Robb Eric S. Moss

Applied Civil Engineering Risk Analysis

Second Edition

Springer

Applied Civil Engineering Risk Analysis

Robb Eric S. Moss

Applied Civil Engineering Risk Analysis

Second Edition

Springer

Robb Eric S. Moss
Department of Civil and Environmental
Engineering
California Polytechnic State University
San Luis Obispo, CA, USA

ISBN 978-3-030-22679-4 ISBN 978-3-030-22680-0 (eBook)
https://doi.org/10.1007/978-3-030-22680-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This is the second edition of Moss (2013). The objective of this text is to introduce fundamental applied probability and reliability methods within one term of study so that, as the course outcome, students can perform first-order probability of failure analysis on Civil Engineering problems. To achieve this outcome I present a simplified or streamlined approach to the material:

- Only basic necessary statistical methods are presented, focusing on an intuitive understanding of the data and what it can tell us about past events.
- Probability is presented in a systematic manner to make solving the often tricky probability problems more tractable.
- Only the normal and lognormal distributions are used in this text because of their useful mathematical properties and because they can often approximate a solution adequately for a first-order analysis.
- Functions of random variables are presented in a canonical manner to make solving straightforward.
- A comprehensive set of reliability methods is presented, with emphasis on methods that can be solved with minimal background in higher math.
- The tools of decision analysis, systems analysis, and other related topics are presented in an applied and usable manner.
- Many examples covering a wide range of problems are presented.

New in this second edition are: more example problems, numerical solutions presented in both MATLAB and R, and new chapters on spatial variability and Bayesian updating.

My hope for this text is that: (1) Readers are left with an appreciation of the uncertainty that exists in Civil Engineering problems and are able to quantify and treat the uncertainty in a proper way to enhance Civil Engineering design, and

(2) readers can tackle any problem and arrive at a first-order estimate of the probability of failure which will enable them to determine if further effort is necessary in refining the analysis.

San Luis Obispo, USA Robb Eric S. Moss, Ph.D. P.E. F.ASCE

Reference

Moss, R. E. S. (2013). *Applied civil engineering risk analysis* (1st edn.). Shedwick Press.

Acknowledgments

The intellectual support for this text has been provided by those whose shoulders I stand upon and those who paved the way in applied probability, Civil Engineering statistics, and structural reliability. Armen Der Kiureghian introduced me to and guided me through the concepts and methods in reliability, and set me on a Bayesian path. Ray Seed provided the initial opportunity and support for diving into this field. Alfredo Ang and Wilson Tang provided the foundation textbooks in probability and applied statistics with their first editions of Volume 1 and (the hard to find) Volume 2, as well as the excellent second edition of Volume 1 that I used in teaching parts of this course. Fred Kulhawy and Paul Mayne with their work in defining uncertainty in geotechnical engineering in its various forms opened my eyes to aleatory and epistemic uncertainty. The seminal textbook by Jack Benjamin and Allin Cornell in which they canonized much of this material from its nascent form, as well as the accessible textbook by Milton Harr which included reliability, both provided a strong background of material to draw from. Gregory Baecher and John Christian and their comprehensive text on uncertainty in geotechnical engineering delved into the specifics of my chosen field and led me to a better understanding of the material.

The logistical support for this book has been provided by Cal Poly in the form of sabbatical to pursue this writing project. Other logistical support came from Montana State University and the Universidad de Concepcion for providing a good place to "hide out" as a visiting professor.

Emotional support for this endeavor has been provided by my family (Aleksandra, Konrad, and Sebastian), my colleagues, my former students (Justin Hollenback, Daniel Wagstaffe, and Jonathan Schmidt in particular), and my friends.

Thanks to you all!

Contents

Chapter 1
Introduction

Uncertainty exists. It exists in Civil Engineering problems, in both the engineering analysis and the engineering design. The key to dealing with uncertainty is properly quantifying it and then addressing it in a systematic manner. Some examples of uncertainty that exist in Civil Engineering[1] are presented in the following chapter, but first we need to define uncertainty which comes in two forms;

1. Aleatory, and
2. Epistemic.

Aleatory uncertainty, with the root "alea" derived from the Latin for rolling of dice, is the inherent or natural randomness, whereas epistemic uncertainty stems from the lack of data, measurement error, improper mathematical modeling, or missing explanatory variables. Aleatory is the irreducible uncertainty in a particular problem. Epistemic is the reducible uncertainty that can be minimized through more study, additional data collection, better modeling, and other steps to better quantify the problem; it is the uncertainty that we as observers introduce into the problem. These two forms of uncertainty are often difficult if not impossible to separate, and can be hard to define depending on the problem and the perspective. Nonetheless the concepts of aleatory and epistemic uncertainty hold true.

To help clarify what may be considered aleatory and what may be considered epistemic, let us evaluate the breaking strength of concrete cylinders in a compression test. If we had single batch of concrete cylinders which we tested using the exact same loading equipment, the variability of the breaking strength values could be considered aleatory, but if there was any bias or imprecision due to the load measuring equipment (e.g., load cell out of calibration) that would be considered epistemic. If a second batch was tested and showed on average markedly different breaking strength, then the manufacturing process could be the culprit. If tighter control could be exerted on the manufacturing process to reduce the variability between batches then this batch

[1] Throughout this text the reference to Civil Engineering is shorthand for referring to Civil (Transportation, Structural, Geotechnical, Water) and related fields such as Environmental Engineering, Construction Engineering, and Engineering Geology. The methods presented in this text can apply to all related fields equally.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_1

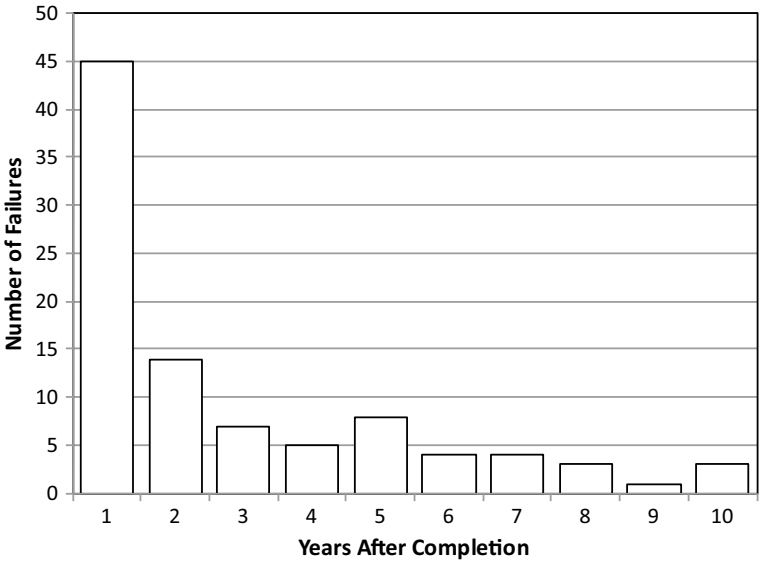


Fig. 1.1 Statistics of dam failure in the USA [after van Gelder (2000)]

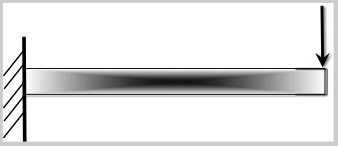
variability would be considered epistemic. And if all the breaking strength results were averaged, and the average value was used in an engineering calculation without including the variability, this would introduce additional epistemic uncertainty into the engineering process.

In most Civil Engineering problems we start out with some data on the system we are to analyze, and in almost all Civil Engineering problems we can frame the problem as one of load versus resistance. Load is what is being required of the system. Resistance is what the system can withstand (Load is synonymous with demand or stress, resistance is synonymous with strength or capacity). We are interested in how reliable the system is, or conversely what is the probability of failure (P_f) in situations we would like to avoid such as bridge collapse, slope failure, grid lock, culvert failure, chemical overload, project incompletion …

The following figures provide examples of the uncertainty that exists within Civil Engineering (Figs. 1.1, 1.2, 1.3, 1.4 and 1.5).

Example: Cantilever Beam

To demonstrate uncertainty in an engineering problem we will look at the deflection of a simple cantilever beam. The deflection at the end of the beam due to a point load is described by:



$$\Delta = \frac{PL^3}{3\text{EI}}$$

Here we will treat the load (P) and the length (L) of the beam as deterministic, meaning there is no uncertainty in the values. But if the stiffness (EI) of the beam is uncertain, due to manufacturing and material variability, then the deflection would also vary.

$$P = 10 \text{ kN}$$
$$L = 1 \text{ m}$$

$$\text{mean EI} = 80 \text{ kN m}^2$$
$$\text{std EI} = 20 \text{ kN m}^2$$

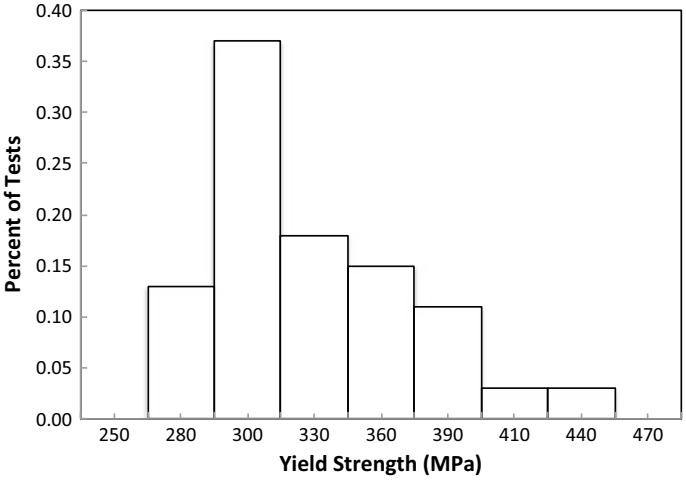


Fig. 1.2 Yield strength of steel reinforcing bars [after Julian (1957)]

The *mean* stands for the typical or average value and *std* for the standard deviation defining the uncertainty (these will be formally defined in subsequent chapters). If we do a simple sensitivity study by including the uncertainty:

$$\text{mean } \Delta = \frac{10 \text{ kN m}^3}{240 \text{ kN m}^2} = 0.042 \text{ m}$$

$$\text{mean} + \text{std } \Delta = \frac{10 \text{ kN m}^3}{300 \text{ kN m}^2} = 0.033 \text{ m}$$

$$\text{mean} - \text{std } \Delta = \frac{10 \text{ kN m}^3}{180 \text{ kN m}^2} = 0.056 \text{ m}$$

We can see that the variability in material properties can have an impact on the deflections. If the structural tolerance for deflection is <0.045 m then there is a probability of unacceptable deformations and/or failure.

The bulk of this text is focused on estimating the influence of uncertainty on the probability of failure by

- performing statistics on the data,
- applying the rules of probability,
- and defining failure within a reliability framework to calculate the probability of failure.

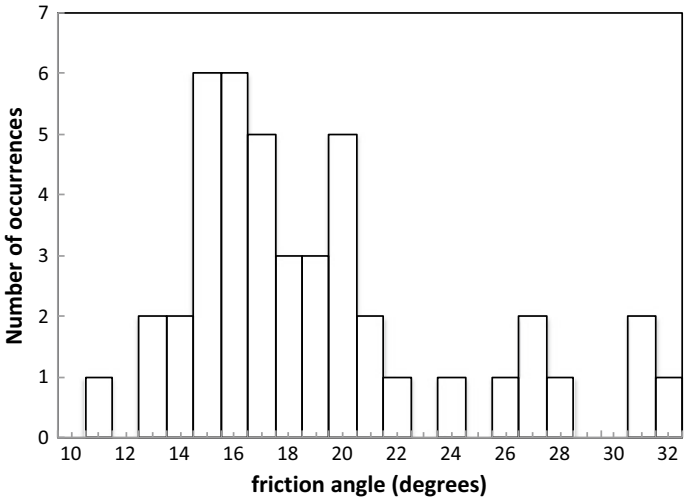


Fig. 1.3 Residual friction angle of mudstone [after Becker et al. (1998)]

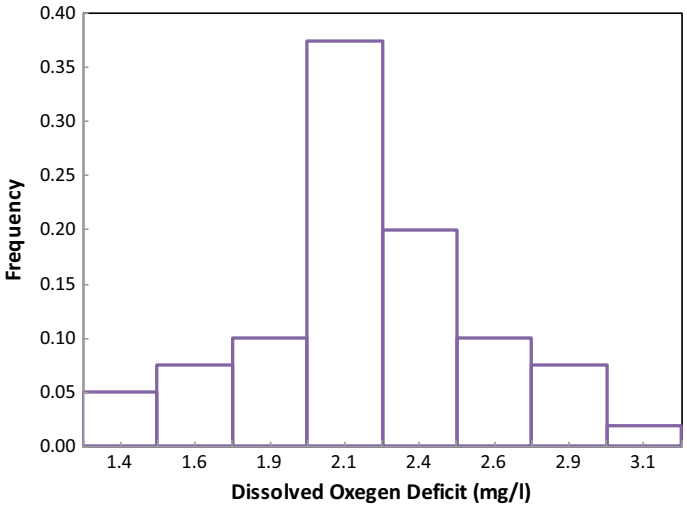


Fig. 1.4 DO deficit in a river [after Kothandaraman and Ewing (1969)]

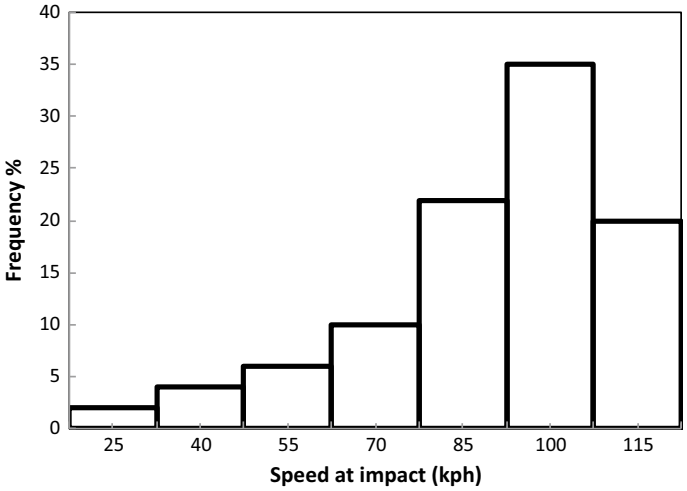


Fig. 1.5 Vehicle impact speed of passenger car accidents [after Viner (1972)]

Statistics is simply a way of evaluating data. It is a means of interrogating past events. Probability is a method of projecting trends, or forecasting the future. Failure is defined subjectively by some engineering criteria; displacement, cracking, collapse, phase transformation, uptake, crossing …

Failure has traditionally been defined in Civil Engineering in an ASD/WSD (allowable stress design/working stress design) formulation using a factor of safety.

$$\text{Factor of Safety} = \text{Strength}/\text{Stress} \tag{1.1}$$

A safe design is when the factor of safety ≥ 1. Here the uncertainty is often not quantified, and when it is quantified will often be lumped into the resistance side of the problem. As will be shown later in this text a more accurate means of defining failure is by using the limit state formulation or margin of safety.

$$\text{Margin of Safety} = \text{Strength} - \text{Stress} \tag{1.2}$$

A safe design is when margin of safety ≥ 0. Here the uncertainty is often quantified and separated into uncertainty contributed from the load and from the resistance, respectively.

The probability of failure is then how near to the failure criteria a particular problem lies. Uncertainty is an inherent part of solving for the probability of failure. Without measuring the uncertainty the probability of failure is meaningless. Probability comes in various forms and the philosophical underpinnings of probability are quite complex and are open to many interpretations. Many authors have delved into the meaning of probability (e.g., Laplace 1814) and readers are encouraged to follow up on this if there is interest, but for our purposes we assume that probability exists as a tool for forecasting future events or trends. Here we will define two camps with respect to probability, the

- frequentist approach (or classic approach) and
- degree-of-belief approach (or Bayesian approach).

The frequentist approach relies on a large number of samples to define probability. Frequentist probability works well for, say, Manufacturing Engineering where man-made objects are produced in large numbers and the probability of a faulty item can be forecast based on a large database of objects tested to failure.

The degree-of-belief approach will result in the same probability if enough data exists for a particular problem but deals with situations where data is scarce, is to expensive or infeasible to collect, and a probability must be arrived at to move forward with the analysis. The types of problems that this approach lends to are the probability of flooding at a particular site, the probability of liquefaction given the soil conditions, the probability of a column buckling under a load, or the probability of a contaminant moving into an aquifer … This form or probability is often necessary in Civil Engineering because engineered features tend to be unique in both the resistance and the load and there may be little or no data to draw from.

In practice the dual nature of probability can aid in achieving an engineering solution. One approach does not negate the other, and both approaches are equally valid and can be proven with mathematical rigor (De Finetti 1972, 1974).

Once we have determined the probability of failure we can frame the problem with risk, where risk is defined in engineering as:

$$\text{Risk} = (\text{Probability of Failure}) \cdot (\text{Consequences}) \tag{1.3}$$

We move through the steps of analyzing the data, assessing the probability, defining failure, estimating the probability of failure, and measuring the consequences. We can then state what the risk is, usually quoted as an order of magnitude within some time frame. The consequences of failure are a scaling metric, often a monetary value or life loss number, used to rank the potential failure. The risk value can then be used to compare with other potential failures or types of loss to achieve a comprehensive balanced assessment of all failure modes and their consequences, a level assessment for making rational decisions.

For example, if we take the annual probability of levee failure in a specific location due to seismic loading to be on the order of 10^{-6}, this means there is a 1/1,000,000 chance of levee failure in any given year due to seismic loading. Let's say that there is an urban area behind the levee, and if levee breach and flooding were to occur, the loss of life is estimated to be on average 10 people per failure. The annual risk of life loss due to seismic failure of a levee is then 10^{-5}. Compare this to the annual risk of dying due to an auto accident of 6.0 (10^{-4}) which is in this example 60 times higher.

There are two distinct forms of risk: voluntary and involuntary. Voluntary risk is something that we subject ourselves to consciously, and involuntary risk is when we are subjected to the risk without complete knowledge. The boundary is quite fuzzy and hard to define, and migrates depending on several factors. For instance, the annual risk of dying in a commercial airplane is less than 10^{-4} which has been held stable for years. In the 70s there were a spate of plane accidents which pushed the risk higher than 10^{-4}. This resulted in public outcry, legislation requiring stricter standards and safety oversight, and the decrease of the risk back down below the 10^{-4} range. Even though we choose to fly, the risk of dying in a commercial flight is considered by society to be involuntary and a 1/10,000 chance of death is what society deems as acceptable.

Contrast this with types of voluntary annualized risk of life loss that we subject ourselves to regularly which fall in the 1/100 and 1/1000 range; risk 100–1000 times greater than involuntary (Table 1.1). It's also important to note that perceived risk is not equal to real risk, as people are in general not good at perceiving how "risky" things are. People tend to be afraid of dying via shark or bear attacks, but it is roughly 10,000 times higher annual odds that you will die in an accident in your home (e.g., falling from a ladder) than being attacked by a large predator.

Following the 9/11 attacks, airplane ridership dropped significantly in the USA as people choose to travel by car instead. It has been estimated (Gigerenzer 2006) that in the year following the attacks almost 1600 more people died in car accidents than would have in an average year. That is roughly a 3% increase in annual car fatalities, but is more than 6 times the number of passengers who died in the four hijacked planes. Perceived versus real risk can have a significant impact on our choices and actions. This can manifest not just in our personal life choices but in our professional engineering decisions as well.

Table 1.1 Risk comparison

Type of risk	Annual risk of death	Annual odds of death
Smoking a pack a day[a]	3.6 (10^{-3})	1 in 278
Himalayan mountaineering[a]	2.8 (10^{-3})	1 in 356
Cancer (all types)[f]	2.6 (10^{-3})	1 in 387
Motor vehicle accidents[c]	6.0 (10^{-4})	1 in 1667
Killed in any war[b]	2.4 (10^{-4})	1 in 4167
Homicide[d]	1.8 (10^{-4})	1 in 5556
Air travel[c]	7.0 (10^{-5})	1 in 14,286
Flood[e]	1.0 (10^{-5})	1 in 100,000
Tornadoes[c]	4.0 (10^{-6})	1 in 250,000
Hurricanes[c]	4.0 (10^{-7})	1 in 2,500,000

[a]Wilson and Crouch (1987), [b]Small and Singer (1982), [c]US Nuclear Regulatory Commission (1975), [d]Holmes and De Burger (1988), [e]Lichtenstein et al. (1982), [f]http://www.hse.gov.uk/education/statistics.htm

1.1 Chapter Summary

- Uncertainty exists in Civil Engineering analysis and design.
- Uncertainty comes in two forms: **aleatory** and **epistemic**.
- Almost all problems in Civil Engineering can be framed as **load** versus **resistance** (synonymous with stress vs. strength or demand vs. capacity).
- Determining the **probability of failure** requires statistics to interrogate the data, probability to forecast future trends, and a definition of failure within a reliability framework.
- Probability can be viewed from a frequentist or degree-of-belief approach. The material in this text leans toward the degree-of-belief or Bayesian approach.
- **Risk** is the product of the probability of failure and the consequences, usually reported as an order of magnitude (e.g., annual risk of life loss on the order of 10^{-4}).
- Risk can be categorized as voluntary or involuntary depending on how it's applied and perceived.

References

Becker, D. E., Burwash, W. J., Montgomery, R. A., & Liu, Y. (1998). Foundation design aspects of the Confederation Bridge. *Canadian Geotechnical Journal*, *35*(5), 750–768.

De Finetti, B. (1972). *Probability, induction and statistics: The art of guessing*. London: Wiley.
De Finetti, B. (1974). *Theory of probability: A critical introductory treatment*. London: Wiley.
Gigerenzer, G. (2006). Out of the frying pan and into the fire. *Risk Analysis, 26*(2), 347–351.
Holmes, R. M., & De Burger, J. (1988). *Serial murder*. Beverly Hills, CA: Sage Publications.
Julian, O. G. (1957). Synopsis of first progress report of committee on factors of safety. *Journal of the Structural Division*, *83*(4), 1–22.
Kothandaraman, V., & Ewing, B. B. (1969). A probabilistic analysis of dissolved oxygen-biochemical oxygen demand relationship in streams. *Journal (Water Pollution Control Federation)*, R73–R90.
Laplace, P. S. (1814). *Essai philosophique sur les probabilities*. Paris: Gauthier-Villars. Modern translated version; Laplace. (1951). *A philosophical essay on probabilities*. Toronto: Dover.
Lichtenstein, S., Fischhoff, B., & Phillips, L. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slivc, & A. Tversky (Eds.), *Judgement under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge Press.
Small, M., & Singer, J. D. (1982). *Resort to arms: International and civil wars 1816–1980*. Beverly Hills, CA: Sage Publications.
U.S. Nuclear Regulatory Commission. (1975). *Reactor safety study: An assessment of accident risks in the U.S. commercial nuclear power plants* (NRC WASH-1400), Washington D.C.
van Gelder, P. H. A. J. M. (2000). Statistical methods for the risk-based design of civil structures. *Communications on hydrailic and geotechnical engineering*. Delft University of Technology.
Viner, J. G. (1972). Recent devolopements in roadside crush cushions. *Journal of Transportation Engineering. ASCE* (98), 71–87.
Wilson, R., & Crouch, E. A. (1987). Risk assessment and comparisons: An introduction. *Science, 236*(4799), 267–270.

Chapter 2
Data Analysis/Sample Statistics

Statistics is a means of interrogating past events. If we have some data on previous events, statistics are a set of tools that allow us to quantify the trends of these past events. Data in Civil Engineering is often scarce due to the unique nature of the engineered features (e.g., bridges, buildings, levees, tunnels, pipelines), but there are situations where we can assume that the data is similar enough in its variability, which is termed homoscedastic, to represent the same type of engineering problem. In these cases we need to estimate the trends of an occurrence of some event related to the engineered feature. How do we quantify uncertainty given some observations? The methods covered in this chapter include

- Frequency plots (aka histograms),
- Sample statistics (for single and multiple variables), and
- Estimation methods.

A variable is defined as something that can assume a mathematical value for engineering calculation purposes.[1] Take for instance the equation of a line, $Y = mX + b$. Assuming that we know the slope, m, and the intercept, b, then X is a variable that can assume a range of mathematical values to calculate Y.

2.1 Histograms

A histogram is a means of plotting up data on a single variable to observe the frequency of occurrence. It is a bar chart showing how often the measured event occurs. Here we will take some annual rainfall data and look at what a histogram can tell

[1] A variable will be denoted in this text by a capital letter, X, whereas a specific numerical value of that variable will be denoted by a lower case letter, x. When a variable assumes a specific value we can describe it as $X = x$.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_2

us about rainfall events. Rainfall data was collected for 49 consecutive years and reported as the annual accumulation (usually in centimeters)[2] as shown in Table 2.1.

The frequency or count for a particular annual rainfall value is then plotted as a histogram (Fig. 2.1). If we label the rainfall data with the variable X, then the data is binned with bin widths of $x + \Delta x$ (where Δ is some small change in the value of x).

As can be seen in the histogram, the most frequent average annual rainfall value is 20, and 20 happened 12 of the 49 years resulting in $12/49 = 0.25$ or 25% of the time. There are no hard rules about selecting the bin width; however, the bin width can have a dramatic impact on what the histogram looks like and what information it conveys. A convention that can be used to determine the bin width is (Benjamin and Cornell 1970):

Table 2.1 Annual rainfall data for 49 consecutive years

18	16	18	26	24	18
24	22	24	20	22	20
22	20	16	26	18	22
20	18	24	20	18	26
16	24	20	16	20	20
22	20	18	18	22	22
26	22	22	18	20	20
24	26	20	16	22	24
18					

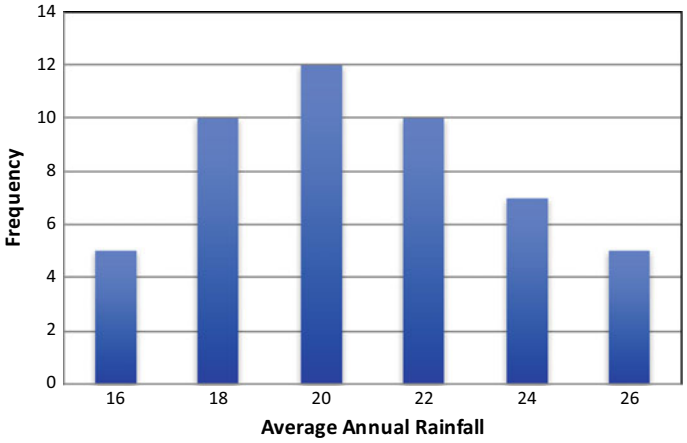


Fig. 2.1 Histogram or frequency plot of rainfall data

[2]Metric units will be used throughout this text, with a preference for SI units. Using SI units is consistent with the rest of the world (only the US still uses English units), helps greatly in minimizing computational errors, and is far easier to use and understand. The goal is that someday the US will move out of the past and adopt the SI system once and for all.

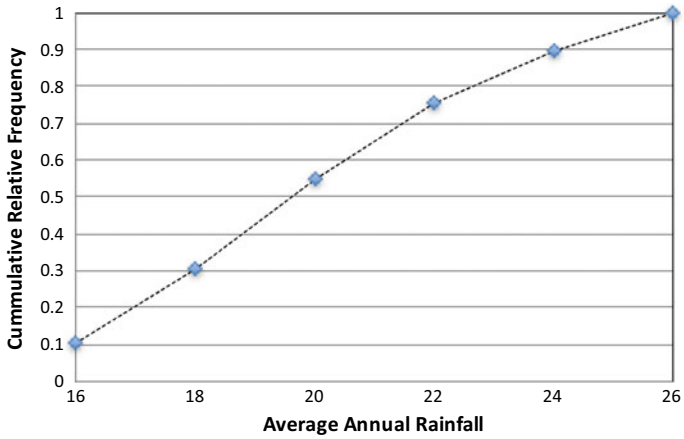


Fig. 2.2 Cumulative relative frequency plot of rainfall data

$$k = 1 + 3.3 \cdot \log(n) \tag{2.1}$$

where k is the number of intervals between the maximum and minimum values observed, n is the number of data observed, and log is the base 10 logarithm.

If we divided the frequency of occurrence, the y-axis on the histogram, by the total number of observations it would then show relative frequency. Relative frequency is useful if we are comparing two different data sets that have different total number of observations; it normalizes the data. We can also sum these relative frequencies at each bin to plot the cumulative relative frequency distribution (Fig. 2.2).

Both the frequency plot and the cumulative frequency plot (normalized or not) provide us with useful information about the data. The frequency plot shows the most likely outcome and how the data clusters around this value. The cumulative plot shows there is a gradual change between bins and the relative frequency for values $\leq x$. The general shape of these plots is going to be important in subsequent chapters when we tackle probability and theoretical probability distributions.

2.2 Central Tendency

The central tendency is measuring the most likely outcome of a data set. The histogram in Fig. 2.1 shows us the most likely value, visually but that can be influenced by the bin size. The most common measure of central tendency of a data set is the average or arithmetic mean, called the **sample mean**. For a data set of observed values $x = x_1, x_2, \ldots, x_n$ the sample mean is:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2.2}$$

If we calculate the sample mean of the rainfall data it is $\bar{x} = 20.77$. The **mode** is the most frequently occurring value, which is 20 for the rainfall data as observed in the histogram. The final common measure of central tendency is the **median** which is the value with 50% on either side, or the middle value in an ordered list (the middle value if n is odd, the average of the two middle values if n is even). For the rainfall data the media is $x_{0.5} = 20$.

Note that when the frequency plot is symmetric the mean, median, and mode are equivalent (mean = mode = median). This is a nice property of symmetric data or symmetric distributions.

2.3 Dispersion

Dispersion is the measure of how the data is dispersed or spread out from the central tendency. This is the variability of a particular problem. The variability is due to aleatory and epistemic uncertainty and has been captured in the measurement of the data set. Ways of describing dispersion include the range, the sample variance, sample standard deviation, and the coefficient of variation.

The **range** is the spread between the maximum and minimum values in a data set. The rainfall data set ranges from 16 to 26; therefore, the range is 10. This measure of dispersion is simple but tends to emphasize the extreme values and neglect the bulk of the data around the central tendency. The **sample variance** conceptually is analogous to the moment of inertia about a center of gravity. It is also analogous to the squares of distance. Calculating the sample variance is accomplished by:

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right) \tag{2.3}$$

Dividing by $n-1$ provides an unbiased estimate of the sample variance (Benjamin and Cornell 1970; Ang and Tang 2007). The dispersion is usually reported as the **sample standard deviation**, s, which is the square root of the sample variance. The rainfall data has a sample standard deviation of $s = 2.97$. A normalized form of dispersion is the **coefficient of variation**, which is the sample standard deviation divided by the sample mean:

$$\delta = \frac{s}{\bar{x}} \tag{2.4}$$

This is quite useful when comparing the dispersion from different data sets. It is often reported in literature and can be estimated based on trends or limited data. The coefficient of variation for the rainfall data is $\delta = 0.14$ or 14%.

The equations of the central tendency and dispersion are presented here but in most cases these will be calculated using built-in functions in common computation software (e.g., Excel, MATLAB, Maple, etc.).

2.4 Estimation of Variance

The dispersion or variance can often be estimated or approximated when there is insufficient data to calculate it by using the methods above. Three types of estimation methods include coefficient of variation published in the literature, the "6 Sigma" approach, and expert consensus.

For most Civil Engineering problems there is precedent, meaning that someone has tackled a problem similar enough that it can be used as an analog. We can find **published coefficient of variation** values in the literature in various places depending on the problem. These are not catalogued in any one place but can be found in codes, specifications, or technical journal publications. Table 2.2 provides a starting point for most geotechnical and structural problems.

Because geotechnical engineering uses natural and not man-made materials, there is depth in the geotechnical literature discussing dispersion and parameter uncertainty. An example of this is the work by Kulhawy and Mayne (1990) where they provide a comprehensive assessment of the coefficient of variation for both laboratory and in situ geotechnical tests. These values were accumulated from many tests over many years and can be applied to any specific problem with some confidence that the dispersion will be reasonably approximated. The coefficient of variation is multiplied by the sample mean of any specific problem to arrive at a defensible sample standard deviation.

The "**6 Sigma**" approach can be used for approximating the dispersion where it can be assumed the data has a relatively symmetric distribution (e.g., previous rainfall histogram). The steps are

- estimate the mean or median value for a variable,
- conceive of the upper and lower extreme values, and
- divide the range by 6 to get the standard deviation.

This gives an estimate of 99.9% of the data which is ± 3 standard deviations on either side of the mean. This is a useful approach, but note that humans are notoriously bad at estimating extreme values. Many failures of Civil projects are due to poor understanding of extreme loads or unforeseen load combinations. So caution must be taken when applying this approach.

The method of **expert consensus** is used in many situations where the dispersion is unquantified and there are many sources that contribute to the uncertainty. This method is just as it sounds, get a group of experts together and have them estimate the dispersion. The experts can be polled to give their best estimate of the standard

Table 2.2 Typical coefficients of variation for geotechnical and structural properties [after Harr (1987)]

Type	Parameter	δ (%)
Soil	Porosity[a]	10
	Specific gravity[b]	2
	Water content-clay[c]	13
	Degree of saturation[c]	10
	Unit weight[d]	3
	Permeability[e]	90–240
	Compression index-sandy clay[f]	26
	Compression index-clay[c]	30
	Friction angle-gravel[g]	7
	Friction angle-sand[g]	12
Structural loads	Dead load[h]	10
	Live load[h]	25
	Snow load[h]	26
	Wind load[h]	27
	Earthquake load[h]	>100
Structural resistance	Steel-tension member-yielding[h]	11
	Steel-tension member-tensile strength[h]	11
	Steel-compression beam-uniform moment[h]	13
	Steel-plate/girder-flexure[h]	12
	Reinforced concrete-grade 60 steel-flexure[h]	11
	Reinforced concrete-grade 40 steel-flexure[h]	14
	Reinforced concrete-cast in place beam-flexure[h]	8.0–9.5
	Reinforced concrete-short columns[h]	12–16
	Wood-compressive strength[i]	19
	Wood-flexural strength[i]	19

[a]Schultze (1972), [b]Padilla and Vanmarcke (1974), [c]Fredlund and Dahlman (1972), [d]Hammit (1966), [e]Nielsen et al. (1974), [f]Lumb (1966), [g]Schultze (1975), [h]Ellingwood et al. (1980), [i]Borri et al. (1983)

deviation, or their best estimate of the range which would then lend to the "6 Sigma" approach. The same caveat of extreme values applies here as well, experts are also susceptible to a limited ability to conceive of extreme values.

2.5 Correlation of Paired Data

Engineering data often comes in pairs. At the beginning of this chapter we discussed the equation of a line where Y was a function of the X. Here X is the independent variable and Y the dependent variable where the slope m and the intercept b are treated as coefficients. We infer the relationship between X and Y through the mathematical function of a straight line. If X and Y are positively correlated that means that as X increases so does Y. Negative correlation would indicate the opposite trend, as X increases Y decreases. We can calculate how X and Y vary together; this is called the sample covariance:

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{2.5}$$

But more useful is to normalize the sample covariance to arrive at the sample correlation coefficient:

$$\rho = \frac{s_{xy}}{s_x s_y} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) \tag{2.6}$$

The correlation coefficient describes the normalized dependence between the two variables and takes a value in the range from $+1$ to -1. If there is one-to-one positive dependence between X and Y then the correlation coefficient is equal to 1. Figures 2.3 and 2.4 give examples of covariance and correlation coefficient.

A quick means of estimating the correlation coefficient of paired data is by using the properties of an ellipse. As shown in Fig. 2.5 the height and the width of the ellipse can be used in Eq. (2.7) to estimate ρ. The correlation coefficient is proportional to the ratio of the ellipse width, h, to the ellipse height, H:

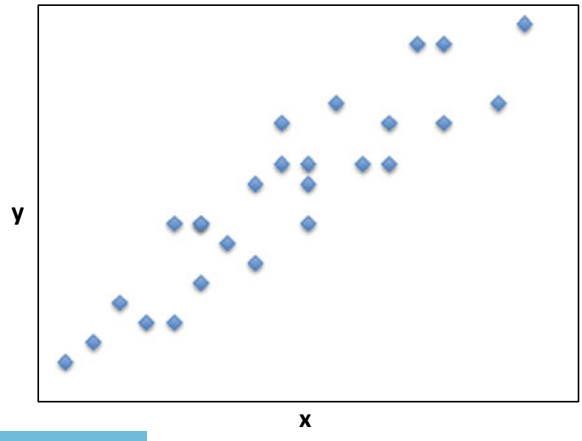


Fig. 2.3 Paired data showing small covariance (s_{xy} = small) and positive correlation ($\rho \sim 1$)

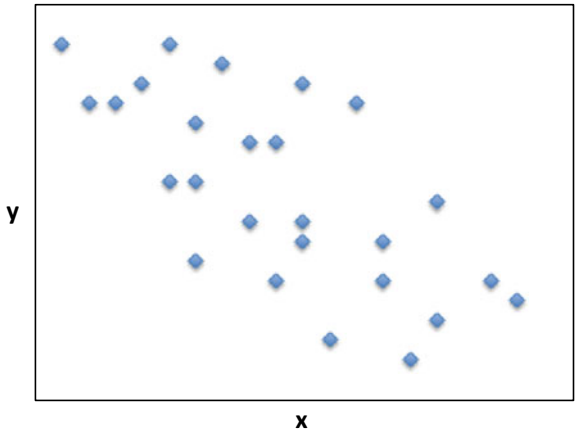


Fig. 2.4 Paired data showing large covariance (s_{xy} = large) and negative correlation ($\rho \sim -1$)

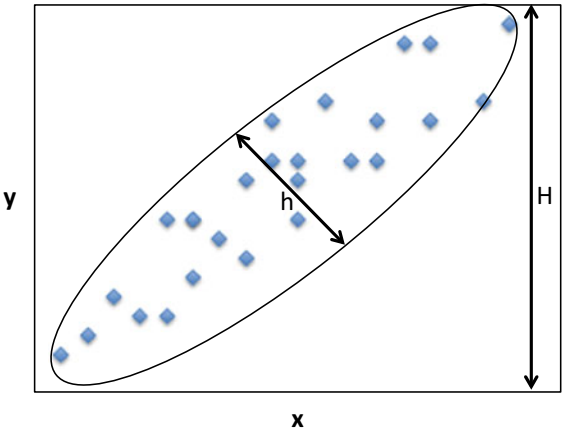


Fig. 2.5 Estimating correlation coefficient using an ellipse

$$\rho = \sqrt{\left(1 - \left(\frac{h}{H}\right)^2\right)} \tag{2.7}$$

Note as the ratio goes to zero the correlation coefficient goes to 1. This ellipse approximation provides a good check on the correlation found using regression presented the next section.

2.6 Basic Linear Regression

The most common approach to determine the correlation coefficient in engineering is by linear regression. Spreadsheet programs have "best fit" functions built into

the graphing operations making this a simple task. Here only the very basics of regression will be covered so that the correlation value can be used later when we get to probability and systems. There are any number of textbooks written about linear, multilinear, nonlinear, orthogonal, Bayesian, and other types of regression (e.g., Ang and Tang 2007; Moss 2009).

Least squares linear regression is a means of best fitting a straight line to paired data, X and Y. The objective is to get the mean or expected value of Y given $X = x$ using a straight line as the mathematical relationship between the two:

$$E(Y|X = x) = \beta x + \alpha \tag{2.8}$$

The slope, β, and the intercept, α, are the coefficients we need to solve for given a particular set of observed paired data. A "best fit" straight line is one that minimizes the error for all observations. The total absolute error for all the data points can be represented by the total cumulative squared error:

$$\Delta^2 = \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 \tag{2.9}$$

The solution for this, a partial differential equation minimizing the squared error in the slope and intercept (Ang and Tang 2007), is:

$$\beta = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \tag{2.10}$$

$$\alpha = \bar{y} - \beta \bar{x} \tag{2.11}$$

The variance of this "best fit" is the sample conditional variance of Y given X:

$$s_{Y|x}^2 = \frac{\Delta^2}{n-2} \tag{2.12}$$

If we normalize the conditional variance of Y given X by the variance of Y alone we arrive at a measure of the reduction in variance due to X because they are correlated and their variances are interrelated:

$$r^2 = 1 - \frac{s_{Y|x}^2}{s_Y^2} \tag{2.13}$$

The r^2 is a common metric used in regression to qualify how well the line fits the data. A high r^2 indicates that there is a greater reduction in the conditional variance associated with the regression which results in a better prediction of Y. It has been shown (Ang and Tang 2007) that the correlation coefficient is related to r^2 as the number of samples, n, becomes large:

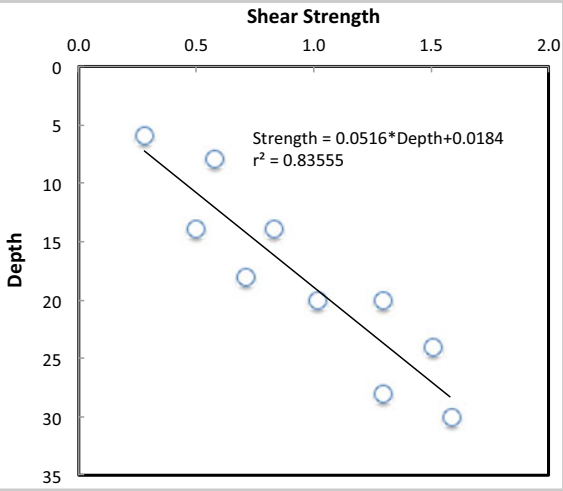
$$\rho = \sqrt{1 - \frac{n-2}{n-1}\frac{s^2_{Y|x}}{s^2_Y}} \approx \sqrt{r^2} \tag{2.14}$$

This then brings us back to paired data and the correlation of X and Y, but here using linear regression to estimate the correlation coefficient. Notice the similarities between Eqs. (2.14) and (2.7); both are the square root of one minus the ratio of the variance of Y given X over the variance of Y alone.

