



THE UNIVERSITY OF NEW SOUTH  
WALES

SCHOOL OF ELECTRICAL ENGINEERING AND  
COMPUTER SCIENCE AND ENGINEERING

*Trabajadores Flotantes  
en Procesos de  
Producción  
Asignaciones  
Dinámicas Óptimas*

Ismael Reyes González

Maestro en Ciencias

Enero del 2009

*Supervisor:*

*Assessor:* Dr. Raúl Montes de Oca M.

# Índice general

Índice general	I
Índice de figuras	III
Índice de cuadros	IV
<b>1. Introducción</b>	<b>1</b>
<b>2. Preliminares Matemáticos</b>	<b>4</b>
2.1. Cadena de Decisión de Markov . . . . .	5
2.1.1. Políticas . . . . .	6
2.2. Distribución Condicional de Costos . . . . .	7
2.3. Criterios de Optimización . . . . .	8
2.4. Método de Aproximación Secuencial . . . . .	10
2.5. Cadena de Decisión de Markov a Tiempo Continuo . . . . .	12
2.5.1. Optimización Costo Promedio de una CDMTC . . . . .	13
<b>3. Una Línea de Producción con K Trabajadores Especializados y un Trabajador Flotante</b>	<b>19</b>
3.1. Descripción del Modelo . . . . .	19
3.2. La Formulación del Modelo como PDMTC . . . . .	20
3.3. Optimalidad, Estabilidad y Acotamiento . . . . .	22
3.4. Ecuación de Iteración de Valores . . . . .	30
<b>4. N Líneas de Producción con NK Trabajadores Especializados y un Trabajador Flotante</b>	<b>32</b>
4.1. Descripción del Problema . . . . .	32
4.2. Formulación del Modelo como un PDMTC . . . . .	33
4.3. Optimalidad, Estabilidad y Acotamiento . . . . .	36
4.4. Ecuación de Iteración de Valores . . . . .	42

---

<b>5. Conclusiones</b>	<b>44</b>
<b>6. Notación y Terminología</b>	<b>46</b>
6.1. Notación . . . . .	46
6.2. Lista de Abreviaturas . . . . .	48
<b>A. Procesos Estocásticos</b>	<b>49</b>
A.1. Conceptos Generales. . . . .	49
A.2. Proceso de Poisson. . . . .	50
A.3. Proceso de Markov . . . . .	54
A.3.1. Proceso de Markov en Tiempo Continuo . . . . .	54
A.3.2. Cadena de Markov a Tiempo Discreto . . . . .	56
A.4. Colas Poissonianas . . . . .	59
A.4.1. El Sistema M/M/1 . . . . .	59
<b>B. Complementos a las Demostraciones</b>	<b>61</b>
B.1. Algoritmo de Iteración de Valores (AIV) . . . . .	62
B.2. Computo de Políticas Costo Promedio Óptimas para Espacio de Estados Infinitos . . . . .	64
B.2.1. Suposiciones (AC) . . . . .	64
B.3. Proceso de Decisión Semi-Markoviano . . . . .	67
<b>C. Equivalencia de PDMs a Tiempo Continuo y a Tiempo Discreto</b>	<b>70</b>
C.1. Equivalencia . . . . .	71
<b>D. Bibliografía</b>	<b>75</b>

# Índice de figuras

# Índice de cuadros

# Capítulo 1

## Introducción

En el desarrollo de la producción en masa, el más visible exponente fue Henry Ford, proponiendo la producción en serie como una arquitectura dentro de la cual muchos sistemas están contruídos [14]. La Teoría Justo a Tiempo (JT), revoluciona y direcciona nuestro pensamiento acerca de las líneas de producción. En medio de otros pensamientos, el JT pone énfasis en la reducción de la producción en proceso y en tener buenos niveles de inventario, reduciendo costos, tiempo de ciclo corto, elevando la calidad, y ampliando el trabajo asignado a los trabajadores. El sistema de producción Toyota, fue el primer en poner en práctica el JT, véase por ejemplo [7].

Una de las reglas del JT es tener trabajadores adicionales, conocidos como trabajadores **flotantes** o **flexibles**. Esto se debe a que a diferencia del trabajador especializado (es decir, que sólo es experto en una sola actividad dentro del sistema de producción), el trabajador flotante conoce de todas las actividades dentro del sistema de producción. La inversión en trabajadores adicionales a los especializados, es una práctica que no solo ayuda a la organización a eficientar las operaciones, sino además se pueden obtener dividendos por la aplicación de estas prácticas. Definimos como estación de trabajo, el lugar donde solo se hace algún tipo de trabajo.

Un ejemplo de estos sistemas de producción son los de manufactura, y en particular los del tipo maquilador. Para ilustrar lo anterior considérese una maquiladora que fabrica calcetas y lo hace de la siguiente manera: se tiene un almacén de materia prima donde llegan calcetas sin coser las cuales pasan a un departamento, llamado

Costura, donde esperan hasta que una costurera, cose la punta de dicha calceta. De ahí pasa a otro departamento llamado Tintorería donde espera hasta que le toca su turno y se blanquea o se tiñe, dependiendo el tipo de calceta. Posteriormente pasa a otro departamento llamado Hormado donde se le da la forma del pie, y por último pasa al departamento de Empaque donde espera hasta que se empaqa y se envía hacia el contenedor. A estos trabajos que permanecen en espera se les conocen como trabajos en proceso (TP). Este ejemplo se ilustra en la figura de abajo.

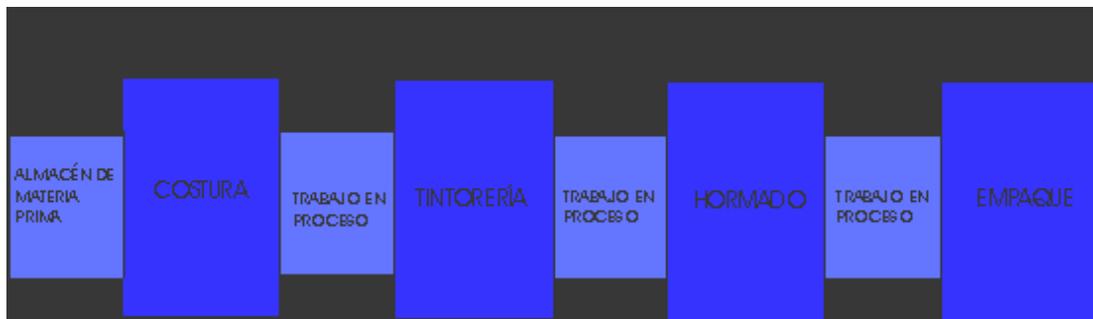


Figura 1

En lo anterior un trabajador especializado es aquel que solo permanece en alguno de los departamentos (Costura, Tintorería, Hormado o Empaque) y el trabajador flotante está capacitado para colaborar en cualquier departamento. [14] desarrolla un estudio sobre este tipo de sistemas para reducir el trabajo en proceso de una sola línea de producción. En este trabajo se hacen las siguientes consideraciones: (a) la no colaboración de los trabajadores en un mismo trabajo, (b) dos trabajadores pueden colaborar en un mismo lugar de trabajo, pero en diferentes actividades, (c) un trabajador flotante puede colaborar en cualquier trabajo, en cualquier estación de trabajo, mientras que los trabajadores especializados solo pueden colaborar en una estación de trabajo, (d) se supone que un trabajador solo atiende un trabajo a la vez, y (e) se considera el costo por el tiempo que permanece un trabajo en proceso.

Es conveniente mencionar que muchos sistemas tienen más de una línea de producción, (por ejemplo la empresa Acamex S. A de C. V., Tepeji del Río, Hgo.). Por lo que esta tesis, como principal objetivo, generaliza los resultados de [14] para diversas

líneas de producción, bajo las mismas consideraciones de este artículo .

El presente trabajo se organiza de la siguiente manera.

En el **capítulo 2** se revisan los antecedentes matemáticos como son: las cadenas de decisión de Markov en tiempo discreto y en tiempo continuo, el método de aproximación secuencial, y demás temas que serán herramientas principales para el desarrollo de nuestro problema.

El **capítulo 3** está dedicado al antecedente de nuestro trabajo, dado en [14]. Es decir, en [14] se trata el caso de una línea de producción con trabajadores especializados y un trabajador flotante. Este caso se modela como un proceso de decisión de Markov a tiempo continuo, en el cual se consideran ciertos costos de producción y se presentan asignaciones dinámicas óptimas para el trabajador flotante.

En el **capítulo 4** abordamos el problema de más de una línea de producción con trabajadores especializados y un trabajador flotante. Bajo los supuesto propuestos en [14], se plantea el problema y se modela como un proceso de decisión de Markov a tiempo continuo. Con las mismas consideraciones en los costos presentadas en [14], se dan asignaciones dinámicas óptimas para el trabajador flotante y damos también un algoritmo para llegar a estas asignaciones óptimas. Demostramos, además, bajo qué condiciones nuestro sistemas es estable o inestable.

El **capítulo 5** presenta una recapitulación de las aportaciones más importantes de nuestro trabajo y sugiere algunas de las líneas de investigación en las que se podría seguir trabajando en el futuro.

Finalmente, se presentan tres Apéndices. El Apéndice A trata de los temas básicos de procesos estocásticos. El Apéndice B presenta algunos detalles matemáticos que complementan las demostraciones de los resultados de la tesis. El Apéndice C aborda la equivalencia entre un proceso de decisión de Markov a tiempo discreto y uno a tiempo continuo. Además, para facilitar la lectura del trabajo, se proporciona una lista con los principales términos y la notación de éste.

# Capítulo 2

## Preliminares Matemáticos

La teoría de los sistemas de control estocástico, en tiempo discreto está motivada por sistemas que surgen en aplicaciones. Estamos interesados en el uso de resultados de esta teoría para el estudio de sistemas de control de colas en tiempo discreto. Un sistema de colas incluye servidores, clientes y usualmente líneas de espera ó colas para que los clientes esperen un servicio. En tiempo discreto, los servidores pueden ser: transmisiones, computadoras o líneas de comunicación, o ellos pueden ser estaciones, o una línea de producción. Los clientes pueden ser mensajes ó grupos de longitud fija de bits, conocidos como paquetes u objetos existentes en los procesos de manufactura. En este trabajo llamaremos a las colas buffers ó almacenes.

Un sistema de inventarios puede suponerse como un ejemplo de un sistema de colas. En un sistema de inventarios los clientes son los artículos en el inventario, y los servidores son demandas externas que remueven estos clientes desde el sistema.

A continuación mostramos un ejemplo común en los sistemas de colas (véase [13]).

### **Ejemplo 1** *Cola con un solo servidor.*

Supongase que paquetes (clientes) entran a cierto almacén. Mientras esperan su turno, son atendidos por un sólo servidor, y al acabar el servicio se marchan del sistema. El servicio es usualmente en el orden de arribo, esto es, primero en entrar, primero en salir (PEPS), pero pueden ser otras disciplinas como: último en entrar, primero en salir (UEPS), ó en orden aleatorio.

Ahora consideremos las opciones de control. Podemos poner un controlador en la entrada de la cola que decida que paquete admitir al buffer (almacén). También podríamos imponer un control en el servicio que ajuste el nivel en el cual los paquetes se atienden.

Ambos métodos de control podrían imponerse simultáneamente.

## 2.1. Cadena de Decisión de Markov

La estructura matemática que a continuación se presenta se conoce como cadena de decisión de Markov (también conocida como proceso de decisión de Markov o un programa dinámico estocástico), el cual nos permite analizar sistemas de control en tiempo discreto, involucrando aspectos aleatorios. El tiempo se particiona en longitudes iguales, llamadas periodos. A continuación damos la definición de Cadena de Decisión de Markov.

**Definición 1** Una *cadena de decisión de Markov (CDM)* consiste de: un espacio de estados, un espacio de acciones, los costos, y una distribución de probabilidad de transición (véase el apéndice A.3.2):

1. El estado del sistema es la información relevante necesaria para describir las condiciones actuales del sistema. *El espacio de estados* es el conjunto de todos los estados del sistema, el cual denotamos por  $S$  y puede ser infinito numerable ó finito.
2. Cuando el sistema está en el estado  $i \in S$ , el controlador tiene varias acciones disponibles. Estas acciones comprenden un conjunto finito  $A_i$  (no vacío). El espacio de acciones es el conjunto de la unión de todas las  $A_i$   $i = 1, 2, \dots, K$ , el cual denotamos por  $A$ .
3. Supongamos que el sistema se encuentra actualmente en el estado  $i$  y una acción  $a \in A_i$ , se elige por el controlador. Entonces se incurre en un costo (finito) no negativo  $C(i, a)$ .
4. Si el sistema está en el estado  $i$  y una acción  $a$  se elige, entonces el estado al comienzo del siguiente periodo es  $j$  con probabilidad  $P_{ij}(a)$ , donde

$\sum_{j \in S} P_{ij}(a) = 1$ .  $P_{ij}(a)$  se llama la *probabilidad de transición*, de ir del estado  $i$  al estado  $j$ , bajo la acción  $a$ .

### 2.1.1. Políticas

Informalmente una política es una regla de operación de una CDM. Sea  $t = 0$  el periodo inicial. La CDM puede operar en dos modos. En el modo de horizonte infinito, el sistema opera en periodos  $t = 0, 1, 2, \dots$ . En el modo de horizonte finito, el sistema opera en  $N$  periodos  $t = 0, 1, \dots, N - 1$ , donde  $N$  es entero positivo fijo.

**Definición 2** *La historia al tiempo  $t = 0$  está dada por  $h_0 = (i)$ ,  $i \in S$ . Se elige una acción inicial de  $A_i$  de acuerdo a la distribución  $\delta(a | i) = \delta(\cdot | h_0)$ ,  $a \in A_i$ . Esta es la distribución de probabilidad de la acción  $a \in A_i$ . Supongamos que  $a_0 \in A_i$  se selecciona. Entonces el estado del sistema en  $t = 1$  está determinado por la distribución de probabilidad de transición asociada con  $i$  y  $a_0$ . Supongamos que ese estado es  $j$ . La historia al tiempo  $t = 1$  está dada por  $h_1 = (i, a_0, j)$ . La acción al tiempo  $t = 1$  se elige de  $A_j$  de acuerdo a la distribución  $\delta(\cdot | i, a_0, j) = \delta(\cdot | h_1)$ . Una vez que esta acción se elige (digamos  $a_1$ ), entonces el estado del sistema en  $t = 2$  está determinado por la distribución asociada con  $j$  en  $a$ . Supongamos que este estado es  $k$ . La historia al tiempo  $t = 2$  está dada por  $h_2 = (i, a_0, j, a_1, k)$ . El proceso continúa de esta forma. Supongamos que el proceso estuvo operando para los periodos  $t = 0, 1, 2, \dots, n - 1$  y que el estado al tiempo  $t = n$  está determinado. La historia al tiempo  $t = n$  es  $h_n = (i, a_0, i_1, a_1, \dots, i_{n-1}, a_{n-1}, i_n)$  de los estados pasados y acciones en el estado actual. Entonces la acción en  $n$  se elige de acuerdo a la distribución de probabilidad  $\delta(\cdot | h_n)$  en el conjunto de acciones asociado con  $i_n$ . Una vez que esta acción se elige, el estado en el tiempo  $n + 1$  puede ser determinado. El proceso continúa de manera infinita.  $H_n$  es el conjunto de historias bajo  $n \in \mathbb{N}$ .*

**Definición 3** *Una política aleatorizada  $\delta$  es una sucesión  $\delta = (\delta_1, \delta_2, \dots, \delta_{N-1})$  donde  $\delta_n$ ,  $n = 1, \dots, N - 1$  especifica una distribución de probabilidad  $q_{\delta_n}(\cdot)$  en el conjunto de acciones. Para una política aleatorizada Markoviana, tal  $\delta_n$  mapea el conjunto de estados sobre el conjunto de distribuciones de probabilidad bajo el espacio de acciones, esto es  $\delta_t : S \rightarrow P(A)$ .*

**Definición 4** Una *política no aleatorizada*  $\delta$  es una sucesión de mapeos  $\delta_n$ ,  $n \in \mathbb{N}$  de  $H_n$  al conjunto de acciones  $A$  tal que  $\delta_n(h_i) \in A_i$  para todo  $i$ .

**Definición 5** Una *política de Markov*  $\delta$  se define por el mapeo  $\delta_n : S \rightarrow A$  tal que  $\delta_n(i) \in A_i$ , es decir el mapeo solamente depende del estado  $i$ .

**Definición 6** Una *política estacionaria*, denotada por  $\delta$  es un simple mapeo de  $\delta : S \rightarrow A$  tal que  $\delta(i) \in A_i$ . Para implementar esta política el controlador solo necesita conocer el estado actual del sistema. Los estados y las acciones pasadas son irrelevantes.

## 2.2. Distribución Condicional de Costos

Supongamos que  $i$  es el estado inicial y que el proceso opera bajo una política  $\delta$ . El estado al tiempo  $t$  es una variable aleatoria denotada por  $X_t$ . Similarmente la acción elegida al tiempo  $t$  es una variable aleatoria, denotada por  $A_t$ . La distribución de probabilidad conjunta de  $(X_t, A_t)$  está dada por:

$$P_\delta(X_t = j, A_t = a \mid X_0 = i), \text{ donde podemos tener que } a \in A_j.$$

En este caso la distribución de probabilidad está bien definida. No demostraremos esto, pero mostraremos el cálculo para  $t = 0, 1, 2$ . Esto será suficiente para indicar las ideas operativas.

$$\text{Sea } P_\delta(A_0 = a \mid X_0 = i) = \delta(a \mid i).$$

Para  $t = 1$ , tenemos:

$$P_\delta(X_1 = j, A_1 = a \mid X_0 = i) = \sum_{b \in A_i} \delta(b \mid i) P_{ij}(b) \delta(a \mid i, b, j).$$

El último término es la probabilidad de elegir la acción  $b$ , transitando a  $j$ , y eligiendo la acción  $a$ ; los términos se suman sobre las acciones  $b \in A_i$ .

Para  $t = 2$ , tenemos:

$$\begin{aligned} P_\delta(X_2 = j, A_2 = a \mid X_0 = i) \\ = \sum_{b \in A_i} \delta(b \mid i) \left( \sum_k P_{ik}(b) \sum_{d \in A_k} \delta(d \mid i, b, k) P_{kj}(d) \delta(a \mid i, b, k, d, j) \right). \end{aligned}$$

Asociamos con el par aleatorio  $(X_t, A_t)$  el costo  $C(X_t, A_t)$ . El costo incurrido al tiempo  $t$  cuando un controlador opera bajo  $\delta$ . Dado que  $C(X_t, A_t)$  es una variable aleatoria. Un camino efectivo para calcular este es empleando la esperanza dada por:

$$E_\delta [C(X_t, A_t) | X_0 = i] = \sum_j \sum_{a \in A_j} C(j, a) P_\delta(X_t = j, A_t = a | X_0 = i), \quad (2.2)$$

y representa el *costo esperado* al tiempo  $t$ . Dado que los costos son no negativos, el valor esperado (2.2) está bien definido. Puede suceder que este valor sea  $+\infty$ .

Consideremos una situación importante, cuando  $\delta$  es una política estacionaria  $f$ . En este caso emplearemos una notación especial, El costo asociado en el estado  $i$  está denotado por  $C(i, f)$ , donde  $f = f(i)$ . Similarmente la probabilidad de transición será denotada por  $P_{ij}(f)$  (véase el apéndice A.3.2).

Entonces  $P_f(X_t = j, A_t = a | X_0 = i)$  es cero al menos que  $a = f(j)$ , y tenemos:

$$\begin{aligned} P_f(X_t = j, A_t = f(j) | X_0 = i) &= P_f(X_t = j | X_0 = i) \\ &= \sum_{k_1 \in S} P_{ik_1}(f) \sum_{k_2 \in S} P_{k_1 k_2}(f) \dots \sum_{k_{t-1} \in S} P_{k_{t-1} j}(f) \\ &: = P_{ij}^{(t)}(f). \end{aligned}$$

Entonces de (2.2) se tiene:

$$\begin{aligned} E_f [C(X_t, A_t) | X_0 = i] &= \sum_j C(j, f) P_f(X_t = j | X_0 = i) \\ &= \sum_j C(j, f) P_{ij}^{(t)}(f). \end{aligned}$$

Este es el costo esperado al tiempo  $t$ , bajo la política estacionaria  $f$ .

## 2.3. Criterios de Optimización

En este trabajo consideramos dos criterios de optimización. El criterio de costo esperado en horizonte finito y el costo promedio esperado a largo plazo. Ambos criterios se basan en el costo promedio  $E_\delta [C(X_t, A_t)]$ , al tiempo  $t$ , como se define en (2.2).

Supongamos que el proceso opera bajo un horizonte finito  $n$  y que existe un costo terminal no negativo  $F(k)$  en el que se incurre cuando el proceso está en el estado  $k$ . Sean  $i$  un estado inicial,  $n$  el horizonte y  $\delta$  una política dada. El costo esperado en un horizonte  $n$  bajo la política  $\delta$  se define como:

$$v_{\delta,n}(i) = E_{\delta} \left[ \sum_{t=0}^{n-1} C(X_t, A_t) + F(X_n) \mid X_0 = i \right] \quad (2.3)$$

$$= \sum_{t=0}^{n-1} E_{\delta} [C(X_t, A_t) \mid X_0 = i] + E_{\delta} [F(X_n) \mid X_0 = i]. \quad (2.4)$$

La última igualdad se sigue por la linealidad del valor esperado.

**Definición 7** La función de valores óptimos en horizonte finito  $n$ , se define como:

$$v_n(i) = \inf_{\delta} v_{\delta,n}(i) \quad (2.5)$$

Ahora damos la definición de política óptima bajo este criterio.

**Definición 8** Sea  $\delta$  una política para el criterio de optimización en horizonte finito. Entonces  $\delta$  es **óptima** para el criterio de costo esperado para el horizonte  $n$  si  $v_{\delta,n}(i) = v_n(i)$  para  $i \in S$ .

