

Chapter 5

Reliability-Based Design Optimization with the Generalized Inverse Distribution Function

Domenico Quagliarella, Giovanni Petrone and Gianluca Iaccarino

Abstract This paper presents an approach to optimization under uncertainty that is very well and naturally suited to reliability-based design optimization problems and it is a possible alternative to traditional approaches to robust design based on the optimization of statistical moments. The approach shown here is based on the direct use of the generalized inverse distribution function estimated using the empirical cumulative distribution function (ECDF). The optimization approach presented is illustrated with the application to some test functions for both robust optimization and reliability-based design optimization. In the robust optimization test case, the bootstrap statistical technique is used to estimate the error introduced by the usage of the ECDF for quantile estimation.

Keywords Optimization under uncertainty · Reliability based design · Robust design · Generalized inverse distribution function · Bootstrap method

5.1 Introduction

Many industrial optimization processes must take account of the stochastic nature of the system and processes to design or re-design and consider the variability of some of the parameters that describe them. Thus it is necessary to characterize the system that is being studied from various points of view related to the treatment of uncertainty.

D. Quagliarella (✉)
Department of Fluid Mechanics, Italian Aerospace Research Center,
Via Maiorise snc, 81043 Capua, Italy
e-mail: d.quagliarella@cira.it

G. Petrone
Aerospace, Automotive and Turbo CFD Team, ANSYS UK Ltd,
Sheffield Business Park, S6 Europa View S9 1XH, UK
e-mail: giovanni.petrone@ansys.com

G. Iaccarino
Department of Mechanical Engineering, Stanford University,
Stanford, CA 94305, USA
e-mail: jops@stanford.edu

© Springer International Publishing Switzerland 2015
D. Greiner et al. (eds.), *Advances in Evolutionary and Deterministic Methods for Design, Optimization and Control in Engineering and Sciences*, Computational Methods in Applied Sciences 36, DOI 10.1007/978-3-319-11541-2_5

In particular, it is necessary to consider the sensitivity of the system to the uncertain parameters and assess its reliability. Having established the ability to characterize the system from this point of view, it is necessary to build an optimization loop that can improve its reliability, or that is capable of providing a robust optimum, or that could withstand acceptably random perturbations of design parameters or operating conditions. The classical approach to this problem is the so-called “robust design optimization” (RDO), which tries to maximize the performance and simultaneously to minimize the performance sensitivity with respect to random parameters. Instead, the “reliability based design optimization” (RBDO) tries to find the optimum design by explicitly assigning a specific level of risk and a given level of reliability. This is equivalent to assigning a quantile of the function to be optimized as the actual objective function and, for example, to minimize its value. Therefore, if the goal is that a given objective function is less than a specific value \bar{q} in 75% of possible cases, this will translate into the following constraint on the corresponding quantile: $q^{0.75} \leq \bar{q}$. Alternatively, the problem can be set as the minimization of $q^{0.75}$ and a function is thus obtained which is less than or equal to the value obtained by optimization of the quantile in 75% of cases. If, instead, the objective is $\min q^1$, then the purpose of the optimization procedure is to protect against the worst case scenario, as it happens when the problem is of minimax type.

In this work an approach to robust and reliable design optimization based on the use of the generalized inverse distribution function is presented. The robust optimization framework is illustrated and the commonly used techniques to face the problem are briefly summarized making reference to the related literature. A very simple evolutionary multi-objective optimization algorithm based on the usage of the inverse cumulative distribution function is illustrated and discussed with the help of some test problems.

5.2 Robust Optimization

Let Z be a metric space and $z \in Z$ the vector of design variables. Let also $X : \Omega \rightarrow \mathcal{E} \subseteq \mathbb{R}$ be a real valued random variable defined in a given (Ω, \mathcal{F}, P) probability space. We want to deal with an optimization problem where an objective is optimized with respect to $z \in Z$ and depends on the realizations x of X . In other terms we have:

$$y(z, X): z \in Z, X \longrightarrow Y(z)$$

with $Y(z)$ a new random variable, e.g. a new mapping of (Ω, \mathcal{F}, P) into \mathbb{R} , that depends on z . Solving an optimization problem involving $Y(z) = y(z, X)$ means that we want to find a value $\bar{z} \in Z$ such that the random variable $Y(\bar{z})$ is optimal. To establish the optimality of a given $Y(\bar{z})$ with respect to all $Y(z)$, $\forall z \in Z$, a ranking criterion must be defined such that for any couple $z_1, z_2 \in Z$ it is possible to state that $Y(z_1)$ is better or worse than $Y(z_2)$ (from now on, $Y(z_1) \preceq Y(z_2)$ will mean that $Y(z_1)$ is better or equivalent to $Y(z_2)$).

Recalling that a random variable is a measurable function, it seems natural to introduce measures that highlight particular features of the function. This leads to the classical and widely used approach of using the statistical moments to define the characteristics of the probability distribution that are to be optimized. More generally, let's consider an operator

$$\Phi_X : Y(z) = y(z, X) \in Z \times (\Omega, \mathcal{F}, P) \longrightarrow \Phi(z) \in V \subseteq \mathbb{R}$$

that translates the functional dependency on the random variable, Y , into a real valued function of z that represents a deterministic attribute of the function, $Y(z)$. This makes possible to formulate the following optimization problem

$$P_\Phi : \min_{z \in Z} \Phi(z)$$

Without loss of generality, it is possible to identify the random variable Y through its distribution function $f_Y(y)$ or its cumulative distribution function $F_Y(y)$. If $\Phi(\cdot)$ is assumed as the expected value of the objective function (\mathbb{E}), the classical formulation of first moment optimization is retrieved:

$$P_{\mathbb{E}} : \min_{z \in Z} \int_{\mathbb{R}} y f_Y(y, z) dy$$

that in terms of the CDF becomes:

$$P_{\mathbb{E}} : \min_{z \in Z} \int_{\mathbb{R}} y dF_Y(y, z)$$

It should be noted that here the distribution function depends also on z , that is the vector of the design variables.

