The Flaw of Averages in Project Management

By Philip Fahringer, John Hinton, Marc Thibault, and Sam Savage

New technologies and management practices offer a better approach to understanding and controlling project risk. Interactive simulation provides intuitive risk dashboards that can be used to detect and manage hidden risks, even for those with no statistical training. The distribution string (DIST) provides the needed standard for communicating project input and output uncertainties, so that experts in the organization can unambiguously share their knowledge of uncertainty in durations, costs, and other factors with managers at all levels.

> "Revolutionary new technologies and business practices make working with distributions easy, intuitive, and compelling."

The Flaw of Averages

How long will it take and how much will it cost?

There isn't just one answer to each of these questions. Each question has several possible answers, each with its own probability of being correct. The answer is not one number—it's a ***probability distribution***. "I remember studying probability distributions in my statistics class," you say, "and, as I recall, I preferred going to the dentist." Don't worry—revolutionary new technologies and business practices make working with distributions easy, intuitive, and compelling.

The common practice of reducing an uncertainty to a single ***best guess*** eliminates a lot of information, which leads to the flaw of averages, a set of systematic errors that occur when a single number, usually an ***average*** (also known as an ***expected value***) is substituted for a distribution. The resulting miscalculations have dire consequences in many areas of commerce and government and are particularly prevalent in project management. In fact, the flaw of averages explains why most projects are behind schedule, beyond budget, and below projection.

We will begin with an actual project that involves aircraft maintenance and for which we will substitute hypothetical data. As airplanes age, they frequently undergo a major overhaul. For the purposes of our high-level discussion, we have reduced the project of refurbishing one plane to the following steps: (1) Removing the wings (2a and b); rebuilding or replacing the left and right wings as necessary; (3) overhauling the fuselage, checking for corrosion, and structural, electrical, and hydraulic issues; (4) reattaching the wings; and (5), final assembly and inspection. Steps 2a, 2b, and 3 occur in parallel, whereas the others occur in a series. Because of the precise alignment required, we will assume that you cannot move on to step 4 until both of the wings and the fuselage have been completed. The dependency graph is shown in Figure 1.

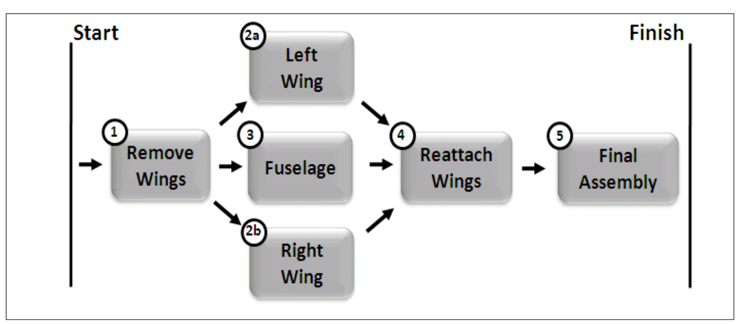

Figure 1: Dependency graph for refurbishing an aircraft.

Now, imagine you are managing a refurbishing project for an entire fleet of a particular type of aircraft. You have many years of experience and have collected plenty of data on the durations of the tasks involved, and you are preparing a bid for all 50 planes in the fleet.

The boss enters your office and asks you how long the project will take.

"I don't know," you say, "there are so many uncertainties involving the numbers, types of landings, and the corrosion." "Give me a number!" barks the boss, "What are we paying you for?"

"Would you settle for averages?" you ask timidly.

"Yes, if that's the best you can do," he responds.

"That's easy," you say, proud of the data you have gathered over the years. "Steps 1, 4, and 5 each consistently takes about one week to complete. The wings and fuselage are where the big surprises come up, but, on average, they each take five weeks. Because we are doing the wings and fuselage in parallel, we should be able to finish one plane every 8 weeks, on average." This time line is often depicted as a GANTT chart and is shown in Figure 2.

