Universidad del País Vasco
Euskal Herriko Unibertsitatea

ZTF-FCT
Zientzia eta Teknologia Fakultatea
Facultad de Ciencia y Tecnología

Two-Stage Stochastic Optimization. An Application in the Third Sector

Final Degree Dissertation
Degree in Mathematics

Sara Laconcepción Moraza

Supervisor:

María Merino Maestre

Leioa, 24^{th} February 2015

Contents

List of Figures

List of Tables

Acknowledgements

Quiero expresar mi agradecimiento a mis padres y a mi hermana, por animarme y apoyarme siempre en todas las decisiones que he tomado. A Ainara, por estar siempre que lo he necesitado, por su apoyo en los momentos más duros y por no dejar que me rinda.

A mi tutora María Merino Maestre, por su ayuda y dedicación; por haberme dado la oportunidad y haber aceptado dirigirme en este Trabajo de Fin de Grado.

Al profesor Fernando Tusell Palmer del departamento de Economía Aplicada III (Econometría y Estadística) de la Facultad de Ciencias Económicas y Empresariales, por facilitarnos el acceso al servidor u107499 del Laboratorio de Economía Cuantitativa (LEC) para poder llevar a cabo la experiencia computacional de la aplicación nutricional de este trabajo.

A Paola Dapena Montes, responsable del proyecto Hazia en la asociación sin ánimo de lucro Sortarazi, por facilitar el acceso al almacén de alimentos y la consulta de información sobre las bases de datos de las familias usuarias del servicio.

A la profesora Ana María Rocandio de Pablo, del área de Nutrición y Bromatología, del departamento de Farmacia y Ciencias de los Alimentos, de la Facultad de Farmacia, por su asesoramiento en los aspectos nutricionales necesarios para realizar este trabajo.

A la profesora Leire Escajedo San Epifanio, investigadora principal del proyecto Universidad-Sociedad US 14/19 "Urban Elika - Elikagaiak denontzat", por su invitación al Congreso Internacional Multidisciplinar "Envisioning a Future without Food Waste and Food Poverty: Societal Challenges", celebrado en noviembre de 2015.

A María del Mar Lledó Sainz de Rozas (subdirectora de Calidad e Innovación) y Amaia Inza Bartolomé, de la Facultad de Relaciones Laborales y Trabajo Social, por facilitar el acceso a material bibliográfico sobre la aplicación de interés social.

A todo el equipo del proyecto de Innovación Educativa HBP-PIE 2014-16 "CTs de Compromiso Ético y Social. Las competencias Transversales de Compromiso Ético y Social en TFG de Responsabilidad Social contra la Pobreza Alimentaria en la CAPV", por el apoyo prestado en la realización de este Trabajo Fin de Grado.

Laburpena

Bizitzako problema gehienak ziurgabetasunaren menpekoak dira. Proiektu honen lehenengo atalean, Programazio Estokastikoaren, bestela esanda, ziurgabetasunaren menpeko Optimizazioaren oinarrizko kontzeptu eta propietateak landu dira. Gainera, hain konplexua da programa estokastikoak konputatzea, hainbat modelo ezberdin ikasi ditugula, esate baterako, *itxaron eta ikusi*, *itxarotako balioa* eta *itxarotako soluzioa itxarotako balioa erabiliz*. Bestetik, *informazio perfektuaren itxaropena* eta *soluzio estokastikoaren balioa* neurriak aurkeztu ditugu, erabaki ahal izateko merezi duen problema estokastikoa ebaztea beste modeloen aurrean. Bigarren zatian, GAMS modelizatzaile eta CPLEX optimizatzailearen bidez produktu ez-galkorren banaketa optimizatzen duen aplikazio bat diseinatu eta inplementatu da, nutrienteen beharrak kostu minimoan bermatzen dituena. Hazia proiektuaren barnean garatu da, Sortarazi elkarteak kudeatuta eta Bizkaiko Elikagai Bankua eta hainbat udalerri bizkaitarren Oinarrizko Zerbitzu Sozialei lotuta.

Resumen

La incertidumbre es inherente a la mayoría de los problemas de la vida real. En la primera parte de la disertación se estudian conceptos y propiedades básicas de Programación Estocástica, también llamada Optimización bajo Incertidumbre. Además, dado que los problemas estocásticos son complejos de resolver, se han presentado otros modelos como *espera y observa*, *el valor esperado* y *el resultado esperado de utilizar la solución que proporciona el valor esperado*. Las medidas del *valor esperado de la información perfecta* y el *valor de la solución estocástica* cuantifican lo que aporta la programación estocástica frente a otro modelos. En la segunda parte, se ha diseñado e implementado con el modelizador GAMS y el optimizador CPLEX una aplicación que optimiza la distribución de productos no perecederos, que garanticen ciertos requerimientos nutricionales al mínimo coste. Se ha desarrollado dentro del proyecto *Hazia*, gestionado por la asociación Sortarazi y vinculado al Banco de Alimentos de Vizcaya y a los Servicios Sociales de Base de varios municipios vizcaínos.

Abstract

It is known that most of the problems applied in the real life present uncertainty. In the first part of the dissertation, basic concepts and properties of the Stochastic Programming have been introduced to the reader, also known as Optimization under Uncertainty. Moreover, since stochastic programs are complex to compute, we have presented some other models such as *wait-and-wee*, *expected value* and *the expected result of using expected value*. *The expected value of perfect information* and *the value of stochastic solution* measures quantify how worthy the Stochastic Programming is, with respect to the other models. In the second part, it has been designed and implemented with the modeller GAMS and the optimizer CPLEX an application that optimizes the distribution of non-perishable products, guaranteeing some nutritional requirements with minimum cost. It has been developed within *Hazia* project, managed by Sortarazi association and associated with Food Bank of Biscay and Basic Social Services of several districts of Biscay.

Chapter 1

Introduction

1.1 Social and Mathematical Motivation

There is not a global definition of food waste and food poverty. The European Commission, the Food & Agricultural Organization of the United Nation (FAO), USDA's Economic Research Service, Smil, the UK by the Waste & Resources Action Programme (WRAP), the Barilla Center of Food and Nutrition and much more professional bodies or Community institutions define waste of food in different ways. However, they all have in common the awareness of the magnitude of the problem and the importance of reduction; at a time that even though it is produced enough food to feed the whole population, 870 million people live in hunger (FAO) and a third of all the food produced, 1.3 billion tones, 180kg per capita in only 27 european countries are wasted every year, see Gjerris & Gaiani 2015 [13] and González Vaqué 2015 [14]. We know that the problem of hunger is, above all, a result of war and massive displacement of refugees. However, there are more and more poor and malnourished families in rich countries due to the lack of access to resources and the inefficiency of the food chain. Spain is the sixth on the list of countries in the European Union that waste more food, around 7.7 million tones, 18% of what is bought by the population for own feeding (FAO). Food waste has numerous causes at every level: overproduction, deterioration, imperfect size/shape of the product or its packaging and problem of appearance or defective packaging and inadequate stock management (marketing rules), among others.

Whether just one-forth of the wasted food could be saved, around 900 million hungry people in the world would be feed. That is the reasoning behind the necessity of reaction and strategies for solutions in order to prevent and avoid food waste:

- Redistribute unsold and discarded products to citizens below the minimum income.
- Reeducate the citizenship providing tips, recipes, messages and graphics (household food waste makes up almost half of all food waste in UK).
- Improve the efficiency of the food supply chain by promoting direct relations between producers and consumers.

- Improve logistic, transport, stock management and packaging, since some food products are produced, transformed and consumed in very different parts of the world.

That is why it should be given preference to agricultural and food products produced as near as possible to the place of consumption. The role of Food Banks is essential in the use of discarded food. Spain is the first european country in Food Bank activities: there are 54, delivering million kilograms of food every year.

Clearly, in view of such a magnitude problem, the cooperation of professionals from various disciplines is welcomed. Multidisciplinary domains can provide solutions to help improving more disadvantage people life condition. Particularly, there are several problems to adress by the area of Operations Research & Management Science. In this project we wanted to focus on optimizing the problem of the distribution of aliments which satifies minimum nutritional requirements.

Diet Problem is considered one of the first problems of linear programming, Stigler 1945 [30]. It was laid out with the intention of optimizing the cost of the soldiers diet before finishing that year the Second World War in USA.

Later on, it was proposed as an alternative the stochastic programming, where the model includes uncertain parameters and some of the decisions must be taken before unceertainty is revealed. See in Vitoriano et al. 2013 [32] a recent book about decision models in disaster management and humanitarian emergency.

1.2 Aims of the project

In this project we will consider two main goals. On one hand, the study of two-stage Stochastic Programming basis, focusing in concepts such as Stochastic Problem (SP), Wait-and-See (WS), Expected Value (EV), Expected result of using Expected Value (EEV) and their relations, as well as the measures Expected Value of Perfect Information (EVPI) and Value of Stochastic Solution (VSS).

On the other hand, the design and implementation of an application in the third sector, also known as social economy. We are interested to supply monthly around 900 people, reaching almost all of the nutritional requirement with only non-perishable products, with or without exceeding a budget. Particularly, we will distinguish four models: (1) minimize cost for feeding all these people with a healthy diet, that is, supplying all the nutritional requirement, (2) improve the percentage reached nowadays by the provision of the association Sortarazi, without exceeding the budget, (3) and (4) previous models but guaranteeing a balanced diet.

1.3 Scientific literature

The theory and the applications of the Stochastic Programming are progressing significantly, which is reflected in the number of nowadays publications. Dantzig 1995 [8] and Beale 1955 [3] are considered the origin of the Optimization under Uncertainty. Some of the fundamental books are Kall et al. 1988 [17], Kall & Wallace 1994 [18], Prekopa 1955 [25], Wallace & Zeimba (eds.) 2005 [33], Shapiro et al. 2009 [29], Birge & Louveaux 2011 [5] and King & Wallace 2012 [21].

One of the first applications of stochastic programming was related to airline planning: a decision on the allocation of aircraft to routes, developed in Ferguson & Dantzig 1956 [10], also collected in King 1988 [20].

Stochastic programming has been applied to a wide variety of areas, such as Production planning which is a major area worth mentioning. We could get good explanations for manufacturing production or machine capacity planning problems, production or machine scheduling and hydrothermal power production among others, see [1], Klein Haneveld & Van der Vlerk 2001 [22].

In the **financial area** there are lots of models with uncertain parameters, which is a good reason for stochastic modeling. We can see many examples such as asset liability management, an option selection model and macroeconomic modeling and planning and network models, among others, see Gassman & Ziemba 2012 [12].

According to **expansion and planning problems** we can assume some examples related to energy planning which has been the focus of many stochastic programming studies such as electricity generation capacity and dairy farm expansion planning (first appeared in deterministic form in Swart et al. 1975 [31]; now, we can find it well explained in stochastic form in Birge & Louveaux 2011 [5]), among others.

Stochastic programming has been applied in many other areas such as sports, design of a multistage truss, traffic assignment, telecommunications, climate change, forestry planning model, the hospital staffing problem, see Kao & Queyranne 1985 [19] and lake level management among others, King 1988 [20], Gassman & Ziemba 2012 [12] and the collection Wallace & Ziemba 2005 [33]. For more information, we can visit the Web Side of the Stochastic Programming Society (SPS) [40].

1.4 Organization of the project

This project is organized as follows: Chapter 2 defines and compares deterministic and stochastic programming, where some basic concepts and properties of the theory of Stochastic Optimization are introduced, also known as Optimization under Undertainty, and ilustrated by its corresponding examples.

Chapter 3 shows some alternative models, known as the wait-and-see, expected value and expected result of using expected value. The expected value of perfect information and value of stochastic solution measures are introduced and some basic inequalities and the relationship between them are given.

In Chapter 4 we have developed an application for the third sector. There is explained the context of the realistic problem, the diet stochastic model and alternative models, the datasets are detailed and the solutions and analysis of the models are explained. Chapter 5 concludes.

In Appendix A and B are shown the GAMS codes implemented corresponding to computational experiences given along the whole project. Finally, the bibliography is presented.

Chapter 2

Optimization Models under Uncertainty

In this chapter we will present and compare deterministic and stochastic programming. In Subsections 2.2.1, 2.2.2 and 2.2.3 there are explained some basic concepts of the theory of Stochastic Optimization, such as probability spaces and random variables, decisions and recourses and non-anticipativity principle. Subsections 2.2.4 and 2.2.5 are focused in two-stage models representations. The examples code is detailed in Appendix A.

2.1 Deterministic Linear Programming

A deterministic linear problem consists of finding a solution that minimizes (or maximizes) a linear function (the objective function), subject to a set of linear constraints, taking into account the certainty of all the parameters. The problem reads as follows:

$$\begin{aligned} Z \quad = \quad \min \quad & c_1X_1 + c_2X_2 + ... + c_nX_n \\ \text{subject to} \quad & \underline{b}_1 \leq a_{11}X_1 + a_{12}X_2 + ... + a_{1n}X_n \leq \bar{b}_1 \\ & \underline{b}_2 \leq a_{21}X_1 + a_{22}X_2 + ... + a_{2n}X_n \leq \bar{b}_2 \\ & \qquad \vdots \\ & \underline{b}_m \leq a_{m1}X_1 + a_{m2}X_2 + ... + a_{mn}X_n \leq \bar{b}_m \\ & X_1, X_2, ..., X_n \geq 0 \end{aligned} \tag{2.1}$$

Hence, if we use the matricial notation, we can express it in this way:

$$\begin{aligned} Z \quad = \quad \min \quad & cX \\ \text{s.t.} \quad & \underline{b} \leq AX \leq \overline{b} \\ & X \geq 0 \end{aligned} \tag{2.2}$$

where X is the decision vector with dimension $n \times 1$ and c, A, $\underline{b}$ and $\overline{b}$ are known data: c is a $1 \times n$ vector of costs, $A \in M_{m \times n}$ is the constraints matrix and $\underline{b}$ and $\overline{b}$ are *left hand side* (LHS) and *right hand side* (RHS), respectively, the vectors of independent items of the constraints of sizes $m \times 1$.

Besides, $z = cX$ is the objective function, while $\{X \mid \underline{b} \leq AX \leq \overline{b}, X \geq 0\}$ defines the set of feasible solutions. A feasible solution X^* is optimal if $cX \geq cX^*$ for any feasible X. Linear programs usually try to find solutions with minimum cost over linear constraints of demand or maximum profit over a situation with limited resources. Since maximizing an objective function z is equivalent to minimizing $-z$, without loss of generality, in this project we will deal with minimization problems.

Example 2.1. Let us consider the following diet problem addapted from NEOS server [35]. The goal of the problem is to select a set of aliments that will satisfy monthly nutritional requirements of 20-49 years old 1000 women at minimum cost. The problem corresponds to the supply of products in a restaurant in order to serve all those clients. We will consider that half of the requirements ($\alpha_1 = 0.5$) must be satisfied with the products at the first day of the month and the rest of them will be purchased after the first two weeks. The problem is formulated as a linear program with the goal of minimizing the cost and the constraints are stated to satisfy the specified nutritional requirements. For the sake of simplification, we will assume that there are two products available (pasta and lentils) and two nutritional requirements (iron and energy) where the cost and nutrients are defined in Table 2.1 and the requirements in Table 2.2. In order to guarantee feasibility, we have relaxed the maximum requirement of iron and reduced the minimum in 5% with respect to the recommendation, see Carbajal 2013 [6].

Table 2.1: Cost and nutrients per aliment

Products	Cost/product (€)	Iron (mg)	Energy (kcal)
Pasta (1kg)	1.98	18.00	3530
Lentils (1kg)	1.58	68.74	3100

Table 2.2: Nutrient requirements per day and person

Nutrients	Minimum	Maximum
Iron (mg)	9.5	-
Energy (kcal)	2185.0	3000

Let us define the variables of the model:

- X_i: amount of product i to be purchased at the first day of the month, $i \in \{1, 2\}$
- Y_i: amount of pruduct i to be purchased after two weeks, $i \in \{1, 2\}$

The Diet Problem can be modeled as follows:

$$\begin{aligned} min \quad & 1.98(X_1 + Y_1) + 1.58(X_2 + Y_2) \\ \text{s.t.} \quad & 142.5 \leq 18X_1 + 68.737X_2 \\ & 32775 \leq 3530X_1 + 3100X_2 \leq 45000 \\ & 285 \leq 18(X_1 + Y_1) + 68.737(X_2 + Y_2) \\ & 65550 \leq 3530(X_1 + Y_1) + 3100(X_2 + Y_2) \leq 90000 \\ & X_i,\ Y_i \geq 0,\ i \in \{1, 2\} \end{aligned}$$

Notice that the bounds are given for 1 month (30 days) and in thousand units (1000 clients). Thus, the optimal solution of this problem is:

$$X^* = (X_1, X_2, Y_1, Y_2) = (0,\ 10573,\ 0,\ 10573) \quad \text{and} \quad Z^* = 33474€$$

This means that we should buy at first day 10573 packages of lentils and purchase two weeks later 10573, with a total cost of 33474 €.

2.2 Stochastic Linear Programming

Most of the optimization problems applied in the real life present uncertain data: production costs and transport depend on fuel price, future demands depend on the uncertain market conditions or crop returns depend on the weather, among others. If we suppose that all the parameters are known, it could be produced not satisfactory result, or even disastrous. Hence, it seems more accurate to model the optimization problems taking into account unknown parameters (unknown by the decisor at the moment of making decisions and out of his/her control). In fact, Stochastic Programming is an alternative to the deterministic problems.

Stochastic linear problems are those linear optimization problems where some of the parameters c, A, $\underline{b}$ and $\overline{b}$ of the model (2.2) are uncertain. So, uncertainty can be defined by random variables in the form of probability distributions, densities or, in general, probability measures.

Definition 2.1. A **stage** of a given planning horizon is a set of consecutive time periods where the realization of one or more stochastic (i.e., uncertain) events take place. At the end of a stage, decisions are taken, considering the specific outcomes of the stochastic events of this and previous stages.

Stochastic programs may be classified according to the amount of stages: *two-stage* problem is composed by two stages and those which has three or more stages are called *multistage* problems.

2.2.1 Probability Spaces and Random Variables

Now, we will describe probabilistic concepts assummed in the progress of the project, which are essential to understand the structure of a stochastic problem.

We will consider the technique called *analysis of scenarios* for modeling uncertainty. This methodology consists of knowing a finite set of values of the stochastic parameters with their corresponding likelihood. That is, the goal of this method is to define a future state of a system known in the present (at least partially) and show the different processes which pass from the present to the future. This situation happens in strategic problems, where possible results are obtained by the opinion of experts and where there are only a discrete and finite number of scenarios.

Definition 2.2. A **scenario** is a realization of the uncertain and deterministic parameters of the model from the first stage until the last one. It can also be defined as the representation of the possible evolution of a system through the future. The scenario will show the hypotetical situation of each constitutive parameter of a system for each period in a particular horizon planning.

Uncertainty is usually characterized by a probability distribution on the random parameters. It can be represented in terms of random experiments where all possible outcomes are denoted by ω and the set of all of them by Ω. The outcomes can be combined in subsets called *events*. Each event $\omega \in \Omega$ determines a *scenario*: $\xi^\omega = (c^\omega, A^\omega, \underline{b}^\omega, \overline{b}^\omega)$ and Ξ is the set of all the scenarios. The collection of random events is denoted by $\mathcal{F}$, which is a tribu or σ-algebra of the parts of Ω. Finally, let define *probability* as an aplication $P : \mathcal{F} \to [0,1]$ so that $P(\Omega) = 1$ and $P(\cup_{n\geq 1} A_n) = \sum_{n\geq 1} P(A_n)$ where $\forall A_i, A_j \in \mathcal{F} : A_i \cap A_j = \emptyset, i \neq j$. The triplet $(\Omega, \mathcal{F}, P)$ is called *probability space*.

