

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
DIVISIÓN DE ESTUDIOS DE POSGRADO



HEURÍSTICO BASADO EN DENSIDADES DE PROBABILIDAD
PARA OPTIMIZACIÓN ESTOCÁSTICA

TESIS PRESENTADA POR:
AYDEÉ LÓPEZ MUÑIZ

CON OPCIÓN AL GRADO
MAESTRÍA EN CIENCIAS EN INGENIERÍA DE SISTEMAS

SAN NICOLÁS DE LOS GARZA, NUEVO LEÓN

JULIO DEL 2009

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
DIVISIÓN DE ESTUDIOS DE POSGRADO



HEURÍSTICO BASADO EN DENSIDADES DE PROBABILIDAD
PARA OPTIMIZACIÓN ESTOCÁSTICA

TESIS PRESENTADA POR:

AYDEÉ LÓPEZ MUÑIZ

CON OPCIÓN AL GRADO

MAESTRÍA EN CIENCIAS EN INGENIERÍA DE SISTEMAS

SAN NICOLÁS DE LOS GARZA, NUEVO LEÓN

JULIO DEL 2009

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
DIVISIÓN DE ESTUDIOS DE POSGRADO

Los miembros del Comité de Tesis recomendamos que la Tesis “Heurístico basado en densidades de probabilidad para optimización estocástica”, realizada por la alumna Aydeé López Muñiz con matrícula 971742, sea aceptada para su defensa como opción al grado de Maestro en Ciencias en Ingeniería de Sistemas.

Dr. José Arturo Berrones Santos

Asesor

Dra. Deniz Özdemir

Revisor

Dr. César Emilio Villarreal Rodríguez

Revisor

Dr. Moisés Hinojosa Rivera

División de Estudios de Posgrado

DEDICATORIA

*Con todo mi amor, por todo el esfuerzo y trabajo,
dedico este proyecto al tesoro más
precioso que tengo en la vida:
mis padres y mi hermano.*

AGRADECIMIENTOS

Gracias a mi familia que adoro y siempre me apoya, y a todos mis amigos.

Gracias a CONACYT por brindarme apoyo a través de esta beca para realizar mis estudios de posgrado. Además, al Posgrado en Ingeniería de Sistemas, a mi asesor por los conocimientos, apoyo y consejos que me brindó, y a los doctores que integran este posgrado pues me abrieron la puerta para ser parte de este bonito grupo de investigación.

Agradezco a Dios por ayudarme a terminar este proyecto en mi vida.

ÍNDICE GENERAL

Dedicatoria

Agradecimientos

Resumen

1. Antecedentes	7
1.1 Complejidad computacional	7
1.2 Optimización lineal	9
1.3 Optimización estocástica	10
1.3.1 Espacio de probabilidades y variables aleatoria	10
1.3.2 Los problemas de optimización con incertidumbre	11
1.3.3 Complejidad de un problema estocástico de dos etapas	15
1.3.4 Complejidad del problema determinista equivalente	15
1.3.5 La complejidad de los problemas de dos etapas con parámetros de aleatoriedad continuamente distribuidos	16
1.3.6 La dificultad del problema estocástico con múltiples variables	17
1.3.7 Método Monte Carlo en el problema de optimización estocástica	19
1.4 Optimización heurística	21
1.4.1 Optimización heurística mediante búsqueda estocástica	24
1.4.2 Adaptación y diversificación	25
2. Objetivo	27
3. Motivación	28
4. Justificación	30

5. Metodología	34
5.1 Principios del método de muestreo de Fokker-Planck y recocido simulado	34
5.2 Principios del algoritmo de muestreo de Fokker-Planck	36
5.3 Algoritmo Fokker-Planck: densidad de probabilidad estacionaria de procesos de búsqueda estocásticos en la optimización global	40
6. Resultados	46
6.1 Problema del vendedor de periódicos	46
6.2 Resultados del algoritmo SFP en el problema del vendedor de periódicos	49
6.2.1 Implementación del método SFP en el problema	49
6.2.2 Problema con demanda unimodal	50
6.2.3 Resultados del algoritmo SFP en el problema del vendedor de periódicos con demanda bimodal	56
6.2.4 Problema del vendedor de periódicos con demanda en dos intervalos	63
7. Aplicaciones	72
7.1 Aplicación del método SFP con demanda de producto 0397	73
7.2 Aplicación del método SFP con demanda de producto 0419	76
7.3 Aplicación del método SFP con demanda de producto 0429	78
7.4 Aplicación del método SFP con demanda de producto 0432	79
7.5 Aplicación del método SFP con demanda de producto 0438	80
7.6 Resultados con algoritmo Monte Carlo	81
8. Conclusiones	87
9. Contribución	88
Apéndice A	89
Apéndice B	90

Bibliografía

RESUMEN

Ing. Aydeé López Muñiz

Candidata para el grado de Maestro en Ciencias en Ingeniería de Sistemas

Universidad Autónoma de Nuevo León

Facultad de Ingeniería Mecánica y Eléctrica

Título del estudio:

HEURÍSTICO BASADO EN DENSIDADES DE PROBABILIDAD PARA OPTIMIZACIÓN ESTOCÁSTICA

Número de páginas: 91

OBJETIVOS Y MÉTODO DE ESTUDIO: Proponer un método heurístico que se basa en la estimación de densidades de probabilidad, combinando adaptación y diversificación, en el proceso de búsqueda estocástica en el espacio de soluciones. En este trabajo de investigación, se introduce este método dentro de la optimización estocástica sin restricciones, donde se resuelve un problema clásico de optimización estocástica y para el cual se puede calcular solución analítica. Por lo que se pueden establecer comparaciones en las soluciones obtenidas. Es un método de muestreo que construye las densidades de probabilidad marginales estacionarias que pueden ser definidas y dibujadas, para obtener la distribución de probabilidad asintótica asociada a la estructura de la función objetivo del problema estocástico, que nos indique la zona que encierra el óptimo global con muy alta probabilidad.

En el proceso de estimación de densidades, el método realiza un número definido de operaciones lineales, lo cual otorga gran ventaja sobre otros métodos, y por lo que promete ser una herramienta eficiente en la solución de problemas estocásticos, incluso de múltiples variables.

CONTRIBUCIONES Y CONCLUSIONES: Se propone este método como una nueva herramienta alternativa en la solución eficiente de problemas de optimización estocástica sin restricciones. Se aplica este método en la solución de un problema conocido de optimización estocástica y se comparan las soluciones analíticas calculadas para el problema con los resultados obtenidos del método heurístico de muestreo. Se observan resultados muy satisfactorios.

Dados los resultados observados, se implementa el método de muestreo en el problema de optimización estocástica para casos en que las variables de incertidumbre tienen comportamientos bimodales. Además, se aplica el método en la solución del problema, con un caso práctico donde no se conoce la distribución de la variable aleatoria, sino que se tienen datos históricos de ella y que se toman como tales. En este caso, se hacen comparaciones de estos resultados con los obtenidos en otro método llamado Monte Carlo.

Los resultados obtenidos en los casos mencionados fueron muy satisfactorios pues se observaron soluciones de muy buena calidad, pues son consistentes con las propiedades de las densidades estacionarias correctas, y donde el tiempo de cómputo necesario de solución mediante el método propuesto, es muy razonable.

Firma del asesor:

Dr. José Arturo Berrones Santos

Capítulo 1

ANTECEDENTES

1.1 COMPLEJIDAD COMPUTACIONAL

En los problemas de optimización existe siempre un procedimiento elemental para determinar la solución óptima buscando: realizar una exploración exhaustiva en el espacio de soluciones. Es decir, generar todas las soluciones factibles (las que satisfacen las restricciones), calcular para cada una el costo asociado y elegir finalmente la que haya dado lugar al mejor de ellos. Sin embargo, de nuevo sucede algo similar a lo que nos ocurría con el simplex: aunque este método teóricamente nos lleva siempre a la solución óptima buscada, no es eficiente, pues el tiempo de cálculo necesario crece exponencialmente con el número de variables del problema [3].

Por ejemplo, el problema de la mochila. El número de subconjuntos del conjunto $\{1..n\}$ es 2^n . Por tanto, si una computadora pudiera, en tan sólo un segundo, generar un millón de esos subconjuntos y evaluara su valor en el objetivo, necesitaría solamente un segundo para hallar la solución de un problema con $n = 20$ variables (es $2^{20} = 1,048,580$); sin embargo serían ya necesarias unas dos semanas para un problema con $n = 40$ variables; y 365 siglos de cálculo para analizar las 2^{60} posibles soluciones de un problema con $n = 60$ [3].

Por tanto, existen problemas para los que no se conocen algoritmos de solución o algunos exactos que produce un tiempo de cálculo alto (exponencialmente) al aumentar el tamaño del problema. Son problemas computacionalmente difíciles de tratar. Por el contrario para otros problemas, sí existen algoritmos que solo crecen polinomialmente, con el tamaño del problema.

Aquellos problemas para los que se conocen algoritmos que necesitan un tiempo polinomial para obtener la solución óptima, se dice que pertenecen a la clase P y se considera que son resolubles eficientemente. Muchos problemas pertenecen a la clase NP , donde no se conoce un algoritmo polinomial para resolver el problema.

Las clases P y NP están compuestas por los correspondientes “problemas de decisión” de los problemas de optimización. Dichos problemas consisten, en determinar si para el problema de optimización existe o no una solución que mejore el costo sobre un cierto valor. Sin embargo, la mayoría de los principales problemas de optimización pertenece a otra clase, la denominada NP , en la cual están incluidos aquellos problemas para los que no se conoce un algoritmo polinomial de solución, aunque dada una solución, si sea posible comprobar en tiempo polinomial, si su costo es mejor que un determinado valor [3].

Solamente los problemas de la clase P se pueden solucionar eficientemente mediante algún algoritmo que lo haga posible. La clase $P \subseteq NP$. Si lo contrario fuera $P \supseteq NP$, entonces significara que para la mayoría de los problemas existen algoritmos eficientes de solución. Hasta la fecha no se ha demostrado que la igualdad $P=NP$ sea cierta, ni tampoco que hay problemas en NP que no están en P . Esta es una cuestión que sigue abierta [3].

Se ha demostrado que hay problemas en NP que son especialmente difíciles: los problemas de la clase NP -completos, que además son NP -difícil, es decir, todos que los problemas en NP pueden ser reducidos polinomialmente a ellos. Lo que significa que, si se puede dar una solución en un tiempo polinomial para uno de ellos (de NP -difícil), se podría dar para todos los de NP , y por lo tanto, sería $P=NP$ [3].

No se han encontrado algoritmos eficientes para resolver los problemas NP -completos. Por lo que se pensaría que, una vez demostrado que un problema es NP -completo es inútil buscar algún algoritmo resolverlo [3].

En la vida cotidiana existen tan variados problemas de decisión dentro del ámbito de la optimización, como métodos para resolverlos de acuerdo a sus características, estructura y condiciones.

1.2 OPTIMIZACIÓN LINEAL

Por lo general, el objetivo al tomar la decisión consiste en llevar a cabo en plan propuesto de manera óptima con los costos máximos o mínimos, los problemas pueden plantearse como un problema de optimización lineal, que consiste en encontrar su solución óptima mediante el modelo general: [1]

$$\begin{aligned} \text{Min } z &= c^T x && (1.1) \\ \text{s. a. } Ax &= b, \\ x &\geq 0, \end{aligned}$$

donde x es un vector de decisión de tamaño $(n \times 1)$, Los datos del problema c , A , b son conocidos y de tamaño $(n \times 1)$, $(m \times n)$ y $(m \times 1)$, respectivamente.

El valor de $z = c^T x$ corresponde a la función objetivo y el conjunto de soluciones factibles está definido como $\{x|Ax = b, x \geq 0\}$. La solución óptima x^* es aquella donde $c^T x \geq c^T x^*$, $\forall x \in \{x|Ax = b, x \geq 0\}$. Para encontrar la solución exacta en este tipo de problemas existen métodos conocidos que permiten encontrar el óptimo global en un tiempo polinomial, como el método simplex [1]. Cuando el conjunto de restricciones y la función objetivo, son lineales, el problema de optimización es lineal.

Un problema de optimización es no-lineal cuando al menos una de las funciones que se involucran en el modelo, no es lineal.

En el caso en que todas las funciones (objetivo y restricciones) son lineales, el algoritmo simplex o algunos otros métodos exactos, pueden ser una herramienta eficiente para solucionar el problema. Sin embargo, cuando ocurre que son muchísimas las variables que intervienen en el modelo (del orden de miles o más), ya no resulta eficiente la aplicación de un método exacto, pues el tiempo de cálculo en operaciones computacionales es

excesivamente largo, crece exponencialmente con el número de variables del problema. No es que el simplex o estos métodos exactos no puedan solucionarlo llegando al óptimo, sino que el tiempo de respuesta no es siquiera operable. Por lo que es necesario buscar una herramienta alternativa a los algoritmos exactos, para resolver los problemas en los que aplicar un método de solución no es eficiente. Como algún algoritmo heurístico.

1.3 OPTIMIZACIÓN ESTOCÁSTICA

En un problema de optimización, diversos aspectos pueden ser considerados inciertos y éstos pueden representarse con variables aleatorias, tales variables pueden ser las demandas futuras de productos que dependen de las condiciones del mercado incierto, o los costos de distribución y producción que dependen de los costos fluctuantes del combustible [1].

1.3.1 Espacio de probabilidades y variables aleatorias

La incertidumbre es representada en términos de experimentos con resultados denotados por ω . El conjunto de todos los resultados se representa con Ω . El conjunto de resultados depende del problema. Es importante ser capaces de definir el impacto de algunas variables aleatorias. Los resultados pueden ser combinados en subconjuntos de Ω llamados eventos. Si denotamos como ω un evento aleatorio, para cada evento $\omega \in \Omega$ está asociado un valor de probabilidad $P(\omega)$, de manera que $0 \leq P(\omega) \leq 1$, $P(\emptyset) = 0$, $P(\Omega) = 1$ [1].

La optimización estocástica se caracteriza por la descripción de variables aleatorias que está relacionada con Ω . En algunos casos, ciertamente, los elementos de $\omega \in \Omega$ son usados para describir algunos estados, es decir, los escenarios. Entonces, todas las variables aleatorias conjuntamente dependen de todos estos escenarios. Es importante reconocer las variables aleatorias en el problema [1].

Para una variable aleatoria particular ξ , se define su distribución acumulada como $F(x) = P(\xi \leq x)$, o bien, $F(x) = P(\{\omega: \xi(\omega) \leq x\})$. Para el caso donde las

variables aleatorias son continuas, pueden ser descritas por su función de densidad f . La probabilidad de ξ en un intervalo $[a, b]$ es obtenida por [1]:

$$P(a \leq \xi \leq b) = \int_a^b f(t)dt,$$

o de igual forma, por su función de distribución de probabilidad, con

$$P(a \leq \xi \leq b) = \int_a^b dF(t), \quad (1.2)$$

donde $F(t)$ es la distribución acumulada de ξ . La probabilidad de un valor $P(\xi = a)$ siempre es cero para una variable aleatoria continua. La distribución $F(t)$ debe ser tal que $\int_{-\infty}^{\infty} dF(t) = 1$.

El valor esperado de la variable aleatoria continua es $E(t) = \int_{-\infty}^{\infty} t F(t)dt$.

1.3.2 Los problemas de optimización con incertidumbre

Los problemas de optimización lineal estocástica son problemas lineales de optimización donde algunos de los datos incorporan incertidumbre, y se representan con variables aleatorias en el modelo general del problema. Los problemas estocásticos de recurso, son aquellos donde algunas decisiones o acciones de recurso, pueden ser tomadas una vez que la incertidumbre se realiza, es decir, los valores particulares que las variables aleatorias toman, serán conocidas una vez que se realicen, es decir, $\xi = \xi(\omega)$. La incertidumbre es caracterizada por una distribución de probabilidad, o su densidad.

Los conjuntos de decisiones en el problema estocástico de recurso son:

En la primer etapa: Las decisiones que tienen que tomarse sin conocer lo que pasará con las variables inciertas durante la incertidumbre. Esas decisiones, como son tomadas en el periodo de la primera etapa del problema, son llamadas decisiones de la primera etapa.

En la segunda etapa: Las decisiones que pueden ser tomadas una vez que se conocen, o se realizan las variables estocásticas. Son las decisiones de la segunda etapa.

Las decisiones de la primer etapa se representan con el vector x , las decisiones de la segunda etapa se representan con el vector $y(\omega)$. La secuencia de las decisiones es como sigue:

$$x \rightarrow \xi(\omega) \rightarrow y(\omega).$$

El tomar las decisiones en la primera y segunda etapas, se relaciona a que se toman antes y después del experimento aleatorio, respectivamente.

Aunque la incertidumbre está definida en la modelación del problema, puede detallarse en escenarios (posibles resultados de los datos) para especificar y precisar las distribuciones de probabilidad conjuntas para esa variable aleatoria en ese escenario.

Cuando algunos datos son aleatorios, el valor objetivo del problema de optimización incluye también estas variables aleatorias. Una forma de modelar el problema sería que al tomar una decisión ahora, la consecuencia que ocurriera sería minimizar el costo esperado (o utilidad), en la función objetivo. A esto se le llama modelo de recurso. Se supone un vector x de decisiones que debemos tomar ahora, $y(\omega)$ es el vector de decisiones que representan las nuevas acciones que realizaron o consecuencias, de tomar esa decisión x . Nótese que será una diferente de y para cada posible resultado de ω .

La formulación para el modelo de dos etapas [1]:

$$\text{Min } z = c^T x + E[\text{Min } q^T y] \quad (1.3)$$

$$\text{s. a. } Ax = b, \quad (1.4)$$

$$T(\omega)x + Wy(\omega) = h(\omega), \quad (1.5)$$

$$x \geq 0, y(\omega) \geq 0. \quad (1.6)$$

Las decisiones de la primera etapa están representadas por el vector x , de tamaño $n_1 \times 1$. Los vectores y matrices correspondientes al vector x en la primera etapa son c, b, A de tamaños $n_1 \times 1$, $m_1 \times 1$, y $m_1 \times n_1$, respectivamente [1].

En la segunda etapa, un resultado aleatorio $\omega \in \Omega$ podría realizarse. Si se realiza el resultado ω , los datos de la segunda etapa $q(\omega), h(\omega)$ y $T(\omega)$ entonces ya son conocidos, q, h y T son vectores de tamaños $n_2 \times 1$, $m_2 \times 1$, y $m_2 \times n_1$, respectivamente. Cada componente de q, h y T hacen posible una realización de la variable aleatoria: q representa el vector de costo, W la matriz tecnológica de recurso, la cual es una matriz de aleatoriedad, pero se supone que es fija, es decir, es recurso fijo. Tenemos que $T_i(\omega)$ es el i -ésimo renglón de $T(\omega)$.

Uniendo todos los componentes estocásticos, obtenemos un vector $\xi^T(\omega) = (q(\omega)^T, h(\omega)^T, T_1(\omega), \dots, T_{m_2}(\omega))$, con $N = n_2 + m_2 + (m_2 n_1)$ componentes [1]. Donde m son los renglones, y n las columnas.

Como se mencionó, un resultado aleatorio (realización) simple ω , influye en diversas variables aleatorias, que son los componentes de ξ . Cuando ω se realiza, los datos de la segunda etapa (q, h, T) se conocen, es entonces cuando la decisión de recurso o correctiva, $y(\omega)$ de la segunda etapa debe tomarse, se refleja en la función objetivo [1].

La función objetivo contiene un término determinista $c^T x$, además de un valor esperado del objetivo en la segunda etapa $q(\omega)^T y(\omega)$ que absorbe las realizaciones del evento aleatorio ω , es decir, los escenarios.

Al tomar la decisión en la primer etapa en el problema estocástico, una realización ω dada de la variable aleatoria ξ en la función objetivo, es:

$$Q(x, \xi(\omega)) = \underset{y}{\text{Min}} \{q(\omega)^T y(\omega) \mid Wy(\omega) = h(\omega) - T(\omega)x, y \geq 0\}, \quad (1.7)$$

donde $Q(x, \xi(\omega))$ es el valor de la función objetivo en la segunda etapa.

Entonces, se define que $Q(x)$ es el valor esperado del costo óptimo en la función objetivo, asociado con una decisión x de la primer etapa. El valor esperado del objetivo en la segunda etapa es:

$$Q(x) = E[Q(x, \xi(\omega))], \quad (1.8)$$

y la forma equivalente del problema determinista (PDE) es [1]:

$$\text{Min } z = c^T x + Q(x) \quad (1.9)$$

$$\text{s. a. } Ax = b, \quad (1.10)$$

$$x \geq 0.$$

Esta representación del modelo PDE del problema lineal estocástico, muestra la diferencia del modelo general del problema de optimización lineal: en el problema lineal estocástico se agrega la parte de la segunda etapa en la función objetivo, que contiene información de las variables estocásticas.

De este término $Q(x)$ en la segunda etapa, viene la mayor dificultad en el modelo estocástico porque, para cada ω , es decir, si ocurre cada realización de la variable estocástica en cada escenario, el valor $y(\omega)$ es una solución para un problema lineal. Por lo que, para lidiar con esto, puede hacerse uso del problema determinista equivalente [1]. De aquí la dificultad del problema estocástico, pues la dimensión del problema se hace cada vez mayor al incrementar el número de variables estocásticas y, si ocurre la realización de cada variable estocástica para cada escenario, se tendría un problema lineal para cada realización, es decir, un problema determinista equivalente (lineal) de gran dimensión.

El problema de optimización estocástico en su forma equivalente se vuelve un problema de gran dimensión, pues por cada una de las realizaciones de cada una de las variables estocásticas, se formula un problema determinista equivalente.

El problema estocástico de dos etapas modela una situación en la que existe información que es incierta en la primera etapa, y tiene que tomarse una decisión en el mismo momento, es hasta entonces que, esa información se completa la durante la segunda etapa.

Las propiedades de la estructura de la función objetivo en un problema lineal estocástico de dos etapas son conocidas: continua, convexa, diferenciable.

1.3.3 Complejidad de un problema estocástico de dos etapas

La complejidad de un problema de optimización, en términos de tiempo o espacio en memoria para resolverlo, está relacionada con el tamaño (número de variables) del problema.

Para cada instancia, el número de operaciones elementales de cómputo requeridas o la cantidad del espacio de almacenamiento requerido para resolver la instancia del problema como función del tamaño de la entrada (número de variables, es decir, dimensión del problema) indica, respectivamente, el tiempo o espacio de solución, que reflejan la complejidad del problema. En el caso del problema estocástico de dos etapas, la complejidad se relaciona con el número de variables de incertidumbre.

La forma en que los parámetros de aleatoriedad están descritos en los problemas estocásticos, tiene un importante impacto en la complejidad del problema. En la forma del problema determinista equivalente, bajo cierto modelo específico en las variables aleatorias, los problemas de optimización estocástica de dos etapas pueden ser *NP*-difíciles [6].

Evaluar la función $Q(x)$ en la segunda etapa del objetivo, en un simple punto de su dominio requiere resolver una integral múltiple, cuando las variables aleatorias son continuamente distribuidas. Esta es una característica de su complejidad, la cual, es un gran obstáculo. En los problemas de dos etapas, la evaluación de $Q(x)$ es *NP*-difícil [6].

1.3.4 Complejidad del problema determinista equivalente

La elección del método de solución para problemas estocásticos, comienza al formular el problema determinista equivalente. Para formular este problema, básicamente, las realizaciones de los parámetros de aleatoriedad se especifican en forma de escenarios. Cada escenario contiene una descripción completa de (q, T, h) valores en una realización. De aquí, los escenarios son enumerados $(q^1, T^1, h^1), \dots, (q^k, T^k, h^k)$, donde k es el número total de posibles realizaciones de (q, T, h) . Cada realización (q^k, T^k, h^k) tiene una probabilidad p^k

de ocurrir [6]. El modelo estocástico en forma de escenarios con sus realizaciones es:

$$\text{Min } c^T x + \sum_{k=1}^K p^k(q^k)y^k$$

$$\text{s.a. } Ax = b$$

$$T^k x + W y^k \leq h^k, \quad k = 1, \dots, K. \quad (1.12)$$

Si cada escenario y su correspondiente probabilidad tiene que ser especificada completamente en el problema, entonces el tamaño del problema es, el tamaño indicado por todas las variables aleatorias que conforman este problema determinista equivalente [6].

Si se considera que todas las variables aleatorias son independientes e idénticamente distribuidas y, usando m_1 para el número de renglones de las T -matrices, hay $K = 2^{n_1+m_1 n+m_1}$ escenarios posibles. El tamaño del problema determinista equivalente crece exponencialmente con el tamaño de las variables estocásticas, y la complejidad con el tamaño del problema es generada cuando se especifican los escenarios en combinación con las realizaciones de las variables aleatorias independientes.

