (1-\alpha)p_1 + (1-\alpha)^2 p_2 + \dots$

Es claro que $p_i > 0$, para al menos un i , y puesto que $\alpha < 1$, se tiene $\mathbb{P}[\alpha \circ X = 0] > 0$, como afirmamos en la subsección anterior.

Proposición 4.24 Algunas propiedades del operador de adelgazamiento Binomial

1. $0 \circ X = 0$.
2. $1 \circ X = X$.
3. $\alpha \circ (\beta \circ X) \stackrel{d}{=} (\alpha\beta) \circ X$, con α y β tales que $0 \leq \alpha\beta \leq 1$.
4. $E[\alpha \circ X] = \alpha E[X]$.
5. $E[(\alpha \circ X)^2] = \alpha^2 E[X^2] + \alpha(1-\alpha)E[X]$.
6. $E[(\alpha \circ X)^3] = \alpha^3 E[X^3] + 3\alpha^2(1-\alpha)E[X^2] + \alpha(1-\alpha)(1-2\alpha)E[X]$.
7. $E[X(\alpha \circ Z)] = \alpha E[XZ]$.
8. $E[X(\alpha \circ Z)^2] = \alpha^2 E[XZ^2] + \alpha(1-\alpha)E[XZ]$.
9. Si X y Z son independientes, $E[(\alpha \circ X)(\beta \circ Z)] = \alpha\beta E[X]E[Z]$.
10. $E[(\alpha \circ X)(\beta \circ Z)] = \alpha\beta E[XZ]$ si las series de conteo de $\alpha \circ X$ y $\beta \circ Z$ son independientes, e independientes de X y Z .
11. $E[(\alpha \circ X)^2(\beta \circ Z)] = \alpha^2\beta E[X^2Z] + \alpha(1-\alpha)\beta E[XZ]$ si las series de conteo de $\alpha \circ X$ y $\beta \circ Z$ son independientes, e independientes de X y Z .
12. $E[XZ(\beta \circ W)] = \beta E[XZW]$.
13. $E[X(\beta \circ Z)(\gamma \circ W)] = \beta\gamma E[XZW]$ si las series de conteo de $\beta \circ Z$ y $\gamma \circ W$ son independientes, e independientes de X , Z y W .
14. $E[(\alpha \circ X)(\beta \circ Z)(\gamma \circ W)] = \alpha\beta\gamma E[XZW]$ si las series de conteo de $\alpha \circ X$, $\beta \circ Z$ y $\gamma \circ W$ son independientes, e independientes de X , Z y W .
15. $\text{Var}(\alpha \circ X) = \alpha^2 \text{Var}(X) + \alpha(1-\alpha)E[X]$.
16. Si X y Z son independientes $\text{Cov}(\alpha \circ X, \beta \circ Z) = 0$.
17. Si las series de conteo de $\alpha \circ X$ y $\beta \circ Z$ son independientes, e independientes de X y Z , entonces, $\text{Cov}(\alpha \circ X, \beta \circ Z) = \alpha\beta \text{Cov}(X, Z)$.

18. $\text{Cov}(X, \alpha \circ Z) = \alpha \text{Cov}(X, Z)$.
19. $\text{Cov}(\alpha \circ X, \alpha \circ X) = \text{Var}(\alpha \circ X)$.
20. $\text{Cov}(X, \alpha \circ X) = \alpha \text{Var}(X)$.

Demostración:

1. Por definición. \square

2. Por definición. \square

3. Sabemos por la Proposición 4.22 que $P_{\alpha \circ X}(s) = P_X(1 - \alpha + \alpha s)$, entonces

$$P_{(\alpha\beta) \circ X}(s) = P_X(1 - \alpha\beta + \alpha\beta s).$$

Por otra parte se sabe que $\alpha \circ X$, también es v.a.. Entonces

$$P_{\alpha \circ (\beta \circ X)}(s)$$

$$= P_{\beta \circ X}(1 - \alpha + \alpha s) = P_X(1 - \beta + \beta\{1 - \alpha + \alpha s\}) = P_X(1 - \alpha\beta + \alpha\beta s). \quad \square$$

4.

$$E \left[\sum_{i=1}^X Y_i \right] = E \left[E \left[\sum_{i=1}^X Y_i \mid X \right] \right] = E \left[X E[Y] \right] = \alpha E[X]. \quad \square$$

5. Sea $A = (\alpha \circ X)^2 = \sum_{i=1}^X Y_i \sum_{j=1}^X Y_j$, que puede escribirse como

$$A = \begin{cases} Y_1^2 & + & Y_1 Y_2 & + & Y_1 Y_3 & + & \cdots & + & Y_1 Y_X \\ + & Y_2 Y_1 & + & Y_2^2 & + & Y_2 Y_3 & + & \cdots & + & Y_2 Y_X \\ + & Y_3 Y_1 & + & Y_3 Y_2 & + & Y_3^2 & + & \cdots & + & Y_3 Y_X \\ + & \cdots & & & & & & & & \\ + & Y_X Y_1 & + & Y_X Y_2 & + & Y_X^2 & + & \cdots & + & Y_X^2 \end{cases}$$

$$= \begin{cases} A_1 \\ + & A_2 \\ + & \cdots \\ + & A_X \end{cases}$$

Recordando que si $Y \sim \text{Bernoulli}(p)$, entonces $E[Y^n] = p$, para toda $n \in \mathbb{N}$, entonces $E[A_1|X] = \alpha + \alpha^2(X - 1)$. Además obsérvese que $E[A_i|X] = E[A_1|X]$ para toda i . Entonces

$$\begin{aligned} E[E[A|X]] &= E\left[E\left[\sum_{i=1}^X A_i \mid X\right]\right] = E\left[\sum_{i=1}^X E[A_i \mid X]\right] \\ &= E[X\{\alpha + \alpha^2(X - 1)\}] = E[\alpha^2 X^2 + \alpha(1 - \alpha)X] \\ &= \alpha^2 E[X^2] + \alpha(1 - \alpha)E[X]. \quad \square \end{aligned}$$

6. Sea $B = (\alpha \circ X)^3 = Y_1 A + Y_2 A + \dots + Y_X A$, con A como en 5.

Obsérvese que A puede ser reescrita como,

$$A = \sum_{i=1}^X Y_i^2 + 2Y_1 \sum_{i=2}^X Y_i + 2Y_2 \sum_{i=3}^X Y_i + 2Y_3 \sum_{i=4}^X Y_i + \dots + 2Y_{X-1} Y_X.$$

Consideremos el primer sumando de B , $Y_1 A$. Éste queda

$$Y_1^3 + Y_1 \sum_{i=2}^X Y_i^2 + 2Y_1^2 \sum_{i=2}^X Y_i + 2Y_1 Y_2 \sum_{i=3}^X Y_i + 2Y_1 Y_3 \sum_{i=4}^X Y_i + \dots + 2Y_1 Y_{X-1} Y_X.$$

Se tiene que

$$\begin{aligned} E[Y_1 A|X] &= \alpha + \alpha(X - 1)\alpha + 2\alpha(X - 1)\alpha + 2\alpha^2(X - 2)\alpha + 2\alpha^2(X - 3)\alpha + \dots + 2\alpha^2 1\alpha \\ &= \alpha + 3\alpha^2(X - 1) + 2\alpha^3 \left[\frac{(X - 2)(X - 1)}{2} \right] = \alpha + 3\alpha^2(X - 1) + \alpha^3(X^2 - 3X + 2) \end{aligned}$$

El lector puede comprobar que $E[Y_i A|X] = E[Y_1 A|X]$ para toda i , entonces aplicando la misma técnica que en 5,

$$\begin{aligned} E[(\alpha \circ X)^3] &= E[B] = E[X\{\alpha + 3\alpha^2(X - 1) + \alpha^3(X^2 - 3X + 2)\}] \\ &= \alpha^3 E[X^3] + 3\alpha^2(1 - \alpha)E[X^2] + \alpha(1 - \alpha)(1 - 2\alpha)E[X]. \quad \square \end{aligned}$$

7.

$$E\left[X \sum_{i=1}^Z Y_i\right] = E\left[X E\left[\sum_{i=1}^Z Y_i \mid Z\right]\right] = E[X Z E[Y]] = \alpha E[X Z].$$

Note que lo anterior es equivalente a sólo escribir

$$E[X(\alpha \circ Z)] = E\left[X E[\alpha \circ Z|Z]\right] = E[X \alpha Z] = \alpha E[X Z]. \quad \square$$

$$\begin{aligned}
8. \quad E[X(\alpha \circ Z)^2] &= E\left[XE[(\alpha \circ Z)^2|Z]\right] = E[X\{\alpha Z^2 + \alpha(1 - \alpha)Z\}] \quad (\text{por 5}) \\
&= \alpha^2 E[XZ^2] + \alpha(1 - \alpha)E[XZ]. \quad \square
\end{aligned}$$

$$\begin{aligned}
9. \quad E[(\alpha \circ X)(\beta \circ Z)] &= E\left[(\alpha \circ X)E[(\beta \circ Z)|Z]\right] \\
&= E[(\alpha \circ X)\beta Z] = E\left[\beta ZE[(\alpha \circ X)|X]\right] = \alpha\beta E[X]E[Z]. \quad \square
\end{aligned}$$

$$\begin{aligned}
10. \quad E[(\alpha \circ X)(\beta \circ Z)] &= E\left[(\alpha \circ X)E[(\beta \circ Z)|Z]\right] = E\left[\beta ZE[(\alpha \circ X)|X]\right] = \alpha\beta E[XZ]. \quad \square
\end{aligned}$$

11. Haciendo el mismo procedimiento que en 10 y usando 5 como en 8, se obtiene el resultado. \square

12. Usando esperanza condicional, condicionando sobre W , se obtiene el resultado. \square

13. Aplicando un procedimiento semejante al de 10 se obtiene el resultado. \square

14. De manera análoga a 10. \square

15. Es consecuencia de 4 y 5. \square

16. De 9 y 4. \square

17. Por 10 y 4. \square

$$18. \quad \text{Cov}(X, \alpha \circ Z) = E[X(\alpha \circ Z)] - E[X]E[\alpha \circ Z] = \alpha E[XZ] - \alpha E[X]E[Z] = \alpha \text{Cov}(X, Z). \quad \square$$

$$19. \text{Cov}(\alpha \circ X, \alpha \circ X) = E[(\alpha \circ X)^2] - E^2[\alpha \circ X] = \alpha^2 E[X^2] + \alpha(1 - \alpha)E[X] - \alpha^2 E^2[X] \\ = \alpha^2 \text{Var}(X) + \alpha(1 - \alpha)E[X] = \text{Var}(\alpha \circ X). \quad \square$$

20. Por 18, con $Z = X$. \square

Proposición 4.25 Si definimos $G_X(s) := P_X(1 - s)$ y $G_\alpha(s) := P_\alpha(1 - s)$ entonces, $G_{\alpha \circ X}(s) = G_X(\alpha s)$.

Demostración: Siguiendo el mismo procedimiento de la Proposición 4.22 se obtiene

$$E[(1 - s)^{\alpha \circ X}] = E\left[E[(1 - s)^{\alpha \circ X} \mid X]\right] = E\left[\left(E[(1 - s)^Y]\right)^X\right] \quad (4.22)$$

$$= E\left[(G_Y(s))^X\right] = E[(1 - \alpha s)^X] = G_X(\alpha s), \quad (4.23)$$

pues

$$G_Y(s) = E[(1 - s)^Y] = \sum_{y=0,1} (1 - s)^y \alpha^y (1 - \alpha)^{1-y} = 1 - \alpha s. \quad \square \quad (4.24)$$

A partir de este resultado, (4.18) es inmediata debido a la independencia en (4.16).

Capítulo 5

Modelos basados en adelgazamiento

El presente capítulo trata de modelos de series de tiempo con valores en los enteros no negativos, basados en *adelgazamiento*. Estudiamos concretamente las familias AR(1) discreta, MA(1) discreta y ARMA(1,1) discreta. Para la primera se ha dedicado un mayor espacio no solo en el presente trabajo sino en la literatura en general, tal es así que ésta recibe el nombre particular de INAR (INteger-valued Autoregressive) de orden uno, INAR(1). Para las tres familias estudiamos el caso de las distribuciones Poisson, Geométrica y Binomial como distribución marginal de los modelos. Para la distribuciones Poisson y Geométrica usamos adelgazamiento Binomial, “o”, el cual ha sido estudiado en detalle en el capítulo anterior. Para el caso de la distribución Binomial por la Proposición 4.20 será necesario hacer modificaciones y usar lo que llamaremos *adelgazamiento Hipergeométrico*.

Por otra parte, para la familia AR(1) discreta también estudiamos la distribución Binomial Negativa, presentando incluso dos modelos distintos con ésta distribución marginal. El primero se obtiene mediante adelgazamiento Binomial y el segundo por una generalización del mismo que denotamos por “•”.

5.1. INAR(1)

Consideremos el modelo AR(1) con distribución marginal no negativa dado por

$$X_n = \alpha X_{n-1} + Z_n, \quad n = 0, \pm 1, \pm 2, \dots \quad (5.1)$$

donde $\alpha \in [0, 1)$, y $\{Z_n\}$ una sucesión de v.a.'s i.i.d. no negativas.

De manera particular hemos estudiado los casos en que el modelo posee distribución marginal Exponencial y Gamma (modelo EAR(1) y GAR(1) respectivamente).

Sin embargo si requerimos que las X_n tomen valores en un espacio de estados discreto, en particular en \mathbb{N}_0 , el modelo (5.1) es inapropiado. Pues es claro que en general Z_n no puede ser independiente del valor de X_{n-1} , ya que el espacio donde X_n toma valores condiciona tal independencia. Además

hemos visto que las soluciones estacionarias para (5.1) están dadas por aquellas distribuciones que son descomponibles-por sí mismas, las cuales se sabe que son absolutamente continuas. No obstante tenemos un concepto de descomponibilidad-por sí misma caso discreto.

Definición 5.1 Sea X una v.a. con valores en \mathbb{N}_0 y f.g.p. P_X . X (o su distribución) se dice que es descomponible-por sí misma discreta (S-D discreta) si para todo $0 < \alpha < 1$,

$$P_Z(s) = \frac{P_X(s)}{P_X(1 - \alpha + \alpha s)}, \quad (5.2)$$

es una función generadora de probabilidad.

En términos de v.a.'s la ec. (5.2) se escribe como

$$X \stackrel{d}{=} \alpha \circ X + Z, \quad (5.3)$$

donde en el lado derecho $\alpha \circ X$ y Z son independientes, y Z es una v.a. en los enteros no negativos. Aquí $\alpha \circ X$ está definida (en distribución) por su f.g.p. $P_X(1 - \alpha + \alpha s)$, o como

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

donde X, Y_1, Y_2, \dots son independientes, y $\mathbb{P}[Y_i = 1] = 1 - \mathbb{P}[Y_i = 0] = \alpha$ para toda i ; si $X = 0$, la suma es cero. Además dado X , $\alpha \circ X$ se distribuye Binomial(X, α), es por ello que el operador “ \circ ”, recibe el nombre de adelgazamiento Binomial.

Ciertamente la ecuación (5.3) nos da un modelo de series de tiempo con valores discretos, que recibe en la literatura el nombre de INAR(1). Éste es

$$X_n = \alpha \circ X_{n-1} + Z_n, \quad n = 0, \pm 1, \pm 2, \dots \quad (5.4)$$

con $0 < \alpha < 1$, y $\{Z_n\}$ una sucesión de v.a.'s i.i.d. con valores en los enteros no negativos, e independientes de todas las series de conteo involucradas en (5.4), las cuales a su vez son mutuamente independientes. A este respecto sería más apropiado escribir $\alpha^{(n)} \circ$, en lugar de $\alpha \circ$, en la ecuación (5.4) para indicar explícitamente que hay una sucesión de v.a.'s, Y_1^n, Y_2^n, \dots , para cada n .

El modelo (5.4) puede interpretarse de la siguiente manera. Los componentes del proceso al tiempo n , X_n , son la suma de aquellos que sobreviven del tiempo $n - 1$ a n , i.e. $\alpha \circ X_{n-1}$, y de aquellos que arriban al sistema en el intervalo $(n - 1, n]$ como término de innovación Z_n .

Por último, nos será más conveniente usar la que ha sido llamada *función alternativa generadora de probabilidad* (f.a.g.p.), en lugar de la f.g.p., que está definida como

$$G_X(s) := P_X(1 - s) = E[(1 - s)^X].$$

En términos de dicha f.a.g.p. sabemos (Capítulo 3) que (5.2) queda como

$$G_Z(s) = \frac{G_X(s)}{G_X(\alpha s)}, \quad (5.5)$$

Por lo tanto si la sucesión $\{Z_n\}$ con f.a.g.p. $G_Z(s)$ dada por (5.5) es usada como innovación en (5.4) (y X_0 tiene f.a.g.p. $G_X(s)$), entonces el proceso $\{X_n\}$ tendrá distribución marginal estacionaria con f.a.g.p. $G_X(s)$.

A continuación presentamos dos resultados que serán de gran ayuda para el resto de la sección.

El modelo (5.4) es una cadena de Markov con matriz de probabilidades de transición $[\mathbb{P}]$, donde

$$\begin{aligned} p(x, y) &= \mathbb{P}[X_n = y | X_{n-1} = x] \\ &= \sum_k \mathbb{P}[\alpha \circ X_{n-1} = k | X_{n-1} = x] \mathbb{P}[Z_n = y - k] \\ &= \sum_{k=0}^{y \wedge x} \binom{x}{k} \alpha^k (1 - \alpha)^{x-k} \mathbb{P}[Z_n = y - k], \end{aligned} \quad (5.6)$$

donde $y \wedge x := \min\{y, x\}$.

Proposición 5.2 Sea $G_{X_n, X_{n+1}}(u, v) := P_{X_n, X_{n+1}}(1-u, 1-v) = E[(1-u)^{X_n} (1-v)^{X_{n+1}}]$, entonces suponiendo estacionariedad para el proceso INAR(1), (5.4), se tiene

$$G_{X_n, X_{n+1}}(u, v) = G_X(u + \alpha v - \alpha uv) \frac{G_X(v)}{G_X(\alpha v)}. \quad (5.7)$$

Demostración:

$$E[(1-u)^{X_n} (1-v)^{X_{n+1}}] = E\left[(1-u)^{X_n} (1-v)^{\alpha \circ X_n} (1-v)^{Z_{n+1}}\right]. \quad (5.8)$$

Calculamos $E[(1-u)^{X_n} (1-v)^{\alpha \circ X_n}]$,

$$\begin{aligned} E\left[(1-u)^{X_n} (1-v)^{\alpha \circ X_n}\right] &= E\left[(1-u)^{X_n} E\left[(1-v)^{\alpha \circ X_n} | X_n\right]\right] \\ &= E\left[(1-u)^{X_n} (1-\alpha v)^{X_n}\right] = E\left[(1-u - \alpha v + \alpha uv)^{X_n}\right], \end{aligned} \quad (5.9)$$

donde en la penúltima igualdad se usó, $E\left[(1-v)^{\alpha \circ X_n} | X_n\right] = \prod_{i=1}^{X_n} E\left[(1-v)^{Y_i}\right] = \prod_{i=1}^{X_n} (1-\alpha v) = (1-\alpha v)^{X_n}$, según vimos en (4.24).

Aplicando la independencia entre Z_{n+1} y X_n en (5.8), suponiendo estacionariedad y usando (5.5), se obtiene el resultado. \square

Estudiaremos ahora distintas distribuciones marginales para (5.4). Es por demás mencionar que éstas deben ser S-D discretas, si se desea que el proceso sea estacionario y con valores discretos.

5.1.1. Poisson

Supongamos que deseamos que el proceso INAR(1),

$$X_n = \alpha \circ X_{n-1} + Z_n, \quad \alpha \in (0, 1), \quad (5.10)$$

con $\{Z_n\}$ es una sucesión de v.a.'s i.i.d. con valores en \mathbb{N}_0 , e independiente de todas las operaciones de adelgazamiento en (5.10), las cuales a su vez se realizan de manera independiente, tenga distribución marginal Poisson(θ).

Puesto que la distribución Poisson es S-D discreta es posible encontrar Z_n tal que el proceso $\{X_n\}$ dado por (5.10) tenga dicha distribución marginal.

La f.a.g.p. de la distribución Poisson(θ) es fácilmente calculada

$$G_X(s) = E[(1-s)^X] = \sum_{x=0}^{\infty} (1-s)^x e^{-\theta} \frac{\theta^x}{x!} = e^{-\theta} e^{\theta(1-s)} = e^{-\theta s}.$$

Entonces la f.a.g.p. de la distribución de la innovación es

$$G_Z(s) = \frac{e^{-\theta s}}{e^{-\theta \alpha s}} = e^{-\theta(1-\alpha)s},$$

que es reconocida inmediatamente como la f.a.g.p. de una v.a. Poisson($\theta\bar{\alpha}$) con $\bar{\alpha} = (1-\alpha)$. Por lo tanto si la sucesión de v.a.'s i.i.d. $\{Z_n\}$ tiene distribución Poisson($\theta\bar{\alpha}$), y X_0 es una v.a independiente Poisson(θ), entonces las $\{X_n\}$ tendrán distribución marginal Poisson(θ).

Esperanza y varianza condicionales Es claro que

$$E[X_n|X_{n-1}] = E[\alpha \circ X_{n-1}|X_{n-1}] + E[Z_n] = \alpha X_{n-1} + (1-\alpha)\theta \quad (5.11)$$

y

$$\text{Var}(X_n|X_{n-1}) = \alpha(1-\alpha)X_{n-1} + (1-\alpha)\theta. \quad (5.12)$$

Correlación serial Puesto que la esperanza condicional $E[X_n|X_{n-1}]$ tiene la forma (5.11), del Corolario 2.17 se tiene

$$\rho(k) = \alpha^k, \quad k = 0, 1, \dots$$

Probabilidad condicional Usando (5.6) se obtiene

$$p(x, y) = x! e^{-\bar{\alpha}\theta} \sum_{k=0}^{y \wedge x} \frac{\alpha^k (1-\alpha)^{x+y-2k} \theta^{y-k}}{k!(x-k)!(y-k)!}, \quad y = 0, 1, 2, \dots$$

Distribución conjunta Usando (5.7) se tiene

$$G_{X_n, X_{n+1}}(u, v) = e^{-\theta(u+\alpha v-\alpha uv)} e^{-\theta(1-\alpha)v} = \exp\{-\theta(u+v-\alpha uv)\}. \quad (5.13)$$

Ya que la f.a.g.p. conjunta (5.13) es simétrica se sigue que el proceso Markoviano en cuestión es reversible en el tiempo.

A manera de ilustración la Figura 5.1 muestra la simulación de dos procesos INAR(1) con distribución marginal Poisson de media 5.

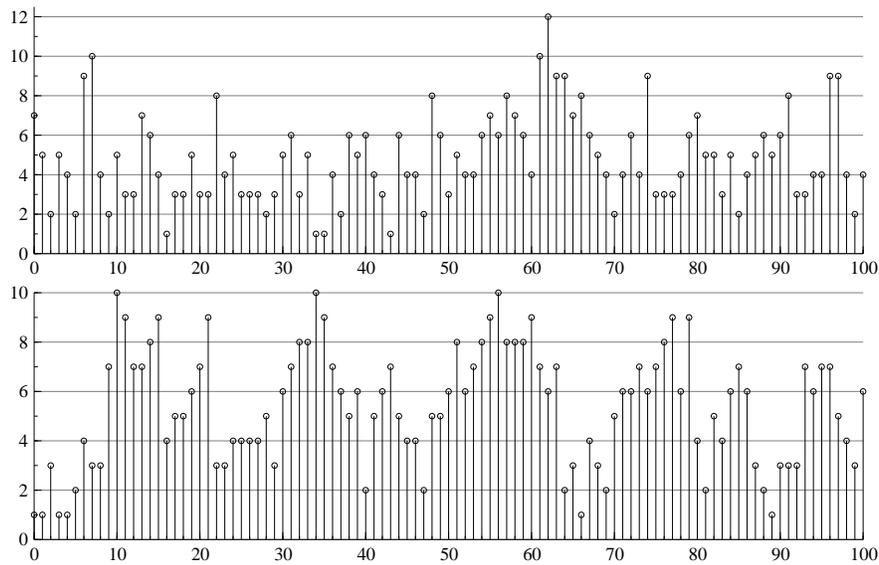


Figura 5.1: Simulación de un proceso Poisson INAR(1), con $\lambda = 5$ y $n = 100$. En la parte superior: $\alpha = 0.3$. En la parte inferior: $\alpha = 0.7$.

5.1.2. Geométrica

Gran parte del trabajo al que nos enfrentaremos al trabajar con esta distribución ya ha sido realizado en capítulos anteriores.

Como sabemos la distribución Geométrica, con función de densidad $p_X(k) = \theta^k(1-\theta)$, $k = 0, 1, \dots$ y $\theta \in (0, 1)$, denotada por $\text{Geom}(\theta)$, es S-D discreta.

La f.a.g.p. de esta distribución es

$$G_X(s) = E[(1-s)^X] = \sum_{x=0}^{\infty} (1-s)^x \theta^x (1-\theta) = \frac{1-\theta}{1-\theta+\theta s}.$$

Si en esta última hacemos $\lambda = \frac{1-\theta}{\theta}$, se tiene

$$G_X(s) = \frac{\lambda}{\lambda + s}. \quad (5.14)$$

¿Conocida? ¡Claro! Es idéntica a la transformada de Laplace de la distribución Exponencial(λ). La razón de la similitud entre la T.L. de una v.a. Exponencial y la f.a.g.p. de una distribución Geométrica es bien conocida, Proposición 5.9. Por lo tanto cualquier resultado distribucional para una combinación lineal de v.a.'s Exponenciales puede ser duplicado para v.a.'s Geométricas si el producto αX es reemplazado por $\alpha \circ X$. Por ejemplo, al comparar (5.5) con (3.3), se tiene que la distribución para la innovación, del proceso INAR(1) con distribución marginal $\text{Geom}(\theta)$, está dada por la mezcla

$$Z_n = \begin{cases} 0 & \text{c.p. } \alpha, \\ G_n & \text{c.p. } (1 - \alpha), \end{cases} \quad (5.15)$$

con $\{G_n\}$ una sucesión de v.a.'s i.i.d. $\text{Geom}(\theta)$.

