

**EL MÉTODO DE REGENERACIÓN PARA
LA ESTIMACIÓN DE LA ESTABILIDAD
DE PROCESOS CONTROLABLES CON
COSTO PROMEDIO**

Tesis que presenta:

✓ Guillermo Pablo García Martínez

✓ Para obtener el grado de Maestro en Matemáticas

Director: Dr. Evgueni I. Gordienko

Universidad Autónoma Metropolitana - Iztapalapa
✓ División de Ciencias Básicas e Ingeniería
Departamento de Matemáticas

Marzo de 1998

A mi abuelita:

Rosaura

A mis padres

A mis hermanos

A mis compañeros

A los niños:

Janet

Kevin

Jenifer

Agradesco al Consejo Nacional de Ciencia y Tecnología por el apoyo brindado para realizar mis estudios de maestría, y al departamento de matemáticas de la UAM-I.

También quiero agradecer, de manera especial al Dr. Evgueni I. Gordienko por aceptar dirigir esta tesis y al Dr. Raúl Montes de Oca Machorro por su ayuda en la revisión del manuscrito de esta tesis.

CONTENIDO

Introducción.....	1
1. Modelo de Control de Markov a Tiempo Discreto con Costo Promedio	
1.1. Definiciones de los Elementos del Modelo.....	9
1.2. Políticas de Control.....	12
1.3. Índice de Funcionamiento - Costo Promedio.....	14
2. Planteamiento del Problema de la Estabilidad con respecto al Costo Promedio.....	17
3. Hipótesis acerca de la clase de procesos de control considerados.....	23
4. Resultados Principales.....	27
5. Demostraciones de los Resultados	
5.1 Demostración del Teorema 1.....	29
5.2 Demostración del Teorema 2.....	38
6. Un ejemplo de un Sistema de Espera con tasa de servicio controlable.....	47
7. Conclusión y Algunos Problemas Abiertos.....	61
Bibliografía	

INTRODUCCIÓN

Esta tesis trata el problema de estabilidad (o robustez en otra terminología) de los procesos de control markovianos a tiempo discreto. La estabilidad se entiende aquí, como, poca sensibilidad del funcionamiento óptimo del sistema de control, con respecto a perturbaciones de los parámetros o distribuciones de los procesos que describen dicho sistema. Gran parte de la literatura (Ver referencias en [5,15,16,24]) sobre procesos de control de Markov, estudia el problema de control óptimo, bajo la suposición de que todas las características de los procesos están dadas precisamente. Sin embargo, ésta es una hipótesis poco realista en la mayoría de los problemas de control aplicados, donde, como regla, hay incertidumbre en la información de entrada. Los orígenes de tal incertidumbre pueden ser: al considerar aproximaciones teóricas del proceso real, o al tener estimaciones de algunos parámetros, o distribuciones desconocidas en el sistema. Esta situación es típica, en particular, en problemas de control adaptado (Ver por ejemplo, [11,15,17]).

Lo que hemos dicho anteriormente resulta que en muchos casos no es posible encontrar una política π_* de control óptimo (en costo promedio, como se especifica en el presente trabajo) para el "proceso real" \mathcal{P} , ya que π_* depende de valores **no**

conocidos exactamente. Pero como regla, se podría usar para optimizar un proceso de control aproximante \tilde{P} que es completamente conocido y que está a disposición del controlador. Una manera más sencilla de aproximar π_* es hallar la política $\tilde{\pi}_*$, óptima en \tilde{P} , y aplicar esta política para controlar el sistema (o proceso) original \mathcal{P} . La respuesta a la pregunta : "¿Es $\tilde{\pi}_*$ una buena aproximación para π_* ?", depende del aumento del costo adicional debido al reemplazamiento de la política óptima π_* por su aproximación $\tilde{\pi}_*$.

Como el índice de funcionamiento de los procesos de control definidos más adelante en (I.1), (I.2), se usa en este trabajo, el costo promedio por etapa en el problema de control en horizonte infinito. Denotaremos los costos promedios, como funciones de la política aplicada π , por $J(\pi)$ y $\tilde{J}(\pi)$, respectivamente, para los procesos (o sistemas) \mathcal{P} y \tilde{P} (cuyas definiciones están dadas en los siguientes capítulos, ver (1.5), (2.5)). El índice de costo promedio es adecuado para evaluar el funcionamiento de sistemas que realizan muchas transiciones durante el período de funcionamiento del sistema. Por ejemplo, podríamos apuntar algunos sistemas de espera controlables en redes de comunicación (ver [6, 27]).

Trabajamos en esta tesis, con sistemas de control presentados por procesos de Markov controlables, a tiempo discreto, $t = 0, 1, 2, \dots$ que son definidos por ecuaciones recurrentes, como sigue:

Sea \mathcal{P} un "proceso real":

$$x_{t+1} = \Phi(x_t, a_t, \xi_t), \quad t = 0, 1, 2, \dots, \quad (\text{I.1})$$

y $\tilde{\mathcal{P}}$ un "proceso aproximante":

$$\tilde{x}_{t+1} = \Phi(\tilde{x}_t, \tilde{a}_t, \tilde{\xi}_t), \quad t = 0, 1, 2, \dots, \quad (\text{I.2})$$

donde en el momento t , x_t y \tilde{x}_t son estados, y, a_t y \tilde{a}_t son controles (o acciones) seleccionados, mientras que, ξ_t y $\tilde{\xi}_t$, son vectores aleatorios (V.A'S) independientes e idénticamente distribuidas (i.i.d). Ambos procesos (I.1) y (I.2) están definidos en el mismo espacio de estados X y de controles A , ambos conjuntos borelianos, además, el costo por una etapa se da por una función acotada $c(x, a)$, $x \in X$, $a \in A$, común para ambos procesos \mathcal{P} y $\tilde{\mathcal{P}}$. La única diferencia entre (I.1) y (I.2) son las funciones de distribución F y \tilde{F} , respectivamente de los V.A'S i.i.d. ξ_t y $\tilde{\xi}_t \in R^n$.

Para las políticas de control π_* y $\tilde{\pi}_*$, óptimas en costo promedio, definimos respectivamente para \mathcal{P} y $\tilde{\mathcal{P}}$, los **costos promedios óptimos** como:

$$J(\pi_*) = \inf_{\pi} J(\pi) ;$$

$$\tilde{J}(\tilde{\pi}_*) = \inf_{\pi} \tilde{J}(\pi) .$$

Introducimos el **índice de estabilidad promedio** Δ_p , siguiendo a [8,9,10], como:

$$\Delta_p := J(\tilde{\pi}_*) - J(\pi_*) \geq 0 . \quad (\text{I.3})$$

El valor de Δ_p denota el aumento (en comparación con el valor óptimo $J(\pi_*)$) del costo promedio debido al uso de la política $\tilde{\pi}_*$, en lugar de la política óptima π_* .

Notamos que bajo las hipótesis de este trabajo, mostraremos la existencia de políticas óptimas estacionarias π_* y $\tilde{\pi}_*$ (y también, el hecho de que los costos promedios $J(\pi)$, $\tilde{J}(\pi)$ no dependen del estado inicial del proceso para políticas π , estacionarias óptimas en costo promedio).

Decimos que el proceso de control \mathcal{P} es **estable en costo promedio** con respecto a una métrica probabilista ν si

$$\Delta_p \rightarrow 0 \quad \text{cuando} \quad \nu(F, \tilde{F}) \rightarrow 0. \quad (\text{I.4})$$

Ejemplos de procesos del tipo (I.1), **no estables** se presentan en [8,9,10], donde se usa $\nu = \sigma$ una métrica tan fuerte como la métrica de variación total (ver [25]).

En los ejemplos mencionados, los procesos del tipo (I.1) no poseen propiedades de ergodicidad (aplicando las políticas de control estacionarias). El objetivo de esta tesis, es proporcionar condiciones acerca de las propiedades de ergodicidad de los procesos (I.1), (I.2), que garanticen la existencia de políticas óptimas estacionarias, y, lo más importante, es que permitan establecer la estabilidad de (I.4) en una forma cuantitativa, a saber, obtener una desigualdad como:

$$\Delta_p \leq \varphi(\sigma(F, \tilde{F})). \quad (\text{I.5})$$

Donde $\varphi : [0, \infty) \rightarrow [0, \infty)$ es una función dada con $\lim_{\omega \rightarrow 0} \varphi(\omega) = 0$. Es claro, que desde el punto de vista de las aplicaciones, (I.5) es mucho más útil, que la afirmación pura de estabilidad como en (I.4).

Al explotar algunas hipótesis de ergodicidad en la forma de "condiciones de deriva"

y condiciones de recurrencia, reduciremos el problema de obtener (I.5), al problema de comparación uniforme en tiempo de procesos regenerativos no controlados. Luego usando métodos y resultados desarrollados para la estimación de la estabilidad de procesos regenerativos (ver [18,19,20]), establecemos una desigualdad de estabilidad para procesos controlables:

$$\Delta_p \leq N \left[\sigma(F, \tilde{F}) \right]^{\frac{s-1}{s}}, \quad (\text{I.6})$$

donde, σ es la distancia de la variación total; $s > 2$ es una constante involucrada en nuestras hipótesis y N es una constante la cual es posible estimar mediante cantidades que aparecen en nuestras hipótesis.

Notamos que en los trabajos [9,10,14] se trata el mismo problema, y por métodos distintos a los de esta tesis, se obtuvo la siguiente cota superior:

$$\Delta_p \leq N_1 \sigma(F, \tilde{F}) \max \left\{ 1, \ln \left(\frac{1}{\sigma(F, \tilde{F})} \right) \right\}. \quad (\text{I.7})$$

Es claro que (I.7) tiene mejor orden de $\sigma(F, \tilde{F})$ en la parte derecha en comparación con (I.6). Sin embargo, la desigualdad (I.6) es aplicable a una clase de procesos más amplia, ya que para deducir (I.7), se usaron condiciones de ergodicidad, que garantizan la convergencia geométrica de las distribuciones de procesos del tipo (I.1), a las distribuciones invariantes, si se usan políticas estacionarias. Por otro lado, vamos a ver en el capítulo 5, que nuestras hipótesis (bajo las cuales se establece (I.6)), corresponden a la convergencia de las distribuciones con rapidez de un orden de potencia $(s - 1)$. Nuestra conjetura es que bajo las hipótesis de este trabajo el

orden de $\sigma(F, \tilde{F})$ en la parte derecha de (I.6) no podría, en general, ser mejor.

Sin dejar de considerar los sistemas de control lineales (ver por ejemplo, [22]), los trabajos en los cuales se estudia la estabilidad (en el sentido aquí usado) son escasos. Podríamos apuntar (aparte de los mencionados [8, 9, 10, 14]), el artículo [1] que trata procesos en espacios finitos (sin estimaciones cuantitativas); los artículos [3, 4, 28], donde, han sido establecidas algunas desigualdades de estabilidad, para modelos de control de Markov con espacios de estados numerables; (nótese que aquí, consideramos, los procesos en espacios generales de Borel). Desafortunadamente en [3, 4, 28], las desigualdades obtenidas, no se expresan en cantidades involucradas directamente con las definiciones de los procesos y con las hipótesis usadas, sino que dichas desigualdades contienen "cantidades derivadas", que se calculan difícilmente, para sistemas de control concretos. Las mencionadas "cantidades derivadas", son cotas para el costo descontado óptimo, uniformes con respecto al factor de descuento $\beta \in (0, 1)$, o cotas superiores, para los números esperados de transiciones para pasar de un estado a otro (ver [3, 4]).

Las aplicaciones potenciales de la desigualdad (I.6) son estimaciones de la estabilidad, en términos de la precisión de aproximación de la distribución F por \tilde{F} , para procesos que satisfacen las hipótesis del capítulo 3. Estos podrían ser, por ejemplo, procesos de control de inventario, de producción (ver [14, 16]), de sistemas de espera controlables (ver [6, 13, 27]), de control de suministro de agua en presas (ver [29]), etc. Nos concentraremos en el capítulo 6, en un ejemplo, de un sistema de espera con servicio controlable, para el cual verificaremos todas nuestras hipótesis (usando,

condiciones sencillas, de la existencia de momentos de orden s , del tiempo de servicio e intervalos de llegada).

