

SPRINGER BRIEFS IN OPTIMIZATION

Urmila Diwekar
Amy David

**BONUS Algorithm
for Large Scale
Stochastic
Nonlinear
Programming
Problems**



Springer

SpringerBriefs in Optimization

Series Editors

Panos M. Pardalos

János D. Pintér

Stephen M. Robinson

Tamás Terlaky

My T. Thai

SpringerBriefs in Optimization showcases algorithmic and theoretical techniques, case studies, and applications within the broad-based field of optimization. Manuscripts related to the ever-growing applications of optimization in applied mathematics, engineering, medicine, economics, and other applied sciences are encouraged.

For further volumes:

<http://www.springer.com/series/8918>

Urmila Diwekar • Amy David

BONUS Algorithm for Large Scale Stochastic Nonlinear Programming Problems

 Springer

Urmila Diwekar
Clarendon Hills
Illinois
USA

Amy David
Krennert School of Business
Purdue University
West Lafayette
Indiana
USA

ISSN 2190-8354

SpringerBriefs in Optimization

ISBN 978-1-4939-2281-9

DOI 10.1007/978-1-4939-2282-6

ISSN 2191-575X (electronic)

ISBN 978-1-4939-2282-6 (eBook)

Library of Congress Control Number: 2014955715

Springer New York Heidelberg Dordrecht London

© Urmila Diwekar, Amy David 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

*To my husband Dr. Sanjay Joag,
who changed his career to support mine,
for his constant love and support.
—Urmila*

Preface

Stochastic programming problems are very difficult to solve as they involve optimization as well as uncertainty analysis. Algorithms for solving large-scale nonlinear stochastic programming problems are very few in number, as are the engineering applications of these problems. This book introduces two algorithms for large-scale stochastic nonlinear problems for both open equation systems and black box models. These algorithms are the Better Optimization of Nonlinear Uncertain Systems (BONUS) algorithm and the L-shaped BONUS algorithm. Real-world applications of these algorithms in the areas of energy and environmental engineering are also detailed. Many have contributed to this book. Researchers who worked with Dr. Diwekar including Dr. Adrian Lee, Dr. Kemal Sahin, Dr. Juan Salazar, and Dr. Yogendra Shastri, as well as collaborators such as Dr. Emil Constantinescu, Dr. Victor Zavala, and Dr. Stephen Zitney have provided the material for this book with their research. Thanks also to our group members Dr. Pahola Benavides, Dr. Berhane Gabreslassie, Dr. Rajib Mukherjee, Shivam Tyagi, and Kirti Yenki who went through the first draft of the book and meticulously pointed out mistakes. Hope you enjoy this work.

Urmila Diwekar and Amy David

Contents

1	Introduction	1
1.1	Stochastic Optimization Problems	1
1.2	Stochastic Nonlinear Programming	4
1.3	Summary	6
	Notations	7
2	Uncertainty Analysis and Sampling Techniques	9
2.1	Specifying Uncertainty Using Probability Distributions	9
2.2	Sampling Techniques	10
2.2.1	Monte Carlo Sampling	11
2.3	Variance Reduction Techniques	13
2.3.1	Importance Sampling	13
2.3.2	Stratified Sampling	16
2.3.3	Quasi-Monte Carlo Methods	19
2.4	Summary	24
	Notations	24
3	Probability Density Functions and Kernel Density Estimation	27
3.1	The Histogram	27
3.2	Kernel Density Estimator	28
3.3	Summary	32
	Notations	34
4	The BONUS Algorithm	35
4.1	Reweighting Schemes	36
4.2	Effect of Sampling on Reweighting	38
4.3	BONUS: The Novel SNLP Algorithm	41
4.4	Summary	54
	Notations	55
5	Water Management Under Weather Uncertainty	57
5.1	Introduction	57

