



Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas
Posgrado en Ciencias Matemáticas

Caracterización de Políticas Óptimas (s, S) para una Clase de Procesos de Decisión de Markov

Tesis

presentada para obtener el título de
Maestría en Ciencias Matemáticas

Presenta
Rubén Blancas Rivera

Director de Tesis
Dr. Hugo Adán Cruz Suárez

Puebla, Puebla. Julio 2018

Agradecimientos

Me gustaría agradecer a todas las personas que de alguna u otra manera formaron parte de la realización de este proyecto.

En primer lugar al Consejo Nacional de Ciencia y Tecnología (CONACYT), por su apoyo al proyecto, que hizo posible la elaboración de esta tesis.

Al Dr. Hugo Adán Cruz Suárez, asesor de tesis, por apoyarme durante la elección y desarrollo del tema, por el tiempo dedicado a la elaboración y revisión, sin su dirección esto no habría sido posible.

A mis sinodales: Dr. Fernando Velasco Luna, Dr. Víctor Hugo Vazquez Guevara, Dr. Francisco Tajonar Sanabria y Dra. Hortensia Josefina Reyes Cervantes, por sus contribuciones y correcciones al trabajo, además de su disponibilidad y deseo de ayudar.

A mi familia, por todo su apoyo. Gracias mamá, papá, abuelita y Rosy.

A ti mi amor Isa, por todo tu apoyo y ser mi complice en todos los proyectos que enfrentamos.

A mis amigos: Yazmin, Ciria, Erick, Toño y Paulino.

A la Vicerrectoría de Investigación y Estudios de Posgrado, por conducto del proyecto con clave BUAP-CA-30, se recibió el apoyo para la impresión de esta tesis.

Introducción

Este trabajo está relacionado con la teoría de los Procesos de Decisión de Markov (PDMs) a tiempo discreto ([12] y [13]). Los PDMs son utilizados para modelar un sistema que es observado de forma discreta en el tiempo y el cual cuenta con la propiedad de Markov. En el trabajo desarrollamos la teoría básica de PDMs para el estudio de un modelo de inventarios, basado en la caminata aleatoria de Lindley.

De manera general, un PDM modela un sistema dinámico cuyos estados son observados de manera periódica y es aplicado un control. El desarrollo de un PDM, a través del tiempo está dado de acuerdo al siguiente procedimiento. En cada tiempo t , $t = 0, 1, \dots$, se elige un control que se aplicará dependiendo del estado del sistema. Entonces como consecuencia del estado actual y de haber aplicado un control se paga un costo y el sistema se traslada a un nuevo estado en el instante de tiempo $t + 1$, mediante una ley de transición. Al ocurrir un estado en $t + 1$, el proceso se repite. De esta manera se obtiene una sucesión de controles a la cual se le denomina política. Con la finalidad de medir la calidad de una política, el PDM está dotado de una función real llamada criterio de rendimiento. Este trabajo se enfoca en el criterio de costo total descontado. En general, el problema de control óptimo consiste en encontrar una política que minimice el criterio de rendimiento, a tal política se le llama óptima, y al criterio de rendimiento evaluado en la política óptima se le llama función de valor.

En este trabajo de tesis se presenta el desarrollo teórico del criterio de costo total descontado para el caso de Borel. Posteriormente, estudiamos un modelo de inventarios, el cual es motivado por la caminata aleatoria de Lindley. Mediante la teoría de los PDMs demostramos la existencia de una política óptima estacionaria markoviana. Después, como resultado principal de este trabajo se demuestra para este modelo generalizado la optimalidad de una clase políticas conocidas en el contexto de sistemas de inventarios como (s, S) . En general, una política del tipo (s, S) consiste en un reordenamiento de hasta S productos si el nivel de almacenaje se encuentra por debajo de s , en caso contrario no se ordena. Es importante señalar que existen diversos trabajos que estudian este tipo de políticas óptimas en modelos matemáticos aplicados a sistemas de inventarios. En este trabajo, presentamos de manera resumida los avances que se han logrado para encontrar optimalidad de políticas (s, S) en diversos modelos de inventario. Posteriormente, con el modelo de inventarios motivado por la caminata aleatoria de Lindley con espacio de estados y acciones no necesariamente compactos probamos la optimalidad de políticas (s, S) .

La tesis está estructurada de la siguiente manera.

En el Capítulo 1, presentamos la teoría básica de los PDMs, basados en la literatura clásica [12] y [13]. En el Capítulo 2, estudiamos los modelos de inventarios en particular los estocásticos. Se tienen tres modelos matemáticos que aproximan el estudio de los sistemas de inventarios más conocidos; un modelo lineal, uno basado en el propuesto por David Lindley en [17] y una generalización del modelo de Lindley. Después realizamos una revisión de los resultados más importantes que demuestran optimalidad de políticas de la clase (s, S) . En el Capítulo 3, se prueba la optimalidad de políticas (s, S) para el modelo propuesto, dicho resultado se considera una de las principales aportaciones de este trabajo de tesis. Por último, presentamos una aplicación de este modelo, para ello se considera una base de datos indicada en [19], para una compañía textil en la venta de poliéster de Indonesia, donde realizamos una aproximación numérica a la política óptima de la clase (s, S) .

Índice general

Introducción	VIII
Índice general	IX
1. Procesos de decisión de Markov	1
1.1. Modelos de control de Markov	1
1.2. Políticas	2
1.3. Construcción del proceso de Markov	3
1.4. Criterio de rendimiento	4
1.5. Programación dinámica	5
1.6. Horizonte finito	5
1.7. Horizonte infinito	8
2. Modelos de inventarios	15
2.1. Modelos matemáticos	16
2.2. Antecedentes	18
2.3. Modelo con demanda no suplida	24
3. Optimalidad de la política (s, S)	31
3.1. Análisis de las funciones de valor óptimo	31
3.1.1. Política miope	38
3.2. Ejemplo numérico	42
3.2.1. Gestión de inventarios en Shanghai Tang Fashion Company	42
3.2.2. Política óptima (s, S)	45
Conclusiones	49
A. Resultados y Definiciones Auxiliares	51
Bibliografía	55

Capítulo 1

Procesos de decisión de Markov

El principal objetivo en este capítulo es introducir formalmente los conceptos básicos y los teoremas más importantes de Procesos de Decisión de Markov a tiempo discreto, que serán requeridos más adelante para la caracterización de la política óptima del modelo que estudiamos.

1.1. Modelos de control de Markov

Un Proceso de Decisión de Markov (PDM) consiste en estudiar un sistema que es observado de forma discreta en el tiempo bajo incertidumbre en su dinámica. Para construir un PDM es necesario considerar las siguientes componentes:

$$(X, A, \{A(x)|x \in X\}, Q, C). \quad (1.1)$$

Los conjuntos X y A son llamados espacios de estados y acciones (o controles), respectivamente, ambos conjuntos son subespacios de Borel pertenecientes a espacios polacos (espacios métricos y completos). La colección de subconjuntos no vacíos $\{A(x)|x \in X\}$ de A , donde cada conjunto es llamado espacio de controles (o acciones) admisibles en el estado $x \in X$. Finalmente, Q es un kernel estocástico o una ley de transición definido en X dado $X \times A$ y $C : X \times A \rightarrow \mathbb{R}$ es una función medible de costo en un paso.

La quintupla en (1.1) es conocida como Modelo de Control de Markov estacionario a tiempo discreto (MCM).

Un MCM estacionario a tiempo discreto es un sistema estocástico controlado que se observa de manera periódica en los tiempos

$t = 0, 1, 2, \dots$. La dinámica que describe este sistema funciona de la siguiente manera:

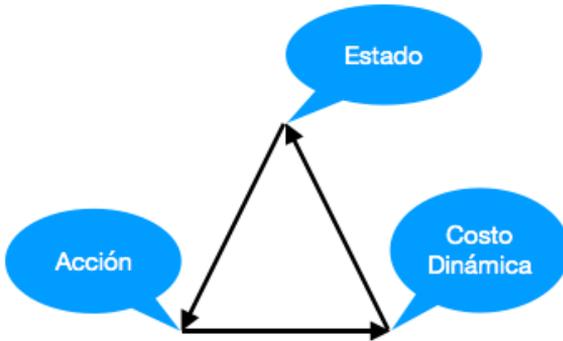


Figura 1.1: Comportamiento de un MCM

Si el sistema se encuentra al tiempo t en el estado $x_t = x \in X$, y la acción $a_t \in A(x) \subset A$ es aplicada, entonces ocurren dos cosas:

- se paga un costo $c(x, a)$,
- el sistema se traslada a un nuevo estado x_{t+1} mediante la ley de transición $Q(\cdot|x, a)$ sobre X , es decir,

$$Q(B|x, a) = P(x_{t+1} \in B|x_t = x, a_t = a),$$

en donde $B \in \mathcal{B}(X)$, con $\mathcal{B}(X)$ la σ -álgebra de Borel en X y $(x, a) \in \mathbb{K} := \{(x, a) \in A \times X | a \in A(x)\}$.

Una vez hecha esta transición a un nuevo estado, se elige una nueva acción y el proceso anterior se repite; como se muestra en la Figura 1.1.

Suposición 1.1.1 *Supongamos que \mathbb{K} contiene la gráfica de una función medible de X a A ; esto es, existe una función medible $f : X \rightarrow A$ tal que $f(x) \in A(x)$ para cada $x \in X$. La familia de tales funciones se denotará por \mathbb{F} .*

1.2. Políticas

En esta sección se va a definir formalmente el concepto de política en el contexto de MCM. Primero considere el espacio de historias

observadas en un proceso de control hasta un tiempo $t = 0, 1, \dots$, denotado por \mathbb{H}_t , como:

$$\begin{aligned}\mathbb{H}_0 &= X, \\ \mathbb{H}_t &= \mathbb{K} \times \mathbb{H}_{t-1} = \mathbb{K}^t \times X, \quad t = 1, 2, \dots\end{aligned}$$

Cada elemento $h_t \in \mathbb{H}_t$, se llama t -historia, el cual es un vector de la forma $(x_0, a_0, x_1, a_1, \dots, a_{t-1}, x_t)$, en donde $(x_i, a_i) \in \mathbb{K}$ para cada $i = 0, 1, \dots, t-1$ y $x_t \in X$.

Definición 1.2.1 Una política es una sucesión $\pi = \{\pi_t\}$ de kernels estocásticos, donde cada π_t está definido sobre A dado \mathbb{H}_t y satisface que: $\pi_t(A(x_t)|h_t) = 1$ para cada $h_t \in \mathbb{H}_t$ con $t = 0, 1, \dots$. El conjunto de todas las políticas se denota por Π .

De acuerdo con esta definición, una política $\pi = \{\pi_t\}$ puede interpretarse como una sucesión $\{a_t\}$ de variables aleatorias sobre A , tales que para cada t -historia y $t = 0, 1, \dots$, la distribución de a_t es $\pi_t(\cdot|h_t)$, la cual está concentrada en el conjunto de acciones admisibles $A(x_t)$. En otras palabras, cuando usamos una política arbitraria, la acción en cualquier tiempo t es una variable aleatoria que depende de todas las t -historias.

Definición 1.2.2 Una política $\pi \in \Pi$ es **Determinista Markoviana**, si existe una sucesión $\{f_t\} \subset \mathbb{F}$, en donde

$$\mathbb{F} = \{f : X \rightarrow A | f \text{ es medible y } f(x) \in A(x) \text{ para cada } x \in X\},$$

con la característica de $\pi(\cdot|h_t)$ está concentrada $f_t(x_t)$ para cada $h_t \in \mathbb{H}_t$ y $t = 0, 1, 2, \dots$

Si existe $f \in \mathbb{F}$ tal que para toda $t = 0, 1, 2, \dots$, $f_t = f$, entonces a π se le conoce como política **Determinista Markoviana Estacionaria**.

Observación 1.2.3 Se entiende que $\pi(\cdot|h_t)$ está concentrada en $f(x)$ si $\pi_t(C|h_t) = I_C(f(x))$ para cada $C \in \mathcal{B}(A)$. En donde $\mathcal{B}(A)$ denota la σ -álgebra de Borel de A e I_C es la función indicadora sobre C .

1.3. Construcción del proceso de Markov

Sea (Ω, \mathcal{F}) el espacio medible que consiste del espacio muestral canónico,

$$\Omega := (X \times A)^\infty$$

con \mathcal{F} su correspondiente σ -álgebra producto. Los elementos de Ω son de la forma $\omega = (x_0, a_0, x_1, a_1, \dots)$ con $x_t \in X$ y $a_t \in A$ para toda $t = 0, 1, \dots$, con las proyecciones x_t y a_t de Ω sobre X y A , llamadas estado y acción, respectivamente.

Sea $\pi = \{\pi_t\}$ una política de control y ν una medida de probabilidad en X . Por el teorema de Ionescu-Tulcea (A.1.7) en el Anexo A, existe una única medida de probabilidad P_ν^π en (Ω, \mathcal{F}) tal que, para cada $B \in \mathcal{B}(X)$, $C \in \mathcal{B}(A)$ y $h_t \in \mathbb{H}_t$ se tiene:

$$\begin{aligned} P_\nu^\pi(x_0 \in B) &= \nu(B), \\ P_\nu^\pi(a_t \in C | h_t) &= \pi(C | h_t), \\ P_\nu^\pi(x_{t+1} \in B | h_t, a_t) &= Q(B | x_t, a_t). \end{aligned}$$

El proceso estocástico $(\Omega, \mathcal{F}, P_\nu^\pi, \{x_t\})$ es llamada **Proceso de Decisión de Markov a tiempo discreto**.

Observación 1.3.1 Cuando ν está concentrada en el estado $x_0 = x \in X$, se utiliza la siguiente notación:

$$\begin{aligned} P_x^\pi &:= P_\nu^\pi \\ E_x^\pi &:= E_\nu^\pi. \end{aligned}$$

1.4. Criterio de rendimiento

Cada PDM está dotado de una función real llamada criterio de rendimiento, la cual mide en algún sentido la calidad de cada política, a través de la sucesión de costos que se generan. A continuación se define el criterio de rendimiento que se utilizará en este trabajo. Para esto, considere un MCM estacionario a tiempo discreto y un conjunto de políticas Π .