Ciertamente, los problemas estocásticos se enfocan en los métodos para tratar esta dificultad, que usualmente, es causada por la especificación de los escenarios y realizaciones de las variables aleatorias independientes [6].

1.3.5 La complejidad de los problemas de dos etapas con parámetros de aleatoriedad continuamente distribuidos

Para problemas estocásticos con variables aleatorias continuamente distribuidas, la complejidad computacional de una evaluación del valor esperado $Q(x)$ de la función, puede ser establecida, incluso bajo las más generales condiciones en las distribuciones [6]. Se ha demostrado la complejidad del problema estocástico de dos etapas:

Teorema 1.3.5.1 La evaluación de $Q(x)$ de un problema de optimización estocástico de dos etapas con parámetros continuamente distribuidos es NP-difícil, incluso si todos los parámetros estocásticos tienen distribución uniforme $[0,1]$.

Por lo que se tiene que, la complejidad de los problemas de optimización de dos etapas es NP-difícil [6].

1.3.6 La dificultad del problema estocástico con múltiples variables

La dificultad de un problema estocástico depende del número de variables estocásticas y las realizaciones en los escenarios. Por ejemplo, si m es el número de escenarios con n variables de incertidumbre en el problema de dos etapas, supóngase que alguna de demanda de un producto ocurre con baja, media y alta demanda. El número de escenarios son tres $m = \{-1, 0, 1\}$, y el número de variables aleatorias del problema es $n = 2$, entonces el número total de escenarios que hay que considerar en el problema son las realizaciones de la demanda del vector de las dos variables, el vector $[n1, n2]$:

$$[-1, -1], [-1, 0], [-1, 1], [0, -1], [0, 0], [0, 1], [1, -1], [1, 0], [1, 1]$$

Por lo que el número de escenarios es $[m]^n = [3]^2$, es decir, el número de escenarios crece exponencialmente con el número de variables estocásticas del problema, de aquí viene que la complejidad computacional del problema estocástico: NP-difícil. El número de escenarios crece exponencialmente con la dimensión del problema, es decir, el número de variables.

La integral de un problema de optimización estocástica involucra una función de muchas variables, con dimensión mayor a uno, y no es fácil de resolver. Por dos razones.

La primera: el número de evaluaciones de la función que se necesitan para muestrear un espacio de N –dimensión, de N variables, se incrementa a la N potencia del número de evaluaciones necesarias sólo para resolver la integral de 1 –dimensión. Por mencionar un ejemplo, si necesitas 30 evaluaciones de

la función para resolver la integral de 1-dimensión, probablemente se necesitarían realizar evaluaciones del orden de 30,000, para resolver una integral de 3-dimensión.

La segunda: la región de integración en un espacio de N -dimensión está definida por una frontera de $N-1$, la cual puede ser terriblemente complicada. En contraste, una integral con una frontera de 1-dimensión consiste de dos números: los límites inferior y superior [4].

Considerando la forma general del problema estocástico de dos etapas, donde $Q(x)$ es el valor esperado de la función objetivo en la segunda etapa, es decir, de la función de recurso, cada valor de $Q(x)$ en el problema de optimización estocástica, puede involucrar una integral multidimensional de dimensiones extremadamente grandes.

La forma básica del problema estocástico es:

$$Q(x) = \int_{\Omega} Q(x, \omega) P(d\omega). \quad (1.13)$$

Nos referimos, como anteriormente, a un problema que no tiene restricciones probabilistas. Donde $P(d\omega)$ indica una medida de convergencia general de la probabilidad. La mayor dificultad es realizar las evaluaciones de la función $Q(x, \omega)$, las cuales involucran una múltiple integración de un número potencialmente grande de variables aleatorias [1].

La mayoría de los métodos de integración numérica están contruidos con fórmulas que aplican en funciones con dimensiones pequeñas; los resultados que se obtienen no son muy precisos o no se ajustarían exactamente al integrando $Q(x, \omega)$. Esta función es la parte fundamental en el problema de dos etapas, que es de recurso simple o fijo, en función de ω .

Cuando se tiene una integral multidimensional, tiene que revisarse si puede ser reducida analíticamente a una dimensión más baja, para poder resolverla.

Por dar un ejemplo, la integral de una función de una sola variable $f(t)$, podría reducirse en integrales de 1-dimensión por la fórmula [4]:

$$\int_0^x dt_n \int_0^{t_n} dt_{n-1} \dots \int_0^{t_3} dt_2 \int_0^{t_2} f(t_1) dt_1 \quad (1.14)$$

$$= \frac{1}{(n-1)!} \int_0^x (x-t)^{n-1} f(t) dt$$

La función en la integral del problema puede tener especial simetría, depende de sus variables independientes. Si la frontera también tiene simetría, entonces la dimensión puede ser reducida.

Por otro lado, se puede aplicar la integración por Monte Carlo en el problema. Este método se puede programar, se necesita solamente conocer una región con simples fronteras que incluya la complicada región de integración, además de un método para determinar si un punto aleatorio está dentro o fuera de la región de integración. La integración de Monte Carlo evalúa la función en una muestra aleatoria de puntos y hace la estimación de su integral basándose en esa muestra. El método de Monte Carlo es por naturaleza asintóticamente lento para converger [4].

Se puede romper la integral en varias regiones, y así, se resuelven por separado. Si no se sabe dónde están estas regiones, la conclusión sería que no se puede resolver, pues no tiene buenas esperanzas aplicar alguna rutina de integración en un enorme espacio N -dimensional [4].

1.3.7 Método Monte Carlo en el problema de optimización estocástica

Como vimos, cada valor de la función en un problema estocástico puede involucrar una integral multidimensional de dimensiones extremadamente grandes. Monte Carlo es un método que ofrece buenos resultados, como una elección para aplicar en problemas estocásticos.

Ahora bien, si consideramos que la forma básica del problema estocástico, en su valor esperado es:

$$(1.15)$$

$$Q(x) = \int_{\mathbb{R}^N} g(x, s) f(s) ds,$$

donde:

$$g(x, \xi(\omega)) = Q(x, \omega).$$

Observemos que, si $\xi(\omega_1) = \xi(\omega_2)$, entonces, $Q(x, \omega_1) = Q(x, \omega_2)$.

En 1.15, el término $g(x, s)$ es el valor de la función objetivo en la segunda etapa, cuando s es el valor de la variable aleatoria ξ . Y esta función depende de los valores de la realización ω . Supóngase que el problema estocástico (en ecuación 1.3) tiene solución óptima x^* .

Para resolver (1.15) se considera hacer una aproximación por medio de un problema derivado por ν muestras de ξ , de manera que permitiría aplicar obtener convergencia de las soluciones óptimas de los ν problemas [1].

Por medio de muestreo, el principal objetivo es reemplazar una muestra usando la distribución de $\xi(\omega)$, con otra que use una distribución alternativa. Se trata de encontrar [1]:

$$Q(x) = \int_{\mathbb{R}^N} g(x, s) f(s) ds \quad (1.16)$$

El método de Monte Carlo genera cada muestra ξ^i de acuerdo a la distribución dada por la densidad f . Donde se utiliza, para realizar el muestreo, la densidad de probabilidad $f(s)$ [1].

Por lo que entonces, Monte Carlo considera hacer muestras ξ^i de observaciones de ξ independientes, que son usadas en el problema de aproximación:

$$\inf_{x \in X} \frac{1}{\nu} \sum_{i=1}^{\nu} g(x, \xi^i). \quad (1.17)$$

En donde, x^ν es el vector aleatorio de soluciones, con muestras aleatorias independientes ξ^i , $i = 1, \dots, \nu$.

La gran cuestión es encontrar una distribución tal de manera que Monte Carlo pueda converger aunque pues se sabe que el método de Monte Carlo es, por naturaleza, asintóticamente lento para converger.

El método de solución que se propone en este trabajo, ofrece ser una herramienta eficiente para encontrar buenas soluciones, encontrando las distribuciones en el problema estocástico, incluso de múltiples variables.

1.4 OPTIMIZACIÓN HEURÍSTICA

En optimización global, no se sabe para muchos problemas si son fáciles de resolver o podría encontrarse una solución para ellos en un tiempo computacional polinomial, para encontrar su solución exacta. Por lo que, los algoritmos heurísticos son de gran ayuda en estos casos, pues nos brindan soluciones aproximadas en un tiempo computacional razonable y se puede modelar el problema con un factor aleatorio de búsqueda en el espacio de soluciones.

Dada la dificultad para resolver problemas de optimización de forma exacta, por otra parte, es necesario ofrecer alguna solución dado su interés práctico.

Comenzaron a aparecer algoritmos que proporcionan soluciones factibles (es decir, que satisfacen las restricciones del problema), las cuales, aunque no optimicen la función objetivo, se supone que al menos se acercan al valor óptimo en un tiempo de cálculo razonable. Podríamos llamarles en lugar de óptimas, “satisfactorias” o “buenas”, pues al menos es de suponer que son lo suficientemente buenas como una solución útil [3].

Este tipo de algoritmos se denominan heurísticos, del griego Heuriskein, encontrar o buscar. Su interés práctico como herramienta útil que da soluciones a problemas reales, fueron logrando aceptación, sobre todo a partir de la mitad de los años setenta, por los resultados en el campo de la complejidad computacional [3].

Una posible manera de definir estos métodos es como “procedimientos simples, a menudo, basados en el sentido común, que se supone ofrecerán una buena solución (aunque no necesariamente la óptima) a problemas difíciles, de un modo fácil y rápido” [3].

Son varios los factores que pueden hacer interesante la utilización de algoritmos heurísticos para la resolución de un problema [3]:

- a. Cuando no existe un método exacto de resolución o éste requiere mucho tiempo de cálculo o memoria. Ofrecer entonces una solución que solo sea aceptablemente buena resulta de interés frente a la alternativa de no tener ninguna solución en absoluto.
- b. Cuando no se necesita la solución óptima. Si los beneficios que adquiere la función objetivo son relativamente pequeños, puede no merecer la pena esforzarse (por el costo en tiempo y dinero), en hallar una solución óptima que, por otra parte, no representa un beneficio importante respecto a una que sea simplemente sub-óptima. En este sentido, si el heurístico puede ofrecer una solución mejor que el actualmente disponible, que se traduzca en beneficios sobre la función objetivo, esto puede ser ya de interés suficiente en muchos casos.
- c. Cuando los datos son poco confiables, o bien, cuando el modelo es una simplificación de la realidad, puede carecer de interés buscar una solución exacta, dado que por sí ésta no será más que una aproximación de la real, al basarse en datos que no son los reales.
- d. Cuando limitaciones de tiempo, espacio (para almacenamiento de datos), etc., obliguen al empleo de métodos de rápida respuesta, sin importar mucho la precisión de la solución.
- e. Como paso intermedio en la aplicación de otro algoritmo. A veces son usadas soluciones heurísticas como punto de partida de algoritmos exactos de tipo iterativo.

Una importante ventaja que presentan los heurísticos respecto a las técnicas que buscan soluciones exactas es que, por lo general, permiten una mayor flexibilidad para el manejo de las características del problema. No suele resultar complejo diseñar algoritmos heurísticos, en lugar de considerar funciones lineales que utilicen no linealidades. Generalmente ofrecen más de una solución, lo cual permite ampliar las posibilidades de elección del que decide, sobre todo cuando existen factores no cuantificables que no han podido ser añadidos en el modelo, pero que también deben ser considerados [3].

Por otra parte, suele ser más fácil de entender la fundamentación de las heurísticas que los complejos métodos matemáticos que utilizan la mayoría de las técnicas exactas [3].

Por el contrario, también presenta inconvenientes el uso de métodos heurísticos. Uno de ellos es que, por lo general, no es posible conocer la calidad de la solución x_{heu} que nos ofrecen, es decir, qué tan cerca está del óptimo x^* .

Si, por ejemplo el problema es de maximización, y lo único que sabemos es que $x_{heu} \leq x^*$. Afortunadamente, existen métodos para realizar acotaciones que nos den una orientación respecto a la calidad de la solución obtenida.

Un procedimiento consiste en relajar el problema (si bien, eliminando algunas de las restricciones o efectuando una relajación Lagrangeana) de modo que así el problema sea más fácil de resolver. Si el óptimo del problema relajado es x' sabemos que $x_{heu} \leq x^* \leq x'$, ya que al eliminar restricciones aumenta el conjunto de soluciones, y puede entonces aparecer un nuevo óptimo mejor que el original. De este modo, valores x' cercanos a x_{heu} nos garantizan que la heurística está dando una buena aproximación [3].

Cuando este tipo de procedimientos evaluadores de la calidad de la heurística no son posibles, siempre cabe utilizar métodos sencillos que detectan simplemente que la heurística no es buena. Así, podrían generarse aleatoriamente varias soluciones y si son similares a la x_{heu} , cabría poner en duda la efectividad de la heurística [3].

Sin embargo, y a pesar de todas sus ventajas, no cabe duda de que una técnica exacta debe ser preferida sobre cualquier tipo de heurístico sobre todo cuando, pequeñas variaciones respecto al óptimo, representen beneficios o pérdidas en millones [3].

1.4.1 Optimización heurística mediante búsqueda estocástica

El heurístico más simple aplicando búsqueda aleatoria en el espacio de solución es (como en una ruleta al tirar dardos), buscar puntos aleatoriamente en el espacio (llenar el espacio de búsqueda con tantos tiros), evaluar cada uno de ellos en la función objetivo del problema, y realizar esto hasta dar con el óptimo global.

Esta es una heurística ciega de búsqueda estocástica que asegura que al pasar el tiempo, encerrará el óptimo global con probabilidad de 1, según teorema de convergencia de búsqueda estocástica, que dice: con probabilidad 1, el algoritmo converge con el óptimo en un número infinito de iteraciones, ver [8], lo cual es una justificación para su uso. Esto es una propiedad del continuo pues las variables aleatorias distribuidas uniformemente en un espacio continuo llenan densamente este espacio continuo, por lo que al pasar el tiempo encontrarás el óptimo global. La gran desventaja es, el tiempo necesario o cuántas veces iterar, para poder atinar por fin al óptimo global.

La búsqueda estocástica encuentra con probabilidad de 1 el óptimo global en un tiempo esperado infinito. Por lo que la búsqueda aleatoria en la heurística miope esperaría encontrar el óptimo global en un tiempo esperado infinito.

Con el propósito de orientar a esta búsqueda ciega debido al tiempo esperado de solución, se trata de construir heurísticos más sistemáticos que combinen aleatoriedad con alguna dirección que califique ciertas regiones en el espacio de búsqueda que sean preferibles sobre otras. Ésta es la esencia de muchos heurísticos. Los métodos de optimización heurística utilizan para este propósito, una fase de diversificación además de otra fase de adaptación.

1.4.2 Adaptación y diversificación

Los algoritmos heurísticos pueden tener dos fases en la búsqueda de soluciones: adaptación y diversificación.

La diversificación es la parte aleatoria dentro de la búsqueda, es como tirar un dardo al azar en el espacio de soluciones (como en el algoritmo de búsqueda ciega), y después de encontrarla, se puede realizar la fase de adaptación para mejorar la solución buscando otra mejor, pero localmente.

Si un heurístico cuenta con ambas características, busca una solución aleatoriamente, luego cuando la encuentra, se “adapta”, es decir, busca una mejor solución local a la obtenida, una vez encontrada, ésta es la nueva mejor solución, y se hace de nuevo la diversificación. Por lo que el algoritmo heurístico es capaz de encontrar una solución, mejorarla localmente y luego salirse de esa zona con la diversificación para ir a explorar otras regiones del espacio de soluciones y no quedar atrapado en un óptimo local en la adaptación.

En la adaptación se encuentra una solución por medio de búsqueda local con algún método determinista que asegure mejorar el valor del objetivo. Un método podría ser el gradiente negativo: la dirección del gradiente negativo es donde localmente las derivadas son decrecientes, es una función de decrecimiento de la función objetivo. Sólo se observa la dirección del gradiente negativo hasta tanto como disminuya el objetivo, y entonces se tiene un óptimo local en la parte adaptativa. Al realizar repetidamente estas fases caemos en una zona de atracción del óptimo global.

Por ejemplo, en un problema de minimización utilizando el gradiente negativo. Al buscar aleatoriamente en la región de soluciones veremos que la función objetivo es continua y diferenciable, si calculamos la dirección del gradiente, éstas son propiedades que comparten muchos problemas de optimización, continuos y diferenciables, en su valor objetivo. Si se disminuye el objetivo en esa dirección, se llegará a un punto tal que sea el de mayor beneficio en la función objetivo hasta disminuir tanto antes que el objetivo vuelva a crecer con la dirección del gradiente, toda esa zona es de atracción para ese punto

óptimo local. En problemas con esta zona de atracción, se espera que el punto caiga en una zona de atracción del óptimo global. Entonces combinando adaptación con diversificación se tiene la esperanza de que en un tiempo no infinito sino menor, se encuentre el óptimo global.

Los heurísticos en general tienen una parte aleatoria o de diversificación en el espacio de soluciones que explora globalmente (fuera de óptimos locales), y otra parte de reglas deterministas que buscan mejorar el objetivo localmente en regiones específicas del espacio de búsqueda.

Ahora que si se cuenta con alguna solución de un espacio de soluciones propuestas y combinas las dos soluciones para obtener otra nueva que sea mejor, la tomas, y así repetidamente esta combinación de soluciones para seguir mejorando, Combinar soluciones para crear otras nuevas, es análogo a moverse en la vecindad de un punto (búsqueda local) combinando soluciones parecidas y se toma el mejor punto nuevo como solución, que es la parte adaptativa. Al encontrar este punto, se hace diversificación escogiendo al azar una región no muy lejana del mejor punto adaptado para buscar mejorar, y luego entonces hacer de nuevo adaptación combinando las dos soluciones.

Al combinar adaptación (mejora local) con diversificación (búsqueda aleatoria en regiones locales) se producen buenos resultados comenzando con la búsqueda para encontrar una zona de atracción buena, esto es lo que utilizan algunos heurísticos con elementos aleatorios.

Los algoritmos heurísticos son de gran ayuda en la solución de problemas de optimización.

Capítulo 2

OBJETIVO

Demostrar que el método de muestreo de Fokker-Planck es una buena alternativa y una herramienta heurística nueva en la solución eficiente de problemas de optimización estocástica.

Capítulo 3

MOTIVACIÓN

Los algoritmos heurísticos son de gran ayuda en la solución de problemas de optimización. Debido al tiempo esperado de solución, con el propósito de orientar la búsqueda ciega, se busca desarrollar o aplicar heurísticos más sistemáticos que combinen aleatoriedad junto con alguna dirección, que califiquen ciertas regiones en el espacio de búsqueda y que sean preferibles sobre otras regiones no tan favorables.

Esta tesis está enfocada en la aplicación de un método heurístico de esa naturaleza, que es de nueva introducción dentro de la optimización estocástica: el método de muestreo de Fokker-Planck (SFP, por las siglas en inglés). Cabe mencionar, que se introduce en optimización estocástica sin restricciones.

Se busca un esquema heurístico que englobe adaptación y diversificación combinadas, no para asegurar optimalidad global en el método, sino para obtener la distribución de probabilidad asintótica asociada a la estructura de la función objetivo del problema que nos indique la zona, que contiene o encierra el óptimo global, con muy alta probabilidad.

Es de gran motivación y relevancia desarrollar métodos alternativos para problemas de optimización estocástica, pues como se explica en capítulos anteriores, aquellos problemas estocásticos que involucran múltiples variables, se vuelven problemas de gran dimensión, lo que se traduce en un intento de resolver una múltiple integral que requiere de un alto esfuerzo computacional.

Por lo que es una importante motivación el contar con un esquema heurístico alternativo, que nos ayude a reducir el espacio en memoria o el tiempo requerido de solución, que englobe adaptación y diversificación combinadas para optimización estocástica.

Capítulo 4

JUSTIFICACIÓN

En este trabajo de investigación se busca utilizar un esquema heurístico que engloba adaptación y diversificación combinadas, no para asegurar optimalidad global en el método, sino para obtener la distribución de probabilidad asintótica asociada a la estructura de la función objetivo del problema que nos indique la zona, que contiene o encierra el óptimo global, con muy alta probabilidad, es lo que busca el algoritmo de muestreo de Fokker-Planck.

Análogamente, como en un heurístico adaptativo con diversificación, si al realizar la búsqueda de soluciones, cada una de ellas cae en una zona de atracción, las soluciones tenderán al valor de un óptimo local mediante la adaptación. Para el algoritmo Fokker-Planck, ésta es una zona donde la probabilidad en el proceso de búsqueda se ve favorecida por esta región local, si uno pintara la densidad de probabilidad para los puntos de esa región sobre la función objetivo, esta densidad será mayor en esta zona. La zona de mayor densidad o más probable es donde está el óptimo local, mientras más cerca se esté de esta zona es más probable encontrarlo pues cualquier punto será atraído por esta zona y cualquier movimiento en dirección contraria será muy poco favorable para la función objetivo.

Entonces al dibujar la densidad de probabilidad completa (no local), nos otorga los máximos asociados con los mínimos de la función objetivo, que al analizarla debe ser equivalente a analizar la densidad de probabilidad que genera un proceso de búsqueda aleatoria en un heurístico en su parte adaptativa.

Por lo que es una justificación para el uso del algoritmo de muestreo de Fokker-Planck, que ya ha sido utilizado para resolver problemas no lineales

con restricciones como Rosenbrock, Schwefel, entre otras. Ver referencia [2]. Ya que es un algoritmo que, como un heurístico, cuenta con una parte adaptativa y otra de diversificación, que busca encerrar el óptimo global construyendo la estructura de la densidad de probabilidad completa que está asociada con los mínimos en una función, es posible aplicar este método e introducirlo en el área de la optimización estocástica. Hablemos del porqué.

De $Q(x)$ en la segunda etapa del problema estocástico, viene la mayor dificultad porque, para cada variable aleatoria, es decir, para cada realización de la variable estocástica en el escenario, se formula un problema lineal. Para lidiar con esto, se hace uso del problema determinista equivalente. De aquí la dificultad del problema estocástico, pues la dimensión del problema se hace cada vez mayor al incrementar el número de variables estocásticas y, si ocurre la realización de cada variable estocástica para cada escenario, entonces, se tendría un problema lineal para cada realización, es decir, un problema determinista equivalente (lineal) de gran dimensión.

El problema de optimización estocástico en su forma equivalente se vuelve un problema de gran dimensión. Por cada una de las realizaciones en los escenarios de cada una de las variables estocásticas, se formula un problema determinista equivalente. El número de escenarios crece exponencialmente con el número de variables aleatorias.

En la optimización estocástica, implica un alto esfuerzo computacional para resolver la integral en el caso de una sola variable estocástica, además, si se tienen múltiples variables estocásticas, implica resolver una múltiple integral de un problema estocástico:

$$\text{Min} \int_{\mathbb{R}^N} Q(\vec{x}, \vec{w}, \vec{\varepsilon}) f(\vec{\varepsilon}) d\vec{\varepsilon}$$

donde \vec{x} es el vector de variables de decisión, \vec{w} el vector de variables no estocásticas, $\vec{\varepsilon}$ el vector de variables estocásticas, $f(\vec{\varepsilon})$ la densidad de probabilidad del vector aleatorio.

Si se conoce dónde están las regiones de integración, se puede romper la integral en varias regiones, así el integrando es más suave en cada una, y se resuelven por separado. Si no se sabe dónde están estas regiones, no tiene buenas esperanzas aplicar alguna rutina de integración en un enorme espacio N -dimensional, y no se puede resolver en un tiempo de cómputo razonable.

La integración de Monte Carlo evalúa la función en una muestra aleatoria de puntos y hace la estimación de su integral basándose en esa muestra. Sólo que el método de Monte Carlo es, por naturaleza, asintóticamente lento para converger.

El método de solución que se propone en este trabajo, ofrece ser una herramienta eficiente para encontrar buenas soluciones en problemas estocásticos, incluso de múltiples variables.

Ya que SFP es un algoritmo que, como un heurístico, cuenta con una parte adaptativa y otra de diversificación, que busca encerrar el óptimo global construyendo la estructura de la densidad de probabilidad completa que está asociada con los mínimos en una función, es posible aplicar este método e introducirlo en el área de la optimización estocástica, incluso en problemas donde se involucran múltiples variables aleatorias, y resolverlo para cada una de ellas.

Con el método SFP se puede cuantificar el beneficio de la región de búsqueda en el espacio de soluciones asociando probabilidades en él. Por las regiones o zonas de atracción, podemos decir que se tiene una probabilidad de ser una buena solución o una mala, de una manera sólida. Se pueden tratar problemas sin restricciones (o con ellas, sólo que no es objeto de discusión en la investigación de esta tesis), y comparar los resultados con métodos deterministas o heurísticos que ya se conocen.

