University of Alberta

Understanding and Improving Input for Quantitative Risk Analysis in the Construction Industry

by

Hala Abou Rizk

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Construction Engineering and Management

Civil and Environmental Engineering

Fall 2013
Edmonton, Alberta

Abstract:

Risk analysis in the construction industry involves identifying risk events that could potentially affect a project and its delivery, quantifying those risks, and developing mitigation strategies to enhance project success. This thesis aims to improve the process of risk analysis through the enhancement of the quantification process (QRA) using Monte Carlo techniques. In particular, this study investigates qualitative verbal expressions utilized when gathering information from experts and the methods by which they are converted into quantitative data for analysis. The effects of the quantitative data used as inputs have been found to affect the resulting values and distributions. In order to enhance this process, a survey is created to better understand verbal expressions used in the risk analysis process. This has resulted in a table of corresponding quantitative values, in terms of deterministic values and beta distributions, which can be utilized for QRA and extended to other areas of academia.

Acknowledgements

Without the support of my incredible family, this could not have been possible. Thank you to my loving husband, who kept me sane throughout this process, doing anything in his power to help me succeed. To my sisters, Jenna and Deema, for consistently being my biggest supporters and always knowing exactly how and when to help. To my mother for always believing in me more than I ever have myself, for keeping me on track to complete when I needed the motivation, and for providing me with every type of support I could require. Finally, to my father, who is not only my supervisor, but the entire reason that I am graduating with a masters in construction management, I cannot thank you enough for being there with me every step of the way, encouraging me, writing with me, teaching me, and making this experience one I will never forget, I am the luckiest student/daughter to get to be in your company.

I owe a great deal of gratitude to Dr. Robinson Fayek for her constant support and wisdom throughout this process, for being a perfect role model, and always making me feel like family. I could not have asked for more.

To my family at SMA (especially to Jesse, Hussien, Fangyi, Mohammed and Holly) for allowing me the time to pursue this degree, for being patient throughout the process, for providing me the experience I needed to understand my studies, and always offering any help and assistance I could need. I am so lucky to work where I do.

Table of Contents

List of Tables

List of Figures

1 Chapter 1: Introduction

1.1 Background and Problem Statement

The development of a capital construction project undergoes various stages including conceptual design, preliminary design, detailed design, construction, and commissioning. During those stages many decision are made based on uncertain information, potential risk events and dynamic market conditions. Risk analysis is generally accepted as a process that can be used to understand uncertainty in the project development lifecycle. It is defined as one of the main processes for project management in the project management body of knowledge (Project Management Institute, 2000). Risk analysis involves identifying risk events (or factors) that could potentially affect the project and its delivery, quantifying those risks, and developing mitigation strategies to enhance project success. This thesis is concerned with one element of risk analysis, namely the quantification process (QRA) for costs using Monte Carlo techniques.

The risk analysis quantification process can generally take a variety of forms. The Construction Industry Institute (CII) Implementation Resource 280-2 outlines three levels of risk analysis based on the level of quantification required. In Level 1[1], risk factors are identified and quantification is limited to a simplistic ranking of the importance of risks. Level 2 requires estimating probabilities and impacts of each risk factor using approximate indicators from tables, for example. Level 3 requires more detailed

[1] These levels identified by CII will be referred to throughout this thesis to represent different risk analysis processes.

analytical techniques such as Monte Carlo simulation modeling (Construction Industry Institute, 2012).

Quantitative risk analysis (QRA) is a term used to refer to Monte Carlo Based analysis of risks (Level 3 in CII-IR 083). It has become popular with the introduction of Monte Carlo Simulation software accessible on a variety of computing devices over the past few decades. QRA models simply represent risk factors in a model with each risk modeled with (1) a probability of occurrence and (2) an impact upon its occurrence (this thesis will focus on impacts in terms of costs, however other impacts including schedule, quality scope etc. are possible). Values are sampled for all risks for both probability and impacts and combined to estimate the overall impact to the project's costs and schedule. Once a Monte Carlo Simulation Model is complete, the project cost (and/or schedule) is represented with a distribution. The decision maker then uses that distribution to determine the project budget by accepting a certain amount of risk. For example, if a distribution of the results of an analysis showed a range between $100 000 and $200 000, the percentiles of the distribution, based on the shape and skew, could be determined. If the 85th percentile of this distribution was determined to be $170 000, this would mean that choosing a budget of $170 000 would result in a 15% possibility of cost overrun. If any other percentile was chosen, this would be the same principle (100% - percentile chosen (e.g.: 35%) = possibility of cost overrun (e.g. 65%). This can provide input to decision makers for contingency computation for the project, for example.

A related term to QRA is Range Estimating which refers to a Monte Carlo Simulation of the estimate itself including risk quantification. The range estimating approach is best described in AACE International Recommended Practice No. 41R-08. In this approach, the simulation model represents the line items in the estimate and their summation. During simulation we simply sample the base values for the individual line items (or their basic parameters like unit rates or quantities), and extend them as required to determine the overall cost distribution of the project.

Comprehensive Monte Carlo Simulation models of a project's cost and schedule can also be created (an extension of, or combination of the above two methods). Models can be created to model an integrated system of all uncertain components of the project. A comprehensive simulation model for a bridge construction project is provided in Chapter 3 to demonstrate such models.

While the approaches of Monte Carlo Simulation are popular, research to support systematic QRA modeling were previously identified and remain unresolved (Galway, 2004). In this thesis the focus is on risk quantification (the determination of probability and impact) in support of systematic QRA modeling.

In most QRA applications, the input values of probabilities of risk events are generally subjectively determined and, therefore, subject to manipulation and incorrect interpretation. The risk impact requires identifying and parameterizing statistical distributions (again mostly subjectively) which is also subject to similar manipulation and incorrect interpretation. The proper conduct of QRA studies requires proper determination of and understanding of the ‘likelihood’ and ‘impact’ of a risk occurring. These parameters provide the foundation of the process of risk analysis, thus if gained inaccurately, can result in unreliable estimates thus limiting the usefulness of the entire exercise. By undertaking this research the intent is to create a practical and reliable method of quantification for QRA.

1.2 Purpose of Study

The main objectives of this study are to:

- Understand the impact of introducing errors in the quantification of subjective information for quantification of probability and impact during a QRA.
- Provide a method for gaining reliable measures for probability of a risk occurring.
- Provide a method for gaining reliable measures for the impact of a risk occurring.
- Provide improvement to the risk analysis process to be used by industry.

1.3 Expected Contributions

1.3.1 Academic Contributions

The expected academic contributions can be summarized as follows:

1. Developing the means to transform verbal expressions of likelihood to numeric equivalents, which can be used in all areas of academia where such expressions are utilized.
2. Developing an understanding of an impact of the shape of the distribution chosen to model a range on the final project cost.

3. Developing a survey that can be used as a means to understanding verbal expressions in terms of quantitative data and provide a guide to others that use verbal expressions and would like to extend this research to other fields (e.g. in FMEA analysis).

1.3.2 Industrial Contributions

The expected industrial contributions can be summarized as follows:

1. Developing a table that can be used during risk analysis to replace verbal expressions both deterministically or in terms of distributions.
2. Providing a guide to utilizing Monte Carlo simulation for risk analysis in the industry

1.4 Research Methodology

This research was conducted using the following methodology:

- Conduct a literature review on the methods of QRA, measures for understanding quantifying verbal expressions, Monte Carlo simulation, and applications of QRA in the construction industry.
- Develop Monte Carlo simulation experiments to understand the sensitivity of decision parameters used in QRA to the subjective input (of probability and impact) estimation by experts.
- Conduct informal interviews with industry professionals to understand the reasoning behind the use of QRA.

- Create a questionnaire based on limitations in current literature and requirements set out from professionals.
- Complete a statistical analysis of results gained through the survey.
- Provide recommendations to improve the quantification of subjective input for QRA that can be used to enhance best practices and become useful in a practical setting.

1.5 Thesis Organization

The thesis is organized as follows: *Chapter 2* provides a summary of the literature and the state of the art. *Chapter 3* provides a case study of QRA to set the stage for the thesis work scope. *Chapter 4* summarizes the results of experiments to investigate the sensitivity of decision parameters used in QRA to the input provided subjectively by experts. *Chapter 5* details the survey undertook to develop a better manner for quantifying subjective input for probability and determination of parameters of the distributions for risk impact. *Chapter 6* provides the analysis of the results, and *Chapter 7* provides recommendations to enhance current practice and conclusions.

2 Chapter 2: Literature Review

2.1 Introduction

Because construction projects innately contain a great deal of uncertainty (Al-Bahar & Crandall, 1990), risk management has become essential to successful project delivery. It allows for proper action to be taken towards the duty of care and for proper management of inevitably unique and consistently changing projects. Risk can affect countless areas of project delivery including "productivity, performance, quality and budget of a construction project" (Kangari, 1995, p. 442). Risk analysis continues to be regarded as an imperative part of successful project delivery. Quantitative risk analysis is a part of the overarching process of risk management and is vital to practical application. By quantifying risk, decision makers are able to see the direct applications to their work including gaining results in terms of monetary values and schedule which could affect the project.

2.2 Definitions

There are numerous definitions of risk. The following is select set of definitions that are mostly commonly encountered:

- ISO Guide 73 (2009) defines risk as the “effect of uncertainty on objectives”. The effect can be a negative or positive deviation from expected objectives (ISO, 2009). The definition emphasizes that in order to complete a risk analysis, objectives of an investment, for example, must be established and the events that could cause a deviation from achieving the defined objectives, along with the likelihood of occurrence of such events and their consequences.

- The Online Oxford Dictionary defines risk as: The possibility of suffering harm or loss; a factor, thing, element, or course involving uncertain danger. Where a hazard is “a situation involving exposure to danger” (Oxford University Press, 2013).
- The Project Management Institute (2000) defines risk as: “An uncertain event or condition that if it occurs has a positive or negative impact on project objective.”
- Kumane and Mahadik (2013) define risk as: “The exposure to the chance of occurrences of events adversely or favorably affecting project objectives as a consequence of uncertainty.”

2.3 Risk Analysis Process Review:

In order to understand quantitative risk analysis, one must first gain an overall understanding of risk management in construction and how it is used.

This process, according to the PMBOK consists of six essential steps:

1. Risk Management Planning
2. Risk Identification
3. Qualitative Risk Analysis
4. Quantitative Risk Analysis
5. Risk Response Planning
6. Risk Monitoring and Control

While the names of the above steps of risk management may vary among the literature, the overall information is consistent. For example, the following will outline the ISO (2009) risk standards and show the associated PMBOK steps:

1. Establishing the Context (step 1);
2. Risk Identification (step 2);
3. Risk Analysis (includes both steps 3 and 4);
4. Risk Evaluation; Risk Treatment (step 5);
5. Monitoring and Review (step 6); and
6. Communication and Consultation (incorporated in all steps).

It is understood among the literature that while these steps can be outlined in general terms as linear, it is a cyclical process that can continue to re-route throughout a project's lifespan. The following section will detail these steps based on the PMBOK outline terminology.

2.3.1 Risk Management Planning

The risk management process in construction begins with risk management planning; this phase outlines the strategy for achieving the objectives and goals for the process. During this phase, owner objectives, scope, and risk criteria are outlined. It is important to consider both the external environment (e.g.: political, economic trends, external stakeholder input) and internal environment (e.g.: organizations expectations and risk tolerance) to understand the overall project expectations (ISO, 2009). The project charter, risk management policies of the given party, roles and responsibilities and the work breakdown structure among other processes are thoroughly investigated during this phase (Project Management Institute, 2000, p. 129). It is imperative to understand the unique intentions, requirements and risk tolerance level of the given party for each project. The result of the risk planning phase is a document consisting of the risk management plan. This plan summarizes the information gained during this phase and creates an outline of key plans and steps to follow for the remainder of the process.

2.3.2 Risk Identification

Once the risk management plan is in place, the risk identification phase can begin. Ideally, during this process, key project personnel, the risk management team and experts in the area (both those affiliated with the project and not affiliated) attend a brainstorming session where all possible project risks can be recorded. During these brainstorming sessions all risks should be incorporated, including those related to quality of work, efficiency, cost and schedule (Al-Bahar & Crandall, 1990). This can then be supplemented with historical information from similar projects, especially including documented projects consisting of lessons learned. Risks identified can be organized in terms of categories that best suit the given project. Categories should strategically unify certain risks and can help with subsequent stages of risk management. The Project Management Institute (2000) provides examples of such categories including "project management risks, organizational risks, and quality risks" (Project Management Institute, 2000, p. 132). The risk identification process is ongoing for the duration of the project. As uncertainty decreases, such as when a project design becomes complete, risks will become clearer. Risks should be updated constantly within the project.

2.3.3 Qualitative Risk Analysis

Qualitative risk analysis refers to when risk factors are analyzed in terms of their likelihood of occurrence on the project, and the impact that this occurrence would have on the project if the risk were to occur. These factors are examined in the qualitative phase, in terms of verbal expressions. This method lends itself best when utilizing expert opinions, as verbal expressions are often innate and understood by the general population. By examining each risk identified, risks can be better understood. The quality of each risk determined in the risk identification phase can be assessed, and some may be omitted. Risks can be organized in terms of importance based on the assessed impact and available methods of mitigation for a risk. The first level of quantification described by the Construction Industry Institute (CII) Implementation Resource 280-2 involves enumerating the risk factors (or events) and then quantifying them by subjectively ranking them in terms of importance. The higher ranked risks are then mitigated and managed on the project. Specialized tools and methods have been developed to examine risks in this way. One such commonly used tool is the risk rating matrix (AbouRizk S. M., Risk Analysis for Construction Projects: A practical guide for engineers and project managers, 2009). This matrix combines the probability and the impact in order to understand the severity of the risk and the actions that should be taken. For example, if a risk is very likely to occur and the impact is disastrous, this would mean that a risk is intolerable for the project. Measures would either have to be taken to eliminate this risk, or the project might be terminated. Where the line is drawn for the severity of the risk depends entirely on the risk tolerance of the decision maker. While some parties are willing to take a great deal of risk, others may have to draw the line much earlier. Once a qualitative assessment of risks has been completed, there is generally a good understanding of the most important risks that will affect project. At this point, certain decisions can be made. Perhaps based on this analysis it is found that most risks occur in one particular location of the project. Redesign of this area may be a possibility and could be investigated. Likewise, a different method of construction may be chosen based on the number of risks impacting the project due to this method. In order to gain further understanding of the cost and schedule impacts that risks will have on the project, quantitative analysis usually follows.

2.3.4 Quantitative Risk Analysis

Quantitative risk analysis (QRA) uses the results determined during qualitative analysis and turns them into quantifiable information that can be useful for the project. This is generally done by assigning probability and impact scores to the verbal expressions, and multiplying them together in order to gain a value for risk severity (Construction Industry Institute, 2012). Quantitative analysis can occur for both cost and schedule; however the focus of this paper will remain on costs. Quantitative analysis of schedules can be undergone using techniques such as the PERT analysis or Monte Carlo Simulation of CPM networks. Quantitative analysis of cost utilizes the work breakdown structure, an analysis of the estimate often using Monte Carlo simulation or other analytical techniques.

After a thorough review of information on quantitative risk analysis, Galway (2003) concluded that the literature is split among many different fields, not necessarily condensed into one area. Similarly, projects utilizing quantitative risk analysis are continuously changing, progressing, and often confidential. There is little incentive to find evidence for the use of risk analysis on a project, often due to limited project resources, thus analysis to determine whether the process worked is often not in the budget or scope of work and therefore difficult to analyze. Galway (2003) has, however, found that the common consensus among users of risk analysis is that it is valuable and desired. Still, empirical studies appear to be non-existent in this area. This report concludes that this is clearly an area where research is needed. The lack of a body of empirical evidence, the often-cited reluctance of managers to use risk analysis techniques in project management, and the ambivalence of risk practitioners themselves over key issues such as applicability all call for a program of evaluation of these techniques and their application, especially in the area of complex, technologically advanced projects." (Galway, 2004, p. 34).

Methods for Quantitative Risk Analysis

While the other phases of the risk management process can be compressed based on literature consensus, quantitative risk analysis is much more diverse. The following portion of this section, therefore, provides an overview of the different methods, most commonly used. While this does not include all of the possible methods of QRA, it should provide a basis for the approach being used in this thesis.

2.3.4.1.1 Quantitative Risk Analysis

The Construction Industry Institute (CII, Resource 280-2), outlines their risk process in terms of three levels. For their second level of quantification, each of the risk factors that were identified in level one (identified here as a qualitative approach in section 2.3), is given a probability and a consequence to costs or schedule. The multiplication of the two provides an indication of risk severity which can then be used to segregate risks into various groups of importance. The third level described by the CII requires more detailed analytical techniques such as Monte Carlo simulation modeling. This is generally referred to as quantitative risk analysis (QRA). Within QRA, each of the risk factors is quantified by estimating a probability of occurrence, and a cost or schedule impact upon the risk's occurrence. In a Monte Carlo simulation, values are sampled for each of the risks for both probability and impacts and combined to estimate the overall impact to the project's costs and schedule. Once a Monte Carlo Simulation Model is complete, the project cost (or schedule) is represented with a distribution. The decision maker then uses that distribution to determine the project budget by accepting a certain amount of risk (Construction Industry Institute, 2012).

2.3.4.1.2 Monte Carlo Simulation

Monte Carlo Simulation is a proven technique used to simulate costs and schedule. Monte Carlo Simulation makes use of probability distributions in place of deterministic values in order to model the uncertainty associated with a particular input and the possible outcomes that can occur, thus allow for better decision making abilities when dealing with such uncertainty (Palisade Corporation, 2010). This technique simulates the real world by randomly selecting data from a defined distribution in order to provide the most probable distribution of the component subject to uncertainty. Monte Carlo Simulation depends on statistical sampling to evaluate outcomes. Therefore the simulation experiment entails taking samples on numerous iterations. The more iterations run, the more realistic the resulting distribution will be. This is because every time an iteration is carried out, a component is randomly selected from its defined cost distribution (or schedule). This varying quality is the reason that the random characteristics can be modeled, which allows for an increase in both reliability and validity of the measure at hand.

The purpose of Monte Carlo Simulation is to derive the true probabilistic distribution for the estimated cost (or schedule) of each component of a project estimate under uncertainty. The nature of the Monte Carlo Simulation allows the random characteristic of the cost of construction items to be modeled. This is achieved by providing numeric estimations of the uncertain features of the cost components. Each of the uncertain components of the estimates is then estimated by a probability distribution of the cost rather than a single number.

A range estimating approach is described in AACE International Recommended Practice No. 41R-08. In this approach, the simulation model represents the line items in the estimate and their summation. During simulation samples of the base values for the individual line items (or their basic parameters like unit rates or quantities) are taken, and extended as required to determine the overall cost distribution of the project.

A more detailed discussion of Monte Carlo Simulation models for Range Estimating and Risk Analysis is given in Chapter 3.

2.3.4.1.3 HAZOP

HAZOP (hazard and operability) analysis is a Process Hazard Analysis (PHA) technique, used to identify safety problems (Rodríguez & Luis de la Mata). HAZOP was originally developed by a chemical company to identify hazardous materials and thus avoid any sort of dangerous accidents, and was officially published as a procedure to identify any potential variation from the design intention, in the 1970s (Dunjóa, Fthenakisb, & Vílcheza, 2010). A HAZOP study is defined as a “formal, systematic critical examination of the process and the engineering intentions of new or existing facilities to assess the potential for malfunctioning of individual pieces of equipment, and the consequential effects on the facility as a whole” (Dunjóa, Fthenakisb, & Vílcheza, 2010, p. 20). This process involves methodically analysing each portion of the design, with a multidisciplinary team, and determining any ways that deviations from the design intent could occur. Guidewords and process parameters are used to expose deviations in each node (a part of the system with a particular purpose, used to distinguish parts of the whole to analyse) (Dunjóa, Fthenakisb, & Vílcheza, 2010). Such guidewords include “no, more, less, higher, lower, part of” while process parameters can include “temperature, flow, pressure, level” etc. For example, the guide word “more” may be paired with “flow” indicating potential flow problems in a system. These deviations are then analysed as to their potential consequences, and the effect of these consequences on the project are assessed. Mitigation measures, or “safeguards” may be developed for areas safety or efficiency could be jeopardized. This method has been found to be very useful in the construction industry during all phases of the project life.

A sample of a HAZOP analysis (one node and one deviation)[2] for water treatment plant is shown below in Table 1.

Table 1: HAZOP analysis example

Node: 2. Cells Outlets		Design Intent: Prevent debris from entering downstream system - max 25 mm discharge (screens)		
Type: Control				
Deviation 1. No Flow or Less Flow				
Causes	Consequences	Safeguards	Recommendations	Responsibility
The Screens are plugged up	The water flow will stop from the supply	Regular maintenance including check and inspection (usually divers are hired to inspect those screens)	Regular maintenance including check and inspection (usually divers are hired to inspect those screens)	Operations

2.3.4.1.4 *FMEA*

FMEA, or failure mode and effects analysis, is a tool used to quantify reliability. FMEA is a subjective analysis tool used for systematic identification of possible Causes and Failure Modes" which show the cause and effect of failures and allows for the assessment of potential risks (Arabian-Hoseynabadi, Oraee, & Tavne, 2010, p. 818) (Rhee & Ishii, 2003). FMEA is recommended by many international

[2] Reproduced with permission from SMA Consulting Ltd.

standards including the "Society of Automotive Engineers, US Military of Defense, Automotive Industry Action Group" (Wang, Cheng, Hu, & Wu, 2012, p. 959). In a typical FMEA study, a comprehensive exercise is undertaken to identify all of the system, every mode of failure of a component, and all possible causes and of failure for each failure mode. This is generally done in a workshop setting, with a group of experts, particularly design and maintenance personnel who are the area to be analysed (Arabian-Hoseynabadi, Oraee, & Tavne, 2010). The FMEA severity, occurrence and detection as measures of the failure mode in order to Risk Priority Number (RPN) (Wang, Cheng, Hu, & Wu, 2012). Each of the its own rating scale and description which is used to qualitatively examine each mode. The RPN is a product of these three measures and depending on the goal of FMEA can be used in multiple ways to analyse the results. If the goal is to compare potential design options for example, the analysis may be structured in a way that RPNs of particular failure modes can be compared. In addition to the evaluation of failure modes and effects, mitigation strategies are typically identified for critical modes. A sample FMEA table is given in

Table 2 below for illustration.

Table 2: FMEA example

1	2	3	4	5	6	7
Mode of Failure	**Cause of Failure**	**Effect of Failure**	**Frequency of Occurrence**	**Degree of Severity**	**Chance of Detection**	**Risk priority number RPN (4)x(5)x(6)**
A conveyor belt break down	Rocks jamming the belt	Crushing operation will cease	Once every two months 7	10	1	70
----	----	---	---	---	---	---

2.3.4.1.5 *Fault tree analysis*

Fault tree analysis (FTA) is a risk analysis technique mainly used for accident scenario assessment (Chiacchio, Compagno, D'Urso, & Mann, 2011). The FTA technique is based on deductive logic. It shows the relationship between system failures and failures of the components of a system through a logical diagram. A risk tree is created based on a defined, undesirable event. In FTA an event can be "any proposition that is true with a certain probability" (Kaisera, Gramlich, & Fo, 2007, p. 1522). Failures leading to the initial event are identified and described as either a primary or intermediate events. A primary event is not used for further analysis, while intermediate events are developed and connected to a gate which connects this to other events (Brooke & Paige, 2003). Examples of gates are the "AND" gate, which specifies that the hazard is caused by two or more factors, and the "OR" gate, which specifies that the hazard is caused by any of the factors on its own (Kaisera, Gramlich, & Fo, 2007). Analysis can be qualitative or quantitative (Dhillon & Singh, 1979). Qualitative analysis can include lists of failure combinations which lead to top-level failures. Quantitative analysis involves calculating probabilities for top-events based on basic events. FTA is generally supplemented with software or other technical systems. (Kaisera, Gramlich, & Fo, 2007). Benefits of FTA, according to Dhillon and Singh (1979) includes: the ability to focus on a single event at time, enhancement of the system design, a useful graphical representation, improved organization, control and flexibility.

2.3.4.1.6 Decision Trees

Decision trees describe a problem or a product as a sequence of decisions and events in a chronological order using a tree structure. The tree is constructed using decision nodes, chance nodes, branches and payoff nodes. Decision trees have the advantage of showing all possible decision operation and chance events in great detail. They show events and decisions as they occur in time which proceeding from left to right (Palisade Corporation, 2010).

A sample decision tree for a budget decision is given in Figure 1 (Courtesy of SMA Consulting Ltd). A decision by an owner has to be made regarding a budget approved by its board of directors at $46 Million while an estimate by the design team shows that the budget needs to be increased to $56 Million.

The decision node (square) on the left most side shows that the decision maker must make a decision about a budget either to keep it at a set value of 46 Million with specific outcomes (subject to chance) or to increase it to 56 Million and with various outcomes subject to chance. The upper branch of the tree shows what happens if he decides to increase the budget to 56 Million and then go to tender. The chance node (circle node) has three branches with outcomes and probabilities on them. There is a chance of 9.7% that the tender will be under the original budget of 46M; a chance of 90.29% that the budget is higher than what was originally budgeted (46 Million) but lower than the revised one (56 Million); and a small chance of 2% that the tender will be higher than the revised budget of 56 Million. The payoffs of each are represented with a negative dollar value in this case (a loss).

The lower branch is structured in a similar way. The tree can be used to select the decisions and branches that produce the most favourable expected value of the outcome. In this case (using Precision Tree from Palisade to construct the problem), the best course of action is to keep the budget at 46 Million, but award base items and, if budget permits, award provisional items (see where the tree says True).

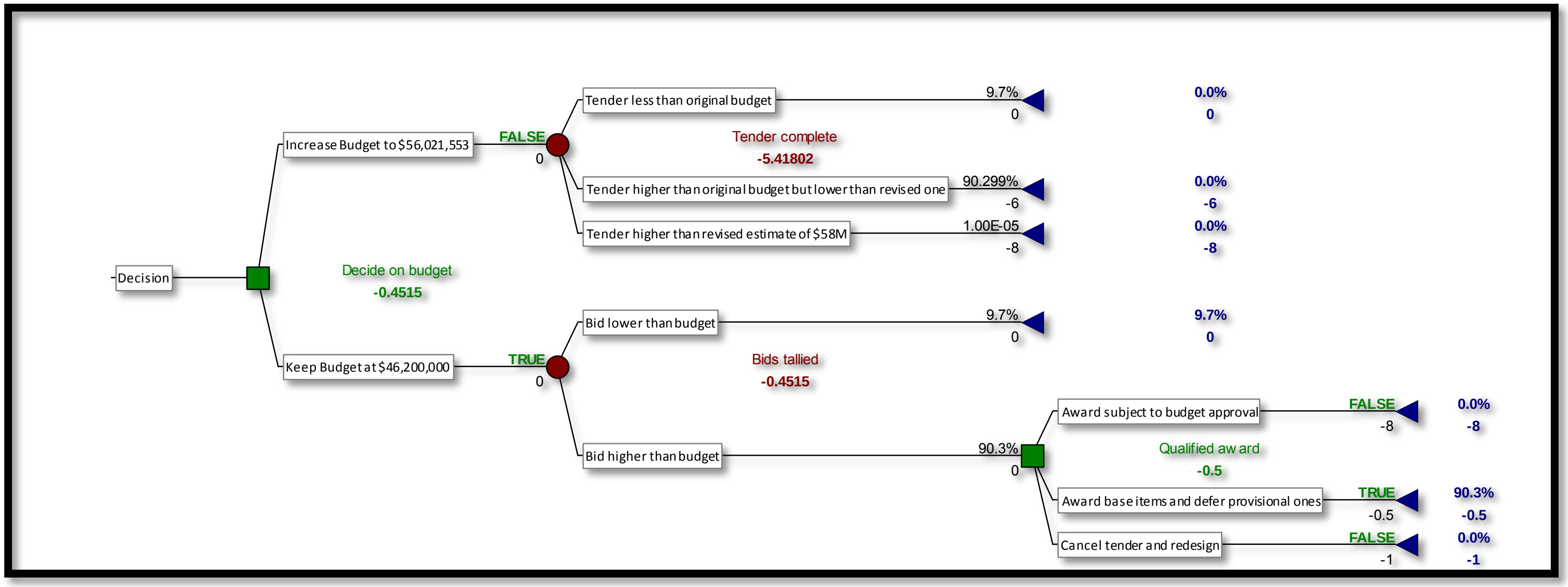


Figure 1: Decision tree example (courtesy of SMA Consulting Ltd.)

2.3.5 Risk Response Planning

Based on the established risks, risk response planning involves determining the most optimal way of mitigation (Project Management Institute, 2000). Response planning may directly follow the qualitative phase, or continue after quantification. This depends on what is defined in phase 1 (risk planning), as goals will vary based on the project. The risk response planning phase involves determining how to avoid a risk, for example changing portions of the design or undertaking a new process which can help to reduce a risk. Such methods are usually derived through a workshop setting and conducted in a hierarchical manner (risks most dangerous to the project first). Historical data can also help to determine the most effective ways to avoid a risk.

2.3.6 Risk Monitoring and Control

Risk monitoring and control, the final phase on risk management, involves continual inspection, examination, review and observation in order to determine if any changes or unexpected outcomes have occurred and to maintain that the established goals are being met (ISO, 2009). This is often done by having a designated individual or group of individuals who are dedicated to checking on the process through interviews with appropriate individuals, on site checks, consistent reporting and thorough documentation of events.

2.4 Input Modelling

Most of the risk analysis methods described above (especially the QRA and RE), require appropriate modeling of the input parameters. For example, most of the reviewed techniques require that one estimate the probability of occurrence of future events. This is often not available through historic data or through analytical modeling and must be estimated through expert judgement. Likewise many of the techniques require that impacts of risk events be estimated through some form to reflect uncertainty. This is generally done through statistical distributions and is known in the simulation literature as ‘input modeling’.

Input modeling in Monte Carlo Simulation, receives significant attention in the simulation literature (Law, 2012). Most of the literature, however focuses on parameterizing a distribution (defining the parameters of a distribution) when historical data exist. AbouRizk (1990) reported on input modeling for construction simulation. In his work he deduced that using flexible distributions to model existing data was best suited to the data he collected. He further recommended the use of the Beta distribution family to model construction duration data due to its flexibility in representing many data sets, as well as it was readily available within most simulation software. Law (2012) shows that the most common pitfalls input modeling include (1) replacing a distribution by its mean (i.e. using the mean value, a deterministic number, to represent the input instead of the distribution itself), and (2) selecting the wrong distribution to represent the data.

Probability

In terms of probability, the literature mostly shows that in the absence of data (as is the case in most QRA analyses), the prevailing approach is either to solicit the probability directly from the expert as a percentage (0% to 100%) or to use terms for subjective probabilities and then convert those terms to probabilities.

The use of percentages is self-explanatory but generally ill-advised as most experts cannot distinguish probabilities on such a continuous scale. For example, the difference between 5% and 7% would be very difficult to state subjectively. When more than one expert is involved (such as the case in QRA) and consensus is required for a probability, this would generally present a challenge.

The use of verbal expressions to specify probability can be found in literature without much consistency or uniformity. Reagan et.al. (1989) show tables with equivalent terms and probabilities that were based on a survey of students. AbouRizk (2009) shows a table that can be used for risk analysis workshop quantification (see Table 3):

Table 3: Probability estimated from linguistic terms (AbouRizk, 2009)

Likelihood table	Low Prob.	High Prob.
HL --Highly Likely: Almost certain that it will happen, > 70%	0.7	1
LI --Likely: More than 50-50 chance 50% - 70%	0.5	0.7
SL (L) --Somewhat Likely: Less than 50-50 chance 15%-50%	0.15	0.5
UN --Unlikely: Small likelihood but could well happen 1% - 15%	0.01	0.15
VU --Very Unlikely: Not expected to happen 0.01% - 1%	0.0001	0.01
EU --Extremely Unlikely: Just possible but would be very surprising < 0.01%	0.0000	0.0001

In the Failure Mode and Effects Analysis (FMEA): A Guide for Continuous Improvement for the Semiconductor Equipment Industry Technology Transfer

#92020963B-ENG (Villacourt, 1992) Table 4 is shown to quantify probability of occurrence of a failure in a system for semiconductor manufacturing:

Table 4: Occurrence likelihood table (SEMATECH 1992)

Value	Frequency	Frequency
1	An extremely unlikely probability of occurrence during the item operating time interval. Extremely Unlikely is defined as about once every 5 to 10 years.	Once every 5 to 10 years
2	An unlikely probability of occurrence during the item operating time interval (i.e. once every year)	Once a year
3-4	A remote probability of occurrence during the item operating time interval (i.e. once every 6 months)	Once every 6 months
5-6	An occasional probability of occurrence during the item operating time interval (i.e. once every 3 months)	Once every 3 months
7-8	A moderate probability of occurrence during the item operating time interval (i.e. once every month)	Monthly
9	A high probability of occurrence during the item operating time interval (i.e. once every week)	Weekly
10	A very high probability of occurrence during the item operating time interval (i.e. once every day)	Daily

The tables described in here are presented to illustrate the wide variety of possibilities for how probabilities are estimated from experts.