For this project we will mainly consider discret variables, where the random variables, $\boldsymbol{\xi}$, take a finite number of values, $\xi^\omega, \omega \in \Omega$ with probability $P(\boldsymbol{\xi} = \xi^\omega) = p^\omega$ so that $\sum_{\omega\in\Omega} p^\omega = 1$. *Cumulative distribution* is defined as $F(\xi) = P(\{\omega \in \Omega | \boldsymbol{\xi} \leq \xi\}) = P(\boldsymbol{\xi} \leq \xi)$. Besides, *expectation* of a random variable can be calculated as $E[\boldsymbol{\xi}] = \sum_{\omega\in\Omega} p^\omega \xi^\omega$ and variance is $Var[\boldsymbol{\xi}] = E[\boldsymbol{\xi} - E[\boldsymbol{\xi}]]^2$.

Here and subsequently, for simplicity of notation, we will use the symbol ω to denote a scenario, instead of ξ^ω and Ω, rather than Ξ as the set of scenarios.

The set of scenarios is usually represented by a tree[1], so called *scenario tree*, whose levels correspond to stages, that is, the different periods of the planning horizon where it is necessary to make a decision. A scenario tree is the representation of a set of scenarios and each branch of the tree will be a possible evolution of the system. That is, scenario trees are used to represent history of decision making.

Although from now on we are going to deal only with two-stage programs, we will show a multistage tree (corresponding to a multistage problem), in order to understand some basic concepts. In Figure 2.1 is ilustrated a tree with 4 stages, $\mathcal{T} = \{1, 2, 3, 4\}$ and 8 scenarios, $\Omega = \{1, 2, ...8\}$, $|\Omega| = 8$, with 18 nodes, $\mathcal{G} = \{1, 2, ...18\}$.

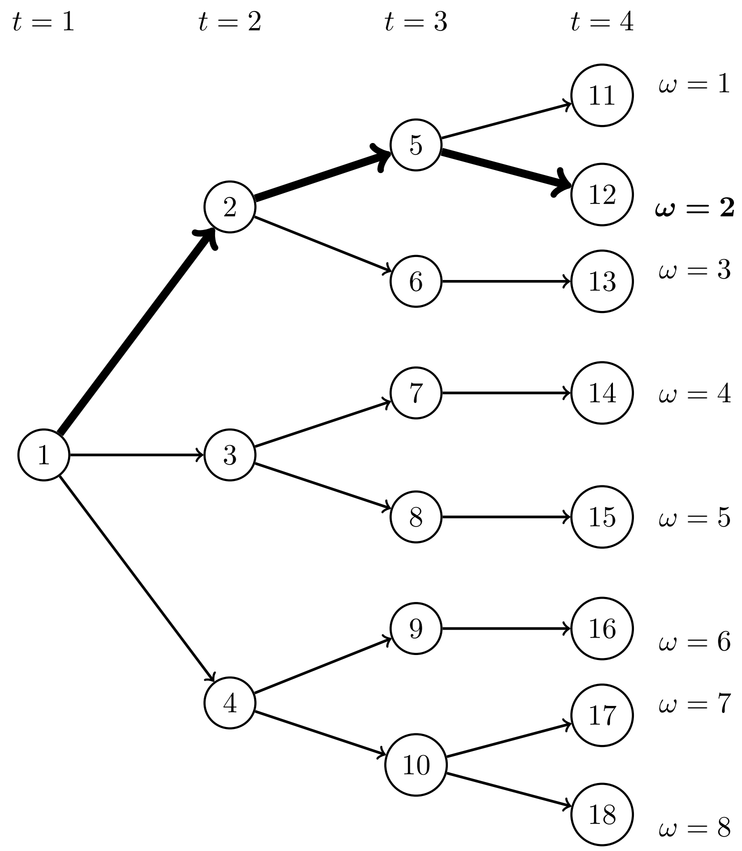


Figure 2.1: Example of a scenario tree

[1]In graph theory, a graph is defined as a representation of a set of objects where some pairs of objects are connected. These objects are called *nodes*, also known as vertices, and the connections, edges. There are many types of graphs, among others, undirected graphs (whose edges have no orientation). There are some important classes of graphs as connected graphs. This is an undirected graph in which every unordered pair of vertices in the graph is connected. Thus, a *tree* is a connected graph with no cycles.

In this tree we can see that the number of nodes in each post-node is equivalent to the realizations of uncertain parameters and in each stage there is enough information to make a decision. Notice that in the first stage there is only one node, called *root node*. A scenario is not one of each possible last state; but, as it can be seen in Figure 2.1, for example, $\omega = 2$ is one of the possible evolution of the system from the first stage to the last hypothetical one. Thus, each path from the root to a leaf is called scenario, a feasible realization of the uncertainty. In each node there is a variable to decide, that is, a decision that has to be made.

Once that the scenario tree is generated, it is necessary to extend the problem modelling, in such a way that it will take the information from that tree. One option consists of solving the deterministic problems according to each scenario $\omega \in \Omega$:

$$\begin{aligned} Z^\omega \quad = \quad \min \quad & c^\omega X^\omega \\ \text{s.t.} \quad & \underline{b}^\omega \leq A^\omega X^\omega \leq \overline{b}^\omega \\ & X^\omega \geq 0 \end{aligned} \tag{2.3}$$

Based on the the previous model (2.3), the way of choosing an optimal solution is not clear. There can be feasible solutions in a scenario that could be non-feasible in another one.

However, the *analysis of scenarios* applied in a optimization problem provides feasible solutions under all the scenarios and optimal expected value over all of them. This happens, as we will see later, due to the optimization of a linear combination of objective functions according to the set of scenarios.

2.2.2 Decisions and recourses

One of the most attractive aspects of the Stochastic Programming is the fact of including changes in the decisions to be taken, whenever information is available throughout the planning horizon. Furthermore, it has sense that, at the beginnig of a process with several decision stages, the first stage decisions must be taken. However, it does not have to happen the same with the decisions of the other stages.

Definition 2.3. A solution is **anticipative** if there is a unique value for each variable, implementable and independent from the random experiment. That is, decisions which must be taken before the uncertainties are resolved. On the contrary decisions that are taken after uncertainty in the parameters has been resolved are called **adaptative** variables.

All stochastic programs have some anticipative variables, since they would otherwise become deterministic. Stochastic programs that include both types of variables are generally called recourse models and according to the decision anticipativity, there is the following classification:

- **Simple recourse model** is that where all the decisions to take have to be fixed from the beginning, without any variation even though in the following periods of time more information of each scenario can be reached. All the decisions have to be taken before the random experiment, and they belong to anticipative variables. This makes easier the representation and resolution of the model. However, conclusions can be far from the real results, since new information reached in each stage is not used.

- **Relatively complete recourse** is the model where decisions of the first r stages are determined in the beginning (implementable periods), and the decisions of the rest are adjusted to the possible changes (non-implementable periods).

- **Complete recourse model** is that where all the decisions are adapted along the time, every time that information of the uncertain parameters is revealed; except for the variables of the first stage, which do not depend on the scenario that happens. That is to say that the solution is formed by a set of only decisions for the first period and an optimal decision for each scenario. Complete recourse is often added to a model to ensure that no outcome can produce infeasible results. It is the most interesting and useful model because the solution provided is a set of decisions which get adapted to the information disposed in each stage, allowing changes and optimizing the expected value of the objective function.

All the stochastic models used in this project are complete recourse models.

2.2.3 Non-anticipativity principle

The principle of non-anticipativity (NAC) was introduced for two-stage problems in 1974 by Wets, see [34] and restated by Rockafellar and Wets in 1991, see [27]. It states that if two scenarios, ω and ω', are equal according to the available information from the first to the r^{th} stage, then the decisions to take from those scenarios until this last stage have to be the same. Now, let us represent in Figure 2.8 the *non-anticipativity principle* according to the multistage example corresponding to the tree in previous Figure 2.1 where decision variables in nodes inside the same dashed ellipse must be the same. That is, there are six sets of variables that must be the same.

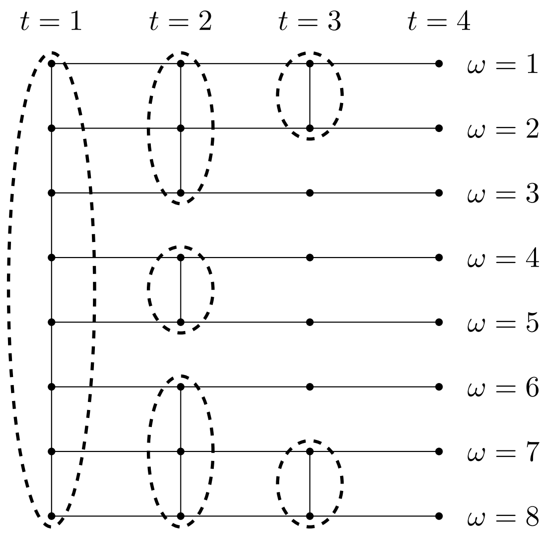


Figure 2.2: *Non-ancitipativity* principle

All the decisions taken in the first stage must be the same under all the scenarios. At the third stage, the decisions of first and second scenarios and seventh and eighth scenarios, must be the same, respectively.

According to the implicit or explicit representation of the NAC constraints, the models can be defined in compact or splitting variable representation.

2.2.4 Two-Stage Model in Compact Representation

Stochastic problems were formulated for the first time in 1955 derivated from the linear optimization by Dantzig and Beale in [8] and [3], respectively. First stage is made up of a common node with the same information, while the second one consists of one node for each scenario. Therefore, first stage decisions will be the same and independent from the scenario, while second decisions are non-anticipative and depend on the scenario.

Let us consider the next two-stage stochastic linear program, where objective function is minimized:

$$\begin{aligned}
SP \quad = \quad \min \quad & cX + E[\min \mathbf{q}\ \mathbf{Y}] && (2.4)\\
\text{s.t.} \quad & \underline{b} \leq AX \leq \overline{b} && (2.5)\\
& \underline{h}^{\omega} \leq T^{\omega}X + W^{\omega}Y^{\omega} \leq \overline{h}^{\omega}, \qquad \forall \omega \in \Omega && (2.6)\\
& X, Y^{\omega} \geq 0, \qquad \forall \omega \in \Omega && (2.7)
\end{aligned}$$

where the $n_1 \times 1$ random vector X represents the first-stage decision vector (decision that has to be taken before the experiment, also called *here-and-now* solution) and c, $\underline{b}$, $\overline{b}$ and A are the first-stage known vectors and matrices corresponding to X, of sizes $1 \times n_1$, $m_1 \times 1$, $m_1 \times 1$ and $m_1 \times n_1$, respectively. c is the row vector composed by the objective function coefficients, $\underline{b}$ and $\overline{b}$ are the column vectors with independent items of the constraints, (known as *left hand side*, LHS and *right hand side*, RHS, respectively) and A is the matrix of the constraints. On the contrary, $Y = (Y^{\omega})$ is the second stage decision vector (the one that can be taken after the experiment) and in this stage some random events $\omega \in \Omega$ can happen. For each realization ω, $\underline{h}^{\omega}$ and $\overline{h}^{\omega}$ with size $m_2 \times 1$ both of them, are the LHS and RHS column vectors, $q = (q^{\omega})$, q^{ω} $1 \times n_2$ vector corresponds to the objective function coefficients and the *technological matrix* $T = (T^{\omega})$, $T^{\omega} \in M_{m_2 \times n_1}$. There is also a matrix called *recourse matrix*, $W^{\omega} \in M_{m_2 \times n_2}$ associated to the recourse Y^{ω} variables.

Let us define $(X_{SP}, (Y^{\omega}_{SP})_{\omega \in \Omega})$ the optimal solution of (2.4)-(2.7) and $Z^{\omega} = cX_{SP} + q^{\omega}Y^{\omega}_{SP}$, then the objective function (2.4) corresponds to the expected value of the random variable of costs $Z = (Z^{\omega})_{\omega \in \Omega}$.

$$E[Z] = \sum_{\omega \in \Omega} P(Z = Z^{\omega})Z^{\omega} = \sum_{\omega \in \Omega} p^{\omega} Z^{\omega} = \sum_{\omega \in \Omega} p^{\omega}(cX_{SP} + q^{\omega}Y^{\omega}_{SP}) =$$

$$= cX_{SP} \sum_{\omega \in \Omega} p^{\omega} + \sum_{\omega \in \Omega} p^{\omega} q^{\omega} Y^{\omega}_{SP} = cX_{SP} + \sum_{\omega \in \Omega} p^{\omega} q^{\omega} Y^{\omega}_{SP}$$

This problem is equivalent to the *Deterministic Equivalent Program* (DEP):

$$\begin{aligned}
\min \quad & cX + E[Q(X, \omega)]\\
\text{s.t.} \quad & \underline{b} \leq AX \leq \overline{b} \qquad (2.8)\\
& X \geq 0
\end{aligned}$$

where

$$
\begin{aligned}
Q(X,\omega) \;=\; \min \quad & q^\omega Y \\
\text{s.t.} \quad & \underline{h}^\omega - T^\omega X \le W^\omega Y \le \overline{h}^\omega - T^\omega X, \quad \forall \omega \in \Omega \\
& Y \ge 0
\end{aligned}
\tag{2.9}
$$

The two-stage stochastic linear program, defined in (2.4) - (2.7), is in *Compact Representation.* In this way, we will use the following variables:

- X, variable vector of first stage
- Y^ω, variable vector of second stage, for each scenario $\omega \in \Omega$

The decisions along the scenario tree and the matrix structure are shown in Figure 2.3 and 2.4, respectively.

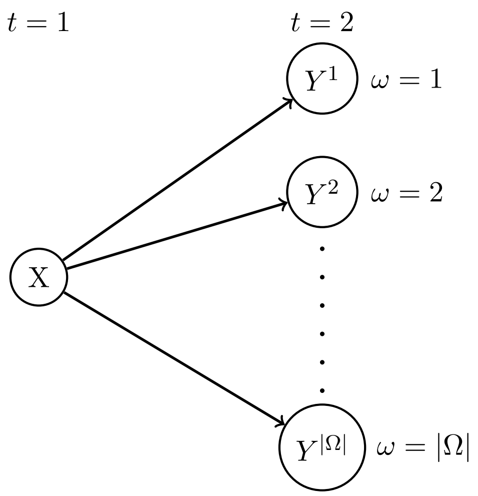


Figure 2.3: Scenario tree in compact representation

$$
\begin{array}{ccccccc}
 & X & Y^1 & Y^2 & \cdots & Y^{|\Omega|} & \\
\underline{b} & \boxed{A} & & & & & \overline{b} \\
\underline{h}^1 & \boxed{T^1} & \boxed{W^1} & & & & \overline{h}^1 \\
\underline{h}^2 & \boxed{T^2} & & \boxed{W^2} & & & \overline{h}^2 \\
\vdots & \vdots & & & \ddots & & \vdots \\
\underline{h}^{|\Omega|} & \boxed{T^{|\Omega|}} & & & & \boxed{W^{|\Omega|}} & \overline{h}^{|\Omega|}
\end{array}
$$

Figure 2.4: Matrix structure in compact representation

2.2.5 Two-Stage Model in Splitting Variable Representation

Alternatively, the *Splitting Variable Representation* of the two-stage stochastic linear program in given in (2.10)-(2.14).

$$
\begin{aligned}
SP \quad = \quad \min \quad & \sum_{\omega \in \Omega} p^{\omega} \left(cX^{\omega} + q^{\omega} Y^{\omega} \right) && && (2.10)\\
\text{s.t.} \quad & \underline{b} \leq AX^{\omega} \leq \overline{b}, && \forall \omega \in \Omega && (2.11)\\
& \underline{h}^{\omega} \leq T^{\omega} X^{\omega} + W^{\omega} Y^{\omega} \leq \overline{h}^{\omega}, && \forall \omega \in \Omega && (2.12)\\
& X^{\omega} = X^{\omega^*}, && \forall \omega \neq \omega^*,\ \omega, \omega^* \in \Omega && (2.13)\\
& X^{\omega}, Y^{\omega} \geq 0, && \forall \omega \in \Omega && (2.14)
\end{aligned}
$$

where one copy of X, first stage variable vector, is considered for each scenario $\omega \in \Omega$, and therefore, non-anticipativity constraints (2.13) are added explicitly, see Rockafellar and Wets [26].

The decisions along the scenario tree and the matrix structure are shown in Figure 2.5 and 2.6, respectively.

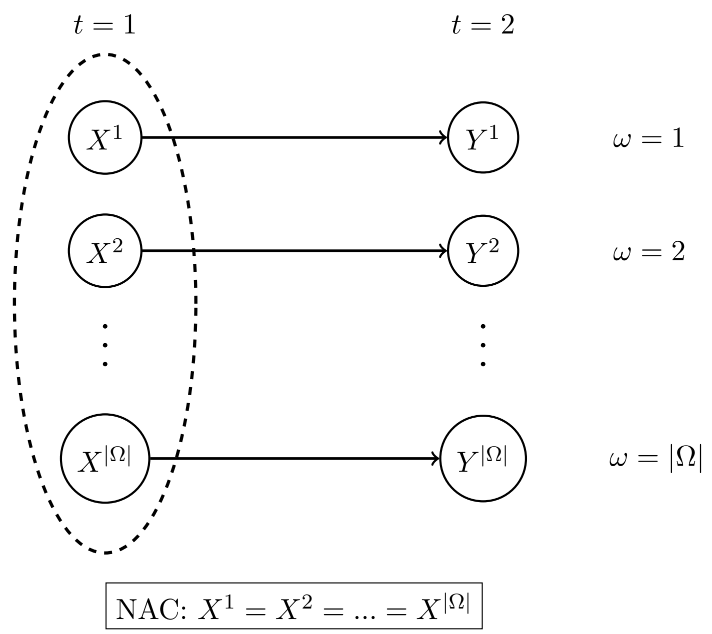


Figure 2.5: Scenario tree in splitting variable representation

$$
\begin{array}{cccccccccc}
 & X^1 & Y^1 & X^2 & Y^2 & \cdots & X^{|\Omega|} & Y^{|\Omega|} & \\
\underline{b} & \boxed{A} & & & & & & & \bar{b} \\
\underline{h}^1 & \boxed{T^1} & \boxed{W^1} & & & & & & \bar{h}^1 \\
\underline{b} & & & \boxed{A} & & & & & \bar{b} \\
\underline{h}^2 & & & \boxed{T^2} & \boxed{W^2} & & & & \bar{h}^2 \\
\vdots & & & & & \ddots & & & \vdots \\
\underline{b} & & & & & & \boxed{A} & & \bar{b} \\
\underline{h}^{|\Omega|} & & & & & & \boxed{T^{|\Omega|}} & \boxed{W^{|\Omega|}} & \bar{h}^{|\Omega|} \\
 & & & & & & & & \\
 & \boxed{I} & & \boxed{-I} & & & & & 0 \\
 & & & \boxed{I} & & & & & 0 \\
 & & & & & \ddots & & & \vdots \\
 & & & & & & \boxed{-I} & & 0
\end{array}
$$

Figure 2.6: Matrix structure in splitting variable representation

This representation is particularly interesting for decomposition algorithms, because the relaxation of NAC breaks the SP into $|\Omega|$ independent scenario-problems, see Figure 2.7. It could be a large number of subproblems, but certainly computationally simpler than SP.

$$
\begin{array}{cccc}
 & X^\omega & Y^\omega & \\
\underline{b} & \boxed{A} & & \bar{b} \\
\underline{h}^\omega & \boxed{T^\omega} & \boxed{W^\omega} & \bar{h}^\omega
\end{array}
$$

Figure 2.7: Matrix structure under one scenario ω

2.2.6 Some two-stage examples

Example 2.2. Let us remember the previous deterministic Diet Problem, in Example 2.1. We have two stage decisions:

- First stage decisions (X): amount of products to buy today for the warehouse, decisions that must be taken here and now.
- Second stage decisions (Y^ω): amount of products needed to be purchased two weeks later, decisions depending on the scenario ω.