Esperanza y varianza condicionales

$$\begin{aligned} E[X_n|X_{n-1}] &= \alpha X_{n-1} + E[Z_n] = \alpha X_{n-1} + (1 - \alpha)E[G_n] \\ &= \alpha X_{n-1} + (1 - \alpha) \left(\frac{\theta}{1 - \theta} \right) = \alpha X_{n-1} + (1 - \alpha) \frac{1}{\lambda}. \end{aligned} \quad (5.16)$$

En tanto que

$$\begin{aligned} \text{Var}(X_n|X_{n-1}) &= \alpha(1 - \alpha)X_{n-1} + \text{Var}(Z_n) \\ &= \alpha(1 - \alpha)X_{n-1} + (1 - \alpha) \frac{\theta}{(1 - \theta)^2} (1 + \alpha\theta). \end{aligned} \quad (5.17)$$

Para el cálculo de $\text{Var}(Z_n)$ obsérvese que $E[Z_n^2] = (1 - \alpha) \left[\frac{\theta}{(1 - \theta)^2} + \frac{\theta^2}{(1 - \theta)^2} \right]$.

Por (5.14) se obtuvo inmediatamente (5.15). Observemos que (5.16) tiene la misma estructura que (3.10), lineal en X_{n-1} . Sin embargo es de recalcar que (5.17) **no** comparte la estructura de (3.11). Esta última es una constante (independiente del valor de X_{n-1}), mientras que la primera es lineal en X_{n-1} . ¿A qué se debe tal diferencia si como ya señalamos la f.a.g.p. (5.14) y la transformada de Laplace de una v.a. Exponencial(λ) son la misma? Hay una gran diferencia entre (5.1) y (5.10). Así, αX_{n-1} es una constante cuando conocemos X_{n-1} , mientras que $\alpha \circ X_{n-1}$ es una v.a., Binomial, aún cuando el valor de X_{n-1} es conocido.

Correlación serial Puesto que $E[X_n|X_{n-1}]$ dada por (5.16) cumple con la estructura establecida en el Corolario 2.17, ya que $E[G_n] = E[X_n]$ para toda n , se tiene

$$\rho(k) = \alpha^k, \quad k = 0, 1, \dots$$

Probabilidad condicional Para calcular ésta usamos (5.6). Sin embargo hay que ser cuidadosos por la estructura que tiene la innovación Z_n .

Para $y = 0, 1, \dots, x$

$$\begin{aligned} p(x, y) &= \alpha \binom{x}{y} \alpha^y (1 - \alpha)^{x-y} + (1 - \alpha) \sum_{k=0}^y \binom{x}{k} \alpha^k (1 - \alpha)^{x-k} (1 - \theta) \theta^{y-k} \\ &= \binom{x}{y} \alpha^{y+1} (\bar{\alpha})^{x-y} + \bar{\alpha} \bar{\theta} \theta^{y-x} \sum_{k=0}^y \binom{x}{k} \alpha^k (\bar{\alpha} \theta)^{x-k}. \end{aligned}$$

Mientras que para $y = x + 1, x + 2, \dots$

$$p(x, y) = (1 - \alpha) \sum_{k=0}^x \binom{x}{k} \alpha^k (1 - \alpha)^{x-k} (1 - \theta) \theta^{y-k}.$$

Esta última se simplifica multiplicando por $\frac{\theta^x}{\theta^x}$:

$$p(x, y) = \bar{\alpha} \bar{\theta} \theta^{y-x} (\alpha + \bar{\alpha} \theta)^x.$$

Distribución conjunta Usando (5.7) y (5.14) se obtiene

$$G_{X_n, X_{n+1}}(u, v) = \frac{\lambda}{\lambda + u + \alpha v - \alpha uv} \frac{\lambda + \alpha v}{\lambda + v}. \quad (5.18)$$

Comparando está última con la transformada de Laplace (3.8) del modelo EAR(1),

$$\Phi_{X_n, X_{n+1}}(s_1, s_2) = \frac{\lambda}{\lambda + s_1 + \rho s_2} \frac{\lambda + \rho s_2}{\lambda + s_2},$$

observamos que esta última difiere de (5.18) únicamente por el termino cruzado αuv ($\rho s_1 s_2$). Además, puesto que (5.18) no es simétrica, el proceso es no reversible en el tiempo.

La figura 5.2 ilustra el comportamiento de un proceso INAR(1) con distribución marginal Geométrica.

5.1.3. Binomial Negativa

Para $r > 0$ y $\theta \in (0, 1)$, sea X una v.a. con función de densidad

$$p_X(k) = \binom{r+k-1}{k} \theta^k (1 - \theta)^r, \quad k = 0, 1, \dots, \quad (5.19)$$

denotada por $\text{BN}(r, \theta)$. Así,

$$\begin{aligned} G_X(s) &= \sum_{x=0}^{\infty} (1 - s)^x \binom{r+x-1}{x} \theta^x (1 - \theta)^r \\ &= (1 - \theta)^r \sum_{x=0}^{\infty} \binom{r+x-1}{x} (\theta - \theta s)^x = (1 - \theta)^r \frac{1}{[1 - (\theta - \theta s)]^r}. \end{aligned}$$

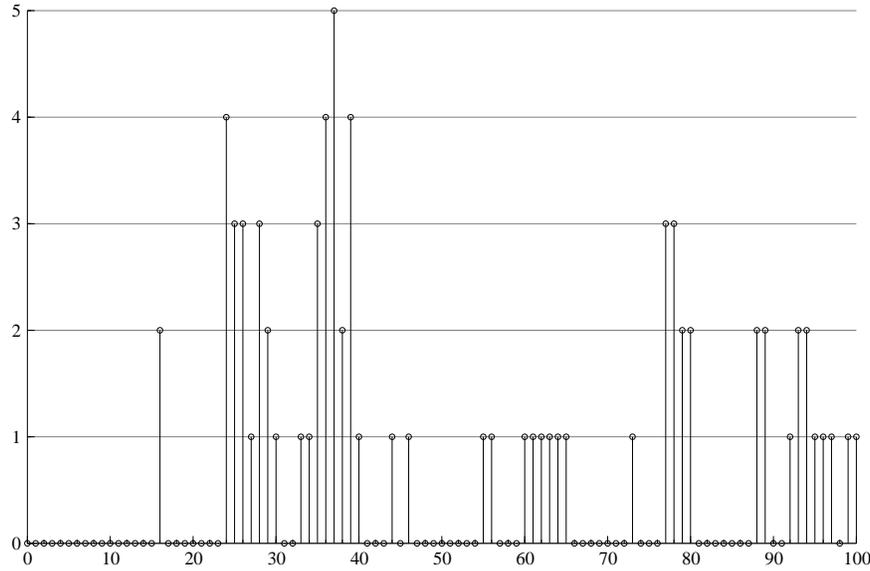


Figura 5.2: Simulación de un proceso Geométrico AR(1) con $\theta = 0.5$, $\alpha = 0.5$ y $n = 100$.

Haciendo $\lambda = \frac{(1-\theta)}{\theta}$ se obtiene

$$G_X(s) = \left(\frac{\lambda}{\lambda + s} \right)^r, \quad (5.20)$$

que es idéntica a la transformada de Laplace de la distribución $\text{Gamma}(r, \lambda)$. Como se mencionó en la subsección anterior, la razón de la similitud entre la T.L. de una v.a. Gamma y la f.a.g.p. de una distribución Binomial Negativa se encuentra en la Proposición 5.9. Por lo tanto cualquier resultado distribucional para una combinación lineal de v.a.'s Gammas puede ser duplicado para v.a.'s Binomiales Negativas si el producto αX es reemplazado por $\alpha \circ X$.

Por (5.5):

$$G_Z(s) = \left(\frac{\lambda + \alpha s}{\lambda + s} \right)^r, \quad (5.21)$$

que es del todo idéntica a (4.9), modelo GAR(1).

Si $r = 1, 2, \dots$, (5.21) se resuelve como una mezcla finita de distribuciones Geométricas (véase (4.9) y (4.10) por ejemplo). Sin embargo, para valores grandes de r , una solución así es impráctica. Más aún, para valores no enteros de r una solución de esa índole no es posible.

Para el modelo GAR(1) se probó una solución general para la innovación. Se probó que la distribución de la innovación puede ser generada como una distribución Poisson compuesta de la forma (4.11). El análogo discreto a (4.11) consistiría en sustituir la multiplicación por escalar por adelgazamiento Binomial y la distribución Exponencial por Geométrica, quedando

$$Z_n = \sum_{r=1}^N (\alpha^{U_r}) \circ Y_r, \quad \text{para cada } n, \quad (5.22)$$

con $\{U_r\}$ una sucesión de v.a.'s i.i.d. con distribución Uniforme(0,1). $\{Y_r\}$ una sucesión de v.a.'s i.i.d. Geom(θ), y N es una v.a. Poisson de media $r \ln(1/\alpha)$. Además, $\{Y_r\}$, $\{U_r\}$ y N son independientes.

¿Representa realmente (5.22) una forma de generar la innovación del proceso INAR(1) con distribución marginal Binomial Negativa?

Proposición 5.3 Z_n como en (5.22) es la innovación para el proceso INAR(1) con distribución marginal BN(r, θ).

Demostración: Siguiendo la estrategia que se utilizó para el modelo GAR(1) lo que haremos es probar que la f.a.g.p. de Z_n como en (5.22) es igual a (5.21).

Sean W y Y dos v.a.'s independientes, entonces

$$E[(1-s)^{W \circ Y}] = E\left[\underbrace{E[(1-s)^{W \circ Y} | W]}_{(*)}\right].$$

La esperanza que tiene una llave y un asterisco debajo no es sino la f.a.g.p. de $W \circ Y$, con W conocido en cuyo caso por (4.23), $\{*\} = G_Y(Ws)$, por lo tanto

$$E[(1-s)^{W \circ Y}] = E[G_Y(Ws)].$$

Si Y se distribuye Geom(θ) y haciendo $\lambda = (1-\theta)/\theta$, por (5.14) la ecuación anterior queda como

$$E[(1-s)^{W \circ Y}] = E\left[\frac{\lambda}{\lambda + sW}\right] = \lambda E\left[\frac{1}{\lambda + sW}\right]. \quad (5.23)$$

Sea $W = \alpha^U$, con $\alpha \in (0, 1)$ y U una v.a. Uniforme(0,1). Por (4.12) tenemos

$$f_W(w) = -\frac{1}{w} \frac{1}{\ln(\alpha)}, \quad w \in (\alpha, 1).$$

Sea $M = \frac{1}{\lambda + sW}$. Haciendo uso del Teorema 2.23 se obtiene

$$f_M(m) = -\frac{1}{\ln(\alpha)} \frac{1}{m} \frac{1}{(1-\lambda m)}, \quad m \in \left(\frac{1}{\lambda+s}, \frac{1}{\lambda+\alpha s}\right),$$

entonces

$$E[M] = -\frac{1}{\ln(\alpha)} \int_{\frac{1}{\lambda+s}}^{\frac{1}{\lambda+\alpha s}} \frac{dm}{1-\lambda m} = \frac{1}{\lambda \ln(\alpha)} \ln \left[\frac{\alpha(\lambda+s)}{\lambda+\alpha s} \right].$$

Sustituyendo en (5.23)

$$E[(1-s)^{W \circ Y}] = \frac{\ln \left[\frac{\alpha(\lambda+s)}{\lambda+\alpha s} \right]}{\ln(\alpha)} = 1 - \frac{\ln \left(\frac{\lambda+\alpha s}{\lambda+s} \right)}{\ln(\alpha)},$$

que no es sino (4.13). El resto de la demostración es igual a la de la Proposición 4.7. \square

Esperanza y varianza condicionales

$$\begin{aligned}
E[X_n | X_{n-1}] &= \alpha X_{n-1} + E[Z_n] \\
&= \alpha X_{n-1} + (1 - \alpha)r \left(\frac{\theta}{1 - \theta} \right) = \alpha X_{n-1} + (1 - \alpha) \frac{r}{\lambda},
\end{aligned} \tag{5.24}$$

pues

$$E[Z_n] = E \left[\sum_{r=1}^N E[\alpha^{U_r} \circ Y_r | N] \right] = E \left[N E[\alpha^U \circ Y] \right]$$

y, $E[\alpha^U \circ Y]$ la obtenemos aplicando esperanza condicional sobre U , y la propiedad 4 de la Proposición 4.24 quedando, $E[\alpha^U \circ Y] = E[\alpha^U]E[Y]$.

Por lo tanto

$$E[Z_n] = E \left[N E[\alpha^U] E[Y] \right] = E[N] E[\alpha^U] E[Y],$$

con $E[\alpha^U] = E[W] = \frac{\alpha-1}{\ln(\alpha)}$; $E[Y] = \frac{\theta}{(1-\theta)} = \frac{1}{\lambda}$, y $E[N] = -r \ln(\alpha)$.

Mientras que

$$\text{Var}(X_n | X_{n-1}) = \alpha(1 - \alpha)X_{n-1} + \text{Var}(Z_n).$$

Correlación serial Puesto que $E[X_n | X_{n-1}]$ tiene la forma (5.24), y $E[X_n] = (r\theta)/(1 - \theta) = r/\lambda$ para toda n , entonces por el Corolario 2.17,

$$\rho(k) = \alpha^k, \quad k = 0, 1, \dots$$

Distribución conjunta Es claro que, como sucede entre los modelo EAR(1) y GAR(1), la f.a.g.p. conjunta del modelo INAR(1) con distribución marginal Binomial Negativa(r, θ) es r veces la de la Geométrica(θ), esto es, r veces (5.18)

$$G_{X_n, X_{n+1}}(u, v) = \left(\frac{\lambda}{\lambda + u + \alpha v - \alpha uv} \frac{\lambda + \alpha v}{\lambda + v} \right)^r, \tag{5.25}$$

con $\lambda = (1 - \theta)/\theta$.

Ya que ésta, no es simétrica se sigue que el proceso INAR(1) con distribución marginal BN(r, θ) es no reversible en el tiempo.

5.1.4. Propiedades generales

Teorema 5.4 El modelo INAR(1), (5.10), posee esperanza y varianza condicionales, de X_n dado X_{n-1} , lineales en X_{n-1} .

Demostración: En (5.10)

$$E[X_n | X_{n-1}] = E[\alpha \circ X_{n-1} | X_{n-1}] + E[Z_n] = \alpha X_{n-1} + \mu_z. \tag{5.26}$$

$$\text{Var}(X_n | X_{n-1}) = \text{Var}(\alpha \circ X_{n-1} | X_{n-1}) + \text{Var}(Z_n) = \alpha(1 - \alpha)X_{n-1} + \sigma_z^2.$$

Corolario 5.5 Si el modelo INAR(1) tiene varianza constante finita, entonces su función de autocorrelación es

$$\text{Corr}(X_n, X_{n+k}) = \rho(k) = \alpha^k, \quad k = 0, 1, 2, \dots$$

Demostración: Aplicando esperanza en (5.26) se tiene $E[X_n] = \alpha E[X_{n-1}] + \mu_z$. Por lo tanto si $\{X_n\}$ posee media constante μ , entonces $\mu(1 - \alpha) = \mu_z$.

Por el Corolario 2.17 se tiene el resultado. \square

Observe que la función de autocorrelación del proceso INAR(1) es siempre positiva.

Teorema 5.6 Si el proceso INAR(1) es estacionario de segundo orden, su esperanza y varianza, μ y σ^2 respectivamente, están dadas en términos de las de la innovación, como

$$\mu = \frac{\mu_z}{(1 - \alpha)}, \quad \sigma^2 = \frac{\alpha\mu_z + \sigma_z^2}{1 - \alpha^2}.$$

Demostración: La esperanza ya la hemos obtenido arriba.

Para la varianza: Por hipótesis el proceso posee varianza constante σ^2 . Aplicando varianza de ambos lados de la ecuación (5.10) y usando la Proposición 4.24, punto 15, se tiene

$$\begin{aligned} \sigma^2 &= \alpha^2 \sigma^2 + \alpha(1 - \alpha)\mu + \sigma_z^2. \\ \sigma^2 &= \frac{\alpha(1 - \alpha)\mu + \sigma_z^2}{1 - \alpha^2} = \frac{\alpha\mu_z + \sigma_z^2}{1 - \alpha^2}. \quad \square \end{aligned}$$

Teorema 5.7 Sean Z y W v.a.'s. Si las series de conteo de $\alpha \circ Z$ y de $\alpha \circ W$ son independientes, e independientes de Z y W , entonces

$$\alpha \circ (Z + W) \stackrel{d}{=} \alpha \circ Z + \alpha \circ W.$$

Demostración: Sabemos que

$$G_{\alpha \circ X}(s) = G_X(\alpha s).$$

De donde

$$G_{\alpha \circ (Z+W)}(s) = G_{Z+W}(\alpha s).$$

Por otra parte

$$\begin{aligned} G_{\alpha \circ Z + \alpha \circ W}(s) &= E\left[(1 - s)^{\alpha \circ Z} (1 - s)^{\alpha \circ W}\right] = E\left[(1 - s)^{\alpha \circ W} E\left[(1 - s)^{\alpha \circ Z} | Z\right]\right] \\ &= E\left[(1 - s)^{\alpha \circ W} (1 - \alpha s)^Z\right] = E\left[(1 - \alpha s)^Z E\left[(1 - s)^{\alpha \circ W} | W\right]\right] \\ &= E\left[(1 - \alpha s)^Z (1 - \alpha s)^W\right] = E\left[(1 - \alpha s)^{Z+W}\right], \end{aligned}$$

i.e.

$$G_{\alpha \circ Z + \alpha \circ W}(s) = G_{Z+W}(\alpha s).$$

Acabamos de probar que $G_{\alpha \circ Z + \alpha \circ W}(s) = G_{\alpha \circ (Z+W)}$.

Por lo tanto

$$\alpha \circ (Z + W) \stackrel{d}{=} \alpha \circ Z + \alpha \circ W. \quad \square$$

Tenemos entonces que el operador de adelgazamiento Binomial, por decirlo así, abre sumas.

Además, es evidente que el resultado es válido para un número arbitrario de sumandos.

Teorema 5.8 La distribución marginal del proceso INAR(1), (5.10), puede ser expresada en términos de la sucesión de innovación $\{Z_n\}$, como

$$X_n \stackrel{d}{=} \sum_{i=0}^{\infty} \alpha^i \circ Z_{n-i}.$$

Demostración: Por la Proposición 4.24, $\alpha \circ (\beta \circ X) \stackrel{d}{=} (\alpha\beta) \circ X$.

Entonces

$$\begin{aligned} X_n &= \alpha \circ X_{n-1} + Z_n \\ &= \alpha \circ (\alpha \circ X_{n-2} + Z_{n-1}) + Z_n \\ &= \alpha \circ \left(\alpha \circ \{ \alpha X_{n-3} + Z_{n-2} \} + Z_{n-1} \right) + Z_n \\ &\vdots \\ &\stackrel{d}{=} \sum_{i=0}^{\infty} \alpha^i \circ Z_{n-i}, \end{aligned}$$

puesto que por definición, las Z_n 's son independientes de todas las series de conteo, las cuales a su vez son mutuamente independientes. \square

Este resultado es el análogo discreto para el modelo continuo AR(1), ecuación (2.5).

Por último, presentamos un resultado pendiente.

Proposición 5.9 Si un proceso Poisson de intensidad unitaria es observado por un período de tiempo X que se distribuye Gamma($r, \frac{1-\theta}{\theta}$), entonces el número de eventos (N) que ocurren en ese lapso de tiempo se distribuye Binomial Negativo(r, θ).

Demostración:

$$\mathbb{P}[N(X) = n | X = x] = e^{-x} \frac{x^n}{n!},$$

puesto que la intensidad del proceso es $\lambda = 1$.

Para simplificar la notación hacemos $\beta = \frac{1-\theta}{\theta}$. Por lo que $X \sim \text{Gamma}(r, \beta)$. Entonces

$$\begin{aligned} \mathbb{P}[N = n] &= \int_0^{\infty} e^{-x} \frac{x^n}{n!} \frac{\beta^r}{\Gamma(r)} x^{r-1} e^{-\beta x} dx \\ &= \frac{\beta^r}{\Gamma(r)} \frac{1}{n!} \int_0^{\infty} x^{n+r-1} e^{-(\beta+1)x} dx \end{aligned}$$

$$\begin{aligned}
&= \frac{\beta^r}{\Gamma(r)} \frac{1}{n!} \frac{\Gamma(n+r)}{(\beta+1)^{n+r}} \\
&= \binom{r+n-1}{n} \left(\frac{1}{\beta+1}\right)^n \left(\frac{\beta}{\beta+1}\right)^r.
\end{aligned}$$

Por lo tanto $N \sim \text{Binomial Negativa}(r, \frac{1}{\beta+1})$, y en términos de θ , $N \sim \text{Binomial Negativa}(r, \theta)$. \square

5.2. Otros AR(1) discretos

Los dos modelos de esta sección pretenden ilustrar formas alternativas al adelgazamiento Binomial que permiten la creación de modelos con distribución marginal discreta, con valores en \mathbb{N}_0 , y estructura AR(1).

Primero presentamos un modelo con distribución marginal Binomial Negativa, que es la versión discreta del modelo de Sim (1990) para un proceso AR(1) con distribución marginal Gamma. Estos dos modelos permiten ilustrar una generalización que se ha hecho del adelgazamiento Binomial. Este operador lo denotamos por “ \bullet ”, y está definido como

$$\phi \bullet X = \sum_{i=1}^{N(X)} W_i, \quad (5.27)$$

donde $N(X)$ es una v.a. discreta no negativa y $\{W_i\}$ es una sucesión de v.a.'s i.i.d., independiente de $N(X)$, tales que $E[N(X)W_i|X] = \phi X$.

El adelgazamiento Binomial es un caso particular, en el que $N(X) = X$, y $\{W_i\}$ son v. a.'s Bernoulli de media ϕ .

El segundo modelo considera la distribución Binomial como distribución marginal. Al tener la distribución Binomial un soporte finito el modelo (5.4) no puede funcionar, e introducimos lo que ha sido llamado adelgazamiento Hipergeométrico, $S(\cdot)$.

5.2.1. Binomial Negativa

Necesitamos hacer algunos cambios en la notación.

En esta subsección 5.2.1 diremos que X tiene distribución Binomial Negativa(r, θ) si su función de densidad está dada por

$$p_X(k) = \binom{r+k-1}{k} \theta^r (1-\theta)^k, \quad k = 0, 1, \dots \quad (5.28)$$

Obsérvese que (5.28) en términos de (5.19), es una distribución Binomial Negativa($r, 1-\theta$).

En éste caso

$$E[X] = r \frac{1-\theta}{\theta}, \quad \text{Var}(X) = r \frac{(1-\theta)}{\theta^2}. \quad (5.29)$$

Además, cambiamos la notación para la f.a.g.p. de X , sólo, y recalcamos **sólo** para esta subsección 5.2.1, por $G_X(\bar{s})$, con

$$G_X(\bar{s}) = G_X(1-s) := E[(1-s)^X].$$

En este caso

$$\begin{aligned} G_X(\bar{s}) &= E[(1-s)^X] = \sum_{x=0}^{\infty} \binom{r+x-1}{x} \theta^r (1-\theta)^x (1-s)^x \\ &= \theta^r \left(\frac{1}{1-(1-s-\theta+\theta s)} \right)^r = \left(\frac{\theta}{s+\theta-\theta s} \right)^r. \end{aligned}$$

Si por ejemplo $\theta = \alpha/(1+\alpha)$, con $\alpha > 0$, entonces

$$E[X] = \frac{r}{\alpha}, \quad \text{Var}(X) = \frac{r(1+\alpha)}{\alpha^2}, \quad G_X(\bar{s}) = \left(\frac{\alpha}{\alpha+s} \right)^r. \quad (5.30)$$

Considere ahora el modelo

$$X_n = p \bullet X_{n-1} + Z_n, \quad p \in (0, 1), \quad (5.31)$$

con

$$p \bullet X = \sum_{i=0}^{N(X)} W_i,$$

donde

- $\{W_i\}$ es una sucesión de v.a.'s i.i.d. Geométricas de parámetro $\frac{\alpha}{1+\alpha}$, $W_0 = 0$.
- Dado X , $N(X)$ es una v.a. Binomial(X, λ), con $\lambda = \alpha p$. Lo que implica $0 < \alpha < 1/p$.
- $\{Z_n\}$ es una sucesión de v.a.'s i.i.d. con distribución Binomial Negativa($r, \frac{\alpha}{1+\alpha}$), $r > 0$, y Z_n es independiente de $p \bullet X_{n-1}$.

La diferencia primordial entre $p \circ X$, y $p \bullet X$, en este caso, se encuentra en que dada X , para el primero se tiene una suma cierta de v.a.'s, mientras que para el segundo tenemos una suma aleatoria de v.a.'s.

A continuación enunciamos el resultado principal de esta subsección.

Proposición 5.10 Si X_0 tiene distribución Binomial Negativa($r, \frac{\alpha \bar{p}}{1+\alpha \bar{p}}$) entonces el modelo (5.31) define un proceso estacionario con distribución marginal Binomial Negativa($r, \frac{\alpha \bar{p}}{1+\alpha \bar{p}}$).

Demostración: Observe que $G_{X_n}(\bar{s}) = G_{X_n}(1-s) = E[(1-s)^{X_n}]$. Si X es fija digamos x , entonces $G_{X_n=x}(1-s) = (1-s)^x$.

