Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

UNIVERSIDAD AUTÓNOMA METROPOLITANA – IZTAPALAPA
DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

ESTIMACIÓN DE LA ESTABILIDAD DE LOS PROCESOS CONTROLABLES DE MARKOV CON RESPECTO A LA METRICA DE PROKHOROV

Tesis que presenta
Jaime Eduardo Martínez Sánchez

Para obtener el grado de:
DOCTOR EN CIENCIAS (MATEMÁTICAS).

Director de tesis: Dr. Evgueni I. Gordienko

Jurado calificador:

Presidente: Dr. Francisco Sergio Salem Silva

Secretario: Dr. Julio César García Corte

Vocal: Dr. Juan Ruiz De Chavez Somoza

Vocal: Dr. Gabriel Escarela Pérez

Ciudad de México, 06 de abril de 2017

Casa abierta al tiempo
UNIVERSIDAD AUTÓNOMA METROPOLITANA

ACTA DE DISERTACIÓN PÚBLICA

No. 00050
Matrícula: 209280188

ESTIMACIÓN DE LA ESTABILIDAD DE LOS PROCESOS CONTROLABLES DE MARKOV CON RESPECTO A LA MÉTRICA DE PROKHOROV.

En la Ciudad de México, se presentaron a las 13:00 horas del día 6 del mes de abril del año 2017 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

DR. FRANCISCO SERGIO SALEM SILVA
DR. JUAN RUIZ DE CHAVEZ SOMOZA
DR. GABRIEL ESCARELA PEREZ
DR. JULIO CESAR GARCIA CORTE

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron a la presentación de la Disertación Pública cuya denominación aparece al margen, para la obtención del grado de:

DOCTOR EN CIENCIAS (MATEMATICAS)

DE: JAIME EDUARDO MARTINEZ SANCHEZ

y de acuerdo con el artículo 78 fracción IV del Reglamento de Estudios Superiores de la Universidad Autónoma Metropolitana, los miembros del jurado resolvieron:

Aprobar.

Acto continuo, el presidente del jurado comunicó al interesado el resultado de la evaluación y, en caso aprobatorio, le fue tomada la protesta.

JAIME EDUARDO MARTINEZ SANCHEZ
ALUMNO

REVISÓ

LIC. JULIO CESAR DE LARA ISASSI
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTOR DE LA DIVISIÓN DE CBI

DR. JOSE GILBERTO CORDOBA HERRERA

PRESIDENTE

DR. FRANCISCO SERGIO SALEM SILVA

VOCAL

DR. JUAN RUIZ DE CHAVEZ SOMOZA

VOCAL

DR. GABRIEL ESCARELA PEREZ

SECRETARIO

DR. JULIO CESAR GARCIA CORTE

DEDICATORIA

Este trabajo de investigación está dedicado a mis padres, Sr. Jaime Martínez y Sra. María Isabel Sánchez, por habeme encauzado por los caminos de la Educación.

A mis hermanos, Ligia Paola y Edgar Vladyvosky, por su entusiasmo y alegría en nuestras infinitas charlas de que este día llegaría!

JM

Abril de 2017

AGRADECIMIENTOS

Agradezco al Dr. Gordienko su confianza al haber aceptado dirigir este trabajo, así como sus consejos académicos y personales a través de estos años de colaboración.

A los sinodales, el Dr. Salem, el Dr. García, el Dr. Ruiz De Chavez, el Dr. Escarela y al Dr. González (q.e.p.d.), agradezco el tiempo y los comentarios realizados para que este trabajo se enriqueciera.

Finalmente, mi agradecimiento con todos los ciudadanos mexicanos fiscalmente responsables que hacen posible la existencia del Conacyt, y de universidades como la UAM-I, para realizar estudios de posgrado.

JM
Abril de 2017

Resumen

En este trabajo de investigación se estudian los procesos controlables de Markov en espacios de Borel, a tiempo discreto con horizonte infinito, con funciones de costos (o de recompensa) acotadas y usando los siguientes criterios: *costo esperado descontado*, *costo promedio*, *recompensa total esperada.*

Se plantea el problema de la estimación de la estabilidad para este tipo de procesos. El objetivo central de este trabajo es el de obtener unas cotas superiores para *el índice de estabilidad*.

Se proporcionan las condiciones suficiente para la existencia de las *desigualdades de estabilidad* para cada uno de los criterios antes mencionados usando la *métrica de Prokhorov*.

PALABRAS CLAVE. Procesos controlables de Markov a tiempo discreto, costo esperado descontado, costo promedio esperado, recompensa total esperada, índice de estabilidad, métricas probabilísticas.

INDICE

Algunos hechos preeliminares.

En este capítulo se presenta los elementos de la teoría y las definiciones necesarias para plantear el problema de control óptimo para los procesos controlables de Markov (PCM), también llamados procesos de decisión de Markov (MDP).

Una vez planteado el problema de control óptimo, se introducen las causas que originan el problema de la estimación del *índice de estabilidad*, ver su definición en (XVI). Las *desigualdades de estabilidad* para dicho índice, como las dadas en (XVII), se expresan en términos de alguna *métrica probabilística*, por lo que se proporciona unos conceptos básicos de la teoría de *métricas probabilísticas* en el espacio de distribuciones, así como una serie de relaciones existentes entre ellas.

Para mayores detalles y demostraciones de los resultados con respecto al problema de control óptimo que se presentan aquí, véase por ejemplo, Bersekas, Shreve (1978), Dynkin, Yushkevich (1979), o bien Hernández-Lerma, Lasserre (1996). Mientras que para una presentación de las numerosas aplicaciones a problemas teóricos y prácticos de las *métricas probabilísticas* véase Kalashnikov, Rachev (1988), Rachev (1991), o bien en Rachev, Rüschendorf (1998).

Al final del capítulo se introduce la definición de las tres métricas (*Prokhorov, Dudley y la de Variación Total* – ver definiciones en (XVIII, XX, XXII)-) con las que se trabaja en los siguientes capítulos de la tesis.

Procesos controlables de Markov a tiempo discreto y criterios de optimalidad.

Un proceso controlable de Markov a tiempo discreto con horizonte infinito, estacionario y homogéneo es una quíntupla:

$$M = \left(X, A, \{A(x) : x \in X\}, p, c\right), \tag{I}$$

donde X y A son espacios borelianos dados, llamados *espacio de estados* y conjunto de *control* (o acciones) respectivamente; $\{A(x) : x \in X\}$ es una familia de subconjuntos no vacíos (medibles), $A(x) \subset A$, con $A(x)$ representando el conjunto de *controles (acciones) admisibles* en el estado $x \in X$. Luego, p es un *kernel estocástico* (ver definición en el apéndice A) en X dado $\mathbb{K}$, donde $\mathbb{K} = \{(x,a) : x \in X, a \in A(x)\}$ es un subconjunto medible del espacio producto $X \times A$. Este *kernel estocástico* especifica la probabilidad de transición $p = \Pr\left(B \mid x,a\right)$, donde B pertenece a la σ-algebra de Borel en X , misma que se denotará por $\mathcal{B}_X$, y $\left(x,a\right) \in \mathbb{K}$. Finalmente, $c : \mathbb{K} \to R$ es una función medible llamada función de *costo de un paso*.

Cabe aclarar que en algunos problemas podría ser más conveniente considerar una función de *recompensa* (beneficio, ingreso) de un paso, dada por $r(x,a)$: $\mathbb{K} \to R$, en lugar de la función de costo en una etapa $c(x,a)$.

Los PCMs como el dado en (I) aparecen en diferentes áreas, desde ingeniería hasta economía, así como en control de poblaciones tales como los problemas en la administración de pesca, control de pestes, de epidemias, control de inventarios, modelos de líneas de espera, modelos de acumulación de capital, en problemas de selección de portafolio, en los procesos de seguros, en administración de recursos renovables y no renovables, etc.

Interpretación. El modelo de control M, dado en (I), representa un sistema estocástico controlable que es observado a un tiempo discreto $t = 0,1,...$. Si se denota por x_t y a_t , el estado del sistema y el control (o acción) aplicado al tiempo t, respectivamente, entonces la evolución del sistema puede ser descrito de la siguiente manera:

Si el sistema está en el estado $x_t = x \in X$ en el tiempo t , y se aplica el control $a_t = a \in A(x)$, entonces sucede lo siguiente:

(i) se incurre en un costo $c(x,a)$ (o bien en una recompensa $r(x,a)$); y

(ii) el sistema se mueve al siguiente estado x_{t+1}, el cual es un vector aleatorio que toma valores en X con distribución $p(B|x,a)$, es decir,

$$p = (B \mid x,a) = \Pr(x_{t+1} \in B \mid x_t = x, a_t = a)\ ,\ B \in \mathcal{B}_X. \qquad \text{(II)}$$

Una vez realizada esta transición al nuevo estado x_{t+1}, se elige un nuevo control y la dinámica del proceso se vuelve a repetir.

Por otro lado, en muchas aplicaciones la evolución del PCM está especificada por una ecuación en diferencias de la forma:

$$x_t = F(x_{t-1}, a_t, \xi_t)\ ,\ t = 1,2...\ \text{, con } x_0 \in X \text{ dado,} \qquad \text{(III)}$$

donde x_0 representa el estado inicial y $\{\xi_t\}$ es una sucesión de vectores aleatorios independientes e idénticamente distribuidos (*i.i.d.*) que toman valores en algún espacio de Borel S con una distribución común D_ξ .

En este caso el *kernel estocástico* p, dado en (II), queda expresado como:

$$p(B \mid x,a) = \int_S I_B[F(x,a,s)] D_\xi(ds)\ ,$$

o bien, $$p(B \mid x,a) = P(F(x,a,\xi_1) \in B).$$

En esta tesis se estudian dos PCMs que evolucionan a través del tiempo de la siguiente forma:

$$x_t = F\left(x_{t-1}, a_t, \xi_t\right) \ , \quad t = 1,\ldots; \qquad \text{(IV)}$$

$$\tilde{x}_t = F\left(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t\right) \ , \quad t = 1,\ldots; \qquad \text{(V)}$$

donde $F : K \times S \to X$ es una función medible; $x_t, \tilde{x}_t \in X$ son estados de los procesos; $a_t \in A\left(x_{t-1}\right)$, $\tilde{a}_t \in A\left(\tilde{x}_{t-1}\right)$ son acciones en el estado correspondiente; y $\{\xi_t\}$, $\{\tilde{\xi}_t\}$ son, respectivamente, dos sucesiones de vectores aleatorios *i.i.d.* con valores en S.

La única diferencia entre los procesos dados en (IV) y en (V) es posiblemente, las diferentes distribuciones D_ξ y $D_{\tilde{\xi}}$ de los vectores aleatorios ξ_t y $\tilde{\xi}_t$ respectivamente.

A lo largo de esta tesis se hará referencia a los procesos anteriores como *el proceso original*, dado en (IV), y como su *aproximación*, dado en (V), éste último también llamado proceso *perturbado*.

Políticas de control. Para cada $t = 0,1,\ldots$ considérese el espacio H_t como el historial admisible hasta el tiempo t. Sea $H_0 := X$, luego H_t para $t = 1,2,\ldots$ consta de los vectores de la forma $h_t = \left(x_0, a_0, \ldots, x_{t-1}, a_{t-1,} x_t\right)$, donde $(x_i, a_i) \in \mathbb{K}$ para $i = 0,1,\ldots t-1$, y $x_t \in X$.

Una *política de control* (randomizada, en general) es una sucesión $\pi = \left\{\pi_t, \quad t = 0,1,\ldots\right\}$ de *kernels estocásticos* π_t en el conjunto de control A dado H_t que satisfacen la siguiente restricción:

$$\pi_t\left(A(x_t)\middle| h_t\right) = 1 , \quad \text{para cada } h_t \in H_t \ , \ t = 0,1,\ldots \ . \qquad \text{(VI)}$$

El conjunto de todas las políticas de control se denotará por Π.

Cuando se usa una política de control randomizada como la dada en (VI), la acción en cualquier tiempo dado es una variable aleatoria, y depende del total t-histórico, h_t. Sin embargo, hay subclases importantes de políticas que determinan acciones que pueden ser usadas como funciones "determinísticas" de h_t y que podrían depender solamente del estado actual x_t. Un tópico central en teoría de control de Markov es encontrar condiciones para la existencia de políticas óptimas en esta clase restrictiva.

Sea $\mathscr{P}(A\,|\,X)$ el conjunto de todos los *kernels estocásticos* definidos en A dado X. Además, sea $\Phi=\{\varphi\in\mathscr{P}(A\,|\,X)\ |\ \varphi(A(x)|A)=1\}$ para toda $x\in X$. Finalmente se denotará por $\mathbb{F}$ al conjunto de todas las funciones medibles $f:X\to A$ satisfaciendo que $f(x)\in A(x)$ para toda $x\in X$.

Las funciones en $\mathbb{F}$ son llamadas selectores del multifuncional $x\mapsto A(x)$, (ver definiciones en apéndice A).

Una *política* $\pi=\{\pi_t\}\in\Pi$ se dice que es *estacionaria determinística*, si existe una función $f\in\mathbb{F}$ tal que $\pi_t(\cdot\,|\,h_t)$ está concentrada en $f(x_t)\in A(x_t)$ para toda $h_t\in H_t$ y $t=0,1,\ldots$

Es importante subrayar que el uso de $f\in\mathbb{F}$, significa la aplicación del control $a_t=f(x_{t-1})\in A(x_{t-1})$ si el proceso ocurrió en el estado x_{t-1}. De tal aplicación resulta un proceso de Markov homogéneo con valores en X, y *probabilidad de transición*

$$P_f(B|x)=P(B|x,f(x)),\ x\in X\ ,\ B\in\mathfrak{B}_X. \tag{VII}$$

De aquí en adelante, en esta tesis se tendrá como regla lo siguiente: Se usará la denotación $\mathbb{F}$ para la clase de políticas *estacionarias determinísticas* y a estas políticas deterministas se les llamará "estacionarias". Además, si una política π de

$\mathbb{F}$ está definida por la función f, es decir, $\pi = \{f, f, \ldots\}$, entonces para denotar a esta, se usará la misma letra f (en lugar de π).

Criterios de Optimalidad. Para la especificación del problema de control óptimo, además del sistema dinámico dado por las ecuaciones (IV) y (V) y un conjunto de políticas, se requiere de un criterio de desempeño (rendimiento) también conocido como función objetivo.

Para el problema de optimización de control con horizonte infinito $(t = 0,1,\ldots)$, entre los criterios de optimalidad que se encuentran en la literatura, en esta tesis se van a usar los siguientes:

(A) *Costo total esperado descontado*, definido por el siguiente funcional:

$$V_\alpha(x,\pi) := E_x^\pi \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t) \ , \tag{VIII}$$

donde $\alpha \in (0,1)$ es un coeficiente de descuento fijo; E_x^π denota la esperanza que corresponde a la distribución del proceso $\{x_t\}$ con el estado inicial $x \in X$ y la política de control $\pi = (a_0, a_1, \ldots, a_t, \ldots) \in \Pi$ aplicada.

(B) *Costo promedio por una unidad del tiempo,* dado por el siguiente funcional:

$$J(x,\pi) := \limsup_{n\to\infty} \frac{1}{n} E_x^\pi \sum_{t=0}^{n-1} c(x_t, a_t) . \tag{IX}$$

(C) *Costo total esperado*, definido por el funcional:

$$V(x,\pi) := \limsup_{n\to\infty} E_x^\pi \sum_{t=0}^{n} c(x_t, a_t) . \tag{X}$$

Si el problema de control involucra una *función de recompensa* $r(x,a)$, en lugar de una función de costo $c(x,a)$, entonces los criterios anteriores quedarían expresados de la siguiente manera:

A´) Recompensa total esperada descontada, definida por el funcional:

$$R_{\alpha}(x,\pi) := E_x^{\pi} \sum_{t=0}^{\infty} \alpha^t r(x_t, a_t)\ . \tag{XI}$$

B´) Recompensa promedio por una unidad del tiempo, dada por el funcional:

$$R_J(x,\pi) := \liminf_{n\to\infty} \frac{1}{n} E_x^{\pi} \sum_{t=0}^{n-1} r(x_t, a_t)\ . \tag{XII}$$

C´) Recompensa total esperada, definida por el funcional:

$$R(x,\pi) := \liminf_{n\to\infty} E_x^{\pi} \sum_{t=0}^{n} r(x_t, a_t)\ . \tag{XIII}$$

Nótese que en (A)-(C) y en A´)-C´) se asume la existencia de todas las esperanzas involucradas.

La motivación para estudiar problemas de costos descontados es principalmente económica. Procesos de acumulación de capital de una economía, problemas de inventario, de administración de portafolio son aplicaciones de este tipo de criterio.

El costo promedio es el criterio preferido en muchas aplicaciones, por ejemplo, en análisis de redes de comunicación y sistemas de líneas de espera.

El criterio de costo total esperado fué probablemente el primer problema de control con horizonte infinito estudiado en la literatura al menos desde principios de 1920. Aparece por ejemplo, en problemas de paro óptimo.

Para más aplicaciones prácticas de cada uno de estos tres criterios enfocados a costos, ver Hernández-Lerma, Lasserre (1996), así como las referencias que ahí se mencionan.

Los criterios referidos a recompensa se han utilizando en aplicaciones en Ciencias Financieras, tales como en problemas de consumo, problemas óptimos de saldos en efectivo (cash balance), problemas de inversión en mercados financieros, problemas de dividendos en teoría de riesgo. Para más aplicaciones ver Bäuerle, Rieder (2011).

Sea ahora $W(x,\pi)$ un criterio fijo seleccionado entre (VIII)–(X).

La función $W_*(x) := \inf_{\pi\in\Pi} W(x,\pi)$ con $x \in X$, se llama *la función de valor*, y una política de control π_* (si existe) se denomina *óptima* (con respecto al criterio W) si cumple lo siguiente:

$$W_*(x) = W(x,\pi_*) = \inf_{\pi\in\Pi} W(x,\pi) \ , \quad x \in X \ . \qquad \text{(XIV)}$$

Nota 0.1:

Si el costo $c(x,a)$ de una etapa es reemplazado por la función de recompensa $r(x,a)$ de una etapa, entonces el problema de control óptimo resultante (dado en (XIV)) es el de maximizar el criterio $W(x,\pi)$, seleccionado entre (XI)-(XIII).

En el desarrollo del tema de la tesis se van a usar las condiciones que garantizan la existencia de una *política óptima estacionaria* f_* para (XIV), (véase Hernández-Lerma, Lasserre (1996)).

El índice de estabilidad y el problema de su estimación. El problema de *estimación (cuantitativa) de estabilidad* (o “continuidad”, o “robustez”) surge cuando hay una incertidumbre acerca de la probabilidad de transición p, definida en (II), para el proceso (I). Supóngase que se tiene fijo un criterio de optimización $W(x,\pi)$ (de entre (VIII)-(X)). La tarea “original” del controlador consiste en la búsqueda (o aproximación) de la política óptima π_* que satisfaga (XIV) para el proceso “original” dado en (IV) con la probabilidad de transición $p = P(B \mid x,a)$ dada en (II). En muchas aplicaciones esta tarea no puede ser cumplida directamente debido a alguna de las siguientes causas:

1) Frecuentemente p o algunos de sus parámetros, son *desconocidos* para el controlador y esta probabilidad de transición se estima usando algunos procedimientos estadísticos (a partir de observaciones). Con los resultados

de estas estimaciones se genera otra probabilidad de transición $\tilde{p}$, que se interpreta como una aproximación accesible a la desconocida p.

2) Hay situaciones en las que p es conocida pero demasiado complicada como para tener una esperanza de resolver el problema de optimización de política de control. En tales casos, a veces p se reemplaza por *"una aproximación teórica"* $\tilde{p}$*,* que resulta en un proceso controlable con una estructura más sencilla.

En ambos casos, en la optimización de políticas el controlador tiene que trabajar con el proceso de Markov controlable $\tilde{M} = \left(X, A, \{A(x) : x \in X\}, \tilde{p}, c\right)$ definido por la probabilidad de transición accesible $\tilde{p}$. Cambiando x_t por $\tilde{x}_t$ en (VIII)–(X), se definen los correspondientes criterios de optimización $\tilde{V}_\alpha$, $\tilde{J}$, $\tilde{V}$ para el proceso aproximado $\tilde{M}$. Al denotar cualquiera de ellos por $\tilde{W}$, supongamos que es posible (al menos teóricamente) encontrar una política óptima $\tilde{\pi}_*$ para el proceso $\tilde{M}$, es decir,

$$\tilde{W}(x, \tilde{\pi}_*) = \tilde{W}_*(x) := \inf_{\pi \in \Pi} \tilde{W}(x, \pi) \ , \ \ x \in X \ \ . \tag{XV}$$

La política de control $\tilde{\pi}_*$ definida en (XV) se usa como la *aproximación* a la política no accesible óptima π_* (suponiendo que existe). En otras palabras, la política $\tilde{\pi}_*$ se usa para controlar el proceso original $M = \left(X, A, \{A(x) : x \in X\}, p, c\right)$ en lugar de la política π_*.

El aumento en el costo por tal aproximación se estima por medio del siguiente *índice de estabilidad*, (véase Gordienko, (1988,1992)):

$$\Delta_w(x) := W(x, \tilde{\pi}_*) - W_*(x) \equiv W(x, \tilde{\pi}_*) - W(x, \pi_*) \geq 0 \ , \ \ x \in X \ \ . \tag{XVI}$$

Nota 0.2:

Si el costo $c(x,a)$ de una etapa es reemplazado por la función de recompensa $r(x,a)$, entonces (XVI) se replantea como:

$$\Delta_w(x) := W_*(x) - W(x,\tilde{\pi}_*) \equiv W(x,\pi_*) - W(x,\tilde{\pi}_*) \geq 0 \ , \ \ x \in X \ .$$

El cual se interpreta como una reducción en la recompensa por el uso de tal aproximación.

El problema de *estimación de la estabilidad* consiste en la búsqueda de unas desigualdades del siguiente tipo (*desigualdades de estabilidad*):

$$\Delta_w(x) \leq K(x)\psi\left[\mu(p,\tilde{p})\right] \ , \ \ x \in X \quad , \qquad \text{(XVII)}$$

donde:

$\mu(p,\tilde{p})$ es cierta “distancia” entre las probabilidades de transición p y $\tilde{p}$ (expresada en los términos de una *métrica probabilística*),

$\psi : [0,\infty] \to [0,\infty]$ es una función continua tal que $\psi(s) \to 0$ cuando $s \to 0$, y $K(x)$, $x \in X$, es una función con valores calculados explícitamente.

La evolución de los procesos controlables $M = \left(X, A, \{A(x) : x \in X\}, p, c\right)$ y $\tilde{M} = \left(X, A, \{A(x) : x \in X\}, \tilde{p}, c\right)$ está especificada por las ecuaciones dadas en (IV) y en (V).

En esta tesis, la distancia $\mu(p,\tilde{p})$ involucrada en (XVII), mide la aproximación entre las distribuciones D_ξ y $D_{\tilde{\xi}}$ de los vectores aleatorios ξ y $\tilde{\xi}$ que aparecen en los procesos dados en (IV) y (V).

El objetivo central de esta tesis es el de proponer el uso de la *métrica de Prokhorov* (π) , véase la definición en (XVIII), para medir esta distancia, es decir, en la desigualdad dada en (XVII) se propone expresar $\mu(p,\tilde{p}) = \pi(D_\xi, D_{\tilde{\xi}})$.

Una parte de los resultados obtenidos en la tesis son *desigualdades de estabilidad* como la dada en (XVII), y son expresados de la siguiente manera:

$$\Delta_w(x) \leq K(x)\pi(D_\xi, D_{\tilde{\xi}}) \ , \ \ x \in X \ ,$$

donde π es la *métrica de Prokhorov* y $K(x)$ es una función con valores calculados explícitamente.

Métricas probabilísticas.

El desarrollo de la teoría de *métricas probabilísticas* empezó con la investigación de problemas relacionados con teoremas de límite en teoría de probabilidad.

En general, la aplicabilidad de teoremas de límite recae en el hecho de que la ley del límite pueda ser considerada como una aproximación para el modelo estocástico considerado, y pueda ser aceptada como una aproximación. La cuestión central recae entonces, en qué tan grande es el error en que caemos al adoptar el modelo aproximado.

Generalmente la teoría de *las métricas probabilísticas* estudia lo siguiente:

- En primer lugar, el problema de medir la distancia entre cantidades aleatorias. Por un lado, proporciona los principios fundamentales para construir *métricas probabilísticas*, es decir, los medios de medición de tales distancias.
- El segundo ámbito de estudio son las relaciones entre varias clases de *métricas probabilísticas*, es decir, problemas que requieren de una métrica particular, mientras que los resultados básicos pueden ser obtenidos en términos de otras métricas. En tales casos, las relaciones entre métricas son de vital importancia.

Sea $\mathfrak{X} = \mathfrak{X}(S)$ el conjunto de todos los vectores aleatorios definidos en el espacio de probabilidad (Ω, F, P) y con valores en el espacio métrico separable (S, r).

Se denotará a $\mathscr{P}(S)$ como el espacio de todas las distribuciones en $(S, \mathcal{B}_S)$ de vectores aleatorios de $\mathfrak{X}$, y $\mathcal{B}_S$ es la σ-álgebra de Borel en S.

Además, sea $\mathfrak{X}_2$ el conjunto de pares (X,Y), (de hecho de clases de equivalencia), de vectores aleatorios de $\mathfrak{X}$; y $\mathfrak{L}_2\mathfrak{X}_2$ denota el conjunto de todas las *distribuciones conjuntas* $P_{X,Y}$ de pares (X,Y) de $\mathfrak{X}_2$.

En el caso general, una *métrica probabilística* μ asigna a cada $P_{X,Y} \in \mathfrak{L}_2\mathfrak{X}_2$ un número de $[0,\infty]$. En la siguiente definición (y en el resto de toda la tesis) será más cómodo para $(X,Y) \in \mathfrak{X}_2$ escribir $\mu(X,Y)$, entendiendo esta expresión como $\mu(X,Y) := \mu(P_{X,Y})$.

Definición A. Un mapeo $\mu : \mathfrak{L}_2\mathfrak{X}_2 \to [0,\infty]$ se llama *métrica probabilística* si para cada uno de los vectores aleatorios X,Y,Z (con correspondientes distribuciones conjuntas $P_{X,Y}, P_{X,Z}$ y $P_{Y,Z}$) se cumplen las siguientes propiedades:

(i) $\mu(X,Y) \geq 0$.

(ii) $\mu(X,Y) = \mu(Y,X)$.

(iii) $\mu(X,Y) \leq \mu(X,Z) + \mu(Z,Y)$.

Definición B. Una *métrica probabilística* se llama *simple* si los valores $\mu(P_{X,Y})$ se determinan únicamente por las distribuciones *marginales* P_X y P_Y (de X y Y respectivamente).

Entonces *una métrica simple* μ define la distancia entre las distribuciones, es decir, elementos de $\mathscr{P}(S)$, y es una función $\mu : \mathscr{P}(S) \times \mathscr{P}(S) \to [0,\infty]$. En estos casos se escribirá $\mu(P_X, P_Y)$ o bien $\mu(X,Y)$.

Más abajo se proporcionan tres ejemplos de *métricas simples* (y solamente *métricas simples* se usan en esta tesis).

Un ejemplo sencillo de métrica que no es simple (sino compuesta) es la siguiente: $\mu(X,Y) := E|X-Y|$. Otro ejemplo, es *la métrica de Ky Fan* (véase su definición en el Apéndice B).

En el desarrollo de esta tesis, para el estudio del *índice de estabilidad* se usarán frecuentemente las siguientes tres métricas simples: la *métrica de Prokhorov* (π), la *métrica de Dudley* (d) y la *métrica de Variación Total* (**Vr**).

Para las siguientes tres definiciones recuerde que $X, Y \in \mathfrak{X}(S)$ con distribuciones de probabilidad $P_X, P_Y \in \mathscr{P}(S)$ respectivamente.

Definición 1. La *métrica de Prokhorov* (a veces, también llamada la *métrica de Levy-Prokhorov*) está definida como:

$$\pi : \mathscr{P}(S) \times \mathscr{P}(S) \to [0, \infty),$$

$$\pi(X,Y) \equiv \pi(P_X, P_Y) := \inf\{\ \varepsilon > 0 : P_X(A) \le P_Y(A^{\varepsilon}) + \varepsilon \quad para\ toda\ A \in \mathfrak{B}_S\}, \quad \text{(XVIII)}$$

donde $$A^{\varepsilon} := \{x : r(x,y) < \varepsilon\ \ para\ un\ \ y \in A\}.$$

Esta métrica fué definida por Prokhorov en 1956 como la análoga de la *métrica de Levy* (ver definición en apéndice B) para espacios más generales. Aún y cuando no es fácil de calcular, esta métrica es teóricamente importante porque metriza la convergencia débil en cualquier espacio métrico separable, ver demostración por ejemplo en Huber (1981).

Algunas de las relaciones que guarda la *métrica de Prokhorov* con otras métricas se pueden encontrar en Huber (1981), como por ejemplo la siguiente:

Para medidas de probabilidad en R, es decir, si $S = R$, se cumple que

$$L(P_X, P_Y) \le \pi(P_X, P_Y)\ ,$$

donde L es la *métrica de Levy*.

Mientras que en Gibbs, Su (2002) se demuestra que las *métricas de Prokhorov* π *y de Kantorovich* κ (ver definición en apéndice B) satisfacen la siguiente relación:

$$\pi^2 \le \kappa \le (diam(S) + 1)\pi\ , \quad \text{(XIX)}$$

donde $diam(S) = \sup\{d(u,v) : u, v \in S\}$ es el diámetro del espacio S.

Definición 2. La *métrica de Dudley* tiene la siguiente definición:

$$d : \mathscr{P}(S) \times \mathscr{P}(S) \to [0,2] \; ,$$

$$d(X,Y) \equiv d(P_X, P_Y) := \sup_{\varphi \in D} | E\varphi(X) - E\varphi(Y) | \quad , \qquad \text{(XX)}$$

donde:

$$D := \{\varphi : S \to R : \|\varphi\|_{\infty} + \|\varphi\|_{L} \le 1\} \;, \quad \text{y}$$

$$\|\varphi\|_{\infty} = \sup_{x \in S} |\varphi(x)| \quad ; \qquad \|\varphi\|_{L} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{r(x,y)} \; .$$

La *métrica de Dudley* es popularmente usada en el contexto de probar la convergencia de *medidas de probabilidad* con respecto a la topología débil.

La relación entre la *métrica de Prokhorov* y la *métrica de Dudley* fué demostrada por R. Dudley (consulte Dudley, (1989)) y está establecida por la siguiente desigualdad:

$$\frac{1}{2} d \le \pi \le \sqrt{\frac{3}{2}} \, d^{\frac{1}{2}} \; . \qquad \text{(XXI)}$$

Por lo tanto, las *métricas de Prokhorov y de Dudley* son equivalentes en este sentido fuerte, y ambas metrizan la convergencia débil.

Definición 3. La *distancia de Variación Total* está definida como:

$$\mathbf{Vr}(X,Y) \equiv \mathbf{Vr}(P_X, P_Y) := \sup_{\varphi \in B} | E\varphi(X) - E\varphi(Y) | \; , \qquad \text{(XXII)}$$

donde $\qquad B = \{\varphi : S \to R : \|\varphi\|_{\infty} \le 1\}.$

De las definiciones de la *métrica de Dudley* y de la *métrica de Variación Total* es inmediato que para $P_X, P_Y \in \mathscr{P}(S)$ se cumple la siguiente desigualdad:

$$d(P_X, P_Y) \le \mathbf{Vr}(P_X, P_Y).$$

De hecho, la *métrica* **Vr** es topológicamente más fuerte que la *métrica de Dudley*, (y por consiguiente, que la *métrica de Prokhorov*), si el espacio S es no-contable.

Huber (1981) prueba la siguiente cota:

$$\pi(X,Y) \leq \mathbf{Vr}(X,Y). \tag{XXIII}$$

La *métrica* **Vr** en $\mathscr{P}(S)$ puede ser considerada como la reducción para $\mathscr{P}(S)$ de la norma de Variación Total, denotada por $\|\cdot\|_{VT}$ **,** definida en el espacio lineal $\mathscr{M}$ de todas las medidas signadas en S por la fórmula ($m \in \mathscr{M}$):

$$\|\cdot\|_{VT} = m^{+}(S) + m^{-}(S) \ , \tag{XXIV}$$

donde

$m := m^{+} - m^{-}$ es la descomposición de Jordan-Hahn de la medida m .

Por lo que, para toda $P_X, P_Y \in \mathscr{P}(S)$ se cumple que:

$$\mathbf{Vr}\left(P_X, P_Y\right) := \| P_X - P_Y \|_{VT} \ . \tag{XXV}$$

Se sabe que:

$$\mathbf{Vr}\left(P_X, P_Y\right) := 2 \sup_{A \in B_S} | P_X(A) - P_Y(A) | \ . \tag{XXVI}$$

Informalmente, (XXVI) mide la mayor diferencia posible entre las probabilidades que dos distribuciones de probabilidad pueden asignar al mismo evento.

Dos casos particulares de la métrica de Variación Total.

