Chapter Six

Basics of Uncertainty Analysis

6.1 Introduction

As shown in Fig. 6.1, analysis models are used to predict the performances or behaviors of a product under design. In what follows, we will call the model $Y = g(\mathbf{X}) = g(X_1, X_2, \cdots, X_n)$ performance function. The performance function $Y = g(\mathbf{X})$ specifies the relationship between the input $\mathbf{X}$ and output Y. As we discussed in Chapter one, there are various uncertainties associated with the model input. The uncertainty of the model input will be propagated through the model to the model output. It is essential to quantify the uncertainty associated with the model output in order to accommodate and manage the uncertainty in the model output.

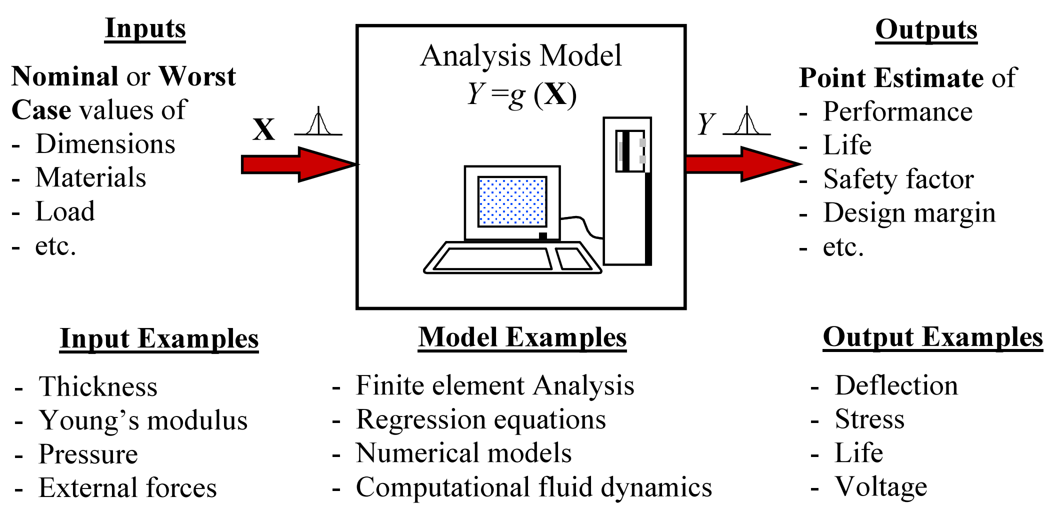

Figure 6.1 Analysis Model with Uncertain Input

The task of uncertainty analysis is to quantify the uncertainty associated with the model output. Uncertainty analysis helps engineers understand how uncertainty of the model input impacts the uncertainty of the model output (product performance). Only with this understanding, engineers are able to manage and mitigate the effects of uncertainty by choosing appropriate design variables (part of model inputs $\mathbf{X}$) during the design process. Through uncertainty analysis on an existing design, engineers can evaluate if the design satisfies all the requirements in the presence of uncertainty. For example, engineers will be able to know if the design is robust and if the design is safe. If the design is not satisfactory, the uncertainty analysis will provide engineers with useful guidance to improving the design. Therefore, uncertainty analysis is an important and imperative stage for design under uncertainty.

Before discussing various methods of uncertainty analysis, we will first introduce important concepts in uncertainty analysis, such as those of reliability and robustness. We will then provide an overview of the uncertainty analysis methods.

6.2 Uncertainty Analysis Model

The primary task of uncertainty analysis is to find the probabilistic characteristics of response variables (model outputs). The probabilistic characteristics include

- Cumulative distribution function (*cdf*)
- Probability density function (*pdf*)
- Moments such as mean, standard deviation, skewness, and correlation coefficient
- Percentile values and median

Mathematically, the uncertainty analysis can be formulated as

Given	1) Distributions of input variables $\mathbf{X} = (X_1, X_2, \cdots, X_n)$, including the *pdf* $f_{X_i}(x_i)$ or *cdf* $F_{X_i}(x_i)$, $i = 1, 2, \cdots, n$, and their joint *pdf* or *cdf*
	2) Performance function $Y = g(\mathbf{X})$
Find	Probabilistic characteristics of Y, including
	1) *cdf* $F_Y(y)$
	2) *pdf* $f_Y(y)$
	3) Moments: mean μ_Y, standard deviation σ_Y, etc
	4) α percentile values Y^α

Not all the probabilistic characteristics specified above need to be identified in uncertainty analysis. Depending on specific applications, different probabilistic characteristics may be used. For example, for robust design, only the mean and standard deviation may be used.