Dados un estado inicial  $i$  y  $\delta$  una política en horizonte infinito, el **costo esperado promedio** a largo plazo se define como:

$$J_{\delta}(i) = \limsup_{n \rightarrow \infty} \frac{1}{n} E_{\delta} \left[ \sum_{t=0}^{n-1} C(X_t, A_t) \right] \quad (2.6)$$

$$= \limsup_{n \rightarrow \infty} \frac{v_{\delta,n}(i)}{n}. \quad (2.7)$$

Ahora definimos la función de valores óptimos para el costo esperado promedio a largo plazo (costo promedio para acertar) como:

$$J(i) = \inf_{\delta} J_{\delta}(i), \quad i \in S. \quad (2.8)$$

donde el ínfimo se toma sobre todas las políticas para el horizonte infinito. En seguida se presenta la definición de una política óptima bajo este criterio.

**Definición 9** Sea  $\delta$  una política para el horizonte infinito. Entonces  $\delta$  es **óptima** para el criterio de optimización costo promedio si  $J_{\delta}(i) = J(i)$  para  $i \in S$ .

## 2.4. Método de Aproximación Secuencial

Supongamos que  $\Delta$  es una cadena de decisión de Markov (a tiempo discreto) con espacio de estados  $S$  numerable. Definiremos una sucesión  $\{\Delta_N\}_{N \geq N_0}$  de cadenas de decisión de Markov que aproxime  $\Delta$ .

El espacio de estados en cada  $\Delta_N$  es finito y se computa una política óptima para  $\Delta_N$ . Bajo ciertas condiciones los resultados de estos cálculos convergerán a la política óptima para  $\Delta$  (véase sección 2.5 de [13]).

**Definición 10** Sea  $N_0$  un entero no negativo. La sucesión  $\{\Delta_N\}_{N \geq N_0}$  es una aproximación secuencial (AS) para  $\Delta$  si existe una sucesión creciente  $(S_N)_{N \geq N_0}$  de subconjuntos finitos no vacíos de  $S$  tal que  $\bigcup_{N=0}^{\infty} S_N = S$ . Cada  $\Delta_N$  es una CDM con espacio de estados  $S_N$  satisfaciendo dos condiciones:

1. Para cada  $i \in S_N$  se tiene un conjunto de acciones  $A_i$  y para cada  $a \in A_i$  un costo  $C(i, a)$ .
2. Para cada  $i \in S_N$  y  $a \in A_i$ ,  $P_{ij}(a; N)$  es la distribución de probabilidad en  $S_N$  tal que:

$$\lim_{N \rightarrow \infty} P_{ij}(a; N) = P_{ij}(a), \quad j \in S. \quad (2.9)$$

Si nos referimos al caso de horizonte finito con un costo terminal  $F$ , entonces este mismo costo terminal se aplica a  $\Delta_N$ . Al entero  $N$ , se le llama *nivel de aproximación*.

La CDM  $\Delta_N$  tiene como espacio de estados un subconjunto de  $S$ . En este subconjunto finito, el conjunto de acciones y los costos para  $\Delta_N$  son exactamente los mismos para  $\Delta$ . Solamente las probabilidades de transición son diferentes. De esta forma las distribuciones en un subconjunto finito convergen a la distribución original en  $\Delta$ . Las distribuciones en la definición se llaman *distribuciones de aproximación*.

Así entonces solo se requieren de dos condiciones para especificar un AS; el subconjunto finito y la distribución de aproximación.

El camino más importante para definir la distribución de aproximación es por medio de un *procedimiento de aumentación*.

Este procedimiento se usa cuando el espacio de estados es multidimensional. La idea es la siguiente:

Supongamos que el proceso está en el estado  $i \in S_N$  y se elige la acción  $a$ . Para  $j \in S_N$  la probabilidad  $P_{ij}(a)$  no se cambia. Ahora supongamos que  $P_{ir} > 0$  para algún  $r \notin S_N$ . Esto es, bajo esta probabilidad, el proceso original transitará al estado  $r$  fuera de  $S_N$ . A esto le llamamos *excesos de probabilidad* asociados con  $(i, a, r, N)$ . Estos excesos de probabilidad se redistribuyen (es decir se envían) a los estados de  $S_N$  de acuerdo a la distribución especificada. Generalmente esta distribución puede depender de  $i, a, r$  y  $N$ ; esta se llama *distribución de aumentación asociada* con  $(i, a, r, N)$ . Se requiere además definir  $P_{ir}(a) = 0$ . A continuación se da la definición formal de un procedimiento de aumentación.

**Definición 11** *La sucesión de aproximación  $\{\Delta_N\}_{N \geq N_0}$  es una aproximación secuencial del tipo de aumentación (ASTA) si la distribución de aproximación cumple lo siguiente:*

Dado  $i \in S_N$  y  $a \in A_i$ , para cada  $r \notin S_N$  existe una distribución de probabilidad  $(q_j(i, a, r, N))_{j \in S_N}$ , llamada la distribución de aumentación asociada con  $(i, a, r, N)$ , tal que:

$$P_{ij}(a; N) = P_{ij}(a) + \sum_{r \in S - S_N} P_{ir}(a) q_j(i, a, r, N), j \in S_N. \quad (2.10)$$

**Proposición 1** *La ecuación (2.10) define una distribución de probabilidad de aproximación en  $S_N$ .*

**Demostración.** Para mostrar que (2.10) define una distribución de probabilidad, note que:

$$\begin{aligned} \sum_{j \in S_N} P_{ij}(a; N) &= \sum_{j \in S_N} P_{ij}(a) + \sum_{j \in S_N} \sum_{r \notin S_N} P_{ir}(a) q_j(i, a, r, N) \\ &= \sum_{j \in S_N} P_{ij}(a) + \sum_{r \notin S_N} P_{ir}(a) \left( \sum_{j \in S_N} q_j(i, a, r, N) \right) \\ &= \sum_{j \in S_N} P_{ij}(a) + \sum_{r \notin S_N} P_{ir}(a) \\ &= \sum_{j \in S} P_{ij}(a) = 1. \end{aligned}$$

El intercambio en la sumatoria en la segunda línea es válida dado que todos los términos son no negativos. La tercera línea se sigue dado que las probabilidades en una distribución de aumentación suman 1. Ahora mostraremos que la distribución en (2.10) satisface (2.9). Primero observemos que:

$$1 = \sum_{j \in S_N} P_{ij}(a) + \sum_{r \notin S_N} P_{ir}(a). \quad (2.11)$$

Dado que el conjunto  $S_N$  crece a  $S$ , este es el caso en el que el primer término en el lado derecho de (2.11) se aproxima a 1 cuando  $N \rightarrow \infty$ . De donde tenemos:

$$\lim_{N \rightarrow \infty} \sum_{r \notin S_N} P_{ir}(a) = 0. \quad (2.12)$$

Ahora fijemos  $j \in S$ , y asumimos que  $N$  es más grande que  $j \in S_N$ . Dado que los términos en (2.10) son no negativos y  $q_j \leq 1$  se sigue que:

$$P_{ij}(a) \leq P_{ij}(a; N) \leq P_{ij}(a) + \sum_{r \notin S_N} P_{ir}(a), \quad (2.13)$$

tomando límite en (2.13) cuando  $N \rightarrow \infty$ , el resultado se sigue de (2.12). ■

## 2.5. Cadena de Decisión de Markov a Tiempo Continuo

En esta sección discutiremos una estructura matemática, llamada Cadena de Decisión de Markov a Tiempo Continuo (CDMTC) que se usa para modelar el control de ciertos sistemas que ocurren a tiempo continuo.

La CDMTC denotada por  $\Psi$ , tiene espacio de estados  $S$ , que es un conjunto numerable. Asociada con cada  $i \in S$  existe un conjunto finito no vacío  $A_i$  de acciones disponible en  $i$ . Asumimos que la acción  $a \in A_i$  se elige. Entonces se incurre en un costo. Este consiste de un costo instantáneo  $G(i, a)$  incurrido inmediatamente y un parámetro de costo  $g(i, a)$ , tal efecto ocurre hasta la siguiente transición.

El tiempo hasta la siguiente transición está distribuido exponencialmente (véase el apéndice A.2) con parámetro  $v(i, a)$ . El nuevo estado de acuerdo a la distribución de probabilidad  $(P_{ij}(a))_{j \in S}$ . La teoría puede ser desarrollada asignando  $P_{ii}(a) > 0$ .

Sin embargo para muchos sistemas se tiene que  $P_{ii}(a) = 0$ , y nosotros suponemos esto último.

### 2.5.1. Optimización Costo Promedio de una CDMTC

En esta sección daremos un conjunto de suposiciones que permiten computar una política estacionaria costo promedio óptima para  $\Psi$ .

De manera intuitiva definimos una política  $\delta$  como una regla no anticipatoria para la elección de acciones.

Esto puede depender de la historia de los procesos a través del estado presente y puede aleatorizar entre las acciones. La historia incluye el estado pasado de los procesos, las acciones elegidas en esos estados, y el tiempo esperado en esos estados.

Una política estacionaria se define como en la sección 2.1.1.

Existen dos definiciones comunes de los costos promedio bajo una política arbitraria. La primera definición considera el costo esperado incurrido bajo la política durante el intervalo  $[0, t)$ , dividiendo sobre  $t$  y tomando límite superior de esta cantidad cuando  $t \rightarrow \infty$ . Nosotros emplearemos la segunda definición, la cual considera el costo esperado incurrido durante  $n$  transiciones, dividido por el tiempo esperado de estas transiciones, y entonces tomamos el límite superior cuando  $n \rightarrow \infty$ . Formalmente sea  $E_\delta[C_n]$  el costo total esperado incurrido bajo  $\delta$  durante los primeros  $n$  periodos de transición. Sea  $E_\delta[T_n]$  el tiempo total esperado tomado bajo  $\delta$  para los  $n$  primeros periodos de transición. Entonces definimos:

$$J_\delta^\Psi(i) = \limsup_{n \rightarrow \infty} \frac{E_\delta[C_n | x_0 = i]}{E_\delta[T_n | x_0 = i]}, \text{ y} \quad (2.14)$$

$$J^\Psi(i) = \inf_\delta J_\delta^\Psi(i), \quad i \in S. \quad (2.15)$$

La igualdad (2.14) está bien definida, lo cual se puede ver en la sección 2.2.

Estamos interesados en las condiciones bajo las cuales  $J^\Psi(i)$  es idénticamente igual a una constante (finita)  $J^\Psi$  y la existencia de una política estacionaria costo promedio óptima que sea computable.

Note que si el proceso está en el estado  $i \in S$  y una acción  $a \in A_i$  se elige, entonces el tiempo esperado hasta un cambio del estado está dado por  $\tau(i, a) = \frac{1}{v(i, a)}$ . A continuación se tiene la siguiente hipótesis:

**Hipótesis 1** Supongamos que el proceso se encuentra en el estado  $i \in S$  y una acción  $a \in A_i$  se elige, entonces, existen constantes  $B$  y  $\tau$  tal que:

$$0 < \tau < \inf_{i,a} \tau(i, a) \leq \sup_{i,a} \tau(i, a) \leq B < \infty. \quad (2.16)$$

**Lema 1** Sea  $\Psi$  una CDMTC satisfaciendo la hipótesis 1, y sea  $e$  una política estacionaria para  $\Psi$ . Supongamos que existe una constante  $Z$  y una función  $z$  acotada por abajo en  $i$  tal que:

$$Z\tau(i, e) + z(i) \geq G(i, e) + g(i, e)\tau(i, e) + \sum_j P_{ij}(e)z(j), \quad i \in S. \quad (2.17)$$

Entonces  $J_e^\Psi(i) \leq Z$  para  $i \in S$ .

**Demostración.** Para evitar confusión con el tiempo continuo, emplearemos  $k$  para el índice en tiempo discreto. Probaremos por inducción que:

$$E_e [z(X_k)] \leq ZB_k + z(i) \text{ para } k \geq 0.$$

De la ecuación (7.6) de [13] se tiene:

$$E_e(G(X_k, e) + g(X_k, e)\tau(X_k, e)) \leq ZE_e [\tau(X_k, e)] + E_e [z(X_k)] - E_e [z(X_{k+1})] \quad k \geq 0. \quad (2.18)$$

Sumando los términos en (2.18) para  $k = 0, 1, 2, \dots, n-1$  y dividiendo por la suma de los tiempos de transición esperado obtenemos:

$$\frac{\sum_{k=0}^{n-1} E_e [G(X_k, e) + g(X_k, e)\tau(X_k, e)]}{\sum_{k=0}^{n-1} E_e [\tau(X_k, e)]} \leq Z + \frac{z(i) - E_e [z(X_n)]}{\sum_{k=0}^{n-1} E_e [\tau(X_k, e)]} \leq Z + \frac{z(i) + L}{n\tau}, \quad (2.19)$$

donde  $-L$  es una cota finita por abajo para  $z$  y  $\tau$  es como en la hipótesis 1. Ahora tomando el límite superior en ambos lados de (2.19). El límite superior de el lado izquierdo de (2.19) es  $J_e^\Psi(i)$ , y se tiene el resultado. ■

El plan ahora es introducir una Cadena de Decisión de Markov auxiliar (a tiempo discreto)  $\Delta$  que conecte estrechamente con el proceso a tiempo continuo  $\Psi$ .

Podemos, entonces, formar una sucesión de aproximación  $\{\Delta_N\}_{N \geq N_0}$  para  $\Delta$  y usar métodos computacionales introducidos en el capítulo 8 de [13] para computar una política estacionaria costo promedio óptima para  $\Delta$ . Bajo ciertas suposiciones esta política también es óptima para  $\Psi$  (véase [15]). El desarrollo puede ser representado esquemáticamente como:

$$\Psi \implies \Delta \implies \{\Delta_N\}_{N \geq N_0},$$

donde  $\Psi$  es el proceso original a tiempo continuo con espacio de estados infinito,  $\Delta$  es el proceso auxiliar a tiempo discreto con espacio de estados infinito, y  $\{\Delta_N\}_{N \geq N_0}$  es la sucesión de aproximación para  $\Delta$  consistiendo de un proceso con espacio de estados finito, para el cual se pueden realizar cálculos.

Ahora definimos la CDM (en tiempo discreto)  $\Delta$ . Los estados y las acciones son los mismos que  $\Psi$ . Sean  $i, j \in S$  y  $a \in A_i$ , las probabilidades de transición (véase el apéndice A.3.2) y los costos están dados por:

$$P_{ij}^*(a) = \begin{cases} \tau v(i, a) P_{ij}(a), & j \neq i, \\ 1 - v(i, a), & j = i, \end{cases} \quad (2.20a)$$

$$C(i, a) = G(i, a)v(i, a) + g(i, a). \quad (2.20b)$$

Note de (2.20a) que  $\tau v(i, a) = \tau/\tau(i, a) \leq 1$ . Si el proceso está en el estado  $i$ , entonces en cada periodo la probabilidad de transitar a  $j \neq i$  es proporcional a la probabilidad en  $\Psi$ . Además la probabilidad de permanecer en  $i$  es cero. Esto puede ser contrastante a  $\Psi$  para el cual  $P_{ii}(a) = 0$ .

Observe que el conjunto de políticas para  $\Delta$  y para  $\Psi$  no son idénticas. Una política para  $\Psi$  solo puede seleccionar una nueva acción cuando una transición ocurre en el estado. Una política para  $\Delta$  puede elegir una nueva acción en cada periodo de tiempo, incluso si permanece en el mismo estado. Además, es fácil ver que el conjunto de políticas estacionarias es idéntico. Para esto es crucial el siguiente lema.

**Lema 2** *Sea  $\Psi$  una CDMTC satisfaciendo la hipótesis 1, sea  $\Delta$  una CDM auxiliar, y sea  $e$  una política estacionaria.*

i) Supongamos que existe una constante (finita)  $Z$  y una función (finita)  $w$  tal que:

$$Z + w(i) \geq C(i, e) + \sum_j P_{ij}^*(e)w(j), \quad i \in S. \quad (2.21)$$

Entonces  $Z$  y  $z = \tau w$  satisfacen (2.17).

ii) Supongamos que existe una constante finita  $Z$  y  $w = z/\tau$  que satisfacen (2.17). Entonces  $Z$  y  $w = z/\tau$  satisfacen (2.21).

**Demostración.** i) Supongamos que (2.21) se cumple. Sustituyendo  $z/\tau$  por  $w$  en (2.21) y usando (2.20a) se tiene (2.17), después de algunas manipulaciones algebraicas. La demostración de (ii) es similar. ■

Ahora daremos una hipótesis que conecta el costo promedio mínimo en  $\Psi$  y en  $\Delta$ .

**Hipótesis 2** *Se tiene que  $J^\Delta(\cdot) \leq J^\Psi(\cdot)$ , donde  $J^\Delta(\cdot)$  es el mínimo costo promedio en  $\Delta$ .*

Supongamos que se tiene una sucesión de aproximación  $\{\Delta_N\}_{N \geq N_0}$  para  $\Delta$ . El siguiente resultado es análogo al Teorema 8.1.1 en [13] y concede el computo de una política estacionaria costo promedio óptima para  $\Psi$ .

**Teorema 1** *Sea  $\Psi$  una CDMTC satisfaciendo la hipótesis 1, sea  $\Delta$  una CDM auxiliar tal que la hipótesis 2 se tiene, y sea  $\{\Delta_N\}_{N \geq N_0}$  una sucesión de aproximación para  $\Delta$  satisfaciendo las suposiciones (AC) de la sección 8.1 de [13]. Note que de 8.1 de [13] se tiene:*

$$J^N + r^N(i) = \min_a \left\{ C(i, a) + \sum_{j \in S_n} P_{ij}^*(a; N)r^N(j) \right\}, \quad i \in S_N, \quad N \geq N_0. \quad (2.22)$$

Entonces:

- i) La cantidad  $J^* = \lim_{N \rightarrow \infty} J^N$  es el costo promedio mínimo en  $\Delta$  y  $\Psi$ .
- ii) Cualquier punto límite  $e^*$  de una sucesión  $\{e^N\}_{N=1}^\infty$  de políticas estacionarias realizando el mínimo en (2.22), es costo promedio óptimo para  $\Delta$  y  $\Psi$ .

**Demostración.** Dadas las suposiciones (AC) de la sección 8.1 de [13], para  $\{\Delta_N\}_{N \geq N_0}$  y  $\Delta$ , la demostración y conclusiones del Teorema 8.1.1 de [13] se tienen

para  $\{\Delta_N\}_{N \geq N_0}$  y  $\Delta$ . De tal resultado concluimos que la cantidad  $J^*$  es el mínimo costo promedio en  $\Delta$ ,  $e^*$  es costo promedio óptimo en  $\Delta$ , de la ecuación 8.3 en [13] se sigue que:

$$J^* + w(i) \geq c(i, e^*) + \sum_j P_{ij}^*(e^*)w(j), \quad j \in S. \quad (2.23)$$

Se sigue del lema 1.2 que  $Z = J^*$  y  $z = \tau\omega$  satisface (2.17). Además de la hipótesis (AC3) de la sección 8.1 de [13] se tiene que  $\omega$  (y por lo tanto  $z$ ) es acotada por abajo en  $i$ . El lema 1.1 implica que

$$J_{e^*}^*(\cdot) \leq J^* \equiv J^\Delta(\cdot) \leq J^\Psi(\cdot) \leq J_{e^*}^\Psi(\cdot),$$

donde la segunda desigualdad se sigue de la hipótesis 2. Por lo tanto esos términos son iguales. Esto prueba que  $e^*$  es costo promedio óptimo para  $\Psi$  con costo promedio constante  $J^*$ . ■

**Comentario 1** *Si la hipótesis 1 se cumple, y sea  $e$  una política estacionaria. Entonces  $e$  induce una CM en  $\Delta$ . Similarmente  $e$  induce lo que es conocido como Cadena de Markov a Tiempo Continuo (CMTC) en  $\Psi$ . Este método usa algunos resultados de la teoría de costo promedio para la cadena de Markov en tiempo continuo. Daremos la idea pero omitimos los antecedentes de este material (véase pág. 247 de [13]).*

**Idea de la Demostración.** Las clases comunicantes (véase el apéndice A.3.2) de CM y CMTC son la misma. Sea  $R$  una clase. Entonces  $R$  es recurrente positiva (véase el apéndice A.3.1 y A.3.2) en la CM si y sólo si esta es recurrente positiva en la CMTC. De esto se puede demostrar que:

$$\pi_i^\Delta(e) = \frac{\tau(i, e)\pi_i^\Psi(e)}{\sum_{j \in R} \tau(j, e)\pi_j^\Psi(e)}, \quad i \in R, \quad (2.24)$$

donde el denominador del lado derecho de (2.24) es una constante normalizada, la cual se denota por  $\gamma$ . Usando esto, podemos demostrar que el costo promedio en  $R$

es el mismo para CM y CMTC. Esto es:

$$\begin{aligned}
 J_{e,R}^{\Delta} &= \sum_{i \in R} C(i, e) \pi_i^{\Delta}(e) \\
 &= \sum_{i \in R} [G(i, e)v(i, e) + g(i, e)] \pi_i^{\Delta}(e) \\
 &= \frac{1}{\gamma} \sum_{i \in R} [G(i, e) + g(i, e)\tau(i, e)] \pi_i^{\Psi}(e) \\
 &= J_{e,R}^{\Psi}.
 \end{aligned}$$

La segunda línea se sigue de (2.20a), y la tercera línea se sigue de (2.24). La última línea se sigue de la teoría de CMTC. Y suponiendo que  $J^{\Delta}(\cdot)$  es una constante  $J^{\Delta}$  y  $J^{\Psi}(\cdot)$  es una constante  $J^{\Psi}$ . Además supongamos que existe una política estacionaria costo promedio óptima  $f$  para  $\Psi$  induciendo una CMTC con una clase recurrente  $R$  (véase el apéndice A.3.1),

$$J^{\Delta} \leq J_{f,R}^{\Delta} = J_{f,R}^{\Psi} = J^{\Psi}. \quad (2.25)$$

Por lo tanto, la hipótesis 2 se cumple. ■

## Capítulo 3

# Una Línea de Producción con $K$ Trabajadores Especializados y un Trabajador Flotante

A continuación mostramos una serie de resultados debido a [14], donde se modela y analiza una línea de producción en serie con especialistas en cada estación y un trabajador flotante (TF) que puede trabajar en cualquier estación. Se formula un Modelo de Decisión de Markov de  $K$ -estaciones de producción en el cual: (1) los trabajadores no colaboran en el mismo trabajo, y (2) dos trabajadores pueden colaborar en el mismo trabajo en alguna estación de trabajo, pero en diferentes tareas al mismo tiempo. Este modelo considera los costos, por el tiempo de permanecer en cada estación. Se propone una política costo promedio óptima para determinar las asignaciones del trabajador flotante.

