



UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO

FACULTAD DE CIENCIAS

Ruido y azar en series de tiempo

**T E S I S**

QUE PARA OBTENER EL TÍTULO DE

M A T E M Á T I C O

P R E S E N T A :

PEDRO REYES PÉREZ

L



DIRECTOR DE TESIS: DR. MANUEL JESÚS FALCONI MAGAÑA

2005

m343630



FACULTAD DE CIENCIAS  
SECCION ESCOLAR



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

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

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



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MÉXICO

**ACT. MAURICIO AGUILAR GONZÁLEZ**  
**Jefe de la División de Estudios Profesionales de la**  
**Facultad de Ciencias**  
**Presente**

Comunicamos a usted que hemos revisado el trabajo escrito:

"Ruido y azar en series de tiempo"

realizado por Pedro Reyes Pérez

con número de cuenta 09653466-3 , quien cubrió los créditos de la carrera de: Matemáticas

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis  
Propietario Dr. Manuel Jesús Falconi Magaña

Propietario Dr. Pedro Eduardo Miramontes Vidal

Propietario Dr. Luis Antonio Rincón Solís

Suplente M. en C. Arturo Hugo Nieva Gochicoa

Suplente M. en C. Jefferson Edwin King Dávalos

Consejo Departamental de Matemáticas

M. en C. Alejandro Bravo Mojica

AGUILAR GONZÁLEZ  
CONSEJO DEPARTAMENTAL  
DE  
MATEMÁTICAS

## **Agradecimientos**

Agradezco en forma muy particular al Dr. Manuel Jesús Falconi Magaña, por haberme hecho participe de sus conocimientos, por su asesoría, su apoyo incondicional y por la dirección de este trabajo, así como por su paciencia y confianza.

A mis sinodales, Dr. Pedro Eduardo Miramontes Vidal, Dr. Luís Antonio Rincón Solís, M. en C. Arturo Hugo Nieva Gochicoa y al M. en C. Jefferson Edwin King Dávalos; gracias por el apoyo que cada uno de ustedes me brindó, por sus comentarios y sugerencias que aportaron para mejorar la realización de éste trabajo.

A la Universidad Nacional Autónoma de México que nos permite introducirnos a un mundo de conocimientos que nos conducen a concluir sueños, metas y ser mejores individuos.

A mis compañeros que sin conocernos han caído en los movimientos en contra de la privatización de la educación y que de que lejos de ser egoístas han luchado por que la gratuidad de la Universidad se mantenga, para que individuos como yo alcancemos el sueño de ser profesionistas.

## Dedicatoria

A mis padres, hermanos y familiares que cada uno de forma diferente me brindaron apoyo para seguir adelante, doy gracias en particular a J6se G., a Brenda Berenice.

A Rosa S., Porque gracias a su insistencia logre conocer un mundo diferente.

A Antonio P., Adri6n S., Miriam T. Maritza P., Santiago L., Vanesa P., Juan M. Agradezco por haberme acompa1ado por los diferentes laberintos del conocimiento y de la vida.

A Srta. Hermila Berm6des Rojas del Dep. De matem6ticas le dedico en especial 6ste trabajo, porque siempre se mostr6 muy interesada en los tr6mites que necesité.

A la Srta. Claudia Rojas por ayudarme a concluir todos los tramites para que 6ste trabajo fuera presentado. L

A todos aquellos compa1eros que por diferentes motivos no lograron la consolidaci6n de sus metas.

*No olvidemos que la vida  
Puede ser tan peque1a  
O tan grande como cada  
uno de nosotros queramos.*

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Introducción . . . . .	1
<b>2. Análisis de índices de dispersión</b>	<b>3</b>
2.1. Introducción . . . . .	3
2.1.1. Índice de homogeneidad . . . . .	4
2.1.2. Análisis del índice IDH . . . . .	4
2.1.3. El índice $\delta$ . . . . .	6
2.1.4. Experimento 1 . . . . .	8
2.1.5. Experimento 2 . . . . .	9
2.1.6. El índice IZU . . . . .	11
2.1.7. Experimento 3 . . . . .	11
<b>3. La no linealidad en una serie de tiempo</b>	<b>13</b>
3.1. Introducción . . . . .	13
3.2. Mapeos uni-dimensionales . . . . .	14
3.2.1. Propiedades en ausencia de ruido . . . . .	14
3.2.2. Análisis gráfico de un mapeo uni-dimensional . . . . .	17
3.2.3. La no linealidad a través de la serie de tiempo . . . . .	18
3.2.4. Un vistazo desde la estadística . . . . .	19
3.2.5. Dimensión de correlación . . . . .	24
<b>4. Programas de simulación</b>	<b>29</b>
4.1. Introducción . . . . .	29
4.2. Índice IDH . . . . .	29
4.3. La proporción de ceros en una sucesión . . . . .	32
4.4. Índice IDH vs índice IZU . . . . .	33
4.5. El mapeo de Ricker . . . . .	34
4.6. Mapeo de Ricker con ruido . . . . .	36
4.7. Histogramas . . . . .	37
4.8. Gráficos de dispersión . . . . .	38

4.9. Aproximaciones polinomiales . . . . .	39
4.10. Dimensión de correlación . . . . .	44
4.11. Dimensión de correlación con ruido . . . . .	47
<b>A.</b>	<b>49</b>
A.1. Glosario . . . . .	49

# Capítulo 1

## Introducción

### 1.1. Introducción

Muchos fenómenos son estudiados usando sucesiones de medidas u observaciones hechas a través del tiempo. Estas series de observaciones, a menudo resumen una parte importante de la información disponible sobre un sistema bajo investigación. Los ejemplos de series de tiempo incluyen registros de temperatura diaria de una región, tasas de retorno en finanzas, ritmos cardíacos, crecimiento poblacional, etcétera. Las series de tiempo observadas a menudo exhiben un comportamiento irregular y aparentemente impredecible. La meta principal de la modelación de series de tiempo es obtener los hechos fundamentales a partir de las irregularidades mostradas e incrementar la comprensión de los procesos que las generan. La modelación clásica se basa en los modelos lineales gaussianos que generan series irregulares a partir de un “ruido aleatorio ambiental”. En contraste, los sistemas dinámicos determinísticos llamados caóticos pueden generar series de apariencia complicada que no dependen de una fuente azarosa. Estos hechos han originado el problema (difícil), para determinar si una serie de tiempo aparentemente aleatoria es realmente azarosa o admite una explicación en términos de un mecanismo determinístico.

Sin lugar a dudas un problema importante tanto desde el punto de vista teórico como práctico, es la determinación del grado de aleatoriedad de una serie de datos experimentales. En el presente trabajo se trata este asunto a través del análisis de los siguientes problemas: 1) Análisis de algunos índices utilizados para determinar la distribución de ceros de una sucesión binaria. 2) Estudio experimental para ver si el ruido aleatorio destruye la no-linealidad presente en una serie de tiempo, en las órbitas de un sistema dinámico.

Respecto al problema (1), en la primera parte de este trabajo se analizarán

algunas de las herramientas que se han creado para medir el grado de aleatoriedad o dispersión de los ceros en una sucesión binaria. Una de las herramientas diseñadas con este fin es el índice  $IDH$ , definida por [mir], para estudiar las variaciones locales de sucesiones del ADN. En [cow] se introduce el índice  $\delta$  para generar sucesiones de ceros y unos con diversos grados de dispersión. Estas sucesiones les permitieron simular la alternancia de dos mapeos  $F_a$  y  $F_b$ . Si aparecía un cero en la sucesión aplicaban el mapeo  $F_a$ ; en el otro caso aplicaban el mapeo  $F_b$ . Posteriormente se estudia el índice  $IZU$ , que igual que el  $IDH$  mide la dispersión de ceros a lo largo de la sucesión. Concluiremos que los tres índices  $IDH$ ,  $IZU$  y  $\delta$  miden exactamente lo mismo y que son parámetros que ayudan a distinguir diferentes grados de aleatoriedad en una sucesión binaria.

En torno al problema (2) se planteó la cuestión de si el ruido puede ocultar la no-linealidad de una serie generada por un sistema determinístico. Para esto se analizaron las órbitas del mapeo de Ricker, distorsionadas con diversos grados de ruido “multiplicativo”. En términos más precisos, dado un valor del parámetro del mapeo de Ricker, se observó cómo son las órbitas generadas con diversos grados de ruido. Se encontró, tanto gráficamente, como con la dimensión de correlación, que la no-linealidad era más fácilmente destruida conforme la periodicidad de la órbita atractora era más pequeña. En cambio, cuando el mapeo es caótico, la no-linealidad se preserva aún para perturbaciones aleatorias bastante grandes. Los detalles de este análisis conforman el capítulo 2.

La cuarta parte del trabajo, contiene los programas de simulación que fueron hechos en *Mathematica*, para las realizar los experimentos de los capítulos 2 y 3.

# Capítulo 2

## Análisis de índices de dispersión

### 2.1. Introducción

En este capítulo se analizará estadísticamente el índice *IDH* que P. Miramontes y colaboradores introdujeron en el trabajo *Structural and thermodynamic properties of DNA uncover different evolutionary histories*, para estudiar el grado de aleatoriedad de una secuencia de ADN (ver [mir]).

Este índice lo utilizaron para cuantificar las variaciones locales dentro de una serie de *ADN*. El índice *IDH* trata de medir la homogeneidad del *ADN* basado en un modelo de distribución binaria, que cuantifica los agregados termodinámicos y estructurales presentes en la arquitectura primaria del *ADN*. Un análisis cuidadoso de sucesiones de ADN con el *IDH* permite descubrir algunas restricciones significativas del *ADN*, diferentes de las obtenidas por medio del uso del codón o de la función proteína. Este índice también sirve para distinguir los diferentes orígenes evolutivos de los organismos y los coloca en dominios ajenos en el espacio de las secuencias de *ADN*.

La estructura de doble hélice del *ADN* fué un gran descubrimiento de la biología estructural moderna. Sin embargo, esta conformación idealizada es tan sólo un promedio de un conjunto grande de variaciones locales. De manera muy simplificada, en una cadena de *ADN* se reconoce que un par CG (citosina-guanina) es un enlace fuerte (F) y el par AT (adenina-timina) es un enlace débil (D). Otro aspecto a distinguir en una secuencia de *ADN* es la distribución de purina (R) y de pirimidina (Y), debido a sus diferentes efectos sobre los fosfatos adyacentes. Finalmente, la adenina y citosina (bases M) modifican la estructura de un modo distinto a la guanina y timina (bases K).

La naturaleza binaria de los “causantes” de la variación estructural del

*ADN* ( i.e., *F* vs. *D*; *Y* vs. *R* y *M* vs. *K* ), sugirió la formulación de un índice que expresara lo inhomogeneo de las sucesiones de *ADN* como función de elementos binarios. Este índice llamado *Índice de homogeneidad del ADN (IDH)*, fué definido en [mir].

### 2.1.1. Índice de homogeneidad

Las dicotomías *YR*, *MK* y *FD* pueden analizarse simultáneamente usando el mismo tratamiento matemático, si la sucesión de nucleótidos original se convierte en una sucesión binaria. Por ejemplo, se le puede asignar el valor cero a una base *Y* y el 1 a una base *R*. En forma análoga, se pueden hacer las conversiones para los casos *MK* y *FD*. Para cuantificar las variaciones locales dentro de una serie de *ADN*, es necesario encontrar el número de ceros y unos dentro de su representación binaria; así como la manera en que están organizados a lo largo de la sucesión. Para ese propósito se define el índice *IDH* del modo siguiente,

$$d = \frac{N_{00}N_{11} - N_{01}N_{10}}{N_0N_1}, \quad (2.1.1)$$

donde  $N_{ij}$  es el número de dinucleótidos  $ij$  ( $ij = 00, 10, 01, 11$ ). El número de ceros y unos en la secuencia se denotan con  $N_0$  y  $N_1$ , respectivamente. Notemos que  $d$  varía entre -1 y 1.

Una sucesión aleatoria de ceros y unos, en promedio mostraría el mismo número de ceros, unos y de cadenas de la forma 00 y 11, por lo que uno espera que en este caso el valor del índice esté alrededor del cero. Si la sucesión muestra cadenas grandes de ceros y de unos, el índice estará cercano a uno y si hay una gran alternancia de ceros y unos, el valor sería cercano a -1.

### 2.1.2. Análisis del índice IDH

En esta sección se hará un análisis del índice  $d$ . Nos interesa estudiar el índice como función del número  $m$  de ceros en la sucesión y de la probabilidad  $P(0)$  de obtener un cero. El número  $m$  es en realidad el valor esperado del número de ceros que aparecen en una sucesión. El propósito es determinar si el índice puede distinguir el grado de dispersión de los ceros, independientemente de  $m$ . Para controlar el grado de dispersión de los ceros en la sucesión, utilizaremos un índice  $\delta$ , que llamamos *índice de agregación*. Valores de  $\delta$  cercanos a -1 producen sucesiones muy alternantes y conforme  $\delta$  se acerca a 1, las sucesiones tienden a mostrar grandes bloques de ceros y de unos. Con ayuda de *Mathematica* generaremos sucesiones de ceros y unos.

Iniciaremos el análisis buscando el tipo de la sucesión que genera el máximo valor del índice  $d$ . Sea  $K$  una sucesión aleatoria de tamaño  $N$ , compuesta de  $m$  ceros y  $n$  unos. Entonces  $N = m + n$ . Se tiene el siguiente resultado.

**Lema 1** *El índice de cualquier sucesión con  $m$  ceros y  $n$  unos es menor o igual a*

$$d^* = \frac{(m-1)(n-1)}{mn}. \quad (2.1.2)$$

La sucesión  $K = \{0, \dots, 0, 1, \dots, 1\}$  es la única sucesión con índice  $d^*$ .

*Demostración.* Sea  $K = \{0, \dots, 0, 1, \dots, 1\}$  con  $m$  ceros y  $n$  unos. Analizaremos primero el caso en que  $n = m$ . Entonces, tenemos que  $N_{10} = 0$ ,  $N_{01} = 1$ ,  $N_{00} = m - 1$  y  $N_{11} = m - 1$  y por lo tanto

$$d^* = \frac{(m-1)^2}{m^2}. \quad (2.1.3)$$

Ahora se demostrará que el índice de cualquier otra sucesión con  $m$  ceros y  $N = 2m$  es menor que  $d^*$ . Sea  $K'$  una sucesión binaria de la forma  $\{1, \dots, 1, 0, \dots, 0, 1, \dots, 1\}$ , de tamaño  $N = 2m$  con  $m$  ceros. Entonces  $N_{10}(K') = 1 = N_{01}(K')$ ,  $N_{11}(K') = m - 2$  y  $N_{00}(K') = m - 1$ , por lo tanto

$$d(K') = \frac{(m-1)(m-2) - 1}{m^2}. \quad (2.1.4)$$

De aquí se obtiene que  $d(K') < d^*$ .

Si  $m = 1$ , las únicas sucesiones son  $\{0, 1\}$  y  $\{1, 0\}$ . Ambas tienen índice cero.

En general, si los ceros no forman un bloque conexo como en la sucesión  $K'$ , habrá un número mayor que cero de parejas 01 y 10. Además, el número de parejas 00 y 11 serán menor a  $(m-1)$ , por lo que el índice  $d = \frac{N_{00}N_{11} - N_{01}N_{10}}{m^2}$  de éste tipo de sucesiones es menor que  $d^* = \frac{(m-1)^2}{m^2}$ .

El caso  $m \neq n$  es completamente análogo al anterior.

□

Antes de calcular el valor mínimo del índice daremos una definición.