El resto del trabajo está organizado de la siguiente manera: En el capítulo 1, daremos una descripción de los elementos del modelo de control de Markov con costo promedio. En el capítulo 2, plantearemos el problema de estimación de la estabilidad en términos precisos. Luego, en el capítulo 3, se introducen las hipótesis, acerca de la clase de procesos de control considerados. En el capítulo 4, formularemos los resultados principales, a saber, el teorema sobre existencia de políticas óptimas estacionarias y la desigualdad (I.6). En el capítulo 5, se dan las demostraciones, de las afirmaciones del capítulo 4, mientras en el capítulo 6, estudiaremos un ejemplo particular de la tasa de servicio en un sistema de espera del tipo $GI | GI | 1 | \infty$. Concluimos con una discusión de algunos problemas abiertos en el último capítulo 7.

CAPÍTULO 1

MODELO DE CONTROL DE MARKOV A TIEMPO DISCRETO CON COSTO PROMEDIO

1.1 Definiciones de los elementos del modelo.

En el próximo capítulo, vamos a plantear el problema de la estabilidad, para procesos controlables a tiempo discreto, que se inducen, por un modelo de control de Markov general, con espacio de estados y de controles, ambos conjuntos Borelianos.

Según las definiciones estándar (ver por ejemplo, [5, 15, 16]), un **modelo de control de Markov** se compone de cuatro objetos (X, A, p, c) donde:

(i) X es un espacio de Borel (i.e. un subconjunto medible de un espacio métrico,

separable y completo) llamado **espacio de estados**. Los elementos $x \in X$, se les llama **estados**.

(ii) A es un espacio de Borel llamado el **conjunto de controles**. Los elementos $a \in A$, se les llama **controles** (o **acciones**). Para cada $x \in X$, asociamos, un subconjunto medible no vacío $A(x)$, de A , cuyos elementos son los **controles admisibles**.

Supondremos que el conjunto

$$K := \{(x, a) : x \in X, a \in A(x)\},$$

de las parejas estado-control admisibles, es un subconjunto medible, del espacio producto $X \times A$.

(iii) p representa la ley de evolución del sistema y es una **probabilidad de transición** sobre X dado K .

Por definición, p satisface las siguientes propiedades :

- (a) $p(\cdot | x, a)$ es una distribución de probabilidad en X para cada $(x, a) \in K$.
- (b) $p(B | \cdot, \cdot)$ es una función medible en K para cada conjunto de Borel $B \in \mathcal{B}(X)$.

Nota 1.1 En esta tesis, entendemos por **medibilidad**, medibilidad con respecto a una σ -álgebra de Borel (i.e. la σ -álgebra mínima que contiene todos los subconjuntos abiertos), las cuales denotaremos como $\mathcal{B}(X)$, $\mathcal{B}(A)$, $\mathcal{B}(Y)$, etc., para espacios de Borel X , A , Y , etc.

(iv) c es una función medible en K que representa la **función de costo por etapa**.

Lo anterior es interpretado, como la representación de un proceso estocástico de Markov controlado, el cual se observa en los tiempos $t = 0, 1, \dots$. El estado y control

al tiempo t se denotan por x_t y a_t respectivamente, y, el proceso evoluciona como sigue. Si el proceso está en un estado $x_t = x \in X$ en el tiempo t y se escoge el control $a_t = a \in A(x)$, entonces, se genera un costo $c(x, a)$, y, el proceso pasa en el tiempo $t + 1$ a un nuevo estado x_{t+1} , de acuerdo a la distribución de probabilidad $p(\cdot | x, a)$, a saber, $P(x_{t+1} \in B | x_t = x, a_t = a) = p(B | x, a)$ para cada $B \in \mathcal{B}(X)$. Una vez que la transición ha ocurrido a un nuevo estado digamos $x_{t+1} = x' \in X$, se escoge un nuevo control $a_{t+1} = a' \in A(x')$, lo cual genera el costo $c(x', a')$, y, se repite el proceso anterior un número infinito de veces (ya que consideramos, el problema de control con horizonte infinito).

Un caso especial, pero suficientemente general, para definir el modelo de control descrito anteriormente, consiste en definir directamente una probabilidad de transición p mediante las siguientes ecuaciones recurrentes:

$$x_{t+1} = \Phi(x_t, a_t, \xi_t), \quad t = 0, 1, \dots \quad (1.1)$$

Donde ξ_0, ξ_1, \dots son vectores aleatorios (V.A'S) i.i.d. con valores en un espacio euclidiano R^n , con una función de distribución común denotado por F ; mientras $\Phi : K \times R^n \rightarrow X$ es una función medible dada. Las ecuaciones (1.1) describen la evolución del proceso en una manera directa. Por otro lado, la probabilidad de transición para el proceso (1.1) se determina por las siguientes relaciones:

$$p(B | x, a) = P(\Phi(x_t, a_t, \xi_t) \in B | x_t = x, a_t = a) = \int_{R^n} I_B[\Phi(x, a, w)] dF(w), \quad (1.2)$$

$t = 0, 1, \dots$; donde I_B es la indicadora del conjunto B .

En esta tesis, vamos a considerar sólo modelos de control dados por ecuaciones recurrentes del tipo (1.1).

1.2 Políticas de control

Primero definimos para $t = 0, 1, \dots$, los espacios de las t -historias como sigue:

$$H_0 := X ; \quad H_t := K^t \times X , t = 1, 2, \dots$$

Para cada t , el conjunto H_t contiene todas las formas posibles en que el proceso de control puede evolucionar hasta el tiempo t . Por esta razón llamaremos un punto de H_t que es un vector de la forma:

$$h_t = (x_0, a_0, x_1, a_1, \dots, x_{t-1}, a_{t-1}, x_t)$$

donde $a_t \in A(x_t)$, como una t -historia.

En términos generales, una **política de control** o simplemente **política** (ver, definición 1.1) es una sucesión $\pi = \{\pi_t\}$, donde para cada t , $\pi_t = \pi_t(h_t)$ es una regla para escoger un control $a_t \in A(x_t)$, si $h_t = (x_0, a_0, x_1, a_1, \dots, x_{t-1}, a_{t-1}, x_t)$. Además, si una política depende sólo del último estado x_t entonces decimos que esta política es Markoviana.

Definición 1.1 (ver [5, 15, 16, 24]) Una **política de control** es una sucesión $\pi = \{\pi_t\}$ tal que para cada t , π_t es una probabilidad transición sobre A dado H_t tal que $\pi_t(A(x_t) | h_t) = 1$ para toda historia $h_t \in H_t$.

El conjunto de todas las políticas lo denotaremos por Π . En muchos casos el problema de control óptimo para los procesos (1.1) es posible restringirse con una subclase particular de Π -de las políticas que se llaman **políticas estacionarias** las cuales son políticas Markovianas, y, además, deterministas, de manera precisa:

Definición 1.2 (ver [5, 15, 16, 24]) Se dice que una política es **estacionaria** si existe una función medible $f : X \rightarrow A$ con $f(x) \in A(x)$ para cada $x \in X$ (llamada selector) tal que para $t = 0, 1, \dots, \pi_t(h_t) = f(x_t)$ si $h_t = (x_0, a_0, x_1, a_1, \dots, x_{t-1}, a_{t-1}, x_t)$.

En otras palabras, en el caso de una política estacionaria la medida $\pi_t(h_t)$ (definición 1.1) está concentrada en el único punto $a_t = f(x_t)$ y por lo tanto, si el proceso en el momento t ocurre en el estado x_t , entonces independientemente de la historia antes de t , la política estacionaria dada por f ordena elegir el control $a_t = f(x_t) \in A(x_t)$. Para simplificar notaciones, en lo que sigue vamos a denotar la política estacionaria dado por la función f por este mismo símbolo f . Denotaremos por S la clase de todas las políticas estacionarias. Es evidente que $S \subset \Pi$.

Del teorema de Ionescu-Tulcea [2] para cada política $\pi \in \Pi$ y x un estado inicial del proceso (i.e $x_0 = x \in X$) existe una distribución de probabilidad denotada por P_x^π , que está definida en el espacio de las historias del proceso hasta $t = \infty$, tal que

$$P_x^\pi(x_{t+1} \in B \mid h_t, a_t) = p(B \mid x_t, a_t), \quad t = 0, 1, \dots; \quad B \in \mathcal{B}(X). \quad (1.3)$$

La esperanza con respecto a la probabilidad P_x^π será denotada por E_x^π . Entonces, para $x_0 = x$, $\pi \in \Pi$ está definido el proceso estocástico $\{x_t\} \equiv \{x_t^{x, \pi}\}$ que llamamos **proceso controlado**.

Nota 1.2 Sean $x \in X$ un estado inicial fijo y $f \in S$ una política estacionaria.

Entonces el proceso correspondiente denotado por $\{x_t^{(f)}\}$ es un **proceso de Markov**, debido a (1.3) y a la definición 1.2. Además si el proceso está dado por (1.1), entonces la probabilidad de transición de $\{x_t^{(f)}\}$ se define por las siguientes fórmulas:

$$\begin{aligned} p(B \mid y, f(y)) &= P(x_{t+1} \in B \mid x_t = y) = \\ P(\Phi(y, f(y), \xi_0) \in B) &= \int_{\mathbb{R}^n} I_B(\Phi(y, f(y), w)) dF(w) \end{aligned} \quad (1.4)$$

$$t = 0, 1, \dots$$

1.3 Índice de funcionamiento-costo promedio

Para considerar el problema de optimización de las políticas de control necesitamos definir un índice de funcionamiento de cada política, que va a servir como un criterio de optimización. Tal índice puede definirse de diversas maneras (ver [24]), pero, en esta tesis usamos el **costo promedio** que se define como:

$$J(x, \pi) := \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} E_x^\pi c(x_t, a_t), \quad (1.5)$$

cuando se aplica la política $\pi \in \Pi$ dado que el estado inicial es $x_0 = x \in X$.

El problema de control óptimo en costo promedio consiste en encontrar una política $\pi_* \in \Pi$ que satisfaga:

$$J(x, \pi_*) = J_*(x) := \inf_{\pi \in \Pi} J(x, \pi) \text{ para todo } x \in X. \quad (1.6)$$

La política que satisface (1.6) se le llama **política óptima en costo promedio**. La función $J_*(x)$, $x \in X$, que describe un mejor funcionamiento del sistema, se le llama la **función de costo óptimo**.

Nota 1.3 El teorema 1 del capítulo 4 afirma, que bajo las hipótesis introducidas en el capítulo 3, para los procesos de control considerados, existen políticas óptimas estacionarias f_* , i.e. tal política $f_* \in S$ cumple que

$$J(x, f_*) = J_*(x) \text{ para cada } x \in X. \quad (1.7)$$

De (1.6) y (1.7) se ve que en este caso $\inf_{\pi \in \Pi} J(x, \pi) = \inf_{f \in S} J(x, f)$, $x \in X$.

Por consiguiente, el análisis del problema de estabilidad de control óptimo puede ser restringido a la clase S de políticas estacionarias.

CAPÍTULO 2

PLANTEAMIENTO DEL PROBLEMA DE ESTABILIDAD CON RESPECTO AL COSTO PROMEDIO

De nuevo, consideremos el modelo de control de Markov que está definido mediante las ecuaciones recurrentes (1.1):

$$x_{t+1} = \Phi(x_t, a_t, \xi_t), \quad t = 0, 1, \dots \quad (2.1)$$

$$x_0 = x \in X,$$

llamaremos al proceso controlado correspondiente (dado por (2.1)) "proceso real", para distinguirlo del siguiente "proceso aproximante" :

$$\tilde{x}_{t+1} = \Phi(\tilde{x}_t, \tilde{a}_t, \tilde{\xi}_t), \quad t = 0, 1, \dots \quad (2.2)$$

$$\tilde{x}_0 = x \in X.$$

Ambos procesos (2.1) y (2.2) están definidos en el mismo espacio de estados X , de controles A (ambos de Borel), con los mismos conjuntos de controles admisibles $A(x)$, $x \in X$, y, ambos tienen la misma función de costo por una etapa $c(x, a)$. Por otro lado en (2.1) los V.A'S i.i.d. ξ_t , $t = 0, 1, \dots$ (con los valores en R^n) tienen la función de distribución común F , mientras en (2.2) los V.A'S i.i.d. $\tilde{\xi}_t$, $t = 0, 1, \dots$ (con los valores en R^n) tienen la función de distribución común \tilde{F} . Interpretamos \tilde{F} como la aproximación dada a la función de distribución F conocida incompletamente.

