

# PROPIEDADES DE LOS TOTALES DE SESIÓN BAJO PROGRAMAS DE RAZÓN ALEATORIA: UN ESTUDIO POR SIMULACIÓN

Dr. Jaime Robles  
Dr. Guillermo Yáber

## RESUMEN

Se presenta un caso de relaciones complejas que pueden ser entendidas a partir de las propiedades de los totales de sesión en un programa de razón aleatoria. Se analiza la postergación, en el contexto de la conducta gobernada por reglas, para la cual algunos enfoques conceptuales enfatizan el proceso que describe la ejecución parcial que ocurre a lo largo del período indicado por la regla. En el estudio, el proceso de la postergación es simulado con un modelo simple de aprendizaje por reforzamiento y una regla de respuesta con umbral aleatorio. Los resultados sugieren que un modelo simple de aprendizaje puede reproducir las propiedades de la relación de control observado en el denominado festón de postergación.

**Palabras clave:** *postergación, conducta gobernada por reglas, simulación*

Las propiedades de la respuesta bajo un programa de reforzamiento pueden ser estudiadas usando al menos dos bases de tiempo distintas: el ensayo, lo cual constituye un estudio intra-sesión; y la sesión, lo que se considera un estudio inter-sesiones. Los estudios intra-sesión permiten analizar la estructura fina de las series de respuestas y sus distintas dimensiones, existiendo trabajos que reportan complejos patrones para varias dimensiones de respuesta (Palya, 1992, Davison y Baum, 2000). Los estudios inter-sesiones, más comunes, permiten hacer

consideraciones acerca de la estabilidad de la respuesta y construir modelos que aluden a invariancias de cierto tipo, como es el caso de la teoría del momentum conductual (Nevin, 1988; Nevin, 1992). No obstante, la complejidad que se puede observar en los estudios intra-sesión y las relaciones de invariancia que se encuentran en los patrones entre-sesiones, pueden dar la idea de que la ventana temporal amplia, usada para estos últimos, exhibe una dinámica menos interesante de lo que realmente puede ser.

El presente trabajo intenta mostrar un caso en el cual existe un patrón complejo de los totales de sesión, usando como unidad temporal la sesión. Para lograr la ilustración mencionada se estudia el caso de la postergación, la cual es entendida como una relación de control entre el número de respuestas y un límite temporal (Michael, 1993). La postergación puede involucrar conceptos y modelos de alto orden, como la conducta gobernada por reglas (CGR).

La aproximación de ventana temporal amplia, entendiendo a los totales de sesión como parte de un patrón, permite modelar la relación de control, usando principios elementales del aprendizaje por reforzamiento.

Las secciones que siguen, desarrollan un modelo basado en el operador de aprendizaje por reforzamiento de Bush y Mosteller (Bush y Mosteller, 1951), previamente desarrollado en el contexto de la CGR (Robles y Yáber, 1998; Robles y Yáber, 2000); para posteriormente presentar un estudio por simulación que brinda evidencia acerca de la adecuación del modelo.

## **POSTERGACIÓN Y PROCESOS COMPLEJOS**

Dentro de las distintas aproximaciones a la conducta gobernada por reglas, uno de los fenómenos que reúne varias de las propiedades más importantes de la CGR —que incluye conceptos como el aprendizaje discriminativo, las operaciones de establecimiento o las contingencias indirectas— es el de la postergación, ilustrada a través del festón de pos-

tergación (Michael, 1993, p.114). Este fenómeno de postergación involucra una relación de control, entre la cantidad de respuestas y un límite temporal, que define un criterio de número de respuestas a alcanzar. Esto puede verse como una relación plausible en un programa de razón aleatoria sostenido a lo largo de varias sesiones (Robles y Yáber, 2000). Una aplicación del modelo de postergación se presenta en el ámbito del comportamiento académico (Michael, 1993). A nivel explicativo, el proceso básico puede ser entendido como parte de un mecanismo escape-evitación (EE) (Mowrer, 1956; Malott, Malott y Trojan, 2000). La hipótesis principal del presente trabajo, es que se puede construir un modelo plausible para fenómenos complejos como la postergación, usando sólo principios elementales del aprendizaje por reforzamiento. Si bien nociones como la CGR o el mecanismo EE pueden servir como principios explicativos, lo que se pretende mostrar es que dicho uso no es estrictamente necesario para la construcción de un modelo.

