

12
29



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ACATLAN**

**LAS SERIES DE TIEMPO EN LA
TEORIA DEL RIESGO**

T E S I S
QUE PARA OBTENER EL TITULO DE
LICENCIADO EN ACTUARIA

P R E S E N T A
MAURICIO IGNACIO GONZALEZ VERGARA

ASESOR:
M. EN C. LUCIO PEREZ RODRIGUEZ



MEXICO, D. F.

DICIEMBRE DE 1993

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México



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 está 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.



UNIVERSIDAD NACIONAL
AVENIDA DE
MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES "ACATLAN"

DIVISION DE MATEMATICAS E INGENIERIA
PROGRAMA DE ACTUARIA Y MAC.

SR. MAURICIO IGNACIO GONZALEZ VERGARA
Alumno de la carrera de Actuaría,
P r e s e n t e .

De acuerdo a su solicitud presentada con fecha 22 de mayo de 1992, me complace notificarle que esta Jefatura tuvo a bien asignarle el siguiente tema de tesis: "LAS SERIES DE TIEMPO EN LA TEORIA DEL RIESGO", el cual se desarrollará como sigue:

INTRODUCCION

CAPITULO 1.

La función monto de reclamos.

CAPITULO 2.

Estimación de la función monto de reclamos acumulados por medio de las series de tiempo.

CAPITULO 3.

Dependencia de la función monto de reclamos acumulados en reaseguro.

CAPITULO 4.

Probabilidad de ruina.

CONCLUSIONES.

Asimismo fue designado como Asesor de Tesis el - M. en C. LUCIO PEREZ RODRIGUEZ, profesor de esta Escuela.

Ruego a usted tomar nota que en cumplimiento de lo especificado en la Ley de Profesiones, deberá presentar servicio social durante un tiempo mínimo de seis meses como requisito básico para sustentar examen profesional, así como de la disposición de la - Coordinación de la Administración Escolar en el sentido de que se imprima en lugar visible de los ejemplares de la tesis el título del trabajo realizado. Esta comunicación deberá imprimirse en el interior de la - tesis.

A T E N T A M E N T E
"POR MI RAZA HABLARA EL ESPIRITU"
Acatlán, Edo. Méx. noviembre 23 de 1993.

ACT. LARRY M. RIVERA BECERRA
Jefe del Programa.

LMRB'og.

Indice de la Tesis.

Introducción.		1
Capítulo 1.	La Función Monto de Reclamos.	22
1.1	El Proceso de Riesgo. Fundamentos.	22
1.2	El Proceso del Número de Reclamos.	24
1.3	El Proceso de Reclamos Acumulados.	39
Capítulo 2.	Estimación de la Función Monto de Reclamos acumulados por medio de las Series de Tiempo.	44
2.1	Análisis de la Serie de Tiempo.	63
2.2	Modelado de la Serie de Tiempo.	73
2.3	Estimación por medio de la Serie de Tiempo.	109
Capítulo 3.	Dependencia de la Función Monto de Reclamos Acumulados en el Reaseguro.	121
3.1	Aspectos Generales.	121
3.2	Distintos tipos de Reaseguro.	122
Capítulo 4.	Probabilidad de Ruina.	146
4.1	Introducción.	146
4.2	Discusión de los diferentes métodos	147
Conclusiones		158

Bibliografía básica.

Pentikäinen T. et al (1984); "Risk Theory"; Third Edition; Chapman and Hall; New York.

Bühlman Hans (1970); "Mathematical Methods in Risk Theory"; Springer-Verlag; Berlin.

Harvey A.C. (1986); "Time Series Model"; Wiley; New York.

Feller William (1989); "Introducción a la Teoría de Probabilidades y sus aplicaciones"; Noriega Editores; México; Volumen I.

Feller William (1989); "Introducción a la Teoría de Probabilidades y sus aplicaciones"; Noriega Editores; México; Volumen II.

Parzen E. (1986); "Teoría Moderna de la Probabilidad y sus aplicaciones"; Wiley; New York.

Parzen E. (1962); "Stochastic Processes"; Holden-Day; New York.

Bartlett M.S. (1955); "An Introduction to Stochastic Processes"; Cambridge University Press; London.

Isaacson Dean L. et al (1976); "Markov Chains and Applications"; Wiley; new York.

Apuntes de la materia "Series de Tiempo" con el profesor Víctor Espinoza. 1989.

nombre: Mauricio Ignacio González Vergara
escuela: Universidad Nacional Autónoma de México
Escuela Nacional de Estudios Profesionales
Acatlán
licenciatura: Actuaría
número de cuenta: 8652317-8

dirección del hogar:
Progreso 32-8. Colonia Industrial. C.P. 07800
Delegación Gustavo A. Madero.
Teléfono: 5 37 43 89

nombre del asesor de Seminario de Tesis:
Actuario Enrique Peña

nombre del asesor de Tesis:
M. en C. Lucio Pérez Rodríguez
dirección (hogar): Emperadores 99-10
Colonia Portales

dirección (trabajo):
Razón social: U.N.A.M. E.N.E.P. Acatlán
Puesto: Investigador.
Dirección: Alcanfores y Sn. Juan
Totoltepec

nombre de la Tesis:
LAS SERIES DE TIEMPO EN LA TEORIA DEL RIESGO.

INTRODUCCION

I.1. VARIABLES ALEATORIAS.

I.1.1. Definición

Las técnicas actuariales convencionales se basan en las frecuencias y promedios del monto de reclamos. Por ejemplo, si un asegurador tiene una cartera de N pólizas de cualquier riesgo y si el promedio esperado del número de reclamos q y el promedio esperado del costo por reclamo es m , entonces la cantidad total del dinero reclamado debido a los siniestros sucedidos es Nqm . Pero, la cantidad real difiere de esta figura y fluctuará alrededor de esta, y se nota que el fenómeno de la fluctuación no se considera. Se toma la cantidad Nqm como un valor determinístico y no probabilístico. Los estudios de las diferentes clases de fluctuaciones que aparecen en una cartera de seguros constituyen el ramo de las matemáticas actuariales determinada como Teoría del Riesgo. Por ello se debe considerar al número y tamaño (costo) de las reclamaciones, así como otras cantidades, como variables aleatorias.

Una variable aleatoria, o más propiamente, una función aleatoria, S es una función definida en un espacio muestral que transforma los resultados del espacio muestral Ω en puntos sobre la recta de los reales. Esta variable aleatoria tiene asociada una función de distribución acumulativa $F_S(s)$ definida como la probabilidad de que la variable aleatoria S toma un valor numérico real, menor o igual que s , que se encuentra en el espacio muestral Ω .

$$F_S(s) = P [S \leq s]$$

Se tomarán en consideración dos tipos de distribución:

- a) funciones de distribución discretas:
- (i) Distribución binomial
 - (ii) Distribución poisson
 - (iii) Distribución binomial negativa
 - (iv) Distribución logarítmica

b) funciones de distribución con funciones de densidad

- i) Distribución normal
- ii) Distribución log-normal
- iii) Distribución gamma
- iv) Distribución beta
- v) Distribución Cauchy
- vi) Distribución Pareto

I.1.2. VALORES ESPERADOS

El concepto de valor esperado, esperanza o promedio de una variable aleatoria es muy útil, como se nota en el caso del reaseguro. Por medio de este valor se dará una idea de cuanto será el monto total de la reclamación que ocurre provocado por un siniestro. En el caso del reaseguro, nos ayuda para saber el número de reclamos que deberá retener la Cia. de seguros para no tener una pérdida catastrófica debido a la cantidad exagerada de riesgos que retuvo.

Se define al valor esperado o esperanza o promedio como sigue:

$$E [g(s)] = \int_{-\infty}^{\infty} g(s) dF_S(s)$$

Interpretación intuitiva de valor esperado.

Tómese en cuenta la variable aleatoria S , que representa la cantidad de pérdida en unidades monetarias, que se produce por que se llevó a cabo un siniestro, como son los daños causados a un edificio por fuego. Sea $g(s)$ el pago del reclamo que se debe hacer por un monto de pérdida de S y sea $F_S(s)$ una función de distribución discreta, entonces:

$$E [g(s)] = \sum_{k=1}^{\infty} g(k) P_k \quad \text{donde } P_k = \text{Prob} [s = k] = \frac{n_k}{n}$$

$$\frac{n_k}{n} = \frac{\text{numero de reclamos de tipo } k}{\text{numero total de reclamos}}$$

Por lo tanto:

$$E[g(s)] = \frac{\sum g(k) n_k}{n} = \frac{\text{Pago de reclamos totales del tipo } k}{\text{numero total de reclamos}}$$

= pago promedio por reclamo

Aplicación práctica en reaseguro

Sea S , el monto perdido y sea $g(s)$ la parte del reclamo cubierto por cuenta de la aseguradora (retención). Diferentes tipos de $g(s)$ pueden representar distintos tipos de reaseguros, por ejemplo:

$g(s) = a S$ ($0 \leq a \leq 1$) Reaseguro proporcional con una cuota retenida de a ;

$g(s) = \begin{cases} S & \text{para } S \leq M \\ M & \text{para } S > M \end{cases}$ Reaseguro por exceso de pérdida con primer riesgo M

En cualquier caso:

$E[g(s)] =$ reclamos esperados dentro de la retención

$E[S - g(s)] =$ reclamos esperados en reaseguro

Propiedades fundamentales del valor esperado empleadas en reaseguro.

$$E_1) \quad E[a S] = a E[S]$$

En palabras: El valor esperado dentro del monto de retención es " a " veces el total esperado en pérdidas.

$$E_2) \quad E[g(s)] + E[S - g(s)] = E[S]$$

En palabras: Los reclamos esperados retenidos más los reclamos esperados reasegurados es igual a los reclamos esperados totales.

E_s) Como $g(s) \leq S$ entonces $E [g(s)] \leq E [S]$

En palabras: Los reclamos esperados retenidos son menores o iguales a los reclamos totales esperados.

I.1.3. CARACTERISTICAS Y FUNCIONES AUXILIARES DE UNA DISTRIBUCION DE PROBABILIDAD.

a) Características:

i) $E [S^k] = \mu_k$: k-ésimo momento no central de $F_S(s)$
Caso especial: $\mu_1 = \mu$. Valor esperado de $F_S(s)$.

ii) $E [(S - \mu)^k] = \alpha_k$: k-ésimo momento central de $F_S(s)$.
Caso especial: $\alpha_2 = E [(S - \mu)^2] = \sigma^2$: varianza
 σ : desv. std.

En adición a estas medidas:

$$\gamma = \frac{\alpha_3}{\alpha_2^{3/2}} = \frac{E [(S - \mu)^3]}{\sigma^3} : \text{asimetría}$$

El valor esperado, esperanza o promedio mide el centro de gravedad de la distribución. Análogamente la varianza se considera como una medida de dispersión o desviación media cuadrática del valor esperado. La asimetría mide la extensión a la cual la masa de la distribución de probabilidades se extiende a la izquierda (asimetría negativa) o a la derecha (asimetría positiva) del valor esperado.

b) Funciones auxiliares.

i) función generatriz de momentos

$$M(t) = E [e^{it}] = \int_{-\infty}^{\infty} e^{it} dF_S(s); t \in \mathbb{R}$$

Si la función generatriz de momentos está definida para la función de distribución, entonces la siguiente relación es válida:

$$M^{(k)}(0) = E [X^k]$$

La relación entre función generatriz de momentos y la función de distribución es una a una y por lo tanto se puede describir una función de distribución usando la función auxiliar.

I.2. SUCCESIONES DE VARIABLES ALEATORIAS

I.2.1. Distribución de varias variables y funciones auxiliares.

Las distribuciones de varias variables se pueden emplear dentro de la teoría del riesgo en las reproducciones de monto de reclamos de un mismo riesgo, ya que cuando se habla de riesgo no solamente se piensa de un sólo conductor de una automóvil quien se asegura contra reclamos por responsabilidad civil o de un solo asegurado bajo una póliza contra accidentes. Se puede incluir estos riesgos sencillos en el sentido usual de la palabra bajo el concepto general de riesgo. Un grupo de pólizas de seguro o un completo trato por reaseguro puede ser igualmente bien referido como un solo riesgo. La separación tradicional de riesgos individuales de los colectivos no es necesaria para la teoría matemática, y en realidad es confuso, por lo tanto se hablará de riesgos sin distinción.

Las leyes de probabilidad regidoras de las sucesiones de variables aleatorias $X_1, X_2, \dots, X_n, \dots$ son descritas por funciones de distribución de varias variables:

$$G_{S_1, S_2, \dots, S_n}(a_1, a_2, \dots, a_n) = \text{Prob} [S_1 \leq a_1, \dots, S_n \leq a_n]$$

que es igual a la probabilidad de que todas las condiciones $S_k \leq a_k$ ($k = 1, 2, \dots, n$) se cumplen simultáneamente.

Las funciones de distribución k-dimensionales $F_{S_k}(s)$ ($k = 1, 2, \dots, m$; $m < n$) se pueden obtener a partir de la función de distribución de varias variables $G_{S_i}(s_i)$ ($i = 1, 2, \dots, n$) por medio de la siguiente relación:

$$F_{S_k}(s) = G_{S_1, S_2, \dots, S_n}(s_1 = \infty, s_2 = \infty, \dots, s_k = s, \dots, s_n = \infty)$$

Otra relación de importancia es la siguiente:

Si los primeros momentos de S_1, \dots, S_n existen, entonces:

$$E[S_1 + S_2 + S_3 + \dots + S_n] = E[S_1] + E[S_2] + \dots + E[S_n]$$

Las funciones auxiliares como:

a) Función generatriz de momentos:

$$M_{S_1, S_2, \dots, S_n}(t_1, t_2, \dots, t_n) = E[e^{t_1 s_1 + \dots + t_n s_n}] =$$

$$\int_{-\infty}^{\infty} e^{t_1 s_1 + \dots + t_n s_n} dG_{S_1, \dots, S_n}(s_1, \dots, s_n)$$

también se asocian con las funciones de distribución de varias variables $G_{S_1, \dots, S_n}(s_1, \dots, s_n)$ en una manera única.

I.3. Probabilidad y Esperanza Condicional.

La probabilidad condicional de un evento A dado el valor de una variable aleatoria S, $P[A | S = s]$, tiene un papel importante dentro de la Teoría de Procesos Estocásticos y estos a su vez son importantes en la Teoría del Riesgo, tal es el caso de las Series de Tiempo.

La función de distribución condicional de una variable aleatoria T, dado (el valor de) una variable aleatoria S:

$$F_{T|S}(t|s) = P[T \leq t | S = s]$$

y la esperanza condicional de una variable aleatoria T, dado (el valor de) una variable aleatoria S:

$$E [T | S=s]$$

desempeñan un rol importante dentro de la teoría de procesos estocásticos.

La probabilidad condicional $P [A|B]$ del evento A, dado el evento B está definida como sigue:

$$P [A|B] = \frac{P [A \cap B]}{P [B]} \quad P [B] > 0$$

En base a lo anterior se puede definir $P [A | S=s]$, la probabilidad condicional de un evento A, dado el evento $[S=s]$ de que el valor observado de la variable aleatoria S es igual a s, como sigue:

$$P [A|S=s] = \frac{P [A \text{ ocurre y } S=s]}{P [S=s]} \quad P [S=s] > 0$$

Por medio de la definición básica anterior, se da lo siguiente:

a) Para funciones de distribución continuas, teniendo funciones de densidad:

$$f_{T|S} (t|s) = \frac{g_{s,T} (s,t)}{f_S (s)} = \frac{g_{s,T} (s,t)}{\int_{-\infty}^{\infty} g_{s,T} (s,t) dt}$$

El valor esperado condicional para cualquier función $h(x)$:

$$E [h(T)|S=s] = \int h(w) f_{T|S} (w|s) dw$$

Por analogía se definen las probabilidades y esperanzas condicionales para distribuciones discretas.

Las siguientes dos propiedades son importantes:

$$P [T=k] = \int P [T=k | S = s] dF_S(s) \quad \text{en el caso de una distribución discreta con } P [T|S];$$

$$f_T(t) = \int f_{T|S}(t|s) dF_S(s)$$

en el caso de una
Distr. condicional con
función de densidad
 $f_{T|S}$

Propiedad fundamental de la esperanza condicional.

$$\int E[h(t)|S=s] dF_S(s) = E[h(t)] = E\{E[h(t)|S]\}$$

I.4. Independencia.

Se define a:

$$S_1, S_2, \dots, S_n, \dots$$

como una sucesión de variables aleatorias independientes si todas las distribuciones de varias variables relacionadas, son de la forma producto, i.e. si:

$$G_{S_1, \dots, S_n}(s_1, \dots, s_n) = F_{S_1}(s_1) F_{S_2}(s_2) \dots F_{S_n}(s_n)$$

La independencia se puede definir también por medio de funciones auxiliares de varias variables.

Criterio: Para la independencia de S_1, \dots, S_n las siguientes condiciones son equivalentes:

a) $G_{S_1, \dots, S_n}(s_1, \dots, s_n) = F_{S_1}(s_1) F_{S_2}(s_2) \dots F_{S_n}(s_n)$

b) $M_{S_1, \dots, S_n}(s_1, \dots, s_n) = M_{S_1}(s_1) M_{S_2}(s_2) \dots M_{S_n}(s_n)$

El inciso b) se refiere a las funciones generatrices de momentos.

Propiedades fundamentales de variables independientes

- a) $E [g(S) h(T)] = E [g(S)] E [h(T)]$
- b) $Cov (S, T) = E [(S - E [S]) (T - E [T])] = 0$ (Covarianza de S y T)
- c) $Var (S+T) = Var(S) + Var (T)$
- d) $M_{S+T} (v) = M_S (v) M_T (v)$
- e) $F_{S+T} (s) = \int F_S (s-v) dF_T (v) = \int F_T (s-v) dF_S (v)$

Este último inciso se llama la convolución de S y T, frecuentemente denotada por el símbolo *. (S*T).

I.4.1. Covarianza y correlación.

Para variables aleatorias independientes S y T se cumple:

$$Cov (S, T) = E [ST] - E [S] E [T] = 0$$

La desigualdad de Schwarz siempre se cumple para variables aleatorias dependientes:

$$\begin{aligned} |Cov (S, T)| &\leq \sqrt{E [(S - E [S])^2] E [(T - E [T])^2]} \\ &= \sqrt{Var (S) Var (T)} = \sigma(S) \sigma(T) \end{aligned}$$

El caso $Cov (S, T) = \pm \sigma(S) \sigma(T)$ caracteriza la dependencia lineal. Esto da lugar a introducir la siguiente medida de dependencia:

$\rho (S, T) = \frac{Cov (S, T)}{\sigma(S) \sigma(T)}$ a esta razón se conoce como coeficiente de correlación de S y T.

Las siguientes propiedades se cumplen para el coeficiente de correlación:

- Si S y T son independientes, entonces $\rho(S, T) = 0$. (Lo inverso no se cumple).
- S y T son linealmente dependientes si y solo si $\rho (S, T) = \pm 1$
- $-1 \leq \rho(S, T) \leq 1$

1.5. La ley de los grandes números.

Esta ley se refiere con el comportamiento asintótico del promedio de "n" variables aleatorias.

Sean S_1, \dots, S_n una sucesión de variables aleatorias:

$$X_n = \sum_{i=1}^n S_i \qquad \bar{X}_n = \frac{X_n}{n} = \frac{\sum_{i=1}^n S_i}{n}$$

La definición de convergencia en probabilidad es:

X_n converge en probabilidad a cero si:

$$P\{|X_n| > \epsilon\} \longrightarrow 0$$

para cualquier $\epsilon > 0$, lo que se explica mediante $X_n \xrightarrow{P} 0$.

Por extensión, $X_n \xrightarrow{P} X$ significa lo mismo que

$$X_n - X \xrightarrow{P} 0.$$

Ley de los grandes números:

"Si S_1, \dots, S_n es una sucesión de variables aleatorias independientes idénticamente distribuidas, entonces \bar{X}_n converge en probabilidad hacia la constante $E[S_1]$ ".

El teorema significa que en el caso de reclamos independientes e idénticamente distribuidas, el reclamo promedio de todos ellos se aproxima al valor esperado de un reclamo sencillo con un alto grado de probabilidad. Este hecho se considera como la piedra angular teórica del seguro. Sin embargo, esta afirmación debe ser calificada como una exageración debido al campo de aplicaciones tan reducido que tienen los supuestos de las variables independientes e idénticamente distribuidas. Por ello se establece un teorema de estabilidad el cual generaliza el resultado clásico en el sentido de que los supuestos de variables aleatorias independientes e idénticamente distribuidas no se requieran. Sin embargo, se requieren supuestos relacionados a la existencia de segundos momentos.

Teorema de estabilidad para riesgos

Sea S_1, S_2, \dots una sucesión de variables aleatorias donde

$\text{Var}(S_i) < \sigma_0^2 \quad \forall i \quad \rho(S_i, S_j) = 0 \quad \text{si } |i-j| > N$
 entonces S_n converge en probabilidad a

$$\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n E[S_i]}{n}$$

estipulado que este último límite exista.

Demostración: Por medio de la desigualdad de Chebyshev se ve que:

$$P \left[\frac{|X_n - E[X_n]|}{n} > \epsilon \right] \leq \frac{\text{Var}(X_n)}{\epsilon^2 n^2}$$

Para demostrarlo se tiene que ver que

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(X_n)}{\epsilon^2 n^2} = 0$$

Entonces, $\text{Var}(X_n) = E \left[\left(\sum_{i=1}^n (S_i - E[S_i]) \right)^2 \right] =$

$$\sum_{i,j} E[(S_i - E[S_i])(S_j - E[S_j])]$$

Como $\rho(S_i, S_j) = 0$ para $|i-j| > N$ entonces la última expresión se reduce a:

$$\text{Var}(X_n) = \sum_{|i-j| \leq N} E[(S_i - E[S_i])(S_j - E[S_j])]$$

donde el número de sumandos es, a lo más, igual a $(n - 2N + 1)$ y entonces $\text{Var}(X_n) \leq (2N+1)n \sigma_0^2$, por lo tanto:

$$\lim_{n \rightarrow \infty} \frac{\text{Var}(X_n)}{\epsilon^2 n^2} \leq \lim_{n \rightarrow \infty} \frac{(2N+1)\sigma_0^2}{\epsilon^2 n} = 0 \quad \text{C.Q.D.}$$

El comportamiento de la convergencia de $\frac{\text{Var}(S_n)}{n^2}$ mide la velocidad con la cual se obtiene la estabilidad.

Una interpretación del teorema de estabilidad para riesgos se debe a la convergencia del reclamo promedio hacia una cantidad fija que se puede esperar -para grandes números- para compensar los efectos del azar (monto de reclamos) por cantidades fijas (primas). La velocidad de esta convergencia tiene un papel decisivo en ésta conexión. En particular, se está constantemente confrontado en la práctica con la pregunta de que número, grande y finito, de variables aleatorias se aproxima al resultado asintótico con la suficiente precisión. La desigualdad de Chebyshev puede dar una respuesta preliminar a esta pregunta.

1.5. Procesos estocásticos.

El proceso de reclamos se describe gráficamente en la figura siguiente. Cada vez que ocurre un reclamo se representa por un escalón vertical, la altura del escalón representa el monto del reclamo. El tiempo se mide a la derecha sobre el eje horizontal y la altura X de la línea escalonada muestra el monto total de reclamos durante el intervalo de tiempo $(0,t)$. El desarrollo es un proceso estocástico compuesto en el sentido de que el tiempo de ocurrencia y el número de reclamos son fenómenos aleatorios y el tamaño de cada reclamo es también una variable aleatoria.

1.5.1. Definiciones y clasificación.

Un proceso estocástico es una familia de variables aleatorias definidas sobre algún espacio muestral Ω . Si hay muchos elementos contables de la familia, el proceso se denotará por X_1, X_2, \dots y se llamará proceso estocástico de tiempo discreto. Si hay muchos elementos no contables de la familia el proceso estocástico se denota por $\{X_t\}_{t \geq 0}$ y es de tiempo continuo.

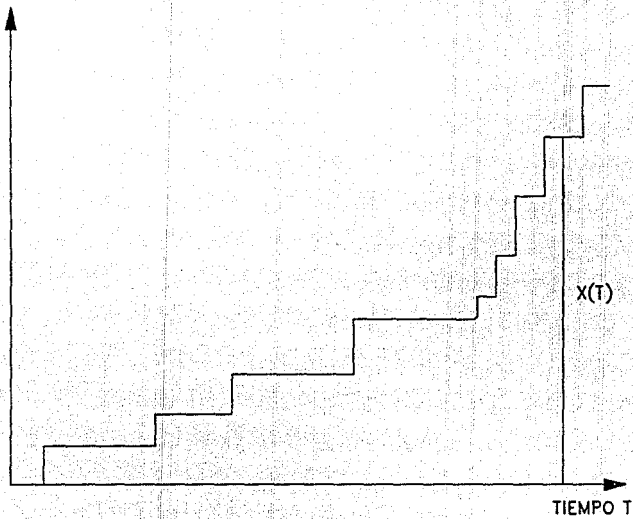
El conjunto de distintos valores que toma un proceso estocástico se llama el espacio de estado. El espacio de estado puede ser, también, discreto o continuo.

En base al tiempo y al espacio de estado se pueden tener 4 tipos de procesos estocásticos:

Procesos	Estocásticos
Tiempo discreto. Espacio discreto.	Tiempo continuo. Espacio discreto.
Tiempo discreto. Espacio continuo.	Tiempo continuo. Espacio continuo.

U.N.A.M.
E.N.E.P ACATLAN
PROCESO DE RECLAMOS

RECLAMOS



1.5.2. Procesos estocásticos con incrementos independientes.

$\{X_t, t \geq 0\}$ es un proceso con incrementos independientes si para todo intervalo $(a_i, b_i]$ que no se traslapan,

$$X_{b_1} - X_{a_1}, X_{b_2} - X_{a_2}, \dots, X_{b_n} - X_{a_n}$$

son variables aleatorias independientes.

$\{X_t, t \geq 0\}$ es un proceso con incrementos estacionarios si

$$F_{X_{t+h} - X_t}(x) \text{ es independiente de } t$$

i.e., si todo incremento del intervalo de la misma longitud tiene la misma distribución. (Homogeneidad).

Los siguientes dos teoremas son de suma importancia:

Teorema I (Proceso de Poisson): Supóngase que $\{X_t, t \geq 0\}$ es un proceso de conteo con incrementos independientes. Si $E[X_t] = \tau$ entonces $m(t)$ se puede introducir siempre como la nueva variable de tiempo de acuerdo con la relación:

$$\bar{X}_t = X_{t/\tau}, \text{ con } t(\tau) = \inf \{s, m(s) = \tau\}$$

asi que

- $\{\bar{X}_t \mid \tau \geq 0\}$ tiene incrementos independientes
- $\{\bar{X}_t \mid \tau \geq 0\}$ tiene incrementos estacionarios
- $P[\bar{X}_t = k] = \exp(-\tau) \frac{\tau^k}{k!} \quad k = 0, 1, 2, \dots$

Nota: $\tau = m(t)$ se llama tiempo operacional. El paso del tiempo ya no se mide en años, horas, segundos, sino en el número de reclamos esperados.

Teorema II (Proceso de Poisson compuesto): Se supone que para el proceso estocástico $(X_t | t \geq 0)$:

- a) $X_0 = 0$
- b) $(X_t | t \geq 0)$ tiene incrementos independientes estacionarios.
- c) $(X_t | t \geq 0)$ tiene solamente funciones de muestreo discretas
- d) (el número esperado de discontinuidades en cada intervalo finito, es finito. Entonces:

$$P[X_t \leq x] = \sum_{k=0}^{\infty} \exp[-\lambda t] \frac{(\lambda t)^k}{k!} F^{*k}(x) \text{ o en términos de}$$

funciones características:

$$P_{X_t}(u) = \exp[\lambda t(x(u)-1)]$$

donde $x(u)$ es la función característica de $F(x)$.

El teorema II determina la ley de distribución a la libre selección de la función $F(x)$ y de λ . El tipo de función de distribución que aparece en el teorema II se llama Poisson compuesta.

Nota: Si se escoge $F(x) = 0$ para $x < 1$ y $F(x) = 1$ para $x \geq 1$ entonces se obtiene la distribución de Poisson para X_t .

Si $E[Y^2] < \infty$, entonces $X(t) = X_t$ tiene segundos momentos finitos dados por:

$$\begin{aligned} E[X(t)] &= \lambda t E[Y] \\ \text{Var}[X(t)] &= \lambda t E[Y^2] \\ \text{Cov}[X(S), X(T)] &= \lambda E[Y^2] \min(s, t) \end{aligned}$$

Y representa la variable aleatoria independiente e idénticamente distribuida, que da origen a la función característica común $x(u)$.

1.5.3. Procesos de Markov.

Los procesos de Markov son una generalización de los procesos con incrementos independientes. La probabilidad de las variables del proceso futuro depende solamente en el último valor conocido del proceso (en el presente o en el pasado). Esta propiedad se conoce como la propiedad de Markov M:

$$P[X_{m+1} \leq x_{m+1}, \dots, X_n \leq x_n | X_1 = x_1, \dots, X_m = x_m] =$$

$$P[X_{m+1} \leq x_{m+1}, \dots, X_n \leq x_n | X_m = x_m] \quad (M)$$

donde la sucesión

$$t_1 < t_2 < t_3 < \dots < t_m < t_{m+1} < \dots < t_n$$

se ordena de acuerdo al orden natural del tiempo.

La ley de probabilidad que gobierna a un proceso de Markov se describe por las probabilidades de transición

$$P[X_t \leq x | X_{t'} = y] = Q(t', t; y, x), \quad t' \leq t$$

De la condición de Markov se sigue que la probabilidad de transición debe satisfacer la ecuación Chapman-Kolmogorov

$$Q(t', t; y, x) = \int_{s=-\infty}^{\infty} dQ(t', t; y, s) Q(t, t; s, x) \quad t' \leq \tau \leq t$$

Cuando el proceso toma solamente un número contable de valores, la probabilidad de transición se escribe de la forma

$$P[X_t = a_k | X_{t'} = a_j] = p_{jk}(t, t')$$

y la ecuación de Chapman-Kolmogorov se escribe

$$p_{jk}(t, t) = \sum_i p_{ji}(t', t) p_{ik}(t, t)$$

Las probabilidades de transición en el caso discreto se supone que tienen las siguientes propiedades de continuidad

a) $P_{jk}(t', t)$ es continua para todo par (t', t)

$$b) \lim_{\tau \rightarrow t} P_{jk}(t, \tau) = P_{jk}(t, t) = \begin{cases} 1 & \text{si } j=k \\ 0 & \text{si } j \neq k \end{cases}$$

Las probabilidades de transición se llaman estacionarias si

$$P_{jk}(t', t) = P_{jk}(t'+h, t+h) \quad \text{para toda } h > 0.$$

1.6. Series de Tiempo.

Las series de tiempo dentro de la teoría del riesgo es una herramienta importante para poder analizar, modelar y predecir como se comportan los reclamos que ocurren al realizarse un siniestro. Es útil porque los reclamos presentan ciclos, tendencias, estacionalidad y fluctuaciones aleatorias, componentes que también debe de tener una serie de tiempo. Para poder tratar con la componente de los ciclos se puede usar la teoría de series de tiempo.

Una aproximación sencilla en el modelado de como se comportan los reclamos, Z , es suponer que el valor $Z(t)$ depende del estado del proceso en los tiempos previos $t-1$, $t-2$, También Z está sujeta a irregularidades más o menos fuertes que pueden ser supuestas como fluctuaciones puramente aleatorias. Estas características dan una base para la construcción de modelos matemáticos para describir el proceso del tipo:

$$Z(t) = f(Z(t-1), Z(t-2), \dots) + a(t) + \theta_1 a(t-1) + \theta_2 a(t-2) + \dots$$

donde la función "descriptoria" o "predictora" f conduce la realimentación de la experiencia pasada al futuro. Los términos $a(t)$, el tan llamado "ruido", representa el efecto aleatorio, el cual se supone normalmente distribuido. $(N(0, \sigma^2))$.

Las funciones lineares son aproximaciones satisfactorias para:

$$Z(t) = \theta_1 Z(t-1) + \theta_2 Z(t-2) + \dots + a(t) + \theta_1 a(t-1) + \theta_2 a(t-2) + \dots$$

frecuentemente, las cuales se conocen como procesos ARMA (Auto-Regresivo Moving Average, Auto-Regresivo Medias Móviles). Los ciclos se generan por medio de una elección adecuada de coeficientes. En términos de la teoría existe una fuerte autocorrelación entre valores de variables consecutivos.

Si $\theta_1 = \theta_2 = \dots = 0$ entonces $Z(t)$ se reduce a un proceso auto-regresivo (Proceso AR). Un beneficio de la aproximación ARMA es que una serie de tiempo estacionaria se puede describir por un modelo que tenga menos parámetros que un proceso AR puro.

Cualquier proceso estacionario discreto se puede expresar como la suma de dos procesos no correlacionados, uno puramente determinístico y otro puramente aleatorio.

1.6.1. Definiciones y generalidades

Un conjunto de observaciones arregladas cronológicamente se llama una serie de tiempo. Las series de tiempo se observan en conexión con muchos fenómenos diversos y por una amplia variedad de investigadores.

Para representar una serie de tiempo se hace lo siguiente. El conjunto de puntos de tiempo al cual las medidas son hechas se llama T . En muchas aplicaciones, T es un conjunto de puntos equidistantes discretos del tiempo (en cuyo caso se escribe $T = \{1, 2, 3, \dots, N\}$, donde N es el número de observaciones) o T es un intervalo del eje real del tiempo (en cuyo caso se escribe $T = \{0 \leq t \leq L\}$, donde L es la longitud del intervalo). La observación hecha al tiempo t se denota por $X(t)$ o X_t . El conjunto de observaciones $\{X(t), t \in T\}$ se llama una serie de tiempo.

La idea básica (ver Barlett [1955]) de la estadística clásica del análisis de una serie de tiempo $\{X(t), t \in T\}$ es considerar a la serie de tiempo como una observación hecha en una familia de variables aleatorias $\{X(t), t \in T\}$, esto es, para cada t en T , la observación $X(t)$ es un valor observado de una variable aleatoria. Una familia de variables aleatorias $\{X(t), t \in T\}$ se llama un proceso estocástico. Habiendo hecho el supuesto que la serie de tiempo observada $\{X(t), t \in T\}$ es una observación (o una realización) de un proceso estocástico $\{X(t), t \in T\}$, la estadística clásica del análisis de series de tiempo intenta inferir desde la serie de tiempo observada, la ley de probabilidad de el proceso estocástico. El método por el cual las series de tiempo trata este problema es similar en esencia (aunque requiere una técnica analítica más complicada) al método por el cual la estadística clásica trata el problema de inferir la ley de probabilidades de una variable aleatoria X , en donde se tiene un número finito de observaciones independientes

$$X_1, X_2, \dots, X_n$$

Hay dos aspectos en el estudio de series de tiempo -análisis y modelado. En orden para analizar una serie de tiempo $\{X(t), t \in T\}$ se debe suponer primero un modelo para $\{X(t), t \in T\}$ para el cual está completamente especificado, excepto para los valores de cierto parámetros (θ 's y ϕ 's) los cuales uno procede a estimar en base a una muestra observada. El objetivo del análisis es resumir las propiedades de una serie y caracterizar sus rasgos notorios. Esto se puede hacer por medio de dos enfoques: el dominio del tiempo o el dominio de las frecuencias. en el dominio del tiempo la atención está enfocada en la relación entre las observaciones hechas a diferentes puntos en el tiempo, mientras en el dominio de las frecuencias son los movimientos cíclicos los cuales se estudian.

La razón principal para modelar una serie de tiempo es permitir pronósticos de valores futuros. La característica que distingue a un modelo de series de tiempo es que no intenta relacionar a $X(t)$ con otras variables. Los movimientos en $X(t)$ se

explican solamente en términos de su propio pasado, o por su posición en relación al tiempo. Los pronósticos se harán entonces por extrapolación.

1.6.2. Naturaleza de los modelos de series de tiempo

Si las observaciones de una serie fueron independientes entre sí, el mejor pronóstico de la siguiente observación en la serie $X(T+1)$ simplemente sería μ . Sin embargo, las observaciones $X(t)$ claramente no son independientes. Cada una tiende a tener un valor el cual se relaciona más con aquellas observaciones inmediatamente adyacentes que con aquellas que están más alejadas. Este tipo de estructura se conoce como la correlación de la serie. Tomando en cuenta el patrón de correlación de la serie, que es común en las series de tiempo, mejores predicciones de observaciones futuras se pueden obtener.

La aproximación estadística para pronosticar se basa en la construcción de un modelo. El modelo define un mecanismo que se considera como capaz de producir las observaciones.

El modelo auto-regresivo de primer orden (AR(1))

$$X(t) = \theta X(t-1) + a(t)$$

es un ejemplo simple de un proceso estocástico. La incertidumbre se produce de la variable $a(t)$. Esto es un disturbio o ruido puramente aleatorio en el sentido de que la correlación entre dos valores cualquiera en diferentes puntos del tiempo es cero. Las demás características del modelo está determinadas por el parámetro θ que se estima por regresión ordinaria de mínimos cuadrados. Dada esta estimación la siguiente observación se pronostica por

$$X(T+1/T) = \hat{\theta} X(T)$$

Restringiendo a θ a los valores entre -1 y 1 en $X(t)$ significa que el proceso es estacionario. Cuando una serie de observaciones se generan por un proceso estacionario, fluctúan alrededor de un nivel constante y no hay tendencia en su

propagación a aumentar o disminuir sobre el tiempo.

Valores retrasados adicionales $X(t-2)$, $X(t-3)$, ... , se pueden añadir a $X(t) = \theta X(t-1) + a(t)$, con lo cual habilita a patrones más complicados de dependencia a ser modelados. En realidad, las propiedades de casi cualquier serie de tiempo estacionaria se puede reproducir introduciendo un número de retrasos suficientemente grande. La desventaja de modelar una serie de esta forma es que cuando un gran número de valores retrasados se necesitan, un gran número de parámetros se estiman. La solución a este problema es ampliar la clase de modelos para permitir valores retrasados en la variable observada y el término del ruido se conoce como proceso auto-regresivo - medias móviles (autoregressive moving average) ARMA. Tales procesos tienen un papel central en el modelado dinámico porque permiten una representación parsimoniosa de una serie de tiempo estacionaria, es decir, un modelo con relativamente pocos parámetros se puede construir.

La noción de estacionariedad es fundamental al análisis de series de tiempo. Los procesos estacionarios ARMA toman la base sobre la cual modelos más generales se pueden construir. El término ruido en cualquier modelo de regresión que usa datos de series de tiempo se pueden concebir como un proceso estocástico estacionario.

Un proceso es débilmente estacionario si:

$$E [X(t)] = \mu \quad \forall t$$

$$\text{Var} [X(t)] = \sigma^2 \quad \forall t$$

$$\begin{aligned} \text{Corr} [X(t), X(t+s)] &= \text{Corr} [X(t-s), X(t)] = \\ &= \frac{E [(X(t) - \mu) (X(t+s) - \mu)]}{\sigma^2} \\ &= \frac{E [(X(t-s) - \mu) (X(t) - \mu)]}{\sigma^2} \\ &= \rho_s \quad \forall t. \end{aligned}$$

Función Monto de Reclamos

1.1. El proceso de riesgos. Fundamentos.

La definición adecuada de riesgo ha sido ampliamente discutida en los círculos del seguro. El actuario acostumbrado al método axiomático está en ventaja en esta discusión, porque caracteriza al riesgo no "por lo que es", sino por las propiedades que tiene. Las características principales de un riesgo residen en sus propiedades como son:

- Futuro: por ocurrir
- Incierto: no se conoce el cuando
- Fortuito: que ocurra por casualidad
- Productor de reclamos
- Pagador de primas

El riesgo debe ser medible y tasable, para que se considere como pagador de primas, por medio de estadísticas, los grandes números y las probabilidades o por un volumen suficientemente grande de casos similares.

