Ways and means with scenarios *

Jitka Dupačová
Dept. of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic
e-mail: dupacova@karlin.mff.cuni.cz

Abstract. We shall survey and classify various ways of approximation of the true probability distribution by a discrete distribution concentrated in a finite number of atoms, called scenarios, which occur with preassigned probabilities. Resistance of the numerical results with respect to the choice of scenarios and/or of their probabilities and also the relation between results obtained for the selected scenarios and those valid for the true underlying distribution are of great practical interest. We shall discuss various approaches to output analysis in the framework of an expectation type stochastic program, bearing in mind the diverse origin of scenarios.

Keywords: Stochastic decision models, generation of scenarios, robustness, output analysis and stability

JEL classification: C44, C61, C81
AMS classification: 90C15, 90C31, 90C59

1. Introduction

The outcome of decision problems depends on the choice of decisions and also on various noncontrollable parameters. One setup of values of these parameters is called scenario. Scenarios have been exploited in various problems, such as

- stochastic immunization and dedication [38];
- tracking models [13];
- scenario analysis used in pilot or feasibility studies;
- evaluation of dynamic investment strategies [53];
- stochastic programming including multiperiod and dynamic models, e.g. [5, 23, 67, 72].

In this paper we accept the probabilistic interpretation of uncertainty with scenarios mostly understood as realizations of random vectors or time series trajectories. Even with this specification, the origin of scenarios can be very diverse; they can

*This research was supported by the Czech Science Foundation under Grants P201/12/0083 and P402/11/0150.

come from a truly discrete known distribution, can be obtained in the course of a discretization/approximation scheme, by simulation or by a limited sample information, they can result from a preliminary analysis of the problem with probabilities of their occurrence that may reflect an ad hoc belief or a subjective opinion of an expert.

To present different scenario generation procedures along with properties of the obtained output we shall focus on a class of stochastic optimization problems leaving aside both more complicated types of stochastic programs and the above mentioned, mostly descriptive applications which are based on evaluation of the outcomes along individual scenarios. We shall assume that in the stochastic optimization problem

$$\text{minimize} \quad F(x, P_0) := Ef(x, \omega) \text{ on a set } \mathcal{X} \tag{1}$$

ω is a random parameter with support $\Omega \subset \mathbb{R}^m$, probability distribution P_0 and the corresponding expectation $E = E_{P_0}$,

$\mathcal{X} \subset \mathbb{R}^n$ is a given nonempty closed set,

$f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ is a given random lower semicontinuous function.

The dimensionalities m, n can be very large.

A typical, even though not quite realistic assumption is that the probability distribution P_0 is known (decision making under risk) and independent of the decision x. The existence of expectations in (1) for all feasible decisions $x \in \mathcal{X}$ is guaranteed by special assumptions from case to case. The optimal value of the objective function in (1) will be denoted $\varphi(P_0)$.

The main stumbling block for algorithmic solution of problem (1) is the necessity to compute repeatedly at least the values of the multidimensional integral of the random objective function $f(x, \omega)$; for continuous distributions this is tractable for low dimensional cases or for simple separable random objective functions $f(x, \bullet)$ and for special types of probability distributions. For discrete probability distributions carried by $\omega_1, \ldots, \omega_S$ with probabilities $p_s > 0$, $s = 1, \ldots, S, \sum_s p_s = 1$ the objective function is the sum $F(x, P_0) := \sum_{s=1}^{S} p_s f(x, \omega_s)$ and its evaluation is relatively easy provided that there is a modest number of scenarios.

To deal with problems of an algorithmic solution of (1) various approximation schemes, both stochastic and deterministic ones, were designed; see for instance [7, 5, 44] and the references therein. *The goal is to get a numerically tractable optimization problem, or a sequence of such problems, whose solution would be acceptable as an approximation of the solution of the true underlying problem (1).*

For approximation of the true probability distribution P_0, one should certainly use the structure of the problem. Stability theorems guarantee closeness of optimal solutions of the true problem (1) and of its approximation if the probability distance of the two considered probability distributions is small; cf. [64] and references therein.

There are relatively many prospects if the approximation of P_0 reduces to an approximation of one-dimensional probability distributions: Besides of an approximation by a discrete distribution, one can use piecewise uniform distributions [3], or approximate the density by kernel estimates [35, 36], minimize the probability distance [60], etc. If there are independent sources of uncertainties the number of

scenarios needed to represent their mutual influence on the results is the product of the number of scenarios used to represent the impact of each source separately.

In the truly multi-dimensional case, approximation by discrete distributions is the prevailing approach. It means that the true distribution P_0 is replaced by a *discrete* distribution $\hat{P}$ concentrated in a finite number of points, say, $\omega_1, \ldots, \omega_S$ with probabilities $p_s > 0 \quad \forall s, \quad \sum_{s=1}^{S} p_s = 1$ that enter the coefficients and/or the functions in a known way. The atoms ω_s of the distribution P are called *scenarios*. For the already fixed scenarios and probabilities the problem to solve is a deterministic mathematical program

$$\text{minimize} \quad F(x, \hat{P}) := \sum_{s=1}^{S} p_s f(x, \omega_s) \text{ on a set } \mathcal{X}. \tag{2}$$

Suitable statistical methods such as principal components analysis help to reduce the dimensionality of the random vector and, at the same time, to preserve the correlation structure. Principal components are uncorrelated and their simulated values together with estimated scores are used for generating scenarios, e.g. for generating the future joint assets returns that enter the considered two-stage multiperiod model designed for investment planning [54]. Instead of principal components and scores one can use factors and factor loadings; this has been suggested e.g. in [8, 11, 43]. An alternative is to model the correlation structure of the data via copulas, cf. [48].