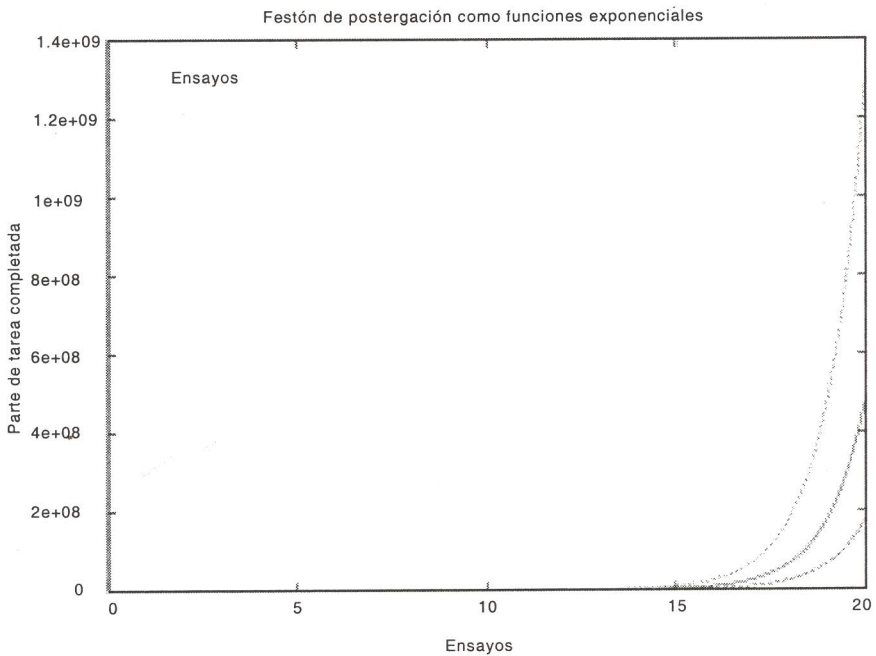


Figura 1. Festón de postergación como función exponencial del tiempo

De esta forma, aunque el proceso puede verse desde distintas perspectivas (Michael, 1982; Cerutti, 1989; Malott, Malott y Trojan, 2000), una de las opciones es aproximarse cuantitativamente al fenómeno. Desde esta aproximación, el presente trabajo tiene como objetivos: (a) Describir paramétricamente la situación del festón de postergación y (b) Probar un modelo matemático para la simulación de este patrón temporal

### DESCRIPCIÓN DEL FESTÓN DE POSTERGACIÓN

Para la descripción cuantitativa de esta relación de control, dentro de una situación concreta, se plantea una tarea acumulativa, que como en el caso de una tarea de estudio acumulativo, en la cual toda la materia hasta una determinada sesión debe ser repasada o cubierta en dicha sesión. Cada punto del eje temporal se considera como un total de respuestas por sesión, siendo los parámetros más importantes:

- $S$  (número de sesiones),
- $K$  (número de ensayos por sesión)
- $N_s$  (total de respuestas por sesión)

#### MODELO

### OPERADOR DE APRENDIZAJE DE BUSH Y MOSTELLER

El modelo de aprendizaje por reforzamiento de Bush y Mosteller de 1955, se basa en un modelo originalmente más simple (Bush y Mosteller, 1951); una versión usada para procesos con conductas competitivas (Davis y cols., 1993), se expresa como:

$$P_{k+1} = aP_k + (1-a)X_k, \quad 0 < a < 1 \quad (1)$$

en donde  $P_k$  puede tomarse como la fuerza de la respuesta en el momento  $k$ ,  $X$  representa la magnitud de estímulo, mientras que  $a$  es un parámetro de persistencia del estado del modelo. Este operador básico ha sido

