ORIGINAL PAPER

CrossMark

Decision-dependent probabilities in stochastic programs with recourse

Lars Hellemo[1] · Paul I. Barton[2] · Asgeir Tomasgard[3]

Abstract
Stochastic programming with recourse usually assumes uncertainty to be exogenous. Our work presents modelling and application of decision-dependent uncertainty in mathematical programming including a taxonomy of stochastic programming recourse models with decision-dependent uncertainty. The work includes several ways of incorporating direct or indirect manipulation of underlying probability distributions through decision variables in two-stage stochastic programming problems. Two-stage models are formulated where prior probabilities are distorted through an affine transformation or combined using a convex combination of several probability distributions. Additionally, we present models where the parameters of the probability distribution are first-stage decision variables. The probability distributions are either incorporated in the model using the exact expression or by using a rational approximation. Test instances for each formulation are solved with a commercial solver, BARON, using selective branching.

1 Introduction

Most practical decision problems involve uncertainty at some level, and stochastic programming was introduced by Dantzig (1955) and Beale (1955) to handle uncertain parameters in optimization models. Their approach was to model a discrete time

✉ Asgeir Tomasgard
asgeir.tomasgard@ntnu.no

Lars Hellemo
Lars.Hellemo@sintef.no

[1] Department of Economics and Technology Management, SINTEF Technology and Society, S. P. Andersens veg 5, 7031 Trondheim, Norway

[2] Process Systems Engineering Laboratory Department of Chemical Engineering, Massachusetts Institute of Technology, 66-470b, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[3] Department of Industrial Economics and Technology Management, NTNU, Alfred Getz vei 3, Trondheim, Norway

Springer

decision process where uncertain parameters are represented by scenarios and their respective probabilities. In a scenario-based stochastic program, decisions are made, and uncertain values are revealed at discrete points in time. Some decisions are made before the actual values of uncertain parameters are known, but the realization of the stochastic parameters is independent of the decisions. This framework will later be referred to as stochastic programs with exogenous uncertainty or stochastic programming with decision-independent uncertainty. In recent years stochastic programs with endogenous uncertainty or decision-dependent uncertainty have received increased attention. Some early examples of papers with decision-dependent uncertainty are Jonsbråten (1998), Jonsbråten et al. (1998) and Goel and Grossmann (2004). The terms decision-dependent uncertainty and endogenous uncertainty are used interchangeably.

The main contribution of our paper is to provide new formulations for endogenous stochastic programming models where the probabilities of future events depend on decision variables in the optimization model, in the following called stochastic programs with decision-dependent probabilities. This is a subclass of endogenous stochastic programming models that has received little attention in the literature. There are some examples in the existing literature of problems where a decision may shift from one predefined set of probabilities to another. To the best of our knowledge, there are no examples in the literature where the relation is modeled as a continuous function. In Sect. 2 a more thorough description of problem classes with endogenous uncertainty is presented together with the choices a problem owner or modeler needs to make. An extended taxonomy for stochastic programs with endogenous uncertainty and a literature review is suggested in Sect. 3. New formulations for models with decision-dependent probabilities are found in Sect. 4. Several test instances of models using these formulations are presented in Sect. 5. Computational results follow in Sect. 6.1 and conclusions in Sect. 7.

2 Decision problems with decision-dependent uncertainty

To discuss the concept of decision-dependent uncertainty, it is useful to first make distinctions between the real world, the description of the real world presented to the modeler as a problem and the actual mathematical model formulation. A problem description belongs to one of these classes:

Deterministic problems are problems where there is no substantial uncertainty, there may for example be available precise measurements of all parameters, or there may be some official values available, such as the prices for today's operations.

Exogenous uncertainty problems are problems with substantial uncertainty, where the distribution of the stochastic parameters is known, for example based on historical data or expert opinion. The information structure and the probability distributions do not depend on any decisions in the model. Rather, the model will seek a solution that does well in expectation. Some models also include different risk attitudes or use a risk measure.

Endogenous uncertainty problems are problems where decisions at one point in time will have a substantial impact on the uncertainty faced later, either in terms of when information about the actual value of a stochastic parameter becomes available, or the
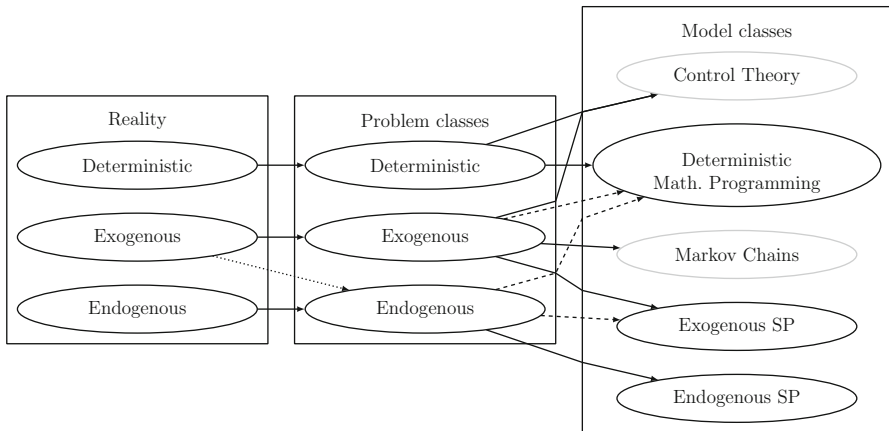
Fig. 1 Classification problem and model classes. Stochastic programming models in black, examples of other modelling paradigms in grey. Dashed line indicates problem relaxation, whereas the full stroke indicates a mapping

probability that a certain realization of a parameter occurs. A problem is classified as having endogenous uncertainty when decisions that are part of the problem to be solved, influences the uncertainty of parameters that are also part of the problem.

Note that there is not a one to one mapping between reality and the problem description or between the problem and the model choice. Figure 1 shows some alternative mappings. In the following, this is illustrated with some examples.

First, consider a river where a dam is to be built and the design parameters of the dam are to be determined. The risk of a dam break must be balanced against the extra cost of further reinforcing it. The stochastic inflow is not influenced by the way the dam is built, rather the dam's resistance to various inflows is. In this case, a problem description may focus on the stochastic inflow, and describe this as a design problem with exogenous uncertainty. The risk of a dam failure would depend on the stochastic inflow, but the design decision would not affect the stochastic parameter in the model. Alternatively, the decision-maker could decide to model the probability of a dam break by linking the uncertainty description to the dam design, making it decision-dependent.

Next consider a petroleum reservoir where there is some uncertainty about the properties of the reservoir, and the decisions are the technology used for drilling wells, where to drill wells, as well as when the wells should be drilled if drilled at all. The actual petroleum content of the reservoir is fixed, but not known precisely. The decision to drill test wells does not change the content of the reservoir as such, but it may provide more information about the reservoir. The information process of a problem is described by the combination of how uncertainty is resolved, and the sequences of decisions made as a response to that. In this case, this information is not revealed unless the owner drills the test wells, which incurs a substantial cost. This is a situation where the underlying reservoir content is deterministic, but unknown to the decision-maker. The decisions affect when information is revealed and what information is revealed. This calls for a model that handles decision-dependent uncer-

tainty, even if the underlying reservoir size is deterministic. In the same reservoir case, the choice between alternative drilling technologies is another, similar, consideration. Some drilling approaches may jeopardize the reservoir itself by introducing leaks between layers in the ground, something that could render part of the resources unrecoverable. In this way, our decisions may change the recoverable volume from said reservoir. Note that now it is not only the information revelation that leads to a decision-dependent uncertainty formulation, now there is another uncertain variable also depending on the decision: the drilling hazard. For this oil reservoir with fixed but unknown petroleum content we may choose to ignore the decision-dependent part of the information structure. The resulting model is a traditional stochastic program with decision-independent recourse, but missing important parts of the decision-maker's problem.

The right-hand part of Fig. 1 shows examples of model classes. Moving on to formulating a specific model to aid the solution of a certain problem, some relaxations or approximations will usually have to be made, often to reduce cognitive load of model users or to improve computational tractability, or both. While this work focuses on stochastic programming, also other modeling paradigms exist such as control theory, game theory and several others that may be considered for stochastic problems. The following literature review and taxonomy is limited to include problems described with endogenous uncertainty and where the model choice is stochastic programming with recourse.