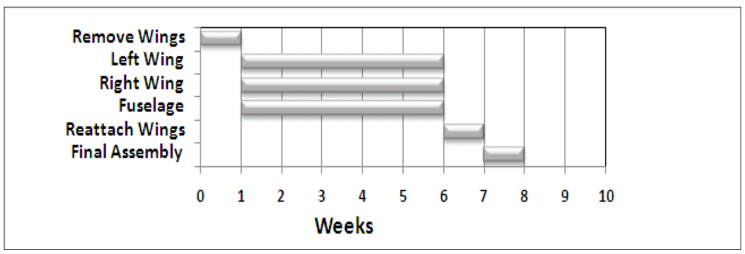

Figure 2: GANTT chart for refurbishing a plane.

Uh, oh! You have just run afoul of the flaw of averages. You cannot reattach the wings until both the wings and fuselage have been completed, so having all three of these elements come in below their five-week average duration is like flipping three heads in a row on a coin and only happens, on the average, once in eight tries! In such a situation, depending on the uncertainty, the average time to complete these three tasks may be significantly greater than the average of any individual task and this just scratches the surface of the problem. The flaw of averages has many other manifestations as well, which can only be detected by explicitly modeling the uncertainty of the tasks, as we will describe later.

In actuality, based on the ***average*** duration per process step, the maintenance firm wrote a firm-fixed price contract that earned it a pre-specified amount per aircraft. There was a penalty if the delivery of a plane did not occur within an agreed-on time frame. On the upside, however, early delivery would save costs and lead to extra profit. The good news is that several years into the contract, the average task durations were roughly as initially predicted. The bad news was that the average time to complete each plane was significantly higher than anticipated—the firm lost millions of dollars, and four vice presidents lost their jobs.

Next, we turn from the actual example with hypothetical data, as described above, to a hypothetical example based on actual data. We will use a residential construction project to show how new technologies can help you better manage the uncertainty.

A married couple and the husband's widowed mother, each has a house on the family's property. The mother is frail, and the couple has decided to build a third, larger house on the same plot of land on which they will all live together. Once the new house is built they will remodel the two older houses and rent them out for income. The dependency graph of the project is shown in Figure 3.

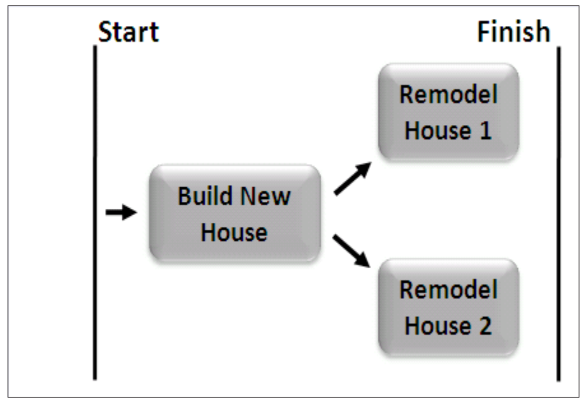

Figure 3: Dependency graph.

Once the new house has been completed, the couple and the husband's mother will move out of their own houses, and the remodeling will begin on the existing houses in parallel.

Again, imagine you are bidding for the job. From historical data, you know that houses of this size take an average of 12 months to complete, at an average cost of US$400,000. These sorts of remodels typically take 6 months and cost US$200,000. Because the remodels were going to be done in parallel, the job would take 18 months and cost US$800,000, based on the averages. The GANTT chart illustrating this is shown in Figure 4.

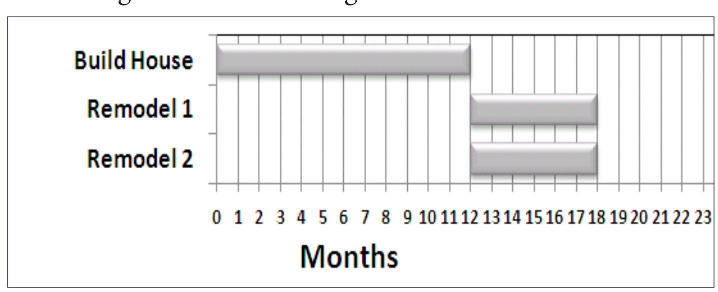

Figure 4: GANTT chart.