utilizado para modelar comportamiento de elección en situaciones de dos alternativas, de forma que existan dos alternativas de respuesta para un momento  $k$ , tales como  $P_{ik}$  y  $P_{jk}$ . En el presente trabajo, se plantea el modelo para una sola respuesta, tomando la fuerza de la respuesta competitiva como un umbral oscilante, tal como se describe en la sección siguiente. Este modelo permite una cuantificación en distintas ventanas temporales, una ventana intrasesión, que corresponde al programa de razón aleatoria, y una ventana intersesión en la que se observa la relación de control esperada de acuerdo al festón de postergación. Esta distinción entre ventanas temporales permite hacer detallados análisis intrasesión, e intersesión sin acudir a términos distintos a los de la dinámica del programa de reforzamiento (Palya, 1992). Es de notar que el patrón intersesión es una resultante de los parámetros intrasesión exclusivamente.

### **ESPECIFICACIÓN DEL MODELO PARA EL EXPERIMENTO POR SIMULACIÓN**

El modelo de Bush y Mosteller se usa como cálculo de la probabilidad de respuesta, la cual tiene como referencia un umbral para la determinación de la ocurrencia de la respuesta. En una formulación anterior del proceso de aprendizaje operante (Bush y Mosteller, 1951), se plantea que la salida del modelo ( $P$ ), puede interpretarse como una probabilidad, cuya escala debe ser corregida o compensada por las probabilidades de las respuestas competitivas. En el caso del presente estudio  $P_i$  es tomada como una medida de la probabilidad de la respuesta  $i$ . La regla de respuesta es:

$$R = 1 \text{ w } P > c \quad (2)$$

Es decir, la respuesta ocurre si  $P$  es mayor al umbral  $c$ . En este caso  $c$  representa la probabilidad de ocurrencia de una respuesta competitiva con  $i$ . Esta partición es similar en forma a la clasificación básica de las respuestas en el caso de la ley de igualación (Herrnstein, 1970). Para cada ensayo, el umbral  $c$ , se estima como  $c_k = c + u_k$  en

donde  $c_k$  es el valor particular de  $c$  para el ensayo  $k$ ,  $u$  es el término de oscilación con distribución uniforme ( $M(u)=0$ ,  $S(u)=c/2$ ). De esta manera, la regla de respuesta puede reexpresarse como:

$$R_{ik}=1 \text{ w } P_{ik} > c_k \quad (3)$$

Para un ensayo  $k$ , la respuesta  $i$  ocurre si su probabilidad  $P$  supera el valor del umbral  $c$  en este ensayo. Es importante señalar que  $c$  representa la probabilidad de una respuesta competitiva, siendo entonces  $R_i$  una respuesta que ocurre en competición con otra respuesta. Esto implica que para el caso en que  $R_i=0$ , no quiere decir que la respuesta sea encubierta, sino que ocurre la respuesta competitiva. Esto separa al modelo de la interpretación original de que el umbral distinguía entre respuestas encubiertas y manifiestas, presente en algunas formulaciones del enfoque de las redes neuronales artificiales. El planteamiento de la respuesta en referencia a una respuesta competitiva, coincide con la concepción de la ejecución en programas de reforzamiento simple (Baum, 1993, Herrnstein, 1970).

## MÉTODO

### **SIMULADOR DE APRENDIZAJE POR REFORZAMIENTO**

En la realización del experimento por simulación se usó un programa en C++ orientado a objetos, incluyendo tres componentes: Clase de objetos C++ para el modelo de Bush y Mosteller, Biblioteca de clases C++ para control de parámetros experimentales y Clases de objetos C++ de generadores de números aleatorios (RNG). Los dos primeros componentes han sido desarrollados por el primer autor (Robles, 1997). Se usaron dos tipos de componentes RNG, uno diseñado por Robles (Robles, 1996) y otro incluido en la biblioteca libg++ del compilador GNU C++ de la Free Software Foundation. El programa final para la simulación del festón de postergación, programado por los autores en C++, maneja los siguientes parámetros:

- Términos  $a$ ,  $X$  y  $P_0$  del modelo de Bush y Mosteller
- Número de ensayos para la primera sesión ( $K_0$ )
- Criterio de la razón aleatoria para la primera sesión ( $w_0$ )

## PROCEDIMIENTO

En la primera sesión  $s=0$ , se calcula  $P$  para cada ensayo  $k$  ( $P_k$ ), para un número predeterminado de ensayos ( $K_0$ ). Para el resto de las sesiones, es decir desde  $s = 1$  hasta  $s = S$ , el número de ensayos queda determinado por la ejecución simulada por el modelo. En las siguientes sesiones se simulan tantos ensayos como sean necesarios para obtener el mismo número de refuerzos que en la primera sesión. De esta forma, a mayor  $w_0$  menor aceleración del festón, ya que hay menor densidad de reforzamiento (también llamada tasa de refuerzo) para la conducta controlada por la regla. Este parámetro permite especificar propiedades importantes de la respuesta en cuanto a su variabilidad a lo largo del tiempo (Nevin, 1992; Cohen, 1998). Para simular el festón, a partir de la sesión  $s = 1$  hasta la última sesión  $s = S$ , se incrementa la razón del programa de reforzamiento, cumpliendo con la relación general que simula el proceso de CGR:  $N_s = f(s)$ , lo que se logra a través de la relación entre la razón y la respuesta, de tal modo que al incrementar la razón se incrementa el número de respuestas por sesión ( $N_s$ ). De esta forma, la simulación se realiza incrementando la razón para cada sesión ( $w_s$ ), teniendo entonces la ecuación básica de la manipulación experimental  $w_s = g(s)$ , en donde  $g(x)$  es una función apropiada para producir el incremento en  $N_s$ , que se ajuste a la relación esperada como es el festón de postergación.

En el experimento realizado, la función  $g(x)$  correspondió a  $\log(x)$ , y el resto de los parámetros:  $K_0=50$ ,  $a=0.65$ ,  $X=1$ ,  $P_0=0.50$ ,  $c=0.25$ ; siendo  $w_0=5$  para un caso y  $w_0=10$  para el otro. Una segunda simulación se realizó con  $c=0.35$ .

RESULTADOS Y DISCUSIÓN

Los resultados indican que el incremento logarítmico en la razón, producen una función ascendente en las respuestas acumuladas por sesión, la cual es bastante similar al planteamiento del festón de postergación. Por otra parte, un retardo en el incremento de la razón, produce una curva de ejecución, que al igual que lo expresado por Michael, no alcanza el mismo total de tarea completada en el límite de tiempo establecido por la regla. (Figura 2)

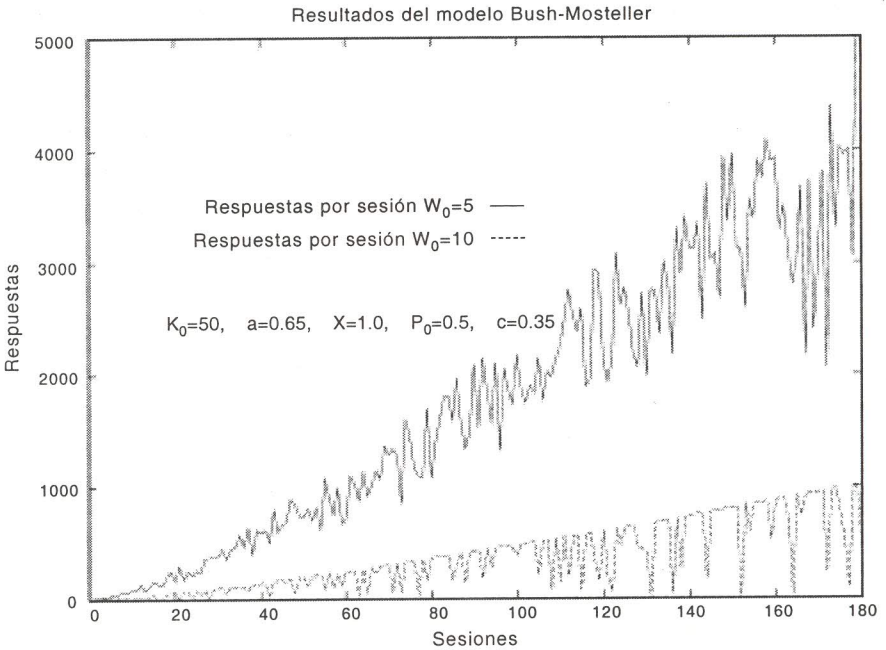


Figura 2. Resultados de la primera simulación, con parámetros:  $K_0=50, a=0.65, X=1, P_0=0.50, c=0.25$