**Definición 1** *Una sucesión con  $m$  ceros y  $n$  unos es alternante si*

1. No tiene cadenas de 2 o más ceros, cuando  $m \leq n$ , o
2. No tiene cadenas de 2 o más unos, cuando  $m > n$ .

**Lema 2** *El índice  $d$  de cualquier sucesión con  $m$  ceros y  $n$  unos ( $m \neq n$ ), es mayor o igual que:*

$$d_* = -\frac{(\min\{m, n\})^2}{mn}. \quad (2.1.5)$$

*El índice  $d_*$  se alcanza en una sucesión alternante cuyos extremos no son ceros, en el caso en que  $m < n$  o que no son unos si  $m > n$ .*

**Demostración.** Supongamos que  $m < n$ , (el caso  $m > n$  se demuestra en forma análoga). Es fácil ver que el índice de una sucesión alternante cuyos extremos no son ceros es igual a  $-mm/mn = -m/n$  y entonces se cumple la ecuación 2.1.5. Si la sucesión es alternante y  $m=n$ , entonces el número de parejas 01 es  $m-1$  y el de 10 es  $m$ , o el número de parejas 01 es  $m$  y el de 10 es  $m-1$ , según sea cero el extremo derecho o el izquierdo, respectivamente. Por lo tanto el índice de estas sucesiones es  $-m(m-1)/mn = -(m-1)/n$ , que obviamente es mayor que  $d_*$ . Finalmente, observamos que si la sucesión no es alternante, entonces tiene cadenas de ceros o cadenas de unos, que hacen que  $N_{01}$  y  $N_{10}$  disminuyan y además  $N_{00}$  o  $N_{11}$  pueden ser diferentes de cero; por lo tanto su índice es mayor que  $d_*$ .

□

Ahora nos enfocaremos a desarrollar un estudio más completo sobre las sucesiones, donde podremos hasta cierto punto manipular los valores de los índices. Esto forzosamente dependerá de la forma en que podamos manejar la colocación de los ceros y unos dentro de la sucesión.

### 2.1.3. El índice $\delta$

Para controlar la distribución de los ceros introducimos, el índice de agregación,  $\delta$ , donde  $\delta = 1 - \frac{P(0|1)}{P(0)}$ , ver [cow],  $P(0)$  es la probabilidad de obtener un cero y  $P(0 | 1)$ , la probabilidad condicional de obtener cero dado que el anterior fué un uno.

Notemos que se satisfacen las siguientes igualdades:

**Lema 3**

$$P(0)P(1 | 0) = P(1)P(0 | 1), \quad (2.1.6)$$

$$P(0) + P(1) = 1, \quad (2.1.7)$$

$$P(0 | 1) + P(1 | 1) = 1, \quad (2.1.8)$$

$$P(1 | 0) + P(0 | 0) = 1. \quad (2.1.9)$$

**Demostración.** Las pruebas de las ecuaciones 2.1.7, 2.1.8 y 2.1.9 no se realizarán ya que son directas. En seguida se prueba la ecuación 2.1.6.

Sea  $q$  la probabilidad de obtener cero y  $1 - q$  la probabilidad de obtener uno,  $s = P(1 | 0)$ ,  $1 - s = P(0 | 0)$ ;  $r = P(0 | 1)$ ,  $1 - r = P(1 | 1)$ . Entonces  $P(1) = qs + (1 - q)(1 - r)$  pero  $P(1) = 1 - q$  entonces  $1 - q = qs + (1 - q)(1 - r)$  factorizando  $(1 - q)$  se tiene  $qs = (1 - q)(1 - (1 - r)) = (1 - q)r$ . Por lo tanto  $(1 - q)r = qs$ . Es decir  $P(1)P(0 | 1) = P(0)P(1 | 0)$

□

De acuerdo con el lema anterior podemos escribir las probabilidades condicionales como funciones de  $P(0)$  y  $\delta$  de la siguiente manera:

**Lema 4**

$$P(0 | 0) = \delta + P(0)(1 - \delta), \quad (2.1.10)$$

$$P(1 | 1) = 1 - P(0)(1 - \delta), \quad (2.1.11)$$

$$P(1 | 0) = [1 - P(0)](1 - \delta), \quad (2.1.12)$$

$$P(0 | 1) = P(0)(1 - \delta). \quad (2.1.13)$$

En seguida se demostrará la ecuación 2.1.10. Sea

$$\begin{aligned} P(0 | 0) &= 1 - P(1 | 0) \\ &= 1 - \frac{P(1)P(0 | 1)}{P(0)} \\ &= 1 - P(1)(1 - \delta) \\ &= P(0) + P(1)\delta \\ &= P(0) + (1 - P(0))\delta \\ &= P(0) + \delta - P(0)\delta \\ P(0 | 0) &= \delta + P(0)(1 - \delta) \end{aligned}$$

Ahora procederemos a la demostración de 2.1.11. Se tiene que:

$$\begin{aligned} P(1 | 1) &= 1 - P(0 | 1) \\ &= 1 - P(0)(1 - \delta). \end{aligned}$$

La demostración de 2.1.12 se obtiene directamente de 2.1.9 y 2.1.10. Para la prueba de 2.1.13, empleamos 2.1.8 y 2.1.11. Los detalles son simples y no se darán.

□

Se debe tomar en cuenta que una vez fijos  $P(0)$  y  $P(1)$  se pueden obtener diferentes grados de aleatoriedad variando  $\delta$ .

Notemos que el parámetro  $\delta$  varía en el intervalo  $[-1, 1]$  y la probabilidad  $P(0)$  está en un intervalo  $I_\delta$  y

**Lema 5**  $P(0)$  está en el intervalo  $I_\delta$ , donde

$$I_\delta = \begin{cases} [0, 1], & \text{si } \delta \in [0, 1], \\ [\frac{\delta}{\delta-1}, \frac{1}{1-\delta}], & \text{si } \delta \in [-1, 0]. \end{cases} \quad (2.1.14)$$

**Demostración.** Caso  $i)$  si  $\leq \delta \leq 0$  de la ecuación 2.1.10 se tiene  $0 \leq P(0 | 0) = \delta + P(0)(1 - \delta)$ . Entonces  $0 \leq \delta + P(0)(1 - \delta)$  y  $-\frac{\delta}{1-\delta} \leq P(0)$  o lo que es lo mismo  $\frac{\delta}{\delta-1} \leq P(0)$ . Por otro lado de la ecuación 2.1.11 se sigue  $0 \leq P(1 | 1) = 1 - P(0)(1 - \delta)$ ; así,  $P(0)(1 - \delta) \leq 1$  y despejando a  $P(0)$  se llega a que  $P(0) \leq \frac{1}{1-\delta}$ . Por lo tanto  $\frac{\delta}{\delta-1} \leq P(0) \leq \frac{1}{1-\delta}$ . Para terminar la demostración de este caso, notemos que  $\frac{\delta}{\delta-1} \leq 0$  y  $\frac{1}{1-\delta} \geq 1$  para toda  $\delta \in [0, 1]$ .

$P(1 | 1) = 1 - P(0)$ ,  $P(1 | 0) = 1 - P(0)$  y  $P(0 | 1) = P(0)$ . De la ecuación 2.1.10 se tiene la siguiente desigualdad  $0 \leq P(1/1) = 1 - P(0)(1 - \delta)$ , se sigue que  $P(0) \leq 1/(1 - \delta)$ . Por otro lado de la ecuación 2.1.12 se tiene que  $0 \leq P(1/0) = (1 - P(0)(1 - \delta)) \leq 1$ , entonces  $0 \leq 1 - P(0) \leq \frac{1}{1-\delta}$  lo cual implica  $1 - \frac{1}{1-\delta} \leq P(0)$  simplificando llegamos a  $-\frac{\delta}{1-\delta} \leq P(0)$ , por lo tanto  $\frac{\delta}{\delta-1} \leq P(0) \leq \frac{1}{1-\delta}$ . Finalmente, observemos que  $\frac{1}{1-\delta} < 1$  si  $\delta < 0$  y  $\frac{\delta}{\delta-1} > 0$  si  $\delta < 0$ .  $\square$

Existen casos especiales que se mencionan a continuación:

1) Cuando  $\delta \rightarrow 1$  las ecuaciones 2.1.10 y 2.1.11 también tienden a 1, es decir  $P(0 | 0)$  y  $P(1 | 1)$  tienden a 1 y  $P(1 | 0)$  y  $P(0 | 1)$  tienden a cero; esto ocasiona que las sucesiones tiendan a presentar grandes bloques de ceros y grandes bloques de unos.

2) El caso  $\delta = 0$  se describe una situación totalmente independiente, ya que el estado en el siguiente instante no depende del estado actual. Esto se debe a que  $P(0 | 0) = P(0 | 1) = P(0)$  y  $P(1 | 1) = P(1 | 0) = 1 - P(0)$

3) Si  $\delta \rightarrow -1$  se tienen sucesiones alternadas ya que  $P(0/0)$  y  $P(1/1)$  tienden a 1, mientras que  $P(1/0)$  y  $P(0/1)$  tienden a 0.

### 2.1.4. Experimento 1

Se tratará de determinar si existe una relación directa entre  $P(0)$  y el promedio de ceros que aparecen en una sucesión. Denotaremos a  $P(0)$  con  $P_o$ . Para realizar el estudio, fijamos  $\delta$ , y se toman probabilidades  $P_{o_j} = j/20$ , y  $j = 0, \dots, 20$ . Para cada  $P_{o_j}$ , se generaron 300 sucesiones de tamaño 100 y se calculó la proporción de ceros en cada una de ellas. Finalmente, se obtuvo el promedio  $q_{o_j}$  de estas proporciones. Con estos datos se realizó la regresión lineal de la forma:  $P_{o_j} = \beta q_{o_j} + \alpha$ . Lo anterior se hace para cada  $\delta \in \{-.9, -.8, \dots, .8, .9\}$ . Los resultados obtenidos se presentan en la tabla 2.1.

$\delta$	$\alpha$	$\beta$	$r$	$r^2$	E. St.	Exp.	No Exp.	F
-0,9	0.089	0.995	0.995	0.990	0.011	0.961	0.00956	1708.47
-0,8	0.075362	0.997	0.99703	0.99407	0.00904	1.024562	0.006103	2854.16
-0,7	0.061987	0.998	0.99843	0.99687	0.00678	1.093308	0.003431	5416.45
-0,6	0.046875	0.999	0.99922	0.99844	0.00494	1.17016	0.001819	10933.8
-0,5	0.034753	1.00	0.99953	0.99906	0.00824	1.22949	0.001156	18074.4
-0,4	0.022309	1.00	0.99981	0.99962	0.00257	1.30055	0.000494	44792.9
-0,3	0.010574	1.00	0.99986	0.99972	0.00222	1.36104	0.000369	62779.7
-0,2	0.002791	1.00	0.99993	0.99987	0.001547	1.40453	0.000178	133831.0
-0,1	0.001393	1.00	0.99996	0.99992	0.001196	1.42008	0.000107	266525.8
0,0	0.0009192	1.00	0.99993	0.99987	0.001523	1.41607	0.000173	139137.2
0,1	0.001487	1.00	0.99993	0.99987	0.001562	1.42452	0.000182	133054.6
0,2	0.000462	1.00	0.99993	0.99987	0.001576	1.42743	0.000185	130952.4
0,3	-0.000242	1.00	0.99984	0.99968	0.002473	1.43099	0.000456	53327.6
0,4	-0.00033	1.00	0.99980	0.99961	0.002704	1.41815	0.000545	44227.43
0,5	-0.001955	1.00	0.99982	0.99965	0.002586	1.42653	0.000499	48620.81
0,6	-0.001767	1.00	0.99980	0.99961	0.002729	1.43507	0.000555	43937.75
0,7	0.002249	1.00	0.99975	0.99951	0.003033	1.42808	0.000686	35383.92
0,8	0.004624	1.00	0.99969	0.99939	0.003375	1.41094	0.000849	28246.71
0,9	-0.000460	0.99	0.99894	0.99788	0.006350	1.42371	0.003013	8031.117
0,95	-0.007399	0.99	0.99856	0.99714	0.007518	1.469558	0.004213	5929.962

Tabla 2.1

Para cada regresión se presentan los valores de  $r$  muy cercanos a uno, con esto se asegura que existe una relación directa entre el promedio de ceros que se presentan en cada sucesión y  $P(0)$ ; las  $\beta_i$  de todas las regresiones son muy cercanos o iguales a uno; además la suma de los errores de cada una es prácticamente cero. Así pues, podemos concluir que la proporción de ceros en una sucesión es estadísticamente igual a la probabilidad  $P(0)$ .

### 2.1.5. Experimento 2

En este estudio se tratará de averiguar si el índice IDH tiene alguna relación con el índice  $\delta$ . Para realizar este análisis se fijó  $\delta$  y se permitió que  $P(0)$  tomara diferentes valores en el intervalo  $(0, 1)$ . Para cada  $P(0)$  se crearon 100 sucesiones de tamaño 300, se calcularon los índices  $IDH$  de cada sucesión y se tomó el promedio  $\overline{IDH}$ . Lo anterior se hizo para  $\delta = -0.9, -0.8, \dots, 0.0, 0.1, \dots, 0.9$  y se encontró que  $\overline{IDH}$  coincidió con  $\delta$ , cuando  $\delta$  es positivo (ver figura 2.1). Sin embargo, esto no sucede si  $\delta < 0$ . En este caso, la igualdad que se presenta entre  $\delta$  y el índice  $IDH$  vale para  $P(0)$  en un intervalo  $\overline{I}_\delta$  alrededor de  $1/2$ , que disminuye conforme  $\delta$  tiende a  $-1$ . De hecho, la longitud que experimentalmente se encuentra de  $\overline{I}_\delta$  se ajusta bien a la longitud de  $I_\delta$  calculada en el lema 5. Esto se observa en la parte inferior de la figura 2.1.

Por los resultados que se han obtenido de las regresiones, podemos asegurar estadísticamente que los valores de los índices dependen solamente de la distribución de los ceros en la sucesión y no de su número.

Ahora compararemos el índice IDH con un índice que hemos llamado IZU y que también se ha utilizado para determinar la aleatoriedad de una sucesión de ceros y unos.

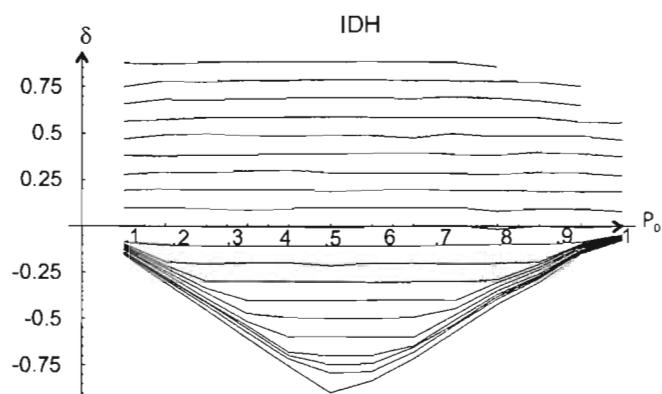


Figura 2.1: Se observa que los índices IDH y  $\delta$  son prácticamente iguales cuando  $0 \leq \delta$ , pero conforme  $\delta$  tiende a  $-1$ , la igualdad se da solo en un intervalo decreciente  $I_\delta$ , con centro en  $P_0=0.5$ . Ver lema 5.