1. Si $S = \{s_1, s_2, ...\}$ es un espacio contable, entonces la definición dada en (XXVI) se puede expresar como:

$$\mathbf{Vr}\left(P_X, P_Y\right) = \sum_{i=1}^{\infty} \left| P_X\left(\{s_i\}\right) - P_Y\left(\{s_i\}\right) \right| . \qquad \text{(XXVII)}$$

2. Si las distribuciones de probabilidad están valuada en los reales $\left(S \equiv R\right)$, y P_X, P_Y tienen densidades g_X, g_Y entonces, (XXVI) se puede expresar como:

$$\mathbf{Vr}\left(P_X, P_Y\right) = \int_R \left| g_X(x) - g_Y(x) \right| \, dx . \qquad \text{(XXVIII)}$$

En los siguientes capítulos se mostrará la forma de acotar el *índice de estabilidad* dado en (XVI), encontrando *desigualdades de estabilidad* como la dada en (XVII) en términos de alguna de estas tres métricas definidas. En el apéndice B se presenta un listado de las *métricas probabilísticas* que comúnmente son usadas en la teoría de probabilidad.

Introducción

Esta tesis trata sobre la aplicación de las *métricas de Prokhorov y de Variación Total* para encontrar cotas superiores para el *índice de estabilidad* de procesos controlables de Markov.

Se estudian los procesos controlables de Markov (PCM) a tiempo discreto con horizonte infinito, estacionarios y homogéneos, donde los conjuntos de estados y acciones son de Borel, dichos procesos los denotaremos por $M = \left(X, A, \{A(x) : x \in X\}, p, c\right)$, ver (I).

Supuestos generales. En el desarrollo de todos los capítulos de esta tesis se asumirá que los componentes del proceso controlable M tienen las siguientes características:

- El espacio de estados X, es un *espacio métrico* con una métrica ρ y $\mathcal{B}_X$ denota la *sigma-algebra* de Borel;
- el espacio de acciones A, es un *espacio métrico* con una métrica ℓ;
- el conjunto de acciones admisibles $A(x)$, es *compacto* para toda $x \in X$;
- el conjunto de pares estado-acciones admisibles $\mathbb{K} = \left\{(x,a) \in X \times A: \quad a \in A(x), \quad x \in X\right\}$ es un subconjunto no vacío de Borel del conjunto $X \times A$, y está equipado con la métrica $\nu = \max\{\rho, \ell\}$;
- Se compararán los procesos de control dados en (IV) y en (V), ver explicación dada en esa sección. Para facilitar la lectura, ambos procesos se transcriben a continuación:

proceso original, $x_t = F(x_{t-1}, a_t, \xi_t)$ para $t = 1,2,\ldots$,

proceso aproximado, $\tilde{x}_t = F(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t)$ para $t = 1,2,\ldots$.

Bajo el problema de estimar la estabilidad, se asume que el proceso $\tilde{x}_t$ es interpretado como una aproximación disponible para el proceso x_t.

- Se considera que los elementos aleatorios $\xi_1, \xi_2, \ldots$ y $\tilde{\xi}_1, \tilde{\xi}_2, \ldots$, de los procesos original y aproximado, respectivamente, son vectores aleatorios independientes e idénticamente distribuídos (*i.i.d.*) que toman sus valores en un espacio métrico separable (S, r) y tienen distribuciones D_ξ y $D_{\tilde{\xi}}$ respectivamente.

En la tesis se usan: el criterio de *costo total esperado descontado*, el criterio de *costo promedio por unidad del tiempo*, el criterio de *costo total esperado*, ver definiciones dadas en (VIII) - (X), así como los criterios de recompensa definidos en (XI) – (XIII).

El problema de la *estimación de la estabilidad* en términos generales se plantea como sigue:

Una vez definido el criterio de optimalidad para el *proceso original*, un primer problema es la búsqueda de una política óptima π_* para dicho proceso.

Sin embargo, este problema no puede ser resuelto directamente cuando la distribución D_ξ del vector aleatorio ξ en (IV) se supone que es desconocida. No obstante, dicha distribución puede ser aproximada por una distribución disponible $D_{\tilde{\xi}}$ del vector "perturbado" $\tilde{\xi}$ definido en el proceso de control "aproximado" dado en (V). Tal y como se explicó en la sección de preliminares, se puede considerar por ejemplo, el caso cuando $D_{\tilde{\xi}}$ es la distribución empírica obtenida de las observaciones independientes de los elementos aleatorios que tienen la distribución desconocida D_ξ.

Bajo ciertas condiciones (véase Hernández-Lerma, Lasserre (1996)) se puede garantizar la existencia de políticas óptimas para ambos procesos: “el original” y “el aproximado”, las cuales se denotarán por π_* y $\tilde{\pi}_*$ respectivamente.

Se admite que la política óptima $\tilde{\pi}_*$ puede ser encontrada para el proceso aproximado y que es usada para controlar el proceso original. El aumento en el costo por el uso de tal aproximación se estima con el *índice de estabilidad* dado en (XVI), cuando se usa el criterio W :

$$\Delta_w(x) := W(x,\tilde{\pi}_*) - W_*(x) \equiv W(x,\tilde{\pi}_*) - W(x,\pi_*) \geq 0 \ , \ x \in X \ .$$

Siguiendo el enfoque usado por ejemplo en los artículos de Gordienko (1988,1992), Gordienko, Yushkevich (2003), Montes-de-Oca, Salem-Silva (2005), Gordienko *et al.* (2008), el problema de *estimación de la estabilidad* (explicado en la sección de preliminares) consiste en la búsqueda de unas desigualdades como las dadas en (XVII) llamadas *desigualdades de estabilidad* :

$$\Delta \leq K\psi\left(\mu(D_\xi, D_{\tilde{\xi}})\right),$$

donde μ es cierta *métrica probabilística*.

Observación. Para dichas *desigualdades de estabilidad*, ¿Cómo se puede estimar (o acotar) la distancia $\mu\left(D_\xi, D_{\tilde{\xi}}\right)$ si D_ξ es desconocida?

En la mayor parte de los resultados obtenidos en esta tesis, la métrica μ es *la métrica de Prokhorov*, es decir, $\mu = \pi$ (véase la definición en (XVIII)).

Una cota superior para $\mu\left(D_\xi, D_{\tilde{\xi}}\right) = \pi\left(D_\xi, D_{\tilde{\xi}}\right)$ puede ser deducida en por lo menos las siguientes dos situaciones:

1) Cuando D_ξ se conoce, pero es reemplazada por alguna aproximación $D_{\tilde{\xi}}$, la cual produce un problema de control óptimo más simple.

2) Cuando D_ξ o alguno de sus parámetros son desconocidos, pero estimados usando algún procedimiento estadístico con observaciones disponibles $\xi_1, \xi_2, \ldots, \xi_n$.

Resultados obtenidos en la tesis.

I. Considerando funciones de costo *acotadas* y bajo el criterio de *costo total esperado descontado*, se obtuvieron las *desigualdades de estabilidad* (cotas superiores para el *índice de estabilidad*) expresadas en los términos de la *métrica de Prokhorov* (π), es decir, estimaciones para desigualdades como la dada en (XVII) con $\mu = \pi$ y con ψ siendo la función identidad.

   Ver capítulo 1.

II. Bajo ciertas condiciones de ergodicidad y de Lipschitz, con funciones de costo acotado, y usando el criterio de *costo promedio por unidad de tiempo*, se encontraron las *desigualdades de estabilidad* expresadas en términos de *la métrica de Prokhorov*. Este resultado fue presentado en un artículo de investigación que fue publicado en Kybernetika. Ver capítulo 2.

III. Considerando un proceso de decisión de Markov (MDP) transitorio y bajo *el criterio de recompensa total esperada* (ver (XIII) y la nota 0.2), se obtuvieron dos resultados para las *desigualdades de estabilidad* expresadas en términos de la *métrica de Prokhorov* y de la *métrica de Variación Total* para *el índice de estabilidad*. El motivo para usar la *métrica de Variación Total* fue el de poder utilizar condiciones menos restrictivas, y por lo tanto considerar clases más amplias de MDPs. Estos resultados fueron publicados en un artículo de investigación en Progress in Probability. Ver capítulo 3.

IV. Para un caso particular de un proceso controlable de Markov, y usando *el criterio de recompensa total esperada descontada*, (ver (XI) y la nota 0.2), se realizaron estimaciones numéricas (usando métodos estadísticos) para estimar el órden del *índice de estabilidad* dado en (XVI) con respecto al término $(1-\alpha)$. Se obtuvo que, bajo ciertas condiciones, el orden del *índice de estabilidad*, ver (XVI), es de $M(1-\alpha)^{-1}$ cuando $\alpha \to 1$. Ver capítulo 4.

Este trabajo de investigación consta de 5 capítulos, y está organizado de la siguiente manera:

En la sección anterior (“Algunos hechos preliminares”) se presentaron los elementos básicos que se requieren para plantear un problema de control óptimo, así como el problema de estimación del *índice de estabilidad*. Se incluyeron además, definiciones, relaciones y ejemplos de *métricas probabilísticas*.

En el capítulo 1 se presenta una revisión de los artículos de investigación que han sido publicados previamente y que están relacionados con el tema de la tesis, así como los resultados que se obtuvieron durante los cursos de Trabajo de Investigación I, II y III del programa doctoral. Estos resultados (ver el Teorema 1.1. y Teorema 1.2.) incluyen las primeras *desigualdades de estabilidad* encontradas en términos de la *métrica de Prokhorov* para un problema planteado usando el criterio de *costo total esperado descontado*. Se anexan las demostraciones de cada uno de los Teoremas que se encontraron, y finalmente se presentan dos ejemplos de aplicación: el primero (ver ejemplo 1.1.), referido a un sistema de líneas de espera bajo el supuesto de que $A(x) = A$ para todo $x \in X$; mientras que el segundo ejemplo de aplicación (ver ejemplo 1.2.) se refiere al proceso de regularización del nivel de agua en una presa, y se resuelve bajo el supuesto de que $A(x)$ puede depender de $x \in X$. En ambos ejemplos se validan los supuestos y luego se aplica el resultado obtenido en cada uno de los dos Teoremas encontrados. En el segundo ejemplo se aplican las relaciones existente entre las *métricas probabilísticas*, para establecer las *desigualdades de estabilidad* encontradas, en términos de diferentes *métricas probabilísticas*.

El capítulo 2 está basado en un artículo de investigación realizado por Martínez, Zaitseva (2015) y que fué publicado en la revista Kybernetica. Se usa el criterio de *costo promedio por unidad de tiempo* y bajo ciertas condiciones de ergodicidad y de *Lipschitz* se encuentra una cota superior que queda en términos de la *métrica de Prokhorov*, para *el índice de estabilidad* de un proceso general de Markov a tiempo discreto. El resultado principal de este capítulo es el Teorema 2.1.

Como ejemplo de aplicación se incluye un modelo de control para el proceso de regularización del nivel del agua de una presa (continuación del ejemplo 1.2. presentado en el capítulo 1). Además, se desarrolla a detalle un contraejemplo tomado de Gordienko *et al.* (2009) para mostrar la importancia de las condiciones de ergodicidad (ver *supuesto 2.1.*), y que cuando no se satisfacen, el resultado puede ser un proceso no estable. Al final del capítulo se incluye la prueba del Teorema 2.1.

En el capítulo 3, se presenta un ejemplo de aplicación acerca de consumo-invesión, por lo que a diferencia de los dos primeros capítulos, en éste es más conveniente usar el criterio de *recompensa total esperada*. Específicamente, se aplica este criterio a un proceso de decisión de Markov (MDP) transitorio. Los resultados presentados en este capítulo están basados en un artículo de investigación realizado por Gordienko, Martínez, Ruiz-de-Chavez (2015) publicado en Progress in Probability, Birkhäuser. Se prueban las desigualdades encontradas para estimar *el índice de estabilidad* con respecto a *la métrica de Variación Total* y a *la métrica de Prokhorov*. Estas desigualdades encontradas, se presentan en los Teoremas 3.1 y 3.2. Se muestran dos ejemplos de aplicación: uno, referido al problema de paro óptimo y el segundo, de consumo-inversión. En ambos ejemplos se busca acotar superiormente el *índice de estabilidad* usando la *métrica de Variación Total*.

En el capítulo 4 se presenta el análisis de estabilidad para un ejemplo particular tomado de Dynkin, Yushkevich (1979). En este caso se usa el criterio de *recompensa esperada descontada*, ya que el ejemplo involucra consumo e ingreso a lo largo del tiempo. Usando las fórmulas explícitas para las políticas estacionarias óptimas y para las funciones de valor, se calcula explícitamente el *índice de estabilidad* para el criterio descontado, y se investiga su comportamiento cuando el coeficiente de descuento se aproxima a uno.

Finalmente, en el capítulo 5 se presentan las conclusiones obtenidas en la tesis y se proporciona una breve reseña de un nuevo criterio de optimización llamado *sensible al riesgo* para el cual, se proponen una serie de problemas que se dejan abiertos para futuras investigaciones.

La tesis incluye tres apéndices. En el apéndice A se presentan algunas definiciones y resultados importantes que se usarán en los procesos controlables, en el apéndice B se muestra un listado de *las métricas probabilísticas* que comunmente se usan en la teoría de probabilidad, mientras que en el apéndice C se presentan las ecuaciones de optimalidad y las condiciones para la existencia de políticas estacionarias óptimas para los criterios usados en la tesis.

Capítulo 1.

Antecedentes y primeros resultados obtenidos para el problema de control óptimo usando el costo total esperado descontado.

En este capítulo, en la sección 1.1. se hace una revisión de la literatura, así como de los artículos de investigación que han sido publicados, y que guardan relación con el tema de la tesis.

Así mismo, en la sección 1.2. se presentan los primeros resultados obtenidos (bajo supuestos muy relajados) durante los cursos de Trabajo de Investigación I, II y III del programa doctoral. Estos resultados encontrados se presentan en el Teorema 1.1. y en el Teorema 1.2. de este capítulo, y muestran las primeras *desigualdades de estabilidad* obtenidas y expresadas en términos de la *métrica de Prokhorov*. En ambos Teoremas mencionados se usó el criterio de *costo total esperado descontado* y las condiciones de *Lipschitz* sobre cierta clase de funciones.

Puesto que la investigación de la estabilidad de procesos controlables tiene "inclinación hasta aplicaciones", una tarea natural de la tesis consiste en la búsqueda de nuevas aplicaciones de las *desigualdades de estabilidad* obtenidas. En otras palabras, es interesante estudiar ejemplos de unos modelos de control "reales" para los cuales se cumplen las condiciones de los Teoremas 1.1. y 1.2. y que sea posible estimar de alguna manera la *distancia de Prokhorov*, $\pi(D_\xi, D_{\tilde{\xi}})$, en las correspondientes desigualdades (como en (XVII)).

Como primer ejemplo, se considera un modelo de líneas de espera con servicio controlable (consulte Kitaev, Rykov (1995), Gordienko *et al.* (2008)). Como segundo ejemplo, se considera un modelo simple para el proceso de regularización del nivel de agua en una presa (consulte por ejemplo, Hernández-Lerma (1989), Asmussen (1987)).

1.1.- Revisión de artículos de investigación publicados.

El estudio sistemático así como el de aplicaciones de *métricas probabilísticas* básicamente inició con el famoso trabajo de Prokhorov (1956) donde, entre muchas otras cosas importantes introdujo la métrica π, ahora conocida como *la métrica de Prokhorov.*

La métrica π, “metriza” la *convergencia débil* y desde que apareció publicada en el artículo de Prokhorov (1956) ha sido usada en diversas aplicaciones, principalmente en la investigación de convergencia y aproximación de procesos estocásticos (con trayectorias interpretadas como vectores en un espacio métrico adecuado).

En los años sesenta, R. Dudley propuso otra famosa métrica probabilística d, ahora llamada *métrica de Dudley* (véase por ejemplo Dudley (1968)).

Dudley demostró (consulte Dudley (1989)) que las *métricas de Prokhorov y de Dudley* son equivalentes en un sentido fuerte.

A pesar de que *la métrica de Prokhorov* se usa con mayor frecuencia en el estudio de aproximación de procesos estocásticos (por ejemplo, por “procesos empíricos”), ocurrió que para propósitos de esta tesis la estructura de la *métrica de Dudley* tiene muchas ventajas (comparando con la de Prokhorov) para resolver los problemas de estimación de la estabilidad planteados en esta tesis. Sin embargo, la relación dada en (XXI) permite pasar a desigualdades correspondientes expresadas en términos de la *métrica de Prokhorov*.

Durante los años 70-90´s aparecieron muchos trabajos en los que se desarrolla la teoría general de *métricas probabilísticas* y sus numerosas aplicaciones a teoremas de límite de probabilidad, véase por ejemplo, Zolotarev (1976b), Senatov (1998), Rachev, Rüschendorf (1998), Rachev (1991), así como en estadística matemática, en particular en problemas de estimación por distribuciones empíricas, que tienen relación con el tema de esta tesis, véase por ejemplo Van der Vaart, Weller (1996), Dudley (1969), Rachev (1991), Gordienko *et al.* (2008). También se publicaron bastantes trabajos que incluyeron la investigación cuantitativa acerca de la continuidad (o “estabilidad”) de distribuciones de procesos estocásticos generales y aplicados bajo algunas *perturbaciones* de unos parámetros o distribuciones. Véase, por ejemplo Zolotarev (1976a, 1976b), Kalashnikov (1983,1995), Borovkov (1977), Rachev (1991).

Paralelamente han sido desarrollados los métodos de estimación de tasas de convergencia a distribuciones invariantes de procesos de Markov a tiempo discreto (no controlables, véase por ejemplo Tweedie (2001), Balaji, Meyn (2000), Meyn, Tweedie (2009), Asmussen (1987), Kartashov (1985)). Resultó que en esa área de investigación más útiles que las “métricas débiles” (como de Prokhorov o de Dudley) surgieron *métricas probabilísticas* que es posible denominar como “fuertes”; estas son del tipo de *la distancia de Variación Total* (o sus versiones ponderadas).

Acerca de procesos estocásticos a tiempo discreto *controlables*, las primeras aplicaciones de *métricas probabilísticas* (para estudiar su estabilidad bajo perturbaciones de distribuciones) aparecen en los trabajos de Van Dijk (1988), Van Dijk, Puterman (1988) para procesos en espacios discretos, y de manera independientemente en Gordienko (1985,1992) para procesos en espacios generales.

En los siguientes años se publicaron artículos sobre estimación cuantitativa de la estabilidad de procesos en espacios finitos, véase por ejemplo Van Dijk, Sladky (1999). Para procesos de Markov en espacios generales (de Borel), las investigaciones (cuantitativas) de estabilidad (“continuidad”, “robustez”, - terminología no establecida formalmente -) bajo perturbaciones de probabilidades de transición han sido realizadas en particular, en los siguientes trabajos: Muller (1997), Montes-de-Oca, Salem-Silva (2005), Favero, Runggaltier (2002) , Montes-de-Oca *et al.* (2003), Gordienko (1994), Godienko, Salem (1998,2000), Gordienko, Yushkevich (2003), Gordienko *et al.* (2008, 2009).

Además de varias *métricas probabilísticas* y sus propiedades, en los trabajos arriba mencionados se usó ampliamente la técnica de operadores contractivos en unos espacios funcionales, por ejemplo, en espacios con “normas ponderadas” (ver Wessels (1977)). Los operadores para los cuales se demuestra que tienen ciertas propiedades de “*contractividad*” normalmente aparecen en ciertas ecuaciones funcionales llamadas “*ecuaciones de optimalidad*”, consulte por ejemplo, Hernández-Lerma (1989), Hernández-Lerma, Lasserre (1996), Vega-Amaya (2003).

Resultó que establecer cotas superiores para el *índice de estabilidad* es relativamente fácil combinando las propiedades de contracción con las propiedades de las *métricas probabilísticas* fuertes. Por esa razón, en todos los artículos previamente mencionados, con excepción de Gordienko *et al.* (2008, 2009) usaron métricas de *Variación Total* o de *Variación Total Ponderada*.

No obstante, en situaciones importantes (y frecuentemente en aplicaciones) *las desigualdades de estabilidad* expresadas en las “métricas fuertes” son poco útiles cuando por ejemplo, las probabilidades de transición del proceso (desconocidas parcial o completamente) se aproximan por una estimación (como distribuciones empíricas) estadística.

Las distribuciones empíricas no convergen a la distribución original en la *métrica de Variación Total*, $\mathbf{Vr}$, pero sí convergen con respecto a la *métrica de Prokhorov*, π , o con respecto a la *métrica de Dudley*, d , y además, con una tasa conocida en muchos casos.

En los artículos Gordienko *et al.* (2008, 2009) se obtienen unas *desigualdades de estabilidad* en los términos de la métrica “débil” llamada la *métrica de Kantorovich,* κ (véase la definición en el Apéndice B). La convergencia en esta métrica equivale a la convergencia débil más la convergencia de los primeros momentos absolutos de vectores aleatorios, (véase Rachev, Rüchevdorf (1998)). El uso de esta métrica permite trabajar con funciones de costo no acotadas e implementar las aproximaciones por distribuciones empíricas.

En estos últimos artículos referidos (Gordienko *et al*. (2008, 2009)) se investiga la estabilidad con respecto a los criterios clásicos de optimalidad: El criterio de *costo total descontado* y el criterio de *costo promedio por unidad de tiempo*.

Procesos de control no estables. Para los criterios de costo descontado y de costo promedio, hay ejemplos de procesos de control como los dados en (IV) y en (V) que son *no estables*, es decir, tales que el *índice de estabilidad* no se aproxima a cero cuando la "distancia" entre las probabilidades de transición p y $\tilde{p}$ convergen a cero (véase Gordienko (1994), Gordienko, Salem (2000), Gordienko *et al.* (2009)).

En la sección 2.3 del siguiente capítulo se desarrolla a detalle un contraejemplo tomado de Gordienko *et al.* (2009) en el que se muestra un proceso de control que no es estable usando la *métrica de Kantorovich* κ. Específicamente este ejemplo muestra que $\Delta_J(x) \geq 1$ cuando $\kappa(D_{\xi}, D_{\tilde{\xi}}) \to 0$; donde $\Delta_J(x)$ es el *índice de estabilidad* usando el criterio de *costo promedio*.

1.2.- Planteamiento y resultados obtenidos usando el criterio de costo total esperado descontado.

En las siguientes secciones se consideran una subclase (de hecho, muy amplia) de procesos controlables de Markov que tienen la estructura explicada en (IV) y (V), las cuales se transcriben a continuación:

$$X_t = F(X_{t-1}, a_t, \xi_t)\ ,\ t = 1,2,... \quad \text{para el } \textit{“proceso original”}\ , \tag{1.1}$$

y su *aproximación accesible*,

$$\tilde{X}_t = F(\tilde{X}_{t-1}, \tilde{a}_t, \tilde{\xi}_t)\ ,\ t = 1,2,... \tag{1.2}$$

La única diferencia entre los procesos (1.1) y (1.2) consiste en las distribuciones diferentes D_ξ y $D_{\tilde{\xi}}$ de vectores aleatorios independientes e idénticamente distribuidos $\xi_1, \xi_2, ...$ y $\tilde{\xi}_1, \tilde{\xi}_2, ...$. Mientras que los espacios de estados y controles, así como los vectores aleatorios cumplen los **supuestos generales** mencionados en la introducción.

Los resultados encontrados y que se presentan a continuación, se obtuvieron bajo el supuesto de que $A(x) \equiv A$ para la sección 1.2.1., mientras que en la sección 1.2.2. se considera que $A(x)$ es general, es decir, puede depender del estado x.

1.2.1.- *Resultados obtenidos para el costo total esperado descontado. (Caso 1. Cuando $A(x) \equiv A$ no depende de los estados $x \in X$).*

Bajo los *supuestos 1.1. y 1.2.* que se introducen más adelante, se asegura la existencia de políticas óptimas estacionarias, $\pi_* = f_*$ y $\tilde{\pi}_* = \tilde{f}_*$, para los procesos dados en (1.1) y (1.2) usando el criterio de *costo total descontado* definido en (VIII).

Las funciones de valor (véase (XIV)) para el proceso original dado en (1.1) y para el proceso aproximado dado en (1.2), quedan definidas por V_* y $\tilde{V}_*$, respectivamente. Por lo que para medir el aumento en el costo por tal aproximación, se usa el *índice de estabilidad* definido en (XVI), y que bajo el criterio de *costo total descontado* se denotará por Δ_{V_α} .

Bajo las condiciones arriba mencionadas se procedió a resolver el problema de *estimación de la estabilidad*. El resultado encontrado fué una *desigualdad de estabilidad* del siguiente tipo:

$$\Delta_{V_\alpha}(x) \leq K_\alpha \pi(D_\xi, D_{\tilde{\xi}}) \quad , \quad x \in X \ , \tag{1.3}$$

donde para cada x ,

K_α : es una constante explícitamente calculable y que no depende del estado inicial x ;

$\pi(D_\xi, D_{\tilde{\xi}})$: es la *distancia de Prokhorov* dada en (XVIII) entre las distribuciones de ξ , $\tilde{\xi}$;

ξ , $\tilde{\xi}$: son vectores aleatorios *i.i.d.* como $\xi_1, \xi_2, \ldots$ y $\tilde{\xi}_1, \tilde{\xi}_2, \ldots$ respectivamente.

En esta tesis la cota dada en (1.3) se logró demostrar bajo ciertas condiciones de *Lipschitz*. La demostración usa las *propiedades contractivas* de operadores involucrados en las correspondientes “ecuaciones de optimalidad”. Es importante mencionar que bajo las condiciones usadas, dichos operadores en espacios funcionales conservan la “continuidad de *Lipschitz*” (véase Gordienko (1985), Hinderer (2005), Laraki, Sudderth (2004)) y esta peculiaridad es apta con la definición de la *métrica de Dudley*.

Además de los **supuestos generales** dados en la introducción, se requiere de los siguientes dos conjuntos adicionales de supuestos.

Supuesto 1.1.

(i) A es compacto y $A(x)=A$ para cada $x\in X$.

(ii) La función de costo por un paso $c(x,a):\mathbb{K}\to\mathrm{R}$ *es acotada*, es decir, $|c(k)|\leq b$ para cada $k\in\mathbb{K}$.

(iii) Para cada $x\in X$, la función de costos $c(x,\cdot)$ es *inferiormente semicontinua* en A.

(iv) Para cada función continua acotada $u:X\to R$, las funciones:

$$u'(x,a):=Eu\left[F(x,a,\xi)\right]\ ;$$

$$u''(x,a):=Eu\left[F(x,a,\tilde{\xi})\right],$$

son continuas en $\mathbb{K}$.

El segundo conjunto de supuestos (más restrictivo) impone las “condiciones de *Lipschitz*” en la función de costos de un paso y en las *probabilidades de transición* de los procesos (1.1) y (1.2).

Supuesto 1.2.

Existen constantes finitas L_1, L_2, L_3 tales que se cumple lo siguiente:

(i) $|c(x,a)-c(x',a)|\leq L_1\rho(x,x')$ para toda $x,x'\in X$, $a\in A$.

(ii) $E\rho\left(F(x,a,\xi),F(x',a,\xi)\right)\leq L_2\rho(x,x')$ para toda $x,x'\in X$, $a\in A$.

(iii) $\alpha L_2<1$.

(iv) $\rho\left(F(k,s),F(k,s')\right)\leq L_3 r(s,s')$ para toda $k\in\mathbb{K}$; $s,s'\in S$.

(v) Para cada $x \in X$, $s \in S$, la función $F(x,\cdot,s)$ es *inferiormente semicontinua* en A.

Recordando la definición (XIV), la *función de valor* para este problema particular usando el criterio de costo total esperado descontado, queda definida de la siguiente manera:

$$V_*(x) := V(x,\pi_*) := \inf_{\pi\in\Pi} V_\alpha(x,\pi) \ , \ x \in X \ , \qquad (1.4)$$

donde $V_\alpha(x,\pi)$ es el funcional dado en (VIII).

Lema 1.1. (Resultado obtenido).

Bajo los *supuestos 1.1. y 1.2.*, la función de valor V_* satisface las condiciones de *Lipschitz* con la constante $L = \dfrac{L_1}{1-\alpha L_2}$.

Demostración:

Sea $V_0 := 0$. Luego, defínase la siguiente función:

$$V_n(x) := \inf_{a\in A}\left\{c(x,a) + \alpha E V_{n-1}\left[F(x,a,\xi)\right]\right\}, \text{ para } \quad n = 1,2,...$$

El siguiente resultado es bien conocido: $V_n(x) \to V_*(x)$ uniformemente en $x \in X$ cuando $n \to \infty$, (véase Hernández-Lerma, Lasserre (1996)).

Por otro lado, bajo el *supuesto 1.1.* inciso (i) y *el supuesto 1.2.* inciso (i) se afirma que $V_1(x) \in Lip(L_1)$, es decir, V_1 satisface la condición de *Lipschitz* con la constante L_1.

Para ver lo anterior, por definición se tiene que

$$V_1(x) := \inf_{a\in A}\left\{c(x,a) + \alpha E V_0\right\} = \inf_{a\in A}\left\{c(x,a)\right\}.$$

Sean $x, y \in X$, y usando *el supuesto 1.2.* inciso (i), se obtiene que

$$|V_1(x) - V_1(y)| \leq \sup_{a \in A} |c(x,a) - c(y,a)| \leq L_1 \rho(x,y) .$$

Por lo que $V_1(x) \in Lip(L_1)$.

Se probará ahora que la función V_2 es también *Lipszchitz* con la constante $(1+\alpha L_2)L_1$.

Por definición, la función V_2 está dada por $V_2(x) := \inf_{a \in A} \{c(x,a) + \alpha E V_1[F(x,a,\xi)]\}$.

Ahora, para $x, y \in X$, se tiene que

$$|V_2(x) - V_2(y)| \leq \sup_{a \in A} \{ \ |c(x,a) - c(y,a)| \ + \alpha \, | E V_1[F(x,a,\xi)] - E V_1[F(y,a,\xi)] \, | \ \},$$

y usando *el supuesto 1.2.* inciso (i), de la desigualdad anterior se obtiene que

$$|V_2(x) - V_2(y)| \ \leq L_1 \rho(x,y) + \alpha \sup_{a \in A} E \, | \ V_1[F(x,a,\xi)] - V_1[F(y,a,\xi)] \ |,$$

y como V_1 es Lipschitz con constante L_1, se tiene que

$$|V_2(x) - V_2(y)| \ \leq L_1 \rho(x,y) + \alpha L_1 \sup_{a \in A} E \rho(F(x,a,\xi), F(y,a,\xi)).$$

Ahora, usando *el supuesto 1.2.* inciso (ii), se obtiene la siguiente desigualdad

$$|V_2(x) - V_2(y)| \ \leq L_1 (1 + \alpha L_2) \rho(x,y) \ , \ x, y \in X.$$

Por lo que $V_2(x) \in Lip((1+\alpha L_2)L_1)$.

Usando inducción, supongamos que para un $n \geq 1$ fijo, se cumple que $V_n(x) \in Lip((1+\alpha L_2 + \cdots + (\alpha L_2)^{n-1})L_1)$. Entoces para $x, y \in X$:

$$|V_{n+1}(x) - V_{n+1}(y)| \leq \sup_{a \in A} \{ \ |c(x,a) - c(y,a)| \ + \alpha \, | E V_n[F(x,a,\xi)] - E V_n[F(y,a,\xi)] \, | \ \}$$

$$\leq L_1 \rho(x,y) + \alpha (1 + \alpha L_2 + \cdots + (\alpha L_2)^{n-1}) L_1 \sup_{a \in A} E \rho(F(x,a,\xi), F(y,a,\xi))$$

$$\leq (1+\alpha L_2+\cdots+(\alpha L_2)^n)L_1\rho(x,y).$$

Por lo que $V_{n+1}(x)\in Lip\big((1+\alpha L_2+\cdots+(\alpha L_2)^n)L_1\big)$, para cada $n\geq 0$.

Usando el hecho de que $V_*(x)=\lim_{n\to\infty}V_{n+1}(x)$ uniformemente en $x\in X$, y por *el supuesto 1.2.* inciso (iii), se obtiene que para $x,y\in X$,

$$|V_*(x)-V_*(y)| \leq L_1\sum_{i=0}^{\infty}(\alpha L_2)^i\rho(x,y) = \frac{L_1}{1-\alpha L_2}\rho(x,y) \text{ , es decir, } V_*(x)\in Lip\left(\frac{L_1}{1-\alpha L_2}\right).$$

Por lo que V_* es *Lipschitz* para la constante $\dfrac{L_1}{1-\alpha L_2}$.

La siguiente definición se usará en la demostración del lema 1.2.

Definición 1.1. Sea $k\in\mathbb{K}$; defínase la función φ_k, como $\varphi_k(\cdot):=V_*[F(k,\cdot)]:S\to R$.

Lema 1.2. (Resultado obtenido)

Bajo los *supuestos 1.1. y 1.2.*, y para cada $k\in\mathbb{K}$, la función $\varphi_k(\cdot)$ satisface la condición de *Lipschitz* con constante $\dfrac{L_1L_3}{1-\alpha L_2}$.

Demostración:

Sean $k,k'\in\mathbb{K}$, $s,s'\in S$; por la definición 1.1., se tiene que

$$|\varphi_k(s)-\varphi_k(s')| = |V_*[F(k,s)]-V_*[F(k,s')]| ,$$

y por el lema 1.1., se llega a la siguiente desigualdad:

$$|\varphi_k(s)-\varphi_k(s')| \leq \frac{L_1}{1-\alpha L_2}\rho[F(k,s),F(k,s')].$$

Ahora, aplicando *el supuesto 1.2.* inciso (iv), de la desigualdad anterior se obtiene lo siguiente:

$$|\varphi_k(s)-\varphi_k(s')| \leq \frac{L_1L_3}{1-\alpha L_2} r(s,s').$$

$$\text{Es decir, } \varphi_k \in Lip\left(\frac{L_1L_3}{1-\alpha L_2}\right) \text{ en } S.$$

Se denotará por $\mathbb{B}$ al espacio de Banach de todas las funciones medibles $u : X \to R$, para las cuales la norma $\|u\|_\infty = \sup_{x\in X}\{\ |u(x)|\ \}$ es finita.

