Basic Concepts of Risk Analysis

M. Vidyasagar

Cecil & Ida Green Chair
The University of Texas at Dallas
Email: M.Vidyasagar@utdallas.edu

October 1, 2011
Outline

1 Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Outline

1 Preliminaries
   - Definitions
     - Illustration of Terms
     - A Baby Example
2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)
3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Definiton of Terms

**Risk:** An issue, item or event that may or may not occur, such that its occurrence has a negative impact.

An ‘opportunity’ can be thought of as ‘negative risk’ – something whose occurrence has a positive benefit.
**Risk Assessment**: Quantifying, to the extent possible based on current knowledge, (i) the probability of occurrence of an event, and (ii) its potential impact.

**Risk Analysis**: The entire *process* of identifying the sources of risk, quantifying them to the extent possible, communicating the risks to all stakeholders, etc.

**Risk Management**: Coming up with one or more *strategies* to reduce, or in some cases (theoretically at least) entirely eliminate the risk.
Outline

1 Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Suppose you are building a large factory, the construction of which will take two years. You need to buy cement at various times during this construction process. (Say you buy once a month for the next month’s construction demands.) You finance your construction by borrowing at a fixed interest rate at the outset of construction. This will be done via issuing bonds at a predetermined interest rate.

What are risk assessment, risk analysis, and risk management in this scenario?
Risk *assessment* is identifying the fact that the price of cement could go up, thereby causing the cost of the project to increase.
Usage of Terms: Example 1 (Cont’d)

Risk *assessment* is identifying the fact that the price of cement could go up, thereby causing the cost of the project to increase.

Risk *analysis* is constructing a time series (random variables indexed by time) of the projected price of cement at monthly intervals for the next two years into the future. The time series would be a series of 24 random variables and their probability distributions. Further, you would analyze the impact on your project cost in terms of expected cost, variance, tail estimate, etc.
Usage of Terms: Example 1 (Cont’d)

Risk *assessment* is identifying the fact that the price of cement could go up, thereby causing the cost of the project to increase.

Risk *analysis* is constructing a time series (random variables indexed by time) of the projected price of cement at monthly intervals for the next two years into the future. The time series would be a series of 24 random variables and their probability distributions. Further, you would analyze the impact on your project cost in terms of expected cost, variance, tail estimate, etc.

Risk *management* is entering into a series of ‘futures’ contracts for cement at predetermined rates each month for the next two years; also determining, on the basis of some algorithm, how much cement to buy, and at what price. Finally, management would also include determining how to price the initial bond issue to finance the project.
Usage of Terms: Example 2

You undertake a large project with several steps in it, as below.

![PERT Chart](image)

You will incur a cost penalty if the time overruns, and will get a bonus if the project is completed ahead of time.
Risk assessment consists of coming up with the above PERT (Program Evaluation Review Technique) chart. Analyzing the overall project, breaking it down into its constituent subprojects, and identifying the dependencies (which package needs to be completed in order to execute which other package) is already a major part of risk assessment.

Realizing that the time taken to complete each work package is random, and that the overall time to completion of the project is also random, completes the risk assessment.
Risk analysis consists of:

- Coming up with plausible distributions for the time needed to complete each work package. (PERT uses a very simple-minded distribution, on which more later.)
- Combining these, either via analytical techniques or via simulation, to come up with the cumulative distribution of the overall project time.
- Mapping this distribution into your likely bonus-penalty distribution.
Risk management consists of some or all of the following:

- Taking insurance against possible cost/time overruns, weighing the insurance premiums against the likelihood of the insurance being invoked.

- Assessing the relative merits and demerits of various ways of speeding up the project (e.g. hiring extra staff, paying workers overtime).
Outline

1 Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Let us return to the flow chart shown earlier.

Fig. 1: PERT Chart

* Numbered rectangles are nodes and represent events or milestones.
* Directional arrows represent dependent tasks that must be completed sequentially.
* Diverging arrow directions (e.g. 1-2 & 1-3) indicate possibly concurrent tasks.
* Dotted lines indicate dependent tasks that do not require resources.
The Bonus/Penalty Function

You get a bonus if the project is completed before 42 months, and incur a penalty if the project takes longer than 60 months. For in-between times, you get neither.

Suppose the bonus/penalty function is

\[ B = \begin{cases} 
42 - T & \text{if } T \leq 42, \\
60 - T & \text{if } T \geq 60, \\
0 & \text{if } 42 \leq T \leq 60. 
\end{cases} \]

Note that bonus is positive if \( T \leq 42 \) and negative (penalty) if \( T \geq 60 \).
The Bonus/Penalty Function: Depiction

\[ B(T) \]
Based on past experience, suppose you come up with minimum and maximum times for each of the various steps. You conclude that the minimum time to complete the project is $T = 36$ months, while the maximum is $T = 72$ months.