A study conducted by Reagan, Mosteller and Youtz (1989), examined the accuracy of probabilities that are expressed in words, rather than numbers. Their findings demonstrate the wide range of interpretations of such terms by different individuals.

Impact or Consequence:
Biller and Guens (2010) overview input modeling for simulation studies. In their paper, they indicate that “when no data are available, the key idea is to use any available information that may help to identify some characteristics of the process”.   Furthermore, they indicate that in the absence of data, expert opinions are widely used to characterize a

distribution. The resulting distribution compiled from the expert depends on how much information can be solicited. The approach represents the current state of practice which depends on soliciting parameters that can be used to characterize the distribution. AbouRizk (1990) shows that in the absence of data, one can solicit the most likely value, and the two end points with some reliability from an expert. Beyond this, the information may become unreliable. Using those three parameters one can fit a triangular distribution and in the case of AbouRizk (1990) he shows how Beta distributions can also be specified from such information. Furthermore, his approach involved presenting a Beta distribution to the expert and then allowed him to manipulate the shape to reflect his best judgement of the underlying distribution.

Law (2012) and AbouRizk (1990) indicate that in the absence of data, and for continuous distributions, a triangular distribution can be specified based on parameters obtained from an expert. While Biller and Guens (2010) argue that "despite being widely used due to their availability in commercial software packages, limited shapes of the standard distributions may not be flexible enough to represent key characteristics of the data". The triangular distribution is one of the standard distributions with limited shapes and hence may not be always adequate to represent underlying data.

2.5 Limitations in the Literature:

The intent of this thesis is to focus on two limitations in input modeling for risk analysis namely, the specifying and quantifying probability (likelihood) and the risk consequences (impact). These two parameters are often defined in qualitative terms (e.g. the probability of an event is "likely to occur"), and then converted to numerical values for simulation (e.g. 60% probability of occurring). Generally, a group of individuals will choose the term for probability that they feel best describes the risk factor at hand. The difficulty with this technique however, is that not all individuals may understand the words provided to have the same meaning. One major gap found in the literature is the method in which input parameters are gained and used. Galway (2004) notes this limitation, explaining that the ability to assess probability distributions for tasks (for schedule and cost) gets inadequate attention in risk analysis literature. In order to achieve the objectives outlined in Chapter 1, (provide a method for gaining reliable measures for probability and impact of a risk occurring, and improve the risk analysis process to be used by industry) this area, will be further analysed and improved.

2.6 Concluding remarks

The literature seems to have created a consensus on the process of risk management and the methods of acquiring the information that is needed. Most stages of the process have a description which has very little variance whether in academia or industry. The quantitative risk analysis process however, has many possible methods which can be used, and there is the least agreement among the state of the art of the most effective method.

A gap in the literature related to converting qualitative likelihood descriptions into probabilities for quantification for risk analysis in QRA, in the absence of historical data, has been identified. There is currently no structured approach that is scientifically derived to accomplish this. While most QRA methods depend on proper solicitation of probability information, a proper foundation for estimation of the risk probability from expert judgement is required.

There is also a gap in the literature regarding the quantification of impact or consequence of risk occurrence. While most simulation literature shows that flexible distributions (such as the Beta distribution) are more appropriate for modeling cost and duration, literature in general shows that the impact of risk is modeled using triangular distributions.

In this thesis I will attempt to understand the implications of the gaps in literature on the outcome of QRA and then develop methods to assist in filling those gaps.

3 Chapter 3: Background of Risk Analysis Process – A case study

3.1 Introduction:

This chapter provides a case study for the risk analysis process for a bridge construction project to illustrate the methods currently carried out in industry. The case study is based on an actual project but has been altered to protect confidential information. The risk analysis is taking place during the preliminary design phase of the project and is being undertaken on behalf of the owner. The risk analysis team is SRANT Consulting (a fictitious name) and is working closely with the design team (TRR) to complete the risk analysis.

The project under considerations is a one way bridge built over the river connecting the south and north side of the City of Champions. Approximately 60% of the project's budget is related to the bridge structure, while the remaining 40% is associated with the construction of and rerouting of existing roads on both ends of the bridge, design fees and administration costs. The project must be completed within a pre-established budget that was determined at concept design and approved by City Council. Care must be taken to ensure all environmental issues are properly dealt with (the bridge is built over the river and therefore is in an environmentally sensitive area), that all potential burial grounds (within Native lands) are preserved and not disturbed, and that traffic is properly accommodated during construction and when the bridge is commissioned.

The project design team is led by a local consultant with extensive transportation experience and is supported by a number of specialty consultants. The focus in this case study is on the bridge structure only for brevity.

The risk analysis process followed on this project can be summarized in the following steps:

1. ***Planning for risk analysis:*** where the risk analysis team understands the background of the project, defines the objectives of the risk analysis, understands the Owner's tolerance for risk and sets up the risk management for the project.
2. ***Complete a risk analysis workshop***: where experts from the design team, the owner, and external parties are asked to participate in a facilitated brainstorming exercise to identify, quantify and suggest mitigation strategies for risks.
3. ***Complete required follow up work***:
    a. Develop a risk register for the project, and complete a risk mitigation strategy.
    b. Develop a Monte Carlo simulation model and simulate the risk occurrence and its impact on the project budget (schedule impact not included in here for brevity but is similar in nature). This step is not always completed for projects (see CII, 2012 as in most cases, teams opt to use Level 2). When this step is completed, more workshops and interviews would be required to solicit input for the quantification.

These steps will be illustrated in the following sections.

3.2 Planning for the Risk Analysis

STRANT held several meetings with key project team members in order to understand and properly define project scope, objectives of the risk analysis and the expected outcomes. Historical information and expert knowledge are also incorporated during these meetings in order to create baseline information for the risk analysis process.

3.2.1 Project Scope

The project scope was discussed and defined as shown below:

- Provide a bridge with 3 lanes connecting the South to the North Side of the City
- Tie in the bridge with the associated roadway system
- Address some of the current congestion issues in the area: traffic delays at the intersections
- Relocation of existing utilities
- Incorporate public art into the project design and construction

3.2.2 Owners' Measure of Success

Measures of success were identified as follows:

- On or under budget
- On schedule
- Positive public, media and stakeholder perception
- Positive interface management with other city departments and projects in the area
- Constructible design for structure and roads

- The structure should be relatively maintenance free and minimize lifecycle cost.

3.2.3 Project Stakeholders

- Transportation Operations (Bridge Engineer) – Inspection and maintenance; want a bridge that is easy to inspect on an annual basis, maintenance is kept to a minimum and kept minor.
- Citizen interest groups
- City Council
- Non-government environmental stakeholders
- Parks and Community Services
- Utility companies

3.2.4 Project Constraints and Limitations

Project constraints were identified as below:

- Schedule: open bridge by February 2020
- Funding is limited
- Existing pedestrian traffic during construction

3.2.5 Project Schedule and Budget

Budget: $100 million

Schedule: Open for traffic at the end of 2020

3.2.6 Definition of Risk Impact

Quantification of risk factors can be done in three different levels as previously described in the CII IR-082. In this case study, we start with a Level 2 analysis which requires approximation tables for the impact of risk. Later on, we complete a Level 3 analysis which quantifies the risk impact by relating the risks to specific line items in the project cost estimate.

In order to quantify the risks using the Level 2 analysis, the budget and schedule impact levels which correspond to the magnitude labels must be identified, based on the perceived acceptable level of risk. This is done in collaboration with the owner. The budget and schedule impact levels which correspond to the impact labels were identified during the preliminary meetings as shown in Table 5.

Table 5- Impact interpretation table (in terms of costs for Level 2 analysis)

Verbal term	Budget Impact
Disastrous	$100 M
Severe	$50 M
Substantial	$10 M
Moderate	$2.5 M
Marginal	$1 M
Negligible	$0.1 M

3.3 Undertake a Facilitated Risk Analysis Workshop

The workshop objectives were to identify project risks, quantify them and to provide recommended approaches for mitigation.

3.3.1 Risk identification during workshop

There are 9 total risks being used for this example. The risks were divided into four categories, as follows:

- **Project Management:** risks which have to do with organization and administration of the project, including schedule, stakeholders, and communication.
- **External:** risks which are a result of the decisions or actions of parties not directly involved in the project (excluding communication risks).
- **Design:** risks which occur during or are a result of the design process.
- **Construction:** risks which occur during or are a result of construction.

The risk factors which were identified in the workshop are shown later in this subsection.

3.3.2 Risk Quantification using Level 1

The following tables (Table 6 and Table 7) were provided to participants to utilize as verbal expressions for the risks, if Level 1 quantification was to be undertaken.

Table 6: Likelihood expressions (for Level 1 analysis)

Verbal expression	Explanation
Highly Likely	HL -Highly Likely: Almost certain that it will happen, > 70%
Likely	LI -Likely: More than 50-50 chance 50% - 70%
Somewhat likely	SL -Somewhat Likely: Less than 50-50 chance 15%-50%
Unlikely	UN -Unlikely: Small likelihood but could well happen 1% - 15%
Very unlikely	VU -Very Unlikely: Not expected to happen 0.01% - 1%
Extremely unlikely	EU -Extremely Unlikely: Just possible but would be very surprising < 0.01%

Table 7: Impact expressions (in terms of description for Level 1 analysis)

Verbal expression	Explanation
Disastrous	The impact is totally unacceptable to the organization –value established in workshop or by owner.
Severe	Serious threat to the organization, public etc.
Substantial	Considerably affects cost
Moderate	Moderately affects costs
Marginal	Small effect on costs
Negligible	Trivial effect on costs

3.3.3 Risk quantification using Level 2

To quantify risk factors using Level 2 analysis the following was used:

- Determination of the likelihood of the factor being encountered (e.g. probability, or a subjective verbal expression) using Table 8.
- Determination of the magnitude of the impact if the factor is encountered (e.g. dollar value or a verbal expression) using Table 9 (in certain cases we may use supplementary tables as appropriate).
- Determination of the overall severity of the factor by multiplying the likelihood (1) by magnitude (2).
- The factors are then grouped based on the overall severity score according to the grouping in Table 10.

Table 8: Assessment of likelihood/probability of risk occurrence (AbouRizk S. M., 2009)

Verbal expression	Explanation	Low	High	Value to use to determine severity[3]
Highly Likely	HL -Highly Likely: Almost certain that it will happen, > 70%	0.7	1	100
Likely	LI -Likely: More than 50-50 chance 50% - 70%	0.5	0.7	50
Somewhat likely	SL -Somewhat Likely: Less than 50-50 chance 15%-50%	0.15	0.5	25
Unlikely	UN -Unlikely: Small likelihood but could well happen 1% - 15%	0.01	0.15	10
Very unlikely	VU -Very Unlikely: Not expected to happen 0.01% - 1%	0.0001	0.01	1
Extremely unlikely	EU -Extremely Unlikely: Just possible but would be very surprising < 0.01%	0.0000	0.0001	0.05

Table 9: Assessment of the impact of risk (AbouRizk S. M., 2009)

Verbal expression	Explanation	Budget Impact	Schedule Impact	Value to use to determine severity[2]
Disastrous	The impact is totally unacceptable to the organization –value established in workshop or by owner.	$100 M	3 seasons	700
Severe	Serious threat to the organization, public etc.	$50 M	2 season	200
Substantial	Considerably affects cost	$10 M	1 season	50
Moderate	Moderately affects costs	$2.5 M	6 months	15
Marginal	Small effect on costs	$1 M	3 months	5
Negligible	Trivial effect on costs	$0.1 M	1 month	1

[3] This value is arbitrary, utilized to undertake the ranking process of the risks based on AbouRizk (2009) works.

Table 10: Assessment of the consequence of a risk factor (AbouRizk S. M., 2009)

Total severity [4]	Category	Response
Over 10,000	Intolerable	Must eliminate or transfer risk, it may jeopardize the entire organization or its cost may be manifold that of the project.
5,001-10,000	Critical	Expected cost to the project is unacceptably high, (more than x% of the total project cost). This risk must be eliminated or transferred before proceeding with the project. Attempt to avoid or transfer risk.
1,001-5,000	Serious	Expected cost is high compared to total project cost. It is probably cost effective to eliminate or transfer this risk.
201-1,000	Important	Consider eliminating or transferring. If this risk is accepted, then it should be managed proactively.
26-200	Acceptable	Accept and manage.
0-25	Negligible	The expected cost of this risk is too small to justify any mitigation effort. Accept.

3.3.4 Summary of the risks resulting from the workshop:

In this workshop, 15 participants (4 females, and 11 males) went through the brainstorming process to determine potential risks and their causes for the project. They were asked to use the verbal expression provided to determine likelihood and magnitude of the identified risks. Next, participants described potential mitigation strategies. An example of the raw form of data gained is as follows:

- Risk: Project delayed by up to 2 seasons
- Cause: constructability challenges, water level, or adverse weather conditions
- Result: project delayed by 1 season or more, costs increased by 10 Million
- "unlikely" to occur
- Impact would be "substantial"

[4] Total severity is “value to use” for probability x “value to use” for impact

- Mitigation: Properly evaluate schedule through better means (i.e. Monte Carlo simulation, constructability review, traffic accommodations plan etc.)

The raw information gathered within the workshop leads to the following data.

3.3.5 Schedule Risk

Risk Factor
If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.
Category
Construction
Likelihood
Unlikely
Magnitude
Substantial
Severity
Important
Mitigation
See Table 11.

Table 11: Action items for risk factor mitigation (Schedule Risk)

Action	Responsible Party	Date
1. Conduct range estimating for the schedule	TRR	May 2018
2. Conduct constructability review and get a schedule estimate with contractors	TRR	June 2018
3. Establish a traffic accommodation and staging plan and confirm if it is doable	TRR	June 2018

3.3.6 Supply of Critical Material

Risk Factor
If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.

Category
Construction
Likelihood
Unlikely
Magnitude
Substantial
Severity
Important
Mitigation
See Table 12.

Table 12: Action items for risk factor mitigation (Supply of Critical Material Risk)

Action	Responsible Party	Date
1. Investigate supplier availability and cost	City	Dec 2018
2. Conduct escalation analysis	TRR	Dec 2018
2. Verify that material is ordered by deadline	TRR	During procurement

3.3.7 Material Cost

Risk Factor
If steel unit costs (material) are higher than estimated due to market escalations steel costs may increase by 10% to 20% (currently about $32M)
Category
Construction
Likelihood
Somewhat Likely
Magnitude
Substantial
Severity
Serious
Mitigation
See Table 13.

Table 13: Action items for risk factor mitigation (Material Cost Risk)

Action	Responsible Party	Date
1. Complete a range estimate to anticipate potential costs	TRR	May 2018
2. Investigate historical data and economic outlooks	TRR	May 2018

3.3.8 Labour Costs

Risk Factor
If steel erection costs (labour) are higher than estimated due to market escalation, costs may increase by 10% to 20% (currently about $8.1M)
Category
Construction
Likelihood
Somewhat Likely
Magnitude
Moderate
Severity
Important
Mitigation
See Table 14

Table 14: Action items for risk factor mitigation (Labor Cost Risk)

Action	Responsible Party	Date
1. Investigate historical data and economic outlooks	TRR	September 2018

3.3.9 Concrete Prices

Risk Factor
If concrete prices escalate by tender time, costs may increase from the estimate by 10% to 20%.
Category
Construction
Likelihood
Somewhat Likely
Magnitude
Moderate
Severity
Important
Mitigation
See Table 15

Table 15: Action items for risk factor mitigation (Concrete Prices Risk)

Action	Responsible Party	Date
1. Investigate historical data and economic outlooks	TRR	September 2018

3.3.10 Missing Estimate Items

Risk Factor
Potential for estimate missing major items or work packages, leading to a cost increase of 5% to 10%.
Category
Design
Likelihood
Unlikely
Magnitude
Moderate
Severity
Acceptable
Mitigation
Accept and manage.

3.3.11 Design Changes to thrust block

Risk Factor
If thrust block design changes to piles (currently design shows thrust blocks will work), costs will increase by $3M.
Category
Design
Likelihood
Unlikely
Magnitude
Moderate
Severity
Acceptable
Mitigation
Accept and manage.

3.3.12 Site Layout/Access/Parking/Transportation

Risk Factor
If the planned site layout, access, parking and material transportation is not resolved during design then the project cost will increase
Category
Design
Likelihood
Very Unlikely

Magnitude
Moderate
Severity
Negligible
Mitigation
Accept and manage.

3.3.13 Tight Schedule

Risk Factor
There is potential risk with the tight schedule.
Category
Design and Construction
Likelihood
Highly Likely
Magnitude
Substantial
Severity
Serious
Mitigation
See Table 16.

Table 16: Action items for risk factor mitigation (Tight Schedule Risk)

Action	Responsible Party	Date
1. Verify design has optimized construction processes	TRR	Dec 2018
2. Create critical path model and determine mitigation methods for schedule delay	TRR	Dec 2018

3.3.14 Risk Response

Once the identification and quantification is completed, the STRANT develops a response plan to all risk factors. This is described below:

- Decide on the actions to be taken in response to key risks. Actions can include:

- Reduce uncertainty by obtaining more information. (This generally leads to a re-evaluation of the likelihood and sometimes the magnitude.)
- Eliminate or avoid the risk factor through means such as partial or complete modifications to the proposed ideas, a different strategy or method etc.
- Transfer the risk element to other parties.
- Insure against the occurrence of the factor if and when possible.
- Abort the project if the risk is intolerable and no other means can be undertaken to mitigate its damages.
- Plan responses to key risks.
- Communicate the mitigation strategy and response plan to the risk review team.

Mitigation action items were identified for the high-risk items during the workshop itself; these action items were detailed further in the follow-up meeting. In addition, during the follow-up meeting all risk factors of severity "Important" or greater had mitigation actions identified, as well as those risk factors which were "Acceptable" or even "Negligible" but whose magnitude was "Substantial" or "Severe." After this process, the probability and impact of each risk can be re-evaluated using the same process.

3.3.15 Risk Monitoring and Control

After the mitigation actions were identified, STRANT Consulting collected the risks and actions into a proprietary special-purpose risk management software package, the STRANT Risk Management Program, categorized them, and developed a Risk Management Plan. The mitigation plan will be carried forward by STRANT Consulting and updated monthly.

The workshop objectives as defined by the participants were to identify the project risks and establish a risk mitigation plan. Nine risks were identified, ranging from Negligible to Serious. One risk was identified as Negligible, 2 risks were identified as Acceptable, 4 as Important, and 2 as Serious (see Table 17). There were no Critical or Intolerable risks.

Table 17: Overview of risks in each category

Severity	Number of Risk Factors
Negligible	**1**
Acceptable	**2**
Important	**4**
Serious	**2**
Critical	**0**
Intolerable	**0**

The serious risks which were identified were related to:

1. Tight Schedule
2. Material Cost (steel)

3.4 Post workshop analysis

3.4.1 Risk Quantification using Level 3:

Example risks that that were identified previously can now be analyzed using Level 3 Monte Carlo simulation. Nine risks are provided in Table 18 below. Each is associated with a Likelihood (as shown in Column G), the work package that this risk affects (as shown in Column I), and the cost of that work package (as given in Column J). Note that column J shows an Excel formula linking the cell to the spreadsheet shown in Table 20 which contains the estimate.

Table 18: Example Risks for Monte Carlo Simulation

A	D	G	I	J
	Risk factors as identified in workshop		Based on the latest project cost estimate	
	Risk Factor	**Likelihood**	**Work package**	**Current estimate**
1	If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.	U (10)	Total costs	=Estimate!F35
2	If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.	U (10)	Steel costs	=SUM(Estimate!F20:F21)
3	If steel unit costs (material) are higher than estimated due to market escalations steel costs may increase by 10% to 20%	SL (25)	Steel costs	=SUM(Estimate!F20:F21)
4	If steel erection costs (labor) are higher than estimated due to market escalation, costs may increase by 10% to 20%	SL (25)	Steel erection	=SUM(Estimate!F23:F26)
5	If concrete prices escalate by tender time, costs may increase from the estimate by 10% to 20%	SL (25)	Concrete	=SUM(Estimate!F28:F30)
6	Potential for estimate missing major items or work packages leading to a cost increase of 5% to 10%.	U (10)	Total costs	=Estimate!F35
7	If thrust block design changes to piles (currently design shows thrust blocks will work), costs will increase by $3M.	U (10)	Thrust blocks	=SUM(Estimate!F13)
8	If the planned site layout, access, parking and material transportation is not resolved during design then the project cost will increase.	VU (1)	Total costs	=Estimate!F35
9	There is potential risk with the tight schedule.	HL (100)	Total costs	=Estimate!F35

Table 19: Base Estimate Sheet

A	B	C	D	E	F
Basic component	Work package	LS / Unit Price	Unit	Quantity	Subtotal
Foundation work					
	EXCAVATION	$18.50	M3	25000	$462,500
	BACKFILL	$48.25	M3	23500	$1,133,875
	SHORING - NORTH ABUTMENT	$1,200.00	M2	2320	$2,784,000
	CIP CONCRETE PILES (750mm & 1200mm)	$28,500.00	EA	25	$712,500
Substructure work					
	THRUST BLOCK	$2,100.00	M3	2400	$5,040,000
	PILE CAP CONCRETE	$750.00	M3	380	$285,000
	ABUTMENT CONCRETE	$1,500.00	M3	350	$525,000
	WINGWALL CONCRETE	$2,100.00	M3	120	$252,000
	TIE BEAM CONCRETE	$1,200.00	M3	12	$14,400
	SUBSTRUCTURE REBAR (BLACK)	$2.15	KG	28800	$61,920
Supply of steel					
	FURNISH STEEL BEAMS	$5,188.16	MTON	2340	$12,140,294
	FURNISH ARCH RIB STEEL	$11,345.20	MTON	1710	$19,400,292
Erection of Steel					
	ERECT/PAINT STEEL BEAMS	$1,400.00	MTON	2880	$4,032,000
	ERECT/PAINT ARCH RIB STEEL	$2,100.00	MTON	1610	$3,381,000
	TEMPORARY ERECTION TOWER	$1,280,040.00	LS	1	$1,280,040
	HANGER SYSTEM (OPTION B - DSI)	$25.24	KG	75123	$1,896,105
Concrete work					

A	B	C	D	E	F
Basic component	**Work package**	**LS / Unit Price**	**Unit**	**Quantity**	**Subtotal**
	CONCRETE DECK	$2,100.00	M3	1750	$3,675,000
	APPROACH SLAB	$650.00	M3	85	$55,250
	SUPERSTRUCTURE REBAR (SS)	$10.25	KG	320050	$3,280,513
Miscellaneous items					
	EXPANSION JOINTS, Bearings, Asphalt overlay, walkways, deck drainage, signs etc.)	$2,100,875.00	M	1	$2,100,875
	BERM ACCESS	$110.00	M3	31000	$3,410,000

The probability for each risk factor is determined by using Table 8. Since the experts would have provided a linguistic term for likelihood as shown in Column G (Table 20 which is based on Table 8), we would simulate this by sampling a uniform random number from the range given in Table 8 for the value specified and use it as the argument for the Binomial distribution reflecting the probability of occurrence. For example a probability of 'unlikely' will be simulated by sampling from a Binomial distribution whose parameters are '1' (representing the number of trials) and 'Uniform (0.01, 0.15)' representing the probability of sampling this risk on this trial. This is shown in Table 20 below in Column K (note that the average value of the Uniform distribution being sampled is shown in this table).

The impact of the risk occurrence will be determined from having the experts describe the impact the risk will have on project costs as a percentage affecting a work package (or any other parameter that can be modeled). The impacts are given in Table 20 (Columns L, M, and N) reflects the impact in the best case scenario of the risk occurring, the impact in the most likely scenario and that in the worst case scenario respectively.

Table 20: Risk Quantification Example Continued

A	D	G	K	L	M	N
	Risk factors as identified in workshop			**If risk occurs it will increase cost of work package by:**		
	Risk Factor	**Likely**	**Probability**	**Low**	**ML**	**High**
1	If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.	U (10)	0.125	2%	5%	13%
2	If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.	U (10)	0.125	5%	12%	15%
3	If steel unit costs (material) are higher than estimated due to market escalations steel costs may increase by 10% to 20%	SL (25)	0.325	5%	8%	20%
4	If steel erection costs (labor) are higher than estimated due to market escalation, costs may increase by 10% to 20%	SL (25)	0.325	10%	25%	30%
5	If concrete prices escalate by tender time, costs may increase from the estimate by 10% to 20%	SL (25)	0.325	5%	12%	20%
6	Potential for estimate missing major items or work packages leading to a cost increase of 5% to 10%.	U (10)	0.125	2%	7%	10%
7	If thrust block design changes to piles (currently design shows thrust blocks will work), costs will increase by $3M.	U (10)	0.125	100%	100%	100%
8	If the planned site layout, access, parking and material transportation is not resolved during design then the project cost will increase.	VU (1)	0.05055	1%	5%	10%

9	There is potential risk with the tight schedule.	HL (100)	0.85	3%	5%	10%

The final model is summarized in Table 21 below. Column R shows that for each simulation iteration, a value is sampled from a Binomial distribution to determine if the risk occurs or not. If it does the Binomial distribution returns a value of "1" otherwise it will be a '0'. The result is multiplied by the impact which is in Column P as a sampled value from a triangular distribution of the percentage increase that is multiplied by Column J (the cost affected).

Table 21: Final Model for Monte Carlo Simulation of Risk Factors Example

A	D	J	K	P	R
	Risk factors as identified in workshop	**Based on the latest project cost estimate**			**Risk quantum**
	Risk Factor	**Current estimate**[5]	**Probability**[6]	**Impact**[7]	**Estimate x Probability x Impact**
1	If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.	=Estimate Value	=RiskUniform(0.1,0.15)	=RiskTriang (L1,M1,N1,)	= RiskBinomial (1,K1)*P1*J1
2	If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.	= Estimate Value	=RiskUniform(0.1,0.15)	=RiskTriang (L2,M2,N2,)	= RiskBinomial (1,K2)*P2*J2
3	If steel unit costs (material) are higher than estimated due to market escalations steel costs may increase by 10% to 20%	= Estimate Value	=RiskUniform(0.15, 0.5)	=RiskTriang (L3,M3,N3,)	= RiskBinomial (1,K3)*P3*J3
4	If steel erection costs (labor) are higher than estimated due to market escalation, costs may increase by 10% to 20%	= Estimate Value	=RiskUniform(0.15, 0.5)	=RiskTriang (L4,M4,N4,)	= RiskBinomial (1,K4)*P4*J4
5	If concrete prices escalate by tender time, costs may increase from the estimate by 10% to 20%	= Estimate Value	=RiskUniform(0.15, 0.5)	=RiskTriang (L5,M5,N5,)	= RiskBinomial(1, K5)*P5*J5
6	Potential for estimate missing major items or work packages leading to a cost increase of 5% to 10%.	= Estimate Value	=RiskUniform(0.1,0.15)	=RiskTriang (L6,M6,N6,)	= RiskBinomial (1,K6)*P6*J6

[5] This value represents a link to the estimated value for a given project subsection.
[6] This formula represents a uniform distribution with the parameters minimum and maximum used to represent probability when embedded in a binomial distribution in column R.
[7] This formula represents a triangular distribution with the parameters minimum, mean and maximum from Table 20.

A	D	J	K	P	R
	Risk factors as identified in workshop	**Based on the latest project cost estimate**			**Risk quantum**
	Risk Factor	**Current estimate**	**Probability**	**Impact**	**Estimate x Probability x Impact**
7	If thrust block design changes to piles (currently design shows thrust blocks will work), costs will increase by $3M.	= Estimate Value	=RiskUniform(0.1,0.15)	=RiskTriang (L10,M10,N10,)	= RiskBinomial (1,K10)*P10*J10
8	If the planned site layout, access, parking and material transportation is not resolved during design then the project cost will increase.	= Estimate Value	=RiskUniform(0.0011,0.1)	=RiskTriang (L11,M11,N11,)	= RiskBinomial (1,K11)*P11*J11
9	There is potential risk with the tight schedule.	= Estimate Value	=RiskUniform(0.7,1)	=RiskTriang (L12,M12,N12,)	= RiskBinomial (1,K12)*P12*J12

The simulation of the risks yields the risk profile shown in Figure 2 below. The uncertainty arising from potential risks on the project can account for somewhere between $0 and $30.2 Million. The expected value is $7.8Million. This value is used in determining the project's final estimate. If a full range estimate is carried out (i.e.: ranges are created for each work package, totalled and the distribution of the risk is added to the total distribution of the cost) the project's cost estimate will also be a distribution to which the risks are added. In this case the output distribution would look similar to the one shown in Figure 3. The estimate is between $60.3 Million and $93.0 Million with a mean value of 72.1 Million. The traditional, deterministic estimate, estimate stripped of all contingency for this project is $65.9 Million.

In general, the project manager, in consultation with the owner can determine the project's budget from the distribution in Figure 3 based on the risk tolerance and other consideration.

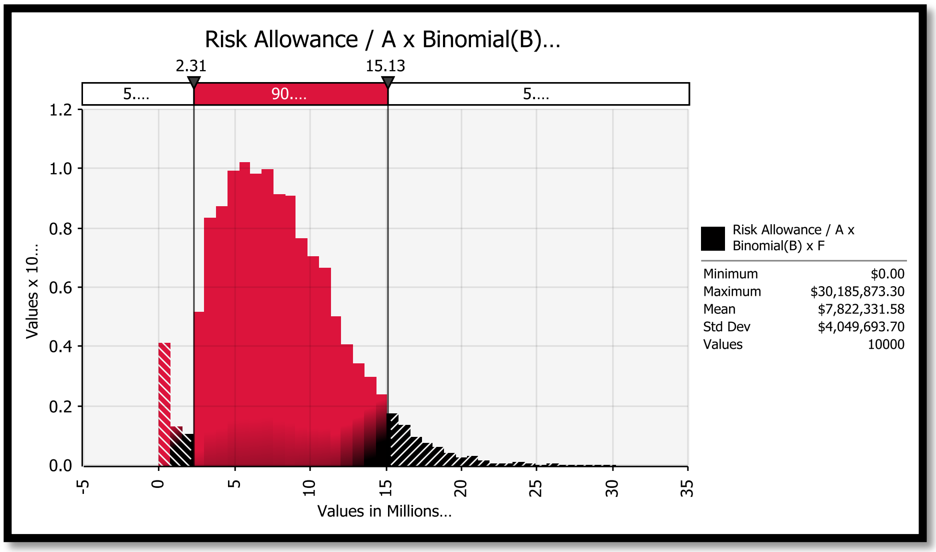


Figure 2: Risk profile for the bridge case study[8].

[8] Please note that the x-axis for a statistical probability density function as shown in this chart represents the random variable being modeled (e.g. in this chart it is the total project estimate). The y-axis represents the relative likelihood for this random variable to take on a given value (e.g. the likelihood of $60M is close to zero while the likelihood of $70M is close to 0.93^{-7}).
The scale of the y-axis depends on the frequency of values of the random variable at a given value of the random variable and is generally dependent on the number of observations collected.

In all subsequent charts of the probability density functions in this thesis, the x-axis represents the random variable (and its meaning is generally reflected in its title) while the y-axis will represent the likelihood of occurrence and its scale is not significant as it simply relates to the relative likelihood of the random variable –only its shape is important.