3 Taxonomy

This section presents a taxonomy and literature review for stochastic programs with decision-dependent uncertainty. Our taxonomy expands previously presented classifications of such problems and is summarized in Fig. 2.
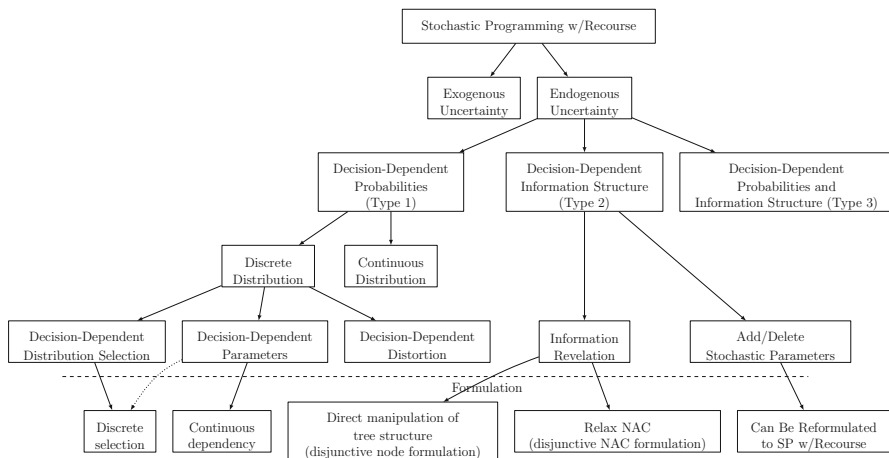

Fig. 2 Classification of endogenous SPs

The literature on endogenous uncertainty in stochastic programming is sparse. This should come as no surprise as one quickly departs from the domains where well performing solution techniques are available, notably for convex programming in general and linear programming in particular, as noted by Varaiya and Wets (1989). Jonsbråten (1998) and Jonsbråten et al. (1998) proposed a generalized formulation of stochastic programs with recourse of which the standard SP is a special case (Eq. 1), and suggested the classification of stochastic programs into two subclasses: endogenous and exogenous uncertainty.

$$\min \ E^p \ f(x) = \int_{\varXi} f(\xi; x) p(d\xi) \text{ such that } (p, x) \in \mathcal{K} \subset \mathcal{P} \times \mathbb{R}^N \qquad (1)$$

$\mathcal{P}$ is a subset of the probability measures on $\varXi$ and $\mathcal{K}$ are the constraints linking the decision x to the choice of p.

The problems discussed in the paper by Jonsbråten et al., concern situations where the time at which information becomes available is determined by the decisions in preceding stages. As an example, they use stochastic production costs. Only after making the decision of which product to make, is the uncertainty of this particular product revealed. The other possible products' true costs remain hidden (stochastic) until a decision to produce them is made.

Several authors (Dupačová 2006; Tarhan et al. 2009) identify two subclasses within endogenous stochastic programs. The first class of problems is where the probabilities are decision-dependent, denoted as *Decision-Dependent Probabilities* or Type 1. Problems with decision-dependent probabilities are discussed further in Sect. 3.2. Equation (2) further generalizes Eq. (1) to include the possibility that the probability measure also depends on x:

$$\min \ E^p \ f(x) = \int_{\varXi} f(\xi; x) p(x; d\xi) \text{ such that } (p, x) \in \mathcal{K} \subset \mathcal{P} \times \mathbb{R}^N \qquad (2)$$

The other subclass is denoted Type 2, and concerns models where the time of the information revelation is decision-dependent. That means that decision variables are used to make realization of uncertain variables known earlier in time, as in buying information or drilling a well. Often type 2 problems are called problems with Decision-Dependent Information Structure. They are discussed further in Sect. 3.1.

Some problems may have both kinds of decision-dependent uncertainty, and we suggest adding a Type 3 to the taxonomy to include such problems. To the best of our knowledge, problems of Type 3 have not yet been discussed in the literature. For an overview of the different problem classes and their subclasses, see Fig. 2.

3.1 Decision-dependent information structure

By decision-dependent information structure we mean all ways of altering the time dynamics of a stochastic program. This includes the time of information revelation, as in endogenous problems of Type 2, as well as the addition of stochastic parameters, and deletion of stochastic parameters. Another example is problems for which the time

when uncertainty is redefined/refined is a decision variable, such as in using sensors or in acquisition of information. This category includes all stochastic programs with endogenous uncertainty were nonanticipativity constraints (NAC) can be manipulated by decision variables, whereas the probabilities remain fixed.

3.1.1 Information revelation

The subcategory of information revelation has received most attention in the literature, following Jonsbråten (1998), Jonsbråten et al. (1998) and Goel and Grossmann (2004). The most used technique is to relax the nonanticipativity constraints of a stochastic program, allowing selection of the times of branching of the tree (when scenarios become distinguishable), see discussion below.

Goel and Grossmann (2004) formulated a model for development of natural gas resources where the time of exploitation can be selected in the model. This introduces endogenous uncertainty as the information revelation depends on which wells are drilled and when, and it is formulated as a disjunctive programming problem where the nonanticipativity constraints depend on the decision variables related to drilling. They first considered a model with pure decision-dependent uncertainty, and later generalized it to a hybrid model including both endogenous and exogenous uncertainty (Goel and Grossmann 2006). This form of endogenous uncertainty arises in multi-stage models, where the decisions to explore a field unravels the true parameter values of the field that is explored, but not the others. As this decision can be made at different times (stages), it is only relevant in a multi-stage environment. Effectively their approach is a model with decision-dependent nonanticipativity constraints, and they develop several theoretical results demonstrating redundancy in the constraints and that the number of nonanticipativity constraints can be reduced accordingly. This improves the practicality of the model by making it more readily solvable. The models are still quite large, though, and they propose a branch and bound solution procedure based on Lagrangian duality.

Solak (2007) presents portfolio optimization problems where the timings of the realizations are dependent on the decisions to invest in the projects. The application is from R&D in the aviation industry where a technology development portfolio is to be optimized. Solak introduces gradual resolution of uncertainty, where the amount invested in a project increases the resolution of the uncertainty regarding that project up to a point where all uncertainty has been resolved. The author proposes solution approaches for the multi-stage stochastic integer programming model with focus on decomposability, sample average approximation and Lagrangian relaxation with lower bounding heuristics.

A model with gradual resolution of information is also presented by Tarhan et al. (2009), another petroleum application with a multi-stage non-convex stochastic model, solved by a duality-based branch and bound method. In a series of papers Colvin and Maravelias (2008), Colvin (2009a) and Colvin and Maravelias (2010) study a stochastic programming approach for clinical trial planning in new drug development, where information revelation depends on decisions. Colvin and Maravelias (2009b) build on the work by Goel and Grossmann. They further improve on a reformulation with redundant nonanticipativity constraints removed and observe that few of the remaining

are binding. They add the constraints only as needed through a customized branch-and-cut algorithm. The model is formulated as a pure MIP. Solak et al. (2010) deal with a project portfolio management problem for selecting and allocating resources to research and development (R&D) projects to design, test and improve a technology, or the process of building a technology.

Boland et al. (2008) also build on the work of Goel and Grossman in their open pit mining application where geological properties of the mining blocks (quality) varies, and there is a mix of already mined blocks, and blocks where the quality is uncertain until the point of development. They find that they can reuse existing variables for nonanticipativity constraints and thus reduce the size of the problem. They exploit the problem structure to omit a significant proportion of the nonanticipativity constraints. Boland et al. implemented a version of their model with "lazy" constraints but found that this did not improve performance for their model instances.

Peeta et al. (2010) address a pre-disaster planning problem that seeks to strengthen a highway network whose links are subject to random failures due to a disaster. Each link may be either operational or non-functional after the disaster. The link failure probabilities are assumed to be known a priori, and investment decreases the likelihood of failure. Escudero et al. (2016) examine a resource allocation model and an algorithmic approach for a three-stage stochastic problem related to managing natural disasters mitigation. The endogenous uncertainty is based on the investment, for getting a better accuracy on the disaster occurrence.

A later improvement on the work by Goel and Grossman is by Gupta and Grossmann (2011), and they also propose new methods for obtaining a more compact representation of the nonanticipativity constraints. In addition, they propose three solution procedures. One is based on a relaxation of the problem in what they call a k-stage constraints problem, where only nonanticipativity constraints for a given number of stages are included. Secondly, they propose an iterative procedure for nonanticipativity constraint relaxation, and third they present a Lagrangian decomposition algorithm. The application is the same as in Goel and Grossmann (2006). Apap and Grossmann (2017) discuss formulations and solution approaches for stochastic programs with Decision-Dependent Information Structure.

An alternative and equivalent way of formulating stochastic programming problems with recourse is using a node formulation of the scenario tree. As an alternative to the disjunctive nonanticipativity constraints (NAC) formulation with relaxation of NAC, problems with decision-dependent information revelation may be formulated using a disjunctive node formulation. However, to our knowledge, such a model has never been presented in the literature.

In an early paper, Artstein and Wets (1994) present a model where a decision-maker can seek more information through a sensor, in a model that allow a redefinition of the probability distribution used in the stochastic program. This refines the decision process in that it acknowledges that the inquiry process may itself introduce errors. They solve an example based on a variant of the newsboy problem where the newsboy may perform a poll/sampling to gain information about the probability distribution, possibly at a cost. They provide a general approach to the situation when the underlying uncertainty is not known, and decisions may influence the accuracy of the uncertainty in a stochastic program.