Example: Shear Strength with Depth

This example (after Moss 2009, Appendix) presents the results of a linear regression analysis. We are interested in how the shear strength changes linearly with depth given the data in the table.

Depth	Shear Strength
6	0.28
8	0.58
14	0.5
14	0.83
18	0.71
20	1.01
20	1.29
24	1.5
28	1.29
30	1.58



The equations for linear regression are presented in this chapter, but it can be carried out easily using the curve fit option in Excel, or using the curve fitting toolbox in MATLAB. The results show a fairly linear relationship between the two variables with an r^2 of 0.84, suggesting that for this problem depth is a reasonably good predictor of shear strength. Note that the variables have been switched in this problem because viewing depth along the vertical axis makes intuitive sense.

2.7 Chapter Summary

- A **histogram** is a useful way of viewing the frequency of data and how it is distributed.

- Frequency data can also be plotted in cumulative form and normalized to provide the relative **cumulative frequency**.
- The **central tendency** of data can be described using the **mean**, **median**, and/or **mode**.
- The **dispersion** of data can be described using the **variance**, **standard deviation**, and/or **coefficient of variation**.
- Estimations of variance can be achieved by looking for reported coefficient of variation published in the literature, using the "6 Sigma" method, and soliciting expert consensus.
- Paired data can be plotted to determine if correlation exists. A rough estimate of the **correlation coefficient** can be made using the ellipse approach.
- Linear least squares regression can also be used to estimate the correlation coefficient of paired data.

References

Ang, A. H.-S., & Tang, W. H. (2007). *Probability concepts in engineering: Emphasis on applications to civil and environmental engineering* (2nd ed.). New York: Wiley.

Benjamin, J. R., & Cornell, C. A. (1970). *Probability, statistics, and decisions for civil engineers*. New York: McGraw-Hill.

Borri, A., Ceccotti, A., & Spinelli, P. (1983, June). Statistical analysis of the influence of timber defects on the static behavior of glue laminated beams. In *Proceedings of 4th International Conference on Applications of Statistics and Probability in Soil and Structural Engineering* (Vol. 1).

Ellingwood, B., Galambos, T. V., MacGregor, J. G., & Cornell, C. A. (1980, June). *Development of a probability-based load criteria on for the American National Standard A58*. Special Pub. 577. Washington D.C.: National Bureau of Standards.

Fredlund, D. G., & Dahlman, A. E. (1972). *Statistical geotechnical properties of glacial Lake Edmonton sediments*. London: Statistics and Probability in Civil Engineering, Oxford University Press.

Hammitt, G. M. (1966). *Statistical analysis of data from a comparative laboratory test program sponsored by ACIL*. Misc. Paper 4-785, U.S. Army Engineering Waterways Experiment Station, Corps of Engineers.

Harr, M. E. (1987). *Reliabilty-based design in civil engineering* (p.290). New York: McGraw-Hill.

Kulhawy, F. H., & Mayne, P. W. (1990, August). *Manual on estimating soil properties for foundation design*. EPRI EL-6800 Research Project 1493-6.

Lumb, P. (1966). Variability of natural soils. *Canadian Geotechnical Journal, Vol. 3*.

Moss, R. E. S. (2009). *Reduced uncertainty of ground motion prediction equations through Bayesian variance analysis*. Pacific Earthquake Engineering Research Center (PEER) Report#105 (Appendix on Bayesian Regression). http://peer.berkeley.edu/publications/peer_reports/reports_2009/reports_2009.html.

Nielsen, D. R., Biggar, J. W., & Erh, K. T. (1974). Spatial variability of field-measured soil-water properties. *Hilgardia A Journal of Agricultural Science (CA Agricultural Experiment Station), 42*, 215–260.

Padilla, J. D., & Vanmarcke, E. H. (1974). *Settlement of structures on shallow foundations: A probabilistic analysis* (Res. Rep. R74-9). Cambridge, MA: MIT.

Schultze, E. (1972). *Frequency distributions and correlations of soil properties*. London: Statistics and Probability in Civil Engineering, Oxford University Press.

Schultze, E. (1975, September). The general significance of statistics for civil engineering. In *Proceedings of 2nd International Conference on Applications of Statistics and Probability in Soil and Structural Engineering*, Germany.

Chapter 3
Elementary Probability and Set Theory

We have interrogated the past in Chap. 2 by applying statistical tools to some data. Now in order to design for future events we need to forecast what is likely to happen during the design life of an engineered feature. This is where probability comes in. Probability is heavily used in Civil Engineering risk analysis because as mentioned before data is often scarce.

Set theory is a logical framework for defining the relationships between events. Venn diagrams provide a visual way of defining these relationships and in defining probability. Probability is itself an independent branch of mathematics that is:

- Logically consistent,
- Founded on 3 axioms, and
- Defines the components of probability (but does not define what probability is or what it means).

Probability theory arose from gambling. A friend of Blaise Pascal's (the renown mathematician of the mid-1600s) posed a dice question to him about the probability of winning a particular dice game with certain combinations. Pascal started a technical discussion with his equally renown colleague Pierre de Fermat, and through a series of letters the two worked out the theory of probability (Bernstein 1998; Gonic and Smith 1993). The 3 axioms that Pascal and Fermat worked out are still the basis for probability theory today:

Axiom 1: For every event (A) there is a probability; $P(A) \geq 0$,
Axiom 2: The probability of all events (S) is one; $P(S) = 1$,
Axiom 3: For two mutually exclusive events, (A) and (B), the probability of the union of the two events is the sum of the probabilities of the individual events; $P(A \cup B) = P(A) + P(B)$.

The first two axioms are relatively self-explanatory and define the range of probability from 0 to 1. The third axiom with the terms mutually exclusive and union will be defined after a discussion of the nature of probability and set theory (Ang and Tang 1975).

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_3

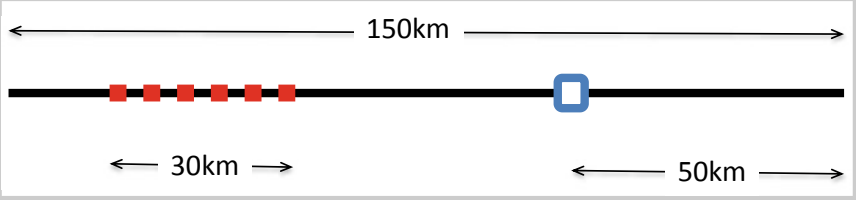
3.1 Nature of Probability

The nature of probability is often vehemently argued in certain circles with the same vigor that is usually reserved for politics or religion. In Civil Engineering we utilize all forms or interpretations of probability to aid in solving our problems. Here probability will be classified into four types:

1. Probability from **relative frequency** (i.e., observed from statistical information). As with the rainfall data in Chap. 2 we observed the most frequent value of 20 occurred 25% of the time (12/49).
2. Probability from **a priori** information, usually defined by elementary or geometric constraints of the problem (e.g., when flipping a coin P(tails) = 50%).
3. Probability **assigned subjectively**, per expert consensus or engineering judgment.
4. Probability from **mixed information**:
(observed) + (a priori) + (subjective). Using all available information like this is often called a Bayesian approach.

Example: Fault Rupture (a priori Probability)

A fault has been discovered under the foundation of a power plant. The fault has been mapped at 150 km long, and the power plant is located 50 km from one end. Seismological investigations indicate that the fault is likely to produce a M_W6.5 event which could result in 30 km of surface fault rupture contained within the 150 km total length. The fault is equally likely to rupture along its entire length. Should the M_W6.5 event occur, what is the probability that the surface fault rupture occurs under the power plant?



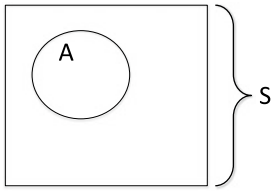
The 30 km surface rupture can equally occur anywhere along the fault length, but is contained within the 150 km fault (150–30 km). We are concerned with the likelihood of it occurring beneath the power plant (30 km).

$$P(\text{rupture beneath plant}) = \frac{30}{(150-30)} = \mathbf{0.25}\ \text{or}\ \mathbf{25\%}$$

3.2 Venn Diagrams and Set Theory

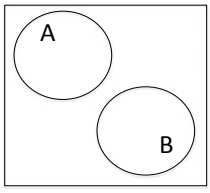
Venn diagrams are an intuitive way of visualizing events and visualizing probabilities. Set theory describes the logical/mathematical relationship between events. The following is the set theory terminology and operators we will be using for events, along with a Venn diagram visually showing the relationships being described.

The Venn diagram below and left shows the entire sample space (S) as represented by the box. The circle (A) represents an area within the sample space that is our event of interest. The sample space (S) is **collectively exhaustive** meaning it contains all possible events or combinations of events; therefore it has a probability of 1. The probability of (A) is equal to the area of (A) with respect to the total area of the sample space (S).



$$P(A) = \frac{P(A)}{P(S)} = \frac{\text{fractional area of } A}{1} \tag{3.1}$$

This concept of area ratios is useful in conceptualizing an event with respect to all possible events for a particular problem. When we have two events that are **mutually exclusive**, meaning that they share no area of sample space in common, then the union is accomplished by adding their fractional areas.

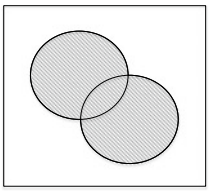


$$P(A \cup B) = P(A) + P(B) \tag{3.2}$$

Union ($\cup$)

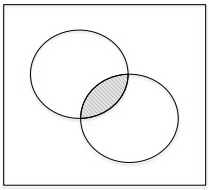
$A \cup B$ is the occurrence of A **or** B or both events. The operative word here is **or** which is the word you want to equate with the union ($\cup$) symbol. As the Venn diagram shows

the union includes the sample space of both events combined. Here the two events are not mutually exclusive because they have some sample space in common.



Intersection ($\cap$)

$A \cap B$ is the occurrence of A **and** B jointly. The operative word here is **and** which is the word you want to equate to the intersection ($\cap$) symbol. As the Venn diagram shows the intersection includes only the area where the two events overlap or intersect. Often the intersection symbol is dropped and the set theory statement becomes simply AB.



Set theory follows the same logical rules as arithmetic, such as the commutative, associative, and distributive rules. The commutative rule means that the order of events in a union or intersection does not affect the outcome, that is:

$$A \cup B = B \cup A \text{ and } AB = BA \tag{3.3}$$

The associative rule means that all events are equally associated with a union or intersection:

$$(A \cup B) \cup C = A \cup (B \cup C) \text{ and } (AB)C = A(BC) \tag{3.4}$$

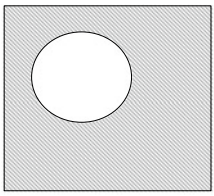
And the distributive rule means that an event can be distributed equally when there is an intersection:

$$(A \cup B)C = (AC) \cup (BC) \tag{3.5}$$

These rules should be familiar and almost intuitive because there are the same rules we learn at an early age with respect to addition and multiplication.

3.3 Complement

The complement of an event is one minus the event; it is the sample space that represents everything but the event. The complement is represented by a bar over the event (e.g., $\bar{A}$). Therefore:

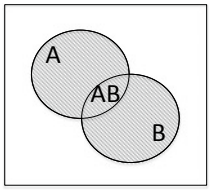


$$P(A) = 1 - P\left(\bar{A}\right) \tag{3.6}$$

$$P\left(\bar{A}\right) = 1 - P(A) \tag{3.7}$$

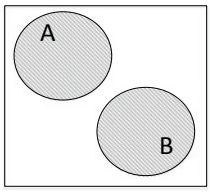
3.4 Addition Rule

The addition rule defines the mathematics of the union of events. The union of event A and event B includes the combined area of A and B. The union states that the outcome could be event A, **or** event B, **or** both. We sum the area of A and the area of B, but have double counted the intersection so must subtract AB to arrive at the combined area:



$$P(A \cup B) = P(A) + P(B) - P(AB) \tag{3.8}$$

If A and B are **mutually exclusive** (**ME**) that means there is no joint area and the intersection is zero, therefore:



$$P(A \cup B) = P(A) + P(B) \quad \text{if ME} \tag{3.9}$$

Extending the addition rule for three events results in:

$$\begin{aligned} P(A \cup B \cup C) = {} & P(A) + P(B) + P(C) \\ & - P(AB) - P(BC) - P(AC) + P(ABC) \end{aligned} \tag{3.10}$$

Here we subtract the double-counted area, but then must add the triple intersection back into arrive at the union. A Venn diagram can aid in visualizing this solution.

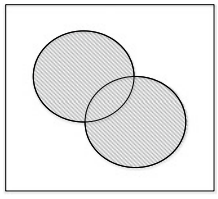
3.5 De Morgan's Rule

What is commonly called De Morgan's rule is not a rule itself but more a subset of the addition rule. Nonetheless it is traditionally called De Morgan's rule and will be referred to as such in this text. De Morgan's is an alternate way of solving the addition rule when the complement is easier to compute. A Venn diagram is often the best way to visualize De Morgan's:

$$P(A \cup B) = 1 - P\left(\overline{A \cup B}\right) = 1 - P\left(\bar{A}\bar{B}\right) \tag{3.11}$$

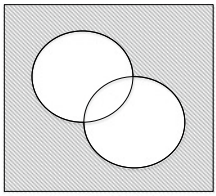
$$\begin{aligned} P(A \cup B \cup \cdots \cup Z) &= 1 - P\left(\overline{A \cup B \cup \cdots \cup Z}\right) \\ &= 1 - P\left(\bar{A}\bar{B} \ldots \bar{Z}\right) \end{aligned} \tag{3.12}$$

The Venn diagram of the union of A and B is shown below:



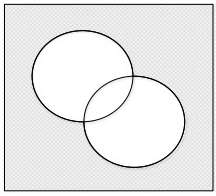
$P(A \cup B)$

The Venn diagram of the complement of the union is everything outside the union (see below), and if we take the complement of that we arrive at the first diagram.



$P(\overline{A \cup B})$

The Venn diagram of the intersection of the complement of A and the complement of B is the double-hatched region (see below), and if we take the complement of that we arrive at the first diagram again.



$P(\bar{A}\bar{B})$

Example: Water and Power Demand

A 6-building complex has recently been constructed, and the buildings are unleased as of yet. The engineer wants to meet the demand for water and power utilities as closely as possible without costly over- or under-targeting. For any one building the electricity can be either 5 or 10 units, and the water can be either 1 or 2 units. The total sample space then looks like:

E_5W_2	$E_{10}W_2$
E_5W_1	$E_{10}W_1$

The owner ascribes the following probabilities (expert consensus) based on previous experience with similar type of buildings:

$$\left.\begin{aligned} P(E_5W_1) &= 0.1 \\ P(E_{10}W_1) &= 0.1 \\ P(E_5W_2) &= 0.2 \\ P(E_{10}W_2) &= 0.6 \end{aligned}\right\} \text{sums to } 1.0\checkmark$$

(a) What is the probability of water demand of 2 units?

$$P(W_2) = P(E_5W_2) + P(E_{10}W_2) = 0.2 + 0.6 = \mathbf{0.8}$$

Note that $P(E_5W_2)$ and $P(E_{10}W_2)$ are mutually exclusive.

(b) What is the probability of power demand of 10 units?

$$P(E_{10}) = P(E_{10}W_1) + P(E_{10}W_2) = 0.1 + 0.6 = \mathbf{0.7}$$

(c) What is the probability of either the water demand is 2 units or the power demand is 10 units?

$$\begin{aligned} P(W_2 \cup E_{10}) &= P(W_2) + P(E_{10}) - P(W_2 \cap E_{10}) \quad \text{addition rule} \\ &= P(W_2) + P(E_{10}) - P(E_{10}W_2) \quad \text{intersection} \\ &= 0.8 + 0.7 - 0.6 = \mathbf{0.9} \end{aligned}$$

Note In solving these problems I will annotate to the right what rule was used for each step to aid in problem-solving clarity.

3.6 Conditional Probability Rule

It's been said that every probability is a conditional probability, and this statement rings true the more one spends working on probability problems. Conditional probability defines the measure of dependence between events; it is large when events occur jointly, small when they don't, and zero when events are mutually exclusive. A typical conditional probability question is "What is the probability of event A given the occurrence of event B?":

$$P(A|B) = \frac{P(AB)}{P(B)} \tag{3.13}$$

Because B has occurred the sample space is limited to the area of A within B; the area is restricted to the conditional event space. The area of A occurring within B is the intersection of A and B. The conditional probability is this intersection normalized by B, the ratio of the intersection AB to the total area of B.

If B has no influence on A then we can say that the two events are **statistically independent** (SI).

$$P(A|B) = P(A) \quad \text{if SI} \tag{3.14}$$

The above statement (Eq. 3.14) is the only means of determining if two events are statistically independent. Statistical independence pertains to the influence of one event on another, which is much different from mutual exclusivity that pertains to the sample space two events have in common. Events A and B may intersect but that does not dictate whether they are statistically independent or not.

Another way to think of conditional probability is through the correlation coefficient introduced in Chap. 2. In both we are describing the influence of one event upon another; however the correlation coefficient is constrained to a linear relationship.

$$\begin{aligned} &\text{if } P(A|B) > P(A) \quad \text{then } \rho = + \\ &\text{if } P(A|B) < P(A) \quad \text{then } \rho = - \\ &\text{if } P(A|B) = P(A) \quad \text{then } \rho = 0 \end{aligned}$$

When the conditioned event has a higher probability than the marginal event then there is a positive correlation, when it has a lower probability than the marginal event then there is a negative correlation, and when the conditioned event has an equal probability to the marginal event the events are statistically independent and the correlation coefficient is zero.

Note that often $P(A)$ and $P(A|B)$ come from an engineering study and the joint probability $P(A \cap B)$ or $P(AB)$ is what is desired, where:

$$\begin{aligned} P(AB) &= P(A|B)P(B) \\ &= P(B|A)P(A) \end{aligned}$$

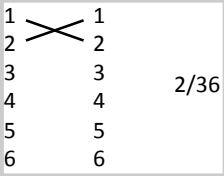
Example: Water and Power Demand II

What's the probability that a building that needs 10 units of power will also require 2 units of water? In this problem the probability of 2 units of water is conditioned on the demand for 10 units of power.

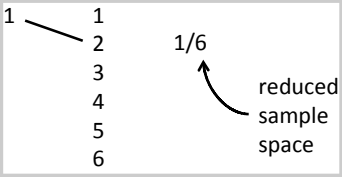
$$P(W_2|E_{10}) = \frac{P(E_{10}W_2)}{P(E_{10})} = \frac{0.6}{0.7} = \mathbf{0.86} \quad \text{conditional probability rule}$$

Example: Rolling Dice

Two die are rolled together. What is the probability the dice sum to 3?



Now one die is rolled and comes up 1. What is the probability that both dice sum to 3?



In the second situation the sample space is reduced because it is conditional.

3.7 Multiplication Rule

The multiplication rule defines the mathematics for the intersection of events. We rearrange the conditional probability rule to solve for the intersection or joint probability:

$$P(AB) = P(A|B)P(B) \tag{3.15}$$

If we determine that A and B are **statistically independent** then the conditional probability statement is reduced and we have the multiplication of two marginal event probabilities:

$$P(AB) = P(A)P(B) \quad \text{for SI} \tag{3.16}$$

For two events we can reverse the statement to check statistical independence:

$$P(A)P(B) = P(AB) \tag{3.17}$$

But this does not necessarily hold true for more than two events (Der Kiureghian 2001) and one should always check statistical independence with the conditional probability rule. For three events the multiplication rule becomes:

$$P(ABC) = P(A|BC)P(B|C)P(C) \tag{3.18}$$

Example: Building Foundation

The foundation of a tall building may fail due to inadequate bearing capacity (B) or excessive settlement (S). The following probabilities are known for this type of building and failure situation:

$$P(B) = 0.001; \quad P(S) = 0.008; \quad P(B|S) = 0.10$$

(a) What is the probability of failure of the foundation? Failure would be B or S or both.

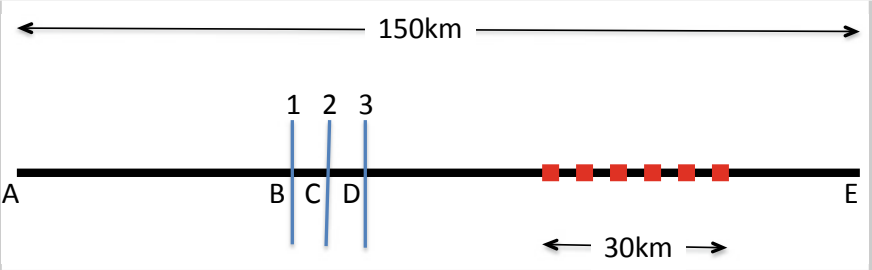
$$\begin{aligned} P(B \cup S) &= P(B) + P(S) - P(BS) \quad \text{addition rule} \\ &= P(B) + P(S) - P(B|S)P(S) \quad \text{conditional probabilty rule} \\ &= 0.001 + 0.008 - 0.10(0.008) = \mathbf{0.0082} \end{aligned}$$

(b) What is the probability the building will experience excessive settlement but not bearing capacity failure?

$$\begin{aligned} P\left(S \cap \bar{B}\right) &= P\left(S\bar{B}\right) = P\left(S|\bar{B}\right)P\left(\bar{B}\right) \quad \text{multiplication rule} \\ &= P\left(\bar{B}|S\right)P(S) \quad \text{commutative property} \\ &= [1 - P(B|S)]P(S) \quad \text{complement} \\ &= [1 - (0.10)]0.008 = \mathbf{0.0072} \end{aligned}$$

Example: Fault Rupture II

In this example we are readdressing the 150 km long fault with the potential for 30 km of surface fault rupture. For this example there are three pipes that cross the fault.



The distance between points: A to $B = 60$ km, B to $C = 10$ km, C to $D = 12.5$ km, and D to $E = 67.5$ km. P_i here denotes a failure of pipe i.

(a) What is the probability of failure of each pipe? The solution is the same as for the power plant.

$$P(P_i) = \frac{30}{(150 - 30)} = \mathbf{0.25} \quad \text{a priori}$$

(b) What is the probability of P_1 and P_2? The 30 km surface rupture must include the 10 km distance between B and C.

$$P(P_1 \cap P_2) = P(P_1 P_2) = \frac{(30 - 10)}{(150 - 30)} = \mathbf{0.167} \quad \text{a priori}$$

(c) What is the probability that 1 break, given that 2 has broken?

$$P(P_1|P_2) = \frac{P(P_1 P_2)}{P(P_2)} \quad \text{conditional probabilty rule}$$
$$= \frac{0.167}{0.25} = \mathbf{0.667}$$

(d) Are P_1 and P_2 statistically independent?

$$P(P_1|P_2) \neq P(P_1) \quad \text{therefore not SI}$$

(e) What is the probability of 1 and 3 breaking?

$$P(P_1P_3) = \frac{(30 - 10 - 12.5)}{(150 - 30)} = \mathbf{0.0625} \quad \text{a priori}$$

(f) What is the probability of at least one of the first two pipes failing?

$$\begin{aligned} P(P_1 \cup P_2) &= P(P_1) + P(P_2) - P(P_1P_2) \quad \text{addition rule} \\ &= 0.25 + 0.25 - 0.167 = \mathbf{0.333} \end{aligned}$$

(g) The first two pipes are redundant and only one needs to survive to maintain continued service (water, phone, other …). What is the probability at least one will survive?

$$\begin{aligned} P\left(\bar{P}_1 \cup \bar{P}_2\right) &= 1 - P(P_1P_2) \quad \text{De Morgan's rule} \\ &= 1 - 0.167 = \mathbf{0.8333} \end{aligned}$$

(h) What is the probability that both will survive?

$$\begin{aligned} P\left(\bar{P}_1\bar{P}_2\right) &= 1 - P(P_1 \cup P_2) \quad \text{De Morgan's rule} \\ &= 1 - [P(P_1) + P(P_2) - P(P_1P_2)] \quad \text{addition rule} \\ &= 1 - [0.25 + 0.25 - 0.167] = 1 - 0.333 = \mathbf{0.667} \end{aligned}$$

3.8 Problem-Solving Rubric

Solving elementary probability problems can be intuitive for some people and opaque for others. However, as the complexity of the problems increases they tend to become opaque to all. In order to aid in solving elementary probability problems there are some commonalities that lend to a routine problem-solving rubric. Practice of this rubric with simple problems makes solving more complex problems tractable and routine.

Rubric

(1) **Declare all knowns**. This can be tricky in some problems because the knowns may be implicit, unobvious, or not directly stated.
(2) **Write the problem statement in set terminology**. This means translating the word statement into unions and/or intersections of events. This is often the

hardest part of any elementary probability problem and takes careful dissection of the problem statement and translation of it into set theory.

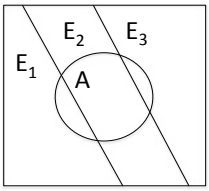
(3) **Reduce the statement to a calculable form by invoking the rules of probability**. If the set theory statement is correct then this step is often a straightforward manipulation of the rules of probability until a mathematical calculation can be performed.

This rubric works for most problems of elementary probability, but not all. There are a few things to watch out for when problems get tricky:

- In certain problems a priori probability is part of the knowns but not explicitly stated.
- Some problems result in complex probability statements that can be simplified by invoking De Morgan's rule.
- When in doubt draw a Venn diagram to clarify what the probability statement is describing.
- Adhere to logic when solving these problems.
- Always perform a "gut check" on the answer when a resulting probability is calculated. Is the answer reasonable and does it answer the word statement?

3.9 Total Probability Theorem

Total probability is a theorem that applies to a marginal event probability that is composed of many conditional events. Equation (3.19) and the Venn diagram below are shown for three condition events, but this theorem can be extended to any number of conditional events (Eq. 3.20).



$$P(A) = P(A|E_1)P(E_1) + P(A|E_2)P(E_2) + P(A|E_3)P(E_3) \tag{3.19}$$

$$P(A) = \sum_{i=1}^{n} P(A|E_i)P(E_i) \tag{3.20}$$

The best way to elucidate this theorem is through example.

Example: Hurricane Damage

Hurricanes are categorized by increasing wind speed, Category 1 through Category 5. Based on historical hurricane data along the Louisiana Coast the annualized probability of each category is:

$$P(C_1) = 0.35, \quad P(C_2) = 0.25, \quad P(C_3) = 0.14,$$
$$P(C_4) = 0.05, \quad P(C_5) = 0.01.$$

In this problem we are interested in the probability of structural damage due to hurricane winds. Reconnaissance of previous hurricane disasters has shown that structural damage can be approximated by the following conditional probabilities:

$$P(D|C_1 \ldots C_5) = 0.05, 0.10, 0.25, 0.60, 1.00$$

Given this information, what is the annual probability of structural damage?

$$\begin{aligned} P(D) &= P(D|C_1)P(C_1) \ldots P(D|C_5)P(C_5) \quad \text{total probability theorem} \\ &= 0.05 \times 0.35 + 0.10 \times 0.25 + 0.25 \times 0.14 + 0.60 \times 0.05 + 1.00 \times 0.01 \\ &= \mathbf{0.1175} \end{aligned}$$

3.10 Bayes' Theorem

If we combine the conditional probability rule and the multiplication rule we arrive at what is called Bayes' Theorem.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \tag{3.21}$$

This theorem (often called Bayes' rule) is a simple enough statement that has a lot of history behind it and has caused more than a few scandals in the realm of statistics and probability. It is named after Thomas Bayes, a minister who worked on probability in his spare time. He perished before publishing his work, but his proofs were published posthumously by his colleague Richard Price circa 1763. The theorem was independently rediscovered and published by Laplace (1812).

Bayes' theorem describes an inverse probability, that of event B given the occurrence of event A. In this form the theorem is often used for updating the probability of something given subsequent information. The following examples demonstrate this Bayesian updating.

Example: Aggregate Supply

In this problem there is aggregate being delivered to a construction site from two sources, A and B. Trucks from sources A deliver 600 loads a day of which 3% is bad (meaning it does not meet the project specifications). Trucks from source B deliver 400 loads a day of which 1% is bad.

(a) What is the probability of bad aggregate?

First we must declare the probability that the aggregate is from each source and declare the conditional probability of bad aggregate.

$$P(A) = \frac{600}{(600+400)} = 0.60 \quad P(B) = \frac{400}{(600+400)} = 0.40$$

$$P(\text{Bad}|A) = 0.03 \quad P(\text{Bad}|B) = 0.01$$

$$P(\text{Bad}) = P(\text{Bad}|A)P(A) + P(\text{Bad}|B)P(B) \quad \text{total probability}$$

$$= 0.03 \times 0.60 + 0.01 \times 0.40 = \mathbf{0.022}$$

(b) If the aggregate is bad, what is the probability it was from source A?

This is the type of inverse probability statement that lends well to using Bayes' theorem. Often in problems like this total probability and Bayes' theorem go hand-in-hand.

$$P(A|\text{Bad}) = \frac{P(\text{Bad}|A)P(A)}{P(\text{Bad})} \quad \text{Bayes' theorem}$$

$$= \frac{0.03 \times 0.60}{0.022} = \mathbf{0.818}$$

Example: Pollution Control

This problem deals with urban air quality. It has been determined that a city's air pollution is primarily due to two sources; industrial (I) and auto (A) emissions. In the next 5 years the chances of controlling these two emissions sources have been estimated at 75 and 60% respectively. If only one of the two sources is controlled, there is an 80% probability of bringing air pollution under control (C).

(a) What is the probability of controlling air pollution?

The probability of controlling air pollution is conditioned on controlling the two sources. It can be divided into four conditional probability statements that are summed together per total probability.

If both sources are controlled then control is obviously achieved, 100%. As stated previously if one of the other sources is controlled, the probability of air quality control is 80%. And if neither source is controlled then there is no

chance of control, 0%. In this problem we are assuming that industrial and auto emissions are statistically independent.

$$\begin{aligned} P(C) &= P(C_1|AI)P(AI) + P\left(C_2|A\bar{I}\right)P\left(A\bar{I}\right) + P\left(C_2|\bar{A}I\right)P\left(\bar{A}I\right) + P\left(C_3|\bar{A}\bar{I}\right) \\ &= 1(0.60 \times 0.75) + 0.80(0.60 \times 0.25) + 0.80(0.40 \times 0.75) + 0(0.4 \times 0.25) \\ &= \mathbf{0.81} \end{aligned}$$

(b) If pollution is not controlled after 5 years, what is the probability it was due entirely to auto pollution?

All the conditional and marginal probabilities needed to answer this inverse probability statement can be found in the total probability statement above.

$$\begin{aligned} P\left(\bar{A}I|\bar{C}\right) &= \frac{P\left(\bar{C}|\bar{A}I\right)P\left(\bar{A}I\right)}{P\left(\bar{C}\right)} \quad \text{Bayes' theorem} \\ &= \frac{0.20 \times 0.30}{0.19} = \mathbf{0.32} \end{aligned}$$

Books, theses, and other documents have been written on the philosophical underpinnings of Bayes' theorem and applications using a Bayesian approach (e.g., McGrayne 2011). The Bayesian approach has been used quite successfully in areas where there is limited data and an estimate of the probability must be made. This procedure of estimating the probability given limited data tends to fly in the face of classical or frequentist statistics and for a period of time, spanning the early to middle 1900s, was considered equivalent to statistical heresy. Nonetheless people found it quite useful for solving difficult problems where there was no classical solution, and it gradually took hold. Since the 1960s, there was movement to provide a robust mathematical basis for the Bayesian approach which succeeded in establishing the validity on equal footing with classical statistics. To this day there are still intellectual battles between the two camps, classical versus Bayesian, but this often has more to do with ones' own philosophical view of determinism than which method solves the problem adequately.

3.11 Chapter Summary

- Probability is founded on three axioms.
- A probability can be of different forms depending on the problem and the way it is framed; relative frequency, a priori, subjective, and Bayesian.
- Set theory is a way of logically expressing the **union** (**or**) or i**ntersection** (**and**) of events.
- Venn diagrams are a means of visually showing how events are related within the sample space.
- Tools for solving probability problems are the rules and theorems that are an extension of the three axioms of probability; **complement**, **addition rule**, **De Morgan's rule**, **conditional probability rule**, **multiplication rule**, **total probability theorem**, and **Bayes' theorem**.
- Using a problem-solving rubric is useful as probability problems become more complex and nonintuitive. The rubric provides a systematic way of solving most problems encountered in probability.
- Two events are **mutually exclusive** when they share no sample space in common. When mutually exclusive, the occurrence of one event precludes the occurrence of the other.
- **Statistical independence** is when the occurrence of one event does not influence the probability of occurrence of another. The dependence between events is measured using the conditional probability rule.

References

Ang, A. H.-S., & Tang, W. H. (1975). *Probability concepts in engineering planning and design: Volume I—Basic principles*. New York: Wiley.

Bernstein, P. L. (1998). *Against the Gods: The remarkable story of risk*. Wiley.

Der Kiureghian, A. (2001). *Structural Reliability CE229 course notes*. UC Berkeley.

Gonick, L., & Smith, W. (1993). *The cartoon guide to statistics*. New York: Harper Collins Publishers Inc.

Laplace, P. S. (1812). *Théorie analytique des probabilities*. Paris: Original tretise.

McGrayne, S. B. (2011). *The theory that would not die. How Bayes' rule cracked the enigma code, hunted down Russian submarines, and emerged triumphant from two centuries of controversy*. Yale University Press.

Chapter 4
Random Variables and Probability Distributions

A random variable is a mathematical tool for describing an event. It maps the possible outcomes of an event onto a number line as shown below. The Venn diagram shows the sample space with events A and B. These are mapped onto a number line to aid in mathematical calculations. In Fig. 4.1, A and B are mutually exclusive, whereas in Fig. 4.2 they have some sample space in common.

The random variable in these figures is denoted by a capital letter, X, and the number that the random variable assumes is the lower case, x. Here event A is mapped to the random variable X and is comprised of all real numbers within that space: $A = \{x_n \leq X \leq x_m\}$.

We can look at the range of values that X can assume and calculate the relative probability of any particular value or range of values; this then defines a probability distribution.

4.1 Probability Distribution

A distribution is classified first by the type of number it describes, whether it is discrete or continuous. An example of a discrete number in Civil Engineering would be a traffic count (e.g., 12 cars made a left-hand turn at the intersection). An example of a continuous number would be a contaminant concentration (e.g., 0.055 g/m^3 of polyvinyl chloride).

The mathematical density of a discrete probability distribution is shown in Fig. 4.3. This is analogous to a histogram, showing the likelihood of any discrete value occurring with respect to all other discrete values. A discrete probability distribution is called a probability mass function, **PMF**, is denoted by the symbol $p_X(x)$, and is equal to the probability of the random variable X assuming a particular value $P(X = x)$ for all x values.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_4

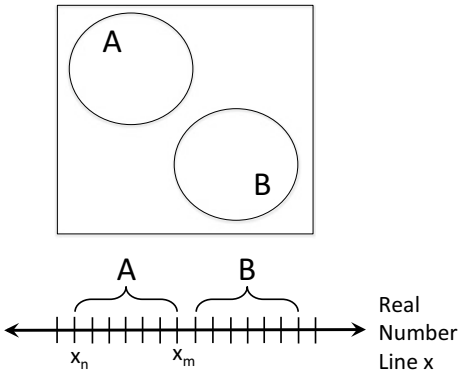


Fig. 4.1 Mapping mutually exclusive events to a real number line

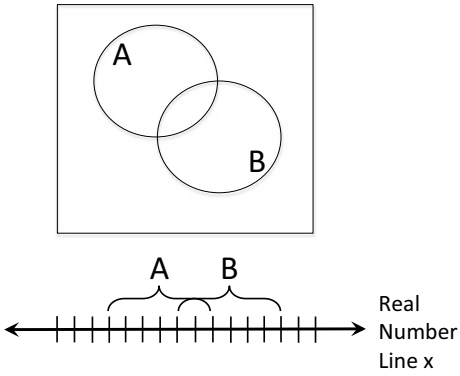


Fig. 4.2 Mapping not-mutually exclusive events to a real number line

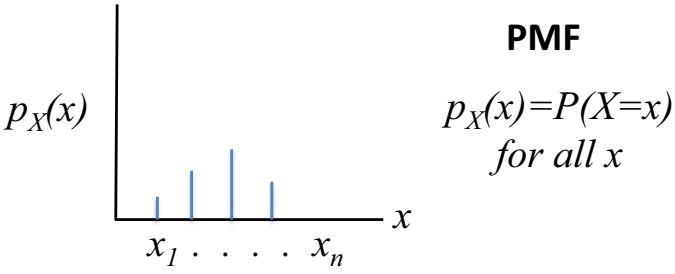


Fig. 4.3 The probability mass function, PMF, is the probability distribution for a discrete number

The cumulative distribution of a discrete number is called a **CDF** or cumulative distribution function. It is denoted by the symbol $F_X(x)$ and is equal to the probability of the random variable X being less than or equal to a particular value:

$$F_X(x) = \sum_{\text{all } x_i \leq x} P(X = x_i) = \sum_{\text{all } x_i \leq x} p_X(x_i) = P(X \leq x) \tag{4.1}$$

Figure 4.4 shows an example of a cumulative distribution function. Notice that the cumulative probability sums to 1.0 at the upper bound.

For continuous numbers we must define a small interval dx to describe the probability density because for a continuous number $P(X = x) = 0$ as the number goes out to an infinite number of decimal places. The probability distribution of a continuous number is called the probability density function, **PDF**, is denoted by the symbol $f_X(x)$, and is equal to the probability of the random variable over the small interval $P(x < X \leq x + \mathrm{d}x)$. Figure 4.5 shows a probability density function, of which the area under the entire curve must sum up to 1.0.

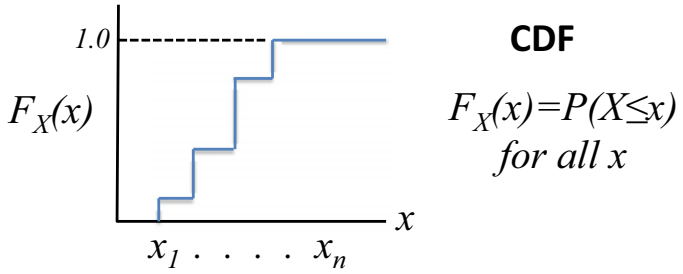


Fig. 4.4 Cumulative distribution function, CDF, for a discrete number

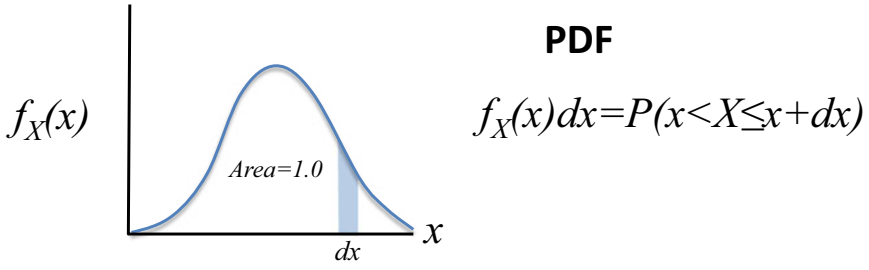


Fig. 4.5 Probability density function, PDF, for a continuous variable

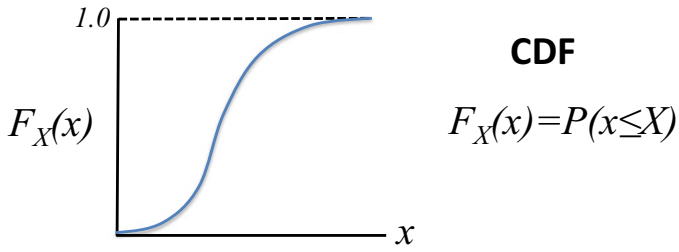


Fig. 4.6 Cumulative distribution function, CDF, for a continuous variable

We can also show the cumulative distribution function, **CDF**, for a continuous variable in the same manner as before, but instead of summing we integrate.

$$F_X(x) = \int P(x \leq X \leq x + \mathrm{d}x)$$
$$= \int_{-\infty}^{x} f_X(x)\mathrm{d}x = P(x \leq X) \tag{4.2}$$

Figure 4.6 shows the cumulative distribution function for a continuous variable, with the upper bound approaching 1.0.

