

Optimización bajo Incertidumbre. Enfoques Estocásticos y Robustos Aplicados al Problema de la Mochila

Optimization under Uncertainty. Stochastic and Robust Frameworks for the Knapsack Problem

Trabajo Fin de Grado en Matemáticas
Universidad de Málaga

Autora: Carolina Berrocal Enríquez.

Área de conocimiento y/o departamento: Área de Estadística e Investigación Operativa. Departamento de Análisis Matemático, Estadística e Investigación Operativa y Matemática Aplicada.

Fecha de presentación: Noviembre de 2025.

Tema: Optimización bajo Incertidumbre.

Tipo: Trabajo de revisión bibliográfica.

Modalidad: Individual.

Número de páginas (sin incluir introducción, bibliografía ni anexos): 48

DECLARACIÓN DE ORIGINALIDAD DEL TFG

D./Dña. *Carolina Berrocal Enríquez*, con DNI (NIE o pasaporte) [REDACTED] estudiante del Grado en *Matemáticas* de la Facultad de Ciencias de la Universidad de Málaga, **DECLARO:**

Que he realizado el Trabajo Fin de Grado titulado “*Optimización bajo Incertidumbre. Enfoques Estocásticos y Robustos Aplicados al Problema de la Mochila*” y que lo presento para su evaluación. Dicho trabajo es original y todas las fuentes bibliográficas utilizadas para su realización han sido debidamente citadas en el mismo.

De no cumplir con este compromiso, soy consciente de que, de acuerdo con la normativa reguladora de los procesos de evaluación de los aprendizajes del estudiantado de la Universidad de Málaga de 23 de julio de 2019, esto podrá conllevar la calificación de suspenso en la asignatura, sin perjuicio de las responsabilidades disciplinarias en las que pudiera incurrir en caso de plagio.

Para que así conste, firmo la presente en Málaga, [REDACTED]

Fdo:.....

Índice general

Resumen	1
Abstract	1
Introducción	1
1. Programación estocástica	10
1. Aproximación por Promedio de Muestras (SAA)	14
2. Programación estocástica con decisiones en el tiempo	20
2.1. Programación estocástica de dos etapas	20
2.2. Programación estocástica multietapa	26
2. Optimización Robusta	34
1. Introducción	34
2. Conjuntos de incertidumbre y reformulaciones	35
2.1. Conjuntos de incertidumbre poliédricos	36
2.2. Conjuntos de incertidumbre tipo presupuesto	42
3. Experimentos computacionales	49
4. Conclusiones	57
A. Teoremas	58
B. Tablas	60
Bibliografía	63

Optimización bajo Incertidumbre. Enfoques Estocásticos y Robustos Aplicados al Problema de la Mochila

Resumen

La incertidumbre es inherente a los problemas reales y su adecuada incorporación en la optimización permite diseñar soluciones más sólidas y aplicables. No obstante, su tratamiento no es una tarea trivial, ya que, en general, no existe una formulación determinista única y tratable sin hipótesis adicionales. En este trabajo se analizan dos marcos fundamentales que reformulan o aproximan problemas con incertidumbre como problemas de optimización lineales (LP), cuyas soluciones óptimas son acordes a las características de la incertidumbre y a la actitud del decisor: la Programación Estocástica y la Optimización Robusta

El Capítulo 1 introduce los enfoques estocásticos, desde el modelo clásico de una etapa hasta formulaciones multietapa. Se definen los árboles de escenarios así como dos vías estándar para hacer el problema estocástico tratable: la formulación determinista equivalente (para incertidumbre discreta y finita) y la aproximación por medias muestrales (SAA) (aplicable de forma general). Todos estos enfoques se aplican a un ejemplo ilustrativo común que actúa como hilo conductor y permite comparar cómo varía el valor objetivo al aumentar el número de etapas. Además, se discuten las principales ventajas y limitaciones de estos enfoques.

El Capítulo 2 aborda la optimización robusta (RO). Se introduce el concepto de conjunto de incertidumbre y se estudian dos familias de especial relevancia: los conjuntos poliédricos (con énfasis en los conjuntos tipo caja como caso particular) y los conjuntos de tipo presupuesto. Se muestra su reformulación lineal y se aplica el marco al ejemplo ilustrativo para facilitar la comprensión y la comparación directa con los enfoques estocásticos.

El Capítulo 3 presenta una adaptación del problema de la mochila fraccionaria donde las utilidades de los objetos están sujetas a incertidumbre. Se implementan y comparan los modelos estudiados con experimentos específicos que permiten analizar cómo influye el nivel de incertidumbre en el valor objetivo para cada modelo: vemos cómo mejora dicho valor al aumentar el grado de adaptabilidad en la programación estocástica multietapa o reducir el número de parámetros que pueden tomar el peor caso en el caso del enfoque de Bertsimas y Sim. Asimismo, se estudia el coste computacional de cada enfoque.

Finalmente, el Capítulo 4 sintetiza las conclusiones del trabajo.

Palabras clave:

PROGRAMACIÓN ESTOCÁSTICA, PROGRAMACIÓN ESTOCÁSTICA MULTIETAPA, APROXIMACIÓN POR PROMEDIO DE MUESTRAS (SAA), OPTIMIZACIÓN ROBUSTA, CONJUNTO DE INCERTIDUMBRE PRESUPUESTARIO, PROBLEMA DE LA MOCHILA FRACCIONARIA.

Optimization under Uncertainty. Stochastic and Robust Frameworks for the Knapsack Problem

Abstract

Uncertainty is intrinsic to most practical problems due to the stochastic nature of real phenomena. Incorporating it into optimization models leads to solutions that are more realistic and perform better under unobserved scenarios. However, solving optimization problems under uncertainty is nontrivial, as there is generally no unique, tractable deterministic formulation without additional assumptions. This work presents, in a rigorous yet accessible manner, tools that reformulate or approximate uncertain problems as linear optimization problems (LP), whose optimal solutions align with both the characteristics of the uncertainty and the decision-maker's risk attitude.

Chapter 1 introduces *stochastic approaches*, from the classical one-stage model to multi-stage formulations. We define scenario trees and present two standard routes to tractability: (i) the deterministic equivalent formulation for finite discrete uncertainty, and (ii) sample average approximation (SAA), applicable more broadly. All approaches are applied to a common illustrative example that serves as a running thread and enables a direct comparison of how the objective value evolves as the number of stages increases. We also discuss the main advantages and limitations of these approaches.

Chapter 2 addresses robust optimization (RO). We introduce the notion of an uncertainty set and study two particularly relevant families: polyhedral sets (emphasizing box sets as a special case) and budgeted sets. We show their linear reformulations and apply the framework to the illustrative example to facilitate understanding and enable a direct comparison with the stochastic approaches.

Chapter 3 presents an adaptation of the fractional knapsack problem in which item utilities are subject to uncertainty, formulated so that it can be solved by all the approaches introduced. We implement and compare the models, designing targeted experiments to analyze how the uncertainty level affects the objective value for each model; how that value improves as adaptivity increases in multi-stage stochastic programming; and how it changes when reducing the number of coefficients allowed to attain their worst case in the Bertsimas–Sim framework. We also examine the computational cost of each approach.

Finally, Chapter 4 synthesizes the main conclusions of the work.

key words:

STOCHASTIC PROGRAMMING, MULTISTAGE STOCHASTIC PROGRAMMING, SAMPLE AVERAGE APPROXIMATION (SAA), ROBUST OPTIMIZATION, BUDGETED UNCERTAINTY SET, FRACTIONAL KNAPSACK PROBLEM.

Introducción

En numerosos enfoques clásicos de la optimización matemática se asume que todos los parámetros del modelo son conocidos con exactitud, es decir, se trabaja con modelos deterministas. Sin embargo, esta hipótesis rara vez se cumple en la práctica. Muchos problemas reales están sujetos a errores de estimación, fluctuaciones en los datos o a la propia naturaleza estocástica de los fenómenos implicados. Ignorar esta incertidumbre puede dar lugar a decisiones subóptimas que, en algunos casos, han estado detrás de consecuencias catastróficas.

Un ejemplo es el desastre del transbordador *Challenger* en 1986, [1]. Los ingenieros de la NASA no incorporaron adecuadamente la incertidumbre asociada a las bajas temperaturas del día del lanzamiento y su efecto sobre algunos componentes del transbordador. A pesar de disponer de datos que indicaban riesgos relevantes, [2], se asumió un comportamiento habitual del sistema. Esta omisión derivó en una fuga de combustible, la explosión del vehículo y la muerte de toda la tripulación.

Casos más recientes refuerzan esta idea. Por ejemplo, durante la crisis financiera de 2008, muchos modelos matemáticos empleados para evaluar el riesgo de las inversiones bancarias resultaron ser excesivamente simplistas. Estos modelos asumían, por un lado, que los activos financieros (como bonos o hipotecas) eran independientes entre sí, es decir, que una caída en el valor de un activo no afectaría al resto. Por otro lado, también suponían que los rendimientos seguían una distribución normal, lo que implica que los eventos extremos son muy improbables. Aunque estas hipótesis pueden ser razonables en contextos de estabilidad, no refleja el comportamiento de los mercados en situaciones de crisis, donde el pánico generalizado provoca caídas simultáneas en numerosos activos.

En este contexto, a lo largo de los años previos a la crisis financiera, y especialmente entre 2003 y 2006, se concedieron numerosas hipotecas de alto riesgo que fueron agrupadas y titulizadas como si se tratara de activos financieros seguros. Los modelos utilizados por las agencias de calificación y las entidades financieras estimaban que era altamente improbable que un número significativo de estas hipotecas no se pagaran de forma simultánea. Sin embargo, a partir de 2007 comenzaron a acumularse los impagos, lo que provocó una rápida depreciación de dichos activos. Esta pérdida de valor generó una profunda desconfianza entre las instituciones financieras, lo que a su vez paralizó el mercado interbancario. La cadena de acontecimientos derivó en quiebras de gran magnitud, como la de Lehman Brothers, y obligó a los gobiernos y bancos centrales a intervenir para evitar el colapso total del sistema financiero, [3]. Como consecuencia, millones de personas perdieron sus empleos, sus viviendas y sus ahorros, y muchas economías entraron en una recesión profunda.

La pandemia de la COVID-19 constituye otro ejemplo paradigmático de las consecuencias de ignorar la incertidumbre. Durante esta crisis, el uso de modelos deterministas en múltiples ámbitos, desde la logística hasta la sanidad, evidenció sus limitaciones ante shocks globales. En la gestión hospitalaria, la falta de planificación frente a escenarios

extremos provocó escasez de camas UCI, respiradores y material de protección individual. En el sector industrial y logístico, la ausencia de márgenes de seguridad en los modelos de inventario conllevó graves disrupciones en las cadenas de suministro, afectando desde productos farmacéuticos hasta bienes de consumo esenciales. Incluso en el ámbito educativo, el diseño de políticas sin contemplar variabilidad en la evolución epidemiológica dificultó la continuidad de la enseñanza.

En todos estos casos, la ausencia de un análisis explícito de la incertidumbre fue un factor determinante en el fallo de los sistemas de decisión. Así, llegamos a una conclusión esencial: incorporar la incertidumbre de forma explícita en los modelos de decisión es crucial para representar con mayor fidelidad la complejidad del mundo real.

Más aún, los modelos que contemplan la incertidumbre no solo proporcionan soluciones más robustas, sino que también permiten identificar qué variables o factores son los principales responsables de dicha incertidumbre. Esta información puede resultar decisiva. Por ejemplo, si se desea predecir el precio de un recurso estratégico a partir de datos históricos, un análisis determinista podría limitarse a estimaciones pasadas. En cambio, un análisis bajo incertidumbre podría revelar que la variabilidad está dominada por decisiones políticas internas del país productor. En tal caso, resultaría más provechoso enfocar los esfuerzos en estudiar escenarios geopolíticos futuros que en mejorar la obtención y/o los análisis de los datos históricos.

Aunque en este trabajo nos centraremos en problemas de optimización lineal bajo incertidumbre, es conveniente presentar en primer lugar la formulación general de un problema de programación lineal determinista, que constituirá la base común hacia la que convergerán la mayor parte de las formulaciones posteriores:

$$\begin{aligned} \min_x \quad & c'x \\ \text{s.a.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{LP}$$

donde $x \in \mathbb{R}^n$ es el vector de variables de decisión, $c \in \mathbb{R}^n$ el vector de coeficientes de la función objetivo, $A \in \mathbb{R}^{m \times n}$ la matriz de restricciones, y $b \in \mathbb{R}^m$ el vector de términos independientes. En este caso, la función objetivo $f : X \rightarrow \mathbb{R}$ viene dada por $f(x) = c'x$, siendo X el conjunto factible definido como:

$$X = \{x \in \mathbb{R}^n : Ax \leq b, x \geq 0\}.$$

Supondremos que X es no vacío y acotado para garantizar la existencia de solución óptima: si X es no vacío y acotado, como además X es cerrado (dado que X es intersección de conjuntos cerrados), entonces, por el teorema de Heine-Borel, (Apéndice A, Teorema 1), X es compacto. Por la continuidad de $c'x$ (por ser lineal), el teorema de Weierstrass, (Apéndice A, Teorema 2), garantiza la existencia de algún mínimo, por lo que existe alguna solución óptima.

En un problema del tipo (LP) se asume que los coeficientes del modelo son conocidos con total precisión. No obstante, como ya hemos mencionado, esta suposición dista mucho de cumplirse en la práctica. El ejemplo que se expone a continuación ilustra cómo pueden variar tanto la solución como el valor óptimo de un problema del tipo (LP) al modificar distintos parámetros del mismo, [4].

Ejemplo 1. *Consideramos una empresa química que comercializa dos disolventes, D_1 y D_2 . El primero se comercializa a 1€ por tonelada y el segundo a 4€ por tonelada.*

La empresa adquiere los disolventes con un día de antelación y los almacena hasta su venta, con el objetivo de maximizar el beneficio total obtenido. No obstante, se enfrenta a varias restricciones operativas. En primer lugar, la capacidad conjunta de almacenamiento de D_1 y D_2 está limitada a 8 toneladas diarias. En segundo lugar, por contrato con su proveedor, la cantidad de D_1 adquirida debe ser exactamente el doble de la cantidad de D_2 . Finalmente, debido a que el disolvente D_2 está fabricado con un compuesto químico escaso, la disponibilidad diaria de D_2 está limitada a un máximo de 5 toneladas. A partir de estas condiciones, se plantea el siguiente problema de optimización, reformulado en forma de minimización en lugar de maximización para ajustarse al formato adoptado en este trabajo:

$$\min_{x_1, x_2} \quad -x_1 - 4x_2 \quad (1a)$$

$$s.a. \quad 2x_1 + x_2 \leq 8 \quad (1b)$$

$$x_2 \leq 5 \quad (1c)$$

$$x_1, x_2 \geq 0 \quad (1d)$$

donde las variables de decisión son x_i , $i = 1, 2$, que representan la cantidad a vender del disolvente D_i . La función objetivo (1a) corresponde a las ganancias totales. La restricción (1b) hace referencia a la capacidad de almacenamiento para D_1 y D_2 mientras que la restricción (1c) impone la limitación de disponibilidad para D_2 . Por su parte, la restricción (1d) garantiza la no negatividad de las variables x_1 y x_2 . El problema anterior es un problema de programación lineal de tipo (LP), el cual puede resolverse mediante técnicas muy conocidas, como el método simplex, [5]. La región factible del problema viene delimitada por las líneas continuas azul, lila, verde y naranja en la Figura 1. La solución óptima es (1.5, 5), representada mediante un círculo rojo en la misma figura. El valor objetivo del problema original de maximización es 21.5€.

Veamos ahora cómo varía la región factible, y, en consecuencia, la solución óptima si consideramos que alguno de los parámetros del modelo varía, es decir, si introducimos incertidumbre en los parámetros.

Supongamos en primer lugar que la empresa ha decidido aumentar el espacio de almacenamiento para el disolvente D_1 y que, por tanto, el coeficiente que acompaña a la variable x_1 en (1b) pasa a ser 3 en lugar del valor original 2. De esta manera, el nuevo problema a resolver queda formulado como:

$$\begin{aligned} \min_{x_1, x_2} \quad & (1a) \\ s.a. \quad & 3x_1 + x_2 \leq 8 \\ & (1c), (1d) \end{aligned} \quad (D1)$$

La región factible de este problema queda delimitada por la línea discontinua verde, y las líneas continuas lila, azul y naranja en la Figura 1. Además, la solución óptima de este nuevo problema es el punto (1, 5) representado por el círculo verde. El valor óptimo de este nuevo problema es de 21€.

En segundo lugar, veamos qué ocurre si, en lugar de considerar incertidumbre en la matriz A de coeficientes de las restricciones, es la función objetivo la que sufre una variación. Por ejemplo, imaginemos que un estudio reciente ha demostrado que el material D_1 tiene un impacto negativo sobre el medio ambiente, pero nuestra empresa tiene que seguir vendiéndolo debido a un contrato vigente con algunos clientes, lo que se traduce en una pérdida de 1€ por tonelada vendida. El nuevo problema resultante es:

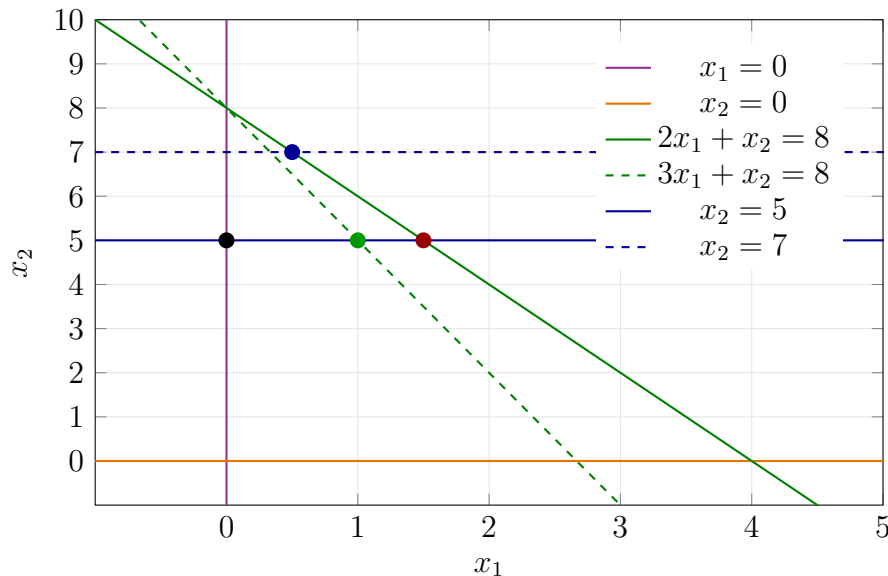


Figura 1: Inclusión de incertidumbre en problemas de programación lineal

$$\begin{aligned}
 & \underset{x_1, x_2}{\text{mín}} && x_1 - 4x_2 \\
 & \text{s. a.} && (1b), (1c), (1d).
 \end{aligned} \tag{D2}$$

La región factible es la misma que en (1), mientras que su solución óptima es, en este caso, $(0, 5)$, representada en la Figura 1 por el punto negro, con un valor óptimo de 20€ .

En tercer y último lugar, veamos qué ocurre si modificamos el término independiente de las restricciones. Supongamos que el dueño ha alquilado el almacén contiguo y ahora se pueden almacenar hasta 7 toneladas del disolvente x_2 . Teniendo en cuenta este cambio, el problema resultante es:

$$\begin{aligned}
 & \underset{x_1, x_2}{\text{mín}} && (1a) \\
 & \text{s. a.} && (1b), (1d), \\
 & && x_2 \leq 7.
 \end{aligned} \tag{D3}$$

La región factible del problema (D3) está delimitada por la línea discontinua azul y las líneas continuas lila, naranja y verde, tal como se muestra en la Figura 1. Su solución óptima es $(0.5, 7)$, representada por el punto azul, con valor óptimo 28.5€ .

Se puede observar que, dependiendo de las modificaciones en los datos del problema, la solución óptima y la región factible pueden verse afectadas. En otras palabras, resulta impredecible determinar cómo impactará la incertidumbre en nuestro problema. Esto plantea la cuestión natural: ¿existe alguna manera de obtener una solución óptima en presencia de incertidumbre que sea razonablemente satisfactoria ante esta falta de información? La respuesta es afirmativa, y precisamente de ahí surge la motivación de este Trabajo de Fin de Grado: estudiar y comparar varios de los principales marcos teóricos desarrollados para incorporar la incertidumbre en problemas de optimización.

Antes de introducir dichos marcos, es necesario establecer cómo vamos a modelizar la incertidumbre. Para ello, se introduce un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, donde Ω representa el conjunto de posibles escenarios o estados, \mathcal{F} es una σ -álgebra sobre Ω , y \mathbb{P} una medida de probabilidad definida sobre \mathcal{F} . $\xi : \Omega \rightarrow \Xi$ será un vector aleatorio de

dimensión p , \mathcal{F} -medible. Asumimos que el conjunto $\Xi \subseteq \mathbb{R}^p$, denominado soporte de ξ , es cerrado, está equipado con la σ -álgebra de Borel $\mathcal{B}(\Xi)$ y contiene todos los valores posibles que esta variable puede tomar.

En este contexto, la función objetivo se generaliza como una función $c : \Xi \rightarrow \mathbb{R}^n$ que depende de la realización de la incertidumbre ξ . Del mismo modo, las restricciones también pueden verse afectadas por ξ , lo que se traduce en una matriz de coeficientes $A(\xi)$ y un vector de términos independientes $b(\xi)$, ambos definidos sobre el espacio de escenarios Ξ .

Así, el problema de optimización lineal bajo incertidumbre puede escribirse, en su forma más general, como:

$$\begin{aligned} \min_x \quad & c(\xi)'x \\ \text{s.a.} \quad & A(\xi)x \leq b(\xi) \quad \forall \xi \in \Xi \\ & x \geq 0 \end{aligned} \tag{LP}_\xi$$

donde asumimos que $c(\xi)$, $A(\xi)$ y $b(\xi)$ son $\mathcal{B}(\Xi)$ -medibles, de modo que el problema está bien definido, [6]. La decisión x debe tomarse sin conocer la realización futura de ξ . En este contexto, una solución factible es aquella que verifica $A(\xi)x \leq b(\xi)$ para todo $\xi \in \Xi$, es decir, debe ser válida frente a cualquier posible realización del vector aleatorio.

En este modelo se considera que la incertidumbre puede afectar a todos los parámetros del problema, es decir, tanto el vector de costes c , como la matriz de restricciones A y el vector de términos independientes b dependen de la variable aleatoria ξ . No obstante, sin pérdida de generalidad, se puede asumir que la incertidumbre reside exclusivamente en la función objetivo o, alternativamente, sólo en la matriz de restricciones. Veamos con más detalle cómo se justifica cada caso.

En primer lugar, con el objetivo de concentrar la incertidumbre del problema $(LP)_\xi$ en la matriz de restricciones, veamos que dicho problema es equivalente a:

$$\begin{aligned} \min_{x,W} \quad & W \\ \text{s.a.} \quad & c(\xi)'x \leq W \\ & A(\xi)x - b(\xi) \leq 0 \\ & x \geq 0. \end{aligned} \tag{LP}_\xi^A$$

Por un lado, hemos trasladado $b(\xi)$ al miembro izquierdo de la desigualdad. Por otro lado, hemos reformulado la función objetivo mediante una variable auxiliar W que actúa como cota superior para $c(\xi)'x$. Incluyendo la restricción $c(\xi)'x \leq W$, aseguramos que W tome un valor mayor o igual que el coste efectivo $c(\xi)'x$. Como se está minimizando W , en el óptimo no es eficiente mantener $W > c(\xi)'x$, ya que podríamos reducir W sin violar la factibilidad. Por tanto, en el óptimo se cumple necesariamente que $W = c(\xi)'x$. Así, $(LP)_\xi$ y $(LP)_\xi^A$ son equivalentes, pues comparten el mismo conjunto de soluciones óptimas. De este modo, toda la dependencia respecto a ξ queda encapsulada en los coeficientes de las restricciones.

En segundo lugar, para concentrar la incertidumbre únicamente en la función objetivo, basta con observar que la solución óptima del problema $(LP)_\xi$ coincide con el límite de la sucesión de soluciones de problemas penalizados de la forma, [7]:

$$\min_x \{c(\xi)'x + s_k \phi(x)\} \tag{LP}_\xi^c$$

donde $\{s_k\}_{k \geq 1}$ es una sucesión creciente de constantes positivas tal que $s_k \rightarrow \infty$ y $\phi : \mathbb{R}^n \rightarrow \mathbb{R}_+$ es una función continua que se anula si y sólo si x es factible para $(LP)_\xi$, es

decir, si x verifica $A(\xi)x \leq b(\xi)$ para todo $\xi \in \Xi$ y $x \geq 0$. A este tipo de funciones se las denomina *funciones de penalización*.

Veamos que (LP_ξ) y (LP_ξ^c) son equivalentes con más detalle: en el problema (LP_ξ^c) , se minimiza la expresión $c(\xi)'x + s_k \phi(x)$. A medida que $s_k \rightarrow \infty$, cualquier solución no factible (es decir, tal que $\phi(x) > 0$) queda asociada a un valor objetivo cada vez mayor. En consecuencia, la sucesión de soluciones óptimas x_k de los problemas penalizados (LP_ξ^c) tiende a aproximarse al conjunto factible del problema original (LP_ξ) . De hecho, si x_k no fuese factible en el límite, entonces $\phi(x_k) > 0$, y dado que $s_k \rightarrow \infty$, el término de penalización $s_k \phi(x_k)$ también tendería a infinito, y con ellos, la función objetivo. En consecuencia, x_k no podría ser óptimo. Además, entre los puntos factibles, el término de penalización se anula, de modo que el problema penalizado coincide con el problema original. Por tanto, si x^* es un punto límite de la sucesión de soluciones de (LP_ξ^c) , x_k , se garantiza que x^* es factible y, al mismo tiempo, minimiza $c(\xi)'x$ dentro del conjunto factible. Es decir, x^* es solución óptima del problema original (LP_ξ) .