You feel you can't win the contract if you take more than a 10% profit margin; so, if there was no uncertainty, you would bid US$880,000 for the job and tell the client you can complete the project in 18 months. However, there **is** uncertainty and, to complicate matters, the owner of the property has expressed concern about the project coming in late and has requested a liquidated damages penalty of US$1000 per day if the job runs beyond the 18-month projected completion time. Needless to say, based on the averages, you don't expect this to happen, but the penalty clause would make anyone a little nervous. So, the questions are:

1. How many months should you expect the project to take?
2. What should you bid?
3. What should you expect in profit?
4. What are the financial risks?

Let's assume you negotiate an additional US$20,000 into the contract to cover the potential penalty, for a total cost of US$900,000, and hope for the best. In the following section we propose a more systematic approach to answering the above questions.

Probability Management—A Cure for the Flaw of Averages

Data management is for storing, retrieving, and auditing numbers, as in "Give me a number so I can perpetuate the flaw of averages." ***Probability management*** is for storing, retrieving, and auditing probability distributions, so you can cure the flaw of averages. But, what do you do with distributions? You feed them into your plans to simulate the uncertainties you may face.

The methods of probability management had their origins in a technique called Monte Carlo simulation, which has been applied to project management in the past, and for which specialized software has been developed. Probability management has been enabled by recent advances in three areas: software technology, information representation, and managerial protocols (described below): together, we believe they will make Monte Carlo simulation far more practical in the context of project management.

- Software—Recent advances in software allow for ***interactive simulation***, in which thousands of scenarios may be run through a model before the user's finger has left the <Enter> key. This provides more insight into uncertainty and risk than the traditional simulation tools and makes a connection between the user's mind (the seat of their intellect) and their gut (the seat of their pants).
- Information—The ***DIST*** standard ***distribution string*** encapsulates thousands of numbers into one data element to represent an uncertainty, which allows probability distributions to be shared across information networks.
- Management—With sharing come responsibility, accountability, and authority. The ultimate authority in an organization in the area of probability falls to the ***chief probability officer (CPO)***. Although we do not expect firms to open up a new C-level position for this function, someone in the organization must ultimately serve as the CPO.

Simulation as a Card Game

Imagine a deck of cards containing different possible *costs* for building the house, and another deck with different possible *durations* for building the house. Now imagine four more decks, containing the possible *costs* and *durations* for each of the two remodels (Figure 5).

One can think of a simulation as a card game, which is played as follows:

1. Draw one card from each of the six decks.
2. Calculate and record the resulting duration and the total cost of the entire project (including late penalties) based on the numbers you've drawn.
3. Return to step one until all cards have been used then proceed to step 4.
4. Average all the costs and durations you have recorded to get an estimate of the true average outcomes of the project. For this project, the flaw of averages ensures that they will be very different from the outcome associated with the average assumptions. You may also analyze the risk of bad outcomes.

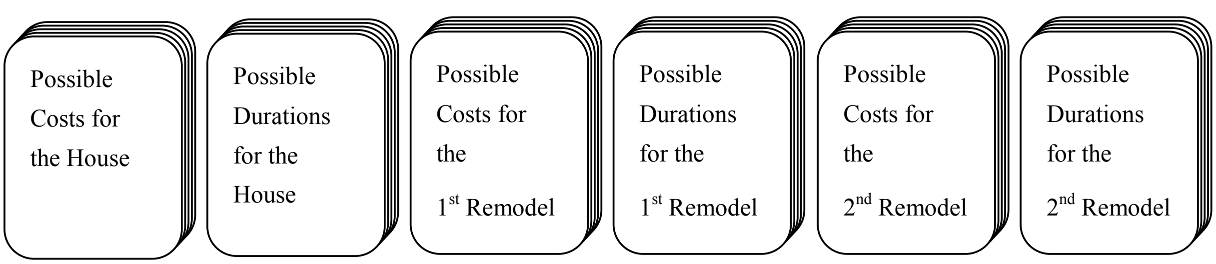

Figure 5: A representation of the uncertainties in the project.