3.1.2 Problems that may be reformulated as ordinary SP

In addition to problems with decision-dependent information revelation, other structures are conceivable, that may be reformulated as stochastic programs with recourse. This includes deleting stochastic variables, adding stochastic variables, and modifying the support. This may be achieved using binary variables. For a recent example, see Ntaimo et al. (2012) where a two-stage stochastic program for wildfire initial attacks is presented. The cost incurred by each wildfire is one of two possible outcomes for each scenario, depending on whether the fire can be contained through an effective attack or not. The model is formulated as a two-stage stochastic (integer) program with recourse, with binary variables to select which set of recourse costs is incurred in stage two based on the selection of attack means available as a consequence of decisions in stage one. The scenarios are based on fire simulations, giving a large number of scenarios. The model size is reduced by applying sample average approximation (SAA).

3.2 Decision-dependent probabilities

The first attempt to model explicitly the relationship between the probability measure and the decision variable was made by Ahmed (2000). He formulates single-stage stochastic programs that are applied to network design, server selection and p-choice facility location. Ahmed uses Luce's choice axiom to develop an expression for the probability that, e.g., a path is used, and this probability depends on the design variables of the network. The resulting model is 0–1 hyperbolic program, which he solves by a binary reformulation and by genetic programming in addition to a customized branch and bound algorithm.

For some problems with decision-dependent probabilities, the decision dependency may be removed through an appropriate transformation of the probability measure, which is called the push-in technique by Rubinstein and Shapiro (1993, 214f), see also Pflug (1996, 143ff). Dupačová (2006) notes that in some cases, dependence of distribution $\mathcal{P}$ on decision variable x can be removed by a suitable transformation of the decision-dependent probability distribution (push-in technique).

Escudero et al. (2014) have developed a multi-stage stochastic model including both exogenous and endogenous uncertainty. They also include risk considerations in the form of stochastic dominance constraints. The resulting model is a mixed-integer quadratic program where the weights (probabilities) of each scenario group and/or outcomes of the stochastic parameters may be determined by decision variables from previous stages. To be able to solve large problem instances the authors apply a customized Branch and Fix Coordination (BFC) parallel algorithm.

For the problems in this section, only probabilities depend on the decision variables, while the information structure is fixed. To be specific, nonanticipativity constraints are not manipulated by decision variables. Dupačová (2006) identifies two fundamental classes of problems with endogenous probabilities. One where the probability distribution is known, and the decisions influence the parameters of the probability distribution, the other where some decision will cause the probability distribution to

be chosen between a finite set of probability distributions. We extend her taxonomy with a third category, decision-dependent distribution distortion.

In principle, both discrete and continuous distributions may be considered, where the use of a finite set of scenarios as an approximation also for continuous distributions is the most used method for modeling such problems. The authors are not familiar with any attempts to model and solve problems with decision-dependent probabilities using continuous probability distributions, and in the following only problems with discrete probability distributions, using a finite set of scenarios, are considered.

3.2.1 Decision-dependent distribution selection

Viswanath et al. (2004) consider the design of a robust transportation network where links can be reinforced by investing in additional measures. By investing, the probability of survival of a disruptive event is improved. The model is an investment model with a choice between a finite number of sets of probabilities, typically two, p_e and q_e where p_e is used if there is investment, q_e otherwise. The random variables take values 0 or 1 with probabilities given above. Dupačová (2006) also discusses the subset of problems where available techniques from binary and integer programming can be can be applied to choose between a finite number of set of probability distributions with fixed parameters.

3.2.2 Decision-dependent parameters

Selection between a discrete number of parameter values can be implemented using a generalization of the technique described above. We suggest some models where parameters are continuous decision variables in this work, see Sect. 4. An example of using the exact expression for a probability distribution is shown in Sect. 4.2.1 and a rational approximation in Sect. 4.2.2. We are not aware of any other attempts to include models of Type 1 where the probability distribution parameters can be set continuously.

3.2.3 Distortion

We also include some models where some prior set of probabilities for a distribution with known parameters are distorted. A distortion of these probabilities controlled by decision variables is introduced. This distortion could be applied in form of a transformation of one set of probabilities or by combining several sets of probabilities. Examples of linear transformations in Sect. 4.1 are given, distorting one set of prior probabilities in Sect. 4.1.1 and using the convex combination of several sets of probabilities in Sect. 4.1.2.

The authors are not aware of any other works that present this kind of model, however Dupačová (2006) makes notes on the stability of optimal solutions. She uses probability distribution contamination to investigate the case where a convex combination of several distributions can be applied for convex problems.

3.2.4 Related work

A bit on the side, Held and Woodruff (2005) consider a multi stage stochastic network interdiction problem. The goal is to maximize the probability of sufficient disruption, in terms of maximizing the probability that the minimum path length exceeds a certain value. They present an exact (full enumeration) algorithm and a heuristic solution procedure.

Another approach to uncertainty in optimization is to search for solutions that are robust in the sense that they are good for the most disadvantageous outcomes of the stochastic parameters. Several research groups are working with robust optimization, going back to Ben-Tal et al. (1994), Ben-Tal and Nemirovski (1998), Bertsimas and Sim (2003, 2004). Also, rather than taking a worst-case approach, introducing some ambiguity to the underlying probability distribution has been demonstrated in the works of Pflug and Wozabal (2007) and Pflug and Pichler (2011). In three recent papers, robust optimization is extended to the situation where uncertainty sets are decision-dependent, see Nohadani and Sharma (2016), Lappas and Gounaris (2016, 2018). This approach should be considered as a possible Type 4 decision-dependent uncertainty and is also described for multi-stage robust programs.

Lejeune and Margot (2017) present a static model for aeromedical battlefield evacuation. Endogenous uncertainty is used to make the availability of ambulances depend on their location and of allocation of patients.

Finally, while the optimization over a finite set of scenarios is the dominant approach within stochastic programming, Kuhn (2009) and Kuhn et al. (2011) optimize linear decision rules over a continuous probability distribution.

4 Decision-dependent probabilities

This section presents several formulations of stochastic programs with decision-dependent probabilities. The formulations allow the probabilities of scenarios $s \in \mathcal{S}$ to be altered by some decision variable y, typically a first-stage variable in a two-stage stochastic program. This section considers the case where the function $p_s : \mathbb{R} \to [0, 1]$ is an affine function.

$$\min c_x^\top x + c_y^\top y + \sum_s p_s(y) q_s^\top z_s$$
$$s.t. \quad x \in X, y \in Y, z_s \in Z_s(x, y). \tag{3}$$

4.1 Affine p_s

This formulation does not directly manipulate the parameters of the probability distribution but applies a transformation to one or more predetermined probability distributions. First consider some special cases where the function p_s is an affine function. An affine function is a linear transformation, followed by a translation, i.e. need not be fixed at the origin as with pure linear functions. This is primarily motivated by computational tractability, as it will yield optimization models where, in the

case where the rest of the model is linear, the only nonlinearities are bilinear terms related to variables controlling scenario probabilities. This can easily be generalized to nonlinear transformations and nonlinear stochastic programs.

4.1.1 Linear scaling

Let $s \in \mathcal{S}$ be scenarios, each with probability $p_{0s} > 0, \sum_{s \in \mathcal{S}} p_{0s} = 1$. For each $s \in \hat{\mathcal{S}} \subset \mathcal{S}$ let the variable y scale the probability linearly, whereas the remaining scenarios $s \in \mathcal{S} \backslash \hat{\mathcal{S}}$ are adjusted:

$$p_s(y) = \begin{cases} p_{0s} \, y, & s \in \hat{\mathcal{S}} \subset \mathcal{S} \\ \frac{1 - y \sum_{s' \in \hat{\mathcal{S}}} p_{0s'}}{\sum_{s' \in \mathcal{S} \backslash \hat{\mathcal{S}}} p_{0s'}} \, p_{0s}, & s \in \mathcal{S} \backslash \hat{\mathcal{S}}. \end{cases} \tag{4}$$

In the special case where the original distribution is uniform, this gives the function p_s:

$$p_s(y) = \begin{cases} \frac{1}{|\mathcal{S}|} y, & s \in \hat{\mathcal{S}} \subset \mathcal{S} \\ \frac{1 - y \sum_{\hat{\mathcal{S}}} \frac{1}{|\mathcal{S}|}}{|\mathcal{S}| - |\hat{\mathcal{S}}|}, & s \in \mathcal{S} \backslash \hat{\mathcal{S}}. \end{cases} \tag{5}$$

This model includes bilinear terms $p_s(y)z_s$ in the objective. In addition, in some cases the z may take binary or integer values, for example representing investments. In any case the models are nonlinear and non-convex.

4.1.2 Convex combination of distributions

Let set $\mathcal{I}$ be discrete distributions with probabilities $p_{i,s}, \sum_{s \in \mathcal{S}} p_{i,s} = 1, \forall i \in \mathcal{I}$ associated to each scenario $s \in \mathcal{S}$.

