

## APPENDIX L. UNCERTAINTY AND VARIABILITY

### TABLE OF CONTENTS

L.1	INTRODUCTION .....	L-1
L.2	UNCERTAINTY .....	L-1
L.3	VARIABILITY .....	L-1
L.4	APPROACHES TO UNCERTAINTY AND VARIABILITY .....	L-2
L.5	PROBABILITY ANALYSIS AND THE USE OF CRYSTAL BALL .....	L-2

### LIST OF FIGURES

Figure L.5.1	Normal Probability Distribution .....	L-3
Figure L.5.2	Triangular Probability Distribution .....	L-3
Figure L.5.3	Uniform Probability Distribution .....	L-3

## **APPENDIX L. UNCERTAINTY AND VARIABILITY**

### **L.1 INTRODUCTION**

Analysis of an energy-efficiency standard involves calculations of impacts, for example, the impact of a standard on consumer life-cycle cost (LCC). In order to perform the calculation, the analyst must first: 1) specify the equation or model that will be used; 2) define the quantities in the equation; and 3) provide numerical values for each quantity. In the simplest case, the equation is unambiguous (contains all relevant quantities and no others), each quantity has a single numerical value, and the calculation results in a single value. However, unambiguity and precision are rarely the case. In almost all cases, the model and/or the numerical values for each quantity in the model are not completely known (i.e., there is uncertainty) or the model and/or the numerical values for each quantity in the model depend upon other conditions (i.e., there is variability).

Thorough analysis involves accounting for uncertainty and variability. While the simplest analysis involves a single numerical value for each quantity in a calculation, arguments can arise about what the appropriate value is for each quantity. Explicit analysis of uncertainty and variability is intended to provide more complete information to the decision-making process.

### **L.2 UNCERTAINTY**

When making observations of past events or speculating about the future, imperfect knowledge is the rule rather than the exception. For example, the energy actually consumed by a particular appliance type (such as the average U.S. commercial air conditioner or heat pump) is not directly recorded, but rather estimated based upon available information. Even direct laboratory measurements have some margin of error. When estimating numerical values expected for quantities at some future date, the exact outcome is rarely known in advance.

### **L.3 VARIABILITY**

Variability means that different applications or situations produce different numerical values when calculating a quantity. Specifying an exact value for a quantity may be difficult because the value depends on something else. For example, the number of hours an air conditioner is operated by a household depends upon the specific circumstances and behaviors of the occupants (e.g., number of persons, personal habits about how comfortable the person wants to be, etc.). Variability makes specifying an appropriate population value more difficult in as much as any one value may not be representative of the entire population. Surveys can be helpful here, and analysis of surveys can relate the variable of interest (e.g., hours of use) to other variables that are better known or easier to forecast (e.g., persons per household).

## **L.4 APPROACHES TO UNCERTAINTY AND VARIABILITY**

This section describes two approaches to uncertainty and variability:

- scenario analysis, and
- probability analysis.

Scenario analysis uses a single numerical value for each quantity in a calculation, then changes one (or more) of the numerical values and repeats the calculation. A number of calculations are done, which provide some indication of the extent to which the result depends upon the assumptions. For example, the life-cycle cost of an appliance could be calculated for energy rates of 2, 8, and 14¢ per kWh.

The advantages of scenario analysis are that each calculation is simple; a range of estimates is used; and crossover points can be identified. (An example of a crossover point is the energy rate above which the life-cycle cost is reduced, holding all other inputs constant. That is, the crossover point is the energy rate at which the consumer achieves savings in operating expense that more than compensate for the increased purchase expense.) The disadvantage of scenario analysis is that there is no information about the likelihood of each scenario.

Probability analysis considers the probabilities within a range of values. For quantities with variability (e.g., electricity rates in different households), surveys can be used to generate a frequency distribution of numerical values (e.g., the number of households with electricity rates at particular levels) to estimate the probability of each value. For quantities with uncertainty, statistical or subjective measures can be used to provide probabilities (e.g., manufacturing cost to improve energy efficiency to some level may be estimated to be  $\$10 \pm \$3$ ).