Por otro lado, si x^* es solución óptima de (LP_ξ) , entonces x^* es, en particular, factible, lo que implica que $\phi(x^*) = 0$. En ese caso, las funciones objetivo de (LP_ξ) y (LP_ξ^c) coincidirían, por lo que x^* también es óptimo de (LP_ξ^c) .

Con todas estas convenciones y explicaciones expuestas, podemos comenzar con las metodologías para aproximar el problema (LP_ξ) . A continuación, en el Capítulo 1, vemos la primera de ellas, la programación estocástica.

Capítulo 1

Programación estocástica

Históricamente, la programación estocástica comenzó a desarrollarse a mediados del siglo XX, impulsada por la necesidad de tomar decisiones óptimas en contextos inciertos. Uno de los trabajos fundacionales fue el de George Dantzig, quien propuso en 1955 un modelo de programación lineal bajo incertidumbre en dos etapas, [8]. Poco después, William Orchard-Hays contribuyó a su formalización computacional. András Prékopa introdujo más adelante los modelos con restricciones probabilísticas, ampliando notablemente el ámbito de aplicación, [9]. Por su parte, Ralph Tyrer Rockafellar, y Roger Jean-Baptiste Robert Wets, desarrollaron un marco teórico basado en análisis convexo y teoría de la medida que permitió avanzar en la formulación y resolución numérica de estos problemas, [10].

La programación estocástica se ha consolidado como una herramienta eficaz en numerosos contextos donde la incertidumbre desempeña un papel crucial. Entre sus aplicaciones más destacadas se encuentran la gestión de la cadena de suministro, la planificación energética, la optimización de carteras financieras y la asignación de recursos en el sector salud [11]. En todos estos casos, su capacidad para modelar decisiones bajo incertidumbre ha contribuido significativamente a mejorar tanto la eficiencia operativa como la resiliencia de los sistemas.

En el caso de un problema de programación estocástica cuya incertidumbre se concentra únicamente en la función objetivo (suposición que, como se vio en la Introducción, no conlleva pérdida de generalidad), la formulación estándar busca minimizar el valor esperado de la misma. De manera formal, este tipo de problema puede escribirse como:

$$\begin{aligned} \min_x \quad & \mathbb{E}_{\mathbb{P}}[c(\xi)'x] \\ \text{s. a.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{SP}$$

donde $\mathbb{E}_{\mathbb{P}}$ denota la esperanza de la variable aleatoria $f(x, \xi) = c(\xi)'x$ con respecto a la medida de probabilidad \mathbb{P} . En adelante, también escribiremos esta esperanza como \mathbb{E} o \mathbb{E}_{ξ} , cuando no haya riesgo de ambigüedad. El objetivo es encontrar una solución óptima x^* dentro del conjunto factible X que minimice el valor esperado de la función objetivo, teniendo en cuenta los posibles escenarios de incertidumbre representados por ξ .

La primera cuestión de fondo respecto al modelo (SP), que puede surgirnos es: ¿tiene sentido considerar este problema? Es decir, ¿tiene sentido considerar la esperanza como medida de riesgo? La respuesta es sí y tiene su justificación en la Ley de los Grandes Números (LGN), [12]. Los enunciados de sus versiones débil y fuerte se presentan en los Teoremas 3 y 4 del Apéndice A. Este resultado garantiza que, si se resolviera el mismo

problema bajo la misma distribución de probabilidad un número indefinido de veces, la solución del modelo (SP) proporcionaría, en promedio, el mejor resultado posible.

Profundicemos en esta idea. Sean $\xi_1, \xi_2, \dots, \xi_N$ variables aleatorias independientes e idénticamente distribuidas (i.i.d.) con distribución de probabilidad \mathbb{P} . A partir de ahora, por simplicidad, denotaremos por $\xi_\ell \in \mathbb{R}^p$ a cada muestra de ξ , reservando la letra ℓ para indexarlas. Entonces, para un x_0 fijo, la evaluación de la función objetivo también constituye una muestra aleatoria de N elementos, $f_\ell := f(x_0, \xi_\ell)$, $\ell = 1, \dots, N$. Entonces, de acuerdo con la Ley de los Grandes Números (LGN), se tiene que

$$\frac{f_1 + f_2 + \dots + f_N}{N} \xrightarrow{N \rightarrow \infty} \mathbb{E}[f],$$

donde $\mathbb{E}[f]$ representa el valor esperado verdadero de la distribución de probabilidad que genera los valores de $f(x_0, \xi)$. Cabe destacar que la convergencia puede ser casi segura o en probabilidad, según se aplique la versión fuerte o débil de la LGN. Con este argumento tenemos que:

$$\min_x \frac{f(x, \xi_1) + \dots + f(x, \xi_N)}{N} \quad \text{es equivalente en términos de convergencia a} \quad \min_x \mathbb{E}[f(x, \xi)]$$

Veamos un ejemplo para ilustrar esto mejor.

Ejemplo 2. Consideremos la función objetivo del Ejemplo 1,

$$f(x_1, x_2) = -x_1 - 4x_2$$

Asumamos ahora que el coeficiente de la variable x_1 es incierto, es decir:

$$f(x_1, x_2, \xi) = \xi x_1 - 4x_2.$$

Además, supongamos que x es fijo al valor $x_0 = (1, 1)$ y que ξ sigue una distribución uniforme en $[0, 10]$, es decir, $\xi \sim U(0, 10)$. Entonces, tenemos que

$$f(x_0, \xi) = \xi - 4$$

y

$$\mathbb{E}[f(x_0, \xi)] = \mathbb{E}[\xi - 4] = \mathbb{E}[\xi] - 4 = \frac{10 + 0}{2} - 4 = 5 - 4 = 1$$

donde la segunda igualdad se debe a la linealidad de la esperanza y la tercera se da porque la esperanza de una distribución uniforme en un intervalo de extremos a y b se calcula como el punto medio de dicho intervalo. Por tanto, si extraemos aleatoriamente una muestra independiente idénticamente distribuida (i.i.d) de ξ , tendremos también una muestra de $f(x_0, \xi)$ cuya media debe ser 1 de acuerdo a la LGN.

Con el fin de ilustrarlo gráfica y numericamente, se ha realizado un experimento representado en la Figura 2, tomando N valores de $\xi \sim U(0, 10)$, con $N \in \{10^i : i = 1, \dots, 5\}$. La línea roja continua de esta figura denota el valor medio teórico, $\mathbb{E}[f(x_0, \xi)]$, que en este caso es 1, mientras que la línea azul discontinua representa el valor medio estimado, $\sum_{\ell=1}^N \frac{f(x_0, \xi_\ell)}{N}$. Se puede observar que, efectivamente, a medida que $N \rightarrow \infty$, el valor medio estimado converge al teórico.

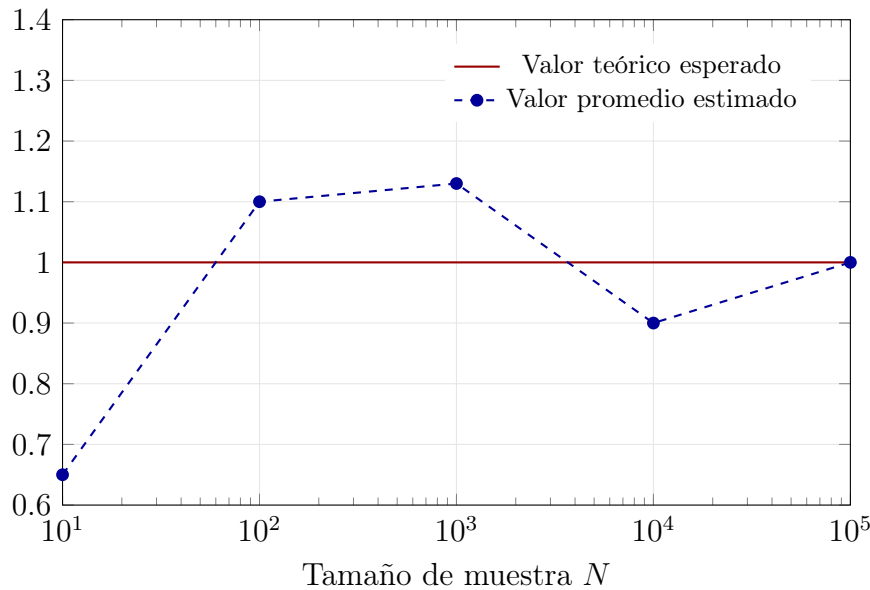


Figura 2: Convergencia del valor medio estimado al valor medio teórico.

Así, considerar la esperanza como criterio de optimización se justifica: garantiza el mejor rendimiento medio cuando el problema se repite múltiples veces bajo la misma distribución.

No obstante, existen otras medidas de riesgo que se pueden considerar en lugar de la esperanza, como la varianza, el valor en riesgo (*Value-at-Risk*, VaR) o el valor en riesgo condicional (*Conditional Value-at-Risk*, CVaR), que permiten capturar distintos aspectos del comportamiento bajo incertidumbre. Para más información sobre estas medidas de riesgo se recomienda consultar [6].

Aunque la formulación matemática de (SP) es de carácter general, resulta fundamental acompañarla de un ejemplo práctico que ilustre con claridad su utilidad. En el caso que se presenta a continuación, se muestra cómo un modelo de programación estocástica permite abordar un problema inicialmente intratable debido a la incertidumbre ξ , reformulándolo como un problema del tipo (LP).

Ejemplo 3. *En este ejemplo vamos a ayudar a Paco, dueño de un restaurante famoso de Málaga, a maximizar las ganancias obtenidas con la venta de su plato estrella: la ensaladilla rusa. Para ello, debe decidir cuánta cantidad preparar para el fin de semana. Sean x_1 , x_2 las raciones de ensaladilla rusa que Paco planea preparar el viernes por la mañana y el sábado por la mañana, respectivamente.*

Sabemos que el viernes hay una alta demanda de ensaladilla rusa, por lo que toda la cantidad que Paco prepare se venderá, obteniendo una ganancia de 25€ por kg. No obstante, el sábado existe incertidumbre en la demanda, debido a la meteorología, y Paco podría incurrir en pérdidas. Más concretamente, si llueve, la mayoría de ensaladilla que se prepara no se vende, generando una pérdida de 23€ por kg, mientras que si hace sol, se vende toda la cantidad preparada, generando una ganancia de 27€ por kg. Sea p la probabilidad de que haga sol el sábado. Entonces, si ξ es la variable aleatoria que modela la meteorología del sábado, tenemos que $\xi \sim Be(p)$, y la utilidad por kg el sábado se modela como:

$$c(\xi) = \begin{cases} 23 & \text{si } \xi = 0 \text{ (lluvia),} \\ -27 & \text{si } \xi = 1 \text{ (sol).} \end{cases}$$

La función objetivo del problema de optimización a resolver para ayudar a Paco queda formulada como:

$$-25x_1 + c(\xi)x_2.$$

Además, Paco se enfrenta a varias restricciones operativas. En particular, dispone de patatas para preparar como máximo 300kg de ensaladilla rusa. Asimismo, cuenta con personal para elaborar, como mucho, 50kg el viernes y 40kg el sábado. Teniendo en cuenta estas limitaciones, el modelo estocástico se expresa como:

$$\begin{aligned} \underset{x_1, x_2}{\text{mín}} \quad & -25x_1 + c(\xi)x_2 \\ \text{s. a.} \quad & x_1 + x_2 \leq 300, \\ & x_1 \leq 50 \\ & x_2 \leq 40 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Dado que no conocemos el valor exacto de la función objetivo (al depender de ξ), el problema no puede resolverse de forma determinista. Sin embargo, sí conocemos la distribución de ξ , y si Paco desea obtener el mejor valor esperado, entonces el problema se formula como un problema del tipo (SP):

$$\begin{aligned} \underset{x_1, x_2}{\text{mín}} \quad & \mathbb{E}[-25x_1 + c(\xi)x_2] \\ \text{s. a.} \quad & x_1 + x_2 \leq 300 \\ & x_1 \leq 50 \\ & x_2 \leq 40 \\ & x_1, x_2 \geq 0 \end{aligned}$$

Supongamos que la probabilidad de que haga sol es de 0.4, es decir, $p = 0.4$. Entonces, $\mathbb{E}[c(\xi)] = (1-p) \cdot 23 + p \cdot (-27) = 0.6 \cdot 23 + 0.4 \cdot (-27) = 13.8 - 10.8 = 3$, y el problema estocástico original se reduce a un problema de programación lineal determinista de la forma (LP):

$$\begin{aligned} \underset{x_1, x_2}{\text{mín}} \quad & -25x_1 + 3x_2 \\ \text{s. a.} \quad & x_1 + x_2 \leq 300 \\ & x_1 \leq 50 \\ & x_2 \leq 40 \\ & x_1, x_2 \geq 0. \end{aligned} \tag{1.1}$$

La solución óptima se alcanza en $x_1^* = 50$, $x_2^* = 0$, es decir, Paco decide producir únicamente el viernes, agotando la capacidad disponible de personal ese día y evitando riesgos asociados a la incertidumbre meteorológica del sábado. El valor óptimo de la función objetivo es $-25 \cdot 50 + 3 \cdot 0 = -1250$, que se traduce en una ganancia de 1250€

En conclusión, este ejemplo ilustra cómo, gracias a la programación estocástica, podemos resolver problemas que antes resultaban intratables debido a la incertidumbre.

Una vez introducida la teoría esencial de la programación estocástica y presentado un ejemplo práctico de su implementación, conviene señalar que su aplicación directa en entornos reales presenta tres debilidades fundamentales:

1. **Complejidad computacional.** Calcular la esperanza exacta (a menudo expresada como una integral multidimensional) suele ser intratable en la práctica.
2. **Falta de realismo temporal.** En la mayoría de los problemas reales, las decisiones no se toman íntegramente *antes* de que se revele la incertidumbre, sino que pueden posponerse y ajustarse conforme se dispone de nueva información.
3. **Insuficiente aversión al riesgo.** Minimizar la esperanza del coste o maximizar la ganancia esperada no tiene en cuenta la posibilidad de eventos extremos, lo que puede conducir a soluciones poco robustas frente a escenarios adversos.

En las secciones y capítulos siguientes analizaremos cada una de estas limitaciones y presentaremos las extensiones más habituales en la literatura que permiten abordarlas. En concreto:

- A continuación, se presenta la técnica de Aproximación por Promedio de Muestras, ampliamente utilizada para aproximar la esperanza mediante la media muestral.
- Posteriormente, se estudia la *programación estocástica en dos etapas y multietapa*, que introduce decisiones adaptativas y permite resolver la falta de realismo temporal.
- Finalmente, en el Capítulo 2 se analiza la *optimización robusta*, una metodología que incorpora explícitamente los eventos extremos en el proceso de decisión.

Comenzamos, por tanto, con el estudio de la Aproximación por Promedio de Muestras.

1. Aproximación por Promedio de Muestras (SAA)

La “Aproximación por Promedio de Muestras” (*Sample Average Approximation (SAA)* en inglés), es un método que aproxima la esperanza matemática de la función objetivo mediante su media muestral. Más formalmente, el método reemplaza el problema estocástico original (SP) por el siguiente problema determinístico:

$$\begin{aligned} \min_x \quad & \frac{1}{N} \sum_{\ell=1}^N f(x, \xi_\ell) \\ \text{s. a.} \quad & Ax \leq b \\ & x \geq 0 \end{aligned} \tag{SAA}$$

donde cada ξ_ℓ representa una muestra independiente del vector aleatorio ξ , que, recordemos, modela la incertidumbre inherente en el problema. A medida que N crece, la media muestral tiende a aproximarse a la verdadera esperanza de la función objetivo, ofreciendo una representación cada vez más precisa del problema original. Esta propiedad es clave, pues garantiza que, bajo condiciones razonables [6], la sucesión de soluciones x_N^* , donde cada x_N^* es la solución óptima del problema (SAA) con N muestras, converge con alta probabilidad a una solución óptima del problema verdadero x^* cuando $N \rightarrow \infty$.

A continuación ilustramos el uso del método SAA y su comportamiento de convergencia mediante un ejemplo, [13]. Se ha elegido una versión del clásico problema del vendedor de periódicos (también conocido como *newsvendor problem*, por su nombre en inglés), [14, 15], cuya solución teórica es conocida, lo que permite comparar los resultados del estimador SAA con el valor óptimo real. Para facilitar la representación gráfica, consideraremos una única variable de decisión x .

Ejemplo 4. *En este ejemplo vamos a ayudar a Paco a determinar la cantidad x de postres a preparar para maximizar las ganancias esperadas. Cada unidad vendida genera un ingreso de $r = 8$ €, mientras que el coste de producción por unidad es de $k = 5$ €. Si sobra producto, este puede venderse a una cafetería vecina por un precio de liquidación de $v = 4$ €. Pero de nuevo, Paco se enfrenta a incertidumbre en la demanda ξ , que sabemos que sigue una distribución exponencial de media 10, es decir, $\xi \sim \text{Exp}(0.1)$. Dado esto, el modelo que se busca resolver para ayudar a Paco es:*

$$\min_{x \in X} \mathbb{E} [kx - r \cdot \min\{x, \xi\} - v \cdot \max\{x - \xi, 0\}] \quad (1.2)$$

donde $X = \mathbb{R}^+$. Como hemos comentado, este problema es una versión del clásico problema del *newsvendor*, cuya solución teórica x^* se sabe que ha de cumplir, [13]:

$$F(x^*) = \frac{r - k}{r - v}$$

donde F es la función de distribución de ξ . Aplicando esta expresión a nuestro caso, se obtiene que:

$$F(x^*) = \frac{8 - 5}{8 - 4} = \frac{3}{4} \Rightarrow x^* = F^{-1}(0.75).$$

En este caso, como $\xi \sim \text{Exp}(0.1)$, entonces

$$F(x^*) = 1 - e^{-\lambda x^*} = 1 - e^{-x^*/10}.$$

Por tanto:

$$x^* = -\frac{1}{\lambda} \log(1 - 0.75) = -10 \cdot \log(0.25) \approx 13.86.$$

Por tanto, el número óptimo de postres a preparar es aproximadamente 14.

Veamos cuál es la solución que obtenemos si aplicamos el método SAA al problema (1.2). Para ello, generamos una muestra de tamaño N de la distribución $\text{Exp}(0.1)$, denotada por $\{\xi_1, \xi_2, \dots, \xi_N\}$. Queremos ver que la sucesión de soluciones óptimas de los problemas aproximados

$$\min_{x \in X} \frac{1}{N} \sum_{\ell=1}^N [kx - r \cdot \min\{x, \xi_\ell\} - v \cdot \max\{x - \xi_\ell, 0\}] \quad (1.3)$$

converge hacia la solución x^* obtenida anteriormente.

Como se observa en la Figura 3, donde se representan, mediante líneas discontinuas de distintos colores, las funciones empíricas

$$\bar{f}_N(x, \boldsymbol{\xi}) = \frac{1}{N} \sum_{\ell=1}^N [kx - r \cdot \min\{x, \xi_\ell\} - v \cdot \max\{x - \xi_\ell, 0\}], \quad x \in [0, 40] \quad (1.4)$$

correspondientes a diferentes tamaños de muestra N , $N \in \{10^i, i = 1, \dots, 5\}$, y mediante una línea continua roja la esperanza de la función objetivo verdadera:

$$\mathbb{E}[f(x, \xi)] := \mathbb{E}[kx - r \cdot \min(x, \xi) - v \cdot \max(x - \xi, 0)], \quad (1.5)$$

vemos que las funciones empíricas (1.4) convergen puntualmente hacia la función verdadera (1.5) a medida que $N \rightarrow \infty$. Para calcular la expresión (1.5), se ha usado que, para cada $x \geq 0$ se verifica que, [16]:

$$\mathbb{E}[\min(\xi, x)] = \int_0^x [1 - F_\xi(t)] dt,$$

donde F_ξ es la función de distribución de ξ . Si $\xi \sim \text{Exp}(\theta)$ con $\theta = 0.1$, entonces:

$$\begin{aligned} \mathbb{E}[\min(\xi, x)] &= \int_0^x e^{-t/\theta} dt = [-\theta e^{-t/\theta}]_{t=0}^{t=x} \\ &= \theta (1 - e^{-x/\theta}) = 10 (1 - e^{-x/10}) \end{aligned}$$

Además, se ha usado la identidad:

$$\max\{x - \xi, 0\} = x - \min\{x, \xi\},$$

que se comprueba distinguiendo los casos $\xi \leq x$ y $\xi > x$:

- Si $\xi \leq x$, entonces $\min\{x, \xi\} = \xi$, y:

$$x - \min\{x, \xi\} = x - \xi = \max\{x - \xi, 0\}.$$

- Si $\xi > x$, entonces $\min\{x, \xi\} = x$, y:

$$x - \min\{x, \xi\} = x - x = 0 = \max\{x - \xi, 0\}.$$

Sustituyendo en (1.5) los valores $k = 5$, $r = 8$ y $v = 4$ obtenemos que

$$\begin{aligned} \mathbb{E}[f(x, \xi)] &= 5x - 8 \cdot 10 (1 - e^{-x/10}) - 4 (x - 10 (1 - e^{-x/10})) \\ &= x - 40 (1 - e^{-x/10}) \end{aligned} \quad (1.6)$$

Además, en la Figura 4 se ha representado, en el eje de abscisas el tamaño de la muestra y en el eje de ordenadas la solución del problema aproximado (1.3). Se observa también que la sucesión de soluciones óptimas obtenidas mediante el método SAA, x_N^* , converge efectivamente a la solución teórica x^* del problema original.

A pesar de su simplicidad, el método SAA es una herramienta ampliamente utilizada en programación estocástica, debido a sus buenas propiedades teóricas, cuando $N \rightarrow \infty$.

El principal inconveniente aparece cuando el tamaño de la muestra N es pequeño, ya que en ese caso, la calidad de la aproximación se ve comprometida, lo que introduce varias desventajas.

Una de las más destacadas es la *maldición del optimizador* (*optimizer's curse* en inglés), [17], que se refiere al hecho de que, debido a que empleamos una muestra finita, el modelo puede sesgarse hacia la representación de dicha muestra en lugar de capturar la estructura subyacente de la distribución verdadera de ξ . Esto genera una tendencia hacia soluciones subóptimas cuando las comparamos con la solución verdadera.

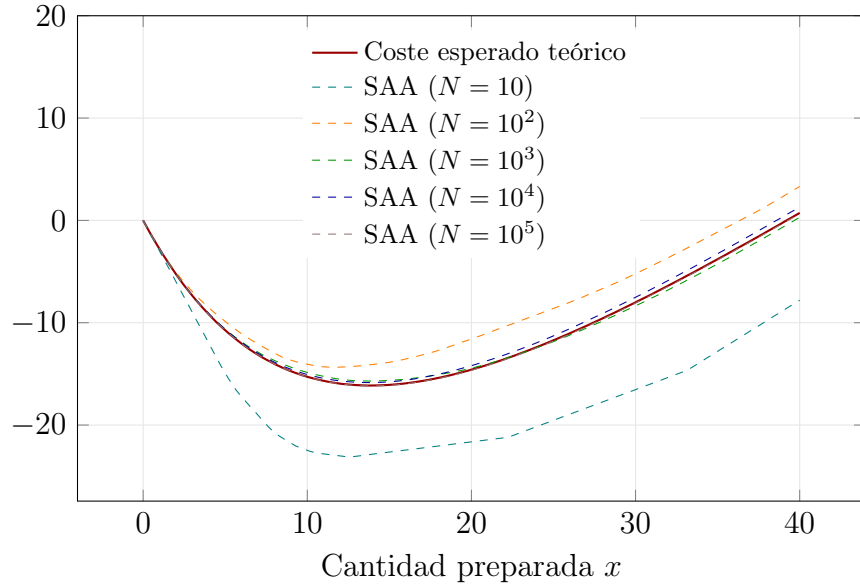


Figura 3: Aproximación de la función $\mathbb{E}[f(x, \xi)]$ mediante SAA para distintos tamaños de muestra.

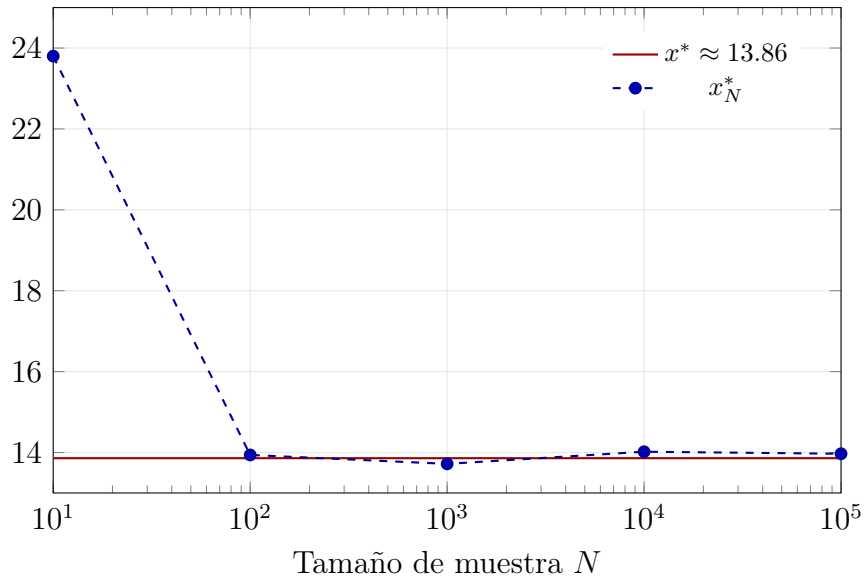


Figura 4: Convergencia de x_N^* hacia x^* al aumentar el tamaño de la muestra.

