From Scenarios to Conditional Scenarios in Two-Stage Stochastic MILP Problems

C. Beltran-Royo[a]

[a]*Computer Science, Statistics and Operations Research, King Juan Carlos University, Madrid, Spain*
E-mail: cesar.beltran@urjc.es

Abstract

The conditional scenario approach was introduced as an effective approximation to the two-stage stochastic mixed-integer linear programming problem. Although the original definition of conditional scenario is general, in practice it is basically suitable for the multivariate normal distribution. In this paper, we propose a new definition of conditional scenario that is suitable for approximating any multivariate distribution (continuous or discrete). This definition allows the approximation of a potentially large set of scenarios using a small set of conditional scenarios, unlike the previous definition. In the computational study, dedicated to solving the portfolio optimization problem with hard real-world constraints, the conditional scenario approach has clearly outperformed the sample average approximation approach in terms of solution time.

Keywords: Stochastic programming, MILP problem, scenario, conditional scenario, portfolio optimization.

1. Introduction

A common approach to model stochastic programming problems is based on scenarios (Chen et al., 2015). One finds this approach in different fields and applications such as air traffic flow management (Chang et al., 2016), portfolio optimization (Tong et al., 2009), forestry planning (Alonso-Ayuso et al., 2018), vehicle routing (Calvet et al., 2019), trading in electricity markets (Morales et al., 2009), e-commerce supply chains (Pages-Bernaus et al., 2019), etc. Given the high dimension and complexity of stochastic programming problems, it is common to use specialized approaches to solve them. This is even more necessary in presence of integer variables, which make the problem size a critical issue. In this case one can use approaches such as Lagrangian relaxation (Takriti and Birge, 2000), Benders decomposition (Santoso et al., 2005), decomposition with branch-and-cut (Sen and Sherali, 2006), parallel computing (Lubin et al., 2012), among others. As an alternative, one can use some heuristic approach in order to obtain good, but suboptimal, solutions. See, for example, (Amorim et al., 2015) for a path relinking method, (Karademir et al., 2014) for a greedy heuristic, (Raba et al., 2020) for a simheuristic, etc.

Scenario based optimization and deterministic optimization represent two extreme choices regarding the capability to model the uncertainty and the computational burden. The conditional scenario (CS) approach was introduced in (Beltran-Royo, 2017) as an effective midpoint between these two choices, since it showed a moderate computational effort and a reasonable capability to model the uncertainty, in the context of two-stage stochastic mixed-integer linear programming (SMILP).

In scenario based optimization one solves the so-called recourse problem, whereas in deterministic optimization one solves the so-called expected value problem (Birge and Louveaux, 2011). The CS problem improves the ability of the expected value problem to deal with uncertainty by considering conditional scenarios instead of the expected scenario. On the other hand, the CS problem reduces the computational burden of the recourse problem by considering a reduced number of conditional scenarios instead of a potentially large number of scenarios. In fact, the expected value problem is an approximation to the CS problem, which, in turn, is an approximation to the recourse problem. Therefore, the CS solution is, in general, suboptimal for the recourse problem but hopefully better than the expected value solution. For this reason, the CS approach should only be used in cases where the scenario approach results impractical regarding problem size or solution time (Figure 1).

In (Beltran-Royo, 2017), the CS approach was applied to approximate the SMILP problem formulated in terms of a multivariate normal random vector which accounted for all the uncertain parameters of the optimization problem. Although the definition of conditional scenario given in (Beltran-Royo, 2017) was general, in practice, it results suitable basically for the multivariate normal distribution. In order to improve the previous results, in this paper, we will try to answer the following questions: a) How could we make the definition of conditional scenario more useful for practitioners? b) How could we approximate any random vector (continuous or discrete) by conditional scenarios? c) How could we approximate a potentially large set of scenarios by a set of conditional scenarios?

To answer the previous questions we proceed as follows: In Section 2 we describe the notation and formulate the recourse problem. In Section 3 we review the key concepts and ideas that will be used in the other sections.In Section 4 a new definition of conditional scenario is introduced, which allows for the approximation of a set of scenarios by a set of conditional scenarios. The CS problem is formulated in Section 5. In Section 6 we compare, from a computational point of view, the CS approach versus the sample average approximation approach by solving the portfolio optimization problem with hard real-world constraints. Finally, Section 7 concludes the paper.

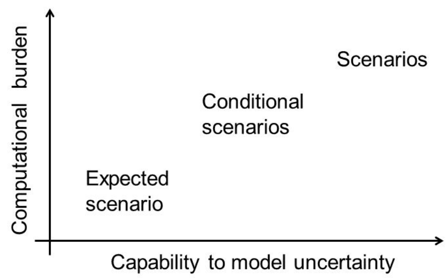

Fig. 1. The 'Conditional scenario' approach represents an effective midpoint between the deterministic optimization (based on the 'Expected scenario') and the 'Scenario' based optimization, regarding the capability to model the uncertainty and the computational burden.

2. Problem formulation

2.1. Notation

Indexes:

t	Decision stages	$t \in \mathcal{T} = \{1, 2\}$
j	Integer components of decision vectors	$j \in \mathcal{J}_t = \{1, \ldots, J_t\},\ t \in \mathcal{T}$
e	Realizations of discrete random variables	$e \in \mathcal{E} = \{1, \ldots, E\}$
r	Components of random vectors	$r \in \mathcal{R} = \{1, \ldots, R\}$
s	Scenarios	$s \in \mathcal{S} = \{1, \ldots, S\}$
re	Index pair for conditional scenarios	$re \in \mathcal{RE} = \mathcal{R} \times \mathcal{E}$

Random vectors:

$\tilde{\boldsymbol{\xi}}$	Random vector with finite support which accounts for all the random parameters of the recourse problem	
$\tilde{\xi}^s$	Scenario or realization of the random vector $\tilde{\boldsymbol{\xi}}$	$s \in \mathcal{S}$
$\tilde{p}^s$	Probability of $\tilde{\xi}^s$, that is, $\tilde{p}^s = P\big(\tilde{\boldsymbol{\xi}} = \tilde{\xi}^s\big)$	$s \in \mathcal{S}$
$\tilde{\boldsymbol{\xi}}_r$	Component r of the random vector $\tilde{\boldsymbol{\xi}}$	$r \in \mathcal{R}$
$\hat{\xi}^{re}$	Conditional scenario	$re \in \mathcal{RE}$
$\hat{p}^{re}$	Probability of $\hat{\xi}^{re}$	$re \in \mathcal{RE}$
$\hat{\boldsymbol{\xi}}^r$	Random vector defined by the conditional scenarios $\{\hat{\xi}^{re}\}_{re\in\mathcal{RE}}$ and the probability values $\{\hat{p}^{re}\}_{re\in\mathcal{RE}}$	$r \in \mathcal{R}$
$\bar{\xi}$	Expectation of $\tilde{\boldsymbol{\xi}}$, that is, $\bar{\xi} = \mathbb{E}[\tilde{\boldsymbol{\xi}}]$, the expected scenario.	

Notice that random vectors and random variables are indicated in boldface, scenarios are indicated by the tilde symbol, e.g. $\tilde{\xi}^s$, conditional scenarios by the hat symbol, e.g. $\hat{\xi}^{re}$, and expected values by the bar symbol, e.g. $\bar{\xi}$.

2.2. The recourse problem

In this paper we consider the following two-stage stochastic mixed-integer linear programming problem with recourse, which we call the recourse problem (RP) for short:

$$\min_{\tilde{x}} \quad z_{RP} = c_1^{\intercal}\tilde{x}_1 + \sum_{s\in\mathcal{S}} \tilde{p}^s\ \tilde{c}_2^{s\intercal}\tilde{x}_2^s \tag{1}$$

$$\text{s.t.} \quad A_1\tilde{x}_1 = b_1 \tag{2}$$

$$\tilde{A}_2^s\tilde{x}_1 + \tilde{B}_2^s\tilde{x}_2^s = \tilde{b}_2^s \qquad s\in\mathcal{S} \tag{3}$$

$$\tilde{x}_1 \geq 0 \tag{4}$$

$$\tilde{x}_2^s \geq 0 \qquad s\in\mathcal{S} \tag{5}$$

$$\tilde{x}_{1j} \text{ integer} \qquad j\in\mathcal{J}_1 \tag{6}$$

$$\tilde{x}_{2j}^s \text{ integer} \qquad s\in\mathcal{S},\ j\in\mathcal{J}_2, \tag{7}$$

where $\mathcal{J}_t$ is the index set for the integer variables at stage t for all $t\in\mathcal{T}$. In this context, scenario $\tilde{\xi}^s$ represents the sth realization of the random parameters, that is, $\tilde{\xi}^s = \text{vec}\big(\tilde{c}_2^s, \tilde{A}_2^s, \tilde{B}_2^s, \tilde{b}_2^s\big)$, where 'vec' is the operator that stacks vectors and matrix columns into a single vector. Then, $\tilde{\boldsymbol{\xi}}$ is the random vector defined by the set of scenarios $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$ and the corresponding probability values are $\{\tilde{p}^s\}_{s\in\mathcal{S}}$, such that $P\big(\tilde{\boldsymbol{\xi}} = \tilde{\xi}^s\big) = \tilde{p}^s$, for all $s\in\mathcal{S}$.

The RP problem (1)-(7) is one of the paradigms in optimization under uncertainty, also known as stochastic programming (Birge and Louveaux, 2011; Shapiro et al., 2009). The main feature of this problem is that it is a two-stage one. This means that the decision-maker takes some decision in the first stage, after which a random event occurs affecting the result of the first-stage decision. A *recourse decision* can then be made in the second stage by taking into account this result. The optimal solution from such a model is a single first-stage decision and a collection of recourse decisions defining which second-stage decision should be taken in response to each possible result.

More precisely, the elements of the first stage are: $\tilde{x}_1$, the decision to be taken before observing the realization of the random vector $\tilde{\boldsymbol{\xi}}$, and the deterministic parameters c_1, A_1 and b_1. On the other hand, the elements of the second stage are: the uncertain parameters $\tilde{c}_2^s, \tilde{A}_2^s, \tilde{B}_2^s, \tilde{b}_2^s$, and $\tilde{x}_2^s$, the recourse decision that will be taken after observing the realization of the random vector $\tilde{\boldsymbol{\xi}}$, that is, scenario $\tilde{\xi}^s$. For this reason there is a recourse decision $\tilde{x}_2^s$ for each possible future scenario $\tilde{\xi}^s$, for all $s\in\mathcal{S}$. Notice that some of the parameters considered uncertain in the second stage, could be deterministic, depending on the application.