El mecanismo (EE), una de las posibles explicaciones al festón, no es introducido en forma explícita en el modelo, por lo que la simulación permite generar el patrón de ejecución a nivel molar, sin introducir suposiciones o términos adicionales a los del proceso operante expresado en el modelo de Bush y Mosteller (Figura 3)

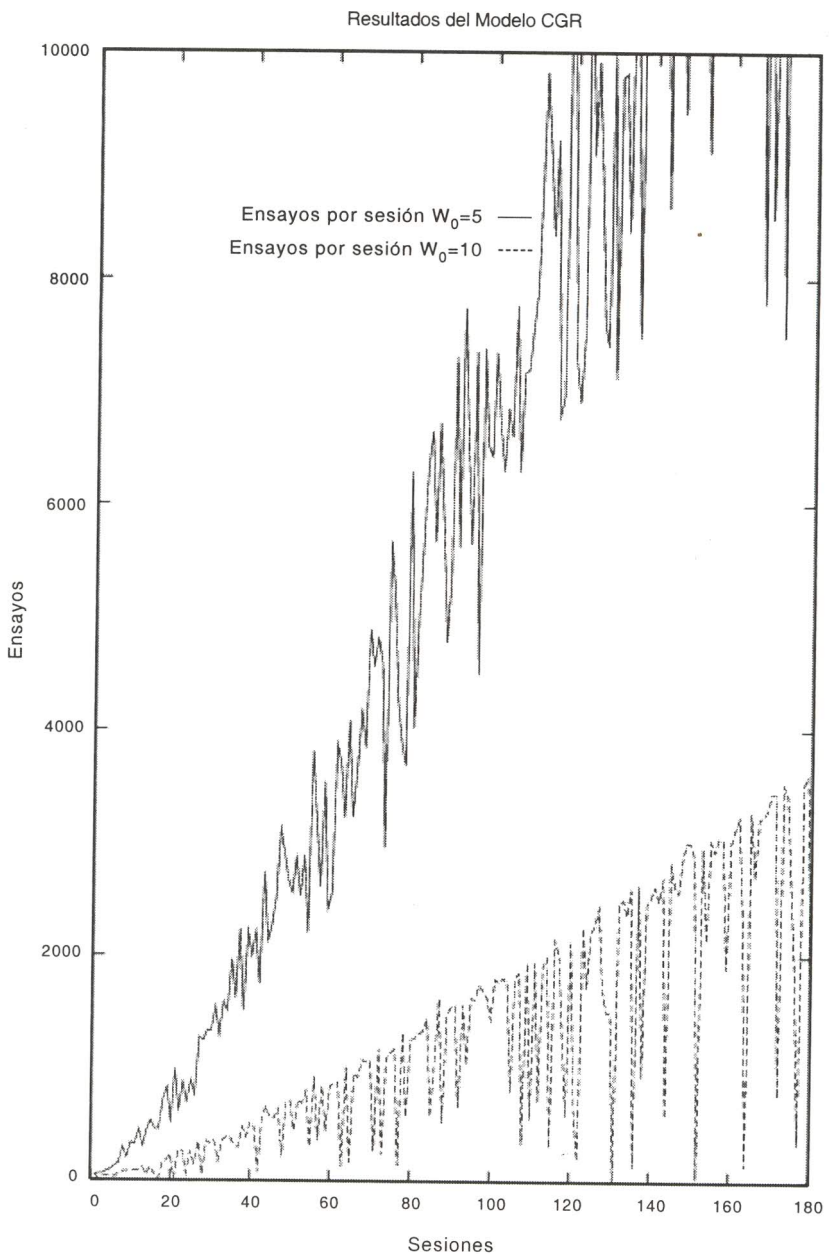
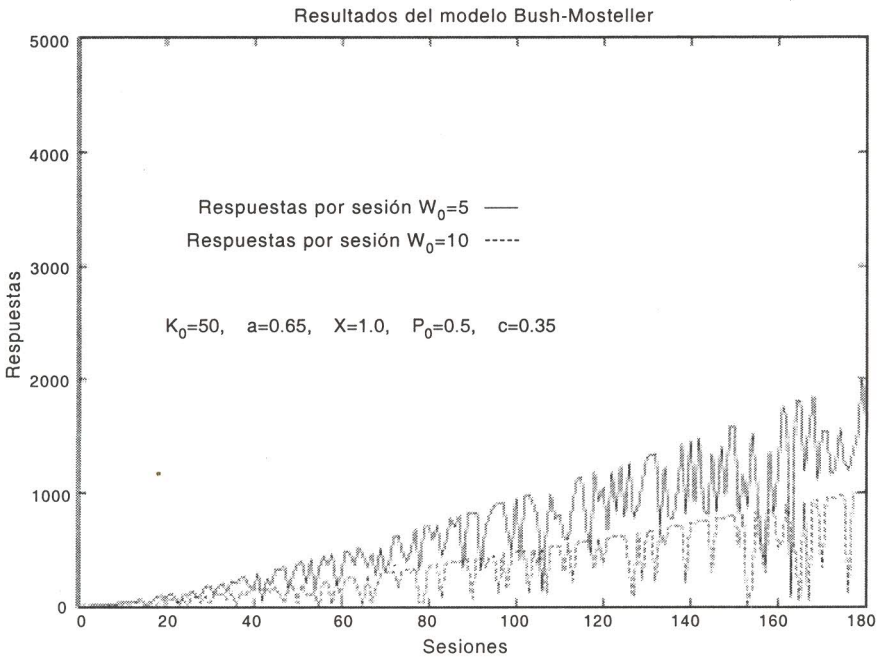