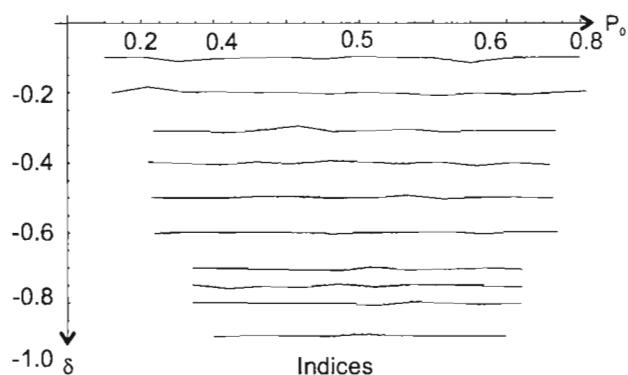


Figura 2.2: Siempre se puede encontrar un intervalo con centro  $P(0)=0.5$ , en el que los índices  $IDH$  y  $\delta$  son iguales, cuando  $\delta \rightarrow -1$ .

### 2.1.6. El índice IZU

En seguida se presenta el índice IZU, que de igual forma que el anterior, mide la dispersión de ceros y unos en una sucesión. Se tratará de encontrar si estadísticamente existe una relación entre las dos fórmulas. Tal índice es:

$$Z = \frac{N_{00} + N_{11}}{N_0 N_1} \tag{2.1.15}$$

### 2.1.7. Experimento 3

Para realizar este estudio se fijó  $P(0)$  y se permitió que  $\delta$  tomara los valores de  $-0.9, -0.8, \dots, 0.8, 0.9$ . Para cada pareja  $(P(0), \delta)$  se generaron 100 sucesiones de tamaño 300 ( $N = 300$ ), se calcularon los índices  $IDH$  e  $Izu$  y se sacó el valor promedio de  $d$  y  $Z$ . De este modo obtuvieron 19 parejas  $(d, Z)$ , para cada valor de  $P(0) = 0.1, 0.2, \dots, 0.9$ . Por último se realizaron regresiones lineales de la forma  $Z = \beta d + \alpha$ . Los resultados se presentan en la tabla 2.2

$P(0)$	$\alpha$	$\beta$	$r$	$r^2$	E. St.	Exp. Reg.	valor P	F
0.09	0.835	0.999	0.999	0.998	0.0008	0.021856	$p < 0,00$	7159,875
0.19	0.689	1.00	0.999	0.999	0.0012	1.136028	$p < 0,00$	15481,86
0.33	0.553	0.998	0.997	0.995	0.0037	0.525616	$p < 0,00$	2942,884
0.40	0.516	0.999	0.998	0.997	0.0028	0.906741	$p < 0,00$	6916,683
0.49	0.496	0.999	0.999	0.998	0.0028	1.371135	$p < 0,00$	8881,256
0.55	0.504	1.00	0.999	0.999	0,0011	1.181598	$p < 0,00$	52879,41
0.60	0.525	0.999	0.999	0.998	0,0201	0.771477	$p < 0,00$	13005,62
0.64	0.544	0.999	0.999	0.998	0.0024	0.569675	$p < 0,00$	7369,259
0.74	0.611	0.999	0.998	0.997	0.0029	0.265934	$p < 0,00$	3685,141
0.78	0.659	0.999	0.999	0.998	0.0020	0.162125	$p < 0,00$	6621,748
0.85	0.734	0.991	0.990	0.981	0.0056	0.098194	$p < 0,00$	522,9672
0.9	0.816	0.998	0.998	0.996	0.0018	0.026924	$p < 0,00$	2036,637

Tabla 2.2

Como se observa en cada una de las regresiones el coeficiente de correlación  $r^2$ , es prácticamente uno, lo mismo sucede con  $\beta$ . El error normal que se presenta en cada una de las regresiones es muy cercano a cero. En resumen, por los resultados que se muestran en la tabla 2.2, se asegura que estadísticamente los índices  $IDH$  e  $IZU$  calculan exactamente lo mismo. Es decir si se requiere calcular índices se puede utilizar cualquiera de las fórmulas.

### Conclusiones

De acuerdo con los resultados obtenidos en el experimento 1, podemos afirmar estadísticamente que  $P(0)$  es el parámetro que determina el número de ceros que aparecerán en cada sucesión, mientras que el índice  $\delta$  es el que determina la forma en que se distribuyen los ceros a lo largo de la sucesión, como se concluye del experimento 2. Por último se observó que los índices  $IDH$ ,  $\delta$  e  $IZU$  miden exactamente lo mismo. Es decir los tres índices miden la dispersión de ceros y unos que aparecen en una sucesión binaria; la única diferencia es que  $IZU$  varí entre 0 y 1.

# Capítulo 3

## La no linealidad en una serie de tiempo

### 3.1. Introducción

Un problema general en la ciencia, es deducir el mecanismo dinámico de un sistema a partir de una serie de tiempo observada. El problema se enuncia en términos sencillos: las irregularidades mostradas por los datos, son debidas al ruido aleatorio o a una dinámica caótica determinística?. No siempre es posible distinguir a simple vista en una serie de datos, la componente determinística de su contraparte estocástica. En los últimos años se ha desarrollado una gran cantidad de herramientas para estudiar este problema.

El propósito de éste trabajo no es analizarlas, sino generar una serie de evidencias experimentales que fortalezcan la intuición y facilite la búsqueda de métodos analíticos adecuados. Para esto, se analizarán los efectos de las perturbaciones aleatorias sobre la dinámica de mapeos unimensionales útiles en la ecología de poblaciones.

Las poblaciones en la naturaleza a menudo fluctúan erráticamente. Los cambios climáticos que las afectan se dan de manera aparentemente aleatoria; la contaminación por ejemplo, son factores que influyen también de modo impredecible. De tal forma que parece útil estudiar el efecto de fluctuaciones aleatorias sobre sistemas dinámicos determinísticos. Puesto de otro modo, las variaciones que uno observa en una población tienen base determinista o reflejan factores estocásticos?. Se ha sugerido (ver [sha]) que atractores extraños gobiernan seguramente el movimiento del mundo real en ecología, así como en los sistemas epidemiológicos. La existencia de un atractor extraño implica que las variaciones observadas en el fenómeno, aunque irregulares, son gene-

radas por un sistema determinístico. Al mismo tiempo se debe reconocer que en los sistemas ecológicos existe ruido. Ya sea por el cambio en las condiciones climáticas, o por la misma intervención humana. Esto debemos tomarlo en cuenta al estudiar los efectos en las fluctuaciones aleatorias sobre los sistemas dinámicos. En particular se desea conocer si las propiedades no lineales persisten o no, en un ambiente ruidoso. Dicho de otro modo, la pregunta es: ¿hasta que punto, un simple observador podría razonablemente deducir la existencia y el carácter cualitativo del elemento determinístico subyacente en la serie de mediciones temporales (serie de tiempo)? Se intentará dar una respuesta preliminar con base en el análisis de algunos modelos unidimensionales en ecología. Concretamente se analizará el llamado modelo de Ricker, (ver [sha]).

## 3.2. Mapeos uni-dimensionales

### 3.2.1. Propiedades en ausencia de ruido

Es importante notar que ecuaciones uni-dimensionales de la forma

$$x_{i+1} = T(x_i, \alpha) , \quad (3.2.1)$$

para ciertos valores del parámetro  $\alpha$  pueden exhibir dinámica complicada, incluyendo caos, cuando  $T$  es una aplicación unimodal (solo hay un máximo) no invertible. Dos de los ejemplos más conocidos de ecuaciones 3.2.1 es la llamada familia logística

$$x_{i+1} = \alpha x_i(1 - x_i) , \quad (3.2.2)$$

y la familia de Ricker,

$$x_{i+1} = x_i e^{\alpha(1-x_i)} , \quad (3.2.3)$$

donde  $x \in [0, 1]$  en el primer caso y  $x \in [0, \infty)$ , en el segundo caso.

En la Figura 3.1 se muestra la gráfica de un mapeo de Ricker. Este tipo de mapeos se pueden ver graficando iteraciones de un punto de la siguiente forma: una vez trazadas la función identidad y el mapeo de Ricker para algún  $\alpha$  fijo, por el punto  $(x_0, 0)$  se traza un segmento vertical que interseca a la gráfica de  $f$  en el punto  $(x_0, f(x_0))$ , por el que se traza un segmento horizontal que interseca a la recta  $y = x$  en un punto  $p = (f(x_0), f(x_0))$ . La proyección de  $p$  sobre el eje  $x$  es el punto  $(x_1, 0)$ , donde  $x_1 = f(x_0)$ . Al repetir este procedimiento se generan los puntos  $x_2, x_3, \dots$  que dependen de  $x_0$ . Es claro que la primera y segunda iteración de  $x_c$  (la preimagen del valor máximo de  $f$ ), define una región *invariante*. En ésta figura 3.1  $x_s$  y  $x_m$  denotan la primera y segunda iteración de  $x_c$  respectivamente.

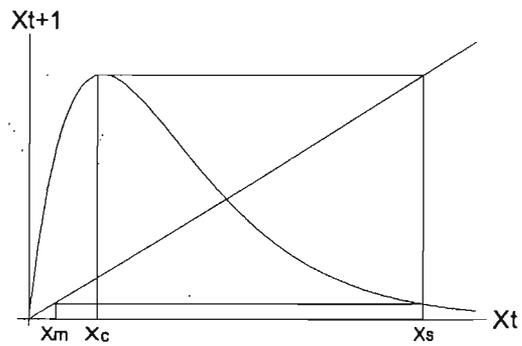


Figura 3.1: La ecuación de Ricker con  $\alpha = 2.7$ , punto inicial  $x=0.35$ . Se realizaron 2 iteraciones.

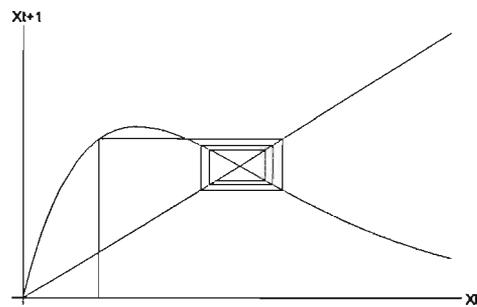
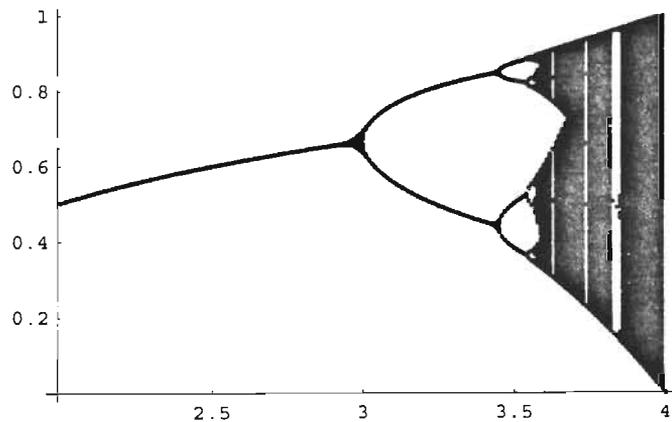


Figura 3.2: Se observa como las trayectorias se aproximan a un punto; los parámetros son:  $x_0 = 0.35$ ,  $\alpha = 1.9$ . Se hicieron 8 iteraciones.



En general la dinámica asintótica varía dependiendo del valor de  $\alpha$ . En la figura 3.3, se muestra el llamado diagrama de bifurcación que resume la dinámica del mapeo de la ecuación (3.2.3) (Este diagrama es típico de las familias de mapeos con un único máximo cuadrático (ver [col])). Cuando  $\alpha < 2$ , la ecuación (3.2.3), tiene un punto fijo estable en  $x = 1$ , ver figura 3.2. Para  $\alpha > 2$ , vemos la familiar cascada de doblamiento de período 4 hasta un valor  $\alpha_c$ , a partir del cual, los mapeos pueden presentar dinámica caótica, figura 3.3.

En la Figura 3.1. Se muestran las regiones invariantes discutidas anteriormente. Es fácil ver que al incrementar el valor de  $\alpha$ , el límite superior  $x_s$  (primera iteración de  $x_c$ ) crece sin restricción, mientras que el límite inferior  $x_m$  (segunda iteración de  $x_c$ ) tiende a cero. Nótese que muchas trayectorias se aproximan a la frontera de la región invariante, de modo que la longitud de esta región da una buena estimación del tamaño de las fluctuaciones de una órbita.

Una forma útil de analizar el comportamiento de la ecuación (3.2.3), es por medio de la *distribución invariante* (ver [lich]). Esta distribución invariante se define como:

$$P_I(A) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \chi_A(x_i) \quad (3.2.4)$$

siempre que este límite exista y no dependa de la condición inicial  $x_0$ , excepto posiblemente por un conjunto de medida cero. Aquí  $\{x_i\}$  es la órbita de  $x_0$  bajo la aplicación  $f : I \rightarrow I$  y  $\chi_A$  denota a la función característica del conjunto  $A \subset I$ . Si  $P_I$  tiene una función de densidad  $f_I$ ,

$$P_I(A) = \int_A f_I(x) dx, \quad (3.2.5)$$

$f_I$  es llamada la “densidad invariante”. Hay que notar que si  $f_I > 0$  en algún intervalo  $[a, b]$ , cada subintervalo de  $[a, b]$  contendrá una infinidad de  $x_i$ , cuando  $N \rightarrow \infty$  y por lo tanto habrá fluctuaciones aperiódicas en la dinámica determinística.

La figura 3.4 muestra algunas densidades invariantes representativas del mapeo de Ricker, que se obtuvieron iterando la ecuación 10,000 veces (3.2.3). Para dinámicas periódicas, la distribución consiste de “picos” igualmente probables (funciones delta), es decir, distribuciones de combinaciones lineales de funciones deltas. Ver figura 3.4 con  $\alpha = 2.6$ . Para valores de  $\alpha$  correspondientes a las bandas caóticas, parece existir una verdadera densidad invariante, i.e.,  $f_I$  es una función ordinaria real. Esto ha sido probado analíticamente bajo condiciones que satisfacen casi todos los valores de  $\alpha$  que corresponden a las bandas caóticas ver [col]).

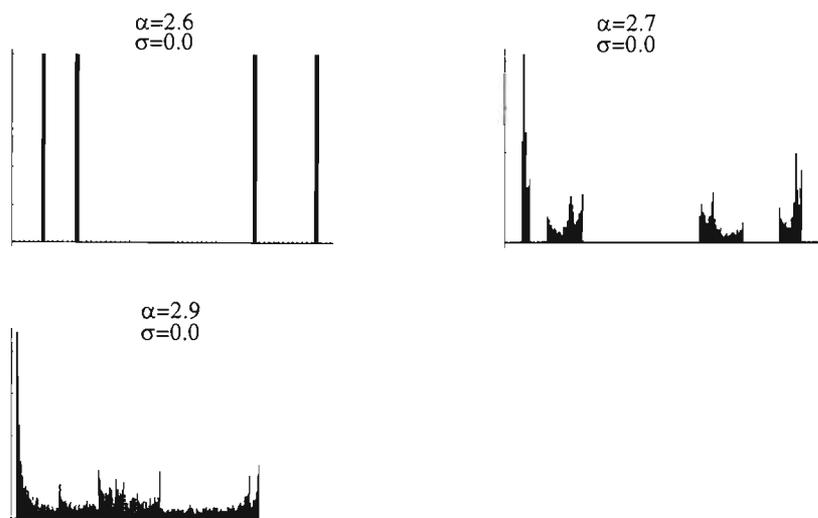


Figura 3.4: Histogramas del mapeo de Ricker sin la presencia de ruido con el parámetro  $\alpha = 2.6, 2.7, 2.9$ .

