

SOLUCIÓN DE UN PROCESO
MARKOVIANO POR PROGRAMACIÓN
DINÁMICA

Pacheco Mendoza Jesús

Julio de 2003

CONTENIDO

1	PRELIMINARES	1
1.1	TOMA DE DECISIONES	1
1.2	PROGRAMACIÓN DINÁMICA	8
1.2.1	INTRODUCCIÓN	8
1.2.2	ELEMENTOS DE UN MODELO DE PD	9
1.2.3	CARACTERÍSTICAS DE LOS PROBLEMAS DE PROGRAMACIÓN DINÁMICA	9
1.2.4	LA PROGRAMACIÓN DINÁMICA TIPO DETERMINÍSTICA	11
1.2.5	LA PROGRAMACIÓN DINÁMICA PROBABILÍSTICA	12
1.3	PROCESOS ESTOCÁSTICOS	14
1.3.1	INTRODUCCIÓN	14
1.3.2	LA PROPIEDAD DE MARKOV Y MATRICES DE TRANSICIÓN	18
1.3.3	ECUACIONES DE CHAPMAN-KOLMOGOROV	19
1.3.4	CLASIFICACIÓN DE ESTADOS EN UNA CADENA DE MARKOV	21
1.3.5	TIEMPOS DE PRIMERA PASADA	24
1.3.6	PROPIEDADES A LARGO PLAZO DE LAS CADENAS DE MARKOV	25
2	TEORÍA DE INVENTARIOS	29
2.1	COMPONENTES DE LOS MODELOS DE INVENTARIOS	29
2.2	MODELOS DETERMINÍSTICOS	31
2.2.1	REVISIÓN CONTINUA, DEMANDA UNIFORME Y NO SE PERMITEN FALTANTES	31
2.2.2	REVISIÓN CONTINUA, DEMANDA UNIFORME Y SE PERMITEN FALTANTES	32
2.3	MODELOS ESTOCÁSTICOS	34
2.3.1	MODELO DE UN PERIODO SIN COSTO FIJO	34
2.3.2	MODELO CON UN INVENTARIO INICIAL	36
2.3.3	DERIVACIÓN DE LA POLÍTICA ÓPTIMA	36
2.3.4	MODELO DE INVENTARIOS DE UN PERIODO CON COSTO DE PREPARACIÓN	38

2.3.5	MODELO DE INVENTARIOS DE DOS PERIODOS SIN COSTO DE PREPARACIÓN	40
2.3.6	MODELO DE VARIOS PERIODOS SIN COSTO DE PREPARACIÓN	43
2.4	PROCESOS DE DECISIÓN	43
2.4.1	MODELO UTILIZADOS PARA PROCESOS DE DECISIÓN MARKOVIANOS	43
2.4.2	MODELO DE ETAPA INFINITA	44
3	APLICACIÓN	51
3.1	INTRODUCCIÓN	51
3.2	SOLUCIÓN DE UN PROBLEMA DE INVENTARIOS	51
3.3	CONCLUSIONES	57

CAPÍTULO 1

PRELIMINARES

1.1 TOMA DE DECISIONES

El término toma de decisiones se refiere a la selección de una alternativa de entre todo un conjunto de ellas. Significa escoger, como tal, la toma de decisiones es sólo un paso dentro de este proceso de selección.

Surgen las siguientes preguntas:

- 1.- ¿Cómo debe actuarse al tomar una decisión?
- 2.- ¿Qué debe hacerse para tomar la mejor decisión?

Una rama de la matemática aplicada dedicada al estudio de la toma de decisiones lo constituye la *investigación de operaciones*. Como su nombre lo indica, la investigación de operaciones significa *hacer investigación sobre las operaciones*. La parte de investigación en el nombre significa que la investigación de operaciones usa un enfoque similar a la manera en que se lleva a cabo la investigación en los campos científicos establecidos. En gran medida se usa el método científico para investigar el problema en cuestión. El proceso comienza por la observación cuidadosa y la formulación del problema, incluyendo la recolección de los datos pertinentes. El siguiente paso es la construcción de un modelo científico (por lo general matemático) que intenta abstraer la esencia del problema real. En este punto se propone la hipótesis de que el modelo es una representación lo suficientemente precisa de las características esenciales de la situación como para que las conclusiones (soluciones) obtenidas de este modelo sean válidas también para el problema real. Después, se llevan a cabo los experimentos adecuados para probar esta hipótesis, modificarla si es necesario y eventualmente verificarla. En cierto modo, la investigación de operaciones incluye la investigación científica creativa de las propiedades fundamentales de las operaciones.

Una característica adicional es que la investigación de operaciones intenta encontrar una mejor solución, llamada solución óptima, para el problema en

consideración. Decimos *una mejor solución* y no *la mejor decisión* porque pueden existir muchas soluciones que empaten como la mejor.

Como ayuda para la toma de decisiones se usan matemáticas, ya que permiten expresar pensamientos complejos de manera concisa. Para ello se hace uso de algún modelo matemático. Las fases de solución de un problema son:

- 1.- Definición del problema.
- 2.- Construcción del modelo.
- 3.- Solución del modelo.
- 4.- Validación del modelo.
- 5.- Implantación de los resultados finales.

La *primera fase* indica tres aspectos principales:

- a) Una descripción de la meta u objetivo del estudio.
- b) Una identificación de las alternativas de decisión del sistema.
- c) Un reconocimiento de las limitaciones, restricciones y requisitos del sistema.

Este proceso de definir el problema es crucial ya que afectará en forma significativa la relevancia de las conclusiones del estudio. Es difícil extraer una respuesta *correcta* a partir de un problema *equivocado*. Determinar los objetivos apropiados vienen a ser un aspecto muy importante de la formulación del problema. Para hacerlo, es necesario primero identificar a la persona o personas de la administración que de hecho tomarán las decisiones concernientes al sistema bajo estudio, y después escudriñar el pensamiento de estos individuos respecto a los objetivos pertinentes. Incluir al tomador de decisiones desde el principio es esencial para obtener su apoyo al realizar el estudio.

En la *segunda fase* se decide el modelo más adecuado para representar el sistema, se deben especificar expresiones cuantitativas para el objetivo y las restricciones del problema en función de sus variables de decisión. La forma convencional en que la investigación de operaciones realiza esto es construyendo un modelo matemático que represente la esencia del problema. Antes de analizar cómo formular los modelos de este tipo, se explora la naturaleza general de los modelos matemáticos. Los modelos matemáticos son representaciones idealizadas que están expresadas en términos de símbolos y expresiones matemáticas. La determinación de los valores apropiados que deben asignarse a los parámetros del modelo es crítica y a la vez un reto dentro del proceso de construcción del modelo. Al contrario de los problemas en los libros en donde se proporcionan los números, la determinación de los valores de los parámetros en los problemas reales requiere la *recolección de los datos relevantes* (véase [5]). La recolección de los datos exactos con frecuencia es difícil. Así, el valor asignado a un parámetro muchas veces es, por necesidad, sólo una estimación.

Aún así, los modelos matemáticos tienen muchas ventajas sobre una descripción verbal del problema. Una ventaja obvia es que el modelo matemático

describe un problema en forma mucho más concisa. Esto tiende a hacer que toda la estructura del problema sea mucho más comprensible y ayude a revelar las relaciones importantes entre causa y efecto. De esta manera, indica con más claridad qué datos adicionales son importantes para el análisis. También facilita simultáneamente el manejo del problema en su totalidad y el estudio de todas sus interrelaciones. Por último, un modelo matemático forma un puente para poder emplear técnicas matemáticas y computadoras de alto poder, para analizar el problema.

Por otro lado, existen obstáculos que deben evitarse al usar modelos matemáticos. Un modelo, como se dijo anteriormente, es una idealización abstracta del problema, por lo que casi siempre se requieren aproximaciones y suposiciones de simplificación si se requiere que el modelo sea manejable. Por lo tanto, debe tenerse cuidado de que el modelo sea siempre una representación válida del problema. El criterio apropiado para juzgar la validez de un modelo es el hecho de si predice o no con suficiente exactitud los efectos relativos de los diferentes cursos de acción, para poder tomar una decisión que tenga sentido. En consecuencia no es necesario incluir detalles sin importancia o factores que tengan aproximadamente el mismo efecto sobre todas las opciones. Ni siquiera es necesario que la magnitud absoluta de la medida de efectividad sea aproximadamente correcta para las diferentes alternativas, siempre que sus valores relativos sean bastantes precisos. Entonces, todo lo que se requiere es que exista una alta correlación entre la predicción del modelo y lo que ocurre en la vida real. Para asegurar que este requisito se cumpla, es importante hacer un número considerable de pruebas del modelo y las modificaciones consecuentes.

Al desarrollar el modelo, se recomienda empezar con una versión sencilla y moverse, en una forma evolutiva, hacia modelos más elaborados que reflejen mejor la complejidad del problema real. Este proceso de enriquecimiento del modelo continúa sólo mientras permanezca manejable. El trueque básico que debe tomarse en cuenta todo el tiempo está entre la precisión y el manejo del modelo.

La *tercera fase* se logra usando técnicas de optimización bien definidas. Una vez formulado el modelo matemático para el problema bajo estudio, la siguiente etapa de un estudio de investigación de operaciones consiste en desarrollar un procedimiento para derivar una solución al problema a partir de este modelo. Puede pensarse que esto debe ser la parte principal del estudio pero, en realidad, en la mayoría de los casos no lo es. De hecho, a veces ésta es una etapa relativamente sencilla, en la que se aplica uno de los algoritmos de investigación de operaciones en una computadora, empleando uno de los paquetes de software disponibles (véase [5]).

Un tema común en investigación de operaciones es la búsqueda de una *solución óptima*, es decir, *la mejor*. Sin duda, se han desarrollado muchos procedimientos para encontrarla en cierto tipo de problemas, pero es necesario

reconocer que estas soluciones son óptimas sólo respecto al modelo que se está utilizando. Como el modelo es una idealización y no una representación exacta del problema real, no puede existir una garantía utópica de que la solución óptima del modelo resulte ser la mejor solución posible que pueda llevarse a la práctica para el problema real. Esto es de esperarse si se toman en cuenta los muchos imponderables e incertidumbres asociados a casi todos los problemas reales, pero si el modelo está bien formulado y verificado, la solución que resulta debe tender a una buena aproximación de un curso de acción ideal para el problema real. Por todo esto, más que enfrascarse en pedir lo imposible, la prueba del éxito de un estudio de investigación de operaciones debe ser el hecho de si proporciona o no una mejor guía en las decisiones que la que se puede obtener por otros medios.

El eminente científico de la administración y premio Nobel de Economía, Herbert Simon(ver [5]), introdujo el concepto de que en la práctica es mucho más frecuente satisfacer que optimizar. Al inventar el término satisfacer como una combinación de satisfacer y optimizar, Simon describe la tendencia de los administradores a buscar una solución que sea *lo suficientemente buena* para el problema que se tiene. En lugar de intentar desarrollar una medida global de eficiencia para conciliar de manera óptima los conflictos entre los diferentes objetivos deseables se puede usar un enfoque más pragmático. Las metas se pueden establecer de manera que marquen los niveles mínimos satisfactorios de eficiencia en las diferentes áreas, basándose quizá en niveles de desempeño anteriores o en los logros de la competencia. Si se encuentra una solución que permita que todas estas metas se cumplan, es posible que se adopte sin más requisitos. Ésta es la naturaleza de satisfacer. La distinción entre optimizar y satisfacer refleja la diferencia entre la teoría y la realidad, diferencia que con frecuencia se encuentra al tratar de implantar esa teoría en la práctica.

La *cuarta fase* indica cuando un modelo es válido, éste debe dar una predicción confiable del funcionamiento del sistema. Para realizar la validez, se puede comparar su funcionamiento con algunos datos pasados disponibles del sistema actual. El modelo será válido si, bajo condiciones similares de entrada, puede reproducir el funcionamiento pasado del sistema.

Es inevitable que la primera versión de un modelo matemático tenga muchas fallas. Sin duda, algunos factores o interrelaciones relevantes no se incorporaron al modelo y algunos parámetros no se estimaron correctamente. Por lo tanto, antes de usar el modelo debe probarse exhaustivamente para intentar identificar y corregir todas las fallas que se pueda. Con el tiempo, después de una larga serie de modelos mejorados, se concluye que el modelo actual produce resultados razonablemente válidos.

Este proceso de prueba y mejoramiento de un modelo para incrementar su validez se conoce como validación del modelo. Es difícil describir cómo se lleva a cabo la validación del modelo porque el proceso depende en gran parte del

problema bajo estudio y del modelo usado. Para ello, después de completar los detalles de la versión inicial del modelo, una buena manera de comenzar las pruebas es observarlo en forma global para verificar los errores u omisiones obvias. El grupo que hace esta revisión debe, de preferencia, incluir por lo menos a una persona que no haya participado en la formulación. Al examinar de nuevo la formulación del problema y compararla con el modelo pueden descubrirse este tipo de errores. También es útil asegurarse de que todas las expresiones matemáticas sean consistentes en las dimensiones de las unidades que emplean. Además, puede obtenerse un mejor conocimiento de la validez del modelo variando los valores de los parámetros de entrada y/o de las variables de decisión, y comprobando que los resultados del modelo se comporten de una manera factible.

Un enfoque más sistemático para la prueba del modelo es emplear una *prueba retrospectiva*. Esta prueba utiliza datos históricos y reconstruye el pasado para determinar si el modelo y la solución resultante hubiera tenido un buen desempeño, de haberse usado. La comparación de la efectividad de este desempeño hipotético con lo que en realidad ocurrió, indica si el uso del modelo tiende a dar mejoras significativas sobre la práctica actual. Puede también indicar áreas en las que el modelo tiene fallas y requiere modificaciones. Lo que es más, al emplear las alternativas de solución y estimar sus desempeños históricos hipotéticos, se pueden reunir evidencias en cuanto a lo bien que el modelo predice los efectos relativos de los diferentes cursos de acción.

Por otra parte, la prueba retrospectiva tiene la desventaja de que usa los mismos datos que sirvieron para formular el modelo. Entonces, la pregunta crucial es si el pasado en realidad representa el futuro. Para salvar esta desventaja de la prueba retrospectiva, a veces es útil continuar con las cosas como están por una temporada. Esto proporcionará datos con los que no se contaba cuando se construyó el modelo.

Es importante documentar el proceso usado para las pruebas de la validación del modelo. Esto ayuda a aumentar la confianza en él, de los usuarios subsecuentes. Más aún, si en el futuro surgen preocupaciones sobre el modelo, esta documentación ayudará a diagnosticar en donde pueden estar los problemas.

La *quinta fase* recae principalmente en los investigadores de operaciones, emitiendo en una forma comprensible a los individuos que administrarán y operarán el sistema. Si el modelo ha de usarse varias veces, el siguiente paso es instalar un sistema bien documentado para aplicar el modelo según lo establecido por la administración. Este sistema incluirá el modelo y el procedimiento de solución y los procedimientos operativos para su implementación. Entonces, aun cuando cambie el personal, el sistema puede consultarse periódicamente para proporcionar una solución numérica específica.

Una vez desarrollado un sistema para aplicar un modelo, la última etapa de un estudio de investigación de operaciones es implementarlo siguiendo lo

establecido por la administración. Esta etapa es crítica, ya que es aquí, y sólo aquí, donde se cosecharán los beneficios del estudio. Por lo tanto, es importante que el equipo de investigación de operaciones participe, tanto para asegurar que las soluciones del modelo se traduzcan con exactitud a un procedimiento operativo, como para corregir cualquier defecto en la solución que salga a la luz en este momento.

Resumimos brevemente los pasos del método científico:

1.- *DEFINICIÓN DEL PROBLEMA*. Se debe tener bien claro cuál es el problema a resolver y delimitarlo completamente ya que no tiene sentido encontrar la mejor solución para un problema equivocado.