2.3. The expected value problem

The RP problem can be approximated by the so-called expected value (EV) problem (Birge and Louveaux, 2011). More precisely, the EV problem corresponds to approximate $\tilde{\boldsymbol{\xi}}$ by $\bar{\xi} = \mathbb{E}[\tilde{\boldsymbol{\xi}}] =$

$\text{vec}\big(\,\bar{c}_2, \bar{A}_2, \bar{B}_2, \bar{b}_2\,\big)$ and can be formulated as follows:

$$
\begin{aligned}
\min_{\bar{x}} \quad & z_{EV} = c_1^{\mathsf{T}} \bar{x}_1 + \bar{c}_2^{\mathsf{T}} \bar{x}_2 \\
\text{s.t.} \quad & A_1 \bar{x}_1 = b_1 \\
& \bar{A}_2 \bar{x}_1 + \bar{B}_2 \bar{x}_2 = \bar{b}_2 \\
& \bar{x}_1 \geq 0 \\
& \bar{x}_2 \geq 0 \\
& \bar{x}_{tj} \text{ integer} \qquad\qquad\qquad\qquad t \in \mathcal{T}, j \in \mathcal{J}_t.
\end{aligned}
$$

The EV problem is the simplest approximation to the RP problem, given that all the parameter uncertainty is summarized by a single scenario, the expected scenario $\bar{\xi}$. The EV problem and the RP problem are optimization models of the so-called deterministic optimization and stochastic optimization, respectively. As already pointed out in the introduction, deterministic optimization and stochastic optimization represent two extreme choices regarding the capability to model the uncertainty and the computational burden. In this context, the conditional scenario problem, formulated in Section 5, represents an effective midpoint between these two choices, since it shows a moderate computational effort and a reasonable capability to model the uncertainty.

3. Conditional scenarios

In this section we review some key concepts and ideas that will be used in the other sections, such as the conditional expectation of a random vector and the definition of conditional scenario introduced in (Beltran-Royo, 2017). The readers familiar with these concepts can skip this section.

3.1. Conditional expectation of a random vector

Given a random vector $\boldsymbol{\xi} = \big(\boldsymbol{\xi}_1 \ldots\ \boldsymbol{\xi}_R\big)^{\mathsf{T}}$, its expectation $\mathbb{E}[\,\boldsymbol{\xi}\,]$ corresponds to the long-run average vector of repetitions of the experiment $\boldsymbol{\xi}$ represents. In the case of having dependent components in $\boldsymbol{\xi}$, it can be advantageous to use the possible knowledge of one component to compute the expectation of the whole random vector. For example, if we observe that $\boldsymbol{\xi}_1 = 100$, the *conditional expectation* of $\boldsymbol{\xi}$ given $\boldsymbol{\xi}_1 = 100$, that is, $\mathbb{E}[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_1 = 100\,]$, will give us a better picture of the long-run average vector of repetitions of the experiment $\boldsymbol{\xi}$ represents, than $\mathbb{E}[\,\boldsymbol{\xi}\,]$. In general we will write $\mathbb{E}[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r = \xi_r\,]$ or for short, $\mathbb{E}[\,\boldsymbol{\xi} \mid \xi_r\,]$ for any $r \in \mathcal{R} = \{1, \ldots, R\}$. For details see (DeGroot and Schervish, 2012). In this subsection we restrict ourselves to the *multivariate normal distribution* (Hardle and Simar, 2012), which is characterized by a random vector $\boldsymbol{\xi}$ with mean μ and covariance matrix $\Sigma \succ 0$ (positive definite). In this case, we write $\boldsymbol{\xi} \sim N_R(\mu, \Sigma)$.

Proposition 1. *(Hardle and Simar, 2012) Let us consider a multivariate normal random vector* $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \dots \boldsymbol{\xi}_R \end{pmatrix}^\intercal \sim N_R(\mu, \Sigma)$*. Then the conditional expectation of* $\boldsymbol{\xi}$ *given* ξ_r *can be calculated as follows:*

$$\mathbb{E}[\, \boldsymbol{\xi} \mid \xi_r \,] = \mu + \frac{\xi_r - \mu_r}{\sigma_r^2}\, \Sigma_{*r} \qquad r \in \mathcal{R}, \tag{8}$$

where Σ_{*r} *is the* r*th column of the covariance matrix* Σ*.*

Instead of only considering an observation ξ_r and computing $\mathbb{E}[\, \boldsymbol{\xi} \mid \xi_r \,]$, we can consider the random variable $\boldsymbol{\xi}_r$ itself and compute $\mathbb{E}[\, \boldsymbol{\xi} \mid \boldsymbol{\xi}_r \,]$ that we denote by $\boldsymbol{\xi}^r$ for all $r \in \mathcal{R}$ and call it the rth conditional expectation of $\boldsymbol{\xi}$. In the multivariate normal case $\mathbb{E}[\, \boldsymbol{\xi} \mid \boldsymbol{\xi}_r \,]$ can be computed by using formula (8).

3.2. Conditional scenarios for a continuous distribution

In this section we review the definition of conditional scenario for a continuous distribution given in (Beltran-Royo, 2017).

Definition 1. *(Conditional scenario) Given a continuous random vector* $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \dots \boldsymbol{\xi}_R \end{pmatrix}^\intercal$ *and any* $r \in \mathcal{R}$*, let us consider the random variable* $\boldsymbol{\xi}_r$ *and a discretization of it, say* $\tilde{\boldsymbol{\xi}}_r$*, with finite support* $\{\tilde{\xi}_{re}\}_{e \in \mathcal{E}}$ *and the corresponding probability values* $\{\tilde{p}_{re}\}_{e \in \mathcal{E}}$*, with* $\mathcal{E} = \{1, \dots, E\}$ *for a given* E*. Then, the conditional scenarios of* $\boldsymbol{\xi}$*, with their probability values, corresponding to the above discretization of* $\boldsymbol{\xi}_r$ *are defined as follows:*

$$\begin{aligned} &\textit{conditional scenarios:} \quad \hat{\xi}^{re} = \mathbb{E}[\, \boldsymbol{\xi} \;\mid\; \boldsymbol{\xi}_r = \tilde{\xi}_{re} \,] && e \in \mathcal{E}, \\ &\textit{probability values:} \quad \hat{p}^{re} = \tilde{p}_{re} && e \in \mathcal{E}, \end{aligned} \tag{9}$$

where the right-hand side in (9) is the conditional expectation of $\boldsymbol{\xi}$ *given* $\boldsymbol{\xi}_r$ *is equal to* $\tilde{\xi}_{re}$*. Furthermore,* $\hat{\boldsymbol{\xi}}^r$ *is the finite support random vector defined by the conditional scenarios* $\{\hat{\xi}^{re}\}_{re \in \mathcal{E}}$ *and the corresponding probability values* $\{\hat{p}^{re}\}_{re \in \mathcal{E}}$*, such that* $P\big(\, \hat{\boldsymbol{\xi}}^r = \hat{\xi}^{re} \,\big) = \hat{p}^{re}$*.*

We illustrate the previous definition with the following example.

Example 1. *(Conditional scenarios of a multivariate normal random vector)* Let us consider the multivariate normal random vector $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 & \boldsymbol{\xi}_2 \end{pmatrix}^\intercal \sim N_2(\mu, \Sigma)$ such that

$$\mu = \begin{pmatrix} 100 & 200 \end{pmatrix}^\intercal \qquad \Sigma = \begin{pmatrix} 400 & 480 \\ 480 & 1600 \end{pmatrix}.$$

Let us approximate $\boldsymbol{\xi}$ by six conditional scenarios based on a discretization of $\boldsymbol{\xi}_1$. First, one computes the corresponding conditional expectation by using (8):

$$\boldsymbol{\xi}^1 = \mathbb{E}[\, \boldsymbol{\xi} \mid \boldsymbol{\xi}_1 \,] = \begin{pmatrix} \boldsymbol{\xi}_1 \\ 80 + 1.2\boldsymbol{\xi}_1 \end{pmatrix}, \tag{10}$$

where $\boldsymbol{\xi}_1 \sim N(\mu_1,\ \sigma_1)$ with $\mu_1 = 100$ and $\sigma_1 = 20$. Second, one discretizes the random variable $\boldsymbol{\xi}_1$ into six representative points (see (Beltran-Royo, 2017) for details):

$$
\begin{array}{ll}
\tilde{\xi}_{1,1} = 53.7 & \tilde{p}_{1,1} = 0.0214 \\
\tilde{\xi}_{1,2} = 72.3 & \tilde{p}_{1,2} = 0.1363 \\
\tilde{\xi}_{1,3} = 90.8 & \tilde{p}_{1,3} = 0.3423 \\
\tilde{\xi}_{1,4} = 109.2 & \tilde{p}_{1,4} = 0.3423 \\
\tilde{\xi}_{1,5} = 127.7 & \tilde{p}_{1,5} = 0.1363 \\
\tilde{\xi}_{1,6} = 146.3 & \tilde{p}_{1,6} = 0.0214.
\end{array}
$$

Third, one computes the conditional scenarios combining the previous discretization of $\boldsymbol{\xi}_1$ with (10). For example the first conditional scenario $\hat{\xi}^{1,1}$ can be computed as follows:

$$
\hat{\xi}^{1,1} = \mathbb{E}[\,\boldsymbol{\xi} \mid \tilde{\xi}_{1,1}\,] = \begin{pmatrix} \tilde{\xi}_{1,1} \\ 80 + 1.2\tilde{\xi}_{1,1} \end{pmatrix} = \begin{pmatrix} 53.7 \\ 80 + 1.2 \cdot 53.7 \end{pmatrix} = \begin{pmatrix} 53.7 \\ 144.4 \end{pmatrix}
$$
$$
\hat{p}^{1,1} = \tilde{p}_{1,1} = 0.0214.
$$

In summary one obtains:

$$
\begin{array}{ll}
\hat{\xi}^{1,1} = \begin{pmatrix} 53.7 & 144.4 \end{pmatrix}^\intercal & \hat{p}^{1,1} = 0.0214 \\
\hat{\xi}^{1,2} = \begin{pmatrix} 72.3 & 166.8 \end{pmatrix}^\intercal & \hat{p}^{1,2} = 0.1363 \\
\hat{\xi}^{1,3} = \begin{pmatrix} 90.8 & 189.0 \end{pmatrix}^\intercal & \hat{p}^{1,3} = 0.3423 \\
\hat{\xi}^{1,4} = \begin{pmatrix} 109.2 & 211.0 \end{pmatrix}^\intercal & \hat{p}^{1,4} = 0.3423 \\
\hat{\xi}^{1,5} = \begin{pmatrix} 127.7 & 233.2 \end{pmatrix}^\intercal & \hat{p}^{1,5} = 0.1363 \\
\hat{\xi}^{1,6} = \begin{pmatrix} 146.3 & 255.6 \end{pmatrix}^\intercal & \hat{p}^{1,6} = 0.0214.
\end{array}
$$

The approximation of $\boldsymbol{\xi}$ by six conditional scenarios derived from the 2nd conditional expectation $\boldsymbol{\xi}^2$ could be done analogously. In total we would obtain twelve conditional scenarios that are represented in Figure 2. As we will see in the Section 5, the conditional scenario problem is formulated in terms of conditional scenarios in contrast with the expected value problem, which is formulated in terms of the expected scenario μ (the square in in the center of Figure 2). Further details can be found in (Beltran-Royo, 2017). ■

4. Conditional scenarios for a general distribution

In this section we analyze and improve the definition of conditional scenario introduced in (Beltran-Royo, 2017). Furthermore, we show how the new definition can be used to approximate a potentially large set of scenarios by a reduced set of conditional scenarios.