Then define

$$p_s = \sum_{i \in \mathcal{I}} p_{i,s} y_i, \forall s \in \mathcal{S}. \tag{6}$$

A distribution defined like this is often called a mixture distribution, see, e.g., Feller (1943), Behboodian (1970), and Frühwirth-Schnatter (2006). One interpretation would be that the final outcome is selected at random from the underlying distributions, with a certain probability y_i associated with each of them. In our model the mixture weights $y_i \geq 0$ are decision variables, but of course the sum of weights need to be 1. See Fig. 3 for some examples of convex combinations of normal distributions.

Mixture distributions are often used when subsets of the data have specific characteristics, for example where subpopulations exist in a population. Our model then gives the opportunity to influence the weights of the different subpopulations, potentially at a cost.

To reduce the number of y-variables, let one y_u be uniquely determined by the remaining $i \in \mathcal{I} \backslash \{u\}$ such that:

$$p_s = \sum_{i \in \mathcal{I} \backslash \{u\}} p_{i,s} y_i + \left(1 - \sum_{i \in \mathcal{I} \backslash \{u\}} y_i \right) p_{u,s}, u \in \mathcal{I}, \forall s \in \mathcal{S}. \tag{7}$$
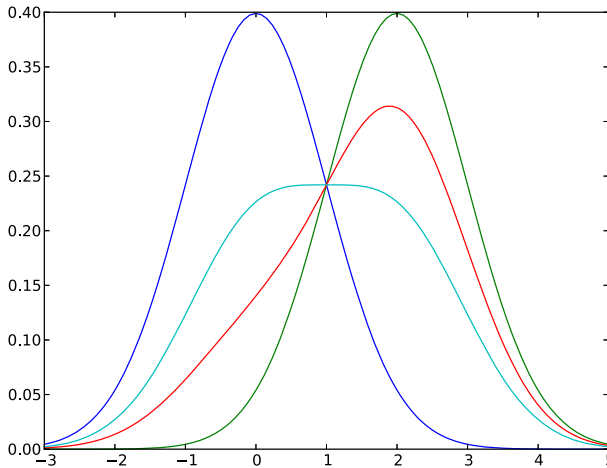
Fig. 3 Example of convex combination of normal distributions

This model includes bilinear terms $p_{i,s} y_i$ in the objective and is nonlinear and non-convex.

4.2 Parameterization of distribution

This formulation changes the parameters of a probability distribution directly, rather than distorting or combining some pre-existing probability distributions. Taking a known probability distribution and letting the model choose the mean, or variability, for example, would allow for a range of interesting applications. This formulation gives the ability to model general properties such as an increase of the expected value or reduction of variability. It is often desirable to apply continuous distributions. To stay within the frameworks of scenario-based recourse models the distribution must be discretized:

For a stochastic parameter x, define an allowed interval $[X^L, X^U]$ which is divided into $|S|$ subintervals, one for each scenario $s \in S$. The subintervals are $[x_{L,s}, x_{U,s}]$, $X^L \leq x_{L,s}, x_{U,s} \leq X^U, \forall s \in S$, using a representative value $x_{M,s}$ for each scenario, normally $x_{M,s} = \frac{x_{L,s}+x_{U,s}}{2}$. The probability of a scenario p_s is given by the cumulative probability (cumulative density function, cdf) of the upper value less the cumulative probability of the lower value of each subinterval: $p_s = cdf(x_{U,s}) - cdf(x_{L,s})$.

We will first give a formulation using a discretization of a probability distribution with closed form cdf Sect. 4.2.1, then a discretization of an approximation of the Normal distribution in Sect. 4.2.2.

4.2.1 Kumaraswamy distribution

The double bounded pdf proposed in Kumaraswamy (1980) was developed to better match observed values in hydrology. In practice it has been in little use, but interest in it

is increasing. This is among others because it is closely related to the Beta distribution and because the Kumaraswamy probability density function has the nice property that both the pdf and the cdf have closed form. With parameters $a, b > 0$, $x \in [0, 1]$, the probability density function is given as

$$f(x|a, b) = abx^{a-1}(1 - x^a)^{b-1}. \tag{8}$$

While the cumulative density function is:

$$F(x|a, b) = 1 - (1 - x^a)^b. \tag{9}$$

Note that the original formulation allows parameters $a, b \geq 0$, but as values $a, b = 0$ would imply situations where the probability of all scenarios equal to 0, this possibility is excluded. Interestingly, the shape of the probability density function changes radically when parameters a or b pass from a value < 1.0 to a value > 1.0, see Fig. 4 for examples.

With the cumulative probability given as a closed form expression, the discretized Kumaraswamy distribution can be directly included in an optimization model as follows, see also example in Sect. 5.1.4:

$$
\begin{aligned}
p_s(a, b) &= F(x_{U,s}|a, b) - F(x_{L,s}|a, b) \\
&= 1 - (1 - x_{U,s}^a)^b - 1 + (1 - x_{L,s}^a)^b \\
&= (1 - x_{L,s}^a)^b - (1 - x_{U,s}^a)^b, \quad \forall s \in \mathcal{S}.
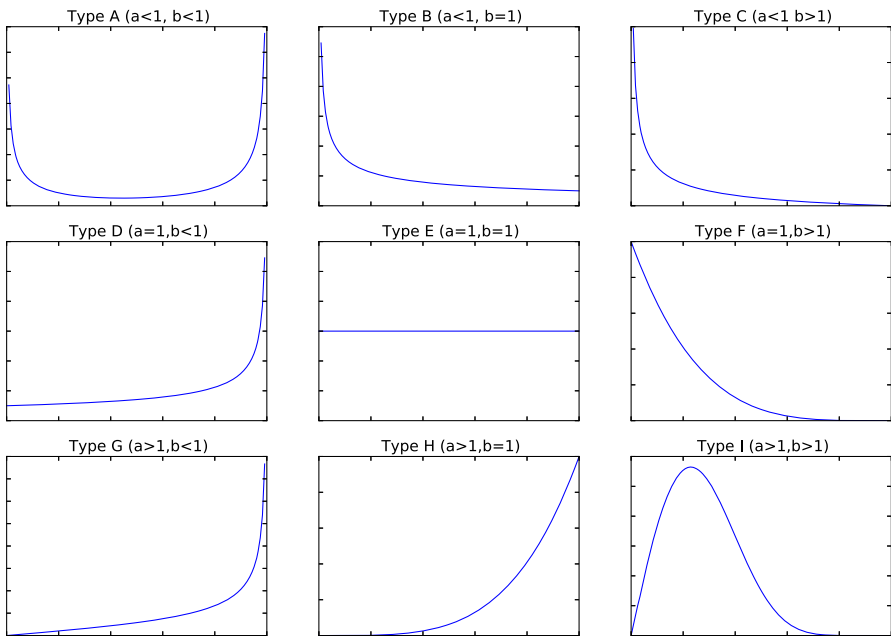\end{aligned} \tag{10}
$$

Fig. 4 Examples of Kumaraswamy probability density function (pdf) with different parameters a and b

This model includes a complex polynomial expression as well as the previously mentioned bilinear terms, resulting in a non-convex nonlinear formulation.

4.2.2 Approximation of normal distribution

The widely applied normal distribution has no closed form cdf, which makes it difficult to apply directly. Fortunately, there are polynomial and rational approximations to the standard normal distribution. For example, the cdf of the standard normal distribution can be approximated for $x \geq 0$ with the following expression (Abramowitz and Stegun 1964, 26.2.19):

$$P(x) = 1 - \frac{1}{2} \left(1 + d_1 x + d_2 x^2 + d_3 x^3 + d_4 x^4 + d_5 x^5 + d_6 x^6 \right)^{-16} + \epsilon(x)$$

$$|\epsilon(x)| < 1.5 \times 10^{-7}$$

$$d_1 = 0.0498673470 \quad d_3 = 0.0032776263 \quad d_5 = 0.0000488906$$

$$d_2 = 0.0211410061 \quad d_4 = 0.0000380036 \quad d_6 = 0.0000053830. \tag{11}$$

This closed form approximation for the normal distribution can be used in the model. To include a normal distribution where the mean is a decision variable, an expression for the cdf with mean a is needed, for example by applying the change of variables $x = x' - a$ (see Fig. 5) to the expression of the standard normal distribution cdf above. As the approximation is only valid for positive x, the symmetry of the standard normal distribution is exploited to use $P^-(x) = P(-x)$ for $x < 0$ to approximate the normal distribution $N(a, 1)$. This disjunctive formulation combining one expression for positive x with another for negative x requires the use of binary variables, yielding a MINLP.
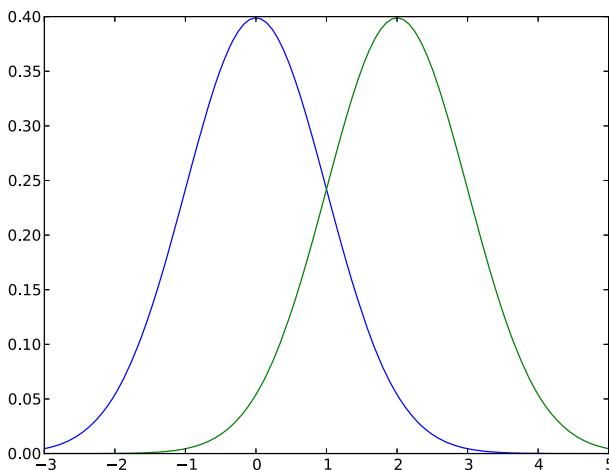