2.- *RECOLECCIÓN DE DATOS*. Se estará más capacitado para resolver problemas si se tiene información sobre los datos de estos. Deberá reunirse información pasada, hechos pertinentes, y soluciones previas a problemas semejantes.

3.- *DEFINICIÓN DE ALTERNATIVAS DE SOLUCIÓN*. El método científico se basa en la suposición de que las soluciones existen. Por lo tanto se buscan las soluciones posibles tratando de caracterizar todas ellas.

4.- *EVALUACIÓN DE ALTERNATIVAS DE SOLUCIÓN*. Una vez caracterizadas todas las alternativas de solución, deberán evaluarse. Esto se puede lograr comparando los beneficios que se obtiene en cada una de ellas mediante un conjunto de criterios de solución u objetivos que se deben cumplir. También puede lograrse estableciendo rangos relativos de las alternativas de acuerdo a factores que sean importantes para la solución. Por lo general, se hacen ambas cosas.

5.- *SELECCIÓN DE LA MEJOR ALTERNATIVA*. Aquí se toma la decisión de cuál de las alternativas cumple mejor con los criterios de solución.

6.- *PUESTA EN PRÁCTICA*. La alternativa seleccionada deberá ponerse en práctica.

Aunque la mayoría de los problemas son diferentes, casi todos se pueden resolver por el método científico descrito anteriormente. Si bien los problemas y métodos pueden variar, es sorprendente el parecido en el proceso de razonamiento, ya que están basados en el método científico.

Por otra parte, el análisis de decisiones proporciona un marco conceptual y una metodología para la toma de decisiones racional. El análisis de decisiones divide la toma de decisiones en dos casos:

- 1.- Sin experimentación
- 2.- Con experimentación

En la toma de decisiones sin experimentación, el tomador de decisiones debe elegir una acción a de un conjunto de acciones posibles. El conjunto contiene todas las alternativas factibles bajo consideración para las distintas formas de proceder en el problema en cuestión. Esta elección de una acción debe hacerse frente a la incertidumbre porque el resultado se verá afectado por factores

aleatorios que se encuentran fuera del control del tomador de decisiones. Estos factores aleatorios determinan en qué situación se encontrará en el momento en que se ejecute la acción. Cada una de estas situaciones posibles se conoce como un *estado de la naturaleza*, que se denotará por θ . Para cada combinación de una acción a y un estado de la naturaleza θ , el tomador de decisiones sabe cuál sería el pago resultante. El pago es una medida cuantitativa del valor de las consecuencias del resultado para el tomador de decisiones. Sea $p(a, \theta)$ = pago al tomar la acción a cuando el estado de la naturaleza es θ .

En general, se usa una tabla de pagos para dar $p(a, \theta)$ para cada combinación de a y θ . Existe una analogía interesante entre los conceptos de análisis de decisiones y el juego de dos personas con suma cero. Un juego de dos personas con suma cero (véase[10]), como su nombre lo indica, es aquel en que participan sólo dos jugadores o adversarios y se le llama con suma cero porque un jugador gana lo que el otro pierde, de manera que la suma de sus ganancias netas es cero. El tomador de decisiones y la naturaleza se pueden ver como dos jugadores de este juego. Las acciones posibles y los estados de la naturaleza posibles se pueden ver como las estrategias disponibles para los respectivos jugadores, donde cada combinación de estrategias da como resultado un pago para el jugador 1 (el tomador de decisiones) . Desde este punto de vista, el marco conceptual del análisis de decisiones se puede resumir como sigue:

- 1.- El tomador de decisiones necesita elegir una de las acciones posibles.
- 2.- La naturaleza elegirá entonces uno de los estados de la naturaleza posibles.
- 3.- Cada combinación de una acción a y un estado de la naturaleza θ da como resultado un pago $p(a, \theta)$, que está dado como uno de los elementos de la tabla de pagos.
- 4.- Esta tabla de pagos debe usarse para encontrar una acción óptima para el tomador de decisiones según un criterio adecuado.

En la teoría de juegos se supone que ambos jugadores son racionales y eligen sus estrategias para promover su propio beneficio. Esta descripción se ajusta al tomador de decisiones pero no se ajusta a la naturaleza. Por el contrario, la naturaleza es un jugador pasivo que elige sus estrategias (estados de la naturaleza) de alguna manera aleatoria. Este cambio significa que el criterio de la teoría de juegos para la forma de elegir una estrategia óptima (acción) no será el más convincente para muchos tomadores de decisiones en el contexto actual.

Es necesario agregar otro elemento a estos conceptos de teoría de decisiones. El tomador de decisiones por lo general tendrá alguna información que debe tomar en cuenta sobre la posibilidad relativa de los estados de la naturaleza posibles. Es común que se pueda traducir esta información a una distribución de probabilidad, si se piensa que el estado de la naturaleza es una variable aleatoria, en cuyo caso esta distribución se conoce como una *distribución a priori*. Las probabilidades individuales para los respectivos estados se llaman *probabilidades a priori*.

En la toma de decisiones con experimentación se pueden realizar pruebas adicionales (experimentación) para mejorar las estimaciones preliminares de las probabilidades de los respectivos estados de la naturaleza que dan las probabilidades a priori. Estas estimaciones mejoradas se llaman probabilidades a posteriori. Para encontrar estas probabilidades a posteriori, sea n = número de estados de la naturaleza posibles; $P(\theta = \theta_i)$ = probabilidad a priori de que el estado de la naturaleza verdadero sea θ_i , para $i = 1, 2, \dots, n$; S = estadístico que resume los resultados de la experimentación (una variable aleatoria); s = una realización (valor posible) de S y $P(\theta = \theta_i | S = s)$ = probabilidad a posteriori de que el estado de la naturaleza verdadero sea θ_i , dado que $S = s$, para $i = 1, 2, \dots, n$.

La pregunta a contestar es: dado $P(\theta = \theta_i)$ y $P(S = s | \theta = \theta_i)$, ¿Cuál es el valor de $P(\theta = \theta_i | S = s)$?

Aplicando la definición de probabilidad condicional y la de probabilidad total tenemos:

$$P(\theta = \theta_i | S = s) = \frac{P(\theta = \theta_i, S = s)}{P(S = s)} \quad (1.1)$$

$$P(S = s) = \sum_{i=1}^n P(\theta = \theta_i, S = s) \quad (1.2)$$

$$P(\theta = \theta_i | S = s) = P(S = s | \theta = \theta_i) P(\theta = \theta_i). \quad (1.3)$$

Por lo tanto, para cada $i = 1, 2, \dots, n$, la fórmula deseada para la probabilidad a posteriori correspondiente es

$$P(\theta = \theta_i | S = s) = \frac{P(S = s | \theta = \theta_i) P(\theta = \theta_i)}{\sum_{i=1}^n P(\theta = \theta_i, S = s)}. \quad (1.4)$$

1.2 PROGRAMACIÓN DINÁMICA

1.2.1 INTRODUCCIÓN

La programación dinámica (PD) determina la solución óptima de un problema de n variables descomponiéndola en n etapas, con cada etapa incluyendo un subproblema de una sola variable.

La ventaja en el aspecto de los cálculos es que optimizaremos una sola variable, en vez de subproblemas de n variables. La principal contribución de la PD es el principio de optimalidad, un marco de referencia para descomponer el problema en etapas. La PD no proporciona los detalles de los cálculos para

optimizar cada etapa. Quien resuelve un problema improvisa y diseña esos detalles.

Existen dos tipos de programación dinámica: La programación dinámica determinística (PDD) y la programación dinámica probabilística (PDP). En la PDD se utilizan datos que se conocen con certeza y en la PDP se usan datos que no se conocen con certeza pero que se determinan a través de distribuciones de probabilidad.

1.2.2 ELEMENTOS DE UN MODELO DE PD

Dentro de un modelo de PD se pueden identificar tres elementos importantes, que son: las etapas, las alternativas en cada etapa y los estados para cada etapa. De estos tres elementos, el más importante es la definición de estado.

El estado del sistema se considera como la información que une cada etapa, de tal manera que se puedan tomar decisiones óptimas para las etapas restantes reexaminando la forma en la cual se llegó a las decisiones para las etapas previas.

PRINCIPIO DE OPTIMALIDAD. Las futuras decisiones para las etapas restantes constituirán una política óptima, sin importar cual haya sido la política adoptada en las etapas previas.

Los cálculos en la PD se hacen recursivamente, en el sentido de que la solución de un subproblema se utiliza como una entrada para el siguiente subproblema. Para el momento en que resolvamos el último subproblema, tendremos a la mano la solución óptima para todo el problema (veáse [10] y [5]). La forma en la cual se hacen los cálculos recursivos depende de la forma en la cual descomponemos el problema original.

Existen dos formas de realizar estos cálculos recursivos. Una forma es *la recursión hacia adelante*, en la cual los cálculos avanzan de la primera etapa hasta llegar a la última etapa. La otra forma es *la recursión hacia atrás*, en la que los cálculos empiezan de la última etapa y terminan con la primera etapa.

1.2.3 CARACTERÍSTICAS DE LOS PROBLEMAS DE PROGRAMACIÓN DINÁMICA

Se resumen brevemente las características de los problemas de PD.

- 1.- El problema se puede dividir en *etapas* que requieren una *política de decisión* en cada una de ellas.
- 2.- Cada etapa tiene cierto número de *estados* asociados con su inicio.
- 3.- El efecto de la política de decisión en cada etapa es *transformar el estado actual en un estado asociado con el inicio de la siguiente etapa* (tal vez de acuerdo a una distribución de probabilidad).

4.- El procedimiento de solución está diseñado para encontrar una *política óptima* para el problema completo, es decir, una receta para la política de decisión óptima en cada etapa para cada uno de los estados posibles.

5.- Dado el estado actual, una *política óptima para las etapas restantes* es *independiente* de la política adoptada en *etapas anteriores*. Por tanto la decisión inmediata óptima depende sólo del estado actual y no de cómo se llegó ahí. Éste es el principio de optimalidad para programación dinámica.

6.- El procedimiento de solución se inicia al encontrar la *política óptima para la última etapa*.

7.- Se dispone de una relación recursiva que identifica la política óptima para la etapa n , dada la política óptima para la etapa $n + 1$.

La notación que se usará se resume a continuación.

N = número de etapas.

n = etiqueta para la etapa actual ($n = 1, 2, \dots, N$)

s_n = estado actual para la etapa n .

x_n = variable de decisión para la etapa n .

x_n^* = valor óptimo de x_n (dado s_n).

$f_n(s_n, x_n)$ = contribución a la función objetivo de las etapas $n, n + 1, \dots, N$, si el sistema se encuentra en el estado s_n en la etapa n , la decisión inmediata es x_n y en adelante se toman decisiones óptimas.

$$f_n^*(s_n) = f_n(s_n, x_n^*)$$

La función recursiva siempre tendrá la forma

$$f_n^*(s_n) = \max_{x_n} \{f_n(s_n, x_n^*)\} \quad \text{o} \quad f_n^*(s_n) = \min_{x_n} \{f_n(s_n, x_n^*)\} \quad (1.5)$$

en donde $f_n(s_n, x_n)$ se escribe en términos de $s_n, x_n, f_{n+1}^*(s_{n+1})$ y tal vez alguna medida de la contribución inmediata de x_n a la función objetivo. Lo que hace que la expresión para $f_n^*(s_n)$ sea una relación recursiva es la inclusión de $f_{n+1}^*(s_{n+1})$ en el lado derecho, de manera que $f_n^*(s_n)$ está definida en términos de $f_{n+1}^*(s_{n+1})$.

8.- Cuando se usa esta relación recursiva, el procedimiento de solución comienza al final y se mueve hacia atrás etapa por etapa (encontrando cada vez la política óptima para esa etapa) hasta que se encuentra la política óptima desde la etapa inicial. Esta política óptima lleva de inmediato a una solución óptima para el problema completo, a saber x_1^* para el estado inicial s_1 , después x_2^* para el estado s_2 que resulta, luego x_3^* para el estado s_3 que resulta, y así sucesivamente hasta x_N^* para el estado s_N resultante.

Para todos los problemas de programación dinámica, se obtiene una tabla como la que se muestra en la Figura 1.1 para cada etapa ($n = N, N - 1, \dots, 1$).

en adelante. La optimización respecto a x_n proporciona entonces $f_n^*(s_n) = f_n(s_n, x_n^*)$. Una vez encontrados x_n y $f_n^*(s_n)$ para cada valor posible de s_n , el procedimiento de solución se mueve hacia atrás una etapa.

Una manera de clasificar los problemas de programación dinámica determinística es por la forma de la función objetivo. El objetivo puede ser minimizar la suma de las contribuciones de cada una de las etapas individuales, o maximizar esa suma, o bien minimizar el producto de los términos, etc. Otra clasificación se puede hacer en términos de la naturaleza del conjunto de estados en las respectivas etapas. En particular, los estados s_n pueden estar representados por una variable de estado discreta, o por una variable de estado continua.

1.2.5 LA PROGRAMACIÓN DINÁMICA PROBABILÍSTICA

La programación dinámica probabilística difiere de la determinística en que el estado de la siguiente etapa no está completamente determinado por el estado y la política de decisión de la etapa actual. En su lugar existe una distribución de probabilidad para determinar cuál será el siguiente estado. Sin embargo, esta distribución de probabilidad sí queda bien determinada por el estado y la política de decisión en la etapa actual. En la Figura 1.3 se describe con un diagrama la estructura básica que resulta en los problemas de programación dinámica probabilística.

En lo que se refiere a este diagrama, sea S el número de estados posibles en la etapa $n+1$ y etiquete estos estados al lado derecho por $1, 2, \dots, S$. El sistema cambia al estado i con probabilidad p_i ($i = 1, 2, \dots, S$) dados el estado s_n y la decisión x_n en la etapa n . Si el sistema cambia al estado i , C_i es la contribución de la etapa n a la función objetivo.

Cuando se expande la Figura 1.3 para incluir todos los estados y las decisiones posibles en todas las etapas, se obtiene lo que con frecuencia se conoce como un *árbol de decisión*. Si este árbol de decisión no es muy grande, proporciona una forma útil de resumir estas posibilidades.

Debido a la estructura probabilística, la relación entre $f_n(s_n, x_n)$ y $f_{n+1}^*(s_{n+1})$ necesariamente es más complicada que para el caso determinístico. La forma exacta de esta relación dependerá de la forma global de la función objetivo.