By defining probability distributions in this manner we can then answer probability questions with ease. If we are dealing with a discrete number, like train crossings, and want to know how many crossings have occurred, then we consult the PMF for that specific number. If we have a continuous number, like axial load on a column, and want to know the probability of exceeding some load, then we consult the CDF for that number and take the complement. We will discuss more problem solving of this type later in the chapter.

4.1.1 Generalities of Probability Distributions

Probability distributions can take on any mathematic form (i.e., any equation) as long as they satisfy the three axioms of probability; the function must be nonnegative and probabilities of all possible values must add up to 1.0. To frame this with a CDF:

(1) $F_X(-\infty) = 0$ and $F_X(\infty) = 1.0$.
(2) $F_X(x) \geq 0$ for all x and is nondecreasing with x.
(3) $F_X(x)$ is continuous to the right.

Therefore any mathematical function satisfying the three axioms is a valid probability distribution. There are fundamentally two reasons for the existence of a probability distribution: mathematical utility (e.g., normal distribution), and conceptual

utility (e.g., Poisson distribution). An entire menagerie of pre-defined probability distributions exists that people have come up with to solve some problem requiring the mathematical or conceptual utility of those particular functions. These can be found in statistics text or reference books (e.g., Ang and Tang 2007). But to reiterate, any mathematical function that satisfies the three axioms is a valid probability distribution. Examples of common probability distributions are:

- Continuous—Uniform (i.e., a straight line), triangular, exponential, normal, lognormal, Gamma, Beta, extreme value, etc.
- Discrete—Binomial, Poisson, geometric, etc.

In this text we will only be using the normal and lognormal, for mathematical ease and simplicity, but all the discussions herein apply to any probability distribution.

Some things to keep in mind when selecting and applying a probability distribution to a particular problem:

- Does the variable of interest represent something that is infinite or does it have a finite upper bound?
- Is it a variable that can assume a negative value, or must it be positive?
- To fit a distribution to a particular problem some tricks that are often used; shifted distributions, compound or mixed distributions, or truncated and renormalized continuous distributions.

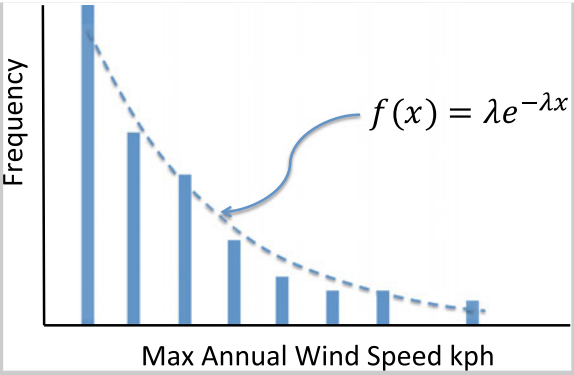
There is nothing magical about any particular mathematical function that is a probability distribution, and you simply want one that best describes the phenomenon of interest and doesn't introduce any more epistemic uncertainty than necessary. That being said the choice of a distribution is often subjective and should be thoughtfully and clearly justified.

Note: It is common to drop the subscripts on the generic names of probability distributions for brevity. Therefore:

$$\begin{aligned} &\text{PMF} \quad p_X(x) = p(x) \\ &\text{PDF} \quad f_X(x) = f(x) \\ &\text{CDF} \quad F_X(x) = F(x) \end{aligned}$$

Example: Wind Loading

A structural engineer is designing a tall tower to withstand wind loads. Based on the maximum annual wind velocity recorded over many years the histogram of the data looks like the figure below. The engineer decided to model the data using a negative exponential function because it provides a reasonable fit to the data, a subjective decision that introduced epistemic uncertainty to the problem but provided mathematical convenience in solving the problem.

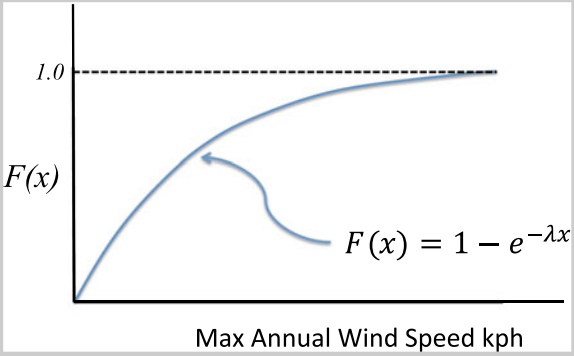


The negative exponential function is now being used as a PDF describing the probability of max annual wind speed $P(x < X \leq x + \mathrm{d}x)$.

(a) What is the CDF for this problem?

The CDF is the integral of the PDF and will allow us to answer questions about less than or greater than a particular value easily.

$$F(x) = \int_0^x f(x)\mathrm{d}x = \int_0^x \lambda \mathrm{e}^{-\lambda x}\mathrm{d}x = -\mathrm{e}^{-\lambda x}\Big|_0^x = 1 - \mathrm{e}^{-\lambda x} \quad \text{where } x \geq 0$$



(b) Data shows that the probability of the max annual wind speed less than 70 kph is 90%. Estimate the coefficient (λ) of the function.

We could use the PDF and integrate from 0 to 70, or the CDF at that value since it is the integral of the PDF.

$$0.9 = P(0 \leq X \leq 70) = P(X \leq 70) = F(x)$$

$$0.9 = 1 - e^{-\lambda 70} \quad e^{-\lambda 70} = 0.1$$

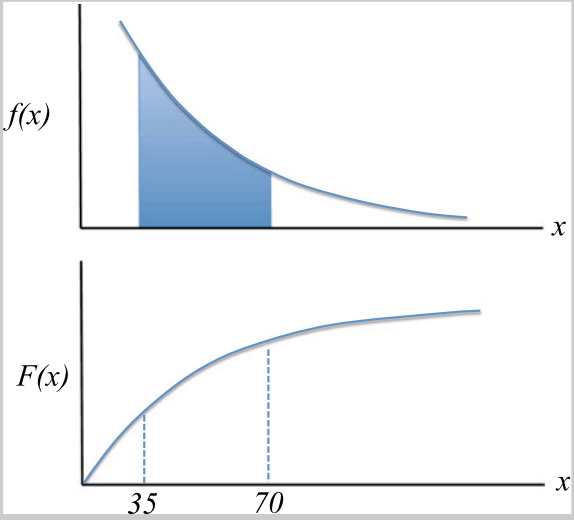
$$-\lambda 70 = \ln(0.1) \quad \lambda = \mathbf{0.033}$$

Now that we have the PDF, the CDF, and the coefficient in the equation we can move forward to answer any number of engineering-related probability questions.

(c) What's the probability of the max annual wind speed between 35 and 70 kph?

Here we want the area under the PDF from 35 to 70 kph. With the CDF this is accomplished by evaluating at 70 and then subtracting at 35 because the CDF is the pre-integrated area under the PDF.

$$P(35 \leq X \leq 70) = F(70) - F(35)$$

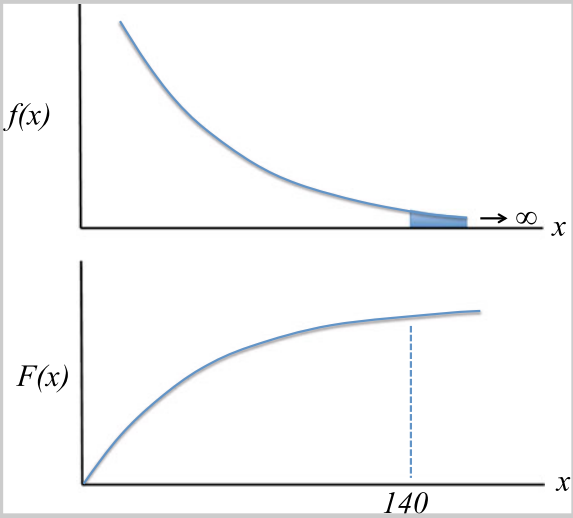


$$P(35 \leq X \leq 70) = \left(1 - e^{-0.033(70)}\right) - \left(1 - e^{-0.033(35)}\right)$$
$$= 0.900 - 0.685 = \mathbf{0.215} \text{ or } \mathbf{21.5\%}$$

(d) What is the probability that the max annual wind speed will be greater than 140 kph?

The area under the PDF that is greater than 140 kph answers this question, which is equivalent to the complement of the value of the CDF at 140 kph.

$$P(X \geq 140) = 1 - P(X \leq 140) = 1 - F(140)$$



$$P(X \geq 140) = 1 - \left(1 - e^{-0.033(14)}\right) = \mathbf{0.0099} \text{ or approximately } \mathbf{1}\%$$

4.2 Expectation and Moments

A probability distribution is a theoretical representation of the frequency of occurrence, analogous to the histograms generated from data as shown in Chap. 2. We can quantify the central tendency and dispersion of probability distributions just the same as we did with sample statistics from empirical data.

4.2.1 Central Tendency

The central tendency can be quantified as the mean, mode, or median. The **mean** for probability distributions, called the **expectation** or **first moment**, for discrete and continuous numbers, respectively, is:

$$E(x) = \mu = \sum_{i=1}^{n} x_i \cdot p(x_i) \tag{4.3}$$

$$E(x) = \mu = \int_{-\infty}^{\infty} x \cdot f(x)\mathrm{d}x \tag{4.4}$$

The **mode** is the most probable value or value of highest probability and is denoted by $\tilde{x}$. The **median**, as before with sample statistics, is the value with 50% on either side, or when the CDF is equal to 50%:

$$x_m = F(x_m) = 50\% \tag{4.5}$$

When a probability distribution is symmetric, the mean, median, and mode are equivalent.

4.2.2 Dispersion

The spread of a probability distribution can be quantified by the **variance** for discrete and continuous, respectively, as:

$$\mathrm{Var}(x) = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot p(x_i) \tag{4.6}$$

$$\mathrm{Var}(x) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)\mathrm{d}x \tag{4.7}$$

The **standard deviation**, often called the **second moment**, is the square root of the variance.

$$\sigma = \sqrt{\mathrm{Var}(x)} \tag{4.8}$$

The normalized form of dispersion is the **coefficient of variation**, which is the standard deviation divided by the mean:

$$\delta = \frac{\sigma}{\mu} \tag{4.9}$$

Higher moments can be calculated but are generally unimportant for the purposes of engineering risk analysis. The equations of the moments are presented here but in most cases these will be calculated using built-in functions in common computation software (e.g., Excel, MATLAB, Maple).

4.3 Multivariate Probability Distribution

We may encounter an engineering problem where we are mapping an event onto two lines to describe two random variables involved in the event. We are not limited to two random variables, but with two we can draw the joint probability in three dimensions. For two variables we can define the joint CDF, as well as the joint PMF for discrete numbers and joint PDF for continuous numbers:

$$\text{CDF:} F(xy) = P(X \leq x, Y \leq y) \tag{4.10}$$

$$\text{PMF:} p(xy) = P(X = x, Y = y) \tag{4.11}$$

$$\text{PDF:} f(xy)\mathrm{d}x\mathrm{d}y = P(x < X \leq x + \mathrm{d}x, y < Y \leq y + \mathrm{d}y) \tag{4.12}$$

The joint PDF can look something like Fig. 4.7, a three-dimensional surface showing the joint probability of X and Y.

Multivariate distributions, of course, follow the rules of probability. The multiplication rule is the joint probability distribution of X and Y which is equal to the conditional probability distribution of X and Y multiplied by the marginal distribution of Y:

$$f(xy) = f(x|y)f(y) \tag{4.13}$$

If X and Y are statistically independent random variables, then the multiplication rule simplifies to the product of their marginal probability distributions:

$$f(xy) = f(x)f(y) \tag{4.14}$$

The marginal distribution can be determined from a joint distribution by integrating out the other marginal distribution:

$$f(y) = \int_{-\infty}^{\infty} f(xy)\mathrm{d}x \tag{4.15}$$

Similarly all the other rules and theorems of probability apply and can be used to help solve engineering problems. If we are dealing with a joint distribution where the marginals are correlated, then we often describe the amount of correlation by the correlation coefficient (as discussed in the previous chapters) here applied to probability distributions:

$$\rho = \frac{\text{Cov}(xy)}{\sigma_x \sigma_y} \tag{4.16}$$

$$\text{Cov}(xy) = E(xy) - E(x)E(y) \tag{4.17}$$

4.4 Theoretical Distributions: Normal and Lognormal

As mentioned there is an entire menagerie of theoretical distributions that achieve some mathematical utility, conceptual utility, or both. In this text we will confine our discussion of specific theoretical distributions to the normal and lognormal distributions because they satisfy both utilities and will provide us with a quick first-order approximation for the majority of problems encountered in Civil Engineering.

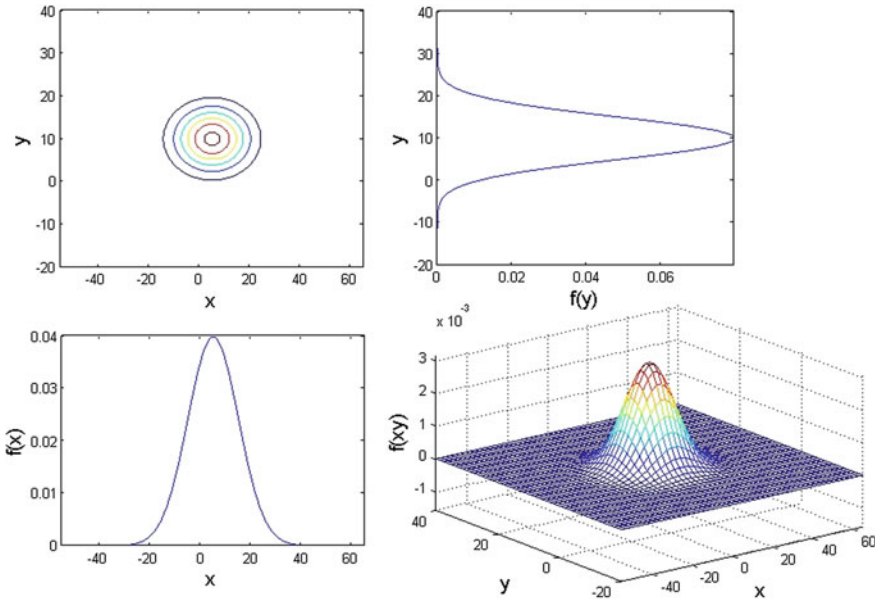


Fig. 4.7 Joint PDF of X and Y showing their marginal distributions, the joint plan view (upper left), and the joint 3D view (lower right)

4.4.1 Normal Distribution

The normal distribution, or Gaussian distribution as it is sometimes called, is defined by the following mathematical function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad \text{for} -\infty < x < \infty \tag{4.18}$$

This function is symmetric, has a "bell-shaped" curve, and is completely defined by the first (μ) and second (σ) moments. The normal distribution has been found to describe the likelihood of many natural phenomena from a very broad range of fields. The history of the normal distribution starts with gambling in the 1700s, traverses through many mathematical treatises for 100s of years, and continues today where it finds almost ubiquitous application in realms utilizing probability and mathematics. The shorthand notation used for the PDF of the normal distribution is $N(\mu, \sigma)$ (Fig. 4.8).

A change in the mean (μ) value will shift the distribution left or right, but leave the breadth of the curve unchanged. A change in the standard deviation (σ) will change the breadth of the curve, thereby altering the height of the curve, but leave the central tendency unshifted.

We can also define what is called the standard normal distribution, a normalized distribution that has a mean of zero ($\mu = 0$) and a unit standard deviation ($\sigma = 1$) which can be shorthanded to $N(0, 1)$. The indefinite integral of the standard normal distribution, that is the CDF, is denoted by the symbol $\Phi(x)$ (Fig. 4.9).

Unfortunately there is no closed-form solution to this indefinite integral; therefore the CDF of the standard normal distribution is usually presented in tabular form as can be found in Appendix A. There are also approximation equations of the CDF, also presented in Appendix A. For computational purposes the tabular or approximate CDF of the standard normal distribution is programmed into almost every handheld calculator and computational software available today (e.g., MATLAB, R, Maple,

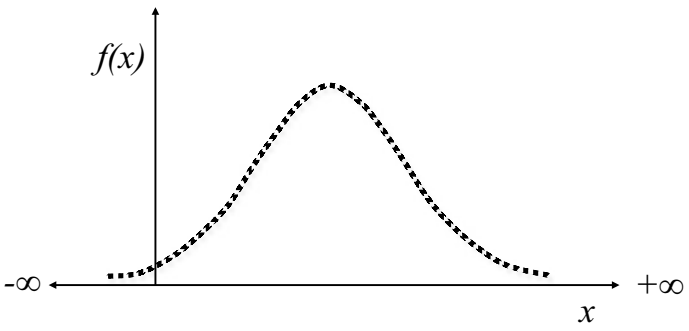


Fig. 4.8 Example PDF of the normal distribution, $N(\mu, \sigma)$

Excel). The tabulated values are given only for positive values because the standard normal CDF is symmetric; therefore:

$$\Phi(-x) = 1 - \Phi(x) \tag{4.19}$$

And the values of x for probabilities less than $p < 0.5$ are negative and can be calculated as:

$$x = \Phi^{-1}(p) = -\Phi^{-1}(1-p) \tag{4.20}$$

For the standard normal distribution the probability (or area) within ± 1 standard deviation is 68.3%, within ± 2 standard deviations is 95.4%, and within ± 3 standard deviations is 99.7%. One can see that three standard deviations on either side of the mean account for almost all of the likelihood for a problem dealing with a normally distributed random variable.

If we have a random variable that is normal (meaning that its probability distribution is described by the normal distribution), or we can reasonably assume that it is normal, we can easily solve probability problems with the tabulated standard normal distribution results:

$$P(a < X \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right) \tag{4.21}$$

It has been shown that the sum of independent random variables asymptotically approaches the normal distribution regardless of the distribution of the underlying variables. This property of the normal distribution will become important in the next chapter where we deal with functions of random variables. The above statement affirms the central limit theorem (CLT) which states that the sum of N random variables approaches normality as N becomes large:

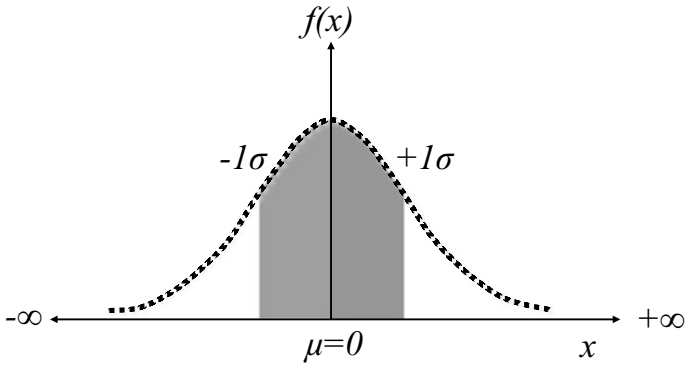


Fig. 4.9 Standard normal distribution $N(0, 1)$

$$z = k_1 + k_2 + \ldots + k_n$$

$$z = \sum k \rightarrow N(\mu_z, \sigma_z)$$

Example: River Flows

A river stage gage (in former Yugoslavia) measured peak annual flows over 39 years. The data appears to be normally distributed with $N(28{,}676;\ 21{,}117)$. Compute the probability that the peak flow will be less than 100,000; 80,000; and 50,000 in any given year, and determine the return period for each of these flows.

$$P(X \leq x) \text{ for } N(\mu, \sigma) \text{ is } \Phi(x) \text{ therefore:}$$

$$P(X \leq 100{,}000) = \Phi\left(\frac{100{,}000 - 28{,}676}{21{,}117}\right) = \Phi(3.38) \cong 0.99964$$

where we are using Appendix A to approximate $\Phi(x)$ values.

$$P(X \leq 80{,}000) = \Phi\left(\frac{80{,}000 - 28{,}676}{21{,}117}\right) = \Phi(2.43) \cong 0.99245$$

$$P(X \leq 50{,}000) = \Phi\left(\frac{50{,}000 - 28{,}676}{21{,}117}\right) = \Phi(1.01) \cong 0.84375$$

The return period (T) is usually given as the annual probability of exceedance in years, and exceedance is the complement of a value less-than-or-equal-to.

$$T = \frac{1}{P(X > x)} = \frac{1}{1 - P(X > x)} = \frac{1}{1 - \Phi(x)} \text{ therefore:}$$

$$T(100{,}000) \cong \frac{1}{1 - 0.99964} = 2777.8 \text{ yrs}$$

$$T(80{,}000) \cong \frac{1}{1 - 0.99245} = 132.5 \text{ yrs}$$

$$T(50{,}000) \cong \frac{1}{1 - 0.84375} = 6.4 \text{ yrs}$$

So for the lowest flow value (50,000) there is a probability of 84% that the flows will be less than this value, and the return period for this flow is roughly six and a half years.

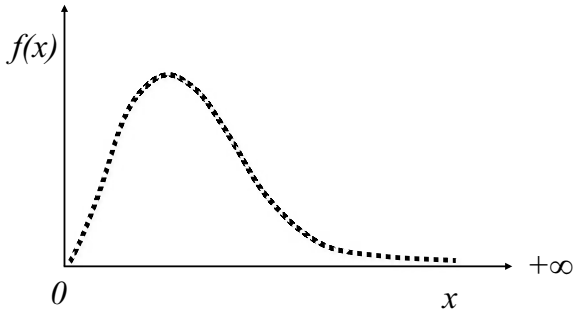


Fig. 4.10 Example PDF of the lognormal distribution, LN(λ, ξ)

4.4.2 Lognormal Distribution

The lognormal distribution is just that the log of the normal distribution. Here log refers to the natural log (ln) or log base e. The mathematical function that describes the lognormal distribution is:

$$f(x) = \frac{1}{\xi x\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln(x) - \lambda}{\xi}\right)^2\right) \quad \text{for } x \geq 0 \tag{4.22}$$

The lognormal distribution starts at zero and goes to infinity which makes this distribution often useful for describing numbers in engineering that cannot take a negative value (e.g., penetration resistance, flow rate) (Fig. 4.10).

The moments for this distribution are the log-transformed first and second moments:

$$\lambda = \ln(\mu) - \frac{1}{2}\xi^2 \tag{4.23}$$

$$\xi^2 = \ln\left(1 + \frac{\sigma^2}{\mu^2}\right) \tag{4.24}$$

For the lognormal distribution it is often more convenient to use the median and coefficient of variation because:

$$\lambda = \ln(x_m) \tag{4.25}$$

$$\xi \approx \delta \quad \text{for } \delta \leq 0.3 \tag{4.26}$$

A lognormal distribution with a median of $x_m = 10$ has a $\lambda \cong 2.303$, and the coefficient of variation δ to ξ is one to one for values less than about 30%. Since the lognormal is just the natural log of the normal, then we can use the standard normal distribution once we log transform the parameters:

$$P(a < X \leq b) = \Phi\left(\frac{\ln(b) - \lambda}{\xi}\right) - \Phi\left(\frac{\ln(a) - \lambda}{\xi}\right) \tag{4.27}$$

Because of the summation properties of the normal distribution and the mathematics of logs, it can be shown that the product of N random variables asymptotically approaches the lognormal distribution. This will also be important in the next chapter.

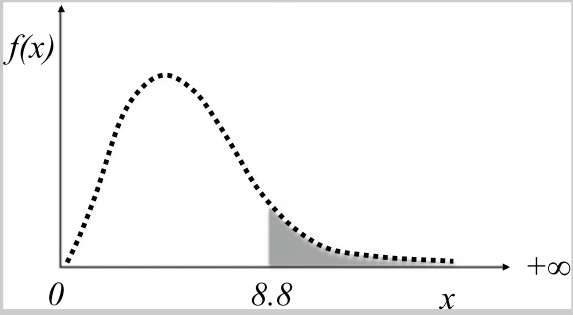
$$z = k_1 k_2 \ldots k_n$$

$$\ln(z) = \ln(k_1) + \ln(k_2) + \cdots + \ln(k_n)$$

$$\ln(z) = \sum \ln(k_i) \rightarrow \mathrm{LN}(\lambda_z, \xi_z)$$

Example: Rainfall Intensity

The average annual rainfall over a 24 h period in San Luis Obispo in the fall months is $\bar{x} = 2.3$ cm. A rather intense storm came through on 13 October 2009 that dropped 8.8 cm within 24 h. What is the probability that there could be a storm with a higher 24 h total? Since the rainfall total cannot be less than zero, we will assume that the distribution is lognormal. The shaded area in the figure represents the question we are asking.



There is no information on the variability, and only the average was reported by the weather service. Here we will use the "6 Sigma" approach to make an estimate of the uncertainty to arrive at a rough first-order estimate of the probability. For this problem we will assume that the maximum intensity could be something like 15 cm and the minimum would be no rain or 0 cm. Dividing this range by 6 gives a sample standard deviation of $s = 2.5$ cm.

Solving for the lognormal parameters given the sample mean and estimated sample standard deviation:

$$\xi^2 = \ln\left(1 + \frac{\sigma^2}{\mu^2}\right) \approx \ln\left(1 + \frac{s^2}{\bar{x}^2}\right) = \ln\left(1 + \frac{2.5^2}{2.3^2}\right) = 0.78$$

$$\xi = \sqrt{0.78} = 0.88$$

$$\lambda = \ln\mu - \frac{1}{2}\xi^2 \approx \ln\bar{x} - \frac{1}{2}\xi^2 = \ln(2.3) - \frac{1}{2}(0.78) = 0.44$$

With the parameters of the lognormal distribution we can now solve for the probability of exceedance.

$$\begin{aligned} P(X > 8.8) &= 1 - P(X \leq 8.8) = 1 - \Phi\left(\frac{\ln x - \lambda}{\xi}\right) \\ &= 1 - \Phi\left(\frac{\ln 8.8 - 0.44}{0.88}\right) = 1 - \Phi(1.97) \approx 1 - 0.975 \\ &= \mathbf{0.025} \end{aligned}$$

So the first-order answer is that there is a 2.5% chance that a storm can dump more rain than the one that occurred on 13 October 2009. This is of course a function of our assumptions. The maximum assumed value may be underestimated, and if this is so we are estimating a lower probability. The assumption that it follows the lognormal distribution which is infinite in the positive direction gives us a higher probability. But given the minimal information we started with, we were able to estimate that this storm was a relatively intense storm and that the odds of a more intense one are not high.

Example: Structural Supports

A structure has three supports A, B, and C. The loads on these supports can be estimated with reasonable accuracy, but the soil conditions are heterogeneous below the support footings. Assume the settlements S_A, S_B, and S_C can be defined as independent normal variates based on subsurface field testing with:

$$\begin{aligned} \mu_{S_A} &= 3.0\,\text{cm} \quad \delta_{S_A} = 0.20 \\ \mu_{S_B} &= 2.5\,\text{cm} \quad \delta_{S_B} = 0.30 \\ \mu_{S_C} &= 3.0\,\text{cm} \quad \delta_{S_C} = 0.25 \end{aligned}$$

What is the probability that the total settlement exceeds 3.5 cm, which would result in the slab of the structure being misaligned with the incoming underground utilities?

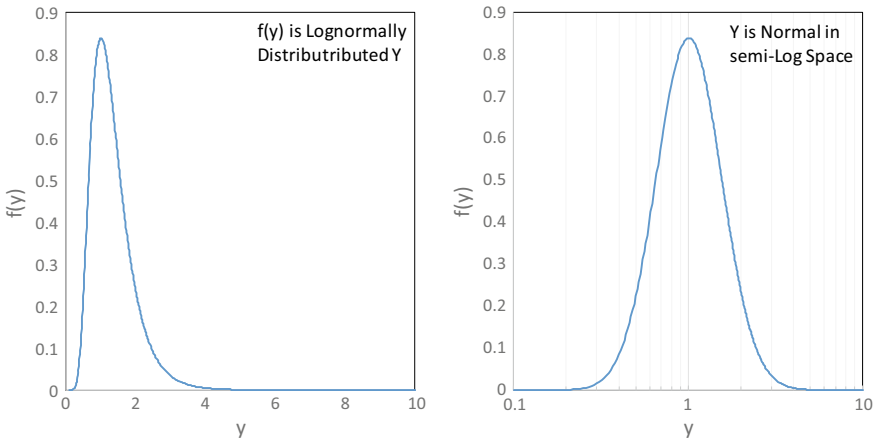


Fig. 4.11 Starting with a lognormally distributed variable (left figure), we replot it in semilog space (right figure)

$$
\begin{aligned}
P(S > 3.5) &= P(S_A > 3.5 \cap S_B > 3.5 \cap S_C > 3.5) \\
&= P(S_A > 3.5)P(S_B > 3.5)P(S_C > 3.5) \quad \text{statistically independent} \\
&= (1 - P(S_A \leq 3.5))(1 - P(S_B \leq 3.5)) \\
&\qquad (1 - P(S_C \leq 3.5)) \quad \text{complement} \\
&= \left(1 - \Phi\left(\frac{3.5 - \mu_A}{\sigma_A}\right)\right)\left(1 - \Phi\left(\frac{3.5 - \mu_B}{\sigma_B}\right)\right)\left(1 - \Phi\left(\frac{3.5 - \mu_C}{\sigma_C}\right)\right) \\
&\qquad \text{standard normal with } \sigma = \delta \cdot \mu \\
&= (1 - \Phi(0.83))(1 - \Phi(1.33))(1 - \Phi(0.67)) \\
&\cong (0.20)(0.09)(0.25) = \mathbf{0.005}
\end{aligned}
$$

The probability of any of the supports exceeding 3.5 cm is roughly 0.5%, essentially negligible from a settlement standpoint.

The lognormal distribution is commonly used in engineering practice for representing many hazards that by their nature must take positive numbers (e.g., ground motions, flood levels, wave heights, fault displacements). We are often concerned with the probability of exceedance; for example, what is the annualized likelihood of exceeding some engineering threshold that could cause damage to our Civil Engineering feature. Below are a series of figures (Figs. 4.11, 4.12, and 4.13) showing how these exceedance curves are generated from a lognormal distribution (after Steve Thompson, LCI).

So we have represented the probability distribution of a hazard with the lognormal distribution. Then following the process above we have arrived at an annualized

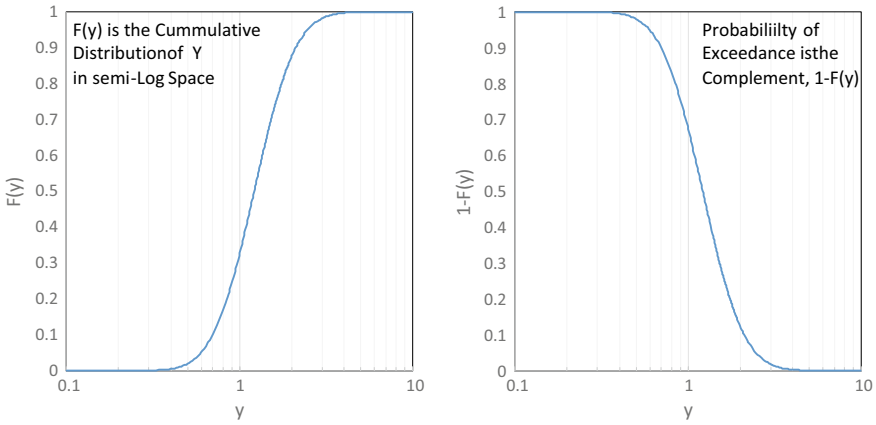


Fig. 4.12 The cumulative distribution of the variable is now presented (left plot), and the complement of that provides the exceedance curve (right plot)

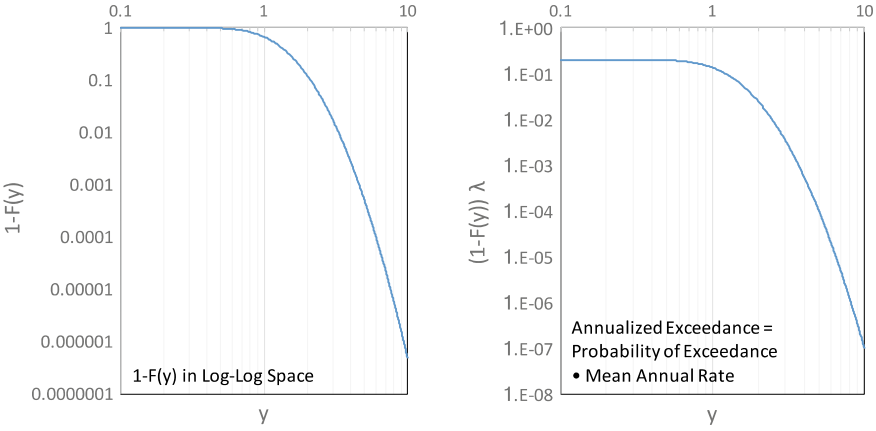


Fig. 4.13 The exceedance curve in log–log space (left figure) gives the characteristic downsloping function, which when multiplied by a mean annual rate will produce an annualized probability of exceedance curve (right figure)

probability of exceedance curve. The mean annual rate would be calculated from some data on the hazard using tools presented in Chap. 2. The final figure (Fig. 4.13 right) provides results that can be compared to other annualized hazards and if multiplied by consequences can be compared to other annualized risks as presented in Table 1.1.

4.5 Chapter Summary

- A random variable is an event mapped to a real number line. This allows for mathematical calculations of likelihood.
- A probability distribution shows all possible values a random variable can assume and the relative likelihoods of those values.
- Numbers can be discrete or continuous, and the probability distributions reflect this. For discrete numbers there is the probability mass function, **PMF**, for continuous numbers there is the probability distribution function, **PDF**, and the sum or integral of these results in the cumulative distribution function, **CDF**. Each of these distributions refers to a specific relationship between the number line and the likelihood of a number or range of numbers.
- Any mathematical function can be a probability distribution as long as it satisfies the three axioms of probability. There are a number of common pre-defined probability distributions, of which we will be focusing on the normal and lognormal distribution.
- The **first** and **second moment** of a distribution refers to the **mean** and **standard deviation** of the distribution.
- When two or more random variables are of interest, we consider a multivariate distribution, where commonly the correlation or independence of the marginal distributions is the primary concern.
- The normal and lognormal distributions satisfy both mathematical and conceptual utilities for most Civil Engineering problems and can be used to quickly make a first-order assessment of the probability.
- No closed-form solution for the integral of the normal distribution exists so it is common to use tabulated results of the standard normal distribution $N(0, 1)$ in its CDF form $\Phi(x)$ found in Appendix A.
- Annualized probability of exceedance curves is regularly generated using the lognormal distribution and when multiplied by consequences can give annualized risk values to compare with other risks.

Reference

Ang, A. H.-S., & Tang, W. H. (2007). *Probability concepts in engineering: Emphasis on applications to civil and environmental engineering* (2nd ed.). New York: Wiley.

Chapter 5
Functions of Random Variables: Error Propagation

This chapter contains the core of engineering risk analysis, the central theme about which the rest of the material pivots. In Civil Engineering almost all designs involve some calculation using an equation that is based on the physics of the problem, empirical data of the phenomenon, or a combination of the two (often called semi-empirical equations). The parameters in these equations are routinely treated as deterministic and it is common to use mean or median values for calculation purposes. But if we treat these parameters as random variables and propagate the uncertainty through the equations we get a much better understanding of the most likely answer, as well as an understanding of how much confidence we should have in that most likely answer. Through this process of error propagation we fully characterize the problem by accounting for the uncertainty of the input variables and their mathematical interrelationship as described in the engineering equation. This sets the stage for determining how accurate the answer is, how much confidence we can have in the mean or median, and how best to proceed to ensure a reliable engineering design.

The phrase **functions of random variables** can be defined as "equations, formula, or mathematical models that contain parameters with uncertainty." The application of this can be demonstrated through the following simple example. Let's say we have the following equation:

$$z = x + y$$

If the input parameters are treated as deterministic, meaning they have no uncertainty or more likely they have uncertainty but it is neglected, then we can simple add the values to determine the results. Let's say that $x = 5$ and $y = 4$, then the solution is $z = 9$. But if we include the uncertainty of the parameters, treating them as random variables, and the probability distributions of the random variables are similar to those shown in Fig. 5.1, then the solution is not evident. How do we sum the discrete distribution of X and the continuous distribution of Y to get the resultant distribution of Z? Should the resultant distribution be discrete or continuous? How are the probabilities combined?

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_5

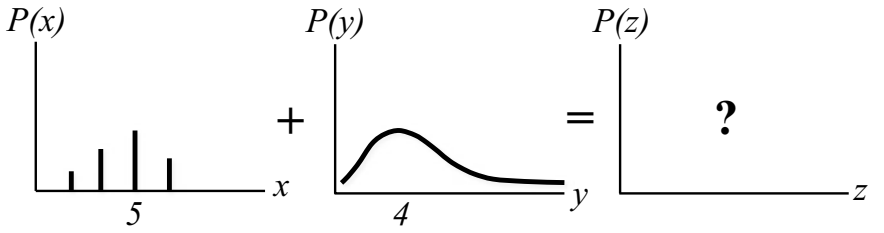


Fig. 5.1 A function of random variables, $Z = X + Y$

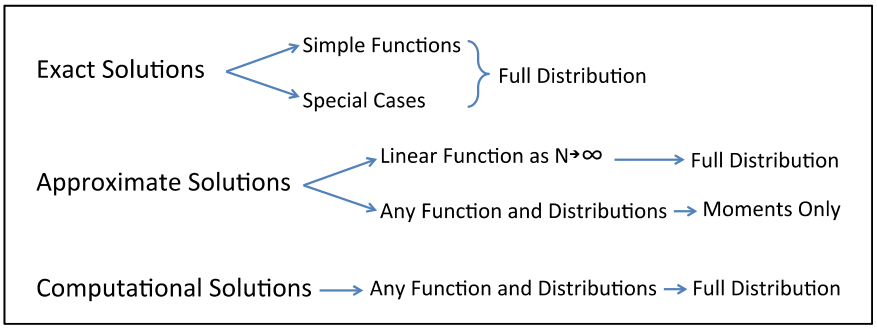
It turns out that the solution depends on:

(a) The type of mathematical function,
(b) The specific distributions of the random variables, and
(c) How much information we need of the resultant.

There are three groups of solutions for functions of random variables depending on the answers to a, b, and c above. The three groups are:

1. Exact solutions,
2. Approximate solutions, and
3. Computational solutions.

The box below shows how these three groups map out as presented in the subsequent discussion.



5.1 Exact Solutions

Exact solutions will provide the full distribution of the resultant but only work when we have a simple function or a special case. Simple functions are one-to-one functions with a single root. Special cases are when the function is (1) a sum or difference of normally distributed random variables or (2) a product or quotient of lognormally distributed random variables.

5.1.1 Simple Function

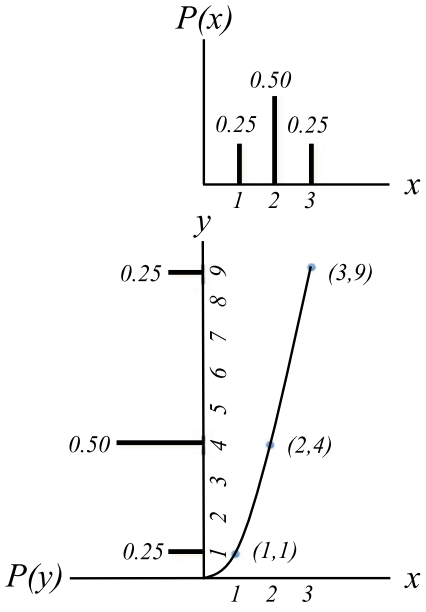
An example of a one-to-one function with a single root is:

$$Y = X^2 \text{ where } x \geq 0$$

In this example there is a direct relationship between the probability of X and the resulting probability of Y. Take the values given in the table below. If we map this we see the direct relationship.

x	$P(x)$	y	$P(y)$
1	0.25	1	0.25
2	0.50	4	0.50
3	0.25	9	0.25
4	0	16	0

For simple functions such as this there is a one-to-one relationship of probability, $p(y) = p(x)$. If we had a continuous number then $f(y) = f(x)$. The CDFs would follow accordingly. Unfortunately there are few situations in engineering where we are working with a simple one-to-one function with a single root.