Para una demostración de la siguiente proposición, ver el teorema 4.2.3 (b) de Hernández-Lerma, Lassarre (1996).

Proposición 1.1.

Para ambos procesos de control (1.1) y (1.2), bajo los *supuestos 1.1. y 1.2.*, existen políticas óptimas estacionarias denotadas por $f_* = \{f_*, f_*, \ldots\}$ y $\tilde{f}_* = \{\tilde{f}_*, \tilde{f}_*, \ldots\}$ respectivamente.

Además, las funciones de valor correspondientes V_* , $\tilde{V}_* \in \mathbb{B}$, y por el lema 1.2. la función V_* es *Lipschitz*.

En particular, para cada $(x,a) \in \mathbb{K}$ fijo, las esperanzas $EV_*[F(x,a,s)]$ y $E\tilde{V}_*[F(x,a,s')]$ están bien definidas.

La proposición 1.1. permite reescribir el *índice de estabilidad* definido en (XVI), $\Delta_{V\alpha}$, de la siguiente manera:

$$\Delta_{V_\alpha}(x) = V_\alpha(x, \tilde{f}_*) - V_\alpha(x, f_*) \ , \ \ x \in X. \tag{1.5}$$

El primer resultado obtenido en la tesis para acotar el *índice de estabilidad* en términos de la *métrica de Prokhorov*, es el siguiente Teorema:

Teorema 1.1.

Bajo los s*upuestos 1.1. y 1.2.*, se tiene que

$$\sup_{x \in X} \ \Delta_{V_\alpha}(x) \leq K_\alpha \pi(D_\xi, D_{\tilde{\xi}}) \ , \tag{1.6}$$

donde:

π : es la *métrica de Prokhorov* definida en (XVIII), y

$$K_\alpha = \frac{4\alpha}{(1-\alpha)^2}\left[\frac{b}{1-\alpha} + \frac{L_1 L_3}{1-\alpha L_2}\right] \ . \tag{1.7}$$

Demostración:

El inicio de la demostración es igual a la primera parte de la prueba del Teorema 1 que aparece en el artículo Gordienko *et al.* (2008), por lo que a continuación se presentan las incorporaciones y desarrollos requeridos para la prueba del Teorema 1.1. de la tesis.

Sean $\pi = \{f, f, \ldots\}$, $\tilde{\pi} = \{\tilde{f}, \tilde{f}, \ldots\}$ las políticas óptimas estacionarias para los procesos (1.1) y (1.2) respectivamente, y V_*, $\tilde{V}_*$ las funciones de valor correspondientes.

Entonces (ver cap. 8 de Hernández-Lerma, Lasserre (1999)) V_* , $\tilde{V}_*$ y f, $\tilde{f}$ satisfacen las siguientes ecuaciones de optimalidad (más aún, V_* y $\tilde{V}_*$ son las únicas soluciones a estas ecuaciones):

$$V_*(x) := \inf_{a \in A(x)} \{c(x,a) + \alpha E V_*[F(x,a,\xi)]\}$$

$$= c\big(x, f(x)\big) + \alpha E V_*\left[F(x, f(x), \xi)\right] , \tag{1.7.1}$$

$$\tilde{V}_*(x) := \inf_{a \in A(x)} \left\{c(x,a) + \alpha E \tilde{V}_*\left[F(x,a,\tilde{\xi})\right]\right\}$$

$$= c\big(x, \tilde{f}(x)\big) + \alpha E \tilde{V}_*\left[F(x, \tilde{f}(x), \tilde{\xi})\right] . \tag{1.7.2}$$

Para toda $(x,a) \in \mathbb{K}$ se definen:

$$H(x,a) := c(x,a) + \alpha E V_*\left[F(x,a,\xi)\right] ,$$

$$\tilde{H}(x,a) := c(x,a) + \alpha E \tilde{V}_*\left[F(x,a,\tilde{\xi})\right] .$$

Como ha sido mostrado en Gordienko *et al.* (2008), el *índice de estabilidad* (1.5) se puede representar como:

$$\Delta_{V_\alpha}(x) := \sum_{i=1}^{\infty} \alpha^{t-1} E_x^{\pi} \Lambda_t \quad , \tag{1.7.3}$$

donde $$\Lambda_t := H\big(x_{t-1}, a_t\big) - \inf_{a \in A(x_{t-1})} H\big(x_{t-1}, a\big), \tag{1.7.4}$$

y $\{x_{t-1}, a_t,\ t \geq 1\}$ es la trayectoria del proceso (1.1) aplicando la política de control $\tilde{\pi} = \{\tilde{f}, \tilde{f}, \ldots\}$.

Por la definición de (1.7.2) y de (1.74) junto con el hecho de que $\tilde{\pi}$ es óptima para el proceso (1.2), se obtiene que

$$\Lambda_t = H\big(x_{t-1}, a_t\big) - \tilde{H}\big(x_{t-1}, a_t\big) + \inf_{a \in A(x_{t-1})} \tilde{H}\big(x_{t-1}, a\big) - \inf_{a \in A(x_{t-1})} H\big(x_{t-1}, a\big),$$

lo cual implica que

$$|\Lambda_t| \ \leq 2\alpha \sup_{a \in A(x_{t-1})} |\ H\big(x_{t-1}, a\big) - \tilde{H}\big(x_{t-1}, a\big)\ | ,$$

y por la definición de la funciones H y $\tilde{H}$, se tiene que

$$|\Lambda_t| \leq 2\alpha \sup_{a \in A(x_{t-1})} |\ E V_*\left[F(x_{t-1}, a, \xi)\right] - E \tilde{V}_*\left[F(x_{t-1}, a, \tilde{\xi})\right] | .$$

Luego, aplicando la desigualdad del triángulo, se llega a que

$$|\Lambda_t| \leq 2\alpha \sup_{a\in A(x_{t-1})} |\, EV_*\left[F(x_{t-1},a,\xi)\right] - EV_*\left[F(x_{t-1},a,\tilde{\xi})\right] | \quad +$$

$$+\, 2\alpha \sup_{a\in A(x_{t-1})} |\, EV_*\left[F(x_{t-1},a,\tilde{\xi})\right] - E\tilde{V}_*\left[F(x_{t-1},a,\tilde{\xi})\right] | \,. \qquad (1.7.5)$$

Sea:

$$\mu\left(D_\xi, D_{\tilde{\xi}}\right) := \sup\left\{ \,|\, EV_*\left[F(x,a,\xi)\right] - EV_*\left[F(x,a,\tilde{\xi})\right] | : (x,a) \in \mathbb{K} \right\}. \qquad (1.7.6)$$

Entonces de (1.7.6) se observa que el primer sumando del lado derecho de (1.7.5) está acotado por $2\alpha\mu\left(D_\xi, D_{\tilde{\xi}}\right)$.

Por otro lado, aplicando el supremo sobre x_{t-1} en (1.7.5), el segundo término del lado derecho queda expresado como

$$2\alpha \sup_{a\in A(x_{t-1})} |\, EV_*\left[F(x_{t-1},a,\tilde{\xi})\right] - E\tilde{V}_*\left[F(x_{t-1},a,\tilde{\xi})\right] | \leq 2\alpha\, \| V_* - \tilde{V}_* \|_\infty \,. \qquad (1.7.7)$$

En Hernández-Lerma, Lasserre (1999) se prueba que los siguientes operadores:

$$Tu(x) := \inf_{a\in A(x)} \left\{ c(x,a) + \alpha Eu\left[F(x,a,\xi)\right] \right\} , \qquad (1.7.8)$$

$$\tilde{T}u(x) := \inf_{a\in A(x)} \left\{ c(x,a) + \alpha Eu\left[F(x,a,\tilde{\xi})\right] \right\} ,$$

son contractivos en $\mathbb{B}$ con módulo α . Por lo que (1.7.1), (1.7.2) se pueden expresar como $V_* = TV_*$ y $\tilde{V}_* = T\tilde{V}_*$.

Ahora, dado que V_* y $\tilde{V}_*$ son puntos fijos de estos operadores, se obtiene que

$$\| V_* - \tilde{V}_* \|_\infty = \| TV_* - \tilde{T}\tilde{V}_* \|_\infty \ ,$$

o bien, aplicando la desigualdad del triángulo

$$\| V_* - \tilde{V}_* \|_\infty \leq \| TV_* - \tilde{T}V_* \|_\infty + \| \tilde{T}V_* - \tilde{T}\tilde{V}_* \|_\infty$$

$$\leq \ \|TV_* - \tilde{T}V_*\|_\infty \ + \ \alpha \, \|V_* - \tilde{V}_*\|_\infty$$

$$\leq \ \frac{1}{1-\alpha} \|TV_* - \tilde{T}V_*\|_\infty$$

$$\leq \ \frac{\alpha}{1-\alpha} \sup_{(x,a)\in K} \left\{ \ | EV_*[F(x,a,\xi)] - EV_*[F(x,a,\tilde{\xi})] | \ \right\} .$$

Usando la definición dada en (1.7.6), se obtiene que la desigualdad anterior se puede expresar como

$$\|V_* - \tilde{V}_*\|_\infty \ \leq \ \frac{\alpha}{1-\alpha} \mu\left(D_\xi, D_{\tilde{\xi}}\right).$$

Sustituyendo esta última expresión en (1.7.7), se obtiene que el segundo término del lado derecho de (1.7.5) está acotado por $\frac{2\alpha}{1-\alpha} \mu\left(D_\xi, D_{\tilde{\xi}}\right)$, y entonces (1.7.5) queda acotado por

$$\sup_X \ |\Lambda_t| \ \leq \ 2\alpha\mu\left(D_\xi, D_{\tilde{\xi}}\right) \ + \ \frac{2\alpha^2}{1-\alpha} \mu\left(D_\xi, D_{\tilde{\xi}}\right),$$

$$\leq \ \frac{2\alpha}{1-\alpha} \mu\left(D_\xi, D_{\tilde{\xi}}\right). \tag{1.7.9}$$

Finalmente, sustituyendo (1.7.9) en (1.7.3) se obtiene que

$$\sup_X \Delta_{V_\alpha}(x) := \sum_{i=1}^{\infty} \alpha^{t-1} \left\{ \frac{2\alpha}{1-\alpha} \mu\left(D_\xi, D_{\tilde{\xi}}\right) \right\}$$

$$\leq \ \frac{2\alpha}{(1-\alpha)^2} \mu\left(D_\xi, D_{\tilde{\xi}}\right). \tag{1.7.10}$$

Ahora, para acotar la parte derecha de (1.7.10) se usará la definición de la *métrica de Dudley* dada en (XX) y la definición dada en (1.7.6). Por el lema 1.2., la función de valor $V_* \in Lip\left(\frac{L_1 L_3}{1-\alpha L_2}\right)$ y por otro lado, es fácil ver que $\|V_*\|_\infty \ \leq \ \frac{b}{1-\alpha}$.

Entonces, por la definición de la *métrica de Dudley* y la definición de (1.7.6) se tiene que (1.7.10) se puede acotar de la siguiente forma:

$$\mu\left(D_{\xi}, D_{\tilde{\xi}}\right) \leq \left[\frac{b}{1-\alpha} + \frac{L_1 L_3}{1-\alpha L_2}\right] d\left(D_{\xi}, D_{\tilde{\xi}}\right).$$

Juntando esta última desigualdad con (1.7.10) y (XXI), finalmente se obtiene que

$$\sup_X \Delta_{V_\alpha}(x) \ \leq \ \frac{4\alpha}{(1-\alpha)^2}\left[\frac{b}{1-\alpha} + \frac{L_1 L_3}{1-\alpha L_2}\right] \pi\left(D_{\xi}, D_{\tilde{\xi}}\right).$$

Lo cuál demuestra el Teorema 1.1.

Ejemplo 1.1.
Sistema de línea de espera del tipo $GI|D|\infty$
con la tasa del servicio controlable

Se considera un proceso de control de Markov definido por la siguiente ecuación:

$$x_t = \left(x_{t-1} + a_t - \xi_t\right)^+ \quad , \qquad \text{para } t = 1,2,... \tag{1.8}$$

donde $x_t \in X := [0,\infty)$ es el tiempo de espera del t -ésimo cliente; $a_t \in A \equiv A(x) = [\gamma_0, \gamma_1] \subset (0,\infty)$, $0 \leq \gamma_0 < \gamma_1$, a_t es el tiempo (controlable) de servicio del t-ésimo cliente, y ξ_t es el intervalo entre llegadas del $(t-1)$-ésimo y el t-ésimo cliente. Se supone que $\xi_1, \xi_2,...$ son variables aleatorias *i.i.d.* positivas. La ecuación (1.8) describe el tiempo de espera en un sistema de colas $GI|D|\infty$ con una tasa de servicio controlable a_t . Si la distribución D_ξ de $\xi \equiv \xi_1$ es desconocida y se le aproxima por una distribución $D_{\tilde{\xi}}$ de $\tilde{\xi} = \tilde{\xi}_1$, entonces resulta *"un proceso aproximado"* que puede expresarse como:

$$\tilde{x}_t = \left(\tilde{x}_{t-1} + \tilde{a}_t - \tilde{\xi}_t\right)^+ \quad , \quad \text{para } t = 1,2,\ldots. \tag{1.9}$$

Se denotará por $c(x,a)$ el costo que se paga si el tiempo de espera es x y la duración del servicio es a. Se asume que c satisface el *supuesto 1.1.*, incisos (ii), (iii), así como el *supuesto 1.2.* inciso (i).

En el contexto del problema de estabilidad, el proceso (1.8) fué estudiado por Gordienko, Salem (1998) usando la *métrica de Variación Total*.

En esta tesis, para este ejemplo se verificarán los *supuestos 1.1.* y *1.2.* para el proceso (1.8) y así poder aplicar el resultado que se presentó en el Teorema 1.1. y poder establecer la estabilidad en términos de la *métrica de Prokhorov*.

Validación de los *supuestos 1.1. y 1.2*. Para comprobar la condición (ii) del *supuesto 1.2.*, obsérvese que la función $\varphi(z) = \max\{0, z\}$, $z \in R$, satisface la condición de *Lipschitz* con la constante 1. Entonces, para cualquier $a, s \in R$ fijos, la función de $x \in [0, \infty)$ dada por $[x + a - s]^+$ es de *Lipschitz* con la constante 1.

Lo anterior implica que $E\rho\left(F(x,a,\xi), F(x',a,\xi)\right) \leq \rho\left(x, x'\right)$, por lo que seleccionando la constante $L_2 = 1$ se cumple el *supuesto 1.2.* inciso ii).

Como $L_2 = 1$, es claro que $\alpha L_2 < 1$, por lo que la condición (iii) del *supuesto 1.2.* también se satisface.

La condición (iv) del *supuesto 1.2.* se verifica análogamente y resulta que $\rho\left(F(k,s), F(k,s')\right) \leq r(s,s')$. Es claro que para una constante $L_3 = 1$ se satisface este inciso (iv). El inciso (v) del *supuesto 1.2.* es evidente. La condición (iv) del *supuesto 1.1.* se verifica fácilmente.

Por lo tanto, aplicando el resultado obtenido en el Teorema 1.1., ver (1.6) y (1.7), se obtiene la siguiente cota para el *índice de estabilidad* en este ejemplo:

$$\Delta_{V_\alpha} \leq \frac{4\alpha\left(b + L_1\right)}{\left(1-\alpha\right)^3} \pi(D_\xi, D_{\tilde{\xi}}) .$$

1.2.2.- Resultados obtenidos para el costo total esperado descontado. (Caso 2. Asumiendo que $A(x)$ puede depender de $x \in X$.)

Para obtener una desigualdad como la dada en el Teorema 1.1., ver (1.6), para los procesos (1.1) y (1.2) bajo esta situación más general, ahora se requieren unas condiciones más restrictivas (comparando con la subsección anterior). Se necesitan, para mostrar que las soluciones de la ecuación de optimalidad (funciónes de valor) pertenecen a la clase de funciones de *Lipschitz*.

Supuesto 1.3.

Existen constantes finitas b , L_0 , L_1 , L_2 , L_3 tales que se cumple lo siguiente:

(i) $|c(k)| \leq b < \infty$ para cada $k \in \mathbb{K}$;

(ii) $|c(k) - c(k')| \leq L_0 \nu(k,k')$ para todo $k, k' \in \mathbb{K}$;

(iii) $h(A(x), A(x')) \leq L_1 \rho(x,x')$ para toda $x, x' \in X$, donde $h(\cdot.\cdot)$ es la *métrica de Hausdorff*, definida en el apéndice B.

(iv) $\mathbf{Vr}(F(k,s), F(k',s)) \leq L_2 \nu(k,k')$ para toda $k, k' \in \mathbb{K}$, $s \in S$ donde $\mathbf{Vr}(\cdot,\cdot)$ es *la métrica de Variación Total* definida en (XX);

(v) $\rho(F(k,s), F(k,s')) \leq L_3 r(s,s')$ para toda $k \in \mathbb{K}$ y $s, s' \in S$;

(vi) Para cada $x \in X$, y la función acotada $u : X \to R$, se tiene que la función $a \to Eu[F(x,a,s')]$ es inferiormente semicontinua en $A(x)$.

Lema 1.3. (resultado obtenido)

Bajo el *supuesto 1.3. ,* la función de valor V_* definida en (1.4) satisface la condición de *Lipschitz* con la constante $L = (L_1 + 1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]$.

Demostración:

Por el *supuesto 1.3.* inciso (i), para cada $\pi \in \Pi$ se tiene que $V(x,\pi) = E_x^f \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)$ está acotada por $\frac{b}{1-\alpha}$, entonces $\|V_*\|_\infty \leq \frac{b}{1-\alpha}$.

Por otro lado, de los operadores definidos en (1.7.8) se seleccionarán los términos que están adentro de los “corchetes” para definir la siguiente función:

$$g(x,a) \equiv g(k) := c(k) + \alpha E V_*[F(k,\xi)] \text{ , } k \in \mathbb{K}.$$

Se afirma que la función $g(k)$ es *Lipschitz* para la constante $\bar{L} = L_0 + \frac{\alpha b L_2}{1-\alpha}$. Para verlo, sean $k = (x,a)$, $k' = (x',a') \in \mathbb{K}$, entonces:

$$|g(k) - g(k')| = | c(k) + \alpha E V_*[F(k,\xi)] - c(k') - \alpha E V_*[F(k',\xi)] | ,$$

$$\leq |c(k) - c(k')| + \alpha | E V_*[F(k,\xi)] - E V_*[F(k',\xi)] |.$$

Aplicando el *supuesto 1.3.* inciso (ii) y el hecho de que $\|V_*\|_\infty \leq \frac{b}{1-\alpha}$ se tiene que

$$|g(k) - g(k')| \leq L_0 \nu(k,k') + \alpha \frac{b}{1-\alpha} \mathbf{Vr}(F(k,\xi), F(k'\xi)).$$

Luego, aplicando el *supuesto 1.3.* inciso (iv), la desigualdad anterior se puede expresar como

$$|g(k) - g(k')| \leq L_0 \nu(k,k') + \frac{\alpha b}{1-\alpha} L_2 \nu(k,k') ,$$

$$\leq \left[L_0 + \frac{\alpha b L_2}{1-\alpha} \right] \nu(k,k') \equiv \bar{L} \nu(k,k') .$$

Entonces, $g(k) \in Lip\left(\bar{L}\right)$ en $\mathbb{K}$.

Ahora, puesto que $V_* = TV_* = \inf_{a\in A(x)}\{g(x,a)\}$ (ver 1.7.8), para ver que la función V_* es *Lipschitz* con la constante $L = (L_1+1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right] \equiv (L_1+1)\bar{L}$, es suficiente probar lo siguiente:

Para la función $g : \mathbb{K} \to R$ tal que $|g(k)-g(k')| \leq \bar{L}\nu(k,k')$, $k,k' \in \mathbb{K}$, se sigue que para todo $x, y \in X$, $a \in A$:

$$\left|\inf_{a\in A(x)} g(x,a) - \inf_{a\in A(y)} g(y,a)\right| \leq \bar{L}(L_1+1)\rho(x,y). \tag{1.10}$$

Nota 1.1.:

Observe que $\left|\inf_{a\in A(x)} g(x,a) - \inf_{a\in A(y)} g(y,a)\right| =$

$$= \left|\inf_{a\in A(x)}\{c(x,a)+\alpha EV_*[F(x,a,\xi)]\} - \inf_{a\in A(y)}\{c(y,a)+\alpha EV_*[F(y,a,\xi)]\}\right|$$

$$= |V_*(x) - V_*(y)| .$$

Por lo que si se cumple la desigualdad dada en (1.10), entonces se cumple que

$$|V_*(x) - V_*(y)| \leq \bar{L}(L_1+1)\rho(x,y).$$

Por lo que se concluiría que V_* es Lipschitz en X .

A continuación se presenta la demostración de (1.10) en base a los argumentos dados en Gordienko *et al.* (2009).

Sea $q(x,y) := \left|\inf_{a\in A(x)} g(x,a) - \inf_{a\in A(y)} g(y,a)\right|$. Luego por la desigualdad del triángulo se obtiene lo siguiente:

$$q(x,y) \leq \left|\inf_{a\in A(x)} g(x,a) - \inf_{a\in A(x)} g(y,a)\right| + \left|\inf_{a\in A(x)} g(y,a) - \inf_{a\in A(y)} g(y,a)\right| ,$$

$$q(x,y) \leq \bar{L}\rho(x,y) + I ,$$

donde:

$$I = \left| \inf_{a\in A(x)} g(y,a) - \inf_{a\in A(y)} g(y,a) \right| . \qquad (1.10.1)$$

Se probará que $I \leq \bar{L} L_1 \rho(x,y)$. (1.10.2)

Si se asume que no se cumple la desigualdad dada en (1.10.2), es decir, que existe un $\varepsilon > 0$ tal que se satisface lo siguiente:

$$I > \bar{L} L_1 \rho(x,y) + \varepsilon . \qquad (1.10.3)$$

Debido a la compacidad de los conjuntos $A(x)$, $A(y)$ y a la continuidad de g, existen elementos $a_x \in A(x)$, $a_y \in A(y)$ para los cuales se alcanzan los ínfimos en I, ver (1.10.1).

Si se admite por ejemplo que

$$I = \inf_{a\in A(x)} g(y,a) - \inf_{a\in A(y)} g(y,a) \; > \; \bar{L} L_1 \rho(x,y) + \varepsilon \; ,$$

$$g(y,a_x) - g(y,a_y) \; > \; \bar{L} L_1 \rho(x,y) + \varepsilon \; ,$$

$$g(y,a_x) \; > \; g(y,a_y) + \bar{L} L_1 \rho(x,y) + \varepsilon . \qquad (1.10.4)$$

Ahora, por el *supuesto 1.3.* inciso (iii), como $h(A(x), A(y)) \leq L_1 \rho(x,y)$, existe $\bar{a}_x \in A(x)$ tal que $\ell(a_y, \bar{a}_x) \leq L_1 \rho(x,y)$ y consecuentemente se obtiene que

$$\begin{aligned} | g(y,\bar{a}_x) - g(y,a_y) | &\leq \bar{L} v(k,k'), \\ &\leq \bar{L} \ell(\bar{a}_x, a_y), \\ &\leq \bar{L} L_1 \rho(x,y). \end{aligned}$$

Lo anterior implica que

$$g(y,\bar{a}_x) \leq g(y,a_y) + \bar{L}L_1\rho(x,y),$$

$$g(y,a_y) \geq g(y,\bar{a}_x) - \bar{L}L_1\rho(x,y),$$

si se sustituye esta última desigualdad en (1.10.4), se obtiene que

$$g(y,a_x) \; > \; g(y,\bar{a}_x) + \varepsilon,$$

lo cual contradice el hecho de que $\bar{a}_x$ es el elemento para el cual el mínimo de $g(y,a)$ sobre $A(x)$ es alcanzado. Por lo tanto, el supuesto hecho en (1.10.3) es falso.

Entonces se tiene que $I \leq \bar{L}L_1\rho(x,y)$, lo cual implica que $q(x,y) \leq \bar{L}\rho(x,y) + \bar{L}L_1\rho(x,y)$. Y por lo tanto

$$q(x,y) \in Lip\left((1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]\right).$$

Finalmente, por lo comentado en la nota 1.1. se tiene que $V_* \in Lip\left((1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]\right)$, lo cuál demuestra el lema 1.3.

La función $\varphi_k(\cdot)$, introducida en la definción 1.1., se usará en el siguiente lema.

Lema 1.4. (Resultado obtenido)

Bajo el *supuesto 1.3.* la función φ_k satisface las condiciones de *Lipschitz* (en S) con una constante $(1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]L_3$.

Demostración: La demostración es parecida a la realizada para el lema 1.2.

Proposición 1.2. (Resultado bien conocido).

Bajo el *supuesto 1.3.* para los procesos de control dados en (1.1) y (1.2) existen políticas de control óptimas estacionarias denotadas por f_* y $\tilde{f}_*$ respectivamente.

La demostración de esta proposición se obtiene como consecuencia del Teorema 8.3.6 de Hernández-Lerma, Laserre (1999).

El segundo resultado obtenido en la tesis para acotar el *índice de estabilidad* en términos de la *métrica de Prokhorov*, es el siguiente Teorema:

Teorema 1.2.

Bajo el *supuesto 1.3., el índice de estabilidad* dado en (1.5) cumple la siguiente desigualdad:

$$\sup_{x\in X} \Delta_{V_\alpha}(x) \le K_\alpha \pi(D_\xi, D_{\tilde{\xi}}), \qquad (1.11)$$

donde:

$$K_\alpha = \frac{4\alpha}{(1-\alpha)^3}\left[b + L_3(1+L_1)(L_0(1-\alpha)+\alpha b L_2)\right]. \qquad (1.12)$$

Demostración:

La demostración de este Teorema es parecida a la dada para el Teorema 1.1. de la sección anterior. De hecho, bajo el *supuesto 1.3.* es perfectamente válido tomar la cota dada en (1.7.10) para el *índice de estabilidad*:

$$\Delta_{V_\alpha}(x) \quad \leq \quad \frac{2\alpha}{(1-\alpha)^2}\mu\left(D_\xi, D\tilde{\xi}\right),$$

donde $\mu\left(D_\xi, D_{\tilde{\xi}}\right)$ está definida en (1.7.6).

Para encontrar una cota para $\mu\left(D_\xi, D_{\tilde{\xi}}\right)$, se usará la definición de *la métrica de Dudley*. Por el lema 1.4. , se tiene que $V_* \in Lip\left((1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]L_3\right)$, y dado que $||V_*||_\infty \leq \frac{b}{1-\alpha}$, entonces el *índice de estabilidad* se puede acotar en términos de *la métrica de Dudley* por la siguiente expresión:

$$\Delta_{V_\alpha}(x) \quad \leq \quad \frac{2\alpha}{(1-\alpha)^2}\left[\frac{b}{1-\alpha} + (1+L_1)\left[L_0 + \frac{\alpha b L_2}{1-\alpha}\right]L_3\right]d\left(D_\xi, D_{\tilde{\xi}}\right).$$

Finalmente, usando la relación entre las *métricas de Dudley* y la de *Prokhorov* dada en (XXI) el Teorema 1.2. queda demostrado.

Ejemplo 1.2.
El proceso de regularización
del nivel del agua en una presa

Una importante aplicación de problemas de control (determinísticos y estocásticos) son los relacionados con operaciones de reserva de agua. Una excelente introducción para muchos de estos problemas, incluyendo la conección entre estos y sistemas de inventarios, es dada en Yakowitz (1982).

En el caso más simple de regularización del nivel de agua en una presa se puede hacer uso de la siguiente modelación para el *proceso original*:

$$x_t = \min\left\{x_{t-1} + \xi_t - a_t;\ U\right\} \quad , \quad t = 1,2,... \tag{1.13}$$

y el respectivo modelo "*aproximado*" queda como:

$$\tilde{x}_t = \min\left\{\tilde{x}_{t-1} + \tilde{\xi}_t - \tilde{a}_t;\ U\right\} \quad , \quad t = 1,2,... \tag{1.14}$$

En este modelo, la variable estado x_t representa el nivel del stock (volumen) de agua que tiene la presa al iniciar el período t; el control a_t es la cantidad de agua que se libera de la presa para consumo familiar, riego, energía eléctrica, etc. durante el período t; y la "perturbación" ξ_t es la cantidad de agua que recibe la presa, en forma aleatoria, vía lluvia por ejemplo.

Para este ejemplo $X = [0,U]$, $S = [0,\infty)$, $A(x) = [0,x]$, con $x \in X$, y donde U es la capacidad máxima de la presa.

Sea $0 \le c(x,a) \le b < \infty$ el costo que se paga por el servicio de agua liberada, por ejemplo, se puede hacer uso de una función de costo dada por $c(x,a) = c_0 a$, proporcional al consumo de agua, y donde c_0 representaría el costo de una unidad de agua.

Validación del *supuesto 1.3.* Para garantizar el cumplimiento del *supuesto 1.3.* para este ejemplo, se admite que se cumplen las siguientes condiciones:

- ✓ C1. El costo por un paso $c(x,a)$ satisface el *supuesto 1.3.* incisos (i), e inciso (ii).
- ✓ C2. La variable aleatoria ξ tiene una densidad g_ξ, la cuál es:
  a) acotada por una constante M_g; y
  b) satisface la condición de *Lipschitz* con una constante L_g.

Para $A(x)=[0,x]$, la condición (iii) en el *supuesto 1.3.* se verifica directamente (usando la definición de la *métrica de Hausdorff*) con la constante $L_1=1$.

Denotando $y:=x-a$, es fácil ver que para cada y fijo, la función $F(x,a,s)=\min\{y+s,U\}$ es *Lipschitz* en S con la constante 1. Entonces la condición (v) de este supuesto, se cumple con $L_3=1$.

A continuación se verificará la condición (iv) del *supuesto 1.3.*
Denotando $y:=x-a$, $y':=x'-a'$, $x,x'\in[0,U]$, $a\in[0,x]$, $a'\in[0,x']$, y considérese las siguientes variables aleatorias:

$$\zeta(y):=\min\{y+\xi,U\},$$

$$\zeta(y'):=\min\{y'+\xi,U\}.$$

Puesto que

$$|y-y'|\le|x-x'|+|a-a'|\le 2v(k,k'),$$

es suficiente demostrar que para una constante $\tilde{L}$ se cumple la siguiente desigualdad:

$$\mathbf{Vr}\ (\zeta(y),\zeta(y'))\le\tilde{L}|y-y'|\ . \qquad (1.15)$$

En su momento se verá que, según la definición de la métrica de *Variación Total* dada en (XXII), para demostrar (1.15) se debe probar que para cada función medible $\varphi:S\to R$, con $\|\varphi\|_\infty\le 1$ se cumple que

$$|E\varphi[\zeta(y)]-E\varphi[\zeta(y')]| \ \le \ \tilde{L}\,|y-y'|\,.$$

Se tiene que

$$E\varphi[\zeta(y)]=E\left\{\varphi[\zeta(y)]\ ;\ \ y+\xi<U\right\}+E\left\{\varphi[\zeta(y)]\ ;\ \ y+\xi\ge U\right\}\ ,$$

donde para una variable aleatoria η, se tiene que $E\left\{\eta\ ;\ \ A\right\}=E\eta I_A$.

Usando la misma representación para $E\varphi[\zeta(y')]$, se obtiene que

$$|E\varphi[\zeta(y)]-E\varphi[\zeta(y')]| \ \le$$

$$\le\ |E\left\{\varphi[\zeta(y)]\ ;\ \ y+\xi<U\right\}-E\left\{\varphi[\zeta(y')]\ ;\ \ y'+\xi<U\right\}|\ +$$

$$+\ \ |E\left\{\varphi[\zeta(y)]\ ;\ \ y+\xi\ge U\right\}-E\left\{\varphi[\zeta(y')]\ ;\ \ y'+\xi\ge U\right\}|\ \ ,$$

o bien,

$$|E\varphi[\zeta(y)]-E\varphi[\zeta(y')]| \ \le\ I_1(y,y')\ +\ I_2(y,y')\ . \tag{1.16}$$

Para el segundo término del lado derecho de la última desigualdad se tiene que

$$I_2(y,y') \ = \ |\ \varphi(U)P(\xi\ge U-y)\ -\ \varphi(U)P(\xi\ge U-y')\ |$$

$$=\ |\ \varphi(U)\,|\ |\int_{U-y}^{\infty}g_\xi(s)ds\ -\ \int_{U-y'}^{\infty}g_\xi(s)ds\ | \tag{1.17}$$

Sea por ejemplo $y>y'$. Entonces de (1.17) y de la condición C2 (véase página 50), se sigue que

$$I_2(y,y')\ \le\ \int_{U-y}^{U-y'}g_\xi(s)ds\ \ \le\ M_g\ |y-y'|\,. \tag{1.18}$$

Sea $z\in(0,U)$ un número arbitrario pero fijo, y dz que denote un intervalo infinitesimal con centro en z .