Ahora, la f.a.g.p. de $p \bullet X$, dado $X = x$ es

$$G_{p \bullet X|X=x}(\bar{s}) = E \left[(1-s)^{\sum_{i=0}^{N(X)} W_i} \middle| X = x \right] = \sum_{n=0}^x \left(\frac{\alpha}{\alpha+s} \right)^n \binom{x}{n} \lambda^n (1-\lambda)^{x-n}$$

$$= \left(\frac{\lambda\alpha}{\alpha+s} + 1 - \lambda \right)^x = \left(1 - \frac{\lambda s}{\alpha+s} \right)^x. \quad (5.32)$$

Por independencia la f.a.g.p. de X_n dado X_{n-1} es

$$G_{X_n|X_{n-1}=x}(\bar{s}) = \left(1 - \frac{\lambda s}{\alpha+s} \right)^x \left(\frac{\alpha}{\alpha+s} \right)^r.$$

Entonces, de la observación primera se tiene

$$G_{X_n}(\bar{s}) = G_{X_{n-1}} \left(1 - \frac{\lambda s}{\alpha+s} \right) \left(\frac{\alpha}{\alpha+s} \right)^r.$$

Esta ecuación se resuelve recursivamente.

$$\begin{aligned} G_{X_{n-1}} \left(1 - \frac{\lambda s}{\alpha+s} \right) &= G_{X_{n-2}} \left(1 - \frac{\lambda \left(\frac{\lambda s}{\alpha+s} \right)}{\alpha + \left(\frac{\lambda s}{\alpha+s} \right)} \right) \left(\frac{\alpha}{\alpha + \left(\frac{\lambda s}{\alpha+s} \right)} \right)^r \\ &= G_{X_{n-2}} \left(1 - \frac{\lambda^2 s}{\alpha^2 + \alpha s + \lambda s} \right) \left(\frac{\alpha}{\alpha + \left(\frac{\lambda s}{\alpha+s} \right)} \right)^r. \end{aligned}$$

En tanto que

$$\begin{aligned} G_{X_{n-2}} \left(1 - \frac{\lambda^2 s}{\alpha^2 + \alpha s + \lambda s} \right) &= G_{X_{n-3}} \left(1 - \frac{\frac{\lambda^3 s}{\alpha^2 + \alpha s + \lambda s}}{\frac{\alpha(\alpha^2 + \alpha s + \lambda s) + \lambda^2 s}{\alpha^2 + \alpha s + \lambda s}} \right) \left(\frac{\alpha}{\alpha + \frac{\lambda^2 s}{\alpha^2 + \alpha s + \lambda s}} \right)^r \\ &= G_{X_{n-3}} \left(1 - \frac{\lambda^3 s}{\alpha^3 + \alpha^2 s + \alpha \lambda s + \lambda^2 s} \right) \left(\frac{\alpha}{\alpha + \frac{\lambda^2 s}{\alpha^2 + \alpha s + \lambda s}} \right)^r \end{aligned}$$

y,

$$\begin{aligned} G_{X_{n-3}} \left(1 - \frac{\lambda^3 s}{\alpha^3 + \alpha^2 s + \alpha \lambda s + \lambda^2 s} \right) \\ &= G_{X_{n-4}} \left(1 - \frac{\lambda^4 s}{\alpha^4 + \alpha^3 s + \alpha^2 \lambda s + \alpha \lambda^2 s + \lambda^3 s} \right) \left(\frac{\alpha}{\alpha + \frac{\lambda^3 s}{\alpha^3 + \alpha^2 s + \alpha \lambda s + \lambda^2 s}} \right)^r. \end{aligned} \quad (5.33)$$

En general el lector puede comprobar que G_{X_n} está en términos de $G_{X_{n-m}}$, $m \geq 1$, con $G_{X_{n-m}}$ evaluada en

$$G_{X_{n-m}} \left(1 - \frac{\lambda^m s}{\alpha^m + \alpha^{m-1} s + \alpha^{m-2} \lambda s + \alpha^{m-3} \lambda^2 s + \dots + \alpha \lambda^{m-2} s + \lambda^{m-1} s} \right). \quad (5.34)$$

Analícemos el denominador

$$\begin{aligned} \alpha^m + \alpha^{m-1} s + \alpha^{m-2} \lambda s + \alpha^{m-3} \lambda^2 s + \dots + \alpha \lambda^{m-2} s + \lambda^{m-1} s \\ = \alpha^m + s[\alpha^{m-1} + \alpha^{m-2} \lambda + \alpha^{m-3} \lambda^2 + \dots + \alpha \lambda^{m-2} + \lambda^{m-1}] \end{aligned}$$

$$\begin{aligned}
&= \alpha^m + s\alpha^{m-1}[1 + p + p^2 + \dots + p^{m-2} + p^{m-1}] \quad (\text{pues } \lambda = \alpha p) \\
&= \alpha^m + s\alpha^{m-1} \left[\frac{1 - p^m}{1 - p} \right], \quad m \geq 1.
\end{aligned} \tag{5.35}$$

Por lo tanto, haciendo $m = n$ en (5.34), es decir, expresando G_{X_n} en términos de G_{X_0} , nos queda que esta última está evaluada en

$$G_{X_0} \left(1 - \frac{\alpha^n p^n s}{\alpha^n + s\alpha^{n-1} \left[\frac{1-p^n}{1-p} \right]} \right) = G_{X_0} \left(1 - \frac{p^n s}{1 + \frac{s}{\alpha} \left[\frac{1-p^n}{1-p} \right]} \right). \tag{5.36}$$

Así,

$$\begin{aligned}
&G_{X_n}(\bar{s}) \\
&= G_{X_{n-1}}(\cdot) \left(\frac{\alpha}{\alpha + s} \right)^r \\
&= G_{X_{n-2}}(\cdot) \left(\frac{\alpha}{\alpha + \left(\frac{\lambda s}{\alpha + s} \right)} \right)^r \left(\frac{\alpha}{\alpha + s} \right)^r \\
&= G_{X_{n-3}}(\cdot) \left(\frac{\alpha}{\alpha + \frac{\lambda^2 s}{\alpha^2 + \alpha s + \lambda s}} \right)^r \left(\frac{\alpha}{\alpha + \left(\frac{\lambda s}{\alpha + s} \right)} \right)^r \left(\frac{\alpha}{\alpha + s} \right)^r \\
&\quad \vdots \\
&= G_{X_{n-m}}(\cdot) \left(\frac{\alpha}{\alpha + \frac{\lambda^{m-1} s}{\alpha^{m-1} + \alpha^{m-2} s + \alpha^{m-3} \lambda s + \dots + \alpha \lambda^{m-3} s + \lambda^{m-2} s}} \right)^r \dots \left(\frac{\alpha}{\alpha + s} \right)^r \\
&= G_{X_{n-m}}(\cdot) \left(\frac{\alpha}{\alpha + \frac{\alpha^{m-1} p^{m-1} s}{\alpha^{m-1} + s\alpha^{m-2} \left[\frac{1-p^{m-1}}{1-p} \right]}} \right)^r \dots \left(\frac{\alpha}{\alpha + s} \right)^r,
\end{aligned} \tag{5.37}$$

haciendo uso de (5.35) para obtener esta última. Es de recalcar que (5.37) aplica para $m \geq 2$.

Ahora

$$\begin{aligned}
\left(\frac{\alpha}{\alpha + \frac{\alpha^{m-1} p^{m-1} s}{\alpha^{m-1} + s\alpha^{m-2} \left[\frac{1-p^{m-1}}{1-p} \right]}} \right)^r &= \left(\frac{\alpha}{\alpha + \frac{p^{m-1} s}{1 + \frac{s}{\alpha} \left[\frac{1-p^{m-1}}{1-p} \right]}} \right)^r \\
&= \left(\frac{\alpha}{\alpha + \frac{\alpha(1-p)p^{m-1}s}{\alpha(1-p) + s(1-p^{m-1})}} \right)^r
\end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\alpha(1-p) + s(1-p^{m-1})}{\alpha(1-p) + s(1-p^{m-1}) + (1-p)p^{m-1}s} \right)^r \\
&= \left(\frac{\alpha(1-p) + s(1-p^{m-1})}{\alpha(1-p) + s(1-p^m)} \right)^r, \quad \text{para } m \geq 2.
\end{aligned}$$

Por lo tanto, por (5.37) y (5.36)

$$\begin{aligned}
G_{X_n}(\bar{s}) &= G_{X_0} \left(1 - \frac{p^n s}{1 + \frac{s}{\alpha} \left[\frac{1-p^n}{1-p} \right]} \right) \prod_{m=2}^n \left(\frac{\alpha(1-p) + s(1-p^{m-1})}{\alpha(1-p) + s(1-p^m)} \right)^r \left(\frac{\alpha}{\alpha+s} \right)^r \\
&= G_{X_0} \left(1 - \frac{p^n s}{1 + \frac{s}{\alpha} \left[\frac{1-p^n}{1-p} \right]} \right) \left(\frac{\alpha(1-p) + s(1-p)}{\alpha(1-p) + s(1-p^2)} \frac{\alpha(1-p) + s(1-p^2)}{\alpha(1-p) + s(1-p^3)} \cdots \right. \\
&\quad \left. \cdots \frac{\alpha(1-p) + s(1-p^{n-1})}{\alpha(1-p) + s(1-p^n)} \frac{\alpha}{\alpha+s} \right)^r.
\end{aligned}$$

Cancelándose todos los términos de en medio y quedando sólo

$$\begin{aligned}
G_{X_n}(\bar{s}) &= G_{X_0} \left(1 - \frac{p^n s}{1 + \frac{s}{\alpha} \left[\frac{1-p^n}{1-p} \right]} \right) \left(\frac{\alpha(1-p) + s(1-p)}{\alpha(1-p) + s(1-p^n)} \frac{\alpha}{\alpha+s} \right)^r \\
&= G_{X_0} \left(1 - \frac{p^n s}{1 + \frac{s}{\alpha} \left[\frac{1-p^n}{1-p} \right]} \right) \left(\frac{\alpha(1-p)}{\alpha(1-p) + s(1-p^n)} \right)^r.
\end{aligned}$$

Haciendo $n \rightarrow \infty$ en esta última, se obtiene que

$$G_{X_n}(\bar{s}) = \left(\frac{\alpha(1-p)}{\alpha(1-p) + s} \right)^r. \quad (5.38)$$

De las ecuaciones (5.38) y (5.30) se tiene que la distribución límite de X_n es la distribución Binomial Negativa($r, \frac{\alpha\bar{p}}{1+\alpha\bar{p}}$).

Por lo tanto, si X_0 tiene distribución Binomial Negativa($r, \frac{\alpha\bar{p}}{1+\alpha\bar{p}}$) entonces, (5.31) define un proceso estacionario con esta distribución marginal. \square

Esperanza y varianza condicionales

$$\begin{aligned}
E[X_n|X_{n-1}] &= E[\alpha \bullet X_{n-1} | X_{n-1}] + E[Z_n] \\
&= E \left[\sum_{i=0}^{N(X_{n-1})} W_i | X_{n-1} \right] + E[Z_n] = E \left[\sum_{i=0}^N W_i \right] + E[Z_n],
\end{aligned}$$

donde $N \sim \text{Binomial}(X_{n-1}, \lambda)$ independiente de toda W_i . Aplicando esperanza condicional se tiene, $E[\sum_{i=0}^N W_i] = E[N]E[W_1]$. Haciendo uso de (5.30) tenemos que

$$E[X_n|X_{n-1}] = \lambda X_{n-1} \frac{1}{\alpha} + \frac{r}{\alpha} = pX_{n-1} + \frac{r}{\alpha}. \quad (5.39)$$

Mientras que

$$\text{Var}(X_n|X_{n-1}) = \text{Var} \left(\sum_{i=0}^N W_i \right) + \text{Var}(Z_n), \quad (5.40)$$

con N como arriba.

Nos ocupamos del primer sumando

$$\begin{aligned}\text{Var}\left(\sum_{i=0}^N W_i\right) &= E\left[\text{Var}\left(\sum_{i=0}^N W_i \mid N\right)\right] + \text{Var}\left(E\left[\sum_{i=0}^N W_i \mid N\right]\right) \\ &= E[N\text{Var}(W_1)] + \text{Var}(NE[W_1]) = \text{Var}(W_1)E[N] + E^2[W_1]\text{Var}(N),\end{aligned}$$

Haciendo uso de (5.30)

$$\text{Var}\left(\sum_{i=0}^N W_i\right) = \frac{(1+\alpha)}{\alpha^2}\lambda X_{n-1} + \frac{1}{\alpha^2}\lambda(1-\lambda)X_{n-1}.$$

Retomando finalmente (5.40),

$$\text{Var}(X_n|X_{n-1}) = (2 + \alpha\bar{p})\frac{p}{\alpha}X_{n-1} + \frac{r(1+\alpha)}{\alpha^2}.$$

Correlación serial Puesto que X_n tiene distribución marginal Binomial Negativa($r, \frac{\alpha\bar{p}}{1+\alpha\bar{p}}$), de (5.29)

$$\mu = E[X_n] = r \frac{1 - \frac{\alpha\bar{p}}{1+\alpha\bar{p}}}{\frac{\alpha\bar{p}}{1+\alpha\bar{p}}} = \frac{r}{\alpha\bar{p}}.$$

Entonces, $E[X_n|X_{n-1}]$ se puede escribir como $E[X_n|X_{n-1}] = pX_{n-1} + (1-p)\mu$. Por lo tanto del Corolario 2.17 se tiene

$$\rho(k) = p^k, \quad k = 0, 1, 2, \dots$$

Distribución conjunta Es claro que en este modelo tanto (5.6) como (5.7) ya no se pueden usar.

En lugar de (5.7), para este modelo se tiene:

Considere X_n y X_{n+1} ,

$$G_{X_n, X_{n+1}}(\bar{u}, \bar{v}) = E[(1-u)^{X_n}(1-v)^{X_{n+1}}] = E\left[(1-u)^{X_n}(1-v)^{\alpha \bullet X_n}(1-v)^{Z_{n+1}}\right]. \quad (5.41)$$

Ahora

$$\begin{aligned}E\left[(1-u)^{X_n}(1-v)^{\alpha \bullet X_n}\right] &= E\left[(1-u)^{X_n}E\left[(1-v)^{\alpha \bullet X_n} \mid X_n\right]\right] \\ &= E\left[(1-u)^{X_n}\left(1 - \frac{\lambda v}{\alpha + v}\right)^{X_n}\right] \quad \text{por (5.32)} \\ &= E\left[\left(1 - u - \frac{\lambda v}{\alpha + v} + \frac{\lambda uv}{\alpha + v}\right)^{X_n}\right] = G_X\left(u + \frac{\lambda v}{\alpha + v} - \frac{\lambda uv}{\alpha + v}\right).\end{aligned}$$

Aplicando independencia en (5.41),

$$G_{X_n, X_{n+1}}(\bar{u}, \bar{v}) = \left(\frac{\alpha\bar{p}}{\alpha\bar{p} + u + \frac{\lambda v}{\alpha + v} - \frac{\lambda uv}{\alpha + v}}\right)^r \left(\frac{\alpha}{\alpha + v}\right)^r$$

$$= \left[\frac{\alpha^2 \bar{p}}{\alpha^2 \bar{p} + \alpha u + \alpha v \bar{p} + uv + \lambda v - \lambda uv} \right]^r = \left[\frac{\alpha^2 \bar{p}}{\alpha^2 \bar{p} + \alpha(u+v) + uv(1-\alpha p)} \right]^r.$$

Puesto que $G_{X_n, X_{n+1}}(\bar{u}, \bar{v})$ es simétrica en u y v , se tiene que el proceso (5.31) es reversible en el tiempo. Podemos usar este resultado puesto que el proceso (5.31) es una cadena de Markov con probabilidades condicionales de la siguiente forma:

Probabilidad condicional En este caso la idea es la misma que en (5.6):

$$p(x, y) = \mathbb{P}[X_n = y | X_{n-1} = x] = \sum_{k=0}^y \mathbb{P}[\alpha \bullet X_{n-1} = k | X_{n-1} = x] \mathbb{P}[Z_n = y - k]. \quad (5.42)$$

Ahora, hay que tener mucho cuidado con la v.a. $(p \bullet X_{n-1} | X_{n-1})$.

Abusando de la notación

$$\mathbb{P}[p \bullet X_{n-1} = k | X_{n-1} = x] = \sum_{m=0}^x \mathbb{P}\left[\sum_{i=0}^N W_i = k | N = m\right] \mathbb{P}[N = m].$$

Donde $N \sim \text{Binomial}(x, \lambda)$, y $\sum_{i=0}^m W_i \sim \text{Binomial Negativa}(m, \frac{\alpha}{1+\alpha})$.

Entonces

$$\mathbb{P}[p \bullet X_{n-1} = k | X_{n-1} = x] = \sum_{m=0}^x \binom{x}{m} \lambda^m (1-\lambda)^{x-m} \binom{m+k-1}{k} \left(1 - \frac{\alpha}{1+\alpha}\right)^k \left(\frac{\alpha}{1+\alpha}\right)^m.$$

Hay una observación adicional que hacer para el caso en que $m = 0$. $\mathbb{P}[N = 0] = (1-\lambda)^x$.

En tal caso a Z_n en (5.42), no le queda otra más que tomar el valor y . Esto es

$$\begin{aligned} \mathbb{P}[X_n = y | X_{n-1} = x] &= (1-\lambda)^x \binom{r+y-1}{y} \left(\frac{1}{1+\alpha}\right)^y \left(\frac{\alpha}{1+\alpha}\right)^r \\ &= \frac{\alpha^r (1-\alpha p)^x}{(1+\alpha)^{r+y}} \binom{r+y-1}{y}, \end{aligned} \quad (5.43)$$

puesto que $\lambda = \alpha p$, y $\left(1 - \frac{\alpha}{1+\alpha}\right) = \frac{1}{1+\alpha}$.

Para el resto de los casos, $m \neq 0$, se tiene

$$\begin{aligned} \mathbb{P}[X_n = y | X_{n-1} = x] &= \sum_{k=0}^y \sum_{m=1}^x \binom{x}{m} \lambda^m (1-\lambda)^{x-m} \\ &\quad \times \binom{m+k-1}{k} \left(\frac{1}{1+\alpha}\right)^k \left(\frac{\alpha}{1+\alpha}\right)^m \binom{r+y-k-1}{y-k} \left(\frac{1}{1+\alpha}\right)^{y-k} \left(\frac{\alpha}{1+\alpha}\right)^r. \end{aligned}$$

Simplificando esta última y uniendo (5.43), tenemos que de manera general

$$\begin{aligned} p(x, y) &= \frac{\alpha^r (1-\alpha p)^x}{(1+\alpha)^{r+y}} \left\{ \binom{r+y-1}{y} \right. \\ &\quad \left. + \sum_{m=1}^x \left(\frac{\alpha}{1+\alpha} \frac{\alpha p}{1-\alpha p}\right)^m \binom{x}{m} \sum_{k=0}^y \binom{m+k-1}{k} \binom{r+y-k-1}{y-k} \right\}. \end{aligned} \quad (5.44)$$

5.2.2. Binomial

Como ya se mencionó al inicio del capítulo, por la Proposición 4.20 la distribución Binomial no es I-D discreta y por lo tanto no es descomponible-por sí misma discreta y entonces no puede ser una distribución marginal para la solución estacionaria del modelo (5.4), INAR(1). Sin embargo como hemos visto en la subsección anterior, pueden existir formas alternativas al modelo (5.4) que permitan obtener modelos con estructura AR(1) y la distribución marginal discreta deseada.

En particular el adelgazamiento Binomial asume que cada elemento del proceso al tiempo $n - 1$, X_{n-1} , es retenido al tiempo n con probabilidad (constante) α , o eliminado con probabilidad $\bar{\alpha} = (1 - \alpha)$. Esta es la suposición que modificaremos. Asumiremos que la probabilidad de retener un elemento de X_{n-1} no es constante sino que varía dependiendo del número de elementos que ya han sido retenidos. Concretamente, dado $X_{n-1} = x$ asumimos que el número de elementos retenidos (Sobrevivientes), $S(X_{n-1})$, tiene distribución Hipergeométrica(N, M, x) -y no Binomial(x, α)-. Esto es

$$\mathbb{P}[S(X_{n-1}) = k | X_{n-1} = x] = \frac{\binom{x}{k} \binom{N-x}{M-k}}{\binom{N}{M}}, \quad (5.45)$$

$$N > M > 0, \quad \text{máx}\{0, M - (N - x)\} \leq k \leq \text{mín}\{x, M\}.$$

Proposición 5.11 Si X_{n-1} se distribuye Binomial(N, p) y, $S(X_{n-1}) | X_{n-1}$ tiene distribución como en (5.45), entonces $S(X_{n-1})$ se distribuye Binomial(M, p).

Más aún, $S(X_{n-1})$ y, $X_{n-1} - S(X_{n-1})$, son independientes.

Demostración: Obsérvese que

$$\mathbb{P}[S(X_{n-1}) = k | X_{n-1} = x] = \frac{\binom{x}{k} \binom{N-x}{M-k}}{\binom{N}{M}} = \frac{\binom{M}{k} \binom{N-M}{x-k}}{\binom{N}{x}}.$$

Entonces

$$\begin{aligned} \mathbb{P}[S(X_{n-1}) = k] &= \sum_{x=0}^N \binom{M}{k} \binom{N-M}{x-k} p^x (1-p)^{N-x} \\ &= \binom{M}{k} p^k (1-p)^{M-k} \sum_{x=k}^N \binom{N-M}{x-k} p^{x-k} (1-p)^{N-M-x+k} \\ &= \binom{M}{k} p^k (1-p)^{M-k} \sum_{i=0}^{N-k} \binom{N-M}{i} p^i (1-p)^{N-M-i}. \end{aligned}$$

Puesto que $N - M \leq N - k$ para cualquier k , se tiene

$$\mathbb{P}[S(X_{n-1}) = k] = \binom{M}{k} p^k (1-p)^{M-k}, \quad k = 0, 1, \dots, M.$$

La independencia se prueba en el siguiente capítulo. \square

Definimos entonces el proceso AR(1) con distribución marginal Binomial(N, p), BAR(1), como

$$X_n = S(X_{n-1}) + Z_n, \quad n = 1, 2, \dots, \quad (5.46)$$

con $\{Z_n\}$ una sucesión de v.a.'s i.i.d. Binomiales($N - M, p$), X_0 una v.a.i. Binomial(N, p), y la distribución de $S(X_{n-1})$, representando los sobrevivientes de X_{n-1} , como en (5.45). Además, dado $X_{n-1} = x$, se asume que $S(X_{n-1})$ es independiente de $\{X_{n-j}; j \geq 2\}$ y sus sobrevivientes, $\{S(X_{n-j}); j \geq 2\}$.

Esperanza y varianza condicionales

$$E[X_n | X_{n-1}] = \frac{M}{N} X_{n-1} + (N - M)p = \alpha X_{n-1} + (N - M)p, \quad (5.47)$$

con $\alpha = M/N$.

$$\begin{aligned} \text{Var}(X_n | X_{n-1} = x) &= \frac{N - M}{N - 1} M \frac{x}{N} \left(1 - \frac{x}{N}\right) + (N - M)p(1 - p) \\ &= \left(\frac{NM - M^2}{N^2}\right) \left(\frac{Nx - x^2}{N - 1}\right) + (N - M)p\bar{p} \\ &= \alpha(1 - \alpha) \frac{N - x}{N - 1} x + (N - M)p\bar{p}. \end{aligned} \quad (5.48)$$

Obsérvese que a diferencia de todos los modelos vistos hasta el momento, la varianza condicional de X_n dado X_{n-1} para este modelo es cuadrática en X_{n-1} .

Correlación serial Sea $\mu = E[X_n] = Np$, entonces

$$\begin{aligned} E[X_n | X_{n-1}] &= \frac{M}{N} X_{n-1} + (N - M)p \\ &= \frac{M}{N} X_{n-1} + \left(1 - \frac{M}{N}\right) Np = \alpha X_{n-1} + (1 - \alpha)\mu. \end{aligned}$$

Por el Corolario 2.17 se tiene

$$\rho(k) = \alpha^k, \quad k = 0, 1, 2, \dots$$

Observe que es conveniente que N no sea demasiado grande. De lo contrario se pierde la dependencia. Además, para cualquier rezago la correlación es siempre no negativa. De hecho es estrictamente positiva.

Probabilidad condicional Usando (5.6) se tiene

$$p(x, y) = \mathbb{P}[X_n = y | X_{n-1} = x]$$

$$= \sum_{k=0}^m \binom{M}{k} \binom{N - M}{x - k} \binom{N - M}{y - k} p^{y-k} (1 - p)^{N - M - (y - k)} \bigg/ \binom{N}{x}.$$

Distribución conjunta Es claro que no podemos usar (5.7).

Solo para recordar, si $W \sim \text{Binomial}(x, p)$,

$$G_W(u) = \sum_{i=0}^x (1-u)^i \binom{x}{i} p^i (1-p)^{x-i} = (1-pu)^x.$$

Además, sea $G_W(u, v) := E[(1-u)^W (1-v)^W]$, entonces

$$G_W(u, v) = \sum_{i=0}^x (1-u-v+uv)^i \binom{x}{i} p^i (1-p)^{x-i} = (1-pu-pv+puv)^x.$$

Entonces para el modelo (5.46) tenemos que

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= E\left[(1-u)^{X_n} (1-v)^{S(X_n)+Z_{n+1}}\right] \\ &= G_Z(v) E\left[(1-u)^{X_n-S(X_n)} (1-v)^{S(X_n)} (1-u)^{S(X_n)}\right] \\ &= G_Z(v) G_{S(X_n)}(u, v) G_{X_n-S(X_n)}(u) \\ &= (1-pv)^{N-M} (1-pu-pv+puv)^M (1-pu)^{N-M}. \end{aligned} \quad (5.49)$$

Puesto que $G_{X_n, X_{n+1}}(u, v)$ es simétrica en u y v , se sigue que el proceso BAR(1) es reversible en el tiempo.

La Figura 5.3 presenta la simulación de dos procesos BAR(1).

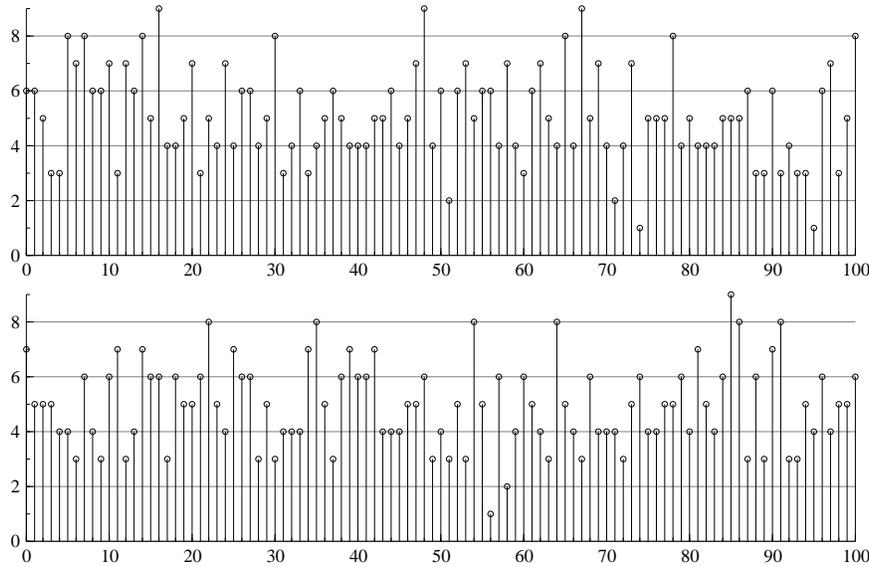


Figura 5.3: Simulación de un modelo Binomial AR(1), con $N = 10$, $p=0.5$, y $n = 100$. En la parte superior: $\alpha = 0.3$. En la parte inferior: $\alpha = 0.7$.

5.3. MA(1) discretos

5.3.1. Poisson

Si bien (5.4) es la versión discreta -en los enteros no negativos- de (2.4), ¿podemos hablar de la versión discreta de (2.3)?

¿Qué pasa si reemplazamos la multiplicación por escalar por adelgazamiento Binomial?