5.1.2 Special Cases

The special cases are a byproduct of the properties of the normal distribution. The first special case is a function that is a sum or difference of normally distributed random variables. If we return to the function we were initially evaluating:

$$Z = X + Y$$

But now X and Y are normally distributed random variables with their respective moments, $N(\mu_x \sigma_x)$ and $N\left(\mu_y \sigma_y\right)$, then:

$$\mu_z = \mu_x + \mu_y \tag{5.1}$$

From this it can be said that "the mean of the sums is the sum of the means." The squared second moment is:

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\rho\sigma_x\sigma_y \tag{5.2}$$

These results show that the resultant Z is a normally distributed random variable with moments, $N(\mu_z, \sigma_z)$. If the random variables of the input parameters are statistically independent, that is $\rho = 0$, then the squared second moment (a proof be found in detail in Appendix B) reduces to:

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 \tag{5.3}$$

So when the input parameters are normally distributed and the function is a sum we see the solution is simply the sum of the means and the sum of the variances. We can expand this to any number of random variables along with their coefficients to broaden the application to any sum or difference of Gaussian random variables.

$$Z = \sum_{i=1}^{n} a_i Q_i \tag{5.4}$$

where the a's are coefficients (fixed deterministic values) and the Q's are normally distributed random variables. Then generically:

$$\mu_z = \sum_{i=1}^{n} a_i \mu_{Q_i} \tag{5.5}$$

$$\sigma_z^2 = \sum_{i=1}^{n} a_i^2 \sigma_{Q_i}^2 + \sum_{i,j=1}^{n} \sum_{i \neq j}^{n} a_i a_j \rho_{Q_i Q_j} \sigma_{Q_i} \sigma_{Q_j} \tag{5.6}$$

Note when we have a difference equation the coefficient of -1 is squared in the first term and becomes a $+1$ having no effect, but does influence the second term that accounts for the correlation. The 2 that we see in front is because of the double sum of the i's and j's (e.g., when there are 2 terms we would see the combinations of 1,2 and 2,1 giving us the factor of two in the double sum). If the random variables are statistically independent, that is $\rho = 0$, then the double sum goes to zero and we are left with:

$$\sigma_Z^2 = \sum_{i=1}^{n} a_i^2 \sigma_{Q_i}^2 \tag{5.7}$$

Notation Clarity

Depending on how i's and j's are treated will dictate how the variance can be written. The following are all equivalent notations resulting in the exact same solution for the squared second moment. The first is the most concise while the third is the most explicit.

$$\sigma^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j \rho_{ij} \sigma_i \sigma_j \tag{5.8a}$$

$$\sigma^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 + \sum_{i,j=1}^{n}\sum_{i\neq j}^{n} a_i a_j \rho_{ij} \sigma_i \sigma_j \tag{5.8b}$$

$$\sigma^2 = \sum_{i=1}^{n} a_i^2 \sigma_i^2 + 2\sum_{i=1}^{n}\sum_{j=i+1}^{n} a_i a_j \rho_{ij} \sigma_i \sigma_j \tag{5.8c}$$

The second special case is where the function is a product or quotient of lognormally distributed random variables. This is because when we take the natural log of a product we have the sum, or when we take the natural log of a quotient we have the difference. And the sum or difference will be of a log transformed normal distribution.

$$Z = X \cdot Y \text{ becomes } \ln Z = \ln X + \ln Y \text{ when we log both sides.}$$
$$Z = X/Y \text{ becomes } \ln Z = \ln X - \ln Y \text{ when we log both sides.}$$

Where: $\ln X$ is $N(\mu_X, \sigma_X)$ and X is $\text{LN}(\lambda_X, \xi_X)$
$\ln Y$ is $N(\mu_Y, \sigma_Y)$ and Y is $\text{LN}(\lambda_Y, \xi_Y)$

Therefore: $\ln Z$ is $N(\mu_Z, \sigma_Z)$ and Z is $\text{LN}(\lambda_Z, \xi_Z)$

We are now back to a sum or difference of normally distributed random variables. When we have the product of X and Y which are lognormally distributed random variables the first and second moments are:

$$\lambda_Z = \lambda_X + \lambda_Y \tag{5.9}$$

$$\xi_Z^2 = \xi_X^2 + \xi_Y^2 + 2\rho\xi_X\xi_Y \tag{5.10}$$

And when the random variables are statistically independent:

$$\xi_Z^2 = \xi_X^2 + \xi_Y^2 \tag{5.11}$$

This can be similarly expanded for any number of random variables with coefficients in the same manner as with normally distributed random variables, using the same equivalent notation as presented in Notation Clarity.

Example: Waste Treatment Facility

The annual operating cost function for waste treatment plants can be written:

$$C = \frac{WF}{\sqrt{E}}$$

where W is waste, F is the cost factor, and E is efficiency on an annualized basis. All these parameters are treated as lognormal random variables in this problem, LN(λ, ξ).

	Median value	Coefficient of variation (%)
W	2000 metric tons/year	20
F	\$20/metric ton	15
E	1.6	12.5

With the given information, what is the probability that the cost will be greater than \$35,000 and the waste contractor must make financial contingency plans for the year?

Because the function is a product/quotient and the random variables are lognormal we can use an exact solution. We first solve for the moments of C. Taking the natural log of both sides of the function:

$$\ln C = \ln W + \ln F - \frac{1}{2}\ln E$$

$$\lambda_C = \lambda_W + \lambda_F - \frac{1}{2}\lambda_E$$

$$= \ln 2000 + \ln 20 - \frac{1}{2}\ln 1.6 = 10.36$$

$$\xi_C^2 = \xi_W^2 + \xi_F^2 + \left(-\frac{1}{2}\right)^2 \xi_E^2$$

$$= (0.20)^2 + (0.15)^2 + \left(-\frac{1}{2}\right)^2 (0.125)^2 = 0.067$$

$$\xi_C = \sqrt{0.067} = 0.259$$

Now that we have the moment of C we can evaluate the probability of exceedance.

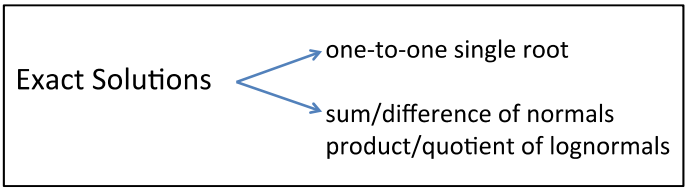
$$P(C > 35{,}000) = 1 - P(C \leq 35{,}000) \text{ complement}$$

$$= 1 - \Phi\left(\frac{\ln 35{,}000 - \lambda_C}{\xi_C}\right) \text{ as } C \text{ is LN}(\lambda_C, \xi_C)$$

$$= 1 - \Phi(0.398) \cong 1 - 0.655 = \mathbf{0.345}$$

There is roughly a 35% chance that the annual cost will exceed $35,000. The waste contractor would most likely prepare for a cost overrun given these odds. If the decision was close the contractor may want to spend more time on the probability analysis by investigating the distributions of the independent variables.

In summary, for the special cases of sum/difference of normally distributed random variables, and product/quotient of lognormally distributed random variables we can solve for the full distribution of the resultant. And in the situation where we have a simple one-to-one function with a single root we can also solve for the full distribution of the resultant. These are the two situations where we have exact solutions.



If, however, we have a function or random variables that don't fit these cases (which is typical in Civil Engineering) then we must resort to other solution techniques.

5.2 Approximate Solutions

There are two common approximate solutions. The first invokes the central limit theorem (CLT), and the second uses a Taylor series expansion.

5.2.1 Central Limit Theorem Approximation

The central limit theorem states that the sum of a large number of individual random components, none of which are significantly more dominant than the others, tends toward a Gaussian distribution as the number of components increases. This holds true regardless of the underlying distribution of the individual components. This would then give us similar results as the sum/difference of normally distributed random variables:

$$Z = \sum_{i=1}^{n} a_i Q_i \tag{5.12}$$

$$\mu_Z \approx \sum_{i=1}^{n} a_i \mu_{Q_i} \tag{5.13}$$

$$\sigma_Z^2 \approx \sum_{i=1}^{n} a_i^2 \sigma_{Q_i}^2 + \sum_{i,j=1}^{n} \sum_{i \neq j}^{n} a_i a_j \rho_{Q_i Q_j} \sigma_{Q_i} \sigma_{Q_j} \tag{5.14}$$

As $i \to \infty$ then Z approaches the normal distribution, $N(\mu_z, \sigma_z)$, even if the Q's are not normally distrusted. The drawback with this approximation is that we don't know how large i needs to be to give a good approximation, and it requires validation in most cases.

Example: Linear Equation

The equation below is a sum of random variables (some with negative coefficients) and the specified correlation.

$$Y = V - T + G - Q \quad \text{where } \rho_{VT} = 0.5$$

Invoking the central limit theorem to approximate the moments of the dependent variable:

$$\mu_Y \approx \mu_V + (-1)\mu_T + \mu_G + (-1)\mu_Q$$

$$\sigma_Y^2 \approx \sigma_V^2 + (-1)^2\sigma_T^2 + \sigma_G^2 + (-1)^2\sigma_Q^2 + 2(1)(-1)0.5\sigma_V\sigma_T$$

The assumption with this solution is that the dependent variable is Gaussian with the above moments.

5.2.2 First Order Second Moment

A more verifiable solution is what is called the mean value first order second moment or more simply the first order second moment (FOSM) approach. This solution works for any type of function and with variables of any distribution. The drawback is that we can only solve for the first and second moments, and not the full distribution.

With this method we can find the moments of any functional form or mathematical equation $g(X)$ of a random variable X. If we say that:

$$Y = g(X) \tag{5.15}$$

Then the first and second moments would be:

$$E(Y) = \int_{-\infty}^{\infty} g(X) \cdot f(x)\mathrm{d}x \tag{5.16}$$

$$\mathrm{Var}(Y) = \int_{-\infty}^{\infty} [g(x) - \mu_x]^2 \cdot f(x)\mathrm{d}x \tag{5.17}$$

To solve for the mean and variance of X we need the PDF $f(X)$. But in many cases the PDF is not available or assuming a PDF introduces too much epistemic uncertainty, which lends to an approximate solution for the mean and the variance. To approximate the moments we expand the function in a Taylor series about the mean value of X:

$$g(X) = g(\mu_X) + (X - \mu_X)\frac{\mathrm{d}g}{\mathrm{d}X} + \frac{1}{2}(X - \mu_X)^2\frac{\mathrm{d}^2 g}{\mathrm{d}X^2} + \cdots \tag{5.18}$$

where the derivatives are also evaluated at the mean value of X. It is common to truncate the series above the linear terms thereby giving a first order approximate of the moments:

$$g(X) \approx g(\mu_X) + (X - \mu_X)\frac{\mathrm{d}g}{\mathrm{d}X} \tag{5.19}$$

$$E(Y) = \mu_Y \approx g(\mu_x) \tag{5.20}$$

$$\mathrm{Var}(Y) = \sigma_Y^2 \approx \sigma_X^2 \left(\frac{\mathrm{d}g}{\mathrm{d}X} \right)^2 \tag{5.21}$$

The first order approximate will provide reasonable accuracy in situations where the function is not highly nonlinear and in situations where the variance is not large. If more accuracy is warranted a second order approximation can be carried out using higher order terms in the Taylor series expansion. Here we reevaluate the mean using a second order approximate:

$$E(Y) \approx g(\mu_x) + \frac{1}{2}\sigma_X^2 \frac{\mathrm{d}^2 g}{\mathrm{d}X^2} \tag{5.22}$$

In most Civil Engineering problems a first-order (FOSM) estimate is sufficient for quantifying the mean and standard deviation. We can expand this discussion to include multiple random variables:

$$Y = g(X_1, X_2, \ldots, X_n) \tag{5.23}$$

$$E(Y) = \mu_Y \approx g\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_n}\right) \tag{5.24}$$

$$\begin{aligned} \mathrm{Var}(Y) = \sigma_Y^2 \approx & \sum_{i=1}^{n} \sigma_{X_i}^2 \left(\frac{\partial g}{\partial X_i} \right)^2 \\ & + \sum_{i,j=1}^{n} \sum_{i \neq j}^{n} \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j} \frac{\partial g}{\partial X_i} \frac{\partial g}{\partial X_j} \end{aligned} \tag{5.25}$$

When the random variables, X's, are statistically independent, that is $\rho = 0$, then the double sum goes to zero and we are left with:

$$\mathrm{Var}(Y) = \sigma_Y^2 \approx \sum_{i=1}^{n} \sigma_{X_i}^2 \left(\frac{\partial g}{\partial X_i} \right)^2 \tag{5.26}$$

If a higher order estimate of the mean for multiple random variables is desired, the second-order Taylor series expansion takes the form:

$$E(Y) = \mu_Y \approx g\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_n}\right) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \sigma_{X_i} \sigma_{X_j} \left(\frac{\partial^2 g}{\partial X_i \partial X_j} \right) \tag{5.27}$$

If the independent variables are uncorrelated then Eq. (5.27) reduces to:

$$E(Y) = \mu_Y \approx g\left(\mu_{X_1}, \mu_{X_2}, \ldots, \mu_{X_n}\right) + \frac{1}{2} \sum_{i=1}^{n} \sigma_{X_i}^2 \left(\frac{\partial^2 g}{\partial X_i^2} \right) \tag{5.28}$$

Example: Equation of a Line

In this example we will evaluate the equation of a line where the parameters are all random variables with their respective means and standard deviations. This would be the case if we performed a linear regression while keeping track of the uncertainty in the input data and the slope and intercept terms. Here we will assume that the random variables are statistically independent but that would obviously not be the case for the slope and intercept terms. Using FOSM to approximate the moments of the dependent variable:

$$y = m \cdot x + b$$

$$\mu_y \approx \mu_m \cdot \mu_x + \mu_b$$

$$\sigma_y^2 \approx \sigma_m^2 \left(\frac{\partial y}{\partial m}\right)^2 + \sigma_x^2 \left(\frac{\partial y}{\partial x}\right)^2 + \sigma_b^2 \left(\frac{\partial y}{\partial b}\right)^2$$

$$\text{where} \quad \frac{\partial y}{\partial m} = x, \quad \frac{\partial y}{\partial x} = m, \quad \frac{\partial y}{\partial b} = 1$$

$$\begin{aligned} \sigma_y^2 &\approx \sigma_m^2 (x)^2 + \sigma_x^2 (m)^2 + \sigma_b^2 \\ &\approx \sigma_m^2 \cdot \mu_x^2 + \sigma_x^2 \cdot \mu_m^2 + \sigma_b^2 \quad \text{about the mean} \end{aligned}$$

One benefit of performing FOSM is that it provides a sensitivity analysis on the function and how the uncertainty of the variables individually influences the resultant. If we look at the variance of a function of random variables we see that the uncertainty from the input parameter propagates to the resultant as the product of the variance and the squared partial derivative of the function with respect to the input parameter. The variance of the input parameter is weighted according to how it participates in the mathematics of the function. When divided by the total resultant uncertainty this provides the percent relative contribution to variance (RCV%) from each random variable:

$$\text{RCV\%} = \sigma_{X_i}^2 \left(\frac{\partial g}{\partial X_i}\right)^2 / \sigma_Y^2 \tag{5.29}$$

Example: Manning's Equation

Manning's equation is an empirical equation that is used to determine the velocity (V) in meters per second of uniform flow in an open channel:

$$V = \frac{R^{2/3} S^{1/2}}{n}$$

where:
R = hydraulic radius (m)
S = slope of the energy line (%)
n = empirical roughness coefficient of the channel.

If the channel is an open concrete rectangular section the mean values and coefficients of variation of the variables are:

Variable	μ	δ
R	2 m	0.05
S	1%	0.10
n	0.013	0.10

If we assume the variables are statistically independent, then the FOSM solution of the mean and variance is:

$$\mu_V \approx \frac{\mu_R^{2/3}\mu_S^{1/2}}{\mu_n} = \frac{(2)^{2/3}(1)^{1/2}}{0.013} = 122.108 \text{ mps}$$

$$\sigma_V^2 \approx \sigma_R^2\left(\frac{2}{3}\frac{\mu_R^{-1/3}\mu_S^{1/2}}{\mu_n}\right)^2 + \sigma_S^2\left(\frac{1}{2}\frac{\mu_R^{2/3}\mu_S^{-1/2}}{\mu_n}\right)^2 + \sigma_n^2\left(-\frac{\mu_R^{2/3}\mu_S^{1/2}}{\mu_n^2}\right)^2$$

$$= (2 \times 0.05)^2\left(\frac{2}{3}\frac{(2)^{-2/3}(1)^{1/2}}{0.013}\right)^2 + (1 \times 0.01)^2\left(\frac{1}{2}\frac{(2)^{2/3}(1)^{-1/2}}{0.013}\right)^2$$

$$+ (0.013 \times 0.10)^2\left(-\frac{(2)^{2/3}(1)^{1/2}}{(0.013)^2}\right)^2$$

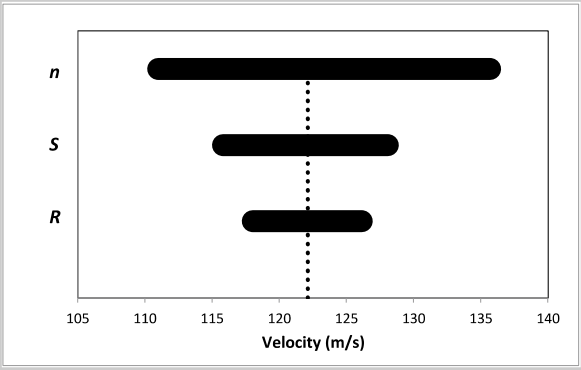
$$= (0.01 \times 2485.052) + (0.01 \times 3727.576) + (0.0000169 \times 88{,}226{,}676.230)$$

$$= 24.851 + 37.276 + 149.103 = 211.229$$

$$\sigma \approx \sqrt{211.23} = 14.534$$

For this problem the relative contribution to the total variance (RCV%) is: ***R*** of 11.76%, ***S*** of 17.65%, and ***n*** of 70.59%. It can be seen that the uncertainty in the input roughness coefficient, n, has the largest impact on the total uncertainty of the resultant velocity, V. This relative uncertainty impact is commonly expressed through a sensitivity study. Each variable within a function is varied about its mean by plus/minus one standard deviation (±sigma), and the effect on the dependent variable is then shown as a "tornado" plot. Visually we see

how the uncertainty of the roughness coefficient, n, has the most influence on the resultant velocity, V, in this problem.



To check the accuracy of the solution we can evaluate the second-order approximation of the mean velocity, V, following Eq. (5.28):

$$\mu_V \approx \mu_V + \frac{1}{2}\left[\sigma_R^2\left(-\frac{2}{9}\frac{\mu_R^{-4/3}\mu_S^{1/2}}{\mu_n}\right) + \sigma_S^2\left(-\frac{1}{4}\frac{\mu_R^{2/3}\mu_S^{-3/2}}{\mu_n}\right) + \sigma_n^2\left(2\frac{\mu_R^{2/3}\mu_S^{1/2}}{\mu_n^3}\right)\right]$$

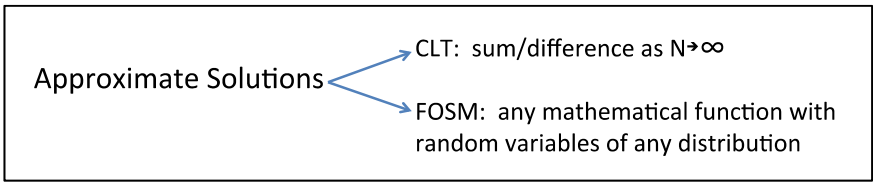
$$= 122.11 + \frac{1}{2}\left[(2\times 0.05)^2\left(-\frac{2}{9}\frac{(2)^{-4/3}(1)^{1/2}}{0.013}\right) + (1\times 0.01)^2\left(-\frac{1}{4}\frac{(2)^{2/3}(1)^{-3/2}}{0.013}\right) + (0.013\times 0.10)^2\left(2\frac{(2)^{2/3}(1)^{1/2}}{(0.013)^3}\right)\right]$$

$$= 122.108 + \frac{1}{2}[-0.068 - 0.305 + 2.442] = 123.142$$

The first-order estimate is less than 1% lower than the second order estimate, providing confidence in the approximate solution of this problem.

In summary, for situations that don't lend to an exact solution we can approximate the solution using the two methods: CLT and FOSM. The central limit theorem approach is limited to a sum/difference function and the resultant distribution is assumed normal, but not verified.

The first order second moment approach works for any type of mathematical function with random variables of any distribution. FOSM is an extremely versatile approach and works when the moments are sufficient information to proceed with the problem. For most Civil Engineering problems this is adequate, therefore FOSM is the "workhorse" of propagating uncertainty.



A means of confirming the approximate solutions and providing confidence in the answer is a computational solution as discussed in the next section.

5.3 Monte Carlo Simulations

Monte Carlo (MC) simulations are a "brute force" computation approach to solve probabilistic problems. MC simulations are used in a wide variety of fields, and not restricted to probability problems, giving rise to a vast depth of literature on different sampling techniques, computational efficiency, and numerical algorithms (e.g., Ripley 1987). For our purposes we will be using MC simulations for functions of random variables to complement the exact or approximate solutions.

The term Monte Carlo comes from the casino in Monaco and is named thus because the method relies on a random simulation of outcomes similar to the process of gambling. In order to solve the simple sum problem that was posed at the beginning of the chapter with a Monte Carlo simulation we

- randomly simulate a realization from the discrete distribution of X,
- randomly simulate a realization from the continuous distribution of Y,
- add the two realizations together as this particular mathematical function calls for, which produces a realization of Z,
- then repeat this process many many times to characterize the full distribution of Z.

We are in essence discretizing the entire problem, generating histograms of the input parameter distributions, and carrying out the function's mathematics a very large number of times until we have fully developed the resultant distribution.

With rapid computational speeds and pre-programmed probability distributions we can accomplish 10,000–1,000,000 simulations in fractions of a second. The resultant distribution is an approximation of the answer but as the number of simulations becomes large the line between approximate and exact becomes blurred, dictated by machine precision or machine epsilon.

Generating random realizations can be accomplished using any number of computational programs (e.g., MATLAB, R, Mathcad, Excel). In this text we will be focusing on MATLAB and R because of its ease and capabilities in simulating numbers, but the discussions pertain to any similar computational program.

There is an entire body of literature that discusses how to generate a random number based on a fixed computer algorithm (e.g., Knuth 1997), which is an inherently contradictory problem. For our purposes we assume that the random number generated is random enough for our calculations.

The best way to discuss Monte Carlo simulations methods is through example.

Example: Sum of Lognormals

We need to analyze the sum function:

$$S = X_1 + X_2 + X_3$$

If the input parameters were normally distributed then we could use an exact solution, but in this problem we are given that the input parameters are lognormally distributed, X_i are LN(λ_i, ξ_i). The sum of lognormal variables does not produce a lognormal resultant.

We can use FOSM to approximate the moments of S, given the moments of the X_i. We can also use Monte Carlo simulations, and then compare the results of the two methods. The information given for the X_i:

	X_1	X_2	X_3
μ	500	600	700
δ	0.50	0.60	0.70

Since no correlation coefficient is given we assume that the X_i are statistically independent.

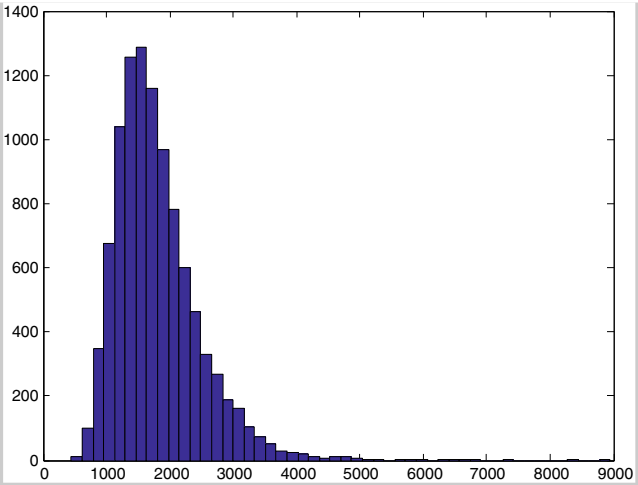
FOSM gives (which is the same answer arrived at using CLT approximation):

$$\mu_S \approx \mu_{X_1} + \mu_{X_2} + \mu_{X_3} = 500 + 600 + 700 = 1800$$

$$\sigma_S^2 \approx \sigma_{X_1}^2\left(\frac{\partial g}{\partial X_1}\right)^2 + \sigma_{X_2}^2\left(\frac{\partial g}{\partial X_2}\right)^2 + \sigma_{X_3}^2\left(\frac{\partial g}{\partial X_3}\right)^2 = \sigma_{X_1}^2 \times 1 + \sigma_{X_2}^2 \times 1 + \sigma_{X_3}^2 \times 1$$
$$\approx (500 \times 0.50)^2 + (600 \times 0.60)^2 + (700 \times 0.7)^2 = 432{,}200$$
$$\sigma_S \approx \sqrt{432{,}200} = 657$$

MC simulations can be performed using MATLAB with the following m-file to carry out the simulations:

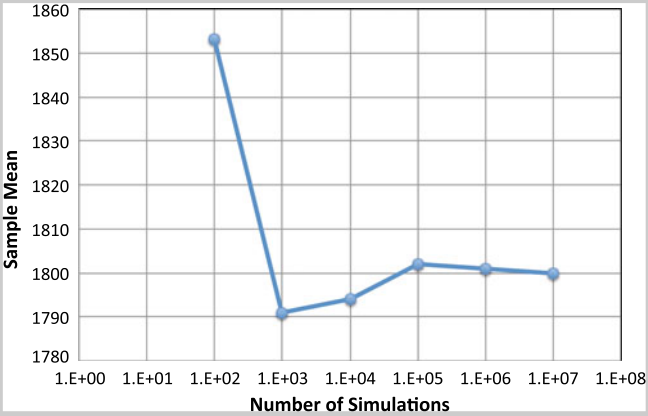
```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Using MC simulations on S=X1+X2+X3 where X's are LN

%first we need to calculate the moments of the
%lognormal distributions from their given means and
%coefficients of variations

xsi_1=sqrt(log(1+0.5^2));
lambda_1=log(500)-0.5*(xsi_1)^2;

xsi_2=sqrt(log(1+0.6^2));
lambda_2=log(600)-0.5*(xsi_2)^2;

xsi_3=sqrt(log(1+0.7^2));
lambda_3=log(700)-0.5*(xsi_3)^2;

%next we simulation the 3 lognormal random variables
n=10000; %number of simulations

x1=lognrnd(lambda_1,xsi_1,n,1);
x2=lognrnd(lambda_2,xsi_2,n,1);
x3=lognrnd(lambda_3,xsi_3,n,1);

%then we perform the calculation sequentially
s=x1+x2+x3;

%Results
hist(s,50) %the discrete results of n simulations
x_bar=mean(s) %sample mean value of n simulations
s=std(s)  %sample standard deviation of n simulations
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```




```
x_bar=1.7942e+03
s=660.2819
```

The histogram shows a skewed distribution but if we were to test it we would find that it does not exactly follow the lognormal distribution (It is beyond the scope of this text but the Gamma distribution is often useful for modeling a skewed distribution like the results here.).

Results will vary slightly for each run of the subroutine, even with the same number of simulations, because the realizations are randomly generated. But as we approach a large number of simulations the results will converge. This is a way of determining the accuracy of the simulations. The figure below shows the change in the sample mean as the number of simulations is increased.



We can see that in this problem the answer converges around 100,000 (1.E + 05) to 1,000,000 (1.E + 06) simulations. The computational cost of increased simulations is measured in runtime, where it took 0.050 s to generate 100,000 simulations compared to 0.284 s for 1,000,000 simulations. Both runs are nominal computationally, but if we increase the simulations in this problem up to 10,000,000 (1.E + 07) the run time is 2.370 s, a noticeable delay for the user. Of course the run time will increase with more complex problems/calculations, and more efficient programming can help decrease runtime.

If we compare the MC simulations with the FOSM approximations we find there is good agreement which provides confidence in our answer. This also demonstrates that for linear functions with input parameters that have relatively symmetric distributions, FOSM can often give accurate results.

FOSM	MC (1.E ± 07) simulations
$\mu = 1800$	`x_bar=1800`
$\sigma = 657$	`s=657`

It is common to use both an approximate and computational method to provide confidence in the answer and to tease out if there is any nonlinearity or other aspects contributing to the results.
Solving the same sum of lognormal problem with coding in R:

```
means <-c(500,600,700) #vector of means
coefv <- c(0.5,0.6,0.7) #vector of coefficents of variations
sd <- means*coefv #R defaults to elementwise, not matrix, operations
n <- 10000
xi_sq <- log(1+sd^2/means^2)
lambda <- log(means) - 1/2*xi_sq #lognormal parameters
xi <- sqrt(xi_sq)
s <- matrix(nrow = n, ncol = 3)
for (i in 1:3)
  s[,i] <- rlnorm(n, lambda[i], xi[i]) #generate random values of s
sum <- s[,1] + s[,2] + s[,3] #add values together
mean(sum)
sd(sum)
hist(sum,50)
```

This produces a similar answer as the code in MATLAB, and as the number of simulations (n) increases the answers converge.

Example: Product/Quotient of Normals

In this problem we will be evaluating the function:

$$P = \frac{X_1 X_3}{X_2}$$

The random variables in this function are all normally distributed with their respective moments, X_i are $N(\mu_i, \sigma_i)$. Note, if these were all lognormally distributed then we could solve using an exact solution, the product/difference of lognormals. Since they are normally distributed we must resort to an approximate solution and/or MC simulations. The information given for the X_i are:

	X_1	X_2	X_3
μ	500	600	700
σ	75	120	210

No correlation coefficient was given so, as before, we will assume that the variables are statistically independent.

The FOSM approximate of this is:

$$\mu_P \approx \frac{\mu_{X_1}\mu_{X_3}}{\mu_{X_2}} = 583$$

$$\sigma_P^2 \approx \sigma_{X_1}^2\left(\frac{X_3}{X_2}\right)^2 + \sigma_{X_2}^2\left(-\frac{X_1 X_3}{X_2^2}\right)^2 + \sigma_{X_3}^2\left(\frac{X_1}{X_2}\right)^2$$

$$\approx 75^2\left(\frac{700}{600}\right)^2 + 120^2\left(-\frac{500\times 700}{600^2}\right)^2 + 210^2\left(\frac{500}{600}\right)^2 = 51{,}892.36$$

$$\sigma_P \approx 227.8$$

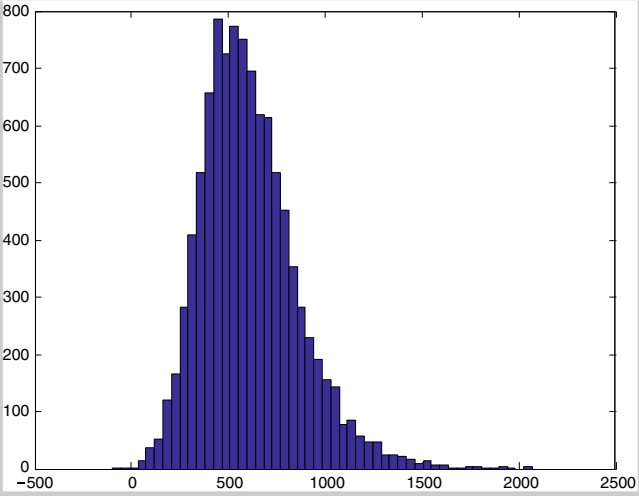
If we now wanted to know the probability that P would exceed 700, then we could approximate the probability using the standard normal distributions.

$$P(P \geq 700) = 1 - P(P \leq 700) \approx 1 - \Phi\left(\frac{700-500}{227.8}\right) = 1 - 0.8105$$

$$\approx \mathbf{0.1894} \text{ or roughly } \mathbf{19\%}$$

Solving the problem using MC simulations in MATLAB the m-file looks like:

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Using MC simulations to solve P=(X1*X3)/X2 where X's
%are normal

%first we declare the mean's and standard deviations
mu_1=500;
sigma_1=75;

mu_2=600;
sigma_2=120;

mu_3=700;
sigma_3=210;

%next we simulate the normal random variables
n=10000;

x1=normrnd(mu_1,sigma_1,n,1);
x2=normrnd(mu_2,sigma_2,n,1);
x3=normrnd(mu_3,sigma_3,n,1);

%then we perform the calculation sequentially.
%note: the dots in the equation specifies that linear,
%not matrix, mathematics be carried out
p=(x1.*x3)./x2;

%Results
hist(p,50)
x_bar=mean(p)
s=std(p)
prob=(sum(p>700))/n

%note: because these are simulations we can sum all the
%realizations that exceed the threshold of 700 and
%divide by the total number of simulations n to
%estimate the probability
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```



Results for (1.E + 07) simulations

```
x_bar=610.2443
s=261.0716
prob=0.3060
```

In this problem it is interesting to observe that the resultant distribution is neither normal nor lognormal but something in between. Now comparing the results of FOSM and MC simulations we find that in this case, with a nonlinear equation, FOSM underestimates the mean, standard deviation, and probability of exceedance. In this particular problem the MC simulations more accurately model the interaction of the input parameter distributions and the functional form of the equation.

FOSM	MC (1.E + 07) simulations
$\mu = 583$	`x_bar=610.2`
$\sigma = 227.8$	`s=261.1`
$P(P \geq 700) \approx 19\%$	`prob=0.3060`

The code for solving this same problem in R can be written as follows:

```
means <-c(500,600,700) #vector of means
sds <- c(75,120,210) #vector of standard deviations
n <- 10000 #number of simulations
x <- matrix(nrow = n, ncol = 3)
for (i in 1:3)
  x[,i] <- rnorm(n, means[i], sds[i]) #generate random values of x
p <- x[,1]*x[,3]/x[,2] #elementwise addition
hist(p,50)
mean(p)
sd(p)
sum(p>700)/n
```

As has been demonstrated, MC simulations for statistically independent random variables are relatively straightforward. When random variables are correlated we need to generate joint distributions to account for that correlation.

The **joint distribution** of two variables can be written with the multiplication rule, here using CDF's:

$$F(xy) = F(y|x)F(x) \tag{5.30}$$

If we have the marginal distribution of X and the conditional distribution of Y given X, then we can solve for the joint distribution. Computationally the joint distribution can be found using the algorithm presented in Ang and Tang (1984) for two random variables:

(1) Generate two uniformly distributed vectors, u_1 and u_2, between 0 and 1.
(2) Random variable X is generated as the inverse CDF of the first uniformly distributed random vector: $X = F^{-1}(u_1)$.
(3) Random variable Y is generated as the inverse CDF of the second uniformly distributed random vector conditioned on X: $Y = F^{-1}(u_2|x)$.

This algorithm applies to any joint distribution and many inverse CDF's are pre-programmed into computational software (e.g., normal, lognormal, Gamma, Beta). For joint distributions that are not pre-programmed the inverse CDF must be solved symbolically. This can be accomplished using MATLAB or R but lies outside the scope of this text and readers are referred to Ang and Tang (2007) for more details. For **joint normal** the distributions of Gaussian X and $Y|X$ are:

$$X = \Phi^{-1}(u_1)\sigma_X + \mu_X \tag{5.31}$$

$$Y = \Phi^{-1}(u_2)\sigma_Y\sqrt{1-\rho^2} + \left(\mu_Y + \rho\left(\frac{\sigma_Y}{\sigma_X}\right)(x-\mu_X)\right) \tag{5.32}$$

The following example illustrates this calculation in MATLAB and R.

Example: Joint Normals

In this problem we are evaluating the function:

$$\text{SQ} = X - Y$$

where X and Y are assumed jointly normal and negatively correlated. We can solve this using an exact solution and then verify using MC simulations. The statistics provided for the input parameters are:

	X	Y
$\bar{x}$	500	600
s	250	360
ρ	−0.33	

An exact solution produces:

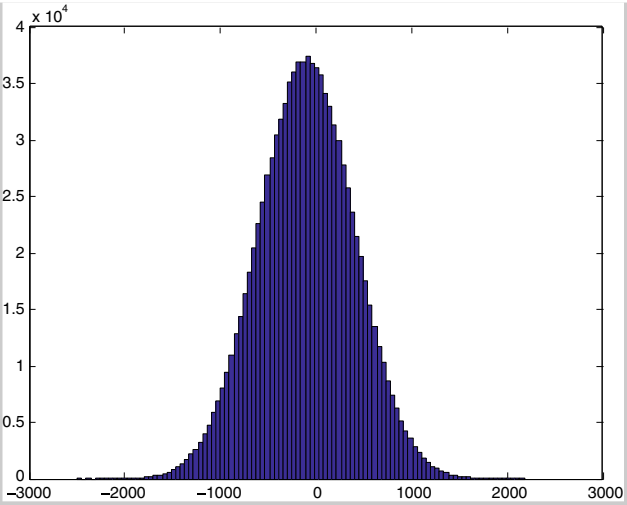
$$\begin{aligned}
\mu_{\text{SQ}} &= \mu_X + \mu_Y = 500 + (-600) = -100 \\
\sigma^2_{\text{SQ}} &= a_X^2\sigma_X^2 + a_Y^2\sigma_Y^2 + 2\rho a_X a_Y \sigma_X \sigma_Y \\
&= (1)^2 250^2 + (-1)^2 360^2 + 2(-0.33)1(-1)250(360) \\
&= 62{,}500 + 129{,}600 + 59{,}400 = 251{,}500 \\
\sigma_{\text{SQ}} &= \sqrt{251{,}500} = 501.5
\end{aligned}$$

If we are interested in the probability that SQ will exceed 500, following the exact solution we find:

$$\begin{aligned}
P(\text{SQ} > 500) &= 1 - P(\text{SQ} \le 500) = 1 - F(500) \\
&= 1 - \Phi\left(\frac{500 - (-100)}{501.5}\right) \approx 1 - 0.8842 = 0.1158
\end{aligned}$$

The solution using MC simulations uses the following MATLAB script:

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Using MC simulations to solve SQ=X-Y
%where X and Y are jointly normal and correlated

%first we declare the knowns
mu_x=500;
sigma_x=250;

mu_y=600;
sigma_y=360;

rho=-0.33;

%next we simulate the uniformly distributed vectors
n=1000000;
u1=unifrnd(0,1,n,1);
u2=unifrnd(0,1,n,1);


%then we generate the normal distribution of X
x=(norminv(u1).*sigma_x)+mu_x;

%and the conditional distribution of Y given X that
%includes correlation
y=(norminv(u2).*sigma_y.*sqrt(1-
   rho^2))+(mu_y+rho*(sigma_y/sigma_x).*(x-mu_x));

%now evaluating the function
sq=x-y;

%Results
hist(sq,100);
x_bar=mean(sq)
s=std(sq)
prob=(sum(sq>500)/n)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```




```
x_bar=−99.9253
s=501.6675
prob=0.1156
```

The histogram shows a normally distributed resultant as would be expected when we have the difference of normal distributions. Comparing the numeric results, we see that MC simulations provide a very close answer to the exact solution. This provides a check on our work and confidence in moving forward with the probability of exceedance for decision purposes.

Exact	MC simulations
$\mu = -100$	`x_bar=−99.9253`
$\sigma = 501.5$	`s=501.6675`
$P(\text{SQ} > 500) \approx 11.58\%$	`prob=0.1156`

The above code can be written using the multivariate function, which eliminates the need for explicitly generating the uniformly distributed random vectors.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Using MC simulations to solve SQ=X-Y where X and Y are
%jointly normal with prescribed correlation coefficient

%declare the means and standard deviations
mu_x=500;
sigma_x=250;

mu_y=600;
sigma_y=360;

rho=-0.33;

%setting up the means vector and covariance matrix
mu_vector=[mu_x mu_y];
%see Appendix C for discussion of the covariance matrix
cov_matrix=[sigma_x^2           rho*sigma_x*sigma_y ;
            rho*sigma_x*sigma_y           sigma_y^2];

%simulate the joint normal distribution
n=1000000;
xy=mvnrnd(mu_vector,cov_matrix,n);

%evaluate the function
sq=xy(:,1)-xy(:,2);

%Results
hist(sq,100);
x_bar=mean(sq)
s=std(sq)
prob=(sum(sq>500)/n)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

The results using the multivariate function (mvnrnd) provides nearly identical results as before.

Exact	MC simulations
$\mu = -100$	x_bar=−99.9746
$\sigma = 501.5$	stdev=502.0354
$P(\text{SQ} > 500) \approx 11.58\%$	prob=0.1159

Solving this problem in R could result in the following code:

```
library(MASS) #this must be installed to include multivariate normal function
#if not, install using install.packages(MASS)
means <-c(500,600) #vector of means
cov <- matrix(c(250^2, -0.33*250*360,-0.33*250*360,360^2),2,2) #covariance
matrix
n <- 10000 #number of simulations
u <- mvrnorm(n,means,cov) #simulate correlated variables
SQ <- u[,1]-u[,2]
hist(SQ,100)
mean(SQ)
sd(SQ)
sum(SQ>500)/n
```

For jointly lognormal parameters the algorithm using uniformly distributed random variables can be similarly used to solve problems. MATLAB does not currently support a multivariate lognormal function at the time of writing, however one is available in the user forum which performs the desired correlated lognormal simulation.[1]

In summary, the discussion of computational methods, by using Monte Carlo simulations we can generate sequential realizations of the random variables of interest and plug those realizations into the function we are evaluating. This provides a simulation as accurate as needed to solve any function of random variables. The normal and lognormal distributions are easy to simulate because these distributions are pre-programmed into computational software. Monte Carlo simulations provide a means of checking approximate solutions and can often be used to tackle more complex functions with difficult partial derivatives that can render FOSM intractable.

5.4 Chapter Summary

- Solving a **function of random variables** is the core material of this text. In almost all Civil Engineering problems we are using some mathematical equation made up of parameters that contain uncertainty. Translating or propagating this uncertainty from the input parameters to the resultant gives us a measure of how accurate the solution is.

[1] http://www.mathworks.com/matlabcentral/fileexchange/6426-multivariate-lognormal-simulation-with-correlation accessed 12/17/2018.

- **Exact solutions** exist for one-to-one single root functions, for a function that is the sum/difference of normally distributed random variables, and for a function that is the produce/difference of lognormally distributed random variables.
- For situations that don't fit the exact solutions we can use **approximation methods** and/or **Monte Carlo (MC) simulations**.
- Two approximate methods presented are the **central limit theorem** (CLT) and the **first order second moment** (FOSM) methods.
- CLT assumes the resultant is normal given a large number of input parameters in a sum/difference function.
- FOSM uses a **Taylor series** expansion about the mean to estimate the moments of a function given the moments of the input parameters. FOSM works on any mathematical function with input parameters having any distribution.
- MC simulations use many realizations to estimate the resulting distribution.
- MC simulations methods are often used in conjunction with other methods to validate the answer.

References

Ang, A. H.-S., & Tang, W. H. (1984). *Probability concepts in engineering planning and design: Volume II—Decision, risk, and reliability*. Published by the authors.

Ang, A. H.-S., & Tang, W. H. (2007). *Probability concepts in engineering: Emphasis on applications to civil and environmental engineering* (2nd ed.). New York: Wiley.

Knuth, D. (1997). Chapter 3—Random numbers. In *The art of computer programming. Vol. 2: Seminumerical algorithms* (3rd ed.).

Ripley, B. D. (1987). *Stochastic simulation*. New York: Wiley.

Chapter 6
Component Reliability Analysis

Reliability at the component level is the probabilistic relationship between load and resistance, or stress and strain, or demand and capacity; these paired terms are all analogous. Reliability is commonly expressed using the reliability index, β, which can in turn be related to the probability of failure, p_f.

Figure 6.1 shows a deterministic view of an engineering problem. Load (Q) and resistance (R) are shown on a number line with respect to each other. Whenever the load is less than the resistance then the design is considered safe. Using a factor of safety formulation where FS $= R/Q$, a "no failure" state is where FS > 1 when Q is less than R. Using a margin of safety formulation where $M = R - Q$, a "no failure" state is where M > 0 when Q is less than R.

When we include the uncertainty of the load and resistance in the analysis, as Fig. 6.2 shows there may be a region where failure can occur. Using factor of safety formulation, $P(\text{FS}) = P(R)/P(Q)$, or margin of safety formulation, $P(M) = P(R) - P(Q)$, and the same mean values as in the deterministic view, if there is sufficient uncertainty in the load and/or resistance then there is a probability of failure.

6.1 Component Reliability Formulation

The fundamental presentation of component reliability can be accomplished using the margin of safety formulation and assuming that the load and resistance are jointly normal.

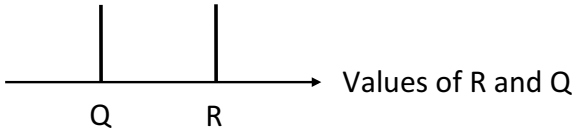


Fig. 6.1 Deterministic view of load (Q) versus resistance (R)

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_6

$$M = R - Q \quad \text{where } R \text{ and } Q \text{ are jointly normal} \tag{6.1}$$

$$\mu_M = \mu_R - \mu_Q \tag{6.2}$$

$$\sigma_M^2 = \sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q \tag{6.3}$$

$$\beta = \frac{\mu_m}{\sigma_m} \tag{6.4}$$

The reliability index is the number of standard deviations the mean is away from failure. For correlated R and Q, substituting from above:

$$\beta = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q}} \tag{6.5}$$

For statistically independent R and Q:

$$\beta = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2}} \tag{6.6}$$

The reliability index, β, is the distance between the mean and the failure point $M = 0$ in units of standard deviation as shown in Fig. 6.3. It is a measure of how far away the most likely value is from failure.

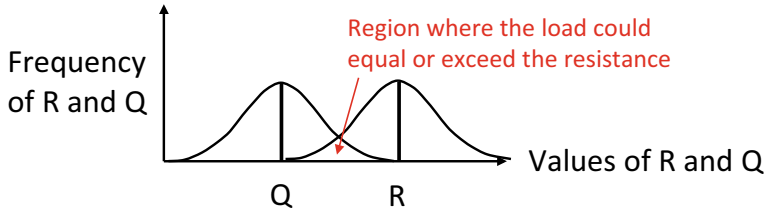


Fig. 6.2 Probabilistic view of load (Q) versus resistance (R)

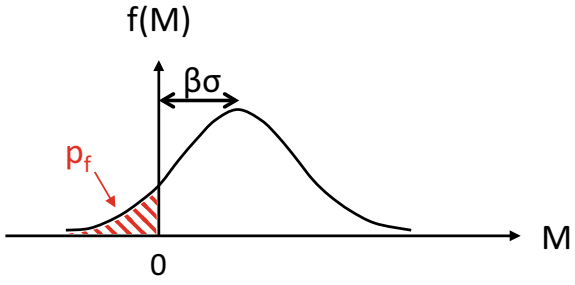


Fig. 6.3 Shown is the distribution of M with the probability of failure (p_f) as the region where $M \leq 0$. The distance between the failure point, $M = 0$, and the mean value (μ) is the reliability index (β) in units of standard deviation (σ)

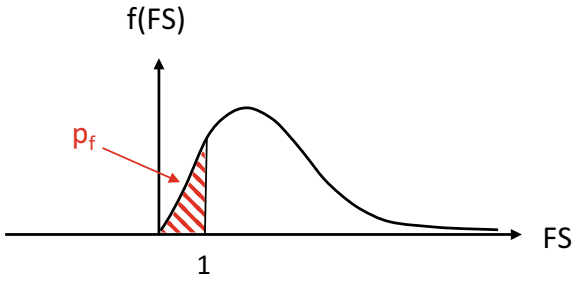


Fig. 6.4 Reliability using the factor of safety formulation

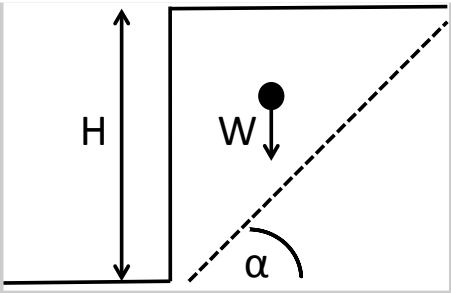
The area under the probability distribution $f(M)$ where $M \leq 0$ is the probability of failure; therefore, we are interested in the CDF to define the failure $\left[F(m) = \int_{-\infty}^{0} f(m)\,\mathrm{d}m\right]$. Since in this derivation Q and R are normally distributed the margin of safety function is a sum of these variables, then M is normally distributed and we can use the standard normal distribution to solve this integral.

$$p_\mathrm{f} = F(M) = \Phi\left(\frac{M-\mu_m}{\sigma_m}\right) = \Phi\left(\frac{0-\mu_m}{\sigma_m}\right)$$
$$= \Phi\left(\frac{-\mu_m}{\sigma_m}\right) = \Phi(-\beta) = 1 - \Phi(\beta) \tag{6.7}$$

By assuming that the load and resistance are normal and that failure is defined using the margin of safety formulation we arrive at an exact solution for the reliability index and the probability of failure. This is the most common and intuitive presentation of the component reliability calculation. Note however if load and resistance are not normal then this solution is not exact. Other solution methods exist when the assumption of normality does not hold true.

Example: Exact Solution for Normals

To demonstrate reliability using an exact solution we will look at a slope stability problem borrowed from Baecher and Christian (2003). This is called the cut slope problem because we are interested in the stability of a vertical cut in cohesive soil. Figure 6.4 shows the geometry. If we assume that the soil is purely cohesive and that the slope will fail along a 45° failure plane (α) then we can set up the load and resistance functions for solving the stability of this slope. For this example the height will be a deterministic value fixed at 10 m.



Cut slope geometry shows H is the vertical height of the cut slope, α is the angle of the failure plane, and W is the weight vector at the center of mass of the potential failure wedge.
$R =$ c cohesion of the soil
$Q = 0.25\gamma H$ from $0.5\gamma H \sin\alpha \cos\alpha$ the solution for a 45° slope.

If c and γ are normally distributed then R and Q are normally distributed because both are one-to-one functions of their independent variables. The first and second moments and correlation coefficient for this problem are given as:

$$\mu_c = 100 \text{ kPa} \quad \sigma_c = 30 \text{ kPa}$$
$$\mu_\gamma = 20 \text{ kN/m}^3 \quad \sigma_\gamma = 2 \text{ kN/m}^3$$
$$\rho_{c\gamma} = 0.5$$

Solving for R and Q which are functions of the random variables c and γ:

$$\mu_R = 100 \text{ kPa}$$
$$\sigma_R = 30 \text{ kPa}$$
$$\mu_Q = 0.25\left(20\frac{\text{kN}}{\text{m}^3}\right)10 \text{ m} = 50 \text{ kPa}$$
$$\sigma_Q = \sqrt{(0.25 \times 10 \text{ m})^2\left(2\frac{\text{kN}}{\text{m}^3}\right)^2} = 5 \text{ kPa}$$

Because these are one-to-one functions relating the dependent and independent variables then R and Q have the same correlation coefficient as c and γ. Using the margin of safety formulation and having normally distributed variables lends to the exact solution:

$$M = R - Q$$

$$\mu_M = \mu_R - \mu_Q = 100 - 50 = 50 \text{ kPa}$$

$$\sigma_M^2 = \sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q = 30^2 + 5^2 - 2(0.5)(30)(5)$$
$$= 900 + 25 - 150 = 775 \text{ kPa}$$

$$\beta = \frac{\mu_m}{\sigma_m} = \frac{50}{\sqrt{775}} = 1.80$$

(i.e., the mean is 1.8 standard deviations from failure)

$$p_f = 1 - \Phi(\beta) = 1 - \Phi(1.80) \simeq 1 - 0.96407 = 0.03593$$

So given:

- The distributions and statistics of the independent variables c and γ,
- The equations relating the independent variables to the dependent variables R and Q, and
- The margin of safety formulation defining failure,

then the probability of failure for this problem is 3.6%. The correlation between the cohesion and the density is physically obvious but let us assume that they are not correlated (i.e., statistically independent where $\rho = 0$) to see how it influences the probability of failure.

$$\mu_M = \mu_R - \mu_Q = 100 - 50 = 50 \text{ kPa}$$

$$\sigma_M^2 = \sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q = 30^2 + 5^2 - 0$$
$$= 900 + 25 - 0 = 925 \text{ kPa}$$

$$\beta = \frac{\mu_m}{\sigma_m} = \frac{50}{\sqrt{925}} = 1.64$$

(i.e., the mean is 1.64 standard deviations from failure)

$$p_f = 1 - \Phi(\beta) = 1 - \Phi(1.64) \simeq 1 - 0.949497 = 0.050503$$

The reliability index is lower and the probability of failure is higher for the uncorrelated situation. Why is this? In a correlated situation the uncertainties in the load and resistance are joint or interrelated, whereas the uncertainties in the uncorrelated situation are independent resulting in larger uncertainty

when propagated through the margin of safety equation and therefore a smaller reliability index.

The reliability index decreases, and probability of failure increases, as the mean moves closer to the failure point. The mantra "Bigger Beta is Better" is a useful one to remember in engineering. It is common to design engineered features to have a reliability index of 2–3 that is a design with the mean 2–3 standard deviations away from the failure point. This concept will be discussed in more detail in the chapter on reliability-based codes (Chap. 9).

6.2 Lognormal Parameters

An alternate exact solution exists if load and resistance are assumed lognormal (i.e., the natural log of load and resistance is normal) and the factor of safety formulation is used to define failure (Fig. 6.5).

This formulation is used often for situations where the load and resistance are both nonnegative, and historically was developed for steel construction.

$$\text{FS} = R/Q \quad \text{where } R \text{ and } Q \text{ are jointly normal} \tag{6.8}$$

By taking the natural log of both sides we find:

$$\ln(\text{FS}) = \ln(R) - \ln(Q) \tag{6.9}$$

$$\lambda_{\text{FS}} = \lambda_R - \lambda_Q \tag{6.10}$$

$$\xi_{\text{FS}}^2 = \xi_R^2 + \xi_Q^2 - 2\rho_{\ln R \ln Q}\xi_R\xi_Q \tag{6.11}$$

$$\beta = \frac{\lambda_{\text{FS}}}{\xi_{\text{FS}}} \tag{6.12}$$

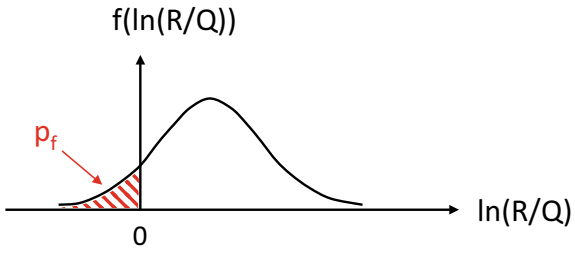


Fig. 6.5 For lognormal load and resistance then the natural log of the FS = R/Q is normally distributed

$$p_f = F(1) = \Phi\left(\frac{\ln(1) - \lambda_{FS}}{\xi_{FS}}\right) = \Phi\left(\frac{0 - \lambda_{FS}}{\xi_{FS}}\right)$$
$$= \Phi\left(\frac{-\lambda_{FS}}{\xi_{FS}}\right) = \Phi(-\beta) = 1 - \Phi(\beta) \tag{6.13}$$

The reliability index can also be written as:

$$\beta = \frac{\mu_{\ln R} - \mu_{\ln Q}}{\sqrt{\sigma^2_{\ln R} + \sigma^2_{\ln Q} - 2\rho_{\ln R \ln Q}\sigma_{\ln R}\sigma_{\ln Q}}}$$
$$= \frac{\ln\left(\frac{\mu_R}{\mu_Q}\sqrt{\frac{1+\delta_Q^2}{1+\delta_R^2}}\right)}{\sqrt{\ln\left(1+\delta_R^2\right)\ln\left(1+\delta_Q^2\right) - 2\ln\left(1+\rho_{RQ}\delta_R\delta_Q\right)}} \tag{6.14}$$

where the mean and standard deviation of the natural log are shown below for R, and follow the same pattern for Q.

$$\mu_{\ln R} = \lambda_R = \ln \mu_R - \frac{1}{2}\xi_R^2 \tag{6.15}$$

$$\sigma^2_{\ln R} = \xi_R^2 = \ln\left(1 + \frac{\sigma_R^2}{\mu_R^2}\right) = \ln\left(1 + \delta_R^2\right) \tag{6.16}$$

An approximation of the reliability index that is commonly used for the lognormal case with given mean and coefficient of variations (Rosenblueth and Esteva 1972) can be written as:

$$\beta \simeq \frac{\ln\left(\frac{\mu_R}{\mu_Q}\right)}{\sqrt{\delta_R^2 + \delta_Q^2 - 2\rho_{RQ}\delta_R\delta_Q}} \tag{6.17}$$

6.3 General Reliability Procedure

Most reliability problems require more sophisticated solutions because the distributions of load and resistance cannot be assumed to be normal (or lognormal). The various solutions however all follow similar steps:

1. Determine the equations, formulas, models that will be used to calculate R and Q. These can be empirical, theoretical, or approximate.
2. Calculate the first and second moments of R and Q. Mean and coefficient of variation are often sufficient, but the full distributions can be used if available.

3. In most cases the margin of safety formulation is used, $M = R - Q$ so the first and second moments of M are calculated. Here the uncertainty from R and Q are propagated to M.
4. Calculate the reliability index, β.
5. Calculate the probability of failure, p_f.

The following is a list of common reliability solution techniques used in Civil Engineering. The asterisk (*) indicates a method that is covered in this text.

- Exact solutions* as described above. These solutions are limited by the assumption of normal or lognormal R and Q and the corresponding failure formulation.
- First order second moment (FOSM)* applies to any distribution of R and Q, but is only approximate.
- Second order second moment (SOSM) has increased accuracy over FOSM, but still an approximate.
- Point-Estimate is an interesting technique similar to Gaussian quadrature integration, but can be rather cumbersome in this age of fast computing (see Rosenbluth 1975; Baecher and Christian 2003).
- First Order Reliability Method (FORM)* is the "standard" of reliability analysis and often considered requisite when doing probability of failure calculations.
- Second Order Reliability Method (SORM)* is particularly useful as the failure surface becomes more nonlinear.
- Monte Carlo simulations (MC)* is the "brute force" approach that provides a robust approximate, often used to confirm results found using other methods.

6.4 FOSM

First order second moment (FOSM) is the same error propagation technique presented for functions of random variables in Chap. 5. Here the function of interest is the margin of safety formulation (or the factor of safety formulation) and the random variables are R and Q. This solution works in any situation, but provides an approximate solution. As the functions of R and Q become more nonlinear the FOSM results can diverge from the true results. The following example uses the same cut slope problem but avoids the assumption of normally distributed R and Q.

Example: FOSM Solution

$$M = R - Q$$

where $R = c$ and $Q = 0.25\gamma H$ and $H = 10$ m therefore

$$M = c - 2.5\gamma$$

$$\mu_M = \mu_c - 2.5\mu_\gamma = 100 - 2.5(20) = 50$$

$$\sigma_M^2 = \sigma_c^2\left(\frac{\partial M}{\partial c}\right)^2 + \sigma_\gamma^2\left(\frac{\partial M}{\partial \gamma}\right)^2 + 2\rho\sigma_c\sigma_\gamma\frac{\partial M}{\partial c}\frac{\partial M}{\partial \gamma}$$

$$\frac{\mathrm{d}M}{\mathrm{d}c} = 1 \quad \frac{\partial M}{\partial \gamma} = -2.5$$

$$\begin{aligned}\sigma_M^2 &= 30^2(1)^2 + 2^2(-2.5)^2 + 2(0.5)30(2)1(-2.5)\\ &= 900 + 25 - 150 = 775\end{aligned}$$

$$\beta = \frac{\mu_m}{\sigma_m} = \frac{50}{\sqrt{775}} = 1.80$$

$$p_\mathrm{f} = 1 - \Phi(\beta) = 1 - \Phi(1.80) \simeq 1 - 0.96407 = 0.03593$$

For uncorrelated load and resistance:

$$\begin{aligned}\sigma_M^2 &= 30^2(1)^2 + 2^2(-2.5)^2 + 2(0)30(2)1(-2.5)\\ &= 900 + 25 = 925\end{aligned}$$

$$\beta = \frac{\mu_m}{\sigma_m} = \frac{50}{\sqrt{925}} = 1.64$$

$$p_\mathrm{f} = 1 - \Phi(\beta) = 1 - \Phi(1.64) \simeq 1 - 0.949497 = 0.050503$$

Second order second moment (SOSM) takes the approximation further using the second order expansion of the Taylor series. This may be more accurate for problems where the second partial derivative of the R and Q have nonzero results.

It should be noted that FOSM (and SOSM by extension) involves some assumptions that can lead to inaccurate results. This was first shown by Hasofer and Lind (1974), and more recently demonstrated using the cut slope problem by Baecher and Christian (2003). The previous example shows the cut slope problem with uncorrelated load and resistance using the margin of safety formulation which results in a FOSM-based reliability index of 1.64. The same problem using a factor of safety formulation will give a FOSM-based reliability index of 1.58, even though $M = 0$ and $\mathrm{FS} = 1$ are mathematically identical. This variability of FOSM-based results as a function of the problem formulation leads researchers to investigate other invariant solutions to the reliability problem.

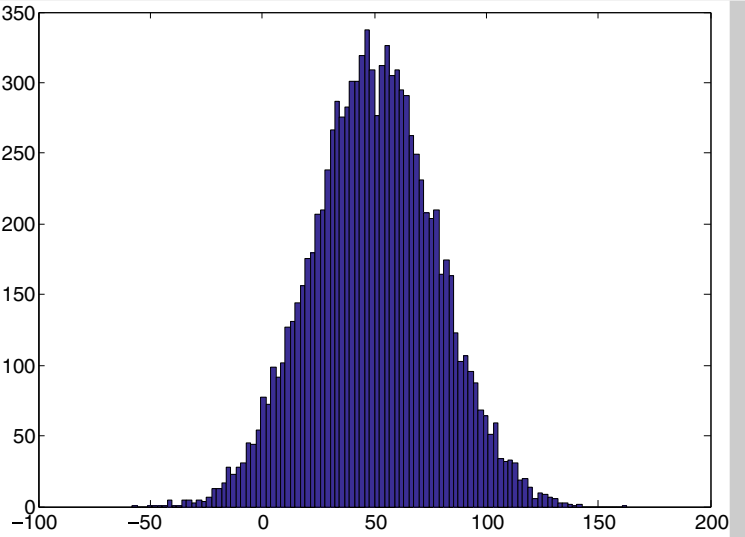
6.5 Monte Carlo Simulations

As with functions of random variables we can solve the margin of safety formulation (or factor of safety formulation) using Monte Carlo simulations. The MATLAB script for the cut slope problem is shown below.

Example: Monte Carlo Approximation

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%MC of vertical cut problem (correlated)

%uniformly distributed random numbers from 0 to 1
u1=unifrnd(0,1,10000,1);
u2=unifrnd(0,1,10000,1);

%simulations of c
c=norminv(u1).*30+100;
%simulations of Gamma given c
Gamma=(norminv(u2)).*(2).*(sqrt(10.50.^2))+((0.50).*(2/
30).*(c-100))+(20);
%margin of safety
M=c-(Gamma.*(10/4));

%results
hist(M,100);
x_bar=mean(M)
s=std(M)
Beta=x_bar/s
pf=normcdf(-Beta,0,1)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

The results are shown below for the correlated example.

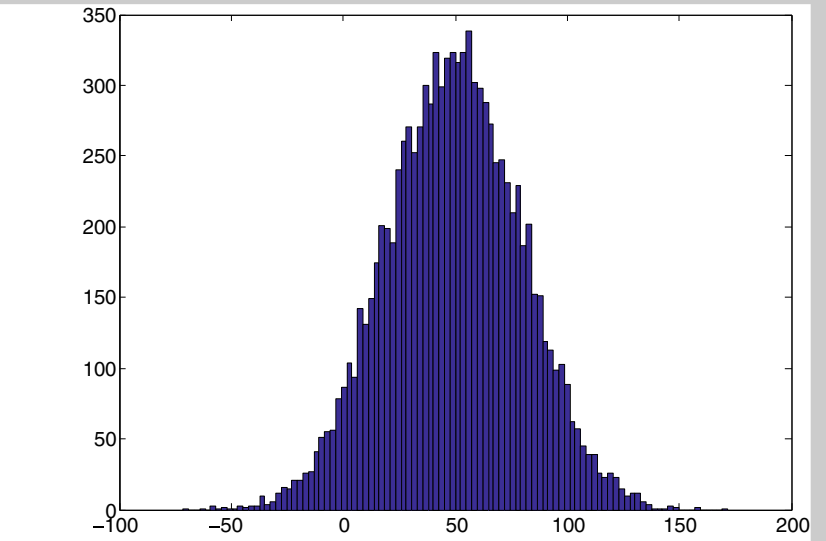



```
x_bar = 49.9724
s = 27.7358
Beta = 1.8017
pf = 0.0358
```

To run the simulations for the uncorrelated example can set the correlation coefficient to zero in the previous script or rewrite the script in a much simpler manner as shown below.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% MC of vertical cut problem (uncorrelated normals)

c=normrnd(100,30,10000,1);
Gamma=normrnd(20,2,10000,1);

M=c-(Gamma.*(10/4));

hist(M,100);
x_bar=mean(M)
s=std(M)
Beta=x_bar/s
pf=normcdf(-Beta,0,1)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

And the results for the uncorrelated load and resistance are:




```
x_bar = 49.4460
s = 30.1155
Beta = 1.6419
pf = 0.0503
```

Note that the results vary depending on how many simulations we run. In this example 10,000 simulations were run. As discussed previously we can run the number of simulations high enough to minimize the error below a certain acceptable tolerance. With Monte Carlo simulations we are estimating the results; however given enough simulations the results become fairly accurate.

The vertical cut problem solved by coding in *R* looks like:

```
#vertical cut mc solution
rm(list = ls()) #clear workspace
#requires MASS library for multivariate normals
library(MASS)
m_gamma <- 20
s_gamma <- 2
m_c <- 100
s_c <- 30
rho <- 0.5
n <- 10000
#uncorrelated normals
c <- rnorm(n,m_c,s_c)
gamma <- rnorm(n,m_gamma,s_gamma)
#calculate saftey margin
M <- c-(gamma*(10/4))
hist(M,100, main = "uncorrelated normals")
mean <- as.character(mean(M))
std <- as.character(sd(M))
beta <- mean(M)/sd(M)
prob <- as.character(pnorm(-beta,0,1))
writeLines(c("mean = ", mean , "std = ", std,"beta = ", as.character(beta),
"prob = ", prob))
#correlated normals
cov <- rbind(c(s_gamma^2,rho*s_gamma*s_c),c(rho*s_gamma*s_c, s_c^2))
means <- c(m_gamma, m_c)
m <- mvrnorm(n,means,cov)
M <- m[,2]-(m[,1]*(10/4))
hist(M,100, main = "correlated normals")
mean <- as.character(mean(M))
std <- as.character(sd(M))
beta <- mean(M)/sd(M)
prob <- as.character(pnorm(-beta,0,1))
writeLines(c("mean = ", mean , "std = ", std,"beta = ", as.character(beta),
"prob = ", prob))
```

6.6 FORM/SORM

A more robust means of determining the probability of failure is using the geometric approach of first order reliability method (FORM) and/or second order reliability method (SORM). As discussed FOSM can give different results dependent on if the problem is formulated using margin of safety or factor of safety. To avoid this the

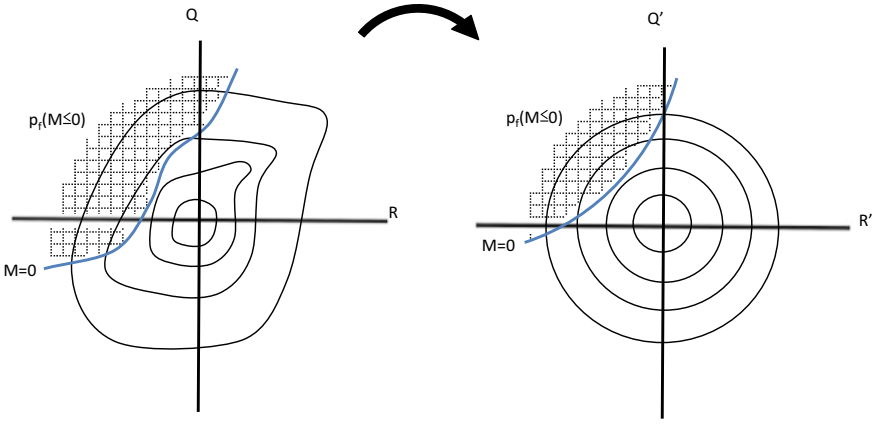
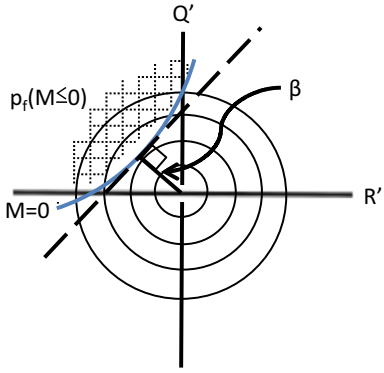


Fig. 6.6 Transforming to standard normal space, from (R, Q) to (R', Q')

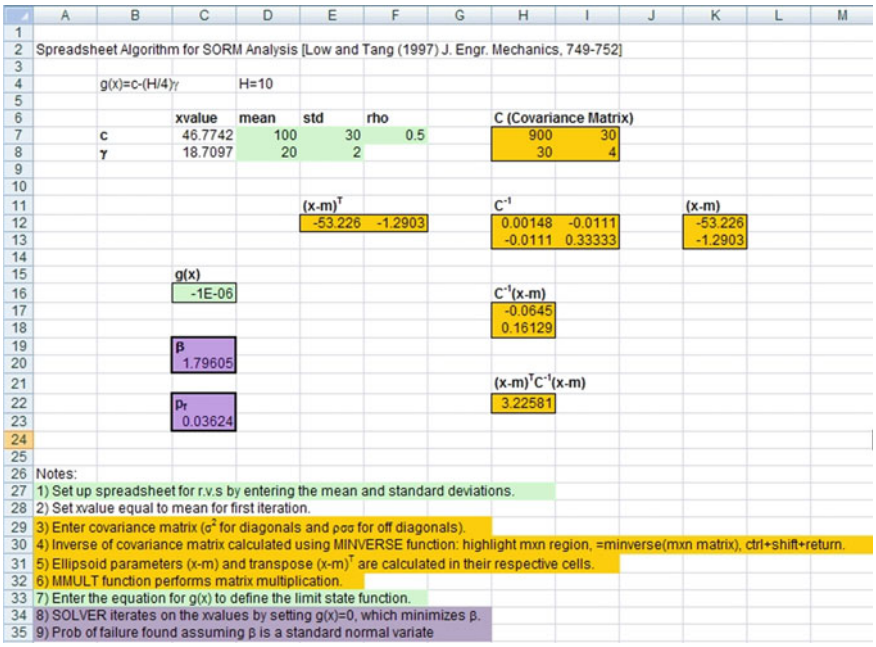
Fig. 6.7 The reliability index, β, is the minimum perpendicular distance from the origin to a straight line tangent at $M = 0$



Hasofer–Lind approach (Hasofer and Lind 1974) to FORM translates the problem into standard normal space and then solves for the distance from the mean to the point of failure, which is invariant regardless of how failure is formulated.

The Hasofer–Lind approach typically follows:

1. Formulate the problem as margin of safety, $M = R - Q$.
2. Transform from (R, Q) space into standard normal space (R', Q') as shown conceptually in Fig. 6.6 (This transformation is a function of the joint distribution of R and Q).
3. Find minimum distance from $M = 0$ to origin using a straight line fit at the tangent point (see Fig. 6.7). This requires an iterative solution which can be accomplished using different approaches.
4. The reliability index is equal to the minimum distance to the tangent, $\beta = \min(\text{distance})$, and the probability of failure is the standard normal distribution of the negative reliability index, $p_f = \Phi(-\beta)$.



Spreadsheet Algorithm for SORM Analysis [Low and Tang (1997) J. Engr. Mechanics, 749-752]

g(x)=c-(H/4)γ H=10

	xvalue	mean	std	rho
c	46.7742	100	30	0.5
γ	18.7097	20	2	

C (Covariance Matrix)

900	30
30	4

$(x-m)^T$: -53.226 -1.2903

C^{-1}

0.00148	-0.0111
-0.0111	0.33333

(x-m)

-53.226
-1.2903

g(x): -1E-06

$C^{-1}(x-m)$

-0.0645
0.16129

β: 1.79605

$(x-m)^T C^{-1}(x-m)$: 3.22581

p_f: 0.03624

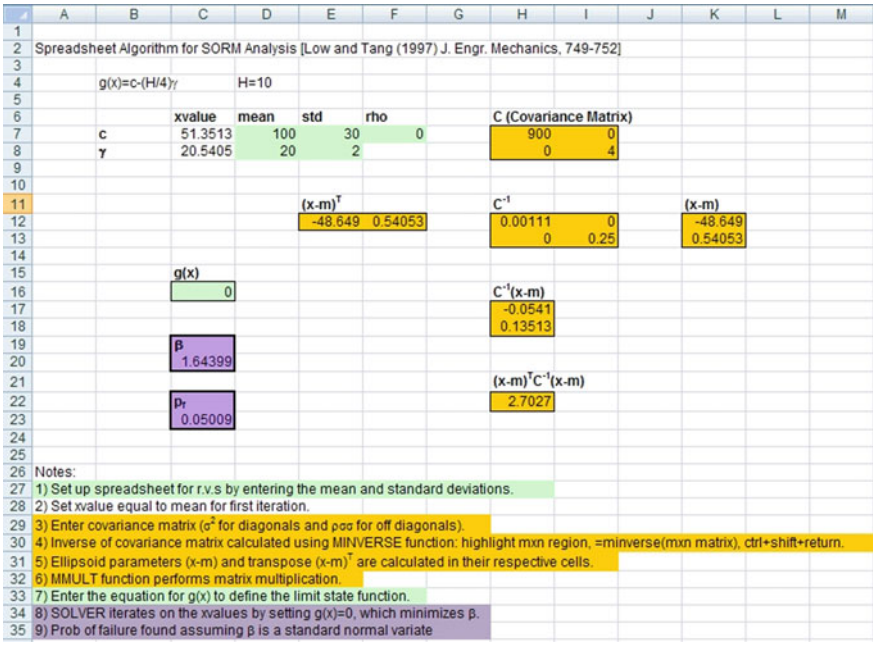
Notes:
1) Set up spreadsheet for r.v.s by entering the mean and standard deviations.
2) Set xvalue equal to mean for first iteration.
3) Enter covariance matrix (σ^2 for diagonals and $\rho\sigma\sigma$ for off diagonals).
4) Inverse of covariance matrix calculated using MINVERSE function: highlight mxn region, =minverse(mxn matrix), ctrl+shift+return.
5) Ellipsoid parameters (x-m) and transpose $(x-m)^T$ are calculated in their respective cells.
6) MMULT function performs matrix multiplication.
7) Enter the equation for g(x) to define the limit state function.
8) SOLVER iterates on the xvalues by setting g(x)=0, which minimizes β.
9) Prob of failure found assuming β is a standard normal variate

Fig. 6.8 Cut slope problem with correlated load and resistance

SORM can provide a more accurate solution as it fits a curve instead of a straight line to $M = 0$ when solving the minimum distance to the tangent (Der Kiureghian et al. 1987). Both FORM and SORM require an iterative solution to determine the minimum distance to the failure surface and the most efficient solutions use matrix manipulations (e.g., Cholesky decomposition, Jacobian matrix, etc.). In Appendix C the "improved" HLRF algorithm (Zhang and Der Kiureghian, 1995) is presented for solving the cut slope problem using FORM.

To simplify the application of FORM/SORM and encourage practicing engineers to use these methods, they have been written into Excel to take advantage of the built-in matrix manipulation and iteration solver functions available in the spreadsheet program. Low and Tang (1997) programed SORM into Excel utilizing the properties of an ellipse and built-in matrix manipulation functions. These spreadsheet reliability methods provide invariant reliability solutions that are readily applicable to Civil Engineering problems.

Figures 6.8 and 6.9 show the cut slope problem solved using the Low and Tang (1997) spreadsheet solution. Notes at the bottom of the spreadsheet describe the solution method. Figure 6.10 shows the equations in each cell and how the Solver is used to iteratively calculate the minimum reliability index. The ease of this solution method renders reliability accessible to any practicing engineer with the interest in performing reliability.



Spreadsheet Algorithm for SORM Analysis [Low and Tang (1997) J. Engr. Mechanics, 749-752]

g(x)=c-(H/4)γ H=10

	xvalue	mean	std	rho
c	51.3513	100	30	0
γ	20.5405	20	2	

C (Covariance Matrix)

900	0
0	4

$(x-m)^T$: -48.649 0.54053

C^{-1}

0.00111	0
0	0.25

(x-m)

-48.649
0.54053

g(x): 0

$C^{-1}(x-m)$

-0.0541
0.13513

β: 1.64399

$(x-m)^T C^{-1}(x-m)$: 2.7027

p_f: 0.05009

Notes:
1) Set up spreadsheet for r.v.s by entering the mean and standard deviations.
2) Set xvalue equal to mean for first iteration.
3) Enter covariance matrix (σ^2 for diagonals and $\rho\sigma\sigma$ for off diagonals).
4) Inverse of covariance matrix calculated using MINVERSE function: highlight mxn region, =minverse(mxn matrix), ctrl+shift+return.
5) Ellipsoid parameters (x-m) and transpose $(x-m)^T$ are calculated in their respective cells.
6) MMULT function performs matrix multiplication.
7) Enter the equation for g(x) to define the limit state function.
8) SOLVER iterates on the xvalues by setting g(x)=0, which minimizes β.
9) Prob of failure found assuming β is a standard normal variate

Fig. 6.9 Cut slope problem with uncorrelated load and resistance

Other spreadsheet solutions exist and subsequent work has expanded these solutions to accommodate many probability distributions and the transformation of those distributions into standard normal space. For further reading the following references are recommended; Low and Tang (2004), Phoon and Nadim (2004), and Low (2005).

6.7 Limit State

Up to this point we have been discussing failure and defining failure primarily using the margin of safety formulation. However we may be interested not in outright failure but some unsatisfactory performance that we would not necessarily label as failure in the breaking, fracturing, or collapsing sense. In defining a performance criterion in this manner it is then often called a limit state. The same mathematics and solution techniques can be used to solve any limit state. We generalize the margin of safety formulation to now encompass a threshold beyond which we can have unsatisfactory performance. The limit state function is usually denoted by g and the independent variables X where $g \leq 0$ is unsatisfactory performance:

$$g = X_1 - X_2 \tag{6.18}$$

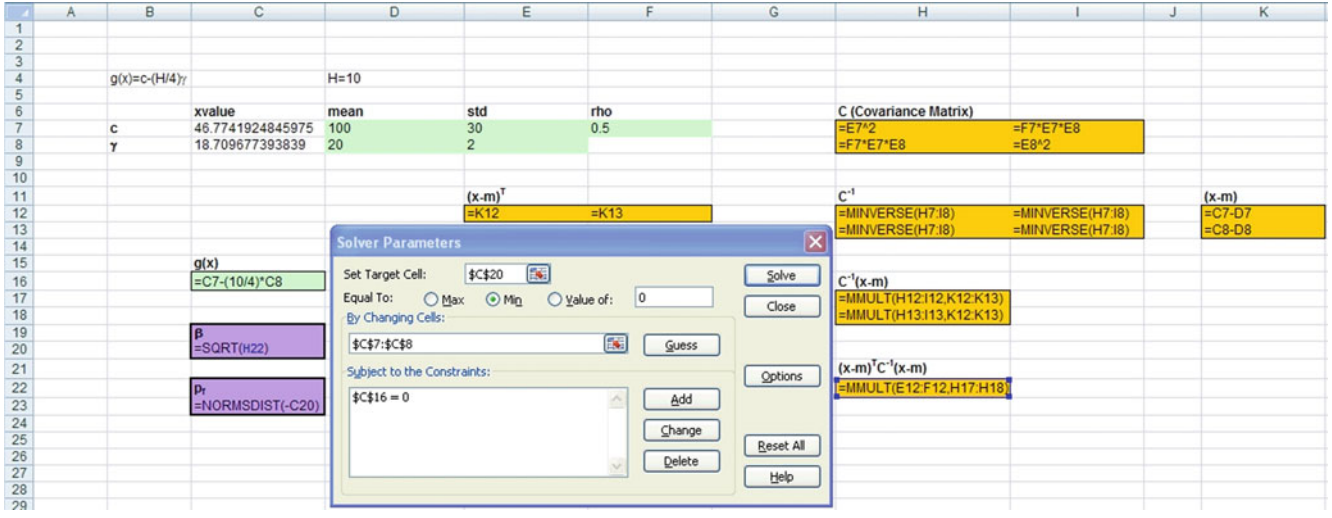


Fig. 6.10 Equations and solver for iteratively calculating the minimum reliability index using SORM spreadsheet solution

6.8 Chapter Summary

- Reliability is solving a specific function of random variables where the variables are **load** and **resistance** and the function is commonly the difference of the two (i.e., margin of safety formulation).
- Reliability is described by the **reliability index**, the number of standard deviations away from failure.
- The reliability index can be related mathematically to the **probability of failure**.
- The solution techniques covered in the previous chapter (Chap. 5 Functions of Random Variables) can be applied in solving reliability problems; **exact solutions** for normal and lognormal, **approximate solutions** using a Taylor series expansion, and **Monte Carlo** simulations.
- Additionally, geometric solution techniques specific to reliability, **FORM** and **SORM**, are used because they are invariant with respect to the function that relates load and resistance.
- A simplified SORM spreadsheet solution is presented to make reliability accessible without the need for a background in matrix mathematics.
- Reliability is generalized to the **limit state** formulation for solving problems involving any unsatisfactory engineering performance, not just failure.
- In engineering the goal is to achieve a safe design that is also economical. The reliability index (β) is the measure of how many standard deviations away from failure the design is, therefore within the economic constraints of the project, **Bigger Beta is Better**.

References

Baecher, G. B., & Christian, J. T. (2003). *Reliability and statistics in geotechnical engineering*. New Jersey: Wiley.

Der Kiureghian, A., Lin, H.-Z., & Hwang, S.-J. (1987). Second-order reliability approximations. *Journal of Engineering Mechanics, ASCE, 113*(8), 1208–1225.

Hasofer, A. M., & Lind, N. (1974). An exact and invariant first-order reliability format. *Journal of Engineering Mechanics, ASCE, 100*(EM1), 111–121.

Low, B. K. (2005). Reliability-based design applied to retaining walls. *Geotechnique, 55*(1), 63–75.

Low, B. K., & Tang, W. H. (1997). Efficient reliability evaluation using spreasheet. *Journal of Engineering Mechanics, 123*(7), 749–752.

Low, B. K., & Tang, W. H. (2004). Reliability analysis using object-oriented constrained optimization. *Structural Safety, 26*(1), 69–89.

Phoon, K.-K, & Nadim, F. (2004, November). *Modeling non-Gaussian random vectors for FORM: State-of-the-art review* (pp. 55–85). Keynote Paper International Workshop on Risk Assessment in Site Characterization and Geotechnical Design, Bangalore, India.

Rosenbluth, E. (1975). Point estimates for probability moments. *Proceedings of National Academy of Science, 72*(10), 3812–3814.

Rosenbluth, E., & Esteva, L. (1972). *Reliability basis for some Mexican codes*. Detroit: American Concrete Institute.

Zhang, Y., & Der Kiureghian, A. (1995). Two improved algorithms for reliability analysis. In *Reliability of and optimization of structural systems. Proceedings 6th IFIP Working Conference on Optimization of Structural Systems* (pp. 297–204).

Chapter 7
System Reliability Analysis

Component reliability was covered in Chap. 6, but we may often have a system made of many components. To analyze a system we can generalize and expand the limit state by considering multiple components and multiple failures that define the system [we will be using the term failure as synonymous with any unacceptable performance as defined by the limit state].

$$g = X_1 - X_2$$

where $g \leq 0$ is failure or unsatisfactory performance

$$g_j(\boldsymbol{X}) = g(x_1, x_2, \ldots, x_n)$$

where $j = 1, 2, \ldots, k$ so there are k potential failures to evaluate.

If we restrict the system to just two independent variables and three limit states [$\boldsymbol{X} = (x_1, x_2)$ and $j = 1, 2, 3$ where $g_j(\boldsymbol{X}) = 0$] we can visualize the multiple limit states in standard normal space as shown in Fig. 7.1.

If the joint PDF of the independent variables is $f(x_1, \ldots, x_n)$ and any component failure results in system failure then the probability of failure for the system in general is the volume integral:

$$p_f = \int_{g(\boldsymbol{X}) \leq 0} \cdots \int f(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n \tag{7.1}$$

Calculation of this multifold integral can often be difficult and in most cases approximations or bounds are used to estimate the range of the probability of system failure.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_7

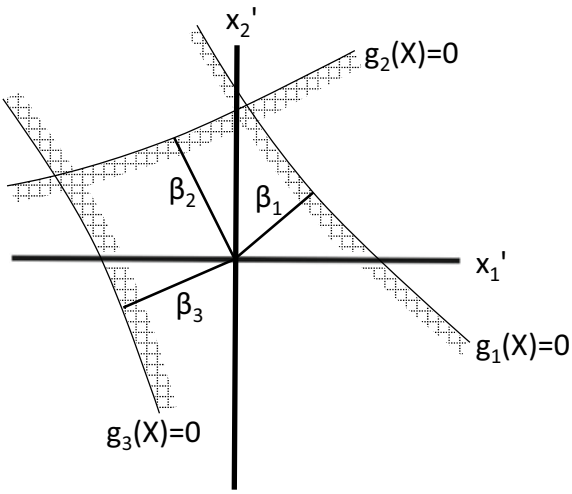


Fig. 7.1 Multiple limit states, $g_j(X)$, defining system failure, $g_j(X) = 0$

7.1 Idealized Systems

We often idealize a system into series or parallel to aid in conceptualizing how components are connected and interrelated.

7.1.1 Series

A series system is one where the entire system fails if any component in the systems fails (i.e., "weak-link-in-the-chain"). This is a nonredundant system and is the type of system that engineers should avoid; unfortunately in Civil Engineering we build many systems that are in series, particularly lifelines such as highways, bridges, pipelines, communication/transmission lines, canals, levees, etc. And even when systems are redundant at one scale, they can often be nonredundant at a different scale (e.g., the electricity grid has redundancy at the local level but when scaled to the regional level is often dependent on a single component or node when connecting to a neighboring regional grid).

The probability of failure of a series system is the joint union of all the failure states of the components.

$$p_{f,\text{series}} = P\left[\bigcup_{j=1}^{k} \left(g_j(X) \leq 0\right)\right] \tag{7.2}$$

7.1.2 Parallel

Parallel systems are redundant systems that only fail when all components in the system fail. These are desirable qualities for an engineered system. Redundancy is often critical when we are addressing collapse or other catastrophic failure modes. The probability of failure of a parallel system is the joint intersection of all the failure states.

$$p_{f,\text{parallel}} = P\left[\bigcap_{j=1}^{k}\left(g_j(X) \leq 0\right)\right] \tag{7.3}$$

7.1.3 General

Some systems can be a combination of series and parallel components in different arrangements (e.g., a water distribution system is better characterized as a general system). To generalize the discussion we can talk about cut sets and link sets; a cut set results in system failure and a link set results in system survival.

- In a series system each component is a cut set and all components together are a link set.
- For a parallel system, all components together form the only cut set and every component is a link set.

The probability of failure for a general system can be described as follows where the cut sets are defined as $[C_1 \ldots C_M]$.

$$p_{f,\text{general}} = P\left[\bigcup_{m=1}^{M}\bigcap_{j\in C_M}\left(g_j(X) \leq 0\right)\right] \tag{7.4}$$

Here we are taking the intersection of the events within the cut sets because these are parallel, and the union of the events between cut sets because these are series.

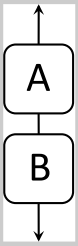
Example: Two-Component System

The following example evaluates a simple two-component system arranged in series and parallel. The two components have the same component probability of failure. The component failures are correlated due to similar construction and materials and/or similar loading:

$$P(A) = P(B) = 0.1 \quad P(A|B) = 0.5$$

If we arrange the components into a series system, then the probability of failure would be the union of the failure events; the failure of component A **or** the failure of component B:

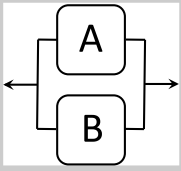
$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(AB) && \text{addition rule} \\ &= P(A) + P(B) - P(A|B)P(B) && \text{multiplication rule} \\ &= 0.1 + 0.1 - (0.5)0.1 = 0.15 \end{aligned}$$



Notice that the conditional probability between the two components reduces the probability of failure of a series system. If the component failures were perfectly correlated the system probability of failure would be 0.1, and if they were statistically independent the system probability of failure would be 0.19 (calculate these yourself to verify).

If we arrange the two components in a parallel system then the system probability of failure would be the intersection of the failure events; the failure of component A **and** the failure of component B:

$$\begin{aligned} P(AB) &= P(A|B)P(B) && \text{multiplication rule} \\ &= (0.5)0.1 = 0.05 \end{aligned}$$



If the components were perfectly correlated the system probability of failure would be 0.1, and if they were statistically independent the system probability of failure would be 0.01 (calculate these yourself to verify). Notice that conditional probability increases the probability of failure of parallel systems.

7.2 System Bounds

Bounds on the system probability of failure provide a simplified way of estimating the range that the probability of failure can take. How narrow the estimate of the range is a function of the complexity of the bounds estimate and how much information is included in estimating the bounds. The first order or unimodal bounds neglect the specific correlation between events and therefore can be rather wide. Higher order bounds account for specific correlation between events and provide narrower estimates as a function of the degree of correlation included.

7.2.1 Unimodal Bounds for Series Systems

Unimodal bounds for positively correlated ($\rho_{ij} > 0$) individual failure events ($E_i = [g_i(x) < 0]$) in a series system can be written as:

$$\left(\max_i P(E_i)\right) \leq p_{f,\text{series}} \leq \left(1 - \prod_{i=1}^{k} (1 - P(E_i))\right) \simeq \sum_{i=1}^{k} P(E_i) \text{ for small component } p_f \tag{7.5}$$

The left side of the inequality states that the lower unimodal bound on the probability of failure for a series system with positively correlated failure events is the max of any single failure event. Whereas the right side of the inequality states that the upper bound is approximately the sum of the probability of failure of all failure events, as the positive correlation between events will result in something less than this.

The complementary probability of safety ($p_s = 1 - p_f$) is then:

$$\prod_{i=1}^{k} P\left(\overline{E}_i\right) \leq p_{s,\text{series}} \leq \min_i P\left(\overline{E}_i\right) \tag{7.6}$$

The left side of this inequality is the product of the complement of the probability of failure of all failure events, and the right side is the minimum of the complement of the probability of failure for any single failure event.

Unimodal bounds for negatively correlated failure events $\left(\rho_{ij} < 0\right)$ in a series system are:

$$\left(1 - \prod_{i=1}^{k} P\left(\overline{E}_i\right)\right) \leq p_{f,\text{series}} \leq 1 \tag{7.7}$$

The left side of the inequality is the complement of the product of all nonfailures, and the right side is one (which is rather uninformative). The probability of safety is then:

$$0 \leq p_{s,\text{series}} \leq \left(\prod_{i=1}^{k} P\left(\overline{E}_i\right) \right) \tag{7.8}$$

The left side of the inequality is zero (again rather uninformative), and the right side is the product of the complement of the probability of failure for all failure events.

Example: Reservoir Problem

A reservoir is designed for both flood control and water supply. Flood control (F) is affected by snow melt (A) and rainfall (B). And a low reservoir (G) can be caused by a dry winter (C) and low rainfall (D). Say we know that snow melt and rainfall are positively correlated (i.e., wet winters lead to wet springs) but do not have a value of the correlation coefficient (ρ_{AB}). And we know that dry winters and low rainfall are positively correlated (i.e., dry winters lead to dry springs) but again we do not have a value of the correlation coefficient (ρ_{CD}). Intuitively we know that flooding and drought are negatively correlated (ρ_{FG}). If we are given the following component probabilities what is the probability of poor reservoir performance?

$$P(A) = 0.15 \; P(B) = 0.20 \; P(C) = 0.10 \; P(D) = 0.20$$

$$p_f = P(F \cup G) = P((A \cap B) \cup (C \cap D))$$

The statement above reads; the probability of failure or probability of poor performance of the reservoir can be due to flood control problems **or** a low reservoir. Flood control problems can be caused by snow melt **and** rainfall, and a low reservoir can be caused by low rainfall **and** a dry winter. We can estimate the bounds of this system using a first-order or unimodal approximation.

$$p_{f,\text{series}} \geq 1 - P\left(\overline{F}\right) P\left(\overline{G}\right)$$

because F and G are negatively correlated.

To determine $P\left(\overline{F}\right)$ we use unimodal bounds for positively correlated events.

$$\left[P\left(\overline{A}\right) P\left(\overline{B}\right)\right] \leq P\left(\overline{F}\right) \leq \left[\min\left(P\left(\overline{A}\right), P\left(\overline{B}\right)\right)\right]$$

$$0.85 \times 0.80 \leq P(\overline{F}) \leq 0.80$$
$$0.68 \leq P(\overline{F}) \leq 0.80$$

Similarly solving for $P(\overline{G})$.

$$[P(\overline{C})P(\overline{D})] \leq P(\overline{G}) \leq [\min(P(\overline{C}), P(\overline{D}))]$$
$$0.90 \times 0.80 \leq P(\overline{G}) \leq 0.80$$
$$0.72 \leq P(\overline{G}) \leq 0.80$$

Therefore

$$p_{f,\text{series}} \geq 1 - 0.80 \times 0.80 = 0.36$$

The probability of unsatisfactory reservoir performance is greater than or equal to 36% which is information that can aid in the decision or planning process for the reservoir as part of a larger a water system.

7.2.2 *Bimodal Bounds for Series Systems*

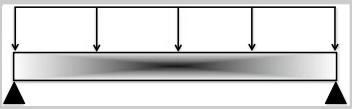
To improve the bounds (i.e., find a narrower approximation of the bounds) on a series system we can use bimodal or higher estimates. Bimodal bounds partially account for the correlation between failure events by using event pairs or joint events (E_i E_j). Ordering of the events can affect the results with higher modal estimates, and those interested are encouraged to read Song and Der Kiureghian (2003) for further details.

A method of bimodal bounds for series systems was derived by Kounias (1968) and Hunter (1976) and simplified by Ditlevsen (1979) by assuming Gaussian variates:

$$P(E_1) + \sum_{i=2}^{k} \max\left(P(E_i) - \sum_{j=1}^{i-1} P(E_i E_j); 0 \right) \leq p_{f,\text{series}}$$
$$\leq P(E_1) + \sum_{i=2}^{k} \left(P(E_i) - \max_{j<i} P(E_i E_j) \right) \tag{7.9}$$

Example: Unimodal versus Bimodal

A simply supported beam subjected to a uniformly distributed load can fail in flexure (f_1), shear (f_2), or both flexure and shear (f_3). Therefore failure of this system happens if any component fails $p_{f,\text{series}} = P(f_1 \cup f_2 \cup f_3)$. Component reliability has been performed for each failure mode and the results are:



	β_i	$p_{s,\,i}$	$p_{f,\,i}$
f_1	1.59	0.9445	0.0555
f_2	1.57	0.9418	0.0582
f_3	1.57	0.9418	0.0582

The failure modes are assumed to be positively correlated. The unimodal bounds of this system are:

$$\begin{aligned}\max(0.0555, 0.0582) \le p_{f,\text{series}} &\le 1 - (0.9445 \times 0.9418 \times 0.9418) \\ &\simeq (0.0555 + 0.0582 + 0.0582)\end{aligned}$$

$$\begin{aligned}0.0582 \le p_{f,\text{series}} &\le 0.1622 \\ &\simeq 0.1719\end{aligned}$$

If some numerical simulations (e.g., finite element) were run to determine how the failure modes are correlated we may have the following results:

$$P(f_1|f_2) = 0.45;\;\; P(f_1|f_3) = 0.99;\;\; P(f_2|f_3) = 0.57$$

The joint probabilities are then:

$$\begin{aligned}&P(AB) = P(A|B)P(B) \quad \text{multiplication rule} \\ &P(f_1 f_2) = 0.0250;\;\; P(f_1 f_3) = 0.0576;\;\; P(f_1 f_2) = 0.0332\end{aligned}$$

The bimodal lower bound is then:

$$\begin{aligned}&P(f_1) + (\max[(P(f_2) - P(f_1 f_2)); 0] \\ &\quad + \max[(P(f_3) - P(f_1 f_3) - P(f_2 f_3)); 0])\end{aligned}$$

$$
\begin{aligned}
&= 0.0555 + (\max[(0.0582 - 0.0250); 0] \\
&\quad + \max[(0.0582 - 0.0576 - 0.0332); 0]) \\
&= 0.0887
\end{aligned}
$$

And the upper bound is:

$$
\begin{aligned}
&P(f_1) + [P(f_1) - P(f_2 f_1)] + [P(f_3) - P(f_3 f_1)] \\
&= 0.0555 + [0.0582 - 0.025] + [0.0582 - 0.0576] \\
&= 0.0893
\end{aligned}
$$

Therefore

$$00887 \leq p_{f,\text{series}} \leq 0.0893$$

The table below shows the comparison between unimodal and bimodal bounds.

	Lower bound	Upper bound
Unimodal	0.0582	0.1622
Bimodal	0.0887	0.0893

In this example the bimodal bounds provided a much narrower range by including the correlation of joint events. The bimodal bounds here may constrain the system probability of failure sufficiently that it can be useful for engineering decision purposes.

7.2.3 Unimodal Bounds for Parallel Systems

Estimating the probability of failure for a parallel system is often not as critical as a series system because parallel systems have redundancy. Nonetheless the first order bounds for a parallel system treat the components as the wholly uncorrelated or perfectly correlated:

$$\prod_{i=1}^{k} P(E_i) \leq p_{f,\text{parallel}} \leq \min_i P(E_i) \tag{7.10}$$

Table 7.1 Table showing the general effect that correlation has on series and parallel systems

	Inc $+\rho$	Dec $+\rho$
Series system	$\downarrow$ Dec p_f	$\uparrow$ Inc p_f
Parallel system	$\uparrow$ Inc p_f	$\downarrow$ Dec p_f

Table 7.2 Table showing the general effect that the number of components has on series and parallel systems

	Inc # components	Dec # components
Series system	$\uparrow$ Inc p_f	$\downarrow$ Dec p_f
Parallel system	$\downarrow$ Dec p_f	$\uparrow$ Inc p_f

These bounds are often very wide and not highly informative. For parallel systems with a small number of components an estimate can be accomplished by using higher order bounds or multifold integration (Ang and Tang 1984).

For structures or other built features redundancy of the components is most often **active** as the components are each carrying load even though they are in a parallel arrangement. An example of active redundancy is structural columns that are designed to carry the full load in the event of failure of a nearby column. This compares to redundancy that is passive where parallel components are standby or backup in case of emergency. An example of **passive** redundancy is backup generators at a hospital that come online when there is a power outage. Parallel systems with active redundancy are markedly different than those with passive redundancy and should be treated accordingly.

7.3 Correlation and Components

Correlation or conditional probability between components can have a dramatic impact on the probability of failure, the difficulty in evaluating this impact often arises in the measuring and accounting for this dependence. Also the number of components in a system can have a large impact on the system probability of failure. Tables 7.1 and 7.2 show the general trends that correlation and number of components have on the probability of failure for a series or parallel system (Hollenback 2013). To more formally calculate this impact some type of simulation is often required. Alternatively, higher order bounds can provide an estimate.

If a series system has positive correlation between components then it is more reliable than a system of statistically independent components. Positive correlation is beneficial to series systems because if one component is in survival then all of the components are likely to be in survival, which leads to system survival. Since the same holds true for one component being in failure state, it might seem that the beneficial effect of positive correlation would be canceled out. However, if one component is in failure then the system fails regardless of other component states.

If a parallel system has positive correlation between components it will be less reliable than a system of statistically independent components. Positive correlation is

detrimental to parallel systems because if one component is in failure then all of the components are likely to fail, which leads to system failure. Since the same holds true for one component being in survival state, it might seem that the detrimental effect of positive correlation would get canceled out. However, if only one component is in survival state then the system survives regardless of the other component states.

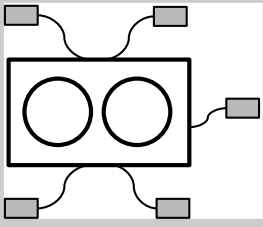
When considering the impact of the number of components of an idealized system we must consider system redundancy. Series systems are nonredundant; therefore, any increase in the number of components increases the likelihood of system failure. Parallel systems however are redundant and increasing the number of components increases the redundancy thereby decreasing the probability of failure.

Defining components of a system can be straight forward for some systems, and ambiguous for other systems. For a bridge system composed of deck sections the components are obvious. The same for a structure with columns supporting a floor slab. For systems such as highways or levees, determining what constitutes a component can be difficult and somewhat arbitrary. For spatially distributed systems such as lifelines we can define a component based on its spatially correlated load and resistance (Hollenback and Moss 2011; Moss and Germeraad 2013).

Note that correlation and conditional probability have been used interchangeable when discussing dependence. Conditional probability is a complete measure of statistical dependence, whereas the correlation coefficient is a partial or incomplete measure. As discussed previously in Chap. 3 in the section on conditional probability, the correlation coefficient is a linear estimate of statistical dependence. However for jointly normal or jointly lognormal random variables the correlation coefficient fully defines the conditional distribution and therefore is a complete measure of dependence in these special cases (personal communication, Armen Der Kiuregian, February 2018).

Example: Parallel and Series Power Generators

A nuclear power plant has five redundant backup power generators. These are designed to withstand strong ground shaking. Any one of the generators can provide sufficient energy to safely shut down the power plant, so if they are wired independently and are distributed around the site then this system might be considered a parallel system.



The component load and resistance, in typical ground shaking units of gravity, are assumed normally distributed with the following moments:

$$Q = N(0.10g, 0.03g) \quad R_i = N(0.20g, 0.05g)$$

The load is the forecast of strong ground shaking associated with some future event (usually estimated in a probabilistic manner). The resistance is the estimated capacity a generator has to resist the strong ground shaking. The limit state of each component is then:

$$g_i(X) = R_i - Q$$

If we assume that load and resistance are statistically independent the component probability of failure is:

$$p_{f_i} = \Phi\left(-\frac{0.20g - 0.10g}{\sqrt{0.05g^2 + 0.03^2}}\right) = 1 - \Phi(1.71) \cong 0.044$$

The system probability of failure for this parallel system is the intersection of all failure events (E_i) because for the system to fail all components must fail:

$$p_f = P(E_1 \cap E_2 \cap \ldots \cap E_5)$$

The bounds on this parallel system are (Eq. 7.10):

$$\prod_{i=1}^{k} P(E_i) \leq p_{f,\text{parallel}} \leq \min_i P(E_i)$$

$$1.65 \times 10^{-7} \leq p_f \leq 0.044$$

The left side of the inequality is for statistically independent component failure events, and the right side is for perfectly correlated component failure events. These bounds are rather large, so Ang and Tang (1984) considered the same problem and included correlation of the ground motion. This is reasonable since strong ground motions can be spatially correlated over typical distances of a power plant footprint. Ang and Tang solved for the correlation coefficient between component limit states ($\rho_{ij} = 0.265$) and then used a multifold integration approach to evaluate the system probability of failure. The exact solution is a fivefold integral of the joint standard normal PDF:

$$p_f = \int_{-\infty}^{-\beta} \int_{-\infty}^{-\beta} \cdots \int_{-\infty}^{-\beta} f_{g_1', g_2', \ldots g_5'} \mathrm{d}g_1' \ldots \mathrm{d}g_5'$$

where to get into standard normal space the limits states are written as:

$$g_i' = \frac{g_i - \mu_{g_i}}{\sigma_{g_i}}$$

To calculate this multifold integral a numerical solution is often the most tractable approach. Ang and Tang used numerical quadrature and found the system probability of failure to be $p_f = 1 \times 10^{-4}$. If correlation between the resistance of the individual components (due to manufacturing similarities or other) is subsequently included, then the system probability of failure would show an additional increase.

Now if we evaluate the same five backup power generators for tsunami loading there are other considerations. The component load and resistance, in units of wave height, are again assumed normally distributed:

$$Q = N(0.5\,\mathrm{m}, 0.5\,\mathrm{m}) \quad R_i = N(1.0\,\mathrm{m}, 0.2\,\mathrm{m})$$

If we assume that load and resistance are statistically independent the component probability of failure is:

$$p_{f_i} = \Phi\left(-\frac{1.0\,\mathrm{m} - 0.5\,\mathrm{m}}{\sqrt{0.2\,\mathrm{m}^2 + 0.5\,\mathrm{m}^2}}\right) = 1 - \Phi(0.928) \cong 0.1762$$

For the loading we might assume that the wave height is nearly perfectly correlated across the site ($\rho_{Q_i Q_j} \approx 1.0$) because of the scale and duration of the wave with respect to the scale of the power plant. Instead of resorting to a complex multifold integral we could use a proxy solution by assuming that the system is now acting in series, if one generator fails due to a tsunami wave then all will fail because of the load correlation. Utilizing bounds of a series system (Eq. 7.5):

$$\left(\max_i P(E_i)\right) \leq p_{f,\mathrm{series}} \leq \left(1 - \prod_{i=1}^{k} (1 - P(E_i))\right)$$

$$0.1762 \leq p_f \leq 0.6206$$

The right side of the inequality is the perfectly correlated bound which is a reasonable estimate of how the system will behave given the loading of a tsunami wave.

7.4 Chapter Summary

- A system is a group of components that make up some engineered feature.
- The probability of failure of a system can be difficult to determine, often requiring multifold integration or simulations.
- Systems can be idealized as **series** systems, where the system fails when any single component fails, or **parallel** systems, where the system fails only if every component fails.
- Bounds can provide an estimate on the upper and lower range of system probability of failure. Presented are **unimodal** series and parallel system bounds and **bimodal** series system bounds.
- **Correlation** of components can have a strong influence on the system probability of failure. Increasing positive correlation between components has a beneficial impact on series systems by increasing the likelihood of overall system survival, and detrimental impact on parallel systems by increasing the likelihood of overall system failure.
- An increasing **number of components** in a system increase the probability of failure for a series system because there are more components to fail, yet decrease the probability of failure for a parallel system because of redundancy.
- Redundancy in parallel systems can be **active** or **passive**. Passive redundancy is usually in the form of a backup system that comes into play when failure occurs. Active redundancy is often in the form of load-bearing components that are designed to carry additional load in the event of nearby component failure.

References

Ang, A. H.-S., & Tang, W. H. (1984). Probability concepts in engineering planning and design: Volume II-decision, risk, and reliability. Published by the authors.

Ditlevsen, O. (1979). Narrow reliability bounds for structural systems. *Journal of Structural Mechanics, 7*(4), 453–472.

Hollenback, J. C. (2013). *Reliability of levee systems under seismic loading* (Dissertation in partial fulfillment of Ph.D.). U.C. Berkeley.

Hollenback, J. C., & Moss, R. E. S. (2011). Bounding the probability of failure for levee systems. In *Georisk 2011 Conference Proceedings,* Atlanta, GA.

Hunter, D. (1976). An upper bound for the probability of a union. *Journal of Applied Probability, 13,* 597–603.

Kounias, E. G. (1968). Bound for the probability of a union, with applications. *American Mathematical Society, 39*(6), 2154–2158.

Moss, R. E. S., & Germeraad, M. (2013). Lifelines annex to the 2013 California state multi-hazard mitigation plan. http://hazardmitigation.calema.ca.gov/plan/state_multi-hazard_mitigation_plan_shmp.

Song, J., & Der Kiureghian, A. (2003). Bounds on system reliability by linear programming. *Journal of Engineering Mechanics, 129*(6), 627–636.

Chapter 8
Introduction to Decision Analysis

Component (Chap. 6) and System (Chap. 7) reliability have been presented for computing the probability of failure, but engineering problems are often more complicated than calculating the component or system reliability. In this chapter some tools for intuitively mapping out a complex engineering problem and moving toward a decision are demonstrated. These tools can be described as trees because visually they are often presented like branches of a tree. These tools are grouped here into four types: event trees, decision trees, logic trees, and fault trees. Schematics of these trees are shown in Figs. 8.1 and 8.2.

An event tree maps the potential outcomes of some situation with the probability of those outcomes explicitly described on each branch. The probabilities must sum to 1.0 vertically down across all branches, thereby capturing the total probability at each step. Decision trees are event trees with the costs or consequences for each branch included. The probability of a particular outcome is multiplied by the cost of that outcome, thereby quantifying the risk of that particular outcome. Decision trees provide risk, a calibrated metric, for engineering decisions as was first discussed in Chap. 1. The probability of each branch can be determined using the various tools described throughout this text, statistics, probability, and reliability. Often one branch of an event or decision tree requires a component (or system) reliability analysis to define the probability of that particular outcome.

A logic tree is often used in situations where a particular problem has multiple solution methods. Rather than arbitrarily selecting one solution method, all methods are used and weighted according to the accuracy of or confidence in a particular method. The probabilities here are often determined using expert consensus. Lacking any prior information equal probabilities should be used. Sensitivity studies can then determine the impact of a particular method on the results. A logic tree provides a robust mean or median estimate by using all available information to defeat epistemic uncertainty and solve the problem. However, the variance from a logic tree is ill defined and there is little agreement as to how best to measure the variance within this type of weighting scheme.

A fault tree does not contain probabilities for each branch because it is a visual tool used for building often nontrivial probability statements by combining the unions

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_8

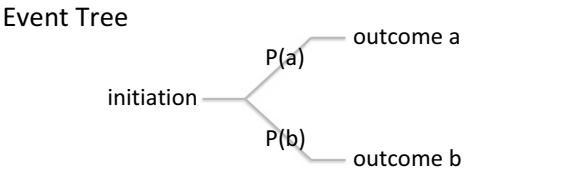
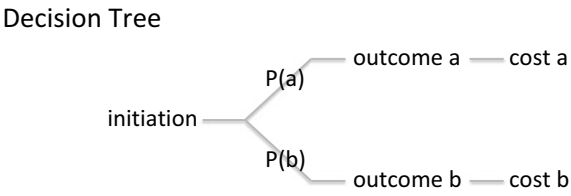


Fig. 8.1 Schematics of event and decision trees



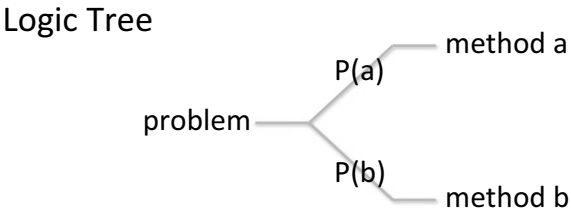
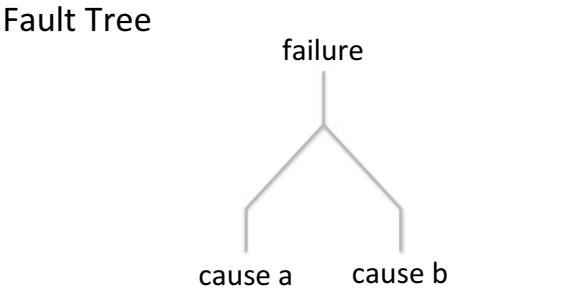


Fig. 8.2 Schematics of logic and fault trees



and intersections of subevents that cause failure. This tree is useful for complex interdependent systems where a total probability statement is not intuitive from the outset.

The best way to fully demonstrate these tools is through example, so the bulk of this chapter will present specific problems where trees are an asset in problem solving and moving the decision process forward.

8.1 Correlation in Decision Analysis