Puesto que

$$\{y+\xi\in dz\}\subset\{y+\xi<U\},$$

$$P(\xi \in dz - y, y + \xi < U) = P(\xi \in dz - y) = g_\xi(z-y)dz\ .$$

Similarmente

$$P(\xi \in dz - y', y' + \xi < U) = g_\xi(z - y')dz\,.$$

Entonces en (1.16) (tomando en cuenta que $g_\xi(x) = 0$ para $x < 0$),

$$I_1(y, y') = \left| \int_y^U \varphi(z) g_\xi(z-y)dz - \int_{y'}^U \varphi(z) g_\xi(z-y')dz \right| ,$$

o bien, suponiendo por ejemplo que $y > y'$ se tiene que

$$I_1(y, y') = \left| \int_y^U \varphi(z) g_\xi(z-y)dz - \int_{y'}^y \varphi(z) g_\xi(z-y')dz - \int_y^U \varphi(z) g_\xi(z-y')dz \right| \le$$

$$\le \left| \int_{y'}^y \varphi(z) g_\xi(z-y')dz \right| + \int_y^U |\varphi(z)|\,|g_\xi(z-y) - g_\xi(z-y')|\,dz \le$$

$$\le M_g |y - y'| + UL_g |y - y'|\ . \tag{1.19}$$

(Aplicando las condiciones C1 y C2).

Juntando (1.16), (1.18) y (1.19) se obtiene la desigualdad (1.15) con $\tilde{L} = 2M_g + UL_g$.

Finalmente se ha establecido que para este ejemplo la condición (iv) del *supuesto 1.3.* se cumple con $L_2 = 2(2M_g + UL_g)$. Siguiendo argumentos similares se puede mostrar que el inciso (vi) del *supuesto 1.3.* también se cumple.

Por lo tanto, en este ejemplo se puede aplicar la desigualdad (1.11) del Teorema 1.2., obteniéndose lo siguiente:

$$\sup_{x\in[0,U]} \Delta_{V_\alpha}(x) \le K_\alpha \pi(D_\xi, D_{\tilde{\xi}})\ . \tag{1.20}$$

donde

$$K_\alpha = \frac{4\alpha}{(1-\alpha)^3}\left[b + 2L_0(1-\alpha) + 4\alpha b(2M_g + UL_g)\right]. \tag{1.21}$$

Para un caso particular, en el que la función de costo esté dada por $c(x,a)=c_0 a$, se obtienen las constantes $b=c_0 U$, $L_0=0$, con las que se satisfacen el *supuesto 1.3.* y la constante K_α dada en (1.21) se convierte en

$$K_\alpha=\frac{4\alpha c_o U}{(1-\alpha)^3}\left[1+4\alpha\left(2M_g+UL_g\right)\right].$$

Por otro lado, la distancia $\pi(D_\xi, D_{\tilde{\xi}})$ dada en (1.20) es muy difícil de calcular. Por lo que, para hacer la parte derecha de la desigualdad (1.20) “más calculable”, se puede hacer uso de las relaciones entre las métricas que fueron mostradas en la sección “Algunos hechos preliminares”.

Por lo que, el resultado dado en (1.20) puede ser expresado en términos de otras *métricas probabilísticas*, como se muestra a continuación:

- Desigualdad (1.20) en términos de la *métrica de Variación Total*.
  Usando las relaciones (XXIII) y (XXVIII) se puede acotar la parte derecha de (1.20) por la siguiente *desigualdad de estabilidad*:

$$\sup_{x\in[0,U]}\Delta_{V_\alpha}(x)\le K_\alpha\int_0^\infty |\, g_\xi(s)-g_{\tilde{\xi}}(s)\,|\, ds \ , \qquad (1.22)$$

  donde la constante K_α está dada en (1.21).

- Desigualdad (1.20) en términos de la métrica de Kantorovich.
  Sean $(s\ge 0)$,

$$G_\xi(s)=\int_0^s g_\xi(z)dz \quad ; \quad G_{\tilde{\xi}}(s)=\int_0^s g_{\tilde{\xi}}(z)dz \ ,$$

  las funciones de distribución de las variables aleatorias ξ y $\tilde{\xi}$, respectivamente.

Entonces de la relación dada en (XIX) se obtiene que (1.20) se puede expresar por a siguiente *desigualdad de estabilidad*:

$$\sup_{x\in[0,U]} \Delta_{V_\alpha}(x) \le K_\alpha \left[\int_0^\infty | G_\xi(s) - G_{\tilde{\xi}}(s) | \, ds \right]^{1/2}. \tag{1.23}$$

donde la constante K_α está dada en (1.21).

La integral en la última desigualdad representa la *métrica de Kantorovich* entre ξ y $\tilde{\xi}$ (véase apéndice B). La desigualdad (1.23) es más informativa comparada con (1.22) puesto que admite la aproximación de G_ξ por las correspondientes funciones de distribución empíricas.

Capítulo 2.

Estimación de la estabilidad de un proceso controlable de Markov usando el criterio de costo promedio por unidad del tiempo.

Este capítulo está basado en el artículo Martínez, Zaitseva (2015), mismo que fué publicado en la revista Kybernetika. En dicho artículo se estudia la estabilidad del control óptimo de un proceso general de Markov a tiempo discreto, usando *el criterio de costo promedio por unidad de tiempo*.

A diferencia del capítulo anterior en el que se usó el criterio de *costo total esperado descontado*, V_α , cuando se utiliza el criterio de *costo promedio por unidad del tiempo*, J , para acotar el *índice de estabilidad* Δ_J , además de *"las condiciones de Lipschitz"* mencionadas en las secciones anteriores, se usarán condiciones bastante fuertes de *ergodicidad geométrica* de los procesos bajo la aplicación de políticas de control estacionarias.

En Gordienko (1994), Gordienko *et al.* (2009) se dan ejemplos de procesos de control *no estables* con respecto al costo promedio, en los casos cuando no se cumplen unas condiciones de ergodicidad.

Bajo ciertas condiciones de ergodicidad y de *Lipschitz*, el resultado que se obtuvo es una *desigualdad de estabilidad*, es decir, se encuentra que *el índice de estabilidad* está acotado por una constante explícita que multiplica a la *distancia de Prokhorov* entre las distribuciones de los vectores aleatorios que determinan el proceso controlado *"original" y el "perturbado"*.

2.1.- Especificación del planteamiento del problema de estabilidad usando el costo promedio.

A lo largo de este capítulo se mantendrán válidos los **supuestos generales** mencionados en la introducción y se compararán dos procesos controlables de Markov cuyas dinámicas están definidas por las ecuaciones (IV) y (V), mismas que se transcriben a continuación:

$$x_t = F(x_{t-1}, a_t, \xi_t) \qquad t = 1,2,... \qquad ; \tag{2.1}$$

$$\tilde{x}_t = F(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t) \qquad t = 1,2,... \; . \tag{2.2}$$

Como antes, las distribuciones de ξ_1 , $\tilde{\xi}_1$ se denotarán por D_ξ , $D_{\tilde{\xi}}$ respectivamente. Sea $c : \mathbb{K} \to \mathbb{R}$ una función mediable y acotada dada, que representa el costo de un paso, y $b := \sup_{(x,a)\in K} |c(x,a)|$.

Recordando la definición dada en (IX), para cada *estado inicial* $x \in X$, y una política de control $\pi \in \Pi$ (ver VI), *el costo promedio por unidad de tiempo* para cada uno de los procesos dados en (2.1) y (2.2) queda expresado, respectivamente, como sigue:

$$J(x,\pi) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_x^\pi c(x_{t-1}, a_t) \quad , \quad x \in X \quad ; \tag{2.3}$$

$$\tilde{J}(x,\pi) := \limsup_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_x^\pi c(\tilde{x}_{t-1}, \tilde{a}_t) \quad , \quad x \in X \; . \tag{2.4}$$

En este capítulo, los funcionales J y $\tilde{J}$ son las magnitudes a ser minimizadas. Una *política óptima* π_* (cuando exista), para el proceso dado en (2.1) es tal que

$$J(x,\pi_*) = \inf_{\pi\in\Pi} J(x,\pi) =: J_*(x) \quad , \quad x \in X \; ,$$

y una *política óptima* $\tilde{\pi}_*$ para el proceso dado en (2.2), es tal que

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

La siguiente proposición es bien conocida, se puede consultar en el corolario 3.6, de Hernández-Lerma (1989), o bien en Gordienko, Hernández-Lerma (1995).

Proposición 2.1.

Bajo los *supuestos 2.1. y 2.2.* que se presentan en la siguiente sección, existen políticas óptimas estacionarias f_* y $\widetilde{f}_*$, tales que $J(x, f_*)$ y $J(x, \widetilde{f}_*)$ no dependen de $x \in X$. Es decir, las funciones de valor $J_*(x) \equiv J_*$, $\widetilde{J}_*(x) \equiv \widetilde{J}_*$ son constantes, y además

$$J_* = J(f^*) = \inf_{\pi \in \Pi} J(x, \pi) \qquad ; \qquad \widetilde{J}_* = \widetilde{J}(f^*) = \inf_{\pi \in \Pi} \widetilde{J}(x, \pi) . \tag{2.5}$$

Por lo tanto, *el índice de estabilidad* dado en (XVI) para este criterio queda definido como

$$\Delta_J(x) := J\left(\widetilde{f}_*\right) - J\left(f_*\right) \geq 0 , \tag{2.6}$$

donde J está definido en (2.3).

El objetivo planteado en este capítulo es el de encontrar una cota superior para el *índice de estabilidad* $\Delta_J(x)$ dado en (2.6) en términos de la *métrica de Prokhorov*, π , es decir, obtener una desigualdad como la dada en (XVII), con una constante $K(x)$ calculada explícitamente y con $\psi(z) = z$, $z \in [0, \infty)$.

Las desigualdades como en (XVII) para el caso en el que $\mu = \mathbf{Vr}(\cdot,\cdot)$ fueron mostradas en Gordienko (1992), Gordienko, Yushkevich (2003), Montes-de-Oca, Salem-Silva (2005). El uso de estas *métricas probabilísticas* “fuertes” no permite estimar el *índice de estabilidad* Δ_J cuando por ejemplo, se usan aproximaciones por distribuciones empíricas.

En el artículo Gordienko *et al.* (2009) fué probado que para procesos con costos c posiblemente no-acotados y bajo el criterio de *costo promedio por unidad de* *tiempo*, *el índice de estabilidad* (XVI) podía acotarse por desigualdades (ver (XVII)) del siguiente tipo:

$$\Delta_J \le \overline{K}\,\kappa(D_{\xi}, D_{\tilde{\xi}}) \ , \tag{2.7}$$

donde $\overline{K}$ es una constante explícitamente calculada, y κ es la *métrica de Kantorovich* en el espacio R^n (ver apéndice B para la defininción), la convergencia en la métrica κ es equivalente a la convergencia débil más la convergencia de los primeros momentos absolutos (ver Rachev, Rüschendorf (1998)). La desigualdad (2.7) fué probada bajo la hipótesis de ergodicidad en términos de funciones estocásticas de *Lyapunov*, y usando ciertas condiciones de *Lipschitz*.

Desafortunadamente las condiciones de *Lyapunov* (ver el supuesto 1 en el artículo Gordienko *et al.* (2009)) usadas en dicho artículo para el *caso particular de un costo acotado* c, resultan en hipótesis de ergodicidad demasiado fuertes (conocidas como “condiciones de minorización”, ver el ejemplo dado en Dynkin, Yushkevich (1979) y en Gordienko (1988)).

El objetivo del artículo Martínez, Zaitseva (2015) y que se presenta en este capítulo bajo condiciones de ergodicidad menos restictivas bien conocidas (ver Arapostathis *et al.* (1993), Hernández-Lerma (1989)) dadas a continuación en este capítulo en el *supuesto 2.1.*, y para el caso de costo *acotado* c, es el de mostrar una *desigualdad de estabilidad* del siguiente tipo:

$$\Delta_J \le \hat{K}\pi(D_{\xi}, D_{\tilde{\xi}}) \ , \tag{2.8}$$

donde π es *la métrica Levy-Prokhorov* definida en (XVIII).

Como se mencionó al final de la sección 1.1. de esta tesis, existen procesos de control como los dados en (2.1) y (2.2) que son *no estables*. El ejemplo 1 presentado en Gordienko *et al.* (2009) ofrece PCMs como los dados en (2.1) y (2.2) con una función de costo acotado c, tal que $\kappa(D_{\xi}, D_{\tilde{\xi}}) \to 0$, donde κ es la *métrica de Kantorovich*, mientras que el *índice de estabilidad* dado en (2.7) se mantiene mayor a 1. Para este contra-ejemplo no se satisface el *supuesto 2.1.* que se introduce en la siguiente sección de este capítulo. Este contrajemplo se desarrolla a detalle en la sección 2.3.

2.2.- Supuestos y el resultado encontrado.

Para cada $k := (x,a) \in K$, $B \in \mathcal{B}_X$, sean:

$$p(B|k) := P(F(x,a,\xi_1) \in B) \quad ;$$

$$\tilde{p}(B|k) := P(F(x,a,\tilde{\xi}_1) \in B) \quad ,$$

las *probabilidades de transición* (kernels), definidas en (II), de los procesos (2.1) y (2.2) respectivamente.

Las condiciones de ergodicidad dadas en el *supuesto 2.1.* han sido extensamente usadas en la literatura de los procesos controlables de Markov bajo el criterio de *costo promedio por unidad de tiempo*, ver por ejemplo Arapostathis *et al.* (1993) y Hernéndez-Lerma (1989).

<u>*Supuesto 2.1.*</u> Condiciones de Ergodicidad.

Existe un número $\delta < 1$ tal que se cumple lo siguiente:

$$\sup_{k,k' \in K} \mathbf{V\!r}\left(p(\cdot \mid k), p(\cdot \mid k')\right) \leq 2\delta \quad ;$$

$$\sup_{k,k' \in K} \mathbf{V\!r}\left(\tilde{p}(\cdot \mid k), \tilde{p}(\cdot \mid k')\right) \leq 2\delta \, .$$

donde $\mathbf{V\!r}$ es la distancia de *Variación Total* definida en (XXII), y K es el conjunto de estados-acciones definido por $K := \{(x,a) \in X \times A, x \in X, a \in A(x)\}$.

El segundo conjunto de supuestos imponen "las condiciones de *Lipschitz*" en la función de costo de un paso y en las *probabilidades de transición* de los procesos (2.1) y (2.2).

<u>*Supuesto 2.2.*</u> Condiciones de *Lipschitz*.

Existen constantes finitas L_0, L, L_1 y L_* tales que, para toda $x, x' \in X$; $k, k' \in K$; $s, s' \in S$, se cumple:

(i) $h(A(x), A(x')) \leq L_0 \rho(x, x')$, donde h es la *métrica de Hausdorff,* definida en el apéndice B.

(ii) $|c(k) - c(k')| \leq L_1 \nu(k, k)$;

(iii) $\mathbf{Vr}\,(F(k, \xi_1), F(k', \xi_1)) \leq L\, \nu(k, k')$;

(iv) $\rho(F(k, s), F(k, s')) \leq L_* r(s, s')$;

(v) Para cada $x \in X$ y una función medible y acotada $u : X \to R$, el mapeo $a \to Eu\left[F\left(x, a, \tilde{\xi}\right)\right]$ es continuo en $A(x)$.

El principal resultado encontrado en el artículo Martínez, Zaitseva (2015) es el siguiente Teorema.

Teorema 2.1.

Bajo los *supuestos 2.1. y 2.2.*,

$$\Delta_J \leq \hat{K} \pi\left(D_\xi, D_{\tilde{\xi}}\right) , \qquad (2.9)$$

donde :

$$\hat{K} = 8\left(1 + \frac{2}{1-\delta}\right)\left[\left(1 + L_0\right)\left(L_1 + \frac{2bL}{1-\delta}\right)L_* + \frac{2b}{1-\delta}\right] , \qquad (2.10)$$

y $\pi(\cdot,\cdot)$ es la *métrica de Prokhorov*.

NOTA 2.1.:

Si el "parámetro de contractividad", δ, involucrado en el *supuesto 2.1.,* tiende a uno, entonces $\hat{K}$ en (2.10) es del orden $M(1-\delta)^{-2}$. Esto es mejor comparado con la constante $\overline{K}$ en la desigualdad de estabilidad obtenida en el artículo Gordienko *et al.* (2009) (usando la *métrica de Kantorovich*), véase (2.7), en donde dicha constante $\overline{K}$ es del orden $M(1-\delta)^{-3}$. Aunque parecería natural que el mejor orden posible debería de ser a lo mucho $M(1-\delta)^{-1}$.

Si se considera el caso cuando $S=R^m$ y usando la distribución empírica $D_{\tilde{\xi}}=\hat{D}_{\tilde{\xi},n}(\xi_1,\ldots,\xi_n)=\frac{1}{n}\sum_{i=1}^{n}\delta_{\xi_i}$, (donde δ_{ξ_i} es la *medida de Dirac*) como "la distribución aproximada" de $\tilde{\xi}_t$ en (2.2). La distribución $\hat{D}_{\tilde{\xi},n}$ se obtiene de realizaciones *i.i.d.* del vector aleatorio ξ_t en el proceso (2.1). El siguiente resultado es una consecuencia directa de la desigualdad dada en (2.9) y de la cota para la tasa de convergencia $E\pi(D_\xi,\hat{D}_{\tilde{\xi},n})$ obtenida en Dudley (1969).

Corolario.

Sea γ un número fijo tal que $\gamma>\max(2,m)$ y $\vartheta=\dfrac{m\gamma}{(\gamma-m)(\gamma-2)}$. Suponga que $E\|\xi_1\|^{\vartheta}<\infty$, y que $D_{\tilde{\xi}}=\hat{D}_{\tilde{\xi},n}(\xi_1,\ldots,\xi_n)$ en (2.9), entonces existe una constante finita $M<\infty$ tal que:

$$E\Delta_J\le Mn^{-1/\gamma} \quad , \quad n=1,2,\ldots$$

2.3.- Ejemplos.

Ejemplo 2.1. Modelo de control para el proceso de regularización del nivel del agua en una presa. (La continuación del ejemplo 1.2.)

En este modelo simplificado (véase Hernández-Lerma, (1989)) el proceso de control del nivel de agua en una presa es descrito de la siguiente manera:

$$x_t = \min\{x_{t-1} - a_t + \xi_t, U\} \quad , \quad t \geq 1 \ ,$$

$$\tilde{x}_t = \min\{\tilde{x}_{t-1} - \tilde{a}_t + \tilde{\xi}_t, U\} \quad , \quad t \geq 1 \ ,$$

donde al igual que en el ejemplo 1.2, $U < \infty$ es la capacidad máxima de la presa, $a_t \in A(x) := [0, x_t]$, $t \geq 1$ representa los niveles de consumo de agua y $\{\xi_t\}$, $\{\tilde{\xi}_t\}$ son variables aleatorias no negativas *i.i.d* que miden la entrada de agua (vía lluvia, por ejemplo). Correspondientemente, $X = [0, U]$, $S = [0, \infty]$ y al igual que en el ejemplo 1.2. se admite que se cumplen las condiciones C1 y C2 de dicho ejemplo (véase pág. 50).

Validación de los *supuestos 2.1. y 2.2.* Previa selección de una función de costo de un paso que cumpla que sea acotada, es decir, $|c(x,a)| \leq b < \infty$, y de *Lipschitz* con constante L_1. Por ejemplo, $c(x,a) := c_0 a$, donde $0 < c_0 < \infty$ es el costo de una unidad de agua.

(a) El *supuesto 2.2.* ya estaba verificado en el desarrollo del ejemplo 1.2.
Tomando las constantes $L_0 = 1$, $L = 2(2M_g + UL_g)$ y $L_* = 1$ se cumple este *supuesto 2.2.*

(b) Se asume que ambas variables aleatorias $\{\xi_1\}$ y $\{\tilde{\xi}_1\}$ tienen densidades g , $\tilde{g}$ correspondientemente, y que satisfacen la condición C2 dada en el ejemplo 1.2., y además son estrictamente positivas en algún intervalo abierto $(0, \Gamma) \supset (0, U)$. Se muestra que existe una constante positiva β tal que

$$p(B|x,a) \geq \beta \delta_U(B) \quad ; \quad \tilde{p}(B|x,a) \geq \beta \delta_U(B) \ , \qquad (2.10.1)$$

para toda $k \equiv (x,a) \in \mathbb{K} := \{(x,a) \in X \times A, x \in X, a \in A(x)\}$ y $B \in \mathcal{B}_{[0,U]}$. Aquí δ_U es la medida de Dirac. Si $U \notin B$, entonces $\delta_U(B) = 0$ y (2.10.1) se cumple para cualquier β .

Sea $U \in B$, es decir, $\delta_U(B) = 1$. Entonces (recuerde que $a \in [0, x]$),

$$P(B \mid x,a) = P(\min\{x - a + \xi, U\} \in B) \geq P(\min\{x - a + \xi, U\} = U) = P(x - a + \xi \geq U)$$

$$\geq P(\xi \geq U - x + a) \geq P(\xi \geq U) \geq \int_U^\Gamma g_\xi(s)ds =: \beta > 0.$$

De la misma manera se demuestra la segunda desigualdad en (2.10.1).

(c) Si se cumple el inciso anterior, entonces el lema 3.3 dado en Hernández-Lerma (1989) implica que el *supuesto 2.1.* se cumple.

Por lo tanto, los *supuestos 2.1. y 2.2.* se cumplen para este ejemplo y por esta razón los resultados del Teorema 2.1. son aplicables. Entonces, el *índice de estabilidad* (Δ_J) para este ejemplo queda acotado superiormente usando la *métrica de Prokhorov* (π) por la desigualdad (2.9), y usando las constantes L_0, L y L_* encontradas para el *supuesto 2.2.*, se obtiene explícitamente la constante dada en (2.10).

Por lo tanto, la *desigualdad de estabilidad* para este ejemplo queda expresada de la siguiente manera:

$$\Delta_J \leq \hat{K}\pi(D_\xi, D_{\tilde{\xi}}), \tag{2.11}$$

donde

$$\hat{K} = 16(3-\delta)\frac{(1-\delta)L_1 + 4b(M_g + UL_g) + 2b}{(1-\delta)^2}. \tag{2.12}$$

Si se usa la relación de métricas dada en (XIX) y (XXIII), entonces el *índice de estabilidad* (Δ_J) dado en (2.11) se puede acotar usando la *métrica de Variación Total* de la siguiente manera:

$$\Delta_J \leq \hat{K}\int_0^\infty |g_\xi(s) - g_{\tilde{\xi}}(s)|ds, \tag{2.13}$$

donde la constante $\hat{K}$ está dada en (2.12), y g_ξ, $g_{\tilde{\xi}}$ son las funciones de denisdad de ξ y $\tilde{\xi}$ respectivamente.

Contraejemplo 2.2. (Que muestra la importancia del *supuesto 2.1.*)

Sean $X = \{0,1,2,...\}$, $A(x) = A = \{0,1\}$, $x \in X$, $S = \{0,1,2,...\}$ y los procesos "original" y "aproximado" dados por las siguientes ecuaciones, respectivamente:

$$x_t = a_t x_{t-1} + \xi_t \quad , \tag{2.14}$$

$$\tilde{x}_t = \tilde{a}_t \tilde{x}_{t-1} + \tilde{\xi}_t \quad , \quad t = 1,2,... \tag{2.15}$$

donde $\{\xi_t, t \geq 1\}$ y $\{\tilde{\xi}_t, t \geq 1\}$ son dos sucesiones de variables aleatorias *i.i.d.* Se denotará $\xi_1 \equiv \xi_\varepsilon$ y $\tilde{\xi}_1 \equiv \tilde{\xi}$. Seleccionando $\tilde{\xi} \equiv 0$ y para cualquier $\varepsilon \in (0,1)$ la variable aleatoria $\xi = \xi_\varepsilon$ tiene la distribución de Poisson con parámetro ε. La función de costos de un paso $c(x,a) = c(x)$, $x \in X$, está definida de la siguiente manera:

$$c(x) = \begin{cases} 1 & si \quad x = 0, \\ 0 & si \quad x = 1, \\ 3 & para \quad x \geq 2. \end{cases}$$

Si se selecciona a $x_0 = \tilde{x}_0 = 1$ como el estado inicial, entonces es evidente que la política estacionaria $\tilde{f}_* = \{1,1,...\}$ es óptima-promedio para el proceso (2.15) con una función de valor de $\tilde{J}_*(1) = \tilde{J}(1, \tilde{f}_*) = 0$.

De hecho, aplicando esa política al proceso (2.15) se tiene que $x_t = 1$, $t = 0,1,2,\ldots$, y en el estado $x = 1$ el costo $c(x)$ toma su valor mínimo absoluto.

Por otro lado, aplicando la misma política $\tilde{f}_* = \{1,1,...\}$ pero ahora al proceso original dado en (2.14) se obtiene lo siguiente:

$$x_t = x_{t-1} + \xi_t = 1 + \xi_1 + \cdots + \xi_t \quad , \quad t \geq 1.$$

Por la ley fuerte de los grandes números $x_t = 1 + \xi_1 + \cdots + \xi_t \to \infty$ casi seguramente. Entonces, casi seguro $c(x_t) \to 3$ con $t \to \infty$; y por el Teorema de Convergencia Dominada $E_1^{\tilde{f}_*} c(x_t) \to 3$.

Por lo tanto, existe el siguiente límite $\lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} E_1^{\tilde{f}_*} c(x_{t-1}) = 3$, lo cual significa que

$$J(1, \tilde{f}_*) = 3. \tag{2.16}$$

Por otro lado, aplicando la política estacionaria $f = (0,0,...)$ al proceso original (2.14), se obtiene que

$$x_0 = 1 \text{ y } x_t = \xi_t \text{ , } t \geq 1.$$

Luego, para $t \geq 2$,

$$\begin{aligned} E_1^f c(x_{t-1}) &= 1 \cdot e^{-\varepsilon} + 3 \sum_{k=2}^{\infty} \frac{\varepsilon^k}{k!} e^{-\varepsilon} \quad , \\ &= e^{-\varepsilon} \left[1 + 3\varepsilon^2 \sum_{m=0}^{\infty} \frac{\varepsilon^m}{(m+2)!} \right], \\ &< 1 + 3\varepsilon^2 \text{ ,} \\ &< 2 \text{ para todo } \varepsilon \in \left(0, \frac{1}{\sqrt{3}} \right). \end{aligned}$$

Por lo anterior,

$$J(1, f) = \limsup_{n\to\infty} \frac{1}{n} \left[1 + \sum_{t=1}^{n} E_1^f c(x_t) \right] \text{ ,}$$

$$J(1, f) \leq 2. \tag{2.17}$$

Comparando (2.16) y (2.17) con la definición del *índice de estabilidad* $\Delta_J(1)$ en (2.6), se obtiene lo siguiente:

$$\Delta_J(1) = J(\tilde{f}_*) - J(f_*) \text{ ,}$$

$$\Delta_J(1) \geq J(\tilde{f}_*) - J(f) \geq 1, \tag{2.18}$$

donde (2.18) se cumple para toda $\varepsilon \in \left(0, \frac{1}{\sqrt{3}} \right)$.

Por otro lado, es fácil ver que $\kappa\left(D_{\tilde{\xi}}, D_{\xi_\varepsilon}\right) \to 0$ cuando $\varepsilon \to 0$. Obsérvese que por la definición de la *métrica de Kantorovich,* dada en el apéndice B, se tiene lo siguiente:

$$\kappa\left(D_{\tilde{\xi}}, D_{\xi_\varepsilon}\right) = \sup_{\varphi\in Lip_1}\left|E\varphi(\tilde{\xi}) - E\varphi(\xi_\varepsilon)\right| \leq \sup_{\varphi\in Lip_1} E\left|\varphi(\tilde{\xi}) - \varphi(\xi_\varepsilon)\right| \leq \sup_{\varphi\in Lip_1} E\left|\tilde{\xi} - \xi_\varepsilon\right| \leq E\xi_\varepsilon = \varepsilon .$$

Por lo que la distancia $\kappa\left(D_{\tilde{\xi}}, D_{\xi_\varepsilon}\right) \to 0$ cuando $\varepsilon \to 0$, mientras que el *índice de estabilidad* $\Delta_J(1)$ dado en (2.18) permanece siempre mayor o igual a 1.

La inestabilidad del proceso en este ejemplo, sucede debido a que el *supuesto 2.1.* no se cumple.

Para verlo, en el proceso "aproximado" (2.15) se selecciona por ejemplo, $k = (0,1)$ y $k' = (1,1)$. Entonces, $\tilde{p}(\cdot\,|\,(0,1)) = \delta_0(\cdot)$, $\tilde{p}(\cdot\,|\,(1,1)) = \delta_1(\cdot)$, (medidas de Dirac). También, $\mathbf{Vr}\,(\delta_0, \delta_1) = 2\sup_{B\in B_X} |\,\delta_0(B) - \delta_1(B)\,| = 2$, ya que por ejemplo, para $B = \{0\}$, $\delta_0(B) = 1$ y $\delta_1(B) = 0$.

Por lo que el *supuesto 2.1.* no se cumple para este ejemplo, lo que produce un proceso inestable.

2.4.- Prueba del teorema.

Para probar la desigualdad dada en (2.9), la base será el método propuesto en Gordienko *et al.* (2009). Sin embargo, se requiere modificar ésta técnica debido a que en esta tesis se usa la *seminorma span* en lugar de la norma uniforme ponderada usada en dicho artículo.

En Gordienko *et al.* (2009) la combinación de ciertas condiciones de *Lyapunov* y el resultado de Vega-Amaya (2003) permiten el uso de propiedades contractivas en los operadores relacionados a las ecuaciones de optimalidad. Resulta que estos operadores son contractivos con respecto a la norma uniforme ponderada en el espacio de las funciones $f : X \to R$.

Bajo el *supuesto 2.1.* se debe usar la *seminorma span* en el espacio de *funciones acotadas* y el hecho bien conocido (ver Hernández-Lerma (1989)) acerca de la contractividad con respecto a esta seminorma.

Recuérdese que $\mathbb{B}$ denota el espacio de todas las funciones medibles acotadas $u : X \to R$. Para $u \in \mathbb{B}$ se usará la norma del supremum $\|u\|_\infty := \sup_{x \in X} |u(x)|$, y la seminorma span $\|u\|_{sp} := \sup_{x \in X} u(x) - \inf_{x \in X} u(x)$. Es claro que $\|u + \beta\|_{sp} = \|u\|_{sp}$ para todo $\beta \in R$, y $\|u\|_{sp} \leq 2\|u\|_\infty$. En lo que resta de esta sección los vectores aleatorios ξ_1 y $\tilde{\xi}_1$ que aparecen en (2.1) y (2.2) serán denotados por ξ y $\tilde{\xi}$.

Defínase los siguientes operadores $T, \tilde{T} : \mathbb{B} \to \mathbb{B}$ como

$$Tu(x) := \inf_{a \in A(x)} \{c(x,a) - J_* + Eu[F(x,a,\xi)]\} \quad , \quad x \in X \tag{2.19}$$

$$\tilde{T}u(x) := \inf_{a \in A(x)} \{c(x,a) - \tilde{J}_* + Eu[F(x,a,\tilde{\xi})]\} \quad , \quad x \in X \ . \tag{2.20}$$

Es bien conocido (ver Hernández-Lerma (1989)) que bajo el *supuesto 2.1.*, para toda $u, v \in \mathbb{B}$, se cumple lo siguiente:

$$\|Tu - Tv\|_{sp} \leq \delta \|u - v\|_{sp} \qquad ; \qquad \|\tilde{T}u - \tilde{T}v\|_{sp} \leq \delta \|u - v\|_{sp} \ , \tag{2.21}$$

y por lo tanto existen funciones $h, \tilde{h} \in \mathbb{B}$ que son soluciones de las ecuaciones de optimalidad (2.19), (2.20), es decir,

$$h = Th \qquad ; \qquad \tilde{h} = \tilde{T}\tilde{h} \ . \tag{2.22}$$

Dado que para cualquier $\beta, \beta' \in R$ se tiene que $h+\beta$ y $\tilde{h}+\beta'$ también son soluciones de (2.22), entonces para el siguiente desarrollo se puede escoger h y $\tilde{h}$ de tal forma que

$$\|h\|_{sp} = \|h\|_{\infty} \quad ; \quad \|h-\tilde{h}\|_{sp} = \|h-\tilde{h}\|_{\infty} \ . \tag{2.23}$$

Luego, de (2.21) se sigue que

$$\|h-\tilde{h}\|_{sp} \leq 2\|Th-\tilde{T}h\|_{\infty} + \delta\|h-\tilde{h}\|_{\infty} \ ,$$

o bien, en vista de (2.19) y (2.20) se obtiene que

$$\|h-\tilde{h}\|_{\infty} \leq \frac{2}{1-\delta}\left[\left|J_* - \tilde{J}_*\right| + Q\right] \ , \tag{2.24}$$

donde

$$Q := \sup_{k\in K}\left|Eh\big(F(k,\xi)\big) - Eh\big(F(k,\tilde{\xi})\big)\right| \ . \tag{2.25}$$

Se denotará por $f := \tilde{\pi}_*$ a la política óptima estacionaria para el proceso dado en (2.2), y por $\{x_t, \ t\geq 0\}$ el proceso de Markov inducido por la aplicación de la política f aplicada al proceso controlable dado en (2.1). Ahora, usando la propiedad de Markov de $\{x_t\}$, y las ecuaciones de optimalidad dadas en (2.22), se obtiene que para todo $t\geq 1$, $x\in X$ se cumple lo siguiente:

$$\begin{aligned} E_x^f h(x) \ = \ & E_x^f h(x_{t-1}) \ - \ E_x^{\pi} c(x_{t-1}, f(x_{t-1})) \ + \ J_* \ + \\ & + \ E_x^f\left[H(x_{t-1}, f(x_{t-1})) - \inf_{a\in A} H(x_{t-1}, a)\right] \ , \end{aligned} \tag{2.26}$$

donde:

$$H(x,a) := c(x,a) + Eh\big[F(x,a,\xi)\big] - J_* \quad , \quad (x,a)\in \mathbb{K}. \tag{2.27}$$

Sumando en la ecuación (2.27) sobre los valores de $t = 1,\dots,n$, luego dividiendo por n y finalmente tomando el límite superior cuando $n\to\infty$, se obtiene que

$$J(x,f)=J_*+\limsup_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}E_x^f\left[H(x_{t-1},f(x_{t-1}))-\inf_{a\in A(x_{t-1})}H(x_{t-1},a)\right]. \qquad (2.28)$$

Similarmente, para (2.27) sea

$$\tilde{H}(x,a):=c(x,a)+E\tilde{h}[F(x,a,\xi)]-J_* \quad , \quad (x,a)\in\mathbb{K}. \qquad (2.29)$$

Ya que para el óptimo del proceso (2.2) en la política $\tilde{\pi}_*$, la acción $a_t=\tilde{f}(x_{t-1})$ alcanza el infimum de $\tilde{H}(x_{t-1},a)$ sobre $a\in A(x_{t-1})$, entonces se obtiene que en (2.28),

$$I_t=H(x_{t-1},f(x_{t-1}))-\inf_{a\in A(x_{t-1})}H(x_{t-1},a)=$$

$$=H(x_{t-1},f(x_{t-1}))-\tilde{H}(x_{t-1},f(x_{t-1}))+\inf_{a\in A(x_{t-1})}\tilde{H}(x_{t-1},a)-\inf_{a\in A(x_{t-1})}H(x_{t-1},a).$$

Usando la última ecuación anterior junto con (2.27) y (2.29), se llega a que

$$|I_t| \le 2|J_*-\tilde{J}_*|+2\sup_{a\in A(x_{t-1})}\left| Eh[F(x_{t-1},a,\xi)]-E\tilde{h}[F(x_{t-1},a,\tilde{\xi})]\right| , \qquad (2.30)$$

y observando (2.25) se tiene que el segundo término del lado derecho de (2.30) es menor que $2Q+2\|h-\tilde{h}\|_\infty$.

