6

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

EL MODELO DE CONTROL MARKOVIANO APLICADO A UN PROBLEMA DE COLAS Y UNO DE SEGUROS

T E S I S

QUE PARA OBTENER EL TÍTULO DE

A C T U A R I A

P R E S E N T A :

M A R Í A I S A B E L A L O N S O G A R C Í A

D R . J U A N G O N Z Á L E Z H E R N Á N D E Z

2 0 0 2

FACULTAD DE CIENCIAS
UNAM

VNIVERSIDAD NACIONAL
AVPNMA DE
MEXICO

M. EN C. ELENA DE OTEYZA DE OTEYZA
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo escrito:

El modelo de control Markoviano aplicado a un problema de colas y uno de seguros.

realizado por María Isabel Alonso García

con número de cuenta 8904787-7 , quién cubrió los créditos de la carrera de Actuaría.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	Dr. Juan González Hernández
Propietario	M. en C. Guadalupe Carrasco Licea
Propietario	Act. Marisa Miranda Tirado
Suplente	Act. Hortensia Cano Granados
Suplente	Fís. Guillermo César Aníbal Zepeda Pérez.

Consejo Departamental de Matemáticas

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

M. en C. José Antonio Flores Díaz

No sė como iniciar...
Me dije un día.
Y sin darme cuenta me ate a esta cometa
y hoy termino.

Por que se que tu me acompañas en todos mis caminos.
Gracias.

Ser Madre y Padre no es fácil,
pero tú lo haces de maravilla;
por toda la esencia que nos das
por el cariño, la confianza, el apoyo que brindas;
simplemente por que te quiero mucho
y que gracias a ti, concluyo un ciclo en mi vida.
Maaaá
Con todo mi corazón te dedico este
Sueño hecho realidad.

A mis queridos hermanos:
Lur Má y David.

Con todo mi cariño a esos angelitos:
Alejandra por crecer a nuestro lado, eres muy especial,
¡Gemelas! ¡Que Sorpresa!, Karla y Daniela
Super Ximena.

A mis primos, en especial a las chicas super-desastrozas,
Diana y Fernanda.

A mi abuelita Sra. Catalina.
A mis tíos; tías y cuñados.
Por alentarme siempre, gracias.

Por tener el privilegio de encontrarnos y compartir tantos
Momentos, secretos, sueños, risas, ideas, ideologías
Por querer arreglar el mundo, lo mismo en la pared que en la calle,
por decir las cosas sin temor.
A perder ese lazo
Que es la AMISTAD
A mis adorables amigos y maestros:
Luz Má, Aida, Teté, Alexina, Marú, Rosa, Gabriel,
Lalo, Ariel, Toño, Raúl, Martha, Marcela, Ara, César...

De manera muy especial, a Usted Maestro, por
Toda su paciencia, comprensión y ayuda.
Gracias: Dr. Juan González Hernández.

En este caos, ha bastado sólo una mirada
Para rendirme ante ti, y dejarte ver mis
sueños una tarde cálida de Invierno,
por ir quitando la maleza de mi camino.
Por ayudarme a poner orden a este sueño,
por todo lo encantador que eres,
por todo lo que haz hecho
a lo largo de este recorrido.
Mil gracias.

Floriberto

El modelo de control Markoviano aplicado a un problema de colas y uno de seguros.

Prólogo

En todas las actividades que desempeña el hombre, ya sea en forma individual o en grupo, posee la capacidad de tomar decisiones en beneficio de él mismo o de la comunidad u organización a la que pertenezca. Sin embargo, en ocasiones el rumbo no es el idóneo y por consiguiente puede perjudicar a los intereses que se desean sostener o lograr.

Al respecto, el tomar la decisión más apegada a lo que se persigue esta influenciada o se efectúa ya sea aplicando el criterio de quien la toma, de acuerdo a la jerarquía o rango, intuición o al contexto que impere en ese momento.

Pero, ¿Cuántos tomadores de decisiones de nuestro país basan sus acciones en modelos matemáticos que les permitan optimizar los objetivos apegados al bien común? A pesar de que los modelos existen y pueden simular de una manera objetiva la realidad, estos son ignorados y no se consideran como una herramienta clave. Pues esto depende en gran medida de la formación y albedrío de quien toma la decisión.

En contraste, en muchas compañías tanto privadas como públicas de algunos países Orientales, Europeos y de Norte América estos modelos gozan de gran anuencia; en tanto que, en países como el nuestro estas herramientas son poco conocidas y por ende muy poco aplicadas; pues quienes están a cargo de la toma de decisiones desconocen la forma de aplicarlos. Simplemente, a varios funcionarios de una prestigiada empresa aseguradora se les cuestionó sobre este conocimiento, declarando desconocerlo o que no es necesario esta clase de herramientas en empresas mexicanas, pues la administración es muy distinta y se aplica lo que "diga el jefe".

Por lo antes expuesto, a lo largo del presente trabajo nos enfocaremos al estudio en específico, de entre todo el universo de modelos matemáticos al Modelo de Control Markoviano.

En esta línea consideraremos dos variantes, abordaremos el Modelo de Control Markoviano sin Restricciones y el Modelo de Control Markoviano Restringido. Lo cual desembocará en la aplicación para un modelo de seguros.

Para ello se ha dividido el trabajo por medio de cinco capítulos que nos acercarán tanto a la Teoría del Control Markoviano como al modelo de seguros.

De esta forma, el Capítulo 1 proporciona un preliminar de algunos resultados de la teoría de la probabilidad; iniciando con la definición de ésta para posteriormente dar el concepto de experimento aleatorio; destacaremos la definición de sigma-álgebra.

Se dedica una sección a la medibilidad, donde establecemos la definición de medida de probabilidad y destacamos sus propiedades, se concluirá con el espacio de probabilidades. En la sección de variables aleatorias se establece la función de distribución y la función de densidad conjunta. Como herramienta básica se tiene el concepto de probabilidad condicional e independencia, a su vez destacaremos el valor esperado, la función generadora de momentos, la función característica y la transformada de Laplace; y la esperanza condicional. Mencionamos la distribución exponencial resaltando la pérdida de la memoria, así como las funciones de tasa de riesgo.

Adicionalmente el capítulo contiene algunas distribuciones típicas como son: la binomial, Bernoulli, geométrica, hipergeométrica, Poisson y la uniforme.

El Capítulo 2, abordará lo relacionado con los procesos estocásticos, el capítulo inicia con la explicación de la caminata aleatoria, con el objeto de motivar el estudio de los procesos estocásticos. Resulta de gran utilidad la sección dedicada a los problemas de la Ruina y la duración esperada del juego.

Posteriormente se presentan las cadenas de Markov, donde se establecen las ecuaciones de Chapman-Kolmogorov, la clasificación de estados, las probabilidades límites; así como el concepto de tiempo de paro.

Habremos de concluir el capítulo con el estudio del proceso Poisson, en el cual definimos los principales resultados de esta teoría.

Entraremos en materia al conocer los principales argumentos de la teoría de los procesos de decisión de Markov, lo cual nos aproximará al Capítulo 3, donde habremos de conocer la descripción del modelo sin restricciones.

Así mismo, aplicaremos los resultados de la existencia de soluciones, ecuación de Bellman y criterio de óptimalidad a el costo total y al costo descontado.

Dedicamos el Capítulo 4 a los problemas de control óptimo con restricciones, habremos de considerar la formulación del problema restringido, se consideran las herramientas para la solución de este tipo de cuestiones, de esta forma se presenta el método de los multiplicadores de Lagrange y los resultados auxiliares que podremos aplicar en el capítulo 5.

Es importante resaltar que a lo largo del desarrollo de los capítulos 3 y 4 se propone a manera de ilustración la solución de un ejemplo de Colas, el cual nos apoyará para fijar ideas. Primero mostrando el problema sin restricciones y posteriormente con restricciones.

El Capítulo 5 muestra la constucción de un modelo de seguros propuesto por Bruno de Finetti en el Congreso Internacional de Actuarios en 1957. En el se aplica la teoría desarrollada a lo largo de este trabajo, para encontrar una política óptima que provea un máximo de dividendos con la restricción de que la compañía dure al menos "n" años. De esta forma mostramos que los modelos son un medio adecuado para bosquejar la realidad.

Pues bien iniciemos...

Índice

Capítulo 1

Preliminares de probabilidad.

1.1 Probabilidad.

La teoría de la probabilidad estudia los métodos de análisis en el tratamiento de un fenómeno aleatorio. Es decir, estudia las propiedades de un fenómeno donde interviene la aleatoriedad.

En este capítulo expondremos los conceptos básicos de la teoría de la probabilidad como son: $\sigma-$*álgebra,* espacios de medida y de probabilidad, variable aleatoria, función de distribución, de densidad, característica y generadora de momentos, probabilidad condicional e independencia. Terminamos el capítulo mencionando brevemente algunas variables aleatorias que usaremos en el texto: exponencial, Bernoulli, binomial, geométrica, hipergeométrica, Poisson y uniforme. Algunas referencias para este capítulo son: [1] [3] [6] [7] y [16].

Una noción básica en la teoría de probabilidad es el concepto de *experimento aleatorio* o *fenómeno aleatorio*, el cual es un suceso empírico en el que no se sabe con exactitud el resultado que se obtendrá. El número de resultados puede ser finito o infinito dependiendo de la naturaleza del experimento. El conjunto de todos los posibles resultados de un experimento es llamado el *espacio muestral*, y lo denotaremos por **S**.

Definición 1.1 *$\sigma-$Álgebra.*

Sea $\mathcal{A}$ una colección de subconjuntos de **S**. Decimos que $\mathcal{A}$ es una *$\sigma-$ álgebra* sí:

a) $\mathbf{S}\in \mathcal{A}$.

b) Si $\mathbf{E}\in \mathcal{A}$, entonces $\mathbf{E}^c \in \mathcal{A}$.

c) Si $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$, $\ldots \in \mathcal{A}$, entonces $\cup_{i=1}^{\infty}\mathbf{E}_i \in \mathcal{A}$.

Un suceso o evento **E** es un subconjunto de **S** tal que $\mathbf{E}\in \mathcal{A}$, donde $\mathcal{A}$ es una *$\sigma-$ álgebra* de **S**.

Sea $\mathbf{S}=\mathcal{R}$ consideremos la clase de intervalos abiertos en $\mathcal{R}$. $\mathcal{B}$ la *$\sigma-$ álgebra* generada por los intervalos abiertos en $\mathcal{R}$. Es decir, sí denotamos por

$$\mathcal{G}=\{(a,\ b)\mid a,\ b\in \mathcal{R}\},$$

entonces

$$\sigma\left(\mathcal{G}\right)=\mathcal{B}\ la\ \sigma-\text{álg}\,ebra\ de\ Borel.$$

A los elementos de $\mathcal{B}$ se les llama borelianos.
En el caso X finito $\mathcal{A}=P\left(X\right)$.

Teorema 1.1 *Sea $\mathcal{F}$ la clase de conjuntos $(-\infty,\ x]$ $-\infty < x < \infty$ en $\mathcal{R}$. Entonces $\mathcal{B}$ es igual a la $\sigma-$álgebra generada por $\mathcal{F}$.*

Una *$\sigma-$ álgebra* es generada por una familia de conjuntos $\mathcal{B}$, si $\mathcal{B}\subseteq\mathcal{P}\left(X\right)$ entonces $\sigma-\mathcal{A}\left(\mathcal{B}\right)=\cap\mathcal{A}$. Donde la intersección se toma sobre todas las *$\sigma-$ álgebras* que contiene a $\mathcal{B}$.

Definición 1.2 *Una $\sigma-$*álgebra *producto es la generada por la familia de "cilindro" $(X_i,\mathcal{A}_i)$. $X=\prod_{i=1}^{\infty}X_i$*

$$\mathcal{B}=\left\{\begin{array}{l} A_1\times A_2\times\cdots\times A_n\times X_{n+1}\times X_{n+2}\times\cdots, \\ para\ A_1\in\mathcal{A}_1, A_2\in\mathcal{A}_2,\ldots A_n\in\mathcal{A}_n, n\in N\end{array}\right\}$$

1.2 Medibilidad.

Definición 1.3 *Espacio medible.*

Una pareja ordenada $(\mathbf{S},\ \mathcal{A})$ *tal que* S *es un conjunto y* $\mathcal{A}$ *es una* $\sigma-$*álgebra de subconjuntos de* S *es llamada el* espacio medible.

Definición 1.4 *Función de conjunto.*

Una función que mapea al conjunto de los reales extendidos $\mathcal{R}\cup\{-\infty,\ \infty\}$, *cuyo dominio es una colección de conjuntos, es llamada una función de conjunto.*

Definición 1.5 *Medida.*

Sea $(\mathbf{S},\ \mathcal{A})$ *un espacio medible. Una* medida *es una función* λ *definida sobre* $\mathcal{A}$ *tal que*

i) $\lambda(\emptyset)=0.$

ii) $\lambda(E)\geq 0$ para cada $\mathbf{E}\in\mathcal{A}$.

iii) λ es numerablemente aditiva, en el sentido de que si $\{E_n\}$ es una sucesión de elementos de $\mathbf{S}$ tal que $\mathbf{E}_i\cap\mathbf{E}_j=\emptyset$ para $i\neq j$, entonces

$$\lambda\left(\bigcup_{n=1}^{\infty}E_n\right)=\sum_{n=1}^{\infty}\lambda(E_n).$$

Definición 1.6 *Espacio de medida.*

Una terna ordenada $(\mathbf{S},\ \mathcal{A},\ \lambda)$, *donde* $(\mathbf{S},\ \mathcal{A})$ *es un espacio medible y* λ *es una medida sobre* $\mathcal{A}$, *es llamada un* espacio de medida.

Definición 1.7 *Medida de probabilidad.*

Sea $(\mathbf{S},\ \mathcal{A},\ P)$ *un espacio de medida entonces la función de conjuntos* P *es una medida de probabilidad, si* $P(S)=1$.

Si $E\in\mathcal{A}$ entonces a $P(E)$ se le llama la probabilidad del evento E.

Propiedades inmediatas de la medida de probabilidad. [1]

i) $P(\emptyset) = 0$.

ii) $0 \leq P(\mathbf{E}) \leq 1$.

iii) Si $E \subset F$ entonces $P(E) \leq P(F)$.

iv) Para cualesquiera $\mathbf{E}$, $\mathbf{F} \in \mathcal{A}$, $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.

v) Sea $\mathbf{E} \in \mathcal{A}$ entonces $P(E) = 1 - P(E^c)$.

vi) Si $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3, \ldots, \mathbf{E}_n \in \mathcal{A}$ tales que $\mathbf{E}_i \cap \mathbf{E}_j = \emptyset$ para $i \neq j$, entonces $P(\cup_{i=1}^{n} E_i) = \sum_{i=1}^{n} P(E_i)$.

vii) Desigualdad de Boole $P(\cup_{i=1}^{\infty} E_i) \leq \sum_{i=1}^{\infty} P(E_i)$, para cualesquiera $E_i \in \mathcal{A}$.

Definición 1.8 *Espacio de probabilidades.*
Un espacio de probabilidad es una terna $(\mathbf{S}, \mathcal{A}, P)$*, donde* $\boldsymbol{S}$ *es el espacio muestral,* $\mathcal{A}$ *es una* $\sigma-$*álgebra definida sobre* $\boldsymbol{S}$ *y* P *es una función de probabilidad sobre* $\mathcal{A}$.

Definición 1.9 *Evento límite.*
Una sucesión de eventos $\{E_n,\ n \geq 1\}$ *se dice que es una sucesión creciente si* $E_n \subset E_{n+1}$, $n \geq 1$ *y se dice decreciente si* $E_n \supset E_{n+1}$, $n \geq 1$.

Si $\{E_n,\ n \geq 1\}$ es una sucesión creciente de eventos, entonces definimos un nuevo evento, denotado por

$$\lim_{n \to \infty} E_n = \cup_{i=1}^{\infty} E_i, \text{ donde } E_n \subset E_{n+1}, n \geq 1.$$

Similarmente, si $\{E_n,\ n \geq 1\}$ es una sucesión decreciente, entonces definimos $\lim\limits_{n \to \infty} E_n$ por

$$\lim_{n \to \infty} E_n = \cap_{i=1}^{\infty} E_i, \text{ donde } E_n \supset E_{n+1}, n \geq 1.$$

Si $\{E_n,\ n \geq 1\}$ es cualquier sucesión creciente o decreciente de eventos entonces decimos que P es continua, si

$$\lim_{n \to \infty} \mathrm{P}(E_n) = \mathrm{P}\left(\lim_{n \to \infty} E_n\right).$$

1.3 Variables aleatorias.

En ocasiones, es posible o necesario cuantificar una característica especial del fenómeno aleatorio a estudiar. Es entonces conveniente calcular la probabilidad sobre el conjunto de eventos que satisfacen dicha propiedad a través de una variable aleatoria.

Definición 1.10 *Función medible.*

Sea (S, $\mathcal{A}$) *un espacio medible. Se dice que una función X:S→ $\mathcal{R}$ es medible si*

$$\{w/w \in \mathcal{A},\ X(w) \leq c\} = X^{-1}(-\infty, c] \in \mathcal{A} \quad \forall\ c \in \mathcal{R}.$$

Definición 1.11 *Variable aleatoria.*

Sea $(X, \mathcal{A}, \lambda)$ *un espacio de probabilidad. Una variable aleatoria X:S→ $\mathcal{R}$ es cualquier función medible.*

Para conjuntos numerables cualquier función X:S→ $\mathcal{R}$ es una variable aleatoria.

Teorema 1.2 *Sea* (S, $\mathcal{A}$, P) *un espacio de probabilidad y sea X:S→ $\mathcal{R}$. X es medible si y sólo si* $X^{-1}(B) \in \mathcal{A}$ *para cada conjunto de Borel,* $\mathcal{B} \in \mathcal{R}$*, es decir* $X^{-1}(\mathcal{B}) \in \mathcal{A}$.

La *función de distribución* F de la variable aleatoria X está definida para todo número real x por

$$F(x) = P\{X \leq x\}.$$

Denotaremos a 1-F(x) por F$^c(x)$, y también F$^c(x) = $ P$\{X > x\}$.

Una variable aleatoria X es discreta si el número de valores que puede tomar es contable (finito o infinito), y si éstos pueden arreglarse en una secuencia que corresponde con los entero positivos. [1].

Sea X una variable aleatoria discreta. La función f definida sobre $\mathcal{R}$ por

$$f(x) = P\{X = x\}$$

es llamada la *función de densidad* o la *función de densidad de probabilidad de X.*

Una variable aleatoria X es llamada (absolutamente) continua si existe una función $f(x)$, llamada la *función de densidad*, tal que

$$F(x) = P\{X \leq x\} = \int_{-\infty}^{x} f(t)\,dt,$$

donde la integral es la integral de Riemann.

Si X es una variable aleatoria continua en el intervalo $(-\infty, \infty)$ su función de densidad de probabilidad (F.D.P.) $f(x)$ debe satisfacer las dos condiciones siguientes:

$$1.\ \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

$$2.\ f(x) \geq 0, -\infty < x < \infty.$$

Si x es una variable aleatoria discreta, su F.D.P. cumple con las siguientes propiedades:

$$1.\ \sum f(x_i) = 1,$$

donde los x_i son los valores donde la densidad es estrictamente positiva $f(x_i) > 0$.

$$2.\ f(x) \geq 0\ \forall x \in \mathcal{R}.$$

La función de distribución de una variable aleatoria cumple con las siguientes propiedades:

1. $\lim_{x \to \infty} F(x) = 1.$

2. $\lim_{x \to -\infty} F(x) = 0.$

3. F(x) es no decreciente $\forall x \in \mathcal{R}$.

4. F(x) es continua por la derecha.

5. Para cualesquiera $a < b \in \mathcal{R}$ se tiene que $P[a < x \leq b] = F(b) - F(a)$ [7].

1.3.1 Función de distribución y función de densidad conjunta.

Cuando lo que interesa en un fenómeno aleatorio no es una sino varias de las características y éstas se pueden cuantificar, entonces en lugar de una variable aleatoria, se considera un vector aleatorio y se calculan las probabilidades sobre este conjunto.

Las notaciones $f(x_1, x_2, \ldots, x_n)$ y $P(x_1, x_2, \ldots, x_n)$ se utilizan para describir la F.D.P. conjunta de n variables aleatorias en los casos continuo y discreto. Estas funciones deben satisfacer las condiciones siguientes:

caso continuo

$$f(x_1, x_2, \ldots, x_n) \geq 0 \; para \; -\infty < x_i < \infty, \; i = 1, 2, \ldots, n, \text{ y}$$

$$\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 \ldots \int_{-\infty}^{+\infty} dx_n f(x_1, x_2, \ldots, x_n) = 1;$$

caso discreto

$$P(x_1, x_2, \ldots, x_n) \geq 0 \; para \; toda \; x_i, \; i = 1, 2, \ldots, n, \text{ y}$$

$$\sum_{toda\ x_1} \sum_{toda\ x_2} \cdots \sum_{toda\ x_n} P(x_1, x_2, \ldots, x_n) = 1.$$

La *función de distribución conjunta* F de dos variables aleatorias X y Y esta definida por

$$F(x, \; y) = P\{X \leq x, \; Y \leq y\}.$$

Las funciones de distribución de X y Y

$$F_X(x) = P(X \leq x) \quad \text{y} \quad F_Y(y) = P(Y \leq y)\,.$$

pueden ser obtenidas de F(x, y) haciendo uso de las propiedades de continuidad. Específicamente, sea $\{y_n\}$, $n \geq 1$, una sucesión creciente convergente a ∞. Entonces como los eventos $\{X \leq x, \; Y \leq y_n\}$, $n \geq 1$, son crecientes y

$$\lim_{n \to \infty} P\{X \leq x, \; Y \leq y_n\} = P \bigcup_{n=1}^{\infty} \{X \leq x, \; Y \leq y_n\} = P\{X \leq x\}$$

esto se sigue de la propiedad de continuidad, ya que

$$\lim_{n\to\infty} P\{X \le x,\ Y \le y_n\} = P\{X \le x\}$$

o, equivalentemente

$$F_X(x) = \lim_{Y\to\infty} F(x,\ y)$$

similarmente

$$F_Y(y) = \lim_{X\to\infty} F(x,\ y).$$

Las variables aleatorias X y Y se dicen (absolutamente) *continuas conjuntamente* si existe una función $f(x,\ y)$, llamada la función de densidad (de probabilidad) conjunta, tal que

$$P\{X \text{ está en } A,\ Y \text{ está en } B\} = \iint_{AB} f(x,\ y)\,dydx$$

para todos los conjuntos A y B.

Similarmente la distribución conjunta de una colección $X_1,\ X_2, \ldots,\ X_n$ de variables aleatorias está definida por

$$F(x_1,\ x_2, \ldots, x_n) = P\{X_1 \le x_1,\ X_2 \le x_2, \ldots, X_n \le x_n\}\ \forall (x_1,\ x_2, \ldots, x_n) \in R^n.$$

1.4 Probabilidad condicional e independencia.

La probabilidad condicional de que ocurra B, dado que haya ocurrido A, (que se denota como $P(B/A)$), es $P(B/A) = P(B \cap A)/P(A)$, si $P(A) > 0$ [6].

Teorema 1.3 *Bayes-Laplace [6]*
Supongamos que se tienen k eventos $A_1,\ A_2, \ldots,\ A_k$ tales que

$$\begin{array}{ll} 1.- & A_1 \cup A_2 \cup \ldots \cup A_k = S \\ 2.- & A_i \cap A_j = \emptyset \quad \forall\ i \ne j \end{array}$$

(estos eventos forman una partición de **S**); entonces para cualquier evento $E \subset S$,

$$P(A_j/E) = \frac{P(A_j)\,P(E/A_j)}{\sum_{i=1}^{k} P(A_j)\,P(E/A_i)}, \qquad j = 1, 2, \ldots, k.$$

Sea X_1, $X_2, \ldots$, X_n un conjunto de variables aleatorias, entonces la función de densidad condicional de X_1, $X_2, \ldots$, X_m dado $X_{m+1}, X_{m+2}, \ldots, X_n$ se define como:

$$f_{X_1,X_2,\ldots,X_m/X_{m+1},X_{m+2},\ldots,X_n}\left(x_1, x_2, \ldots, x_m/x_{m+1}, x_{m+2}, \ldots, x_n\right)$$

$$= P\left[\bigcap_{i=1}^{m}\{X_i = x_1\} \diagup \bigcap_{i=m+1}^{n}\{X_i = x_1\}\right], \text{ si } P\left[\bigcap_{i=m+1}^{n}\{X_i = x_1\}\right] > 0.$$

1.4.1 Noción de independencia.

Cuando en un fenómeno aleatorio, la ocurrencia (o no) de un evento A no influye en la ocurrencia (o no) de un evento B, se dice que los eventos A y B son independientes entre sí.

Dos eventos A y B son independientes si y sólo si $P(A \cap B) = P(A)P(B)$ [6].

A y B son mutuamente excluyentes si y sólo si $A \cap B = \emptyset$ [6].

A, B y C son independientes, si y sólo si

1. $P(A \cap B) = P(A)P(B)$,
2. $P(A \cap C) = P(A)P(C)$,
3. $P(B \cap C) = P(B)P(C)$,
4. $P(A \cap B \cap C) = P(A)P(B)P(C)$ [6].

En general, sea C una colección de eventos. Los eventos en C se dice que son independientes si la probabilidad de la intersección de cualquier número finito de ellos es igual al producto de sus probabilidades.

Sea X_1, $X_2, \ldots$, X_n, un conjunto de variables aleatorias. Se dice que estas variables aleatorias son independientes, sí para todas las posibles secciones de pares de números reales $(a_1,\ b_1)$, $(a_2,\ b_2), \ldots$, $(a_n,\ b_n)$, con $a_i \leq b_i\ \forall i$, los sucesos $[a_1 \leq X_1 \leq b_1]$, $[a_2 \leq X_2 \leq b_2]$, ..., $[a_n \leq X_n \leq b_n]$ son independientes.

Las variables aleatorias X y Y son *independientes* si y sólo si

$$F(x,\ y) = F_X(x) \cdot F_Y(y) \text{ para toda } x \text{ y } y.$$

Similarmente, las n variables aleatorias son independientes si y sólo si

$$F(x_1,\ x_2, \ldots, x_n) = F_{X_1}(x_1)\, F_{X_2}(x_2) \ldots F_{X_n}(x_n),$$

donde $F_{X_i}(x_i) = \lim\limits_{x_j \to \infty} F(x_1,\ x_2, \ldots, x_n)$, $j \neq i$; es la distribución marginal de x_i.

1.5 Valor esperado.

A menudo es deseable resumir las características de una distribución de probabilidad por medidas que tengan sentido, de las cuales puedan sacarse conclusiones generales sobre la variable aleatoria. Estas medidas usualmente se especifican como el *valor esperado* de ciertas funciones de la variable aleatoria.

La esperanza o media de una variable aleatoria X, denotada por $E(X)$, esta definida por

$$E[X] = \int_{-\infty}^{\infty} xdF(x),$$
$$\begin{cases} \int_{-\infty}^{\infty} xf(x)\,dx & \text{si } X \text{ es continua} \\ \sum_{x} xP[X=x] & \text{si } X \text{ es discreta} \end{cases} \tag{1.1}$$

A condición de que la integral de arriba exista.

La ecuación (1.1) también define la esperanza de cualquier función de X, digamos $h(X)$.

Donde $h(X)$ es en si misma una variable aleatoria, se sigue de (1.1) que

$$E[h(X)] = \int_{-\infty}^{\infty} xdF_h(x)$$

donde F_h es la función de distribución de $h(X)$.

Sin embargo, puede demostrarse que esto es idéntico a $\int_{-\infty}^{\infty} h(x)\,dF(x)$.

Esto es

$$E[h(X)] = \int_{-\infty}^{\infty} h(x)\,dF(x). \tag{1.2}$$

Esta ecuación es a veces conocida como la *ley del estadístico inconsciente* (desde que los estadísticos habían sido acusados de usar la cantidad (1.2) sin percatarse que esto no es una definición) [16].

La varianza de una variable aleatoria X está definida por

$$VarX = E[(X - E[X])^2] = E(X^2) - E^2(X).$$

La distribución conjunta de dos variables aleatorias X y Y se dicen ser no correlacionadas si su covarianza, definida por

$$Cov(X,Y) = E[(X - EX)(Y - EY)] = E[XY] - E(X)E(Y)$$

es cero. Se deduce que variables aleatorias independientes, son no correlacionadas. Sin embargo, lo contrario no necesita ser verdadero, excepto para las variables con distribución normal.

Una propiedad importante de esperanzas es que la esperanza de la suma de variables aleatorias es igual a la suma de las esperanzas

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E\left[X_i\right].$$

La correspondiente propiedad para la varianza es tal

$$Var\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} Var\left(X_i\right) + 2\sum_{i<j} Cov\left(X_i,\ X_j\right).$$

1.6 Función generadora de momentos y característica.

La función generadora de momentos de X esta definida por

$$\psi\left(t\right) = E\left[e^{tX}\right].$$

Todos los momentos de X pueden ser obtenidos sucesivamente por diferenciación de ψ y evaluando en $t = 0$, (siempre y cuando las derivadas existan en una vecindad del cero). Es decir,

$$\begin{aligned}
\psi'\left(t\right) &= E\left[Xe^{tX}\right], \\
\psi''\left(t\right) &= E\left[X^2 e^{tX}\right], \\
&\vdots \\
\psi^n\left(t\right) &= E\left[X^n e^{tX}\right],
\end{aligned}$$

$$\begin{aligned}
\text{evaluando en } t &= 0 \\
\psi^n\left(0\right) &= E\left[X^n\right],\ n \geq 1.
\end{aligned}$$

Cuando existe una función generadora de momentos, bajo en condiciones muy generales, ella sola determina la distribución. Esto es muy importante porque nos permite caracterizar la distribución de probabilidad de una variable aleatoria por su función generadora.

Puesto que la función generadora de momentos de una variable aleatoria X no necesariamente existe, es teóricamente conveniente definir la *función característica* de X por

$$\phi(t) = E\left[e^{itX}\right], \qquad -\infty < t < \infty,$$

$$\text{donde } i = \sqrt{-1}.$$

Se puede demostrar que ϕ siempre existe y, que ésta determina la distribución de X bajo condiciones no restrictivas [16].

Similarmente, podemos definir también la generadora de momentos conjunta de las variables aleatorias $X_1, X_2, \ldots, X_n$ por

$$\psi(t_1, \ \ldots, \ t_n) = E\left[\exp\left\{\sum_{j=1}^{n} t_j X_j\right\}\right],$$

y la función característica conjunta por

$$\phi(t_1, \ \ldots, \ t_n) = E\left[\exp\left\{\sum_{j=1}^{n} t_j X_j\right\}\right].$$

1.7 Esperanza condicional.

Si X y Y son variables aleatorias discretas, la función de densidad de probabilidad de X, dado Y=y, está definida, para toda y tal que $P\{Y = y\}>0$, por

$$f(X/y) = P\{X = x/Y = y\} = \frac{P\{X = x, \ Y = y\}}{P\{Y = y\}}.$$

Si X y Y tienen una función de densidad de probabilidad conjunta $f(x, y)$, la función de densidad de probabilidad condicional de X, dado $Y = y$, está definida para toda y tal que $f_Y(y) \geq 0$ por

$$f(x/y) = \frac{f(x, \ y)}{f_Y(y)}.$$

Sean X y Y variables aleatorias con función de densidad $f_X(x)$ y $f_Y(y)$, respectivamente. Si la función de densidad condicional de X dado $Y = y$ (donde $f_Y(y) \neq 0$) está dada por $f(x/Y = y)$ entonces X y Y son variables aleatorias independientes si y sólo si $f_X(x) = f(x/Y = y)$ [7]

La función de distribución condicional de X dado $Y = y$ está definida por

$$F(x/y) = P\{X \leq x/Y = y\} = \int_{-\infty}^{x} f(x/y)\,dx, \quad \text{caso continuo,}$$

$$F(x/y) = P\{X \leq x/Y = y\} = \sum_{x_i \leq x} f(x_i/y), \quad \text{caso discreto,}$$

donde los x_i son los valores con probabilidad positiva.

Y la esperanza condicional de X dado $Y = y$, por

$$E[X/Y = y] = \int x dF(x/y), \quad \text{caso continuo,}$$

$$E[X/Y = y] = \sum_{x} xP\{X = x/Y = y\}, \quad \text{caso discreto.}$$

Si X y Y tienen una función de densidad de probabilidad conjunta $f(x, y)$, la función de densidad de probabilidad condicional de X, dado $Y = y$, está definida para toda y tal que $f_Y(y) > 0$ por

$$f(x/y) = \frac{f(x,\ y)}{f_Y(y)}.$$

Por consiguiente todas las definiciones son exactamente como en el caso incondicional excepto que todas las probabilidades son ahora condicionales sobre el evento que $Y = y$.

Denotemos por $E[X/Y]$ a la función de la variable aleatoria Y cuyo valor $Y = y$ es $E[X/Y = y]$. Una propiedad extremadamente útil de la esperanza condicional es que para todas las variables aleatorias X y Y se cumple que

$$E[X] = E[E[X/Y]] = \int E[X/Y = y]\,dF_Y(y) \qquad (1.3)$$

cuando la esperanza existe [16].

Si Y es una variable aleatoria discreta, entonces tenemos de (1.3)

$$E[X]=\sum_{y}E[X/Y=y]P\{Y=y\}.$$

Mientras que si Y es continua con densidad $f(y)$, entonces la ecuación (1.3) dice

$$E[X]=\int_{-\infty}^{\infty}E[X/Y=y]f(y)\,dy.$$

La ecuación (1.3) muestra que $E[X]$ es un peso promedio del valor esperado condicional de X dado que $Y=y$, cada uno de los términos $E[X/Y=y]$ inicia con un peso por la probabilidad del evento sobre el cual esta condicionado.

1.8 La distribución exponencial, pérdida de la memoria, y funciones de tasa de riesgo.

Se dice que una variable aleatoria continua X tiene una distribución exponencial con parámetro λ, $\lambda>0$, si la función de densidad de probabilidad está dada por

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & \text{si } x\geq 0\\ 0, & \text{si } x<0\end{cases}.$$

O, equivalentemente, si su distribución es

$$F(x)=\int_{-\infty}^{x}f(y)\,dy=\begin{cases}1-\lambda e^{-\lambda x}, & \text{si } x\geq 0\\ 0, & \text{si } x<0\end{cases}.$$

Generalmente esta distribución es muy usada en Teoría de Colas.

La función generadora de momentos de la distribución exponencial está dada por

$$E\left[e^{tX}\right]=\int_{0}^{\infty}e^{tx}\lambda e^{-\lambda x}dx=\frac{\lambda}{\lambda-t}. \tag{1.4}$$

Todos los momentos de x pueden ser obtenidos por diferenciación de (1.4) , se tiene que

$$E[X] = 1/\lambda, \quad Var(X) = 1/\lambda^2.$$

La utilidad de la variable aleatoria exponencial deriva del hecho que ella posee, la propiedad de pérdida de la memoria (sin memoria, o carente de memoria), que significa

$$P\{X > t + s / X > t\} = P\{X > s\}, \text{ para } s,\ t \geq 0. \tag{1.5}$$

Si pensamos a X como la vida de algún instrumento, entonces (1.5) es la probabilidad de que la vida del instrumento dure por lo menos $s + t$ horas, dado que ha sobrevivido t horas; esto es, como la probabilidad inicial de que al menos viva s horas. En otras palabras, si el instrumento está vivo en el tiempo t, entonces la distribución de su vida restante es la distribución de su vida original. La condición (1.5) es equivalente a

$$\vec{F}(s+t) = \vec{F}(s)\,\vec{F}(t).$$

Esto se satisface cuando F es la exponencial, y vemos que tales variables aleatorias son desmemoriadas.

Como dato adicional, existe una relación importante entre la distribución de Poisson (a la cual se le dedica un capítulo) y la distribución exponencial. Si la distribución de Poisson describe el número de fracasos por unidad de tiempo entre dos fracasos sucesivos entonces la distribución exponencial puede obtenerse de la distribución de Poisson [16].

1.9 Algunas distribuciones típicas.

1.9.1 Variables aleatorias binomal y Bernoulli.

Definición 1.12 *Una prueba de Bernoulli es un experimento que tiene solamente dos resultados posibles, a los cuales con frecuencia se les llama éxito y fracaso.*

En general, se denotará por $S = \{e, f\}$ al espacio muestral para una prueba de Bernoulli. Por ejemplo: si se lanza al aire una sola moneda (sol o águila), el vuelo de un proyectil (si sólamente se dice que es un éxito o no), etcétera. Se puede considerar como prueba de Bernoulli a cualquier mecanismo de azar cuyos resultados se puedan agrupar en dos clases ajenas.

Una notación que se emplea con frecuencia es

$$P(\{e\}) = p$$
$$P(\{f\}) = q = 1 - p,$$

desde luego, la cantidad p puede tomar cualquier valor en el intervalo de 0 a 1 inclusive, para varios tipos de pruebas de Bernoulli. Si se supone que se tiene un experimento formado por la repetición de n pruebas independientes de Bernoulli, entonces el espacio muestral para este experimento es el producto Cartesiano de los espacios muestrales de las pruebas individuales, expresado como $S = S_1 \times S_2 \times \cdots \times S_n$, en que $S_i = \{e, f\}$, $i = 1, 2, \ldots, n$ y $P_i(\{e\}) = p$ para toda i. Esta es la variable aleatoria binomial y se define como sigue.