Definición 1.4.1 Se define para cada $x \in X$ y $\pi \in \Pi$ el **criterio de costo total α -descuento** por

$$v_{\alpha, N}(\pi, x) := E_x^\pi \left[\sum_{t=0}^{N-1} \alpha^t C(x_t, a_t) \right], \quad (1.2)$$

en donde $\alpha \in (0, 1)$ es llamado factor de descuento y N es un entero positivo conocido, llamado horizonte del problema de optimización. Cuando $N = \infty$, se tiene el criterio de costo total α -descuento con horizonte infinito y se denota el criterio total α descuento con horizonte infinito como v_α .

Definición 1.4.2 La función de valor óptimo α -descuento con horizonte finito se define como:

$$V_{\alpha, N}^*(x) := \inf_{\pi \in \Pi} v_{\alpha, N}(\pi, x), \quad \forall x \in X. \quad (1.3)$$

El problema de control óptimo α -descuento con horizonte N , consiste en encontrar una política $\pi_N^* \in \Pi$ tal que

$$V_{\alpha,N}^*(x) = v_{\alpha,N}(\pi_N^*, x), \quad \forall x \in X. \quad (1.4)$$

La política $\pi_N^* \in \Pi$ es llamada política óptima.

Ahora bien, se define la función de valor óptimo α descuento con horizonte infinito como:

$$V_\alpha^*(x) := \inf_{\pi \in \Pi} v_\alpha(\pi, x), \quad \forall x \in X. \quad (1.5)$$

El problema de control óptimo α -descuento con horizonte infinito consiste en encontrar una política $\pi_* \in \Pi$ tal que

$$V_\alpha^*(x) := v_\alpha(\pi_*, x), \quad \forall x \in X. \quad (1.6)$$

1.5. Programación dinámica

En esta sección se presenta una herramienta esencial para resolver el problema de control óptimo, conocida en la literatura como Programación dinámica [2]. Este procedimiento permite determinar la función de valor óptimo para horizonte finito e infinito, así como la política óptima. Bajo condiciones adecuadas sobre la función de costo en un paso y la ley de transición se caracterizan las funciones de valores óptimos α descuento V_α^* y $V_{\alpha,N}^*$ mediante una ecuación funcional. Más aún, el conocimiento de V_α^* y $V_{\alpha,N}^*$ permite obtener una política óptima determinista markoviana estacionaria. La función objetivo a considerar es el criterio de rendimiento α -descuento con horizonte $N \in \mathbb{N} \cup \{\infty\}$ con estado inicial del proceso fijo $x_0 = x \in X$.

Primero presentamos una breve introducción con horizonte finito del problema.

1.6. Horizonte finito

La idea central de Programación Dinámica consiste en dividir el problema de optimizar el costo desde el tiempo 0 hasta N , en subproblemas más pequeños, en los que se requiere optimizar el costo en un número de periodos menor a N . De manera intuitiva, los subproblemas son más sencillos de resolver en la medida de que involucra menos intervenciones por parte del controlador, que el problema original.

Considere como criterio de rendimiento al costo total α -descuento $v_{\alpha,N}$ con horizonte finito $N \in \mathbb{N}$.

Definición 1.6.1 Sean v_0, v_1, \dots, v_N funciones definidas en X como:

$$v_0(x) := 0$$

y para cada $t \in \{1, 2, \dots, N\}$,

$$v_t(x) := \min_{a \in A(x)} \left\{ C(x, a) + \alpha \int_X v_{t-1}(y) Q(dy|x, a) \right\}, \quad (1.7)$$

en donde cada función v_t es medible.

La ecuación funcional dada en (1.7) se conoce como **Ecuación de Programación Dinámica (EPD)**, junto con su condición inicial $v_0(x) = 0$ y su nombre es debido a Bellman [2].

Una consecuencia interesante que determina el algoritmo es que la política óptima encontrada resulta ser determinista markoviana. De esta manera se deduce que para seleccionar una acción, a_t en el tiempo t sólo es necesario conocer al estado actual x_t , para lograr todo esto, es necesaria la suposición de existencia de selectores $f \in \mathbb{F}$, los cuales minimizan la ecuación de programación dinámica en cada etapa. La suposición presentada a continuación es conocida como **condición de selección medible**.

Suposición 1.6.2 Supongamos para cada $t = 1, 2, \dots, N$, existe un selector $f_t \in \mathbb{F}$ tal que $f_t(x) \in A(x)$ alcanza el mínimo en (1.7) para cada $x \in X$, es decir,

$$v_t(x) = C(x, f_t(x)) + \alpha \int_X v_{t-1}(y) Q(dy|x, f_t(x)), \quad \forall x \in X \quad (1.8)$$

Note que la sucesión $\{f_1, f_2, \dots, f_{N-1}\}$ genera una política determinista markoviana.

Para poder garantizar la condición de selección medible se buscan condiciones sobre el modelo, las cuales presentamos a continuación. Antes, es necesario presentar los siguientes conceptos.

Definición 1.6.3 Sean X el espacio de estados, (X, d) un espacio métrico y $v : X \rightarrow \mathbb{R} \cup \{\infty\}$ una función medible tal que $v(x) < \infty$ para algún $x \in X$. La función v se dice ser **inferiormente semicontinua** (l.s.c., por sus siglas en inglés) en $x \in X$, si

$$\liminf_{n \rightarrow \infty} v(x_n) \geq v(x),$$

para cualquier sucesión $\{x_n\}$ en X convergente a $x \in X$.

Si v es l.s.c. para cada $x \in X$, se llama **inferiormente semicontinua**.

Definición 1.6.4 Una función $v : \mathbb{K} \rightarrow \mathbb{R}$ se llama *inf-compacta* sobre \mathbb{K} , si el conjunto $\{a \in A(x) | v(x, a) \leq \lambda\}$ es compacto para cada $\lambda \in \mathbb{R}$

Definición 1.6.5 Un kernel estocástico $Q \in P(X|A)$ (ver Definición A.1.3 en el Anexo) es fuertemente continuo, si la función

$$w(a) = \int_X v(y)Q(dy|x, a), \quad (1.9)$$

es continua y acotada en A para cada v , función continua y acotada sobre X y para cada $x \in X$.

Suposición 1.6.6 Las siguientes condiciones se cumplen:

1. La función de costo en un paso C es acotada inferiormente e inf-compacta sobre \mathbb{K} .
2. El kernel estocástico Q es fuertemente continuo.

A continuación se presenta un lema, que demuestra la relación entre la Suposición 1.6.2 y la Suposición 1.6.6.

Lema 1.6.7 La Suposición 1.6.2 implica la Suposición 1.6.6.

Demostración. Ver [12]. ■

Teorema 1.6.8 (Teorema de Programación Dinámica)

Suponga que la Suposición 1.6.2 se satisface entonces la política determinista markoviana $\pi^* = \{f_0, \dots, f_{N-1}\}$ es óptima y se tiene que

$$V_{\alpha, N}^*(x) = v_{\alpha, N}(\pi^*, x) = v_N(x), \quad \forall x \in X. \quad (1.10)$$

Demostración. Ver [12] ■

El teorema anterior muestra que v_t es el óptimo del subproblema desde el tiempo 0 al tiempo t , es decir,

$$v_t(x) = \inf_{\pi \in \Pi} v_{\alpha, t}(\pi, x),$$

para cada $x \in X$ y $t = 0, 1, \dots, N$.

Con la finalidad de garantizar la existencia de selectores medibles, se pueden consultar en [12] y [13] condiciones sobre las componentes del modelo de control, las cuáles garantizan la Suposición 1.6.2. Cabe destacar que una consecuencia interesante que determina el algoritmo, es que la política óptima es determinista markoviana. Entonces para seleccionar una acción a_t en el tiempo t solo es necesario conocer al estado actual x_t , es decir, no es indispensable emplear mecanismos aleatorios para seleccionar las acciones.

1.7. Horizonte infinito

Esta sección trabaja el problema de control óptimo cuando el horizonte del problema es infinito. El criterio de rendimiento a considerar es el costo total α descontado, denotado por v_α . Para poder resolver el problema, se aproxima a V_α^* con la funciones de valor óptimo $V_{\alpha,n}^*$. Uno de los resultados principales en esta sección es la convergencia puntual de la sucesión de funciones $\{V_{\alpha,n}^*\}$ a V_α^* .

Definición 1.7.1 *Las funciones de iteraciones de valores óptimos se definen de la forma siguiente, para cada $x \in X$ y $n = 1, 2, \dots$,*

$$V_n(x) = \inf_{a \in A(x)} \left\{ C(x, a) + \alpha \int_X V_{n-1}(y) Q(dy|x, a) \right\}, \quad (1.11)$$

con $V_0(x) = 0$.

Suponga que la función de costo C es no negativa, entonces por el Teorema de Convergencia Dominada [1], se tiene que

$$\lim_{n \rightarrow \infty} v_{\alpha,n}(\pi, x) = v_\alpha(\pi, x). \quad (1.12)$$

Probaremos que la función de valor óptimo α descontado V_α^* , es solución de la ecuación de programación dinámica, es decir,

$$V_\alpha^*(x) = \min_{A(x)} \left\{ C(x, a) + \alpha \int_X V_\alpha^*(y) Q(dy|x, a) \right\}, \quad \forall x \in X, \quad (1.13)$$

y además que

$$V_\alpha^*(x) = \lim_{n \rightarrow \infty} V_n(x), \quad \forall x \in X.$$

Por definición de cada iteración y lo visto en la sección anterior V_n es el valor óptimo α descontado con horizonte $n \in \mathbb{N}$, así $V_n = V_{\alpha,n}^*$. De esta forma

$$V_\alpha^*(x) = \lim_{n \rightarrow \infty} V_{\alpha,n}^*(x) = \lim_{n \rightarrow \infty} \inf_{\pi \in X} v_{\alpha,n}(\pi, x), \quad \forall x \in X. \quad (1.14)$$

Si logramos justificar el intercambio entre límite e ínfimo, al aplicar (1.12) tendríamos las igualdades en (1.14). Para esto, se requieren las siguientes suposiciones al modelo.

Suposición 1.7.2 *1. La función de costo es no negativa, l.s.c. e inf-compacta sobre \mathbb{K} .*

2. La ley de transición Q es fuertemente continua.

Para proseguir con los resultados, es importante la siguiente suposición.

Suposición 1.7.3 *Existe una política π tal que $v_\alpha(\pi, x) < \infty$ para cada $x \in X$.*

Se denota por Π_0 al conjunto de todas las políticas que cumplen la Suposición 1.7.3.

El siguiente lema da como resultado el intercambio de ínfimo y límite bajo condiciones sobre la sucesión de funciones.

Lema 1.7.4 *Sean u y u_n , $n = 1, 2, \dots$ funciones l.s.c., acotadas inferiormente e inf compactas en \mathbb{K} . Si $u_n \uparrow u$ entonces,*

$$\lim_{n \rightarrow \infty} \min_{a \in A(x)} u_n(x, a) = \min_{a \in A(x)} u(x, a), \quad \forall x \in X.$$

Demostración. Para cada $x \in X$, se define

$$l(x) := \lim_n \min_{a \in A(x)} u_n(x, a) \quad \text{y} \quad u^*(x) := \min_{a \in A(x)} u(x, a).$$

Como $u_n \uparrow u$, se tiene que $l(x) \leq u^*(x)$. Para probar la otra desigualdad, sea $x \in X$ fijo y considere los siguientes conjuntos para cada $n \in \mathbb{N}$:

$$\begin{aligned} A_n &:= \{a \in A(x) \mid u_n(x, a) \leq u^*(x)\} \quad \text{y} \\ A_0 &:= \{a \in A(x) \mid u(x, a) = u^*(x)\}. \end{aligned}$$

Observe que cada uno de estos conjuntos es no vacío debido a la definición de u^* y la convergencia de u_n . Por otro lado son compactos por que u_n son inf – compactos. También, los conjuntos A_n decrecen a un conjunto compacto, ya que si $a \in A_{n+1}$ entonces $u_{n+1}(x, a) \leq u^*(x)$ pero $u_n(x, a) \leq u_{n+1}(x, a)$, de esto que $a \in A_n$, así $A_n \downarrow A_0$.

Para cada $n \geq 1$, sea $a_n \in A_n$ tal que $u_n(x, a_n) = \min_{a \in A(x)} u_n(x, a)$, se garantiza la existencia de a_n por la Proposición A.1.5. Utilizando que $A_n \downarrow A_0$ y que son conjuntos compactos, existen $a_0 \in A_0$ y una subsucesión $\{a_{n_i}\} \subset \{a_n\}$ tal que $a_{n_i} \rightarrow a_0$.

Así

$$u_{n_i}(x, a_{n_i}) \geq u_n(x, a_{n_i}), \quad \forall n_i > n,$$

haciendo tender $n \rightarrow \infty$

$$\begin{aligned} &\lim_{n \rightarrow \infty} u_{n_i}(x, a_{n_i}) \geq u_n(x, a_{n_i}) \\ \Rightarrow &\lim_{n \rightarrow \infty} \min_{a \in A(x)} u_n(x, a_{n_i}) \geq u_n(x, a_0) \\ &\Rightarrow l(x) \geq u_n(x, a_0), \end{aligned}$$

y como $u_n \uparrow u$, se tiene que $l(x) \geq u(x, a_0) = u^*(x)$ y esto sucede para cualquier $x \in X$, de esta forma queda probado el lema. ■

Definición 1.7.5 Se denota al espacio de las funciones medibles no negativas en X como $M(X)^+$ y para cada $u \in M(X)^+$ se define el siguiente operador sobre el espacio $M(X)^+$

$$Tu(x) := \min_{a \in A(x)} \left\{ C(x, a) + \alpha \int_X u(y)Q(dy|x, a) \right\}, \quad \forall x \in X.$$

Lema 1.7.6 Suponga que se cumple la Suposición 1.7.2 entonces $Tu \in M(X)^+$ para cada $u \in M(X)^+$, más aún existe un selector $f \in \mathbb{F}$ tal que

$$Tu(x) = C(x, f) + \alpha \int_X u(y)Q(dy|x, f), \quad \forall x \in X.$$

Demostración. Ver [2]. ■

Observe que usando el operador T se reescribe la ecuación (1.11) y (1.17) de la forma siguiente

$$V_\alpha^* = TV_\alpha^* \quad y \quad V_n = TV_{n-1}, \quad \forall n \in \mathbb{N}.$$

y $V_0 = 0$.

Lema 1.7.7 Suponga que se cumplen las Suposiciones 1.7.2 y 1.7.3 entonces

a) Si $u \in M(X)^+$ con $u \geq Tu$ entonces $u \geq V_\alpha^*$.

b) Sea $u : X \rightarrow \mathbb{R}$ una función medible tal que Tu está bien definida, con $u \leq Tu$ y

$$\lim_{n \rightarrow \infty} \alpha^n E[u(x_n)] = 0, \quad \forall x \in X, \quad \pi \in \Pi_0,$$

entonces $u \leq V_\alpha^*$.