Consideremos el modelo

$$X_n = Z_n + \beta \circ Z_{n-1}, \quad (5.50)$$

con $\beta \in [0, 1]$ y, $\{Z_n\}$ una sucesión de v.a.'s i.i.d. e independiente de todas las series de conteo involucradas en (5.50), las cuales a su vez son mutuamente independientes.

Algunos autores dan el nombre de INMA(1) (INteger-valued Moving Average, de orden 1), al proceso (5.50).

En particular nos planteamos el problema en el que (5.50) sea estacionario con distribución marginal Poisson(θ). Puesto que la suma de v.a.'s Poisson es Poisson, supongamos que las Z_n se distribuyen Poisson(μ_z). Aplicando la f.a.g.p. de ambos lados de la ecuación, nos queda

$$e^{-\theta s} = e^{-\mu_z s} e^{-\mu_z \beta s}, \quad (5.51)$$

de donde se concluye que $\mu_z = \theta/(1 + \beta)$.

Esto es, si $\{Z_n\}$ es una sucesión de v.a.'s i.i.d. Poisson($\theta/(1 + \beta)$), entonces X_n como en (5.50), define un proceso estacionario MA(1) discreto Poisson(θ).

Correlación serial

$$\text{Cov}(X_n, X_{n+k}) = \text{Cov}(Z_n + \beta \circ Z_{n-1}, Z_{n+k} + \beta \circ Z_{n+k-1}).$$

Para $k = 1$,

$$\text{Cov}(X_n, X_{n+1}) = \text{Cov}(Z_n, \beta \circ Z_n) = \beta \text{Var}(Z_n) = \beta \frac{\theta}{(1 + \beta)}.$$

Mientras que para $k > 1$ se puede observar que $\text{Cov}(X_n, X_{n+k}) = 0$.

Por lo tanto

$$\rho(k) = \begin{cases} 1 & k = 0, \\ \beta/(1 + \beta) & k = 1, \\ 0 & k > 1. \end{cases}$$

Distribución conjunta

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= E[(1 - u)^{X_n} (1 - v)^{X_{n+1}}] \\ &= E[(1 - u)^{Z_n + \beta \circ Z_{n-1}} (1 - v)^{Z_{n+1} + \beta \circ Z_n}] \\ &= E[(1 - u)^{\beta \circ Z_{n-1}}] E[(1 - u)^{Z_n} (1 - \beta v)^{Z_n}] E[(1 - v)^{Z_{n+1}}] \end{aligned}$$

$$\begin{aligned}
&= G_Z(\beta u) G_Z(u + \beta v - \beta uv) G_Z(v) \\
&= \exp \left\{ -\frac{\theta}{1 + \beta} [\beta u + u + \beta v - \beta uv + v] \right\} \\
&= \exp \left\{ -\frac{\theta}{1 + \beta} (1 + \beta) \left[u + v - \frac{\beta}{1 + \beta} uv \right] \right\} \\
&= \exp \{-\theta(u + v - \rho uv)\}, \tag{5.52}
\end{aligned}$$

con $\rho = \rho(1) = \beta/(1 + \beta)$.

Puesto que la f.a.g.p. de orden dos es simétrica en u y v , se tiene que el proceso es reversible en el tiempo, McKenzie (1988a).

Por otra parte, al comparar las ecuaciones (5.52) y (5.13), se concluye que la distribución conjunta de (X_n, X_{n+1}) para el proceso Poisson MA(1), tiene exactamente la misma forma que la correspondiente para el proceso Poisson AR(1). Ya que la distribución conjunta es la misma se sigue que las distribuciones condicionales y sus momentos también lo son, preservándose así las ecuaciones (5.11) y (5.12) -con ρ en lugar de α -. Sin embargo, es muy importante notar que aunque la distribución bivariada sea la misma para los dos procesos, la estructura distribucional de los dos procesos es completamente distinta.

La Figura 5.4 muestra la simulación de un proceso INMA(1) con distribución marginal Poisson de media 5.

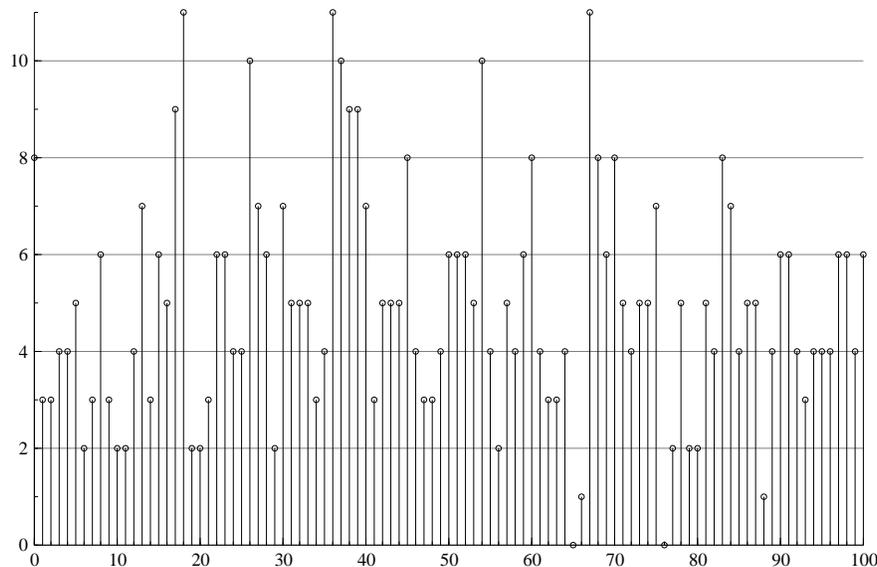


Figura 5.4: Simulación de un proceso Poisson INMA(1), con $\lambda = 5$, $\beta=0.5$ y $n=100$.

5.3.2. Geométrico

Como hemos mencionado en secciones anteriores, la similitud entre la T.L. de una v.a. Exponencial y la f.a.g.p. de una distribución Geométrica se fundamenta en la Proposición 5.9 y, permite que cualquier resultado distribucional para una combinación lineal de v.a.'s Exponenciales puede ser duplicado para v.a.'s Geométricas si el producto αX es reemplazado por $\alpha \circ X$. Por lo tanto, el análogo discreto del modelo (3.12), EMA(1), es

$$X_n = \begin{cases} \beta \circ Z_n & \text{c.p. } \beta, \\ \beta \circ Z_n + Z_{n-1} & \text{c.p. } (1 - \beta), \end{cases} \quad (5.53)$$

$$0 \leq \beta \leq 1, \quad n = 0, \pm 1, \pm 2, \dots,$$

con $\{Z_n\}$ una sucesión de v.a.'s i.i.d. Geom(θ) ($p_Z(k) = \theta^k(1 - \theta)$, $k = 0, 1, \dots$, y $\theta \in (0, 1)$).

Ya hemos calculado en apartados anteriores

$$E[Z_n] = \frac{\theta}{1 - \theta}, \quad \text{Var}(Z_n) = \frac{\theta}{(1 - \theta)^2}, \quad G_Z(s) = \frac{1 - \theta}{1 - \theta + \theta s}.$$

Estacionariedad El proceso (5.53) tiene distribución marginal Geométrica.

$$\begin{aligned} G_{X_n}(s) &= E[(1 - s)^{X_n}] = \beta E[(1 - s)^{\beta \circ Z_n}] + (1 - \beta) E[(1 - s)^{\beta \circ Z_n + Z_{n-1}}] \\ &= \beta G_Z(\beta s) + (1 - \beta) G_Z(\beta s) G_Z(s) \\ &= \beta \frac{1 - \theta}{1 - \theta + \theta \beta s} + (1 - \beta) \frac{1 - \theta}{1 - \theta + \theta \beta s} \frac{1 - \theta}{1 - \theta + \theta s} = \frac{1 - \theta}{1 - \theta + \theta s}. \end{aligned}$$

Por lo tanto las $\{X_n\}$ tienen la misma distribución que las $\{Z_n\}$.

Correlación serial La distribución de $X_n X_{n+1}$ es

$$X_n X_{n+1} = \begin{cases} (\beta \circ Z_n)(\beta \circ Z_{n+1}) & \text{c.p. } \beta^2, \\ (\beta \circ Z_n)(\beta \circ Z_{n+1}) + (\beta \circ Z_n)Z_n & \text{c.p. } \beta(1 - \beta), \\ (\beta \circ Z_n)(\beta \circ Z_{n+1}) + (\beta \circ Z_{n+1})Z_{n-1} & \text{c.p. } \beta(1 - \beta), \\ (\beta \circ Z_n)(\beta \circ Z_{n+1}) + (\beta \circ Z_n)Z_n \\ + (\beta \circ Z_{n+1})Z_{n-1} + Z_{n-1}Z_n & \text{c.p. } (1 - \beta)^2. \end{cases}$$

Entonces,

$$\begin{aligned} E[X_n X_{n+1}] &= E[(\beta \circ Z_n)(\beta \circ Z_{n+1})] + (1 - \beta)E[(\beta \circ Z_n)Z_n] \\ &\quad + (1 - \beta)E[(\beta \circ Z_{n+1})Z_{n-1}] + (1 - \beta)^2 E[Z_{n-1}Z_n]. \end{aligned}$$

Utilizando las propiedades 7 y 9 de la Proposición 4.24, y la independencia entre las Z_n , tenemos

$$E[X_n X_{n+1}] = \beta^2 \frac{1}{\lambda^2} + \beta(1 - \beta) \left(\frac{\theta}{(1 - \theta)^2} + \frac{1}{\lambda^2} \right) + \beta(1 - \beta) \frac{1}{\lambda^2} + (1 - \beta)^2 \frac{1}{\lambda^2},$$

con $\lambda = (1 - \theta)/\theta = 1/\mu = E[X_n] = E[Z_n]$. Agrupando términos

$$E[X_n X_{n+1}] = \frac{1}{\lambda^2} + \beta(1 - \beta) \frac{\theta}{(1 - \theta)^2}.$$

Entonces

$$\rho(1) = \text{Corr}(X_n, X_{n+1}) = \beta(1 - \beta)$$

y el lector puede probar que

$$\rho(k) = 0, \quad \text{si } k > 1.$$

Distribución conjunta

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= E[(1 - u)^{X_n} (1 - v)^{X_{n+1}}] \\ &= \beta^2 E[(1 - u)^{\beta \circ Z_n} (1 - v)^{\beta \circ Z_{n+1}}] \\ &\quad + \beta(1 - \beta) E[(1 - u)^{\beta \circ Z_n} (1 - v)^{\beta \circ Z_{n+1} + Z_n}] \\ &\quad + \beta(1 - \beta) E[(1 - u)^{\beta \circ Z_n + Z_{n-1}} (1 - v)^{\beta \circ Z_{n+1}}] \\ &\quad + (1 - \beta)^2 E[(1 - u)^{\beta \circ Z_n + Z_{n-1}} (1 - v)^{\beta \circ Z_{n+1} + Z_n}]. \end{aligned}$$

Este cálculo es semejante al que se realizó en el modelo EMA(1) para obtener la transformada de Laplace de la distribución conjunta (X_n, X_{n+1}) . Por ello usaremos una notación similar a la que se usó en ese momento.

Sea $\psi(\rho, w) := E[(1 - w)^{\rho \circ Z_n}] = G_{\rho \circ Z_n}(w)$ para cualquier n , con $0 \leq w \leq 1$ y $0 < \rho \leq 1$. Obsérvese que $\psi(1, w) = E[(1 - w)^{Z_n}]$.

Entonces nos queda que

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \psi(\beta, v) \left[\beta^2 \psi(\beta, u) + \beta(1 - \beta) E[(1 - u)^{\beta \circ Z_n} (1 - v)^{Z_n}] \right. \\ &\quad \left. + \beta(1 - \beta) \psi(\beta, u) \psi(1, u) + (1 - \beta)^2 E[(1 - u)^{\beta \circ Z_n} (1 - v)^{Z_n}] \psi(1, u) \right]. \end{aligned}$$

De la demostración de la Proposición 5.2 se sabe que

$$E[(1 - u)^{\beta \circ Z_n} (1 - v)^{Z_n}] = E[(1 - \beta u)^{Z_n} (1 - u)^{Z_n}] = \psi(1, v + \beta u - \beta uv)$$

Sustituyendo este resultado y agrupando términos

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \psi(\beta, v) \left\{ \psi(\beta, u) \left[\beta^2 + \beta(1 - \beta) \psi(1, u) \right] \right. \\ &\quad \left. + \psi(1, v + \beta u - \beta uv) \left[\beta(1 - \beta) + (1 - \beta)^2 \psi(1, u) \right] \right\} \\ &= \frac{\lambda}{\lambda + \beta v} \left\{ \frac{\lambda}{\lambda + \beta u} \frac{\beta(\lambda + \beta u)}{\lambda + u} + \frac{\lambda}{\lambda + v + \beta u - \beta uv} \frac{\lambda + \beta u - \beta \lambda - \beta^2 u}{\lambda + u} \right\} \\ &= \frac{\lambda}{\lambda + \beta v} \frac{\lambda}{\lambda + u} \frac{\lambda + \beta u + \beta v - \beta^2 uv}{\lambda + v + \beta u - \beta uv}. \end{aligned} \tag{5.54}$$

Comparando esta última con (3.15), se tiene que sólo difieren por los términos cruzados uv .

5.3.3. Binomial

De manera análoga al proceso Binomial AR(1), definimos el proceso Binomial MA(1),

$$X_n = Z_n + S(Z_{n-1}), \quad (5.55)$$

con $\{Z_n\}$ una sucesión de v.a.'s i.i.d. Binomial $(N - N_1, p)$, con $N_1 \leq N/2$. Dado $Z_{n-1} = x$, $S(Z_{n-1})$ se distribuye Hipergeométrica $(N - N_1, N_1, x)$. Por lo tanto, Proposición 5.11, $S(Z_{n-1})$ se distribuye Binomial (N_1, p) , y se tiene entonces que (5.55) genera un proceso estacionario con distribución marginal Binomial (N, p) .

Correlación serial Para $k = 1$,

$$\begin{aligned} \text{Cov}(X_n, X_{n+1}) &= \text{Cov}(X_n, Z_{n+1} + S(Z_n)) \\ &= \text{Cov}(X_n, Z_{n+1}) + \text{Cov}(X_n, S(Z_n)), \end{aligned}$$

el primer sumando es cero. Así

$$\text{Cov}(X_n, X_{n+1}) = \text{Cov}(X_n, S(Z_n)) = \text{Cov}([Z_n + S(Z_{n-1})], S(Z_n)) = \text{Cov}(S(Z_n), Z_n).$$

Ahora,

$$\begin{aligned} E[S(Z_n)Z_n] &= E\left[Z_n E[S(Z_n)|Z_n]\right] = E\left[Z_n \frac{N_1}{N - N_1} Z_n\right] = \frac{N_1}{N - N_1} E[Z_n^2] \\ &= \frac{N_1}{N - N_1} (N - N_1)p[\bar{p} + (N - N_1)p] = N_1p[\bar{p} + (N - N_1)p], \end{aligned}$$

restando $E[S(Z_n)]E[Z_n] = N_1p(N - N_1)p$, obtenemos que

$$\text{Cov}(X_n, X_{n+1}) = \text{Cov}(S(Z_n), Z_n) = N_1p\bar{p}. \quad (5.56)$$

Finalmente, puesto que $\text{Var}(X_n) = Np\bar{p}$, se tiene que

$$\rho(1) = \frac{N_1}{N}.$$

En tanto que para $k > 1$, el lector puede corroborar que

$$\rho(k) = 0.$$

Distribución conjunta

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= E[(1 - u)^{Z_n + S(Z_{n-1})} (1 - v)^{Z_{n+1} + S(Z_n)}] \\ &= G_{S(Z_{n-1})}(u) G_{Z_{n+1}}(v) E\left[(1 - u)^{Z_n - S(Z_n)} (1 - v)^{S(Z_n)} (1 - u)^{S(Z_n)}\right] \\ &= (1 - pu)^{N_1} (1 - pv)^{N - N_1} G_{Z_n - S(Z_n)}(u) G_{S(Z_n)}(u, v) \\ &= (1 - pu)^{N - N_1} (1 - pv)^{N - N_1} (1 - pu - pv + puv)^{N_1}. \end{aligned}$$

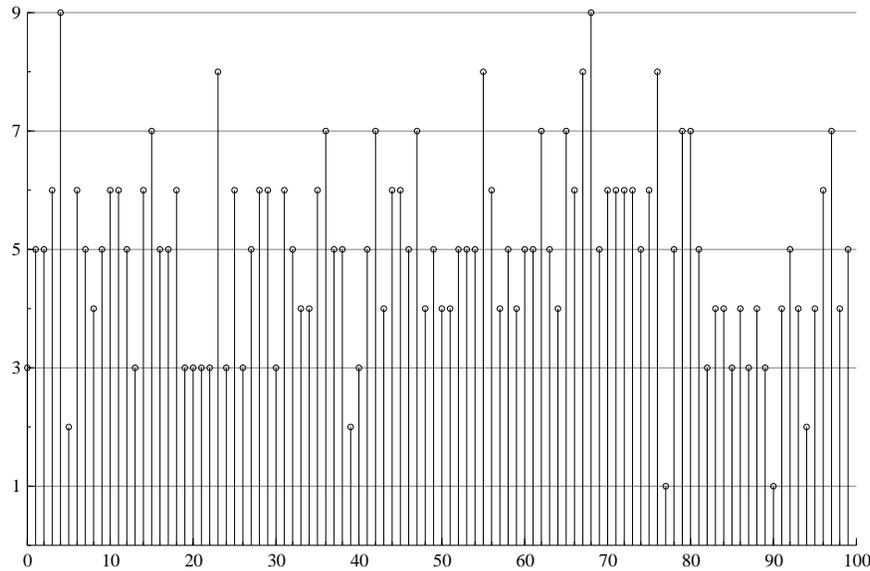


Figura 5.5: Simulación de un proceso Binomial MA(1) con $N=10$, $p=0.5$ y $\rho(1)=0.5$ y $n=100$.

Puesto que la f.a.g.p. conjunta de X_n y X_{n+1} del proceso Binomial MA(1) tiene la misma forma que la correspondiente del proceso Binomial AR(1), ecuación (5.49), entonces la distribución conjunta es la misma, al igual que las distribuciones condicionales y sus momentos, ecuaciones (5.47) y (5.48). Sin embargo, es importante notar que aunque la distribución bivariada es la misma para los dos procesos, la estructura distribucional de los dos procesos es completamente distinta.

La Figura 5.5 muestra la simulación de un proceso Binomial MA(1).

5.4. ARMA(1,1) discretos

5.4.1. Poisson

Definimos el proceso ARMA(1,1) con distribución marginal Poisson combinando los dos procesos Poisson vistos anteriormente, INAR(1) y MA(1) discreto. La idea consiste en enlazar los dos procesos por medio de un proceso de innovación, $\{Z_n\}$, común. En concreto

$$X_n = \beta \circ Z_n + M_{n-1}, \quad \beta \in [0, 1], \quad (5.57)$$

con

$$M_n = \alpha \circ M_{n-1} + Z_n, \quad \alpha \in [0, 1], \quad (5.58)$$

donde $\{Z_n\}$ es una sucesión de v.a.'s i.i.d., e independientes de todas las series de conteo involucradas tanto en (5.57) como en (5.58), las cuales son todas mutuamente independientes.

Estacionariedad De la sección 5.1.1 tenemos que si $\{Z_n\}$ son v.a.'s i.i.d. $\text{Poisson}(\bar{\alpha}\theta)$, y M_0 una v.a. independiente $\text{Poisson}(\theta)$, entonces $\{M_n\}$ es un proceso estacionario AR(1) Poisson de media θ . Con tales condiciones $\{X_n\}$ es un proceso estacionario con distribución marginal $\text{Poisson}(\theta(1+\bar{\alpha}\beta))$ (véase por ejemplo la ecuación (5.51)).

Correlación serial

$$\begin{aligned} c(k) &= \text{Cov}(X_n, X_{n+k}) = \text{Cov}(\beta \circ Z_n + M_{n-1}, \beta \circ Z_{n+k} + M_{n+k-1}) \\ &= \text{Cov}(\beta \circ Z_n, M_{n+k-1}) + \text{Cov}(M_{n-1}, M_{n+k-1}). \end{aligned}$$

Para $k = 1$,

$$\begin{aligned} c(1) &= \text{Cov}(X_n, X_{n+1}) = \text{Cov}(\beta \circ Z_n, M_n) + \text{Cov}(M_{n-1}, M_n) \\ &= \text{Cov}(\beta \circ Z_n, \alpha \circ M_{n-1} + Z_n) + \text{Cov}(M_{n-1}, \alpha \circ M_{n-1} + Z_n) \\ &= \text{Cov}(\beta \circ Z_n, Z_n) + \text{Cov}(M_{n-1}, \alpha \circ M_{n-1}) \\ &= \beta \text{Var}(Z) + \alpha \text{Var}(M) \quad (\text{Proposición 4.24}) \\ &= \beta \bar{\alpha} \theta + \alpha \theta = \theta(\alpha + \beta \bar{\alpha}). \end{aligned}$$

Para $k > 1$

$$\begin{aligned} c(k) &= \text{Cov}(X_n, X_{n+k}) = \text{Cov}(\beta \circ Z_n, M_{n+k-1}) + \text{Cov}(M_{n-1}, M_{n+k-1}) \\ &= \text{Cov}(\beta \circ Z_n, \alpha \circ M_{n+k-2} + Z_{n+k-1}) + \text{Cov}(M_{n-1}, \alpha \circ M_{n+k-2} + M_{n+k-1}) \\ &= \text{Cov}(\beta \circ Z_n, \alpha \circ M_{n+k-2}) + \text{Cov}(M_{n-1}, \alpha \circ M_{n+k-2}) \\ &= \alpha \left[\text{Cov}(\beta \circ Z_n, M_{n+k-2}) + \text{Cov}(M_{n-1}, M_{n+k-2}) \right] \\ &= \alpha c(k-1) = \alpha^{k-1} c(1) = \alpha^{k-1} \theta(\alpha + \beta \bar{\alpha}). \end{aligned}$$

Por lo tanto

$$\rho(k) = \begin{cases} 1 & k = 0, \\ (\alpha + \bar{\alpha}\beta)/(1 + \bar{\alpha}\beta) & k = 1, \\ \alpha^{k-1} \rho(1) & k > 1. \end{cases}$$

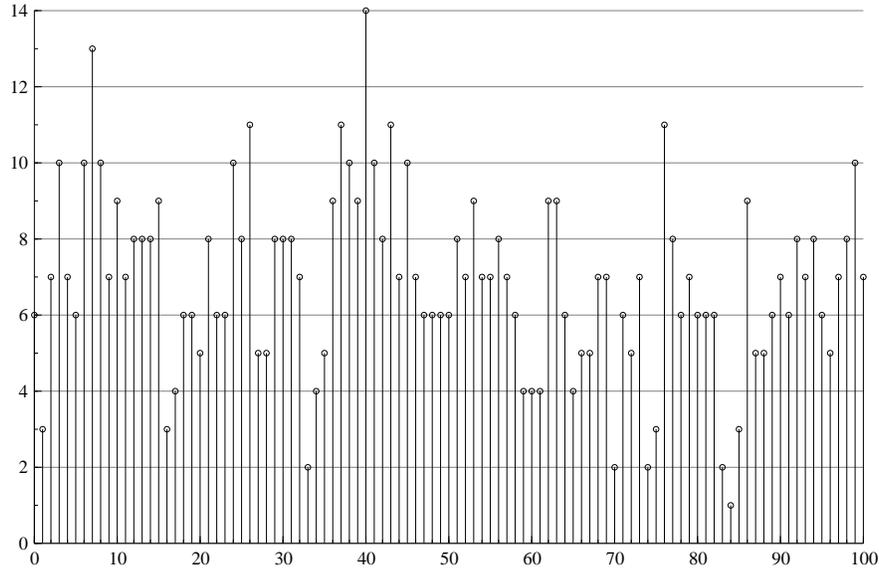


Figura 5.6: Poisson ARMA(1,1), con $\lambda = 5$, $\alpha = 0.5 = \beta$ y $n = 100$.

Distribución conjunta

$$\begin{aligned}
 G_{X_n, X_{n+1}}(u, v) &= E[(1-u)^{X_n} (1-v)^{X_{n+1}}] \\
 &= E\left[(1-u)^{\beta \circ Z_n + M_{n-1}} (1-v)^{\beta \circ Z_{n+1} + M_n}\right] \\
 &= E\left[(1-u)^{\beta \circ Z_n} (1-v)^{\beta \circ Z_{n+1}} (1-u)^{M_{n-1}} (1-v)^{\alpha \circ M_{n-1} + Z_n}\right] \\
 &= E\left[(1-v)^{\beta \circ Z_{n+1}} (1-u)^{\beta \circ Z_n} (1-v)^{Z_n} (1-u)^{M_{n-1}} (1-v)^{\alpha \circ M_{n-1}}\right] \\
 &= G_Z(\beta v) G_Z(v + \beta u - \beta uv) G_M(u + \alpha v - \alpha uv) \\
 &= \exp\{-\bar{\alpha}\theta(\beta v + v + \beta u - \beta uv)\} \exp\{-\theta(u + \alpha v - \alpha uv)\} \\
 &= \exp\{-\theta\beta\bar{\alpha}(u + v - uv) - \theta(u + v - \alpha uv)\}.
 \end{aligned}$$

La Figura 5.6 es la simulación de un proceso ARMA(1,1) con distribución marginal Poisson de media 5.

5.4.2. Geométrico

Consideremos esta vez, lo que sería la versión discreta del modelo Exponencial ARMA(1,1), EARMA(1,1), dada por

$$X_n = \begin{cases} \beta \circ Z_n & \text{c.p. } \beta, \\ \beta \circ Z_n + A_{n-1} & \text{c.p. } (1 - \beta), \end{cases} \quad (5.59)$$

$$0 \leq \beta \leq 1, \quad n = 1, 2, \dots,$$

con

$$A_n = \begin{cases} \alpha \circ A_{n-1} & \text{c.p. } \alpha, \\ \alpha \circ A_{n-1} + Z_n & \text{c.p. } (1 - \alpha), \end{cases} \quad (5.60)$$

$$0 \leq \alpha < 1, \quad n = 1, 2, \dots$$

con $\{Z_n\}$ una sucesión de v.a.'s i.i.d. $\text{Geom}(\theta)$.

Obviamente de los procesos INAR(1) y MA(1) con esta distribución marginal, se tiene que la condición de estacionariedad para el procesos descrito por (5.59) y (5.60), Geométrico ARMA(1,1), es $A_0 = Z_0$, esto es, A_0 es una v.a. $\text{Geom}(\theta)$ independiente del resto de las Z_n .