Veamos esto con más detalle. Supongamos que \mathbb{P} es la verdadera distribución generadora de datos y que ξ_ℓ son muestras independientes obtenidas de \mathbb{P} , es decir, $\xi_\ell \sim \mathbb{P}$ donde $\ell \in \{1, 2, \dots, N\}$. El valor objetivo resultante de un problema de optimización empírica (SAA) se define como:

$$Z_N(\xi_1, \dots, \xi_N) := \min_{x \in X} \frac{1}{N} \sum_{\ell=1}^N f(x, \xi_\ell) = \min_{x \in X} \bar{f}_N(x, \xi) \quad (1.7)$$

donde $\xi = (\xi_1, \dots, \xi_N)$, X es el conjunto de posibles soluciones y $f(x, \xi_\ell)$ es la función objetivo asociada con la solución x y a un elemento de la muestra $\{\xi_\ell\}_{\ell=1}^N$. Denotemos por \mathbb{P}^N la distribución conjunta de las muestras $\xi_1, \xi_2, \dots, \xi_N$. Entonces, para cualquier

$x \in X$, se cumple, [18]:

$$\begin{aligned} \mathbb{E}_{\mathbb{P}^N} [Z_N(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N)] &= \mathbb{E}_{\mathbb{P}^N} \left[\min_{x \in X} \frac{1}{N} \sum_{\ell=1}^N f(x, \boldsymbol{\xi}_\ell) \right] \\ &\leq \mathbb{E}_{\mathbb{P}^N} \left[\frac{1}{N} \sum_{\ell=1}^N f(x, \boldsymbol{\xi}_\ell) \right] = \frac{1}{N} \sum_{\ell=1}^N \mathbb{E}_{\mathbb{P}} [f(x, \boldsymbol{\xi}_\ell)] = \mathbb{E}_{\mathbb{P}} [f(x, \xi)] \end{aligned}$$

donde en la última igualdad desaparece el subíndice de la esperanza porque los datos $\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N$ son independientes e idénticamente distribuidos (i.i.d.).

Como se cumple que

$$\mathbb{E}_{\mathbb{P}^N} [Z_N(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N)] \leq \mathbb{E}_{\mathbb{P}} [f(x, \xi)] \quad \forall x \in X,$$

se concluye que, [18] :

$$\mathbb{E}_{\mathbb{P}^N} [Z_N(\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N)] = \mathbb{E}_{\mathbb{P}^N} \left[\min_{x \in X} \frac{1}{N} \sum_{\ell=1}^N f(x, \boldsymbol{\xi}_\ell) \right] \leq \min_{x \in X} \mathbb{E}_{\mathbb{P}} [f(x, \xi)]$$

es decir,

$$\mathbb{E}_{\mathbb{P}^N} \left[\min_{x \in X} \bar{f}_N(x, \boldsymbol{\xi}) \right] \leq \min_{x \in X} \mathbb{E}_{\mathbb{P}} [f(x, \xi)] \quad (\text{OC}_{SAA})$$

En este contexto, se concluye que la esperanza del valor objetivo empírico es siempre menor o igual que la esperanza del valor verdadero de la función objetivo: esta es la llamada *maldición del optimizador*.

Nota: En la Figura 3 puede observarse que, en algunos puntos de algunas muestras de tamaño N , la función empírica se sitúa por encima de la verdadera, mientras que para otras se encuentra por debajo. Esto no contradice la desigualdad (OC_{SAA}), ya que dicha desigualdad indica que, *en promedio*, el valor óptimo estimado mediante (SAA) tiende a subestimar el valor óptimo real. Es decir, se cumple que:

$$\forall N \in \mathbb{N}, \quad \mathbb{E} \left[\min_{x \in X} \bar{f}_N(x, \boldsymbol{\xi}) \right] \leq \min_{x \in X} \mathbb{E}[f(x, \xi)].$$

Pero no se cumple que:

$$\forall x \in X, \quad \mathbb{E}[\bar{f}_N(x, \boldsymbol{\xi})] \leq \mathbb{E}[f(x, \xi)].$$

Para ilustrar este fenómeno, retomamos el Ejemplo 4. Se ha realizado un experimento numérico considerando distintos tamaños de muestra $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N\}$, $N \in \{10, 30, 100, 300, 10^3\}$ (no se ha llegado hasta $N = 10^5$ como en experimentos previos por limitaciones computacionales). Fijado un tamaño de muestra N , se generan $M = 500$ réplicas independientes, esto es, 500 conjuntos distintos de N muestras i.i.d. de ξ . Sobre cada conjunto se resuelve el problema de optimización (1.3), obteniéndose en total, para dicho N , 500 valores óptimos aproximados.

$$Z_N^j := \min_{x \geq 0} \frac{1}{N} \sum_{\ell=1}^N [kx - r \cdot \min\{x, \boldsymbol{\xi}_\ell^j\} - v \cdot \max\{x - \boldsymbol{\xi}_\ell^j, 0\}], \quad j = 1, \dots, M.$$

En la Figura 5, los puntos azules representan la media

$$\bar{Z}_N = \frac{1}{M} \sum_{j=1}^M Z_N^j$$

para cada N , mientras que las barras verticales azules indican la desviación típica:

$$\sigma_N = \sqrt{\frac{1}{M} \sum_{j=1}^M (Z_N^j - \bar{Z}_N)^2}.$$

La línea roja muestra el valor esperado teórico:

$$\begin{aligned} \mathbb{E}[Z^*] &= \mathbb{E}[kx^* - r \min\{x^*, \xi\} - v \max\{x^* - \xi, 0\}] \\ &= x^* - 40(1 - e^{-x^*/10}) = 13.86 - 40 + 40e^{-13.86/10} \approx -16.14 \end{aligned}$$

donde $x^* = 13.86$ es el número óptimo de postres a preparar:

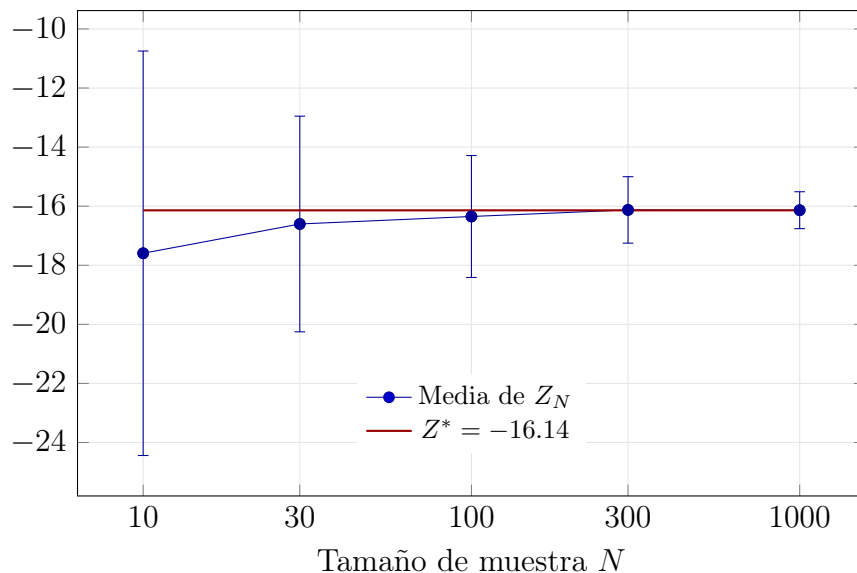


Figura 5: Media y desviación típica de 500 réplicas de Z_N para distintos N frente al valor esperado teórico Z^* .

Se observa que, para tamaños de muestra pequeños, \bar{Z}_N se sitúa por debajo del valor verdadero $\mathbb{E}[Z^*]$ lo que refleja la presencia de la *maldición del optimizador*. No obstante, a partir de $N = 300$, la diferencia entre \bar{Z}_N y el valor teórico esperado se vuelve despreciable, indicando la convergencia del estimador.

Pese a este sesgo inicial, el enfoque SAA constituye una herramienta de gran utilidad ya que permite aproximar problemas de optimización estocástica a partir de datos muestrales, sin requerir conocimiento explícito de la distribución subyacente. Su simplicidad y la posibilidad de aplicar técnicas deterministas lo convierten en una metodología ampliamente utilizada en la práctica.

A continuación, abordamos el caso en que las decisiones se toman de forma secuencial en el tiempo, adaptándose a la información que se va revelando. Para ello, extendemos el marco de la programación estocástica hacia formulaciones multietapa, que modelan explícitamente la dinámica temporal de la incertidumbre y la estructura recursiva de las decisiones.

2. Programación estocástica con decisiones en el tiempo

Para superar la segunda limitación del modelo (SP), introducimos la **programación estocástica multietapa**, un marco en el que el proceso de decisión se estructura a lo largo de varios instantes temporales, y cada acción puede tomarse en función de la información disponible hasta ese momento.

Comenzaremos con el caso de dos etapas, presentado en la Sección 2.1, suficiente para ilustrar la lógica del recurso y los elementos fundamentales del modelo. Posteriormente, en la Sección 2.2 generalizaremos el esquema a un número arbitrario de etapas, reflejando mejor la naturaleza secuencial de muchos problemas reales.

2.1. Programación estocástica de dos etapas

La programación estocástica de dos etapas se aplica a problemas en los que las decisiones deben tomarse en dos momentos distintos: una primera decisión x^1 *here-and-now*, o, de *aquí y ahora* que se adopta antes de conocer la incertidumbre, y una segunda decisión $x^2(\xi)$ *wait-and-see*, o, de *espera y observación* que se toma una vez observada la realización de un vector aleatorio ξ . Como se indicó al comienzo del capítulo, dicho vector modeliza la incertidumbre presente en el problema. En la literatura es habitual emplear la misma notación ξ tanto para el vector aleatorio como para una de sus realizaciones concretas (llamadas también escenarios), siendo el significado preciso deducible a partir del contexto. En este trabajo se adoptará esa misma convención notacional.

La diferencia esencial entre ambos tipos de decisiones radica en el momento en que se toman. El proceso temporal puede resumirse en la siguiente secuencia:

1. Se adopta la decisión de primera etapa x^1 ;
2. Se revela la incertidumbre ξ ;
3. Se adopta la decisión de segunda etapa $x^2(\xi)$.

El esquema de decisiones descrito anteriormente puede visualizarse de manera intuitiva mediante un *árbol de escenarios*, [19], como el que se muestra en la Figura 6, donde se representa un proceso de decisión con dos posibles realizaciones del vector aleatorio ξ : ξ_1 y ξ_2 .

Gráficamente, un árbol de escenarios está compuesto por un conjunto de nodos y ramas. Los nodos representan los distintos estados del problema en momentos concretos, es decir, los puntos en los que se toman decisiones. Cada nodo tiene un único predecesor y puede tener varios sucesores. El primer nodo se denomina *raíz* y corresponde al inicio del horizonte de planificación. En este nodo se adoptan las decisiones de primera etapa x^1 .

Los nodos conectados directamente con la raíz son los nodos de segunda etapa, que representan los puntos en los que se adoptan las decisiones *wait-and-see* o de *espera y observación*. En un problema de dos etapas, el número de nodos de segunda etapa coincide con el número de escenarios considerados y reciben el nombre de *hojas*. Las ramas del árbol representan las distintas posibles realizaciones del vector aleatorio ξ , capturando así la evolución de la incertidumbre a lo largo del proceso de decisión.

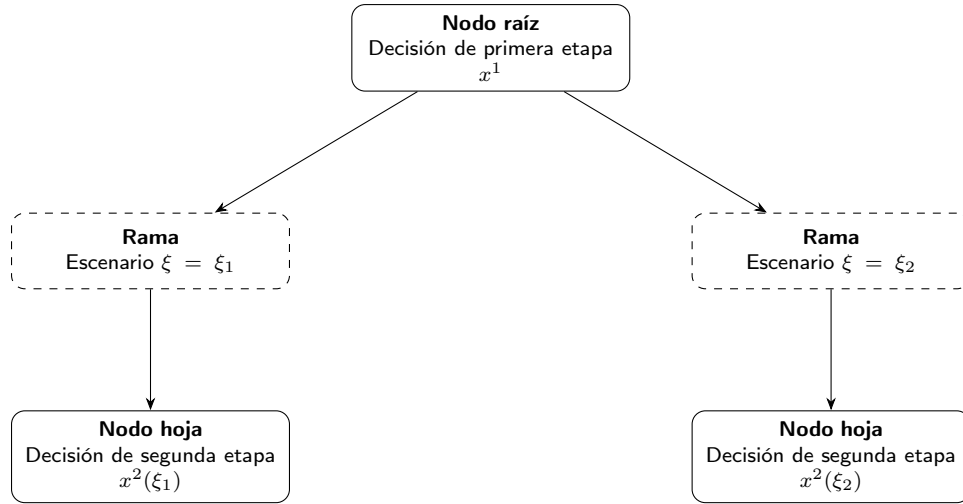


Figura 6: Representación vertical de un árbol de escenarios con dos etapas y dos posibles realizaciones para la incertidumbre ξ .

La formulación estándar de este modelo viene dada por (SP2). Nótese que hemos eliminado la notación de traspuestas para simplificar notación.

$$\begin{aligned}
 \min_{x^1} \quad & c^1 x^1 + \mathbb{E}_\xi [Q(x^1, \xi)] \\
 \text{s.a.} \quad & A^0 x^1 \leq b^0 \\
 & x^1 \geq 0
 \end{aligned} \tag{SP2}$$

donde $Q(x^1, \xi)$ se define $\forall \xi \in \Xi$ como:

$$\begin{aligned}
 Q(x^1, \xi) = \min_{x^2} \quad & c^2(\xi) x^2 \\
 \text{s.a.} \quad & A_1^1(\xi) x^1 + A_2^1(\xi) x^2 \leq b^1(\xi) \\
 & x^2 \geq 0
 \end{aligned} \tag{SP2_Q}$$

En esta formulación, $c^1 x^1$ representa el coste determinista asociado a las decisiones de primera etapa. El término $\mathbb{E}_\xi [Q(x^1, \xi)]$ corresponde al valor esperado del coste adicional que implica adaptarse a cada posible escenario ξ mediante las decisiones de segunda etapa. La función $Q(x^1, \xi)$ incluye este coste de recurso, resolviendo un problema paramétrico condicionado tanto al escenario como a la decisión previa x^1 .

Las restricciones deterministas de primera etapa se expresan como $A^0 x^1 \leq b^0$, donde $A^0 \in \mathbb{R}^{m_1 \times n_1}$ y $b^0 \in \mathbb{R}^{m_1}$, siendo n_1 el número de variables de primera etapa y m_1 el número de restricciones en la primera etapa. Se ha utilizado deliberadamente la notación A^0 y b^0 en lugar de A y b para enfatizar que en este contexto estas matrices hacen referencia *exclusivamente* a restricciones que solo involucran a las variables de primera etapa x^1 . Además, no se ha escrito directamente A^1 y b^1 por coherencia con la literatura, donde el superíndice 1 suele reservarse para las restricciones de segunda etapa. En este trabajo se ha optado por mantener la letra A para las matrices de restricciones y b para los términos independientes, por su claridad semántica. Nótese que estas restricciones son las únicas del modelo que no dependen de la incertidumbre.

Por otro lado, las restricciones de segunda etapa se dan por $A_1^1(\xi) x^1 + A_2^1(\xi) x^2(\xi) \leq b^1(\xi)$, con $A_1^1(\xi) \in \mathbb{R}^{m_2 \times n_1}$, $A_2^1(\xi) \in \mathbb{R}^{m_2 \times n_2}$, $b^1(\xi) \in \mathbb{R}^{m_2}$ y $x^2(\xi) \in \mathbb{R}^{n_2}$, donde n_2 es el número de variables de segunda etapa y m_2 el número de restricciones de segunda etapa.

Estas restricciones enlazan ambas etapas y dependen directamente de la realización del vector aleatorio ξ , así como de la decisión de primera etapa, x^1 .

Cuando el conjunto de escenarios Ξ es discreto y finito, el modelo estocástico de dos etapas puede reformularse como un problema determinista de gran escala, conocido como *Deterministic Equivalent Formulation* (DEF). Esta reformulación permite escribir de forma explícita la esperanza matemática como una suma finita ponderada de los costes, afectando tanto a la función objetivo como a las restricciones.

Para ello, se introduce una copia independiente de las variables de segunda etapa para cada posible escenario ξ_k , con $k = 1, \dots, K$, denotadas por $x^2(\xi_k) = x_k^2$. La probabilidad de cada escenario se representa como $\pi_k = \mathbb{P}(\xi = \xi_k)$.

De esta forma, la función objetivo se obtiene sustituyendo la esperanza por la suma ponderada de costes sobre cada escenario:

$$c^1 x^1 + \sum_{k=1}^K \pi_k \cdot c^2(\xi_k) x_k^2.$$

Las restricciones de segunda etapa también se particularizan por escenario y están acopladas únicamente a través de x^1 , que debe ser idéntica en todos ellos al haberse fijado antes de la observación de ξ :

$$A_1^1(\xi_k) x^1 + A_2^1(\xi_k) x_k^2 \leq b^1(\xi_k), \quad \forall k = 1, \dots, K.$$

A partir de esta estructura, se obtiene el siguiente modelo determinista equivalente:

$$\begin{aligned} \min_{x^1, x_1^2, \dots, x_K^2} \quad & c^1 x^1 + \sum_{k=1}^K \pi_k \cdot c^2(\xi_k) x_k^2 \\ \text{s. a.} \quad & A^0 x^1 \leq b^0 \\ & A_1^1(\xi_k) x^1 + A_2^1(\xi_k) x_k^2 \leq b^1(\xi_k), \quad \forall k = 1, \dots, K \\ & x^1 \geq 0, \quad x_k^2 \geq 0, \quad \forall k = 1, \dots, K \end{aligned} \tag{SP2}_{DEF}$$

Al igual que en la formulación general, el objetivo es encontrar una política de decisión que minimice el coste total esperado, asegurando la factibilidad tanto antes como después de la revelación de la incertidumbre. Nótese que en este caso se minimizan explícitamente las variables x^1 y x_k^2 , a diferencia del modelo (SP2), donde únicamente se optimiza de forma explícita la variable x^1 .

Llegados a este punto, cabe preguntarse cómo se resolvería el modelo (SP2) si ξ tiene una distribución continua, un soporte infinito, o ambas cosas. En estos casos, lo que se suele hacer en la práctica es recurrir al método de aproximación por muestreo (SAA), que ya presentamos en la sección anterior para el caso de una sola etapa.

El procedimiento para aplicar SAA guarda similitudes con el seguido para obtener la formulación (SP2)_{DEF}, con la salvedad de que, en lugar de considerar el conjunto completo de posibles realizaciones de ξ , se trabaja con una muestra finita extraída de su distribución, $\boldsymbol{\xi} = \{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_N\}$.

En primer lugar, para cada $\boldsymbol{\xi}_\ell$, con $\ell \in \{1, \dots, N\}$, se introduce una copia independiente de la variable de segunda etapa, $x^2(\boldsymbol{\xi}_\ell)$. En este caso mantenemos explícita la dependencia de x^2 respecto a $\boldsymbol{\xi}_\ell$ para diferenciarla con el modelo (SP2)_{DEF} y enfatizar que se trata de una realización concreta de una muestra de ξ .

En segundo lugar, sustituimos la esperanza $\mathbb{E}[\mathcal{Q}(x^1, \xi)]$ por su versión empírica:

$$\mathbb{E}[Q(x^1, \xi)] = \frac{1}{N} \sum_{\ell=1}^N c^2(\xi_\ell)' x^2(\xi_\ell),$$

y, particularizando las restricciones para cada muestra $\xi_\ell \in \xi$, el modelo (SP2) se aproxima mediante el siguiente problema determinista:

$$\begin{aligned} \min_{x^1, x^2(\xi_1), \dots, x^2(\xi_N)} \quad & c^1 x^1 + \sum_{\ell=1}^N \frac{1}{N} \cdot c^2(\xi_\ell) x^2(\xi_\ell) \\ \text{s. a.} \quad & A^0 x^1 \leq b^0 \\ & A_1^1(\xi_\ell) x^1 + A_2^1(\xi_\ell) x^2(\xi_\ell) \leq b^1(\xi_\ell), \quad \forall \ell = 1, \dots, N \\ & x^1 \geq 0, \quad x^2(\xi_\ell) \geq 0, \quad \forall \ell = 1, \dots, N \end{aligned} \tag{SP2}_{SAA}$$

Así, mediante SAA podemos tratar (SP2) también cuando la incertidumbre sigue una variable aleatoria continua o tiene soporte infinito.

Una vez introducida la formulación general del modelo estocástico junto con dos variantes clave para su resolución (la formulación empírica y la aproximación por muestreo), presentamos una variante del Ejemplo 3, con los siguientes objetivos:

- Mostrar cómo se aplica en la práctica un modelo de programación estocástica en dos etapas.
- Comparar esta aproximación con la resolución del problema suponiendo únicamente uno de los escenarios posibles (enfoque determinista ingenuo).
- Compararla también con la resolución mediante una programación estocástica estándar (de una sola etapa), que optimiza la función objetivo esperada sin distinguir decisiones por escenario, vista en la Sección 1.

Ejemplo 5. Como hemos comentado antes, retomemos el Ejemplo 3, pero esta vez, renombramos las variables de decisión x_1 y x_2 por x^1 y x^2 . El problema de optimización a resolver para ayudar a Paco se expresa como:

$$\begin{aligned} \min_{x^1, x^2} \quad & -25x^1 + c(\xi) \cdot x^2 \\ \text{s. a.} \quad & x^1 + x^2 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2 \geq 0 \end{aligned} \tag{1.8}$$

donde $\xi \sim Be(p)$ con $p = 0.4$, y

$$c(\xi) = \begin{cases} 23 & \text{si } \xi = 0 \quad (\text{lluvia}), \\ -27 & \text{si } \xi = 1 \quad (\text{sol}). \end{cases}$$

Asumamos también que la actitud de Paco es optimista y supone que siempre se da el mejor caso posible (es decir, siempre hace sol ($\xi = 1$), y por tanto $c(\xi) = -27$). Entonces,

ayudar a Paco es equivalente a resolver el problema de tipo (LP):

$$\begin{aligned} \min_{x^1, x^2} \quad & -25x^1 - 27x^2 \\ \text{s.a.} \quad & x^1 + x^2 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2 \geq 0 \end{aligned}$$

La solución óptima es $x^{*1} = 50$, $x^{*2} = 40$, con un valor objetivo de -2330 . Sin embargo, si esta solución se aplica sistemáticamente en todos los escenarios, el valor esperado real que se obtiene es:

$$\begin{aligned} c_{\text{real}} &= -25 \cdot x^{*1} + \mathbb{P}(\xi = 0) \cdot 23 \cdot x^{*2} + \mathbb{P}(\xi = 1) \cdot (-27) \cdot x^{*2} \\ &= -25 \cdot 50 + 0.6 \cdot 23 \cdot 40 + 0.4 \cdot (-27) \cdot 40 = -1130. \end{aligned}$$

En cambio, si Paco se levanta pesimista, y aplica sistemáticamente la solución obtenida para el caso más desfavorable, que se da cuando $\xi = 0$, estaría resolviendo el siguiente problema determinista:

$$\begin{aligned} \min_{x^1, x^2} \quad & -25x^1 + 23x^2, \\ \text{s.a.} \quad & x^1 + x^2 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2 \geq 0 \end{aligned}$$

La solución óptima de este problema es $x^{*1} = 50$, $x^{*2} = 0$ con valor objetivo de -1250 . Al aplicar esta solución en el mundo real, se obtendría un valor objetivo real dado por:

$$\begin{aligned} c_{\text{real}} &= -25 \cdot x^{*1} + \mathbb{P}(\xi = 0) \cdot 23 \cdot x^{*2} + \mathbb{P}(\xi = 1) \cdot (-27) \cdot x^{*2} \\ &= -25 \cdot 50 + 0.6 \cdot 23 \cdot 0 + 0.4 \cdot (-27) \cdot 0 = -1250 \end{aligned}$$

Pasemos al caso en el que sí se tiene en cuenta la incertidumbre y Paco quiere optimizar el valor esperado, pero no se permite adaptar las decisiones a las realizaciones de la incertidumbre. Entonces, retomamos el problema (1.1) cuya solución óptima era $x^{*1} = 50$, $x^{*2} = 0$ con valor objetivo esperado de -1250 .

No obstante, si admitimos el caso en que se permite adaptar la decisión x^2 en función de la realización de ξ , es decir, primero se decide cuánta cantidad de ensaladilla rusa se va a preparar el viernes, x^1 , luego se observa la meteorología del sábado, y finalmente se decide cuánta cantidad de ensaladilla rusa, x^2 , preparar el sábado, entonces, nos encontramos ante un problema de programación estocástica de dos etapas del tipo (SP2):

$$\begin{aligned} \min_{x^1} \quad & -25x^1 + \mathbb{E}_\xi[Q(x^1, \xi)] \\ \text{s.a.} \quad & x^1 \leq 50 \\ & x^1 \geq 0 \end{aligned}$$

donde $Q(x^1, \xi)$ es, para $\xi \in \{0, 1\}$:

$$\begin{aligned}
Q(x^1, \xi) &= \min_{x^2(\xi)} c(\xi) x^2(\xi) \\
\text{s.a. } &x^1 + x^2(\xi) \leq 300 \\
&x^2(\xi) \leq 40 \\
&x^2(\xi) \geq 0
\end{aligned}$$

Como el espacio de escenarios es finito, $\Xi = \{0, 1\}$, el problema (1.9) admite una formulación determinista equivalente del tipo (SP2_{DEF}). En consecuencia, para cada posible realización de la incertidumbre, introducimos el vector completo de decisiones de ajuste, que en nuestro caso será de dimensión 2,

$$x^2 = (x_0^2, x_1^2),$$

donde identificamos $x^2(0) := x_0^2$ y $x^2(1) := x_1^2$, es decir, x_0^2 es la decisión que se toma si $\xi = 0$, y x_1^2 si $\xi = 1$.