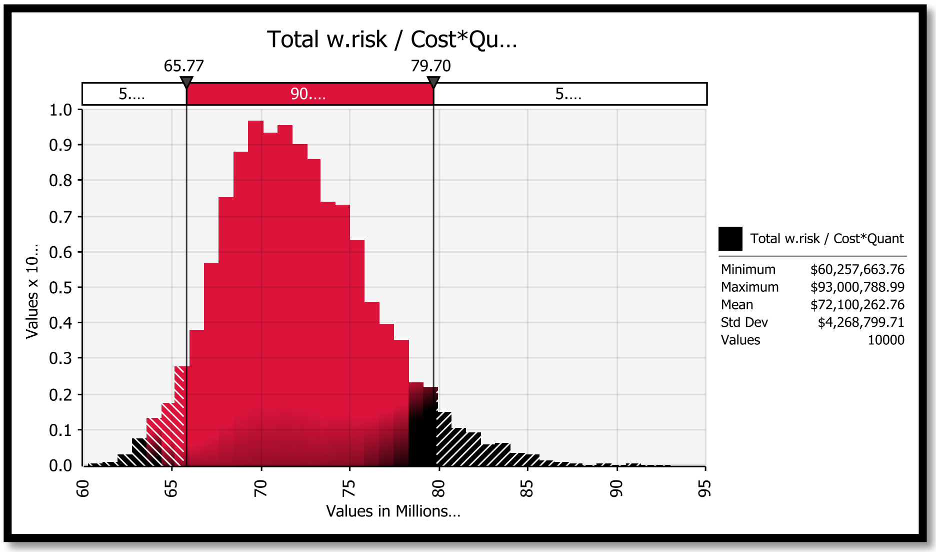


Figure 3: Total project estimate including the risk allowance determined in Figure 2

3.5 Risk Management

The risk register formalizes the risk response by defining specific tasks to be undertaken which will mitigate the risk, assigning responsibility and timelines for the tasks and following up on the risk factors on a regular basis until the project is complete. The tasks and responsible parties created form the basis of the risk updates. STRANT Consulting will continue to edit and revise all tasks created throughout the process. Task responsibility does not in any way imply risk allocation between parties. It is used to assign the responsibility to the project team member most suited for the task. The quantification determined in the previous section often determines the detail of mitigation that is required for this process. Similarly, the updates provided during risk management can allow for adjusted results in the simulation. The simulation can be re-run at many stages of the project to determine the updated risk allowance and project cost based on successful mitigation, or further unforeseen events. An example of the risk register for the project follows.

This classification will be used to determine and represent the Percent Complete column in the Risk Register:

- 5% Base work conducted
- 15% Task started
- 50% Progressing
- 75% Near completion
- 95% Review for completion
- 100% Complete

For purpose of brevity, a select number of the risk factors transferred into the risk management plan are given below for illustration in Tables 22 and 23.

Table 22: Risk Register for Risk 1

Risk ID	Description	Responsibility	Start Date	End Date	Percent Complete	Updates
Risk Category: Construction						
Risk 1: If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.						
1.1	Conduct range estimating for the schedule	TRR	May, 2018	June, 2018	0%	
1.2	Conduct constructability review and get a schedule estimate with contractors	TRR	June, 2018	July, 2018	0%	
1.3	Establish a traffic accommodation and staging plan and confirm if it's doable	TRR	June, 2018	August, 2018	0%	

Table 23: Risk Register for Risk 3

Risk ID	Description	Responsibility	Start Date	End Date	Percent Complete	Updates
Risk Category: Construction						
Risk 2: If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.						
2.1	Investigate supplier availability and cost	City	Dec, 2018	April, 2019	0%	
2.2	Conduct escalation analysis	TRR	Dec, 2018	April, 2019	0%	
2.3	Verify that material is ordered by deadline	TRR	During procurement	End of procurement	0%	

3.6 Concluding remarks

This chapter provided an overview of the risk analysis process through a case study that was slightly altered for confidentiality. The case study demonstrated the various elements of the risk analysis process discussed in Chapter 2. Through this example, the area of study can be understood in terms of a practical basis, particularly the use of verbal expressions and how they are fit into the QRA process. This section also provides a basis for Chapter 4 which investigates the influence of input data in the analysis.

4 Chapter 4 Understanding the Influence of Input Data

4.1 Introduction:

Regardless of whether one is undertaking Quantitative Risk Analysis (QRA), or building and experimenting with a Range Estimating (RE) Model, current practice (as well as existing literature) demonstrates that accurately modeling the input to a simulation model is critical to deriving accurate recommendations. While the scope of this paper is to address the process of QRA, RE tends to go hand in hand with this process as risk impact is often related to a portion of the cost estimate. While RE is meant to understand the "General Uncertainty" apparent in an estimate (i.e.: unit rates/quantity), QRA addresses the project risks. In order to understand the full impact of QRA on a project, it is important to understand it in the context of a full RE.

Input modeling in Monte Carlo Simulation receives significant attention in the simulation literature (Law, 2012); however there is a gap in the project risk analysis literature, as noted in Chapter 2, with regards to assessing probability distributions for risk. Most of the literature, in simulation, focuses on parameterizing a distribution (defining the parameters of a distribution) when historical data exist. AbouRizk (1990) reported on input modeling for construction simulation. In his work he deduced that using flexible distributions to model existing data was best suited to the data he collected. He further recommended the use of the Beta distribution family due to its flexibility in representing many data sets, as well as for its advantage of being readily available within most simulation software (in other words, using more powerful distributions that are not supported in simulation tools is generally not value adding). It seems that this logic can be utilized for furthering the process of QRA.

4.2 Understanding the Variables involved in QRA and RE

Given our focus on simulation as it applies to QRA and RE it would be prudent to first outline the random variables we are interested in modeling.

In general terms, the following variables are encountered in QRA and RE models:

1. Probability of a risk event (p),
2. Impact of a risk event on cost or schedule (I_c, I_s).
3. For a given scope of work:
    a. Unit costs (u),
    b. Quantity (q),
    c. Duration (d), and
    d. Escalation rates for costs over time (e)

Other basic variables can form part of a simulation model for QRA and RE but are not included in here, for brevity, especially since they can be treated in the same manner as the variables we outlined above. Such variables, for example, may include labour productivity, crew composition, hourly rates, and physical properties of work packages (i.e. dimensions) etc. which are used to generate the variables listed above.

In the next section we review the properties of these random variables and discuss how they are generally modeled in QRA and RE.

4.3 Modeling Random Variables Encountered in QRA and RE

4.3.1 Probability of a risk event p

The probability of a risk event occurring can be modeled using a Binomial distribution (Binomial (1, p)) (Palisade Corporation, 2010) (Megill, 1984). The Binomial distribution is used to model the number of successes in *n* independent draws/trials with probability *p* of success on each draw (Law & Kelton, 1991). The general form of the distribution is Binomial (n, p) where n is the number of draws and p is the probability of success from each draw. The binomial distribution is a discrete distribution (only integer values greater than or equal to zero) corresponding to the number of successes in the trial. This makes it appropriate for modelling the probability of a risk occurrence. For QRA and RE applications, Palisades (2013) states the following regarding the use of the Binomial distribution:

> "The most important modelling application is when n=1, so that there are two possible outcomes (0 or 1), where the 1 has a specified probability p, and the 0 has probability 1-p. With p=0.5, it is equivalent to tossing a fair coin. For other values of p, the distribution can be used to model event risk i.e. the occurrence or not of an event, and to transform registers of risks into simulation models in order to aggregate the risks."

As a result of the above, of interest in this paper, is the distribution 'Binomial' with parameters (1, p) where a risk event with a probability p=10% of occurring is modeled using Binomial (1, 0.1), for example. In such a case, we expect that for 100 simulation runs, we will observe the event (i.e. Binomial (1, 0.1) = 1) 10% of runs or 10 times out of 100 runs as shown in Figure 4.

To illustrate how the model works in QRA, consider that the risk event is R1 with a probability of 10% occurring. When the risk occurs, it will increase costs by 30% to 60% for the excavation work which is estimated at $100,000. To simulate the impact of the risk we simply do the following:

1. Excavation work = $100,000
2. Risk probability = Binomial (1, 0.1); Risk impact = Uniform (0.3, 0.6) x $100,000.
3. Expected value of risk occurrence (risk allowance) = Binomial (1, 0.1) x Uniform (0.3, 0.6) x $100,000.
4. Total expected cost of excavation work = Excavation work(1) + Expected value of risk occurrence (3)

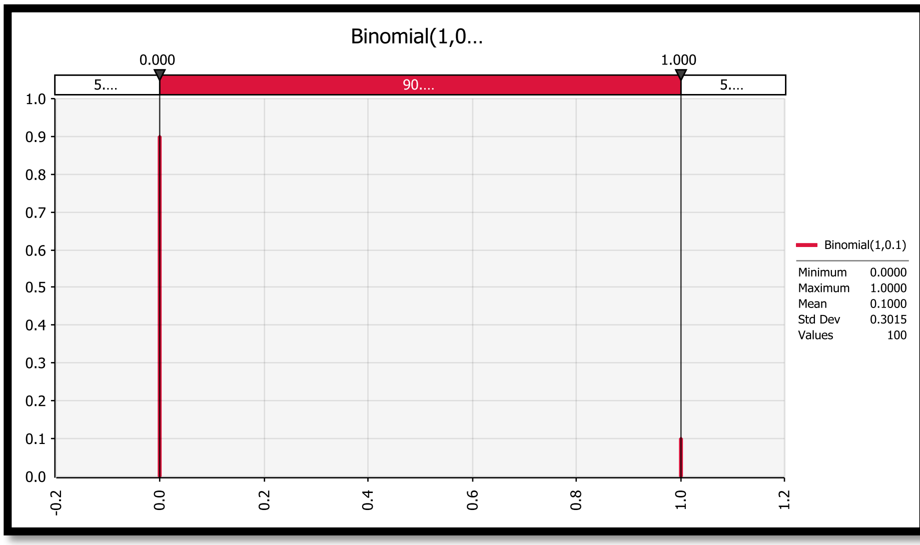


Figure 4: Binomial distribution B (1, 0.1) showing 100 runs

Figure 5 shows the probability as a Binomial (1, 0.1), Figure 6 shows the impact of the occurrence of the risk as a Uniform (30%, 60%) distribution of $100,000, Figure 7 shows the risk allowance and Figure 8 shows the total estimated cost. The risk contributes on average just under $4500 to the project's cost (Figure 7) and can be as low as $0 to as high as $60,000. Since the cost of the work package was fixed at $100,000, we notice from Figure 8 that the distribution of cost we are predicting ranges from $100,000 to approximately $160,000 (as expected from the numbers we used). The mean cost is $104,500.

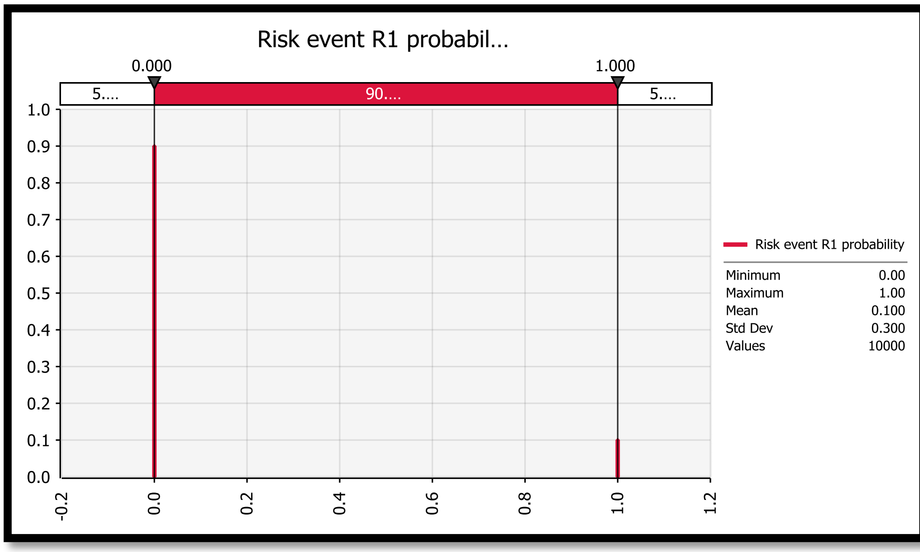


Figure 5: Probability of risk event occurring

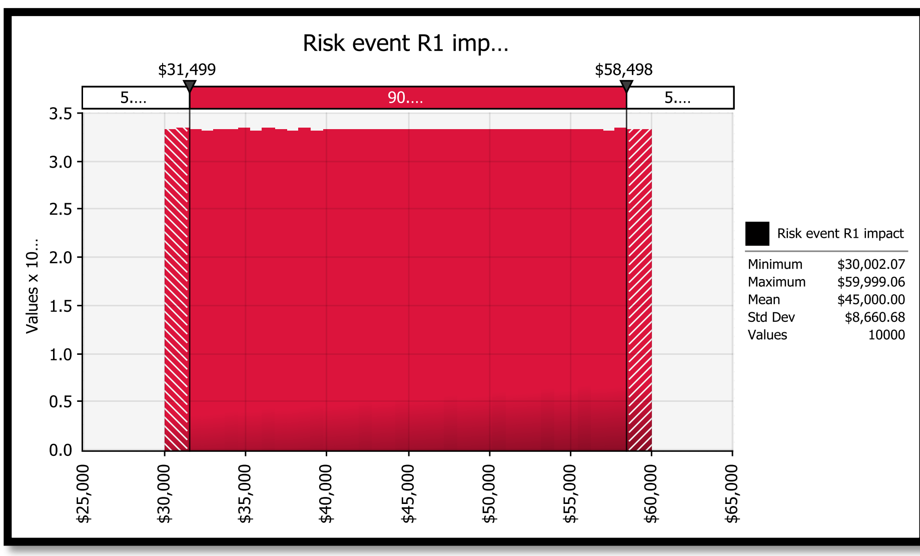


Figure 6: Impact of risk when it occurs

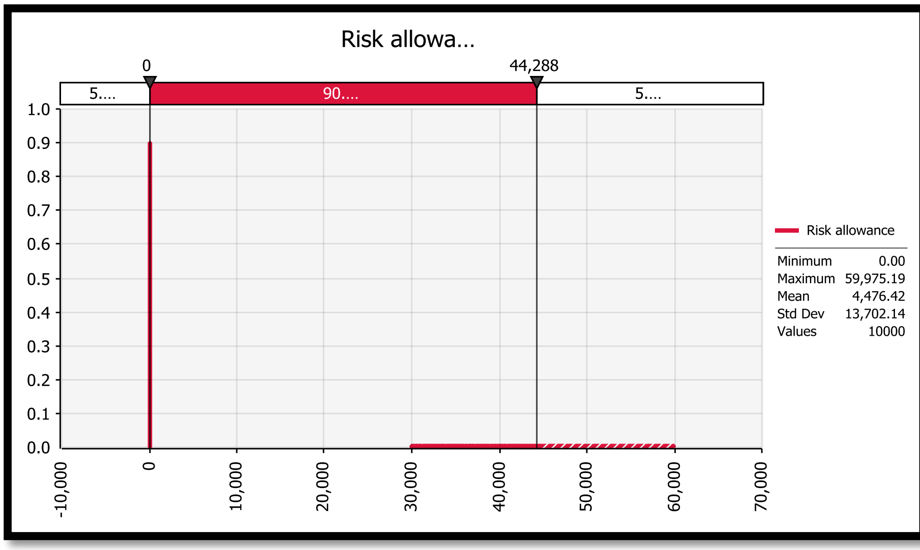


Figure 7: Risk allowance

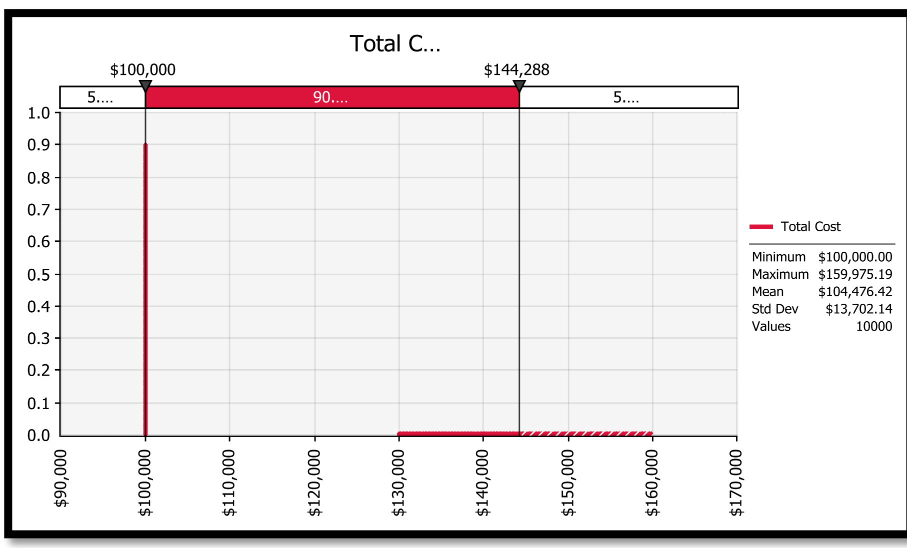


Figure 8: Total cost of work package

4.3.2 Impact of a risk event on cost or schedule (I_c, I_s).

Random variables that represent the impact of a risk event on cost or schedule (I_c, I_s), unit costs (u), quantities (q), duration (d), or escalation rate (e) are generally represented with a unimodal statistical distribution (i.e. a standard statistical distribution that has one defined mode). Comprehensive discussion of unimodal statistical distributions can be found in Law and Kelton, (1991). In this section we review the distributions that are most often used to represent impact in QRA, and represent general uncertainty in RE.

Triangular distribution: The triangular distribution with its three parameters, the minimum, mode, and maximum is given as Triangular (Min, Mode, Max) and is demonstrated in Figure 9. The triangular distribution is quite common in construction simulation and project risk analysis because it is simple to understand seems to work well in the absence of historical data. It only requires three parameters to define it and those parameters are generally readily available from experts as the most likely estimate of a variable, the pessimistic estimate and the optimistic estimate. It has been revealed however that triangular distributions, while simple to use, do not provide optimal results. Results in fact seem to create an upward bias in the probability of exceeding the most likely estimate, leading to a bias with a magnitude of about 20 percentage points (Chau, 1995). Chau (1995) suggests that while this distribution seems simple to use, it is not optimal for use in a Monte Carlo simulation.

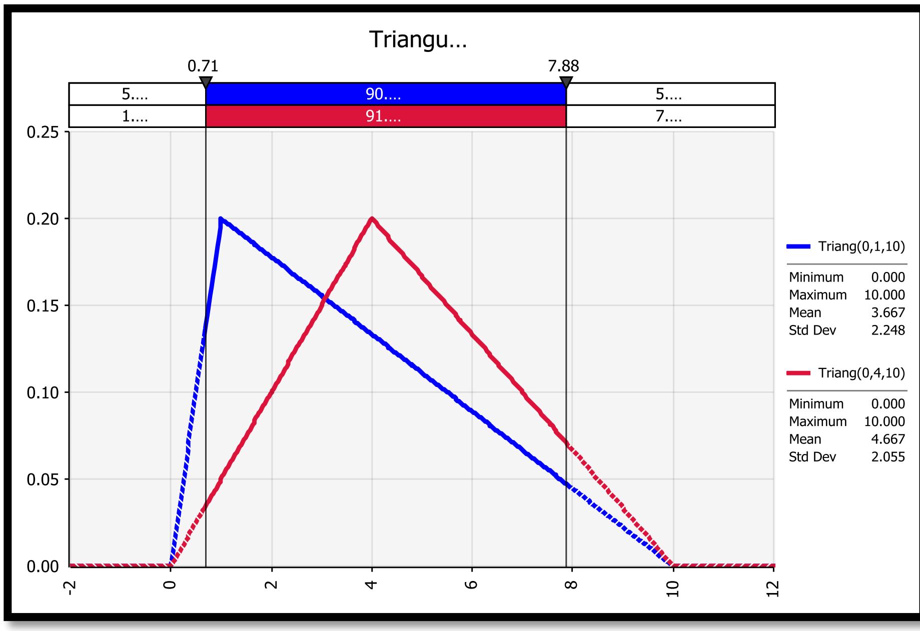


Figure 9: Triangular Distribution

Normal distribution: The Normal distribution is defined by its mean μ and variance σ2 and generally given as Normal (μ, σ2). The distribution is unbound and symmetric. The normal distribution is readily identifiable and understood by most risk analysts, project managers and engineers due to its significance in modeling many phenomena in statistics. The Normal distribution is demonstrated in Figure 10.

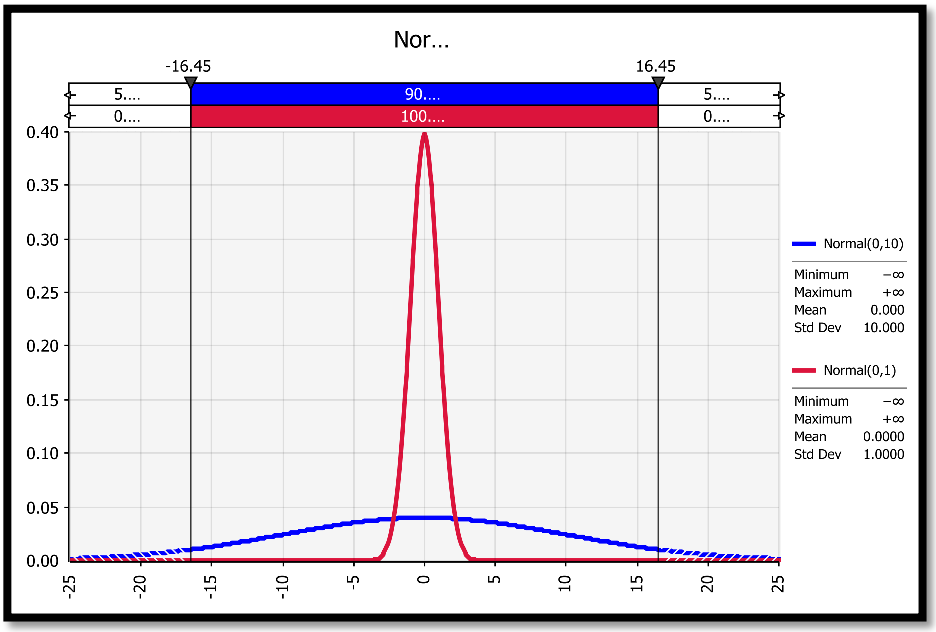


Figure 10: Normal Distribution

LogNormal: The lognormal distribution is defined by its mean μ and variance σ2 and generally given by LogNormal (μ, σ2). The LogNormal distribution is bound on the left side and unbound on the right side. The LogNormal distribution was used in construction simulation to model truck haul times (AbouRizk, 1990). "It is often used in the oil industry as a model of reserves following geological studies whose results are uncertain" (Palisades 2013). Sample LogNormal distributions are shown in Figure 11.

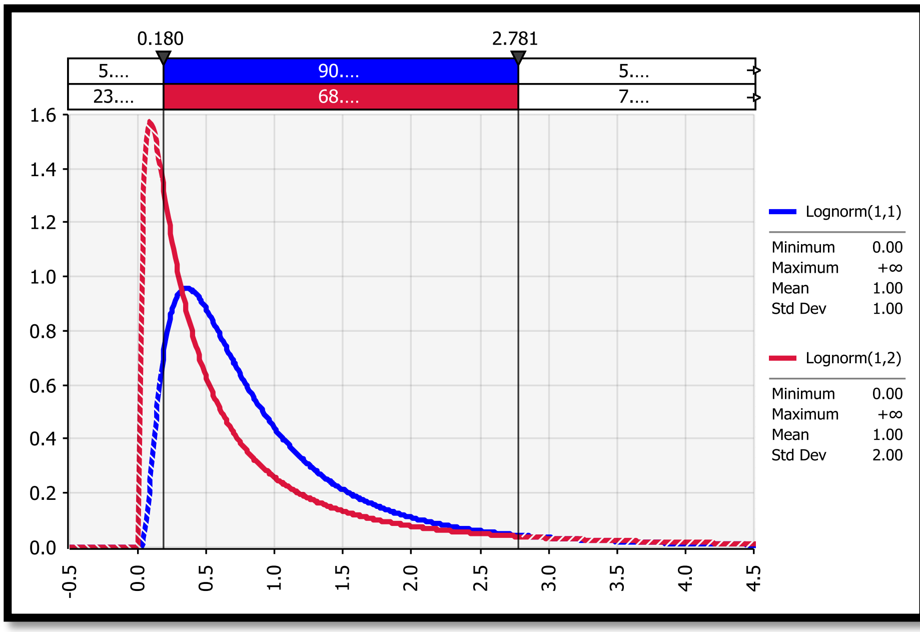


Figure 11: LogNormal distribution

Beta distribution: Beta (α, β, Min, Max). The beta distribution is defined by its lower and upper bounds (Min and Max) and its two shape parameters (α and β). The beta distribution was shown in AbouRizk (1990) to be effective in modeling construction duration data. Given its shape flexibility it is generally capable of modeling many random variables as demonstrated in Figure 12.

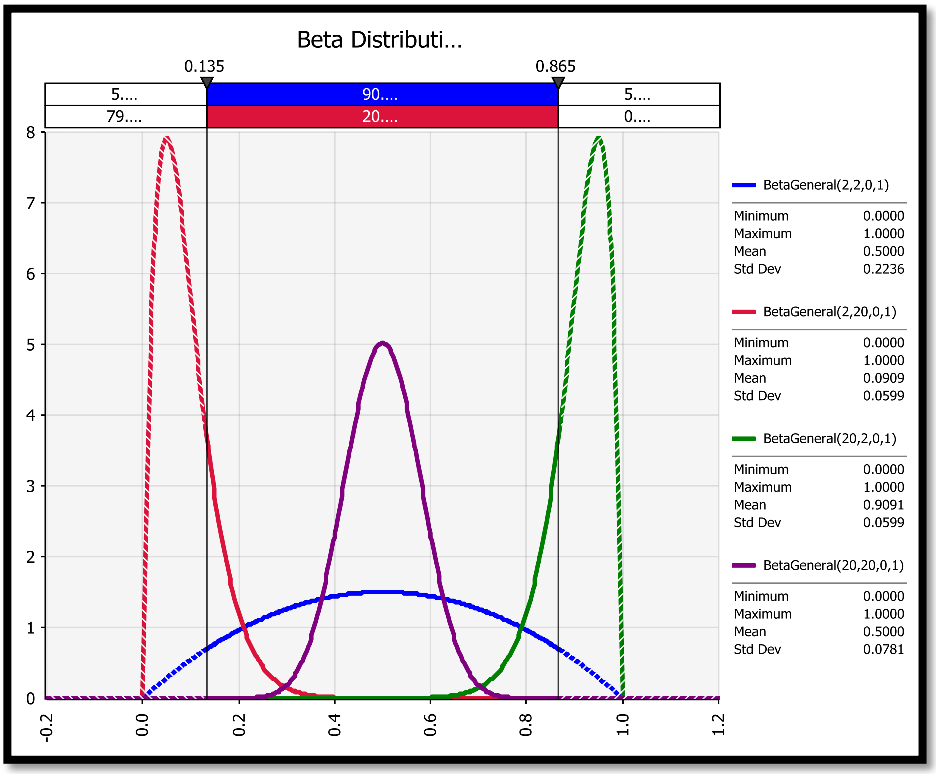


Figure 12: Beta distribution

4.4 Challenges with using statistical distributions to model random variables encountered in QRA and RE

4.4.1 Challenges with modeling probability of risk events:

In most practical applications of QRA or RE, historical data is generally not available to define the probability of various risk events. There are, at least two logical reasons for the lack of historical data of risk events encountered in construction projects:

1. Records of risk events are not standardized in the industry (definition of risk varies depending on who defines it) and, therefore, the definition of probability from historic data (if they existed) would not be practical.

2. The occurrences of risk events on a project are not generally publically shared or documented in a manner that lends itself to public dissemination (Galway, 2004). Although risk events can be inferred from changed orders, construction claims, and other project daily reports, efforts to collect such data will be monumental and convincing participants to share such information may be difficult.

4.4.2 Challenges with modeling impact of risk occurrence:

Defining statistical probability distributions for the impact of a risk occurring is quite challenging. In general terms, most QRA analysis will represent the impact of a risk occurring as either an absolute value of cost or duration (e.g. the impact of this risk is $1M, or three months delay) or a percentage impact to one or more component of the project's work packages' cost or schedule (e.g. the impact of this risk event if its occurs is a cost increase of somewhere between 10% and 20% of the estimated work of the underground work package). The impact has two variables: the distribution as a percentage and the work package estimated costs. Both of these can be random variables and neither would be normally supported with historical records simply because most work packages are uniquely defined for a given project due to lack of standardization of work package definitions, and due to uniqueness of the construction projects (generally one of a kind).

4.4.3 Modeling unit costs:

Defining statistical distributions for the unit costs can sometimes be supported with historical records (e.g. RS Means). Unfortunately distributions are not provided in this or other standards, and only specific statistics are given (e.g. mean, low or high values).

4.4.4 Modeling quantity:

Defining distributions for quantities greatly depends on the level of effort that went into the design and its details as well as the nature of the work package. For example, excavation quantities cannot be always be precisely determined from drawings due to over excavations, for example. Similarly the quantities of concrete for a slab on the ground cannot be neatly calculated as the ground is generally not even.

4.5 Understanding the Implications of Input Parameters:

The discussion in the previous section (see for example sections 2.3.3, 3.1, 3.3.3, Table 8 and Table 9) and the literature (Diekmann, 1983) notes that most QRA and RE studies depend on subjective estimation of input models. It is common to gain input for areas of uncertainty from an expert (as shown in Chapter 2), and as such the accuracy of reflecting an expert's opinion in the form of a statistical distribution in the simulation becomes very significant.

The objectives of the remainder of this chapter, therefore, is to try to understand the implications of selecting and parameterizing distributions (based on experts opinion) on the results of the QRA or RE simulation.

It is assumed that the solicitation of input from an expert is properly conducted (not within the scope of this paper) thus resulting is basic statistics and subjective indicators that are commonly acquired in QRA and RE studies including:

- Most pessimistic value for the random variable (high end point of a distribution)
- Most optimistic value for the random variable (low end point of a distribution)
- Most likely value for the random variable
- A judgement related to whether the distribution of the underlying random variable we are trying to estimate is skewed one way or the other (e.g. “I know the costs will be on the high side”)
- A judgement related to how spread the distribution of the underlying random variable is (e.g. “I am very confident of my estimate and, therefore, I believe that we should be close to my optimistic value”).
- The probability of a risk event occurring in the form of a verbal expression (e.g. it is unlikely that this event will occur).

The above information/statistics are generally used to define statistical distributions to model the random variables we previously identified. In this paper we limit the distributions we will investigate to the ones shown in the previous section and only as they may practically apply (e.g. using a Beta distribution or a Triangular distribution is more appropriate for cost and quantity input models because they are bounded between two limits as opposed to the Normal and lognormal since the two variables of costs and quantity have lower bounds and upper bounds). As noted in Section 4.3.1, using the

Binomial distributions is suitable for QRA (Palisade Corporation, 2010). The use for the binomial distribution in terms of probability is discussed in AbouRizk et. al. (2011).

The approach we will use to achieve our objective is a set of experiments which will be set up to investigate the effect of selecting and parameterizing input distributions for the random variables we identified on:

I. One element of the QRA or RE
II. An entire project's estimated risk allowance and its overall cost distribution (based on a real project).

The approach we will use is to set up Monte Carlo Simulation models and conduct sensitivity analyses as follows:

Analyzing input models for probability:

A limitation in assessing the probability of a risk event (likelihood) from subjective input solicited from an expert using subjective terms such as those given in AbouRizk (2011) is that not all individuals may understand the words provided to have the same meaning. A study conducted by Reagan, Mosteller and Youtz (1989), examined the accuracy of probabilities that are expressed in words, rather than numbers. They provided a set of terms and what they mostly mean to individuals which AbouRizk (2009) based his table on as shown in Table 24.

Table 24: Probability estimated from linguistic terms (AbouRizk, 2009)

Likelihood table	Low	High
HL --Highly Likely: Almost certain that it will happen, > 70%	0.7	1
LI --Likely: More than 50-50 chance 50% - 70%	0.5	0.7
SL (L) --Somewhat Likely: Less than 50-50 chance 15%-50%	0.15	0.5
UN --Unlikely: Small likelihood but could well happen 1% - 15%	0.01	0.15
VU --Very Unlikely: Not expected to happen 0.01% - 1%	0.0001	0.01
EU --Extremely Unlikely: Just possible but would be very surprising < 0.01%	0.0000	0.0001

1. Given that an expert has selected a term for the probability from Table 24 how much influence will there be (if any) on the evaluation of risk allowance if we use:
    a. The lower value of the range provided in Table 24 in the Binomial distribution
    b. The upper value of the range provided in Table 24 in the Binomial distribution
    c. The midpoint (commonly used as the best approximation)
    d. A random number from a uniform distribution to represent the range associated with the verbal expression being used.

What is the effect of the expert selecting one lower or higher term from Table 1 (i.e. examine the effect of the gradation in Table 24). If the expert provides one lower or higher range, what difference will there be?