5.2	The Pulverized Coal Power Plant	57
5.3	Parameter Uncertainty	60
5.4	Problem Formulation	61
5.5	Selection of Decision Variables	62
5.6	Implementation of BONUS Algorithm	63
5.7	Results	64
5.8	Summary	65
	Notations	65
6	Real-Time Optimization for Water Management	67
6.1	Introduction	67
6.2	Power Plant Operations	67
6.3	Formulation of the Stochastic Problem	70
6.4	Solution Approach	70
6.5	Weather Forecasting and Uncertainty Quantification	72
6.5.1	Ensemble Initialization	72
6.5.2	Ensemble Propagation	73
6.5.3	Validation of Weather Forecast	74
6.6	Application to Pulverized Coal Power Plant	75
6.7	Summary	78
	Notations	79
7	Sensor Placement Under Uncertainty for Power Plants	81
7.1	Introduction	81
7.1.1	The Integrated Gasification Combined Cycle Power Plant ...	81
7.1.2	Measurement Uncertainty	83
7.2	Fisher Information and Its Use in the Sensor-Placement Problem ...	84
7.3	Computation of Fisher Information	84
7.3.1	Reweighting Using the BONUS Method	85
7.3.2	Calculating the Fisher Information from Kernel Density Estimation	86
7.4	The Optimization Problem	87
7.4.1	Defining the Objective Function	87
7.4.2	The IGCC Power Plant	88
7.4.3	Problem Approach	90
7.4.4	Results	91
7.5	Summary	93
	Notations	93
8	The L-Shaped BONUS Algorithm	95
8.1	The L-Shaped BONUS Algorithm	100
8.2	Illustrative Example 1: The Farmer's Problem	102
8.2.1	Problem Formulation	102
8.2.2	Problem Solution	105
8.2.3	Results of the Farmer's Problem	107

- 8.3 Illustrative Example 2: The Blending Problem 108
 - 8.3.1 Problem Formulation 109
 - 8.3.2 Simulations and Results 111
- 8.4 Summary 112
- Notations 113

- 9 The Environmental Trading Problem 117**
 - 9.1 Introduction 117
 - 9.2 Basics of Pollutant Trading 117
 - 9.3 Christina Watershed Nutrient Management 118
 - 9.4 Trading Problem Formulation 119
 - 9.5 Results 124
 - 9.6 Summary 126
 - Notations 126

- 10 Water Security Networks 127**
 - 10.1 Introduction 127
 - 10.2 Motivation and Prior Work 128
 - 10.3 Solution Methodology 130
 - 10.3.1 Use of BONUS Reweighting for Pattern Estimation 131
 - 10.3.2 Back Estimation of Flow Patterns 132
 - 10.4 Results 133
 - 10.5 Summary 136
 - Notations 137

- References 139**

- Index 143**

List of Figures

Fig. 1.1	Pictorial representation of the numerical optimization framework [7].....	2
Fig. 1.2	Different probabilistic performance measures (PDF) [7].....	3
Fig. 1.3	Different probabilistic performance measures (CDF) [7].....	3
Fig. 1.4	Pictorial representation of the stochastic programming framework.....	4
Fig. 1.5	Example of a dual block angular structure (LP), each diagonal block is a realization of a random variable(scenario or sample) .	5
Fig. 1.6	Effect of changes in decision variables	6
Fig. 2.1	The stochastic modeling framework.....	10
Fig. 2.2	Examples of probabilistic distribution functions for stochastic modeling.....	10
Fig. 2.3	PDF for a lognormal distribution. <i>PDF probability density function</i>	12
Fig. 2.4	Sample placement on the CDF. <i>CDF cumulative density function</i>	13
Fig. 2.5	(<i>Left hand side</i>) 100 pseudorandom numbers on a unit square, (<i>right hand side</i>) 250 pseudorandom numbers on a unit square obtained by the linear congruential generator developed by Wichmann and Hill [60].....	14
Fig. 2.6	The function behavior.....	15
Fig. 2.7	Lognormal distribution with a mean $\mu = 1$ and a standard deviation of $\sigma = 1.7$	16
Fig. 2.8	Distribution and stratification for variable X_1	17
Fig. 2.9	Distribution and stratification for variable X_2	18
Fig. 2.10	Two-dimensional representation of a possible Latin hypercube sample of size 5 using X_1 and X_2	19
Fig. 2.11	Generation of 100 Hammersley points in 2 dimension	21
Fig. 2.12	Generation of 100 points on a unit square from various sampling techniques	22
Fig. 2.13	Generation of 100 correlated points on a unit square from various sampling techniques	23