Por tanto, la función objetivo quedaría formulada como:

$$-25x^1 + \sum_{i=0}^1 \mathbb{P}(\xi = i) c(i) x_i^2 = -25x^1 + 0.6 \cdot 23x_0^2 + 0.4 \cdot (-27x_1^2) = -25x^1 + 13.8x_0^2 - 10.8x_1^2.$$

Adaptando las restricciones del problema original para cada escenario, el nuevo problema completo se formula como:

$$\begin{aligned}
\min_{x^1, x_0^2, x_1^2} & -25x^1 + 13.8x_0^2 - 10.8x_1^2 \\
\text{s.a. } &x^1 + x_0^2 \leq 300 \\
&x^1 + x_1^2 \leq 300 \\
&x^1 \leq 50 \\
&x_0^2 \leq 40 \\
&x_1^2 \leq 40 \\
&x^1, x_0^2, x_1^2 \geq 0
\end{aligned}$$

La solución óptima es $x^{*1} = 50$, $x_0^{*2} = 0$, $x_1^{*2} = 40$, con un valor esperado de -1682 . Esta solución debe interpretarse del siguiente modo: si resolvemos el problema de partida (1.8) infinitas veces, y cada vez que $\xi = 0$ elegimos la solución (x^1, x_0^2) y cada vez que $\xi = 1$ elegimos la solución (x^1, x_1^2) , entonces obtendríamos en promedio el mejor valor objetivo posible.

Hagamos ahora una comparativa del valor objetivo obtenido utilizando la programación estocástica en dos etapas, de -1682 , frente a las otras alternativas expuestas:

- **Escenario favorable ($\xi = 1$):** si se resuelve únicamente el escenario optimista ($c_2 = -27$) y se aplica su solución en todos los escenarios, el valor esperado real obtenido es de -1250 . Es decir, considerar la programación estocástica en dos etapas frente a este enfoque aumenta la función objetivo en un

$$\frac{-1682 + 1130}{1130} \cdot 100 \approx -48.85\% \quad (\text{es decir, una disminución del } \mathbf{48.85\%}).$$

- **Escenario desfavorable** ($\xi = 0$): si se resuelve únicamente el escenario pesimista ($c_2 = 23$) y se aplica su solución en todos los escenarios, el valor esperado real es de -1250 . En este caso, vemos un aumento de la función objetivo en un:

$$\frac{-1682 + 1250}{1250} \cdot 100 \approx -34.56\% \quad (\text{es decir, una mejora del } \mathbf{34.56\%}).$$

- **Programación estocástica**: esta estrategia también conduce a un valor esperado de -1250 , por lo que la mejora respecto al modelo completo en dos etapas vuelve a ser del $\mathbf{34.56\%}$.

Nota: En este ejemplo, como la solución que elige la programación estocástica coincide con la del escenario desfavorable, el valor esperado obtenido en ambos casos es el mismo. Sin embargo, esto no ocurre en general. Por ejemplo, si se añade la restricción $x_1 + 2x_2 \geq 100$, entonces la programación estocástica selecciona la misma solución que en el escenario favorable.

En resumen, la programación estocástica en dos etapas permite incorporar explícitamente la incertidumbre y adaptar las decisiones de la segunda etapa a la realización observada del escenario. Este enfoque mejora de forma significativa los resultados frente a resolver un único escenario de manera aislada y también frente a la programación estocástica tradicional de una sola etapa, ya que aprovecha la flexibilidad de ajustar las decisiones tras revelarse la incertidumbre. Sin embargo, su alcance es aún limitado, pues solo contempla una decisión inicial y una única corrección posterior. En muchos problemas reales, la incertidumbre se revela de manera progresiva en el tiempo, lo que exige decisiones secuenciales que se adapten dinámicamente a la nueva información disponible. Para capturar esta situación más general, introducimos a continuación el marco de la programación estocástica multietapa.

2.2. Programación estocástica multietapa

En esta sección presentamos la generalización natural del modelo de dos etapas al caso en el que se consideran T etapas, es decir, la *programación estocástica multietapa*.

En este contexto, se toman decisiones en T momentos distintos, condicionadas a $T - 1$ realizaciones del vector de incertidumbre ξ . Para identificar la etapa en la que se toma cada decisión, utilizamos el superíndice: es decir, x^t , con $t = 1, \dots, T$, denota la decisión tomada en la etapa t , mientras que ξ^t , con $t = 1, \dots, T - 1$, representa la incertidumbre revelada justo después de tomar x^t .

Utilizaremos además indistintamente ξ^t para referirnos tanto a la variable aleatoria que modela la incertidumbre en la etapa t , como a su realización concreta. El significado quedará claro en cada caso por el contexto. La decisión x^t , $t \in \{2, \dots, T\}$ depende de las realizaciones observadas hasta la etapa $t - 1$, es decir, ξ^1, \dots, ξ^{t-1} , lo que formalmente se expresa como:

$$x^t : \Xi^1 \times \dots \times \Xi^{t-1} \longrightarrow X^t.$$

donde X^t denota el conjunto factible del problema en la etapa t . Para $t = 1$, el conjunto factible es determinista y viene dado por $A^0 x^1 \leq b^0$. En cambio, para $t \geq 2$, el conjunto factible X^t depende de la realización de la incertidumbre observada hasta la etapa t , es decir, de $\xi^{(t-1)}$. En nuestro caso, se define como:

$$X^t(\xi^{(t-1)}) := \left\{ x^t \in \mathbb{R}^{n_t} \mid \sum_{s=1}^t A_s^{t-1}(\xi^{(t-1)}) x^s \leq b^{t-1}(\xi^{(t-1)}), x^t \geq 0 \right\}$$

donde $A_s^{t-1}(\xi^{(t-1)}) \in \mathbb{R}^{m_t \times n_s}$ y $b^{t-1}(\xi^{(t-1)}) \in \mathbb{R}^{m_t}$ para cada $s = 1, \dots, t$, siendo m_t el número de restricciones en la etapa t y n_s el número de variables de decisión en la etapa s . Para abreviar notación antes de presentar la formulación general del problema, denotaremos por $\xi^{(t)}$ al historial observado de la incertidumbre hasta la etapa t , es decir:

$$\xi^{(t)} := (\xi^1, \dots, \xi^t).$$

De forma análoga, representaremos por $x^{(t)}$, $x^{(t)}(\xi^{(t-1)})$ o $x^{(t)}(\xi^1, \dots, \xi^{t-1})$ al vector de decisiones tomadas hasta la etapa t . Es decir, $x^{(t)} = (x^1, x^2(\xi^1), \dots, x^t(\xi^1, \dots, \xi^{t-1}))$. Nótese que, con esta notación: $x^{(1)} = x^1$ y $\xi^{(1)} = \xi^1$. En lo que respecta al proceso de decisión, este se estructura según el siguiente esquema, [19]:

- Se adopta la decisión de primera etapa x^1 .
- Se revela la incertidumbre de la primera etapa ξ^1 .
- Se toma la decisión de segunda etapa $x^2(\xi^1)$, adaptada a la realización observada.
- Se revela la incertidumbre de la segunda etapa ξ^2 .
- ⋮
- Se revela la incertidumbre de la etapa $T-1$ ξ^{T-1} .
- Finalmente, se toma la decisión de la última etapa $x^T(\xi^1, \dots, \xi^{T-1})$.

Este marco de decisiones puede representarse de forma intuitiva mediante un **árbol de escenarios**, como el mostrado en la Figura 7, que ilustra un ejemplo de problema con tres etapas, donde las variables de decisión son x^1, x^2, x^3 , todas unidimensionales.

El nodo inicial corresponde a la etapa 1, en la que se toma la decisión inicial x^1 , antes de que se revele ninguna información sobre la incertidumbre. A continuación, se observa la realización de la variable aleatoria ξ^1 , que en este ejemplo es $\xi^1 \sim \text{Be}(p_1)$, representada en el árbol mediante rectángulos discontinuos.

Cada posible realización de ξ^1 conduce a un nuevo nodo en el que se toma la decisión x^2 . Tras ello, se realiza la incertidumbre de segunda etapa $\xi^2 \sim \text{Be}(p_2)$, también representada mediante rectángulos discontinuos, y se procede a tomar la decisión final x^3 , ya en la tercera etapa.

Cada trayectoria desde la raíz hasta una hoja representa un escenario completo, es decir, una secuencia particular de realizaciones de la incertidumbre y decisiones asociadas. Este esquema se repetiría sucesivamente en problemas con un mayor número de etapas. El número total de nodos en la etapa final coincide con el número total de escenarios posibles, que en este ejemplo es $|\Xi^1| \times |\Xi^2| = 2 \cdot 2 = 4$. A estos nodos terminales se les denomina *hojas*.

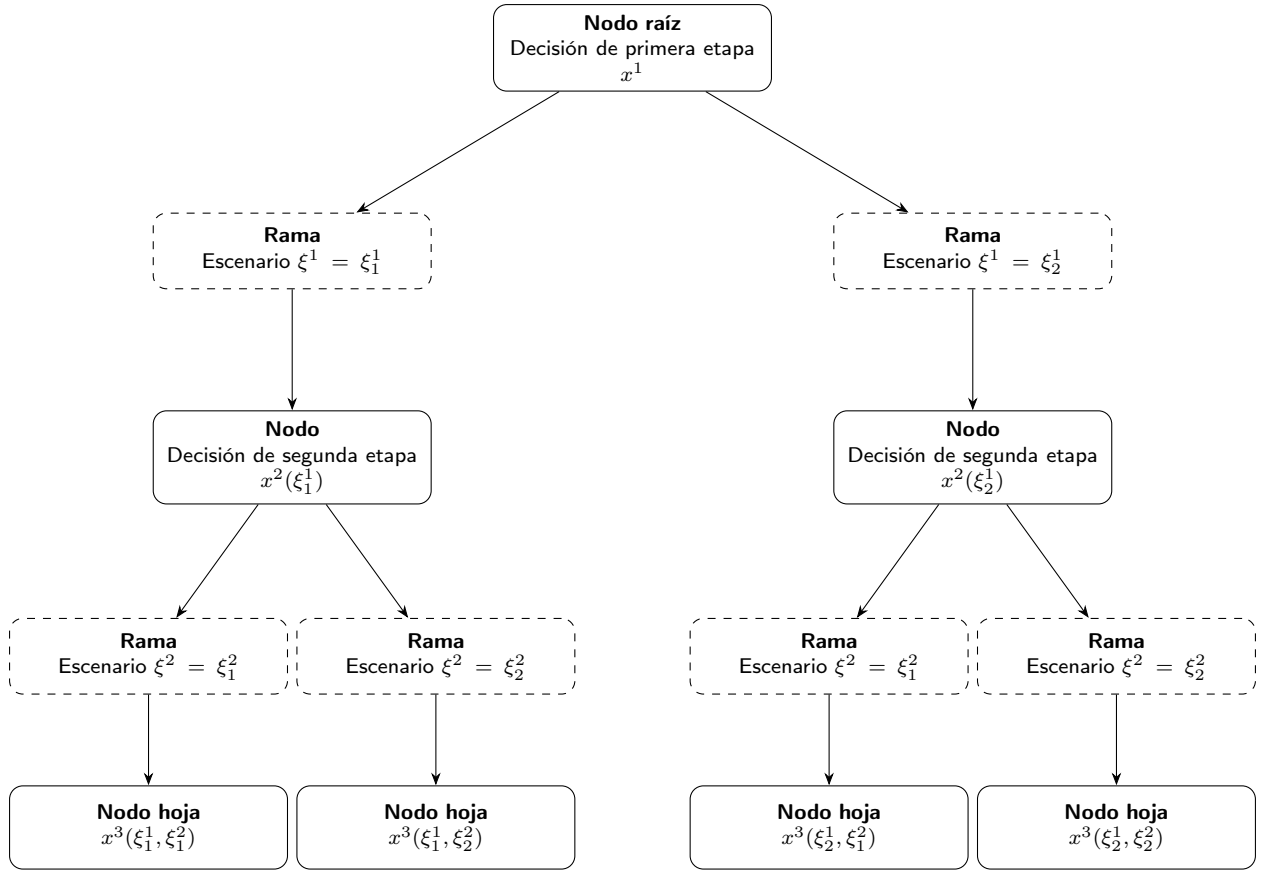


Figura 7: Representación vertical de un árbol de escenarios con tres etapas y dos posibles realizaciones por variable de incertidumbre.

A continuación presentamos la formulación del modelo. Matemáticamente, el problema multietapa se expresa como:

$$\begin{aligned}
 \min_{x^1} \quad & c^1 x^1 + \mathbb{E}_{\xi^1} [Q^1(x^1, \xi^1)] \\
 \text{s.a.} \quad & A^0 x^1 \leq b^0 \\
 & x^1 \geq 0
 \end{aligned} \tag{SPT}$$

donde $Q^1(x^1, \xi^1)$, se define para todo $\xi^1 \in \Xi^1$ como:

$$\begin{aligned}
 \min_{x^2} \quad & c^2(\xi^1) x^2 + \mathbb{E}_{\xi^2} [Q^2(x^2, \xi^2)] \\
 \text{s.a.} \quad & A_1^1(\xi^1) x^1 + A_2^1(\xi^1) x^2 \leq b^1(\xi^1) \\
 & x^2 \geq 0
 \end{aligned}$$

a su vez, la función $Q^2(x^2, \xi^2)$, se define $\forall (\xi^1, \xi^2) \in \Xi^1 \times \Xi^2$ como:

$$\begin{aligned}
 \min_{x^3} \quad & c^3(\xi^2) x^3 + \mathbb{E}_{\xi^3} [Q^3(x^3, \xi^3)] \\
 \text{s.a.} \quad & A_1^2(\xi^2) x^1 + A_2^2(\xi^2) x^2 + A_3^2(\xi^2) x^3 \leq b^2(\xi^2) \\
 & x^3 \geq 0.
 \end{aligned}$$

En general, para $t = 1, \dots, T-2$, $Q^t(x^t, \xi^t)$ se define $\forall (\xi^1, \dots, \xi^t) \in \Xi^1 \times \dots \times \Xi^t$

como:

$$\begin{aligned} \min_{x^{t+1}} \quad & c^{t+1}(\xi^{(t)}) x^{t+1} + \mathbb{E}_{\xi^{t+1}} [Q^{t+1}(x^{(t+1)}, \xi^{(t+1)})] \\ \text{s. a.} \quad & \sum_{s=1}^{t+1} A_s^t(\xi^{(t)}) x^s \leq b^t(\xi^{(t)}) \\ & x^{t+1} \geq 0 \end{aligned}$$

Finalmente, en la etapa $T-1$, la función $Q^{T-1}(x^{(T-1)}, \xi^{(T-1)})$ se define para todo $\xi^{(T-1)} \in \Xi^1 \times \dots \times \Xi^{T-1}$ como:

$$\begin{aligned} \min_{x^T} \quad & c^T(\xi^{(T-1)}) x^T \\ \text{s. a.} \quad & \sum_{s=1}^T A_s^{T-1}(\xi^{(T-1)}) x^s \leq b^{T-1}(\xi^{(T-1)}) \\ & x^T \geq 0 \end{aligned}$$

Nótese que la función de valor en la etapa $T-1$, $Q^{T-1}(x^{(T-1)}, \xi^{(T-1)})$, no lleva asociado un término de esperanza. Esto se debe a que, en dicha etapa, ya se ha revelado toda la incertidumbre del problema: el vector $\xi^{(T-1)} = (\xi^1, \dots, \xi^{T-1})$ está completamente observado, y no quedan futuras etapas en las que puedan aparecer nuevas realizaciones aleatorias. Así, la decisión de última etapa $x^T(\xi^{(T-1)})$ se toma en un entorno completamente determinista.

Al igual que en la programación estocástica de dos etapas, cuando el espacio de escenarios es **finito y discreto** también es posible reformular el modelo multietapa (SPT) como un único problema determinista de gran escala. El procedimiento es análogo: para cada $t \in \{1, \dots, T\}$ y cada posible realización $\xi_k^t \in \Xi^t$, se introduce una variable de decisión $x^t(\xi_k^t)$, la cual denotaremos por simplicidad x^t , que representa la acción tomada en la etapa t tras observar dicha realización. Denotamos por $\pi_k^t = \mathbb{P}(\xi^t = \xi_k^t)$ la probabilidad correspondiente. Sustituyendo las esperanzas anidadas por sumas ponderadas de los costes en cada etapa y particularizando las restricciones a cada escenario, se obtiene la reformulación determinista del modelo. No obstante, para simplificar la notación, en la formulación utilizaremos un escenario genérico $\xi^t \in \Xi^t$ y denotaremos su probabilidad asociada como π^t , omitiendo el subíndice que distingue cada realización individual. Con estos elementos, el modelo multietapa se expresa como:

$$\begin{aligned} \min_{x^1, \dots, x^T} \quad & c^1 x^1 + \sum_{\xi^1 \in \Xi^1} \pi^1 \left[c^2(\xi^1) x^2 + \sum_{\xi^2 \in \Xi^2} \pi^2 [c^3(\xi^2) x^3 + \dots \right. \\ & \left. + \sum_{\xi^{T-1} \in \Xi^{T-1}} \pi^{T-1} c^T(\xi^{(T-1)}) x^T \right] \\ \text{s. a.} \quad & \sum_{s=1}^t A_s^t(\xi^{(t)}) x^s \leq b^t(\xi^{(t)}), \quad \forall t = 1, \dots, T, \forall \xi^{(t)} \in \Xi^1 \times \dots \times \Xi^t \\ & x^t \geq 0, \quad \forall t = 1, \dots, T \end{aligned} \tag{SPT}_{DEF}$$

Esta formulación muestra cómo, bajo incertidumbre discreta, es posible tratar el problema multietapa como un problema determinista clásico, aunque de gran tamaño. De hecho, si denotamos por $K^t = |\Xi^t|$ el número de escenarios en la etapa t , entonces el número total de posibles trayectorias de escenarios es:

$$\prod_{j=1}^T K^j,$$

y el número total de variables del modelo:

$$\sum_{t=1}^T \left(\prod_{j=1}^t K^j \right) \cdot n_t,$$

lo que evidencia un crecimiento exponencial respecto al número de etapas.

De forma análoga al caso de dos etapas, cuando las variables aleatorias ξ^{t-1} , $t \in \{2, \dots, T\}$ tienen soporte continuo o infinito, la resolución directa del modelo multietapa (SPT) no es posible en la práctica en la mayoría de casos. En estos contextos, se recurre habitualmente al *método de aproximación empírica* (SAA). El procedimiento consiste en generar, para cada etapa t , una muestra finita $\{\xi_1^t, \dots, \xi_{N_t}^t\}$ de tamaño N_t . A cada trayectoria muestral se asocia una variable de decisión $x^t = x^t(\xi_\ell^t)$. Con ello, las esperanzas anidadas pueden aproximarse por sus contrapartidas empíricas, y, particularizando las restricciones a cada posible realización muestral, el problema multietapa se transforma en un modelo determinista aproximado con estructura explícita:

$$\begin{aligned} \min_{x^1, x^2, \dots, x^T} \quad & c^1 x^1 + \frac{1}{N_2} \sum_{\ell_2=1}^{N_2} \left[c^2(\xi_{\ell_2}^{(2)}) x^2 + \frac{1}{N_3} \sum_{\ell_3=1}^{N_3} \left[c^3(\xi_{\ell_3}^{(3)}) x^3 + \dots + \frac{1}{N_T} \sum_{\ell_T=1}^{N_T} c^T(\xi_{\ell_T}^{(T)}) x^T \right] \right] \\ \text{s.a.} \quad & A^0 x^1 \leq b^0 \\ & A_1^1(\xi_{\ell_2}^{(2)}) x^1 + A_2^1(\xi_{\ell_2}^{(2)}) x^2 \leq b^1(\xi_{\ell_2}^{(2)}), \quad \forall \ell_2 \in \mathcal{I}_2 \\ & A_1^2(\xi_{\ell_3}^{(3)}) x^1 + A_2^2(\xi_{\ell_3}^{(3)}) x^2 + A_3^2(\xi_{\ell_3}^{(3)}) x^3 \leq b^2(\xi_{\ell_3}^{(3)}), \quad \forall (\ell_2, \ell_3) \in \mathcal{I}_3 \\ & \vdots \\ & \sum_{j=1}^T A_j^{T-1}(\xi_{\ell_T}^{(T)}) x^j \leq b^{T-1}(\xi_{\ell_T}^{(T)}), \quad \forall (\ell_2, \dots, \ell_T) \in \mathcal{I}_T \\ & x^1, x^t \geq 0, \quad t = 2, \dots, T \end{aligned} \tag{SPT}_{SAA}$$

donde

$$\mathcal{I}_t := \prod_{s=2}^t \{1, \dots, N_s\}, \quad t = 2, \dots, T,$$

representa el conjunto de ramas posibles hasta la etapa t del árbol de escenarios, y

$$\xi_{\ell_t}^{(t)} := (\xi^1, \xi_{\ell_2}^2, \dots, \xi_{\ell_t}^t),$$

es la historia de incertidumbre acumulada hasta la etapa t .

Tras presentar el marco teórico de la programación estocástica multietapa, retomamos el Ejemplo 5 con el fin de ilustrar de manera explícita la estructura de esta formulación y compararla con los modelos de una y dos etapas. Para ello, introducimos una nueva variable de tercera etapa, x^3 , que representa la cantidad de ensaladilla rusa a preparar el domingo. Esta decisión se toma tras la revelación de la incertidumbre $\xi^2 \sim Be(p_2)$, con $p_2 = 0.3$, que modela la meteorología del domingo por la mañana.

Ejemplo 6. *Supongamos que las ganancias o pérdidas de la ensaladilla rusa preparada el domingo dependen de la meteorología del sábado (modelada por la variable $\xi^1 \sim Be(p_1)$) y del domingo (modelada por la variable $\xi^2 \sim Be(p_2)$) como sigue:*

- Si llueve ambos días ($\xi^1 = 0, \xi^2 = 0$), nadie pide ensaladilla y Paco pierde 28 €/kg preparado.
- Si llueve el sábado pero no el domingo ($\xi^1 = 0, \xi^2 = 1$), toda la producción se vende con ganancia de 26 €/kg.
- Si no llueve el sábado pero llueve el domingo ($\xi^1 = 1, \xi^2 = 0$), también se vende todo, pero con ganancia de 24 €/kg.
- Si hace sol ambos días ($\xi^1 = 1, \xi^2 = 1$), la clientela prefiere ir a los chiringuitos, provocando una pérdida de 22 €/kg.