The interactive simulation model in Excel that accompanies this article contains the equivalent of the six decks, with 1000 cards in each deck. Thanks to recent technology, each deck fits into one cell on the worksheet!

The Distribution String

The ***distribution string***, or ***DIST***, is a new data type used for storing large numbers of scenarios. We began the article by describing several possible answers, each with its own probability of being right. A DIST represents this concept. Instead of a single value, it can store thousands of equally likely potential values in one data element. DISTs allow experts in an organization to collect their knowledge for use by others in such activities as estimating and bidding. This is a fundamental change from the traditional Monte Carlo simulation, in which distributions cannot be easily shared between applications and users.

The Data

The statistical data used in this example came from De Mattei Construction, a large residential contractor in San Jose, California, USA, and is based on 10 new home construction projects and 29 remodels. As we will see, even a small data set like this can be invaluable in project planning. Most construction firms keep records of the actual costs and durations of previous jobs. De Mattei Construction, however, takes an additional crucial step; they also record what they had originally predicted for these projects, and this enables them to track and improve the accuracy of their forecasts, which are critical to cash flow projections and staffing decisions. Storing both forecasts and actual results is a critical function of a good CPO, yet many large organizations ignore this step.

We applied De Mattei's data to create DISTs of both the cost and duration for our hypothetical structures. One hand of the card game is known as a trial of the simulation. The first three and last three trials are shown in Table 1. For example, in the second trial of the simulation, the house cost US$449,000, and took 20 months to complete. Portions of the actual DISTs are displayed in Table 2.

The histograms of these values along with the average values appear in Figure 6. We assume that each of the remodels is an independent draw from the same cost and duration distributions so they have the same histogram. Note that the average is where the histogram would balance if it was made of a solid material. Although averages are often used as a substitute for the entire shape, this can be very misleading.

Also, note that the relative durations of remodels are much more uncertain than those of new dwellings. A contractor must make major assumptions about the general health of the existing building before he or she starts. After the contractor opens up the walls, dry rot, termites, construction code violations, mold, or unsupported foundations may be discovered. Many of these problems require a redesign, an approved inspection by an engineer, and

Value	Cost of House in Thousands	Duration of House in Months	Cost of 1st Remodel in Thousands	Duration of 1st Remodel in Months	Cost of 2nd Remodel in Thousands	Duration of 2nd Remodel in Months
1	US$243	28	US$104	4	US$126	2
2	US$449	20	US$157	11	US$277	4
3	US$343	12	US$380	8	US$118	3
:	:	:	:	:	:	:
:	:	:	:	:	:	:
:	:	:	:	:	:	:
998	US$501	13	US$180	12	US$118	3
999	US$391	8	US$284	4	US$247	4
1000	US$391	8	US$181	4	US$145	3

Table 1: Selected values of cost and duration uncertainties.


```
<dist name="House Cost" avg="4.00000000E+002" min="2.43253009E+002" max="6.66594544E+002" count="1000" type="Double" ver="1.1">AAB8iDxITVlcaFll
<dist name="House Duration" avg="1.20000000E+001" min="7.18770000E+000" max="2.83877000E+001" count="1000" type="Double" ver="1.1">//+ftDpPDiM
<dist name="Remodel 1 Cost" avg="2.00000000E+002" min="1.03823177E+002" max="4.37161622E+002" count="1000" type="Double" ver="1.1">AAApGNPCFg
<dist name="Remodel 1 Duration" avg="6.00000000E+000" min="5.46376191E-001" max="1.18797095E+001" count="1000" type="Double" ver="1.1">RTLp1qW
<dist name="Remodel 2 Cost" avg="2.00000000E+002" min="1.03823177E+002" max="4.37161622E+002" count="1000" type="Double" ver="1.1">EL6FMgq4bhs
<dist name="Remodel 2 Duration"6.00000000E+000" min="5.46376197E-001" max="1.18797095E+001" count="1000" type="Double" ver="1.1">K4ZWsTuWTaj/
```


Table 2: The six DIST standard distribution strings.