Comparando (2.6), (2.24), (2.28) y (2.30) se obtiene que

$$\Delta_J\le 2\left(1+\frac{2}{1-\delta}\right)\left[\ |J_*-\tilde{J}_*|+Q\ \right]. \qquad (2.31)$$

El siguiente paso es encontrar una cota superior para $|J_*-\tilde{J}_*|$ expresada en términos de Q. Escogiendo una sucesión $\{\alpha_n\}\subset(0,1)$, tal que $\alpha_n\uparrow 1$ y para cada n se define *el costo total esperado descontado* de la siguiente forma:

$$V_{\alpha_n}(x,\pi):=E_x^\pi\sum_{t=1}^{\infty}\alpha_n^{t-1}c(x_{t-1},a_t) \quad , \quad x\in X \quad , \quad \pi\in\Pi \quad ,$$

para el proceso (2.1). Sea $V_{\alpha_n}^*$ la función de valor correspondiente y $h_n(x) := V_{\alpha_n}(x) - V_{\alpha_n}(z)$, donde $z \in X$ es un estado arbitrario, pero fijo.

Aplicando la aproximación estándar del desvanecimiento de descuento (vanishing discount) y argumentos similares a los usados en la última parte de la prueba de la sección 4 dada en Gordienko *et al.* (2009), se llega a lo siguiente:

$$\left| J_* - \tilde{J}_* \right| \leq \limsup_{n \to \infty} \sup_{k \in K} \left| Eh_n \left[F(k, \xi) \right] - Eh_n \left[F(k, \tilde{\xi}) \right] \right|. \tag{2.32}$$

Es bien conocido (ver Arapostathis *et al.* (2008)) que el *supuesto 2.1.* implica que para cualquier política estacionaria f' y para cada $x \in X$, $t = 1,\ldots$ se cumple que

$$\left| E_x^{f'} c(x_{t-1}, f'(x_{t-1})) - \int_X c(x, f'(x)) q_{f'}(dx) \right| \leq 2b\delta^{t-1} \ , \tag{2.33}$$

donde $q_{f'}$ es la correspondiente probabilidad invariante.

De (2.33) es fácil ver que para cualquier $n \geq 1$, se tiene que $\|h_n\| \leq \dfrac{4}{1-\delta}$.

Además, es bien conocido (ver Hernández-Lerma (1989)) que para toda $x \in X$, $h_n(x) \to h'(x)$ uniformemente, donde h' es una solución para la ecuación de optimalidad $h' = Th'$ (ver (2.19)). Entonces, usando el teorema de convergencia acotada y (2.32) se obtiene fácilmente que

$$\left| J_* - \tilde{J}_* \right| \leq \sup_{k \in K} \left| \ Eh' \left[F(k, \xi) \right] - Eh' \left[F(k, \tilde{\xi}) \right] \ \right|. \tag{2.34}$$

Dado que la solución de la ecuación de optimalidad es única, incluyendo la adición de una constante arbitraria, comparando (2.25), (2.31) y (2.34) se obtiene que

$$\Delta_J \leq 4\left(1 + \frac{2}{1-\delta}\right) Q \ . \tag{2.35}$$

Ahora, usando (2.22), (2.23), (2.24) se obtiene que

$$\|h\|_\infty = \|h\|_{sp} \leq \|Th - T0\|_{sp} + \|T0\|_{sp}$$, o bien

$$\|h\|_\infty \leq \frac{1}{1-\delta}\left\|\inf_{a\in A(\cdot)} c(\cdot,a) - J_*\right\|_{sp} = \frac{1}{1-\delta}\left\|\inf_{a\in A(\cdot)} c(\cdot,a)\right\|_{sp} \leq \frac{2b}{1-\delta}. \qquad (2.36)$$

Entonces, en la definición de Q dada en (2.25) la función h está acotada por la constante $\frac{2b}{1-\delta}$.

El siguiente paso es mostrar que en (2.25) y por lo tanto en la desigualdad (2.35), las funciones $h[F(k,\cdot)]: S \to R$ satisfacen la condición de *Lipschitz* para una constante independiente de $k \in \mathbb{K}$.

En base a (2.19) se define la función g como

$$g(k) := c(k) + Eh[F(k,\xi)] \quad , \quad k \in \mathbb{K}.$$

Aplicando la definición dada en (XXII) para $\mathbf{Vr}(p, p')$ junto con el *supuesto 2.2.* incisos (ii) y (iii), se encuentra que para todo $k, k' \in \mathbb{K}$,

$$|\, g(k) - g(k') \,| \leq \left[L_1 + \frac{2bL}{1-\delta}\right]\nu(k,k'). \qquad (2.37)$$

Ahora, la definición del operador dado en (2.19) y la primera ecuación de (2.22) sugiere que para toda $x \in X$,

$$h(x) = \inf_{a\in A(x)}\left\{\, g(x,a) - J_* \,\right\}. \qquad (2.38)$$

En el artículo Gordienko *et al.* (2009) se pueba el siguiente lema:

Bajo la condición (2.37) y el *supuesto 2.2.* inciso (i), se tiene que para la función h dada en (2.38) se cumple que

$$|\, h(x) - h(x') \,| \leq (1+L_0)\left[L_1 + \frac{2bL}{1-\delta}\right]\rho(x,x') \text{ , para cada } x, x' \in X. \qquad (2.39)$$

Ahora, para $s, s' \in S$ fijos, por (2.39) y *el supuesto 2.2.* inciso (iv) se obtiene (ver (2.25)) que

$$\left| h(F(k,s)) - h(F(k,s')) \right| \leq (1+L_0)\left[L_1 + \frac{2bL}{1-\delta} \right] L_* r(s,s') \quad , \tag{2.40}$$

por lo tanto, para cada $k \in \mathbb{K}$, la función $h[F(k,\cdot)]: S \to R$ satisface la condición de de *Lipschitz* con la constante $\overline{L} = (1+L_0)\left[L_1 + \frac{2bL}{1-\delta} \right] L_*$, y por (2.36) está acotada por la constante $\frac{2b}{1-\delta}$, por lo que, usando la definición dada en (XX) se puede acotar *el índice de estabilidad* dado en (2.35) usando l*a métrica de Dudley*, obteniendo que

$$\Delta_J \leq 4\left(1 + \frac{2}{1-\delta}\right)\left(\overline{L} + \frac{2b}{1-\delta}\right) d\left(D_\xi, D_{\tilde{\xi}}\right).$$

Finalmente y para obtener la desigualdad (2.9) y la constante (2.10), se usa la relación dada en (XXI).

Capítulo 3.

Estimación de la estabilidad de un proceso de decisión de Markov transitorio usando el criterio de recompensa total esperada.

Los resultados presentados en este capítulo están basados en el artículo Gordienko, Martínez, Ruiz-de-Chávez (2015) de reciente publicación en Progress in Probability, Birkhäuser.

Se considera un proceso de decisión de Markov (MDP) absorbente (también llamado transitorio) a tiempo discreto con *recompensa total esperada*. Ver definición de proceso absorbente en James, Collins (2006) y en el *supuesto 3.1.* de esta tesis. Se prueban *desigualdades de estabilidad* con respecto a *la norma de Variación Total* y a *la métrica de Prokhorov*. La demostración de la desigualdad en términos de la métrica de *Variación Total* usa unas condiciones menos restrictivas comparadas con las que se usan en la demostración del caso de la *métrica de Prokhorov*.

Los principales resultados encontrados en el artículo están dados en los Teoremas 3.1. y 3.2. de este capítulo.

A lo largo de este capítulo se denotará a $r(x,a)$ como la función de recompensa de un paso, mientras que $R(x,\pi)$ denotará el criterio de optimización de *recomenpensa total esperada*, definido en (XIII). Finalmente, se denotará como Δ_R al *índice de estabilidad*, definido en (XVI), cuando se usa este criterio.

3.1. Motivación y planteamiento del problema.

En este capítulo se estudia la clase de modelos de control llamados transitorios, los cuales fueron introducidos por Veinott (1969) para el caso en el que X , el espacio de estados, así como A, el conjunto de acciones, son finitos.

Esta clase de proceso incluyen a los modelos descontados (ver proposición 9.6.3 de Hernández-Lerma (1999)) y bajo ciertas condiciones adecuadas (ver proposición 9.6.4 de Hernández-Lerma (1999)) están contenidos en la clase de modelos convergentes.

Algunos autores han extendido las definiciones y resultados de Veinott para el caso de cuando X , el espacio de estados, es numerable, pero muchas de estas extensiones dependen de la numerabilidad de X .

Los procesos de decisión de Markov (MDPs) transitorios estudiados en este capítulo pueden considerarse como "procesoso absorbentes". Tal terminología es adecuada puesto que en los artículos de James, Collins (2006) y en Kallenberg (1983), está demostrado que para cada política de control estacionaria que induce un proceso de Markov transitorio, la dinámica está dada por un kernel pseudo-probabilístico. Tal kernel se puede ampliar a un *kernel probabilístico* añadiendo un estado absorbente.

El uso del término "transitorio" es más común en la literatura sobre MDPs, véase por ejemplo, Hernández-Lerma *et al.* (1999), Hernández-Lerma, Lasserre (1999), Hordijk (1974), James, Collins (2006), Kallenberg (1983), Pliska (1978), y Veinott (1969).

En su artículo, Hinderer, Waldmann (2005) usan el término "absorbente" para MDPs numerables. Bertsekas, Tsitsiklis (1991) estudian MDPs numerables similares pero con diferente terminología.

En este capítulo se presentan ejemplos de aplicación tales como el de consumo-inversión, y un ejemplo clásico acerca del problema de paro óptimo, por lo que conviene trabajar usando el criterio de *recompensa total esperada* en un paso, definida en (XIII), en lugar del costo por un paso usado en capítulos anteriores.

Considérese MDPs como el definido en (I) del capítulo preliminar. Se asumirá que sus componentes satisfacen **los supuestos generales** dados en la introducción. De nuevo, la evolución del MDP original y del aproximado estarán especificados como en las ecuaciones dadas en (IV) y (V), mismas que se transcriben a continuación:

$$x_t = F(x_{t-1}, a_t, \xi_t) \ , \quad t = 1, \ldots; \tag{3.1}$$

$$\tilde{x}_t = F(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t) \ , \quad t = 1, \ldots; \tag{3.2}$$

Sea $r : \mathbb{K} \to R$ una función de *recompensa de un paso*, donde r es asumida medible y acotada en $\mathbb{K}$. Si $x \in X$, representa el estado inicial del proceso y $\pi \in \Pi$ es la política aplicada a los proceso (3.1) y (3.2), entonces los funcionales que se desean maximizar bajo el criterio de *recompensa total esperada* están definidos en (XIII), los cuales se transcriben a continuación:

$$R(x,\pi) = \liminf_{n\to\infty} E_x^{\pi}\left[\sum_{i=1}^{n} r(x_{t-1}, a_t)\right], \tag{3.3}$$

$$\tilde{R}(x,\pi) = \liminf_{n\to\infty} \tilde{E}_x^{\pi}\left[\sum_{i=1}^{n} r(\tilde{x}_{t-1}, \tilde{a}_t)\right], \tag{3.4}$$

donde E_x^{π} y $\tilde{E}_x^{\pi}$ representan la esperanza con respecto a las medidas de probabilidad, P_x^{π} y $\tilde{P}_x^{\pi}$ generadas por la aplicación de π a los procesos (3.1), (3.2) con el estado inicial x.

Las funciones de valor correspondientes (véase (XIV)) para los procesos (3.1) y (3.2), quedan definidas como:

$$R_*(x) = \sup_{\pi\in\Pi} R(x,\pi) \quad , \quad x \in X \quad , \tag{3.5}$$

$$\tilde{R}_*(x) = \sup_{\pi\in\Pi} \tilde{R}(x,\pi) \quad , \quad x \in X \quad , \tag{3.6}$$

En las siguientes secciones se introducirán los supuestos bajo los cuales, las funciones de valor R_* y $\tilde{R}_*$ son acotadas, y además aseguran la existencia de políticas estacionarias óptimas f_* y $\tilde{f}_*$ para los procesos (3.1) y (3.2) respectivamente, es decir, que satisfacen lo siguiente:

$$R(x, f_*) = R_*(x) \quad ; \quad \tilde{R}(x, \tilde{f}_*) = \tilde{R}_*(x) \quad , \quad x \in X \ .$$

El *índice de estabilidad* (con respecto a R) se define como:

$$\Delta_R(x) := R(x, f_*) - R(x, \tilde{f}_*) \quad , \quad x \in X \ .$$

3.2. Supuestos y existencia de políticas óptimas estacionarias.

Como antes, $\mathbb{F}$ denotará el conjunto de todas las políticas estacionarias. Para cada f , P_f y $\tilde{P}_f$ denotan las *probabilidades de transición* de los procesos de Markov (3.1) y (3.2) cuando $a_t = f(x_{t-1})$, $\tilde{a}_t = f(\hat{x}_{t-1})$, es decir, cuando se aplica la política de control f .

El siguiente supuesto es una variante del supuesto 2 usado en James, Collins, (2006).

<u>Supuesto 3.1.</u> Condiciones de transitoriedad o absorbente.

Existe un conjunto $\Theta \in \mathcal{B}_X$ tal que:

(a) Si $f \in \mathbb{F}$, entonces se cumple alguno de los siguientes dos enunciados:

$$\Theta \text{ es absorbente bajo } P_f \text{ , y } \sup_{x\in X} E_x^f \tau_{x,f}(\Theta) \le M_f < \infty , \tag{3.7}$$

o bien

$$\inf_{x\in X} R(x, f) = -\infty .$$

(b) Si $f \in \mathbb{F}$, entonces se cumple alguno de los siguientes dos enunciados:

$$\Theta \text{ es absorbente bajo } \tilde{P}_f \text{ , y } \sup_{x\in X} \tilde{E}_x^f \tilde{\tau}_{x,f}(\Theta) \leq \tilde{M}_f < \infty , \tag{3.8}$$

o bien

$$\inf_{x\in X} \tilde{R}(x, f) = -\infty .$$

(c) El conjunto de políticas estacionarias que satisfacen (3.7) es no vacío, y el conjunto de políticas estacionarias que satisfacen (3.8) es no vacío. En (3.7) y (3.8) los términos $\tau_{x,f}(\Theta)$, $\tilde{\tau}_{x,f}(\Theta)$ son tiempos de primera-entrada (para $t \geq 0$) en el conjunto Θ de los procesos (3.1) y (3.2) respectivamente, cuando se aplica la política f , y el estado inicial es x .

(d) Las funciones de valor R_* y $\tilde{R}_*$ son superiormente acotadas.

Sea $\mathbb{B}$ el espacio de todas las funciones medibles y acotadas $u : X \to R$, tal que $u(x) = 0$, $x \in \Theta$. El espacio lineal $\mathbb{B}$ está equipado con la norma del supremum $\|\cdot\|_\infty$.

<u>*Supuesto 3.2.*</u> Condiciones de continuidad y de acotamiento.

(a) La función de recompensa de un período $r(x,a)$ es acotada en $\mathbb{K}$; $r(x,a) = 0$ para toda $x \in \Theta$, $a \in A(x)$. Para toda $x \in X$, la función $a \to r(x,a)$ es semicontinua superiormente en $A(x)$.

(b) Para cada $x \in X$, $A(x)$ es compacto.

(c) Para cada $u \in \mathbb{B}$, y $x \in X_0 = X \setminus \Theta$, las funciones siguientes:

$$a \to Eu[F(x,a,\xi)] \text{ ,}$$
$$a \to Eu[F(x,a,\tilde{\xi})]$$

son continuas en $A(x)$.

El Lema 3.1. que se presenta en la sección 3.5, muestra que una política estacionaria f que satisfaga (3.7), es transitoria lo cual significa que (véase James y Collins, 2006):

$$\left\| \sum_{t=0}^{\infty} Q_f^t \right\|_0 < \infty \ , \tag{3.9}$$

donde Q_f es la restricción del kernel P_f a $\left(X_{0,} \ \mathcal{B}_{X_0}\right)$, $\|\cdot\|_0$ es el operador de la norma y nuevamente $X_0 = X \setminus \Theta$. Lo mismo es cierto para cualquier política estacionaria que satisfaga (3.8) (reemplazando P_f por $\tilde{P}_f$ y Q_f por $\tilde{Q}_f$). Por lo tanto, la siguiente afirmación es una consecuencia directa del Teorema 1 dado en James y Collins (2006).

Proposición 3.1.

Si se asume que los *supuestos 3.1 y 3.2* se cumplen, entonces:

(a) Existen políticas estacionarias f_* y $\tilde{f}_*$ que son óptimas para los procesos (3.1) y (3.2) respectivamente.

(b) f_* satisface (3.7), y $\tilde{f}_*$ satisface (3.8).

(c) $R_*, \tilde{R}_* \in \mathbb{B}$, y en particular son cero en Θ.

Observación 3.1. Si una política estacionaria f satisface (3.7), entonces de (3.9) se sigue que la recompensa total en (3.3) es

$$R_f(x) \equiv R(x, f) = E_x^f \sum_{t=1}^{\infty} r\left(x_{t-1}, f(x_{t-1})\right) \ , \tag{3.10}$$

y además,

$$R_f \in \mathbb{B} \ . \tag{3.11}$$

Similarmente, si f satisface (3.8), entonces:

$$\tilde{R}_f(x) \equiv \tilde{R}(x, f) = \tilde{E}_x^f \sum_{t=1}^{\infty} r(\tilde{x}_{t-1}, f(\tilde{x}_{t-1})) \ , \tag{3.12}$$

y

$$\tilde{R}_f \in \mathbb{B}. \tag{3.13}$$

3.3. Estimación de la estabilidad con respecto a la métrica de Variación Total.

Se requerirá el siguiente supuesto para hacer que la constante K en la *desigualdad de estabilidad* (3.15), dada más adelante, esté completamente determinada por las características del MDP original (3.1), y las constantes mencionadas en el *supuesto 3.3.* En tales casos, fijando D_ξ y considerando $D_{\tilde{\xi}} \equiv D_{\tilde{\xi},n}$ tal que $\mathbf{V}\mathbf{r}(D_\xi, D_{\tilde{\xi},n}) \to 0$, la constante K en (3.15) no depende de $\mathbf{V}\mathbf{r}$ $(D_\xi, D_{\tilde{\xi},n})$.

<u>*Supuesto 3.3.*</u>

(a) El conjunto Θ en el *supuesto 3.1.* es absorbente bajo P_{f_*} y $P_{\tilde{f}_*}$.

(b) Existen constantes conocidas $M < \infty$, $\gamma < 1$ y un entero $m \geq 1$ tales que

$$\sup_{x \in X_0} \tilde{E}_x^{f_*} \left[\tilde{\tau}_{x, f_*}(\Theta) \right] \leq M \ ;$$

$$\sup_{x \in X_0} \tilde{E}_x^{\tilde{f}_*} \left[\tilde{\tau}_{x, \tilde{f}_*}(\Theta) \right] \leq M \ , \text{ y}$$

$$\sup_{x \in X_0} P_*^{\tilde{f}_*} \left[\tau_{x, \tilde{f}_*}(\Theta) > m \right] \leq \gamma \ . \tag{3.14}$$

Como antes, en este capítulo D_ξ y $D_{\tilde{\xi}}$ denotarán las distribuciones de ξ y de $\tilde{\xi}$ respectivamente, definidas en $(S, \mathcal{B}_S)$.

Teorema 3.1.

Se asume que los *supuestos 3.1, 3.2. y 3.3.* se cumplen. Entonces existe una constante $K<\infty$ determinada de forma única por las propiedades del MDP de (3.1) y las constantes M,γ,m del *supuesto 3.3.*, tales que:

$$\sup_{x\in X}\Delta_R(x)\leq K\ \mathbf{Vr}(D_\xi,D_{\tilde{\xi}}), \tag{3.15}$$

donde $\mathbf{Vr}$ es la *métrica de Variación Total* definida en (XXII).

Al final de este capítulo se proporciona la demostración de este Teorema.

Observación 3.2.

- La prueba del Teorema 3.1. usa el método de operadores contractivos tal y como se hace en los artículos de Gordienko, Salem (2000), y en Zaitseva (2008). Si se omite la frase "determinada de forma única por…", entonces la desigualdad (3.15) sigue siendo cierta sin asumir el supuesto 3.3.
- Siguiendo las ideas expuestas en Gordienko, Salem (2000), se puede obtener (usando una prueba similar) una versión de (3.15) en un contexto más general del problema de *estimación de la estabilidad*: en la definición del MDP de (3.2) la función F (denotada ahora, digamos por $\tilde{F}$) podría ser diferente de la función F en (3.1) (pero en cierto sentido cerrada).

Ejemplo 3.1. Problema de paro óptimo.

Considere el problema usual de paro óptimo de un proceso de Markov a tiempo discreto (ver por ejemplo, Shiryayev (1978)):

$$y_t=\Phi(y_{t-1},\xi_t)\ ,\ t\geq 1. \tag{3.16}$$

El proceso (3.16) se aproxima por el siguiente proceso:

$$\tilde{y}_t = \Phi\left(\tilde{y}_{t-1}, \tilde{\xi}_t\right),\ t \geq 1. \tag{3.17}$$

Ambos procesos toman valores en un espacio-fase de Borel Y. Note que si Y es un espacio métrico separable, entonces cada proceso de Markov a tiempo discreto homogéneo puede ser representado en la forma (3.16) con vectores aleatorios *i.i.d.* $\xi_1, \xi_2, \ldots$ en algún espacio de Borel S (ver Borovkov, Foss (1992)).

Considérese una extensión estándar, (ver por ejemplo Ross (1992), o bien Shiryayev (1978)), de (3.16) y de (3.17) a los MDPs en $X = Y \cup *$, donde * es un estado absorbente y donde el proceso “vive” después de ser parado. El conjunto de acciones es $A(x) = A = \{0,1\}$, donde la acción 0 significa que se continúa con las observaciones y la acción 1 “para” el proceso.

Por lo tanto, se puede definir la función F para obtener dos MDPs dados por:

$$x_t = F\left(x_{t-1}, a_t, \xi_t\right),\ t \geq 1, \tag{3.18}$$

$$\tilde{x}_t = F\left(\tilde{x}_{t-1}, \tilde{a}_t, \tilde{\xi}_t\right),\ t \geq 1, \tag{3.19}$$

y el problema de paro óptimo es reducido a la maximización de *la recompensa total esperada* para los procesos (3.18) y (3.19) con la siguiente función de retorno a un paso:

$$r(x,a) = \begin{cases} \hat{R}(x) & si \quad a = 1, x \in Y; \\ -c(x) & si \quad a = 0, x \in Y; \\ 0 & si \quad x = *, a \in \{0,1\}. \end{cases} \tag{3.20}$$

En (3.20), $c, \hat{R} : Y \to \mathbb{R}$ son funciones reales y acotadas dadas, tal que representan los pagos para la continuación de las observaciones, y la recompensa por parar, respectivamente. Por la definición de las probabilidades de transición correspondientes a los procesos (3.18) y (3.19), el conjunto $\Theta = \{*\}$ es absorbente para cualquier política $\pi \in \Pi$. La situación más simple en la que se cumplen los *supuestos 3.1., 3.2. y 3.3.* para este ejemplo, es el caso donde se satisface la siguiente condición:

$$\inf_{y \in Y} c(y) \geq \alpha_0 > 0. \tag{3.21}$$

El *supuesto 3.2.* se satisface porque A es finito. Las condiciones (3.7) y (3.8) en el *supuesto 3.4.* se cumplen, por ejemplo para la política estacionaria: "parar en el momento $t = 0$ " (en un estado inicial). Debido a (3.21), para cada política estacionaria f con tiempo promedio infinito hasta detenerse $R(x, f) = -\infty$. El *supuesto 3.3.* se satisface, y la desigualdad (3.14) se cumple con $\gamma = 0$ para todo $m \geq \frac{1}{\alpha_0} \sup_{y \in Y} \hat{R}(y) + 1$.

Observe que bajo la condición (3.21) es muy simple encontrar una cota superior de la constante K en (3.15), y esta cota podría elegirse para que sólo dependa del $\sup_{y \in Y} \hat{R}(y)$, y de α_0.

Observación 3.3. La condición (3.21) hace del problema de estimación de la estabilidad "casi trivial" (ya que todo es reducido a un horizonte de tiempo finito). Mientras que la condición (3.21) puede ser relajada de varias maneras, no se puede omitir por completo en el caso general. De hecho, problemas de paro inestables son presentados en Zaitseva (2008). En esos problemas $\mathbf{Vr}(D_{\xi}, D_{\tilde{\xi}}) \to 0$ cuando $\varepsilon \to 0$, pero para un cierto $y \in Y$ *el índice de estabilidad* es más grande que cierta constante positiva. En particular, en tales ejemplos $\inf_{y \in Y} c(y) = 0$. En Gordienko, Novikov (2014), así como en Zaitseva (2008, 2010), para obtener la desigualdad como en (3.15) (el problema de paro óptimo sin la condición (3.21)), fueron impuestas fuertes condiciones de ergodicidad en los procesos (3.16) y (3.17).

Ejemplo 3.2. Modelo simplificado de consumo-inversión, con inversión de "mucho riesgo". Sean $X = [0,\infty]$, $A(x) = [0,x]$, $x \in X$, $S = R$, y los procesos,

$$x_t = [(x_{t-1} - a_t)\xi_t]^+ \quad , \quad t \geq 1; \tag{3.22}$$

$$\tilde{x}_t = [(\tilde{x}_{t-1} - \tilde{a}_t)\tilde{\xi}_t]^+ \quad , \quad t \geq 1. \tag{3.23}$$

En este modelo si $x_{t-1} = x_t$ es el capital actual, entonces la cantidad $(x - a)$ es usada para inversión, y el resto de dinero a se consume.

Sea $\Theta = \{0\}$ y r una función de retorno que satisface el *supuesto 3.2.* inciso (a). La siguiente suposición significa que en este modelo las inversiones son "muy riesgosas":

Existe una constante $\beta \geq 0$ tal que $P(\xi \leq 0) \geq \beta$ y $P(\tilde{\xi} \leq 0) \geq \beta$. (3.24)

Si se asume que en (3.22) y (3.23) las variables aleatorias ξ y $\tilde{\xi}$ tienen densidades continuas g y $\tilde{g}$ que desaparecen lo suficientemente rápido en $\pm\infty$, entonces el *supuesto 3.2.* inciso (c) se satisface. De (3.22) y (3.23) es claro que el estado $x = 0$ es absorbente para toda $f \in \mathbb{F}$.

También $P_x^f(\tau_{x,f}(\{0\}) > n) \leq P(\xi_1 > 0, \xi_2 > 0, ..., \xi_n > 0) \leq (1-\beta)^n$ debido a (3.24).

Por lo tanto, se cumplen los *supuestos 3.1. y 3.3.*, por lo que se puede aplicar (3.15), para acotar superiormente el *índice de estabilidad* que para este ejemplo. Dicho índice queda expresado como (véase (XXVIII)) :

$$\sup_{x \geq 0} \Delta(x) \leq K \int_{-\infty}^{+\infty} |g(s) - \tilde{g}(s)| ds .$$

3.4.- Estimación de la estabilidad con respecto a la métrica de Prokhorov.

Dado que *la métrica de Variación Total* es demasiado fuerte, la *desigualdad de estabilidad* dada en (3.15) no siempre es aplicable.

En esta sección se fortalece el *supuesto 3.1.* y se asumen ciertas condiciones de *Lipschitz*. Esto permitirá probar una versión de la desigualdad (3.15) pero usando *la métrica de Prokhorov*, π, en lugar de la de *Variación Total*, **Vr,** del lado derecho.

En lugar de los *supuestos 3.2. y 3.3.*, se introduce el siguiente supuesto:

<u>Supuesto 3.4.</u> (Uniform transitory or absorbing conditions).

Existe un conjunto $\Theta \in \mathcal{B}_X$ tal que:

(a) Para cada $f \in \mathbb{F}$, el conjunto Θ es absorbente bajo P_f y $\widetilde{P}_f$;

(b) Existe una constante $M < \infty$ tal que

$$\sup_{f \in F} \sup_{x \in X} E_x^f \tau_{x,f}(\Theta) \leq M \quad , \tag{3.25}$$

$$\sup_{f \in F} \sup_{x \in X} \widetilde{E}_x^f \widetilde{\tau}_{x,f}(\Theta) \leq M \quad , \tag{3.26}$$

El siguiente supuesto sustituye al *supuesto 3.2.*:

<u>Supuesto 3.5</u>. Condiciones de *Lipschitz*.

(a) La *función de recompensa* $r(x,a)$ es acotada en $\mathbb{K}$; $r(x,a)=0$ para cada (x,a) con $x \in \Theta$; además, existe una constante finita L_0 tal que

$$|r(k)-r(k')| \leq L_0 \nu(k,k') \quad ; \quad k,k' \in \mathbb{K}. \tag{3.27}$$

(b) Para cada $x \in X$, el conjunto $A(x)$ es compacto y existe una constante finita L_1 tal que

$$h(A(x), A(x')) \leq L_1 \rho(x,x') \quad , \quad x,x' \in X_0, \tag{3.28}$$

donde h es *la métrica de Hausdorff*.

(c) Existen constantes finitas $L, \tilde{L}$ tales que para toda $u \in \mathbb{B}$ con $\|u\|_\infty \leq 1$,

$$\left| \; Eu[F(k,\xi)] - Eu[F(k',\xi)] \; \right| \leq L\,\nu(k,k') \;;$$

$$\left| \; Eu[F(k,\tilde{\xi})] - Eu[F(k',\tilde{\xi})] \; \right| \leq \tilde{L}\,\nu(k,k') \quad , \quad k,k' \in \mathbb{K}. \tag{3.29}$$

(d) Existe una constante finita L_* tal que, para toda $k \in \mathbb{K}$,

$$\rho[F(k,s), F(k,s')] \leq L_* r(s,s') \quad , \quad s,s' \in S\,. \tag{3.30}$$

Teorema 3.2.

Se asume que los *supuestos 3.4. y 3.5.* se satisfacen. Entonces, existe una constante finita, $\overline{K}$, que depende solamente de las características del MDP (3.1) y de las constantes involucradas en los *supuestos 3.4., 3.5.*, tal que:

$$\sup_{x \in X} \Delta_R(x) \leq \overline{K}\pi\left(D_\xi, D_{\tilde{\xi}}\right)\max\left\{1, \log\left(\frac{1}{\pi\left(D_\xi, D_{\tilde{\xi}}\right)}\right)\right\}, \tag{3.31}$$

donde π es la *métrica de Prokhorov*.

Al final de este capítulo se proporciona la demostración.

Ejemplo 3.3. Sean $X = [0, H]$, donde H es un número finito dado, $A(x) = [0, x]$, $x \in X$, y $S = R$. Considérese el siguiente par de MDPs:

$$x_t = \min\left\{H, \left[x_{t-1} - a_t - \min\left\{\frac{1}{x_{t-1} - a_t}, \vartheta\right\} + \xi_t\right]^+\right\} \;, \tag{3.32}$$

$$\tilde{x}_t = \min\left\{H, \left[\tilde{x}_{t-1} - \tilde{a}_t - \min\left\{\frac{1}{\tilde{x}_{t-1} - \tilde{a}_t}, \vartheta\right\} + \tilde{\xi}_t\right]^+\right\} \;, \tag{3.33}$$

$$t = 1,2,...$$

En (3.32) y (3.33), los términos $\{\xi_t\}$, $\{\tilde{\xi}_t\}$ son dos sucesiones de variables aleatorias *i.i.d.*; y ϑ es un número positivo dado que está involucrado en el *supuesto 3.6.* que se introduce más abajo.