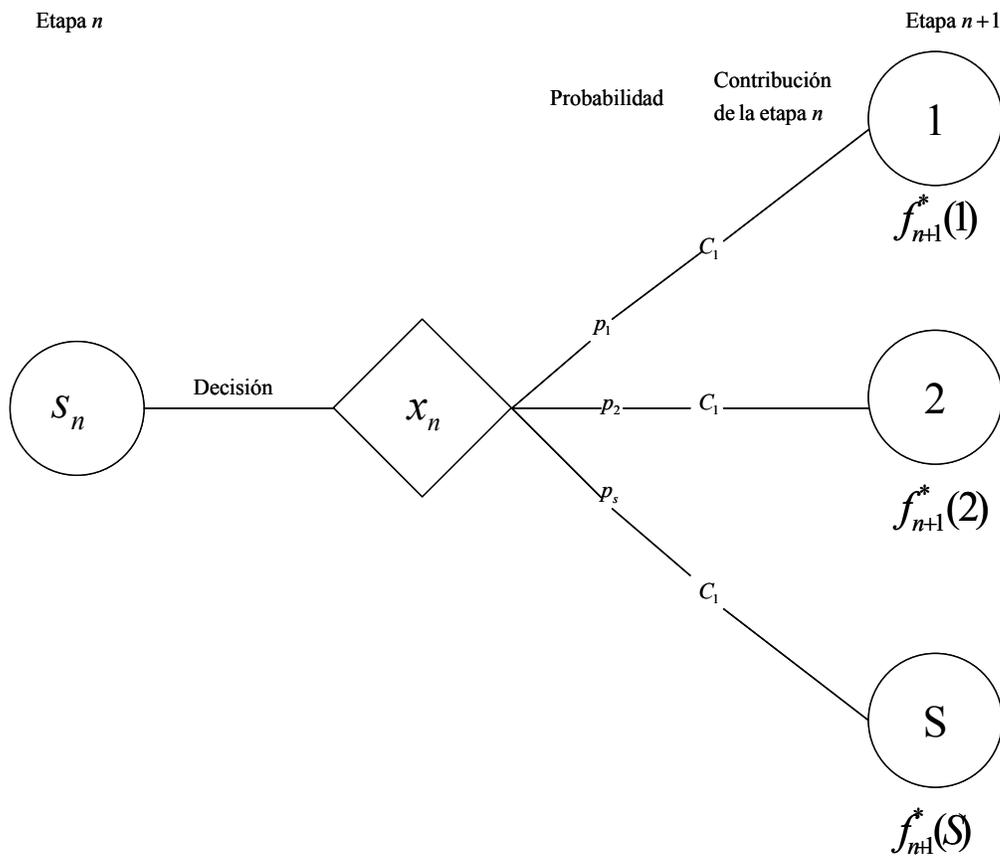


Figure 1.3: Estructura básica para programación dinámica probabilística

1.3 PROCESOS ESTOCÁSTICOS

1.3.1 INTRODUCCIÓN

Definición 1 *El conjunto de todos los posibles resultados de un experimento se llama espacio muestral del experimento aleatorio. El espacio muestral será denotado por el símbolo Ω y un elemento de Ω será denotado por ω . A los elementos de Ω se le llaman puntos muestrales.*

Cada vez que un experimento es ejecutado, exactamente uno de los posibles resultados ocurrirá. Usualmente no se conoce cuál de los posibles resultados ocurrirá. El investigador puede sentir que cada resultado tiene *una buena oportunidad de ocurrir* como cualquier otro o puede sentir que algunos resultados tienen una *mejor* oportunidad de ocurrir que otros.

Para trasladar *esta forma de sentir* en términos matemáticos precisos, el investigador puede asignar probabilidades a cada resultado. Matemáticamente, se puede definir una medida de probabilidad P sobre Ω de tal manera que sea consistente con las siguientes condiciones:

- 1) Para cada $\varpi_i \in \Omega$ existe un número no negativo asociado $p_i = P(\varpi_i)$.
- 2) $\sum_{\varpi_i \in \Omega} P(\varpi_i) = 1$.

Definición 2 *a) Si Ω es un espacio muestral finito o infinitamente numerable y si P es una medida de probabilidad sobre Ω , entonces a (Ω, P) se le llama espacio de probabilidad discreto.*

b) Sea (Ω, P) es un espacio de probabilidad, entonces un evento o suceso E es cualquier subconjunto de Ω . Si $E = (\varpi_1, \varpi_2, \dots, \varpi_k)$, entonces la probabilidad de E se define como

$$P(E) = \sum_{i=1}^k P(\varpi_i). \quad (1.6)$$

c) Se dice que dos eventos, A y B , son mutuamente excluyentes si no pueden ocurrir juntos. Esto se expresa escribiendo $A \cap B = \phi$.

A continuación presentamos lo tres principales axiomas de la probabilidad y algunos teoremas que nos ayudarán para mostrar resultados importantes para la siguiente sección.

AXIOMAS DE LA PROBABILIDAD

Axioma 1 *Para cada evento $A \subset \Omega$, $0 \leq P(A) \leq 1$.*

Axioma 2 $P(\Omega) = 1$.

Axioma 3 Si A y B son eventos mutuamente excluyentes, entonces

$$P(A \cup B) = P(A) + P(B)$$

Del axioma 3 se deduce que $P(\phi) = 0$.

Teorema 1 Si $A = A_1 \cup A_2 \cup \dots \cup A_n$ y A_1, A_2, \dots, A_n son mutuamente excluyentes, entonces

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_n). \quad (1.7)$$

Demostración. La demostración será por inducción.

Para $n = 2$, el caso se reduce al axioma 3.

Supongamos que el resultado es válido para $n = k$, es decir, si

$$A = A_1 \cup A_2 \cup \dots \cup A_k$$

y A_1, A_2, \dots, A_k son mutuamente excluyentes, entonces

$$P(A) = P(A_1) + P(A_2) + \dots + P(A_k).$$

Demostraremos que el resultado se cumple para $n = k + 1$.

Sea A_{k+1} un evento tal que $A_{k+1} \cap A_i = \phi$ para $i = 1, \dots, k$ y

$$B = A_1 \cup A_2 \cup \dots \cup A_k \cup A_{k+1} = A \cup A_{k+1}.$$

Es claro que A_{k+1} y A son mutuamente excluyentes, entonces, por el axioma 3 y la hipótesis de inducción, tenemos

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_k \cup A_{k+1}) &= P(B) \\ &= P(A \cup A_{k+1}) \\ &= P(A) + P(A_{k+1}) \\ &= P(A_1) + P(A_2) + \dots + P(A_k) + P(A_{k+1}). \end{aligned}$$

Con esto se concluye la demostración. ■

El espacio de probabilidad (Ω, P) contiene la información necesaria para estudiar las propiedades probabilísticas del experimento. Sin embargo, el análisis podría ser complicado por el hecho de que la descripción de los puntos muestrales podría estar en una forma que es familiar al científico experimental pero muy extraña para el probabilista. Para aliviar este problema de comunicación, los puntos en Ω son frecuentemente mapeados en los números reales. De esta manera, un número puede ser usado para representar un resultado experimental.

Definición 3 Una función que mapea un espacio muestral sobre los números reales se llama *variable aleatoria*. Las letras X, Y, Z serán usadas para denotar variables aleatorias.

De la definición anterior, vemos que definir una variable aleatoria sobre un espacio muestral puede servir para dos propósitos:

- i) Cada resultado es renombrado como un número real.
- ii) Alguna información ineficaz contenida en el resultado es convenientemente perdida en el mapeo de Ω en R .

Definición 4 *Un proceso estocástico es una familia de variables aleatorias definidas sobre algún espacio muestral Ω . Si hay una cantidad numerable de miembros en la familia, el proceso será denotado por X_1, X_2, X_3, \dots . Si hay una cantidad no numerable de miembros de la familia, el proceso será denotado por $\{X_t : t \geq 0\}$ o $\{X_t\}_{t \geq 0}$. En el primer caso, el proceso es llamado un proceso de tiempo discreto mientras que en el segundo caso es llamado un proceso de tiempo continuo.*

Un proceso estocástico se considera como una función de dos variables, $X_t(\varpi) = X(t, \varpi)$. Para t fija, la función es una variable aleatoria. Para ϖ fija, el resultado es una función de valores reales de t , llamada camino muestral.

Definición 5 *Al conjunto de valores distintos que asume un proceso estocástico se le llama espacio de estado. Si el espacio de estado de un proceso estocástico es numerable o finito, el proceso será llamado una cadena. El espacio de estado será denotado por S .*

Probabilidad condicional

La noción de probabilidad condicional es un instrumento básico de la teoría de probabilidades y, por desgracia, su gran simplicidad se ve a veces oscurecida por una terminología singularmente inadecuada.

Definición 6 *Sea (Ω, P) un espacio de probabilidad, E, F dos subconjuntos de Ω y $P(E \cap F)$ la probabilidad de que ambos eventos E y F ocurran. Si $P(F) \neq 0$, la probabilidad de E dado que F ha ocurrido se define como*

$$P(E|F) = \frac{P(E \cap F)}{P(F)}. \quad (1.8)$$

La cantidad así definida, $P(E|F)$, se llamará probabilidad de E bajo la hipótesis F (ver [6]). Cuando todos los puntos muestrales tienen probabilidades iguales, $P(E|F)$, es la razón del número de puntos muestrales que E y F tienen en común al número de puntos de F . La probabilidad condicional queda indefinida cuando $P(F)$ es cero. Esto no tiene consecuencias en el caso de espacios muestrales discretos, pero es importante en la teoría general.

Calcular probabilidades condicionales de varios eventos con respecto a una hipótesis particular F equivale a elegir F como un nuevo espacio muestral con probabilidades proporcionales a las originales; el factor de probabilidad $P(F)$ es necesario para que la probabilidad total del nuevo espacio muestral sea la unidad.

Esta formulación nos muestra que todos los teoremas generales de probabilidad también son válidos para probabilidades condicionales con respecto a cualquier hipótesis particular F .

Así tenemos dos maneras de calcular la probabilidad condicional $P(E|F)$:

a) Directamente considerando la probabilidad de E con respecto al espacio muestral reducido F .

b) Usando la definición anterior, donde $P(E \cap F)$ y $P(F)$ se calculan con respecto al espacio muestral original Ω .

A continuación daremos la definición de *partición* de un espacio muestral y algunos resultados preliminares para demostrar el "*Teorema de Bayes*", el cual involucra a la probabilidad condicional.

Definición 7 Decimos que los eventos B_1, B_2, \dots, B_K representan una *partición del espacio muestral* Ω si:

a) $B_i \cap B_j = \phi$ para todo $i \neq j$.

b) $\bigcup_{i=1}^k B_i = \Omega$.

c) $P(B_i) > 0$ para todo i . En otras palabras: cuando se efectúa un experimento, *ocurre uno y sólo uno* de los eventos B_i .

Teorema 2 (*Teorema de la probabilidad total*) Sea A un evento de Ω y sea B_1, B_2, \dots, B_K una *partición* de Ω . Entonces

$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_k)P(B_k). \quad (1.9)$$

Demostración. Por definición de *partición*, $B_i \cap B_j = \phi$ para todo $i \neq j$, de aquí obtenemos que $(A \cap B_i) \cap (A \cap B_j) = \phi$ para todo $i \neq j$. Ahora, podemos escribir A como sigue

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^k B_i \right) = (A \cap B_1) \cup (A \cap B_2) \cup \dots \cup (A \cap B_k).$$

Entonces, aplicando el Teorema 1.1 obtenemos

$$\begin{aligned} P(A) &= P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_k) \\ &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_k)P(B_k) \end{aligned}$$

Con esto concluimos la demostración. ■

Teorema 3 (Teorema de Bayes). Sea B_1, B_2, \dots, B_K una partición del espacio muestral Ω y sea A un evento asociado con Ω . Entonces

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^k P(A|B_j)P(B_j)}. \quad (1.10)$$

Demostración. Por el teorema anterior,

$$\begin{aligned} P(A) &= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \dots + P(A|B_k)P(B_k) \\ &= \sum_{j=1}^k P(A|B_j)P(B_j). \end{aligned}$$

Por lo tanto

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A \cap B_i)}{\sum_{j=1}^k P(A|B_j)P(B_j)} = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^k P(A|B_j)P(B_j)}$$

Así, queda demostrado el teorema de Bayes. ■

1.3.2 LA PROPIEDAD DE MARKOV Y MATRICES DE TRANSICIÓN

A un proceso estocástico le podemos poner ciertas restricciones como:

- 1) Que el proceso sea de tiempo discreto.
- 2) Que tenga un espacio de estado numerable o finito.
- 3) Que el proceso satisfaga la propiedad de Markov.

A los procesos estocásticos que satisfacen las 3 restricciones dadas se les llama Cadenas de Markov

Definición 8 Un proceso estocástico $\{X_t\}$ $t = 0, 1, 2, \dots$, con espacio de estado $S = \{0, 1, 2, 3, \dots\}$ se dice que satisface la propiedad de Markov si para cada t y todos los estados i_0, i_2, \dots, i_t se cumple que

$$P[X_{t+1} = j | X_t = i, X_{t-1} = i_{t-1}, \dots, X_0 = i_0] = P[X_{t+1} = j | X_t = i]. \quad (1.11)$$

Esta propiedad de Markov es equivalente a establecer que la probabilidad condicional de cualquier evento futuro dado cualquier evento pasado y el estado actual $X_t = i$, es independiente del evento pasado y sólo depende del estado actual del proceso. Las probabilidades condicionales $P[X_{t+1} = j | X_t = i]$ se llaman probabilidades de transición de un paso.

Definición 9 Si para cada i y j

$$P[X_{t+1} = j | X_t = i] = P[X_1 = j | X_0 = i] \quad \text{para toda } t = 0, 1, \dots, \quad (1.12)$$

entonces se dice que las probabilidades de transición (de un paso) son estacionarias y se denotan por p_{ij} .

Así, tener probabilidades de transición estacionarias implican que las probabilidades de transición (de un paso) no cambian con el tiempo. La existencia de probabilidades de transición estacionarias (de un paso) también implican que para i, j y n ($n = 0, 1, 2, \dots$)

$$P[X_{t+n} = j | X_t = i] = P[X_n = j | X_0 = i] \quad \text{para toda } t = 0, 1, \dots \quad (1.13)$$

Generalmente se denotan por $p_{ij}^{(n)}$ y se llaman probabilidades de transición de n pasos. Para $n = 0$, $p_{ij}^{(0)}$ es sólo $P(X_0 = j | X_0 = i)$ y así es igual a 1 cuando $i = j$ y es 0 cuando $i \neq j$. Para $n = 1$, $p_{ij}^{(1)}$ es simplemente la probabilidad de transición de un paso, p_{ij} .

Una forma de representar las probabilidades de transición de n pasos es la forma matricial

$$P^{(n)} = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ M \end{array} \begin{array}{c} 0 \quad \dots \quad M \\ \begin{array}{|c|c|c|} \hline p_{00}^{(n)} & \dots & p_{0M}^{(n)} \\ \hline p_{10}^{(n)} & \dots & p_{1M}^{(n)} \\ \hline \cdot & & \cdot \\ \hline \cdot & & \cdot \\ \hline p_{M0}^{(n)} & \dots & p_{MM}^{(n)} \\ \hline \end{array} \end{array} \quad (1.14)$$

o, en forma equivalente,

$$P^{(n)} = \begin{pmatrix} p_{00}^{(n)} & \dots & p_{0M}^{(n)} \\ \dots & \dots & \dots \\ p_{M0}^{(n)} & \dots & p_{MM}^{(n)} \end{pmatrix}. \quad (1.15)$$

1.3.3 ECUACIONES DE CHAPMAN-KOLMOGOROV

Las ecuaciones de Chapman-Kolmogorov proporcionan un método para calcular probabilidades de transición de n pasos:

$$p_{ij}^{(n)} = \sum_{k=0}^M p_{ik}^{(m)} p_{kj}^{(n-m)} \quad \text{para toda } i, j \text{ y } n \text{ con } 0 \leq m \leq n. \quad (1.16)$$

Estas ecuaciones señalan que al ir del estado i al estado j en n pasos, el proceso estará en algún estado k después de exactamente m ($m < n$) pasos.

Así, $p_{ik}^{(m)} p_{kj}^{(n-m)}$ es sólo la probabilidad condicional de que si se comienza en el estado i , el proceso vaya al estado k después de m pasos y después al estado j en $n - m$ pasos. A continuación analizaremos los casos especiales cuando $m = 1$ y $m = n - 1$.

Para $m = 1$

$$p_{ij}^{(n)} = \sum_{k=0}^M p_{ik} p_{kj}^{(n-1)}. \quad (1.17)$$

Para $m = n - 1$

$$p_{ij}^{(n)} = \sum_{k=0}^M p_{ik}^{(n-1)} p_{kj} \quad (1.18)$$

para toda i, j y n .