Fig. 5 Examples of Normal distribution shifted mean by change of variables

To express the split formulation of Eq. (11), split the expression into denominators divL_s^+ and divU_s^+ for $x_{M,s} - a > 0$ and divL_s^- and divU_s^- for $x_{M,s} - a \leq 0$:

$$\mathrm{divL}_s^+ = (1 + d_1(x_{L,s} - a) + d_2(x_{L,s} - a)^2 + d_3(x_{L,s} - a)^3$$
$$+ d_4(x_{L,s} - a)^4 + d_5(x_{L,s} - a)^5 + d_6(x_{L,s} - a)^6)^{16} \tag{12}$$

$$\mathrm{divU}_s^+ = (1 + d_1(x_{U,s} - a) + d_2(x_{U,s} - a)^2 + d_3(x_{U,s} - a)^3$$
$$+ d_4(x_{U,s} - a)^4 + d_5(x_{U,s} - a)^5 + d_6(x_{U,s} - a)^6)^{16} \tag{13}$$

$$\mathrm{divL}_s^- = (1 + d_1(-x_{L,s} + a) + d_2(-x_{L,s} + a)^2 + d_3(-x_{L,s} + a)^3$$
$$+ d_4(-x_{L,s} + a)^4 + d_5(-x_{L,s} + a)^5 + d_6(-x_{L,s} + a)^6)^{16} \tag{14}$$

$$\mathrm{divU}_s^- = (1 + d_1(-x_{U,s} + a) + d_2(-x_{U,s} + a)^2 + d_3(-x_{U,s} + a)^3$$
$$+ d_4(-x_{U,s} + a)^4 + d_5(-x_{U,s} + a)^5 + d_6(-x_{U,s} + a)^6)^{16} \tag{15}$$

In combination, the expressions above give the resulting interval probabilities for p_s^- and p_s^+:

$$p_s^- \leq (1 - \delta_s) \frac{1}{2} \left(\frac{1}{\mathrm{divL}_s^-} - \frac{1}{\mathrm{divU}_s^-} \right) \tag{16}$$

$$p_s^+ \leq \delta_s \frac{1}{2} \left(\frac{1}{\mathrm{divU}_s^+} - \frac{1}{\mathrm{divL}_s^+} \right) \tag{17}$$

For all scenarios $s \in \mathcal{S}$ with corresponding possible realization of the variable $x_{M,s} \in [x_{L,s}, x_{U,s}]$ use $x_{M,s}$ and binary variables $\delta_s \in \{0, 1\}, \forall s \in \mathcal{S}$ for determining the location of the interval. This will give some inaccuracy for the interval spanning both definitions. For improving accuracy, separate indicator variables may be used for upper and lower interval values, doubling the number of binary variables.

Note that to calculate the cumulative probabilities correctly for the tail scenarios, extreme values for the end points $x_{L,1}$ and $x_{U,|S|}$ can be used.

$$p_s = p_s^- + p_s^+, \tag{18}$$
$$p_s^+ \leq 1 - \delta_s, \tag{19}$$
$$p_s^- \leq \delta_s. \tag{20}$$

Ensure appropriate δ_s is set to 1 with big M constraints using constants M^+ and M^-:

$$x_{M,s} - a \leq (1 - \delta_s)M^+, \tag{21}$$
$$x_{M,s} - a \geq \delta_s M^-. \tag{22}$$

Bound probabilities to 1:

$$\sum_s p_s = 1, \sum_s p_s^+ \leq 1, \sum_s p_s^- \leq 1, \forall s \in \mathcal{S}. \tag{23}$$

Only allow one shift from negative to positive:

$$\delta_s \leq \delta_{s-1}, \forall s = 2 \ldots S. \tag{24}$$

This model includes a complex polynomial expression as well as binary variables, resulting in a non-convex mixed integer nonlinear formulation.

5 Test instances and example

We have implemented a few test instances to investigate how hard they will be to solve. All test models are implemented as GAMS models and can be downloaded from http://iot.ntnu.no/users/hellemo/DDP/. Tests include data sets with different numbers of scenarios. The results of these experiments can be seen in Sect. 6.1. Our test case looks at capacity expansion of power generation. The investor seeks to minimize the cost of meeting a given demand. Either unit cost or demand is stochastic. In addition to the available production technologies, it is possible to invest in an activity or technology that will alter the probabilities of the scenarios occurring. By investing in such a technology or activity, it is possible to alter the probability distribution as discussed in Sect. 4.

5.1 Test instances

The mathematical formulations of each test model follow here, first the base model in Sect. 5.1.1, then in the following subsections the deviations from the base model in accordance with the models discussed above. These modifications mostly concern the objective function.

5.1.1 Base model

B Total investment budget,
$\mathcal{G}$ Set of probability distributions or subset of scenarios (index g),
$\mathcal{I}$ Set of available technologies (index i),
$\mathcal{J}$ Set of modes of electricity demand (index j),
$\mathcal{S}$ Set of scenarios (index s),
p_{gs} Probability of scenario s for probability distribution g,
π_{js} Price of electricity in mode j in scenario s,
x_i New capacity of i, decided in first stage,
c_i Unit investment cost of i,
c Unit investment cost of increasing weight to a subset of scenarios,
c_g Unit investment cost of increasing weight to probability distribution g,
d_{js} Electricity demand in mode j in scenario s (if stochastic),
q_{is} Unit production cost of i in scenario s (if stochastic)
y_g Weight assigned to distribution g in a mixed distribution formulation,
y Scaling factor for a subgroup of scenarios in the scaling formulation,

z_{ijs} Production rate from i for mode j in scenario s,
$\overline{X_i}$ Upper bound on x_i,
$\underline{X_i}$ Lower bound on x_i,
$\overline{Y_g}$ Upper bound on y_g,
$\underline{Y_g}$ Lower bound on y_g,
$\overline{Z_{ij}}$ Upper bound on z_{ij},
$\underline{Z_{ij}}$ Lower bound on z_{ij}.

$$\min \sum_{i \in \mathcal{I}} c_i x_i + \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs}, \tag{25}$$

subject to:

$$\sum_{i \in \mathcal{I}} z_{ijs} = d_{js} \quad \forall s \in \mathcal{S}, j \in \mathcal{J}, \tag{26}$$

$$\sum_{j \in \mathcal{J}} z_{ijs} \leq x_i \quad \forall s \in \mathcal{S}, i \in \mathcal{I}, \tag{27}$$

$$\sum_{i \in \mathcal{I}} c_i x_i \leq B, \tag{28}$$

$$\underline{X_i} \leq x_i \leq \overline{X_i}, \forall i \in \mathcal{I}, \tag{29}$$

$$\underline{Z_{ij}} \leq z_{ijs} \leq \overline{Z_{ij}}, \forall i \in \mathcal{I}, j \in \mathcal{J}. \tag{30}$$

This model takes inspiration from the model of Louveaux and Smeers (1988), an investment problem from the electricity sector. There are I technologies available to invest in to generate electricity in order to meet demand. The demand for electricity in mode $j \in \mathcal{J}$ is given by the parameter d_{js} (alternatively this could be considered as demand in a location j). The model is formulated as a two-stage stochastic recourse model. As before the scenario tree is defined by scenarios $s \in S$.

New capacity of technology i is decided upon and installed in the first stage, determined by variables x_i. The objective function minimizes the aggregated costs of investments $c_i x_i$ for all technologies i and expected operational cost over all scenarios s, represented by the unit costs q_{is}, unit income π_{js} and production z_{ijs}. Demand at location (or mode) d_{js} is met by production at locations i allocated to mode j, z_{ijs} as described in Eq. (26). The total capacity available for technology i in stage two is limited by the investments in the first stage x_i by Eq. (27). The investments in technologies x_i are limited by the budget B (Eq. 28).

To enforce relatively complete recourse, Louveaux and Smeers (1988) make sure there is a technology $i_{\mathrm{rcr}} \in \mathcal{I}$ with high production cost which simulates purchases in the market to balance supply $\sum_{i \in \mathcal{I}} z_{ijs}$ and demand d_{js}. All variables are bounded, Eqs. (29) and (30).

5.1.2 Scalable subsets of scenarios

We present here an extension of the base model in Sect. 5.1.1 with scalable decision-dependent probabilities. This is an example of the linear scaling with uniform distributions in Sect. 4.1.1. Let as before $s \in \mathcal{S}$ be scenarios, each with equal probability. For each $s \in \hat{\mathcal{S}} \subset \mathcal{S}$ let the variable y represent the possibility to invest in scaling the probability linearly, whereas the remaining scenarios $s \in \mathcal{S} \setminus \hat{\mathcal{S}}$ are adjusted proportionally in the opposite direction. The practical interpretation is that by investing in a technology or activity, it is possible to increase the probability of some scenarios, while reducing the probability of the remaining scenarios, or vice versa.