### 3.1. Descripción del Modelo

Este modelo presenta una línea de producción en serie con  $K + 1$  trabajadores y  $K$  estaciones etiquetadas como  $1, 2, \dots, K$ , cada estación tiene un buffer infinito de trabajos. Las estaciones pueden arreglarse en una línea. En cada estación hay dos estaciones de trabajo, una que es atendida por un trabajador especializado y la otra está disponible para el TF cuando se necesite. El tiempo para completar un

trabajo en una estación de trabajo en la estación  $s$  ( $s = 1, 2, \dots, K$ ) está distribuida exponencialmente (véase el apéndice A.2) con parámetro  $\mu_s$  considerando si el especialista o el TF está trabajando ahí. Este modelo es un sistema de producción abierto, con arribos de acuerdo a un proceso de Poisson con parámetro  $\lambda$  (véase el apéndice A.2).

Si el TF se mueve a la estación  $s$ , él puede comenzar a servir un segundo trabajo en la estación  $s$ , siempre y cuando haya un segundo trabajo en tal estación. El trabajador TF no colabora en un mismo trabajo con un especializado, si hay solo un trabajo en la estación, el TF no trabajará en éste. El tiempo de servicio de un segundo trabajo en la estación  $s$  está distribuido exponencialmente con parámetro  $\mu_s$ .

### 3.2. La Formulación del Modelo como PDMTC

A continuación se construye el modelo como un Proceso de Decisión de Markov en Tiempo Continuo (PDMTC) denotado como  $\Psi$ . Para el PDMTC, tenemos:

1. El *Espacio de Estados*. El Espacio de estados está dado por

$$S = \{(i_1, i_2, \dots, i_K) \mid i_s \in \mathbb{N}, s = 1, 2, \dots, K\},$$

donde  $i_s$  denota el número de trabajos en la estación  $s$ ,  $s = 1, 2, \dots, K$ .

2. *Épocas de Decisión*. Las épocas de decisión suceden cuando justamente un trabajo entra al sistema (siempre entra al buffer 1) o un trabajo se ha completado y si esto ocurre en la estación  $s$  entonces tal trabajo será sustraído del buffer  $s$  y se agregará al buffer  $s + 1$ ,  $s = 1, 2, \dots, K$ .
3. *Espacio de Acciones*: El espacio de acciones está dado por  $A = \{1, 2, \dots, K\}$ , donde elegir  $a \in A$ , significa que el TF se moverá (instantáneamente) a la estación  $a \in A$  (o permanecerá en ella, si él está ahí).
4.  $H_s > 0$ , será el parámetro de costo por cada unidad de tiempo que el trabajo este en la estación  $s$ ,  $s = 1, 2, \dots, K$ . Dado un estado  $i$  y una acción  $a$  se elige

entonces se incurre en un costo dado por:  $g(i, a) = \sum_s H_s i_s = h(i)$ . Estos costos no son instantáneos.

5. El tiempo de transición de un estado a otro está distribuido exponencialmente con parámetro:

$$v(i, a) = \lambda + \sum \mu_s I(i_s \geq 1) + \mu_a I(i_a \geq 2), \quad (3.1)$$

donde  $I(E)$  es la función indicadora para el evento  $E$ . El primer término del lado derecho de (3.1) es el parámetro de arribo. El segundo término de (3.1) es el parámetro de servicio de los trabajadores especializados y el último término es el parámetro de servicio del TF.

Para propósitos computacionales transformamos  $\Psi$  bajo un Proceso de Decisión de Markov (PDM),  $\Delta$  usando uniformización (véase sección 2.5).

El espacio de estados para  $\Delta$  es el mismo que para  $\Psi$ . Ahora elegimos un  $\tau$  fijo, el periodo de transición del sistema uniformizado (en unidades de tiempo por transición), satisfaciendo:

$$\tau = (\lambda + \sum \mu_s + \text{máx} \{\mu_s\})^{-1}.$$

Si el sistema está en el estado  $i$  al inicio de un periodo y se toma una decisión  $a$  al comienzo del siguiente periodo, el sistema tendrá una transición a  $i + e_1$  (donde  $e_s$  es el vector cuya componente  $s$  es 1 y las demás son cero,  $s = 1, 2, \dots, K$ ) con probabilidad  $\tau\lambda$ , tendrá una transición a  $i - e_s + e_{s+1}$  con probabilidad  $\tau\mu_s [I(i_s \geq 1) + I(a = s, i_s \geq 2)]$ , y de (3.1) permanecerá en  $i$  con probabilidad  $1 - \tau v(i, a)$ . Definimos  $e_{k+1}$ , como el vector que controla en caso de completar un trabajo en la última estación. El costo en que se incurre al comenzar un periodo de tiempo es  $C(i, a) = g(i)$ .

El problema clave aquí es que  $\Delta$  tiene un buffer infinito, y podemos buscar un camino para implementar cálculos rigurosos con solamente un espacio de estados finito. Empleamos el método de aproximación secuencial (AS) introducido en la sección 1.3 al reemplazar  $\Delta$  con una sucesión  $\{\Delta_N\}_{N \geq N_0}$  de espacio de estados finito PDM en el cual se hacen cálculos.

El espacio de estados de  $\Delta_N$  consiste de vectores  $i = (i_1, i_2, \dots, i_K)$  tal que  $i_s \leq N$   $s = 1, 2, \dots, K$  ( el máximo tamaño del buffer es de  $N$  trabajos). El método AS no es un simple esquema de truncamiento. Se distribuyen los ‘excesos de probabilidad’ que causa que el sistema salga fuera de la cota. El sistema puede solamente escapar a  $\Delta_N$  desde un estado acotado para el cual algún buffer está lleno. Si  $i_1 = N$ , un sobreflujo ocurrirá en el buffer 1 si un nuevo trabajo arriba, el cual ocurre con probabilidad  $\tau\lambda$ . Este exceso es asignado al estado  $i$ , efectivamente se le niega la entrada a un nuevo trabajo en la estación 1 cuando el buffer está en el nivel  $N$ , hasta que se completa un trabajo. Similarmente, si  $i_s = N$ , para algún  $s$ ,  $2 \leq s \leq K$ , entonces con probabilidad  $\tau\mu_{s-1} [I(i_{s-1} \geq 1) + I(a = s - 1, i_{s-1} \geq 2)]$  el siguiente estado es  $j$ , donde  $j_{s-1} = i_{s-1} - 1$ , y las demás coordenadas de  $j$  son iguales a las de  $i$ . El efecto de esta asignación es remover los trabajos completados en  $s - 1$  desde el buffer previo y descartar éstos. El costo en  $\Delta_N$  es el mismo que  $\Delta$ .

En resumen se reduce el PDMTC  $\Psi$  al asociado espacio de estados infinito en tiempo discreto PDM  $\Delta$ , el cual aproximamos por una sucesión de espacio de estados finito  $\{\Delta_N\}_{N \geq N_0}$  para propósitos de cómputo.

### 3.3. Optimalidad, Estabilidad y Acotamiento

El objetivo es determinar una política dinámica que especifique el mínimo costo promedio esperado por unidad de tiempo. Supongamos que  $\Psi$  está controlada bajo una política  $\delta$ . Sea  $F(t)$  el costo total incurrido durante un intervalo  $[0, t)$ . Para el estado inicial  $i$ , el costo promedio bajo  $\delta$  y el costo promedio óptimo están definidas respectivamente como:

$$J_\delta^{(1)}(i) = \limsup_{t \rightarrow \infty} E_\delta^{(i)} \left[ \frac{F(t)}{t} \right]; \quad J^{(1)}(i) = \inf_\delta J_\delta^{(1)}(i). \quad (3.2)$$

Una política estacionaria (como opuesta a una política aleatorizada o política variando en el tiempo) toma la misma decisión en el estado  $i$ , recurrente en cualquier estado de tiempo. Esto puede mostrar que nuestro modelo es ergódico bajo nuestras condiciones de estabilidad (véase el comentario 3 abajo). En resumen buscamos una constante  $J^{(1)}$  y una política estacionaria  $f$  tal que:

$$J^{(1)} \equiv J_f^{(1)} = J^{(1)}(i) \quad \forall i. \quad (3.3)$$

**Comentario 2** Una CM se llama **Estable** si es ergódica. Y se llama **Inestable** si es débilmente ergódica.

En el caso especial en que  $H_s \equiv 1$ , entonces  $J^{(1)}$  es igual al mínimo número promedio de trabajos en la línea, y una política óptima garantiza el promedio mínimo de trabajos en proceso. Se tienen las siguientes definiciones:  $r_s = \lambda/\mu_s$ , el cual mide la carga ofertada en la estación  $s$ , (véase el apéndice A.4.1) sea  $q_k = \sum_{k=1}^K r_k = \sum_{k=1}^K \lambda/\mu_k$  el cual se interpreta como la carga ofertada de todo el sistema. Sea  $\rho$  el conjunto de índices de las colas con  $r_s \geq 1$ , y sea  $|\rho|$  el número de colas del conjunto. En adición a la condición de estabilidad, (véase el apéndice A.3.1) es importante encontrar una política que estabilice el sistema. Nos enfocaremos a la política ‘Cola más Larga’ (CL), la cual en muchos casos es fácil de implementar y no requiere complicados parámetros. En cualquier época de decisión, definimos la política CL que asigna al TF a la cola más larga. Si dos o más colas empatan a ser las más largas se elige a la que esté más cerca de éste.

**Hipótesis 3**  $q < |\rho| + 1$ , donde:  $q = \sum_{s \in \rho} r_s = \sum_{s \in \rho} (\lambda/\mu_s)$ , donde  $|\rho|$  es el número de colas del conjunto  $\rho$ , y  $\rho = \{s \mid r_s \geq 1 \text{ y } s \in \{1, 2, \dots, K\}\}$ .

**Teorema 2** El modelo puede ser estabilizado por la política CL si la hipótesis 3 se cumple. Además, modelo no puede ser estabilizado por cualquier política si  $q > |\rho| + 1$ .

**Demostración.** La demostración tiene dos partes: en la parte 1, probamos que si la hipótesis 3 se cumple, entonces CL estabiliza el sistema, y en la parte dos, probaremos que si  $q > |\rho| + 1$ , entonces el sistema es inestable.

**Parte 1.** Mostraremos que si se cumple la hipótesis 3 entonces la política CL puede estabilizar el sistema. Probaremos la parte 1 demostrando lo contrario, es decir: Si el modelo es inestable bajo la política CL, entonces la hipótesis 3 no se cumple. Sea  $\beta_s^{TF}$  la porción de tiempo a largo plazo en la cola  $s$  del TF, ya sea ocupado o desocupado como sigue:

Definimos  $I_{TF}^s(t)$   $t \geq 0$ ,  $s \in \{1, 2, \dots, K\}$  como el proceso estocástico en el cual la función indicadora que toma el valor de 1 cuando TF está sirviendo un trabajo ó simplemente está desocupado en la estación  $s$ . Esto es:

$$I_{TF}^s(t) = \begin{cases} 1 & \text{si TF está en la estación } s \text{ ya sea ocupado,} \\ & \text{o desocupado,} \\ 0 & \text{otro caso.} \end{cases}$$

Entonces:

$$\beta_s^{TF} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I_{TF}^s(t) dt.$$

De forma análoga definimos  $\gamma_s^{TF}$  la porción de tiempo a largo plazo del TF en la cola  $s$  ocupado (es decir trabajando). Esto es:

$$\tilde{I}_{TF}^s(t) = \begin{cases} 1 & \text{si TF está en la estación } s \text{ ocupado trabajando,} \\ 0 & \text{otro caso.} \end{cases}$$

Entonces:

$$\gamma_s^{TF} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \tilde{I}_{TF}^s(t) dt,$$

se tiene que:

$$\gamma_s^{TF} \leq \beta_s^{TF}.$$

Iniciaremos con un análisis preliminar que verifique que el sistema es lo suficientemente bien comportado para nuestra aproximación. Bajo la política CL, la cadena de Markov estacionaria describiendo el comportamiento del sistema es fácilmente reconocido a ser irreducible (véase el apéndice A.3.1). Note que la Cadena de Markov Controlada se establece basada en la regla de que los especialistas nunca están ociosos cuando un trabajo está presente. Esto garantiza que la Cadena de Markov es aperiódica (véase el apéndice A.3.2).

De la definición 3.11 y la sección 3.5.3 de [7], la Cadena de Markov es ergódica (débilmente ergódica en el caso de inestabilidad) y la distribución de estado estacionario (véase el apéndice A.3.1) en la Cadena de Markov existe. Dado que  $\beta_s^{TF}$  puede ser expresado en términos de la probabilidad de inicio en un conjunto de estados en la Cadena de Markov, además existe (como una constante casi seguramente) y está en  $[0, 1]$ .

Conociendo que  $\beta_s^{TF}$  existe, ahora retomamos la demostración. Esto es, probaremos que si nuestro modelo es inestable bajo una política CL, entonces la hipótesis 3 no se cumple.

La demostración se presenta en dos casos:

**Parte 1, caso 1.** Nuestra hipótesis será que todas las colas son inestables y mostraremos que la hipótesis 3 no se cumple.

Agregaremos rigor comprendiendo las implicaciones de la utilización del servicio (especialistas y TF), usaremos la Ley de Little (véase el apéndice A.3.1) aplicando individualmente, a los servidores especializados y al trabajador flotante solamente en  $s$ .

Para el caso especialista nos enfocaremos al modelo del subsistema que toma solamente el servidor (en cualquier cola) el tiempo promedio de espera es  $\mu_s^{-1}$ .

El número promedio, de trabajos a largo plazo en nuestro subsistema es  $L_{sub}^{sp}$ , poniendo la porción de tiempo a largo plazo del servidor (especializado) ocupado en  $s$ , el cual llamaremos  $\gamma_s^{sp}$ .

Es claro que cuando el parámetro de llegadas crece, y tal parámetro de flujo de los especialistas se aproxima a  $\mu_s$ , por abajo.

La Ley de Little (véase el apéndice A.3.1) prueba que  $\gamma_s^{sp}$  existe, es finita y se aproxima por abajo al valor 1 (mediante el 100% de los especialistas ocupados).

En otras palabras, repitiendo este ejercicio y enfocandonos en el proceso de completar los trabajos por TF, revela que TF se generará un parámetro en  $s$  dado por  $\delta_s^{TF} = \gamma_s^{TF} \mu_s$ , trabajos por unidad de tiempo.

Así entonces, combinando los parámetros de salida en  $s$ , se tiene:

$$\delta_s^{sp} + \delta_s^{TF} = (1 + \gamma_s^{TF})\mu_s.$$

Dado que la dinámica del sistema de colas entero, no permite que el parámetro de trabajos de salida en cualquier estación de trabajo pueda exceder el parámetro de llegadas al sistema entero. Esto es:

$$(1 + \gamma_s^{TF})\mu_s \leq \lambda. \tag{3.4}$$

Bajo CL, el TF nunca puede estar ocioso en el sistema cuando dos ó más trabajos están presentes en al menos una cola. Cuando todas las colas en  $\rho$  son inestables,

éstas pueden crecer, de manera no acotada, sobre una longitud de tiempo infinito bajo la política CL, la porción de tiempo esperado a largo plazo del TF ocupado trabajando en las colas en  $\rho$  es 1, y:

$$\gamma_s^{TF} = \beta_s^{TF}. \quad (3.5)$$

Combinando (3.4) y (3.5), multiplicando ambos lados por  $\mu_s^{-1}$  y usando la definición de  $r_s$  se tiene:

$$\begin{aligned} (1 + \beta_s^{TF})\mu_s &\leq \lambda, \\ \mu_s^{-1}(1 + \beta_s^{TF})\mu_s &\leq \lambda\mu_s^{-1}, \\ (1 + \beta_s^{TF}) &\leq \frac{\lambda}{\mu_s}, \\ \beta_s^{TF} &\leq \frac{\lambda}{\mu_s} - 1, \\ \beta_s^{TF} &\leq r_s - 1. \end{aligned}$$

Además, por definición:

$$\sum_{s \in \rho} \beta_s^{TF} = 1,$$

y se tiene también que las colas en  $\{1, 2, \dots, k\} - \rho$  son estabilizadas solamente por especialistas y no reciben ayuda del TF ‘en la cola más larga’ cuando las colas en  $\rho$  son inestables (véase el apéndice A.3.1), se tiene:

$$1 = \sum_{s \in \rho} \beta_s^{TF} \leq \sum_{s \in \rho} (r_s - 1) = q - |\rho|,$$

esto es:  $q \geq |\rho| + 1$ , la cual es la negación de la hipótesis 3.

**Parte 1, caso 2.** Ahora consideremos la posibilidad de que bajo la política CL algunas de las colas en  $\rho$  son inestables (véase el apéndice A.4) (esto es, existe un conjunto no vacío de colas inestables,  $U \subset \rho$ ), mientras que las otras colas en  $\rho$  son estables (esto es, existe un conjunto no vacío de colas estables,  $Y = \rho - U$ ). Mostraremos que esto no puede pasar bajo la política CL, del caso 1 se establece el resultado.

Se sigue que  $(r_s - 1) \geq 0$  para  $s \in Y$  y la hipótesis de que las colas en  $Y$  son estables, bajo cualquier política (incluyendo CL) TF gasta una porción promedio a

largo plazo positiva de su tiempo en cualquier cola en  $Y$  al estabilizar este ( $\gamma_s^{TF} > 0 : s \in Y$ ). Esto es posible que TF gaste tiempo en las colas en  $\{1, 2, \dots, k\} - \rho$  pero no necesariamente porque todas las colas en este conjunto pueden ser estabilizadas por los especialistas sin ayuda del TF. En contraste, cualquier cola en  $U$  es inestable y tendrá una longitud de cola más larga que cualquier cola en  $Y$  ó en  $\{1, 2, \dots, k\} - \rho$  sobre una longitud de tiempo no acotada.

Esto implica que CL nunca asignará a TF a cualquier cola en  $Y$  (esto es  $\gamma_s^{TF} = 0, s \in Y$ ), ¡lo cual es una contradicción!

**Parte 2.** Demostraremos que: si  $q > |\rho| + 1$ , entonces el sistema es inestable.

De igual manera que en la demostración anterior, probaremos lo contrario, esto es:

Si el sistema es estable, entonces:

$$q \geq |\rho| + 1.$$

Nos referiremos a la aproximación usando la Ley de Little (véase el apéndice A.3.1). Cuando el sistema es estable, todas las estaciones de trabajo son estables y no hay largos términos en cualquier cola. Aquí, la combinación del parámetro de salida desde cualquier estación  $s \in \rho$ , (es decir,  $\delta_s^{sp} + \delta_s^{TF}$ ) pueden ser igual a lo más al parámetro de llegadas esto es

$$\delta_s^{sp} + \delta_s^{TF} = \lambda.$$

Porque la longitud promedio del modelo del subsistema de los servidores especializados solamente puede ser a lo más 1.

El parámetro de salida del especialista puede ser a lo más  $\mu_s$ , es decir,  $\delta_s^{sp} \leq \mu_s$ , de donde se tiene:

$$\begin{aligned} \beta_s^{TF} \mu_s &\geq \delta_s^{TF} \geq \lambda - \mu_s, \\ \beta_s^{TF} &\geq (\lambda - \mu_s) / \mu_s = r_s - 1. \end{aligned} \tag{3.7}$$

Observe que esta es una restricción básica que TF asigne exactamente 100% de su tiempo (trabajando u ocioso) al sistema de colas entero. Bajo nuestra definición de

$\beta_s$  y usando (3.7) se tiene:

$$\begin{aligned} \sum_{s=1}^K \beta_s^{TF} &= 1, \\ \sum_{s \in \rho} \beta_s^{TF} &\leq 1, \\ \sum_{s \in \rho} (r_s - 1) &\leq 1, \\ \sum_{s \in \rho} r_s &\leq 1 + \sum_{s \in \rho} 1, \\ q &\leq |\rho| + 1. \end{aligned}$$

■

El siguiente teorema es el resultado principal del capítulo. Demuestra que una política óptima para un tiempo continuo, del modelo de estados infinito  $\Psi$  existe, y una política control y un costo pueden ser determinados precisamente por computo, usando iteración de valores en  $\Delta_l$  el modelo de espacio de estados finito, en tiempo discreto.

**Teorema 3** *Supóngase que la hipótesis 3 se cumple. Entonces el costo promedio en  $\Delta_l$  es una constante  $J_l$  y una política óptima estacionaria  $e_l$  puede ser calculada usando iteración de valores. Además,  $J_l \rightarrow J^{(1)}$  y la política límite de  $e_l$ , denotada por  $e$ , es óptima para  $\Psi$ .*

**Demostración.** Seguiremos la aproximación de la sección 2.5 excepto que aquí trataremos con el costo promedio como se definió en las ecuaciones (3.2). La última definición es como sigue. Dada una política  $\delta$  y un estado inicial  $i$ , sea  $F_n$  el costo total incurrido durante los primeros  $n$  periodos de transición, y sea  $T_n$  el tiempo total de estas transiciones. Entonces definimos:

$$G_\delta^{(1)}(i) = \limsup_{n \rightarrow \infty} \frac{E_\delta^{(i)}[F_n]}{E_\delta^{(i)}[T_n]},$$

y sea  $G_\delta^{(1)}(i)$  el ínfimo sobre todas las políticas.

El argumento tiene varios pasos. Primero emplearemos la Proposición 12 del apéndice B para probar que (3.3) se tiene para una política estacionaria óptima.

La política estacionaria usada para verificar esta hipótesis de este resultado es la política CL.

Tal regla induce una Cadena de Markov recurrente positiva irreducible (véase el apéndice A.3.1) en todo el espacio de estados y produce un sistema estable con longitud de cola, con media finita. De este modo, el costo promedio es finito. Este es entonces el caso en el que una política estacionaria existe, satisfaciendo (2.17) con  $Z$  reemplazado por  $J^{(1)}$ . Esto implica que  $G_f^{(1)}(i) \leq J^{(1)}$ .

Además, por la Proposición 11 del apéndice B, se tiene el caso que  $G_f^{(1)}(i) \leq J_f^{(1)}(i)$ .