El método SFP localiza las buenas regiones de solución con bases probabilistas, promete ser una buena herramienta para los problemas donde no se conoce la estructura del objetivo, problemas *no* -convexos, o cuando la distribución de la(s) variable(s) estocásticas no se conoce, por lo cual sería de

gran ayuda utilizarlo para resolver un problema estocástico en tiempo razonable.

El método SFP asegura converger la distribución de probabilidad correcta de la estructura de la función objetivo, lo que los heurísticos sólo evolucionan explorando regiones de búsqueda sin tener certeza de que será una buena zona en el espacio de soluciones o que se encontrará la solución “buena” en un tiempo determinado de cómputo.

Con SFP se conoce la tasa en que el método converge con la distribución hasta que sea la correcta. Se puede programar en paralelo para un problema de n -variables por lo que se utiliza un tiempo de n -operaciones de cómputo, de manera independiente para cada variable. Por lo que el tiempo de solución es proporcional o linealmente con el número de variables del problema. Pueden utilizarse pruebas de convergencia con estadística, pruebas paramétricas y confianzas, diferente a los heurísticos o métodos deterministas que encuentran cotas.

Los resultados obtenidos con el método SFP son determinados de manera que puede decirse que una solución sería el óptimo local o llegar al global con cierta probabilidad o confianza, o llegar con cierta certeza a una buena región en el espacio de soluciones, por medio de la densidad de probabilidad que adopta una forma conocida.

Por lo anterior, es justificable y de gran motivación, implementar el algoritmo de muestreo de Fokker-Planck en optimización estocástica, como una nueva herramienta alternativa en la solución eficiente de este tipo de problemas.

Capítulo 5

METODOLOGÍA

5.1 PRINCIPIOS DEL MÉTODO DE MUESTREO DE FOKKER-PLANCK Y RECOCIDO SIMULADO

El principio básico del algoritmo de recocido simulado es: si se mejora el objetivo se acepta tal solución, si no, se acepta con una cierta probabilidad, de tal manera que siempre hay una probabilidad de salirse de la región buena o empeorar en el objetivo, y esta probabilidad, está dada por una función que describe su densidad asociada con un proceso de búsqueda aleatorio.

En el proceso fisico-térmico de recocido, se reblandece el sólido mediante su calentamiento a una temperatura elevada, y luego lo va enfriando lentamente hasta que las partículas se van colocando por sí mismas en el estado fundamental del sólido, en el que las partículas forman retículas perfectas y el sistema está en su más bajo nivel de energía. Para cada temperatura durante el proceso de recocido, el sólido puede alcanzar el equilibrio térmico si el enfriamiento se produce muy lentamente. Si el sólido se enfría demasiado rápido, puede llegar a estados meta-estables donde existen defectos en forma de estructuras de alta energía [3].

El algoritmo de Metrópolis utiliza las técnicas de Monte Carlo para simular la evolución de un sistema físico en equilibrio térmico a una determinada temperatura, que se consigue mediante la generación de un elevado número de transiciones utilizando una distribución (de Boltzman) para describir el equilibrio.

El algoritmo realiza el paso de un estado a otro según las reglas siguientes: si el estado generado posee una energía menor que el actual, entonces se acepta el estado generado como el estado actual; en caso contrario, el estado generado se aceptará con una determinada probabilidad. Esta probabilidad de aceptación es función de la temperatura y de la diferencia entre los dos niveles

de energía. Cuando menor sea la temperatura, menor será la probabilidad de transformación en un estado de mayor energía, y cuanto mayor sea la energía del nuevo estado, menor será la probabilidad de que sea aceptado. Por lo tanto, cada estado tiene una posibilidad de ser alcanzado, pero con diferente probabilidad a diferentes temperaturas [3].

Los estados del sistema corresponden a las soluciones del problema de optimización: la energía de los estados con el criterio de evaluación de la calidad de la solución (el costo), el estado fundamental con la solución óptima, y los estados meta-estables son los equivalentes de los óptimos locales. El papel de la temperatura sería un parámetro de control.

Recocido simulado es un proceso iterativo del algoritmo de Metrópolis que se va ejecutando con valores decrecientes del parámetro de control. Es un método de búsqueda por entornos, el criterio de elección son las reglas de transición del algoritmo de Metrópolis: el algoritmo selecciona aleatoriamente un candidato del entorno de la solución actual, si el candidato es mejor que esa solución, se acepta como actual y si no, se aceptará con una probabilidad que decrece según crezca la diferencia en los costos de la solución candidata y la actual; cuando el candidato no sea aceptado, el algoritmo selecciona aleatoriamente otro candidato y se repite el proceso [3].

La aleatoriedad en la selección de la siguiente solución se realiza de manera que se reduce la probabilidad de quedar atrapado en un óptimo local. Se ha demostrado que el recocido simulado puede encontrar el óptimo global con probabilidad 1. Sin embargo, la optimalidad se alcanza tras un número infinito de iteraciones, en el peor de los casos. Por lo que, sólo puede ser aproximada la convergencia asintótica del algoritmo, que afortunadamente se calcula en tiempo polinomial [3].

Este proceso es básico: la convergencia del mínimo global en la energía observada en sistemas físicos en equilibrio térmico, tiende a cero, es decir, con el nivel de energía estacionaria en las moléculas.

5.2 PRINCIPIOS DEL ALGORITMO DE MUESTREO DE FOKKER-PLANCK

El método SFP utiliza un principio básico de equilibrio, así como en recocido simulado. En contraste, el algoritmo de muestreo SFP considera la densidad de los puntos en las regiones del espacio de soluciones, y busca converger con el óptimo global que se encuentra dentro de la zona de mayor probabilidad, dentro de la zona completa de densidad de probabilidad estacionaria, la cual, es positiva y con valor igual a 1, que tiende a 1.

Además, en la parte adaptativa y determinista del algoritmo, el método SFP considera la información obtenida por el gradiente de la función objetivo del problema, para hacer una búsqueda local de soluciones, orientando dicha búsqueda hacia una dirección.

El algoritmo del gradiente se basa en considerar la información obtenida al calcular el gradiente (denotado con el símbolo “nabla” ∇) de la función objetivo en el problema. El gradiente es un vector que tiene magnitud y sentido. Las componentes de este vector gradiente resultan de evaluar las derivadas parciales de la función objetivo en diferentes puntos del espacio de soluciones. Al evaluar el gradiente en la función objetivo, se obtiene información de la dirección de ascenso o descenso (descenso, en el caso de minimización) en puntos particulares, en ciertas zonas de la función en el espacio de soluciones, el gradiente nos indica la dirección de máximo ascenso o descenso.

Por lo que, información de la dirección del gradiente se usa en algoritmos como el de máximo descenso (en inglés, steepest-descent), con el gradiente negativo:

- Paso 1. Se tiene una solución actual x_0 , para la función objetivo $f(x)$
Hacer iteraciones desde 1 hasta n:
- Paso 2. Calcular el gradiente negativo de la función objetivo:
$$-grad(f(x_n)) = -\nabla(f(x_n))$$
- Paso 3. Minimizar en dirección del gradiente negativo, para definir un óptimo local nuevo x_{n+1} .

Paso 4. Ir al paso 2 en la siguiente iteración.

Con el método SFP, mediante el gradiente negativo, cuando exista una zona de “decrecimiento” en la función objetivo, se explora la zona local en dirección del gradiente negativo, buscando un óptimo local, minimizando de esta forma la función objetivo, siempre, y hasta, que ya no se pueda buscar en esa dirección, es decir, antes que la función objetivo comience a crecer en su valor.

Después de encontrar un óptimo local en la parte adaptativa de SFP, se realiza la exploración de otras zonas del espacio de soluciones mediante la parte de diversificación, para salir del óptimo local y no quedar atrapados en esa misma región, es decir, diversificar en busca de mejores soluciones. Con ambas partes, en la búsqueda del algoritmo SFP se puede explorar en el espacio de soluciones siempre con una dirección “buena” que asegura (con el gradiente negativo) encontrar mínimos locales y mejorar el objetivo, además de otras mejores o peores soluciones, pero en diversas regiones del espacio de solución.

De esta manera, el muestreo del algoritmo SFP encuentra buenas soluciones, cuando las encuentra en cada iteración, construye la densidad de probabilidad de esos puntos, dibujando completamente las zonas de alta y baja probabilidad de encerrar el óptimo global, hasta converger con la densidad estacionaria y correcta.

El principio de convergencia del método SFP se explica a continuación.

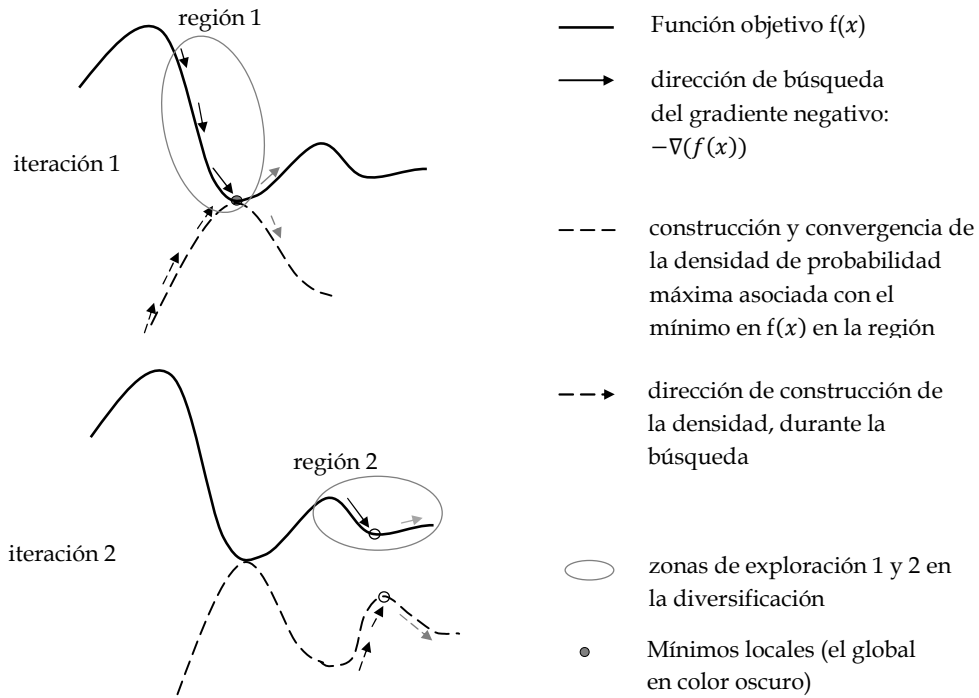


Fig. 5.2.1 Principios de la convergencia del algoritmo de muestreo de Fokker-Planck a través de las iteraciones. La densidad de probabilidad es la correcta y estacionaria.

Al maximizar la densidad $f(\vec{x}) = e^{-\frac{f(\vec{x})}{D}}$, donde D es un parámetro. Aquí se minimiza la función objetivo $f(\vec{x})$. Como $f(\vec{x})$ es negativa (mínima), $f(\vec{x})$ es máxima, donde la función exponencial negativa decae con su argumento, es monótona. Si el argumento es negativo la función será monótonamente creciente, si el argumento es positivo, será monótonamente decreciente. Recordemos que una función monótona se refiere a que conserva el orden que tiene y se observa en la forma de su gráfica, por ejemplo, si es monótona creciente significa que conserva el orden creciente en su dominio. Las funciones monótonas tienen importante aplicación en probabilidad.

Supongamos que si $f(\vec{x})$ es positiva, un mínimo global es cuando $f(\vec{x}) = 0$ entonces $f(\vec{x}) = e^0 = 1$. Cualquier otro valor de la función objetivo que no es cero sino positivo por ejemplo, $f(\vec{x}) = 1$, $f(\vec{x}) = e^{-\frac{1}{D}}$ es un número chico, y si $f(\vec{x}) = 2$ entonces $f(\vec{x}) = e^{-\frac{2}{D}}$ es más pequeño por lo que vemos que es decreciente. Entonces el valor mínimo de $f(\vec{x})$ es equivalente a maximizar $f(\vec{x})$

y tenemos un problema de optimización global, es decir, transformar el problema de minimización de la función objetivo en uno de maximizar una densidad de probabilidad.

Sería más fácil trabajar con un problema de maximización de densidad de probabilidad por sus propiedades como que es positiva, integrable a 1, y con la integral de esta función de densidad, que es la función de distribución acumulada, que es monótona, creciente, continua y diferenciable; de esta manera obtener los máximos de la densidad de probabilidad para tener la estructura de los mínimos en la función objetivo, incluso cuando ésta sea una función sin estructura o no se conozca nada sobre ella. Las propiedades de la densidad de probabilidad nos permiten trabajar mejor con esta función no importa cómo sea el objetivo del problema.

Definición 5.2.1 Una función $f(x)$ es diferenciable en $x = a$, si existe el límite $\frac{f(x)-f(a)}{x-a}$, este límite es la derivada de $f(x)$ en $x = a$.

Teorema 5.2.1 Si una función $f(x)$ tiene un extremo local en $x = a$, y $f(x)$ es diferenciable en $x = a$, entonces $f(a)' = 0$.

Se puede analizar la función de densidad sin tratar directamente con el problema multidimensional porque una densidad $f(\vec{x})$ es la densidad conjunta de $f(\vec{x}_1, \dots, \vec{x}_n)$ de x_1, \dots, x_n , que son muchas variables de decisión, donde hay procesos estocásticos que convergen a una representación donde se tiene una densidad estacionaria para cada una de las variables y se puede encontrar entonces un proceso estocástico que puede converger a esta densidad conjunta, estacionaria, para cada una de las variables. Entonces, se pueden tratar las n variables por separado, como n -problemas de una dimensión cada uno, y evitar trabajar con el problema multidimensional. Se aprovecharían las densidades estacionarias o marginales de cada variable para hacer varios problemas unidimensionales, en lugar de trabajar con la densidad conjunta y entonces, el proceso de búsqueda estocástica se realizaría en operaciones computacionales para cada problema por separado y el tiempo de solución sería mucho menor, por lo que sería más fácil resolverlo por medio de las n densidades de probabilidad.

Si tuviéramos una densidad bimodal de una variable x_n resulta que es muy probable tener un conjunto de máximos potenciales de densidad de las n – variables de decisión donde la probabilidad de encontrar mínimos en el objetivo es alta, así se definen las zonas de atracción donde hay probabilidad de encontrar buenas soluciones o una buena región en el espacio de soluciones.

Análogamente trabajan de esta forma los métodos heurísticos: ir a una buena región en el espacio de soluciones, descartar las soluciones malas, y usar regiones de búsqueda ya que no se tiene una manera exacta de cuantificar el beneficio en el objetivo. El método de muestreo de Fokker-Planck utiliza una parte adaptativa y una fase de diversificación en la búsqueda de soluciones.

5.3 ALGORITMO FOKKER-PLANCK: DENSIDAD DE PROBABILIDAD ESTACIONARIA DE PROCESOS DE BÚSQUEDA ESTOCÁSTICOS EN LA OPTIMIZACIÓN GLOBAL

En optimización global, los métodos estocásticos son esenciales para muchas técnicas heurísticas usadas en problemas complejos y sin estructura. En los algoritmos de optimización global, una característica común es la reducción progresiva del espacio de soluciones hasta que el óptimo global se ha alcanzado con cierta precisión [2]. Análogamente, los algoritmos estocásticos tienen deficiencia en esta propiedad.

Por lo que la contribución de esta tesis es utilizar un método para estimar la densidad de probabilidad asintótica de un proceso general de búsqueda estocástica, como nueva introducción en el ámbito de optimización estocástica, y para analizar los resultados de la implementación, utilizaremos un problema clásico de optimización estocástica: el problema del vendedor de periódicos (newsvendor, en inglés) que es un problema donde las variables aleatorias son continuas.

El método SFP es un método donde se construyen las expresiones analíticas para las densidades marginales estacionarias en el proceso de búsqueda estocástico de la solución del problema del vendedor de periódicos [2]. Este método no había sido aplicado en problemas de optimización estocástica, por lo que se introduce su aplicación en el problema del vendedor de periódicos donde la demanda tiene comportamientos diferentes.

La convergencia de la densidad asintótica estimada ayuda a localizar la región en el espacio de soluciones con cotas probabilistas confiables, donde se encierra el óptimo global con alta probabilidad. Para hacer la estimación de la densidad, el procedimiento requiere de un número controlado de operaciones lineales y un número bien definido de evaluaciones en la función objetivo con un costo computacional por iteración que crece linealmente con el número de variables del problema [2]. El método será utilizado en este problema continuo, como lo es problema del vendedor de periódicos. El método se puede aplicar en problemas discretos agregando algunas restricciones.

Los resultados del método muestran regiones del espacio de búsqueda que con bases probabilistas, pueden ser descartados. Esta propiedad puede ser implementada para mejorar algoritmos que ya existen o para desarrollar otros nuevos, ya que se pueden construir bases de optimalidad probabilista [2], por ejemplo, en la solución de problemas de optimización lineal estocástica.

El principal objetivo de este método no es converger con el óptimo global, sino la aproximación de la densidad de probabilidad después de un largo tiempo de exploración del espacio de soluciones, manteniendo fijo un parámetro de aleatoriedad [2].

Si se considera la función de costo $U(x_1, x_2, \dots, x_n, \dots, x_n)$ con un espacio de búsqueda definido sobre $L_{1,n} \leq x_n \leq L_{2,n}$, el proceso de búsqueda estocástico a través del tiempo para este problema está modelado como sigue:

$$\dot{x}_n = -\frac{\partial U}{\partial x_n} + \varepsilon(t), \quad (5.1)$$

donde (5.1) tiene propiedades básicas de la estrategia de búsqueda estocástica general donde se agrega $\varepsilon(t)$ que es un ruido blanco con media cero. Si existe

un ruido Gaussiano no-correlacionado de intensidad constante, la ecuación (5.1) presenta una búsqueda por difusión, en cambio, una intensidad del ruido variando en el tiempo lentamente, la ecuación sigue el proceso de recocido simulado. Cuando el ruido toma valores de amplitud infinita, la influencia dinámica de la función de costo sobre el proceso exploración se pierde, lo que conlleva a una búsqueda ciega [2].

La evolución en el tiempo de la densidad de probabilidad del sistema en la ecuación (5.1) en presencia de ruido aditivo blanco Gaussiano, está descrita por la ecuación lineal diferencial de Fokker-Planck:

$$\dot{f} = \sum_{n=1}^N \frac{\partial}{\partial x_n} \left[\frac{\partial U}{\partial x_n} f \right] + \sum_{n=1}^N \sum_{m=1}^N D_{n,m} \frac{\partial^2}{\partial x_n \partial x_m}, \quad (5.2)$$

donde D es una constante proporcional a la intensidad del ruido. El uso de esta ecuación en optimización implica el cálculo de integrales de gran dimensión para propósitos de optimización [2].

Sin embargo, resulta numéricamente menos demandante considerar la siguiente proyección de 1-dimensión de la ecuación (5.2), considerando la densidad condicional:

$$f(x_n | \{x_{j \neq n}\}) = \frac{f(x_1, \dots, x_N)}{\int f(\{x_{j \neq n}\} | x_n) f(x_n) dx_n}, \quad (5.3)$$

la cual está relacionada con la densidad marginal $f(x_n)$:

$$f(x_n) = \int f(x_n | \{x_{j \neq n}\}) f(\{x_{j \neq n}\}) d\{x_{j \neq n}\}. \quad (5.4)$$

Bajo condiciones muy generales, e.g. ausencia de valores infinitos de costo, la ecuación (5.2) tiene una solución estacionaria sobre el espacio de búsqueda con sus cotas reflejadas [2].

Introducir la ecuación (5.3) en (5.2) y considerar las variables fijas $\{x_{j \neq n} = x_j^*\}$, resulta que la densidad de probabilidad condicional estacionaria satisface la ecuación unidimensional de Fokker-Planck [2]:

$$D \frac{\partial(x_n | \{x_{j \neq n} = x_j^*\})}{\partial x_n} + f(x_n | \{x_{j \neq n} = x_j^*\}) \frac{\partial U}{\partial x_n} = 0. \quad (5.5)$$

Una característica importante en (5.5) es que la densidad marginal $f(x_n)$ puede ser muestreada dibujando puntos de la condicional $f(x_n | \{x_{j \neq n} = x_j^*\})$ por medio del muestreo de Gibbs [2]. Debido a la linealidad de la ecuación de Fokker-Planck, puede construirse una forma particular del muestreo de Gibbs, de manera que no solo es posible muestrear la densidad marginal, sino también dar una expresión analítica aproximada para ella. De la ecuación (5.5) siguiendo la ecuación lineal diferencial de segundo orden para la distribución acumulada $y(x_n | \{x_{j \neq n} = x_j^*\}) = \int_{-\infty}^{x_n} f(x'_n | \{x_{j \neq n} = x_j^*\}) dx'_n$,

$$\frac{d^2 y}{dx_n^2} + \frac{1}{D} \frac{\partial U}{\partial x_n} \frac{dy}{dx_n} = 0, \quad (5.6)$$

$$y(L_{1,n}) = 0, \quad y(L_{2,n}) = 1.$$

Desviaciones aleatorias pueden dibujarse a partir de la densidad condicional $f(x_n | \{x_{j \neq n} = x_j^*\})$ por el método de inversión, basándose en que y es una variable aleatoria distribuida uniformemente en el intervalo $y \in [0,1]$. Vista como una función de la variable aleatoria x_n , $y(x_n | \{x_{j \neq n} = x_j^*\})$ puede ser aproximada mediante una combinación lineal de funciones de un conjunto completo que satisfaga las condiciones de las restricciones en el intervalo de interés [2],

$$\hat{y}(x_n | \{x_{j \neq n} = x_j^*\}) = \sum_{l=1}^L a_{n,l} \varphi_l(x_n). \quad (5.7)$$

Digamos que una instancia fuese, una base en la que $\varphi_l(0) = 0$, los coeficientes L están definidos únicamente por la evaluación de la ecuación (5.6) en $L - 1$ puntos interiores. De esta forma, la aproximación de y se lleva a cabo resolviendo un conjunto de L ecuaciones lineales algebraicas, involucrando $L - 1$ evaluaciones de la derivada de U [2].

El procedimiento del método se basa en la iteración de los siguientes pasos:

- a. Fijar las variables $x_{j \neq n} = x_j^*$ y aproximar $y(x_n | \{x_{j \neq n} = x_j^*\})$ usando (5.6) y (5.7).

- b. Usar $\hat{y}(x_n|\{x_{j \neq n}\})$ y construir una tabla, esto para generar desviación de x_n^* dibujada desde la distribución estacionaria $f(x_n|\{x_{j \neq n} = x_j^*\})$.
- c. Actualizar $x_n = x_n^*$ y repetir pasos a, b, c para una nueva variable $x_{j \neq n}$.

Al iterar los tres pasos, se ejecuta el algoritmo para estimar la distribución de equilibrio del proceso de búsqueda estocástico descrito por (5.1). Se obtiene una representación convergente para $f(x_n)$ después de tomar el promedio de los coeficientes a 's en la expansión (5.7) sobre las iteraciones [2].

Para verlo, en la ecuación (5.4) se sigue que la marginal $y(x_n)$ está dada por el valor esperado de la condicional $y(x_n|\{x_{j \neq n}\})$ sobre el conjunto $\{x_{j \neq n}\}$,

$$y(x_n) = E_{\{x_{j \neq n}\}}[y(x_n|\{x_{j \neq n}\})]. \quad (5.8)$$

Toda la información del conjunto $\{x_{j \neq n}\}$ se almacena en los coeficientes de expansión (5.7). Entonces:

$$\langle \hat{y} \rangle = \sum_{l=1}^L \langle a_{n,l} \rangle \varphi_l(x_n) \rightarrow y(x_n), \quad (5.9)$$

donde $\langle \rangle$ representa el promedio sobre las iteraciones en el procedimiento de estimación de la densidad.

La selección de los parámetros L y D es fundamental en el procedimiento de estimación de la densidad. La constante de difusión D “suaviza” la densidad, esto lo vemos cuando el límite $D \rightarrow \infty$ en (5.6), lo que implicaría una densidad uniforme en el dominio. Por otro lado, el número de funciones base L define la capacidad del algoritmo para “aprender” la estructura de la densidad. De ahí que para un valor dado D , el número L debe ser lo suficientemente grande al menos para asegurar que el algoritmo hará la estimación de las distribuciones $y(x_n|\{x_{j \neq n}\})$ válidas. Una distribución válida debe ser una función continua creciente y monótona que satisfaga las condiciones del problema [2].

El parámetro L determina el costo computacional del procedimiento, pues en cada iteración un sistema de tamaño αL de ecuaciones algebraicas lineales se resuelve N veces. Por lo que el usuario es capaz de controlar el costo

computacional mediante la interacción de estos dos parámetros básicos: para una D más grande será estimada una densidad más suave, entonces debe usarse un parámetro L más pequeño [2].

Este método de estimación de densidad es una buena herramienta en la generación de poblaciones de puntos iniciales para algoritmos de optimización, además en la construcción de cotas probabilistas confiables en problemas de optimización [2].