In some cases, we do not have adequate information about the distributions of input random variables, and we only know the moments (for example, mean and standard deviation). Then the uncertainty analysis problem becomes: Find the moments of the response variable $Y = g(X_1, X_2, \cdots, X_n)$ given the moments of the input random variables $\mathbf{X} = (X_1, X_2, \cdots, X_n)$.

6.3 Reliability

Reliability is the probability of a product performing its intended function over its specified period of usage, and under specified operating conditions, in a manner that meets or exceeds customer expectations.

The reliability methods can be roughly classified into two major types – *math-based reliability* and *physics-based reliability*. In the math-based reliability, the reliability of systems or components is evaluated based on testing. Systems and especially components are typically tested until they fail; the time to failure and failure modes are then recorded. The life related information is also commonly collected from field data. Then the statistical analysis is used to evaluate the system or component reliability, assess the risk, and improve reliability. The "empirical" and "testing" approaches of math-based reliability and their practical applications have existed in engineering fields for many years.

Mathematically, the math-based reliability is defined by the function $R(t)$, which represents the probability of success for a component or system in the time interval $(0,t)$. In other words, the reliability at the instant of time t is the probability of the life T greater than t, namely,

$$R(t) = \mathrm{P}\left(T > t\right) \tag{6.1}$$

The probability of failure is the complement of the reliability, i.e.

$$p_f(t) = 1 - R(t) \tag{6.2}$$

Examples 2.5 and 2.6 in Chapter 2 presented the applications of math-based reliability in system reliability analysis. The reliability of each of the individual components is estimated from testing or filed data. Then the system reliability is estimated according to the component reliability and the logical relationship between the system and its components.

With the fact that engineering design has evolved considerably, we are able to have a new paradigm shift – *physics-based reliability* with the aid of today's synthesis and simulation approaches. In this sense, reliability can be computationally evaluated with physical equations (models) or computer simulations that specify the state of failure. For example, as demonstrated in Fig. 6.2 for an automotive vehicle in a "virtual" world, simulation models such as linear and nonlinear finite element analyses are created to predict the behaviors (including the failure events and crashes), and the reliability can be actually calculated based on the simulation models. We will primarily focus on physics-based reliability in this book.

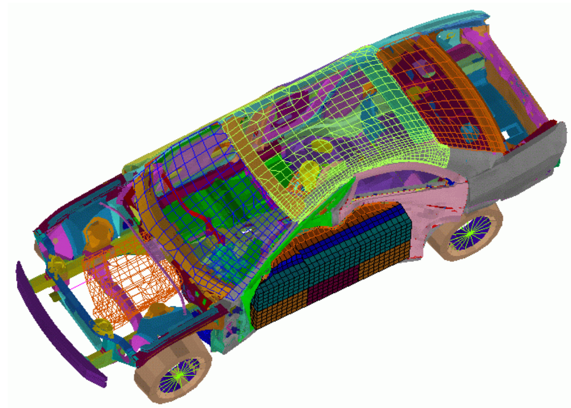

Figure 6.2 A Vehicle Impact FEA Model (Courtesy of Ford Motor Company)

As mentioned previously, the model $Y = g(\mathbf{X})$ specifies the relationship between the performance (response) Y and the input random variables $\mathbf{X}$. When the performance reaches a certain threshold, the state of the component or system will change from safety to failure. The threshold value is called a limit state. If we use a threshold of zero as the limit state, then $Y = g(\mathbf{X}) = 0$ divides the random variable space into safe and failure (unsafe) regions. When $g(\mathbf{X}) > 0$, the product is considered safe, and when $g(\mathbf{X}) < 0$, the product can no longer fulfill the function for which it was designed and therefore is considered as unsafe. Because of the above reason, performance function $Y = g(\mathbf{X})$ is also called a *limit-state function* in the area of reliability analysis and reliability-based design. For convenience, we will use a threshold value of zero in the following discussions. Fig. 6.3 shows the limit state for a two dimensional problem.