Sea $\Theta = \{0\}$, y $r(k)$ con $k \in K$ la función de recompensa de un paso que satisface el *supuesto 3.5.* inciso (a).

<u>*Supuesto 3.6.*</u> (Aplicado solamente para este ejemplo).

(a) Las variables aleatorias ξ y $\tilde{\xi}$ tienen densidades continuas acotadas, g y $\tilde{g}$ respectivamente. Además $g(x) = \tilde{g}(x) = 0$ para toda $x \geq \vartheta$.

(b) Las densidades g, $\tilde{g}$ son diferenciables casi en todas partes, y sus derivadas están acotadas (donde existan).

(c) $E\xi < 0$; $E\tilde{\xi} < 0$.

En primer lugar, se demostrará que para cualquier $f \in \mathbb{F}$ el conjunto $\Theta = \{0\}$ es absorbente bajo P_f y $\tilde{P}_f$. En efecto, si $x_{t-1} = 0$, entonces $a_t = 0$, $\min\left\{\dfrac{1}{x_{t-1} - a_t}, \vartheta\right\} = \vartheta$. Pero en (3.32) la variable aleatoria $\xi_t - \vartheta$ es no-positiva, y por lo tanto $x_t = 0$. Por lo que el *supuesto 3.4.* inciso (a) se cumple.

Por otra parte, gracias al *supuesto 3.6.*, (c) se cumplen las condiciones (3.25) y (3.26). De (3.32), para cada $f \in \mathbb{F}$ con probabilidad 1 se tiene que $x_t \leq x'_t$, $t = 0,1,...$, donde $x'_t = [x'_{t-1} + \xi_t]^+$, $t \geq 1$ es la caminata aleatoria absorbente en $x = 0$, y para esta caminata $E\xi_t < 0$. Es bien sabido (ver Meyn y Tweedie (1993)) que para tal caminata aleatoria, el tiempo promedio para entrar en $\{0\}$ está uniformemente acotado por los estados iniciales de un intervalo acotado. Por lo tanto, el *supuesto 3.4.* se cumple para este ejemplo.

Para $A(x)=[0,x]$ se puede verificar que (3.28) se satisface para $L_1=1$, y de (3.32), la desigualdad (3.30) se cumple para $L_*=1$.

Finalmente, sólo queda verificar el *supuesto 3.5.* inciso (c). Si $y=x-a$, $\varphi(y):=y-\min\left\{\frac{1}{y},\vartheta\right\}$ es la función de *Lipschitz* en $[0,\infty]$. Sea u una función arbitraria de $\mathbb{B}$ con $\|u\|_\infty \le 1$. Para cualquier $y,y'\in[0,H]$, se tiene que

$$I(y):=Eu\left[\min\left\{H,[\varphi(y)+\xi]^+\right\}\right]$$

$$=u(H)P(\varphi(y)+\xi\ge H)+\int_0^H u(\varphi(y)+t)g(t)dt$$

$$=u(H)\int_{H-\varphi(y)}^{\infty} g(z)dz+\int_{\varphi(y)}^{\varphi(y)+H} u(z)g(z-\varphi(y))dz\,. \tag{3.34}$$

Para ver que $|I(y)-I(y')|\le L|y-y'|$, obsérvese que el primer término del lado derecho de (3.34) es función de *Lipschitz* ya que g es acotada y su soporte está acotado superiormente. Sean por ejemplo $y>y'$; entonces $\varphi(y)>\varphi(y')$ y la diferencia del segundo término del lado derecho de (3.34) es menor que

$$\int_{\varphi(y')+H}^{\varphi(y)+H}|u(z)|g(z-\varphi(y))dz+\int_{\varphi(y')}^{\varphi(y)}|u(z)|g(z-\varphi(y'))dz+$$

$$+\int_{\varphi(y)}^{\varphi(y')+H}|u(z)|\sup_z|g'(z)||\varphi(y)-\varphi(y')|dz\,. \tag{3.35}$$

Como $\varphi(y')$, $\varphi(y)\in[-\vartheta,H]$, el último sumando en (3.35) es menor que $(2M+\vartheta)\|u\|_\infty\sup_z|g'(z)||\varphi(y)-\varphi(y')|$, por lo que φ también es Lipschitz. De la acotación de g se sigue que el primero y el segundo término de (3.35) es menor que $L|y-y'|$.

Ahora, sea $y=0$. De (3.34) y el *supuesto 3.6.* inciso (a), se tiene que $I(y')=u(0)=0$. Para un valor de y lo suficientemente pequeña, $\varphi(y)=y-\vartheta$, y como $\xi-\vartheta\leq 0$, en (3.34),

$$|I(y)|\leq \|u\|_\infty P(y+\xi-\vartheta>0)=\int_{\vartheta-y}^{\vartheta} g(z)dz\leq const\cdot y\ .$$

Observación 3.4. En Gordienko, Salem (2000) se discute un ejemplo de inestabilidad de un MDP con respecto a la métrica de **Ѵr**. En dicho ejemplo, se usa el criterio de optimización descontado (con una función de costo de un paso que es no-acotada). El ejemplo puede ser modificado para el caso de una función de costo acotada. Por otro lado, en Hernández-Lerma, Lasserre (1999) está demostrado que un MDP con el criterio de *costo total descontado* puede ser transformado en un MDP transitorio con una *recompensa total esperada*.

3.5. Demostración de los teoremas.

El siguiente lema establece la conexión entre el *supuesto 3.1.* y la definición de política transitoria dada en James, Collins (2006) y en Pliska (1977).

Lema 3.1.

Sea Θ como en el *supuesto 3.1.*, $X_0=X\setminus\Theta$ para $f\in\mathbb{F}$, Q_f es la restricción del kernel P_f para $(X_0,\mathcal{B}_{X_0})$. Si se satisface (3.7), entonces:

$$\left\|\sum_{t=0}^{\infty}Q_f^t\right\|_0\leq M_f<\infty\,. \tag{3.36}$$

En (3.36) $\|\cdot\|_0$ es la norma del operador correspondiente a la norma del supremo $\|\cdot\|_\infty$ en $\mathbb{B}$ y Q_f^t es la *t-ésima* potencia del operador que corresponde al kernel Q_f . Lo último es la restricción de P_f a $(X_0,\mathcal{B}_{X_0})$.

NOTA:

En el resto de este capítulo para simplificar la nomenclatura, para $u \in \mathbb{B}$, en lugar del símbolo $\|u\|_\infty := \sup_{x \in X} |u(x)|$, se usará el símbolo $\|u\|$.

<u>Demostración.</u>

Como Q_f^t es un operador monótono, se tiene que

$$\left\| \sum_{t=0}^{\infty} Q_f^t \right\|_0 = \left\| \sum_{t=0}^{\infty} Q_f^t I \right\| = \sup_{x \in X_0} \left| \sum_{t=0}^{\infty} Q_f^t I(x) \right|. \tag{3.37}$$

Para cada $t \geq 0$, se cumple que

$$Q_f^t I(x) = P_x^f (x_t \in X_0) = P_x^f (\tau_{x,f}(\Theta) > t), \tag{3.38}$$

ya que Θ es un conjunto absorbente para P_f.

Por lo que, de (3.37), (3.38) se tiene,

$$\left\| \sum_{t=0}^{\infty} Q_f^t \right\|_0 = \sup_{x \in X_0} \sum_{t=0}^{\infty} P_x^f (\tau_{x,f}(\Theta) > t) \leq M_f \ .$$

Demostración del Teorema 3.1.

Sean $f_*, \tilde{f}_*$ las políticas estacionarias óptimas introducidas en la proposición 3.1. y $F_* := \{f_*, \tilde{f}_*\}$.

Bajo los *supuestos 3.1. y 3.3.*, para cada $f \in F_*$ las recompensas correspondientes $R_f \equiv R(x, f)$, $\tilde{R}_f \equiv \tilde{R}(x, f)$ son funciones acotadas y pueden ser re-escritas como

$$R_f(x) = E_x^f \left[\sum_{t=1}^{\infty} r(x_{t-1}, f(x_{t-1})) \right] , \tag{3.39}$$

$$\tilde{R}_f(x) = \tilde{E}_x^f\left[\sum_{t=1}^{\infty} r(\tilde{x}_{t-1}, f(\tilde{x}_{t-1}))\right]. \tag{3.40}$$

De la proposición 3.1. y el *supuesto 3.3.*, los siguientes operadores G_f , $\tilde{G}_f$ $(f \in F_*)$,

$$G_f u(x) := r(x, f(x)) + Eu[F(x, f(x), \xi)] \ , \tag{3.41}$$

$$\tilde{G}_f u(x) := r(x, f(x)) + Eu[F(x, f(x), \tilde{\xi})] \tag{3.42}$$

actuan de $\mathbb{B}$ a $\mathbb{B}$.

Usando (3.39), (3.40) y la propiedad de Markov, se encuentra que para $f \in F_*$,

$$R_f = G_f R_f \quad \text{y} \quad \tilde{R}_f = \tilde{G}_f \tilde{R}_f . \tag{3.43}$$

Para el *índice de estabilidad* dado en (XVI) y de los resultados en los artículos de Gordienko, Salem (2000), y de Zaitseva (2008) se tiene que

$$\Delta_R(x) \leq 2 \max_{f \in F_*} \left| R(x, f) - \tilde{R}(x, f) \right| . \tag{3.44}$$

Primero, omitiendo el subíndice *, es decir, sea $f = f_*$. Entonces, por (3.43), para cada $n \geq 1$:

$$\left| R(x, f) - \tilde{R}(x, f) \right| \leq \left\| R_f - \tilde{R}_f \right\| = \left\| G_f^n R_f - \tilde{G}_f^n \tilde{R}_f \right\|$$

$$\leq \left\| G_f^n R_f - G_f^n \tilde{R}_f \right\| + \left\| G_f^n \tilde{R}_f - \tilde{G}_f^n \tilde{R}_f \right\| . \tag{3.45}$$

De la proposición 3.1. inciso (b), la política $f = f_*$ satisface (3.7). Por lo tanto, por el Lema 3.1. y los resultados correspondientes dados en James, Collins (2006), existe un entero $n \geq 1$ tal que, el operador G_f^n es contractivo en $\mathbb{B}$ con algún módulo $\delta < 1$. Entonces, de (3.45) se tiene lo siguiente:

$$\left\| R_f - \tilde{R}_f \right\| \leq \frac{1}{(1-\delta)} \left\| G_f^n \tilde{R}_f - \tilde{G}_f^n \tilde{R}_f \right\| . \tag{3.46}$$

Teniendo en cuenta (3.41), (3.42) y aplicando argumentos usados en Zaitseva (2008) se obtiene,

$$\left\|G_f^n\widetilde{R}_f-\widetilde{G}_f^n\widetilde{R}_f\right\|\le n\left\|\widetilde{R}_f\right\|\sup_{x\in X_0,a\in A(x)}\sup_{B\in B_{X_0}}\left|P\big(F(x,a,\xi)\in B\big)-P\big(F(x,a,\widetilde{\xi})\in B\big)\right|. \tag{3.47}$$

El último término del lado derecho de (3.47) es menor que $\frac{1}{2}\,\mathbf{Vr}(\xi,\widetilde{\xi})$.

Por otro lado, ya que $r\equiv 0$ en Θ, del *supuesto 3.3.* se tiene que

$$\left\|\widetilde{R}_f\right\|\equiv\left\|\widetilde{R}_{f_*}\right\|=\sup_{x\in X_0}\left|\widetilde{E}_x^f\sum_{t=1}^{\infty}r\big(\widetilde{x}_{t-1},f_*(\widetilde{x}_{t-1})\big)I_{\{\widetilde{\tau}_{x,f_*}(\Theta)>t-1\}}\right|\le bM\ , \tag{3.48}$$

donde $b=\sup_{k\in K}|r(k)|$ y M es la constante involucrada en el *supuesto 3.3.*

Segundo, en (3.44) sea $f=\widetilde{f}_*$. Ahora, en la desigualdad (3.45) se sustituye $f=\widetilde{f}_*$. Sea $m\ge 1$ y $\gamma<1$ las constantes mencionadas en el *supuesto 3.3.* Luego, de (3.38) y (3.14) del *supuesto 3.3.*, se tiene que $\left\|Q_f^m\right\|\le\gamma$.

Dado que el conjunto Θ es absorbente bajo $P_f=P_{\widetilde{f}}$ (ver el *supuesto 3.3.*), e iterando (3.41), para cada $u,v\in\mathbb{B}$: $\left\|Q_f^m u-Q_f^m v\right\|\le\gamma\|u-v\|$. Entonces, de (3.45) se sigue que

$$\left\|R_f-\widetilde{R}_f\right\|\le\frac{1}{(1-\gamma)}\left\|G_f^m\widetilde{R}_f-\widetilde{G}_f^m\widetilde{R}_f\right\|.$$

Procediendo como en (3.47) y (3.48) (con $f=\widetilde{f}_*$ en lugar de $f=f_*$), y aplicando el *supuesto 3.3.* inciso (b), para alguna constante K se tiene

$$\left\|R_{\widetilde{f}_*}-\widetilde{R}_{\widetilde{f}_*}\right\|\le K\,\mathbf{Vr}\big(D_\xi,D_{\widetilde{\xi}}\big). \tag{3.49}$$

Para concluir la demostración del Teorema 3.1., basta reunir las desigualdades (3.44), (3.46), (3.47), (3.48) y (3.49), con lo que se obtiene explícitamente la siguiente constante

$$K = \max\left\{\frac{n}{1-\delta}, \frac{m}{1-\gamma}\right\} bM \ . \tag{3.49.1}$$

Demostración del Teorema 3.2.

Sean f_* , $\tilde{f}_* \in \mathbb{F}$ políticas estacionarias óptimas para los MDPs (3.1), (3.2), respectivamente, y $R_* = R_{f_*}$, $\tilde{R}_* = R_{\tilde{f}_*}$ las funciones de valor corresponientes. La existencia de f_* y $\tilde{f}_*$ fue asegurada en la proposición 3.1. Del *supuesto 3.4.* inciso (a), para cada $f \in \mathbb{F}$, las recompensas totales correspondientes R_f y $\tilde{R}_f$ (ver (3.10)-(3.13)) son cero en Θ . Particularmente $R_*(x) = \tilde{R}_*(x) = 0$, para $x \in \Theta$. Por lo tanto, para $f \in \mathbb{F}$, se pueden considerar todas las funciones R_f y $\tilde{R}_f$, como elementos del espacio $\mathbb{B}$ (teniendo en cuenta su acotación, según el *supuesto 3.4.*).

De manera habitual, se introducen los operadores de la programación dinámica, $T, \tilde{T} : \mathbb{B} \to \mathbb{B}$, :

$$Tu(x) := \sup_{a \in A(x)} \{r(x,a) + Eu[F(x,a,\xi)]\} \ , \ x \in X \ , \tag{3.50}$$

$$\tilde{T}u(x) := \sup_{a \in A(x)} \{r(x,a) + Eu[F(x,a,\tilde{\xi})]\} \ , \ x \in X \ , \tag{3.51}$$

Del *supuesto 3.5.* incisos (a) y (b), se sigue que para cada $u \in \mathbb{B}$ existe una política estacionaria (selector), f_u , tal que,

$$\sup_{a \in A(x)} \{r(x,a) + Eu[F(x,a,\xi)]\} = r(x, f_u(x)) + Eu[F(x, f_u(x), \xi)]$$

$$= r(x, f_u(x)) + E_x^{f_u} u(x) \ , \ x \in X \ .$$

Ahora, por el *supuesto 3.4.* inciso (a), para $x \in \Theta$, se tiene $Tu(x) = 0$, y además $T\,\mathbb{B} \subseteq \mathbb{B}$. (Similarmente $\tilde{T}\,\mathbb{B} \subseteq \mathbb{B}$).

Tal y como se demostró en James, Collins (2006) y en Pliska (1978) el cumplimiento de los *supuestos 3.4. y 3.5.* es suficiente para la validez la siguiente proposición.

Proposición 3.2.

(a) $R_* = TR_*$, $\tilde{R}_* = \tilde{T}\tilde{R}_*$.

(b) La política óptima f_* es un selector en el lado derecho de (3.50) con $u = R_*$; y la política óptima $\tilde{f}_*$ es un selector en el lado derecho de (3.51) con $u = \tilde{R}_*$

(c) Existe un número entero $m \geq 1$ tal que el operador T^m es contractivo en $\mathbb{B}$ con algún módulo $\delta < 1$.

Para cualquier $(x,a) \in \mathbb{K}$, se define:

$$H(x,a) = r(x,a) + ER_*\left[F(x,a,\xi)\right] \ , \tag{3.52}$$

$$\tilde{H}(x,a) = r(x,a) + E\tilde{R}_*\left[F(x,a,\tilde{\xi})\right]. \tag{3.53}$$

Para simplificar la notación, sea $f = \tilde{f}_*$. Similarmente a lo realizado en Gordienko *et al.* (2008), sea $\Gamma_t = \{x, a_1, x_1, a_2, ..., x_{t-1}, a_t\}$ $(t \geq 1)$ la parte de una trayectoria del proceso (3.1) bajo la política de control $f = \{f, f, ...\}$ (con el estado inicial $x \in X_0$).

Por la propiedad de Markov, se tiene que,

$$\begin{aligned} \zeta_t &= E^f\left[R_*(x_t) \middle| \Gamma_t\right] \\ &= H(x_{t-1}, a_t) - r(x_{t-1}, a_t) - \sup_{a \in A(x_{t-1})} H(x_{t-1}, a) + \sup_{a \in A(x_{t-1})} H(x_{t-1}, a). \end{aligned} \tag{3.54}$$

Por (3.50), (3.52) y la proposición 3.2. inciso (a) se obtiene que

$$\zeta_t = H(x_{t-1}, a_t) - \sup_{a \in A(x_{t-1})} H(x_{t-1}, a) - r(x_{t-1}, a_t) + R_*(x_{t-1})$$

$$= \Lambda_t - r(x_{t-1}, a_t) + R_*(x_{t-1}) \ , \tag{3.55}$$

donde

$$\Lambda_t := \sup_{a \in A(x_{t-1})} H(x_{t-1}, a) - H(x_{t-1}, a_t) \geq 0 \,. \tag{3.56}$$

De (3.54) y (3.55) se obtiene lo siguiente:

$$E_x^f R_*(x_t) = E_x^f \Lambda_t - E_x^f r(x_{t-1}, a_t) + E_x^f R_*(x_{t-1}) \,.$$

Sumando la última ecuación sobre $t \in [1, n]$, se obtiene que

$$E_x^f \sum_{t=1}^{n} r(x_{t-1,} a_t) = R_*(x) - E_x^f R_*(x_n) - \sum_{t=1}^{n} E_x^f \Lambda_t \,. \tag{3.57}$$

Como $r(x,a)$, $R_* \in \mathbb{B}$, y bajo el *supuesto 3.4.* inciso (b) cuando $n \to \infty$, se tiene que $E_x^f \sum_{t=1}^{n} r(x_{t-1,} a_t) \to E_x^f \sum_{t=1}^{n} r(x_{t-1}, a_t) = R(x, f(x))$, (ver 3.10), y $E_x^f R_*(x_n) = [Q_f^n R_*](x) \to 0$, donde Q_f es el kernel definido en el lema 3.1.

Así que se puede pasar al límite en (3.57) para encontrar lo siguiente:

$$\Delta_R(x) = R_*(x) - R(x, f) \leq \limsup_{n \to \infty} \sum_{t=1}^{n} E_x^f \Lambda_t \ , \tag{3.58}$$

De manera similar al lema 3.1., está demostrado que (3.25) implica lo siguiente:

$$\left\| \sum_{t=0}^{\infty} Q_f^t \right\|_0 \leq M \text{ , para cada } f \in \mathbb{F}. \tag{3.59}$$

Por otro lado, en James, Collins (2006) se muestra que $\|R_*\| \leq M \|r\|$, similarmente $\|\tilde{R}_*\| \leq M \|r\|$. De la primera de estas desigualdades se sigue que (ver (3.52), (3.53)) en (3.58) el término Λ_t es una función de x_{t-1} (un estado bajo la política f) acotada por

$$2\|r\|(1+M)=:b \quad . \tag{3.60}$$

De la proposición 3.2. inciso (a), junto con (3.52) y (3.56) se tiene que

$$\Lambda_t(x_{t-1}) = R_*(x_{t-1}) - r(x_{t-1}, f(x_{t-1})) - ER_*(x_t) \qquad ,$$

y por el *supuesto 3.4.*, si $x_{t-1} \in \Theta$, entonces $x_t \in \Theta$, y por lo tanto (como $r(x,a)$ y R_* son cero en Θ) $\Lambda_t(x_{t-1}) = 0$ cuando $x_{t-1} \in \Theta$. Por consiguiente, $\Lambda_t = \Lambda(x_{t-1})$, donde Λ es una función de $\mathbb{B}$.

En Pliska (1978) se demuestra que bajo el *supuesto 3.4.* existen constantes $c < \infty$ y $\theta < 1$ tales que para cada $f \in \mathbb{F}$:

$$\left\|Q_f^n\right\|_0 \leq c\theta^n \quad , \quad n = 1,2,... \tag{3.61}$$

Por otra parte, de acuerdo a las propiedades anteriores de Λ, el lado derecho de (3.58) puede ser reescrito de la siguiente manera:

Sea $N \geq 1$ un entero arbitrario fijo (por ahora); entonces,

$$I(x) := \lim_{n\to\infty} \sup \sum_{t=1}^{n} E_x^f \Lambda_t = \sum_{t=1}^{\infty} E_x^f \Lambda_t$$

$$= \sum_{t=1}^{N} E_x^f \Lambda_t + \sum_{t>N} E_x^f \Lambda_t \quad , \tag{3.62}$$

y de (3.61), se tiene

$$\sup_{x\in X_0} \left| \sum_{t>N} E_x^f \Lambda_t \right| = \left\| \sum_{t>N} Q_f^t \Lambda_t \right\| \leq \sum_{t>N} \left\| Q_f^t \Lambda_t \right\| \leq \frac{bc}{1-\theta} \theta^{N+1} \quad . \tag{3.63}$$

combinando (3.58), (3.62) y (3.63), se obtiene la desigualdad siguiente:

$$\Delta_R(x) \leq \sum_{t=1}^{N} E_x^f \Lambda_t + \frac{bc}{1-\theta} \theta^{N+1} \quad . \tag{3.64}$$

Ahora, de la definición de Λ_t en (3.56), (3.51)-(3.53) y de la proposición 3.2. inciso (a) se tiene que

$$\Lambda_t = \sup_{a\in A(x_{t-1})} H(x_{t-1},a) - \sup_{a\in A(x_{t-1})} \tilde{H}(x_{t-1},a) + \tilde{H}(x_{t-1},a) - H(x_{t-1},a)$$

$$\leq 2 \sup_{a\in A(x_{t-1})} \left| H(x_{t-1},a) - \tilde{H}(x_{t-1},a) \right|$$

$$\leq 2 \sup_{a\in A(x_{t-1})} \left| ER_*\left[F(x_{t-1},a,\xi)\right] - E\tilde{R}_*\left[F(x_{t-1},a,\tilde{\xi})\right] \right| , \tag{3.65}$$

donde las esperanzas son interpretadas como esperanzas condicionales con x_{t-1} siendo fija.

De (3.65) se obtiene que

$$\Lambda_t \leq 2 \sup_{a\in A(x_{t-1})} \left| ER_*\left[F(x_{t-1},a,\xi)\right] - ER_*\left[F(x_{t-1},a,\tilde{\xi})\right] \right| +$$

$$+ 2 \sup_{a\in A(x_{t-1})} \left| ER_*\left[F(x_{t-1},a,\tilde{\xi})\right] - E\tilde{R}_*\left[F(x_{t-1},a,\tilde{\xi})\right] \right|$$

$$\Lambda_t \leq 2 \sup_{k\in K} \left| ER_*\left[F(k,\xi)\right] - ER_*\left[F(k,\tilde{\xi})\right] \right| + 2\left\| R_* - \tilde{R}_* \right\| . \tag{3.66}$$

De la proposición 3.2. inciso (c), existen enteros $m \geq 1$ y $\delta < 1$ tales que el operador T^m es contractivo con modulo $\delta < 1$. Entonces, nuevamente usando la proposición 2, se tiene que

$$\left\| R_* - \tilde{R}_* \right\| = \left\| T^m R_* - \tilde{T}^m \tilde{R}_* \right\| \leq \left\| T^m R_* - T^m \tilde{R}_* \right\| + \left\| T^m \tilde{R}_* - \tilde{T}^m \tilde{R}_* \right\| ,$$

o bien,

$$\left\| R_* - \tilde{R}_* \right\| \leq \frac{1}{1-\delta} \left\| T^m \tilde{R}_* - \tilde{T}^m \tilde{R}_* \right\| . \tag{3.67}$$

Ahora, dado que el operador T es no-expansivo, por inducción se tiene

$$\left\|T^m\tilde{R}_* - \tilde{T}^m\tilde{R}_*\right\| \le \left\|TT^{m-1}\tilde{R}_* - T\tilde{T}^{m-1}\tilde{R}_*\right\| + \left\|T\tilde{T}^{m-1}\tilde{R}_* - \tilde{T}\tilde{T}^{m-1}\tilde{R}_*\right\|$$

$$\le \left\|T^{m-1}\tilde{R}_* - \tilde{T}^{m-1}\tilde{R}_*\right\| + \left\|T\tilde{R}_* - \tilde{T}\tilde{R}_*\right\|$$

$$\le m\left\|T\tilde{R}_* - \tilde{T}\tilde{R}_*\right\|$$

$$\le m\sup_{k\in K}\left| E\tilde{R}_*\left[F(k,\xi)\right] - E\tilde{R}_*\left[F(k,\tilde{\xi})\right]\right| . \qquad (3.68)$$

De (3.51) y la proposición 3.2. inciso (a), se tiene que

$$\tilde{R}_*(x) = \sup_{a\in A(x)}\left\{r(x,a) + E\tilde{R}_*\left[F(x,a,\tilde{\xi})\right]\right\} . \qquad (3.69)$$

Como $\tilde{R}_*$ está acotada por $M\|r\|$, del *supuesto 3.5.* incisos (a) y (c), en (3.69) se tiene que la función bajo el supremo es *Lipschitz* con respecto a $k=(x,a)$. Entonces, tal y como se muestra en Gordienko *et al.* (2009) y bajo el *supuesto 3.5.* inciso (b), prueba que la función $\tilde{R}_*$ en (3.69) es *Lipschitz.* Por lo tanto, aplicando (3.30) en el *supuesto 3.5.* inciso (d), para la función $s \to \tilde{R}_*\left[F(k,s)\right]$ en (3.68) se obtiene que esta función satisface las condiciones de *Lipschitz* con una constante que no depende de k .

De manera análoga, usando el *supuesto 3.5.* inciso (c), se puede confirmar que la función $s \to \tilde{R}_*\left[F(k,s)\right]$ es *Lipschitz*.

Finalmente, combinando las desigualdades (3.66), (3.67) y (3.68), se tiene que Λ_t en (3.64) es menor que $\sup\left|E\varphi(\xi) - E\varphi(\tilde{\xi})\right|$ sobre cierta clases de funciones φ , las cuales son acotadas por la misma constante $\bar{b}$ y que satisfacen las condiciones de *Lipschitz* con la misma constante $\bar{L}$ (y esta constante depende solamente de m , δ y las constantes involucradas en los *supuestos 3.4. y 3.5.*).

Por lo tanto,

$$\Lambda_t \leq 2(\bar{b}+\bar{L})\left[1+\frac{m}{1-\delta}\right]d(D_{\xi},D_{\tilde{\xi}}) \ ,$$

$$\leq 4(\bar{b}+\bar{L})\left[1+\frac{m}{1-\delta}\right]\pi(D_{\xi},D_{\tilde{\xi}}) \ , \qquad (3.70)$$

donde $d(D_{\xi},D_{\tilde{\xi}})$, $\pi(D_{\xi},D_{\tilde{\xi}})$ denotan *la distancia de Dudley*, y de *Prokhorov*, respectivamente, entre las distribuciones de los vectores aleatorios ξ y $\tilde{\xi}$. En las últimas dos desigualdades se usó (XX) y (XXI) para la definición de *la métrica de Dudley*, y la relación entre *las métricas de Dudley* (d) y de *Prokhorov* (π).

Si hacemos $\overline{K} = 4(\bar{b}+\bar{L})\left[1+\frac{m}{1-\delta}\right]$, entonces de (3.70) y (3.64) se tiene que

$$\sup_{x\in X_0} \Delta_R(x) \leq N\,\overline{K}\,\pi(D_{\xi},D_{\tilde{\xi}}) + \frac{bc}{1-\theta}\theta^{N+1} . \qquad (3.71)$$

Finalmente, la desigualdad deseada (3.31) en el Teorema 2 se obtiene de (3.71) si cambiamos

$$N = \left[\max\left\{1, \log_{\theta}\left(\frac{1}{\pi(\mu,\tilde{\mu})}\right)\right\}\right] + 1,$$

la constante $\overline{K}$ que aparece en (3.31) queda explícitamente calculada como,

$$\overline{K} = (\bar{b}+\bar{L})\left[1+\frac{m}{1-\delta}\right] \ , \qquad (3.71.1)$$

con $\quad \bar{b} = 2\,|\,r\,|\,(1+M)$,

y $\quad \bar{L} = (L_1+1)L_*\left[L_0 + \frac{L}{1-\delta}\max\left\{1, \frac{M|r|}{1-\delta}\right\}\right].$

Capítulo 4.

Evaluaciones asintóticas del índice de estabilidad con recompensa descontada cuando el coeficiente de descuento se aproxima a uno.

En este capítulo se considera un ejemplo particular de un *proceso de control de Markov* para el cual las políticas estacionarias óptimas pueden calcularse explícitamente. La función de recompensa en un paso se denotará por $r(x,a)$, y la optimalidad se realiza con respecto al criterio de *recompensa total* α*-descontada* $(\alpha \in (0,1))$, el cual se definió en (XI), y que se transcribe a continuación para el proceso original:

$$R_\alpha(x,\pi) := E_x^\pi \sum_{t=1}^{\infty} \alpha^{t-1} r(x_{t-1}, a_t), \quad x \in X \ , \quad \pi \in \Pi. \tag{4.1}$$

Por lo que la función de valor definida en (XIV), queda expresada para este caso, de la siguiente manera:

$$R^*_\alpha(x) := R_\alpha(x,\pi_*) = \sup_{\pi\in\Pi} R_\alpha(x,\pi), \quad x \in X \ .$$

De nuevo, reemplazando $\tilde{x}$ en lugar de x en (4.1) se obtiene la expresión del criterio de de *recompensa total* α*-descontada* ($\tilde{R}_\alpha(x)$) para el proceso aproximado, y de igual forma su función de valor $\tilde{R}^*_\alpha(x)$.

La forma explícita de las políticas estacionarias $\pi_* = f_*$ (para el proceso original, véase (4.5)) y de $\tilde{\pi}_* = \tilde{f}_*$ (para el proceso aproximado, dado en (4.6)) del ejemplo presentado en este capítulo, permite calcular explícitamente el *índice de estabilidad*, para este ejemplo.

Dicho *índice de estabilidad* bajo este criterio, se denotará como Δ_{R_α} y queda expresado, según la definición dada en (XVI), de la siguiente manera:

$$\Delta_{R_\alpha}(x) := R^*{}_\alpha(x, f_*) - R^*{}_\alpha(x, \tilde{f}_*) \geq 0\,, \tag{4.2}$$

el propósito de este capítulo es el de estudiar el comportamiento asintótico del *índice de estabilidad* $\Delta_{R_\alpha}(x)$ cuando $\alpha \uparrow 1$.

De (4.1), es claro que cuando $\alpha \uparrow 1$ se espera un crecimiento no acotado de R_α en (4.1), y posiblemente de $\Delta_{R_\alpha}(x)$ en (4.2).

Usando cálculos directos y aproximaciones numéricas de funciones, se muestra (en un caso particular) la existencia de constantes ς , γ tales que el *índice de estabilidad* $\Delta_{R_\alpha}(x)$ se puede expresar como una función que depende del término $(1-\alpha)$:

$$\Delta_{R_\alpha} = \Delta_{R_\alpha}\big((1-\alpha)\big) \sim \frac{\varsigma}{(1-\alpha)^\gamma} \quad \text{, cuando} \quad \alpha \to 1, \tag{4.3}$$

$$\text{donde} \qquad \gamma \leq 1. \tag{4.4}$$

Es importante resaltar que en el ejemplo considerado, para cada $\alpha < 1$ fijo, se tiene que $\Delta_{R_\alpha}(x) \to 0$ cuando $\kappa\big(D_\xi, D_{\hat{\xi}}\big) \to 0$, donde κ es la *métrica de Kantorovich* (véase apéndice B). En otras palabras, se considera el caso de MDP´s *estables*.

Ahora, compárese (4.3) y (4.4) con las *desigualdades de estabilidad* (1.7) y (1.12) que aparecen en los Teoremas 1.1. y 1.2. dados en el capítulo 1. Las constantes de la parte derecha de estas desigualdades crecen al infinito como $\dfrac{C}{(1-\alpha)^3}$ cuando $\alpha \uparrow 1$, y C es una constante. Esto es una señal de que las condiciones y los métodos de demostraciones usados en los Teoremas 1.1. y 1.2. no permiten obtener desigualdades bastantes buenas para el caso en el que el coeficiente de descuento esté cercano a 1.

Un enfoque que permite evitar este problema fue sugerido en Gordienko, Salem (2000). Sin embargo, “el precio” que se paga por el mejoramiento de las constantes en la parte derecha de las *desigualdades de estabilidad* hasta obtener $\frac{C}{(1-\alpha)}$ cuando $\alpha \uparrow 1$, es alto: Hay que usar condiciones restrictivas de ergodicidad del proceso aplicando políticas de control estacionarias.