For the purposes of the definition of the problem, it is not necessary to know exactly the distribution f_Y (or F_Y). Indeed, it is possible, as will be shown below, to use an estimate of the distribution having the required accuracy. In particular, the Empirical Cumulative Distribution Function (ECDF) will be used in this work as statistical estimator of the CDF.

The first order moment method is also called mean value approach, as the mean is used as objective to reduce the dependency on Y . This method is widely used, mostly because the mean is the faster converging moment and relatively few samples are required to obtain a good estimate. Often, however, the mean alone is not able to capture and represent satisfactorily the uncertainties embedded in a given design optimization problem. To overcome this drawback, a possible approach is the introduction in the objective function of penalization terms that are function of higher order moments. The drawback of this technique is that the ideal weights of the penalization terms are often unknown. Furthermore, in some cases, an excessive number

of higher order moments may be required to adequately capture all the significant aspect of the uncertainty embedded into a given problem. Finally, a wrong choice of the penalties may lead to a problem formulation that does not have any feasible solution. Instead of penalization terms, explicit constraints can be introduced in the robust optimization problem, and the same considerations apply for the advantages and the drawbacks of the technique.

Another possibility is the minimax criterion, very popular in statistical decision theory, according to which the worst case due to uncertainty is the objective for optimization. This ensures protection against worst case scenario, but it is often excessively conservative.

The multi-objective approach [7] based on constrained optimization is also widely adopted. Here different statistical moments are used as independent trade-off objectives. The obtained Pareto front allows an a-posteriori choice of the optimal design between a set of equally ranked candidates. In this case a challenge is posed by the increase in the dimensionality of the Pareto front when several statistical moments are used. The research related to the multi-objective method has led to several extensions of the classical Pareto front concept. In [10], for example, the Pareto front exploration in presence of uncertainties is faced introducing the concept of *probabilistic dominance*, which is an extension of the classical *Pareto dominance*. While in [6], a probabilistic ranking and selection mechanism is proposed that introduces the *probability of wrong decision* directly in the formula for rank computation.

An interesting approach, similar in some aspects to the one here described, is found in [5] where a quantile based approach is coupled with the probability of *Pareto nondominance* (already seen in [6]). Here, contrary to the cited work, the optimization technique introduced relies on direct estimation of the quantile function obtained through the Empirical Cumulative Distribution Function.

5.3 The Generalized Inverse Distribution Function Method

In the methodology presented herein, the operator that is used to eliminate the dependence on random variables is the quantile function of the objective function to be optimized, calculated in one or more points of its domain of definition.

Before going into the details of the exposure, the definitions of *Cumulative Distribution Function* (CDF) and *Generalized Inverse Distribution Function* (GIDF) that will be used are reported.

The “cumulative distribution function” (CDF), or just “distribution function”, describes the probability that a real-valued random variable Q with a given probability distribution will be found at a value less than or equal to q . Intuitively, it is the “area so far” function of the probability distribution. The CDF is one of the most precise, efficient and compact ways to represent information about uncertainty, and a new CDF based approach to robust optimization is described.

If the CDF is continuous and strictly monotonic then it is invertible, and its inverse, called quantile function or inverse distribution function, returns the value below

which random draws from the given distribution would fall, $s \times 100$ percent of the time. That is, it returns the value of q such that

$$F_Q(q) = \Pr(Q \leq q) = s \tag{5.1}$$

Hence $F^{-1}(s)$, $s \in [0, 1]$ is the unique real number q such that $F_Q(q) = s$.

Unfortunately, the distribution does not, in general, have an inverse. If the probability distribution is discrete rather than continuous then there may be gaps between values in the domain of its CDF, while, if the CDF is only weakly monotonic, there may be “flat spots” in its range. In general, in these cases, one may define, for $s \in [0, 1]$, the “generalized inverse distribution function” (GIDF)

$$q^s = Q(s) = F_Q^{-1}(s) = \inf \{q \in \mathbb{R} : F(q) \geq s\}$$

that returns the minimum value of s for which the previous probability statement (5.1) holds. The infimum is used because CDFs are, in general, weakly monotonic and right-continuous (see [15]).

Now that the CDF and the GIDF have been introduced, it becomes easy to define, within the framework of multi-objective optimization, a robust optimization problem in terms of an arbitrary number of quantiles to optimize:

$$P_{Q(s_i)}: \min_{z \in Z} q^{s_i}(z) = \min_{z \in Z} \inf \{q(z) \in \mathbb{R} : F_Q(q(z)) \geq s_i\} \quad i = 1, \dots, n \tag{5.2}$$

where n is the number of objectives chosen. The approach, then, can be further extended by introducing objectives that are arbitrary functions of quantiles.

Of course, the problem now is focused on how to satisfactorily calculate the quantiles required by the method. In this work the *Empirical Cumulative Distribution Function* (ECDF) is used for this purpose. The definition of ECDF, taken from [16], is reported here for the sake of completeness.