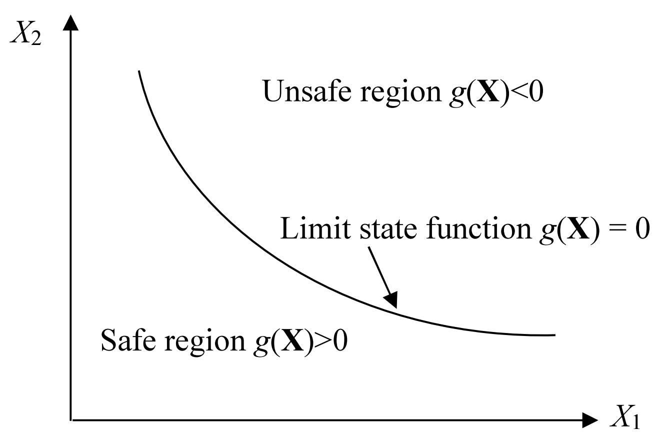

Figure 6.3 The Concept of Limit State

For example, if the performance function $Y = g(\mathbf{X}) = X_2 - X_1$ where X_2 and X_1 are strength and stress, respectively. $Y = g(\mathbf{X}) = X_2 - X_1 = 0$ specifies the limit state,

$Y = g(\mathbf{X}) = X_2 - X_1 > 0$ (strength > stress) defines the safe region, and $Y = g(\mathbf{X}) = X_2 - X_1 < 0$ (strength < stress) defines the failure region.

In physics-based reliability, the reliability is expressed by

$$R = \mathrm{P}\left[g(\mathbf{X}) \geq 0\right] \tag{6.3}$$

and the probability of failure is expressed by

$$p_f = 1 - R = \mathrm{P}\left[g(\mathbf{X}) < 0\right] \tag{6.4}$$

Since in engineering applications, reliability is usually high and the probability of failure is usually low, both of them are related to the left tail of the performance function. This is illustrated in Fig. 6.4.

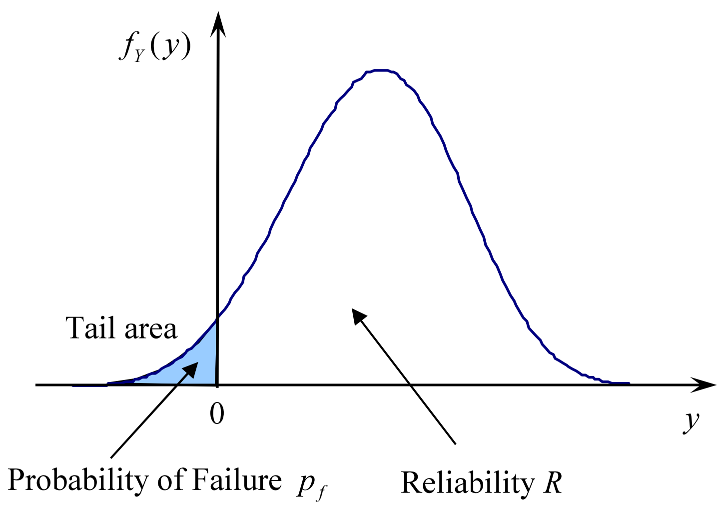

Figure 6.4 Reliability Concept

For the aforementioned performance function $Y = g(\mathbf{X}) = X_2 - X_1$, the reliability is

$$R = \mathrm{P}\left[g(\mathbf{X}) \geq 0\right] = \mathrm{P}\left(X_2 - X_1 \geq 0\right) \tag{6.5}$$

which is the probability of strength X_2 grater than the stress X_1.

The input random variables may vary with time. (In this case, they called stochastic processes.) For example, the material strength may deteriorate with time and the loading of a structure may be a function of time. Therefore, the physics-based reliability may also be time dependent. In this sense, it is possible to relate the math-based reliability with physics-based reliability. When the reliability is time dependent, Eq. 6.5 becomes

$$R(t) = \mathrm{P}\{g([\mathbf{X}(t)] \geq 0\} \tag{6.6}$$

Table 1 summarizes the differences between the math-based reliability and physics-based reliability.