4.1. Un proceso controlable de consumo-inversión y su aproximación.

Este ejemplo se presenta en el capítulo 6, sección 9, página 189 de Dynkin, Yushkevich (1979).

Sean $X=[0,\infty)$; $A=[0,\infty)$; $A(x)=[0,x]$, $x \in X$. La dinámica del proceso original está dada (como en (IV)) por las siguientes ecuaciones:

$$x_t = a_t \xi_t \quad , \quad \text{para } t = 1,2,... \quad ; \tag{4.5}$$

y para el proceso aproximado, por

$$\tilde{x}_t = \tilde{a}_t \tilde{\xi}_t \quad , \quad \text{para } t = 1,2,... \quad , \tag{4.6}$$

donde $\{\xi_t, t \geq 1\}$ y $\{\tilde{\xi}_t, t \geq 1\}$ son dos sucesiones de variables aleatorias *i.i.d.* no negativas.

En este modelo x_{t-1} se interpreta como el capital corriente. La cantidad $a_t \in [0, x_{t-1}]$ representa lo que se invierte en activos (tales como acciones, bonos, etc.), los cuales generan una ganancia/pérdida dada por $a_t \xi_t$. El resto del capital $(x_{t-1} - a_t)$ se dedica al consumo, y la satisfacción (o beneficio) de este consumo se estima por la función de utilidad dada por $(x_{t-1} - a_t)^p$, donde $0 < p < 1$ es un parámetro dado.

Por lo tanto, la función de recompensa por un paso involucrada en (4.1) está dada por

$$r(x_{t-1}, a_t) = \left(x_{t-1} - a_t\right)^p , \tag{4.7}$$

Con respecto a las variables aleatorias en (4.5) y (4.6) se asume que satisfacen lo siguiente:

$$\lambda := E\xi_1^p < \frac{1}{\alpha} \quad ; \quad \tilde{\lambda} \equiv \lambda_\varepsilon := E\tilde{\xi}_1^p < \frac{1}{\alpha} . \tag{4.8}$$

Bajo estas condiciones en el capítulo 6, sección 9, página 189 de Dynkin, Yushkevich (1979) se demuestra lo siguiente:

(a) La política estacionaria óptima para el proceso dado en (4.5) está dada por el siguiente selector:

$$f_*(x) = \left(\alpha\lambda\right)^{\frac{1}{1-p}} x \quad , \quad x \in [0, \infty) . \tag{4.9}$$

(b) La *función de valor* correspondiente al proceso original es:

$$R_*(x) \equiv R_\alpha(x, f_*) = \frac{1}{\left[1 - (\alpha\lambda)^{\frac{1}{1-p}}\right]^{1-p}} x^p \quad , \quad x \in [0, \infty) . \tag{4.10}$$

(c) La política estacionaria óptima para el proceso aproximado dado en (4.6), está dada por el siguiente selector:

$$\tilde{f}_*(x) = \left(\alpha\lambda_\varepsilon\right)^{\frac{1}{1-p}} x \quad , \quad x \in [0, \infty) . \tag{4.11}$$

Para efectos de estimaciones numéricas realizdas en este capíulo, considérese por ejemplo, que $\xi \sim \exp(\theta)$ y $\tilde{\xi} \sim \exp(\tilde{\theta})$, con $\tilde{\theta} = \theta(1-\varepsilon)$, donde los valores de ε miden la aproximación entre ambas densidades $\left(0 < \varepsilon < 1\right)$.

Entonces, de (4.8) se tiene que

$$\lambda := E\xi_1^p = \int_0^\infty \xi^p \frac{e^{-\xi/\theta}}{\theta} d\theta \,,$$

de donde se obtiene

$$\lambda = \theta^p \Gamma(p)\,. \tag{4.12}$$

Similarmente, para la variable aleatoria perturbada se encuentra que

$$\tilde{\lambda} \equiv \lambda_\varepsilon = \tilde{\theta}^{\,p} \Gamma(p)\,,$$

y dado que $\tilde{\theta} = \theta(1-\varepsilon)$, entonces de la igualdad anterior se obtiene que

$$\lambda_\varepsilon = \lambda(1-\varepsilon)^p\,. \tag{4.13}$$

4.2. Cálculo del índice de estabilidad para el ejemplo.

En esta sección se calcula el *índice de estabilidad* (Δ_{R_α}) cuando se aplican las políticas óptimas dadas en (4.9) y (4.11) a los procesos (4.5) y (4.6) respectivamente. Se tomará como estado inicial a $x = \tilde{x} = 1$.

Recordando de (4.2), el *índice de estabilidad* (Δ_{R_α}) queda definido como:

$$\Delta_{R_\alpha} = R_\alpha(1, f_*) - R_\alpha(1, \tilde{f}_*)\,. \tag{4.14}$$

El primer término del lado derecho de (4.14) está calculado explícitamente en (4.10) con $x = 1$.

Falta por calcular el segundo término del lado derecho de (4.14).

Para ello, se sustituye la política óptima $\tilde{f}_*$ obtenida para el proceso aproximado (que está dada en (4.11)) en la función dada en (4.1), tal y como se muestra a continuación:

$$R_\alpha(1,\tilde{f}_*) := E_1^{\tilde{f}_*}\sum_{t=1}^{\infty}\alpha^{t-1}(\tilde{x}_{t-1}-\tilde{a}_t)^p = \sum_{t=1}^{\infty}\alpha^{t-1}E_1^{\tilde{f}_*}\left[\tilde{x}_{t-1}-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\tilde{x}_{t-1}\right]^p = \sum_{t=1}^{\infty}\alpha^{t-1}E_1^{\tilde{f}_*}\left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^p\tilde{x}_{t-1}^p$$

$$R_\alpha(1,\tilde{f}_*) = \left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^p\sum_{t=0}^{\infty}\alpha^t E_1^{\tilde{f}_*}\tilde{x}_t^p\ . \qquad (4.15)$$

De (4.6), la evolución del proceso aproximado se puede representar de la siguiente manera:

$$\tilde{x}_1 = \tilde{a}_1\tilde{\xi}_1 = (\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}x\tilde{\xi}_1 = (\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\tilde{\xi}_1\ ,$$

$$\tilde{x}_2 = \tilde{a}_2\tilde{\xi}_2 = (\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\tilde{x}_1\tilde{\xi}_2 = (\alpha\lambda_\varepsilon)^{\frac{2}{1-p}}\tilde{\xi}_1\tilde{\xi}_2\ ,$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\tilde{x}_t = \tilde{a}_t\tilde{\xi}_t = (\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\tilde{x}_{t-1}\tilde{\xi}_t = (\alpha\lambda_\varepsilon)^{\frac{t}{1-p}}\tilde{\xi}_1\tilde{\xi}_2\cdots\tilde{\xi}_t\ .$$

Si la última igualdad anterior se eleva a la potencia p en ambos lados, se obtiene que

$$\tilde{x}_t^p = (\alpha\lambda_\varepsilon)^{\frac{pt}{1-p}}\tilde{\xi}_1^p\tilde{\xi}_2^p\cdots\tilde{\xi}_t^p\,.$$

Ahora, si se toma la esperanza en ambos lados de la última igualdad y dado que los elementos aleatorios son *i.i.d.*, se llega a que

$$E_1^{\tilde{f}_*}\tilde{x}_t^p = (\alpha\lambda_\varepsilon)^{\frac{pt}{1-p}}E_1^{\tilde{f}_*}(\xi_1^p)E_1^{\tilde{f}_*}(\xi_2^p)\cdots E_1^{\tilde{f}_*}(\xi_t^p)\,.$$

Luego, usando la condición impuesta en (4.8), resulta que

$$E_1^{\tilde{f}_*}\tilde{x}_t^p = (\alpha\lambda_\varepsilon)^{\frac{pt}{1-p}}\lambda_\varepsilon^t\,. \qquad (4.16)$$

Sustituyendo (4.16) en (4.15), y después de realizar algunos cálculos directos, se obtiene lo siguiente:

$$R_\alpha(1,\tilde{f}_*) = \left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^p \sum_{t=0}^{\infty}\left[(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^t . \tag{4.17}$$

La condición impuesta en (4.8) garantiza que $\alpha\lambda_\varepsilon < 1$. El hecho de que $0 < p < 1$, valida que $1 < \frac{1}{1-p}$. Las dos razones anteriores garantizan que $(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}} < 1$. Por lo tanto, (4.17) se puede expresar como

$$R_\alpha(1,\tilde{f}_*) = \left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^p \frac{1}{1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}} ,$$

o bien

$$R_\alpha(1,\tilde{f}_*) = \frac{1}{\left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^{1-p}} . \tag{4.18}$$

Luego, para obtener *el índice de estabilidad* se sustituye (4.10) y (4.18) en (4.14), obteniéndose que

$$\Delta_{R_\alpha} = \frac{1}{\left[1-(\alpha\lambda)^{\frac{1}{1-p}}\right]^{1-p}} - \frac{1}{\left[1-(\alpha\lambda_\varepsilon)^{\frac{1}{1-p}}\right]^{1-p}} . \tag{4.19}$$

Si se sustituye en (4.19) la expresión de λ_ε dada en (4.13), se obtiene lo siguiente:

$$\Delta_{R_\alpha} = \frac{1}{\left[1-(\alpha\lambda)^{\frac{1}{1-p}}\right]^{1-p}} - \frac{1}{\left[1-(\alpha\lambda)^{\frac{1}{1-p}}(1-\varepsilon)^{\frac{p}{1-p}}\right]^{1-p}} . \tag{4.20}$$

En esta última expresión el *índice de estabilidad* depende del factor de descuento α y del parámetro p de la función de recompensa descontada, donde $0<\alpha<1$ y $0<p<1$.

Para cada p fijo se puede escoger θ en (4.12) de tal modo que $\lambda=1$.

Entonces (4.20) puede reescribirse como:

$$\Delta_{R_\alpha} = \frac{1}{\left[1-\alpha^{\frac{1}{1-p}}\right]^{1-p}} - \frac{1}{\left[1-\alpha^{\frac{1}{1-p}}(1-\varepsilon)^{\frac{p}{1-p}}\right]^{1-p}} . \qquad (4.21)$$

Es fácil ver que para $\varepsilon>0$, el segundo término en la parte derecha de (4.21) está acotado cuando $\alpha\uparrow 1$. Por lo que,

$$\Delta_{R_\alpha} \sim \frac{1}{\left[1-\alpha^{\frac{1}{1-p}}\right]^{1-p}} , \text{ con } \alpha\uparrow 1 . \qquad (4.22)$$

4.3. El estudio del comportamiento asintótico del índice de estabilidad.

Sea por ejemplo $p=\frac{1}{2}$. Entonces de (4.22), :

$$\Delta_{R_\alpha} \sim \frac{1}{(1-\alpha)^{\frac{1}{2}}(1+\alpha)^{\frac{1}{2}}} \sim \frac{1}{(1-\alpha)^{\frac{1}{2}}} \quad , \quad \text{cuando } \alpha\uparrow 1.$$

Del mismo modo, para los valores de p cercanos a cero, de (4.22) Δ_{R_α} se comporta aproximadamente como $\frac{1}{1-\alpha}$ (con $\alpha\uparrow 1$).

Para observar un panorama más amplio para otros valores de p, se realizaron los siguientes experimentos numéricos.

Para varios valores fijos de p (ver la tabla más abajo) se aproximó la siguiente función (véase 4.22):

$$\varphi(\alpha) = \frac{1}{\left[1-\alpha^{\frac{1}{1-p}}\right]^{1-p}} \quad , \text{ con } \alpha \in [0.5,1), \tag{4.23}$$

por una función

$$\psi((1-\alpha)) = \frac{\varsigma}{(1-\alpha)^{\gamma}} \quad , \text{ con } \alpha \in [0.5,1), \tag{4.24}$$

donde las constantes ς y γ se ajustan por el método de mínimos cuadrados (considerando desviaciones cuadráticas entre $\varphi(\alpha)$ y $\psi((1-\alpha))$ para números bastante grandes del argumento α).

Los resultados obtenidos se muestran en la siguiente tabla:

Valor del parámetro p aplicado en (4.23)	Estimación numérica de γ, dada en (4.24)
$1/100$	0.989
$1/50$	0.98
$1/20$	0.95
$1/10$	0.90
$1/9$	0.88
$1/5$	0.78
$1/4$	0.74
$3/10$	0.67

$^{2}/_{5}$	0.57
$^{1}/_{2}$	0.46
$^{3}/_{5}$	0.35
$^{7}/_{10}$	0.25
$^{3}/_{4}$	0.24
$^{4}/_{5}$	0.14
$^{8}/_{9}$	0.10
$^{9}/_{10}$	0.08
$^{99}/_{100}$	0.06

Notas:

(a) Es interesante que para valores de p cercanos a uno, asintóticamente el *índice de estabilidad* depende muy poco de α (cuando $\alpha \uparrow 1$).
La naturaleza de este fenómeno todavía no es clara.

(b) Para $p = \frac{1}{2}$, $\Delta_{R_\alpha} \sim \frac{1}{(1-\alpha)^{\frac{1}{2}}}$ con $\alpha \uparrow 1$. De los resultados obtenidos numéricamente, se sigue que $\Delta_{R_\alpha} \approx \frac{0.81}{(1-\alpha)^{0.46}}$.

Tal diferencia se explica por el hecho de que se hace la aproximación en el intervalo de $\alpha \in [0.5, 1)$.

Capítulo 5.

Conclusiones e investigaciones futuras.

5.1. Conclusiones.

Pese a la basta literatura que existe acerca del tema de procesos controlables de Markov, son escazos los trabajos desarrollados en el tema de estimación de la estabilidad.

El estudio de la estabilidad para procesos controlables de Markov representa un reto, tanto desde el punto de vista teórico como práctico. Proponer *métricas probabilísticas* apropiadas para lograr las llamadas *desigualdades de estabilidad* es todo un esfuerzo adicional.

En la literatura revisada se han propuesto el uso de diferentes tipos de métricas para este logro, sin embargo algunas de ellas, o son bastante restrictivas (las llamadas "métricas fuertes") o bien, se obtienen *desigualdades de estabilidad* en términos de otra métrica no deseada. Es en este caso, cuando tiene relevancia el conocimiento de las relaciones que hay entre los diferentes tipos de *métricas probabilísticas*, para cambiar la desigualdad del índice a la métrica deseada.

Este tabajo de investigación pretende contribuir al estudio de la estabilidad. Particularmente se propone el uso de la *métrica de Prokhorov*.

En esta tesis se estudia el problema de estimación del *índice de estabilidad* para cierta clase de procesos estocásticos controlables de Markov usando la *métrica de Prokhorov*.

La importancia de poder usar la *métrica de Prokhorov*, reside en el hecho de que para problemas de aplicación permite realizar estimaciones del *índice de estabilidad* bajo el uso de distribuciones empíricas para los elementos aleatorios, ya que éstas convergen débilmente bajo esta métrica a las distribuciones que se pretende estimar (a diferencia de las llamadas "métricas fuertes").

En la tesis se trabaja con tres criterios de optimalidad comúnmente estudiados en la literatura de control óptimo: *costo descontado*, *costo promedio* y *recompensa total*. Usando estos criterios mencionados y bajo condiciones generales se obtienen nuevas *desigualdades de estabilidad* con el uso de la *métrica de Prokhorov*.

Los resultados obtenidos usando el *criterio de costo descontado* (Teoremas 1.1 y 1.2.) están basados en la aplicación de ciertas condiciones de *Lipschitz*, así como ciertas propiedades de contractividad (en los operadores involucrados en las ecuaciones de optimalidad). Por otro lado, en el caso del *criterio de costo promedio*, los resultados encontrados están basados en el uso de la semi-norma span, así como en las condiciones de ergodicidad y de *Lipschitz* para obtener el Teorema 2.1. Finalmente, usando el *criterio de de recompensa total*, el resultado encontrado y que es presentado en el Teorema 3.2. fue deducido bajo ciertas condiciones de transitoriedad y de *Lipschitz* (para la función de recompensa, para el *kernel estocástico*, etc.), así como el uso de ciertos operadores contractivos.

Las desigualdades encontradas para el *índice de estabilidad* para procesos transitorios con *recompensa total esperada* obtenidas en el capítulo 3 no se encuentran en la literatura, hasta ahora. Esta clase de procesos controlables de Markov tienen muchas aplicaciones importantes (en problemas de paro óptimo, por ejemplo).

Junto con la demostración de estabilidad con respecto de las *métricas de Prokhorov* y de *Dudley* (que permiten usar aproximaciones empíricas), también se obtuvieron desigualdades en términos de la distancia de *Variación Total*. Estas últimas, permiten relajar las condiciones usadas y por ello, ampliar la clase de aplicaciones potenciales.

Los resultados del capítulo 2 - *desigualdades de estabilidad* usando el criterio promedio con respecto a la *métrica de Prokhorov* - son desarrollos de métodos del artículo Gordienko *et al* (2009), en donde se obtuvieron cotas parecidas en términos de la *métrica de Kantorovich* para costos no acotados (ver apéndice B para la definición de la *métrica de Kantorovich*). Dicho artículo y la tesis presente usan al parámetro $\delta \in (0,1)$, al que es posible llamar "parámetro de ergodicidad" del proceso y que en ejemplos frecuentemente toma un valor cercano a 1. El uso de la *métrica de Prokhorov* y funciones de costo acotado permitieron mejorar el órden de la constante en la parte derecha de la *desigualdad de estabilidad* correspondiente. En Gordienko *et al* (2009) esta constante es del orden $O(1-\delta)^{-3}$ con $\delta \uparrow 1$, mientras que el resultado obtenido en la tesis (ver (2.10), capítulo 2) la constante tiene un orden de $O(1-\delta)^{-2}$.

Se resalta nuevamente que los principales resultados de este trabajo de investigación son *desigualdades de estabilidad* obtenidas en los capítulos 2 y 3 (con ejemplos correspondientes).

Los resultados de los capítulos 1 y 4 sirven como una fuente de los ejemplos y apuntaciones de las direcciones para investigaciones futuras.

5.2. Investigaciones futuras planeadas.

5.2.1. Criterio Sensible al Riesgo.

En las últimas décadas ha recibido mucha atención un criterio de optimalidad (en parte, en relación con algunos procesos controlables en matemáticas financieras) llamado *criterio sensible al riesgo*.

El estudio de la teoría de control bajo riesgo sensible es un área significativa en teoría de control estocástico.

El estudio de los procesos de decisión de Markov (MDPs) incorporando *el criterio sensible al riesgo* para medir el rendimiento de la política de control fueron considerados primero por Bellman (1957) para modelos con espacios de estado finitos. Un profundo análisis apareció en Howard, Matheson (1972) para el caso de espacios de estados finitos, donde cada cadena controlada es irreducible y aperiódica. El caso general con espacios de estados finitos fué tratado en Rothblum (1984), Fleming, Hernández-Hernández (1997), Cavazos-Cadena, Fernández-Gaucherand (1999) y las referencias que ahí se mencionan.

Problemas con *costo descontado sensible al riesgo* son estudiados en Chung, Sobel (1987), un resultado sorprendente es que, para problemas con horizonte infinito y costos descontados, en general, la política óptima no necesariamente es estacionaria.

Problemas con *costos promedio sensible al riesgo* en horizonte infinito han sido estudiados en Fleming, Hernández-Hernández (1997a, 1997b), Hernández-Hernández, Marcus (1996,1997), Howard, Matheson (1972), Borkar, Meyn (2002), Cavazos-Cadena, Fernández-Gaucherand (1999), Whittle (1990) para espacios de estados numerables y en Di-Masi, Stettner (2000a, 2000b), Dupuis, Ellis (1997) para espacios de estados generales.

En el artículo de Cavazos-Cadena, Montes-de-Oca (2000) se estudian cadenas de Markov controladas con espacio de estados finitos, recompensa no-negativa y el rendimiento usado de una política de control es medido por el criterio de *recompensa total esperada sensible al riesgo*. Los autores estudian las condiciones para la existencia de políticas óptimas estacionarias.

Criterios de optimalidad sensibles al riesgo. Sea λ un factor de riesgo dado. Para un estado inicial $x \in X$ y una política $\pi \in \Pi$, definamos los siguientes criterios de optimización sensibles al riesgo:

A. Criterio de costo descontado sensible al riesgo.

$$V_\alpha(\lambda, x, \pi) := \frac{1}{\lambda} \log E_x^\pi \exp\left\{\lambda \sum_{t=0}^{\infty} \alpha^t c(x_t, a_t)\right\}, \tag{5.1}$$

B. Criterio de *costo promedio sensible al riesgo*.

$$J(\lambda, x, \pi) := \limsup_{n\to\infty} \frac{1}{\lambda n} \log E_x^\pi \exp\left\{\lambda \sum_{t=0}^{n-1} c(x_t, a_t)\right\}, \tag{5.2}$$

C. Criterio de *costo total esperado sensible al riesgo*.

$$V(\lambda, x, \pi) := \limsup_{n\to\infty} \frac{1}{\lambda} \log E_x^\pi \exp\left\{\lambda \sum_{t=0}^{\infty} c(x_t, a_t)\right\}, \tag{5.3}$$

Un agente tomador de decisiones se dice que tiene adversión al riesgo si $\lambda > 0$, mientras que si es arriesgado (risk seeking) entonces $\lambda < 0$. Finalmente si $\lambda = 0$, el agente es neutral al riesgo.

En general, los artículos sobre criterios *sensibles al riesgo* estudian *las condiciones que garanticen la existencia de políticas óptimas* para los criterios (5.1) - (5.3). Hasta donde se llegó en esta investigación, en la litertura existe un número limitado de artículos que traten el tema de estabilidad usando este nuevo criterio.

Uno de ellos, Montes-de-Oca, Zaitseva (2014) se ofrece una estimación cuantitativa de la estabilidad del *costo total esperado sensible al riesgo* (usando $\lambda < 0$, en (5.3)) para un problema de paro òptimo de una cadena de Markov en espacios de Borel. Se proporciona una cota superior para el *índice de estabilidad* expresada en términos de la *distancia de Variación Total*. Además, en dicho artículo se presenta un ejemplo sencillo de un problema de optimización inestable, es decir,

$\mathbf{Vr}\ (p, \tilde{p}) \to 0$ mientras que el índice $\Delta \geq M > 0$, donde M es una constante arbitrariamente grande, pero fija.

Avila-Godoy (1998) obtiene condiciones para la existencia de políticas óptimas para modelos con espacio de estados y de acciones finitos. En Jáskiewicz (2007) se estudian MDP´s a tiempo discreto en espacios de estados generales, con función de costo no-negativa y bajo *el criterio de costo promedio sensible al riesgo*. La autora establece la desigualdad de optimización y obtiene políticas óptimas estacionarias.

En Cavazos-Fernández-Gaucherand (1999) se trabajan cadenas de Markov con espacio de estados numerables, función de costo acotado, y usando el criterio de *costo promedio sensible al riesgo*, asociado a una función de utilidad exponencial. Bajo las condiciones de Doeblin, los autores encuentran que si el factor de riesgo es lo suficientemente pequeño, entonces la ecuación de optimización asociada tiene una solución. Así mismo, se muestra que bajo condiciones de continuidad-compacidad, las políticas óptimas estacionarias existen.

En Cavazos-Cadena, Hernández-Hernández (2005) se usa *el criterio de costo promedio sensible al riesgo* y se prueba una caracterización de la función de valor óptimo.

En Borkar, Meyn (2002) se trabajan modelos con el criterio de *costo promedio sensible al riesgo* y funciones de costo acotadas, los autores imponen formas fuertes de *ergodicidad uniforme* para mostrar que existen funciones de valor que están acotadas, también establecen las condiciones suficientes (con $\lambda \to 0$) para la existencia de políticas óptimas estacionarias.

Finalmente vale la pena mencionar los siguientes dos hechos: Uno, que el control *sensible al riesgo* tiene aplicaciones naturales en administración de portafolios, donde el objetivo es maximizar el crecimiento de la utilidad esperada de la riqueza, ver por ejemplo Bielecki, Hernández-Hernández, Pliska (1999), Bielecki, Pliska (1999), Stettner (1999). Dos, que uno de los resultados clave que han sido explorados es la relación que existe entre control estocástico sensible al riesgo y la teoría de juegos, consulte Dai-Pra, Meneghini, Runggaldier (1996), Fleming, McEneaney (1995), Runolfsson (1994).

5.2.2. Problemas propuestos.

1. **Cotas para el *índice de estabilidad* en el caso de *criterio promedio sensible al riesgo*.**

   Este criterio de optimalidad está definido en (5.2). Se plantea el problema de encontrar las condiciones de ergodicidad convenientes que permitan probar unas propiedades de contracción de operadores involucrados en las “ecuaciones de optimalidad” para este criterio. Luego, bajo unas condiciones de *Lipschitz* acerca de $F(\cdot,\cdot,\xi)$, $F(x,a,\cdot)$ y $A(x)$ con $x \in X$, es posible que se pueda establecer que tales operadores preservan continuidad de Lipschitz. Parece también que bajo condiciones de ergodicidad el *índice de estabilidad* $\Delta_{J_\lambda}(x) = J(\lambda, x, \tilde{\pi}_*) - J(\lambda, x, \pi_*)$ no depende de x.

   Tomando en cuenta tales observaciones se plantea el problema de lograr demostrar una *desigualdad de estabilidad* (véase (XVII)) como la siguiente:

   $$\Delta_{J_\lambda} \leq B(\lambda)\pi(D_\xi, D_{\tilde{\xi}}), \tag{5.4}$$

   donde la constante $B(\lambda)$ depende del parámetro de sensibilidad $\lambda > 0$. Es de notarse que debido a la estructura de $J(\lambda, x, \pi)$ en (5.2), el problema de establecer una *“desigualdad de estabilidad*” de la forma (5.4) es más complicada que el problema clásico (riesgo neutral). Esta afirmación se apoya por la dificultad del problema (considerada en la literatura) de búsqueda de políticas óptimas con respecto al criterio $J(\lambda, x, \pi)$.

2. **Experimentos numéricos sobre la dependencia de los *índices de estabilidad* Δ_{V_α} y Δ_{J_λ} con respecto a los parámetros α y λ.**

De la desigualdad obtenida para el *costo descontado* $V_\alpha(x,\pi)$ (cuando $A(x)$ no depende de x), se sigue que para cada $x \in X$ fijo, el valor de la constante K_α en la parte derecha de (1.12) tiene un orden de $M(1-\alpha)^{-3}$ cuando el coeficiente de descuento $\alpha \to 1$. En el capítulo 4, a partir de un proceso particular (y sencillo) de control, se realizaron estimaciones numéricas para obtener que el *índice de estabilidad* dado en (4.22) tiene un orden de $M(1-\alpha)^{-1}$ cuando el coeficiente de descuento $\alpha \to 1$ y $p \to 0$.

Por lo que el problema que se plantea es en este caso, primero obtener una *desigualdad de estabilidad* como la dada en (5.4), y posteriormente determinar la forma en que la constante B(λ) dependerá del parámetro de sensibilidad al riesgo λ .

En el artículo de Montes-de-Oca, Zaitseva (2014) se usa $\lambda < 0$ (risk seeking) y el criterio (5.3). En el Teorema 3.2. de dicho artículo, se presenta una desigualdad como la mostrada en (5.4) con la diferencia de que los autores usan la *distancia de Variación Total* **Vr**, en lugar de π, *la métrica de Prokhorov* como se propone en (5.4). En dicho Teorema 3.2, la constante $B(\lambda)$ tiende a infinito exponencialmente cuando $\lambda \to -\infty$ (y en este orden es correcto, como muestran los ejemplos).

Para $\lambda > 0$, el problema sigue abierto.

Por otro lado, se plantea el problema de hacer experimentos numéricos y simulación (usando ejemplos simples) para ver la manera en que el *índice de estabilidad* Δ_{J_λ} depende de λ > 0.

Referencias

ARAPOSTATHIS, A., BORKAR, V.S., FERNANDEZ-GAUCHERAND, E., GHOSH, M.K., MARCUS, S.I. (1993). Discrete-time controlled Markov processes with average cost criterion: A survey *SIAM J. Control and Optimization*, 31, 282-344.

ASMUSSEN, S. (1987). *Applied Probability and Queues*. Wiley, N.Y.

AVILA-GODOY, M.G. (1998). Controlled Markov chains with exponential risk-sensitive criteria: modularity, structured policies and explications, *Ph.D. Dissertation*, Dept. of Math., Univ. of Arizona, Tucson, AZ.

BALAJI, S., MEYN, S.P. (2000). Multiplicative ergodicity and large deviations for an irreducible Markov chain. *Stoch. Process Appl*, 90, 123-144.

BÄURLE, N., RIEDER, U. (2011). Markov Decision Processes with Applications to Finance. Springer-Verlang Berlin Heidelberg.

BELLMAN, R. (1957). Dynamic Proggraming. Princenton University Press, Princenton, N.J.

BERGE, C. (1963). Topological Spaces. Dover Publications, Inc. Mineola, New York.

BERSEKAS, D.P., TSITSIKLIS, J.N. (1991). An analysis of stochastic shortest path problems. *Math. Oper. Res.*, 16, 3, 580-595.

BIELECKI, T., PLISKA, S. (1999). Risk-sensitive dynamic asset management. *Appl. Math. Optim.*, 39, 337-360.

BIELECKI, T., HERNÁNDEZ-HERNÁNDEZ, D., PLISKA, S. (1999). Risk-sensitive control of finite state Markov chains in discrete time, with applications to portfolio management. *Math. Methods Oper. Res.* 50, 167-188.

BORKAR, V.S., MEYN, S.P. (2002). Risk-sensitive optimal control for Markov decision processes with monotone cost. *Math. Oper. Res.*, 27, 192-209.

BOROVKOV, A.A. (1977). Some estimates of the rate convergences in stability theorems. *Theory Probab. Appl.*, 22, 668-678.

BOROVKOV, A.A., FOSS, S.G. (1992). Stochastically recursive sequences and their generalization. *Siberian and Adv. Math.*, 2, 16-18.

CAVAZOS-CADENA, R., FERNÁNDEZ-GAUCHERAND, E. (1999). Controlled Markov chains with risk-sensitive criteria: Averange cost, optimal equations and optimal solutions. *Math. Methods Oper. Res.*, 49, 299-324.

CAVAZOS-CADENA, R., HERNÁNDEZ-HERNÁNDEZ, D. (2005). A characterization of the optimal risk-sensitive average cost in finite controlled Markov chains. *The Annals of Appl. Probab.*, 15, 175-212.

CAVAZOS-CADENA, R., MONTES-DE-OCA, R. (2000). Optimal stationary policies in risk-sensitive dynamic programs with finite state space and nonnegative rewards. *Applic. Math.*, 27, 167-185.

CHUNG, K.J., SOBEL, M.J. (1987). Discounted MDP´s: Distribution functions and exponential utility maximization. *SIAM J. Control and Optim*., 25, 49-62.

DAI-PRA, P., MENEGHINI, L., RUNGGALDIER, W.J. (1996). Some connections between stochastic control and dynamic games. *Mathematics of Control, Signals and Systems*, 9, 303-326.

DI-MASI,G.B., STETTNER, L. (2000a). Risk-sensitive control of discrete-time Markov processes with infinite horizon. *SIAM J. Control Opt.*, 38. 61-78.

DI-MASI,G.B., STETTNER, L. (2000b). Infinite horizon risk sensitive control of discrete time Markov processes with small risk. *Systems Control Lett.*, 40, 15-20.

DUDLEY, R.M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.*, 39, 1563-1572.

DUDLEY, R.M. (1969). The speed of mean Glivenko-Cantelli convergence. *Ann. Math. Stat.*, 40, 40-50.

DUDLEY, R.M. (1989). *Real Analysis and Probability*. Wadsworth & Brooks/Cole, Pacific Grove.

DUDLEY, R.M., (2002). Real analysis and probability, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press., Cambride. Revised reprint of the 1989 original.

DUNCAN, T.E., PASIK-DUNCAN, B., STETTNER, L. (2001). Risk-sensitive adaptive control of discrete time markov processes. *Probab. Math. Statist.*, 21, 493-512.

DUPUIS, P., ELLIS, R.S. (1997). A Weak Convergence Approach to the Theory of Large Deviations. Wiley, New York.

DYNKIN, E.B., YUSHKEVICH, A.A. (1979). Controlled Markov Processes. Springer-Verlang, New York.

ENGL, H.W., HOFINGER, A., KINDERMANN, S. (2005). Convergence rates in the Prokhorov metric for assessing uncertainty in ill-posed problems, *Inverse problems, 21, pp. 399-412.*

FAVERO, G., RUNGGALDIER W.J. (2002). A robustness result for stochastic control. *Syst. Contr. Lett.*, 46, 91-97.

FLEMING, W.H., McENEANEY, W.M. (1995). Risk-sensitive control and differential games. *Springer Lecture Notes in Control and Info. Sci*. No. 184, 185-197.

FLEMING, W.H., McENEANEY, W.M. (1992). Risk-sensitive control on an infinite horizon. *SIAM J. Control and Optim.*, 33, 1881-1915.

FLEMING, W.H., HERNÁNDEZ-HERNÁNDEZ, (1997a). Risk sensitive control of finite state machines on an infinite horizon I. *SIAM J. Control and Optim.*, 35, 1790-1810.

FLEMING, W.H., HERNÁNDEZ-HERNÁNDEZ, (1997b). Risk sensitive control of finite state machines on an infinite horizon II. Technical Report, *Division of Applied Mathematics*, Brown Univeristy.