La estructura financiera de una compañía de seguros depende de los costos de administración e inversión del capital además del aspecto de los reclamos que está sujeta a una fluctuación aleatoria.

Por esto un riesgo se puede describir por un par funcional (P_t, S_t) donde:

P_t es la prima ganada en el intervalo de tiempo $(0, t)$

S_t es la suma de monto de reclamos ocurridos en $(0, t)$

Tanto P_t como S_t pueden ser funciones aleatorias (procesos estocásticos) o funciones no dependientes del azar. Normalmente se piensa a P_t como independiente del azar y a S_t estocástico. En el caso de que las primas dependan de la experiencia que haya en reclamaciones (estadísticas), P_t se vuelve un proceso estocástico.

La diferencia $P_t - S_t$ de las dos funciones definidas es esencial para el análisis matemático del riesgo. Se debe suponer siempre que P_t es una función determinística. Si P_t depende explícitamente de la experiencia del reclamo en el tiempo t (i.e.

sobre todo S_t , donde $t \leq t$ y por lo tanto es estocástico, entonces se definen nuevas funciones.

$$P^*(t) = E [P(t)]$$

$$S^*(t) = S(t) - P(t) + E [P(t)]$$

donde E representa la operación de valor esperado. Por lo tanto se tiene:

$$P^*(t) - S^*(t) = P(t) - S(t)$$

donde $P^*(t)$ es ahora determinístico, i.e. no depende más del azar. Se llama a P_t la función prima del riesgo.

El proceso aleatorio S_t se llama el proceso de reclamos acumulados. (Este proceso se refiere frecuentemente al proceso de riesgo). Usualmente se descompone en un proceso de número de reclamos N_t , el cual da el número de reclamos que pueden ocurrir al tiempo t , y un monto de reclamo aleatorio Y_t relacionado a un reclamo ocurrido al tiempo t . Los reclamos acumulados al tiempo t se pueden representar como una suma aleatoria

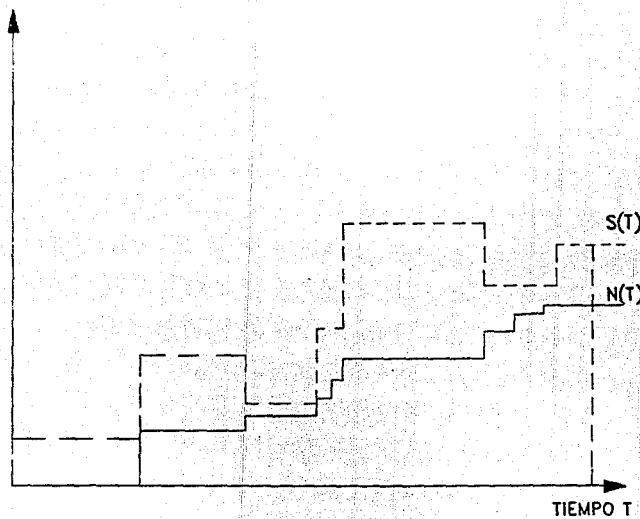
$$S_t = \sum_{t_i \in A_t} Y_{t_i}$$

donde A_t se entiende como el conjunto aleatorio de tiempo al cual los reclamos ocurren.

La siguiente figura ilustra lo que se ha dicho.

U.N.A.M.
E.N.E.P ACATLAN
PROCESO DE RIESGO

RECLAMOS



1.2. El Proceso del Número de Reclamos.

Aquí se ven dos modelos matemáticos que tratan de encontrar la función de probabilidad del número de reclamos que surgen en un riesgo colectivo, i.e. una función $p[k;t]$ que de la probabilidad de que el número de reclamos N_t en el tiempo t sea igual a k :

$$p[k;t] = \text{prob} [N_t = k]$$

El primer modelo hace la suposición más débil y más realista de que $\{N_t; t \geq 0\}$ es un proceso de Markov. Intuitivamente se hace la suposición de que el número total de reclamos determinados en el pasado tienen un papel importante en cuanto al número de reclamos futuros sucedan; no es significativo, sin embargo, donde han ocurrido estos reclamos en el pasado.

Se describe la ley de probabilidad del proceso del número de reclamos $\{N_t; t \geq 0\}$ por la probabilidad de transición

$$P_{n,n+k}(t, t+h)$$

que es la probabilidad de que k reclamos entre t (exclusivo) y $t+h$ (inclusivo) ocurran dado que n reclamos han ocurrido ya al punto de tiempo t (inclusivo).

Directamente de la naturaleza del proceso del número de reclamos y del proceso puro de nacimiento se hacen las siguientes suposiciones

a) $P_{nm}(t, t+h) = 0$ para $m < n$

b) $\lim_{h \rightarrow 0} \frac{P_{n,n+k}(t, t+h)}{h} = q_{n,n+k} = \frac{d}{dh} P_{n,n+k}(0) = 0 \quad k \geq 2$

c) $\lim_{h \rightarrow 0} \frac{P_{n,n+1}(t, t+h)}{h} = \lambda_{n+1}(t) = q_{n+1}$

d) $\lim_{h \rightarrow 0} \frac{1 - P_{nn}(t, t+h)}{h} = q_n = -\frac{d}{dh} P_{nn}(0) = -\lambda_n$

En palabras la suposición a) quiere decir que el proceso del número de reclamos nunca es decreciente, siempre está en continuo crecimiento, la probabilidad de que el proceso tenga un número de reclamos negativos es nula.

Las relaciones b) y c) significan intuitivamente que para pequeños intervalos de tiempo la probabilidad de que un reclamo ocurra es proporcional a la longitud de éstos intervalos y las funciones de muestreo del proceso del número de reclamos tienen solamente saltos positivos de monto 1. A estas relaciones se las conoce como intensidades de transición que se definen en términos de derivadas en 0 de la función de transición de probabilidad.

La función λ_n se llama la intensidad de frecuencia del n-ésimo reclamo (transición de n-1 a n reclamos). Conociendo éstas intensidades se permite calcular todas las probabilidades de transición usando el sistema de ecuaciones diferenciales prospectivas de Kolmogorov:

$$\frac{\partial}{\partial t} p_{jk}(s, t) = \sum_i p_{jk}(s, t) a_{ik}(t) = -q_k(t) + \sum_{i \neq k} p_{ji}(s, t) q_k(t)$$

Este sistema de ecuaciones diferenciales se obtiene en base a la multiplicación de matrices $P(t)$ y $A(t)$, donde los valores de estas matrices satisfacen las hipótesis antes presentadas. Por lo tanto:

$$P(t) = \begin{bmatrix}
 P_{00}(t) \dots P_{0,m-k-1}(t) & P_{0,m-k}(t) & P_{0,m-k+1}(t) & \dots & P_{0,m}(t) \\
 P_{10}(t) \dots P_{1,m-k-1}(t) & P_{1,m-k}(t) & P_{1,m-k+1}(t) & \dots & P_{1,m}(t) \\
 P_{20}(t) \dots P_{2,m-k-1}(t) & P_{2,m-k}(t) & P_{2,m-k+1}(t) & \dots & P_{2,m}(t) \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 P_{m-k-10}(t) \dots P_{m-k-1,m-k-1}(t) & P_{m-k-1,m-k}(t) & P_{m-k-1,m-k+1}(t) & \dots & P_{m-k-1,m}(t) \\
 P_{m-k,0}(t) \dots P_{m-k,m-k-1}(t) & P_{m-k,m-k}(t) & P_{m-k,m-k+1}(t) & \dots & P_{m-k,m}(t) \\
 P_{m-k+1,0}(t) \dots P_{m-k+1,m-k-1}(t) & P_{m-k+1,m-k}(t) & P_{m-k+1,m-k+1}(t) & \dots & P_{m-k+1,m}(t) \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 P_{m0}(t) \dots P_{m,m-k-1}(t) & P_{m,m-k}(t) & P_{m,m-k+1}(t) & \dots & P_{m,m}(t)
 \end{bmatrix}$$

de acuerdo con la suposición a) $P_{nm}(t, t+h) = 0$ para $m < n$, se obtendrá una matriz triangular superior:

$$P(t) = \begin{bmatrix}
 P_{00}(t) \dots P_{0,m-k-1}(t) & P_{0,m-k}(t) & P_{0,m-k+1}(t) & \dots & P_{0,m}(t) \\
 0 \dots P_{1,m-k-1}(t) & P_{1,m-k}(t) & P_{1,m-k+1}(t) & \dots & P_{1,m}(t) \\
 0 \dots P_{2,m-k-1}(t) & P_{2,m-k}(t) & P_{2,m-k+1}(t) & \dots & P_{2,m}(t) \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 \dots P_{m-k-1,m-k-1}(t) & P_{m-k-1,m-k}(t) & P_{m-k-1,m-k+1}(t) & \dots & P_{m-k-1,m}(t) \\
 0 \dots 0 & P_{m-k,m-k}(t) & P_{m-k,m-k+1}(t) & \dots & P_{m-k,m}(t) \\
 0 \dots 0 & 0 & P_{m-k+1,m-k+1}(t) & \dots & P_{m-k+1,m}(t) \\
 \vdots & \vdots & \vdots & \ddots & \vdots \\
 0 \dots 0 & 0 & 0 & \dots & P_{mm}(t)
 \end{bmatrix}$$

También se supone que se tiene una matriz $A(t)$ que cumple con las condiciones b), c), d):

$$A(t) = \begin{bmatrix} -\lambda_0 & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & & & & & & \\ 0 & \dots & -\lambda_{m-k-1} & \lambda_{m-k-1} & 0 & \dots & 0 \\ 0 & \dots & 0 & -\lambda_{m-k} & \lambda_{m-k+1} & \dots & 0 \\ 0 & \dots & 0 & 0 & -\lambda_{m-k+1} & \dots & 0 \\ \vdots & & & & & & \\ \vdots & & & & & & \\ 0 & \dots & 0 & 0 & 0 & \dots & -\lambda_m \end{bmatrix}$$

$$\frac{\partial}{\partial t} P(t) = P(t) A(t)$$

$$\frac{\partial}{\partial t} P_{00}(t) = -\lambda_0 P_{00}(t)$$

$$\frac{\partial}{\partial t} P_{01}(t) = \lambda_0 P_{00}(t) - \lambda_1 P_{01}(t)$$

⋮
⋮
⋮

$$\frac{\partial}{\partial t} P_{mm}(t) = -\lambda_m P_{mm}(t)$$

$$\frac{\partial}{\partial t} P_{mn}(t) = \lambda_{m-k-1} P_{m-k-1, m-k-1}(t) - \lambda_{m-k} P_{m-k-1, m-k}(t)$$

para la columna n sea $m-k = n$ y para el renglón $m-k-1 = m$

$$\frac{\partial}{\partial t} P_{mn}(t) = \lambda_{n-1} P_{m, n-1}(t) - \lambda_n P_{mn}(t)$$

Este sistema de ecuaciones diferenciales, aunque parecen parciales, en realidad son ordinarias, ya que m, n permanecen constantes y la única que es variable es t . Por ello el sistema de ecuaciones diferenciales:

$$\frac{\partial}{\partial t} p_{mm}(t) = -\lambda_m p_{mm}(t)$$

$$\frac{\partial}{\partial t} p_{mn}(t) = \lambda_{n-1} p_{m,n-1}(t) - \lambda_n p_{m,n}(t)$$

se puede resolver en base a la solución general de una ecuación diferencial no homogénea de primer orden lineal:

Si $g(t)$ es la solución de una ecuación diferencial

$$g'(t) + \nu g(t) = h(t) \quad a \leq t \leq b$$

Además el sistema cumple con las siguientes condiciones iniciales:

$$p_{j,k} = \begin{cases} 1 & \text{si } t=s \text{ } k=j \\ 0 & \text{si } t=s \text{ } k \neq j \end{cases}$$

Para:

$$\frac{\partial}{\partial t} p_{mm}(t) = -\lambda_m p_{mm}(t) \quad \text{se observa que}$$

$$g(t) = p_{mm}(t) \quad \nu = \lambda_m \quad h(t) = 0$$

$$g(a) = g(0) = 1 \quad t=s \quad k=j$$

$$g(t) = \int_a^t \exp[-\nu(t-s)] h(s) ds + g(a) \exp[-\nu(t)]$$

sustituyendo los valores

$$g(t) = \int_0^t \exp[-\lambda_m(t-s)] (0) ds + \exp[-\lambda_m t] = \exp[-\lambda_m t]$$

Para:

$$\frac{\partial}{\partial t} P_{mn}(t) = \lambda_{n-1} P_{m,n-1}(t) - \lambda_n P_{m,n}(t)$$

$$g(t) = P_{mn}(t) \quad \nu = \lambda_n \quad h(t) = \lambda_{n-1} P_{m,n-1}(t)$$

$$g(a) = g(0) = 0 \quad t=s \quad k \neq j$$

sustituyendo los valores

$$\begin{aligned} g(t) &= \int_0^t \exp[-\lambda_n(t-s)] \lambda_{n-1} P_{m,n-1}(s) ds = \\ &= \exp[-\lambda_n t] \int_0^t \exp[\lambda_n s] \lambda_{n-1} P_{m,n-1}(s) ds \end{aligned}$$

La condición de estandarización

$$\sum_m P_{nm}(t,s) = 1 \quad \forall \quad n, t, s \geq t$$

no se cumple en las soluciones de este sistema de ecuaciones diferenciales.

Quando los coeficientes λ_n se incrementan rápidamente sucede que

$$\sum_m P_{nm}(t,s) < 1$$

El que los coeficientes λ_n se incrementen rápidamente indica que muchos reclamos pueden ocurrir en un intervalo de tiempo finito (VER FELLER VOL. I PAG. 447-449).

A partir de que

$$\sum_m p_{nm}(t,s) \leq 1$$

se observan dos casos diferentes

a) $\sum_m p_{nm}(t,s) = 1$ caso normal

b) $\sum_m p_{nm}(t,s) < 1$ caso de una explosión de número de reclamos

Para que no se tenga un "proceso divergente de nacimiento" o un "proceso patológico", es decir, que $\sum_m p_{nm}(t,s) < 1$ se tiene que cumplir que la serie $\sum_n \lambda_n^{-1}$ diverja.

En base a las funciones de transición de probabilidad

$$p_{mm}(t) = \exp[-\lambda_m t]$$

$$p_{mn}(t) = \exp[-\lambda_n t] \int_0^t \exp[\lambda_n s] \lambda_{n-1} p_{m,n-1}(s) ds$$

se puede obtener la función de transición de probabilidad $\{N(t), t \geq 0\}$ con una tasa de crecimiento constante $\lambda_n = \nu$ dada por:

$$p_{mn}(t) = \exp[-\nu t] \begin{cases} \frac{(\nu t)^{n-m}}{(n-m)!} & \text{para } n \geq m \\ 0 & \text{en otro caso} \end{cases}$$

Esta función de probabilidades, que es el proceso de Poisson, indica que un salto desde el estado m en t hasta el estado n en s significa que ocurrieron $n - m$ reclamos entre s y t .

Se debe observar que el proceso puede comenzar en cualquier estado m y cumplir con las condiciones

$$p_m(0) = 1 \quad p_n(0) = 0 \quad n \neq 1$$

El problema que se ha de tratar de resolver se puede hacer de distintas formas. Un método es comenzar considerando el portafolio en cuestión como integrado por un número de pólizas individuales, cada una de las cuales tiene una probabilidad cierta de reclamo (e.g. en el seguro de vida se supone que la

probabilidad de que una persona con edad x fallezca dentro de un año es q_x). Entonces el número total de reclamos es la suma de las contribuciones de las pólizas individuales y las probabilidades $p(k;t)$ se pueden obtener por medio del teorema de adición del cálculo de probabilidad de las probabilidades primarias. Básicamente las probabilidades son binomiales pero llevar a cabo ésta suma de una manera rigurosa conduce a cálculos muy complicados y envuelve algunas suposiciones restrictivas.

Una aproximación alternante, que ha conducido a un desarrollo provechoso, es según el método colectivo adoptado por Lundberg. En este método la estructura de la póliza individual no es tomada en cuenta y en vez el portafolio se considera como un todo, i.e. un "proceso" es tomado en cuenta en los casos los cuales el número de eventos (reclamos) son registrados y cuando ninguna atención se toma a pólizas particulares de las cuales los reclamos surgen. Comenzando con algunas condiciones generales que el proceso aleatorio tiene que obedecer, se puede deducir que el proceso toma la bien conocida forma de Poisson.

En este proceso, el del número de reclamos, se supone que cumple con las siguientes condiciones, en base a la teoría de riesgo colectivo.

i) Los eventos que ocurren en dos intervalos de tiempo diferentes son independientes (Independencia de incrementos).

ii) El número de eventos en un intervalo de tiempo (t_1, t_2) es dependiente solamente de la longitud del intervalo $t = t_2 - t_1$ y no del valor inicial t_1 (Incrementos estacionarios).

iii) La probabilidad de que más de un evento ocurrirá al mismo tiempo y la probabilidad de que un número infinito de eventos ocurran en algún intervalo de tiempo finito son ambas cero. (Exclusión de eventos múltiples)

En base a estas suposiciones y aplicando el teorema I o II de la introducción se ve (después de cambiar al tiempo operacional) que N_t es un proceso de Poisson, i.e.

$$P_{mn}(t) = P(k;t) = P_k(\lambda t) = P[N_1=k] = \exp[-\lambda t] \frac{(\lambda t)^k}{k!}$$

para cada $t > 0$, donde $\lambda \geq 0$ es un parámetro indicando el promedio de número de reclamos en una unidad de tiempo. El acontecimiento de un evento de reclamo depende del número de casos (unidades de riesgo) expuestos al riesgo y también a la intensidad del riesgo, i.e. la oportunidad de que un caso particular de lugar a una reclamación. El proceso de Poisson surge como un producto de estas componentes.

Discusión de supuestos

a) Incrementos independientes.- Este supuesto significa que un evento (e.g. fuego) no puede dar lugar a cualquier otro evento o reclamo (eliminación de las reacciones en cadena). En la práctica, sin embargo, un fuego frecuentemente puede dar lugar de un riesgo a otro (e.g. explosión) en contradicción con esta condición.

El supuesto de incrementos independientes se puede cumplir por medio de la definición de una unidad de riesgo, como es costumbre en la práctica del reaseguro, como una combinación de todos aquellos riesgos que se parecen, entre los cuales la contaminación o contagio es posible (e.g. toda propiedad en una construcción sin tomar en cuenta si está formalmente asegurada por una o varias pólizas o si tiene un dueño o varios). De la misma forma un barco y su carga se consideran como una unidad de riesgo. Sin embargo, no siempre es posible construir unidades de riesgo de tal manera que la contaminación externa no ocurra. Tal es el caso con las enfermedades contagiosas en el seguro de accidentes y enfermedades o epidemias en el seguro de vida. La función de Poisson no se puede aplicar, al menos no sin algunas modificaciones.

b) Incrementos estacionarios.- Este supuesto da a entender que el flujo colectivo de los reclamos es estacionario, i.e. no es constantemente creciente o decreciente ni tampoco más oscilatorio

que aquel producido por una fluctuación aleatoria normal. En otras palabras, la intensidad de los reclamos es constante. Este es el caso usual en el seguro, particularmente durante periodos cortos, cuando el número de pólizas u otras circunstancias no están sujetas a cambios marcados. Esto implica que el portafolio es tan grande que la salida de las pólizas individuales por razones de reclamos o por otras causas y la entrada de nuevos casos no puede afectar el flujo colectivo de los eventos a algún nivel significativo.

Frecuentemente, sin embargo, hay situaciones donde la estacionariedad no se aplica estrictamente, por ejemplo, puede haber variaciones estacionales en las intensidades de reclamos. Entonces el intervalo de tiempo referido puede ser dividido en subintervalos de tal forma que el correspondiente subproceso tenga (al menos aproximadamente) intensidades constantes y por lo tanto sean Procesos de Poisson. En consecuencia, el número total de reclamos durante todo el intervalo se distribuye como Poisson, debido a que la suma de variables Poisson independientes es de nuevo una variable Poisson. Así que, si solamente el número total y su comportamiento son de interés, entonces las variaciones estacionales se pueden no tomar en cuenta, con tal que los cambios en las intensidades de reclamos son determinísticos, i.e. las variaciones estacionales recurren de tal forma que su predicción es posible a través de la experiencia.

Lo anterior también es aplicable para casos en los cuales las intensidades están cambiando de acuerdo con alguna tendencia, en el sentido de que sea determinísticamente predecible. El número total de reclamos permanece todavía distribuida como Poisson pero el parámetro λt se calcula como una suma de los parámetros relacionados a los subintervalos. Otro método que se utiliza con frecuencia para obtener el mismo resultado es introducir el concepto de "tiempo operacional".

Sin embargo, hay otras circunstancias que surgen en la práctica donde los supuestos de incrementos independientes y estacionarios no se pueden cumplir. Por ejemplo, el seguro contra incendio puede ser grandemente afectado por las condiciones meteorológicas y un periodo largo, seco y soleado puede dar lugar

a numerosos fuegos anormales; en algunos países, los huracanes u otras catástrofes naturales pueden provocar un enorme incremento en los reclamos. También es conocido que las condiciones económicas tienen una influencia considerable en muchos de las ramas de la operación del seguro de daños. Tiempos de explosiones o recesiones económicas da lugar a un aumento o decrecimiento considerable en el número de accidentes por tráfico o de trabajo, así como también influye en el negocio de seguro de crédito. Tales circunstancias son tan generales que la aplicación de la función Poisson elemental está grandemente limitada, y así hay una necesidad de desarrollar la teoría omitiendo las condiciones de independencia y estacionariedad. En vista de estas limitaciones, la función Poisson da al menos una primera buena aproximación, particularmente para intervalos de tiempo pequeños. Es la base de distribuciones más generalizadas. Aún más, el riesgo de cambios y de variaciones que afecten la estacionariedad pueden ser tratadas simplemente añadiendo un monto precautorio al parámetro λ .

c) Exclusión de eventos múltiples.- En primera instancia parecería que este supuesto no siempre es válido. Por ejemplo, en el seguro automóviles, dos vehículos pueden chocar, dando lugar a un doble evento. Incidentes similares pueden ocurrir en el seguro marítimo y en algunas otras ramas. Este obstáculo puede ser resuelto por un adecuado cambio de definición, por ejemplo considerando el caso de que el choque entre dos automóviles es un sólo reclamo. Esto significa que la suma de ambas partes se usa cuando se construye las estadísticas de la distribución del tamaño de un reclamo, el cual se considera aparte como otra variable aleatoria.

Como una aplicación práctica del proceso de Poisson en el seguro se darán los siguientes ejemplos.

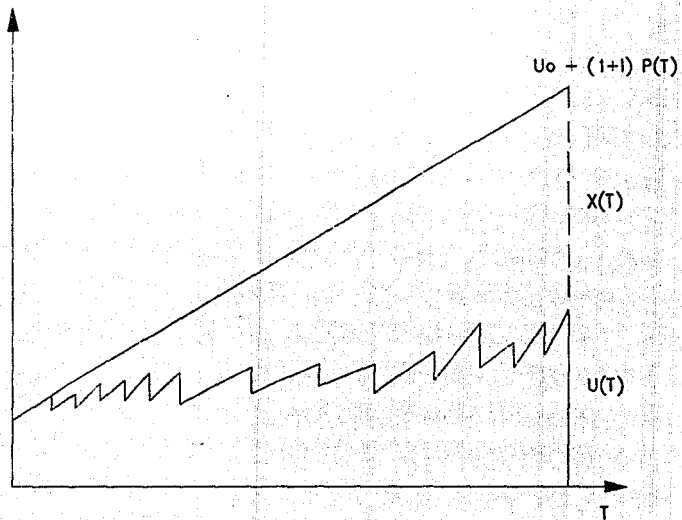
Una sociedad mutualista tiene 1000 miembros. En el caso de que una muerte ocurra una suma fija de $S = \$ 1000$.- se paga. El valor medio de la tasa de mortalidad es 0.01, la prima $P_p = (1 + \rho) S E [k]$, donde $\rho = 0.1$ es un margen de seguridad y k es el número de fallecimientos. El estado actuarial de la

sociedad se examina cada año. ¿Qué tan grande una reserva de seguridad U_0 debe tener la sociedad para estar segura, a un nivel de probabilidad del 99%, que el balance no muestre algún déficit?

Si se grafica como se comporta las primas, las reservas y los reclamos se obtendría lo siguiente:

U.N.A.M.
E.N.E.P ACATLAN
RESERVA DE RIESGO

RESERVA DE RIESGO $U(T)$



La prima de riesgo P junto con un margen de seguridad ρ está continuamente creciendo, esta se acumula en una reserva de riesgo U de una cantidad inicial U_0 , así que el ingreso se representa como una línea ascendente a la derecha. Los reclamos que se consideran como un ingreso negativo, se pagan de ésta reserva, y se representa como escalones hacia abajo. La diferencia

$$U(t) - U_0 = P(1+\lambda)t - X(t)$$

da la ganancia neta (positiva o negativa) que surge durante el tiempo t .

En base a ésta diferencia se calculará la reserva inicial U_0 , para que a un nivel de significancia del 99% el balance no muestre algún déficit. Ahora se calcula $E[k]$, el número esperado de fallecimientos

$$E[k] = \lambda t N = .01 \times 1 \times 1000 = 10$$

El número esperado de fallecimientos durante el transcurso de ese año es igual a 10, este valor se obtiene de multiplicar la tasa de mortalidad promedio, que es de 0.01, la longitud del intervalo que es de un año, igual a una unidad, y el número de miembros de la sociedad que es de 1000.

Una vez calculado el número de muertos promedio durante el año, se puede calcular la prima

$$P_\rho = (1 + \rho) E[k] S = 1.1 \times 10 \times 1000 = 11,000$$

dónde $\rho = 0.1$ y $S = 1000$, efectuando operaciones, se obtiene el valor de la prima que se ha de juntar entre todos los miembros de la sociedad, para obtener el beneficio de \$ 1000.- por miembro de la sociedad que fallezca.

En base al planteamiento del problema, se nota que

$U(t) = U(1) =$ reserva al final del año $= 0$ y

$X(t) = X(1) =$ monto total de reclamos al final del año $= kS =$ número de fallecimientos esperado por la suma fija de dinero que se paga por cada fallecimiento, se tiene que:

$$U(t) - U_0 = P(1 + \rho)t - X(t)$$

$$0 \leq U_0 + P(1 + \rho)t - X(t)$$

$$0 \leq U_0 + P(1 + \rho)t - kS$$

$$P[U_0 + P(1 + \rho)t - kS \geq 0] \geq .99$$

$$P[kS \leq U_0 + P(1 + \rho)t] \geq .99$$

$$P\left[k \leq \frac{U_0 + P(1 + \rho)t}{S}\right] \geq .99$$

Cómo $k \sim P(10)$ entonces, buscando en tablas o realizando operaciones se nota que:

$P[k \leq 17] = .986$ y $P[k \leq 18] = .993$, tomando a $k = 18$ entonces:

$$U_0 = X(t) - P_\rho = kS - P_\rho = 18(1000) - 11000 = 7000$$

El valor de la reserva inicial para que el balance no muestre algún déficit debe ser de \$ 7000.-

¿Cuántos miembros debe tener la sociedad del ejemplo anterior para que no sea necesaria la reserva de seguridad bajo las condiciones mencionadas?

Se nota en este ejemplo que el total de primas que aportan los miembros deben de ser en una cantidad suficiente, de tal manera que éstas alcancen a cubrir el número esperado de reclamos, es decir:

$$P_t = X_t$$

También se nota que se utiliza el proceso de Poisson para encontrar el número de miembros, sería muy tedioso realizar todas las operaciones, por lo cual se utiliza el teorema del Límite

Central para variables distribuidas como Poisson. Este teorema explica que cuando el número de casos es muy grande ($n \rightarrow \infty$) entonces la función de distribución Poisson tiende a la función de distribución Normal con media = $E[k] = \lambda N t$ y desviación estándar $\sigma = \sqrt{\lambda N t}$, es decir, $F(k) \approx N \left[\frac{k - E[k]}{\sigma} \right]$

En este caso $N =$ número de miembros se desconoce, entonces
 $\mu = E[k] = \lambda N t = 0.01 N$ y $\sigma = \sqrt{\lambda N t} = \sqrt{0.01 N}$

Se desea que $P \left[\frac{k - E[k]}{\sigma} \right] = .99$, como en este caso N es muy grande y utilizando el teorema del Límite central se desea obtener el valor cuantil Z correspondiente, de tal manera que $P[Z] = 0.99$, buscando en tablas se observa que $P[Z \leq 2.33] = 0.99$, entonces $z = 2.33$.

Ahora de la ecuación $P_i = X_i$ y del Teorema del Límite Central se procede a calcular N :

$P_i = X_i$	$P \left[\frac{k - E[k]}{\sigma} \leq z \right] = P[Z \leq z] = P[Z \leq 2.33]$
$(1.1)(0.01) NS = kS$	$\frac{k - E[k]}{\sigma} = \frac{0.01N - 0.01N}{\sqrt{0.01N}} = 2.33$
$0.011N = k$	$0.001 N = 0.233 \sqrt{N}$
$E[k] = 0.01 N$	$0.000001 N^2 = 0.054289 N$
$\sigma = \sqrt{0.01N}$	$N = 54,289$

Se obtiene el valor de $N =$ número de miembros que requiere la sociedad para que el monto total de las primas obtenidas sirvan para cubrir el monto total de reclamaciones y no se requiera de una reserva inicial.

Por último se verá como se aplica el Proceso de Poisson en el reaseguro, para determinar el monto esperado de reclamos.

Una sociedad mutualista ofrece gastos de funeral cuando ocurre la muerte de un miembro, siendo éste beneficio otorgado por medio de una suma fija de \$ 100.-. El número promedio de reclamo

$\lambda t = 1$. La sociedad tiene un reaseguro por pérdida estacionaria de acuerdo con el cual, si el número de muertes es mayor de dos, el reasegurador paga el tercero y demás fallecimientos. ¿Cuál es la prima del riesgo (= monto esperado de reclamos) para el reaseguro?

En la introducción se vio que

$$E[g(s)] + E[S - g(s)] = E[S]$$

Donde $E[S - g(s)]$ es el monto esperado de reclamos para el reaseguro. Utilizando este principio, se obtiene que:

$$\frac{P}{1000} = \sum_{k=3}^{\infty} (k-2) P_k ; \quad P_k = e^{-1} \frac{1^k}{k!}$$

$$\frac{P}{1000} = \sum_{k=3}^{\infty} (k-2) e^{-1} \frac{1^k}{k!} = e^{-1} \left[\sum_{k=3}^{\infty} \frac{k}{k!} - 2 \sum_{k=3}^{\infty} \frac{1}{k!} \right]$$

$$= e^{-1} \left[\sum_{k=3}^{\infty} \frac{1}{(k-1)!} - 2 \sum_{k=3}^{\infty} \frac{1}{k!} \right]$$

Haciendo operaciones, se ve que el monto esperado P es igual a \$ 103.36.-

1.3. El Proceso de Reclamos Acumulados.

Definiciones.- Ahora se extiende la consideración de proceso del número de reclamos a procesos que opera los reclamos acumulados, refiriéndose tanto a los reclamos individuales como sus sumas, los reclamos acumulados. Un bloque constructor primario es el tamaño variante aleatorio Z de un reclamo individual, i.e. la suma que paga el asegurador al momento de que suceda un incendio, accidente o cualquier otro evento contra el cual se asegura. Se supone que los tamaños del reclamo Z son mutuamente independientes e idénticamente distribuidas, teniendo una función de densidad:

$$S(Z) = P [Z \leq z]$$

$$1 e^{-x} = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

De acuerdo a la aproximación colectiva que se vio anteriormente no se toma en cuenta la unidad de riesgo (póliza) de la cual el reclamo surge. La función de densidad S describe la variabilidad de tamaños del flujo continuo de reclamos. En la sección de discusión de supuestos sugiere que los pagos debidos a uno y al mismo evento deben unirse como un reclamo sin considerar que los pagos formalmente conciernen a diferentes pólizas (e.g. diferentes dueños de un propiedad en un complejo industrial dañado o en un vehículo). Aún más, si dos o más reclamos de la misma unidad de riesgo se consideran como reclamos diferentes si se provocan por diferentes eventos.

La variable del tamaño del reclamo introduce al proceso referido a un nuevo estrato de estocasticidad en adición a la variación del número de reclamos. El proceso que se constituye por una variación estocástica tanto en el número de reclamos como el tamaño del reclamo se llama proceso compuesto.

La existencia de una función S está en concordancia con la experiencia general, al menos en la medida que se consideren periodos de longitud moderada y estipulando que los efectos de cambios en valores monetarios se eliminan. Los reclamos actuales pueden ser registrados y estimaciones numéricas se pueden obtener para S . Por ahora se supondrá que la función existe y es conocida. Posteriormente los detalles de su cálculo práctico se considerarán.

Reclamos pendientes.- Para mantener los modelos teóricos del riesgo dentro de dimensiones reales, usualmente se supone que los reclamos se pagan inmediatamente en el momento en que ocurren. En la práctica, inevitablemente hay un retraso de tiempo entre el ocurrimiento del evento que da lugar a un reclamo y su liquidación, ya sea por aspectos menores de administración o por problemas legales tales como la determinación de la suma asegurada, o por retrasos en notificar el siniestro a la Cia aseguradora, por lo tanto, en adición a los problemas de fijar el monto probable de reclamos que han sido notificados, alguna concesión debe hacerse para los casos esperados que se notifican tarde.

En principio, cualquier error que se estime de los reclamos pendientes se corregirán finalmente cuando los reclamos son finalmente establecidos, pero puede dar lugar a una serie de beneficios o pérdidas entre algunos años y en esa forma afectar también a la fluctuación de riesgo, sin embargo, serán de mucho menor tamaño que las fluctuaciones ordinarias. Se deberá notar que el supuesto relacionado al pago inmediato de reclamos no funciona en la eliminación de la estimación de errores de los reclamos pendientes del modelo. Como cualquier otra inexactitud del supuesto básico, da lugar a fluctuaciones extras en los resultados asegurados, probablemente de una manera periódica. Cuando los parámetros del modelo se calibran sobre la base de fluctuaciones actuales observadas, el efecto de estas inexactitudes serán tomadas en cuenta automáticamente. Aún más, no hay algún obstáculo esencial a la introducción de reclamos pendientes o su error de estimación como una entrada particular al modelo.

Del otro lado, sobre o sub estimaciones sistemáticas, pueden dar lugar a sesgos considerables en la hoja de balance y por lo tanto en la valuación de los márgenes de solvencia actuales (reservas de riesgo). La consideración de estos, así como muchas clases de aspectos no estocásticos, e.g. riesgos incalculables arriesgando la existencia de los aseguradores tales como fallas mayores en inversiones o evaluación del riesgo, malos manejos administrativos, etc. son partes esenciales de la solvencia general de la industria aseguradora.

Los análisis del desarrollo de las estimaciones de reclamos es una parte normal de la rutina del negocio e indicará la necesidad de algunas provisiones extras, se puede pensar, por ejemplo, que un margen se necesita tratar con los cambios inflacionarios con respecto tanto al nivel promedio de inflación así como la necesidad de medidas de emergencia en los casos cuando la tasa de inflación pueda ocasionalmente irse muy arriba. Tales ajustes se pueden necesitar para fijar los parámetros del modelo completo y tendrán que ser tratados en circunstancias individuales.

Distribución Compuesta del reclamo agregado.

Una cartera asegurada se considera de nuevo y se desea encontrar la distribución de probabilidad del monto total de reclamos X , o brevemente el reclamo agregado, que ocurren durante un intervalo de tiempo. La probabilidad p_k de que el número de reclamos sea igual a k y la función de distribución S de un reclamo, se suponen que son conocidos.

La distribución requerida $F(X)$ proporciona la probabilidad del evento $X \leq x$. Este evento puede ocurrir en las siguientes formas alternantes:

- i) En el intervalo de tiempo no ocurren reclamos
- ii) El número de reclamos = 1 y el monto de reclamos es $\leq x$
- iii) El número de reclamos = 2 y la suma de los montos de éstos es $\leq x$.
- iv) El número de reclamos = 3 y la suma de los montos de éstos es $\leq x$.

ETC.

La probabilidad condicional de que, si el número de reclamos es exactamente k , la suma de éstos k reclamos es $\leq X$ se denota por $S_k(X)$. Usando las reglas de suma y multiplicación combinadas de probabilidad, se sigue que

$$F(X) = \sum_{k=0}^{\infty} p_k S_k(x)$$

Se supone que los montos de los reclamos son mutuamente independientes, la función $S_k(x)$ se conoce bien del cálculo de probabilidades como la k -ésima convolución de la función de distribución $S(X)$, que se puede calcular de la fórmula de recurrencia:

$$S_k(x) = \int_0^x s_{k-1}(X-Z) dS(Z) = s^{(k-1)*} * S(X) = S^{k*}(X)$$

y la siguiente importante fórmula se obtiene para la función de densidad de los reclamos agregados:

$$F(X) = \sum_{k=0}^{\infty} p_k S_k(x)$$

Terminología.- La función de distribución $F(X) = \sum_{k=0}^{\infty} p_k S_k(x)$ se llama compuesta, refiriéndose al proceso compuesto que hay detrás de ella. Si el proceso del número de reclamos relacionada a X es un proceso de Poisson, entonces X se llama proceso de Poisson compuesto y la función $F(X)$ es la función de densidad de Poisson compuesta, p_k es la probabilidad de Poisson simple.

CAPITULO 2. Estimación de la función monto de reclamos acumulados por medio de las series de tiempo.

Aspectos generales.- En casi todas las teorías del riesgo es necesario conocer la función de distribución del tamaño de reclamo S más o menos aproximado. La aproximación debe estar formada en base a los datos que existen, en general, empíricos. Las Cias. aseguradoras siempre tienen archivos de datos conteniendo información detallada de tanto pólizas como de reclamos y muchas clases de estas estadísticas se producen para contabilidad, medir tasas o razones y otros propósitos. La construcción de distribuciones del tamaño del reclamo y otros datos que se necesitan para el análisis teórico del riesgo se pueden obtener directamente o por medio de algunas modificaciones como multiplicaciones de éstos procesos de datos.

Extensión de los supuestos básicos.- Se vio anteriormente que los parámetros básicos del proceso de reclamo son sujetos continuamente a alteraciones que son parcialmente reveladas como tendencias más o menos regulares y frecuentes cambios cíclicos. El efecto de éste fenómeno es tan significativo que no se pueden omitir en consideraciones de etapas largas. Los supuestos concernientes al proceso de Poisson compuesto se extienden ahora. Primero, el parámetro n (el número de reclamos esperado) se discute, y después, en secciones subsecuentes, algunas otras bases se toman en cuenta.

c) Las tendencias en el parámetro n se deben parcialmente al cambio gradual del tamaño del portafolio y se afectan también parcialmente por las alteraciones en la exposición del riesgo dentro del portafolio. Se pueden tomar en cuenta suponiendo que el parámetro del modelo n es dependiente del tiempo. Esto se puede hacer convenientemente por medio de un factor de

crecimiento, $r_g(t) = \frac{n(t)}{n(t-1)}$ o equivalentemente

$$n(t) = n \prod_{\tau=1}^T r_g(\tau)$$

donde $n = n(0)$. La razón correspondiente $r_g - 1$ se denota por i_g .

La tasa de crecimiento no necesita ser la misma de un año a otro. Sin embargo, si se consideran periodos de tiempo cortos, su constancia se puede suponer. La fórmula puede ser considerablemente simplificada en esta forma. En consecuencia, las tendencias se introducen en el modelo suponiéndolas exponenciales y poniendo

$$n(t) = n r_g^t$$

Otra aproximación puede ser usando la fórmula lineal

$$n(t) = n + r_g^t$$

Los ciclos.- Además de las tendencias hay variaciones periódicas en las frecuencias de los reclamos. Aun cuando las variaciones concernientes son frecuentemente irregulares en la práctica, estas se llaman ciclos. Se distinguen del término 'oscilación'. Está compuesta de variaciones que aparecen como ondas que se expanden sobre dos o más años. La forma más sencilla es encontrar alguna fórmula determinística para indicar las

desviaciones relativas denotada por $Z(t) = \frac{\Delta n(t)}{n(t)}$ de $n(t)$ del

flujo de la tendencia. Una forma senoidal determinística

$$Z(t) = z_m \cos(\omega t + \phi)$$

se supone como un ejemplo donde z_m es la amplitud, ϕ la fase y ω es el coeficiente de la frecuencia

$$\omega = \frac{2\pi}{T_z}$$

T_z viene siendo la longitud de la onda.