Analysing input models for "impact"
In order to analyze the input models for "impact" determination in QRA or line item estimation in RE the following analysis will be conducted:

2. Will there be a difference in the values for "impact" if:
    a. a triangular distribution is used vs. a Beta distribution
    b. different shape parameters are used for a Beta distribution (in other words, the impact of reflecting additional judgements from the expert related to skewness and spread of data on the outcome)

Analysing the effects of input for QRA and RE within a Case Study
How will the input parameters used affect a full project estimate?

3. Analyze the input and model's effects (both probability and impact of occurrence) on the risk allowance for a typical project based on a real case study.

4.6 Experiment 1: Effect of using various input parameters for probabilities using a Binomial Distribution in QRA

Given the definition of a Binomial distribution, we normally use it to produce a 1 or a 0 in the QRA simulation where 1 represents the fact that a risk occurred while a 0 representing non-occurrence. Therefore for a Binomial (1, 0.1) we will observe 10% of the sampled Binomial values to be 1 while 90% are 0. The successful event with a '1' is multiplied by the impact of the risk factor resulting in a cost incurred (or a duration extension).

Given that in most QRA work we use large number of runs (e.g. 1,000) the impact of selecting a probability value for the Binomial distribution will be directly proportional to the probability itself. For the same example, if we have 1000 runs then we will get 0.1 x 1000 or 100 of those runs with a '1'. Likewise if we chose 0.25, we will get 250 runs with "1". Therefore the error introduced from incorrectly identifying the probability is directly proportional to the probability itself. This raises immediate questions regarding Table 24 particularly in the range of SL where the probability can be between 0.15 and 0.5 which can produce results differing by as much as 35% depending on what value is used from the table.

A way to deal with this (assuming that Table 24 is accurate) would be to either use the midpoint of the range or to sample a uniform value from that range during the experiment and then use that value as the input to the Binomial probability. The effects of this are demonstrated in Table 25 below, which is based on the output of the simulation of one work package with $1000 impact.

Table 25: Summary of the experiment demonstrating impact of selection of the Binomial parameters

	Binomial parameters (from probability table)						Mean value of Allowance			
Verbal expression	n	P(low)	P(high)	P(Mid)	Uniform	Impact ($)	P(low) ($)	P(high) ($)	P(Mid) ($)	Uniform($)
HL	1	0.7	1	0.85	0.85	1,000	700.0	1000.0	850.0	847.8
L	1	0.5	0.7	0.6	0.6	1,000	500.0	700.0	600.0	602.3
SL	1	0.15	0.5	0.325	0.325	1,000	150.0	500.0	325.0	323.4
U	1	0.01	0.15	0.08	0.08	1,000	10.0	150.0	80.0	80.9

VU	1	0.0001	0.01	0.0055	0.00505	1,000	0.1	10.0	5.0	5.2
EU	1	0	0.0001	0.0005	0.00005	1,000	0.0	0.1	0.1	0.0

Table 26: Output results of Experiment 1 for the Uniform distribution input

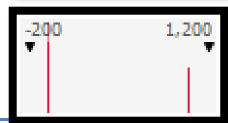
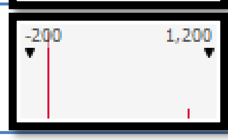
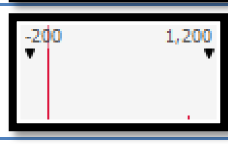
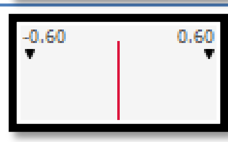
Name	Graph	Min	Mean	Max
HL / Unif		$0	$848	$1,000
L / Unif		$0	$602	$1,000
SL / Unif		$0	$323	$1,000
U / Unif		$0	$81	$1,000
VU / Unif		$0	$5	$1,000
EU / Unif		$0	$0	$0

Table 25 and Table 26 demonstrate how the mean value of the risk allowance varies depending on what is selected from the probability range (of Table 25). The following is readily noticeable:

1. When we choose to use the P (low), P (High), P (Mid) or P (Uniform), we get different mean values for the risk allowance. Since decisions are normally based on the distribution (and quite often on the mean value) once can conclude that the risk allowance is quite sensitive to the input chosen. Since Table 25 offers a

range of values, a simulationist is left with making a choice that will affect the output and produce consequences unintended by the person providing the input even within the same value of the verbal expression (e.g. HL has 0.7, 1, 0,85 and Uniform (0.7-1.0) resulting in means of 700, 1000, 850 and 847.8). Therefore, for a simulationist, the most plausible approach given the limitations of Table 25 is to use the P (uniform) since we only know the extent of the range of values for the probability and therefore any one number between those values should have a chance of being sampled in the simulation.

2. The output of risk allowance is sensitive to the choice of the verbal expression regardless of the values chosen for the simulation. For example if we chose Somewhat Likely instead of the Unlikely (expert cannot tell the difference, for example) the results can vary from a risk allowance of 150 to 10 (for P (Low)) and 323.4 to 80.9 (for P (Unif)), for example. The same is true for all other ranges.

   Since we used a unit of 1000 for simplicity, and since the Binomial distribution's probability will be proportional to the inputted parameter of probability, we can note that selecting a proper value to use to represent the range is important and also that selecting the range itself is also significant.

4.7 Experiment 2: Effect of using Triangular or Beta Distribution for Impact (in QRA)

Consider a simple example of a common excavation work package of 1000 m^3 estimated at $10/$m^3$. The subtotal cost for this package will be $10,000. The expert was asked about the uncertainty related to this estimate and she indicated that the unit price can be up to 10% lower and up to 20% higher than the $\$10/m^3$ she provided. The quantity is more certain but can be 5% lower and up to 10% higher than estimated. The example is simple to solve as the boundaries for the total costs can be established through multiplication for the low and high values as shown in Table 27 below:

Table 27: Boundaries for cost example

Work package	Lump Sum Unit Rate	Quantity	Subtotal	Unit Rate low	Unit Rate High	Quantity Low	Quantity High	Unit Rate x Quantity (Low)	Unit Rate x Quantity (High)
EXCAVATION	$10	1000	$10,000	$9	$12	950	1100	$8,550	$13,200

The values established in Table 27 are equivalent to using a Uniform distribution for the unit cost and quantities. A simulation with such input using a Uniform distribution (i.e.: uniform distribution of cost x uniform distribution of quantity) provides the cost distribution as shown in Figure 13.

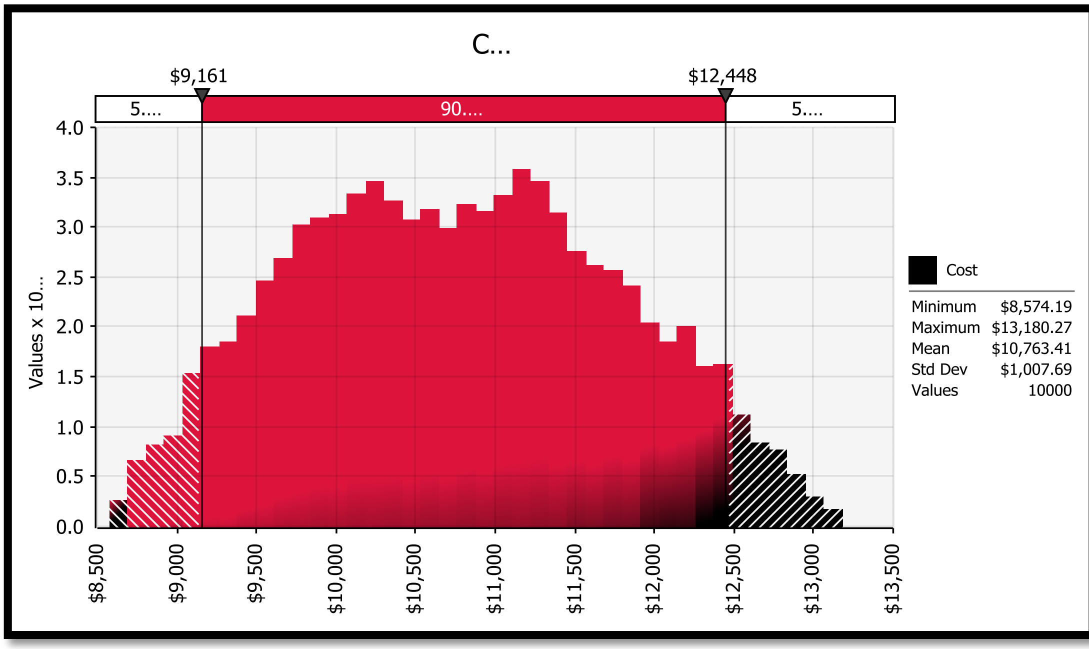


Figure 13: Boundaries for Cost using a Uniform Distribution Simulated

In order to simulate this scenario using a triangular distribution, assume that this distribution is used to define the input in three scenarios:

1. The most likely value is the one provided by the estimator ($10/m3) and the minimum and maximum values as discussed above for unit costs (-10%, +20%) and quantities (-5%, +10%).
2. The most likely values are near to the minimum values, with all other values remaining the same (i.e. positive skew).
3. The most likely values are near to the maximum values, with all other values remaining the same (i.e. negative skew).

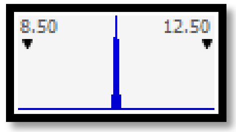
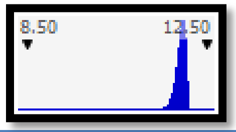
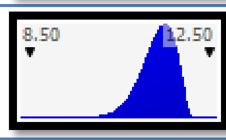
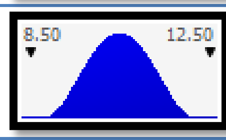
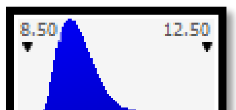
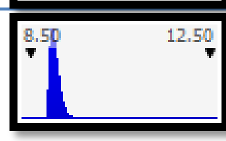
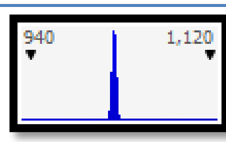
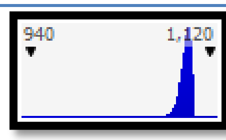
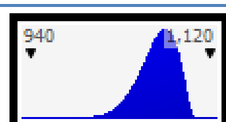
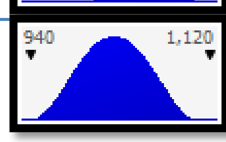
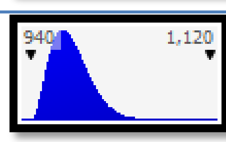
Input parameters are shown for these experiments in Figure 14.

In order to simulate this scenario using a beta distribution, assume that we used this distribution to reflect various properties as summarized in Table 28. Note that when $\alpha^1 = \alpha^2$, the distributions are symmetrical. When $\alpha^1 < \alpha^2$, the skew is positive, When $\alpha^1 > \alpha^2$, the skew is negative. The larger the α^1 and α^2 values, the higher the kurtosis.

Table 28: Distribution properties for a Beta Distribution

Distribution	α^1	α^2	α^1	α^2
Estimator is very certain about the most likely value she provided.	1000	1000	1000	1000
Estimator is certain that it would be on high side for cost and quantity	50	3	50	3
Estimator is somewhat certain that it would be on high side for cost and quantity	10	3	10	3
Estimator uncertain (could go either way)	3	3	3	3
Estimator is somewhat certain that it would be on low side for cost and quantity	3	10	3	10
Estimator is very certain that it would be on low side for cost and quantity	3	50	3	50

The input distribution reflecting the parameters for the uniform distributions, the triangular distributions and the beta distributions (Table 28) noted above are given in Figure 14.

Name	Graph	Min	Mean	Max
Beta (1000,1000)/Unit Price		$9.0	$10.5	$12.0
Beta (50,3)/Unit Price		$9.0	$11.8	$12.0
Beta (10, 3)/Unit Price		$9.0	$11.3	$12.0
Beta (3,3)/Unit Price		$9.0	$10.5	$12.0
Beta (3,10)/ Unit Price		$9.0	$9.7	$12.0
Beta (3,50)/Unit Price		$9.0	$9.2	$12.0
Beta (1000,1000)/ Quantity		950	1025	1100
Beta (50,3)/ Quantity		950	1092	1100
Beta (10,3)/ Quantity		950	1065	1100
Beta (3,3)/ Quantity		950	1025	1100
Beta (3,10)/ Quantity		950	985	1100

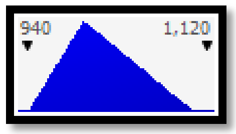
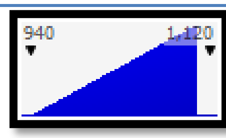
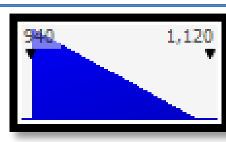
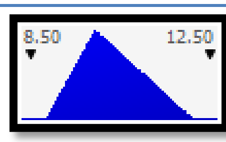
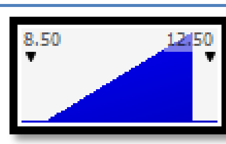
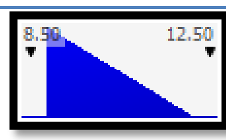
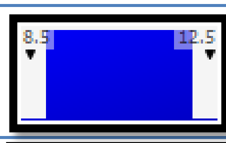
Name	Graph	Min	Mean	Max
Beta (3,50)/ Quantity		950	958	1100
Triangular Mid/ Quantity		950	1017	1100
Triangular Negative Skew/Quantity		950	1050	1100
Triangular Positive Skew/Quantity		950	1000	1100
Triangular Mid/Unit Price		$9.0	$10.3	$12.0
Triangular Negative Skew/Unit Price		$9.0	$11.0	$12.0
Triangular Positive Skew/Unit Price		$9.0	$10.0	$12.0
Uniform, Unit Price		9	10.5	12
Uniform, Quantity		950	1025	1100

Figure 14 Input Distributions for the experiment

The results of the simulation for the Scenarios noted above are summarized in Figure 15. They are also summarized in Table 29 and Figure 16. The results show how much of an impact the type of distribution and its parameters will have on a simple model. The following observations can be made:

- Figure 15 shows that there is a wide difference in simulation output resulting from using the same basic parameters as input from the expert, with different distributions and/or different parameters for each distribution chosen.

Figure 16 and Table 29 show, in numeric terms, the difference in values observed in the output. In particular: the triangular distributions generally result in a much wider spread of the results compared to choosing Beta distributions (except for when we chose (3, 3) as the two shape parameters. They also demonstrate that if we desire more control over how the data is spread, the Uniform and Triangular distributions will not be practical to use and as suspected, a flexible distribution like the Beta distribution allows us to better represent the skewness of data and its spread as shown in Figure 15.

In summary, even though we started with soliciting three parameters from the expert (minimum, maximum and most likely) and augmented those with some judgement regarding the underlying distribution the estimator has in mind, we arrived a differing conclusions based on what we selected in a distribution type and its parameters. Therefore, it is important to structure the process of solicitation and specification of both distributions and parameters to ensure that what we are observing in the output reflects what was specified in the input.

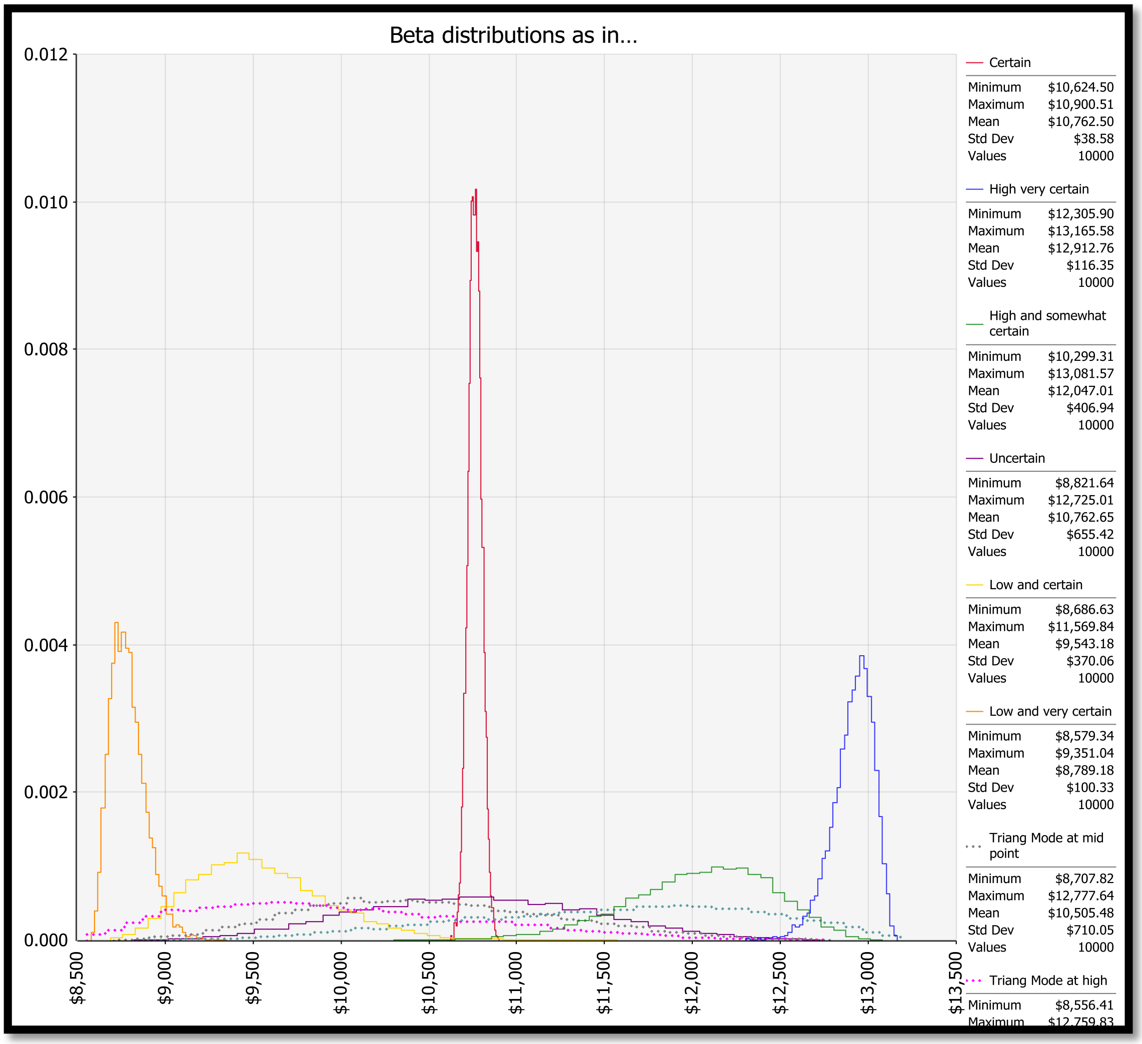


Figure 15: Resulting output distributions from the experiment

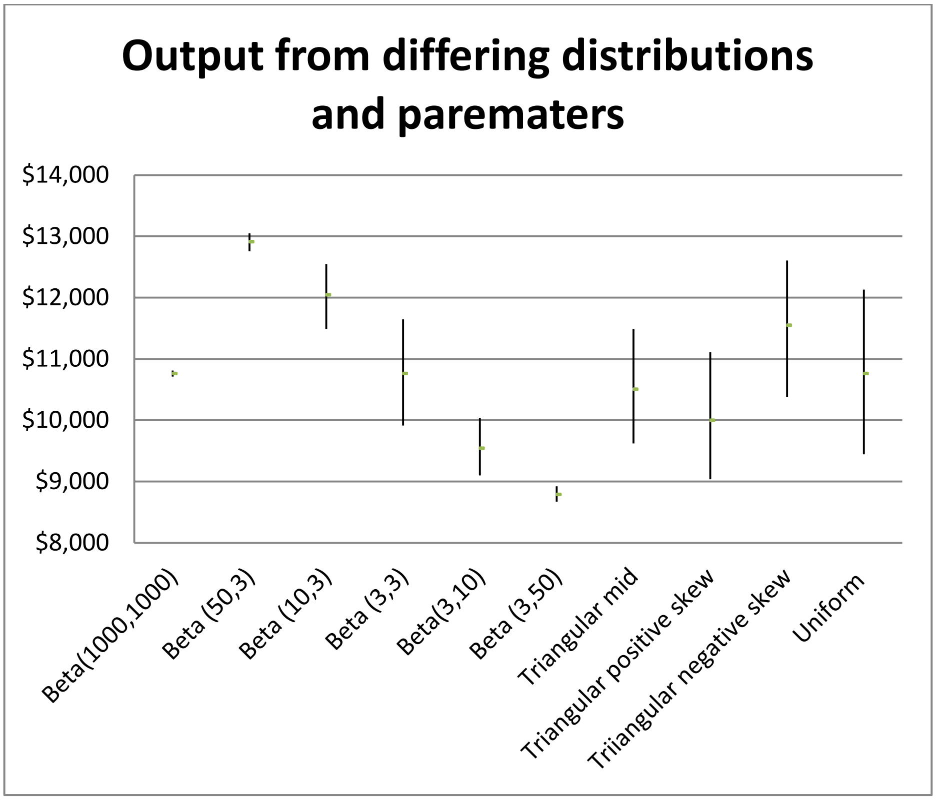


Figure 16: Box plot showing the mean and percentiles of the resulting output from Figure 15

Table 29: Summary of the results

Distribution	90th percentile	10th percentile	Mean	Min	Max	Spread 10th-90th percentiles	Mean-95	Mean-10	Standard Deviation
Beta(1000,1000)	$10,812	$10,713	$10,762	$10,625	$10,901	$99	$50	$49	$39
Beta (50,3)	$13,047	$12,756	$12,913	$12,306	$13,166	$291	$134	$157	$116
Beta (10,3)	$12,544	$11,497	$12,047	$10,299	$13,082	$1,047	$497	$550	$407
Beta (3,3)	$11,628	$9,911	$10,763	$8,822	$12,725	$1,717	$865	$851	$655
Beta(3,10)	$10,054	$9,096	$9,543	$8,687	$11,570	$958	$510	$447	$370
Beta (3,50)	$8,924	$8,672	$8,789	$8,579	$9,351	$252	$135	$117	$100
Triangular Mid	$11,486	$9,623	$10,505	$8,708	$12,778	$1,863	$980	$883	$710
Triangular Positive Skew	$11,128	$9,033	$10,000	$8,556	$12,760	$2,095	$1,128	$967	$793
Triangular Negative Skew	$12,605	$10,366	$11,550	$8,870	$13,183	$2,238	$1,054	$1,184	$840
Uniform	$12,139	$9,439	$10,763	$8,589	$13,193	$2,700	$1,376	$1,325	$1,005

4.8 Experiment 3: Effect of input models on the risk allowance derivation for a project

The bridge project, based on a real case study, used in introducing the concepts of risk analysis has an engineer's estimate at the preliminary design level that is summarized in Table 30 below[9]. The risk factors identified for the project (abridged over here for brevity to nine factors) are given in Table 31.

[9] Please note that the estimate has been slightly altered in the work breakdown structure and that all numbers have been scaled to preserve confidentiality of the data. The changes do not affect the experiment that is presented in this chapter, however.

Table 30: Estimate for Bridget Project

A	B	C	D	E	F	G	H	I	J	Q
Basic component	Work package	LS / Unit Price	Unit	Quantity	Subtotal	Unit price low	Unit Price High	Quantity Low	Quantity High	Cost*Quantity
Foundation work										
	EXCAVATION	$ 18.50	M3	25000	$462,500	85%	110%	90%	115%	$462,211
	BACKFILL	$48.25	M3	23500	$1,133,875	85%	110%	90%	115%	$1,133,166
	SHORING - NORTH ABUTMENT	$1,200.00	M2	2320	$2,784,000	90%	115%	95%	105%	$2,853,600
	CIP CONCRETE PILES (750mm & 1200mm)	$28,500.00	EA	25	$712,500	85%	110%	100%	100%	$694,688
Substructure work										
	THRUST BLOCK	$ 2,100.00	M3	2400	$5,040,000	80%	110%	85%	115%	$4,788,000
	PILE CAP CONCRETE	$ 750.00	M3	380	$285,000	80%	110%	85%	115%	$270,750
	ABUTMENT CONCRETE	$1,500.00	M3	350	$525,000	80%	110%	85%	115%	$498,750
	WINGWALL CONCRETE	$2,100.00	M3	120	$252,000	80%	110%	85%	115%	$239,400
	TIE BEAM CONCRETE	$1,200.00	M3	12	$14,400	80%	110%	85%	115%	$13,680
	SUBSTRUCTURE REBAR (BLACK)	$2.15	KG	28800	$61,920	80%	110%	85%	115%	$58,824
Supply of steel										
	FURNISH STEEL BEAMS	$5,188.16	MTON	2340	$12,140,294	95%	105%	85%	110%	$11,836,787
	FURNISH ARCH RIB STEEL	$11,345.20	MTON	1710	$19,400,292	95%	105%	85%	110%	$18,915,285

Erection of Steel										
	ERECT/PAINT STEEL BEAMS	$ 1,400.00	MTON	2880	$4,032,000	90%	105%	85%	110%	$3,832,920
	ERECT/PAINT ARCH RIB STEEL	$2,100.00	MTON	1610	$3,381,000	90%	105%	85%	110%	$3,214,063
	TEMPORARY ERECTION TOWER	$1,280,040.00	LS	1	$1,280,040	90%	105%	85%	110%	$1,216,838
	HANGER SYSTEM (OPTION B - DSI)	$ 25.24	KG	75123	$1,896,105	90%	105%	85%	110%	$1,802,484
Concrete work										
	CONCRETE DECK	$2,100.00	M3	1750	$3,675,000	80%	110%	95%	105%	$3,491,250
	APPROACH SLAB	$650.00	M3	85	$55,250	80%	110%	95%	105%	$52,488
	SUPERSTRUCTURE REBAR (SS)	$10.25	KG	320050	$3,280,513	80%	110%	95%	105%	$3,116,487
Miscellaneous items										
	EXPANSION JOINTS, Bearings, Asphalt overlay, walkways, deck drainage, signs etc.)	$2,100,875.00	M	1	$2,100,875	95%	105%	85%	125%	$2,205,919
	BERM ACCESS	$110.00	M3	31000	$3,410,000	95%	105%	85%	125%	$3,580,500
				Total	**$65,922,563**					
									Total	**$64,278,089**
									Risk Allowance	**$6,891,814**
									Project cost	**$71,169,903**

Table 31: Risk Factors for Bridge Project

No.	Risk Factor	Likelihood	Magnitude	Work package	A Current estimate	B Probability	C Low	D Most Likely	E High	I A*B*G
1	If the schedule for construction is not met due to constructability challenges, water level, or adverse weather conditions then the project will be delayed and its costs may increase due to claims.	U (10)	SUB (50)	Total costs	$65,922,563	0.08	2%	5%	13%	$351,587.00
2	If supply of critical material is delayed by economic or other circumstances, then the project will be delayed, and costs may be increased.	U (10)	SUB (50)	Steel costs	$31,540,586	0.08	5%	12%	15%	$269,146.34
3	If steel unit costs (material) are higher than estimated due to market escalations steel costs may increase by 10% to 20% (currently about $32M)	SL (25)	SUB (50)	Steel costs	$31,540,586	0.325	5%	8%	20%	$1,127,575.9
4	If steel erection costs (labour) are higher than estimated due to market escalation, costs may increase by 10% to 20% (currently about $8.1M)	SL (25)	MOD (15)	Steel erection	$10,589,145	0.325	10%	25%	30%	$745,652.26
5	If concrete prices escalate by tender time, costs may increase from the estimate by 10% to 20%	SL (25)	MOD (15)	Concrete	$7,010,763	0.325	5%	12%	20%	$281,014.73

No.	Risk Factor	Likelihood	Magnitude	Work package	A Current estimate	B Probability	C Low	D Most Likely	E High	I A*B*G
6	Potential for estimate missing major items or work packages leading to a cost increase of 5% to 10%.	U (10)	MOD (15)	Total costs	$65,922,563	0.08	2%	7%	10%	$334,007.65
7	If thrust block design changes to piles (currently design shows thrust blocks will work), costs will increase by $3M.	U (10)	MOD (15)	Thrust blocks	$5,040,000	0.08	100%	100%	100%	$403,200.00
8	If the planned site layout, access, parking and material transportation is not resolved during design then the project cost will increase.	VU (1)	MOD (15)	Total costs	$65,922,563	0.005	1%	5%	10%	$17,579.35
9	There is potential risk with the tight schedule. The current schedule indicates the tendering is in October 2012 and construction finishes in November 2014. The tender award time will be most likely in winter 2012, and construction cannot be started right away in winter. This poses potential risks to delivering project on time.	HL (100)	SUB (50)	Total costs	$65,922,563	0.85	3%	5%	10%	$3,362,050.7
						Risk Allowance				**$6,891,814**

The approach used above calculates the risk impact as a distribution of costs.

Assumptions made:

1. Risks on average occur (e.g. scope change will take place and its impact will be 3%, 5%, 10%, see Table 31.)
2. Base costs are deterministic. (i.e. total capital will be $65.9M)
3. Use triangular distributions for impact percent. Then simulate by multiplying this impact (Triangular) by the Deterministic cost to get the risk allowance for each line item
4. Sum up the risk allowances for each item to get the total risk allowance to add to the estimate (e.g. 65.9M+7.5M) to get final cost. In here the distribution of risk is 0M to 25M but we use the mean in the analysis

Notes and limitations:

1. There is no need to actually simulate if all that is required is the mean value of the risk allowance. Just use the mean value for the triangular distribution (L+M+H)/3, multiply by deterministic cost and get sum as in Column I)
2. The deterministic values have inherited uncertainty that can change the above outcome, beyond the risk occurrence.
3. The appropriate methods would be to model the base cost as a distribution reflecting the uncertainty with the exception of risk, then model the risk occurrence as a probability (e.g. binomial), and an impact as a distribution to get

the full impact then add the base costs for base distribution and the risk allowance as a secondary distribution.

The following experiments were undertaken to study the sensitivity of the decision making statistics of total project costs and the risk allowance to errors introduced in 1) the transfer of the likelihood verbal expression to probabilities in the risk analysis and 2) the shape of the distributions of the unit rates and quantities for the estimate.

The distribution of the costs without the risk allowance is summarized in Figure 17 below:

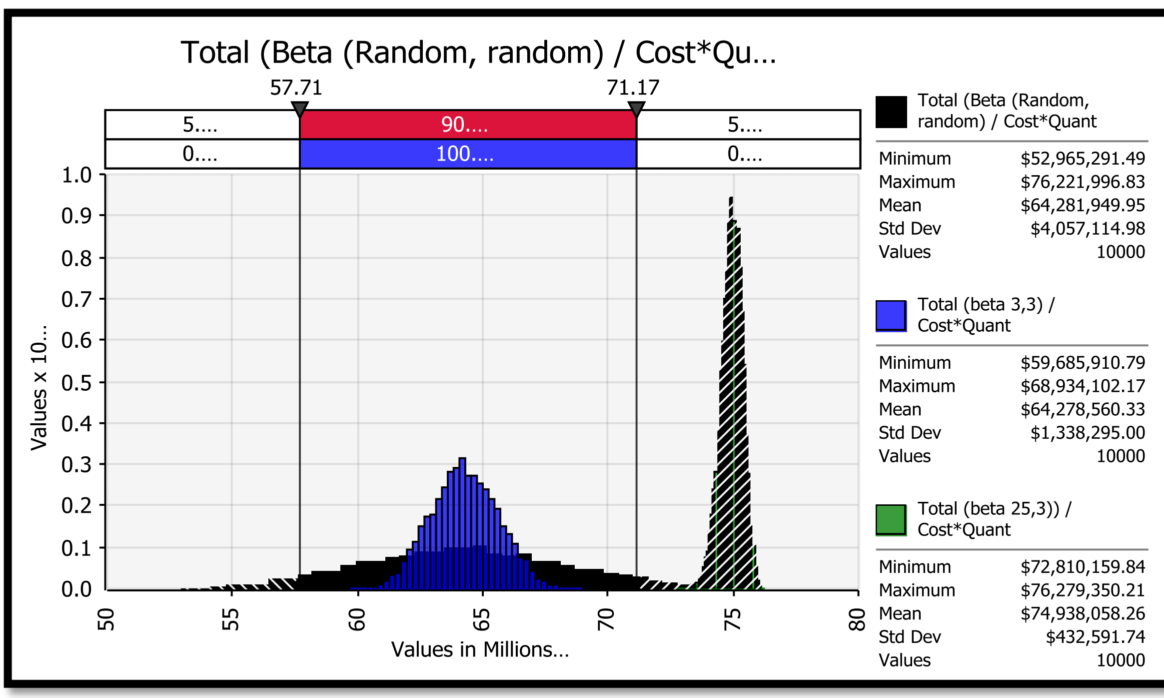


Figure 17: Distribution of Costs using differing input parameters

The distributions of the costs include:

1. The base case is shown in blue where all work package costs had a symmetric distribution. The Beta distribution parameters being (α_1= 3 and α_2 = 3).
2. The green distribution represents the case where all work package costs were very negatively skewed. The Beta distribution parameters being (α_1= 3 and α_2 = 25).
3. The red distribution represents the case where we sample the shape of the distribution randomly on each of the 10,000 iterations between the values of the Beta distribution parameters being (α_1= 2 and α_2 = 50.)