Otro comportamiento interesante se presenta, cuando la gráfica de  $f_I$  tiene varias componentes (por ejemplo, cuando  $\alpha = 2.7$  en la figura 3.4). En este caso, se presenta lo que se llama *periodicidad ruidosa* de período bajo. Dentro de cada componente la órbita se mueve en una forma regular, pero su visita a las bandas es errática.

### 3.2.2. Análisis gráfico de un mapeo uni-dimensional

Considérese ahora la ecuación 3.2.3 con un ruido aleatorio multiplicativo (se crean números aleatorios entre cero y uno que multiplican a cada iterada del mapeo).

$$x_{t+1} = x_t e^{\alpha(1-x_t)}(1 + \sigma z_t), \quad (3.2.6)$$

donde  $(z_t)_{t=0}^{\infty}$  son variables aleatorias normales independientes con media cero y varianza uno.

Cuando las densidades invariantes son calculadas para  $\sigma > 0$  la estructura del diagrama de bifurcación se hace difuso y finalmente desaparece, esto si el nivel del ruido es suficientemente grande, (es decir  $0 < 1 + \sigma z_t$ ). Para  $\alpha < \alpha_c$ , las órbitas periódicas son reemplazadas por bandas gruesas. La figura 3.4, muestra los efectos al agregar ruido a una órbita estable de período cuatro, con  $\alpha = 2.6$ . Inicialmente, la órbita es reemplazada por una densidad invariante con picos cerca de los puntos del ciclo, en una especie de ciclo ruidoso de período 4.

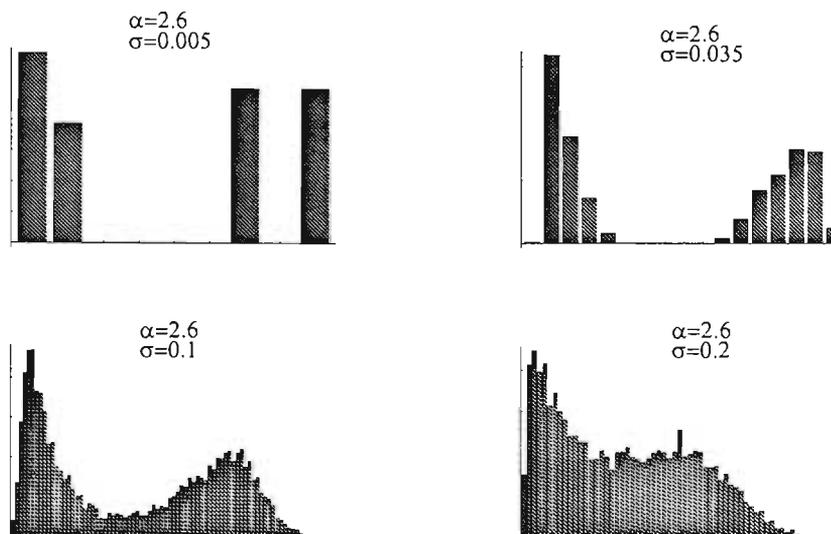


Figura 3.5: Histogramas con pequeñas cantidades de ruido.

Si  $\sigma$  se incrementa surgen picos adyacentes, de modo que para  $\sigma = 0.035$  semeja un ciclo ruidoso de período. Finalmente cuando  $\sigma > 0.1$ , se observa una sóla banda. Para ciclos en la región caótica, el cambio a una distribución aparentemente continua, es mucho más rápido.

### 3.2.3. La no linealidad a través de la serie de tiempo

Dado que pequeñas cantidades de ruido destruyen la estructura de las órbitas periódicas y producen fluctuaciones similares a las aleatorias, resulta interesante analizar el problema de cómo recuperar o reconocer las componentes determinísticas de la dinámica que genera una serie de datos contaminados con ruido. El siguiente teorema muestra que si  $\sigma$  es pequeña,  $X_{t+1}$  se puede determinar de  $X_t$ , con buena aproximación.

**Teorema 3.2.1** *Sea  $X_t$  dado por la Ec 3.2.6. Entonces*

$$\begin{aligned}\mathbb{E}(X_{t+1}) &= X_t \exp(\alpha(1 - X_t)), \\ \text{Var}(X_{t+1}) &= \sigma^2 (X_t \exp(\alpha(1 - X_t)))^2.\end{aligned}\quad (3.2.7)$$

**Demostración.** Claramente,

$$\mathbb{E}(X_{t+1}) = X_t \exp(\alpha(1 - X_t)) \mathbb{E}(1 + Z_t) = X_t \exp(\alpha(1 - X)),$$

ya que  $Z_t$  tiene media cero. Por otra parte,  $\text{Var}(Y_t) = \mathbb{E}(Y_t^2) - (\mathbb{E}(Y_t))^2$ . Sea  $Y_t = X_t \exp(\alpha(1 - X_t))(1 + Z_t)$ , entonces

$(Y_t)^2 = \{(X_t \exp(\alpha(1 - X_t)))^2\} (1 + 2Z_t + Z_t^2)$ . Por lo tanto

$$\begin{aligned} \mathbb{E}(Y_t^2) &= \{(X_t \exp(\alpha(1 - X_t)))^2\} \mathbb{E}((1 + 2Z_t + Z_t^2)) \\ &= \{(X_t \exp(\alpha(1 - X_t)))^2\} \mathbb{E}(Z_t^2) \\ &= \{(X_t \exp(\alpha(1 - X_t)))^2\} \sigma^2. \end{aligned}$$

□

Esto sugiere que los componentes determinísticos de la dinámica, se pueden observar gráficamente los datos  $X_{t+1}$  contra  $X_t$ . Como hemos visto, la eficiencia de este método gráfico, depende del tipo del componente determinístico.

En las figuras 3.6, 3.7, 3.8, se muestran los resultados que se obtuvieron al realizar entre 500 y 800 iteraciones del mapeo de Ricker con ruido (ecuación 3.2.6). Las perturbaciones se tomaron de una distribución normal con media cero y desviación estándar  $\sigma = 0.05, 0.1, 0.2$ . Se observa que hasta la aparición de un ciclo de periodo 2, la componente determinística de la dinámica se nota, es decir, si  $0 < \alpha < 2$ , los puntos de la órbita periódica correspondiente pueden estar tan cercanos que el ruido produzca una mancha indistinguible de un punto fijo, y esta no nos proporciona ninguna información sobre el modelo determinístico del que proviene, esto se ilustra en la figura 3.6. A partir de esto conforme  $\alpha$  crece, es más clara la forma unimodal del mapeo. Obviamente, el punto en el que se empieza a notar la estructura determinística, depende del mapeo. Que la estructura determinística surja depende fuertemente de la estabilidad de la órbita periódica. En la figura 3.6, el mapeo con  $\alpha = 1.5$  tiene un punto fijo con valor propio 0.5 y produce una mancha más alargada que el mapeo correspondiente a  $\alpha = 1.0$  cuyo valor propio es 0. La naturaleza lineal de una pequeña porción del mapeo, es más clara para  $\alpha = 1.5$  que para  $\alpha = 1$ .

Específicamente en las simulaciones anteriores hemos notado:

1. Añadir ruido a mapeos que presentan un punto fijo atractor, produce una nube de puntos que no proporciona información.
2. Agregar ruido a mapeos con ciclos de período dos, muestra una parte de los componentes determinísticos. Este efecto se observa más claramente cuando se pasa de un mapeo con periodo dos, a uno con periodo más grande.
3. Cuando el mapeo es caótico, en los componentes determinísticos se observan claramente, cuando el ruido es nulo o significativamente grande.

### 3.2.4. Un vistazo desde la estadística

Las conclusiones anteriores se basan en una gran cantidad de iteraciones del mapeo, lo cual contrasta con la situación usual en Ecología donde las series

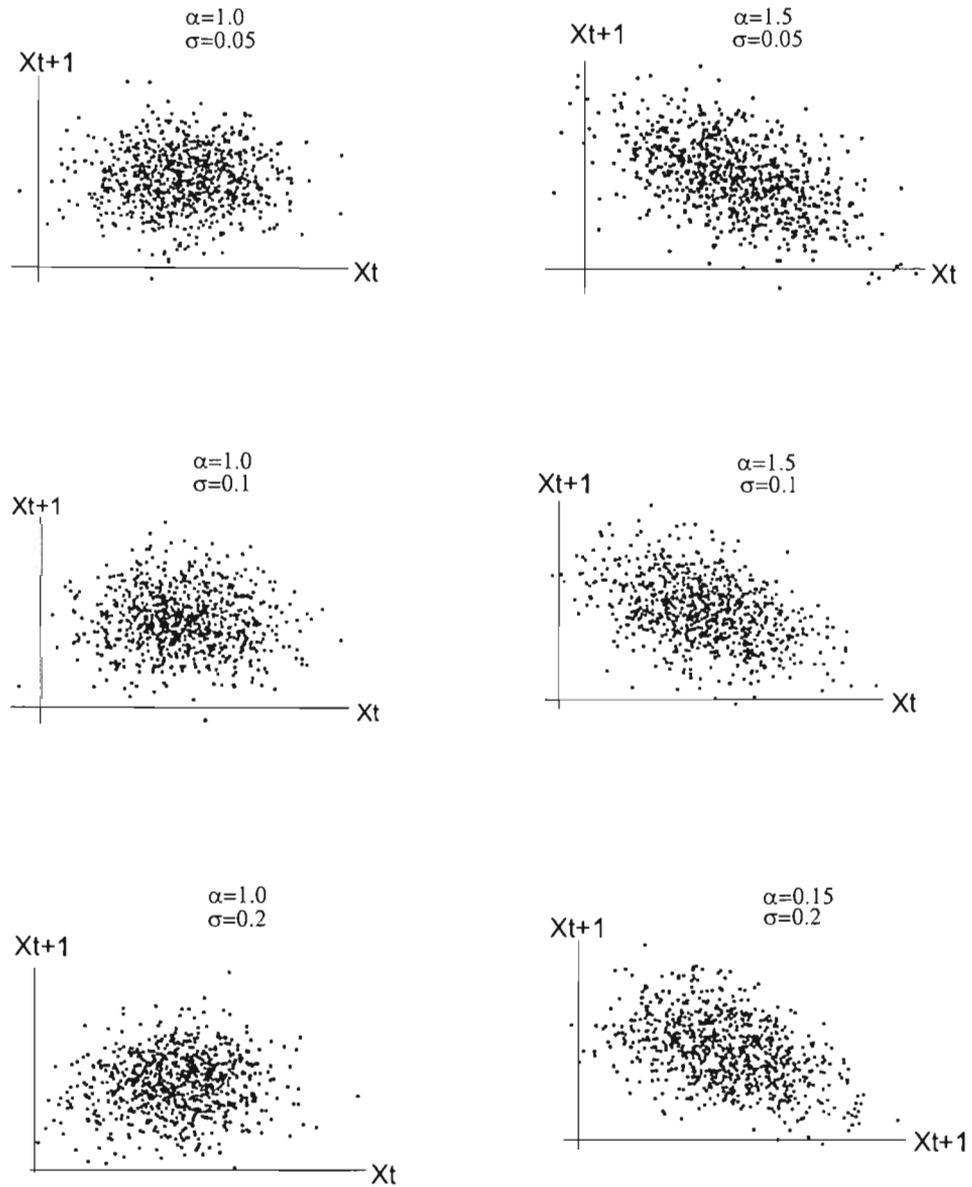


Figura 3.6: Se observa sólo una nube de puntos cuando hay un punto atractor ( $\alpha < 2$ ), con la presencia de ruido.

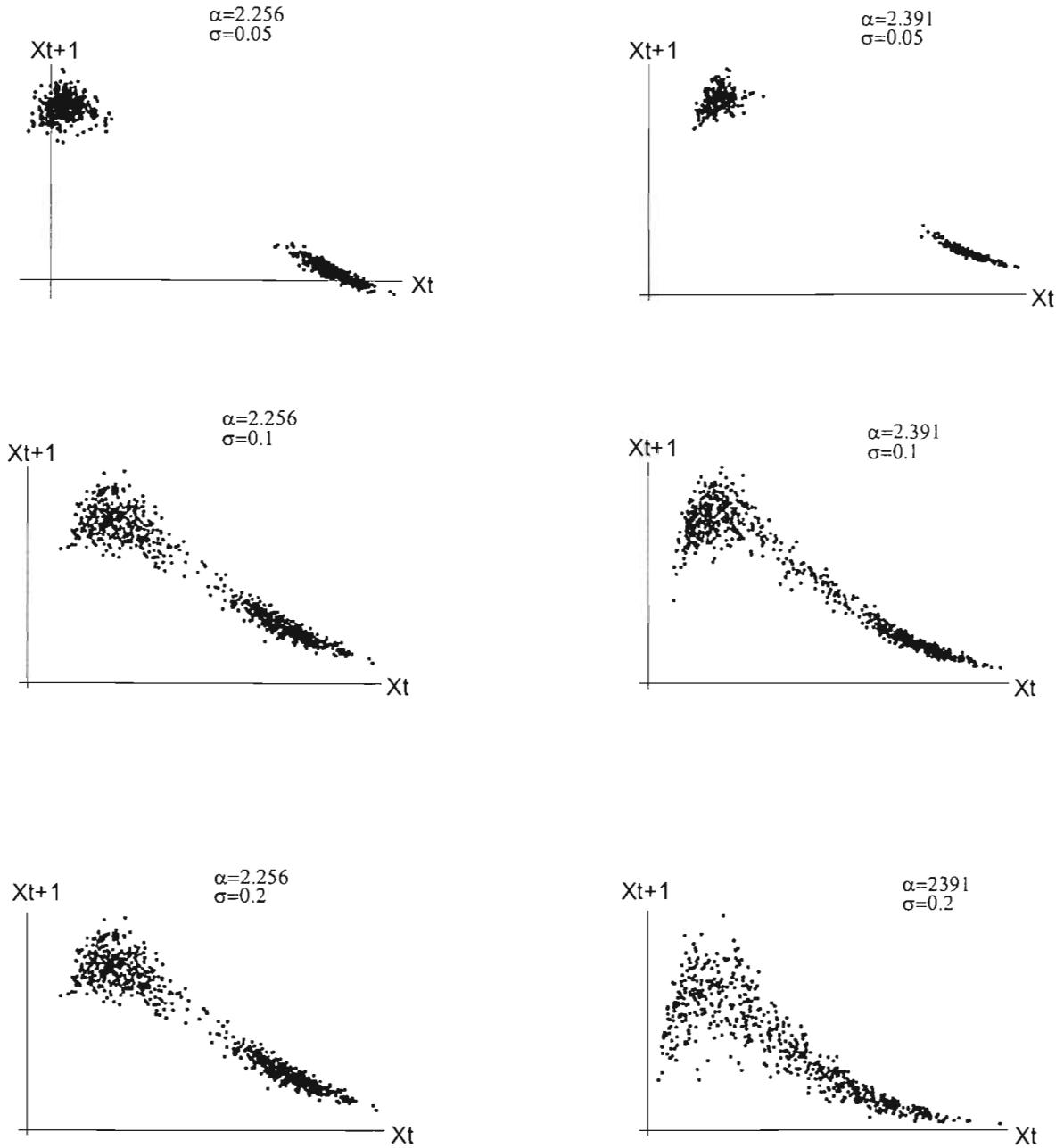


Figura 3.7: Si  $2 < \alpha$  los puntos se empiezan a comportar como el modelo determinístico, aún con la presencia de ruido.