Correlación serial El modelo Geométrico ARMA(1,1) posee la misma correlación serial que su análogo continuo.

Escribiendo la distribución de $X_n X_{n+k}$, como se hizo para el modelo MA(1) Geométrico, y tomando esperanza nos queda:

Para $k \neq 0$,

$$E[X_n X_{n+k}] = E[(\beta \circ Z_n)(\beta \circ Z_{n+k})] + (1 - \beta)E[(\beta \circ Z_n)A_{n+k-1}] \\ + (1 - \beta)E[(\beta \circ Z_{n+k})A_{n-1}] + (1 - \beta)^2 E[A_{n-1}A_{n+k-1}].$$

Aplicando 7 y 9 de la Proposición 4.24, la independencia entre las variables y haciendo $\lambda = (1 - \theta)/\theta$,

$$E[X_n X_{n+k}] = \beta^2 \frac{1}{\lambda^2} + \beta(1 - \beta)E[Z_n A_{n+k-1}] + \beta(1 - \beta) \frac{1}{\lambda^2} + (1 - \beta)^2 E[A_{n-1}A_{n+k-1}] \\ = \beta(1 - \beta)E[Z_n A_{n+k-1}] + (1 - \beta)^2 E[A_{n-1}A_{n+k-1}] + \frac{\beta}{\lambda^2}. \quad (5.61)$$

Compárese esta última con (3.18). Puesto que $E[(\alpha \circ A_{n+i})Z_n] = \alpha E[A_{n+i}Z_n]$ para cualquier i , varios de los cálculos que se realizaron para el modelo EARMA(1,1) se preservan aquí. Entre ellos

$$E[Z_n A_n] = E[\rho \circ A_{n-1} Z_n] + (1 - \rho)E[Z_n^2] \\ = \rho E[A_{n-1} Z_n] + (1 - \rho)E[Z_n^2] \\ = \rho \frac{1}{\lambda^2} + (1 - \rho) \left[\frac{\theta}{(1 - \theta^2)} + \frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2} + (1 - \alpha) \frac{\theta}{(1 - \theta)^2}.$$

Y para $k \geq 1$,

$$E[Z_n A_{n+k}] = \alpha^k E[Z_n A_n] + \frac{(1 - \alpha^k)}{\lambda^2} \quad (\text{por (3.19)}) \\ = \alpha^k \left[\frac{1}{\lambda^2} + (1 - \alpha) \frac{\theta}{(1 - \theta)^2} \right] + \frac{(1 - \alpha^k)}{\lambda^2} = \alpha^k (1 - \alpha) \frac{\theta}{(1 - \theta)^2} + \frac{1}{\lambda^2}.$$

Así mismo puesto que $\{A_n\}$ es un proceso INAR(1) con distribución marginal $\text{Geom}(\theta)$, $\rho_A(k) = \alpha^k$,

y

$$E[A_n A_{n+k}] = \alpha^k \frac{\theta}{(1 - \theta)^2} + \frac{1}{\lambda^2}.$$

Sustituyendo estos resultados en (5.61),

$$\begin{aligned} E[X_n X_{n+k}] &= \beta(1-\beta) \left[\alpha^{k-1}(1-\alpha) \frac{\theta}{(1-\theta)^2} + \frac{1}{\lambda^2} \right] + (1-\beta)^2 \left[\alpha^k \frac{\theta}{(1-\theta)^2} + \frac{1}{\lambda^2} \right] + \frac{\beta}{\lambda^2} \\ &= \beta(1-\beta)(1-\rho)\rho^{k-1} \frac{\theta}{(1-\theta)^2} + (1-\beta)^2 \rho^k \left(\frac{\theta}{(1-\theta)^2} \right) + \frac{1}{\lambda^2}. \end{aligned}$$

Restando $E^2[X_n]$ y dividiendo entre la $\text{Var}(X_n)$,

$$\begin{aligned} \rho(k) &= \beta(1-\beta)(1-\rho)\rho^{k-1} + (1-\beta)^2 \rho^k \\ &= \bar{\beta}(\beta\bar{\rho} + \rho\bar{\beta})\rho^{k-1}, \quad k = 1, 2, \dots \end{aligned}$$

Distribución conjunta Siguiendo el mismo procedimiento que se utilizó en el modelo Geométrico MA(1) y usando la misma notación, se tiene

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \beta^2 \psi(\beta, u) \psi(\beta, v) + \beta(1-\beta) \left[\alpha \psi(\beta, u) \psi(\beta, v) \phi(\alpha, v) \right. \\ &\quad \left. + (1-\alpha) E[(1-u)^{\beta \circ Z_n} (1-v)^{Z_n}] \psi(\beta, v) \phi(\alpha, v) \right] + \beta(1-\beta) \psi(\beta, u) \phi(1, u) \psi(\beta, v) \\ &\quad + (1-\beta)^2 \left[\alpha \psi(\beta, u) \psi(\beta, v) E[(1-u)^{A_n-1} (1-v)^{\alpha \circ A_n-1}] \right. \\ &\quad \left. + (1-\alpha) E[(1-u)^{\beta \circ Z_n} (1-v)^{Z_n}] \psi(\beta, v) E[(1-u)^{A_n-1} (1-v)^{\alpha \circ A_n-1}] \right]. \end{aligned}$$

De la demostración de la Proposición 5.2 se sabe que

$$E[(1-u)^{\beta \circ Z_n} (1-v)^{Z_n}] = E[(1-\beta u)^{Z_n} (1-u)^{Z_n}] = \psi(1, v + \beta u - \beta uv)$$

$$\text{y, } E[(1-u)^{A_n} (1-v)^{\alpha \circ A_n}] = E[(1-u)^{A_n} (1-\alpha u)^{A_n}] = \psi(1, u + \alpha v - \alpha uv).$$

Usando estos resultados y agrupando términos

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \beta \psi(\beta, u) \psi(\beta, v) \left[\beta + (1-\beta) \phi(1, u) \right] \\ &\quad + \left[\alpha \psi(\beta, u) + (1-\alpha) \psi(1, v + \beta u - \beta uv) \right] \psi(\beta, v) \left[\beta(1-\beta) \phi(\alpha, v) \right. \\ &\quad \left. + (1-\beta)^2 \psi(\beta, v) \psi(1, u + \alpha v - \alpha uv) \right]. \end{aligned}$$

Sustituyendo

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \beta \psi(\beta u) \psi(\beta, v) \left(\frac{\lambda + \beta u}{\lambda + u} \right) \\ &\quad + \left(\frac{\lambda(\lambda + \beta u + \alpha v - \alpha \beta uv)}{(\lambda + \beta u)(\lambda + \beta u + v - \beta uv)} \right) \psi(\beta, v) \left(\frac{(1-\beta)\lambda(\lambda + \beta u + \alpha v - \alpha \beta uv)}{(\lambda + \alpha v)(\lambda + u + \alpha v - \alpha uv)} \right). \end{aligned}$$

Finalmente

$$\begin{aligned} G_{X_n, X_{n+1}}(u, v) &= \left(\frac{\lambda}{\lambda + \beta u} \right) \left(\frac{\lambda}{\lambda + \beta v} \right) \\ &\quad \times \left[\frac{\beta(\lambda + \beta u)}{(\lambda + u)} + \frac{(1-\beta)\lambda(\lambda + \beta u + \alpha v - \alpha \beta uv)^2}{(\lambda + \beta u + v - \beta uv)(\lambda + \alpha v)(\lambda + u + \alpha v - \alpha uv)} \right]. \end{aligned}$$

Compárese con (3.23). De nuevo, la diferencia entre la versión discreta y la continua son sólo los términos cruzados, uv .

Además, si

$\beta = 0$, entonces se tiene la f.a.g.p. conjunta del modelo INAR(1) con distribución marginal Geométrica, ecuación (5.18), y cuando

$\alpha = 0$, tenemos la f.a.g.p. conjunta del modelo Geométrico MA(1), ecuación (5.54).

5.4.3. Binomial

Definimos el proceso Binomial ARMA(1,1), $\{X_n\}$, como

$$X_n = Y_{n-1} + S(Z_n), \quad (5.62)$$

con

$$Y_n = S(Y_{n-1}) + Z_n. \quad (5.63)$$

Sea $\{Z_n\}$ una sucesión de v.a.'s i.i.d. con distribución Binomial($N - M, p$), Y_0 una v.a. independiente Binomial(N, p) y $(S(Y_{n-1}) \mid Y_{n-1} = y)$ con distribución Hipergeométrica(N, M, y), entonces $\{Y_n\}$ es un proceso AR(1) estacionario con distribución marginal Binomial(N, p). Además, dado $Z_n = x$, suponemos que $(S(Z_n) \mid Z_n) \sim \text{Hipergeométrica}(N - M, N_1, x)$, con $N_1 < N - M$, entonces $S(Z_n)$ se distribuye Binomial(N_1, p), y por tanto el proceso $\{X_n\}$ es estacionario con distribución marginal Binomial($N + N_1, p$).

Correlación serial

$$\begin{aligned} g(1) &= \text{Cov}(X_n, X_{n+1}) = \text{Cov}(Y_{n-1} + S(Z_n), Y_n + S(Z_{n+1})) \\ &= \text{Cov}(Y_{n-1}, Y_n) + \text{Cov}(S(Z_n), Y_n) \\ &= \text{Cov}(Y_{n-1}, Y_n) + \text{Cov}(S(Z_n), S(Y_{n-1}) + Z_n) \\ &= \text{Cov}(Y_{n-1}, Y_n) + \text{Cov}(S(Z_n), Z_n). \end{aligned}$$

Para el primer sumando puesto que $\{Y_n\}$ es un proceso estacionario BAR(1), $\rho(1) = \alpha = M/N$. Mientras que para el segundo, seguimos un procedimiento idéntico al que se realizó para obtener (5.56). Quedando así

$$\begin{aligned} \text{Cov}(X_n, X_{n+1}) &= \alpha \text{Var}(Y_n) + N_1 p \bar{p} \\ &= \alpha N p \bar{p} + N_1 p \bar{p} = (M + N_1) p \bar{p}. \end{aligned}$$

Para $k > 1$,

$$\begin{aligned} \text{Cov}(X_n, X_{n+k}) &= \text{Cov}(Y_{n-1} + S(Z_n), Y_{n+k-1} + S(Z_{n+k})) \\ &= \text{Cov}(Y_{n-1}, Y_{n+k-1}) + \text{Cov}(S(Z_n), Y_{n+k-1}). \end{aligned}$$

El primer sumando se obtiene recursivamente

$$c(k) = \text{Cov}(Y_{n-1}, Y_{n+k-1}) = \alpha c(k-1) = \alpha^{k-1} c(1) = \alpha^k \text{Var}(Y) = \left(\frac{M}{N}\right)^k N p \bar{p}. \quad (5.64)$$

Para el segundo:

Por demostrar

$$\text{Cov}(S(Z_n), Y_{n+k}) = \left(\frac{M}{N}\right)^k N_1 p \bar{p}. \quad (5.65)$$

Para $k = 1$,

$$\text{Cov}(S(Z_n), Y_{n+1}) = \text{Cov}(S(Z_n), S(Y_n) + Z_{n+1}) = \text{Cov}(S(Z_n), S(Y_n)).$$

Ahora bien

$$\begin{aligned} \text{Cov}(S(Z_n), S(Y_n)) &= E[S(Z_n)E[S(Y_n) | Y_n]] - E[S(Z_n)]E[E[S(Y_n) | Y_n]] \\ &= \frac{M}{N} E[S(Z_n)Y_n] - \frac{M}{N} E[S(Z_n)]E[Y_n], \end{aligned} \quad (5.66)$$

puesto que $S(Y_n) | Y_n \sim \text{Hipergeométrica}(N, M, Y_n)$. Entonces

$$\begin{aligned} \text{Cov}(S(Z_n), S(Y_n)) &= \frac{M}{N} \text{Cov}(S(Z_n), Y_n) \\ &= \frac{M}{N} \text{Cov}(S(Z_n), S(Y_{n-1}) + Z_n) = \frac{M}{N} \text{Cov}(S(Z_n), Z_n) = \frac{M}{N} N_1 p \bar{p}. \end{aligned}$$

Por lo tanto

$$\text{Cov}(S(Z_n), Y_{n+1}) = \frac{M}{N} N_1 p \bar{p}.$$

Supongamos que la relación (5.65) es válida para $k - 1$,

$$\begin{aligned} \text{Cov}(S(Z_n), Y_{n+k}) &= \text{Cov}(S(Z_n), S(Y_{n+k-1}) + Z_{n+k}) \\ &= \text{Cov}(S(Z_n), S(Y_{n+k-1})) = \frac{M}{N} \text{Cov}(S(Z_n), Y_{n+k-1}), \end{aligned}$$

esta última se obtiene de manera totalmente análoga a (5.66). Luego, aplicando la hipótesis de inducción

$$\text{Cov}(S(Z_n), Y_{n+k}) = \frac{M}{N} \left(\frac{M}{N}\right)^{k-1} N_1 p \bar{p} = \left(\frac{M}{N}\right)^k N_1 p \bar{p}.$$

L.Q.Q.D.

Entonces de (5.64) y (5.65),

$$\text{Cov}(X_n, X_{n+k}) = \left(\frac{M}{N}\right)^{k-1} (M p \bar{p} + N_1 p \bar{p}) = \left(\frac{M}{N}\right)^{k-1} g(1).$$

De estos cálculos se tiene

$$\rho(k) = \begin{cases} (M + N_1)/(N + N_1) & \text{si } k = 1, \\ \left(\frac{M}{N}\right)^{k-1} \rho(1) & k > 1. \end{cases}$$

Haciendo $\alpha = M/N$, $\beta = N_1/(N - M)$ y $d = N_1/N$,

$$\rho(k) = \begin{cases} (\alpha + \bar{\alpha}\beta)/(1 + d) & k = 1, \\ \alpha^{k-1} \rho(1) & k > 1. \end{cases}$$

Distribución conjunta

$$\begin{aligned}
G_{X_n, X_{n+1}}(u, v) &= E[(1-u)^{Y_{n-1}+S(Z_n)} (1-v)^{Y_n+S(Z_{n+1})}] \\
&= E[(1-u)^{Y_{n-1}+S(Z_n)} (1-v)^{S(Y_{n-1})+Z_n+S(Z_{n+1})}] \\
&= E[(1-u)^{Y_{n-1}-S(Y_{n-1})} (1-v)^{S(Y_{n-1})} (1-u)^{S(Y_{n-1})}] \\
&\quad \times E[(1-v)^{Z_n-S(Z_n)} (1-u)^{S(Z_n)} (1-v)^{S(Z_n)}] E[(1-v)^{S(Z_{n+1})}] \\
&= (1-pu)^{N-M} (1-pu-pv+puv)^M (1-pv)^{N-M-N_1} (1-pu-pv+puv)^{N_1} (1-pv)^{N_1} \\
&= (1-pu)^{N-M} (1-pu-pv+puv)^{M+N_1} (1-pv)^{N-M}.
\end{aligned}$$

Capítulo 6

Modelo AR(1) con marginal ED

Harry Joe en 1996 publicó un artículo sobre series de tiempo considerando como marginal familias paramétricas infinitamente-divisibles cerradas bajo convolución. Dicho trabajo unifica y generaliza trabajos previos que habían abordado la cuestión de modelos de series de tiempo con distribución marginal no-Gaussiana, como Lewis (1983) y Lewis *et al.* (1989) para el caso de la distribución Gamma, McKenzie (1986) para la distribución Binomial Negativa, McKenzie (1988b) para la Poisson y Al-Osh y Alzaid (1991) para la Binomial, en lo que se refiere a la estructura AR(1). Así mismo, su enfoque permite la creación de un modelo AR(1) con distribución marginal Normal distinto al de la ecuación (2.4).

Dos años después Bent Jørgensen y Peter Xue-Kun Song publican un artículo que trata modelos de series de tiempo con marginal en la clase de *modelos de dispersión exponencial* infinitamente-divisibles, la cual es un subconjunto de la clase tratada por Joe (1996). Los modelos construidos por Jørgensen y Song (1998) están basados en el operador de adelgazamiento construido por Joe (1996), el cual se reduce al operador de adelgazamiento Binomial y al adelgazamiento Hipergeométrico cuando la distribución que se trata es la Poisson y la Binomial respectivamente.

Aunque ambos artículos tratan procesos de la forma ARMA, nosotros nos limitamos al estudio del proceso AR(1).

6.1. Modelos de dispersión exponencial

Sea $X \sim ED^*(\theta, \lambda)$ que denota un modelo de dispersión exponencial con función de densidad de probabilidad

$$f_X(x; \theta, \lambda) = c(x; \lambda) \exp\{\theta x - \lambda k(\theta)\}, \quad x \in \mathbb{R}, \quad (6.1)$$

donde θ es llamado el parámetro canónico cuyo dominio es el conjunto $\Theta = \{\theta \in \mathbb{R}; k(\theta) < \infty\}$, λ el *parámetro índice* con $\lambda \in \Lambda \subseteq \mathbb{R}^+$ y, $k(\theta) = \ln \int_{\mathbb{R}} c(x; \lambda) e^{\theta x} d\eta(x) / \lambda$, donde η denota la medida de Lebesgue o la medida de conteo, en el último caso la integral se reemplaza por una suma.

Observe que (6.1) es una familia exponencial cuando λ es conocida.

Se sabe además que el modelo es infinitamente-divisible si y solo si el dominio del parámetro índice es $\Lambda = \mathbb{R}^+$.

Puesto que (6.1) es una función de densidad

$$\int_{\mathbb{R}} c(x; \lambda) e^{\theta x} d\eta(x) = e^{\lambda k(\theta)}. \quad (6.2)$$

Calculemos ahora la función generadora de momentos (f.g.m.) de $X \sim ED^*(\theta, \lambda)$,

$$\begin{aligned} M_X(t) &= E[e^{tX}] = e^{-\lambda k(\theta)} \int_{-\infty}^{\infty} c(x; \lambda) e^{(\theta+t)x} d\eta(x) \\ &= \exp \left\{ \lambda \{k(\theta+t) - k(\theta)\} \right\}, \quad \text{por (6.2)}. \end{aligned}$$

De donde podemos obtener fácilmente que

$$E[X] = \lambda k'(\theta) \quad \text{y} \quad \text{Var}(X) = \lambda k''(\theta). \quad (6.3)$$

Sean Z_1 y Z_2 dos variables aleatorias independientes tales que

$$Z_i \sim ED^*(\theta, \lambda_i), \quad i = 1, 2.$$

Y sea $Z = Z_1 + Z_2$, entonces

$$M_{Z_1+Z_2}(t) = M_{Z_1}(t) M_{Z_2}(t) = e^{\lambda_1 + \lambda_2 \{k(\theta+t) - k(\theta)\}} = M_Z(t),$$

por lo que $Z \sim ED^*(\theta, \lambda_1 + \lambda_2)$.

Es decir, se tiene que

$$ED^*(\theta, \lambda_1) \otimes ED^*(\theta, \lambda_2) = ED^*(\theta, \lambda_1 + \lambda_2). \quad (6.4)$$

Entonces cualquier modelo de dispersión exponencial satisface la fórmula (6.4), se dice entonces que la familia es cerrada bajo convolución.

6.1.1. Ejemplos

Una forma de precisar un modelo de dispersión exponencial es especificando la función $k(\theta)$, llamada el *cumulante generador de* (6.1), o equivalentemente la función $c(\cdot; \lambda)$.

Veamos algunos ejemplos.

Ejemplo 6.1 Sea $k(\theta) = \theta^2/2$, entonces

$$f_X(x; \theta, \lambda) = c(x; \lambda) \exp\{\theta x - \lambda \theta^2/2\},$$

lo que implica que

$$c(x; \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{-\frac{x^2}{2\lambda}\right\}, \quad \lambda > 0.$$

Así, $f_X(x; \theta, \lambda)$ toma la forma

$$f_X(x; \theta, \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp\left\{-\frac{(x - \theta\lambda)^2}{2\lambda}\right\},$$

i.e.

$$X \sim ED^*(\theta, \lambda) = \text{Normal}(\theta\lambda, \lambda). \quad \square$$

Ejemplo 6.2 Sea $k(\theta) = -\ln(-\theta)$, con $\theta < 0$,

$$f_X(x; \theta, \lambda) = c(x; \lambda) \exp\{\theta x + \lambda \ln(-\theta)\} = c(x; \lambda) e^{\theta x} (-\theta)^\lambda,$$

entonces

$$c(x; \lambda) = \frac{x^{\lambda-1}}{\Gamma(\lambda)}, \quad \lambda > 0,$$

quedando que

$$f_X(x; \theta, \lambda) = \frac{(-\theta)^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{\theta x}.$$

Por lo tanto

$$X \sim ED^*(\theta, \lambda) = \text{Gamma}(\lambda, -\theta), \text{ con } \theta < 0. \quad \square$$

Ejemplo 6.3 Sea $k(\theta) = -\ln(1 - e^\theta)$, con $\theta < 0$,

$$f_X(x; \theta, \lambda) = c(x; \lambda) e^{\theta x} (1 - e^\theta)^\lambda,$$

implicando que

$$c(x; \lambda) = \binom{\lambda + x - 1}{\lambda - 1} = \frac{\Gamma(\lambda + x)}{\Gamma(\lambda) x!}, \quad \lambda > 0.$$

Sustituyendo

$$f_X(x; \theta, \lambda) = \binom{\lambda + x - 1}{\lambda - 1} e^{\theta x} (1 - e^\theta)^\lambda,$$

entonces

$$X \sim ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta), \text{ con } \theta < 0,$$

en términos de la notación del Capítulo 1. \square

Ejemplo 6.4 Sea $k(\theta) = e^\theta$,

$$f_X(x; \theta, \lambda) = c(x; \lambda) \exp\{\theta x - \lambda e^\theta\},$$

lo que implica que

$$c(x; \lambda) = \frac{\lambda^x}{x!}, \quad \lambda > 0.$$

Se tiene así que

$$f_X(x; \theta, \lambda) = \frac{(\lambda e^\theta)^x}{x!} \exp\{-\lambda e^\theta\},$$

i.e.

$$X \sim ED^*(\theta, \lambda) = \text{Poisson}(\lambda e^\theta). \quad \square$$

Ejemplo 6.5 Sea $k(\theta) = \ln(1 + e^\theta)$,

$$f_X(x; \theta, \lambda) = c(x; \lambda) e^{\theta x} (1 + e^\theta)^{-\lambda},$$

entonces

$$c(x; \lambda) = \binom{\lambda}{x}, \quad \lambda \in \mathbb{N}.$$

Sustituyendo

$$f_X(x; \theta, \lambda) = \binom{\lambda}{x} \left(\frac{e^\theta}{1 + e^\theta} \right)^x \left(\frac{1}{1 + e^\theta} \right)^{\lambda - x},$$

por lo tanto

$$X \sim ED^*(\theta, \lambda) = \text{Binomial}(\lambda, e^\theta / (1 + e^\theta)). \quad \square$$

La clase de modelos de dispersión exponencial (infinitamente-divisibles) es muy extensa. De hecho cualquier distribución (I-D) con f.g.m. genera un modelo de dispersión exponencial (I-D) (Jørgensen 1986).

Joe (1996) considera la clase más general de distribuciones infinitamente-divisibles de convolución cerrada que incluye distribuciones sin momentos. Sin embargo estas distribuciones pueden ser vistas como la familia de modelos de dispersión exponencial con $\Theta = \{0\}$, esto es, $\{ED^*(0, \lambda); \lambda > 0\}$.

6.2. Operador de adelgazamiento de Joe

Sean Z_1 y Z_2 dos variables aleatorias independientes tales que

$$Z_i \sim ED^*(\theta, \lambda_i), \quad i = 1, 2. \quad (6.5)$$

Hemos visto que $Z = Z_1 + Z_2 \sim ED(\theta, \lambda)$, con $\lambda = \lambda_1 + \lambda_2$. Ya que la suma es suficiente para el parámetro θ , la distribución condicional de Z_1 dado $Z = x$ no depende de θ y es denotada por $Z_1 | Z_1 + Z_2 \sim G(\lambda_1, \lambda_2, x)$, cuya función de densidad es

$$\begin{aligned} f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) &= \frac{f_{Z_1}(w; \theta, \lambda_1) f_{Z_2}(x-w; \theta, \lambda_2)}{f_Z(x; \theta, \lambda_1 + \lambda_2)} \\ &= \frac{c(w; \lambda_1) \exp\{\theta w - \lambda_1 k(\theta)\} c(x-w; \lambda_2) \exp\{\theta(x-w) - \lambda_2 k(\theta)\}}{c(x; \lambda_1 + \lambda_2) \exp\{\theta x - (\lambda_1 + \lambda_2) k(\theta)\}} \\ &= \frac{c(w; \lambda_1) c(x-w; \lambda_2)}{c(x; \lambda)}. \end{aligned} \quad (6.6)$$

La distribución $G(\lambda_1, \lambda_2, x)$ es llamada la *contracción* correspondiente a $ED^*(\theta, \lambda)$.

Definimos ahora la operación de adelgazamiento, $A(X; \alpha)$, con $\alpha \in [0, 1]$.

Sea X una v.a. con distribución $ED^*(\theta, \lambda)$, un modelo de dispersión exponencial infinitamente-divisible, y sea $\bar{\alpha} = 1 - \alpha$. Definimos la v.a. $A(X; \alpha)$ por la distribución condicional

$$A(X; \alpha) | X = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x), \quad (6.7)$$

entonces la distribución marginal de $A(X; \alpha)$ es $ED^*(\theta, \alpha\lambda)$ (Proposición 6.6), y decimos que $A(X; \alpha)$ es el *adelgazamiento de X por una proporción α* . Aunque bien podría llamarse el *adelgazamiento de Joe*. Se pide que X sea un modelo de dispersión exponencial infinitamente-divisible para asegurar que tanto $\alpha\lambda$ como $\bar{\alpha}\lambda$ sean valores válidos para el parámetro índice para cualquier valor de $\alpha \in (0, 1)$ (véase al respecto el Ejemplo 6.12). En los casos extremos, $\alpha = 0, 1$, definimos $A(X; 1) = 1$, y $A(X; 0) = 0$.

Proposición 6.6 Sea $ED^*(\theta, \lambda)$ un modelo de dispersión exponencial con conjunto índice Λ , y supongamos que $X \sim ED^*(\theta, \lambda)$, y $Y | X = x \sim G(\lambda_1, \lambda_2, x)$, donde $\lambda_1, \lambda_2, \lambda \in \Lambda$, y $\lambda = \lambda_1 + \lambda_2$. Entonces Y y $X - Y$ son independientes, $Y \sim ED^*(\theta, \lambda_1)$ y $X - Y \sim ED^*(\theta, \lambda_2)$.