Esto implica que (3.3) se cumple tal que  $f$  es costo promedio óptima para  $\Psi$ . Se sigue entonces del Lema 2 y Lema 3 del apéndice B que  $J_f^\Delta(i) \leq J^{(1)}$ , donde  $J_f^\Delta$  es el costo promedio usual por unidad de tiempo en  $\Delta$  bajo  $f$ .

Ahora turnamos nuestra atención a  $\Delta$  y verificamos que las suposiciones (AC) del apéndice B se cumplen.

Estas suposiciones garantizan la existencia de una política estacionaria óptima con costo promedio constante en  $\Delta_l$ ; que el límite de estos costos promedio convergen al costo promedio en  $\Delta$ , y que cualquier punto límite de políticas estacionarias computadas por el algoritmo de iteración de valores es óptima en  $\Delta$ .

La hipótesis (AC1) del apéndice B afirma que existe una solución de la ecuación de iteración de valores en  $\Delta_l$ . Para demostrar esto, es suficiente, por la Proposición 8 del apéndice B mostrar que cualquier política estacionaria tiene clases recurrentes positivas y aperiodicas (véase el apéndice A.3.2). Dado que la creación de  $\Delta$  (y por lo tanto de  $\Delta_l$ ) siempre incluye transiciones desde un estado inmediato anterior a sí mismo, esto es claro. Para verificar las otras suposiciones, emplearemos los pasos 2-4 de la Proposición 9 del apéndice B.

Para la política en el paso 2 tomamos la regla CL. Podemos verificar bajo esta regla que si las colas están limitadas a  $l$ , como en  $\Delta_l$ , entonces la distribución estacionaria (véase el apéndice A.3.2) y costo promedio converge para la regla CL. Esto se sigue de las ideas de la Proposición 10 del apéndice B. Como nuestro ‘punto base’ tomamos el sistema vacío, con el TF en la estación 1. Necesitamos comparar el tiempo esperado y la participación del costo esperado desde el ‘sobreflujo’, un estado atrás al punto base con las cantidades respectivas en el sistema restringido.

La forma puede exceder la última. Esto es intuitivamente claro, dado si los trabajos son ejecutados desde el sistema, entonces pensamos sólo en recibir el mejor.

Al verificar el paso 3, argumentamos que el mínimo costo esperado de operar en  $\Delta_l$  en  $l$  pasos no excede la cantidad respectiva en  $\Delta$ . Nuevamente esto es intuitivamente claro.

Esto verifica que la hipótesis AC se cumple y por lo tanto una política estacionaria óptima  $e$  para  $\Delta$  puede ser computada como un punto límite desde la política estacionaria óptima en  $\Delta_l$ . Además el costo promedio  $J^\Delta$  en  $\Delta$  y  $e$  satisface (2.21). Pero esto implica que  $G_e^{(1)}(i) \leq J^\Delta$ .

Poniendo estos puntos se obtiene que:

$$G_e^{(1)}(i) \leq J^\Delta \leq J_f^\Delta(i) \leq J^{(1)} \leq J_e^{(1)}(i).$$

Ahora argumentamos intuitivamente que la política  $e$  puede inducir una cadena de Markov recurrente positiva (véase el apéndice A.3.2) en todo el espacio de estados, y además los costos son no acotados. Esto implica por ([10], Teorema 7.5), que el término exterior de la serie de la desigualdad anterior son iguales. Esto prueba que  $e$  es costo promedio óptimo para  $\Psi$  con costo promedio constante  $J^\Delta$ . Por lo tanto, los procedimientos computacionales son válidos. ■

### 3.4. Ecuación de Iteración de Valores

Sean  $i$  el estado del proceso  $n$ , el número de iteración, y  $s$  el número de estación. El algoritmo de iteración de valores involucra a una sucesión  $u_n(i)$  y  $w_n(i)$ . Comenzamos poniendo  $u_0 \equiv 0$  como una condición inicial. Elegimos el vector 0 como nuestro punto base ó referencia, y definimos

$$w_n(i) = g(i) + \tau \lambda u_n(i + e_1) + \tau \sum \mu_s I(i_s \geq 1) u_n(i - e_s + e_{s+1}) \\ + \underset{a}{\text{mín}} \{ \tau \mu_a I(i_a \geq 2) u_n(i - e_a + e_{a+1}) + (1 - \tau v(i, a)) u_n(i) \}, \quad (3.8)$$

actualizamos  $u_{n+1}(i) = w_n(i) - w_n(0)$ .

Modificando  $\Delta$  al crear  $\Delta_l$  es fácil como sigue. Si  $i_1 = l$ , entonces reemplazamos  $e_1$  en (3.8) con el vector cero, y tal nuevo trabajo es turnado fuera. El caso de

completar un servicio se resuelve similarmente, tal trabajo es eliminado siempre que esta transferencia causa un sobreflujo en la cola.

Los programas a calcular numéricamente el PDM, se calculan usando (3.8), modificados para  $\Delta_l$ . Este es el caso en que  $w_n(0) \rightarrow J_l$  y la política estacionaria (dependiente de  $n$ ) realiza el mínimo en (3.8) convergiendo a  $e_l$ , una política óptima para  $\Delta_l$ . Si deseamos entonces uno puede recomputar para valores grandes de  $l$  y usar el Teorema 2.2 para determinar  $J^{(1)}$  y una política óptima para  $\Psi$ . Este es usualmente el caso en que una política óptima es determinada para valores pequeños de  $N$ .

# Capítulo 4

## N Líneas de Producción con NK Trabajadores Especializados y un Trabajador Flotante

### 4.1. Descripción del Problema

El siguiente modelo describe  $N$  líneas de producción en serie e independientes entre sí, con  $NK + 1$  trabajadores y  $NK$  estaciones etiquetadas como:  $(i, j)$ , para  $i = 1, 2, \dots, N$  y  $j = 1, 2, \dots, K$  cada estación tiene un almacén infinito de trabajos.

Las estaciones están formadas en  $N$  líneas. En cada estación hay dos estaciones de trabajo, una que es atendida por un trabajador especializado y la otra está disponible para un TF, cuando se necesite. El tiempo de completar un trabajo en una estación de trabajo  $(i, s)$ ,  $i = 1, 2, \dots, N$ ;  $s = 1, 2, \dots, K$ , está distribuido exponencialmente con parámetro  $\mu_{is}$  considerando si el especialista ó el TF está trabajando ahí.

Este modelo es un sistema de producción abierto con llegadas de acuerdo a un proceso de Poisson con parámetro  $\lambda_i$  (véase el apéndice A.2) con  $i = 1, 2, \dots, N$  para las líneas 1, 2, ...,  $N$  respectivamente. Las líneas no comparten trabajos, es decir, los trabajos que llegan a la línea 1 se terminan en la línea 1, así mismo para la línea 2, sucesivamente hasta la línea  $N$ .

Si el trabajador flotante se mueve a la estación  $(i, s)$ , él puede iniciar a servir un

segundo trabajo en la estación  $(i, s)$ , habiendo un segundo trabajo en tal estación.

El TF no colabora en el mismo trabajo con un especializado, si hay solo un trabajo en la estación, TF no trabajará en éste. El tiempo de servicio de un segundo trabajo en la estación  $(i, s)$  está distribuido exponencialmente con parámetro  $\mu_{is}$ . A continuación se presenta el modelo.

## 4.2. Formulación del Modelo como un PDMTC

A continuación presentamos el modelo como un Proceso de Decisión de Markov en Tiempo Continuo denotado como  $\Psi$ .

1. El *Espacio de estados*: Sea  $M$  una matriz de dimensión  $N \times K$ .

$$S = \{M \mid M = (m_{is}), i = 1, 2, \dots, N \text{ y } s = 1, 2, \dots, K, m_{is} \in \mathbb{N}\},$$

donde las entradas de los elementos de  $S$ ,  $m_{is}$  denotan el número de trabajos en la línea  $i$  y la estación  $s$ .

2. *Épocas de decisión*: Las épocas de decisión ocurren cuando un trabajo entra al sistema (siempre entran al almacén  $(i, 1)$ , para  $i = 1, 2, \dots, N$ ) ó un trabajo se ha completado y si ocurre en la estación  $(i, s)$  entonces tal trabajo será sustraído del almacén y se agregará al almacén  $(i, s + 1)$ .

3. *Espacio de acciones*: El espacio de acciones son:

$A = \{(i, s) \mid i = 1, 2, \dots, N; s = 1, 2, \dots, K\}$  donde elegir  $a \in A$ , significa que el TF se moverá (instantáneamente) a la estación  $a$  (o permanecerá en ella si él se encuentra ahí).

4.  $H_{is} > 0$ , será el parámetro de costo por unidad de tiempo que el trabajo esté en la línea  $i$  y la estación  $s$ . Si está en el estado  $M$  y una acción  $a$  se elige, entonces se incurre en un costo dado por:

$$g(M, a) = \sum_{i=1}^N \sum_{s=1}^K H_{is} m_{is} = h(M).$$

Estos costos no son instantáneos.

5. El tiempo de transición de un estado a otro está distribuido exponencialmente con parámetro:

$$v(M, a) = \sum_{i=1}^N \lambda_i + \sum_{i=1}^N \sum_{s=1}^K \mu_{is} I(m_{is} \geq 1) + \mu_a I(m_a \geq 2), \quad (4.1)$$

donde  $I(E)$  es la función indicadora para el evento  $E$ . El primer término de lado derecho es el nivel (parámetro) de llegadas. El segundo término es el parámetro de servicio de los trabajadores especializados y el último término es el parámetro de servicio de TF.

**Demostración de 5.** Partimos del hecho de que una transición ocurre cuando un trabajo entra o cuando un trabajo se ha terminado en alguna de las estaciones de trabajo. Así entonces:

Sean  $Y_i$  variables aleatorias independientes exponencialmente distribuidas con parámetros  $\lambda_i$  respectivamente; y  $X_{is}, F_a$ , variables aleatorias independientes, exponencialmente distribuidas con parámetros  $\mu_{is}, \mu_a$ ,  $i = 1, 2, \dots, N$ ,  $s = 1, 2, \dots, K$  respectivamente, entonces:

Sea  $T = \min \{Y_i, X_{is}, F_a \mid i = 1, 2, \dots, N, s = 1, 2, \dots, K\}$ . Entonces  $T$  está distribuido exponencialmente con parámetro  $v(E, a)$ . Lo anterior se sigue de la proposición 4 del apéndice A.2. ■

Para propósitos computacionales transformamos  $\Psi$  bajo un Proceso de Decisión de Markov (PDM)  $\Delta$  usando uniformización (véase sección 2.5).

El espacio de estados para  $\Delta$  es el mismo que para  $\Psi$ . Ahora elegimos un  $\tau$  fijo, el periodo de transición del sistema uniformizado (en unidades de tiempo por transición), satisfaciendo:

$$\tau = \left( \sum_{i=1}^N \lambda_i + \sum_{i=1}^N \sum_{s=1}^K \mu_{is} + \max \{ \mu_{is} \} \right)^{-1}. \quad (4.2)$$

Supongamos que el sistema está en el estado  $M = (m_{is}), i = 1, 2, \dots, N$  y  $s = 1, 2, \dots, K$  al inicio de un periodo y se toma una decisión  $a$ . Sea  $E_{ij}$  la matriz de dimensión  $N \times K$  donde la componente  $e_{ij} = 1$  y los demás son cero. Así entonces al comienzo del siguiente periodo el sistema tendrá una transición a:

$M + E_{i1}$  con probabilidad  $\tau\lambda_i$ ; tendrá una transición a  $M - E_{is} + E_{is+1}$  con probabilidad  $\tau\mu_{is} [I(m_{is} \geq 1) + I(a = m_{is}, m_{is} \geq 2)]$  para  $i = 1, 2, \dots, N$

Y de 4.1 permanecerá en  $M$  con probabilidad  $1 - \tau v(M, a)$

Definimos la matriz  $E_{i,K+1}$ , que controla el caso de completar un trabajo en la línea  $i$  y en la última estación. El costo en que se incurre al comenzar el siguiente periodo de tiempo es:

$$C(M, a) = h(M).$$

El problema clave aquí es que  $\Delta$  tiene un almacén infinito, y podemos buscar un camino para implementar cálculos rigurosos con solamente un espacio de estados finito. Empleamos el método de la aproximación secuencial (*AS*) introducido en [13] al reemplazar  $\Delta$  con una sucesión  $\{\Delta_l\}_{l \geq N_0}$  de espacio de estados finito PDM en el cual se hacen cálculos.

El espacio de estados  $\Delta_l$  consiste de matrices  $M$  donde la componente  $m_{is} \leq N$ ,  $i=1,2,\dots$ , (el máximo número de trabajos en el almacén  $(i, s)$  es  $l$ ).

El método *AS* no es un simple esquema de truncamiento. Distribuimos los ‘excesos de probabilidad’ que causa que el sistema salga fuera de cota. El sistema puede escapar solamente a  $\Delta_l$  desde un estado acotado para el cual algún almacén está lleno. Si  $m_{i1} = l$ , un sobreflujo ocurrirá en el almacén 1 de la línea  $i$ , si un nuevo trabajo llega, el cual ocurre con probabilidad  $\tau\lambda_i$ ,  $i = 1, 2, \dots, N$ .

Este exceso es asignado al estado  $M$ , efectivamente se le niega la entrada a un nuevo trabajo en la línea  $i$  y en la estación 1, cuando el almacén está en el nivel  $l$ , hasta que se completa un trabajo. Similarmente, si  $m_{is} = l$  para  $i = 1, 2, \dots, N$ , para algún  $s$ ,  $s = 1, \dots, K$ , entonces con probabilidad

$$\tau\mu_{is-1} [I(m_{is-1} \geq 1) + I(a = (i, s - 1), m_{is-1} \geq 2)].$$

El siguiente estado es  $M' = (m'_{is}), i = 1, 2, \dots, N$  y  $s = 1, 2, \dots, K$ , donde  $m'_{is-1} = m_{is-1} - 1$ , y las demás coordenadas de  $M'$  son iguales a las de  $M$ .

El efecto de esta asignación es remover los trabajos completados en  $(i, s - 1)$  desde el almacén previo y descartar estos. El costo de  $\Delta_l$  es el mismo que  $\Delta$ .

En resumen reducimos el PDMTC  $\Psi$  al asociado espacio de estados infinito en tiempo discreto PDM  $\Delta$  el cual aproximamos por una sucesión de espacio de estados finito  $\{\Delta_l\}_{l \geq N_0}$  para propósitos computacionales.

Similarmente a la hipótesis 3, tenemos:

**Hipótesis 4**  $q < |\rho| + 1$ , donde:  $q = \sum_{is \in \rho} r_{(i,s)} = \sum_{is \in \rho} \frac{\lambda_i}{\mu_{(i,s)}}$ , y donde  $|\rho|$  es el número de colas del conjunto  $\rho$ ,  $\rho = \{(i,s) \mid r_{(i,s)} \geq 1, \text{ y } i = 1, 2, \dots, N; s = 1, 2, \dots, K\}$ .

### 4.3. Optimalidad, Estabilidad y Acotamiento

**Teorema 4** *El modelo puede ser estabilizado por una política CL si la hipótesis 4 se cumple. Además TF no puede ser estabilizado por cualquier política si  $q > |\rho| + 1$ .*

**Demostración.** De manera análoga al Teorema 2, de los antecedentes, la prueba se hará en dos partes; en la parte 1, probamos que si la hipótesis 4 se cumple, entonces CL puede estabilizar el sistema, y en la parte 2, probaremos que si  $q > |\rho| + 1$ , entonces el sistema es inestable.

**Parte 1.** Demostraremos: ‘si se cumple la hipótesis 4 entonces CL puede estabilizar el sistema’.

Probaremos la parte 1 demostrando lo contrario, es decir: Si el modelo es inestable bajo la política CL, entonces la hipótesis 4 no se cumple.

Sea  $\beta_{(i,s)}^{TF}$  la porción de tiempo a largo plazo en la cola  $(i, s)$  del TF, ya sea ocupado o desocupado como sigue: Definimos  $I_{TF}^{(i,s)}(t)$  como el proceso estocástico en el cual la función indicadora que toma el valor de 1 cuando TF está sirviendo un trabajo ó simplemente está desocupado en la estación  $(i, s)$ . Esto es:

$$I_{TF}^{(i,s)}(t) = \begin{cases} 1 & \text{si TF está en la estación } (i, s) \text{ ya sea ocupado} \\ & \text{o desocupado,} \\ 0 & \text{otro caso.} \end{cases}$$

Entonces:

$$\beta_{(i,s)}^{TF} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T I_{TF}^{(i,s)}(t) dt.$$

De forma análoga definimos  $\gamma_{(i,s)}^{TF}$  la porción de tiempo a largo plazo del TF en la cola  $(i, s)$  trabajando. Esto es:

$$\tilde{I}_{TF}^s(t) = \begin{cases} 1 & \text{si TF está en la estación } (i, s) \text{ ocupado trabajando,} \\ 0 & \text{otro caso.} \end{cases}$$

se tiene que:

$$\gamma_{(i,s)}^{TF} \leq \beta_{(i,s)}^{TF}.$$

Iniciaremos con un análisis preliminar que verifique que el sistema es lo suficientemente bien comportado para nuestra aproximación.

Bajo CL, la cadena de Markov estacionaria describiendo el comportamiento del sistema es fácilmente reconocido a ser irreducible (véase el apéndice A.3.1). Note que la cadena de Markov controlada se establece basada en la regla de que los especialistas nunca están ociosos cuando un trabajo está presente. Esto garantiza que la cadena de Markov es aperiódica.

De la definición 3.11 y la sección 3.5.3 del [7], la cadena de Markov es ergódica (débilmente ergódica en el caso de inestabilidad) y la distribución de estado estacionario en la cadena de Markov existe. Dado que  $\beta_{(i,s)}^{TF}$  puede ser expresado en términos de la probabilidad de inicio en un conjunto de estados en la cadena de Markov, además existe (como una constante casi seguramente) y está en  $[0, 1]$ .

Conociendo que  $\beta_{(i,s)}^{TF}$  existe, ahora retomamos la demostración. Esto es, probaremos que si el modelo es inestable bajo una política CL, entonces la hipótesis 4 no se cumple.

La demostración se presenta en dos casos:

**Parte 1, caso 1.** Nuestra hipótesis será que todas las colas son inestables (véase el apéndice A.3.1) y probaremos que la hipótesis 4 no se cumple.

Agregaremos rigor comprendiendo las implicaciones de la utilización del servicio (especialistas y TF), usaremos la Ley de Little (véase por ejemplo Teorema 7.4 de [7]) aplicando individualmente a los servidores especializados solamente y al trabajador flotante solamente en  $(i, s)$ .

Para el caso especialista nos enfocaremos al modelo del subsistema que toma solamente el servidor (en cualquier cola) el tiempo promedio de espera es  $\mu_{(i,s)}^{-1}$ .

El número promedio de trabajos a largo plazo, de nuestro subsistema es  $L_{sub}^{sp}$ , poniendo la porción promedio de tiempo a largo plazo del servidor (especializado) ocupado en  $(i, s)$ , el cual llamaremos  $\gamma_{(i,s)}^{sp}$ .

Es claro que cuando el parámetro de llegadas crece, y tal parámetro de flujo de los especialistas se aproxima a  $\mu_{(i,s)}$  por abajo.

#### 4. N Líneas de Producción con NK Trabajadores Especializados y un Trabajador Flotante 38

---

La Ley de Little (véase el apéndice A.3.1) prueba que  $\gamma_{(i,s)}^{sp}$  existe, es finita y se aproxima por abajo al valor 1 (mediante el 100 % de los especialistas ocupados).

En otras palabras, repitiendo este ejercicio y enfocandonos en el proceso de completar los trabajos por el TF, revela que el TF generará un parámetro en  $(i, s)$  de

$$\delta_{(i,s)}^{TF} = \gamma_{(i,s)}^{TF} \mu_{(i,s)},$$

trabajos por unidad de tiempo.

Así entonces, combinando los parámetros de salida en  $(i, s)$ , se tiene:

$$\delta_{(i,s)}^{sp} + \delta_{(i,s)}^{TF} = (1 + \gamma_{(i,s)}^{TF}) \mu_{(i,s)}.$$

Dado que la dinámica del sistema para cada línea no permite que el número de trabajos de salida en cualquier estación de trabajo pueda exceder el parámetro de llegadas de cada línea, esto es:

$$(1 + \gamma_{(i,s)}^{TF}) \mu_{(i,s)} \leq \lambda_i \quad i = 1, 2, \dots, N. \quad (4.3)$$

Bajo CL, TF nunca puede estar ocioso en el sistema cuando dos ó más trabajos están presentes en al menos una cola.

Cuando todas las colas en  $\rho$  son inestables, (véase el apéndice A.3.1) éstas pueden crecer de manera no acotada sobre una longitud de tiempo infinito, bajo la política CL, la fracción de tiempo esperado más larga del TF ocupado trabajando en las colas en  $\rho$  es 1, y:

$$\gamma_{(i,s)}^{TF} = \beta_{(i,s)}^{TF}. \quad (4.4)$$

Combinando (4.3) y (4.4), multiplicando ambos lados por  $\mu_{(i,s)}^{-1}$  y usando la definición de  $r_{(i,s)}$  se tiene:

$$\begin{aligned} (1 + \beta_{(i,s)}^{TF}) \mu_{(i,s)} &\leq \lambda_i, \\ \mu_{(i,s)}^{-1} (1 + \beta_{(i,s)}^{TF}) \mu_{(i,s)} &\leq \lambda_i \mu_{(i,s)}^{-1}, \\ (1 + \beta_{(i,s)}^{TF}) &\leq \frac{\lambda_i}{\mu_{(i,s)}}, \\ \beta_{(i,s)}^{TF} &\leq \frac{\lambda_i}{\mu_{(i,s)}} - 1, \\ \beta_{(i,s)} &\leq r_{(i,s)} - 1. \end{aligned}$$

Además por definición:

$$\sum_{(i,s) \in \rho} \beta_{(i,s)}^{TF} = 1,$$

y dado que las colas en  $\{1, 2, \dots, k\} - \rho$  son estabilizadas (véase apéndice A.3.1) solamente por especialistas y no reciben ayuda del TF ‘en la más larga’ cuando las colas en  $\rho$  son inestables, se tiene:

$$1 = \sum_{(i,s) \in \rho} \beta_{(i,s)}^{TF} \leq \sum_{(i,s) \in \rho} (r_{(i,s)} - 1) = q - |\rho|,$$

esto es:  $q \geq |\rho| + 1$ , la cual es la negación de la hipótesis 4.