Definición 1.13 *Sea X el número total de éxitos en las n pruebas independientes y repetidas de Bernoulli con probabilidad p de éxito en una prueba dada. A X se le llama variable aleatoria* binomial *con parámetros n y p.*

El rango de la variable aleatoria X lo constituyen los enteros 0, 1, 2, ... n; por tanto, X es una variable aleatoria discreta.

Teorema 1.4 *Si X es una binomial con parámetros n y p, entonces*

$$p_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x}, & x = 0, 1, 2, \ldots, n, \\ 0, & \textit{en los demás casos.} \end{cases}$$

$p_X(k)$ es una función de probabilidad ya que

$$p_X(k) \geq 0, \qquad \text{para toda } k.$$

La suma de $p_X(k)$ en el rango de X es

$$\begin{aligned} \sum_{k=0}^{n} p_X(k) &= \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \\ &= (p+q)^n \\ &= (p+1-p)^n = 1. \end{aligned}$$

Así observamos que la variable aleatoria de Bernoulli es un caso especial de la binomial cuando $n = 1$. La función generadora de momentos es

$$\varphi_X(t) = q + e^t p,$$

y tiene $\mu_X = p$, $\sigma_X^2 = pq$ [6].

1.9.2 Variables aleatorias geométrica e hipergeométrica.

Definición 1.14 *Se realizan pruebas sucesivas de Bernoulli, independientes hasta obtener un éxito. La probabilidad de éxito en cada prueba es p, con $0 < p \leq 1$. Sea Y el número de pruebas necesarias (para obtener el primer éxito). A Y se le denomina la variable aleatoria* geométrica *con parámetro p.*

El rango de Y es el conjunto de enteros positivos, ya que el número de pruebas necesarias puede ser igual a cualquiera de esos valores. Consecuentemente, Y también es una variable aleatoria discreta.

Teorema 1.5 *Si Y es una variable aleatoria geométrica con parámetro p, entonces*

$$p_Y(k) = \begin{cases} q^{k-1}p, & k = 1, 2, 3, \ldots \\ 0, & \text{para los demás casos.} \end{cases}$$

A continuación mencionaremos la función generadora de momentos de una variable aleatoria geométrica y las primeras dos derivadas de ésta, así como su media y varianza:

$$\varphi_Y(t) = \frac{pe^t}{1 - qt},$$

$$\varphi_Y^{(1)}(t) = \frac{pe^t}{(1 - qt)^2},$$

$$\varphi_Y^{(2)}(t) = \frac{pe^t(1 + qe^t)}{(1 - qt)^3}.$$

Media de la variable aleatoria geométrica:

$$\mu_Y = \frac{1}{p}.$$

Y la varianza esta dada por:

$$\sigma_Y^2 = \frac{q}{p^2}.$$

En seguida se estudiará la distribución hipergeométrica. Esta variable aleatoria obtiene su nombre del hecho de que su función de probabilidad toma valores en una serie hipergeométrica. La variable aleatoria hipergeométrica es semejante a una variable aleatoria binomial, pero con variables Bernoulli que no son independientes.

Definición 1.15 *En una población de M ratas, de las cuales W son blancas. Definir Z como el número de ratas blancas que salen en una muestra de n ratas tomadas sin reemplazo y al azar de la población. A Z se le denomina la variable aleatoria* hipergeométrica.

Ya que el número de ratas blancas sacadas sólo toma valores enteros, Z es otro ejemplo de una variable aleatoria discreta. Si el muestreo de la población se hubiera hecho con reemplazo en vez de sin reemplazo, Z sería una variable aleatoria binomial con parámetros n y $p = W/M$; ésta es la similitud entre las variables aleatorias binomial e hipergeométrica señalada arriba. En seguida se proporciona la función de probabilidad para Z.

Teorema 1.6 *Si Z es la variable aleatoria dada en la definición 1.15, entonces*

$$p_Z(k) = \frac{\binom{W}{k}\binom{M-W}{n-k}}{\binom{M}{n}}, \qquad k = 0, 1, 2, \ldots, n,$$

en que se usa la convención $\binom{b}{a} = 0$ para $a > b$ [6].

1.9.3 La variable aleatoria de Poisson.

Llamada así en honor de Siméon Denis Poisson probabilista francés del siglo XIX, quien fue el primero en describirla, es otra distribución discreta de probabilidad muy útil en la que la variable aleatoria representa el número de eventos independientes por unidad que ocurren a una tasa constante. Muchos eventos aleatorios ocurren de manera independiente con una velocidad constante en el tiempo o en el espacio. Algunos ejemplos típicos son el número de personas que llegan a una tienda de autoservicio en un tiempo determinado, el número de defectos en piezas similares para el material, el número de bacterias en un cultivo, el número de solicitudes de seguro procesadas por una compañía en un periodo específico, etc. De hecho, la distribución de Poisson es el principal modelo de probabilidad empleado para analizar problemas de líneas de espera. Además, ofrece una aproximación excelente a la función de probabilidad binomial cuando p es pequeño y n grande.

Definición 1.16 *Sea X una variable aleatoria que representa el número de eventos aleatorios independientes que ocurren a una rapidez constante sobre el tiempo o el espacio. Se dice entonces que la variable aleatoria X tiene una distribución de Poisson con función de probabilidad:*

$$p_X(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda}, & k = 0, 1, 2, \ldots \\ 0, & \text{en los demás casos.} \end{cases}$$

Se comprueba que $p_X(k)$ es una función de probabilidad. En primer lugar,

$$\frac{(\lambda)^k}{k!} e^{-\lambda} > 0$$

para todos los valores de k de manera que $p_X(k) \geq 0$. En segundo lugar,

$$\begin{aligned} \sum_{k=0}^{\infty} p_X(k) &= \sum_{k=0}^{\infty} \frac{(\lambda)^k}{k!} e^{-\lambda} \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda)^k}{k!} \\ &= e^{-\lambda} e^{\lambda} = 1. \end{aligned}$$

Por tanto $p_X(k)$ satisface los requerimientos de una función de probabilidad.

1.9.4 Variable aleatoria uniforme.

La variable aleatoria continua más sencilla posible se llama variable aleatoria uniforme, misma que se define como sigue.

Definición 1.17 *X es una variable aleatoria uniforme en el intervalo (a, b) si*

1. El rango de X es el intervalo (a, b).
2. $f_X(x)$ es constante para $x \in R_X$.

(Se dice que X está distribuida uniformemente en (a, b)).

Teorema 1.7 *Si X está distribuida uniformemente en (a, b), entonces,*

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \textit{en los demás casos.} \end{cases}$$

La variable aleatoria uniforme deriva su nombre de que su función de densidad es uniforme (constante) en el intervalo (a,b) en el sentido de que conjuntos contenidos en (a,b) con la misma longitud tienen la misma probabilidad. La función de distribución para una variable aleatoria uniforme es

$$\begin{aligned} F_X(t) &= \int_{-\infty}^{t} f_X(x)\,dx \\ &= 0, \qquad t<a \\ &= \frac{t-a}{b-a}, \qquad a \le t \le b \\ &= 1, \qquad t>b; \end{aligned}$$

por lo que F_X crece en forma lineal en el intervalo (a,b).

Suponga que X está distribuida uniformemente en (a,b). Entonces

$$\mu_X = E[X] = \int_a^b \frac{x}{b-a}dx = \frac{b+a}{2},$$

$$E\left[X^2\right] = \int_a^b \frac{x^2}{b-a}dx = \frac{b^2+ab+a^2}{3},$$

de manera que

$$\sigma_X^2 = E\left[X^2\right] - \mu_X^2 = \frac{(b-a)^2}{12}.$$

Capítulo 2

Procesos estocásticos.

Por proceso estocástico se entiende una familia de variables aleatorias $\{X_t\}$, donde t es un punto en un espacio T, llamado espacio parametral, y donde para cada $t \in T$, X_t es un punto en un espacio de probabilidad S, llamado espacio de estados. Es decir, los valores que puede tomar X_t son llamados sus estados y los cambios en el valor de X_t reciben el nombre de transiciones entre sus estados.

Se puede imaginar la familia $\{X_t\}$ como la trayectoria de una partícula que se mueve "al azar" en el espacio S, siendo X_t su posición en el instante t. Un registro de estas trayectorias se conoce como realización del proceso.

Las leyes del proceso aleatorio $\{X_t\}$, $t \in T$, se determinan por las distribuciones conjuntas de probabilidad de sus valores $\{X_{t_1}\}, ...\{X_{t_n}\}$ para los distintos $t_1, t_2, ..., t_n$ (ellas se llaman distribuciones de dimensión finita del proceso estocástico dado).

En este capítulo expondremos los conceptos básicos de los procesos estocásticos. Comenzamos por el ejemplo de la caminata aleatoria y el de la ruina del jugador. Después abordamos las cadenas de Markov y mencionamos las ecuaciones de Chapman-Kolmogorov, la clasificación de los estados de la cadena, las probabilidades límites y el concepto de tiempo de paro. Terminamos con un resumen del proceso Poisson. Algunas referencias para este capítulo son: [3] y [17].

2.1 Caminata aleatoria.

2.1.1 Primer paso.

Una fuente muy importante de la teoría de los procesos estocásticos es la caminata aleatoria simple. Ya un rico inesperado y elegante fenómeno, la caminata aleatoria también guía inexorablemente al desarrollo del movimiento Browniano, la teoría de la difusión y a innumerables aplicaciones importantes en finanzas, economía, y ciencias físicas, entre otras.

La caminata aleatoria simple provee un modelo para la acumulación de la riqueza de una persona, quien hace lanzar una moneda y apuesta una cantidad fija en cada volado. Sea $\{X_i : 1 \leq i \leq \infty\}$ la sucesión de variables aleatorias independientes que tiene distribución de probabilidad dada por

$$P(X_i = 1) = P(X_i = -1) = \frac{1}{2}.$$

En lo que sigue, sea S_0 un entero arbitario fijo y que representa la riqueza con la que inicia el jugador, y para $1 \leq n \leq \infty$, sea S_n la suma de S_0 más la suma parcial de las X_i :

$$S_n = S_0 + X_1 + X_2 + \cdots + X_n.$$

Donde $S_n - S_0$ son las ganancias netas después de n apuestas justas de un peso cada una, así nos preguntamos ¿cuál será la probabilidad de que el jugador gane A pesos antes de perder B pesos? ponemos esta pregunta en una notación útil. Consideremos favorable el primer tiempo τ en el cual la suma parcial de S_n alcanza el nivel A o el nivel $-B$:

$$\tau = \min\{n \geq 0 : S_n = A \text{ o } S_n = -B\}.$$

En el tiempo aleatorio τ, tenemos $S_\tau = A$ o $S_\tau = -B$, así nuestro problema básico es determinar $P(S_\tau = A \mid S_0 = 0)$. Aquí, por supuesto, permitiremos a la riqueza del jugador idealizado llegar a ser negativa.

Análisis del primer paso.

La solución de este problema puede ser obtenido de varias maneras, pero quizás el método más general es el *análisis del primer paso.*

Para nuestro problema inmediato, el análisis del primer paso sugiere que consideremos la situación del jugador después del primer juego. Vemos que la riqueza crece o decrece en un peso. Consideremos la relación de recursión dada por la función

$$f(k) = P(S_\tau = A \mid S_0 = k), \text{ donde } -B \leq k \leq A.$$

En esta notación, $f(0)$ es precisamente la probabilidad deseada de ganar A pesos antes de perder B pesos.

Como consecuencia del primer paso, encontramos la recursión deseada por $f(k)$,

$$f(k) = \frac{1}{2}f(k-1) + \frac{1}{2}f(k+1) \text{ para } -B < k < A, \tag{2.1}$$

y esta recursión será determinada únicamente por f cuando se contemplan las restricciones

$$f(A) = 1 \text{ y } f(-B) = 0.$$

La solución no es de sorprendernos. Por ejemplo, si designamos $f(-B+1) = \alpha$ y substituimos los valores de $f(-B)$ y $f(-B+1)$ en la ecuación (2.1), encontramos que $f(-B+2) = 2\alpha$. Si entonces substituimos los valores de $f(-B+1)$ y $f(-B+2)$ en la ecuación (2.1) tenemos $f(-B+3) = 3\alpha$; por que $f(-B+k) = k\alpha$ para toda $0 \leq k \leq A+B$.

Finalmente, determinamos que $\alpha = 1/(A+B)$ de la condición $f(A) = 1$ y del hecho que para $k = A+B$ nuestra fórmula para f requiere $f(A) = (A+B)\alpha$. Con lo cual, llegamos a la fórmula:

$$P(S_n \text{ alcanzar } A \text{ antes de } -B \mid S_0 = 0) = \frac{B}{A+B}. \tag{2.2}$$

La caminata aleatoria simple se puede describir como sigue. Una partícula se mueve a lo largo de una línea por pasos aleatorios; cada paso es de una unidad esta puede ir a la derecha ó a la izquierda con probabilidades p y q=*1-p* respectivamente donde $0<p<1$. Supongamos que cada paso toma una unidad de tiempo así que el *n-ésimo* paso se ha producido instantáneamente en el tiempo n; además supongamos que las posibles posiciones de la partícula son el conjunto de todos los enteros sobre el eje coordenado. Este conjunto es frecuentemente referido como los "enteros enrejados" sobre $R\cup(-\infty,\infty)$ y será denotado por I. Así la partícula efectúa una caminata sobre la reja, atrás y adelante, y continua al infinito. Si marcamos su posición X_n como una función de tiempo n, su camino es una línea zigzag (una de las cuales se puede apreciar en la figura 2.1).

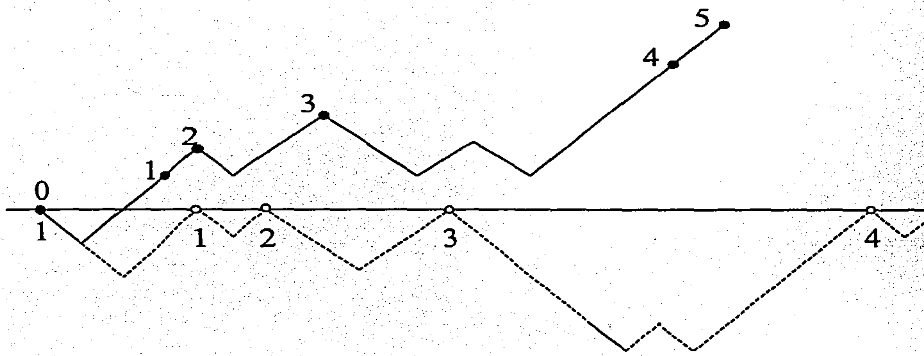

Figura 2.1:

En un lenguaje más coloquial imaginemos a la partícula como a un vagabundo o un borracho y la línea como una calle infinita dividida en cuadras. En cada unidad de tiempo digamos 5 minutos, el camina una cuadra desde la esquina de una calle a otra esquina, en cada esquina el puede escoger ir adelante ó regresar con probabilidades p ó q. Entonces está tomando un camino aleatorio y su camino puede ser trazado en la calle con una parte de doblando y redoblando.

2.2 Problema de la ruina del jugador.

Consideremos al jugador que gana o pierde un peso con probabilidades respectivas p y q. Sea su capital inicial z y sea $a-z$ el capital inicial del jugador contrario, de manera que el capital combinado será a. El juego continúa hasta que el capital del jugador se reduce a cero o se incrementa hasta a, es decir, hasta que uno de los jugadores quede arruinado. Estamos interesados en la probabilidad de la ruina del jugador y en la distribución de probabilidades de la duración del juego. Este es el *problema clásico de la ruina.*

Tomemos en cuenta una partícula que parte de la posición inicial z y a intervalos de tiempo regulares, se mueve un paso unitario en dirección positiva o negativa, según sea el ensayo correspondiente un éxito ó un fracaso. La posición de la partícula después de n pasos representa el capital del jugador al terminar el $n-$ *ésimo* ensayo. Los ensayos terminan cuando la partícula alcanza por primera vez 0 o a, lo cual se puede describir si decimos que la partícula realiza una caminata aleatoria con barreras absorbentes en 0 y a. Esta caminata aleatoria se restringe a las posiciones posibles 1, 2, ...,$a-1$.

Sea q_z la probabilidad de que al final[1] el jugador se arruine si actualmente cuenta con una riqueza de z pesos, y sea p_z la probabilidad de que gane el juego. En la terminología de las caminatas aleatorias, q_z y p_z son las probabilidades respectivas de que una partícula que comienza en z sea absorbida en 0 o a. Demostraremos que $p_z + q_z = 1$.

Después del primer ensayo, la fortuna del jugador es $z-1$ o $z+1$ y, por lo tanto, debemos tener

$$q_z = pq_{z+1} + qq_{z-1} \tag{2.3}$$

a condición de que $1 < z < a-1$. Si $z = 1$, el primer ensayo puede conducir a la ruina, y (2.3) se reemplaza por $q_1 = pq_2 + q$. De manera parecida, para $z = a-1$, el primer ensayo puede conducir a la victoria, y, por lo tanto, $q_{a-1} = qq_{a-2}$ para unificar estas ecuaciones, definimos

$$q_0 = 1, \qquad q_a = 0. \tag{2.4}$$

Con esta convención, la probabilidad q_z, de ruina, satisface (2.3) en z =1, 2, ...,$a-1$.

Los sistemas de la forma (2.3) se conocen como ecuaciones de diferencias, y (2.4) representa las condiciones de frontera sobre q_z. Ahora obtendremos una expresión explícita de q_z.

Supongamos en primer lugar que $p \neq q$. Se puede comprobar fácilmente que las ecuaciones de diferencias de (2.3) admiten las dos soluciones particulares $q_z = 1$ y $q_z = \left(\frac{q}{p}\right)^z$. Resulta que, para constantes arbitrarias A y B, la sucesión

$$q_z = A + B\left(\frac{q}{p}\right)^z \tag{2.5}$$

representa una solución formal de (2.3). Las condiciones de frontera (2.4) se cumplen si, y sólo si, A y B satisfacen las dos ecuaciones lineales $A + B = 1$

[1]En un sentido estricto, la probabilidad de ruina está definida en un espacio muestral de juegos que se prologan infinitamente, pero podemos trabajar con el espacio muestral de n ensayos. La probabilidad de ruina en menos de n ensayos se incrementa con n y tiene, un límite. Llamaremos a este límite "probabilidad de ruina".

y $A + B\,(q/p)^a = 0$. De este modo,

$$q_z = \frac{(q/p)^a - (q/p)^z}{(q/p)^a - 1} \tag{2.6}$$

es solución formal de la ecuación de diferencia (2.3) que cumple con las condiciones de frontera (2.4). Para demostrar que (2.6) es la probabilidad de ruina, sólo nos falta verificar que la solución es única, es decir que todas las soluciones de (2.3) son de la forma (2.5). Ahora bien, dada una solución arbitaria de (2.3), las dos constantes A y B pueden elegirse de manera que (2.5) coincida con ella en $z = 0$ y en $z = 1$. Con estos dos valores se obtienen todos los demás al substituir sucesivamente en (2.3) z =1, 2, 3,.... Por lo tanto, dos soluciones que coinciden en $z = 0$ y $z = 1$, son idénticas y, de este modo, cualquier solución tendrá la forma (2.5).

Es así como hemos demostrado que la probablidad requerida de la ruina del jugador está dada por (2.6), si $p \neq q$. La probabilidad p_z de que el jugador gane el juego es igual a la probabilidad de ruina del jugador contrario y, por consiguiente, se obtiene con las mismas fórmulas al reemplazar respectivamente, p, q y z por q, p y $a - z$. Se ve $p_z + q_z = 1$ [3].

A continuación, investigaremos el efecto de los cambios de apuesta. La probabilidad correspondiente de ruina q_z^* se obtiene de (2.6), si se reemplaza z por $2z$ y a por $2a$:

$$q_z^* = \frac{(q/p)^{2a} - (q/p)^{2z}}{(q/p)^{2a} - 1} = q_z \frac{(q/p)^a - (q/p)^z}{(q/p)^a + 1}. \tag{2.7}$$

Para $q > p$, la última fracción es mayor que la unidad y $q_z^* > q_z$. Enunciaremos esta conclusión de otra manera: si se duplican las apuestas, mientras que los capitales iniciales no se alteran, la probabilidad de ruina decrece para el jugador cuya probabilidad de éxito es $p < 1/2$, y se incrementa para el adversario.

El caso límite en que $a = \infty$ corresponde a un juego contra un adversario infinitamente rico. Si tomamos $a \to \infty$ en (2.6), obtenemos

$$q_z = \begin{cases} 1 & \text{si } p \leq q \\ (q/p)^z & \text{si } p > q. \end{cases} \tag{2.8}$$

Interpretamos a q_z como la probabilidad de ruina a largo plazo de un jugador con capital inicial z, que juega contra un adversario infinitamente rico. En la terminología de caminatas al azar, q_z es la probabilidad de que una partícula que parte de $z > 0$ alcance alguna vez el origen. Es más natural expresar este resultado de la siguiente manera: En una caminata al azar que empieza en el origen, la probabilidad de alcanzar alguna vez la posición $z > 0$, es igual a uno si $p \geq q$, y es igual a $(p/q)^z$ cuando $p < q$.

2.3 La duración esperada del juego.

Siguiendo con el problema anterior consideremos por dado el hecho de que la duración del juego tiene la esperanza finita D_z.

Si el primer ensayo resulta en éxito, el juego continúa como si la posición inicial hubiese sido $z + 1$. Por consiguiente, la esperanza condicional de la duración, si suponemos que hay éxito en el primer ensayo, es $D_{z+1} + 1$. Este argumento nos demuesta que la duración esperada D_z satisface la ecuación de diferencia

$$D_z = pD_{z+1} + qD_{z-1} + 1, \qquad 0 < z < a \tag{2.9}$$

con las condiciones de frontera

$$D_0 = 0, \qquad D_a = 0. \tag{2.10}$$

La aparición del término 1 hace que la ecuación de diferencia (2.9) sea no homogénea. Si $p \neq q$, entonces $D_Z = z/(q-p)$ es una solución formal de (2.9). La diferencia, Δ_z, de dos soluciones cualesquiera de (2.9), satisface las ecuaciones homogéneas $\Delta_z = p\Delta_{z+1} + q\Delta_{z-1}$, y ya sabemos que todas las soluciones de esta ecuación son de la forma $A + B(q/p)^z$. Resulta que, cuando $p \neq q$, todas las soluciones son de la forma

$$D_z = \frac{z}{q-p} + A + B\left(\frac{q}{p}\right)^z. \tag{2.11}$$

Las condiciones de frontera (2.10) requieren que

$$A + B = 0, \qquad A + B(q/p)^a = -a/(q-p).$$

Al resolver para A y B encontramos que

$$D_z = \frac{z}{q-p} - \frac{a}{q-p} \cdot \frac{1-(q/p)^z}{1-(q/p)^a}. \tag{2.12}$$

como se indicó al final de la sección anterior, podemos pasar al límite $a \to \infty$ y considerar el juego contra un adversario infinitamente rico. Cuando $p > q$ el juego puede continuar indefinidamente y, en este caso, la duración esperada es infinita. Cuando $p < q$, obtenemos que la duración esperada es $z\,(q-p)^{-1}$.

2.4 Cadenas de Markov.

Sea un proceso estocástico $\{X_n, n = 0, 1, 2, ...\}$ que toma valores sobre un conjunto numerable. A menos que se mencione lo contrario, este conjunto de posibles valores del proceso se denotará por el conjunto de enteros no negativos $\{0, 1, 2, ...\}$. Si $X_n = i$, entonces el proceso se dice estar en el estado i al tiempo n. Cuando el proceso está en el estado i, hay una probabilidad fija P_{ij} de que el proceso pasará al estado j. Es decir

$$P\{X_{n+1} = j/X_n = i,\ X_{n-1} = i_{n-1}, ..., X_1 = i_1,\ X_0 = i_0\} = P_{ij} \tag{2.13}$$

$$= P\{X_{n+1} = j/X_n = i\} \tag{2.14}$$

Para todos los estados i_0,$i_1, ..., i_{n-1}, i, j$ y toda $n \geq 0$.

Tal proceso estocástico es conocido como una cadena de Markov. La ecuación (2.13) puede ser interpretada como la probabilidad condicional de un estado futuro X_{n+1}, dado los estados pasados X_0, X_1, ..., X_{n-1} y el estado presente X_n, sea independiente de los estados pasados y sólo dependa del estado presente. A esta propiedad se le conoce como propiedad de Markov.

Por otra parte, si las probabilidades en (2.14) no dependen de n se dice que la *Cadena de Markov* es *homogénea en el tiempo* o *estacionaria*, en cuyo caso, a las probabilidades definidas por

$$P_{ij} := P\{X_{n+1} = j/X_n = i\},$$

se les llama *probabilidades de transición* en un paso.

De aquí al final de este capítulo consideraremos cadenas de Markov estacionarias y denotaremos por $\{P_{ij}\}$ a sus probabilidades de transición.

El valor P_{ij} representa la probabilidad que el proceso pueda, cuando se encuentra en el estado i, continuar al estado j. Dado que las probabilidades son no negativas y dado el proceso debe hacer una transición en algún estado, tenemos que

$$P_{ij} \geq 0, \qquad i,\ j \geq 0;$$
$$\sum_{j=0}^{\infty} P_{ij} = 1, \qquad i = 0,\ 1, \ldots$$

Sea $\mathbb{P}$ la matriz del primer-paso de probabilidades de transición P_{ij} así tenemos

$$\mathbb{P} = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \ldots \\ P_{10} & P_{11} & P_{12} & \ldots \\ \vdots & & & \\ P_{i0} & P_{i1} & P_{i2} & \ldots \end{bmatrix}.$$

2.4.1 Ecuaciones de Chapman-Kolmogorov.

Ahora definiremos, las probabilidades de transición en n-pasos, P_{ij}^n es la probabilidad que un proceso del estado i pasará al estado j después de n transiciones adicionales. Esto es

$$P_{ij}^n = \{X_{n+m} = j / X_m = i\} \qquad n \geq 0,\ i,\ j \geq 0$$

también definimos

$$P_{ij}^0 := \begin{cases} 1 \text{ si } i = j \\ 0 \text{ si } i \neq j \end{cases}.$$

Usando la Propiedad de Markov (2.14) y la hipótesis de estacionariedad, tenemos

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \qquad \text{para toda } n,\ m \geq 0 \text{ y toda } i,\ j. \tag{2.15}$$

A esta expresión se le conoce como las *Ecuaciónes de Chapman-Kolmogorov*, y parte de su importancia reside en que nos permite calcular de forma recursiva a las probabilidades de transición en $n-pasos$ partiendo de las probabilidades de transición en un paso.

Las ecuaciones son establecidas observando que

$$\begin{aligned}
P_{ij}^{n+m} &= P\{X_{n+m} = j / X_0 = i\} \\
&= \sum_{k=0}^{\infty} P\{X_{n+m} = j,\ X_n = k / X_0 = i\} \\
&= \sum_{k=0}^{\infty} P\{X_{n+m} = j /\ X_n = k,\ X_0 = i\} P\{X_0 = k / X_0 = i\} \\
&= \sum_{k=0}^{\infty} P_{kj}^{m} P_{ik}^{n}.
\end{aligned}$$

Sea $P^{(n)}$ la matriz de las probabilidades de transición de $n - pasos$ P_{ij}^n, entonces la ecuación (2.15) afirma que

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)}$$

donde el punto representa la multiplicación matricial.

Hasta ahora, todas las probabilidades que hemos considerado son probabilidades condicionales. Por ejemplo P_{ij}^n es la probabilidad que el estado al tiempo n es j dado que el estado inicial al tiempo 0 es i. Si la distribución incondicional del estado al tiempo n es requerida, es necesario especificar la distribución de probabilidad del estado inicial. Denotaremos esto por

$$\alpha_i \equiv P\{X_0 = i\} \text{ para } i \in E, \text{ además } \sum_{i=0}^{\infty} \alpha_i = 1.$$

Todas las probabilidades incondicionales pueden ser calculadas bajo la condición del estado inicial.

Esto es,

$$\begin{aligned}
P\{X_n = j\} &= \sum_{i=0}^{\infty} P\{X_n = j / X_0 = i\} P\{X_0 = i\} \\
&= \sum_{i=0}^{\infty} P_{ij}^{n} \alpha_i.
\end{aligned}$$

2.4.2 Clasificación de estados.

El estado j se dice ser *accesible* para el estado i sí P_{ij}^n>0 para alguna $n \geq$0. Notemos que esto implica que el estado j es accesible desde el estado i si y sólo si, estando en i, es posible que el proceso entre al estado j. Si j es no accesible desde i, entonces tenemos que

$$\begin{aligned} P\{\text{entre alguna vez a } j/\text{inicia en } i\} &= P\left\{\bigcup_{n=0}^{\infty}\{X_n = j\}/X_0 = i\right\} \\ &\leq \sum_{n=0}^{\infty} P\{X_n = j/X_0 = i\} \\ &= \sum_{n=0}^{\infty} P_{ij}^n = 0. \end{aligned}$$

Si dos estados i y j son accesibles mutuamente se dice que son comunicantes (o que se comunican), y escribimos $i \leftrightarrow j$.

Cualquier estado, es comunicante consigo mismo, por definición

$$P_{ii}^0 = \{X_0 = i/X_0 = i\} = 1.$$

La relación de comunicación satisface las siguientes tres propiedades:

(i) El estado i se comunica con el estado i, $\quad \forall i \geq$0.

(ii) Si el estado i se comunica con el estado j, entonces el estado j se comunica con el estado i.

(iii) Si el estado i se comunica con el estado j, y el estado j se comunica con el estado k, entonces el estado i se comunica con el estado k [17].

Dos estados que se comunican se dicen estar en la misma *clase.*

Cualesquiera 2 clases son idénticas o ajenas. En otras palabras, el concepto de comunicación divide el espacio de estados en clases separadas. La cadena de Markov se dice que es *irreducible* si hay sólo una clase, esto es, si todos los estados se comunican con cada uno de los otros estados.

Para algún estado i denotamos a f_i como la probabilidad, iniciada en el estado i, de que el proceso siempre regresará al estado i. El estado i se dice ser *recurrente* si f_i=1 y *transitorio* si f_i<1.

Suponga que el proceso inicia en el estado i, el cual es recurrente. De donde, con probabilidad 1, el proceso eventualmente regresará a i. Sin embargo, por la definición de una cadena estacionaria de Markov, se sigue que el proceso será iniciado otra vez cuándo el estado i reinicie y, por lo tanto, el estado i eventualmente será visitado otra vez. Una repetición continua de este argumento guía a la conclusión que si el estado i es recurrente entonces, iniciando en el estado i, el proceso reiniciará en el estado i de nuevo y otra vez y otra vez; en realidad, infinitamente.

Ahora, supongamos que el estado i es transitorio. Es decir, en cada momento el proceso entrará al estado i con probabilidad 1-f_i, y entonces (en ese caso) nunca se iniciará otra vez al estado. Por lo tanto, iniciando en el estado i, la probabilidad de que el proceso llegue a estar en el estado i exactamente al tiempo n del período es igual a $f_i^{n-1}(1\text{-}f_i)$, $n \geq 1$. En otras palabras, si el estado i es transitorio entonces, comenzando en el estado i, el número de período en el que el proceso estará en el estado i tiene una distribución geométrica con media finita $1/(1\text{-}f_i)$.

De lo mencionado anteriormente, se sigue que el estado i es recurrente si y sólo si, iniciando en el estado i, el número esperado de períodos al tiempo que el proceso esta en el estado i es infinito. Sea

$$A_n = \begin{cases} 1, & \text{si } X_n = i \\ 0, & \text{si } X_n \neq i \end{cases}$$

tenemos que $\sum\limits_{n=0}^{\infty} A_n$ representa el número de períodos en los que el proceso esta en el estado i. También

$$\begin{aligned} E\left[\sum_{n=0}^{\infty} A_n / X_0 = i\right] &= \sum_{n=0}^{\infty} E\left[A_n / X_0 = i\right] \\ &= \sum_{n=0}^{\infty} P\left\{X_n = i / X_0 = i\right\} \\ &= \sum_{n=0}^{\infty} P_{ii}^n. \end{aligned}$$

Proposición 2.1 *[17] El estado i es*

$$\text{recurrente si } \sum_{n=1}^{\infty} P_{ii}^n = \infty,$$
$$\text{transitorio si } \sum_{n=1}^{\infty} P_{ii}^n < \infty.$$

La propiedad de ser recurrente o transitorio es una propiedad de las clases. Si el estado i es recurrente, y este se comunica con el estado j, entonces el estado j es recurrente.

2.4.3 Probabilidades límites.

Abordaremos el comportamiento a "largo plazo" de las probabilidades de transición P_{ij}^n, es decir a que valor converge cuando $n \to \infty$.

El estado i tiene *período* d si $P_{ii}^n \geq 0$,donde d se define como el máximo común divisor de los naturales n. Un estado con período 1 se dice ser *aperíodico.* Se puede mostrar que la periodicidad es una propiedad de clase. Es decir, si el estado i tiene período d, y los estados i y j se comunican entre sí, entonces el estado j también tiene período d.

Si el estado i es recurrente, entonces se dice ser *recurrente positivo* si, iniciando en i, el tiempo esperado hasta regresar el proceso al estado i es finito. Se puede mostrar que la recurrencia positiva es una propiedad de clase. Puede ser mostrado que en un espacio de estados finitos de la cadena de Markov todos los estados recurrentes son recurrentes positivos. Los estados aperíodicos y recurrentes positivos son llamados ergódicos.

Teorema 2.1 *Para una cadena de Markov irreducible ergódica* $\lim_{n\to\infty} P_{ij}^n$ *existe y es independiente de i. Más precisamente, sea*

$$\pi_j = \lim_{n\to\infty} P_{ij}^n,\ j \geq 0$$

entonces π_j *es la única solución no-negativa de*

$$\pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij},\ j \geq 0,$$

$$\sum_{j=0}^{\infty} \pi_j = 1.$$

2.4.4 Tiempo de paro.

Dada la cadena de Markov homogénea $\{X_{n,},\ n \in N\}$ una variable aleatoria T se dice ser un Tiempo de Paro si para cada m el evento $\{T = m\}$ está determinado sólo por $\{X_0, X_1, \ldots X_m\}$.

Llamaremos tiempo aleatorio T al tiempo de paro para el proceso X_n. Si para toda $t \geq 0$, la función indicadora del evento $\{T \leq t\}$ es una función de los valores $\{N(s) : s \leq t\}$ del proceso al tiempo t.

2.5 El Proceso Poisson.

Un proceso estocástico $\{N(t),\ t \geq 0\}$ se dice que es un proceso de conteo si $N(t)$ representa el número total de "eventos" que han ocurrido hasta el tiempo t. Por lo tanto, un proceso de conteo $N(t)$ debe satisfacer:

(i) $N(t) \geq 0$.

(ii) $N(t)$ es valor entero.

(iii) Si $s < t$, entonces $N(s) \leq N(t)$.

(iv) Para $s < t$, $N(s) - N(t)$ es igual al número de eventos que han ocurrido en el intervalo $(s, t]$.

Un proceso de conteo se dice que posee *incrementos independientes* si el número de eventos que ocurren en intervalos de tiempo ajenos son independientes. Por ejemplo, esto significa que el número de eventos que han ocurrido en el tiempo t (esto es, $N(t)$) debe ser independiente del número de eventos ocurridos entre el tiempo t y $t+s$ (esto es $N(t+s) - N(t)$):

Un proceso de conteo se dice que posee incrementos estacionarios si la distribución del número de eventos que corren en un intervalo de tiempo depende solamente de la longitud del intervalo de tiempo. En otras palabras, el proceso tiene incrementos estacionarios si el número de eventos en el intervalo $(t_{1+s}, t_{2+s}]$ (esto es, $N(t_{2+s}) - N(t_{1+s})$) tiene alguna distribución como el número de eventos en el intervalo $(t_1,\ t_2]$ (esto es, $N(t_2) - N(t_1)$) para toda $t_1 < t_2$, y $s > 0$.

Uno de los tipos más importantes de proceso de conteo es el proceso Poisson, el cual se define como sigue:

Definición 2.1 *El proceso de conteo $\{N(t), t > 0\}$ se dice que es un proceso Poisson con razón λ, $\lambda > 0$, si:*

(i) $N(0) = 0$.

(ii) El proceso tiene incrementos independientes.

(iii) El número de eventos en cualquier intervalo de longitud t se distribuye Poisson con media λt. Esto es, para toda s, $t \geq 0$,

$$P\left\{N\left(t+s\right) - N\left(s\right) = n\right\} = \frac{\left(\lambda t\right)^n e^{-\lambda t}}{n!}, \quad n = 0,\ 1, ...$$

Note que se sigue de la condición (iii) que un proceso Poisson tiene incrementos estacionarios y también que

$$E[N(t)] = \lambda t.$$

Lo cual explica porque λ es llamada la razón o intensidad del proceso.