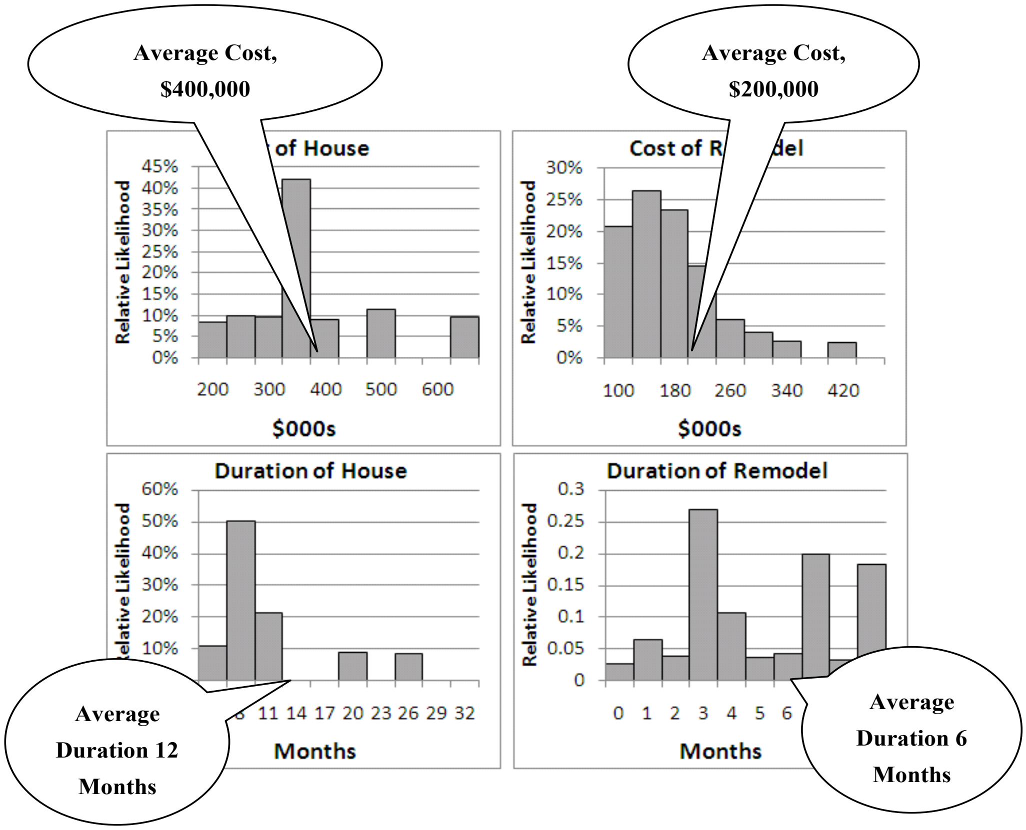

Figure 6: Distributions of costs and durations.

building code approval, which explains the big spikes at the right end of the remodel duration histogram.

The Simulation

Construction Project.xlsm contains an interactive simulation model of the project in which the six uncertain costs and durations are linked to DISTs. If you want to follow the discussion below, go to ProbabilityManagement.org and download the workbook from the models page (note that you must have macros enabled in Microsoft Excel for Windows). Figure 7 displays the model with average values of the uncertain inputs, which resulted in finishing the project in 18 months, with a cost of US$800,000, no penalty, and a profit of US$100,000.