Serie de Tiempo Autocorrelativas. - Una forma más sofisticada de manejar los ciclos y las tendencias es haciendo uso de la Teoría de Series de Tiempo económicas. La variable

ciclo $Z = \frac{\Delta n}{n}$ periódicamente sube y luego baja por otro periodo hasta que otra subida ocurre.

Una aproximación simple en modelar z es suponer que su valor $z(t)$ depende del estado del proceso en los tiempos previos $t-1$, $t-2$, Si z es alto en estos tiempos, $z(t)$ es también alto, pero si durante varios años continúa creciendo, puede ser indicación de que un estancamiento está por venir. Además z está sujeta a irregularidades más o menos fuertes que se pueden suponer fluctuaciones meramente aleatorias. Estas características proveen una base para construir modelos matemáticos para describir el proceso del tipo

$$z(t) = f(z(t-1), z(t-2), \dots) + a(t) + \theta_1 a(t-2) + \dots$$

donde la función 'predictora' o 'descriptoria' f coordina la retroalimentación de la experiencia pasada con el futuro. Los términos $a(t)$, llamados ruido, representan el efecto aleatorio, que se supone está distribuido normalmente ($N(0, \sigma^2)$).

Las funciones lineales son aproximaciones satisfactorias para la ecuación anterior

$$z(t) = \theta_1 z(t-1) + \theta_2 z(t-2) + \dots + a(t) + \theta_1 a(t-1) + \theta_2 a(t-2) + \dots$$

y se conocen como procesos ARMA (Auto-Regresivo Medias Móviles). Los ciclos se pueden generar por una adecuada decisión de coeficientes como se puede demostrar analíticamente. En términos de la teoría existe una fuerte autocorrelación entre valores de variables consecutivas.

Método Montecarlo.

Debido a la dificultad que existe para poder conseguir estadísticas de reclamos en las aseguradoras, se estudiará este método, que es un modelo de simulación, para poder aplicar las series de tiempo y poder analizar y predecir como se comportan los reclamos.

El método Montecarlo es sobretodo una herramienta muy poderosa; se puede aplicar a modelos muy complicados debido a la técnica del algoritmo la cual procede paso por paso del tiempo t a $t+1$. Las reglas de transición se pueden escoger en maneras flexibles y las suposiciones referentes a tendencias, ciclos, oscilaciones de corto tiempo, función de densidad del tamaño del reclamo, reaseguro, etc., se pueden integrar al modelo actual sin la necesidad urgente de simplificar, aproximar o atenuarlos en alguna forma analítica que sea conveniente para los cálculos, que es el mayor inconveniente de la mayoría de los métodos directos.

La desventaja principal de la simulación es que solo da muestras como resultados, y esto está envuelto con las inexactitudes de la muestra, lo cual puede hacer necesario aumentar el tamaño de la muestra tanto, que el método se vuelve costoso o aun intratable. Sin embargo, puede proveer de una aproximación satisfactoria del límite estocástico y de una buena impresión de la estructura del proceso con tamaños de muestra moderados.

Números aleatorios

Números aleatorios distribuidos uniformemente. La técnica de Monte Carlo utiliza los números aleatorios, los cuales son sucesiones de variables aleatorias independientes distribuidas idénticamente o muestras (realizaciones) de tales sucesiones.

Considérese un ejemplo de una variable aleatoria r , la cual está distribuida uniformemente sobre el intervalo $(0,1)$. Su función de distribución es:

$$R(r) = \begin{cases} 0 & \text{para } r \leq 0 \\ r & \text{para } 0 < r \leq 1 \\ 1 & \text{para } r > 1 \end{cases}$$

La distribución se llama rectangular debido a la forma de su función de densidad. Una sucesión de números aleatorios distribuidos uniformemente se obtiene o es generada tomando un conjunto de variables r_1, r_2, \dots , las cuales están distribuidas uniformemente y son mutuamente independientes. Tales sucesiones se necesitan como bloques de construcción básicos para las simulaciones de Monte Carlo como se verá en lo sucesivo. varias técnicas se sugieren para su generación. Funcionan como la lotería, produciendo una lista de números ganadores, cada número dentro del intervalo (0,1) (redondeando a cierto número de decimales) debe de tener exactamente la misma probabilidad de aparecer en la sucesión. Debido a que los métodos prácticos pueden generar sucesiones los cuales solamente satisfacen aproximadamente los supuestos teóricos se llaman números pseudoaleatorios.

Porque el Método de Monte Carlo necesita un número grande de números aleatorios, entonces tienen que ser generados por una computadora y con la ayuda de algún paquete (STATGRAPH).

Números aleatorios distribuidos arbitrariamente.- Sea r una variable aleatoria distribuida uniformemente y sea F una función de distribución dada y que transforma r en una nueva variable aleatoria $X = F^{-1}(r)$. F no necesita ser estrictamente creciente, puede ser, por ejemplo, una función discreta discontinua, por lo tanto $F^{-1}(r)$ se define como el valor de X más pequeño que satisface la desigualdad $F(X) \geq r$ donde $0 < r < 1$.

Entonces

$$\text{prob} \{X \leq x\} = \text{prob} \{F^{-1}(r) \leq x\} = \text{prob} \{r \leq F(x)\} = F(x)$$

lo que significa que F es la función de distribución de X . Esta observación da (en principio) un procedimiento simple de como generar números aleatorios teniendo una función de distribución F dada.

Método Monte Carlo.- Para ilustrar la técnica de simulación del método de Monte Carlo, la convolución de dos funciones de distribución F_1 y F_2 se calcula. La idea es hacer uso del bien conocido hecho de que la suma de dos variables aleatorias

teniendo funciones de distribución F_1 y F_2 se distribuye como $F_1 * F_2$. Las primeras dos sucesiones de números aleatorios ($X_{11}, X_{12}, \dots, X_{1n}$) y ($X_{21}, X_{22}, \dots, X_{2n}$) se generan, la primera se distribuye de acuerdo a F_1 y la segunda de acuerdo con F_2 . por ejemplo la técnica presentada anteriormente se puede aplicar. Entonces poniendo $X_i = X_{i1} + X_{i2}$ se obtiene una sucesión de números aleatorios distribuidos de acuerdo con la función $F_1 * F_2$. Para obtener valores numéricos de esta función de distribución, o mejor dicho, una estimación de esta, los valores de X_i de X_i dadas por el generador de números aleatorios se reordenan de acuerdo con sus magnitudes y, denotados por $c(X)$ "el número contador" de los $X_i \leq X$, el estimado $F(X) \approx \frac{c(X)}{S}$ se obtiene inmediatamente.

Se verá que el método de simulación es, de hecho, exactamente equivalente al método de observar los valores actuales que aparecen en algún experimento y construir el estimado estadístico de la función de distribución. Ningún experimento físico se lleva a cabo, en vez de ello "se juega" o "simula" por medio de números aleatorios. Este método ha aumentado su uso en conexión con varios proyectos de investigación, particularmente cuando el cálculo directo de las funciones de distribución es complicado. También es muy útil aplicarlo en el campo de la teoría de riesgo, especialmente donde otros métodos son intratables debido a la complejidad del modelo.

Evaluación de exactitud.- Debido a que el valor de $\hat{p} = \frac{c(X)}{S}$ se obtiene como una muestra aleatoria, es sujeta a una inexactitud como cualquier otra muestra. De hecho se distribuye binomialmente, teniendo desviación estándar

$$\sigma = \sqrt{\frac{p(1-p)}{S}}$$

donde p es el valor verdadero desconocido de $F(X)$. Usando la aproximación normal, el error máximo al nivel de confianza $1 - \epsilon$

$$\Delta_s = y_\epsilon \sigma = y_\epsilon \sigma \sqrt{\frac{p(1-p)}{S}} \leq \frac{y_\epsilon}{2\sqrt{S}}$$

donde $y_\epsilon = N^{-1} \left[\frac{1 - \epsilon}{2} \right]$. El límite Δ_s depende del tamaño de la muestra S .

Simulación directa de la función de Poisson Compuesta.

La simulación para el cálculo de la función de distribución del monto de reclamo

$$F(X) = \sum_{k=0}^{\infty} P_k S^{k*}(X)$$

para una X dada, o simultáneamente para un conjunto dado X_j ($j = 1, 2, \dots$) de X valores, se puede obtener repitiendo los siguientes pasos a veces. La distribución del número y el tamaño de reclamos se suponen dados.

i) Genérese el número de reclamos k utilizando algún paquete estadístico o con la ayuda de algún programa y en base a la distribución del número de reclamos supuesta.

ii) Genérese k números Z_1, Z_2, \dots, Z_k cada uno distribuidos de acuerdo con la función de distribución dada y sumesen

$$X_{s,m} = \sum_{i=1}^k Z_i$$

En base a esta técnica y con la ayuda del paquete estadístico STATGRAPH, se crean los siguientes datos estadísticos, tomando en cuenta una tasa de crecimiento del número de reclamos del 20% y la tasa de crecimiento del monto de reclamos del 15%.

En el inciso i) se utiliza el paquete STATGRAPH para generar los números aleatorios del número de reclamos, donde se supone que están distribuidos como Poisson con $\lambda=10$, solamente se generan 12 números aleatorios, uno por cada mes del año. Al final de cada año se incrementa el número de reclamos en un 20%, es decir, para el siguiente año, $\lambda = 10 (1.2) = 12$ y así sucesivamente.

En el inciso ii), también se vale de la ayuda del STATGRAPH para generar los números aleatorios del número de reclamos, que se suponen distribuidos como una Log-normal con $\mu=10,000$ y $\sigma=4000$. Una vez generados éstos números, se suman tantos números aleatorios lognormales de acuerdo con el número aleatorio del

número de reclamos se indique, con la ayuda del paquete LOTUS 123. La media (μ) y la desviación estándar (σ) también irán creciendo en un 15% cada año, obteniéndose de esta forma lo siguiente:

AÑO 1:	Número de reclamos: (Poisson) $\lambda = 10$
	Monto de reclamos: (Log-normal) $\mu=10,000 \sigma=4,000$
AÑO 2:	Número de reclamos: (Poisson) $\lambda = 12$
	Monto de reclamos: (Log-normal) $\mu=11,500 \sigma=4,600$
AÑO 3:	Número de reclamos: (Poisson) $\lambda = 14$
	Monto de reclamos: (Log-normal) $\mu=13,225 \sigma=5,290$
AÑO 4:	Número de reclamos: (Poisson) $\lambda = 17$
	Monto de reclamos: (Log-normal) $\mu=15,209 \sigma=6,084$
AÑO 5:	Número de reclamos: (Poisson) $\lambda = 21$
	Monto de reclamos: (Log-normal) $\mu=17,490 \sigma=6,995$
AÑO 6:	Número de reclamos: (Poisson) $\lambda = 25$
	Monto de reclamos: (Log-normal) $\mu=20,114 \sigma=10,493$
AÑO 7:	Número de reclamos: (Poisson) $\lambda = 30$
	Monto de reclamos: (Log-normal) $\mu=23,131 \sigma=12,067$
AÑO 8:	Número de reclamos: (Poisson) $\lambda = 38$
	Monto de reclamos: (Log-normal) $\mu=26,601 \sigma=13,877$

Como se vio anteriormente, las series de tiempo son una importante herramienta para poder realizar modelos de comportamiento de algún fenómeno aleatorio y en base a esto poder realizar pronósticos de un futuro. Se puede notar que tanto el modelo como los pronósticos están basados en información histórica pasada, es decir, en una serie de observaciones de algún fenómeno que se desea estudiar.

Tienen un enfoque distinto que el de análisis de regresión, que aunque los dos sirvan para hacer pronósticos y modelos, como es el caso de la inflación, déficit financiero, déficit petrolero, como se comporta los valores de una casa de bolsa y, en el caso especial que se está tratando, los reclamos que tiene una compañía de seguros durante el transcurso de varios años, para poder explicar cuales son las causas de ese comportamiento y sabiendo como se comporta, poder controlar a ese

fenómeno, en el caso de análisis de regresión se considera que las observaciones que se realizan son independientes, logrando de esta forma reflejar solamente la tendencia del fenómeno estudiado. En las series de tiempo, las observaciones son dependientes entre sí, esto quiere decir que una observación depende de la anterior, por ejemplo, dentro del fenómeno de la inflación el que se tenga un índice de inflación alto en el mes de enero, puede inducir a pensar que en el mes de febrero también se tenga un índice de inflación alto.

El fenómeno que se desea estudiar deberá ser estacionario, es decir, que deberá de presentar un comportamiento homogéneo durante todo el proceso que se lleve a cabo el estudio, esto se representa por medio de que el fenómeno no presenta tendencias ni ciclos y tampoco estacionalidades, deberá ser un proceso homogéneo, sin que algún dato se dispare muy lejos del valor promedio observado. En el caso de que se presenten tendencias y estacionalidades se tendrán que realizar ciertas transformaciones de los valores observados originales (Transformación estabilizadora de la media y la varianza) de tal manera que el fenómeno transformado sea un proceso homogéneo y estacionario.

Las funciones de autocorrelación y sus respectivos correlogramas de cualquier proceso ARIMA (p,d,q) son muy importantes ya que por medio de ellos se podrá "diagnosticar", dar una primera aproximación del modelo al cual el fenómeno que se está estudiando, se parezca.

Al estudiar los procesos ARIMA (p,d,q) de manera teórica por medio de sus funciones de autocorrelación y correlogramas, se podrán encontrar propiedades de estos procesos que ayudarán para poder explicar de la mejor manera posible como se comporta el fenómeno estudiado, así como el de poder hacer pronósticos de lo más confiables posibles (invertibilidad, estacionariedad, parsimonia, memoria finita o infinita).

Después de haber identificado a que tipo de modelo teórico se parece el modelo práctico se pasa a la fase de estimación de los parámetros p y q. Para hallar el valor de estos parámetros, se valdrá de la ayuda del método de máxima verosimilitud y de las

ecuaciones de Yule-Walker.

Una vez estimado los parámetros p y/o q se requiere que los valores de estos parámetros sean sometidos a una serie de pruebas (Diagnóstico del modelo) basadas en el análisis de residuales y el modelo que pase el mayor número de pruebas, será el modelo más adecuado.

Pasando la etapa de diagnóstico se tiene el modelo más adecuado que servirá para poder realizar pronósticos lo más confiables posibles.

Modelos de series de tiempo.

En este apartado se tratará el modelado de series de tiempo como procesos estocásticos. Cada observación en un proceso estocástico es una variable aleatoria, y las observaciones se desarrollan en el tiempo de acuerdo con ciertas leyes probabilísticas. Entonces un proceso estocástico se puede definir como una colección de variables aleatorias que se ordenan en el tiempo. (ver definición de series de tiempo)

El modelo define el mecanismo por el cual las observaciones se generan. Un ejemplo simple es el proceso de medias móviles de primer orden,

$$z(t) = a_t + \theta a_{t-1}$$

donde a_t es una sucesión de variables aleatorias no correlacionadas con media cero y varianza constante y θ es un parámetro. Un conjunto particular de valores de a_0, a_1, \dots, a_T resulta en una sucesión de observaciones correspondiente, z_1, \dots, z_T . Teniendo un conjunto diferente de valores de a_0, a_1, \dots, a_T , se obtiene un conjunto diferente de observaciones y el modelo anterior se puede considerar como capaz de generar un conjunto infinito de tales realizaciones sobre el periodo $t = 1, \dots, T$. Entonces el modelo efectivamente define una distribución conjunta para las variables aleatorias z_1, \dots, z_T .

Los momentos de un proceso estocástico se definen con respecto a la distribución de la variable aleatoria z_1, \dots, z_T . La media del proceso al tiempo t es

$$\mu(t) = E [Z_t], \quad t = 1, \dots, T$$

y se puede interpretar como un valor promedio de Z_t tomado sobre todas las posibles realizaciones. Los segundos momentos tienen una interpretación similar. La varianza al tiempo t se define por

$$\text{Var} [z_t] = E [(z_t - \mu_t)^2], \quad t = 1, \dots, T$$

mientras la covarianza entre z_t y $z_{t-\tau}$ se da por

$$\text{Cov} [z_t, z_{t-\tau}] = E [(z_t - \mu_t) (z_{t-\tau} - \mu_{t-\tau})] \quad t = \tau+1, \dots, T$$

Si se disponen de varias realizaciones, las cantidades anteriores se pueden estimar con promedios 'conjuntos'. Por ejemplo,

$$\hat{\mu}_t = m^{-1} \sum_{i=1}^m z_t^{(i)} \quad t = 1, \dots, T$$

donde $z_t^{(i)}$ representa la i -ésima observación sobre z_t y m es el número de realizaciones. Sin embargo, en muchas series de tiempo, solamente se dispone de una serie de observaciones. En estas circunstancias, no se pueden hacer inferencias significativas acerca de las cantidades definidas anteriormente, al menos que algunas restricciones sean puestas en el proceso que se suponen seguir para generar los pronósticos. Esto conduce al concepto de estacionariedad.

Estacionariedad.

Quando se dispone solamente de una realización de observaciones, la atención se debe cambiar desde el conjunto de observaciones en puntos particulares del tiempo, hacia el promedio de observaciones sobre el tiempo. Esto es solamente

posible si el proceso generador de datos es tal que la media, varianza y covarianza de z_t son independientes del tiempo. Entonces, por ejemplo, si $\mu_t = \mu$ para $t = 1, \dots, T$, se puede estimar tomando el promedio de las observaciones z_1, \dots, z_T .

Para que un proceso sea estacionario, las siguientes condiciones se deben cumplir para todos los valores de t :

$$\begin{aligned} E [z_t] &= \mu \\ E [(z_t - \mu)^2] &= \gamma(0) \\ E [(z_t - \mu)(z_{t-\tau} - \mu)] &= \gamma(\tau) \quad \tau = 1, 2, \dots \end{aligned}$$

Las expresiones anteriores definen la media, la varianza y la autocovarianza en el retraso τ del proceso generador de datos.

La media, varianza y autocovarianza se pueden estimar de una sencilla serie de observaciones como sigue:

$$\begin{aligned} \bar{\mu} &= \bar{z} = T^{-1} \sum_{i=1}^T z_i \\ \bar{\gamma}(0) &= c(0) = T^{-1} \sum_{i=1}^T (z_i - \bar{z})^2 \\ \bar{\gamma}(\tau) &= c(\tau) = T^{-1} \sum_{i=\tau+1}^T (z_i - \bar{z})(z_{i-\tau} - \bar{z}) \end{aligned}$$

Si el proceso es ergódico, estos estadísticos se vuelven estimadores consistentes de la media, varianza y autocovarianzas. La ergodicidad no se definirá formalmente aquí, pero lo que requiere básicamente es que las observaciones que están lo suficientemente alejadas entre sí deben estar casi no correlacionadas. Para los modelos que se verán, la estacionariedad implica ergodicidad.

Las condiciones para la varianza, media y autocovarianza definen la estacionariedad débil.

El ejemplo más sencillo de un proceso estocástico estacionario es una sucesión de variables aleatorias no correlacionadas con media y varianza constante. Un proceso de este tipo se conoce como 'ruido blanco'. El símbolo a_t siempre denotará una variable ruido blanco, y a menos que se establezca

de otra forma, tal variable tendrá media igual a cero y varianza igual a σ^2 . Debido a que las variables en una sucesión de ruido blanco no están correlacionadas, las autocovarianzas en retrasos distintos de cero son todas nulas. Entonces

$$E [a_t a_{t-\tau}] = \begin{cases} \sigma^2 & \tau = 0 \\ 0 & \tau \neq 0 \end{cases}$$

Funciones de autocovarianza y autocorrelación.

Quando un proceso estocástico es estacionario, sus propiedades se pueden resumir graficando $\gamma(\tau)$ contra τ . Esto se conoce como función de la autocovarianza. Como $\gamma(\tau) = \gamma(-\tau)$, no es necesario extender la gráfica sobre valores negativos de τ .

Las autocovarianzas se pueden estandarizar dividiendo entre la varianza del proceso, que son las autocorrelaciones

$$\rho(\tau) = \gamma(\tau)/\gamma(0) \quad \tau = 0, \pm 1, \pm 2, \dots$$

Una gráfica de $\rho(\tau)$ contra valores no negativos de τ da la función de autocorrelación. Nótese que $\rho(0) = 1$ por definición.

Las propiedades del modelo de medias móviles son fácilmente de derivar. Como a_t es una variable de ruido blanco con media cero y desviación estándar σ , se sigue que:

$$\mu = E [a_t] + \theta E [a_{t-1}] = 0$$

mientras

$$\begin{aligned} \gamma(0) &= E [(a_t + \theta a_{t-1}) (a_t + \theta a_{t-1})] \\ &= E [a_t^2] + 2 E [\theta a_t a_{t-1}] + \theta^2 E [a_{t-1}^2] \\ &= (1 + \theta^2) \sigma^2 \end{aligned}$$

Similarmente

$$\begin{aligned} \gamma(1) &= E [(a_t + \theta a_{t-1}) (a_{t-1} + \theta a_{t-2})] \\ &= E [a_t a_{t-1}] + \theta E [a_t a_{t-2}] + \theta E [a_{t-1}^2] + \theta^2 E [a_{t-1} a_{t-2}] \\ &= \theta E [a_{t-1}^2] \\ &= \theta \sigma^2 \end{aligned}$$

y

$$\gamma(\tau) = 0 \quad \tau = 2, 3, 4$$

Se nota que la media, la varianza y covarianzas son independientes de t y por lo tanto el proceso es estacionario.

Las funciones de autocovarianza y autocorrelación tienen exactamente la misma forma y dan la misma información sobre la naturaleza del proceso. Es más usual graficar la función de autocorrelación debido a que no tiene dimensiones. Estandarizando $\gamma(1)$, resulta

$$\rho(1) = \theta/(1+\theta^2)$$

Cuando θ es positivo, los valores sucesivos de z_i están correlacionados positivamente y así el proceso tenderá a ser más uniforme que las series aleatorias, a_i . Por otro lado, un valor negativo de θ producirá una serie la cual es más irregular que una serie aleatoria, en el sentido de que valores positivos de z_i tienden a ser seguidos por valores negativos y viceversa. Esto se refleja en la función de autocorrelación, cuando $\rho(1)$ es negativo para $\theta < 0$.

El correlograma.

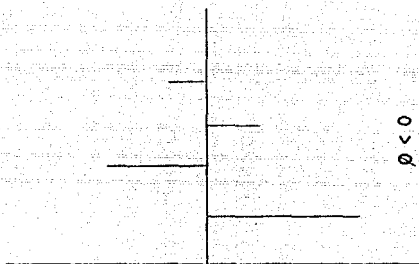
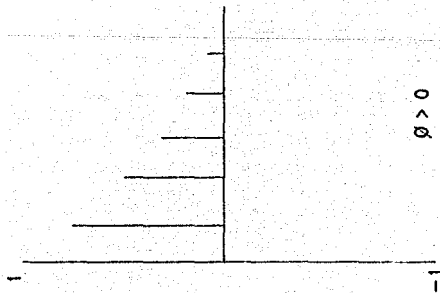
Los estimadores de la media, la varianza y la autocovarianza son la media muestral, la varianza muestral y la autocovarianza muestral respectivamente. Las autocovarianzas muestrales se pueden estandarizar de la misma forma que las autocovarianzas teóricas. Esto produce las autocorrelaciones muestrales

$$r(\tau) = c(\tau)/c(0) \quad \tau = 1, 2, \dots$$

y una gráfica de $r(\tau)$ contra valores no negativos de τ se conoce como la función de autocorrelación muestral o correlograma.

Las autocorrelaciones muestrales son estimados de las autocorrelaciones teóricas correspondientes para el proceso estocástico el cual se supone que genera los datos. Por lo tanto

U.N.A.M.
E.N.E.P. ACATLAN
CORRELOGRAMA



estarán sujetos a la variabilidad muestral, y aunque el correlograma tenderá a reproducir las propiedades de la función de autocorrelación teórica, no las reproducirá exactamente. Las autocorrelaciones muestrales de un proceso de ruido blanco, por ejemplo, estarán cerca de cero, pero en general no serán idénticamente iguales a cero.

El correlograma es la herramienta principal para analizar las propiedades de una serie de observaciones en el dominio del tiempo. Sin embargo, para poder interpretar el correlograma, es necesario conocer algo, primeramente acerca de la variabilidad muestral de las autocorrelaciones estimadas, y secundariamente acerca de las funciones de autocovarianzas de diferentes procesos estocásticos. En lo sucesivo, se examinarán la naturaleza de la función de autocorrelación de varios casos especiales dentro de la clase de modelos autorregresivos-medias móviles. Esto provee la base para la estrategia de construcción de modelos.

El operador retraso.

El operador retraso, L , tiene un rol extremadamente útil en realizar manipulaciones algebraicas en el análisis de series de tiempo. Se define por la transformación

$$Lz_t = z_{t-1}$$

Aplicando L a z_{t-1} resulta $Lz_{t-1} = z_{t-2}$. Sustituyendo la transformación da $L(Lz_{t-1}) = L^2 z_t = z_{t-2}$ y así en general,

$$L^T z_t = z_{t-T}$$

El operador retraso se puede manipular de una forma similar a cualquier cantidad algebraica. Considérese un proceso de medias móviles infinito en el cual el coeficiente para a_{t-j} es ϕ^j para $j = 0, 1, 2, \dots$, y $|\phi| < 1$. el modelo se puede escribir

$$z_t = \sum_{j=0}^{\infty} (\phi L)^j a_t$$

y si a L se considera que tiene la propiedad $|L| \leq 1$, se sigue que $|\phi L| < 1$ y así la serie $1, \phi L, (\phi L)^2, \dots$ se puede sumar como una progresión geométrica infinita, entonces

$$z_t = \sum_{j=0}^{\infty} (\phi L)^j a_t = a_t / (1 - \phi L)$$

y esto se puede reorganizar para dar el proceso auto regresivo de primer orden

$$z_t = \phi z_{t-1} + a_t$$

El operador de primer diferencia, Δ , se puede manipular de una forma similar que el operador de retraso, desde que $\Delta = 1 - L$. La relación entre los dos operadores se puede explotar útilmente muy seguido. Por ejemplo,

$$\Delta^2 z_t = (1 - L)^2 z_t = (1 - 2L + L^2) z_t = z_t - 2z_{t-1} + z_{t-2}$$

Procesos Auto-Regresivos-Medias Móviles

Una clase general de procesos estocásticos se pueden formar introduciendo un número infinito de retrasos en un promedio móvil. Esto produce la representación

$$z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

donde $\psi_0, \psi_1, \psi_2, \dots$ son parámetros. La condición

$$\sum_{j=1}^{\infty} \psi_j^2 < \infty$$

se debe poner para asegurar que el proceso tiene varianza finita. Cualquier modelo que se pueda escribir de esta forma se dice que es un proceso lineal o determinístico.

Un proceso lineal es estacionario y sus propiedades se pueden expresar en términos de la función de autocovarianzas. Estas propiedades se pueden aproximar, a cualquier nivel de

exactitud, por un modelo dibujado de la clase de procesos autorregresivos-medias móviles. Un proceso autorregresivo-medias móviles de orden (p,q), ARMA (p,q), se escribe como:

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

Algunas restricciones se deben establecer para que tal modelo sea estacionario.

Un proceso ARMA se puede escribir más concisamente definiendo polinomios asociados con el operador retraso. Si

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$$

y

$$\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q$$

el modelo z_t se convierte

$$\phi(L)z_t = \theta(L)a_t$$

Esta representación sirve más para ventajas técnicas que de notación. Por ejemplo, si una media distinta de cero, μ , se introduce en un modelo estacionario se vuelve

$$z_t = \mu + \phi^{-1}(L)\theta(L)a_t$$

En lo sucesivo se exploran las propiedades de procesos puros autorregresivos, procesos puros de medias móviles y procesos mixtos. Se supone que los procesos tienen media cero, $\mu=0$. Esto es simplemente por conveniencia y no se pierde generalidad.

Procesos Autorregresivos

Un proceso autorregresivo de orden p se escribe como

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t, \quad t = 1, \dots, T$$

Esto se denota escribiendo $z_t \sim \text{AR}(p)$. Los procesos autorregresivos son populares porque tienen una interpretación natural y porque son más fáciles de estimar que un proceso de medias móviles o mixto.

El primer punto a establecer es el de la estacionariedad del modelo, es decir, que condiciones debe cumplir el modelo para que sea estacionario. Esto se reduce a determinar si se puede escribir en la forma

$$z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

y cumple con la condición

$$\sum_{j=1}^{\infty} \psi_j^2 < \infty$$

Una vez hecho esto, la función de autocorrelación se puede derivar.

Estacionariedad para el modelo de primer orden.

El proceso $\text{AR}(1)$ es

$$z_t = \phi z_{t-1} + a_t, \quad t = 1, 2, 3, \dots, T$$

Aunque la serie se observa primero al tiempo $t = 1$, el proceso se considera como comenzado en cualquier punto del remoto pasado. Sustituyendo repetidamente para valores retrasados de z_t da

$$z_t = \sum_{j=0}^{J-1} \phi^j a_{t-j} + \phi^J z_{t-J}$$

El lado derecho de la ecuación anterior consiste de dos partes, la primera la cual es un promedio móvil de valores retrasados de la variable de ruido blanco que conduce el proceso. La segunda parte depende del valor de z_t al tiempo $t - J$. Tomando esperanzas y tratando a z_{t-J} como un valor fijo, resulta

$$E [z_t] = E \left[\sum_{j=0}^{J-1} \phi^j a_{t-j} \right] + E [\phi^J z_{t-J}] = \phi^J z_{t-J}$$

Si $|\phi| \geq 1$, el valor promedio del proceso depende del valor inicial, z_{t-J} . La ecuación $z_t = \sum_{j=0}^{J-1} \phi^j a_{t-j} + \phi^J z_{t-J}$ por lo tanto contiene un componente determinístico y un conocimiento de z_{t-J} permite hacer predicciones no triviales para los valores futuros de la serie, sin importar que tan alejados estén. Si por el otro lado ϕ es menor que uno en valor absoluto, este componente determinístico es despreciable si J es grande. Cuando $J \rightarrow \infty$, este componente efectivamente desaparece y así el proceso se considera como habiendo comenzado en algún punto del remoto pasado, es muy legítimo escribir $z_t = \phi z_{t-1} + a_t$ en la forma

$$z_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}, \quad t = 1, \dots, T$$

Comparando $z_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}$ con $z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$, se puede notar que un proceso AR(1) con $|\phi| < 1$ es indeterminístico, ya que sumando los coeficientes cuadrados como una progresión geométrica produce:

$$\sum_{j=0}^{\infty} \phi^{2j} = 1 / (1 - \phi^2)$$

La esperanza de z_t es cero para toda t , mientras

$$\begin{aligned} \gamma(0) &= E[z_t^2] = E\left[\sum_{j=-\infty}^{\infty} \phi^j a_{t-j}\right]^2 = \sum_{j=-\infty}^{\infty} \phi^{2j} E[a_{t-j}^2] = \\ &= \sigma^2 \sum_{j=0}^{\infty} \phi^{2j} = \sigma^2 / (1-\phi^2) \end{aligned}$$

Estacionariedad para el modelo de orden p .

El modelo AR(p), $z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t$, será estacionario si las raíces de la ecuación característica

$$x^p - \phi_1 x^{p-1} - \dots - \phi_p = 0$$

son menores que uno en valor absoluto, i.e. si caen dentro del círculo unitario. Una forma alternativa de expresar esta condición es en términos del polinomio asociado

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$$

La ecuación $1 - \phi_1 L - \dots - \phi_p L^p = 0$ es similar en forma a

$$x^p - \phi_1 x^{p-1} - \dots - \phi_p = 0$$

excepto que x se reemplaza por $1/L$ y toda la ecuación se multiplica por L^p . La condición de estacionariedad es que las raíces de $1 - \phi_1 L - \dots - \phi_p L^p = 0$ deben estar fuera del círculo unitario.

Funciones de autocovarianza y autocorrelaciones

Cuando $|\phi| < 1$, el modelo AR(1) tiene media cero y varianza dada por $\sigma^2 / (1-\phi^2)$. La autocovarianza al retraso τ se puede derivar expresando z_t como una combinación lineal de z_t y $a_t, a_{t-1}, \dots, a_{t-\tau+1}$. Esto se logra sustituyendo $J = \tau$ y así

$$\gamma(\tau) = E [z_t z_{t-\tau}] = E \left[\left(\phi^\tau z_{t-\tau} + \sum_{j=0}^{\tau-1} \phi^j a_{t-j} \right) z_{t-\tau} \right] =$$

$$\phi^\tau E [z_{t-\tau}^2] = \phi^\tau \gamma(0), \quad \tau = 1, 2, \dots$$

Las autocovarianzas solo dependen de τ , confirmando que el proceso es estacionario.

Las autocovarianzas se pueden derivar de una forma más directa si la estacionariedad se supone desde el principio. Multiplicando ambos lados de $z_t = \phi z_{t-1} + a_t$ por $z_{t-\tau}$ y tomando esperanzas da

$$E [z_t z_{t-\tau}] = \phi E [z_{t-1} z_{t-\tau}] + E [a_t z_{t-\tau}], \quad \tau = 0, 1, 2, \dots$$

Para un proceso estacionario,

$$E [z_{t-1} z_{t-\tau}] = E [z_t z_{t-\tau+1}] = \gamma(\tau-1) \text{ y, si } \tau > 0, \text{ el último}$$

término es cero, en la medida que a_t no está correlacionado con valores pasados. Por lo tanto

$$\gamma(\tau) = \phi \gamma(\tau-1), \quad \tau = 1, 2, \dots$$

Esta es una ecuación en diferencias de primer orden con una

solución dada por $\phi^\tau E [z_{t-\tau}^2] = \phi^\tau \gamma(0)$. La expresión para la varianza del proceso se puede obtener de una manera similar notando que $E [a_t z_t] = \sigma^2$.

La función de autocorrelación toma la forma $\rho(\tau) = \phi^\tau$ y para valores positivos de ϕ la función exhibe un decaimiento exponencial como se muestra en la figura. Cuando ϕ es negativo, la función de autocorrelación de nuevo decae exponencialmente pero oscila entre valores positivos y negativos como se muestra en la otra figura. Cuando ϕ es positiva, la serie esta cambiando suavemente en que las diferencias entre observaciones sucesivas tienden a ser relativamente pequeñas. Un valor negativo de ϕ conduce a patrones irregulares, desde que valores cercanos entre

si, están correlacionados negativamente.

La misma técnica se puede emplear para derivar las propiedades de cualquier proceso AR. Multiplicando

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t$$

por $z_{t-\tau}$ tomando esperanzas y dividiendo por $\gamma(0)$ da una ecuación en diferencias de orden p

$$\rho(\tau) = \phi_1 \rho(\tau-1) + \dots + \phi_p \rho(\tau-p), \quad \tau = 1, 2, \dots$$

Cuando $\tau = 0$, la siguiente expresión para la varianza se obtiene despues de algunos arreglos

$$\gamma(0) = \sigma^2 / [1 - \rho(1)\phi_1 - \dots - \rho(p)\phi_p]$$

Para todo proceso estacionario AR(p) la función de autocorrelación decae, en el sentido de que $\rho(\tau)$ tiende a cero cuando $\tau \rightarrow \infty$. El comportamiento actual lo exhibe, por ejemplo con respecto a movimientos ciclicos, depende de las raices de la ecuación característica

$$x^p - \phi_1 x^{p-1} - \dots - \phi_p = 0$$

Procesos de medias móviles

Un proceso de medias móviles de orden q se escribe como

$$z_t = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad t = 1, \dots, T$$

y se denota por $z_t \sim MA(q)$. Un proceso de medias móviles finito

siempre es estacionario. La condición $\sum_{j=0}^{\infty} \psi_j^2 < \infty$ se satisface

automáticamente y las covarianzas se pueden derivar sin la necesidad de poner alguna restricción sobre los parámetros $\theta_1, \dots, \theta_q$. Sin embargo algunas veces se necesita expresar un proceso MA en forma autorregresiva. Si esto se debe hacer, los parámetros del proceso MA(q) deben de satisfacer las condiciones similares a aquellas que se impusieron a los parámetros del proceso AR(p) para asegurar la estacionariedad. Si estas condiciones se cumplen, el proceso MA se dice que es invertible.

Funciones de autocovarianza y autocorrelación.

Tomando esperanzas de

$$z_t = a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad t = 1, \dots, T$$

inmediatamente se muestra que z_t tiene media cero, mientras

$$\gamma(0) = E [z_t^2] = (1 + \theta_1^2 + \dots + \theta_q^2) \sigma^2$$

Las autocovarianzas se obtienen por

$$\gamma(\tau) = \begin{cases} (\theta_\tau + \theta_1 \theta_{\tau+1} + \dots + \theta_{q-\tau} \theta_q) \sigma^2 & \tau \leq q \\ 0 & \tau > q \end{cases}$$

A partir que las autocovarianzas en retrasos mayores que q son cero, la función autocovarianza, y por lo tanto la función de

autocorrelaciones tiene un distinto valor en $\tau = q$. Esto contrasta con la función autocovarianza de un proceso AR el cual gradualmente decae hacia cero. Si q es pequeño, el valor máximo de $|\rho(1)|$ está muy por debajo de la unidad. Se puede mostrar que

$$|\rho(1)| \leq \cos [\pi/(q+2)]$$

Invertibilidad.

El proceso MA(1) $z_t = a_t + \theta_1 a_{t-1}$, se puede expresar en términos de valores retrasados de z_t sustituyendo repetidamente por valores retrasados de a_t . Esto produce

$$z_t = \theta z_{t-1} - \theta^2 z_{t-2} + \dots - (-\theta)^J z_{t-J} + a_t - (-\theta)^{J+1} a_{t-J-1}$$

Si z_t no depende de un sobrealto del sistema que surja de algún punto del remoto pasado, θ debe ser menor que uno en valor absoluto. Si J tiende a infinito, el último término en la ecuación anterior desaparece y z_t se puede escribir como un proceso autorregresivo infinito con pesos declinantes, i.e.

$$z_t = - \sum_{j=1}^{\infty} (-\theta)^j z_{t-j} + a_t$$

Un modelo MA(1) con $|\theta| > 1$ no es invertible, pero aun es estacionario. Sin embargo, su función de autocorrelación se puede reproducir exactamente por un proceso invertible con parámetro $1/\theta$. Esto se puede ver sustituyendo en $\rho(1) = \theta/(1+\theta^2)$

$$\rho(1) = \frac{1/\theta}{1 + (1/\theta)^2} = \frac{\theta}{1 + \theta^2}$$

Excepto para el caso de $|\theta| = 1$, entonces, una función de autorrelación particular será compatible con dos parámetros, donde solamente uno es invertible. Poniendo atención en los procesos invertibles resuelve el problema de identificabilidad, pero la razón principal para no considerar los procesos no invertibles es que dan lugar a predicciones ineficientes.

Un proceso MA(1) con $\theta = 1$ es algo anómalo, desde que puede

ser únicamente identificado desde la función de autocorrelación. Aunque tales procesos no son estrictamente invertibles, no se pueden dejar de tomar en cuenta junto con los procesos en que $|\theta| > 1$, debido a que se pueden hacer predicciones sensibles con ellos.

El concepto de invertibilidad se extiende a proceso MA de orden mayor. Las condiciones necesarias para invertibilidad se pueden expresar en términos del polinomio asociado MA

$$\theta(L) = 1 + \theta_1 L + \dots + \theta_q L^q = 0$$

requiriendo que las raíces caigan fuera del círculo unitario.

Procesos Mixtos.