Demostración.

a) Sea $u \in M(X)^+$ con $u \geq Tu$ y considere $f \in \mathbb{F}$ tal que

$$u(x) \geq C(x, f) + \alpha \int_X u(y)Q(dy|x, f), \quad \forall x \in X, \quad (1.15)$$

iterando esta desigualdad, se tiene que

$$\begin{aligned}
u(x) &\geq C(x, f) + \alpha \int_X \left(C(x, f) + \alpha \int_X u(z)Q(dz|y, f) \right) Q(dy|x, f) \\
&= C(x, f) + \alpha \int_X C(y, f)Q(dy|x, f) \\
&\quad + \alpha^2 \int_X \int_X u(z)Q(dz|y, f)Q(dy|x, f) \\
&= E_x^f \left[\sum_{t=0}^1 \alpha^t C(x_t, f) \right] + \alpha^n E_x^f [u(x_2)],
\end{aligned}$$

procediendo de manera inductiva, se obtiene que

$$u(x) \geq E_x^f \left[\sum_{t=0}^{n-1} \alpha^t C(x_t, f) \right] + \alpha^n E_x^f [u(x_n)],$$

en donde

$$E_x^f [u(x_n)] = \int_X u(y)Q^n(dy|x, f),$$

y Q^n denota el n -ésimo paso del Proceso de Markov usando la estrategia f . Como u es no negativa se tiene la siguiente relación

$$u(x) \geq E_x^f \left[\sum_{t=0}^{n-1} \alpha^t C(x_t, f(x_t)) \right],$$

para cada $n \in \mathbb{N}$ y $x \in X$, haciendo tender $n \rightarrow \infty$

$$u(x) \geq v_\alpha(f, x) \geq V^*(x), \quad \forall x \in X.$$

Así, se ha demostrado el inciso a) del lema .

b) Sean $\pi \in \Pi$, $x \in X$ y considere que $Tu \geq u$, entonces por la propiedad de Markov se prosigue de la manera siguiente:

$$\begin{aligned}
E_x^\pi [\alpha^{t+1}u(x_{t+1})|h_t, a_t] &= \alpha^{t+1} \int_X u(y)Q(dy|x_t, a_t) \\
&= \alpha^t \left(C(x_t, a_t) + \alpha \int_X u(y)Q(dy|x_t, a_t) \right) \\
&\quad - \alpha^t C(x_t, a_t) \\
&\geq \alpha^t (u(x_t) - C(x_t, a_t)).
\end{aligned}$$

Así,

$$\begin{aligned}
\alpha^t C(x_t, a_t) &\geq -E_x^\pi [\alpha^{t+1}u(x_{t+1}) - \alpha u(x_t)|h_t, a_t] \\
&= E_x^\pi [\alpha^t u(x_t) - \alpha^{t+1}u(x_{t+1})|h_t, a_t].
\end{aligned}$$

Tomando esperanzas y sumando desde $t = 0$ a $n - 1$, se obtiene que

$$\begin{aligned} E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t C(x_t, a_t) \right] &\geq E_x^\pi \left[\sum_{t=0}^{n-1} \alpha^t u(x_t) - \alpha^{t+1} u(x_{t+1}) \middle| h_t, a_t \right] \\ &= E_x^\pi [u(x_0) - \alpha^n u(x_n)] \\ &= u(x) - \alpha^n E_x^\pi [u(x_n)], \end{aligned}$$

haciendo $n \rightarrow \infty$ y considerando las hipótesis de este inciso, se tiene que

$$E_x^\pi \left[\sum_{t=0}^{\infty} \alpha^t C(x_t, a_t) \right] \geq u(x),$$

y esto fue para x y π arbitrarios, de esto que

$$V_\alpha^* \geq u.$$

■

Lema 1.7.8 *Suponga que se cumplen las Suposiciones 1.7.2 y 1.7.3, entonces $V_n \uparrow V_\alpha^*$ y V_α^* satisface la ecuación de programación dinámica (1.13).*

Demostración. Para cada $x \in X$, $\pi \in \Pi$ y $n = 0, 1, \dots$ se tiene que

$$V_n(x) \leq v_{\alpha, n}(\pi, x) \leq v_\alpha(\pi, x).$$

Por lo tanto,

$$V_n(x) \leq V^*(x). \quad (1.16)$$

Claramente si $u_1, u_2 \in M(X)^+$ con $u_1 \geq u_2$ entonces $Tu_1 \geq Tu_2$. De esta forma T es monótono. Como $V_0 = 0$ y $V_n = TV_{n-1}$ para cada $n \geq 1$, las funciones de iteración de valores forman una sucesión de funciones no decreciente en $M(X)^+$, lo cual implica que $V_n \uparrow V_\alpha^*$, para alguna $V_\alpha^* \in M(X)^+$.

Considere

$$\begin{aligned} u_n(x, a) &:= C(x, a) + \alpha \int_X V_n(y) Q(dy|x, a), \\ u(x, a) &:= C(x, a) + \alpha \int_X V_\alpha^*(y) Q(dy|x, a), \end{aligned}$$

para cada $(x, a) \in \mathbb{K}$. Luego, por el Teorema de Convergencia Monótona [1], se tiene que $u_n \uparrow u$.

Por otro lado, las funciones u y u_n para cada $n \in \mathbb{N}_0$ son *l.s.c.* e *inf*–*compactas* en \mathbb{K} , así por el Lema 1.7.6, se tiene que

$$V^* = \lim_{n \rightarrow \infty} V_n = \lim_{n \rightarrow \infty} TV_{n-1} = TV^*,$$

es decir, V^* es solución de la ecuación de programación dinámica. Falta ver que $V^* = V_\alpha^*$. Por el Lema 1.7.6, $V^* \geq V_\alpha^*$ y la otra desigualdad se sigue por (1.16). ■

El siguiente teorema es el resultado principal de esta sección.

Teorema 1.7.9 *Suponga que se cumplen las Suposiciones 1.7.2 y 1.7.3 entonces*

- a) *La función de valor α descontado es la solución mínima de la ecuación de optimalidad de costo α descontado, es decir,*

$$V_\alpha^*(x) = \min_{a \in A(x)} \left\{ C(x, a) + \alpha \int_X V_\alpha^*(y) Q(dy|x, a) \right\}, \quad (1.17)$$

y si u es otra solución de la ecuación de optimalidad, entonces $u(x) \geq V_\alpha^(x)$ para cada $x \in X$.*

- b) *Existe un selector $f_* \in \mathbb{F}$ tal que $f_* \in A(x)$ alcanza el mínimo de (1.17), es decir,*

$$V_\alpha^*(x) = C(x, f_*) + \alpha \int_X V_\alpha^*(y) Q(dy|x, f_*), \quad (1.18)$$

para cada $x \in X$ y la política determinista estacionaria f_ es óptima; de manera recíproca si $f_* \in \Pi_{DS}$ es óptima, entonces satisface (1.18).*

- c) *Si π^* es una política tal que $v_\alpha(\pi^*, \cdot)$ es solución de la ecuación de optimalidad y satisface*

$$\lim_{n \rightarrow \infty} \alpha^n E_x^\pi [v_\alpha(\pi^*, x)] = 0, \quad \forall x \in X,$$

entonces, $v_\alpha(\pi^, \cdot) = V^*(\cdot)$, así π^* es óptima si y sólo si $v_\alpha(\pi^*, \cdot)$ satisface la ecuación de optimalidad de costo α –descontado.*

- d) *Si existe una política óptima α –descontado entonces existe una que es determinista estacionaria.*

Demostración. a) Del Lema 1.7.8, V_α^* es una solución de la ecuación de programación dinámica de costo descontado y por el Lema 1.7.6, V_α^* es la mínima solución de la misma ecuación. De esta manera $u \geq Tu$, entonces $u \geq V_\alpha^*$

b) La existencia de un selector $f_* \in \mathbb{F}$ se asegura por el Lema 1.7.6. Para cada $x \in X$ y $n \geq 1$ se tiene

$$\begin{aligned} V_\alpha^*(x) &= E_x^{f_*} \left[\sum_{t=0}^{\infty} \alpha^t C(x_t, f_*(x_t)) \right] \\ &= E_x^{f_*} \left[C(x_0, f_*) + \sum_{t=1}^{\infty} \alpha^t C(x_t, f_*(x_t)) \right] \\ &= C(x, f_*(x)) + E_x^{f_*} \left[\sum_{t=1}^{\infty} \alpha^t C(x_t, f_*(x_t)) \right]. \end{aligned}$$

Usando la propiedad de Markov podemos reescribir de la manera siguiente

$$\begin{aligned} E_x^{f_*} \left[\sum_{t=0}^{\infty} \alpha^t C(x_t, f_*(x_t)) \right] &= \int_X E_x^{f_*} \left[\sum_{t=1}^{\infty} \alpha^t C(x_t, f_*(x_t)) \right] Q(dy|x, f_*) \\ &= \int_X v_\alpha(f_*, y) Q(dy|x, f_*). \end{aligned}$$

Particularmente, si f_* es óptima entonces $V^*(x) = v_\alpha(f_*, x)$ para cada $x \in X$.

c) Si $v_\alpha(\pi^*, x)$ satisface la ecuación de programación dinámica y por Lema 1.7.7 se tiene que

$$v_\alpha(\pi^*) = V_\alpha^*(\cdot).$$

d) Finalmente, podemos concluir que si una política óptima existe, entonces existe una política determinista estacionaria, ver [12]. ■

Capítulo 2

Modelos de inventarios

El control y mantenimiento de los inventarios de bienes físicos es un problema común a todas las empresas en cualquier sector de una economía determinada. No obstante las tiendas minoristas son las que deben administrar sus inventarios. El mantenimiento de los inventarios es necesario para cualquier empresa que se ocupe de productos físicos, incluidos fabricantes, mayoristas y minoristas. Por ejemplo, los fabricantes necesitan inventarios de los materiales necesarios para fabricar sus productos, ellos también necesitan inventarios de los productos terminados en espera de envío, del mismo modo, ambos mayoristas y minoristas necesitan mantener los inventarios de los productos para estar disponibles para la compra de los clientes.

Los costos anuales asociados con el almacenamiento del inventario son muy grandes, tal vez hasta una cuarta parte del valor del inventario. Por lo tanto, los costos que se incurren para el almacenamiento de inventario en México se encuentran con los cientos de miles de millones de pesos anualmente. Reducir los costos de almacenamiento evitando inventarios innecesariamente grandes puede mejorar la competitividad de cualquier empresa.

Muchas compañías en otras partes del mundo también han estado modernizando la forma en que administran sus inventarios. La aplicación de técnicas de investigación de operaciones en esta área proporciona una herramienta poderosa para obtener una ventaja competitiva. La investigación de operaciones comprende de los siguientes pasos:

- Formular un modelo matemático que describa el comportamiento del sistema de inventario.
- Buscar una política de inventario con respecto a este modelo.
- Usar un sistema de procesamiento de información computari-

zado para mantener un registro recurrente de los niveles de inventario.

- Con este registro de niveles de inventario actuales, aplicar la política de inventario óptima para señalar cuándo y cuánto reponer de inventario.

Los modelos matemáticos de inventarios utilizados con este enfoque se pueden dividir en dos amplias categorías: modelos deterministas y modelos estocásticos, de acuerdo con la predictibilidad de la demanda involucrada. La demanda de un producto en inventario es la cantidad de unidades que deberán retirarse del inventario para algún uso durante un período específico. Si la demanda en periodos futuros puede pronosticarse con considerable precisión, es razonable usar una política de inventario que asuma que todas las ordenes siempre son exactas. Este es el caso de la demanda conocida donde un modelo determinista sería utilizado. Sin embargo, cuando la demanda no es conocida, es necesario utilizar un modelo de inventario estocástico en donde la demanda en cualquier período es una variable aleatoria en lugar de una constante conocida. En este trabajo estamos interesados en el estudio de modelos estocásticos.

2.1. Modelos matemáticos

Los modelos matemáticos estocásticos más importantes que aproximan el estudio de sistemas de inventarios se expresan por medio de las siguientes ecuaciones en diferencias.

Modelo Clásico

$$\boxed{x_{t+1} = x_t + a_t - \xi_{t+1}.} \quad (2.1)$$

Modelo de Lindley

$$\boxed{x_{t+1} = (x_t + a_t - \xi_{t+1})^+.} \quad (2.2)$$

Modelo de Lindley con una Generalización

$$\boxed{x_{t+1} = (x_t + \eta_t a_t - \xi_{t+1})^+.} \quad (2.3)$$

En dichos modelos se estudia el problema a tiempo discreto y las variables representan lo siguiente: en cada etapa $t = 0, 1, \dots$, x_t es

el nivel del almacén de un producto, a_t la producción u orden, ξ_{t+1} la demanda al finalizar el periodo y η_t es una variable del tipo Bernoulli, la cual modela el arribo de producto ordenado o producido en el almacén.

El modelo (2.1) es de los más usuales que se utilizan para el estudio de sistemas de inventarios ([8], [15], [21], [22], [24], [29]). Por otro lado, el modelo (2.2) tiene su aparición por primera vez en [17], propuesto por David Lindley en 1959. En el contexto de sistemas de inventarios, dicho sistema se conoce por modelo con demanda no suplida ([10], [30]). Por último, el modelo (2.3) es una generalización de (2.2). Algunos trabajos que estudian este modelo son [6] y [11].

De acuerdo a la definición de cada modelo estocástico, cada $x_t \in \mathbb{R}$ para (2.1), mientras que en (2.2) y (2.3), $x_t \in \mathbb{R}^+$. Por otro lado, el espacio de acciones y acciones admisibles deben ser subconjuntos de \mathbb{R}^+ , ya que no es posible tener en el sistema producción u orden negativa.

Ahora bien, (2.1) y (2.2) pueden ser sustituidos por

$$x_{t+1} = J(x_t, a_t, \xi_{t+1}), \quad (2.4)$$

en donde J es una función continua. Luego, si Q es la ley de transición para el proceso estocástico (2.4) entonces para cada $B \in \mathcal{B}(\mathbb{R}^+)$

$$\begin{aligned} Q(B|x, a) &= P(x_{t+1} \in B|x_t, a_t) \\ &= P(J(x_t, a_t, \xi_{t+1}) \in B|x_t = x, a_t = a) \\ &= \int_{\mathbb{R}^+} I_B(J(x, a, s)) dF_{t+1}(s) \\ &= E[I_B J(x, a, \xi_{t+1})], \end{aligned}$$

en donde F_{t+1} es la función de distribución de la variable aleatoria ξ_{t+1} . Análogamente, podemos describir la ley de transición para (2.3).