Table 1 Differences between Two Types of Reliability

Math-based reliability	**Physics-based reliability**
Reliability is related to life – the time to failure.	Reliability is related to the limit state.
The state change is observed.	The state change can be mathematically modeled.
Reliability evaluation relies on testing or field data.	Reliability can be evaluated from physical equations (mdels).
The reliability is defined by $R(t) = \mathrm{P}(T > t)$	The reliability is defined by $R = \mathrm{P}[g(\mathbf{X}) \geq 0]$
The reliability is time dependent.	The reliability may or may not be time dependent.
Typical methods include ▪ Fault tree analysis (FTA) ▪ Event tree analysis (ETA) ▪ Failure models, effects, and criticality analysis (FMECA) ▪ Markov process ▪ Monte Carlo simulation	Typical methods include ▪ First order second moment (FOSA) method ▪ First order reliability (FORM) method ▪ Second order reliability (SORM) method ▪ Design of Experiments (DOE) ▪ Monte Carlo simulation

The concept of physics-based reliability can be generally considered as the probability of success, probability of customer satisfaction, and so on. If a design has reliability less than the required level, conceptually, there are multiple ways to improve reliability or decrease the probability of failure, including

1) Shrinking the distribution,
2) Shifting the distribution, or
3) Both.

The idea is demonstrated in Fig. 6.5. Shrinking the distribution without changing the mean makes the distribution narrower and the area underneath the *pdf* in the failure region (Y<0) smaller; therefore, the probability of failure can be reduced. Shifting the mean of the entire distribution without changing the standard deviation also makes the probability of failure smaller. To improve the reliability through the above means, appropriate decisions on changing design variables are required. It is the task of reliability-based design.

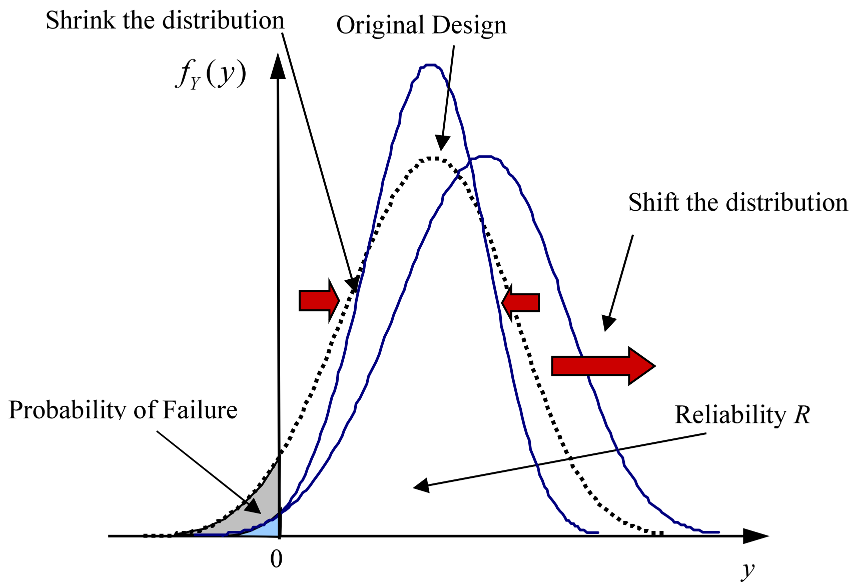

Figure 6.5 Reliability Improvement

6.4 Robustness

The robustness of a system or component is the degree to which its properties (performances) are not affected by the uncertainties of input variables or uncertainties of environmental conditions. It measures the insensitivity of system or component properties to parameter variation and uncertainties in environment.

Robustness is usually measured with the variance or standard deviation of the performance function $Y = g(\mathbf{X})$. For two designs with the same mean value as shown in Fig. 6.6, Design 1 is more robust than Design 2 since the former has a narrower distribution (lower variance).