Nota 2.1 La probabilidad de transición del proceso de control (2.2), denotado por \tilde{p} , se determina de forma semejante a (1.2):

$$\begin{aligned} \tilde{p}(B | x, a) &= P(\Phi(\tilde{x}_t, \tilde{a}_t, \tilde{\xi}_t) \in B | \tilde{x}_t = x, \tilde{a}_t = a) = \\ &P(\Phi(x, a, \tilde{\xi}_0) \in B) = \int_{R^n} I_B[\Phi(x, a, w)] d\tilde{F}(w). \end{aligned} \quad (2.3)$$

También, para cada política estacionaria $f \in S$ fija, se cumplen las fórmulas análogas a (1.4), pero para el proceso de Markov $\{\tilde{x}_t^{(f)}\}$ definido según (2.2) por las ecuaciones:

$$\tilde{x}_{t+1} = \Phi(\tilde{x}_t, f(\tilde{x}_t), \tilde{\xi}_t), \quad t = 0, 1, \dots \quad (2.4)$$

$$\tilde{x}_0 = x \in X.$$

Introducimos el costo promedio $\tilde{J}(x, \pi)$ para el proceso de control (2.2) (análogamente a (1.5)). A saber, si para (2.2) se aplica la política $\pi \in \Pi$ con el estado inicial $x_0 = x \in X$, entonces

$$\tilde{J}(x, \pi) := \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} E_x^\pi c(\tilde{x}_t, \tilde{a}_t). \quad (2.5)$$

Del mismo modo como en (1.6), la política $\tilde{\pi}_*$ se llama óptimo (en costo promedio) para el proceso de control (2.2), si $\tilde{\pi}_*$ satisface lo que sigue:

$$\tilde{J}(x, \tilde{\pi}_*) = \tilde{J}_*(x) := \inf_{\pi \in \Pi} \tilde{J}(x, \pi), \text{ para todo } x \in X. \quad (2.6)$$

Si, también $\tilde{f}_* \in S$ y

$$\tilde{J}(x, \tilde{f}_*) = \tilde{J}_*(x) \text{ para cada } x \in X, \quad (2.7)$$

entonces como en (1.7) decimos que \tilde{f}_* es una política estacionaria **óptima para el proceso (2.2)**.

Nota 2.2 (i) Para las políticas óptimas y óptimas estacionarias del proceso de control (2.1) usamos las denotaciones π_* y f_* , como en las definiciones (1.6) y (1.7).

(ii) Observamos que, en general, $J(x, \pi) \neq \tilde{J}(x, \pi)$ y $J_*(x) \neq \tilde{J}_*(x)$ ya que las distribuciones de (x_t, a_t) y $(\tilde{x}_t, \tilde{a}_t)$ dependen, respectivamente de F y \tilde{F} , por consiguiente son diferentes.

(iii) A pesar de que usamos las políticas óptimas estacionarias f_* y \tilde{f}_* en las definiciones de este capítulo, la demostración de su existencia la posponemos hasta en el capítulo 5.

Siguiendo a [8, 9, 10] definimos el **índice de estabilidad promedio** Δ_p como:

$$\Delta_p(x) := J(x, \tilde{f}_*) - J(x, f_*), \quad (2.8)$$

donde $J(x, \tilde{f}_*)$ es el costo promedio que corresponde a la aplicación de la política \tilde{f}_* para el proceso original (2.1), mientras $J(x, f_*) \equiv J_*(x)$, es el costo óptimo para este proceso.

Nuestro problema es demostrar la existencia de las políticas estacionarias óptimas f_* , \tilde{f}_* , y establecer estimaciones cuantitativas de la estabilidad del proceso (2.1) que garanticen: $\Delta_p(x) \rightarrow 0$ cuando $\sigma(F, \tilde{F}) \rightarrow 0$ para todo $x \in X$, i.e. conseguir una desigualdad del tipo:

$$\Delta_p(x) \leq \varphi[\sigma(F, \tilde{F})], \quad x \in X, \quad (2.9)$$

con una función φ tal que $\varphi(\omega) \rightarrow 0$ si $\omega \rightarrow 0$. En (2.9) σ denota la **métrica de variación total**, cuya definición la damos enseguida.

Definición 2.1 Sean ξ y $\tilde{\xi}$ vectores aleatorios en R^n con las distribuciones F y \tilde{F} , respectivamente, se define la métrica de variación total σ por:

$$\sigma(\xi, \tilde{\xi}) \equiv \sigma(F, \tilde{F}) := \sup_{h \in H} |Eh(\xi) - Eh(\tilde{\xi})|, \quad (2.10)$$

donde H es la clase de todas las funciones medibles $h : R^n \rightarrow R$, con la propiedad de que $\sup_{w \in R^n} |h(w)| \leq 1$.

Una definición de σ equivalente a (2.10) está dado por (ver [25]):

$$\sigma(\xi, \tilde{\xi}) \equiv \sigma(F, \tilde{F}) = 2 \sup_{B \in \mathcal{B}(R^n)} |P(\xi \in B) - P(\tilde{\xi} \in B)|. \quad (2.11)$$

CAPÍTULO 3

HIPÓTESIS ACERCA DE LA CLASE DE PROCESOS DE CONTROL CONSIDERADOS

Primero introduciremos algunas hipótesis generales que nos sirven para establecer la existencia de políticas óptimas estacionarias para los procesos de control (2.1) y (2.2).

Hipótesis H.1

- (i) El conjunto $A(x)$ es compacto para cada estado $x \in X$.
- (ii) El costo $c(x, a)$ en una etapa es acotado, a saber, existe una constante $k < \infty$ tal que:

$$|c(x, a)| \leq k \quad \text{para todo} \quad (x, a) \in K. \quad (3.1)$$

además la aplicación $a \rightarrow c(x, a)$ es semicontinua inferiormente para cada $x \in X$.

- (iii) Las probabilidades de transición dadas en (1.2) y (2.3) son fuertemente continuas,

i.e. para cualquier función $g : X \rightarrow R$ medible y acotada, y $x \in X$, las aplicaciones

$$\begin{aligned} a &\rightarrow \int_X g(y) p(dy | x, a) , \\ a &\rightarrow \int_X g(y) \tilde{p}(dy | x, a) , \end{aligned}$$

son continuas en $A(x)$.

Nota 3.1 Para la verificación de H.1 (iii), es preferible tener algunas condiciones suficientes en términos de las propiedades de la función Φ y las distribuciones de ξ y $\tilde{\xi}$ en (2.1) y (2.2). Se conocen algunas condiciones de este tipo (ver [16, apéndice C]) que, en palabras generales, requieren la existencia de densidades "suficientemente suaves" de F y \tilde{F} . El ejemplo del uso de tales condiciones las consideraremos en el capítulo 6.

No es muy probable esperar la existencia de buenas cotas para Δ_p , sin ninguna de las condiciones que aseguren la ergodicidad de los procesos (2.1), (2.2), cuando se aplican las políticas estacionarias. Esta afirmación se apoya por los contraejemplos en [8, 9, 10], y también por razones de que el costo promedio $J(x, \pi)$ (ver (1.5)) depende fuertemente del comportamiento de (x_t, a_t) para $t \rightarrow \infty$. Nuestras hipótesis de ergodicidad se fórmulan en los términos de existencia de un estado $x_* \in X$ exclusivo y la función de Lyapunov V , tales que se cumplen condiciones llamadas "condiciones de deriva" (ver [18, 23]).

Cuando aplicamos cualquier política estacionaria $f \in S$ a los procesos (2.1) y (2.2) recibimos, como se observó en las Notas 1.2 y 2.1 dos procesos de Markov dados por las siguientes ecuaciones:

$$x_{t+1} = \Phi(x_t, f(x_t), \xi_t), \quad t = 0, 1, \dots \quad (3.2)$$

$$\tilde{x}_{t+1} = \Phi(\tilde{x}_t, f(\tilde{x}_t), \tilde{\xi}_t), \quad t = 0, 1, \dots \quad (3.3)$$

con algunos estados iniciales x_0, \tilde{x}_0 , y las probabilidades de transición denotados respectivamente por $p(\cdot | \cdot, f(\cdot))$ y $\tilde{p}(\cdot | \cdot, f(\cdot))$ descritos en las Notas 1.2 y 2.1. En el resto del trabajo denotaremos el proceso en (3.2) por $\{x_t^{(f)}\}$, y el proceso en (3.3) por $\{\tilde{x}_t^{(f)}\}$.

Hipótesis H.2

Existe un estado $x_* \in X$ y un número $\gamma > 0$ tales que cumplen lo siguiente:

- (i) $\inf_{f \in S} p(\{x_*\} | x_*, f(x_*)) \geq \gamma$,
- (ii) $\inf_{f \in S} \tilde{p}(\{x_*\} | x_*, f(x_*)) \geq \gamma$,

Hipótesis H.3

Existe una función no negativa $V(x)$ y números positivos α, b y $s > 2$, tales que las siguientes condiciones se satisfacen para cada política estacionaria $f \in S$.

- (i) $\sup_{x \neq x_*} \left[\int_X V(y) p(dy | x, f(x)) - V(x) \right] \leq -\alpha$,
- (ii) $\sup_{x \in X} E_x^f \left\{ |V(x_1^{(f)}) - V(x)|^s \mid x_0^{(f)} = x \right\} \leq b$.

Hipótesis H.4

Las condiciones (i) y (ii) en H.3 son válidas para el proceso $\{\tilde{x}_t^{(f)}\}$, $f \in S$, con una función \tilde{V} y constantes positivas $\tilde{\alpha}, \tilde{b}$, pero, con la misma constante s en H.3 (ii).

Nota 3.2 (i) Es importante subrayar que las constantes $s, \alpha, b, \tilde{\alpha}, \tilde{b}$, y las funciones V y \tilde{V} no dependen de la política $f \in S$.

(ii) Al analizar la demostración del teorema 2, dada en el capítulo 5, podríamos ver que las hipótesis H.2-H.4 se pueden cambiar por las mismas condiciones, pero para su realización se requiere no para todas las políticas estacionarias, sino sólo para políticas de una subclase $S_* \subset S$. Esta subclase puede ser arbitraria pero tal que S_* incluya las políticas óptimas f_* y \tilde{f}_* . Así, la verificación de las hipótesis H.2-H.4, para procesos de control concretos, puede simplificarse, buscando una subclase S_* suficientemente pequeña.

No obstante, si en H.2-H.4 utilizáramos una subclase $S_* \neq S$ de políticas estacionarias, entonces la demostración del teorema 1 en el capítulo 5 ya no sería válida. Por lo tanto, en esta situación tenemos que postular la existencia de políticas óptimas estacionarias.

CAPÍTULO 4

RESULTADOS PRINCIPALES

El primer teorema justifica la existencia de políticas estacionarias óptimas en costo promedio (ver las definiciones (1.6), (1.7) y (2.6), (2.7)).

Teorema 1 Las hipótesis H.1, H.2, H.3, H.4 son suficientes para la existencia de las políticas estacionarias f_* , $\tilde{f}_* \in S$ óptimas, respectivamente, para los procesos de control (2.1) y (2.2). Además, bajo estas hipótesis los costos promedios $J(x, f)$ y $\tilde{J}(x, f)$ en (1.5) y (2.5) no dependen del estado inicial si $f \in S$.

De lo último, vamos a escribir $J(f)$ en lugar de $J(x, f)$ y $\tilde{J}(f)$ en lugar de $\tilde{J}(x, f)$. También, teniendo en cuenta la definición (2.8) el índice de estabilidad promedio $\Delta_p(x)$ no depende de $x \in X$, así escribiremos en lo que sigue Δ_p . El resultado más importante de este trabajo se da en el siguiente teorema.

Teorema 2 Suponiendo que las hipótesis H.1, H.2, H.3, H.4 se satisfacen. Entonces, para cualesquiera estados iniciales x_0 y \tilde{x}_0 de los procesos (2.1) y (2.2) se tiene:

$$\Delta_p \leq kN \left[\sigma(F, \tilde{F}) \right]^{\frac{s-1}{s}}, \quad (4.1)$$

donde k es la cota para la función de costo en H.1 (ii), y $N = N(s, \gamma, \alpha, b, \tilde{\alpha}, \tilde{b})$ es una constante, que puede calcularse mediante las constantes involucradas en H.2-H.4.