The results are as expected, with the base case showing a moderate normal spread with a bell shaped curve, the skewed distributions with high work package costs showing a high cost distribution and the randomly sampled distribution shapes providing a significantly wider spread.

The experiments are summarized in Table 32 below: All risks have the same effect as in the original case study in Table 25, the likelihood is as shown in Table 25, but when quantified we use differing distributions for “p” in Binomial (1,p) as follows:

Table 32: Experiments to Study Sensitivity

No.	Range Estimate	Risk Analysis
1.a	Base estimate with all work packages having a distribution that is symmetric around the mean (Beta (3,3))	Binomial distribution is used for the likelihood as follows: Binomial (1, Midpoint of the values related to chosen verbal expression based on Table 25). *This is typically what has been used.*
1.b	Base estimate with all work packages having a distribution that is symmetric around the mean (Beta (3,3))	Binomial distribution is used for the likelihood as follows: Binomial (1, Uniform (low, high of the Range chosen for the verbal expression based on Table 25)). *This should be the approach used since we do know that the range is what the expert was presented with (or had in mind).*
2.a	Base estimate with all work packages having a distribution that is skewed with (Beta (25,3))	Binomial distribution is used for the likelihood as follows: Binomial (1, Midpoint of the values related to chosen verbal expression based on Table 25).
2.b	Base estimate with all work packages having a distribution that is skewed with (Beta (25,3))	Binomial distribution is used for the likelihood as follows: Binomial (1, Uniform (low, high of the Range chosen for the verbal expression based on Table 25)).
3.a	Base estimate with all work packages having a distribution that is symmetric around the mean with shape parameters random (Beta (Uniform 2, 50),Uniform 2,50)) *This allows the beta distribution definition to be random –a point of comparison.*	Binomial distribution is used for the likelihood as follows: Binomial (1, Midpoint of the values related to chosen verbal expression based on Table 25).
3.b	Base estimate with all work packages having a distribution that is symmetric around the mean (Beta (Uniform (2, 50), Uniform (2, 50)). *This allows the beta distribution definition to be random –a point of comparison.*	Binomial distribution is used for the likelihood as follows: Binomial (1, Uniform (low, high of the Range given for the chosen verbal expression based on Table 25)).

4.9 Experiment 1.a and 1.b:

The impact to project cost (through a change in risk allowance) when "p" in the Binomial distribution (1, p) is a midpoint value or a uniform distribution, is shown in Figure 18 and Figure 19 below. Ranges were added to the entire base estimate using a distribution Beta (Low, High, 3, 3), a symmetric distribution.

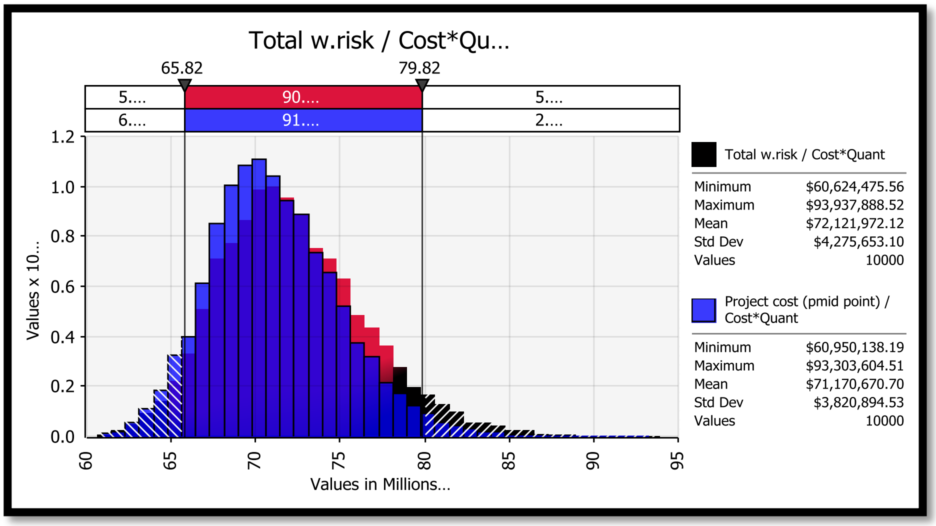


Figure 18: Impact to Project Cost inclusive of Risk Allowance Experiment 1

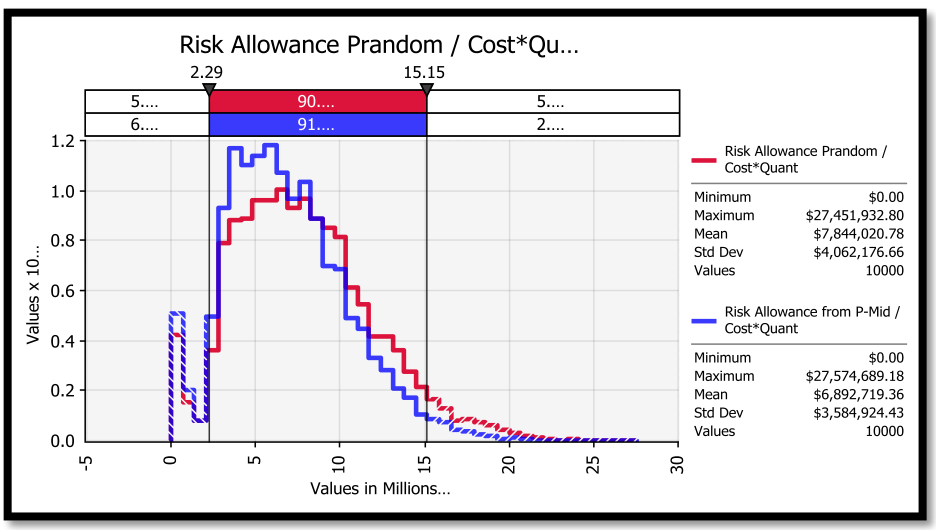


Figure 19: Impact to Risk Allowance Experiment 1

4.10 Experiment 2.a and 2.b:

The impact to project cost (through a change in risk allowance) when "p" in the Binomial distribution (1, p) is a midpoint value or a uniform distribution, is shown in Figure 20 and Figure 21 below. Ranges were added to the entire base estimate using a distribution Beta (Low, High, 25, 3), a distribution skewed right.

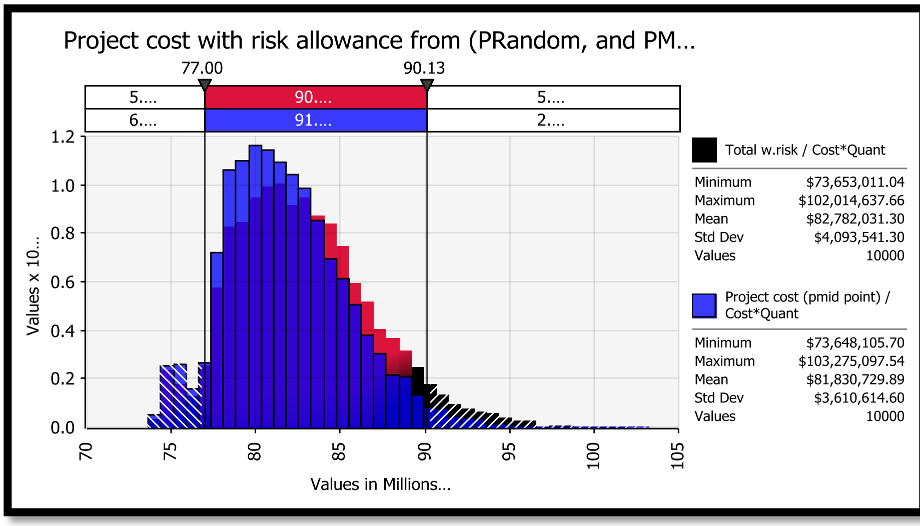


Figure 20: Impact to Project Cost inclusive of Risk Allowance Experiment 2

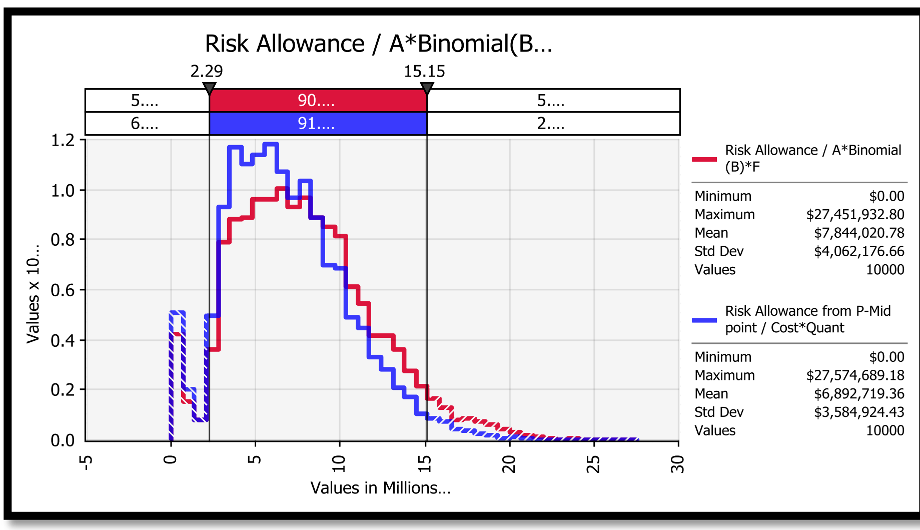


Figure 21: Impact to Risk Allowance Experiment 2

4.11 Experiment 3.a and 3.b:

The impact to project cost (through a change in risk allowance) when “p” in the Binomial distribution (1, p) is a midpoint value or a uniform distribution, is shown in Figure 22 and Figure 23 below. Ranges were added to the entire base estimate using a distribution Beta

(Low, High, Uniform (2, 50), Uniform (2, 50)), a distribution which is composed of random values.

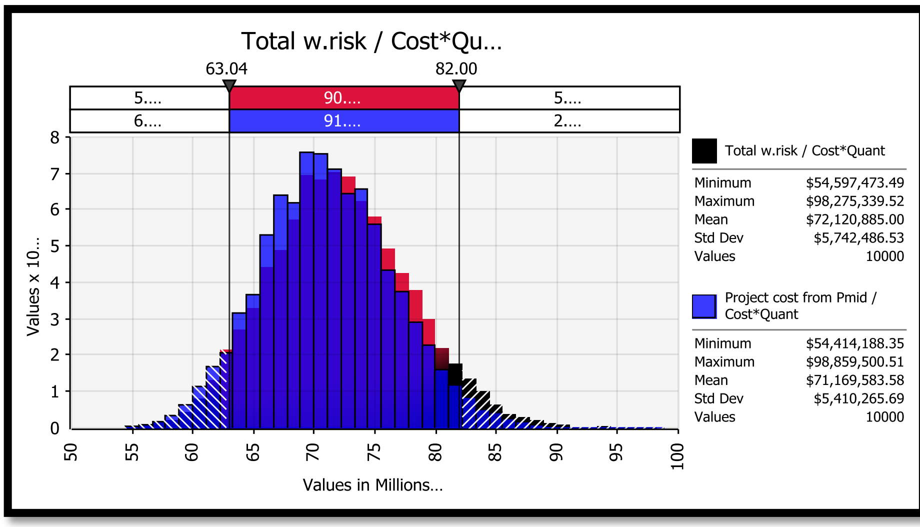


Figure 22: Impact to Project Cost inclusive of Risk Allowance Experiment 3

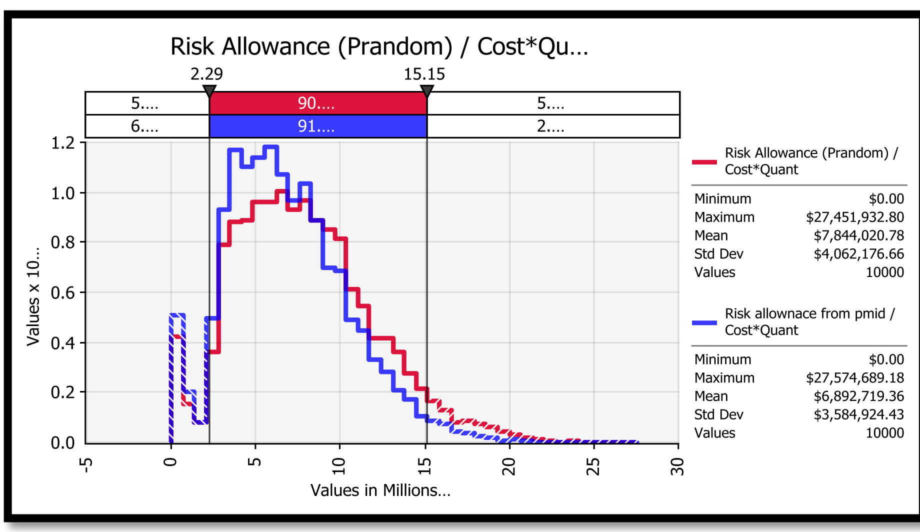


Figure 23: Impact to Risk Allowance Experiment 3

The results of the simulation shown above demonstrate that for the sample project, the project risk allowance, and the final cost estimate result in different distributions when the input values are varied as specified in Table 32. Different inputs for the range estimate were all compared against the two scenarios for risk input: Binomial (midpoint of range) and Binomial (Uniform (min, max of the range)). The range estimate input values included a central range, a skewed range and a random range which are linked to the impact of the risk (risks are an impact of a given portion of the estimate).

The results of the sensitivity of the decision variables (total cost and risk allowance) as evaluated using their mean and variance is shown in Table 33. The reference point we are measuring against is case 1.a. We notice that the effect of changing the probability of a risk has minor impact on the mean value of the project cost (only about 1%). The effect on the standard deviation is more significant at about 11%. The impact of the risk allowance is also significant for both mean and standard deviation at approximately 11%.

In summary, the mean of the total project cost absorbs the variability induced by the selection of the Binomial distribution parameters. The changes to the Binomial parameters, however, affect the standard deviation of the total cost and the risk allowance mean and standard deviation.

Table 33: Effects of Probability Selection (midpoint vs. uniform)

Experiment	Mean project cost	Standard Deviation Project cost	Mean Risk allowance	Standard Deviation Risk allowance cost	Effect on Mean Project Cost	Effect on Project Cost Standard Deviation	Effect on Mean Risk Allowance	Effect on Risk Allowance Standard Deviation
1.a	71.2	3.8	6.9	3.6	1.2%	11.6%	11.5%	12.2%
1.b	72.1	4.3	7.8	4.1				
2.a	81.8	3.6	6.9	3.6	1.2%	12.2%	11.5%	12.2%
2.b	82.8	4.1	7.8	4.1				
3.a	71.2	5.4	6.9	3.6	1.2%	5.3%	11.5%	12.2%
3.b	72.1	5.7	7.8	4.1				

The second sensitivity test was to see the effect that the shape of a distribution has on the results if the probability of risk event is kept the same. This experiment compared Experiment 1 results against experiment 2 and 3. Table 34 shows that changing the shape of the distribution to a skew to the right (experiment 2) changes the mean of the total cost by 13% (expected as the distribution simply is shifted to the right) and the variance reduced by 5.6% (expected since the skewed distribution has a very high kurtosis –the shape parameter of α=25 makes it tight distribution).

Changing the shape of the distribution but keeping it symmetric essentially has no impact on the mean (since the base case is also symmetric). The impact of the standard deviation is significant at 30%. These observations can be seen graphically in Figure 17.

Table 34 Effect of Input parameters for Impact (differing ranges)

Experiment	Mean project cost	Standard Deviation Project cost	Mean Risk allowance	Standard Deviation Risk allowance cost	Effect on Mean Project Cost	Effect on Project Cost Standard Deviation	Effect on Mean Risk Allowance	Effect on Risk Allowance Standard Deviation
1.a	71.2	3.8	6.9	3.6				
1.b	72.1	4.3	7.8	4.1				
2.a	81.8	3.6	6.9	3.6	13.0%[10]	-5.6%[11]	N/A	N/A
2.b	82.8	4.1	7.8	4.1				
3.a	71.2	5.4	6.9	3.6	0.0%[12]	29.6%[13]	N/A	N/A
3.b	72.1	5.7	7.8	4.1				

Although the experimentation was limited to the said scenarios, the results demonstrate that the decision making parameters of cost and risk allowance are sensitive to the input error we may introduce in the probability and/or shape of the distribution used even though we may be using the same verbal expression for the probability and the same end points for the ranges of unit price and quantity.

4.12 Concluding remarks

The experiments in this chapter show that regardless of whether we are using a single risk factor with a single unit value for impact, or a single work package estimate with unit values for unit cost and quantity, or a real project, the input parameters used effect the results. In the following chapters effort will be spent to improve the accuracy by which

[10] These values are showing the comparison between experiment 1 and experiment 2 for the mean.
[11] These values are showing the comparison between experiment 1 and experiment 2 for the standard deviation.
[12] These values are showing the comparison between experiment 1 and experiment 3for the mean.
[13] These values are showing the comparison between experiment 1 and experiment 3 for the standard deviation.

the verbal expression can be converted to probability values in such a way that they would reflect the expert's judgement, and secondly to enhance the way we solicit the shape parameters for the distribution.

5 Chapter 5 Understanding Verbal Expressions through Survey Data

5.1 Survey Goals:

This research aims to enhance the area of risk management in the construction industry by improving the current approach being used to quantify risks. The goal is to provide enhancements which are practical and applicable for a construction project today. The focus of this research is on quantitative risk analysis, a portion of the overarching process of risk management. In quantitative risk analysis, risk statements and qualitative data gained are transformed to quantifiable data that can be used to assess risks in terms of applicable project data such as cost and schedule. The process in which this is done currently is quite subjective. The goal of this study is to create a more systematic approach to deriving data. This survey aimed to do this by understanding certain verbal expressions in terms of their quantitative meaning. In other words, the survey aimed to link a verbal expression such as the term "likely" with a quantitative value such as "80% probability of occurrence". The survey will be provided in Appendix A, for reference and as a tool which can be replicated for future studies. It should be noted that the results gained are context specific and can only be generalized to construction industry personnel within Edmonton.

5.2 Sample Population Solicited:

This study was carried out within Edmonton, with the target population being construction engineering personnel. Individuals responding to the questionnaire were to be in some way affiliated with the construction industry, whether it is through academia or practice, in order to generalize back to the population where such analysis will be of use (risk analysis in the construction industry). The sample population used for this study included undergraduate students in Civil Engineering (completing 400 level courses) at the University of Alberta, graduate students in the Hole School of Construction Engineering at the University of Alberta and Construction Industry Personnel within Edmonton. Data collection was conducted either in a University of Alberta classroom, or emailed to construction personnel within Edmonton, Alberta. Participants were recruited either by email, or, if a student, approached during class time. All participation was voluntary. Some companies who have previously worked with the researcher were approached, however there was no hierarchical relationship and no reason that individuals would feel obliged to participate. All participants had the capacity to give free and informed consent. Consent was implied by overt action (i.e.: completion of the questionnaire). If a participant wished to withdraw, end or modify their participation, they could do so at any time before they had submitted their questionnaire. Their data would be safely destroyed and not used in any way in the study. For results on participant demographics, refer to section 5.6.4.

5.3 Research Procedures:

This study involved a questionnaire, and database. This was conducted by use of "Survey Monkey", an online survey tool that can be accessed via the web. Access to survey data is protected by login information. The full version of the online survey tool (which allows the researcher to create surveys and maintain/analyse responses) can only be accessed by researchers who have signed a confidentiality agreement provided to them by their supervisor (in this case the principal researcher's supervisor). The privacy and security of the information contained in this database is kept through the use of confidentiality agreements. The principle researcher maintained responsibility for the survey within the program.

5.3.1 Data collection:

Participants that were asked to complete the survey online were provided a standard link which gave the individual access to the survey questions through Survey Monkey, which was used to store their responses. Participants provided the survey in the classroom were given a hard copy of the survey to fill in, which was later recorded by the principle researcher. Identifying information was not stored in any case; only demographic information asked in the survey was kept. Questions were not of a sensitive nature (i.e.: they did not involve questions that could cause emotional or physical distress).The data collection and analysis was completed by the principle researcher and kept confidential. Data analysis was conducted using Microsoft Excel and Survey Monkey. Results were generated in terms of statistical values or visual representations (graphs/charts etc.). Data gained was handled and analysed only by the principle researcher, and only aggregated statistics are shared as part of the research. The data collection process started on March 25, 2013 and continued until April 04, 2013.

5.4 Study Summary:

In order to improve the input data for risk analysis, this study will provide a survey which asks participants to explain their understanding of commonly used verbal expressions for quantitative risk analysis. In particular, this survey will investigate the likelihood terms discussed previously in Table 8, and will also look to gain information on impact parameters. While likelihood terms have been previously studied, impact parameters for risk analysis have little documented information, therefore the goal for these parameters is to gain data that can be utilized and improved upon for future research. After study of the range estimating process in practise, it appears that two factors are required to understand the input: the confidence and the conservatives of the estimator. If an estimator is confident for example, this would mean that the impact distribution should be tight, if the estimator is unsure about the estimate, the distribution should be wider. In other words, confidence levels can be used to determine the amount of variance of a distribution. In order to understand the skew, the level of conservatives or riskiness perceived by the estimator can be investigated. If the estimator believes that the estimate is risky for example, the distribution should show the data as being more likely to be in the higher end of the distribution.

This survey is a modified version of that conducted in Reagan, Mosteller and Youtz (1989) where quantitative meanings of verbal probability expressions were investigated.

While this paper aimed to understand 18 probability expressions using the stems "probably, likely, possible and chance", this research focuses on using the results of words with the stem 'likely', as this is what is used to find the probability of a risk occurring in quantitative risk analysis. Findings from Reagan et. al. (1989) suggests the following ranges for those with the stem "likely": Very unlikely (2% to 15%); Unlikely (10% to 25%); Likely (65% to 85%); Very Likely (75% to 90%). Mainly missing from these percentages is a value between 25% and 65%. To account for this range, the term "somewhat likely" will be added to the expressions list. The term "very unlikely" will be replaced with "extremely unlikely" in order to attempt to differentiate between the results that Reagan et. al (1989) gained for the terms very unlikely and unlikely. The expectation is that given these five expressions, these percentages will differ slightly from those found in the Reagan et. al. (1989) paper. Following the logic in this paper, both word to number (verbal expression is presented and numerical value is gained) and number to word (numerical value is presented and verbal expression is gained) expressions will be investigated, however while the interest of Reagan et. al. (1989) lied in the ability to interchange numbers and words, this study will integrate this to avoid response error, an error which occurs usually if respondents do not correctly understand questions or concepts referred to in the survey, or if participants fill in the survey without being fully attentive, variables were correlated within the same survey (Camburn, Huff, Goldring, & May, 2010). In Section 1, parts A, B and D correlate with one another and in Section 2 parts A and B are correlated with one another. Essentially these questions ask for the same information in different ways. This provides evidence to enhance the validity of the survey. The method of word to number and number to word expressions will then be extended to the expressions used to define confidence in risk analysis to find out which corresponding ranges these refer to. The terms for confidence include: very confident, confident, somewhat confident, unsure and extremely unsure. The options for ranges are chosen based on informal interviews with industry professionals who have determined the following to be the most common values used for construction ±5%, ±10%, ±20%, ±25%, ±50%, ±75%, ±95%. An option is left at the end for any other percentage that an individual may find more appropriate. Lastly, this study is interested in understanding the quantitative meaning for conservativeness expressions including: conservative, somewhat conservative, close, somewhat risky and risky. Participants will be asked to provide their interpretation of these expressions based on a short scenario. This scenario provides an estimate of a construction project including the maximum, minimum and most likely values expected to occur. Participants will then be prompted with a conservativeness expression and asked which of the values provided would best depict their estimate given the expression.

The survey contained 46 questions, mainly multiple choice questions, and some option for short answer.

5.5 Methods:

The quantitative meanings of verbal expressions are of interest in this study. In order to measure this, participants were asked numerous questions which will be described within this section.

First demographic information was gained including age range and gender. Participants were asked whether they were a full time student (and if so, their program of study) as well as their current occupation (if applicable).

In section A (qualitative expressions in terms of single values) participants were asked to answer by providing a single value to describe their understanding of 5 predefined terms: very likely, likely, somewhat likely, unlikely and extremely unlikely. Participants were then provided the same terms, but asked to provide a range of values to describe each term. Following this, their preference for responding in terms of a single value, or a range of values was examined.

In section B (understanding likelihood): Participants were asked which of the predefined expressions (very likely, likely, somewhat likely, unlikely and extremely unlikely) they would use to describe a given probability (options included: 99%, 90%, 80%, 60%, 50%, 40%, 30%, 20%, 10%, 1%). Participants could check as many or as few that seem appropriate. Values for the probabilities provided were given based on most commonly used percentage expressions based on expert opinion. Regrettably, the value 70% was excluded from the survey in this portion. This did not compromise the results being investigated however, as this question was utilized as a validity check.

In section C (understanding confidence), participants were given the following scenario "An estimate has been created for the building of a 1 km tunnel. You are a tunneling expert who has been asked to assess the level of confidence you have regarding the line items of the current estimate. The risk analyst asks you how confident you are about line item A: Excavation of the working shaft for which the estimate you provided was $200,000.” Five predefined word responses (very confident, confident, somewhat confident, unsure and extremely unsure), were provided and participants were asked to respond by checking as many or as few of the distributions outlined to describe that probability. An example of a possible response is as follows: $200 000 ±20% (i.e. $160 000- $240 000). Next, a predefined probability value such as a “95%” is provided. Options included choosing as many or as few of the predefined expressions (very confident, confident, somewhat confident, unsure and extremely unsure) as appropriate.

In section D (understanding conservativeness), the following scenario was provided "An estimate has been created for the building of a 1 km tunnel. The cost of the tunnel can range between a minimum cost of $200 000 and a maximum cost of $300 000. The most likely cost of the tunnel is $250 000. You, an estimator, have been called upon to conduct your own estimate of the tunnel." Participants were then told that they had estimated the cost of the tunnel and that they believed their estimate was one of 5 conservativeness expressions (conservative, somewhat conservative, close, somewhat risky, and risky). Options for answers included values between $200 000 and $300 000, in intervals of $10 000.

5.6 Results

5.6.1 Survey Objectives

The goal of the survey was to understand what people believe likelihood expressions relate to in terms of quantitative data, in other words finding word to number correlations.

Of the 218 surveys sent out, 151 responses were completed. Of these responses, 142 were found to be valid. The next section describes the inclusion criteria required for survey validity.

5.6.2 Inclusion Criteria

Responses were not to be included in the study if one or more of the following occurred:

1. Instructions were not correctly followed (e.g.: respondent did not choose 2 or more responses when required).
2. Instructions were not correctly interpreted, or respondent did not understand terminology used.
3. Responses did not correlate with one another when expected (see Section 5.4 for validity check information).
4. A response was dependent on a previous response, and one of the responses was missing or incorrectly interpreted.

5.6.3 Assumptions:

A few assumptions were used in the analysis of the survey, firstly for Question 5 in the survey, respondents were asked to choose the most appropriate range of values that would best represent the probability of that outcome occurring. Many respondents answered this question by choosing two values. It is assumed that in this case, if the two values chosen were not continuous, the respondent chose a minimum and maximum value for the range. In this case, all boxes that were not checked within this range were checked on behalf of the respondent, by the principle researcher. For example: if a respondent checked the values 80 and 95, boxes 85 and 90 were checked on behalf of the respondent. Next, it was assumed that for Question 10, respondents did not understand the question if they believed that a conservative estimate was a higher value than a somewhat conservative, close, somewhat risky or risky estimate. In this case, answers for this question were removed.

5.6.4 Demographics:

Participants were asked to volunteer their time in order to further a study looking to enhance current techniques being used for quantitative risk analysis in construction. As noted above, a total of 218 questionnaires were distributed to industry professionals and civil/construction engineering students. One hundred fifty one responses were completed (a response rate of 69%). Data from 142 was used for the study based on assumptions used to determine validity. Of the 142 respondents, 24.3% were female and 75.7% were male, one respondent chose other, and another did not respond to this question. The majority of respondents were between the ages 18-24 (59.2%), followed by respondents between the ages of 25-34 (21.8%) and 35-44 (8.5%). The remaining 10% of participants were between the ages of 45-54 (5.6%) and 55-64 (4.9%). Thirty-nine respondents were industry employed while 102 were full time engineering students. Forty percent of the students were completing an after degree in Construction Management, while 60% were completing a Bachelor of Science in Civil Engineering. There were no statistically significant differences in results between the different demographics in this study.

5.6.5 Qualitative expressions for probability (word to number):

Qualitative expressions for probability were first investigated by asking respondents to answer in terms of single values and in terms of a range of values for the following terms: very likely, likely, somewhat likely, unlikely, extremely unlikely. Respondents could choose values between 0 and 100 in increments of 5. The question was also asked in terms of number to word responses in order to verify validity of responses. Number to word associations were presented for all terms and yielded agreeing results (i.e. there were no discrepancies between the two ways respondents answered the same question).

The results of the survey are summarized in

Table 35 for the two approaches (1) where the respondent was asked to provide one single answer for each of the subjective terms and (2) where the respondent was asked to provide a range of values for each of the subjective terms. They are also shown in the form of distributions (histograms but shown as 'line' plots for illustrative purposes only, as comparison through the overlay of the graphs using a line plot is more clearly observable) in Figure 24 for single values and in Figure 25 for range of values.

To illustrate what is given in the

Table 35, for the term "very likely" the mean response was 87.3%, with a mode of 90%, a minimum of 60% and maximum of 100%, when asked to choose a single value. When asked to choose a range of values the mean was 87.1%, with a mode of 90% and a range between 50% and 100%.

Table 35: Summary of the results for probability

	Single value response (percent)				Range of values response (percent)			
Term	**Mean**	**Mode**	**Minimum**	**Maximum**	**Mean**	**Mode**	**Minimum**	**Maximum**
very likely	87.3%	90%	60%	100%	87.3%	90%	60%	100%
Likely	69.3%	70%	50%	85%	70.6%	70%	40%	100%
Somewhat likely	52.8%	60%	20%	70%	52%	50%	0%	80%
Unlikely	26.5%	30%	5%	50%	28.9%	30%	0%	65%
Extremely unlikely	7.7%	5%	0%	50%	11.1%	5%	0%	55%

The histograms of the responses for each of the subjective terms are given in Appendix 2 for the terms of ‘very likely’, ‘likely’, ‘somewhat likely’, ‘unlikely’, and ‘extremely unlikely’ respectively.