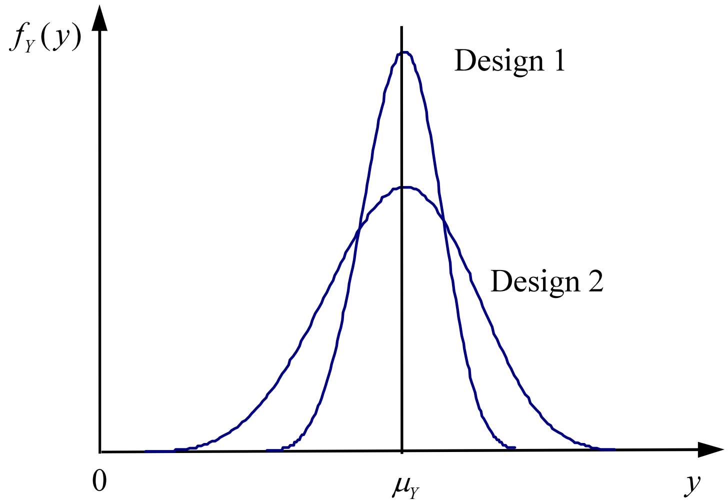

Figure 6.6 Robustness of Two Designs

Both of reliability and robustness are critical constituents of high quality. However, reliability and robustness are conceptually different. As shown in Fig. 6.7, the differences are:

1) Reliability is concerned with the performance distribution at the tails of the probability density function, while robustness is concerned with the performance distribution around the mean of the performance function.

2) Reliability is more related to safety for the avoidance of extreme catastrophic events, while robustness deals with the everyday fluctuations and is more related to the avoidance of quality loss.

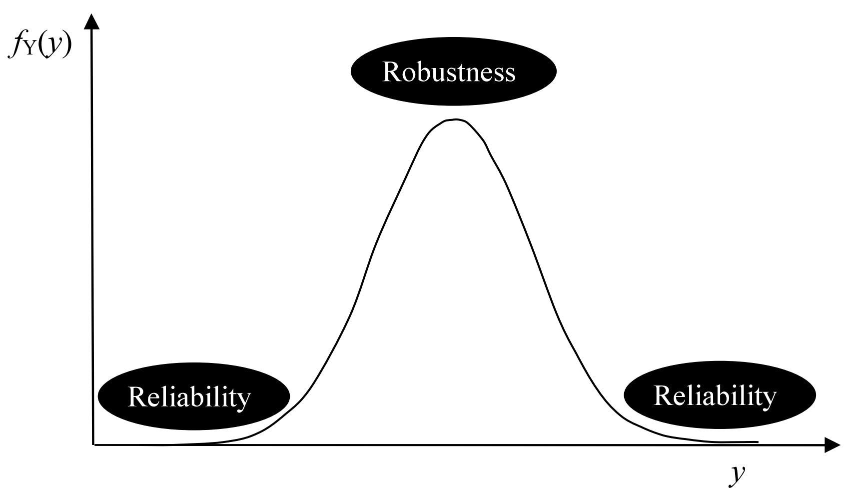

Figure 6.6 Reliability versus Robustness

Because of the difference between reliability and robustness, there are two different design methodologies dealing with reliability and robustness, namely, reliability-based design and robust design. As suggested by the name, reliability-based design makes a design reliable or ensures the probability of failure less than the required level. Robust design makes a design not sensitivity to uncertainty or reduces the variations of design performance. At the analysis level, we will first discuss how to assess reliability and robustness for a given design, and then at the design level, we will discuss both reliability –based design and robust design based on the results from reliability analysis and robustness assessment.

The concepts of reliability and robustness can be further demonstrated from the domains of applications. Fig. 6.7 [1, 2] demonstrates the domains of applications. Two aspects are considered: the likelihood of events and the consequences of the events. No matter what consequences may be resulted in, reliability-based design (regions 2 and 4 in Fig. 6.7) is applied to make sure that extreme events will be unlikely. An example of reliability-based design is to ensure that the probability of the downfall of a bridge be invariably small. Robust design (region 1 in Fig. 6.7) deals with everyday fluctuations so that the design is insensitive to such fluctuations. The everyday fluctuations may not lead to catastrophic consequences, but may lead to quality or performance losses, such as costly warranty and poor customer satisfactions. No design is acceptable if everyday fluctuations lead to catastrophe (region 3 in Fig. 6.7).