Corolario 1 Bajo las hipótesis del teorema 1 el proceso de control (2.1) es estable, a saber, para cada estado inicial del proceso

$$\Delta_p \rightarrow 0 \quad \text{cuando} \quad \sigma(F, \tilde{F}) \rightarrow 0. \quad (4.2)$$

Nota 4.1 (i) Según (4.1) la rapidez de convergencia en (4.2) depende de la constante $s > 2$ involucrado en la "condición de momentos" en H.3 (ii), como se puede ver en [19; 18, cap.5], el valor de s en H.3 (ii), esta relacionado con "momentos del tiempo de regreso al estado x_* ", y, en su turno, con la tasa de convergencia en el teorema de ergodicidad (ver, la demostración del teorema 1 en el capítulo 5).

(ii) (comentario sobre la constante N en (4.1)) El método de estimación de N se puede obtener de los resultados en [18, caps.6,7]. Esta constante se expresa por medio de los momentos de la potencia s del "tiempo de cruce" de los procesos $\{x_t^{(f)}\}$ y $\{\tilde{x}_t^{(f)}\}$ vistos, posteriormente en este trabajo, como procesos regenerativos. El tiempo de cruce (ver [18, cap.6]) es el tiempo de primera regeneración común de ambos procesos $\{x_t^{(f)}\}$ y $\{\tilde{x}_t^{(f)}\}$. Estimar los momentos del tiempo de cruce, no es un problema sencillo. Sin embargo, para procesos de control concretos la constante N puede estimarse por el método dado en [18], por lo menos con ayuda de un programa para computadora.

CAPÍTULO 5

DEMOSTRACIONES

DE LOS RESULTADOS

5.1 Demostración del Teorema 1.

Mostraremos que el teorema 1 es una consecuencia del teorema 2.6 en [12] y del lema 2.10 en [13], donde, se afirma la existencia de políticas estacionarias óptimas en promedio y la independencia del costo promedio del estado inicial, bajo la hipótesis H.1 y algunas condiciones de Lyapunov sobre la deriva del proceso cuando se aplican las políticas estacionarias. En nuestro caso no hemos supuesto tales condiciones como en [12], sin embargo del análisis de la demostración del teorema 2.6 en [12], se ve que las condiciones de Lyapunov mencionadas se usan en [12] sólo para demostrar la siguiente propiedad de ergodicidad de los procesos controlables.

Condición E.R. Existen una función medible $W : X \rightarrow [0, \infty)$ y constantes α_t , $t = 0, 1, \dots$ con $\sum_{t=0}^{\infty} \alpha_t < \infty$ tales que para cada política estacionaria $f \in S$ el proceso

$\{x_t^{(f)}\}$ con estado inicial x satisface la siguiente condición de ergodicidad:

$$\sigma(x_t^{(f)}, x_\infty^{(f)}) \leq W(x) \alpha_t, \quad t = 0, 1, \dots \quad (5.1)$$

donde $x_\infty^{(f)}$ es el elemento aleatorio con valores en X , que tiene la distribución invariante q_f para el proceso $\{x_t^{(f)}\}$ a saber (ver [23]):

$$q_f(B) = \int_X p(B | x, f(x)) q_f(dx), \quad B \in \mathcal{B}(X).$$

Nota 5.1 Es importante (ver [12]) que W y α_t deben ser independientes de la política $f \in S$.

Entonces, para demostrar el teorema 1 (aplicando el teorema 2.6 en [12]) bastará verificar que las hipótesis H.2-H.4 implican la condición E.R. para los procesos $\{x_t^{(f)}\}$, $\{\tilde{x}_t^{(f)}\}$, $f \in S$. Vamos a hacerlo sólo para $\{x_t^{(f)}\}$ (ya que $\{\tilde{x}_t^{(f)}\}$ se trata análogamente), usando el resultado sobre ergodicidad de un proceso de Markov regenerativo en [19]. Primero, entonces, hay que verificar que el proceso $\{x_t^{(f)}\}$ (definido en (3.2)) con estado inicial $x_0 = x_*$ es un proceso de regeneración para cada política estacionaria $f \in S$.

Definición 5.1 (El caso particular de [18, cap.7]) Un proceso z_t a tiempo discreto $t = 0, 1, \dots$ con estados en un espacio de Borel Z se llama **proceso regenerativo** si existen variables aleatorias (v.a's) i.i.d. τ_1, τ_2, \dots con valores naturales, tales que si $\eta_n = \sum_{i=1}^n \tau_i$, $n = 1, 2, \dots$ entonces, para cada $n \geq 1$ "La parte de la trayectoria"

$$T_n = \{z_{\eta_n}, z_{\eta_n+1}, z_{\eta_n+2}, \dots\}, \quad (5.2)$$

no depende de z_t con $t < \eta_n$, y el proceso T_n tiene la misma distribución como la de $\{z_0, z_1, z_2, \dots\}$.

En nuestro caso las v.a's τ_n serán intervalos entre tiempos de regreso del proceso $\{x_t^{(f)}\}$ en el estado inicial $x_0 = x_*$. Pero antes, necesitamos estudiar las propiedades del tiempo de alcance $\tau_{x,f}$ en el conjunto $\{x_*\}$. Sean, ahora $f \in S$ y $\{x_t^{(f)}\}$ el proceso como en (1.4), (3.2) con estado inicial $x \in X$ arbitrario fijo. Definamos:

$$\tau_{x,f} = \begin{cases} \inf \{t > 0 : x_t^{(f)} = x_*\} \\ \infty & \text{si no existe tal } t. \end{cases} \quad (5.3)$$

Lema 5.1 Existen constantes $d, r < \infty$ independientes del estado inicial $x \in X$ y la política estacionaria $f \in S$ tales que:

$$(i) P_x^f(\tau_{x,f} < \infty) = 1,$$

(ii)

$$E_x^f \tau_{x,f} \leq \begin{cases} \left[d + \frac{2V(x)}{\alpha} \right]^s & \text{si } x \neq x_* \\ r & \text{si } x = x_* \end{cases} \quad (5.4)$$

donde V y α fueron definidos en H.3.

Nota 5.2 Las constantes d y r en (5.4) dependen sólo de las cantidades en H.3, a saber, $d = d(s, \alpha, b, V(x_*))$, $r = r(s, \alpha, b, V(x_*))$, (ver teorema 2 en [18 cap.5]).

Demostración

La desigualdad (5.4) se sigue como consecuencia directa del teorema 2, sección 5.2.1 en [18 cap.5]. Lo último establece la desigualdad (5.4) usando las condiciones (ver [18, cap.5]) que siguen de la hipótesis H.3 .

La afirmación (i) del lema 5.1 se sigue de (ii), pues si $P_x^f(\tau_{x,f} = \infty) > 0$, entonces

$$E_x^f \tau_{x,f}^s = E_x^f \left\{ \tau_{x,f}^s; \tau_{x,f} < \infty \right\} + E_x^f \left\{ \tau_{x,f}^s; \tau_{x,f} = \infty \right\} = \infty,$$

lo que contradice al punto (ii). ■

Lema 5.2 Sea $f \in S$ arbitraria fija. Si el estado inicial del proceso $\{x_t^{(f)}\}$ es x_* como en H.2, entonces $\{x_t^{(f)}\}$ es un proceso regenerativo con momentos de regreso en el estado x_* como los momentos de regeneración.

Lo mismo se cumple para el proceso $\{\tilde{x}_t^{(f)}\}$.

Demostración Definamos (ver (5.3) y la definición 5.1) por inducción:

$$\begin{aligned} \tau_1 &= \tau_{x_*,f}; \\ \eta_1 &= \tau_1, \\ \eta_2 &= \inf \{t > \eta_1 : x_t^{(f)} = x_*\}, \\ \eta_3 &= \inf \{t > \eta_2 : x_t^{(f)} = x_*\}, \dots \text{ etc.} \end{aligned} \tag{5.5}$$

Por el lema (5.1) (i) las v.a's η_1, η_2, \dots son finitos $P_{x_*}^f$ -casi seguramente, haciendo:

$$\tau_2 = \eta_2 - \eta_1, \quad \tau_3 = \eta_3 - \eta_2, \dots \quad \text{etc};$$

entonces $\eta_n = \sum_{i=1}^n \tau_i$, $n = 1, 2, \dots$. Mostremos que las v.a's enteras η_n , dadas en (5.5) (momentos de regreso del proceso al estado x_*), satisfacen la definición 5.1 (i.e. son "momentos de regeneración"). Se ve fácilmente que η_n es el tiempo de paro para $\{x_t^{(f)}\}$ (ver las definiciones en [18 cap.5]). Por otro lado, cualquier proceso de Markov a tiempo discreto posee la propiedad fuerte de Markov (ver [18, 23]). Esto implica que

dado el valor $x_{\eta_n}^{(f)}, x_t^{(f)}$, con $t < \eta_n$ no depende de $x_t^{(f)}$ con $t \geq \eta_n$. En particular, esto significa que las v.a's τ_1, τ_2, \dots son independientes. Luego, por (5.5) (ver la definición 5.1)

$T_n = \{x_*, x_{\eta_n+1}^{(f)}, x_{\eta_n+2}^{(f)}, \dots\}$, y la evolución del proceso $\{x_t^{(f)}\}$ en los momentos $\eta_n + 1, \eta_n + 2, \dots$ se determinan únicamente por la probabilidad de transición $p(\cdot | \cdot, f(\cdot))$.

Lo mismo se puede decir para $\{x_0^{(f)}, x_1^{(f)}, x_2^{(f)}, \dots\} = \{x_*, x_1^{(f)}, x_2^{(f)}, \dots\} = T_0$.

Por lo tanto, T_n y T_0 tienen la misma distribución. ■

Nota 5.3 La demostración del lema 5.2 para el proceso $\{\tilde{x}_t^{(f)}\}$, $f \in S$, es análoga a la anterior, pero para realizarla, falta asegurar, como en el lema 5.1 (i), que los tiempos de regreso a x_* son finitos con probabilidad 1.

Considerando $\{\tilde{x}_t^{(f)}\}$, $x \in X$ un estado inicial arbitrario fijo, definamos como en (5.3) lo siguiente:

$$\tilde{\tau}_{x,f} = \begin{cases} \inf \{t > 0 : \tilde{x}_t^{(f)} = x_*\} \\ \infty \end{cases} \quad \text{si no existe tal } t. \quad (5.6)$$

Lema 5.3 Bajo H.4 se establece:

(i) $P_x^f(\tilde{\tau}_{x,f} < \infty) = 1$,

(ii)

$$\sup_{f \in S} E_x^f \tilde{\tau}_{x,f}^s \leq \begin{cases} \left[\tilde{d} + \frac{2\tilde{V}(x)}{\tilde{\alpha}} \right]^s & \text{si } x \neq x_* \\ \tilde{r} < \infty & \text{si } x = x_* \end{cases} \quad (5.7)$$

donde $\tilde{d} = \tilde{d}(s, \tilde{\alpha}, \tilde{b}, \tilde{V}(x_*))$, $\tilde{r} = \tilde{r}(s, \tilde{\alpha}, \tilde{b}, \tilde{V}(x_*))$.

La demostración, se hace del mismo modo como la demostración del lema 5.1.

Para aplicar los resultados de ergodicidad de [19], es necesario establecer "buenas propiedades aritméticas" de las distribuciones de los tiempos de regreso en (5.3) y (5.6).

Definición 5.2 Sea τ una v.a que toma los valores naturales $1, 2, \dots$, y sea G_τ su función de distribución. Decimos que $G_\tau \in \text{UNP}(N, a)$, donde $N \in \{1, 2, 3, \dots\}$, $a > 0$ si

$$\text{MCD} \{k: k \leq N \text{ y } P(\tau = k) \geq a\} = 1 .$$

Aquí: MCD es la abreviatura para el "máximo común divisor", y UNP significa "uniformemente no periódica". (ver [18 pag.181]).

Lema 5.4 Bajo la hipótesis H.2 las distribuciones de las v.a's $\tau_{x_*, f}$ y $\tilde{\tau}_{x_*, f}$ en (5.3), (5.6) están en la clase $\text{UNP}(1, \gamma)$.

Demostración: En efecto

$$\begin{aligned} P(\tau_{x_*, f} = 1) &= P(x_1^{(f)} = x_* \mid x_0^{(f)} = x_*) \\ &= p(\{x_*\} \mid x_*, f(x_*)) \geq \gamma \end{aligned}$$

por H.2 (i), lo que significa que la distribución de $\tau_{x_*, f}$ está en $\text{UNP}(1, \gamma)$.

Lo mismo es válido para $\tilde{\tau}_{x_*, f}$. ■

Para el siguiente lema 5.5 necesitamos dar dos definiciones acerca de procesos de Markov a tiempo discreto generales (ver [23]). Sea $\{z_t\}$ un proceso de Markov en un espacio de Borel Z .