Figura 3. Número de ensayos por sesión, con parámetros:  $K_0=50$  ,  $a=0.65$  ,  $X=1$  ,  $P_0=0.50$  ,  $c=0.25$

El mecanismo complejo EE y la relación con las operaciones de establecimiento (Mowrer, 1956; Michael, 1993; Malott, Malott y Trojan, 2000), son manejadas como principios explicativos para conceptualizar la situación y fijar parámetros como el número de sesiones y el procedimiento de fijación del número de reforzamientos. En este sentido se incorporan en el modelo general. No obstante, los resultados indican que el proceso básico del festón de postergación puede generarse usando relaciones de control basadas en programas de reforzamiento, simulado la ejecución sin introducción de eventos internos o conceptos distintos al del control de las respuestas por el programa de reforzamiento. Uno de los determinantes principales de la aceleración de la curva de ejecución  $w_0$ , se relaciona en realidad con la densidad de reforzamiento, lo que especifica paramétricamente la idea de un comportamiento más previsorio ( $w_0$  más baja, mayor densidad de reforzamiento, mayor aceleración). Estos resultados pueden examinarse en las figuras 4 y 5.



*Figura 4.* Resultados de la segunda simulación, con parámetros:  $K_0=50$ ,  $a=0.65$ ,  $X=1$ ,  $P_0=0.50$ ,  $c=0.35$