We will consider three cases, depending on the uncertainty sources:

Case 1 , where prices are now unknown for the second stage and depend on the selected market.

Case 2 , where nutrients and prices can be altered depending on the own-brand taken from each market, see Mulvey et al. 1955 [24] collected in Censor & Zenios 1997 [7].

Case 3 , where bounds on nutrient requirements depend on the characteristics of different potential clients.

Let us detail the three models, where stochasticity appears in several ways.

Case 1. Stochasticity in objective function coefficients, $\boldsymbol{\xi^\omega = (q^\omega)}$.

Now, prices in two weeks are unknown. We have considered three markets where we can go shopping (ω_1: Eroski, ω_2: Simply and ω_3: Mercadona). So, depending where we buy, the price of each product will be different (stochasticity). In this case, uncertainty is only considered in the cost.

Table 2.3: Price of each product depending on the market

	First Stage		Second Stage			
	c		q^ω	Eroski (ω_1)	Simply (ω_2)	Mercadona (ω_3)
Pasta	c_1	1.98	q_1^ω	2.00	2.25	1.25
Lentils	c_2	1.58	q_1^ω	1.29	2.48	1.25

The scenario tree for modeling Case 1 is detailed in Figure 2.8.

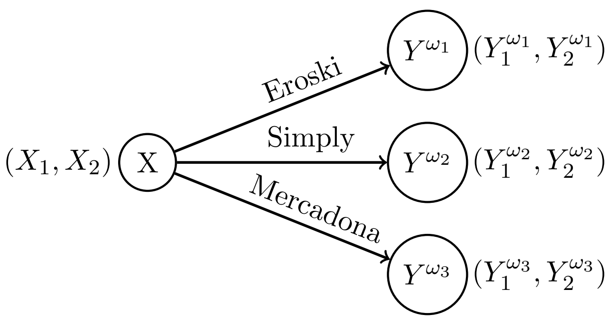


Figure 2.8: Scenario tree for Case 1 and Case 2

Let us model the problem taking into account that all values are not equally likely. In Vitoria city, there are 23 Eroski, 9 Simply and 4 Mercadona, i.e., $(p^{\omega_1}, p^{\omega_2}, p^{\omega_3}) = \left(\frac{23}{36}, \frac{9}{36}, \frac{4}{36}\right)$. So, the model is given in (2.15):

$$\begin{aligned}
&min && 1.98X_1 + 1.58X_2 + \sum_{\omega\in\Omega} p^{\omega}\,(q_1^{\omega}\ q_2^{\omega}) \begin{pmatrix} Y_1^{\omega} \\ Y_2^{\omega} \end{pmatrix} \\
&\text{subject to} && 142.5 \leq 18X_1 + 68.737X_2 \\
& && 32775 \leq 3530X_1 + 3100X_2 \leq 45000 \\
& && 285 \leq 18(X_1 + Y_1^{\omega}) + 68.737(X_2 + Y_2^{\omega}), \qquad \forall\omega \in \{1,2,3\} \\
& && 65550 \leq 3530(X_1 + Y_1^{\omega}) + 3100(X_2 + Y_2^{\omega}) \leq 90000, \quad \forall\omega \in \{1,2,3\} \\
& && X_i,\ Y_i^{\omega} \geq 0,\ i \in \{1,2\} \text{ and } \ \omega \in \{1,2,3\}
\end{aligned} \tag{2.15}$$

where q_i^{ω} values are given in Table 2.3.

Hence, the optimal solution of the problem (2.15) is given in Table 2.4, with a total expected cost of $Z_{SP} = 31963$€:

Table 2.4: Optimal SP solution for Case 1

	First Stage		Second Stage			
	X		Y^{ω}	ω_1	ω_2	ω_3
Pasta	X_1	0	Y_1^{ω}	0	9285	0
Lentils	X_1	10573	Y_2^{ω}	10573	0	9285

Case 2. Stochasticity in objective function coefficients and recourse matrix, $\boldsymbol{\xi^{\omega}} = (\boldsymbol{q^{\omega}}, \boldsymbol{W^{\omega}})$.

Since we buy in different markets the nutrients of each product could also change. Let us show all the nutrients expressed in Table 2.5:

Table 2.5: Nutrients of each product depending on the market

	W^{ω}	Eroski (ω_1)	Simply (ω_2)	Mercadona (ω_3)
Iron (mg)	w_{11}^{ω}	17	16	19.0
	w_{12}^{ω}	68	69	68.6
Energy (kcal)	w_{21}^{ω}	3540	3440	3590
	w_{22}^{ω}	2810	2810	2807

Let us model the problem in this case:

$$
\begin{aligned}
\textit{min} \quad & 1.98X_1 + 1.58X_2 + \sum_{\omega \in \Omega} p^{\omega} \left(q_1^{\omega} \ q_2^{\omega}\right) \begin{pmatrix} Y_1^{\omega} \\ Y_2^{\omega} \end{pmatrix} \\
\text{subject to} \quad & 142.5 \leq 18X_1 + 68.737X_2 \\
& 32775 \leq 3530X_1 + 3100X_2 \leq 45000 \\
& 285 \leq 18X_1 + 68.737X_2 + \left(w_{11}^{\omega} \ w_{12}^{\omega}\right) \begin{pmatrix} Y_1^{\omega} \\ Y_2^{\omega} \end{pmatrix}, \qquad \forall \omega \in \{1,2,3\} \\
& 65550 \leq 3530X_1 + 3100X_2 + \left(w_{21}^{\omega} \ w_{22}^{\omega}\right) \begin{pmatrix} Y_1^{\omega} \\ Y_2^{\omega} \end{pmatrix} \leq 90000, \quad \forall \omega \in \{1,2,3\} \\
& X_i, \ Y_i^{\omega} \geq 0, \ i \in \{1,2\} \text{ and } \ \omega \in \{1,2,3\}
\end{aligned}
\tag{2.16}
$$

where $(p^{\omega_1}, \ p^{\omega_2}, \ p^{\omega_3}) = \left(\frac{23}{36}, \frac{9}{36}, \frac{4}{36}\right)$, q^{ω} values are defined in the Table 2.3 and nutrients w_{ij}^{ω} in Table 2.5. Thus, the optimal solution of the problem (2.16) is given in Table 2.6, with a total expected cost of $Z_{SP} = 32977$€:

Table 2.6: Optimal SP solution for Case 2

	First Stage		Second Stage			
	X		Y^{ω}	ω_1	ω_2	ω_3
Pasta	X_1	0	Y_1^{ω}	0	9285	9130
Lentils	X_1	10573	Y_2^{ω}	11664	0	0

Case 3. Stochasticity in second stage bounds, $\boldsymbol{\xi^{\omega} = (\underline{h}^{\omega}, \overline{h}^{\omega})}$.

Let us define three different scenarios according to the age and sex of the person: ω_1: 100% women (20-49), ω_2: 50% women (20-49) and 50% men (20-49), ω_3: 100% men (20-49). In this case, LHS and RHS are uncertain where here-and-now decisions must be taken.

Let us show the bounds in the Table 2.7:

Table 2.7: Bounds according to each nutrient and scenario

	h^{ω}	100% W(ω_1)	50% W(ω_2)	0% W(ω_3)
Min Iron	$\underline{h}_1^{\omega}$	494	387.20	285
Min Energy	$\overline{h}_2^{\omega}$	65550	75744.45	85500
Max Energy	$\underline{h}_2^{\omega}$	69000	79731.00	90000

Case 3 can be expressed in the following way by a decision tree:

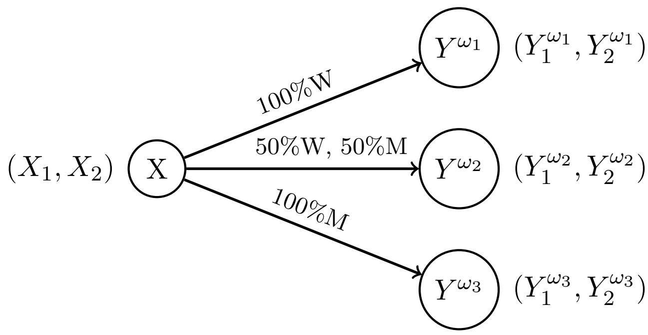


Figure 2.9: Scenario tree for Case 3

Let us model the problem taking into account that all scenarios are almost equally likely (in Alava there are 66740 men of 20-49 years old and 63823 women of 20-49 years old, according to EUSTAT [39]), i.e., $(p^{\omega_1},\ p^{\omega_2},\ p^{\omega_3}) = (0.33, 0.33, 0.34)$. Let us assume that prices and nutrients are fixed to first stage values, detailed in Table 2.1:

$$
\begin{aligned}
\mathit{min} \quad & 1.98X_1 + 1.58X_2 + 1.98Y_1^{\omega} + 1.58Y_2^{\omega} \\
\text{subject to} \quad & 142.5 \le 18X_1 + 68.737X_2 \\
& 32775 \le 3530X_1 + 3100X_2 \le 45000 \\
& \underline{h}_1^{\omega} \le 18(X_1 + Y_1^{\omega}) + 68.737(X_2 + Y_2^{\omega}), \qquad \forall \omega \in \{1,2,3\} \\
& \underline{h}_2^{\omega} \le 3530(X_1 + Y_1^{\omega}) + 3100(X_2 + Y_2^{\omega}) \le \overline{h}_2^{\omega}, \qquad \forall \omega \in \{1,2,3\} \\
& X_i,\ Y_i^{\omega} \ge 0,\ i \in \{1,2\} \text{ and } \ \omega \in \{1,2,3\}
\end{aligned}
\tag{2.17}
$$

Thus, the optimal solution of the problem (2.17) is given in Table 2.8, with a total cost of $Z_{SP} = 38681$€.:

Table 2.8: Optimal SP solution for Case 3

	First Stage		Second Stage			
	X		Y^{ω}	ω_1	ω_2	ω_3
Pasta	X_1	0	Y_1^{ω}	0	0	13065
Lentils	X_1	14516	Y_2^{ω}	6629	9918	0

Chapter 3

The Value of Perfect Information and the Stochastic Solution

Stochastic programs as, real world problems, are often computationally difficult to solve. Before solving the stochastic model, we could be tempted to solve simpler problems: for example, we could simplify the real imprecise data, replacing the unknown parameters with expected values of those random variables and solve the obtained deterministic problem or alternatively, we could solve all related scenario submodels and compute the expectation of these different solutions.

The main issue about these alternatives is that sometimes the solution can be nearly optimal, totally inexact or even non implementable. The way to know if the simplified model is good enough is calculating these two measures: the expected value of perfect information (EVPI) and the value of the stochastic solution (VSS), see Birge & Louveaux 2011 [5].

In this chapter we will explain these two concepts for two-stage models. From Section 3.1 to 3.3 there are shown essencial models, known as Wait-and-See (WS), Expected Value (EV) and Expected result of using Expected Value (EEV). Sections 3.4 and 3.5 provide the expected value of perfect information and the value of stochastic solution. Some basic inequalities and the relationship between EVPI and VSS are given in Sections 3.6 and 3.7, respectively. For more general definitions and inequalities, extended to the multistage stochastic model, see [9]. The example code is detailed in Appendix A.

3.1 Wait-and-See solution (WS)

W*ait-and-****S****ee* models assume that the decision maker is able to wait till the uncertainty is over, before calculating the optimal solutions. Therefore, this aproximation is based on the perfect information along the horizon planning. Then, the problem can be defined as follows:

$$\begin{aligned} WS \quad = \quad \min \quad & \sum_{\omega\in\Omega} p^\omega (cX^\omega + q^\omega Y^\omega) \\ \text{subject to} \quad & \underline{b} \le AX^\omega \le \bar{b}, && \forall \omega \in \Omega \\ & \underline{h}^\omega \le T^\omega X^\omega + W^\omega Y^\omega \le \overline{h}^\omega, && \forall \omega \in \Omega \\ & X^\omega, Y^\omega \ge 0, && \forall \omega \in \Omega \end{aligned} \tag{3.1}$$

This problem can be decomposed in $|\Omega|$ problems, Z^ω_{WS}, one for each scenario $\omega \in \Omega$:

$$\begin{aligned} Z^\omega_{WS} \quad = \quad \min \quad & cX^\omega + q^\omega Y^\omega \\ \text{subject to} \quad & \underline{b} \le AX^\omega \le \bar{b} \\ & \underline{h}^\omega \le T^\omega X^\omega + W^\omega Y^\omega \le \overline{h}^\omega \\ & X^\omega, Y^\omega \ge 0 \end{aligned} \tag{3.2}$$

Consequently, Wait-and-See solution, WS, is defined as the expected value of the random variable $Z_{WS} = (Z^\omega_{WS})_{\omega\in\Omega}$ where Z^ω_{WS} are the optimal solutions of the problems (3.2) and it is equivalent to the optimal solution of (3.1):

$$WS = E[Z_{WS}] = \sum_{\omega\in\Omega} p^\omega Z^\omega_{WS}$$

Example 3.1. Let us continue with the example presented in Chapter 2 and calculate the WS in the three cases defined before.

Case 1. Stochasticity in objective function coefficients, $\boldsymbol{\xi^\omega = (q^\omega)}$.

Let us solve these three problems for $\omega \in \{\omega_1, \omega_2, \omega_3\}$:

$$\begin{aligned} Z^\omega_{WS_1} \quad = \quad min \quad & 1.98X^\omega_1 + 1.58X^\omega_2 + (q^\omega_1\ q^\omega_2)\begin{pmatrix} Y^\omega_1 \\ Y^\omega_2 \end{pmatrix} \\ \text{subject to} \quad & 142.5 \le 18X^\omega_1 + 68.737X^\omega_2 \\ & 32775 \le 3530X^\omega_1 + 3100X^\omega_2 \le 45000 \\ & 285 \le 18(X^\omega_1 + Y^\omega_1) + 68.737(X^\omega_2 + Y^\omega_2) \\ & 65550 \le 3530(X^\omega_1 + Y^\omega_1) + 3100(X^\omega_2 + Y^\omega_2) \le 69000 \\ & X^\omega_i,\ Y^\omega_i \ge 0,\ i \in \{1,2\} \end{aligned} \tag{3.3}$$

where second stage stochastic prices q are given in Table 2.3.

Thus, the optimal solutions of the problems (3.3) according to each scenario are given in Table 3.1:

Table 3.1: Optimal WS solutions for Case 1

	ω_1	ω_2	ω_3
X_1^*	0	0	0
X_2^*	10573	14513	10573
Y_1^*	0	5822	9085
Y_2^*	10573	0	0
Z^*	30376€	36078€	28343€

It follows that the expected cost of products under the wait and see approach is

$$WS_1 = \frac{23}{36} \cdot 30376 + \frac{9}{36} \cdot 36078 + \frac{4}{36} \cdot 28343 = 31575€$$

Case 2. Stochasticity in objective function coefficients and recourse matrix, $\boldsymbol{\xi^\omega} = (\boldsymbol{q^\omega}, \boldsymbol{W^\omega})$.

Let us solve these three problems for $\omega \in \{\omega_1, \omega_2, \omega_3\}$:

$$\begin{aligned}
Z_{WS_2}^{\omega} \quad = \quad min \qquad & 1.98X_1^\omega + 1.58X_2^\omega + (q_1^\omega \ q_2^\omega)\begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
\text{subject to} \qquad & 142.5 \leq 18X_1^\omega + 68.737X_2^\omega \\
& 32775 \leq 3530X_1^\omega + 3100X_2^\omega \leq 45000 \\
& 285 \leq 18X_1^\omega + 68.737X_2^\omega + (w_{11}^\omega \ w_{12}^\omega)\begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
& 65550 \leq 3530X_1^\omega + 3100X_2^\omega + (w_{21}^\omega \ w_{22}^\omega)\begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \leq 69000 \\
& X_i^\omega, \ Y_i^\omega \geq 0, \ i \in \{1, 2\}
\end{aligned} \tag{3.4}$$

where q values are defined in the Table 2.3 and nutrients W in Table 2.5.

Thus, the optimal solutions of problems (3.4) according to each scenario are in Table 3.2:

Table 3.2: Optimal WS solutions of Case 2

	ω_1	ω_2	ω_3
X_1^*	0	0	0
X_2^*	10573	14516	10573
Y_1^*	0	5974	9130
Y_2^*	11664	0	0
Z^*	31783€	36421€	28149€

It follows that the expected cost of products under the wait and see approach is

$$WS_2 = \frac{23}{36} \cdot 31783 + \frac{9}{36} \cdot 36421 + \frac{4}{36} \cdot 28149 = 32539€$$

Case 3. Stochasticity in second stage bounds, $\boldsymbol{\xi^\omega = (\underline{h}^\omega, \overline{h}^\omega)}$.

Let us solve these three problems for $\omega \in \{\omega_1, \omega_2, \omega_3\}$:

$$\begin{aligned}
Z^\omega_{WS_3} \quad = \quad min \qquad & 1.98X_1^\omega + 1.58X_2^\omega + 1.98Y_1^\omega + 1.58Y_2^\omega \\
\text{subject to} \qquad & 142.5 \leq 18X_1^\omega + 68.737X_2^\omega \\
& 32775 \leq 3530X_1^\omega + 3100X_2^\omega \leq 45000 \\
& \underline{h}_1^\omega \leq 18(X_1^\omega + Y_1^\omega) + 68.737(X_2^\omega + Y_2^\omega) \\
& \underline{h}_2^\omega \leq 3530(X_1^\omega + Y_1^\omega) + 3100(X_2^\omega + Y_2^\omega) \leq \overline{h}_2^\omega \\
& X_i^\omega,\ Y_i^\omega \geq 0,\ i \in \{1, 2\}
\end{aligned} \tag{3.5}$$

where second stage stochastic LHS and RHS are given in Table 2.7

Thus, the optimal solutions of the problems (3.5) according to each scenario are given in Table 3.3:

Table 3.3: Optimal WS solutions of Case 3

	ω_1	ω_2	ω_3
X_1^*	0	0	0
X_2^*	14516	14516	14516
Y_1^*	0	0	0
Y_2^*	6629	9918	13065
Z^*	33474€	38680€	43662

It follows that the expected cost of products under the wait and see approach is

$$WS_3 = 0.33 \cdot 33474 + 0.33 \cdot 38680 + 0.34 \cdot 436262 = 38681€$$

3.2 Expected Value problem (EV)

In practice, many people either do not know stochastic modeling or believe that finding stochastic solution is too difficult. A usual simplification consists of replacing each random variable with an estimation, for example, the expected value. This easier problem is called *expected value problem*, also *mean value problem*, and its solution, *Expected Value*, EV. Let show it in the deterministic problem (3.6):

$$
\begin{aligned}
EV \quad = \quad & \min \quad cX + E[\boldsymbol{q}]Y \\
& \text{subject to} \quad \underline{b} \leq AX \leq \overline{b} \\
& \qquad E[\underline{\boldsymbol{h}}] \leq E[\boldsymbol{T}]X + E[\boldsymbol{W}]Y \leq E[\overline{\boldsymbol{h}}] \\
& \qquad X, Y \geq 0
\end{aligned} \tag{3.6}
$$

Example 3.2. Let us calculate the Expected Value solution on the previous three cases.