Definición 5.3 Sea ψ una medida σ -finita en $(Z, \mathcal{B}(Z))$. El proceso $\{z_t\}$ se llama ψ -irreducible si para cada $z \in Z$, $B \in \mathcal{B}(Z)$ con $\psi(B) > 0$, se encuentra un número n natural tal que

$$P_z(z_n \in B \mid z_0 = z) > 0.$$

Definición 5.4 El proceso $\{z_t\}$ ψ -irreducible se llama **Harris-recurrente**, si para cada $C \in \mathcal{B}(Z)$ con $\psi(C) > 0$, $P_z(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{z_n \in C\}) = 1$ para todo estado inicial $z \in Z$.

Lo que sigue es la última preparación para aplicar el teorema 1 de [19].

Lema 5.5 Bajo las hipótesis H.2, H.3 y H.4 para cada $f \in S$ y cada estado inicial x , los procesos $\{x_t^{(f)}\}$, $\{\tilde{x}_t^{(f)}\}$ en (3.2) y (3.3) son δ_{x_*} -irreducibles, aperiódicos y Harris-recurrentes. Aquí δ_{x_*} denota la distribución de Dirac, a saber

$$\delta_{x_*}(B) = \begin{cases} 1 & \text{si } x_* \in B \\ 0 & \text{si } x_* \notin B \end{cases}$$

Demostración

Sea f una política estacionaria arbitraria fija. Demostramos el lema sólo para $\{x_t^{(f)}\}$ (ya que para $\{\tilde{x}_t^{(f)}\}$ el razonamiento es el mismo). De H.2 (i), obtenemos que para todo $x \in X$, $B \in \mathcal{B}(X)$

$$p(B \mid x, f(x)) \geq \gamma I_{\{x_*\}}(x) \delta_{x_*}(B). \quad (5.8)$$

En efecto, cuando $x_* \notin B$ o $x \neq x_*$ la parte derecha de (5.8) es cero. Por otro lado, si $x = x_*$ y $x_* \in B$, entonces la parte derecha de (5.8) es γ , y de H.2 (i).

$$p(B \mid x_*, f(x_*)) \geq p(\{x_*\} \mid x_*, f(x_*)) \geq \gamma.$$

La desigualdad (5.8) nos dice que el proceso $\{x_t^{(f)}\}$ es estrictamente aperiódico, y, por lo tanto, es aperiódico (ver [23 cap.5 sección 5.4.3]).

Sea $x \in X$ el estado inicial de $\{x_t^{(f)}\}$. Según el lema 5.1 el tiempo de alcance $\tau_{x,f}$ al conjunto $\{x_*\}$ es P_x^f -casi seguramente finito. Denotaremos ahora este tiempo por $t_1(x)$. Es claro que $t_1(x)$ es tiempo de paro, y, por la propiedad fuerte de Markov, el proceso $\{x_t^{(f)}\}$ tiene después de $t_1(x)$ la misma estructura probabilista como después de $t = 0$ con el estado inicial $x_0 = x_*$. Otra vez aplicando el lema 5.1, aseguramos la existencia con $P_x^f = 1$ del momento $t_2(x) - t_1(x)$ tal que $x_{t_2(x)} = x_*$ (de hecho $t_2(x) - t_1(x)$ tiene la misma distribución que $\tau_{x_*,f}$ en (5.3)). Al repetir estos argumentos, por inducción, establecemos que:

$$\begin{aligned} P_x^f (x_t^{(f)} = x_* \text{ un número infinito de veces}) &= \\ P_x^f \left(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{x_n^{(f)} = x_*\} \right) &= 1 \end{aligned} \quad (5.9)$$

para x arbitrario (así para todo $x \in X$).

Como $P_x^f (t_1(x) < \infty) = 1$ se sigue que existe un $n \geq 1$ tal que:

$$P_x^f (t_1(x) = n) > 0. \quad (5.10)$$

Ahora, (ver, la definición 5.3 y la definición de δ_{x_*}), si $\delta_{x_*}(B) > 0$, entonces $x_* \in B$, y por (5.10)

$$P_x^f(x_n^{(f)} \in B \mid x_0^{(f)} = x) \geq P_x^f(x_n^{(f)} = x_* \mid x_0^{(f)} = x) = P_x^f(t_1(x) = n) > 0$$

lo cual demuestra que $\{x_t^{(f)}\}$ es δ_{x_*} -irreducible.

Acerca de Harris-recurrencia tenemos lo siguiente. Como $\delta_{x_*}(C) > 0$ significa que $x_* \in C$, se sigue de (5.9) que:

$$P_x^f\left(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{x_n^{(f)} \in C\}\right) \geq P_x^f\left(\bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \{x_n^{(f)} \in \{x_*\}\}\right) = 1$$

para cada $x \in X$.

La última desigualdad establece que $\{x_t^{(f)}\}$ es Harris-recurrente. ■

Finalmente, para concluir la demostración del teorema 1, volvamos a la condición E.R. Para cualquier $f \in S$ las desigualdades (5.4), (5.7) y los lemas 5.2, 5.4, 5.5 aseguran la hipótesis del corolario del teorema 1 en [19], de lo último se sigue que existe un elemento aleatorio $x_{\infty}^{(f)}$ con la distribución invariante para $p(\cdot \mid \cdot, f(\cdot))$ tal que, cuando el estado inicial es x , se tiene

$$\sigma(x_t^{(f)}, x_{\infty}^{(f)}) \leq Wt^{1-s}, \quad t = 1, 2, \dots; \quad (5.11)$$

donde $s > 2$ aparece en la hipótesis H.3 (ii), y W depende sólo del estado inicial (por medio de $V(x)$) y de las constantes s, γ, d, α, r en los lemas 5.1 y 5.4. Por consiguiente, W en (5.11) no depende de la política $f \in S$, y como $\sum_{t=1}^{\infty} t^{1-s} < \infty$, se sigue que la condición E.R se satisface.

Por lo tanto, aplicando el teorema 2.6 de [12] y el lema 2.10 de [13], establecemos la existencia de la política estacionaria f_* óptima para (2.1) y la independencia del costo promedio $J(x, f)$ del estado inicial x para cada $f \in S$.

De la misma manera demostramos la existencia de la política $\tilde{f}_* \in S$ óptima en costo promedio para el proceso controlado (2.2), y además, el hecho:

$$\tilde{J}(x, f) \equiv \tilde{J}(f), \quad f \in S.$$

Así, el teorema 1 queda demostrado. ■

5.2 Demostración del teorema 2.

De hecho, en el párrafo anterior hicimos la mayoría de las preparaciones para establecer el teorema 2. Los siguientes dos lemas nos permiten deducir la desigualdad (4.1) por medio de las cotas conocidas en el problema de comparación (uniforme en t) de procesos regenerativos.

Lema 5.6 Si $J(x, f) \equiv J(f)$, $\tilde{J}(x, f) \equiv \tilde{J}(f)$ los costos promedios definidos en (1.5), (2.5) para los procesos controlables (2.1), (2.2), entonces:

$$\Delta_p \leq 2 \sup_{f \in S} |J(f) - \tilde{J}(f)|. \quad (5.12)$$

Demostración

Sean $f_*, \tilde{f}_* \in S$ las políticas óptimas del teorema 1. Escribiendo Δ_p para $\Delta_p(x)$ en (2.8), (por independencia de x) tenemos (ver (1.6), (1.7) y (2.6), (2.7)):

$$\begin{aligned}
\Delta_p &= J(\tilde{f}_*) - J(f_*) \\
&= J(\tilde{f}_*) - \tilde{J}(\tilde{f}_*) + \tilde{J}(\tilde{f}_*) - J(f_*) \\
&= J(\tilde{f}_*) - \tilde{J}(\tilde{f}_*) + \inf_{f \in S} \tilde{J}(f) - \inf_{f \in S} J(f), \tag{5.13}
\end{aligned}$$

ya que

$$\begin{aligned}
|\inf \tilde{\varphi}(u) - \inf \varphi(u)| &= |\sup(-\tilde{\varphi}(u)) - \sup(-\varphi(u))| \\
&\leq \sup |\tilde{\varphi}(u) - \varphi(u)|.
\end{aligned}$$

De (5.13),

$$\begin{aligned}
\Delta_p &\leq |J(\tilde{f}_*) - \tilde{J}(\tilde{f}_*)| + \sup_{f \in S} |\tilde{J}(f) - J(f)| \\
&\leq 2 \sup_{f \in S} |J(f) - \tilde{J}(f)|. \blacksquare
\end{aligned}$$

Lema 5.7 Bajo la hipótesis H.1 (ii),

$$\Delta_p \leq 2k \sup_{f \in S} \sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}); \tag{5.14}$$

donde k es la constante en H.1 (ii), mientras $\{x_t^{(f)}\}$, $\{\tilde{x}_t^{(f)}\}$ son los procesos de Markov definidos en (3.2), (3.3) con los estados iniciales $x_0^{(f)} = \tilde{x}_0^{(f)} = x_*$.

Demostración

Debido a que J y \tilde{J} no dependen de x podríamos reescribir (5.12) como:

$$\Delta_p \leq 2 \sup_{f \in S} |J(x_*, f) - \tilde{J}(x_*, f)|, \quad (5.15)$$

luego por (1.5), (2.5)

$$\begin{aligned} |J(x_*, f) - \tilde{J}(x_*, f)| &= \left| \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} \left\{ E_{x_*}^f c(x_t, f(x_t)) - E_{x_*}^f c(\tilde{x}_t, f(\tilde{x}_t)) \right\} \right| \\ &\leq k \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} \left| E_{x_*}^f \frac{c(x_t, f(x_t))}{k} - E_{x_*}^f \frac{c(\tilde{x}_t, f(\tilde{x}_t))}{k} \right| \\ &\leq k \limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{t=0}^{n-1} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}), \end{aligned} \quad (5.16)$$

debido a la definición (2.10). De (5.16) obtenemos

$$|J(x_*, f) - \tilde{J}(x_*, f)| \leq k \sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}),$$

por esta desigualdad y (5.15) obtenemos (5.14). ■

Utilizaremos (5.14) para establecer (4.1) estimando $\sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$ por unas cotas superiores que no dependen de la selección de la política $f \in S$. De nuevo, sean $f \in S$ una política estacionaria y $\{x_t^{(f)}\}, \{\tilde{x}_t^{(f)}\}$, la pareja de procesos como en (3.2), (3.3) con estados iniciales x_* . Primero establecemos una estimación de la distancia σ entre $\{x_t^{(f)}\}$ y $\{\tilde{x}_t^{(f)}\}$ uniforme sobre un intervalo finito $t = 0, 1, \dots, T$.