Correlation as a concept is rather straightforward, but dealing with correlation in a specific and quantifiable manner can be at times quite difficult. The term correlation can describe many different interrelationships and phenomena. The probability of each branch of a tree can be correlated via causal dependence, probabilistic correlation, spatial/temporal autocorrelation, and/or statistical correlation. This list may not be complete but tries to encompass the bulk of correlation. Below is described each of these types of correlation with examples from a levee system to provide concrete context.

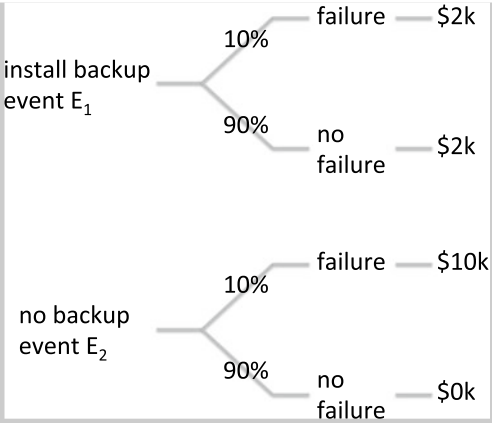
Causal dependence is where one event causes another, and this is sequential failure which is by nature a conditional probability (e.g., foundation settlement of a levee increases the probability of a flood wave overtopping the levee).
Probabilistic correlation is where two uncertainties may depend on a third uncertainty (e.g., the density of soil is an independent variable that can impact two other variables important to levee performance, soil piping, and excess pore pressure generation).
Spatial or temporal autocorrelation is where two uncertainties are a function of space or time (e.g., various soil properties of the levee foundation material are spatially correlated due to the depositional nature of the soil. Seismic loading of a levee system is temporally correlated due to the finite nature of earthquake fault rupture).
Statistical correlation is where two uncertainties are estimated from a data set that is influenced by a common variance (e.g., the cohesion and friction angle of levee embankment and levee foundation soil are estimated through a linear regression of the Mohr–Coulomb failure envelope to laboratory data and are negatively correlated).

8.2 Decision Analysis Examples

The following are examples demonstrating the utility of different decision tools.

Example: Power Failure Decision Tree

A power company has the option to install a backup power system at a cost of $2k per year. If power goes out the company will take a $10k hit due to penalties and fines. Based on an internal reliability analysis of the system the annual probability of failure is 10%.



Calculating the risk of events E_1 and E_2:

$$R(E_1) = 0.1(\$2\text{k}) + 0.9(\$2\text{k}) = \$2\text{k}$$

$$R(E_2) = 0.1(\$10\text{k}) + 0.9(\$0\text{k}) = \$1\text{k}$$

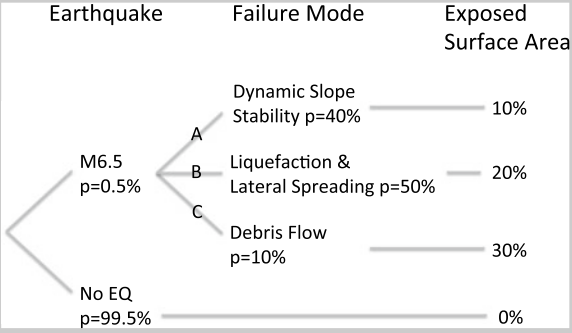
Based on this risk assessment the least expensive option would be to not install the backup system. Of course this assumes that the costs and the probabilities have been accurately assessed. [As a side note, previous case histories have demonstrated time and again that people are notoriously bad at estimating the consequences of a failure or catastrophe before hand (Moss and Germeraad 2013); therefore in this problem the $10k hit the company may experience if no backup is installed may be an optimistically low estimate of the consequences.]

Example: Shell Mound Decision Tree

Offshore drilling operations will leave rock cuttings from the drilling process on the ocean floor. These cuttings are composed partially of petroleum bearing rock. The pile of cuttings left after the drilling is complete is called shell mounds because sea organisms with shells (among other animals) take up residence on the cuttings. Off the coast of central California an oil company had decommissioned the drilling platforms and wanted to be recused of further liability associated with the shell mounds they created. The question that the local citizens asked was "Should the shell mounds be removed, or should they stay in place in perpetuity? And if they stay in place, what is the risk given that the area is prone to earthquakes?"

This is a relatively complex problem that can be broken down into a decision tree for finding the optimal solution. Here the consequences were quantified in terms of surface area of the shell mound exposed, which is the percent area of a shell mound where the petroleum-bearing rock is re-exposed due to seismic-related deformations. This information was then combined into a larger decision tree that incorporated costs associated with sea life, fishing operations, beach tourism, and long-term resident health.

A decision tree provided the framework and guidance for getting a first-order estimate of the risk due to earthquakes. The conditional probability of exposed surface area of the shell mound for a particular failure mode was calculated given the annual probability of exceedance of a large damaging earthquake. Based on offshore sampling and laboratory data, engineering calculations were combined with expert consensus to provide a relative assessment.



The large damaging earthquake (M6.5) is the maximum expected event for that area, and the annual probability of exceedance was calculated using a logic tree common in PSHA (probabilistic seismic hazard analysis; McGuire 2004). The risk associated with each failure mode is:

$$R(A) = 0.005(0.40)0.10 = 0.00020$$
$$R(B) = 0.005(0.50)0.20 = 0.00050$$
$$R(C) = 0.005(0.10)0.50 = 0.00025$$

And the total annual risk due to an earthquake is the sum of the risk from each failure mode:

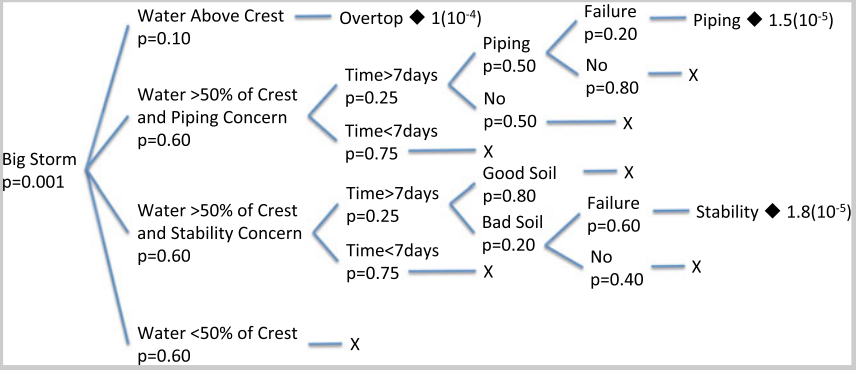
$$R(\text{EQ}) = 0.0002 + 0.0005 + 0.00025 = 0.00095$$

Compare this to 100% exposure in the year of removal if the shell mound were to be removed.

Armed with this rough analysis and subsequent details provided by a team of economists, the citizens made a decision to leave the shell mound in place and absolve the oil company of any future liability. The risk was effectively communicated to the stake holders, and a rational decision was made that considered all the available information.

Example: Levee Failure Event Tree

The following event tree describes flooding potential associated with levee performance. Here we map the potential of poor planning that results in levee overtopping and poor design that results in levee failure.

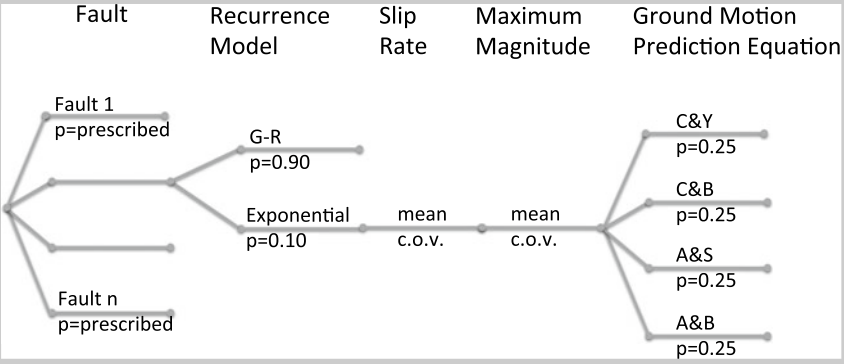


The event tree shows the relationship between the possible failure modes for this levee system. A branch of the tree that ends in an **X** denotes no failure. A branch of a tree that ends in failure has a ◆ and the associated probability.

In this example, the probability of a big storm, usually reported as an annual probability, is 0.1%, which is probably based on historical weather and stream flow data for the region. The overtopping here has the highest probability of occurrence when compared to piping and stability failure. This means that the peak water level (i.e., loading) for which the levee was designed was not adequately characterized, as opposed to piping and stability failures where the engineering of the levee (resistance) is faulty.

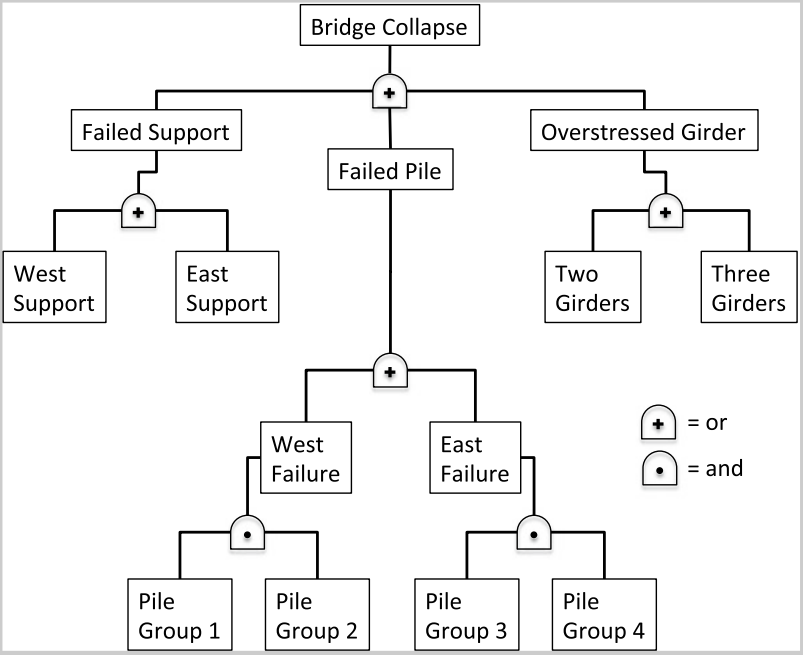
Example: Seismic Hazard Logic Tree

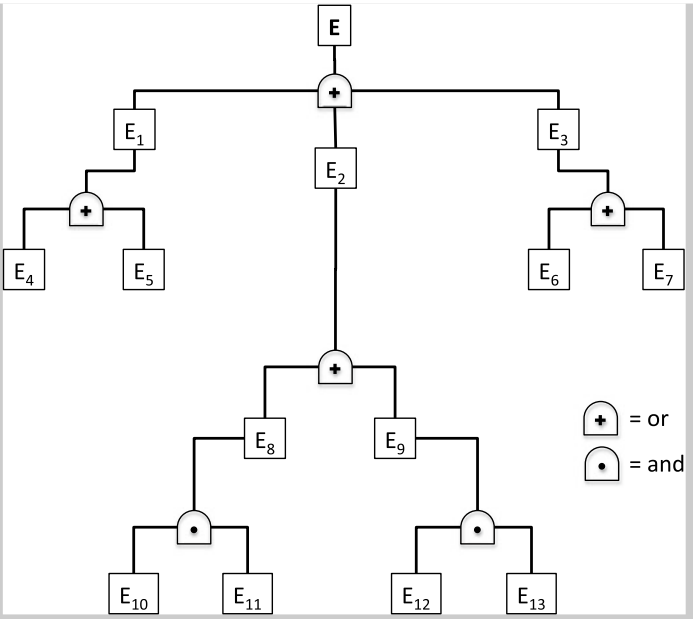
Logic trees are used frequently for the problem of assessing seismic hazard. The annual probability of exceeding some ground shaking level is calculated as a multifold integral of conditional probabilities of earthquake rupture, magnitude, and distance. There are many competing models for calculating the seismic hazard that each contains some epistemic uncertainty. To defeat the epistemic uncertainty these competing models are used and weighted according to accuracy, confidence, or some other scheme. The following logic tree shows one path of a seismic hazard calculation that would be performed using Monte Carlo simulations to generate many realizations. The fault branch can include many faults in a region that could rupture and the probability associated with the fault represents the confidence in how viable the fault is in producing a rupture (this can include alternate fault geometry, fault segmentation, or other fault characteristics). The recurrence model branch includes two common relationships weighted according to how well they describe that particular fault. The slip rate and maximum magnitude branches are represented by a single model each with some defined uncertainty. And the ground motion prediction equation branch includes four equally viable models that are all included to defeat epistemic uncertainty in the individual models and produce a robust mean or median result.



Example: Bridge Failure Fault Tree

In this example we are trying to quantify the probability of failure of an event that comprised of many subevents. The fault tree allows us to intuitively diagram the different subevents to then produce a statement of the total probability of the event. The different subevents, here different failure modes of a bridge, are shown connected by symbols that represent **or** (indicating union) or **and** (indicating intersection). As shown in the following fault tree, bridge collapse can be caused by a failed support **or** a failed pile **or** an overstressed girder. These events are then broken down into further subevents. We can then replace the description of the event and subevents with event numbers to aid us in writing the total probability statement. The order of the numbering is unimportant; it is just a shorthand way of keeping track of the subevents.





Using the numbered event tree we can write a probability statement describing the system as shown below. At the first level we have:

$$P(E) = P(E_1 \cup E_2 \cup E_3)$$

We take that to the second level of subevents:

$$P(E) = P((E_4 \cup E_5) \cup (E_8 \cup E_9) \cup (E_6 \cup E_7))$$

And the final level of subevents for this event tree is:

$$P(E) = P((E_4 \cup E_5) \cup ((E_{10}E_{11}) \cup (E_{12}E_{13})) \cup (E_6 \cup E_7))$$

The event tree allowed us to systematically piece together the complex probability statement of bridge collapse including several different failure modes. Component reliability analysis would inform us as to the probability of failure for each subevent which can then be plugged into this complex probability statement to arrive at the total probability of failure of the bridge, $P(E)$.

8.3 Chapter Summary

- Complex or multicomponent systems can often be intuitively mapped using different trees.
- Trees can allow for a broader perspective on a risk or probability of failure problem, thereby providing a framework for the component or system reliability analysis.
- Event trees and decision trees map the sequence of events that lead up to failure. Event trees quantify the probability of failure for a system, and decision trees include consequences to provide risk for a system.
- Logic trees are useful when multiple models can be used to solve a problem. The tree provides a weighting framework for including all models, thereby minimizing epistemic uncertainty in each individual model.
- Fault trees can allow for a much clearer understanding of multiple failure modes in a system and can provide a means of determining a total probability statement for the system.

References

McGuire, R. K. (2004). *Seismic hazard and risk analysis*. Earthquake Engineering Research Institute (EERI) Monograph MNO-10.

Moss, R. E. S., & Germeraad, M. (2013). *Lifelines annex to the 2013 California State multi-hazard mitigation plan*. http://hazardmitigation.calema.ca.gov/plan/state_multi-hazard_mitigation_plan_shmp.

Chapter 9
Reliability-Based Codes (LRFD)

There are many design codes used in engineering design today. Originally most codes used some form of an allowable or working stress design (ASD/WSD) approach.

$$\frac{R}{\mathrm{FS}} \geq \sum Q_i \tag{9.1}$$

R is the resistance, FS is the factor of safety, and Q_i are the loads (e.g., dead, live, extreme…). The factor of safety represents the lumped uncertainty from load and resistance and the desire for overdesign and safety in one factor. This approach has been used to check and ultimate limit state (i.e., failure, collapse) and/or a service limit state (i.e., deformation, deflection).

Within the last few decades design codes have been switching to a load and resistance factor design (LRFD) approach. The basis of LRFD is reliability, but the reason for the switch is because using a reliability-based approach provides a more accurate answer which often results in cost savings for the project.

$$\varphi R \geq \sum \alpha_i Q_i \tag{9.2}$$

R and Q_i are the same resistance and loads as described above, but instead of one lumped factor of safety we have φ the resistance factor and α_i the load factors. The resistance is factored down and the loads factored up to achieve a desired design level. The design level in LRFD is based on the reliability index, thereby pinning the reliability on the number of standard deviations away from failure the design should be. This provides a more rational relationship between load and resistance by accounting for the respective uncertainties in each.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_9

9.1 Reliability Formulation

There are two methods for determining the load and resistance factors for LRFD design:

(1) Optimizing the LRFD equation using empirical data and reliability methods (FOSM and/or FORM). A target reliability index is used, often in the range of 2.0–3.0.
(2) Calibrating the LRFD equation to previous ASD/WSD design practice using codified FS.

The first method is preferred, but there is often insufficient data to carry this out. The second method is comforting because it relies on past practice but may not achieve the goal of a more accurate answer and cost savings that the first method can. Examples of each are shown below.

9.1.1 Calibrated LRFD (Method 2)

We will start with method 2 where we are calibrating load and resistance factors from past codified factors of safety. If we combine Eqs. (9.1) and (9.2) and solve for the resistance factor:

$$\varphi \geq \frac{\sum \alpha_i Q_i}{\mathrm{FS} \sum Q_i} \tag{9.3}$$

If only dead and live load are considered:

$$\varphi = \frac{\alpha_D Q_D + \alpha_L Q_L}{\mathrm{FS}(Q_D + Q_L)} \tag{9.4}$$

By dividing through by Q_L we get an equation in terms of the load ratio:

$$\varphi = \frac{\alpha_D \frac{Q_D}{Q_L} + \alpha_L}{\mathrm{FS}\left(\frac{Q_D}{Q_L} + 1\right)} \tag{9.5}$$

If we fix the load factors (that means fixing the uncertainty from dead and live loads) then we can back-calculate the resistance factor for a given factor of safety. Typical load factors from AASHTO (American Association of State Highway Transportation Officials who have implemented LRFD in highway bridge design) are $\alpha_D = 1.25$ and $\alpha_L = 1.75$. Fixing the load factors means that the remaining uncertainty in the factor of safety is attributed to the resistance; we are in essence partitioning the lumped uncertainty in the factor of safety into load and resistance sides of the equation (Table 9.1).

Table 9.1 Factor of safety calibrated resistance factors (φ) for different dead (Q_D) versus live load (Q_L) ratios

Calibrated φ factors from past FS			
FS	$Q_D/Q_L = 1$	$Q_D/Q_L = 2$	$Q_D/Q_L = 3$
1.5	1.00	0.94	0.92
2.0	0.75	0.71	0.69
2.5	0.60	0.57	0.55
3.0	0.50	0.47	0.46

The loads are to be factored up by 125 and 175%, respectively. If our ratio of dead to live load is 1 and the ASD/WSD design traditionally used a FS = 3.0 then we will factor our resistance down by 50% to proceed with the design using LRFD (Eq. 9.2).

9.1.2 Optimized LRFD (Method 1)

If there is sufficient data to perform statistics then the load and resistance factors can be evaluated directly from the data, and not back-calculated from prior factors of safety. This is the preferred method but requires sufficient data and careful study of the particular engineering problem.

For this method we will use a shallow foundation design problem (borrowed from Baecher and Christian (2003)) and FOSM to illustrate the application. The LRFD Eq. (9.2) (rewritten here):

$$\varphi R \geq \sum \alpha_i Q_i$$

We treat R and Q as random variables and in this problem will assume we are dealing with a single load to simplify the example:

$$\varphi \mu_R \geq \alpha \mu_Q \tag{9.7}$$

And if both R and Q are Gaussian (to allow for an exact solution) we know from Chap. 6 on component reliability that the margin of safety and reliability index are:

$$M = R - Q$$

$$\beta = \frac{\mu_M}{\sigma_M} = \frac{\mu_R - \mu_Q}{\sqrt{\sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q}}$$

Solving for μ_R and substituting into the single load LRFD Eq. (9.7), we arrive at an expression for the resistance factor:

$$\varphi = \frac{\alpha\mu_Q}{\mu_R} = \frac{\alpha\mu_Q}{\beta\sqrt{\sigma_R^2 + \sigma_Q^2 - 2\rho_{RQ}\sigma_R\sigma_Q} + \mu_Q} \tag{9.8}$$

The reliability index is set at a target value, usually 2 or 3 depending on the number of standard deviations away from failure that is deemed acceptable for the particular engineering design. A target reliability index, like a factor of safety, is determined via consensus usually through building code panels or other expert consensus forums.

That leaves determining the moments of the load and resistance in addition to the load factor in order to solve for the resistance factor. Note that Eq. (9.8) applies to any similar engineering design situation where we are dealing with a single load and we have normally distributed load and resistance.

In this example we are analyzing the bearing capacity of a shallow foundation footing. For this problem R is the soil resistance or in geotechnical terms the ultimate bearing capacity (q_{ult}) as calculated using Terzaghi's method, Q is the single dead load due to the foundation often called the bearing pressure (q), and the load factor found in AASHTO foundation design recommendations, as discussed previously.

Terzaghi's method for calculating the soil resistance uses the following equation:

$$R = q_{\text{ult}} = cN_c + \sigma'_{ZD}N_q + \frac{1}{2}\gamma BN_\gamma$$

We will assume it is a strip footing ($B = 1m$) with its base located at the soil surface ($\sigma'_{ZD} = 0$) resting on a sandy soil ($c = 0$) with a unit weight (γ) of 18.5 kN/m^3 and a friction angle (ϕ) of 35°. For readers unfamiliar with bearing capacity analysis the details of the equation are unimportant for this discussion, only the fact that the equation is a function of a random variable representing resistance is important. The bearing capacity equation reduces when we input the assumptions to:

$$R = q_{\text{ult}} = \frac{1}{2}\gamma BN_\gamma$$

The above equation is a function of the unit weight (γ), the footing dimension (B), and an empirical factor (N_γ). N_γ is itself a function of the friction angle (ϕ), the solution geometry, and the physical assumptions behind Terzaghi's method. If we treat N_γ as a random variable we can propagate the uncertainty it contains through the bearing capacity equation. N_γ is usually determined using linear regression which is a "best fit" to data that has a certain amount of scatter as shown in Fig. 9.1.

Using results from Terzaghi's (1943) original analysis for $\phi = 35°$ the statistics of the empirical factor are $\mu_{N_\gamma} = 82$ and $\sigma_{N_\gamma} = 17$. Using FOSM to estimate how this uncertainty propagates to q_{ult}:

$$\mu_{q_{\text{ult}}} \approx \tfrac{1}{2}\gamma B\mu_{N_\gamma} = 760\,\text{kPa}$$
$$\sigma_{q_{\text{ult}}} \approx \sqrt{\sigma_{N_\gamma}^2\left(\tfrac{1}{2}\gamma B\right)^2} = 157\,\text{kPa}$$
$$\delta_{q_{\text{ult}}} \approx 21\%$$

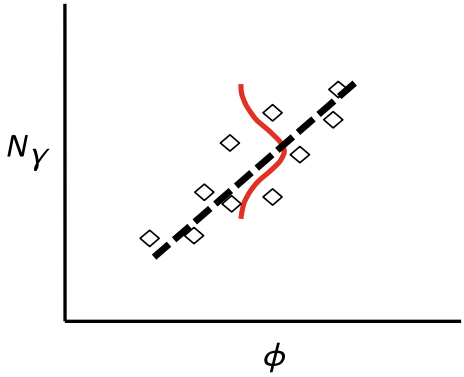


Fig. 9.1 Schematic showing linear regression results of friction angle versus the empirical bearing capacity factor. The scatter or uncertainty in the relationship is evident, and a hypothetical distribution is overlaid

We now have the resistance side of the problem defined by its first and second moments. For the load side of the problem the bearing pressure is usually calculated as a function of the column load (P), area (A), self weight (W) of the footing, and any uplift pore pressure (u) caused by a water table above the footing base:

$$q = \frac{P + W}{A} - u$$

For this example we will assume that the bearing pressure is roughly 250 kPa per linear meter of the strip footing. To quantify the uncertainty of this load we look up a coefficient of variation reported in the structural codes, which gives $\delta_q \approx 15\%$. The codes also provide the load factor, and as mentioned we are just analyzing for a dead load $\alpha_D = 1.25$. Typical shallow foundation design would have a target reliability index of $\beta = 3.0$.

Neglecting any correlation between load and resistance we can calculate the resistance factor using Eq. (9.8):

$$\varphi = \frac{1.25(250)}{3.0\sqrt{157^2 + 37.5^2} + 250} = \frac{312.5}{734.3} = 0.43$$

This resistance factor is a function of the uncertainty, and here the uncertainty is from the method used to determine the bearing capacity. Table 9.2 shows a comparison of this calculated value to typical values (Barker et al. 1991; Withiam et al. 2001).

A semi-empirical method uses field test results directly for the resistance calculation (e.g., SPT blow counts). A rational method uses engineering properties that were converted from field or laboratory test results for the resistance calculation (e.g., friction angle from SPT blow counts).

For shallow foundation design, the soil type, method of analysis, and type of test used to measure the soil properties all affect the amount of uncertainty in the resultant, which is reflected in the value of the resistance factor. In LRFD design, the

Table 9.2 Typical resistance factors for shallow foundation bearing capacity analysis

Soil type	Method + test	Resistance factor ϕ	FS ASD	FS LRFD
Sand	Semi-empirical + SPT	0.45	3.0	3.2
	Semi-empirical + CPT	0.55	2.5	2.6
	Rational + SPT	0.35	2.5	4.1
	Rational + CPT	0.45	2.5	3.2
Clay	Semi-empirical + CPT	0.50	2.5	2.9
	Rational + lab	0.60	2.5	2.4
	Rational + field vane	0.60	2.5	2.4
	Rational + CPT	0.50	2.5	2.9

higher the resistance factor, the higher the confidence in the answer because there is less uncertainty in the variables and/or the resultant.

Shown in Table 9.2 are the factors of safety typically used for ASD design and the true or LRFD-equivalent factors of safety that include the uncertainty in the analysis. In most cases, shallow foundations are overdesigned using ASD because the uncertainty is not properly quantified. This is the impetus for quantifying the uncertainty and the basis of LRFD design: to better define the threshold of failure and properly design against it. Overdesign is often costly, both in time and materials, and by performing a more accurate analysis by including the uncertainty a better more cost-effective design can often be achieved.

9.2 Chapter Summary

- LRFD is a design code framework that includes quantified uncertainty in the load and the resistance.
- The basis for LRFD is reliability, the same methods that were presented in Chap. 6.
- Load and resistance factors can be derived from statistics of the problem (Method 1) or back-calculated from a previous factor of safety (Method 2).
- Using LRFD provides a more accurate answer to the design problem, which translates to a safer design that also often results in cost savings.

References

Baecher, G. B., & Christian, J. T. (2003). *Reliability and statistics in geotechnical engineering*. New Jersey: Wiley.

Barker, R. M., Duncan, J. M., Rojiani, K. B., Ooi, P. S. K., Tan, C. K., & Kim, S. G. (1991). *Load factor design criteria for highway structure foundations* (Final Report, NCHRP Project 24-4). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Terzaghi, K. (1943). *Theoretical soil mechanics*. New York: Wiley.
Withiam, J. L., et al. (2001). *Load and resistance factor design (LRFD) for highway bridge substructures* (Report No. FHWA HI-98-032). Washington, D.C.: Federal Highway Administration.

Chapter 10
Spatial Variability

All uncertainty addressed in this text so far has been in the form of point estimates, or uncertainty that can be attributed to a specific location. Uncertainty is also a function of differences in properties from location to location. This is commonly referred to as spatial variability and is the focus of this chapter. Uncertainty can also be a function of the progression of time (temporal uncertainty), but that will not be addressed here.

Spatial variability is considered and quantified in many fields of engineering but was first done so for developing location-dependent probability distributions of ore grades for mining operations (Isaaks and Srivastava 1989). In mining and related areas the means of quantifying spatial variability are often called geostatistics and applications can be found in the fields of geology, petroleum exploration, hydrogeology, oceanography, geography, agriculture, geographic information systems, epidemiology, logistics, and others where spatial variability is of import.

In our discussion of spatial variability we will focus on the correlation of performance between individual components within a Civil Engineering system and how that influences system reliability. We are primarily interested in how component performance parameters are jointly distributed and how the correlation structure of this joint distribution is dependent on the relative spacing of the components. Other applications of geostatistics focus on interpolation and simulation of random variable realizations with distance, which is often referred to as kriging.

To frame the discussion let's take a linear engineered system such as a railroad corridor that is founded on an embankment (Fig. 10.1). If we want to determine the probability of failure of this entire corridor, we are interested in the probability of failure of the components that make up this system. In reliability terms we are interested in the joint distribution of load and the resistance as a function of distance along the railroad corridor. The rails and the embankment are man-made and will have properties associated with their fabrication and construction. But construction along this corridor took time and maybe the conditions changed throughout the building process resulting in the variability of the resistance with distance (e.g., temperature/weather changes, labor supply changes, differential exposure of materials, variable labor quality with time/distance).

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_10

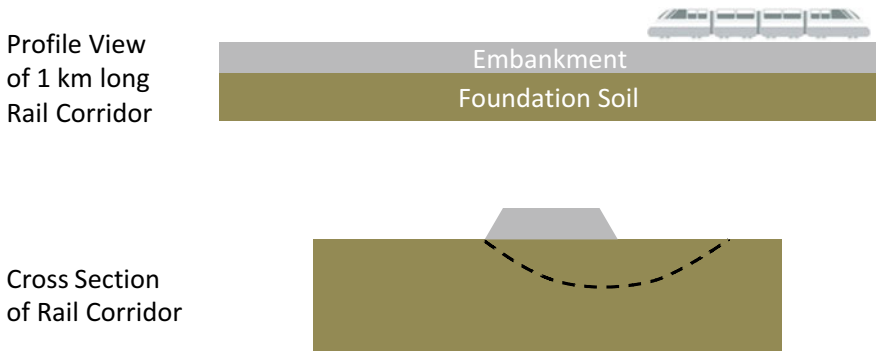


Fig. 10.1 Conceptual image of a rail corridor showing the profile and cross section. Failure mechanism of interest is bearing capacity failure

The foundation soil below the embankment has variability with distance as a function of the depositional environment, past and present stress conditions, land use, etc.

For this corridor if we divide it up into three segments and treat the segments as components, then the system probability of failure for this series would follow the addition rule:

$$M = R - Q \quad \text{for each component}$$

$$\begin{aligned} P(f_{\text{system}}) &= P(\cup M_i) = P(M_1 \cup M_2 \cup M_3) \\ &= P(M_1) + P(M_2) + P(M_3) - P(M_1 M_2) - P(M_1 M_3) \\ &\quad - P(M_2 M_3) + P(M_1 M_2 M_3) \end{aligned}$$

Without the joint probabilities we could bound the system unimodally as in Chap. 7 to estimate the system probability of failure. But the joint probability terms are informed by the spatial variability (in this case mainly the resistance), and if we quantify the spatial variability, then we can better estimate the probability of system failure as well as statistically define the length that constitutes a normative size of a segment given the physics of the problem.

As we have done with stationary problems in this textbook, spatial problems will start with statistics (interrogating the past) and then proceed toward probability (forecasting the future).

10.1 Empirical Spatial Statistics

The properties of something can have spatial dependence in its; overall trend, extreme value locations, degree of continuity, and correlation. The univariate (central ten-

dency and dispersion) and bivariate (correlation) statistics we developed in Chap. 1 do not apply to these spatial properties.

Overall trends may be important. Think of a surface topography of the land, algae concentrations on a pond, and texture of a concrete slab. These can be mapped as contour maps, symbol maps (categories or bins), or indicator maps (Boolean).

Extreme values or anomalies are often of interest, and regions of high variability can be important when quantifying spatial variability. The term heteroskedasticity means having regions of differing variability, and homoscedasticity means having regions of the same variability. Moving window statistics is a means of binning the data with distance. In Fig. 10.2 we see two different situations: the first (top plot) showing the central tendency and the dispersion that are relatively constant with distance, whereas the second (bottom plot) showing both the central tendency and dispersion trending with distance.

For spatial continuity, values that are at a closer distance to each other are more similar than values that are at further distance from each other. This is common with soil properties or fluid concentrations or load interfaces or other similar situations. In Fig. 10.3 we show plots of some value at location x with respect to another value at location $x + h$, where h is a fixed separation (or lag) distance between points. As we increase h from 1 to 10 we see an increase in scattering indicating a spatial dependence of correlation. Using ellipses as in Chap. 1 we can visualize the ellipse getting fatter as we go from $h = 1$ to $h = 10$ which means the correlation coefficient is decreasing with increasing separation distance.

The plots above can be generated for as many separation distances as feasible with a given dataset spread out over some distance. Paired data (x and $x + h$) "hopscotch" along at increasing separation distances until there are a diminishing number of pairs (Fig. 10.4). For an accurate characterization of spatial continuity there generally needs to be a minimum of 100 values with distance (Webster and Oliver 1992). The

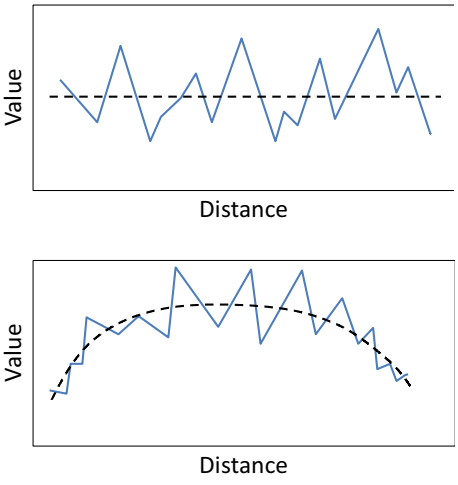


Fig. 10.2 Central tendency and dispersion binned and plotted as a function of distance for homoscedastic and heteroskedastic situations

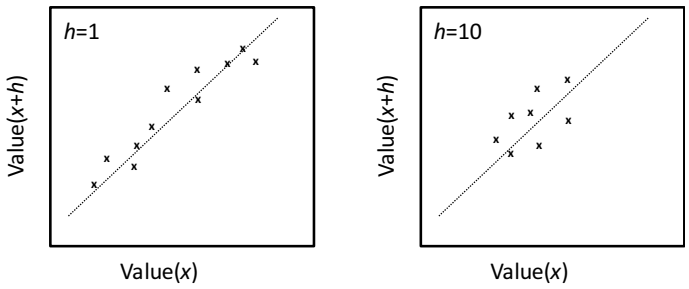


Fig. 10.3 Plots of spatial continuity with separation distance (h). As the separation distance increases, the spatial continuity and spatial correlation decrease

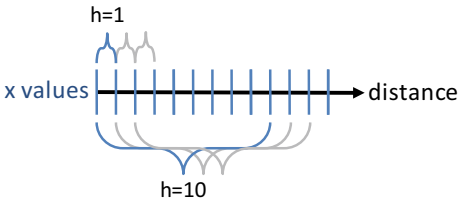


Fig. 10.4 Schematic of value x with distance measured at increasing separation distances (h)

sampling interval is important to properly characterize what is called the range, or the distance over which values are spatially correlated.

The aggregate results of spatial continuity plots can be used to evaluate the spatial continuity in three common formats:

1. Correlation function,
2. Covariance function,
3. Moment of Inertia.

The correlation function (also called autocorrelation function) is found by calculating the statistical correlation for each separation distance (h) and then plotting these as a function of separation distance. The covariance (aka autocovariance) as you may recall is the numerator of correlation (Eq. 2.6) and is calculated spatially and plotted as a function of separation distance. The moment of inertia, which is also called the semi-variance in geostatistics, is the "fatness" of the ellipse for each separation distance. The semi-variogram quantifies when the spatial dependence of data scatter reaches a maximum. Spatial correlation and spatial covariance functions quantify when the spatial dependence reaches a minimum. Therefore, for a standard normal assumption the covariance function and semi-variogram are complements of each other [i.e., $C(h) = \mathrm{Var}(h) - \gamma(h) = 1 - \gamma(h)$]. The correlation function is simply a scaled version of the covariance function so that it has a maximum value of 1.0 (Fig. 10.5).

The semi-variogram can be calculated using Eq. (10.1);

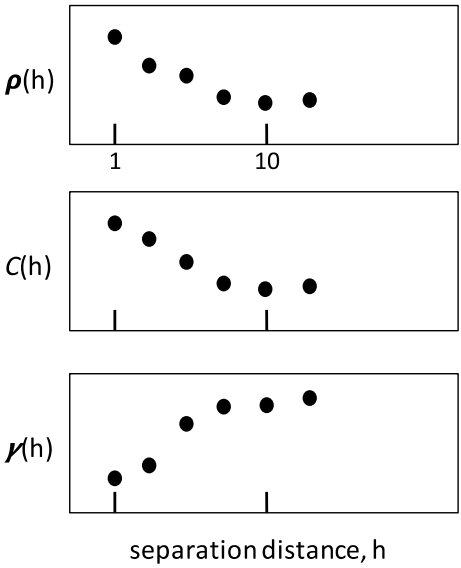


Fig. 10.5 Conceptual plots of spatial correlation $\rho(h)$, spatial covariance $C(h)$, and spatial semi-variance $\gamma(h)$

$$\gamma(\boldsymbol{h}) = \frac{1}{2N(\boldsymbol{h})} \sum_{(i,j)|h_{ij}=h} (x_i - x_j)^2 \tag{10.1}$$

where the data values are $x_i, \ldots, x_n$ and the summation of the $N(\boldsymbol{h})$ whose locations are separated by $\boldsymbol{h}$. The $\boldsymbol{h}$ is bolded to represent multiple separation distances. The semi-variogram represents half of the average squared difference between data pairs and can be shown to equal the variance of the values being estimated. Similarly, equations can be found for calculating the covariance and correlation functions, but for our purposes, we will stick to the semi-variogram.

Semi-variograms are more commonly used in geostatistics because they tend to filter out the influence of spatially varying mean values. The covariance function and correlation function require second-order stationarity (constant variance), whereas the semi-variance only requires what has been termed "intrinsic" stationarity (Matheron 1965; Cressie 1992). What this means in practice is that the semi-variogram tends to be a more stable spatial statistical tool that reliably quantifies the spatial variability.

Calculating semi-variograms by hand or spreadsheet can be rather tedious. It is commonly done in computational or statistical programs like MATLAB and R.

10.2 Theoretical Spatial Distributions

Theoretical models are commonly fit to the empirical results to aid in interpreting and forecasting spatial dependence. Some typical "basic" models that are used to fit

spatial data include: linear, exponential, spherical, and Gaussian. Whether using these or other models the mathematical functions must be positive definite to guarantee that the estimations exist are unique and are stable.

Figure 10.6 shows these three models. The separation distance at which the model reaches a plateau is called the range (a). The semi-variance at the range value is called the sill (c). All three models shown are for spatial data that exhibit a plateau.

The equation for a spherical model is:

$$\gamma(\boldsymbol{h}) = \begin{cases} c\left(\frac{3\boldsymbol{h}}{2a} - \frac{\boldsymbol{h}^3}{2a^3}\right) & \text{when } \boldsymbol{h} \leq a \\ c & \text{when } \boldsymbol{h} > a \end{cases} \tag{10.2}$$

The equation for an exponential model is:

$$\gamma(\boldsymbol{h}) = c\left[1 - \exp\left(-\frac{3\boldsymbol{h}}{a}\right)\right] \tag{10.3}$$

The equation for the Gaussian model is:

$$\gamma(\boldsymbol{h}) = 1 - \exp\left(-\frac{3\boldsymbol{h}^2}{a^2}\right) \tag{10.4}$$

Figure 10.7 shows an example of spherical and Gaussian models fit to some spatial variability data. Least squares regression is typically used to fit the theoretical models to the empirical data, with R-squared being the common metric for the goodness of fit.

Note that both models start at a nonzero value at short separation distances and they both approach a plateau at long separation distances. The nonzero start has to do with measurement uncertainty at a point and is called the nugget in geostatistics literature. The plateau or range indicates the separation distance at which data pairs are maximally dissimilar. The range is important for lifeline systems analysis because it shows how far apart things need to be from each other to achieve spatial

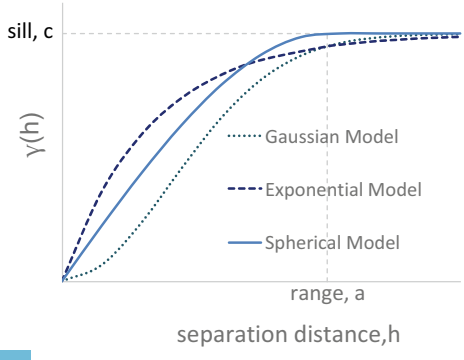


Fig. 10.6 Schematic showing three "basic" theoretical models used to fit spatial variability data with a plateau

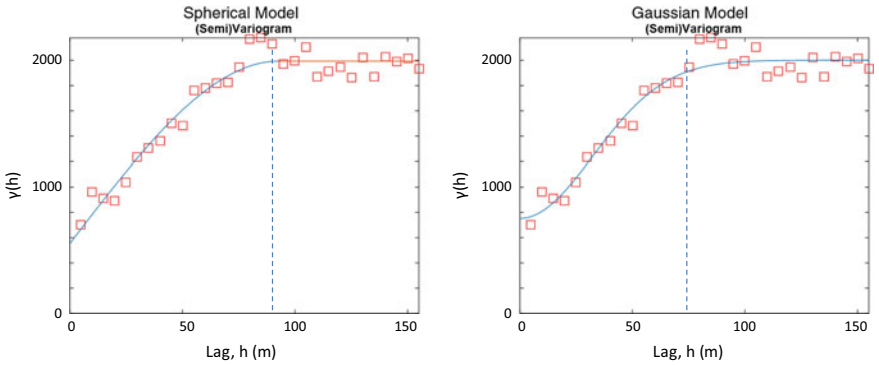


Fig. 10.7 Spherical and Gaussian models fit to semi-variograms of repeated shear wave velocity measurements over a 300 m stretch of quaternary alluvium (after Moss and Wagstaffe 2019)

independence. If we normalize the semi-variogram to create a general relative semi-variogram, then we can see in Fig. 10.8 that the nugget is the squared coefficient of variation $\left(\delta^2\right)$ from point estimate uncertainty which is the nonspatial uncertainty addressed in the previous chapters of this book.

$$\gamma_{\mathrm{GR}}(h) = \frac{\gamma(h)}{\bar{x}(h)^2} \tag{10.5}$$

In Eq. (10.5) the denominator $\bar{x}(h)^2$ is the squared mean of all values used to calculate the semi-variogram. Alternatively it can be the squared localized mean as well (Isaaks and Srivastava 1989).

To properly characterize the range Oliver and Webster (2015) recommend a sample interval that allows at least five estimates of $\gamma(h)$ before the range. Fewer than that can render the estimate of the range ambiguous. If the goal of mapping spatial variability is to define the range, then care must be taken to sample with sufficient spatial density to remove ambiguity.

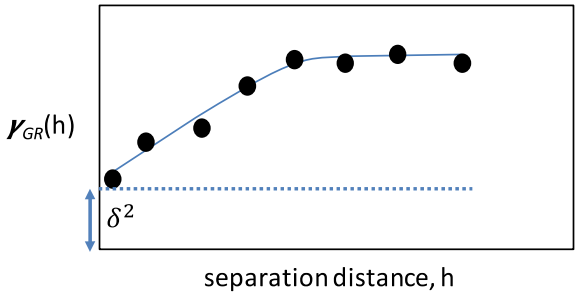
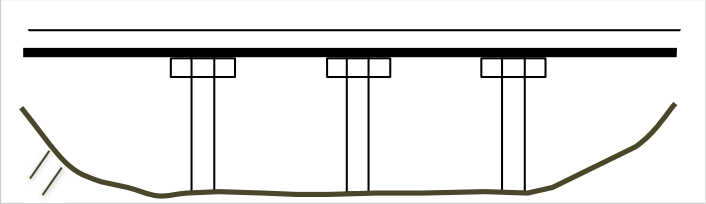


Fig. 10.8 Schematic of a general relative semi-variogram showing how a point estimate of uncertainty is connected to the spatial variability

Example: Bridge System Analysis (after Moss and Hollenback 2015)

To demonstrate the influence of spatial variability for non redundant lifelines, a bridge with three supports spanning an alluvial flood plain is used. The failure mode that the three supports will be analyzed for is foundation failure due to liquefaction. Typically for bridge support failure we would need to analyze the conditional probability of vertical or lateral displacement given liquefaction, but for this example failure will be simplified to only liquefaction initiation.



For liquefaction initiation the resistance is $R = \text{CRR}$, the cyclic resistance ratio, and the load is $Q = \text{CSR}$, the cyclic stress ratio. CRR is usually defined by a semi-empirical relationship between penetration resistance and CSR, and CSR is the seismic-induced shear stress needed to liquefy the soil normalized by the initial vertical effective stress (e.g., Moss et al. 2006; Yazdi and Moss 2016). Both CRR and CSR are a function of the stress conditions so they are not statistically independent, but for the purposes of exploring spatial correlation we will treat them as if they are independent. For this example we will also assume that the probability of failure for each bridge support due to liquefaction is 5%: $P(M_A) = P(M_B) = P(M_C) = 0.05$. This is the probability of failure for one component of the bridge system. Unimodal bounds of this series problem give 15 and 5% probability of system failure (Eq. 7.5).

As in the beginning of this chapter, for a three-component series system the probability of bridge failure is defined by the addition rule:

$$\begin{aligned} P(\text{Bridge}) &= P(\cup M_i) = P(M_1 \cup M_2 \cup M_3) \\ &= P(M_1) + P(M_2) + P(M_3) - P(M_1 M_2) - P(M_1 M_3) \\ &\quad - P(M_2 M_3) + P(M_1 M_2 M_3) \end{aligned}$$

Solving the problem assuming no spatial correlation, which means that the failure of each support is not dependent the others, the calculation would be carried out by applying the multiplication rule to each joint event:

$$\begin{aligned} P(\text{bridge failure}) &= P(M_A) + P(M_B) + P(M_C) - P(M_A)P(M_B) \\ &\quad - P(M_A)P(M_C) - P(M_B)P(M_C) + P(M_A)P(M_B)P(M_C) \end{aligned}$$

The probability of bridge failure for spatially uncorrelated supports is then:

$$P(\text{bridge failure}) = 3 \times 0.05 - 3 \times 0.05^2 + 0.05^3 = 0.1426$$

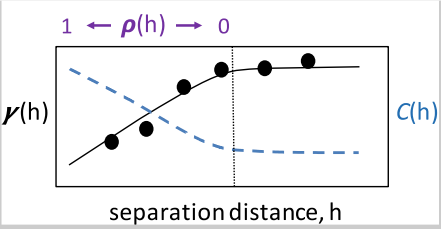
For perfectly spatially correlated supports all the conditional probabilities are 1.0 and the multiplication rule for each joint event results in the following:

$$\begin{aligned} P(\text{bridge failure}) &= P(M_A) + P(M_B) + P(M_C) - P(M_A|M_B)P(M_B) \\ &\quad - P(M_A|M_C)P(M_C) - P(M_B|M_C)P(M_C) \\ &\quad + P(M_A|M_B M_C)P(M_B|M_C)P(M_C) \\ P(\text{bridge failure}) &= 3(0.05) - 3(1)0.05 + 1(1)0.05 = 0.05 \end{aligned}$$

A probability of bridge failure spread of 14–5% may be too ambiguous for an engineering decision.