El método de estimación de densidad de Fokker-Planck puede ser usado en la construcción de cotas probabilísticas para localizar el óptimo global en problemas de gran dimensión. La estimación de densidad se realiza en un buen número definido de operaciones elementales. La teoría desarrollada y los experimentos numéricos indican que cualquier precisión deseada en las cotas puede ser alcanzada con algunos valores finitos de los parámetros básicos, realizando un número finito de iteraciones. El costo computacional total por iteración crece linealmente con el tamaño del problema. El algoritmo estima la densidad marginal de cada variable por separado, lo cual hace posible la implementación en paralelo [2].

Estas características indican que el método de estimación de densidad propuesto es una herramienta prometedora pues puede ser usado para desarrollar nuevas heurísticas o mejorar las existentes.

Por lo que es importante además una motivación, el proponer este algoritmo como una nueva técnica que nos ayude a resolver problemas de optimización estocástica.

En este sentido, la contribución de esta tesis está enfocada en la implementación del método de muestreo de Fokker-Planck (SFP-sampling Fokker-Planck, por las siglas en inglés), un método de estimación de densidad estacionaria, en el ejemplo del problema del vendedor de periódicos. Este problema se escogió por ser uno de los más fáciles en optimización estocástica por lo que es de gran utilidad para hacer comparaciones entre la solución exacta que ya existe con los resultados del método SFP.

Capítulo 6

RESULTADOS

6.1 PROBLEMA DEL VENDEDOR DE PERIÓDICOS

En este capítulo nos introducimos dentro en la optimización estocástica a través de un ejemplo de gran aplicación, que nos ayudará a entender la estructura de un problema de optimización con incertidumbre, el proceso de decisión de acuerdo al objetivo y su relación con la variable aleatoria. Los resultados obtenidos en un modelo de optimización estocástica pueden llevar a significantes ahorros, lo que es una motivación para proponer la solución por medio de esta técnica.

Un problema básico en la optimización lineal estocástica es, el del vendedor de periódicos, que tiene propiedades como un problema de dos etapas.

En este general problema, el vendedor visita diariamente una editorial por las mañanas para comprar x cantidad de periódicos a un precio c por cada uno. Esta cantidad a comprar está limitada normalmente por la capacidad económica del vendedor o por un límite u establecido por la editorial para los vendedores. Entonces el vendedor camina por las calles para vender tantos periódicos como sea posible a un precio de venta p . La demanda de periódicos varía día con día y se describe por una variable aleatoria d . Se toma en cuenta que los lectores sólo están interesados en la última edición impresa. Además, el vendedor no podría regresar a la editorial y comprar más periódicos durante su recorrido en el caso de que la demanda fuera mayor. Por lo cual es de gran importancia decidir cuántos periódicos comprar diariamente por las mañanas dado que la demanda no se conoce con certeza. Las ganancias por la venta efectiva de periódicos están definidas por la variable y . El objetivo es minimizar los costos al decidir comprar cierta cantidad de periódicos diariamente sin tener la certeza del comportamiento de la demanda día con día [1].

El problema puede formularse como sigue:

$$\text{Min } cx + E[Q(x, d)]$$

$$\text{s. a. } 0 \leq x \leq u$$

$$\text{donde } Q(x, d) = \min -p y(d)$$

$$\text{s. a. } y(d) \leq d,$$

$$0 \leq y(d) \leq x,$$

donde E denota el valor esperado respecto a la demanda d . En esta notación, $Q(x, d)$ es la ganancia en ventas si la demanda es d , por lo que tenemos entonces que $E[Q(x, d)]$ es el valor esperado de las ganancias en ventas cuando la demanda es d [1].

El modelo muestra las dos etapas para el problema del vendedor de periódicos, es decir, tomar una decisión de compra antes de tener cualquier información dada por la demanda de periódicos. Cuando se conoce ya la demanda, en la segunda etapa se representa el fin del periodo de venta para la edición de periódicos, las ganancias pueden ser calculadas. Esto se realiza según la regla simple [1]:

$$y^*(d) = \min(d, x).$$

Las ventas nunca pueden exceder el número de periódicos disponibles o la demanda. Por lo que el valor esperado de la función de la segunda etapa es simplemente:

$$E[Q(x, d)] = E[-p \min(d, x)]$$

Esta función es continua, convexa y diferenciable cuando la demanda d es un vector aleatorio continuo. En este caso, existe la solución analítica para el problema [1]:

$$E[Q(x, d)] = \int_{-\infty}^x (-pd) dF_d + \int_x^{\infty} (-px) dF_d$$

$$= -px + p \int_{-\infty}^x F(d) d(d)$$

$$= -px + p[F(x)]$$

$$\text{Min } cx + [-px + p[F(x)]],$$

donde F_d es la función de distribución de probabilidad acumulada de la demanda d como es de igual manera $F(x)$, la variable x la cantidad de periódicos. La solución óptima es [1]:

$$c - p + p[F(x)] = 0$$

$$x^* = F^{-1}\left(\frac{p - c}{p}\right)$$

El problema del vendedor de periódicos es un modelo matemático de gran aplicación en la investigación de operaciones y economía, pues es usado para determinar el nivel óptimo de inventario. Si es p el precio de venta del producto, c es el costo de producción o inventario, la solución indica el nivel de inventario o cantidad óptima del producto a pedir x^* , dada una demanda que depende de las condiciones cambiantes del mercado económico. De acuerdo al nivel de inventario, puede presentarse el caso en que las unidades de demanda excedan este nivel, por lo que llevarían a pérdidas en ganancias por la escasez de producto traducidas en ventas perdidas a un precio p . De igual forma, si las unidades de demanda se encuentran por debajo del inventario, se generan costos innecesarios de producción, además de costo por almacenaje-espacio y pérdidas en ganancias del rango $(p - c)$ por ventas perdidas, debido al sobre inventario.

Si el precio de venta es menor al costo de inventario (o producción) $p < c$, el numerador resultaría en un negativo, indica que no es beneficioso mantener unidades en inventario.

Dado que existe una solución analítica para el problema del vendedor de periódicos podemos aplicar el algoritmo SFP y demostrar que funciona encontrando asintóticamente la densidad de probabilidad máxima y correcta

asociada con el mínimo para el problema. En las siguientes secciones se muestran los resultados del método para este problema.

6.2 RESULTADOS DEL ALGORITMO SFP EN EL PROBLEMA DEL VENDEDOR DE PERIÓDICOS

6.2.1 Implementación del método SFP en el problema

En el problema del vendedor de periódicos como se explicó en la sección anterior, se refiere a tomar diariamente la decisión de comprar cierta cantidad de periódicos, tomando en cuenta que existe incertidumbre en la demanda y no se sabe con certeza cómo se comportará día con día.

El algoritmo de muestreo de Fokker-Planck o SFP, no había sido probado en problemas de optimización estocástica, es probado por primera vez en problemas de optimización estocástica.

Para hacer implementación del método, se utilizaron varios métodos con librerías en lenguaje de programación java. La función objetivo con un espacio de búsqueda definido en un intervalo dado, se encuentra en un método del programa que contiene el modelo del problema, el cual necesita y manda llamar valores aleatorios (de demanda) generados en otro método. El método que contiene la función objetivo y los valores de demanda son llamados por el método de SFP.

El método SFP describe con la ecuación diferencial de Fokker-Planck, la evolución de la densidad de probabilidad del sistema dinámico de ecuaciones lineales que evoluciona con ruido blanco Gaussiano; el sistema dinámico de ecuaciones lineales viene de la ecuación del proceso de búsqueda estocástica a través del tiempo. El método se basa en la representación de las condicionales acumuladas de la solución estacionaria de esta ecuación diferencial. El método SFP manda llamar el método que contiene la función objetivo y los valores de demanda generados. Se hacen las evaluaciones para generar un sistema de ecuaciones algebraicas lineales comenzando desde el primer valor aleatorio generado de demanda, y SFP calcula las densidades marginales para

dibujar un punto nuevo que ahora tiene información de la función objetivo. Haciendo este procedimiento en la implementación, al realizarse las iteraciones se puede estimar la distribución marginal acumulada para cada variable, que en este caso es sólo una. El parámetro L y D , vienen del número de evaluaciones del sistema de ecuaciones lineales algebraicas usadas por la ecuación de Fokker-Planck y de la evaluación de la ecuación diferencial estacionaria para calcular las condicionales acumuladas [7].

Dichos parámetros L y D , involucran la parte de adaptación y de diversificación en el proceso de búsqueda estocástica dentro del método SFP, respectivamente.

Como resultado del procedimiento del algoritmo SFP con los parámetros adecuados, se obtienen puntos nuevos por medio de la densidad estacionaria y correcta, es decir, la distribución marginal acumulada de cada variable, que refleja información de la función objetivo del problema.

Por lo que al observar esta densidad dibujada, en la zona de más alta probabilidad encontraremos la zona que contiene el óptimo global mínimo para el problema del vendedor de periódicos. Se probará el método SFP en los casos donde la demanda tiene diferentes comportamientos a continuación.

Este método es una nueva aplicación en problemas de optimización estocástica. La contribución de esta tesis está enfocada en introducir el método de muestreo de Fokker-Planck en la optimización estocástica. Para esto, el método se implementa en el problema del vendedor de periódicos sin restricciones, cuyo caso es un problema clásico, además, se puede calcular su solución analítica. Esto nos permite establecer comparaciones entre dicha solución y la solución del método SFP, para varias instancias.

6.2.2 Problema con demanda unimodal

Para el problema con demanda unimodal, es decir que se comporta de un solo modo, se puede calcular la solución analítica que indica la cantidad óptima a pedir. El problema del vendedor de periódicos es un problema de minimización

donde se analiza una sola variable (cantidad de periódicos a ordenar). Con el método SFP se espera obtener una densidad de probabilidad, es decir, una distribución que contiene en la zona de más alta probabilidad el óptimo global.

La acumulada inversa de la solución analítica nos proporciona el mínimo valor esperado, debido a que la distribución de la demanda es simétrica como una Gaussiana en este caso, entonces al aplicar el método SFP obtenemos una densidad de probabilidad máxima, es decir, la distribución de los valores objetivo que casi coincide con el mínimo valor esperado que nos da la acumulada inversa.

En la tabla 6.2.2 se muestran por columnas los detalles de las instancias y las comparaciones de ambas soluciones: las columnas con la letra correspondiente de cada instancia desde A hasta G, precios de venta y costo por unidad, el intervalo que indica las cantidades entre las que ocurre la demanda, la cantidad óptima calculada en la solución analítica y en seguida, la solución del método SFP con los parámetros L y D que se usaron en el algoritmo para la parte adaptativa y de diversificación, respectivamente. Más adelante, se muestran las gráficas resultantes de cada instancia.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. Opt. SOL . AN. (unidades)	Cant. SOL SFP (unid)	Param L-D
A	p = 11.0 c = 8.0	[20,90]	39.090	39.8619	50-90
B	p = 20.0 c = 11.0	[5,120]	56.75	56.9148	50-100
C	p = 500.0 c = 125.0	[15,200]	153.75	153.914	72-2000
D	p = 20,000.0 c = 14,000.0	[250,400]	295	295.342	72-70000
E	p = 140,000.0 c = 75,000.0	[14,820]	388.214	389.01	265-500500
F	p = 40.0 c = 32.0	[32,180]	61.6	61.5906	68-70
G	p = 2,300.0 c = 1,650.0	[10,140]	46.739	46.8247	100-2000

Tabla 6.2.2 Instancias para problema del vendedor de periódicos con demanda unimodal. Se compara la solución analítica con la solución resultante del algoritmo SFP, y parámetros.

En el eje horizontal se indica el intervalo en el que se encuentra la demanda, en el eje vertical, la correspondiente densidad asintótica generada por el algoritmo. Los gráficos muestran que se genera asintóticamente la densidad correcta y máxima que está asociada con el mínimo de la función objetivo.

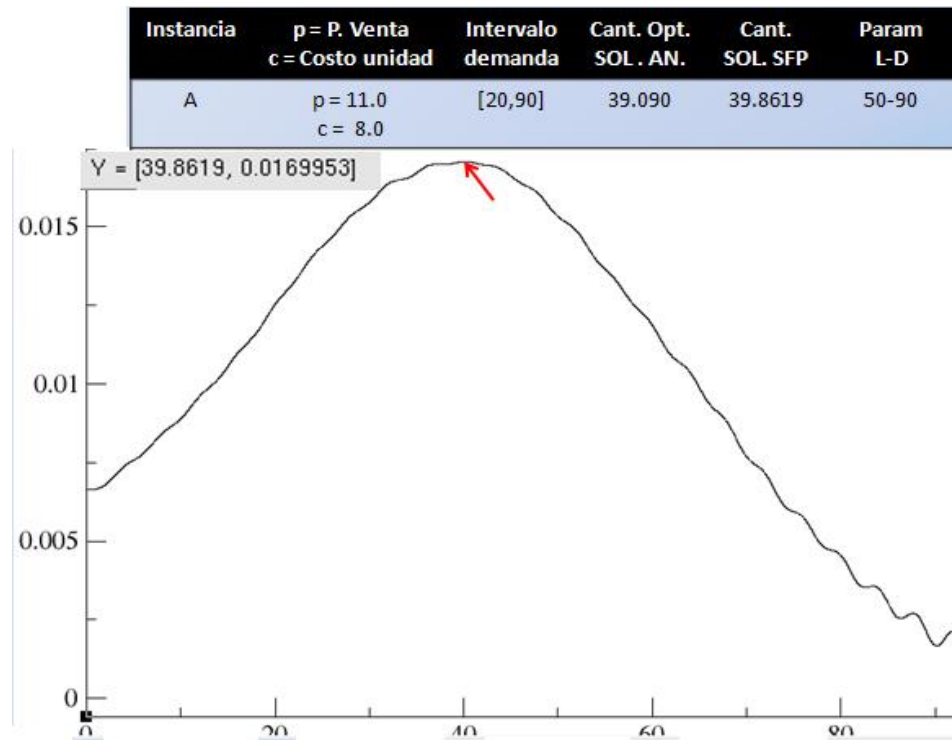


Figura 6.2.2a. Instancia A. En problema del vendedor de periódicos. La densidad máxima correcta generada por el algoritmo SFP refleja el mínimo en la solución del problema.

Analizando cada una de las curvas, se observa que la solución del algoritmo SFP refleja la estructura de la función objetivo en el problema, es decir, donde existe la zona de máxima densidad de probabilidad en la curva, ahí se encerrará el mínimo valor global del objetivo.

Para el caso del problema con demanda unimodal, como se explica en la sección anterior, la solución analítica es conocida y se puede calcular el óptimo global y como vemos, la solución que se obtiene a través del método SFP es muy aproximada al óptimo global.

Instancia	p = P. Venta c = Costo unidad	Intervalo demanda	Cant. Opt. SOL. AN.	Cant. SOL. SFP	Param L-D
B	p = 20.0 c = 11.0	[5,120]	56.75	56.9148	50-100

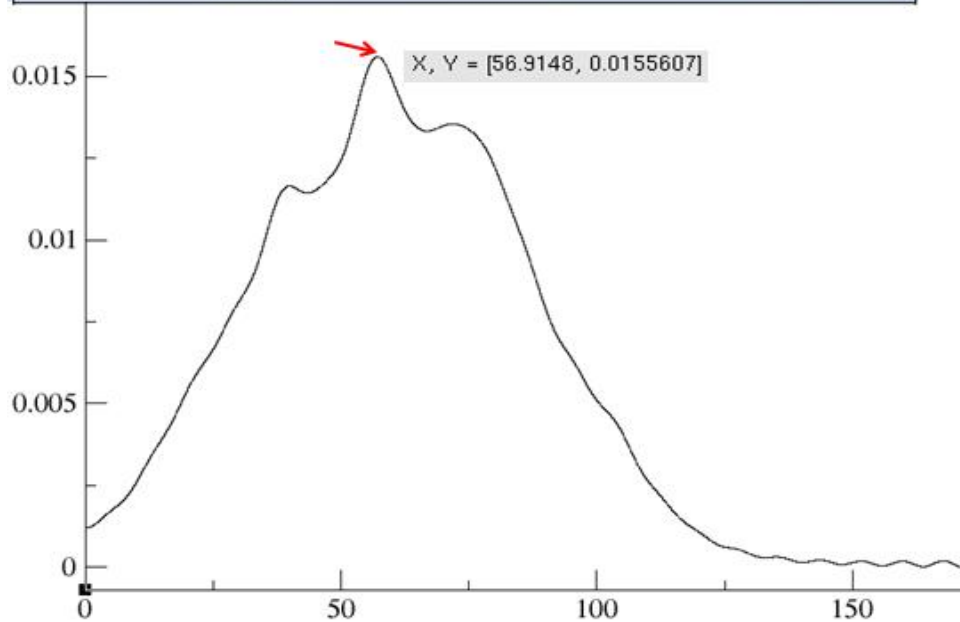


Figura 6.2.2b. Instancia B. Densidad generada para el problema del vendedor de periódicos. En los parámetros, se observa que D es mayor que L para el algoritmo SFP.

Instancia	p = P. Venta c = Costo unidad	Intervalo demanda	Cant. Opt. SOL. AN.	Cant. SOL. SFP	Param L-D
C	p = 500.0 c = 125.0	[15,200]	153.75	153.914	72-2000

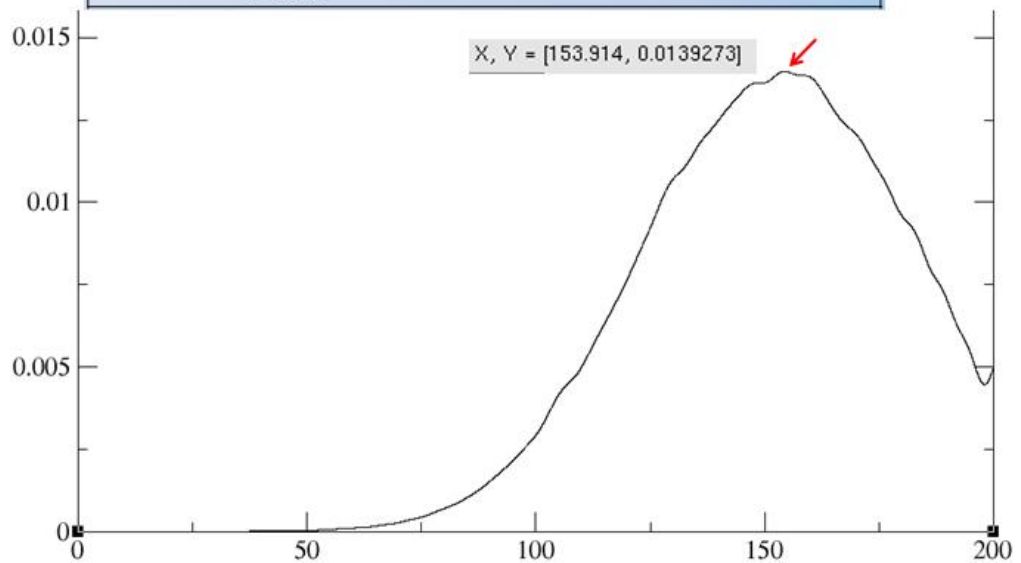


Figura 6.2.2c. Instancia C. En problema del vendedor de periódicos. Densidad generada por algoritmo SFP. Si la demanda se encuentra entre el intervalo [15, 200], la cantidad de periódicos a pedir por el vendedor está alrededor es 153 ó 154.

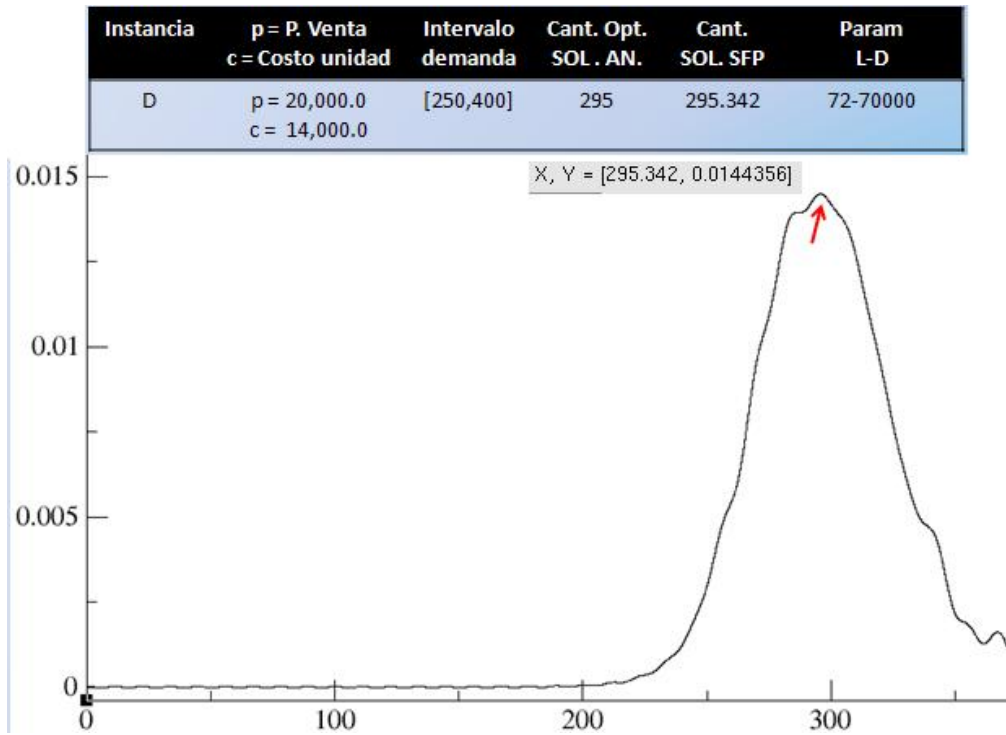


Figura 6.2.2d. Instancia D. En problema del vendedor de periódicos.

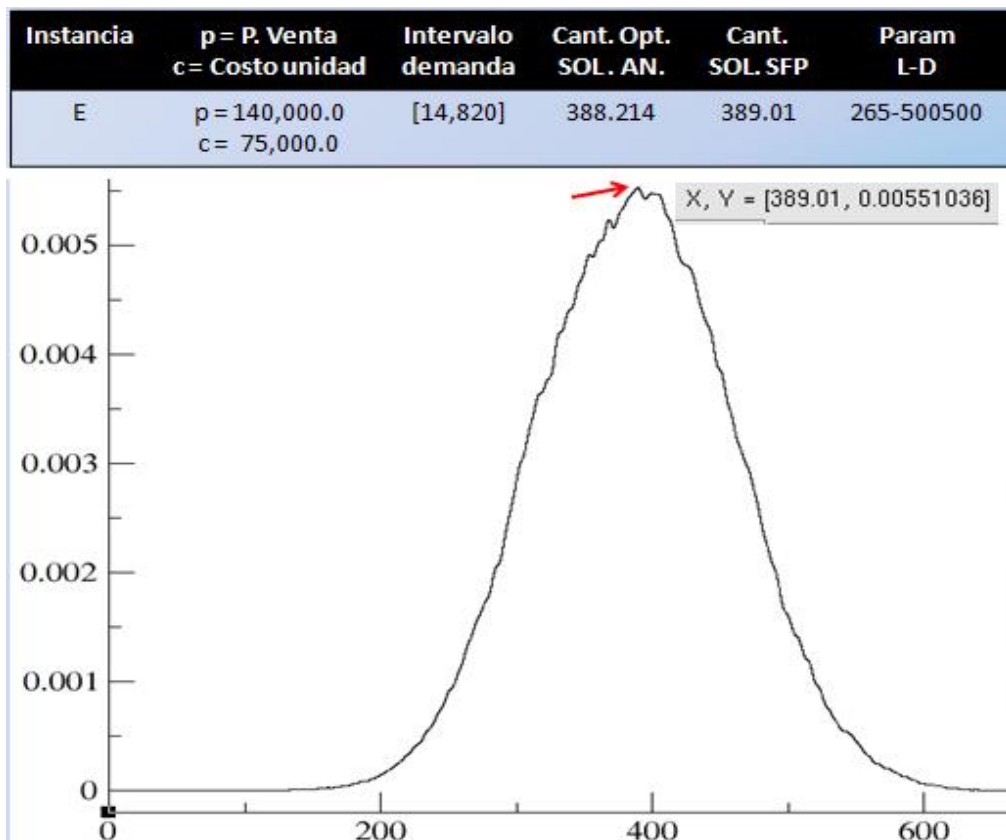


Figura 6.2.2e. Instancia E. En problema del vendedor de periódicos.