Proposición 1 *La matriz de probabilidades de transición de n pasos se puede obtener de $P^{(n)} = P^n$.*

Demostración. Sea

$$P^{(n)} = \begin{pmatrix} p_{00}^{(n)} & p_{01}^{(n)} & \cdots & p_{0M}^{(n)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0}^{(n)} & p_{M1}^{(n)} & \cdots & p_{MM}^{(n)} \end{pmatrix}$$

la matriz de transición de n pasos y

$$P = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0M} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0} & p_{M1} & \cdots & p_{MM} \end{pmatrix}$$

la matriz de transición de un paso.

La demostración será por inducción.

i) Para $n = 1$ se tiene que

$$P^{(1)} = \begin{pmatrix} p_{00}^{(1)} & p_{01}^{(1)} & \cdots & p_{0M}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0}^{(1)} & p_{M1}^{(1)} & \cdots & p_{MM}^{(1)} \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0M} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0} & p_{M1} & \cdots & p_{MM} \end{pmatrix} = P$$

Por lo tanto, se cumple el resultado para $n = 1$.

ii) Supongamos que se cumple la relación para $n = k$, es decir, $P^{(k)} = P^k$.

Demostraremos que la relación es verdadera para $n = k + 1$.

$$\begin{aligned}
P^{k+1} &= P^k P = P^{(k)} P \\
&= \begin{pmatrix} p_{00}^{(k)} & p_{01}^{(k)} & \cdots & p_{0M}^{(k)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0}^{(k)} & p_{M1}^{(k)} & \cdots & p_{MM}^{(k)} \end{pmatrix} \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0M} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0} & p_{M1} & \cdots & p_{MM} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{j=0}^M p_{0j}^{(k)} p_{j0} & \sum_{j=0}^M p_{0j}^{(k)} p_{j1} & \cdots & \sum_{j=0}^M p_{0j}^{(k)} p_{jM} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{j=0}^M p_{Mj}^{(k)} p_{j0} & \sum_{j=0}^M p_{Mj}^{(k)} p_{j1} & \cdots & \sum_{j=0}^M p_{Mj}^{(k)} p_{jM} \end{pmatrix} \\
&= \begin{pmatrix} p_{00}^{(k+1)} & p_{01}^{(k+1)} & \cdots & p_{0M}^{(k+1)} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ p_{M0}^{(k+1)} & p_{M1}^{(k+1)} & \cdots & p_{MM}^{(k+1)} \end{pmatrix} = P^{(k+1)}
\end{aligned}$$

esto último se debe a las ecuaciones de Chapman-Kolmogorov.

Así, tenemos que $P^n = P^{(n)}$, para todo $n \geq 1$. ■

Si se desea la probabilidad incondicional $P[X_n = j]$, es necesario que se especifique la distribución de probabilidad del estado inicial. Tenemos así que

$$P[X_n = j] = P[X_0 = 0]p_{0j}^{(n)} + P[X_0 = 1]p_{1j}^{(n)} + \dots + P[X_0 = M]p_{Mj}^{(n)} \quad (1.19)$$

1.3.4 CLASIFICACIÓN DE ESTADOS EN UNA CADENA DE MARKOV

A continuación daremos los principales conceptos de los tipos de estados así como de algunos resultados que se desprenden de ellos.

Definición 10 Se dice que el estado j es accesible desde el estado i si $p_{ij}^{(n)} > 0$ para alguna $n \geq 0$.

El que el estado j sea accesible desde el estado i significa que es posible que el sistema llegue eventualmente al estado j si comienza en el estado i . En general, una condición suficiente para que todos los estados sean accesibles es que exista un valor de n para el que $p_{ij}^{(n)} > 0$ para todo i y j .

Definición 11 Si el estado j es accesible desde el estado i y el estado i es accesible desde el estado j , entonces se dice que los estados i y j se comunican.

En general: 1) Cualquier estado se comunica consigo mismo; 2) si el estado i se comunica con el estado j , entonces el estado j se comunica con el estado i y 3) Si el estado i se comunica con el estado j y el estado j se comunica con el estado k , entonces el estado i se comunica con el estado k .

Como resultado de estas propiedades de comunicación, se puede hacer una partición del espacio de estados en clases ajenas, en donde se dice que dos estados que se comunican pertenecen a la misma clase. Así, los estados de una cadena de Markov pueden construir una o más clases ajenas (una clase puede consistir en un solo estado). Así, tenemos la siguiente definición.

Definición 12 *Se dice que una cadena de Markov es irreducible si existe sólo una clase, es decir, si todos los estados se comunican.*

Con frecuencia es útil hablar sobre si un proceso que comienza en el estado i regresará alguna vez a este estado, por tal razón introducimos los conceptos de estado recurrente y transitorio.

Definición 13 *Sea f_{ii} la probabilidad de que el proceso regrese al estado i dado que comienza en el estado i . El estado i se llama estado recurrente si $f_{ii} = 1$ y es transitorio si $f_{ii} < 1$.*

Un caso especial de estado recurrente es un estado absorbente.

Definición 14 *Se dice que un estado i es absorbente si la probabilidad de transición p_{ii} sea igual a 1.*

Determinar si un estado es recurrente o transitorio evaluando f_{ii} no es sencillo. Sin embargo, es posible determinar algunas propiedades de f_{ii} que pueden ayudar a determinar su valor. Si un proceso de Markov se encuentra en el estado i y este estado es recurrente, la probabilidad de que el proceso regrese al estado i es 1. Como el proceso es una cadena de Markov, esto es equivalente a que el proceso comience una vez más en el estado i y que con probabilidad de 1, regrese una vez más a ese estado. La repetición de este argumento lleva a la conclusión de que llegará al estado i un número infinito de veces. Así, un estado recurrente tiene la propiedad de que el número esperado de periodos que el proceso está en el estado i es infinito.

Si un proceso de Markov se encuentra en el estado i , y este estado es transitorio, entonces la probabilidad de que regrese al estado i es f_{ii} y la probabilidad de que no regrese es $1 - f_{ii}$. Se puede demostrar que el número esperado de periodos que el proceso se encuentra en el estado i es finito y está dado por

$$\frac{1}{1 - f_{ii}}. \quad (1.20)$$

Por lo tanto, se concluye que el estado i es recurrente si y sólo si el número esperado de periodos que el proceso se encuentra en el estado i es infinito, dado que el proceso comenzó en el estado i .

Con el fin de calcular el número esperado de periodos que el proceso se encuentra en el estado i dado que $X_0 = i$, se define

$$B_n = \begin{cases} 1 & \text{si } X_n = i \\ 0 & \text{si } X_n \neq i \end{cases} . \quad (1.21)$$

La cantidad

$$\sum_{n=1}^{\infty} (B_n | X_0 = i) \quad (1.22)$$

representa el número de periodos que el proceso está en i dado que $X_0 = i$. Por lo tanto, su valor esperado está dado por

$$\begin{aligned} E \left(\sum_{n=1}^{\infty} B_n | X_0 = i \right) &= \sum_{n=1}^{\infty} E (B_n | X_0 = i) \\ &= \sum_{n=1}^{\infty} P(X_n = i | X_0 = i) \\ &= \sum_{n=1}^{\infty} p_{ii}^{(n)} . \end{aligned} \quad (1.23)$$

Se ha demostrado que el estado i es recurrente si, y sólo si,

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty . \quad (1.24)$$

Este resultado se puede usar para demostrar que la recurrencia es una propiedad de clase, es decir, todos los estados en una clase son recurrentes o transitorios. Más aun, en una cadena de Markov de estado finito, no todos los estados pueden ser transitorios. Entonces todos los estados de una cadena de Markov de estado finito irreducible son recurrentes.

Otra propiedad útil de las cadenas de Markov es el de las periodicidades.

Definición 15 *El estado j tiene periodo d si las siguientes dos condiciones se cumplen:*

- (i) $p_{jj}^{(n)} = 0$ a menos que $n = md$ para algún entero positivo m y
- (ii) d es el máximo entero con esta propiedad.

El estado j es llamado *aperiódico* cuando $d = 1$.

Al igual que la recurrencia es una propiedad de clase, se puede demostrar que la periodicidad también es una propiedad de clase. Esto es, si el estado i tiene periodo t , todos los estados en esa clase tienen periodo t .

Una última propiedad de las cadenas de Markov pertenece a una nueva clasificación de los estados recurrentes.

Definición 16 *Se dice que un estado i es recurrente positivo si, comenzando en el estado i , el tiempo esperado para que el proceso regrese al estado i es finito. De igual manera, un estado recurrente i , es recurrente nulo si, comenzando en el estado i el tiempo esperado para que el proceso regrese al estado i es infinito. Los estados recurrentes positivos que son aperiódicos se llaman estados ergódicos. Se dice que una cadena de Markov es ergódica si todos sus estados son ergódicos.*

1.3.5 TIEMPOS DE PRIMERA PASADA

Con frecuencia es conveniente poder hacer afirmaciones en términos de probabilidades sobre el número de transiciones que hace el proceso al ir de un estado i a un estado j por primera vez. Este lapso se llama *tiempo de primera pasada* al ir del estado i al estado j . Cuando $j = i$, este tiempo de primera pasada es justo el número de transiciones hasta que el proceso regrese al estado inicial i . En este caso, el tiempo de primera pasada se llama *tiempo de recurrencia* para el estado i .

En general, los tiempos de primera pasada son variables aleatorias y, por lo tanto, tienen una distribución de probabilidad asociada a ellos. Estas distribuciones de probabilidad dependen de las probabilidades de transición del proceso. En particular, $f_{ij}^{(n)}$ denota la probabilidad de que el tiempo de primera pasada del estado i al j sea igual a n . Se puede demostrar que estas probabilidades satisfacen las siguientes relaciones recursivas:

$$\begin{aligned} f_{ij}^{(1)} &= p_{ij}^{(1)} = p_{ij} \\ f_{ij}^{(2)} &= p_{ij}^{(2)} - f_{ij}^{(1)} p_{jj} \\ &\vdots \\ f_{ij}^{(n)} &= p_{ij}^{(n)} - f_{ij}^{(1)} p_{jj}^{(n-1)} - f_{ij}^{(2)} p_{jj}^{(n-2)} - \dots - f_{ij}^{(n-1)} p_{jj}. \end{aligned} \quad (1.25)$$

Entonces se puede calcular la probabilidad de un tiempo de primera pasada del estado i al j en n pasos, de manera recursiva, a partir de las probabilidades de transición de un paso.

Para i y j fijos, las $f_{ij}^{(n)}$ son números no negativos tales que

$$\sum_{n=1}^{\infty} f_{ij}^{(n)} \leq 1. \quad (1.26)$$

Desafortunadamente, esta suma puede ser estrictamente menor que 1, lo que significa que un proceso que al iniciar se encuentra en el estado i puede no llegar nunca al estado j . Cuando la suma sí es igual a 1, las $f_{ij}^{(n)}$ (para $n = 1, 2, \dots$) pueden considerarse como una distribución de probabilidad para la variable aleatoria, el tiempo de primera pasada.

Mientras que puede ser difícil calcular $f_{ij}^{(n)}$ para toda n , es relativamente sencillo obtener el tiempo esperado de primera pasada del estado i al estado j . Sea μ_{ij} este valor esperado, que se define como

$$\mu_{ij} = \begin{cases} \infty & \text{si } \sum_{n=1}^{\infty} f_{ij}^{(n)} < 1 \\ \sum_{n=1}^{\infty} n f_{ij}^{(n)} & \text{si } \sum_{n=1}^{\infty} f_{ij}^{(n)} = 1 \end{cases}. \quad (1.27)$$

Siempre que $\sum_{n=1}^{\infty} f_{ij}^{(n)} = 1$, entonces μ_{ij} satisface, de manera única, la ecuación

$$\mu_{ij} = 1 + \sum_{k \neq j} p_{ik} \mu_{kj} \quad (1.28)$$

Cuando $i = j$, μ_{ij} se llama *tiempo esperado de recurrencia*.

1.3.6 PROPIEDADES A LARGO PLAZO DE LAS CADENAS DE MARKOV

Probabilidades de estado estable

Para analizar el concepto de *probabilidades de estado estable* estableceremos las siguientes condiciones:

Para una cadena de Markov irreducible ergódica el $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$ existe y es independiente de i . Más aun, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j > 0$, en donde las π_j satisfacen de manera única las siguientes *ecuaciones de estado estable*

$$\begin{aligned} \pi_j &= \sum_{i=0}^M \pi_i p_{ij} \quad \text{para } j = 0, 1, \dots, M, \\ \sum_{j=0}^M \pi_j &= 1. \end{aligned} \quad (1.29)$$

Las π_j se llaman *probabilidades de estado estable* de la cadena de Markov y son iguales al inverso del tiempo esperado de recurrencia, es decir,

$$\pi_j = \frac{1}{\mu_{jj}}, \quad \text{para } j = 0, 1, \dots, M. \quad (1.30)$$

El término probabilidad de *estado estable* significa que la probabilidad de encontrar el proceso en un cierto estado, por ejemplo j , después de un número grande de transiciones tiende al valor π_j y es independiente de la distribución de probabilidad inicial definida para los estados. Es importante hacer notar que la probabilidad de estado estable no significa que el proceso se establezca en un estado. Por el contrario, el proceso continúa haciendo transiciones de un estado a otro y en cualquier paso n la probabilidad de transición del estado i al estado j es todavía p_{ij} .

También se pueden interpretar las π_j como probabilidades estacionarias (sin que se confundan con las probabilidades de transición estacionarias) en el siguiente sentido. Si la probabilidad absoluta de encontrarse en el estado j está dada por π_j (esto es $P(X_0 = j) = \pi_j$) para toda j , entonces la probabilidad absoluta de encontrar el proceso en el estado j en el tiempo $n = 1, 2, \dots$ también está dada por π_j (es decir, $P(X_n = j) = \pi_j$).

Debe observarse que las ecuaciones de estado estable consisten en $M + 2$ ecuaciones con $M + 1$ incógnitas. Como el sistema tiene una solución única, al menos una de las ecuaciones debe ser redundante, por lo que se puede eliminar. No puede ser la ecuación

$$\sum_{j=0}^M \pi_j = 1 \quad (1.31)$$

porque $\pi_j = 0$ para toda j satisface las otras $M + 1$ ecuaciones. Es más, las ecuaciones de las otras $M + 1$ ecuaciones de estado estable tienen una solución única con una constante multiplicativa y es la ecuación final la que fuerza la solución a ser una distribución de probabilidad.

Existen otros resultados importantes respecto a las probabilidades de estado estable. En particular, si i y j son estados recurrentes que pertenecen a clases distintas, entonces $p_{ij}^{(n)} = 0$, para toda n .

De manera parecida, si j es un estado transitorio, entonces $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$, para toda i . Este resultado significa que la probabilidad de encontrar el proceso en un estado transitorio después de un número grande de transiciones es cero.