Con esta información, el coste asociado a x^3 , c^3 , se define como:

$$c^3(\xi^1, \xi^2) = \begin{cases} 28, & \text{si } \xi^1 = 0, \xi^2 = 0, \\ -26, & \text{si } \xi^1 = 0, \xi^2 = 1, \\ -24, & \text{si } \xi^1 = 1, \xi^2 = 0, \\ 22, & \text{si } \xi^1 = 1, \xi^2 = 1. \end{cases}$$

y el problema de optimización se puede expresar $\forall \xi^1 \in \{0, 1\}, \forall \xi^2 \in \{0, 1\}$ como:

$$\begin{aligned} \underset{x^1, x^2, x^3}{\text{mín}} & \quad -25x^1 + c^2(\xi^1)x^2 + c^3(\xi^1, \xi^2)x^3 \\ \text{s.a.} & \quad x^1 + x^2 + x^3 \leq 300 \\ & \quad x^1 \leq 50 \\ & \quad x^2 \leq 40 \\ & \quad x^1, x^2, x^3 \geq 0 \end{aligned} \tag{1.9}$$

Si Paco busca la solución óptima en promedio, con la flexibilidad de adaptar sus decisiones a medida que se revela la incertidumbre, nos encontramos ante un problema de programación estocástica en tres etapas, expresado como:

$$\begin{aligned} \underset{x^1}{\text{mín}} & \quad c^1x^1 + \mathbb{E}_{\xi^1} [Q(x^1, \xi^1)] \\ \text{s.a.} & \quad x^1 \leq 50 \\ & \quad x^1 \geq 0 \end{aligned} \tag{1.10}$$

donde $Q(x^1, \xi^1)$ se define para todo $\xi^1 \in \Xi^1$ como:

$$\begin{aligned} \underset{x^2}{\text{mín}} & \quad c^2(\xi^1)x^2 + \mathbb{E}_{\xi^2} [Q^2(x^1, \xi^1, x^2, \xi^2)] \\ \text{s.a.} & \quad x^2 \leq 40 \\ & \quad x^2 \geq 0 \end{aligned} \tag{1.11}$$

y $Q^2(x^1, \xi^1, x^2, \xi^2)$ se define para todo $(\xi^1, \xi^2) \in \Xi^1 \times \Xi^2$ como:

$$\begin{aligned} \underset{x^3}{\text{mín}} & \quad c^3(\xi^1, \xi^2)x^3 \\ \text{s.a.} & \quad x^1 + x^2 + x^3 \leq 300 \\ & \quad x^3 \geq 0. \end{aligned} \tag{1.12}$$

Dado que el espacio de soluciones es discreto, el problema (1.10) puede expresarse en la forma determinista equivalente (SPT_{DEF}). Para ello, introducimos variables de segunda etapa para cada realización de la incertidumbre ξ^1 : $x_0^2 := x^2(\xi^1 = 0)$ y $x_1^2 := x^2(\xi^1 = 1)$, con costes asociados $c_0^2 = c^2(\xi^1 = 0)$ y $c_1^2 = c^2(\xi^1 = 1)$, respectivamente. En la tercera etapa, introducimos las variables $x_{ij}^3 := x^3(\xi^1 = i, \xi^2 = j)$, con $i, j \in \{0, 1\}$, que representan la cantidad que se decide preparar el domingo, condicionada a la información de ambos días: $\xi^1 = i$ y $\xi^2 = j$. Los términos $c_{ij}^3 = c^3(\xi^1 = i, \xi^2 = j)$ reflejan el coste o beneficio por kilogramo de ensaladilla rusa según cada escenario meteorológico combinado. Así, la función objetivo del modelo sería:

$$\begin{aligned} & -25x^1 + (1-p_1)c_0^2x_0^2 + p_1c_1^2x_1^2 + (1-p_1)(1-p_2)c_{00}^3x_{00}^3 + (1-p_1)p_2c_{01}^3x_{01}^3 \\ & + p_1(1-p_2)c_{10}^3x_{10}^3 + p_1p_2c_{11}^3x_{11}^3 = -25x^1 + 0.6 \cdot 23x_0^2 + 0.4 \cdot (-27)x_1^2 \\ & + 0.6 \cdot 0.7 \cdot 28x_{00}^3 + 0.6 \cdot 0.3 \cdot (-26)x_{01}^3 + 0.4 \cdot 0.7 \cdot (-24)x_{10}^3 + 0.4 \cdot 0.3 \cdot 22x_{11}^3 \\ & = -25x^1 + 13.8x_0^2 - 10.8x_1^2 + 11.76x_{00}^3 - 4.68x_{01}^3 - 6.72x_{10}^3 + 2.64x_{11}^3 \end{aligned}$$

Particularizando, además, las restricciones originales a cada nueva variable introducida por escenario, obtenemos el problema:

$$\begin{aligned} & \min_{x^1, x_0^2, x_1^2, x_{00}^3, x_{01}^3, x_{10}^3, x_{11}^3} && -25x^1 + 13.8x_0^2 - 10.8x_1^2 + 11.76x_{00}^3 - 4.68x_{01}^3 - 6.72x_{10}^3 + 2.64x_{11}^3 \\ & \text{s.a.} && x^1 + x_0^2 + x_{00}^3 \leq 300 \\ & && x^1 + x_0^2 + x_{01}^3 \leq 300 \\ & && x^1 + x_1^2 + x_{10}^3 \leq 300 \\ & && x^1 + x_1^2 + x_{11}^3 \leq 300 \\ & && x^1 \leq 50 \\ & && x_0^2 \leq 40 \\ & && x_1^2 \leq 40 \\ & && x^1, x_0^2, x_1^2, x_{00}^3, x_{01}^3, x_{10}^3, x_{11}^3 \geq 0. \end{aligned}$$

Este problema es un problema determinista de tipo (LP), cuya solución óptima es $x^{*1} = 50$, $x_0^{*2} = 0$, $x_1^{*2} = 40$, $x_{00}^{*3} = 0$, $x_{10}^{*3} = 210$, $x_{01}^{*3} = 250$, $x_{11}^{*3} = 0$, con un valor esperado de la función objetivo igual a -4263.2 . Esto se interpreta de la siguiente forma: Paco prepara el viernes $x^{*1} = 50$ kg de ensaladilla rusa. A continuación, observa la meteorología del sábado: si llueve no prepara nada ($x_0^{*2} = 0$), mientras que si hace sol prepara $x_1^{*2} = 40$ kg. El domingo vuelve a decidir en función de la meteorología conjunta del fin de semana. Si el tiempo es el mismo en ambos días, no prepara ensaladilla ($x_{00}^{*3} = 0, x_{11}^{*3} = 0$); si llovió el sábado y el domingo hace sol, produce $x_{01}^{*3} = 250$ kg; y si hizo sol el sábado y llueve el domingo, produce $x_{10}^{*3} = 210$ kg. De este modo, Paco implementa una política que minimiza el valor esperado de los costes, evitando producir en escenarios de baja demanda y concentrando la producción en los escenarios más rentables.

Para establecer una breve comparación con el planteamiento de dos etapas, asumamos que x^1 y x^2 se toman conjuntamente en la primera etapa, mientras que x^3 queda para la segunda. En tal caso, el parámetro c^2 deja de depender de la incertidumbre y se reemplaza

por su valor esperado, $\mathbb{E}[c^2(\xi^1)] = 3$. La formulación determinista resultante es:

$$\begin{aligned} \min_{x^1, x^2, x^3, x_{00}^3, x_{01}^3, x_{10}^3, x_{11}^3} & -25x^1 + 3x^2 + 11.76x_{00}^3 - 4.68x_{01}^3 - 6.72x_{10}^3 + 2.64x_{11}^3 \\ \text{s.a.} & x^1 + x^2 + x_{00}^3 \leq 300 \\ & x^1 + x^2 + x_{01}^3 \leq 300 \\ & x^1 + x^2 + x_{10}^3 \leq 300 \\ & x^1 + x^2 + x_{11}^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x_{00}^3, x_{01}^3, x_{10}^3, x_{11}^3 \geq 0. \end{aligned}$$

La solución óptima de este problema viene dada por $x^{*1} = 50$, $x^{*2} = 0$, $x_{00}^{*3} = 0$, $x_{01}^{*3} = 250$, $x_{10}^{*3} = 250$, $x_{11}^{*3} = 0$, con valor objetivo igual a -4100 . La interpretación de la solución resulta análoga a la expuesta en el párrafo anterior.

En comparación con el modelo de dos etapas, el valor de la función objetivo obtenido con el modelo de tres etapas mejora en $\frac{-4100+4682}{4100} \cdot 100 = -3.98\%$, lo que equivale a una reducción de los costes de casi un 4%.

Como se muestra en el ejemplo, la capacidad de ajustar las decisiones a medida que se revela la incertidumbre ofrece una ventaja cuantificable. Los modelos multietapa superan en rendimiento a los de una o dos etapas, dado que incorporan más puntos de decisión y, por tanto, mayor adaptabilidad. No obstante, esta flexibilidad viene acompañada de un incremento en el número de variables y restricciones, lo que puede dificultar su resolución en instancias de gran tamaño. Resolver estos modelos no es una tarea trivial, sin embargo, la literatura propone varios enfoques que reducen la carga computacional que implica la resolución de un modelo multietapa, manteniendo la calidad de la solución. Para profundizar más en estas técnicas, se recomienda consultar [11].

Por nuestra parte, presentamos en el próximo capítulo la Optimización Robusta, un enfoque complementario que evita la necesidad de estimar distribuciones, presenta en general una carga computacional menor y se centra en garantizar soluciones viables frente al peor escenario posible.

Capítulo 2

Optimización Robusta

1. Introducción

Como se ha comentado en la sección anterior, la programación estocástica resulta una herramienta poderosa en contextos donde los parámetros de entrada pueden modelarse como variables aleatorias con distribuciones de probabilidad conocidas, la toma de decisiones se realiza con frecuencia y/o la tolerancia al riesgo es moderada.

Sin embargo, cuando la distribución de probabilidad es desconocida, la incertidumbre es elevada o la tolerancia al riesgo es reducida, su aplicabilidad se ve limitada. En estos casos, la dependencia exclusiva de la esperanza matemática puede no capturar adecuadamente los riesgos asociados a escenarios extremos, comprometiendo la fiabilidad de la solución. A continuación, se presenta un ejemplo que ilustra esta limitación y motiva la necesidad de enfoques alternativos.

Ejemplo 7. *Consideremos el problema de encontrar la ruta más rápida para desplazarse al trabajo (una decisión de alta frecuencia, ya que se toma a diario, y sin consecuencias críticas en caso de desviaciones adversas), donde el objetivo es minimizar el tiempo total de viaje T . Este tiempo puede modelarse como una variable aleatoria que depende de factores inciertos, como el tráfico o el estado de los semáforos. Minimizar el valor esperado de T es una estrategia razonable en este contexto, ya que las consecuencias de una desviación adversa pueden limitarse, por ejemplo, a una reprimenda del jefe. Ahora bien, consideremos un escenario distinto: en lugar de ir al trabajo, se conduce un camión de bomberos que debe llegar con urgencia a un edificio en llamas. Supongamos que existen dos rutas disponibles: la ruta A , cuyo tiempo de viaje $T \sim U([3, 8])$, y la ruta B , cuyo tiempo de viaje $T \sim U([6, 7])$. Si se elige la ruta basándose solo en el tiempo esperado, se podría preferir la ruta A (con valor esperado $(3 + 8)/2 = 5.5$ minutos) frente a la ruta B (con valor esperado $(6 + 7)/2 = 6.5$ minutos). Sin embargo, una solución robusta considera el peor escenario: la ruta A , pese a tener un tiempo de viaje medio menor a la ruta B , podría requerir hasta 8 minutos, mientras que la ruta B tomaría como máximo 7 minutos.*

En conclusión, este ejemplo ilustra cómo en contextos con alta aversión al riesgo (es decir, cuando se prioriza evitar los peores escenarios) no basta con optimizar el valor esperado. Además, cuando el número de variables inciertas es elevado, los árboles de escenarios utilizados en programación estocástica pueden crecer exponencialmente, lo que hace que los problemas resultantes sean computacionalmente intratables en la práctica.

En respuesta a estas limitaciones, surge un nuevo paradigma conocido como *Optimización Robusta*, propuesto originalmente por Scarf en 1958, [20]. La filosofía de este

paradigma ante un problema con incertidumbre es la de optimizar el peor caso posible que pueda tomar la incertidumbre, garantizando la factibilidad de las restricciones para cualquier posible realización de la misma. Matemáticamente, en su forma más general, la optimización robusta plantea, siguiendo esta filosofía de peor caso, el siguiente **problema robusto**, [21]:

$$\begin{aligned} & \underset{x}{\text{mín}} \quad \underset{c \in \mathcal{C}}{\text{máx}} \quad c'x \\ & \text{s.a.} \quad Ax \leq b \quad \forall (A, b) \in \mathcal{U}_{(A,b)} \\ & \quad \quad x \geq 0 \end{aligned} \tag{RO}$$

donde \mathcal{C} y $\mathcal{U}_{(A,b)}$ son los conjuntos de incertidumbre que agrupan todas las posibles realizaciones del vector de costes c y del par (A, b) , respectivamente.

Para el desarrollo teórico que se presenta a continuación, es conveniente introducir el concepto de *problema nominal*, que servirá como referencia a lo largo de este capítulo. Dicho problema se obtiene fijando los parámetros inciertos en un valor fijo representativo. En nuestro caso, adoptaremos como referencia el valor esperado. Formalmente, el problema nominal asociado a nuestro modelo con incertidumbre de partida (LP_ξ) se define como:

$$\begin{aligned} & \underset{x}{\text{mín}} \quad \bar{c}'x \\ & \text{s.a.} \quad \bar{A}x \leq \bar{b} \\ & \quad \quad x \geq 0 \end{aligned} \tag{LP}_{nom}$$

donde \bar{c} , \bar{A} y \bar{b} representan los valores esperados de los parámetros inciertos del problema original.

Además, supondremos que el modelo (RO) tiene una función objetivo determinista, centrado únicamente la incertidumbre en la matriz de coeficientes de restricciones A , puesto que nos centraremos en reformular de manera lineal restricciones del tipo:

$$a'_i x \leq b \quad \forall a_i \in \mathcal{U}_i \tag{RR}$$

donde a_i es la fila i de la matriz A y \mathcal{U}_i es su conjunto de incertidumbre. Como se explicó en la Introducción, donde demostrábamos que el problema (LP_ξ) es equivalente al problema (LP_ξ^A), esta hipótesis no supone una pérdida de generalidad.

Tras analizar detenidamente la forma de (RR), resulta evidente que la definición del conjunto de incertidumbre es el elemento que determina qué escenarios serán considerados en el análisis y, por tanto, el grado de protección que ofrece la solución robusta. La forma concreta de este conjunto también influye de manera decisiva en la posibilidad de obtener reformulaciones lineales y eficientes, tema que abordaremos en la siguiente sección.

2. Conjuntos de incertidumbre y reformulaciones

El concepto de conjunto de incertidumbre es el núcleo de la optimización robusta. En esta sección presentaremos dos tipos de conjuntos de incertidumbre, cada uno con sus ventajas y limitaciones. Comenzaremos con conjuntos de tipo poliedro, definidos mediante sistemas de inecuaciones lineales. A continuación, abordaremos el enfoque propuesto por Bertsimas y Sim, [22], que introduce una parametrización adicional para controlar el grado de protección frente a la incertidumbre manteniendo al mismo tiempo la eficiencia computacional.

Cada caso se abordará introduciendo primero la definición teórica del conjunto de incertidumbre, después su reformulación determinista equivalente y, para cerrar, un ejemplo que ilustre de manera concreta su funcionamiento.

2.1. Conjuntos de incertidumbre poliédricos

Los *conjuntos de incertidumbre poliédricos* son una clase especialmente relevante de conjuntos de incertidumbre por su generalidad y por permitir reformulaciones tratables. Con el objetivo de construir la intuición de forma progresiva, comenzaremos analizando un caso particular de este tipo de conjuntos ampliamente utilizados en la literatura: los conjuntos tipo *caja*. A partir de ellos, generalizaremos posteriormente al caso poliédrico general.

Conjuntos de incertidumbre tipo caja

Si cada coeficiente a_{ij} de la matriz de restricciones $A \in \mathbb{R}^{m \times n}$, con $i \in \{1, \dots, m\}$, y $j \in \{1, \dots, n\}$ puede variar de forma independiente dentro de un intervalo centrado en su valor nominal \bar{a}_{ij} , con amplitud \hat{a}_{ij} , es decir,

$$a_{ij} \in [\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}],$$

entonces decimos que el vector $a_i = (a_{i1}, \dots, a_{in})$ reside en un conjunto de incertidumbre de tipo caja [21]:

$$\mathcal{U}_{\text{caja}} := \{a_i \in \mathbb{R}^n : \exists \xi \in \mathbb{R}^n \text{ tal que } a_i = \bar{a}_i + \Delta_i \xi, \|\xi\|_\infty \leq 1\}, \quad (2.1)$$

donde $\bar{a}_i = (\bar{a}_{i1}, \dots, \bar{a}_{in}) \in \mathbb{R}^n$ representa el vector nominal de la fila i , $\Delta_i = \text{diag}(\hat{a}_{i1}, \dots, \hat{a}_{in})$ es la matriz diagonal de desviaciones asociadas a cada componente de dicha fila, y $\|\cdot\|_\infty$ denota la norma infinito.

Nuestro objetivo es ahora reformular la restricción robusta asociada a este conjunto de incertidumbre de manera lineal, es decir, reformular la restricción, donde, por simplicidad, se ha omitido el subíndice i que denota la fila:

$$a'x \leq b \quad \forall a \in \mathcal{U}_{\text{caja}}. \quad (\text{RRC})$$

A partir de la definición del conjunto $\mathcal{U}_{\text{caja}}$ es claro que (RRC) es equivalente a:

$$\max_{\|\xi\|_\infty \leq 1} \{\bar{a}'x + \xi'(\Delta x)\} \leq b, \quad (\text{RRC}_{\|\cdot\|_\infty})$$

Ahora bien, dado que las normas $\|\cdot\|_\infty$ y $\|\cdot\|_1$ son duales entre sí, es decir, que para cualquier vector $v \in \mathbb{R}^n$, [23], se tiene que

$$\max_{\|\xi\|_\infty \leq 1} \xi'v = \|v\|_1,$$

entonces, $(\text{RRC}_{\|\cdot\|_\infty})$ es equivalente a la siguiente inecuación determinista:

$$\bar{a}'x + \|\Delta x\|_1 \leq b. \quad (\text{RRC}_{\|\cdot\|_1})$$

Finalmente, aplicando la definición de $\|\cdot\|_1$ a $(\text{RRC}_{\|\cdot\|_1})$ se tiene:

$$\bar{a}'x + \sum_{i=1}^n |\hat{a}_i x_i| \leq b, \quad (\text{RRC}_{\Sigma|\cdot|})$$

En nuestro modelo base (LP) suponemos que las variables de decisión x son siempre no negativas, por tanto, la restricción robusta (RRC) queda finalmente como:

$$\bar{a}'x + \sum_{i=1}^n |\hat{a}_i|x_i \leq b, \quad (\text{RRC}_{lin})$$

Nota: Si el modelo a considerar tomase variables negativas, para obtener una formulación completamente lineal de $(\text{RRC}_{\Sigma|\cdot})$, introducimos variables auxiliares $y_j \geq 0$ para cada $j = 1, \dots, n$, que satisfacen:

$$y_j \geq \hat{a}_j x_j, \quad y_j \geq -\hat{a}_j x_j.$$

Es decir, $y_j \geq |\hat{a}_j x_j|$, y en la solución óptima se alcanzará la igualdad, ya que la suma $\sum_j y_j$ aparecerá en la función objetivo o en el lado izquierdo de una inecuación, y por tanto se minimizará.

De este modo, la restricción robusta $(\text{RRC}_{\Sigma|\cdot})$ puede reformularse como:

$$\begin{cases} \bar{a}'x + \sum_{j=1}^n y_j \leq b, \\ -\hat{a}_j x_j \leq y_j \leq \hat{a}_j x_j, \quad \forall j = 1, \dots, n. \end{cases} \quad (\text{RRC}_{lin\pm})$$

Para ilustrar el uso de esta reformulación en la práctica, retomemos el Ejemplo 6, pero ahora considerando que $c^2 \in [-27, 23]$ y $c^3 \in [-26, 28]$. Buscamos una solución robusta frente a estas variaciones paramétricas, es decir, que mantenga su factibilidad y optimalidad para cualquier realización de los parámetros dentro de dichos intervalos. Obsérvese que, aunque en el desarrollo teórico hemos supuesto que la incertidumbre reside en las restricciones, en este ejemplo la situamos en la función objetivo con el fin de mantener la coherencia narrativa a lo largo del trabajo.

Ejemplo 8. Sea $c = (c^2, c^3)$, entonces, el conjunto de incertidumbre de c es de tipo caja:

$$\mathcal{U}_{caja} = \{c \in \mathbb{R}^2 \mid \exists \xi \in \mathbb{R}^2 \text{ tal que } c = \bar{c} + \Delta\xi, \|\xi\|_\infty \leq 1\}$$

$$\text{con } \bar{c} = \left(\frac{-27+23}{2}, \frac{-26+28}{2}\right) = (-2, 1) \text{ y } \Delta = \text{diag}\left(\frac{23-(-27)}{2}, \frac{28-(-26)}{2}\right) = \text{diag}(25, 27).$$

Por tanto, el problema (1.8) con esta modificación quedaría formulado como un problema de tipo (RO):

$$\begin{aligned} \min_{x^1, x^2, x^3} \quad & -25x^1 + \max_{c \in \mathcal{U}_{caja}} c^2 x^2 + c^3 x^3 \\ \text{s. a.} \quad & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3 \geq 0 \end{aligned} \quad (2.2)$$

Lo primero que debemos hacer es bajar la incertidumbre a las restricciones, así obtenemos el problema:

$$\min_{W, x^1, x^2, x^3} \quad W \quad (2.3)$$

$$\text{s. a.} \quad c^2 x^2 + c^3 x^3 \leq W + 25x^1 \quad \forall c \in \mathcal{U}_{caja}, \quad (2.4)$$

$$x^1 + x^2 + x^3 \leq 300 \quad (2.5)$$

$$x^1 \leq 50 \quad (2.6)$$

$$x^2 \leq 40 \quad (2.7)$$

$$x^1, x^2, x^3 \geq 0 \quad (2.8)$$

Es claro que por la definición de \mathcal{U}_{caja} , la restricción (2.4) se reescribe como:

$$\max_{\|\xi\|_\infty \leq 1} \left[(-2, 1) \begin{pmatrix} x^2 \\ x^3 \end{pmatrix} + (\xi_1, \xi_2) \begin{pmatrix} 25 & 0 \\ 0 & 27 \end{pmatrix} \begin{pmatrix} x^2 \\ x^3 \end{pmatrix} \right] \leq W + 25x^1$$

donde, por simplicidad, escribimos $\xi = (\xi_1, \xi_2)$. Esta restricción se reescribe a su vez, gracias a la dualidad entre $\|\cdot\|_\infty$ y $\|\cdot\|_1$ como:

$$(-2, 1) \begin{pmatrix} x^2 \\ x^3 \end{pmatrix} + \left\| \begin{pmatrix} 25 & 0 \\ 0 & 27 \end{pmatrix} \begin{pmatrix} x^2 \\ x^3 \end{pmatrix} \right\|_1 \leq W + 25x^1.$$

Aplicando la definición de la norma 1 y desarrollando queda:

$$-2x^2 + x^3 + 25|x^2| + 27|x^3| \leq W + 25x^1.$$

Como x^2 y x^3 son siempre positivas, podemos quitar el valor absoluto, obteniendo la restricción:

$$23x^2 + 28x^3 \leq W + 25x^1.$$

Así, nuestro problema de optimización se vería como:

$$\begin{aligned} & \min_{W, x^1, x^2, x^3} && W \\ & \text{s.a.} && 23x^2 + 28x^3 \leq W + 25x^1 \\ & && x^1 + x^2 + x^3 \leq 300 \\ & && x^1 \leq 50 \\ & && x^2 \leq 40 \\ & && x^1, x^2, x^3 \geq 0. \end{aligned}$$

En este caso, podemos omitir la variable W observando que la restricción $W + 25x^1 - 23x^2 - 28x^3 \geq 0$ junto con la minimización de W fuerza la igualdad $W = -25x^1 + 23x^2 + 28x^3$. Por tanto, el problema se simplifica a:

$$\begin{aligned} & \min_{x^1, x^2, x^3} && -25x^1 + 23x^2 + 28x^3 \\ & \text{s.a.} && x^1 + x^2 + x^3 \leq 300 \\ & && x^1 \leq 50 \\ & && x^2 \leq 40 \\ & && x^1, x^2, x^3 \geq 0. \end{aligned} \tag{2.9}$$

Este último problema es un problema de programación lineal del tipo (LP) con solución óptima $x^{*1} = 50$, $x^{*2} = 0$, $x^{*3} = 0$ y valor objetivo -1250 . Es decir, Paco opta por preparar únicamente 50 kg de ensaladilla rusa el viernes y, ante la incertidumbre asociada al fin de semana, decide no producir nada más en las etapas posteriores, asegurando así una estrategia conservadora frente a posibles escenarios adversos.

Podemos comparar el resultado obtenido al resolver este problema con un enfoque estocástico, si suponemos que $c^2 \sim U([-27, 23])$ y $c^3 \sim U([-26, 28])$. En este caso, obtendríamos que el problema queda formulado como:

$$\begin{aligned} & \min_{x^1, x^2, x^3} && \mathbb{E}[-25x^1 + c^2x^2 + c^3x^3] \\ & \text{s.a.} && x^1 + x^2 + x^3 \leq 300 \\ & && x^1 \leq 50 \\ & && x^2 \leq 40 \\ & && x^1, x^2, x^3 \geq 0 \end{aligned} \tag{2.10}$$

Por la linealidad de la esperanza y sabiendo que $\mathbb{E}[c^2] = -2$ y $\mathbb{E}[c^3] = 1$, el problema se simplifica a:

$$\begin{aligned} \min_{x^1, x^2, x^3} \quad & -25x^1 - 2x^2 + x^3 \\ \text{s. a.} \quad & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3 \geq 0 \end{aligned} \tag{2.11}$$

cuya solución óptima es $x^{*1} = 50$, $x^{*2} = 40$, $x^{*3} = 0$, con valor objetivo -1330 . En términos prácticos, esto significa que Paco prepara 50 kg de ensaladilla rusa el viernes y añade otros 40 kg el sábado, pero decide no producir nada el domingo. En comparación con el modelo nominal, el valor objetivo empeora en torno a un 6.4% ($\frac{-1330+1250}{1250} = 6.4\%$). A cambio, se obtiene una solución robusta, capaz de proteger frente al peor escenario posible.

En conclusión, hemos reducido una restricción robusta con infinitas realizaciones posibles a un conjunto finito de cinco inecuaciones lineales, manteniendo la resolubilidad del problema y cubriendo el peor escenario. A continuación, consideramos el caso más general donde se engloban los conjuntos de tipo caja: los conjuntos de incertidumbre poliédricos.

Conjuntos de incertidumbre poliédricos

La siguiente restricción robusta que pretendemos reformular es aquella cuyos coeficientes residen en un conjunto de incertidumbre poliédrico. Un poliedro se define como un conjunto convexo que puede describirse como la intersección de un número finito de semiespacios definidos por inecuaciones lineales (**H-representación**). Así, un poliedro en su H-representación se expresa como, [21]:

$$\mathcal{U}_{\text{poliedro}} := \{a \in \mathbb{R}^n : Ba \leq d\},$$

donde $B \in \mathbb{R}^{k \times n}$ y $d \in \mathbb{R}^k$ definen un conjunto de restricciones lineales. La desigualdad $Ba \leq d$ se interpreta componente a componente.