So far so good, but now let the card games begin! The input scroll bar lets you work your way through all 1000 trials, one at a time (which we don't recommend), but it is helpful to review a few of these trials. Trial 2, for example (Figure 8), has the project finishing in 31.3 months, with a cost of US$884,000 and a penalty of US$399,000, leaving the contractor with a loss of US$383,000. This is a shame!

Trial 7 displays a more positive picture (Figure 9), with the project coming in on time and below budget, with a profit of US$202,000.

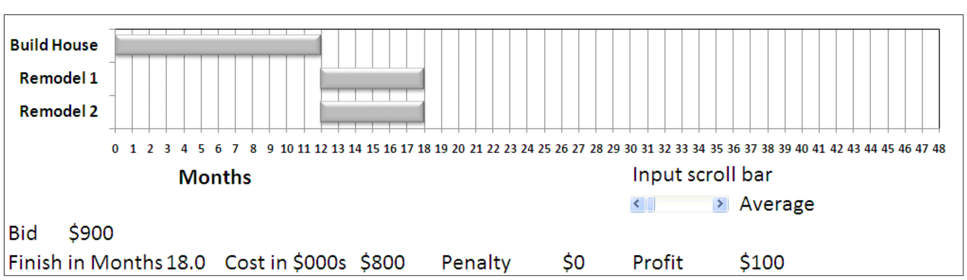

Figure 7: Model with average inputs.

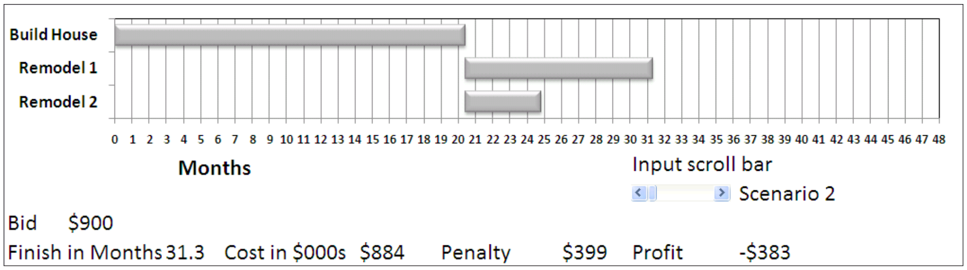

Figure 8: Results of trial 2.

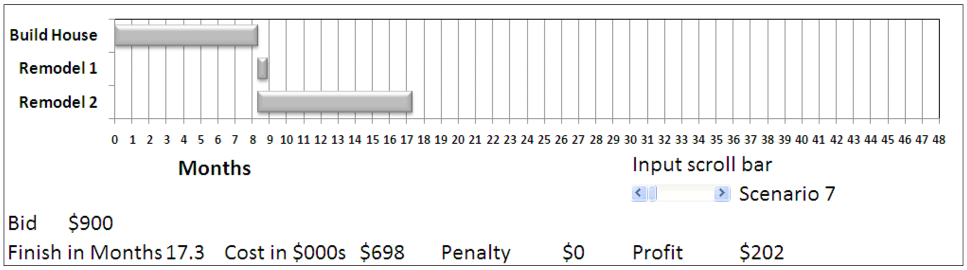

Figure 9: Results of trial 7.

Keeping in mind that each of these results is as likely as any other, the degree of uncertainty, based on real data, is sobering; however, the real lesson of simulation comes from viewing the aggregate results over all 1000 trials. A large construction firm can survive one or two profit-losing jobs, as long as it makes enough money in the long term. Table 3 displays the results of the average inputs, along with the average results, taking into account all 1000 trials.

Note that the average time to complete the project is almost 20 months, rather than the anticipated 18 months. The average cost is, indeed, US$800,000, but the average penalty caused by the delay is US$102,000, which leads to an average loss of US$2000 instead of the expected US$100,000 profit. There is no way to make up in volume of business, an average loss of $2000 per job. This is a classic case of the *flaw of averages*.