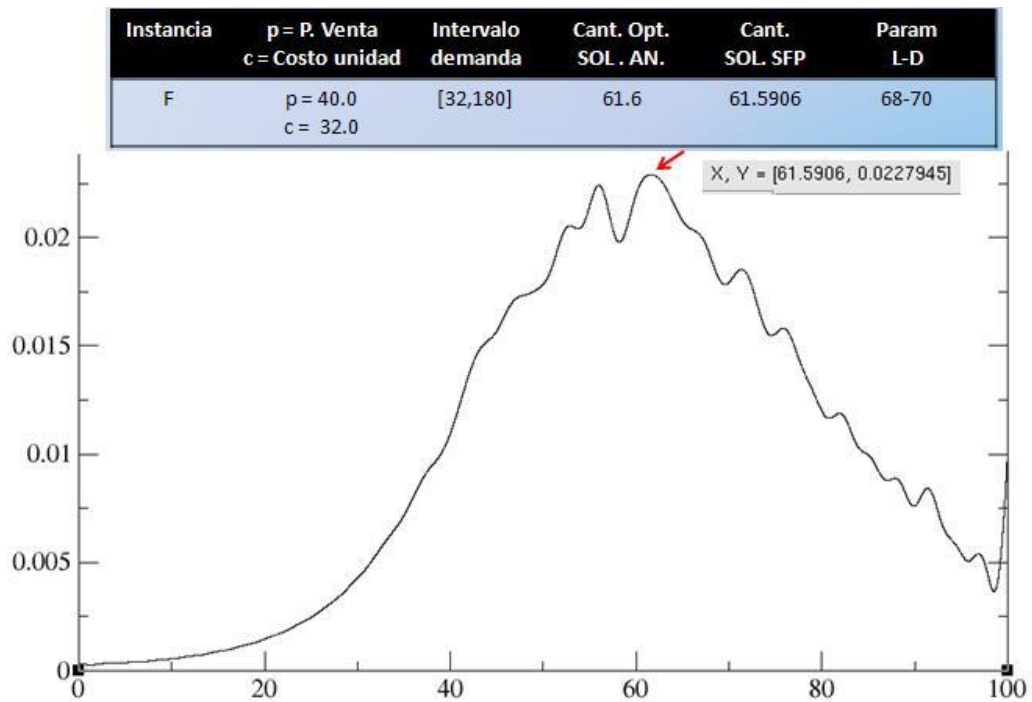


Figura 6.2.2f. Instancia F. Densidad generada por algoritmo SFP en problema del vendedor de periódicos.

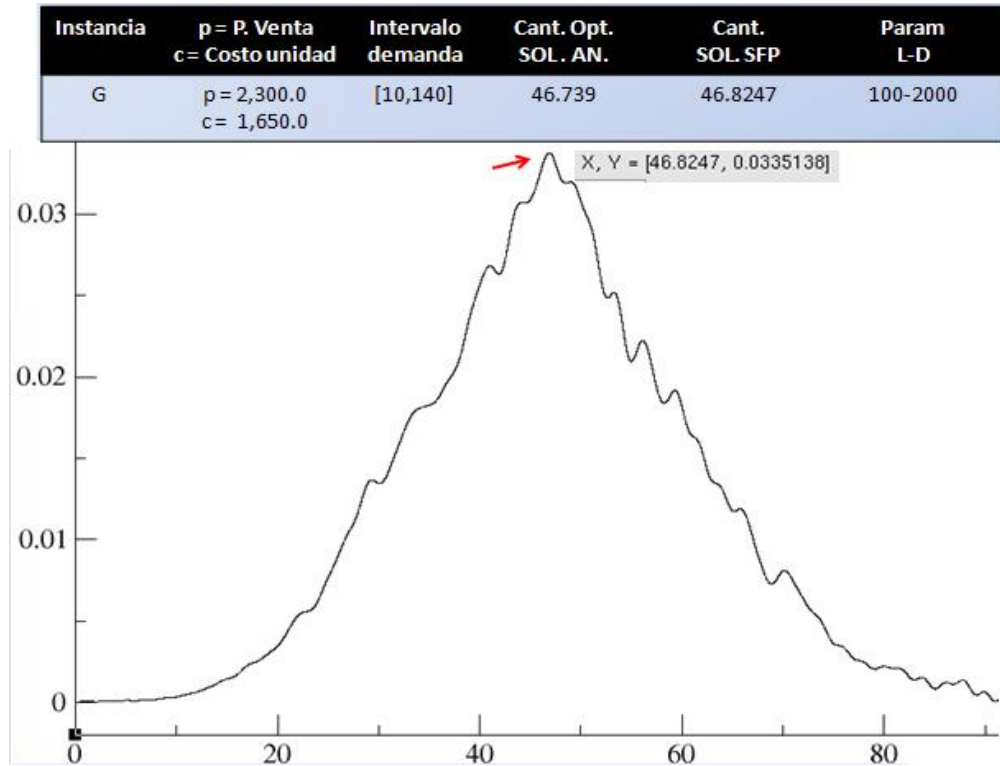


Figura 6.2.2g. Instancia G. Densidad generada por algoritmo SFP en problema del vendedor de periódicos.

El algoritmo de muestreo de Fokker-Planck o SFP (Sampling Fokker-Planck, por las siglas en inglés) no había sido probado en problemas de optimización estocástica. La solución obtenida con el método SFP tiene buena calidad. Si calculamos paralelamente la solución analítica como se demostró, funciona eficientemente en el problema con demanda unimodal del problema clásico del vendedor de periódicos obteniendo la solución muy cercana al óptimo global.

Por lo que se puede concretar que es una herramienta eficiente para resolver este problema.

6.2.3 Resultados del algoritmo SFP en el problema del vendedor de periódicos con demanda bimodal

Cuando hablamos del caso en que la demanda es bimodal en el problema, nos referimos a un comportamiento que tiene dos formas. Para el problema del vendedor, diariamente la demanda bimodal de periódicos es diferente durante la mañana que durante la tarde.

Este modelo de optimización estocástica puede hacer alusión a muchas aplicaciones en la vida cotidiana, como los casos donde la demanda de algún bien o producto tiene comportamiento diferente durante las temporadas del año, por ejemplo la demanda alta de los helados en tiempos de verano y demanda baja en invierno; o bien, el tiempo de uso de la red telefónica en horas pico que es alto y más bajo en las horas no-pico durante el día. En problemas donde no se conoce la distribución de la demandas, que son altas y más probables que otras en algún tiempo, es decir, demandas estacionales en periodos de tiempo donde es más viable vender que en otros periodos con poca demanda, donde no se recomienda tener inventarios altos por las bajas ventas que traen demandas bajas.

El algoritmo de muestreo de Fokker-Planck o SFP, no había sido probado en problemas de optimización estocástica y como vimos en la sección anterior, funciona eficientemente en la solución del problema con demanda unimodal.

Este problema con demanda bimodal del vendedor de periódicos es muy interesante por sus aplicaciones, para el cual no se tiene la solución analítica, razones por las que es de gran motivación encontrar un método eficiente que sea herramienta para resolver este tipo de problemas estocásticos. El valor esperado del costo, es decir el promedio, es la mejor decisión a largo plazo para el caso unimodal donde se observa que la forma de la densidad de probabilidad (distribución) es como la de una Gaussiana. Sin embargo, esto no sucede para el caso bimodal. El promedio o valor esperado para este caso no necesariamente es la mejor decisión a largo plazo debido a que el valor esperado podría caer en medio de una zona que significaría un valor muy poco probable de ocurrencia de la variable de decisión.

Conocer toda la distribución que generan los valores en el objetivo dada este tipo de demanda incierta de dos modos es difícil, y el método SFP estima esa distribución, por lo que en esta sección se probará aplicar el método SFP en el caso bimodal del vendedor de periódicos, se ha supuesto que la demanda se comporta de dos formas en el intervalo de $[70,90]$, ya sea entre $[70,80]$ ó $[80,90]$.

La función objetivo es de minimizar los costos de comprar periódicos diariamente dado que no se sabe ciertamente cómo será la demanda dentro de estos valores.

Cuando se minimiza el valor esperado, como en el problema con demanda unimodal, esta solución será una decisión óptima a largo plazo. Para el caso bimodal, esta solución del mínimo valor esperado nos da la mejor pero a corto plazo, porque tiene mayor probabilidad de que ocurra una demanda que alguna otra. Lo que queremos conocer en este problema es tomar la decisión óptima global a largo plazo de la cantidad de periódicos que el vendedor debe pedir si se tiene este comportamiento en la demanda donde esta solución sea la mejor tanto para corto como para largo plazo sin tener riesgos de que, si tomamos la decisión que nos minimice los costos con buen beneficio no se tengan pérdidas por no satisfacer demandas altas que pudiera haber en un futuro o pérdidas por tener demasiado inventario de periódicos si habrá poca demanda.

Aplicaremos el método SFP en diferentes instancias donde la demanda se comporta más alta en un periodo, que en otro.

La tabla 6.3 muestra las instancias con precios de venta y costo por unidad de periódico, la demanda se presenta entre las [70,90] unidades sin conocerse el valor que tendrá durante el día. Se muestra la solución obtenida con el algoritmo SFP y los parámetros utilizados L para adaptación y D para diversificación.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo demanda	Cant. Opt. SOL. AN. (unidades)	Cant. SOL. SFP (unid)	Param L-D	Gráfico
A	p = 11.0 c = 11.0	[70,90]		68.0696	10-40	a
B	p = 120.0 c = 10.0	[70,90]		85.8715 83.2305	50-30 50-40	b1 b2
C	p = 120.0 c = 10.0	[70,90]		77.4334 76.8571	20-80 20-100	c1 c2

Tabla 6.2.3 Instancias para problema del vendedor de periódicos con demanda bimodal en el intervalo indicado. Se muestra la evolución de la densidad al cambiar los parámetros. En números grises, se indica la mejor solución para la instancia con la densidad correcta en la curva.

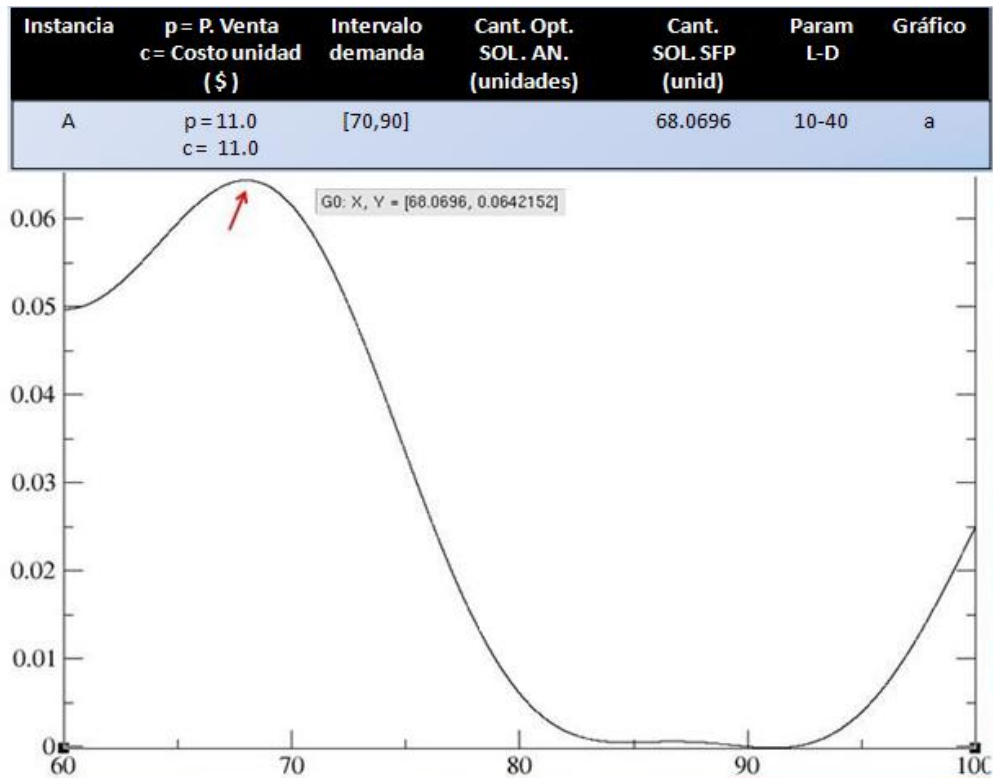


Figura 6.2.3A Instancia A. Precio de venta igual al costo. Solución de SFP con la densidad estacionaria correcta que indica que no es conveniente comprar periódicos.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo demanda	Cant. Opt. SOL. AN. (unidades)	Cant. SOL. SFP (unidad)	Param L-D	Gráfico
B	p = 120.0 c = 10.0	[70,90]		85.8715	50-30	b1
				83.2305	50-40	b2

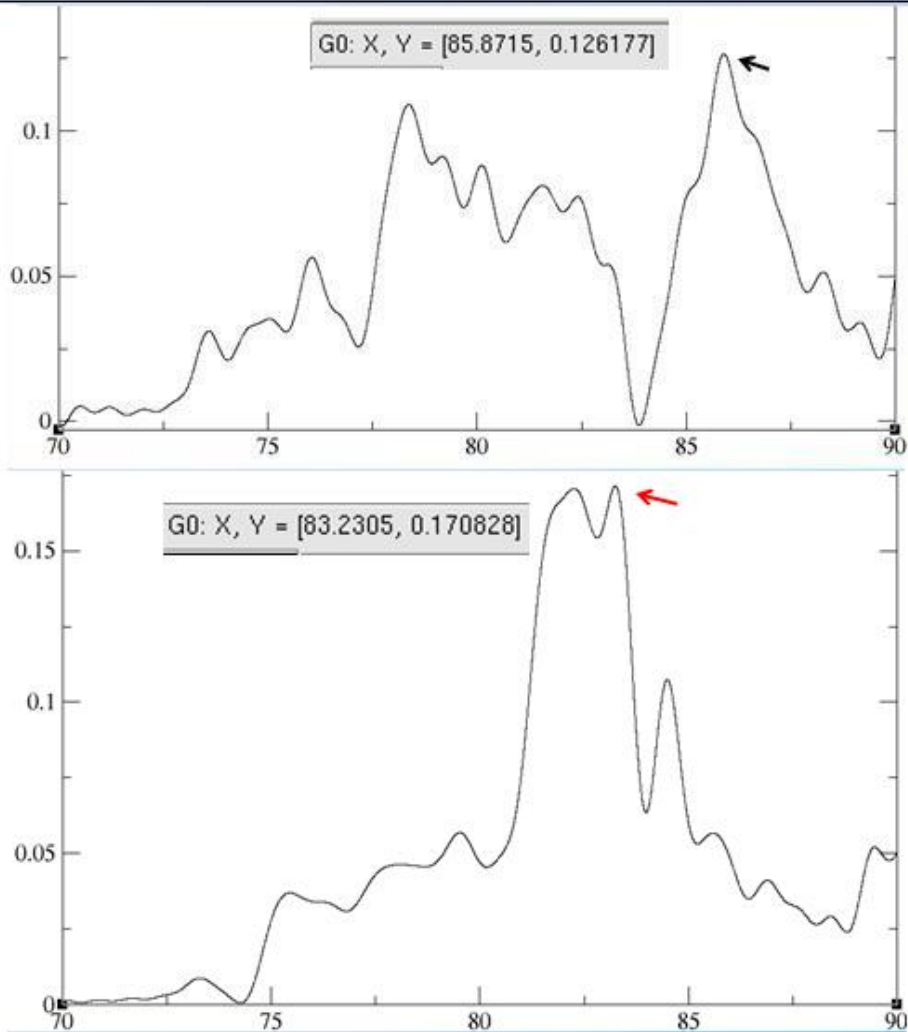


Figura 6.2.3B. Se muestra el progreso en la convergencia de la densidad de probabilidad. En la gráfica b1, se observa una densidad aún no correcta con los parámetros L y D. El gráfico b2 muestra la densidad estacionaria y parámetros correctos, se indica la solución $x=83.2305$.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo demanda	Cant. Opt. SOL . AN. (unidades)	Cant. SOL. SFP (unid)	Param L-D	Gráfico
C	p = 120.0 c = 10.0	[70,90]	77.4334	77.4334	20-80	c1
			76.8571	76.8571	20-100	c2

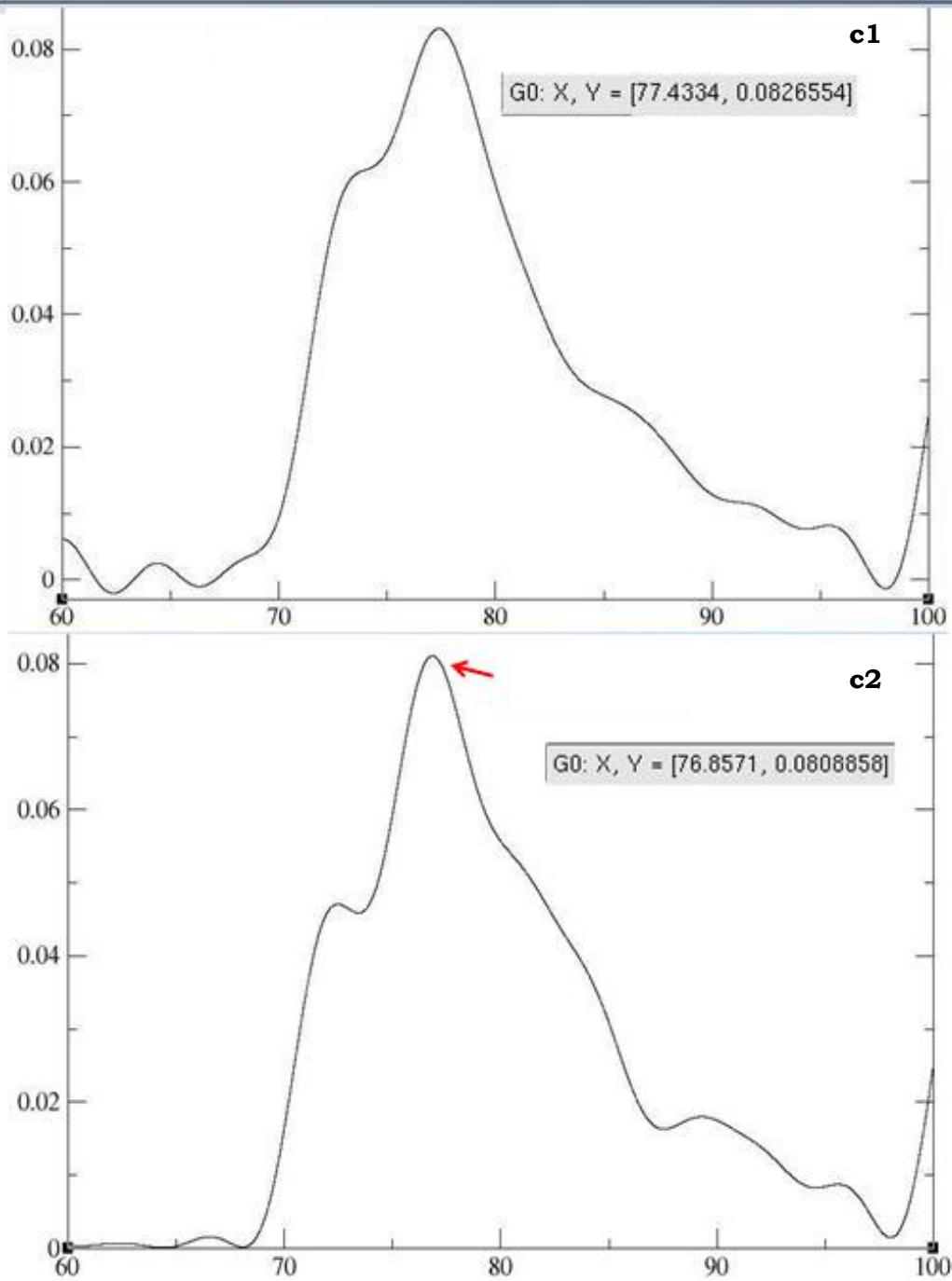


Figura 6.2.3C. En el gráfico c1 se observa una densidad no correcta (negativa) con los parámetros no adecuados, solución con letras negras. El gráfico c2 muestra la densidad estacionaria y parámetros correctos, se indica la solución $x=76.8571$.

En este problema del vendedor de periódicos con demanda bimodal, la demanda ocurre en los valores entre 70 y 90 pero se comporta de dos maneras en este intervalo: tiene más alta probabilidad de presentar ciertos valores dentro del intervalo, mayor probabilidad de que ocurran demandas entre algunos valores en el intervalo, y a su vez existe otra probabilidad más baja de que ésta se presente con los otros valores en ese intervalo.

Como en la instancia B, observamos que en el intervalo [70,90] se tiene una muy baja probabilidad de que la demanda ocurra entre [70,80], en cambio es muy alta para que ocurra entre [80,90].

En la instancia C, se observa que en el intervalo [70,90] existe una mayor probabilidad de que los valores de demanda se encuentren entre [70,80] y una probabilidad muy baja de que esté entre [80,90].

Por lo que tenemos lidiar con tomar la mejor decisión pues no se sabe con certeza entre qué valores se presentará la demanda con mayor o menor probabilidad. Se quiere entonces, encontrar la mejor decisión a largo plazo de manera que exista el mejor compromiso entre el comportamiento de alta y baja probabilidad de ocurrencia de la demanda para no caer en desabasto de periódicos o un exceso en inventario, la mejor decisión ahora y para un futuro. Sólo que esta solución, no minimiza el costo a largo plazo.

Si queremos saber qué será lo mejor a corto y largo plazo, por medio de *sfp* se genera la distribución de probabilidad completa y se pueden muestrear soluciones a partir de esta distribución, en la cual, el pico de mayor probabilidad para el caso bimodal contiene la mejor decisión a corto plazo, es decir, es la más probable a corto plazo. Sólo que esta decisión no minimiza el costo a largo plazo.

El valor esperado del costo, es decir el promedio, es la mejor decisión a largo plazo para el caso unimodal donde se observa que la forma de la densidad de probabilidad (distribución) es como la de una Gaussiana. Sin embargo, esto no sucede para el caso bimodal. El promedio o valor esperado para este caso no necesariamente es la mejor decisión a largo plazo debido a que la forma de la distribución (o densidad) tiene “dos montañas” y el valor esperado podría caer

en medio de la zona donde se encuentra el llano entre esas montañas, lo que significaría un valor muy poco probable de ocurrencia de la variable de decisión.

La variable aleatoria (la demanda de periódicos) se presenta en términos de ocurrencia a corto plazo. El valor esperado de la segunda etapa del problema, representa el promedio sobre los posibles valores en que la variable aleatoria puede ocurrir. De aquí que, obtener el valor esperado implica tomar en cuenta las realizaciones de la variable aleatoria que podrían no ocurrir junto con las realizaciones que sí podrían. Por ejemplo, si se tiene una demanda incierta que ocurre diariamente o semanalmente y, la función objetivo en periodos anuales, entonces con el valor esperado en el objetivo del modelo estocástico se tendría un beneficio anual [1]. Entonces, el promedio obtenido de esta forma en la segunda etapa de la función objetivo, no refleja realmente los resultados en términos de los periodos de ocurrencia de la variable aleatoria.

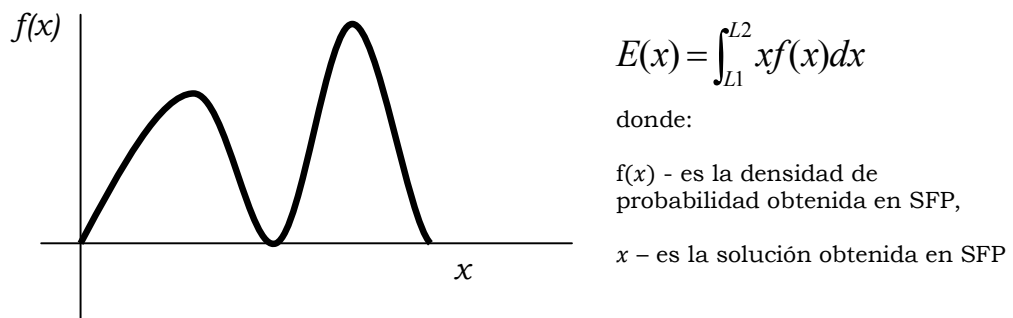


Figura 6.2.4C. El valor de la función objetivo en la segunda etapa, dada una decisión en la primer etapa, es el valor esperado.

Por lo que, el método SFP podríamos adaptarlo para minimizar el valor esperado o promedio y obtener una solución mejor para largo plazo pues en el método funciona de manera que, dada la ocurrencia de una demanda (ya sea diaria, semanal, mensual, anual, etc.), se calcula la función objetivo en el gradiente y SFP genera la distribución de probabilidad que proviene de este proceso, para construir la densidad de probabilidad que funciona para una solución a corto plazo (diaria, semanal, mensual, anual, etc.). La adaptación

del método para largo plazo sería: en la parte donde ocurre la demanda, en lugar de hacer la evaluación de la función objetivo con esta demanda a corto plazo para que SFP calcule el gradiente, se introduce en su lugar, el promedio de unas N demandas, y este promedio de demanda se calcularía en la función objetivo dentro del gradiente para que el método SFP genere la distribución (esta densidad) que proviene del proceso de calcular el gradiente en el objetivo que ocupa la “demanda promedio”, que significaría una solución para el largo plazo, que minimice el costo a largo plazo.