Demostración: Supongamos que Z_1 y Z_2 son dos v.a.'s independientes con distribución (6.5), y sea $Z = Z_1 + Z_2$. Entonces la distribución condicional de $Z_1 | Z$ y la distribución marginal de Z son las mismas que la distribución condicional de $Y | X$ y la distribución marginal de X , respectivamente. Por lo tanto la distribución conjunta de $(Z_1, Z_1 + Z_2)$ es la misma que la distribución conjunta de (Y, X) , lo que a su vez implica que la distribución conjunta de (Z_1, Z_2) es la misma que la distribución conjunta de $(Y, X - Y)$.

Por lo tanto Y y $X - Y$ son independientes, $Y \sim ED^*(\theta, \lambda_1)$ y $X - Y \sim ED^*(\theta, \lambda_2)$. \square

Vale la pena mencionar que con el resultado anterior queda probada la independencia entre $S(X_{n-1})$ y $X_{n-1} - S(X_{n-1})$, en la Proposición 5.11, la cual hasta el momento no se había demostrado, y sin embargo se utilizó de manera intensa en aquel apartado.

En adelante, asuma $\lambda_1 = \alpha\lambda$, $\lambda_2 = \bar{\alpha}\lambda$, y $\alpha \in (0, 1)$, entonces $\lambda = \lambda_1 + \lambda_2$.

A continuación presentamos el modelo AR(1) de Joe.

Definición 6.7 El modelo estacionario AR(1) construido por Joe (1996), y puesto en el contexto de ED por Jørgensen and Song (1998), es

$$X_n = A_n(X_{n-1}; \alpha) + \epsilon_n, \quad n = 1, 2, \dots, \quad (6.8)$$

donde el operador aleatorio $A(\cdot; \alpha)$ es independiente de la variable de innovación ϵ_n y está definido como en (6.7), el subíndice n (en $A_n(\cdot; \alpha)$) enfatiza que el adelgazamiento se realiza de manera independiente para cada n . Las innovaciones $\{\epsilon_n\}$ son v.a.'s i.i.d. con distribución $ED^*(\theta, (1 - \alpha)\lambda)$. Entonces la distribución marginal de X_n es $ED^*(\theta, \lambda)$ (siempre que $X_0 \sim ED^*(\theta, \lambda)$). Por lo tanto (6.8) constituye un proceso estacionario con marginal $ED^*(\theta, \lambda)$.

6.2.1. Ejemplos

Revisemos desde esta perspectiva parte del trabajo hecho en el capítulo anterior. Consideremos las distribuciones Gamma, Binomial Negativa, Poisson y Binomial, y echemos un vistazo a la Normal,

de acuerdo a sus correspondientes modelos de dispersión exponencial.

Ejemplo 6.8 Sean Z_1 y Z_2 , dos v.a.'s independientes tales que

$$Z_i \sim ED^*(\theta, \lambda_i) = \text{Normal}(\theta\lambda_i, \lambda_i), \quad i = 1, 2.$$

La función de densidad de Z_1 dado $Z = Z_1 + Z_2 = x$, como hemos visto es

$$f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) = \frac{c(w; \lambda_1) c(x-w; \lambda_2)}{c(x; \lambda)}.$$

En este caso, del Ejemplo 6.1, tenemos

$$\begin{aligned} f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) &= \frac{(2\pi\lambda_1)^{-1/2} \exp\{-w^2/2\lambda_1\} (2\pi\lambda_2)^{-1/2} \exp\{-(x-w)^2/2\lambda_2\}}{(2\pi\lambda)^{-1/2} \exp\{-x^2/2\lambda\}} \\ &= \frac{\lambda^{1/2}}{\sqrt{2\pi\lambda_1\lambda_2}} \exp\{-(w-\alpha x)^2/2\alpha(1-\alpha)\lambda\}, \end{aligned}$$

con $\alpha = \lambda_1/\lambda$. Es decir, se tiene que

$$Z_1 | Z_1 + Z_2 = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x) = \text{Normal}(\alpha x, \alpha(1-\alpha)\lambda). \quad (6.9)$$

Por lo tanto el proceso AR(1) (6.8), con distribución marginal infinitamente-divisible $ED^*(\theta, \lambda) = \text{Normal}(\theta\lambda, \lambda)$ puede ser representado mediante la expresión

$$X_n = \alpha X_{n-1} + \omega_n + \epsilon_n, \quad (6.10)$$

donde $\{\omega_n\}$ es una sucesión de v.a.'s i.i.d. con distribución $\text{Normal}(0, \alpha\bar{\alpha}\lambda)$, y $\{\epsilon_n\}$ es una sucesión de v.a.'s i.i.d. con distribución $ED^*(\theta, (1-\alpha)\lambda) = \text{Normal}(\theta\bar{\alpha}\lambda, \bar{\alpha}\lambda)$.

Pues, observe que

$$\alpha X_{n-1} + \omega_n | X_{n-1} = x \sim \text{Normal}(\alpha x, \alpha\bar{\alpha}\lambda),$$

que es la distribución que obtuvimos en (6.9).

Además

$$\begin{aligned} \alpha X_{n-1} + \omega_n &\sim \text{Normal}(\alpha\theta\lambda, \alpha^2\lambda) + \text{Normal}(0, \alpha\bar{\alpha}\lambda) \\ &= \text{Normal}(\theta\alpha\lambda, \alpha\lambda) = ED^*(\theta, \alpha\lambda). \end{aligned}$$

Y por último,

$$\alpha X_{n-1} + \omega_n + \epsilon_n \sim \text{Normal}(\theta\alpha\lambda, \alpha\lambda) + \text{Normal}(\theta\bar{\alpha}\lambda, \bar{\alpha}\lambda) = \text{Normal}(\theta\lambda, \lambda),$$

que es la distribución deseada para el modelo.

Por lo tanto la representación (6.10) es válida. \square

Ejemplo 6.9 Sean Z_1 y Z_2 , dos v.a.'s independientes con distribución

$$Z_i \sim ED^*(\theta, \lambda_i) = \text{Gamma}(\lambda_i, -\theta), \quad \theta < 0; \quad i = 1, 2.$$

Por el Ejemplo 6.2 tenemos que la función de densidad de Z_1 dado $Z_1 + Z_2 = x$, es

$$\begin{aligned} f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) &= \frac{w^{\lambda_1-1}}{\Gamma(\lambda_1)} \frac{(x-w)^{\lambda_2-1}}{\Gamma(\lambda_2)} \Big/ \frac{x^{\lambda_1+\lambda_2-1}}{\Gamma(\lambda_1+\lambda_2)} \\ &= \frac{\Gamma(\lambda_1+\lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} \frac{w^{\lambda_1-1}(x-w)^{\lambda_2-1}}{x^{\lambda_1+\lambda_2-1}} \\ &= \frac{\Gamma(\lambda_1+\lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} \left(1 - \frac{w}{x}\right)^{\lambda_2-1} \left(\frac{w}{x}\right)^{\lambda_1-1} \frac{1}{x}, \quad w \in (0, x). \end{aligned}$$

Esta última es la función de densidad de una v.a. W tal que $W = xA$, con x una constante positiva y A una v.a. Beta(λ_1, λ_2). El lector puede usar el Teorema 2.23 para corroborar esta afirmación, con $W = \varphi(A) = xA$.

Así que,

$$Z_1 | Z_1 + Z_2 = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x) = x\text{Beta}(\alpha\lambda, \bar{\alpha}\lambda).$$

Por lo tanto el modelos AR(1) (6.8), con distribución marginal infinitamente-divisible $ED^*(\theta, \lambda) = \text{Gamma}(\lambda, -\theta)$, $\theta < 0$, se expresa en la forma

$$X_n = A_n X_{n-1} + \epsilon_n, \quad (6.11)$$

donde $\{A_n\}$ es una sucesión de v.a.'s i.i.d. Beta($\alpha\lambda, \bar{\alpha}\lambda$), en tanto que $\{\epsilon_n\}$ es una sucesión de v.a.'s i.i.d. $ED^*(\theta, (1-\alpha)\lambda) = \text{Gamma}(\bar{\alpha}\lambda, -\theta)$.

Este proceso AR(1) con coeficiente aleatorio y distribución marginal Gamma es en esencia completamente distinto al modelo GAR(1) que estudiamos en el Capítulo 3. Fue introducido por Lewis (1983), y profundizado su estudio en Lewis *et al.* (1989), llamándole: Beta-Gamma AR(1), teniendo como punto de partida para su creación el hecho de que el producto de una v.a. Beta($\alpha\lambda, \lambda - \alpha\lambda$) con una v.a. Gamma($\lambda, -\theta$), independientes, es una v.a. Gamma($\alpha\lambda, -\theta$). Resultado que en estos momentos lo obtenemos directamente de la Proposición 6.6, pues según ésta

$$A_n X_{n-1} \sim ED^*(\theta, \alpha\lambda) = \text{Gamma}(\alpha\lambda, -\theta). \quad \square$$

McKenzie (1986, pg.690) propuso el análogo discreto a (6.11), esto es, un proceso AR(1) de coeficiente aleatorio y distribución marginal Binomial Negativa (BN). El cual está definido de la siguiente manera

$$X_n = A_n \circ X_{n-1} + \epsilon_n, \quad (6.12)$$

con $\{A_n\}$ una sucesión de v.a.'s i.i.d. Beta($\alpha\lambda, \lambda - \alpha\lambda$),

$A_n \circ X_{n-1} = \sum_{i=1}^{X_{n-1}} B_i(A_n)$ donde $\{B_i(A_n)\}$ es una sucesión de v.a.'s i.i.d. Bernoulli de media A_n e independientes de X_{n-1} para cada n , y $\{\epsilon_n\}$ una sucesión de v.a.'s i.i.d. Binomial

Negativa($\lambda - \alpha\lambda, p$).

Para el proceso definido en (6.12) McKenzie (1986) encuentra que si A es una v.a. Beta(λ_1, λ_2), entonces $A \circ X$ dado $X = x$ se distribuye Beta-Binomial(λ_1, λ_2, x). Así, si X se distribuye Binomial Negativa($\lambda_1 + \lambda_2, p$), $A \circ X$ se distribuye Binomial Negativa(λ_1, p). Por lo tanto, $\{X_n\}$ es un proceso estacionario con distribución marginal Binomial Negativa($\lambda_1 + \lambda_2, p$).

Revisemos ahora la distribución Binomial Negativa desde la perspectiva de modelos de dispersión exponencial y veamos como este enfoque nos lleva al proceso (6.12).

Ejemplo 6.10 Sean Z_1 y Z_2 , dos v.a.'s independientes tales que

$$Z_i \sim ED^*(\theta, \lambda_i) = \text{BN}(\lambda_i, 1 - e^\theta), \quad \theta < 0; \quad i = 1, 2.$$

Usando el Ejemplo 6.3, la función de densidad de Z_1 dado $Z_1 + Z_2 = x$, es

$$\begin{aligned} f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) &= \frac{\Gamma(\lambda_1 + w)}{\Gamma(\lambda_1)w!} \frac{\Gamma(\lambda_2 + x - w)}{\Gamma(\lambda_2)(x - w)!} \bigg/ \frac{\Gamma(\lambda_1 + \lambda_2 + x)}{\Gamma(\lambda_1 + \lambda_2)x!} \\ &= \frac{\Gamma(\lambda_1 + \lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)\Gamma(\lambda_1 + \lambda_2 + x)} \binom{x}{w} \Gamma(\lambda_1 + w)\Gamma(\lambda_2 + x - w), \end{aligned}$$

que es una distribución Beta-Binomial(λ_1, λ_2, x). Es decir, tenemos que

$$Z_1 | Z_1 + Z_2 = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x) = \text{Beta-Binomial}(\alpha\lambda, \bar{\alpha}\lambda, x).$$

Por lo tanto el proceso AR(1) (6.8), con distribución marginal infinitamente-divisible $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$, $\theta < 0$, corresponde al proceso (6.12), con $p = 1 - e^\theta$.

Por último note que al igual que en el ejercicio anterior la Proposición 6.6, nos permite obtener directamente

$$A_n \circ X_{n-1} \sim ED^*(\theta, \alpha\lambda) = \text{BN}(\alpha\lambda, 1 - e^\theta)$$

y

$$\epsilon_n \sim ED^*(\theta, (1 - \alpha)\lambda) = \text{BN}((1 - \alpha)\lambda, 1 - e^\theta). \quad \square$$

Ejemplo 6.11 Sean Z_1 y Z_2 , dos v.a.'s independientes con distribución

$$Z_i \sim ED^*(\theta, \lambda_i) = \text{Poisson}(\lambda_i e^\theta), \quad i = 1, 2.$$

La densidad de Z_1 dado $Z_1 + Z_2$ es, del Ejemplo 6.4,

$$\begin{aligned} f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) &= \frac{\lambda_1^w}{w!} \frac{\lambda_2^{x-w}}{(x-w)!} \bigg/ \frac{(\lambda_1 + \lambda_2)^x}{x!} \\ &= \binom{x}{w} \alpha^w (1 - \alpha)^{x-w} \end{aligned}$$

con $\alpha = \lambda_1/(\lambda_1 + \lambda_2)$. Es decir, tenemos que

$$Z_1 \mid Z_1 + Z_2 = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x) = \text{Binomial}(x, \alpha).$$

Por lo tanto el proceso AR(1) (6.8), con distribución marginal infinitamente-divisible $ED^*(\theta, \lambda) = \text{Poisson}(\lambda e^\theta)$ no es otro sino nuestro conocidísimo modelo INAR(1) con distribución marginal Poisson,

$$X_n = \alpha \circ X_{n-1} + \epsilon_n,$$

con $\alpha \circ X = Y_1 + Y_2 + \dots + Y_X$, donde $\{Y_i\}$ es una sucesión de v.a.'s i.i.d. Bernoulli de media α , e independientes de X , y $\{\epsilon_n\}$ una sucesión de v.a.'s i.i.d.. De la Proposición 6.6

$$\alpha \circ X_{n-1} \sim ED^*(\theta, \alpha\lambda) = \text{Poisson}(\alpha\lambda e^\theta)$$

y

$$\epsilon_n \sim ED^*(\theta, \bar{\alpha}\lambda) = \text{Poisson}(\bar{\alpha}\lambda e^\theta). \quad \square$$

Si nos olvidamos de la condición infinitamente-divisible para la distribución de X en (6.7), podemos entonces considerar la distribución Binomial.

Ejemplo 6.12 Sean Z_1 y Z_2 , dos v.a.'s independientes con distribución

$$Z_i \sim ED^*(\theta, \lambda_i) = \text{Binomial}(\lambda_i, e^\theta/(1 + e^\theta)), \quad \lambda_i \in \mathbb{N}; \quad i = 1, 2.$$

La densidad de Z_1 dado $Z_1 + Z_2$ es, del Ejemplo 6.5

$$f_{Z_1|Z}(w|x; \lambda_1, \lambda_2) = \frac{\binom{\lambda_1}{w} \binom{\lambda_2}{x-w}}{\binom{\lambda_1 + \lambda_2}{x}}. \quad (6.13)$$

Entonces, si restringimos α a la forma k/λ , $k \in \mathbb{N}$ (y $k < \lambda$), nos queda

$$Z_1 \mid Z_1 + Z_2 = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x) = \text{Hipergeométrica}(\alpha\lambda, \bar{\alpha}\lambda, x).$$

Por lo tanto el proceso AR(1) (6.8), con distribución marginal **no** infinitamente-divisible $ED^*(\theta, \lambda) = \text{Binomial}(\lambda, e^\theta/(1 + e^\theta))$ no es otro sino el modelo BAR(1), (5.46), visto en el capítulo anterior

$$X_n = S(X_{n-1}) + \epsilon_n,$$

donde $S(X_{n-1}) \mid X_{n-1} = x$ se distribuye como en (6.13), y por la Proposición 6.6

$$S(X_{n-1}) \sim ED^*(\theta, \alpha\lambda) = \text{Binomial}(\alpha\lambda, e^\theta/(1 + e^\theta))$$

y

$$\epsilon_n \sim ED^*(\theta, \bar{\alpha}\lambda) = \text{Binomial}(\bar{\alpha}\lambda, e^\theta/(1 + e^\theta)).$$

En relación a como se escribió el proceso BAR(1), en la sección 5.2.2, se tiene que $\lambda = N$ y $k = M$. Por último, es importante notar que el conjunto índice Λ para la distribución no infinitamente-divisible $ED^*(\theta, \lambda) = \text{Binomial}(\lambda, e^\theta/(1 + e^\theta))$, es $\Lambda = \{1, 2, \dots\}$. Es por ello que se restringe a α , a ser un múltiplo de λ^{-1} , de lo contrario $\alpha\lambda$ y $\bar{\alpha}\lambda$ podrían no pertenecer a Λ , lo cual es una hipótesis de la Proposición 6.6. \square

6.2.2. Propiedades generales

Proposición 6.13 Si la distribución $ED^*(\theta, \lambda)$ para X_n en (6.8) tiene momento de segundo orden, entonces su autocorrelación de orden k es α^k .

Demostración: Sean Z_1, Z_2 , dos v.a.'s como en (6.5) y sea $Z = Z_1 + Z_2$.

Por (6.3), $E[Z_1] = \lambda_1 k'(\theta) = \alpha \lambda k'(\theta) = \alpha E[Z]$.

Entonces

$$E[Z_1] = E_Z \left[E_{Z_1}[Z_1|Z] \right] = \alpha E_Z[Z] = E_Z[\alpha Z],$$

por lo tanto

$$E_{Z_1}[Z_1|Z] = \alpha Z,$$

el subíndice en las esperanzas enfatiza respecto a quién se calculan.

Se tiene entonces que

$$E[A(X; \alpha) | X = x] = \alpha x,$$

si X posee primer momento.

De donde, la esperanza condicional del proceso (6.8) es

$$E[X_n | X_{n-1} = x] = \alpha x + (1 - \alpha)\mu,$$

entonces aplicando el Corolario 2.17, su correlación de orden k es

$$\rho(k) = \alpha^k, \quad k = 0, 1, \dots \quad (6.14)$$

si el proceso tiene momento de segundo orden. \square

Sean Z_1, Z_2 , dos v.a.'s como en (6.5). Denotamos la distribución condicional de $Z = Z_1 + Z_2$ dado $Z_1 = w$, por $B(\theta, \lambda_2, w)$, cuya función de densidad es

$$\begin{aligned} f_{Z|Z_1}(x|w; \theta, \lambda_2) &= \frac{f_{Z_1}(w; \theta, \lambda_1) f_{Z_2}(x - w; \theta, \lambda_2)}{f_{Z_1}(w; \theta, \lambda_1)} \\ &= f_{Z_2}(x - w; \theta, \lambda_2) \\ &= c(x - w; \lambda_2) \exp\{\theta(x - w) - \lambda_2 k(\theta)\}. \end{aligned} \quad (6.15)$$

Así entonces, escribiendo (6.7) en términos de la notación (6.8) tenemos que

$$A_n(X_{n-1}; \alpha) | X_{n-1} = x \sim G(\alpha\lambda, \bar{\alpha}\lambda, x)$$

y

$$X_n | A_n(X_{n-1}; \alpha) = w \sim B(\theta, \bar{\alpha}\lambda, w),$$

con funciones de densidad (6.6) y (6.15) respectivamente.

El proceso (6.8) es de Markov (de orden uno) por construcción. Más aún, sus probabilidades de transición pueden ser calculadas como

$$\begin{aligned}
p(x_n, x_{n+1}) &= \mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n] \\
&= \int_w f(x_{n+1}, x_n, w) d\eta / f(x_n) \\
&= \int_w f(x_{n+1} \mid w, x_n) f(w \mid x_n) f(x_n) d\eta / f(x_n) \\
&= \int_w f_{Z|Z_1}(x_{n+1}|w) f_{Z_1|Z}(w|x_n) d\eta \\
&= \frac{\exp\{\theta x_{n+1} - \lambda_2 k(\theta)\}}{c(x_n; \lambda)} \int_w c(w; \lambda_1) c(x_n - w; \lambda_2) c(x_{n+1} - w; \lambda_2) e^{-\theta w} d\eta, \tag{6.16}
\end{aligned}$$

donde $f(x_{n+1} \mid w, x_n) = f(x_{n+1} \mid w)$ se justifica por la independencia condicional de X_{n+1} y X_n dado $A_{n+1}(X_n; \alpha) = w$.

Si el modelo $ED^*(\theta, \lambda)$ tiene soporte en los enteros no negativos entonces, la integral se reemplaza por $\sum_{w=0}^{x_n \wedge x_{n+1}}$.

Proposición 6.14 El proceso AR(1) descrito por (6.8), es reversible en el tiempo.

Demostración: De los cálculos anteriores la densidad conjunta, h , de (X_n, X_{n+1}) es

$$\begin{aligned}
h(x_n, x_{n+1}) &= \exp\{\theta(x_n + x_{n+1}) - (\lambda_2 + \lambda)k(\theta)\} \\
&\quad \times \int_w c(w; \lambda_1) c(x_n - w; \lambda_2) c(x_{n+1} - w; \lambda_2) e^{-\theta w} d\eta,
\end{aligned}$$

puesto que esta última es simétrica en x_n y x_{n+1} , el proceso de Markov (6.8) es reversible en el tiempo. \square

Capítulo 7

Aplicación

El presente capítulo pretende ilustrar con un ejemplo la utilidad e importancia de la teoría desarrollada en capítulos anteriores.

Para esto usaremos los casos mensuales de poliomielitis reportados por los Centros para el Control y Prevención de Enfermedades -CDC- de los Estados Unidos entre los años 1970-1983. Estos datos fueron utilizados por vez primera dentro del contexto estadístico por Zeger (1988), para ilustrar la aplicación de un modelo de regresión: *parameter-driven model*.

Encontramos en los datos una dependencia de Markov de orden uno. Ajustamos un modelo con distribución $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$ visto en el capítulo anterior. Los datos poseen una observación que se encuentra fuera del rango que prevalece en la serie. Nos preguntamos sobre la influencia de este dato en el ajuste obtenido, por lo que realizamos ajustes con y sin él. Así mismo, diferenciamos los casos, α , el factor que pesa la dependencia de Markov, variable y α constante ($=\hat{\rho}(1)$). Realizamos un ajuste más excluyendo un total de dos observaciones obteniendo mejores resultados. El método que utilizamos para la estimación de los parámetros es el de máxima verosimilitud (MV).

7.1. Poliomielitis en U.S. de 1970 a 1983

El Cuadro 7.9 (al final del capítulo) contiene los casos mensuales de poliomielitis en Estados Unidos reportados por los Centros para el Control y Prevención de Enfermedades -CDC- de ese país entre los años de 1970 a 1983 y publicados en el *Morbidity and Mortality Weekly Report Annual Summary* (1970-1983).

La Figura 7.1 muestra la gráfica de los datos.

Llama particularmente la atención la observación del mes de Noviembre de 1972 por su magnitud respecto al resto de las observaciones. Este caso será de suma importancia en nuestro análisis.

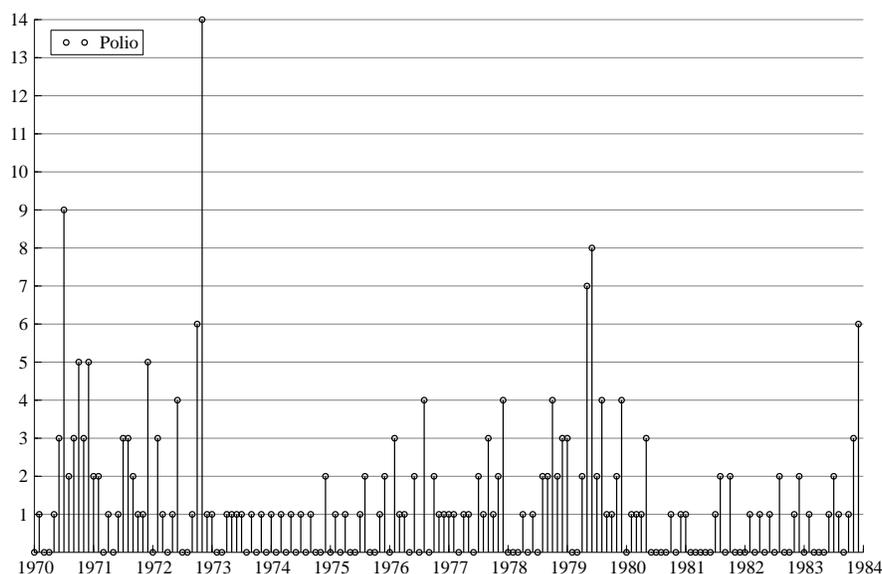


Figura 7.1: Número de casos mensuales de poliomielitis en U.S. entre los años 1970-1983.

	T	Media	Mediana	Moda	Varianza	Suma
Polio	168	1.333333	1	0	3.504990	224

Cuadro 7.1: Principales estadísticos de los datos.

Un análisis de los datos se presenta a continuación.

7.1.1. Distribución

Es claro que los datos provienen de una distribución discreta. A lo largo del trabajo hemos tratado con tres distribuciones: Poisson, Binomial y Binomial Negativa, por lo que intentaremos ajustar una de estas tres a los datos.

El histograma de los datos se encuentra en la Figura 7.2. Como podemos observar los datos provienen de una distribución asimétrica. Además, (del Cuadro 7.1) la varianza de los datos es mayor a la media, por lo que los datos podrían provenir de una distribución Binomial Negativa(λ, p). Realizamos una prueba de bondad de ajuste usando la prueba χ^2 con un nivel de significancia del 5%, e hipótesis nula: los datos provienen de una distribución Binomial Negativa($1, \hat{p}$) con $\hat{p} = 1/(\bar{X} + 1) = 0.42857$. Obteniendo un p-valor(= 0.17638) mayor a 0.05. Por lo tanto no rechazamos la hipótesis nula. Así, la distribución que decidimos ajustar a los datos es, en efecto, una distribución Binomial Negativa.