Naturally, one is interested in results for the true underlying problem. However, one cannot rely on a complete knowledge of the underlying probability distribution. Generation of scenarios and a validation of obtained results is then even more demanding. An important item is robustness of the obtained approximate optimal solution and optimal value: The procedure should be robust in the sense that small perturbations of the input, i.e., of the chosen scenarios and of their probabilities, should impair the outcome only slightly so that the obtained results remain close to the unperturbed ones and that somewhat larger perturbations do not cause a catastrophe. The importance of robust procedures increases with the complexity of the model and with its dimensionality.

The purpose of this paper is to survey and classify various approaches to scenario generation (Section 2) and to give an idea about the methodological advances developed for stability and postoptimality studies for the scenario based approximate problems (Section 3).

2. Scenarios and their generation

We have introduced scenarios as atoms of the discrete probability distribution $\hat{P}$ that is used to approximate the underlying (true, reference) probability distribution P_0. However, the main goal of scenario generation procedures is to get a representative set of scenarios which supports sensible decisions and is acceptable for the user. This is an ambitious task in which compromise is needed between the precision of the approximation and the size of the approximate problem, and which often requires a specific form of the input (e.g., a scenario tree for multistage problems). Hence, scenario generation is problem specific; for example, in financial models there are evident differences when considering prediction, pricing or decision making. Moreover,

the procedure should reflect both the problem structure and the available information that comes from theory, historical data and experience: To generate scenarios it is natural to use historical data (if any) in conjunction with an assumed background model, to apply suitable estimation, simulation and sampling procedures, heuristics, and also to reflect the opinion of experts based on their knowledge or sentiment, see e.g. [29, 50] or Chapter 2 of [77]. These are reasons why generation of scenarios cannot be reduced solely to forecasting the future development of the complex system under consideration.

Concerning the level of the available information, we can distinguish four types of problems:

- *The distribution P_0 is fully specified.*

  If this assumption is true, one needs to generate a discrete distribution $\hat{P}$ which approximates P_0. The scenarios, atoms of $\hat{P}$, can be obtained by sampling from P_0 or by application of a discretization or simulation scheme. Related to the chosen approximation technique, there are various possibilities how to draw conclusions about the optimal solution of the original problem.

  In its pure form this situation appears mostly in the context of testing the designed models and/or the performance of newly developed solvers; this was the case of the first papers on stochastic programming, [3, 33], as well as of applications surveyed in [49], and the assumed fully known distribution was mostly uniform [3], discrete [33] or normal. For application purposes, however, one should address also the modeling error due to misspecification of P_0; it composes then with errors due to the applied approximation method.

- *P_0 is known to belong to a specified parametric family.*

  The true probability distribution P_0 belongs to $\mathcal{P} = \{P_\theta, \theta \in \Theta\}$ and is identified by an unknown parameter value, say, $\theta_0 \in \Theta$. The assumed parametric form of the distribution P_θ should be preferably based on a theoretical model. If this assumption is accepted, the problem of an incomplete knowledge of distribution is transferred into a problem of estimation of the parameter θ_0 from the available data. The choice of the parametric form of the probability distribution or of the stochastic process corresponds to the choice of the model, the estimation of parameters to calibration of the model and a subsequent simulation, sampling or discretization procedure follows similarly as before. An example is sampling from the estimated model for loads and water quality in [74] or for electricity demand and oil prices in [32].

  The procedure is to estimate the parameter θ_0, say by $\hat{\theta}$, to generate scenarios using the estimated probability distribution $P_{\hat{\theta}}$ and to exploit the quantitative stability results in the output analysis. The quality of the parametric approach depends on the right choice of the parametric model and on the applied estimation procedure: Besides of the modeling error, there is also an estimation error by estimation of parameters and errors due to simulation from the already estimated probability distribution $P_{\hat{\theta}}$.

  This type of information appears frequently in stochastic programming problems in finance and water resources management and planning, partly due to

the fact that the relevant stochastic models of interest rates, cf. [1], and assets prices or those of water inflows came to the attention relatively early and have been well developed and supported by historical data. At present, it seems to be an increasing interest in building appropriate models for probability distributions of demand in service network design or for energy problems [9, 32], technological processes [26], etc.

Sometimes, the type of the parametric family of distributions may reflect the fact that, for the sake of the numerical tractability, only a specific family of probability distributions is taken into account. Frequently it is the multinormal distribution which is *inter alia* consistent with common models of multidimensional time series of observed data. However, this assumption may influence the results substantially; we refer to [75] which illustrates differences in results obtained for Gaussian and stable non-Gaussian distributions.

- *Only a sample information about P_0 is available.*

  Such information is mostly based on observed past data, e.g. [35, 36, 68]. For large sets of available data that are homogeneous enough, independent, identically distributed, the straightforward possibility is to use the empirical distribution P_ν based on a sample of size ν, the corresponding empirical objective function $F(x, P_\nu)$, the set of empirical optimal solutions and the optimal empirical value $\varphi(P_\nu)$.