Por otro lado, consideremos una función de costo denotada por C , la cual generalmente se encuentra definida en $\mathbb{R}^+ \times \mathbb{R}^+$, con rango en \mathbb{R}^+ .

De este modo hemos enunciado los elementos que conforman un PDM. Con lo cual se tiene la relación que tienen estos modelos con la teoría vista en el Capítulo 1.

En el capítulo anterior, se menciona el problema de control óptimo y las condiciones que se requieren para encontrar una política óptima. Una clase de políticas interesantes en el estudio de modelos de inventarios son frecuentemente conocidas como punto de reorden, política con cantidad de orden, o (s, S) . En la práctica ocurre lo

siguiente con esta política: Cuando el inventario tiene un nivel de productos por debajo de s , se ordena una cantidad de no más de S unidades para estar por arriba de s .

Formalmente la política queda definida de la siguiente manera. En la siguiente sección realizamos una revisión de los trabajos más importantes que se han hecho para demostrar que esta clase de políticas son óptimas para los modelos de inventarios que se han presentado en esta sección.

2.2. Antecedentes

Como se mencionó en la sección anterior, en esta sección presentamos un repaso de la literatura que estudia el control de inventarios con políticas (s, S) óptimas bajo los modelos estocásticos (2.1), (2.2) y (2.3). Este tipo de resultados tienen sus inicios en el año 1966, con un resultado importante presentado por Scarf [22].

En la siguiente tabla mostramos una recopilación de los trabajos, en donde clasificamos por referencia, un resumen breve del problema que se plantea, las suposiciones que utilizan y los resultados que se obtuvieron en cada uno de los trabajos de investigación, con el fin de poder diferenciar los resultados nuevos que presentamos en este trabajo de tesis.

Referencia	Resumen del problema	Suposiciones	Resultados
Scarf (1959) [22]	Este es el primer trabajo encontrado que considera el problema de caracterizar a la política óptima del tipo (s, S) con horizonte finito al modelo (2.1).	Se considera una función de costo compuesta por un costo de almacenaje y escasez las cuales son funciones convexas, además es agregado un costo fijo de ordenamiento K .	Se tiene que la política óptima es del tipo (s_n, S_n) , donde $n \in \mathbb{N}$ es el horizonte del problema.
Iglehart (1963) [15]	En base a los resultados presentados en [22], este trabajo estudia el comportamiento límite de las sucesiones $\{s_n\}$ y $\{S_n\}$.	Se tienen las mismas suposiciones del problema dadas en [22].	Se demuestra la existencia de los límites s y S de las sucesiones $\{s_n\}$ y $\{S_n\}$, respectivamente, en donde además tales límites forman una política óptima del tipo (s, S) , para el problema de control con horizonte infinito.
Veinnot y Wagner (1965) [28]	Se desarrolla un enfoque computacional completo para encontrar las políticas de inventarios óptimas (s, S) para el modelo (2.2).	Se utilizan costos no negativos, un factor de descuento $0 < \alpha \leq 1$ y un tiempo de entrega.	El método se deriva de la teoría de renovación y el análisis estacionario. Se establecen nuevos límites superior e inferior en los valores óptimos de s y S .
Veinott (1966) [27]	En este trabajo se dan nuevas condiciones de unimodalidad a la función de costo que no implican ni están implicadas las hipótesis dadas en [22] y [15] y se busca demostrar la optimalidad de las políticas (s, S) .	Se estudia el modelo con horizonte infinito. Los costos son unimodales y se supone que los mínimos absolutos de los costos esperados de un período están aumentando con el tiempo.	Se demuestra la optimalidad de la política del tipo (s, S) con estas nuevas condiciones que resultan ser más débiles que las presentadas en [22].

Referencia	Resumen de el problema	Suposiciones	Resultados
Porteus (1971) [21]	Se examina el modelo (2.1) a horizonte de tiempo finito con hipótesis sobre las funciones de costo más débiles.	Se utilizan funciones de costos no lineales y cóncavas. Se requiere una generalización de las funciones K -convexas y cuasi convexas para funciones K -cuasiconvexas. Además, se supone que las densidades de probabilidades de la demanda deben ser densidades de Pólya unilaterales.	Se muestra que una política generalizada (s, S) es óptima.
Ehrhardt (1979) [7]	Para el modelo (2.1) se presenta una nueva aproximación analítica para las políticas (s, S) .	Se utilizan costos no lineales y con un tiempo de espera de reaprovisionamiento fijo y retraso de la demanda no satisfecha.	Las fórmulas de aproximación se derivan utilizando los resultados existentes de la teoría de renovación asintótica para caracterizar el comportamiento de s y S .
Sheng (1991) [29]	Se busca un algoritmo para la compilación de las políticas óptimas (s, S) .	Se suponen funciones de costos no negativas y unimodales. Se tiene una demanda con un único periodo.	Se encuentra un algoritmo que determina a las políticas (s, S) a horizonte infinito y se aplica a una revisión periódica en los sistemas de inventario con revisión continua.

Referencia	Resumen de el problema	Suposiciones	Resultados
Iglehart (1995) [15]	En base a los resultados presentados en [22], este trabajo estudia el comportamiento límites de la sucesiones $\{s_n\}$ y $\{S_n\}$.	Se tienen las mismas suposiciones del problema dadas en [22].	Se demuestra la existencia de los límites s y S de las sucesiones $\{s_n\}$ y $\{S_n\}$, respectivamente, donde además tales límites forman una política óptima del tipo (s, S) , para el problema de control con horizonte infinito.
Suresh (1995) [24]	El trabajo estudia el modelo (2.1) bajo las mismas hipótesis que [22] pero con demandas no necesariamente independientes, entonces se demuestra la optimalidad de las políticas (s, S) .	Se estudia el problema con horizonte de tiempo finito e infinito, con funciones de costo semejantes a las utilizadas en [22] pero con demandas no necesariamente independientes pero si con un comportamiento markoviano estacional o periódico.	Se demuestra la optimalidad de las políticas del tipo (s, S) para este tipo de generalización.
Daduna (1999) [6]	Es el primer trabajo encontrado que estudia el modelo (2.2). Se busca la optimalidad de las políticas (s, S) para este modelo.	Se requieren funciones de costo no negativas y además la función de valor óptimo es acotada.	Se obtiene la optimalidad de las políticas tipo (s, S) , y además con un método diferente a lo que se ha trabajado anteriormente.

Referencia	Resumen de el problema	Suposiciones	Resultados
Özekici-Parlar (1999) [20]	Se consideran modelos de inventarios de horizonte infinito con proveedores poco confiables y se busca una política óptima de nivel de orden dependiente del medio ambiente.	Se supone que el proceso ambiental sigue una cadena de Markov. Las ecuaciones de flujo de existencias del sistema de inventarios están sujetas a fluctuaciones ambientales, se representan mediante un proceso estocástico bidimensional.	Se encuentra la política de nivel de orden dependiente del medio ambiente, la cual resulta ser óptima cuando el costo de orden es lineal. Además cuando hay un costo fijo de pedido, muestran que las políticas del tipo (s, S) son óptimas.
Chen y Simchi-Levi (2004) [5]	Se analiza un modelo de inventario con un solo producto. El objetivo es maximizar el costo total descontado con horizonte infinito o el costo promedio esperado y encontrar una política estacionaria óptima (s, S) .	Las demandas en diferentes periodos son independientes entre sí y sus distribuciones dependen del precio del producto.	Se demuestra la optimalidad de la política (s, S) y el precio del producto se determina en función de la posición del inventario al comienzo de cada periodo.
Guo y Zhu [11] (2006)	El criterio a minimizar es de costo promedio esperado y se estudia la existencia de una política óptima para el modelo (2.3).	La función de costo no es necesariamente acotada inferiormente y el espacio de acciones admisibles es compacto.	Se proporcionan dos desigualdades de optimalidad promedio de direcciones opuestas y se dan las condiciones para la existencia de soluciones. Luego, utilizando las dos desigualdades, se asegura la existencia de una política estacionaria óptima promedio bajo supuestos adicionales de continuidad-compacidad.

Referencia	Resumen de el problema	Suposiciones	Resultados
Feinberg [8] (2015)	Este trabajo retoma todo lo realizado en los anteriores artículos, estudia el modelo (2.2) y busca la optimalidad de las políticas tipo (s, S) pero se agregan nuevas condiciones y se debilitan las hipótesis que se utilizaron en los trabajos previos.	Se supone funciones de costo K -convexas y K -inf compactas, también son necesarios criterios de unimodalidad sobre estas funciones.	Se demuestra la optimalidad de las políticas (s, S) y además se demuestra que también son óptimas con el criterio de costo promedio.
Feinberg [9] (2016)	Con los resultados encontrados en [8], se busca dar algunas alternativas para demostrar la optimalidad de las políticas (s, S) , y además el modelo propuesto por Veinott [27] tiene una política óptima del mismo tipo.	Se tienen las mismas suposiciones dadas en [8].	Se tienen nuevos resultados y algunos teoremas auxiliares para poder demostrar la optimalidad de las políticas (s, S) en algún modelo que cumpla las condiciones que plantea este autor.

Mediante esta revisión bibliográfica del estudio de la optimalidad de políticas (s, S) en problemas de control de inventario, encontramos pocos resultados de optimalidad de este tipo de políticas en el modelo (2.3). Como se presentó en el resumen, Daduna [6] demuestra la optimalidad pero utiliza espacio de estados y acciones compactos. De esta manera, la aportación que realizamos en este trabajo de tesis es demostrar la optimalidad de las políticas (s, S) con la misma función de costo que se utiliza en [6] pero con espacio de estados y acciones no necesariamente compacto.

2.3. Modelo con demanda no suplida

En esta sección retomamos el modelo (2.3) para su estudio. Recordemos que la variable x_t representa el inventario al final del período t , a_t es la decisión de la cantidad de producto que se debe ordenar, supongamos que η_t es una variable aleatoria Bernoulli con parámetro p , donde $\{\eta_t = 1\}$ representa el evento de que el producto solicitado llegue al inventario con probabilidad p y $\{\eta_t = 0\}$, lo contrario, es decir, con probabilidad $1 - p$, el producto no llegó al inventario, y la variable ξ_{t+1} es la demanda al finalizar el periodo t .

El modelo tiene el siguiente escenario de toma de decisiones: se observa el sistema en cada etapa, o tiempo. Asumimos que existe una probabilidad positiva de que el producto en producción u ordenado llegue en el periodo con éxito, luego los productos están disponibles de inmediato para satisfacer la demanda. La demanda se genera pero si no es suficiente la capacidad del inventario para satisfacer está, se considera como no suplida, y el nivel de stock es cero.

Ahora bien, de acuerdo al modelo (2.3), el espacio de estados del proceso es $X = \mathbb{R}^+$ y el espacio de acciones o controles es $A = \mathbb{R}^+$. Consideremos el espacio de acciones admisibles como $A(x) = A$ para cada $x \in \mathbb{R}^+$. De acuerdo a la notación utilizada en el capítulo anterior, el conjunto de parejas estado-acción está dado por $\mathbb{K} = \mathbb{R}^+ \times \mathbb{R}^+$.

La ecuación en diferencias dada en (2.3), induce el kernel estocástico Q o ley de transición de la siguiente forma.

Para cada $B \in \mathcal{B}(\mathbb{R}^+)$ se tiene que

$$\begin{aligned} Q(B|x, a) &= P_x^\pi(x_{t+1} \in B | x_t = x, a_t = a) \\ &= P_x^\pi((x + \eta_t a - \xi_t)^+ \in B). \end{aligned}$$

Si $B = (-\infty, y]$ con $y \in \mathbb{R}$, entonces

$$\begin{aligned} Q((-\infty, y]|x, a) &= P_x^\pi((x + \eta_t a - \xi_t)^+ \leq y) \\ &= P_x^\pi(x + \eta_t a - \xi_t \leq 0, y \geq 0) \\ &\quad + P_x^\pi(0 \leq x + \eta_t a - \xi_t \leq y). \end{aligned}$$

Si $y < 0$, se sigue que

$$Q((-\infty, y]|x, a) = 0.$$

Si $y \geq 0$, entonces

$$\begin{aligned} Q((\infty, y]|x, a) &= P_x^\pi(x + \eta_t a - \xi_t \leq 0) + P_x^\pi(0 \leq x + \eta_t a - \xi_t \leq y) \\ &= pP_x^\pi(x + a \leq \xi_t) + pP_x^\pi(y \leq \xi_t \leq x + a) \\ &\quad + (1 - p)P_x^\pi(x \leq \xi_t) + (1 - p)P_x^\pi(y \leq \xi_t \leq x + a) \\ &= p(1 - F_t(x + a)) + (1 - p)(1 - F_t(x)) \\ &\quad + p(F_t(x + a) - F_t(y)) + (1 - p)(F_t(x) - F_t(y)), \end{aligned}$$

en donde la función F_t representa la función de distribución de probabilidad de la variable aleatoria ξ_t .

Suposición 2.3.1 *Supongamos lo siguiente:*

- *La sucesión de demandas, $\{\xi_t\}_{t \in \mathbb{N}}$ son variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) con función de distribución de probabilidad F y función de densidad continua Δ .*
- *Si ξ representa un elemento genérico de la sucesión $\{\xi_t\}_{t \in \mathbb{N}}$, entonces $\mu := E[\xi] < \infty$.*
- *La sucesión de variables $\{\eta_t\}_{t \in \mathbb{N}}$ son independientes e idénticamente distribuidas con distribución Bernoulli con parámetro p .*

Como se mencionó anteriormente, en cada etapa se generan costos. En este trabajo consideramos gastos de almacenaje, producción y demanda no suplida. Este costo se representa mediante la siguiente función definida sobre \mathbb{K} ,

$$C(x, a) = g(a) + H(x, a), \quad (2.5)$$

en donde la función $g : \mathbb{R} \rightarrow \mathbb{R}^+$, representan los costos de producción y la función $H : \mathbb{K} \rightarrow \mathbb{R}^+$ los costos de almacenaje y demanda no suplida.