Let X_1, \dots, X_n be random variables with realizations $x_i \in \mathbb{R}$, the empirical distribution function is an indicator function that estimates the true underlying CDF of the points in the sample. It can be defined by using the order statistics $X_{(i)}$ of X_i as:

$$\widehat{F}_n(x, \omega) = \begin{cases} 0 & \text{if } x < x_{(1)}; \\ \frac{1}{n} & \text{if } x_{(1)} \leq x < x_{(2)}, 1 \leq k < 2; \\ \frac{2}{n} & \text{if } x_{(2)} \leq x < x_{(3)}, 2 \leq k < 3; \\ \vdots & \\ \frac{i}{n} & \text{if } x_{(i)} \leq x < x_{(i+1)}, i \leq k < i + 1; \\ \vdots & \\ 1 & \text{if } x \geq x_{(n)}; \end{cases}$$

where $x_{(i)}$ is the realization of the random variable $X_{(i)}$ with outcome (elementary event) $\omega \in \Omega$.

From now on, therefore, when the optimization algorithm requires the calculation of the $F_Q(s)$, it will use instead its estimator $\widehat{F}_{Q_n}(s)$, where n indicates the number of samples used to estimate this ECDF.

Note that each indicator function, and hence the ECDF, is itself a random variable. This is a very delicate issue to consider. Indeed, if the EDCF is used to approximate the deterministic operator $Q(s)$, a direct residual influence of the random variables that characterize the system under investigation remains on $P_{Q(s)}$. In other words $Q(s)$ behaves as a random variable, but with the important difference that its variance tends to zero when the ECDF approximates the CDF with increasing precision. It is possible to demonstrate that the estimator $\widehat{F}_{Q_n}(s)$ is consistent, as it converges almost surely to $F_Q(s)$ as $n \rightarrow \infty$, for every value of s [11]. Furthermore, for the Glivenko-Cantelli theorem [8], the convergence is also uniform over s . This implies that, if the ECDF is calculated with sufficient accuracy, it can be considered and treated as a deterministic operator. On the other hand, if the number of samples, or the estimation technique of the ECDF, do not allow as such, one can still correlate the variance of the ECDF with the precision of the obtained estimate. Of course, if the ECDF is estimated in a very precise way, it is possible to use an algorithm conceived for deterministic problems, provided that it has a certain resistance to noise. Conversely, if the ECDF is obtained from a coarse sample, its practical use is only possible with optimization algorithms specifically designed for that purpose.

For the same reason, it is often convenient, especially in applications where the ECDF is defined with few samples, to use q^ε instead of q^0 , with $\varepsilon > 0$ and small, but such that a not excessive variance of the estimate of q^ε is ensured.

5.4 A Robust Optimization Test Case

The function reported in Table 5.1, taken from [13], is used as a benchmark to test the GIDF based approach to robust optimization. With respect to the function reported in the reference, the following changes have been introduced: the ranges of design and uncertain parameters have been changed as reported in table, and a multiplicative factor equal to $1/n$ has been introduced to make easier the result comparison when the dimension of the parameter space changes. The random variables \mathbf{u} have a uniform distribution function. Table 5.2 reports the solutions to the optimization problems

$$\min_{\mathbf{d} \in D, \mathbf{u} \in U} f(\mathbf{d}, \mathbf{u})$$

$$\min_{\mathbf{d} \in D} \max_{\mathbf{u} \in U} f(\mathbf{d}, \mathbf{u})$$

over the cartesian product of D and U . The first problem represents the best possible solution obtainable if the \mathbf{u} are considered as design parameters varying in

Table 5.1 Benchmark functions table

ID	Function	Ranges	Dimension
MV4	$f = \frac{1}{n} \sum_{i=1}^n (2\pi - u_i) \cos(u_i - d_i)$	$\mathbf{u} \in [0, 3]^n, \mathbf{d} \in [0, 2\pi]^n$	1, 2 and 6

Table 5.2 Benchmark functions table results

ID	$\min_{\mathbf{d} \in D, \mathbf{u} \in U} f(\mathbf{d}, \mathbf{u})$			$\min_{\mathbf{d} \in D} \max_{\mathbf{u} \in U} f(\mathbf{d}, \mathbf{u})$		
	\mathbf{d}	\mathbf{u}	f	\mathbf{d}	\mathbf{u}	f
MV4	$[3.1416]^n$	$[0]^n$	$-6.283185 \dots$	$[4.6638]^n$	$[0]^n$	$-0.305173 \dots$

U . The second one, instead, minimizes the maximum possible loss or, alternatively, maximizes the minimum gain, according to the framework of decision theory [12]. These solutions have been obtained analytically and verified by exhaustive search for $n = 1$. It is worth to note that these particular optimal solutions are the same whatever is the dimension of the search space.

The optimization algorithm used here is a simple multi-objective genetic algorithm not specially conceived for optimization under uncertainty. The algorithm is based on the Pareto dominance concept and on local random walk selection [14]. The crossover operator is the classical *one-point* crossover which operates at bit level, while mutation operator works at the level of the design vector parameters (which are real numbers). A parameter, called *mutation rate* controls the operator activation probability for each variable vector element, while a further parameter, called *strength*, is the maximum relative value for the uniform word mutation. The word mutation value is given by $\text{strength} \cdot (r - 0.5)(u - l)$ where $r \in [0, 1]$ is the uniform random number, u is the upper variable bound and l is the lower variable bound. An elitist strategy was adopted in the optimization runs. It consists in replacing 20% of the population calculated at each generation with elements taken at random from the current Pareto front. Obviously, the elements of the population are used to update the current Pareto front before the replacement, in order to avoid losing non-dominated population elements.