A continuación damos una segunda definición de un proceso Poisson, definiremos el concepto de una función f a partir de $o(h)$.

Definición 2.2 *La función f se dice $o(h)$ si*

$$\lim_{h\to 0}\frac{f(h)}{h}=0.$$

Otra definición del proceso Poisson.

Definición 2.3 *El proceso continuo $\{N(t), t\geq 0\}$ se dice que es un proceso Poisson con razón λ, $\lambda>0$, si:*

(i) $N(0)=0$.
(ii) El proceso tiene incrementos estacionarios e independientes.
(iii) $P\{N(h)=1\}=\lambda h+o(h)$.
(iv) $P\{N(h)\geq 2\}=o(h)$.

2.5.1 Intervalos de llegada y distribución de tiempos de espera.

Considere un proceso Poisson, y sea X_1, el tiempo de arribo del primer evento. Sea X_n, con $n\geq 1$, el tiempo entre el $(n-1)-$*ésimo* evento y el $n-$*ésimo* evento.

La sucesión $\{X_n, n\geq 1\}$ es llamada la sucesión de tiempos de intervalos de llegada.

Ahora determinaremos la distribución de los X_n. Primero veremos que el evento $\{X_1>t\}$ toma lugar si y sólo si los eventos del proceso Poisson ocurren en el intervalo [0, t], y esto es $P\{X_1>t\}=P\{N(t)=0\}=e^{-\lambda t}$.

Por lo tanto, X_1 tiene una distribución exponencial con media $1/\lambda$. Para la obtención de la distribución de X_2 bajo la condición de X_1. Consideramos:

$$\begin{aligned}P\{X_2>t/X_1=s\} &= P\left\{0 \text{ eventos en } (s,\ s+t]\,\big/ X_1=s\right\}\\ &= P\{0 \text{ eventos en } (s,s+t]\}\\ &\ \text{(por incrementos independientes)}\\ &= e^{-\lambda t}\\ &\ \text{(por incrementos estacionarios).}\end{aligned}$$

De lo anterior concluimos que X_2 es también una variable aleatoria exponencial con media $1/\lambda$, y más aún, que X_2 es independiente de X_1. Repitiendo el mismo argumento obtenemos lo siguiente.

Proposición 2.2 *Sea X_n, $n = 1, 2, \ldots$, la sucesión de v.a. independientes exponenciales idénticamente distribuidas e independientes entonces se distribuyen en forma exponencial con parámetro $1/\lambda$.*[2]

Otra cantidad de interés es S_n, el tiempo de llegada del $n-$*ésimo* evento, también llamado el tiempo de espera hasta el enésimo evento.

Donde

$$S_n = \sum_{i=1}^{n} X_i, \ n \geq 1.$$

La Proposición 2.2 implica que S_n tiene una distribución gama con parámetros n y λ. Es decir su densidad de probabilidad es

$$f(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \qquad t \geq 0.$$

Lo anterior pudo también haber sido derivado considerando que el $n-$*ésimo* evento ocurre antes o en el tiempo t si y sólo si el número de eventos ocurridos en el tiempo t es al menos n. Esto es

$$N(t) \geq n \Longleftrightarrow S_n \leq t.$$

Donde,

$$P\{S_n \leq t\} = P\{N(t) \geq n\} = \sum_{j=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^j}{j!}.$$

Aplicando la diferenciación se observa que la función de densidad de S_n es

$$\begin{aligned} f(t) &= -\sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^j}{j!} + \sum_{j=n}^{\infty} \lambda e^{-\lambda t} \frac{(\lambda t)^{j-1}}{(j-1)!} \\ &= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

[2] La proposición puede no sorprendernos. La suposición de incrementos estacionarios e independientes es equivalente a afirmar que, en algún punto es independiente de todo lo que ha ocurrido previamente (por incrementos independientes), y también alguna distribución como el proceso original (por incrementos estacionarios). En otras palabras. El proceso no ha tenido memoria, y los intervalos de tiempo exponencial son lo esperado.

La proposición 2.2 también nos da otra manera de definir un proceso Poisson. Suponiendo que iniciamos con una sucesión $\{X_n, n\geq 1\}$ de variables aleatorias independientes e idénticamente distribuidas en forma exponencial cada una tiene media $1/\lambda$. Ahora definiremos un proceso de conteo, decimos que el $n-$*ésimo* evento de este proceso ocurre al tiempo S_n, donde

$$S_n \equiv X_1 + X_2 + ... + X_n.$$

El proceso continuo $\{N(t),\ t\geq 0\}$ resulta ser Poisson con razón λ.

2.5.2 Distribución condicional y tiempos de llegada.

Supongamos que en un proceso Poisson exactamente un evento ha sucedido hasta el tiempo t, y nos preguntamos como determinar la distribución del tiempo en el cual ocurrió el evento. Puesto que un proceso Poisson posee incrementos estacionarios e independientes, parece razonable que cada intervalo en [0, t] de igual longitud debería tener alguna probabilidad de contener el evento. En otras palabras, el tiempo del evento debería tener una distribución uniforme sobre [0, t]. Esto es fácil de verificar, para $s \leq t$,

$$\begin{aligned} P\{X_1 < s/N(t) = 1\} &= \frac{P\{X_1 < s,\ N(t) = 1\}}{P\{N(t) = 1\}} \\ &= \frac{P\{1 \text{ evento en } [0,\ s),\ 0 \text{ eventos en } [s,\ t)\}}{P\{N(t) = 1\}} \\ &= \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\ &= \frac{s}{t}. \end{aligned}$$

2.5.3 Procesos Poisson no homogéneos.

En esta sección generalizaremos el proceso Poisson considerando la razón de llegadas al tiempo t como una función de t.

Definición 2.4 *El proceso contable* $\{N(t), t \geq 0\}$ *se dice que es un proceso Poisson no-estacionario o no-homogéneo con función de intensidad* $\lambda(t)$, $t \geq 0$ *si:*

(i) $N(0) = 0$.
(ii) $\{N(t), t \geq 0\}$ tiene incrementos independientes.
(iii) $P\{N(t+h) - N(t) \geq 2\} = o(h)$.
(iv) $P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$.

Si $m(t) = \int_0^t \lambda(s)\, ds$, entonces esto puede demostrar que

$$\begin{aligned} P\{N(t+s) - N(t) = n\} &= \exp\{-(m(t+s) - m(t))\} \frac{[m(t+s) - m(t)]^n}{n!} \\ con \quad n &\geq 0. \end{aligned}$$

Esto es, $N(t+s) - N(t)$ se distribuye Poisson con media $m(t+s) - m(t)$.

La importancia del proceso Poisson no-homogéneo reside en el hecho que no requiere incrementos estacionarios, y así considera la posibilidad de que los eventos puedan ser más probables de ocurrir en cierto tiempo que en otros momentos.

Cuando la función de intensidad $\lambda(t)$ está acotada, podemos pensar en un proceso no homogéneo como iniciado con una muestra aleatoria de un proceso Poisson homogéneo.

Específicamente, sea λ tal que

$$\lambda(t) \leq \lambda \qquad \forall t \geq 0$$

y considere un proceso Poisson con razón λ. Ahora si suponemos que un evento del proceso Poisson que ocurre en el tiempo t está contenido con probabilidad $\frac{\lambda(t)}{\lambda}$, entonces el proceso de contar los eventos es un proceso Poisson no-homogéneo con función de intensidad $\lambda(t)$.

Capítulo 3

Modelo de control Markoviano.

3.1 Introducción.

La teoría de control óptimo trata con sistemas dinámicos cuyo comportamiento puede modificarse o influenciarse mediante la elección de alguna de las variables del sistema llamada variable de control, acción o decisión. Un problema de control óptimo consiste de un modelo que contiene los elementos que describen el comportamiento del sistema y además un índice de funcionamiento o criterio de optimalidad el cual es una función que mide de alguna forma la respuesta del sistema a los controles aplicados. Este índice por lo general representa un costo o una ganancia y el objetivo es optimizar el índice de funcionamiento mediante la aplicación de los controles adecuados.

Los modelos para estudiar problemas de control óptimo se clasifican en:

i) estocásticos o determinísticos, dependiendo si se consideran o no componentes aleatorios,

ii) en tiempo continuo si los controles pueden elegirse en cualquier tiempo o en tiempo discreto si los controles se aplican en un conjunto discreto (finito o numerable).

En esta sección abordaremos los procesos de decisión de Markov los cuales son procesos estocásticos que describen la evolución de sistemas dinámicos controlados por sucesiones de decisiones o acciones. Así mismo, aplicaremos la programación dinámica para la solución de un proceso de decisión pues el objetivo es descomponer en etapas el problema a través del uso de cálculos recursivos. La evolución del sistema será el resultado de la interacción entre las "leyes de movimiento" del sistema y la sucesión de acciones que se tomen en el tiempo. Los diferentes caminos del sistema tendrán asociados consecuencias económicas; el objetivo es escoger la sucesión de acciones que conduzcan a una evolución óptima de acuerdo al índice que controle el sistema. "Óptimamente" será entendido de acuerdo a los objetivos que se pretendan alcanzar.

La aplicación de la programación dinámica nos conduce a la solución de un proceso de decisión estocástico que se puede describir a través de un número finito de etapas; las probabilidades de transición juegan un papel importante entre los puntos a ser considerados. La estructura de remuneración del proceso la describe el ingreso (o costo) resultante de la elección de una acción. Los resultados dependen de las alternativas de decisión disponibles para la persona que toma la decisión.

En el resto del capítulo se definen los elementos de un modelo de control Markoviano, el proceso estocástico canónico, los índices de funcionamiento de políticas, descontado y promedio, el problema de control óptimo (sin restricciones) y el algoritmo de programación dinámica. A lo largo del presente y siguiente capítulo se considera para fijar ideas un caso partícular de teoría de colas, el cual se va resolviendo. Para este caso se contempla que en la cola existe un controlador el cual acepta o rechaza al solicitante del servicio con el objeto de que el prestador del servicio esté desocupado u ocupado. Al final del capítulo proponemos una manera de abordar el ejemplo cuando en efecto se constituye una cola. Referencias para este capítulo son las siguientes: [2], [4], [9], [10] y [12].

3.2 Descripción del modelo.

Un *modelo de control Markoviano* (MCM) consiste de cinco elementos: estados, acciones, probabilidades de transición, recompensas y épocas de decisión, (X, A, P, r_t, T) que a continuación se describen:

a) El *espacio de estados* X, el cual suponemos que es un conjunto finito o numerable.

El estado inicial del proceso puede ser independiente de quien tome la decisión o bien el tomador de decisiones puede elegir una distribución inicial, su realización esta designada por $x_0 \in X$.

b) El *espacio de acciones* A que suponemos es un conjunto finito o númerable.

La *familia de acciones posibles* para cada estado $x \in X$ tiene asociado un conjunto A(x)⊂A. Denotamos por

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

al conjunto de todas las parejas estado-acción que son posibles. También es un conjunto finito o numerable.

c) La *probabilidad de transición* $P(y/x, a)$ es una distribución de probabilidad en X, para cada $(x, a) \in K$.

d) Las *funciones de costo por etapa* o *de recompensa* $r_t(x, a)$ $t = 0, 1, ...$ son funciones de K en R.

e) El *conjunto de épocas de decisión* T. Puede ser discreto o continuo. Cuando es discreto, las decisiones se efectuan en todas las épocas de decisión. Cuando es continua, las decisiones pueden ser

1 todas las épocas de decisión,

2 puntos aleatorios de tiempo cuando ciertos eventos ocurren,

3 escoger tiempos oportunos para efectuar la decisión.

Cuando las decisiones son hechas en forma continua, los problemas de decisión secuencial son analizados mejor usando métodos de teoría de control basados en sistemas de ecuaciones diferenciales estocásticas.

En problemas de tiempo discreto, al tiempo entre dos épocas de decisión consecutivas se le llama *período o etapa.* El conjunto de épocas de decisión puede ser finito, en tal caso $T \equiv \{1, 2, \ldots, N\}$ para algún entero $N < \infty$, o infinito, en esté caso $T \equiv \{1, 2, \ldots\}$. Los elementos de T (épocas de decisión) serán denotados por t y nos referiremos a él como al "tiempo t". Cuando N es finito, el problema de decisión será llamado problema de *horizonte finito*; en el otro caso un problema de *horizonte infinito.*

Un modelo de control Markoviano representa un sistema que evoluciona de la siguiente manera: en cada tiempo $t = 0, 1,...$ llamado *épocas de decisión*, el sistema se encuentra en el estado $x_t = x$ y se elige un control $a_{t+1} \in A(x)$ que produce lo siguiente:

1) se incurre en un costo $r_t(x, a)$ y

2) el sistema evoluciona al estado x_{t+1} de acuerdo a la distribución de probabilidad

$$P\,(y/x, a) = P\,(x_{t+1} = y/x_t = x,\ a_{t+1} = a)$$

una vez que el sistema se encuentra en el estado $x_{t+1} = x'$, se elige un nuevo control $a' \in A(x')$ y el proceso anterior se repite.

Nota: En algunas ocasiones se consideran recompensas por etapa en lugar de costos por etapa, una recompensa puede verse como un costo negativo.

3.2.1 Ejemplo: sistema de colas.

Consideraremos un sistema simple de colas formado por un servidor o una línea de espera. Al inicio del período, se observa el estado del sistema. El estado x_0 puede ser 0 (desocupado), es decir, que nadie se encuentra con el servidor; ó 1 (ocupado) cuando en el servidor está prestando el servicio; con base en esta observación, se elige una acción que puede ser 0 si se decide mantener un servicio lento o 1 que consiste en seleccionar un servicio rápido.

Por lo que:

- $X = \{0, 1\}$
- $A = \{0, 1\}$

Suponemos que el servicio concluye en el tiempo $t = 1$ y que la transición al estado x_1, se realiza de la siguiente forma:

(a) si $x_0 = 1$ y $a_1 = 1$ entonces la transición al estado $x_1 = 0$ ó $x_1 = 1$ sucede con probabilidad q^1 y $1 - q^1$; respectivamente.

(b) si $x_0 = 1$ y $a_1 = 0$ entonces la transición al estado $x_1 = 0$ ó $x_1 = 1$ sucede con probabilidad q^0 y $1 - q^0$; respectivamente.

Supondremos que $0<q^0<q^1<1$, donde q^0 es la probabilidad para el servicio lento y q^1 la probabilidad para el servicio rápido.

(c) si $x_0 = 0$ entonces independientemente del valor de a_1 la transición al estado $x_1 = 1$ ó $x_1 = 0$ sucede con probabilidades p y $1-p$, respectivamente, donde $p \in (0, 1)$ es la probabilidad de que un cliente llegue en el intervalo $(0, 1]$. Si el sistema está ocupado al llegar un cliente este se puede perder y no afecta el progreso futuro del proceso. Así contamos con ocho resultados elementales del tipo de sucesiones (x_0, a_1, x_1) : $(0, 0, 0)$, $(0, 0, 1)$, $(0, 1, 0)$,

$(1,0,0)$, $(0,1,1)$, $(1,0,1)$, $(1,1,0)$, $(1,1,1)$. Las probabilidades iniciales del sistema, serán

$$P_0(0),\ P_0(1) \geq 0 \text{ donde } P_0(0) + P_0(1) = 1.$$

La probabilidad de transición en el paso t esta expresada por la fórmula

$$p_t(y \mid x, a) = \begin{cases} p, & \text{si } x = 0,\ y = 1 \\ 1 - p, & \text{si } x = 0,\ y = 0 \\ q^a, & \text{si } x = 1,\ y = 0 \\ 1 - q^a, & \text{si } x = 1,\ y = 1, \end{cases}$$

donde p es la probabilidad de que un cliente llegue en el intervalo $(t-1, t]$. q^0 (q^1) es la probabilidad de terminar el servicio entre el tiempo $t-1$ y el momento t para el servicio lento (rápido); $q^0 < q^1$. La pérdida en el intervalo $(t-1, t]$ está formada por los siguientes términos:

a) la función de costo por el tipo de servicio dado e_a, donde $e_0 < e_1$ son números dados;

b) la penalidad causada por la pérdida de una orden la cual es igual a $I\{x_{t-1} = 1\} \cdot c$[1]. Se incurre en esta penalidd sólo si el cliente al llegar encuentra el sistema ocupado.

Así

$$r_t(x_{t-1}, a) = e_a + c \cdot I\{x_{t-1} = 1\}$$

aplicando el valor esperado

$$\begin{aligned} r_t(x, a) &= e_a + c \cdot E[I\{x_{t-1} = 1\}] \\ &= e_a + c \cdot [P(x = 1) \cdot 1 + P(x = 1) \cdot 0] \end{aligned}$$

es decir

$$r_t(x, a) = e_a + cxp$$

para $t = 1, 2, \ldots T$.

Se tiene además un costo terminal:

$$r_T(x) = d.$$

Como mencionamos, $P_0(0)$ y $P_0(1)$ son probabilidades dadas al inicio del sistema cuando está desocupado y ocupado al instante $t = 0$;

$$P_0(0) + P_0(1) = 1.$$

[1] Estamos considerando la función indicadora.

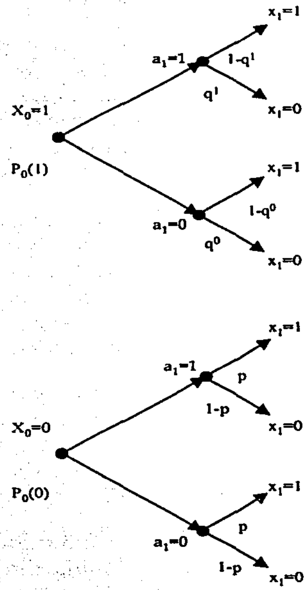

Figura 3.1: El diagrama de operación para un sistema de colas controlado.

Para el cálculo de las probabilidades $P(x_1/x_0, a_1)$: describiremos una regla para elegir las acciones $a_1 \in A$ en todos los posibles $x_0 \in X$. A continuación definiremos tal regla de decisión pero antes abordaremos el siguiente concepto. En la figura 3.1 se incluyen las posibles evoluciones del sistema para cada uno de los estados iniciales.

3.2.2 Reglas de decisión.

Denotamos a H_t como el conjunto de todas las historias h_t, y definimos:

$$H_0 \ : \ = X \text{ y}$$

$$H_t \ : \ = X \times (A \times X)^t \ , \ t = 1, 2, \ldots$$

Una *regla de decisión* o *selector* es un procedimiento para elegir una acción en cada estado en una época de decisión especificada y tomando en cuenta la historia. Las reglas de decisión pueden ir desde las Markovianas deterministas hasta las reglas de decisión aleatorizadas y con dependencia de las historias, dependiendo de como incorporan la información pasada y como seleccionan las acciones.

Las reglas Markovianas y deterministas, son funciones $f_t : X \longrightarrow A$, que especifican la acción a seguir cuando el sistema ocupa el estado x en la época de decisión t. Para cada $x \in X$, $f_t(x) \in A(X)$. Esta regla de decisión se llama *Markoviana* (desmemoriada) porque depende de los estados y de las acciones previas sólo a través del último estado del sistema, y *determinista* porque la elección de una acción se efectuá con certeza.

Llamaremos a una regla de decisión determinista *dependiente de la historia* si para elegir la acción se toma en cuenta la historia ocurrida hasta ese momento representada por la sucesión de estados y acciones previas. Es decir, φ_t es una función de la historia $h_t = (x_1, a_1, \ldots, x_{t-1}, a_{t-1}, x_t)$ donde x_i y a_i denotan el estado y acción del sistema en la época de decisión i.

Una regla de decisión *aleatorizada* φ_t especifica una distribución de probabilidad $q_{\varphi_t}(\cdot)$ sobre el conjunto de acciones. Las reglas de decisión Markovianas aleatorizadas mapean el conjunto de estados en el conjunto de las distribuciones de probabilidad sobre el espacio de acciones, esto es $\varphi_t : X \longrightarrow P(A)$, con la restricción $P(A(x)) = 1$. Una regla de decisión determinista se puede considerar como un caso especial de una regla de decisión aleatorizada.

En resumen, podemos clasificar las reglas de decisión como dependientes de la historia y aleatorizadas (HA), dependientes de la historia y deterministas (HD), Markovianas y aleatorizadas (MA), o Markovianas y deterministas (MD) en función de su grado de dependencia sobre la información anterior y sobre su método de selección de la acción.

3.2.3 Políticas.

Una *política, plan de contingencia* o *estrategia* especifica la regla de decisión a usar en todas las épocas de decisión. Es decir una política π es una sucesión de reglas de decisión $\pi = (\varphi_0, \varphi_1, \ldots \varphi_T)$ con $T \leq \infty$. Sea $\prod$ el conjunto de todas la políticas.

Llamaremos a una política *estacionaria* si en cada época de decisión aplica la misma regla, es decir, si $\pi = (\varphi, \varphi, \ldots)$. Con $\prod^E$ designaremos al conjunto de políticas estacionarias.

Si las φ_t son markovianas, la política $\pi = (\varphi_0, \varphi_1, \ldots, \varphi_T)$ con $T \leq \infty$ se dice ser *Markoviana* (tiene la propiedad de la pérdida de la memoria).Sea $\prod^M$ el conjunto de la políticas markovianas.

Sea $\prod^{EM}$ el conjunto de políticas markovianas que son invariantes en el tiempo, se conocen como *markovianas estacionarias.*

Una política $\pi = (\varphi_0, \varphi_1, \ldots \varphi_T)$ con $T \leq \infty$ es *determinista* si φ_t para toda $t \in T$ es determinista. Y sea el conjunto $\prod^D$ el formado por las políticas deterministas, es decir son iguales sin importar el instante del tiempo en el que se encuentre el sistema. De esta forma $\prod^{ED}$ es el conjunto de políticas *estacionarias deterministas.*

3.2.4 Índices de funcionamiento.

Definición 3.1 *Sean $x \in X$, $\pi \in \prod$, y $r_T : X \to R$, con $T < \infty$*

(a) Se define el costo total esperado con costo terminal r_T como

$$J_T(\pi, x) := E_x^\pi \left\{ \sum_{t=0}^{T-1} r_t(x_t, a_{t+1}) + r_T(x_T) \right\}.$$

(b) Para $\beta \in (0,1)$, el costo esperado $\beta -$ descontado está dado por

$$V_T(\pi, x) := E_x^\pi \left\{ \sum_{t=0}^{T-1} \beta^t r_t(x_t, a_{t+1}) + \beta^T r_T(x_T) \right\}.$$

La motivación para el estudio de problemas con índice de costo descontado proviene de factores económicos. El factor de descuento β permite encontrar el "valor presente" de una cantidad en la época de decisión t: un costo de L unidades en el tiempo t equivale a un costo presente de $\beta^t L$ unidades.

Si $T = \infty$ y $r_t(x, a)$ es un costo estacionario es decir $r(x, a)$, tenemos:

Definición 3.2 *Para $x \in X$, $\pi \in \prod$ y $\beta \in (0,1)$, se define el* costo total esperado *$\beta -$ descontado como*

$$V(\pi, x) := E_x^\pi \left[\sum_{t=0}^{\infty} \beta^t r(x_t, a_{t+1}) \right].$$

Definición 3.3 *Para $x \in X$ y $\pi \in \prod$, se define el costo promedio esperado por*

$$J(\pi, x) := \limsup_{T \to \infty} \frac{1}{T} E_x^{\pi} \sum_{t=0}^{T-1} r(x_t, a_{t+1}).$$

En algunas ocasiones consideramos al tiempo como aleatorio asumiendo $\tau = \min \left\{ t \geq 0, \ x_t \notin \widehat{X} \right\}$ que es la primer salida del subconjunto $\widehat{X} \subset X$ que está dado. En este caso es conveniente introducir una nueva función de costo a el modelo de Markov.

Definición 3.4 *Costo Total hasta un tiempo aleatorio τ*

$$J_\tau(\pi, x) := E_{x,\tau}^{\pi} \left\{ \sum_{t=0}^{\tau-1} r_t(x_t, a_{t+1}) + r_\tau(x_\tau) \right\}.$$

De igual manera establecemos el costo esperado $\beta -$ descontado hasta un tiempo aleatorio τ por

$$V_\tau(\pi, x) := E_{x,\tau}^{\pi} \left\{ \sum_{t=0}^{\tau-1} \beta^t r_t(x_t, a_{t+1}) + \beta^\tau r_\tau(x_\tau) \right\}.$$

3.2.5 Problema de control Markoviano.

Un *problema de control Markoviano* está formado por un modelo controlado y un índice de funcionamiento. El objetivo que se persigue es encontrar una política bajo la cual el índice de funcionamiento alcance su valor óptimo, es decir, si $v(\pi, x)$ es el índice de funcionamiento entonces el objetivo es encontrar una política π^* tal que

$$v(\pi^*, x) = \inf_{\pi \in \prod} v(\pi, x) = v(x), \ \forall x \in X \tag{3.1}$$

A esta política se le llama una política óptima y a $v(x)$ se le llama la función de valor.

Observación 3.1 *En la expresión (3.1) se ha usado el ínfimo para optimizar el índice de funcionamiento ya que el modelo que se esta exponiendo incluye funciones de costo. En el caso que se trate de optimizar recompensas o ganancias se usa supremo.*

3.2.6 Proceso estocástico canónico.

Un modelo de probabilidad consiste de tres elementos: un espacio muestral Ω, un $\sigma-$*álgebra* de subconjunto de Ω, $B(\Omega)$, y una medida de probabilidad P sobre $B(\Omega)$.

En un proceso de decisión de Markov con horizonte finito, tenemos

$$\Omega = X \times A \times X \times A \times \cdots \times A \times X = \{X \times A\}^{N-1} \times S,$$

y en un modelo de horizonte infinito, $\Omega = \{X \times A\}^{\infty}$. Un elemento típico $\omega \in \Omega$ consiste de una sucesión de estados y acciones, esto es

$$\omega = (x_0, a_1, x_1, a_2, \ldots, a_N, x_N),$$

y, en un modelo de horizonte infinito,

$$\omega = (x_0, a_1, x_1, a_2, \ldots).$$

Nos referiremos a ω como *un camino muestral.* En modelos con horizonte finito, el sigma álgebra está dado por $B(\Omega) = B\left(\{X \times A\}^{N-1} \times X\right)$ y en modelos de horizonte infinito, $B(\Omega) = B(\{X \times A\}^{\infty})$.

Definimos las variables aleatorias X_t y Y_t las cuales toman los valores en X y A, respectivamente, por

$$X_t(\omega) = x_t \text{ y } Y_t(\omega) = a_{t+1} \tag{3.2}$$

para $t = 0, 1, 2, \ldots, N$, $N \leq \infty$. Esto significa que cuando la sucesión observada de estados y acciones es ω, la variable aleatoria X_t denota el estado al tiempo t, y Y_t denota la acción al tiempo t. Definimos el proceso de la historia Z_t por

$$Z_0(\omega) = x_0 \text{ y } Z_t(\omega) = (x_0, a_1, \ldots x_t) \text{ para } 0 \leq t \leq N; \ N \leq \infty,$$

es decir, sea la distribución de probabilidad $P_0(\cdot)$ que denota la *distribución inicial* del estado del sistema.

Una política aleatorizada dependiente de la historia $\pi = (\varphi_1, \varphi_2, \ldots, \varphi_{N-1})$, $N \leq \infty$, induce a una probabilidad P^{π} sobre $(\Omega, B(\Omega))$ a través de

$$P^{\pi}\{X_0 = x\} = P_0(x),$$

$$P^{\pi}\{Y_t = a / Z_t = h_t\} = \varphi_t(a/h_t),$$

$$P^{\pi}\left\{X_{t+1}=x/Z_t=\left(h_{t-1},a_{t-1},x_t\right),Y_t=a_{t+1}\right\}=p_t\left(x\mid x_t,a_{t+1}\right)$$

de manera que la probabilidad de un camino muestral $\omega=(x_0,a_1,x_1,\ldots,x_N)$ esta dado por

$$\begin{aligned}P^{\pi}\left(x_0,a_1,x_1,a_2,\ldots,,x_N\right) &= P_0\left(x_0\right)\varphi_0\left(a_1/h_0\right)p_1\left(x_1\mid x_0,a_1\right)\varphi_1\left(a_2/h_1\right)\\ &\quad\ldots\varphi_{N-1}\left(a_{N-1}/h_N\right)p_{N-1}\left(x_N\mid x_{N-1},a_N\right).\end{aligned}$$

La existencia de esta medida de probabilidad está garantizada por Ionescou Tulcea [12]

Denotamos por E_x^{π} el valor esperado correspondiente a P_x^{π}, es decir, si W es una variable aleatoria definida en Ω_T, su valor esperado está dado por

$$E_x^{\pi}\left(W\right)=\sum W\left(x_0,a_1,\ldots,x_{T-1},a_T,x_T\right)P_x^{\pi}\left(x_0,a_1,\ldots,x_{T-1},a_T,x_T\right)$$

donde la suma se toma sobre todas las trayectorias de Ω_T.

(a) Si W es una función de x_t, $a_{t+1},\ldots,x_T$, y $h_t\in H_t$ entonces

$$E_x^{\pi}\left(W\left(x_t,a_{t+1},\ldots,x_T\mid h_t\right)\right)=\sum W\left(x_t,a_{t+1},\ldots,x_T\right)P_x^{\pi}\left(x_t,a_{t+1},\ldots,x_T\right.$$

(b) Si v es una función de x_{t+1} entonces

$$E_x^{\pi}\left(v\left(x_{t+1}\mid h_t\right)\right)=\sum_{y\in Y}v\left(y\right)P\left(y\mid x,f_t\left(h_t\right)\right)\tag{3.3}$$

3.3 Costo total.

En esta sección abordaremos la teoría para la solución de un problema con el índice de costo total con horizonte finito, por medio de la programación dinámica cuya característica fundamental consiste en "reducir". un problema de optimización en T etapas a T problemas de optimización en una etapa.

Como mencionamos anteriormente el índice en costo total esperado esta dado por $J_T\left(\pi,x\right)$, es decir se trata de encontrar una política $\pi^*\in\prod$ tal que

$$J_T\left(\pi^*,x\right)=\inf_{\pi\in\prod}J_T\left(\pi,x\right)=J_T\left(x\right).\tag{3.4}$$

A la función J_T definida en (3.4) se le llama la *función de costo óptimo.*

3.3.1 Algoritmo de programación dinámica.

Para $t = 0, 1, 2, \ldots, T$ se definen las funciones de programación dinámica v_t en X recursivamente por

$$v_T(x) = r_T(x) \tag{3.5}$$

$$v_t(x) = \min_{a \in A(x)} \left\{ r_t(x, a) + \sum_{y_i} v_{t+1}(y_i) \, p_t(y_i \mid x, a) \right\}, \; t = 1, 2, \ldots, T. \tag{3.6}$$

las cuales son llamadas *ecuaciones de Bellman.* La función $v_t(x)$ es también llamada la *función de Bellman.*

Teorema 3.1 *(Algoritmo de la Programación Dinámica) [9]. Si para cada $t = T-1, T-2, \ldots, 0$, existe $f_t^*(x) \in A(X)$ tal que*

$$v_t(x) = r_t(f_t^*(x), x) + \sum_{y_i} v_{t+1}(y) \, p_t(y_i \mid x, f_t^*(x))$$

entonces

i) $v_0(x) = J_T(x)$.

ii) La política $\pi^* = (f_0^*, f_1^*, \ldots, f_{T-1}^*)$ es óptima, es decir

$$J_T(\pi^*, x) = J_T(x) = v_0(x)$$

ostración: Sea $\pi \in \prod$ una política arbitraria y definamos las funciones

$$u_t(h_t) = E_x^\pi \left[\sum_{n=t}^{T-1} r_n(x_n, a_{n+1}) + r_T(x_T) \mid h_t \right], \; t = 0, \ldots, T-1, \tag{3.7}$$

$$u_T(h_T) = E_x^\pi \left[r_T(x_T) \mid h_T \right] \tag{3.8}$$

Nótese que u_t representa el costo total esperado desde la época de decisión t hasta la época T, dada la historia h_t.

Probaremos que para todo $t = 0, 1, \ldots, T$,

$$v_t(x_t) \le u_t(h_t) \tag{3.9}$$

y en el caso en que $\pi = \pi^*$, se cumple que $v_t(x_t) = u_t(h_t)$. Esto es suficiente para la demostración ya que entonces tenemos

$$v_0(x) \le J_T(\pi, x) \text{ y } v_0(x) = J_T(\pi^*, x)$$

y puesto que π es una política arbitraria tendremos

$$J_T(\pi^*, x) = J_T(x) = v_0(x).$$

La demostración de (3.9) será por inducción. Para $t = T$, se sigue de (3.5) y (3.8) que

$$u_T(h_T) = r_T(x_T) = v_T(x_T).$$

Supóngase que

$$v_{t+1}(x_{t+1}) \leq u_{t+1}(h_{t+1}). \tag{3.10}$$

Entonces de (3.7) se tiene

$$u_t(h_t) = E_x^\pi[r_t(x_t, a_t) \mid h_t] + E_x^\pi\left[\sum_{n=t+1}^{T-1} r_n(x_n, a_n) + r_T(x_T) \mid h_t\right]$$

y usando propiedades de la esperanza condicional, la hipótesis de inducción y (3.3) se obtiene

$$\begin{aligned} u_t(h_t) &= r_t(x_t, f_t(h_t)) + E_x^\pi[u_{t+1}(h_{t+1}) \mid h_t] \\ &\geq r_t(x_t, f_t(h_t)) + E_x^\pi[v_{t+1}(h_{t+1}) \mid h_t] \qquad (3.11) \\ &= r_t(x_t, f_t(h_t)) + \sum_{y\in X} v_{t+1}(y) P_t(dy \mid x, f(h)) \end{aligned}$$

por lo que

$$u_t(h_t) \geq \min_{a\in A(x_t)}\left\{r_t(x_t, a) + \sum_{y\in X} v_{t+1}(y) P_t(dy \mid x, a)\right\} = v_t(x_t)$$

con lo que hemos probado (3.9). Por otra parte, si se cumple la igualdad en (3.10) y $\pi = \pi^*$, entonces cuando la igualdad en (3.11) se cumple tenemos:

$$u_t(h_t) = r_t(x_t, a) + \sum_{y\in X} v_{t+1}(y) P_t(dy \mid x, a) = v_t(x_t)$$

$$u_t(h_t) = v_t(x_t).$$

Si consideramos $t = 0$ en (3.7) y en (3.9) vemos que, para algún estado inicial $x_0 = x$,

$$J_0(x) = E_x^{\pi}[J_0(x_0)] \leq E_x^{\pi}[u_0(h_0)] = J(\pi, x), \qquad (3.12)$$

que, para cualquier π, se cumple $J_0(x) \leq \min_{\pi} J(\pi, x) = J^*(x)$. Finalmente, cuando la igualdad $\pi = \pi^*$ se cumple tenemos la igualdad en (3.12), concluimos que $J_0(x) = J(\pi^*, x) = J^*(x)$. ♠

3.3.2 Ejemplo: sistema de colas con costo total.

Retomando el ejemplo mencionado en 3.2.1. Sea $X = \{0, 1\}$ donde $x_t = 0$ ($x_t = 1$) si el sistema esta desocupado (ocupado) al momento t; $A = \{0, 1\}$ donde $a_t = 0$ ($a_t = 1$) representa elegir el servicio menos intensivo (más intensivo) en el intervalo $(t-1, t]$. A continuación aplicaremos el algoritmo de programación dinámica para nuestro ejemplo.