Case 1. Stochasticity in objective function coefficients, $\boldsymbol{\xi^\omega} = (\boldsymbol{q^\omega})$.

First of all, we only need the expected value of the prices of the products

$$E[\mathbf{q}] = \left(\frac{23}{36} \cdot (2,\ 1.29) + \frac{9}{36} \cdot (2.25,\ 2.48) + \frac{4}{36} \cdot (1.25,\ 1.25)\right) = (1.98,\ 1.58)$$

.

Now, let us solve the first EV problem:

$$
\begin{aligned}
EV_1 \quad = \quad & min \quad 1.98X_1 + 1.58X_2 + 1.98Y_1 + 1.58Y_2 \\
& \text{subject to} \quad 142.5 \leq 18X_1 + 68.737X_2 \\
& \qquad 32775 \leq 3530X_1 + 3100X_2 \leq 45000 \\
& \qquad 285 \leq 18(X_1 + Y_1) + 68.737(X_2 + Y_2) \\
& \qquad 65550 \leq 3530(X_1 + Y_1) + 3100(X_2 + Y_2) \leq 69000 \\
& \qquad X_i,\ Y_i \geq 0,\ i \in \{1, 2\}
\end{aligned} \tag{3.7}
$$

Thus, the optimal solution of the EV_1 problem (3.7) is

$$(X^*_{EV_1}, Y^*_{EV_1}) = (X_1, X_2, Y_1, Y_2) = (0,\ 14516,\ 0,\ 6629) \quad \text{and} \quad Z^*_{EV_1} = EV_1 = 33474€$$

Case 2. Stochasticity in objective function coefficients and recourse matrix, $\boldsymbol{\xi}^{\boldsymbol{\omega}} = (\boldsymbol{q}^{\boldsymbol{\omega}}, \boldsymbol{W}^{\boldsymbol{\omega}})$.

First of all, we need the expected value of the prices of the products (calculated in the previous case) and the expected value of the nutrients, W, the recourse matrix

$$E[\mathbf{W}] = \frac{23}{36}\begin{pmatrix} 17 & 68 \\ 3540 & 2810 \end{pmatrix} + \frac{9}{36}\begin{pmatrix} 16 & 69 \\ 3440 & 2810 \end{pmatrix} + \frac{4}{36}\begin{pmatrix} 19 & 68.6 \\ 3590 & 2807 \end{pmatrix} = \begin{pmatrix} 16.97 & 68.32 \\ 3520.56 & 2809.67 \end{pmatrix}$$

Now, let us solve the second EV problem:

$$\begin{aligned} EV_2 \quad = \quad min \quad & 1.98X_1 + 1.58X_2 + 1.98Y_1 + 1.58Y_2 \\ \text{subject to} \quad & 142.5 \le 18X_1 + 68.737X_2 \\ & 32775 \le 3530X_1 + 3100X_2 \le 45000 \\ & 285 \le 18X_1 + 68.737X_2 + 16.97Y_1 + 68.32Y_2 \\ & 65550 \le 3530X_1 + 3100X_2 + 3520.56Y_1 + 2809.67Y_2 \le 69000 \\ & X_i,\ Y_i \ge 0,\ i \in \{1,2\} \end{aligned} \tag{3.8}$$

Thus, the optimal solution of the EV_2 problem (3.8) is

$$(X^*_{EV_2}, Y^*_{EV_2}) = (X_1, X_2, Y_1, Y_2) = (0,\ 14516,\ 5837,\ 0) \quad \text{and} \quad EV_2 = 34533€$$

Case 3. Stochasticity in second stage bounds, $\boldsymbol{\xi}^{\boldsymbol{\omega}} = (\underline{\boldsymbol{h}}^{\boldsymbol{\omega}}, \overline{\boldsymbol{h}}^{\boldsymbol{\omega}})$.

First of all, we only need the expected value of the bounds on nutrient requirements depending on the clients distribution:

$$\begin{aligned} E[\boldsymbol{h}] &= 0.33\begin{pmatrix} 494 & \infty \\ 65550 & 69000 \end{pmatrix} + 0.33\begin{pmatrix} 387.20 & \infty \\ 75744.45 & 79731 \end{pmatrix} + 0.34\begin{pmatrix} 285 & \infty \\ 85500 & 90000 \end{pmatrix} \\ &= \begin{pmatrix} 387.70 & \infty \\ 75697.17 & 79681.23 \end{pmatrix} \end{aligned}$$

Now, let us solve the third EV problem:

$$\begin{aligned} EV_3 \quad = \quad min \quad & 1.98(X_1+Y_1)+1.58(X_2+Y_2) \\ \text{subject to} \quad & 18X_1+68.737X_2 \geq 142.5 \\ & 327750 \leq 3530X_1+3100X_2 \leq 45000 \\ & 18(X_1+Y_1)+68.737(X_2+Y_2) \geq 387.7 \\ & 75697.17 \leq 3530(X_1+Y_1)+3100(X_2+Y_2) \leq 79681.23 \\ & X_i,\ Y_i \geq 0,\ i \in \{1,2\} \end{aligned} \tag{3.9}$$

Thus, the optimal solution of the EV_3 problem (3.9):

$$(X^*_{EV_3}, Y^*_{EV_3}) = (X_1, X_2, Y_1, Y_2) = (0,\ 14516,\ 0,\ 9918) \quad \text{and} \quad EV_3 = 38681€$$

3.3 Expected result of using the EV solution (EEV)

It is interesting to evaluate the optimal solution reached in the *mean value problem* along the scenario tree. That is, fixing the first-stage solution of the EV problem, X_{EV}, in the stochastic program (2.4) - (2.7) gives the ***E**xpected result of using the **EV** solution.* Then, the problem can be defined as follows:

$$\begin{aligned} EEV \quad = \quad cX_{EV} \quad + \quad \min \quad & \textstyle\sum_{\omega\in\Omega} p^\omega q^\omega Y^\omega \\ \text{s.t.} \quad & \underline{h}^\omega - T^\omega X_{EV} \leq W^\omega Y^\omega \leq \overline{h}^\omega - T^\omega X_{EV}, \qquad \forall\omega\in\Omega \\ & Y^\omega \geq 0, \qquad \forall\omega\in\Omega \end{aligned} \tag{3.10}$$

This problem can be decomposed in $|\Omega|$ different problems, Z^ω_{EEV} depending on the scenario ω:

$$\begin{aligned} Z^\omega_{EEV} \quad = \quad cX_{EV} \quad + \quad \min \quad & q^\omega Y^\omega \\ \text{s.t.} \quad & \underline{h}^\omega - T^\omega X_{EV} \leq W^\omega Y^\omega \leq \overline{h}^\omega - T^\omega X_{EV} \\ & Y^\omega \geq 0 \end{aligned} \tag{3.11}$$

Therefore, the optimal solution of the problem (3.10) is equivalent to the expected value over the random variable $Z_{EEV} = (Z^\omega_{EEV})_{\omega\in\Omega}$, where Z^ω_{EEV} are the solutions of the problems (3.11):

$$EEV = E[Z_{EEV}] = \textstyle\sum_{\omega\in\Omega} p^\omega Z^\omega_{EEV}$$

Example 3.3. Let us continue with the Diet Problem. We will analyse the effect of implementing the optimal solution given by EV, $X_{EV} = (X_1, X_2)$, along the corresponding scenario tree.

Case 1. Stochasticity in objective function coefficients, $\boldsymbol{\xi^\omega = (q^\omega)}$.

Let us define three different problems according to each scenario:

$$\begin{aligned}
Z^{\omega}_{EEV_1} &= 1.98X_1 + 1.58X_2 \ + \ \min \quad (q_1^\omega \ q_2^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
\text{subject to} \quad & 142.5 - (18X_1 + 68.737X_2) \leq 18Y_1^\omega + 68.737Y_2^\omega \\
& 65550-(3530X_1+3100X_2) \leq 3530Y_1^\omega+3100Y_2^\omega \leq 69000-(3530X_1+3100X_2) \\
& X_{EV_1} = (X_1, X_2) = (0, \ 14516) \\
& Y_i^\omega \geq 0, \ i \in \{1, 2\}
\end{aligned}$$

where q values are defined in Table 2.3.

Equivalently, if we fix the solution $X_{EV} = (X_1, X_2) = (0, \ 14516)$ from EV_1 model:

$$\begin{aligned}
Z^{\omega}_{EEV_1} &= 22935.48 \ + \ \min \quad (q_1^\omega \ q_2^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
\text{subject to} \quad & 0 \leq 18Y_1^\omega + 68.737Y_2^\omega \\
& 20550 \leq 3530Y_1^\omega + 3100Y_2^\omega \leq 24000 \\
& Y_i^\omega \geq 0, \ i \in \{1, 2\}
\end{aligned} \tag{3.12}$$

Thus, the optimal solutions of the $Z^{\omega}_{EEV_1}$ problems (3.12) according to each scenario ω are given in Table 3.4.

Table 3.4: Optimal EEV solutions for Case 1

	ω_1	ω_2	ω_3
Y_1^*	0	5822	5822
Y_2^*	6629	0	0
Z^*	31531€	36078€	30257€

It follows that the expected solution using the expected value is

$$EEV_1 = \frac{23}{36} \cdot 31531 + \frac{9}{36} \cdot 36078 + \frac{4}{36} \cdot 30257 = 32526€$$

Case 2. Stochasticity in objective function coefficients and recourse matrix, $\boldsymbol{\xi^\omega} = (\boldsymbol{q^\omega}, \boldsymbol{W^\omega})$.

Let us define three different problems according to each scenario:

$$
\begin{aligned}
Z^\omega_{EEV_2} &= 1.98X_1 + 1.58X_2 \;+\; \min \quad (q_1^\omega \; q_2^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
\text{subject to} \quad & 142.5 - (18X_1 + 68.737X_2) \le (w_{11}^\omega \; w_{12}^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
& 65550-(3530X_1+3100X_2) \le (w_{21}^\omega \; w_{22}^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \le 69000-(3530X_1+3100X_2) \\
& X_{EV_2} = (X_1, X_2) = (0, \; 14516) \\
& Y_i^\omega \ge 0, \; i \in \{1,2\}
\end{aligned}
$$

where q values are defined in the Table 2.3 and W in the Table 2.5.

Equivalently, if we fix the solution $X_{EV} = (X_1, X_2) = (0, \; 14516)$ from model EV_2.

$$
\begin{aligned}
Z^\omega_{EEV_2} &= 22935.48 \;+\; \min \quad (q_1^\omega \; q_2^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
\text{subject to} \quad & 0 \le (w_{11}^\omega \; w_{12}^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \\
& 20550 \le (w_{21}^\omega \; b_{22}^\omega) \begin{pmatrix} Y_1^\omega \\ Y_2^\omega \end{pmatrix} \le 24000 \\
& Y_i^\omega \ge 0, \; i \in \{1,2\}
\end{aligned}
\tag{3.13}
$$

Thus, the optimal solutions of the $Z^\omega_{EEV_2}$ problems (3.13) according to each scenario ω are given in Table 3.5

Table 3.5: Optimal EEV solutions for Case 2

	ω_1	ω_2	ω_3
Y_1^*	0	5974	5724
Y_2^*	7313	0	0
Z^*	32414€	36421€	30135€

It follows that the expected solution using the expected value is

$$EEV_2 = \frac{23}{36} \cdot 32414 + \frac{9}{36} \cdot 36421 + \frac{4}{36} \cdot 30135 = 33162€$$

Case 3. Stochasticity in second stage bounds, $\boldsymbol{\xi^\omega = (\underline{h}^\omega, \overline{h}^\omega)}$.

Let us define three different problems according to each scenario:

$$
\begin{aligned}
Z^\omega_{EEV_3} \;=\; & 1.98X_1 + 1.58X_2 + \min\{1.98Y^\omega_1 + 1.58Y^\omega_2\} \\
\text{subject to} \quad & \underline{h}^\omega_1 - (18X_1 + 68.737X_2) \le 18Y^\omega_1 + 68.737Y^\omega_2 \\
& \underline{h}^\omega_2 - (3530X_1 + 3100X_2) \le 3530Y^\omega_1 + 3100Y^\omega_2 \le \overline{h}^\omega_2 - (3530X_1 + 3100X_2) \\
& X_{EV_3} = (X_1, X_2) = (0,\ 14516) \\
& Y^\omega_i \ge 0,\ i \in \{1, 2\}
\end{aligned}
$$

Equivalently, if we fix the solution $X_{EV} = (X_1, X_2) = (0,\ 14516)$ from EV_3 model

$$
\begin{aligned}
Z^\omega_{EEV_3} = & \; 22935.49 \; + \; \min \quad 1.98Y^\omega_1 + 1.58Y^\omega_2 \\
\text{s.t.} \quad & 0 \le 18Y^\omega_1 + 68.737Y^\omega_2 \\
& \underline{h^\omega_2} - 45000 \le 3530Y^\omega_1 + 3100Y^\omega_2 \le \overline{h^\omega_2} - 45000 \\
& Y^\omega_i \ge 0,\ i \in \{1, 2\}
\end{aligned}
\tag{3.14}
$$

where $\underline{h}$ and $\overline{h}$ values are defined in the Table 2.7.

Thus, the optimal solutions of the $Z^\omega_{EEV_3}$ problems (3.14) according to each scenario ω given in Table 3.6:

Table 3.6: Optimal EEV solutions for Case 3

	ω_1	ω_2	ω_3
Y^*_1	0	0	0
Y^*_2	6629	9918	13065
Z^*	33474€	38680€	43662

It follows that the expected solution using the expected value is

$$EEV_3 = 0.33 \cdot 33474 + 0.33 \cdot 38680 + 0.34 \cdot 43662 = 38681€$$

3.4 The Expected Value of Perfect Information (EVPI)

The *Expected Value of Perfect Information* is the measure of how much a decision maker would be willing to pay in order to obtain perfect information about the future (complete and accurate), see Schlaifer & Raiffa 1961 [28].

The EVPI is defined as the difference between the *wait-and-see*, WS, and the *here-and-now* solution, SP. In minimization problems:

$$EVPI = SP - WS \tag{3.15}$$

The bigger the difference, the more importance will have the uncertainty.

Example 3.4. Since wait-and-see and here-and-now solutions have been calculated before in the three different cases, the expected value of perfect information is summarized in Table 3.7:

Table 3.7: EVPI solutions for Cases 1, 2 and 3

	Case 1	Case 2	Case 3
EVPI	387	438	0

This means that, in Case 1 and 2, the maximum amount of money that we are willing to pay in order to know previously the market where we have to buy is 387€ and 438€, respectively. On the contrary, in Case 3 the decision maker thinks that is not worthy to pay nothing for real information of the future, but WS solutions are not implementable in general, since they do not always provide unique first stage decisions.

3.5 The Value of Stochastic Solutions (VSS)

The *Value of the Stochastic Solution* allows us to obtain more precisely the goodness of the expected value solution against the stochastic problem solution. VSS represents the cost of ignoring uncertainty in the moment of making a decision, the expected loss of using the deterministic solution, therefore the importance of solving the stochastic model, SP.

The VSS is defined as the difference between the *Expected result of using EV* problem and the *here-and-now* solution, SP. In minimization problems:

$$VSS = EEV - SP \tag{3.16}$$

In some cases the value of the stochastic solution and the expected value of the mean solution can be the same whereby VSS=0. In this occasion, calculating this measure, it would be unecessary to solve such a hard problem as the stochastic one. Nevertheless, it is really difficult to know if the uncertainty is important before solving the problem.

If we remember the definition of EVPI, we can see that VSS seems similar to that. The difference is that EVPI is the maximum price that the decision maker should pay in order to know the uncertainty, and VSS, on the contrary, is the real cost of ignoring it.

Example 3.5. Since the expected result of using the expected value and here-and-now solutions have been calculated before in the three different cases, the value of stochastic solution is summarized in Table 3.8.

Table 3.8: VSS solutions for Cases 1, 2 and 3

	Case 1	Case 2	Case 3
VSS	564	185	0

This means that Case 3 is the only one where it has not been worthy to calculate the stochastic solution. However, we cannot predict the result before the implementation of both models. Notice that for Case 1 and Case 2, the result is remarkable.

3.6 Main Inequalities

The relations between the solutions and measures defined in the previous sections were established by Madansky in 1960 [23].

Proposition 3.1. *For the minimization lineal models, the following inequalities are satisfied:*

$$WS \leq SP \leq EEV \tag{3.17}$$

Obviously, in the maximization models the inequalities are the opposite.

Proof. On one hand, since the optimal solution of the *Stochastic Problem* (2.4) - (2.7), also known as *Recourse Problem* is feasible solution of the *Wait-and-See* problem (3.1), it is directly proved the first part of the inequality: $WS \leq SP$.

On the other hand, since the optimal solution of the *Expected result of using EV* problem (3.6) is feasible solution of the *Stochastic Problem* (2.4) - (2.7), the second inequality is reached: $SP \leq EEV$

□

Example 3.6. We just need to compare the three values obtained before. As we can see in Table 3.9, Proposition 3.1 is verified in strict inequality for Cases 1 and 2:

Table 3.9: Proposition 3.1 verified for Cases 1, 2 and 3

	WS		SP		EEV
Case 1	31575	$<$	31963	$<$	32526
Case 2	32539	$<$	32977	$<$	33162
Case 3	38681	$=$	38681	$=$	38681

Proposition 3.2. *In stochastic programs of minimization with fixed objective coefficients and fixed recourse matrix W:*

$$EV \leq WS \tag{3.18}$$

Proof. First, note that $EV = \min z(X, E(\boldsymbol{\xi}))$ and $WS = E_{\boldsymbol{\xi}}[\min z(X, \boldsymbol{\xi})]$. This means that we can base the proof in Jensen's inequality, see Jensen 1906 [16]. It states that for any convex function $f(\boldsymbol{\xi})$ of $\boldsymbol{\xi}$: $Ef(\boldsymbol{\xi}) \geq f(E(\boldsymbol{\xi}))$. Since $\boldsymbol{\xi} = (\xi^{\omega})_{\omega \in \Omega}$, we need to show that $f(\xi) = Z(X^*, \xi) = Z^{\xi}_{WS}$ is a convex function of ξ.