Observe que existen una gran variedad de funciones de costo, por ejemplo, funciones lineales, polinomiales, exponencias etc., pero no

todas pueden modelar los costos que hemos involucrado. En [6], se presenta una propuesta de función de costo, definida de la siguiente manera:

$$g(a) := K1_{\{a \neq 0\}} + ca,$$

y

$$H(x, a) := hE[x + \eta a] + lE[(\xi - (x + \eta a))^+],$$

en donde

- K es el costo fijo por producción, también puede indicar un impuesto,
- c es el costo por unidad producida,
- h es el costo por unidad almacenada,
- l es el costo por unidad faltante.

En la Sección 2.2, mencionamos que las hipótesis utilizadas en [6] manejaron espacio de estados y acciones compactos. En este trabajo se estudia la optimalidad de las políticas (s, S) usando la misma función de costo dada en [6] y espacio de estados y acciones no necesariamente compacto.

Ahora procederemos a demostrar la existencia de una política óptima markoviana como primer resultado importante en esta tesis. Primero veamos que la Suposición 1.6.6 se cumple, utilizando el siguiente lema.

Lema 2.3.2 *Sea Q el kernel estocástico inducido por la dinámica (2.3) entonces Q es fuertemente continuo, es decir, para cada función u acotada y medible se tiene que la integral de medida siguiente,*

$$\int_X u(y)Q(dy|x, a)$$

es continua y acotada sobre el espacio de acciones A .

Demostración. Sea $u \in \mathbb{B}(X)$, entonces se define la siguiente función:

$$w(x, a) := \int_X u(y)Q(dy|x, a), \quad \forall (x, a) \in \mathbb{K}. \quad (2.6)$$

Para probar que el kernel Q es fuertemente continuo, debemos demostrar que w es continua en A .

Ahora bien, consideremos $x \in X$ fijo y $\{a_n\}$ una sucesión definida

en A . Supongamos que la sucesión converge a algún $a \in A$. De esta forma, procedemos a realizar el siguiente cálculo

$$\begin{aligned}
\lim_{n \rightarrow \infty} w(x, a_n) &= \lim_{n \rightarrow \infty} \int_X u(y)Q(dy|x, a_n) \\
&= \lim_{n \rightarrow \infty} E[u(x + a_n\eta - \xi)^+] \\
&= p \lim_{n \rightarrow \infty} \int_X u((x + a_n - s)^+)dF(s) \\
&\quad + (1 - p) \lim_{n \rightarrow \infty} \int_X u((x - s)^+)dF(s) \\
&= p \lim_{n \rightarrow \infty} u(0)F(x + a_n) \\
&\quad + p \lim_{n \rightarrow \infty} \int_0^{x+a_n} u(x + a_n - s)dF(s) \\
&\quad + (1 - p) \int_X u((x - s)^+)dF(s).
\end{aligned}$$

Utilizando la continuidad de la función de distribución F y que la función u es acotada, por el Teorema de Convergencia Dominada A.1.6, se tiene que

$$\begin{aligned}
\lim_{n \rightarrow \infty} w(x, a_n) &= pu(0)F(x + a) + p \int_0^{x+a} u(x + a - s)dF(s) \\
&\quad + (1 - p) \int_X u((x - s)^+)dF(s) \\
&= w(x, a).
\end{aligned}$$

Por lo tanto, w es una función continua en A y concluimos que el kernel Q es fuertemente continuo. ■

Lema 2.3.3 *La función de costo definida en (2.5) es inf-compacta y l.s.c. sobre \mathbb{K} .*

Demostración. Para demostrar que es l.s.c., sólo basta con demostrar que la función $H(x, a)$ es una función l.s.c., debido a que

$$K1_{\{a \neq 0\}}(a) + ca,$$

sí lo es, ya que es continua para $a > 0$ y l.s.c. para $a = 0$. Desarrollando la función H , considerando que la variable aleatoria η tiene distribución Bernoulli con parámetro p , entonces

$$\begin{aligned}
H(x, a) &= hp(x + a) + h(1 - p)x + lp \int_{x+a}^{\infty} (s - (x + a))dF(s) \\
&\quad + l(1 - p) \int_x^{\infty} (s - x)dF(s).
\end{aligned}$$

Sean $\{x_n\}$ y $\{a_n\}$ sucesiones convergentes con límite $x \in X$ y $a \in A$, respectivamente, entonces

$$\begin{aligned} \lim_{n \rightarrow \infty} H(x_n, a_n) &= \lim_{n \rightarrow \infty} (hp(x_n + a_n) + h(1-p)x_n) \\ &\quad + \lim_{n \rightarrow \infty} lp \int_{x_n + a_n}^{\infty} (s - (x_n + a_n)) dF(s) \\ &\quad + \lim_{n \rightarrow \infty} l(1-p) \int_{x_n}^{\infty} (s - x_n) dF(s). \end{aligned}$$

Luego,

$$\begin{aligned} \lim_{n \rightarrow \infty} (hp(x_n + a_n) + h(1-p)x_n) &= hp \lim_{n \rightarrow \infty} (x + a) + h(1-p) \lim_{n \rightarrow \infty} x_n \\ &= hp(x + a) + h(1-p)x, \end{aligned}$$

y

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left(lp \int_{x_n + a_n}^{\infty} (s - (x_n + a_n)) dF(s) + l(1-p) \int_{x_n}^{\infty} (s - x_n) dF(s) \right) \\ &= lp \lim_{n \rightarrow \infty} \int_0^{\infty} I_{[x_n + a_n, \infty)}(s - (x_n + a_n)) dF(s) \\ &\quad + l(1-p) \lim_{n \rightarrow \infty} \int_0^{\infty} I_{[x_n, \infty)}(s - x_n) dF(s), \end{aligned}$$

note que la función $I_{[x_n + a_n, \infty)}(s - (x_n + a_n)) \leq s$, y además se satisface la siguiente propiedad

$$\liminf[0, x_n] \subset \limsup[0, x_n] \subset [0, x],$$

se tiene que $I_{[0, x_n]}$ converge a $I_{[0, x]}$ casi seguramente, análogamente $I_{[0, x_n + a_n]}$ converge casi seguramente a $I_{[0, x + a]}$, así por el Teorema de Convergencia Dominada A.1.6,

$$\begin{aligned} &\lim_{n \rightarrow \infty} \left(lp \int_{x_n + a_n}^{\infty} (s - (x_n + a_n)) dF(s) + l(1-p) \int_{x_n}^{\infty} (s - x_n) dF(s) \right) \\ &= lp \int_{x+a}^{\infty} (s - (x + a)) dF(s) + l(1-p) \int_x^{\infty} (s - x) dF(s). \end{aligned}$$

Por lo tanto, H es continua para $a > 0$ y *l.s.c.* para $a = 0$.

Para probar que la función C es *inf*-compacta sobre \mathbb{K} , primero vamos a demostrar que la función $K1_{\{a \neq 0\}}(a) + ca$ sea *inf*-compacta sobre $A = \mathbb{R}^+$. En efecto, considere $\lambda \in \mathbb{R}$ y el conjunto de nivel asociado a $\lambda \geq K$

$$D_f(\lambda, A) = \{a \geq 0 \mid K1_{\{a \neq 0\}}(a) + ca \leq \lambda\} = \left[0, \frac{\lambda - K1_{\{a \neq 0\}}}{c} \right],$$

el cual es cerrado y acotado en \mathbb{R} , por tanto compacto. Si $\lambda < K$ se tiene que

$$D_f(\lambda, A) = \{0\},$$

un conjunto cerrado y acotado, así compacto. Por tanto, los conjuntos de nivel de la función $K1_{a \neq 0}(a) + ca$ son compactos, así la función es inf-compacta, luego por el Lema A.1.8 se tiene que la función de costo C es inf-compacta sobre \mathbb{K} ya que cada término de la función de costo C es positivo y *l.s.c.*. ■

Con la función de costo definida en (2.5) se tiene el siguiente resultado para la función de valor óptimo V_α^* .

Lema 2.3.4 $V_\alpha^*(x) < \infty$ para cada $x \in X$.

Demostración. Consideremos la política estacionaria de nunca ordenar:

$$f(x) = 0, \quad \forall x \in X.$$

Luego $V_\alpha^*(x) \leq v_\alpha(f, x)$, si probamos que $v_\alpha(f, x) < \infty$ para cada $x \in X$, se tiene el resultado buscado.

Denotemos a x_t^f a la dinámica (2.3) utilizando la política f . Sea $x_0^f = x \geq 0$, entonces

$$x_{t+1}^f = (x_t - \xi_t)^+.$$

Consideremos el siguiente proceso de renovación

$$N(x) := \sup\{t | S_t \leq x\}, \quad (2.7)$$

donde $S_0 = 0$ y $S_t = \sum_{j=1}^t \xi_j$. Observe que $E[N(x)] < \infty$ para cada $x \geq 0$. Luego

$$\begin{aligned} v_\alpha(f, x) &= E_x^f \left[\sum_{t=0}^{\infty} \alpha^t C(x_t^f, 0) \right] \\ &= hx + lE[(\xi - x)^+] \\ &+ E_x^f \left[\sum_{t=1}^{N(x)} \alpha^t (h(x - S_t) + E[(\xi_t - (x - S_t))^+]) \right] \\ &+ E_x^f \left[\sum_{t=N(x)+1}^{\infty} \alpha^t E[\xi_k] \right] \\ &\leq hx + l\mu + E \left[\sum_{t=1}^{\infty} \alpha^t (hx + \mu) \right] + \sum_{t=0}^{\infty} \alpha^t \mu \\ &= hx + l\mu + \frac{\alpha}{1-\alpha} (hx + \mu) + \frac{1}{1-\alpha} \mu < \infty. \end{aligned}$$

En la ecuación anterior se utilizó la suposición del modelo sobre la esperanza de la demanda finita. Así, V_α^* es finita. ■

Los lemas demostrados anteriormente garantizan que las Suposiciones 1.7.2 y 1.7.3 se cumplen para el modelo (2.3). Así se tiene el siguiente resultado.

Teorema 2.3.5 . *Existe una política óptima determinista estacionaria para el problema de control óptimo con horizonte finito e infinito bajo el modelo (2.3). Además las funciones de iteración de valor óptimo $V_n \uparrow V_\alpha^*$.*

Demostración. Por el Teorema 1.7.9 se tiene la existencia de la política óptima determinista estacionaria y por el Lema 1.7.8 se tiene que $V_n \uparrow V_\alpha^*$. ■

Capítulo 3

Optimalidad de la política (s, S)

En este capítulo se presentan los resultados principales de este trabajo de tesis. Mediante las funciones de iteración de valor óptimo se demuestra que la política óptima que resuelve el problema de control es de la forma (s, S) .

3.1. Análisis de las funciones de valor óptimo

En esta sección se busca caracterizar a la política óptima markoviana estacionaria garantizada en el Teorema 2.3.5. Para poder lograrlo, es necesario analizar el comportamiento de las funciones de iteración de valor óptimo (1.11), ya que es una herramienta útil que proporciona el método de programación dinámica para ir resolviendo el problema de control óptimo por etapas.

Sea $n \in \mathbb{N}$ y considerando la función de costo definida en (2.5), se tiene lo siguiente para cada estado inicial $x \geq 0$ y función de iteración V_n .

$$\begin{aligned} V_{n+1}(x) &= \inf_{a \geq 0} \{K 1_{\{a \neq 0\}}(a) + ca + hE[x + \eta a] + lE[(\xi - (x + \eta a))^+]\} \\ &\quad + \alpha E[V_n((x + \eta a - \xi)^+)] \\ &= \min\{K + G_n(x) + H_n(x), \inf_{a > 0} \{G_n(x + a) + H_n(x)\}\} \\ &= \min\{K + G_n(x), \inf_{a > 0} \{G_n(x + a)\}\} + H_n(x). \end{aligned}$$

donde

$$G_n(x+a) := (c+hp)(x+a) + \alpha p E[V_n((x+a-\xi)^+)] + p E[(\xi - (x+a))^+],$$

y

$$H_n(x) := (h(1-p) - c)x + \alpha(1-p) E[V_n((x-\xi)^+)] + (1-p) E[(\xi - x)^+].$$

Realizando el cambio de variable $y := x + a$, el problema de optimización se reestructura de la manera siguiente:

$$V_{n+1}(x) = \min\{K + G_n(x), \inf_{y>x} \{G_n(y)\}\} + H_n(x), \quad x \in X. \quad (3.1)$$

Para proceder con los siguientes resultados, recordemos la definición de una función convexa.

Definición 3.1.1 Sea $f : I \rightarrow \mathbb{R}$ donde I es un intervalo o es un subconjunto convexo de algún espacio vectorial. Se dice que f es convexa cuando verifica la siguiente condición:

$$a, b \in I, a < b \Rightarrow f((1-\lambda)a + \lambda b) \leq (1-\lambda)f(a) + \lambda f(b), \quad \forall \lambda \in [0, 1]. \quad (3.2)$$

A continuación presentamos el siguiente resultado sobre cada función V_n .

Lema 3.1.2 Para cada $n = 1, 2, \dots$, la función V_n cumple:

1. V_n es convexa sobre X ,
2. $\lim_{x \rightarrow \infty} V_n(x) = \infty$.

Demostración. Primero, veamos que G_0 es convexa. Por (3.1) y dado que $V_0 = 0$ se tiene que

$$G_0(y) = (c + hp)y + pE[(\xi - y)^+], \quad y \geq 0. \quad (3.3)$$

Luego, $(c + hp)y$ es convexa porque es una función lineal sobre la variable y , y por Lema 3.1.4 la función $E[(\xi - y)^+]$ es convexa. Así G_0 es convexa, además $G_0(y) \rightarrow \infty$ cuando $y \rightarrow \infty$, pues

$$0 \leq E[(\xi - y)^+] = \int_y^\infty sf(s)ds - y(1 - F(y)) \leq E[\xi I_{(y, \infty)}(\xi)],$$

entonces $E[(\xi - y)^+] \rightarrow 0$ cuando $y \rightarrow \infty$. Análogamente se observa que H_0 satisface las mismas condiciones que G_0 tomando en cuenta la hipótesis que $c \leq h(1 - p)$.