El término del servicio para cualquier estado es

$$v_T(x) = r_T(x) = d \qquad \text{para } x \in \{0, 1\}.$$

Para el caso $T - 1$, tenemos

$$\begin{aligned}
v_{T-1}(x) &= \min_{a\in A}\left\{r_{T-1}(x,a) + \sum_X v_T(y)\, p_{T-1}(dy \mid x,a)\right\} \\
&= \min_{a\in A}\{r_{T-1}(x,a) + E[v_T(y) \mid x,a]\} \\
&= \min_{a\in A}\{r_{T-1}(x,a) + d\} = \min_{a\in\{0,1\}}\{e_a + xpc + d\} \\
&= \min\{e_0 + xpc + d,\ e_1 + xpc + d\}, \\
\text{como } e_0 &< e_1 \\
&= e_0 + pcx + d.
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
v_{T-1}(x) &= e_0 + pcx + d \\
v_{T-1}(x) &= e_0 + pcx + v_T(x).
\end{aligned}$$

Para $T - 2$

$$v_{T-2}(x) = \min_{a\in A}\{r_{T-2}(x,a) + E[v_{T-1}(x_{T-1}) \mid x,a]\}$$

$$
\begin{aligned}
&= \min_{a\in A}\left\{r_{T-2}(x,a)+e_0+d+pcE\left[x_{T-1}\mid x,a\right]\right\}\\
&= \min_{a\in A}\left\{e_a+cpx+e_0+d+pcE\left[x_{T-1}\mid x,a\right]\right\}
\end{aligned}
$$

Caso $x=0$

$$
\begin{aligned}
v_{T-2}(0) &= \min\left\{\begin{array}{c} e_0+cp(0)+e_0+d+pcE\left[x_{T-1}\mid 0,0\right], \\ e_1+cp(0)+e_0+d+pcE\left[x_{T-1}\mid 0,1\right] \end{array}\right\}\\
&= \min\left\{\begin{array}{c} 2e_0+d+pc\left[1(p)+0(1-p)\right], \\ e_1+e_0+d+pc\left[1(p)+0(1-p)\right] \end{array}\right\}\\
&= \min\left\{2e_0+d+p^2c,\ e_1+e_0+d+p^2c\right\},
\end{aligned}
$$

es decir

$$v_{T-2}(0)=2e_0+d+p^2c.$$

Ahora, $x=1$

$$
\begin{aligned}
v_{T-2}(1) &= \min\left\{\begin{array}{c} e_0+cp(1)+e_0+d+pcE\left[x_{T-1}\mid 1,0\right], \\ e_1+cp(1)+e_0+d+pcE\left[x_{T-1}\mid 1,1\right] \end{array}\right\}\\
&= \min\left\{\begin{array}{c} e_0+cp+e_0+d+pcE\left[x_{T-1}\mid 1,0\right], \\ e_1+cp+e_0+d+pcE\left[x_{T-1}\mid 1,1\right] \end{array}\right\},
\end{aligned}
$$

calculando

$$
\begin{aligned}
E\left[x_{T-1}/1,0\right] &= 0\cdot q^0+1\cdot\left(1-q^0\right)=\left(1-q^0\right),\\
E\left[x_{T-1}/1,1\right] &= 0\cdot q^1+1\cdot\left(1-q^1\right)=\left(1-q^1\right),\\
v_{T-2}(1) &= \min\left\{\begin{array}{c} 2e_0+pc+d+pc\left(1-q^0\right), \\ e_1+e_0+pc+d+pc\left(1-q^1\right) \end{array}\right\}.
\end{aligned}
$$

Para elegir el mínimo supongamos por un momento que es válida la siguiente desigualdad

$$
\begin{aligned}
2e_0+pc+d+pc\left(1-q^0\right) &< e_1+e_0+pc+d+pc\left(1-q^1\right),\\
e_0+pc\left(1-q^0\right) &< e_1+pc\left(1-q^1\right),\\
pc\left(q^1-q^0\right) &< e_0-e_1,\\
pc &< \frac{e_0-e_1}{q^1-q^0}.
\end{aligned}
$$

Con lo cual obtenemos:

$$v_T(0)=v_T(1)=d,$$

$$v_{T-1}(x) = \min_{a\in A} \{r_{T-1}(x,a) + E[v_T(x_T)/x.a]\}.$$

Si $x = 0$,

$$\begin{aligned}
v_{t-1}(0) &= \min_{a\in\{0,1\}} \{r_{T-1}(0,a) + E[v_T(x_T)/0.a]\} \\
&= \min \left\{ \begin{array}{c} e_0 + v_T(1) P(1/0,0) + v_T(0) P(0/0,0), \\ e_1 + v_T(1) P(1/0,1) + v_T(0) P(0/0,1) \end{array} \right\} \\
&= \min\{e_0 + dp + d(1-p), e_1 + dp + k(1-p)\} \\
& o \\
&= \min\{e_0 + v_T(1)p + v_T(0)(1-p), e_1 + v_T(1)p + v_T(0)(1-p)\} \\
&= e_0 + v_T(1)p + v_T(0)(1-p),
\end{aligned}$$

por lo tanto

$$v_{t-1}(0) = e_0 + v_T(1)p + v_T(0)(1-p).$$

Ahora calculemos $x = 1$

$$v_{t-1}(1) = \min_{a\in\{0,1\}} \{r_T(1,a) + E[v_T(x_T)/1,a]\}$$

(donde $r_T(1,a) = e_a + pc$)

$$\begin{aligned}
&= \min \left\{ \begin{array}{c} e_0 + pc + v_T(1) P(1/1,0) + v_T(0) P(0/1,0), \\ e_1 + pc + v_T(1) P(1/1,1) + v_T(0) P(0/1,1) \end{array} \right\} \\
&= \min \left\{ \begin{array}{c} e_0 + pc + v_T(1)(1-q^0) + v_T(0) q^0, \\ e_1 + pc + v_T(1)(1-q^1) + v_T(0) q^1 \end{array} \right\},
\end{aligned}$$

para conocer el mínimo, supongamos que la siguiente desigualdad es válida

$$\begin{aligned}
e_0 + pc + v_T(1)\left(1-q^0\right) + v_T(0) q^0 &< e_1 + pc + v_T(1)\left(1-q^1\right) + v_T(0) q^1, \\
-v_T(1) q^0 + v_T(0) q^0 &< e_1 - e_0 - v_T(1) q^1 + v_T(0) q^1, \\
-q^0 (v_T(1) - v_T(0)) &< e_1 - e_0 - q^1 (v_T(1) + v_T(0)), \\
(v_T(1) - v_T(0))\left(q^1 - q^0\right) &< e_1 - e_0, \\
\therefore\ v_T(1) - v_T(0) &< \frac{e_1 - e_0}{(q^1 - q^0)}.
\end{aligned}$$

De lo cual deducimos la siguiente regla

$$\begin{aligned}
\text{si } v_T(1) - v_T(0) &< \frac{e_1 - e_0}{(q^1 - q^0)}, \qquad \text{si } a = 0, \\
v_T(1) - v_T(0) &> \frac{e_1 - e_0}{(q^1 - q^0)}, \qquad \text{si } a = 1,
\end{aligned}$$

si se cumple la igualdad es indistinto tomar una u otra acción.

En resumen tenemos:

$$\begin{aligned} v_T(0) &= v_T(1) = d, \\ v_{t-1}(0) &= e_0 + v_t(0)(1-p) + v_t(1)p, \\ v_{t-1}(1) &= \min\left\{ \begin{array}{c} e_0 + pc + v_t(1)(1-q^0) + v_t(0)q^0, \\ e_1 + pc + v_t(1)(1-q^1) + v_t(0)q^1 \end{array} \right\}. \end{aligned} \tag{3.13}$$

Para el caso cuando $x = 0$, la igualdad se deduce de la inecuación $e_0 < e_1$, esto es obvio pues es mejor usar un sistema menos intenso, cuando no hay clientes para atender.

Cuando $x = 1$ seguimos la siguiente regla:

si $v_t(1) - v_t(0) > \frac{e_1 - e_0}{(q^1 - q^0)}$, entonces el mínimo es alcanzado cuando $a = 1$;

si $v_t(1) - v_t(0) < \frac{e_1 - e_0}{(q^1 - q^0)}$, entonces el mínimo es alcanzado cuando $a = 0$.

Si se da la igualdad entre ambas ecuaciones es indistinto elegir una u otra acción.

Definamos $\Lambda = \frac{e_1 - e_0}{(q^1 - q^0)}$, el cual es el punto que divide a la región factible en 2 partes ver gráficas (3.2, 3.3 y 3.4); $\Delta_t = v_t(1) - v_t(0)$, entonces

$$\Delta_T = v_T(1) - v_T(0) = d - d = 0, \quad \therefore \Delta_T = 0.$$

Hagamos la deducción para Δ_{t-1},
$\Delta_{t-1} = v_{t-1}(1) - v_{t-1}(0)$,
para $v_{t-1}(1)$, tenemos:

$$\begin{aligned} v_{t-1}(1) &= \min\left\{ \begin{array}{c} e_0 + pc + v_t(1)(1-q^0) + v_t(0)q^0, \\ e_1 + pc + v_t(1)(1-q^1) + v_t(0)q^1 \end{array} \right\} \\ &= pc + \min\left\{ \begin{array}{c} e_0 + v_t(1)(1-q^0) + v_t(0)q^0, \\ e_1 + v_t(1)(1-q^1) + v_t(0)q^1 \end{array} \right\}. \\ &= pc + e_0 + \min\left\{ \begin{array}{c} v_t(1) - v_t(1)q^0 + v_t(0)q^0, \\ (e_1 - e_0) + v_t(1) - v_t(1)q^1 + v_t(0)q^1 \end{array} \right\} \\ &= pc + e_0 + v_t(1) + \min\left\{ \begin{array}{c} -q^0(v_t(1) - v_t(0)), \\ (e_1 - e_0) - q^1(v_t(1) - v_t(0)) \end{array} \right\}, \end{aligned}$$

por lo tanto

$$v_{t-1}(1) = pc + e_0 + v_t(1) + \min\left\{-\Delta_t q^0, (e_1 - e_0) - \Delta_t q^1\right\}.$$

Para $v_{t-1}(0)$, tenemos:

$$
\begin{aligned}
v_{t-1}(0) &= e_0 + v_t(0)(1-p) + v_t(1)p \\
&= e_0 + v_t(0) - v_t(0)p + v_t(1)p \\
&= e_0 + v_t(0) + p(v_t(1) - v_t(0)),
\end{aligned}
$$

por lo cual

$$v_{t-1}(0) = e_0 + v_t(0) + p\Delta_t,$$

restando $v_{t-1}(1)$ y $v_{t-1}(0)$

$$
\begin{aligned}
v_{t-1}(1) - v_{t-1}(0) &= pc + e_0 + v_t(1) + \min\left\{-\Delta_t q^0, (e_1 - e_0) - \Delta_t q^1\right\}. - e_0 - v_t(\\
&= pc + (1-p)\Delta_t + \min\left\{-\Delta_t q^0, (e_1 - e_0) - \Delta_t q^1\right\}.
\end{aligned}
$$

En resumen

$$
\left.
\begin{aligned}
\Delta_T &= 0, \\
\Delta_{t-1} &= pc + (1-p)\Delta_t + \min\{-\Delta_t q^0, (e_1 - e_0) - \Delta_t q^1\}.
\end{aligned}
\right\} \tag{3.14}
$$

Consecuentemente calculando $\Delta_T, \Delta_{T-1}, \ldots, \Delta_1$

A manera de ilustración

$$
\begin{aligned}
\Delta_T &= 0 \\
\Delta_{T-1} &= pc + 0 \cdot (1-p) + \min\left\{-(0)q^0, e_1 - e_0 - (0)q^1\right\} \\
&= pc + \min\{0, e_1 - e_0\} = pc + 0 \\
&= pc,
\end{aligned}
$$

por lo tanto

$$\Delta_{T-1} = pc$$

$$
\begin{aligned}
\Delta_{T-2} &= pc + (pc)(1-p) + \min\left\{-pcq^0, e_1 - e_0 - pcq^1\right\} \\
&= pc(2-p) + \min\left\{-pcq^0, e_1 - e_0 - pcq^1\right\},
\end{aligned}
$$

para encontrar el mínimo, consideremos válida la siguiente desigualdad

$$
\begin{aligned}
-pcq^0 &< e_1 - e_0 - pcq^1, \\
pc\left(q^1 - q^0\right) &< e_1 - e_0, \\
pc &< \frac{e_1 - e_0}{q^1 - q^0}, \\
i.e.\ \Delta_{T-1} &< \Lambda
\end{aligned}
$$

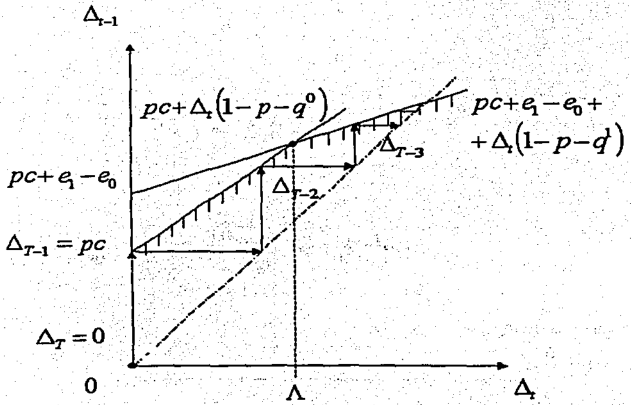

Figura 3.2: La Δ_t dinámica en $1 - p - q^1 > 0$ y $\Lambda < pc/(p + q^0)$.

En conclusión tenemos si $\Delta_t > \Lambda$, entonces elegimos $a = 1$, en caso contrario elegimos la acción $a = 0$.

Así hemos construido la regla de decisión de Markov

$$f^*(t, x_{t-1}) = x_{t-1} \cdot I\{\Delta_t > \Lambda\},$$

lo cual es una solución a nuestra tarea de acuerdo con el Teorema 3.1.

Claramente, la característica dinámica de la variable Δ_t, $t = T, T - 1, \ldots, 1$ establece la relación entre los valores de p, q^0, q^1, e_0 y e_1. Algunas ilustraciones son presentadas en las fig. 3.2-3.4; la línea diagonal que corre hacia arriba es la bisectriz del primer cuadrante.

Para concretar ideas:

Sea $T = 2$, $1 - p - q^0 > 0$.

Construyamos las deltas:

Δ_2 por los cálculos efectuados y por definición tenemos:

$$\Delta_2 = 0,$$

para Δ_1 es igual a

$$\Delta_1 = pc,$$

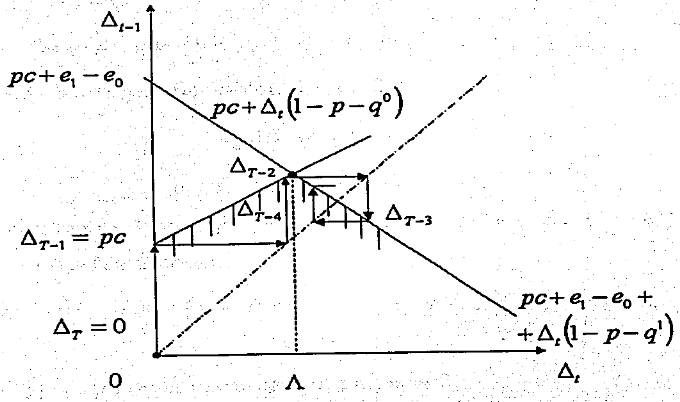

Figura 3.3: La Δ_t dinámica en $1-p-q^1<0$ y $\Lambda<pc/(p+q^0)$.

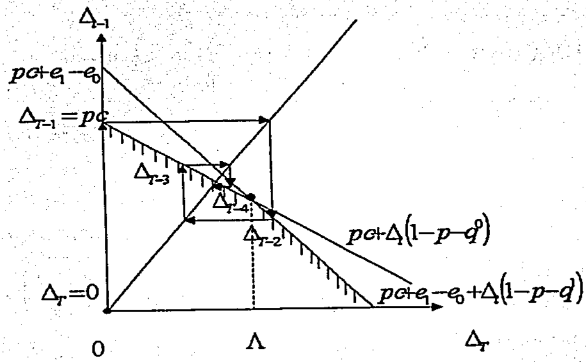

Figura 3.4: La Δ_t dinámica en $1-p-q^1<0$ y $\Lambda>pc/(p+q^0)$.

para Δ_0 es igual a

$$\Delta_0 = pc + pc\,(1-p) + \min\left\{-pcq^0, e_1 - e_0 - pcq^1\right\}.$$

Aplicando los resultados obtenidos en 3.3.1

$$v_2\,(0) = v_2\,(1) = d,$$

$$v_1\,(0) = e_0 + (1-p)\,d + pd,$$

$$v_1\,(1) = \min\left\{e_0 + pc + q^0 d + \left(1-q^0\right)d; e_1 + pc + q^1 d + \left(1-q^1\right)d\right\},$$

veamos que estrategia elegimos

$$\begin{aligned} e_0 + pc + q^0 d + \left(1-q^0\right)d &< e_1 + pc + q^1 d + \left(1-q^1\right)d, \\ e_0 &< e_1, \end{aligned}$$

como $e_0 < e_1$ elegimos la estrategia 0, i.e., $a = 0$.

$$v_1\,(1) = e_0 + pc + q^0 d + \left(1-q^0\right)d = e_0 + pc + d,$$

por lo tanto

$$v_1\,(1) = e_0 + pc + d.$$

Para $v_0\,(0)$

$$\begin{aligned} v_0\,(0) &= e_0 + (1-p)\left[e_0 + (1-p)\,d + pd\right] + p\left[e_0 + pc + q^0 d + (1-q^0)\,d\right] \\ &= e_0 + (1-p)\left[e_0 + d\right] + p\left[e_0 + pc + d\right] \\ &= e_0 + e_0 + d - pe_0 - pd + pe_0 + p^2 c + pd \\ &= 2e_0 + p^2 cd, \end{aligned}$$

$$v_0\,(0) = 2e_0 + p^2 cd.$$

Para $v_0\,(1)$, tenemos:

$$v_0\,(1) = \min\left\{\begin{array}{l} e_0 + pc + q^0\left[e_0 + (1-p)\,d + pd\right] + (1-q^0)\left[e_0 + pc + q^0 d + (1-q^0)\,d\right] \\ e_1 + pc + q^1\left[e_0 + (1-p)\,d + pd\right] + (1-q^1)\left[e_0 + pc + q^0 d + (1-q^0)\,d\right] \end{array}\right.$$

para conocer el mínimo supongamos válida la siguiente desigualdad y después hagamos la deducción:

$$\begin{aligned} &e_0 + pc + q^0\left[e_0 + (1-p)\,d + pd\right] + (1-q^0)\left[e_0 + pc + q^0 d + (1-q^0)\,d\right] < \\ &\qquad e_1 + pc + q^1\left[e_0 + (1-p)\,d + pd\right] + (1-q^1)\left[e_0 + pc + q^0 d + (1-q^0)\,d\right], \end{aligned}$$

$$-q^0\left[\left(e_0+pc+q^0d+\left(1-q^0\right)d\right)-\left(e_0+\left(1-p\right)d+pd\right)\right]<$$
$$e_1-e_0-q^1\left[\left(e_0+pc+q^0d+\left(1-q^0\right)d\right)-\left(e_0+\left(1-p\right)d+pd\right)\right],$$
$$\left[\left(e_0+pc+q^0d+\left(1-q^0\right)d\right)-\left(e_0+\left(1-p\right)d+pd\right)\right]\left(q^1-q^0\right)<e_1-e_0$$

$$\begin{aligned} pc\left(q^1-q^0\right) &< e_1-e_0, \\ pc &< \frac{e_1-e_0}{\left(q^1-q^0\right)}, \\ \Delta_1 &< \Lambda. \end{aligned}$$

Nota: por ambos lados llegamos, a la misma inecuación $\Delta_1<\Lambda$.

Para encontrar la respuesta consideremos dos casos

(a) Si $\Delta_1=pc>\Lambda$,

$$\Delta_0=pc+pc\left(1-p\right)+e_1-e_0-pcq^1.$$

Sabemos que

$$f\left(t,x_{t-1}\right)=x_{t-1}I\left\{\Delta_T>\Lambda\right\},$$

$$f\left(2,0\right)=0I\left\{\Delta_2>\Lambda\right\}=0,\quad f\left(1,0\right)=0I\left\{\Delta_1>\Lambda\right\}=0,$$
$$f\left(2,1\right)=1I\left\{\Delta_2>\Lambda\right\}=0,\quad f\left(1,1\right)=1I\left\{\Delta_1>\Lambda\right\}=1.$$

Por lo tanto

$$f^*\left(1,1\right)=1,$$

es decir en la etapa 1 el mínimo se obtiene cuando el estado es 1.

Ahora para $v_0\left(1\right)$,

$$\begin{aligned} v_0\left(1\right) &= e_1+pc+q^1\left[e_0+\left(1-p\right)d+pd\right]+\left(1-q^1\right)\left[e_0+pc+q^0d+\left(1-q^0\right)d\right] \\ &= e_1+pc+q^1\left[e_0+d\right]+\left(1-q^1\right)\left[e_0+pc+d\right] \\ &= e_1+pc+q^1e_0+q^1d+e_0+pc+d-q^1e_0-q^1pc-q^1d \\ &= e_1+e_0+2pc+d-q^1pc, \end{aligned}$$

por lo tanto

$$v_0\left(1\right)=e_1+e_0+2pc+d-q^1pc.$$

(b) Si $\Delta_1=pc<\Lambda$, entonces $\Delta_0=pc+pc\left(1-p\right)-pcq^0$,

$$f^*\left(t,x_{t-1}\right)\equiv 0.$$

$$\begin{aligned} v_0\left(1\right) &= e_0+pc+q^0\left[e_0+\left(1-p\right)d+pd\right]+\left(1-q^0\right)\left[e_0+pc+q^0d+\left(1-q^0\right)d\right] \\ &= e_0+pc+q^0\left[e_0+d\right]+\left(1-q^0\right)\left[e_0+pc+d\right] \\ &= e_0+pc+q^0e_0+q^0d+e_0+pc+d-q^0e_0-q^0pc-q^0d \\ &= 2e_0+2pc+d-q^0pc, \end{aligned}$$

por lo tanto

$$v_0(1) = 2e_0 + 2pc + d - q^0 pc,$$

(c) Si $\Delta_1 = pc = \Lambda$ y $P_0(0) \neq 1$, entonces sólo las estrategias óptimas π^* son óptimas para lo cual

$$\pi_2^*(0/x_0, a_1, 1) = 1;\ \pi_2^*(0/0) = \pi_2^*(0/x_0, a_1, 0) = 1;$$

la distribución de probabilidad $\pi_1^*(\cdot/1)$ puede ser arbitraria. El valor $v_0(1)$ esta expresado por alguna de las fórmulas presentadas anteriormente.

La pérdida mínima es igual a

$$v_0(0)\, P_0(0) + v_0(1)\, P_0(1).$$

3.4 Costo descontado.

En esta sección construiremos una estrategia o política que provea el mínimo para la expresión:

$$V(\pi, x) := E_x^\pi \left[\sum_{t=1}^{\infty} \beta^t r(x_t, a_{t+1}) \right]. \tag{3.15}$$

El coeficiente $\beta \in (0,1)$ tiene diferentes interpretaciones. Si la pérdida es expresada en dinero, entonces β es el coeficiente de devaluación: la pérdida de una unidad ahora es equivalente a la pérdida de β unidades al siguiente tiempo.

La función

$$v(x_0, a_1, \ldots, x_t) = \sum_{\theta=0}^{t-1} \beta^\theta r(x_\theta, a_{\theta+1}) + \beta^t \cdot v(x_t) \tag{3.16}$$

sobre X satisface la ecuación

$$v(x) = \min_{a \in A(X)} \left\{ r(x, a) + \beta \sum_X v(y)\, p(dy \mid x, a) \right\}. \tag{3.17}$$

La ecuación (3.17) es llamada la *ecuación de Bellman* (para el modelo descontado de Markov homogéneo); la función $v(x)$ es también llamada la *función de Bellman.* Su significado es intuitivamente evidente: $\beta^t \cdot v(x)$ es la pérdida mínima.

Sea φ una regla de decisión estacionaria arbitraria. La función

$$v^{\varphi}(x_0, a_1, \ldots x_t) = \sum_{\theta=0}^{t-1} \beta^{\theta} r(x_\theta, a_{\theta+1}) + \beta^t v^{\varphi}(x_t). \tag{3.18}$$

donde $v^{\varphi}(x)$ sobre X satisface la ecuación

$$v^{\varphi}(x) = r(x, d(x)) + \beta \sum_X v^{\varphi}(y) p(dy \mid x, d(x)). \tag{3.19}$$

Si $r(\cdot) \geq 0$ entonces la estrategia estimada φ es de la forma (3.18), donde $v^{\varphi}(\cdot)$ es la solución a la ecuación (3.19). En este caso la función $v^{\varphi}(\cdot)$ es a menudo llamada la regla de decisión estacionaria φ *estimada.*

Corolario 3.1 *Sea v_t las soluciones de la ecuación de Bellman*

$$\begin{aligned} v_T(x) &= r_T(x) \\ v_{t-1}(x) &= \min_{a \in A(X)} \left\{ r_{t-1}(x, a) + \beta \sum_{y_i} v_t(y_i) p_{t-1}(y_i \mid x, a) \right\} \end{aligned} \tag{3.20}$$

Sean $f_t(x)$ el valor de la acción entonces $v_0(x) = v(x)$ y $\pi^* = \{f_i\}$ es una política óptima.

Observación 3.2 *En el modelo estacionario $p_t = p$, $r_t = r$. Con horizonte infinito las Ecuaciones de Bellman toman la forma*

$$v(x) = \min_{a \in A(X)} \left\{ r_t(x, a) + \beta \sum_{y_i} v(y_i) p(y_i \mid x, a) \right\}. \tag{3.21}$$

3.4.1 Ejemplo: sistema de colas con costo descontado.

Para fijar ideas, retomemos el ejemplo del sistema de colas. Usaremos las Ecuaciones de Bellman (3.21) de la observación 3.2.

Si $x = 0$

$$v(0) = \min_{a \in \{0,1\}} \left\{ r(0, a) + \beta \sum_{y=0}^{1} v(y) p(dy \mid 0, a) \right\}$$

$$
\begin{aligned}
&= \min\left\{\begin{array}{c} r(0,0)+\beta\left[v(0)\,p(0\mid 0,0)+v(1)\,p(1\mid 0,0)\right]; \\ r(0,1)+\beta\left[v(0)\,p(0\mid 0,1)+v(1)\,p(1\mid 0,1)\right] \end{array}\right\},
\end{aligned}
$$

$$
\begin{aligned}
\text{como } r(x,a) &= e_a + xpc \\
&= \min\left\{\begin{array}{c} e_0+\beta\left[v(0)(1-p)+v(1)\,p\right]; \\ e_1+\beta\left[v(0)(1-p)+v(1)\,p\right] \end{array}\right\} \\
&= e_0+\beta v(0)(1-p)+\beta v(1)\,p,
\end{aligned}
$$

por lo tanto

$$
v(0)=e_0+\beta v(0)(1-p)+\beta v(1)\,p.
$$

Si $x=1$

$$
\begin{aligned}
v(1) &= \min_{a\in\{0,1\}}\left\{r(1,a)+\beta\sum_{y=0}^{1}v(y)\,p(dy\mid 1,a)\right\} \\
&= \min\left\{\begin{array}{c} r(1,0)+\beta\left[v(0)\,p(0\mid 1,0)+v(1)\,p(1\mid 1,0)\right]; \\ r(1,1)+\beta\left[v(0)\,p(0\mid 1,1)+v(1)\,p(1\mid 1,1)\right] \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0+pc+\beta\left[v(0)\,q^0+v(1)\left(1-q^0\right)\right]; \\ e_1+pc+\beta\left[v(0)\,q^1+v(1)\left(1-q^1\right)\right] \end{array}\right\},
\end{aligned}
$$

por lo tanto

$$
v(1)=\min\left\{\begin{array}{c} e_0+pc+\beta\left[v(0)\,q^0+v(1)\left(1-q^0\right)\right]; \\ e_1+pc+\beta\left[v(0)\,q^1+v(1)\left(1-q^1\right)\right] \end{array}\right\}.
$$

En resumen tenemos:

$$
\begin{aligned}
v(0) &= e_0+\beta v(0)(1-p)+\beta v(1)\,p. \\
v(1) &= \min\left\{\begin{array}{c} e_0+pc+\beta\left[v(0)\,q^0+v(1)\left(1-q^0\right)\right]; \\ e_1+pc+\beta\left[v(0)\,q^1+v(1)\left(1-q^1\right)\right] \end{array}\right\}.
\end{aligned} \tag{3.22}
$$

Para verificar los casos de la última igualdad hagamos:

sí

$$
\begin{aligned}
e_0+pc+\beta\left[v(0)\,q^0+v(1)\left(1-q^0\right)\right] &\leq e_1+pc \\
&\quad +\beta\left[v(0)\,q^1+v(1)\left(1-q^1\right)\right] \\
\beta v(0)\,q^0+\beta v(1)\left(1-q^0\right)-\beta v(0)\,q^1-\beta v(1)\left(1-q^1\right) &\leq e_1-e_0 \\
\beta v(0)\,q^0-\beta v(1)\,q^0-\beta v(0)\,q^1+\beta v(1)\,q^1 &\leq e_1-e_0 \\
-\beta v(0)\left(q^1-q^0\right)+\beta v(1)\left(q^1-q^0\right) &\leq e_1-e_0 \\
\left(v(1)-v(0)\right)\left[\beta(q^1-q^0)\right] &\leq e_1-e_0 \\
v(1)-v(0) &\leq \frac{e_1-e_0}{\beta(q^1-q^0)}.
\end{aligned}
$$

Antes de proseguir, calculemos: $\Delta = v(1) - v(0)$.

Supongamos que elegimos o mejor dicho el mínimo es para

$$v(1) = e_0 + pc + \beta\left[v(0)\,q^0 + v(1)\left(1 - q^0\right)\right].$$

Por lo tanto

$$\begin{aligned}
\Delta = \quad & e_0 + pc + \beta v(0)\,q^0 + \beta v(1) - \beta v(1)\,q^0 \\
& -(e_0 + \beta v(0) - \beta v(0)\,p + \beta v(1)\,p)
\end{aligned}$$

$$\begin{aligned}
\Delta &= pc + \beta\left(v(1) - v(0)\right) - \beta q^0\left(v(1) - v(0)\right) - \beta p\left(v(1) - v(0)\right) \\
&= pc + \beta\Delta - \beta q^0\Delta - \beta p\Delta \\
&= pc + \Delta\left(\beta - \beta q^0 - \beta p\right) \\
\Delta &= \frac{pc}{1 - \beta + \beta q^0 + \beta p}.
\end{aligned}$$

Ahora supongamos que se cumple "$\geq$ " es decir

$$v(1) = e_1 + pc + q^1\beta v(0) + \left(1 - q^1\right)\beta v(1).$$

Prosiguiendo como arriba tenemos:

$$\begin{aligned}
\Delta = \quad & e_1 + pc + \beta v(0)\,q^1 + \beta v(1) - \beta v(1)\,q^1 \\
& -(e_0 + \beta v(0) - \beta v(0)\,p + \beta v(1)\,p)
\end{aligned}$$

$$\begin{aligned}
\Delta &= e_1 - e_0 + pc + \beta\left(v(1) - v(0)\right) - \beta q^1\left(v(1) - v(0)\right) - \beta p\left(v(1) - v(0)\right) \\
&= e_1 - e_0 + pc + \beta\Delta - \beta q^1\Delta - \beta p\Delta \\
&= e_1 - e_0 + pc + \Delta\left(\beta - \beta q^1 - \beta p\right) \\
\Delta &= \frac{e_1 - e_0 + pc}{1 - \beta + \beta q^1 + \beta p}.
\end{aligned}$$

En resumen

$$\Delta = \begin{cases} \frac{pc}{1-\beta+\beta q^0+\beta p}, & \text{si } \frac{pc}{1-\beta+\beta q^0+\beta p} \leq \frac{e_1-e_0}{\beta(q^1-q^0)}; \\ \frac{e_1-e_0+pc}{1-\beta+\beta q^1+\beta p}, & \text{si } \frac{pc}{1-\beta+\beta q^0+\beta p} \geq \frac{e_1-e_0}{\beta(q^1-q^0)}. \end{cases} \tag{3.23}$$

Los valores de $v(0)$ y $v(1)$ pueden ser calculados ahora con la ayuda de la primera ecuación de (3.22).

$$v(0) = e_0 + \beta v(0)(1 - p) + \beta v(1)\,p$$

$$\begin{aligned}
&= e_0 + \beta v(0) - \beta v(0) p + \beta v(1) p \\
&= e_0 + \beta v(0) + p\beta (v(1) - v(0)) \\
&= e_0 + \beta v(0) + p\beta\Delta \\
v(0) - \beta v(0) &= e_0 + p\beta\Delta \\
v(0)(1-\beta) &= e_0 + p\beta\Delta \\
v(0) &= \frac{e_0 + p\beta\Delta}{(1-\beta)} .
\end{aligned}$$

Ahora para $v(1)$

$$\begin{aligned}
v(1) &= \min \left\{ \begin{array}{c} e_0 + pc + \beta [v(0) q^0 + v(1)(1-q^0)] ; \\ e_1 + pc + \beta [v(0) q^1 + v(1)(1-q^1)] \end{array} \right\} \\
&= e_0 + pc + \min \left\{ \begin{array}{l} \beta v(0) q^0 + \beta v(1) - q^0 \beta v(1) ; \\ e_1 - e_0 + \beta v(0) q^1 + \beta v(1) - q^1 \beta v(1) \end{array} \right\} \\
&= e_0 + pc + \min \left\{ \begin{array}{l} -\beta q^0 \Delta + \beta v(1) ; \\ e_1 - e_0 - \beta q^1 \Delta + \beta v(1) \end{array} \right\} \\
&= e_0 + pc + \beta v(1) + \min \left\{ -\beta q^0 \Delta ; e_1 - e_0 - \beta q^1 \Delta \right\} .
\end{aligned}$$

A partir de esta expresión consideremos

1) caso

$$\min \left\{ -\beta q^0 \Delta ; e_1 - e_0 - \beta q^1 \Delta \right\} = -\beta q^0 \Delta,$$

por lo tanto

$$\begin{aligned}
v(1) &= e_0 + pc + \beta v(1) - \beta q^0 \Delta \\
&= \frac{e_0 + pc - \beta q^0 \Delta}{1-\beta} \\
&= \frac{e_0 + p\beta\Delta}{1-\beta} + \frac{pc - p\beta\Delta - \beta q^0 \Delta}{1-\beta} .
\end{aligned}$$

Si

$$\begin{aligned}
\Delta &= \frac{pc - p\beta\Delta - \beta q^0 \Delta}{1-\beta} \\
\Delta \left[(1-\beta) + p\beta + \beta q^0 \right] &= pc \\
\Delta &= \frac{pc}{(1-\beta) + p\beta + \beta q^0},
\end{aligned}$$

por lo tanto

$$v(1) = v(0) + \Delta\,.$$

Para el segundo caso. Sea

$$\min\left\{-\beta q^0\Delta; e_1 - e_0 - \beta q^1\Delta\right\} = e_1 - e_0 - \beta q^1\Delta,$$

por lo tanto

$$\begin{aligned} v(1) &= e_0 + pc + \beta v(1) + e_1 - e_0 - \beta q^1\Delta \\ &= \frac{e_0 + pc + e_1 - e_0 - \beta q^1\Delta}{1-\beta} \\ &= \frac{e_0 + p\beta\Delta}{1-\beta} + \frac{pc + e_1 - e_0 - \beta q^1\Delta - p\beta\Delta}{1-\beta} \end{aligned}$$

sea

$$\begin{aligned} \Delta &= \frac{pc + e_1 - e_0 - \beta q^1\Delta - p\beta\Delta}{1-\beta} \\ \Delta\left[(1-\beta) + \beta q^1 + p\beta\right] &= pc + e_1 - e_0 \\ \Delta &= \frac{pc + e_1 - e_0}{(1-\beta) + \beta q^1 + p\beta}, \end{aligned}$$

por lo tanto

$$v(1) = v(0) + \Delta\,.$$

En resumen

$$v(0) = \frac{e_0 + p\beta\Delta}{(1-\beta)}; \qquad v(1) = v(0) + \Delta. \tag{3.24}$$

De acuerdo con el criterio de óptimalidad del Corolario 3.1 la regla de decisión estacionaria φ^* provee el mínimo en la ecuación de Bellman. Y la respuesta es la siguiente:

$$\begin{aligned} \varphi^*(0) &= 0; \\ \varphi^*(1) &= 0, \text{ si } \frac{pc}{1-\beta+\beta q^0+\beta p} < \frac{e_1 - e_0}{\beta(q^1 - q^0)}; \\ \varphi^*(1) &= 1, \text{ si } \frac{pc}{1-\beta+\beta q^0+\beta p} > \frac{e_1 - e_0}{\beta(q^1 - q^0)}. \end{aligned}$$

En el caso
$$\frac{pc}{1-\beta+\beta q^0+\beta p}=\frac{e_1-e_0}{\beta(q^1-q^0)},$$
todas las estrategias son óptimas en H por lo cual $\pi_t\left(0 \mid x_0, a_1, x_1, \ldots, a_{t-1}, 0\right)=1$; en $h_t=(x_0, a_1, x_1, \ldots, a_{t-1}, 1)$ esto es irrelevante para la acción que se ha elegido.

Capítulo 4

Modelo de control Markoviano con restricciones.

4.1 Introducción.

En este capítulo estudiaremos los procesos de decisión de Markov con restricciones. Para resolver el problema de control óptimo con restricciones usaremos la programación convexa, lo podemos hacer pues la esperanza matemática es un operador lineal y por lo tanto convexo. El método consiste en introducir los multiplicadores de Lagrange, se optimiza la función de Lagrange considerando multiplicadores fijos. Esta optimización de la función de Lagrange la hacemos vía la programación dinámica desarrollada en el capítulo anterior. Después optimizamos sobre los multiplicadores. En otras palabras estamos encontrando el punto silla. En nuestro caso la esperanza del costo es lineal y por lo tanto el óptimo se alcanza en la frontera de la región factible, es decir la región definida por las restricciones.

En el capítulo formulamos los conceptos análogos al capítulo anterior pero con restricciones. Definimos el Teorema de Kuhn-Tucker, la función de Lagrange y el algoritmo de la programación convexa. A lo largo del capítulo continuamos con el ejemplo de colas desarrollado en el capítulo anterior para fijar ideas. Algunas referencias para este capítulo son: [2] y [14].

4.2 Formulación del problema.

4.2.1 El problema de control Markoviano con una restricción.

Sea (X, A, P, r, T) el modelo de control Markoviano propuesto en el capítulo anterior. Sean I^1 e I^2 índices de funcionamiento como definimos en 3.1, 3.2, 3.3 y 3.4.