Let us consider two different vectors, ξ^1 and ξ^2, and some convex combination: $\xi^{\lambda} = \lambda\xi^1 + (1-\lambda)\xi^2, \lambda \in (0,1)$. Let $Z_1^* = Z(X_1^*, \xi^1)$ and $Z_2^* = Z(X_2^*, \xi^2)$ be some optimal solutions of $min\{cX + E[\min q^{\xi} Y^{\xi} | \underline{h}^{\xi} \leq T^{\xi}X + W^{\xi}Y^{\xi} \leq \overline{h}^{\xi}, Y^{\xi} \geq 0]\}$, s.t. $\underline{b} \leq AX \leq \bar{b},\ X \geq 0$ for $\xi = \xi^1$ and $\xi = \xi^2$, respectively. Then, $\lambda Z_1^* + (1-\lambda) Z_2^*$ is a feasible solution of $min\{cX + E[\min q^{\xi^{\lambda}} Y^{\xi^{\lambda}} \,|\, \underline{h}^{\xi^{\lambda}} \leq T^{\xi^{\lambda}} X + W^{\xi^{\lambda}} Y^{\xi^{\lambda}} \leq \overline{h}^{\xi^{\lambda}},\ Y^{\xi^{\lambda}} \geq 0]\}$, s.t. $\underline{b} \leq AX \leq \bar{b},\ X \geq 0$. Now, let Z_{λ}^* be an optimal solution of the last problem. We thus have

$$\begin{aligned} f(\lambda\xi^1 + (1-\lambda)\xi^2) = f(\xi^{\lambda}) = Z_{\lambda}^* = \min z(X, \lambda\xi^1 + (1-\lambda)\xi^2) \leq Z(\lambda(X_1^*, \xi^1) + (1-\lambda)(X_2^*, \xi^2)) \\ \leq \lambda Z(X_1^*, \xi^1) + (1-\lambda) Z(X_2^*, \xi^2) = \lambda Z_1^* + (1-\lambda) Z_2^* = \lambda f(\xi^1) + (1-\lambda) f(\xi^2) \end{aligned}$$

.

This stablishes convexity of $f(\xi)$, so according to Jensen's inequality

$$\mathbf{EV} = \min_X z(X, E(\boldsymbol{\xi})) = f(E(\boldsymbol{\xi})) \leq Ef(\boldsymbol{\xi}) = E_{\boldsymbol{\xi}}[\min_X z(X, \boldsymbol{\xi})] = \mathbf{WS}$$

$\square$

Notice that, the previous proposition is not true for all stochastic programs. Since we have already seen, it can only be uncertainty in the independent terms of the constraints or in the technological matrix T. Consequently, it is enough choosing a stochastic program being q the only non-fixed value. In this case, z would be a concave function of ξ, so Jensen's inequality cannot be applied.

Example 3.7. We just need to compare the two values calculated in previous examples:

Table 3.10: Proposition 3.2 verified for Cases 1, 2 and 3

	EV		WS
Case 1	33474	$\not\leq$	31575
Case 2	34533	$\not\leq$	32539
Case 3	38681	$=$	38681

The first two cases do not fulfill the inequality of the Proposition 3.2, notice that recourse matrix and objective coefficients are not fixed.

3.7 The Relationship between EVPI and VSS

The values of EVPI and VSS are usually different. This section describes the relationships between these two measures of uncertainty effects.

Proposition 3.3. *For any stochastic program:*

$$0 \leq EVPI \tag{3.19}$$

$$0 \leq VSS \tag{3.20}$$

Proof. It can be proved directly using Proposition 3.2.

□

Example 3.8. Note the satisfaction of Proposition 3.3 in Tables 3.7 and 3.8.

Proposition 3.4. *For stochastic programs with fixed recourse matrix and objective coefficients,*

$$EVPI \leq EEV - EV \tag{3.21}$$

$$VSS \leq EEV - EV \tag{3.22}$$

Proof. The proof is direct based on Proposition 3.3.

□

This means that EVPI and VSS are always nonnegative and besides, depending on the problem, these quantities are bounded above by $EEV - EV$, an easily computable value. Hence, whether $EV = EEV$, the value of both measures, $EVPI$ and VSS, is null. This will happen if $X(\boldsymbol{\xi})$, any feasible solution, is independent of $\boldsymbol{\xi}$. This means that optimal solution will be in the same situation and if we find optimal solution for one ξ, we will reach the same result, so it would be unnecessary to solve a recourse problem.

Example 3.9. Using the same example and the values reached before, we can directly see that

Table 3.11: Proposition 3.4 verified for Cases 1, 2 and 3

	EVPI		EEV-EV	VSS		EEV-EV
Case 1	387	$\nleq$	32526-33474=-948	564	$\nleq$	32526-33474=-948
Case 2	438	$\nleq$	33162-34533=-1371	185	$\nleq$	33162-34533=-1371
Case 3	0	=	38681-38681=0	0	=	38681-38681=0

The first two cases do not fulfill the inequalities 3.21 and 3.22 of the Proposition 3.4, notice again that recourse matrix and objective coefficients are not fixed.

Let us end this section by showing some other examples, see Birge & Louveaux 2011 [5], where one of the two previous measures vanish.

Example 3.10. $\boldsymbol{EVPI = 0}$ **and** $\boldsymbol{VSS \neq 0}$

Let us define the following problem with the continuous random variable $\boldsymbol{\xi}$ uniformly distributed over [1, 3] interval:

$$\begin{aligned} Z(X, \boldsymbol{\xi}) = {} & X_1 + 4X_2 + \min\{Y_1 + 10Y_2^+ + 10Y_2^-\} && (3.23)\\ \text{s.t.} \quad & X_1 + X_2 = 1 && (3.24)\\ & Y_1 + Y_2^+ - Y_2^- = \boldsymbol{\xi} + X_1 - 2X_2 && (3.25)\\ & X_i \geq 0, Y_1 \leq 2, Y_i \geq 0, i \in \{1,2\} && (3.26) \end{aligned}$$

Notice that $Y_2^+ + Y_2^- = |Y_2|$, $Y_2^+ - Y_2^- = Y_2$, $Y_2 \in \mathbb{R}$. Since we want to keep $Y_2^{+/-}$ as small as possible in order to minimize $\{Y_1 + 10Y_2^+ + 10Y_2^-\} = \{Y_1 + 10|Y_2|\}$, let us consider three different cases:

- If $Y_2 = 0$, then $Y_2^+ - Y_2^- = 0$, $Y_1 = \xi + X_1 - 2X_2$. In addition, since $0 \leq Y_1 \leq 2$, $0 \leq \xi + X_1 - 2X_2 \leq 2$ (first region)
- If $Y_2^- = 0$, then $Y_1 + Y_2^+ = \xi + X_1 - 2X_2$, since $Y_1 \leq 2$, $\xi + X_1 - 2X_2 \leq 2 + Y_2^+$. As it is a minimizing problem, $Y_2^+ = \xi + X_1 - 2X_2 - 2$ and $Y_1 = 2$. In this case, since $Y_2^+ \geq 0$, $Y_1 + Y_2^+ = 2 + Y_2^+ \geq 2$, so, $2 + \xi + X_1 - 2X_2 - 2 = \xi + X_1 - 2X_2 \geq 2$ (second region)
- If $Y_2^+ = 0$, then $Y_1 - Y_2^- = \xi + X_1 - 2X_2$, since $Y_1 \geq 0$, $\xi + X_1 - 2X_2 \geq 0 - Y_2^-$. As it is a minimizing problem, $Y_2^- = 2X_2 - X_1 - \xi$ and $Y_1 = 0$. In this case, $Y_1 - Y_2^- = 2X_2 - X_1 - \xi \geq 0$, so, $\xi + X_1 - 2X_2 \leq 0$ (third region)

That is,

$$Y^*(X,\xi) = (Y_1, Y_2^+, Y_2^-) = \begin{cases} (\xi + X_1 - 2X_2,\ 0,\ 0) & \text{if } 0 \leq \xi + X_1 - 2X_2 \leq 2 \\ (2,\ \xi + X_1 - 2X_2 - 2,\ 0), & \text{if } \xi + X_1 - 2X_2 \geq 2 \\ (0,\ 0,\ 2X_2 - \xi - X_1), & \text{if } \xi + X_1 - 2X_2 \leq 0 \end{cases}$$

Therefore,

$$Z(X,\xi) = \begin{cases} 2X_1 + 2X_2 + \xi, & \text{if } 0 \leq \xi + X_1 - 2X_2 \leq 2 \\ -18 + 11X_1 - 16X_2 + 10\xi, & \text{if } \xi + X_1 - 2X_2 \geq 2 \\ -9X_1 + 2X_2 - 10\xi, & \text{if } \xi + X_1 - 2X_2 \leq 0 \end{cases} \tag{3.27}$$

Given the first-stage constraint (3.24) $X_1 + X_2 = 1$, in the first region $Z(X,\xi) = 2(X_1 + X_2) + \xi = 2 + \xi$. Now, using the second-stage constraint (3.25), $Y_1 + Y_2^+ - Y_2^- = \xi + X_1 - 2X_2$, $Y_1 + 10Y_2^+ + 10Y_2^- \geq \xi + X_1 - 2X_2$. So, for an optimal Y in the second and third region, applying (3.24), $Z(X,\xi) \geq X_1 + 4X_2 + (\xi + X_1 - 2X_2) = 2(X_1 + X_2) + \xi = 2 + \xi$. Therefore, any $\widehat{X} \in \{(X_1, X_2) | X_1 + X_2 = 1, X \geq 0\}$ is an optimal solution of the problem (3.23) - (3.26) for $-X_1 + 2X_2 \leq \xi \leq -X_1 + 2X_2 + 2$, and applying (3.24), equivalently,
$-X_1 + 2(1 - X_1) \leq \xi \leq -X_1 + 2(1 - X_1) + 2 \quad \Leftrightarrow \quad 2 - 3X_1 \leq \xi \leq 4 - 2X_1$.

Since ξ follows a uniform distribution over [1,3], let us define three different cases:

- If $\xi \geq 1$, $2 - 3X_1 = 1 \Leftrightarrow X_1 = \frac{1}{3}$ and $X_2 = \frac{2}{3}$. Then, $\xi \leq 3$, so, $(\frac{1}{3}, \frac{2}{3})$ is an optimal solution for all ξ.
- For $X_1 = 1$, $-1 \leq \xi \leq 1$. Therefore, (1,0) is optimal in $\xi \in \{1\}$.
- For $X_1 = 0$, $2 \leq \xi \leq 4$. Thus, (0,1) is optimal in $\xi \in [2, 3]$.

Taking $X^*(\xi) = (\frac{1}{3}, \frac{2}{3})$ for all ξ, since all the solutions will be the same, we can conclude that $WS = RP = 2 + \xi = 2 + \frac{1+3}{2} = 2 + 2 = 4$, so $EVPI = RP - WS = 0$. On the other hand, solving $Z(X, E(\xi) = 2)$, we can reach another solution: $X^*(2) = (0, 1)$, then $EV = \min z(X, E(\xi)) = 2 + 2 = 4$.

In that case, since ξ is uniform over $[1, 3]$, $P(\xi) = 1/2$, $\forall \xi \in [1, 2]$ and $P(\xi) = 1/2$, $\forall \xi \in [2, 3]$. Whereas, since ξ is a continuous random variable, $E[X] = \int_{-\infty}^{+\infty} x f(x) dx$ where $f(x)$ is density function and

$$EEV = E_{\boldsymbol{\xi} \leq 2}(24 - 10\xi) + E_{\boldsymbol{\xi} \geq 2}(2 + \xi) = \int_1^2 (24 - 10\xi) \cdot \frac{1}{2} d\xi \ + \ \int_2^3 (2 + \xi) \cdot \frac{1}{2} d\xi =$$

$$\frac{1}{2} \left[24\xi - 10 \, \frac{\xi^2}{2} \bigg|_1^2 + 2\xi + \frac{\xi^2}{2} \bigg|_2^3 \right] = \frac{1}{2} \left(24 - 5 \cdot 3 + 2 + \frac{9}{2} - 2 \right) = \frac{27}{4}.$$

Thus, $VSS = EEV - RP = \frac{27}{4} - 4 = \frac{11}{4}$.

Example 3.11. $EVPI \neq 0$ and $VSS = 0$

Let us consider again the previous problem (3.23)-(3.26), where ξ is discrete random variable, $\xi \in \{0, \frac{3}{2}, 3\}$, with each event occurring with same probability, $\frac{1}{3}$. Taking into account the optimal solution reached in the previous example:

- If $\xi = 0$, $2 - 3X_1 \leq 0 \leq 4 - 3X_1 \;\Rightarrow\; \begin{cases} 2 - 3X_1 \leq 0 \\ 4 - 3X_1 \geq 0 \end{cases}$

  So, $X^*(0) = \{X | X_1 + X_2 = 1,\ \frac{2}{3} \leq X_1 \leq \frac{4}{3}\}$

- If $\xi = \frac{3}{2}$, $2 - 3X_1 \leq \frac{3}{2} \leq 4 - 3X_1 \;\Rightarrow\; \begin{cases} 1 - 6X_1 \leq 0 \\ 5 - 6X_1 \geq 0 \end{cases}$

  So, $X^*(\frac{3}{2}) = \{X | X_1 + X_2 = 1,\ \frac{1}{6} \leq X_1 \leq \frac{5}{6}\}$

- If $\xi = 3$, $2 - 3X_1 \leq 3 \leq 4 - 3X_1 \;\Rightarrow\; \begin{cases} -1 - 3X_1 \leq 0 \\ 1 - 3X_1 \geq 0 \end{cases}$

  So, $X^*(3) = \{X | X_1 + X_2 = 1,\ 0 \leq X_1 \leq \frac{1}{3}\}$

Let us take $X^*(3/2) = (2/3, 1/3)$ as optimal solution of EV. Then, $EV = Z((2/3, 1/3), 3/2) = 2 + 3/2 = 7/2$. Since for $\xi \in \{0, \frac{3}{2}\}$ we are in the first region and for $\xi = 3$ in the second one,

$$EEV = E_{\xi \in \{0, \frac{3}{2}, 3\}}[Z((2/3, 1/3), \xi)] = 2 + E_{\xi \in \{0, \frac{3}{2}\}}[Y^*((2/3, 1/3), \xi)] + E_{\xi = 3}[Y^*((2/3, 1/3), \xi)]$$

$$= 2 + E_{\xi \in \{0, \frac{3}{2}\}}[\xi] + E_{\xi = 3}\left[2 + 10\left(\xi + \frac{2}{3} - 2\,\frac{1}{3} - 2\right)\right] = 2 + \frac{1}{3}\left(0 + \frac{3}{2}\right) + \frac{1}{3}\,12 = \frac{13}{2}.$$

There is not just one optimal solution for the three cases, so $RP \neq WS$ and therefore, $EVPI \neq 0$. In Wait-and-See solution, it is possible to get a different optimal solution depending on the case (being all of them in the first region):

- $X^*(0) = (1, 0)$, so $Z((1, 0), 0) = 2 + 0 = 2$
- $X^*(3/2) = (1/2, 1/2)$, so $Z((1/2, 1/2), 3/2) = 2 + \frac{3}{2} = \frac{7}{2}$
- $X^*(3) = (0, 1)$, so $Z((0, 1), 3) = 2 + 3 = 5$

Hence,

$$WS = \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot \frac{7}{2} + \frac{1}{3} \cdot 5 = \frac{7}{2}.$$

So that we can reach the recourse solution, we will solve the stochastic program $\min E[Z(X, \boldsymbol{\xi})]$, so $(2/3, 1/3)$ is the optimal solution of SP. Thus

$$EV = 7/2 = WS \leq RP = 13/2 = EEV$$

and $EVPI = RP - WS = 3$ while $VSS = EEV - RP = 0$.

Chapter 4

An Application in the Third Sector: Hazia Project

In this chapter a third sector application is described, implemented and analysed. The third sector, also known as social economy or community sector, is the economic sector undertaken by non-governmental organizations and other non-profit organizations.

The computational experience has been carried out on a UPV/EHU server with the following features: 64 bits Linux Debian 8.3. operative system, Intel E5-2670 processor and 20 cores, 40 such virtualized. It also contains two hard drives: a RAID one with 300Gb and a solid state with 120Gb (128Gb usable RAM).

The application is detailed in Appendix B. The codes are implemented with the modeling system for mathematical programming and optimization, GAMS 22.8.1 [41] and solved by the optimizer solver CPLEX 11.1.1 [43]. The figures shown in this chapter have been obtained from the software environment for statistical computing and graphics, R-project [42].

Section 4.1 explains the context, Section 4.2 details the models, Section 4.3 includes all the datasets needed for the model, Section 4.4 shows and analyses the results and Section 4.5 concludes.

4.1 Background

Sortarazi association [37], with recognition of Public Utility, was created in 1991 in the neighbourhood of San Francisco (Bilbao) in order to promote and contribute in the development of men and women in risk of social exclusion. Nowadays, it is extended to some other places in Bizkaia, above all, in areas or neighbourhoods with socio-economical disadvantages, such as Erandio, Getxo, Leioa, Astrabudua and Bilbao.

Sortarazi has several social projects, but we will focus in the one called *Hazia project.* It pretends to help people and families improving their personal, familiar and social situation so that they can improve their social and labor integration possibilities. They want to treat specifically: orientation and emotional support, individual and group psicological support, advance of municipal help, material basic necessities (payment of gas, water and electricity supply, rents, school expenses) and basic necessities of feeding (food delivery and baby set). Among all the services that are offered, in this project we will be dealing with the last support, food delivering, for those families that do not receive social benefits neither any other income.

In 2013 there were 545 families assisted, adding up to 1404 people. Every family unit is summoned monthly making a week delivery in tuesday, wednesday or thursday in the afternoon, where there can be assisted no more than 40 families per day.

The service continues thanks mainly to the Food Bank of Bizkaia, public and private support, as well as individual cooperations.

4.2 Diet stochastic models and alternative models

Let us explain the application model, based on the well known Diet Problem where the horizon planning is a month. As we have said before, there are two main goals in this application: (1) minimize the cost of a monthly nutrition of around 900 users, supplying the maximum of the nutrients requirement and (2) compute the minimum cost of a monthly nutrition of around 900 users, improving the percentage of the nutrients requirement, without exceeding the budget. The nutrients analized are those which appear in the product labels: energy, fat, saturated fat (SF), carbohydrates (CH), fibre, proteins and salt.

Notice that we can consider three sources of uncertainty: prices and nutrients can change according to the available brand in the market and the number of users can vary slightly from the expected demand.