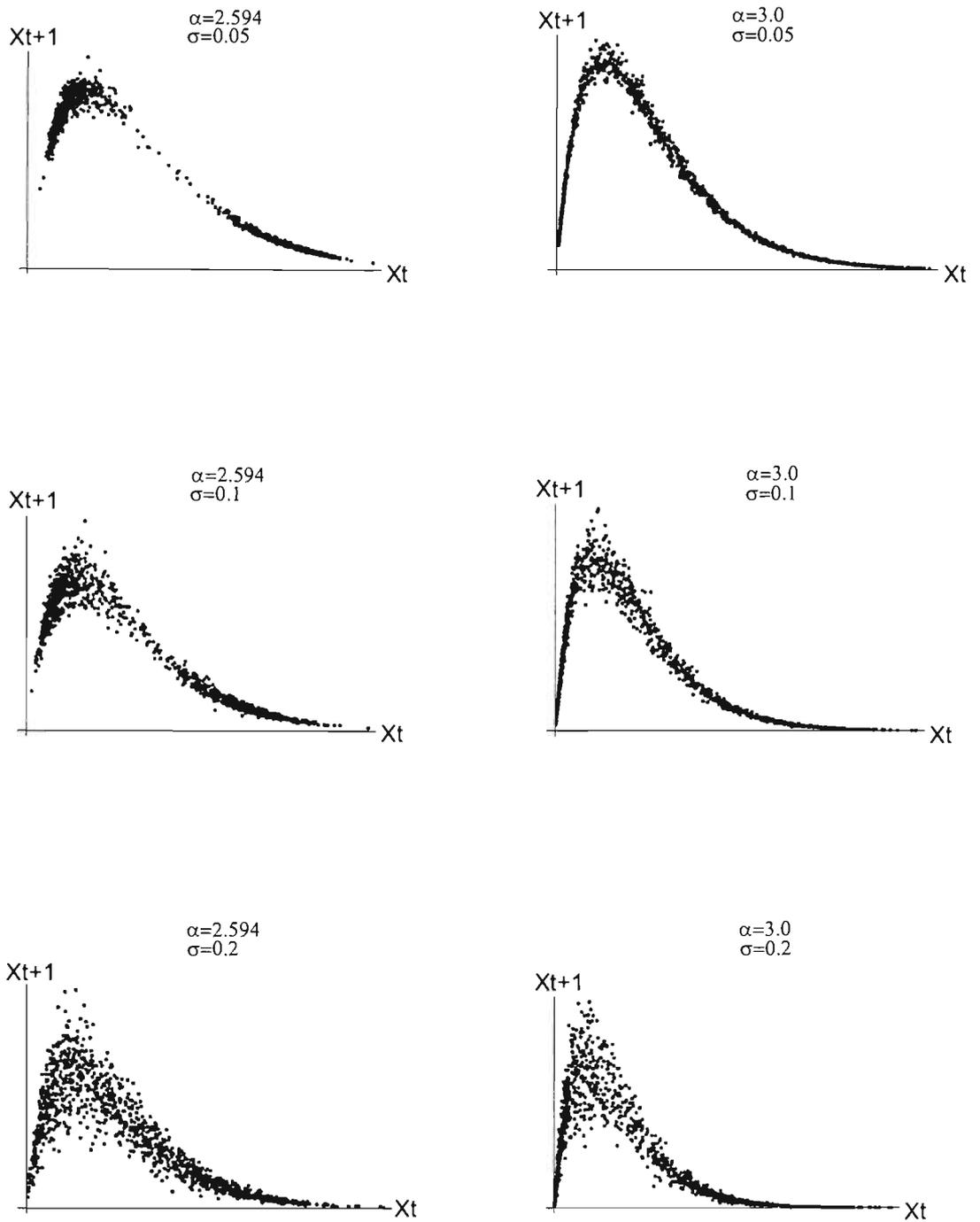


Figura 3.8: Se aprecia que conforme el caos aumenta, los datos se aproximan cada vez más a la curva del modelo determinístico.

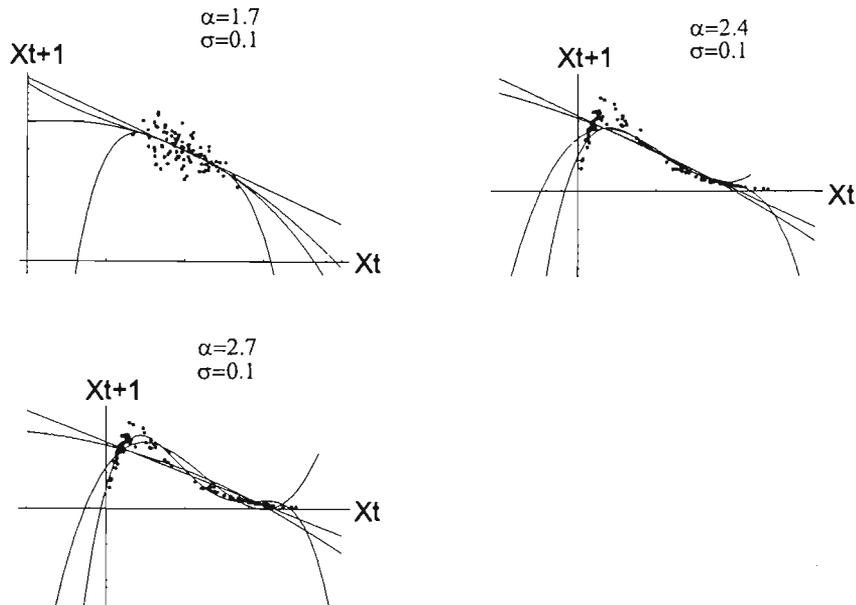


Figura 3.9: Se ajustaron polinomios de primero a cuarto grado manteniendo fija la varianza.

de datos son pequeñas. Por lo tanto es válido preguntarse, si los patrones observados en series grandes se pueden detectar sobre intervalos de tiempo más cortos. En particular, se puede plantear el problema de ¿Cómo usar el análisis estadístico para inferir la existencia de una componente no-lineal en la dinámica a partir de la relación entre los tamaños sucesivos de una población?

Aunque actualmente el análisis estadístico de series de tiempo no-lineales está en pleno desarrollo, sus métodos más efectivos se restringen a unos pocos modelos. Aquí simplemente expondremos un enfoque que puede ser útil a un ecólogo, como una primera aproximación. Se propone a  $x_{t+1}$  como la variable dependiente y se hace una regresión no-lineal tomando a  $x_t$  como variable independiente. Los datos ecológicos es una serie de 100 iteraciones del mapeo correspondiente, y un valor dado del parámetro  $\alpha$ , con ruido (varianza) igual a 0.1. Para cada serie se ajustaron polinomios de primer a cuarto grado. Se observó que para valores cercanos a 1.5, no era significativo agregar términos de órdenes superiores a 1. La adición de términos cúbicos fué significativo a partir de  $\alpha = 2.1$  y muy significativo a partir de  $\alpha = 2.4$ . De este modo, la presencia de la componente no lineal se nota cuando la dinámica es suficientemente complicada. Si el ecólogo trabajara con una población cuya dinámica estuviera determinada por la ecuación de Ricker, podría observar la no linealidad sólo si la población mostrara un ciclo de periodo dos o mayor. Ver figura 3.9.

### 3.2.5. Dimensión de correlación

Supongamos que se tiene un sistema dinámico con espacios de estado  $M$ , donde  $F : M \rightarrow M$  es la ley de cambios de estados. Sea  $h : M \rightarrow \mathbb{R}$  una función suave; es decir,  $h$  es un *observable*. Si al tiempo  $n$  se tiene el estado  $y_n$ , la observación  $x_n$  es  $h(y_n)$ . Definimos las transformaciones  $\Phi_m : M \rightarrow \mathbb{R}$  como

$$\Phi_m(y) = (h(y), h(F(y)), \dots, h(F^{m-1}(y))), \quad (3.2.8)$$

que mandan el estado  $y_n$  en  $(x_n, \dots, x_{n+m-1})$ . De acuerdo con el Teorema de encaje de Takens ver ([tak]), para  $m$  suficientemente grande  $\Phi_m$  es un mapeo suave invertible de  $D$  en  $\mathbb{R}^m$ , donde  $D$  es el atractor del sistema dinámico. De este modo si  $y_n$  está en  $D$ , el vector  $(x_n, \dots, x_{n+m-1})$  está en una imagen fiel del atractor en  $\mathbb{R}^m$ . Por lo tanto, si se tiene una sucesión de observaciones  $\{x_n\}$  del sistema dinámico, existe  $m$  tal que los vectores  $X_n = (x_n, \dots, x_{n+m-1})$  se acercan asintóticamente a la imagen del atractor en  $\mathbb{R}^m$ . En el apéndice se enuncia rigurosamente el Teorema. En [gras] se define la dimensión para distinguir el ruido de un sistema caótico, que se conoce con el nombre de dimensión de correlación. Consideremos el conjunto de puntos  $\{X_i\}_{i=1}^N$  en el (imagen del) atractor. Si el sistema es caótico, debido a la divergencia exponencial de las trayectorias, muchas de las parejas  $(X_i, X_j)$  con  $i \neq j$  serán dinámicamente parejas de puntos esencialmente aleatorios no correlacionados. Pero los puntos están sobre un atractor y entonces estarán espacialmente correlacionados. Esta correlación espacial se mide con la integral de correlación  $C(g, m)$ , definida por:

$$C(g, m) = \lim_{N \rightarrow \infty} \frac{1}{N(N-1)} \sum_{i \neq j}^N \theta(g - \|X_i - X_j\|), \quad (3.2.9)$$

donde

$$\theta(x) = \begin{cases} 1, & t \in (0, \infty), \\ 0 & t \in (-\infty, 0], \end{cases}$$

y  $g$  es un real positivo. Grassberger y Procaccia ([gras]) demostraron que para  $g$  pequeña y  $m$  suficientemente grande,  $C(g, m)$  crece como una potencia, es decir

$$C(g, m) \sim kg^\nu. \quad (3.2.10)$$

Para ver si se puede reconocer la presencia de componentes determinísticas en una serie de tiempo, realizamos el siguiente experimento. Se generaron series de tiempo con y sin ruido para el mapeo de Ricker para  $\alpha = 1, 2.0, 2.7, 3.0, 3.5$ . Como en la sección anterior, el ruido fue Gaussiano. También se consideró una serie de tiempo puramente ruidosa. Se calcularon las dimensiones  $C(g, m)$  de

cada una de las series para las dimensiones de inmersión  $m = 1, 2, 3, 4$  y  $g = 1/100, 2/200, \dots, 1$ . La manera en que se obtuvieron los cálculos de  $C(g, m)$  es la siguiente. Si tomamos la igualdad y aplicamos el logaritmo en la ecuación 3.2.10 se obtiene la relación lineal.

$$\log(C(g, m)) = \log(k) + \nu \log(g). \tag{3.2.11}$$

Por medio de una regresión lineal entre  $\log(g)$  y  $\log(C(g, m))$ , se calcula una aproximación de  $\nu$ . Los cálculos se realizaron con el mapeo de Ricker, para una dinámica de punto fijo atractor ( $\alpha = 1, 2$ ) y cuando se presenta caos ( $\alpha = 2.7, 3.0, 3.5$ ). Los resultados que se obtuvieron se presentan en la tabla 1.

$\sigma = 0.00$			$\sigma = 0,05$		
$\alpha=1.0$ (punto fijo)			$\alpha=1.0$ (punto fijo)		
M	$r^2$	$\nu$	M	$r^2$	$\nu$
1	0.773	0.01030	1	0.99	0.9419
2	0.773	0.01035	2	0.99	1.9172
3	0.773	0.01040	3	0.99	2.9479
4	0.773	0.01046	4	0.98	3.8714
$\alpha=2.0$ (punto fijo)			$\alpha=2.0$ (punto fijo)		
1	0.227	0.04070	1	0.99	0.9906
2	0.227	0.04071	2	0.99	1.9484
3	0.227	0.04072	3	0.99	2.8893
4	0.227	0.04073	4	0.98	3.8076
$\alpha=2.7$ (caos)			$\alpha=2.7$ (caos)		
1	0.99	0.7228	1	0.99	0.8821
2	0.99	0.8667	2	0.99	1.2195
3	0.99	0.8829	3	0.99	1.6373
4	0.99	0.9167	4	0.99	2.0047
$\alpha=3.0$ (caos)			$\alpha=3.0$ (caos)		
1	0.99	0.8160	1	0.99	0.7648
2	0.99	1.0822	2	0.99	1.0603
3	0.99	1.1416	3	0.95	1.3245
4	0.98	1.1674	4	0.92	1.3730
$\alpha=3.5$ (caos)			$\alpha=3.5$ (caos)		
1	0.99	0.5431	1	0.99	0.5237
2	0.99	0.8343	2	0.99	0.8582
3	0.99	1.0057	3	0.99	1.1641
4	0.99	1.0900	4	0.99	1.2939

(a) Series sin ruido

(b) Series con ruido

Tabla 1.

**Conclusiones:** Si hay un punto atractor (casos  $\alpha = 1$  y  $2$ ) y no hay ruido, entonces la dimensión de correlación es cero, pero al agregar ruido, la dimensión de correlación aumenta casi como la dimensión de inmersión. En este caso, la serie se comporta como una serie puramente ruidosa y la parte determinística no se puede detectar. Si consideramos el modelo caótico ( $\alpha = 2.7, 3.0, 3.5$ ), pero sin la presencia de ruido, se observa que la dimensión de correlación ya no crece tanto como la dimensión de inmersión, sino que converge a un número; esto mismo ocurre aún si agregamos ruido. Por lo tanto, la dimensión de correlación podría detectar la existencia de componentes determinísticas si éstas son caóticas. A continuación se realizará un estudio con el mapeo logístico para comparar los resultados con los del mapeo de Ricker. Como en el caso anterior se realizaron cálculos con ruido y sin la presencia de ruido. Los resultados que se obtuvieron se presentan en la tabla 2.

$\sigma = 0.00$			$\sigma = 0,05$		
$\alpha=3.2$			$\alpha=3.2$		
M	$r^2$	$\nu$	M	$r^2$	$\nu$
1	0.669	0.2804	1	0.99	0.8983
2	0.669	0.2804	2	0.99	1.4296
3	0.669	0.2804	3	0.99	2.1765
4	0.669	0.2804	4	0.99	2.2584
$\alpha=3.6$			$\alpha=3.6$		
1	0.99	0.6731	1	0.99	0.8494
2	0.99	0.6915	2	0.99	1.4023
3	0.99	0.7564	3	0.99	2.1802
4	0.99	0.8324	4	0.99	2.2222
$\alpha=3.82$			$\alpha=3.82$		
1	0.99	0.7247	1	0.99	0.8530
2	0.99	0.8348	2	0.99	1.0766
3	0.99	0.8415	3	0.98	1.4163
4	0.99	0.8508	4	0.99	1.4630
$\alpha=4.0$			$\alpha=4.0$		
1	0.99	0.7539	1	0.99	0.8284
2	0.99	0.9170	2	0.99	1.0534
3	0.99	0.9598	3	0.99	1.2832
4	0.99	0.9902	4	0.99	1.4828

(a) Series sin ruido

(b) Series con ruido

Tabla 2.

Como en el caso anterior (mapeo de Ricker), la dimensión de correlación tiende a un número, conforme la dimensión de inversión aumenta (de 1 a 4). De forma análoga, la dimensión de correlación podría detectar la existencia de componentes determinísticos si éstos son caóticos, ya que la presencia del ruido no oculta la no linealidad. Recordemos que si el ruido dominara, encontraríamos que la dimensión de correlación crecería tanto como la de inmersión.