Costo promedio esperado por unidad de tiempo

En la subsección anterior se vio las cadenas de Markov cuyos estados son ergódicos (recurrentes positivos y aperiódicos). Si se relaja el requerimiento de que los estados sean aperiódicos, entonces el límite $\lim_{n \rightarrow \infty} p_{ij}^{(n)}$ puede no existir. Sin embargo, el siguiente límite siempre existe para una cadena de Markov irreducible con estados recurrentes positivos

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} \right) = \pi_j \quad (1.32)$$

en donde las π_j satisfacen las ecuaciones de estado estable dadas en la subsección anterior. Este resultado es muy importante al calcular el costo promedio a la larga por unidad de tiempo, asociado a una cadena de Markov. Suponga que se incurre en un costo (u otra función de penalización) $C(X_t)$ cuando el proceso se encuentra en el estado X_t en el tiempo t , para $t = 0, 1, 2, \dots$. Note que $C(X_t)$ es una variable aleatoria que toma cualquiera de los valores $C(0), C(1), \dots, C(M)$ y que la función $C(\cdot)$ es independiente de t . El costo promedio esperado en el que se incurre a lo largo de los primeros n periodos está dado por la expresión

$$E \left[\frac{1}{n} \sum_{t=1}^n C(X_t) \right]. \quad (1.33)$$

Usando el resultado de que

$$\lim_{n \rightarrow \infty} \left\{ \frac{1}{n} \sum_{k=1}^n p_{ij}^{(k)} \right\} = \pi_j \quad (1.34)$$

se puede demostrar que el costo promedio esperado por unidad de tiempo (a la larga), está dado por

$$\lim_{n \rightarrow \infty} E \left[\frac{1}{n} \sum_{t=1}^n C(X_t) \right] = \sum_{j=0}^M \pi_j C(j). \quad (1.35)$$

Otra medida del costo promedio esperado (a la larga) por unidad de tiempo es el costo promedio real por unidad de tiempo (a la larga). Se puede demostrar que esta última medida está dada por

$$\lim_{n \rightarrow \infty} \left[\frac{1}{n} \sum_{t=1}^n C(X_t) \right] = \sum_{j=0}^M \pi_j C(j) \quad (1.36)$$

para casi todas las trayectorias del proceso. Así, ambas medidas llevan al mismo resultado. Estos resultados también se pueden usar para interpretar el significado de las π_j . Para hacer esta interpretación, sea

$$C(X_t) = \begin{cases} 1 & \text{si } X_t = j \\ 0 & \text{si } X_t \neq j. \end{cases} \quad (1.37)$$

La fracción esperada del número de veces (a la larga) que el sistema se encuentra en el estado j está dada entonces por

$$\lim_{n \rightarrow \infty} E \left[\frac{1}{n} \sum_{t=1}^n C(X_t) \right]$$
$$= \lim_{n \rightarrow \infty} E (\text{fracción de veces que el sistema está en el estado } j) = \pi_j. \quad (1.38)$$

De igual manera, π_j se puede interpretar también como la fracción o porcentaje real (a la larga) del número de veces que el sistema se encuentra.

CAPÍTULO 2

TEORÍA DE INVENTARIOS

2.1 COMPONENTES DE LOS MODELOS DE INVENTARIOS

Las políticas de inventarios afectan las ganancias. La elección entre una política y otra depende de su rentabilidad relativa. Algunos de estos costos que determinan esta rentabilidad son:

- 1.- Los costos de ordenar o fabricar.
- 2.- Los costos de mantener o almacenar.
- 3.- Los costos de penalización por faltantes o demanda insatisfecha.

Otros costos relevantes incluyen:

- 4.- Los ingresos.
- 5.- Los costos de recuperación o salvamento.
- 6.- Las tasas de descuento.

El *costo de ordenar* o *fabricar* una cantidad z se puede representar por una función $c(z)$ (véase [5] y [10]). La forma más sencilla de esta función es aquella que es directamente proporcional a la cantidad ordenada o producida, es decir, cz , donde c es el precio unitario pagado. Otra suposición es que $c(z)$ se compone de dos partes: un término que es directamente proporcional a la cantidad ordenada o producida y un término que es una constante K para $z > 0$ y 0 para $z = 0$.

$$c(z) = \begin{cases} 0 & \text{si } z = 0 \\ K + cz & \text{si } z > 0 \end{cases} \quad (2.1)$$

El *costo de mantener inventario (costo de almacenar)* representa los costos asociados con el almacenamiento del inventario hasta que se vende o se usa. Puede incluir el costo del capital invertido, del espacio, seguros, protección e impuestos atribuibles al almacenamiento.

El *costo de penalización por faltantes (costo de la demanda insatisfecha)* surge cuando la cantidad que se requiere de un bien (demanda) es mayor que el inventario disponible.

Este costo depende de cuál de los dos casos siguientes se aplica.

En un caso, llamado *con faltantes* la demanda excesiva no se pierde sino que se queda pendiente hasta que se pueda satisfacer con el siguiente reabastecimiento.

En un segundo caso, llamado *sin faltantes*, si ocurre un exceso de demanda sobre el inventario disponible, el distribuidor no puede esperar a la siguiente entrega normal para reabastecer el inventario. Ya que 1) el exceso de demanda se satisfaga mediante un envío prioritario o 2) no se cumpla. En la situación 1), el costo por faltantes se puede interpretar como el costo de envío prioritario. Para la situación 2), este costo por faltantes se puede ver como la pérdida en la que se incurre por no satisfacer la demanda, más el costo de perder negocios futuros debido a la pérdida de la buena voluntad.

El *ingreso* puede o no incluirse en el modelo. Si no se incluye, entonces la pérdida del ingreso debe incluirse en el costo de penalización por faltantes siempre que la empresa no pueda cumplir con esa demanda y se pierda la venta.

El *valor de recuperación o salvamento de un producto* es el valor de un artículo sobrante cuando no se requiere más del inventario. El negativo del valor de recuperación se llama *costo de recuperación*. Se supondrá en adelante que cualquier costo de recuperación se incorpora al costo de mantener.

Los modelos de inventarios, por lo general, se clasifican según si se conoce la demanda para el periodo (demanda determinística) o si se trata de una variable aleatoria que tiene una distribución de probabilidad conocida (demanda no determinística o aleatoria). Otra clasificación posible se relaciona con la forma en que se revisa el inventario, ya sea continua o periódicamente. En un sistema de revisión continua, se hacen los pedidos en el momento en que el inventario baja del punto de reorden que se determine, mientras que en el caso de revisión periódica se verifica el nivel de inventario en intervalos discretos (al final de cada semana, mes, etc.)

2.2 MODELOS DETERMINÍSTICOS

2.2.1 REVISIÓN CONTINUA, DEMANDA UNIFORME Y NO SE PERMITEN FALTANTES

El modelo del lote económico es un modelo sencillo que representa el problema de inventario en el que los niveles de existencias se reducen con el tiempo y después se reabastecen con la llegada de nuevas unidades.

Se supone que los artículos bajo consideración se sacarán en forma continua a una tasa constante conocida denotada por a unidades por unidad de tiempo al mes. Se supone también que el inventario se reabastece produciendo u ordenando un lote de tamaño fijo (Q unidades) y que las Q unidades llegan juntas en el tiempo deseado. Los únicos costos que se considerarán son:

K = Costo de preparación para producir u ordenar un lote.

c = El costo de producir o comprar cada unidad.

h = El costo de mantener el inventario por unidad, por unidad de tiempo.

El objetivo consiste en determinar con qué frecuencia y en qué cantidad reabastecer el inventario de manera que se minimice la suma de estos costos por unidad de tiempo. Se supondrá revisión continua, por lo que el inventario se puede reabastecer cuando el nivel baje lo suficiente. Con la tasa de demanda fija, se pueden evitar los faltantes reabasteciendo el inventario cada vez que el nivel baje a cero, y esto también minimizará el costo de mantener.

Para iniciar el análisis, calculamos la longitud del ciclo. La longitud del ciclo se calcula por $\frac{Q}{a}$. El costo total T por unidad de tiempo se obtiene a partir de las siguientes componentes.

$$\text{Costo por ciclo de producción u ordenar} = K + cQ \quad (2.2)$$

El nivel de inventario promedio durante un ciclo es $\frac{Q+0}{2} = \frac{Q}{2}$ unidades por unidad de tiempo y el costo correspondiente es $h\frac{Q}{2}$ por unidad de tiempo. Como la longitud del ciclo es $\frac{Q}{a}$ entonces tenemos que

$$\text{Costo por ciclo de mantener inventario} = \frac{hQ^2}{2a} \quad (2.3)$$

Por lo tanto

$$\text{Costo total por ciclo} = K + cQ + \frac{hQ^2}{2a} \quad (2.4)$$

y el costo total T por unidad de tiempo es

$$T = \frac{K + cQ + hQ^2/2a}{Q/a} = \frac{aK}{Q} + ac + \frac{hQ}{2} \quad (2.5)$$

El valor de Q , digamos Q^* , que minimiza a T se encuentra estableciendo la primera derivada igual a cero y observando que la segunda derivada sea positiva.

Así

$$\frac{dT}{dQ} = -\frac{aK}{Q^2} + \frac{h}{2} = 0 \quad (2.6)$$

de manera que

$$Q^* = \sqrt{\frac{2aK}{h}} \quad (2.7)$$

que es la familiar fórmula del lote económico. De igual manera, el tiempo que toma obtener este valor de Q^* , llámese t^* , está dado por

$$t^* = \frac{Q^*}{a} = \sqrt{\frac{2K}{ah}}. \quad (2.8)$$

2.2.2 REVISIÓN CONTINUA, DEMANDA UNIFORME Y SE PERMITEN FALTANTES

Sea

p = costo del faltante por unidad de demanda insatisfecha por unidad de tiempo.

S = nivel de inventario justo después de recibir un lote de Q unidades.

$Q - S$ = faltante en inventario justo antes de recibir un lote de Q unidades.

Ahora, el costo total por unidad de tiempo se obtiene a partir de las siguientes componentes

$$\text{Costo de producir u ordenar por ciclo} = K + cQ \quad (2.9)$$

Durante el ciclo, el nivel de inventario es positivo durante un tiempo $\frac{S}{a}$. El nivel de inventario promedio durante este tiempo es $\frac{S+0}{2} = \frac{S}{2}$ artículos por unidad de tiempo y el costo correspondiente es $\frac{hS}{2}$ por unidad de tiempo. Entonces

$$\text{Costo de mantener el inventario por ciclo} = \frac{hS}{2} \left(\frac{S}{a}\right) = \frac{hS^2}{2a} \quad (2.10)$$

De manera similar, los faltantes ocurren durante un tiempo $\frac{Q-S}{a}$. La cantidad promedio del faltantes durante este tiempo es $\frac{0+Q-S}{2} = \frac{Q-S}{2}$ artículos por

unidad de tiempo y el costo correspondiente es $p(\frac{Q-S}{2})$ por unidad de tiempo. Así

$$\text{Costo de faltantes por ciclo} = \frac{p(Q-S)}{2} \frac{Q-S}{a} = \frac{p(Q-S)^2}{2a}. \quad (2.11)$$

Por lo tanto

$$\text{Costo total por ciclo} = K + cQ + \frac{hS^2}{2a} + \frac{p(Q-S)^2}{2a} \quad (2.12)$$

y el costo total T por unidad de tiempo es

$$\begin{aligned} T &= \frac{K + cQ + hS^2/2a + p(Q-S)^2/2a}{Q/a} \\ &= \frac{aK}{Q} + ac + \frac{hS^2}{2Q} + \frac{p(Q-S)^2}{2Q}. \end{aligned} \quad (2.13)$$

Este modelo tiene dos variables de decisión S y Q , y los valores óptimos se encuentran estableciendo las derivadas parciales $\frac{\partial T}{\partial S}$ y $\frac{\partial T}{\partial Q}$ igual a cero. Entonces

$$\begin{aligned} \frac{\partial T}{\partial S} &= \frac{hS}{Q} - \frac{p(Q-S)}{Q} = 0 \\ \frac{\partial T}{\partial Q} &= -\frac{aK}{Q^2} - \frac{hS^2}{2Q^2} + \frac{p(Q-S)}{Q} - \frac{p(Q-S)^2}{2Q^2} = 0. \end{aligned} \quad (2.14)$$

Al resolver estas ecuaciones simultáneas se obtiene

$$S^* = \sqrt{\frac{2aK}{h}} \sqrt{\frac{p}{p+h}} \quad Q^* = \sqrt{\frac{2aK}{h}} \sqrt{\frac{p+h}{p}}. \quad (2.15)$$

La longitud óptima del ciclo t^* está dada por

$$t^* = \frac{Q^*}{a} = \sqrt{\frac{2K}{ah}} \sqrt{\frac{p+h}{p}} \quad (2.16)$$

El faltante máximo es

$$Q^* - S^* = \sqrt{\frac{2aK}{p}} \sqrt{\frac{h}{p+h}}. \quad (2.17)$$

2.3 MODELOS ESTOCÁSTICOS

2.3.1 MODELO DE UN PERIODO SIN COSTO FIJO

Se describirán brevemente los supuestos del modelo.

- 1.- Se está realizando la planeación nada más para un periodo.
- 2.- La demanda D en este periodo es una variable aleatoria con una distribución de probabilidad conocida.
- 3.- No hay inventario inicial.
- 4.- La decisión a tomar es el valor de y , el número de unidades que se deben comprar o producir al principio del periodo para el inventario.
- 5.- El objetivo es minimizar el costo total esperado, donde los componentes del costo son:

c = costo unitario de comprar o producir cada unidad.

h = costo de mantener por unidad que queda al final del periodo (incluye el costo de almacenaje menos el valor de recuperación).

p = costo por faltantes por unidad de demanda no satisfecha (incluye el rendimiento perdido y el costo de la pérdida de la buena voluntad del cliente)

Análisis del Modelo.

La decisión sobre el valor de y , la cantidad de inventario a adquirir, depende fuertemente de la distribución de probabilidad de la demanda D . Lo que se busca es superar la demanda esperada, pero sin alcanzar la demanda máxima posible. Para ello es necesario tener un balance o trueque entre:

- 1.- El riesgo de una escasez que implica costos por faltantes.
- 2.- El riesgo de tener un excedente e incurrir en los costos desaprovechados de ordenar y almacenar más unidades de las necesarias.

Esto se logra minimizando el valor esperado de las sumas de estos costos.

La cantidad vendida está dada por:

$$\min\{D, y\} = \begin{cases} D & \text{si } D < y \\ y & \text{si } D \geq y \end{cases} \quad (2.18)$$

Si la demanda es D y se tiene almacenado y , el costo en el que se incurre está dado por

$$C(D, y) = cy + p\max\{0, D - y\} + h\max\{0, y - D\} \quad (2.19)$$

Como la demanda es una variable aleatoria (con distribución de probabilidad $P_D(d)$), este costo también es una variable aleatoria. El costo esperado está dado por $C(y)$, en donde

$$\begin{aligned}
C(y) &= E[C(D, y)] = \sum_{d=0}^{\infty} [cy + p \max\{0, d - y\} + h \max\{0, y - d\}] P_D(d) \\
&= cy + \sum_{d=y}^{\infty} p(d - y) P_D(d) + \sum_{d=0}^{y-1} h(y - d) P_D(d). \tag{2.20}
\end{aligned}$$

La función $C(y)$ depende de la distribución de probabilidad de D . Con frecuencia se dificulta encontrar una representación de esta distribución de probabilidad cuando la demanda tiene un gran número de valores posibles. Así, muchas veces esta variable aleatoria discreta se aproxima por una variable aleatoria continua. Cuando la demanda tiene un gran número de valores posibles, esta aproximación casi siempre llevará a un valor muy cercano a la cantidad óptima del inventario.

Para la variable aleatoria D , sea

$$\begin{aligned}
\varphi_D(\xi) &= \text{función de densidad de probabilidad de } D \\
\Phi(a) &= \text{función de distribución acumulada de } D
\end{aligned}$$

de manera que

$$\Phi(a) = \int_0^a \varphi_D(\xi) d\xi. \tag{2.21}$$

Haciendo la aproximación, el costo esperado $C(y)$ se expresa como

$$\begin{aligned}
C(y) &= E[C(D, y)] = \int_0^{\infty} C(\xi, y) \varphi_D(\xi) d\xi \\
&= \int_0^{\infty} [cy + p \max\{0, \xi - y\} + h \max\{0, y - \xi\}] \varphi_D(\xi) d\xi \\
&= cy + \int_y^{\infty} p(\xi - y) \varphi_D(\xi) d\xi + \int_0^y h(y - \xi) \varphi_D(\xi) d\xi \\
&= cy + L(y) \tag{2.22}
\end{aligned}$$

en donde $L(y)$ se conoce como costo esperado del faltante y almacenaje.