As pointed out in (Beltran-Royo, 2017), the two-stage stochastic mixed-integer linear programming problem formulated in terms of a continuous random vector, say $\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{\xi}_1 \dots \ \boldsymbol{\xi}_R \end{pmatrix}^\intercal$, is in general numerically intractable. To address this difficulty one can approximate $\boldsymbol{\xi}$ by a random vector, say $\tilde{\boldsymbol{\xi}} = \begin{pmatrix} \tilde{\boldsymbol{\xi}}_1 \dots \ \tilde{\boldsymbol{\xi}}_R \end{pmatrix}^\intercal$, with a discrete support consisting of S scenarios, and solve the corresponding

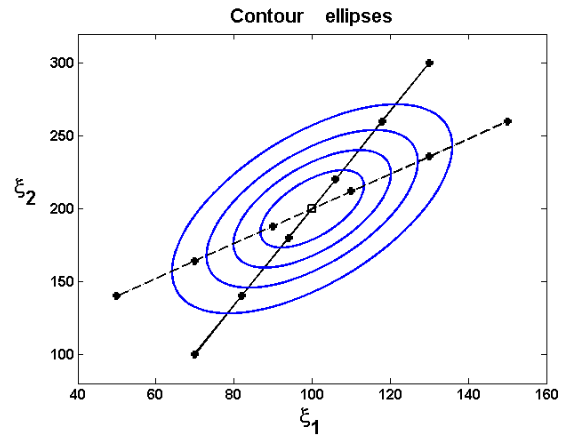

Fig. 2. Equiprobability contour ellipses of the multivariate normal random vector $\boldsymbol{\xi}$ of Example 1. The dashed and the solid lines correspond to the first ($\boldsymbol{\xi}^1$) and the second ($\boldsymbol{\xi}^2$) conditional expectations, respectively. The dots represent the twelve conditional scenarios of Example 1 and the square represents the 'expected scenario' μ.

RP problem (1)-(7). Some technique to compute a representative set of scenarios is normally used: moment matching methods (Hoyland and Wallace, 2001; Hoyland et al., 2003; King and Wallace, 2012), the Sample Average Approximation (SAA) method (Homem-de-Mello and Bayraksan, 2014; Kleywegt et al., 2002; Shapiro et al., 2009), approaches based on probability metrics (Dupacova et al., 2003; Heitsch and Romisch, 2003; Pflug and Pichler, 2014), among others. If the computational effort of using $\tilde{\boldsymbol{\xi}}$ is too high, one can approximate it by conditional scenarios.

Let us analyze Definition 1, which corresponds to the definition of conditional scenario given in (Beltran-Royo, 2017), and point out its drawbacks. According to Definition 1, the set of conditional scenarios discards much of the multivariate structure of the original probability distribution of $\boldsymbol{\xi}$. However, the basic uncertainty model based on conditional scenarios can result effective in cases where the uncertainty model based on scenarios results computationally infeasible. See (Beltran-Royo, 2017), where Definition 1 was successfully used within the procedure to approximate a multivariate normal random vector by means of conditional scenarios. This is an important case since the multivariate normal distribution is one of the most relevant ones mainly because of the multivariate central limit theorem (Jacod and Protter, 2004). The procedure introduced in (Beltran-Royo, 2017) requires to compute the conditional expectation (9), which in the multivariate normal case it is easily done by means of the corresponding analytical expression (see Proposition 1). Thus, if our objective is to approximate a continuous random vector by means of conditional scenarios, the first drawback of the procedure based on Definition 1 is that it requires the analytical expression of the conditional expectation (9), which is not known or may be difficult to compute for a general multivariate distribution (the multivariate normal distribution is an exception).

Notice that Definition 1 could easily be adapted for a discrete random vector, say $\tilde{\boldsymbol{\xi}}$, with a finite support given by a set of scenarios $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$. However, in this case appears the se second drawback since given a particular value of $\tilde{\boldsymbol{\xi}}_r$, say $\tilde{\xi}_{re_0}$, it is likely that only one scenario, say $\tilde{\xi}^{s_0}$, would attain this value in the rth component and therefore

$$\hat{\xi}^{re_0} = \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \;\mid\; \tilde{\boldsymbol{\xi}}_r = \tilde{\xi}_{re_0}\,] = \tilde{\xi}^{s_0}.$$

Thus, in this case, it is likely that the set of conditional scenarios would correspond to the original set of scenarios. In summary, taking into account the previous two drawbacks one can conclude that, although Definition 1 is a general one, in practice, it is suitable basically for the multivariate normal distribution.

Below we give a new definition of conditional scenario that overcomes the drawbacks of Definition 1. Furthermore, the new definition is equivalent to the previous one for the multivariate normal distribution (see Proposition 4). To this end we need the following assumption:

Assumption 1. *Given a random vector $\boldsymbol{\xi} = \big(\boldsymbol{\xi}_1 \dots \boldsymbol{\xi}_R\big)^{\intercal}$ (continuous or discrete):*

a) The interval $\mathcal{I}_r = [a_r, b_r] \subset \mathbb{R}$, with $-\infty \le a_r < b_r \le +\infty$, is any interval that contains the support of $\boldsymbol{\xi}_r$, for all $r \in \mathcal{R} = \{1, \dots, R\}$.

b) $\mathcal{I}_r$ is partitioned into finitely many subintervals $\mathcal{I}_{re}$ such that $\mathcal{I}_r = \cup_{e\in\mathcal{E}}\mathcal{I}_{re}$, where $\mathcal{I}_{re} = [a_{re}, b_{re})$, for all $e \in \mathcal{E} \setminus \{E\}$ and $\mathcal{I}_{rE} = [a_{rE}, b_{rE}]$, that is, all the subintervals are right-open except the last one.

The new definition of conditional scenarios is as follows (in this paper the term 'conditional scenario' will refer to this definition):

Definition 2. *(Conditional scenario) Let us consider the random vector $\boldsymbol{\xi} = \big(\boldsymbol{\xi}_1 \dots \boldsymbol{\xi}_R\big)^{\intercal}$ (continuous or discrete). Given any $r \in \mathcal{R}$, let us also consider an interval $\mathcal{I}_r$ and a partition of it $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$ that fulfill Assumption 1. Then, the conditional scenarios of $\boldsymbol{\xi}$ corresponding to the above partition of the support of $\boldsymbol{\xi}_r$, are defined as follows:*

$$\begin{aligned} \hat{\xi}^{re} &:= \mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \;\mid\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big] && e \in \mathcal{E} \\ \hat{p}^{re} &:= P\big(\,\boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big) && e \in \mathcal{E}, \end{aligned} \tag{11}$$

where the right-hand side in (11) is the conditional expectation of $\boldsymbol{\xi}$ given $\boldsymbol{\xi}_r \in \mathcal{I}_{re}$.

In this context, we can define the CS random vector $\hat{\boldsymbol{\xi}}^r = \big(\hat{\boldsymbol{\xi}}^r_1 \dots \hat{\boldsymbol{\xi}}^r_R\big)^{\intercal}$ such that its support is $\{\hat{\xi}^{re}\}_{e\in\mathcal{E}}$ and such that $P\big(\,\hat{\boldsymbol{\xi}}^r = \hat{\xi}^{re}\,\big) = \hat{p}^{re}$ for all $e \in \mathcal{E}$. As proposed in (Beltran-Royo, 2017), one can approximate a given random vector $\boldsymbol{\xi}$ by the CS random vectors $\hat{\boldsymbol{\xi}}^r$, for all $r \in \mathcal{R}$, each one taken with probability $1/R$, and then formulate the CS problem in terms of these random vectors (this problem will be defined in Section 5). Let us see some properties regarding conditional expectation and conditional scenarios.

Proposition 2. *Let us consider the random vector $\boldsymbol{\xi} = \big(\boldsymbol{\xi}_1 \dots \boldsymbol{\xi}_R\big)^{\intercal}$ (continuous or discrete). Given any $r \in \mathcal{R}$, let us also consider an interval $\mathcal{I}_r$ and a partition of it $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$ that fulfill Assumption 1. In this case one has that*

$$\mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \;\mid\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big] = \mathbb{E}_{\boldsymbol{\xi}_r}\Big[\,\mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,\big] \;\mid\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\Big] \qquad re \in \mathcal{RE},$$

where, on the one hand, it is assumed that the expectation on the left-hand side of the equation is finite and, on the other hand, $\mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,\big]$ is the conditional expectation of $\boldsymbol{\xi}$ given $\boldsymbol{\xi}_r \in \mathcal{I}_r$.

Notice that the previous result is a generalization of the law of total probability for expectations ((DeGroot and Schervish, 2012), Theorem 4.7.1) and it can be proved in a similar way.

Proposition 3. *If* $L(\boldsymbol{\xi}_r) = \mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,\big]$ *is a linear function, then*

$$\mathbb{E}_{\boldsymbol{\xi}_r}\Big[\,\mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,\big] \;\big|\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\Big] = \mathbb{E}_{\boldsymbol{\xi}}\Big[\,\boldsymbol{\xi} \;\big|\; \mathbb{E}_{\boldsymbol{\xi}_r}\big[\,\boldsymbol{\xi}_r \mid \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big]\Big] \qquad re \in \mathcal{RE}.$$

In this case, the conditional scenarios can be computed as follows:

$$\hat{\xi}^{re} = \mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \;\big|\; \bar{\xi}_{re}\,\big] \qquad re \in \mathcal{RE},$$

where $\bar{\xi}_{re} = \mathbb{E}_{\boldsymbol{\xi}_r}\big[\,\boldsymbol{\xi}_r \mid \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big]$.

Proof: The first statement is clear considering that the expectation is a linear operator, i.e., for any linear function $L(\boldsymbol{\xi}_r)$ one has $\mathbb{E}[\,L(\boldsymbol{\xi}_r)\,] = L(\mathbb{E}[\,\boldsymbol{\xi}_r\,])$. The second statement is also straightforward:

$$\begin{aligned}
\hat{\xi}^{re} &= \mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \;\big|\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big] && re \in \mathcal{RE} \\
&= \mathbb{E}_{\boldsymbol{\xi}_r}\Big[\,\mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \mid \boldsymbol{\xi}_r\,\big] \;\big|\; \boldsymbol{\xi}_r \in \mathcal{I}_{re}\Big] && \text{(by Proposition 2)} \\
&= \mathbb{E}_{\boldsymbol{\xi}}\Big[\,\boldsymbol{\xi} \;\big|\; \mathbb{E}_{\boldsymbol{\xi}_r}\big[\,\boldsymbol{\xi}_r \mid \boldsymbol{\xi}_r \in \mathcal{I}_{re}\,\big]\Big] && \text{(by the first statement)} \\
&= \mathbb{E}_{\boldsymbol{\xi}}\big[\,\boldsymbol{\xi} \;\big|\; \bar{\xi}_{re}\,\big].
\end{aligned}$$