El problema de control con restricciones en el que estamos interesados [14] es

$$\begin{gathered}\min_{\pi \in \Pi} \mathbf{I}^1 (\pi, x), \\ \text{sujeta a } \mathbf{I}^2 (\pi, x) \leq d.\end{gathered} \tag{4.1}$$

Definición 4.1 *Una estrategia π es llamada admisible si se satisface la desigualdad*

$$\mathbf{I}^2 (\pi, x) \leq d. \tag{4.2}$$

4.2.2 Ejemplo: sistema de colas con restricciones.

Retomaremos el ejemplo planteado en Item 3.3.2; un sistema de colas en una cadena de Markov con rechazo. La función $\mathbf{r}(\cdot)$ representa el consumo del servicio y la función $\mathbf{s}(\cdot)$ designará la penalidad por cliente no atendido. (A lo largo de esta sección abordaremos el caso $N = 1$) Así, designaremos

$$\begin{array}{ll} r_t (x, a) = e_a; & s_t (x, a) = xpc; \\ r_T (x) \;\; = 0; & s_T (x) \;\; = -d; \end{array}$$

Con los índices de funcionamiento:

$$\begin{aligned} J_T^1 (\pi, x) &= E_x^\pi \left[\sum_{t=1}^{T-1} r_t (x_{t-1}, a_t) + r_T (x_T) \right]; \\ J_T^2 (\pi, x) &= E_x^\pi \left[\sum_{t=1}^{T-1} s_t (x_{t-1}, a_t) + s_T (x_T) \right]. \end{aligned}$$

Se trata de mínimizar $J_T^1 (\pi, x)$ sujeto a la restricción de que $J_T^2 (\pi, x) \leq d$; es decir minimizar el consumo del servicio bajo la restricción de que la penalidad por la pérdida de la demanda no sea más grande que d. El valor $d > 0$ es elegido con anterioridad.

4.3 Condiciones necesarias y suficientes de optimalidad.

4.3.1 La función de Lagrange y resultados similares.

Introduciremos la función de Lagrange $L(\cdot)$ por la fórmula

$$L(\pi, l) = J_T^1(\pi, x) + l J_T^2(\pi, x). \tag{4.3}$$

Donde l es llamado el *multiplicador de Lagrange* para el problema de optimización.

Definición 4.2 $g(l) = \inf_{\pi \in \prod} L(\pi, l)$ *es la función sobre* $\mathbf{R}_+$ *llamada la función dual*

$$\omega(l) = \inf \left\{ J_T^1(\pi, x) : \pi \in \prod, \ J_T^2(\pi, x) \leq l \right\}$$

es la función sobre $\mathbf{R}$ *llamada la función primal.*[1]

Además, la siguiente relación dual es válida

$$\inf_{\pi \in \prod} \sup_{l \in \mathbf{R}_+} L(\pi, l) = \sup_{l \in \mathbf{R}_+} \inf_{\pi \in \prod} L(\pi, l) \ \ [14]. \tag{4.4}$$

Kuhn-Tucker (1951) demostró que un punto silla asociado al Lagrangiano L provee una solución a el problema de programación convexa. Su resultado a continuación es presentado.

Teorema 4.1 *Un punto* $\pi^* \in \prod$ *es una solución al problema principal de programación convexa (4.1) si y sólo si existe un número real* $l \in \mathbf{R}_+$ *para el cual una de las dos siguientes afirmaciones equivalentes se cumple:*

(a) el par (π^*, l_*) es un punto silla de la función de Lagrange:

$$L(\pi^*, l) \leq L(\pi^*, l_*) \leq L(\pi, l_*) \tag{4.5}$$

en algún $(\pi, l) \in \prod \times \mathbf{R}_+$.

(b) $J_T^2(\pi^*, x) \leq d$; $L(\pi^*, l_*) = \min_{\pi \in \prod} L(\pi, l_*)$ y la 'condición de holgura complementaria'

$$l J_T^2(\pi, x) = d$$

es válida [2] [14].

[1] Es la equivalente a la función de valor para el caso restringido, con l número real.

4.4 Algoritmo para resolver el problema principal de programación convexa.

En esta sección abordaremos la solución para el problema de control de Markov con restricciones, apoyándonos en el algoritmo de la programación convexa. Es conocido, de la programación convexa, que la solución puede ser planteada, primero como un problema sin restricciones de la siguiente forma:

$$\mathbf{M}(\pi, x) \quad = \quad l_1 J_T^1(\pi, x) + l_2 J_T^2(\pi, x) . \tag{4.6}$$

considerando l_1, l_2 constantes y después optimizamos sobre estos multiplicadores.

Cabe mencionar que se supone conocido un método para resolver el problema no restringido (4.6), en nuestro caso la programación dinámica.

4.4.1 Resultados auxiliares.

Si consideramos un modelo con horizonte finito o el modelo descontado de Markov entonces en el caso de un espacio de estado finito X usaremos el *método de la función de penalidad* para la mininización de la función $\mathbf{M}(\cdot)$ del tipo (4.6). A continuación explicamos el método que usaremos para resolver el problema de control con restricciones, costo descontado y horizonte infinito. Los índices de funcionamiento son:

$$\begin{aligned} v^1(\pi, x) &= E_x^\pi \left[\sum_{t=1}^{\infty} \beta^{t-1} r(x_t, a_t) \right]; \\ v^2(\pi, x) &= E_x^\pi \left[\sum_{t=1}^{\infty} \beta^{t-1} s(x_t, a_t) \right], \end{aligned}$$

donde suponemos que la función $r(\cdot)$ es semicontinua inferiormente y acotada por abajo, y la función $s(\cdot)$ es continua. Sea

$$\widehat{v}(x) = \min_{\pi} \left\{ E_x^\pi \left[\sum_{t=0}^{\infty} \beta^{t-1} m(x_t, a_t) \right] \right\}$$

la función de valor para el índice v. Entonces el problema de encontrar $\pi^* \in \prod$ tal que

$$\widehat{v}(x) = \mathbf{M}(\pi, x)$$

puede ser obtenida con la ayuda del método de programación dinámica, formula (3.21) .

Algoritmo para el problema de control con restricciones.

A continuación describiremos la manera en la que aplicaremos estas ideas a el problema de colas. Cabe mencionar, que todo el procedimiento es recursivo en caso de que haya varias restricciones.

1o. Con la ayuda de la Función de Langrange, deduciremos las Ecuaciones de Bellman.

2o. En la última etapa, para conocer la política óptima consideramos una política Λ, la cual proveera la política mínima en esa etapa.

3o. A partir de este momento conoceremos las reglas de decisión óptimas de Markov.

4o. Deduciremos el mínimo de la función de Lagrange (4.6) suponiendo l_1 y l_2 fijos. Una vez obtenido el resultado procedemos a maximizarlo sobre los l_1 y l_2.

5o. Introducimos el conjunto $\mathcal{D}^*$, el cual esta formado por la combinación convexa de las políticas óptimas.

6o. Construiremos una política $\pi^+ \in \Pi^*$ tal que $J^2(\pi^*) \geq 0$. E introducimos la función $\widetilde{J^2}(\pi) = -J^2(\pi)$ y resolvemos el problema

$$\min_{\pi \in \Pi^*} \left\{ \widetilde{J^2}(\pi) \right\}$$

sin restricciones.

7o. Los valores de los escalares de la combinación convexa estaran designados por:

$$\lambda^* = \frac{-J^2\left(\pi^{\varphi_1^*}\right)}{J^2\left(\pi^{\varphi_0^*}\right) - J^2\left(\pi^{\varphi_1^*}\right)}$$

Con lo anterior habremos solucionado el problema, restringido.

4.5 Ejemplo: sistema de colas con restricciones.

Presentaremos la solución completa del problema formulado en la subsección 4.2.2 para el caso $T = 2$. A pesar de su simplicidad este ejemplo establece la ilustración de todas las ideas teóricas y el método para resolver el problema restringido presentado en la Sección 4.4.

Considere la función de Lagrange

$$L(\pi, Y) = \mathbf{J}_T^1(\pi, x) + l\mathbf{J}^2(\pi, x).$$

En este ejemplo $N = 1$, $l \in \mathbf{R}_+$,

$$\begin{aligned} L(\pi, l) &= E^\pi\left[\sum_{t=1}^{T} j_t^1(x_{t-1}, a_t) + j_T^1(x_T)\right] + lE^\pi\left[\sum_{t=1}^{T} j_t^2(x_{t-1}, a_t) + j_T^2(x_T)\right] \\ &= E^\pi\left[\sum_{t=1}^{T}\left(j_t^1(x_{t-1}, a_t) + lj_t^2(x_{t-1}, a_t)\right) + j_T^1(x_T) + lj_T^2(x_T)\right] \end{aligned}$$

$$\text{como } j_T^1(x_T) = 0$$

tenemos:

$$L(\pi, l) = E^\pi\left[\sum_{t=1}^{T}\left(j_t^1(x_{t-1}, a_t) + lj_t^2(x_{t-1}, a_t)\right) + lj_T^2(x_T)\right]$$

y para toda l fija en $\mathbf{R}_+$ el problema es

$$\min_{\pi\in\prod}\{L(P, l)\}. \tag{4.7}$$

Antes de continuar hagamos un paréntesis para verificar el comportamiento de las Ecuaciones de Bellman, como hemos visto en capítulos anteriores, tenemos:

$$\begin{aligned} v_T(x) &= j_T^1(x) + lj_T^2(x) \\ v_{t-1}(x) &= \min_{a\in A}\left\{j_t^1(x, a) + lj_t^2(x, a) + \sum_{y=0}^{1} v_t(y)\, p(y/x, a)\right\} \end{aligned}$$

de donde

$$\begin{aligned}
v_T(x) &= -ld \\
v_{t-1}(0) &= \min\left\{\begin{array}{c} j_2^1(0,0) + lj_2^2(0,0) + \sum_{y=0}^{1} v_t(y)\, p(y/0,0), \\ j_2^1(0,1) + lj_2^2(0,1) + \sum_{y=0}^{1} v_t(y)\, p(y/0,1) \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0 + v_2(0)\, P(0/0,0) + v_2(1)\, P(1/0,0), \\ e_1 + v_2(0)\, P(0/0,1) + v_2(1)\, P(1/0,1) \end{array}\right\} \\
&= \min\{e_0 - ld(1-p) - ldp, e_1 - ld\} \\
&= -ld + \min\{e_0, e_1\} = -ld + e_0.
\end{aligned}$$

Si $x = 1$

$$\begin{aligned}
v_{t-1}(0) &= \min\left\{\begin{array}{c} j_2^1(1,0) + lj_2^2(1,0) + \sum_{y=0}^{1} v_t(y)\, p(y/1,0), \\ j_2^1(1,1) + lj_2^2(1,1) + \sum_{y=0}^{1} v_t(y)\, p(y/1,1) \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0 + lpc + \sum_{y=0}^{1} v_t(y)\, p(y/1,0), \\ e_1 + lpc + \sum_{y=0}^{1} v_t(y)\, p(y/1,1) \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0 + lpc + v_t(0)\, p(0/1,0) + v_t(1)\, p(1/1,0), \\ e_1 + lpc + v_t(0)\, p(0/1,1) + v_t(1)\, p(1/1,1) \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0 + lpc - ldq^0 - ld(1-q^0), \\ e_1 + lpc - ldq^1 - ld(1-q^1) \end{array}\right\} \\
&= lpc + \min\{e_0 - ld, e_1 - ld\} \\
&= lpc - ld + e_0
\end{aligned}$$

para $t = 1$ y $x = 0$

$$\begin{aligned}
v_0(0) &= \min\left\{\begin{array}{c} j_1^1(0,0) + lj_1^2(0,0) + \sum_{y=0}^{1} v_1(y)\, p(y/0,0), \\ j_1^1(0,1) + lj_1^2(0,1) + \sum_{y=0}^{1} v_1(y)\, p(y/0,1) \end{array}\right\} \\
&= \min\left\{\begin{array}{c} e_0 + v_1(0)\, P(0/0,0) + v_1(1)\, P(1/0,0), \\ e_1 + v_1(0)\, P(0/0,1) + v_1(1)\, P(1/0,1) \end{array}\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \min\left\{ \begin{array}{l} e_0 + (-ld + e_0)(1-p) + (lpc - ld + e_0)\,p, \\ e_1 + (-ld + e_0)(1-p) + (lpc - ld + e_0)\,p \end{array} \right\} \\
&= e_0 + (-ld + e_0)(1-p) + (lpc - ld + e_0)\,p \\
&= e_0 - ld + e_0 + ldp - e_0 p + yp^2 c - ldp + e_0 p \\
&= 2e_0 - ld + lp^2 c
\end{aligned}
$$

es decir en la alternativa 0 en la etapa 0 y $x = 0$.

Para $t = 1$ y $x = 1$

$$
\begin{aligned}
v_0(1) &= \min\left\{ \begin{array}{l} j_1^1(1,0) + lj_1^2(1,0) + \sum\limits_{y=0}^{1} v_1(y)\,p(y/1,0), \\ j_1^1(1,1) + lj_1^2(1,1) + \sum\limits_{y=0}^{1} v_1(y)\,p(y/1,1) \end{array} \right\} \\
&= \min\left\{ \begin{array}{l} e_0 + lpc + v_1(0)\,P(0/1,0) + v_1(1)\,P(1/1,0), \\ e_1 + lpc + v_1(0)\,P(0/1,1) + v_1(1)\,P(1/1,1) \end{array} \right\} \\
&= \min\left\{ \begin{array}{l} e_0 + lpc + (-ld + e_0)\,q^0 + (lpc - ld + e_0)\left(1 - q^0\right), \\ e_1 + lpc + (-ld + e_0)\,q^1 + (lpc - ld + e_0)\left(1 - q^1\right) \end{array} \right\} \\
&= lpc + \min\left\{ \begin{array}{l} e_0 + lpc - ld + e_0 - lpcq^0, \\ e_1 + lpc - ld + e_0 - lpcq^1 \end{array} \right\} \\
&= 2lpc - ld + e_0 + \min\left\{ e_0 - lpcq^0, e_1 - lpcq^1 \right\}
\end{aligned}
$$

para conocer cual política conviene, hacemos:

$$
\begin{aligned}
e_0 - lpcq^0 &< e_1 - lpcq^1 \\
-lpc\left(q^0 - q^1\right) &< e_1 - e_0 \\
lpc\left(q^1 - q^0\right) &< e_1 - e_0 \\
lpc &< \Lambda.
\end{aligned}
$$

En resumen

$$v_T(x) = j_T^1(x) + lj_T^2(x)$$

$$v_2(x) = -ld \text{ si } x = 0, 1.$$

$$v_1(0) = -ld + e_0;\ v_1(1) = lpc - ld + e_0.$$

$$
\begin{aligned}
v_0(0) &= 2e_0 - ld + p^2 lc; \\
v_0(1) &= 2lpc - ld + e_0 + \min\left\{ e_0 - lpcq^0, e_1 - lpcq^1 \right\}
\end{aligned}
$$

de esta última igualdad se desprenden dos casos.

(*a*) Si $lpc > \mathbf{\Lambda}$

$$v_0(1) = 2lpc - ld + e_0 + e_1 - lpcq^1.$$

(*b*) Si $lpc < \mathbf{\Lambda}$

$$\begin{aligned} v_0(1) &= 2lpc - ld + e_0 + e_0 - lpcq^0 \\ &= 2lpc - ld + 2e_0 - lpcq^0. \end{aligned}$$

Considere que $1 - p - q^0 > 0$; $P_0(0) \in (0,1)$. La solución al problema (4.7) toma la siguiente forma (donde $\mathbf{\Lambda} = (e_1 - e_0)/(q^1 - q^0)$):

(a) Si $plc > \mathbf{\Lambda}$, entonces

$$\begin{aligned} \min_{\pi\in\Pi} L(\pi, l) &= g(l) \qquad (4.8) \\ &= P_0(0)\,v_0(0) + P_0(1)\,v_0(1) \\ &= P_0(0)\left[2e_0 - ld + p^2lc\right] + P_0(1)\left[2lpc - ld + e_0 + e_1 - lpcq^1\right] \\ &= -ld + e_0 + P_0(0)\left(e_0 + lp^2c\right) + P_0(1)\left[e_1 + 2lpc - lpcq^1\right] \end{aligned}$$

el mínimo es alcanzado sólo por $\pi^* = \pi^{\varphi_1^*}$, donde φ_1^* es la regla de decisión de Markov definida por la fórmula:

$$\varphi_1^*(1,1) = 1, \quad \varphi_1^*(2,1) = 0, \quad \varphi_1^*(1,0) = \varphi_1^*(2,0) = 0.$$

(b) Si $plc < \mathbf{\Lambda}$, entonces

$$\begin{aligned} g(l) &= P_0(0)\left[2e_0 - ld + p^2lc\right] + P_0(1)\left[2lpc - ld + 2e_0 - lpcq^0\right] \qquad (4.9) \\ &= 2e_0 - ld + P_0(0)\,p^2lc + P_0(1)\left(2plc - q^0plc\right); \end{aligned}$$

el mínimo es alcanzado sólo por $\pi^* = \pi^{\varphi_0^*}$, donde φ_0^* es el selector de Markov

$$\varphi_0^*(t,x) \equiv 0.$$

(c) Si $plc = \mathbf{\Lambda}$, entonces $g(l)$ esta expresada por ambas fórmulas (4.8) y (4.9): sus valores coinciden. El mínimo en el problema (4.7) es alcanzado por estas y sólo estos puntos $\widehat{\pi}$ los cuales corresponden a las estrategias del siguiente tipo

$$\widehat{\pi}_2(0 \mid h_1) = 1; \quad \widehat{\pi}_1(0 \mid 0) = 1.$$

Note que las desigualdes

$$P^{\pi}\{x_t = 0\} > 0, \qquad P^{\pi}\{x_t = 1\} > 0, \qquad t = 0, 1, 2$$

se tienen para alguna estrategia π. (recordemos que $P_0(0) \in (0,1)$.) Por lo tanto, según el criterio de óptimalidad, la estrategia medible presentada en Items (a)-(c) completamente exhaustivas al conjunto de soluciones del problema (4.7). Ciertamente, el caso (c) es el más interesante: aquí las estrategias $\widehat{\pi}$ generan (este es un prisma de seis dimensiones) los puntos extremos que corresponden a los selectores π_0^* y π_1^*.

Existen cinco alternativas diferentes de equivalencias de los cambios de la función $g(l) = \min_{\pi \in \prod} L(\pi, l)$ dependiendo de los valores de los parámetros. Antes de mostrarlas las deduciremos de (4.8) y (4.9). De (4.8) tenemos:

$$\begin{aligned}
\min_{\pi \in \prod} L(\pi, l) &= g(l) \\
&= P_0(0)\, v_0(0) + P_0(1)\, v_0(1) \\
&= P_0(0)\left[2e_0 - ld + p^2 lc\right] + P_0(1)\left[2lpc - ld + e_0 + e_1 - lpcq^1\right] \\
&= -ld + e_0 + P_0(0)\left(e_0 + lp^2 c\right) + P_0(1)\left[e_1 + 2lpc - lpcq^1\right] \\
&= 2e_0 P_0(0) - P_0(0)\, ld + P_0(0)\, lp^2 c + P_0(1)\, e_0 + P_0(1)\, e_1 - \\
&\quad P_0(1)\, ld + P_0(1)\, 2lpc - P_0(1)\, lpcq^1 \\
&= 2e_0 P_0(0) + e_0 P_0(1) + P_0(1)\, e_1 \\
&\quad + l\left[-P_0(0)\, d + P_0(0)\, p^2 c - P_0(1)\, d + P_0(1)\, 2pc - P_0(1)\, pcq^1\right] \\
&= 2e_0 P_0(0) + e_0 P_0(1) + P_0(1) \\
&\quad e_1 + l\left[-d + pc\left(pP_0(0) + 2P_0(1) - P_0(1)\, q^1\right)\right].
\end{aligned}$$

De (4.9), tenemos:

$$\begin{aligned}
g(l) &= P_0(0)\left[2e_0 - ld + p^2 lc\right] + P_0(1)\left[2lpc - ld + 2e_0 - lpcq^0\right] \\
&= 2e_0 - ld + P_0(0)\, p^2 lc + P_0(1)\left(2plc - q^0 plc\right) \\
&= 2e_0 + l\left[-P_0(0)\, d + P_0(0)\, p^2 c + P_0(1)\, 2pc - P_0(1)\, d - P_0(1)\, pcq^0\right] \\
&= 2e_0 + l\left[-d + P_0(0)\, p^2 c + P_0(1)\, 2pc - P_0(1)\, pcq^0\right] \\
&= 2e_0 + l\left[-d + pc\left[P_0(0)\, p + P_0(1)\, 2 - P_0(1)\, q^0\right]\right] \\
&= 2e_0 + l\left[-d + pc\left[P_0(0)\, p + P_0(1)\left(2 - q^0\right)\right]\right].
\end{aligned}$$

En resumen, a continuación mostramos los cinco casos:

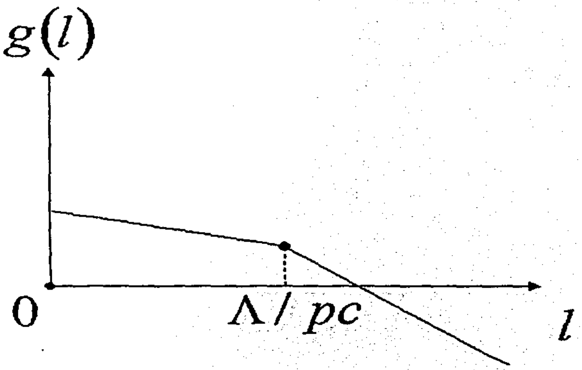

Figura 4.1: Alternativa (1).

(1) $d > pc\left[P_0\left(0\right)p + P_0\left(1\right)\left(2 - q^0\right)\right]$;
(2) $d = pc\left[P_0\left(0\right)p + P_0\left(1\right)\left(2 - q^0\right)\right]$;
(3) $pc\left[pP_0\left(0\right) + P_0\left(1\right)\left(2 - q^1\right)\right] < d < pc\left[P_0\left(0\right)p + P_0\left(1\right)\left(2 - q^0\right)\right]$;
(4) $d = pc\left[pP_0\left(0\right) + P_0\left(1\right)\left(2 - q^1\right)\right]$;
(5) $d < pc\left[pP_0\left(0\right) + P_0\left(1\right)\left(2 - q^1\right)\right]$.

Los diagramas correspondientes son presentados en las figuras 4.1–4.5. En el primero de los cuatro casos el problema

$$\max_{l \in \mathbb{R}_+} \left\{g\left(l\right)\right\}$$

puede simplificar la solución; en el quinto caso este problema es irresoluble. Investigaremos esta alternativa con más detalle. Obviamente,

$$\min_{\pi \in \prod} \mathbf{J}^2\left(\pi, x\right) = pc\left[pP_0\left(0\right) + P_0\left(1\right)\left(2 - q^1\right)\right] - d$$

y su mínimo es alcanzado por la regla de decisión $\varphi^*\left(t, x\right) = 1$ por ejemplo. Ahora, esta claro que la alternativa (5) corresponde a el caso en el cual no existen estrategias admisibles óptimas:

$$\left\{\pi \in \prod : \mathbf{J}^2\left(\pi\right) \leq 0\right\} = \emptyset.$$

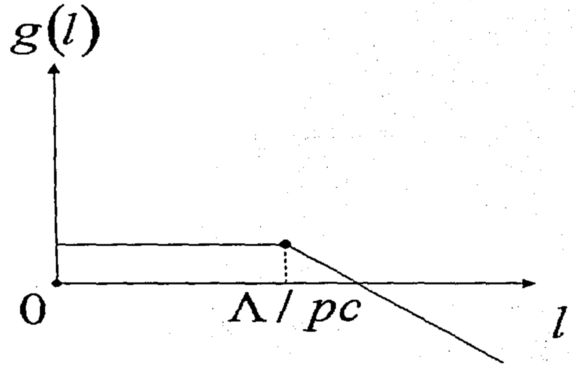

Figura 4.2: Alternativa (2).

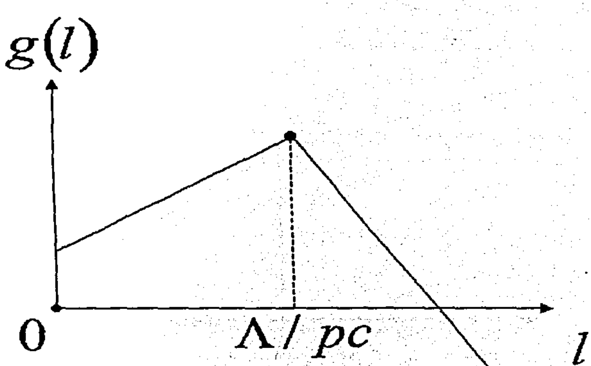

Figura 4.3: Alternativa (3).

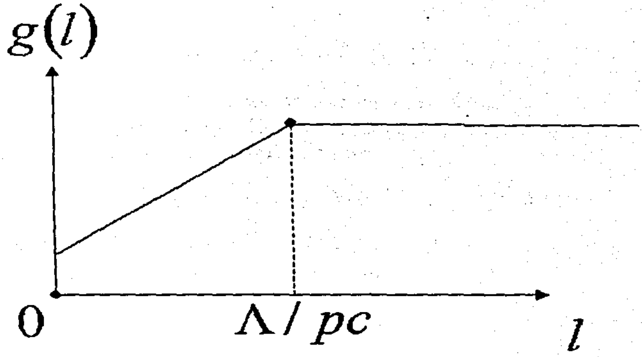

Figura 4.4: Alternativa (4).

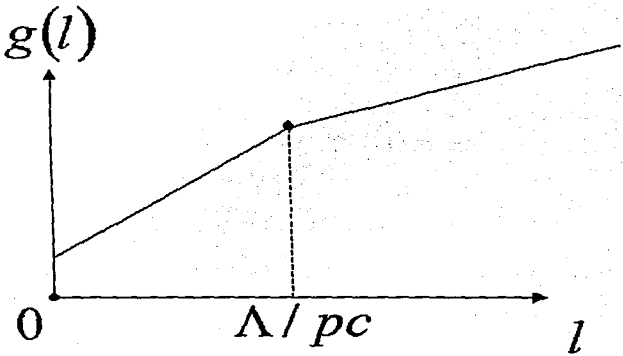

Figura 4.5: Alternativa (5).

Es por lo que la quinta alternativa no será considerada en lo siguiente. Hablando ampliamente, la igualdad

$$d = pc\left[pP_0(0) + P_0(1)\left(2 - q^1\right)\right]$$

no es válida sin embargo, no descartaremos esté caso.

Finalizaremos investigando la alternativa (3). Donde $l_* = \Lambda/pc > 0$; De acuerdo al caso (c) estudiado arriba, tenemos

$$\mathcal{D}^* = \left\{\lambda\pi^{\varphi_0^*} + (1-\lambda)\pi^{\varphi_1^*},\ \lambda \in [0,1]\right\}.$$

Construiremos un punto $\pi^+ \in \prod^*$ tal que $\mathbf{J}^2(\pi^+) \geq 0$. E introducimos la función $\widetilde{\mathbf{J}}^2(\pi) = -\mathbf{J}^2(\pi)$ y resolvemos el problema:

$$\min_{\pi\in\prod^*}\left\{\widetilde{\mathbf{J}}^2(\pi)\right\} \tag{4.10}$$

sin restricciones.De acuerdo a el método de la función de penalidad descrito en la Subsección 4.4.1, tenemos

$$\begin{array}{ll}
j_2^2(x,x) = d & \\
j_1^2(0,0) = 0 & j_1^2(0,1) = +\infty, \\
j_1^2(1,0) = -pc & j_1^2(1,1) = +\infty, \\
j_0^2(0,0) = 0 & j_0^2(0,1) = +\infty, \\
j_0^2(1,0) = +\infty & j_0^2(1,1) = -pc
\end{array}$$

por lo que

$$\widetilde{j}_t^2(0,1) = +\infty, \quad \widetilde{j}_2^2(1,1) = +\infty$$

recordemos que los otros valores son

$$\widetilde{j}_t^2(0,0) = 0; \quad \widetilde{j}_2^2(1,0) = -pc$$

no cambian. Cuando resolvemos el problema (4.10) en la clase $\pi \in \prod$ por el método de programación dinámica obtenemos

$\widetilde{v}_2(x) \equiv d$;

para $\widetilde{v}_1(0)$ hacemos

$$\widetilde{v}_1(0) \quad = \quad \min_{a\in A}\left\{j_1^2(0,a) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 0,a)\right\},$$

$$\begin{aligned}
&= \min\left\{ j_1^2(0,0) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 0,0), \right. \\
&\qquad \left. j_1^2(0,1) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 0,1) \right\} \\
&= \min\{ j_1^2(0,0) + j_2^2(0)\, P(0 \mid 0,0) + j_2^2(1)\, P(1 \mid 0,0), \\
&\qquad j_1^2(0,1) + j_2^2(0)\, P(0 \mid 0,1) + j_2^2(1)\, P(1 \mid 0,1)\}. \\
&= \min\{ d(1-p) - dp, \infty \} = d \text{ (alternativa 0)}
\end{aligned}$$

para $\widetilde{v}_1(1)$

$$\begin{aligned}
\widetilde{v}_1(1) &= \min_{a \in A}\left\{ j_1^2(1,a) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 1,a) \right\}, \\
&= \min\left\{ s_1^2(1,0) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 1,0), \right. \\
&\qquad \left. j_1^2(1,1) + \sum_{y=0}^{1} j_2^2(y)\, P(y \mid 1,1) \right\} \\
&= \min\{ j_1^2(1,0) + j_2^2(0)\, P(0 \mid 1,0) + j_2^2(1)\, P(1 \mid 1,0), \\
&\qquad j_1^2(1,1) + j_2^2(0)\, P(0 \mid 1,1) + j_2^2(1)\, P(1 \mid 1,1)\} \\
&= \min\{ -pc + dq^0 + d(1-q^0), \infty \} = -pc + dq^0 + d - dq^0 \\
&= -pc + d. \text{ (alternativa 0)}.
\end{aligned}$$

Ahora para $\widetilde{v}_0(0)$

$$\begin{aligned}
\widetilde{v}_0(0) &= \min_{a \in A}\left\{ j_0^2(0,a) + \sum_{y=0}^{1} j_1^2(y)\, P(y \mid 0,a) \right\}, \\
&= \min\left\{ j_0^2(0,0) + \sum_{y=0}^{1} j_1^2(y)\, P(y \mid 0,0), \right. \\
&\qquad \left. j_0^2(0,1) + \sum_{y=0}^{1} j_1^2(y)\, P(y \mid 0,1) \right\} \\
&= \min\{ j_0^2(0,0) + j_1^2(0)\, P(0 \mid 0,0) + j_1^2(1)\, P(1 \mid 0,0), \\
&\qquad j_0^2(0,1) + j_1^2(0)\, P(0 \mid 0,1) + j_1^2(1)\, P(1 \mid 0,1)\}
\end{aligned}$$

$$
\begin{aligned}
&= \min\{0 + d(1-p) + (-pc+d)\,p, +\infty\} \\
&= d - p^2 c.
\end{aligned}
$$

Para $\widetilde{v}_0(1)$

$$
\begin{aligned}
\widetilde{v}_0(0) &= \min_{a\in A}\left\{ j_0^2(1,a) + \sum_{y=0}^{1} v_1(y)\, P(y \mid 1,a) \right\}, \\
&= \min\left\{ j_0^2(1,0) + \sum_{y=0}^{1} v_1(y)\, P(y \mid 1,0), \right. \\
&\qquad \left. j_0^2(1,1) + \sum_{y=0}^{1} v_1(y)\, P(y \mid 1,1) \right\} \\
&= \min\{ j_0^2(1,0) + v_1(0)\, P(0 \mid 1,0) + v_1(1)\, P(1 \mid 1,0), \\
&\qquad j_0^2(1,1) + v_1(0)\, P(0 \mid 1,1) + v_1(1)\, P(1 \mid 1,1)\} \\
&= \min\left\{-pc + dq^0 + (-pc+d)\left(1-q^0\right), \infty\right\} \\
&= \min\left\{-pc + dq^0 - pc + pcq^0 + d - dq^0, \infty\right\} \\
&= -pc + dq^0 - pc + pcq^0 + d - dq^0 \\
&= d + pcq^0 - 2pc. \text{ (alternativa 0).}
\end{aligned}
$$

En resumen

$$
\begin{array}{ll}
\widetilde{v}_2(x) \equiv d; & \\
\widetilde{v}_1(0) = d; & \widetilde{v}_1(1) = -pc + d; \\
\widetilde{v}_0(0) = d - p^2 c; & \widetilde{v}_0(1) = d - 2pc + q^0 pc.
\end{array}
$$

Concurrentemente vemos que el punto $P^+ = P^{\varphi_0^*}$ es la única solución al problema (4.10):

$$
\min_{\pi\in\prod^*} \widetilde{\mathbf{J}}^2(\pi) = \widetilde{\mathbf{J}}^2\left(\pi^{\varphi_0^*}\right) = d - pc\left[P_0(0)\,p + P_0(1)\left(2-q^0\right)\right].
$$

En el caso considerado

$$
d < pc\left[P_0(0)\,p + P_0(1)\left(2-q^0\right)\right],
$$

de esta manera la siguiente desigualdad es valida: $\widetilde{\mathbf{J}}^2(\pi^+) < 0$.

De manera similar construiremos el punto $\pi^- \in \prod^*$ para el cual $\mathbf{J}^2(\pi^-) < 0$; de hecho el punto será $\pi^- = \pi^{\varphi_1^*}$,

$\widetilde{v}_2(x) \equiv -d;$
para $\widetilde{v}_1(0)$ hacemos

$$\begin{aligned}
\widetilde{v}_1(0) &= \min_{a\in A}\left\{j_1^2(0,a)+\sum_{y=0}^{1} v_2(y)\,P(y\mid 0,a)\right\}\\
&= \min\left\{j_1^2(0,0)+\sum_{y=0}^{1} v_2(y)\,P(y\mid 0,0),\right.\\
&\qquad\left. j_1^2(0,1)+\sum_{y=0}^{1} v_2(y)\,P(y\mid 0,1)\right\}\\
&= \min\{j_1^2(0,0)+v_2(0)\,P(0\mid 0,0)+v_2(1)\,P(1\mid 0,0),\\
&\qquad j_1^2(0,1)+v_2(0)\,P(0\mid 0,1)+v_2(1)\,P(1\mid 0,1)\}\\
&= \{-d(1-p)-dp,-d(1-p)-dp\}=-d
\end{aligned}$$

con $\widetilde{v}_1(1)$

$$\begin{aligned}
\widetilde{v}_1(1) &= \min_{a\in A}\left\{j_1^2(1,a)+\sum_{y=0}^{1} v_2(y)\,P(y\mid 1,a)\right\}\\
&= \min\{j_1^2(1,0)+v_2(0)\,P(0\mid 1,0)+v_2(1)\,P(1\mid 1,0),\\
&\qquad j_1^2(1,1)+v_2(0)\,P(0\mid 1,1)+v_2(1)\,P(1\mid 1,1)\}\\
&= \min\left\{pc-dq^0-d\left(1-q^0\right),pc-dq^1-d\left(1-q^1\right)\right\}\\
&= -d+pc
\end{aligned}$$

para $\widetilde{v}_0(0)$

$$\begin{aligned}
\widetilde{v}_0(0) &= \min_{a\in A}\left\{j_0^2(0,a)+\sum_{y=0}^{1} v_1(y)\,P(y\mid 0,a)\right\}\\
&= \min\{j_0^2(0,0)+v_1(0)\,P(0\mid 0,0)+v_1(1)\,P(1\mid 0,0),\\
&\qquad j_0^2(0,1)+v_1(0)\,P(0\mid 0,1)+v_1(1)\,P(1\mid 0,1)\}\\
&= \min\{-d(1-p)+(pc-d)\,p,-d(1-p)+(pc-d)\,p\}\\
&= -d+p^2c
\end{aligned}$$

y para $\widetilde{v}_0(1)$

$$\widetilde{v}_0(1) = \min_{a\in A}\left\{j_0^2(1,a)+\sum_{y=0}^{1} v_1(y)\,P(y\mid 1,a)\right\}$$

$$
\begin{aligned}
&= \min\left\{j_0^2(1,0) + v_1(0)\,P(0 \mid 1,0) + v_1(1)\,P(1 \mid 1,0)\,,\right. \\
&\qquad \left. j_0^2(1,1) + v_1(0)\,P(0 \mid 1,1) + v_1(1)\,P(1 \mid 1,1)\right\} \\
&= \min\left\{pc - dq^0 + (pc-d)\left(1-q^0\right),\right. \\
&\qquad \left. pc - dq^1 + (pc-d)\left(1-q^1\right)\right\}. \\
&= \min\left\{2pc - d - q^0pc, 2pc - d - q^1pc\right\},
\end{aligned}
$$

para conocer el mínimo, suponemos la siguiente desigualdad

$$
\begin{aligned}
2pc - d - q^0pc &< 2pc - d - q^1pc \\
-q^0pc &< q^1pc \\
\text{como } q^0 &< q^1 \\
\text{entonces -}q^0 &> -q^1
\end{aligned}
$$

por lo tanto

$$\widetilde{v}_0(1) = 2pc - d - q^1pc.$$

Considerando la función de óptimalidad, tenemos:

$$P_0(0)\,\widetilde{v}_0(0) + P_0(1)\,\widetilde{v}_0(1)$$

$$
\begin{aligned}
&P_0(0)\left(-d+p^2c\right) + P_0(1)\left(2pc-d-q^1pc\right) \\
&-d + p^2cP_0(0) + 2pcP_0(1) - q^1pcP_0(1) \\
&-d + pc\left(pP_0(0) + 2P_0(1) - q^1P_0(1)\right) \\
&-d + pc\left[pP_0(0) + P_0(1)\left(2-q^1\right)\right].
\end{aligned}
$$

Es decir

$$\min_{\pi\in\Pi^*} \mathbf{J}^2(\pi) = \mathbf{J}^2\left(\pi^{\varphi_1^*}\right) = -d + pc\left[P_0(0)\,p + P_0(1)\left(2-q^1\right)\right].$$

Así tenemos las siguientes igualdades para los puntos extremos del segmento $\mathcal{D}^*$:

$$
\begin{aligned}
\mathbf{J}^2\left(\pi^{\varphi_0^*}\right) &= pc\left[P_0(0)\,p + P_0(1)\,(2-q^0)\right] - d; \\
\mathbf{J}^2\left(\pi^{\varphi_1^*}\right) &= pc\left[P_0(0)\,p + P_0(1)\,(2-q^1)\right] - d;
\end{aligned}
\tag{4.11}
$$

los diagramas son presentados en la figura 4.6.