  It is necessary to say that the available historical data are not always satisfactory for various reasons; they need not be observed in regular time instances; non-equilibrium markets or change-points are other examples. In such situation, an additional expert knowledge [50] and preprocessing can help to build sensible scenarios. We refer to preprocessing procedures (e.g. in [28, 35] for electricity demand or [41] for water inflows) or to adjustments to preserve specific values of sample moments [8, 51, 56].

  For small sample sizes it pays to exploit all available information to get an approximation which is a good substitute of the true problem. Several examples how to use such "soft" information were delineated in [76].

- The above mentioned procedures fail if there are *no reliable data.* Under such circumstances, scenarios and their probabilities are mostly based on experts forecasts (see e.g. [6] for predictions of future economic development, [31] for demand outlooks in production planning or in power generation planning, [26] for optimization of a technological process, [73] for financial models). Based on observed data or on sample moments, simple extremal scenarios can be constructed; branching to "upper" and "lower" cases and the expected value scenario are typical examples; see e.g. [29, 58] for an application and [13] for a discussion. Even though it is impossible to draw conclusions about the optimal solution of the true underlying problem, these scenarios are useful for obtaining bounds for the optimal value and it pays to analyse robustness and stability of the approximate solutions with respect to changes of these scenarios and their probabilities or inclusion of additional scenarios.

In majority of applications one can trace interactions of the four information levels and to use all available information is the best thing to do.

Different information levels can apply to distinct parameters of the model separately. In portfolio management, for instance, different classes of securities require different treatment, cf. [8, 53], inflation should be taken into account [53], deposits and liabilities can be driven by external factors such as mortality data or demographic situation [24] and also user intervention may play a substantial role. Similarly, the inflow scenarios can be based on a model fitted to historical data whereas scenarios of future water and electricity demand are roughly forecasted.

Also the structure of the problem influences essentially the possibilities of an adequate scenario generation. The easy case is the simple recourse problem with random right hand sides for which scenario generation reduces to a discretization of one-dimensional marginal distributions of the random right-hand sides. The most complicated problems are multistage stochastic programs with interstage dependent coefficients. They have been solved mainly for scenario trees designed by experts or constructed from individual scenario paths, cf. [21].

Monte Carlo simulation from the assumed probability distribution or directly from the data often enters the last step of the scenario generation procedure. The general strategy of scenario generation is concerned with the number of scenarios to be employed because of the evident trade-off between the dimensionality of the resulting problem and the information preserved. Importance sampling is used to increase the precision and efficiency of the crude Monte Carlo techniques, e.g. [11, 12, 40, 52, 57]. Another suggestion is to adjust the sample so that the sample moments are equal to the reference values, see e.g. [8, 39, 51]. Instead of using random sampling, Quasi Monte Carlo methods exploit function evaluations in points selected according to a specific non-random scheme. For a succinct discussion of differences in Monte Carlo sampling versus Quasi Monte Carlo methods, and other numerical techniques see [65] and Chapter 9 of [5].

Using simulation approaches one may generate very large sets of scenarios. The question is how to select a smaller, representative set of scenarios for which the optimization problem (2) will be manageable. There exist examples of successfully applied heuristic ideas, see e.g. [4, 32]. Scenario reduction technique based on quantitative stability results was initiated in [22]; see a brief summary in [65].

3. How to draw inference about the solution of the true problem?

As we shall see, there are various general results and specific approaches ready for the sought output analysis. Their application depends on the solved problem, is nontrivial and evaluation of their efficiency requires extensive numerical tests and experiments. It is necessary to take into account both the structure of the stochastic optimization problem and the probabilistic specification and to reflect properties of the input data. Generally speaking, the precision of the output cannot be higher than that of the input and it is easier to get conclusions concerning the optimal value than those for optimal decisions.

Scenario generation approaches are frequently composed of the following steps:

- choice of an appropriate probabilistic model of random parameters;
- calibration of the model (estimation of parameters) from data;
- generation scenarios using the calibrated model by means of simulation, discretization, etc.;
- evaluation of model coefficients or function values for each of considered scenarios.

Each of these steps may introduce a specific error into the final results – a modeling, estimation and sampling or simulation error. To quantify them one may use approaches of asymptotic or robust statistics and of parametric programming.

To simplify the exposition, *we shall focus on stability of optimal values and we shall assume in addition that in (1) $\mathcal{X}$ is convex, compact and $\mathcal{F}$ is a class of convex functions $f(\bullet, \omega)$ on $\mathcal{X}$ such that all expectations of $f(x, \omega)$ are finite.*

3.1. Use of parametric programming

Stability results for (1) with respect to the probability distribution, cf. [64], imply that under suitable assumptions, the optimal values and sets of optimal solutions of the true problem (1) with probability distribution P_0 and of its approximation with probability distribution $\hat{P}$ are close to each other if the distance of the two probability distributions is small enough. There are various distances or semi-distances of probability distributions, see e.g. [62], and an adequate choice is related to the structure of the stochastic program. For our problem (1) a natural possibility is

$$d_{\mathcal{F}}(P, \hat{P}) = \sup_{f \in \mathcal{F}} \left| \int_{\Omega} f(\omega) P(d\omega) - \int_{\Omega} f(\omega) \hat{P}(d\omega) \right|$$

with $\mathcal{F}$ a suitable class of measurable functions from Ω to $\mathbb{R}$ and $P, \hat{P}$ belonging to a class $\mathcal{P}$ of probability distributions on Ω. Here, $\mathcal{F}$ should contain all random objectives $f(x, \bullet)$ that may be considered.