■

Proposition 4. *(Conditional scenarios of a multivariate normal random vector) Let us consider the multivariate normal random vector* $\boldsymbol{\xi} \sim N_R(\mu, \Sigma)$. *Let us also consider, for each* $r \in \mathcal{R}$, *an interval* $\mathcal{I}_r$ *and a partition of it, say* $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$, *that fulfill Assumption 1. Then, the conditional scenarios of* $\boldsymbol{\xi}$ *corresponding to the above partitions can be computed as follows:*

$$\hat{\xi}^{re} = \mu + \frac{\bar{\xi}_{re} - \mu_r}{\sigma_r^2}\,\Sigma_{*r} \qquad re \in \mathcal{RE},$$

where

$$\bar{\xi}_{re} = \mu_r + \sigma_r\,\frac{\phi(\alpha_{re}) - \phi(\beta_{re})}{\Phi(\beta_{re}) - \Phi(\alpha_{re})} \qquad re \in \mathcal{RE} \tag{12}$$

$$\alpha_{re} = (a_{re} - \mu_r)/\sigma_r \;\; \text{and} \;\; \beta_{re} = (b_{re} - \mu_r)/\sigma_r, \tag{13}$$

scalars μ_r *and* σ_r^2 *are the mean and variance, respectively, of the random variable* $\boldsymbol{\xi}_r$, *functions* ϕ *and* Φ *are the density and distribution function, respectively, of the standard normal distribution and vector* Σ_{*r} *is the* r*-th column of the covariance matrix* Σ. *The corresponding probability values can be computed as follows:*

$$\hat{p}^{re} = \Phi(\beta_{re}) - \Phi(\alpha_{re}) \qquad re \in \mathcal{RE}.$$

Proof: In the case of a multivariate normal random vector, the conditional expectation $\mathbb{E}_{\boldsymbol{\xi}}\big[\, \boldsymbol{\xi} \mid \boldsymbol{\xi}_r \,\big]$ can be computed as follows (Hardle and Simar, 2012):

$$L(\boldsymbol{\xi}_r) = \mathbb{E}_{\boldsymbol{\xi}}\big[\, \boldsymbol{\xi} \mid \boldsymbol{\xi}_r \,\big] = \mu + \frac{\boldsymbol{\xi}_r - \mu_r}{\sigma_r^2}\, \Sigma_{*r} \qquad\qquad r \in \mathcal{R},$$

which is linear in $\boldsymbol{\xi}_r$. Thus, by Proposition 3, one has

$$\hat{\xi}^{re} = \mathbb{E}_{\boldsymbol{\xi}}\big[\, \boldsymbol{\xi} \mid \bar{\xi}_{re} \,\big] = \mu + \frac{\bar{\xi}_{re} - \mu_r}{\sigma_r^2}\, \Sigma_{*r},$$

where $\bar{\xi}_{re} = \mathbb{E}_{\boldsymbol{\xi}_r}\big[\, \boldsymbol{\xi}_r \mid \boldsymbol{\xi}_r \in \mathcal{I}_{re} \,\big]$. This expression corresponds to the expectation of a truncated normal distribution which can be computed by equations (12)–(13) (Johnson et al., 1994). Finally,

$$\hat{p}^{re} = P\big(\, \boldsymbol{\xi}_r \in \mathcal{I}_{re} \,\big) = \Phi(\beta_{re}) - \Phi(\alpha_{re}) \qquad re \in \mathcal{RE}. \qquad \blacksquare$$

Notice that this way to compute the conditional scenarios for the multivariate normal case is equivalent to Method 2 in (Beltran-Royo, 2017). This is true for any set of partitions of the intervals $\mathcal{I}_1, \ldots, \mathcal{I}_R$, provided that both Proposition 4 and Method 2 use the same set of partitions. Otherwise said, the new definition of conditional scenario given in this paper and the definition given in (Beltran-Royo, 2017) are equivalent for the multivariate normal case.

Proposition 5. *(Conditional scenarios of a finite support random vector) Let us consider a random vector $\tilde{\boldsymbol{\xi}}$ with finite support $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$ and probability mass function $f(\tilde{\xi}^s) = \tilde{p}^s$ for all $s \in \mathcal{S}$. Let us also consider, for each $r \in \mathcal{R}$, an interval $\mathcal{I}_r$ and a partition of it, say $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$, that fulfill Assumption 1, then the conditional scenarios of $\tilde{\boldsymbol{\xi}}$ corresponding to the above partitions can be computed as follows:*

$$\hat{\xi}^{re} = \sum_{s\in\mathcal{S}_{re}} \frac{\tilde{p}^s}{\sum_{s\in\mathcal{S}_{re}} \tilde{p}^s}\, \tilde{\xi}^s \qquad\qquad re \in \mathcal{RE},$$

where $\mathcal{S}_{re} = \{s \mid \tilde{\xi}^s_r \in \mathcal{I}_{re}\}$. Furthermore, the corresponding probability values can be computed as follows:

$$\hat{p}^{re} = \sum_{s\in\mathcal{S}_{re}} \tilde{p}^s \qquad\qquad re \in \mathcal{RE}.$$

Proof: First, the probability mass function of $\tilde{\boldsymbol{\xi}}$ given $\tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}$ corresponds to

$$f(\tilde{\xi}^s \mid\ \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}) = \frac{f(\tilde{\xi}^s)}{\sum_{s\in\mathcal{S}_{re}} f(\tilde{\xi}^s)} = \frac{\tilde{p}^s}{\sum_{s\in\mathcal{S}_{re}} \tilde{p}^s} \qquad\qquad s \in \mathcal{S}_{re}$$

and $f(\tilde{\xi}^s \mid\ \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}) = 0$ for all $s \notin \mathcal{S}_{re}$.

Table 1
Each column corresponds to a scenario.

Scenario number	1	2	3	4	5	6	7	8	9	10
$\tilde{\xi}_1^s$	0.0	3.2	1.3	1.5	1.9	1.6	2.5	2.7	0.6	4.0
$\tilde{\xi}_2^s$	1.0	6.7	2.1	3.9	5.5	4.5	6.1	7.9	3.3	9.0
$\tilde{\xi}_3^s$	4.0	13.9	7.4	8.3	9.6	10.5	11.1	12.5	5.8	16.0

Second, by definition of conditional scenario

$$\begin{aligned}\hat{\xi}^{re} &= \mathbb{E}_{\tilde{\boldsymbol{\xi}}}\big[\,\tilde{\boldsymbol{\xi}} \;\mid\; \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}\,\big] && re \in \mathcal{RE}\\ &= \sum_{s\in\mathcal{S}_{re}} \tilde{\xi}^s\, f(\tilde{\xi}^s \mid \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re})\\ &= \sum_{s\in\mathcal{S}_{re}} \frac{\tilde{p}^s}{\sum_{s\in\mathcal{S}_{re}} \tilde{p}^s}\, \tilde{\xi}^s.\end{aligned}$$

Finally,

$$\hat{p}^{re} = P\big(\boldsymbol{\xi}_r \in \mathcal{I}_{re}\big) = \sum_{s\in\mathcal{S}_{re}} \tilde{p}^s \qquad re \in \mathcal{RE}.$$

■

4.1. *From scenarios to conditional scenarios*

In this section we show how Proposition 5 can be used to approximate a set of scenarios by a set of conditional scenarios. As a first step in this direction, let us illustrate Proposition 5 by means of the following example.

Example 2. *(From scenarios to conditional scenarios: Calculations)*

A company has forecast the next season demand of three products in order to plan the corresponding production. This demand forecast can be represented by means of ten equiprobable scenarios that can be found in Table 1. We use vector $\tilde{\xi}^s$ to denote scenario s, where component $\tilde{\xi}^s_r$ represents the demand of the rth product according to scenario s for all $r \in \mathcal{R} = \{1, 2, 3\}$ and for all $s \in \mathcal{S} = \{1, \dots, 10\}$. Thus, the uncertain demand can be modeled by means of the random vector of demands $\tilde{\boldsymbol{\xi}}$ with support $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$ and probability values $\{\tilde{p}^s\}_{s\in\mathcal{S}}$, where $\tilde{p}^s = 1/10$ for all $s \in \mathcal{S}$.

The simplest approximation of $\tilde{\boldsymbol{\xi}}$ is the expected scenario, that is $\bar{\xi} = \begin{pmatrix} 1.9 & 5.0 & 9.9\end{pmatrix}^\mathsf{T}$, which corresponds to the average of the previous ten scenarios. A more informative approximations can be obtained by means of conditional scenarios. For example, one can compute a set of two conditional scenarios considering low and high forecast demands of Product 1. We will use vectors $\hat{\xi}^{1,1}$ and $\hat{\xi}^{1,2}$ to denote the conditional scenarios associated to low and high forecast demands, respectively, of Product 1. To compute the two conditional scenarios, first, we observe in Table 1 that the forecast demands of Product

Table 2
The scenarios have been classified according to the demand of the first product into two groups: low and high demand.

	Low demand (group 1)						High demand (group 2)			
Scenario number	1	9	3	4	6	5	7	8	2	10
$\tilde{\xi}_1^s$	*0.0*	*0.6*	*1.3*	*1.5*	*1.6*	*1.9*	*2.5*	*2.7*	*3.2*	*4.0*
$\tilde{\xi}_2^s$	1.0	3.3	2.1	3.9	4.5	5.5	6.1	7.9	6.7	9.0
$\tilde{\xi}_3^s$	4.0	5.8	7.4	8.3	10.5	9.6	11.1	12.5	13.9	16.0

1 are in the interval $\mathcal{I}_1 = [0.0,\ 4.0]$. Second, we partition $\mathcal{I}_1$ into two subintervals:

$$
\begin{aligned}
\mathcal{I}_{1,1} &= [0.0,\ 2.0) \qquad \text{(low demand)}\\
\mathcal{I}_{1,2} &= [2.0,\ 4.0] \qquad \text{(high demand),}
\end{aligned}
$$

and group the scenarios according to this partition of the Product 1 demand in Table 2. Notice that in Table 2 entries in the row corresponding to $\tilde{\xi}_1^s$ are in italic font and sorted in ascending order, such that in the first group $\tilde{\xi}_1^s$ is in $\mathcal{I}_{1,1}$ and in the second group $\tilde{\xi}_1^s$ is in $\mathcal{I}_{1,2}$. The number of scenarios is six and four in the first and second groups, respectively.