**Parte 1, caso 2.**

Ahora consideremos la posibilidad de que bajo la política CL algunas de las colas en  $\rho$  son inestables (esto es, existe un conjunto no vacío de colas inestables,  $U \subset \rho$ ), mientras que las otras colas en  $\rho$  son estables (esto es, existe un conjunto no vacío de colas estables,  $Y = \rho - U$ ).

Mostraremos que esto no puede pasar bajo la política CL, del caso 1 se establece el resultado.

Se sigue que  $(r_{(i,s)} - 1) \geq 0$  para  $(i, s) \in Y$  y la hipótesis de que las colas en  $Y$  son estables, bajo cualquier política (incluyendo CL) TF gasta una fracción promedio más larga positiva de su tiempo en cualquier cola en  $Y$  al estabilizar este ( $\gamma_{(i,s)}^{TF} > 0$ ;  $(i, s) \in Y$ ). Esto es posible que TF gaste tiempo en las colas en  $\{1, 2, \dots, k\} - \rho$  pero no necesariamente porque todas las colas en este conjunto pueden ser estabilizadas por los especialistas sin ayuda del TF. En contraste, cualquier cola en  $U$  es inestable y tendrá una longitud de cola más larga que cualquier cola en  $Y$  ó en  $\{1, 2, \dots, k\} - \rho$  sobre una longitud de tiempo no acotada.

Esto implica que CL nunca asignará al TF a cualquier cola en  $Y$  (esto es  $\gamma_{(i,s)}^{TF} = 0$ ,  $(i, s) \in Y$ ), ¡lo cual es una contradicción!

**Parte 2.** Demostraremos que: si  $q > |\rho| + 1$ , entonces el sistema es inestable.

De igual manera que en la demostración anterior, probaremos lo contrario, esto es:

Si el sistema es estable, entonces:

$$q \geq |\rho| + 1.$$

Nos referiremos a la aproximación usando la Ley de Little.

Cuando el sistema es estable, todas las estaciones de trabajo son estables y no hay largos términos en cualquier cola. Aquí, la combinación del parámetro de salida desde cualquier estación  $(i, s) \in \rho$ , (es decir,  $\delta_{is}^{sp} + \delta_{is}^{TF}$ ) pueden ser igual a lo más al parámetro de llegadas

$$\delta_{(i,s)}^{sp} + \delta_{(i,s)}^{TF} = \lambda.$$

Porque la longitud promedio del modelo del subsistema de los servidores especializados solamente puede ser a lo más 1.

El parámetro de salida del especialista puede ser a lo más  $\mu_{(i,s)}$ , es decir,  $\delta_{is}^{sp} \leq \mu_{(i,s)}$ , de donde se tiene:

$$\begin{aligned} \beta_{is}^{TF} \mu_{is} &\geq \delta_{is}^{TF} \geq \lambda - \mu_{is}, \\ \beta_{is}^{TF} &\geq (\lambda - \mu_{is}) / \mu_{is} = r_{is} - 1. \end{aligned} \tag{4.5}$$

Observe que esta es una restricción básica que TF asigne exactamente 100% de su tiempo (trabajando u ocioso) al sistema entero de la cola. Bajo nuestra definición de  $\beta_{(i,s)}$  y usando (4.5) se tiene:

$$\begin{aligned} \sum_{s=1}^K \sum_{i=1}^N \beta_{(i,s)}^{TF} &= 1, \\ \sum_{(i,s) \in \rho} \beta_{(i,s)}^{TF} &\leq 1, \\ \sum_{(i,s) \in \rho} (r_{(i,s)} - 1) &\leq 1, \\ \sum_{(i,s) \in \rho} r_{(i,s)} &\leq 1 + \sum_{(i,s) \in \rho} 1, \\ q &\leq |\rho| + 1. \end{aligned}$$

■

**Teorema 5** *Supóngase que la hipótesis 4 se cumple. Entonces el costo promedio en  $\Delta_l$  es una constante  $J_l$  y una política óptima estacionaria  $e_l$  puede ser calculada usando iteración de valores. Además,  $J_l \longrightarrow J^{(2)}$  y la política límite de  $e_l$ , denotada por  $e$ , es óptima para  $\Psi$ .*

**Demostración.** Seguiremos la aproximación de ([13, pags. 243-248]) excepto que nosotros trataremos con el costo promedio como se definió en las ecuaciones (3.2). La última definición es como sigue. Dada una política  $\delta$  y un estado inicial  $M$ , sea  $F_n$  el costo total incurrido durante los primeros  $n$  periodos de transición, y sea  $T_n$  el tiempo total de estas transiciones. Entonces definimos:

$$G_\delta^{(2)}(M) = \limsup_{n \rightarrow \infty} \frac{E_\delta^{(M)} [F_n]}{E_\delta^{(M)} [T_n]},$$

y sea  $G_\delta^{(2)}(M)$  el ínfimo sobre todas las políticas.

El argumento tiene varios pasos. Primero emplearemos la proposición 12 del apéndice B para probar que (3.3) se tiene para una política estacionaria óptima. La política estacionaria usada para verificar esta hipótesis de este resultado es la política CL.

Tal regla induce una Cadena de Markov recurrente positiva irreducible (véase el apéndice A.3.1) en todo el espacio de estados y produce un sistema estable con longitud de cola con media finita. De este modo, el costo promedio es finito. Este es entonces el caso en el que una política estacionaria existe, satisfaciendo (2.17) con  $Z$  reemplazado por  $J^{(2)}$ . Esto implica que  $G_f^{(2)}(M) \leq J^{(2)}$ .

Además, por la proposición 11 del apéndice B, se tiene el caso que  $G_f^{(1)}(M) \leq J_f^{(2)}(M)$ .

Esto implica que (3.3) se cumple tal que  $f$  es costo promedio óptima para  $\Psi$ . Se sigue entonces del Lema 2 y Lema 3 del apéndice B que  $J_f^\Delta(M) \leq J^{(2)}$ , donde  $J_f^\Delta$  es el costo promedio usual por unidad de tiempo en  $\Delta$  bajo  $f$ .

Ahora turnamos nuestra atención a  $\Delta$  y verificamos que las suposiciones (AC) del apéndice B se cumplen.

Estas suposiciones garantizan la existencia de una política estacionaria óptima con costo promedio constante en  $\Delta_l$ ; que el límite de estos costos promedio convergen al costo promedio en  $\Delta$ , y que cualquier punto límite de políticas estacionarias computadas por el algoritmo de iteración de valores es óptima en  $\Delta$ .

La hipótesis (AC1) del apéndice B afirma que existe una solución de la ecuación de iteración de valores en  $\Delta_l$ . Para demostrar esto, es suficiente, por la proposición 8 del apéndice B mostrar que cualquier política estacionaria tiene clases recurrentes positivas y aperiodicas (véase el apéndice A.3.2). Dado que la creación de  $\Delta$  (y por

lo tanto de  $\Delta_l$ ) siempre incluye transiciones desde un estado inmediato anterior a si mismo, esto es claro. Para verificar las otras suposiciones, emplearemos los pasos 2-4 de la Proposición 9 del apéndice B.

Para la política en el paso 2, tomamos la regla CL. Podemos verificar bajo esta regla que si las colas están limitadas a  $l$ , como en  $\Delta_l$ , entonces la distribución estacionaria (véase el apéndice A.3.2) y costo promedio converge para la regla CL. Esto se sigue de las ideas en la Proposición 10 del apéndice B. Como nuestro ‘punto base’ tomamos el sistema vacío, con el TF en la estación  $(1, 1)$ . Necesitamos comparar el tiempo esperado y la participación del costo esperado desde ‘sobreflujo’, un estado atrás al punto base con las cantidades respectivas en el sistema restringido. La forma puede exceder la última. Esto es intuitivamente claro, dado si los trabajos son ejecutados desde el sistema, entonces pensamos sólo en recibir el mejor.

Al verificar el paso 3, argumentamos que el mínimo costo esperado de operar en  $\Delta_l$  en  $n$  pasos no excede la cantidad respectiva en  $\Delta$ . Nuevamente esto es intuitivamente claro.

Esto verifica que la hipótesis AC se cumple y por lo tanto una política estacionaria óptima  $e$  para  $\Delta$  puede ser computada como un punto límite desde la política estacionaria óptima en  $\Delta_l$ . Además el costo promedio  $J^\Delta$  en  $\Delta$  y  $e$  satisface (2.21). Pero esto implica que  $G_e^{(2)}(M) \leq J^\Delta$ .

Poniendo estos puntos se obtiene que:

$$G_e^{(2)}(M) \leq J^\Delta \leq J_f^\Delta(M) \leq J^{(2)} \leq J_e^{(2)}(M).$$

Ahora argumentamos informalmente que la política  $e$  puede inducir una cadena de Markov recurrente positiva (véase el apéndice A.3.2) en todo el espacio de estados, y además los costos son no acotados. Esto implica por ([10], Teorema 7.5), que el término exterior de la serie de la desigualdad son iguales. Esto prueba que  $e$  es costo promedio óptimo para  $\Psi$  con costo promedio constante  $J^\Delta$ . Por lo tanto los procedimientos computacionales son válidos. ■

#### 4.4. Ecuación de Iteración de Valores

Sean  $M = (m_{is})$ ,  $i = 1, 2, \dots, N$  y  $s = 1, 2, \dots, K$ , el estado del proceso,  $n$ , el número de iteración y  $(i, s)$  el número de línea y estación respectivamente. El

algoritmo de iteración de valores involucra a una sucesión  $u_n(M)$  y  $w_n(M)$ . Comenzamos poniendo  $u_0 \equiv 0$  como una condición inicial. Elegimos la matriz 0 como nuestro punto base o referencia y definimos

$$w_n(M) = g(M) + \sum_{i=1}^N \tau \lambda_i u_n(M + E_{i1}) + \tau \sum_{i=1}^N \sum_{s=1}^K \mu_{is} I(E_{is} \geq 1) u_n(M - E_{is} + E_{i,s+1}) \\ + \min_a \{ \tau \mu_a I(E_a \geq 2) u_n(M - M_a + M_{a+1}) + (1 - \tau v(M, a)) u_n(M) \} \quad (4.6)$$

actualizamos  $u_{n+1}(E) = w_n(E) - w_n(0)$ .

Modificando  $\Delta$  al crear  $\Delta_l$  es fácil como sigue. Si  $E_{i1} = l$ , entonces reemplazamos  $E_1$  en (4.6) con el vector cero, y tal nuevo trabajo es turnado fuera. El caso de completar un servicio se resuelve similarmente, tal trabajo es eliminado siempre que esta transferencia causa un sobreflujo en la cola.

Los programas a computar numéricamente el PDM son computados usando (4.6), modificados para  $\Delta_l$ . Este es el caso en que  $w_n(0) \rightarrow J_l$  y la política estacionaria (dependiente de  $n$ ) realiza el mínimo en (4.6) convergiendo a  $e_l$ , una política óptima para  $\Delta_l$ . Si deseamos entonces uno puede recomputar para valores grandes de  $l$  y usar el Teorema 3.2 para determinar  $J^{(1)}$  y una política óptima para  $\Psi$ . Este es usualmente el caso en que una política óptima es determinada para valores pequeños de  $l$ .

# Capítulo 5

## Conclusiones

El objetivo de este trabajo ha sido el estudio de líneas de producción en serie, con trabajadores especializados, donde se dispone de un trabajador flotante. Nuestro interés en este tipo de problemas está motivado en el hecho de que muchos sistemas de manufactura están contruidos de esta forma. Por otro lado, cabe mencionar que es muy importante dar una política óptima de asignación del trabajador flotante a alguna de las líneas de producción, pues esto ayuda a hacerlas más eficientes.

Hemos considerado líneas de producción independientes entre sí, y la no colaboración de los trabajadores en un mismo trabajo, aunque se permite que el trabajador flotante trabaje en una misma estación, pero en diferente trabajo. Consideramos los costos por tener trabajo en proceso entre las estaciones de trabajo, y nuestro objetivo fue minimizar el promedio de dichos costos. También partimos del supuesto que el tiempo de llegadas cumple con la distribución de Poisson y el tiempo de servicio con la distribución exponencial.

En el capítulo 3 expusimos los antecedentes, sobre nuestro trabajo, es decir el caso de una sola línea de producción como se expone en [14].

En el capítulo 4 hicimos una extensión para diversas líneas de producción. Además dimos las suposiciones necesarias para su planteamiento, y posteriormente, se presentó la construcción del modelo de matemático correspondiente. Este modelo se construyó usando la teoría de los Procesos de Decisión de Markov a Tiempo Continuo (PDMTC). Después hicimos una discretización del mismo, presentandolo como un Proceso de Decisión de Markov a Tiempo Discreto (PDMTD). Aún con

este modelo a tiempo discreto, teníamos limitaciones ya que también suponíamos que el número de trabajos en cualquier estación podría ser infinito y esto generaba un número infinito de estados. Por esto dimos un esquema de truncamiento, llamado método de Aproximación Secuencial (AS), el cual nos permitió que el espacio de estados fuese finito y con esto dar un algoritmo iterativo para encontrar la política óptima. Propusimos, como política óptima, la política: ‘cola más larga’ (CL) esto es que el trabajador flotante se asigna a la estación con mayor número de trabajos. Cabe mencionar que esta política fue primero mostrada en [14]. Demostramos además que bajo la política CL el modelo puede ser estabilizado siempre que la siguiente condición se cumpla: la carga del sistema entero (véase el apéndice A.3.1) es estrictamente menor que el número de estaciones cuya carga es mayor que cierto número. Más aún, probamos que ninguna política puede estabilizar el modelo si esta condición no se cumple. Por otra parte dimos un teorema que nos permite calcular la política óptima a través de una ecuación de iteración de valores para el PDMTD con espacio de estados finito, y en dicho Teorema demostramos que esta política converge a la política óptima del PDMTC. Cabe destacar la importancia del resultado ya que esta práctica de tener trabajadores flotantes en un sistema de producción en serie, son uno de los principios de la teoría Justo a Tiempo. En este trabajo hemos mostrado que, bajo la política CL, se pueden generar más utilidades y una reducción en el trabajo en proceso.

Para futuros trabajos se podría proponer el problema de asignación de un trabajador flotante, permitiendo la interacción entre las diferentes líneas de producción (es decir, cuando los trabajos en proceso pueden ser enviados a cualquiera de las líneas de producción).

Otro cambio en el modelo sería proponer a más de un trabajador flotante. Y más aún, determinar el número óptimo de trabajadores flotantes para alguna organización.

Cabe mencionar que en la presente tesis no se consideró el hecho de que la asignación del trabajador flotante sólo se pudiera hacer cuando éste haya terminado un trabajo. Esto motiva otro problema abierto donde se propondría esta situación.

# Capítulo 6

## Notación y Terminología

### 6.1. Notación

$C(i, a)$  : Costo en el que se incurre cuando el estado es  $i$  y una acción  $a$  se elige.

$P_{ij}(a)$  : Probabilidad de transición, de ir del estado  $i$  al estado  $j$ , bajo la acción  $a$ .

$S$  : Espacio de estados.

$A$  : Espacio de acciones.

$E_f [C(X_t, A_t)]$  : Costo esperado al tiempo  $t$ , bajo la política estacionaria  $f$ .

$\nu_n(i)$  : Función de valores óptimos en horizonte finito  $n$ .

$J(i)$  : Función de valores óptimos para el costo promedio a largo plazo.

$J_\delta(i)$  : Costo esperado promedio a largo plazo en un estado inicial  $i$  y una política  $\delta$ .

$\Delta$  : Cadena de Decisión de Markov a tiempo discreto.

$\{\Delta_N\}_{N \geq N_0}$  : Sucesión de Cadenas de Decisión de Markov.

$\Delta_N$  : Cadena de Decisión de Markov con espacio de estados  $S_N$ .

$\Psi$  : Cadena de Decisión de Markov a tiempo continuo.

$E_\delta [C_n]$  : Costo total esperado incurrido bajo  $\delta$  durante los primeros  $n$  períodos de transición.

$A_i$  : Subconjunto de acciones.

$J^\Delta(\cdot)$  : El mínimo costo promedio en  $\Delta$ .

$J^\Psi(\cdot)$  : El mínimo costo promedio en  $\Psi$ .

$\mu_k$  : Parámetro de servicio en la estación  $k$ .

$H_s$  : Parámetro de costo por cada unidad de tiempo que el trabajo esté en la estación  $s$ .

$\nu(i, a)$  : Tiempo de transición de un estado a otro cuando se toma la acción  $a$ .

$I(E)$  : Función indicadora para el evento  $E$ .

$r_s$  : Carga ofertada en la estación  $s$ .

$\rho$  : Conjunto de índices de las colas.

$\lambda_i$  : Parámetro de entrada de la línea  $i$ ,  $i = 1, 2, \dots, N$ .

$(i, s)$  : Línea  $i$ , estación  $s$ ,  $i = 1, 2, \dots, N$ ,  $s = 1, 2, \dots, k$ .

$M$  : Matriz de dimensión  $N \times K$ . Donde  $N$  es el número de líneas y  $K$  el número de estaciones.

$H_{is}$  : Parámetro de costo por unidad de tiempo que el trabajo esté en la línea  $i$  y la estación  $s$ .

$g(M, a)$  : Costo en que se incurre cuando el proceso se encuentra en el estado  $M$  y una acción  $a$  se elige.

$\nu(M, a)$  : Parámetro de tiempo de transición de un estado a otro.

$C(M, a)$  : Costo en que se incurre al comenzar el siguiente período de tiempo, cuando el estado actual es  $M$  y una acción  $a$  se toma.

$\mu_{is}$  : Parámetro de servicio de la estación  $(i, s)$  donde  $i = 1, 2, \dots, N$  y  $s = 1, 2, \dots, K$ .

$r_{(i,s)}$  : Carga de la estación  $(i, s)$   $i = 1, 2, \dots, N$  y  $s = 1, 2, \dots, K$ .

$\tau$  : Período de transición del sistema uniformizado.

## 6.2. Lista de Abreviaturas

JT: Justo a tiempo

PEPS: Primeras en Entrar Primeras en Salir.

UEPS: Últimas en entrar Primeras en Salir.

CDM: Cadena de Decisión de Markov.

AS: Aproximación secuencial.

ASTA: Aproximación secuencial del tipo de Aumentación.

CDMTC: Cadena de Decisión de Markov a Tiempo Continuo.

TF: Trabajador flotante.

PDM: Proceso de Decisión de Markov.

CM: Cadena de Markov.

CL: Cola más Larga.

NM: Proceso de Nacimiento y Muerte.

# Apéndice A

## Procesos Estocásticos

En términos generales, un proceso estocástico es una colección de eventos registrados por leyes probabilísticas. Sus aplicaciones ocurren en áreas tan diversas como demografía, economía, ingeniería, psicología, y muchas otras disciplinas.

### A.1. Conceptos Generales.

**Definición 12** Sea  $T \subset \mathbb{R}$  un subconjunto no vacío. Un **proceso estocástico** es una familia  $X = \{X(t), t \in T\}$  de variables aleatorias. Es decir, para cada  $t \in T$ , en donde  $T$  es el conjunto de parámetros o conjunto de índices,  $X(t)$  es una variable aleatoria definida sobre algún espacio de probabilidad  $(\Omega, F, P)$ ;  $X(t) : \Omega \rightarrow R$ .

**Definición 13** El **espacio de estados** del proceso es el conjunto  $S$  de todos los valores posibles que pueden tomar las variables  $X(t)$  para  $t \in T$ . Cada punto en  $S$  se llama un estado del proceso. Si  $S$  es finito o numerable decimos que el proceso es **discreto**; si  $S$  es un intervalo, decimos que el proceso es **continuo**.

El proceso también se clasifica de acuerdo a la naturaleza del conjunto de parámetros  $T$ . Si  $T$  es numerable, por ejemplo,  $T = \{0, 1, 2, \dots\}$ ,  $T = \{0, \pm 1, \pm 2, \dots\}$ , etc., decimos que  $X$  es un proceso con **parámetro** (o en tiempo) **discreto**. Si  $T$  es un intervalo, e.g.,  $T = [a, b]$ ,  $T = [0, \infty]$ ,  $T = (-\infty, \infty)$ , etc.,  $X$  es un proceso con parámetro (o en tiempo) **continuo**.

Si  $X = \{X(t), t \in T\}$  es un proceso en tiempo discreto, digamos  $T = \{0, 1, 2, \dots\}$ , escribimos  $X(t) = X_t$ , y  $X$  no es otra cosa que una sucesión de variables aleatorias  $X = \{X_0, X_1, \dots\}$ .

**Definición 14** Un **proceso de conteo** es un proceso  $\{N(t), t \geq 0\}$ , con espacio de estados  $S = \{0, 1, \dots\}$  y tal que para cada  $t \geq 0$ ,  $N(t)$  cuenta el número de veces

que ha ocurrido cierto evento  $E_*$  durante el intervalo  $(0, t]$ . Es decir,  $\{N(t), t \geq 0\}$  es un proceso de conteo si satisface que:

1.  $N(0)=0$ ,
2.  $N(t)$  sólo toma valores enteros no negativos,
3.  $N(t)$  es no decreciente:  $N(s) \leq N(t)$  si  $s < t$ ,
4.  $N(t) - N(s)$  es el número de veces que  $E_*$  ha ocurrido durante el intervalo  $(s, t]$ .

**Definición 15** *Considérese el proceso estocástico  $X = \{X(t), t \geq 0\}$ . Decimos que:*

1.  $X$  es un proceso con **incrementos independientes** si para cualquier selección de índices  $0 \leq t_0 < t_1 < \dots < t_n$ , se satisface que las variables aleatorias  $X(t_1) - X(t_0)$ ,  $X(t_2) - X(t_1)$ , ...,  $X(t_n) - X(t_{n-1})$  son independientes.
2.  $X$  es un proceso con **incrementos estacionarios** si la distribución del incremento  $X(t+h) - X(t)$  depende sólo de  $h$ , es decir,  $X(t+h) - X(t) \sim X(\tau+h) - X(\tau)$  para todo  $t, \tau \geq 0$  y  $h > 0$ .