Then we have

Theorem 1 (Theorem 1 in [22]). *Under the assumptions of the model (1), whenever $P, \hat{P} \in \mathcal{P}$ then there exists $\varepsilon > 0$ such that for $d_{\mathcal{F}}(P, \hat{P}) < \varepsilon$ the optimal values*

$$|\varphi(P) - \varphi(\hat{P})| < d_{\mathcal{F}}(P, \hat{P}).$$

More stringent assumptions concerning the structure of the problem (1) are needed to get qualitative and quantitative stability results for optimal solutions; see Theorems 5 and 9 of [64].

There are scenario generation methods based on the above quantitative stability results, e.g. [60, 62]. The idea is to construct scenarios and probabilities so that the distance of the discrete approximation $\hat{P}$ from the true probability distribution P_0 is as small as possible.

Another question is how far is the optimal output from the true one if the true probability measure P_0 has been misspecified. A possibility is to use the contamination technique to study the influence of specific deviations from the assumed probability distribution P_0: For an alternative fixed probability distribution $\overline{P}$ one uses in (1) instead of P_0 the contaminated probability distribution $P_\lambda = (1-\lambda)P_0 + \lambda\overline{P}$ with $0 \leq \lambda \leq 1$. The considered perturbation is thus represented only by a scalar parameter λ which enters the objective function in a linear way. This opens various possibilities for exploitation of parametric programming results. The optimal value function $\varphi(P_\lambda)$ is concave for $\lambda \in [0,1]$ and global lower and upper bounds for its value can be constructed. For the development of the theory see [15, 17], for applications e.g. [20, 25].

3.2. Empirical approximations

To a certain extent stability results of the preceding section apply also to approximations of P_0 by an empirical distribution P_ν. However, there is one main difference, namely, P_ν are *random* probability distributions and then also distances $d(P_0, P_\nu)$ are random and the results are valid almost surely. This can be formulated as follows:

Consider a sample space $(Z, \mathcal{Z}, \mu)$ of infinite sequences ζ, with an increasing sequence of σ-fields $(\mathcal{Z}_\nu)_{\nu=1}^\infty$ contained in $\mathcal{Z}$. For an increasing sample size the sample path $\zeta \in Z$ leads to a sequence of $\mathcal{Z}_\nu$-measurable probability distributions $\{P_\nu(\bullet, \zeta), \nu = 1, 2, \dots\}$ on $(\Omega, \mathcal{B})$ based on information collected up to ν.

The optimal value $\varphi(P_\nu)$ and optimal solutions $x^*(P_\nu)$ of the approximate stochastic program

$$\min_{x \in \mathcal{X}} F(x, P_\nu) := E_{P_\nu} f(x, \omega) \tag{3}$$

based on $P_\nu(\bullet, \zeta)$ depend on the used sample path ζ and *all presented results hold true for almost all sample paths* ζ, i.e., μ-a.s.

The empirical probability distributions are a special case with the sample path $\zeta = \{\omega_1, \omega_2, \dots\}$ obtained by the simple random sampling from $(\Omega, \mathcal{B}, P)$, $\mu = P_\infty$.

The *asymptotic properties* of empirical estimates, such as consistence, rates of convergence and asymptotic distributions have been studied in a host of papers; see e.g. [27, 61, 70] or Chapter 6 of [66] and in selected papers listed in [19]. For empirical distributions and for continuous, bounded random objectives $f(x, \bullet)$ the pointwise convergence of expected values $F(x, P_\nu) \to F(x, P_0)\, \forall x \in \mathcal{X}$ follows directly from the Law of Large Numbers. If $\mathcal{X}$ is compact and the convergence of expectations $F(x, P_\nu)$ is *uniform* on $\mathcal{X}$ then μ-a.s. convergence of optimal values $\varphi(P_\nu) \to \varphi(P_0)$ follows. It implies convergence in probability, i.e. for all $t > 0$

$$\mu\{\zeta \,:\, |\varphi(P_\nu) - \varphi(P_0)| > t\} \to 0 \text{ for } \nu \to \infty.$$

General consistency results are based on the notion of epi-convergence of lower semicontinuous (lsc.) functions, cf. [27]. The main step is to prove that the empirical objective functions $F(x, P_\nu)$ epi-converge to the true objective function $F(x, P_0)$, which in turn implies convergence results for optimal values and for *sets of optimal solutions*, see [63].

Consistency results are of a limited use for the sought output analysis as they do not provide any information about the rate of convergence of the empirical optimal values to the true one. For empirical distributions, Central Limit Theorem implies pointwise asymptotic normality of random objectives provided that the variance σ_x^2 of $f(x,\omega)$ exists. It means that

$$\sqrt{\nu}\left(\frac{1}{\nu}\sum_{1=1}^{\nu} f(x,\omega_i) - Ef(x,\omega)\right) \rightarrow^d \mathcal{N}(0,\sigma_x^2), \tag{4}$$

that is $\frac{1}{\nu}\sum_{i=1}^{\nu} f(x,\omega_i)$ is asymptotically normal $\mathcal{N}(Ef(x,\omega), \frac{\sigma_x^2}{\nu})$. The accuracy of the normal approximation can be estimated via the Berry-Esseén theorem provided that there exist higher moments of $f(x,\omega)$. See [69] and an application in [14].

If the optimal solution of the true problem (1) is unique and the objective function is Lipschitz continuous, also asymptotic normality of the empirical optimal values $\varphi(P_\nu)$ can be proved; see e.g. Theorem 5.7 of [72]. Such type of results allows to construct approximate confidence intervals for the true optimal value.