El proceso general autorregresivo-medias móviles se definió como

$$z_t = \phi_1 z_{t-1} + \dots + \phi_p z_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q}$$

y tales procesos se indican por la notación $z_t \sim \text{ARMA}(p, q)$. Los modelos AR(p) y MA(q) son casos especiales y es muy legítimo denotarlos por ARMA(p,0) y ARMA(0,q) respectivamente.

Estacionariedad e invertibilidad.

El que un proceso mixto sea estacionario depende solamente de la parte autorregresiva. Las condiciones se pueden expresar en términos del polinomio autorregresivo asociado $\phi(L)z_t = \theta(L)a_t$ requiriendo que las raíces de $\phi(L) = 0$ caigan fuera del círculo unitario. De una forma similar, la condición de invertibilidad es exactamente la misma que para un modelo MA(q), siendo que las raíces de $\theta(L) = 0$ caigan fuera del círculo unitario.

La razón por la cual la estacionariedad depende de la parte AR del modelo surge en la medida en que un intento se hace para

expresar un proceso mixto como un promedio móvil infinito. El proceso ARMA(1,1)

$$z_t = \phi_1 z_{t-1} + a_t + \theta_1 a_{t-1}$$

es el proceso mixto más simple y podría ser posible sustituir por valores retrasados de z_t repetidamente como en el proceso AR(1). Una forma alternativa de definir la estacionariedad de un proceso mixto es escribir el modelo ARMA(1,1) en la forma

$$(1 - \phi_1 L)z_t = (1 + \theta_1 L)a_t$$

y se divide ambos lados por $(1 - \phi_1 L)$. Esto produce

$$z_t = \frac{a_t}{1 - \phi_1 L} + \frac{a_{t-1}}{1 - \phi_1 L}$$

Si L se estima que satisface la condición $|L| \leq 1$, el término $1/(1-\phi_1 L)$ se puede considerar como la suma de una progresión geométrica infinita $1, \phi_1 L, (\phi_1 L)^2, \dots$ cuando $|\phi| < 1$ y así la ecuación anterior se puede reescribir como

$$\begin{aligned} z_t &= \sum_{j=0}^{\infty} (\phi_1 L)^j a_t + \theta \sum_{j=0}^{\infty} (\phi_1 L)^j a_{t-1} \\ &= \sum_{j=0}^{\infty} \phi^j a_{t-j} + \theta \sum_{j=0}^{\infty} \phi^j a_{t-j-1} \\ &= a_t + \sum_{j=1}^{\infty} (\theta \phi^{j-1} + \phi^j) a_{t-j} \end{aligned}$$

Cuando $\theta = 0$ esta expresión se reduce al del modelo AR(1). Dado que $|\phi| < 1$, los pesos en la ecuación anterior declinan lo suficientemente rápido para que el proceso tenga varianza finita y que las autocovarianzas existan.

Para $p > 1$, la representación del MA infinito se puede obtener de una artificio similar el cual impone factorizar $\phi(L)$ y

expandir $\phi^{-1}(L)\theta(L)$ en fracciones parciales. Una forma más conveniente de proceder es igualar los coeficientes de las potencias de L en la expresión

$$\theta(L) = \phi(L) \psi(L)$$

Después de algunos arreglos, la ecuación anterior se convierte en

$$\psi_0 = 1$$

$$\psi_j = \theta_j + \sum_{i=1}^{\min(j,p)} \phi_i \psi_{j-i} \quad j \leq q$$

$$\psi_j = + \sum_{i=1}^{\min(j,p)} \phi_i \psi_{j-i} \quad j > q$$

Para $j \geq \max(p, q+1)$, las ψ_j 's se determinan por la ecuación

en diferencias $\psi_j = \sum_{i=1}^{\min(j,p)} \phi_i \psi_{j-i}$, con valores iniciales dados

por los valores previos p de ψ_j .

Ejemplo. Para el modelo ARMA(1,1), las ψ_j 's se calculan de la ecuación en diferencias

$$\psi_j = \phi \psi_{j-1} \quad j \geq 2$$

con valor inicial

$$\psi_1 = \theta + \phi \psi_0 = \theta + \phi$$

Funciones de autocovarianza y autocorrelaciones.

Las propiedades de los modelos mixtos incorporan características de los modelos AR y MA. Esto se puede ilustrar claramente por medio del modelo ARMA(1,1). Multiplicando

$$z_t = \phi_1 z_{t-1} + a_t + \theta_1 a_{t-1}$$

por $z_{t-\tau}$ y tomando esperanzas resulta

$$\gamma(\tau) = \phi \gamma(\tau-1) + E [a_t z_{t-\tau}] + \theta E [a_{t-1} z_{t-\tau}] \quad \tau = 0, 1, \dots$$

Las últimas dos esperanzas son cero para $\tau > 1$. Para $\tau = 1$, el último término se convierte

$$E [a_{t-1} z_{t-1}] = E [a_{t-1} (\phi z_{t-2} + a_{t-1} + \theta a_{t-2})] = \sigma^2$$

Aunque el segundo término de $\gamma(\tau)$ siga siendo cero. Cuando $\tau = 0$, ambas esperanzas son distintas de cero y son dadas por

$$E [a_t z_t] = \sigma^2$$

y

$$\begin{aligned} E [a_{t-1} z_t] &= E [a_{t-1} (\phi z_{t-1} + a_t + \theta a_{t-1})] \\ &= \phi E [a_{t-1} z_{t-1}] + \theta \sigma^2 \\ &= (\phi + \theta) \sigma^2 \end{aligned}$$

La función de autocovarianza es por lo tanto

$$\gamma(0) = \phi \gamma(1) + \sigma^2 + \theta \phi \sigma^2 + \theta^2 \sigma^2$$

$$\gamma(1) = \phi \gamma(0) + \theta \sigma^2$$

$$\gamma(\tau) = \phi \gamma(\tau - 1) \quad \tau = 2, 3, \dots$$

Sustituyendo $\gamma(1)$ en $\gamma(0)$ se obtiene

$$\gamma(0) = \frac{1 + \theta^2 + 2\phi\theta}{1 - \phi^2} \sigma^2$$

Y sustituyendo $\rho(0)$ en $\rho(1)$ resulta

$$\rho(1) = \frac{(1 + \phi\theta)(\phi + \theta)}{1 - \phi^2} \phi^2$$

Para obtener las funciones de autocorrelación $\rho(1)$ y $\rho(\tau)$ entre $\rho(0)$ lo que da lugar

$$\rho(1) = \frac{(1 + \phi\theta)(\phi + \theta)}{1 + \theta^2 + 2\phi\theta}$$

$$\rho(\tau) = \phi \rho(\tau-1)$$

Analizando la función de autocorrelación, se puede ver que, para $\tau > 1$, su comportamiento se gobierna por la ecuación en

diferencia $\rho(\tau) = \phi \rho(\tau-1)$. Entonces las autocorrelaciones muestran un decaimiento exponencial, con oscilaciones cuando ϕ es negativo. Esto es igual como en el caso AR(1). Sin embargo, mientras que el valor inicial para el proceso AR(1) es $\rho(0) = 1$, en el modelo ARMA(1,1) el valor inicial está dado por $\rho(1)$. Como se puede ver en la ecuación para $\rho(1)$, este depende de ϕ y θ , y el signo para $\rho(1)$ dependerá del signo de $(\phi + \theta)$.

Las propiedades de modelos ARMA de mayor orden se pueden derivar de una manera similar. El patrón general el cual emerge para la función de autocorrelación es que las primeras q autocorrelaciones dependen de tanto los parámetros de las medias móviles y autorregresivos. Las autocorrelaciones de mayor orden se dan por una ecuación en diferencias de orden p de la forma

$$\rho(\tau) = \phi_1 \rho(\tau - 1) + \dots + \phi_p \rho(\tau - p)$$

con $\rho(q)$, $\rho(q-1)$, ..., $\rho(q-p+1)$ como valores iniciales

Factores comunes

Si los polinomios AR y MA en $\phi(L) z_t = \theta(L) a_t$ tienen raíces las cuales son las mismas, se dice que tienen factores comunes. En este caso el modelo está sobreparametrizado, desde que un

modelo con idénticas propiedades se puede construir reduciendo tanto p y q por uno. Es importante reconocer los factores comunes, debido a que si existen el modelo no será identificable y problemas de cálculo pueden surgir.

A continuación se dan unas gráficas de funciones de autocorrelaciones que servirán para poder indentificar un modelo ARMA(p,q).

2.1. Análisis de la serie de tiempo.

Como se explicó anteriormente, los datos observados del monto y número del reclamo se generaron en base al Método MonteCarlo. Para ello se valió de la ayuda del paquete estadístico STATGRAPH y Lotus, donde el procedimiento a seguir es el siguiente:

1^o Se generan numeros aleatorios distribuidos como log-normales con media igual a 10,000 y desviación estandar igual a 4,000. Esto se hace en el paquete estadístico STATGRAPH, realizando los siguientes pasos:

- a) Cargar en la computadora el paquete estadístico.
- b) Aparecerá en pantalla 6 grupos de alternativas, cada grupo con su respectivo menu, los cuales son:

DATA HANDLING AND SYSTEM UTILITIES
(Manejo de Datos y utilidades del sistema)
PLOTING AND DESCRIPTIVE STATISTICS
(Gráficas y Estadísticas descriptivas)
ANOVA AND REGRESSION ANALYSIS
(Análisis ANOVA y de Regresión)
TIME SERIES PROCEDURES
(Procedimientos de Series de Tiempo)
ADVANCED PROCEDURES
(Procedimientos Avanzados)
MATHEMATICAL PROCEDURES
(Procedimientos Matemáticos)

El grupo que se utilizará para la generación de números aleatorios es el de PLOTTING AND DESCRIPTIVE STATISTICS, que tiene las siguientes alternativas a escoger

- | | |
|------------------------------|--------------------------|
| E. Plotting Functions | (Funciones de Graficado) |
| F. Descriptive Methods | (Metodos Descriptivos) |
| G. Estimation and Testing | (Estimación y Pruebas) |
| H. Distribution Functions | (Funciones de Distr.) |
| I. Exploratory Data Analysis | (Análisis de Datos Exp) |

De donde se empleara H. Distribution Functions. La función H del paquete, así como cualquier otro procedimiento o rutina, se escoge por medio del manejo del cursor, posicionando este en la alternativa deseada y una vez seleccionada presionar ENTER.

Al seleccionar la opción H. Distribution Functions, se presenta otra subrutina con otro menú:

1. Distribution Fitting
2. Distribution Plotting
3. Tail Area Probabilities
4. Inverse Distribution Functions
5. Random Number Generation

Del menú anterior se escoge la opción 5. Random Number Generation. Cuando se haga, aparecerá en pantalla las distribuciones por medio de las cuales se pueden generar números aleatorios:

THE FOLLOWING DISTRIBUTIONS ARE AVAILABLE

(Las siguientes distribuciones son disponibles)

DISCRETE DATA

- | | |
|---------------------|----------------------|
| 1. Bernoulli | 4. Geometric |
| 2. Binomial | 5. Negative Binomial |
| 3. Discrete Uniform | 6. Poisson |

Continuous Data

- | | |
|-----------------|-----------------|
| 7. Beta | 13. Lognormal |
| 8. Chi-Square | 14. Normal |
| 9. Erlang | 15. Student's T |
| 10. Exponential | 16. Triangular |
| 11. F | 17. Uniform |
| 12. Gamma | 18. Weibull |

Data Multivariate

19. Multivariate Normal

ENTER THE NUMBER OF YOUR CHOICE (1):

En la última línea se pide que tipo de distribución es la que se utilizará para generar los números aleatorios. Para nuestro caso se selecciona la opción 13. Lognormal. Posteriormente se introducirá la media de la distribución, que es 10000, y la desviación estándar que es 4000, el número de observaciones, 200, y por último el nombre de la variable en el cual guardará los números aleatorios: LOG1. Esta rutina sirve para generar el monto de los reclamos.

Un procedimiento semejante se utilizará para generar el número de reclamos, que se distribuyen como Poisson:

- i) Se selecciona la opción 5. Random Number Generation
- ii) Se selecciona el tipo de distribución 6. Poisson
- iii) Se introduce la media o λ : 10
- iv) Se introduce el número de muestras a generar: 12
- v) Se introduce el nombre de la variable donde se salvarán los números aleatorios: POI1
- vi) Se presiona ENTER para continuar

Para poder regresar al menú principal donde se presentan los 6 grupos principales de procedimientos, o para regresar al menú

previo donde se encuentra uno, se presiona la tecla F10.

Una vez realizado estos procedimientos, se tendrán que salvar los datos para que se puedan procesar en LOTUS. Para ello se utiliza el grupo principal de rutinas llamado

DATA HANDLING AND SYSTEM UTILITIES.

Al seleccionar este grupo se observaran las siguientes opciones

- A. DATA MANAGEMENT
- B. SYSTEM ENVIRONMENT
- C. REPORT WRITER AND GRAPHICS REPLAY
- D. PLOTTER INTERFACE

De las cuales se escogera la alternativa A. DATA MANAGEMENT, con lo cual aparecerá un menú:

- 1. Manipulate Defined Variables
- 2. Full Screen Data Editor
- 3. Read Variable Definitions from SG file
- 4. Write Variable Definitions from SG file
- 5. Import Data from ASCII Data File
- 6. Export Data to ASCII Data File
- 7. Import Data from DIF File
- 8. Export Data TO DIF File
- 9. Import LOTUS 1-2-3 Worksheet
- 10. Export LOTUS 1-2-3 Worksheet
- 11. Recode Missing Values

De estas opciones se escoge la numero 4. Write To SG file, que se usa para escribir variables a un archivo. El archivo puede ya existir anteriormente o un nuevo archivo se puede generar. El primer menú muestra todos los archivos que existen en el disco de trabajo. Se puede escribir (W) a un archivo ya existente. Renombrar (R), Copiar (C), Borrar (B) o crear (N) un nuevo archivo. Si se copia o renombra un archivo anterior o se crea uno nuevo, se debe poner un nombre de archivo de 8 caracteres. Si se

escoge escribir a un archivo ya existente, un segundo panel mostrará los nombres de las variables que existen en ese archivo. Para añadir una nueva variable al archivo, se presiona N y se pone el nombre de la variable a ser escrita o salvada en el archivo. Esta variable deberá estar presente en la memoria o deberá haber sido leída previamente de algún archivo. Una variable que ya esté en el archivo podrá ser actualizada (U) presionando U con lo cual salva su valor actual en el archivo. También se puede desplegar (D), añadir un comentario (C), borrarla (E) o reemplazarla (R)

En nuestro caso se generará un nuevo archivo llamado YEAR1 y se salvarán nuevas variables LOG1 y POI1 en este nuevo archivo. Con lo explicado anteriormente, se tendrá que hacer lo siguiente:

- a) Entrar a la opción 4.
- b) Generar un nuevo archivo, presionando la tecla N
- c) Teclar el nombre del nuevo archivo (YEAR1)
- d) Escribir las nuevas variables en el archivo YEAR1, presionando la tecla W.
- e) Salvar las nuevas variables, primero LOG1, presionando la tecla N, y escribir un comentario si se desea hacerlo, después se hace exactamente lo mismo con la variable POI1.

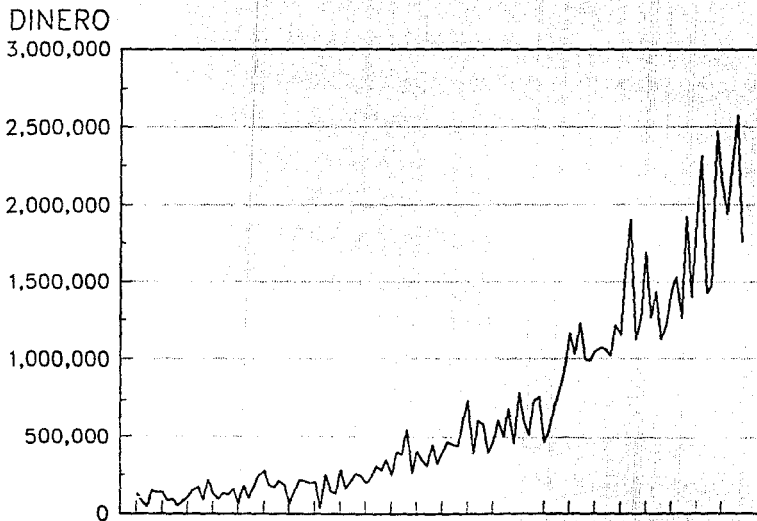
Una vez salvado las variables en el archivo YEAR1, se utiliza el procedimiento Export Data To Lotus 1-2-3 worksheet, el cual permite escribir variables de un archivo SG a un archivo de LOTUS. Para exportar estos datos se necesitan:

- a) Nombre del archivo SG (YEAR1.ssf)
- b) Nombre de la hoja de trabajo a crear (YEAR1.wk1)

Después de generarse los números aleatorios, de haberlos salvado en un archivo de SG y de haberlos exportado a un archivo de Lotus, se procesa la información en el paquete Lotus, debido a que los números aleatorios distribuidos como Lognormales representan los reclamos individuales por siniestro y los números aleatorios distribuidos como Poisson representan el número de reclamos que ocurren durante un mes, por ello con el paquete Lotus se sumaran el número de reclamos ocurridos durante un mes por medio de la función @SUM(rango). Posteriormente se grafican estos números, lo que da por resultado la siguiente gráfica

DATOS ESTADISTICOS QUE SE UTILIZARON PARA EL MODELADO Y ANALISIS DE LA SERIE

	ANO 1	ANO 2	ANO 3	ANO 4	ANO 5	ANO 6	ANO 7	ANO 8	ANO 9	ANO 10
ENE	128,045	170,034	240,496	353,415	284,534	396,193	510,545	933,748	1,598,184	1,922,180
FEB	77,914	90,963	278,253	254,702	348,212	469,320	690,807	1,163,802	1,906,944	1,402,091
MAR	48,003	219,124	183,406	142,618	252,385	448,409	457,529	1,034,356	1,121,856	1,883,019
ABR	149,585	128,154	171,230	128,702	398,748	437,991	790,377	1,229,831	1,262,861	2,318,090
MAY	133,919	94,646	211,867	286,309	383,791	605,170	622,075	993,638	1,697,753	1,428,227
JUN	136,644	132,357	184,242	166,149	551,376	743,259	519,165	993,523	1,265,632	1,480,443
JUL	86,458	118,693	71,018	209,894	268,212	391,238	744,546	1,046,764	1,430,587	2,475,483
AGO	92,912	158,495	138,874	259,945	403,098	612,455	765,781	1,072,671	1,127,860	2,158,871
SEP	51,352	71,040	215,867	249,063	350,436	594,380	468,196	1,061,770	1,223,197	1,938,503
OCT	81,698	176,402	210,936	199,018	311,671	396,288	556,840	1,021,675	1,431,738	2,264,098
NOV	105,344	99,716	197,591	236,507	448,007	461,335	702,898	1,218,948	1,534,328	2,581,352
DIC	148,808	173,450	207,707	312,957	326,042	619,711	814,001	1,155,302	1,262,803	1,768,782
TOTAL	1,240,682	1,633,074	2,311,487	2,799,279	4,326,512	6,176,749	7,642,780	12,926,028	16,863,743	23,621,139



TIEMPO

MONTO DE RECLAMOS

Como se podrá notar en la gráfica, los datos no presentan un proceso estacionario, por lo que se tendrá que estacionarizar, es decir, se tendrá que hacer constante a la media y a la varianza.

Para poder hacer constante a la media se tendrá que diferenciar la serie, en base a la función de retraso definida anteriormente (LZ). Para saber hasta que grado de retraso o que número de diferencias se necesitan, se tendrá que calcular la varianza de la serie original y el de la serie diferenciada y la varianza que sea menor de ellas, esa será la serie que tenga "estabilizada" la media, además de que esta debe ser aproximadamente igual a cero.

Para estabilizar la varianza se tendrá que realizar una transformación llamada Box-Cox que se define como

$$z_t = \begin{cases} z_t^\lambda & \lambda \neq 0 \\ \ln z_t & \lambda = 0 \end{cases}$$

donde $-2 \leq \lambda \leq 2$. El valor de λ se escoge en base al siguiente procedimiento.

- 1) Separar las observaciones en H grupos del mismo tamaño y homogéneos
- 2) Calcular la media (\bar{z}) y desviación estandar de cada grupo (s)

$$3) \text{ Cálculo de } O_{h,\lambda} = \frac{s_h}{\bar{z}_h^{1-\lambda}} = \bar{z}_h \lambda \frac{s_h}{\bar{z}_h} \quad h = 1, \dots, H$$

$$\lambda = -1, -.5, 0, .5, 1$$

$$4) \text{ Cálculo de } \text{med}(\lambda) = \frac{1}{H} \sum_{h=1}^H O_{h,\lambda}$$

$$\text{des}(\lambda) = \left[\frac{1}{H-1} \sum_{h=1}^H (O_{h,\lambda} - \text{med}(\lambda))^2 \right]^{1/2}$$

$$CV(\lambda) = \frac{\text{des}(\lambda)}{\text{med}(\lambda)}$$

- 5) Se escoge λ^* tal que $\lambda^* = \min (CV(\lambda))$

Con los datos presentados, las observaciones se separan en diez grupos cada uno con 12 observaciones, formándose así la siguiente tabla:

En la tabla anterior se puede observar que la transformación que necesita la serie es la de $z_1^* = \ln z_1$.

Una vez encontrada la transformación que se realizará a la serie, se captura la serie original dentro del paquete SYSTAT para transformar la serie y poder analizarla, con lo cual se tiene que realizar los siguientes pasos

- 1) Cargar el paquete SYSTAT en la computadora.
- 2) Entrar al módulo DATA
- 3) Teclar las siguientes instrucciones

```
> SAVE (Nombre de archivo)
> INPUT (Variable 1, Variable 2, ...)
> RUN
INPUT DATA ONE CASE AT A TIME AFTER PROMPT ARROW
```

ANALISIS DE VARIANZA, PARA PODEL REDUCIRLA DEL MODELO

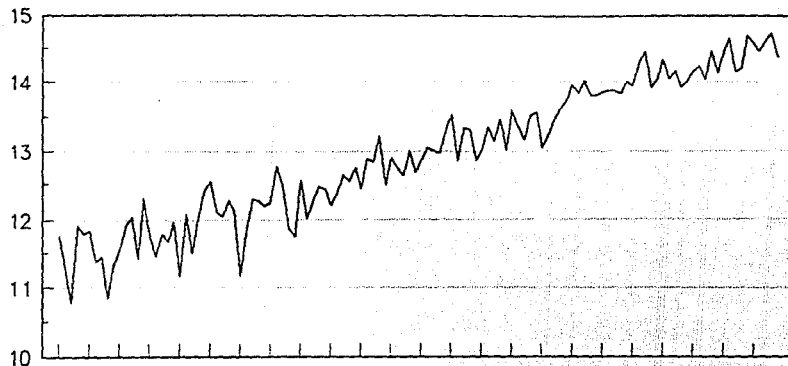
GRUPO	media	desv. std.	-1	-0.5	0	0.5	1
AÑO 1	103390	34324	3.2110038E-06	0.001032476297	0.331985685269	106.7477243841	34324
AÑO 2	136090	42089	2.2725643E-06	0.000838357562	0.309273275039	114.0920806785	42089
AÑO 3	192624	49508	1.3343034E-06	0.000585811756	0.257018855387	112.6028789193	49508
AÑO 4	233273	64811	1.1910220E-06	0.000575243885	0.277833268317	134.1888667247	64811
AÑO 5	360543	80473	6.1906472E-07	0.000371718854	0.223199451938	134.0206308588	80473
AÑO 6	514729	110315	4.1636795E-07	0.000298721591	0.214316659835	153.7606657427	110315
AÑO 7	636898	124989	3.0812881E-07	0.000245904777	0.196246494729	156.6162607449	124989
AÑO 8	1077169	90210	3.0812881E-07	0.000245904777	0.196246484729	156.6162607449	124989
AÑO 9	1405312	231470	3.0812881E-07	0.000245904777	0.196246484729	156.6162607449	124989
AÑO 10	1968428	385088	9.9367598E-08	0.000139441473	0.195637330906	274.4804899579	385088
		media	1.0066100E-06	0.000457928575	0.239500401088	149.9942129499	114158.5
		desv. std.	9.6895000E-07	0.000278203606	0.048814805438	45.48621534795	86484.20615857
		c.v.	0.962398122142	0.607526197277	0.202729458405	0.30325313526	0.845177592195

DATOS ESTADISTICOS CON LA TRANSFORMACION LOGARITMICA

	AÑO 1	AÑO 2	AÑO 3	AÑO 4	AÑO 5	AÑO 6	AÑO 7	AÑO 8	AÑO 9	AÑO 10
ENE	11.76	12.04	12.39	12.78	12.56	12.89	13.14	13.75	14.28	14.47
FEB	11.26	11.42	12.54	12.45	12.76	13.06	13.45	13.97	14.46	14.15
MAR	10.78	12.30	12.12	11.87	12.44	13.01	13.03	13.85	13.93	14.45
ABR	11.92	11.76	12.05	11.77	12.90	12.99	13.58	14.02	14.05	14.66
MAY	11.80	11.46	12.26	12.56	12.86	13.31	13.34	13.81	14.34	14.17
JUN	11.83	11.79	12.12	12.02	13.22	13.52	13.16	13.81	14.05	14.21
JUL	11.37	11.68	11.17	12.25	12.50	12.88	13.52	13.86	14.17	14.72
AGO	11.44	11.97	11.84	12.47	12.91	13.33	13.55	13.89	13.94	14.59
SEP	10.85	11.17	12.28	12.43	12.77	13.30	13.06	13.88	14.02	14.48
OCT	11.31	12.08	12.26	12.20	12.65	12.89	13.23	13.84	14.17	14.63
NOV	11.56	11.51	12.19	12.37	13.01	13.04	13.46	14.01	14.24	14.76
DIC	11.91	12.06	12.24	12.65	12.69	13.34	13.61	13.96	14.05	14.39
TOTAL	137.79	141.25	145.48	147.82	153.26	157.55	160.13	166.64	169.71	173.67

U.N.A.M.
E.N.E.P. ACATLAN
RECLAMOS CON VAR. CONST., MEDIA NO CONSTANTE

DINERO



TIEMPO

MONTO DE RECLAMOS

Al aparecer este mensaje en pantalla se comienza a capturar la serie ORIGINAL un caso a la vez. Al terminar de capturar los datos, se tecldea el siguiente comando: > QUIT

Con las instrucciones anteriores solamente se crea una sola variable (YEAR), por ello se tendrá que crear otras dos nuevas variables: N = variable tiempo y la transformada de la serie original, llamada VARTRA. Para crear estas variables se escriben los siguientes comandos:

```
> LET N=CASE
```

```
> LET VARTRA=LOG(YEAR)
```

Ya que se transformó la serie, estabilizando la varianza, se grafica la serie, ahora llamada VARTRA. Esto se hace por medio del modulo PLOT del paquete estadístico SYSTAT. Cuando se haya cargado este modulo en la memoria RAM de la computadora, se tecldean los siguientes comandos:

```
> USE YEAR (YEAR es el nombre del archivo que contiene las variables YEAR, N, VARTRA)
```

```
> OUTPUT @ (Sirve para que la grafica salga por la impresora y no en pantalla)
```

```
> PLOT VARTRA*N (Grafica la serie tomando al eje y los valores de VARTRA, que es la serie transformada, y al eje de las abscisas a N, que es el tiempo)
```

```
> OUTPUT * (Con este comando hace que salgan los mensajes y graficas por pantalla)
```

Al mostrarse la gráfica se notará que tiene una tendencia, por lo cual se tendra que realizar otra transformación para estabilizar a la media, que sera la de tomar una o dos diferencias de la serie llamada VARTRA.

Al tomar una diferencia en la serie VARTRA y calculando sus estadísticos se obtiene

N OF CASES	119
MINIMUM	-.953
MAXIMUM	1.137
MEAN	0.022
STANDARD DEV	0.371

Al tomar dos diferencias se observa lo siguiente

N OF CASES	118
MINIMUM	-1.480
MAXIMUM	1.712
MEAN	0.001
STANDARD DEV	0.626

Con los resultados anteriores se observa que la serie solamente necesita de una diferencia.

Para realizar los procedimientos anteriores, se valió de la ayuda del paquete estadístico SYSTAT.

Para obtener la 1ª y 2ª diferencia de la serie VARTRA se hace con el módulo SERIES. Dentro de este módulo existe un comando llamado DIFFERENCE, con el cual se obtiene una nueva variable ya transformada o diferenciada y se le llama TRANSF por default. Así que los pasos a seguir para obtener los resultados anteriores son:

- 1) Entrar al módulo SERIES
- 2) Dar la instrucción >USE YEAR

Este archivo tiene las variables YEAR, que es la serie original, N, y VARTRA que es la serie con la transformación $VARTRA = \text{LOG}(YEAR)$

- 3) Teclar el comando >SAVE YEAR1

Con este comando salvara la variable llamada transf es el archivo YEAR1, que es la variable ya transformada con una diferencia.

- 4) Escribir el comando >DIFFERENCE VARTRA

Con lo cual hara la operación $z_t = z_t - z_{t-1}$ y a z_t la llama TRANSF.

Para poder realizar la segunda diferencia se hace lo siguiente:

- 1) >USE YEAR1
- 2) > SAVE YEAR2
- 3) > DIFFERENCE TRANSF

Estos dos nuevos archivos, YEAR1 y YEAR2, se tienen que unir con YEAR cada uno respectivamente para generar otros dos nuevos archivos llamados YEAR3 y YEAR4, para poder graficar las nuevas series y poder calcular sus estadísticos.

Al diferenciar la serie VARTRA por primera y segunda vez, se generan dos variables con distintos valores pero nombre igual (TRANSF). Para saber cual es la serie con una sola diferencia, al momento de graficar y al calcular sus estadísticos hará los cálculos y la gráfica en base a 119 valores y no en base a los 120 valores originales. Esto es debido a que al momento de realizar la diferencia se pierde una observación. Lo mismo sucederá para la variable que tiene dos diferencias, pero en este caso se pierden dos observaciones, por lo tanto los estadísticos y la gráfica se hacen en base a 118 valores.

Para poder unir dos o mas archivos distintos se hace por medio del modulo DATA con las siguientes instrucciones

```
> SAVE YEAR3
> USE YEAR YEAR1
> RUN
```

Con estas instrucciones se generan el nuevo archivo, YEAR3, que contiene a la serie con una diferencia

```
> SAVE YEAR4
> USE YEAR YEAR2
> RUN
```

Este archivo, YEAR4, tendrá la serie con dos diferencias.

Para graficar la serie con una diferencia, se hace en el modulo PLOT, con las siguientes instrucciones:

```
> USE YEAR3
> OUTPUT @
> PLOT TRANSF*N
> OUTPUT *
```

La primera instrucción es para que utilice el archivo que

tiene la serie con una diferencia, la siguiente instrucción sirve para que la gráfica salga por la impresora y no por pantalla. La tercera instrucción sirve para que realice la grafica tomando al eje Y a la variable TRANSF y al eje X los valores de N. La última instrucción sirve para mandar de nuevo las impresiones y mensajes a pantalla.

Lo mismo se hará para graficar la serie con dos diferencias, lo único que cambia es la primer instrucción por > USE YEAR4

Para calcular los estadísticos de las dos series se realizan en el módulo STATS del SYSTAT, con las siguientes instrucciones

```
> USE YEAR3
> OUTPUT @
> STATS TRANSF
> OUTPUT *
```

Lo mismo se hará con el archivo que tiene la serie con dos diferencias (YEAR4)

Anteriormente se observó que la serie solo necesita de una diferencia para que sea estacionaria. Con esto ya se podrá identificar el modelo.

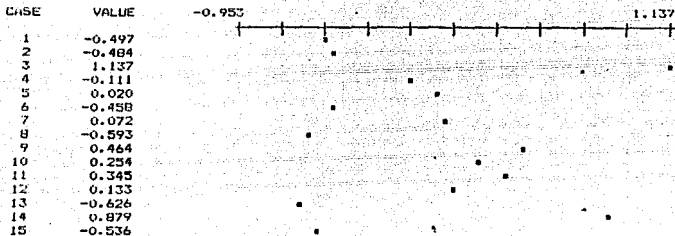
2.3. Modelado de la Serie de Tiempo

Aunque los procesos ARMA se diseñaron para series de tiempo estacionarias, es necesario extenderlas para formar una clase que englobe un rango amplio de comportamientos no estacionarios. Esto se hace diferenciando la serie original una o dos veces como se hizo anteriormente. Un modelo ARMA se adapta, no en base a las observaciones originales, sino a la serie ya diferenciada una o dos veces. El resultado es un proceso Autorregresivo-Medias Móviles Integrado (ARIMA.- AutoRegresive Integrated Moving Average).

El desarrollo de una metodología apropiada para adaptar modelo ARIMA se debe a G.E.P. Box y G.M. Jenkins. De hecho toda la teoría se refiere frecuentemente al método BOX-JENKINS.

PLOT OF TRANSF
 NUMBER OF CASES = 119
 MEAN OF SERIES = 0.022
 STANDARD DEVIATION OF SERIES = 0.369

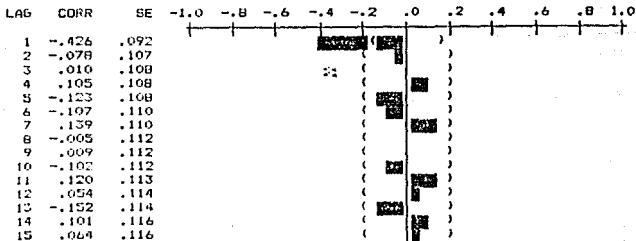
SEQUENCE PLOT OF SERIES



GRAFICA DE LA SERIE CON VAR. ESTABILIZADA Y UNA DIFERENCIA

PLOT OF TRANSF
 NUMBER OF CASES = 119
 MEAN OF SERIES = 0.022
 STANDARD DEVIATION OF SERIES = 0.369

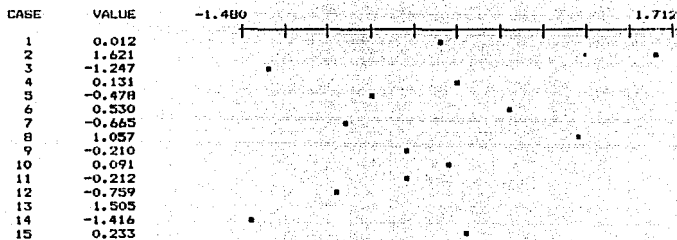
PLOT OF AUTOCORRELATIONS



CORRELOGRAMA DE LA SERIE CON UNA DIFERENCIA

PLOT OF TRANSF
NUMBER OF CASES = 118
MEAN OF SERIES = 0.001
STANDARD DEVIATION OF SERIES = 0.623

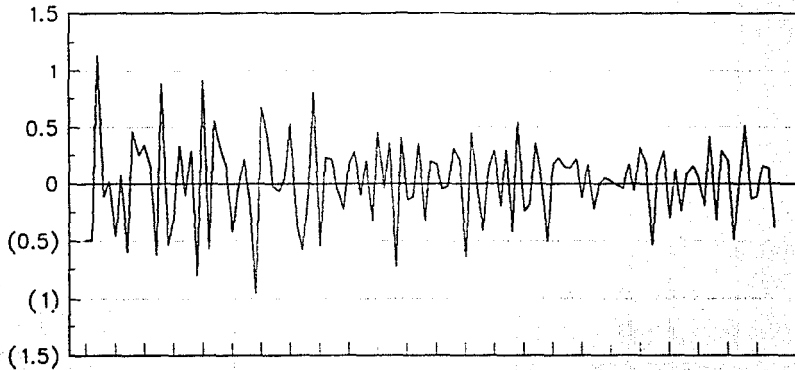
SEQUENCE PLOT OF SERIES



GRAFICA DE LOS DATOS CON VAR. ESTABILIZADA Y DOS DIFERENCIAS

U.N.A.M.
E.N.E.P. ACATLAN
RECLAMOS CON VAR. CONST., MEDIA CONSTANTE

DINERO



TIEMPO

MONTO DE RECLAMOS

Básicamente consiste de una estrategia de selección del modelo que se divide en tres partes. La primera parte consiste en seleccionar un modelo tentativo. Esto significa decidir el grado apropiado de diferenciación, en nuestro caso es un grado de diferenciación, así como valores fijos de p y q . La segunda etapa es la estimación de los parámetros ϕ 's y θ 's. Finalmente el modelo se somete a un diagnóstico, que consiste en una serie de pruebas que se hace al o a los modelos, y el que pase el mayor número de ellas, será el modelo más apropiado. La secuencia de identificación, estimación y diagnóstico se ve como un ciclo que se repite hasta que un modelo satisfactorio se obtiene.

Prueba de aleatoriedad.

Un modelo exitoso tenderá a capturar los movimientos sistemáticos de los datos. Si este objetivo se alcanza, entonces lo que sobra, llamados los residuales, deberán ser esencialmente aleatorios. En otras palabras no deberán contener componentes sistemáticos predecibles. Aquí se introduce las consideraciones básicas que surgen de determinar la aleatoriedad de una serie de tiempo observada y en la sección de diagnóstico se extienden en las pruebas de los residuales de un modelo ARMA simulado.

Sea z_1, \dots, z_t una realización de un proceso de ruido blanco con media μ y varianza σ^2 . Las autocorrelaciones teóricas para tal proceso son, por definición, idénticas a cero para todo intervalo distinto de cero. Un conjunto de muestras de autocorrelaciones, por el otro lado, tenderán a enseñar alguna divergencia de este patrón debido al error de muestreo. Por lo tanto, si un establecimiento de aleatoriedad se basa en las muestras de las correlaciones, primero es necesario investigar su distribución de muestreo. Afortunadamente, un resultado muy simple se dispone para muestras muy grandes, sabiendo que para $\tau = 0$, las $r(\tau)$'s se distribuyen normal e independientemente con una media de cero y una varianza de $1/T$. Esto se obtiene directamente del teorema Mann-Wald, desde que $r(\tau)$ es aproximadamente igual al coeficiente de $z_{t-\tau}$ en la regresión de z_t sobre $z_{t-\tau}$. Por que la serie es aleatoria, este coeficiente

tiene una distribución normal limitante cuando se multiplica por \sqrt{T} , con media igual a cero y varianza igual a:

$$\sigma^2 (\text{plim } T^{-1} \sum z_{1-r}^2)^{-1} = 1$$

Usando el resultado anterior, la hipótesis de que $\rho(r) = 0$ puede ser formalmente probada para cualquier valor particular de r tomando a $T^{1/2}r(r)$ como una normal estandarizada. Al nivel del 5% de significancia, la hipótesis nula se rechaza si $|T^{1/2}r(r)| > 1.96$. Sin embargo, tal prueba solamente es válida si el retraso que se prueba se especifica por adelantado. Esto implica un conocimiento a priori de la naturaleza de la serie. Por ejemplo, para datos trimestrales una prueba de la significancia de $r(4)$ sería claramente relevante. Excepto para el caso de un efecto estacional, sin embargo, tal conocimiento a priori es más bien la excepción mas que la regla, y los procedimientos para pruebas formales se restringen generalmente a las autocorrelaciones de primer orden, $r(1)$. Por esta razón algun esfuerzo se ha puesto para desarrollar procedimientos de pruebas de hipótesis exactos referentes a $\rho(1)$.

Sin tomar en cuenta las notas anteriores, es muy útil graficar dos líneas sobre el correlograma uno a la altura de $2 / \sqrt{T}$ sobre el eje horizontal y el otro a la misma distancia por debajo. Esto se puede usar como un criterio para asignar desviaciones de la aleatoriedad. Si la serie que es estudiada es ruido blanco, la mayoría de las autocorrelaciones muestrales caerán dentro de estos límites. Los valores que caigan fuera de estos límites es una indicación de no aleatoriedad, aunque para un proceso de ruido blanco, cerca de una en veinte autocorrelaciones será significante.

Selección del modelo.