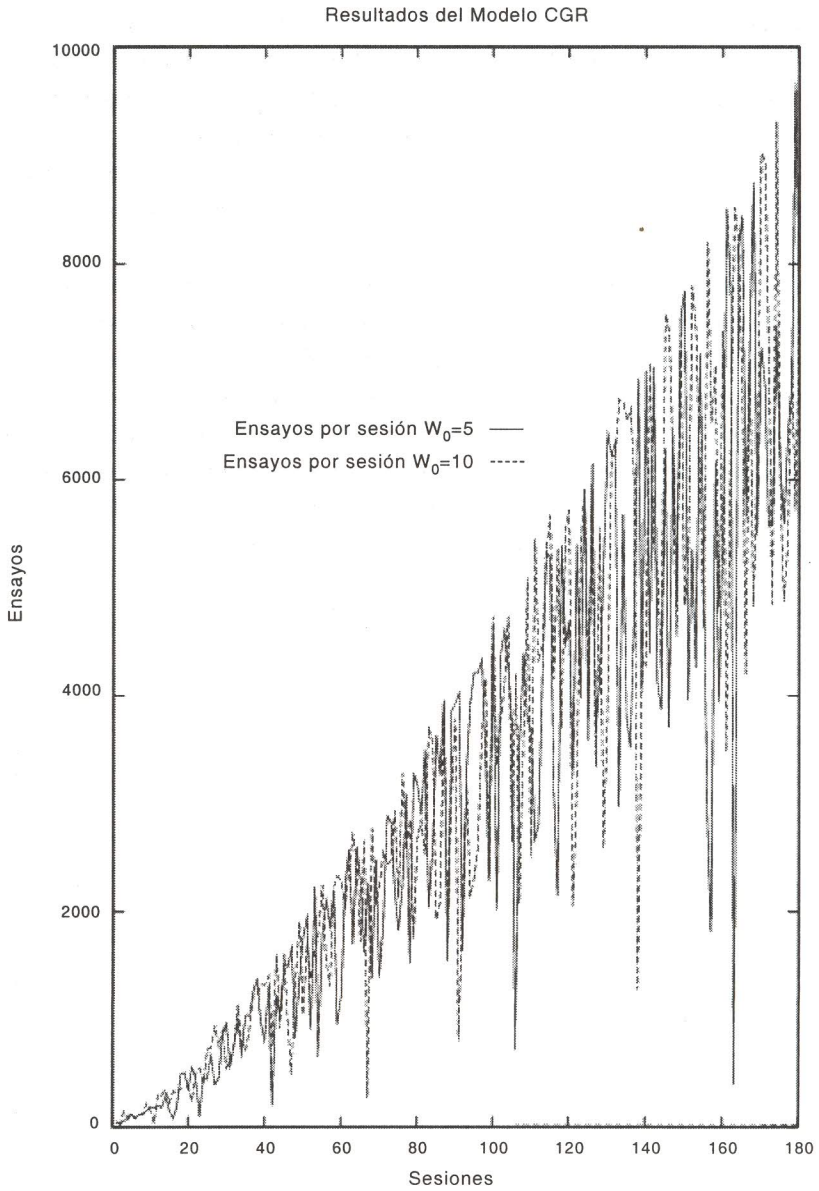


Figura 5. Número de ensayos por sesión, con parámetros:  $K_0=50$ ,  $a=0.65$ ,  $X=1$ ,  $P_0=0.50$ ,  $c=0.35$

El patrón de ejecución muestra un esperado aumento de la variabilidad con el aumento de  $w$  a lo largo de los ensayos. No obstante esta variabilidad, se cumple siempre la diferencia entre el patrón de alta densidad de reforzamiento ( $w_0 = 5$ ) y el de baja densidad ( $w_0 = 10$ ), observándose la propiedad citada por Michael de que el comportamiento más previsor, con mayor aceleración y un inicio de pendiente más temprano, alcanza una mayor magnitud de tarea completada al momento del límite de tiempo establecido por la regla. Esta diferencia se hace más pequeña al aumentar la probabilidad de la conducta competitiva ( $c = 0.35$ ), cumpliéndose también que el número de respuestas alcanzado es menor en términos absolutos. Aun tomando en cuenta que se requiere una exploración más completa del espacio paramétrico; el enfoque cuantitativo propuesto, aunque basado en principios más simples, coincide en sus resultados con otros enfoques de las propiedades cuantitativas del reforzamiento (Killeen y Bizo, 1998). Una relación de control compleja, como la postergación, que puede llevar a elaboraciones conceptuales de otra índole, como interpretaciones cognitivas o de comportamiento social. Con menos separación de los datos, puede involucrar discusiones teóricas sobre la CGR, pero aun con menos compromisos conceptuales, puede ser entendida como una resultante de las propiedades de los totales de sesión, como un patrón que describen los datos, que es apreciable cuando se usa la ventana temporal amplia.