So you may get a bonus of up to $6$ million, or face a penalty of up to $12$ million, or get neither.

Is this information enough for risk management?

No – you need more detailed information.
Minimum and maximum times to complete each step are a crude substitute for the *probability distribution* of the time to complete that step.

If you had probability distributions for the time to complete each step (even really naive ones like uniform, or triangular), how can you combine them to produce the probability distribution of the overall time to complete the project?

This is the focus of various techniques such as PERT, Monte Carlo simulation, etc.
Suppose we have combined the probability distributions of the individual steps in the project to come up with the overall project time to completion, as shown below.

What is the distribution function of the bonus?
From the above, we know that

\[
\Pr\{T \leq 42\} = 0.1, \ Pr\{T \geq 60\} = 0.2, \ \text{so} \ \Pr\{42 \leq T \leq 60\} = 0.7.
\]

Mapping this into the bonus, we conclude that

\[
\Pr\{B \geq 0\} = 0.1, \ \Pr\{B \leq 0\} = 0.2, \ \Pr\{B = 0\} = 0.7.
\]

The distribution function of the bonus \(B\) looks as shown on the next slide.
Note that the expected value of the bonus is negative! This is because of the ‘long tail’ of the time to completion.

So risk management becomes essential.
Outline

1 Preliminaries
   • Definitions
   • Illustration of Terms
   • A Baby Example

2 Various Methods for Project Management
   • Gantt Chart
   • PERT Charts
   • Critical Path Method (CPM)

3 Simulation Techniques
   • Monte Carlo Simulation
   • Joint Distributions and Copulas
The Gantt chart was invented by Henry Gantt in the 1910’s. It represents each activity in a project as a horizontal bar, with start and end times indicated, as shown on the next slide.
Example of a Gantt Chart

Project Development Schedule

<table>
<thead>
<tr>
<th>Project Steps</th>
<th>Qtr 1</th>
<th>Qtr 2</th>
<th>Qtr 3</th>
<th>Qtr 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore Market Need</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop Concept for Product</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Begin Development Cycle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop GUI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User Interface Test Evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha Version Release</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality Assurance Testing Phase 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fix Outstanding Problems from Alpha</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beta Version Release</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality Assurance Testing Phase 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fix Outstanding Problems from Beta</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design Box and CD Labels</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Begin Advance Advertising Campaign</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCS Preparation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final Quality Assurance Testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCS Release</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Production and Packaging</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Development | QA Testing | Marketing | Box Art | Milestones

M. Vidyasagar Basic Concepts of Risk Analysis
Pros and Cons of Gantt Chart

**Advantage:** Preparing a Gantt chart forces you to think through all the work components and to make a proper schedule.

**Disadvantages:**

- It cannot capture *sequential dependence*, i.e. Step A must be completed before Step B, etc.
- No scope whatsoever for ‘randomness’ (or variations) in schedule.

Nevertheless, it was a very useful concept in its day!
Outline

1 Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
PERT = Project Evaluation Review Technique

Starting point: Break down overall project into a sequence of sub-projects.

Graphical representation also captures *sequencing information* — a vast improvement over Gantt charts!

In fact what we saw earlier was a PERT chart!
Various Methods for Project Management
Simulation Techniques

PERT Chart Example (Reprise)

Fig. 1:
PERT Chart

* Numbered rectangles are nodes and represent events or milestones.
* Directional arrows represent dependent tasks that must be completed sequentially.
* Diverging arrow directions (e.g. 1-2 & 1-3) indicate possibly concurrent tasks.
* Dotted lines indicate dependent tasks that do not require resources.

M. Vidyasagar
Basic Concepts of Risk Analysis
PERT is not just about capturing sequential information. For each stage, ‘experts’ were also asked to estimate the maximum time (or cost), minimum time, and average time. Then a simple distribution (usually a ‘triangular’ distribution’ shown on next slide) is fit to the time to complete that stage.

If only minimum and maximum times are available, one can also use a ‘uniform’ probability distribution (shown in next after next slide).

Graphical structure makes it easy to combine individual estimates into an overall estimate.
Triangular Distributions Used in PERT

\[ \Phi_T(T) \]

\[ \phi_T(T) \]

\[ T_{min} \quad T_{avg} \quad T_{max} \]
Uniform Distributions Used in PERT

\[ \Phi_T \]

\[ T_{\min} \quad T_{\max} \]

\[ \phi_T \]

\[ T_{\min} \quad T_{\max} \]
Limitations of PERT

- PERT was evolved for an era when computation was difficult! Today we need not worry about doing complex computations!
- PERT cannot take into account *dependence among different steps* (dependence among different random variables). In other words, PERT assumes that all random variables are independent.
If Step from 1 to 3 takes longer than expected, wouldn’t we expect the step from 3 to 5 also to take longer than expected?
Outline

1. Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2. Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3. Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Critical Path Method (CPM)

In this method, one again lays out the components of the overall project, and works out the average (expected) time needed to complete each task.