Cuando se modela una serie de tiempo particular p y q no se conocen. Por lo tanto, antes de realizar una estimación un modelo adecuado se debe escoger. Esto involucra escoger a p y q de tal

forma que una buena adaptación se obtiene con un número mínimo de parámetros. Box y Jenkins se refiere con esto como el principio de Parsimonia.

Algunos principios generales para la selección del modelo se darán a continuación. La aproximación óptima es empezar del modelo más general y probar sucesivamente más y más restricciones severas. Esta aproximación podría ser viable en el contexto presente si fuera razonable considerar un proceso puro AR o MA como la formulación más general. Entonces, dado un modelo AR(p), la primera hipótesis a probar es que el proceso es de orden p. La siguiente hipótesis a probar es ver si el proceso es de orden p-1 y así sucesivamente. El orden del modelo se fija cuando una hipótesis en particular se rechaza. Desafortunadamente este procedimiento tiene dos objeciones, una menor y otra mayor. La objeción menor es que el conjunto de hipótesis ordenadas que se han definido no pueden ser siempre apropiadas. El mejor modelo puede tener un orden mayor que el número de parámetros distintos de cero; por ejemplo, un modelo AR(2) puede tener $\phi_1 = 0$. Sin embargo, excepto en los casos estacionales, tales formulaciones no son comunes, y una vez que p se ha determinado no es difícil detectar coeficientes igual a cero en algún caso. La objeción mayor básica es que puede tomar un número relativamente grande de parámetros AR para aproximar un proceso lineal particular del todo bien. Esto tiene dos implicaciones. Primero, el modelo inicial general se debe escoger con un orden suficientemente grande. Segundo, una vez que el modelo se selecciona por un procedimiento de pruebas secuencial bien podría contener un número muy grande de parámetros, violando de esta forma el principio de parsimonia. Un modelo mixto ARMA tiene la ventaja que puede usualmente aproximar un modelo lineal con relativamente pocos parámetros. El problema con los modelos mixtos es que no hay un orden natural de las hipótesis. Aún más, una dificultad técnica surge cuando tal modelo está sobreparametrizado. Como un ejemplo simple, supóngase que un modelo ARMA(1,1)

$$z_t = \phi_1 z_{t-1} + \theta_1 a_{t-1} + a_t$$

se adapta a un conjunto de observaciones de un proceso de ruido

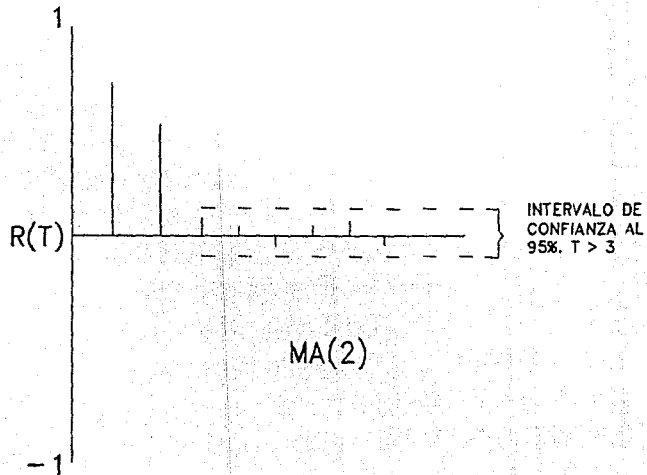
blanco. El modelo anterior se reduce a ruido blanco cuando $\phi = -\theta$, pero cuando esta limitante se impone la matriz de información asintótica es singular. Esto puede provocar problemas computacionales, desde que los procedimientos para el cálculo de estimadores de máxima verosimilitud se basan en una matriz la cual es asintóticamente equivalente a la matriz de información. Si la rutina no falla, pero converge a estimar que $\phi = -\theta$, una importante pista de sobreparametrización caerá en los errores estándar estimados, los cuales debido a la cercana singularidad de la matriz de información estimada tenderán a ser altos.

Las consideraciones anteriores conducen a Box y Jenkins a establecer una estrategia para la selección del modelo basada en un procedimiento iterativo de tres etapas. En la primer etapa, la identificación, un modelo tentativo se selecciona en la base de la apariencia del correlograma y otras estadísticas relacionadas. Dados p y q , entonces los parámetros se estiman. Como un resultado de la etapa de estimación, se calculan los residuales y estos se pueden usar para la etapa de diagnóstico, una desviación significativa de la aleatoriedad indica que el modelo es, en algún sentido, no adecuado. Un regreso a la etapa de identificación se debe hacer, y el ciclo completo se repite hasta que una formulación adecuada se encuentra.

Identificación

El correlograma provee el medio más importante por el que un proceso ARMA (p, q) adecuado se puede identificar. Tomando en cuenta que la muestra es razonablemente grande, el correlograma deberá mostrar la función de autocorrelación teórica del proceso a tratar. Anteriormente el comportamiento de la función de autocorrelación se examinó y ciertas características se notaron para procesos particulares. Entonces, un proceso puro MA(q) exhibe un corte en su función de autocorrelación para retrasos mayores que q . Por el otro lado, la autocorrelación de un proceso AR o mixto gradualmente desciende a cero. Un correlograma del tipo mostrado en la figura podría conducir a pensar que es un modelo MA(2). Mientras que las autocorrelaciones después de $r=2$

U.N.A.M.
E.N.E.P. ACATLAN
IDENTIFICACION DEL MODELO



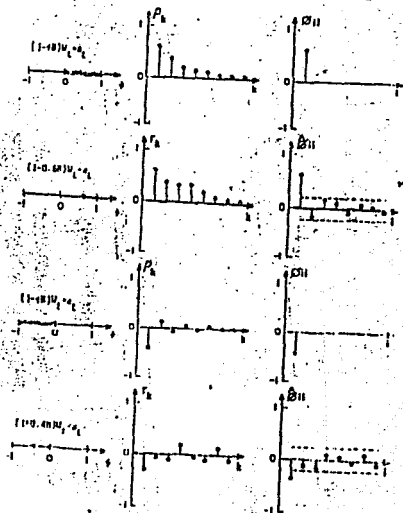
REFERENCIA: APUNTES DE SERIES DE TIEMPO CON EL PROFESOR VICTOR ESPINOZA.

FUNCIÓN DE AUTOCORRELACIÓN Y DE AUTOCORRELACIÓN PARCIAL
TEÓRICAS Y MUESTRALES PARA MODELOS ADMISIBLES AR (1)

MODELO

FAC

FACP

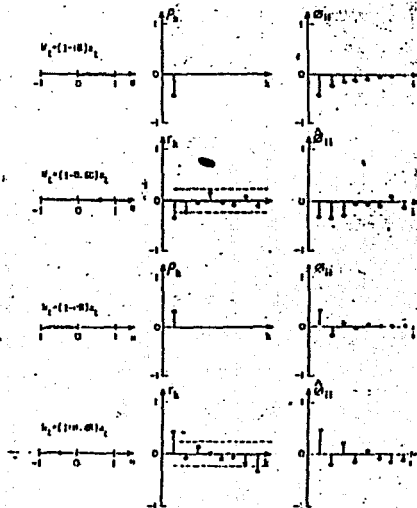


FUNCIÓN DE AUTOCORRELACIÓN Y DE AUTOCORRELACIÓN PARCIAL
TEÓRICAS Y MUESTRALES PARA MODELOS ADMISIBLES MA (1)

MODELO

FAC

FACP



ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

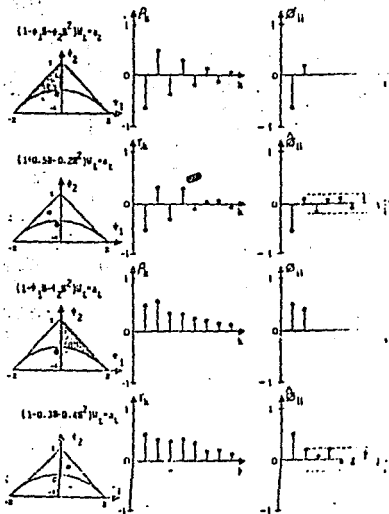
REFERENCIA: APUNTES DE SERIES DE TIEMPO CON EL PROFESOR VICTOR ESPINOZA

FUNCIONES DE AUTOCORRELACION Y DE AUTOCORRELACION PARCIAL TEORICAS Y MUESTRALES PARA MODELOS ADMISIBLES AR (2)

MODELO

FAC

FACP

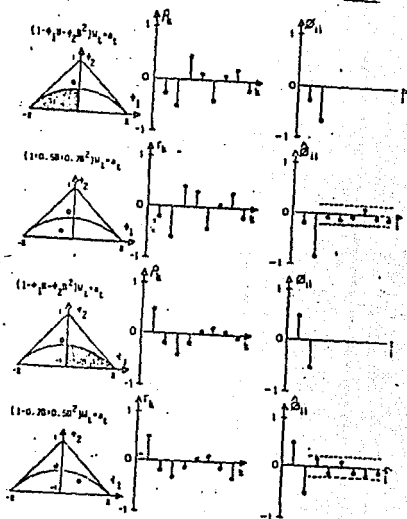


FUNCIONES DE AUTOCORRELACION Y DE AUTOCORRELACION PARCIAL TEORICAS Y MUESTRALES PARA MODELOS ADMISIBLES AR (2)

MODELO

FAC

FACP



REFERENCIA: APUNTES DE SERIES DE TIEMPO CON EL PROFESOR VICTOR ESPINOZA.

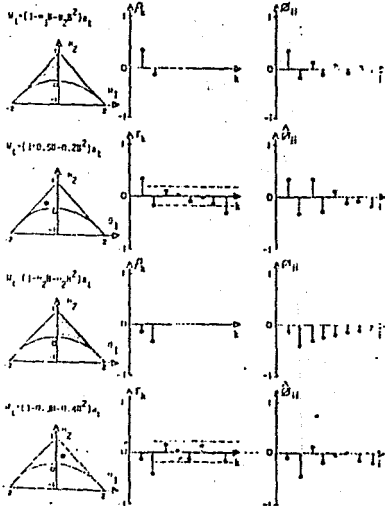
26

FUNCIÓN DE AUTOCORRELACIÓN Y DE AUTOCORRELACIÓN PARCIAL
TEÓRICAS Y MUESTRALES PARA MODELOS ADMISIBLES MA (2)

MODELO

FAC

FACP

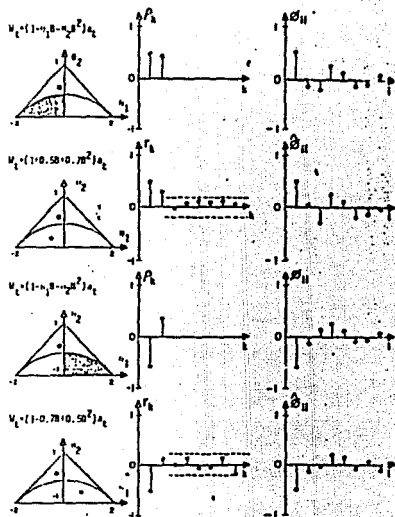


FUNCIÓN DE AUTOCORRELACIÓN Y DE AUTOCORRELACIÓN PARCIAL.
TEÓRICAS Y MUESTRALES PARA MODELOS ADMISIBLES MA (2)

MODELO

FAC

FACP



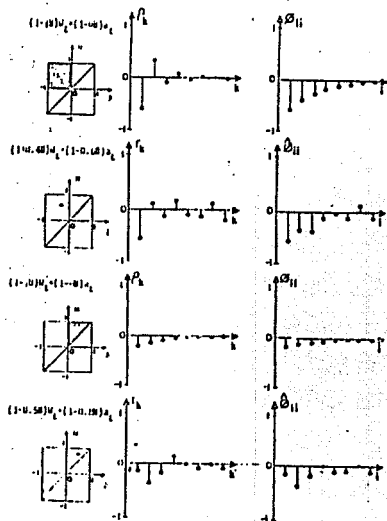
REFERENCIA: APUNTES DE SERIES DE TIEMPO CON EL PROFESOR VICTOR ESPINOZA

FUNCIONES DE AUTOCORRELACION Y DE AUTOCORRELACION INICIAL
TEORICAS Y MUESTRALES PARA MODELOS ADMISIBLES ARMA (1,1)

MODELO

FAC

FACP

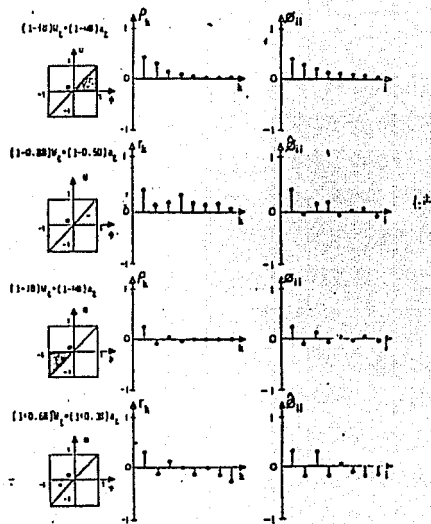


FUNCIONES DE AUTOCORRELACION Y DE AUTOCORRELACION INICIAL
TEORICAS Y MUESTRALES PARA MODELOS ADMISIBLES ARMA (1,1)

MODELO

FAC

FACP



no son idénticamente iguales a cero, son suficientemente pequeños en comparación con $r(1)$ y $r(2)$ para sugerir que su tamaño se debe a un error de muestreo.

Si las autocorrelaciones teóricas más allá de cierto punto, q , son cero, se puede demostrar que las autocorrelaciones muestrales son aproximadamente distribuidas normalmente con media cero y varianza

$$\text{Avar} [r(\tau)] = \frac{1 + 2 \sum_{j=1}^q \rho^2(j)}{T} \quad \tau > q$$

Un estimado de $\text{Avar} [r(\tau)]$ se puede obtener remplazando las autocorrelaciones teóricas en la ecuación anterior por las correspondientes autocorrelaciones muestrales. Las líneas punteadas en la figura indican intervalos al 95% para cada $r(\tau)$, $\tau > 2$, bajo el supuesto de que el proceso es en realidad MA(2).

El orden de un proceso puro AR es más difícil de determinar del autocorrelograma, excepto tal vez cuando $p=1$. Un procedimiento complementario se lleva a cabo para la identificación. Este se basa en la función de autocorrelaciones muestrales parcial. Denótese a $\hat{\phi}(\tau)$ los coeficientes estimados de $z_{t-\tau}$ en un modelo AR(τ). La función de autocorrelación muestral parcial se define por una gráfica de $\hat{\phi}(\tau)$ contra τ . El punto importante a notar a cerca de esta función es que su comportamiento es lo opuesto a lo que muestra la función de autocorrelación. Para un proceso AR puro, las autocorrelaciones parciales teóricas son cero en retrasos más allá de p , mientras que para un proceso MA estas descienden gradualmente. Si las observaciones se generan por un proceso AR(p), las autocorrelaciones parciales muestrales que están más allá de p , se distribuyen normalmente con media cero y varianza

$$\text{Avar} [\hat{\phi}(\tau)] = 1/T$$

Si los datos se pueden aproximar razonablemente bien por un modelo AR o MA puro, tampoco debe de provocar dificultad alguna

seleccionar un valor adecuado de p o q examinando la función muestral de autocorrelación o autocorrelación parcial. La identificación tanto de p como de q en un modelo mixto es de alguna forma más problemática. Ni la función de autocorrelación, ni la función de autocorrelación parcial tienen un punto de corte definido, y una considerable habilidad se puede necesitar para interpretar los patrones obtenidos. En estas circunstancias el ciclo completo de la identificación, estimación y diagnóstico pueden ser repetidas varias ocasiones antes de que un modelo adecuado se pueda encontrar.

Las dificultades anteriores se pueden componer por error de muestreo.

De acuerdo al correlograma y al correlograma parcial de las autocorrelaciones de la serie que se está estudiando, se puede notar que es un modelo MA(1) o un AR(6).

Se observa en el correlograma parcial que el valor más alto se obtiene en $r(1)$ y los demás valores van tendiendo a uno, es por ello que se supone que el modelo es un MA(1)

En cuanto a la sospecha de un modelo AR(6), se basa en el correlograma parcial, donde los valores de $r(1)$, $r(2)$, $r(3)$, $r(4)$, $r(5)$, $r(6)$ son distintos de cero, y de $R(6)$ en adelante los valores de $R(t)$ van desvaneciéndose.

Una vez identificado el modelo, se procederá a estimar los parámetros ϕ 's y/o θ 's.

La mayoría de los métodos usados para estimar los parámetros en los modelos de series de tiempo se basan en el principio de máxima verosimilitud. La ventaja de adoptar un procedimiento de máxima verosimilitud es que los estimadores resultantes son estadísticamente eficientes en muestras grandes. Por otro lado, algún tipo de procedimiento iterativo será usualmente necesario para calcularlos, como es el caso cuando se utiliza una computadora y un paquete estadístico, como es el caso que se está tratando. Por ejemplo en el paquete estadístico SYSTAT utiliza un método de optimización numérica como es el de GAUSS-JORDAN.

Máxima verosimilitud.

La función de densidad conjunta de las observaciones z_1, \dots, z_T se supone que depende de un conjunto de n parámetros desconocidos en el vector $\psi = (\psi_1, \dots, \psi_n)$ que se denotará por $L(z_1, \dots, z_T; \psi)$. Una vez que la muestra ha sido dibujada, $L(z_1, \dots, z_T; \psi)$ se puede reinterpretar como una función de verosimilitud que indica la plausibilidad de valores diferentes de ψ , dada la muestra. Entonces la verosimilitud es una función de ψ , y el estimador de máxima verosimilitud, $\hat{\psi}$, da el valor de ψ el cual maximiza $L(\psi)$.

El método usual de encontrar el estimador de máxima

verosimilitud es diferenciar la función logarítmica de verosimilitud, $\text{Log}(L)$, con respecto a cada uno de los parámetros $\psi_1, \psi_2, \dots, \psi_m$. Igualando las derivadas a cero, da como resultado las ecuaciones de verosimilitud

$$\frac{\delta \text{Log}(L)}{\delta \psi} = 0$$

Como una regla, las ecuaciones de verosimilitud son no lineales y así los estimados de máxima verosimilitud se deben calcular por algún procedimiento iterativo.

Se supondrá que las observaciones están normalmente distribuidas. El principio de la estimación por máxima verosimilitud conduce a procedimientos que están esencialmente basados en minimizar una función de suma de cuadrados de residuales. También se supondrá que el modelo verdadero es tanto estacionario como invertible. Sin embargo es importante notar que los modelos en los cuales algunas de las raíces del polinomio MA caen dentro del círculo unitario son perfectamente válidas y pueden surgir en la práctica cuando una sobrediferenciación ha sido hecha. Como un simple ejemplo suponga que $z_t = a_t$ donde $a_t \sim \text{NID}(0, \sigma^2)$. Entonces

$$\Delta z_t = a_t + \theta a_{t-1}$$

donde $\theta = -1$, y así Δz_t se genera por un proceso MA(1) el cual es, estrictamente hablando, no invertible. Aunque esto es sensible a demandar que los estimados de los parámetros AR deben satisfacer las condiciones estacionarias, lo mismo no es necesariamente cierto en cuanto a la invertibilidad.

La función exacta de verosimilitud se construye del error de la descomposición del error de predicción, que es un resultado fundamental en el análisis de series de tiempo que permite escribir la función de densidad conjunta de las observaciones de una forma manejable.

Considérese un conjunto de T observaciones dependientes con media μ y matriz de covarianzas, $\sigma^2 V$, obtenida de una

distribución normal multivariada, i.e. $z \sim N(\mu, \sigma^2 V)$. La función conjunta de densidad de las observaciones es

$$\begin{aligned} \log L(z) = & -(T/2) \log 2\pi - (T/2) \log \sigma^2 - (1/2) \log |V| \\ & - 1/2 \sigma^{-2} (z-\mu)' V^{-1} (z-\mu) \end{aligned}$$

Se puede factorizar en dos partes escribiendo

$$\log L(z) = \log L(z_1, \dots, z_{T-1}) + \log l(z_T/z_{T-1}, \dots, z_1)$$

El segundo término del lado derecho de la ecuación anterior es la distribución de la última observación, condicional a todas las observaciones previas. Escribir la función de densidad conjunta de esta forma se obtiene de la definición de función de densidad de probabilidad condicional. Si $p(x, z)$ es la función de densidad conjunta de dos variables aleatorias, x y z , entonces

$$p(x, z) = f(x/z)h(z)$$

donde $f(x, z)$ es la función de densidad de probabilidad de x condicional a z y $h(z)$ es la función de densidad de probabilidad marginal de z .

Ahora considérese el problema de estimar z_T , dado que z_{T-1}, \dots, z_1 se conocen. Si $\hat{z}_{T/T-1}$ es un estimador de z_T construido sobre esta base, el error de estimación, o predicción, se puede dividir en dos partes:

$$\hat{z}_{T/T-1} - z_T = [z_T - E(z_{T/T-1}, \dots, z_1)] + [E(z_{T/T-1}, \dots, z_1) - \hat{z}_{T/T-1}]$$

donde $E(z_{T/T-1}, \dots, z_1)$ es la media de la distribución de z_T condicional sobre z_{T-1}, \dots, z_1 . Se sigue de la ecuación anterior que el estimador de la media cuadrática de $\hat{z}_{T/T-1}$, denotado como $MSE(\hat{z}_{T/T-1})$ es:

$$\text{MSE}(\bar{z}_{T/T-1}) = \text{Var}(z_{T/T-1}, \dots, z_1) \\ + E\{[\bar{z}_{T/T-1} - E(z_{T/T-1}, \dots, z_1)]^2\}$$

El primer término del lado derecho de la ecuación anterior es independiente de $z_{T/T-1}$, y así el estimador mínimo de la media cuadrática de z_T condicional a z_{T-1}, \dots, z_1 es

$$\bar{z}_{T/T-1} = E(z_{T/T-1}, \dots, z_1)$$

La varianza del error de predicción asociado con $\bar{z}_{T/T-1}$ es $E\{[z_T - \bar{z}_{T/T-1}]^2\}$, esto es idéntico a $\text{MSE}(\bar{z}_{T/T-1})$, el cual en vista de la construcción de $\bar{z}_{T/T-1}$ es igual a

$\text{Var}(z_{T/T-1}, \dots, z_1)$, esta cantidad se denotará por $\sigma^2 f_T$.

Desde que las observaciones se distribuyen normalmente ambos componentes de la ecuación

$$\log L(z) = \log L(z_1, \dots, z_{T-1}) + \log l(z_T/z_{T-1}, \dots, z_1)$$

son normales, el segundo término se puede escribir

$$\log l(z_T/z_{T-1}, \dots, z_1) = - (1/2) \log 2\pi - (1/2) \log \sigma^2 \\ - (1/2) \log f_T - (1/2) \sigma^{-2} (z_T - \bar{z}_{T/T-1})^2 / f_T$$

y se interpreta como la distribución del error de predicción $z_T - \bar{z}_{T/T-1}$.

La descomposición de

$$\log L(z) = \log L(z_1, \dots, z_{T-1}) + \log l(z_T/z_{T-1}, \dots, z_1)$$

se puede repetir con respecto a la verosimilitud de las primeras T-1 observaciones, y repetir las hasta que se obtiene

$$\log L(z) = \sum_{t=2}^T \log L(z_t/z_{t-1}, \dots, z_1) + \log l(z_1)$$

Para T=2, ..., T la media de z_t condicional sobre z_{t-1} debe ser igual a $\bar{z}_{t|t-1}$, el estimador de la media cuadrática mínima de z_t dadas las observaciones anteriores. Cada una de las distribuciones condicionales es por lo tanto la distribución del error asociado con el predictor óptimo, mientras que $l(z_1)$ es la distribución incondicional de z_1 . Sin embargo, si μ_1 se considera como el estimador de la media cuadrática mínimo de z_1 , dado que no hay observaciones previas, el término $z_1 - \mu_1$ es el error de predicción asociado con z_1 . Es por lo tanto apropiado denotar la varianza de z_1 por $\sigma^2 f_1$.

La expresión

$$\log L(z) = \sum_{t=2}^T \log L(z_t/z_{t-1}, \dots, z_1) + \log l(z_1)$$

permite a la función de máxima verosimilitud descomponerse en la distribución conjunta de T errores de predicción independientes

$$v_t = z_t - \bar{z}_{t|t-1}, \quad t = 1, \dots, T$$

donde $\bar{z}_{1|0} = \mu_1$. Cada error de predicción tiene media cero y varianza $\sigma^2 f_t$ y así

$$\begin{aligned} \log L(z) = & -(T/2) \log 2\pi - (T/2) \log \sigma^2 - (1/2) \log |V| \\ & - 1/2 \sigma^{-2} (z-\mu)' V^{-1} (z-\mu) \end{aligned}$$

se convierte

$$\log L(z) = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} \sigma^2 \sum_{t=1}^T v_t^2 / f_t$$

La descomposición del error de predicción se puede interpretar en términos de una descomposición de Cholesky de V^{-1} . Si \bar{L} es una matriz triangular inferior con unos en la diagonal principal, V^{-1} se puede factorizar como $V^{-1} = \bar{L}' D \bar{L}$, donde $D = \text{diag}(f_1^{-1}, \dots, f_T^{-1})$. Esta factorización es única y los errores de predicción en $\nu_t = z_t - \bar{z}_{t|t-1}$ se dan por la transformación $\nu = \bar{L} z$. Como el jacobiano de esta transformación es $|\bar{L}| = 1$, la distribución conjunta de los elementos en el vector $\nu = (\nu_1, \dots, \nu_T)'$ es como en

$$\log L(z) = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} \sigma^2 \sum_{t=1}^T \nu_t^2 / f_t$$

Notese que $|V^{-1}| = |\bar{L}'| |D| |\bar{L}| = |D|$ y así

$$\log |V| = \sum \log f_t$$

La descomposición de Cholesky se puede usar como la base de factorizar V^{-1} numéricamente y calcular los errores de predicción directamente de la transformación $\bar{L}z$.

Simplificaciones considerables pueden ser usualmente obtenidas en la descomposición del error de predicción si se observa que una aproximación a la función de verosimilitud se adecua en la práctica. Como una regla general, una aproximación se puede justificar en campos teóricos haciendo algunas suposiciones acerca de las observaciones iniciales. Esto usualmente resulta en errores de predicción con varianza constante, σ^2 , y así para $f_t = 1$ para toda t , la función de verosimilitud se convierte

$$\log L = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sigma^{-2} \sum a_t^2$$

donde el cambio en la notación de ν_t a a_t sirve para indicar que cada error de predicción coincidirá con el disturbio del ruido

blanco del modelo. Si la función de verosimilitud se puede reducir a una forma como la anterior, hay importantes implicaciones para la estimación. Maximizar la función de verosimilitud es equivalente a minimizar la función de sumas de cuadrados:

$$S(\psi) = \sum a_i^2$$

donde ψ denota todos los parámetros del modelo aparte de σ^2 .

El modelo autorregresivo de primer orden

$$z_t = \phi z_{t-1} + a_t, \quad |\phi| < 1$$

donde $a_t \sim \text{NID}(0, \sigma^2)$ se dio a conocer anteriormente. Construyendo la función de verosimilitud, se supondrá que el proceso comienza en cualquier punto del pasado, pero solamente se observa en los tiempos $t = 1, \dots, T$.

Para cualquier t , la distribución de z_t , dadas todas las observaciones anteriores, tiene una media de ϕz_{t-1} y una varianza de σ^2 . Entonces las últimas $T-1$ errores de predicción son idénticos a los últimos $T-1$ disturbios desde que

$$v_t = z_t - \bar{z}_{t|t-1} = z_t - \phi z_{t-1} = a_t, \quad t = 2, \dots, T$$

Correspondiendo a la ecuación

$$\text{Log } L(z) = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} \sigma^2 \sum_{t=1}^T v_t^2 / f_t$$

la función logarítmica de verosimilitud se puede descomponer en las funciones de densidad de a_2, \dots, a_T más la función de densidad incondicional de z_1 . Esta tiene distribución normal con media cero y varianza $\sigma^2 / (1-\phi^2)$. El primer error de predicción es por lo tanto igual a z_1 , en vez de que sea igual al disturbio a_1 . Poniendo $v_1 = z_1$, la función logarítmica de verosimilitud para todas las observaciones T se pueden escribir de la forma

$$\text{Log } L(z) = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sum_{t=1}^T \log f_t - \frac{1}{2} \sigma^2 \sum_{t=1}^T v_t^2 / f_t$$

i.e.

$$\begin{aligned} \log L(z) = & -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 + \frac{1}{2} \sum_{i=1}^T \log (1-\phi^2) \\ & - \frac{1}{2} \sigma^{-2} (1-\phi^2) z_1^2 - \frac{1}{2} \sigma^{-2} \sum_{i=2}^T (z_i - \phi z_{i-1})^2 \end{aligned}$$

Nótese que $f_1 = 1$ para $t = 2, \dots, T$, mientras que $f_1 = (1-\phi^2)^{-1}$.

El tercer y cuarto término de la ecuación anterior tienden a cero asintóticamente, y si no se toman en cuenta en la función de verosimilitud se simplifica a la forma

$$\log L = -\frac{T}{2} \log 2\pi - \frac{T}{2} \log \sigma^2 - \frac{1}{2} \sigma^{-2} \sum z_i^2$$

Una justificación formal de hacer esto es considerar a z_1 como un valor fijo. El estimador de máxima verosimilitud de ϕ es lineal, siendo obtenido por una regresión de z_i sobre z_{i-1} .

Como se comentó anteriormente, la función de verosimilitud exacta se construye de la descomposición del error de predicción. Esto representa alguna desviación de la práctica usual. Sin embargo, la descomposición del error de predicción es una aproximación natural desde un punto de vista teórico. Aparte de clarificar la estructura de la función de verosimilitud exacta, hace completamente aparente la naturaleza de la aproximación envuelta en la función de verosimilitud condicional. La función de verosimilitud condicional se obtiene fijando ciertos disturbios iniciales y/o observaciones, y maximizar esta función equivale a minimizar una función de suma de cuadrados correspondiente. Los estimadores resultantes, los cuales se conocen como estimadores de suma de cuadrados condicionales, se toman generalmente para servir como aproximaciones adecuadas a los estimadores de la función de máxima verosimilitud exacta.

Una alternativa a la máxima verosimilitud es basar los estimados en el correlograma. Para un modelo AR puro, los estimados que se toman del correlograma son completamente eficientes. Aunque esto no sucede para el caso de modelos MA puro y mixtos, los estimadores obtenidos del correlograma pueden

proveer útiles valores iniciales para un procedimiento iterativo de optimización.

Modelo Autorregresivos.

La función de verosimilitud para un proceso AR(1) es derivado anteriormente. La misma técnica se puede utilizar para un proceso AR(p), el primer paso es llevar a cabo la descomposición del error de predicción

$$\log L(z) = \sum_{t=p+1}^T \log(z_t/z_{t-1}, \dots, z_1) + \log L(z_p)$$

El primer término del lado derecho de la ecuación anterior se puede interpretar como el logaritmo de la distribución conjunta de a_{p+1}, \dots, a_T , mientras que $L(z_p)$ es la distribución conjunta de las primeras p observaciones, $z_p = (z_1, \dots, z_p)'$,

Si la matriz de covarianzas de z_p se denota por $\sigma^2 V_p$, la función logarítmica de verosimilitud se puede escribir como

$$\log L = -(1/2) T \log 2\pi - (1/2) T \log \sigma^2 - (1/2) \log |V_p| \\ - (1/2) \sigma^{-2} \left[z_p' V_p^{-1} z_p + \sum_{t=p+1}^T (z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p})^2 \right]$$

Estimación exacta de máxima verosimilitud

El parámetro σ^2 se puede concentrar fuera de la función de verosimilitud. Sin embargo la función resultante es todavía no

lineal en ϕ , y la estimación de máxima verosimilitud se debe llevar a cabo por una optimización numérica. La matriz V_p^{-1} se puede obtener por una serie de métodos disponibles.

Para el modelo AR(1), dejando $\frac{d}{d\phi} \log L = 0$ da como resultado una ecuación de tercer grado sobre ϕ . Esta tiene una raíz real única en el rango $[-1, 1]$.

Regresión

La estimación de modelos AR se puede simplificar considerablemente considerando las primeras p observaciones, z_1, \dots, z_p , como fijas. Esto provee una justificación teórica por la cual el término $\log L(z_p)$ no se tomó en cuenta, con el resultado de que la maximización de la función de verosimilitud se vuelve equivalente a minimizar la función de suma de cuadrados

$$s(\phi) = \sum_{t=p+1}^T (z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p})^2$$

El estimador de máxima verosimilitud de ϕ por lo tanto se obtiene por una regresión de mínimos cuadrados ordinarios (OLS) de z_t sobre sus valores z_{t-1}, \dots, z_{t-p} .

En muestras grandes se hace muy pequeña la diferencia si la estimación se lleva a cabo por la función de verosimilitud exacta o por regresión. En el modelo AR(1), se puede observar que la única característica distintiva de la función exacta de verosimilitud es la función que involucra a z_1^2 y $\log(1-\phi^2)$. Estos términos no se toman en cuenta por el resto de la función de verosimilitud si T es suficientemente grande, y la distribución asintótica de los estimadores de ϕ y σ^2 no es afectada si se omiten.

La introducción de un media distinta de cero en un modelo

AR(1), resulta de que la función de sumas de cuadrados se convierte

$$S(\phi) = \sum_{t=2}^T [z_t - \mu - \phi(z_{t-1} - \mu)]^2$$

Diferenciando con respecto a μ da como resultado

$$\hat{\mu} = \frac{\sum_{t=2}^T z_t - \phi \sum_{t=1}^{T-1} z_t}{(T-1)(1-\phi)} \approx \bar{z}$$

Esta μ se puede estimar independientemente de ϕ

Las ecuaciones de Yule-Walker

Los estimadores de los parámetros sobre ϕ se pueden obtener del correlograma. Para un modelo AR(p) puro, las autocorrelaciones se dan por la ecuación en diferencias de orden p

$$\rho(\tau) = \phi_1 \rho(\tau-1) + \dots + \phi_p \rho(\tau-p) \quad \tau = 1, 2, \dots$$

Escribiendo esta ecuación para $\tau = 1, \dots, p$ da

$$\begin{bmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix}$$

Estas expresiones se conocen como las ecuaciones de Yule-Walker. Reemplazando a $\rho(\tau)$ por $r(\tau)$ resulta un conjunto de ecuaciones lineales las cuales se pueden resolver directamente para obtener estimados de ϕ_1, \dots, ϕ_p .

Procesos de Medias Móviles y Mixtos

La función de verosimilitud para cualquier proceso ARMA(p,q) se puede construir de la descomposición del error de predicción. Encontrar la función de verosimilitud de alguno de estos procesos puede tomar mucho tiempo, particularmente si $p + q$ es grande.

Si algunas suposiciones se hacen acerca de los valores iniciales de los disturbios y/u observaciones, una función de verosimilitud condicional se obtiene. En el caso del proceso AR puro conduce a un estimador de máxima verosimilitud lineal de ϕ . Para procesos MA y mixtos el estimador de máxima verosimilitud es todavía no lineal, pero el cálculo de la función de verosimilitud se simplifica considerablemente. Aún más, derivados analíticos son disponibles y estos pueden ser importantes para mejorar la eficiencia del procedimiento de optimización.

La suma de cuadrados condicional para el modelo MA(1)

El modelo MA(1)

$$z_t = a_t + \theta a_{t-1}$$

provee la ilustración más sencilla de las técnicas usadas en la estimación de modelos. Supóngase, por el momento, que θ se conoce. Si al tiempo $t-1$, a_{t-1} se conoció, el estimador de la media cuadrática mínimo (MMSE) de z_t podría ser $\hat{z}_{t|t-1} = \theta a_{t-1}$ y el error de predicción asociado sería

$$a_t = z_t - \theta a_{t-1}$$

Los errores de predicción futuros, cada uno idéntico al término de disturbio correspondiente se puede calcular de la recursión de la expresión anterior. Desafortunadamente hay una dificultad con esta aproximación que es la de que al calcular el

conjunto completo de los errores de predicción, a_1, \dots, a_T , el disturbio inicial, a_0 , se debe conocer. Una solución al problema es suponer que a_0 es igual a cero. El conjunto completo de los errores de predicción se pueden entonces calcular, y si se sigue suponiendo que a_0 es fijo en cero para todas las realizaciones, la función logarítmica de verosimilitud es de la forma

$$\log L = -(T/2) \log 2\pi - (T/2) \log \sigma^2 - (1/2) \sigma^{-2} \sum a_i^2$$

Para casos más prácticos tiene poca importancia si la suposición $a_0 = 0$ es verdadera o no. Sin embargo, si a_0 es aleatorio, es importante notar que la aproximación implícita en la verosimilitud condicional solamente se vale si el proceso es invertible, i.e. si $|\theta| < 1$. De la misma forma como se mostró en este ejemplo, la diferencia entre las funciones de verosimilitud exacta y condicional se convierte negligible en la medida que T aumenta.

Maximizando la función de verosimilitud exacta es equivalente a minimizar la función de suma de cuadrados condicional $S(\theta) = \sum a_i^2$ con respecto a θ . Es importante notar que a_i ya no es más un disturbio, sino que es un residual cuyo valor depende de θ . La notación $a_i(\theta)$ confunde este punto. Sin embargo, usualmente no será necesario hacer esta modificación, desde que la interpretación de a_i será clara del contexto.

Para un valor dado de T , un conjunto de T residuales se pueden generar de la recursión

$$a_t = z_t - \theta a_{t-1} \quad t = 1, \dots, T$$

con $a_0 = 0$. Los residuales se pueden usar para construir $S(\theta)$ y el problema que falta es encontrar un método adecuado para minimizar esta función con respecto a θ . Como solamente un solo parámetro se involucra, una búsqueda sobre el rango de $[-1, 1]$ se puede llevar a cabo. Para procesos más generales esta suposición puede ser no muy viable y el algoritmo obvio a adoptar es el de GAUSS-NEWTON, o una modificación adecuada de esta. Para el modelo MA(1), diferenciar la ecuación anterior produce

$$\frac{\partial a_t}{\partial \theta} = -\theta \frac{\partial a_{t-1}}{\partial \theta} - a_{t-1} \quad t = 1, \dots, T$$

Como $a_0 = 0$ se obtiene que $\frac{\partial a_t}{\partial \theta} = 0$, entonces las derivadas se producen por una recursión paralela a

$$a_t = z_t - \theta a_{t-1} \quad t = 1, \dots, T$$

con la inicialización manejada de una forma similar. Dado el

estimado de θ , el algoritmo procede a calcular a_t y $\frac{\partial a_t}{\partial \theta}$ y

entonces actualiza el estimado de la regresión de a_t sobre $\frac{\partial a_t}{\partial \theta}$

La suma de cuadrados condicional en el caso general.

Para modelos de mayor orden MA, la función condicional de verosimilitud se obtiene tomando a_{1-q}, \dots, a_0 iguales a cero en todas las realizaciones. Los residuales usados para calcular la función de suma de cuadrados condicional se obtiene de la recursión

$$a_t = z_t - \theta_1 z_{t-1} - \dots - \theta_q z_{t-q} \quad t = 1, \dots, T$$

con $a_{1-q} = a_{2-q} = \dots = a_0 = 0$

Procedimientos similares se pueden adoptar para modelos mixtos, aunque en tales casos existe el problema adicional de manejar las observaciones iniciales. Considérese el modelo ARMA(1,1). $z_t = \phi z_{t-1} + a_t + \theta a_{t-1}$. Si z_t se toma como un valor fijo, los errores de predicción se pueden calcular de la recursión

$$a_t = z_t - \phi z_{t-1} - \theta a_{t-1} \quad t = 2, \dots, T$$

donde $a_1 = 0$. Una aproximación alternativa podría ser comenzar la recursión en $t=1$, con $z_0 = a_0 = 0$. Sin embargo, aunque esto produce T , más que $T-1$, residuales, esto no es recomendable, debido a que al dejar arbitrariamente $z_0 = 0$ introduce una distorsión en los cálculos. En general, la aproximación adecuada para un modelo ARMA(p,q) es calcular $T-p$ errores de predicción de una recursión de la forma

$$a_t = z_t - \phi_1 z_{t-1} - \dots - \phi_p z_{t-p} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

$$t = p+1, \dots, T$$

$$\text{con } a_p = a_{p-1} = \dots = a_{p-q+1} = 0$$

Las derivadas que se necesitan para implementar el algoritmo de GAUSS-NEWTON se puede calcular por recursiones. Para el modelo ARMA(1,1):

$$\frac{\delta a_t}{\delta \phi} = -z_{t-1} - \theta \frac{\delta a_{t-1}}{\delta \phi} \quad t = 2, \dots, T$$

$$\frac{\delta a_t}{\delta \theta} = -a_{t-1} - \theta \frac{\delta a_{t-1}}{\delta \theta} \quad t = 2, \dots, T$$

con $\delta a_1 / \delta \phi = \delta a_1 / \delta \theta = 0$. En general se necesitarán de $p+q$ derivadas. Sin embargo estas se pueden obtener sobre la base de solamente dos recursiones, una para los componentes AR y otra para los componentes MA.