## A.2. Proceso de Poisson.

Comenzaremos con recordar la definición y algunas propiedades básicas de la distribución exponencial.

**Definición 16** *Una variable aleatoria  $T$  tiene una **distribución exponencial** con parámetro  $\lambda$  ó  $T \sim \text{exp}(\lambda)$  si:*

$$P(T \leq t) = 1 - e^{-\lambda t} \quad \forall t \geq 0$$

*Describimos la distribución dada por la función de distribución:  $F(t) = P(T \leq t)$ ,  $t \in \mathbb{R}$ . Se define además la función de densidad  $f_T(t)$ ,  $t \in \mathbb{R}$ , la cual es la derivada de la función de distribución:*

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{para } t \geq 0 \\ 0 & \text{para } t < 0 \end{cases}$$

además:

$$E(T) = \frac{1}{\lambda} \text{ y } Var(T) = E(T^2) - (ET)^2 = \frac{1}{\lambda^2}.$$

A continuación se presenta una importante propiedad de la distribución exponencial, conocida como **Pérdida de Memoria**.

**Proposición 2** Sea  $T$  una variable aleatoria con distribución exponencial, con parámetro  $\mu$ , entonces  $P(T > t_1 + t_2 \mid T > t_2) = P(T > t_1)$ , para todo  $t_1$  y  $t_2$  números positivos fijos.

**Demostración.** Sean  $t_1$  y  $t_2$  dos números positivos fijos. Entonces,

$$\begin{aligned} P(T > t_1 + t_2 \mid T > t_2) &= \frac{P(T > t_1 + t_2, T > t_2)}{P(T > t_2)} \\ &= \frac{P(T > t_1 + t_2)}{P(T > t_2)} = \frac{e^{-\mu(t_1+t_2)}}{e^{-\mu t_2}} \\ &= e^{-\mu t_1} = P(T > t_1). \end{aligned}$$

La primera igualdad se cumple, por la definición de probabilidad condicional y la tercera es simple aplicación de la distribución exponencial. Por lo tanto  $P(T > t_1 + t_2 \mid T > t_2) = P(T > t_1)$ . ■

El resultado anterior nos dice que la probabilidad de ocurrencia de eventos presentes ó futuros no depende de los que hayan ocurrido en el pasado.

**Proposición 3** Sean  $T_1$  y  $T_2$  variables aleatorias independientes, distribuidas exponencialmente, con parámetros  $\mu_1$  y  $\mu_2$ , respectivamente. Entonces, se tienen las siguientes afirmaciones:

i  $P(T_1 < T_2) = \frac{\mu_1}{\mu_1 + \mu_2}$

ii Además sea  $Y = \min(T_1, T_2)$ . Entonces  $Y$  tiene una distribución  $\exp(\mu_1 + \mu_2)$ .

**Demostración.** (i) Sean  $T_1$  y  $T_2$  variables aleatorias independientes distribuídas exponencialmente y sea  $t$  número positivo fijo. Aplicando la ley de probabilidad total se tiene:

$$\begin{aligned} P(T_1 < T_2) &= \int_0^{\infty} P(T_1 < T_2 \mid T_1 = t) f_{T_1}(t) dt \\ &= \int_0^{\infty} P(T_2 > t) f_{T_1}(t) dt \\ &= \int_0^{\infty} e^{-\mu_2 t} (\mu_1 e^{-\mu_1 t}) dt = \frac{\mu_1}{\mu_1 + \mu_2}, \end{aligned}$$

donde  $f_T$ , es la densidad de la variable aleatoria  $T_1$ .

(ii) Sean  $T_1$  y  $T_2$ , variables aleatorias independientes distribuídas exponencialmente. Sea  $t$  número positivo fijo y  $Y = \min(T_1, T_2)$ . Entonces, para  $t > 0$  :

$$\begin{aligned} P(Y > t) &= P(T_1 > t, T_2 > t) \\ &= P(T_1 > t) P(T_2 > t) \\ &= e^{-\mu_1 t} e^{-\mu_2 t} = e^{-(\mu_1 + \mu_2)t} \end{aligned}$$

La segunda desigualdad se debe a la independencia de  $T_1$  y  $T_2$ . ■

A continuación se muestra el caso general para el inciso (ii) de la proposición anterior.

**Proposición 4** Sean  $T_1, \dots, T_n$ , variables aleatorias independientes distribuídas exponencialmente con parámetro  $\lambda_i$ , donde  $i = 1, 2, \dots, n$ , y  $t$  es un número positivo fijo. Entonces

$$P(\min(T_1, \dots, T_n) > t) = e^{-(\lambda_1 + \dots + \lambda_n)t}.$$

**Demostración.** Sean  $T_1, \dots, T_n$  variables aleatorias distribuídas exponencialmente con parámetro  $\lambda_i$ , donde  $i = 1, 2, \dots, n$ , y  $t$  número positivo fijo entonces:

$$\begin{aligned} P(\min(T_1, \dots, T_n) > t) &= P(T_1 > t, \dots, T_n > t) \\ &= P(T_1 > t)P(T_2 > t)\dots P(T_n > t) \\ &= \prod_{i=1}^n P(T_i > t) = \prod_{i=1}^n e^{-\lambda_i t} = e^{-(\lambda_1 + \dots + \lambda_n)t}. \end{aligned}$$

La segunda igualdad se debe a la independencia de las  $T_i$ ,  $i = 1, 2, \dots, n$  y la tercera se sigue de la definición de la distribución exponencial. ■

El proceso de Poisson es un ejemplo de un proceso de conteo, el cual, como veremos se puede definir en varias formas equivalentes.

**Definición 17** Un proceso de conteo  $X = \{N(t), t \geq 0\}$  es un **proceso de Poisson** con intensidad media  $\lambda$  ( $\lambda > 0$ ) si  $P\{X(t+h) - X(t) = k\} = e^{-\lambda h} (\lambda h)^k / k!$ ,  $k = 0, 1, \dots$ . De aquí se sigue, en particular, que

$$E[X(t+h) - X(t)] = \lambda h \text{ y } \text{var}(X(t+h) - X(t)) = \lambda h, t \geq 0, h > 0.$$

De la primera ecuación vemos que  $\lambda = \frac{1}{h} E[X(t+h) - X(t)]$  lo cual justifica el nombre de intensidad (o tasa) media para  $\lambda$ ; es decir,  $\lambda$  se puede interpretar como el número promedio de veces que ocurre el evento de interés, por unidad de tiempo.

**Notación 1** La expresión ' $f = o(h)$  cuando  $h \rightarrow 0$ ', significa que  $f(\cdot)$  es una función tal que  $\lim_{h \rightarrow 0} (f(h)/h) = 0$ ; es decir, para cualquier  $\varepsilon > 0$ , existe  $\delta > 0$  tal que  $|f(h)/h| < \varepsilon$  para todo  $0 < |h| < \delta$ .

Ahora veremos una caracterización equivalente de un proceso de Poisson.

**Teorema 6**  $X = \{X(t), t \geq 0\}$  es un proceso de Poisson con intensidad  $\lambda(> 0)$  si y sólo si  $X$  es un proceso de conteo tal que:

1.  $X$  tiene incrementos independientes y estacionarios;
2.  $P\{X(t) = 1\} = \lambda t + o(t)$  cuando  $t \rightarrow 0+$ ;
3.  $P\{X(t) \geq 2\} = o(t)$  cuando  $t \rightarrow 0+$ .

Nótese que 1 y 3 implican que

$$P\{X(t) = 0\} = 1 - P\{X(t) \geq 1\} = 1 - \lambda t + o(t) \text{ cuando } t \rightarrow 0+.$$

Otra forma de definir el Proceso de Poisson es la siguiente:

**Definición 18** Sean  $\eta_1, \eta_2, \dots$  variables aleatorias independientes distribuidas exponencialmente con parámetro  $\lambda$ . Sea  $T_n = \eta_1 + \eta_2 + \dots + \eta_n$  para  $n \geq 1$ ,  $n \in \mathbb{N}$ ,  $T_0 = 0$ , y definimos  $N(s) = \max\{n \in \mathbb{N} : T_n \leq s\}$ ,  $s > 0$ .  $N(s)$   $s > 0$  se llama **proceso de Poisson**.

**Teorema 7** El proceso  $N(s)$ ,  $s > 0$ , definido anteriormente tiene una distribución de Poisson con parámetro  $\lambda s$ .

**Demostración.** Sea  $s > 0$  y  $n \in \mathbb{N}$ . Notemos que  $N(s) = n$  si y sólo si  $T_n \leq s < T_{n+1}$ . Sea  $t$  número positivo fijo. Tomando:  $T_n = t$  y notando que:  $s < T_{n+1}$  se tiene entonces que:  $t_{n+1} > s - t$  y  $t_{n+1}$  es independiente de  $T_n$ , así entonces:

$$P(N(s) = n) = P(T_n \leq s).$$

Aplicando la ley de probabilidad total, se tiene, para  $t$  y  $s$  positivos:

$$\begin{aligned} &= \int_0^s P(T_n = t)P(T_{n+1} > s | T_n = t)dt \\ &= \int_0^s P(T_n = t)P(T_{n+1} > s - t)dt. \end{aligned}$$

Sustituyendo la distribución de Poisson, se tiene:

$$\begin{aligned} &= \int_0^s \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} e^{-\lambda(s-t)} dt \\ &= \frac{\lambda^n}{(n-1)!} e^{-\lambda s} \int_0^s t^{n-1} dt = e^{-\lambda s} \frac{(\lambda s)^n}{n!}. \end{aligned}$$

La cual es una distribución de Poisson con parámetro  $\lambda s$ . ■

### A.3. Proceso de Markov

**Definición 19** *Un proceso estocástico  $X = \{X(t), t \in T\}$ , con  $T \subset \mathbb{R}$ , es un proceso de Markov si para cualquier  $x \in \mathbb{R}$  y  $s, t \in T$ , con  $s \leq t$ , se satisface que;*

$$P(X(t) \leq x \mid X(r), r \leq s) = P(X(t) \leq x \mid X(s)) \tag{A.1}$$

*Esta ecuación se conoce como la propiedad de Markov.*

Podemos ver que interpretando  $s$  como el tiempo ‘presente’,  $t > s$  como el ‘futuro’ y  $r < s$  como el ‘pasado’, la ecuación (A.1) nos dice que dado el presente  $X(s)$  del sistema, el futuro  $X(t)$ ,  $t \geq s$ , es independiente del pasado  $X(r)$ ,  $r < s$ .

Los procesos de Markov se pueden clasificar de acuerdo al conjunto  $T$  de parámetros y al espacio  $S$  de estados, como sigue:

Parámetro ( $T$ )	Espacio de estados $S$	
	Discreto	Continuo
Discreto	Cadena de Markov en tiempo discreto	Proceso de Markov en tiempo discreto
Continuo	Cadena de Markov en tiempo continuo	Proceso de Markov en tiempo continuo

#### A.3.1. Proceso de Markov en Tiempo Continuo

Por el momento nos restringiremos a estudiar exclusivamente cadenas de Markov con espacio de estados  $S = \{0, 1, 2, \dots\}$  y en tiempo continuo, con conjunto de parámetros  $T = [0, \infty)$ .

**Probabilidad de transición.**

Sea  $X = \{X(t), t \geq 0\}$  una cadena de Markov en tiempo continuo. Como espacio de estados podríamos tomar cualquier conjunto numerable  $S = \{x_0, x_1, x_2, \dots\}$ , pero para simplificar un poco la notación, supondremos que  $S = \{0, 1, 2, \dots\}$ .

Si  $i, j \in S$  y  $0 \leq s < t$ , definimos la probabilidad de transición  $p_{ij}(s, t) = P\{X(t) = j \mid X(s) = i\}$ , cuya interpretación es obvia. De hecho, sólo consideramos procesos homogéneos en el tiempo, es decir, procesos para los cuales, para  $i$  y  $j$  fijos,

la probabilidad de transición  $p_{ij}(s, t)$  sólo depende de  $t - s$ , o sea,  $p_{ij}(t, s) = p_{ij}(t - s)$ . En este caso, escribimos

$$p_{ij}(t) = P \{X(t) = j \mid X(0) = i\} \quad (\text{A.2})$$

$$= P \{X(s + t) = j \mid X(s) = i\}, \quad (\text{A.3})$$

para todo  $s \geq 0$ ,  $i, j \in S$ .

Estas probabilidades de transición satisfacen que  $p_{ij}(t) \geq 0$  para todo  $t \geq 0$ ,  $i, j \in S$ , y

$$\sum_{j=0}^{\infty} p_{ij}(t) = 1 \text{ para todo } t \geq 0 \text{ y } i \in S \text{ fijos;} \quad (\text{A.4})$$

$$p_{ij}(t + s) = \sum_{k=0}^{\infty} P_{ik}(s) p_{kj}(t) \text{ para todo } t, s \geq 0, i, j \in S. \quad (\text{A.5})$$

La ecuación (A.5) se llama ecuación de

#### Distribuciones estacionarias.

Recuérdese que una sucesión  $\{\pi_k, k = 0, 1, \dots\}$  es una distribución de probabilidad si  $\pi_k \geq 0$  para todo  $k$ , y  $\sum_{k=0}^{\infty} \pi_k = 1$ .

**Definición 20** Decimos que una cadena de Markov  $\{X(t), t \geq 0\}$  con espacio de estados  $S = \{0, 1, \dots\}$  tiene una distribución estacionaria si existe una distribución de probabilidad  $\{\pi_k, k = 0, 1, \dots\}$  tal que para todo  $i$  y  $k \in S$ ,

$$\lim_{t \rightarrow \infty} p_{ik}(t) = \pi_k. \quad (\text{A.6})$$

**Definición 21** Dos estados  $j$  y  $k \in S$  se **comunican** si existen tiempos  $t_1$  y  $t_2$ , tales que  $p_{jk}(t_1) > 0$ .

**Definición 22** Decimos que la cadena de Markov es **irreducible** si cualesquiera dos estados se comunican.

Se puede demostrar (véase, por ejemplo: [2]) que en una cadena de Markov irreducible los límites  $\pi_k$  en (A.6) siempre existen y son independientes del estado inicial de la cadena. Además, los límites  $\{\pi_k, k = 0, 1, \dots\}$  satisfacen que, o se anulan idénticamente,

$$\pi_k = 0 \text{ para todo } k = 0, 1, \dots,$$

o bien, son todos positivos y forman una distribución de probabilidad,

$$\pi_k > 0 \text{ para todo } k, \text{ y } \sum_{k=0}^{\infty} \pi_k = 1.$$

Si se cumple esto último, decimos entonces que la cadena irreducible es recurrente positiva.

### A.3.2. Cadena de Markov a Tiempo Discreto

En particular, si  $X$  es un proceso de Markov en el que tanto  $S$  como  $T$  son conjuntos numerables, digamos:

$$S = \{x_0, x_1, x_2, \dots\}, T = \{t_0, t_1, t_2, \dots\},$$

entonces  $X$  se llama una cadena de Markov en tiempo discreto. En tal caso, para simplificar la notación, podemos suponer que  $S$  es el conjunto de los enteros no negativos,  $S = \{0, 1, 2, \dots\}$ , y escribiendo  $X_n = X(t_n)$ , podemos suponer que también  $T = \{0, 1, 2, \dots\}$ . Con esta notación, podemos escribir la cadena de Markov  $X = \{X(t), t \in T\}$  como una sucesión:  $X = \{X_n, n = 0, 1, 2, \dots\}$ .

#### Probabilidades de transición

Para una cadena de Markov  $\{X_n\}$  en tiempo discreto, las probabilidades de transición:

$$p_{ij}(s, t) = P \{X(t) = j \mid X(s) = i\}$$

definidas anteriormente, se sustituyen por probabilidades de transición en un paso

$$p_{ij}(n, n+1) = P \{X_{n+1} = j \mid X_n = i\}, \quad n, i, j = 0, 1, 2, \dots$$

Las probabilidades de transición satisfacen (A.4)

1.  $p_{ij} \geq 0$  para todo  $i, j$ ;
2.  $\sum_{j=0}^{\infty} p_{ij} = 1, i = 0, 1, \dots$ ;
3.  $p_{ij}^{(n+m)} = \sum_{k=0}^{\infty} p_{ik}^n p_{kj}^m$  para todo  $n, m, i, j \geq 0$ ,

La última ecuación se llama **Ecuación de Chapman-Kolmogorov**.

#### Cadenas de Markov irreducibles

**Definición 23** Sean  $i, j$  estados de una cadena  $\{X_n\}$ . Decimos que  $j$  es alcanzable desde el estado  $i$  si  $p_{ij}^{(n)} > 0$  para algún  $n \geq 0$ ; es decir, con probabilidad positiva se puede ir de  $i$  a  $j$  en algún número finito ( $n$ ) de pasos.

**Definición 24** Si  $j$  es alcanzable desde  $i$  ( $i \rightarrow j$ ) y  $i$  es alcanzable desde  $j$  ( $j \rightarrow i$ ), decimos que  $i$  y  $j$  son estados **comunicantes** y escribimos  $i \leftrightarrow j$ . Dicho en otras palabras,  $i$  y  $j$  son comunicantes si existen enteros no negativos  $n$  y  $m$  tales que  $p_{ij}^{(n)} > 0$  y  $p_{ji}^{(m)} > 0$ .

La relación  $i \leftrightarrow j$  es una relación de equivalencia.

Un estado  $i$  es **absorbente** si  $p_{ii} = 1$ . Es obvio que si  $i$  es un estado absorbente, entonces  $\{i\}$  es una clase comunicante.

Una cadena de Markov cuyo espacio de estados consiste de una, y sólo una, clase comunicante se llama una cadena irreducible.

**Definición 25** Definimos el **período**  $d(i)$  de un estado  $i$  como el máximo común divisor de todos los enteros  $n \geq 1$  para los cuales  $p_{ii}^{(n)} > 0$ . Si  $p_{ii}^{(n)} = 0$  para todo  $n \geq 1$ , definimos  $d(i) = 0$ . Si  $d(i) > 1$ , decimos que  $i$  es un estado periódico con período  $d(i)$ , mientras que si el período es 1,  $d(i) = 1$ , decimos que  $i$  es un estado aperiódico.

### Recurrencia y Distribuciones Estacionarias.

Sean  $i$  y  $j$  estados de una cadena de Markov  $X_n$ . Definimos  $f_{ij}^n$  como la probabilidad de que empezando en  $i$  ( $x_0 = i$ ) la primera transición al estado  $j$  ocurre en  $n$  pasos exactamente. Es decir,

$$f_{ij}^n = P \{X_n = j, X_k \neq j \text{ para } k = 1, 2, \dots, n-1 \mid X_0 = i\}.$$

Por supuesto,  $f_{ij}^1 = p_{ij}$ , y se puede ver que la probabilidad  $p_{ij}^{(n)}$  de ir de  $i$  a  $j$  en  $n$  pasos satisface

$$p_{ij}^{(n)} = \sum_{k=0}^{n-1} f_{ij}^k p_{jj}^{(n-k)}, \quad (f_{ij}^0 = \delta_{ij}). \quad (\text{A.7})$$

Ahora, definamos  $f_{ij}$  como la probabilidad de que la cadena haga una transición al estado  $j$  dado que  $x_0 = i$ . (Nótese que para  $i \neq j$ ,  $f_{ij} > 0$  si y sólo si  $i \rightarrow j$ .) Es decir,

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n.$$

Decimos que  $i$  es un estado recurrente si  $f_{ii} = 1$ , y que  $i$  es transitorio si  $f_{ii} < 1$ .

**Proposición 5**  $i$  es recurrente si y sólo si  $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$ .

Si  $j$  es un estado transitorio, el número esperado de transiciones de  $i$  a  $j$  para cualquier estado inicial  $x_0 = i$ , es finito, o sea:

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty \text{ para todo } i.$$

De aquí sigue que:

Si  $j$  es transitorio,  $p_{ij}^{(n)} \rightarrow 0$  cuando  $n \rightarrow \infty$  para todo  $i$ .

Definamos

$$\mu_{ij} = \begin{cases} \infty & \text{si } j \text{ es transitorio,} \\ \sum_{n=1}^{\infty} n f_{ij}^n & \text{si } j \text{ es recurrente.} \end{cases}$$

Se llama el **tiempo medio de recurrencia** de  $j$ .

**Teorema 8** Si  $i \leftrightarrow j$ , entonces:

1.  $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} = 1/\mu_{jj}$ .
2.  $j$  aperiódico implica que  $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 1/\mu_{jj}$ .
3.  $j$  periódico con período  $d$  implica que  $\lim_{n \rightarrow \infty} p_{ij}^{(nd)} = \frac{d}{\mu_{jj}}$ .

Si  $j$  es un estado recurrente, decimos que  $j$  es recurrente positivo si  $\mu_{jj} < \infty$ , y que es recurrente nulo si  $\mu_{jj} = \infty$ . Si, además,  $d = d(j) \geq 1$  es el período de  $j$ , y definimos:  $\pi_j = \lim_{n \rightarrow \infty} p_{jj}^{(nd)}$ , entonces,  $j$  es recurrente positivo si  $\pi_j > 0$ , y recurrente nulo si  $\pi_j = 0$ .

**Definición 26** Un estado aperiódico y recurrente positivo se dice que es **ergódico**. En el caso de que sea recurrente nulo se llama **débilmente ergódico**.

**Proposición 6** En una misma clase comunicante se satisface una, y sólo una, de las siguientes proposiciones:

1. Todos los estados son recurrentes positivos,
2. Todos los estados son recurrentes nulos.

**Definición 27** Una CM es **recurrente (transitoria, positiva recurrente, recurrente nula)** si todos sus estados son recurrentes (transitorios, positivos recurrentes, recurrentes nulos).