In the presence of constraints $x \in \mathcal{X}$ asymptotic normality of empirical optimal solutions $x^*(P_\nu)$ cannot be in general expected even when all solution sets $\mathcal{X}^*(P)$ and $\mathcal{X}^*(P_\nu)\,\forall\nu$ are singletons. It is possible to prove that under reasonable assumptions, the asymptotic distribution of unique consistent empirical optimal solutions $x^*(P_\nu)$ is conically normal being projection of normal distribution on a convex cone. Additional assumptions are needed to get asymptotic normality. It may hold true when the true optimal solution $x^*(P_0)$ is an interior point of $\mathcal{X}$ or when the problem reduces on a neighborhood of the true optimal solution $x^*(P_0)$ to one with affine constraints. This occurs e.g. when $\mathcal{X}$ is a convex polyhedron with nondegenerated vertices and the strict complementarity conditions are valid at the true solution; see e.g. [16] and additional references in [47].

Another field of interest is a relaxation of the assumed independence of sampled data connected with the assumed simple random sampling, see e.g. [47] and references therein, and of the existence of second order moments or of the moment generating function which would allow to get asymptotics also for the heavy tailed distributions; see [47, 75].

3.3. Non-asymptotic results and small sample asymptotics

Asymptotic results support the idea that reliable approximate solutions can be obtained via a tractable numerical procedure assuming that the probability distribution is known and an arbitrarily large sample can be used. The required sample size, however, is not indicated. Hence, non-asymptotic confidence bounds valid for any sample size are of interest, namely in situations when the probability distribution is not known and just a finite sample from this distribution is available. They can be based on various probability inequalities. One can use inequalities for differences of the true and empirical expectations (e.g., Hoeffding inequality applied already in [46] or more general results by [59, 61] applied in [62]) or to rely on the quantitative stability results combined with bounds on the distance $d(P_0, P_\nu)$ of the true and the empirical probability distribution, see e.g. [37].

For example, using the Chernoff bound, exponential rate of convergence of consistent empirical objective functions $F(x, P_\nu)$ to $F(x, P_0)$ holds true: For arbitrary real numbers a, t,

$$P(F(x, P_\nu) - F(x, P_0) \geq a) \leq e^{-ta}[M(\frac{t}{\nu})]^\nu$$

provided that the moment generating function $M(t)$ of deviations $f(x, \omega) - F(x, P_0)$ is bounded. Under additional assumptions, exponential convergence

$$P(F(x^*(P_\nu), P_0) - \varphi(P_0) \geq \epsilon) \leq \alpha e^{-\beta\nu} \tag{5}$$

can be proved and extended also to exponential convergence for deviations of the empirical optimal values from the true optimum and deviations of unique minimizers, cf. [10]. A general possibility is to apply the Large Deviations Theorems, cf. [45, 61, 70, 71]. However, when using the inequalities based on the large deviation theory one accepts the assumption that the moment generating function exists. It means that their straightforward application to heavy tailed distributions may be misleading.

The above results provide information about the quality of the already obtained solution and may support construction of stopping rules or indicate the need for larger sample sizes, cf. [70, 71].

Instead of increasing the sample size one can try to decrease the error in the asymptotically normal approximation (4) using higher order expansions, e.g. Edgeworth expansions or small sample asymptotics, cf. [34]. Quite precise results valid also for very small sample sizes have been reported for approximations of expectations and for M-estimators. Optimal solutions of (1) are related with M-estimators, hence, there is an open possibility of application of the small sample asymptotics results within stability and output analysis of (1). However, it will be necessary to deal with presence of constraints in (1) and with the large dimensionality of the decision vector x, etc.

3.4. Parametric families

Assume now that the true probability distribution P_0 is known to belong to a parametric family $\mathcal{P} = \{P_\theta\}$ of probability distributions, indexed by a parameter vector θ belonging to an open set $\Theta \subset \mathbb{R}^p$. It means that stability and sensitivity analysis of (1) with respect to parameters can be treated via techniques of parametric programming; this was done e.g. in the early paper [2]. In the context of stochastic programming the true parameter values are often estimated from sample data, and the estimates θ_ν usually enjoy quite convenient statistical properties such as consistency or asymptotic normality. If the optimal value function is continuous in θ on a neighborhood of the true parameter value θ_0 properties of transformed random sequences, cf. [69], imply consistency of the estimated optimal value $\varphi(\theta_\nu)$. Asymptotic normality of $\varphi(\theta_\nu)$ follows by δ-theorem if φ is continuously differentiable at the true parameter value θ_0 with non-zero gradient. See [14] for details.

An example where an explicit manageable form of the optimal value function φ can be obtained is maximization of expected utility function $1 - \exp\{-a.\omega^\top x\}$ with

$a > 0$ and normal $\mathcal{N}(\theta, \Sigma)$ distribution of ω over a convex polyhedral set $\mathcal{X}$. The resulting minimization problem of the form (1) is the quadratic program

$$\min_{x \in \mathcal{X}} -\frac{1}{a}\theta^\top x + \frac{1}{2}x^\top \Sigma x \tag{6}$$

with coefficients θ, Σ. If Σ is assumed known, the optimal value $\varphi(\theta)$ is continuous in θ, hence for consistent θ_ν consistency of optimal values of $\varphi(\theta_\nu)$ holds true.