The major disadvantage of the probability approach is that it requires more information, namely information about the shapes and magnitudes of the variability and uncertainty of each quantity. The advantage of the probability approach is that it provides greater information about the outcome of the calculations, that is, it provides the probability that the outcome will be in a particular range.

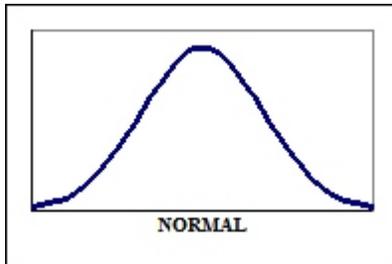
Scenario and probability analysis provide some indication of the robustness of the policy given the uncertainties and variability. A policy is robust when the impacts are acceptable over a wide range of possible conditions.

## **L.5 PROBABILITY ANALYSIS AND THE USE OF CRYSTAL BALL**

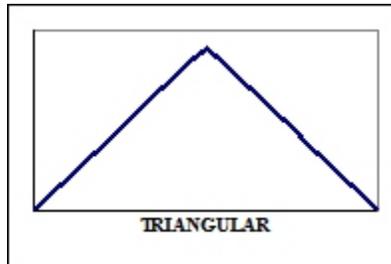
To quantify the uncertainty and variability that exist in inputs to the engineering, LCC, and payback period (PBP) analyses, the Department used Microsoft Excel spreadsheets combined with Crystal Ball, a commercially available add-in, to conduct probability analyses. The probability analyses used Monte Carlo simulation and probability distributions.

Simulation refers to any analytical method meant to imitate a real-life system, especially when other analyses are too mathematically complex or too difficult to reproduce. Without the aid of simulation, a spreadsheet model will only reveal a single outcome, generally the most likely or average scenario. Spreadsheet risk analysis uses both a spreadsheet model and simulation to automatically analyze the effect of varying inputs on outputs of the modeled system. One type of spreadsheet simulation is Monte Carlo simulation, which randomly generates values for uncertain variables again and again to simulate a model. Monte Carlo simulation was named for Monte Carlo, Monaco, where the primary attractions are casinos containing games of chance. Games of chance such as roulette wheels, dice, and slot machines, exhibit random behavior. The random behavior in games of chance is similar to how Monte Carlo simulation selects variable values at random to simulate a model. When you roll a die, you know that either a 1, 2, 3, 4, 5, or 6 will come up, but you do not know which for any particular roll. It's the same with the variables that have a known range of values but an uncertain value for any particular time or event (e.g., equipment lifetime, discount rate, and installation cost).

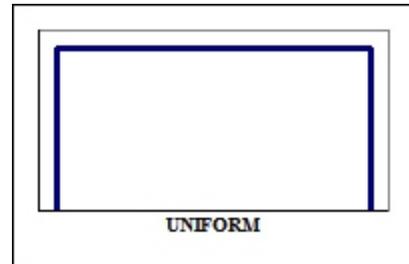
For each uncertain variable (one that has a range of possible values), possible values are defined with a probability distribution. The type of distribution selected is based on the conditions surrounding that variable. Probability distribution types include:



**Figure L.5.1**    **Normal  
Probability  
Distribution**



**Figure L.5.2**    **Triangular  
Probability  
Distribution**



**Figure L.5.3**    **Uniform  
Probability  
Distribution**

During a simulation, multiple scenarios of a model are calculated by repeatedly sampling values from the probability distributions for the uncertain variables and using those values for the cell. Crystal Ball simulations can consist of as many trials (or scenarios) as desired—hundreds or even thousands. During a single trial, Crystal Ball randomly selects a value from the defined possibilities (the range and shape of the probability distribution) for each uncertain variable and then recalculates the spreadsheet.