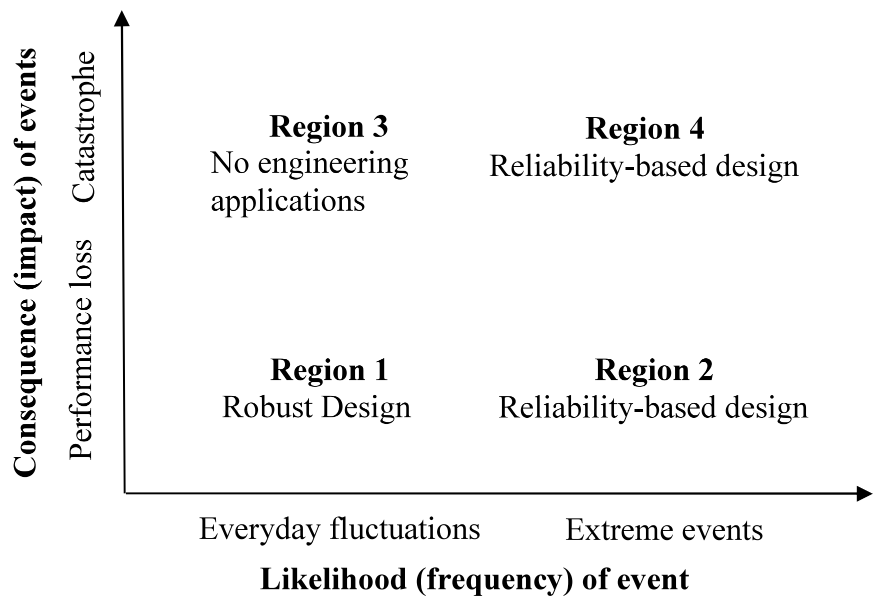

Figure 6.7 Reliability-Based Design versus Robust Design

6.5 Overview of Uncertainty Analysis

The purpose of uncertainty analysis is to find the probabilistic characteristics of a performance function $Y = g(\mathbf{X})$ given the distributions of random variables $\mathbf{X}$.

The *cdf* of $Y = g(\mathbf{X})$ is given by

$$F_Y(y) = \mathrm{P}\left[g(\mathbf{X}) \le y\right] \tag{6.7}$$

According to Eq. 6.1, the probability of failure is the *cdf* of Y at $y = 0$, i.e.

$$p_f = \mathrm{P}\left[g(\mathbf{X}) \le 0\right] = F_Y(0) \tag{6.8}$$

Theoretically, $F_Y(y)$ can be evaluated by the following integral

$$F_Y(y) = \int \cdots \int_{g(\mathbf{x}) \le y} f_{X_1, X_2, \cdots, X_n}(x_1, x_2, \cdots, x_n) dx_1 dx_2 \cdots dx_n \tag{6.9}$$

where $f_{X_1, X_2, \cdots, X_n}$ is the joint *pdf* of **X**.

If all the random variables are independent,

$$f_{X_1, X_2, \cdots, X_n} = \prod_{i=1}^{n} f_{X_i}(x_i) \tag{6.10}$$

in which f_{X_i} is the *pdf* of X_i. $F_Y(y)$ is then given by

$$F_Y(y) = \int \cdots \int_{g(\mathbf{x}) \le y} \prod_{i=1}^{n} f_{X_i}(x_i) dx_1 dx_2 \cdots dx_n \tag{6.11}$$

If changing the values of y sequentially in Eq. 6.9, we can generate a complete *cdf* of the performance function. Based on the *cdf*, other probabilistic characteristics of the performance function can be easily obtained.

There are many complexities in the evaluation of the integration in Eq. 6.9. The performance function $g(\mathbf{X})$ is usually a nonlinear function of **X**; therefore, the integration boundary $g(\mathbf{x}) = y$ is nonlinear. Since the number of random variables in practical applications is usually high, multidimensional integration is involved. In many cases, the evaluation of the performance function is computationally expensive. For example, the performance function is a black box, such as those of finite analysis and computational fluid dynamics. Due to the complexities, there is rarely a closed-form solution to the probability integration. It is also often difficult to evaluate the probability with numerical methods. Therefore, numerical approximation methods have been developed to solve the probability integration. We will discuss some of them in the following chapters.