The conditional scenario $\hat{\xi}^{1,1}$ corresponds to the conditional expectation of $\tilde{\boldsymbol{\xi}}$ given $\tilde{\boldsymbol{\xi}}_1 \in [0,\ 2)$, that is,

$$
\begin{aligned}
\hat{\xi}^{1,1} &= \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_1 \in \mathcal{I}_{1,1}\,]\\
&= \sum_{s\in\mathcal{S}_{1,1}} \frac{\tilde{p}^s}{\sum_{s\in\mathcal{S}_{1,1}} \tilde{p}^s}\,\tilde{\xi}^s = \sum_{s\in\mathcal{S}_{1,1}} \frac{1/10}{6/10}\,\tilde{\xi}^s \qquad\qquad \text{(by Proposition 5)}\\
&= \frac{1}{6}\begin{pmatrix}0.0\\1.0\\4.0\end{pmatrix} + \frac{1}{6}\begin{pmatrix}0.6\\3.3\\5.8\end{pmatrix} + \cdots + \frac{1}{6}\begin{pmatrix}1.9\\5.5\\9.6\end{pmatrix} = \begin{pmatrix}1.2\\3.4\\7.6\end{pmatrix}.
\end{aligned}
$$

Analogously, the conditional scenario $\hat{\xi}^{1,2}$ can be computed as the average of the four scenarios in the second group:

$$
\hat{\xi}^{1,2} = \frac{1}{4}\begin{pmatrix}2.5\\6.1\\11.1\end{pmatrix} + \frac{1}{4}\begin{pmatrix}2.7\\7.9\\12.5\end{pmatrix} + \cdots + \frac{1}{4}\begin{pmatrix}4.0\\9.0\\16.0\end{pmatrix} = \begin{pmatrix}3.1\\7.4\\13.4\end{pmatrix}.
$$

From these conditional scenarios we can define the random vector $\hat{\boldsymbol{\xi}}^1$ with support $\{\hat{\xi}^{1,e}\}_{e\in\mathcal{E}}$ and the corresponding probability values $\hat{p}^{1,1} = 6/10$ and $\hat{p}^{1,2} = 4/10$. Notice that arbitrarily we have only computed two conditional scenarios but more conditional scenarios could have been computed by considering more intervals for Product 1 demand. Furthermore, to compute the conditional scenarios we have focused on Product 1 demand. Analogously we could focus on the demands of Products 2 and 3 and construct the corresponding conditional scenarios whose associated random vectors would be $\hat{\boldsymbol{\xi}}^2$ and

$\hat{\boldsymbol{\xi}}^3$, respectively.

In summary, we observe that with the set of ten scenarios $\tilde{\xi}^s$ with probability $1/10$ each, we have a probabilistic model which can be used to forecast the future demand. This probabilistic model can be approximated by the single expected scenario $\bar{\xi} = \begin{pmatrix} 2.0 & 5.0 & 9.9 \end{pmatrix}^\intercal$. But, it can also be approximated by two conditional scenarios $\hat{\xi}^{1,1} = \begin{pmatrix} 1.2 & 3.4 & 7.6 \end{pmatrix}^\intercal$ and $\hat{\xi}^{1,2} = \begin{pmatrix} 3.1 & 7.4 & 13.4 \end{pmatrix}^\intercal$ with probability values 6/10 and 4/10, respectively, which define the CS random vector $\hat{\boldsymbol{\xi}}^1$. Conditional scenarios $\hat{\xi}^{1,1}$ and $\hat{\xi}^{1,2}$ represent the demands for the three products that can be expected when demand for Product 1 is low and high, respectively. ∎

The following method summarizes the steps performed in Example 2 and can be used to approximate a potentially large set of scenarios into a small set of conditional scenarios.

Method 1. *(From scenarios to conditional scenarios)*

- *Objective:* To approximate a set of scenarios by a set of conditional scenarios.
- *Input:*
  a) A set of scenarios $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$ with probability $\tilde{p}^s = 1/S$, for all $s \in \mathcal{S} = \{1, \dots, S\}$.
  b) E, the number of conditional scenarios for coordinate r, for each $r \in \mathcal{R} = \{1, \dots, R\}$.
- *Output:* For each $r \in \mathcal{R}$, a set of conditional scenarios $\{\hat{\xi}^{re}\}_{e\in\mathcal{E}}$ and the corresponding probability values $\{\hat{p}^{re}\}_{e\in\mathcal{E}}$ with $\mathcal{E} = \{1, \dots, E\}$.
- *Steps:* For each $r \in \mathcal{R}$:
  1) Define the interval $\mathcal{I}_r = [a_r, b_r]$ such that

  $$a_r = \min_{s\in\mathcal{S}}\{\tilde{\xi}^s_r\} \qquad\qquad b_r = \max_{s\in\mathcal{S}}\{\tilde{\xi}^s_r\}.$$

  2) Partition the interval $\mathcal{I}_r$ into subintervals $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$, according to Assumption 1.b.
  3) Classify the scenarios such that the index set $\mathcal{S}_{re}$ contains the indexes of the scenarios that fulfill the condition $\tilde{\xi}^s_r \in \mathcal{I}_{re}$ for all $e \in \mathcal{E}$.
  4) For all $e \in \mathcal{E}$:
     i) Set S_{re} as the cardinality of $\mathcal{S}_{re}$.
     ii) Compute the corresponding conditional scenario and its probability value:

     $$\hat{\xi}^{re} = \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}\,] = \frac{1}{S_{re}} \sum_{s\in\mathcal{S}_{re}} \tilde{\xi}^s$$

     $$\hat{p}^{re} = S_{re}/S.$$

∎

Notice that:

- In this paper, for each $r \in \mathcal{R}$, we will consider subintervals $\mathcal{I}_{re}$ of the same length, given by $(b_r - a_r)/E$ (although this method could also consider subintervals of different length).
- The previous method considers equiprobable scenarios, but it could easily be adapted to the non equiprobable case by using Proposition 5.

- The idea in the previous method is to approximate the scenarios in group $\mathcal{S}_{re}$ by the conditional scenario $\hat{\xi}^{re}$ such that its probability is proportional to the number of scenarios of the group. This generalizes the idea of expected scenario, where one considers only one scenario group which is approximated by $\bar{\xi}$.
- The previous method is distribution free since it only requires a set of scenarios.
- This approach can also be used when the the distribution of $\boldsymbol{\xi}$ is known. The procedure is as follows: Sample a set of scenarios from $\boldsymbol{\xi}$ and proceed with Steps 1 to 4.
- This procedure can be used to effectively approximate a possibly large set of scenarios by a small set of conditional scenarios.

4.2. Randomized Conditional Scenarios

Prior to defining the Conditional Scenario (CS) problem in Section 5, let us explain the Randomized Conditional Scenario (RCS) concept. Given a finite support random vector $\tilde{\boldsymbol{\xi}}$, the CS random vector $\hat{\boldsymbol{\xi}}^r$ is an approximation to $\tilde{\boldsymbol{\xi}}$ for all $r \in \mathcal{R}$. Since it would not have sense to consider all these approximations simultaneously, they are considered one each time (randomly), that is, each $\hat{\boldsymbol{\xi}}^r$ approximates $\tilde{\boldsymbol{\xi}}$ with probability $1/R$. For this reason, it is considered a random index $\mathbf{r}$ with support $\mathcal{R}$ and uniform probability $P\big(\mathbf{r} = r\big) = 1/R$ for all $r \in \mathcal{R}$, to define the RCS random vector $\hat{\boldsymbol{\xi}}^{\mathbf{r}}$ (notice the $\mathbf{r}$, being a random variable, is written in boldface). Therefore, $\hat{\boldsymbol{\xi}}^{\mathbf{r}}$ approximates the random vector $\tilde{\boldsymbol{\xi}}$ by a set of R random vectors of dimension R, that is, $\{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$, each one taken with probability $1/R$.

Example 3. *(Randomized conditional scenarios: Calculations)*

Let us continue Example 2 where the future demand was approximated by two conditional scenarios considering low an high demands for Product 1, that is, $\tilde{\boldsymbol{\xi}}$ was approximated by $\hat{\boldsymbol{\xi}}^1$. Now, let us see how $\tilde{\boldsymbol{\xi}}$ could be approximated by conditional scenarios with high and low demand for each product (not only for Product 1 as in Example 2).

All we need to do is to approximate $\tilde{\boldsymbol{\xi}}$ by a set of randomized conditional scenarios: In the first step, one computes the conditional scenarios ξ^{re} for $r \in \mathcal{R} = \{1, 2, 3\}$ and $e \in \mathcal{E} = \{1, 2\}$, by using Method 1 (see Example 2). In the second step, the probabilities thus calculated are multiplied by $1/R$ (in this case $1/3$). The corresponding randomized conditional scenarios and probabilities are as follows (the first column corresponds to $\hat{\boldsymbol{\xi}}^1$ and son on):

$$\hat{\xi}^{1,1} = \begin{pmatrix} \mathbf{1.2} \\ 3.4 \\ 7.6 \end{pmatrix} \qquad \hat{\xi}^{2,1} = \begin{pmatrix} 1.0 \\ \mathbf{3.0} \\ 7.2 \end{pmatrix} \qquad \hat{\xi}^{3,1} = \begin{pmatrix} 1.1 \\ 3.2 \\ \mathbf{7.0} \end{pmatrix}$$

$$\hat{p}^{1,1}/R = 6/30 \qquad \hat{p}^{2,1}/R = 5/30 \qquad \hat{p}^{3,1}/R = 5/30$$

$$\hat{\xi}^{1,2} = \begin{pmatrix} \mathbf{3.1} \\ 7.4 \\ 13.4 \end{pmatrix} \qquad \hat{\xi}^{2,2} = \begin{pmatrix} 2.9 \\ \mathbf{7.0} \\ 12.6 \end{pmatrix} \qquad \hat{\xi}^{3,2} = \begin{pmatrix} 2.8 \\ 6.8 \\ \mathbf{12.8} \end{pmatrix}$$

$$\hat{p}^{1,2}/R = 4/30 \qquad \hat{p}^{2,2}/R = 5/30 \qquad \hat{p}^{3,2}/R = 5/30,$$

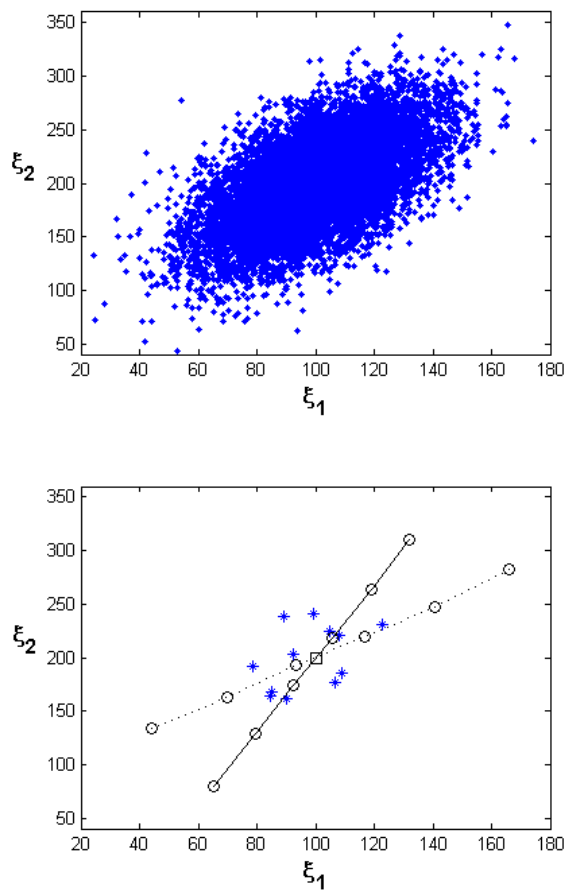

Fig. 3. The initial set of $10,000$ scenarios (top plot) is approximated by a random sample of 12 scenarios labelled by '*' (bottom plot) and by a set of 12 randomized conditional scenarios labelled by 'o' (the lines are plotted to show their trajectories). Conditional scenarios on the dotted line and the solid line correspond to $\hat{\boldsymbol{\xi}}^1$ and $\hat{\boldsymbol{\xi}}^2$, respectively. The square in the center approximates the initial set of scenarios by its expectation.

where we have indicated the low and high demands in boldface for each product. ■

In the following example, we show the result of approximating a large set of scenarios by a small set of randomized conditional scenarios.