La λ^* estará designada por:

$$\lambda^* = \frac{-\mathbf{J}^2\left(\pi^{\varphi_1^*}\right)}{\mathbf{J}^2\left(\pi^{\varphi_0^*}\right) - \mathbf{J}^2\left(\pi^{\varphi_1^*}\right)}; \tag{4.12}$$

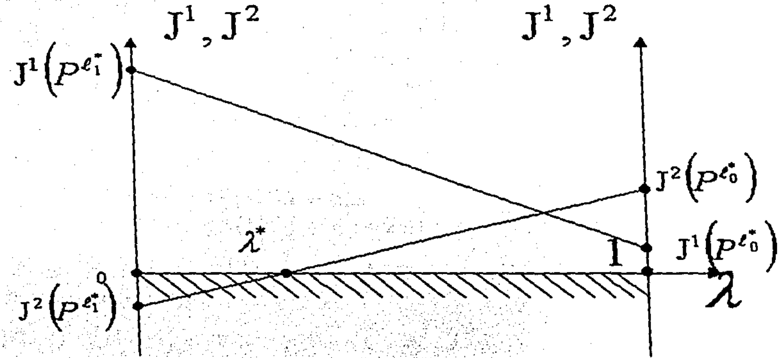

Figura 4.6: Las relaciones entre $\mathbf{J}^2\left(\lambda P^{\varphi_0^*} + (1-\lambda) P^{\varphi_1^*}\right)$ y $\mathbf{J}^1\left(\lambda P^{\varphi_0^*} + (1-\lambda) P^{\varphi_1^*}\right)$ para la alternativa (3).

entonces el punto

$$\pi^* = \lambda^* \pi^{\varphi_0^*} + (1-\lambda^*)\,\pi^{\varphi_1^*} \tag{4.13}$$

resuelve el problema inicial. Los valores $\mathbf{J}^1\left(\pi^{\varphi_0^*}\right)$, $\mathbf{J}^1\left(\pi^{\varphi_1^*}\right)$ son mostrados en la fig. 4.6 se pueden calcular fácilmente, de la siguiente manera:

$$\begin{aligned}
\mathbf{J}^1\left(\pi^{\varphi_0^*}\right) &= E^{\pi}\left[j^1(x_0,a_0) + j^1(x_1,a_1)\right] \\
&= P_0(0)\,j^1(0,0) + P_0(1)\,j^1(1,0) + P_0(0)\,P(1/0,0)\,j^1(1,0) \\
&\quad + P_0(0)\,P(0/0,0)\,j^1(0,0) + P_0(1)\,P(1/1,0)\,j^1(1,0) \\
&\quad + P_0(1)\,P(0/1,0)\,j^1(0,0) \\
&= P_0(0)\,e_0 + P_0(1)\,e_0 + P_0(0)\,pe_0 + P_0(0)\,(1-p)\,e_0 \\
&\quad + P_0(1)\left(1-q^0\right)e_0 + P_0(1)\,q^0 e_0 \\
&= P_0(0)\left[e_0 + pe_0 + (1-p)\,e_0\right] + P_0(1)\left[e_0 + \left(1-q^0\right)e_0 + q^0 e_0\right] \\
&= P_0(0)\,2e_0 + P_0(1)\,2e_0. = 2e_0.
\end{aligned}$$

En este caso las reglas de decisión se toman de acuerdo a la etapa.

Ahora para $\mathbf{J}^1\left(\pi^{\varphi_1^*}\right)$

$$\mathbf{J}^1\left(\pi^{\varphi_1^*}\right) = E^{\pi}\left[j^1(x_0,a_0) + j^1(x_1,a_1)\right]$$

$$
\begin{aligned}
&= P_0(0)\left[j^1(0,0)+P(1/0,0)\,j^1(1,0)+P(0/0,0)\,j^1(0,0)\right] \\
&\quad +P_0(1)\left[j^1(1,1)+P(1/1,1)\,j^1(1,0)+P(0/1,1)\,j^1(0,0)\right] \\
&= P_0(0)\left[e_0+pe_0+(1-p)\,e_0\right]+P_0(1)\left[e_1+\left(1-q^1\right)e_0+q^1e_0\right] \\
&= P_0(0)\,2e_0+P_0(1)\left[e_1+e_0\right] \\
&= e_0+P_0(0)\,e_0+P_0(1)\,e_1.
\end{aligned}
$$

En resumen tenemos:

$$
\begin{aligned}
\mathbf{J}^1\left(\pi^{\varphi_0^*}\right) &= 2e_0. \\
\mathbf{J}^1\left(\pi^{\varphi_1^*}\right) &= e_0+P_0(0)\,e_0+P_0(1)\,e_1.
\end{aligned}
\qquad (4.14)
$$

El valor minímo $\mathbf{J}^1(\pi)$, $\pi\in\prod$ bajo la restricción $\mathbf{J}^2(\pi)\leq 0$ se cálcula de la siguiente forma:

$$
\begin{aligned}
g(l^*) &= J^1(\pi^*)=\lambda^*J^1(\varphi_0^*)+(1-\lambda^*)\,J^1(\varphi_1^*) \qquad (4.15)\\
&= pcP_0(1)\left[2q^1e_0-\left(2+q^0\right)e_0+\left(2-q^0\right)\left[P_0(0)\,e_0+P_0(1)\,e_1\right]+pP_0(0)\,e_1\right] \\
&\quad +pc\left[pP_0(0)\,e_0\left(1+P_0(0)\,e_0\right)\right]-d\left(e_0+P_0(0)\,e_0+P_0(1)\,e_1\right).
\end{aligned}
$$

Claramente, en el caso en consideración (alternativa (3)) es satisfecho.Por lo que, $-l_*=-\Lambda/pc$. La función $\phi(Y)$ será construida como sigue; considerando a la restricción $\mathbf{J}^2(\pi)\leq 0$ como esencial. Obviamente, si las desigualdades

$$
pc\left[P_0(0)\,p+P_0(1)\left(2-q^1\right)\right]<d+l<pc\left[P_0(0)\,p+P_0(1)\left(2-q^0\right)\right]
$$

sostienen, que son equivalentes a $\mathbf{J}^2\left(\pi^{\varphi_1^*}\right)<l<\mathbf{J}^2\left(\pi^{\varphi_0^*}\right)$, entonces la solución a el problema

$$
\min\left\{\mathbf{J}^1(P)\right\},\qquad \pi\in\prod,\ \mathbf{J}^2(\pi)\leq l
$$

es realizada como arriba. Mientras esto tiene lugar la desigualdad se reduce a la igualdad, y el mínimo valor del criterio $\mathbf{J}^1(\cdot)$ coincide con $\phi(l)$ por definición esta expresado por la fórmula

$$
\phi(l)=2e_0+\frac{e_1+e_0}{q^1+q^0}\left[P_0(0)\,p+P_0(1)\left(2-q^0\right)-\frac{d+l}{pc}\right],
$$

lo cual esta mostrada en la figura 4.7. Como ha sido mencionado, $\min\limits_{\pi\in\prod}\mathbf{J}^2(\pi)=\mathbf{J}^2\left(\pi^{\varphi_1^*}\right)$, así la función primal es igual a '$+\infty$' en $l<\mathbf{J}^2\left(\pi^{\varphi_1^*}\right)$. Si $l=$

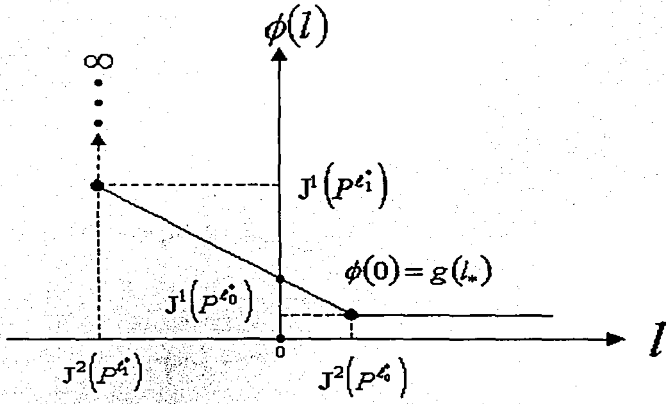

Figura 4.7: Gráfica de la función primal para la alternativa (3).

$J^2(\pi^{\varphi_1^*})$ entonces es demasiado obvio que $\phi(l) = J^1(\pi^{\varphi_1^*})$. La prueba correcta puede ser realizada despuês de introducir el conjunto

$$\mathcal{D}^1 = \left\{\widehat{\pi} : J^2(\widehat{\pi}) = J^2(\pi^{\varphi_1^*})\right\}$$

y resolver el problema $\min\limits_{\pi\in\prod^1} \{J^1(P)\}$, por ejemplo con la ayuda de el método de la función de penalidad. Más precisamente; uno puede rápidamente notar que $J^2(\pi^{\varphi_0^*}) = \max\limits_{\pi\in\prod} J^2(\pi)$; $J^1(\pi^{\varphi_0^*}) = \min\limits_{\pi\in\prod} J^1(\pi)$; así $\phi(l) = J^1(\pi^{\varphi_0^*})$ en $l \geq J^2(\pi^{\varphi_0^*})$. La investigación de la alternativa (3) esta completa.

Consideremos la alternativa (1). Mencionamos anteriormente que

$$\max_{\pi\in\prod} J^2(\pi) = J^2(\pi^{\varphi_0^*}) .$$

Ahora, esta claro que en esta situación uno puede ignorar la restricción, y la solución del problema inicial esta dada por la regla de decisión $\varphi_0^*(t,x) \equiv 0$ para la cual $J^1(\pi^{\varphi_0^*}) = \min\limits_{\pi\in\prod} J^1(\pi)$. A continuación presentamos ilustraciones similares a las figuras 4.6 y 4.7, las cuales se muestran en la fig. 4.8.

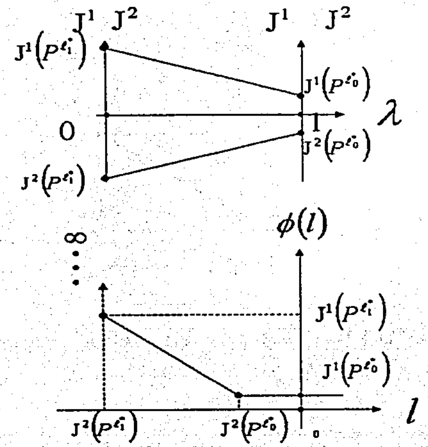

Figura 4.8: Las relaciones entre $\mathbf{J}^2\left(\lambda P^{\varphi_0^*} + (1-\lambda)\, P^{\varphi_1^*}\right)$ y $\mathbf{J}^1\left(\lambda P^{\varphi_0^*} + (1-\lambda)\, P^{\varphi_1^*}\right)$ y la gráfica de la función primal de la alternativa (1).

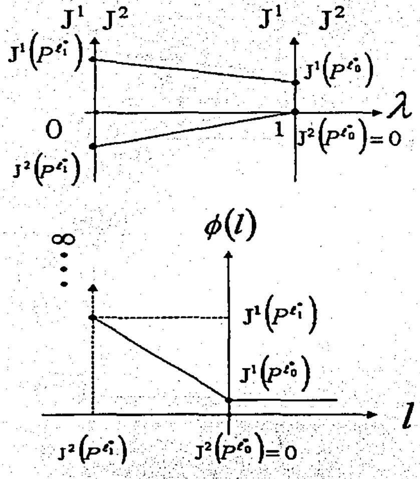

Figura 4.9: Las relaciones entre $\mathrm{J}^2\left(\lambda P^{\varphi_0^*} + (1-\lambda) P^{\varphi_1^*}\right)$ y $\mathrm{J}^1\left(\lambda P^{\varphi_0^*} + (1-\lambda) P^{\varphi_1^*}\right)$ y la gráfica de la función primal de la alternativa (2).

Considere la alternativa (2). Este no es conceptualmente diferente a el caso considerado anteriormente: la restricción es inecesaria, y la solución al problema inicial esta dado por la regla de decisión $\varphi_0^*(t,x) \equiv 0$. Uno puede tomar algún valor en el segmento $[0, \Lambda/pc]$ para l_* y dar la respuesta $\pi^{\varphi_0^*}$, pero en diferentes maneras. Por ejemplo, si tomamos $l_* = \Lambda/pc$ entonces todos los calculos serán similares a los presentados o realizados anteriormente en la investigación de la alternativa (3). El caso $l_* = 0$ se presenta en la ilustración de la fig. 4.9.

Sólo la alternativa (4) resta por investigar. Excepto para la alternativa (5) pues esta no se cumple en el sentido teórico, los resultados son inaplicables. Considerar, por ejemplo, $l_* > \Lambda/pc$. De acuerdo, a $\mathcal{D}^* = \left\{\pi^{\varphi_1^*}\right\}$, de esta forma obtenemos el punto $\widetilde{\pi} = \pi^{\varphi_1^*}$ para el cual $\mathrm{J}^2(\widetilde{\pi}) = 0$. De donde el cero pertenece a el hiper-octante (en nuestro caso -ambos semi-ejes), así la

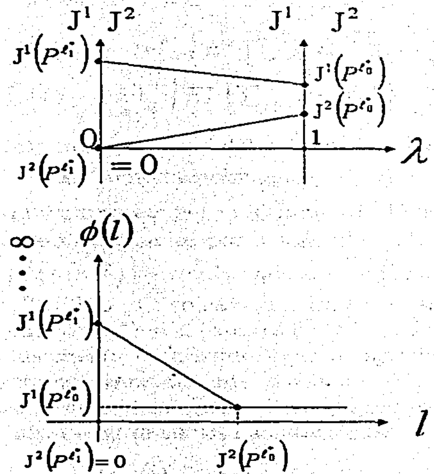

Figura 4.10: Las relaciones entre $\mathbf{J}^2\left(\lambda P^{\varphi_0^*} + (1-\lambda)\, P^{\varphi_1^*}\right)$ y $\mathbf{J}^1\left(\lambda P^{\varphi_0^*} + (1-\lambda)\, P^{\varphi_1^*}\right)$ y la gráfica de la función primal de la alternativa (4).

medida $\pi^{\varphi_1^*}$ es una solución a el problema inicial.

Esta respuesta era de esperarse pues el $\min\limits_{\pi} \mathbf{J}^2(\pi) = 0$ y es alcanzado por todas las estrategias π para el cual $\pi_1(1 \mid h_0) = 1$ si $h_0 = 1$. Esto recuerda elegir una estrategia tal que provea el valor mínimo de la funcional $\mathbf{J}^1(\cdot)$; y que es exactamente φ_1^* .

La función primal $\phi(\cdot)$ puede ser construida como antes. Ilustraciones de la alternativa (4) son presentadas en la figura 4.10. Obviamente, la restricción es esencial en este caso, por defición.

El problema inicial tiene una solución única en cada uno de los casos (1)-(4).

Asumimos que la constante d no esta fija o consideramos el problema de multicriterio

$$\left.\begin{array}{r}\overline{\mathbf{J}^1}(\pi)=\inf\limits_{\pi\in\prod^K}\left\{E^\pi\left[\sum_{t=1}^T e_{a_t}\right]\right\},\\ \overline{\mathbf{J}^2}(\pi)=\inf\limits_{\pi\in\prod^K}\left\{E^\pi\left[\sum_{t=1}^T x_{t-1}pc\right]\right\}.\end{array}\right\}$$

Esta solución se completa con la ayuda del método de restricciones presentado anteriormente para $T=2$. Como mencionamos, $\overline{\mathbf{J}^2}(\pi)\in\left[\overline{\mathbf{J}^2}\left(\pi^{\varphi_1^*}\right),\overline{\mathbf{J}^2}\left(\pi^{\varphi_0^*}\right)\right]$, donde $\overline{\mathbf{J}^2}\left(\pi^{\varphi_{0,1}^*}\right)$ estan expresados por la fórmula (4.11) en $d=0$. Los valores $\overline{\mathbf{J}^1}\left(\pi^{\varphi_0^*}\right)$ y $\overline{\mathbf{J}^1}\left(\pi^{\varphi_1^*}\right)$ estan expresados por la fórmula (4.14). De hecho reescribiendo $d\in\left[\overline{\mathbf{J}^2}\left(\pi^{\varphi_1^*}\right),\overline{\mathbf{J}^2}\left(\pi^{\varphi_0^*}\right)\right]$ y resolviendo el problema restringido obtenemos las alternativas (2), (3) o (4); en cada uno de estos casos existe una única solución π^* para la cual $\overline{\mathbf{J}^2}(\pi^*)=d$ (esto es $\mathbf{J}^2(\pi^*)=\overline{\mathbf{J}^2}(\pi^*)=-d=0$).

Esto será particularmente enfatizado, que la distribución de probabilidad inicial $P_0(\cdot)$ afecta substancialmente la solución del problema. Además, cuando consideramos el caso más significativo (alternativa (3)), la solución π^*corresponde a la estrategia de control aleatorizada:

$$\pi_1^*(0\mid 0)=1;\quad \pi_1^*(0\mid 1)=\lambda^*;\quad \pi_1^*(1\mid 1)=1-\lambda^*;\quad \pi_2(0\mid h_1)=1,$$

y la clase de reglas de decisión deterministas no son suficientes para resolver el problema.

Capítulo 5

Construcción de un modelo de seguros.

5.1 Objetivos y decisiones óptimas en seguros.

Aquí habremos de examinar algunos de los problemas de decisión que ocurren en la dirección de una compañía de seguros. Nos aproximaremos a la materia mediante la construcción de un modelo matemático que describa el ambiente en el cual la compañía opera, y de los objetivos que ésta quiera lograr. La mejor estrategia, o decisión óptima será entonces la que trae a la firma los objetivos más apegados a ella.

Un modelo matemático usualmente será una enorme simplificación de la vida real, pero aun así frecuentemente un modelo simple puede capturar los elementos esenciales de la situación que deseamos estudiar. En tales casos podemos usar el modelo para calcular la decisión óptima y recomendar esto a la dirección. Puede suceder de cualquier modo que la recomendación sea rechazada por algún funcionario, el cual prefiera tomar una decisión que, acorde al modelo, esté lejos de ser la óptima. Podemos pensar que el funcionario crea que el modelo esta equivocado debido a su simplicidad, sin embargo mas bien es que existe en el medio de los seguros todavía mucha reticencia a hacer uso de modelos matemáticos.

En lo siguiente estudiaremos un modelo que a primera vista parece razonable y aceptable. Mostraremos que induce decisiones que están lejos de las observaciones no consideradas como óptimas por los funcionarios en la vida real. Entonces generalizaremos los modelos que inducen más de conformidad con estos que podemos observar en la práctica, notaremos que los modelos pierden algo de su substancia intuitiva en este proceso.

En este capítulo optimizaremos los dividendos y la vida esperada de una compañía de seguros, desafortunadamente no podemos optimizar ambos simultaneamente, pues son inconsistentes. así que optimizaremos uno de ellos con restriccipon sobre el otro, y viceversa. Para esto usamos la teoría de la programación convexa desarrollada en el capítulo anterior para resolver estos dos problemas restringidos. También usaremos la caminata aleatoria. En lo que queda del capítulo definimos el modelos de seguros, planteamos los problemas de control de políticas óptimas en seguros y con restricciones, y finalmente los resolveremos. Algunas referencias para este Capítulo son: [5] y [12].

5.2 Algunas formulaciones simples de los objetivos.

Nos concentraremos en los modelos que hacen una asignación para el pago de dividendos a los socios o a los clientes de una compañía de seguros. El valor total de los dividendos a saldar y el tiempo esperado de vida de la compañía de seguros juegan el papel de criterios para la ejecución de políticas.

Así inciaremos, considerando un modelo simple que presentó Bruno de Finetti en el Congreso Internacional de Actuarios en 1957. El modelo representa una generalización substancial de los modelos usados en las viejas Teorías del Riesgo, que pueden ser descritas por los elementos siguientes:

(i) ζ_t son los reclamos de la cartera asegurada por la compañía de seguros en el período de operación t. Sea $\{\zeta_t\}_{t=1}^{\infty}$ la sucesión de reclamos, consideramos que son variables aleatorias independientes idénticamente distribuidas con distribución $\mathbf{F}(\cdot)$

(ii) $h_t \geq 0$ es el monto total de las primas captadas por la compañía en el período t.

(iii) $\xi_{t-1} \geq 0$ representa la reserva de la compañía al inicio del período $t-1$, $\xi_{t-1} \in (t-1, t]$, antes de pagar dividendos. El valor inicial de la reserva $\xi_0 = x_0 \geq 0$ esta dado.

(iv) η_t representa los dividendos pagados por la compañía al inicio del período t. Cuando los resultados del período $t-1$ se conocen, $\eta_t \in [0, \xi_{t-1}]$; es decir, η_t está en función de las reservas que la compañía genera en el período $(t-1, t]$ por lo que se puede disponer de cualquier cantidad entre 0, que no se hayan generado reservas, y ξ_{t-1} para el pago de dividendos.

Esto nos da lo que un físico podría llamar la "*Ley de Movimiento*":

$$\xi_t = \xi_{t-1} - \eta_t + \mathbf{h}_t - \zeta_t. \tag{5.1}$$

donde $\zeta_t \geq 0$ son las indemnizaciones por los siniestros ocurridos en el período t.

Con esta formulación, el problema de decisión de la compañía consiste sólo en determinar η_t, el monto con el cual se debería iniciar como dividendo antes de iniciar el nuevo período de operación.

Una "regla del juego" razonable es:

Si para alguna t, $\xi_t < 0$ la compañía esta arruinada, y esto no permite la operación en los siguientes períodos.

Por lo que De Finetti sugiere encontrar la *política óptima* del pago de dividendos la cual máximiza la suma descontada esperada de los pagos de dividendos los cuales la compañía efectuara antes de ocurrir la ruina.

De lo cual, uno de los problemas que se podrían abordar es entonces

$$\max \left\{ \sum_{t=0}^{\infty} \beta^t E\left\{\eta_t\right\} \right\} \tag{5.2}$$

- Para resolver el problema que hemos trazado debemos considerar la información sobre el futuro asegurador de la compañía, es decir conocer las sucesiones $\mathbf{F}_0(x), \mathbf{F}_1(x), \ldots$ (distribución de los reclamos en el período t) y $\mathbf{h}_0, \mathbf{h}_1, \ldots, \mathbf{h}_t$ (monto total de primas). No puede ser díficil hacer pronósticos razonables en el corto plazo para estas 2 sucesiones. Podemos por ejemplo asumir que el volumen de los negocios aumentará, parcialmente porque el número de contratos aumentará, y parcialmente porque las sumas aseguradas bajo cada contrato se incrementa con la inflación. Sin embargo por simplicidad no se tomará encuenta esta última característica.

Nuestra propuesta es presentar características esenciales del modelo. Asumiremos que $\mathbf{F}_t(x) = \mathbf{F}(x)$ y $\mathbf{h}_t = \mathbf{h}$ para toda t, es decir consideramos que la compañía asegurará carteras idénticas en todos los periodos futuros de operación.

Las variables aleatorias $\{\zeta_t\}_{t=1}^{\infty}$ son consideradas mutuamente independientes con la distribución de probabilidad $F(d\zeta)$.

El valor inicial de la reserva $\boldsymbol{\xi}_0 \geq 0$ se supone está dado.

Sea $\tau = \inf\{t : \boldsymbol{\xi}_t < 0\}$. Cuando se pagan los dividendos, la compañía puede dedicarse a difentes objetivos. Por ejemplo, uno puede maximizar su valor total

$$\max\left\{E^{\pi}\left\{\sum_{t=1}^{\tau}\boldsymbol{\eta}_t\right\}\right\}, \tag{5.3}$$

o maximizar el tiempo esperado de vida hasta el colapso

$$\max\{E^{\pi}\{\tau\}\}. \tag{5.4}$$

Donde π es una estrategia de control (política), $\boldsymbol{\xi}_t$ es el proceso controlado definido por la ecuación dinámica (5.1). Lo descrito anteriormente es un modelo de proceso de decisión de Markov con la variable aleatoria de tiempo de vida $[0,\tau]$ donde τ es el momento del primer evento desfavorable para la compañía, esto para el conjunto $\hat{\mathrm{X}}=\mathrm{R}_+$.

Una solución al problema (5.4) es $\boldsymbol{\eta}_t \equiv 0$ (no se pagan dividendos). Pero el problema de multi-criterio (5.3), (5.4) es con frecuencia inconsistente pues estariamos planteando que la compañía aseguradora no ha repartido utilidades durante el tiempo de vida que esta lleve, lo cual no tiene razón de ser. Nos apoyaremos en los métodos desarrollados en el capítulo anterior para problemas con restricciones. Para ello introduciremos la distribución de probabilidad más simple $F(\cdot)$.

Consideremos:

- El monto de las primas en el tiempo t es $h = 1$,
- Las indemnizaciones ζ_t, se definen por:

$$\zeta_t = \begin{cases} p, & \text{si } \zeta_t = 0 \\ q, & \text{si } \zeta_t = 2 \end{cases} \quad \text{con } p + q = 1.$$

- El valor de los dividendos se considera que es entero, es decir, $\eta_t \in A = \{0, 1, 2, \ldots\}$.
- A medida que avanza el sistema el capital o reserva toma el espacio de estados del proceso de control ξ_t, que es el conjunto $X = \{-1, 0, 1, 2, \ldots\}$.

En lo siguiente nos avocaremos al problema de maximizar el valor de la compañía, es decir nos enfocaremos al problema (5.3). Como los dividendos no pueden exceder el capital disponible tenemos:

$$j^1(x,a) = \begin{cases} -a, & \text{si } x \geq 0;\ a \in [0,x]; \\ +\infty, & \text{si } x \geq 0;\ a > x; \\ 0, & \text{si } x = -1. \end{cases} \tag{5.5}$$

Donde el valor $+\infty$ es una penalización para excluir cuando no se cumple la restricción. Entonces la expresión (5.3) toma la forma estandar:

$$\mathbf{J}^1(P^\pi) = \min_\pi \left\{ E^\pi \left[\sum_{t=1}^{\infty} j^1(\xi_{t-1}, \eta_t) \right] \right\}. \tag{5.6}$$

Lema 5.1 *(a) Si $p \leq q$ entonces $\inf_\pi \mathbf{J}^1(P^\pi) = -x_0 - p/q$ y el selector estacionario $\varphi^1(x) = x$ es una solución al problema (5.6).*

(b) Si $p > q$ entonces $\inf_\pi \mathbf{J}^1(P^\pi) = -\infty$. Sea φ^T la sucesión de selectores de la forma $\varphi^T(t,x) = I\{t = T\}\,x$; entonces $\lim_{T\to\infty} \mathbf{J}^1\left(P^{\varphi^T}\right) = -\infty$.

Prueba:
(a) Obviamente,

$$\mathbf{J}^1(P^\pi) = \lim_{T\to\infty} E^\pi \left[\sum_{t=1}^{T} j^1(\xi_{t-1}, \eta_t) \right] \text{ con } T < \infty.$$

Resolviendo así el problema siguiente

$$\min_\pi \left\{ E^\pi \left[\sum_{t=1}^{\infty} j^1(\xi_{t-1}, \eta_t) \right] \right\}$$

con $T < \infty$.

Consideramos:

$$X = \{-1, 0, 1, 2, \ldots, x_0, x_0 + 1, \ldots, x_0 + T\} ;$$

$$A = \{0, 1, 2, \ldots, x_0 + T\} .$$

De la ecuación dinámica (5.1) , tenemos:

$$y = x - a + 1 - \begin{cases} 0 \\ 2 \end{cases} ,$$

y las probabilidades de transición

$$p\left(y/x, a\right) = \begin{cases} p & \text{sí } y = x - a + 1, \quad x = 0, 1, 2, \ldots ; \\ q & \text{sí } y = x - a - 1, \quad x = 0, 1, 2, \ldots ; \\ 1 & \text{sí } y = -1 \qquad\qquad\quad x = -1. \end{cases}$$

De lo cual la función de óptimalidad es:

$v_T\left(x\right) \equiv 0,$

$$\begin{aligned} v_{T-1}\left(x\right) &= \min_A \left\{ j^1\left(x, a\right) + \sum_y v_T\left(y\right) p\left(y/x, a\right) \right\} \\ &= \min \left\{ \begin{array}{c} j^1\left(x, 0\right) + v_T\left(x+1\right) p\left(x + 1/x, 0\right) + v_T\left(x-1\right) p\left(x - 1/x, 0\right), \\ j^1\left(x, 1\right) + v_T\left(x\right) p\left(x/x, 1\right) + v_T\left(x-2\right) p\left(x - 2/x, 1\right), \\ \cdots \\ j^1\left(x, x-1\right) + v_T\left(2\right) p\left(2/x, x-1\right) + v_T\left(0\right) p\left(0/x, x-1\right), \\ j^1\left(x, x\right) + v_T\left(1\right) p\left(1/x, x\right) + v_T\left(-1\right) p\left(-1/x, x\right), \\ j^1\left(x, x+1\right) + v_T\left(0\right) p\left(0/x, x+1\right) + v_T\left(-2\right) p\left(x/x, x+1\right), \\ j^1\left(x, x+2\right) + v_T\left(-1\right) p\left(-1/x, x+2\right) + v_T\left(-3\right) p\left(-3/x, x+2\right), \\ \cdots \end{array} \right\} \\ &= \min \left\{0, -1, \ldots - x + 1, -x, \infty, \infty, \ldots\right\} . \end{aligned}$$

Comparando

$$\begin{aligned} -x + 1 &< -x \\ 1 &< 0\,! \end{aligned}$$

por lo tanto:

$$v_{T-1}\left(x\right) = -x \quad \text{con} \quad \varphi^{T-1}\left(x\right) = x.$$

Para

$$\begin{aligned}
v_{t-2}(x) &= \min_{A}\left\{ j^{1}(x,a) + \sum_{y} v_{T-1}(y)\, p(y/x,a) \right\} \\
&= \min \left\{ \begin{array}{c}
j^{1}(x,0) + v_{T-1}(x+1)\, p(x+1/x,0) + v_{T-1}(x-1)\, p(x-1/x,0), \\
j^{1}(x,1) + v_{T-1}(x)\, p(x/x,1) + v_{T-1}(x-2)\, p(x-2/x,1), \\
j^{1}(x,2) + v_{T-1}(x-1)\, p(x-1/x,2) + v_{T-1}(x-3)\, p(x-3/x,2), \\
\cdots \\
j^{1}(x,x-1) + v_{T-1}(2)\, p(2/x,x-1) + v_{T-1}(0)\, p(0/x,x-1), \\
j^{1}(x,x) + v_{T-1}(1)\, p(1/x,x) + v_{T-1}(-1)\, p(-1/x,x), \\
\infty \\
\cdots
\end{array} \right. \\
&= \min \left\{ \begin{array}{c}
0-(x+1)\,p-(x-1)\,q, -1-xp-(x-2)\,q, \\
-2-(x-1)\,p-(x-3)\,q, \ldots, -(x-1)-2p, \\
-x-p, \infty, \infty, \infty, \ldots
\end{array} \right\}.
\end{aligned}$$

Como

$$v_{T-1}(-1) \equiv 0$$

y para conocer el mínimo, tomemos un término general:

$$\begin{aligned}
-b-(x-b+1)\,p-(x-b-1)\,q &= -b-(xp-bp+p)-(xq-bq-q) \\
&= -b-xp+bp-p-xq+bq+q \\
&= -b-x+b-p+q \\
&= -x-p+q
\end{aligned}$$

comparando con el último término:

$$\begin{aligned}
-x-p+q &< -x-p \\
q &< 0!.
\end{aligned}$$

Por lo tanto:

$$\begin{aligned}
v_{t-2}(x) &= -x-p \quad \text{con} \quad \varphi^{t-2}(x) = x. \\
&= -(x+p)
\end{aligned}$$

Para

$$
\begin{aligned}
v_{t-3}(x) &= \min_A \left\{ j^1(x,a) + \sum_y v_{t-2}(y)\, p(y/x,a) \right\} \\
&= \min \left\{ \begin{array}{c}
j^1(x,0) + v_{t-2}(x+1)\, p(x+1/x,0) + v_{t-2}(x-1)\, p(x-1/x,0), \\
j^1(x,1) + v_{t-2}(x)\, p(x/x,1) + v_{t-2}(x-2)\, p(x-2/x,1), \\
j^1(x,2) + v_{t-2}(x-1)\, p(x-1/x,2) + v_{t-2}(x-3)\, p(x-3/x,2), \\
\cdots \\
j^1(x,x-1) + v_{t-2}(2)\, p(2/x,x-1) + v_{t-2}(0)\, p(0/x,x-1), \\
j^1(x,x) + v_{t-2}(1)\, p(1/x,x) + v_{t-2}(-1)\, p(-1/x,x), \\
\infty \\
\cdots
\end{array} \right\} \\
&= \min \left\{ \begin{array}{c}
0-(x+1+p)\,p-(x-1+p)\,q, -1-(x+p)\,p-(x-2+p)\,q, \\
-2-(x-1+p)\,p-(x-3+p)\,q, \ldots, -(x-1)-(2+p)\,p, \\
-x-(1+p)\,p, \infty, \infty, \infty, \ldots
\end{array} \right\}.
\end{aligned}
$$

Consideremos un término general, excluyendo $-x-(1+p)\,p$, de lo cual tenemos:

$$
\begin{aligned}
-k-(x-k+1+p)\,p-(x-k-1+p)\,q &= -k-xp+kp-p-p^2-xq+kq+q-pq \\
&= -k+k-x-p-p^2+q-pq \\
&= -x-p-p^2+q-pq \\
&= -x-p-p^2+q-p(1-p) \\
&= -x-p-p^2+q-p+p^2 \\
&= -x-2p+q
\end{aligned}
$$

Suponiendo que este término es el menor con respecto a $-x-(1+p)\,p$, tenemos:

$$
\begin{aligned}
-x-2p+q &< -x-(1+p)\,p \\
-2p+q &< -(1+p)\,p \\
-2p+1-p &< -p-p^2 \\
1-2p+p^2 &< 0 \\
(1-p)^2 &< 0 \\
q^2 &< 0\,!.
\end{aligned}
$$

Por lo tanto el mínimo es el último término; es decir:

$$v_{t-3}(x) \begin{array}{l} = -x-p-p^2 \\ = -(x+p+p^2) \end{array} \quad \text{con} \quad \varphi^{t-3}(x) = x.$$

Ahora:

$$\begin{aligned} v_{t-4}(x) &= \min_{A}\left\{ j^1(x,a) + \sum_{y} v_{t-3}(y)\, p(y/x,a) \right\} \\ &= \min \left\{ \begin{array}{c} j^1(x,0) + v_{t-3}(x+1)\, p(x+1/x,0) + v_{t-3}(x-1)\, p(x-1/x,0), \\ j^1(x,1) + v_{t-3}(x)\, p(x/x,1) + v_{t-3}(x-2)\, p(x-2/x,1), \\ j^1(x,2) + v_{t-3}(x-1)\, p(x-1/x,2) + v_{t-3}(x-3)\, p(x-3/x,2), \\ \cdots \\ j^1(x,x-1) + v_{t-3}(2)\, p(2/x,x-1) + v_{t-3}(0)\, p(0/x,x-1), \\ j^1(x,x) + v_{t-3}(1)\, p(1/x,x) + v_{t-3}(-1)\, p(-1/x,x), \\ \infty \\ \cdots \end{array} \right. \\ &= \min \left\{ \begin{array}{c} 0-(x+1+p+p^2)\,p-(x-1+p+p^2)\,q, \\ -1-(x+p+p^2)\,p-(x-2+p+p^2)\,q, \\ -2-(x-1+p+p^2)\,p-(x-3+p+p^2)\,q, \\ \ldots, \\ -(x-1)-(2+p+p^2)\,p, \\ -x-(1+p+p^2)\,p, \infty, \infty, \infty, \ldots \end{array} \right\}. \end{aligned}$$

Consideremos un término general, excluyendo $-x-(1+p+p^2)\,p$, de lo cual tenemos:

$$\begin{aligned} & -k-(x-k+1+p+p^2)\,p-(x-k-1+p+p^2)\,q \\ = & -k-xp+kp-p-p^2-p^3-xq+kq+q-pq-p^2q \\ = & -k-x+k-p-p^2-p^3+q-pq-p^2q \\ = & -x-p-p^2+q-p. \end{aligned}$$

Comparando, suponiendo que este término es el mayor con respecto a $-x-(1+p+p^2)\,p$, tenemos:

$$\begin{aligned} -x-p-p^2+q-p. &> -x-p-p^2-p^3 \\ q-p &> -p^3 \end{aligned}$$

como $p \leq q \Longrightarrow 0 \leq q-p$, por lo tanto:

$$0 > -p^3.$$

De lo cual el mínimo es el último término; es decir:

$$v_{t-4}(x) \quad \begin{array}{l} = -x - p - p^2 - p^3 \\ = -(x + p + p^2 - p^3) \end{array} \quad \text{con} \quad \varphi^{t-4}(x) = x.$$

A partir de lo anterior mostraremos el término general.

$$\begin{aligned} v_{t-T}(x) &= -\left(x + p + p^2 + p^3 + \cdots + p^T\right) \\ &= -x - \left(\frac{p - p^{T+1}}{1-p}\right). \end{aligned}$$

La ecuación de Bellman tiene la forma

$$\left.\begin{array}{ll} v_T(x) & = 0; \\ v_{t-1}(x) & = \min\limits_{A} \{-a + pv_t(x - a + 1) + qv_t(x - a - 1)\} \end{array}\right\}$$

con $v_t(-1) \equiv 0$. Y como hemos visto anteriormente la solución a la ecuación de Bellman tiene la forma $v_t(x) = -x - \left(p - p^{T-t}\right) / (1-p)$ y el mínimo es alcanzado por la acción $a_t = \varphi^1(x_{t-1}) = x_{t-1}$.