Such explicit results are rare and the common procedure is to estimate θ_0 as $\hat{\theta}$ and to solve (1) with the objective function $F(x, P_{\hat{\theta}})$ by a usual discretization technique. We refer to [66] for discussion of assumptions under which such approach is legitimate. They include consistency of the estimate and also continuity properties of the random objective function $f(x, \omega)$ and of densities $g(\omega, \theta)$ of the distributions $P_\theta, \theta \in \Theta$. A natural question is how much the sampling error interferes with the error due to estimation. The impact of the applied estimation procedure on the results is discussed in [55].

An alternative is to apply an adaptive Bayesian approach which exploits a sample information about the true probability distribution P_{θ_0}. It was suggested in [42] and elaborated for a special type of problem (1).

3.5. Sample information about the true probability distribution

If the data is homogeneous enough it can be used directly for constructing an empirical probability distribution P_ν based on a sample of size ν, the corresponding empirical objective function $F(x, P_\nu)$, the set of empirical optimal solutions and the optimal empirical value $\varphi(P_\nu)$. For large sample sizes it is possible to think about sampling from a finite population and to apply asymptotic results as in 3.4.

For small sample sizes one can analyse robustness of results with respect to possible changes in scenarios, their probabilities or inclusion of additional scenarios by the contamination method delineated in section 3.1. Inclusion of an additional scenario ω^* corresponds to degenerated distribution $\overline{P}$ concentrated in ω^*, parameter $\lambda \in (0, 1)$ reflects the importance of the additional scenario; see [20, 25].

4. Hints for extensions

Scenario generation is a crucial task for applications of stochastic decision models and we listed only selected papers dealing with scenario generation procedures; see [18] for additional references. Papers dealing with approximation of the expectation in (1) by means of various bounding schemes were not included; consult e.g. Chapter 8 of [5] and references therein.

The output analysis that should follow applications of scenario based models is rather involved. It cannot be reduced only to ad hoc numerical experiments and backtesting for a specific application. The general ideas were explained here for static expectation type stochastic programs with the set $\mathcal{X}$ of feasible decisions independent of the assumed probability distribution. Possible extension of the results to multistage problems, to sets of feasible decisions depending on the probability distribution and to risk objectives is at present an active area of research. Another

field of interest concerns the probabilistic assumptions where the focus is on relaxation of the independence assumption, on developing output analysis techniques valid for heavy tailed distributions and also on attempts to get non-asymptotic output analysis results valid for small samples. The interplay of the theoretical results and requirements of real-life applications is evident and plays an important role.

References

[1] Abaffy, J., Bertocchi,, M., Dupačová, J., and Moriggia, V.: On generating scenarios for bond portfolios, *Bull. Czech Econometric Soc.* **11** (2000), 3–37.

[2] Armacost, R. L. and Fiacco, A. V.: Computational experience in sensitivity analysis for nonlinear programming, *Math. Progr.* **6** (1974), 301–326.

[3] Beale, E.: On minimizing a convex function subject to linear inequalities, *JRSS Ser. B* **17** (1955), 173-184.

[4] Beltratti, A., Consiglio, A., and Zenios, S. A.: Scenario modeling in the management of international bond portfolios, *Annals of Oper. Res.* **85** (1999), 227–247.

[5] Birge, J. R. and Louveaux, F.: *Introduction to Stochastic Programming,* 2nd edition, Springer, New York, 2011.

[6] Birge, J. R. and Rosa, C. H.: Modeling investment uncertainty in the costs of global CO_2 emission policy, *EJOR* **83** (1995), 466–488.

[7] Birge, J. R. and Wets, R. J.-B.: Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse, *Math. Progr. Study* **27** (1986), 54–102.

[8] Cariño, D. R., Myers D. H., and Ziemba, W. T.: Concepts, technical issues, and uses of the Russell - Yasuda Kasai financial planning model, *Oper. Res.* **46** (1998), 450–462.

[9] Conejo, A. J., Carrión, M., and Morales, J. M.: *Decision Making under Uncertainty in Electricity Markets,* Springer, New York, 2010.

[10] Dai, L., Chen, C. H., and Birge, J.L.: Convergence properties of two-stage stochastic programming, *J. Optim. Theory Appl.* **106** (2000), 489–509.

[11] Dantzig, G. B. and Infanger, G.: Multi-stage stochastic linear programs for portfolio optimization, *Annals of Oper. Res.* **45** (1993), 59–76.

[12] Dantzig, G. B. and Infanger, G.: Approaches to stochastic programming with application to electric power systems. In: *Optimization in Planning and Operation of Electric Power Systems* (K. Frauendorfer et al., eds.), Physica-Verlag, Heidelberg, 1993, 125–138.

[13] Dembo, R. S.: Scenario optimization, *Annals of Oper. Res.* **30** (1991), 63–80.

[14] Dupačová, J.: Stability in stochastic programming with recourse - Estimated parameters, *Math. Progr.* **28** (1984), 72–83.

[15] Dupačová, J.: Stability and sensitivity analysis for stochastic programming, *Annals of Oper. Res.* **27** (1990), 115–142.

[16] Dupačová, J.: On non-normal asymptotic behavior of optimal solutions for stochastic programming problems and on related problems of mathematical statictics, *Kybernetika* **27** (1991), 38–52.

[17] Dupačová, J.: Scenario based stochastic programs: Resistance with respect to sample, *Annals of Oper. Res.* **64** (1996), 21–38.

[18] Dupačová, J.: Stochastic programming: Approximation via scenarios, *Aportaciones Matematicás, Ser. Communicaciones* (1998), 77–94.