Por el momento, este trabajo de investigación no se enfoca en desarrollar el caso para largo plazo, pero como se explica anteriormente, el método SFP podría modificarse para obtener soluciones a largo plazo también.

6.2.4 Problema del vendedor de periódicos con demanda en dos intervalos

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. SOL SFP (unid)	Param L-D	Alfa	Gráfico
A	p = 11.0 c = 4.0	[70,90] U [120,150]	125.566	35-45	0.1	a
B			77.3373	50-100	0.9	b1
			72.0072	60-100		b2
			70.7587	30-100		b3
C			78.1597 U 125.218	50-200	0.5	c1
			82.3613 U 123.370	20-200		c2
			73.4538 U 127.067	13-150		c3

Tabla 6.2.4 Instancias para problema del vendedor de periódicos con demanda en dos intervalos. Se muestra la evolución de la densidad con los parámetros. Alfa indica probabilidad de ocurrencia de la demanda. Los valores en negritas, indican la mejor solución para la instancia con la densidad correcta en la curva.

Las instancias y resultados se muestran a continuación. Cabe recordar que la densidad de probabilidad completa converge en su valor correcto y es estacionaria, es decir, es positiva y converge a 1.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. SOL. SFP (unid)	Param L-D	Alfa	Gráfico
A	p = 11.0 c = 4.0	[70,90] U [120,150]	125.566	35-45	0.1	a

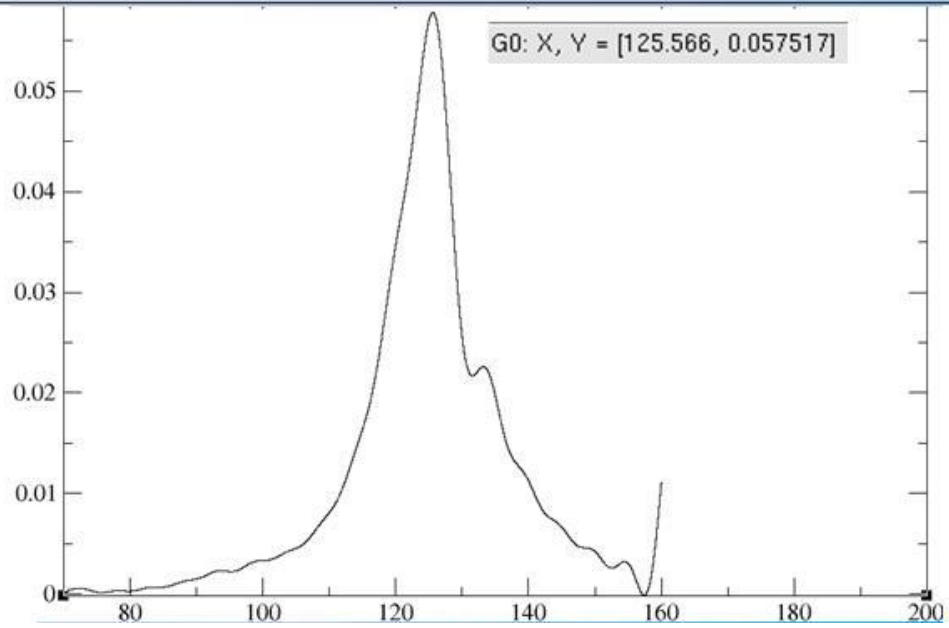


Figura 6.2.4-1 Instancia A. La demanda podría presentarse entre [70,90] ó [120,150]. La mejor solución por el comportamiento de la demanda en este caso, es comprar 125 periódicos.

En la instancia A, se observa que la zona de mayor probabilidad está en la segunda parte del intervalo. En esa zona se encuentra el mínimo global.

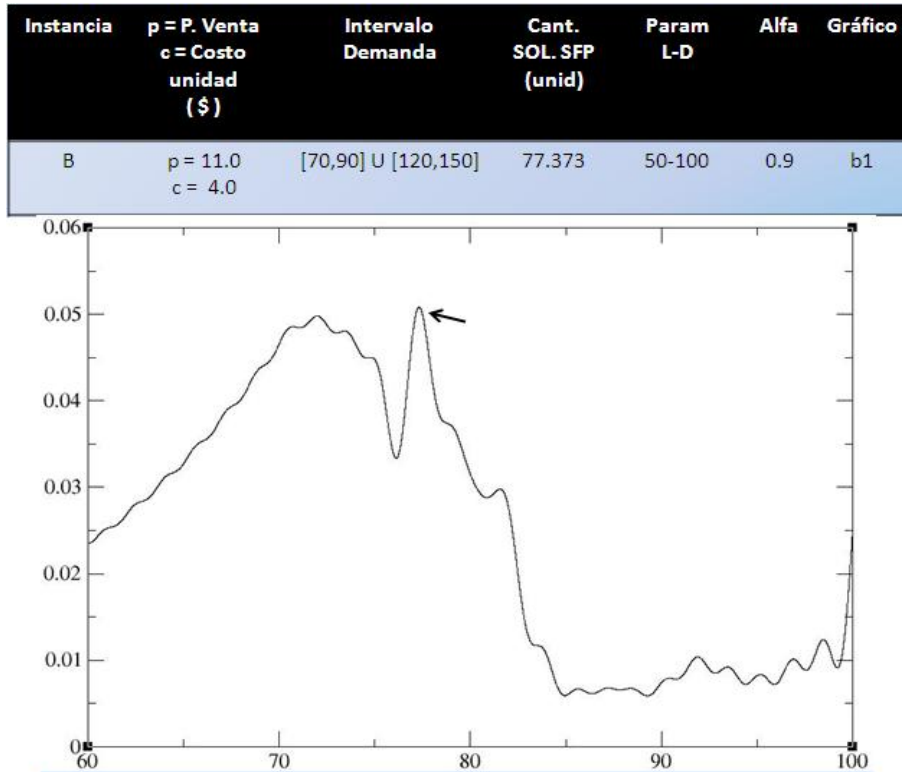


Figura 6.2.4-2 Instancia B. Gráfico b1. Se observa que la solución se presenta dentro del intervalo [70,90]. Comienza a formarse la zona de densidad donde se encontrará la solución.

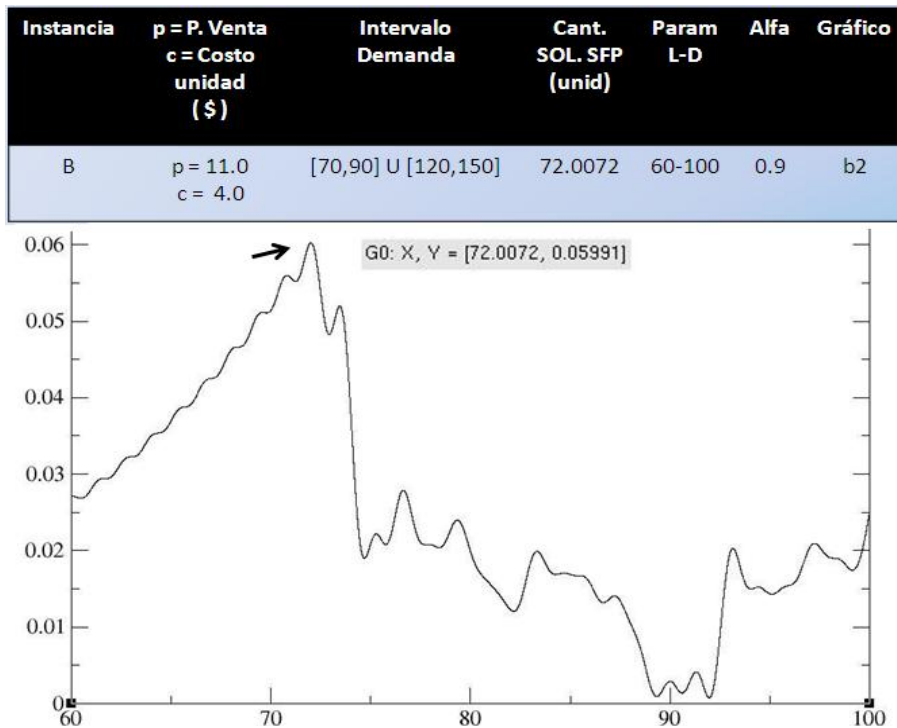


Figura 6.2.4-3 Instancia B. Gráfico b2 que muestra que para un alfa de 0.9 la solución se ubica entre [70,90] con los parámetros indicados.

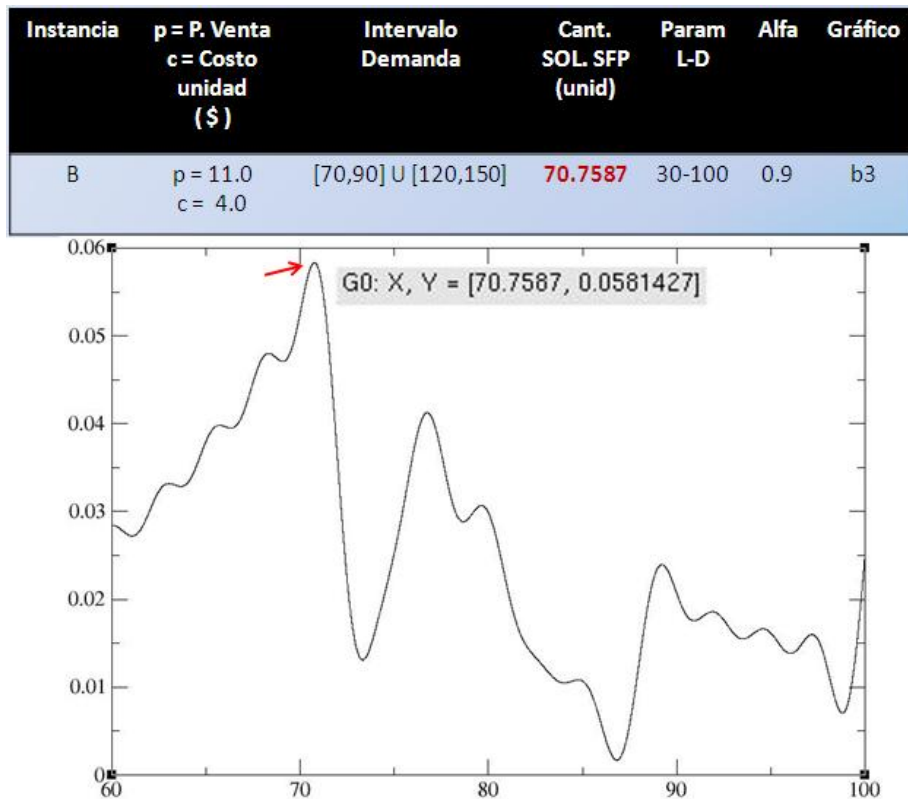


Figura 6.2.4-4 Gráfico b3 que indica la densidad de más alta probabilidad con un alfa de 0.9, se observa la densidad estacionaria y correcta. La solución mejor es $x=70.7587$.

En el problema del vendedor puede darse el caso en que la demanda de periódicos incierta ocurra de manera que adopte con muy alta probabilidad, valores entre cierto rango y además, con una probabilidad más baja que la anterior otros valores en algún otro rango muy diferente al primero. Pero no se sabe con certeza en qué intervalo se presentará. Tal caso se describe en esta sección.

Para el problema se realizaron instancias donde la demanda ocurre y toma valores entre el intervalo [70,90] y [120,150]. El método SFP se implementa en cada caso.

En la instancia A se analiza el caso donde se sabe que la demanda podría ocurrir entre los valores de $[70,90]$ y $[120,150]$, pero con mayor probabilidad se presentó dentro del rango de valores de $[120,150]$, en el intervalo $[70,90]$ se presentó con menor probabilidad.

Por lo que indica que para la solución del problema el vendedor, se debería invertir en comprar periódicos para su venta dentro del rango con mayor probabilidad en $[120,150]$ y no arriesgarse optando por comprar periódicos dentro del rango de $[70,90]$ dado que, aunque si existe una demanda que puede ocurrir entre estos valores, se presenta con mucha menos probabilidad y entonces tendría un riesgo de desabasto.

En la instancia B se tiene el caso donde, si bien, la demanda incierta podría realizarse entre los intervalos $[70,90]$ y $[120,150]$ con alguna probabilidad, ocurre la demanda con mayor probabilidad dentro de $[70,90]$ y ocurre también pero con menor probabilidad, dentro del rango $[120,150]$.

Dado este resultado, la solución que proporciona el método SFP es que debería comprar periódicos entre el rango de $[70,90]$ pues si el vendedor opta por comprar cantidades dentro del intervalo de $[120,150]$ aunque si ocurre una demanda en este rango, ocurre con mucha menos probabilidad y caería en un riesgo de incurrir en exceso de inventario, que se traduce en costos.

Por otro lado, en la instancia C se analiza el caso en que la demanda incierta podría presentarse con igual probabilidad dentro de los dos intervalos $[70,90]$ y $[120,150]$.

La solución resulta en la mejor cantidad a comprar dentro de ambos intervalos, lo que indica que, dado que la demanda ocurre con igual probabilidad, no se incurre en riesgos por optar alguna decisión sobre la otra.

Por lo que es decisión del vendedor, ya sea por su capacidad financiera u otros condicionantes, comprar cantidades de periódicos entre $[70,90]$ y vender solamente en los rangos de esa cantidad o comprar periódicos en la cantidad que se indica dentro del intervalo $[120,150]$ y obtener mayores beneficios por las altas ventas.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. SOL. SFP (unidad)	Param L-D	Alfa	Gráfico
C	p = 11.0 c = 4.0	[70,90] U [120,150]	78.1597 U 125.218 82.3613 U 123.370 73.4538 U 127.067	50-200 20-200 13-150	0.5	c1 c2 c3

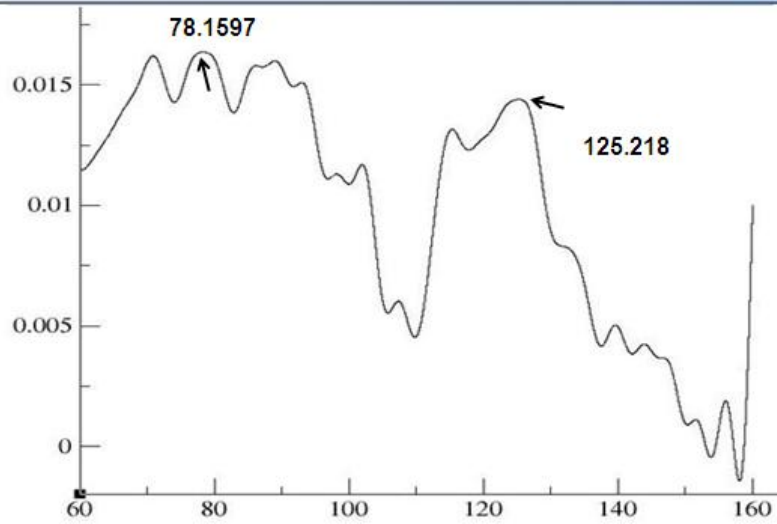


Figura 6.2.4-5 Instancia C. Gráfico c1 con un alfa 0.5. La demanda se comporta de dos formas: entre cualquier valor dentro de [70,90] y [120,150]. La densidad aún no es correcta.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. SOL. SFP (unidad)	Param L-D	Alfa	Gráfico
C	p = 11.0 c = 4.0	[70,90] U [120,150]	78.1597 U 125.218 82.3613 U 123.370 73.4538 U 127.067	50-200 20-200 13-150	0.5	c1 c2 c3

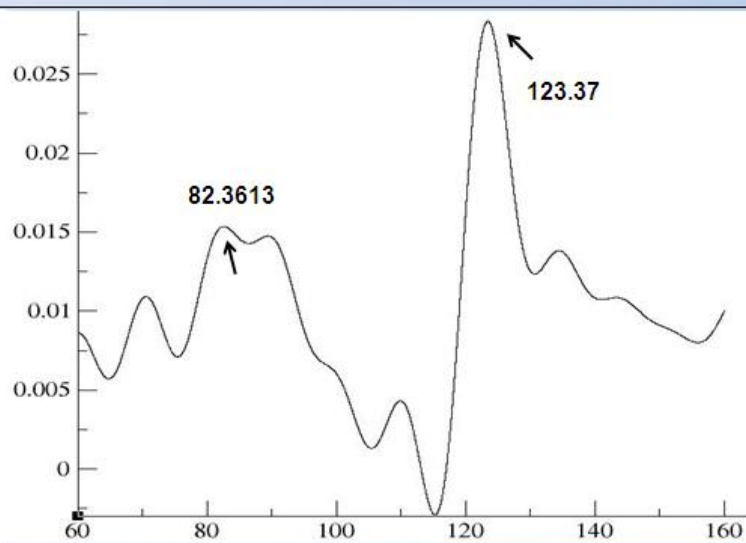


Figura 6.2.4-6 Instancia C. Gráfico c2. La densidad (negativa) no es correcta con los parámetros. Con alfa 0.5 la solución se observa dentro de ambos intervalos.

Instancia	p = P. Venta c = Costo unidad (\$)	Intervalo Demanda	Cant. SOL. SFP (unid)	Param L-D	Alfa	Gráfico
C	p = 11.0 c = 4.0	[70,90] U [120,150]	78.1597 U 125.218 82.3613 U 123.370 73.4538 U 127.067	50-200 20-200 13-150	0.5	c1 c2 c3

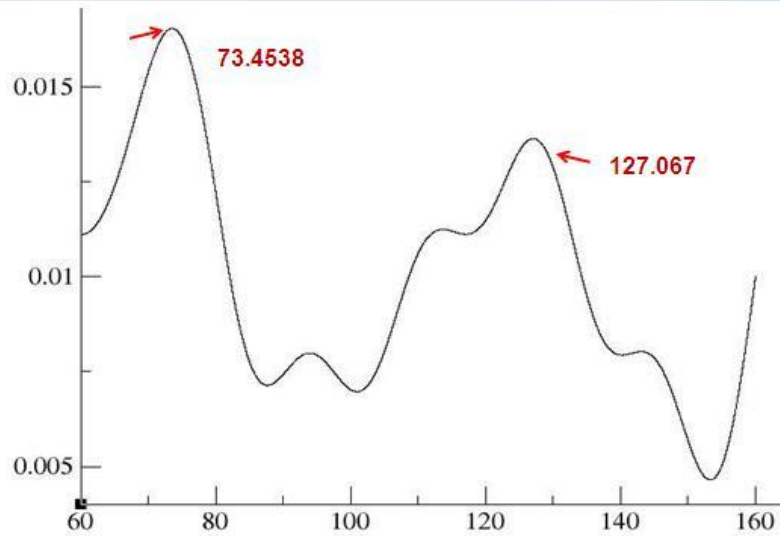


Figura 6.2.4-7 Instancia C. Gráfico c3 con parámetros y densidad estacionaria correctos. La mejor solución es $x=73.4538$ y $x=127.067$. La cantidad de periódicos a pedir son ambas cantidades.

En problemas estocásticos como éste con demandas inciertas de comportamientos como los descritos en esta y anteriores secciones, tomar una decisión en el presente implica mucho riesgo pues ¿cómo saber si será la decisión adecuada también para un futuro a largo plazo si no se conoce con certeza la demanda?

El valor esperado, es decir el promedio, entre estas decisiones es el mejor compromiso para las dos condiciones de la demanda. Pero no necesariamente es la mejor decisión para el caso en dos intervalos de demanda.

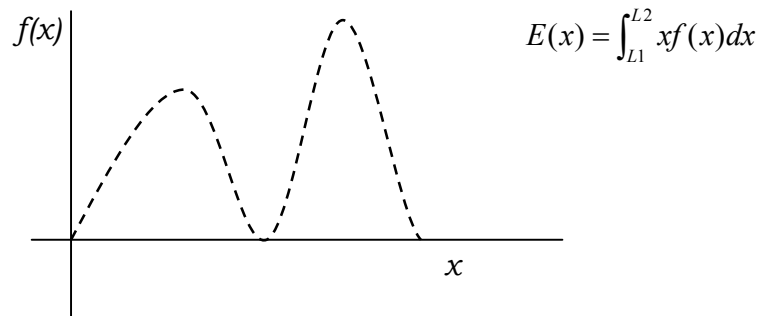


Figura 6.2.4-8 El valor esperado o promedio, no es la mejor decisión a corto plazo para el caso donde las demandas son bimodales.

El valor esperado del costo, es decir el promedio, es la mejor decisión a largo plazo para el caso unimodal donde se observa que la forma de la densidad de probabilidad (distribución) es como la de una Gaussiana. Sin embargo, esto no sucede para el caso bimodal. El promedio o valor esperado para este caso no necesariamente es la mejor decisión a largo plazo debido a que la forma de la distribución (o densidad) tiene “dos montañas” y el valor esperado podría caer en medio de la zona donde se encuentra el llano entre esas montañas, lo que significaría un valor muy poco probable de ocurrencia de la variable de decisión, y no coincide para nada el valor más probable con el del valor esperado.

Con el método SFP podría hacerse una adaptación para minimizar el valor esperado o promedio y obtener una solución mejor para largo plazo pues en el método funciona de manera que, dada la ocurrencia de una demanda (ya sea diaria, semanal, mensual, anual, etc.), se calcula la función objetivo en el gradiente y SFP genera la distribución que proviene de este proceso para construir la densidad de probabilidad que funciona para una solución a corto plazo (diaria, semanal, mensual, anual, etc.). La adaptación del método para largo plazo sería: en la parte donde ocurre la demanda, en lugar de hacer la evaluación de la función objetivo con esta demanda a corto plazo para que SFP calcule el gradiente, se introduce en su lugar, el promedio de unas N demandas, y este promedio de demanda se calcularía en la función objetivo dentro del gradiente para que el método SFP genere la distribución (esta

densidad) que proviene del proceso de calcular el gradiente en el objetivo que ocupa la “demanda promedio”, que significaría una solución para el largo plazo, que minimice el costo a largo plazo.

Por el momento, este trabajo de investigación no se enfoca en desarrollar el caso para largo plazo, pero como se explica anteriormente, con el método SFP podría modificarse para obtener soluciones a corto y largo plazo.

Capítulo 7

APLICACIONES

Por los resultados que se observaron en las secciones anteriores, se demostró que el método de muestreo de Fokker-Planck SFP es una buena herramienta en la solución del problema de tipo estocástico como lo es el caso del problema del vendedor de periódicos. Esta nueva técnica nos muestra la estructura del objetivo del problema dado que se tiene la variable incierta de demanda, estimando la distribución o densidad estacionaria que refleja la zona que encierra el óptimo global con alta probabilidad. Se ha demostrado en los casos que ya vimos.

Por lo que en esta sección se muestran los resultados obtenidos de la implementación del método SFP en un problema cotidiano donde los datos reales de demanda mensual son kg de productos cárnicos. Los datos son las demandas proporcionadas de una fuente productora de carnes embutidas y de las cuales no se conoce su distribución de probabilidad pero es un vector de datos históricos que ya se realizaron durante algunos meses, ya se conocen.

En este caso, se necesita saber cuál será la cantidad óptima de inventario mensual a producir dado que se presenta una demanda incierta cada mes, los costos de producción y el precio de venta por kg son conocidos. Se trata de un problema de minimización de costos de producir kg de cárnicos para mantener un inventario mensualmente.

En seguida, se muestran los resultados del método SFP en implementaciones para 5 productos diferentes los cuales se denominan por nombres numéricos 0397, 0419, 0429, 0432, 0438, respectivamente. En las figuras se muestran los gráficos con las estructuras de las densidades estimadas por el método donde se refleja la zona que encierra el óptimo global con mayor probabilidad.

Se muestran en cada producto, los precios de venta y costo de producción por kg, los parámetros del método SFP y la solución en cada caso. Con propósito ilustrativo, se muestran también el avance hacia la convergencia de la densidad estacionaria durante la estimación del método SFP para algunos de los productos, con los parámetros que se usan en él.

En el apéndice B se muestran los datos históricos de los productos cárnicos.

7.1 APLICACIÓN DEL MÉTODO SFP CON DEMANDA DE PRODUCTO 0397

Producto 0397	p = Precio venta c= costo producir (por kg)	Solución SFP (kg)	Parámetros L - D	Gráfico
	p = 11.0 c = 8.0			
3525.09			80 - 2000	b
3496.28			50 - 2000	c
3573.11			30 - 10000	d
3563.51			30 - 11000	e

Tabla. 7-1 Producto 0397. Muestra la solución que genera SFP con los parámetros indicados hasta llegar a la densidad correcta.

Se muestran a continuación, los resultados en las curvas de SFP con los parámetros y densidades correctos.