Ahora procedemos de manera inductiva sobre n , para $n = 1$,

$$V_1(x) = \min\{K + G_0(x), \inf_{y \geq x} \{G_0(y)\}\} + H_0(x),$$

por el Lema 3.1.4 la función $\inf_{y \geq x} \{G_0(y)\}$ es convexa y la función min es también convexa. Por otro lado,

$$\lim_{x \rightarrow \infty} V_1(x) = \lim_{x \rightarrow \infty} \inf_{y \geq x} \{G_0(y)\} + \lim_{x \rightarrow \infty} H_0(x),$$

para poder obtener el resultado, es necesario demostrar que

$$\lim_{x \rightarrow \infty} \min\{K + G_0(x), \inf_{y \geq x} \{G_0(y)\}\} = \infty,$$

para esto, tenemos que G_0 es convexa y $G_0(y) \rightarrow \infty$ cuando $y \rightarrow \infty$, así existe $y_0^* \geq x$ para $x \in X$ fijo, tal que

$$\inf_{y \geq x} \{G_0(y)\} = G_0(y_0^*),$$

y por el cambio de variable tenemos que existe $a_0^* \in A$ tal que

$$\inf_{a \geq 0} \{G_0(x + a)\} = G_0(x + a_0^*),$$

de esta forma, cuando $x \rightarrow \infty$

$$\inf_{a \geq 0} \{G_0(x + a)\} \rightarrow \infty.$$

Por lo tanto, se tiene lo buscado para V_1 .

Supongamos que para n se cumple que V_n es convexa y además

$$\lim_{x \rightarrow \infty} V_n(x) = \infty,$$

entonces G_n por definición y H_n son convexas. Por tanto, por Lema 3.1.4 se tiene que V_{n+1} es convexa. Por otro lado,

$$\begin{aligned} \lim_{x \rightarrow \infty} G_n(y) &= \lim(G_0(y) + \alpha p E[V_n((y - \xi)^+)]) = \infty, \\ \lim_{x \rightarrow \infty} H_n(x) &= \lim_{x \rightarrow \infty} (H_0(x) + \alpha(1 - p)E[V_n((x - \xi)^+)]) = \infty. \end{aligned}$$

Lo anterior, por la continuidad *l.s.c.* de la función V_n y por la hipótesis inductiva.

Así, existe $y_n^* \geq x$ tal que

$$\inf_{y \geq x} \{G_n(y)\} = G_n(y_n^*),$$

luego por el cambio de variable propuesto a y se tiene que existe $a_n^* \in A$ tal que $\inf_{y \geq x} \{G_n(x + a_n^*)\} = G_n(x + a_n^*)$. Por tanto,

$$\lim_{x \rightarrow \infty} V_{n+1}(x) = \infty. \quad (3.4)$$

■

Se demostró en el Lema 3.1.2 que para cada $t = 0, 1, \dots$ la función G_t y H_t son convexas y además, $G_t(y) \rightarrow \infty$ y $H_t(y) \rightarrow \infty$ cuando $y \rightarrow \infty$.

Por otro lado, en el capítulo anterior se demostró que la función V_n es la función de valor óptimo α descontada con horizonte n , donde $n = 1, 2, \dots$. Por la igualdad dada en (3.1), es importante conocer el comportamiento de las funciones H_n y G_n , para cada $n = 1, 2, \dots$, con el fin de poder intercambiar ínfimo por mínimo y encontrar el minimizador de la función.

El siguiente concepto es más general que la definición de función convexa.

Definición 3.1.3 Una función real g es K -convexa, con $K \geq 0$, si

$$K + g(z + y) \geq g(y) + z \left(\frac{g(y) - g(y - c)}{c} \right), \quad (3.5)$$

para todo $z \geq 0$, $c > 0$, $y \geq 0$.

El siguiente resultado se encuentra en el contexto de funciones K -convexas y resulta ser muy importante para demostrar la optimalidad de políticas (s, S) para el modelo.

Lema 3.1.4 (a) Una función convexa g es también 0-convexa, más aún es K -convexa para cada $K \geq 0$.

(b) Sean $g_1 : \mathbb{R}^+ \rightarrow \mathbb{R}$ función K -convexa y $g_2 : \mathbb{R}^+ \rightarrow \mathbb{R}$ función L -convexa entonces

$$\alpha g_1(y) + \beta g_2(y),$$

es $(\alpha K + \beta L)$ -convexa para cada $\alpha > 0$ y $\beta > 0$.

(c) Si $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ es K -convexa y w es una variable aleatoria con rango en \mathbb{R}^+ , entonces $E_w[g(y - w)]$ es también K -convexa siempre que $E_w[|g(y - w)|] < \infty$ para cada $y \geq 0$.

(d) Si $g : \mathbb{R}^+ \rightarrow \mathbb{R}$ es una función K -convexa y $g(y) \rightarrow \infty$ cuando $|y| \rightarrow \infty$, entonces existen escalares s y S con $s \leq S$ tal que

I) $g(S) \leq g(y)$ para cada $y \geq 0$,

II) $g(S) + K = g(s) < g(y)$ para cada $y \geq 0$,

III) $g(y)$ es una función decreciente en $(-\infty, s)$,

IV) $g(y) \leq g(z) + K$ para cada y, z con $s \leq y \leq z$.

Demostración. Para demostrar la parte (a) del lema, considere $\lambda \in [0, 1]$ y escribamos $x = (1 - \lambda)a + \lambda b = a + \lambda(b - a)$, así $x \in [a, b] \subset I$. Recíprocamente, para cada $x \in [a, b]$, basta tomar

$$\lambda = \frac{x - a}{b - a},$$

en donde $\lambda \in [0, 1]$ y $x = (1 - \lambda)a + \lambda b$. Por tanto, la ecuación (3.2) se expresa equivalentemente de la siguiente forma para cada $a, b \geq 0$ con $a < b$.

$$f(x) \leq \frac{b - x}{b - a}f(a) + \frac{x - a}{b - a}f(b), \quad \forall x \in [a, b]. \quad (3.6)$$

Si ahora, $a := y - c$, $b := y + z$ y $x := y$ y utilizando (3.6) se tiene que

$$g(y) \leq \frac{z}{z + c}g(y - c) + \frac{c}{z + c}g(y + z),$$

luego,

$$g(y)z + g(y)c \leq zg(y - c) + cg(y + z),$$

de esta forma

$$g(z + y) \geq g(y) + z \left(\frac{g(y) - g(y - c)}{c} \right),$$

por lo cual se tiene que g es una función 0-convexa. Además, $K \geq 0$ de esta forma g es K -convexa.

La parte (b) del lema se sigue de la ecuación (3.5) utilizando que $\alpha > 0$ y $\beta > 0$.

Para demostrar la parte (c), utilizamos que la función g es K -convexa y $E_w[g(y - w)] < \infty$ junto con las propiedades de la esperanza, E_w , se tiene

$$\begin{aligned} K + E_w[g(z + y - w)] &\geq E_w[g(y - w)] \\ &+ z \left(\frac{E_w[g(y - w)] - E_w[g(y - c - w)]}{c} \right), \end{aligned}$$

por lo tanto, la función $E_w[g(y - w)]$ es K -convexa. Por último procederemos a demostrar la parte (d) del lema. Utilizando por hipótesis que g es continua y $g(y) \rightarrow \infty$ cuando $|y| \rightarrow \infty$, entonces existe un punto mínimo de g . Sea S tal punto.

Ahora, sea s el escalar más pequeño tal que $z \leq S$ y $g(S) + K = g(z)$. Para cada $y \geq 0$ con $y < s$, se tiene de la ecuación (3.5)

$$K + g(S) \geq g(s) + \frac{S - s}{s - y}(g(s) - g(y)).$$

Debido a que $K + g(S) - g(s) = 0$, obtenemos que $g(s) - g(y) \leq 0$. Si $y < s$ y s es el escalar más pequeño para el cual $g(S) + K = g(s)$ se sigue que $g(s) < g(y)$ y II) se ha demostrado. Para demostrar III), observe que para $y_1 < y_2 < s$, se tiene que

$$K + g(S) \geq g(y_2) + \frac{S - y_2}{y_2 - y_1} (g(y_2) - g(y_1)).$$

Utilizando II),

$$g(y_2) > g(S) + K,$$

se sigue de las desigualdades anteriores

$$0 > \frac{S - y_2}{y_2 - y_1} (g(y_2) - g(y_1)),$$

con lo cual obtenemos que $g(y_1) > g(y_2)$, demostrando así la parte III). Para demostrar IV), note que $y = z$, $y = S$ o $y = s$. se satisface IV). Las otras dos posibilidades restantes son: $s < y < S$ y $S < y < z$. Si $S < y < z$, entonces por la K -convexidad

$$K + g(z) \geq g(y) + \frac{z - y}{y - S} (g(y) - g(S)) \geq g(y),$$

y por tanto, IV) queda demostrado. Si $s < y < S$, entonces por la K -convexidad

$$g(s) = K + g(S) \geq g(y) + \frac{S - y}{y - s} (g(y) - g(s)),$$

de lo cual

$$\left(1 + \frac{S - y}{y - s}\right) g(s) \geq \left(1 + \frac{S - y}{y - s}\right) g(y),$$

y $g(s) \geq g(y)$. Por lo tanto,

$$g(z) + K \geq g(S) + K = g(s),$$

se sigue que $g(z) + K \geq g(y)$. De esta manera IV) queda demostrado para cualquier caso. ■

Por lo tanto, se sigue con la siguiente proposición.

Proposición 3.1.5 *Para cada función G_n , donde $n = 0, 1, 2, \dots$, existe s_n y S_n tales que cumplen las siguientes condiciones:*

- I) $G_n(S_n) \leq G_n(y)$ para cada y ,
- II) $G_n(S_n) + K = G_n(s_n) < G_n(y)$ para cada y ,

III) $G_n(y)$ es una función decreciente en $(-\infty, s_n)$,

IV) $G_n(y) \leq G_n(z) + K$ para cada y, z con $s_n \leq y \leq z$.

Demostración. Por el Lema 3.1.2, cada función G_n es convexa y además $G_n(x) \rightarrow \infty$ cuando $x \rightarrow \infty$. Así por el Lema 3.1.4, se tiene la existencia de s_n y S_n que cumplen las condiciones I) a IV). ■

Por las condiciones anteriores dadas en la Proposición 3.1.5 y por el cambio de variable $y = x + a$, se tiene que la política (s_n, S_n) es óptima para el problema con criterio de costo α -descontado con horizonte de tiempo n .

Solo resta demostrar que la política óptima para el criterio de costo α -descontado con horizonte infinito es del tipo (s, S) . Primero, considere la función de valor óptimo α -descontado, V_α definido en el Capítulo 1. Por el Teorema 1.7.9, se tiene que la función V_α es la única solución de la ecuación de programación dinámica. Así, sustituyendo el costo definido en (2.5) en la ecuación de programación dinámica se tiene que

$$V_\alpha^*(x) = \min\{K + G(x), \inf_{x>0} \{G(x+a)\}\} + H(x),$$

en donde

$$G(x+a) = (c+hp)(x+a) + pE[(\xi - (x+a))^+] + \alpha pE[V_\alpha^*((x+a-\xi)^+)],$$

y

$$H(y) = (h(1-p) - c)x + (1-p)E[(\xi - x)^+] + \alpha(1-p)E[V_\alpha^*((x-\xi)^+)].$$

En el siguiente lema se tienen algunas consecuencias de la función V_α^* .

Lema 3.1.6 *La función de valor óptimo α -descontado, V_α^* es convexa y $V_\alpha^*(x) \rightarrow \infty$ cuando $x \rightarrow \infty$.*

Demostración. Por el Lema 1.7.8, se tiene que las funciones de iteración de valor óptimo cumplen que $V_n \leq V_\alpha^*$. Así

$$\lim_{x \rightarrow \infty} V_\alpha^*(x) = \infty.$$

y además $V_n \uparrow V_\alpha^*$, así V_α^* es una función convexa. ■

Además, por la convergencia de $V_n \uparrow V_\alpha^*$ se tiene que H_n y G_n convergen a H y G , respectivamente. Por último se tiene la optimalidad de la política (s, S) para el problema con horizonte infinito, la cual se presenta en el siguiente resultado.

Proposición 3.1.7 *Existen s y S que cumplen las siguientes condiciones:*

- I) $G(S) \leq G(y)$ para cada y ,
- II) $G(S) + K = G(s) < G(y)$ para cada y ,
- III) $G(y)$ es una función decreciente en $(-\infty, s)$,
- IV) $G(y) \leq G(z) + K$ para cada y, z con $s \leq y \leq z$,

por lo tanto, la política (s, S) es óptima para el criterio de costo total α -descontado con horizonte infinito.

Demostración. Por el Lema 3.1.6, la función G satisface (d) del Lema 3.1.4 entonces se tienen las condiciones (I) a (IV) de esta proposición. Luego por el cambio de variable $y = x + a$ y (I) a (IV) se tiene que la política (s, S) es óptima. ■

Proposición 3.1.8 Para cada s_n y S_n con $n \geq 1$ encontradas en la Proposición 3.1.5, se tiene que convergen a algunos s^* y S^* , respectivamente.

Demostración. Por el Teorema 1.7.9, se tiene que $V_n \uparrow V_\alpha^*$, así, $G_n \uparrow G$, entonces

$$G_n(y) \leq G_{n+1}(y) \leq G(y),$$

para cada $y \geq 0$. Observe que s_n y S_n , se encuentran en el conjunto

$$\{y \geq 0 : G(y) \leq \min_{y \geq 0} G(y) + K\},$$

el cual es compacto, por la continuidad inferior de la función G que se hereda de la continuidad inferior de la función V_α^* y además el conjunto en (3.1) es acotado por el inciso IV de la Proposición 3.1.7. De esta forma existe el punto límite de s_n y S_n los cuales se denotan por s^* y S^* , respectivamente. ■

3.1.1. Política miope

En esta sección, vamos a determinar un tipo de política más simple que del tipo (s, S) , llamada miope. La definición formal de política miope es la siguiente

Definición 3.1.9 Una política estacionaria, determinista $\pi = \{f_n\} \subset \mathbb{F}$ se llama miope si

$$f_n = f_1, \quad \forall n \geq 1.$$

En la literatura (ver [4] y [?]) se estudia la optimalidad de políticas miopes para el modelo de Lindley original (2.2).

En esta parte, mediante el modelo (2.3) con una función de costo

semejante a la utilizada anteriormente y espacio de estados y acciones admisibles compactos, demostramos la optimalidad de una política miope para este modelo de Lindley generalizado.

Consideremos el modelo (2.3), con las mismas suposiciones sólo que el espacio de estados es $X = [0, L]$ y espacio de acciones admisibles $A(x) = [R, Q]$ y la función de costo siguiente:

$$C_1(x, a) = ca + hE[(x + \eta a)] - lE[\min(x + \eta a, \xi)], \quad (3.7)$$

en donde

- c el costo de producción por unidad,
- h el costo de almacenaje por unidad,
- l el precio de venta del producto.