**Conclusiones:** Según los resultados que se han dado en las tablas 1 y 2, tanto en la del mapeo de Ricker como en el logístico, con y sin la presencia de ruido, observamos que al agregar ruido a un mapeo caótico, la dimensión de correlación puede detectar la presencia de componentes no lineales de un modelo determinístico.



# Capítulo 4

## Programas de simulación

### 4.1. Introducción

Las siguientes funciones son para las simulaciones que se realizaron en el capítulo 1 Secc. 2.1.1

Antes de empezar a enunciar las funciones para los distintos programas de simulación, listaremos los paquetes estadísticos que se requieren. Estos paquetes deben ser cargados en forma separada, como se muestra a continuación.

```
<< Statistics`LinearRegression`  
<< Statistics`DataManipulation`  
<< Graphics`Graphics`  
<< Statistics`NormalDistribution`
```

### 4.2. Índice IDH

La función `sucesión` genera sucesiones de ceros y unos de tamaño  $n$  y calcula su IDH (ver sección 2.1.1). Para generar cada una de las sucesiones se introduce el índice  $\delta$  (ver Cap.1 Secc. 2.13). La probabilidad de obtener cero en cada una de las sucesiones es `pcero`, al parámetro  $\delta$  la llamamos `agregacion` (ver capítulo 1 sección 2.1.1); la longitud de cada sucesión es  $n$ . Hay que recordar que si `agregacion` tiende a  $-1$ , entonces la sucesión presentará la distribución de ceros alternada; pero si `agregacion` es cero, entonces la sucesión es completamente aleatoria; y si `agregacion` tiende a 1 las sucesiones tendrán grandes bloques de ceros o grandes bloques de unos.

Para generar una sucesión se procede de la siguiente manera. Se da una probabilidad `pcero` fija, después se calcula un número aleatorio entre cero y uno; si el número aleatorio es menor que `pcero`, entonces se pone un cero, en caso contrario se pone uno. Esto da el primer término de la sucesión; para calcular los siguientes términos se usa la probabilidad condicional correspondiente al término presente de la sucesión. Es decir, si  $\alpha \in \{0, 1\}$  es el término  $i$ ésimo, entonces el término  $(i+1)$ ésimo se calcula así: se elige un número aleatorio entre 0 y 1; si el número es menor que  $P(\alpha|0)$  se pone un cero y un uno en caso contrario.

Las funciones `N00`, `N11`, `N10` y `N01`, cuentan el número de parejas cero-cero, uno-uno, uno-cero y cero-uno; que hay en una sucesión dada de ceros y unos. Por ejemplo en la sucesión  $M = \{1, 1, 0, 0, 0, 0, 1, 0\}$ , hay una pareja de unos, tres parejas de ceros, dos parejas de uno-cero y una pareja de cero-uno. Las funciones `N1` y `N0`, cuentan el número de ceros y unos respectivamente, que aparecen en cada sucesión.

La función `pares` genera parejas de datos de la forma  $(a_i, a_{i+1})$ , de una sucesión dada  $l = \{a_1, a_2, \dots, a_n\}$ .

```

sucesion[n_, pcero_, agregacion_] :=
Module[{a, p00, p11, p10, p01, puno}, puno = 1 - pcero;
  a = {}; p00 = agregacion + pcero (1 - agregacion);
  p11 = 1 - pcero (1 - agregacion); p10 = (1 - pcero) (1 - agregacion);
  p01 = pcero (1 - agregacion); If[Random[] ≤ pcero, a = Append[a, 0],
  a = Append[a, 1]]; For[i = 1, i < n, i++, If[a[[i]] == 0,
  If[Random[] ≤ p00, a = Append[a, 0], a = Append[a, 1]],
  If[Random[] ≤ p11, a = Append[a, 1], a = Append[a, 0]]]]; a]

```

A continuación se dan ejemplos donde se muestran sucesiones con diferentes distribuciones de ceros y unos. Para esto se utilizan diversas combinaciones de los parámetros `agregacion` y `pcero`. Para crear una sucesión corremos la función `sucesion`. En los siguientes ejemplos el tamaño de la sucesión es 20 ( $n=20$ ) las diferentes probabilidades para obtener cero se eligen en la segunda entrada de la función; en los ejemplos, la probabilidad de obtener cero es `pcero = 0.5`; el tercer parámetro es `agregacion` (que varía de -1 a 1) y que en los ejemplos toma los valores de -0.8, 0.0 y 0.8.

```

sucesion[20, .5, -0.8]
{0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0}

```

```

sucesion[20, .5, 0.0]
{0, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1}

```

```

sucesion[20, .5, 0.8]
{0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1}

```

La función `Indice` calcula el índice *IDH* de una sucesión *l*. Ver Cap 2 Secc. 2.1.1.

```

Indice[l_] := Module[ {s}, s = pares[l];
  (N00[s] N11[s] - N01[s] N10[s]) / (N0[1] N1[1])

```

`TablaDeIndices` calcula una tabla de índices *IDH*. La primera entrada `tamanotabla` indica el número de sucesiones que se deseén calcular; la segunda `longsecuencia` es para el tamaño de las sucesiones. Los últimos dos parámetros `pcero` y `agregacion` se definieron antes.

```

TablaDeIndices[tamanotabla_, longsecuencia_, pcero_, agregacion_] :=
  Table[Indice[cadena[longsecuencia, pcero, agregacion]],
    {j, 1, tamanotabla}]

```

Veamos un ejemplo. La siguiente tabla muestra los índices de 20 sucesiones de tamaño 100. La probabilidad `pcero` es 0.5 y `agregacion` es 0 (la distribución de ceros es totalmente aleatoria).

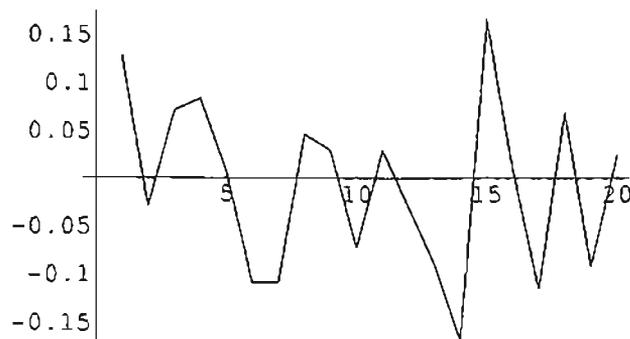
```

TabInd = N[TablaDeIndices[20, 100, 0.5, 0]]
{0.127404, -0.0296, 0.0709914, 0.0833333, 0.00738916,
 -0.109244, -0.109244, 0.0461662, 0.0296118, -0.073063,
 0.0296, -0.0296, -0.0892357, -0.1684, 0.165395,
 0.00960384, -0.115942, 0.0692277, -0.091931, 0.0249597}

```

La función `lisplot` muestra la gráfica de los índices *IDH* unidos por poligonales.

```
ListPlot[TabInd, PlotJoined → True]
```



```
- Graphics -
```

La función `datos` genera una lista de 100 sucesiones, una por cada probabilidad  $p_{\text{cero}} = \frac{i}{100}$ , con  $i = 1, 2, \dots, 100$ . El parámetro  $\delta$  de las sucesiones está dado por `agregacion` y cada sucesión tiene longitud `long`. El tamaño predeterminado de `long` es 200.

```
datos[agregacion_, long_: 200] :=  
  Table[sucesion[long, i/100, agregacion], {i, 1, 20}]
```

### 4.3. La proporción de ceros en una sucesión

`ProbProporcion` muestra la tabla de entradas  $(x, y)$ ,  $x$  es la probabilidad de obtener cero ( $p_{\text{cero}} = \frac{i}{20}$ ) e  $y$  es el promedio de ceros que aparecen en cada sucesión. Todas las sucesiones están generadas con el mismo valor de `agregacion`. La longitud predeterminada de cada sucesión es 200.

```
ProbProporcion[agregacion_, long_: 200] :=  
  Table[{i/20, N[(long - Apply[Plus, datos[agregacion, long][[i]])] /  
    long]}, {i, 1, 20}]
```

`ProbProporcion` muestra la tabla de entradas  $(x, y)$  con `agregacion=0.3`.

```
ll = N[ProbProporcion[0.3]]  
  
{ {0.05, 0.025}, {0.1, 0.015}, {0.15, 0.03}, {0.2, 0.02}, {0.25, 0.02},  
  {0.3, 0.055}, {0.35, 0.06}, {0.4, 0.09}, {0.45, 0.1}, {0.5, 0.105},  
  {0.55, 0.075}, {0.6, 0.12}, {0.65, 0.13}, {0.7, 0.16}, {0.75, 0.155},  
  {0.8, 0.175}, {0.85, 0.175}, {0.9, 0.17}, {0.95, 0.195}, {1., 0.205}}
```

La función `regress` calcula la regresión lineal entre las parejas (x, y) de la tabla que se generó con la función `ProbProporcion`

```
Regress[l1, {1, x}, x]
```

```
ParameterTable →
```

	"Estimate"	"SE"	"TStat"	"PValue"
1	-0.0075	0.0066	-1.1384	0.2698
x	0.21248	0.0110	19.1842	1.9739

```
RSquared → 0.9533, AdjustedRSquared → 0.9507,  
EstimatedVariance → 0.0002,
```

```
ANOVATable →
```

□	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	1	0.0750	0.0750	368.0349	1.9739
"Error"	18	0.0036	0.0002	" "	" "
"Total"	19	0.0787	" "	" "	" "

## 4.4. Índice IDH vs índice IZU

Las siguientes funciones son para analizar la relación que podría existir entre el índice *IDH* y el índice *IZU*. Para esto es necesario correr las funciones listadas anteriormente: `sucesion`, `N00`, `N11`, `N10`, `N01` y `pares`.

Las funciones `Idh` e `Izu` calculan los índices *IDH* e *IZU* respectivamente (ver Cap.1 Secc. 2.1.2 y 2.1.5)

```
Idh[l_] := Module[ {s}, s = pares[l];  
  (N00[s] N11[s] - N01[s] N10[s]) / (N0[l] N1[l]) ]
```

```
Izu[l_] :=  
  Module[ {s}, s = pares[l]; (N00[s] + N11[s]) / (N0[l] + N1[l]) ]
```

La función `TablaDeIdhIzu`, genera una tabla de parejas (x, y), donde x es el índice *IDH* e y es el índice *IZU*. `tamanotabla` es el número de sucesiones, `lonsecuencia` es el tamaño de cada sucesión, `pzero` y `agregacion` son parámetros ya definidos.

```

TablaDeIdhIzu[tamanotabla_, longsecuencia_, pcero_, agregacion_] :=
  Table[{Idh[sucesion[longsecuencia, pcero, agregacion]],
        Izu[sucesion[longsecuencia, pcero, agregacion]]},
        {j, 1, tamanotabla}]

```

A continuación se calcula una tabla con esta función, los parámetros que se utilizan son: `tamanotabla=10`, `longsecuencia=50`, `pcero=0.5`, y `agregacion=0.1`.

```

12 = N[TablaDeIdhIzu[10, 50, 0.5, .1]]

{{0.251603, 0.58}, {0.176, 0.52}, {0.0576923, 0.38},
 {0.0344828, 0.36}, {0.0592, 0.7}, {0.216, 0.5}, {0.0977564, 0.52},
 {-0.10467, 0.56}, {0.0976, 0.46}, {-0.099359, 0.46}}

```

La función `Regress` calcula la regresión lineal entre el índice *IDH* y el *IZU*, de los datos que se generaron con la función `TablaDeIdhIzu` y que se encuentran en la lista `12`.

```
Regress[12, {1, x}, x]
```

ParameterTable →

	"Estimate"	"SE"	"TStat"	"PValue"
1	0.4938	0.0399	12.3730	1.6972
x	0.1286	0.2904	0.44276	0.6696

RSquared → 0.0239, AdjustedRSquared → -0.0980,

EstimatedVariance → 0.0107

ANOVATable →

	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	1	0.0021	0.0021	0.1960	0.6696
"Error"	8	0.0857	0.0107	" "	" "
"Total"	9	0.0878	" "	" "	" "

## 4.5. El mapeo de Ricker

Las siguientes funciones son para analizar el mapeo de Ricker cuando se presenta en forma simple, así y cuando es perturbado por agentes externos a lo que le llamamos ruido.

La función `f` introduce el mapeo de Ricker (ver Cap.3 Secc.3.2.1), en este caso con  $\alpha = 1.9$ .

```
f[x_] := x Exp[1.9 (1 - x)]
```

La función `lister` genera una lista de iteradas del mapeo, la primera entrada es para declarar la función de Ricker, la segunda es para el punto inicial de las iteraciones y la tercera entrada es para el número de iteradas que se necesiten hacer. En el siguiente ejemplo el punto inicial  $x_0$  es 0.35 y se calculan 6 iteradas.

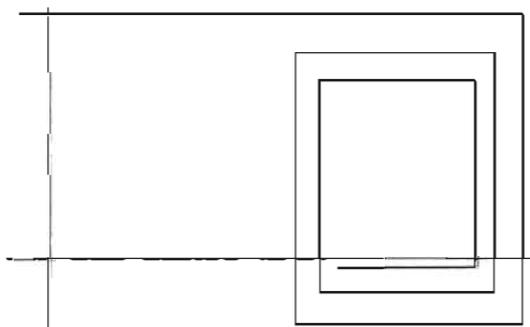
```
lister = NestList[f, 0.35, 6]
{0.35, 1.20343, 0.817631, 1.15622, 0.859278, 1.12267, 0.889263}
```

`lpuntos` genera una lista de parejas de iteradas de la siguiente forma  $(a_i, a_{i+1}), (a_{i+1}, a_{i+2}), \dots, (a_{n-1}, a_n)$

```
lpuntos = Flatten[Table[{{lister[[i]], lister[[i + 1]]},
  {lister[[i + 1]], lister[[i + 1]]}}, {i, 1, 6}], 1]
{{0.35, 1.20343}, {1.20343, 1.20343}, {1.20343, 0.817631},
 {0.817631, 0.817631}, {0.817631, 1.15622}, {1.15622, 1.15622},
 {1.15622, 0.859278}, {0.859278, 0.859278}, {0.859278, 1.12267},
 {1.12267, 1.12267}, {1.12267, 0.889263}, {0.889263, 0.889263}}
```

La función `ListPlot` une por medio de polinomiales, a las parejas de iteradas que se formaron con la función `lpuntos`. A esta gráfica la llamamos `h1`.

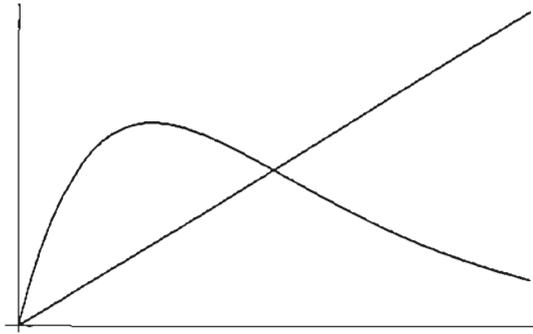
```
h1 = ListPlot[lpuntos, PlotJoined -> True, Ticks -> None]
```



- Graphics -

`Plot`, grafica el mapeo de Ricker junto con la función identidad. A esta gráfica la llamamos `h2`. Los números 0 y 2 forman el intervalo  $(0, 2)$ , dominio de ambas funciones.