The multi-objective runs were performed using 100% crossover activation probability and word mutation with *mutation rate* equal to 50% and *strength* equal to 0.06. The initial population was obtained using the quasi-random low-discrepancy Sobol sequence [1]. The ECDF used to estimate the CDF was obtained with 2,500 Montecarlo samples in all runs. The population size was set to 4,000 elements for all runs, while the number of generations was set to 10 for $n = 1$, 200 for $n = 2$ and 1,000 for $n = 6$. The problem solved was $\min_{\mathbf{z} \in Z} (q^\varepsilon, q^{1-\varepsilon})$.

Figure 5.1 reports the Pareto fronts and the deterministic min and min max solutions obtained for the MV4 test case at different values of the design space size n . It can be easily observed that, in the case $n = 1$, the extremes of the front are practically coincident with the deterministic solutions, while, in the case $n = 2$, the

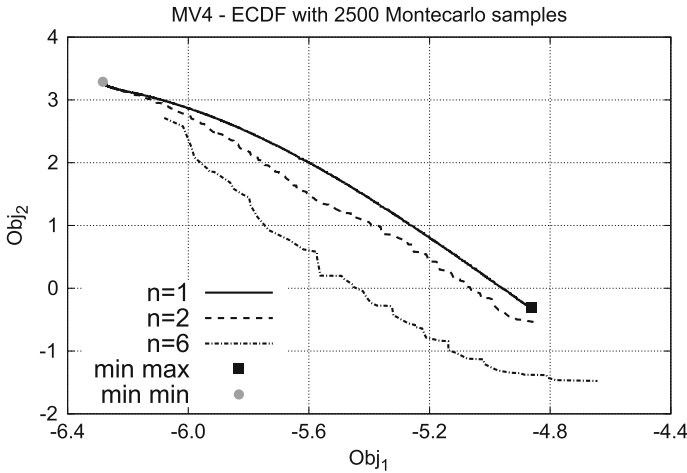


Fig. 5.1 Pareto fronts and deterministic min and min max solutions for the MV4 test case

solution of the Pareto front which minimizes the second objective ($q^{1-\varepsilon}$) underestimates the min max solution. The trend is even more evident in the case $n = 6$, where also the extreme of the front that minimizes the first goal (q^ε) overestimates the value obtained from the min problem. This can be explained by the fact that the two deterministic solutions are located in correspondence with the extremes of variation of the random variables of the problem. Therefore, as the number of random variables increases, in accordance with the central limit theorem [9], it becomes less likely that all random variables are located in correspondence of one of their limits of variation. However, as illustrated in Fig. 5.2, when the Pareto front obtained with the sample size m equal to 2,500 is re-evaluated with a larger Montecarlo sample, the curve is a quite acceptable approximation of the Pareto front obtained with $m = 100,000$.

Figures 5.3 and 5.4 show the ECDF corresponding to the extremes of the Pareto front, respectively for the cases $n = 1$ and $n = 6$. It is noted, again in accordance with the central limit theorem, that, in the case $n = 6$, the ECDF curves are very close to those related to a Gaussian distribution.

5.5 Evaluating and Improving the Quantile Estimation

The example in the previous section shows very clearly that the results of the proposed method may depend in an essential way on the quality of estimation of quantiles that is obtained through the ECDF. This leads in a natural way to deal with two issues: how to evaluate the quality of the estimation of the quantiles used in the multi-objective optimization problem, and how to possibly get a better quantile estimate with a given computational effort.

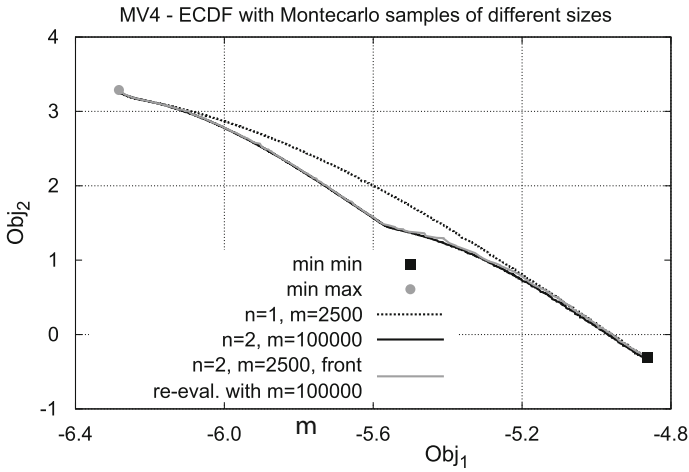


Fig. 5.2 Pareto fronts for the MV4 test case obtained with different sizes for Montecarlo sampling

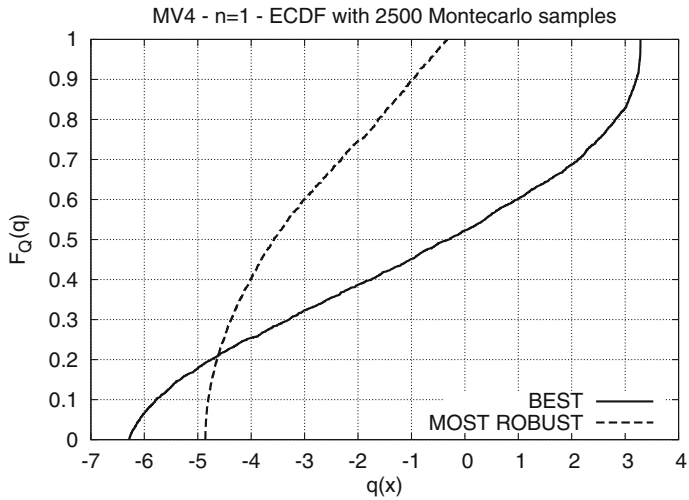


Fig. 5.3 Optimal ECDF curves for the MV4 with $n = 1$

The approach here proposed for assessing the quality of the estimate of the quantile is based on the bootstrap method introduced by Efron in 1977 [3, 4].