IMPORTANTE: Cabe destacar que las demandas utilizadas en el problema de minimización de costos de producción son datos históricos mensuales, no se conoce la forma en específico de la distribución de probabilidad de demanda. Se quiere saber cuál es el inventario óptimo mensual. Se conoce la historia mensual de demanda de los cinco productos.

En la figura 7-1-1 se muestra el progreso de convergencia de la densidad de probabilidad estacionaria para el producto 0397. Con los diferentes parámetros, al cambiar los parámetros se observa que la densidad aún no es la correcta, pues tiene valores negativos.

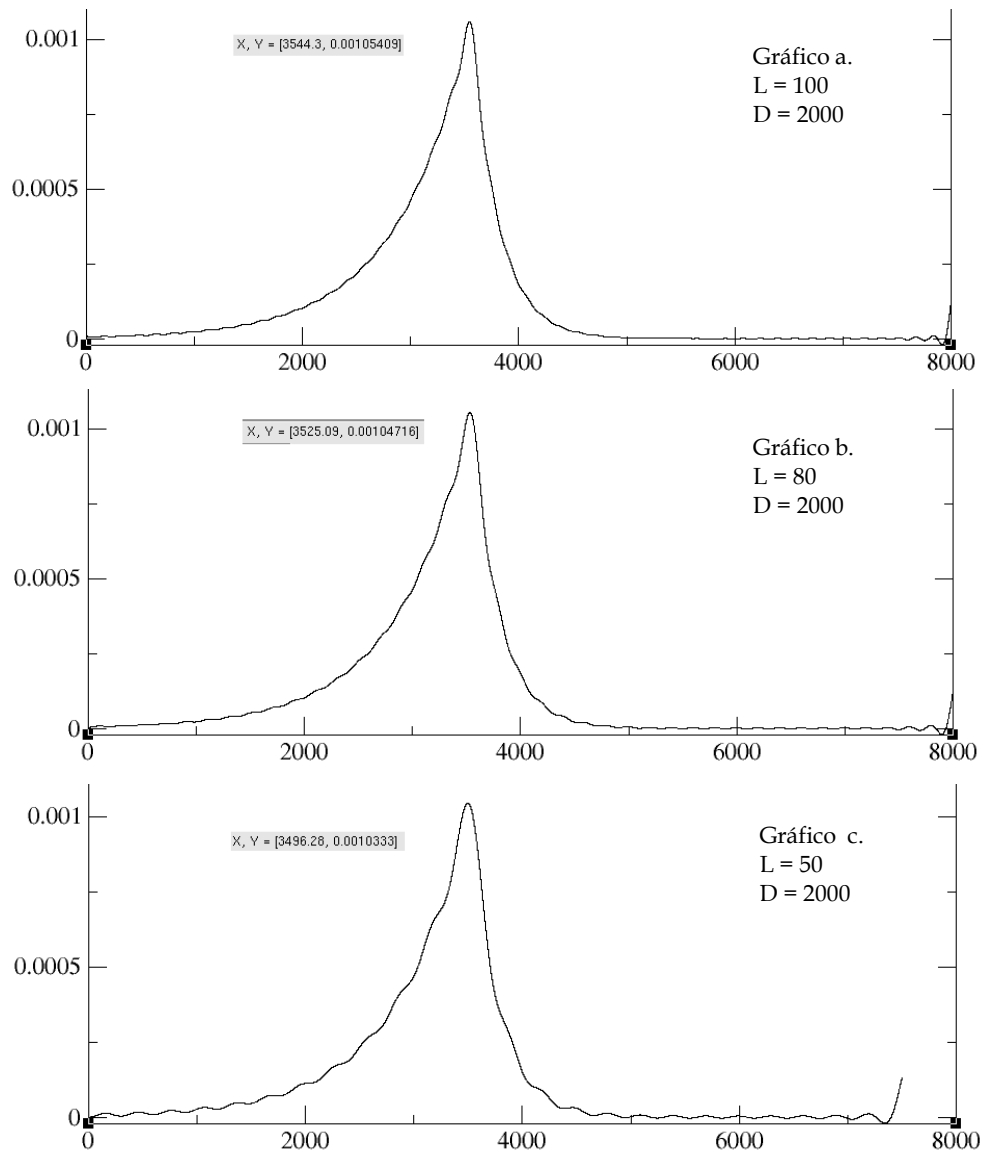


Fig. 7-1-1 Gráfico a, b, c. Respectivamente, las densidades no-correctas para la instancia A. Se buscan los parámetros L y D adecuados. Se disminuye L buscando la densidad estacionaria.

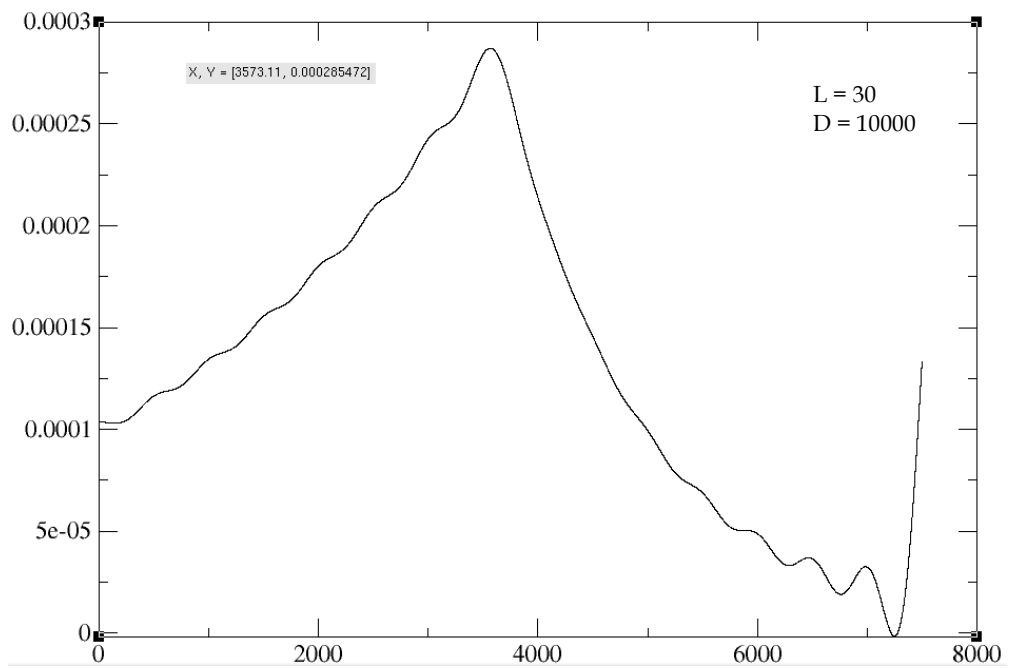


Fig. 7-1-2 Gráfico d. Se observa que el valor de la densidad aumenta, no es correcta aún (negativa). Se incrementa el parámetro D para evaluar en el método.

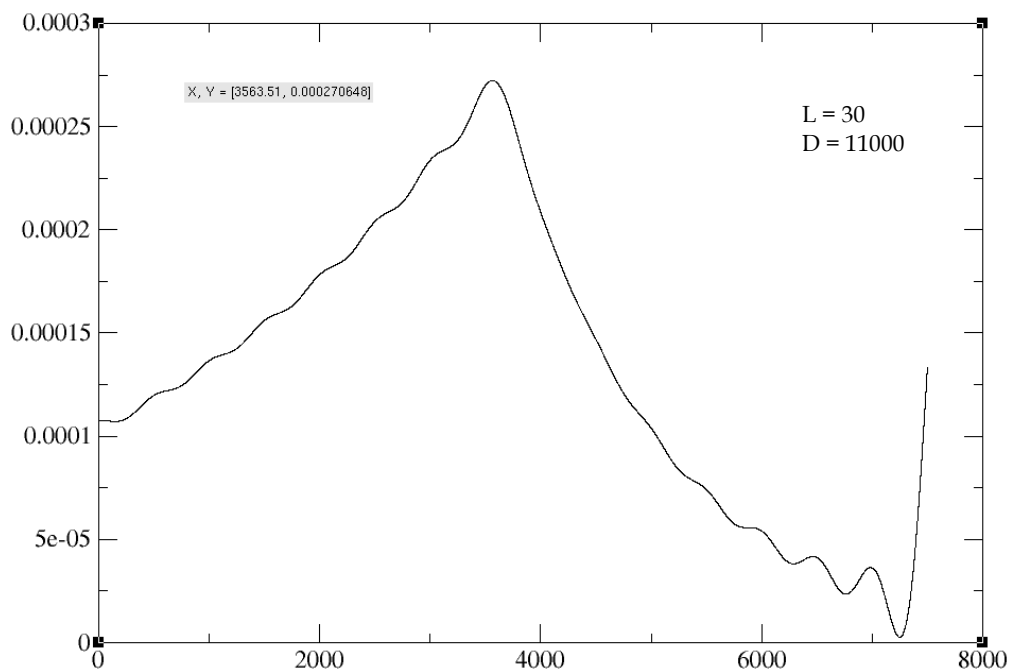


Fig. 7-1-3 Gráfico e. La densidad es correcta (positiva). Los parámetros L y D son adecuados. La solución indica 3563.51 kg para producir en el mes del producto 0397.

En la figura 7-1-3 se muestra que la densidad de probabilidad estacionaria es la correcta, es positiva. Los parámetros correctos son $L=30$, $D=11000$.

7.2 APLICACIÓN DEL MÉTODO SFP CON DEMANDA DE PRODUCTO 0419

Producto 0419	p = Precio venta c= costo producir (\$ / kg)	Solución SFP (kg)	Parámetros L - D	Gráfico
	p = 42.0 c = 19.0			
		19701.1	30 - 50000	b
		19533	100 - 10000	c
		19485	35 - 100000	d

Mejor valor ->

Tabla. 7-2 Producto 0419. Muestra la solución que genera SFP con los parámetros indicados para llegar a la densidad correcta.

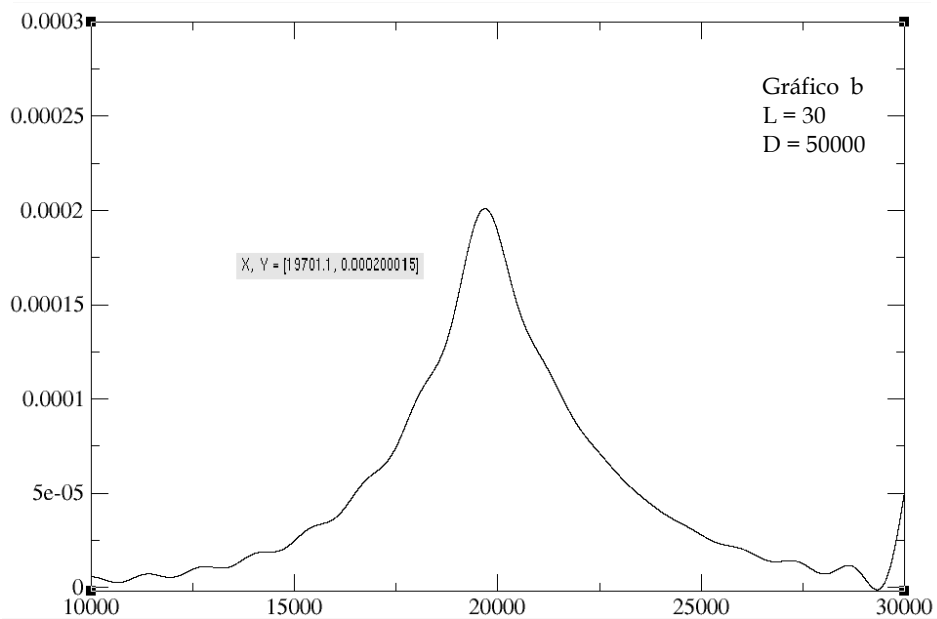
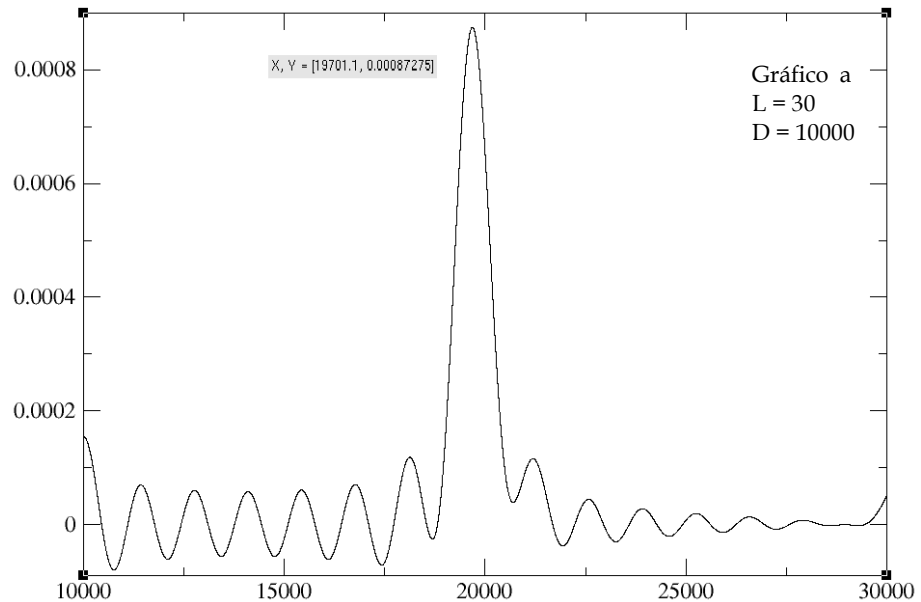


Fig. 7-2-1 Gráfico a, b. La densidad no es correcta. Se incrementa el parámetro D. Se observa una mejor estructura en la densidad.

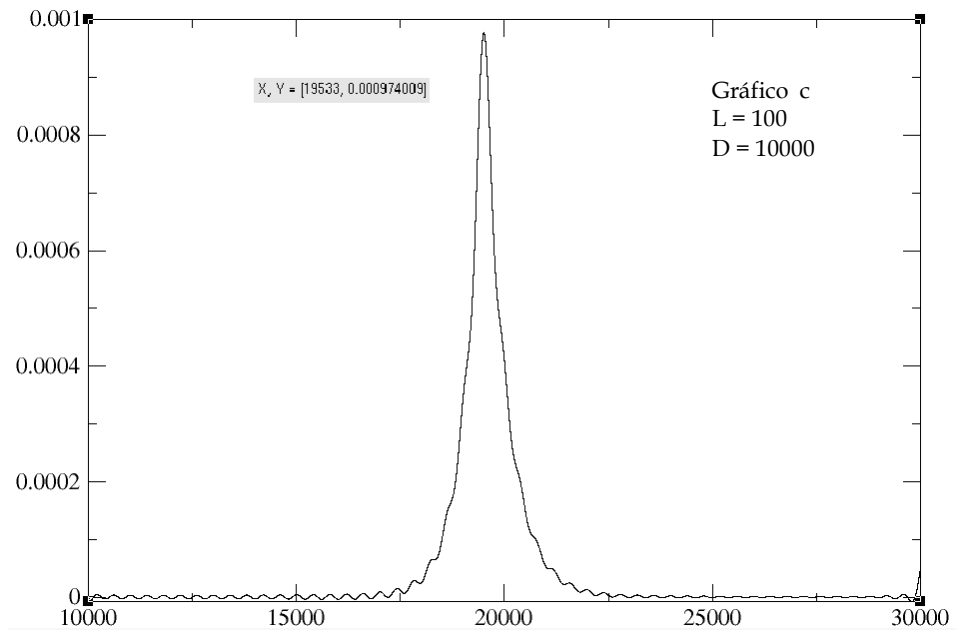


Fig. 7-2-2 Gráfico c. La densidad no es correcta. Se incrementa el parámetro D. Se observa una mejor estructura en la densidad.

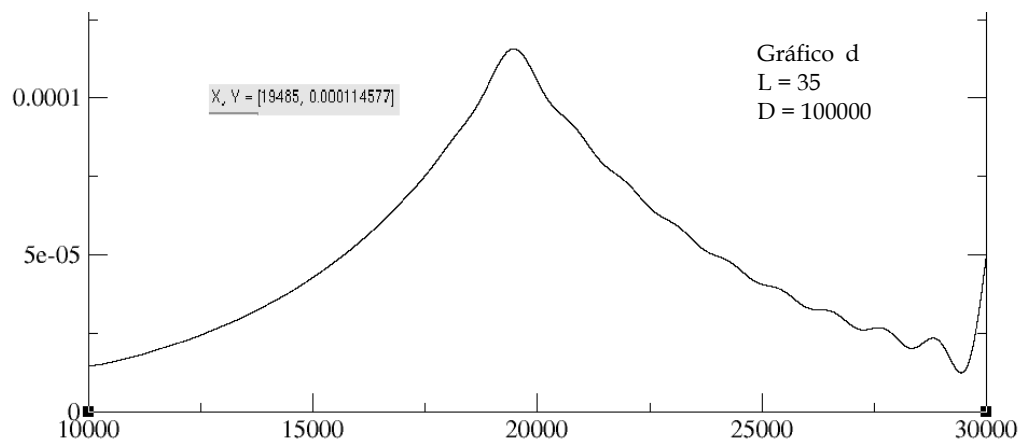


Fig. 7-2-3 Gráfico d. La densidad estacionaria correcta (positiva). Indica la solución mejor de 19485 kg para producir.

7.3 APLICACIÓN DEL MÉTODO SFP CON DEMANDA DE PRODUCTO 0429

Producto	p = Precio venta c= costo producir (\$ / kg)		Solución SFP (kg)	Parámetros L - D	Gráfico
0429	p = 68.0 c = 30.0	Mejor valor ->	1590.94	30 - 30000	a
			1596.94	30 - 35000	b

Tabla. 7-3 Producto 0429. Muestra la solución que genera SFP con los parámetros indicados para llegar a la densidad correcta.

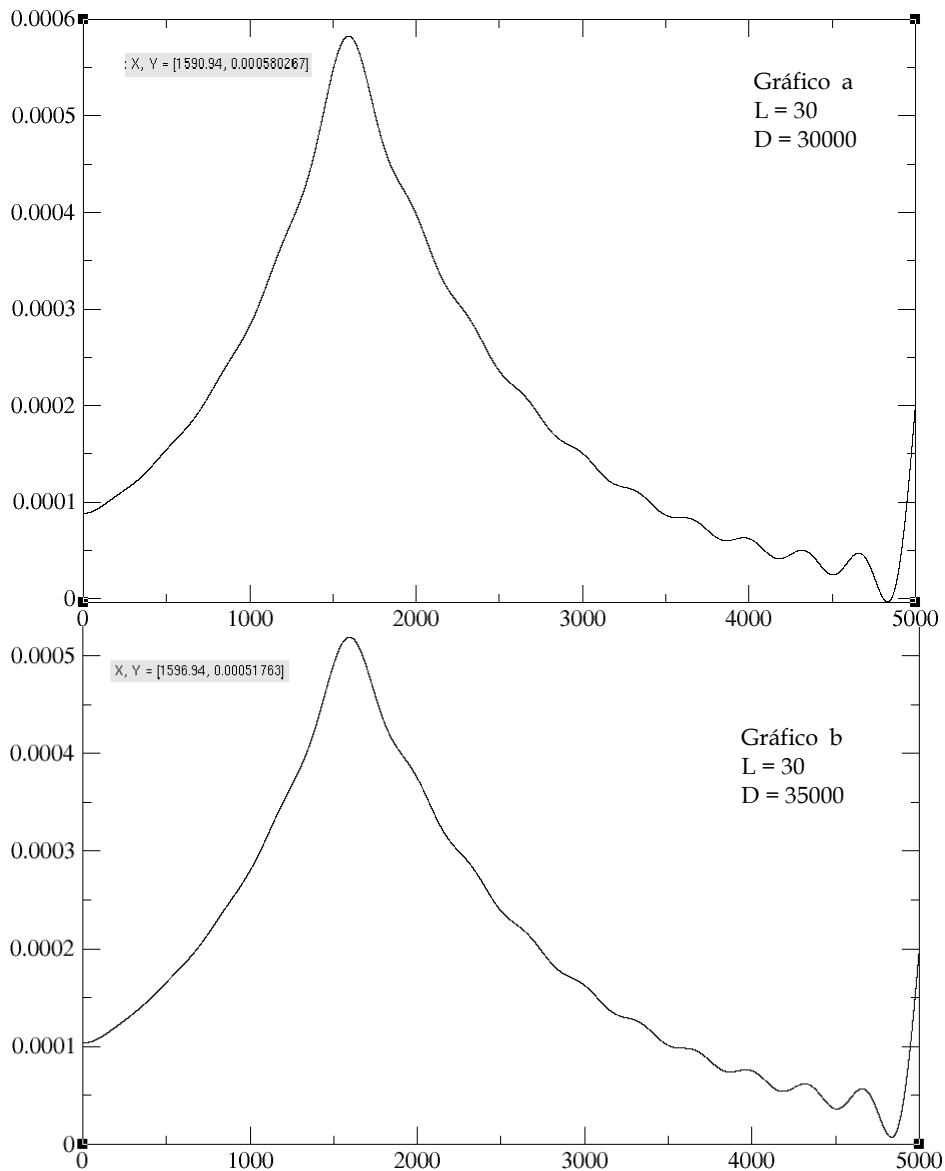


Fig. 7-3-1 Gráfico a, b. La densidad no es correcta en el gráfico a. Se incrementa el parámetro D, se observa en el gráfico b una mejor estructura en la densidad la cual es correcta.

7.4 APLICACIÓN DEL MÉTODO SFP CON DEMANDA DE PRODUCTO 0432

Producto 0432	p = Precio venta c= costo producir (\$ / kg)	Mejor valor ->	Solución SFP (kg)	Parámetros L - D	Gráfico
	p = 56.0 c = 21.0		10799.2	30 - 35000	a
		10817.2	30 - 80000	b	

Tabla. 7-4 Producto 0432. Muestra la solución que genera SFP con los parámetros indicados para llegar a la densidad correcta.

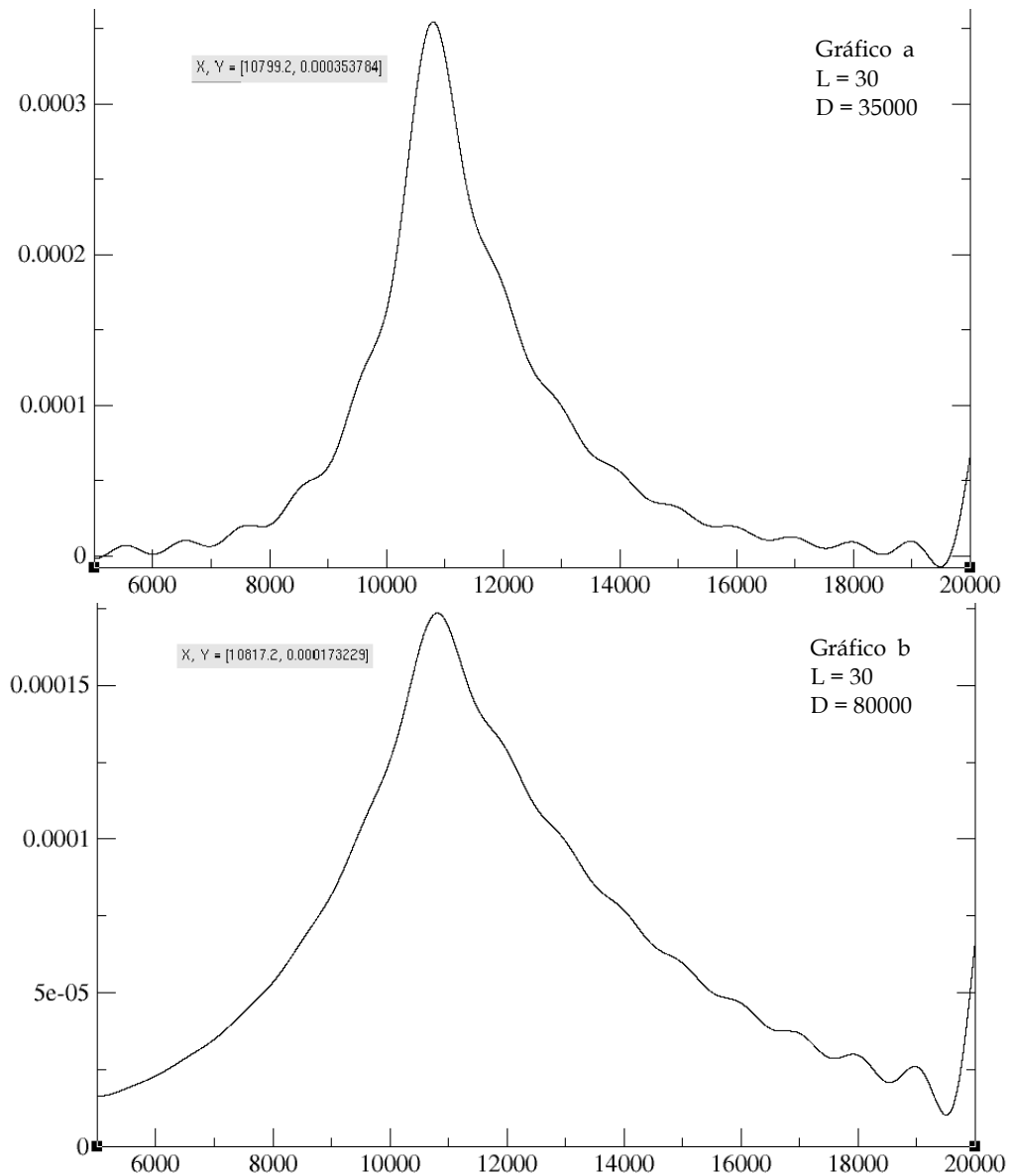


Fig. 7-4-1 Gráfico a, b. La densidad no es correcta en el gráfico a. Se incrementa el parámetro D, se observa en el gráfico b una mejor estructura en la densidad la cual es correcta.