The problem is composed by the following sets, parameters and variables:

Sets:

Ω, the set of scenarios which represents uncertainty

I, the set of products supplied by Hazia Project and

J, the set of nutrients analized from each product

Deterministic parameters:

$c = (c_i)$, the price vector of the products i needed to be purchased in the initial period of planning horizon, $i \in I$

$A = (a_{ij})$, the first stage constraints matrix of the nutrient j corresponding to the product i, $i \in I$ and $j \in J$

$T = (t_{ij})$, the technological matrix of the nutrient j of the product i, $i \in I$ and $j \in J$

$\underline{b}^0 = (\underline{b}_j^0)$, the minimum monthly nutrients requirement for a healthy diet per person

$\overline{b}^0 = (\overline{b}_j^0)$, the maximum monthly nutrients requirement for a healthy diet per person

N^0, the average amount of people demanding the service monthly

$\underline{b} = (\underline{b}_j)$, the minimum nutrients requirement for a healthy diet monthly for N^0 people, $\underline{b} = \underline{b}^0 \cdot N^0$

$\overline{b} = (\overline{b}_j)$, the maximum nutrients requirement for a healthy diet monthly for N^0 people, $\overline{b} = \overline{b}^0 \cdot N^0$

α_1, the minimum percentage of the requirement that must be satisfied in the initial period of planning horizon

α_2, the maximum percentage of the requirement that must be satisfied in the initial period of planning horizon

$\beta = (\beta_j)$, the minimum percentages of nutrient requiment j that must be satisfied for the planning horizon

$\delta = (\delta_i)$, $\delta' = (\delta'_i)$, ratios for determining lower bounds for X and Y variables, respectively

Z^0, the initial budget

Stochastic parameters:

$p = (p^\omega)$, the likelihood of each scenario ω, $\omega \in \Omega$

$q = (q_i^\omega)$, the price vector of the products i needed to be purchased in the middle of planning horizon, $i \in I$

$W = (w_{ij}^\omega)$, the recourse matrix of the nutrient j of the product i, $i \in I$ and $j \in J$

$N = (N^\omega)$, the amount of people demanding the service in a month, $\omega \in \Omega$

$\underline{h} = (\underline{h}_j^\omega)$, the minimum nutrients requirement j for a healthy diet, $j \in J$ and $\omega \in \Omega$, $\underline{h} = \underline{b}^0 \cdot N$

$\overline{h} = (\overline{h}_j^\omega)$, the maximum nutrients requirement j for a healthy diet, $j \in J$ and $\omega \in \Omega$, $\overline{h} = \overline{b}^0 \cdot N$

Variables:

$X = (X_i)$, the vector of the products i needed to be purchased at the initial period of planning horizon, $i \in I$

$Y = (Y_i^\omega)$, the vector of the products i needed to be purchased in the middle of planning horizon depending on the scenario ω, $i \in I$ and $\omega \in \Omega$

The two-stage diet Stochastic Problem reads as follows:

$$
\begin{aligned}
\min \quad & \sum_{i\in I} c_i X_i + \sum_{\omega\in\Omega} p^{\omega} \left(\sum_{i\in I} q_i^{\omega} Y_i^{\omega} \right) && && (4.1)\\
\text{s.t.} \quad & \alpha_1 \cdot \underline{b}_j \le \sum_{i\in I} a_{ij} X_i \le \alpha_2 \cdot \overline{b}_j, && \forall j \in J && (4.2)\\
& \beta_j \cdot \underline{h}_j^{\omega} \le \sum_{i\in I} t_{ij} X_i + \sum_{i\in I} w_{ij}^{\omega} Y_i^{\omega} \le \overline{h}_j^{\omega}, && \forall j \in J,\ \forall \omega \in \Omega && (4.3)\\
& \sum_{i\in I} c_i X_i + \sum_{\omega\in\Omega} p^{\omega} \left(\sum_{i\in I} q_i^{\omega} Y_i^{\omega} \right) < Z^0 && && (4.4)\\
& X_i \ge \delta_i \cdot \sum_{i\in I} \overline{X}_i, && \forall i \in I && (4.5)\\
& Y_i^{\omega} \ge \delta_i' \cdot \sum_{i\in I} \overline{Y}_i^{\omega}, && \forall i \in I,\ \forall \omega \in \Omega && (4.6)\\
& X_i, Y_i^{\omega} \ge 0, && \forall i \in I,\ \forall \omega \in \Omega && (4.7)
\end{aligned}
$$

where (4.1) represents the objective function to optimize, the minimum cost, (4.2) is the set of constraints of the first stage, (4.3) are the global nutrient requirement, (4.4) defines the budget limitation, (4.5) and (4.6) denote the bounds for the variables X and Y^{ω}, where $\overline{X}_i$ and $\overline{Y}_i^{\omega}$ are solutions obtained from non-balanced models, respectively and (4.7) declares the non-negative of the variables.

We will consider four models, according to different goals:

Model 1.a. The goal is to obtain the product distribution that minimizes the cost in order to satisfy entirely the nutrition requirement. In this case, $Z^0 = +\infty$, $\beta_j = 1, \forall j \in J$ and $\delta_i = \delta_i' = 0, \forall i \in I$. That is, (4.4)-(4.6) are inactive.

Model 2.a. The goal is to improve the association standard nutritional bounds, without exceeding the initial budget, Z^0. $\delta_i = \delta_i' = 0, \forall i \in I$, that is, (4.5) and (4.6) are inactive.

Model 1.b. The goal is the same as Model 1.a., but mantaining a balanced diet. In this case, (4.5) is applied, but $Z^0 = \infty$, that is, (4.4) is inactive.

Model 2.b. The goal is the same as Model 1.b., but mantaining a balanced diet. In this case, (4.5) is applied. All the constraints are active.

For comparison purposes, we will also calculate the problems WS(4.8), EV(4.9) and EEV(4.10):

- $WS = E_\xi[min\ z(X, \xi)] = E_\xi[Z(X^*(\xi), \xi)]$,

$$\text{where } X^*(\xi) \text{ corresponds to the optimal solution of WS} \tag{4.8}$$

- $EV = min\ z(X, \bar{\xi})$,

$$\text{where } \xi^* = E[\xi] \tag{4.9}$$

- $EEV = E_\xi[Z(X^*(\xi^*), \xi)]$,

$$\text{where } X^*(\xi^*) \text{ corresponds to the optimal solution of EV} \tag{4.10}$$

4.3 Dataset

The data for the models has been collected from the Sortarazi's storehouse and from standard nutritional references.

4.3.1 First stage prices and nutrients

The values of energy, macronutrients and minerals (Na) for 100g of each product, have been taken from AESAN/BEDCA, the spanish database of food composition (2010), see [36], taking into account the product sets, sizes and prices of the association warehouse, see Table 4.1. These are, the values for first stage A matrix, second stage T technology matrix and c vector.

Table 4.1: Energy, macronutrients and minerals of each product by BEDCA

Reference	BEDCA	Size	Energy (kcal)	Fat (g)	SF (g)	CH (g)	Fibre (g)	Proteins (g)	Salt (g)	Price (€)
2541	Oil	1l	8870.00	999.00	106.20	0.00	0.00	0.00	0.00	2.85
1117	Pasta	500g	1765.00	7.25	0.95	354.50	25.00	62.50	0.06	1.03
1020	Chickpea	1kg	3360.00	63.00	4.28	492.50	149.70	193.10	0.75	2.89
2675	Maria Biscuits	200g	958.00	38.00	19.40	138.00	6.20	15.60	1.09	0.46
2660	Rice	1kg	3870.00	9.00	2.10	860.00	14.00	70.00	0.15	2.05
2648	Cacao	1kg	3900.00	40.00	2.50	810.20	35.00	58.80	12.60	2.09
2131	Tuna	80g	164.00	9.68	1.44	0.00	0.00	19.05	0.74	0.75
2632	Sardine	115g	272.55	18.11	3.22	0.00	0.00	27.52	1.05	1.85
2560	Fried Tomato	390g	327.60	24.96	2.42	20.28	6.24	5.85	3.32	0.66
2493	Whole milk	1l	650.00	38.00	23.00	47.00	0.00	30.60	1.20	0.95
1109	Coffee	250g	277.50	0.75	0.30	20.00	0.00	46.25	0.25	1.20

4.3.2 Second stage prices and nutrients

Comparing products by brand, we have notified that difference in nutrients could be relevant. In order to analyse it, we have considered two brands for each product, collecting the nutritional information from the warehouse and corresponding prices have been collected from EROSKI web side [38], see in Table 4.2. These are the values needed for q^0 (reference second stage price) and W^ω recourse matrix (nutrients for second stage).

Notice that $q^\omega \in \{q^0 \cdot 0.9,\ q^0 \cdot 1.1\},\ \omega \in \Omega$ are equally likely, since in the second stage price there is a growth of -10% or 10%.

Table 4.2: Energy, macronutrients and minerals of each product taken from the Food Bank

	Brand	Size	Energy (kcal)	Fat (g)	SF (g)	CH (g)	Fibre (g)	Proteins (g)	Salt (g)	Price (€)
Oil	1	1l	9000.00	1000.00	130.00	0.00	0.00	0.00	0.00	1.99
	2		8240.00	920.00	140.00	0.00	0.00	0.00	0.00	3.99
Pasta	1	500g	1720.00	10.00	2.50	340.00	15.00	60.00	0.15	0.87
	2		1745.00	7.50	1.50	350.00	18.00	60.00	0.15	0.60
Chickpea	1	1000g	3320.00	70.00	10.00	360.00	266.00	179.00	0.00	1.49
	2		3480.00	58.00	11.00	460.00	140.00	210.00	0.30	2.89
Maria Biscuits	1	200g	872.00	20.00	10.00	156.00	5.80	14.00	1.46	1.12
	2		842.00	16.00	7.60	158.00	4.80	14.00	1.66	1.29
Rice	1	1000g	3540.00	5.00	1.00	795.00	28.00	65.00	0.00	1.52
	2		3470.00	11.00	3.00	740.00	35.00	84.00	0.00	1.19
Cacao	1	1000g	3800.00	38.00	28.00	780.00	85.00	65.00	0.00	2.39
	2		3760.00	24.00	15.00	780.00	78.00	68.00	1.40	6.25
Tuna	1	80g	158.40	8.00	1.20	0.00	0.00	21.60	0.96	0.66
	2		201.60	14.40	2.32	0.48	0.00	17.60	0.50	2.04
Sardine	1	115g	241.50	13.80	2.99	1.15	0.00	28.75	1.73	0.94
	2		236.90	13.23	2.99	1.04	23.00	28.29	0.46	1.32
Fried Tomato	1	390g	284.70	13.65	1.56	35.49	0.00	4.29	5.85	0.74
	2		284.70	13.65	1.56	31.20	1.95	4.68	3.51	0.40
Whole milk	1	1l	630.00	36.00	24.00	46.00	0.00	31.00	1.00	0.59
	2		630.00	36.00	25.00	46.00	0.00	30.00	1.30	0.79
Coffee	-	250g	245.10	10.50	3.50	14.00	24.50	17.50	0.00	1.20

To determine the scenario tree for W^ω, we have compared all the values of Table 4.2 in order to choose the most different pair of products. First of all, the relative difference of each product with regard to the real value of BEDCA have been calculated. Then, the absolute difference of each pair of relative differences of each product have been computed. Finally, the mean ($\overline{x}$), standard deviation (s) and the Coefficients of Variation (CV) have been obtained, see Table 4.3.

Table 4.3: Selection of the most different products

Product	$\overline{x}$	s	CV
Oil	4.09	4.54	**111.16**
Pasta	22.77	33.47	146.95
Chickpea	27.87	26.59	**95.40**
Maria Biscuits	11.77	11.96	**101.58**
Rice	14.86	17.24	**116.02**
Cacao	35.62	61.55	172.82
Tuna	32.64	54.05	165.57
Sardine	4.34	7.02	161.84
Fried Tomato	19.62	24.25	**123.62**
Whole milk	4.30	7.41	172.35

The criteria for selecting a pair of products to be considered in the scenario tree is as follows. The lower the CV, the bigger the convenience of introducing the product nutrients in the scenario tree. In effect, the bigger the mean and the smaller the standard deviation, the more relevant to take into account the differences. The five most different pairs of products, according to the lowest CV are Chickpea, Maria Biscuits, Oil, Rice and Fried Tomato.

4.3.3 Nutricional bounds

For determing bounds on nutritional requirements, we must take into account nutritive properties. For first stage, around half of the requirements should be satisfied (determined by α_1, α_2 parameters near to 0.5). Moreover, for the second stage (almost) all of them, where low bound is fixed by β parameter near to 1.

As stated in BEDCA [36], fat, like carbohydrates and proteins, is considered a major macronutrient for the body due to the energy they provide (9Kcal/g) and the quantity of functions of the organism in which participates. However, it should be only 30%-35% of the energy. It is not only that important the quantity but the quality of the fat. So, less than 7%-8% of the fat must be saturated.

Carbohydrates contain monosaccharides (glucose, fructose and galactose), disaccharides (sacarose, lactose) and polisaccharides (starch and glycogen). They provide 4Kcal/g and they should be 50%-60% of the calories consumed per day.

Fibre is a plant material that cannot be digested but helps us to digest other food. For a healthy diet, it must be consumed more than 14g/1000Kcal.

Protein is a building material of our skin, bones, muscles and other tissues in the body among other things. In fact, the vast majority of the biological functions are carried out by proteins.

Therefore, the daily intake of these nutrients is essencial for a healthy diet. It provides 4Kcal/g and it must be 10%-15% of the daily total calories.

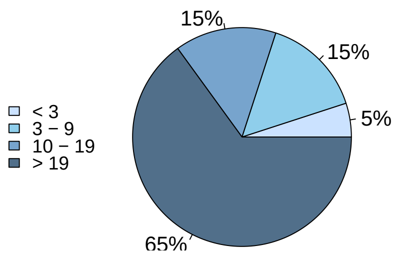
Sodium (Na) is the major positive ion (cation) in fluid outside of cells. When it is combined with chloride (Cl), the resulting substance is table salt (NaCl). Excess sodium can cause cells to malfunction, so 5g/day or less of salt is good enough for a healthy diet, where 5g of salt corresponds to 2000mg of Na.

For computing bounds, first of all we need to estimate the kcal per person and day. Taking into account the user age distribution (see 4.1) and according to Carbajal 2013 [6] and Spanish Nutritional Foundation (FEN) 2013 [11], the estimated energy requirement is shown in Table 4.4.

Table 4.4: Energy estimation

Age	Percentage	Energy (kcal/pers)
< 3	5 %	800
$3 - 9$	15 %	1737.5
$10 - 19$	15 %	2560
> 19	65 %	2428.22
Total	100 %	2263

Figure 4.1: Pie chart of users for age range



Hence, bounds of 2200-2600 kcal of energy have been stated and bounds of every nutrient is summarized in Table 4.5. These are the values needed for b^0. Notice that $(\underline{b}, \overline{b}) = (\underline{b}^0, \overline{b}^0) \cdot N^0$ and $(\underline{h}, \overline{h}) = (\underline{b}^0, \overline{b}^0) \cdot N$, where $N^0 = 900$ users and $N \in \{800, 900, 1000\}$ with probability $\{0.15, 0.7, 0.15\}$, respectively.

Table 4.5: Energy, macronutrients and mineral requirement bounds

	Energy(kcal)	Fat(g)	Sat.Fat(g)	CH(g)	Fibre(g)	Proteins(g)	Salt(g)
Minimum	2200	73.7	-	275	30.8	55	-
Maximum	2600	87.1	6.97	325	-	65	5

4.3.4 Data preliminar analysis

As a preliminar studio, let us consider the nutritional requirement satisfaction with respect to nowadays association products distribution. Table 4.6 and Figures 4.2 and 4.3 show the estimated frequency and percentage of each family type and users number demanding the services of Hazia project monthly, from set $np = 1$ to set $np = 6$, where np is the number of users per family.

Table 4.6: Monthly users distribution

np	# families	% families	# people	% people
1	180	44%	180	20%
2	81	20%	162	18%
3	72	18%	216	24%
4	45	11%	180	20%
5	22	5%	108	12%
6	9	2%	54	6%
Total	409	100%	900	100%

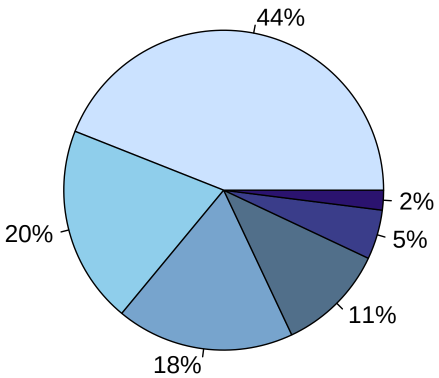


Figure 4.2: Pie chart for families

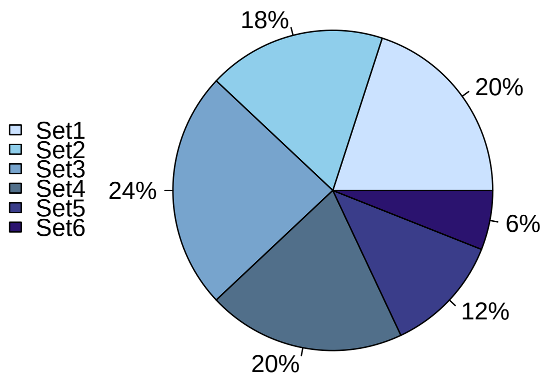


Figure 4.3: Pie chart for users

In Table 4.7 are given the sets of non-perishable products that are distributed monthly according to the number of people and their respective costs. The data can change depending on the product availability and the demand. We have considered the families distribution shown previously for generating the last set, this called **representative set**, for $N^0 = 900$ users and its estimation of monthly products distribution. Notice that δ_i^0 is the percentage of products units, to be considered in the balanced diet models.

Table 4.7: Monthly set of products for sets of np users

Product	Size	Price(€)	Set						Rep. Set	%
			1	2	3	4	5	6	900	δ_i^0
Oil	1l	2.85	1	1	2	3	3	4	641	2
Pasta	500g	1.03	2	4	5	6	7	9	1546	6
Chickpea	1000g	2.89	1	2	3	4	5	6	900	3
Maria Biscuits	200g	0.46	2	3	4	5	6	7	1309	5
Rice	1000g	2.05	2	2	3	3	3	4	974	4
Cacao	1000g	2.09	-	-	-	-	-	-	-	-
Tuna	80g	0.75	18	36	54	72	90	108	16200	58
Sardine	115g	1.85	2	2	3	4	5	6	1080	4
Fried Tomato	400g	0.66	3	4	6	8	10	12	1980	7
Whole milk	1l	0.95	3	6	9	12	15	18	2700	10
Coffee	250g	1.20	1	1	1	1	1	1	409	1
Total		27200€							27738	100%

Consequently, the average of nutrients, per person and day, satisfied by the sets described previously, according to BEDCA, are given in Table 4.8. In Table 4.8 and Figure 4.4 is also shown the percentages of the minimum nutrients satisfied by the sets defined in Table 4.7 according to the recommended healthy diet from Table 4.5. Later on, we will compare the nutrients satisfied by the representative set of 900 people of users and the model proposed in this project.