Lema 5.8 Para cada $T = 1, 2, \dots$

$$\max_{0 \leq t \leq T} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq T \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(\cdot | x, a), \tilde{p}(\cdot | x, a)). \quad (5.17)$$

Demostración

Previamente mostraremos que

$$\begin{aligned} \sigma(x_{t+1}^{(f)}, \tilde{x}_{t+1}^{(f)}) &\leq \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) + \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(\cdot | x, a), \tilde{p}(\cdot | x, a)) \quad (5.18) \\ t &= 0, 1, \dots \end{aligned}$$

Si denotamos por $p^t(B | x, f(x))$ y $\tilde{p}^t(B | x, f(x))$ las probabilidades de transición por t etapas, respectivamente, de los procesos $\{x_t^{(f)}\}$ y $\{\tilde{x}_t^{(f)}\}$, entonces

$$P(x_t^{(f)} \in B | x_0 = x) = p^t(B | x, f(x)),$$

$$P(\tilde{x}_t^{(f)} \in B | x_0 = x) = \tilde{p}^t(B | x, f(x)), \quad B \in \mathcal{B}(X), x \in X.$$

De aquí y de la definición (2.10) tenemos

$$\begin{aligned} I_{t+1} &:= \sigma(x_{t+1}^{(f)}, \tilde{x}_{t+1}^{(f)}) \\ &= \sup_{h \in H} \left| \int_X h(x) \{p^{t+1}(dx | x_*, f(x_*)) - \tilde{p}^{t+1}(dx | x_*, f(x_*))\} \right| \\ &= \sup_{h \in H} \left| \int_X h(x) \int_X p^t(dy | x_*, f(x_*)) p(dx | y, f(y)) - \right. \\ &\quad \left. \int_X h(x) \int_X \tilde{p}^t(dy | x_*, f(x_*)) \tilde{p}(dx | y, f(y)) \right| \end{aligned}$$

por la ecuación de Chapman-Kolmogorov. Luego, por el teorema de Fubini:

$$\begin{aligned}
I_{t+1} &= \sup_{h \in H} \left| \int_X p^t(dy | x_*, f(x_*)) \int_X h(x) p(dx | y, f(y)) - \right. \\
&\quad \left. \int_X \tilde{p}^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) + \right. \\
&\quad \left. \int_X p^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) - \right. \\
&\quad \left. \int_X \tilde{p}^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) \right| \\
&\leq \sup_{h \in H} \left| \int_X p^t(dy | x_*, f(x_*)) \int_X h(x) p(dx | y, f(y)) - \right. \\
&\quad \left. \int_X p^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) \right| + \\
&\quad \sup_{h \in H} \left| \int_X p^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) - \right. \\
&\quad \left. \int_X \tilde{p}^t(dy | x_*, f(x_*)) \int_X h(x) \tilde{p}(dx | y, f(y)) \right| \leq \\
&\quad \sup_{h \in H} \sup_{y \in X} \left| \int_X h(x) p(dx | y, f(y)) - \int_X h(x) \tilde{p}(dx | y, f(y)) \right| + \\
&\quad \sup_{h \in H} \left| \int_X p^t(dx | y, f(y)) \bar{h}(y) - \int_X \tilde{p}^t(dx | y, f(y)) \bar{h}(y) \right|. \quad (5.19)
\end{aligned}$$

Donde $\bar{h}(y) = \int_X h(x) \tilde{p}(dx | y, f(y)) \in H$ para cada $h \in H$,

ya que:

$$|\bar{h}(y)| \leq \int_X |h(x)| \tilde{p}(dx | y, f(y)) \leq \int_X \tilde{p}(dx | y, f(y)) = 1.$$

Entonces, el último sumando en (5.19) es menor o igual que

$$\sup_{\bar{h} \in H} \left| \int_X \bar{h}(y) p^t(dy | x_*, f(x_*)) - \int_X \bar{h}(y) \tilde{p}^t(dy | x_*, f(x_*)) \right| = \sigma(x_t^{(f)}, \tilde{x}_t^{(f)})$$

por la definición (2.10).

Reemplazando $\sup_{h \in H}$ y $\sup_{y \in X}$ por otro sumando en la parte derecha de (5.19), obtenemos que este sumando es

$$\begin{aligned} &\leq \sup_{y \in X} \sigma(p(\cdot | y, f(y)), \tilde{p}(\cdot | y, f(y))) \\ &\leq \sup_{y \in X} \sup_{a \in A(y)} \sigma(p(\cdot | y, a), \tilde{p}(\cdot | y, a)) \end{aligned}$$

ya que $f(y) \in A(y)$ para cada $f \in S$.

Por consiguiente,

$$I_{t+1} \leq \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) + \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(\cdot | x, a), \tilde{p}(\cdot | x, a))$$

que es la desigualdad (5.18).

Las desigualdades (5.17) se obtienen de (5.18) por inducción si tomamos en cuenta que

$$\sigma(x_0^{(f)}, \tilde{x}_0^{(f)}) = \sigma(x_*, x_*) = 0. \blacksquare$$

Lema 5.9 Para los procesos en (2.1) y (2.2) se establece que

$$\sup_{x \in X} \sup_{a \in A(x)} \sigma(p(\cdot | x, a), \tilde{p}(\cdot | x, a)) \leq \sigma(F, \tilde{F}).$$

Demostración

Para cada $(x, a) \in K$ definimos la aplicación medible $G_{x,a} : R^n \rightarrow X$ mediante la fórmula: $G_{x,a}(w) = \Phi(x, a, w)$, $w \in R^n$.

Así, (ver. (1.2) y (2.3)) para $B \in \mathcal{B}(X)$, $x \in X$, $a \in A(x)$

$$p(B | x, a) = P(\Phi(x, a, \xi_0) \in B) = P(\xi_0 \in G_{x,a}^{-1}(B)),$$

$$\tilde{p}(B | x, a) = P(\tilde{\Phi}(x, a, \tilde{\xi}_0) \in B) = P(\tilde{\xi}_0 \in G_{x,a}^{-1}(B)).$$

Entonces (ver (2.11))

$$\begin{aligned} & \sup_{x \in X} \sup_{a \in A(x)} \sigma(p(\cdot | x, a), \tilde{p}(\cdot | x, a)) = \\ & \sup_{x \in X} \sup_{a \in A(x)} \sup_{B \in \mathcal{B}(X)} 2 |p(B | x, a) - \tilde{p}(B | x, a)| = \\ & \sup_{x \in X} \sup_{a \in A(x)} \sup_{B \in \mathcal{B}(X)} 2 \left| P(\xi_0 \in G_{x,a}^{-1}(B)) - P(\tilde{\xi}_0 \in G_{x,a}^{-1}(B)) \right| \leq \\ & \sup_{D \in \mathcal{B}(R^n)} 2 \left| P(\xi_0 \in D) - P(\tilde{\xi}_0 \in D) \right| = \sigma(F, \tilde{F}), \end{aligned}$$

ya que $G_{x,a}^{-1}(B) \in \mathcal{B}(R^n)$ para cada x, a, B . ■

Corolario 5.1 Si $x_0^{(f)} = \tilde{x}_0^{(f)} = x_*$ entonces:

$$\max_{0 \leq t \leq T} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq T \sigma(F, \tilde{F}), \quad T = 1, 2, \dots \quad (5.20)$$

Falta hacer el último paso para acabar la demostración de (4.1) en el teorema 2. En el lema 5.2, demostramos que si $x_0^{(f)} = \tilde{x}_0^{(f)} = x_*$, entonces los procesos $\{x_t^{(f)}\}$ y $\{\tilde{x}_t^{(f)}\}$ son procesos regenerativos con ciclos de regeneración distribuidos respectivamente como las v.a's $\tau_{x_*,f}$ y $\tilde{\tau}_{x_*,f}$ definidos en (5.3) y (5.6). Según el lema 5.4 las distribuciones de los ciclos de regeneración pertenecen a la clase UNP(1, γ). Además, para las v.a's $\tau_{x_*,f}$, $\tilde{\tau}_{x_*,f}$ se cumplen las desigualdades (5.4), (5.7) (con las partes derechas independientes de $f \in S$). Todo lo dicho antes junto con las desigualdades (5.20) aseguran las hipótesis del corolario 1 (caso A), sección 7.4.2 en [18, cap.7]. Este corolario da la estimación de la distancia σ uniforme en $t = 0, 1, \dots$ entre procesos

regenerativos que satisfacen las condiciones descritas anteriormente. En nuestro caso este corolario garantiza que:

$$\sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq N' [\sigma(F, \tilde{F})]^{\frac{\epsilon-1}{\epsilon}}, \quad (5.21)$$

donde la constante N' , debido a las estimaciones de los lemas 5.1, 5.3, 5.4 puede ser escogida independientemente de las políticas $f \in S$, y, además, puede ser estimado por medio de las constantes en los lemas mencionados. Por lo tanto, de (5.21) obtenemos la desigualdad:

$$\sup_{f \in S} \sup_{t \geq 0} \sigma(x_t^{(f)}, \tilde{x}_t^{(f)}) \leq N' [\sigma(F, \tilde{F})]^{\frac{\epsilon-1}{\epsilon}},$$

que junto con (5.14) da la desigualdad (4.1) en el teorema 2 (con $N = 2N'$), lo cual demuestra lo que queríamos. ■

CAPÍTULO 6

EJEMPLO DE UN SISTEMA DE ESPERA CON TASA DE SERVICIO CONTROLABLE.

En este capítulo discutiremos un modelo de control particular, con la idea de ilustrar la posibilidad de aplicación de la estabilidad (4.1) y del teorema de existencia de políticas óptimas. Este ejemplo es un sistema de espera de tipo general $GI | GI | 1 | \infty$ con un servidor controlable, a saber, en la cual el controlador puede elegir y variar la velocidad del servicio de los clientes dependiendo de los tiempos de espera en la cola. Por ejemplo, las aplicaciones importantes de sistemas de control de este tipo, en redes de comunicación, se discuten en [6, 27].

Normalmente, el costo relacionado con un sistema de espera es una función creciente del tiempo de espera y de la velocidad de servicio, mientras la reducción de la velocidad resulta, como regla, un aumento de la cola, y por eso crece el tiempo

de espera. El objetivo del controlador es lograr "un arreglo" que dependa del estado (tiempo de espera) corriente. No existen métodos para encontrar una política de control óptimo para sistemas generales $GI | GI | 1 | \infty$ con las distribuciones arbitrarias de los tiempos de llegada de clientes y del servicio. Un enfoque razonable, para abordar el problema de control óptimo, es aproximar el sistema original por un sistema de control más sencillo, por ejemplo, por el sistema de espera $M | M | 1 | \infty$ con las distribuciones exponenciales del servicio y de los intervalos entre llegadas. Para sistemas exponenciales han sido desarrollados algunos métodos de control óptimo (ver por ejemplo, [21]). Por lo tanto, si una política de control óptimo en el sistema aproximante $M | M | 1 | \infty$ se aplica para el control del sistema original $GI | GI | 1 | \infty$, entonces surge el problema de estimación de la precisión de tal aproximación. La desigualdad (4.1) se puede usar para estimar la calidad de aproximación en los términos de la desviación de las distribuciones en $GI | GI | 1 | \infty$ de las distribuciones exponenciales.

Nuestra tarea en este capítulo es verificar para el ejemplo dado enseguida, las hipótesis H.1-H.4 y de ilustrar el modo de uso de la desigualdad (4.1).

Nota 6.1 El mismo ejemplo fue usado en [13] con propósitos distintos.

Los procesos considerados en este capítulo están definidos en el espacio de estados $X = [0, \infty)$, y en el conjunto de controles A , que es cualquier subconjunto compacto del segmento $(0, \theta]$ tal que $\theta \in A$. Supongamos que para cada $x \in X$ el conjunto de controles admisibles es A , a saber, $A(x) = A$, $X \in [0, \infty)$.

Los procesos de control están dados por las siguientes ecuaciones recurrentes:

$$x_{t+1} = (x_t + a_t \eta_t - \xi_t)^+, \quad t = 0, 1, \dots, \quad (6.1)$$

$$x_0 = 0;$$

$$\tilde{x}_{t+1} = (\tilde{x}_t + \tilde{a}_t \tilde{\eta}_t - \tilde{\xi}_t)^+, \quad t = 0, 1, \dots, \quad (6.2)$$

$$\tilde{x}_0 = 0;$$

donde: $x^+ = \max\{0, x\}$ para $x \in R$; $(\eta_0, \xi_0), (\eta_1, \xi_1), \dots$ son V.A'S i.i.d. en R^2 con los componentes no negativos y con la función de distribución común F , mientras, $(\tilde{\eta}_0, \tilde{\xi}_0), (\tilde{\eta}_1, \tilde{\xi}_1), \dots$ son V.A'S i.i.d. en R^2 con los componentes no negativos y con la función de distribución común \tilde{F} . Como siempre en este trabajo, interpretamos el proceso (6.2) como la aproximación al "proceso real" (6.1), y, correspondientemente, \tilde{F} como la aproximación conocida a F .

Las ecuaciones (6.1) describen, en particular, el modelo de control de la tasa de servicio en un sistema de espera del tipo $GI | GI | 1 | \infty$ con un único servidor y tiempo de espera no acotado. En este caso las variables que aparecen en (6.1) se interpretan como sigue:

- x_t denota el tiempo de espera en la fila del t -ésimo cliente;
- a_t denota el valor recíproco a la velocidad de servicio asignada para el t -ésimo cliente;
- $a_t \eta_t$ denota el tiempo de servicio del t -ésimo cliente;

Así la v.a η_t se interpreta como "el volumen de trabajo que requiere el t -ésimo cliente"; en algunos sistemas η_t puede ser, por ejemplo, la longitud de un mensaje transmitido en canal de una red de comunicación;

- ξ_t denota el tiempo entre llegadas del t -ésimo al $t + 1$ -ésimo cliente.

Las variables en (6.2) se interpretan de la misma manera.

Para concluir, la descripción de los sistemas de control, falta introducir la función de costo por una etapa $c(x, a)$ que se supone acotada por una constante k y semicontinua inferiormente en a .