This method represents a major step forward in the statistical practice because it allows to accurately assess the variability of any statistical estimator without making any assumption about the type of distribution function involved. Suppose that a statistic $T(x_1, x_2, \dots, x_n)$ is given, evaluated on a set of data x_1, x_2, \dots, x_n belonging to an assigned space X . The bootstrap essentially consists in the repeated recalculation of the statistic T employing a tuple of new samples $x_1^*, x_2^*, \dots, x_n^*$

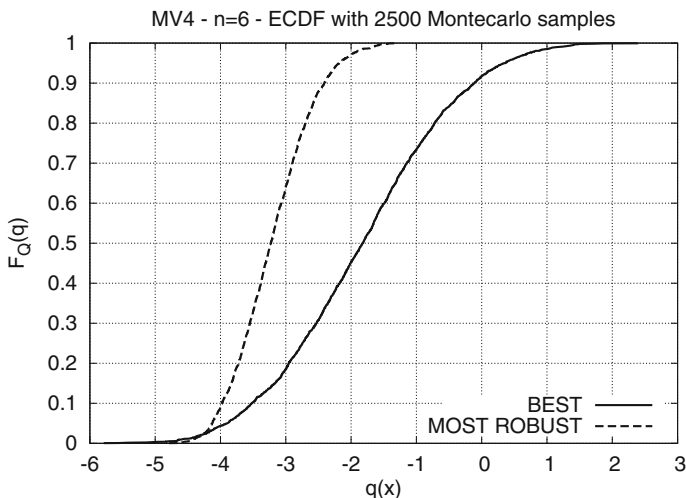


Fig. 5.4 Optimal ECDF curves for the MV4 with $n = 6$

obtained by selecting them from the collection $\{x_1, x_2, \dots, x_n\}$ by replacement. The repeated calculation of $T(x_1^*, x_2^*, \dots, x_n^*)$ gives a set of values that is a good indication of the distribution of T .

Therefore, to calculate the accuracy of a generic quantile q^s , obtained by the estimator $\widehat{F}_{Q_n}(s)$, the bootstrap procedure can be applied to the samples that define the estimator. This allows to calculate the corresponding distribution of q^s for a fixed value of s .

Figure 5.5 reports the ECDF related to the solution labeled as “MOST ROBUST” in Fig. 5.4. The bootstrap was applied to this ECDF repeating the sampling process 2,000 times. The area in gray color represents the superposition of all the curves obtained in this way. From the bootstrap data it is then possible to evaluate the accuracy of a given quantile estimate. According to [3], an accuracy measure for q^s can be obtained considering the central 68% of bootstrap samples. These values lay between an interval $[q_\ell^s, q_u^s]$ centered on the observed value q^s . Half the length of this interval gives a measure of the accuracy for the observed value that corresponds to the traditional concept of “standard error”. Here this value is indicated with \widehat{SE} to distinguish it from the true standard error SE .

Table 5.3 reports the computed accuracy values for the considered quantiles for the above cited “MOST ROBUST” solution obtained from an ECDF with 2,500 Montecarlo samples. The fourth column reports, finally, the maximum estimated error \widehat{ME} .

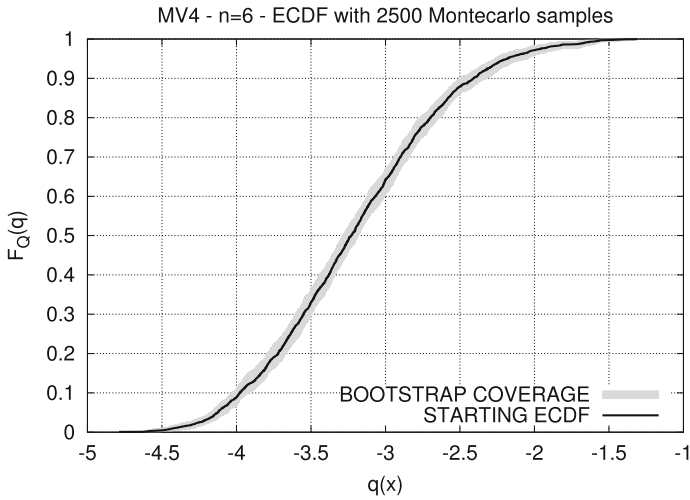


Fig. 5.5 ECDF corresponding to the most robust solution and related bootstrap coverage

Table 5.3 Quantile estimates and related accuracy for MV4 “MOST ROBUST” solution with $n = 6$

s	q^s	\widehat{SE}	\widehat{ME}
0.001000	-4.630433	± 0.090423	± 0.117169
0.500000	-3.230388	± 0.018834	± 0.054983
0.999000	-1.425868	± 0.013192	± 0.136330

5.6 Single and Multi-objective Reliability Optimization Tests

A very complete review of reliability-based optimization and of the solution techniques based on evolutionary algorithms can be found in [2].