Ahora es necesario obtener el valor de y^0 que minimiza $C(y)$. La cantidad óptima a ordenar, y^0 , es el valor que satisface

$$\Phi(y^0) = \frac{p - c}{p + h}. \tag{2.23}$$

En la sección 2.3.3 se mostrará la forma de deducir este resultado.

2.3.2 MODELO CON UN INVENTARIO INICIAL

Suponga que el nivel de inventario inicial está dado por x y que la decisión que debe tomarse es el valor de y , el nivel de inventario después del reabastecimiento de la orden (o la producción) de unidades adicionales (veáse [7]). Así, deberá ordenarse $y - x$ unidades, de manera que:

$$\text{cantidad disponible}(y) = \text{cantidad inicial}(x) + \text{cantidad ordenada}(y-x).$$

La ecuación de costo que se presentó antes permanece igual excepto que el término cy se convierte ahora en $c(y-x)$, de modo que el costo esperado mínimo está dado por

$$\min_{y \geq x} \left\{ c(y-x) + \int_y^\infty p(\xi-y)\varphi_D(\xi)d\xi + \int_0^y h(y-\xi)\varphi_D(\xi)d\xi \right\} \quad (2.24)$$

Debe agregarse la restricción $y \geq x$, puesto que el nivel de inventario después del reabastecimiento no puede ser menor que el nivel inicial x . La política óptima de inventarios se describe como sigue:

$$\text{Si } x \begin{cases} < y^0 & \text{se ordena } y^0 - x \text{ para subir el nivel de inventario a } y^0 \\ \geq y^0 & \text{no se ordena} \end{cases} \quad (2.25)$$

en donde y^0 satisface

$$\Phi(y^0) = \frac{p-c}{p+h}. \quad (2.26)$$

2.3.3 DERIVACIÓN DE LA POLÍTICA ÓPTIMA

Se supondrá que el nivel de inventario inicial es cero. Para cualesquiera constantes positivas c_1 y c_2 se define $g(\xi, y)$ como

$$g(\xi, y) = \begin{cases} c_1(y-\xi) & \text{si } y > \xi \\ c_2(\xi-y) & \text{si } y \leq \xi \end{cases} \quad (2.27)$$

y sea

$$G(y) = \int_0^\infty g(\xi, y)\varphi_D(\xi)d\xi + cy \quad (2.28)$$

en donde $c > 0$. Entonces $G(y)$ se minimiza en $y = y^0$, donde y^0 es la solución de

$$\Phi(y^0) = \frac{c_2-c}{c_2+c_1}. \quad (2.29)$$

Para ver por qué este valor de y^0 minimiza $G(y)$, observe que, por definición

$$G(y) = c_1 \int_0^y (y - \xi) \varphi_D(\xi) d\xi + c_2 \int_y^\infty (\xi - y) \varphi_D(\xi) d\xi + cy. \quad (2.30)$$

Tomando la derivada e igualando a cero se obtiene

$$\frac{dG(y)}{dy} = c_1 \int_0^y \varphi_D(\xi) d\xi - c_2 \int_y^\infty \varphi_D(\xi) d\xi + c = 0 \quad (2.31)$$

Esto implica que $c_1 \Phi(y^0) - c_2 [1 - \Phi(y^0)] + c = 0$ puesto que $\int_0^\infty \varphi_D(\xi) d\xi = 1$. Al resolver esta expresión queda

$$\Phi(y^0) = \frac{c_2 - c}{c_2 + c_1}. \quad (2.32)$$

La solución de esta ecuación minimiza $G(y)$ ya que $\frac{d^2 G(y)}{dy^2} = (c_1 + c_2) \varphi_D(y) \geq 0$ para toda y .

Para aplicar este resultado, es suficiente demostrar que

$$C(y) = cy + \int_y^\infty p(\xi - y) \varphi_D(\xi) d\xi + \int_0^y h(y - \xi) \varphi_D(\xi) d\xi$$

tiene la forma de $G(y)$. Es evidente que $c_1 = h$, $c_2 = p$ y $c = c$, por lo que la cantidad óptima a ordenar y^0 es el valor que satisface

$$\Phi(y^0) = \frac{p - c}{p + h}. \quad (2.33)$$

Para obtener los resultados para el caso en que el inventario inicial sea $x > 0$, es necesario resolver la relación

$$\min_{y \geq x} \{-cx + [\int_y^\infty p(\xi - y) \varphi_D(\xi) d\xi + \int_0^y h(y - \xi) \varphi_D(\xi) d\xi + cy]\}. \quad (2.34)$$

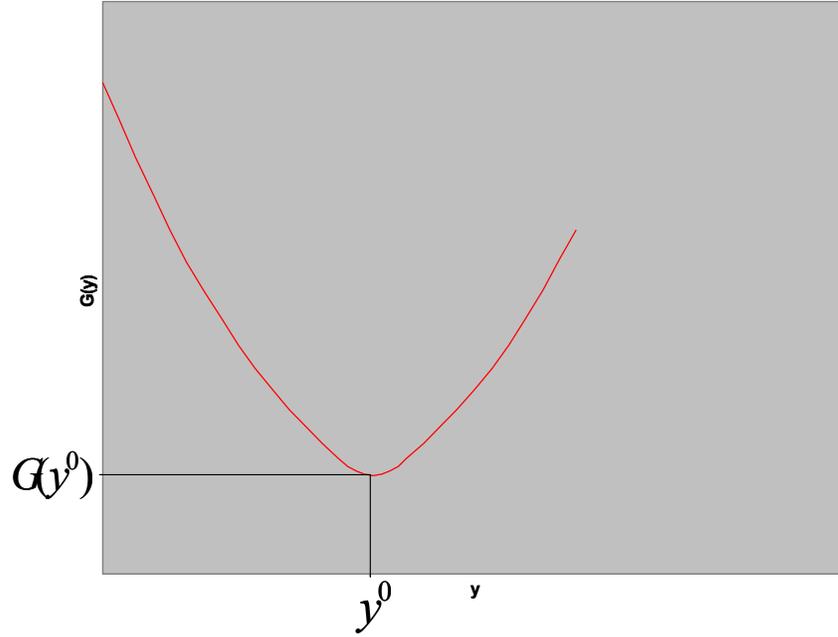
La expresión entre corchetes tiene la forma $G(y)$, $c_1 = h$, $c_2 = p$ y $c = c$. Entonces la función de costo que se tiene que minimizar se puede escribir como

$$\min_{y \geq x} \{-cx + G(y)\}. \quad (2.35)$$

Es claro que $-cx$ es una constante, de manera que es suficiente encontrar el valor de y que satisface la expresión

$$\min_{y \geq x} \{G(y)\}. \quad (2.36)$$

Así, el valor de y^0 que minimiza $G(y)$ satisface $\Phi(y^0) = \frac{p-c}{p+h}$. Más aún, $G(y)$ debe ser una función convexa porque $\frac{d^2 G(y)}{dy^2} \geq 0$. También se observa que

Figure 2.1: Gráfica de $G(y)$

$$\begin{aligned} \lim_{y \rightarrow 0} \frac{dG(y)}{dy} &= c - p \text{ el cual es negativo} \\ \lim_{y \rightarrow \infty} \frac{dG(y)}{dy} &= h + c \text{ el cual es positivo} \end{aligned} \quad (2.37)$$

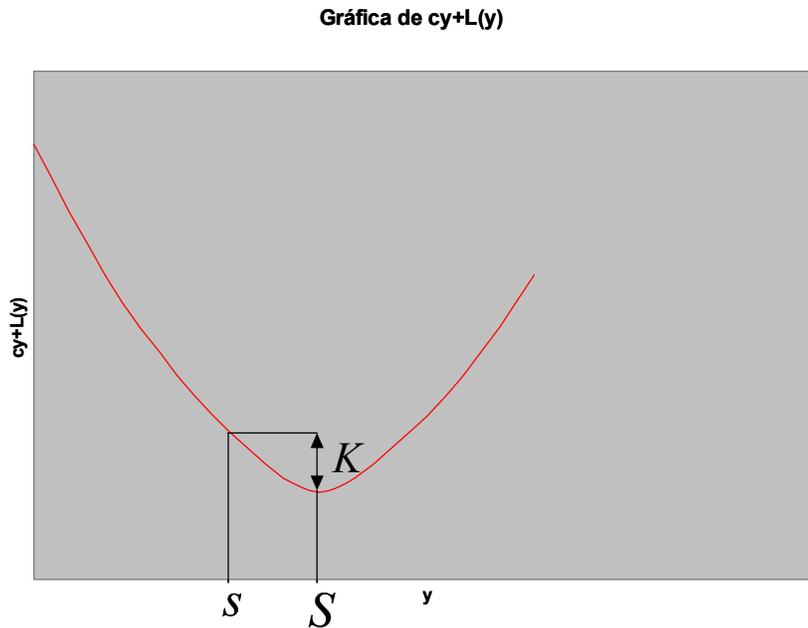
Entonces $G(y)$ debe tener la forma que se muestra en la Figura 2.1 y la política óptima está dada como sigue:

Si $x < y^0$, ordenar $y^0 - x$, para subir el nivel de inventario a y^0 , ya que y^0 se puede alcanzar junto con el valor mínimo de $G(y^0)$.

Si $x \geq y^0$, no ordenar, ya que con $y > x$, cualquier $G(y)$ será mayor que $G(x)$.

2.3.4 MODELO DE INVENTARIOS DE UN PERIODO CON COSTO DE PREPARACIÓN

Se denotará por K el costo de preparación. Cada uno de los costos de almacenaje y por faltantes se supondrá lineal. Su efecto resultante estará dado por

Figure 2.2: Gráfica de $cy+L(y)$

$$L(y) = p \int_y^{\infty} (\xi - y) \varphi_D(\xi) d\xi + h \int_0^y (y - \xi) \varphi_D(\xi) d\xi. \quad (2.38)$$

Así, el costo total esperado en el que se incurre al elevar el nivel de inventario a y unidades está dado por

$$\begin{cases} K + c(y - x) & \text{si } y > x \\ L(x) & \text{si } y = x. \end{cases} \quad (2.39)$$

Observe que $cy + L(y)$ es el mismo costo esperado que se consideró antes cuando se omitió el costo de preparación. Si $cy + L(y)$ se bosqueja como una función de y , tendrá la forma que se muestra en la Figura 2.2.

Defina S como el valor de y que minimiza $cy + L(y)$ y defina s como el valor más pequeño de y para el que $cs + L(s) = K + cS + L(S)$.

Según la figura anterior, se puede ver que:

si $x > S$, entonces

$$K + cy + L(y) > cx + L(x) \quad \text{para toda } y > x$$

de manera que

$$K + c(y - x) + L(y) > L(x).$$

El lado izquierdo de la última desigualdad representa el costo total esperado de ordenar $y - x$ unidades para elevar el nivel de inventario a y unidades, y el lado derecho de esta desigualdad representa el costo total esperado si no se ordena. La política óptima indica que si $x > S$, no se ordene.

Si $s \leq x \leq S$, de nuevo se ve de la figura 2.2 que

$$K + cy + L(y) \geq cx + L(x) \text{ para toda } y > x,$$

de manera que

$$K + c(y - x) + L(y) \geq L(x).$$

Una vez más, no ordenar es menos costoso que ordenar. Por último, si $x < s$, en la Figura 2.2 se ve que

$$\min_{y \geq k} \{K + cy + L(y)\} = K + cS + L(S) < cx + L(x)$$

$$\min_{y \geq k} \{K + c(y - x) + L(y)\} = K + c(S - x) + L(S) < L(x)$$

de manera que es mejor ordenar.

Se incurre en un costo mínimo elevando el nivel de inventario a S unidades. Así, la política óptima de inventario es la siguiente:

$$Si \ x \begin{cases} < s & \text{se ordena } S - x \text{ para elevar el nivel de inventario a } S \\ \geq s & \text{no se ordena.} \end{cases} \quad (2.40)$$

Este valor de S se obtiene a partir de

$$\Phi(S) = \frac{p - c}{p + h} \quad (2.41)$$

y s es el valor más pequeño que satisface la expresión

$$cs + L(s) = K + cS + L(S). \quad (2.42)$$

2.3.5 MODELO DE INVENTARIOS DE DOS PERIODOS SIN COSTO DE PREPARACIÓN

Los supuestos del modelo son los siguientes:

1.- La planeación se hace para dos periodos, en donde la demanda insatisfecha en el periodo 1 se acarrea para satisfacerla en el periodo 2, pero no se permite acarrear faltantes del periodo 2.

2.- Las demandas D_1 y D_2 para los periodos 1 y 2 son variables aleatorias independientes e idénticamente distribuidas. Su distribución de probabilidad

común tiene la función de densidad de probabilidad $\varphi_D(\xi)$ y la función de distribución acumulada $\Phi(\xi)$.

3.- El nivel de inventario inicial (antes de reabastecer) al principio del periodo 1 es $x_1 \geq 0$.

4.- El objetivo es minimizar el costo total esperado para ambos periodos, en donde los componentes del costo para cada periodo son:

c = costo unitario al comprar o producir cada unidad.

h = costo de mantener inventario por unidad que queda al final de cada periodo.

p = costo por faltantes por unidad de demanda insatisfecha al final de cada periodo.

Para comenzar el análisis, sea

y_i^0 = valor óptimo de y_i para $i = 1, 2$.

$C_1(x_1)$ = costo total esperado para ambos periodos cuando se sigue la política óptima dado que x_1 es el nivel de inventario (antes de reabastecer) al principio del periodo 1.

$C_2(x_2)$ = costo total esperado sólo para el periodo 2 cuando se sigue la política óptima dado que x_2 es el nivel de inventario (antes de reabastecer) al principio del periodo 2.

Para usar el enfoque de programación dinámica, primero se obtiene $C_2(x_2)$ y y_2^0 , donde se tiene sólo un periodo por analizar. Después se usarán estos resultados para encontrar $C_1(x_1)$ y y_1^0 . De los resultados del modelo de un sólo periodo, y_2^0 se encuentra resolviendo la ecuación

$$\Phi(y_2^0) = \frac{p - c}{p + h}. \quad (2.43)$$

Dado x_2 , entonces la política óptima que resulta es la siguiente

$$\text{Si } x_2 \begin{cases} < y_2^0 & \text{ordenar } y_2^0 - x_2 \text{ para elevar el inventario hasta } y_2^0 \\ \geq y_2^0 & \text{no ordenar.} \end{cases} \quad (2.44)$$

El costo de esta política óptima se puede expresar como

$$C_2(x_2) = \begin{cases} L(x_2) & \text{si } x_2 \geq y_2^0 \\ c(y_2^0 - x_2) + L(y_2^0) & \text{si } x_2 < y_2^0 \end{cases} \quad (2.45)$$

en donde $L(z)$ es el costo esperado de almacenaje y faltantes para un sólo periodo cuando existen z unidades en inventario (después de reabastecer). Ahora $L(z)$ se puede expresar como

$$L(z) = \int_z^\infty p(\xi - z)\varphi_D(\xi)d\xi + \int_0^z h(z - \xi)\varphi_D(\xi)d\xi. \quad (2.46)$$

Cuando se consideran ambos periodos, los costos consisten en el costo de compra $c(y_1 - x_1)$, el costo esperado de almacenaje y faltantes $L(y_1)$ y los costos

asociados a seguir una política durante el segundo periodo (veáse [5] y [10]). Así, el costo esperado si se sigue una política óptima en los dos periodos está dado por

$$C_1(x_1) = \min_{y_1 \geq x_1} \{c(y_1 - x_1) + L(y_1) + E[C_2(x_2)]\} \quad (2.47)$$

en donde $E[C_2(x_2)]$ se obtiene como sigue.