Table 4.8: Nutrients satisfied per person and day

np	Energy		Fat		SF	CH		Fibre		Proteins		Salt	Na
	(kcal)	(%)	(g)	(%)	(g)	(g)	(%)	(g)	(%)	(g)	(%)	(g)	(mg)
1	1071	49	58	78	9	114	41	9	29	35	63	1.09	434
2	753	34	37	50	6	82	30	8	26	30	55	0.91	365
3	776	35	41	56	7	77	28	8	25	29	53	0.90	362
4	755	34	43	59	7	68	25	7	24	28	51	0.90	360
5	684	31	38	51	7	62	22	7	23	27	50	0.90	358
6	716	33	40	54	7	65	24	7	24	28	50	0.90	358
900	812	37	43	59	7	81	29	8	26	30	54	0.94	376

Table 4.9: Satisfied energy, macronutrients and mineral

	Energy(%)	Fat(%)	CH(%)	Fibre(%)	Proteins(%)
1	49	78	41	29	63
2	34	50	30	26	55
3	35	56	28	25	53
4	34	59	25	24	51
5	31	51	22	23	50
6	33	54	24	24	50
900	37	59	29	26	54

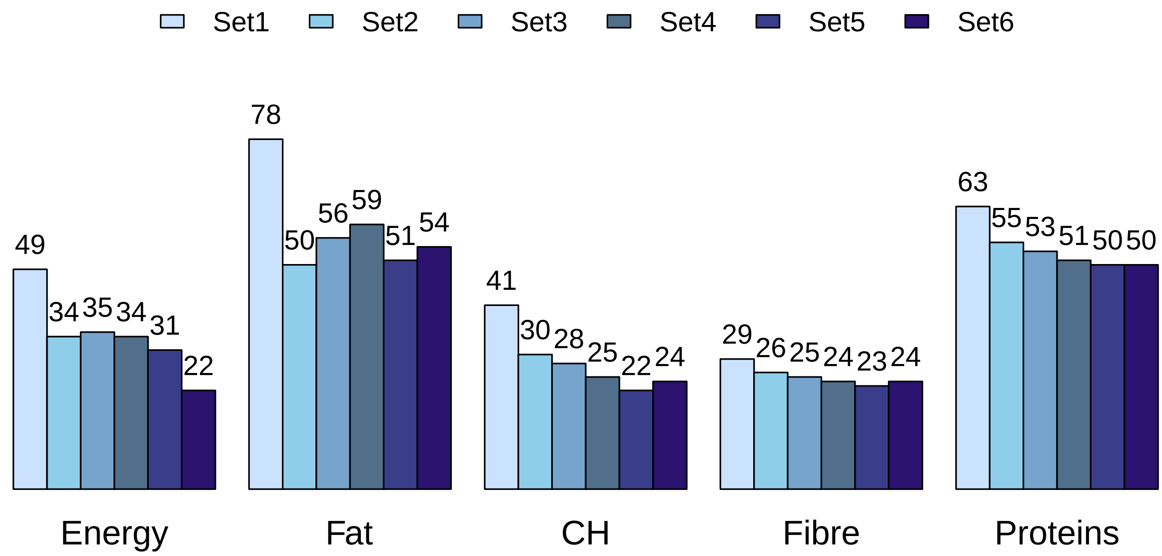


Figure 4.4: Energy, macronutrients and mineral % satisfaction depending on the set

We can observe that the nutritional requirements are not at all satisfied, and specially, proteins satisfaction is poorly guaranteed.

4.4 Analysis and Results

During this project we have analysed four different scenario trees. The scenario tree generation considers a triplet random vector: $\xi^\omega = (W^\omega, q^\omega, h^\omega),\ \omega \in \Omega$, where $|\Omega| = |\Omega_1| \cdot |\Omega_2| \cdot |\Omega_3| = 2^n \cdot 2 \cdot 3$.

Ω_1 represents the scenario subset for W, the nutrient matrix for second stage, where n is the number of product pairs to consider because of its nutritional differences, according to the CV, see Table 4.3, that is, $n \in \{2, 3, 4, 5\}$. Notice that they are equally likely.

Ω_2 is the scenario subset for q, prices random vector, where a -10% and 10% growth rates have been considered equally likely from a reference value. For those product pairs considered due to its nutritional difference, see Table 4.2, and for the others, the mean values have been considered.

Ω_3 is the scenario subset for N, the random variable for the users number demanding this service. Let us consider 800, 900 or 1000 users with probability $\{0.15, 0.7, 0.15\}$, respectively.

The, let us assume that those $\frac{|\Omega|}{3}$ scenarios with 900 users have probability $0.5^n \cdot 0.5 \cdot 0.70$ and the other $\frac{2}{3} \cdot |\Omega|$ scenarios occur with probability $0.5^n \cdot 0.5 \cdot 0.15$.

Depending on n pair of products, we will consider

- $|\Omega| = 24$, Chickpea and Maria Biscuits ($n = 2$)
- $|\Omega| = 48$, Oil added ($n = 3$)
- $|\Omega| = 96$, Rice added ($n = 4$)
- $|\Omega| = 192$, Fried Tomato added ($n = 5$)

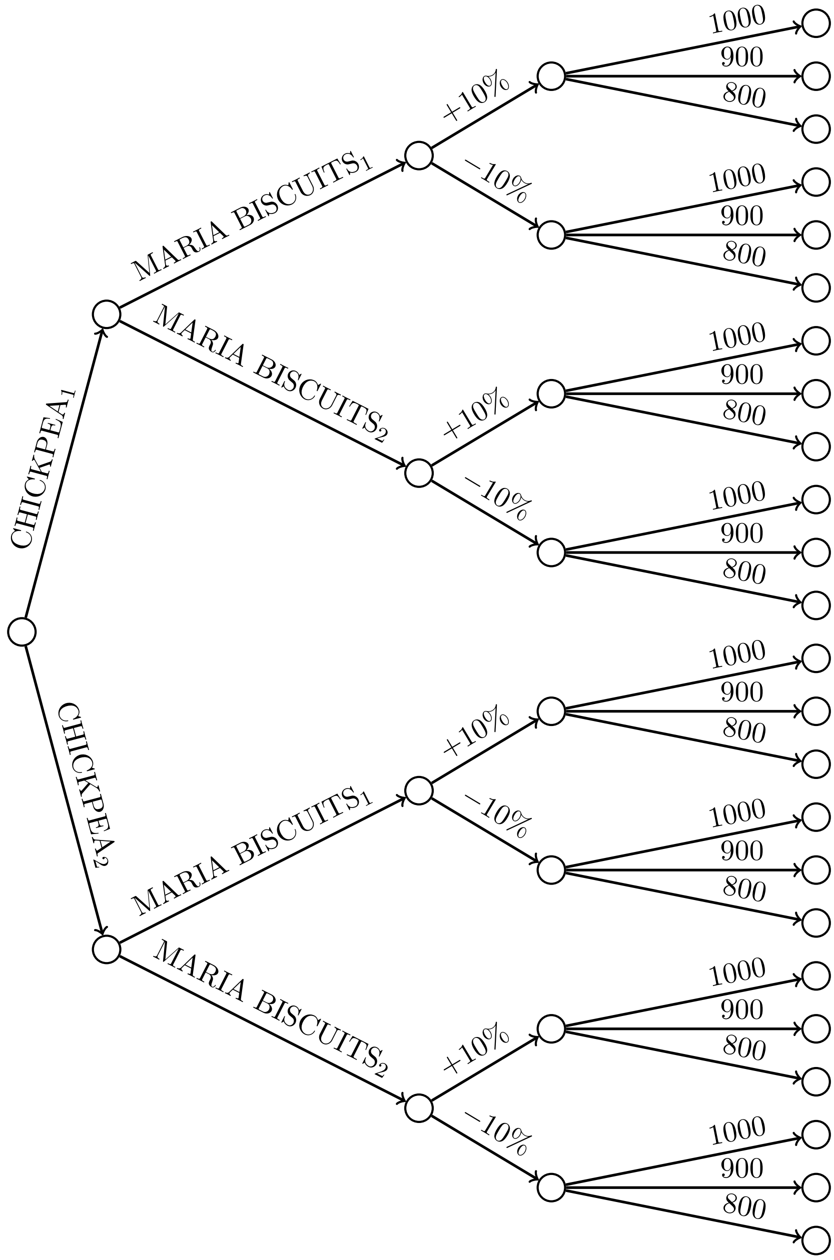
4.4.1 Solutions for Model 1.a.

Remember that the goal of Model 1.a. is to obtain the product distribution that minimizes the cost in order to satisfy entirely the nutrition requirement and the model is (4.1)-(4.3), (4.6). We have considered $\alpha_1 = 0.5$ and $\alpha_2 = 0.7$ for first stage nutritional bounds in all the models. Depending on the number of scenarios, Table 4.10 shows the solutions obtained from the Stochastic Problem (SP), Wait-and-See (WS) model (4.8), the Expected Value problem (4.9) and Expected result of using EV solution (EEV) (4.10), and EVPI and VSS measures.

Table 4.10: Comparison of values EV, WS, SP, EEV, EVPI and VSS for Model 1.a.

	EV	WS	SP	EEV	EVPI	VSS
$\|\Omega\| = 24$	30435	29659	31287	∞	1628	∞
$\|\Omega\| = 48$	30435	30934	32302	∞	1368	∞
$\|\Omega\| = 96$	30435	29642	31279	∞	1636	∞
$\|\Omega\| = 192$	30435	29642	31279	∞	1636	∞

Figure 4.5 shows Z^{ω}_{WS} and Z^{ω}_{SP} for Model 1.a. and $|\Omega| = 24$.



		Z^{ω}_{WS}	Z^{ω}_{SP}
Z^{1}	=	23100	24714
Z^{2}	=	26163	27158
Z^{3}	=	29867	31911
Z^{4}	=	24329	25804
Z^{5}	=	27857	28791
Z^{6}	=	32060	34600
Z^{7}	=	23100	24714
Z^{8}	=	26163	27158
Z^{9}	=	29867	31911
Z^{10}	=	24329	25804
Z^{11}	=	27857	28791
Z^{12}	=	32060	34600
Z^{13}	=	39192	42547
Z^{14}	=	33087	35225
Z^{15}	=	27638	30776
Z^{16}	=	36234	38413
Z^{17}	=	31210	32422
Z^{18}	=	26348	28782
Z^{19}	=	39192	42547
Z^{20}	=	33087	35225
Z^{21}	=	27638	30776
Z^{22}	=	36234	38413
Z^{23}	=	31210	32422
Z^{24}	=	26348	28782

Figure 4.5: Optimal WS, SP and EEV solutions for $|\Omega| = 24$ scenarios

Notice that in all the cases it is really worthy to consider uncertainty. EEV is ∞ because the solution provided by the Expeted Value is not implementable under all the set of scenarios. Moreover, we can see that the case with more scenarios has higher EVPI. We can also remark that with 96 and 192 scenarios, solutions are the same. This happens because the last product branched, fried tomato, has no relevance.

Table 4.11 reveal the variables $X + Y$ in the best and the worst scenarios accoding to each model ($|\Omega| = 192$), that is, those with minimum and maximum cost, respectively. It expresses the amount of each product that must be purchased in the initial period of planning horizon and in the middle of planning horizon, according to the brand. Since EEV problem is infeasible, we have omited it from the following tables. Non-displayed products have not been selected.

Table 4.11: Total decisions for Model 1.a.

Model 1.a.	$X_{EV} + Y_{EV}$	$X_{WS}^{best} + Y_{WS}^{best}$	$X_{WS}^{worst} + Y_{WS}^{worst}$	$X_{SP} + Y_{SP}^{best}$	$X_{SP} + Y_{SP}^{worst}$
Oil	1057	1051	1219	1018	83
Chickpea	6642	5697	5416	-	8448
Rice	3642	4471	2814	2227	2102
Cacao	3519	2137	5234	4351	-
Fried Tomato	-	-	-	-	5636
Z (€)	30435€	23003€	39985€	24653€	43999€

The selection of products according to their brands is as follows: for the best scenario in SP, we will choose Oil 'Koipesol', Chickpeas 'Eroski', Rice 'La Cigala' and Fried Tomato 'Eroski', for the worst, the opposite brands, for the best scenario in WS, it is chosen the same as in SP except Oil and for the worst, the same as SP except Fried Tomato.

Figures 4.6 and 4.7 show the histograms obtained for the random variable Z (cost) where WS and SP models are solved, respectively.

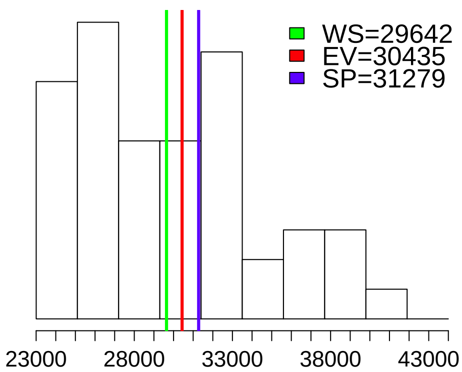


Figure 4.6: Histogram of Z_{WS} for Model 1.a.

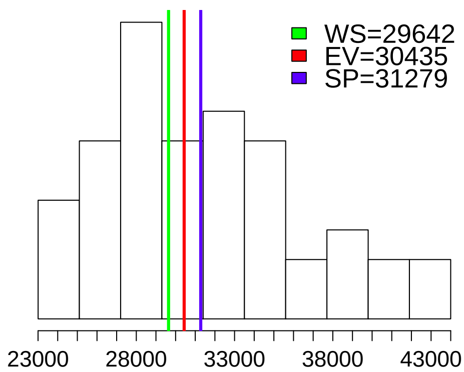


Figure 4.7: Histogram of Z_{SP} for Model 1.a.

Notice that means of the variable cost, Z, according to each problem WS, SP and EV, are expressed by the vertical lines.

In the Figure 4.7 we can observe that the inequalities of the Proposition 3.1 are strictly achieved for minimization: $WS < SP < EEV$. This means that the problem WS offers better results than SP, whereas EEV has the highest cost, in this case ∞, because it is not implementable in the 2% of the scenarios. However, Proposition 3.2 is not satisfied, because $EV \nleq WS$, remember that objective coefficients q and recourse matrix W are not fixed. The distance between both vertical lines, WS and SP, is the value of EVPI, the longer the distance is, the more importance the uncertainty has.

The products distribution in the worst situation, that is, expensive prizes (+10%) and maximum users demand (1000 people) is expressed in Table 4.12. It guarantees to cover all the nutrients required, with a total cost of 33148€.

Table 4.12: Monthly set of products for np people for Model 1.a.

Product	Brand	Unit	Price(€)	Set					
				1	2	3	5	6	7
Oil	Koipesol	1l	2.85	2	3	2	2	4	9
Chickpea	Eroski	1000g	2.89	7	15	22	29	36	44
Rice	La Cigala	1000g	2.05	4	7	11	14	18	21
Cacao	-	1000g	2.09	4	9	13	17	22	26
Fried Tomato	Eroski	400g	0.66	6	11	17	23	28	34

4.4.2 Solutions for Model 2.a.

Remember that the goal of Model 2.a. is to improve the association standard nutritional bounds, without exceeding the initial budget and the model is (4.1)-(4.4), (4.6). We have considered Z^0 equal to the budget given in Table 4.7 and $\underline{\beta}_j$ at least 5% more than the highest value in Table 4.9, according to each nutrient except proteins, which will be satisfied completely. Since it is considered the most important macronutrient in the diet. Table 4.13 shows the solutions obtained from the SP, WS, EV and EEV models and EVPI and VSS measures.

Table 4.13: Comparison of values EV, WS, SP, EEV, EVPI and VSS for Model 2.a.

	EV	WS	SP	EEV	EVPI	VSS
$\|\Omega\| = 24$	22474	22063	22439	∞	376	∞
$\|\Omega\| = 48$	22474	22151	22922	∞	771	∞
$\|\Omega\| = 96$	22474	21564	22299	∞	735	∞
$\|\Omega\| = 192$	22474	21564	22299	∞	735	∞

As stated in previous models, Table 4.14 reveals the decision of the amount of each product that must be purchased in the initial and middle of planning horizon. EV again is not always implementable.

Table 4.14: Total decisions for Model 2.a.

Model 2.a.	$X_{EV}+Y_{EV}$	$X_{WS}^{best}+Y_{WS}^{best}$	$X_{WS}^{worst}+Y_{WS}^{worst}$	$X_{SP}+Y_{SP}^{best}$	$X_{SP}+Y_{SP}^{worst}$
Oil	1165	994	1053	1177	959
Pasta	16052	16565	12669	16160	23419
Chickpea	1849	1396	4205	1598	1909
Rice	2320	1422	3564	842	842
Z	22474	17384	27129	18322	27185

In Model 2.a. the best scenarios follow the same steps as the best scenario in SP for Model 1.a., and the worst, the same as the worst scenario in WS for Model 1.a.

The nutrients satisfaction by the representative set of 900 users of Sortarazi compared with the one created by this Model 2.a. is shown in Figure 4.8.

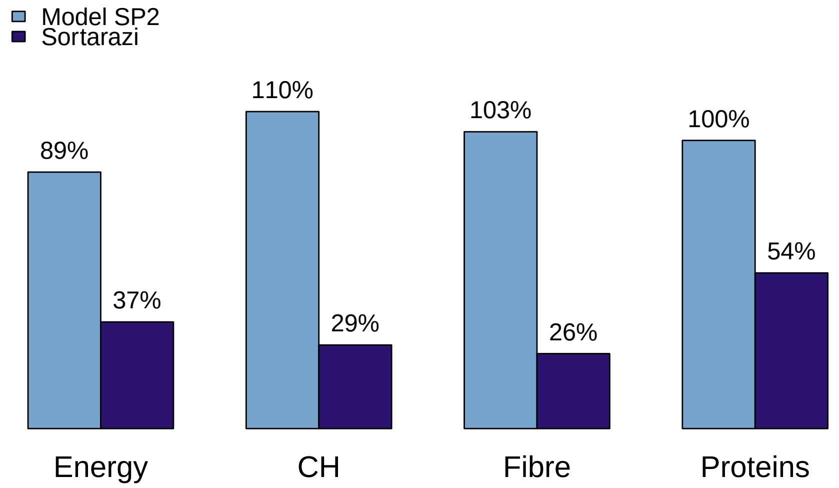


Figure 4.8: Comparison of nutrients requirement satisfaction for Model 2.a.

Since nutrients satisfaction have improved, the new recommendation of products supplyment is shown in Table 4.15, with a total cost of 23276€, even smaller than the initial budget.

Table 4.15: Monthly set of products for np people for Model 2.a.

Product	Brand	Unit	Price(€)	Set					
				1	2	3	4	5	6
Oil	Koipesol	1l	2.85	1	3	4	6	7	8
Pasta	-	500g	1.03	16	32	48	65	81	97
Chickpea	Eroski	1000g	2.89	2	4	6	9	11	13
Rice	La Cigala	1000g	2.05	3	6	9	12	15	19

4.4.3 Solutions for Model 1.b.

Remember that the goal of Model 1.b. is to obtain the product distribution that minimizes the cost and mantains a balanced diet satisfying all the nutrition requirement and the model is (4.1)-(4.6). Although Models 1.a. and 2.a. reach our goals, we have realized that the new product distribution would not imply a balanced diet, but the minimum cost for the nutritional bounds fixed. So, we have decided to create another two models, with the same goals as before, but adding two new constraints (4.5) and (4.6), where $\delta_i = 0.15 \cdot \delta_i^0$ for X and $\delta_i' = 0.10 \cdot \delta_i^0$ for Y^ω. δ_i^0 is the percentage of each product and it is shown in Table 4.7. Notice that we have chosen 0.15 and 0.10 in order to mantain feasibility.

For the sake of simplification, we are going to show only the analysis of all the problems for the biggest scenario set, $|\Omega| = 192$, in the balanced models. Table 4.16 summarizes the optimal objective function for SP, WS, EV and EEV, and EVPI and VSS measures.

Table 4.16: Comparison of values EV, WS, SP, EEV, EVPI and VSS for Model 1.b.

	EV	WS	SP	EEV	EVPI	VSS
$\lvert\Omega\rvert = 192$	36021	34542	36930	∞	2389	∞

The total amount of products, that must be provided in the best and worst situations, is given in Table 4.17. Since some Z_{EEV} problems are infeasible, we have omited EEV from the solutions.

Table 4.17: Best and worst cases depending on the problem for Model 1.b.