Es claro que los conjuntos de tipo caja son casos particulares de poliedros. Su H-representación sería, [24]:

$$\begin{pmatrix} I \\ -I \end{pmatrix} a \leq \begin{pmatrix} \bar{a} + \hat{a} \\ -\bar{a} + \hat{a} \end{pmatrix}$$

donde I es la matriz identidad de dimensión $n \times n$. Además, recordemos que $\bar{a} \in \mathbb{R}^n$ es el vector nominal y $\hat{a} \in \mathbb{R}^n$ representa las desviaciones máximas que puede tomar a . Veamos a continuación cómo convertir la restricción robusta:

$$a'x \leq b, \quad \forall a \in \mathcal{U}_{\text{poliedro}} \tag{RRP}$$

en una restricción lineal determinista equivalente. De la definición de $\mathcal{U}_{\text{poliedro}}$ se deduce que, para $x_0 \in \mathbb{R}^n$ fijo (RRP) es equivalente a:

$$\left\{ \begin{array}{l} \text{máx} \\ a \\ \text{s. a.} \end{array} \quad \begin{array}{l} a'x_0 \\ Ba \leq d \end{array} \right\} \leq b. \tag{RRP}_{\text{def}}$$

Considerando el problema de optimización lineal contenido entre llaves en la expresión anterior y tomando su dual:

$$\begin{aligned}
& \underset{\lambda}{\text{mín}} && \lambda'd \\
& \text{s.a.} && B'\lambda = x_0 \\
& && \lambda \geq 0
\end{aligned} \tag{RRP}_{dual}$$

Recordando que nuestra función objetivo original es $\text{mín } c'x$, si el modelo incluye una restricción robusta del tipo $a'x \leq b$ con $a \in \mathcal{U} = \{a : Ba \leq d\}$, entonces el problema completo se escribe como un problema de programación anidada:

$$\begin{aligned}
& \underset{x}{\text{mín}} && c'x \\
& \text{s.a.} && \left\{ \begin{array}{l} \underset{\lambda}{\text{mín}} \quad \lambda'd \\ \text{s.a.} \quad B'\lambda = x_0, \\ \lambda \geq 0 \end{array} \right\} \leq b.
\end{aligned} \tag{ROP}_{min-min}$$

Notése que es posible unificar ambos mínimos, el de la función objetivo y el derivado de la reformulación robusta, en un único problema de optimización, incorporando λ como variable auxiliar adicional y dejando de fijar x en x_0 , obteniéndose el modelo:

$$\begin{aligned}
& \underset{x, \lambda}{\text{mín}} && c'x \\
& \text{s.a.} && \lambda'd \leq b \\
& && B'\lambda = x \\
& && x, \lambda \geq 0
\end{aligned} \tag{ROP}_{min}$$

Es importante observar que en $(\text{ROP}_{min-min})$ se escribe x_0 porque, al resolver el subproblema interno (RRP_{def}) , el vector x ya está fijado por el problema externo y actúa como parámetro. En cambio, al integrar dicho subproblema en el modelo completo, se optimiza sobre todas las decisiones posibles, de modo que x vuelve a ser una variable de decisión, como se refleja en (ROP_{min}) .

Comprobemos ahora la equivalencia entre $(\text{ROP}_{min-min})$ y (ROP_{min}) . Supongamos que tenemos un par (x, λ) óptimo para (ROP_{min}) . Entonces, x es factible para $(\text{ROP}_{min-min})$, ya que cumple la restricción robusta (en efecto, para ese x se tiene $\text{mín}_{\lambda \geq 0, B'\lambda = x} \lambda'd \leq \lambda'd \leq b$). Además, el valor de la función objetivo es el mismo, pues en ambos casos se evalúa $c'x$.

Recíprocamente, supongamos que tenemos un x óptimo para $(\text{ROP}_{min-min})$. Dado que x es factible para $(\text{ROP}_{min-min})$, existe un λ que verifica la restricción dual correspondiente al problema interno, es decir, $B'\lambda = x$ y $\lambda \geq 0$. Por tanto, el par (x, λ) es factible para (ROP_{min}) , y el valor de la función objetivo sigue siendo $c'x$. De este modo, ambos problemas tienen la misma región factible proyectada sobre x y el mismo valor óptimo. Esto justifica que podemos resolver directamente (ROP_{min}) como un problema de programación lineal estándar sin necesidad de mantener la estructura min-min.

Así, la dualidad nos ha permitido resolver un problema de programación lineal robusta con incertidumbre politópica simplemente resolviendo un problema de programación lineal estándar.

En resumen, una restricción asociada a un conjunto de incertidumbre poliédrico permite reformular el problema completo de manera tratable como un programa lineal clásico. Esta propiedad constituye una de las principales ventajas de los conjuntos poliédricos en el marco de la optimización robusta.

Una vez presentada la teoría sobre cómo reformular una restricción del tipo (RRP), ilustramos este procedimiento mediante un ejemplo numérico.

Ejemplo 9. *El problema expuesto en el Ejemplo 8 es pertinente para esta parte teórica, pues en él se tiene que c pertenece al siguiente conjunto poliédrico:*

$$\mathcal{U}_{\text{poliedro}} = \{c \in \mathbb{R}^2 \mid Bc \leq d\},$$

donde:

$$B = \begin{pmatrix} I \\ -I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad d = \begin{pmatrix} 23 \\ 28 \\ 27 \\ 26 \end{pmatrix}.$$

Para reformularlo, introducimos nuevamente la variable W , que permite incorporar la incertidumbre en las restricciones y obtener el siguiente problema:

$$\begin{aligned} \text{mín} \quad & W \\ \text{s.a.} \quad & c^2 x^2 + c^3 x^3 \leq W + 25x^1 \quad \forall c \in \mathcal{U}_{\text{poliedro}}, \end{aligned} \quad (2.12)$$

$$x^1 + x^2 + x^3 \leq 300 \quad (2.13)$$

$$x^1 \leq 50 \quad (2.14)$$

$$x^2 \leq 40 \quad (2.15)$$

$$x^1, x^2, x^3 \geq 0 \quad (2.16)$$

Por la definición de $\mathcal{U}_{\text{poliedro}}$, la restricción (2.12) es equivalente a:

$$\left\{ \begin{array}{l} \text{máx}_{c \in \mathcal{U}} \quad c^2 x^2 + c^3 x^3 \\ \text{s.a.} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c^2 \\ c^3 \end{pmatrix} \leq \begin{pmatrix} 23 \\ 28 \\ 27 \\ 26 \end{pmatrix} \end{array} \right\} \leq W + 25x^1$$

la cual queda desarrollada como:

$$\left\{ \begin{array}{l} \text{máx}_{c \in \mathcal{U}} \quad c^2 x^2 + c^3 x^3 \\ \text{s.a.} \quad c^2 \leq 23 \\ \quad \quad c^3 \leq 28 \\ \quad \quad -c^2 \leq 27 \\ \quad \quad -c^3 \leq 26 \end{array} \right\} \leq W + 25x^1.$$

Tomando el dual de este problema de maximización, obtenemos el siguiente problema de minimización:

$$\left\{ \begin{array}{l} \text{mín}_{\lambda_1, \lambda_2, \lambda_3, \lambda_4} \quad 23\lambda_1 + 28\lambda_2 + 27\lambda_3 + 26\lambda_4 \\ \text{s.a.} \quad \lambda_1 - \lambda_3 = x^2 \\ \quad \quad \lambda_2 - \lambda_4 = x^3 \\ \quad \quad \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0 \end{array} \right\} \leq W + 25x^1.$$

Incluyendo esta reformulación en nuestro problema original, llegamos al problema equivalente:

$$\begin{aligned}
 & \underset{W, x^1, x^2, x^3, \lambda_1, \lambda_2, \lambda_3, \lambda_4}{\text{mín}} && W \\
 \text{s. a.} & && 23\lambda_1 + 28\lambda_2 + 27\lambda_3 + 26\lambda_4 \leq W + 25x^1 \\
 & && \lambda_1 - \lambda_3 = x^2 \\
 & && \lambda_2 - \lambda_4 = x^3 \\
 & && x^1 + x^2 + x^3 \leq 300 \\
 & && x^1 \leq 50 \\
 & && x^2 \leq 40 \\
 & && x^1, x^2, x^3, \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0.
 \end{aligned}$$

Nótese que podemos omitir la variable W , obteniendo el siguiente problema de programación lineal:

$$\begin{aligned}
 & \underset{x^1, x^2, x^3, \lambda_1, \lambda_2, \lambda_3, \lambda_4}{\text{mín}} && -25x^1 + 23\lambda_1 + 28\lambda_2 + 27\lambda_3 + 26\lambda_4 \\
 \text{s. a.} & && \lambda_1 - \lambda_3 = x^2 \\
 & && \lambda_2 - \lambda_4 = x^3 \\
 & && x^1 + x^2 + x^3 \leq 300 \\
 & && x^1 \leq 50 \\
 & && x^2 \leq 40 \\
 & && x^1, x^2, x^3, \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0.
 \end{aligned}$$

La solución óptima de este modelo es $x^{*1} = 50$, $x^{*2} = 0$, $x^{*3} = 0$, $\lambda_1^* = 0$, $\lambda_2^* = 0$, $\lambda_3^* = 0$, $\lambda_4^* = 0$, con un valor objetivo de -1250 . Como era de esperar, tanto las soluciones óptimas como el valor óptimo coinciden con los obtenidos en la reformulación basada en el conjunto de incertidumbre tipo caja, dado que ambos modelos representan el mismo problema subyacente. Asimismo, el hecho de que todas las variables duales λ_i sean nulas indica en este caso que ninguna de las restricciones que definen el conjunto de incertidumbre poliédrico es activa en el óptimo; es decir, la solución se determina sin que estas cotas adicionales sobre la incertidumbre resulten determinantes en el valor final de la función objetivo [25]. Si, por el contrario, alguna variable dual resultara estrictamente positiva, ello significaría que la restricción correspondiente del conjunto poliédrico está activa, contribuyendo de manera efectiva a delimitar el peor caso y, en consecuencia, influyendo directamente en el valor de la solución robusta.

En esta sección hemos visto que los conjuntos poliédricos permiten modelar la incertidumbre mediante restricciones lineales, garantizando la tratabilidad del problema. No obstante, este enfoque suele ser excesivamente conservador, pues asume que todos los parámetros inciertos pueden desviarse simultáneamente a sus peores valores. Para superar esta limitación surgen los conjuntos de incertidumbre *presupuestarios*, introducidos por Bertsimas y Sim, [22], una idea que veremos a continuación con más detalle.

2.2. Conjuntos de incertidumbre tipo presupuesto

En los conjuntos presupuestarios se introduce un parámetro, denotado por Γ , que controla cuántos coeficientes inciertos pueden desviarse adversamente de sus valores nominales. Este mecanismo mantiene la estructura lineal del modelo y permite modular el grado de conservadurismo, ofreciendo un equilibrio flexible entre robustez y rendimiento.

A continuación, presentamos la formulación matemática del modelo. Para ello, introduciremos algunas definiciones, notación e hipótesis necesarias. Posteriormente, ilustraremos su funcionamiento mediante un ejemplo numérico concreto.

Sea J_i el conjunto de *índices* correspondientes a los coeficientes inciertos en la fila i de la matriz A . Supondremos, de ahora en adelante, que cada a_{ij} , con $j \in J_i$, está sujeto a incertidumbre y se modela como una variable aleatoria a_{ij} simétrica y acotada, que toma valores en el intervalo $[\bar{a}_{ij} - \hat{a}_{ij}, \bar{a}_{ij} + \hat{a}_{ij}]$, donde \bar{a}_{ij} denota el valor nominal de la incertidumbre asociada a a_{ij} y \hat{a}_{ij} la máxima desviación posible de dicha incertidumbre.

Para cada fila i , introducimos además el parámetro $\Gamma_i \in [0, |J_i|]$, que en nuestro caso, por simplicidad expositiva, supondremos entero. Este parámetro, como ya veníamos adelantando, representa el número de coeficientes de la fila i . Definimos de igual manera, para cada fila i , la función de protección $\beta_i : X \times \mathbb{N} \cap [0, |J_i|] \rightarrow \mathbb{R}$ como:

$$\beta_i(x, \Gamma_i) := \max_{\substack{S_i \subseteq J_i \\ |S_i| = \Gamma_i}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} x_j \right\}. \quad (\beta_{i,def})$$

Esta función cuantifica el peor caso posible en la restricción i , tomando las Γ_i desviaciones más adversas dentro de los coeficientes inciertos. En particular, $\beta_i(x, 0) = 0$ para todo $x \in X$, es decir, no se permite ninguna desviación (y por tanto no habría incertidumbre), y $\beta_i(x, |J_i|)$ recoge el peor caso posible en el que todos los parámetros inciertos de la fila i toman su valor más adverso.

A pesar de que la función de protección $\beta_i(x, \Gamma_i)$ se define mediante un máximo no lineal, es posible reformularla de manera exacta como un problema de programación lineal. En concreto, se verifica que:

$$\begin{aligned} \beta_i(x, \Gamma_i) &:= \max_y \sum_{j \in J_i} \hat{a}_{ij} x_j y_j \\ \text{s. a.} \quad &\sum_{j \in J_i} y_j \leq \Gamma_i \\ &y_j \leq 1 \quad \forall j \in J_i \\ &y_j \geq 0. \quad \forall j \in J_i \end{aligned} \quad (\beta_{i,linmax})$$

Nótese que, en esta reformulación, las variables x_j pasan a formar parte de los coeficientes del objetivo, mientras que se introducen nuevas variables auxiliares y_j que son las variables que se maximizan.

Veamos esta reformulación con más detalle. Por un lado, es claro que la solución óptima del problema $(\beta_{i,linmax})$ tomará $y_j = 1$ para exactamente Γ_i índices. Además, seleccionará aquellos valores de $j \in J_i$ para los cuales el producto $\hat{a}_{ij} x_j$ sea lo más grande posible, ya que la función objetivo es de maximización. Esto equivale a elegir el subconjunto $S_i \subseteq J_i$ de tamaño Γ_i que maximiza la suma $\sum_{j \in S_i} \hat{a}_{ij} x_j$, que es precisamente la definición original de $\beta_i(x, \Gamma_i)$ dada en $(\beta_{i,def})$. Por tanto, ambos problemas son equivalentes.

Una vez hemos visto cómo la función de protección $\beta_i(x, \Gamma_i)$ selecciona las Γ_i desviaciones de la incertidumbre más desfavorables, solo queda incorporarla en el problema (LP_{nom}) . De este modo, obtenemos el siguiente modelo no lineal:

$$\begin{aligned}
 & \underset{x}{\text{mín}} && c'x \\
 & \text{s.a.} && \sum_{j=1}^n \bar{a}_{ij}x_j + \beta_i(x, \Gamma_i) \leq b_i \quad \forall i \\
 & && x \geq 0.
 \end{aligned} \tag{BS}_{nl}$$

A continuación, veamos cómo podemos reformularlo de manera lineal. Por un lado, consideramos el dual del problema $(\beta_{i,\text{linmax}})$:

$$\begin{aligned}
 \beta_i(x, \Gamma_i) = & \underset{z_i, \lambda_{ij}}{\text{mín}} && \Gamma_i z_i + \sum_{j \in J_i} \lambda_{ij} \\
 & \text{s.a.} && z_i + \lambda_{ij} \geq \hat{a}_{ij}x_j, \quad \forall j \in J_i \\
 & && z_i, \lambda_{ij} \geq 0. \quad \forall j \in J_i
 \end{aligned} \tag{\beta_{i,dual}}$$

Por el teorema de dualidad fuerte, (véase Teorema 5, Apéndice A) dado que el problema primal $(\beta_{i,\text{linmax}})$ es factible y acotado para todo $\Gamma_i \in [0, |J_i|]$, entonces su problema dual $(\beta_{i,dual})$ también lo es, y ambos comparten el mismo valor óptimo. Es decir, $\beta_i(x, \Gamma_i)$ coincide exactamente con el valor óptimo del problema $(\beta_{i,dual})$. Por tanto, basta con sustituir la expresión dual en el modelo no lineal (BS_{nl}) , para posteriormente unificar los dos min min tal como se hizo en $(\text{ROP}_{\text{min-min}})$ - $(\text{ROP}_{\text{min}})$ obteniéndose la reformulación del modelo:

$$\begin{aligned}
 & \underset{x, z, \lambda}{\text{mín}} && c'x \\
 & \text{s.a.} && \sum_{j=1}^n \bar{a}_{ij}x_j + \Gamma_i z_i + \sum_{j \in J_i} \lambda_{ij} \leq b_i, \quad \forall i = 1, \dots, m \\
 & && z_i + \lambda_{ij} \geq \hat{a}_{ij}x_j, \quad \forall i = 1, \dots, m, \quad j \in J_i \\
 & && x_j \geq 0, \quad \forall j; \quad z_i \geq 0, \quad \forall i = 1, \dots, m; \quad \lambda_{ij} \geq 0, \quad \forall i = 1, \dots, m, \quad j \in J_i
 \end{aligned} \tag{BS}_{lin}$$

Si las variables x_j pudiesen tomar valores negativos, podrían introducirse variables auxiliares y_j tales que $-y_j \leq x_j \leq y_j$, de forma análoga a lo realizado en $(\text{RRC}_{\Sigma|\cdot})$ - $(\text{RRC}_{\text{lin}\pm})$.

La formulación final es completamente lineal y computacionalmente tratable. Consta de $n + k + 1$ variables y $m + k + n$ restricciones, donde n es el número de variables originales, m el número de restricciones originales y $k := \sum_i |J_i|$ el número total de coeficientes sujetos a incertidumbre en el modelo.

El caso en el que Γ_i no es entero es completamente análogo, pero requiere definir la función de protección como:

$$\beta_i(x, \Gamma_i) := \underset{\substack{S_i \subseteq J_i, |S_i| = \lfloor \Gamma_i \rfloor \\ t_i \in J_i \setminus S_i}}{\text{máx}} \left\{ \sum_{j \in S_i} \hat{a}_{ij}x_j + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i}x_{t_i} \right\} \tag{\beta_{i,frac}}$$

En este caso, la interpretación de la función es la siguiente: se seleccionan $\lfloor \Gamma_i \rfloor$ coeficientes que adoptan su peor desviación completa, y un coeficiente adicional que lo hace de forma parcial, en proporción a $\Gamma_i - \lfloor \Gamma_i \rfloor$. En la reformulación de esta función como un problema de programación lineal $(\beta_{i,\text{linmax}})$, esto implica que, para cada i , la variable y_j toma el valor 1 en $\lfloor \Gamma_i \rfloor$ índices y un valor fraccionario igual a $\Gamma_i - \lfloor \Gamma_i \rfloor$ en un único índice adicional.

Para ilustrar el uso práctico del enfoque de Bertsimas y Sim, retomamos el Ejemplo 8, ya que la incertidumbre presente en dicho problema se ajusta a las condiciones requeridas para esta formulación.

Ejemplo 10. Los coeficientes de coste c^2 y c^3 son parámetros inciertos que se modelan como variables acotadas y simétricas, con valores dentro de los intervalos $[\bar{c}^2 - \hat{c}^2, \bar{c}^2 + \hat{c}^2]$ y $[\bar{c}^3 - \hat{c}^3, \bar{c}^3 + \hat{c}^3]$, respectivamente, donde los valores nominales y sus respectivas desviaciones son:

$$\bar{c}^2 = \frac{-27 + 23}{2} = -2, \quad \bar{c}^3 = \frac{-26 + 28}{2} = 1, \quad \hat{c}^2 = \frac{23 - (-27)}{2} = 25, \quad \hat{c}^3 = \frac{28 - (-26)}{2} = 27.$$

La reformulación del problema (1.9) según (BS_{nl}) adopta por tanto la forma:

$$\begin{aligned} \text{mín} \quad & W \\ \text{s.a.} \quad & \bar{c}^2 x^2 + \bar{c}^3 x^3 + \beta(x, \Gamma) \leq W + 25x^1 \\ & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3 \geq 0 \end{aligned} \tag{2.17}$$

donde $\beta(x, \Gamma) = \max_{\substack{S \subseteq \{2,3\} \\ |S| \leq \Gamma}} \sum_{j \in S} \hat{c}^j x^j$. Como únicamente una restricción está sujeta a incertidumbre, omitiremos por simplicidad en la notación, el subíndice i , que identifica a la restricción afectada.

Pasamos ahora a resolver este problema para distintos niveles de robustez, es decir, para $\Gamma = 0$, $\Gamma = 1$ y $\Gamma = 2$.

Si $\Gamma = 0$, entonces $\beta(x, 0) = 0$, y el problema (2.17) se reduce a:

$$\begin{aligned} \text{mín}_{W, x^1, x^2, x^3} \quad & W \\ \text{s.a.} \quad & -2x^2 + x^3 \leq W + 25x^1 \\ & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3 \geq 0 \end{aligned} \tag{2.18}$$

es decir, un problema de programación lineal determinista del tipo (LP). Siguiendo el mismo procedimiento aplicado en secciones anteriores, podemos eliminar la variable W para obtener el problema equivalente:

$$\begin{aligned} \text{mín}_{W, x^1, x^2, x^3} \quad & -25x^1 - 2x^2 + x^3 \\ \text{s.a.} \quad & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3 \geq 0 \end{aligned}$$

Como era de esperar, al tomar $\Gamma = 0$ se recupera el problema (2.11) cuya solución óptima era $x^{*1} = 50$, $x^{*2} = 40$, $x^{*3} = 0$, con un valor objetivo igual a -1330 .

Supongamos ahora que $\Gamma = 1$. En este caso, el término de protección $\beta(x, 1)$ se define como:

$$\beta(x, 1) = \max_{\substack{S \subseteq \{2,3\} \\ |S|=1}} \sum_{j \in S} \hat{c}^j x^j.$$

De acuerdo con la reformulación equivalente del término de protección ($\beta_{i,\text{linmax}}$), podemos expresar $\beta(x, 1)$ como el óptimo del siguiente problema lineal:

$$\begin{aligned} \max_{y_2, y_3} \quad & 25x^2y_2 + 27x^3y_3 \\ \text{s.a.} \quad & y_2 + y_3 \leq 1 \\ & y_2 \leq 1 \\ & y_3 \leq 1 \\ & y_2, y_3 \geq 0 \end{aligned}$$

tomando el dual de este problema, obtenemos:

$$\begin{aligned} \beta(x, 1) := \min_{z, \lambda_2, \lambda_3} \quad & z + \lambda_2 + \lambda_3 \\ \text{s.a.} \quad & z + \lambda_2 \geq 25x^2 \\ & z + \lambda_3 \geq 27x^3 \\ & z, \lambda_2, \lambda_3 \geq 0 \end{aligned} \tag{2.19}$$

Sustituyendo la expresión dual de $\beta(x, 1)$ obtenida en (2.19) dentro del problema (2.17), y unificando ambos mínimos se obtiene el siguiente problema de programación lineal equivalente:

$$\begin{aligned} \min_{x, W, z, \lambda_2, \lambda_3} \quad & W \\ \text{s.a.} \quad & -2x^2 + x^3 + z + \lambda_2 + \lambda_3 \leq W + 25x^1, \\ & z + \lambda_2 \geq 25x^2 \\ & z + \lambda_3 \geq 27x^3 \\ & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3, z, \lambda_2, \lambda_3 \geq 0 \end{aligned} \tag{2.20}$$

Observamos que la variable W aparece únicamente en la función objetivo y en una restricción, de forma lineal. Por tanto, podemos sustituir su valor óptimo directamente en la función objetivo y eliminarla del modelo. El problema resultante es el siguiente:

$$\begin{aligned} \min_{x, z, \lambda_2, \lambda_3} \quad & -25x^1 - 2x^2 + x^3 + z + \lambda_2 + \lambda_3 \\ \text{s.a.} \quad & z + \lambda_2 \geq 25x^2 \\ & z + \lambda_3 \geq 27x^3 \\ & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3, z, \lambda_2, \lambda_3 \geq 0 \end{aligned}$$

La solución óptima de este problema es $x^{*1} = 50$, $x^{*2} = 0$, $x^{*3} = 0$, $z^* = 0$, $\lambda_2^* = 0$, $\lambda_3^* = 0$, con un valor objetivo igual a -1250 . Observamos que este valor es ligeramente mayor que el obtenido con $\Gamma = 0$. A cambio, se obtiene una solución más robusta, que se protege frente al caso en el que uno de los parámetros inciertos toma su peor valor. Supongamos ahora que $\Gamma = 2$. En este caso, el término de protección se convierte en:

$$\beta(x, 2) = \max_{\substack{S \subseteq \{2,3\} \\ |S|=2}} \sum_{j \in S} \tilde{c}^j x^j = 25x^2 + 27x^3.$$

La reformulación equivalente de $\beta(x, 2)$ según $(\beta_{i,\text{linmax}})$ adopta la forma:

$$\begin{aligned} \max_{y_2, y_3} \quad & 25x^2 \cdot y_2 + 27x^3 \cdot y_3 \\ \text{s.a.} \quad & y_2 + y_3 \leq 2 \\ & y_2 \leq 1 \\ & y_3 \leq 1 \\ & y_2, y_3 \geq 0 \end{aligned} \tag{2.21}$$

Aplicando la reformulación dual correspondiente, obtenemos:

$$\begin{aligned} \beta(x, 2) = \min_{z, \lambda_2, \lambda_3} \quad & 2z + \lambda_2 + \lambda_3 \\ \text{s.a.} \quad & z + \lambda_2 \geq 25x^2 \\ & z + \lambda_3 \geq 27x^3 \\ & z, \lambda_2, \lambda_3 \geq 0 \end{aligned} \tag{2.22}$$

Finalmente, al introducir esta reformulación en el problema original y unificar ambos mínimos, se obtiene el siguiente problema de programación lineal equivalente:

$$\begin{aligned} \min_{x, y_2, y_3, W, z, \lambda_2, \lambda_3} \quad & W \\ \text{s.a.} \quad & -2x^2 + x^3 + 2z + \lambda_2 + \lambda_3 \leq W + 25x^1 \\ & z + \lambda_2 \geq 25x^2 \\ & z + \lambda_3 \geq 27x^3 \\ & x^1 + x^2 + x^3 \leq 300 \\ & x^1 \leq 50 \\ & x^2 \leq 40 \\ & x^1, x^2, x^3, z, \lambda_2, \lambda_3 \geq 0 \end{aligned} \tag{2.23}$$

Al omitir la variable auxiliar W , siguiendo el mismo procedimiento aplicado en los casos anteriores, se obtiene el siguiente problema de programación lineal equivalente, de

tipo (LP):

$$\begin{aligned}
 & \underset{x, z, \lambda_2, \lambda_3}{\text{mín}} && -25x^1 - 2x^2 + x^3 + 2z + \lambda_2 + \lambda_3 \\
 & \text{s. a.} && z + \lambda_2 \geq 25x^2 \\
 & && z + \lambda_3 \geq 27x^3 \\
 & && x^1 + x^2 + x^3 \leq 300 \\
 & && x^1 \leq 50 \\
 & && x^2 \leq 40 \\
 & && x^1, x^2, x^3, z, \lambda_2, \lambda_3 \geq 0
 \end{aligned} \tag{2.24}$$

La solución óptima de este problema es $x^{*1} = 50$, $x^{*2} = 0$, $x^{*3} = 0$, $z^* = 0$, $\lambda_2^* = 0$, $\lambda_3^* = 0$, con un valor de la función objetivo igual a -1250 . Como era de esperar, se obtiene la misma solución que cuando resolvíamos este problema mediante el enfoque robusto clásico basado en conjuntos de incertidumbre tipo caja o poliedro (2.9). La función objetivo en este caso es la misma que para $\Gamma = 1$, es decir, aumenta ligeramente respecto al caso $\Gamma = 0$, pero a cambio se garantiza robustez frente al caso en que todos los parámetros inciertos toman el peor caso posible.