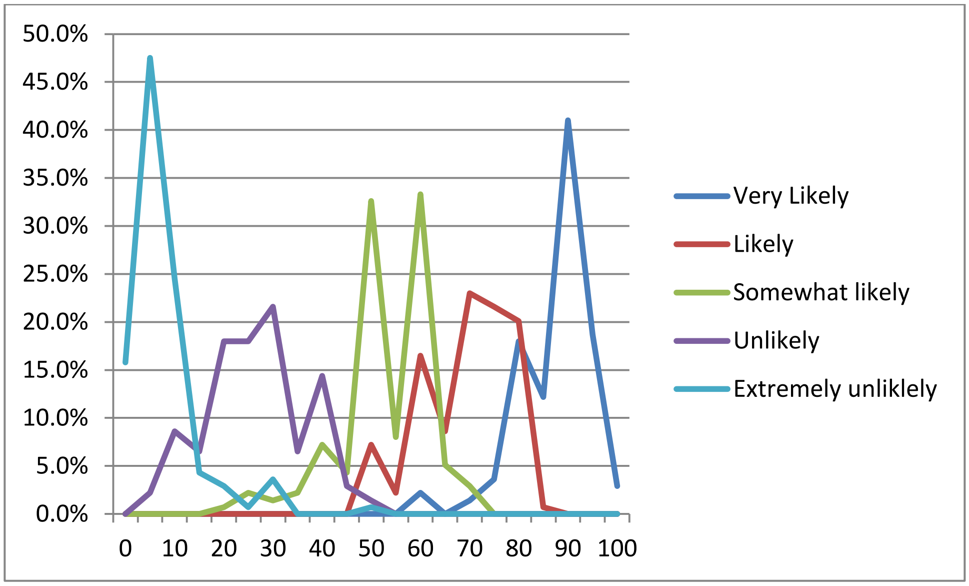


Figure 24: Probability Values in terms of Single Value Responses

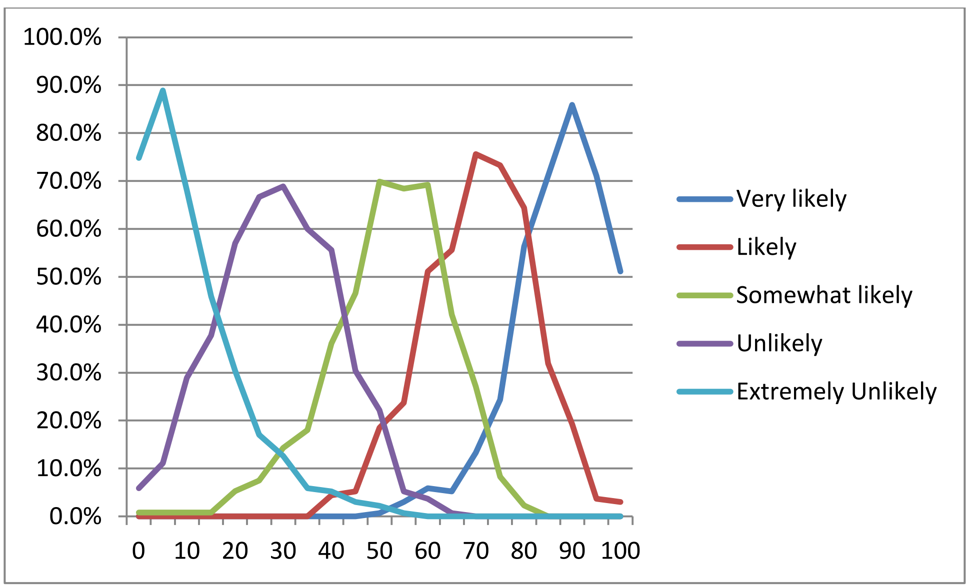


Figure 25: Probability expressions in terms of range responses

5.6.6 Preference: Answering in a range or a single number

When asked whether respondents preferred to answer in terms of a single value, a range of values, or had no preference, 66.7% of the respondents chose a range of values as shown in Figure 26.

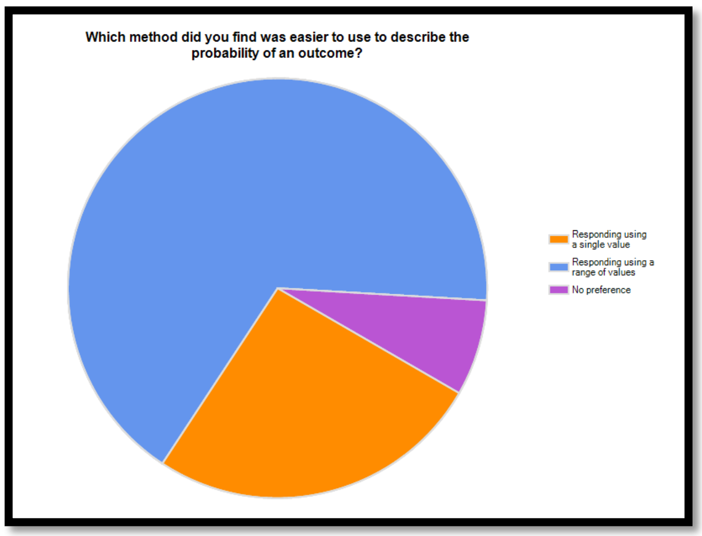


Figure 26: Single Value vs. Range of Values

5.6.7 Confidence

The investigation into what respondents mean when they use a qualitative term to reflect how confident they are about their estimate (e.g. in case of the impact of the risk factor). A summary of the results is given in Figure 27.

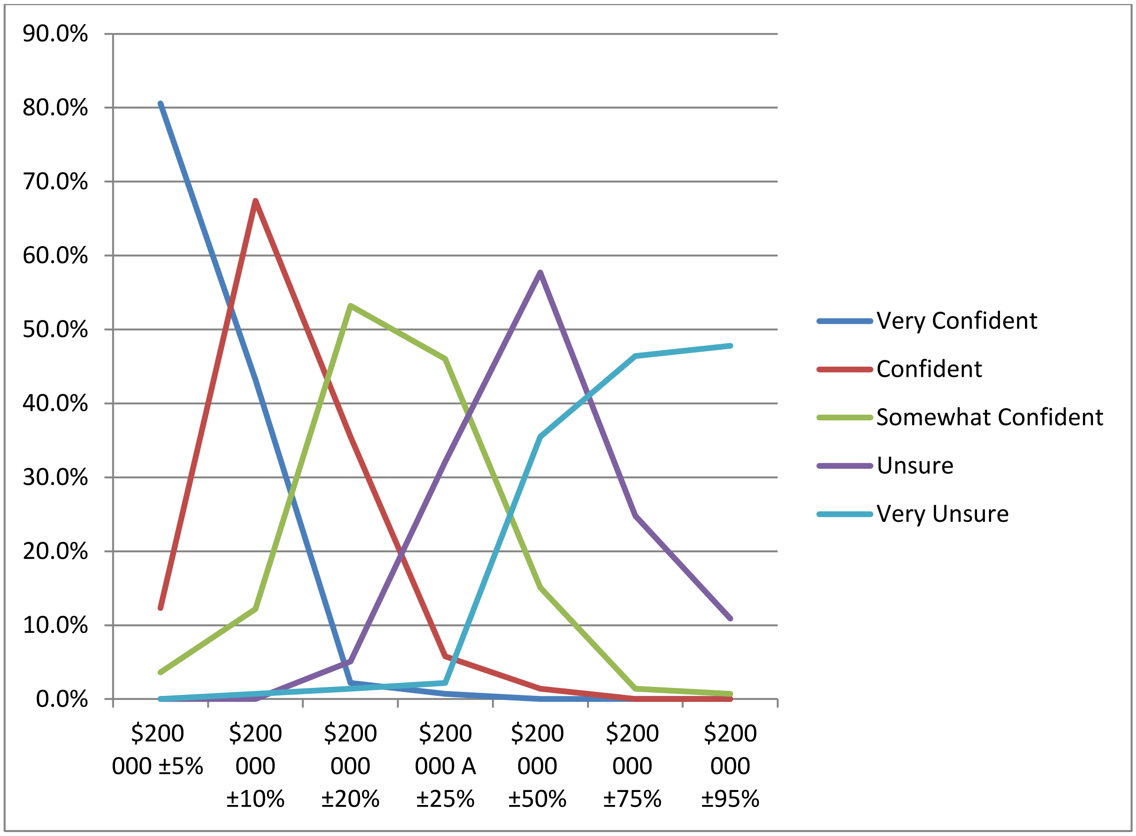


Figure 27: Summary of the results for confidence.

A sample of the individual histograms for 'very confident' is given in Figure 28. The histogram corresponds to the dark blue line in Figure 27.

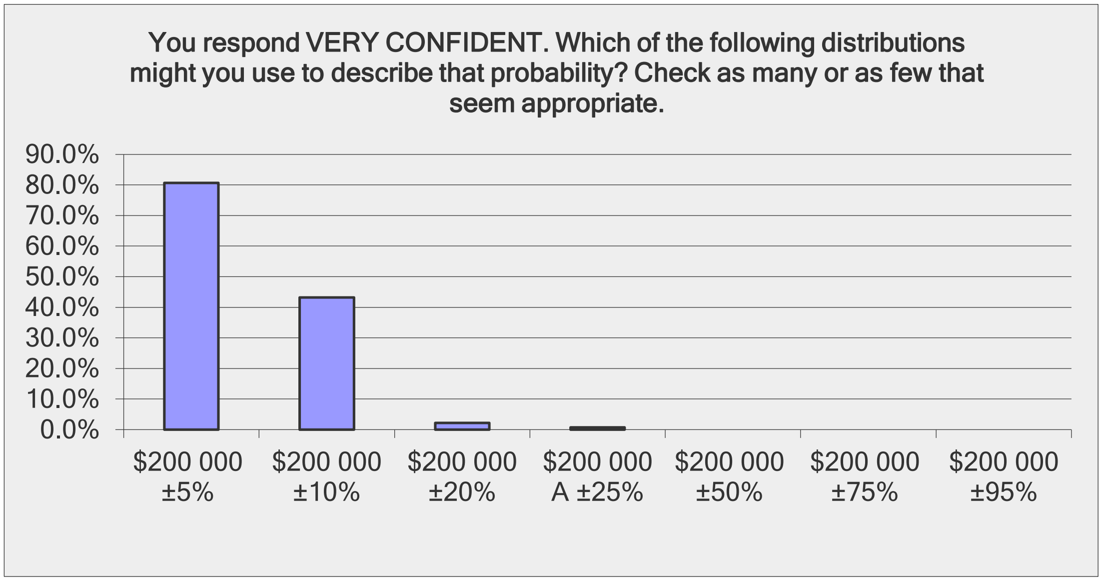


Figure 28: Histogram Results for "Very Confident"

The following observations can be made:

- The term "very confident" was associated most strongly with estimate ±5% (80.6% of respondents) followed by estimate ±10% (43.2% of respondents). Other responses included estimate ±20% (2.2% of respondents) and estimate ±25% (0.7% of respondents).
- The term "confident" was associated most strongly with estimate ±10% (67.4% of respondents), followed by estimate ±20% (35.5% of respondents), and estimate ±5% (12.3% of respondents). Other responses included estimate ±25% (5.8% of respondents), estimate ±50% (1.4% of respondents).
- The term "somewhat confident" was associated most strongly with estimate ±20% (53.2% of respondents), followed closely by estimate ±25 (46.0% of respondents). Other responses included estimate ±50% (15.1% of respondents), estimate ±10%

(12.2% of respondents), ±5% (3.6% of respondents), ±75% (1.4% of respondents), and ±95% (0.7% of respondents).

- The term “unsure” was associated most strongly with ±50% (57.7% of respondents), followed by ±25% (32.1% of respondents) and ±75% (24.8% of respondents). Other responses included ±95% (10.9% of respondents) and ±20% (5.1% of respondents).
- The term “very unsure” was associated most strongly with ±95% (47.8% of respondents) and ±75% (46.4% of respondents). This was followed by ±50% (35.5%), ±25% (2.2% of respondents), ±20% (1.4% of respondents) and ±10% (0.7% of respondents).

5.6.8 Conservativeness

To investigate what respondents mean when they use a term to reflect how conservative their estimate is we used a scenario to arrive at the word to number responses:

“An estimate has been created for the building of a 1 km tunnel. The cost of the tunnel can range between a minimum cost of $200 000 and a maximum cost of $300 000. The most likely cost of the tunnel is $250 000. You, an estimator, have been called upon to conduct your own estimate of the tunnel.”

The results of the survey are summarized in Figure 29.

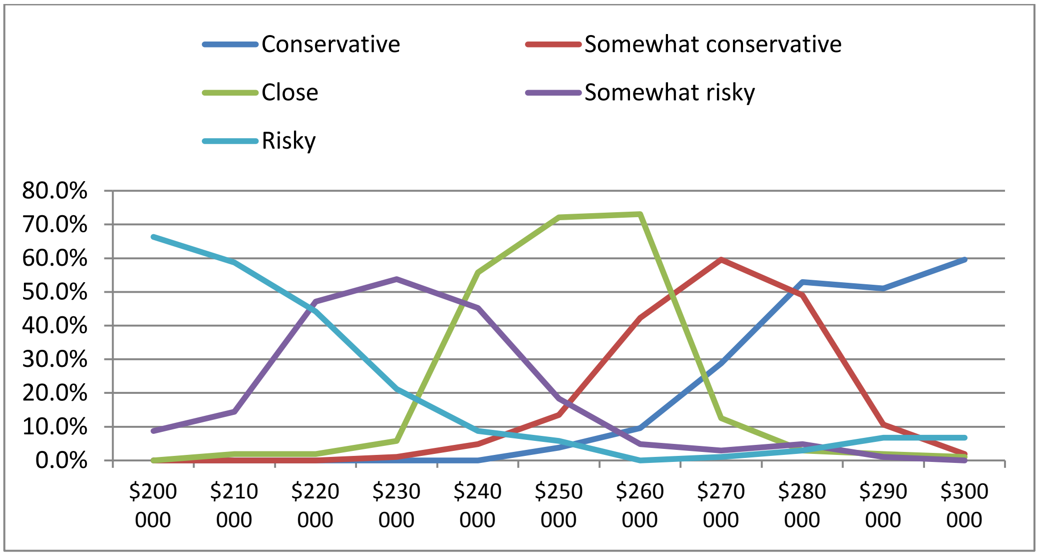


Figure 29: summary of results for Conservativeness

The following results were learned:

- A "conservative" estimate was seen as ranging from $250 000 to $300 000. The most agreement was shown for $300 000 (59.6% of respondents) followed closely by $280 000 (52.9% of respondents) and $290 000 (51% of respondents).
- A "somewhat conservative" estimate ranged from $230 000 to $300 000. The mode was $270 000 (59.6% of respondents) followed closely by 280 000 (49.0% of respondents)
- A "close" estimate ranged from $210 000 to $300 000. Responses showed that $260 000 (73.1% of respondents) and $250 000 (72.1% of respondents) were most strongly associated with this term, followed by $240 000 (55.8% of respondents).

- A “somewhat risky” estimate ranged from $200 000 to $290 000. The mode was $230 000 (53.8% of respondents). Following this was $220 000 (47.1% of respondents) and $240 000 (45.2% of respondents).
- A risky estimate ranged from $200 000 to $300 000. The mode was $200 000 (66.3%), followed by $210 000 (58.7% of respondents) and $220 000 (44.2% of respondents).

5.7 Analysis of the Survey Results for probability

5.7.1 Specifying probability of event occurrence using verbal expressions

Responses to questions 11-15, which ask participants to provide their opinion regarding the range of values that best represent the probability of occurrence for a given probability expression, were analysed using the process described in the three steps below. Once the analysis was complete, the results were summarized in two forms: (1) the basic statistics which can be readily used in Level 2 risk analysis, and (2) fitted statistical distribution which can be used in Level 3 analysis.

Step 1: Aggregate the data and determine the basic statistics:
Data for each expression was aggregated and the basic statistical values given in Table 36 were determined using @risk for Excel.

Table 36: Statistics evaluated from the sample

Statistic
Minimum value
Maximum value
Mean
Mode
Standard deviation
Coefficient of skewness
Kurtosis
First quartile (25th percentile)
Third quartile (75th percentile)

The statistics referred to in Table 36 can be used in risk analysis to determine the probability value to use in Level 2 risk analysis (CII, deterministic analysis) or in Level 3 (CII -Monte Carlo Simulation). While customary to use the mean value in Level 2, the results obtained through the survey (after they have been cleaned up and summarized) as reflected in Table 1 provide more options for the risk analyst (e.g. the mode, any of the quartiles etc.).

Step 2: Outlier data is removed:

A quick review of the raw data previously shown demonstrates that there were various occasions where the answers given were not reasonable. As customary in sampling, an attempt is made to remove any data that is considered to be outliers. In this work, the interquartile range (IQR) method was used to identify and disregard outliers (Vankeerberghen, Vandenbosch, Smeyers-Verbeke, & Massart, 1991). The IQR method is a simple yet effective approach for identifying outliers in a sample of data. The IQR is calculated by subtracting the 25th percentile from the 75th percentile. The boundaries are determined based on calculating the upper and lower bounds by defining the acceptable range to be a spread from the median as follows:

$$Upper\ bound = Median + 1.5(IQR)$$

$$Lower\ bound = Median - 1.5(IQR)$$

Any values that do not fall within the range determined by the lower to upper bound, are considered to be outliers and removed for the purpose of creating a distribution based on overall consensus.

Step 3 Fitting Distributions to the data:

Data for each expression were aggregated and the basic statistics were evaluated as previously given in Table 36. For Level 3 analysis, it would make sense to use a statistical distribution to model the probability given the range of values for each verbal expression we analyzed. In this work we used a simple moment matching approach (described below) coupled with trial and error to find the most suitable Beta distribution. The choice of a Beta distribution was simply based on finding a flexible family of distributions that can match the shapes of the histograms we observed in the previous chapter and while at the same time be available in simulation tools such as @Risk.

In order to fit a Beta distribution to the data, we follow the method proposed by AbouRizk et.al. (1991). If we assume that the limits of the distribution are defined by the Minimum and Maximum observations, then we will only need two parameters to solve for the shape parameters of the Beta Distribution (two equations to solve for the two unknowns). The formulation from AbouRizk et. al. (1991) is replicated below for convenience:

Given $X_{(1)}$ be the minimum value observed in the data and $X_{(n)}$ be the maximum value observed, $\overline{X}$ is the mean of collected data and S^2 its variance while L, U, μ, σ^2 are the minimum, maximum, Mean and Variance of the Beta distribution with shape parameters (a and b) to be fitted to this set, the procedure of moment matching would be as follows:

Set $L = X_{(1)}$ and $U = X_{(n)}$

To use moment matching we set the mean and variance of the sample equal to those of the theoretical mean and variance of the Beta distribution.

The mean and variance of the theoretical beta distribution are given by:

Mean:

$$\mu = L + (U - L)\frac{a}{a + b} \qquad \text{(A)}$$

Variance:

$$\sigma^2 = (U - L)^2 \frac{ab}{(a + b)^2(a + b + 1)} \qquad \text{(B)}$$

Note that L and U were previously set to the end points of the sample. We set $\mu = \overline{X}$ and $\sigma^2 = S^2$ in equations A and B and solve for the estimates of the shape parameters a and b to yield:

$$a = \frac{\overline{X}\, b}{1 - \overline{X}}$$

$$b = \frac{1 - \overline{X}}{S^2} \left[\overline{X}(1 - \overline{X}) - S^2\right]$$

The fitted distribution from the above formulation provides a starting point from which trial and error can be used to match the shape of the distribution to the shape of the data (matching coefficient of skewness and kurtosis) and/or the quartiles. We used @Risk for Excel for this process. Since we are using a trial and error process, the final results may not present the absolute best fit of a Beta Distribution, but would still provide a reasonable approximation given the errors inherited in the analysis. Once this process is complete, the Beta distribution parameters are noted.

5.8 Summary of the Results:

5.8.1 Results for Probability

The results from this section can be used to replace the verbal expression for likelihood in terms of a distribution for risk analysis. If using Level 2 risk analysis, information shown as "summary of the statistics" for each expression can be utilized. Following the logic in Chapter 3, a binomial distribution can be used, along with the mean value found through the survey. If a level 3 risk analysis is taking place, fitted distributions are provided below, along with their information, which can be used for a Monte Carlo Simulation.

Very Likely

A Beta distribution fitted to the statistics given in Table 2 has the parameters summarized in Table 37.

Table 37: Results of the distribution fitted to "very likely"

Distribution	Beta (3.5,2,70,100)
Shape Parameter 1	3.5
Shape Parameter 2	2.0
Minimum	70.0%
Maximum	100.0%

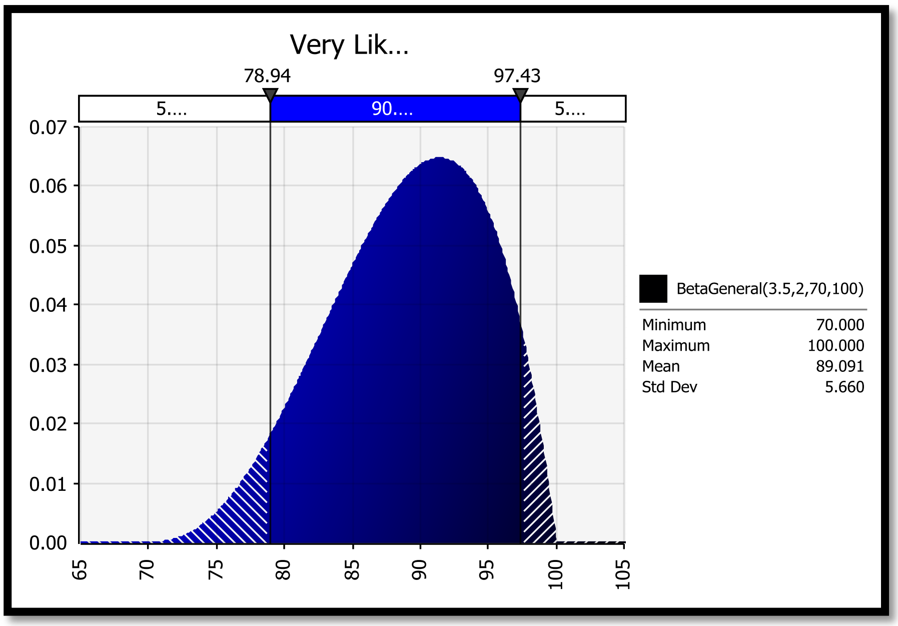


Figure 30: Fitted Beta distribution for "Very Likely":

A comparison between the fitted distribution's statistics and those of the collected data from the survey are shown in Table 38 below for comparison.

Table 38: Summary of the statistics for "very likely" and those of the fitted distribution

Statistics	Statistic from collected Data	Statistics from fitted Beta distribution
Mean	88.2%	88.9%
Standard Deviation	8.0%	5.7%
First quartile (25%)	80.0%	84.9%
Third quartile (75%)	95.0%	93.4%
Minimum	70.0%	70.0%
Maximum	100.0%	100.0%

Likely

For the verbal expression "Likely" the following Beta distribution in Table 39 reflects the best fit from the input modeling experiment.

Table 39: Results of the distribution fitted to "likely"

Distribution	Beta (2.6,2.4,50,90)
Shape Parameter 1	2.6%
Shape Parameter 2	2.4%
Minimum	50.0%
Maximum	90.0%

A comparison between the fitted distribution's statistics and those of the collected data from the survey are shown in Table 40 for comparison while the fitted distribution is shown in Figure 31.

Table 40: Summary of the statistics for "likely" and those of the fitted distribution

Statistics	Statistic from collected Data	Statistics from fitted Beta distribution
Mean	70.9%	70.8%
Standard Deviation	10.1%	8.2%
First quartile (25%)	65.0%	64.7%
Third quartile (75%)	80.0%	77.0%
Minimum	50.0%	50.0%
Maximum	90.0%	90.0%

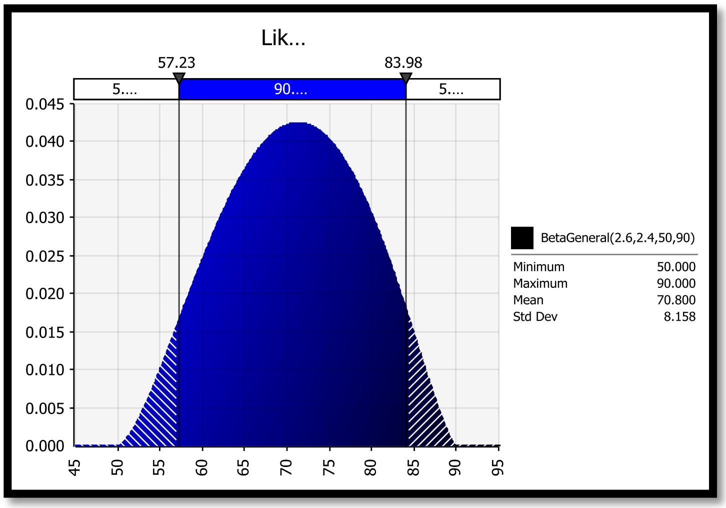


Figure 31: Fitted Beta distribution for "Likely"

Somewhat Likely:

For the verbal expression "Somewhat Likely" the following Beta distribution in Table 41 reflects the best fit from the input modeling experiment.

Table 41: Results of the distribution fitted to "somewhat likely"

Distribution	Beta (2.7,2.8,35,75)
Shape Parameter 1	2.7%
Shape Parameter 2	2.8%
Minimum	35.0%
Maximum	75.0%

A comparison between the fitted distribution's statistics and those of the collected data from the survey are shown in Table 42 below for comparison while the fitted distribution is shown in Figure 32.

Table 42: Summary of the statistics for "somewhat likely" and those of the fitted distribution

Statistics	Statistic from collected Data	Statistics from fitted Beta distribution
Mean	54.0%	54.6%
Standard Deviation	9.8%	7.8%
First quartile (25%)	45.0%	48.7%
Third quartile (75%)	60.0%	60.5%
Minimum	35.0%	35.0%
Maximum	75.0%	75.0%

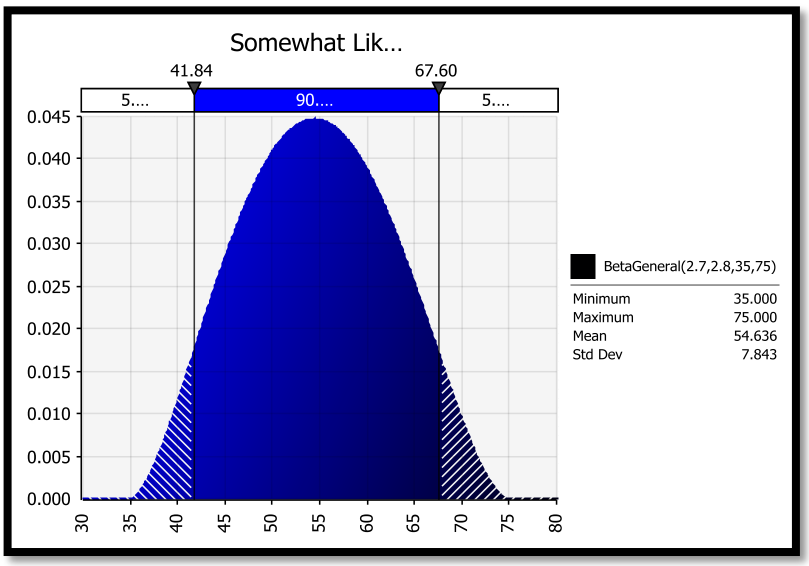


Figure 32: Fitted Beta distribution for "Somewhat Likely"

Unlikely:

For the verbal expression "Unlikely" the following Beta distribution in Table 43 reflects the best fit from the input modeling experiment.

Table 43: Results of the distribution fitted to "unlikely"

Distribution	Beta (4,3.99,0,60)
Shape Parameter 1	4.0%
Shape Parameter 2	3.99%
Minimum	0%
Maximum	60%

A comparison between the fitted distribution's statistics and those of the collected data from the survey are shown in Table 44 below for comparison while the fitted distribution is shown in Figure 33.

Table 44: Summary of the statistics for "unlikely" and those of the fitted distribution

Statistics	Statistic from collected Data	Statistics from fitted Beta distribution
Mean	28.9%	30.0%
Standard Deviation	12.3%	10.0%
First quartile (25%)	20.0%	22.8%
Third quartile (75%)	40.0%	37.3%
Minimum	0.0%	0.0%
Maximum	60.0%	60.0%

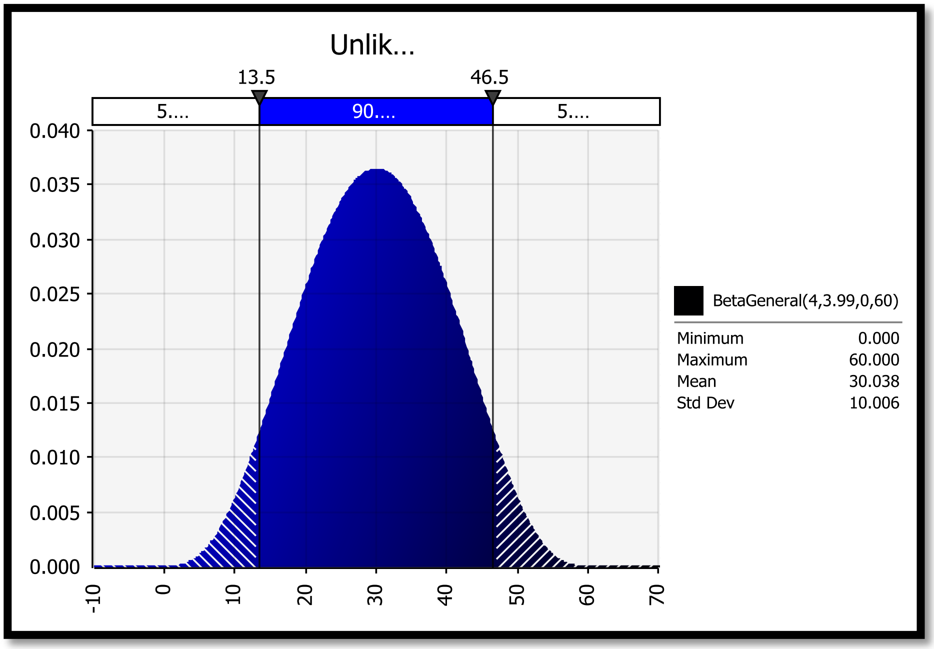


Figure 33: Fitted Beta distribution for "Unlikely"

Extremely Unlikely

For the verbal expression "Extremely Unlikely" the following Beta distribution in Table 45 reflects the best fit from the input modeling experiment.

Table 45: Results of the distribution fitted to "extremely unlikely"

Distribution	Beta (2,3.6,0,25)
Shape Parameter 1	2.0%
Shape Parameter 2	3.6%
Minimum	0.0%
Maximum	25.0%

A comparison between the fitted distribution's statistics and those of the collected data from the survey are shown in Table 46 below for comparison while the fitted distribution is shown in Figure 34.

Table 46: Summary of the statistics for "extremely unlikely" and those of the fitted distribution

Statistics	Statistic from collected Data	Statistics from fitted Beta distribution
Mean	8.8%	8.9%
Standard Deviation	7.3%	4.7%
First quartile (25%)	5%	5.3%
Third quartile (75%)	15%	12.2%
Minimum	0.0%	0.0%
Maximum	25.0%	25.0%

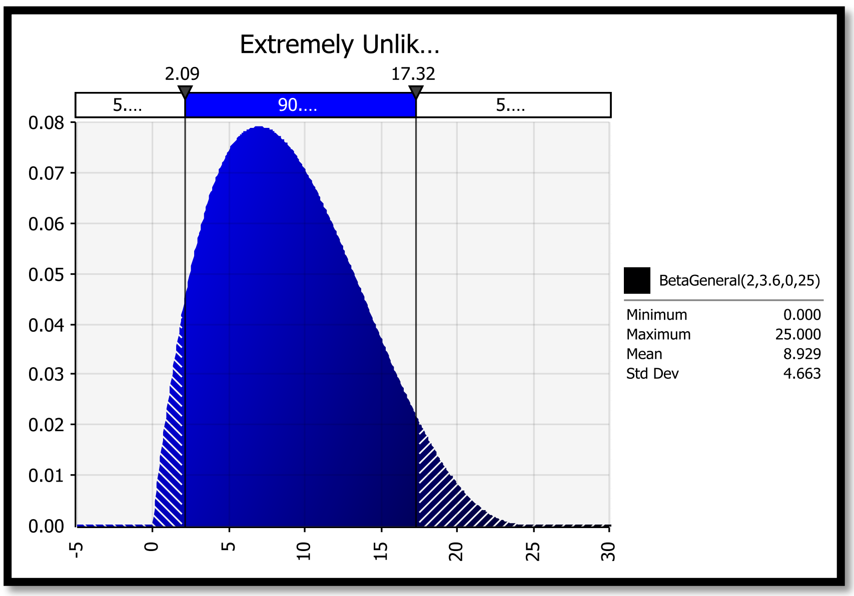


Figure 34: Fitted Beta distribution for "Extremely Unlikely"

5.9 New tables for probability:

Tables to use for Probability (Deterministic):
Table 47 shows values that can be utilized for a risk analysis using deterministic values.

Table 47: Probability Values (Deterministic)

Verbal Expression	Probability (Deterministic)
Very Likely	88.2%
Likely	70.9 %
Somewhat Likely	54.0 %
Unlikely	28.9 %
Extremely Unlikely	8.8 %

Tables to use for Probability (Probabilistic using a binomial, beta embedded distribution):

Table 48 can be utilized for a probabilistic risk analysis based on the survey data. This distribution utilizes the binomial distribution, where n=1, and the probability of the risk occurring is a distribution of values based on the meanings of different verbal expressions from the survey results. Different distributions can be derived, if preferred, using the methods outlined in Section 5.7.1 and the data shown in section 5.8 for likelihood results.