A reliability based single or multi-objective problem can be written as follows:

$$\begin{cases} \min_{\mathbf{z}, \mathbf{x}} \mathbf{f}(\mathbf{x}, \mathbf{z}) \\ \text{s. to: } g_i(\mathbf{z}, \mathbf{x}) \leq 0 \quad i = 1, \dots, n \\ \mathbf{x} \in \mathbf{X}, \quad \mathbf{z} \in \mathbf{Z} \end{cases} \quad (5.3)$$

with \mathbf{x} representing deterministic and \mathbf{z} stochastic design variables. In order to find a solution to this problem, a reliability measure R is introduced that means that the probability of having an infeasible solution instance will be at most $(1 - R)$, and the problem is reduced to the following one:

$$\begin{cases} \min_{\mathbf{z}, \mathbf{x}} \mathbf{f}(\mathbf{x}, \mu_{\mathbf{z}}) \\ \text{s. to: } P_i(g_i(\mathbf{z}, \mathbf{x}) \leq 0) \geq R \quad i = 1, \dots, n \\ \mathbf{x} \in \mathbf{X}, \quad \mathbf{z} \in \mathbf{Z} \end{cases} \quad (5.4)$$

where $\mu_{\mathbf{z}}$ is the mean vector of the uncertain variables \mathbf{z} , and $P_i()$ is the probability that the i th constraint is feasible. The constraints in problem (5.4) can be immediately expressed in terms of generalized inverse distribution function:

$$\begin{cases} \min_{\mathbf{z}, \mathbf{x}} \mathbf{f}(\mathbf{x}, \mu_{\mathbf{z}}) \\ \text{s. to: } q^R(g_i(\mathbf{z}, \mathbf{x})) \leq 0 \quad i = 1, \dots, n \\ \mathbf{x} \in \mathbf{X}, \quad \mathbf{z} \in \mathbf{Z} \end{cases} \quad (5.5)$$

Problem 5 can be further transformed by introducing penalty functions into:

$$\min_{\mathbf{z}, \mathbf{x}} \mathbf{f}(\mathbf{x}, \mu_{\mathbf{z}}) + \sum_{i=1}^n \Gamma(q^R(g_i(\mathbf{z}, \mathbf{x}))) \quad (5.6)$$

with Γ defined, for example, as:

$$\Gamma(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ kx^2, \quad k > 0 & \text{if } x > 0 \end{cases} \quad (5.7)$$

Considering that $\Gamma(q^R = 0) = q^R(\Gamma(g_i(\mathbf{z}, \mathbf{x})))$ it is possible to write, finally:

$$\min_{\mathbf{z}, \mathbf{x}} q^R\left(\mathbf{f}(\mathbf{x}, \mu_{\mathbf{z}}) + \sum_{i=1}^n \Gamma(g_i(\mathbf{z}, \mathbf{x}))\right) \quad (5.8)$$

In the subsequent examples, problem (5.8) is solved using a simple multi-objective genetic algorithm for unconstrained problems. Of course, more sophisticated and efficient algorithms could be used to reduce the computational effort.

The first example considered is taken from [2]. It involves two variables, x and y , and two objectives, f_1 and f_2 :

$$\begin{cases} \min & f_1 = x \\ \min & f_2 = \frac{1+y}{x} \\ \text{s. to:} & y + 9x - 6 \geq 0 \\ & -y + 9x - 1 \geq 0 \\ & 0.1 \leq x \leq 1, 0 \leq y \leq 5 \end{cases} \quad (5.9)$$

with both variables uncertain and characterized by a normal distribution with $\sigma = 0.03$.

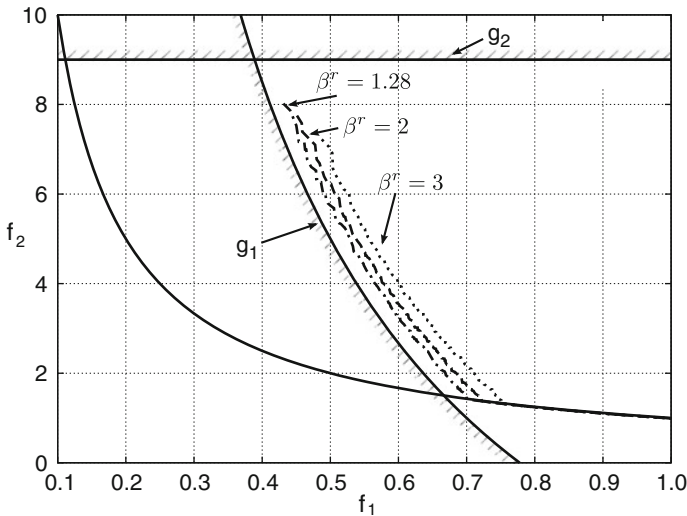


Fig. 5.6 Pareto front for the simple two-objectives reliability problem obtained with three different reliability indexes

The optimization runs have been carried out with three different reliability levels, namely 90 % ($\beta^r = 1.28$), 97.725 % ($\beta^r = 2.00$) and 99.875 % ($\beta^r = 3.00$), where β^r is the reliability index computed according to $R = 50 \left(1 + \operatorname{erf} \left(\beta^r / \sqrt{2} \right) \right)$. The results are reported in Fig. 5.6. The multi-objective genetic algorithm crossover and mutation settings are the same that have been used in the previous test case. For each different reliability level, a population of 100 individuals evolved for 400 generations. For $\beta^r = 1.28$ and $\beta^r = 2.00$ the Montecarlo sample size was set to 1,000, while for the $\beta^r = 3.00$ case the sample size was set to 10,000.