Verificaremos las hipótesis H.1-H.4 del capítulo 3 suponiendo para este ejemplo lo siguiente.

Por sencillez, denotaremos por (ξ, η) el V.A. genérico con la distribución F y por $(\tilde{\xi}, \tilde{\eta})$ el V.A. genérico con la distribución \tilde{F} .

Hipótesis H.E.1

(i) Las variables aleatorias ξ y η son independientes y tienen las densidades ρ_ξ y ρ_η acotadas y continuas en sus soportes.

(ii) Las variables aleatorias $\tilde{\xi}$ y $\tilde{\eta}$ son independientes y tienen las densidades $\rho_{\tilde{\xi}}$ y $\rho_{\tilde{\eta}}$ acotadas y continuas en sus soportes.

Hipótesis H.E.2

(i) $\theta E\eta < E\xi < \infty$;

(ii) $\theta E\tilde{\eta} < E\tilde{\xi} < \infty$;

Hipótesis H.E.3

Existen constantes $s > 2$ y $d, d_1, \tilde{d}, \tilde{d}_1$ tales que

$$(i) E\eta^s \leq d, E\xi^s \leq d_1;$$

$$(ii) E\tilde{\eta}^s \leq \tilde{d}, E\tilde{\xi}^s \leq \tilde{d}_1.$$

Nota 6.2 (i) La hipótesis H.E.2 expresa la condición de ergodicidad conocida (ver [18]) en el caso de aplicación de la política $f = (\theta, \theta, \theta, \dots)$ con servicio más lento $\theta\eta_t$. (sin tal condición $x_t \rightarrow \infty$ casi seguramente, a saber, no hay el régimen estacionario).

(ii) En la hipótesis H.E.3 se afirma que los s -ésimos momentos de los tiempos de servicio y de llegada son finitos. Desde el punto de vista de la práctica esta no es una condición restrictiva.

Nota 6.3 La verificación de las hipótesis H.1-H.4 para los procesos de control (6.1) y (6.2) se hace del mismo modo, ya que los procesos satisfacen las mismas condiciones. Por esta razón sólo lo haremos para (6.1).

Primero, las condiciones H.1 (iii), han sido verificadas para el modelo (6.1) en la proposición 5.4 en [13].

Luego, la hipótesis H.2 es la consecuencia simple de H.E.2. En efecto si elegimos $x_* = 0$ como estado exclusivo en H.2, entonces tenemos para cada política $f \in S$ (ver (1.4)):

$$\begin{aligned} p(\{0\} \mid 0, f(0)) &= P((f(0)\eta - \xi)^+ = 0) \\ &= P(f(0)\eta - \xi \leq 0) \geq P(\theta\eta \leq \xi), \end{aligned}$$

ya que $f(a) \in A \subset (0, \theta]$.

Debido a H.E.2 (i), existe un $\gamma > 0$ tal que

$$P(\theta\eta \leq \xi) \geq \gamma.$$

En efecto, si suponemos lo contrario,

entonces $P(\xi - \theta\eta \leq 0) = 1$ lo que implica $E(\xi - \theta\eta) \leq 0$,

lo cual contradice H.E.2.

Pasamos a examinar la hipótesis H.3 del capítulo 3.

Proposición 6.1 Bajo las hipótesis H.E.2, H.E.3 se cumple H.3.

Demostración

Al principio, observamos que existe una constante $h > 0$ tal que si

$$\xi_h = \min \{h, \xi\}, \text{ entonces } E\xi_h > \theta E\eta. \quad (6.3)$$

De hecho, es evidente que $0 \leq \xi_h \leq \xi$, $\xi_h \rightarrow \xi$ cuando $h \rightarrow \infty$ y $E\xi < \infty$.

Entonces, del teorema de convergencia dominada de Lebesgue

$$\lim_{h \rightarrow \infty} E\xi_h = E\xi > \theta E\eta,$$

lo que establece la existencia de h con la propiedad (6.3).

Como anteriormente, elegimos $x_* = 0 \in X$, además definamos "la función de prueba" $V : [0, \infty) \rightarrow R$ como sigue:

$$V(x) = \begin{cases} 0 & \text{si } x = 0 \\ x + h & \text{si } x > 0 \end{cases} \quad (6.4)$$

donde el número h es de (6.3).

Sean $f \in S$ y $x \neq x_*$ (es decir, $x > 0$) una política estacionaria y un estado inicial arbitrarios fijos. Denotaremos $a = f(x)$. Por (1.4)

$$\begin{aligned} I(x) &= \int_X V(y) p(dy | x, a) - V(x) \\ &= E[V(x + a\eta - \xi)^+] - V(x) \leq E[V(x + a\eta - \xi_h)^+] - V(x), \end{aligned} \quad (6.5)$$

debido a los hechos: $-\xi_h \geq -\xi$; y las funciones x^+ y V son no decrecientes.

Luego de (6.4) y (6.5)

$$\begin{aligned} I(x) &\leq E\{V(x + a\eta - \xi_h)^+; \xi_h \geq x + a\eta\} + E\{V(x + a\eta - \xi_h)^+; \xi_h < x + a\eta\} - V(x) \\ &= E\{x + h + a\eta - \xi_h; \xi_h < x + a\eta\} - (x + h) = \end{aligned}$$

$$\begin{aligned} &(x + h) P(\xi_h < x + a\eta) + E\{a\eta - \xi_h; \xi_h < x + a\eta\} \\ &- (x + h) P(\xi_h < x + a\eta) - (x + h) P(\xi_h \geq x + a\eta), \end{aligned} \quad (6.6)$$

porque $P(D) + P(D^c) = 1$ para cualquier evento D .

Pues

$$E\{a\eta - \xi_h; \xi_h < x + a\eta\} = E(a\eta - \xi_h) - E\{a\eta - \xi_h; \xi_h \geq x + a\eta\} \text{ y}$$

$xP(\xi_h \geq x + a\eta) \geq 0$, de (6.6) obtenemos:

$$I(x) \leq E(a\eta - \xi_h) - E\{(a\eta - \xi_h); \xi_h \geq x + a\eta\} - hP(\xi_h \geq x + a\eta). \quad (6.7)$$

Luego por (6.3)

$$\begin{aligned} E\{(a\eta - \xi_h); \xi_h \geq x + a\eta\} &\geq E\{-\xi_h; \xi_h \geq x + a\eta\} \\ &\geq E\{-h; \xi_h \geq x + a\eta\} \\ &= -hP(\xi_h \geq x + a\eta). \end{aligned} \quad (6.8)$$

Comparando (6.7) y (6.8) establecemos que

$$I(x) \leq E(a\eta - \xi_h).$$

Pero $a \leq \theta$ por la definición de A , entonces de (6.3):

$$I(x) \leq E(\theta\eta - \xi_h) < 0. \quad (6.9)$$

Por consiguiente, si hacemos $\alpha = -E(\theta\eta - \xi_h)$, entonces α es positivo y no depende de $x \in [0, \infty)$.

Entonces, por (6.9) $\sup_{x>0} I(x) \leq -\alpha$ lo cual significa que se cumple la condición H.3 (i).

Respecto de la condición H.3 (ii), hacemos lo siguiente.

Sean $x > 0$, $f \in S$ un estado y política arbitrarios, y $a = f(x)$.

Como $x_1^{(f)} = (x + a\eta - \xi)^+$ se sigue :

$$\begin{aligned}
I(x, f) &:= E_x^f \left\{ \left| V(x_1^{(f)}) - V(x) \right|^s \mid x_0^{(f)} = x \right\} \\
&= E \left\{ \left| V[(x + a\eta - \xi)^+] - V(x) \right|^s \right\} \\
&= E \left\{ \left| V[(x + a\eta - \xi)^+] - V(x) \right|^s ; \xi \geq x + a\eta \right\} + \\
&\quad E \left\{ \left| V[(x + a\eta - \xi)^+] - V(x) \right|^s ; \xi < x + a\eta \right\} \\
&= (x + h)^s P(\xi \geq x + a\eta) + \\
&\quad E \left\{ |a\eta - \xi|^s ; \xi < x + a\eta \right\} \leq \\
&\quad (x + h)^s P(\xi \geq x) + E \left\{ |a\eta - \xi|^s \right\}
\end{aligned} \tag{6.10}$$

Se conoce bien que

$$|a + b|^s \leq 2^{s-1} (|a|^s + |b|^s), \text{ para } s \geq 1, a, b \in R. \tag{6.11}$$

Entonces, de (6.10) obtenemos

$$\begin{aligned} I(x, f) &\leq 2^{s-1} \{[x^s + h^s] P(\xi \geq x) + a^s E\eta^s + E\xi^s\} \\ &\leq 2^{s-1} x^s P(\xi \geq x) + 2^{s-1} \{h^s + \theta^s E\eta^s + E\xi^s\}. \end{aligned} \quad (6.12)$$

Por H.E.3 (i), $E\eta^s$ y $E\xi^s$ son finitos, en particular que $x^s P(\xi \geq x) \rightarrow 0$ cuando $x \rightarrow \infty$. En vista de (6.12) esto significa que $\sup_{x>0} I(x, f)$ es acotada por una constante que no depende de $f \in S$.

Finalmente, para probar la hipótesis H.3 (ii), establecemos que $\sup_{f \in S} I(0, f) < \infty$.

Como $V(0) = 0$, se sigue:

$$\begin{aligned} I(0, f) &= E_0^f \{ |V(x_1^{(f)}) - V(0)|^s \mid x_0^{(f)} = 0 \} \\ &= E \{ |a\eta - \xi + h|^s ; \xi < a\eta \} + E \{ V[(a\eta - \xi)^+] ; \xi \geq a\eta \} \\ &= E \{ |a\eta - \xi + h|^s ; \xi < a\eta \} \\ &\leq r(s) \{ a^s E\eta^s + E\xi^s + h^s \} \\ &\leq r(s) \{ \theta^s E\eta^s + E\xi^s + h^s \}, \end{aligned} \quad (6.13)$$

porque (como en (6.11)) $|a + b + c|^s \leq r(s) \{ |a|^s + |b|^s + |c|^s \}$ con un número $r(s)$.

La parte derecha de (6.13) no depende de $f \in S$. En consecuencia, el proceso controlable (6.1) satisface la hipótesis H.3.

La proposición 6.1 queda demostrado. ■

Los argumentos descritos hasta de la proposición 6.1 y de esta proposición misma justifican que todas las hipótesis de esta tesis se satisfacen para los procesos controlables (6.1) y (6.2). Por consiguiente, podremos aplicar los teoremas 1 y 2 para estos sistemas. Según el teorema 1 existen las políticas estacionarias óptimas en promedio para los procesos (6.1) y (6.2), mientras, debido al teorema 2 para estos procesos se puede aplicar la desigualdad de estabilidad (4.1). Veremos que para el presente ejemplo (4.1) puede ser escrito en la forma siguiente, más adecuado para cálculos prácticos :

$$\Delta_p \leq kN \left\{ \left[\int_0^\infty |\rho_\xi(w) - \rho_{\tilde{\xi}}(w)| dw \right]^{\frac{s-1}{s}} + \left[\int_0^\infty |\rho_\eta(w) - \rho_{\tilde{\eta}}(w)| dw \right]^{\frac{s-1}{s}} \right\}, \quad (6.14)$$

la desigualdad (6.14) es consecuencia de (4.1) y de lo siguiente. (ya que $(a + b)^\beta \leq a^\beta + b^\beta$ para cada $a, b \geq 0$, $\beta \leq 1$).