Estimados iniciales.

Valores iniciales se necesitan para las iteraciones de Gauss-Newton. Para un modelo MA(1) podría ser posible comenzar con $\hat{\theta} = 0$, pero una estrategia similar no podría adoptarse para un modelo mixto. Poniendo $\hat{\phi} = \hat{\theta} = 0$ en el modelo ARMA(1,1) podría dar como resultado que las dos derivadas sean idénticas. Aquí el algoritmo podría romperse inmediatamente debido a la multicolinealidad de la regresión de a_t sobre $\partial a_t / \partial \phi$ y $\partial a_t / \partial \theta$.

Una mejor forma de comenzar las iteraciones es comenzar con estimados consistentes de los parámetros. Una posibilidad es obtener estimados del correlograma. Considere un modelo MA(1). Si $\rho(1)$ se reemplaza por $r(1)$ en $\rho(1) = \theta / (1 + \theta^2)$, un estimador de θ se obtiene resolviendo la ecuación cuadrática

$$\hat{\theta}^2 r(1) - \theta + r(1) = 0$$

Esta expresión tiene dos soluciones, pero como

$$r^{-1}(1) = \hat{\theta} + \hat{\theta}^{-1}$$

una raíz es claramente la recíproca de otra. El problema de decidir cuál raíz seleccionar se resuelve por la condición de invertibilidad. El estimador es por lo tanto

$$\hat{\theta}_c = (1 - [1 - 4r^2(1)]^{1/2}) / 2r(1)$$

La otra solución, $\hat{\theta}_c$ se desecha debido a que tiene un valor absoluto mayor que 1.

El único caso cuando $r^{-1}(1)$ no tiene dos diferentes soluciones es cuando $r(1) = \pm 0.5$. El estimador es entonces $\hat{\theta}_c = \pm 1$. Si $|r(1)| > 0.5$, no hay raíz real en $r^{-1}(1)$, desde que el autocorrelación teórica de primer orden no puede exceder 0.5 para un modelo MA(1).

Una técnica similar se puede emplear para estimar los parámetros en un modelo ARMA(1,1). La función de autocorrelación se da por

$$\rho(1) = \frac{(1+\hat{\phi}) (\hat{\phi}+\theta)}{1 + \theta^2 + 2\phi\theta}$$

$$\rho(\tau) = \phi \rho(\tau - 1)$$

la cual sugiere la estimación de θ por

$$\hat{\phi}_c = r(2)/r(1)$$

Si $\rho(1)$ y ϕ se reemplazan por $r(1)$ y $\hat{\phi}_c$ en $\rho(\tau) = \phi \rho(\tau - 1)$, un estimador se da de nuevo por la solución de una ecuación cuadrática. Esta aproximación se puede generalizar a modelos de mayor orden, aunque esto comienza a complicarse.

Ahora se darán las instrucciones para que por medio del paquete estadístico SYSTAT estime los parámetros de los modelos ya identificados, que es un modelo MA(1) o un AR(6).

```
> USE YEAR3
> SAVE MA1
> ARIMA TRANSF / Q=1
```

Estas tres instrucciones sirven para estimar los parámetros del modelo MA(1) y que los residuales los salve en el archivo MA1. Para el modelo AR(6), se utilizan las siguientes instrucciones:

```
> USE YEAR3
> SAVE AR6
> ARIMA TRANSF / P=6
```

Una vez calculados los parámetros del modelo adecuado se pasa a la etapa de diagnóstico.

El diagnóstico del modelo se puede llevar a cabo examinando los residuales de las desviaciones de la aleatoriedad. Pruebas formales se pueden basar en el correlograma, aunque ninguna de estas se deben ver como sustitutos de una graficación directa de

los residuales. Citando a Box-Jenkins: "No se puede enfatizar demasiado fuerte que la inspección visual de una gráfica de los mismos residuales es un primer paso indispensable en el proceso de chequeo".

Aunque la graficación y examinado de residuales es un ejercicio extremadamente valuable, algún cuidado se debe tomar en la interpretación de estos resultados. Las pruebas asociadas con estos procedimientos frecuentemente se construyen en la suposición de que los residuales tienen las mismas propiedades como los disturbios cuando el modelo se especifica correctamente. Desafortunadamente esta suposición no es válida para los residuales de modelos ARMA, aun cuando el tamaño de la muestra es grande. Por ejemplo, si un modelo AR(1) se acomode a los datos, se puede mostrar que la autocorrelación de primer orden en los residuales tiene una varianza asintótica de σ^2/T . Este puede ser sustancialmente menor que $1/T$, la cual es la varianza de $r(1)$ para un proceso de ruido blanco. Sin embargo, para retrasos de mayor orden los sesgos en la varianza son menos serios. Esto indica el comportamiento de las autocorrelaciones de los residuales para cualquier proceso ARMA(p,q). En todos los casos, una reducción en la varianza tiende a ocurrir en los retrasos bajos. Aun más, las $r(r)$'s en estos retrasos pueden estar altamente correlacionados. Aunque estos efectos usualmente desaparecen a retrasos mas altos, seguidamente son suficientemente importantes para impartir un sesgo severo a la estadística. El resultado que estos procedimientos de prueba tienden a subestimar la significancia de aparentes discrepancias.

Aunque las pruebas asociadas con los procedimientos gráficos no son generalmente válidos en esta situación, pueden proveer útiles guías a pesar de lo anterior.

Algunas de las pruebas de los residuales se darán a conocer enseguida:

SUPUESTO: $E [a_t] = 0$

VERIFICACION:

$$t = \frac{\sqrt{n(n-1)} \bar{a}}{\sqrt{\sum (a_i - \bar{a})^2}}$$

Se rechaza $E[a_i] = 0$ si $|t| > t_{n-1}(1-\alpha/2)$

Si esto ocurre, existe una tendencia determinística o aleatoria. Si es aleatoria, se puede requerir de una diferencia adicional y como

$$\forall z_i = z_i - z_{i-1} \sim \text{AR}(1) \quad \phi = 1$$

tal vez solo requiera un componente AR extra.

Si no es aleatoria, deberá incluirse una constante δ en el modelo, que podría ser $\delta = \bar{a}$

SUPUESTO: $\text{Var}(a_i) = \sigma^2$

VERIFICACION: Gráfica de residuales contra el tiempo.

Si no es constante, revisar o corregir la transformación estabilizadora.

SUPUESTO: $a_i \sim N$

VERIFICACION: Coeficiente de asimetría y curtosis
95% de las observaciones en $[\mu \pm 2\sigma]$

$$\text{Coeficiente de asimetría: } \frac{1}{n} \sum_{i=1}^n \frac{(\hat{a}_i - \bar{a})^3}{\sigma^3} \approx 0$$

$$\text{Coeficiente de curtosis: } \frac{1}{n} \sum_{i=1}^n \frac{(\hat{a}_i - \bar{a})^4}{\sigma^4} \approx 3$$

Si no cumple con el supuesto se deberá hacer una transformación normalizante

$$z_t^{(\lambda)} = \begin{cases} \frac{z_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln z_t & \lambda = 0, z_t > 0 \end{cases}$$

SUPUESTO: $\text{Cov}(a_t, a_{t+k}) = 0$ equivale a $H_0: \rho_k = 0 \forall k$

VERIFICACION: En la gráfica de residuales contra el tiempo no hay comportamiento
Correlograma, todos los $r_k \cong 0$, con la prueba de Bartlett

$$\text{Var}(r_k) \cong \frac{1}{n} \left[1 + 2 \sum_{j=1}^q r_j^2 \right] \text{ donde } q = 0$$

$$\ast \text{Var}(r_k) \cong \frac{1}{n}$$

Sin embargo, $\text{Var}(r_k)$ es un poco menor que $\frac{1}{n}$ para k pequeños ($k \leq 5$), por lo cual se requiere una prueba conjunta, llamada de Box Ljung

$$Q(k) = (n-d-p)(n-d-p+2) \sum_{j=1}^k \frac{r_j^2}{(n-d-p-j)}$$

$$Q(k) \sim \chi_{(k-p-q)}^2 \text{ conviene } k > 20 \text{ ya que } p-q=0$$

Se rechaza si $Q(k) > \chi_{(1-\alpha), (k-p-q)}^2$

SUPUESTO: NO ABERRANTES O DISCREPANTES

Se verifica con la gráfica de residuales contra el tiempo, donde el 99% de las observaciones esta dentro de $\pm 3\sigma$

Se debe revisar caso por caso y tal vez se requiera un componente estacional.

SUPUESTO: $E[a_t, z_{t-k}] = 0 \quad k > 0$

VERIFICACION: Por medio del correlograma cruzado se debe observar el siguiente comportamiento:

$$\hat{\rho}_{z_1 z_2}(k) = \frac{\text{Cov}(a_t, z_{t+k})}{\sqrt{\text{Var}(a_t) \text{Var}(z_t)}} \quad k = \dots -2, -1, 0, 1, 2, \dots$$

$$r_{z_1 z_2}(k) \cong 0 \quad k < 0$$

$$r_{z_1 z_2}(k) \neq 0 \quad k \geq 0$$

$$r_{z_1 z_2}(k) > r_{z_2 z_1}(k) \quad k > 0$$

Si no sigue el comportamiento anterior debe incluirse otros factores en el modelo. Ej:

Modelo correcto ARMA(1,1) pero se modela MA(1)

$$z_t = \theta_1 a_{t-1} + a_t \quad \text{donde } a_t = \phi_1 z_{t-1} + a_t$$

entonces $r_{z_1 z_2}(k)$ no seguirá el comportamiento mas adecuado pues a_t aun no es ruido blanco.

Los siguientes supuestos se refieren al modelo estimado:

ADMISIBILIDAD.- Los parámetros estimados deben caer en las regiones que lo hacen estacionario e invertible.

Por ejemplo:

$$\text{AR}(1) \quad |\phi| < 1$$

$$\text{MA}(1) \quad |\theta| < 1$$

$$\text{AR}(2) \quad \phi_1 + \phi_2 < 1$$

$$\phi_1 - \phi_2 < 1$$

$$\phi_2 > -1$$

Si no caen en la región, pueden considerarse intervalos de confianza para recorrer "ligeramente" las estimaciones.

ESTACIONARIEDAD.- Las raíces de los polinomios $\Phi(B)$ y $\Theta(B)$ deben estar fuera del círculo unitario. Si alguna raíz de $\Phi(B)$ está dentro o cerca del círculo unitario, se debe diferenciar la serie. Si alguna raíz de $\Theta(B)$ está dentro o cerca del círculo unitario, la serie está sobrediferenciada. Ejemplo:

$$z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

que admite una raíz igual a 1; entonces se puede factorizar

$$w_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_{q-1} B^{q-1}) a_t$$

MA(q-1)

$$\text{donde } (1-B) w_t = w_t - w_{t-1} = z_t$$

$$(1-B) w_t = z_t$$

$$w_t = z_t + w_{t-1} \quad (\text{SOBREDIFERENCIACION})$$

PARSIMONIA.- Se prueba con el intervalo $\hat{\theta}_1 \pm 1.96 \sqrt{\text{Var}(\hat{\theta}_1)}$ y si contiene al cero y es el de mayor orden, se elimina y vuelve a estimar el modelo. Si no es el de mayor orden debe analizarse la estabilidad.

ESTABILIDAD.- Si $\text{Corr}(\hat{\phi}_1, \hat{\phi}_2) \cong \pm 1$ (entre cualquier combinación de θ_1 con θ_2 o con ϕ_1 o de ϕ_1 con ϕ_2 o con θ_2), implica que puede darse una variación en el valor de uno de los parámetros estimados y ser compensada con un cambio en el otro, sin que se altere la suma de cuadrados (Inestabilidad). Por lo tanto es recomendable eliminar alguno de los parámetros, siguiendo conjuntamente el criterio para lograr parsimonia, o el de factores comunes que se verá enseguida.

MATRIZ DE VARIANZAS Y COVARIANZAS DE ALGUNOS MODELOS

$$\text{AR}(2): \hat{V}(\hat{\phi}_1, \hat{\phi}_2) \cong \begin{bmatrix} 1 - \hat{\phi}_2^2 & -\hat{\phi}_1(1 + \hat{\phi}_2) \\ -\hat{\phi}_1(1 + \hat{\phi}_2) & 1 - \hat{\phi}_1^2 \end{bmatrix} \frac{1}{n}$$

$$\text{Corr}(\hat{\phi}_1, \hat{\phi}_2) = - \frac{\hat{\phi}_1}{1 - \hat{\phi}_2}$$

$$\text{MA}(2): \hat{V}(\hat{\theta}_1, \hat{\theta}_2) \cong \begin{bmatrix} 1 - \hat{\theta}_2^2 & -\hat{\theta}_1(1 + \hat{\theta}_2) \\ -\hat{\theta}_1(1 + \hat{\theta}_2) & 1 - \hat{\theta}_1^2 \end{bmatrix} \frac{1}{n}$$

$$\text{Corr}(\hat{\theta}_1, \hat{\theta}_2) = - \frac{\hat{\theta}_1}{1 - \hat{\theta}_2}$$

$$\text{ARMA}(1,1): \hat{V}(\hat{\phi}, \hat{\theta}) = \begin{bmatrix} (1-\hat{\phi}^2)(1-\hat{\theta}) & (1-\hat{\phi}^2)(1-\hat{\theta}^2) \\ (1-\hat{\phi}^2)(1-\hat{\theta}^2) & (1-\hat{\phi}^2)(1-\hat{\theta}) \end{bmatrix} \frac{1}{n} \frac{1-\hat{\phi}\hat{\theta}}{(\hat{\phi}-\hat{\theta})^2}$$

$$\text{Corr}(\hat{\phi}, \hat{\theta}) = \frac{(1-\hat{\theta}^2)(1-\hat{\phi}^2)}{\sqrt{(1-\hat{\phi}\hat{\theta})^2(1-\hat{\theta}^2)(1-\hat{\phi}^2)}} = \frac{\sqrt{(1-\hat{\phi}^2)(1-\hat{\theta}^2)}}{(1-\hat{\phi}\hat{\theta})}$$

NO FACTORES COMUNES.- Ejemplo:

$$z_t = .5 z_{t-1} + 0.18 z_{t-2} - 0.43 a_{t-1} + a_t$$

$$\text{S.C.} = 1387.5 \quad \text{Corr}(\hat{\phi}_1, \hat{\theta}_1) \approx 1$$

$$z_t = .8 z_{t-1} + 0.18 z_{t-2} - 0.93 a_{t-1} + a_t$$

$$\text{S.C.} = 1387.5 \quad \text{Corr}(\hat{\phi}_1, \hat{\theta}_1) \approx 1$$

Los dos modelos son equivalentes.

Se observará una cosa semejante en el supuesto de no factores comunes

$$\text{ARMA}(p,q) \quad \Phi(B) z_t = \Theta(B) a_t$$

$$(1-zB) \Phi(B) z_t = (1-zB) \Theta(B) a_t$$

$$\Phi'(B) z_t = \Theta'(B) a_t \quad \text{ARMA}(p+1, q+1)$$

Nuevamente los dos modelos son equivalentes pero el segundo es más complicado innecesariamente. Lo que esto resalta es

que debe buscarse factores comunes en los polinomios de rezagos del modelo estimado (i.e. se comparan las raíces de los polinomios) y si algunas son iguales a pares, se reestima el modelo con un grado menor en la parte AR y MA

SOBREPAREMETRIZACION.- Dado el modelo, proponer uno mayor que sea factible de mejorar la explicación de la serie de tiempo sin tener deficiencias.

Los dos modelos seleccionados serán sometidos a las pruebas anteriormente mencionadas. Primero se comenzara con el modelo MA(1):

$$E[a_t] = 0$$

$$t = \frac{\hat{a}}{\hat{\sigma}} = \frac{.072}{.304} = 0.2368421$$

$$t_{n-1, (1-\alpha/2)} = t_{118, (1-.025)} = t_{118, 0.975} = 1.96$$

$$|t| = 0.2368421 < 1.96 = t_{118, 0.975} \therefore \text{se acepta el supuesto.}$$

$$\text{Var}[a_t] = \sigma^2$$

Se puede observar en la gráfica del modelo MA(1) el comportamiento que tiene la varianza y se observa que mas del 95% de las observaciones se encuentran dentro del rango de $\pm 2\sigma$, razón por demas suficiente para tomar por valido el supuesto.

Para poder encontrar los valores de los estadísticos y de la gráfica del modelo se valió de la ayuda del paquete estadístico SYSTAT, en donde para ello se valió de las siguientes instrucciones:

```
Dentro del modulo 3) STATS
>USE YEARS
>OUTPUT @
>STAT RESIDUAL / MEAN, SD, SKEWNESS, KURTOSIS
```

El primer comando sirve para dar la instrucción al paquete estadístico de que use el archivo que contiene las variables del modelo (YEAR, valores originales de la serie, N, tiempo, VARTRA, serie original transformada para estabilizar la serie, TRANSF, serie transformada con una diferencia con varianza estabilizada, RESIDUAL, residuales del modelo). La siguiente instrucción sirve para que los resultados salgan a la impresora en vez de la pantalla y el último comando sirve para calcular la media (MEAN), desviación estandar (SD), asimetría (SKEWNESS) y curtosis (KURTOSIS) de la variable RESIDUAL, con ello se obtienen los resultados que se observan en la parte inferior de la gráfica.

Para poder realizar la gráfica se hace dentro del modulo 2) PLOT del SYSTAT, con las siguientes instrucciones:

```
>USE YEAR5
>OUTPUT @
>PLOT RESIDUAL*N
```

Las instrucciones son semejantes que las anteriores la única diferencia consiste en la última instrucción que sirve para indicar al paquete que realice la gráfica de la variable residual (eje Y) contra N (eje X).

$$a_i \sim N$$

Para verificar este supuesto se utiliza la prueba PEARSON (χ^2) que consiste en lo siguiente:

- Dividase el espectro de la distribución en m intervalos (no necesariamente iguales)
- Calcúlese las probabilidades que la distribución teórica (NORMAL) asigna a cada uno de estos intervalos; sea p_i la probabilidad del i -ésimo intervalo y $E_i = Np_i$ el número esperado de sus observaciones en ese intervalo, siendo N el tamaño muestral
- Sea O_i el número real de observaciones en el i -ésimo intervalo.

El estadígrafo

$$\chi^2 = \sum_{i=1}^m \frac{(\hat{O}_i - E_i)^2}{E_i}$$

tiene aproximadamente una distribución χ^2 con $m-1$ grados de libertad.

Si el valor calculado de χ^2 excede el fractil apropiado de la distribución χ^2 se rechaza la hipótesis nula que la muestra proviene en efecto de la distribución en cuestión, en el caso expuesto se refiere a al NORMAL, al nivel de confianza elegido.

El procedimiento anterior se muestra en la siguiente tabla:

En la tabla se muestra que el estadígrafo es menor que el fractil de la distribución χ^2 al 95% de confianza, por ello se acepta el supuesto de que el modelo se distribuye como normal.

$$\text{Cov}(a_i, a_{i,k}) = 0$$

Para poder probar este supuesto se hará la prueba de BOX-LJUNG, y se valdrá de la ayuda del paquete SYSTAT y de LOTUS. Con el Paquete SYSTAT se calculan las primeras 20 autocorrelaciones del modelo MA(1) y con el paquete LOTUS se llevará a cabo el proceso para calcular el estadígrafo $Q(k)$ y poder compararlo con el fractil de la distribución χ^2 , con lo cual se llega a la conclusión de que no existe relación alguna entre las observaciones, como se observa en la gráfica del residual*N y en la siguiente tabla:

CDRR	R^2		
-0.081	0.006561	1	118 0.000056
-0.081	0.006561	2	118 0.000056
-0.051	0.002601	3	118 0.000022
-0.086	0.007396	4	118 0.000064
-0.083	0.006889	5	118 0.000060
-0.038	0.001444	6	118 0.000012
-0.093	0.008649	7	118 0.000077
0.01	0.0001	8	118 0.000000
0.077	0.005929	9	118 0.000054
-0.043	0.001849	10	118 0.000017
0.085	0.007225	11	118 0.000067
0.13	0.0169	12	118 0.000159
-0.05	0.0025	13	118 0.000023
0.06	0.0036	14	118 0.000034
0.011	0.000121	15	118 0.000001
-0.148	0.021904	16	118 0.000214
-0.102	0.010404	17	118 0.000103
0.035	0.001225	18	118 0.000012
-0.053	0.002809	19	118 0.000028
0.03	0.0009	20	118 0.000009

14160 0.001078 15.27087 = Q(K)

X^2 AL 95 % Y 19 GRADOS= 30.15

SE ACEPTA EL SUPUESTO YA QUE X^2 < Q(K)

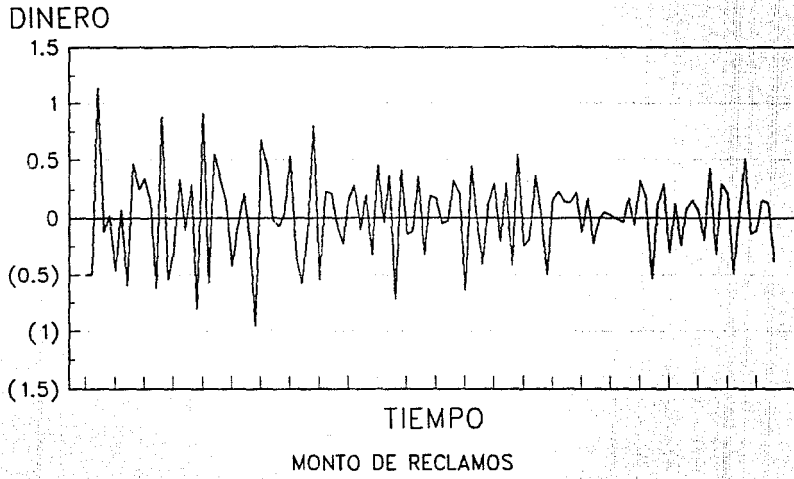
CORR	R^2			
-0.031	0.000961	1	118	0.000008
-0.12	0.0144	2	118	0.000124
-0.05	0.0025	3	118	0.000021
-0.014	0.000196	4	118	0.000001
-0.208	0.043264	5	118	0.000382
-0.161	0.025921	6	118	0.000231
0.084	0.007056	7	118	0.000063
0.032	0.001024	8	118	0.000009
0.014	0.000196	9	118	0.000001
-0.025	0.000625	10	118	0.000005
-0.151	0.022801	11	118	0.000213
0.095	0.009025	12	118	0.000085
-0.081	0.006561	13	118	0.000062
0.079	0.006241	14	118	0.000060
0.023	0.000529	15	118	0.000005
-0.178	0.031684	16	118	0.000310
-0.121	0.014641	17	118	0.000144
0.062	0.003844	18	118	0.000038
-0.051	0.002601	19	118	0.000026
-0.01	0.0001	20	118	0.000001

14160 0.001797 25.45633 = 0(K)

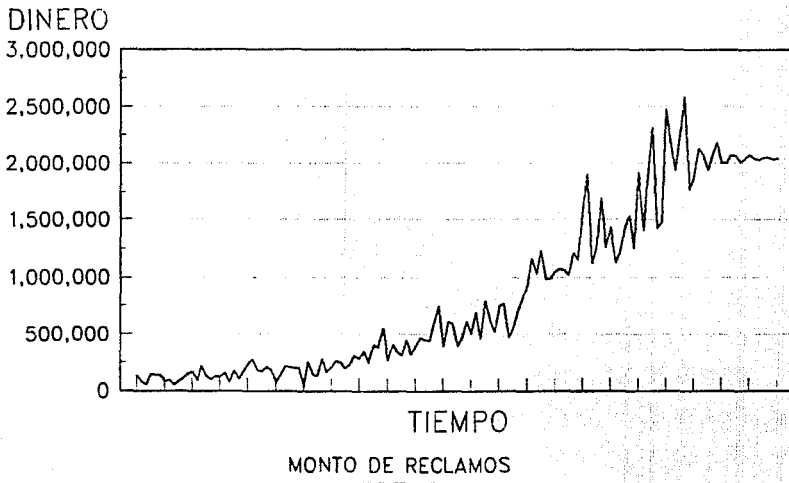
X^2 AL 95 % Y 19 GRADOS= 30.15

SE ACEPTA EL SUPUESTO YA QUE X^2 < 0(K)

U.N.A.M.
E.N.E.P. ACATLAN
RESIDUALES DEL MA(1)



U.N.A.M.
E.N.E.P. ACATLAN
RECLAMOS CON PRONOST. DEL MOD. AR(6)



NO ABERRANTES O DISCREPANTES

También en la gráfica se observa que no existen aberrantes o discrepantes ya que todos los residuales se encuentran dentro del intervalo $-\bar{3}\bar{S}$.

$$E [a_t, Z_{t-k}] = 0$$

Al calcular el correlograma cruzado de los residuales contra las observaciones transformadas (TRANSF) por medio del modulo SERIES del paquete SYSTAT con las siguientes instrucciones:

```
>OUTPUT @
>USE YEARS
>CCF RESIDUAL TRANSF
```

Se observa que el comportamiento de este sigue el del modelo teórico para aceptar el supuesto de que los residuales no tienen relación alguna con las observaciones, por ello se acepta el supuesto.

ADMISIBILIDAD

Para poder admitir el modelo MA(1), el parámetro estimado debiera ser menor, en valor absoluto, que uno, lo cual se cumple:

$$|\theta| < 1$$

$$|0.678| < 1$$

0.678 < 1 \therefore se admite el modelo.

ESTACIONARIEDAD

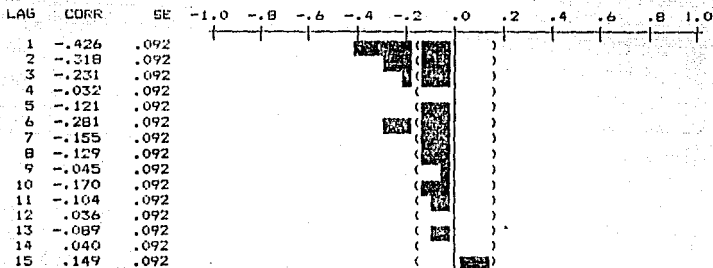
El modelo para que sea estacionario, la raíz de la ecuación $(1 - \theta_1 B)$ tiene que ser mayor que uno.

$$z_1 = (1 - \theta_1 B) a_t$$

$$z_1 = (1 - 0.678 B) a_t$$

PLOT OF TRANSF
 NUMBER OF CASES = 119
 MEAN OF SERIES = 0.022
 STANDARD DEVIATION OF SERIES = 0.369

PLOT OF PARTIAL AUTOCORRELATIONS

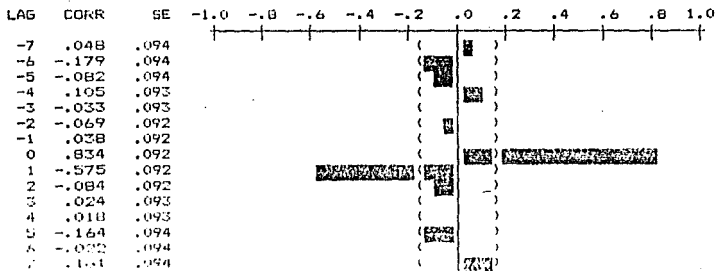


CORRELOGRAMA PARCIAL DEL MODELO CON UNA DIFERENCIA

NOTE

NOTE

PLOT OF CROSS CORRELATIONS



$$(1 - 0.676 B) = 0$$

$$B = 1.47 > 1$$

∴ es estacionario.

PARSIMONIA

Para probar este supuesto se hace uso de los intervalos de confianza, teniendo como valores principales el valor del parámetro calculado y su desviación estandar con lo cual se observa lo siguiente:

$$(\theta - 1.96 \sigma(\theta), \theta + 1.96 \sigma(\theta))$$

$$(0.676 - 1.96(.035), 0.676 + 1.96 (.035))$$

$$(0.6074, 0.7446)$$

Como no contiene al cero el intervalo, entonces se acepta el supuesto de que el modelo tiene parsimonia.

El modelo tambien cumple con los demás supuestos de estabilidad, no tener factores comunes y tampoco tiene parámetros de más.

Para el caso del modelo AR(6) tambien se lleva acabo exactamente lo mismo que se hizo anteriormente notandose que este modelo tambien cumple con todos los supuestos.

Ahora, para establecer cual de los dos modelos es el mejor se escoge el que tenga menos parámetros, en este caso es el MA(1), esto es debido al principio de PARSIMONIA, el cual establece que se escoge el modelo que tenga menos parámetros y que siga el mismo comportamiento de la serie original. En este ejemplo en particular se observa tambien que se prueba el supuesto teórico de que a todo modelo MA(1) le corresponde un AR(∞).

2.3. Estimación por medio de la serie de tiempo

En la práctica, las predicciones se hacen casi invariablemente con ϕ y θ remplazados por sus estimados. Esto crea un origen adicional de variabilidad que se debe incorporar en la expresión para la predicción de error cuadrático medio.

Considérese un proceso AR(1). Cuando ϕ se conoce, el error medio cuadrático mínimo para 1 periodos se da por

$$\hat{z}_{T+1/T} = \phi^1 z_T$$

Cuando ϕ se desconoce reemplaza por su estimador de máxima verosimilitud, $\hat{\phi}$, o por un estimador el cual es asintóticamente equivalente. El predictor actual es por lo tanto

$$\hat{z}_{T+1/T}^* = \hat{\phi}^1 z_T$$

En casos mas generales $\hat{z}_{T+1/T}^*$ se puede calcular por una ecuación en diferencias teniendo la siguiente forma:

$$\hat{z}_{T+1/T}^* = \phi_1 \hat{z}_{T+1-1/T}^* + \dots + \phi_p \hat{z}_{T+1-p/T}^* + a_{T+1/T} + \dots + \theta_q a_{T+1-q/T} \quad l = 1, 2, 3, \dots$$

donde $\hat{z}_{T+1-j/T}^* = z_{T+j}$ para $j \leq 0$, y

$$a_{T+1-j/T} = \begin{cases} 0 & \text{para } j > 0 \\ a_{T+j} & \text{para } j \leq 0 \end{cases}$$

El error de predicción para

$$\hat{z}_{T+1/T}^* = \phi^1 z_T$$

se puede descomponer en dos partes escribiendo

$$z_{T+1} - \hat{z}_{T+1/T}^* = (z_{T+1} - \hat{z}_{T+1/T}) + (\hat{z}_{T+1/T} - \hat{z}_{T+1/T}^*)$$

El primer término del lado derecho de la ecuación anterior es el error de predicción cuando $\hat{\phi}$ se conoce, mientras que el segundo término representa el error que surge de la estimación de ϕ . Esta descomposición es apropiada para todos los modelos ARMA. En el presente caso se especializa a:

$$z_{T+1} - \hat{z}_{T+1|T}^* = (z_{T+1|T} - \hat{z}_{T+1|T}) + (\hat{\phi}^1 - \phi^1) z_T$$

en vista de la ecuación $\hat{z}_{T+1|T}^* = \hat{\phi}^1 z_T$.

Ahora considérese el predictor de un paso adelante. El error medio cuadrático se puede escribir como

$$\text{MSE} (\hat{z}_{T+1|T}^*) = \text{MSE} (\hat{z}_{T+1|T}^*) + z_T^2 E [(\hat{\phi} - \phi)^2]$$

En la formulación de la contribución del error de estimación en la ecuación anterior, z_T se toma como un valor fijo, mientras que $\hat{\phi}$ es una variable aleatoria. Esto puede parecer contradictorio, en la medida que z_t se usa actualmente para construir $\hat{\phi}$. Sin embargo la ecuación del error medio cuadrático provee una definición sensible de error medio cuadrático en éste contexto, desde que cualquier predicción siempre se hace condicional sobre unas observaciones muestrales. Reemplazando $E [(\hat{\phi} - \phi)^2]$ por su varianza asintótica da una aproximación al error cuadrático medio, i.e.

$$\text{MSE} (\hat{z}_{T+1|T}^*) \cong \sigma^2 + z_T^2 (1 - \phi^2) / T$$

Aplicando la fórmula de regresión usual para estimar el error cuadrático medio de $\hat{z}_{T+1|T}^*$ da:

$$\text{MSE} (\hat{z}_{T+1|T}^*) = s^2 (1 + z_T^2 / \sum_{t=2}^T z_{t-1}^2)$$

La fórmula anterior está cercanamente relacionada a

$\text{MSE} (\hat{z}_{T+1|T}^*) \cong \sigma^2 + z_T^2 (1 - \phi^2) / T$ desde que es válido

$$\sum_{i=2}^T z_{i-1}^2 \cong T \sigma^2 / (1 - \phi^2)$$

en muestras grandes.

Cuando $l > 1$, el último término en

$$\text{MSE}(\hat{z}_{T+1/T}) = \text{MSE}(\hat{z}_{T+1/T}^*) + z_T^2 E[(\hat{\phi} - \phi)^2]$$

es $z_T^2 E[(\hat{\phi}^1 - \phi^1)^2]$. Escribiendo

$$\phi^1 - \hat{\phi}^1 = \phi^1 - \phi^1 \left[1 - \frac{(\phi - \hat{\phi})}{\phi} \right]^l$$

y expandiendo el término del corchete nos da:

$$\phi^1 - \hat{\phi}^1 \cong l \phi^{l-1} (\phi - \hat{\phi})$$

cuando los términos de mayor orden se ignoran.

Para el modelo AR(1) $y_{t+j} = \phi^{1+j} a_{t-j}$, y así (empleando el principio de parsimonia)

$$z_{T+1/T} = \sum_{j=0}^{\infty} \phi^{1+j} a_{T-j} = \phi^1 \sum_{j=0}^{\infty} \phi^j a_{T-j} = \phi^1 z_T$$

El error medio cuadrático (MSE) de $z_{T+1/T}$ en el modelo AR(1) se da por

$$\text{MSE}(z_{T+1/T}) = (1 + \phi^2 + \dots + \phi^{2(l-1)}) \sigma^2$$

$$= \frac{1 - \phi^{2l}}{1 - \phi^2} \sigma^2$$

Cuando $l \rightarrow \infty$, la expresión anterior tiende a $\sigma^2/(1-\phi^2)$ que es simplemente la varianza de z_1 .

Tomando en cuenta el resultado anterior y que

$$\phi^l - \hat{\phi}^l \approx l \phi^{l-1} (\phi - \hat{\phi}) \text{ produce}$$

$$\text{MSE}(\hat{z}_{T+l|T}^*) \approx \sigma^2 \left[\frac{1 - \phi^{2l}}{1 - \phi^2} + \frac{l^2 \phi^{2(l-1)}}{T} \right], \quad l = 1, 2, \dots$$

Para el caso especial de $l = 1$

$$\text{AMSE}(\hat{z}_{T+1|T}^*) = \sigma^2(1 + T^{-1})$$

En las dos ecuaciones anteriores, la contribución que surge del error de la estimación de ϕ es un término de $O(T^{-1})$. Esto se dominará por la expresión del error medio cuadrático del predictor óptimo cuando ϕ se conoce. Aunque al ignorar el efecto de estimar ϕ conducirá a un subestimado de la variabilidad del error de predicción, los sesgos no son parecidos a ser severos a menos que T sea pequeño. Estas averiguaciones conducen a modelos más generales.

Al tiempo $T+1$, la ecuación para un proceso MA(1) es de la forma

$$z_{T+1} = a_{T+1} + \theta a_T$$

Desde que a_{T+1} se desconoce, se deja igual a cero en la ecuación de predicción correspondiente la cual es

$$\hat{z}_{T+1|T} = \theta a_T$$

Para $l > 1$, $\hat{z}_{T+l|T} = 0$, y así el conocimiento del proceso generador de datos no es de mayor ayuda más que para predecir más que un solo periodo adelante.

Para el caso del modelo MA(1), el error medio cuadrático MSE, de $\hat{z}_{T+l|T}$ se puede obtener inmediatamente. De la ecuación

$$z_{T+1} = a_{T+1} + \theta a_T$$

y de la expresión

$$z_{T+l/T} = \theta a_T$$

se obtiene que

$$\text{MSE}(z_{T+l/T}) = E[(z_{T+l} - z_{T+l/T})^2] = E[a_{T+l}^2] = \sigma^2$$

Para $l > 1$, la solución es trivial, desde que $z_{T+l/T} = 0$ y

$$\text{MSE}(z_{T+l/T}^*) = \text{Var}(z_{T+l}) = (1 + \theta^2) \sigma^2.$$

Para cualquier modelo MA(q) no es difícil ver que el cálculo de la predicción es hacia adelante.

Como se vio anteriormente, el caso práctico no es estacionario y se tuvo que estacionarizar realizando dos transformaciones, la primera transformación consistió en estabilizar a la varianza, lo cual se hizo tomando logaritmos naturales a la serie original, llamada YEAR, y la serie con varianza constante se llama VARTRA. La segunda transformación consistió en estabilizar la media, con lo cual se sacó una diferencia a la serie llamada VARTRA. Al realizar ésta diferencia el proceso se convierte en un modelo ARIMA (proceso Auto-Regresivo Medias Móviles Integrado).

Una característica básica de las predicciones es que éstas tienden hacia la media de la serie en la medida que el tiempo l aumenta. Si l es grande, la estructura del modelo es irrelevante. Esto no sucede con los procesos integrados, donde la función de predicción contiene un componente determinístico el cual depende tanto del grado de diferenciación y el modelo ARMA escogido para simular las observaciones diferenciadas.

La mecánica de hacer predicciones de los modelos ARIMA son exactamente las mismas que aquellas que se utilizaron en los modelos ARMA. El término $\phi(L)\Delta^d$ en la siguiente ecuación:

$$\phi(L)\Delta^d z_t = \theta(L) a_t$$

se desarrolla para dar como resultado un polinomio de orden $p+d$, i.e.

$$\phi(L) \Delta^d = \varphi(L) = 1 - \rho L - \dots - \rho_{p+d} L^{p+d}$$

Las predicciones se hacen de la ecuación en diferencias

$$\hat{z}_{T+1/T} = \rho_1 \hat{z}_{T+1-1/T} + \dots + \rho_{p+d} \hat{z}_{T+1-p-d/T} +$$

$$\hat{a}_{T+1/T} + \dots + \theta_q \hat{a}_{T+1-q/T}$$

En el modelo ARIMA(0,1,1)

$$\Delta z_t = a_t + \theta a_{t-1}$$

los pronósticos se constuyen de la ecuación en diferencias

$$\hat{z}_{T+1/T} = z_{T+1-1/T} + \hat{a}_{T+1/T} + \theta \hat{a}_{T+1-1/T} \quad i = 1, 2, \dots$$

Como las esperanzas de valores futuros de a_t son cero, se sigue que

$$\hat{z}_{T+1/T} = z_T + \theta a_T$$

y

$$\hat{z}_{T+1/T} = z_{T+1-1/T} \quad i = 2, 3, \dots$$

Entonces los pronósticos hechos al tiempo T siguen una línea horizontal derecha.