Table 48: Beta distributions to use for verbal expressions

Verbal Expression	Binomial (1, Beta Distribution Beta(α1, α2, min, max))
Very Likely	Binomial (1, Beta (3.5,2,0.7,1))
Likely	Binomial (1, Beta (2.6,2.4,0.50,0.90))
Somewhat Likely	Binomial (1, Beta (2.7,2.8,0.35,0.75))
Unlikely	Binomial (1, Beta (4,3.99,0,0.60))
Extremely Unlikely	Binomial (1, Beta (2,3.6,0,0.25))

5.10 Results for Confidence and Conservativeness

Measures of confidence and conservativeness have been investigated to provide a basis for understanding distribution information that can be used to understand the impact of a risk occurring. The following information encompasses the information learned.

5.10.1 Analysis of the measure of confidence

In order to understand the results consider the response to the question: "You respond VERY CONFIDENT. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate", which are tabulated in Table 49 below and provided as a histogram in Figure 35.

Table 49: Response to "very confident"

You respond VERY CONFIDENT. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.	
Answer Options	Response Count
$200 000 ±5%	112
$200 000 ±10%	60
$200 000 ±20%	3
$200 000 A ±25%	1
$200 000 ±50%	0
$200 000 ±75%	0
$200 000 ±95%	0
Other (please specify)	

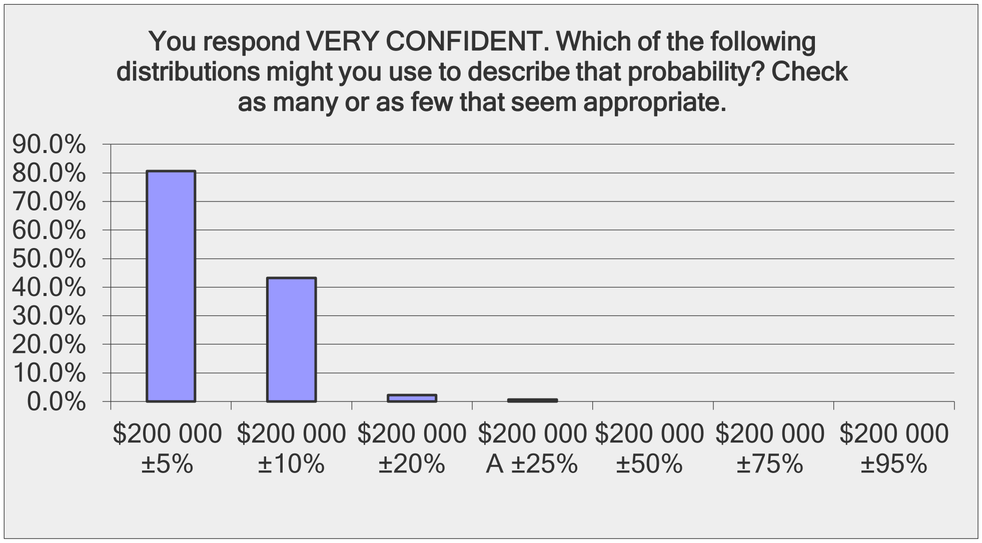


Figure 35 Histogram of responses to "very confident"

The majority of respondents indicated that their preference is ±5% (112) with some preferring ±10% (60).

Given that the data we are analyzing is considered to be nominal (discrete categories ranges increments by ±5%), the mode is considered to be the most appropriate representative of the sample (Driscoll, Lecky, and Crosby, 2000). Therefore, the response to the question 'very confident' as summarized in Table 49 and Figure 35 represented by the mode value which corresponds to the confidence ±5%. remaining responses for 'confident' (Figure 36), somewhat confident (

Figure 37), unsure (Figure 38), and extremely unsure (Figure 39) are analyzed.

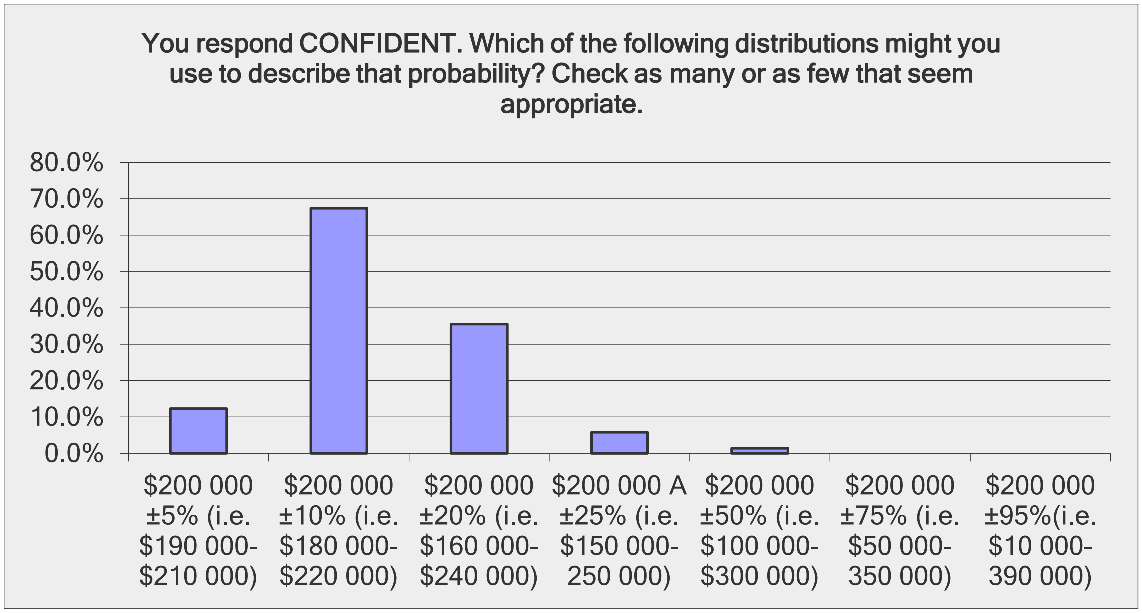


Figure 36 Histogram of responses to "confident"

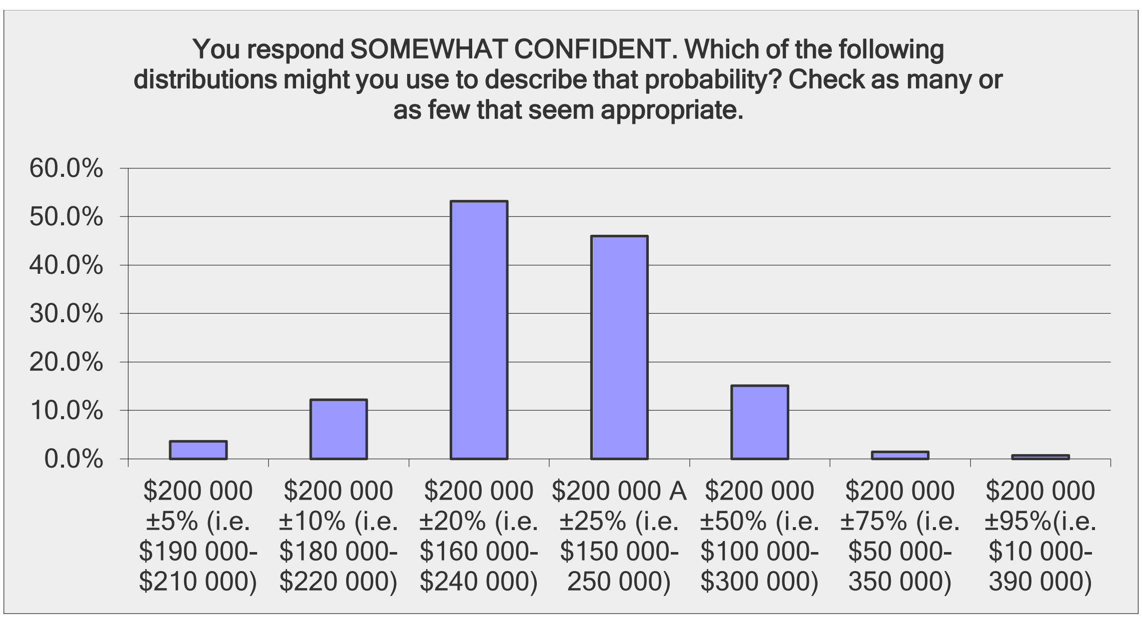


Figure 37: Histogram of responses to "somewhat confident"

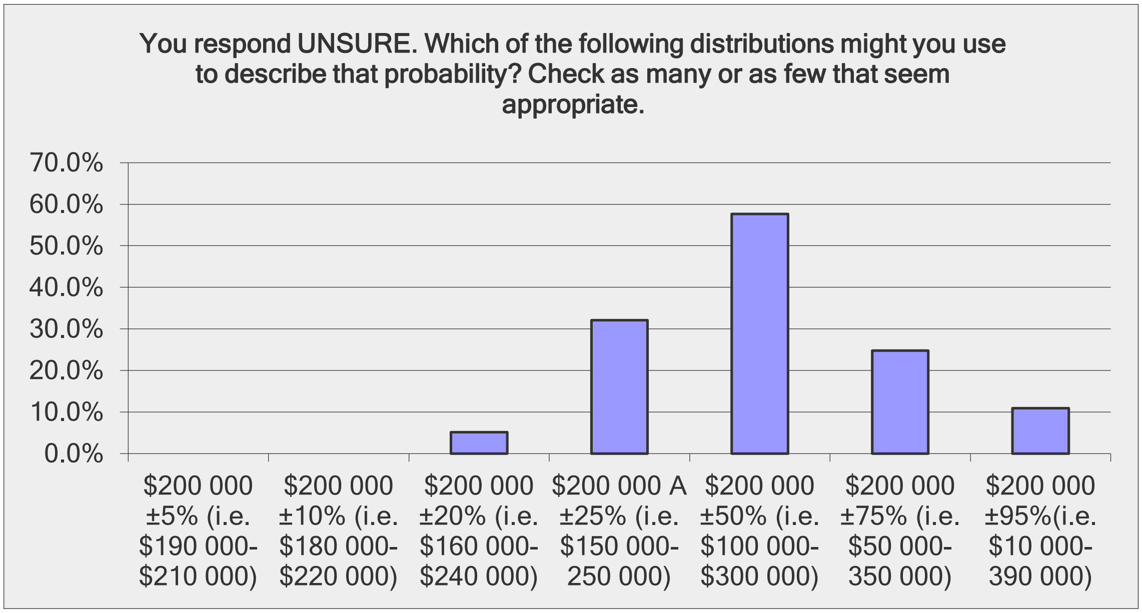


Figure 38: Histogram of responses to "unsure"

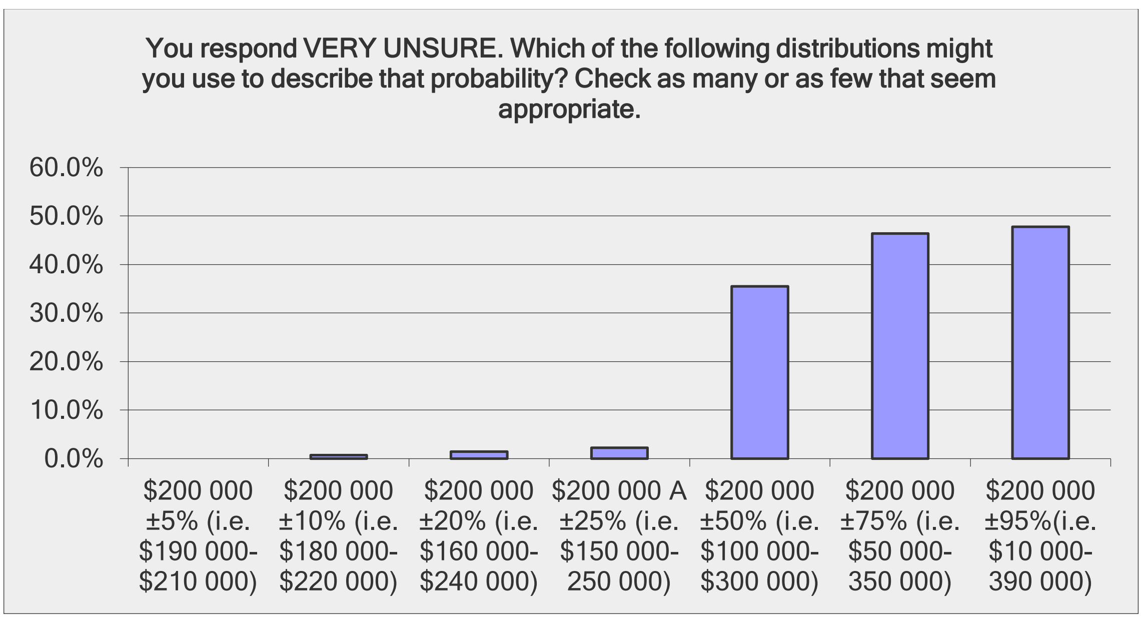


Figure 39: Histogram of responses to "extremely unsure"

A summary of our recommended numerical equivalence for verbal expression of confidence based on the histograms shown in Figure 35 to Figure 39 is given in Table 50.

These measures can be used to understand the tightness of the distribution for impact. These can be understood as a guideline that can be utilized during a Monte Carlo simulation (e.g. a response of very confident means X% of my data should be between ±5% of the estimate provided), and should serve as starting point for further analysis.

Table 50: Verbal expressions and quantitative data for Confidence

Verbal Expression for Confidence	Numerical equivalent
Very confident	±5%
Confident	±10%
Somewhat Confident	±20%
Unsure	±50%
Extremely Unsure	±85%

Analysis of the measure of conservativeness

The results of the survey for conservativeness were summarized in terms of their distribution in Figure 29, and are individually shown in terms of a histogram of results in Figure 40 (somewhat conservative), Figure 41 (close), Figure 42 (conservative), Figure 43 (somewhat risky) and Figure 44 (risky). In order to understand the results consider the response to the question: "You have estimated the cost of the tunnel (which is most likely to be $250,000) and believe that your estimate is SOMEWHAT CONSERVATIVE. Which of the following dollar values best depicts your estimate? Check as many or as few that seem appropriate", which are tabulated in

Table 51 below and provided as a histogram in Figure 40.

Table 51: Results for Somewhat Conservative

You have estimated the cost of the tunnel and believe that your estimate is SOMEWHAT CONSERVATIVE. Which of the following dollar values best depicts your estimate? Check as many or as few that seem appropriate.		
Answer Options	**Response Percent**	**Response Count**
$200 000	0.0%	0

$210 000	0.0%	0	
$220 000	0.0%	0	
$230 000	1.0%	1	
$240 000	4.8%	5	
$250 000	13.5%	14	
$260 000	42.3%	44	
$270 000	59.6%	62	
$280 000	49.0%	51	
$290 000	10.6%	11	
$300 000	1.9%	2	
	answered question		104
	skipped question		38

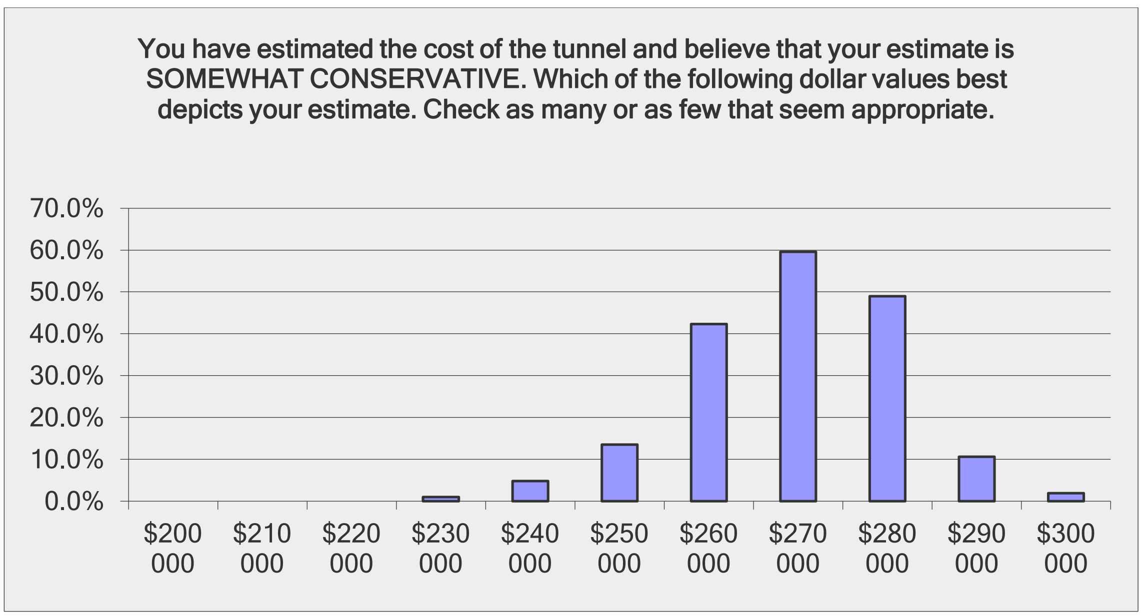


Figure 40: Histogram of Results for Somewhat Conservative

The majority of respondents indicated that their preference is '270,000' (62 responses). Again, given that the data we are analyzing is considered to be nominal (specific terms in this case), the mode is considered to be the most appropriate representative of the sample (Driscoll, Lecky, and Crosby, 2000). Therefore, the response to the question 'somewhat conservative' is best represented by the mode value which corresponds to the value

270,000. In order to generalize the results so they can be used generically, we can consider the percentile corresponding to the answers rather than the answer itself. In Table 52 we determined the percentile corresponding to each of the possible answers. We can see that the 270,000 corresponds to the 70th percentile. When one answers "somewhat conservative," therefore, we can assume that their estimate is close to the 70th percentile. Furthermore one can say that the possible range of values is 60% to 80%.

Table 52: Results for Somewhat Conservative in terms of Percentiles

You have estimated the cost of the tunnel and believe that your estimate is SOMEWHAT CONSERVATIVE. Which of the following dollar values best depicts your estimate? Check as many or as few that seem appropriate.			
Answer Options	**Response Percent**	**Response Count**	**%**
$200 000	0.0%	0	0
$210 000	0.0%	0	10
$220 000	0.0%	0	20
$230 000	1.0%	1	30
$240 000	4.8%	5	40
$250 000	13.5%	14	50
$260 000	42.3%	44	60
$270 000	59.6%	62	70
$280 000	49.0%	51	80
$290 000	10.6%	11	90
$300 000	1.9%	2	100
	answered question	**104**	
	skipped question	**38**	

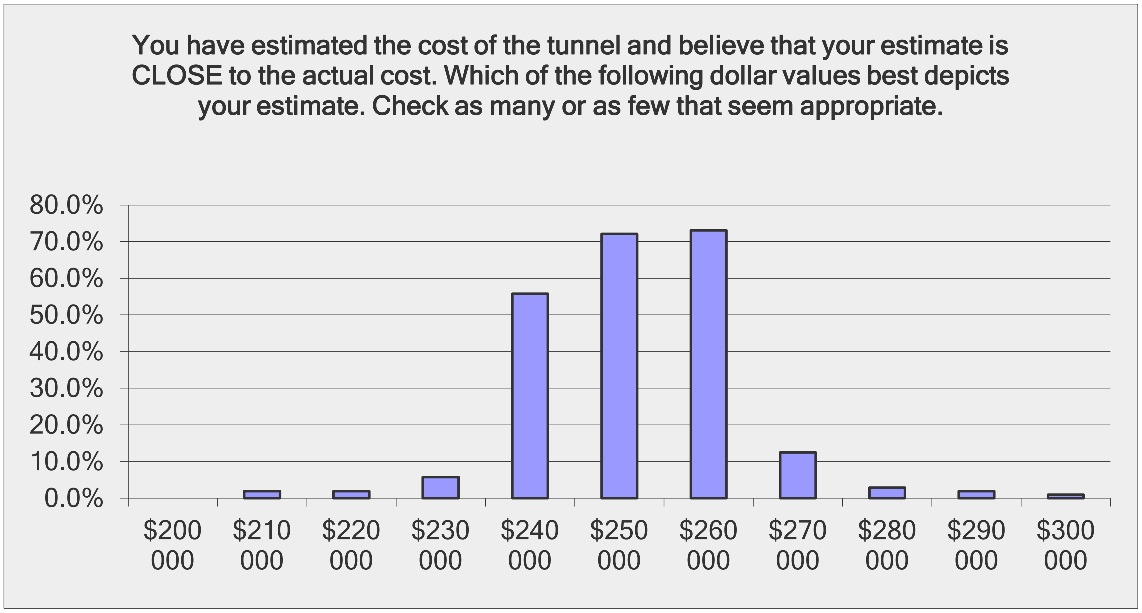


Figure 41: Histogram of Results for Close

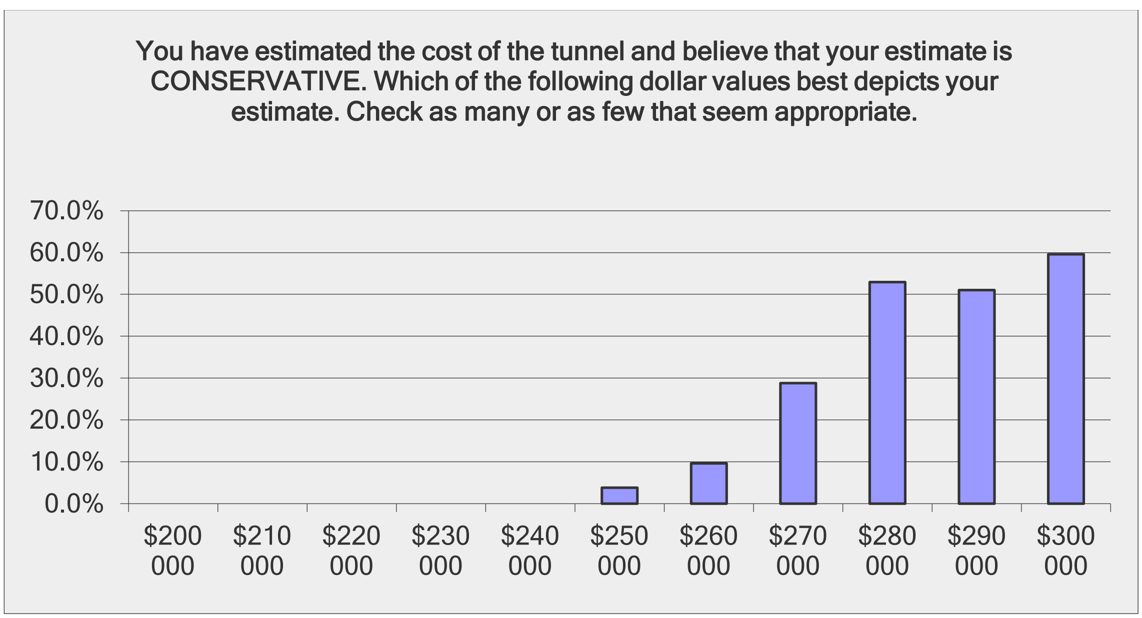


Figure 42: Histogram of Results for Conservative

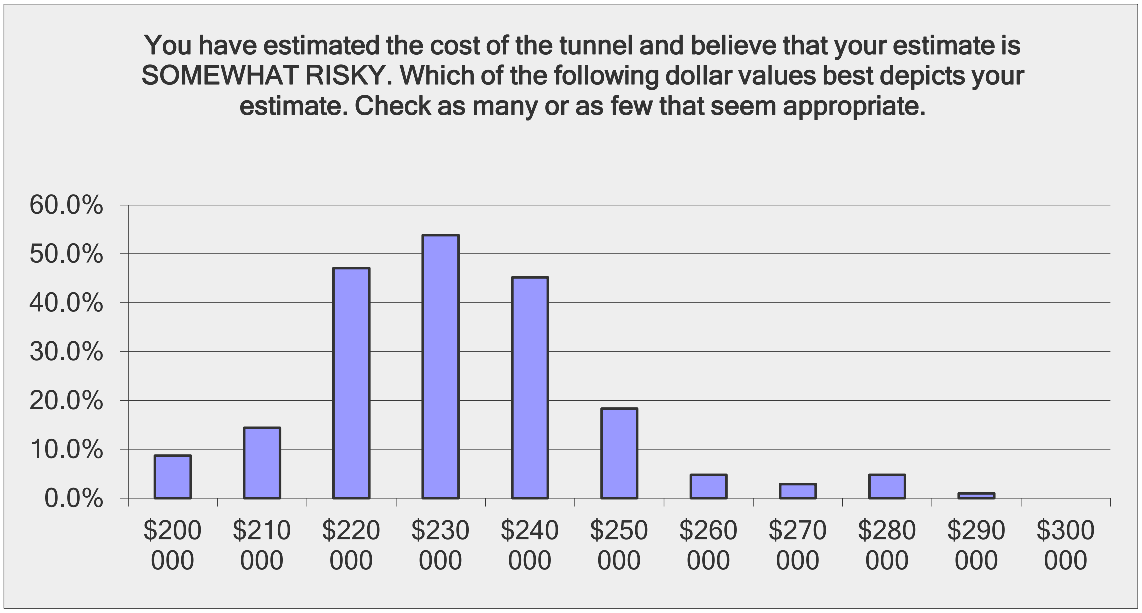


Figure 43: Histogram of Results for Somewhat Risky

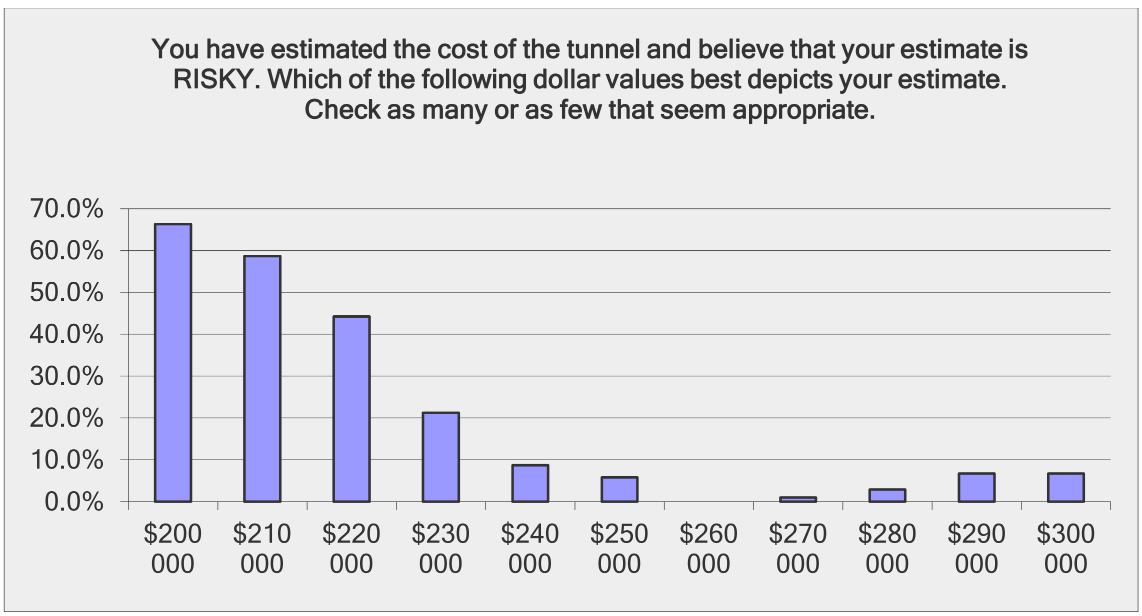


Figure 44: Histogram of Results for Risky

The results shown in Table 53 can be used to understand the skew of the distribution for impact. These can be understood as a guideline that can be utilized during a Monte Carlo simulation (e.g. a response of conservative means X% of my data should be leaning closer to the 80th-100th percentile of the range determined based on confidence), and should serve as starting point for further analysis.

Table 53: Ranges for the conservativeness measure

Verbal Expression for conservativeness	Numerical equivalent
Conservative	80th percentile-100th percentile
Somewhat Conservative	60th percentile-80th percentile
Close	40th percentile- 60th percentile
Somewhat Risky	20th percentile-50th percentile
Risky	0- 20th percentile

6 Chapter 6: Conclusions

6.1 Summary of the work

This thesis aimed to improve the quantitative risk analysis process currently used in the construction industry by understanding the significance of input parameters, improving the method of transcribing verbal expressions used during a risk analysis workshop into optimal quantification data, and understanding on how probability distributions can be developed during Monte Carlo simulation. The start of this thesis established the objectives, the approach to be followed and an overview of the state of the art. It was determined that there is a gap in the current literature relating to transcribing qualitative expressions into quantitative probabilities in the absence of historical data as well as in the quantification of impact of risk occurrence.

In order to understand the objectives, an example of the process of risk analysis was examined through a case study, allowing the reader to understand the practical need for verbal expressions and their role in the full risk management process.

In order to establish the need for accurate input parameters, Monte Carlo simulation experiments were conducted to understand the sensitivity of the project cost distribution to errors in the input of probability of risk events occurrence and the impact of such occurrence. The results showed that regardless of whether we are using a single risk factors with a single unit value for impact, or a single work package estimate with unit values for unit cost and quantity, or a real project, what we use to convert the verbal expression is important and likewise what we use to describe the shape of the distribution is important.

To optimize the input used for shape parameters, a survey was conducted soliciting information on likelihood verbal expressions along with verbal expressions for confidence and conservativeness. While likelihood expressions have been previously investigated for probability in risk analysis, utilizing confidence and conservativeness terms for impact is newly proposed and meant to be used during CII Level 3 analysis to more accurately describe distributions based on expert input. From this survey, a table was created outlining the statistical parameters of the likelihood expressions investigated in terms of probabilities. These probabilities can be used in both CII Level 2 and CII Level 3 analysis as raw data and statistical information from the survey was provided, along with fitted distributions based on these results. The tables of probability we recommend using for the various levels of risk analysis based on the survey we conducted are replicated in here in Tables 54 and 55, from Chapter 5 for convenience of reference.

Table 54: Summary of the results for probability for deterministic analysis (the mode based on Single value response is recommended for Level 2 analysis)

	Single value response (percent)				Range of values response (percent)			
Term	**Mean**	**Mode**	**Minimum**	**Maximum**	**Mean**	**Mode**	**Minimum**	**Maximum**
Very likely	87.3%	90%	60%	100%	87.3%	90%	60%	100%
Likely	69.3%	70%	50%	85%	70.6%	70%	40%	100%
Somewhat likely	52.8%	60%	20%	70%	52%	50%	0%	80%
Unlikely	26.5%	30%	5%	50%	28.9%	30%	0%	65%
Extremely unlikely	7.7%	5%	0%	50%	11.1%	5%	0%	55%

Table 55: Summary of the results for probability for Monte Carlo based QRA

Verbal Expression	Binomial (1, Beta Distribution Beta(α1, α2, min, max))
Very Likely	Binomial (1, Beta (3.5,2,0.70,1))
Likely	Binomial (1, Beta (2.6,2.4,0.50,0.90))
Somewhat Likely	Binomial (1, Beta (2.7,2.8,0.35,0.75))
Unlikely	Binomial (1, Beta (4,3.99,0,0.60))
Extremely Unlikely	Binomial (1, Beta (2,3.6,0,0.25))

For confidence, results were gained in terms of a range which can be utilized to describe the variance of an impact distribution. Conservativeness information was gained in terms of percentiles which can be used to determine the skew of an impact distribution. The confidence and conservativeness can be used together as guideline for creating probability distributions to model the impact of risk occurrence. Our recommendations are summarized in Tables 56 and 57 (replicated from Chapter 5 for convenience of reference).

Table 56: Verbal expressions and quantitative data for Confidence

Verbal Expression for Confidence	Numerical equivalent
Very confident	±5%
Confident	±10%
Somewhat Confident	±20%
Unsure	±50%
Extremely Unsure	±85%

Table 57: Verbal expressions and quantitative data for Conservativeness

Verbal Expression for conservativeness	Numerical equivalent
Conservative	80^{th} percentile-100^{th} percentile
Somewhat Conservative	60^{th} percentile-80^{th} percentile
Close	40^{th} percentile- 60^{th} percentile
Somewhat Risky	20^{th} percentile-50^{th} percentile
Risky	0- 20^{th} percentile

6.2 Overall Conclusions

In summary, prior to this work, deterministic as well as quantitative risk analysis greatly depended on estimates of probability of risk occurrence that was not scientifically based. Most literature assumed that experts are capable of providing an estimate of the probability of a risk event occurring on a continuous scale of 0.0 to 1.0 or through verbal expressions of likelihood values that are then converted to numeric equivalents using approximate means. The results of the thesis provided a practical approach for soliciting risk event probability from experts using five verbal expressions and the statistical underpinning required to convert these expressions into numeric values for QRA. Furthermore, the thesis provided additional guidelines to assist in determining the shape of the probability distributions to model risk impact based on soliciting from the expert input related to how conservative and how confident they believe their estimates are. The results were presented in tables that can be readily used by practitioners. This new approach provide for a next step in the evolution of quantitative risk analysis. Previously, tables being utilized for risk analysis were based on information and analysis not specific to the needs of the construction industry, by completing this survey we have made the approach more reliable by providing a survey which is based on information that is contextually applicable. These tables can continue to be expanded upon and researched to further verify the reliability of this method.