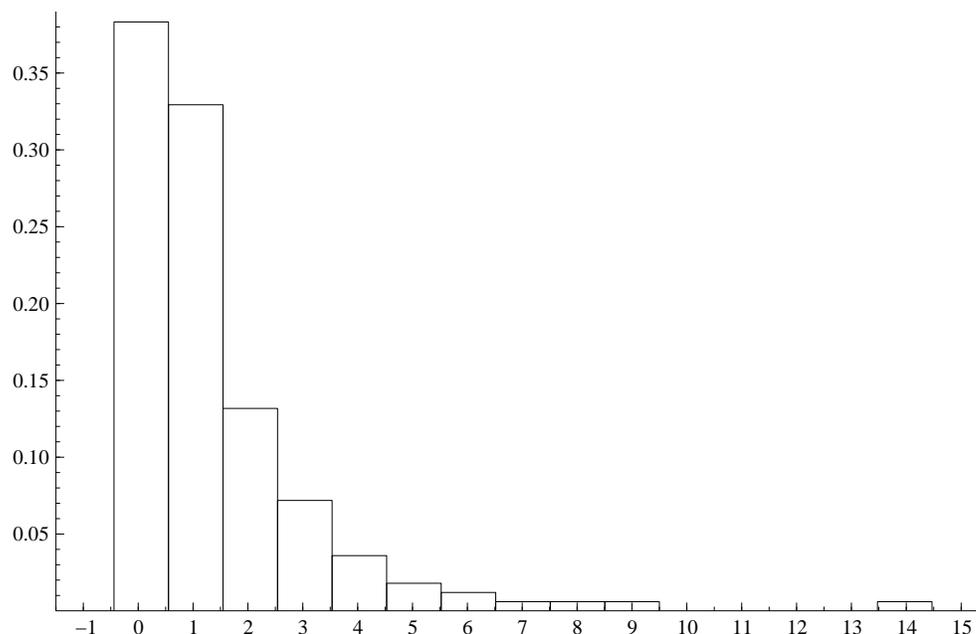


Figura 7.2: Histograma de los datos.

7.1.2. Modelo

Procedemos a inspeccionar las gráficas de la ac.f. y pac.f. muestrales de los datos (Figura 7.3).

Como podemos observar, tanto en la ac.f. como en la pac.f., el único rezago significativo es el primero. Por lo que podría tratarse de un modelo AR(1), un MA(1), o incluso un ARMA(1,1), discretos.

Decidimos ajustar un modelo AR(1) porque creemos que existe una dependencia entre las personas que tienen poliomieltis hoy y las que lo tuvieron ayer. Recordemos que una de las causas de poliomieltis es el contagio.

Tenemos al menos tres modelos AR(1) con distribución marginal Binomial Negativa:

1. Modelo INAR(1), (5.4), con innovación (5.22), McKenzie (1986).
2. El modelo descrito por (5.31), Al-Osh y Aly (1992).
3. El modelo AR(1) de Joe (1996) con distribución marginal $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$, ec. (6.12).

El método que usaremos para la estimación de los parámetros es el de MV. Puesto que consideramos una dependencia de Markov de orden 1, el cálculo de las probabilidades de transición, $p(x_n, x_{n+1})$, juega un papel primordial.

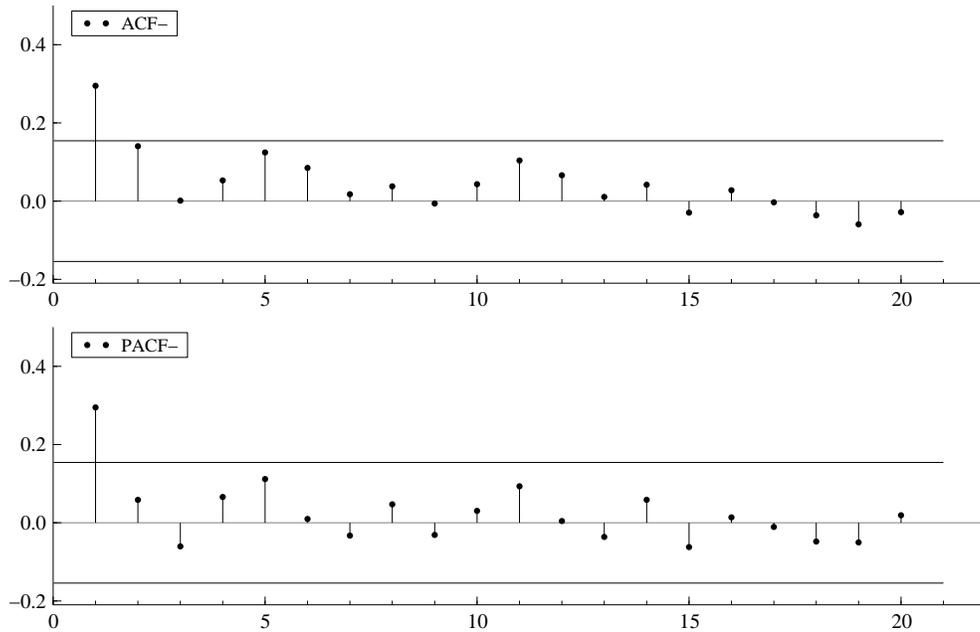


Figura 7.3: ac.f. y pac.f. de los datos de poliomiélitis del Cuadro 7.9.

El modelo 1. posee una innovación bastante compleja (con f.a.g.p. (5.21), que pertenece a una distribución Poisson compuesta de la forma (5.22) que dificulta el cálculo de (5.6).

Para el modelo 2. calculamos explícitamente $p(x_n, x_{n+1})$, ecuación (5.44). El ajuste de una distribución Binomial Negativa por medio de este modelo será el objeto de nuestro estudio en otro momento y lugar.

Escogemos el modelo 3. para ejemplificar (en un caso particular: $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$) la utilidad del trabajo hecho por Joe (1996), y puesto en el contexto de modelos de dispersión exponencial por Jørgensen y Song (1998). Podemos calcular $p(x_n, x_{n+1})$ vía la ec. (6.16) ó, incluso, la construcción del proceso es tan generosa que, para el caso discreto, podemos obtenerlas directamente de la definición (Definición 6.7),

$$\mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n]$$

$$= \sum_{w=0}^{x_n \wedge x_{n+1}} \mathbb{P}[A_{n+1}(X_n; \alpha) = w \mid X_n = x_n] \mathbb{P}[\epsilon_n = x_{n+1} - w],$$

donde

$$A_{n+1}(X_n; \alpha) \mid X_n = x_n \sim G(\alpha\lambda, \bar{\alpha}\lambda, x_n),$$

la contracción correspondiente a $ED^*(\theta, \lambda)$, y

$$\epsilon_n \sim ED^*(\theta, \bar{\alpha}\lambda). \quad \square$$

Realizamos un diagrama de dispersión entre los datos $\{x_1, x_2, \dots, x_{T-1}\}$ Vs. $\{x_2, x_3, \dots, x_T\}$. La Figura 7.4 muestra que la relación que guardan los dos conjuntos de datos sería mínimamente explicada por una línea recta. Esto es, existe una cuestión no lineal en los datos que esperamos que nuestro modelo pueda captar.

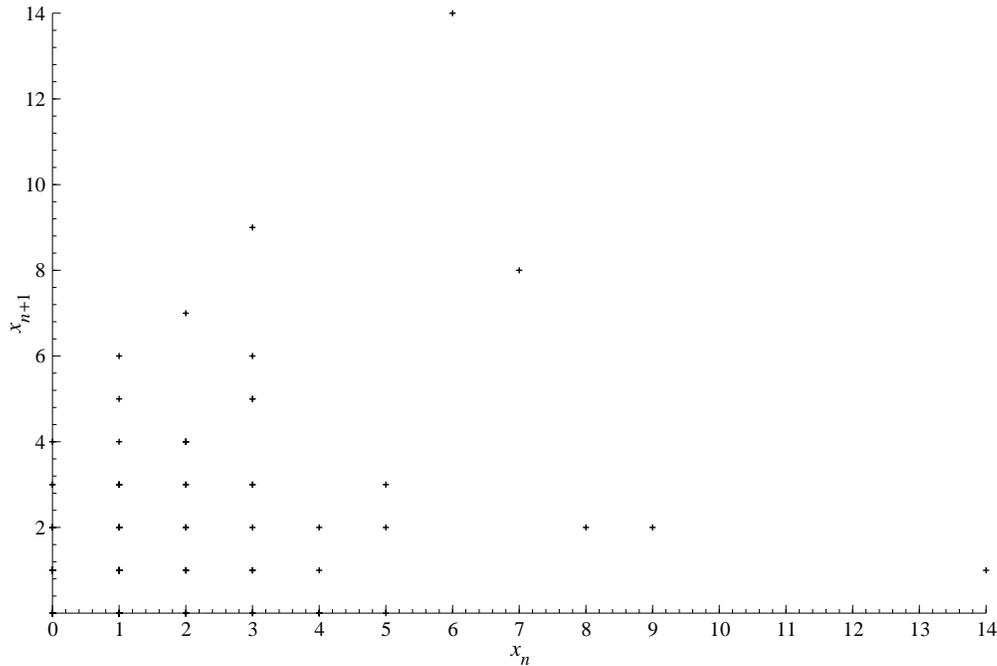


Figura 7.4: Diagrama de dispersión entre $\{x_1, x_2, \dots, x_{T-1}\}$ y $\{x_2, x_3, \dots, x_T\}$, para los datos de poliomielitis.

Resumiendo tenemos que para los datos de poliomielitis del Cuadro 7.9

- ajustaremos una distribución Binomial Negativa,
- creemos poseen una dependencia de Markov de orden 1,
- el modelo que usaremos para la estimación de los parámetros está descrito por (6.12).

7.2. Parámetros

Los detalles analíticos sobre la estimación de los parámetros presentados aquí, se encuentran en la subsección, 7.3. Estimación.

Si bien por la prueba de bondad de ajuste de la subsección 7.1.1, no rechazamos que los datos provinieran de una distribución Binomial Negativa con parámetros $(\hat{\lambda}, \hat{p}) = (1, 0.42857)$, no significa

que éstos sean los mejores. De hecho el estimador $\hat{p} = 1/(\bar{X} + 1)$ supone que los datos provienen de una muestra aleatoria. Sin embargo, el supuesto de independencia no es asumido por nosotros, sino todo lo contrario, como hemos dicho creemos que existe una dependencia de Markov en los datos. No obstante, $(1, 0.42857)$ serán nuestro punto de partida. En contexto con la notación del modelo (7.2), nuestros parámetros son λ, θ , y no λ, p , donde θ y p guardan la relación $1 - e^\theta = p$, por lo que nuestros puntos iniciales serán $(\hat{\lambda}, \hat{\theta}) = (1, -0.56)$.

Realizamos cuatro ajustes, usando el método de MV en todos.

Como podemos observar de la ec. (7.5) las probabilidades de transición, $p(x_n, x_{n+1})$, del proceso están determinadas por tres parámetros λ, θ y α . Así entonces, nuestros dos primeros ajustes difieren en la consideración siguiente:

Ajuste 1 Maximiza la log-verosimilitud, el logaritmo de la función descrita por (7.3), asumiendo que $\hat{\alpha} = \hat{\rho}(1) = 0.29480$.

Ajuste 2 Maximiza la log-verosimilitud tomando a α como una variable. Esto significa que maximizamos el logaritmo de (7.3) con respecto a λ, θ y α .

La log-verosimilitud pasa de -268.16 en el punto $(\hat{\lambda}, \hat{\theta}, \hat{\alpha}) = (1, -0.56, 0.29480)$ a -266.78 en su máximo $(1.4694, -0.70965, 0.29480)$ para el Ajuste 1. Mientras que para el Ajuste 2, la log-verosimilitud pasa de la misma cantidad (-268.16) en el punto inicial $(1, -0.56, 0.29480)$ a -264.61 en su máximo $(\hat{\lambda}, \hat{\theta}, \hat{\alpha}) = (1.3829, -0.71684, 0.14998)$. Resulta interesante observar que si bien $\hat{\lambda}, \hat{\theta}$ experimentan un cambio, la variación mayor se da en $\hat{\alpha}$, al disminuir a casi la mitad de su valor inicial. Así, la log-verosimilitud alcanza un valor mayor al hacerla una función de α también.

Comentario: Al decidir ajustar un modelo AR(1) (estacionario) a los datos, su correlación de orden uno debiera ser $\rho(1) = \alpha$. Sin embargo $\hat{\rho}(1)$ es sólo una estimación de $\rho(1)$ y por lo tanto de α . No significa que $\hat{\rho}(1) = \alpha$. Por lo tanto $\hat{\rho}(1) = 0.29480$ es una mera estimación de α como lo es $\hat{\alpha} = 0.14998$ en el Ajuste 2.

Como mencionamos al principio del capítulo la observación del mes de Noviembre de 1972 ($x_{35} = 14$) es bastante mayor que el resto de las observaciones. Los ajustes tres y cuatro son el Ajuste 1 y el Ajuste 2, respectivamente, con la exclusión de este dato:

Ajuste 3 La correspondiente log-verosimilitud pasa de -260.32 en el punto $(1, -0.56, 0.29480)$, a -257.86 en el máximo $(1.8146, -0.85402, 0.29480)$.

Ajuste 4 La log-verosimilitud pasa de -260.32 a, -255.34 en el máximo $(1.7075, -0.86255, 0.13867)$. Por lo que al igual que entre el Ajuste 1 y el Ajuste 2, la log-verosimilitud alcanza un valor mayor

Número de casos	frecuencia relativa de los datos	Ajuste 1	Ajuste 2
0	0.38095	0.36986	0.39592
1	0.32738	0.18190	0.19333
2	0.13095	0.08946	0.09440
3	0.07143	0.04400	0.04610
4	0.03571	0.02164	0.02251
5	0.01786	0.01064	0.01099
6	0.01190	0.00523	0.00537
7	0.00595	0.00257	0.00262
8	0.00595	0.00127	0.00128
9	0.00595	0.00062	0.00062
14	0.00595	0.00002	0.00002

Cuadro 7.2: Estimaciones del Ajuste 1 y Ajuste 2.

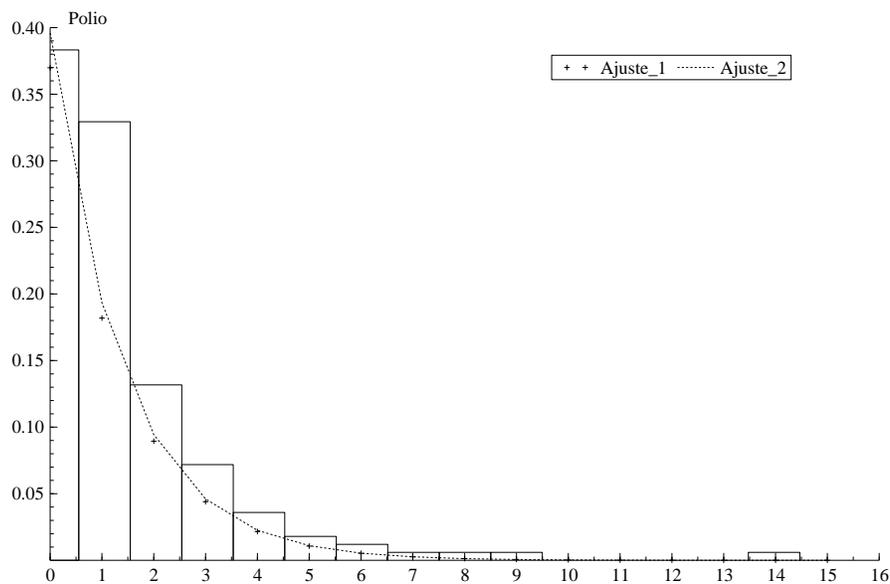
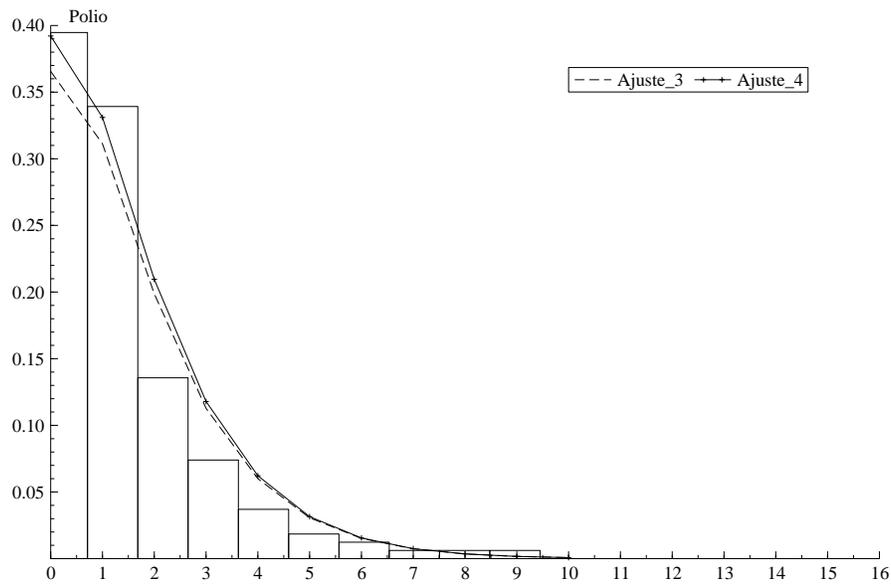


Figura 7.5: Ajustes 1 y 2. Consideran los 168 datos.

Número de casos	frecuencia relativa de los datos	Ajuste 3	Ajuste 4
0	0.38323	0.36554	0.39209
1	0.32934	0.31122	0.33099
2	0.13174	0.19873	0.20956
3	0.07186	0.11280	0.11794
4	0.03593	0.06002	0.06222
5	0.01796	0.03066	0.03152
6	0.01198	0.01523	0.01552
7	0.00599	0.00741	0.00749
8	0.00599	0.00355	0.00356
9	0.00599	0.00168	0.00167

Cuadro 7.3: Estimaciones del Ajuste 3 y Ajuste 4.

Figura 7.6: Ajustes 3 y 4. Excluyen el punto $x_{35} = 14$.

al hacerla una función de α también.

Observemos como $\hat{\alpha}$ en el último ajuste ha disminuido aún más. No obstante, $\hat{\alpha} = 0.13867$ es suficiente para afirmar que las observaciones de poliomielitis del Cuadro 7.9 no son independientes, sino que guardan una dependencia de Markov de orden 1, y que por lo tanto no estamos trabajando vanamente.

7.2.1. Discusión

Definimos el error absoluto del Ajuste i en el caso k ($\text{EAbs}_k^{(i)}$), como la distancia entre la frecuencia estimada por el ajuste correspondiente para k casos de polio y la frecuencia relativa (real) de los datos. Directamente de los Cuadros 7.2 y 7.3, obtenemos:

Número de casos k	$\text{EAbs}_k^{(1)}$	$\text{EAbs}_k^{(2)}$	$\text{EAbs}_k^{(3)}$	$\text{EAbs}_k^{(4)}$
0	0.01109	0.01497	0.01769	0.00886
1	0.14548	0.13405	0.01812	0.00165
2	0.04149	0.03655	0.06699	0.07782
3	0.02743	0.02533	0.04094	0.04608
4	0.01407	0.01321	0.02409	0.02630
5	0.00721	0.00687	0.01270	0.01355
6	0.00667	0.00654	0.00325	0.00354
7	0.00338	0.00333	0.00142	0.00150
8	0.00469	0.00467	0.00244	0.00243
9	0.00533	0.00533	0.00431	0.00432
14	0.00593	0.00594		
$\text{EAbs}_t^{(i)}$	0.27278	0.25678	0.19197	0.18606

Cuadro 7.4: Errores absolutos de cada ajuste.

donde $\text{EAbs}_t^{(i)}$, que denota el error absoluto total del Ajuste i , es igual a la suma de los $\text{EAbs}_k^{(i)}$, con $k = 0, 1, \dots$

Como se puede observar, la observación $x_{35} = 14$ es un punto influyente puesto que las características de los ajustes cambian al retirar esta observación.

Mientras los ajustes uno y dos tienen una tendencia a dar frecuencias por debajo de la frecuencia real de los datos (Cuadro 7.2), los ajustes tres y cuatro tienden a sobreestimar.

La mayor deficiencia por parte de los ajustes uno y dos, es su estimación para un caso de polio. De

hecho, resulta importante notar que más del 50% del error absoluto total de estos ajustes se debe, únicamente, a la pésima estimación para un caso de polio. El Ajuste 2 manifiesta que es necesario rectificar (aumentar) la estimación para este caso pero el dejar variar α , tomarla como variable, no es suficiente. Al retirar la observación $x_{35} = 14$, el Ajuste 3 sí puede dar una mejor estimación para un caso de polio (aún cuando α es constante).

El Ajuste 4 es el ajuste con menor error absoluto total. Su principal virtud son sus excelentes estimaciones para cero y un caso de polio. Sin embargo, esta misma ventaja es causante de que para los casos siguientes la distribución Binomial Negativa correspondiente no alcance a descender lo suficiente, propiciando así una severa sobreestimación para los casos dos, tres, cuatro y cinco de poliomielitis. (Algo similar ocurre con el Ajuste 3). De hecho, los errores absolutos para estos casos de polio por parte de los ajustes tres y cuatro, son casi el doble que los correspondientes de los ajustes uno y dos (Cuadro 7.4).

Calculamos la prueba χ^2 de Pearson de los cuatro ajustes, con un nivel de significancia del 5% e hipótesis nula: los datos se distribuyen Binomial Negativa $(\hat{\lambda}_i, \hat{\theta}_i)$, con $\hat{\lambda}_i, \hat{\theta}_i$ los obtenidos en el ajuste i .

Los estadísticos de prueba de cada uno de los ajustes son:

Ajuste	Estadístico de prueba
1	38.48326
2	33.91558
3	8.56623
4	10.04396

Cuadro 7.5: Estadísticos de prueba para la prueba χ^2 de Pearson de cada ajuste.

Estos estadísticos son comparados con el cuantil 0.95 de una distribución ji-cuadrada con $4(=6-2)$ grados de libertad, el cual es 9.488. Por lo tanto la hipótesis nula es rechazada para los ajustes uno, dos y cuatro, en tanto que para el Ajuste 3 no. En base a este criterio podríamos quedarnos con este último. Sin embargo, es importante subrayar que en nuestro ejemplo esta prueba no es un criterio definitivo sino simplemente un criterio de referencia, pues la prueba χ^2 de Pearson supone que los datos provienen de una muestra aleatoria, mientras que en nuestro caso la suposición de correlación serial distinta de cero es un supuesto básico en el desarrollo del problema.

Comentario: El estadístico de prueba del Ajuste 4 está sólo un poco por arriba del cuantil, siendo que la prueba es meramente una referencia y que su valor es mucho más chico que el de los dos primeros ajustes, el Ajuste 4 también podría ser adecuado.

De la gráfica de los datos podemos observar que existen otros tres casos cuya frecuencia es de uno. Éstos son: siete, ocho y nueve casos de polio. Explorando posibilidades, si retiramos el caso $x_7 = 9$ (y $x_{35} = 14$) obtenemos un modelo cuya estimación para cero casos de polio es excelente, para un casos de polio, sin llegar a ser tan atinada como la del Ajuste 4, permite que la distribución Binomial Negativa correspondiente pueda descender lo suficiente de manera tal que las estimaciones para el resto de los casos sean bastante buenas. Denotamos a este ajuste como el Ajuste 5, el cual maximiza la log-verosimilitud con respecto a λ , θ y α . Sus estimaciones se encuentran en el Cuadro 7.7.

La prueba χ^2 de Pearson de este ajuste es la siguiente,

Casos de de poliomieltis	frecuencia de los datos fr	frecuencia estimada fe	Estadístico _{j}
0	64	65	0.01538
1	55	49	0.73469
2	22	27	0.92593
3	12	14	0.28571
4	6	6	0
5	3	3	0.8
6	2	1	
7	1	1	
8	1	0	
Total	166		2.76171

Cuadro 7.6: Prueba χ^2 de Pearson del Ajuste 5.

Donde

$$\text{Estadístico}_j = \frac{(fr_j - fe_j)^2}{fe_j}$$

y

$$\text{Estadístico de prueba} = \sum_{j=1}^c \text{Estadístico}_j.$$

En este caso $c = 6$, pues como criterio de la prueba se debe evitar que haya clases con frecuencia menor a cinco observaciones. Por lo que los casos de poliomieltis, cinco, seis, siete y ocho, con frecuencias tres, dos, uno y uno respectivamente, se suman, formando una sola clase. (Los estadísticos

Número de casos	frecuencia (relativa) de los datos	Estimación del Ajuste 5	$E\text{Abs}_k^{(5)}$
0	0.38554	0.39147	0.00593
1	0.33133	0.29361	0.03772
2	0.13253	0.16516	0.03263
3	0.07229	0.08258	0.01026
4	0.03614	0.03871	0.00257
5	0.01807	0.01742	0.00065
6	0.01205	0.00762	0.00443
7	0.00602	0.00327	0.00276
8	0.00602	0.00138	0.00465
$E\text{Abs}_t^{(5)}$			0.10159

Cuadro 7.7: Estimaciones del Ajuste 5.

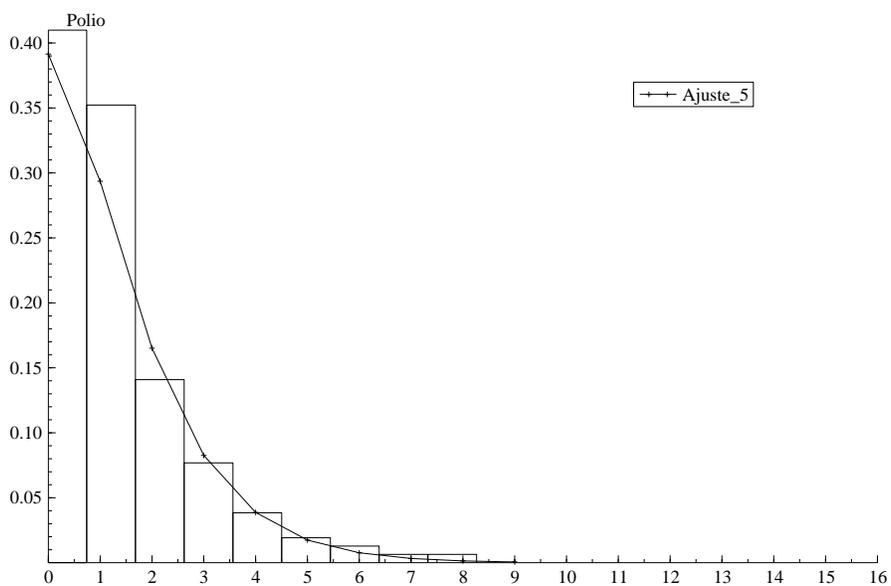


Figura 7.7: Ajuste 5. Ajuste óptimo.

de prueba del Cuadro 7.5 se obtuvieron de esta misma manera.)

El estadístico de prueba del Ajuste 5 es bastante menor en comparación al del Ajuste 3, y al del resto de los ajustes; además, sus estimaciones para cero, dos, tres, cuatro y cinco casos de polio son las más cercanas a las frecuencias (relativas) de los datos (Cuadro 7.7). Por lo que optamos por este ajuste como nuestro ajuste óptimo. Los parámetros estimados por el Ajuste 5 son:

λ	θ	α
1.9953	-0.98080	0.14055

Cuadro 7.8: Parámetros óptimos.