7.5 APLICACIÓN DEL MÉTODO SFP CON DEMANDA DE PRODUCTO 0438

Producto 0438	p = Precio venta c= costo producir (\$ / kg)	Mejor valor ->	Solución SFP (kg)	Parámetros L - D
	p = 84 c = 39		3738.38	30 - 80000

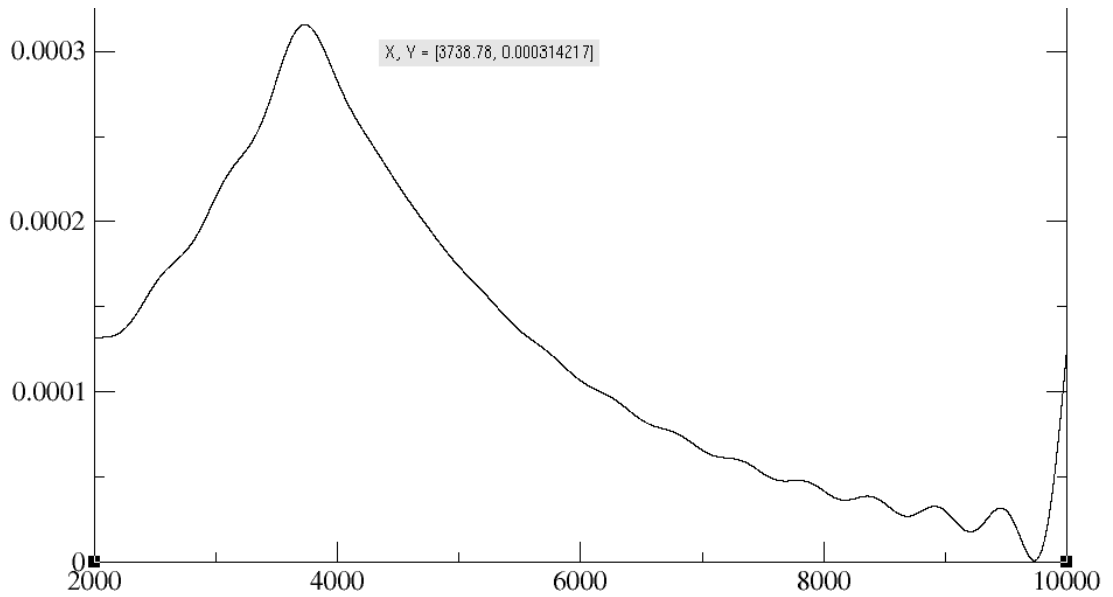


Fig. 7-5-1 La densidad no es correcta en el gráfico. Se observa la mejor estructura en la densidad la cual es correcta, es decir, estacionaria.

Las demandas utilizadas en el problema de minimización de costos de producción son datos históricos mensuales que ya se presentaron, no se conoce la forma en específico de la distribución de probabilidad de demanda. Se quiere saber cuál es el inventario óptimo mensual. Se conoce la historia mensual de demanda de los cinco productos.

El método SFP converge con la densidad positiva, estacionaria, que muestra la zona donde con mayor probabilidad se encierra la solución óptima global.

7.6 RESULTADOS CON ALGORITMO DE MONTE CARLO

El método de Monte Carlo es una herramienta de muestreo aleatorio que puede usarse para simular probabilidades marginales y conjuntas, estimadores estadísticos y en aplicaciones estadísticas donde se hacen estimaciones basadas en distribuciones de probabilidad de múltiples variables o funciones de densidad de probabilidad que está inmersa en alguna constante desconocida. El método obtiene buenos resultados en la generación de muestras desde distribuciones conjuntas basadas en un muestreo más fácil de las distribuciones condicionales.

En el problema general de Monte Carlo se supone que hay un proceso que genera un vector aleatorio X en alguna función f y se quiere encontrar el valor esperado de la función $E[f(X)]$, el vector aleatorio X es continuo con una función de densidad de probabilidad asociada $f(x)$. Se analizan muestras dependientes con la distribución limitante correspondiente a la distribución de la variable de interés. Se describe el método, en forma general: [8]:

$$E[f(X)] = \int f(x)f(x)dx, \quad (7.1)$$

donde los límites de la integral están en el dominio de X . La densidad representa la distribución de la(s) variable(s) aleatorias de interés. En algunos casos, la densidad concierne a un subconjunto de elementos del vector aleatorio X (e.g. puede representar la densidad marginal solo para la primer componente del vector X). La aproximación de la integral puede ser mediante posibles muestras dependientes X_k [8]. Sólo que resolver esta integral si se tiene un problema de múltiples variables es difícil, y computacionalmente tardado.

Un método Monte Carlo para aproximar la integral es dibujar las N muestras independientes, idénticamente distribuidas de X , refiérase como X_k con $k = 1, \dots, N$, de la función de densidad $f(x)$. Entonces tenemos un promedio basado en estas muestras independientes [8]:

$$E[f(X)] \approx \frac{1}{N} \sum_{k=1}^N f(X_k) \quad (7.2)$$

Debido a que las muestras son independientes e idénticamente distribuidas, con el tamaño de muestras se asegura la aproximación del método, por lo que se alcanza la precisión deseada al incrementar N . Aunque no siempre es factible dibujar muestras para x , de la densidad $f(x)$, la densidad puede ser muy complicada y a veces no conocida analíticamente.

No obstante, el algoritmo SFP nos brinda la estructura de esta densidad, y se conoce. Como vimos, el método funciona para el problema del vendedor de periódicos con demanda unimodal, el cual es muy conocido y se obtiene la estructura de la densidad de probabilidad (distribución) asociada al objetivo del problema.

Por lo que, se aplica el método también en un problema real con demandas históricas mensuales de cinco productos cárnicos, estas demandas son datos que provienen de una fuente productora de productos embutidos. El método SFP obtuvo la estructura de la densidad donde se dibuja la zona de mayor probabilidad que encierra el óptimo para el problema de minimización de costos de producción, y se refleja en la función objetivo. El vector aleatorio X de Monte Carlo en este caso, es el vector de demandas y es presentado como datos históricos de demanda que ya han tomado su realización durante varios meses para esta empresa. La información histórica de estos datos es conocida ahora.

Entonces para fines comparativos, se aplica el método Monte Carlo tomando $N = 105$ muestras independientes, idénticamente distribuidas de estas demandas. El resultado del método Monte Carlo se compara con el obtenido de SFP haciendo promedios. La solución x^* que arroja la estructura de la densidad dibujada por medio de SFP se evalúa haciendo promedio con estas N demandas. Se toman 15 soluciones locales o sub-óptimas, cercanas a la

solución óptima de SFP y se realiza Monte Carlo haciendo las 15 iteraciones, una por cada solución de SFP, en cada una se evalúan las N demandas en la función objetivo con x y se hace promedio. Se hacen las comparaciones de $f(x^*)$ y $f(x)$. Esto para cada uno de los cinco productos.

El algoritmo queda como sigue:

1. Obtener x^* de SFP y hacer $f(x^*)$ haciendo promedio sobre N demandas (v.i.i.d.). Guardar.
2. Obtener $x = 15$ soluciones locales (cercanas a x^*) de sfp.
3. Para cada x hacer $f(x)$: evaluar x en función objetivo f con las N demandas.
4. Hacer el promedio sobre N con la $f(x)$ del paso 3.
5. Si $f(x) < f(x^*)$, guardar $f(x)$ como actual (comparando con resultado de SFP).
6. Regresar al paso 3.

Las tablas comparativas se muestran a continuación para cada uno de los productos. El resultado de SFP y el obtenido con Monte Carlo.

Producto 0397			
Precio venta (\$/kg) $p = 11.0$		Costo producir \$/kg) $c = 8.0$	
Solución SFP óptima: x^*		Solución SFP local (cerca a x^*): x	
Número de muestras = 15		Demandas a promediar $N = 105$	
Método SFP		Algoritmo Monte Carlo	
Solución SFP (x^*)	Valor objetivo $f(x^*)$	Soluciones de SFP locales (x)	Valor objetivo $f(x)$ Monte Carlo
3563.51	-10690.53	3563.09	-10689.27
		3564.64	-10693.92
		3566.96	-10700.88
		3565.41	-10696.23
		3539.92	-10619.76
		3541.47	-10624.41
		3543.01	-10629.03
		3545.33	-10635.99
		3547.65	-10642.95
		3551.51	-10654.53
		3554.60	-10663.80
		3556.91	-10670.73
		3558.46	-10675.38
		3560.78	-10682.34
		3562.32	-10686.96

Tabla 7.6-1 Tabla comparativa de resultados con método SFP y Monte Carlo para producto 0397. En problema del vendedor de periódicos. Se hace una muestra de 15 soluciones locales. La solución resaltada en el recuadro es la mejor (mínima) para Monte Carlo.

Producto 0419			
Precio venta (\$/kg) $p = 42.0$		Costo producir (\$/kg) $c = 19.0$	
Solución SFP óptima: x^*		Solución local (cerca a x^*): x	
Número de muestras = 15		Demandas a promediar $N = 105$	
Método SFP		Algoritmo Monte Carlo	
Solución SFP (x^*)	Valor objetivo $f(x^*)$	Soluciones SFP locales (x)	Valor objetivo $f(x)$ Monte Carlo
19485	-378604.2	19413	-378348.6
		19437	-378449.4
		19461	-378532.2
		19509	-378676.2
		19533	-378748.2
		19389	-378247.8
		19482	-378595.2
		19490	-378619.2
		19497	-378640.2
		19493	-378628.2
		19497	-378640.2
		19500	-378649.2
		19504	-378661.2
		19508	-378673.2
		19501	-378652.2

Tabla 7.6-2 Tabla comparativa de resultados con método SFP y Monte Carlo producto 0419. El resultado marcado con un cuadro en negro, es la mejor (mínima) para Monte Carlo.

Producto 0429			
Precio venta (\$/kg) $p = 68.0$		Costo producir (\$/kg) $c = 30.0$	
Sol. SFP óptima: x^*		Solución local (cercana a x^*): x	
Número de muestras = 15		Demandas a promediar $N = 105$	
Método SFP		Algoritmo Monte Carlo	
Solución SFP (x^*)	Valor objetivo $f(x^*)$	Soluciones SFP locales (x)	Valor objetivo $f(x)$ Monte Carlo
1719.57	-63971.08	1710.68	-63667.8072
		1711.15	-63683.8410
		1712.10	-63716.2495
		1712.57	-63732.2832
		1713.04	-63748.3170
		1714.94	-63813.1341
		1715.41	-63829.1678
		1716.83	-63877.6101
		1717.77	-63909.6775
		1718.72	-63942.0861
		1719.66	-63974.1535
		1720.14	-63990.5284
		1721.08	-64022.5958
		1722.03	-64055.0044
		1723.44	-64103.1055

Tabla 7.6-3 Tabla comparativa de resultados con método SFP y Monte Carlo para producto 0429. El resultado marcado con un cuadro en negro, es la mejor (mínima) para Monte Carlo.

Producto 0432			
Precio venta (\$/kg) $p = 56.0$		Costo producir (\$/kg) $c = 21.0$	
Solución SFP óptima: x^*		Solución SFP local (cercana a x^*): x	
Número de muestras = 15		Demandas a promediar $N = 105$	
Método SFP		Algoritmo Monte Carlo	
Solución SFP (x^*)	Valor objetivo $f(x^*)$	Soluciones SFP locales (x)	Valor objetivo $f(x)$ Monte Carlo
10817.2	-372251.60	10818.4	-372284.0
		10819.5	-372313.7
		10820.6	-372343.4
		10821.7	-372373.1
		10823.9	-372432.5
		10816.2	-372224.6
		10815.1	-372194.9
		10814.0	-372165.2
		10812.9	-372135.5
		10811.8	-372105.8
		10810.7	-372076.1
		10809.6	-372046.4
		10808.5	-372016.7
		10826.1	-372491.9
		10825.0	-372462.2

Tabla 7.6-4 Tabla comparativa de resultados con método SFP y Monte Carlo para producto 0432. El resultado marcado con un cuadro en negro, es la mejor (mínima) para Monte Carlo.

Producto 0438			
Precio venta (\$/kg) $p = 84.0$		Costo producir (\$/kg) $c = 39.0$	
Sol. SFP óptima: x^*		Solución local (cerca a x^*): x	
Número de muestras = 15		Demandas a promediar $N = 105$	
Método SFP		Algoritmo Monte Carlo	
Solución SFP (x^*)	Valor función objetivo $f(x^*)$	Soluciones SFP locales (x)	Valor objetivo $f(x)$ Monte Carlo
3730	-167733.20	3729.60	-167715.840
		3728.79	-167680.686
		3728.38	-167662.892
		3727.98	-167645.532
		3727.58	-167628.172
		3726.77	-167593.018
		3725.96	-167557.864
		3725.16	-167523.144
		3724.75	-167505.350
		3730.40	-167750.560
		3731.21	-167785.714
		3731.61	-167803.074
		3732.02	-167820.868
		3732.82	-167855.588
		3733.63	-167890.742

Tabla 7.6-5 Tabla comparativa de resultados con método SFP y Monte Carlo para producto 0438. El resultado marcado con un cuadro en negro, es la mejor (mínima) para Monte Carlo, para ese mínimo local.

Las soluciones locales tomadas del método SFP fueron usadas para calcular la aproximación mediante Monte Carlo, y por los resultados que se obtuvieron, se puede observar que la solución del método SFP es de muy buena calidad comparada con los valores del objetivo evaluados con Monte Carlo.

Por lo que se concluye que el método SFP obtiene soluciones muy satisfactorias, para este caso donde se tienen valores históricos de demanda para cada uno de los productos cárnicos.

Capítulo 8

CONCLUSIONES

Se ha demostrado que el método de muestreo de Fokker-Planck funciona eficientemente en la solución del problema estocástico clásico del vendedor de periódicos con demanda unimodal, se aplicó en el caso con demanda bimodal y con demanda bimodal en dos intervalos, y se observaron resultados muy satisfactorios.

Por los resultados obtenidos, se aplicó el método en la solución de un caso aplicado, donde no se conoce información de la distribución de las demandas de varios productos, en cambio, se tiene información de las demandas reales históricas, es decir, ya se realizaron con alguna distribución no conocida y se cuenta con esos datos. Aplicando el método SFP en este caso, se observa que funciona, se obtiene la densidad correcta y estacionaria que indica la zona de mayor probabilidad donde se encuentra el mínimo valor de la variable de decisión. Luego, se compararon resultados con los del algoritmo Monte Carlo y se observa que la solución de SFP es muy buena.

Se observa que por medio de SFP, la construcción de la densidad de probabilidad de la variable de solución, refleja la estructura del beneficio en la función objetivo y encierra la zona donde con alta probabilidad se encuentra el óptimo global.

Por lo que, de acuerdo a los resultados observados en el problema estocástico del vendedor de periódicos sin restricciones, con demandas de comportamiento unimodal y bimodal, y con datos históricos de demanda, se puede concluir que el método de muestreo de Fokker-Planck puede ser una herramienta eficiente en la solución de problemas de optimización estocástica.

Se puede concluir que el método de muestreo de Fokker-Planck es una herramienta eficiente en la solución de problemas estocásticos sin restricciones.

Capítulo 9

CONTRIBUCIÓN

La contribución de este trabajo de investigación está enfocada en la aplicación del método de muestreo de Fokker-Planck en un problema de optimización estocástica, donde se involucran variables con incertidumbre y sin restricciones.

Este método no se había implementado en este ramo de la optimización, por lo que su introducción en el área estocástica se desarrolla en esta tesis. Se aplica el método SFP en el problema clásico del vendedor de periódicos, además de un caso con demandas reales históricas donde no se conocía su distribución y se observan resultados muy satisfactorios.

Los principales resultados de esta tesis han sido publicados en: Learning probability densities of optimization problems with constraints and uncertainty. En *Advances in Artificial Intelligence, Lecture Notes in Computer Science*: Vol. 5317, ISBN: 978-3-540-88635-8, Springer. Para el 7th Mexican International Conference on Artificial Intelligence, MICAI 2008 (27 al 31 octubre).

Así también, se presentó este trabajo de investigación en el programa de seminarios del Posgrado en Ingeniería de Sistemas en febrero y octubre del 2008.

APÉNDICE A

Derivada direccional de funciones.

Sea $f(x, y)$ una función de dos variables y $\mathbf{u} = (\cos\theta, \operatorname{sen}\theta)$, $0 \leq \theta \leq 2\pi$, un vector unitario. Se llama *derivada direccional* de f en (a, b) en la dirección de \mathbf{u} al siguiente límite (si existe):

$$D_{\mathbf{u}}f(a, b) = D_{\theta}f(a, b) = \lim_{h \rightarrow 0} \frac{f(a + h\cos\theta, b + h\operatorname{sen}\theta) - f(a, b)}{h}$$

Cuando la función es diferenciable en el punto, la derivada direccional se puede expresar en función de las derivadas parciales:

$$D_{\mathbf{u}}f(a, b) = D_{\theta}f(a, b) = f_x(a, b)\cos\theta + f_y(a, b)\operatorname{sen}\theta$$

Gradiente de funciones de dos variables

Se le llama *gradiente* de la función diferenciable f , al vector cuyas componentes son las derivadas parciales:

$$\nabla f(x, y) = (f_x(x, y), f_y(x, y))$$

Usando el *gradiente*, la derivada direccional se puede expresar mediante el producto escalar:

$$D_{\mathbf{u}}f(a, b) = \nabla f(a, b) \cdot \mathbf{u}$$

Propiedades del *gradiente* de una función de dos variables:

Si $\nabla f(a, b) = 0$, entonces $D_{\mathbf{u}}f(a, b) = 0$, para todo \mathbf{u} .

La derivada direccional en (a, b) es máxima en la dirección del vector *gradiente* $\nabla f(a, b)$ (dirección de máximo incremento de f), siendo $\|\nabla f(a, b)\|$ en su valor máximo.

La derivada direccional en (a, b) es mínima en la dirección del vector *gradiente* $-\nabla f(a, b)$ (dirección de mínimo incremento de f), siendo $-\|\nabla f(a, b)\|$ en su valor mínimo.

La derivada direccional de una función de tres variables $f(x, y, z)$ en (a, b, c) en la dirección del vector unitario $\mathbf{u} = (u_1, u_2, u_3)$, es: $\nabla f(a, b, c) \cdot \mathbf{u}$, donde $\nabla f(a, b, c) = (f_x(a, b, c), f_y(a, b, c), f_z(a, b, c))$, es el vector *gradiente*, que tiene las mismas propiedades que el caso de dos variables.

APÉNDICE B

TABLA DE DATOS HISTÓRICOS. DEMANDA MENSUAL DE PRODUCTOS CÁRNICOS (KG). Proporcionados por empresa particular.

* Estos datos fueron utilizados en el caso práctico del capítulo 7.

PRODUCTO				
0397 *	0419 *	0429 *	0432 *	0438 *
6676	12295	1133	9138	5089
5033	13134	2896	11986	6809
5647	10722	2890	10734	6686
6074	11907	2834	12035	6024
7072	13899	3497	11406	6886
6378	14314	3400	9860	7373
6003	12533	3400	10154	5835
5729	15351	4232	10910	6445
5897	15671	3734	10608	6857
4773	13847	3191	9447	5934
5410	13698	4661	11208	6297
4602	11907	4678	9443	5668
5057	12577	4462	10831	6102
5242	14334	4283	10889	6671
4201	12716	2959	9640	5288
5194	12378	3444	10004	5899
5618	16015	4882	12498	7051
5694	14048	3818	10510	6765
5455	16744	4149	11408	5937
5356	15620	3285	10697	6157
5436	19451	3201	10420	5905
5376	19062	3927	13323	6302
5062	13554	3884	10868	5325
5259	17338	3509	12120	5789
4167	16700	2593	11518	5736
6043	16421	2123	11439	5793
5761	19538	2874	13276	5420
5414	21509	3276	14445	6694
4527	16760	3898	13381	5570
4795	17360	3303	12667	6293
5781	22103	3928	14875	7426
4425	17453	3690	13649	5820
4793	20695	3547	12684	6036
4404	18239	3353	12137	5444
5069	15814	3814	12171	7012
4855	19897	3258	12966	5596
5552	21237	4662	11728	6687
4860	16482	3275	10877	5340
5618	25294	3384	12874	6373
5105	22448	3867	12203	6091
4821	19448	3313	11597	6226
4290	16052	3125	10861	5471
4962	17139	3360	12196	6057
5376	25254	3819	13189	6485
4869	21365	3611	12940	6512
4003	16885	2679	11447	5646
4509	15352	2914	10997	5863
5036	19439	3028	13250	6493
5623	22688	3091	14104	6425

5113	21774	2248	12926	5209
6670	26704	1586	12816	6563
4494	18206	859	8488	3667
5419	21351	1811	12307	4898
5340	20106	1453	14084	7170
5367	18276	3058	14337	6812
5661	16935	4779	14190	6120
6049	26412	3400	14307	7025
5236	24637	5719	14605	5870
4859	24053	3657	12773	5597
4659	20080	3308	12993	5538
5048	21684	3152	14802	6783
4926	20682	2927	13623	6292
4985	27529	3404	14072	5957
4542	22530	3425	12534	5729
5103	23656	2844	13270	6094
4367	21100	2365	13583	4966
5188	22186	2406	10622	6332
4544	19330	3279	15159	5688
5543	26584	3312	15830	7132
4002	28283	2889	10829	5793
5767	25640	4154	17180	6013
4220	22260	2936	15077	5691
3986	16287	2215	13659	5719
5050	20953	2566	16747	5964
5504	22042	2610	17020	6679
4417	19832	1911	15497	6392
4777	17377	1880	14804	4805
5131	18802	2085	15456	5794
5676	21434	2189	16176	5962
5113	21141	2390	16129	5805
5397	18485	2344	15724	5184
4907	19103	2385	15065	5619
5412	21388	2716	13887	6299
4863	21887	2942	16325	6083
5040	21268	2762	13892	6427
5178	22202	3004	13855	6266
4685	18471	2766	14639	6670
5362	21001	3946	15745	7231
5525	20177	2704	14245	7096
5997	18824	3455	14152	7876
6074	20345	3036	16110	6074
5206	22014	3777	14918	7609
4734	21516	3324	14515	6339
5192	17618	2606	12394	5482
5042	22617	3126	13844	6495
5763	29374	3433	16120	6671
5732	23319	3729	17221	5880
5009	23307	2673	12333	6296
4935	23452	2615	13178	6022
5330	24703	3371	15645	6338
5555	24445	2537	15254	5921
6393	22904	3020	15220	6294
7200	29605	2160	17148	6521
6669	28134	1533	14124	5332
3578	19576	1634	10586	3647

BIBLIOGRAFÍA

- [1] Birge J. R., Louveaux F.: Introduction to stochastic programming- Springer series in operations research, Springer, New York, 1997.
- [2] Berrones, Arturo: Stationary probability density of stochastic search processes in global optimization, Journal of Statistical Mechanics, 2008.
- [3] Glover, F., Ghaziri H.: Optimización heurística y redes neuronales, Paraninfo, España, 1996.
- [4] Press W., Teukolsky S., Vetterling W. and Flannery B.: Numerical recipes in C++, the Art of Scientific Computing, Cambridge University Press, Cambridge, 2005.
- [5] Bazaraa M., Jarvis J., Sherali H. : Programación lineal y flujo en redes, Limusa, 2004.
- [6] Dyer M., Stougie L. : Computational complexity of stochastic programming problems, Springer-Verlag, United Kingdom - Amsterdam, 2005.
- [7] López A., Velazco J., Berrones A. : Learning probability densities of optimization problems with constraints and uncertainty, MICAI 2008: Advances in Artificial Intelligence, Lecture Notes in Computer Science, Vol. 5317, ISBN: 978-3-540-88635-8, Springer.
- [8] Spall, J., Introduction to stochastic search and optimization - Estimation, simulation and control, Wiley-Interscience, New Jersey, 2003.