Fig. 2.14	One dimensional uniformity of various sampling techniques	24
Fig. 3.1	A typical histogram.....	28
Fig. 3.2	Probability density function from a normal KDE. <i>KDE</i> kernel density estimation	29
Fig. 3.3	Effect of h on PDF. <i>PDF</i> probability density function.....	29
Fig. 3.4	A bivariate distribution using KDE. <i>KDE</i> kernel density estimation	30
Fig. 3.5	PDF for the output variable Z . <i>PDF</i> probability density function	32
Fig. 4.1	Pictorial representation of the stochastic programming framework.....	36
Fig. 4.2	Density estimation approach to optimization under uncertainty .	37
Fig. 4.3	Variance calculation for different sampling techniques	40
Fig. 4.4	Optimization under uncertainty: The BONUS algorithm.....	41
Fig. 4.5	Nonisothermal CSTR	47
Fig. 4.6	Optimization progress for traditional SNLP approach	50
Fig. 4.7	Optimization progress in reducing product variance using BONUS.....	51
Fig. 4.8	Comparison of optimization progress.....	54
Fig. 5.1	Schematic of PC power plant	58
Fig. 5.2	Probability density functions of four seasons (fall to winter from <i>top</i>) for dry-bulb temperature and relative humidity in eight US Midwestern cities.....	60
Fig. 6.1	Schematic representation of the interface between the generation and cooling systems.....	68
Fig. 6.2	Forecast (<i>thick line</i>) and ensemble profiles (<i>thin lines</i>) for dry-bulb temperature and relative humidity in Midwest US for June 1, 2006. Dots are real measurements from meteorological stations.....	75
Fig. 6.3	Fit of ambient temperature WRF ensembles to lognormal distribution	76
Fig. 6.4	Fit of relative humidity WRF ensembles to uniform distribution.	76
Fig. 6.5	Maximum power profiles for deterministic and stochastic approaches.....	78
Fig. 7.1	The integrated gasification combined cycle power plant.....	82
Fig. 7.2	Potential sensor locations in the IGCC power plant. <i>IGCC</i> integrated gasification combined cycle.....	89
Fig. 8.1	L-shaped method decomposition strategy [7].....	96
Fig. 8.2	Recourse function term as a function of the decision variable ...	99
Fig. 8.3	The proposed L-shaped BONUS algorithm structure	101
Fig. 8.4	Variation of objective function with sample size for farmer's problem.....	107
Fig. 8.5	Variation of iteration requirement with sample size for farmer's problem.....	108
Fig. 8.6	Variation of objective function with sample size for blending problem.....	113

Fig. 8.7	Variation of iteration requirement with sample size for blending problem	114
Fig. 9.1	Variation of the objective function result with sample size	123
Fig. 9.2	Variation of decision variable $L(8,1)$ with sample size	124
Fig. 9.3	Variation of computational time with sample size	125
Fig. 10.1	Example water network	128
Fig. 10.2	Estimation error as a function of sample size	132
Fig. 10.3	Placement of low-cost sensors for different methods	135
Fig. 10.4	Placement of high-cost sensors for different methods	136

List of Tables

Table 2.1	Sample generation	13
Table 2.2	The estimation of the integral by using uniform random sampling and importance sampling.....	16
Table 2.3	Possible selection of values for a Latin hypercube sample of size 5 for the random variable X_1	18
Table 2.4	Possible selection of values for a Latin hypercube sample of size 5 for the random variable X_2	18
Table 2.5	Pairing X_1 and X_2 and generating samples.....	19
Table 2.6	Generation of 10 Hammersley points in 2 dimensions.....	21
Table 3.1	Various kernel density functions	30
Table 3.2	Values of decision variables and objective function for ten samples.....	31
Table 3.3	Kernel density estimation for the objective function Z	33
Table 4.1	Calculations for KDE efficiency analysis.....	39
Table 4.2	Bounds for base (uniform) and estimated (normal) distributions	39
Table 4.3	Percentage error in variance estimation for 3-dimensional analysis using 250 samples.....	40
Table 4.4	Base sample	44
Table 4.5	Base sample kernel density estimates	45
Table 4.6	Sample-optimization iteration 1.....	45
Table 4.7	Optimization iteration 1-KDE.....	46
Table 4.8	Optimization progress at $N = 100$	46
Table 4.9	Parameters and their values in CSTR study	48
Table 4.10	Decision variables for optimization	49
Table 4.11	Uncertain variables in capacity expansion case.....	53
Table 4.12	Constants for capacity expansion case	53
Table 4.13	Decision variables in capacity expansion case.....	54
Table 5.1	Partial rank correlation coefficients (PRCC) for relationship between potential decision variables and water consumption ..	63
Table 5.2	Minimization of average water consumption under uncertain air conditions for a 548 MW PC power plant located in the	