Lema 6.1

$$\sigma(F, \tilde{F}) \leq \int_0^\infty |\rho_\xi(w) - \rho_{\tilde{\xi}}(w)| dw + \int_0^\infty |\rho_\eta(w) - \rho_{\tilde{\eta}}(w)| dw. \quad (6.15)$$

Demostración

Por independencia de ξ y η , también, de $\tilde{\xi}$ y $\tilde{\eta}$ obtenemos para las densidades de F y \tilde{F} que $\rho_{(\xi, \eta)} = \rho_\xi \rho_\eta$ y $\rho_{(\tilde{\xi}, \tilde{\eta})} = \rho_{\tilde{\xi}} \rho_{\tilde{\eta}}$. Se conoce bien (ver por ejemplo, [7, vol.II]) que para las distribuciones con densidades la distancia de variación total se representa como:

$$\sigma(F, \tilde{F}) = \int_{\mathbb{R}^2} \left| \rho_{(\xi, \eta)}(w_1, w_2) - \rho_{(\tilde{\xi}, \tilde{\eta})}(w_1, w_2) \right| dw_1 dw_2.$$

Entonces, sumando y restando $\rho_\xi(w_1) \rho_{\tilde{\eta}}(w_2)$, se obtiene:

$$\begin{aligned} \sigma(F, \tilde{F}) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \rho_\xi(w_1) \rho_\eta(w_2) - \rho_{\tilde{\xi}}(w_1) \rho_{\tilde{\eta}}(w_2) \right| dw_1 dw_2 \\ &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \rho_\xi(w_1) \rho_\eta(w_2) - \rho_\xi(w_1) \rho_{\tilde{\eta}}(w_2) \right| dw_1 dw_2 + \\ &\quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left| \rho_\xi(w_1) \rho_{\tilde{\eta}}(w_2) - \rho_{\tilde{\xi}}(w_1) \rho_{\tilde{\eta}}(w_2) \right| dw_1 dw_2 \\ &= \int_{-\infty}^{\infty} \rho_\xi(w_1) dw_1 \int_{-\infty}^{\infty} \left| \rho_\eta(w_2) - \rho_{\tilde{\eta}}(w_2) \right| dw_2 + \\ &\quad \int_{-\infty}^{\infty} \rho_{\tilde{\eta}}(w_2) dw_2 \int_{-\infty}^{\infty} \left| \rho_\xi(w_1) - \rho_{\tilde{\xi}}(w_1) \right| dw_1 \\ &= \int_{-\infty}^{\infty} \left| \rho_\eta(w_2) - \rho_{\tilde{\eta}}(w_2) \right| dw_2 + \int_{-\infty}^{\infty} \left| \rho_\xi(w_1) - \rho_{\tilde{\xi}}(w_1) \right| dw_1. \quad \blacksquare \end{aligned}$$

Ahora, como ilustración, hallemos las cotas superiores para la parte derecha en (6.15) en el caso particular siguiente:

(i) las v.a's η y $\tilde{\eta}$ tienen las distribuciones uniformes, respectivamente, en $[0, r]$, $[0, r + \varepsilon]$;

(ii) las v.a's ξ y $\tilde{\xi}$ tienen las distribuciones exponenciales, respectivamente, con parámetros μ y $\mu + \varepsilon$.

Aquí $r > 0, \mu > 0$ son números dados, y $\varepsilon \geq 0$ es un "parámetro pequeño" que representa "precisión de aproximación".

Por cálculos directos:

$$\begin{aligned} \int_0^\infty \left| \rho_\eta - \rho_{\tilde{\eta}} \right| dw &= \int_0^r \left(\frac{1}{r} - \frac{1}{r + \varepsilon} \right) dw + \\ \int_r^{r+\varepsilon} \frac{1}{r + \varepsilon} dw &= \frac{2\varepsilon}{r + \varepsilon} \leq \frac{2\varepsilon}{r}. \end{aligned} \quad (6.16)$$

Por otro lado, haciendo $w_* = \frac{1}{\varepsilon} \ln \left(\frac{\mu + \varepsilon}{\varepsilon} \right)$, tenemos

$$\int_0^{\infty} |(\mu + \varepsilon) \exp(-(\mu + \varepsilon)w) - \mu \exp(-\mu w)| dw = I_1 + I_2, \quad (6.17)$$

donde

$$\begin{aligned} I_1 &= \int_0^{w_*} [(\mu + \varepsilon) \exp(-(\mu + \varepsilon)w) - \mu \exp(-\mu w)] dw \\ &= -\left(\frac{\mu + \varepsilon}{\mu}\right)^{-\frac{\mu + \varepsilon}{\varepsilon}} + \left(\frac{\mu + \varepsilon}{\mu}\right)^{-\frac{\mu}{\varepsilon}} \\ &= \frac{\varepsilon}{\mu} \left[1 - \frac{1}{\frac{\mu + \varepsilon}{\varepsilon}}\right]^{\frac{\mu + \varepsilon}{\varepsilon}} \\ &\leq \frac{\varepsilon}{\mu} \exp(-1) \\ &\leq \frac{\varepsilon}{\mu} \end{aligned} \quad (6.18)$$

Desarrollando del mismo modo

$$I_2 = \int_{w_*}^{\infty} [\mu \exp(-\mu w) - (\mu + \varepsilon) \exp(-(\mu + \varepsilon)w)] dw$$

se puede ver que

$$I_2 \leq \frac{\varepsilon}{\mu}. \quad (6.19)$$

Juntando (6.14), (6.15), (6.17)-(6.19) establecemos que:

$$\Delta_p \leq 2^{\frac{s-1}{s}} kN \left[r^{\frac{1-s}{s}} + \mu^{\frac{1-s}{s}} \right] \varepsilon^{\frac{s-1}{s}}.$$

CAPÍTULO 7

CONCLUSIÓN Y ALGUNOS PROBLEMAS ABIERTOS

En esta tesis se buscaron las condiciones que permitieran establecer una desigualdad de estabilidad con "tasa de convergencia suficientemente buena". Los resultados son: Hipótesis formuladas en términos de los s -ésimos momentos de la "deriva" de los procesos, y, una cota superior para el índice de estabilidad en costo promedio (ver (4.1)). Las hipótesis dadas, también, permiten demostrar la existencia de políticas óptimas estacionarias. En contrario a algunos resultados establecidos, nuestra desigualdad de estabilidad es aplicable no sólo cuando los procesos tienen la propiedad de "ergodicidad geométrica", sino también, cuando admite procesos con "tasa de ergodicidad con una potencia $s - 1$ ". Por otro lado, si la potencia s en la hipótesis H.3 (ii), es suficientemente grande, entonces la potencia $\frac{s-1}{s}$ en la parte derecha de la desigualdad (4.1) puede ser cercana a 1, lo cual, da sentido a la frase del inicio de este párrafo: "tasa de convergencia suficientemente buena".

Discutiremos brevemente algunos problemas que no se abordan en este trabajo, pero que aparecen lógicamente como desarrollo de los resultados obtenidos. Parece interesante desarrollar un programa para computadora que pueda calcular cotas superiores para la constante N en (4.1). El paso siguiente puede ser escribir un programa que estime la precisión de la cota (4.1) en los sistemas de control particulares. i.e. por un lado calcular la parte derecha de la desigualdad (4.1), por otro lado, para algunos ejemplos calcular el valor exacto del índice de estabilidad Δ_p ; después comparar estas dos cantidades. Calcular exactamente los costos promedios $J(f_*)$, $J(\tilde{f}_*)$, y, por consiguiente, el valor de Δ_p ; esto último es posible, por ejemplo, para algunos modelos de control de inventario (ver [26], donde, se obtuvieron las fórmulas para las políticas estacionarias óptimas).

Otro problema abierto consiste en relajar las condiciones en las hipótesis H.2-H.4. Parece, que la existencia de un estado fijo x_* , con las propiedades descritas en H.2-H.4, es una condición demasiado restrictiva. A pesar, de que esta condición se satisface para algunos modelos de control: como por ejemplo, en un sistema de espera, en inventarios, en presas, etc; esta condición no se cumple, por ejemplo, para modelos lineales del tipo (2.1) (i.e. con la función Φ lineal, ver [22, 23]). Una idea razonable es tratar de cambiar H.2 por algunas condiciones de recurrencia, y, también, tratar de reemplazar $\{x_*\}$ en H.3 por un "conjunto pequeño" (ver [23]).

Es bien conocido, que en varios sistemas de control (lineales, de espera, etc.) es más conveniente escribir los costos por etapa $c(x, a)$ mediante funciones no acotadas. Los trabajos [3.10] contiene algunos resultados sobre estabilidad en promedio con

costos no acotados, pero, estos resultados en algunos sentidos son insatisfactorios. Por esta causa, es interesante extender las hipótesis y los resultados de esta tesis a procesos con costos no acotados.

El problema de generalizar estimaciones de la estabilidad, a otras clases de procesos controlables (por ejemplo, semi-markovianos) puede considerarse también.

BIBLIOGRAFÍA

- [1] Abbad, M. and Filar, J.A (1992), Perturbation and Stability theory for Markov control problems. **IEEE Trans. Automat. Contr.** **37**, 1415- 1420.
- [2] Ash, R.B (1972), **Real Analysis and Probability**, Academic Press, New York
- [3] Van Dijk , N.M. (1988), Perturbation theory for unbounded Markov reward processes with applications to queueing. **Adv. Appl. Prob.** **20**, 99-111.
- [4] Van Dijk , N.M and Puteman M. L. (1988) Perturbation theory for Markov reward processes with applications to queueing systems. **Adv. Appl. Prob.** **20** , 79-98
- [5] Dynkin, E. B. and Yushkevich, A.A. (1979), **Controlled Markov Processes**. Springer-Verlag, New York.
- [6] Ephremides, A and Verdú, S. (1989), Control and optimization methods in communication networks. **IEEE Trans. Automat. Contr.** **34**, 930-942.
- [7] Feller , W. (1978), Introducción en la Teoría de Probabilidades y sus Aplicaciones . Vol. II Ed. Limusa, México.
- [8] Gordienko, E. I. (1988), Stability estimates for controlled Markov chains with a minorant. **J. Soviet Math.** **40**, 481-486.
- [9] Gordienko, E. I (1992), An estimate of the stability of optimal control of certain stochastic and deterministic systems. **J. Soviet Math.** **59**, 891-899.
- [10] Gordienko, E.I. (1994), **Lectures Notes on Stability Estimation in Markov Decision Processes**. Universidad Autónoma Metropolitana, México D.F.
- [11] Gordienko, E.I. and Minjáles-Sosa A.J. (1988), Adaptive control for discrete-

time Markov processes with unbounden costs: average criterion. To appear in **Math Methods of Oper Research**.

[12] Gordienko, E. and Hernández-Lerma, O. (1995) Average cost Markov control processes with weighted norms: existence of canonical policies. **Appl Math** **23**, 199-128.

[13] Gordienko, E. and Hernández-Lerma, O. (1995) Average cost Markov control processes with weighted norms : value iteration. **Appl. Math.** **23**, 219-237.

[14] Gordienko, E.I., Isauro Martinez, M.E. and Marcos Carrillo, R.M. (1997), Estimation of stability in controlled storage systems. **Reporte de investigación**, 04.0405I.01.001.97, UAM-I, México D.F.

[15] Hernández-Lerma, O. (1989), **Adaptive Markov Control Processes**. Springer-Verlag, New York.

[16] Hernández-Lerma, O. and Lasserre, J:B. (1996), **Discrete-Time Markov Control Processes:Basic Optimality Criteria**. Springer-Verlag, New York

[17] Hernández-Lerma, O. and Cavazos -Cadena, R. (1990), Density estimation and adaptive control of Markov processes: average and discounted criteria. **Acta Appl. Math.** **20**, 285-307.

[18] Kalashnikov, V. (1994), **Mathematical Methods in Queueing Theory**, Kluwer Acad. Publ., Dordrecht.

[19] Kalashnikov, V.V. (1994), Regeneration and general Markov chains. **J. Appl. Math. Stoch. Analysis**, **7**, 357-371.

[20] Kalashnikov, V. V. (1994), Crossing and Comparison of regenerative processes. **Acta Appl Math.** **34**, 151-172.

- [21] Kitaev, M.Yu. and Rykov V.V (1995), **Controlled Queueing Systems**, CRC Press, Boca Raton.
- [22] Kumar , P.R. and Varaiya, P. (1986), **Stochastic Systems Estimation, Identification and Adaptive Control**.
- [23] Meyn, S. P. and Tweedie, R.L. (1993), **Markov Chains and Stochastic Stability**, Springer-Verlag, London.
- [24] Puterman, M.L. (1994), **Markov Decision Processes**, Wiley, New York.
- [25] Rachev, S.T. (1991), **Probability Metrics and the Stability of Stochastic Models**, Wiley, New York.
- [26] Vega- Amaya, O and Montes de Oca, R. (1996), Application of average dynamic programming to inventory systems . To appear in **Math. Methods Oper. Res.**
- [27] Weber, R.R. and Stidham, S. jr. (1987) Optimal control of service rates in networks of queues. **Adv. Appl. Prob. 19**, 202-218.
- [28] Whitt, W. (1978), Approximations of dynamic programs. I. **Math. Oper. Res. 3**, 231-243.
- [29] Yakowitz, S. (1982) Dynamic programming applications in water resources. **Water Resour. Res. 18**, 673-699.