The second problem presented is a car side-impact problem related to vehicle crash-worthiness and dealing with vehicle safety rating scores related to human safety issues. The problem is reported in [17] and, in a slightly different form, in [2]. The reader is referred to the references above for more details on the physical nature of the model used and the role and significance of variables and parameters. In extreme synthesis, eleven random variables are involved in the optimization problem, whose characteristics and nature is reported in Table 5.4 for the sake of completeness. The random variables z_{10} and z_{11} are not regarded as design variables, because they are related to aspects that define the type of the problem but which can not be controlled in any way by the designer. The objective function and constraints of the problem are given below:

Table 5.4 Properties of design and random parameters of vehicle side impact model

Random variable	Std dev.	Distr type.	Lower limit	Upper limit
z_1 (B-pillar inner)	0.030	Normal	0.500	1.500
z_2 (B-pillar reinforce)	0.030	Normal	0.500	1.500
z_3 (Floor side inner)	0.030	Normal	0.500	1.500
z_4 (Cross member)	0.030	Normal	0.500	1.500
z_5 (Door beam)	0.030	Normal	0.500	1.500
z_6 (Door belt line)	0.030	Normal	0.500	1.500
z_7 (Roof rail)	0.030	Normal	0.500	1.500
z_8 (Mat. B-pillar inner)	0.006	Normal	0.192	0.345
z_9 (Mat. floor side inner)	0.006	Normal	0.192	0.345
z_{10} (Barrier height)	10.0	Normal	-30.0	30.0
z_{11} (Barrier hitting)	10.0	Normal	-30.0	30.0

$$\begin{aligned}
 f(\mathbf{z}) &= 1.98 + 4.9z_1 + 6.67z_2 + 6.98z_3 + 4.01z_4 + 1.78z_5 + 2.73z_7 \\
 g_1(\mathbf{z}) &= 1.16 - 0.3717z_2z_4 - 0.00931z_2z_{10} - 0.484z_3z_9 \\
 &\quad + 0.01343z_6z_{10} \leq 1 \text{ kN} \\
 g_2(\mathbf{z}) &= 0.261 - 0.0159z_1z_2 - 0.188z_1z_8 - 0.019z_2z_7 \\
 &\quad + 0.0144z_3z_5 + 0.0008757z_5z_{10} + 0.08045z_6z_9 + 0.00139z_8z_{11} \\
 &\quad + 0.00001575z_{10}z_{11} \leq 0.32 \text{ m/s} \\
 g_3(\mathbf{z}) &= 0.214 + 0.00817z_5 - 0.131z_1z_8 - 0.0704z_1z_9 \\
 &\quad + 0.03099z_2z_6 - 0.018z_2z_7 + 0.0208z_3z_8 + 0.121z_3z_9 \\
 &\quad - 0.00364z_5z_6 + 0.0007715z_5z_{10} \\
 &\quad - 0.0005354z_6z_{10} + 0.00121z_8z_{11} \leq 0.32 \text{ m/s} \\
 g_4(\mathbf{z}) &= 0.74 - 0.61z_2 - 0.163z_3z_8 \\
 &\quad + 0.001232z_3z_{10} - 0.166z_7z_9 + 0.227z_2z_2 \leq 0.32 \text{ m/s} \\
 g_5(\mathbf{z}) &= 28.98 + 3.818z_3 - 4.2z_1z_2 + 0.0207z_5z_{10} \leq 32 \text{ mm} \\
 &\quad + 6.63z_6z_9 - 7.7z_7z_8 + 0.32z_9z_{10} \\
 g_6(\mathbf{z}) &= 33.86 + 2.95z_3 + 0.1792z_{10} - 5.057z_1z_2 \\
 &\quad - 11z_2z_8 - 0.0215z_5z_{10} - 9.98z_7z_8 + 22z_8z_9 \leq 32 \text{ mm} \\
 g_7(\mathbf{z}) &= 46.36 - 9.9z_2 - 12.9z_1z_8 + 0.1107z_3z_{10} \leq 32 \text{ mm} \\
 g_8(\mathbf{z}) &= 4.72 - 0.5z_4 - 0.19z_2z_3 - 0.0122z_4z_{10} + 0.009325z_6z_{10} \\
 &\quad + 0.000191z_{11}z_{11} \leq 4 \text{ kN} \\
 g_9(\mathbf{z}) &= 10.58 - 0.674z_1z_2 - 1.95z_2z_8 + 0.02054z_3z_{10} \\
 &\quad - 0.0198z_4z_{10} + 0.028z_6z_{10} \leq 9.9 \text{ mm/ms} \\
 g_{10}(\mathbf{z}) &= 16.45 - 0.489z_3z_7 - 0.843z_5z_6 + 0.0432z_9z_{10} \\
 &\quad - 0.0556z_9z_{11} - 0.000786z_{11}z_{11} \leq 15.7 \text{ mm/ms}
 \end{aligned}
 \tag{5.10}$$

The goal is to minimize the weight of the structure.