Observe que $x_2 = y_1 - D_1$ de manera que x_2 es una variable aleatoria al principio del periodo 1. Entonces

$$C_2(x_2) = C_2(y_1 - D_1) = \begin{cases} L(y_1 - D_1) & \text{si } y_1 - D_1 \geq y_2^0 \\ C(y_2^0 - y_1 + D_1) + L(y_2^0) & \text{si } y_1 - D_1 < y_2^0. \end{cases} \quad (2.48)$$

Así, $C_2(x_2)$ es una variable aleatoria y su valor esperado está dado por

$$\begin{aligned} E[C_2(x_2)] &= \int_0^\infty C_2(y_1 - \xi) \varphi_D(\xi) d\xi \\ &= \int_0^{y_1 - y_2^0} L(y_1 - \xi) \varphi_D(\xi) d\xi \\ &\quad + \int_{y_1 - y_2^0}^\infty [c(y_2^0 - y_1 + \xi) + L(y_2^0)] \varphi_D(\xi) d\xi. \end{aligned} \quad (2.49)$$

Entonces

$$C_1(x_1) = \min_{y_1 \geq x_1} \left\{ \begin{aligned} &c(y_1 - x_1) + L(y_1) + \int_0^{y_1 - y_2^0} L(y_1 - \xi) \varphi_D(\xi) d\xi \\ &+ \int_{y_1 - y_2^0}^\infty [c(y_2^0 - y_1 + \xi) + L(y_2^0)] \varphi_D(\xi) d\xi \end{aligned} \right\} \quad (2.50)$$

Se puede demostrar que $C_1(x_1)$ tiene un mínimo único y que el valor óptimo de y_1 , denotado por y_1^0 , satisface la ecuación

$$-p + (p+h)\Phi(y_1^0) + (c-p)\Phi(y_1^0 - y_2^0) + (p+h) \int_0^{y_1^0 - y_2^0} \Phi(y_1^0 - \xi) \varphi_D(\xi) d\xi = 0 \quad (2.51)$$

Entonces, la política óptima que resulta para el periodo 1 es la siguiente

$$\text{Si } x_1 \begin{cases} < y_1^0 \\ \geq y_1^0 \end{cases} \begin{cases} \text{ordenar } y_1^0 - x_1 \text{ para elevar el nivel de inventario hasta } y_1^0 \\ \text{no ordenar.} \end{cases} \quad (2.52)$$

2.3.6 MODELO DE VARIOS PERIODOS SIN COSTO DE PREPARACIÓN

Ahora, consideremos la extensión del problema anterior de dos periodos a n periodos, donde $n > 2$, con suposiciones idénticas. La única diferencia es que usaremos un factor de descuento α con $0 < \alpha < 1$, para calcular el costo total esperado para n periodos. El problema sigue siendo encontrar números críticos $y_1^0, y_2^0, \dots, y_n^0$ que describan la política óptima de inventario. Al igual que en el modelo de dos periodos, es difícil obtener estos valores numéricos, pero se puede demostrar que la política óptima tiene la siguiente forma.

Para cada periodo i , ($i = 1, 2, \dots, n$) con x_i como nivel de inventario al iniciar este periodo (antes de reabastecer) se hace lo siguiente:

$$\text{Si } x_i \begin{cases} < y_i^0 & \text{ordenar } y_i^0 - x_i \text{ para elevar el nivel de inventario hasta } y_i^0 \\ \geq y_i^0 & \text{no ordenar en el periodo } i. \end{cases} \quad (2.53)$$

Lo que es más

$$y_n^0 \leq y_{n-1}^0 \leq \dots \leq y_2^0 \leq y_1^0 \quad (2.54)$$

Para el caso de un número infinito de periodos, todos estos números críticos y_1^0, y_2^0, \dots son iguales. Sea y^0 este valor constante. Se puede demostrar que y^0 satisface la ecuación

$$\Phi(y^0) = \frac{p - c(1 - \alpha)}{p + h}. \quad (2.55)$$

2.4 PROCESOS DE DECISIÓN

2.4.1 MODELO UTILIZADOS PARA PROCESOS DE DECISIÓN MARKOVIANOS

Muchos sistemas importantes se pueden modelar como una cadena de Markov de tiempo discreto o de tiempo continuo. Es útil describir el comportamiento de tales sistemas con el fin de evaluar su desempeño. Sin embargo, puede ser aún más útil diseñar la operación del sistema para optimizar su desempeño. Así, nos enfocaremos a la manera en que se puede diseñar la operación de una cadena de Markov de tiempo discreto para optimizar su desempeño. Por lo tanto, en lugar de aceptar en forma pasiva el diseño de la cadena de Markov y su matriz de transición fija correspondiente, ahora se actuará. Para cada estado posible de la cadena de Markov se tomará una decisión sobre cuál de las diferentes acciones alternativas debe tomarse en ese estado. La acción elegida afecta las probabilidades de transición al igual que los costos inmediatos (o beneficios) y los costos subsecuentes (o beneficios) por operar el sistema. Se requiere elegir las acciones óptimas para los respectivos estados considerando tanto los costos

inmediatos como los subsecuentes. El proceso de decisión para hacer esto se conoce como un *proceso de decisión markoviano*.

El modelo para los procesos markovianos de decisión considerados se resumen a continuación:

- 1.- Se observa el estado i de una cadena de Markov de tiempo discreto después de cada transición ($i = 0, 1, \dots, M$).
- 2.- Después de cada observación, se selecciona una decisión (acción) k de un conjunto de K decisiones posibles ($k = 1, 2, \dots, K$).
- 3.- Si se elige la decisión $d_i = k$ en el estado i , se incurre en un costo inmediato que tiene un valor esperado C_{ik} .
- 4.- La decisión $d_i = k$ en el estado i determina cuáles serán las probabilidades de transición para la siguiente transición desde el estado i . Denote estas probabilidades de transición por $p_{ij}(k)$, para $j = 0, 1, \dots, M$.
- 5.- Una especificación de las decisiones para los estados respectivos (d_0, d_1, \dots, d_M) prescribe una política para el proceso markoviano de decisión.
- 6.- El objetivo es encontrar una política óptima de acuerdo a algún criterio de costo que considere tanto los costos inmediatos como los subsecuentes que resulten de la evolución futura del proceso. Un criterio común es minimizar el costo promedio esperado por unidad de tiempo (a la larga).

La descripción de una política óptima implica dos propiedades convenientes que se supondrán. Una propiedad es que una política es *estacionaria*; es decir, siempre que el sistema se encuentre en el estado i , la regla para tomar la decisión siempre es la misma sin importar el valor del tiempo actual t . La segunda propiedad es que una política es *determinística*; esto es, siempre que el sistema se encuentre en el estado i , la regla para tomar la decisión definitivamente selecciona una decisión específica. A este tipo de política se le llama *política determinística estacionaria*.

2.4.2 MODELO DE ETAPA INFINITA

A la larga el comportamiento de un proceso markoviano se caracteriza por su independencia del estado inicial del sistema. En este caso se dice que el sistema ha llegado al *estado estable*. Por lo tanto, nos interesa principalmente evaluar políticas para las cuales las cadenas de Markov asociadas permitan la existencia de una solución de estado estable.

Nos interesa determinar la política óptima de largo alcance de un problema de decisión markoviano. Es lógico basar la evaluación de una política en la maximización (minimización) del ingreso (costo) esperado por periodo de transición.

Existen dos métodos para resolver el problema de etapa infinita (para mayor información de éstos véase [10] y [11]). El primer método recomienda la enumeración de todas las políticas estacionarias posibles del problema de decisión.

Al evaluar cada política, se puede determinar la solución óptima. Esto es básicamente equivalente a un proceso de enumeración exhaustiva y sólo se puede emplear si el número total de políticas estacionarias es razonablemente chico para realizar operaciones de cálculo prácticas.

El segundo método, que recibe el nombre de iteración de política, aligera las dificultades de cálculo que pudieran presentarse en el procedimiento de enumeración exhaustiva. El nuevo método es eficiente, en general, en el sentido de que determina la política óptima en un número pequeño de iteraciones. Ambos métodos nos deben llevar a la misma solución óptima.

Antes de iniciar el análisis de cada método, expresemos el problema como un modelo de programación dinámica de estado finito de la siguiente manera. Supongamos que el número de estados para cada etapa es m y definamos

$f_n(i)$ = ingreso esperado óptimo de las etapas $n, n + 1, \dots, N$, dado que el estado del sistema al inicio de la etapa n es i .

La ecuación recursiva hacia atrás que relaciona a f_n y f_{n+1} puede escribirse como

$$f_n(i) = \max_k \left\{ \sum_{j=1}^m p_{ij}^k [r_{ij}^k + f_{n+1}(j)] \right\}, \quad n = 1, 2, \dots, N \quad (2.56)$$

donde $f_{N+1}(j) = 0$ para toda j .

Una justificación para la ecuación es que el ingreso acumulado, $r_{ij}^k + f_{n+1}(j)$ que resulta de llegar al estado j en la etapa $n + 1$ desde el estado i en la etapa n ocurre con probabilidad p_{ij}^k . Si v_i^k representa el rendimiento esperado resultante de una transición desde el estado i dada la alternativa k , entonces v_i^k puede expresarse como

$$v_i^k = \sum_{j=1}^m p_{ij}^k r_{ij}^k. \quad (2.57)$$

Así, la ecuación recursiva de la programación dinámica puede escribirse como

$$\begin{aligned} f_N(i) &= \max_k \{v_i^k\} \\ f_n(i) &= \max_k \left\{ v_i^k + \sum_{j=1}^m p_{ij}^k f_{n+1}(j) \right\}, \quad n = 1, 2, \dots, N - 1 \end{aligned} \quad (2.58)$$

Método de enumeración exhaustiva.

Supóngase que el problema de decisión tiene un total de S políticas estacionarias y también que P^s y R^s son las matrices de transición e ingreso (de un paso) asociadas con la k -ésima política $s = 1, 2, \dots, S$. Los pasos del método de enumeración exhaustiva son los siguientes:

Paso 1. Calcule ν_i^s el ingreso esperado de un paso (un periodo) de la política s dado el estado i , $i = 1, 2, \dots, m$.

Paso 2. Calcule π_i^s las probabilidades estacionarias, a la larga, de la matriz de transición P^s asociadas con la política s . Estas probabilidades, cuando existen, se determinan a partir de las ecuaciones

$$\begin{aligned}\pi^s P^s &= \pi^s \\ \pi_1^s + \pi_2^s + \dots + \pi_m^s &= 1\end{aligned}\quad (2.59)$$

donde $\pi^s = (\pi_1^s, \pi_2^s, \dots, \pi_m^s)$

Paso 3. Determine E^s , el ingreso esperado de la política s por paso (periodo) de transición, mediante el uso de la fórmula

$$E^s = \sum_{i=1}^m \pi_i^s \nu_i^s \quad (2.60)$$

Paso 4. La política óptima s^* se determina de tal forma que

$$E^{s^*} = \max\{E^s\} \quad (2.61)$$

Método de iteración de política sin descuento.

El método de enumeración exhaustiva no es práctico para problemas grandes. El método de iteración de política, que está basado en la ecuación recursiva de PD es, según se demuestra, más eficiente en términos de cálculo que el método de enumeración exhaustiva, ya que normalmente converge en un número de iteraciones pequeño.

El método de iteración de política está basado principalmente en el desarrollo siguiente. Para cualquier política específica, el rendimiento total esperado en la etapa n se expresa a través de la ecuación recursiva

$$f_n(i) = v_i + \sum_{j=1}^m p_{ij} f_{n+1}(j), \quad i = 1, 2, \dots, m. \quad (2.62)$$

Esta ecuación recursiva es la base para el desarrollo del método de iteración de política. Sin embargo, la forma presente se debe modificar ligeramente de manera que nos permita estudiar la conducta asintótica del proceso. En esencia,

definimos η como el número de etapas que faltan por considerar. Esto sucede en contraste con n en la ecuación, que define la n -ésima etapa. Por lo tanto, la ecuación recursiva se escribe como

$$f_\eta(i) = v_i + \sum_{j=1}^m p_{ij} f_{\eta-1}(j), \quad i = 1, 2, \dots, m. \quad (2.63)$$

Obsérvese que f_η es el ingreso esperado acumulado dado que η es el número de etapas que faltan por considerar. Con la nueva definición, el comportamiento asintótico del proceso se puede estudiar haciendo que $\eta \rightarrow \infty$.

Dado que $\pi = (\pi_1, \pi_2, \dots, \pi_m)$ es el vector de probabilidad de estado estable de la matriz de transición $P = (p_{ij})$ y $E = \sum_{j=1}^m \pi_j v_j$ es el ingreso esperado por etapa, se puede probar que para η muy grande, $f_\eta(i) = \eta E + f(i)$ donde $f(i)$ es un término constante que representa la intersección asintótica de $f_\eta(i)$ dado el estado i .

Como $f_\eta(i)$ es el rendimiento óptimo acumulado de η etapas dado el estado i y E es el ingreso esperado por etapa, podemos advertir en forma intuitiva por que $f_\eta(i)$ es igual a ηE , más un factor de corrección $f(i)$ que contribuye al estado específico i . Este resultado, desde luego, supone que η es muy grande.

Utilizando esta información, la ecuación recursiva se escribe como

$$\eta E + f(i) = v_i + \sum_{j=1}^m p_{ij} \{(\eta - 1)E + f(j)\}, \quad i = 1, 2, \dots, m. \quad (2.64)$$

Al simplificar esta ecuación, se obtiene

$$E = v_i + \sum_{j=1}^m p_{ij} f(j) - f(i), \quad i = 1, 2, \dots, m \quad (2.65)$$

que genera m ecuaciones y $m+1$ incógnitas, donde las incógnitas son $f(1)$, $f(2)$, ..., $f(m)$ y E .

Nuestro objetivo final es el de determinar la política óptima que genere el valor máximo de E . Como hay m ecuaciones y $m+1$ incógnitas, el valor óptimo de E no se puede determinar en un paso. En cambio se utiliza un enfoque iterativo que, al comenzar con una política arbitraria, determinará entonces una nueva política que genere un mejor valor de E . El proceso iterativo termina cuando dos políticas sucesivas son idénticas.

El proceso iterativo consta de dos componentes básicas, llamadas paso de determinación del valor y paso de mejora de la política.

1.- *Determinación del valor.* Elíjase una política arbitraria s . Mediante el uso de sus matrices asociadas P^s y R^s y suponiendo arbitrariamente que $f^s(m) = 0$, resuélvanse las ecuaciones

$$E^s = v_i^s + \sum_{j=1}^m p_{ij}^s f^s(j) - f^s(i), \quad i = 1, 2, \dots, m \quad (2.66)$$

con las incógnitas $E^s, f^s(1), f^s(2), \dots, f^s(m-1)$. Diríjase al paso de mejora de la política.

2.- *Mejora de la política.* Para cada estado i , determínese la opción k que genere

$$\max_k \left\{ v_i^k + \sum_{j=1}^m p_{ij}^k f^s(j) \right\}, \quad i = 1, 2, \dots, m. \quad (2.67)$$

Las decisiones óptimas resultantes k para los estados $1, 2, \dots, m$ constituyen la nueva política t . Si s y t son idénticos, deténgase; t es óptimo. En caso contrario, hágase $s = t$ y regrésese al paso de determinación del valor.

El problema de optimización del paso de mejora de la política necesita ser aclarado. Nuestro objetivo en este paso es obtener $\max\{E\}$. Según se indica,

$$E = v_i + \sum_{j=1}^m p_{ij} f(j) - f(i). \quad (2.68)$$

Ya que $f(i)$ no depende de las opciones k , se deduce que la maximización de E sobre las opciones k es equivalente al problema de maximización dado en el paso de mejora de la política.