GIBBS, A.L., SU F.E. (2002). On choosing and bounding probability metrics. *Int. Stat. Rev., pp. 419-435.*

GORDIENKO, E.I. (1985). Adaptive strateges for certain classes of controlled Markov processes. *Theory Prob. Appl.*, 29, 504-518.

GORDIENKO, E.I. (1988). Stability estimates for controlled Markov chains with a minorant. J*. Soviet Math.* 40, 481-486.

GORDIENKO, E.I. (1992). An estimate of the stability of optimal control of certain stochastic and deterministic systems. *J. Soviet Math.*, 59, 891-899 (Traslation from the Russian edition of 1989).

GORDIENKO, E.I. (1994). Lectures Notes on Stability Estimation in Markov Decision Processes. UAM-I, México D.F.

GORDIENKO, E.I., HERNANDEZ-LERMA, O.(1995). Averange cost Markov control processes with weight norms existence of canonical policies. *Appl. Math.* 23:199-218.

GORDIENKO, E.I., SALEM, F.S. (1998). Robustness inequality for Markov control processes with unbounded costs. *Syst. Contr. Lett.*, 33, 125-130.

GORDIENKO, E.I., SALEM, F. (2000). Estimates of stability of Markov controlled processes with unbounded costs. *Kybernetika*, 36, 195-210.

GORDIENKO, E.I., YUSHKEVICH, A. (2003). Stability estimates in the problem of optimal switching of a Markov chain*. Math. Meth. Oper. Res.*, 57, 345-365.

GORDIENKO, E.I., LEMUS-RODRIGUEZ, E., MONTES-DE-OCA, R. (2008). Discounted cost optimality problema: stability with respect to weak metrics. *Math. Meth. Oper. Res.*, 68, 77-96.

GORDIENKO, E.I., LEMUS-RODRIGUEZ, E., MONTES-DE-OCA, R. (2009). Average cost Markov control processes: stability with respect to the Kantorovich metric. *Math. Meth. Oper. Res.*,70,13-33.

GORDIENKO, E.I., NOVIKOV, A. (2014). Characterization of optimal policies in a general stopping problema and stability estimating. *Probability in the Enginneering and Informational Sciences*, 28, 03, 335-352.

GORDIENKO, E.I., MARTINEZ, J., RUIZ-DE-CHAVEZ, J. (2015). Stability estimation of transition Markov decisión processes. *Chapter XI Symposium on Probability and Stochastic Processes. Vol.69 of the series Progress in Probability pp157-176.*

GUJARATI, D.N. (2010). Econometría. Mc-Graw Hill.

HENSEN, L.P., SARGENT, T.J. (1995). Discounted linear exponential quadratic Gaussian control, *IEEE Transactions on Automatic Control*, 40, 968-971.

HERNÁNDEZ-HERNÁNDEZ, D., MARCUS, S.I. (1996). Risk-sensitive control of Markov processes in countable state spaces. *Systems and Control Lett*., 29, 147- 155.

HERNÁNDEZ-HERNÁNDEZ, D., MARCUS, S.I. (1997). Existence of risk sensitive optimal stationary policies for controlled Markov processes. *Appl. Math. Optim.*, 40, 273-285.

HERNANDEZ-LERMA, O. (1989). *Adaptive Markov Control Processes*. Springer-Verlag, New York.

HERNANDEZ-LERMA, O., CARRASCO, G., PEREZ-HERNANDEZ, R. (1999). Markov control processes with the expected total cost criterion: optimality stability, and transient models. *Acta Appl. Math.*, 59, 3, 229-269.

HERNÁNDEZ-LERMA, O., LASSERRE, J.B. (1996). *Discrete-Time Markov Control Processes: Basic Optimality Criteria*. Springer, N.Y.

HERNANDEZ-LERMA, O., LASSERRE, J.B. (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer, N.Y.

HINDERER, K. (2005). Lipschitz continuity of value functions in markovian decision processes. *Math. Meth. Oper. Res.*, 62, 3-22.

HINDERER, K., WALDMANN K.H. (2005). Algorithms for countable state Markov decision models with an absorbing set. *SIAM J. Control Optim.*, 43, 6, 2109-2131 (electronic).

HORDIJK, A. (1974). Dynamic programming and Markov potential theory, volume No. 51 of *Mathematical Centre Tracts*. Mathematisch Centrum, Amsterdam.

HOWARD, R.A., MATHERSON, J.E (1972). Risk Sensitive Markov decision processes. *Management Sci .* 8, 356-369.

HUBER, P.J. (1981). Robust statistics, *John Wiley and Sons, New York.*

INTRILIGATOR, M.D. (1996). Econometric models, techniques, and applications. Pie Impren Upper Saddle River, NJ: Prentice Hall.

JACOBSON, D.H. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Transactions on Automatic Control*, 18, 124-131.

JAMES, H.W., COLLINS, E.J. (2006). An analyisis of transient Markov decision processes. *J. Appl. Probab.*, 43, 3, 603-621.

JASKIEWICZ, A. (2007). Averange optimality for risk-sensitive control with general state space. *Ann. Appl. Probab.*, 17, 654-675.

KALASHNIKOV, V. (1983). The analysis of continuity of queueing systems. In: *Lecture Notes in Math.*, Springer.

KALASHNIKOV, V., RACHEV,S.T. (1988). Mathematical Methods for Constructions of Stochastic Queuing Models. *Nauka, Moscu. (Engl. transl. (1990) Wadsworth, Brooks-Cole, Pacific Grive, California).*

KALASHNIKOV, V. (1995). Quantitative estimates in queueing., In: Advances in Queueing Theory, Methods and open Problems. (Ed. J.H. Dshalalow) *CRC Press*, Boca Raton, 404-427.

KALLENBERG, L.C.(1983). Linear programming and finite Markovian control problems, volume 148 of *Mathematical Centre Tracts. Mathematish Centrum, Amsterdam*.

KARTASHOV, N.V. (1985). Inequalities in theorems of ergodicity and stability of Markov chains with common phase space, II.*Theory Probab. Appl.*, 30, 507-515.

KITAEV, M., RYKOV, V. (1995). *Controlled Queueing Systems*. *CRC Press*, Boca Raton, Florida.

KREPS, D.M., PORTEUS, E.L. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, 46, 185-200.

LARAKI, R., SUDDERTH, W.D. (2004). The preservation of continuity and Lipschitz continuity by optimal reward operations. *Math. Oper. Res.*, 29, 672-685.

MARTINEZ, J., ZAITSEVA, E. (2015). Note on Stability estimation in average Markov Control processe. *Kybernetika*, Vol. 51, No. 4, 629- 638.

MEYN, S.P., TWEEDIE, R.L. (2009). *Markov Chain and Stochastic Stability*. Cambridge University Press, Cambridge.

MONTES-DE-OCA, R., SAKHAVENKO, A., SALEM-SILVA, F. (2003). Estimates for perturbations of general discounted Markov control chains. *Appl. Math.*, 30, 287-304.

MONTES-DE-OCA, R., SALEM-SILVA, F. (2005). Estimates for perturbations of average Markov decisión processes with a minimal state and upper bounded by stochastically ordered Markov chains. *Kybernetika*, 41, 757-772.

MONTES-DE-OCA, R., ZAITSEVA, E. (2014). About stability of risk seeking optimal stopping. *Kybernetika*, 50, 378-392.

MULLER, A. (1997). How does the value function of a Markov decision process depend on the transition probabilities?. *Math. Oper. Res.*, 22, 872-885.

PLISKA, S.R. (1978). On the transient case for Markov decision chains with general state space. *In Dynamic programming and its applications (Proc. Conf. Univ. British Columbia, Vancuver, B.C., 1977) pages 335-349 Academic Press, New-York-London.*

PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorem of probability theory. *Theory Probab. Appl.,* 2, 177-238 (Russian Publication).

RACHEV, S.T., RÜSCHENDORF, L. (1998). *Mass Transportation Problem*. Vol II: Applications Springer, New York.

RACHEV, S.T. (1991). *Probability Metrics and the Stability of Stochastic Models*. Wiley, Chichester.

ROTHBLUM, U.G. (1984). Multiplicative Markov decision chains. *Math. Oper. Res.* 9, 6-24.

ROSS, S.M. (1992). Applied probability models with optimizations. *Dover Publications, Inc., New York. Reprint of the 1970 original.*

RUNOLFSSON, T. (1994). The equivalence between infinite horizon control of stochastic systems with exponential-of-integral performance index and stochastic differential games. *IEEE Transactions on Automatic Control* 39, 1551-1563.

SENATOV, V.V. (1998). *Normal Aproximation: New Results, Methods and Problems*. VSP, Utrecht.

SHIRYAYEV, A.N., (1978). Optimal stopping rules. *Springer-Verlang, New York-Heidelberg translated from Russian by A.B. Aries, Applications of Mathematics, Vol. 8.*

TWEEDIE, R.L. (2001). Markov chains: Structure and applications. In: *Handbook of Statist.*, Amsterdam.

STETTNER, L. (1999). Risk sensitive portfolio optimization. *Math. Methods Oper. Res.*, 50, 463-474.

VAN DIJK, N.M. (1988). Perturbation theory for unbounded Markov reward processes with applications to queueing. *Adv. Appl. Probab.*, 20, 99-111.

VAN DIJK, N.M., PUTERMAN, M.L. (1988). Perturbation theory for Markov reward processes with applications to queueing systems. *Adv. Appl. Probab.*, 20, 79-98.

VAN DER VAART, A.W., WELLNER, J.A. (1996). *Weak Convergence and Empirical Processes*. Springer, N.Y.

VAN DIJK, N.M, SLADKY, K. (1999). Error bounds for nonnegative dynamic models. *J. Optim. Theory Appl.*, 101, 449-474.

VEGA-AMAYA, O. (2003). The average cost optimality equation: a fixed point approach. *Bol. Soc. Math. Mexicana,* 9, 185-195.

VEINOTT, A.F., (1969). Discrete dynamic programming with sensitive discount optimality criteria. *Ann. Math. Statist.* 40:1635-1660.

WESSELS, J. (1977). Markov Programming by successive approximation with respect to weighted supremum norms. *J. Math. Anal. Appl.*, 58, 326-335.

WHITTLE, P. (1990). Risk Sensitive Optimal Control. Jhon Wiley & Sons, New York.

YAKOWITZ, S. (1982). Dynamic programming applications in water resources. *Water Resources Res.*, 18, 673-696.

ZAITSEVA, E. (2008). Stability estimating in optimal stopping problem. *Kybernetika,* 44(3):400-415.

ZAITSEVA, E. (2010). Robustness estimating of optimal stopping problem with unbounded revenue and cost functions. *Int. J. Pure Math.*, 59(3):291-306.

ZOLOTAREV, V. M. (1976a). On stochastic continuity of queueing systems of type G/G/1. *Theory Probab. Appl.*, 21, 250-269.

ZOLOTAREV, V.M. (1976b). Metric distances in space of random variables and their distributions. *Math. URSS Sbornik*, 30, 373-401.

APENDICE A
Algunas definiciones y resultados utilizados en la tesis.

Espacio de Borel. Sea $\mathcal{X}$ un espacio métrico, separable y completo. La mínima σ-algebra $\mathcal{B}_{\mathcal{X}}$ de subconjuntos de $\mathcal{X}$ que contiene todos los subconjuntos abiertos (cerrados) se llama *σ-algebra de Borel*. Cada $\mathcal{B} \in \mathcal{B}_{\mathcal{X}}$ se llama *conjunto de Borel*.

Por otro lado, X se llama *el espacio de Borel* si existe un espacio métrico, separable y completo $\mathcal{X}$ tal que $X \in \mathcal{B}_{\mathcal{X}}$. Es claro que en particular $X = \mathcal{X}$ es de Borel.

Algunos ejemplos importantes, son los siguientes:

(i) Subconjuntos de Borel de $\mathcal{R}^n$;

(ii) espacio numerable con la topología discreta;

(iii) el espacio de medidas de probabilidad $\mathcal{M}$ definidasen la σ-algebra de Borel de un espacio métrico, separable, completo S.

En el último ejemplo, $\mathcal{M}$ se equipa con la *métrica de Prokhorov* definida en (XVIII) del capítulo preliminar así como en el apéndice B.

Distribuciones de vectores alatorios y su convergencia débil. Sea (Ω, F, P) un espacio de probabilidad, y (S, r) un espacio métrico separable con σ-algebra de Borel $\mathcal{B}_S$. La aplicación medible $X : \Omega \to S$ se llama vector aleatorio con valores en S.

La distribución de X es la siguiente medida de probabilidad P_X en $(S, \mathcal{B}_S)$, $P_X(\mathcal{B}) := P(X^{-1}(\mathcal{B}))$, $(:= P(X \in \mathcal{B}))$.

Sean X, X_n, $n \geq 1$, vectores aleatorios en S con distribuciones correspondientes P_X, P_{X_n}, $n = 1,2,\dots$. Se dice que X_n *converge* a X *débilmente* (o P_{X_n} *converge* a

P_X *débilmente*) si para cada $f \in C_b(S)$, $Ef(X_n) \equiv \int_S f dP_n \to \int_S f dP \equiv Ef(X)$. Donde $C_b(S)$ es el espacio de todas las funciones continuas y acotadas $f : S \to \mathfrak{R}$.

En tal caso se escribe $X_n \Rightarrow X$, o bien , $P_{X_n} \Rightarrow P_X$.

Es bien conocido (Prokhorov (1956)) que $X_n \Rightarrow X$ si y sólo si $\pi(X_n, X) \to 0$. Lo anterior es equivalente a $d(X_n, X) \to 0$, donde d es la métrica de Dudley definida en (XX) en el capítulo preliminar, así como en el apéndice.

Distribuciones empíricas y sus convergencias. Sean $X_1, X_2, \ldots X_n$ vectores aleatorios *i.i.d.* definidos en un mismo espacio de probabilidad (Ω, F, P). La distribución empírica es la siguiente *medida de probabilidad* aleatoria:

$$\hat{P}_n := \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k} \quad ,$$

definida en (S, $\mathcal{B}_S$). Además, δ_{X_k} es la medida de Dirac, definida de la siguiente manera:

Para $B \in \mathcal{B}_S$,

$$\delta_{X_k}(B) = \begin{cases} 1 & si \quad X_k \in B. \\ 0 & si \quad X_k \notin B. \end{cases}$$

Se sabe que, (consulte, por ejemplo Rachev (1991)) con probabilidad 1, $\pi(\hat{P}_n, P) \to 0$; pero evidentemente $\mathbf{Var}(P_n, P) = 2$ para $n = 1,2,\ldots$

Kernel estocástico. Sean X y Y dos espacios de Borel. Un *kernel estocástico* q (Borel medible) en X dado Y es una función tal que para cada $y \in Y$, $q(\cdot | y)$ es una medida de probabilidad en (X , $\mathcal{B}_X$), y para cada $B \in \mathcal{B}_X$, $q(B | \cdot)$ es una función medible de X a [0,1].

Para $X = Y$ a cada kernel estocástico q corresponde una familia de procesos de Markov homogéneos a tiempo discreto $X_0, X_1, \ldots X_t$ (con diferentes distribuciones iniciales) con la probabilidad de transición (por un paso) dada por $P(X_{t+1} \in B \mid X_t = x) = q(B \mid x)$, $B \in \mathcal{B}_X$, $x \in X$.

La probabilidad de transición por n - pasos se define como:

$P(X_{t+n} \in B \mid X_t = x) = q^{(n)}(B \mid x) := q \otimes q^{(n-1)}(B \mid x)$, $B \in \mathcal{B}_X$, $x \in X$,

donde para dos kernels q y $\tilde{q}$ se tiene que $q \otimes \tilde{q}(B \mid x) = \int_X q(B \mid x)\tilde{q}(dy \mid x)$.

Sean $r : X \to \mathcal{R}$ una función medible y acotada, y $x \in X$ un estado inicial del proceso de Markov con kernel q. Es demostrable que para cada $n \geq 1$ se cumple que $E\big(r(X_n) \mid X_0 = x\big) = \int_X r(y)q^{(n)}(dy \mid x)$.

De la misma manera se define el operador $Q^{(n)} : \mathbb{B} \to \mathbb{B}$, donde $\mathbb{B}$ es el espacio de Banach de todas las funciones acotadas $\varphi : X \to \mathcal{R}$ con la norma del supremo.

Por definición:

$$Q^{(n)}\varphi(x) : \int_X \varphi(y)q^{(n)}(dy \mid x) \ , \ \varphi \in \mathbb{B} \ , \ x \in X \ .$$

Se dice que p en (X , $\mathcal{B}_X$) es una probabilidad invariante para el kernel q , si $p(B) = \int_X q(B \mid x)p(dx)$ para toda $B \in \mathcal{B}_X$.

Si p existe y la distribución inicial (de X_0) es p , entonces $P(X_n \in B) = P(B)$ para toda $B \in \mathcal{B}_X$, $n = 1,2,\ldots$.

Función inferiormente (superiormente) semicontinua. Sea X un espacio métrico y $f : X \to \mathcal{R} \cup \{\infty\}$. Se dice que f es *inferiormente semicontinua* en $x \in \mathcal{R}$ si $\liminf_{n\to\infty} f(x_n) \geq f(x)$ para cualquier sucesión $x_n \to x$. (Para superiormente continua: $\limsup_{n\to\infty} f(x_n) \leq f(x)$).

La función f es llamada inferiormente semicontinua si es semicontinua para todo $x \in X$.

Multifuncional y selectores medibles. Sean (X, r) y (A, d) dos espacios métricos y 2^A el conjunto de todos los subconjuntos de A.

Un multifuncional $F : X \to 2^A$ se llama Borel medible si para cada $B \in \mathcal{B}_A$, el conjunto $F^{-1}(B) := \{x \in X \mid F(x) \cap B \neq \phi\}$ pertenece a $\mathcal{B}_X$.

El siguiente resultado se usa frecuentemente para justificar la existencia de políticas de control estacionarias óptimas.

Se denotará por $\mathcal{C} \subset 2^A$ a la colección de todos los subconjuntos compactos en A, y sea $F : X \to \mathcal{C}$ un multifuncional Borel medible.

Sean también $\mathbb{K} := \{(x,a) \in X \times A \mid a \in F(x), x \in X\}$ y $u : \mathbb{K} \to \mathcal{R}$ una función medible.

<u>Proposición</u>. Supongase que para cada $x \in X$, la función $u(x,\cdot)$ es inferiormente continua en $F(x)$. Entonces:

(a) Existe una función medible $f : X \to A$ tal que,

- $f(x) \in F(x)$ para toda $x \in X$.
- $u(x, f(x)) = \inf_{a \in F(x)} u(x,a)$, $x \in X$.

(b) La función $u_*(x) := \inf_{a \in F(x)} u(x,a)$, $x \in X$ es medible.

A la función f a veces se le llama también selector.

APENDICE B
Algunas métricas probabilísticas simples.

Las primeras tres métricas usadas en la tesis ya se definieron en el capítulo preliminar. Aquí simplemente se recuerdan.

Sean (S,r) un espacio métrico separable, X, Y vectores aleatorios con valores en S. Recuérdese que una métrica μ simple entre X y Y significa la métrica (distancia) entre las distribuciones P_X y P_Y de X y Y. En este sentido, se aplica la denotación $\mu(X,Y)=\mu(P_X,P_Y)$.

1. **Métrica de Prokhorov, π.** (Véase (XVIII)).

$$\pi(P_X,P_Y):=\inf\left\{\varepsilon>0: P_X(B)\le P_Y(A^\varepsilon)+\varepsilon \text{ , para toda } B\in \mathscr{B}_S\right\},$$

y A^ε es "la ε - vecindad de A".

2. **Métrica de Dudley, d.**

$$d(X,Y):=\sup_{\varphi\in D}|E\varphi(x)-E\varphi(y)| \ ,$$

donde $$D:=\left\{\varphi: S\to \mathscr{R}: \|\varphi\|_\infty+\|\varphi\|_L\le 1\right\} .$$

(Véase (XX) en el capítulo preliminar).

3. **Métrica de Lévy, L.**

$$L(X,Y):=\inf\left\{\varepsilon>0: F_X(x-\varepsilon)-\varepsilon\le F_Y(x)\le F_X(x+\varepsilon)+\varepsilon \text{ para toda } x\in\mathscr{R}\right\}$$

Se sabe que $X_n\Rightarrow X$ (débilmente) si y sólo si $\pi(X_n,X)\to 0$ si y sólo si $d(X_n,X)\to 0$ si y sólo si $L(X_n,X)\to 0$, para $S=\mathscr{R}$ en el último caso.

4. Métrica uniforme (o de Kolmogorov), ρ.

Para $S = R^n$,

$$\rho(X,Y) := \sup_{x \in R^n} | F_X(x) - F_Y(x) | .$$

Se sabe que $\rho(X_n, X) \to 0$ implica la convergencia débil, y para X absolutamente continuos, $X_n \Rightarrow X$ resulta $\rho(X_n, X) \to 0$.

5. Métrica Kantorovich (o de Wasserstein), κ.

$$\kappa(X,Y) := \sup_{\varphi \in Lip_1} | E\varphi(x) - E\varphi(y) | ,$$

donde $Lip_1 : \{\varphi : S \to \mathcal{R} : | \varphi(x) - \varphi(y) | \leq r(x,y) \quad x, y \in \mathcal{R} \ \}$.

La convergencia en κ: $\kappa(X_n, X) \to 0$ es equivalente a $X_n \Rightarrow X$, y además $Er(X_n, a) \to Er(x,a)$ para un $a \in S$.

Se sabe que para $S = \mathcal{R}$, $\kappa(X,Y) = \int_{-\infty}^{\infty} |F_X(x) - F_Y(x)| \, dx$.

6. Métrica de Variación Total, $\mathbf{Vr}$.

$$\mathbf{Vr}(X,Y) := \sup_{\varphi \in B_1} | E\varphi(x) - E\varphi(y) | ,$$

donde $B_1 := \{\varphi \to \mathcal{R} : \|\varphi\|_\infty \leq 1\}$.

Se sabe que:

- $\mathbf{Vr}(X,Y) = 2 \sup_{B \in B_S} | P(X \in B) - P(Y \in B) |$.
- Cuando $S = \mathcal{R}$ y X, Y son absolutamente continuas con densidades f_X y f_Y, entonces:

  $$\mathbf{Vr}(X,Y) = \int_{-\infty}^{\infty} | f_X(x) - f_Y(x) | \, dx .$$

7. Un par de ejemplos de métricas probabilísticas compuestas.

7.1. Métrica de Ky Fan, K.

$$K(X,Y) := \inf \{\varepsilon > 0 : P(r(X,Y) > \varepsilon) < \varepsilon\} .$$

7.2. Métrica L_1.

$$L_1(X,Y) := E\,|\,X - Y\,|\,.$$

Los valores de K y L_1 dependen de las distribuciones conjuntas $P_{X,Y}$ del par de vectores aleatorios X, Y.

8.- **La métrica de Hausdorff**. Sea (A,d) un espacio métrico y $\mathcal{C}$ la familia de todos los conjuntos compactos en A.

La *métrica de Hausdorff* h, se define en $\mathcal{C}$ por la siguiente regla ($B,C \in \mathcal{C}$):

$$h(B,C) := \max\left\{\sup_{x\in B} d(x,C), \sup_{y\in C} d(y,B)\right\},$$

donde $d(x,C) = \inf_{Z\in C} d(x,z)$.

Dos resultados importantes:

Teorema (Strassen).

$\inf\{K(X,Y)$ sobre todas las $P_{X,Y}$ con distribuciones marginales fijas P_X y $P_Y\}$ =

$= \pi(P_X, P_Y)$.

Teorema (Kantorovich-Dudley).

$\inf\{L_1(X,Y)$ sobre todas las $P_{X,Y}$ con distribuciones marginales fijas P_X y $P_Y\}$ =

$= \kappa(P_X, P_Y)$.

APENDICE C
Algunas métricas probabilísticas simples.

En este apéndice se proporcionan las ecuaciones de optimalidad para los criterios usados en esta tesis, así como las condiciones para la existencia de políticas estacionarias óptimas en cada uno de ellos

.

Criterio de Costo Total Esperado Descontado.

El problema de control de interés en esta sección es el de la minimización del *costo total esperado descontado* con horizonte infinito. Sea el modelo de control $M=\left(X, A, \{A(x): x\in X\}, q, c\right)$, como el dado en (I) . El criterio a ser minimizado es el dado en (VIII), mismo que se transcribe a continuación:

$$V_\alpha(x,\pi)=E_x^\pi\left[\sum_{t=0}^{\infty}\alpha^t c(x_t,a_t)\right] \quad , \quad \pi\in\Pi \quad , \quad x\in X \ , \qquad \text{(C.1)}$$

donde $\alpha\in(0,1)$ es el factor de descuento, y se asume que es dado.

Una política π_* que satisfaga

$$V_\alpha(x,\pi_*)=\inf_{\pi\in\Pi}V_\alpha(x,\pi)=V_\alpha^*(x) \quad , \text{ para toda } x\in X \ , \qquad \text{(C.2)}$$

se llama política α - óptima descontada, y V_α^* es llamada función de valor α -descontada.

Las ecuaciones de optimalidad costo descontadas. Una función $v: X\rightarrow R$ se dice que es una solución para las ecuaciones de optimalidad de costo α -descontado (α *-DCOE*), si satisface:

$$v(x)=\min_{a\in A(x)}\left[\ c(x,a)+\alpha\int_X v(y)q\left(dy \mid x,a\right)\ \right] \quad , \text{ para toda } x\in X \ . \qquad \text{(C.3)}$$

Por otro lado, las funciones iteradas de α -valor se definen como:

$$v_n(x) := \min_{a \in A(x)} \left[\ c(x,a) + \alpha \int_X v_{n-1}(y) q\left(dy \middle| x, a\right) \ \right] ,$$

Para toda $x \in X$ y $n = 1,2,\ldots$, con $v_0 \equiv 0$.

Entonces,

$$V_\alpha^*(x) = \lim_{n \to \infty} v_n(x) \quad , \text{ para toda } x \in X \text{ .}$$

<u>*Supuestos 1.*</u>

(a) La función de costo por un paso, c , es inferiormente semicontinua, no negativa e inferiormente compacta en $\mathbb{K} := \{(x,a) \in X \times A \mid a \in F(x), x \in X\}$.

(b) El kernel q es fuertemente continuo.

<u>*Supuestos 2.*</u>

Existe una política π tl que $V_\alpha(x,\pi) < \infty$ para toda $x \in X$.

Teorema 1. Si se asume que los supuestos 1 y 2 se cumplen, entonces:

(i) La función de valor α - descontada dada en (C.2) de este apéndice, es la solución mínima para la α *-DCOE*. Es decir,

$$V_\alpha^*(x) = \min_{a \in A(x)} \left[\ c(x,a) + \alpha \int_X V_\alpha^*(y) q\left(dy \middle| x, a\right) \ \right] \text{ , para toda } x \in X \text{ .} \tag{C.4}$$

Y si u es otra solución para la α *-DCOE*, entonces $u(\cdot) \geq V_\alpha^*$.

(ii) Existe un selector $f_* \in \mathbb{F}$ tal que $f_* \in A(x)$ y alcanza el mínimo en (C.4), es decir,

$$V_\alpha^*(x) = c(x, f_*) + \alpha \int_X V_\alpha^*(y) q\left(dy \middle| x, f_*\right) \text{ , para toda } x \in X \text{ .} \tag{C.5}$$

Y la política estacionaria determinística f_*^∞ es α - óptima descontada.

(iii) Si una política α - óptima descontada existe, entonces existe una que es estacionaria determinística.

Para una demostración de este Teorema véase la sección 4.2, capìtulo 4 de Hernandez-Lerma , Lasserre (1996).

NOTA:

Los supuestos 1 y 2 dados en este apéndice aseguran la existencia de una política estacionaria $f_* \in \Pi$ que satisfaga la ecuación (C.5). Más aún, bajo estos suuestos, la parte entre paréntesis rectangulares de (C.4), es decir, $c(x,a)+\alpha\int_X V_\alpha^*(y)q(dy|x,a)$, es una función medible en $(x,a)\in \mathbb{K}$, y continua en $a\in A(x)$ para toda $x\in X$, donde $A(x)$ es un conjunto compacto.

Costo Promedio

Sea $(X, A, \{A(x)|x\in X\}, q, c)$ un modelo de Markov controlable como el dado en (I), y

$$J_n(x,\pi) := E_x^\pi\left[\sum_{t=0}^{n-1} c(x_t, a_t)\right]. \tag{C.6}$$

El costo total esperado en n etapas cuando se aplica la política π, dado el estado inicial $x_0 = x$.

El *costo promedio* (*AC*) esperado de largo plazo cuando se usa $\pi\in\Pi$, dado $x_0 = x$, es

$$J(x,\pi) := \limsup_{n\to\infty} \frac{J_n(x,\pi)}{n} \ . \tag{C.7}$$

El problema *AC* es encontrar una política π_* tal que,

$$J(x,\pi_*) = \inf_{\pi\in\Pi} J(x,\pi) =: J^*(x) \quad , \text{ para toda } x\in X \ . \tag{C.8}$$

Una política que satisfaga (C.8) se dice que es *AC*- óptima y J^* es llamada función de valor *AC*.

Sean $\rho, h: X \to R$ funciones medibles, y $f \in \mathbb{F}$ un selector dado. Entonces, (ρ, h, f) se dice que es una triada canónica, si para toda $x \in X$ y $n = 0,1,...$ se cumple que

$$J_n\left(f^{\infty}, x, h\right) = J_n^*(x,h) = n\rho(x) + h(x) \quad ,$$

donde

$$J_n^*(x,h) = \inf_{\pi \in \Pi} J_n(x,\pi,h) = \inf_{\pi \in \Pi}\left\{J_n(x,\pi) + E_x^{\pi} h(x_n)\right\}.$$

En el caso especial en el cual para una triada canónica (ρ, h, f), se tenga que $\rho(\cdot)$ es una constante, digamos $\rho(x) = \rho^*$ para toda $x \in X$, entonces las ***ecuaciones de optimalidad de costo promedio*** (*ACOE*) está definidas por:

$$\rho^* + h(x) = \inf_{a \in A(x)}\left[\; c(x,a) + \int_X h(y) q(dy|x,a) \;\right],$$

y si $f \in \mathbb{F}$, entonces se satisface que,

$$\rho^* + h(x) = c(x, f) + \int_X h(y) q(dy|x, f)\,.$$

Por otro lado, recuérdese la ecuación de costo α -óptimo descontado dado en (C.4), misma que se transcribe a continuación:

$$V_\alpha^*(x) = \min_{a \in A(x)}\left[\; c(x,a) + \alpha \int_X V_\alpha^*(y) q\left(dy|x,a\right) \;\right].$$

<u>*Supuestos 3.*</u>

Existe un estado $z \in X$, un número $\beta \in (0,1)$ y $M \geq 0$ tal que:

(a) $(1-\alpha) V_\alpha^*(z) \leq M$ para toda $\alpha \in [\beta, 1)$, donde V_α^* está dada en (C.4).

(b) Existe una constante $N \geq 0$ y una función no negativa $b(\cdot)$ en X tal que $-N \leq h_\alpha(x) \leq b(x)$ para toda $x \in X$ y $\alpha \in [\beta,1)$, donde $h_\alpha(x) := V_\alpha^*(x) - V_\alpha^*(z)$.

El supuesto 3 implica que $V_\alpha^* < \infty$ para toda $x \in X$ y $\alpha \in (0,1)$. En efecto, por a) $V_\alpha^*(z) < \infty$ para todo $\alpha \in [\beta,1)$, así que, por el lado derecho de la desigualdad dada en b), $V_\alpha^*(x) \leq b(x) + V_\alpha^*(z)$ para toda $x \in X$ y $\alpha \in [\beta,1)$. Por lo tanto, el mapeo $\alpha \mapsto V_\alpha^*(x)$ es no decreciente en $\alpha \in (0,1)$, para toda $x \in X$.

Lema. Si se asume el supuesto 3, entonces existe una constante ρ^*, con $0 \leq \rho^* \leq M$, y una sucesión del factor de descuento $\alpha(n) \uparrow 1$ tal que satisface $\lim_{n\to\infty}(1-\alpha(n))V_{\alpha(n)}^*(z) = \rho^*$.

Para una demostración de este Lema, véase sección 5.4, capítulo 4 de Hernandez-Lerma , Lasserre (1996).

<u>*Supuestos 4.*</u>

Si se asume el supuesto 3, y además:

(a) La función $b(\cdot)$ (dada en el supuesto 3, inciso (b)) es medible y tal que, para cada $x \in X$ y $a \in A(x)$: $\int_X b(y)q(dy|x,a) < \infty$.

(b) La sucesión $\{h_{\alpha(n)}\}$ es equicontinua.

Teorema 2. Si se asumen los supuestos 1 y 4, entonces existen una constante ρ^*, una función continua $h(\cdot)$ en X, y un selector $f \in \mathbb{F}$ que satisfacen la *ACOE*, es decir,

$$\rho^* + h(x) = \min_{a\in A(x)}\left[\ c(x,a) + \int_X h(y)q(dy|x,a)\ \right], \tag{C.9}$$

y

$$\rho^* + h(x) = c(x,f) + \int_X h(y)q(dy|x,f)\,. \tag{C.10}$$

Más aún, f^{∞} es *AC*-óptima y $J^{*}(x)=J(f^{\infty},x)=\rho^{*}$ para toda $x\in X$; de hecho, cada política estacionaria determinística f^{∞}, para la cual f satisfaga (C.10) es *AC*-óptima.

Para una demostración de este Lema, véase sección 5.5, capítulo 4 de Hernandez-Lerma , Lasserre (1996).