Example 4. *(From scenarios to randomized conditional scenarios: Geometry)* In order to illustrate the geometry of the randomized conditional scenarios, let us draw a random sample of $10,000$ scenarios from the bivariate normal distribution $N_2(\mu, \Sigma)$ such that

$$\mu = \begin{pmatrix} 100 & 200 \end{pmatrix}^\intercal \qquad \Sigma = \begin{pmatrix} 400 & 480 \\ 480 & 1600 \end{pmatrix},$$

and approximate them by 12 randomized conditional scenarios (see Figure 3). ∎

5. The conditional scenario problem

In this context, the CS problem can be defined as the approximation of the RP problem that is based on the approximation of the random vector $\tilde{\boldsymbol{\xi}}$ by the RCS random vector $\hat{\boldsymbol{\xi}}^{\mathbf{r}}$, that is, by the set of random vectors $\{\hat{\boldsymbol{\xi}}^r\}_{r\in\mathcal{R}}$, each one taken with probability $1/R$. Otherwise said, the set of scenarios $\{\tilde{\xi}^s\}_{s\in\mathcal{S}}$ is approximated by the set of conditional scenarios $\{\hat{\xi}^{re}\}_{re\in\mathcal{RE}}$, in order to define the CS problem:

$$
\begin{aligned}
\min_{\hat{x}} \quad & z_{CS} = c_1^\intercal \hat{x}_1 + \frac{1}{R} \sum_{re\in\mathcal{RE}} \hat{p}^{re}\, \hat{c}_2^{re\intercal} \hat{x}_2^{re} && && (14)\\
\text{s.t.} \quad & A_1\hat{x}_1 = b_1 && && (15)\\
& \hat{A}_2^{re}\hat{x}_1 + \hat{B}_2^{re}\hat{x}_2^{re} = \hat{b}_2^{re} && re \in \mathcal{RE} && (16)\\
& \hat{x}_1 \geq 0 && && (17)\\
& \hat{x}_2^{re} \geq 0 && re \in \mathcal{RE} && (18)\\
& \hat{x}_{1j} \text{ integer} && j \in \mathcal{J}_1 && (19)\\
& \hat{x}_{2j}^{re} \text{ integer} && re \in \mathcal{RE},\ j \in \mathcal{J}_2. && (20)
\end{aligned}
$$

In this context

$$\hat{\xi}^{re} = \mathbb{E}[\,\tilde{\boldsymbol{\xi}} \mid \tilde{\boldsymbol{\xi}}_r \in \mathcal{I}_{re}\,] = \mathrm{vec}\big(\,\hat{c}_2^{re}, \hat{A}_2^{re}, \hat{B}_2^{re}, \hat{b}_2^{re}\,\big) \qquad re \in \mathcal{RE},$$

such that $\{\mathcal{I}_{re}\}_{e\in\mathcal{E}}$ is a partition of $\mathcal{I}_r$ that fulfills Assumption 1 for each $r \in \mathcal{R}$. Notice that, the CS problem is always based on 'randomized conditional scenarios'. However, we will normally drop the term 'randomized' and simply say 'conditional scenarios'. Finally, it is important to mention that although the CS problem has the same structure as the RP problem (1)–(7), it gives suboptimal solutions for the RP problem since it is an approximation.

6. Computational study

Let us assume that the RP problem (1)-(7) has been formulated in terms of a large set of S_1 scenarios. Approximating the RP problem by using a small set of S_2 randomly sampled scenarios can be considered as a simple version of the Sample Average Approximation (SAA) method (Kleywegt et al., 2002). The SAA method is a Monte Carlo sampling method which, according to (Henrion and Romisch, 2017), is the preferred and mostly used scenario generation method for the numerical solution of stochastic programming problems. The objective of this section is to compare the SAA and CS problems as approximations to the RP problem intended to reduce its computational burden. We will try to answer the following questions: How does the CS problem, formulated in terms of a given number of conditional scenarios, compare with the SAA counterpart formulated in terms of the same number of (randomly sampled) scenarios? That is, do the two approaches obtain solutions of similar quality? Are the corresponding solution times similar? Computations have been conducted on a PC using Windows 7 (64 bits), with a Intel Core i5 processor, 2.67GHz and 8 GB of RAM. The corresponding MILP problems have been solved by CPLEX 12.6 with default parameter values.

In this computational study, we will solve the well known risk-return portfolio optimization problem where one decides how much to invest in each asset taking into account its random returns $\tilde{\mathbf{r}}_j$ (Abdelaziz and Masmoudi, 2014; Konno, 2011). The model here solved corresponds to the mean semi-absolute deviation (MSAD) model with hard real-world constraints (Cesarone et al., 2015), namely, a limit on the number of assets to be included in the portfolio (a cardinality constraint) and lower and upper limits on their weights (a buy-in threshold). The MSAD corresponds to the mean of the deviations below the portfolio expected rate of return and can be computed as follows:

$$MSAD(x) = \mathbb{E}[\ [\tilde{\mathbf{r}}^\intercal x - \mu^\intercal x]_-\ ], \tag{21}$$

where $x = \begin{pmatrix} x_1 \dots x_J \end{pmatrix}^\intercal$ is the portfolio selection weights, $\tilde{\mathbf{r}}^\intercal x$ is the (uncertain) portfolio return, $\mu^\intercal x$ is the portfolio expected rate of return and $[\alpha]_- = -\min\{0, \alpha\}$ for any number α. It can be shown that the MSAD model is equivalent to the mean absolute deviation (MAD) model but with half the number of variables (Konno, 2011; Speranza, 1996). According to (Cesarone et al., 2015), when hard real-world constraints are incorporated into portfolio optimization models, most authors prefer to use MILP models as, for example, the MSAD model instead of quadratic ones as, for example, the mean-variance model (Markowitz, 1952), to improve the computational solvability.

The objective is to obtain an expected return of $b\%$ while attaining the lowest possible MSAD, by investing in 13 assets, at most, from an initial set of 50 candidate assets: 25 stocks and 25 commodities (corn, milk, oil, copper, gold, etc.). The investor considers these two types of assets since commodities are known to move in the opposite direction of the stock market, on average. That is, commodities are, in general, negatively correlated with stocks. In this way the investor can invest in both to diversify and hedge his portfolio (stabilize his portfolio from volatility).

6.1. The SAA portfolio optimization problem

The SAA portfolio optimization problem, or for short, the SAA problem, corresponds to solve the portfolio optimization problem with recourse formulated in terms of a small set of randomly sampled scenarios. The deterministic parameters, the random parameters and the decision variables are described in Tables 3, 4 and 5, respectively. To formulate the SAA problem we consider $S_2 = 800$ scenarios randomly sampled from the initial set of $S_1 = 10^5$ scenarios that model the random return rates. We consider 800 scenarios in order to match the number of conditional scenarios that we will consider in the CS counterpart. For simplicity, the initial set of equiprobable scenarios were drawn from the multivariate normal distribution $N_J(\mu, \Sigma)$ with values for μ_j and σ_{j_1,j_2} reported in Table 4. However, the CS approach described in this paper could be used for any set of scenarios (either corresponding to historical data or

Table 3
Indexes and deterministic parameters of the SAA problem.

Parameter	Value	Unit	Description
j		-	Index for assets (stocks and commodities)
J	50	assets	Number of assets
$\mathcal{J}$	$\{1,\ldots,J\}$	-	Index set for assets
k	-	-	Index for stocks
K	25	stocks	Number of stocks
$\mathcal{K}$	$\{1,\ldots,K\}$	-	Index set for stocks
l	-	-	Index for commodities
L	25	commodities	Number of commodities
$\mathcal{L}$	$\{1,\ldots,L\}$	-	Index set for commodities
$\overline{rs}_k$	$0.05+0.10k/K$	%	Expected return rate of stock k
$\overline{rc}_l$	$0.10+0.10l/L$	%	Expected return rate of commodity l
$\bar{r}$	$\mathrm{vec}\big(\overline{rs},\overline{rc}\big)$	-	Vector of expected return rates
b	$\big(\sum_{j\in\mathcal{J}}\bar{r}_j\big)/J$	%	Required level of expected return for the portfolio
U	$\lceil J/4\rceil=13$	assets	Maximum number of assets in the portfolio
lb	$0.5/U=\ \ 3.85$	%	Buy-in lower bound
ub	$1.5/U=11.54$	%	Buy-in upper bound

generated in the computer by Monte Carlo sampling techniques).

The SAA portfolio optimization problem can be written as follows:

$$\min_{\tilde{u},\tilde{x},\tilde{y}} \quad z_{SAA}=\sum_{s\in\mathcal{S}_2}\tilde{p}^s\tilde{y}^s \tag{22}$$

$$\text{s.t.}\quad lb\,\tilde{u}_j\le\tilde{x}_j \qquad j\in\mathcal{J} \tag{23}$$

$$\tilde{x}_j\le ub\,\tilde{u}_j \qquad j\in\mathcal{J} \tag{24}$$

$$\sum_{j\in\mathcal{J}}\tilde{u}_j\le U \tag{25}$$

$$\sum_{j\in\mathcal{J}}\tilde{x}_j=1 \tag{26}$$

$$\sum_{j\in\mathcal{J}}\bar{r}_j\tilde{x}_j=b \tag{27}$$

$$\sum_{j\in\mathcal{J}}\tilde{r}_j^s\tilde{x}_j+\tilde{y}^s\ge b \qquad s\in\mathcal{S}_2 \tag{28}$$

$$\tilde{u}_j\in\{0,1\} \qquad j\in\mathcal{J} \tag{29}$$

$$\tilde{x},\tilde{y}^s\ge 0 \qquad s\in\mathcal{S}_2. \tag{30}$$

This problem corresponds to the RP problem (1)–(7) formulated in terms of 800 randomly sampled scenarios. Notice that the MSAD defined in equation (21) is computed by means of equations (22), (27), (28) and (30). After solving this SAA problem by using the MILP solver we have obtained $z^*_{SAA}=$

Table 4
Random parameters of the SAA problem.

Parameter	Value	Unit	Description
$\tilde{\mathbf{r}}$	$\left(\tilde{\mathbf{r}}_1 \ldots \tilde{\mathbf{r}}_J\right)^\intercal$	%	Random return rates represented by a random sample of $S_1 = 10^5$ equiprobable scenarios $\{\tilde{r}^s\}_{s\in\mathcal{S}_1}$
$\tilde{r}^s$	$\left(\tilde{r}_1^s \ldots \tilde{r}_J^s\right)^\intercal$	%	Realization of $\tilde{\mathbf{r}}$ (scenario)
μ_j	$\bar{r}_j$	%	Expected return rate of asset j
σ_j	$\bar{r}_j$	%	Standard deviation of $\tilde{\mathbf{r}}_j$
$\rho^s_{k_1,k_2}$	0.6	-	Correlation between return rates of stocks k_1 and k_2, $\quad k_1 \neq k_2$
$\rho^c_{l_1,l_2}$	0.6	-	Correlation between return rates of commodities l_1 and l_2, $\quad l_1 \neq l_2$
$\rho^{sc}_{k_1,l_2}$	-0.4	-	Correlation between return rates of stock k_1 and commodity l_2
ρ	$\begin{pmatrix} \rho^s & \rho^{sc} \\ \rho^{sc\,\intercal} & \rho^c \end{pmatrix}$		Correlation matrix
σ_{j_1,j_2}	$\rho_{j_1,j_2}\,\sigma_{j_1}\sigma_{j_2}$	$(\%)^2$	Covariance between $\tilde{\mathbf{r}}_{j_1}$ and $\tilde{\mathbf{r}}_{j_2}$, $\quad j_1 \neq j_2$
R	J	-	Number of random parameters
E	16	-	Number of conditional scenarios per component of $\tilde{\mathbf{r}}$
S_1	10^5	-	Initial number of scenarios
$\mathcal{S}_1$	$\{1,\ldots,S_1\}$	-	Index set for the initial set of scenarios
S_2	800	-	Number of scenarios used in the SAA problem
$\mathcal{S}_2$	$\{1,\ldots,S_2\}$	-	Index set for the scenarios used in the SAA problem
s	-	-	Index for scenarios

Table 5
Decision variables of the SAA problem.