Método de aproximaciones sucesivas.

Ahora dirigiremos nuestra atención a un enfoque llamado *método de aproximaciones sucesivas* para encontrar rápidamente al menos una aproximación a una política óptima. El análisis se hará minimizando el costo esperado.

Se ha supuesto que el proceso de decisión markoviano operará indefinidamente y se ha buscado una política óptima para tal proceso. La idea básica del método de aproximaciones sucesivas es encontrar una política óptima para las decisiones que se toman en el primer periodo cuando sólo quedan n periodos de operación para el proceso antes de terminar, comenzando con $n = 1$, después $n = 2$, después $n = 3$, etc. Conforme n crece, las políticas óptimas correspondientes convergen a una política óptima para el problema de periodo infinito de interés. Entonces las políticas obtenidas para $n = 1, 2, 3, \dots$ proporcionan aproximaciones sucesivas que llevan a la política óptima deseada.

La razón por la que este enfoque es atractivo es que se cuenta con un método rápido para encontrar una política óptima cuando sólo quedan n periodos de operación, a saber, el de programación dinámica probabilística.

En particular, para $i = 0, 1, \dots, M$, sea

V_i^n = costo descontado total esperado por seguir una política óptima, dado que el proceso comienza en el estado i y le quedan sólo n periodos de operación

Por el principio de optimalidad para programación dinámica, las V_i^n se obtienen de la relación recursiva,

$$V_i^n = \min_k \left\{ C_{ij} + \alpha \sum_{j=0}^M p_{ij}(k) V_j^{n-1} \right\}, \quad \text{para } i = 0, 1, \dots, M. \quad (2.69)$$

El valor de k que minimiza proporciona la decisión óptima que se debe tomar en el primer periodo cuando el proceso inicia en el estado i . Para comenzar, con $n = 1$, todas las $V_i^0 = 0$, de manera que

$$V_i^1 = \min_k \{C_{ik}\}, \quad \text{para } i = 0, 1, \dots, M \quad (2.70)$$

Aunque es posible que el método de aproximaciones sucesivas no conduzca a una política óptima para el problema de infinitos periodos después de unas cuantas iteraciones, tiene una ventaja especial sobre las técnicas de mejoramiento de una política y de programación lineal: nunca requiere que se resuelva un sistema de ecuaciones simultáneas, con lo que cada iteración se puede realizar en forma sencilla y rápida. Lo que es más, si en realidad sólo quedan n periodos en el proceso de decisión markovianos, definitivamente n iteraciones de este método llevarán a una política óptima.

CAPÍTULO 3

APLICACIÓN

3.1 INTRODUCCIÓN

El objetivo de este capítulo es mostrar una aplicación de la programación dinámica a un proceso markoviano en teoría de inventarios, que es el objetivo primordial de la tesis. En los capítulos anteriores se desarrolló, en detalle, la teoría necesaria para resolver el problema de aplicación y se pretendió que no faltara ni excediera información para comprenderlo completamente.

El problema que se plantea aquí consiste en encontrar una política óptima de inventario para una tienda de ventiladores. Como en este problema los parámetros que se manejan son los costos (que se mencionaron en el capítulo 2), nuestro objetivo será encontrar la política de inventario que produce el costo mínimo.

3.2 SOLUCIÓN DE UN PROBLEMA DE INVENTARIOS

Una compañía fabricante de ventiladores en la ciudad de Puebla, llamada Siclima, desea controlar su inventario de uno de sus productos: un ventilador que es utilizado por lo regular en hoteles para recámaras de tamaño normal. Analizando la demanda, se ha observado que ésta es variable, teniéndose una mayor demanda en los meses de Abril, Mayo, Junio y Julio. En lo que sigue, deduciremos una política de inventario para el periodo Abril-Julio basándonos en la demanda observada en el año anterior (2002). El análisis se hará por semanas. En total tenemos 17 semanas. Se incurre en un costo fijo de \$400 cada vez que produce un pedido de ventiladores. El costo de almacenamiento por ventilador es de \$120. La penalización por quedar sin existencia se estima es de \$500. Además, después de analizar la demanda del año pasado la compañía decide no tener más de 40 ventiladores en inventario.

SEMANA, i	DEMANDA, x_i
1	23
2	21
3	12
4	12
5	31
6	28
7	15
8	30
9	41
10	17
11	21
12	24
13	31
14	12
15	18
16	23
17	24

Sea D_t la demanda de ventiladores en la semana t . Esta es una variable aleatoria que se ha decidido tomar con una distribución de Poisson. Calculando el parámetro λ tenemos

$$\lambda = \frac{\sum_{i=1}^{17} x_i}{17} = 22.5. \quad (3.1)$$

Así, la distribución de probabilidades para la demanda de ventiladores es

$$P(D_t = k) = \exp(-22.5) \frac{(22.5)^k}{k!}. \quad (3.2)$$

Para facilitar los cálculos y el análisis, la demanda se tomó en decenas y así

$$\lambda = \frac{\sum_{i=1}^{17} x_i}{17} = 2.25. \quad (3.3)$$

Entonces tenemos la distribución de probabilidad de la siguiente manera.

$$P(D_t = k) = \exp(-2.25) \frac{(2.25)^k}{k!}. \quad (3.4)$$

Como el fabricante no permite tener más de 40 ventiladores en inventario, tomaremos como el conjunto de estados posibles números enteros, es decir, $\{0, 1, 2, 3, 4\}$. Utilizando la ecuación (3.4) tenemos

Demanda	$P(D_t = k)$
0	0.1054
1	0.2371
2	0.2668
3	0.2001
4	0.1906

Sea X_t nivel de inventario al final de la semana $t = 0, 1, 2, \dots$. Así,

$$X_{t+1} = \begin{cases} \max\{(4 - D_{t+1}), 0\} & \text{si } X_t < y^0 \\ \max\{(X_t - D_{t+1}), 0\} & \text{si } X_t \geq y^0. \end{cases} \quad (3.5)$$

A continuación obtendremos las matrices de transición y de costos para cada una de las políticas.

Si $\begin{cases} X_t < 1 & \text{ordenar} \\ X_t \geq 1 & \text{no ordenar} \end{cases}$, tenemos que la matriz de transición es

$$P^1 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \\ \left(\begin{array}{ccccc} 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.8946 & 0.1054 & 0 & 0 & 0 \\ 0.6575 & 0.2371 & 0.1054 & 0 & 0 \\ 0.3907 & 0.2668 & 0.2371 & 0.1054 & 0 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \end{array} \right) \end{array}$$

y la matriz de costo es

$$C^1 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{ccccc} \overbrace{0 \quad 1 \quad 2 \quad 3 \quad 4} \\ \left(\begin{array}{ccccc} 495.3 & 424.012 & 432.016 & 428.452 & 412.648 \\ 447.3 & 12.648 & 0 & 0 & 0 \\ 328.75 & 28.452 & 12.648 & 0 & 0 \\ 195.35 & 32.016 & 28.452 & 12.648 & 0 \\ 95.3 & 24.012 & 32.016 & 28.452 & 12648 \end{array} \right) \end{array}$$

Si $\begin{cases} X_t \leq 1 & \text{ordenar} \\ X_t > 1 & \text{no ordenar} \end{cases}$, tenemos que la matriz de transición es

$$P^2 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{array}{ccccc} 0 & 1 & 2 & 3 & 4 \\ \left(\begin{array}{ccccc} 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.6575 & 0.2371 & 0.1054 & 0 & 0 \\ 0.3907 & 0.2668 & 0.2371 & 0.1054 & 0 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \end{array} \right) \end{array}$$

y su matriz de costo es

$$C^2 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 495.3 & 424.012 & 432.016 & 428.452 & 412.648 \\ 495.3 & 424.012 & 432.016 & 428.452 & 412.648 \\ 328.75 & 28.452 & 12.648 & 0 & 0 \\ 195.35 & 32.016 & 28.452 & 12.648 & 0 \\ 95.3 & 24.012 & 32.016 & 28.452 & 12.648 \end{pmatrix}}$$

Si $\begin{cases} X_t \leq 2 & \text{ordenar} \\ X_t > 1 & \text{no ordenar} \end{cases}$ tenemos que la matriz de transición es

$$P^3 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \\ 0.3907 & 0.2668 & 0.2371 & 0.1054 & 0 \\ 0.1906 & 0.2001 & 0.2668 & 0.2371 & 0.1054 \end{pmatrix}$$

y su matriz de costo es

$$C^3 = \begin{array}{c} \text{Estado} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{array} \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 \\ 495.3 & 424.012 & 432.016 & 428.452 & 412.648 \\ 495.3 & 424.012 & 432.016 & 428.452 & 412.648 \\ 495.3 & 424.012 & 432.016 & 428.452 & 412.64 \\ 195.35 & 32.016 & 28.452 & 12.648 & 0 \\ 95.3 & 24.012 & 32.016 & 28.452 & 12.648 \end{pmatrix}}$$

Haciendo los cálculos de v_i^s tenemos que

	v_i^s				
s	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
1	439.59	401.4877	224.2322	92.9442	39.59
2	439.59	439.59	224.2322	92.9442	39.59
3	439.59	439.59	439.59	92.9442	39.59

Continuando con los cálculos, seguiremos el algoritmo propuesto en la sección 2.4.2

*ITERACIÓN 1*1) *Determinación del valor.*

Elegimos la política $s = 1$. Utilizando la fórmula 2.66 tenemos que.

$$E = 439.59 + 0.1906f(1) + 0.2001f(2) + 0.2668f(3) + 0.2371f(4) + 0.1054f(5) - f(1)$$

$$E = 401.4877 + 0.8946f(1) + 0.1054f(2) - f(2)$$

$$E = 224.2322 + 0.6575f(1) + 0.2371f(2) + 0.1054f(3) - f(3)$$

$$E = 92.9442 + 0.3907f(1) + 0.2668f(2) + 0.2371f(3) + 0.1054f(4) - f(4)$$

$$E = 39.59 + 0.1906f(1) + 0.2001f(2) + 0.2668f(3) + 0.2371f(4) + 0.1054f(5) - f(5)$$

Al resolver este sistema de ecuaciones, considerando $f(5) = 0$, obtenemos $E = 327.17$, $f(1) = 400$, $f(2) = 483.07$, $f(3) = 306.95$, $f(4) = 138.29$

2) *Mejora de la política.*

Utilizando la fórmula 2.67 obtenemos los valores de $f(i)$

i	$k = 1$	$k = 2$	$k = 3$	$f(i)$	k^*
1	727.18	727.18	727.18	727.18	1, 2 o 3
2	810.24	727.18	727.18	727.18	2 o 3
3	634.12	634.12	727.18	634.12	1 o 2
4	465.46	465.46	465.46	465.46	1, 2 o 3
5	327.18	327.18	327.18	327.18	1, 2 o 3

Como la política óptima común en cualquier estado es $k^* = 2$, tomaremos su matriz de transición y matriz de costo para continuar con el algoritmo

*ITERACIÓN 2*1) *Determinación del valor.*

Ahora tenemos el sistema de ecuaciones:

$$E = 439.59 + 0.1906f(1) + 0.2001f(2) + 0.2668f(3) + 0.2371f(4) + 0.1054f(5) - f(1)$$

$$E = 439.59 + 0.1906f(1) + 0.2001f(2) + 0.2668f(3) + 0.2371f(4) + 0.1054f(5) - f(2)$$

$$E = 224.2322 + 0.6575f(1) + 0.2371f(2) + 0.1054f(3) - f(3)$$

$$E = 92.9442 + 0.3907f(1) + 0.2668f(2) + 0.2371f(3) + 0.1054f(4) - f(4)$$

$$E = 39.59 + 0.1906f(1) + 0.2001f(2) + 0.2668f(3) + 0.2371f(4) + 0.1054f(5) - f(5)$$

Al resolverlo, obtenemos la solución $f(1) = 400.0$, $f(2) = 400.0$, $f(3) = 305.129$, $f(4) = 133.143$, $f(5) = 0$ y $E = 308.196$.

2) *Mejora de la política.*

Los cálculos del paso de mejora de la política se da en la siguiente tabla.

i	$k = 1$	$k = 2$	$k = 3$	$f(i)$	k^*
1	708.96	708.96	708.96	708.96	1, 2 o 3
2	801.49	708.96	708.96	708.96	2 o 3
3	614.25	614.25	708.96	614.25	1 o 2
4	442.39	442.39	442.39	442.39	1, 2 o 3
5	308.96	308.96	308.96	308.96	1, 2 o 3

Como podemos observar, la nueva política es idéntica a la política obtenida anteriormente, por lo tanto hasta aquí termina el proceso iterativo. Como la política común es la política $k^* = 2$, tenemos que esta es la política óptima.

Notemos que este método llegó muy rápido a la solución óptima. Si se hubiese resuelto el problema con el método de enumeración exhaustiva se hubiesen hecho muchos cálculos para llegar a esta solución. Esta es una de las ventajas de aplicar el enfoque de la programación dinámica a la solución de problemas donde se cumple el principio de optimalidad.

3.3 CONCLUSIONES

Los procesos de decisión markovianos son bastantes importantes, ya que aparecen con bastante frecuencia en la práctica. Debido a que éstos utilizan muchas herramientas matemáticas para su solución, fue uno de los motivos que nos impulsó a tomarlos como objetos de estudio para mostrar el alcance de la matemática aplicada.

En este trabajo de tesis damos una introducción sobre la teoría de decisiones, un breve resumen de la herramienta matemática que se necesita para la solución de problemas de decisión y resolvimos un problema de éstos en el área de inventarios. Si bien la teoría expuesta aquí se puede encontrar en los libros que se mencionan en la bibliografía, el mérito en esta tesis (si es que existe) es la concentración y el tratar de explicar en forma clara y concisa la interrelación entre teoría de decisiones, cadenas de Markov y programación dinámica.

En nuestro país estos métodos de solución no son utilizados actualmente pero se espera que en un futuro si lo sean, ya que en países desarrollados han probado ser efectivos en problemas de minimización de costos y maximización de utilidades.

La importancia de la herramienta matemática se manifiesta en la aplicación de las cadenas de Markov en el modelamiento del problema y de la programación dinámica en el proceso de solución para encontrar la mejor alternativa. El problema resuelto aquí, aunque pequeño, es original y muestra lo que se ha mencionado. Es así como hemos logrado el objetivo inicial.

BIBLIOGRAFÍA

- [1] Bertsekas, D. P.; Dynamic Programming Deterministic and Stochastic Models; Prentice-Hall; Englewood cliffs; NJ; 1987.
- [2] Canavos C., George; Probabilidad y Estadística. Aplicaciones y Métodos; Mc Graw Hill; 1988.
- [3] Denardo, E. V.; Dynamic Programing. Theory and Applications; Prentice-Hall; Englewood cliffs; NJ; 1982.
- [4] Feller, William; Introducción a la Teoría de Probabilidades y sus Aplicaciones; 3a. ed. Limusa; 1993.
- [5] Hillier, Frederick S. y G. J. Lieberman; Investigación de Operaciones; Mc Graw Hill; 2001.
- [6] Isaacson L., Dean y Madsen W., Richard; Markov Chains. Theory and applications; Springer-Verlag; 1975.
- [7] Liu, B. y A. O. Esogbue; Decision Criteria and Optimal Inventory Processes; Kluwer Academic Publisher; Boston; 1999
- [8] Meyer L., Paul; Probabilidad y Aplicaciones Estadísticas; Fondo Educativo Interamericano; 1973.
- [9] Spiegel R., Murray; Probabilidad y Estadística; Mc Graw Hill; 1995.
- [10] Taha, A. Hamdy; Investigación de Operaciones; Alfaomega; 1995.
- [11] White, D. J.; Real Applications of Markov Decision Processes; interfaces; 15(6): 73-83; November-December; 1985.