En el siguiente lema, presentamos las consecuencias del modelo con estas nuevas suposiciones.

Lema 3.1.10 *La función de costo es:*

- *Acotada inferiormente,*
- *Inferiormente semicontinua sobre \mathbb{K} ,*
- *inf – compacta.*

Demostración. La función de costo es acotada inferiormente, debido a que el espacio de estados y acciones son acotados, así para cada $x \in [0, L]$ y $a \in [R, Q]$ se tiene,

$$\begin{aligned} C_1(x, a) &\geq -l(x + a) \\ &\geq -l(M + L). \end{aligned}$$

Luego, los dos primeros términos de la función de costo C_1 son los mismos que tiene la función de costo C definida en (2.5) y en el Lema 2.3.3 se probó que son inferiormente semicontinuas. Para el tercer término, se tiene que la función \min es inferiormente semicontinua en $[0, L] \times [R, Q]$. Así, se tiene que la función de costo C_1 es inferiormente semicontinua en $[0, L] \times [R, Q]$. La función es *inf – compacta*, debido a que el espacio de estados y acciones admisibles es compacto. ■

Por último, sabemos que el kernel inducido por la dinámica (2.3) es fuertemente continuo, por el Lema 2.3.2. Así se tiene el siguiente resultado.

Teorema 3.1.11 *El modelo (2.2), satisface las Suposiciones 1.7.2 y 1.7.3 por tanto, se cumplen las conclusiones del Teorema 1.7.9.*

Ahora, en el siguiente Lema mostramos el comportamiento de las funciones de iteración de valor óptimo.

Lema 3.1.12 *Las funciones de iteración de valor óptimo V_n , con $n = 0, 1, \dots$, son no decrecientes en $[0, L]$, y además la función de valor óptimo V_α es no decreciente en $[0, L]$.*

Demostración. Por definición,

$$V_0 \equiv 0,$$

así, es no decreciente en $[0, L]$ y para $n \geq 1$,

$$\begin{aligned} V_{n+1}(x) &= \min_{a \geq 0} \{C_1(x, a) + \alpha E [V_n((x + \eta a - \xi)^+)]\} \\ &= \min_{a \geq 0} \{g(x + a) + \alpha p E [V_n((x + a - \xi)^+)]\} + h(x) \\ &\quad + \alpha(1 - p)E [V_n((x - \xi)^+)], \end{aligned}$$

en donde

$$\begin{aligned} g(x + a) &= (c + hp)(x + a) + lpE[\min\{x + a, \xi\}] \\ &= (c + hp)(x + a) + lpy(1 - F(x + a)) + lp \int_0^{x+a} s dF(s), \end{aligned}$$

y

$$\begin{aligned} h(x) &= (h(1 - p) - c)x + l(1 - p)E[\min\{x, \xi\}] \\ &= (h(1 - p) - c)x + l(1 - p)x(1 - F(x)) + lp \int_0^x s dF(s), \end{aligned}$$

entonces para cada $x \in [0, L]$, la derivada de h

$$h'(x) = (h(1 - p) - c) + l(1 - p)(1 - F(x)) > 0. \quad (3.8)$$

Por tanto, la función h es no decreciente en $[0, L]$.

Realizando el cambio de variable $y = x + a$, y calculando la derivada de la función g ,

$$g'(y) = c + hp + lp(1 - F(y)) > 0, \quad (3.9)$$

por tanto g es no decreciente en $[0, L]$. Así la primera iteración es,

$$\begin{aligned} V_1(x) &= \min_{y \in [x+R, Q+x]} \{g(y)\} + h(x) \\ &= g(x + R) + h(x). \end{aligned}$$

De esta forma, la función V_1 es no decreciente, y

$$f_1(x) = R, \quad x \in [0, L]$$

La segunda iteración,

$$\begin{aligned} V_2(x) &= \min_{y \in [R+x, Q+x]} \{g(y) + \alpha p E [V_1((y - \xi)^+)]\} \\ &\quad + h(x) + \alpha(1 - p) E [V_1((x - \xi)^+)] \\ &= g(x + R) + \alpha p E [V_1((x + R - \xi)^+)] + h(x) \\ &\quad + \alpha(1 - p) E [V_1((x - \xi)^+)]. \end{aligned}$$

La última igualdad se debe a que la función V_1 es no decreciente. Así,

$$f_2(x) = R, \quad x \in [0, L].$$

Supongamos que para $n \geq 2$, la función V_n es no decreciente, veamos que V_{n+1} es no decreciente. Sabemos que

$$\begin{aligned} V_{n+1}(x) &= \min_{y \in [R+x, Q+x]} \{g(y) + \alpha p E [V_n((y - \xi)^+)]\} + h(x) \\ &\quad + \alpha(1 - p) E [V_n((x - \xi)^+)], \end{aligned}$$

así, utilizando que la función V_n es no decreciente se tiene que

$$\begin{aligned} V_{n+1}(x) &= g(x) + \alpha p E [V_n((x + R - \xi)^+)] + h(x) \\ &\quad + \alpha(1 - p) E [V_n((x - \xi)^+)]. \end{aligned}$$

Nuevamente, utilizando que V_n es no decreciente, y como se demostró anteriormente que las funciones h y g son no decrecientes se tiene que V_{n+1} es no decreciente. Además, la estrategia en el n -ésimo paso es

$$f_n(x) = R, \quad x \in [0, L].$$

Por el Teorema 1.7.9, $V_n \uparrow V_\alpha$, en donde se demuestra que las funciones de iteración son no decrecientes, entonces la función de valor óptimo α -descontada es no decreciente. ■

Por la ecuación de optimalidad, para cada $x \in [0, L]$,

$$\begin{aligned} V_\alpha(x) &= \min_{y \in [R+x, Q+x]} \{g(y) + \alpha p E [V_\alpha((y - \xi)^+)]\} + h(x) \\ &\quad + \alpha(1 - p) E [V_\alpha((x - \xi)^+)], \end{aligned}$$

y por el lema anterior las funciones V_α y h son no decrecientes, entonces la política óptima estacionaria para el problema con horizonte infinito es

$$f(x) = R, \quad x \in [0, L].$$

La cual resulta ser miope. Por lo tanto, para este modelo, todas las políticas óptimas estacionarias para el problema con horizonte finito e infinito son miopes.

Por último, señalamos que el espacio de estados se puede extender a $X = [0, \infty)$, pero restringiendo la dinámica de la siguiente forma,

$$x_{t+1} = \min\{(x_t + \eta_t a_t - \xi_{t+1}), L\}. \quad (3.10)$$

Esta dinámica se estudia en [18], utilizando una función de costo

$$c(x, a) = [L - x]^+.$$

3.2. Ejemplo numérico

En esta sección se considera una base de datos encontrados en [19] de una compañía llamada Shanghai Tang Fashion Company que modelaron la gestión del inventario de manera determinista. Esta compañía tiene giro en la industria textil y se encarga de fabricar dos marcas de poliéster (Benhu, Donker). En la base de datos tenemos información de costos de almacenaje, producción y escasez de cada una de las dos marcas que se producen. También tenemos datos de la demanda de cada poliéster en el año 2014. Con estos datos de la demanda, realizamos un ajuste a una distribución de probabilidad. Luego modelamos el sistema de inventarios mediante (2.3), así vemos que se cumplen las Suposiciones 2.3.1, con el fin de recomendar la utilización de estrategias (s, S) para esta compañía. Así, de manera numérica determinamos cada factor s y S utilizando el software Mathematica 11.0.

3.2.1. Gestión de inventarios en Shanghai Tang Fashion Company

Shanghai Tang Fashion Company es una empresa que tiene giro en la industria textil. Este tipo de industrias en uno de los sectores manufactureros más importantes de Indonesia, en donde tiene sucursales esta compañía. Datos relevantes que se mencionan en [19], es que esta rama de industrias generaron 1.1 millones de empleos en 2012, haciendola una de las principales proveedores de campo de trabajo.

También se menciona, que la industria textil de Indonesia todavía se enfrenta a una serie de problemas internos y externos, especialmente en la actualidad. Estos consisten principalmente en la utilización excesiva de la capacidad y la ineficiencia de maquinaria con incertidumbre de planificación debido a las fluctuaciones del tipo de cambio que conllevan el aumento de costos de producción y escasez de recursos humanos calificados. Shanghai Tang Fashion Company se enfrenta a un problema similar en términos de inexactitud de gestión de inventarios, donde se tiene incertidumbre en la

demanda, por ejemplo; la cantidad, color y tipo de productos de la empresa. Los productos que maneja la compañía son los siguientes:

- Rebaba de poliéster Benhu,
- Rebaba de poliéster Donker.

En cada producto se desea encontrar una política de gestión de inventario para poder minimizar costos. El inventario se revisa al inicio de la primera semana del año y se toma la decisión de la cantidad de poliéster en kilogramos que se va a utilizar. Al finalizar la semana, se cuenta la cantidad de producto vendida en kilogramos y se guarda el nivel de inventario. Para esto, consideramos las siguientes variables:

1. La cantidad de poliéster Benhu que se tiene en el inventario en el mes t se denota por x_t^1 , para el poliéster Donker es x_t^2 .
2. La cantidad de poliéster Benhu producido en el mes t es denotado por a_t^1 , para Donker es a_t^2 .
3. El éxito de haber logrado colocar en el almacén y a la venta el poliéster Benhu es η_t^1 , para el poliéster Donker η_t^2 .
4. La demanda que se tiene al finalizar el periodo t para el poliéster Benhu es ξ_{t+1}^1 , para el Donker es ξ_{t+1}^2 .

De esta forma, modelaremos el inventario de cada producto que se maneja en la compañía con las siguientes ecuaciones en diferencias:

Benhu

$$x_{t+1}^1 = (x_t^1 + a_t^1 \eta_t^1 - \xi_{t+1}^1). \quad (3.11)$$

Donker

$$x_{t+1}^2 = (x_t^2 + a_t^2 \eta_t^2 - \xi_{t+1}^2). \quad (3.12)$$

Para cada $t = 0, 1, 2, \dots$, la capacidad del inventario y de producción es muy grande, entonces $x_t^1, x_t^2 \in [0, \infty)$ y $a_t^1, a_t^2 \in [0, \infty)$.

En [19], se presenta un estudio del inventario durante el año 2014. Primero se observó que la eficiencia de colocar en el almacén lo que se requiere producir en cada mes. Los resultados se observan en la Tabla 3.1.

Así, supongamos que cada éxito o fracaso de colocar el producto

Producto	Benhu	Donker
Eficiencia	85 %	80 %

Cuadro 3.1: Eficiencia de colar producto en el almacén.

en el almacén son independientes y además cada realización, η_t^1

tiene una distribución Bernoulli con parámetro 0.85, para η_i^2 son Bernoulli con parámetro 0.80.

Ahora, para cada tipo de poliéster se tienen demandas independientes en cada mes. En base a los datos presentados en [19], recolectamos las demandas de las primeras 50 semanas del año 2014. En la Tabla 3.2 se observa las demandas en algunas semanas.

Demanda (kg)	Benhur	Donker
Semana 1	229	632
Semana 2	253.9	554.4
Semana 3	266	495.9
Semana 5	248.5	407.1
Semana 10	191.11	380.3
Semana 20	200.7	350.5
Semana 30	306.5	392.8
Semana 40	332.5	461.7
Semana 50	130	356.8

Cuadro 3.2: Demanda en el año 2014.

Con los datos obtenidos de las demandas en el año 2014, realizamos un ajuste a una distribución de probabilidad. Así utilizando el Test Kolmogorov-Smirnov, se ajusta la demanda de cada poliéster a una distribución mixta Normal /Uniforme con los siguientes parámetros dados en la Tabla 3.3.

Demanda / Poliéster	Benhur	Donker
Uniforme (a, b)	$a = 234.8, 268$	$a = 426.8, b = 625.1$
Normal (μ, σ)	$\mu = 266.3, \sigma = 1.7$	$\mu = 452.3, \sigma = 16.8$
Nivel de Confianza	94%	97%

Cuadro 3.3: Ajuste de la Demanda.

Ahora vamos a definir la función de costo de cada poliéster. En la Tabla 3.4 presentamos los costos fijos de producción, por unidad ordenada o producida de almacenaje y por unidad de escasez (cada una de ellas medidas en kg). Estos datos pueden encontrarse en [19] para cada tipo de poliéster. Así nuestra función de costo para tipo de poliéster se expresan con las siguientes ecuaciones.

Benhur

$$C^1(x, a) = K^1 1_{\{a \neq 0\}} + c^1 a + h^1 E[(x + \eta^1 a)] + l^1 E[(\xi^1 - (x + \eta^1 a))^+].$$

Producto	Benhur	Donker
Costo fijo orden	$K^1 := 52,3$	$K^2 := 95,6$
Costo por orden	$c^1 := 144,3$	$c^2 := 332,5$
Costo por almacenaje	$h^1 := 321,2$	$h^2 := 340,8$
Costo por escasez	$l^1 := 197,2$	$l^2 := 169,2$

Cuadro 3.4: Costos por kg de Poliéster

Donker

$$C^2(x, a) = K^2 1_{\{a \neq 0\}} + c^2 a + h^2 E[(x + \eta^2 a)] + l^2 E[(\xi^2 - (x + \eta^2 a))^+].$$

3.2.2. Política óptima (s, S)

En la parte anterior, se adaptó al modelo (2.3) a cada tipo de poliéster, en las ecuaciones (3.11) y (3.12). Además, se definió la función de costo. Además, se tiene que se cumplen las Suposiciones 2.3.1, de esta forma se cumple el Teorema 2.3.5 y se tiene que para cada tipo de poliéster tenemos un política óptima. Además, se cumple la Proposición 3.1.5, con lo cual cada tipo de poliéster tiene una política óptima de reorden o del tipo (s, S) .

Con el programa Mathematica 11.0, realizamos las estimaciones de las funciones de valor óptimo en cada etapa, $n = 1, 2, \dots, 52$ y los valores s_n y S_n que caracterizan a la política óptima. Recordar que S_n es el mínimo de la función G_n definida en (3.1.2). El valor de s_n se calcula resolviendo la ecuación:

$$G_n(S_n) + K = G_n(s_n). \quad (3.13)$$

Dichas condiciones están dadas en la Proposición 3.1.5. En las Tablas 3.5 y 3.6 mostramos los resultados utilizando un factor de descuento $\alpha = 0.9$ para cada producto. Por lo tanto, se observa la convergencia de s_n y S_n que se garantiza en la Proposición 3.1.8. De acuerdo a los resultados, tenemos las siguientes recomendaciones para obtener el costo óptimo al fin de año.