A través de los ejemplos anteriores, hemos ilustrado cómo varía la solución óptima y el valor de la función objetivo al aumentar el nivel de protección Γ en el enfoque robusto presupuestario de Bertsimas y Sim. Hemos comprobado que, conforme crece Γ , se obtiene una solución más conservadora y un coste mayor, reflejando así el compromiso entre rendimiento esperado y robustez frente a escenarios adversos. Cabe destacar que, en estos ejemplos, el número de variables y restricciones se ha mantenido reducido para facilitar la interpretación y la resolución explícita de los modelos. Para analizar de forma más general este comportamiento, y comparar cuantitativamente los distintos marcos de optimización bajo incertidumbre (programación estocástica de una y varias etapas, y optimización robusta), llevamos a cabo a continuación una serie de experimentos computacionales sobre instancias del problema de la mochila considerando un mayor número de variables y escenarios.

Capítulo 3

Experimentos computacionales

En este capítulo se analiza empíricamente el comportamiento de las metodologías presentadas en los Capítulos 1 y 2, aplicadas a una variante del problema clásico de la mochila. Comenzamos presentando brevemente el problema base sobre el que posteriormente se introducirá la incertidumbre.

El *problema de la mochila* (*Knapsack Problem*, KP), [26] se define como sigue: se dispone de un conjunto de n objetos, indexados por $j \in \{1, \dots, n\}$, donde cada objeto j ocupa un espacio $a_j > 0$ en la mochila y proporciona una utilidad $c_j > 0$. La mochila tiene una capacidad máxima $b > 0$. El objetivo es seleccionar un subconjunto de objetos cuya suma de tamaños no exceda la capacidad disponible, y que maximice la utilidad total. Este problema puede formularse como el siguiente modelo de programación entera:

$$\begin{aligned} \max_x \quad & \sum_{j=1}^n c_j x_j \\ \text{s. a.} \quad & \sum_{j=1}^n a_j x_j \leq b, \\ & x_j \in \{0, 1\}, \quad \forall j = 1, \dots, n. \end{aligned} \tag{KP}$$

Cada variable x_j indica si el objeto j se incluye en la mochila ($x_j = 1$) o no ($x_j = 0$).

Por coherencia con el resto del TFG, en el que se han estudiado modelos de programación lineal continua, se considerará la relajación $x_j \in [0, 1]$, obteniendo el denominado *problema de la mochila fraccionaria* (*Fractional Knapsack Problem*), [27].

Para incorporar incertidumbre, se aborda una variante en la que las utilidades c_j son parámetros inciertos. Concretamente, se modela cada utilidad como una variable aleatoria simétrica y acotada en el intervalo $[\bar{c}_j - \hat{c}_j, \bar{c}_j + \hat{c}_j]$, donde \bar{c}_j es el valor nominal y \hat{c}_j su desviación. Denominaremos a esta formulación *Uncertain Knapsack Problem* (UKP), [28].

Cuando resolvamos el problema usando enfoques de programación estocástica, dado que es necesario conocer la distribución de probabilidad de la incertidumbre, supondremos que las utilidades c_j siguen una distribución uniforme continua sobre dicho intervalo.

Hasta donde alcanza nuestro conocimiento, no existen conjuntos de instancias estandarizados para esta variante. Por ello, se han seleccionado las 20 primeras instancias para $n = 50, 100, 200, 500$ del tipo 00Uncorrelated y subcategoría R1000 de la librería `Kplib`, [29]. Dichas instancias se han resuelto empleando Python 3.10.12, concretamente con la librería MIP y el solver CBC, utilizando un MacBook Air 2018 con 8 GB de RAM.

A partir de las utilidades que aparecen en cada instancia considerada del repositorio (que serán nuestras utilidades nominales \bar{c}_j), se han definido desviaciones proporcionales

de la forma $\hat{c}_j = \rho \bar{c}_j$, donde según el parámetro ρ diferenciamos los siguientes niveles de incertidumbre:

- Incertidumbre muy baja: $\rho = 0.1$.
- Incertidumbre baja: $\rho = 0.5$.
- Incertidumbre media: $\rho = 1.0$.
- Incertidumbre alta: $\rho = 1.5$.
- Incertidumbre muy alta: $\rho = 2.0$.

Así, se han resuelto $20 \times 4 \times 5 = 400$ variantes del problema de la mochila con los siguientes enfoques:

- **Programación Estocástica:** Dentro de este marco se consideran dos variantes:
 - *Número de etapas:* Se resuelven diferentes instancias con modelos estocásticos de una a cuatro etapas (SP, SP2, SP3 y SP4). En el enfoque de una etapa, el 100 % de las variables corresponden a la primera etapa y, como hemos supuesto que la incertidumbre en las utilidades sigue una distribución uniforme centrada en el valor nominal, este enfoque equivale a resolver el problema nominal. En el caso de dos etapas, un 50 % de las variables corresponden a la primera y un 50 % a la segunda. En tres etapas, el esquema es 25 % en la primera, 50 % en la segunda y 25 % en la tercera. En cuatro etapas, un 0 % en la primera, un 50 % en la segunda y sendos 25 % en la tercera y cuarta. Este esquema permite analizar cómo varía el valor objetivo y el coste computacional al aumentar el número de etapas, garantizando además que las decisiones de etapas posteriores se basan únicamente en información ya observada (no hay filtración de información del futuro).
 - *Variación de la proporción en dos etapas:* Se considera exclusivamente el enfoque de dos etapas, variando la proporción de variables de segunda etapa. Concretamente, se analizan los casos en que el 25 %, 50 % y 75 % de las variables corresponden a la segunda etapa, denotados como SP2₂₅, SP2₅₀ y SP2₇₅, respectivamente. El resto de variables corresponden a la primera etapa. Nótese que SP2 y SP2₅₀ corresponden al mismo enfoque.

Un resumen de las diferentes configuraciones usadas para los enfoques estocásticos se encuentra en la Tabla 3.1.

- **Optimización Robusta:** se aplica la reformulación de Bertsimas y Sim para distintos niveles de protección Γ , que cuantifica cuántos parámetros de utilidad pueden desviarse simultáneamente hacia su peor caso. Se consideran los valores $\Gamma = 0\%$ (caso nominal), 5 %, 10 %, 25 %, 50 % y 100 % (enfoque robusto clásico) de los parámetros que pueden tomar el peor caso.

Además, la elección de modelar la incertidumbre de esta forma no es arbitraria, ya que en los enfoques robustos nos permitirá analizar, para cada porcentaje de parámetros que toman el peor caso posible, cuán cerca queda el valor objetivo respecto al valor objetivo bajo la reducción máxima que puede darse (es decir, el valor objetivo considerando $\Gamma =$

Enfoque	1ª etapa	2ª etapa	3ª etapa	4ª etapa
SP	100 %	0 %	0 %	0 %
SP2 ₂₅	75 %	25 %	0 %	0 %
SP2/SP2 ₅₀	50 %	50 %	0 %	0 %
SP2 ₇₅	25 %	75 %	0 %	0 %
SP3	25 %	50 %	25 %	0 %
SP4	0 %	50 %	25 %	25 %

Tabla 3.1: Distribución de variables de decisión entre etapas en los diferentes enfoques estocásticos.

100 % de parámetros en el peor caso). Expliquemos esto último con más detalle: cuando $\Gamma = 100\%$, es decir, cuando todas las utilidades toman su peor caso posible, la función objetivo del problema UKP es

$$\sum_{j=1}^n (\bar{c}_j - \hat{c}_j)x_j = \sum_{j=1}^n (\bar{c}_j - \rho\bar{c}_j)x_j = (1 - \rho) \sum_{j=1}^n \bar{c}_j x_j$$

y por tanto, el problema UKP quedaría formulado como:

$$\begin{aligned} \max_x \quad & (1 - \rho) \sum_{j=1}^n \bar{c}_j x_j \\ \text{s. a.} \quad & \sum_{j=1}^n a_j x_j \leq b \\ & 0 \leq x_j \leq 1 \quad \forall j \in \{1, \dots, n\} \end{aligned} \tag{KP}_{RO}$$

donde $\rho \in \{0.1, 0.5, 1.0, 1.5, 2.0\}$. En cambio, el problema nominal es:

$$\begin{aligned} \max_x \quad & \sum_{j=1}^n \bar{c}_j x_j \\ \text{s. a.} \quad & \sum_{j=1}^n a_j x_j \leq b \\ & 0 \leq x_j \leq 1 \quad \forall j \in \{1, \dots, n\} \end{aligned} \tag{KP}_{nom}$$

Denotando por Z_{nom} el valor óptimo de (KP_{nom}) y sabiendo que $\bar{c}_j > 0$ para todo j , distinguimos dos situaciones:

- Si $\rho \geq 1$, entonces $(1 - \rho)\bar{c}_j \leq 0$ y la solución óptima de (KP_{RO}) elegirá $x_j = 0 \forall j$. La variación porcentual del valor objetivo respecto a (KP_{nom}) es

$$\frac{0 - Z_{nom}}{Z_{nom}} \cdot 100 = -100\%.$$

- Si $0 < \rho < 1$, la función objetivo de (KP_{RO}) es un reescalado positivo de la nominal, y como se preserva el mismo conjunto factible, ambos problemas tienen la misma

solución óptima. Por tanto, la variación porcentual del valor objetivo de (KP_{RO}) frente a (KP_{nom}) es

$$\frac{(1 - \rho)Z_{nom} - Z_{nom}}{Z_{nom}} \cdot 100 = -\rho \cdot 100 \%.$$

De este modo, al aumentar el número de parámetros en peor caso y estudiar el valor objetivo correspondiente, cuanto más cercana sea la reducción porcentual a $-\rho \cdot 100\%$ o a -100% , más próximo estará el problema al peor caso posible, indicando que la reducción de Γ no aporta mejoras respecto al escenario extremo.

Una vez aclarada esta nota, pasamos a comentar la presentación de los resultados. El análisis se organiza en torno a tres ejes principales:

- Impacto de la incertidumbre en el valor objetivo.
- Dependencia del tamaño del problema en el valor objetivo.
- Relación entre el tamaño del problema y el tiempo de cómputo.

Para abordar estos aspectos se han elaborado diversas tablas y figuras, con los datos más representativos. Los resultados completos, junto con el código empleado para su generación, están disponibles en un repositorio de GitHub, [30]. Concretamente, los datos correspondientes a los enfoques estocásticos se recogen en el archivo `ExperimentosComputacionalesSP.xlsx`, generado a partir del notebook `ExperimentosComputacionalesSP.ipynb`, mientras que los resultados de los enfoques robustos se encuentran en `ExperimentosComputacionalesRO.xlsx`, junto con el notebook `ExperimentosComputacionalesRO.ipynb`.

Pasamos ahora al primer eje de estudio: el impacto de la incertidumbre en el valor objetivo. Para ello, considérense las Tablas B.1, B.2 y B.3 del Apéndice B, junto con sus gráficas asociadas, reflejadas en la Figura 8, que muestran, respectivamente, los resultados para la primera y segunda variante del enfoque estocástico, así como para los modelos robustos.

Más concretamente, estas gráficas muestran los incrementos medios del valor objetivo con respecto al problema nominal para los distintos niveles de incertidumbre, representados mediante líneas de diferentes colores (muy baja en azul, baja en verde, media en naranja, alta en morado y muy alta en rojo), cada una, acompañada de sus intervalos de confianza al 95 %.

Se observa que, en los enfoques estocásticos, aparece un patrón claro: al aumentar la incertidumbre, el valor objetivo aumenta de manera significativa respecto al nominal. Este efecto es especialmente acusado cuando se incrementa el número de etapas o la proporción de variables de segunda etapa, mostrando en este último caso un crecimiento prácticamente lineal. Estos resultados evidencian que una mayor capacidad de adaptación frente a la incertidumbre, particularmente en escenarios más exigentes, conduce a mejoras sustanciales en el valor objetivo. Conviene señalar, además, que los intervalos de confianza tienden a ensancharse ligeramente en niveles elevados de incertidumbre, lo que se explica porque los incrementos relativos son mayores. No obstante, la tendencia global refleja un crecimiento sostenido.

En contraste, los enfoques robustos muestran la dinámica opuesta: a mayor incertidumbre, menor es el valor objetivo alcanzado. En niveles muy altos, se observa que reducir el porcentaje de parámetros que toman el peor caso pierde eficacia a partir de $\Gamma = 50\%$,

pues el deterioro del valor objetivo resulta prácticamente equivalente al del peor caso posible. De forma similar, en niveles de incertidumbre intermedios, los valores de $\Gamma = 75\%$ y $\Gamma = 100\%$ producen deterioros casi indistinguibles. En consecuencia, a partir de dicho umbral, reducir el presupuesto de robustez aporta únicamente mejoras marginales prácticamente despreciables.

Pasamos ahora al segundo punto de análisis: la dependencia del tamaño del problema en el valor objetivo. Para estudiar este aspecto se ha fijado el nivel de incertidumbre en alto (los demás casos son análogos) y se ha analizado cómo varía el incremento del valor objetivo al aumentar el número de ítems considerados ($n \in \{50, 100, 200, 500\}$). Para cada enfoque se representan barras agrupadas: cada grupo corresponde a un valor de n y, dentro de cada grupo, se muestran los distintos modelos considerados en dicho enfoque (SP2–SP4 en la variante A, SP2₂₅, SP2₅₀ y SP2₇₅ en la variante B, y los valores de Γ en el caso robusto), indicando siempre el incremento porcentual respecto al problema nominal.

Este comportamiento puede observarse en las Tablas B.4, B.5 y B.6 del Apéndice B, así como en la Figura 9, donde se aprecia que el aumento del tamaño del problema no produce variaciones significativas en el incremento del valor objetivo: los grupos de barras se mantienen prácticamente uniformes, sin evidenciar tendencias crecientes ni decrecientes, lo que confirma la estabilidad de este indicador frente al tamaño del problema. En dicha figura, el eje X representa el tamaño del problema y el eje Y el incremento del valor objetivo respecto al valor nominal.

Por último, procedemos a analizar la variación del tiempo de cómputo según el tamaño del problema. En este análisis se ha considerado el tiempo medio de resolución de cada instancia, promediado sobre todos los niveles de incertidumbre. Se estudia cómo evoluciona este tiempo al aumentar el número de ítems considerados, con $n \in \{50, 100, 200, 500\}$.

Este comportamiento puede observarse en la Figura 10, donde se muestran los resultados obtenidos para los distintos enfoques estocásticos y robustos. Los datos correspondientes se recogen en las Tablas B.7, B.8 y B.9 del Apéndice B.

En la primera gráfica de la Figura 10 se aprecia que, para los modelos multietapa, el tiempo de cómputo presenta un crecimiento exponencial. Así, en SP4 el tiempo medio alcanza algo más de 50 segundos para $n = 500$, frente a menos de 10 segundos para $n = 200$, o alrededor de 2 segundos para SP3 con $n = 500$. En la variante B, era esperable que el tiempo creciese con el número de variables de segunda etapa; sin embargo, este incremento resulta mucho más moderado. De hecho, las diferencias entre las variantes de SP2 se sitúan en décimas de segundo para $n = 500$, y en centésimas para instancias de $n = 100$ o $n = 50$, lo que confirma la plena tratabilidad computacional en este caso.

En los enfoques robustos, el mayor tiempo se alcanza en torno a $\Gamma = 75\%$. Para $\Gamma = 100\%$ se observa una ligera reducción, ya que el problema se simplifica al resolverse directamente en el peor caso, lo que constituye una especie de “atajo” computacional. No obstante, las diferencias se sitúan únicamente en décimas de segundo.

En conjunto, todas las gráficas de la Figura 10 muestran que el tiempo de cómputo aumenta de manera significativa con el tamaño del problema, reflejando el mayor coste computacional asociado a instancias de mayor dimensión.

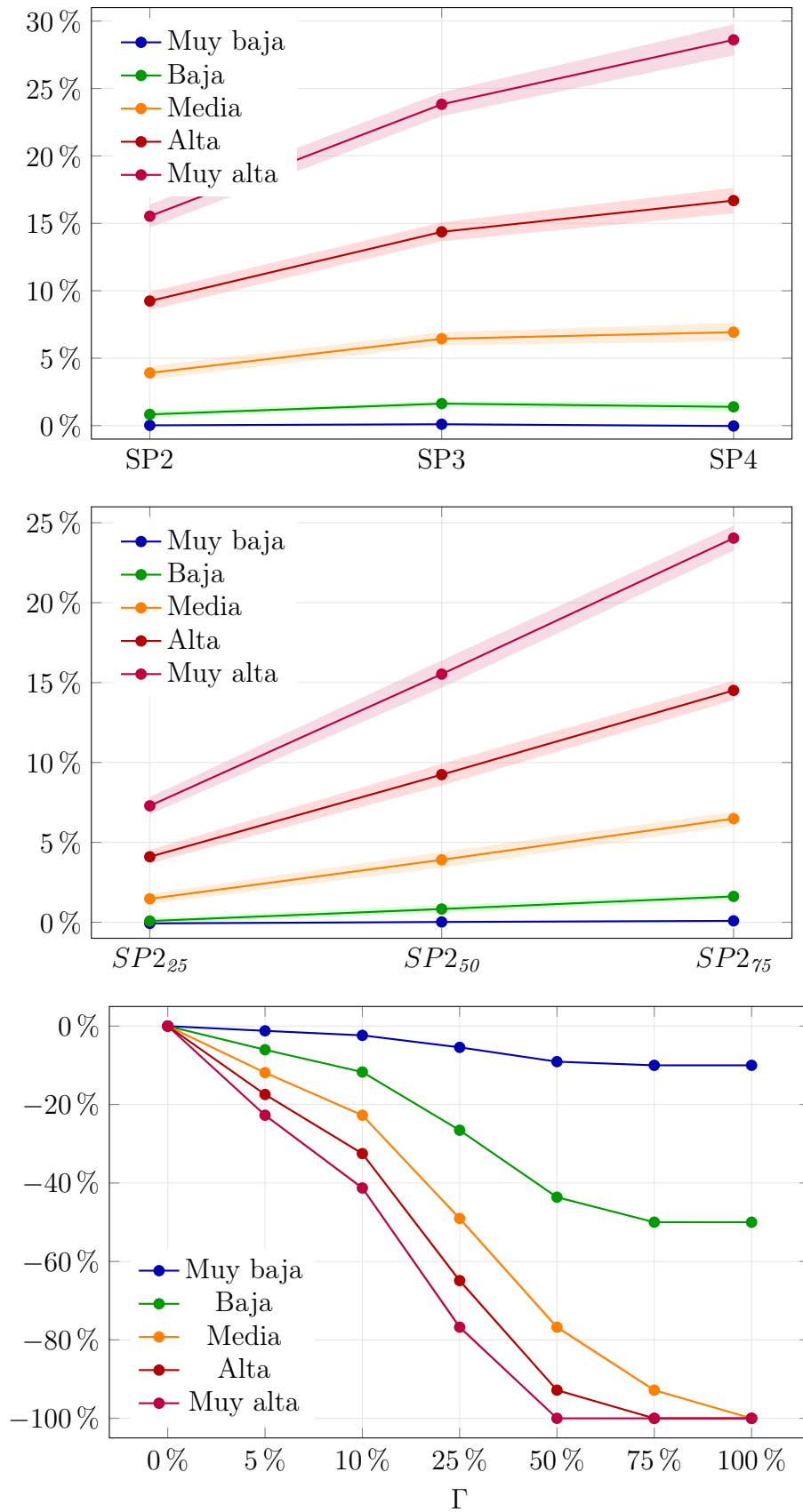


Figura 8: Conjunto de resultados: incrementos medios del valor objetivo frente al nominal (SP y RO) para distintos niveles de incertidumbre, con bandas de IC95 %.

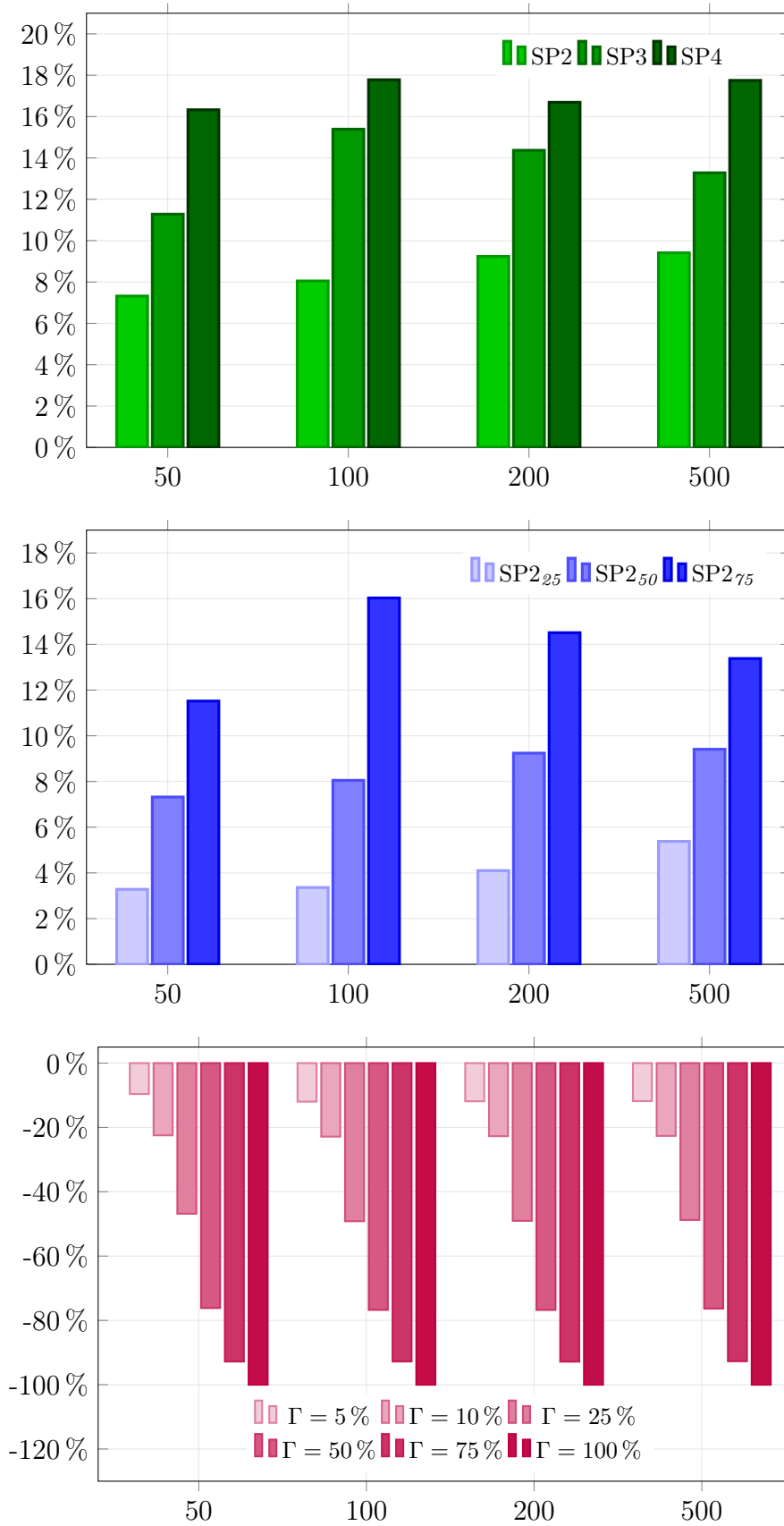


Figura 9: Incremento medio (%) por tamaño n en (a) SP2-SP4, (b) SP2₂₅-SP2₇₅ y (c) RO por Γ .