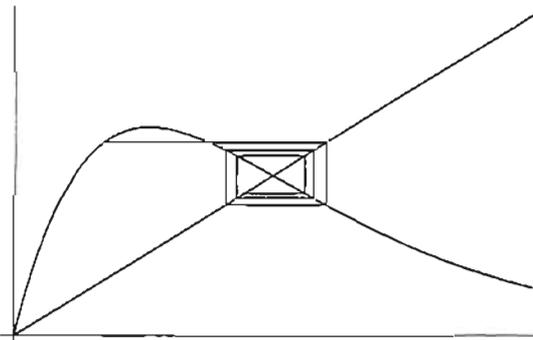
```
h2 = Plot[{f[x], x}, {x, 0, 2}, Ticks -> None]
```



- Graphics -

La función **Show** presenta en el mismo plano coordenado las gráficas **h1** y **h2**.

```
gra = Show[h1, h2]
```



- Graphics -

## 4.6. Mapeo de Ricker con ruido

Para introducir el ruido, se utilizará la función de distribución normal ( $z[s]$ ), con media cero y varianza uno. Con esta función se generarán números aleatorios entre cero y uno, cada número generado multiplica a cada entrada de la serie de tiempo del mapeo de Ricker. De esta forma se introduce el ruido en el mapeo. Ver Cap. 3 Secc. 3.2.2.

```
z[s_] := NormalDistribution[0, s]
```

`fRMul`, introduce el mapeo de Ricker con la presencia de ruido (ecu.3.2.6). En la función hay tres variables; la primera ( $\alpha$ ) es para el parámetro de Ricker, la segunda es la varianza ( $s$ ) de la función de distribución normal y el tercero ( $x$ ) la variable independiente.

```
fRMul[α_, s_][x_] := (x Exp[α (1 - x)]) (1 + Random[z[s]])
```

`literMul`, crea una lista de iteradas de la Ecu. 3.2.6. En la primera entrada se declaran los distintos valores de  $\alpha$ ; la segunda es para la varianza  $s$ ; la tercera es para el punto inicial  $x_0$  y `numiter` es para el número de iteradas que se quieran calcular.

```
literMul[α_, s_, pinicial_, numiter_] :=  
  NestList[fRMul[α, s], pinicial, numiter]
```

A continuación se da un ejemplo de la función `literMul`, con las siguientes condiciones:  $\alpha = 1,5$ ;  $s = 0,2$ ;  $x_0 = 0,3$  y se realizan 10 iteraciones. A la lista la llamamos `lis`.

```
lis = literMul[1.5, 0.2, 0.3, 10]  
  
{0.3, 0.993997, 0.766781, 0.940261, 0.841465,  
 1.28143, 0.855957, 1.01176, 1.16662, 1.06044, 0.866947}
```

## 4.7. Histogramas

Las siguientes instrucciones son para crear histogramas de frecuencia con los datos generados de las iteraciones del mapeo de Ricker que se encuentran en la lista `lis`. La función `frec`, cuenta el número de frecuencias que ocurren en un intervalo dado.

```
frec[lis_, extizq_, extder_, ancho_] :=  
  BinCounts[lis, {extizq, extder, ancho}]
```

La función `histograma` crea el histograma de frecuencias, en él se debe dar el límite inferior y superior del intervalo en que se necesite hacer el histograma.

```
histograma[lis_, extizq_, extder_, ancho_] :=  
  BarChart[frec[lis, extizq, extder, ancho]]
```

Las funciones `Max` y `Min` proporcionan el valor máximo y mínimo de las

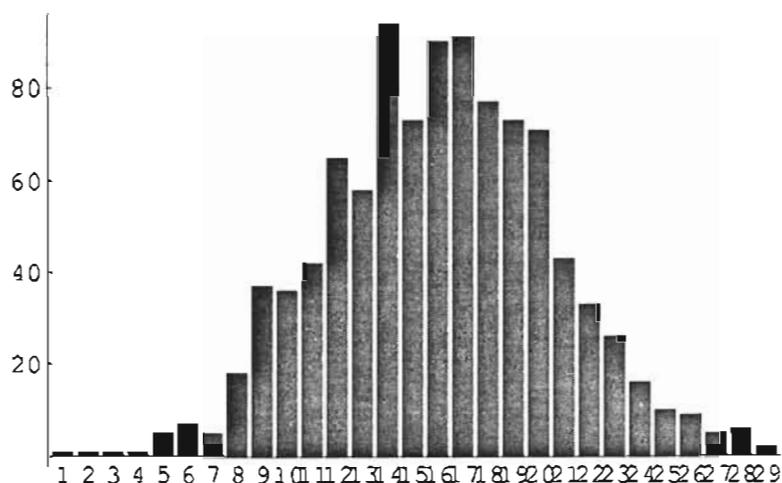
iteradas; con estos datos se crea un intervalo, el cual se divide en n-partes para crear subintervalos, en ellos se cuenta el número de ocurrencias de las iteraciones, con éstos datos se crea el histograma de frecuencias.

```
{Min[lis], Max[lis]}
```

```
{0.2, 1.6}
```

Para crear el histograma se deben de introducir en la función todos los datos que se han estado recabando, como se presenta a continuación. La primera entrada se refiere a la lista *lis* que se generó con la función *frec*; los números 0,2 y 1,6 son los valores mínimos y máximos de las iteraciones; 0,05 es el ancho de cada barra del histograma.

```
h7 = histograma[lis, 0.2, 1.6, 0.05]
```



- Graphics -

Nota si se quieren graficar histogramas con diferentes datos, es necesario cambiar los parámetros en la función *literMul*, y despues correr nuevamente *Min*, *Max* y *h7=histograma*.

## 4.8. Gráficos de dispersión

En seguida se enuncian las funciones para realizar gráficos de dispersión del mapeo de Ricker, posteriormente se buscarán polinomios de grado *n*, que mejor se ajusten a los puntos generados por las iteraciones del mapeo.

Para esto es necesario tener la lista `lis` que se generó al correr las funciones `(z[s])`, `fRMul`, `lis`, `litterMul` y `lis=litterMul`.

La función `Partition` genera una tabla de parejas de la lista `lis`, de la forma  $(x_0, x_1), (x_1, x_2), \dots, (x_{n-1}, x_n)$ .

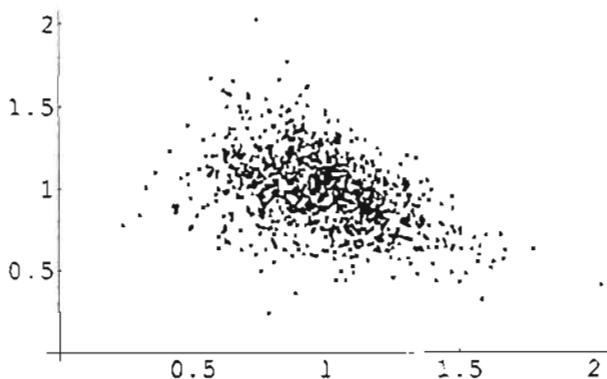
```
P1 = Partition[lis, 2, 1];
```

A continuación se presenta como ejemplo, una pareja formada con la primera y segunda iterada de la lista `lis`, que se creó con la función `Partition`.

```
P1[[1]]
{0.3, 0.834639}
```

`ListPlot`, muestra el gráfico de dispersión de la tabla de parejas de iteradas que se formaron con la función `Partition`. Al gráfico le llamamos `P0a2p5`. Nota la lista de datos que se maneja a partir de esta función es de 1000, esto se hace para que los calculos sean parecidos a los que se manejan en el estudio realizado del capítulo 2.

```
P0a2p5 = ListPlot[P1]
```



- Graphics -

## 4.9. Aproximaciones polinomiales

Dado el gráfico de dispersión, el trabajo que se presenta ahora es encontrar polinomios que mejor se aproximen a los puntos de la iteraciones del mapeo. Para conocer los coeficientes de cada poinomio, se utilizará la función

Regress, con esta calcularemos regresiones polinomiales con diferente grado  $n$ . Por ejemplo si se necesita crear una regresión lineal de la tabla de parejas P1, se corre la función Regress de la siguiente forma:

```
Regress[P1, {1, x, x^2}, x]
```

```
ParameterTable →
```

"	"Estimate"	"SE"	"TStat"	"PValue"
1	1.7145	0.0211	80.9910	2.1695
x	-0.550	0.0463	-11.888	1.5240
x <sup>2</sup>	-0.134	0.0208	-6.4796	1.1545

```
RSquared → 0.7133, AdjustedRSquared → 0.7130,
```

```
EstimatedVariance → 0.1099
```

```
ANOVATable →
```

"	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	2	546.414	273.20	2484.5	0.
"Error"	1997	219.591	0.1099	"	"
"Total"	1999	766.006	"	"	"

Si se quiere calcular una regresión cuadrática, es necesario declararlo en la segunda entrada de la función Regress como se muestra a continuación.

```
Regress[P1, {1, x, x^2}, x]
```

```
ParameterTable →
```

"	"Estimate"	"SE"	"TStat"	"PValue"
1	1.7145	0.0211	80.9910	2.1695
x	-0.550	0.0463	-11.888	1.5240
x <sup>2</sup>	-0.134	0.0208	-6.4796	1.1545

```
RSquared → 0.7133, AdjustedRSquared → 0.7130,
```

```
EstimatedVariance → 0.1099
```

```
ANOVATable →
```

"	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	2	546.414	273.20	2484.5	0.
"Error"	1997	219.591	0.1099	"	"
"Total"	1999	766.006	"	"	"

Siguiendo la forma como se crearon los dos ejemplos anteriores, se pueden crear regresiones polinomiales de diferente grado, de la tabla de iteradas del mapeo introducido. A continuación se realizan regresiones de tercer y

cuarto grado de la misma tabla P1, para obtener todos los coeficientes de los polinomios de primero a cuarto grado.

```
Regress[P1, {1, x, x^2, x^3}, x]
```

```
ParameterTable →
```

"	"Estimate"	"SE"	"TStat"	"PValue"
1	1.1884	0.0255	46.484	2.7224
x	1.9505	0.0955	20.423	0.
x <sup>2</sup>	-2.770	0.0935	-29.62	3.6141
x <sup>3</sup>	0.7556	0.0263	28.690	0.

```
RSquared → 0.7970, AdjustedRSquared → 0.7967,  
EstimatedVariance → 0.0778
```

```
ANOVATable →
```

"	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	3	610.530	203.51	2612.6	2.0801
"Error"	1996	155.475	0.0778	"	"
"Total"	1999	766.006	"	"	"

```
Regress[P1, {1, x, x^2, x^3, x^4}, x]
```

```
ParameterTable →
```

"	"Estimate"	"SE"	"TStat"	"PValue"
1	0.6906	0.0306	22.5012	0.
x	5.2597	0.1620	32.4612	0.
x <sup>2</sup>	-8.440	0.2510	-33.614	1.1609
x <sup>3</sup>	4.2365	0.1474	28.7332	0.
x <sup>4</sup>	-0.695	0.0291	-23.906	2.8050

```
RSquared → 0.8422, AdjustedRSquared → 0.8419,  
EstimatedVariance → 0.0605
```

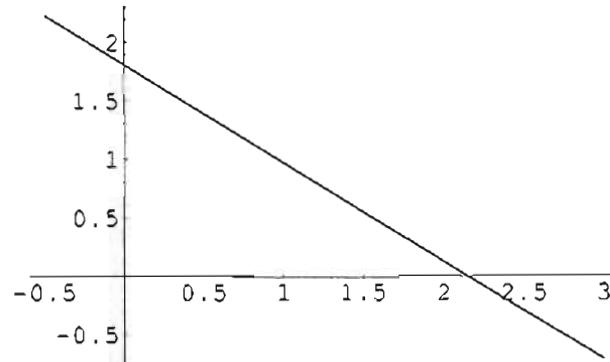
```
ANOVATable →
```

"	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	4	645.152	161.28	2662.4	9.0552
"Error"	1995	120.853	0.0605	"	"
"Total"	1999	766.006	"	"	"

Con los coeficientes de los polinomios que se obtuvieron de las distintas regresiones, nos disponemos a gráficarlos con la función `Plot`, en ésta se introducen los coeficientes de la regresión lineal, cuadrática, etc. Como se

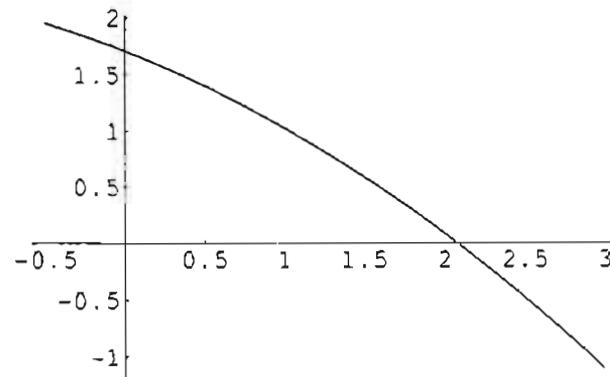
muestra a continuación. Se debe notar que a cada gráfico se le asigna un nombre, de acuerdo al grado del polinomio con el que fué calculado.

```
P1a2p5 = Plot[1.81 - 0.84 x, {x, -0.5, 3}]
```



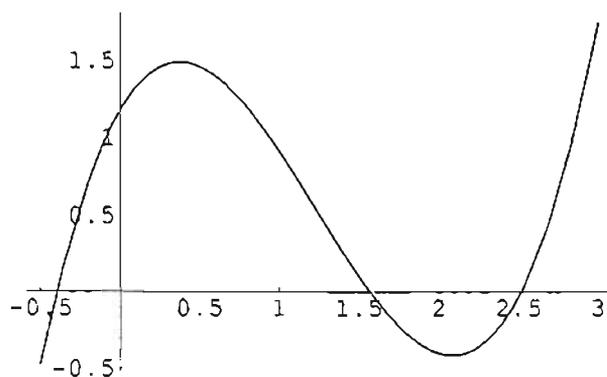
- Graphics -

```
P2a2p5 = Plot[1.71 - 0.55 x - 0.13 x^2, {x, -0.5, 3}]
```



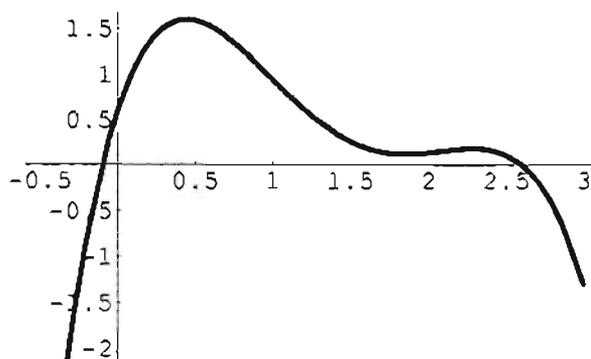
- Graphics -

```
P3a2p5 = Plot[1.18 + 1.75 x - 2.77 x^2 + 0.75 x^3, {x, -0.5, 3}]
```



- Graphics -

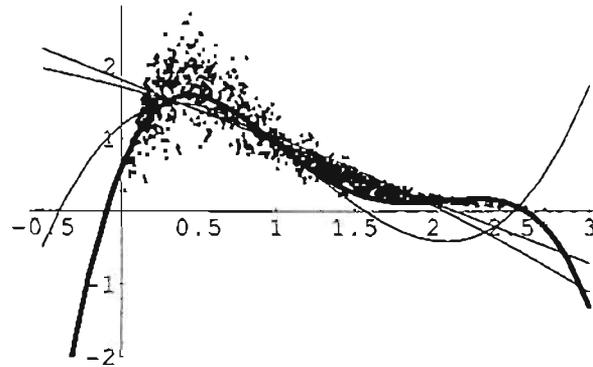
```
P4a2p5 = Plot[0.59 + 5.25 x - 8.44 x^2 + 4.23 x^3 - 0.69 x^4,  
{x, -0.5, 3}, PlotStyle -> Thickness[0.01]]
```



- Graphics -

Una vez que se tienen los gráficos de los polinomios de primero a cuarto grado, así como al de dispersión (P0a2p5), la función `Show` puede mostrar a todos juntos si se enuncian como se muestra a continuación.

```
Poli2p5 = Show[P0a2p5, P1a2p5, P2a2p5, P3a2p5, P4a2p5]
```



- Graphics -

Como se puede observar el polinomio de cuarto grado (el que está re-marcado) es el que mejor se aproxima al conjunto de puntos de la tabla de iteradas del mapeo de Ricker.