	Midwestern US for four different seasons. (Water consumption estimates are reported in millions of pounds per hour).....	64
Table 7.1	Input process variables	89
Table 7.2	Intermediate and output process variables	91
Table 7.3	Computed objective values, f_j^B , for each sensor type.....	92
Table 7.4	Measurement variation of the integrated gasification combined cycle (<i>IGCC</i>) power production and gasifier performance using the optimal sensor network versus no sensors deployed .	93
Table 8.1	Weekly demand uncertainties	97
Table 8.2	Data for farmer’s problem	103
Table 8.3	Data for chemicals in blending problem	111
Table 8.4	Data for blend products	112
Table 9.1	Point source details for Christina River Basin.....	122
Table 9.2	Point source details of NPS emission and treatment.....	123
Table 10.1	Comparison of estimated cost and risk for different solution methods	134
Table 10.2	Comparison of actual cost and risk for different solution methods	134

Chapter 1

Introduction

A general optimization problem can be stated as follows.

$$\begin{aligned} \text{Optimize} \quad & Z = z(x) && (1.1) \\ & x \end{aligned}$$

subject to

$$h(x) = 0 \tag{1.2}$$

$$g(x) \leq 0 \tag{1.3}$$

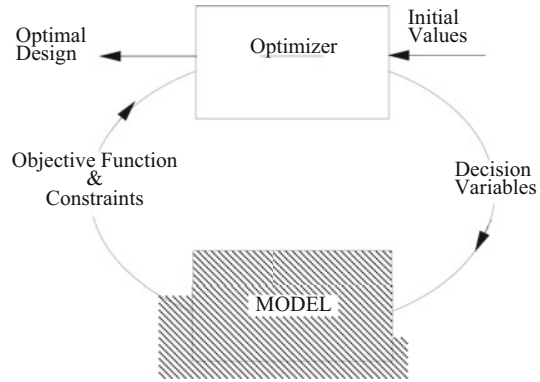
The goal of an optimization problem is to determine the decision variables x that optimize the objective function Z (Eq. 1.1), while ensuring that the model operates within established limits enforced by the equality constraints h (Eq. 1.2) and inequality constraints g (Eq. 1.3).

Figure 1.1 illustrates schematically the iterative procedure employed in a numerical optimization technique. As seen in the figure, the optimizer invokes the model with a set of values of decision variables x . The model simulates the phenomena and calculates the objective function and constraints. This information is utilized by the optimizer to calculate a new set of decision variables. This iterative sequence is continued until the optimization criteria pertaining to the optimization algorithm is satisfied. If the objective function and constraints are linear and the decision variables involved are scalar and continuous, then it is a linear programming (LP) problem. However, if the objective function and/or constraints are nonlinear then it is a nonlinear programming (NLP) problem. An NLP problem involving integers is a mixed integer nonlinear programming (MINLP) problem.

1.1 Stochastic Optimization Problems

Stochastic optimization gives us the ability to optimize systems in the face of uncertainties. A stochastic optimization or a stochastic programming (SP) problem requires that the objective function and constraints be expressed in terms of some

Fig. 1.1 Pictorial representation of the numerical optimization framework [7]



probabilistic representation (e.g., expected value, variance, fractiles, most likely values). For example, in chance constrained programming, the objective function is expressed in terms of expected value, and the constraints are expressed in terms of fractiles (probability of constraint violation), and in Taguchi's offline quality control method ([55], Diwekar and Rubin 1991), the objective is to minimize variance. These problems can be further classified as stochastic linear programming (SLP), stochastic nonlinear programming (SNLP), and stochastic mixed integer linear and nonlinear programming problems. The latter problems are the focus of this book.

A generalized stochastic optimization problem, where the decision variables and uncertain parameters are separated, can then be viewed as:

$$\text{Optimize } J = P_1(j(x, u)) \quad (1.4)$$

x

subject to

$$P_2(h(x, u)) = 0 \quad (1.5)$$

$$P_3(g(x, u) \geq 0) \geq \alpha \quad (1.6)$$

where u is the vector of uncertain parameters and P represents the cumulative distribution functional such as the expected value, mode, variance, or fractiles. Figures 1.2 and 1.3 show the expected value (mean), mode, variance, and fractiles for a probabilistic distribution function.

Unlike the deterministic optimization problem, in stochastic optimization one has to consider the probabilistic functional of the objective function and constraints. The generalized treatment of such problems is to use probabilistic or stochastic models instead of the deterministic model inside the optimization loop. Figure 1.4 represents the generalized solution procedure, where the deterministic model is replaced by an iterative stochastic model with a sampling loop representing the discretized uncertainty space. The uncertainty space is represented in terms of the moments such as the mean, or the standard deviation of the output over the sample space of N_{samp} as given by the following equations (Eqs. 1.7 and 1.8).