[19] Dupačová, J.: Notes on asymptotic properties of approximated stochastic programs, *Bull. Czech Econometric Soc.* **19** (2012), 112–123.

[20] Dupačová, J., Bertocchi, M., and Moriggia, V.: Postoptimality for scenario based financial models with an application to bond portfolio management. In: *World Wide Asset and Liability Modeling* (W. Ziemba and J. Mulvey, eds.), Cambridge Univ. Press 1998, pp. 263–285.

[21] Dupačová, J., Consigli, G., and Wallace, S. W.: Scenarios for multistage stochastic programs, *Annals of Oper. Res.* **100** (2000), 25–53.

[22] Dupačová, J., Gröwe-Kuska, N., and Römisch, W.: Scenario reduction in stochastic programming. An approach using probability metrics, *Math. Progr.* **Ser. A 95** (2003) 493–511.

[23] Dupačová, J., Hurt, J., and Štěpán, J.: *Stochastic Modeling in Economics and Finance*, Kluwer, Dordrecht, 2002.

[24] Dupačová, J. and Polívka, J.: Asset-liability mamagement for Czech pension funds using stochastic programming, *Annals of Oper. Res.* **165** (2009), 5–28.

[25] Dupačová, J. and Polívka, J.: Stress testing for VaR and CVaR, *Quantitative Finance* **7** (2007), 411–421.

[26] Dupačová, J. and Popela, P.: Melt control: Charge optimization via stochastic programming. In: *Applications of Stochastic Programming*, (S. W. Wallace and W. T. Ziemba, eds.), MPS-SIAM Series in Optimization, SIAM, Philadelphia, 2005, pp. 277–297.

[27] Dupačová, J. and Wets, R. J.-B.: Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems, *Ann. Stat.* **16** (1988), 1517–1549.

[28] Eichhorn, A., Römisch, W., and Wegner, I.: Mean-risk optimization of electricity portfolios using multiperiod polyhedral risk measures. In: *IEEE St. Petersburg Power. Tech. Proceedings*, 2004.

[29] Eppen, G. D., Martin, R. K., and Schrage, L.: A scenario approach to capacity planning, *Oper. Res.* **37** (1989), 517–527.

[30] Ermoliev, Yu. and Wets, R. J.-B. (eds.): *Numerical Techniques for Stochastic Optimization Problems*, Springer, Berlin, 1988.

[31] Escudero, L. F., Kamesam, P. V., King, A. J., and Wets, R. J.-B.: Production planning via scenario modeling, *Annals of Oper. Res.* **43** (1993), 301–335.

[32] Feng, Y. and Ryan, S. M.: Scenario construction and reduction applied to stochastic power generation expansion planning, *Computers & Operations Research* **40** (2013), 9–23.

[33] Ferguson, A. and Dantzig, G. B.: The allocation of aircraft to routes: An example of linear programming under uncertain demand, *Manag. Sci.* **3** (1956), 45–73.

[34] Field, Ch. and Ronchetti, E.: *Small Sample Asymptotics*, LN – Monograph series of Inst. of Math. Statistics **13**, IMS, Hayward, California, 1990.

[35] Gröwe, N. and Römisch, W.: A stochastic programming model for optimal power dispatch: Stability and numerical treatment. In: *Stochastic Optimization - Numerical Methods and Technical Applications* (K. Marti, ed.), LN in Economics and Math. Systems Vol. 379, Springer, Berlin, 1992, pp. 111–139.

[36] Gröwe, N., Römisch, W., and Schultz, R.: A simple recourse model for power dispatch under uncertain demand, *Annals of Oper. Res.* **59** (1995), 135–164.

[37] Gülten, S. and Ruszczyński, A.: Two-stage portfolio optimization with higher-order conditional measures of risk, *Annals of Oper. Res.* to appear.

[38] Hiller, R. S. and Schaack, Ch.: A classification of structured bond portfolio modeling techniques, *J. of Portfolio Management*, Fall 1990, 37–48.

[39] Hoyland, A., Kaut, M., and Wallace, S. W.: A heuristic for moment-matching scenario generation, *Optimization and Applications* **24** (2003), 169–185.

[40] Infanger, G.: Monte Carlo (Importance) sampling within a Benders decomposition algorithm for stochastic linear programs, *Annals of Oper. Res.* **39** (1992), 69–95.

[41] Jacobs, J., Freeman, G., Grygier, J., Morton, D., Schultz, G., Staschus, K., and Stedinger, J.: SOCRATES: A system for scheduling hydroelectric generation under uncertainty, *Annals of Oper. Res.* **59** (1995), 99–133.

[42] Jagannathan, R.: Use of sample information in stochastic recourse and chance-constrained programming problems, *Manag. Sci.* **31** (1985), 96–108.

[43] Jamshidian, F. and Yu Zhu: Scenario simulation: Theory and methodology, *Finance and Stochastics* **1** (1997), 43-67.

[44] Kall, P., Ruszczyński, A., and Frauendorfer, K.: Approximation techniques in stochastic programming. In: [30], pp. 33–64.

[45] Kaniovski, Y. M., King, A. J., and Wets, R. J.-B.: Probabilistic bounds (via large deviations) for the solutions of stochastic programming problems, *Annals of Oper. Res.* **56** (1995), 189–208.

[46] Kaňková, V.: An approximate solution of a stochastic optimization problem. In: *Trans. 8th Prague Conf.*, Academia, Prague, 1978, pp. 349–353.