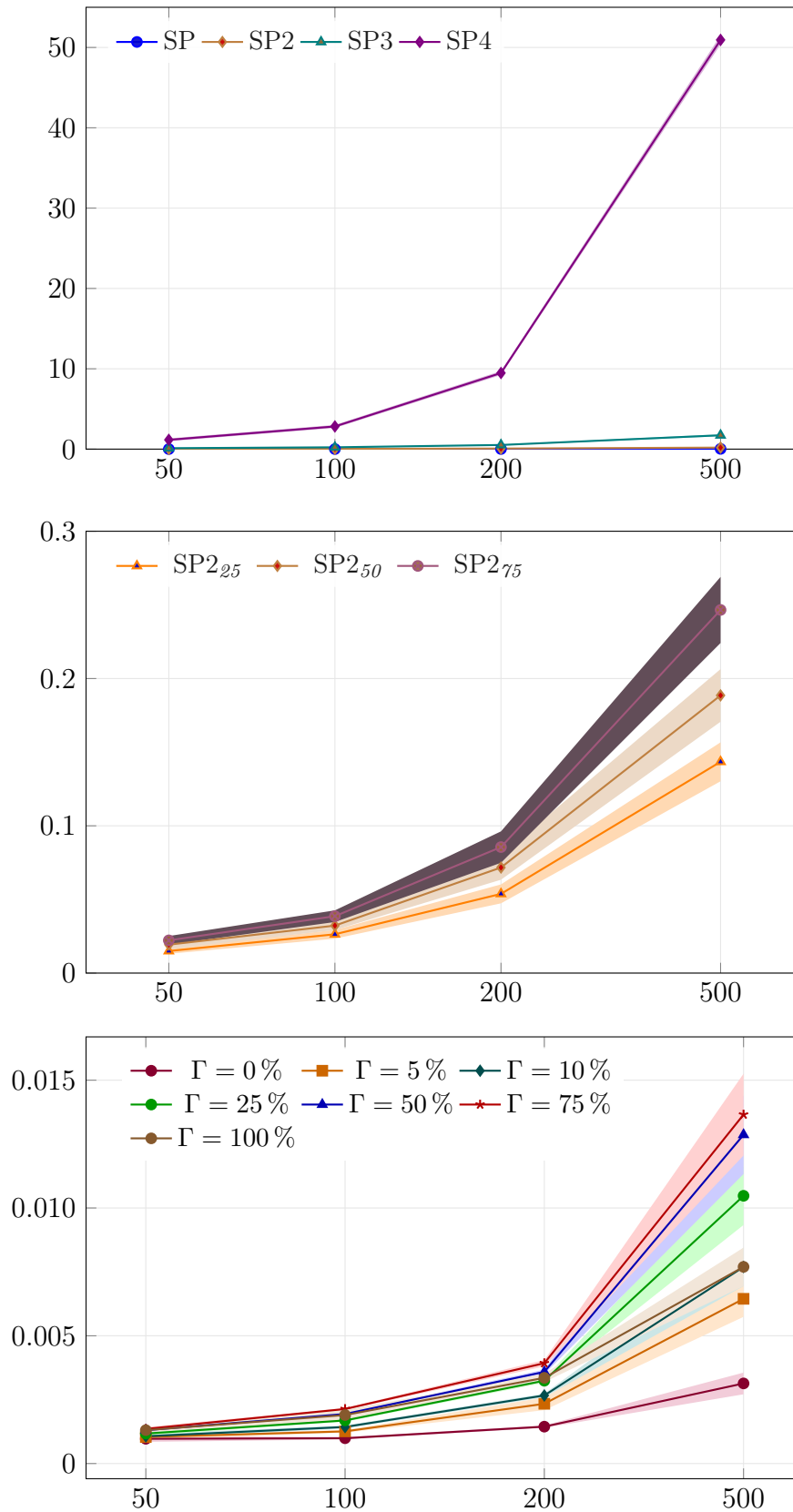


Figura 10: Tiempo de resolución por instancia frente a n : (a) SP/variantes multietapa, (b) SP2₂₅-SP2₇₅ y (c) enfoques robustos por Γ . Cada curva muestra la **media** y las bandas el **IC95 %**.

Capítulo 4

Conclusiones

El objetivo principal de este trabajo ha sido profundizar en la teoría que sustenta los modelos de optimización bajo incertidumbre, concretamente la programación estocástica (SP) y la optimización robusta (RO).

En la parte teórica se ha desarrollado un marco riguroso que formaliza estas metodologías, desde la formulación general de los enfoques estocásticos de una y varias etapas hasta la optimización robusta con distintos conjuntos de incertidumbre, con especial énfasis en el de tipo presupuesto. Un aspecto fundamental ha sido mostrar que, aunque las formulaciones iniciales pueden ser no lineales, siempre se han presentado reformulaciones equivalentes de modo que el problema final es lineal. Como hilo conductor se ha utilizado un ejemplo ilustrativo que permite comprender el impacto de la incertidumbre en las decisiones y los compromisos entre valor esperado, robustez y flexibilidad.

La parte experimental se ha centrado en el problema de la mochila fraccional con utilidades inciertas. Los experimentos han mostrado que el valor esperado de la solución mejora al aumentar la adaptabilidad en los enfoques estocásticos, mientras que en los robustos se mitiga la pérdida relativa respecto al modelo nominal al limitar los parámetros que alcanzan el peor caso. Asimismo, se ha evidenciado el crecimiento exponencial del tiempo de resolución con el número de etapas en SP, confirmando la clásica tensión entre calidad de la solución y coste computacional.

En conjunto, este trabajo subraya la relevancia práctica de la optimización bajo incertidumbre y la necesidad de seleccionar, para cada problema, el enfoque más adecuado según la información disponible, el grado de adaptabilidad, los recursos computacionales y la actitud del decisor frente al riesgo.

Este estudio abre además la puerta a líneas futuras: la extensión a problemas no lineales, la aplicación de enfoques de optimización distribucionalmente robusta (DRO) y el análisis computacional de instancias de mayor dimensión.

Apéndice A

Teoremas

Teorema 1 (Heine–Borel en \mathbb{R}^n). *Sea $K \subset \mathbb{R}^n$ un conjunto. Son equivalentes:*

1. K es compacto.
2. K es cerrado y acotado.

Teorema 2 (Weierstrass o del valor extremo). *Sea $K \subset \mathbb{R}^n$ un conjunto no vacío y compacto, y sea $f : K \rightarrow \mathbb{R}$ continua. Entonces f alcanza sus extremos absolutos en K ; es decir, existen $x_{\min}, x_{\max} \in K$ tales que $f(x_{\min}) = \min_{x \in K} f(x)$ y $f(x_{\max}) = \max_{x \in K} f(x)$.*

Teorema 3 (Ley de los Grandes Números (Versión Débil)). *Sea $\{X_i\}_{i=1}^{\infty}$ una secuencia de variables aleatorias independientes e idénticamente distribuidas con $\mathbb{E}[X_i] = \mu$ y $\text{Var}(X_i) = \sigma^2 < \infty$. Entonces, para cualquier $\epsilon > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{n} \sum_{i=1}^n X_i - \mu \right| \geq \epsilon \right) = 0.$$

Esto es equivalente a

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{p} \mu,$$

donde \xrightarrow{p} denota convergencia en probabilidad.

Teorema 4 (Ley de los Grandes Números (Versión Fuerte)). *Sea $\{X_i\}_{i=1}^{\infty}$ una secuencia de variables aleatorias independientes e idénticamente distribuidas con $\mathbb{E}[X_i] = \mu$. Entonces,*

$$\mathbb{P} \left(\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n X_i = \mu \right) = 1.$$

Esto es equivalente a

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{c.s.} \mu,$$

donde $\xrightarrow{c.s.}$ denota convergencia casi segura.

Teorema 5 (Dualidad fuerte). *Considérese el problema primal (P)*

$$\begin{aligned} & \max_x c'x \\ & \text{s.a. } Ax \leq b, \\ & \quad x \geq 0, \end{aligned}$$

y su dual (D)

$$\begin{aligned} & \min_y b'y \\ & \text{s.a. } A'y \geq c, \\ & \quad y \geq 0. \end{aligned}$$

Si (P) es factible y tiene valor óptimo finito, entonces (D) es factible y ambos admiten solución óptima (x^*, y^*) con igualdad de valores óptimos:

$$c'x^* = b'y^*.$$

Recíprocamente, si (D) es factible y tiene valor óptimo finito, entonces (P) es factible y ambos admiten solución óptima con el mismo valor objetivo. En particular, cuando alguno de los dos problemas tiene valor óptimo finito, no hay brecha de dualidad.

Apéndice B

Tablas

Incertidumbre	Método	Inc. Medio (%)	IC95 %
Muy baja	SP2	0.02 %	[-0.02 %, 0.05 %]
Muy baja	SP3	0.10 %	[0.06 %, 0.14 %]
Muy baja	SP4	-0.03 %	[-0.11 %, 0.04 %]
Baja	SP2	0.83 %	[0.58 %, 1.08 %]
Baja	SP3	1.63 %	[1.43 %, 1.83 %]
Baja	SP4	1.39 %	[1.03 %, 1.75 %]
Media	SP2	3.91 %	[3.42 %, 4.40 %]
Media	SP3	6.44 %	[5.96 %, 6.93 %]
Media	SP4	6.93 %	[6.24 %, 7.61 %]
Alta	SP2	9.24 %	[8.56 %, 9.92 %]
Alta	SP3	14.37 %	[13.67 %, 15.07 %]
Alta	SP4	16.69 %	[15.74 %, 17.63 %]
Muy alta	SP2	15.53 %	[14.68 %, 16.39 %]
Muy alta	SP3	23.83 %	[22.96 %, 24.71 %]
Muy alta	SP4	28.61 %	[27.47 %, 29.76 %]

Tabla B.1: Incremento medio del valor objetivo respecto al nominal e intervalos de confianza de los enfoques SP2–SP4 para los distintos niveles de incertidumbre.

Incertidumbre	Método	Inc. Medio (%)	IC95 %
Muy baja	SP2 ₂₅	-0.07 %	[-0.10 %, -0.03 %]
Muy baja	SP2 ₅₀	0.02 %	[-0.02 %, 0.05 %]
Muy baja	SP2 ₇₅	0.09 %	[0.04 %, 0.14 %]
Baja	SP2 ₂₅	0.08 %	[-0.10 %, 0.27 %]
Baja	SP2 ₅₀	0.83 %	[0.58 %, 1.08 %]
Baja	SP2 ₇₅	1.62 %	[1.40 %, 1.83 %]
Media	SP2 ₂₅	1.47 %	[1.15 %, 1.79 %]
Media	SP2 ₅₀	3.91 %	[3.42 %, 4.40 %]
Media	SP2 ₇₅	6.49 %	[6.06 %, 6.91 %]
Alta	SP2 ₂₅	4.10 %	[3.68 %, 4.53 %]
Alta	SP2 ₅₀	9.24 %	[8.56 %, 9.92 %]
Alta	SP2 ₇₅	14.51 %	[13.90 %, 15.13 %]
Muy alta	SP2 ₂₅	7.29 %	[6.76 %, 7.83 %]
Muy alta	SP2 ₅₀	15.53 %	[14.68 %, 16.39 %]
Muy alta	SP2 ₇₅	24.04 %	[23.25 %, 24.82 %]

Tabla B.2: Incremento medio del valor objetivo respecto al nominal e intervalos de confianza de los enfoques (SP2₂₅, SP2₅₀, SP2₇₅) para los distintos niveles de incertidumbre.

Incertidumbre	% Γ	Inc. Medio (%)	IC95 %
Muy baja	5 %	-1.21 %	[-1.23 %, -1.19 %]
Muy baja	10 %	-2.36 %	[-2.39 %, -2.32 %]
Muy baja	25 %	-5.42 %	[-5.47 %, -5.36 %]
Muy baja	50 %	-9.07 %	[-9.13 %, -9.01 %]
Muy baja	75 %	-10.00 %	[-10.00 %, -10.00 %]
Muy baja	100 %	-10.00 %	[-10.00 %, -10.00 %]
Baja	5 %	-6.03 %	[-6.13 %, -5.94 %]
Baja	10 %	-11.71 %	[-11.88 %, -11.55 %]
Baja	25 %	-26.55 %	[-26.82 %, -26.28 %]
Baja	50 %	-43.62 %	[-43.96 %, -43.29 %]
Baja	75 %	-49.99 %	[-50.00 %, -49.98 %]
Baja	100 %	-50.00 %	[-50.00 %, -50.00 %]
Media	5 %	-11.87 %	[-12.06 %, -11.69 %]
Media	10 %	-22.72 %	[-23.02 %, -22.43 %]
Media	25 %	-49.04 %	[-49.58 %, -48.50 %]
Media	50 %	-76.75 %	[-77.43 %, -76.06 %]
Media	75 %	-92.82 %	[-93.17 %, -92.46 %]
Media	100 %	-100.00 %	[-100.00 %, -100.00 %]
Alta	5 %	-17.43 %	[-17.68 %, -17.19 %]
Alta	10 %	-32.51 %	[-32.89 %, -32.13 %]
Alta	25 %	-64.89 %	[-65.56 %, -64.22 %]
Alta	50 %	-92.82 %	[-93.17 %, -92.46 %]
Alta	75 %	-100.00 %	[-100.00 %, -100.00 %]
Alta	100 %	-100.00 %	[-100.00 %, -100.00 %]
Muy alta	5 %	-22.72 %	[-23.02 %, -22.43 %]
Muy alta	10 %	-41.27 %	[-41.73 %, -40.81 %]
Muy alta	25 %	-76.75 %	[-77.43 %, -76.06 %]
Muy alta	50 %	-100.00 %	[-100.00 %, -100.00 %]
Muy alta	75 %	-100.00 %	[-100.00 %, -100.00 %]
Muy alta	100 %	-100.00 %	[-100.00 %, -100.00 %]

Tabla B.3: Incremento medio del valor objetivo respecto al nominal e intervalos de confianza de los enfoques robustos

n	Método	Inc. Medio (%)	IC95 %
50	SP2	7.32 %	[5.81 %, 8.84 %]
50	SP3	11.28 %	[9.41 %, 13.15 %]
50	SP4	16.33 %	[14.56 %, 18.09 %]
100	SP2	8.05 %	[7.12 %, 8.97 %]
100	SP3	15.39 %	[14.38 %, 16.41 %]
100	SP4	17.78 %	[17.01 %, 18.55 %]
200	SP2	9.24 %	[8.56 %, 9.92 %]
200	SP3	14.37 %	[13.67 %, 15.07 %]
200	SP4	16.69 %	[15.74 %, 17.63 %]
500	SP2	9.41 %	[8.99 %, 9.83 %]
500	SP3	13.28 %	[12.70 %, 13.87 %]
500	SP4	17.75 %	[17.05 %, 18.45 %]

Tabla B.4: Incremento medio del valor objetivo frente al problema nominal e intervalos de confianza al 95 % para SP2–SP4 en distintos tamaños de ítems n .

n	Método	Inc. Medio (%)	IC95 %
50	SP2 ₂₅	3.28 %	[2.56 %, 3.99 %]
50	SP2 ₅₀	7.32 %	[5.81 %, 8.84 %]
50	SP2 ₇₅	11.52 %	[9.98 %, 13.06 %]
100	SP2 ₂₅	3.36 %	[2.76 %, 3.97 %]
100	SP2 ₅₀	8.05 %	[7.12 %, 8.97 %]
100	SP2 ₇₅	16.03 %	[15.28 %, 16.79 %]
200	SP2 ₂₅	4.10 %	[3.68 %, 4.53 %]
200	SP2 ₅₀	9.24 %	[8.56 %, 9.92 %]
200	SP2 ₇₅	14.51 %	[13.90 %, 15.13 %]
500	SP2 ₂₅	5.38 %	[5.08 %, 5.68 %]
500	SP2 ₅₀	9.41 %	[8.99 %, 9.83 %]
500	SP2 ₇₅	13.38 %	[12.82 %, 13.93 %]

Tabla B.5: Incremento medio e intervalos de confianza al 95 % (en porcentaje) para el Experimento B, con SP2₂₅, SP2₅₀ y SP2₇₅ en distintos tamaños de ítems n .

n	Γ	Incremento medio (%)	IC95 %
50	5 %	-9.61 %	[-9.90 %, -9.32 %]
50	10 %	-22.45 %	[-23.03 %, -21.86 %]
50	25 %	-46.83 %	[-47.79 %, -45.87 %]
50	50 %	-76.15 %	[-77.26 %, -75.05 %]
50	75 %	-92.78 %	[-93.58 %, -91.99 %]
50	100 %	-100 %	[-100 %, -100 %]
100	5 %	-11.97 %	[-12.35 %, -11.59 %]
100	10 %	-22.86 %	[-23.51 %, -22.21 %]
100	25 %	-49.16 %	[-50.03 %, -48.29 %]
100	50 %	-76.73 %	[-77.56 %, -75.90 %]
100	75 %	-92.77 %	[-93.35 %, -92.18 %]
100	100 %	-100 %	[-100 %, -100 %]
200	5 %	-11.87 %	[-12.06 %, -11.69 %]
200	10 %	-22.72 %	[-23.02 %, -22.43 %]
200	25 %	-49.04 %	[-49.58 %, -48.50 %]
200	50 %	-76.75 %	[-77.43 %, -76.06 %]
200	75 %	-92.82 %	[-93.17 %, -92.46 %]
200	100 %	-100 %	[-100 %, -100 %]
500	5 %	-11.82 %	[-11.92 %, -11.71 %]
500	10 %	-22.63 %	[-22.81 %, -22.44 %]
500	25 %	-48.77 %	[-49.08 %, -48.46 %]
500	50 %	-76.32 %	[-76.69 %, -75.94 %]
500	75 %	-92.69 %	[-92.97 %, -92.42 %]
500	100 %	-100 %	[-100 %, -100 %]

Tabla B.6: Incremento medio del valor objetivo respecto al nominal para distintos tamaños n y niveles de Γ .

n	Método	Tiempo medio (s)	IC95 %
50	SP	0.0058	[0.0048, 0.0067]
100	SP	0.0075	[0.0068, 0.0083]
200	SP	0.0131	[0.0116, 0.0146]
500	SP	0.0299	[0.0265, 0.0332]
50	SP2	0.0193	[0.0163, 0.0223]
100	SP2	0.0322	[0.0292, 0.0352]
200	SP2	0.0717	[0.0634, 0.0799]
500	SP2	0.1885	[0.1707, 0.2064]
50	SP3	0.1150	[0.0979, 0.1321]
100	SP3	0.2267	[0.2056, 0.2478]
200	SP3	0.5156	[0.4620, 0.5691]
500	SP3	1.7215	[1.5977, 1.8453]
50	SP4	1.1531	[0.9770, 1.3293]
100	SP4	2.8221	[2.6302, 3.0139]
200	SP4	9.4746	[9.2256, 9.7235]
500	SP4	50.9370	[50.3640, 51.5099]

Tabla B.7: Tiempos de cómputo: media e intervalos de confianza al 95 % (IC95 %) en segundos.

n	Método	Tiempo medio (s)	IC95 %
50	SP2 ₂₅	0.0150	[0.0131, 0.0168]
50	SP2 ₅₀	0.0193	[0.0163, 0.0223]
50	SP2 ₇₅	0.0222	[0.0193, 0.0251]
100	SP2 ₂₅	0.0263	[0.0234, 0.0293]
100	SP2 ₅₀	0.0322	[0.0292, 0.0352]
100	SP2 ₇₅	0.0386	[0.0347, 0.0426]
200	SP2 ₂₅	0.0538	[0.0474, 0.0602]
200	SP2 ₅₀	0.0717	[0.0634, 0.0799]
200	SP2 ₇₅	0.0856	[0.0749, 0.0962]
500	SP2 ₂₅	0.1435	[0.1302, 0.1567]
500	SP2 ₅₀	0.1885	[0.1707, 0.2064]
500	SP2 ₇₅	0.2466	[0.2241, 0.2690]

Tabla B.8: Tiempos de cómputo para SP2₂₅, SP2₅₀ y SP2₇₅: medias e intervalos de confianza al 95 % (IC95 %) en segundos para distintos tamaños de ítems n .

n	Γ	Tiempo medio (s)	IC95 %
50	0 %	0.0010	[0.0008, 0.0011]
100	0 %	0.0010	[0.0009, 0.0010]
200	0 %	0.0014	[0.0014, 0.0015]
500	0 %	0.0031	[0.0027, 0.0036]
50	5 %	0.0010	[0.0009, 0.0011]
100	5 %	0.0013	[0.0012, 0.0013]
200	5 %	0.0023	[0.0021, 0.0026]
500	5 %	0.0064	[0.0057, 0.0072]
50	10 %	0.0011	[0.0010, 0.0012]
100	10 %	0.0014	[0.0014, 0.0015]
200	10 %	0.0027	[0.0026, 0.0027]
500	10 %	0.0077	[0.0069, 0.0084]
50	25 %	0.0012	[0.0011, 0.0012]
100	25 %	0.0017	[0.0016, 0.0017]
200	25 %	0.0032	[0.0032, 0.0033]
500	25 %	0.0105	[0.0093, 0.0116]
50	50 %	0.0013	[0.0012, 0.0014]
100	50 %	0.0019	[0.0019, 0.0020]
200	50 %	0.0036	[0.0035, 0.0037]
500	50 %	0.0129	[0.0113, 0.0144]
50	75 %	0.0014	[0.0013, 0.0014]
100	75 %	0.0021	[0.0021, 0.0022]
200	75 %	0.0039	[0.0038, 0.0041]
500	75 %	0.0137	[0.0121, 0.0153]
50	100 %	0.0013	[0.0012, 0.0014]
100	100 %	0.0019	[0.0017, 0.0021]
200	100 %	0.0034	[0.0032, 0.0035]
500	100 %	0.0077	[0.0069, 0.0084]

Tabla B.9: Tiempos de cómputo para los modelos robustos (RO): medias e intervalos de confianza al 95 % (IC95 %) en segundos para distintos tamaños de ítems n y niveles de presupuesto Γ .

Bibliografía

- [1] Forbes. La historia del fatídico accidente del challenger (y las lecciones de liderazgo que se pueden aprender de ella), 2024. Último acceso: 19 de junio de 2025.
- [2] NASA SMA. Safety messages, 2025. Disponible en https://sma-nasa-gov.translate.google.com/news/safety-messages/page/6?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=rq.
- [3] G. Gorton. *Slapped in the face by the invisible hand: Banking and the panic of 2007*. Oxford University Press, 2010.
- [4] U. M. Diwekar. *Introduction to applied optimization*, volume 22. Springer Nature, 2020.
- [5] G. B. Dantzig. *Linear programming and extensions: a report prepared for US Air Force Project RAND*. RAND Corporation, 1963.
- [6] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming: Modeling and theory*. SIAM, 2009.
- [7] B. Póczos and R. Tibshirani. Lecture 16: Penalty methods, 2013. Disponible en <https://www.stat.cmu.edu/~ryantibs/convexopt-F13/scribes/lec16.pdf>.
- [8] G. B. Dantzig. Linear programming under uncertainty. *Management Science*, 1(3-4):197–206, 1955.
- [9] A. Prékopa. Probabilistic programming. In Princeton Symposium, editor, *Mathematical Programming*, pages 113–138. Princeton University Press, 1970.
- [10] R. Tyrrell Rockafellar and Roger J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147, 1991.
- [11] J. R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer, 2 edition, 2011.
- [12] A. Shapiro and A. Philpott. A tutorial on stochastic programming. *Manuscript*, 17, 2007. Disponible en: <http://www2.isye.gatech.edu/ashapiro/publications.html>.
- [13] T. Homem-de Mello and G. Bayraksan. Monte carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, 2014.

- [14] E. L. Porteus. *Foundations of Stochastic Inventory Theory*. Stanford University Press, 2002.
- [15] K. J. Arrow, T. Harris, and J. Marschak. Optimal inventory policy. *Econometrica*, 1951.
- [16] V. Kargin. Life contingency models i, 2018. Disponible en https://www2.math.binghamton.edu/lib/exe/fetch.php/people/kargin/publications/actuarial_1_ln_master.pdf.
- [17] J. E. Smith and R. L. Winkler. The optimizer's curse: Skepticism and postdecision surprise in decision analysis. *Management Science*, 2006.
- [18] D. Z. Long, M. Sim, and M. Zhou. Robust satisficing. *Operations Research*, 2023.
- [19] A. J. Conejo, M. Carrión, and J. M. Morales. *Decision making under uncertainty in electricity markets*, volume 153 of *International Series in Operations Research & Management Science*. Springer, 2010.
- [20] H. E. Scarf. A min-max solution to an inventory problem. Stanford University Press, 1958.
- [21] X. A. Sun and A. J. Conejo. *Robust optimization in electric energy systems*. International Series in Operations Research and Management Science. Springer, 2021.
- [22] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 2004.
- [23] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- [24] A. A. Ahmadi. Lecture 16, 2013. Disponible en https://www.princeton.edu/~aaa/Public/Teaching/ORF523/ORF523_Lec16.pdf.
- [25] D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*. Athena Scientific, Belmont, Massachusetts, 1997.
- [26] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer, Berlin, Heidelberg, 2004.
- [27] G. B. Dantzig. Discrete-variable extremum problems. *Operations Research*, 1957.
- [28] M. Monaci. Exact solution of the robust knapsack problem. *Computers & Operations Research*, 2013.
- [29] T. Koch, M. Lübbecke, S. Wäppler, and A. Martin. Kplib: A library of knapsack problem instances, 2011. Disponible en: <https://github.com/likr/kplib>.
- [30] C. Berrocal Enriquez. Optimization algorithms for uncertain knapsack problem, 2025. Disponible en <https://github.com/carolinaenriqz/Optimization-Algorithms-for-Uncertain-Knapsack-Problem>.
- [31] A. Ben-Tal and A. Nemirovski. Robust optimization—methodology and applications. 2002.