Starting with the base model in Sect. 5.1.1, the objective Eq. (25) is replaced with Eq. (31):

$$
\min \sum_{i \in \mathcal{I}} c_i x_i + cy + \sum_{s \in \hat{\mathcal{S}}} \frac{1}{|\mathcal{S}|} y \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs}
$$
$$
+ \sum_{s \in \mathcal{S} \setminus \hat{\mathcal{S}}} \frac{1 - y \sum_{s \in \hat{\mathcal{S}}} \frac{1}{|\mathcal{S}|}}{|\mathcal{S}| - |\hat{\mathcal{S}}|} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs}. \tag{31}
$$

The investment must still stay within the budget so replace Eq. (28) with:

$$
\sum_{i \in \mathcal{I}} c_i x_i + cy \leq B. \tag{32}
$$

Apart from this the base model is unchanged.

5.1.3 Convex combination of probabilities

Here the mixture distribution is applied, modeling the possibility to change the weights of the underlying probability distributions for the subsets of outcomes. The decision-maker can invest to change the weight of each probability distribution $g \in \mathcal{G}$ represented by y_g, and the associated cost is given by parameter c_g. This can be used to model heterogeneous populations where the relative size of each subpopulation $g \in \mathcal{G}$ can be influenced by a decision variable y_g, determining the relative probability p_{gs} for each scenario $s \in \mathcal{S}$. The sum of the weights to all probability distributions $\sum_{g \in \mathcal{G}} y_g$ must equal 1:

$$
\sum_{g \in \mathcal{G}} y_g = 1, \tag{33}
$$

Replace Eq. (25) with Eq. (34):

$$
\min \sum_{i \in \mathcal{I}} c_i x_i + \sum_{g \in \mathcal{G}} c_g y_g
$$
$$
+ \sum_{s \in \mathcal{S}} \left\{ \sum_{g=1}^{G-1} p_{gs} y_g \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs} + \left[1 - \sum_{g=1}^{G-1} y_g \right] p_{Gs} \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs} \right\}. \tag{34}
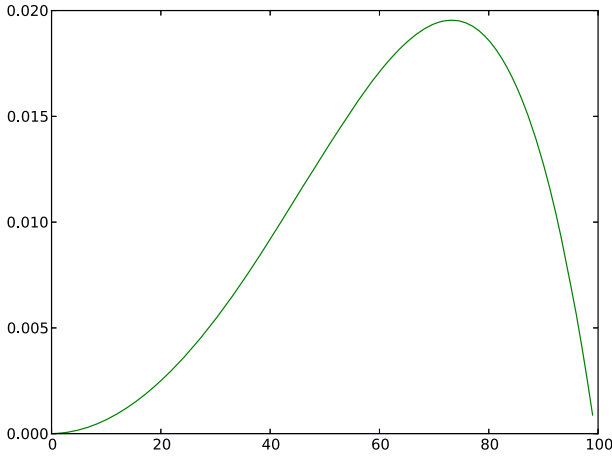$$

Fig. 6 Probabilities from solution of GAMS implementation of a model using 100 scenarios and Kumaraswamy distribution

As before, the budget must stay within the limit so replace Eq. (28) with:

$$\sum_{i\in\mathcal{I}} c_i x_i + \sum_{g=1}^{G-1} c_g y_g \leq B. \tag{35}$$

5.1.4 Kumaraswamy

In this formulation, the decision-maker can change directly parameters a and b in the distribution, possibly at a cost. For this specific problem, it can for example be interpreted as changing the characteristics of the cost uncertainty. See Fig. 6 for an example of scenario probabilities with parameters chosen in the example model. Replace the expression for p_s in Eq. (25) with Eq. (36):

$$
\begin{aligned}
\min &\sum_{i\in\mathcal{I}} c_i x_i + c_a a + c_b b + \sum_{s\in\mathcal{S}} p_s(a,b) \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} (q_{is} - \pi_{js}) z_{ijs} \\
&= \sum_{i\in\mathcal{I}} c_i x_i + c_a a + c_b b + \sum_{s\in\mathcal{S}} \left[(1 - x_{L,s}^a)^b - (1 - x_{U,s}^a)^b \right] \sum_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} (q_{is} - \pi_{js}) z_{ijs}
\end{aligned}
\tag{36}
$$

As explained in Subsection Sect. 4.2 we use a discrete approximation of the continuous distribution. This leads to a nonlinear non-convex formulation due to the polynomial distribution function (degree depends on decision variables a and b) and its multiplication with the continuous variable z.

Also replace the budget constraint Eq. (28) with Eq. (37):

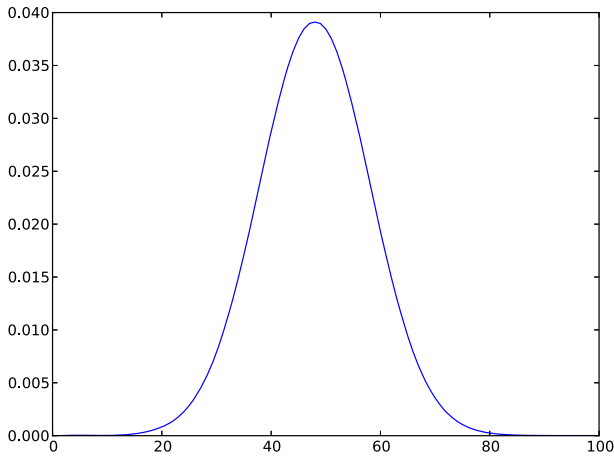$$\sum_{i\in\mathcal{I}} c_i x_i + c_a a + c_b b \leq B. \tag{37}$$

Fig. 7 Probabilities from optimal solution of GAMS implementation of a model using 100 scenarios and approximation of a normal distribution

Parameters a and b should be positive:

$$a, b > 0. \tag{38}$$

5.1.5 Approximation of normal distribution

The cdf of a standard distribution with mean a can be found through a change of variables $x = x' - a$. Using $P^-(x) = P(-x)$ for $x < 0$, the normal distribution $N(a, 1)$ can be approximated. See Fig. 7 for an example of resulting probabilities in the test model.

Replace the objective function given in Eq. (25) with Eq. (39) and use the scenarios $s \in \mathcal{S}$ with corresponding possible realization of the variable $x_{M,s} \in [x_{L,s}, x_{U,s}]$:

$$\min \sum_{i \in \mathcal{I}} c_i x_i + c_a a + \sum_{s \in \mathcal{S}} p_s \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} (q_{is} - \pi_{js}) z_{ijs} \tag{39}$$

where p_s follows the definition from Sect. 4.2.2 and is defined by Eq. (12) to Eq. (24).

Also replace the budget constraint Eq. (28) with Eq. (40):

$$\sum_{i \in \mathcal{I}} c_i x_i + c_a a \leq B. \tag{40}$$

Parameter a should be positive:

$$a > 0. \tag{41}$$

This model includes a complex polynomial expression due to the decision variable for the mean a, a bilinear term where probability p_s is multiplied with continuous

variable z as well as binary variables, resulting in a non-convex mixed integer nonlinear formulation.

5.2 Example of the effects of DDP

To illustrate the effects of decision-dependent probabilities in our models, we will look at the results from one test instance with the approximation of the Normal distribution from Sect. 5.1.5. This instance has stochastic demand, and the demand can be increased by engaging in some activity, for example by investing in campaigning, improving the safety or by reducing emissions from production if the demand is sensitive to these parameters. In the model, demand is influenced by shifting the mean of the probability distribution by a.

In this instance the mean may be shifted by $a \in [-1.0, 0]$. The mean is shifted in the opposite direction from the original model, hence $a \leq 0$. The uncertain parameters are discretized with 10 scenarios. The outcomes for the stochastic parameters are fixed for each scenario, while the probabilities for each scenario occurring are determined by selecting the mean of the distribution. The investment decisions are whether to invest in any of the 10 available technologies x_i, $i \in 1, 2, \ldots, 10$.

The results are summarized in Fig. 8. The figure shows the optimal expected profit for different values of a in the upper pane, while the corresponding investment levels of technologies x_8, x_9 and x_{10} are shown in the lower pane. Expected profit increases with more negative a (increasing demand), and so does investment in the different technologies. As demand shifts it becomes profitable to invest in more technologies, also the ones with higher operating costs as the maximum investment level is reached for technologies with lower operating cost. See also Table 1 for details.