Una vez establecido la teoría de la predicción en base a una serie de tiempo, se darán las instrucciones necesarias del paquete SYSTAT realizar la predicción con el modelo MA(1) y posteriormente se valdrá de la ayuda de LOTUS para realizar el cálculo y graficado de los pronósticos del modelo ARIMA (0,1,1)

Las instrucciones que se utilizaron en SYSTAT son, dentro del modulo SERIES

> USE YEAR5

> ARIMA TRANSF / Q=1, FORECAST=5

De esta forma solamente se obtiene un solo pronóstico del modelo MA(1), que se explica con las razones anotadas anteriormente. Ahora con Lotus se procederá a volver a sus

valores reales de la serie estacionarizada, elevando con la base e el valor de la serie estacionarizada y realizar el graficado de la serie original con sus respectivos pronósticos. También se graficará el modelo AR(6) donde se observa que este modelo, aunque mas complejo que el MA(1), tiene la misma tendencia a pronosticar solamente un periodo adelante y los demás pronósticos casi siguen una línea recta como en el modelo MA(1). En este caso se observa que el principio de parsimonia se cumple.

Estacionalidad.

Quando las observaciones son disponibles en bases mensuales o trimestrales algunas concesiones se deben hacer por los efectos estacionales. Una aproximación podría trabajar con datos ajustados estacionalmente pero esto por lo general no es deseable. Esta sección se refiere a las formas en las cuales la estacionalidad se debe incorporar a los modelos de series de tiempo.

Hay dos aspectos en la estacionalidad. La primera se refiere a los patrones en las observaciones los cuales se repiten mas o menos regularmente año con año. Estas características son permanentes, y se pueden considerar que se originan en factores tal como la temperatura. Modelar tales efectos envuelve consideraciones similares a aquellas que surgieron en las tendencias del modelo. Un componente estacional determinístico es adecuado si se siente que el patrón estacional es constante año con año, mientras que un componente estocástico se puede introducir si se siente que el patrón lentamente está cambiando.

ITERATION	SUM OF SQUARES	PARAMETER VALUES
0	.1503378D+02	.100
1	.1149428D+02	.662
2	.1149009D+02	.678
3	.1149000D+02	.676
4	.1149000D+02	.676
5	.1149000D+02	.676
6	.1149000D+02	.676
7	.1149000D+02	.676

ESTIMATED COEFFICIENTS

INDEX	TYPE	ESTIMATE	STANDARD ERROR
1	MA	0.676	0.035

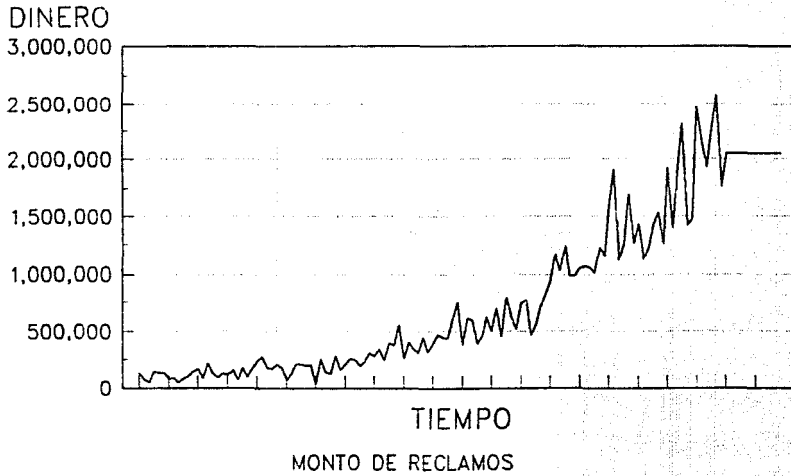
FORECAST VALUES

PERIOD	LOWER95	FORECAST	UPPER95
120	-0.463	0.148	0.760
121	-0.738	0.000	0.738
122	-0.738	0.000	0.738
123	-0.738	0.000	0.738
124	-0.738	0.000	0.738

ESTIMACION Y PRONOSTICO DEL MODELO MA 1

NOTE

U.N.A.M.
E.N.E.P. ACATLAN
RECLAMOS CON PRONOST. DEL MOD. MA(1)



ITERATION	SUM OF SQUARES	PARAMETER VALUES						
0	.1844196D+02	.100	.100	.100	.100	.100	.100	.100
1	.1390912D+02	-.667	-.051	.117	.272	-.127	-.148	
2	.1284131D+02	-.738	-.294	.067	.183	-.085	-.288	
3	.1170170D+02	-.822	-.510	-.279	.006	-.179	-.252	
4	.1151540D+02	-.839	-.520	-.286	-.112	-.227	-.243	
5	.1133478D+02	-.774	-.490	-.319	-.173	-.334	-.335	
6	.1097360D+02	-.666	-.430	-.322	-.171	-.222	-.271	
7	.1090656D+02	-.633	-.451	-.301	-.155	-.232	-.236	
8	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	
9	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	
10	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	
11	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	
12	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	
13	.1090655D+02	-.633	-.451	-.301	-.155	-.232	-.235	

ESTIMATED COEFFICIENTS

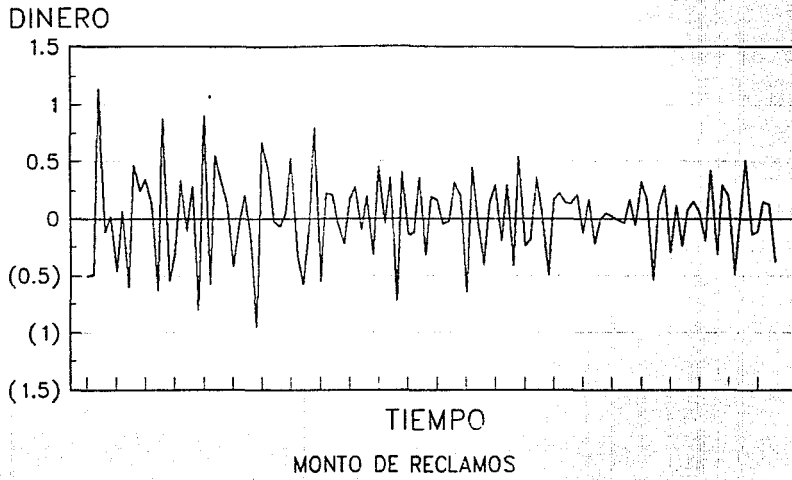
INDEX	TYPE	ESTIMATE	STANDARD ERROR
1	AR	-0.633	0.159
2	AR	-0.451	0.205
3	AR	-0.301	0.261
4	AR	-0.155	0.113
5	AR	-0.232	0.102
6	AR	-0.235	0.085

FORECAST VALUES

PERIOD	LOWER95	FORECAST	UPPER95
120	-0.548	0.061	0.670
121	-0.595	0.126	0.844
122	-0.746	-0.024	0.697
123	-0.790	-0.068	0.653
124	-0.658	0.064	0.786
125	-0.675	0.053	0.782
126	-0.810	-0.082	0.647
127	-0.746	-0.005	0.736
128	-0.707	0.036	0.778
129	-0.745	-0.003	0.740
130	-0.770	-0.028	0.715
131	-0.728	0.015	0.758
132	-0.725	0.019	0.762
133	-0.761	-0.017	0.727
134	-0.750	-0.006	0.739
135	-0.734	0.010	0.755
136	-0.743	0.001	0.746
137	-0.754	-0.009	0.736
138	-0.742	0.002	0.747
139	-0.739	0.006	0.751

ESTIMACION Y PRONOSTICO DEL MODELO AR 6

U.N.A.M.
E.N.E.P. ACATLAN
RESIDUALES DEL AR(6)



Ambos modelos proyectan un patrón estacional regular hacia el futuro, pero cuando un modelo estocástico se usa este patrón depende mucho más en las observaciones pasadas de la serie. Hay un paralelo directo con el concepto de una tendencia local en una serie no estacional. (Diferenciación de la serie)

El otro aspecto de la estacionalidad se refiere a los efectos que surgen en la ausencia de, o en adición a, el tipo de estacionalidad descrita en el párrafo anterior. Por ejemplo, una huelga en el muelle en marzo del año pasado, puede influir en los objetivos de producción de este año, si las firmas creen que hay una alta probabilidad de que tal evento suceda de nuevo. Sin embargo, a menos que las huelgas en los muelles se conviertan en una ocurrencia regular, el efecto de la huelga original será transitoria. Entonces, mientras se parece que las observaciones en las mismas estaciones de diferentes años serán correlacionadas, estas correlaciones se pueden esperar que sean pequeñas si los años están muy separados. Estas consideraciones sugieren un modelo estacionario, en el cual cualquier patrón estacional tiende a desaparecer en la medida en que el tiempo del pronóstico aumenta.

Procesos estacionales ARMA.

Considerese una serie estacionaria de observaciones trimestrales. Una forma muy sencilla de capturar un efecto estacional es por medio de un proceso autorregresivo estacional de cuarto orden de la forma

$$z_t = \phi_4 z_{t-4} + a_t, \quad |\phi_4| < 1$$

Este modelo es un caso especial de un proceso AR(4), pero con el restringimiento de que $\phi_1 = \phi_2 = \phi_3 = 0$. Las propiedades del modelo anterior son muy similares a aquellas del proceso estacionario AR(1) y usando las técnicas desarrolladas, no es difícil demostrar que la función de autocorrelación se da por

$$\rho(\tau) = \begin{cases} \phi_4^{\tau/4} & \tau = 0, 4, 8, \dots \\ 0 & \text{otro caso} \end{cases}$$

Lo más cercano es $|\phi_s|$ a la unidad. Mas fuerte es el patrón estacional. Sin embargo, en la medida que ϕ_s permanezca dentro del círculo unitario, el efecto estacional es no determinístico, con el patrón estacional gradualmente disminuyendo en la medida en que las predicciones se hacen mas y más hacia adelante.

El modelo anterior se puede extender a tener tanto términos AR y MA en otros rezagos estacionales. Si s denota el número de estaciones en el año, una formulación general es como sigue:

$$\phi^\dagger(L^s) z_t = \theta^\dagger(L^s) \zeta_t$$

donde

$$\phi^\dagger(L^s) = 1 - \phi_1^\dagger L^s - \dots - \phi_p^\dagger L^{ps}$$

$$\theta^\dagger(L^s) = 1 - \theta_1^\dagger L^s - \dots - \theta_q^\dagger L^{qs}$$

y ζ_t es un término de disturbio de ruido blanco. El valor de s será cuatro o doce en aplicaciones económicas, correspondientes a observaciones trimestrales o mensuales respectivamente. La expresión

$$\phi^\dagger(L^s) z_t = \theta^\dagger(L^s) \zeta_t$$

define un proceso estacional puro de orden $(P, Q)_s$. En el caso de la ecuación anterior no tendrá vacíos en retrasos no estacionales. Sin embargo, a menos de que los movimientos estacionales se sienta que sean solamente características predecibles de la serie, tal modelo no será apropiado. Con datos mensuales, por ejemplo, parece razonable suponer que una observación en marzo se relacionara a una observación en febrero, así como a la observación de marzo del año pasado.

Hay dos formas de construir modelos que hacen conexiones para movimientos estacionales y no estacionales. El primero simplemente llena los vacíos en el proceso estacional. Entonces un retraso de primer orden puede incorporarse a

$$z_t = \phi_s z_{t-s} + a_t, \quad |\phi_s| < 1$$

para producir

$$z_t = \phi_1 z_{t-1} + \phi_4 z_{t-4} + a_t$$

Una segunda forma es reemplazar el ruido blanco en

$$\phi^{\dagger}(L^4) z_t = \theta^{\dagger}(L^4) \zeta_t$$

por un proceso no estacional ARMA(P,Q), u_t : i.e.

$$\phi(L)u_t = \theta(L)a_t$$

donde $\phi(L)$ y $\theta(L)$ son los polinomios asociados definidos anteriormente. Combinando

$$\phi^{\dagger}(L^4) z_t = \theta^{\dagger}(L^4) \zeta_t$$

y

$$\phi(L)u_t = \theta(L)a_t$$

produce

$$\phi^{\dagger}(L^4) \phi(L) z_t = \theta^{\dagger}(L^4) \theta(L) \zeta_t$$

Este es un proceso ARMA estacional multiplicativo ARMA de orden $(p,q) \times (P,Q)$. Como un ejemplo simple, considerese que un proceso AR(1) se incorpora a

$$z_t = \phi_4 z_{t-4} + a_t, \quad |\phi_4| < 1$$

El modelo se convierte en

$$(1 - \phi_4 L) (1 - \phi_4 L^4) = a_t$$

Esto se puede reescribir como

$$z_t = \phi_1 z_{t-1} + \phi_4 z_{t-4} + \phi_5 z_{t-5} + a_t$$

donde $\phi_5 = -\phi_1 \phi_4$.

Las formas multiplicativas se usan ampliamente en el modelado de series de tiempo. Sin embargo, no hay algo desarrollado que indique porque

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_3 z_{t-3} + \phi_4 z_{t-4} + a_t$$

se debe preferir a

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + a_t$$

Esto se debe suponer que los modelos multiplicativos surgen más naturalmente cuando la diferenciación ha sido hecha y se pueda identificar la estacionalidad por medio del correlograma.

Una clase general de modelos.

El material presentado anteriormente sugiere una aproximación general a la construcción de modelos basado en la diferenciación. Esta es la filosofía adoptada por Box y Jenkins. Ellos proponen que los operadores de diferenciación convencional y estacional se apliquen hasta que la serie sea estacionaria y la serie estacionaria se modele por un proceso ARMA estacional multiplicativo de la forma

$$\phi^*(L^s) \phi(L) \Delta^d \Delta_n^D z_t = \theta^*(L^s) \theta(L) \zeta_t$$

donde D y d son enteros que denotan el número de veces que los operadores de primeras diferencias y diferencias estacionales se aplican respectivamente. Esto se conoce como un proceso ARIMA estacional multiplicativo de orden $(p, d, q) \times (P, D, Q)_n$.

Una formulación más general se obtiene relajando la suposición de que la parte estacionaria del modelo es un proceso multiplicativo. Un modelo de esta forma se puede escribir

$$\phi(L) \Delta^d \Delta_n^D z_t = \theta(L) a_t$$

donde los polinomios $\phi(L)$ y $\theta(L)$ tendrán parámetros de retrasos estacionales y no estacionales, pero pueden tener vacíos.

Capítulo 3. La dependencia de la función S en reaseguro.

Retenciones

El problema de la retención

General

Imagínese una cartera de seguro que se tomó sobre varios portafolios, cada uno de distintas colectividades. Se debe llamar a la suma de varios portafolios retenidos por la cartera del asegurador como su masa de riesgo.

Más adelante se debe suponer el portafolio individual y que la experiencia de los reclamos de los riesgos individuales en cada portafolio son estocásticamente independientes entre sí. Esto significa que las experiencias de reclamos $S_i^{(t)}$ para todo riesgo i en la masa de riesgo entera son independientes entre sí, aún cuando los parámetros que caracterizan el riesgo son conocidos o no.

El problema de la retención consiste en determinar para una masa de riesgo la porción retenida que es de provecho para la cartera del asegurador, y la porción restante que debe ser cedida a otra cartera de riesgo, el reasegurador:

$$Z = Z_{TOT} - Z_{re}$$

Z_{re} .- Parte del reclamo total tomada por el reasegurador

Z_{TOT} .- Reclamo total

El reclamo está decisivamente influido por el costo que el reasegurador demanda por tomar a su cargo la parte cedente.

Considérese una masa de riesgo dada que consiste de N riesgos, donde el i -ésimo riesgo se caracteriza por ser el reclamo acumulado $S_i^{(t)}$. Si se puede determinar una función g_i , que se le pueda asignar una porción retenida $R_i^{(t)}$ con la función de muestreo $S_i^{(t)}$ del proceso del reclamo acumulado al tiempo t , entonces se habla de reaseguro individual (reaseguro por riesgo).

Simbólicamente:

$$R_i^{(1)} = g_i [S_i^{(1)}] \quad (i = 1, 2, 3, 4, \dots, N)$$

En contraste con lo anterior, la retención bajo un sistema de reaseguro global se determina solamente con respecto a la experiencia de reclamos acumulados para la masa de riesgo completa. En símbolos:

$$R_i = g \left[\sum_{i=1}^N S_i^{(1)} \right]$$

3.2. Distintos tipos de reaseguro

La retención bajo reaseguro proporcional y no proporcional

En la práctica hay dos formas usuales de reaseguro individual

1) El caso proporcional

$$g_i [S_i^{(1)}] = a_i S_i^{(1)} \quad (0 \leq a_i \leq 1)$$

El número a_i se llama el porcentaje de retención y la retención es $a_i S_i^{(1)}$

En este caso el precio demandado por el reasegurador de aceptar la cesión del riesgo proporcional es

$$P_i^{(1)} = (1 - a_i) P_i^{(1)}$$

donde $P_i^{(1)}$ denota el costo de tomar el riesgo completo. Esta forma de reaseguro se conoce como reaseguro por excedente. (Surplus Reinsurance).

ii) El caso no proporcional

$$g_i[S_i^{(ii)}] = \sum_{\text{saltes en } (0,t]} \min[\text{salto del monto}, M_i]$$

La conexión entre $S_i^{(ii)}$ y $R_i^{(ii)}$ se puede ver más claramente en el diagrama. Cada función g_i es única, caracterizada una vez que el límite M_i se da.

- M_i se llama el primer riesgo o retención máxima

- El costo de este tipo de reaseguro se da en la forma de la función $P_i^{(ii)}(M_i) = P_i^{(ii)}(g_i)$, donde $P_i^{(ii)}(0) = P_i^{(ii)}$ denota el costo de tomar completamente a $S_i^{(ii)}$. Esta forma de reaseguro se conoce como "Exceso de pérdida".

En ambos casos, proporcional y no proporcional, el problema de la retención consiste en determinar para cada riesgo en una masa de riesgo un número correspondiente (a_i o M_i). En esta conexión, estos números se descomponen como sigue:

$$a_i = C \lambda_i \quad M_i = k \delta_i \quad (i = 1, 2, \dots, N)$$

El problema de determinar los factores λ_i o δ_i se conoce como el problema de la retención relativa: Si el objetivo es determinar las constantes C o k se trata con el problema de la retención absoluta. Esta distinción se puede expresar de otra forma diciendo que el problema de la retención relativa se refiere en encontrar el grado de las retenciones, mientras que el problema absoluto busca determinar el nivel de las retenciones involucradas.

Reaseguro Proporcional

El objetivo es determinar los números característicos a_i para el reaseguro proporcional (retenciones proporcionales) como

constantes multiplicativas. Sin duda el instrumento más conveniente para determinar a_i es el principio de mínima varianza. Se procede de una masa de riesgo de N riesgos para el cual se tiene para el i -ésimo riesgo ($i = 1, 2, \dots, N$)

$S^{(i)}$ = variable aleatoria que denota el reclamo acumulado en el periodo de cálculo $(0,1)$.

$P^{(i)}$ = el costo del reaseguro por asumir el 100% de $S^{(i)}$

Aplicando las retenciones proporcionales a_i , la cantidad

$$Z = \sum_{i=1}^N a_i P^{(i)} - \sum_{i=1}^N a_i S^{(i)}$$

representa una variable estocástica la cual mide el "flujo de beneficios salvados" sobre la porción retenida de la masa de riesgo. El principio de la mínima varianza requiere que para un valor esperado dado de este flujo de beneficios salvado, la varianza sea tan pequeña como sea posible. En símbolos

$$\text{Var} [z] = \text{mínima}$$

bajo la condición adicional que

$$E [z] = \text{constante}$$

Por el método de multiplicadores de Lagrange se encuentra este mínimo en aquellos puntos donde las derivadas parciales con respecto a a_j ($j = 1, 2, \dots, N$) de la función

$$\phi = \text{Var} [z] + \lambda E [z]$$

donde

$$\text{Var} [z] = \sum_{i=1}^N a_i^2 \text{Var} [S^{(i)}]$$

y

$$E [z] = \sum_{i=1}^N a_i (P^{(i)} - E [S^{(i)}]) = \sum_{i=1}^N a_i L^{(i)}$$

Entonces la función ϕ se puede escribir explícitamente como

$$\phi(a_1, \dots, a_N) = \sum_{i=1}^N a_i^2 \text{Var} [S^{(i)}] + \lambda \sum_{i=1}^N a_i L^{(i)}$$

de dónde

$$\frac{\partial \phi}{\partial a_j} = 2 a_j \text{Var} [S^{(j)}] + \lambda L^{(j)} = 0 \quad \text{para toda } j$$

se sigue entonces que

$$a_j = C \frac{L^{(j)}}{\text{Var} S^{(j)}}$$

y por lo tanto

$$\lambda_j = \frac{L^{(j)}}{\text{Var} S^{(j)}} \quad \text{para } j = 1, \dots, N$$

Este resultado se puede interpretar como sigue:

- Dependiendo de la elección de la constante C, la fórmula puede producir un valor de a_j el cual exceda de 1. En este caso a_j se obliga ser 1.
- De Finetti demostró que, sujeto a ésta regla, el a_j calculado de la fórmula son soluciones al problema como se presenta anteriormente.
- La constante C depende de que tan grande el valor esperado de lo "salvado" debe ser.

Es importante notar que hasta aquí se ha dado pequeña importancia a la pregunta si $S^{(j)}$ es el reclamo acumulado de un riesgo con parámetro de riesgo conocido, si es así, entonces

$$E [S^{(i)}] = \mu_i(\phi)$$

y

$$\text{Var} [S^{(i)}] = \sigma_i^2(\phi)$$

entonces se tiene para los factores λ_j

$$\lambda_j = \frac{P^{(j)} - \mu_j(\phi)}{\sigma_j^2(\phi)}$$

Se deberá llamar a la graduación que resulta de los factores proporcionales desarrollados aquí una solución de Finetti del problema de la retención relativa.

Reaseguro no proporcional.

a) Para el reaseguro no proporcional las suposiciones pueden ser mejor formuladas como sigue: Sea

- $S^{(i)}$ la variable aleatoria que describe el reclamo acumulado del riesgo i .

- M_i el límite del reclamo (primer riesgo) que se busca para el i -ésimo riesgo, bajo el cual los reclamos se pagan por el asegurador directo y

- $P^{(i)}(M_i)$ la prima del reasegurador por aceptar el reaseguro proporcional del riesgo i caracterizado por M_i .

Supóngase que la masa de riesgo consiste de nuevo de N riesgos. En adición se debe suponer que $S^{(i)}$ es una suma aleatoria de variables independientes no negativas $Y_j^{(i)}$ con idéntica función de distribución $F^{(i)}(x)$. Simbólicamente

$$S^{(i)} = \sum_{j=0}^{A_i} Y_j^{(i)} \quad (A_i = \text{variable estocástica contable del número de reclamos del riesgo } i)$$

El arreglo del reaseguro caracterizado por M_i determina entonces la variable de retención

$$g_i = g_i[S_i] = \sum_{j=0}^{A_i} \min[M_i, Y_j^{(i)}]$$

Se tiene entonces la fórmula general para el valor esperado

$$E [g_j] = E [A_j] E [Y_j^{(1)} (M_j)]$$

y para la varianza

$$\text{Var} [g_j] = E [A_j] \text{Var} [Y_j^{(1)} (M_j)] + \text{Var} [A_j] E^2 [Y_j^{(1)} (M_j)]$$

Solo falta determinar $E [Y_j^{(1)} (M_j)]$ y $\text{Var} [Y_j^{(1)} (M_j)]$.

Escribiendo simplemente F por $F^{(1)}$, se tiene

$$E [Y_j^{(1)} (M_j)] = \int_0^{M_j} x dF(x) + M_j \int_{M_j}^{\infty} dF(x) = \int_0^{M_j} [1 - F(x)] dx$$

Similarmente

$$E [Y_j^{(1)} (M_j)^2] = \int_0^{M_j} x^2 dF(x) + M_j^2 \int_{M_j}^{\infty} dF(x) = M_j^2 - 2 \int_0^{M_j} x F(x) dx$$

Entonces

$$E [g_j] = E [A_j] \int_0^{M_j} [1 - F(x)] dx$$

$$\begin{aligned} \text{Var} [g_j] = E [A_j] & \left[M_j^2 - 2 \int_0^{M_j} x F(x) dx - \left(\int_0^{M_j} [1 - F(x)] dx \right)^2 \right] \\ & + \text{Var} [A_j] \left[\int_0^{M_j} [1 - F(x)] dx \right]^2 \end{aligned}$$

Por último se debe suponer que la prima de reaseguro se calcula de acuerdo al principio de valor esperado

$$P^{(1)} (g_j) = (1 + \alpha_j) E [S^{(1)} - g_j]$$

Se usará el principio de la mínima varianza. La variable aleatoria

$$Z = \sum_{i=1}^N \{ \tilde{P}^{(i)} - P^{(i)}(g_i) - g_i(S^{(i)}) \}$$

denota la "ganancia remanente dentro de la retención", donde $\tilde{P}^{(i)}$ es la prima que el asegurador directo recibe por el riesgo con la variable de reclamo acumulado $S^{(i)}$.

Como las g_i son independientes, se tiene

$$\begin{aligned} E[Z] &= \sum_{i=1}^N \{ \tilde{P}^{(i)} - (1 + \alpha_i) E[S^{(i)} - g_i] - E[g_i] \} \\ &= \sum_{i=1}^N \{ \tilde{P}^{(i)} - (1 + \alpha_i) E[S^{(i)} - g_i] + \alpha_i E[g_i] \} \end{aligned}$$

y

$$\text{Var}[Z] = \sum_{i=1}^N \text{Var}[g_i]$$

El objetivo es de nuevo hacer $\text{Var}[Z]$ tan pequeña como sea posible para un $E[Z]$ dado.

La función de Lagrange tiene la forma

$$\phi = \text{Var}[Z] + \lambda E[Z]$$

$$= \sum_{i=1}^N \text{Var}[g_i] + \lambda \sum_{i=1}^N \{ \tilde{P}^{(i)} - (1 + \alpha_i) E[S^{(i)} - g_i] + \alpha_i E[g_i] \}$$

y por lo tanto

$$\frac{\delta \phi}{\delta M_j} = \frac{\delta}{\delta M_j} \text{Var}[g_j] + \lambda \alpha_j \frac{\delta}{\delta M_j} E[g_j] = 0$$

$$\frac{\delta}{\delta M_j} \text{Var}[g_j] = \lambda k \alpha_j \frac{\delta}{\delta M_j} E[g_j]$$

c) Solo falta introducir las expresiones para $E [g_j]$ y $\text{Var} [g_j]$ encontradas en el inciso a) en la relación anterior

$$\frac{\delta}{\delta M_j} E [g_j] = E [A_j] [1 - F(M_j)]$$

$$\begin{aligned} \frac{\delta}{\delta M_j} \text{Var} [g_j] &= 2 M_j E [A_j] [1 - F(M_j)] \\ &+ 2 [1 - F(M_j)] [\text{Var} [A_j] - E [A_j]] \int_0^{M_j} [1 - F(x)] dx \end{aligned}$$

Sustituyendo ésto en

$$\frac{\delta}{\delta M_j} \text{Var} [g_j] = \lambda k \alpha_j \frac{\delta}{\delta M_j} E [g_j]$$

se obtiene

$$M_j E [A_j] + (\text{Var} [A_j] - E [A_j]) \int_0^{M_j} [1 - F^{(p)}(x)] dx = k \alpha_j E [A_j]$$

$$M_j = k \alpha_j - \frac{\text{Var} [A_j] - E [A_j]}{E [A_j]} \int_0^{M_j} [1 - F^{(p)}(x)] dx$$

Usando la abreviatura

$$\zeta_j = \frac{\text{Var} [A_j] - E [A_j]}{E [A_j]}$$

se tiene

$$M_j = k \alpha_j - \zeta_j \int_0^{M_j} [1 - F^{(p)}(x)] dx$$

Se debe notar que $\epsilon_j = 0$ para una variable A_j , que es el número de reclamos, distribuida como Poisson, con el parámetro de riesgo conocido.

Si se examinan las suposiciones hechas bajo b), se encuentra que la representación

$$g_i = \sum_{j=0}^{A_i} Y_j^{(i)} (M_j) \text{ con } Y_j^{(i)} (M_j) \text{ independientes e idénticamente}$$

distribuidas es solamente posible si el parámetro de riesgo para $Y_j^{(i)} (M_j)$ y aquel de la función de distribución $F^{(i)}(x)$ se conocen. Por ello, hagase la siguiente suposición:

La variable aleatoria $S^{(i)} = \sum_{j=0}^{A_i} Y_j^{(i)}$ se describe

- $P [A_j = k] = P_k^{(i)}(\phi)$
- $P [Y_j^{(i)} \leq x] = F^{(i)}(x)$ independiente de ϕ
- La convención de que $Y_j^{(i)}$ ($j = 1, 2, \dots$) son no negativas, independiente e idénticamente distribuidas.

Haciendo válida la suposición, se obtiene

$$M_j = k \alpha_j - \frac{\tau_j^2(\phi) - \nu_j(\phi)}{\nu_j(\phi)} \int_0^{M_j} [1 - F^{(j)}(x)] dx$$

$$\text{donde } \nu_j(\phi) = \sum_{k=0}^{\infty} k P_k^{(i)}(\phi)$$

$$\text{y } \tau_j^2(\phi) = \sum_{k=0}^{\infty} [k - \nu_j(\phi)]^2 P_k^{(j)}(\phi)$$

Esta fórmula es de particular interés si la variable de conteo de reclamo de riesgo se distribuye como Poisson, en éste caso, $\tau_j^2(\phi) = \nu_j(\phi)$ y entonces

$$M_j = k \alpha_j$$

Si de determinan los grados que resultan de la solución de Finetti al problema de la retención relativa en el reaseguro proporcional si se conocen todos los parámetros de riesgo y la prima de reaseguro se calcula como

$$P^{(r)} = (1 + \lambda) \mu_j(\phi)$$

se obtiene que:

$$\lambda_j = \frac{P^{(r)} - \mu_j(\phi)}{\sigma_j^2(\phi)} = \frac{(1 + \lambda) \mu_j(\phi) - \mu_j(\phi)}{\sigma_j^2(\phi)} = \frac{\mu_j(\phi) (1 + \lambda - 1)}{\sigma_j^2(\phi)}$$

$$= \frac{\lambda \mu_j(\phi)}{\sigma_j^2(\phi)}$$

$$a_j = C \lambda_j = C \frac{\lambda \mu_j(\phi)}{\sigma_j^2(\phi)}$$

El problema de la retención absoluta

Siguiendo la aproximación vista anteriormente en los aspectos generales se tiene que determinar el nivel de retención para una graduación dada que se calcula con el principio de Finetti. Esto se llama el problema de la retención absoluta.

Formúlese una graduación como sigue:

a) Para el reaseguro proporcional esto se da por

$$V(a_j, S^{(p)}) = C r_j$$

donde

V = función que asocia un número con la distribución de probabilidad del proceso del reclamo acumulado $S^{(p)}$ y la retención proporcional a_j .

γ_j = cargo de porcentaje sobre el valor esperado contenido en la prima de reaseguro. i.e. $P^{(j)} - E[S^{(j)}] = L^{(j)} = \gamma_j E[S^{(j)}]$
 C = constante por determinar

b) Para el reaseguro no proporcional esto es

$$W(M_j, S^{(j)}) = K \alpha_j$$

donde

W = función que asocia un número con la distribución de probabilidad del proceso de reclamo acumulado $S^{(j)}$ y el límite M_j .

α_j = Cargo de porcentaje sobre el valor esperado contenido en la prima de reaseguro

K = constante por determinar.

En estas fórmulas el problema de la retención absoluta consiste simplemente en determinar los valores apropiados de C y de K.

Es fácil ver como la graduación óptima de acuerdo con los principios de la retención relativa concuerda con la formulación dada aquí, para escoger V y W como sigue

$$V^*(a_j, S^{(j)}) = \text{Var} [a_j, S^{(j)}] / E [a_j, S^{(j)}]$$

y

$$W^*(M_j, S^{(j)}) = M_j + \int_0^{M_j} [1 - F^{(j)}(x)] dx$$

Para escoger V y W de una masa de riesgo que parte es reasegurada sobre una base proporcional y parte sobre una no proporcional, las dos constantes C y K son idénticas. Por ello es fácil ver que dos formas de reaseguro coinciden para un riesgo el cual siempre produce el monto de reclamo máximo (completo), si un reclamo ocurre.

Sea $S^{(j)}$ igual a $\sum_{i=0}^A Y_i^{(j)}$, donde $Y_i^{(j)}$ es un número fijo (ya

no es más una variable aleatoria) independiente de t , y escójase a_j y M_j tal que $a_j Y_i^{(j)} = M_j$. Entonces

$$V^*(a_j, S^{(j)}) = \frac{\text{Var} \left[\sum_{i=0}^{A_j} a_j Y_i^{(j)} \right]}{E \left[\sum_{i=0}^{A_j} a_j Y_i^{(j)} \right]} = \frac{\text{Var} [A_j] M_j^2}{E [A_j] M_j} = M_j (\tau_j + 1)$$

Del otro lado

$$W^*(M_j, S^{(j)}) = M_j + \tau_j M_j = M_j (\tau_j + 1)$$

donde $\int_0^{M_j} [1 - F^{(j)}(x)] dx = M_j$

En orden para obtener el riesgo reasegurado exactamente en la misma forma bajo reaseguro no proporcional y proporcional, K debe ser igual a C .

Finalmente se llega a la siguiente formulación del problema de la retención absoluta.

- Sea V y W las funciones anteriores para reaseguro proporcional y no proporcional respectivamente.

- Sea γ_j y α_j los cargos de porcentajes sobre el valor esperado contenido en las primas de reaseguro proporcional y no proporcional respectivamente.

Entonces sean dadas las siguientes graduaciones de retenciones por

$$V(a_j, S^{(j)}) = C \gamma_j$$

$$W(M_j, S^{(j)}) = K \alpha_j$$

Es característico del problema de la retención absoluta que ninguna solución es posible sin alguna formulación de los objetivos de negocios de la cartera de riesgo. Esto es en contraste con las consideraciones sobre graduaciones de

retenciones (problema de retención relativa) donde se emplee el principio de mínima varianza que podría ser aceptable dentro de cualquier contexto razonable de objetivos del negocio.

Antes de formular estos objetivos se deben describir las consecuencias de varios niveles de retención. Tales consecuencias surgen del hecho de que las reservas libres de la cartera de riesgo fluctúan de acuerdo con estos niveles de retención. Se habla de estas fluctuaciones como la caminata aleatoria de las reservas libres de la cartera de riesgo. Para una masa de N riesgos la última tiene la siguiente apariencia:

Sea $S_i^{(t)}$ ($i = 1, 2, \dots, N$) el proceso de reclamos acumulados del riesgo i . Por simplicidad se supondrá que para el parámetro de riesgo conocido este proceso es Poisson compuesto, esto es

$$S_i^{(t)} = \sum_{j=1}^{N_j} Y_i^{(j)}$$

Para el riesgo con parámetro de riesgo ϕ se tiene

$$G_i^{(t)}(x|\phi_i) = \sum_{k=0}^{\infty} \exp[-\mu_i(\phi_i)t] \frac{(\mu_i(\phi_i)t)^k}{k!} F^{k * x}(x|\phi_i)$$

que es la función de distribución del reclamo acumulado $S_i^{(t)}$, donde

$$\psi_i^{(t)}(u|\phi_i) = \exp[\mu_i(\phi_i)t(x_i(u|\phi_i) - 1)]$$

representa la función característica del reclamo acumulado $S_i^{(t)}$.

Sobre las retenciones se ha supuesto que los procesos de los reclamos acumulados son independientes unos de otros. Se seguirá ahora el argumento usando la hipótesis para el caso en el cual todos los parámetros de riesgo se conocen:

Primero que nada, se investigará la distribución de probabilidad del proceso del reclamo acumulado para la masa de riesgo completa, i.e. para

$$S_t = \sum_{i=1}^N S_t^{(i)}$$

como las $S_t^{(i)}$ son independientes se tiene

$$\begin{aligned} \psi_{S_t}(u) &= \prod_{i=1}^N \psi_t^{(i)}(u) = \prod_{i=1}^N \exp[\mu_i(\phi_i)t(x_i(u|\phi_i) - 1)] \\ &= \exp\left[\sum_{i=1}^N \mu_i(\phi_i)t(x_i(u|\phi_i) - 1)\right] = \\ &= \exp\left\{\mu^*(\phi_1, \dots, \phi_N)t(x^*(u, \phi_1, \dots, \phi_N) - 1)\right\} \end{aligned}$$

donde

$$\mu^*(\phi_1, \dots, \phi_N) = \sum_{i=1}^N \mu_i(\phi_i)$$

y

$$x^*(u, \phi_1, \dots, \phi_N) = \sum_{i=1}^N \frac{\mu_i(\phi_i)}{\mu^*} x_i(u|\phi_i)$$

Entonces se tiene el siguiente resultado

Teorema: Si el proceso de reclamos es una Poisson compuesta e independientes son entre si, entonces el proceso del reclamo acumulado de la masa de riesgo completa es tambien una Poisson compuesta.

Las dos relaciones anteriores da la intensidad (condicional) y la distribución del monto de reclamo acumulado (condicional) para este proceso.

En comparación con el proceso de reclamo acumulado S_t se tiene el ingreso por primas acumulado

$$P_t^* = \sum_{i=1}^N P_t^{(i)}$$

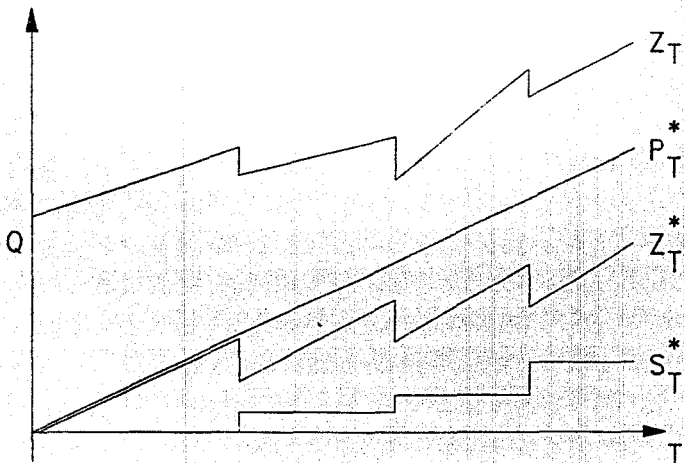
donde $P_t^{(i)}$ es la prima acumulada de la cartera de riesgo i en el intervalo de tiempo $[0, t]$. La diferencia

$$Z_t^* = P_t^* - S_t$$

representa la caminata aleatoria de la ganancia realizada sobre el riesgo en cuestión. Si tambien se toma en cuenta las reservas libres disponibles de la cartera de riesgo Q al tiempo $t = 0$, entonces

$$Z_t = Q + Z_t^*$$

U.N.A.M.
E.N.E.P. ACATLAN
REASEGURO. RETENCION ABSOLUTA



que es la caminata aleatoria de las reservas libres de la cartera de riesgos. El siguiente diagrama hace más claras estas diferencias

Aunque se definió el problema de la retención absoluta, no se resolvió porque esto solamente se puede hacer dando una formulación de los objetivos del negocio. Si se escoge la probabilidad de ruina como el criterio de estabilidad, se define entonces un objetivo de negocio. Intuitivamente, esta opción se puede explicar mejor diciendo que la seguridad se tomara como el objetivo más grande de la empresa.

Las suposiciones dadas anteriormente son dada una graduación de retenciones

a) $V(a_j, S^{(j)}) = K \gamma_j$ Para el reaseguro proporcional

b) $W(M_j, S^{(j)}) = K \alpha_j$ Para el reaseguro no proporcional

se requiere una constante K tal que para el proceso de reclamo acumulado de la masa de riesgo completa remanente dentro de la retención la probabilidad de ruina satisfice

$$\psi(u) \leq P_0$$

para una u dada (usualmente grande) y P_0 (usualmente pequeña)

Es intuitivamente obvio que, previniendo que los cargos en la porción retenida no son cambiados, mientras mas pequeña sea K, mas pequeña es la probabilidad de ruina $\psi(u)$. Como hay un segundo interés de retener lo mas posible de riesgos y su consiguiente ganancia esperada, el caso $\psi(u) = P_0$ es de particular significancia en la práctica.