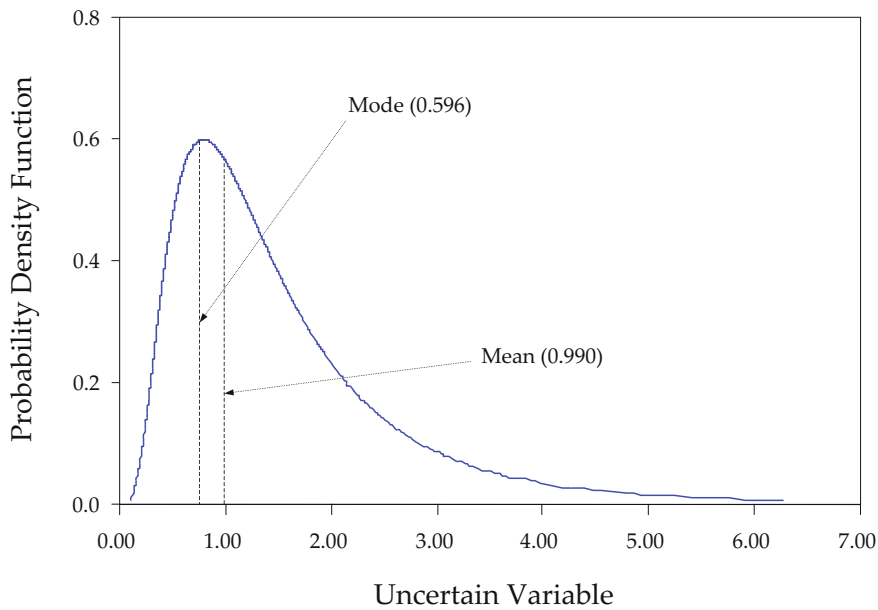


Fig. 1.2 Different probabilistic performance measures (PDF) [7]

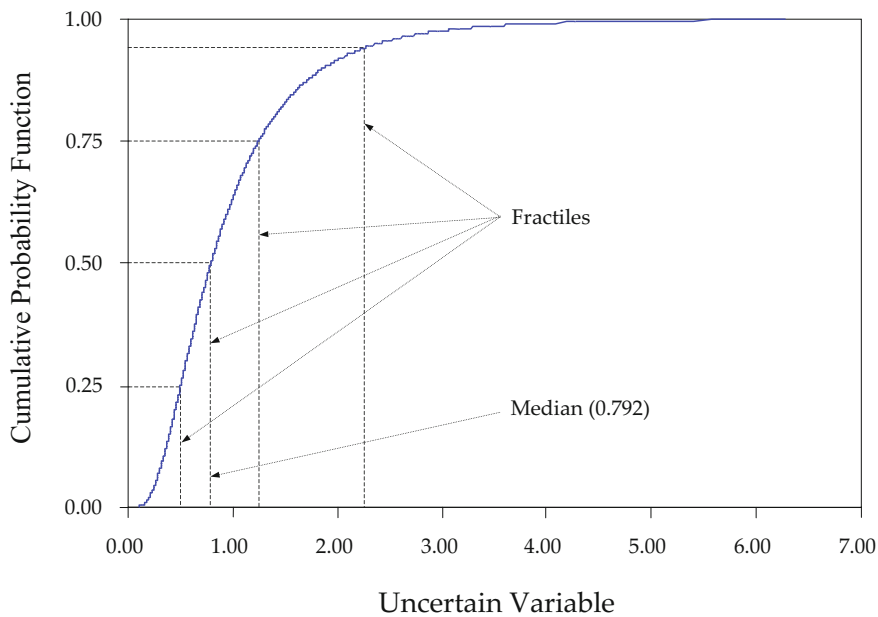
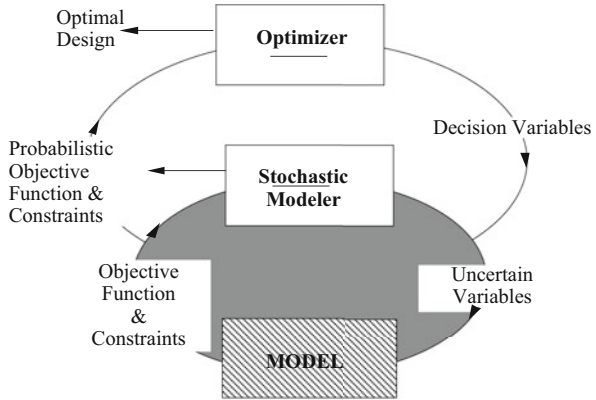