[47] Kaňková, V.: Empirical estimates in optimization problems via heavy tails and dependent samples, *Bull. Czech Econometric Soc.* **19** (2012), 92–111.

[48] Kaut, M. and Wallace, S. W.: Shape-based scenario generation using copulas, *CMS* **8** (2011), 181–199.

[49] King, A. J.: Stochastic programming problems: Examples from the literature. In [30], pp. 543–565.

[50] Koskosides, Y. A. and Duarte, A. M.: A scenario-based approach to active asset allocation, *J. of Portfolio Management*, Winter 1997, 74–85.

[51] Kouwenberg, R.: Scenario generation and stochastic programming models for asset liability management, *EJOR* **134** (2001), 279–292.

[52] Kozmík, V. and Morton, D.: Evaluating policies in risk-averse multi-stage stochastic programming, *Math. Progr.* submitted 2013.

[53] Mulvey, J. M.: Generating scenarios for the Towers Perrin investment system, *Interfaces* **26.2** (1996), 1–15.

[54] Mulvey, J. M. and Vladimirou, H.: Stochastic network optimization models for investment planning, *Annals of Oper. Res.* **20** (1989), 187–217.

[55] Mulvey, J. M., Rosenbaum, D. P., and Shetty, B.: Parameter estimation in stochastic scenario generation systems, *EJOR* **118** (1999), 563–577.

[56] Mulvey, J. M. and Zenios, S. A.: Capturing the correlations of fixed-income instruments, *Manag. Sci.* **40** (1994), 1329–1342.

[57] Nielsen, S. S.: Importance sampling in lattice pricing models. Chapter 12 in: *Interfaces in Computer Science and Operations Research: Advances in Metaheuristics, Optimization and Stochastic Modeling Technologies* (R. S. Barr, R. V. Helgason, and J. Kennington, eds.), Kluwer, Dordrecht, 1997.

[58] Pereira, M. V. F. and Pinto, L. M. V. G.: Multi-stage stochastic optimization applied to energy planning, *Math. Progr.* **52** (1991), 359–375.

[59] Pflug, G. Ch.: Stochastic programs and statistical data, *Annals of Oper. Res.* **85** (1999), 59–78.

[60] Pflug, G. Ch.: Scenario tree generation for multiperiod financial optimization by optimal discretization, *Math. Progr.* **89** (2001), 251–271.

[61] Pflug, G. Ch.: Stochastic optimization and statistical inference. Chapter 7 in [67], pp. 427–482.

[62] Pflug, G. Ch. and Pichler, A.: Approximations for probability distributions and stochastic optimization problems. In: *Stochastic Optimization Methods in Finance and Energy* (M. Bertocchi, G. Consigli, and M. A. H. Dempster, eds.), Springer, New York, 2011, pp. 343-387.

[63] Rockafellar, R. T. and Wets, R. J.-B.: *Variational Analysis*, Springer, New York, 1997.

[64] Römisch, W.: Stability of stochastic programming problems. Chapter 8 in [67], pp. 483–554.

[65] Römisch, W.: Scenario generation in stochastic programming. In: *Wiley Encyclopedia on Operations Research and Management Science* (J. J. Cochran, ed.), Wiley, Hoboken, 2010.

[66] Rubinstein, R. Y. and Shapiro, A.: *Discrete Event Systems. Sensitivity Analysis and Stochastic Optimization by the Score Function Method*, Wiley, Chichester, 1993.

[67] Ruszczyński, A. and Shapiro, A. (eds.): *Stochastic Programming.* Handbooks in OR & MS **10**, Elsevier, Amsterdam, 2003.

[68] Sen, S., Doverspike, R. D., and Cosares, S.: Network planning with random demand, *Telecommunications Systems* **3** (1994), 11–30.

[69] Serfling, R. J.: *Approximation Theorems in Mathematical Statistics,* Wiley, New York, 1980.

[70] Shapiro, A.: Monte Carlo Sampling Methods. Chapter 6 in: [67], pp. 353–425.

[71] Shapiro, A. and Homem-de-Mello, T.: On rate of convergence of optimal solutions of Monte Carlo approximations of stochastic programs, *SIAM J. Optimization* **11** (2001), 70–86.

[72] Shapiro, A., Dentcheva, D., and Ruszczyński, A.: *Lectures on Stochastic Programming: Modeling and Theory*, SIAM-Society for Industrial and Applied Mathematics, Philadelphia, 2009.

[73] Shapiro, J. F.: Stochastic programming models for dedicated portfolio selection. In: *Mathematical Models for Decision Support* (B. Mitra, ed.), NATO ASI Series, **Vol. F48**, Springer, Berlin, 1988, pp. 587–611.

[74] Somlyódy, L. and Wets, R. J.-B.: Stochastic optimization models for lake eurotrophication management, *Oper. Res.* **36** (1988), 660–681.

[75] Tokat, Y., Rachev, S. F., and Schwartz, E. S.: The stable non-Gaussian asset allocation: a comparison with the classical Gaussian approach, *J. of Economic Dynamics and Control* **27** (2003), 937–969.

[76] Wets, R. J.-B.: Statistical estimation from an optimization viewpoint, *Annals of Oper. Res.* **85** (1999), 79–101.

[77] Ziemba, W. T.: *The Stochastic Programming Approach to Asset, Liability, and Wealth Management,* The Research Foundation of the Association for Investment Management and Research, 2003.