This example shows how the inclusion of decision-dependent probabilities changes the problem. Note that for fixed a, the resulting problem is a traditional stochastic
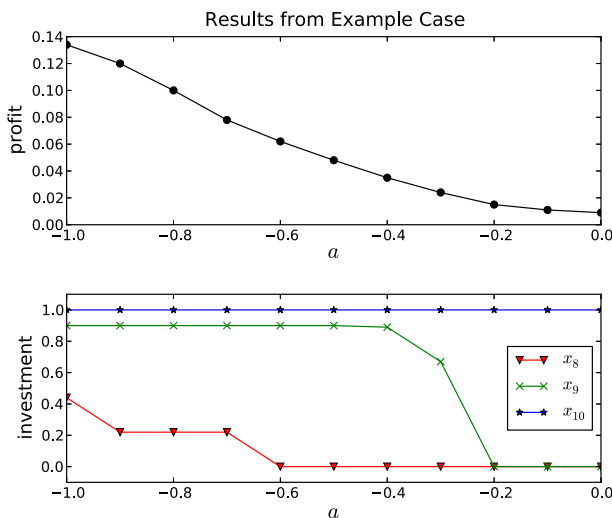

Fig. 8 Results from example model. Investments and profit increase with more negative a

Table 1 Summarized results from one test instance with different values of a, the objective function value, and values for investment in technologies x_8, x_9 and x_{10}

a	PROFIT	x_8	x_9	x_{10}
0	0.009	0	0	1
-0.1	0.011	0	0	1
-0.2	0.015	0	0	1
-0.3	0.024	0	0:67	1
-0.4	0.035	0	0:89	1
-0.5	0.048	0	0:9	1
-0.6	0.062	0	0:9	1
-0.7	0.078	0:22	0:9	1
-0.8	0.1	0:22	0:9	1
-0.9	0.12	0:22	0:9	1
-1.0	0.134	0:44	0:9	1

program with recourse. While finding the optimal solution of the problem with DDP is easy to do by inspection for this simple example, this is of course in general not a practical solution approach for such non-convex models where decision-dependencies are linked to several variables.

The test instances with computational results presented in the next section, are all based on synthetic data. Aggregated results from a series of test instances are provided to illustrate the computational difficulty of this class of problems.

6 Computational results

In this section the computational results from all four variations of the base model are presented. The models are all implemented in GAMS. We first present our solution strategy, followed by a summary of the computational results.

6.1 Solution strategy

All the formulations presented above introduce a continuous decision variable for the probability in a scenario multiplied with a decision variable for some activity, leading to a non-convex bilinear program. If the activities are continuous, this will be in the class of continuous bilinear non-convex nonlinear programs. In addition, other nonlinear terms may be needed in the corresponding optimization problems to represent probability distributions or approximations of these. Many of the potential applications of such models involve investment decisions. Fixed investment costs often require the use of discrete variables. Hence, the models where these modeling techniques should be applied will often already have integer variables, yielding a deterministic equivalent that is a mixed integer nonlinear non-convex model. In all the formulations, the complicating factor lies in the probability distribution and its multiplication with an activity.

Global optimization techniques must be applied to guarantee an optimal solution. BARON is the state-of-the-art global optimization solver, using convex relaxations for non-convex terms. A widely applied technique is to use McCormick relaxations to construct convex relaxations of factorable functions. BARON also applies techniques for constraint propagation to reduce the search space Tawarmalani and Sahinidis (2002). Tests were performed using BARON with three different approaches, Baron1: the problem were fed into BARON without information about structure and bilinear expressions expanded; Baron2: the same as the previous but using a selective branching strategy on the complicating variables motivated by Epperly and Pistikopoulos (1997); Baron3: The problem was fed into BARON using the original un-expanded bilinear expressions in GAMS and solved directly. In addition, the instances were tested using an approach combining relaxations of algorithms (Mitsos et al. 2009) and generalized Benders decomposition (Benders 1962; Geoffrion 1972) implemented for the purpose (GGBD).

We observed initially what appeared to be good results decomposing these stochastic programs based on GGBD and comparing it to the Baron1 approach. The Baron3 approach with a GAMS implementation of the same model showed much better performance, though. Baron2 using a selective branching strategy inferred from the problem structure, achieved the same behaviour as Baron3.

The selective branching strategy that was implemented, was to use the decision variables y in the decision-dependent probabilities $p(y)$ as the complicating variables and branching first on these in a continuous branch and bound scheme. Note that when fixing the variables that influence the probabilities of the scenarios, the resulting sub problems are much easier to solve. For the affine formulations given in Sect. 4.1, the remaining problem is a standard linear or mixed integer stochastic program. Our conclusion is that solution times for these problems can be dramatically improved by using this selective branching strategy. Such selective branching can be readily implemented through setting branching priority in BARON. Interestingly, using the original, un-expanded formulation achieved similar results to the selective branching strategy.

6.2 Solution times for test instances

In Table 2 results from running our test instances are presented, for the problems: scalable probabilities (Subsets), convex combination (Combination), Kumaraswamy and normal distribution (Rational). Each test instance is run with different numbers of scenarios. The resulting problem sizes, both in terms of number of rows, columns and number of discrete and nonlinear variables are all reported in the table. All problems were run with a time limit of one hour, and most test instances were close to optimal after one hour, although not as close as the stopping criterion of a relative gap $< 1 \times 10^{-5}$. All numbers presented are from Baron2 (Baron3 gave similar results).

Our numerical experiments show that BARON is generally able to solve the instances of the convex combination of probabilities from Sect. 5.1.3 as well as the scalable subsets of scenarios from Sect. 5.1.2 to optimality or close to optimal. The instances using the approximation of the normal distribution from Sect. 5.1.5 and the

Table 2 Full results from benchmarks

Model	#Scen.	Status	CPU	#NL	#D	Rows	Cols	LB	UB
Subsets	10	Unfinished	3600	602	0	1013	203	$-3.2 \cdot 10^{-1}$	$-3.2 \cdot 10^{-1}$
Combination	10	Optimal	0.5	901	0	1013	213	$-6.8 \cdot 10^{-1}$	$-6.8 \cdot 10^{-1}$
Kumaraswamy	10	Unfinished	3600	1020	0	1034	232	$-5.9 \cdot 10^{-1}$	$-5.9 \cdot 10^{-1}$
Rational	10	Optimal	4.27	1110	10	1092	334	$-5.2 \cdot 10^{-1}$	$-5.2 \cdot 10^{-1}$
Subsets	20	Optimal	2.25	1302	0	2013	403	$-3.4 \cdot 10^{-1}$	$-3.4 \cdot 10^{-1}$
Combination	20	Optimal	1.48	1901	0	2013	423	$-6.6 \cdot 10^{-1}$	$-6.6 \cdot 10^{-1}$
Kumaraswamy	20	Unfinished	3600.01	2040	20	2054	462	$-5.9 \cdot 10^{-1}$	$-5.9 \cdot 10^{-1}$
Rational	20	Unfinished	3601.02	2220	0	2172	664	$-5.9 \cdot 10^{-1}$	$-5.2 \cdot 10^{-1}$
Subsets	30	Unfinished	3600	1902	0	3013	603	$-3.4 \cdot 10^{-1}$	$-3.4 \cdot 10^{-1}$
Combination	30	Optimal	2.98	2901	0	3013	633	$-6.6 \cdot 10^{-1}$	$-6.6 \cdot 10^{-1}$
Kumaraswamy	30	Unfinished	3600.02	3060	0	3074	692	$-5.9 \cdot 10^{-1}$	$-5.9 \cdot 10^{-1}$
Rational	30	Unfinished	3600.02	3330	30	3252	994	$-6.8 \cdot 10^{-1}$	$-5.2 \cdot 10^{-1}$
Subsets	50	Unfinished	3600.04	3302	0	5013	1003	$-3.5 \cdot 10^{-1}$	$-3.5 \cdot 10^{-1}$
Combination	50	Optimal	9.08	4901	0	5013	1053	$-6.5 \cdot 10^{-1}$	$-6.5 \cdot 10^{-1}$
Kumaraswamy	50	Unfinished	3600.03	5100	0	5114	1152	$-5.9 \cdot 10^{-1}$	$-5.9 \cdot 10^{-1}$
Rational	50	Optimal	93.21	5550	50	5412	1654	$-5.2 \cdot 10^{-1}$	$-5.2 \cdot 10^{-1}$
Subsets	100	Optimal	66.59	6602	0	10013	2003	$-3.5 \cdot 10^{-1}$	$-3.5 \cdot 10^{-1}$
Combination	100	Optimal	45.27	9801	0	10,013	2103	$-6.5 \cdot 10^{-1}$	$-6.5 \cdot 10^{-1}$
Kumaraswamy	100	Unfinished	3600	10,200	0	10,214	2302	-5.3	$-5.9 \cdot 10^{-1}$
Rational	100	Optimal	429	11,100	100	10,812	3304	$-5 \cdot 10^{-1}$	$-5 \cdot 10^{-1}$

Model, Number of scenarios, Solver Status, Cputime in seconds, Number of nonlinear terms, number of discrete terms, rows and columns, and lower and upper bound after stop criterion

Kumaraswamy models from Sect. 5.1.4 proved harder to solve, and while the solver has found a good solution, optimality remains to be proved within the time limit. BARON is able to solve relatively large problems in reasonable time if the problem formulation provides enough structure for the solver to choose an efficient solution strategy. In cases where we provided an unstructured problem without a selective branching strategy, BARON would often end up doing a lot of unnecessary branching, which made convergence very slow and in general slower than our GGBD (Results not included).