It is worthwhile to note that an uncertainty analysis is much more computationally expensive than a deterministic analysis. To obtain the probabilistic characteristics of the performance function, many deterministic analyses have to be conducted in the vicinity of a design point that is under consideration. This is demonstrated in Fig. 6.8.

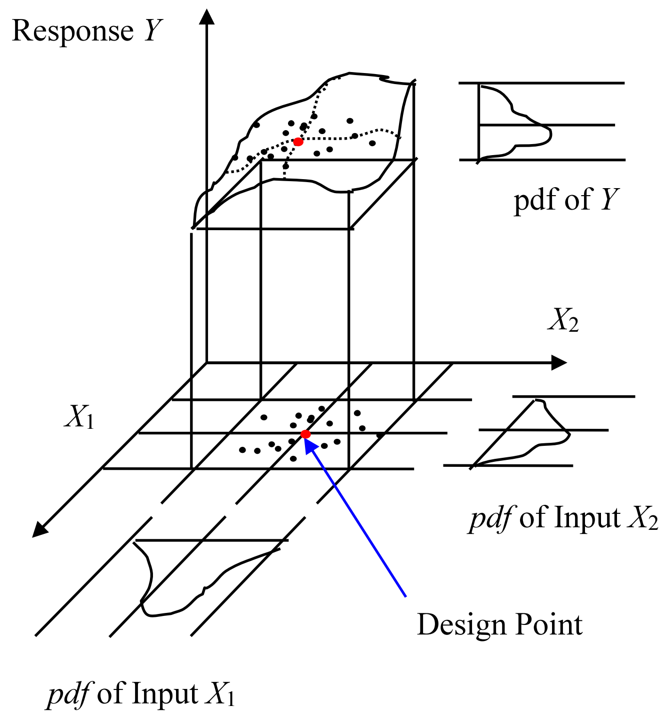

Figure 6.8 Uncertainty Analysis Needs Many Deterministic Analyses

Three categories of methods exist for the task of uncertainty analysis. The first category of uncertainty analysis methods originated from structural reliability analysis. In these methods, the performance function $g(\mathbf{X})$ is approximated such that the analytical probability integration can be easily obtained. We will discuss reliability based methods in Chapter 7 and other approximation methods in Chapter 8.

The second category is the sampling-based approach such as Monte Carlo simulation. Computational simulation is conducted to generate a sufficient large number of samples of the performance, and then the samples are analyzed statistically to get the probabilistic characteristics. Monte Carlo simulation will be discussed in Chapter 9.

The third category is the use of surrogate models to replace the performance function $g(\mathbf{X})$ by Design of Experiments. The major reason of using surrogate models is the evaluation of the original performance function is computationally expensive in many engineering problems. The use of surrogate models will alleviate the computational burden in uncertainty analysis. We will discuss this type of methods in Chapter 10.

6.6 Conclusive Remarks

We have discussed the task of uncertainty analysis and introduced two important concepts: reliability and robustness. We have also given the mathematical formulation of the *cdf* of a performance function. Since the mathematical formulation of the *cdf* of the performance function is in a form of multidimensional integral with nonlinear integration region, computational difficulties are unavoidable. To overcome the difficulties, three types of methods, reliability analysis based methods, simulation methods, and surrogate models based methods, are typically used in engineering analysis and design. All of these methods will be discussed in the following chapters.

[1] Zang, T.A., Hemsch, M.J., Hilburger, M.W., Kenny, S.P., Luckring, J.M., Maghami P., Padula, S.L., and Stroud, W.J., 2002, "Needs and Opportunities for Uncertainty-Based Multidisciplinary Design Methods for Aerospace Vehicles," NASA/TM-2002-211462.

[2] Huyse, L. 2001, "Solving Problems of Optimization under Uncertainty as Statistical Decision Problems," 42nd AIAA Structures, Structural Dynamics and Materials Conference, Seattle, WA, April 16-19, 2001.