Again, the same settings were used for crossover and mutation and, in all cases analyzed, the size of the population and was fixed to 30 individuals that have evolved over 3,000 generations. The Montecarlo sample size was set to 100,000. Figure 5.7

Best f as a function of $R(\%)$

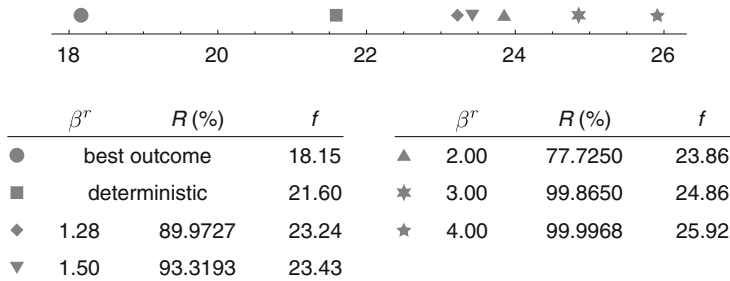


Fig. 5.7 Best optimization result as a function of the reliability index for the crash-worthiness vehicle side impact problem

shows, both in tabular and number line format, the values of the objective function for different reliability values (expressed both in percentage and in β^r terms). These values are compared both with the deterministic case, calculated by setting the parameters z_{10} and z_{11} to zero (average value), that with the case here called “best outcome”, obtained by considering also the two parameters z_{10} and z_{11} as deterministic variables of the problem.

5.7 Conclusions

An alternative approach to the optimization under uncertainty has been introduced and illustrated with examples related to both robust and reliability based design optimization. Furthermore, a method, based on the well known bootstrap statistical technique, has been introduced to estimate the error introduced by the usage of the ECDF instead of the true CDF.

The algorithm used for optimization is a classical genetic algorithm, but, to further improve the proposed procedure, an optimization algorithm capable of accounting for the errors in the estimation of the CDF has to be conceived. This is a very important topic and it will be subject of next research work. In particular, to reduce the curse of dimensionality, the effect of different sampling methodologies, like stochastic collocation, on the estimation of the ECDF will be considered in future works. Indeed the possibility to use the error on the ECDF estimator to properly refine the probability space using adaptive uncertainty quantification algorithms will be explored.

References

1. Bratley P, Fox BL (1988) Algorithm 659: implementing sobol's quasirandom sequence generator. *ACM Trans Math Softw* 14(1):88–100. doi:[10.1145/42288.214372](https://doi.org/10.1145/42288.214372), <http://doi.acm.org/10.1145/42288.214372>
2. Deb K, Gupta S, Daum D, Branke J, Mall A, Padmanabhan D (2009) Reliability-based optimization using evolutionary algorithms. *Evol Comput IEEE Trans* 13(5):1054–1074
3. Diaconis P, Efron B (1983) Computer intensive methods in statistics. Technical Report 83, Division of Biostatistics, Stanford University
4. Efron B (1979) Bootstrap methods: another look at the jackknife. *Ann Stat* 7(1):1–26. <http://projecteuclid.org/euclid.aos/1176344552>
5. Filomeno CR, Bouillard P (2011) Multi-objective reliability-based optimization with stochastic metamodels. *Evol Comput* 19(4):525–560
6. Hughes E (2001) Evolutionary multi-objective ranking with uncertainty and noise. *Evolutionary multi-criterion optimization*. Springer, Berlin, pp 329–343
7. Poloni C, Padovan L, Parussini L, Pieri S, Pediroda V (2004) Robust design of aircraft components: a multi-objective optimization problem. In: Von Karman Institute for Fluid Dynamics, Lecture Series. Von Karman Institute for Fluid Dynamics, Lecture Series 2004–07
8. Serfling RJ (2008) Approximation theorems of mathematical statistics. Wiley. doi:[10.1002/9780470316481](https://doi.org/10.1002/9780470316481), <http://dx.doi.org/10.1002/9780470316481.indsub>
9. Sobol IM (1994) A primer for the monte Carlo method. CRC Press, Boca Raton
10. Teich J (2001) Pareto-front exploration with uncertain objectives. *Evolutionary multi-criterion optimization*. Springer, Heidelberg, pp 314–328
11. van der Vaart AW (1998) Asymptotic statistics. Cambridge University Press, Cambridge. <http://dx.doi.org/10.1017/CBO9780511802256>
12. von Neumann J, Morgenstern O (1953) *Theory of games and economic behavior*. Princeton University Press, Princeton
13. Vasile M, Minisci E (2011) An evolutionary approach to evidence-based multi-disciplinary robust design optimisation. In: Poloni C, Quagliarella D, Periaux J, Gauger N, Giannakoglou K (eds.) EUROGEN 2011 PROCEEDINGS—Evolutionary and deterministic methods for design, optimization and control with applications to industrial and societal problems, ECCOMAS thematic conference. ECCOMAS, CIRA, Capua, Italy
14. Vicini A, Quagliarella D (1997) Inverse and direct airfoil design using a multiobjective genetic algorithm. *AIAA J* 35(9):1499–1505
15. Wikipedia: Cumulative distribution function. http://en.wikipedia.org/wiki/Cumulative_distribution_function
16. Woo C, Empirical distribution function (version 4). <http://planetmath.org/EmpiricalDistributionFunction.html>
17. Youn B, Choi K, Yang RJ, Gu L (2004) Reliability-based design optimization for crashworthiness of vehicle side impact. *Struct Multi Optim* 26(3–4):272–283. doi:[10.1007/s00158-003-0345-0](https://doi.org/10.1007/s00158-003-0345-0), <http://dx.doi.org/10.1007/s00158-003-0345-0>