For larger problems in the harder categories, specialized solution techniques may be necessary, and we hope that our test instances may come of use in future research in this area.

7 Conclusions and further work

Little work has been done on stochastic programming problems with decision-dependent probabilities. This work extends previous taxonomies of stochastic programming problems with decision-dependent uncertainty and presents some examples of models with decision-dependent probabilities. Our contribution is to show how direct or indirect manipulation of probability distribution can be incorporated in stochastic programs with recourse. The work demonstrates that such problems may be solved by the commercial solver BARON, using selective branching in the complicating variables. For the test instances, a selective branching strategy for the scenario probability variables proved much more efficient than the decomposition method implemented and tested. We provide a set of test cases for this class of problems.

As the models and analysis only considered linear dependency between cost and a change on the underlying probability distribution, an extension would be to introduce some nonlinear cost such as diminishing return to scale.

Our test cases were based on a risk neutral approach. Investigating the effects of different risk attitudes on decision-dependent probabilities is another area of research that would be very interesting to pursue.

Finally, as these large scale non-convex problems grow more complex, finding good and robust decomposition techniques would greatly improve the scale at which such techniques could be applied. We hope that the test problems provided can be a starting point for further research on solution methods for stochastic programming problems with decision-dependent probabilities.

Appendix: Hardware and software used

All computations were performed on a Six-Core AMD Opteron processor 2431 with 24Gb memory. The computer was running Linux 2.6.18 (Rocks 5.3).

GAMS versions 23.6.2 and 23.7.2 with BARON using CPLEX in combination with CONOPT or MINOS.

References

Abramowitz M, Stegun I (1964) Handbook of mathematical functions with formulas, graphs, and mathematical tables, vol 55. Dover Publications, New York

Ahmed S (2000) Strategic planning under uncertainty—stochastic integer programming approaches. Ph.D. thesis, Graduate College, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Apap RM, Grossmann IE (2017) Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties. Comput Chem Eng 103:233–274

Artstein Z, Wets R (1994) Stability results for stochastic programs and sensors, allowing for discontinuous objective functions. SIAM J Optim 4:537–550

Beale EML (1955) On minimizing a convex function subject to linear inequalities. J R Stat Soc B 17(2):173–184

Behboodian J (1970) On the modes of a mixture of two normal distributions. Technometrics 12(1):131–139

Benders JF (1962) Partitioning procedures for solving mixed-variables programming problems. Numer Math 4(1):238–252

Ben-Tal A, Nemirovski A (1998) Robust convex optimization. Math Oper Res 23(4):769–805

Ben-Tal A, Eiger G, Gershovitz V (1994) Global minimization by reducing the duality gap. Math Program 63(1):193–212

Bertsimas D, Sim M (2003) Robust discrete optimization and network flows. Math Program 98(1):49–71

Bertsimas D, Sim M (2004) The price of robustness. Oper Res 52(1):35–53

Boland N, Dumitrescu I, Froyland G (2008) A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology. In: 7th joint Australia-New Zealand Mathematics Convention (ANZMC2008), Christchurch, New Zealand

Colvin M, Maravelias CT (2008) A stochastic programming approach for clinical trial planning in new drug development. Comput Chem Eng 32(11):2626–42

Colvin M, Maravelias CT (2009a) Scheduling of testing tasks and resource planning in new product development using stochastic programming. Comput Chem Eng 33(5):964–976

Colvin M, Maravelias C (2009b) A branch and cut framework for multi-stage stochastic programming problems under endogenous uncertainty. Comput Aided Chem Eng 27:255–260

Colvin M, Maravelias CT (2010) Modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming. Eur J Oper Res 203(1):205–15

Dantzig GB (1955) Linear programming under uncertainty. Manag Sci 1:197–206

Dupačová J (2006) Optimization under exogenous and endogenous uncertainty. In: Lukáš L (ed) Proceedings of MME06, University of West Bohemia in Pilsen, pp 131–136

Epperly TG, Pistikopoulos EN (1997) A reduced space branch and bound algorithm for global optimization. J Global Optim 11(3):287–311

Escudero LF, Garín MA, Merino M, Pérez G (2014) On multistage mixed 0–1 optimization under a mixture of Exogenous and Endogenous Uncertainty in a risk averse environment. Working paper

Escudero L, Garin A, Monge J, Unzueta A (2016) On preparedness resource allocation planning for natural disaster relief by multistage stochastic mixed 0–1 bilinear optimization based on endogenous uncertainty and time consistent risk averse management. Working paper

Feller W (1943) On a general class of "contagious" distributions. Ann Math Stat 14(4):389–400

Frühwirth-Schnatter S (2006) Finite mixture and Markov switching models. Springer, New York

Geoffrion A (1972) Generalized benders decomposition. J Optim Theory Appl 10(4):237–260

Goel V, Grossmann I (2004) A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. Comput Chem Eng 28(8):1409–1429

Goel V, Grossmann I (2006) A class of stochastic programs with decision dependent uncertainty. Math Program 108(2):355–394

Gupta V, Grossmann IE (2011) Solution strategies for multistage stochastic programming with endogenous uncertainties. Comput Chem Eng 35(11):2235–2247

Held H, Woodruff D (2005) Heuristics for multi-stage interdiction of stochastic networks. J Heuristics 11(5):483–500

Jonsbråten T (1998) Oil field optimization under price uncertainty. J Oper Res Soc 49(8):811–818

Jonsbråten T, Wets R, Woodruff D (1998) A class of stochastic programs with decision dependent random elements. Ann Oper Res 82:83–106

Kuhn D (2009) An information-based approximation scheme for stochastic optimization problems in continuous time. Math Oper Res 34(2):428–444

Kuhn D, Wiesemann W, Georghiou A (2011) Primal and dual linear decision rules in stochastic and robust optimization. Math Program 130(1):177–209

Kumaraswamy P (1980) A generalized probability density function for double-bounded random processes. J Hydrol 46(1):79–88

Lappas N, Gounaris C (2016) Multi-stage adjustable robust optimization for process scheduling under uncertainty. AIChE Journal 62:1646–1667. https://doi.org/10.1002/aic.15183

Lappas N, Gounaris CE (2018) Robust optimization for decision-making under endogenous uncertainty. Comput Chem Eng. https://doi.org/10.1016/j.compchemeng.2018.01.006

Lejeune M, Margot F (2017) Aeromedical battle field evacuation under endogenous uncertainty in casualty delivery times. Manag Sci. https://doi.org/10.1287/mnsc.2017.2894

Louveaux FV, Smeers Y (1988) Optimal investments for electricity generation: a stochastic model and a test problem. In: Ermoliev Y, Wets RJB (eds) Numerical techniques for stochastic optimization. Springer, Berlin, pp 445–454

Mitsos A, Chachuat B, Barton P (2009) McCormick-based relaxations of algorithms. SIAM J Optim 20(2009):573–601

Nohadani O, Sharma K (2016) Optimization under decision-dependent uncertainty. arXiv:1611.07992

Ntaimo L, Arrubla JAG, Stripling C, Young J, Spencer T (2012) A stochastic programming standard response model for wildfire initial attack planning. Can J For Res 42(6):987–1001

Peeta S, Salman FS, Gunnec D, Viswanath K (2010) Pre-disaster investment decisions for strengthening a highway network. Comput Oper Res 37(10):1708–1719

Pflug G (1996) Optimization of stochastic models: the interface between simulation and optimization. Kluwer Academic, Boston

Pflug GC, Pichler A (2011) Approximations for probability distributions and stochastic optimization problems. Stochastic optimization methods in finance and energy. Springer, New York, pp 343–387

Pflug G, Wozabal D (2007) Ambiguity in portfolio selection. Quant Finance 7(4):435–442

Rubinstein RY, Shapiro A (1993) Discrete event systems: sensitivity analysis and stochastic optimization by the score function method, vol 346. Wiley, New York

Solak S (2007) Efficient solution procedures for multistage stochastic formulations of two problem classes. Ph.D. thesis, Georgia Institute of Technology, Atlanta

Solak S, Clarke JP, Johnson E, Barnes E (2010) Modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming. Eur J Oper Res 207(1):420–433

Tarhan B, Grossmann I, Goel V (2009) Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. Ind Eng Chem Res 48(6):3078–3097

Tawarmalani M, Sahinidis NV (2002) Convexification and global optimization in continuous and mixed-integer nonlinear programming. Kluwer Academic Publishers, Norwell

Varaiya P, Wets RJB (1989) Stochastic dynamic optimization, approaches and computation. In: Iri M, Tanabe K (eds) Mathematical programming, recent developments and applications. Kluwer Academic Publisher, Boston, pp 309–332

Viswanath K, Peeta S, Salman SF (2004) Investing in the links of a stochastic network to minimize expected shortest path length. Tech. rep., Purdue University, Department of Economics, West Lafayette