En términos de la función característica, se tiene

$$V_{S_1}(u) = \int_{\Theta} \exp\{\mu^*(\theta) t (x^*(u) - 1)\} dU(\theta)$$

para el proceso de reclamo acumulado, donde θ aparece como un parámetro general. Sobre la base de A se tiene

$$\mu^*(\theta) = \sum_{i=1}^N \mu_i(\theta)$$

y

$$x^*(u) = \sum_{i=1}^N \frac{\mu_i(\theta)}{\mu^*(\theta)} x_i(u)$$

Para el proceso del reclamo acumulado retenido $g_k(S_1)$ se tiene (el índice k es para recordar la constante usada en la graduación de retenciones):

$$V_{g_k(S_1)}(u) = \int_{\Theta} \exp\{\mu^*(\theta) t (x^*(u|k) - 1)\} dU(\theta)$$

donde $x^*(u|k) = \sum_{i=1}^N \frac{\mu_i(\theta)}{\mu^*(\theta)} x_i(u|k)$ con $x_i(u|k) = x_i(a_i u)$

en el caso de reaseguro proporcional y

$$x_i(u|k) = \int_0^{M_i} e^{-ux} dF_i(x) + \exp\{uM_i\} \int_{M_i}^{\infty} dF_i(x)$$

en el caso de reaseguro no proporcional (exceso de pérdida), siendo a_i la graduación de retención proporcional resultante de la constante k en la graduación de retención y M_i el primer riesgo para el i-ésimo riesgo resultante de la misma constante k.

Los procesos S_1 y $g_k(S_1)$ son sujetos a la misma ley de

probabilidades incluyendo la sustitución $x_i(u|k)$ para $x^*(u)$.

Por último se adherirá el supuesto del ingreso de la prima lineal

$$P_i^* = ct$$

Notese que P_i^* en el problema de la retención absoluta se deberá interpretar como la "prima remanente" despues de la deducción de la prima de reaseguro.

La graduación óptima de retenciones.

En lo subsecuente se trabajara con la regla óptima más que con las reglas generales:

$$V^*(a_j, S^{jp}) = \text{Var} [a_j S^{jp}] / E [a_j S^{jp}] = k \gamma_j$$

$$W^*(M_j, S^{jp}) = M_j + \zeta_j \int_0^{M_j} [1 - F_j(x)] dx = k \alpha_j$$

Recuérdese la relación $\zeta_j = \frac{\text{Var} [A_j] - E [A_j]}{E [A_j]}$

donde A_j denota la variable de conteo del número de reclamos en el periodo de calculo [0,1].

En base al supuesto de que los montos de reclamos de riesgos son independientes, se puede emplear la fórmula:

$$E [a_j S^{jp}] = E [A_j] E [a_j Y^{jp}] = a_j E [A_j] E [Y^{jp}]$$

donde Y_i^{jp} es el i-ésimo monto de reclamos del riesgo j y

$$\begin{aligned} \text{Var} [a_j S^{jp}] &= E [A_j] \text{Var} [a_j Y^{jp}] + \text{Var} [A_j] E^2 [a_j Y^{jp}] \\ &= a_j^2 E [A_j] \{ \text{Var} [Y^{jp}] + (1 + \zeta_j) E^2 [Y^{jp}] \} \\ &= a_j^2 E [A_j] E [(Y^{jp})^2] + a_j^2 \zeta_j E [A_j] E^2 [Y^{jp}] \end{aligned}$$

$$V^*(a_j, S^{(j)}) = a_j \frac{E[(Y^{(j)})^2] + r_j E^2[Y^{(j)}]}{E[Y^{(j)}]} = k r_j$$

En el caso del riesgo con parámetro conocido se caracteriza por $r_j = 0$ (Variable de conteo Poisson)

La condición de estabilidad

Ahora se tiene el problema de estimar la probabilidad de ruina $\psi(u|k)$ para el proceso de reclamo retenido acumulado $g_k(S_1)$ de la masa de riesgo total. Una constante K es admisible si

$$\psi(u, k) \leq P_0 = \exp [P_k(\theta)u]$$

y considérese k como admisible si la condición más estricta

$$\int \exp [P_k(\theta)u] du(\theta) \leq 0$$

se cumple.

Es claro, en virtud de

$$\psi(u, k) \leq P_0 = \exp [P_k(\theta)u]$$

que el valor de k que es admisible en los términos de

$$\int \exp [P_k(\theta)u] du(\theta) \leq 0$$

garantiza una probabilidad de ruina $\leq P_0$. Lo inverso no se cumple, sin embargo, así que en el caso límite

$$\int \exp [P_k(\theta)u] du(\theta) = 0$$

se entenderá en general una constante K para el cual $\psi(u|k) < P_0$. Desde el punto de vista de la seguridad, sin embargo, no se está en el lado seguro, se debe aceptar el hecho que se puede obtener una fórmula clara y laborable solamente al precio de tener valores de k que son algunas veces demasiado bajos.

Determinando la retención absoluta cuando el parámetro de riesgo se conoce

a) General

El proceso de reclamo retenido acumulado $g_k(S)$ de la masa de riesgo entera tiene la forma

$$v_{g_k(S)}^{(\theta)}(u) = \exp(\mu^*(\theta) \cdot (x^*(u|k) - 1))$$

Se establece una constante $R^{(\theta)}$ determinada por la ecuación

$$1 + \frac{C}{\mu^*} R - x^*(R|k) = 0$$

$x^*(u|k)$ se entiende como la función generadora de momentos. Se tiene que

$$x^*(u|k) = \sum_{i=1}^N \frac{\mu_i}{\mu^*} x^*(R|k) = \sum_{i=1}^N \frac{\mu_i}{\mu^*} \int_{-\infty}^{\infty} e^{ux} dF_i(x|k)$$

De la expansión

$e^{Rx} = 1 + Rx + \frac{(Rx)^2}{2} + (\text{términos en tercera y demás potencias})$
se encuentra que, si no se toman en cuenta los términos entre paréntesis,

$$x^*(u|k) \approx 1 + R \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(1)}|k] + \frac{R^2}{2} \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(1,2)}|k]$$

y sustituyendo lo anterior en la ecuación para obtener R:

$$1 + \frac{C}{\mu^*} R = x^*(R|k)$$

$$1 + \frac{C}{\mu^*} R = 1 + R \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(1)}|k] + \frac{R^2}{2} \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(1,2)}|k]$$

$$R = 2 \frac{C - \sum \mu_i E [Y^{(i)} | k]}{\sum \mu_i E [(Y^{(i)})^2 | k]}$$

R = 0 si el resultado anterior produce un valor negativo. En este caso la estimación da un valor 1 para la probabilidad de ruina, independientemente de las reservas libres iniciales u.

b) Para el reaseguro proporcional

Como se trata con parametros de riesgo dados, se tiene $\gamma_j = 0$ para toda j y entonces

$$\delta_j \frac{E [(Y^{(j)})^2]}{E [Y^{(j)}]} = \frac{E [(Y^{(j)})^2 | k]}{E [Y^{(j)} | k]} = k \gamma_j$$

Por lo tanto

$$E [(Y^{(j)})^2 | k] = k \gamma_j E [Y^{(j)} | k]$$

y sustituyendo lo anterior en

$$R = 2 \frac{C - \sum \mu_i E [Y^{(i)} | k]}{\sum \mu_i E [(Y^{(i)})^2 | k]}$$

se produce

$$R = 2 \frac{C - \sum \mu_i E [Y^{(i)} | k]}{k \sum \mu_i \gamma_i E [(Y^{(i)} | k]}$$

R se puede calcular de esta fórmula. La relación se vuelve más clara si se representa con δ_i el cargo proporcional añadido al valor esperado en la prima de seguro retenida (después de la deducción de la prima de reaseguro) y por γ_i como el cargo añadido al valor esperado en la prima de reaseguro.

Se tiene

$$P_i^* = ct = \sum (1 + \delta_i) \mu_i E [Y^{(i)} | k] t$$

y sustituyendo se obtiene

$$R = \frac{2}{k} \frac{\sum \mu_i \delta_i E [Y^{(i)} | k]}{\sum \mu_i \gamma_i E [Y^{(i)} | k]}$$

Se debe notar que $2/k$ se multiplica por un factor el cual representa las ganancias esperadas del reaseguro (lo que el reasegurador alcanzaria por aceptar la retención).

Si se deja $\gamma_i = \rho \delta_i$ (razón fija de cargo proporcional en la porción retenida en relación con el reaseguro) se tiene de inmediato

$$R = \frac{2}{\rho k}$$

La expresión anterior es válida aun sin la última hipótesis si se define

$$\rho = \frac{\sum \mu_i \gamma_i E [Y^{(i)} | k]}{\sum \mu_i \delta_i E [Y^{(i)} | k]} = \frac{\text{Ganancias del reaseguro esperadas si la retención fuera tomada a cuenta.}}{\text{Ganancias retenidas esperadas}}$$

Se encuentra entonces que

$$\psi^{(0)}(u) \leq \exp [-2u/\rho k]$$

y de la relación $\exp [(-2u/\rho) k] = P_0$ se obtiene

$$k = 2u/[\rho |\ln P_0|]$$

donde

u = reservas iniciales libres

ρ = razón de cargos añadidos al valor esperado en la prima de reaseguro y prima retenida

c) Para el reaseguro no proporcional (exceso de pérdida)

Como el parámetro del riesgo es dado para cada riesgo se tiene:

$M_j = k \alpha_j$ (α_j denota el cargo sobre el valor esperado en la prima

de reaseguro)

Si se usa el estimado

$$E[(Y^{(i)})^2 | k] \leq M_i E[Y^{(i)} | k]$$

en la expresión

$$R \geq 2 \frac{C - \sum \mu_i E[Y^{(i)} | k]}{k \sum \mu_i \gamma_i E[(Y^{(i)})^2 | k]}$$

y la misma definición para δ_i como en la sección precedente, del cargo de porcentaje sobre el valor esperado en la retención, se encuentra que

$$R \geq 2 \frac{\sum \mu_i \delta_i E[Y^{(i)} | k]}{\sum \mu_i M_i E[(Y^{(i)})^2 | k]} = \frac{2}{k} \frac{\sum \mu_i \delta_i E[Y^{(i)} | k]}{\sum \mu_i \alpha_i E[(Y^{(i)})^2 | k]}$$

si se supone que

$$\alpha_i = \tau \delta$$

entonces la relación

$$R \geq (2/\tau) k$$

se obtiene automáticamente de la última expresión.

Escogiendo $R = (2/\tau) k$ y haciendo uso de la igualdad $\exp[(-2u/\tau) k] = P_0$ se tiene

$$k = 2u/[\tau |\ln P_0|]$$

donde

u = reservas iniciales libres

\dot{r} = razón de cargos añadidos al valor esperado en la prima de reaseguro y prima retenida

P_0 = probabilidad de ruina admisible

d) Para el reaseguro mixto.

Supóngase que una porción de la masa de riesgo, los riesgos de 1 a N_p , se aseguran sobre una base proporcional y que la otra parte, los riesgos N_{p+1} a N , se aseguran sobre una base no proporcional.

Se puede descomponer los riesgos en dos partes descritas y proceder como en las secciones anteriores. Sea

E_p = ganancia esperada del asegurador en la retención de los riesgos que son reasegurados proporcionalmente y

E_N = Ganancia esperada del asegurador en la retención de los riesgos que son reasegurados no proporcionalmente.

Usando la misma notación que la anterior se tiene

$$k = \frac{2u}{|\ln P_0|} \frac{E_p + E_N}{\rho E_p + r E_N}$$

e) Estimando el error de la aproximación

En a) se dejó

$$e^{Rx} \approx 1 + Rx + \frac{(Rx)^2}{2}$$

El error envuelto aquí es a lo más (= remanente de los términos de la serie de Taylor)

$$(1/6) \exp[|Rx|] R^3 |x|^3$$

Suponiendo $F_i(0|k) = 0$ para toda i y cualquier valor de k (reclamos no negativos), el error en $x^*(R|k)$ suma a lo más

$$\zeta(R) = \sum_{i=1}^N \frac{\mu_i}{\mu^*} \frac{R^2}{8} \int_0^{\infty} \exp[Rx] x^3 dF_i(x|k)$$

De la ecuación

$$1 + \frac{C}{\mu^*} R = 1 + R \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(1)}|k] + \frac{R^2}{2} \sum_{i=1}^N \frac{\mu_i}{\mu^*} E[Y^{(2)}|k] + \zeta(R)$$

se encuentra una nueva aproximación para R , que se llamara R_1 . Si la ecuación sin el término de error $\zeta(R)$ produce el valor estimado R_0 usado hasta ahora, entonces se conoce que para la raíz correcta R de

$$1 + \frac{C}{\mu^*} R - \chi^*(R|k) = 0$$

$$R_1 \leq R \leq R_0$$

El límite R_1 se debiera calcular cuando se trabaje con un ejemplo concreto. En el caso general la fórmula se vuelve muy complicada para ser fácilmente manejada.

Capítulo 4. Probabilidad de Ruina

4.1. Introducción

Un problema en el riesgo del seguro.

Considèrese un riesgo de seguro de una Cia. con capital inicial b , primas ascendentes μ por unidad de tiempo, y los reclamos ocurren aleatoriamente a una tasa promedio ν . El monto de los reclamos sigue la distribución con función generadora de momentos $M(\theta)$. Se requiere conocer la probabilidad F_b de que quiebre la Cia.

Este problema es una generalización del problema de la ruina de un jugador. Aquí se tienen sucesiones aleatorias en donde la variable X_r en el tiempo t_r es independiente del conjunto previo entero de X 's. El interés estadístico de tales sucesiones reside en las propiedades de variables o sucesiones derivadas, tales como las sumas acumuladas

$$S_r = X_1 + X_2 + \dots + X_r$$

El proceso S_r se llama caminata aleatoria, en la medida en que represente la posición al tiempo t_r de una persona dando un paso aleatorio X_r independientemente de los previos. Como un caso especial se tiene

$$P\{X_r = 1\} = p \quad P\{X_r = -1\} = 1-p$$

Este ejemplo representa el problema de la ruina de un jugador, donde, supóngase, que dos jugadores A y B juegan una sucesión de juegos, con la probabilidad de que A gane un juego igual a p . Si gana el juego, adquiere una cantidad de B y si pierde, perderá una unidad de su dinero.

La distribución de la suma S_r es la r -ésima convolución de la función de distribución $F(x)$ de cada X_r , o lo que es lo mismo,

en términos de funciones cumulantes (Parzen E. Teoría Moderna de Probabilidades", pag 438)

$$K_r = rK$$

donde K_r es la función cumulante de S_n y K es la de cada X_r . La función de distribución de S_r (con media cero y variancia finita), tiende a la normalidad en la medida en que r aumente, siempre y cuando la media m y variancia σ^2 (> 0) de la distribución común de que cada X_r exista.

4.2. Discusión de los diferentes métodos.

En el problema original del juego, un punto de interés era la naturaleza de la sucesión S_r si esta era limitada por el capital inicial de cada jugador, y, en particular, la duración del juego. Por lo tanto, considérese los límites $a < 0$ y $b > 0$, tal que si S_r alcanza primero o sale de los límites del intervalo (a, b) en el tiempo t_r , el proceso se termina. Denótese la función de distribución modificada de S_n por $F_n(x)$, i.e.

$$F_n(x) = P \{S_n \leq x \text{ y } a < S_r < b \quad (r = 1, 2, \dots, n-1)\}$$

entonces se tiene la relación de recurrencia

$$F_n(x) = \int_a^{-b} F(x-y) dF_{n-1}(y)$$

La probabilidad P_n de alcanzar el límite b por primera vez en n (ganar el juego) es:

$$P_n = F_n(\infty) - F_n(b-)$$

Similarmente para a se tiene (perder el juego):

$$P_n' = F_n(a) - F_n(-\infty)$$

Si se pone

$$P(\lambda) = \sum_{r=1}^{\infty} \lambda^{r-1} p_r \qquad P'(\lambda) = \sum_{r=1}^{\infty} \lambda^{r-1} p'_r$$

entonces

$$P(\lambda) = G(\infty, \lambda) - G(b-, \lambda), \qquad P'(\lambda) = G(a, \lambda) - G(-\infty, \lambda)$$

donde

$$G(x, \lambda) = \sum_{r=1}^{\infty} \lambda^{r-1} F_r(x)$$

satisface la ecuación integral de Fredholm

$$G(x, \lambda) = F(x) + \lambda \int_a^{b-} F(x-y) dG(y, \lambda)$$

Esta ecuación, debida a Samuelson (1948), se puede resolver en base a los principios antes mencionados, como se puede notar en el siguiente desarrollo:

$$\begin{aligned} G(x, \lambda) &= \sum_{r=1}^{\infty} \lambda^{r-1} F_r(x) = F(x) + \sum_{r=2}^{\infty} \lambda^{r-1} F_r(x) = \\ &= F(x) + \lambda \sum_{r=2}^{\infty} \lambda^{r-2} F_r(x) = \\ &= F(x) + \lambda \sum_{r=1}^{\infty} \lambda^{r-1} \int_a^{b-} F(x-y) dF_r(y) = \\ &= F(x) + \lambda \int_a^{b-} F(x-y) \sum_{r=1}^{\infty} \lambda^{r-1} dF_r(y) = \\ &= F(x) + \lambda \int_a^{b-} F(x-y) dG(y, \lambda) \end{aligned}$$

Ahora se procederá con otro método más general para resolver el caso de la ruina de un jugador, debido a Wald. Primero se establecen dos lemmas:

Lema I.- Denótese el valor aleatorio de n al cual S_n alcanza por primera vez los límites (a, b) por N . Entonces

$$n_0^k P(N \geq n_0) \rightarrow 0 \quad \text{cuando } n_0 \rightarrow \infty$$

para cualquier k finita. Como $\sigma^2 > 0$, $E[S_n^2] \rightarrow \infty$ cuando $n \rightarrow \infty$. En consecuencia, hay una k tal que

$$P\{S_k^2 < (b-a)^2\} < 1 \leq 1 - \lambda$$

De aquí que para $n_0 = hk$

$$P\{(S_{k(r+1)} - S_{kr})^2 < (b-a)^2 \text{ para toda } r = 0, \dots, h-1\} \leq (1-\lambda)^{n_0/k}$$

debido a que los incrementos en X_n son independientes.

Cualquiera de las h desigualdades $|S_{k(r+1)} - S_{kr}| \geq (b-a)$ asegura que $N < n_0$. Entonces finalmente se demostró que

$$P(N \geq n_0) \leq (1-\lambda)^{n_0/k} = O(\exp[-\mu n_0]) = o(n_0^{-\alpha}) \text{ cuando } n_0 \rightarrow \infty$$

La probabilidad de que S_n alcance por primera vez y rebase el límite n_0 tiende a cero cuando este límite tiende a ∞ , es decir, la probabilidad de que continúe el juego después de haber alcanzado este límite es nula.

Lemma II.- Denótese la función generadora de momentos de cada X_r por $M(\theta)$, suponiendo que existe para todo real θ , y que tiene la propiedad

$$M(\theta) \rightarrow \infty \quad \text{cuando } \theta \rightarrow \frac{1}{\lambda}$$

Esta última propiedad se cumple si

$$P\{e^X < 1 - \delta\} > 0 \quad \text{y} \quad P\{e^X > 1 + \delta\} > 0 \quad \text{para alguna } \delta > 0$$

Tambien es válido que

$$M'(\theta) = E \{ X e^{\theta X} \}$$

$$M''(\theta) = E \{ X^2 e^{\theta X} \}$$

donde las comillas significan diferenciación con respecto a θ . Entonces si $m = M'(0) \neq 0$, hay una y solamente una raíz real $\theta_0 \neq 0$ tal que $M(\theta_0) = 1$.

Se nota que $M''(\theta) = E \{ X^2 e^{\theta X} \} > 0$ para todo real θ , en consecuencia, como $M(\theta) \rightarrow \infty$ cuando $\theta \rightarrow \pm \infty$, entonces $M(\theta)$ tiene solamente un mínimo θ_1 . Pero $M'(0) \neq 0$, en consecuencia $\theta_1 \neq 0$ y $M(\theta_1) < M(0) = 1$. El resultado requerido se obtiene

IDENTIDAD DE WALD

$$E \{ \exp \{ \theta S_N \} \} = E \{ [M(\theta)]^{-N} \exp \{ \theta S_N \} \} = 1 \quad (|M(\theta)| \geq 1)$$

Para demostrar la identidad de Wald, denótese a j como una constante entera y $P_j = P \{ N \leq j \}$. Entonces

$$E \{ \exp \{ \theta S_j \} \} = P_j E_j \{ \exp \{ \theta S_j \} \} + Q_j E_j' \{ \exp \{ \theta S_j \} \}$$

donde $Q_j = 1 - P_j$. E_j denota la esperanza condicional de que $N > j$. Ahora para cualquier fijo $j > N$, $S_j - S_N$ es independiente de S_N , y

$$E_j \{ \exp \{ \theta S_j \} \} = E_j \{ \exp \{ \theta S_j \} [M(\theta)]^{j-N} \}$$

$$E \{ \exp \{ \theta S_j \} \} = [M(\theta)]^j; \quad S_j = S_1 + \dots + S_j$$

En consecuencia, por lo anterior, se tiene

$$P_j E_j \{ \exp \{ \theta S_N \} [M(\theta)]^{-N} \} + Q_j E_j' \{ \exp \{ \theta S_j \} [M(\theta)]^{-j} \} = 1$$

Ahora $Q_j \rightarrow 0$ cuando $j \rightarrow \infty$ y $E_j'(\exp[\theta S_j])$ está limitado, (en la medida que esta esperanza es condicional sobre $N > j$, así que $|M(\theta)| \geq 1$, se obtiene:

$$\lim_{j \rightarrow \infty} P_j E_j (\exp[\theta S_N] [M(\theta)]^{-N}) + Q_j E_j' (\exp[\theta S_j] [M(\theta)]^{-j}) = 1$$

$$\lim_{j \rightarrow \infty} P_j = P(N < j) = P(N < \infty) = 1$$

$$\lim_{j \rightarrow \infty} E_j (\exp[\theta S_N] [M(\theta)]^{-N}) = E (\exp[\theta S_N] [M(\theta)]^{-N})$$

$$\lim_{j \rightarrow \infty} Q_j = 0 \quad (\text{Lema I})$$

$$\lim_{j \rightarrow \infty} P_j E_j (\exp[\theta S_N] [M(\theta)]^{-N}) = E (\exp[\theta S_N] [M(\theta)]^{-N}) = 1 \text{ IDENTIDAD}$$

DE WALD.

En la prueba anterior se hizo uso del lema I para $s = 0$, se puede demostrar que la identidad de Wald se puede derivar respecto a θ cualquier número de veces bajo el signo de esperanza. También se debe notar que la identidad de Wald es válida si $a \rightarrow -\infty$, previendo que $P_j \rightarrow 1$ y que la parte real de θ es no negativa. Un caso especial es cuando $X \leq 0$: aquí las condiciones del lema II no son consistentes, y entonces no hay raíz $\theta_0 \neq 0$ para el cual $M(\theta) = 1$.

Para aplicar el resultado anterior al caso donde una u otra limitante es exactamente alcanzada en la etapa N se pone $S_N = a$ con probabilidad P_a y $S_N = b$ con probabilidad $P_b = 1 - P_a$. Entonces

$$P_a E_a (\exp[a\theta] [M(\theta)]^{-N}) + P_b E_b (\exp[b\theta] [M(\theta)]^{-N}) = 1$$

donde E_a denota la esperanza condicional $S_N = a$. Si en la ecuación anterior se pone $\theta = \theta_0$, se obtiene

$$P_a E_a (\exp [a\theta_0] [M(\theta_0)]^{-N}) + P_b E_b (\exp [b\theta_0] [M(\theta_0)]^{-N}) = 1$$

$$P_a E_a (\exp [a\theta] [M(\theta)]^{-N}) + (1 - P_a) E_b (\exp [b\theta] [M(\theta)]^{-N}) = 1$$

$$P_a E_a (\exp [a\theta] [M(\theta)]^{-N}) + E_b (\exp [b\theta] [M(\theta)]^{-N})$$

$$- P_a E_b (\exp [b\theta] [M(\theta)]^{-N}) = 1$$

$$P_a \exp [a\theta_0] + \exp [b\theta_0] - P_a (\exp [b\theta_0]) = 1$$

$$P_a (\exp [a\theta_0] - \exp [b\theta_0]) = 1 - \exp [b\theta_0]$$

$$P_a = \frac{1 - \exp [b\theta_0]}{\exp [a\theta_0] - \exp [b\theta_0]}$$

Por ejemplo, en el problema del juego se tiene

$$M(\theta) = p \exp[\theta] + q \exp[-\theta]$$

$$P_X(x) = \begin{cases} p & \text{si } x = 1 \\ q & \text{si } x = -1 \end{cases}$$

$$M(\theta) = E [\exp[\theta X]] = \sum_{x=-1}^1 \exp[\theta X] P(x) = \exp[-\theta] q + p \exp[\theta]$$

$$\exp[-\theta_0] q + p \exp[\theta_0] = 1$$

$$\theta_0 = \log (q/p)$$

$$P_a = \frac{1 - (q/p)^b}{(q/p)^a - (q/p)^b}$$

De acuerdo al lema II (nótese que $M(\theta)$, como una función de la variable compleja θ , es no singular en 0 o θ_0), la ecuación

$$- \log M(\theta) = i\phi$$

tiene dos raíces $\theta_1(\phi)$, $\theta_2(\phi)$ tales que $\theta_1(\phi) \rightarrow 0$ y $\theta_2(\phi) \rightarrow \theta_0$ cuando $\phi \rightarrow 0$. En consecuencia

$$P_a(\exp[a\theta_1(\phi) C_a(\phi)] + P_b(\exp[b\theta_1(\phi) C_b(\phi)] = 1$$

$$P_a(\exp[a\theta_2(\phi) C_a(\phi)] + P_b(\exp[b\theta_2(\phi) C_b(\phi)] = 1$$

donde $C_a(\phi)$ es la función característica de N condicional a $S_n = a$, así que

$$E(\exp[i\phi N]) = C(\phi) = P_a C_a(\phi) + P_b C_b(\phi)$$

Las ecuaciones

$$P_a = \frac{1 - \exp[b\theta]}{\{\exp[a\theta_0] - \exp[b\theta_0]\}}$$

$$P_a(\exp[a\theta_1(\phi) C_a(\phi)] + P_b(\exp[b\theta_1(\phi) C_b(\phi)] = 1$$

$$P_a(\exp[a\theta_2(\phi) C_a(\phi)] + P_b(\exp[b\theta_2(\phi) C_b(\phi)] = 1$$

$$E(\exp[i\phi N]) = C(\phi) = P_a C_a(\phi) + P_b C_b(\phi)$$

determinan la distribución completa de N , el número de etapas requeridas para alcanzar los límites.

La importancia de esta fórmula recae en su aplicabilidad aproximada aun cuando los límites se pasan, más que cuando son precisamente alcanzados. Como una aproximación se puede tratar a X como una variable normal, desde que se sabe que la distribución irrestricta de S_n tiende a la normalidad, y esto se aplicará también al resultado de varios pequeños pasos. En el límite cuando los pasos individuales se vuelven infinitesimalmente pequeños, esta fórmula se convierte exacta.

Ahora para X distribuida normalmente, la ecuación

$$-\log M(\theta) = i\phi$$

se convierte

$$M(\theta) = \exp[m\theta + (1/2) \sigma^2 \theta^2]$$

$$\log M(\theta) = -i\phi$$

$$\log \{\exp[m\theta + (1/2) \sigma^2 \theta^2]\} = -i\phi$$

$$m\theta + (1/2) \sigma^2 \theta^2 = -1\phi$$

$$\theta = \frac{-m + \sqrt{m^2 - 2\sigma^2 1\phi}}{\sigma^2}$$

$$\theta_1(\phi) = \frac{-m + \sqrt{m^2 - 2\sigma^2 1\phi}}{\sigma^2}$$

$$\theta_2(\phi) = \frac{-m - \sqrt{m^2 - 2\sigma^2 1\phi}}{\sigma^2}$$

$$\lim_{\phi \rightarrow \infty} \theta_1(\phi) = \lim_{\phi \rightarrow \infty} \frac{-m + \sqrt{m^2 - 2\sigma^2 1\phi}}{\sigma^2} = \frac{-m + m}{\sigma^2} = 0$$

$$\lim_{\phi \rightarrow \infty} \theta_2(\phi) = \lim_{\phi \rightarrow \infty} \frac{-m - \sqrt{m^2 - 2\sigma^2 1\phi}}{\sigma^2} = \frac{-m - m}{\sigma^2} = \frac{-2m}{\sigma^2} = \theta_0$$

Como $\frac{-2m}{\sigma^2} = \theta_0$, sustituyendo valores en

$$P_a = \frac{1 - \exp [b\theta]}{\{\exp [a\theta_0] - \exp [b\theta_0]\}}$$

se tiene que

$$P_a = \frac{1 - \exp [-2mb/\sigma^2]}{\{\exp [-2ma/\sigma^2] - \exp [-2mb/\sigma^2]\}}$$

La raíz apropiada en $\theta(\phi)$ ($m > 0$) es aquella con el signo +. Como $P_a \rightarrow 0$, se obtiene que

$$P_0 \{\exp [a\theta_1(\phi)] C_0(\phi) + P_b \{\exp [b\theta_1(\phi)] C_b(\phi) = 1$$

$$C(\phi) = \exp [-b\theta_1(\phi)]$$

$$\begin{aligned} \text{donde } \log C(\phi) &= -b\theta_1(\phi) = -b \left[\frac{-m + \sqrt{m^2 - 2\sigma^2 \phi}}{\sigma^2} \right] \\ &= \frac{bm}{\sigma^2} - \frac{bm}{\sigma^2} \left[1 - \frac{2\sigma^2 \phi}{m^2} \right]^{1/2} \end{aligned}$$

Expandiendo lo anterior en potencias de ϕ se obtienen los cumulantes de N :

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(c)}{2!} (x-c)^2 + \dots + \frac{f^{(n)}(c)}{n!} (x-c)^n + \dots$$

$$k_1 = \left. \frac{d \log C(\phi)}{d\phi} \right|_{\phi=0} = bm$$

$$k_2 = \left. \frac{d^2 \log C(\phi)}{d\phi^2} \right|_{\phi=0} = \frac{b\sigma^2}{m^3}$$

Regresando a la identidad de Wald, se obtendrán algunas fórmulas de momentos por derivación. Derivando una vez,

$$E \{ [M(\theta)]^{-N} \exp [\theta S_N] \} = 1$$

$$\frac{d}{d\theta} E \{ [M(\theta)]^{-N} \exp [\theta S_N] \} =$$

$$E \{ S_N [M(\theta)]^{-N} \exp [\theta S_N] - N [M(\theta)]^{-N-1} M'(\theta) \exp [\theta S_N] \} = 0$$

También se tiene que para $\theta = 0$

$$E \{ S_N [M(\theta)]^{-N} \exp [\theta S_N] - N [M(\theta)]^{-N-1} M'(\theta) \exp [\theta S_N] \} = 0$$

$$M'(0) = E [X] = m$$

$$\exp [\theta S_N] = 1$$

$$M(\theta) = 1$$

$$E\{S_N - N M'(\theta)\} = 0$$

$$E\{S_N\} - E\{N M'(\theta)\} = 0$$

$$E\{S_N\} = E\{N M'(\theta)\}$$

$$E\{S_N\} = E\{N\} E\{X\}$$

En las suposiciones aproximadas, esto da

$$E\{N\} = \frac{a P_a + b P_b}{m} \quad m \neq 0$$

Si $m=0$, se obtiene de una segunda derivación de la identidad de Wald

$$E\{S_N^2\} = E\{N\} E\{X^2\}$$

Ahora en $P_a = \frac{1 - \exp[b\theta]}{\exp[a\theta] - \exp[b\theta]}$ cuando $m \rightarrow 0$ y

$\theta_0 \rightarrow 0$, P_b tiene un límite definido $\frac{-a}{b-a}$

$$P_b = 1 - P_a = 1 - \frac{1 - \exp[b\theta]}{\exp[a\theta] - \exp[b\theta]} =$$

$$\frac{\exp[a\theta] - \exp[b\theta] - 1 + \exp[b\theta]}{\exp[a\theta] - \exp[b\theta]}$$

$$\lim_{\theta_0 \rightarrow 0} P_b = \lim_{\theta_0 \rightarrow 0} \frac{\exp[a\theta_0] - \exp[b\theta_0] - 1 + \exp[b\theta_0]}{\exp[a\theta_0] - \exp[b\theta_0]} = \frac{0}{0}$$

∴ se utiliza la regla de L'Hopital

$$\lim_{\theta_0 \rightarrow 0} P_b = \lim_{\theta_0 \rightarrow 0} \frac{a \exp [a\theta_0]}{a \exp [a\theta_0] - b \exp [b\theta_0]} = \frac{-a}{b-a}$$

$E\{S_N^2\}$ se convierte en $-ab$

$$\lim_{\theta_0 \rightarrow 0} P_a = \lim_{\theta_0 \rightarrow 0} \frac{1 - \exp [b\theta_0]}{\exp [a\theta_0] - \exp [b\theta_0]} = 0$$

∴ se utiliza la regla de L'Hopital

$$\lim_{\theta_0 \rightarrow 0} P_a = \lim_{\theta_0 \rightarrow 0} \frac{-b \exp [b\theta_0]}{a \exp [a\theta_0] - b \exp [b\theta_0]} = \frac{b}{b-a}$$

$$E\{S_N^2\} = a^2 P_a + b^2 P_b = a^2 \frac{b}{b-a} - b^2 \frac{a}{b-a} = \frac{a^2 b - b^2 a}{b-a}$$

$$= \frac{ab(a-b)}{b-a} = -ab$$

En consecuencia

$$E(N) = \frac{-ab}{\sigma^2} \quad m = 0$$

Continuando con el caso del riesgo en el seguro, se tomará el otro límite $a \rightarrow -\infty$. Se supone $V M'(0) - \mu < 0$ (de otra forma $P_b = 1$), y la relación de P_b es entonces

$$P_a E_a \{\exp [X(T)\theta_0]\} + P_b E_b \{\exp [X(T)\theta_0]\} = 1$$

donde $\theta_0 (> 0)$ es la raíz de la ecuación

$$\exp \{V(M(\theta) - 1) - \mu\theta\} = 1$$

Entonces

$$P_b = 1 - P_a = \frac{\exp [a\theta_0] - 1}{\exp [a\theta_0] - \exp [b\theta_0]}$$

$$\text{Cuando } \lim_{a \rightarrow -\infty} P_b = \frac{1}{\exp [b\theta_0]}$$

donde P_b es la probabilidad de ruina de la Cia. aseguradora.

C O N C L U S I O N

El número y el monto de reclamos dentro de una compañía de seguros es vital y muy importante para ella el tratar de llevar un control de los mismos, por que el número y el monto de reclamos representan los costos de una compañía de seguros y si existe un mayor número y monto de reclamos producirá que los costos de la compañía se eleven y si a esto se auna que existen pocas ganancias, provoca que la compañía sufra una pérdida de dinero, que se reflejarán en la merma de sus reservas iniciales y en consecuencia, si esto surge durante varios años provocará la ruina de la compañía de seguros.

El monto y número de reclamos dentro de una compañía es tan importante como el seguro lo debe de ser dentro de la economía, debido a que es una forma de estar alentando el ahorro y de esta forma poder fomentar tambien la inversión.

En México existe un claro subdesarrollo dentro de estos servicios y prefieren dar más auge a los servicios bancarios que a los servicios aseguradores, sin darse cuenta que en algunas ocasiones es mucho más importante la cuestión aseguradora que los servicios bancarios, y que si no se atiende en estos momentos, ante la inminente firma al Tratado de Libre Comercio con Estados Unidos y Canadá estará en plena desventaja este tipo de servicios al encontrarse en subdesarrollo y no darle una plena importancia como lo debe de ser.

El subdesarrollo de la industria aseguradora se da debido a muchas causas, como son la falta de apoyo del gobierno, el no tener personas capacitadas dentro del servicio asegurador, entre otras.

Una forma eficiente y eficaz de tratar de llevar un control sobre el número y monto de reclamos es por medio de las series de tiempo. Con los modelos de las series de tiempo se puede simular de una forma considerable como se comporta en la realidad el número y monto de reclamos y se pueda realizar de manera confiable pronósticos de la forma de cómo se comportan estos, ya que las series de tiempo presentan la forma de tratar las tendencias y ciclos, cuestiones que presentan en la realidad tanto el número como el monto de reclamos.

En el modelo tratado se puede observar una clara tendencia hacia la alta del número y monto de reclamos, situación que puede presentar una compañía de seguros durante los primeros años de sobrevivencia, pero se puede observar también que se pronostica una estabilidad para el número y monto de reclamos, cuestión que puede ser de alivio para la compañía porque puede ser un aviso de mejoría para estos, ya que si el número y monto de reclamos se reducen y el cobro por concepto de primas aumenta provocará un superávit, una ganancia durante ese año para la compañía, situación que se puede tener plena seguridad que sea así, si se tiene el mismo tamaño de cartera de clientes o que aumente un poco esta.

Otro punto importante de una compañía de seguros es el saber reasegurarse, el poder compartir una parte de su riesgo con otra compañía de seguros, y que con la parte que retenga se obtenga el mayor beneficio. El saber manejar el reaseguro es muy importante para la compañía por que, como se mencionó anteriormente, el seguro es cobrar primas, reinvertir estas y obtener un monto y número de reclamos mínimo, y mientras mayor dinero se tenga por concepto del pago de primas, mayores ganancias se obtendrán por concepto de inversión y si por concepto del pago de reclamos es mínimo se obtendrá una utilidad neta muy buena.

Pero al retener una parte considerable de una prima también implica el tener un mayor riesgo dentro de la compañía, ya que si sucede el siniestro tendrá que pagar una mayor parte del reclamo y si la compañía no alcanza a sufragar este gasto, puede llevarla a la quiebra. Por estas razones es muy importante el saber hasta que límite una compañía de seguros puede retener la parte del seguro y de esta forma obtener las mayores ganancias bajo un mínimo de riesgo.

Otro punto importante para una compañía de seguros es saber cual va a ser la política que ha de seguir para poder mantener una estabilidad adecuada y poder asegurar su sobrevivencia, es decir, cual va a ser su principio u objetivo principal que deberá seguir para poder permanecer en una situación sana. Uno de estos principios puede ser el de que la seguridad de una empresa es el objetivo más importante para una compañía de seguros, entonces se habla de que se debe de tratar con la probabilidad de ruina de una compañía de seguros.

La probabilidad de ruina es una política de estabilidad para una compañía de seguros, es el tratar de saber hasta que grado puede retener el mayor número de primas con la menor probabilidad de ruina al asumir el consiguiente riesgo

Bibliografía básica.

Pentikäinen T. et al (1984): "Risk Theory"; Third Edition; Chapman and Hall; New York.

Bühlman Hans (1970); "Mathematical Methods in Risk Theory"; Springer-Verlag; Berlin.

Harvey A.C. (1986); "Time Series Model"; Wiley; New York.

Feller William (1989); "Introducción a la Teoría de Probabilidades y sus aplicaciones"; Noriega Editores; México; Volumen I.

Feller William (1989); "Introducción a la Teoría de Probabilidades y sus aplicaciones"; Noriega Editores; México; Volumen II.

Parzen E. (1988); "Teoría Moderna de la Probabilidad y sus aplicaciones"; Wiley; New York.

Parzen E. (1962); "Stochastic Processes"; Holden-Day; New York.

Bartlett M.S. (1955); "An Introduction to Stochastic Processes"; Cambridge University Press; London.

Isaacson Dean L. et al (1976); "Markov Chains and Applications"; Wiley; new York.

Apuntes de la materia "Series de Tiempo" con el profesor Victor Repinoza. 1989.