Decision variable	Unit	Description
$\tilde{u}_j$	-	$\tilde{u}_j = 1$, if one invests in the jth asset $\tilde{u}_i = 0$, otherwise
$\tilde{x}_j$	%	Proportion of the funds to be placed in the jth asset
$\tilde{y}^s$	%	Negative part of the deviation from the expected return $\tilde{y}^s = [\tilde{r}^{s\intercal}\tilde{x} - \bar{r}^\intercal\tilde{x}]_-$ under scenario s

1.57% and

$$\tilde{x}_1^{*\intercal} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10.52 & 11.54 & 8.95 & 0 & 0 & 0 & 9.24 & 9.80 & 8.38 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5.02 & 0 & 6.48 & 0 & 0 & 0 & 5.89 & 8.24 & 0 & 0 \\ 5.74 & 0 & 0 & 0 & 4.70 & 0 & 0 & 5.50 & 0 & 0 \end{pmatrix} \text{ (in \%)}.$$

We have broken vector $\tilde{x}_1^{*\intercal}$ into rows given its long length, such that the first row corresponds to the first ten assets.

6.2. The CS portfolio optimization problem

To formulate the CS problem, first we calculate the conditional scenarios that approximate the initial set of $S_1 = 10^5$ scenarios by using Method 1, with $R = 50$ random parameters and $E = 16$ conditional scenarios per random parameter (in total: $R\cdot E = 800$ conditional scenarios). Notice that the CS problem has the same structure as the SAA problem (22)-(30). The only difference is the way to approximate the random vector $\tilde{\mathbf{r}}$ (conditional scenarios $\hat{r}^{re}$ versus randomly sampled scenarios $\tilde{r}^s$). Therefore, the resulting MILP problems have the same dimensions for the two approaches. After solving the CS problem we have obtained $z^*_{CS} = 0.52\%$ and

$$\hat{x}_1^{*\mathsf{T}} = \begin{pmatrix} 0 & 0 & 0 & 9.04 & 0 & 0 & 0 & 11.54 & 0 & 11.54 \\ 0 & 0 & 11.54 & 6.35 & 0 & 11.54 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11.54 & 0 & 0 \\ 0 & 0 & 0 & 3.85 & 0 & 0 & 0 & 11.54 & 0 & 11.54 \end{pmatrix} \text{ (in \%)}.$$

6.3. Comparing the CS and SAA solutions

Let us call CS the optimal cost obtained by solving the CS problem, that is CS $= z^*_{CS}$ (notice that the value CS corresponds to the optimal MSAD). As pointed out in (Beltran-Royo, 2017), the expected cost *predicted* by CS is, in general, different from the expected cost *achieved* by the CS solution. The achieved expected cost is known in the literature as the 'expected result of using the CS solution' (E-CS) and can be computed by solving the E-CS problem, which is nothing but the RP problem (1)–(7) with the additional constraint $\tilde{x}_1 = \hat{x}^*_1$, that is, fixing the first stage decision to the first stage CS solution. Notice that the E-CS replicates the concept E-EV (expected result of using the EV solution) as described in (Birge and Louveaux, 2011). Analogously for SAA and E-SAA.

To formulate the E-CS and E-SAA problems we have used the initial set of $S_1 = 10^5$ scenarios. In Table 6 we observe that the predicted and the achieved MSAD values are different for both methods. Notice that, for practical purposes the achieved values are more relevant that the predicted values, since the former are the ones that measure the actual performance of the corresponding optimal solutions. For the SAA method this difference is small and could be reduced practically to zero by taking a sufficiently large number of scenarios (Shapiro et al., 2009) (although this could make the resulting SAA instance intractable). The reason is that the SAA problem is nothing but the RP problem (1)–(7) formulated in terms of a sample of scenarios of arbitrary cardinality. However, for the CS approach this difference is larger and a relevant difference would remain even for a very large number of conditional scenarios. The reason is that the CS problem is only an approximation to the RP problem (thus, the optimal CS solution is, in general, suboptimal for the RP problem). Therefore, as pointed out in the introduction, the CS approach is intended to be used only in cases where the scenario approach results impractical regarding the solution time.

In Table 6 we also observe that the SAA solution time is more than eight times the CS solution time. We investigate this large difference in the next section.

Table 6
Comparing the CS and the SAA solutions.

	Predicted MSAD		Achieved MSAD	
	CS	SAA	E-SAA	E-CS
MSAD (%)	0.52	1.57	1.84	1.98
Solution time (s)	7.38	63.29	0.07	0.07

Table 7
Size of the MILP instances. The CS and the SAA problems have the same sizes.

Instance	J	Scenarios	Rows	Columns
1	20	320	363	360
2	40	640	723	720
3	60	960	1,083	1,080
4	80	1,280	1,443	1,440
5	100	1,600	1,803	1,800
6	120	1,920	2,163	2,160
7	140	2,240	2,523	2,520
8	160	2,560	2,883	2,880
9	180	2,880	3,243	3,240
10	200	3,200	3,603	3,600

6.4. *Comparing the CS and SAA performances*

In this section we compare the SAA and CS performances by solving the portfolio optimization problem with hard real-world constraints for different number of assets. Table 7 reports the number of: assets J (50% stocks and 50% commodities), scenarios S_2 and rows/columns of the constraint matrix of the corresponding MILP instances. The initial number of scenarios is constant in all the instances ($S_1 = 10^5$). The other parameters are summarized in Tables 3 and 4.

The CS and SAA performances are compared in Figures 4 to 6. Figure 4 shows that the achieved MSAD is similar for the two approaches (slightly better for SAA). Notice that the SAA instances 5 to 10 could not be solved within a solution time limit of 3,600 s. Figure 5 corresponds to the ‘solution time’. In Table 9, we observe that the average solution time is 51 seconds and over 2,295 seconds for the CS and SAA approaches, respectively. That is, the CS approach was, on average, over 45 times faster than the SAA approach. Notice that the CS solution time, also accounts for the time to compute the conditional scenarios (see Table 9).

Finally, Figure 6 plods to the ‘LP gap’ (the relative gap between the optimal MILP cost and the optimal cost of the corresponding LP relaxation). In Table 10, we observe that the LP gap of the first four instances is, on average, 0.0354% and 5.4175% for the CS and SAA approaches, respectively. That is, the LP gap of the first four CS instances was, on average, 153 times smaller than the LP gap of the corresponding SAA instances. Therefore, a possible reason for the faster performance of the CS approach is its smaller LP gap (it is well known that the difficulty of MILP problems is strongly related to the LP gap).

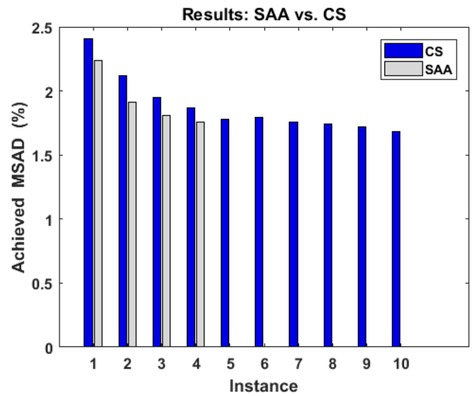

Fig. 4. The achieved MSAD obtained by the CS and SAA approaches are similar (slightly better for the SAA approach). The SAA instances 5 to 10 could not be solved within the solution time limit (3,600 s). For details see Tables 8 and 10.

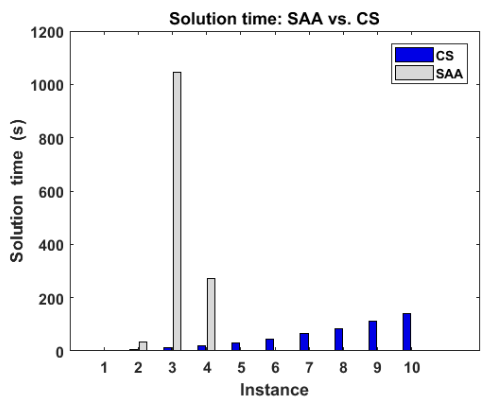

Fig. 5. For the first four instances the CS approach was, on average, over 30 times faster than the SAA approach. For details see Table 9.

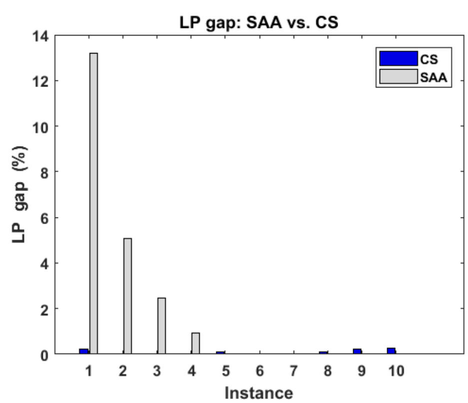

Fig. 6. The LP gap of the first four CS instances was, on average, 153 times smaller than the LP gap of the corresponding SAA instances. For details see Table 10.

Table 8
Comparing the E-CS and E-SAA. The SAA instances 5 to 10 could not be solved within a solution time limit of 3,600 s.

Instance	E-CS (%)	E-SAA (%)
1	2.41	2.24
2	2.12	1.91
3	1.95	1.81
4	1.87	1.76
5	1.78	-
6	1.79	-
7	1.76	-
8	1.74	-
9	1.72	-
10	1.68	-
Average	1.88	-

Table 9
Comparing the CS and SAA solution times. CS_1 = 'Time to compute the conditional scenarios', CS_2 = 'Time to solve the CS problem', Total = CS_1 + CS_2

Instance	CS problem			SAA problem
	CS_1	CS_2	Total (s)	Total (s)
1	2	0.1	2	1
2	5	0.2	5	33
3	10	0.5	11	1047
4	18	1.2	19	273
5	29	1.5	30	$>$ 3,600
6	43	2.6	46	$>$ 3,600
7	60	4.0	64	$>$ 3,600
8	80	5.4	85	$>$ 3,600
9	103	7.7	111	$>$ 3,600
10	131	10.3	141	$>$ 3,600
Average	48	3	51	$>$ 2,295

Tables 8-10 report the figures used to plot Figures 4-6, respectively.