El lector podría preguntarse que hay acerca de un posible Ajuste 6 análogo al Ajuste 5 pero con α constante. Invitamos al lector a que modifique el código de la subsección 7.3.1 y corrobore que las estimaciones de ambos ajustes son muy parecidas pero el estadístico de prueba del Ajuste 6 es, ligeramente, mayor al correspondiente del Ajuste 5.

Conclusiones: En base a la discusión anterior concluimos que al ajustar a los datos de poliomiélitis del Cuadro 7.9 un modelo del tipo (7.1) con distribución marginal $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$,

- la log-verosimilitud, el logaritmo de la función descrita por (7.3), alcanza un máximo mayor al considerarla una función de λ , θ y α , y no sólo de λ y θ .
- Los dos primeros ajustes, Ajuste 1 y Ajuste 2, poseen un pésima estimación para un casos de polio. Situación que se corrige, sobresalientemente, al excluir el punto $x_{35} = 14$. Sin embargo esta corrección es causante de una fuerte sobreestimación en los casos subsiguientes.
- La prueba de bondad de ajuste χ^2 de Pearson (que supone independencia en los datos), de entre los ajustes uno, dos, tres y cuatro, opta por el Ajuste 3.
- Un mejor ajuste se obtiene al excluir las observaciones $x_{35} = 14$ y $x_7 = 9$.

7.3. Estimación

Considere el modelo AR(1) construido por Joe (1996)

$$X_n = A_n(X_{n-1}; \alpha) + \epsilon_n, \quad n = 1, 2, \dots \quad (7.1)$$

En el caso en que $\{X_n\}$ constituya un proceso estacionario con distribución marginal $ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$, hemos visto -Ejemplo 6.10- que éste toma la forma

$$X_n = A_n \circ X_{n-1} + \epsilon_n, \quad (7.2)$$

con $\{A_n\}$ una sucesión de v.a.'s i.i.d. Beta($\alpha\lambda, \bar{\alpha}\lambda$), $A_n \circ X_{n-1} = \sum_{i=1}^{X_{n-1}} B_i(A_n)$ donde $\{B_i(A_n)\}$ es una sucesión de v.a.'s i.i.d. Bernoulli de media A_n e independientes de X_{n-1} para cada n , y $\{\epsilon_n\}$ una sucesión de v.a.'s i.i.d. Binomial Negativa($\bar{\alpha}\lambda, 1 - e^\theta$), con $\bar{\alpha} = 1 - \alpha$ como siempre.

Al existir una dependencia de Markov de orden 1, la función de verosimilitud $L_{\mathbf{X}}(\lambda, \theta)$ para un conjunto de datos $\mathbf{x} = \{x_1, x_2, \dots, x_T\}$, es

$$L_{\mathbf{X}}(\lambda, \theta) = f_X(x_1; \theta, \lambda) \prod_{n=1}^{T-1} p(x_n, x_{n+1}), \quad (7.3)$$

donde $p(x_n, x_{n+1})$ denota las probabilidades de transición, $\mathbb{P}[X_{n+1} = x_{n+1} \mid X_n = x_n]$, de la cadena.

Al considerar un modelo de la forma (7.1), las $p(x_n, x_{n+1})$ pueden ser obtenidas fácilmente mediante la formula dada por la ec. (6.16).

En este caso

$$p(x_n, x_{n+1}) = \frac{x_n! \Gamma(\lambda) e^{\theta x_{n+1}} (1 - e^\theta)^{\lambda_2}}{\Gamma(\lambda + x_n)} \\ \times \sum_{w=0}^{x_n \wedge x_{n+1}} \frac{\Gamma(\lambda_1 + w) \Gamma(\lambda_2 + x_n - w) \Gamma(\lambda_2 + x_{n+1} - w)}{\Gamma(\lambda_1) w! \Gamma(\lambda_2) (x_n - w)! \Gamma(\lambda_2) (x_{n+1} - w)!} e^{-\theta w},$$

pues si $X \sim ED^*(\theta, \lambda) = \text{Binomial Negativa}(\lambda, 1 - e^\theta)$, entonces $c(x; \lambda) = \Gamma(\lambda + x)/\Gamma(\lambda)x!$ y $k(\theta) = -\ln(1 - e^\theta)$, Ejemplo 6.3. Además $\lambda_1 = \alpha\lambda$, $\lambda_2 = \bar{\alpha}\lambda$, con $\alpha \in (0, 1)$, y entonces $\lambda = \lambda_1 + \lambda_2$.

Realizando estas sustituciones nos queda que

$$p(x_n, x_{n+1}) = \frac{x_n! \Gamma(\lambda) e^{\theta x_{n+1}} (1 - e^\theta)^{\bar{\alpha}\lambda}}{\{\Gamma(\bar{\alpha}\lambda)\}^2 \Gamma(\alpha\lambda) \Gamma(\lambda + x_n)} \quad (7.4)$$

$$\times \sum_{w=0}^{x_n \wedge x_{n+1}} \frac{e^{-\theta w} \Gamma(\bar{\alpha}\lambda + x_{n+1} - w) \Gamma(\bar{\alpha}\lambda + x_n - w) \Gamma(\alpha\lambda + w)}{w! (x_{n+1} - w)! (x_n - w)!}. \quad (7.5)$$

El método numérico que decidimos usar para maximizar la log-verosimilitud es el Broyden-Fletcher-Goldfarb-Shanno (BFGS). Este método se encuentra implementado en el software `0x` mediante la función `MaxBFGS`, que requiere la librería `maximize`. Los detalles de su procedimiento se encuentran fuera de nuestros objetivos. Información sobre el algoritmo puede obtenerse en Press *et al.* (1992).

7.3.1. Código

```
1 /*****
2 /***** Ajuste_2 *****/
```

```

3  /***** Librerías requeridas *****/
4  #include <oxstd.h>
5  #import<maximize>
6  #include <oxdraw.h>
7  /***** Variables globales *****/
8  static decl Glb_X;
9  /***** Sumandos *****/
10 Numxyz(const nx, const z, const vP, const ny)
11 {
12 return
13 ((exp(-vP[1]*ny))./factorial(ny)).*
14 ((gammafact((1-vP[2])*vP[0]+nx-ny)).*
15 gammafact((1-vP[2])*vP[0]+z-ny)).*
16 gammafact(vP[2]*vP[0]+ny))./
17 (factorial(nx-ny).*factorial(z-ny)));
18 }
19 /***** Calcula p(x_{i},x_{i+1}) *****/
20 pt(const vp, const x)
21 {
22 decl w, i, T, ptrant;
23 T=rows(x); //tamaño de la muestra
24 ptrant=zeros(1,T);
25 for (i=0; i<T-1; ++i)
26 {
27 w=range(0,min(x[i+1],x[i]));
28 ptrant[i]=((factorial(x[i])*gammafact(vp[0])*
29 exp(vp[1]*x[i+1])*((1-exp(vp[1]))^((1-vp[2])*vp[0])))/
30 ((gammafact((1-vp[2])*vp[0]))^(2))*
31 gammafact(vp[2]*vp[0])*gammafact(vp[0]+x[i])))*)
32 sumr(Numxyz(x[i+1],x[i],vp,w));
33 }
34 ptrant[T-1]=(1-exp(vp[1]))^vp[0];
35 return ptrant;
36 }
37 /***** Log-verosimilitud *****/
38 LogVS(const VP, const serie)
39 {

```

```

40 return sumr(log(pt(VP,serie)));
41 }
42 LogVSapp(const Vp,const adFunc,const avScore,const amHess)
43 {
44 adFunc[0]=LogVS(Vp,Glb_X);
45 return 1;
46 }
47 /***** Estimador Máximo Verosímil *****/
48 MV(const vp_ini, const Pe, const serie)
49 {
50 decl EVecP, dfunc;
51 Glb_X=serie;
52 EVecP=vp_ini;
53 MaxControl(-1,100);
54 MaxControlEps(1e-5,1e-5);
55 MaxBFGS(LogVSapp, &EVecP, &dfunc, 0, 1);
56 Pe[0]=EVecP;
57 return 1;
58 }
59 /***** función de densidad *****/
60 /** Binomial Negativa(\lambda, p), con  $p=1-\exp(-\theta)$  */
61 fX(const x, const Vlt)
62 {
63 return
64 binomial(Vlt[0]+x-1,Vlt[0]-1).*
65 ((exp(Vlt[1]))^x).*((1-exp(Vlt[1]))^Vlt[0]);
66 }
67 /***** Programa principal *****/
68 main()
69 {
70 decl s_polio=loadmat("DatosPolio.xls");
71 decl emv, r;
72 MV(<1;-0.56;0.29480>, &emv, s_polio);
73 r=range(0,15);
74 DrawAdjust(ADJ_AREA_X,0,0,16);
75 DrawDensity(0,s_polio', "Polio",0,1,0,0,0,15,1);
76 DrawXMatrix(0,fX(r,emv),{"Ajuste_2"},r,"",2,1);

```

```

77 SaveDrawWindow("Ajuste_2.gwg");
78 println("BN(lambda,1-exp(theta))=",fX(r,emv));
79 }

```

El código anterior pertenece al Ajuste 2 de la sección 7.2. A continuación una breve descripción del mismo:

La función `Numxyz` situada en la línea 10, calcula los sumandos de la suma que se encuentra en (7.5), con `nx` jugando el papel de x_{n+1} , `z` el de x_n y `ny` el de w ; `vP` denotará el vector de parámetros en el orden $(\lambda, \theta, \alpha) = (\text{vP}[0], \text{vP}[1], \text{vP}[2])$. Observe que `0x` indexa comenzando en cero, no en uno.

La función `pt` en la línea 20 efectúa el cálculo completo de (7.4) y (7.5). `ptrant[i]` (líneas 28 a 32) calcula $p(x_i, x_{i+1})$. Por ejemplo, la línea 32 realiza la suma que está en (7.5). Para esto se usó el hecho de que la función `Numxyz(...)` admite forma matricial para evaluarla sobre un rango dado (`w=range(0,min(x[i+1],x[i]))`). En el último elemento del vector `ptrant` guardamos el término $f_X(x_1; \theta, \lambda)$ de la ec. (7.3), línea 34. Que en nuestro caso es $(1 - e^\theta)^\lambda$, puesto que $x_1 = 0$.

La función `LogVS` recibe dos argumentos, `VP` y `serie`, y calcula la log-verosimilitud, esto es, el logaritmo de (7.3). La función `LogVSapp` es una función apropiada declarada por necesidad. Recibe un total de cuatro argumentos, las variables a estimar, `Vp` (λ, θ y α), `adFunc`, `avScore` y `amHess`. De entrada `adFunc` recibe una dirección y en la salida da la función evaluada en el punto `Vp` (línea 44 y 45). “Score” y el “Hessiano”, `avScore` y `amHess`, no se han especificado analíticamente, por lo que `MaxBFGS` deberá usar las derivadas numéricas para calcular tales cantidades. Así lo indican el número 0 y 1 en el cuarto y quinto argumento de `MaxBFGS(...)` (línea 55).

La rutina encargada de la maximización está dada en la línea 48 por `MV(const vp_ini, const Pe, const serie)`, donde `vp_ini` proporciona un vector de parámetros iniciales para el algoritmo (BFGS), `Pe` sirve para guardar el resultado y `serie` provee los datos con los cuales se hará la estimación de los estimadores máximo verosímiles, línea 72.

La función `fX(const x, const Vlt)` es la densidad de una distribución Binomial Negativa($\lambda, 1 - e^\theta$). Esta función nos permite graficar y obtener los valores estimados del ajuste, líneas 76-78.

El comando `loadmat` en la línea 70 permite a `0x` leer archivos `*.xls`, `*.mat`, `*.dat`, entre otros. El archivo `DatosPolio.xls` contiene las 168 observaciones de poliomielitis (Cuadro 7.9) en una columna.

El Ajuste 5 requiere un archivo que no contenga las observaciones treinta y cinco con valor 14, ni

la séptima con valor 9.

Mediante la función `DrawDendity` en la línea 75 obtenemos el histograma de los datos de poliomielitis (las 168 observaciones). Mientras que con `DrawXmatrix` graficamos el ajuste y con `SaveDrawWindow` guardamos la figura generada en `GiveWin`. Obteniendo una gráfica como la de la Figura 7.7 pero con el Ajuste 2.

Una explicación más detallada sobre los argumentos de las funciones implementadas por `Ox` puede consultarse en la ayuda del programa, situada generalmente en el archivo

`C:\Archivos de programa\Ox\doc\index.html`.

7.4. Datos

Cuadro 7.9: Número de casos mensuales de poliomielitis en U.S. de 1970 a 1983.

	Ene.	Feb.	Mar.	Abr.	May.	Jun.	Jul.	Ago.	Sep.	Oct.	Nov.	Dic.
1970	0	1	0	0	1	3	9	2	3	5	3	5
1971	2	2	0	1	0	1	3	3	2	1	1	5
1972	0	3	1	0	1	4	0	0	1	6	14	1
1973	1	0	0	1	1	1	1	0	1	0	1	0
1974	1	0	1	0	1	0	1	0	1	0	0	2
1975	0	1	0	1	0	0	1	2	0	0	1	2
1976	0	3	1	1	0	2	0	4	0	2	1	1
1977	1	1	0	1	1	0	2	1	3	1	2	4
1978	0	0	0	1	0	1	0	2	2	4	2	3
1979	3	0	0	2	7	8	2	4	1	1	2	4
1980	0	1	1	1	3	0	0	0	0	1	0	1
1981	1	0	0	0	0	0	1	2	0	2	0	0
1982	0	1	0	1	0	1	0	2	0	0	1	2
1983	0	1	0	0	0	1	2	1	0	1	3	6

Capítulo 8

Conclusiones

Series de tiempo con valores discretos significa para nosotros una apasionante área de estudio de la estadística y la probabilidad aplicada.

En esta tesis hemos presentado y estudiado los principales modelos de series de tiempo con valores discretos que surgieron entre los años 1978 y 1998. Centrándonos de manera exhaustiva en procesos Autoregresivos y en ligera proporción en procesos con estructuras MA y ARMA, todos, en su gran mayoría, de orden uno.

Modelos muy variados, de las maneras más ingeniosas, se desarrollaron durante estos 20 años. Desde una mezcla probabilística (1978) hasta la distribución condicional de una v.a. dada esta misma más la suma de otra v.a.i., recibiendo esta distribución condicional el nombre de *contracción*. (1996 y 1998). Ambas ideas, aunque sencillas, ostentan un gran contenido. La primera se utilizó por vez primera, en el contexto de series de tiempo, en modelos continuos con distribución marginal Exponencial. Modelos que poseen un análogo discreto, con distribución marginal Geométrica, al reemplazar la multiplicación escalar por adelgazamiento Binomial, herramienta fundamental para el desarrollo de series de tiempo con valores discretos. Además, este adelgazamiento, al igual que el adelgazamiento Hipergeométrico, puede obtenerse como un caso particular de la segunda idea, *contracción*.

Los modelos de series de tiempo con espacio de estados discreto aquí tratados, exhiben una elegancia teórica admirable. Por ejemplo, para los procesos AR(1), propiedades como la estacionariedad, en la mayoría de los casos, se da por la construcción del modelo, radicando la dificultad del proceso en atinar la distribución que sigue la sucesión de innovación. En tanto que propiedades como la correlación serial se ha obtenido de manera análoga para todos los procesos AR(1) discretos, basándonos en la esperanza condicional lineal que poseen.

Para otras estructuras como la MA(1) y la ARMA(1,1), el adelgazamiento Binomial e Hipergeo-

métrico han permitido la creación de procesos estacionarios con tales estructuras, y distribuciones marginales como la Poisson, la Geométrica y la Binomial Negativa.

No obstante, no existe una familia de modelos con estructura simple, práctica y flexible, que pudiese ser provechosa para una variedad de tipos datos (con estructuras AR, MA o ARMA), como la familia ARMA Gaussiana en el caso de series de tiempo con valores continuos. Lo que se tiene es un amplio rango de familias, cada una con definición, características y, limitantes propias y distintas. El trabajo de Joe (1996) consideramos es un valioso aporte, sin precedentes, que unifica varios de los modelos publicados anteriormente, conocidos como procesos AR(1) con marginal no-Gaussiana, tales como: Lewis (1983) y Lewis *et al.* (1989) para la distribución Gamma, McKenzie (1986) para la Binomial Negativa, Al-Osh y Alzaid (1987) y McKenzie (1988b) para la Poisson, y Al-Osh y Alzaid (1991) para la distribución Binomial. (El trabajo de Joe (1996) incluye como posible marginal la distribución Normal, y de manera general distribuciones en las familias paramétricas infinitamente-divisibles cerradas bajo convolución.) Sin embargo, esta unificación es más a nivel teórico que práctico.

Aunque la aplicación de los modelos discretos es uno de los puntos más explorados hoy día, las publicaciones que tratan sobre éste son realmente pocas, y el número que trata con datos no simulados se reduce aún más. En este trabajo hemos presentado un ejemplo en el que usamos los casos mensuales de poliomielitis en EE.UU. entre los años 1970 y 1983. No existe una metodología clara, formal y de carácter universal para el análisis de series de tiempo con valores discretos. En nuestro caso hicimos lo siguiente:

1. Conocer la(s) posible(s) distribución(es) que siguen los datos.
2. Identificar la posible correlación-dependencia que guardan los datos; escoger –si es el caso– entre las diversas familias que tienen esa estructura y distribución una de éstas.
3. Estimación de los parámetros.
4. Elección del modelo que mejor se ajusta a los datos, ajuste óptimo.

Varias consideraciones deben hacerse al respecto:

Ante la gran variedad de familias, cuyos modelos poseen expresiones de acuerdo a la distribución asignada, los modelos discretos aquí expuestos no han encontrado un espacio en los paquetes estadísticos. Así, para llevar a cabo la estimación de parámetros, punto 3, es necesaria la implementación personal.

Para datos que exhiben una posible estructura AR(1), y usando el método de MV, tal implementación es relativamente sencilla. Hemos presentado las probabilidades de transición de prácticamente todos los procesos AR(1) discretos vistos aquí, por lo que sólo necesitamos implementar

dichas probabilidades y una rutina de maximización. Este fue el procedimiento que nosotros desarrollamos.

Se desconocen las propiedades de los estimadores MV de cada modelo. Más aún, determinar éstas, en algunos casos, podría no ser inmediato, pues los estimadores MV podrían no tener una expresión analítica. Sin embargo, la determinación de las mismas podría ayudar como un criterio en la elección de familia que debe hacerse en el punto 2. Otro elemento que podría ayudar, sobremanera, a hacer dicha elección sería el contar con un criterio análogo al de Akaike para series de tiempo con valores continuos.

Nuestra determinación del modelo que mejor se ajusta a los datos, ajuste óptimo, punto 5, carece de un fundamento teórico que vaya más allá de la inspección gráfica y, finalmente, de la consideración de la distancia entre la frecuencia estimada por el modelo-ajuste y la frecuencia real de los datos. Pruebas de diagnóstico son necesarias, éste es un punto en el que se debe profundizar.

Así, el análisis de series de tiempo con valores discretos es un área de investigación en pleno desarrollo. Continúan (y continuarán) surgiendo publicaciones sobre el tema. Sin embargo, creemos que su estudio se encuentra aún lejos de alcanzar una unificación y generalización que le permita tener una metodología formal y práctica, de fácil e inmediato acceso que pudiese ser provechosa en las aplicaciones para todo usuario interesado en el tema.

Bibliografía

- Al-Osh, M.A. and Aly, E.-E.A.A. (1992) First order autoregressive time series with negative binomial and geometric marginals. *Commun. Statist.-Theory Meth.* **21**, 2483-2492.
- Al-Osh, M.A. and Alzaid, A.A. (1987) First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **8**, 261-275.
- Al-Osh, M.A. and Alzaid, A.A. (1991) Binomial autoregressive moving average models. *Stoch. Models* **7**, 261-282.
- Alzaid, A.A. and Al-Osh, M.A. (1990) An integer-valued pth-order autoregressive structure (INAR(p)) process. *J. Appl. Prob.* **27**, 314-324.
- Anderson, O.D. (1976) *Time Series Analysis and Forecasting: The Box-Jenkins approach*. Butterworths.
- Chatfield, C. (2003) *The Analysis of Time Series: An introduction*. Chapman & Hall.
- Da Silva, M.E. and Oliveira, V.L. (2004) Difference equations for the higher-order moments and cumulants of the INAR(1) model. *J. Time Series Anal.* **25**, 317-333.
- Du, J.-G. and Li, Y. (1991) The Integer Valued Autoregressive (INAR(p)) Model. *J. Time Series Anal.* **12**, 129-142.
- Gaver, D.P. and Lewis, P.A.W. (1980) First order autoregressive gamma sequences and point processes. *Adv. Appl. Prob.* **12**, 724-745.
- Grunwald, G.K., Hyndman, R.J., Tedesco, L.M. and Tweedie, R.L. (2000) Non-Gaussian conditional linear AR(1) models. *Australian and New Zealand Journal of Statistics*, **42**(4), 479-495.
- Jacobs, P.A. and Lewis, P.A.W. (1977) A mixed autoregressive-moving average exponential sequence and point process, EARMA(1,1). *Adv. Appl. Prob.* **9**, 87-104.
- Jacobs, P.A. and Lewis P.A.W. (1978a) Discrete Time Series generated by Mixtures I: Correlational and Runs Properties. *J.R. Statist. Soc. (B)* **40**, 94-105.

- Jacobs, P.A. and Lewis, P.A.W. (1978b) Discrete Time Series generated by Mixtures II: Asymptotic Properties. *J.R. Statist. Soc. (B)* **40**, 222-228.
- Jacobs, P.A. and Lewis, P.A.W. (1978c) Discrete Time Series generated by Mixtures III: Autoregressive Processes (DAR(p)). Naval Postgraduate School Technical Report NPS55Lw 73061A.
- Jacobs, P.A. and Lewis, P.A.W. (1983) Stationary Discrete Autoregressive-Moving Average Time Series generated by Mixtures. *J. Time Ser. Anal.* **4**, 19-36.
- Joe, H. (1996) Time series models with univariate margins in the convolution-closed infinitely divisible class. *J. Appl. Prob.* **33**, 664-677.
- Jørgensen, B. (1986) Some properties of exponential dispersion models. *Scand. J. Statist.* **13**, 187-198.
- Jørgensen, B. (1992) Exponential dispersion models and extensions: a review. *Int. Statist. Rev.* **60**, 5-20.
- Jørgensen, B. y Song, P.X.-K. (1998) Stationary time series models with exponential dispersion model margins. *J. Appl. Prob.* **35**, 78-92.
- Kanter, M. (1975) Autoregression for discrete processes mod 2. *J. Appl. Prob.* **12**, 371-375.
- Latour, A. (1998) Existence and Stochastic Structure of a non-negative integer-valued autoregressive process. *J. Time Series Anal.* **19**, 439-455.
- Lawrance, A.J. (1982) The innovation distribution of a gamma distributed autoregressive process. *Scand. J. Statist.* **9**, 234-236.
- Lawrance, A.J. and Lewis, P.A.W. (1977) An Exponential Moving Average sequence and point process, EMA(1). *J. Appl. Prob.* **14**, 98-113.
- Lawrance, A.J. and Lewis, P.A.W. (1980) The Exponential Autoregressive-Moving Average EAR-MA(p,q) Process. *J.R. Statist. Soc. (B)* **42**, 150-161.
- Lawrance, A.J. and Lewis, P.A.W. (1985) Modelling and residual analysis of nonlinear autoregressive time series in exponential variables (with discussion). *J.R. Statist. Soc. (B)* **47**, 165-183.
- Lewis, P.A.W. (1983) Generating negatively correlated gamma variates using the beta-gamma transform. In *Proc. 1983 Winter Simulation Conf.* ed. S. Roberts, J. Banks and B. Schmeiser. IEEE Press, New York. pp. 175-176.
- Lewis, P.A.W., McKenzie, E. and Hugus, D.K. (1989) Gamma Processes. *Commun. Statist. Stoch. Models* **5**, 1-30.

-
- Littlejohn, R.P. (1992) Discrete minification processes and reversibility. *J. Appl. Prob.* **29**, 82-91.
- MacDonald, I.L. and Zucchini, W. (1997) *Hidden Markov and Other Models for Discrete-valued Time Series*. Chapman & Hall, London.
- McKenzie, E. (1986) Autoregressive-moving average processes with negative-binomial and geometric marginal distributions. *Adv. Appl. Prob.* **18**, 679-705.
- McKenzie, E. (1988a) The distributional structure of finite moving average processes. *J. Appl. Prob.* **25**, 313-321.
- McKenzie, E. (1988b) Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Prob.* **20**, 822-835.
- McKenzie, E. (2003) Discrete variate time series. In *Stochastic processes: modelling and simulation*, volume **21** of Handbook of Statist., 573-606. North-Holland, Amsterdam.
- Mena, R.H. (2004) Un ejemplo en $\mathbb{O}x$. *DATOS* **26**, 4-7.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical recipes in C: The Art of Scientific Computing* (2nd edition). Cambridge University Press.
- Sim, C.H. (1990) First-order autoregressive models for gamma and exponential variables. *J. Appl. Prob.* **27**, 325-332.
- Stutel, F.W. and van Harn, K. (1979) Discrete analogues of self-decomposability and stability. *Ann. Prob.* **7**, 893-899.
- Stutel, F.W. y van Harn, K. (2004) *Infinite divisibility of probability distributions on the real line*. New York, M. Dekker.
- Zeger, S.L. (1988) A regression model for time series of counts. *Biometrika* **75**, 621-629.

Notación y abreviaturas

\mathbb{N}_0	$\mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$
\square	Indica el final de la prueba.
$U \stackrel{d}{=} V$	Igualdad en distribución de las variables aleatorias U y V .
$y \wedge x$	El mínimo entre y y x .
$:=$	Definido como.
$X \sim F$	X se distribuye F .
\otimes	Operador de convolución.
\mathbb{P}	Medida de probabilidad.
\mathbb{R}^+	El conjunto de los reales positivos.
T.L.	Transformada de Laplace.
c.p.	con probabilidad
e.o.c.	en otro caso
ec.	ecuación
v.a.	variable aleatoria
i.i.d.	independientes e idénticamente distribuidas.
f.d.p.	función de densidad de probabilidad.
S-D	Descomponible-por sí misma (por sus siglas en Inglés).
I-D	Infinito-divisible
ED	Modelo de dispersión Exponencial.
L.Q.Q.D.	Lo que queríamos demostrar.
MV	Máxima verosimilitud.