Model 1.b.	$X_{EV} + Y_{EV}$	$X_{WS}^{best} + Y_{WS}^{best}$	$X_{WS}^{worst} + Y_{WS}^{worst}$	$X_{SP} + Y_{SP}^{best}$	$X_{SP} + Y_{SP}^{worst}$
Oil	949	870	1024	885	875
Pasta	332	286	281	269	322
Chickpea	5786	7133	7080	5184	6859
Maria Biscuits	112	81	78	70	105
Rice	1166	1334	1633	784	1837
Cacao	6047	4960	6376	5748	5757
Tune	1625	417	1129	1008	1523
Sardine	112	81	78	70	105
Fried Tomato	1638	142	4325	4289	16883
Whole Milk	280	203	195	174	263
Coffe	28	89	19	17	26
Z(€)	36021€	26981€	53709€	29802€	52207€

The selection of products according to their brands follows the same steps as Models 1.a. and 2.a., except for Fried Tomato, that it is always chosen the first brand.

Finally, the new recommendation of products supplyment in the worst situation is shown in Table 4.18, with a total cost of 41641€.

Table 4.18: Monthly set of products for np people for Model 1.b.

Product	Brand	Unit	Price(€)	Set					
				1	2	3	4	5	6
Oil	Koipesol	1l	2.85	1	2	3	4	4	5
Pasta	-	500g	1.03	-	1	1	1	2	2
Chickpea	Eroski	1000g	2.89	7	14	21	27	34	41
Maria Biscuits	Gullon	200g	0.46	-	-	-	-	1	1
Rice	La Cigala	1000g	2.05	2	4	6	7	9	11
Cacao	-	1000g	2.09	6	12	17	23	29	35
Tuna	-	80g	0.75	2	3	5	6	8	9
Sardine	-	115g	1.85	-	-	-	-	1	1
Fried Tomato	Eroski	390g	0.66	17	34	51	68	84	101
Whole milk	-	1l	0.95	-	1	1	1	1	2

4.4.4 Solutions for Model 2.b.

Once again, we will compare all the situations according to the biggest number of scenarios for Model 2.b., considering that the budget is 27200€.

In Table 4.19 is given the optimal objective function for each problem: SP, WS, EV and EEV and EVPI and VSS measures.

Table 4.19: Comparison of values EV, WS, SP, EEV, EVPI and VSS for Model 2.b.

	EV	WS	SP	EEV	EVPI	VSS
$\|\Omega\| = 192$	23305	22760	22998	∞	239	∞

The total amount of products, that must be provided in the best and worst situations, is given in Table 4.20. Since some Z_{EEV} problems are infeasible, we have omited EEV from the solutions.

Table 4.20: Best and worst cases depending on the problem for Model 2.b.

Model 2.b.	$X_{EV} + Y_{EV}$	$X_{WS}^{best} + Y_{WS}^{best}$	$X_{WS}^{worst} + Y_{WS}^{worst}$	$X_{SP} + Y_{SP}^{best}$	$X_{SP} + Y_{SP}^{worst}$
Oil	58	678	58	542	83
Pasta	17884	15530	16930	15506	19295
Chickpea	951	1505	1803	1356	1118
Maria Biscuits	8974	144	12468	1875	10718
Rice	115	115	331	122	167
Cacao	29	29	29	36	42
Tune	1675	1675	1675	1778	2420
Sardine	115	115	115	123	167
Fried Tomato	202	202	202	215	292
Whole Milk	289	289	289	307	417
Coffe	29	29	29	31	42
Z(€)	21555€	18353€	24952€	18542€	27076€

The selection of products according to their brands follows the same steps as Model 1.b.

Finally we are going to compare the nutrients satisfied by the representative set of 900 people of Sortarazi with the one created by this application in Figure 4.9.

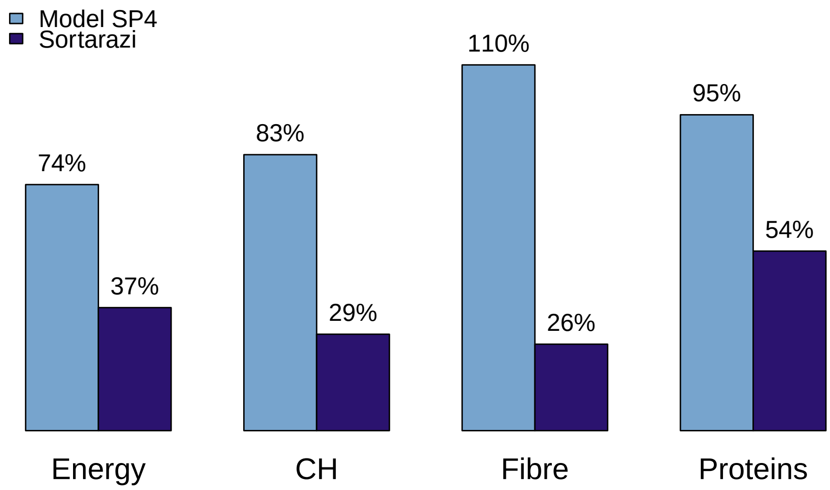


Figure 4.9: Comparison of percentages of nutrients satisfied for Model 2.b.

Since nutrients satisfaction have improved, the new recommendation of products supplyment is shown in Table 4.21, with a total cost of 23283€.

Table 4.21: Monthly set of products for np people for Model 2.b.

Product	Brand	Unit	Price(€)	Set					
				1	2	3	4	5	6
Oil	Koipesol	1l	2.85	1	2	3	3	4	5
Pasta	-	500g	1.03	16	31	47	62	78	93
Chickpea	Eroski	1000g	2.89	3	6	8	11	14	17
Maria Biscuits	Gullon	200g	0.46	2	4	6	8	9	11
Rice	La Cigala	1000g	2.05	1	1	2	3	3	4
Tuna	-	80g	0.75	2	4	6	8	10	11
Sardine	-	115g	1.85	-	-	-	1	1	1
Fried Tomato	Eroski	390g	0.66	-	-	1	1	1	1
Whole milk	-	1l	0.95	-	1	1	1	2	2

4.5 Conclusions and suggestions

In this project we have had two main goals: (1) obtain the product distribution that minimizes the cost in order to satisfy entirely the nutrition requirement and (2) obtain the product distribution which improves the association standard nutritional bounds, without exceeding the initial budget. We have reached our goals with the first two models, however, owing to the simplicity of the previous solutions, the few variety of providing products, we have decided to design two more models.

Let us compare the four models previously analysed, checking the last product distribution in Tables 4.12, 4.15, 4.18 and 4.21.

Model 1.a. Diet based in chickpeas and rice.

Model 2.a. Although it is provided some chickpeas and rice, this diet is basically based on pasta.

Model 1.b. Above all it is supplied chickpeas, rice, fried tomato and cacao. However, big families reach a bit more variety.

Model 2.b. It is given chickpeas and tune, among others, but above all it is provided pasta.

Nutritional satisfaction and the budget of each model are summarized in Table 4.22.

Table 4.22: Nutritional satisfaction and budget of generated models

	Energy (%)	CH (%)	Fibre (%)	Proteins (%)	Budget (€)
Model 1.a.	100	107	263	116	33148
Model 2.a.	89	105	102	98	23276
Model 1.b.	100	109	266	116	41641
Model 2.b.	74	83	110	95	23283

If we compare models according to non or balanced diet, that is, Model 1.a. with 1.b., and 2.a. with 2.b., we realize that the variety increases the price, but not the nutrional satisfaction (at least of the nutrients that we have analysed). And if we compare models according to objectives, that is, Model 1.a. with 2.a., and 1.b. with 2.b. we can consider that pasta is a cheap and satiating product and chickpeas has a lot of fibre.

Another way to have meal variety is altering standard deviation of the products amount, see [2]. It would be needed a new constraint that changes the problem from being linear to non-linear optimization.

What we want to express with this is that the variety is expensive, but maybe the minimum requirement satisfaction not that much. To sum up, we can say that a basic healthy diet is based on chickpeas and rice or pasta, as it is well known in social dinings.

Chapter 5

Conclusions

In this report basic concepts of the Stochastic Optimization have been studied. This discipline belongs to the Applied Mathematical area of Operations Research and Management Science, and deals with mathematical models under uncertainty.

First of all, we have defined and compared deterministic and stochastic programming in order to show the relevance of the latter. It is known that most of the Optimization problems in the real life present uncertain data and this is why Stochastic Programming is an alternative to model them. This project is focused in two-stage Stochastic Programming, where first stage decisions must be taken here-and-now, before uncertainty is revealed, and second stage-ones after it, providing decisions for the set of scenarios. There are explained some basic properties such as probability spaces and random variables, decisions and recourses and non-anticipativity principle. Stochastic Problems, satisfy explicitly or implicitly this principle, that is, there is a unique first-stage decision vector, X, whereas there are $|\Omega|$ second-stage decisions, Y^{ω}, one for each scenario. The problem is feasible for each scenario and X is optimal over all scenarios, since all the first stage decisions are made simultaneously.

Appart from the two-stage Stochastic Problem, some other alternative models are described. Wait-and-See model, WS, is an approximation based on the perfect information and it is equivalent to the SP relaxing anticipativity principle. It is also defined the Expected Value problem, EV, a simplification by the replace of all random variables with the expected value. If we fix the optimal solution of the first-stage in SP, the Expected result of using Expected Value, EEV is reached.

We have designed expressly a diet problem example in three different stochasticity cases in order to explain all the theorical models exposed and compare them. We have realized that it is not easy to get explanatory solutions with a simplified problem. In other words, the examples given in the literature books can contain painstakingly selected data.

According to the introduced models, two measures have been defined: the expected value of perfect information (EVPI) and the value of stochastic solutions (VSS). EVPI is the maximum price that the decision maker should pay in order to know the uncertainty in advance and it is defined as the difference between WS and SP, whereas VSS is the real cost of ignoring it and it is defined as the difference between EEV and SP. Based on these measures, we have concluded that sometimes is worthwhile to calculate the stochastic solution, but we cannot know it previously. Main inequalities of the measures have been verified. Proposition 3.1, $WS \leq SP \leq EEV$ has been checked with the three cases of the diet example, metioned before. The third case is the only one that fulfills the inequality of the Proposition 3.2, $EV \leq WS$, because recourse matrix and objective coefficients are fixed. Therefore, we can confirm that it is absolutely inadecuate to trust on the solution of the expected value problem, EV, since it is expected a better value than the one under perfect information, WS. We have also studied some relations between EVPI and VSS. As it can be seen in the definition, EVPI and VSS will never be negative, see Proposition 3.3. If we verify Proposition 3.4, $EVPI \leq EEV - EV$ and $VSS \leq EEV - EV$, for the three cases, as we expected, the third one is the unique that satisfies it.

After theory explanation, a real problem application has been considered. Motivated on the malnutrition, poverty and wasted food all around the world, Spain included, we realized that it would be interesting to apply maths in the third sector, also known as social economy. In particular, to optimize the food distribution in a local Food Bank, managed by a nonprofit organization. We have modeled a two-stage stochastic Diet Problem with 24, 48, 96 and 192 scenarios. A general model have been described with two different goals: (1) obtain the product distribution that minimizes the cost in order to satisfy entirely the nutrition requirement, with or without balanced diet and (2) improve the association standard nutritional bounds, without exceeding the initial budget and with or without balanced diet. Once data from the food storehouse was taken, we have decided which pair of products are nutritionally more different in order to create the scenario tree. In all the cases solutions provided by the expected value problem have been infeasible in some of the scenarios. Therefore they get the worst cost, that is, provided decisions are not implementable. This means that in all the cases it is really worthy to consider uncertainty. Since reporting 192 different decisions obtained for each model, SP, WS, EV and EEV would be too long, the best and the worst solutions of both models have been shown. After that, the food distributions depending on families have been summarized. The difference of the nutritional requirement satisfied, between their product sets and our results are remarkable. The optimal solutions provide few meal variation. After, the solutions of the four models have been compared. As it is well known in hostels, canteens or social restaurants, rice with legumes is a full nutritional meal. It is logical that the results depends on the goal. For a balanced diet, also a constraint with standard deviation bounded could be added, see [2], but the linear problem would change to a quadratic one. A varied set of products increase quite a lot the cost.

To conclude, we would like to remark that we just have given some suggestions thanks to this simplified application results analysis, considering some of the nutrients needed in a healthy diet. However, from now on, this application will be given over to specialists: nutricionists, health professional, social services and the people in charge of association, among others; obviously, in order to help in the decision making, since they have the final say.

Appendix A

GAMS CODE: Examples

The GAMS codes of examples in Chapters 2 and 3, for Cases 1, 2 and 3, can be downloaded from `https://ehubox.ehu.eus/index.php/s/8LP7rmcqvb5w1T3`.

The code of the program *examples.gms* obtains the solutions of the deterministic problem (2.2) and SP problem (2.4)-(2.7) in Chapter 2, and WS (3.2), EV (3.6) and EEV (3.11) models of the examples in Chapter 3. Moreover, EVPI (3.15) and VSS (3.16) measures are computed.

Appendix B

GAMS CODE: Application in the third sector

The GAMS codes of Model 1.a. for $|\Omega| = 192$ in Chapter 4 can be download from the link `https://ehubox.ehu.eus/index.php/s/hvljPEjbhibANAl`.

The code of the program *application.gms* obtains the solutions of the SP problem (4.1)-(4.6) in Chapter 4, and WS (4.8), EV (4.9) and EEV (4.10) models of the application of Model 1.a. for $|\Omega| = 192$ in Chapter 4. Moreover, EVPI and VSS measures are computed.

Bibliography

[1] A. Alonso-Ayuso, L.F. Escudero, A. Garín, M.T. Ortuo & G. Pérez. On the product selection and plant dimensioning under uncertainty. *Omega*, 33:307-318, 2005.

[2] Prof. Armbruster, F. Ferrero, E. Hsieh & A. Wagner. Diet Optimization Problem. *IEMS 310 Spring*, 2009.

[3] E. M. L. Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society*. Series B (*Methodological*), 173184, 1955.

[4] J. R. Birge. Stochastic Programming Computation and Applications. *Journal of Computing*, 9(2):111-133, 1997.

[5] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Science & Business Media, 2011.

[6] A. Carbajal Azcona. "Ingestas recomendadas de energía y nutrientes" en: M. T. García-Arias & M. C. García-Fernández, Eds., Nutrición y dietética, 2:27-44, 2003. Updated in 2013.

[7] Y. Censor & S. A. Zenios. *Parallel Optimization. Theory, Algorithms and Applications*. Oxford University Press, 1997.

[8] G. B. Dantzig. Linear programming under uncertainty. *Management science*, 1(3-4):197206, 1955.

[9] L. F. Escudero, A. Garín, M. Merino & G. Pérez. The Value of the Stochastic Solution in Multistage Problems. *Sociedad de Estadstica e Investigacin Operativa*, 15:48-64, 2007.

[10] A. Ferguson & G.B. Dantzig. The allocation of aircraft to routes: an example of linear programming under uncertain demands, *Management Science*, 3:45-73, 1956.

[11] Fundación Española de la Nutrición (FEN). Libro Blanco de la Nutrición en España, 2013.

[12] H.I. Gassman & W.T. Ziemba. *Stochastic Programming: applications in finance, energy, planning and logistics*. World Scientific, 2012.

[13] M. Gjerris & S. Gaiani. 'Values in the trash: ethical aspects of food waste' in: L. Escajedo San-Epifanio & M. De Renobales Scheifler, Eds., Envisioning a future without food waste and food poverty, 1:23-31, 2015.

[14] L. González Vaqué. 'Food loss and waste: some short- and medium-term proposal for the European Union' in: L. Escajedo San-Epifanio & M. De Renobales Scheifler, Eds., Envisioning a future without food waste and food poverty, 1:23-31, 2015.

[15] S. C. Graves & W. C. Jordan. Principles of the Benefits of Manufacturing Process Flexibility, *General Motors Laboratories Research Publication*, 1991.

[16] J. L. W. V. Jensen. Sur les fonctions convexes et les ingalits entre les valeurs moyennes. *Acta Mathematica*, 30(1):175193, 1906.

[17] P. Kall, A. Ruszczynski & K. Frauendorfer. Approximation techniques in stochastic programming. *Numerical techniques for stochastic optimization*, 3364, 1988.

[18] P. Kall & S.W. Wallace. *Stochastic Programming*. John Wiley, 1994.

[19] E. Kao & M. Queyranne. Budgeting costs of nursing in a hospital, *Management Science* 31:373-395, 1985.

[20] A. King. "Stochastic programming problems: Examples from the literature" in: Y.Ermoliev & R. West, Eds., *Numerical Techniques for Stochastic Optimization* (Springer-Verlag, Berlin), 543-567, 1988.

[21] A. J. King & S. W. Wallace. *Modeling with Stochastic Programming*. Springer Series in Operations Research and Financial Engineering, 2012.

[22] W.K. Klein Haneveld & M.H. van der Vlerk. Optimizing electricity distribution using two-stage integer recouse models. In S. Uryasev & P.M Pardalos, editors. *Stochastic Optimization: Algorithms and Applications*, Kluwer Academic Publisher, 137-154, 2001.

[23] A. Madansky. Inequalities for stochastic linear programming problems. *Management science*, 6(2):197-204, 1960.

[24] J. M. Mulvey, R. J. Vanderbei & S. A. Zenios. Robust Optimization of Large-Scale Systems, *Operations Research*, 2:264-281, 1995.

[25] A. Prekopa. *Stochastic Programming*. Kluwer Academic Publishers, Dordrecht, 1995.

[26] R. T. Rockafellar & R. J. B. Wets. Non-anticipativity and $\mathcal{L}'$-Martingales in Stochastic Optimization Problems, *Stochastic Systems: Modeling, Identification, and Optimization, Math. Programming Study*, 6:170-187, 1976.

[27] R. T. Rockafellar & R. J-B Wets. Scenario and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16:119-147, 1991.

[28] R. Schlaifer and H. Raiffa. *Applied statistical decision theory* (Harvard University, Boston, MA, 1961).

[29] A. Shapiro, D. Dentcheva & A. Ruszczyński. *Lectures on stochastic programming: Modeling and theory*. MPS-SIAM Book Series on Optimization 9, 2009.

[30] G. J. Stigler. The cost of the subsistence. *Journal of farm economics*, 27(2):303-314, 1945.

[31] W. Swart, C. Smith, & T. Holderby. "Expansion Planning for a Large Dairy Farm," Chapter 8 in *Studies in Linear Programming*, H. Salkin and J. Saha, eds. North Holland, Amsterdam, 1975.

[32] B. Vitoriano, J. Montero de Juan & D. Ruan. Decisions aid models for disaster management and emergency. *Atlantis Press*, 2013.

[33] S.W. Wallace & W.T. Ziemba, editors. *Applications of Stochastic Programming*: MPS-SIAM Book Series in Optimization, 2005.

[34] R. J-B Wets. Stochastic programs with fixed recourse: the equivalent deterministic problem. *SIAM Review*, 16:309-339, 1974 .

[35] NEOS Server. *Diet Problem*: `http://www.neos-guide.org/content/diet-problem`

[36] BEDCA - Base española de composición de alimentos: `http://www.bedca.net`

[37] Hirugarren sektoreko Sortarazi elkartearen web orria: `http://www.sortarazi.org/`

[38] Shopping in EROSKI online: `https://www.compraonline.grupoeroski.com/`

[39] Eustat - Euskal Estatistika Erakundea - Instituto Vasco de Estadística: `http://www.eustat.eus/`

[40] SPS - Stochastic Programming Society: `http://stoprog.org`

[41] GAMS modeling system for mathematical programming and optimization: `https://www.gams.com/`

[42] R-project software for statistical computing and graphics: `https://www.r-project.org/`

[43] IBM ILOG CPLEX Optimizer: `http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/`