Now this is where the spatial correlation of load and resistance comes into play. The spatial correlation of penetration resistance has been quantified in different situations and is typically in the range of 10s–100s of meters (e.g., Moss et al. 2010; Huber 2013). The spatial correlation of strong ground shaking that would cause liquefaction is a function of the frequency and/or the phase of the motion of interest and varies from 10s to 1000s of meters (e.g., Jayaram and Baker 2009). For liquefaction the phase is not necessarily as important as the amplitude at a frequency that corresponds to the resonant frequency of the site, and the spatial correlation length of amplitude is generally in the 100s to a few 1000 m range.

If will treat the random variables and subsequently the performance functions as jointly normal then the conditional probability $P(M_h|M_{h+1})$ and the correlation coefficient $\rho(h)$ are identical. The spatial variability for our particular bridge problem using a semi-variogram may look like the figure below. Let's say that the range for both load and resistance comes out to roughly 100 m (not an unreasonable for this type of problem). The plot shows the semi-variogram, which is the complement of the covariance function, which when normalized becomes the correlation function; all three can be represented on a single plot. At the range (shown on the plot as the vertical dashed line) the semi-variance is at a maximum, the covariance a minimum, and the spatial correlation is approximately zero; $\rho(h) \cong 0$. At zero separation distance there is perfect spatial correlation; $\rho(h) \cong 1$.



If the bridge supports in this problem are 50 m apart, then we are a half of the way to the range of 100 m and the spatial correlation coefficient, and hence, the conditional probabilities are:

$$P(M_A|M_B) = P(M_B|M_C) = 0.50$$
$$P(M_A|M_C) = 0$$

A lag distance of 50 m gives a conditional probability of 0.50 (half way from 1 to 0), and a lag distance of 100 m gives a conditional probability of 0. We need to make an assumption about the triple-conditional probability. It would be reasonable to assume that it is at least identical to the double-conditional probability of 0.50 but could be slightly higher empirically.

$$(M_A|M_B M_C) = 0.50$$

From the spatial variability mapped above, and the resulting conditional probabilities, we see that the probability of bridge failure is then:

$$P(\text{bridge failure}) = 3(0.05) - 2(0.50)(0.05) - 0(0.05) + 0.50(0.50)(0.05) = 0.1125$$

This ~11% system probability failure is much more precise and actionable than the unimodal bounds that gave a range of 15–5% probability of failure, and the perfectly spatially uncorrelated and correlated results of ~14 and 5% probability of failure.

In general for series systems the inclusion of spatial correlation results in a decreased system probability of failure from the no spatial correlation or neglected spatial correlation assumption. This, in and of itself, is a compelling reason for including spatial correlation in systems analysis.

10.3 Chapter Summary

- Uncertainty as a function of distance is termed spatial variability.
- This spatial variability can be quantified statistically most easily using a semi-variogram.
- The spatial relationship can be modeled using different theoretical functions; Gaussian, spherical, exponential.
- At zero separation distance we see perfect spatial correlation, and at the range or maximum semi-variance we see zero spatial correlation.
- Quantifying spatial variability in this manner can lead to accurate estimates of system probability of failure for lifelines and other long linear engineered facilities.

References

Cressie, N. (1992). Statistics for spatial data. *Terra Nova, 4*(5), 613–617.

Huber, M. (2013). *Soil variability and its consequences in geotechnical engineering* (Ph.D. dissertation). Universität Stuttgart, Germany.

Isaaks, E., & Srivastava, R. (1989). *An introduction to applied geostatistics*. New York: Oxford University Press.

Jayaram, N., & Baker, J. W. (2009). Correlation model for spatially distributed ground-motion intensities. *Earthquake Engineering & Structural Dynamics, 38*(15), 1687–1708.

Matheron, G. (1965). Principles of geostatistics. *Economic Geology, 58,* 1246–1266.

Moss, R. E., Seed, R. B., Kayen, R. E., Stewart, J. P., Der Kiureghian, A., & Cetin, K. O. (2006). CPT-based probabilistic and deterministic assessment of in situ seismic soil liquefaction potential. *Journal of Geotechnical and Geoenvironmental Engineering, 132*(8), 1032–1051.

Moss, R. E. S., & Hollenback, J. C. (2015, November). The influence of spatial variability on lifelines probability of failure. In *15th Pan-American Conference on Soil Mechanics and Geotechnical Engineering*, Buenos Aires, Argentina.

Moss, R. E. S., Hollenback, J. C., & Ng, J. (2010, May). Spatial variability of levees as measured using the CPT. In *2nd International Symposium on Cone Penetration Testing*, Huntington Beach, CA.

Moss, R. E. S., & Wagstaffe, D. (2019). Spatial variability of shear stiffness and its effects on lifelines in quaternary alluvium. *Soil Dynamics and Earthquake Engineering*, submitted under review.

Oliver, M., & Webster, R. (2015). *Basic steps in geostatistics: The variogram and kriging*. Springer.

Webster, R., & Oliver, M. A. (1992). Sample adequately to estimate variograms of soil properties. *Journal of Soil Science, 40,* 493–496.

Yazdi, J. S., & Moss, R. E. S. (2016). Nonparametric liquefaction triggering and postliquefaction deformations. *Journal of Geotechnical and Geoenvironmental Engineering, 143*(3), 04016105.

Chapter 11
Bayesian Updating

Bayes' theorem (or Bayes' rule) is frequently used as a means of estimating and updating probability given incomplete information. There are many forms this updating can take, and it has been applied to many problems in data science, engineering, astronomy, economics, biology, sociology, and many other disciplines. This chapter provides a cursory overview of some updating techniques and applications. There are many other texts that delve deep into the Bayesian approach that readers are encouraged to explore (e.g., Gelman et al. 2013).

Bayes' theorem, as introduced in Chap. 3, solves for the conditional probability by combining prior knowledge to update the existing knowledge.

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)} \tag{11.1}$$

Knowing the marginal probabilities of A and B and their interrelationship, we can then determine an updated relationship. The best way to illustrate this is by example.

Example: Landslide Example

Southern California hill slopes are prone to landslides (L) when the rainfall intensity (R) is high. In previous years, data collected for a particular aspect of the San Gabriel Mountains has found that for any given 24 h period during the rainy season the probability of a landslide $P(L)$ is 0.001 (0.1%). When a landslide does occur the probability of rainfall intensity being greater than 10 cm in 24 h is 0.30 (30%), this is $P(R|L)$. Rainfall data averaged over many rainy seasons shows that the probability of 24 h rainfall greater than 10 cm during the rainy season $P(R)$ is 0.005 (0.5%). So if we observe a storm that brings greater than 10 cm of rain, what is the probability of a landslide $P(L|R)$?

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0_11

$$P(L|R) = \frac{P(R|L)P(L)}{P(R)} = \frac{0.30 * 0.001}{0.005} = 0.06$$

Bayes' theorem has been used here to update the probability of landsliding $P(L)$ given information on a conditional effect $P(R|L)$. The marginal probability of landsliding is 0.1% but the conditional probability of landsliding given intense rainfall is 6%.

This can equally be applied to probability distributions. Let's take a variable that can be represented by a probability distribution with its central tendency and dispersion. As new information on the variable becomes available (e.g., more measurements or observations) we revise the estimate by updating the probability distribution, thereby refining the estimate and increasing our knowledge. Figure 11.1 shows the initial and revised estimates of a lognormally distributed variable given new information. Both the initial and revised estimates have the same median, but the dispersion is much reduced in the revised estimate.

Expanding Bayes' theorem to probability distributions and any number of observations (Ang and Tang 2007):

$$P(\Theta = \theta_i|\varepsilon) = \frac{P(\varepsilon|\Theta = \theta_i)P(\Theta = \theta_i)}{\sum_{i=1}^{k} P(\varepsilon|\Theta = \theta_i)P(\Theta = \theta_i)} i = 1, 2, \ldots, k \qquad (11.2)$$

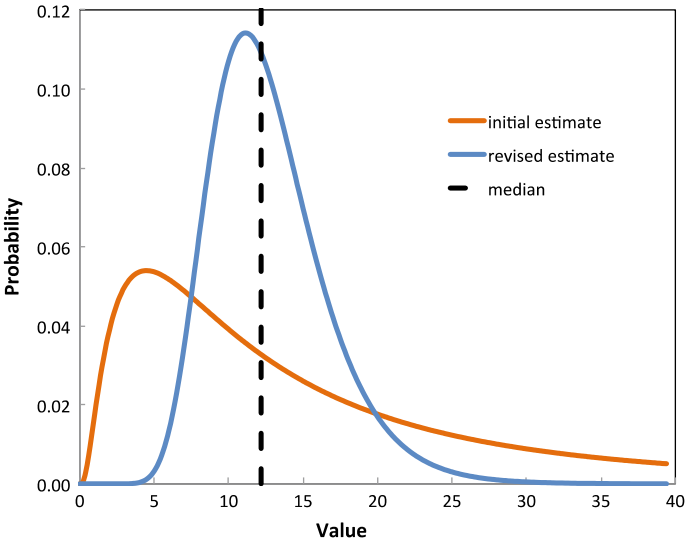


Fig. 11.1 Initial and revised (updated) lognormal distributions of some variable with identical median values. The revised or updated distribution has a reduced uncertainty

In the equation above θ is our parameter of interest and can take on a set of discrete values $i = 1, 2, \ldots, k$. The symbol Θ is the random variable that represents all possible values of parameter θ. The symbol ε represents an observed outcome from a test or measurement. As information becomes available the prior assumptions on the parameter may be modified or updated. The terms in Eq. 11.2 can be described as follows:

- $P(\varepsilon|\Theta = \theta_i)$ is called the likelihood which is the conditional probability of a particular observed outcome,
- $P(\Theta = \theta_i)$ is called the prior which is the marginal probability prior to observing an outcome,
- $P(\Theta = \theta_i|\varepsilon)$ is called the posterior which is the conditional probability that has been updated or revised given the observation.
- The denominator reduces to a constant by summing over the reduced sample space of the observation and is called the normalizing constant.

When using continuous probability distributions, Bayes' rule for updating can be rewritten in a concise form as shown below (Box and Tiao 1992):

$$f(\theta) = c \cdot L(\theta) \cdot p(\theta) \tag{11.3}$$

where $f(\theta)$ is the posterior distribution, $L(\theta)$ is the likelihood function, $p(\theta)$ is the prior, and c is the normalizing constant calculated as:

$$c = \left[\int_{-\infty}^{+\infty} L(\theta) \cdot p(\theta) \cdot \mathrm{d}\theta \right]^{-1} \tag{11.4}$$

If the experimental outcome is an observed set of values (x_1, x_2, …, x_n) that represent a random sample from a population (X) with an underlying PDF ($f(x)$), the probability of observing this particular set of values is the likelihood function with a distribution parameter (θ):

$$L(\theta) = \prod_{i=1}^{n} f(x_i|\theta) \tag{11.5}$$

It should be noted that this likelihood function is identical to that derived using the classical statistics approach of maximum likelihood estimation (Ang and Tang 2007).

For a Gaussian population with known standard deviation, the likelihood function for the mean is:

$$L(\mu) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2 \right] \tag{11.6}$$

It can be shown (e.g., Tang 1971) that the product of normal PDFs with their respective means and standard deviations is also a normal PDF with its sample mean and standard deviation:

$$L(\mu) = N_\mu\left(\bar{x}, \frac{\sigma}{\sqrt{n}}\right) \tag{11.7}$$

Note that the standard deviation divided by the square root of the number of samples is often called the standard error.

If measurement error (epistemic) is quantified it can be combined with the inherent variability (aleatory) using the total probability theorem to capture the total uncertainty:

$$L(\mu) = N_\mu\left(\bar{x}, \sqrt{\sigma_e^2 + \sigma_a^2/n}\right) \tag{11.8}$$

Here the subscript e is for epistemic and a for aleatory uncertainty. In this way we have a comprehensive measurement of the uncertainty present in the problem.

11.1 Conjugate Distribution

The normalizing constant (Eq. 11.4) is often the difficult part of the updating process, and it is common to numerically solve this integral, particularly with sequential updating problems. One simple solution is an analytical method called conjugate distributions which provides closed-form solutions for the posterior if the likelihood and prior are of specific paired distributions (And and Tang 1984; Der Kiureghian 2001). If we assume we have a Gaussian likelihood, then this can be conjugate with a Gaussian prior and posterior.

$$\text{If } p(\theta) = N(\mu, \sigma) \text{ and } L(\theta) = N\left(\mu', \sigma'\right) \text{ then } f(\theta) = N\left(\mu'', \sigma''\right) \tag{11.9}$$

Equation 11.9 says that if the prior and likelihood are normally distributed with their respective means and standard deviations, then the posterior is normally distributed with its mean and standard deviation. Here we are treating the standard deviations as known and are solving for the posterior mean. For a problem of sequential observations the recursive posterior statistics can be found with the following equations (Der Kiureghian 2001):

$$\mu'' = \frac{\mu'/\sigma'2 + \mu/\sigma^2}{1/\sigma'2 + 1/\sigma^2} \text{ and } \sigma'' = \sqrt{\frac{1}{1/\sigma'2 + 1/\sigma^2}} \tag{11.10}$$

Other conjugate priors exist and can be found in Ang and Tang (2007) and other references. The Gaussian is the only recursive conjugate prior that can be used for repeated observations because it always results in a Gaussian posterior.

Example: Normal Conjugates Example

In this example we are evaluating an underwater site for load-bearing foundation purposes [a top secret installation]. No subsurface information exists other than the soil is probably clayey; therefore, the prior distribution of undrained strength of clayey soil ($c = s_u$) could be within a large range from as low as 10 kPa up to 160 kPa or higher (Phoon and Kulhawy 1999). In assuming normal conjugates this means we must use a symmetric distribution, so we take the range of 10–160 and divide by 6 to give three standard deviations on either side of the mean. The estimated statistics we arrive at are $\mu = 85$ and $\sigma = 25$.

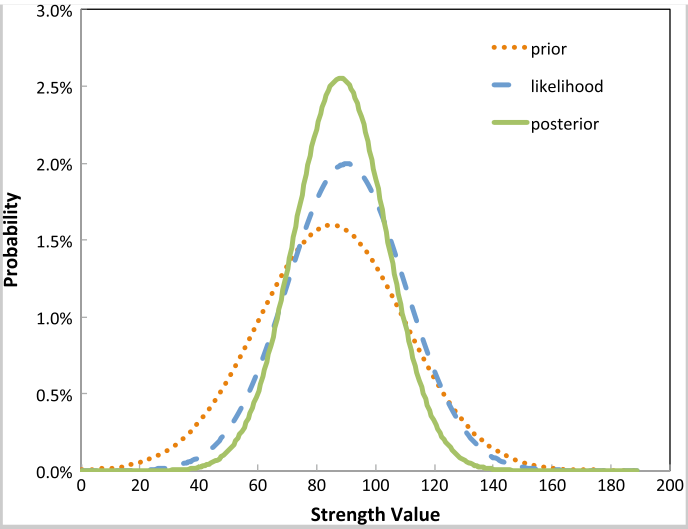
To improve this subjective estimate we consult a surficial geology map for this underwater location and find that the proposed foundation location plots within a unit of lacustrine clay. Based on our previous experience with this surficial geologic unit we tighten the bound by ascribing the likelihood with the statistics of $\mu = 90$ and $\sigma = 15$.

Running these numbers through Eq. 11.6 gives us:

$$\mu'' = \frac{90/15^2 + 85/25^2}{1/15^2 + 1/25^2} = 88.1 \text{ kPa}$$

$$\sigma'' = \sqrt{\frac{1}{1/15^2 + 1/25^2}} = 15.6 \text{ kPa}$$

Plotting these results we see that making an observation (i.e., determining the surficial geologic unit) has resulted in reduced uncertainty for the undrained strength.



Example: Bridge Survey (after Ang and Tang 2007)

The elevation measurements of the top of a bridge bent are important for proper bridge alignment. Survey is a precise tool; however, there is uncertainty that exists in the process due to temperature fluctuations, air quality, operator variability, equipment variability, etc. For a particular bent one surveying company made five measurements during construction relative to a project benchmark, 20.45, 20.38, 20.51, 20.42, 20.46 m. It has been determined through years of analyzing survey data that the typical COV of the inherent variability is 0.4%. Assume that the measurement error for the surveying equipment is normal with a zero mean and standard deviation of 0.01. There was also a prior survey performed by a different company that provided an estimate 20.42 m and a typical standard error of 0.020 m.

What is a likelihood estimate of the actual elevation based on the first survey? What is an updated estimate of the actual elevation given the information from the prior survey?

The sample mean and known standard deviation based on the first company's surveying is:

$$\bar{x} = 20.444 \text{ m}$$
$$\sigma = 0.082 \text{ m}$$

Including the measurement uncertainty using Eq. 11.8:

$$\sigma_{\text{total}} = \sqrt{0.010^2 + 0.082^2/5} = 0.038$$

The estimate by the first company following Eq. 11.7 is then $N(20.444,0.038)$ m.

Now to incorporate the prior information from the other company we use conjugate distributions to update the Gaussian PDF. The total uncertainty for the other company is:

$$\sigma_{\text{total}} = \sqrt{0.010^2 + 0.020^2} = 0.022$$

Solving for the posterior mean and standard deviation using Eq. 11.10:

$$\mu'' = \frac{(20.420)(0.038)^2 + (20.444)(0.022)^2}{(0.038)^2 + (0.022)^2} = 20.587 \text{ m}$$

$$\sigma'' = \sqrt{\frac{(0.038)^2(0.022)^2}{(0.038)^2 + (0.022)^2}} = 0.023 \text{ m}$$

The updated estimate gives us a more robust mean and a reduced uncertainty because there is more information informing the estimate. We can continue this updating process with each measurement to resolve a more and more accurate estimate of the mean value.

11.2 Bayesian Search

Application of Bayesian updating is often found in sequential form like what is used for search and rescue operations by the US Coast Guard, the US Air Force, US Navy, and others (Stone 1975). This type of application is often termed Bayesian search theory and has proven very successful in locating overboard seamen, lost munitions, sunken submarines, and other missing items.

If we are searching for say a missing submarine, we start with all the information on the submarine's last known position. We combine that with information such as the last known heading and speed along with knowledge of ocean currents, water temperatures, wind patterns, water depth, ocean bottom bathymetry, and other pertinent known or estimated variables. All this is used to form an estimate of the submarine's current location, which is inherently uncertain and becomes the prior distribution. We

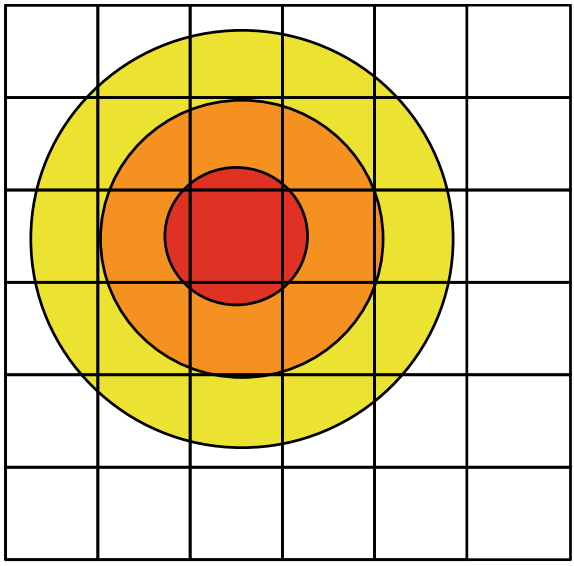


Fig. 11.2 Prior probability distribution of location of a lost item (e.g., submarine) with a georeferenced search grid overlayed on the area

then overlay a georeferenced search grid on this prior distribution and focus search efforts on the most likely grid square first (Fig. 11.2).

Search operations are imperfect because speed and time are critical; thoroughness is sacrificed for expediency when lives or money are at stake. The likelihood of locating a lost item within a grid square if that item is in fact there is less than 100% because of the many factors that can diminish the accuracy of the search (e.g., poor weather, low visibility, large search area, search speed/altitude, etc.). For example the probability of a search team detecting an overboard crewman in a grid square if they are in that grid square is around 80% based on forensic studies by the Coast Guard.

If a team searches a particular grid square and does not locate the lost item then the probability for the grid square can be updated using Bayes' rule. The grid squares that were not searched can also be updated given the new information. The posterior probabilities of the searched and not searched grid squares can then be used to develop an updated distribution of the most likely area where the lost item is located. This is repeated until the lost item is found.

In each grid square if the prior probabilty of the lost item being located there is p, and the probability of detecting the lost item if it is in fact there is q, then we can rewrite Eq. 11.1 for each grid square. If $P(B)$ the prior is p, $P(A|B)$ the likelihood of not detecting the lost item in the grid square is $(1 - q)$ the complement of finding it, and $P(A)$ the sample space is $((1 - p) + p(1 - q))$, then $P(B|A)$ the posterior p'_s can be found as:

$$p'_s = \frac{p(1-q)}{(1-p)+p(1-q)} = p\frac{1-q}{1-pq} < p \tag{11.11}$$

For the grid squares not searched the posterior p'_n becomes:

$$p'_n = p\frac{1}{1-pq} > p \tag{11.12}$$

The numerator in Eq. 11.12 is 1 because the probability of not finding the lost item when no search is conducted in that grid square is 100%.

A benefit of this type of search method is that by using probability updating along with search team costs, the likelihood, and cost of finding the lost item can be estimated giving a risk estimate for the search and rescue (e.g., after five days the probability of success will be 95% with a total cost of $1.2 USD).

In search and rescue the first hurdle is in creating a prior that is as informative as possible by incorporating all available information, which is a Bayesian approach in and of itself. The second hurdle is mobilizing search teams that can cover the prior distribution region as efficiently as possible, maximizing the likelihood of detection while minimizing time and cost of the search. The updating process itself is straight forward and only requires the sequential updating of information as it is reported back from the search teams. The following example demonstrates this process.

Example: Search and Rescue

The following grid shows the prior probability estimates of the location of a shipmate who was swept overboard in rough waters. Note the total probability of the grid squares sums to 1.0; 4 by 0.10 plus 12 by 0.05.

	A	B	C	D
1	0.05	0.05	0.05	0.05
2	0.05	0.10	0.10	0.05
3	0.05	0.10	0.10	0.05
4	0.05	0.05	0.05	0.05

The search team investigates square $B2$ and does not find the overboard shipmate. If this search team is expected to have 80% probability of detecting a lost person when one is present then the updated probability of that particular grid square is:

$$p_s' = 0.10\frac{1-0.80}{1-(0.10)0.80} = 0.022$$

For the other more likely grid squares that were not searched:

$$p_n' = 0.10\frac{1}{1-(0.10)0.80} = 0.109$$

And for the less likely grid squares that were not searched:

$$p_n' = 0.05\frac{1}{1-(0.05)0.80} = 0.052$$

We see a slight increase in the probability of the unsearched grid squares and a large decrease in the probability of the grid square that was searched. The revised grid shows the updated probabilities and refocuses the search team efforts on the region of the high likelihood grid squares.

	A	B	C	D
1	0.052	0.052	0.052	0.052
2	0.052	0.022	0.109	0.052
3	0.052	0.109	0.109	0.052
4	0.052	0.052	0.052	0.052

Some keys to real search operations are discretizing the area as finely as possible while considering the search team(s) capacity, developing the most informative prior possible, and searching in an efficient and rapid manner.

11.3 Chapter Summary

- Bayesian updating provides a formal means of incorporating prior information to improve an estimate of a distribution.
- Updating can be performed one time or sequentially depending on the nature of the problem and the availability of the information.
- Conjugate distributions provide a closed-form solution for Bayesian updating in certain situations. For a normal likelihood distribution and a normal prior distribution, the posterior distribution is also normal.

References

Ang, A. H.-S., & Tang, W. H. (2007). *Probability concepts in engineering: emphasis on applications to civil and environmental engineering* (2nd ed.). New York: Wiley.

Ang, A. H-S., & Tang, W. H. (1984). *Probability concepts in engineering planning and design: Volume II-decision, risk, and reliability*. Published by the authors.

Box, G. E. P., & Tiao, G. C. (1992). *Bayesian inference in statistical analysis*. London: Wiley.

Der Kiureghian, A. (2001). *Structural reliability course notes*. UC Berkeley.

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. Chapman and Hall/CRC.

Phoon, K. K., & Kulhawy, F. H. (1999). Evaluation of geotechnical property variabaility. *Canadian Geotechnical Journal, 36,* 625–639.

Stone, L. D. (1975). *The theory of optimal search*. Operations Research Society of America.

Tang, W. H. (1971). A Bayesian evaluation of information for foundation engineering design. In *Geotechnical safety and reliability* (pp. 214–214).

Appendix A
Standard Normal Probability Tables

There is no closed-form integral for the normal distribution. Because it is so frequently used, approximations of the CDF are commonly found in most computational devices. The following pages provide tabular results of the standard normal cumulative distribution function, $N(0,1)$. In the tables, x is the value of interest and $\Phi(x)$ is the integral of the standard normal PDF at that value. The x represents the number of standard deviations away from the mean the value is located, so a value of 1.00 is approximately the 84th percentile or 84% probability, the value 2.00 is approximately the 98th percentile or 98% probability, and so on.

The standard normal CDF can also be approximated using many equations that can be found in the literature. A simple function based on a Taylor series expansion is shown below (Marsaglia 2004) with its accuracy a function of the number of terms used in the expansion:

$$\Phi(x) \cong \frac{1}{2} + \phi(x)\left(x + \frac{x^3}{3} + \frac{x^5}{3\times 5} + \frac{x^7}{3\times 5\times 7} + \frac{x^9}{3\times 5\times 7\times 9} + \cdots\right)$$

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.77	0.7794	1.54	0.9382	2.31	0.9896
0.01	0.5040	0.78	0.7823	1.55	0.9394	2.32	0.9898
0.02	0.5080	0.79	0.7852	1.56	0.9406	2.33	0.9901
0.03	0.5120	0.80	0.7881	1.57	0.9418	2.34	0.9904
0.04	0.5160	0.81	0.7910	1.58	0.9429	2.35	0.9906
0.05	0.5199	0.82	0.7939	1.59	0.9441	2.36	0.9909
0.06	0.5239	0.83	0.7967	1.60	0.9452	2.37	0.9911
0.07	0.5279	0.84	0.7995	1.61	0.9463	2.38	0.9913
0.08	0.5319	0.85	0.8023	1.62	0.9474	2.39	0.9916
0.09	0.5359	0.86	0.8051	1.63	0.9484	2.40	0.9918
0.10	0.5398	0.87	0.8078	1.64	0.9495	2.41	0.9920

(continued)

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0

(continued)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.11	0.5438	0.88	0.8106	1.65	0.9505	2.42	0.9922
0.12	0.5478	0.89	0.8133	1.66	0.9515	2.43	0.9925
0.13	0.5517	0.90	0.8159	1.67	0.9525	2.44	0.9927
0.14	0.5557	0.91	0.8186	1.68	0.9535	2.45	0.9929
0.15	0.5596	0.92	0.8212	1.69	0.9545	2.46	0.9931
0.16	0.5636	0.93	0.8238	1.70	0.9554	2.47	0.9932
0.17	0.5675	0.94	0.8264	1.71	0.9564	2.48	0.9934
0.18	0.5714	0.95	0.8289	1.72	0.9573	2.49	0.9936
0.19	0.5753	0.96	0.8315	1.73	0.9582	2.50	0.9938
0.20	0.5793	0.97	0.8340	1.74	0.9591	2.51	0.9940
0.21	0.5832	0.98	0.8365	1.75	0.9599	2.52	0.9941
0.22	0.5871	0.99	0.8389	1.76	0.9608	2.53	0.9943
0.23	0.5910	1.00	0.8413	1.77	0.9616	2.54	0.9945
0.24	0.5948	1.01	0.8438	1.78	0.9625	2.55	0.9946
0.25	0.5987	1.02	0.8461	1.79	0.9633	2.56	0.9948
0.26	0.6026	1.03	0.8485	1.80	0.9641	2.57	0.9949
0.27	0.6064	1.04	0.8508	1.81	0.9649	2.58	0.9951
0.28	0.6103	1.05	0.8531	1.82	0.9656	2.59	0.9952
0.29	0.6141	1.06	0.8554	1.83	0.9664	2.60	0.9953
0.30	0.6179	1.07	0.8577	1.84	0.9671	2.61	0.9955
0.31	0.6217	1.08	0.8599	1.85	0.9678	2.62	0.9956
0.32	0.6255	1.09	0.8621	1.86	0.9686	2.63	0.9957
0.33	0.6293	1.10	0.8643	1.87	0.9693	2.64	0.9959
0.34	0.6331	1.11	0.8665	1.88	0.9699	2.65	0.9960
0.35	0.6368	1.12	0.8686	1.89	0.9706	2.66	0.9961
0.36	0.6406	1.13	0.8708	1.90	0.9713	2.67	0.9962
0.37	0.6443	1.14	0.8729	1.91	0.9719	2.68	0.9963
0.38	0.6480	1.15	0.8749	1.92	0.9726	2.69	0.9964
0.39	0.6517	1.16	0.8770	1.93	0.9732	2.70	0.9965
0.40	0.6554	1.17	0.8790	1.94	0.9738	2.71	0.9966
0.41	0.6591	1.18	0.8810	1.95	0.9744	2.72	0.9967
0.42	0.6628	1.19	0.8830	1.96	0.9750	2.73	0.9968
0.43	0.6664	1.20	0.8849	1.97	0.9756	2.74	0.9969
0.44	0.6700	1.21	0.8869	1.98	0.9761	2.75	0.9970
0.45	0.6736	1.22	0.8888	1.99	0.9767	2.76	0.9971
0.46	0.6772	1.23	0.8907	2.00	0.9772	2.77	0.9972
0.47	0.6808	1.24	0.8925	2.01	0.9778	2.78	0.9973
0.48	0.6844	1.25	0.8944	2.02	0.9783	2.79	0.9974
0.49	0.6879	1.26	0.8962	2.03	0.9788	2.80	0.9974
0.50	0.6915	1.27	0.8980	2.04	0.9793	2.81	0.9975

(continued)

(continued)

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.51	0.6950	1.28	0.8997	2.05	0.9798	2.82	0.9976
0.52	0.6985	1.29	0.9015	2.06	0.9803	2.83	0.9977
0.53	0.7019	1.30	0.9032	2.07	0.9808	2.84	0.9977
0.54	0.7054	1.31	0.9049	2.08	0.9812	2.85	0.9978
0.55	0.7088	1.32	0.9066	2.09	0.9817	2.86	0.9979
0.56	0.7123	1.33	0.9082	2.10	0.9821	2.87	0.9979
0.57	0.7157	1.34	0.9099	2.11	0.9826	2.88	0.9980
0.58	0.7190	1.35	0.9115	2.12	0.9830	2.89	0.9981
0.59	0.7224	1.36	0.9131	2.13	0.9834	2.90	0.9981
0.60	0.7257	1.37	0.9147	2.14	0.9838	2.91	0.9982
0.61	0.7291	1.38	0.9162	2.15	0.9842	2.92	0.9982
0.62	0.7324	1.39	0.9177	2.16	0.9846	2.93	0.9983
0.63	0.7357	1.40	0.9192	2.17	0.9850	2.94	0.9984
0.64	0.7389	1.41	0.9207	2.18	0.9854	2.95	0.9984
0.65	0.7422	1.42	0.9222	2.19	0.9857	2.96	0.9985
0.66	0.7454	1.43	0.9236	2.20	0.9861	2.97	0.9985
0.67	0.7486	1.44	0.9251	2.21	0.9864	2.98	0.9986
0.68	0.7517	1.45	0.9265	2.22	0.9868	2.99	0.9986
0.69	0.7549	1.46	0.9279	2.23	0.9871	3.00	0.9987
0.70	0.7580	1.47	0.9292	2.24	0.9875	…	…
0.71	0.7611	1.48	0.9306	2.25	0.9878	3.50	0.99977
0.72	0.7642	1.49	0.9319	2.26	0.9881	…	…
0.73	0.7673	1.50	0.9332	2.27	0.9884	4.00	0.99997
0.74	0.7704	1.51	0.9345	2.28	0.9887		
0.75	0.7734	1.52	0.9357	2.29	0.9890		
0.76	0.7764	1.53	0.9370	2.30	0.9893		

Reference

Marsaglia, G. (2004). Evaluating the normal distribution. *Journal of Statistical Software, 11*(4).

Appendix B
Function Mean and Variance Proof

The following derivation is for the moments (mean and variance) of a function of random variables.

Take a function of two continuous random variables:

$$Z = g(X, Y)$$

The CDF of Z is then the integral of the joint distribution of X and Y considering the functional relationship of the variables:

$$F(z) = \iint_{g(x,y)\le z} f(x,y)\mathrm{d}x\mathrm{d}y = \int_{-\infty}^{\infty}\int_{-\infty}^{g^{-1}(z,y)} f(x,y)\mathrm{d}x\mathrm{d}y$$

To include the functional relationship we solve the function for x, $g^{-1} = g^{-1}(z, y)$ and change the variable of integration from x to z:

$$F(z) = \int_{-\infty}^{\infty}\int_{-\infty}^{z} f\left(g^{-1}, y\right)\left|\frac{\partial g^{-1}}{\partial z}\right|\mathrm{d}z\mathrm{d}y$$

To find the PDF of Z we take the derivative of the CDF with respect to z:

$$f(z) = \int_{-\infty}^{\infty} f\left(g^{-1}, y\right)\left|\frac{\partial g^{-1}}{\partial z}\right|\mathrm{d}y$$

Alternately we could solve the functional form for y instead of x, $g^{-1} = g^{-1}(x, z)$ and find:

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0

$$f(z) = \int_{-\infty}^{\infty} f\left(x, g^{-1}\right) \left| \frac{\partial g^{-1}}{\partial z} \right| \mathrm{d}x$$

The above results are applicable to any function with random variables of any distribution. Solving for this PDF (i.e., propagating the error from the independent variables to the dependent variable) can be simple in certain situations as pointed out in Chap. 5 or can be difficult when we have a function that is difficult to find the inverse of and/or when we have random variables with distributions that are intractable.

In the case where our function is a sum, $Z = X + Y$, we find:

$$x = z - y$$

$$\frac{\partial g^{-1}}{\partial z} = \frac{\partial x}{\partial z} = 1$$

Resulting in the PDF:

$$f(z) = \int_{-\infty}^{\infty} f(z - y, y) \mathrm{d}y$$

or alternatively:

$$f(z) = \int_{-\infty}^{\infty} f(x, z - x) \mathrm{d}x$$

If X and Y are statistically independent, then we can evaluate using their marginal distributions:

$$f(z) = \int_{-\infty}^{\infty} f(z - y) f(y) \mathrm{d}y$$

or alternatively:

$$f(z) = \int_{-\infty}^{\infty} f(x) f(z - x) \mathrm{d}x$$

Now if X and Y are Gaussian (i.e., normally distributed), the PDF becomes:

$$f(z) = \frac{1}{2\pi\sigma_X\sigma_Y}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{z-y-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2}\left(\frac{x-\mu_Y}{\sigma_Y}\right)^2\right]\mathrm{d}y$$

$$= \frac{1}{2\pi\sigma_X\sigma_Y}\exp\left[-\frac{1}{2}\left\{\left(\frac{\mu_Y}{\sigma_Y}\right)^2 + \left(\frac{z-\mu_X}{\sigma_X}\right)^2\right\}\right]\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(uy^2 - 2vy\right)\right]\mathrm{d}y$$

where $u = \frac{1}{\sigma_X^2} + \frac{1}{\sigma_Y^2}$ and $v = \frac{\mu_Y}{\sigma_Y^2} + \frac{z-\mu_X}{\sigma_X^2}$
substituting $w = y - \frac{v}{u}$ the integral becomes

$$\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(uy^2 - 2vy\right)\right]\mathrm{d}y$$

$$= \mathrm{e}^{v^2/2u}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}uw^2\right)\mathrm{d}w = \sqrt{2\pi/u}\,\exp\left(v^2/2u\right)$$

Reducing the algebra in the full distribution reveals:

$$f(z) = \frac{1}{\sqrt{2\pi(\sigma_X^2+\sigma_Y^2)}}\exp\left[-\frac{1}{2}\left(\frac{z-(\mu_X+\mu_Y)}{\sqrt{\sigma_X^2+\sigma_Y^2}}\right)^2\right]$$

which is the definition of the joint normal PDF with the moments:

$$\mu_Z = \mu_X + \mu_Y$$
$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$$

This can be solved for in a similar manner with correlated X and Y resulting in the moments:

$$\mu_Z = \mu_X + \mu_Y$$
$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y$$

If the function contains constants or coefficients in front of statistically independent variables, $Z = aX + bY$, the moments are then:

$$\mu_Z = a\mu_X + b\mu_Y$$
$$\sigma_Z^2 = a^2\sigma_X^2 + b^2\sigma_Y^2$$

Similarly we can evaluate a difference function, $Z = X - Y$, which is the same but with a negative coefficient in front of the second variable:

$$\mu_Z = \mu_X - \mu_Y$$
$$\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$$

This can be expanded to accommodate multiple random variables with their respective coefficients,

$$Z = \sum_{i=1}^{n} a_i X_i$$
$$\mu_Z = \sum_{i=1}^{n} a_i \mu_{X_i}$$
$$\sigma_Z^2 = \sum_{i=1}^{n} a_i^2 \sigma_{X_i}^2$$

The canonical form of these equations can be found in Chap. 5 in the discussion on Notation Clarity.

For a product function, $Z = XY$, where X and Y are lognormal with their respective moments λ and ξ, we can use the same derivation and find:

$$\ln Z = \ln X + \ln Y$$

$$\lambda_Z = \lambda_X + \lambda_Y$$

$\xi_Z^2 = \xi_X^2 + \xi_Y^2$ for statistically independent
$\xi_Z^2 = \xi_X^2 + \xi_Y^2 + 2\rho\xi_X\xi_Y$ with correlation

And similarly for a quotient function, $Z = X/Y$, where X and Y are lognormal with their respective moments λ and ξ:

$$\ln Z = \ln X - \ln Y$$
$$\lambda_Z = \lambda_X - \lambda_Y$$

$\xi_Z^2 = \xi_X^2 + \xi_Y^2$ for statistically independent
$\xi_Z^2 = \xi_X^2 + \xi_Y^2 - 2\rho\xi_X\xi_Y$ with correlation

They also can be easily expanded to multiple lognormal random variables with their respective coefficients.

Appendix C
FORM HLRF Algorithm

The following solution using FORM (first order reliability method) demonstrates the "improved" HLRF (Hasofer–Lind–Rackwitz Fiessler) algorithm as presented in Zhang and Der Kiureghian (1995). The solution uses matrix mathematics for efficient calculations. The matrices often used in this or similar solutions are shown below in reliability notation. These matrices, of course, can be expanded to any number of variables when doing multivariate limit state analysis.

$$\text{Mean vector} \quad M = \begin{bmatrix} \mu_R \\ \mu_Q \end{bmatrix}$$

$$\text{Covariance matrix} \quad \Sigma = \begin{bmatrix} \sigma_R^2 & \rho_{RQ}\sigma_R\sigma_Q \\ \rho_{RQ}\sigma_R\sigma_Q & \sigma_Q^2 \end{bmatrix}$$

$$\text{Standard deviation matrix} \quad D = \begin{bmatrix} \sigma_R & 0 \\ 0 & \sigma_Q \end{bmatrix}$$

$$\text{Correlation matrix} \quad R = \begin{bmatrix} 1 & \rho_{RQ} \\ \rho_{RQ} & 1 \end{bmatrix}$$

Shown in the following pages is the correlated and uncorrelated cut slope problem which can be compared with other solution methods shown in Chap. 6. This algorithm can be applied equally to any reliability problem by defining the limit state function, $g(x)$, transforming the variables to standard normal space, and solving for the gradient of $g(x)$.

Setup: We initialize the problem with the following information.

The limit state function $g(x)$ for the cut slope problem contains cohesion and Gamma as correlated or joint normals with a correlation coefficient of 0.5.

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0

$$g(x) = c - \frac{H}{4}\gamma$$

R is the correlation matrix, and D is the standard deviation matrix.

$$R = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 30 & 0 \\ 0 & 2 \end{pmatrix}$$

We find the L matrix using Cholesky decomposition.

$$L = \text{cholesky}(R) \quad L = \begin{pmatrix} 1 & 0 \\ 0.5 & 0.866 \end{pmatrix}$$

M is the vector of mean values. We set the initial trial of this iterative solution, X, to be equal to the mean values. H is a constant in this analysis.

$$M = \begin{pmatrix} 100 \\ 20 \end{pmatrix} \quad X = M \quad H = 10$$

Transformation: The variables are transformed into standard normal space, from x to u, as follows.

$$u = L^{-1}D^{-1}(X - M)$$

$$g(x) = M_0 - \frac{H}{4}M_1$$

The gradient vector has the partial derivatives of $g(x)$.

$$\nabla g(x) = \frac{1}{-H/4}$$

HRLF Algorithm: We solve the Jacobian matrix as shown and follow the steps of the HLRF algorithm.

$$G(u) = g(x) \quad Jux = L^{-1}D^{-1} \quad Jxu = Jux^{-1}$$

$$\nabla G(u) = Jxu^{\mathrm{T}}\nabla g(x) = \begin{pmatrix} 27.50 \\ -4.33 \end{pmatrix}$$

$$\alpha = \frac{\nabla G(u)}{|\nabla G(u)|} = \begin{pmatrix} -0.988 \\ -0.156 \end{pmatrix}$$

Direction Vector:

$$d = \left(\frac{g(x)}{|\nabla G(u)|} + \alpha * u\right)\alpha - u = \begin{pmatrix} -1.774 \\ 0.279 \end{pmatrix}$$

$$\text{step} = 1$$

New Point:

$$u_{\text{new}} = u + \text{step} * d = \begin{pmatrix} -1.774 \\ 0.279 \end{pmatrix}$$

Check:

$$\frac{|u|}{|\nabla G(u)|} = 0$$

$$\varsigma = 1$$

$$m_1 = 0.5 * (|u|)^2 + \varsigma * |g(x)| = 50$$

Evaluate $g(x)$ for new point.

$$m_2 = 0.5 * (u_{\text{new}})^2 + \varsigma * g(x) = \begin{pmatrix} 51.574 \\ 50.039 \end{pmatrix}$$

Convergence is achieved if the ratio of the limit state function at the new and old mean values is below some prescribed tolerance threshold. The results in standard normal space and original space are then:

$$u_{\text{new}} = \begin{pmatrix} -1.774 \\ -0.279 \end{pmatrix}$$

$$\alpha = \begin{pmatrix} -0.988 \\ 0.156 \end{pmatrix}$$

$$\beta = \alpha^{\text{T}} * u_{\text{new}} = 1.796$$

$$p_f = \Phi(-|\beta|, 0, 1) = 0.0362$$

The reliability index and the probability of failure can be checked against the exact, approximate, and numerical solution for this same problem in Chap. 6.

Now solving the cut slope problem, this time for uncorrelated normals.

Setup:

$$g(x) = c - \frac{H}{4}\gamma$$

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 30 & 0 \\ 0 & 2 \end{pmatrix}$$

$$L = \text{cholesky}(R) \quad L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$M = \begin{pmatrix} 100 \\ 20 \end{pmatrix} \quad X = M \quad H = 10$$

Transformation:

$$u = L^{-1}D^{-1}(X - M)$$

$$g(x) = M_0 - \frac{H}{4}M_1$$

$$\nabla g(x) = \frac{1}{-H/4}$$

HRLF Algorithm:

$$G(u) = g(x) \quad Jux = L^{-1}D^{-1} \quad Jxu = Jux^{-1}$$

$$\nabla G(u) = Jxu^{\mathrm{T}}\nabla g(x) = \begin{pmatrix} 30 \\ -5 \end{pmatrix}$$

$$\alpha = \frac{\nabla G(u)}{|\nabla G(u)|} = \begin{pmatrix} -0.986 \\ 0.164 \end{pmatrix}$$

Direction Vector:

$$d = \left(\frac{g(x)}{|\nabla G(u)|} + \alpha * u\right)\alpha - u = \begin{pmatrix} -1.622 \\ 0.27 \end{pmatrix}$$

$$\text{step} = 1$$

New Point:

$$u_{\text{new}} = u + \text{step} * d = \begin{pmatrix} -1.622 \\ 0.270 \end{pmatrix}$$

Check:

$$\frac{|u|}{|\nabla G(u)|} = 0$$

$$\varsigma = 1$$

$$m_1 = 0.5 * (|u|)^2 + \varsigma * |g(x)| = 50$$

Evaluate $g(x)$ for new point.

$$m_2 = 0.5 * (u_{\text{new}})^2 + \varsigma * g(x) = \begin{pmatrix} 51.315 \\ 50.037 \end{pmatrix}$$

$$u_{\text{new}} = \begin{pmatrix} -1.622 \\ 0.270 \end{pmatrix}$$

$$\alpha = \begin{pmatrix} -0.986 \\ 0.164 \end{pmatrix}$$

$$\beta = \alpha^{\text{T}} * u_{\text{new}} = 1.644$$

$$p_f = \Phi(-|\beta|, 0, 1) = 0.0501$$

Reference

Zhang, Y., & Der Kiureghian, A. (1995). Two improved algorithms for reliability analysis. In *Reliability of and optimization of structural systems. Proceedings of 6th IFIP Working Conference on Optimization of Structural Systems* (pp. 297–304).

Index

R. E. S. Moss, *Applied Civil Engineering Risk Analysis*,
https://doi.org/10.1007/978-3-030-22680-0