**Teorema 9** Una cadena de Markov aperiódica e irreducible pertenece a una, y sólo una, de las dos siguientes categorías:

1. Todos los estados son transitorios o todos son recurrentes nulos; en este caso,  $p_{ij}^{(n)} \rightarrow 0$  cuando  $n \rightarrow \infty$  para todo  $i$  y  $j$ , y no existe una distribución estacionaria;
2. Todos los estados son recurrentes positivos, es decir,

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^n > 0.$$

**Teorema 10** Sea  $X = \{X_n\}$  una cadena de Markov irreducible con espacio de estados  $S = \{0, 1, 2, \dots\}$ . Entonces:

1.  $X$  es recurrente si y sólo si  $f_{i0} = \sum_{n=1}^{\infty} f_{i0}^n = 1$  para todo  $i \neq 0$ .
2.  $X$  es transitoria (es decir, todos los estados son transitorios) si y sólo si el sistema de ecuaciones  $x_i = \sum_{j=1}^{\infty} p_{ij} x_j$ ,  $i = 1, 2, \dots$  tiene una solución acotada distinta de cero.
3. Supóngase que, además de ser irreducible, la cadena  $X$  es aperiódica. Entonces  $X$  es recurrente si existe una solución de las desigualdades.  $x_i \geq \sum_{j=1}^{\infty} p_{ij} x_j$ ,  $i = 1, 2, \dots$  con la propiedad de que  $x_i \rightarrow \infty$  cuando  $i \rightarrow \infty$ .

## A.4. Colas Poissonianas

Muchos ejemplos importantes de sistemas de espera se pueden describir como procesos de nacimiento y muerte (NM) (véase [6]). Sea  $\{N(t), t \geq 0\}$  un proceso NM, en donde  $N(t)$  es el número de clientes en el sistema. La llegada de un cliente se interpreta como un nacimiento y la salida de un cliente después de completar su servicio se interpreta como una muerte. Distintos tipos de sistemas de espera se pueden obtener seleccionando adecuadamente los parámetros de nacimiento  $\lambda_n$  y muerte  $\mu_n$ .

### A.4.1. El Sistema M/M/1

El sistema de espera M/M/1 es un sistema con un servidor, en el que tanto la distribución  $A(t)$  de tiempos entre arribos como la distribución  $B(t)$  del tiempo de servicio son ambas exponenciales. Es decir, los tiempos entre arribos  $\tau_1, \tau_2, \dots$  son variables aleatorias independientes idénticamente distribuídas, con distribución exponencial con parámetro  $\lambda$ :

$$A(t) = P\{\tau \leq t\} = \lambda e^{-\lambda t}, t \geq 0.$$

El proceso NM con parámetros ( $\lambda_n = \lambda$  para  $n = 0, 1, 2, \dots$ , y  $\mu_n = \mu$  para  $n = 1, 2, \dots$ ), tiene una distribución estacionaria  $\{\pi_n, n = 0, 1, \dots\}$  si y sólo si la intensidad de tráfico o carga de la cola  $r = \lambda/\mu < 1$ , y en tal caso,  $\pi_n$  es una distribución geométrica

$$\pi_n = (1 - r)r^n, n = 0, 1, \dots \quad (\text{A.8})$$

En estado estacionario, el número esperado de clientes en el sistema (es decir, la longitud de la cola, incluyendo los clientes recibiendo servicio) es:

$$L \equiv EN = \sum_{n=0}^{\infty} n\pi_n = (1-r)r \sum_{n=1}^{\infty} nr^{n-1},$$

o sea,

$$L = EN = r/(1-r), \quad r = \lambda/\mu < 1. \quad (\text{A.9})$$

La varianza de  $N$  está dada por:

$$\text{var}(N) = E(N^2) - (EN)^2 = r/(1-r)^2. \quad (\text{A.10})$$

Ahora en estado estacionario, se tienen las siguientes:

$W$  = Tiempo esperado en el sistema;

$W_q$  = Tiempo esperado en la cola, sin incluir el tiempo de servicio;

$L_q$  = Número esperado de clientes en la cola, sin incluir los que están en servicio.

En base a lo anterior se tiene el siguiente resultado, demostrado por Little en [8]:

$$L = \lambda W. \quad (\text{A.11})$$

La demostración de esta fórmula es un poco complicada y no la haremos (véase [8]).

**Definición 28** *Una estación de servicio con capacidad ilimitada se llama **estable** si:*

$$\sum_{j=0}^{\infty} p_j = 1, \quad (\text{A.12})$$

donde  $p_j = \lim_{t \rightarrow \infty} P(X(t) = j)$  y  $X(t)$  es el número de clientes en la línea de espera en el tiempo  $t$ .

**Teorema 11 (Condición de Estabilidad)** *Considere la línea de espera de una estación de servicio con  $s$  servidores y capacidad infinita. Suponga que el cliente llega con parámetro  $\lambda$  y con media de tiempo de servicio  $\tau$  entonces la línea de espera es estable si:*

$$\lambda\tau < s. \quad (\text{A.13})$$

Del Teorema se sigue que la línea de espera es estable si  $r < 1$  y es inestable si  $r \geq 1$ .

# Apéndice B

## Complementos a las Demostraciones

A continuación trataremos el criterio de costo promedio cuando la CDM tiene espacio de estados finito  $S$ . Además daremos algunos resultados que nos ayudan a computar la ecuación de optimalidad costo promedio.

**Proposición 7** *Sea  $\Delta$  una CDM con espacio de estados finitos  $S$ .*

1. Supongamos que se tiene una constante (finita)  $F$  y una función (finita)  $r$  tal que

$$F + r(i) \geq \left\{ C(i, a) + \sum_j P_{ij}(a) r(j) \right\}, \quad i \in S. \quad (\text{B.1})$$

Si  $e$  es una política estacionaria realizando el mínimo en (B.1), entonces  $J_e(i) \leq F$  para  $i \in S$ . Por lo tanto, si  $F$  está acotada por abajo por el costo promedio, entonces el costo promedio es igual a la constante  $F$ , y  $e$  es costo promedio óptima.

2. Supongamos que tenemos una constante (finita)  $F$  y una función  $r$  tal que

$$F + r(i) = \min_a \left\{ C(i, a) + \sum_j P_{ij}(a) r(j) \right\}, \quad i \in S. \quad (\text{B.2})$$

Entonces el mínimo costo promedio es igual a la constante  $F$ , y cualquier política estacionaria realizando el mínimo en B.2 es costo promedio óptimo.

Si  $h$  es como en el Teorema 6.4.2 de [13], entonces  $r(i) = h(i) + r(z_k)$  para  $i \in R_k$ . (Recordemos que éste último es una clase recurrente bajo  $f$  con espacio de estados distinguidos  $z_k$ ) Para  $i$  transitorio bajo  $f$  se cumple que

$$r(i) \leq h(i) + \sum_k P_k(i) [r(z_k) - h(z_k)].$$

Ahora daremos la siguiente hipótesis:

**Hipótesis 5 (OPA)** *Sea  $e$  una política estacionaria óptima. Entonces cualquier clase recurrente positiva en la CM inducida por  $e$  es aperiódica.*

**Proposición 8** *Supongamos que el costo promedio mínimo es una constante  $J$  y la hipótesis OPA se cumple. Sea  $x$  un estado distinguido en  $S$ . Entonces el  $\lim_{n \rightarrow \infty} [\nu_n(x) - \nu_{n-1}(x)] = J$  y el  $\lim_{n \rightarrow \infty} f_n(i) =: r(i)$  existe. Por lo tanto la ecuación de optimalidad en horizonte finito dada por*

$$[\nu_n(x) - \nu_{n-1}(x)] + r_n(i) = \min_a \left\{ C(i, a) + \sum_j P_{ij}(a) r_{n-1}(j) \right\} \quad n \geq 1, i \in S. \quad (\text{B.3})$$

*Puede ser usada para computar una solución para la ecuación de optimalidad costo promedio (B.2). Cualquier punto límite de la política estacionaria óptima en horizonte finito  $f'_n$  realiza el mínimo en (B.2), y por lo tanto es costo promedio óptima.*

A continuación daremos el algoritmo de iteración de valores para obtener la solución de (B.2).

## B.1. Algoritmo de Iteración de Valores (AIV)

Sea  $\Delta$  una CDM con espacio de estados finito. Supongamos que el mínimo costo promedio es una constante y que la hipótesis OPA se cumple. Sea  $x$  un estado distinguido y  $\varepsilon$  un número pequeño positivo.

### AIV Versión 1:

1. Sea  $n = 0$  y  $u_0 \equiv 0$ ,

2. Sea  $w_n(i) = \min_a \left\{ e(i, a) + \sum_j P_{ij}(a) u_n(i) \right\}$ ,

3. Si  $n = 0$ , sea  $\delta = 1$ . Si  $n \geq 1$ , entonces sea  $\delta = |w_n(x) - w_{n-1}(x)|$ . Si  $\delta < \varepsilon$  se continua con el paso 6.
4. Sea  $u_{n+1}(i) = w_n(i) - w_n(x)$ .
5. Ir al paso 2 y reemplazar  $n$  por  $n + 1$ .

Imprimir  $w_n(x)$  y una política estacionaria realiza el

$$\min_a \left\{ C(i, a) + \sum_j P_{ij}(a) u_n(j) \right\}.$$

**AIV Versión 2:**

1. La versión 1, pero sea  $\delta = \max_{i \in S} |w_n(i) - w_{n-1}(i)|$  en el paso 3.

**Lema 3** *Sea  $e$  una política estacionaria. Sea  $S$  numerable. Supongamos que existe una constante (finita)  $J$  y una función (finita)  $h$  acotada por abajo en  $i$  tal que*

$$J + h(i) \geq C(i, e) + \sum_j P_{ij}(e)h(j), \quad i \in S. \tag{B.4}$$

Entonces  $J_e(i) \leq J$  para  $i \in S$ .

**Demostración.** Por hipótesis, existe una constante (finita) no negativa  $L$  tal que  $h(i) \geq -L$  para  $i \in S$ . Sea  $X_0 = i, X_1, X_2, \dots$  una sucesión de valores del proceso operando bajo una política  $e$  y suprimido el estado inicial en el que sigue. Entonces de (B.4) se sigue que

$$J + h(X_t) \geq C(X_t, e) + E_e [h(X_{t+1}) | X_t], \quad t \geq 0. \tag{B.5}$$

Es claro que  $E_e [h(X_t)] < \infty$ , y para mostrar esto, probando por inducción en  $t$  que  $E_e [h(X_t)] \leq tJ + h(i)$ . Esto es claro para  $t = 0$ . Supongamos que es cierto para  $t$ . Entonces de (B.5) se sigue que

$$E_e [h(X_{t+1}) | X_t] \leq J + h(X_t).$$

Tomando el valor esperado en ambos lados y usando la propiedad del valor esperado (es decir, que  $E(E(X | Y)) = E(X)$ ), encontramos que

$$E_e [h(X_{t+1})] \leq J + E_e(h(X_t)) \leq J + tJ + h(i) = (t + 1)J + h(i).$$

La segunda desigualdad se sigue de la hipótesis de inducción. Esto completa la inducción. Ahora tomemos el valor esperado en ambos lados de (B.5) y obtenemos

$$E_e [C(X_t, f)] \leq J + E_e [h(X_t)] - E_e [h(X_{t+1})], \quad t \geq 0. \quad (\text{B.6})$$

Lo que se ha demostrado nos asegura que no tenemos la forma indeterminada  $\infty - \infty$ . Sumando los términos en (B.6), para  $t = 0$  hasta  $n - 1$ , y dividiendo por  $n$  se tiene

$$\frac{\nu_{e,n}(i)}{n} \leq J + \frac{h(i) - E_e [h(X_n)]}{n} \leq J + \frac{h(i) + L}{n}. \quad (\text{B.7})$$

Tomando límite superior en ambos lados de (B.7) se tiene el resultado. ■

## **B.2. Computo de Políticas Costo Promedio Óptimas para Espacio de Estados Infinitos**

### **B.2.1. Suposiciones (AC)**

Sea  $\Delta$  una CDM con espacio de estados numerable. El objetivo es computar una política estacionaria costo promedio óptima. Esto se obtiene computando una política estacionaria óptima en una sucesión de aproximación y mostrando que cualquier punto límite de éstas políticas es costo promedio para  $\Delta$ .

A continuación damos un conjunto de suposiciones.

Supongamos que se tiene una aproximación secuencial  $\{\Delta_N\}_{N \geq N_0}$  para  $\Delta$ , para  $N_0$  fijo. Las suposiciones (AC) son las siguientes.

AC1 Existe una constante  $J^N$  (finita) y una función  $r^N$  (finita) en  $S_N$  tal que

$$J^N + r^N(i) = \min_a \left\{ C(i, a) + \sum_{j \in S_N} P_{ij}(a, N) r^N(j) \right\}, \quad i \in S_N, \quad N \geq N_0. \quad (\text{B.8})$$

AC2 Se tiene que  $\limsup_{N \rightarrow \infty} r^N(i) < \infty$  para  $i \in S$ .

AC3 Existe una constante (finita) no negativa  $Q$  tal que  $-Q \leq \liminf_{N \rightarrow \infty} r^N(i)$  para  $i \in S$ .

AC4 Se tiene que el  $\limsup_{N \rightarrow \infty} J^N = J^* < \infty$  y  $J^* \leq J(i)$  para  $i \in S$ .

Ahora damos un resultado importante que utiliza el lema 3.

**Teorema 12** *Supongamos que las suposiciones (AC) se cumplen. Entonces:*

1. La cantidad  $J^* = \lim_{N \rightarrow \infty} J^N$  es el mínimo costo promedio en  $\Delta$ .
2. Cualquier punto límite  $e^*$  de una sucesión  $e^N$  de políticas estacionarias realizando el mínimo en (B.8) es costo promedio óptima para  $\Delta$ .

**Proposición 9** *Sea  $\{\Delta_N\}_{N \geq N_0}$  una sucesión de aproximación secuencial para  $\Delta$ , y sea  $x$  un estado distinguido (podemos asumir que  $x \in S_N$  para todo  $N$ ). Bajo los siguientes 4 pasos se justifica el uso del algoritmo de iteración de valores en  $\Delta_N$  y verifica que las suposiciones (AC) se cumplen para la función*

$$r^N(\cdot) = \lim_{n \rightarrow \infty} (r_n^N(\cdot) - r_n^N(x)).$$

*Paso 1. Demostrar que cualquier política estacionaria para  $\Delta_N$  induce una unicadena MC con clase recurrente positiva aperiodica conteniendo  $x$ .*

*Paso 2. Demostrar que existe una  $x$  estándar política  $d$  para  $\Delta$  tal que la AS es conforme en  $d$ , (véase A.C.4 de [13]).*

*Paso 3. Realizar uno de los siguientes:*

- a) *Demostrar que  $\nu_n^N(i) \leq \nu_n(i)$  para todo  $n, N$  y  $i \in S_N$ .*
- b) *Demostrar que el mínimo costo promedio en  $\Delta$  es una constante, y que existe una política estacionaria costo promedio óptima  $f$  inducida en la CM con una clase recurrente positiva  $R_f$  tal que AS es conforme en  $R_f$ .*

*Paso 4. Realizar uno de los siguientes:*

- a) *Demostrar que  $\nu_n^N(i) \geq \nu_n^N(x)$  para todo  $n, N$ , y  $i \in S_N$*
- b) *Demostrar que existe un conjunto no vacío y finito  $G$  tal que  $\nu_n^N$  toma un mínimo en  $G$  para todo  $n$  y  $N$ . Además existe una política estacionaria  $g$  induciendo una CM con una clase recurrente positiva  $R_g \supset G \cup \{x\}$  teniendo un costo promedio finito y tal que AS es conforme en  $R_g$ .*

**Proposición 10** *Sea  $\Gamma z$  estándar (véase apéndice C.5 de [13]). Supóngase que se tiene un ATAS y un entero no negativo  $N^*$  tal que la distribución de aumentación satisface*

$$\sum_{j \in S_n - \{z\}} q_j(i, r, N) m_{jz} \leq m_{rz}, i \in S_N, r \notin S_N, N \geq N^*, \quad (\text{B.9})$$

y

$$\sum_{j \in S_N - \{z\}} q_i(i, r, N) C_{jz} \leq C_{rz}, i \in S_N, r \notin S_N, N \geq N^*. \quad (\text{B.10})$$

Entonces el ATAS es conforme.

**Demostración.** La idea básica de (B.9) es que la combinación convexa del primer tránsito de tiempo en  $\Gamma$  corresponde al exceso de probabilidad  $P_{ir}$  no puede exceder los primeros tránsitos de tiempo asociados con  $r$ . Este es un tipo de propiedad estructural. Un comentario similar se cumple para (B.10). Observe que

$$\sum_{j \in S_N - \{z\}} P_{ij}(N) m_{jz} = \sum_{j \in S_N - \{z\}} P_{ij} m_{jz} \quad (\text{B.11})$$

$$\begin{aligned} &+ \sum_{r \notin S_N} P_{ir} \left( \sum_{j \in S_N - \{z\}} q_j(i, r, N) m_{jz} \right). \quad (\text{B.12}) \\ &\leq \sum_{j \in S_N - \{z\}} P_{ij} m_{jz} + \sum_{r \notin S_N} P_{ir} m_{rz} \\ &= m_{iz} - 1, \quad i \in S_N \end{aligned}$$

La primera línea se sigue de la ecuación (C.27) del apéndice C de [13]. La segunda línea se sigue de la ecuación (B.9), y la última línea es consecuencia de la ecuación (C.4) del apéndice C de [13]. La hipótesis del Corolario C.16 del apéndice C de [13] se satisfacen para  $\Gamma_n$  (con  $y(i) = m_{iz}$  para  $i \in S_N - \{z\}$  y  $y(z) = 0$ ), y por lo tanto se sigue que  $m_{iz}(N) \leq m_{iz}$  para  $i \in S_N - \{z\}$ . La proposición C.4.2 (ii) del apéndice C de [13] implica que  $m_{iz}(N) \rightarrow m_{iz}$  para  $i \neq z$ . Ahora

$$m_{zz}(N) = 1 + \sum_{j \in S_N - \{z\}} P_{zj}(N) m_{jz}(N). \quad (\text{B.13})$$

Queremos aplicar el Teorema A.2.6 del apéndice A de [13] con función acotada  $m_{jz}$ . La hipótesis se tendrá si se demuestra que

$$\lim_{N \rightarrow \infty} \sum_{j \in S_N - \{z\}} P_{zj}(N) m_{jz} = \sum_{j \neq z} P_{zj} m_{jz}. \quad (\text{B.14})$$

Si (B.14) puede ser mostrada, entonces el Teorema A.2.6 del apéndice A de [13] se tiene

$$m_{zz}(N) \rightarrow 1 + \sum_{j \neq z} P_{zj} m_{jz} = m_{zz}.$$

Así entonces mostremos (B.14). Se sigue de (B.11) y la proposición A.1.8 del apéndice A de [13] que

$$\begin{aligned} m_{zz} - 1 &\geq \limsup_{N \rightarrow \infty} \sum_{j \in S_N - \{z\}} P_{zj}(N) m_{jz} \geq \liminf_{N \rightarrow \infty} \sum_{j \in S_N - \{z\}} P_{zj}(N) m_{jz} \\ &\geq \sum_{j \neq z} P_{zj} m_{jz} = m_{zz} - 1. \end{aligned}$$

Por lo tanto todos los términos son iguales y (B.14) se cumple. La demostración para el costo es similar y se omite. ■

A continuación se dan algunos resultados debido a [12].

### B.3. Proceso de Decisión Semi-Markoviano

Un Proceso de Decisión Semi-Markoviano opera como sigue. El proceso se observa en el tiempo  $t = 0$  en algún estado  $i \in S$ , donde  $S$  es un conjunto numerable. En tal tiempo se toma una acción  $a \in A_i$ , donde  $A_i$  es un conjunto finito de acciones disponibles en el estado  $i$ .

Si  $i$  el estado actual y una acción  $a$  se elige, entonces (1) el proceso se mueve al estado  $j$  con probabilidad  $P_{ij}(a)$ , y (2) condicionado en el presente estado  $i$  y el siguiente estado  $j$ , el tiempo hasta la siguiente transición es una variable aleatoria con distribución  $F_{ij}(t | a)$ . Este proceso se repite indefinidamente.

La estructura del costo es como sigue. Si en el estado  $i$  una acción  $a$  se elige, se incurre en un costo inmediato  $D(i, a)$ , y un costo de  $d(i, a)$  por unidad de tiempo, se incurre hasta la siguiente transición. Supongamos que ambos costos son acotados y no negativos.

A continuación daremos algunas definiciones e hipótesis requeridas. Sea  $H_i(t | a) = \sum_j P_{ij}(a) F_{ij}(t | a)$  la función de distribución del tiempo de estancia en el

estado  $i$ , cuando una acción  $a$  se elige. Sea  $\tau(i, a) = \sum_j P_{ij}(a) \int_0^\infty t dF_{ij}(t | a)$ , el tiempo medio de estancia en el estado  $i$  cuando una acción  $a$  se eligió.

Una política es una regla no anticipatoria para elegir acciones. Esto puede depender de los estados anteriores del proceso, las acciones elegidas en esos estados, el tiempo de estancia en esos estados, y pueden ser aleatorizadas. Una política estacionaria (no aleatorizada)  $f$  tiene la propiedad, que cuando el proceso está en el estado  $i$ , la acción  $f(i)$  siempre se elige. Abusando de la notación denotamos  $D(i, f(i))$  como  $D(i, f)$  y similarmente con otros costos.

Para un factor de descuento  $\alpha > 0$ , sea  $C_\alpha(i, a)$  el costo esperado descontado incurrido en una etapa cuando el proceso se encuentra en el estado  $i$  y una acción  $a$  se elige, y sea

$$C(i, a) = D(i, a) + d(i, a)\tau(i, a), \quad (\text{B.15})$$

el costo sin descuento incurrido durante una transición. Entonces

$$C_\alpha(i, a) = D(i, a) + d(i, a) \sum_j P_{ij}(a) \int_0^\infty \int_0^t e^{-\alpha s} ds dF_{ij}(t | a) \uparrow C(i, a) \text{ si } \alpha \downarrow 0 \quad (\text{B.16})$$

Ahora fijemos una política  $\theta$  y sean  $X_n$   $a_n$   $n \geq 0$ , respectivamente, el estado después de la  $n$ -ésima transición y la acción elegida en tal estado (el estado inicial es  $X_0$ ). Sea  $t_n$  el tiempo de la  $n$ -ésima transición ( $t_0 = 0$ ) y sea  $\delta_n = t_n - t_{n-1}$ ,  $n \geq 1$ , el tiempo de estancia en el estado  $(n - 1)$ .