- Para el poliéster Benhur, si se tiene una capacidad menor a 287.5 kg, se solicita producir más kg de poliéster hasta tener 300.3 kg. En caso contrario no se solicita producción.
- Para el poliéster Donker, se produce más kg de este tipo de poliéster hasta tener 637.1 kg en el almacén siempre que el stock se encuentre por debajo de 431.2 kg. En caso contrario no se solicita producción.

Semana	s_n	S_n	
1	287.5	298.3	
2	287.8	298.9	
10	289.6	299.2	
Benhur	20	289.8	299.3
	30	290.1	299.7
	40	290.3	299.9
	50	290.5	300.1
	52	290.7	300.3

Cuadro 3.5: Estrategia (s,S) para Benhur.

Semana	s_n	S_n	
1	431.2	632.87	
2	433.5	633.8	
10	431.9	634.1	
Donker	20	432.1	635.9
	30	432.4	636.2
	40	432.6	636.7
	50	433.2	636.9
	52	434.5	637.0

Cuadro 3.6: Estrategia (s,S) para Donker.

Poliéster/Costo	Benhur	Donker
Costo Total existente	241 428.6	500 414.7
Costo modelo estocástico	186 321.3	329 361.6
Costo modelo determinista	204 062.738	394 273.8

Cuadro 3.7: Comparativo de los costos.

En la Tabla 3.7, mostramos la comparativa de los costos en 52 semanas que utilizan con un modelo determinista en [19], nuestra propuesta con el modelo estocástico e implementación de la política óptima (V_{52}), y el costo total existente en la compañía. Se observa que la aplicación a un modelo estocástico y la implementación de la política (s, S) reduce los costos a comparación de utilizar un modelo determinista.

Por último, gráficamente presentamos una simulación de la gestión de inventario para cada tipo de poliéster utilizando la política óptima (s, S) .

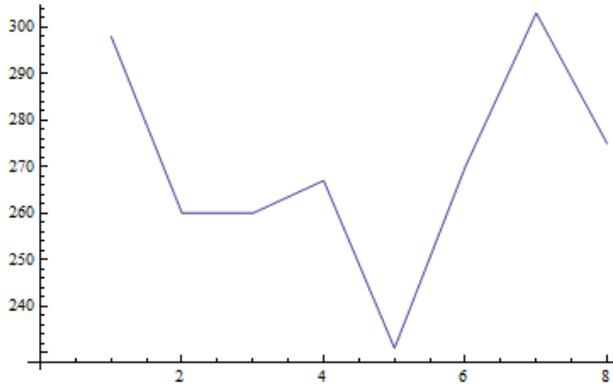


Figura 3.1: Simulación Benhur

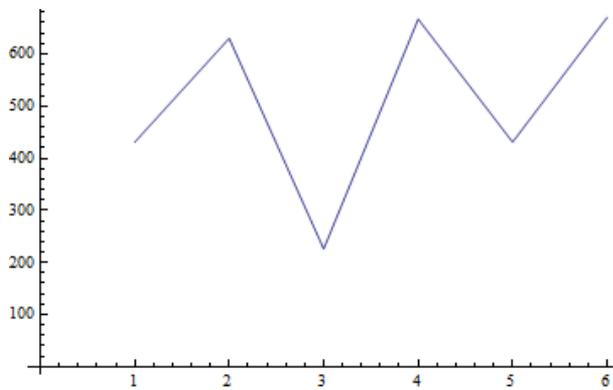


Figura 3.2: Simulación Donker

En las Figuras 3.1 y 3.2 se observa el comportamiento del inventario bajo un control entre 287.4 kg y 303.3 kg para Benhur y 431.2 kg y 637.1 kg para Donker.

Conclusiones

En el trabajo se presentó un resumen de las aportaciones más importantes y avances que se han hecho para demostrar optimalidad de políticas (s, S) en modelos de inventarios. También se realiza el desarrollo de la teoría básica de los PDMs para aplicarla a un modelo de inventario basado en la caminata aleatoria de Lindley. Se presenta una aportación a lo que se ha hecho anteriormente, la cual fue demostrar optimalidad de políticas (s, S) en el modelo generalizado de Lindley bajo nuevas condiciones en el espacio de estados y acciones admisibles. La forma como se resolvió el problema fue a través de programación dinámica y el estudio de funciones convexas. También agregamos un caso especial, el cual es la política miope que tiene como finalidad ser más simple, para la cual logramos demostrar que bajo otras condiciones a la función de costo y espacio de estados y acciones admisibles se tiene la optimalidad de este tipo de política.

Finalmente, se presenta un ejemplo numérico con datos de una compañía de textil en Indonesia para mostrar el comportamiento numérico de políticas (s, S) .

Se han derivado algunos problemas que se pretenden trabajar a futuro, tales como:

- Aplicar estos nuevos resultados teóricos para optimizar la gestión de un inventario en México. Cabe destacar la dificultad que existe para poder encontrar datos reales.
- El modelo presentado de Lindley generalizado, tiene como hipótesis que la sucesión de variables aleatorias *i.i.d.*, $\{\eta_t\}$ tienen distribución Bernoulli con parámetro p , entonces se busca generalizar la distribución de estas variables con el fin de seguir garantizando la optimalidad de políticas (s, S) .
- Estudiar la optimalidad de las políticas (s, S) pero utilizando el criterio de costo promedio. Existen ya avances al respecto para otros modelos de inventarios [8], pero en la literatura no hemos encontrado estudios para el modelo de Lindley analizado.

Apéndice A

Resultados y Definiciones Auxiliares

Definición A.1.1 Sea (X, τ) un espacio topológico, la mínima σ -álgebra que contiene a τ es la σ -álgebra de Borel, es decir, la σ -álgebra generada por τ se denota por $B(X)$.

Definición A.1.2 X es un espacio de Borel, si X es un subconjunto de Borel de un espacio métrico, separable y completo.

Definición A.1.3 Un kernel estocástico definido sobre X dado Y es una función $Q(\cdot|\cdot)$ tal que:

1. $Q(\cdot|y)$ es una medida de probabilidad en X , para cada $y \in Y$.
2. $Q(B|\cdot)$ es una variable aleatoria (función medible) en Y para cada $B \in B(X)$.

La familia de todos los kernels estocásticos lo denotaremos por $P(X|Y)$.

Considere X y A espacios de Borel no vacíos.

Definición A.1.4 Una multifunción (también llamada correspondencia o mapeo conjunto valuado) $\psi : X \rightarrow A$, es una función tal que $\psi(x)$ es un subconjunto no vacío de A , para cada $x \in X$.

Se define el grafo de la multifunción ψ es el subconjunto de $X \times A$ dado por

$$\mathbb{K} := \{(x, a) | x \in X \text{ y } a \in \psi(x)\}.$$

Teorema A.1.5 Suponga que \mathbb{K} es un subconjunto de Borel de $X \times A$, ν es una función l.s.c., acotada inferiormente e inf-compacta sobre \mathbb{K} . Entonces

1. Existe un selector $f^* \in \mathbb{F}$ tal que

$$v(x, f^*(x)) = v^*(x) = \min_{a \in A(x)} v(x, a).$$

2. Si multifunción $x \rightarrow A(x) := \{a \in A(x) | v^*(x) = v(x, a)\}$ es l.s.c. entonces v^* es l.s.c.. Más aún, si v es continua entonces v^* también lo es.

Demostración. Ver [12]. ■

Teorema A.1.6 Teorema de convergencia dominada Sean $\{f_n\}$ una sucesión de funciones medibles y g una función medible y μ -integrable definidas en X , tales que $|f_n| \leq g$, para toda n . Si $f_n \rightarrow f$ casi seguramente, entonces, f es μ -integrable y

$$\int_X f d\mu = \lim_{n \rightarrow \infty} \int_X f_n d\mu.$$

Demostración. Ver [1]. ■

Proposición A.1.7 Teorema de Ionescu-Tulcea Sea X_0, X_1, \dots , una sucesión de espacios de Borel, y para cada $n = 0, 1, \dots$, se define $Y_n := X_0 \times X_1 \times \dots \times X_n$ y $Y_n := \prod_{n=0}^{\infty} X_n$. Sea ν una medida de probabilidad arbitraria sobre X_0 y para cada $n = 0, 1, \dots$, $P(dx_{n+1}|y_n)$ es un kernel estocástico sobre X_{n+1} dado Y_n . Entonces existe una única medida de probabilidad P_ν sobre Y tal que, para cada rectángulo medible $B_0 \times \dots \times B_n$ en Y_n ,

$$P_\nu(B_0 \times \dots \times B_n) = \int_{B_0} \nu(dx_0) \int_{B_1} P_0(dx_1|x_0) \\ \dots \int_{B_n} P_{n-1}(dx_n|x_0, \dots, x_{n-1}).$$

Además, para cualquier función u medible y no negativa sobre Y , la función

$$x \rightarrow \int u(y) P_x(dy)$$

es medible en X_0 , donde P_x representa a P_ν cuando ν es la probabilidad concentrada en $x \in X_0$.

Demostración. Ver [1]. ■

Lema A.1.8 Sea $C : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$, la cual se puede descomponer por dos funciones $H : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ y $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ de la siguiente forma:

$$C(x, a) = H(x, a) + g(a).$$

Si la función g es inferiormente semicontinua l.s.c. e inf-compacta en $\mathbb{R}^+ \times \mathbb{R}^+$ y la función H es inferiormente semicontinua entonces la función H es inf-compacta sobre $\mathbb{R}^+ \times \mathbb{R}^+$.

Demostración. Ver [8] ■

Bibliografía

- [1] Ash R. B. and Doléans-Dade C.A., Probability and Measure Theory Academic Press Elsevier, San Diego, 2005.
- [2] Bellman R., Dynamic Programming. Dover, 2003.
- [3] Bertsekas D.P., Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Inc., New Jersey, 1987.
- [4] Bensoussan A., Dynamic Programming and Inventory Control, Studies in Probability, Optimization and Statistics. Vol. 3, Amsterdam, 2011.
- [5] Chen X. y Simchi-Levi D., Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: the infinite horizon case. Operations Research Center, Vol 29, pp. 698-723, 2004.
- [6] Daduna H., Knopov P. and Tur L., Optimal Strategies for an Inventory System with Cost Functions of General Form. Cybernetic and Systems Analysis, Springer, 1999.
- [7] Ehrhardt R., The Power Approximation for Computing (s, S) Inventory Policies. Management Science, pp. 777-786, 1979.
- [8] Feinberg E., Optimality Conditions for Inventory Control, arxiv:1606.00957v1, 2016.
- [9] Feinberg E. y Lewis M., On the Optimality of (s, S) Policies. arxiv:1507.05125v1, 2015.
- [10] Goldberg D., Dmitriy A., Yingdong L., Sharma M., Squillante M. Asymptotic Optimality of Constant-Order Policies for Lost Sales Inventory Models with Large Lead Times. Septiembre 2014. arxiv:1211.4063v2.
- [11] Guo X y Zhu Q. Average optimality for Markov decision processes in borel spaces: a new conditions and approach. Journal of Applied Probability, Vol. 49, 2006.

- [12] Hernández-Lerma O. and Lasserre J. B., Discrete-Time Markov Control. Processes: Basic Optimality Criteria, Springer-Verlag, New York, 1996.
- [13] Hernández-Lerma O. and Lasserre J. B., Further Topics on Discrete-Time Markov Control Processes, Springer-Verlag, New York, 1999.
- [14] Iglehart D. Capital Accumulation and Production for the Firm: Optimal Dynamic Policies. *Management Science*, Vol. 12, No. 3, pp. 193-205. Nov, 1995.
- [15] Iglehart D., Optimality of (s,S) Policies in the Infinite Horizon Dynamic Inventory Problem. *Institute for Operations Research and the Management Sciences*. pp. 259-267, 1963.
- [16] Johansen G. and Shaler S., Control of Arrivals to a Stochastic Input-Output System. *Advances in Applied Probability*, Vol. 12, No. 4, pp. 972-999, 1980.
- [17] Lindley D., The Theory of Queues with Single Server, *Proc. Cambridge Philos. Soc.* 48, pp. 277-289, 1952.
- [18] Montes-de-Oca R., Saavedra P., Zacarías G. and Cruz D., Optimal Policies for Payment of Dividend through fixed Barrier at Discrete Time. *Science and Technology Publications*. pp. 140-149, 2017.
- [19] Margono S. and Dwi Y., Determine the Appropriate Inventory Model in Tang Company. *Journal of Business and Management*. Vol.4, No. 4, 2015.
- [20] Özekici S. y Parlar M., Inventory models with unreliable suppliers in a random environment. *Operations Research*, pp.123-136, 1999.
- [21] Porteus E., On the Optimality of Generalized (s,S) Policies. *Management Science*, 1971.
- [22] Scarf H., The Optimality of (S,s) Policies in the Dynamic Inventory Problem. *Applied Mathematics and Statistics Laboratory*, Standford University, California, No. 11, 1959.
- [23] Schal M., On The Optimality of (s,S)- Policies in Dynamic Inventory Models with Finite Horizon. *Siam J. Appl. Math.*, Vol. 30, No. 3, 1976.
- [24] Suresh P. and Feng C., Optimality of (s,S) Policies in Inventory Models with Markovian Demand. *Operations Research*. Vol. 4, No. 6, pp. 931-940, 1995.

-
- [25] Sebnem S. and Samyuhta R., A Markov Decision Process Based Policy Characterization Approach for a Stochastic Inventory Control Problem with Unreliable Sourcing. *International Journal of Production Economics*. Vol. 144, No. 2, 2013.
- [26] Vega O. and Montes-de-Oca R., Application of Average Dynamic Programming to Inventory Systems. *Mathematical Methods of Operations Research*. Vol. 47, pp. 451-471, 1997.
- [27] Veinott A., On the Optimality of (s,S) Inventory Policies: New Conditions and New Proof. *SIAM J. Appl. Math.* No. 14, pp. 1067-1089, 1966.
- [28] Veinott A. and Wagner H. Computing Optimal (s,S) policies. *Management Science*. Vol. 11, No. 5, pp. 525-552, 1965.
- [29] Yu-Sheng Z. and Federgruen A., Finding Optimal (s,S) Policies is about as Simple as Evaluating a Single Policy. *Operations Research Society of America*. Vol. 39, pp. 654.665, 1991.
- [30] Zipkin P., Old and New Methods for Lost Sales Inventory Systems *Operations Research*. Vol. 56, No. 5, pp. 1256-1263, 2008.