6.3 Contributions

The following was achieved through this research:

- Showing that the decision parameters of risk allowance and project costs are sensitive to the errors introduced during the quantification of subjective probability of risk occurrence, or the shape of the distribution used for quantifying the impact of those factors.

  In most QRA work the prevalent approach to providing probability estimates are (1) subjective and either utilize tables of verbal expression of likelihood (which are converted to probability values) or (2) a straight approximation of probability on a scale of 0 to 100%. In the first case we have shown that the gradation that is currently used is crude and leads to wide fluctuation in the decision parameters. In the second approach we know from literature that an expert cannot properly estimate probability on such a wide scale and therefore this leads to potential errors that propagate into the decision parameters. This was shown in Chapter 4 through Monte Carlo simulation studies.

- The thesis provided a method for gaining reliable measures for probability of a risk occurring. This was achieved through building off a previous study (Reagan, Mosteller, & Youtz, 1989) and a survey that was described in Chapter 5.
- The thesis provided a method for gaining reliable measures for the impact of a risk occurring by using a common flexible distribution (Beta) and through a survey to support various measures that lead to characterizing the distribution's shape.

The contributions are then in two categories:

1. Showing through a set of experiments that (a) the transformation of verbal expressions of probability from experts, and (b) the type of and the shape of the distribution chosen to model the impact of risk can potentially produce errors in the simulation of project costs that lead to incorrect decisions being made.
2. Establishing a survey to (a) derive the proper quantification approach to verbal expressions and distribution shapes, and (b) provide a guide to others that use verbal expressions on how to accomplish such survey (e.g. in FMEA analysis).

6.4 Recommended further research

The following is recommended for further research:

1. The distribution of costs for impact of risk occurrence or the distribution of costs for work package is generally subjectively solicited. In this thesis we attempted

to quantify verbal expressions of confidence and conservativeness to assist the simulation modeller in determining the proper shape of the distribution. While we have uncovered specific properties, we also note that there may be many other factors that determine the shape of the distribution and further research is required in this area.

2. Our work focused on quantifying risks using Level 2 (deterministic) and Level 3 (Monte Carlo simulation). While these are two prevalent approaches in industry there are other approaches that are proposed in the literature including fuzzy logic. The results of the survey can be modeled using fuzzy logic rather than statistical distribution. Further analysis is required to analyze the data for such purpose, however.
3. A computerized structured model for risks analysis may be beneficial as opposed to a spread sheet with an "add on" tool for risk analysis. The benefits may be in structuring the approach and minimizing the potential errors associated with manipulating data in a spreadsheet especially when the models are large.
4. This approach could be refined by understanding how to combine responses/data from multiple experts with differing levels of expertise since not all responses should be equally weighted.
5. Fuzzy membership functions can be used to address overlap of linguistic terms for example there is overlap among the percentages for the terms very likely and likely which can at times lead to confusion in choosing a variable. Addressing this could lead to an enhanced process.

6. Understanding how to automatically calibrate distributions following the implementation of a mitigation strategy. Currently, the process of risk quantification must be conducted from the beginning each time risks are mitigated or altered. Having an approach which can provide a means for a more continuous process could greatly enhance productivity.

7 Bibliography

AbouRizk, H., Kostelyk, J., AbouRizk, S. M., Smid, B., & Griffith, B. (2011). Improving Risk Analysis in the Construction Industry. *CSCE Annual Conference.* Ottawa, ON.

AbouRizk, S. M. (2009). *Risk Analysis for Construction Projects: A practical guide for engineers and project managers.* Edmonton: NSERC Industrial Research Chair in Construction Engineering and Management.

AbouRizk, S. M., Halpin, D. W., & Wilson, J. R. (1991). Visual Interactive Fitting of Beta Distributions. *Journal of Construction Engineering and Management*, 589-605.

Al-Bahar, J., & Crandall, K. (1990). Systematic Risk Management Approach for Construction Projects. *Journal of Construction Engineering and Management*, 533-546.

Arabian-Hoseynabadi, H., Oraee, H., & Tavne, P. (2010). Failure Modes and Effects Analysis (FMEA) for wind turbines. *Electrical Power and Energy Systems*, 817-824.

Biller, B., & Gunes, C. (2010). INTRODUCTION TO SIMULATION INPUT MODELING. *Winter Simulation Conference*, (pp. 49-58).

Brooke, P. J., & Paige, R. F. (2003). Fault trees for security system design and analysis. *Computers & Security*, 256-264.

Camburn, E. M., Huff, J. T., Goldring, E. B., & May, H. (2010). Assessing the Validity of an Annual Survey for Measuring Principal Leadership Practice. *The Elementary School Journal*, 314-335.

Chau, K. W. (1995). The validity of the triangular distribution assumption in Monte Carlo simulation of construction costs: empirical evidence from Hong Kong. *Construction Management and Economics*, 15-21.

Chiacchio, F., Compagno, L., D'Urso, D., & Mann, G. (2011). Dynamic fault trees resolution:A conscious trade-off between analytical and simulative approaches. *Reliability Engineering and System Safety*, 1515–1526.

Construction Industry Institute. (2012). *Applying Probabilistic Risk Management in Design and Construction Projects.* Austin, Texas: Construction Industry Institute.

Dhillon, B. S., & Singh, C. (1979). ON FAULT TREES AND OTHER RELIABILITY EVALUATION METHODS. *Microelectronics and Reliability*, 57-63.

Diekmann, J. E. (1983). PROBABILISTIC ESTIMATING: MATHEMATICS AND APPLICATIONS. *Journal of Construction Engineering and Management*, 297-306.

Dunjóa, J., Fthenakisb, V., & Vílcheza, J. A. (2010). Hazard and operability (HAZOP) analysis. A literature review. *Journal of Hazardous Materials*, 19-32.

Galway, L. (2004). *Quantitative Risk Analysis for Project Management.* RAND.

Han, S. H., & Diekmann, J. E. (2001). APPROACHES FOR MAKING RISK-BASED GO/NO-GO DECISION FOR INTERNATIONAL PROJECTS. *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, 300-308.

ISO. (2009). *ISO Guide 73.* Switzerland: ISO.

Kaisera, B., Gramlich, C., & Fo, M. (2007). State/event fault trees—A safety analysis model for software-controlled systems. *Reliability Engineering and System Safety*, 1521-1537.

Kamane, S. K., & Mahadik, S. A. (2013). Risk Management in Construction Industry. *Journal of Mechanical and Civil Engineering* , 59-65.

Kangari, R. (1995). Risk Management Perceptions and Trends of U.S. Construction. *Journal of Construction Engineering and Management*, 422-429.

Kwak, Y. H., & Ingall, L. (2007). Exploring Monte Carlo Simulation Applications for Project Management. *Risk Management*, 44 -57.

Law, A. M. (2012). A TUTORIAL ON HOW TO SELECT SIMULATION INPUT PROBABILITY DISTRIBUTIONS. *Winter Simulation Conference*, (pp. 1-15).

Law, A. M. (2012). A Tutorial on How to Select Simulation Input Probability Distrubutions. *Winter Simulation Conference*, (pp. 1-15).

Law, A. M., & Kelton, D. W. (1991). *Simulation Modeling & Analysis Second Edition.* McGraw-Hill Inc.

Megill, R. E. (1984). *An introduction to risk analysis.* Tulsa: Pennwell Books.

Oxford University Press. (2013). *Oxford Dictionaries Online.* Retrieved 08 20, 2013, from http://oxforddictionaries.com/definition/english/risk

Palisade Corporation. (2010). *Palisade.*

Project Management Institute. (2000). *A Guide to the Project Management Body of Knowledge.* Pennsylvania: Project Management Institute.

Reagan, R. T., Mosteller, F., & Youtz, C. (1989). Quantitative Meanings of Verbal Probability Expressions. *Journal of Applied Psychology*, 433-442.

Rhee, S. J., & Ishii, K. (2003). Using cost based FMEA to enhance reliability and serviceability. *Advanced Engineering Informatics*, 179-188.

Rodríguez, M., & Luis de la Mata, J. (n.d.). Automating HAZOP studies using D-higraphs.

Vankeerberghen, P., Vandenbosch, C., Smeyers-Verbeke, J., & Massart, D. (1991). Some robust statistical procedures applied to the analysis of chemical data. *Chemometrics and Intelligent Laboratory Systems*, 3-13.

Villacourt, M. (1992). *Failure Mode and Effects Analysis (FMEA): A Guide for Continuous Improvement for the Semiconductor Equipment Industry*. SEMATECH.

Wang, Y., Cheng, G., Hu, H., & Wu, W. (2012). Development of a risk-based maintenance strategy using FMEA for a continuous catalytic reforming plant. *Journal of Loss Prevention in the Process Industries*, 958-965.

8 APPENDIX 1: Survey

Understanding verbal expressions in terms of quantitative data

1. Demographic Information

1. What is your gender?

- Female
- Male

2. What is your age?

- 18 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64
- 65 to 74
- 75 or above

3. Are you currently a full time student?

- Yes
- No

4. What is your program of study

- MSc program in Construction Management
- MEng program in Construction Management
- PhD program in Construction Management

Other (please specify)

5. What is your current occupation? (If applicable)

Understanding verbal expressions in terms of quantitative data

2. Qualitative Expressions in terms of Single Values

Choose the appropriate range of values that you believe best represents the probability of occurrence for each descriptor. Please answer the following questions by checking one of the following options.

6. If someone told you an outcome was VERY LIKELY, what value would best represent the probability of that outcome occurring?

- ○ 0
- ○ 5
- ○ 10
- ○ 15
- ○ 20
- ○ 25
- ○ 30
- ○ 35
- ○ 40
- ○ 45
- ○ 50
- ○ 55
- ○ 60
- ○ 65
- ○ 70
- ○ 75
- ○ 80
- ○ 85
- ○ 90
- ○ 95
- ○ 100

7. If someone told you an outcome was LIKELY, what value would best represent the probability of that outcome occurring?

- ○ 0
- ○ 5
- ○ 10
- ○ 15
- ○ 20
- ○ 25
- ○ 30
- ○ 35
- ○ 40
- ○ 45
- ○ 50
- ○ 55
- ○ 60
- ○ 65
- ○ 70
- ○ 75
- ○ 80
- ○ 85
- ○ 90
- ○ 95
- ○ 100

8. If someone told you an outcome was SOMEWHAT LIKELY, what value would best represent the probability of that outcome occurring?

- ○ 0
- ○ 5
- ○ 10
- ○ 15
- ○ 20
- ○ 25
- ○ 30
- ○ 35
- ○ 40
- ○ 45
- ○ 50
- ○ 55
- ○ 60
- ○ 65
- ○ 70
- ○ 75
- ○ 80
- ○ 85
- ○ 90
- ○ 95
- ○ 100

Understanding verbal expressions in terms of quantitative data

9. If someone told you an outcome was UNLIKELY, what value would best represent the probability of that outcome occurring?

◯ 0	◯ 35	◯ 70
◯ 5	◯ 40	◯ 75
◯ 10	◯ 45	◯ 80
◯ 15	◯ 50	◯ 85
◯ 20	◯ 55	◯ 90
◯ 25	◯ 60	◯ 95
◯ 30	◯ 65	◯ 100

10. If someone told you an outcome was VERY UNLIKELY, what value would best represent the probability of that outcome occurring?

◯ 0	◯ 35	◯ 70
◯ 5	◯ 40	◯ 75
◯ 10	◯ 45	◯ 80
◯ 15	◯ 50	◯ 85
◯ 20	◯ 55	◯ 90
◯ 25	◯ 60	◯ 95
◯ 30	◯ 65	◯ 100

Understanding verbal expressions in terms of quantitative data

3. Qualitative Expressions in terms of Ranges

Choose the appropriate range of values that you believe best represents the probability of occurrence for each descriptor. Please answer the following questions by checking TWO OR MORE boxes.

11. If someone told you an outcome was VERY LIKELY, what range of values would best represent the probability of that outcome occurring?

☐ 0	☐ 35	☐ 70
☐ 5	☐ 40	☐ 75
☐ 10	☐ 45	☐ 80
☐ 15	☐ 50	☐ 85
☐ 20	☐ 55	☐ 90
☐ 25	☐ 60	☐ 95
☐ 30	☐ 65	☐ 100

12. If someone told you an outcome was LIKELY, what range of values would best represent the probability of that outcome occurring?

☐ 0	☐ 35	☐ 70
☐ 5	☐ 40	☐ 75
☐ 10	☐ 45	☐ 80
☐ 15	☐ 50	☐ 85
☐ 20	☐ 55	☐ 90
☐ 25	☐ 60	☐ 95
☐ 30	☐ 65	☐ 100

13. If someone told you an outcome was SOMEWHAT LIKELY, what range of values would best represent the probability of that outcome occurring?

☐ 0	☐ 35	☐ 70
☐ 5	☐ 40	☐ 75
☐ 10	☐ 45	☐ 80
☐ 15	☐ 50	☐ 85
☐ 20	☐ 55	☐ 90
☐ 25	☐ 60	☐ 95
☐ 30	☐ 65	☐ 100

Understanding verbal expressions in terms of quantitative data

14. If someone told you an outcome was UNLIKELY, what range of values would best represent the probability of that outcome occurring?

- [ ] 0
- [ ] 5
- [ ] 10
- [ ] 15
- [ ] 20
- [ ] 25
- [ ] 30
- [ ] 35
- [ ] 40
- [ ] 45
- [ ] 50
- [ ] 55
- [ ] 60
- [ ] 65
- [ ] 70
- [ ] 75
- [ ] 80
- [ ] 85
- [ ] 90
- [ ] 95
- [ ] 100

15. If someone told you an outcome was VERY UNLIKELY, what range of values would best represent the probability of that outcome occurring?

- [ ] 0
- [ ] 5
- [ ] 10
- [ ] 15
- [ ] 20
- [ ] 25
- [ ] 30
- [ ] 35
- [ ] 40
- [ ] 45
- [ ] 50
- [ ] 55
- [ ] 60
- [ ] 65
- [ ] 70
- [ ] 75
- [ ] 80
- [ ] 85
- [ ] 90
- [ ] 95
- [ ] 100

4. Preference for Qualitative Expressions

16. Which method did you find was easier to use to describe the probability of an outcome?

- ◯ No preference
- ◯ Responding using a range of values
- ◯ Responding using a single value

Understanding verbal expressions in terms of quantitative data

5. Understanding Likelihood

Answer the following by checking as many or as few boxes that seem appropriate.

17. Suppose someone told you an event had a 99% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

18. Suppose someone told you an event had a 90% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

19. Suppose someone told you an event had an 80% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

Understanding verbal expressions in terms of quantitative data

20. Suppose someone told you an event had a 60% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

21. Suppose someone told you an event had a 50% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

22. Suppose someone told you an event had a 40% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

Understanding verbal expressions in terms of quantitative data

23. Suppose someone told you an event had a 30% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

24. Suppose someone told you an event had a 20% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

25. Suppose someone told you an event had a 10% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] Very Likely
- [ ] Likely
- [ ] Somewhat Likely
- [ ] Unlikely
- [ ] Extremely Unlikely

Understanding verbal expressions in terms of quantitative data

26. Suppose someone told you an event had a 1% chance of occurring; which of the following expressions might you use to describe that probability? Check as many or as few that seem appropriate.

- ☐ Very Likely
- ☐ Likely
- ☐ Somewhat Likely
- ☐ Unlikely
- ☐ Extremely Unlikely

6. Understanding Confidence

Scenario: An estimate has been created for the building of a 1 km tunnel. You are a tunnelling expert who has been asked to assess the level of confidence you have regarding the line items of the current estimate. The risk analyst asks you how confident you are about line item A: Excavation of the working shaft for which the estimate you provided was $200,000.

27. You respond VERY CONFIDENT. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.

- ☐ $200 000 ±5% (i.e. $190 000- $210 000)
- ☐ $200 000 ±10% (i.e. $180 000- $220 000)
- ☐ $200 000 ±20% (i.e. $160 000- $240 000)
- ☐ $200 000 A ±25% (i.e. $150 000- 250 000)
- ☐ $200 000 ±50% (i.e. $100 000- $300 000)
- ☐ $200 000 ±75% (i.e. $50 000- 350 000)
- ☐ $200 000 ±95%(i.e. $10 000- 390 000)

Other (please specify)

28. You respond CONFIDENT. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.

- ☐ $200 000 ±5% (i.e. $190 000- $210 000)
- ☐ $200 000 ±10% (i.e. $180 000- $220 000)
- ☐ $200 000 ±20% (i.e. $160 000- $240 000)
- ☐ $200 000 A ±25% (i.e. $150 000- 250 000)
- ☐ $200 000 ±50% (i.e. $100 000- $300 000)
- ☐ $200 000 ±75% (i.e. $50 000- 350 000)
- ☐ $200 000 ±95%(i.e. $10 000- 390 000)

Other (please specify)

29. You respond SOMEWHAT CONFIDENT. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] $200 000 ±5% (i.e. $190 000- $210 000)
- [ ] $200 000 ±10% (i.e. $180 000- $220 000)
- [ ] $200 000 ±20% (i.e. $160 000- $240 000)
- [ ] $200 000 A ±25% (i.e. $150 000- 250 000)
- [ ] $200 000 ±50% (i.e. $100 000- $300 000)
- [ ] $200 000 ±75% (i.e. $50 000- 350 000)
- [ ] $200 000 ±95%(i.e. $10 000- 390 000)

Other (please specify)

30. You respond UNSURE. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] $200 000 ±5% (i.e. $190 000- $210 000)
- [ ] $200 000 ±10% (i.e. $180 000- $220 000)
- [ ] $200 000 ±20% (i.e. $160 000- $240 000)
- [ ] $200 000 A ±25% (i.e. $150 000- 250 000)
- [ ] $200 000 ±50% (i.e. $100 000- $300 000)
- [ ] $200 000 ±75% (i.e. $50 000- 350 000)
- [ ] $200 000 ±95%(i.e. $10 000- 390 000)

Other (please specify)

Understanding verbal expressions in terms of quantitative data

31. You respond VERY UNSURE. Which of the following distributions might you use to describe that probability? Check as many or as few that seem appropriate.

- [ ] $200 000 ±5% (i.e. $190 000- $210 000)
- [ ] $200 000 ±10% (i.e. $180 000- $220 000)
- [ ] $200 000 ±20% (i.e. $160 000- $240 000)
- [ ] $200 000 A ±25% (i.e. $150 000- 250 000)
- [ ] $200 000 ±50% (i.e. $100 000- $300 000)
- [ ] $200 000 ±75% (i.e. $50 000- 350 000)
- [ ] $200 000 ±95%(i.e. $10 000- 390 000)

Other (please specify)

7. Understanding Confidence

You provided a most likely value for an estimate for the shaft at $265 000 and indicated the following:

32. There was a 95% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- [ ] Very confident
- [ ] Confident
- [ ] Somewhat confident
- [ ] Unsure
- [ ] Extremely unsure

33. There was a 85% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- [ ] Very confident
- [ ] Confident
- [ ] Somewhat confident
- [ ] Unsure
- [ ] Extremely unsure

34. There was a 75% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- [ ] Very confident
- [ ] Confident
- [ ] Somewhat confident
- [ ] Unsure
- [ ] Extremely unsure

Understanding verbal expressions in terms of quantitative data

35. There was a 65% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

36. There was an 55% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

37. There was an 45% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

Understanding verbal expressions in terms of quantitative data

38. There was an 35% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

39. There was a 25% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

40. There was a 15% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- ☐ Very confident
- ☐ Confident
- ☐ Somewhat confident
- ☐ Unsure
- ☐ Extremely unsure

Understanding verbal expressions in terms of quantitative data

41. There was a 5% chance that the estimate is between $250 000 and $275 000? Which of the following expressions would you use to describe your confidence of the estimate? Check as many or as few that seem appropriate.

- [ ] Very confident
- [ ] Confident
- [ ] Somewhat confident
- [ ] Unsure
- [ ] Extremely unsure

8. Understanding Conservativeness

Scenario: An estimate has been created for the building of a 1 km tunnel. The cost of the tunnel can range between a minimum cost of $200 000 and a maximum cost of $300 000. The most likely cost of the tunnel is $250 000. You, an estimator, have been called upon to conduct your own estimate of the tunnel

42. You have estimated the cost of the tunnel and believe that your estimate is CONSERVATIVE. Which of the following dollar values best depicts your estimate. Check as many or as few that seem appropriate.

- ☐ $200 000
- ☐ $210 000
- ☐ $220 000
- ☐ $230 000
- ☐ $240 000
- ☐ $250 000
- ☐ $260 000
- ☐ $270 000
- ☐ $280 000
- ☐ $290 000
- ☐ $300 000

43. You have estimated the cost of the tunnel and believe that your estimate is SOMEWHAT CONSERVATIVE. Which of the following dollar values best depicts your estimate. Check as many or as few that seem appropriate.

- ☐ $200 000
- ☐ $210 000
- ☐ $220 000
- ☐ $230 000
- ☐ $240 000
- ☐ $250 000
- ☐ $260 000
- ☐ $270 000
- ☐ $280 000
- ☐ $290 000
- ☐ $300 000

44. You have estimated the cost of the tunnel and believe that your estimate is CLOSE to the actual cost. Which of the following dollar values best depicts your estimate. Check as many or as few that seem appropriate.

- ☐ $200 000
- ☐ $210 000
- ☐ $220 000
- ☐ $230 000
- ☐ $240 000
- ☐ $250 000
- ☐ $260 000
- ☐ $270 000
- ☐ $280 000
- ☐ $290 000
- ☐ $300 000

Understanding verbal expressions in terms of quantitative data

45. You have estimated the cost of the tunnel and believe that your estimate is SOMEWHAT RISKY. Which of the following dollar values best depicts your estimate. Check as many or as few that seem appropriate.

- [ ] $200 000
- [ ] $210 000
- [ ] $220 000
- [ ] $230 000
- [ ] $240 000
- [ ] $250 000
- [ ] $260 000
- [ ] $270 000
- [ ] $280 000
- [ ] $290 000
- [ ] $300 000

46. You have estimated the cost of the tunnel and believe that your estimate is RISKY. Which of the following dollar values best depicts your estimate. Check as many or as few that seem appropriate.

- [ ] $200 000
- [ ] $210 000
- [ ] $220 000
- [ ] $230 000
- [ ] $240 000
- [ ] $250 000
- [ ] $260 000
- [ ] $270 000
- [ ] $280 000
- [ ] $290 000
- [ ] $300 000

Understanding verbal expressions in terms of quantitative data

9. Understanding Conservativeness

Scenario: You, an estimator, have been called upon to conduct an estimate of a 5 km bridge. You have estimated the cost to be $500 000.

47. You believe that there is a 95% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

48. You believe that there is a 75% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

49. You believe that there is a 50% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

Understanding verbal expressions in terms of quantitative data

50. You believe that there is a 25% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- [ ] Conservative
- [ ] Somewhat Conservative
- [ ] Close
- [ ] Somewhat Risky
- [ ] Risky

51. You believe that there is a 15% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- [ ] Conservative
- [ ] Somewhat Conservative
- [ ] Close
- [ ] Somewhat Risky
- [ ] Risky

52. You believe that there is a 5% chance that the actual cost of the estimate is UNDER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- [ ] Conservative
- [ ] Somewhat Conservative
- [ ] Close
- [ ] Somewhat Risky
- [ ] Risky

10. Understanding Conservativeness

Scenario: You, an estimator, have been called upon to conduct an estimate of a 5 km bridge. You have estimated the cost to be $500 000.

53. You believe that there is a 5% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

54. You believe that there is a 15% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

55. You believe that there is a 25% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

Understanding verbal expressions in terms of quantitative data

56. You believe that there is a 50% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

57. You believe that there is a 75% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

58. You believe that there is a 95% chance that the actual cost of the estimate is OVER your estimated cost. Which of the following expressions best depicts your level of conservativeness. Check as many or as few that seem appropriate.

- ☐ Conservative
- ☐ Somewhat Conservative
- ☐ Close
- ☐ Somewhat Risky
- ☐ Risky

9 APPENDIX 2: Histogram of Responses for “Likely” Stemmed Expressions

The histograms of the responses for each of the subjective terms are for the terms of ‘very likely’, ‘likely’, ‘somewhat likely’, ‘unlikely’, and ‘extremely unlikely’ are shown below:

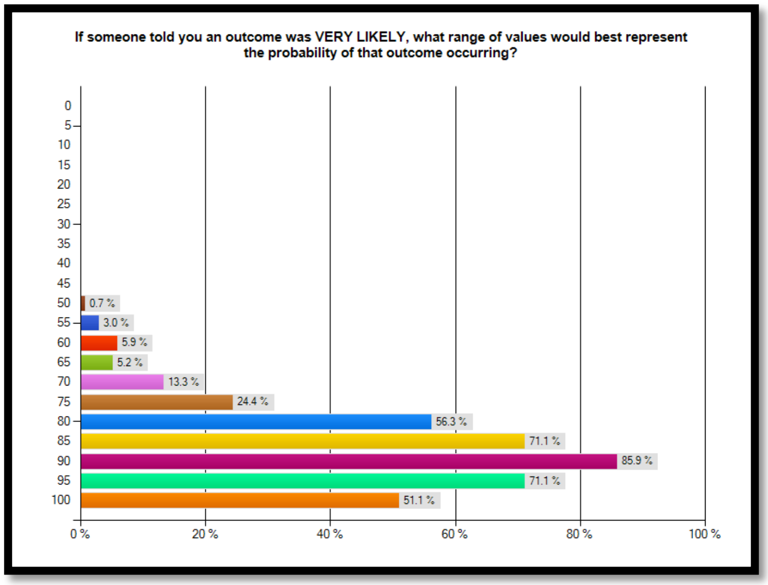


Figure 45: "very likely" in terms of quantitative data

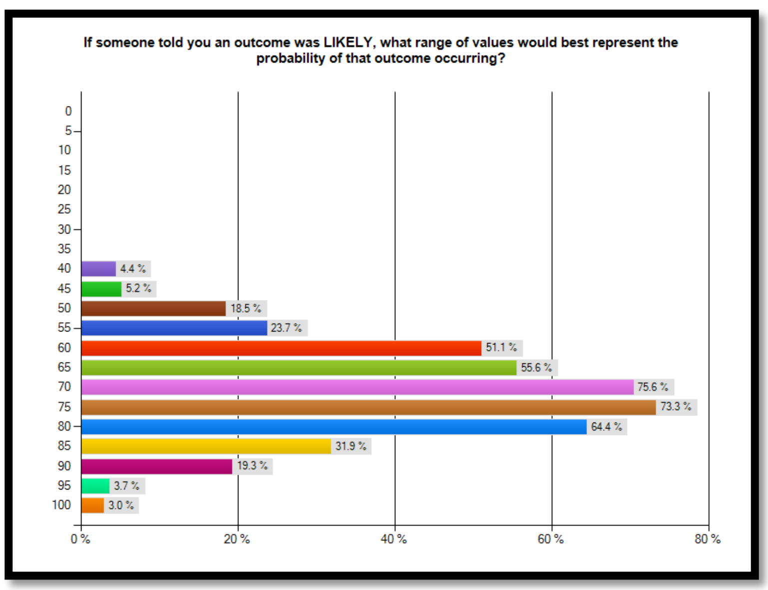


Figure 46: "likely" in terms of quantitative data

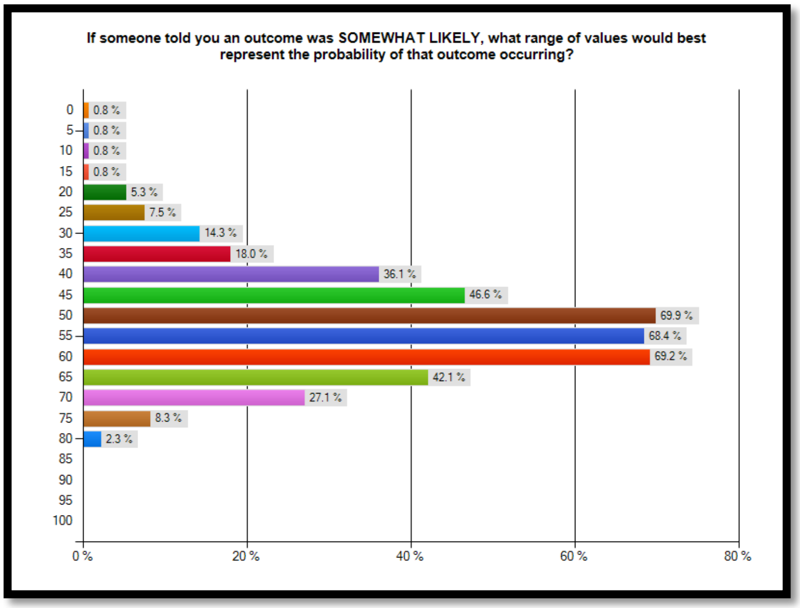


Figure 47: "somewhat likely" in terms of quantitative data

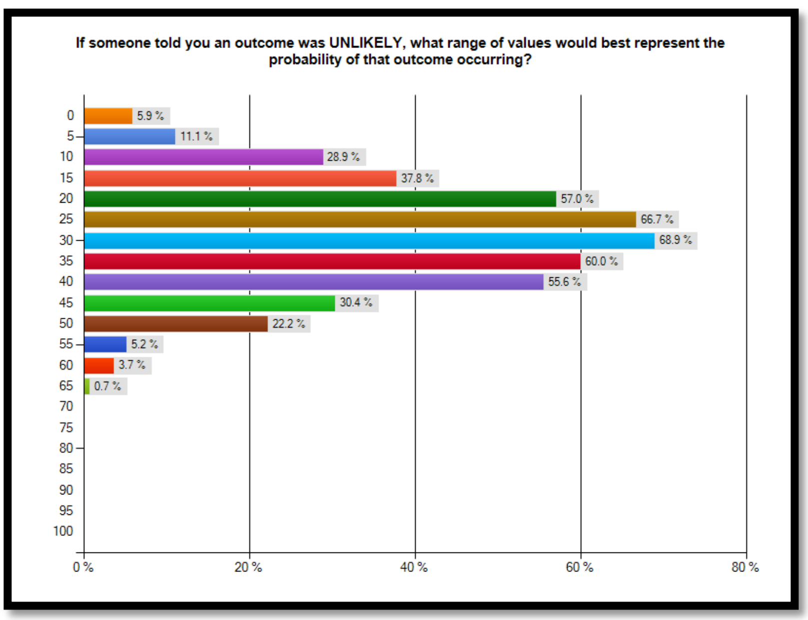


Figure 48 "unlikely" in terms of quantitative data

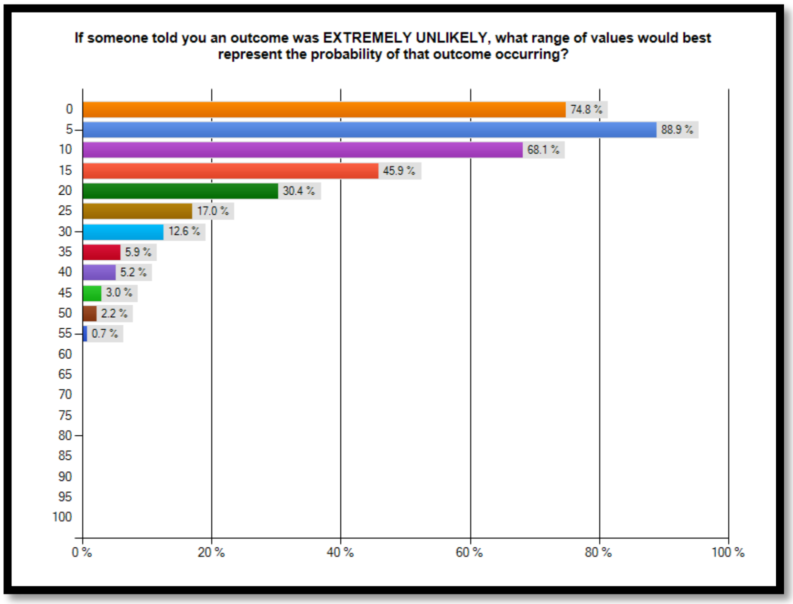


Figure 49 "extremely unlikely" in terms of quantitative data