El costo esperado total  $\alpha$ -descontado bajo la política  $\theta$ , con un estado inicial  $i$ , esta dado por:

$$V_{\theta, \alpha}(i) = E_\theta \left\{ \sum_{n=0}^{\infty} e^{-\alpha t_n} \left( D(X_n, a_n) + d(X_n, a_n) \int_0^{\delta_{n-1}} e^{-\alpha t} dt \right) \mid X_0 = i \right\}.$$

Definimos  $V_\alpha(i) = \inf_\theta V_{\theta, \alpha}(i)$ .

Existen dos definiciones comunes del costo promedio esperado incurrido bajo una política  $\theta$ , sea  $Z(t)$  el costo total incurrido al tiempo  $t$ , y sea  $Z_n$  el costo total incurrido durante los primeros  $n$  intervalos de transición ( $n \geq 1$ ). Definimos

$$\phi_\theta(i) = \limsup_{t \rightarrow \infty} \frac{E_\theta(Z(t))}{t}, \quad \psi_\theta(i) = \limsup_{n \rightarrow \infty} \frac{E_\theta(Z_n)}{E_\theta(t_n)}. \quad (\text{B.17})$$

Definimos  $h_\alpha(i) = V_\alpha(i) - V_\alpha(0)$ , donde 0 es cualquier estado distinguido. Ahora damos algunas hipótesis.

**Hipótesis 6** Existe  $\delta > 0$  y  $\epsilon > 0$  tal que  $1 - H_i(\delta | a) \geq \epsilon$  para todo  $i, a$ .

**Hipótesis 7** Existe  $B$  tal que  $\tau(i, a) \leq B$  para cualquier  $i$  y  $a$ .

**Hipótesis 8** Suponga que  $V_\alpha(i) < \infty$  para todo  $i$  y  $\alpha$ .

**Hipótesis 9** Existe  $\alpha_0 > 0$  y un número no negativo  $M_i$  tal que  $h_\alpha(i) \leq M_i$  para cualquier  $i$  y  $0 < \alpha < \alpha_0$ . Para cualquier  $i$ , existe una acción  $a(i)$  tal que  $\sum_j P_{ij}(a(i))M_j < \infty$ .

**Hipótesis 10** Existe  $\alpha_0 > 0$  y un entero no negativo  $N$  tal que  $-N \leq h_\alpha(i)$  para cualquier  $i$  y  $0 < \alpha < \alpha_0$ .

**Proposición 11** Supongamos que se cumplen las hipótesis 6-10 y sea  $d(i) = \min_a \{d(i, a)\}$  para  $i \in S$ . Sean  $f$  una política estacionaria,  $g$  una constante y  $h(i)$ , con  $-N \leq h(i) \leq M_i$  tal que

$$g\tau(i, f) + h(i) \geq C(i, f) + \sum_j P_{ij}(f)h(j), \quad i \in S \quad (\text{B.18})$$

Supongamos además que existe  $\epsilon > 0$  y un subconjunto finito  $G$  tal que  $d(i) \geq B(g + \epsilon)/\tau$  para  $i \notin G$ . Entonces  $\phi_f(i) = \psi_f(i)$  para todo  $i$  y por lo tanto  $f$  es costo promedio esperado óptima.

Consideremos las siguientes hipótesis:

**Hipótesis 11** Supongamos que existe una política estacionaria  $f$ , con  $S_f = S$ , donde  $S_f = \{i \in S \mid P_{ji}(f) > 0 \text{ para algún estado } j\}$ . Sea  $c_{ij}(f)$  el costo promedio en un paso de ir de  $i$  a  $j$ , satisfaciendo que  $c_{ij}(f) < \infty$ ,  $i, j \in S_f$ .

**Hipótesis 12** Supongamos que en un paso en el proceso semi-Markoviano inducido por  $f$  en  $S_f$  son nonlattice.

**Hipótesis 13** Supongamos que existe una función  $u : S \rightarrow S_f$  tal que  $P_{ij}(f) = P_{u(i)j}(f)$  para todos los estados  $i, j$ .

**Hipótesis 14** Supongamos que  $\sum_{i \in S_f} \pi_i \tau(i, f) < \infty$ . Donde  $\pi_i$  es la probabilidad de estado estacionario en el estado  $i$ .

**Hipótesis 15** Sea  $d(i) = \min_a \{d(i, a)\}$ , supongamos que dado cualquier número  $M$  positivo, existe un conjunto  $G$  tal que  $d(i) \geq M$  para  $i \notin G$ .

**Proposición 12** Supongamos que las hipótesis 6 y 7 se cumplen y además las hipótesis 11-15. Entonces existe una política estacionaria costo promedio óptima.

# Apéndice C

## Equivalencia de PDMs a Tiempo Continuo y a Tiempo Discreto

A continuación presentamos la equivalencia entre un PDMTD y un PDMTC. Dicho resultado es debido a Serfozo [15]. Antes de presentar el resultado daremos algunas definiciones.

**Definición 29** Sea  $Y = \{Y(t) : t \geq 0\}$  un proceso de Markov en tiempo continuo con espacio de estados numerable y probabilidad de transición  $Q(i, j)$ , y cuya estancia en el estado  $i$  tiene una distribución exponencial con parámetro  $0 \leq \lambda_i < \infty$  ( $\lambda_i = 0$  cuando  $i$  es estado absorbente). (Véase Cinlar [3] ó Gihman y Skorohod [4]). El generador infinitesimal de  $Y$  está dado por:

$$A(i, j) = \lim_{t \rightarrow 0} \frac{d}{dt} P(Y(t) = j \mid Y(0) = i) = \begin{cases} -p_i \lambda_i & \text{si } i = j \\ \lambda_i Q(i, j) & \text{si } i \neq j, \end{cases} \quad (\text{C.1})$$

donde  $p_i = 1 - Q(i, i)$ .

**Comentario 3** Sea  $Y = \{Y(t) : t \geq 0\}$  como en la definición anterior y supongamos que  $c = \sup_i p_i \lambda_i < \infty$ , es decir los parámetros de transición de  $Y$  están acotados. Ahora consideremos otro proceso de Markov en tiempo continuo  $Y' = \{Y'(t) : t \geq 0\}$  con matriz de transición:

$$Q'(i, j) = \begin{cases} 1 - p_i \lambda_i / c & \text{si } i = j \\ \lambda_i Q(i, j) / c & \text{si } i \neq j, \end{cases} \quad (\text{C.2})$$

y el parámetro  $\lambda'_i = c$  para todo  $i$ .

Es fácil demostrar que el generador de  $Y'$  es el mismo que el de  $Y$ . De aquí se sigue que  $Y$  es igual en distribución a  $Y'$ . Es decir sus distribuciones finito dimensionales son iguales cuando tienen la misma distribución inicial.

La equivalencia entre  $Y$  y  $Y'$  se usa para simplificar ciertos cálculos para  $Y$  (véase [1]): la  $Y'$  es una cadena subordinada a un proceso de Poisson y es más manejable que  $Y$ .

## C.1. Equivalencia

Consideremos un PDMTC como sigue. Supongamos que el proceso está en el estado  $i \in S$ , una acción  $a \in A$  se elige y una recompensa  $r(i, a)$  se obtiene. Para simplificar sean  $S$  y  $A$  conjuntos numerables. El proceso permanece en el estado  $i$  un tiempo aleatorio el cual está distribuido exponencialmente con parámetro  $\lambda(i, a) \geq 0$  y pasa al estado  $j \in S$  con probabilidad  $p(i, a, j)$ . Esta serie de eventos se repite indefinidamente. Note que el parámetro  $\lambda(i, a) = 0$  significa que  $i$  es un estado absorbente bajo la acción  $a$ . Además  $p(i, a, i) > 0$  significa que el proceso puede tomar un salto ‘falso’ desde  $i$  a si mismo.

Sea  $f$  una política estacionaria la cual elige una acción  $f(i)$  cuando el proceso está en el estado  $i$ . Supongamos que un descuento y una recompensa bajo  $f$  existe. Denotemos por

$$W_f(i) = E_f\left(\sum_{n=0}^{\infty} e^{-\beta T_n} r(Y_n, a_n) \mid Y_0 = i\right), \quad (\text{C.3})$$

donde  $\beta > 0$  es el factor de descuento.

Además

$$\Psi_f(i) = \lim_{t \rightarrow 0} E_f\left(t^{-1} \sum_{n=0}^{N_t} r(Y_n, a_n) \mid Y_0 = i\right). \quad (\text{C.4})$$

Aquí el proceso de Markov controlado  $Y = \{Y(t) : t \geq 0\}$  está dado por  $Y(t) = Y_n$  si  $T_n \leq t < T_{n+1}$ , donde  $T_n$  es el tiempo de  $n$ -ésimo salto de  $Y$  al estado  $Y_n$ , la acción  $a_n = f(Y_n)$  y  $N_t = \max\{n : t_n \leq t\}$  es el número de saltos que ocurren en un tiempo  $t$ . Por convención usamos que  $t_k = 0$  para  $k \geq n + 1$  cuando  $Y$  es absorbido en el estado  $Y_n$ .

Denotemos este proceso de Markov controlado como:  $Y = (S, A, r, \lambda, p, \beta)$ .

Consideramos además a  $r(i, a)$  como una recompensa total recibida al comienzo de un estado  $i$  cuando una acción  $a$  se elige: esta es una recompensa compuesta asociada con la estancia y el siguiente salto. En aplicaciones típicas,  $r$  es de la

forma:

$$\begin{aligned} r(i, a) &= E_f \left[ \rho_0(i, Y_i) + \int_0^{T_1} e^{-\beta t} \rho_1(i, Y_1) dt + e^{-\beta T_1} \rho_2(i, Y_1) \mid Y_0 = i \right] \quad (C.5) \\ &= \sum_j p(i, a, j) [\rho_0(i, j) + (\rho_1(i, j) + \rho_2(i, j)\lambda(i, a))/(\beta + \lambda(i, a))] \quad (C.6) \end{aligned}$$

donde  $a = f(i)$ . Aquí  $\rho_0(i, j)$  y  $\rho_2(i, j)$  son las recompensas totales recibidas durante la estancia cuando en el siguiente estado  $j$ .

Consideremos una cadena de Markov controlada en tiempo discreto, como sigue. Arribando a un estado  $i \in S$ , una acción  $a \in A$  se toma, y una recompensa  $r(i, a)$  se recibe, y entonces el proceso salta a un estado  $j \in S$  con probabilidad  $p(i, a, j)$ . Esta serie de eventos se repite indefinidamente. Por lo anterior sea  $f$  una política estacionaria y supongamos que las recompensas promedio y descontadas bajo  $f$  existen. Denotemos por:

$$V_f(i) = E_f \left( \sum_{n=0}^{\infty} \alpha^n r(X_n, a_n) \mid X_0 = i \right), \quad (C.7)$$

donde  $0 < \alpha < 1$  es un factor de descuento, y

$$\Phi_f(i) = \lim_{n \rightarrow \infty} E_f \left( n^{-1} \sum_{k=0}^{n-1} r(X_k, a_k) \mid X_0 = i \right). \quad (C.8)$$

Aquí  $X_n$  es el  $n$ -ésimo estado del proceso y  $a_n = f(X_n)$  es la  $n$ -ésima acción elegida. Denotamos esta cadena de Markov controlada por  $X = (S, A, r, p, \alpha)$ .

El siguiente resultado afirma que si un proceso de Markov controlado tiene parámetros de transición uniformemente acotados, entonces se puede construir una cadena de Markov controlada la cual es equivalente a ésta.

**Teorema 13** *Sea  $\hat{Y} = (S, A, \hat{r}, \hat{\lambda}, \hat{p}, \beta)$  un proceso de Markov controlado con  $(1 - \hat{p}(i, a, i))\hat{\lambda}(i, a) \leq c < \infty$  para toda acción  $a$  y estado  $i$ . Sea  $Y = (S, A, r, \lambda, p, \beta)$  un proceso de Markov controlado donde*

$$r(i, a) = \begin{cases} \frac{\hat{r}(i, a)(\beta + \hat{\lambda}(i, a))}{(\beta + c)} & \text{cuando se consideran recompensas descontadas} \\ \frac{\hat{r}(i, a)\hat{\lambda}(i, a)}{c} & \text{cuando se consideran recompensas promedio} \end{cases} \quad (C.9)$$

$$p(i, a, j) = \begin{cases} \frac{1 - \hat{\lambda}(i, a)(1 - \hat{p}(i, a, i))}{c}, & \text{si } i = j, \\ \hat{\lambda}(i, a) \hat{p}(i, a, j)/c & \text{si } i \neq j, \end{cases} \quad (C.10)$$

y  $\lambda(i, a) = c$  para todo  $a$  y  $i$ . Sea  $X = (S, A, r, p, c/(\beta + c))$  una cadena de Markov controlada, donde  $r$  y  $p$  están definidas por C.9. Si  $\hat{Y}$ ,  $Y$  y  $X$  están controladas por una política estacionaria  $f$ , entonces  $\hat{W}_f = W_f = V_f$  y  $\hat{\Psi}_f = \Psi_f = c\Phi_f$ .

**Comentario 4** *Supongamos que los procesos de decisión  $\hat{Y}$ ,  $Y$  y  $X$  tienen una política estacionaria óptima dentro de estas clases respectivas de políticas no estacionarias aleatorizadas. Es válido por ejemplo, cuando  $A$  es finito y  $\hat{r}$  es acotado. En efecto  $\hat{Y}$  y  $Y$  tienen una mayor clase de políticas que  $X$  pero los tres tienen las mismas clases de políticas estacionarias (deterministas). Del resultado anterior se sigue que si una política estacionaria  $f$  es óptima descontada para cualesquiera  $\hat{Y}$ ,  $Y$  ó  $X$ , entonces esta es óptima descontada para las otras dos. Un comentario similar aplica para la política promedio óptima,  $\epsilon$ -descontada ó  $\epsilon$ -política promedio óptima. En este sentido, los procesos de decisión  $\hat{Y}$ ,  $Y$  y  $X$  son equivalentes.*

**Demostración.** Bajo una política  $f$  sea  $\hat{Y}$  un proceso de Markov con parámetro de estancia exponencial  $\hat{\lambda}_i = \hat{\lambda}(i, f(i))$ , y probabilidad de transición

$$Q(i, j) = p(i, f(i), j) = \begin{cases} 1 - \hat{\lambda}_i(1 - \hat{Q}(i, i))/c & \text{si } i = j, \\ \hat{\lambda}_i \hat{Q}(i, j)/c & \text{si } i \neq j. \end{cases} \quad (\text{C.11})$$

Claramente  $\hat{Y}$  y  $Y$  tienen el mismo generador infinitesimal y entonces son iguales en distribución (cuando ambos tienen la misma distribución inicial). Entonces para probar que  $\hat{W}_f = W_f$ , es suficiente demostrar que la recompensa esperada descontada asociada con  $\hat{Y}$  y  $Y$  para el momento ‘actual’ en el estado  $i$  son iguales para cada  $i$ . Considérese en el momento actual en el estado  $i$  por el proceso  $Y$ . El proceso  $Y$  puede tomar un salto falso desde  $i$  a si mismo con probabilidad  $Q(i, i)$ . Consecuentemente, el número  $v$  de saltos falsos desde  $i$  en si mismo, antes de un nuevo estado es alcanzado, es una variable aleatoria geométrica con  $P(v = n) = Q(i, i)^n(1 - Q(i, i))$  para  $n \geq 0$ . Aquí  $v = 0$  cuando  $Q(i, i) = 0$ , y  $v = \infty$  cuando  $Q(i, i) = 1$ . Entonces la recompensa esperada recibida por  $Y$  para una estancia actual en el estado  $i$  es:

$$E \sum_{n=0}^v e^{-\beta \tau_n} r(i, f(i)) = r(i, f(i)) E \sum_{n=0}^v \gamma^n = \frac{r(i, f(i))}{(1 - \gamma Q(i, i))}. \quad (\text{C.12})$$

Aquí  $\tau_n$  denota el tiempo del  $n$ -ésimo salto falso en el estado  $i$ : este tiene una distribución gamma con parámetros  $\gamma = c/(\beta + c)$  y  $n$ , y es independiente de  $v$ . Por un argumento similar se sigue que la recompensa esperada recibida por  $\hat{Y}$  para una estancia actual en  $i$  es:

$$E \sum_{n=0}^{\hat{v}} e^{-\beta \hat{\tau}_n} \hat{r}(i, f(i)) = \hat{r}(i, f(i)) / (1 - \hat{\gamma}_i \hat{Q}(i, i)), \quad (\text{C.13})$$

donde  $\hat{\tau}_n$  tiene una distribución gamma con parámetro  $\hat{\gamma}_i = \hat{\lambda}_i / (\beta + \hat{\lambda}_i)$  y  $n$ . Sustituyendo (C.9) y (C.11) en (C.12) se puede ver que la recompensa en (C.12) y (C.13)

son iguales. De donde se tiene que  $\hat{W}_f = W_f$ . Además, por un obvio uso de la esperanza condicional en  $W_f$ , se sigue que  $W_f = V_f$ . Ahora considere la recompensa promedio  $\hat{\Psi}_f$ ,  $\Psi_f$  y  $\Phi_f$  bajo la política  $f$  para  $\hat{Y}$ ,  $Y$  y  $X$  respectivamente. Por el conocido Teorema Abeliano [16, pág. 182][12,p.182],  $\hat{\Psi}_f(i) = \lim_{\beta \rightarrow 0} \beta \hat{W}_f(i) = \lim_{\beta \rightarrow 0} \beta W_f(i) = \Psi_f(i)$  para todo  $i$ . Además, Teoremas Abelianos para sumas e integrales,

$$\Psi_f(i) = \lim_{\beta \rightarrow 0} \beta W_f(i) \tag{C.14}$$

$$= \lim_{\beta \rightarrow 0} \beta E_f \left( \sum_{n=0}^{\infty} (c/(\beta + c))^n r(X_n, f(X_n)) \mid X_0 = i \right) \tag{C.15}$$

$$= \lim_{\alpha \rightarrow 1} c\alpha^{-1}(1 - \alpha) E_f \left( \sum_{n=0}^{\infty} \alpha^n r(X_n, f(X_n)) \mid X_0 = i \right) \tag{C.16}$$

$$= c\Phi_f(i). \tag{C.17}$$

■

# Apéndice D

## Bibliografía

- [1] Bertsekas D. P. Dynamic Programming: Deterministic and Stochastic Models. Prentice-Hall, Inc., 1987.
- [2] Chung, K. L. Markov Chains with Stationary Transition Probabilities. Berlin: Springer-Verlag, 1967.
- [3] Cinlar E., Introduction to Stochastic Processes, Prentice-Hall, Englewood Cliffs, N. J., 1975.
- [4] Gihman E. and A. Skorohod. The Theory of Stochastic Processes II, Springer-Verlag, New York, 1975.
- [5] Hay J. Edward. The Just-in-Time Breakthrough: Implementing the New Manufacturing Basics. John Wiley & Sons, 1988.
- [6] Hernández L. O., Procesos Estocásticos: Introducción a la Teoría de Colas, Cinvestav Dpto. de Matemáticas, México 1981.
- [7] Kulkarni, V.G. Modeling and Analysis of Stochastic Systems, Chapman and Hall, London, 1995.
- [8] Little J.D.C. A proof of the queueing formula  $L = \lambda W$ , Operations Research 9, pp. 383-387, 1961.
- [9] Puterman, M.L. Markov Decision Processes. John Wiley, New York, 1994.
- [10] Ross, S.M., Applied Probability Models with Optimization Applications. Holden-Day, San Francisco, CA, 1970.

- 
- [11] Schweitzer, P. Iterative solution of the functional equation of undiscounted Markov renewal programming. *Journal of Mathematical Analysis and Applications* 34, 495–501, 1971.
  - [12] Sennott, L.I. Average cost semi-Markov decision processes and the control of queueing systems. *Probability in the Engineering and Informational Sciences* 3, 247–272, 1989.
  - [13] Sennott, L.I. *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley, New York, 1999.
  - [14] Sennott L. I. Van Oyen M. P., Irvani S. M. Optimal dynamic assignment of a flexible worker on an open production line with specialists. *European Journal of Operational Research* 170, 541-566, 2006.
  - [15] Serfozo R. F. An equivalence between continuous and discrete time Markov decision processes. *Operations Research* vol. 27, No. 3, May-June 1979.
  - [16] D. Widder. *The Laplace Transform*, Princeton University Press, Princeton N. J., 1941.
  - [17] Van Oyen, M.P., Gel, E.G.S., Hopp, W.J. Performance opportunity for workforce agility in collaborative and non-collaborative work systems. *IIE Transactions* 33 (9), 761–777, 2001.
  - [18] Yasuhiro Monden. *Toyota Production System: An Integrated Approach to Just-in-time*. Industrial Engineering and Management Press, 1993.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

# ACTA DE EXAMEN DE GRADO

No. 00028

Matricula: 205280021

TRABAJADORAS FLOTANTES EN  
PROCESOS DE PRODUCCION:  
ASIGNACIONES DINAMICAS  
OPTIMAS.

En México, D.F., se presentaron a las 11:00 horas del día 27 del mes de marzo del año 2009 en la Unidad Iztapalapa de la Universidad Autónoma Metropolitana, los suscritos miembros del jurado:

- DR. JUAN GONZALEZ HERNANDEZ
- DR. JULIO CESAR GARCIA CORTE
- DR. EVGUENI GORDIENKO
- DR. JOSE RAUL MONTES DE OCA MACHORRO

Bajo la Presidencia del primero y con carácter de Secretario el último, se reunieron para proceder al Examen de Grado cuya denominación aparece al margen, para la obtención del grado de:

MAESTRO EN CIENCIAS (MATEMÁTICAS APLICADAS E INDUSTRIALES)

DE: ISMAEL REYES GONZALEZ



ISMAEL REYES GONZALEZ  
ALUMNO

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

*Aprobar*

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

REVISÓ

LIC. JULIO CESAR DE LARA ISASSI  
DIRECTOR DE SISTEMAS ESCOLARES

DIRECTORA DE LA DIVISIÓN DE CBI

DRA. VERÓNICA MEDINA BAÑUELOS

PRESIDENTE

DR. JUAN GONZALEZ HERNANDEZ

VOCAL

DR. JULIO CESAR GARCIA CORTE

VOCAL

DR. EVGUENI GORDIENKO

SECRETARIO

DR. JOSE RAUL MONTES DE OCA  
MACHORRO