The longest path through the graph is the **critical path**, because there is no slack in it at all.

Ergo: Pay attention to the critical path and don’t worry about the rest (until they become critical!).

M. Vidyasagar
The top path is the critical path.
Outline

1 Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2 Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3 Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Suppose we have a complex event, a combination of individual random variables, and no hope of ‘closed-form formulae’ to come up with the probability distribution of the event. What can we do? We can just ‘simulate’ by generating a large number of samples of each random variable, computing the outcome of the event, and plotting. We illustrate with the same PERT diagram.
**PERT Diagram (Reprise)**

*Numbered rectangles are nodes and represent events or milestones.*
*Directional arrows represent dependent tasks that must be completed sequentially.*
*Diverging arrow directions (e.g., 1-2 & 1-3) indicate possibly concurrent tasks.*
*Dotted lines indicate dependent tasks that do not require resources.*

Fig. 1: PERT Chart
Suppose the path $1 \rightarrow 2$ is random with a triangular distribution, the path $1 \rightarrow 3$ is random with a uniform distribution, and so on. (These would need to be generated by collating expert opinion.) How can we come up with an overall probability distribution of the time to completion?
First Naive Approach

For each step, *generate at random* a set of completion times that are distributed according to that distribution.

Alternately, just choose some step size $\epsilon$, say 10% or 0.01, and choose the corresponding percentile values.

If we choose ten percentile, then we get 10 samples for each of the random variables.

We illustrate the percentile approach, also known as ‘gridding’, with uniform and triangular distributions.

Assign these values in all possible combinations, compute the time to complete the project for each combination, and we get an overall histogram of the time to complete.
The preimages of the percentiles of distribution function are the grid points.
Percentile Approach to Sampling (Cont’d)

The grid points are no longer uniformly spaced.

\[ \Phi_T \]

\[ T_{min} \]

\[ T_{max} \]
Combinatorial explosion!

There are 10 paths in the PERT diagram (and it’s a toy example). If we use 10 samples for each path, we need $10^{10}$ samples!

Monte Carlo simulation is meant precisely to cope with this combinatorial explosion.

*Sample uniformly* amongst all these points, and use only those samples.

Practical detail: Actually we don’t even need to generate the $10^{10}$ samples first.

Theory allows us to say just how many samples we need to draw, to get a desired level of accuracy of the estimate, with a given level of confidence.
Outline

1. Preliminaries
   - Definitions
   - Illustration of Terms
   - A Baby Example

2. Various Methods for Project Management
   - Gantt Chart
   - PERT Charts
   - Critical Path Method (CPM)

3. Simulation Techniques
   - Monte Carlo Simulation
   - Joint Distributions and Copulas
Suppose $X$ is a r.v. with distribution function $\Phi_X$. If $U$ is uniformly distributed over the unit interval $[0, 1]$, then $\Phi_X^{-1}(U)$ is ‘distributionally equal’ to $X$.

What this means is that if we take equally spaced samples of $U$, say $u_1, \ldots, u_n$, and invert them via $x_i = \Phi_X^{-1}(u_i)$, then the $x_i$ are the grid points of $X$.

This is what we did in the earlier example.
Inverting the Distribution Function to Get Grid Points

\[ \Phi_T \]

\[ T_{min} \]

\[ T_{max} \]
If $X, Y$ are \textit{independent} random variables, then we can generate grid points in the $X, Y$ space by generating grid points of $X$ and of $Y$ separately.

But what if they are \textit{coupled} random variables?

This is where the copula is useful.
A two-dimensional copula is a function $C : [0, 1]^2 \rightarrow [0, 1]$, or a bivariate distribution function, which we can think of as the joint distribution of two uniformly distributed r.v.’s $U, V$ that are not necessarily independent.

If $X, Y$ are the r.v.’s of interest, the copula satisfies

$$
\Phi_{X,Y} = C(\Phi_X, \Phi_Y).
$$

In other words

$$
\Pr\{X \leq x, Y \leq y\} = \Pr\{U \leq \Phi_X(x), V \leq \Phi_Y(y)\}.
$$
Use of the Copula in Simulation

Many packages (e.g. excel) have only a limited number of distributions if at all.

By using the above reasoning, we can generate coupled random variables with arbitrary distributions.