## 4.10. Dimensión de correlación

Las instrucciones siguientes sirven para calcular la dimensión de correlación (ver Cap. 4 Secc4.10).

La función **Dis** genera dos vectores  $\bar{x}, \bar{y}$  de tamaño  $n$ , de la forma:  $\bar{x} = \{a_1, a_2, \dots, a_n\}$ ,  $\bar{y} = \{b_1, b_2, \dots, b_n\}$ , y calcula la distancia entre cada pareja de en-eadas, i.e.  $Dis = |a, b|$ .

```
Dis[x_, y_] := Max[Abs[x - y]]
```

La función **CE**, calcula la integral de correlación muestral de la serie de tiempo dada.

```
CE[g_, x_, m_] :=
Module[{l, a, VecRec}, l = Length[x]; VecRec = Partition[x, m, 1];
a =  $\frac{1}{l(l-1)}$  Sum[Sum[UnitStep[g - Dis[VecRec[[i]]], VecRec[[j]]]],
{i, j+1, l-m+1}], {j, 1, l-m}]; N[a]
```

**NestList** genera una lista de iteradas del modelo determinístico, que en

éste caso es el mapeo de Ricker. A continuación se presenta un ejemplo de esta función, los parámetros son:  $\alpha = 3,5$ , el punto inicial es  $x_0 = 0.2$  y se realizan 20 iteraciones. A la lista se le llama **Ric3p5**, para hacer referencia que se trata del mapeo del Ricker y  $\alpha = 3,5$ .

```
Ric3p5 = NestList[# Exp[3.5 (1 - #)] &, .2, 20]

{0.2, 3.28893, 0.00109097, 0.0359902, 1.05077,
 0.879697, 1.34028, 0.407337, 3.24205, 0.00126717,
 0.0417772, 1.19527, 0.603456, 2.41771, 0.0169215, 0.528138,
 2.75418, 0.00593728, 0.192572, 3.25019, 0.00123468}
```

La función **Cor3p5dim1=LogTable**, genera una tabla de iteradas de la integral de correlación con el mapeo de Ricker y con dimensión de inversión 1.

```
Cor3p5dim1 = Log[Table[CE[i / 20, Ric3p5, 1], {i, 1, 20}]]

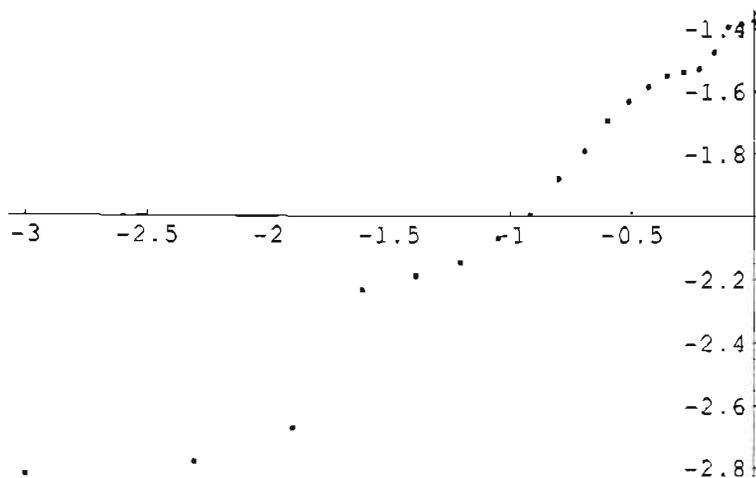
{-2.82138, -2.78216, -2.67296, -2.23359, -2.19011, -2.14843,
 -2.06996, -1.9972, -1.88137, -1.79176, -1.69645, -1.63354, -1.58591,
 -1.55162, -1.54045, -1.5294, -1.47591, -1.39586, -1.38629, -1.37682}
```

La función **TableLog** calcula una tabla de parejas de iteradas  $(x,y)$ ;  $x$  es el logaritmo de  $\frac{i}{150}$ , con  $i = 1, \dots, 150$ , e  $y$  es la integral de correlación  $(\nu)$  con mapeo de Ricker y toma los últimos  $K$  datos de la tabla **Logtable**. En el ejemplo que se presenta se toman los 20 generados. A la tabla le llamamos **lista1**. Nota al correr ésta función el computador tardará algunos minutos.

```
lista1 = Table[{Log[i / 20.], Cor3p5dim1[[i]]}, {i, 1, 20}];
```

Como ya es sabido con la ayuda de la función **LisPlot** graficamos a **lista1**.

```
f1 = ListPlot[lista1]
```



```
- Graphics -
```

Con la Función **Regress** obtenemos la regresión lineal de las parejas (x,y) de la lista1.

```
Regress[lista1, {1, x}, x]
```

```
ParameterTable ->
```

"	"Estimate"	"SE"	"TStat"	"PValue"
1	-1.3750	0.0062	-219.70	5.3300
x	0.54318	0.0046	117.772	0.

```
RSquared -> 0.9894, AdjustedRSquared -> 0.9893,  
EstimatedVariance -> 0.0028
```

```
ANOVATable ->
```

"	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	1	39.2425	39.242	13870	0.
"Error"	148	0.41872	0.0028	"	"
"Total"	149	39.6612	"	"	"

A partir de los datos de la regresión lineal concluimos que la dimensión de correlación para  $\alpha = 3,5$  con dimensión de inversión 1 es: 0.560312. Es decir la dimensión de correlación es precisamente el valor de  $\beta$  de la regresión lineal.

Si se necesitan hacer los mismos calculos pero con diferente dimensión de inversión, habrá que cambiar el el número 1 en la tercera entrada de la función **Cor3p5dim1=LogTable**, tambien es recomendable cambiar los nombres de

las funciones que se corren para evitar confuciones. Por ejemplo para hacer los calculos con dimensión de inversión 2, se deberán correr las funciones como se enlistan a continuación. Los ejemplos ya no se darán.

```
Cor3p5dim2 = Log[Table[CE[i / 150, Ric3p5, 2], {i, 1, 150}]];
li2 = Table[{Log[i / 150.], Cor3p5dim2[[i]]}, {i, 1, 150}];
f2 = ListPlot[lista2]
Regress[lista2, {1, x}, x]
```

## 4.11. Dimensión de correlación con ruido

Para introducir el ruido en la dimensión de correlación es necesario correr las funciones `fRMul` `z[s]`, `Dis`, `CE` y `fRMul`, a ésta última con los parámetros que se necesiten. A la lista generada al correr `fRMul` la llamamos `RicRuido3p5`.

A continuación se da un ejemplo de esta función, los parámetros con los que se calcula son:  $\alpha = 3.5$ ,  $s = 0.05$ ,  $x_0 = 0.2$ , y se calculan 10 iteraciones.

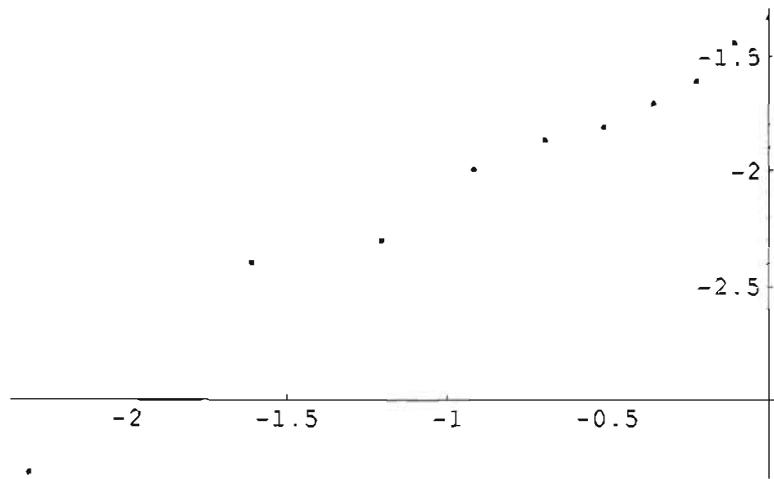
```
RicRuido3p5 = literMul[3.5, .05, .2, 10]
{0.2, 3.29733, 0.00102368, 0.0316891, 0.983323,
 1.07884, 0.840842, 1.42486, 0.345999, 3.4744, 0.000620086}
```

La función `Cor3p5dim1=LogTable`, calcula una tabla de parejas de iteradas  $(x,y)$ ,  $x$  es el logaritmo de  $\frac{i}{150}$ , con  $i = 1, \dots, 150$ , e  $y$  de la integral de correlación con el mapeo de Ricker alterado con Ruido.

```
RCorr3p5dim1 = Log[Table[CE[i / 10, RicRuido3p5, 1], {i, 1, 10}]]
{-3.31419, -2.3979, -2.30259, -1.99243, -1.86727,
 -1.81011, -1.70475, -1.60944, -1.44238, -1.33318}
```

La función `TableLog` toma las últimas parejas de la tabla que se generó con la función `Cor3p5dim1=LogTable`, a ésta lista la llamamos `lis1`. En seguida graficamos las parejas de la lista `lis1`, con la función `LisPlot`.

```
lis1 = Table[{Log[i / 10.], RCorr3p5dim1[{i]}], {i, 1, 10}};
gl = ListPlot[lista1]
```



- Graphics -

Por último realizamos la regresión lineal de las parejas de datos de la lista `lis1`, con la función `Regress`

```
Regress[lis1, {1, x}, x]
```

ParameterTable →

	"Estimate"	"SE"	"TStat"	"PValue"
1	-1.3611	0.0518	-26.240	4.7790
x	0.77801	0.0492	15.8102	2.5611

RSquared → 0.9689, AdjustedRSquared → 0.9651,

EstimatedVariance → 0.0117

ANOVATable →

	DF	SumOfSq	MeanSq	FRatio	PValue
"Model"	1	2.9272	2.9272	249.96	2.5611
"Error"	8	0.0936	0.0117	" "	" "
"Total"	9	3.0208	" "	" "	" "

De la regresión lineal se tiene que la dimensión de correlación para  $\alpha = 3,5$ , con dimensión de inversión 1 y desviación estandar igual a 0.05; es 0,778014.

# Apéndice A

## A.1. Glosario

Sea  $F : I \rightarrow I$  una aplicación continua definida en el intervalo  $I$ . Denotamos con  $F^n$  a la composición de  $F$  con ella misma  $n$  veces, i.e.  $F^n = \overbrace{F \circ \dots \circ F}^{n\text{-veces}}$ . La órbita de  $x_0 \in I$  es el conjunto  $Orb(x_0) = \{x_0, F(x_0), \dots, F^k(x_0), \dots\}$ .

**Definición 2** Sea  $F : I \rightarrow I$ , se dice que un  $x^* \in I$  es periódico si  $f^n(x^*) = x^*$ . La  $n$  más pequeña para la cual  $f^n(x^*) = x^*$  se llama período de  $x^*$ . Un punto periódico  $x^* \in I$  es atractor si existe un intervalo  $J$  alrededor de  $x^* \in I$ , tal que la órbita  $Orb(x)$  tiene una subsucesión convergente a  $x^* \in I$ . Se dice que  $x_*$  es repulsor si existe un intervalo  $J$  alrededor de  $x^*$  y  $k \in \mathbb{N}$ , tal que  $f^i(x_*) \notin J$  para  $i \geq k$ .

Un conjunto  $A$  es *invariante respecto a  $F$* , si  $Orb(x) \subset A$ , para todo  $x \in A$ . Un conjunto invariante cerrado se dice que es *atractor* si existe un abierto  $O$  tal que  $A \subset O$  y  $dist(f^n(x), A)$  tiende a cero, cuando  $n$  tiende a  $\infty$ . La dinámica es *periódica* si el atractor es la cerradura de una órbita periódica. Un atractor es *extraño* si no es la cerradura de una órbita periódica. Para los fines de este trabajo decimos que un sistema dinámico es *caótico* si tiene un atractor extraño. Cuando el sistema es caótico uno observará órbitas que fluctúan de manera errática al acercarse al atractor extraño. Sin intentar ser precisos, entenderemos por *ruido* de una una dinámica a cualquier fluctuación aleatoria de alguno de los parámetros de la aplicación  $F$  que la genera. Si el parámetro que fluctúa aparece como factor de la aplicación el ruido será llamado *multiplicativo*.

En 1981, F. Takens demuestra un teorema fundamental en la teoría de los sistemas dinámicos que da condiciones para que un sistema dinámico a tiempo discreto se pueda reconstruir a partir de mediciones parciales de sus estados. Este teorema se conoce como el Teorema del encaje de Takens.

**Teorema A.1.1** *Sea  $M$  una variedad compacta de dimensión  $m$ , donde están definidos un difeomorfismo  $\phi : M \rightarrow M$  de clase  $\mathcal{C}^k$  y una función real  $h : M \rightarrow \mathbb{R}$ . Entonces (genéricamente) la transformación  $\Phi : M \rightarrow \mathbb{R}^{2m+1}$ , definida por*

$$\Phi(x) = (x, h(\phi(x)), \dots, h(\phi^{2m}(x)))$$

*es un encaje. Aquí  $k \geq 2$ .*

La demostración de este importante teorema se puede ver en [tak]. Para más detalles ver [noa].

# Bibliografía

- [col] Collet, P., Eckman, J.-P.: *Iterated maps on the interval as dynamical systems*. Bostom: Birkhauser 1980.
- [cow] J. M. Cowley, *An approximate theory of order in alloys*, Phys. Rev. **71**,5 (1950), 669–675.
- [cee] C. Diks, *Nonlinear time series analysis*, World Scientific P. (1999).
- [gras] Grassberger, P., Procaccia, I.: *Measuring the strangeness of strange attractors*. Physica **9D**, 189-208 (1983b).
- [kap] D. T. Kaplan and L. Glass, *Direct test for determinism in a time series*, Phys. Rev. L. **68**, 4 (1992), 427– 430.
- [lich] Lichteberg, A. J., Lieberman, M. A.: *Regular and stochastic motion*. New York: Springer 1983.
- [sha] W. M. Schaffer, S. Ellner and M. Kot, *Effects of noise on some dynamical models in ecology*, J. Math Biol. 24 (1986), 479–523.
- [mar] M. Markus, B. Hess, J. Rössler and M. Kiwi, *Populations under periodically and randomly varying growth conditions*, In Chaos in biological systems, eds. H. Degn, A.V. Holden, L.F. Olsen, Plenum Press, NATO ASI Series, (1986), 267–278.
- [mir] P. Miramontes, L. Medrano, C. Cerpa, R. Cedergren, G. Ferbeyre, G. Cocho, *Structural and thermodynamic properties of DNA uncover different evolutionary histories*, J. Mol. Evol. 40 (1995), 698–704
- [noa] L. Noakes, *The Takens embedding theorem*, Int. J. Bif. Ch. **1**(4) (1991), 867–872.
- [tak] Takens, F.: *Detecting strange attractors in turbulence*. In: Rand, D., Young, L.-S. (eds). *Dynamical system and turbulence*. pp. 366-381. War-wich 1980. Berlin Heidelberg New York: Springer 1981.