Fig. 1.3 Different probabilistic performance measures (CDF) [7]

Fig. 1.4 Pictorial representation of the stochastic programming framework



$$E(z(x, u)) = \sum_{k=1}^{N_{samp}} \frac{z(x, u_k)}{N_{samp}} \quad (1.7)$$

$$\sigma^2(z(x, u)) = \sum_{k=1}^{N_{samp}} \frac{(z(x, u_k) - \bar{z})^2}{N_{samp}} \quad (1.8)$$

where \bar{z} is the average value of z . E is the expected value and σ^2 is the variance.

1.2 Stochastic Nonlinear Programming

There are two fundamental approaches used to solve SNLP problems. The first set of techniques identify problem specific structures and transforms the problem into a deterministic NLP problem. For instance, chance constrained programming [4] replaces the constraints that include uncertainty with the appropriate probabilities expressed in terms of moments. The major restrictions in applying the chance constrained formulation include that the uncertainty distributions should be stable distribution functions, the uncertain variables should appear in the linear terms in the chance constraint, and that the problem needs to satisfy the general convexity conditions. The advantage of the method is that one can apply the deterministic optimization techniques to solve the problem.

Decomposition techniques like L-shaped decomposition [2]) divide the problem into stages and generate bounds on the objective function by changing decision variables and solving subproblems that determine the recourse action with respect to the uncertain variables. However, these methods also require convexity conditions and/or dual-block angular structures like the one shown in Fig. 1.5, and are only applicable to discrete probability distributions. For example, Lagrangian-based approaches have been applied to nonlinear SP formulations. The Lagrangian dual ascent

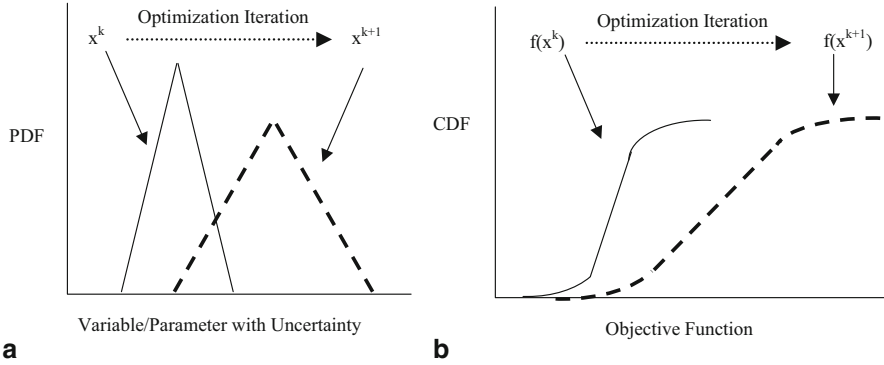


Fig. 1.6 Effect of changes in decision variables

After the model is run, the corresponding output distribution $cdf_{out}(x^k, u^k)$ is generated, shown as the solid line in Fig. 1.6b. As optimization progresses to the next iteration, $k + 1$, moments such as mean, variance, and the probability function can change for the uncertain variables, resulting in a new $\widehat{pdf}_{in}(x^{k+1}, u^{k+1})$, indicated by the dashed line in Fig. 1.6a.

The goal is to identify rapid and efficient techniques that determine an approximation of the properties of the new output distribution, $\widehat{cdf}_{out}(x^{k+1}, u^{k+1})$, given as the dashed cumulative distribution function in Fig. 1.6b. The advantage of this approach is its bypassing of the model evaluations for successive sampling (the inner loop in Fig. 1.4), which is computationally the most intensive task for optimization under uncertainty. The BONUS algorithm only uses sampling for the first iteration. Details of the algorithm are described in Chap. 4. This is followed by three chapters on application of the BONUS to real world systems. Chapter 8 presents a variant of the BONUS algorithm called L-Shaped BONUS that exploits specific structure problems. Applications of this variant are presented in the Chaps. 9 and 10.

1.3 Summary

A generalized way of solving SNLP is to use sampling-based methods. BONUS exploits the advantages of traditional NLP methods based on derivative information. It uses efficient sampling techniques and uses sampling only for the first iteration. A variant of the the BONUS algorithm, namely, the L-shaped BONUS algorithm uses specific structure of the problem for efficiency improvement. The BONUS algorithm and its variant allows for solution of large scale real world problems.