6.5. CS performance regarding the number of scenarios (sensitivity analysis)

In this section we perform a sensitivity analysis of the CS approach regarding solution quality and solution time. With this objective in mind, instance 10, the largest instance in Table 7, is solved for an increasing number of scenarios: from 800 to 5600 as reported in Table 11. The solution quality is assessed by means of the achieved MSAD which is given by the E-CS value. In Table 11 one observes that the solution quality improves by augmenting the number of scenarios up to 3200 (best value E-CS

Table 10
Comparing the LP gaps.

Instance	CS problem LP bound (%)	z^*_{CS} (%)	LP gap (%)	SAA problem LP bound (%)	z^*_{SAA} (%)	LP gap (%)
1	0.6097	0.6104	0.1235	1.5438	1.7476	13.20
2	0.5360	0.5360	0.0000	1.5058	1.5823	5.08
3	0.5156	0.5157	0.0115	1.5558	1.5941	2.46
4	0.5004	0.5004	0.0065	1.4676	1.4813	0.93
5	0.4960	0.4962	0.0441	1.4883	-	-
6	0.4985	0.4985	0.0115	1.5652	-	-
7	0.4961	0.4961	0.0000	1.5353	-	-
8	0.4896	0.4899	0.0596	1.4235	-	-
9	0.4779	0.4784	0.1104	1.4791	-	-
10	0.4868	0.4875	0.1398	1.4844	-	-
Average	0.5107	0.5109	0.0506	1.5049	-	-

Table 11
Solution quality and solution time for an increasing number of scenarios. CS_1 = 'Time to compute the conditional scenarios', CS_2 = 'Time to solve the CS problem', Total = CS_1 + CS_2.

Number of conditional scenarios	Solution quality E-CS (%)	Solution time CS_1	CS_2	Total (s)
800	1,82	128	1	129
1600	1.79	130	3	133
2400	1.75	130	6	136
3200	1.68	131	10	141
4000	1.72	132	15	147
4800	1.71	131	20	151
5600	1.71	131	27	158

= 1.68%). Then, with more than 3200 scenarios the results are similar (slightly worse).

Regarding the solution time, one observes in the same table that the time to compute the conditional scenarios (CS_1) is almost the same for all the cases. This is because the total number of scenarios processed in each case is the same ($S_1 = 10^5$). Roughly speaking, as the number of conditional scenarios increases from 800 to 5600, the computational work per conditional scenario decreases proportionally (see Method 1). On the other hand, the time to solve the corresponding MILP instance (CS_2) increases faster than the number of scenarios. All in all, regarding solution quality and solution time, 3200 scenarios seems the best choice in Table 11, which corresponds to $E = 16$ the value used in previous sections (E is the number of conditional scenarios per random parameter of the problem). For practical purposes it is advisable to tune parameter E (16 represents a good starting point). Of course, the approach in this section has been empirical and the impact of the parameter E is a matter that deserves further research.

7. Conclusions

Although the definition of conditional scenario (CS) given in (Beltran-Royo, 2017) is general, in practice, it is suitable basically for the multivariate normal distribution. The main contributions of this paper have been: First, to give a new definition of conditional scenario that is suitable for any continuous or discrete multivariate distribution. Second, to study some theoretical properties of this new definition. In this respect, Proposition 3 shows that the new definition and the former definition are equivalent for the multivariate normal case. Third, to derive an effective method, based on the new definition, to approximate a potentially large set of scenarios by a reduced set of conditional scenarios. Such a procedure was not possible with the former definition of conditional scenario. Forth, to perform a computational study to test this new method.

In the computational study, we have solved the portfolio optimization problem with hard real-world constraints to compare the performances of the CS approach and the sample average approximation (SAA) approach, with the same number of scenarios. In this comparison, we have observed that the solution performance, in terms of the achieved MSAD (mean semi-absolute deviation), is similar for the two approaches (slightly better for the SAA approach).

Regarding the solution time, the CS approach has shown to be much faster than the SAA counterpart. Furthermore, the CS instances have shown a much smaller LP gap than the corresponding SAA instances, which could explain the shorter CS solution times. This hypothesis deserves further research. Besides studying the LP gap, there remain other research questions regarding the CS approach, as for example, its performance in multistage stochastic optimization (Bruno et al., 2016; Shi et al., 2011) and in non-linear stochastic optimization (Tarhan et al., 2013), among others.

All in all, the CS approach, a basic approximation of the recourse problem, could be a good choice for practitioners interested in fast and good solutions for the portfolio problem with hard real-life constraints in cases where the solution time of the SAA approach would result infeasible for practical purposes.

Acknowledgments

We thank the editor and reviewers for their constructive and stimulating suggestions. We are also grateful to the Spanish Ministry of Economy, Industry and Competitiveness (grants MTM2015-63710-P and RTI2018-094269-B-I00).

References

Abdelaziz, F.B., Masmoudi, M., 2014. A multiple objective stochastic portfolio selection problem with random beta. *International Transactions in Operational Research* 21, 6, 919–933.

Alonso-Ayuso, A., Escudero, L.F., Guignard, M., Weintraub, A., 2018. Risk management for forestry planning under uncertainty in demand and prices. *European Journal of Operational Research* 267, 1051–1074.

Amorim, P., Costa, A.M., Almada-Lobo, B., 2015. A hybrid path-relinking method for solving two-stage stochastic integer problems. *International Transactions in Operational Research* 22, 1, 113–127.

Beltran-Royo, C., 2017. Two-stage stochastic mixed-integer linear programming: The conditional scenario approach. *Omega - The International Journal of Management Science* 70, 31–42.

Birge, J.R., Louveaux, F.V., 2011. *Introduction to Stochastic Programming (Second Edition)*. Springer (New York).

Bruno, S., Ahmed, S., Shapiro, A., Street, A., 2016. Risk neutral and risk averse approaches to multistage renewable investment planning under uncertainty. *European Journal of Operational Research* 250, 3, 979–989.

Calvet, L., Wang, D., Juan, A., Bove, L., 2019. Solving the multidepot vehicle routing problem with limited depot capacity and stochastic demands. *International Transactions in Operational Research* 26, 2, 458–484.

Cesarone, F., Scozzari, A., Tardella, F., 2015. Linear vs. quadratic portfolio selection models with hard real-world constraints. *Computational Management Science* 12, 3, 345–370.

Chang, Y.H., Solak, S., Clarke, J.P.B., Johnson, E.L., 2016. Models for single-sector stochastic air traffic flow management under reduced airspace capacity. *Journal of the Operational Research Society* 67, 1, 54–67.

Chen, M., Mehrotra, S., Papp, D., 2015. Scenario generation for stochastic optimization problems via the sparse grid method. *Computational Optimization and Applications* 62, 3, 669–692.

DeGroot, M.H., Schervish, M.J., 2012. *Probability and Statistics (Fourth Edition)*. Addison Wesley (Boston, USA).

Dupacova, J., Growe-Kuska, N., Romisch, W., 2003. Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming* 95, 3, 493–511.

Hardle, W., Simar, L., 2012. *Applied Multivariate Statistical Analysis (Third Edition)*. Springer, Berlin.

Heitsch, H., Romisch, W., 2003. Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications* 24, 2-3, 187–206.

Henrion, R., Romisch, W., 2017. Optimal scenario generation and reduction in stochastic programming, Humboldt-Universitat zu Berlin, Institut fur Mathematik. Technical report.

Homem-de-Mello, T., Bayraksan, G., 2014. Monte Carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science* 19, 1, 56–85.

Hoyland, K., Kaut, M., Wallace, S.W., 2003. A heuristic for moment-matching scenario generation. *Computational Optimization and Applications* 24, 2-3, 169–185.

Hoyland, K., Wallace, S.W., 2001. Generating scenario trees for multistage decision problems. *Management Science* 47, 2, 295–307.

Jacod, J., Protter, P., 2004. *Probability Essentials*. Springer.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1994. *Continuous Univariate Distributions (Second Edition), Vol. 1*. New York: John Wiley and Sons.

Karademir, S., Kong, N., Prokopyev, O.A., 2014. On greedy approximation algorithms for a class of two-stage stochastic assignment problems. *Optimization Methods and Software* 29, 1, 42–67.

King, A.J., Wallace, S.W., 2012. *Modeling with Stochastic Programming*. Springer Science and Business Media, New York.

Kleywegt, A.J., Shapiro, A., Homem-de-Mello, T., 2002. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization* 12, 2, 479–502.

Konno, H., 2011. Stochastic Programming: The state of the Art in Honor of George B. Dantzig, chap. Mean-Absolute Deviation Model, Springer. pp. 239–255.

Lubin, M., Petra, C.G., Anitescu, M., 2012. The parallel solution of dense saddle-point linear systems arising in stochastic programming. *Optimization Methods and Software* 27, 4-5, 845–864.

Markowitz, H., 1952. Portfolio selection. *The Journal of Finance* 7, 1, 77–91.

Morales, J.M., Pineda, S., Conejo, A.J., Carrion, M., 2009. Scenario reduction for futures market trading in electricity markets. *IEEE Transactions on Power Systems* 24, 2, 878–888.

Pages-Bernaus, A., Ramalhinho, H., Juan, A.A., Calvet, L., 2019. Designing e-commerce supply chains: a stochastic facility-location approach. *International Transactions in Operational Research* 26, 2, 507–528.

Pflug, G.C., Pichler, A., 2014. *Multistage Stochastic Optimization*. Springer, Cham (Switzerland).

Raba, D., Estrada-Moreno, A., Panadero, J., Juan, A.A., 2020. A reactive simheuristic using online data for a real-life inventory routing problem with stochastic demands. *International Transactions in Operational Research, DOI: 10.1111/itor.12776*

Santoso, T., Ahmed, S., Goetschalckx, M., Shapiro, A., 2005. A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research* 167, 1, 96–115.

Sen, S., Sherali, H.D., 2006. Decomposition with branch-and-cut approaches for two-stage stochastic mixed-integer programming. *Mathematical Programming* 106, 2, 203–223.

Shapiro, A., Dentcheva, D., Ruszczynski, A., 2009. *Lectures on Stochastic Programming: Modeling and Theory*, Vol. 9. Society for Industrial Mathematics, Philadelphia.

Shi, Y., Wu, F., Chu, L.K., Sculli, D., Xu, Y.H., 2011. A portfolio approach to managing procurement risk using multi-stage stochastic programming. *Journal of the Operational Research Society* 62, 11, 1958–1970.

Speranza, M.G., 1996. A heuristic algorithm for a portfolio optimization model applied to the Milan stock market. *Computers and Operations Research* 23, 5, 433–441.

Takriti, S., Birge, J.R., 2000. Lagrangian solution techniques and bounds for loosely coupled mixed-integer stochastic programs. *Operations Research* 48, 1, 91–98.

Tarhan, B., Grossmann, I.E., Goel, V., 2013. Computational strategies for non-convex multistage minlp models with decision-dependent uncertainty and gradual uncertainty resolution. *Annals of Operations Research* 203, 1, 141–166.

Tong, X., Wu, F., Qi, L., 2009. Worst-case CVaR based portfolio optimization models with applications to scenario planning. *Optimization Methods and Software* 24, 6, 933–958.