La construcción del selector φ^1 es la estrategia óptima para el problema con horizonte finito, es decir

$$\min_{\pi} E^{\pi}\left[\sum_{t=1}^{T} j^1(\xi_{t-1}, \eta_t)\right] = -x_0 - \frac{p - p^T}{1-p}.$$

Obviamente, $\forall \pi$

$$\mathbf{J}^1(P^{\pi}) \geq -x_0 - \frac{p}{1-p};$$

de otra manera

$$\mathbf{J}^1\left(P^{\varphi^1}\right) = \lim_{T\to\infty} E^{\varphi^1}\left[\sum_{t=1}^{T} j^1(\xi_{t-1}, \eta_t)\right] = -x_0 - \frac{p}{1-p}.$$

Por lo tanto, el selector φ^1 es la estrategia óptima para el problema (5.6) y $\min\limits_{\pi} \mathbf{J}^1(P^{\pi}) = -x_0 - p/(1-p)$.

Con lo anterior hemos probado la parte (a) del Lema, ahora abordaremos el inciso (b).♣

(b) Escogemos una constante arbitraria C, la cual satisface la desigualdad

$$-\frac{1}{q} < C < -\frac{1}{p}.$$

(en el caso degenerado $q = 0$ uno puede asignar $-\frac{1}{q} = -\infty$.)

Introducimos el costo terminal

$$j_T^1(x) = \begin{cases} C, & \text{sí } x \neq -1; \\ 0, & \text{sí } x = -1 \end{cases}$$

y resolvemos el problema

$$\min_{\pi} \left\{ E^{\pi} \left[\sum_{t=1}^{T} j^1(\xi_{t-1}, \eta_t) + j_T^1(\xi_T) \right] \right\}$$

con $T < \infty$. Deduciremos las ecuaciones de Bellman.

Para el valor de salida:

$v_T(x) = j_T^1(x) = C$, sí $x \neq -1$. $v_t(-1) \equiv 0$;

$$\begin{aligned}
v_{t-1} &= \min_A \left\{ j^1(x,a) + \sum_y v_T(y)\, p(y/x,a) \right\} \\
&= \min \left\{ \begin{array}{c} j^1(x,0) + v_T(x+1)\, p(x+1/x,0) + v_T(x-1)\, p(x-1/x,0), \\ j^1(x,1) + v_T(x)\, p(x/x,1) + v_T(x-2)\, p(x-2/x,1), \\ j^1(x,2) + v_T(x-1)\, p(x-1/x,2) + v_T(x-3)\, p(x-3/x,2), \\ \cdots \\ j^1(x,x-1) + v_T(2)\, p(2/x,x-1) + v_T(0)\, p(0/x,x-1), \\ j^1(x,x) + v_T(1)\, p(1/x,x) + v_T(-1)\, p(-1/x,x), \\ \cdots \end{array} \right\} \\
&= \min \left\{ \begin{array}{c} 0 + Cp + Cq, -1 + Cp + Cq, -2 + Cp + Cq, \\ \cdots \\ -(x-1) + Cp + Cq, -x + Cp + 0, \infty, \cdots \end{array} \right\} \\
&= \min \{0 + C, -1 + C, -2 + C, \ldots, -(x-1) + C, -x + Cp\}
\end{aligned}$$

Veamos cual es el elemento que proporciona el mínimo con respecto a los términos de esta última igualdad:

$$\begin{aligned}
-(x-1) + C &> -x + Cp \\
1 + C &> Cp \\
1 + C - Cp &> 0
\end{aligned}$$

$$
\begin{aligned}
1+C(1-p) &> 0\\
1-Cq &> 0\\
C &> -\frac{1}{q}.
\end{aligned}
$$

Por lo que el mínimo es:

$$v_{t-1}(x) = -x+Cp \quad \text{con} \quad \varphi^{t-1}(x)=x.$$

Para $t-2$

$$
\begin{aligned}
v_{t-2}(x) &= \min_A \left\{ j^1(x,a) + \sum_y v_{t-1}(y)\, p(y/x,a) \right\}\\
&= \min \left\{ \begin{array}{c}
j^1(x,0) + v_{t-1}(x+1)\, p(x+1/x,0) + v_{t-1}(x-1)\, p(x-1/x,0)\\
j^1(x,1) + v_{t-1}(x)\, p(x/x,1) + v_{t-1}(x-2)\, p(x-2/x,1),\\
j^1(x,2) + v_{t-1}(x-1)\, p(x-1/x,2) + v_{t-1}(x-3)\, p(x-3/x,2)\\
\dots\\
j^1(x,x-1) + v_{t-1}(2)\, p(2/x,x-1) + v_{t-1}(0)\, p(0/x,x-1),\\
j^1(x,x) + v_{t-1}(1)\, p(1/x,x) + v_{t-1}(-1)\, p(-1/x,x),\\
\infty\\
\dots
\end{array} \right.\\
&= \min \left\{ \begin{array}{c}
0 + (-(x+1)+Cp)\,p + (-(x-1)+Cp)\,q,\\
-1 + (-x+Cp)\,p + (-(x-2)+Cp)\,q,\\
-2 + (-(x-1)+Cp)\,p + (-(x-3)+Cp)\,q,\\
\dots,\\
-(x-1) + (-2+Cp)\,p + (Cp)\,q,\\
-x + (-1+Cp)\,p, \infty, \infty, \infty, \dots
\end{array} \right\}.
\end{aligned}
$$

Excluyendo el término $-x+(-1+Cp)\,p$ para compararlo con los demás, obtengamos un término general de la siguiente forma:

$$
\begin{aligned}
&-k + (-(x-k+1)+Cp)\,p + (-(x-k-1)+Cp)\,q\\
&-k - xp + kp - p + Cp^2 - xq + kq + q + Cpq\\
&-k - x + k - p + Cp^2 + q + Cpq\\
&\therefore \; -x-p+q+Cp \text{ es el término general.}
\end{aligned}
$$

Ahora supongamos la siguiente desigualdad válida:

$$
\begin{aligned}
-x-p+q+Cp &< -x-p+Cp^2\\
q+Cp &< Cp^2
\end{aligned}
$$

$$
\begin{aligned}
q + Cp - Cp^2 &< 0\\
q + Cp(1-p) &< 0\\
q + Cpq &< 0\\
q(1+Cp) &< 0\\
C &< -\frac{1}{p}.
\end{aligned}
$$

Por lo tanto la estrategia puede ser cualquiera de las anteriores menos la última.

De ahí

$$
\begin{aligned}
v_{t-2} &= -k + (-(x-k+1) + Cp)\,p + (-(x-k-1) + Cp)\,q\\
&= -x + Cp - (p-q).
\end{aligned}
$$

Veamos el caso $t-3$

$$
\begin{aligned}
v_{t-3}(x) &= \min_A \left\{ j^1(x,a) + \sum_y v_{t-2}(y)\,p(y/x,a) \right\}\\
&= \min \left\{ \begin{array}{c}
j^1(x,0) + v_{t-2}(x+1)\,p(x+1/x,0) + v_{t-2}(x-1)\,p(x-1/x,0),\\
j^1(x,1) + v_{t-2}(x)\,p(x/x,1) + v_{t-2}(x-2)\,p(x-2/x,1),\\
j^1(x,2) + v_{t-2}(x-1)\,p(x-1/x,2) + v_{t-2}(x-3)\,p(x-3/x,2),\\
\cdots\\
j^1(x,x-1) + v_{t-2}(2)\,p(2/x,x-1) + v_{t-2}(0)\,p(0/x,x-1),\\
j^1(x,x) + v_{t-2}(1)\,p(1/x,x) + v_{t-2}(-1)\,p(-1/x,x),\\
\infty\\
\cdots
\end{array} \right\}\\
&= \min \left\{ \begin{array}{c}
0 + (-(x+1) + Cp - (p-q))\,p + (-(x-1) + Cp - (p-q))\,q,\\
-1 + (-x + Cp - (p-q))\,p + (-(x-2) + Cp - (p-q))\,q,\\
-2 + (-(x-1) + Cp - (p-q))\,p + (-(x-3) + Cp - (p-q))\,q,\\
\ldots,\\
-(x-1) + (-2 + Cp - (p-q))\,p + (Cp - (p-q))\,q,\\
-x + (-1 + Cp - (p-q))\,p, \infty, \ldots
\end{array} \right\}.
\end{aligned}
$$

Procedamos de la misma manera que arriba, sacando un término general, menos el siguiente $-x + (-1 + Cp - (p-q))\,p$;

$$
\begin{aligned}
&-k + (-(x-k+1) + Cp - (p-q))\,p + (-(x-k-1) + Cp - (p-q))\,q\\
&-k - xp + kp - p + Cp^2 - p^2 + qp - xq + kq + q + Cpq - pq + q^2\\
&-x - p(1-q) + Cp^2 - p^2 + q + Cpq - pq + q^2\\
&-x - p^2 + Cp(p+q) - p^2 + q - q(p-q)
\end{aligned}
$$

$$-x-p^2+Cp-p^2+q(1-p)+q^2$$
$$-x-p^2+Cp-p^2+q^2+q^2$$
$$-x-2p^2+2q^2+Cp$$
$\therefore \quad -x-2\left(p^2-q^2\right)+Cp$ es el término general.

Ahora procedamos a comparar, suponiendo verdadera la siguiente desigualdad:

$$\begin{aligned}
-x-2\left(p^2-q^2\right)+Cp &< -x+(-1+Cp-(p-q))\,p \\
-2p^2+2q^2+Cp &< -p+Cp^2-p^2+qp \\
Cp-Cp^2 &< -p-p^2+qp+2p^2-2q^2 \\
C\left(p-p^2\right) &< -p+p^2+qp-2q^2 \\
Cp\,(1-p) &< -2q^2 \\
Cp &< -2q \\
C &< -2\frac{q}{p}.
\end{aligned}$$

Como

$$-\frac{1}{q}<C<-\frac{1}{p}$$

supongamos

$$\begin{aligned}
-\frac{1}{p} &< -2\frac{q}{p} \\
1 &> 2q \\
\frac{1}{2} &> q
\end{aligned}$$

Por lo que:

$$v_{t-3}(x) = -x-2(p-q)+Cp.$$

Con $\varphi^{t-3}(x)$ cualquier estrategia menos la última.

Para $t-4$, tenemos:

$$\begin{aligned}
v_{t-4}(x) &= \min_{A}\left\{ j^{1}(x,a)+\sum_{y} v_{t-3}(y)\,p(y/x,a)\right\} \\
&= \min\left\{\begin{array}{c}
j^{1}(x,0)+v_{t-3}(x+1)\,p(x+1/x,0)+v_{t-3}(x-1)\,p(x-1/x,0),\\
j^{1}(x,1)+v_{t-3}(x)\,p(x/x,1)+v_{t-3}(x-2)\,p(x-2/x,1),\\
j^{1}(x,2)+v_{t-3}(x-1)\,p(x-1/x,2)+v_{t-3}(x-3)\,p(x-3/x,2),\\
\cdots\\
j^{1}(x,x-1)+v_{t-3}(2)\,p(2/x,x-1)+v_{t-3}(0)\,p(0/x,x-1),\\
j^{1}(x,x)+v_{t-3}(1)\,p(1/x,x),\\
\infty\\
\cdots
\end{array}\right\} \\
&= \min\left\{\begin{array}{c}
(-(x+1)-2(p-q)+Cp)\,p+(-(x-1)-2(p-q)+Cp)\,q,\\
-1+(-x-2(p-q)+Cp)\,p+(-(x-2)-2(p-q)+Cp)\,q,\\
-2+(-(x-1)-2(p-q)+Cp)\,p+(-(x-3)-2(p-q)+Cp)\,q,\\
\ldots,\\
-(x-1)+(-2-2(p-q)+Cp)\,p+(-2(p-q)+Cp)\,q,\\
-x+(-1-2(p-q)+Cp)\,p,\infty,\ldots
\end{array}\right.
\end{aligned}$$

Para conocer el mínimo, consideremos el término general

$$\begin{aligned}
& -k+(-(x-k+1)+Cp-2(p-q))\,p+(-(x-k-1)-2(p-q)+Cp)\,q \\
= & -k-xp+kp-p-Cpp-2(p-q)\,p-xq+kq+q-2(p-q)\,q+Cpq \\
= & -x-p+q-2(p-q)+Cp \\
= & -x-3(p-q)+Cp.
\end{aligned}$$

Es decir el término general es $-x-3(p-q)+Cp$, ahora procederemos a comparar este término con $-x+(-1-2(p-q)+Cp)\,p$; supongamos válida la siguiente desigualdad:

$$\begin{aligned}
-x-3(p-q)+Cp &< -x+(-1-2(p-q)+Cp)\,p \\
-3(p-q)+Cp &< -p-2(p-q)\,p+Cp^{2} \\
-3(p-q)+2(p-q)\,p+Cp &< -p+Cp^{2} \\
(-3+2p)(p-q)+Cp-Cpp &< -p \\
(-3+2p)(p-q)+Cp(1-p) &< -p \\
(-3+2p)(p-q)+Cpq &< -p \\
Cpq &< -p-(-3+2p)(p-q) \\
C &< \frac{-p+(3-2p)(p-q)}{pq}
\end{aligned}$$

como

$$C < -\frac{1}{p}$$

supondremos

$$\begin{aligned}
-\frac{1}{p} &< \frac{-p + (3-2p)(p-q)}{pq} \\
-1 &< \frac{-p + (3-2p)(p-q)}{q} \\
-q &< -p + (3-2p)(p-q) \\
p - q &< (3-2p)(p-q)
\end{aligned}$$

como $p > q$ entonces $p - q > 0$

$$\begin{aligned}
1 &< 3 - 2p \\
1 - 3 &< -2p \\
-2 &< -2p \\
1 &> p
\end{aligned}$$

de esto último concluimos que el mínimo es alcanzado cuando, se escoge cualquier estrategia φ^{t-4} menos la última. Por lo tanto

$$v_{t-4}(x) = -x - 3(p-q) + Cp.$$

Resumiendo tenemos, la ecuación de Bellman

$$\left.\begin{array}{l}
v_T(x) = j_T^1(x); \qquad v_t(-1) \equiv 0; \\
v_{t-1}(x) = \min\limits_{A} \{-a + pv_t(x-a+1) + qv_t(x-a-1)\}
\end{array}\right\}$$

la cual como hemos visto su solución en general es

$$v_t(x) = -x + pC - (T-t-1)(p-q) \qquad (t < T,\ x \geq 0)$$

y el mínimo es alcanzado por la acción

$$a_t = \varphi^T(t, x_{t-1}) = I\{t = T\} \cdot x.$$

Ahora es claro que

$$\begin{aligned}
E^{\varphi^T}\left[\sum_{t=1}^{\infty} j^1(\xi_{t-1}, \eta_t)\right] &\leq E^{\varphi^T}\left[\sum_{t=1}^{T} j^1(\xi_{t-1}, \eta_t) + j_T^1(\xi_T)\right] + |C| \\
&= -x_0 + pC - (T-t-1)(p-q) + |C|.
\end{aligned}$$

que es lo que se queria demostrar.♠

Como es evidente de lo anterior, el valor total esperado de los dividendos (5.3) se puede hacer tan grande como se quiera en el caso $p > q$ retrazando el tiempo T de sus pagos. Por lo tanto, ambas funciones (5.3) y (5.4) se incrementan simultaneamente con un aumento de T cuando las estrategias φ^T del Lema 5.1 son implementadas. De cualquier modo, notamos que el selector $\varphi^0(t,x) \equiv 0$, de igual manera para φ^∞, se da una solución para el problema (5.4) pero el valor total de los dividendos $E^{\varphi^0}\left[\sum_{t=1}^{T}\eta_t\right] = 0$ esta lejos del máximo. Como una regla el factor de descuento $\beta \in (0,1)$ es introducido en el caso $p > q$ y el problema

$$\max_\pi \left\{ E^\pi \left[\beta^{t-1}\eta_t\right] \right\}$$

es estudiado en lugar de (5.3). Sólo mencionaremos que su solución es la estrategia $\varphi(x) = I\left\{x > \prod\right\}(x - \prod)$.

En lo que sigue el caso $p < q$ es estudiado. Aqui el valor total de los dividendos (5.3) es máximo si un desembolso es efectuado al capital inicial x_0 en el primer momento. No acumular capital debería hacerse en el futuro: $\varphi^1(x) = x$. Esta respuesta es intuitivamente clara porque los negocios asegurados no generan algún beneficio en el caso $p < q$. De esta manera es improductivo acumular capital: es más probable pagar siniestros. De otra forma, la estrategía $\varphi^1(x)$ es lo peor con respecto al tiempo esperado de vida hasta el colapso (5.4). Así el problema de multicriterio (5.3), (5.4) es significativo.

Como hemos visto anteriormente, tenemos $X = \{-1,0,1,2,\ldots\}$; $A = \{0,1,2,\ldots\}$;

$$p_t(y/x,a) = \begin{cases} p, & \text{sí } y = x - a + 1; \quad x \geq 0; \\ q, & \text{sí } y = x - a - 1; \quad x \geq 0; \\ 1, & \text{sí } y = x = -1; \\ 0, & \text{en otro caso.} \end{cases}$$

Sea

$$j^2(x,a) = \begin{cases} -1, & \text{sí } x \geq 0; \quad a \in [0,x]; \\ +\infty & \text{sí } x \geq 0; \quad a > x; \\ 0, & \text{sí } x = -1 \end{cases} \tag{5.7}$$

y escribimos la función (5.4) en la forma estándar

$$\mathbf{J}^2\left(P^\pi\right)=\min_\pi\left\{E^\pi\left[\sum_{t=1}^{\infty}j^2\left(\xi_{t-1},\eta_t\right)\right]\right\}. \tag{5.8}$$

Como es usual el método de restricciones para el problema (5.6), (5.8) se puede realizar de dos formas.

I. Escoger una costante d y encontrar una estrategia que provea el mínimo a la expresión (5.8) bajo la restricción

$$\mathbf{J}^1\left(P^\pi\right)\leq d. \tag{5.9}$$

II. Escoger una constante d y encontrar una estrategia que provea el mínimo a la expresión (5.6) bajo la restricción

$$\mathbf{J}^2\left(P^\pi\right)\leq d. \tag{5.10}$$

5.3 Resultados generales.

Obviamente, la función (5.8) es siempre finita: su valor mínimo es alcanzado por $\varphi^0(x)\equiv 0$ y coincide con la duración de la caminata aleatoria sobre un enrejado entero en una-dimensión con acumulación en el punto $x=-1$, o con la duración del juego contra un oponente infinitamente rico en el juego clásico de la ruina.

Es decir:

$$\begin{aligned}
d(z,a) &= \frac{z}{q-p}-\frac{a}{q-p}\cdot\frac{1-\left(\frac{q}{p}\right)^z}{1-\left(\frac{q}{p}\right)^a},\\
\lim_{a\to\infty}d(z,a) &= \frac{z}{q-p}-\frac{1-\left(\frac{q}{p}\right)^z}{q-p}\cdot\lim_{a\to\infty}\frac{a}{1-\left(\frac{q}{p}\right)^a},
\end{aligned}$$

aplicando L'Hospital para

$$\lim_{a\to\infty}\frac{a}{1-\left(\frac{q}{p}\right)^a}=0.$$

Por lo que

$$d(z,a) = \frac{z}{q-p}.$$

En nuestro caso

$$\min_{\pi} \mathbf{J}^2 (P^{\pi}) = \mathbf{J}^2 \left(P^{\varphi^0}\right) = -\frac{x_0 - (-1)}{q-p} = -\frac{x_0+1}{q-p}. \tag{5.11}$$

La importancia de la función (5.6) ha sido establecida en el Lema 5.1, Item (a). Por lo tanto las variables duales $y_1 \geq 0$ y $y_2 \geq 0$ de la función Lagrangiana puede ser insertada bajo el símbolo de la sumatoria sobre t :

$$\begin{aligned} L(P^{\pi}, y_1, y_2) &= y_1 \mathbf{J}^1 (P^{\pi}) + y_2 \mathbf{J}^2 (P^{\pi}) - y_1 y_2 d \\ &= E^{\pi} \left[\sum_{t=1}^{\infty} \left\{ y_1 j^1 (\xi_{t-1}, \eta_t) + y_2 j^2 (\xi_{t-1}, \eta_t) \right\} \right] - y_1 y_2 d. \end{aligned}$$

Resolveremos el problema de optimización incondicional

$$\mathrm{g}(y_1, y_2) = \min_{\pi} L(P^{\pi}, y_1, y_2). \tag{5.12}$$

Los siguientes casos son posibles dependiendo de los valores de los parámetros.

Para abordar esta sección deduciremos cuatro casos para posteriormente elaborar un resumen con los resultados propuestos.

(a) Consideraremos el caso $1 > y(q-p)$ para el problema con horizonte finito $T < \infty$:

$$\min_{\pi} \left\{ E^{\pi} \left[\sum_{t=1}^{T} \left\{ y_1 j^1 (\xi_{t-1}, \eta_t) + y_2 j^2 (\xi_{t-1}, \eta_t) \right\} - y_2 \frac{\xi_T + 1}{q-p} \right] \right\}.$$

Asumiremos que $X = \{-1, 0, 1, 2, \ldots, x_0, x_0+1, \ldots, x_0+T\}$;

$A = \{0, 1, 2, \ldots, x_0, x_0+1, \ldots, x_0+T\}$. Y deduciremos las ecuaciones de Bellman. Para ello analizaremos

$$\mathrm{J}^2\left(P^\pi\right)=\min_\pi\left\{E^\pi\left[\sum_{t=1}^{T} j^2\left(\xi_{t-1},\eta_t\right)\right]\right\}$$

s. a.

$$\mathrm{J}^1\left(P^\pi\right)\leq d.$$

Para la solución abordaremos el caso sin restricciones, por lo que el Lagrangiano toma la forma:

$$\begin{aligned} l\left(y,P^\pi\right) &= s+yr \\ &= -1+y\left(-a\right). \end{aligned}$$

Como ya hemos visto:

$$v_T\left(-1\right)\equiv 0.$$

Valor de salida

$$v_T\left(x\right)=-\frac{x+1}{q-p}.$$

Para el tiempo $T-1$,

$$\begin{aligned} v_{T-1}\left(x\right) &= \min_A\left\{l\left(x,a\right)+\sum_z v_T\left(z\right)p\left(z/x,a\right)\right\} \\ &= \min_A\left\{-1-ay+pv_T\left(x-a+1\right)+qv_T\left(x-a-1\right)\right\} \\ &= \min\left\{\begin{array}{c} -1+p\left(-\frac{x+2}{q-p}\right)+q\left(-\frac{x}{q-p}\right), \\ -1-y+p\left(-\frac{x+1}{q-p}\right)+q\left(-\frac{x-1}{q-p}\right), \\ -1-2y+p\left(-\frac{x}{q-p}\right)+q\left(-\frac{x-2}{q-p}\right), \\ \cdots, \\ -1-\left(x-1\right)y+p\left(-\frac{3}{q-p}\right)+q\left(-\frac{1}{q-p}\right), \\ -1-xy+p\left(-\frac{2}{q-p}\right). \end{array}\right\}. \end{aligned}$$

Ahora para conocer el mínimo, comparemos dos términos consecutivos k y $k+1$

$-1-ky+p\left(-\frac{x-k+1+1}{q-p}\right)+q\left(-\frac{x-k-1+1}{q-p}\right)<-1-(k+1)\,y+p\left(-\frac{x-(k+1)+1+1}{q-p}\right)+q\left(-\frac{x-(k+1)-1+1}{q-p}\right)$

$$-ky+p\left(-\frac{x-k+2}{q-p}\right)+q\left(-\frac{x-k}{q-p}\right)<-ky-y+p\left(-\frac{x-k+1}{q-p}\right)+q\left(-\frac{x-k-1}{q-p}\right)$$

$-p(x-k+2)-q(x-k)<-y(q-p)-p(x-k+1)-q(x-k-1)$
$-x+k-2p<-y(q-p)-x+k-p+q$
$-2p<-y(q-p)-p+q$
$-2p+p-q<-y(q-p)$
$-p-q<-y(q-p)$
$-1<-y(q-p)$
$1>y(q-p)$
por lo tanto el mínimo es el primer término.
Ahora comparemos el primero y último término

$$\begin{aligned}
-1+p\left(-\frac{x+2}{q-p}\right)+q\left(-\frac{x}{q-p}\right) &< -1-xy+p\left(-\frac{2}{q-p}\right)\\
-p(x+2)-q(x) &< -xy(q-p)-2p\\
-x-2p &< -xy(q-p)-2p\\
-1 &< -y(q-p)\\
1 &< y(q-p)
\end{aligned}$$

por lo que el mínimo es el primer término, es decir:

$$v_{T-1}(x)=-1-p\left(\frac{x+2}{q-p}\right)-q\left(\frac{x}{q-p}\right)$$

$$v_{T-1}(x)=-\frac{x+1}{q-p}$$

con $\varphi_{T-1}(x)=0=a.$

Para $T-2$

$$\begin{aligned}
v_{T-2}(x) &= \min_A\left\{l(x,a)+\sum_z v_{T-1}(z)\,p(z/x,a)\right\}\\
&= \min_A\left\{-1-ay+pv_{T-1}(x-a+1)+qv_{T-1}(x-a-1)\right\}\\
&= \min\left\{\begin{array}{c}
-1+p\left(-\frac{x+2}{q-p}\right)+q\left(-\frac{x}{q-p}\right),\\
-1-y+p\left(-\frac{x+1}{q-p}\right)+q\left(-\frac{x-1}{q-p}\right),\\
-1-2y+p\left(-\frac{x}{q-p}\right)+q\left(-\frac{x-2}{q-p}\right),\\
\ldots,\\
-1-(x-1)y+p\left(-\frac{3}{q-p}\right)+q\left(-\frac{1}{q-p}\right),\\
-1-xy+p\left(-\frac{2}{q-p}\right)
\end{array}\right\}
\end{aligned}$$

Y como podemos apreciar es idéntica a la etapa $T-1$. Por lo que la solución esta expresada por la forma:

$$v_{T-2}(x) = -\frac{x+1}{q-p}$$

$$\text{con } \varphi_{T-2}(x) = 0 = a.$$

De lo anterior concluimos que la solución esta dada por

$$v_t(x) = -\frac{x+1}{q-p} \ (x \geq 0),$$

y el mínimo es alcanzado por la acción $a = \varphi^0(x) \equiv 0$.

Por lo tanto, $\forall \pi$

$$E^{\pi}\left[\sum_{t=1}^{\infty}\left\{yj^1(\xi_{t-1},\eta_t) + j^2(\xi_{t-1},\eta_t)\right\}\right] \geq -\frac{x_0+1}{q-p}.$$

Pero la estrategia de control $\varphi^0(x) \equiv 0$ proporciona

$$E^{\varphi^0}\left[\sum_{t=1}^{\infty}\left\{yj^1(\xi_{t-1},\eta_t) + j^2(\xi_{t-1},\eta_t)\right\}\right] = E^{\varphi^0}\left[\sum_{t=1}^{\infty} j^2(\xi_{t-1},\eta_t)\right].$$

Esta esperanza matemática está expresada por la fórmula (5.11). Así en este caso

$$\mathrm{g}(y,1) = -\frac{x_0+1}{q-p} - yd. \tag{5.13}$$

(b) Ahora consideremos el siguiente caso $1 < y(q-p)$

Abordaremos el problema con horizonte finito $T < \infty$:

$$\min_{\pi}\left\{E^{\pi}\left[\sum_{t=1}^{T}\left\{yj^1(\xi_{t-1},\eta_t) + j^2(\xi_{t-1},\eta_t)\right\}\right]\right\}.$$

Hagamos la deducción de las ecuaciones de Bellman y consideremos el problema sin restricciones y el Lagrangiano es

$$l(y,P^{\pi}) = -1 - ay.$$

Para

$$v_t(-1) \equiv 0;$$

y el valor de salida

$$v_T(x) = 0, \text{ si } x \geq 0;$$

$$\begin{aligned} v_{T-1}(x) &= \min_A \left\{ l(x,a) + \sum_z v_T(z)\, p(z/x,a) \right\} \\ &= \min\{-1, -1-y, -1-2y, \ldots, -1-(x-1)y, -1-xy\} \end{aligned}$$

Comparemos dos términos consecutivos

$$\begin{aligned} -1-ky &< -1-(k+1)y \\ 0 &< -y\,! \end{aligned}$$

Por lo tanto

$$v_{T-1}(x) = -1-xy,$$

$$\text{con } \varphi_{T-1}(x) = x = a.$$

Para $T-2$

$$\begin{aligned} v_{T-2}(x) &= \min_A \left\{ l(x,a) + \sum_z v_{T-1}(z)\, p(z/x,a) \right\} \\ &= \min_A \{-1-ay+pv_{T-1}(x-a+1)+qv_{T-1}(x-a-1)\} \\ &= \min \left\{ \begin{array}{c} -1+p(-1-(x+1)y)+q(-1-(x-1)y), \\ -1-y+p(-1-xy)+q(-1-(x-2)y), \\ -1-2y+p(-1-(x-1)y)+q(-1-(x-3)y), \\ \ldots, \\ -1-(x-1)y+p(-1-2y)+q(-1), \\ -1-xy+p(-1-y) \end{array} \right\} \end{aligned}$$

Comparando dos consecutivos k y $k+1$

$-1-ky+p(-1-(x-k+1)y)+q(-1-(x-k-1)y) < -1-(k+1)y+$ $p(-1-(x-(k+1)+1)y)+q(-1-(x-(k+1)-1)y)$

$-ky+p(-1-xy+ky-y)+q(-1-xy+ky+y) < -ky-y+p(-1-xy+ky+y-y)+$ $q(-1-xy+ky+y+y)$

$-ky+p(-1-xy+ky-y)+q(-1-xy+ky+y) < -ky-y+p(-1-xy+ky+y-y)+$ $q(-1-xy+ky+y+y)$

$-1-xy+ky-yp+yq < -y-1-xy+ky+y-yp+yq$

$-yp+yq < -yp+yq$ de lo cual deducimos que es indistinto el término que se eliga.

Pero veamos que sucede con el primero y último términos:

$$
\begin{aligned}
-1+p\left(-1-(x+1)\,y\right)+q\left(-1-(x-1)\,y\right) &< -1-xy+p\left(-1-y\right)\\
p\left(-1-xy-y\right)+q\left(-1-xy+y\right) &< -xy-p-py\\
-1-xy-py+qy &< -xy-p-py\\
-1+qy &< -p\\
-1 &< -p-qy\\
q &< -qy\\
1 &< -y\ !
\end{aligned}
$$

De lo cual el mínimo es el último término

$$
\begin{aligned}
v_{T-2}\left(x\right) &= -1-xy+p\left(-1-y\right)\\
&= -1-xy-p-py\\
&= -1-p-y\left(-x-p\right)
\end{aligned}
$$

$$\text{con } \varphi_{T-2}\left(x\right)=x=a.$$

Para $T-3$

$$
\begin{aligned}
v_{T-3}\left(x\right) &= \min_{A}\left\{l\left(x,a\right)+\sum_{z}v_{T-2}\left(z\right)p\left(z/x,a\right)\right\}\\
&= \min_{A}\left\{-1-ay+pv_{T-2}\left(x-a+1\right)+qv_{T-2}\left(x-a-1\right)\right\}\\
&= \min\left\{\begin{array}{c}
-1+p\left(-1-p+y\left(-\left(x+1\right)-p\right)\right)+q\left(-1-p+y\left(-\left(x-1\right)-p\right)\right),\\
-1-y+p\left(-1-p+y\left(-x-p\right)\right)+q\left(-1-p+y\left(-\left(x-2\right)-p\right)\right),\\
-1-2y+p\left(-1-p+y\left(-\left(x-1\right)-p\right)\right)+q\left(-1-p+y\left(-\left(x-3\right)-p\right)\right),\\
\ldots,\\
-1-\left(x-1\right)y+p\left(-1-p+y\left(-2-p\right)\right)+q\left(-1-p+y\left(-p\right)\right),\\
-1-xy+p\left(-1-p+y\left(-1-p\right)\right)
\end{array}\right\}
\end{aligned}
$$

Comparando dos consecutivos k y $k+1$

$-1-ky+p\left(-1-p+y\left(-\left(x-k+1\right)-p\right)\right)+q\left(-1-p+y\left(-\left(x-k-1\right)-p\right)\right)<$
$-1-\left(k+1\right)y+p\left(-1-p+y\left(-\left(x-\left(k+1\right)+1\right)-p\right)\right)+q\left(-1-p+y\left(-\left(x-\left(k+1\right)-1\right)-p\right.\right.$

$-ky+p\left(-1-p+y\left(-x+k-1-p\right)\right)+q\left(-1-p+y\left(-x+k+1-p\right)\right)<$
$-ky-y+p\left(-1-p+y\left(-x+k+1-1-p\right)\right)+q\left(-1-p+y\left(-x+k+1+1-p\right)\right)$

$p\left(-1-p-yx+yk-y-yp\right)+q\left(-1-p-yx+yk+y-yp\right)<-y+$
$p\left(-1-p-yx+yk-yp\right)+q\left(-1-p-yx+yk+2y-yp\right)$

$-1-p-yx+yk-yp+yq-yp<-y-1-p-yx+yk-yp+2qy$

$yq-yp<-y+2qy$

$yq-2qy-yp<-y$

$-qy - yp < -y$

$-y < -y$. Cualquier término nos proporciona el mínimo pero ahora consideremos la comparación entre el primer y último términos:

$$\begin{aligned}
-1 + p\left(-1 - p + y\left(-\left(x+1\right) - p\right)\right) + q\left(-1 - p + y\left(-\left(x-1\right) - p\right)\right) &< \\
&-1 - xy + p\left(-1 - p + y\left(-1 - p\right)\right) \\
-1 + p\left(-1 - p + y\left(-x - 1 - p\right)\right) + q\left(-1 - p + y\left(-x + 1 - p\right)\right) &< \\
&-1 - xy + p\left(-1 - p - y - yp\right) \\
-1 + p\left(-1 - p - yx - y - yp\right) + q\left(-1 - p - yx + y - yp\right) &< \\
&-1 - xy - p - p^2 - py - yp^2
\end{aligned}$$

$$\begin{aligned}
-1 - 1 - p - yx - yp + yq - yp &< -1 - xy - p - p^2 - py - yp^2 \\
-1 + yq - yp &< -p^2 - yp^2 \\
-1 + y\left(q - p\right) &< -p^2 - yp^2
\end{aligned}$$

como $-1 + y\left(q-p\right) > 0$ tenemos $0 < -p^2 - yp^2$!, de lo cual el mínimo es el último término, es decir:

$$\begin{aligned}
v_{T-3}\left(x\right) &= -1 - xy + p\left(-1 - p + y\left(-1 - p\right)\right) \\
&= -\frac{1-p^3}{q} - y\left(x + \frac{p-p^3}{q}\right).
\end{aligned}$$

$$\text{con } \varphi_{T-3}\left(x\right) = x = a.$$

De lo expuesto arriba determinamos que la solución esta dada por la fórmula:

$$v_t\left(x\right) = -y\left(x + \frac{p - p^{T-t}}{q}\right) - \left(\frac{1-p^{T-t}}{q}\right) \text{ con } \left(t < T,\ x \geq 0.\right)$$

y el mínimo es alcanzado por la acción $a = \varphi^1\left(x\right) = x$. Donde, $\forall \pi$

$$E^{\pi}\left[\sum_{t=1}^{\infty}\left\{yj^1\left(\xi_{t-1},\eta_t\right) + j^2\left(\xi_{t-1},\eta_t\right)\right\}\right] \geq -y\left(x_0 + \frac{p}{q}\right) - \left(\frac{1}{q}\right)$$

esto debido a que $\lim_{T\to\infty}\left(x_0+\frac{p-p^{T-t}}{q}\right)=\left(x_0+\frac{p}{q}\right)$; de igual manera $\lim_{T\to\infty}\left(\frac{1-p^{T-t}}{q}\right)=\left(\frac{1}{q}\right)$ pero la estrategia de control proporciona

$$E^{\varphi^1}\left[\sum_{t=1}^{\infty}\left\{yj^1\left(\xi_{t-1},\eta_t\right)+j^2\left(\xi_{t-1},\eta_t\right)\right\}\right]=-y\left(x_0+\frac{p}{q}\right)-\left(\frac{1}{q}\right).$$

Así en este caso, tenemos que

$$\mathbf{g}\left(y,1\right)=-y\left(x_0+\frac{p}{q}\right)-\frac{1}{q}-yd. \tag{5.14}$$

Ahora analizaremos, el siguiente problema

$$\mathbf{J}^1\left(P^\pi\right)=\min_{\pi}\left\{E^\pi\left[\sum_{t=1}^{\infty}j^1\left(\xi_{t-1},\eta_t\right)\right]\right\}$$

s. a.

$$\mathbf{J}^2\left(P^\pi\right)\leq d.$$

Nuevamente consideraremos la solución sin restricciones por lo que el Lagrangiano es

$$\begin{aligned} l\left(y,P^\pi\right) &= j^1+yj^2 \\ &= -a-y \end{aligned}$$

(a) Abordaremos el primer caso $y>(q-p)$.