## **CONCLUSIONES**

Es posible reproducir propiedades complejas de un fenómeno como la postergación usando un modelo simple de aprendizaje por reforzamiento. Las relaciones de control entre los límites temporales y el número de respuestas, son modeladas adecuadamente por medio de un operador tipo Bush-Mosteller y una regla simple de emisión de respuesta. Los totales de sesión de un programa de razón aleatoria, pueden ser usados para la construcción de patrones complejos, que permitan el modelamiento de fenómenos normalmente asociados con

relaciones complejas como la de la conducta gobernada por reglas. El patrón complejo de totales de sesión puede construirse sin el uso de nociones más allá de los principios elementales del aprendizaje por reforzamiento, lo cual apoya la idea de que dichos principios siguen siendo una buena base para una teoría de la conducta (Davison y Nevin, 1999).

## REFERENCIAS

- Baum, W. M. (1993). Performances on ratio and interval schedules of reinforcement: data and theory. *Journal of the Experimental Analysis of Behavior*, 59, 245-264.
- Bush, R. R. y Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, 58, 313-323.
- Cerutti, D. T. (1989). Discrimination theory of rule-governed behavior. *Journal of the Experimental Analysis of Behavior*, 51, 259-276.
- Cohen, S. L. (1998). Behavioral momentum: the effects of temporal separations of rates of reinforcement. *Journal of the Experimental Analysis of Behavior*, 69, 29-47.
- Davis, D. G. S., Staddon, J. E. R., Machado, A. y Palmer, R. G. (1993). The process of recurrent choice. *Psychological Review*, 100, 320-341.
- Davison, M. y Baum, W. M. (2000). Choice in a variable environment: every reinforcer counts. *Journal of the Experimental Analysis of Behavior*, 74, 1-24.
- Davison, M. y Nevin, J. A. (1999). Stimuli, reinforcers, and behavior: an integration. *Journal of the Experimental Analysis of Behavior*, 71, 439-482.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Killeen, P. R. y Bizo, L. A. (1998). The mechanics of reinforcement. *Psychonomic Bulletin & Review*, 5, 221-238.

- Malott, R. W., Malott, M. y Trojan, E. (4 ed.). (2000). *Elementary principles of behavior*. New Jersey: Prentice Hall.
- Michael, J. (1982). Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior*, 37, 149–155.
- Michael, J. L. (1993). *Concepts and principles of behavior analysis*. Society for the Advancement of Behavior Analysis, Kalamazoo.
- Mowrer, O. H. (1956). Two-factor learning theory reconsidered, with special reference to secondary reinforcement and the concept of habit. *Psychological Review*, 63, 114–128.
- Nevin, J. A. (1988). Behavioral momentum and the partial reinforcement effect. *Psychological Bulletin*, 103, 44–56.
- Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, 57, 301–316.
- Palya, W. L. (1992). Dynamics in the fine structure of schedule-controlled behavior. *Journal of the Experimental Analysis of Behavior*, 57, 267–278.
- Robles, J. y Yáber, G. (1998, Noviembre). *Un modelo matemático aplicado a la conducta gobernada por reglas*. Sesión de cartel presentado en el Congreso Internacional de Conductismo y Ciencias de la Conducta, Sevilla, España.
- Robles, J. y Yáber, G. (2000, Mayo). *A quantitative approach to rule-governed behavior*. Sesión de cartel presentado en la convención anual de la Asociación para el Análisis Conductual, Washington, DC, USA
- Robles, J. (1996). Prs: Polytomous response simulator-polytomous item generation according to the common factor model. *Applied Psychological Measurement*, 20, 140.
- Robles, J. (1997). *Biblioteca de objetos C++ para experimentos psicológicos*. Laboratorio de Psicología Experimental, UCAB, Caracas.