De nueva cuenta deduciremos las ecuaciones de Bellman y el caso finito $T<\infty$.

Como ya hemos visto:

$$v_T\left(-1\right)\equiv 0.$$

El valor de salida es

$$v_T\left(x\right)=-y\frac{x+1}{q-p}.$$

Para $T-1$,

$$
\begin{aligned}
v_{T-1}(x) &= \min_{A}\left\{ l(x,a) + \sum_{z} v_T(z)\, p(z/x,a) \right\} \\
&= \min_{A}\left\{ -a-y+pv_T(x-a+1)+qv_T(x-a-1) \right\} \\
&= \min \left\{ \begin{array}{c}
-y+p\left(-y\frac{x+1+1}{q-p}\right)+q\left(-y\frac{x-1+1}{q-p}\right), \\
-1-y+p\left(-y\frac{x+1}{q-p}\right)+q\left(-y\frac{x-2+1}{q-p}\right), \\
-2-y+p\left(-y\frac{x-1+1}{q-p}\right)+q\left(-y\frac{x-3+1}{q-p}\right), \\
\ldots, \\
-(x-1)-y+p\left(-\frac{2+1}{q-p}\right)+q\left(-y\frac{1}{q-p}\right), \\
-x-y+p\left(-y\frac{1+1}{q-p}\right)
\end{array} \right\}
\end{aligned}
$$

Comparemos dos términos consecutivos k y $k+1$

$$-k-y+p\left(-y\tfrac{x-k+1+1}{q-p}\right)+q\left(-y\tfrac{x-k-1+1}{q-p}\right) < -(k+1)-y+p\left(-y\tfrac{x-(k+1)+1+1}{q-p}\right)+ q\left(-y\tfrac{x-(k+1)-1+1}{q-p}\right)$$

$$
\begin{aligned}
&-p\left(y\tfrac{x-k+2}{q-p}\right)-q\left(y\tfrac{x-k}{q-p}\right) < -1-p\left(y\tfrac{x-k-1+1+1}{q-p}\right)-q\left(\tfrac{x-k-1-1+1}{q-p}\right) \\
&-py(x-k+2)-qy(x-k) < -(q-p)-py(x-k+1)-qy(x-k-1) \\
&-p(yx-yk+2y)-q(yx-yk) < -(q-p)-p(yx-yk+y)-q(yx-yk-y) \\
&-yx-yk-2yp < -(q-p)-yx-yk-py+qy \\
&-2yp+py-qy < -(q-p) \\
&-yp-qy < -(q-p) \\
&-y < -(q-p) \\
&y > (q-p).
\end{aligned}
$$

De aqui concluimos que el término menor es el primero pero ahora, verifiquemos el último y primero de los términos.

$$
\begin{aligned}
-y-p\left(y\frac{x+1+1}{q-p}\right)-q\left(y\frac{x-1+1}{q-p}\right) &< -x-y-p\left(y\frac{2}{q-p}\right) \\
-p(yx+2y)-qyx &< -x(q-p)-2py \\
-pyx-2yp-qyx &< -x(q-p)-2py \\
-xy &< -x(q-p) \\
-y &< -(q-p) \\
y &> (q-p).
\end{aligned}
$$

por lo tanto el mínimo es el primer término con $\varphi_{T-1}(x) = 0 = a.$

$$v_{T-1}(x) = -y - p\left(y\frac{x+1+1}{q-p}\right) - q\left(y\frac{x-1+1}{q-p}\right)$$

$$v_{T-1}(x) = -y\frac{x+1}{q-p}$$

con $\varphi_{T-1}(x) = 0 = a.$

Ahora para $T-2$

$$\begin{aligned}
v_{T-2}(x) &= \min_A \left\{ l(x,a) + \sum_z v_{T-1}(z)\, p(z/x,a) \right\} \\
&= \min_A \{-a - y + p v_T(x-a+1) + q v_T(x-a-1)\} \\
&= \min \left\{ \begin{array}{c}
-y + p\left(-y\frac{x+1+1}{q-p}\right) + q\left(-y\frac{x-1+1}{q-p}\right), \\
-1 - y + p\left(-y\frac{x+1}{q-p}\right) + q\left(-y\frac{x-2+1}{q-p}\right), \\
-2 - y + p\left(-y\frac{x-1+1}{q-p}\right) + q\left(-y\frac{x-3+1}{q-p}\right), \\
\ldots, \\
-(x-1) - y + p\left(-\frac{2+1}{q-p}\right) + q\left(-y\frac{1}{q-p}\right), \\
-x - y + p\left(-y\frac{1+1}{q-p}\right)
\end{array} \right\}
\end{aligned}$$

Comparemos dos términos consecutivos k y $k+1$

$$-k-y+p\left(-y\frac{x-k+1+1}{q-p}\right)+q\left(-y\frac{x-k-1+1}{q-p}\right) < -(k+1)-y+p\left(-y\frac{x-(k+1)+1+1}{q-p}\right)+ q\left(-y\frac{x-(k+1)-1+1}{q-p}\right)$$

$$\begin{aligned}
&-p\left(y\frac{x-k+2}{q-p}\right) - q\left(y\frac{x-k}{q-p}\right) < -1 - p\left(y\frac{x-k-1+1+1}{q-p}\right) - q\left(\frac{x-k-1-1+1}{q-p}\right) \\
&-py(x-k+2) - qy(x-k) < -(q-p) - py(x-k+1) - qy(x-k-1) \\
&-p(yx-yk+2y) - q(yx-yk) < -(q-p) - p(yx-yk+y) - q(yx-yk-y) \\
&-yx - yk - 2yp < -(q-p) - yx - yk - py + qy \\
&-2yp + py - qy < -(q-p) \\
&-yp - qy < -(q-p) \\
&-y < -(q-p)
\end{aligned}$$

$y > (q-p)$. De lo cual el mínimo es el término $k-\textit{ésimo}$.

Comparando entre el término de la primera acción y el de la última

$$-y - p\left(y\frac{x+1+1}{q-p}\right) - q\left(y\frac{x-1+1}{q-p}\right) \;<\; -x - y - p\left(y\frac{2}{q-p}\right)$$

$$\begin{aligned}
-p\left(yx+2y\right)-qyx &< -x\left(q-p\right)-2py \\
-pyx-2yp-qyx &< -x\left(q-p\right)-2py \\
-xy &< -x\left(q-p\right) \\
-y &< -\left(q-p\right) \\
y &> \left(q-p\right).
\end{aligned}$$

por lo que el mínimo es el primer término.

Por lo que la solución esta expresada por la forma:

$$\begin{aligned}
v_{T-2}\left(x\right) &= -y-py\frac{x+1+1}{q-p}-qy\frac{x-1+1}{q-p} \\
&= -y\frac{x+1}{q-p}.
\end{aligned}$$

$$\text{con } \varphi_{T-2}\left(x\right)=0=a.$$

Por lo que las ecuaciones de Bellman son:

$$\left.\begin{array}{c}
v_T\left(x\right)=-y\frac{x+1}{q-p},\ \text{si } x\geq 0;\ v_t\left(-1\right)\equiv 0; \\
v_{t-1}\left(x\right)=\min\limits_{A}\left\{-a-y+pv_t\left(x-a+1\right)+qv_t\left(x-a-1\right)\right\}
\end{array}\right\}.$$

Y como hemos visto su solución esta expresada por

$$v_t\left(x\right)=-y\frac{x+1}{q-p}\ \ \left(x\geq 0\right),$$

y el mínimo es alcanzado por la acción $a=\varphi^0\left(x\right)\equiv 0$.

Donde, $\forall\pi$

$$E^{\pi}\left[\sum_{t=1}^{\infty}\left\{j^1\left(\xi_{t-1},\eta_t\right)+yj^2\left(\xi_{t-1},\eta_t\right)\right\}\right]\geq -y\frac{x_0+1}{q-p}.$$

Pero la estrategia de control $\varphi^0\left(x\right)\equiv 0$ proporciona

$$E^{\varphi^0}\left[\sum_{t=1}^{\infty}\left\{j^1\left(\xi_{t-1},\eta_t\right)+yj^2\left(\xi_{t-1},\eta_t\right)\right\}\right]=yE^{\varphi^0}\left[\sum_{t=1}^{\infty}j^2\left(\xi_{t-1},\eta_t\right)\right]$$

y esta esperanza matématica esta expresada por la fórmula (5.11). Así en este caso

$$g(1,y) = -y\frac{x_0+1}{q-p} - yd. \tag{5.15}$$

(*b*) Abordaremos el caso $y < q-p$
Valor de salida:

$$v_T(x) = 0, \quad (x \geq 0).$$

De lo cual, veamos que sucede en la etapa $T-1$:

$$\begin{aligned} v_{T-1}(x) &= \min_A \left\{ l(x,a) + \sum_z v_T(z)\, p(z/x,a) \right\} \\ &= \min_A \left\{ -a - y + pv_T(x-a+1) + qv_T(x-a-1) \right\} \\ &= \min \left\{ -y, -1-y, \ -2-y, \ldots, -(x-1)-y, \ -x-y \right\} \end{aligned}$$

Comparando dos términos consecutivos k y $k+1$:

$$\begin{aligned} -k-y &< -(k+1)-y \\ 0 &< -1\,! \end{aligned}$$

de ello, el menor es el término $k+1$, ahora si consideramos el primer y último términos:

$$\begin{aligned} -y &< -x-y \\ 0 &< -x\,! \end{aligned}$$

Por lo tanto el mínimo es el último término, es decir:

$$v_{T-1}(x) = -x-y \text{ con } \varphi_{T-1}(x) = x = a.$$

Para $T-2$

$$\begin{aligned} v_{T-2}(x) &= \min_A \left\{ l(x,a) + \sum_z v_{T-1}(z)\, p(z/x,a) \right\} \\ &= \min_A \left\{ -a - y + pv_{T-1}(x-a+1) + qv_{T-1}(x-a-1) \right\} \\ &= \min \left\{ \begin{array}{c} -y + p(-(x+1)-y) + q(-(x-1)-y), \\ -1-y+p(-x-y)+q(-(x-2)-y), \\ -2-y+p(-(x-1)-y)+q(-(x-3)-y), \\ \ldots, \\ -(x-1)-y+p(-2-y)+q(-0-y), \\ -x-y+p(-1-y) \end{array} \right\} \end{aligned}$$

Comparando dos consecutivos k y $k+1$

$-k-y+p\left(-\left(x-k+1\right)-y\right)+q\left(-\left(x-k-1\right)-y\right)<-\left(k+1\right)-y+$
$p\left(-\left(x-\left(k+1\right)+1\right)-y\right)+q\left(-\left(x-\left(k+1\right)-1\right)-y\right)$
$-k-y+p\left(-x+k-1-y\right)+q\left(-x+k+1-y\right)<-k-1-y+p\left(-\left(x-k-1+1\right)-y\right)+$
$q\left(-\left(x-k-1-1\right)-y\right)$
$-k-y-x+k-p+q-y<-k-1-y+p\left(-x+k+1-1-y\right)+$
$q\left(-x+k+1+1-y\right)$
$-k-2y-x+k-p+q<-k-1-y+p\left(-x+k-y\right)+q\left(-x+k+2-y\right)$
$-2y-x-p+q<-k-1-y-x+k-y+2q$
$-2y-x-p+q<-1-2y-x+2q$
$-p-2q+q<-1$
$-1<-1.$

De ello, podemos ver que cualquier término es indistinto, por lo que procedemos a comparar el primer y últimos términos:

$$\begin{aligned}
-y+p\left(-\left(x+1\right)-y\right)+q\left(-\left(x-1\right)-y\right) &< -x-y+p\left(-1-y\right)\\
-y+p\left(-x-1-y\right)+q\left(-x+1-y\right) &< -x-y+p\left(-1-y\right)\\
-y-x-p+q-y &< -x-y-p-py\\
-2y-x-p+q &< -x-y-p-py\\
-y+q &< -py\\
q &< y-py\\
q &< y\left(1-p\right)\\
q &< yq\\
1 &< y
\end{aligned}$$

veamos que sucede con esta desigualdad, como $(q-p)>0$ múltipliquemos esta de lo cual tenemos:

$$(q-p)<(q-p)\,y$$

pero sabemos que $y<(q-p)$, por lo que:

$$\begin{aligned}
y &< (q-p)<(q-p)\,y\\
y &< (q-p)\,y\\
1 &< (q-p)!.
\end{aligned}$$

De lo cual concluimos que el mínimo es el último término,

$$\begin{aligned}
v_{T-2}\left(x\right) &= -x-y+p\left(-1-y\right)\\
&= -\left(x+p\right)-y\left(1+p\right).
\end{aligned}$$

$$\text{con } \varphi_{T-2}(x) = x = a.$$

Consideremos ahora $T-3$:

$$\begin{aligned}
v_{T-3}(x) &= \min_{A} \left\{ l(x,a) + \sum_{z} v_{T-2}(z)\, p(z/x,a) \right\} \\
&= \min_{A} \left\{ -a - y + p v_{T-2}(x-a+1) + q v_{T-2}(x-a-1) \right\} \\
&= \min \left\{ \begin{array}{c}
-y + p\left[-(x+1+p) - y(1+p)\right] + q\left[-(x-1+p) - y(1+p)\right], \\
-1 - y + p\left[-(x+p) - y(1+p)\right] + q\left[-(x-2+p) - y(1+p)\right], \\
-2 - y + p\left[-(x-1+p) - y(1+p)\right] + q\left[-(x-3+p) - y(1+p)\right], \\
\ldots, \\
-(x-1) - y + p\left[-(2+p) - y(1+p)\right] + q\left[-(0+p) - y(1+p)\right], \\
-x - y + p\left[-(1+p) - y(1+p)\right].
\end{array} \right.
\end{aligned}$$

Volvamos a comparar dos términos consecutivos k y $k+1$

$$\begin{aligned}
&-k-y+p\left[-(x-k+1+p) - y(1+p)\right]+q\left[-(x-k-1+p) - y(1+p)\right] < \\
&-(k+1)-y+p\left[-(x-k-1+1+p) - y(1+p)\right]+q\left[-(x-k-1-1+p) - y(1+p)\right] \\
&-k-y+p\left[-x+k-1-p-y(1+p)\right]+q\left[-x+k+1-p-y(1+p)\right] < \\
&-k-1-y+p\left[-x+k+1-1-p-y(1+p)\right]+q\left[-x+k+1+1-p-y(1+p)\right] \\
&-k-y-x+k-p+q-p-y(1+p) < -k-1-y-x+k+1-p+q-p-y(1+p) \\
&-y-x-p+q-p-y(1+p) < -y-x-p+q-p-y(1+p) \\
&-y-x-2p+q-y(1+p) < -y-x-2p+q-y(1+p) \\
&-y-x-2p+q-y(1+p) < -y-x-2p+q-y(1+p) \\
&0 < 0.
\end{aligned}$$

De nueva cuenta es indistinto tomar cualquier término distinto al último, por lo que procedamos a comparar los términos arrojados por las acciones $x-1$ y x, de lo cual tenemos:

$$\begin{aligned}
-(x-1) - y + p\left[-(2+p) - y(1+p)\right] + q\left[-p - y(1+p)\right] &< \\
&-x - y + p\left[-(1+p) - y(1+p)\right] \\
-x + 1 - y + p\left[-2 - p - y(1+p)\right] + q\left[-p - y(1+p)\right] &< \\
&-x - y + p\left[-1 - p - y(1+p)\right]
\end{aligned}$$

$$\begin{aligned}
-x + 1 - y - 2p - p - y(1+p) &< -x - y - p - p^2 - py(1+p) \\
1 - 2p - y(1+p) &< -p^2 - py(1+p) \\
1 - 2p + p^2 &< y(1+p) - py(1+p) \\
(1-p)^2 &< y(1+p)\, q
\end{aligned}$$

$$
\begin{aligned}
q^2 &< y(1+p)\,q \\
q &< y(1+p)
\end{aligned}
$$

analizemos esta última desigualdad, sabemos que $y < (q-p)$ de aqui, tenemos:

$$q < y(1+p) < (q-p)(1+p)$$

$$q < (q-p)(1+p)\ !$$

Por lo que concluimos que el mínimo es el término proporcionado por la acción x. Es decir:

$$
\begin{aligned}
v_{T-3}(x) &= -x-y+p\left[-(1+p)-y(1+p)\right] \\
&= -\left[x+\frac{p-p^3}{q}\right]-y\left[\frac{1-p^3}{q}\right].
\end{aligned}
$$

$$\text{con } \varphi_{T-3}(x)=x=a.$$

De lo cual podemos ver que la solución esta expresada por la fórmula:

$$v_t(x)=-\left[x+\frac{p-p^{T-t}}{q}\right]-y\left[\frac{1-p^{T-t}}{q}\right]\quad (t<T,\ x\geq 0),$$

y el mínimo es alcanzado por la acción $a=\varphi^1(x)=x$.

Donde, $\forall\pi$

$$E^{\pi}\left[\sum_{t=1}^{\infty}\left\{j^1(\xi_{t-1},\eta_t)+yj^2(\xi_{t-1},\eta_t)\right\}\right]\geq-\left(x_0+\frac{p}{q}\right)-\frac{y}{q}.$$

Pero la estrategía de control $\varphi^1(x)=x$ da:

$$E^{\varphi^1}\left[\sum_{t=1}^{\infty}\left\{j^1(\xi_{t-1},\eta_t)+yj^2(\xi_{t-1},\eta_t)\right\}\right]=-\left(x_0+\frac{p}{q}\right)-\frac{y}{q}.$$

Así en este caso

$$\mathrm{g}(1,y)=-\left(x_0+\frac{p}{q}\right)-\frac{y}{q}-yd. \tag{5.16}$$

Presentamos las principales propiedades de las estrategias $\varphi^0(x)\equiv 0$ y $\varphi^1(x)=x$:

$$
\left.
\begin{array}{ll}
\min\limits_{\pi}\mathbf{J}^2(P^{\pi})=\mathbf{J}^2\left(P^{\varphi^0}\right)=-\frac{x_0+1}{q-p}; & \max\limits_{\pi}\mathbf{J}^2(P^{\pi})=\mathbf{J}^2\left(P^{\varphi^0}\right)=0; \\
\min\limits_{\pi}\mathbf{J}^1(P^{\pi})=\mathbf{J}^1\left(P^{\varphi^1}\right)=-x_0-\frac{p}{q}; & \max\limits_{\pi}\mathbf{J}^2(P^{\pi})=\mathbf{J}^2\left(P^{\varphi^1}\right)=-\frac{1}{q}.
\end{array}
\right\} \tag{5.17}
$$

5.4 Resolviendo el problema (alternativa I)

Encontremos el $\max_{y\geq 0} g(y,1)$. Iniciemos nuestro análisis considerando $d \geq 0$ entonces la restricción (5.9) es inesencial: la desigualdad $\mathrm{J}^1(P^\pi) \leq 0$ es válida para alguna estrategia de control. Si $d < -x_0 - p/q$ entonces ahí no existen estrategias admisibles para las cuales la restricción (5.9) es satisfecha: $\min_\pi \mathrm{J}^1(P^\pi) = g(1,0) = -x_0 - p/q$. (Esto se deriva del Lema 5.1 y de las expresiones (5.17).)

Consideraremos el caso $-x_0 - p/q < d < 0$. Asumiremos $\varphi^1(x) = x$. Entonces $\mathrm{J}^1\left(P^{\varphi^1}\right) = g(1,0) = -x_0 - p/q$ y $\mathrm{J}^2\left(P^{\varphi^1}\right) = -1/q$: obviamente, $\mathrm{J}^2\left(P^{\varphi^1}\right)$ coincide, en el signo correcto, con el número esperado de pasos antes de la primera implementación del valor $\zeta_t = 2$. De esta manera podemos usar el Teorema 4.1.

Veamos donde alcanza el máximo la función $\max_{y\geq 0} g(y,1)$, como

$$\begin{aligned}
\max_{0\leq y\leq \frac{1}{q-p}} g(y,1) &= \max_{0\leq y\leq \frac{1}{q-p}} \left\{-\frac{x_0+1}{q-p} - yd\right\} \\
&= -\frac{x_0+1}{q-p} - \max_{0\leq y\leq \frac{1}{q-p}} \{yd\} \\
&= -\frac{x_0+1}{q-p} - \frac{d}{q-p} \\
&= -\frac{x_0+1+d}{q-p}.
\end{aligned}$$

Claramente, el máximo es alcanzado por $y^* = 1/(q-p)$ y todas las estrategias estan dadas por $L(P^\pi, y^*, 1) = \min_\pi L(P^\pi, y^*, 1) = -(x_0+1+d)/(q-p)$. Esto recuerda la construcción de una estrategia π para la cual la igualdad $\mathrm{J}^1(P^\pi) = d$ es válida.

Restringiremos nuestras consideraciones a el conjunto de estrategias estacionarias aleatorizadas del siguiente tipo:

$$\begin{array}{lll}
\pi(0|x) = 1, & & \text{si } x \leq \prod; \\
\pi(x-\prod|x) = \gamma; & \pi(x-\prod|x) = \delta, & \text{si } x > \prod;
\end{array}$$

donde $\prod$ es algún número entero no negativo; $\delta \geq 0$, $\gamma \geq 0$, $\delta+\gamma = 1$.

Obviamente, $J^1(P^\pi) = w(x_0)$, donde la función $w(x) \neq \infty$ satisface el siguiente conjunto de ecuaciones

$$\left.\begin{array}{lcll} w(-1) & = & 0; & \\ w(x) & = & w(x+1)p + w(x-1)q, & \text{si } 0 \leq x \leq \prod; \\ & & -\gamma(x - \prod) - \delta(x - \prod - 1) & \\ w(x) & = & +\gamma[w(\prod + 1)p + w(\prod - 1)q] & \\ & & +\delta[w(\prod + 2)p + w(\prod)q], & \text{si } x > \prod. \end{array}\right\} \quad (5.18)$$

La solución esta expresada por la fórmula

$$w(x) = \begin{cases} \frac{p(1-\delta q)}{(1-\delta p)(q-p)}\left[\left(\frac{p}{q}\right)^{\prod+1} - \left(\frac{p}{q}\right)^{\prod - x}\right], & \text{si } x \leq \prod + 1; \\ w(\prod + 1) - x + \prod + 1, & \text{si } x > \prod + 1. \end{cases}$$

La función $[p(1-\delta q)]/[(1-\delta q)(q-p)]$ disminuye cuando $\delta \in [0,1]$ crece. Por lo tanto como $\prod$ varía desde 0 hasta $+\infty$ y δ varia desde 0 hasta 1, la función $J^1(P^\pi) = w(x_0)$ crece monótonamente desde $-x_0 - p/q$ hasta cero y llega a ser igual a d en alguna $\prod^*$ y δ^*. La correspondiente estrategia estacionaria aleatorizada π^* es óptima en el problema (5.8), (5.9) por el Teorema 4.1, Item (b). El valor mínimo de la función $J^2(\cdot)$ bajo la restricción (5.9) es igual a

$$g(y^*, 1) = -\frac{x_0 + 1 + d}{q - p}. \quad (5.19)$$

Esto es fácil de entender que si $d = -x_0 - p/q$ entonces una solución del problema (5.8), (5.9) esta dada por la regla de decisión $\varphi^1(x) = x$. La restricción (5.9) es siempre esencial si $-x_0 - p/q \leq d < 0$ de conformidad con la fórmula (5.11) y (5.19). En el caso $d \geq 0$ una solución esta dada por la regla de decisión φ^0.

5.5 Resolviendo el problema (alternativa II)

Abordaremos la alternativa $\max_{y \geq 0} g(1, y)$. De acuerdo a el Lema 5.1(a)

$$\inf_\pi J^1(P^\pi) = J^1\left(P^{\varphi^1}\right) = -x_0 - \frac{p}{q}$$

y la restricción (5.10) es inesencial si $d \geq \mathbf{J}^2\left(P^{\varphi^1}\right) = -1/q$. En el caso $d < -(x_0+1)/(q-p)$ ahí no existen estrategias admisibles, de acuerdo con (5.17).

Consideraremos el caso $-(x_0+1)/(q-p) < d < -1/q$. Entonces $\mathbf{J}^2\left(P^{\varphi^0}\right) = g(0,1) = -(x_0+1)/(q-p)$ y $\mathbf{J}^1\left(P^{\varphi^0}\right) = 0$.

Como antes, veamos donde alcanza el máximo la función $\max_{y\geq 0} g(1,y)$, como

$$
\begin{aligned}
\max_{0\leq y\leq q-p} g(y,1) &= \max_{0\leq y\leq q-p}\left\{-\left(x_0+\frac{p}{q}\right)-\frac{y}{q}-yd\right\}\\
&= -\left(x_0+\frac{p}{q}\right)-\max_{0\leq y\leq q-p}\left\{\frac{y}{q}+yd\right\}\\
&= -\left(x_0+\frac{p}{q}\right)-\frac{q-p}{q}-(q-p)\,d\\
&= -x_0-\frac{p}{q}-1+\frac{p}{q}-(q-p)\,d\\
&= -x_0-1-(q-p)\,d\\
&= -(x_0+1)-(q-p)\,d.
\end{aligned}
$$

Claramente, el máximo de $\max_{y\geq 0} g(1,y)$ es alcanzado por $y^* = q-p$ y todas las estrategias estan dadas por $L(P^\pi,1,y^*) = \min_\pi L(P^\pi,1,y^*) = -(x_0+1)-(q-p)\,d$. Esto nos recuerda construir una estrategia π para la cual la igualdad $\mathbf{J}^2(P^\pi) = d$ es válida.

Nuestra consideración se restringe a el conjunto de estrategias estacionarias aleatorizadas del siguiente tipo:

$$
\begin{array}{lll}
\pi(0|\,x) = 1, & & \text{si } x \leq \prod;\\
\pi(x-\prod|\,x) = \gamma; & \pi(x-\prod-1|\,x) = \delta, & \text{si } x > \prod,
\end{array}
$$

donde $\prod$ es algún número entero no negativo; $\delta \geq 0$, $\gamma \geq 0$, $\delta+\gamma = 1$.

Obviamente, $\mathbf{J}^2(P^\pi) = w(x_0)$, donde la función $w(x) \neq \infty$ satisface el siguiente conjunto de ecuaciones

$$
\left.
\begin{array}{lcll}
w(-1) &=& 0; & \\
w(x) &=& -1+w(x+1)\,p+w(x-1)\,q, & \text{si } 0\leq x\leq \prod;\\
w(x) &=& \begin{array}{l}-1+\gamma\left[w(\prod+1)\,p+w(\prod-1)\,q\right]\\ +\delta\left[w(\prod+2)\,p+w(\prod)\,q\right],\end{array} & \text{si } x > \prod.
\end{array}
\right\}
\tag{5.20}
$$

La solución esta expresada por la fórmula

$$w(x)=\begin{cases} -\frac{p(1-\delta q)}{(1-\delta p)(q-p)^2}\left[\left(\frac{p}{q}\right)^{\prod+1}-\left(\frac{p}{q}\right)^{\prod-x}\right]-\frac{x+1}{q-p}, & \text{si } x\leq \prod+1; \\ w\left(\prod+1\right), & \text{si } x>\prod+1. \end{cases}$$

La función $[p(1-\delta q)]/\left[(1-\delta q)(q-p)^2\right]$ disminuye cuando $\delta\in[0,1]$ crece. Por lo tanto como $\prod$ varía desde 0 hasta $+\infty$ y δ varia desde 0 hasta 1, la función $\mathbf{J}^2(P^\pi)=w(x_0)$ crece monótonamente desde $-1/q$ hasta $-(x_0+1)/(q-p)$ y llega a ser igual a d en alguna $\prod^*$ y δ^*. La correspondiente estrategia estacionaria aleatorizada π^* es óptima en el problema (5.6), (5.10) por el Teorema 4.1, Item (b). El valor mínimo de la función $\mathbf{J}^2(\cdot)$ bajo la restricción (5.10) es igual a

$$\mathrm{g}(1,y^*)=-x_0-1-d(q-p) \tag{5.21}$$

Es fácil de entender que si $d=-(x_0+1)/(q-p)$ entonces una solución a el problema (5.6), (5.10) esta dada por la regla de decisión $\varphi^0(x)\equiv 0$. La restricción (5.10) es siempre esencial si $-(x_0+1)/(q-p)\leq d<-1/q$ de conformidad con la fórmula (5.17) y (5.21). En el caso $d\geq -1/q$ una solución esta dada por la regla de decisión φ^1.

El modelo considera la inecuación $p<q$ válida, esto es, la esperanza matemática de las responsabilidades $2q$ es tan grande que la prima asegurada $h=1$. Tal situación ocurre en algunos tipos de seguros obligatorios donde a falta de la parte el pago es remunerado a la compañía aseguradora por el estado. Es útil resaltar en esta conección que el valor $y^*=q-p$ coincide con el subsidio mínimo del estado para el cual la aseguradora está libre de pérdidas (en la media). Si la compañía recibe tal subsidio que es renumerado a los socios en paralelo con los dividendos entonces el criterio ejecutado

$$E^\pi\left[\sum_{t=1}^{\tau}\eta_t+y^*\tau\right] \tag{5.22}$$

no depende sobre la estrategia π y el estado puede esperar que la firma estará ocupada en sus negocios asegurados. Si el subsidio es ligeramente grande el valor crítico $q-p$ entonces la estrategia óptima φ^0 provee el máximo a la expresión (5.22) también proporciona el máximo para el tiempo esperado de vida de la compañía aseguradora hasta el colapso.

A

Conclusión

A lo largo del presente trabajo sentamos los argumentos tanto elementales como complejos en cuestiones de probabilidad, como los sustenta el Capítulo 1; para dar paso a los resultados de los procesos estocásticos lo cual se dejo apreciar en el Capítulo 2.

Con dichas herramientas nos introducimos a el modelo de control Markoviano, estudiando el modelo sin restricciones y el modelo restringido, lo cual se manifesto en los Capítulos 3 y 4 respectivamente.

Durante este recorrido logramos reconocer como esta conformado primeramente un modelo de Control en partícular el Markoviano; a través de ello resolvimos un ejemplo de "colas", con el cual fuimos de lo sencillo a lo complejo con el objetivo de fijar ideas. En este sentido logramos apreciar los resultados que se pueden obtener cuando se aplica un modelo no restringido y restringido, básicamente los resultados económicos de no tener una adecuada "administración" en una cola o fila pues aunque parece simple, el llegar a una formación o estar en ella implica para ambos interesados pérdidas o ganancias y es que no se trata de encontrar una fila vacia sino de óptimizar de la mejor manera los recursos en tiempo y dinero. Al respecto el modelo de control proporcionó la mejor estrategía para equilibrar la óptimalidad para los interesados.

Continuando en la misma línea, llegamos al problema concreto de plantear un modelo de control a una compañía de seguros, el cual, el Capítulo 5 es el encargado de mostrar; apreciamos que los modelos matemáticos, para nuestro caso los de control, representan de manera objetiva la situación de un fenómeno. Empero, vimos que otro factores tienen lugar en la toma de decisiones y por consiguiente un modelo matemático sí es aplicable. Es por ello que este capítulo deja de manifiesto que la compañía aseguradora puede vivir tanto en función de los dividendos que esta proporcione a sus accionistas. A pesar de ello esto no se logra del todo cuando intervienen las instituciones de gobierno y no sólo eso sino también la libertad de los accionistas o funcionarios de las compañías, los cuales en lugar de tomar esto como una herramienta lo ven como un desplazamiento a sus capacidades de tomar decisiones.

Como nos hemos percatado en ambas situaciones se han proporcionado resultados orientadores de políticas gracias a el modelo de control, desafortunadamente estas herramientas escapan de su aplicación en países como el nuestro pues aquí impera más lo que digan los "jefes". Sin embargo el presente trabajo demuestra la gran útilidad y apoyó que estos pueden mostrar haciendo una aplicación adecuada. Por lo que es trabajo de los profesionistas obtener y aplicar modelos que proporcionen estimaciones confiables para de esta forma incursionar y aplicar esta cultura matemática.

Bibliografía

[1] Canavos, George C.

Probabilidad y Estadística.

Editorial McGraw-Hill.

México, 1993.

Primera Edición.

PP. 651

[2] Cameron, Neil

Introduction to linear and convex programming.

University of Cambridgel.

Great Britain, 1985.

First published.

PP. 149

[3] Feller, William.

Introducción a la Teoría de Probabilidades y sus aplicaciones (volumen I).

Editorial Limusa.

México, 1973.

Primera Edición.

PP. 504

[4] Hamdy A. Taha.

Investigación de Operaciones.

Editorial Alfaomega
México, D.F., 1991.
Segunda Edición.
PP. 989

[5] Hans Bühlmann.
Mathematical Methods in Risk Theory.
Ed. Springer-Verlan Berlin.
Germany, 1996.
Second Printing.
PP. 210

[6] Harold J. Larson.
Introducción a la Teoría de la Probabilidades e Inferencia Estadística..
Editorial Limusa.
México, 1993.
Décima reimpresión.
PP. 466

[7] Hoel P.
Introduction to Probability Theory
Honghton Mifflin, Co.
1971.

[8] Karl H. Borch.
Economics of Insurance.
Ed. North-Holland.
Printed in the Netherlands, 1990.
Second Edition.
PP. 402

[9] Luque Vásquez Fernando.
Introducción a la Teoría de Control Estocástico
Ed. Universidad de Sonora.
Impreso en Hermosillo, Sonora. 1996.
PP. 79

[10] Mokhtar S. Bazaraa
Nonlinear Programming Theory and Algorithms.
Ed. John Wiley & Sons, Inc.
United States of America, 1993.
Second Edition.
PP. 637

[11] Peressini, A. L.
The Mathematics of Nonlinear Programming
Ed. Springer.
New York, 1988.
PP. 269

[12] Piunovskiy, A. B.
Optimal Control of Random Sequences in Problems with Constrainsts.
Ed. Kluwe Academic Publishers.
Printed in the Netherlands, 1997.
PP. 345

[13] Puterman, Martin L.
Markov Decision Processes.
Ed. John Wiley & Sons, Inc.
United States of America, 1994.
First Edition.
PP. 648

[14] Rockafellar, R. Tyrrell
Convex Analysis
Princeton Univ. Press.
Princeton (N.Y.) 1970
First Edition.
PP. 451

[15] Rodney Coleman.
Procesos Estocásticos (Volumen 14).
Editorial Limusa.
México, 1976.
Primera Edición.
PP. 131

[16] Ross S.
A First Course in Probability.
Macmillan
New York, 1976.

[17] Ross S.
Introduction to Probability Models.
Academic Press, Inc.
Florida, 1980.

[18] Rusell L. Ackoff.
Fundamentos de Investigación de Operaciones.
Editorial Limusa.
México, D.F., 1987.
Séptima Reimpresión.
PP. 502

[19] Suen Danp.
Nonlinear and Dynamic Programming An Introduction.
Ed. Springer-Verlag.
Printed in Austria, 1975.
PP. 164

[20] T. Pentikäinen and E.Personen.
Risk Theory; The Stochastic Basis of Insurance.
Ed. Chapman and Hall.
Great Britain, 1977.
Third Edition.
PP. 408