Next, let's take a look at the full distribution of profit, which is also based on all 1000 trials (Figure 10). The histogram displays the relative likelihood of different possible levels of profit and is quite asymmetric. It is bad enough to lose an average US$2000, but the fact that the downside extends to a loss of US$600,000, whereas the upside is only about US$300,000 should be a warning sign before starting this job. The cumulative distribution displays the chance that

Finish in Months	18.0	Cost in $000s	$800	Penalty	$0	Profit	$100
Average Finish	19.9	Average Cost	$800	Avg. Penalty	$102	Avg. Profit	-$2

Table 3: The results of the average inputs versus the average results.

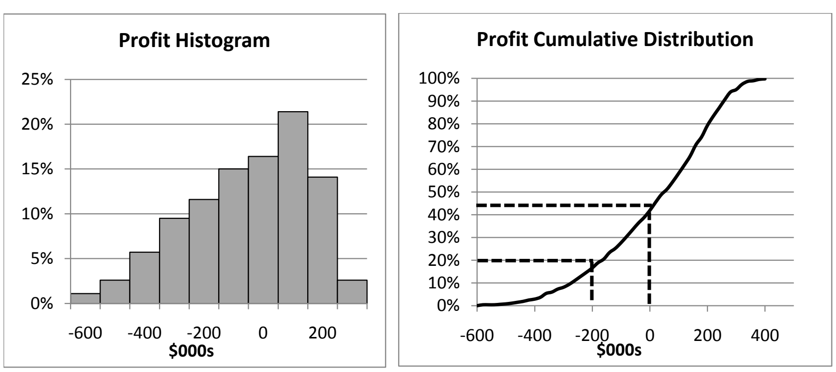

Figure 10: The histogram and cumulative distribution of profit.

the profit will come in below any particular level. We can see, for example, that there is roughly a 45% chance of losing money and a 20% chance of losing US$200,000 or more. You may use the scroll bar in the model to explore the full range of profit percentiles.

Interactive Simulation—Changing the Game

We have already highlighted one of the differences between probability management and traditional Monte Carlo simulation, which is, the ability to store and retrieve probability distributions as data. As mentioned earlier, another important difference is the emergence of interactive simulation, which is like being able to adjust the rules of the card game in real time and then instantly viewing the results of playing 1000 hands.

In this example, let's assume you haven't actually submitted your disastrous bid and the client is still open to some negotiation. How can you change the game? Certainly, you could raise your bid but you are afraid that your competitors will then underbid you. You can also try to negotiate a less stringent late penalty.

Initially, the Excel model has the parameters as displayed in Table 4, which is a penalty of US$1000 for every day the project exceeds 18 months in duration.

Penalty Parameters	
Deadline (Months)	18
Daily Penalty 000s	$1.00

Table 4: The penalty parameters.

Suppose you can get the client to accept either reducing the daily penalty to US$500 or pushing back the penalty deadline to 24 months, but not both. What would each of these alternatives be worth to you on average? One of these changes will raise your average profit to US$57,000 and the other to US$49,000; to find out which one is better, download ***Construction Project.xlsm*** (the macros must be enabled in Microsoft Excel) and try different values of the penalty parameters. Every time you change a value in cell K15 or K16, all 1000 trials are instantly run through the model.

Probability management provides the tools to not just estimate the outcome of a project but to manage the risk in your estimates as well.

The DIST Data Type

The distribution strings stored on their own tab in ***Construction Project.xlsm***, encapsulate thousands of numbers into one data element stored in what is known as an *XML format*. The basic idea involves translating numbers into characters through a code table, as shown in Figure 11. [1]

These characters are then stored with header information in a long string; an actual DIST from this model is displayed in Figure 12. The string begins with header information containing the name of the DIST, its average value, and other meta-data such as the maximum, minimum, and number of trials. This is followed by the body of the DIST, which contains the numerical trials encoded as characters.

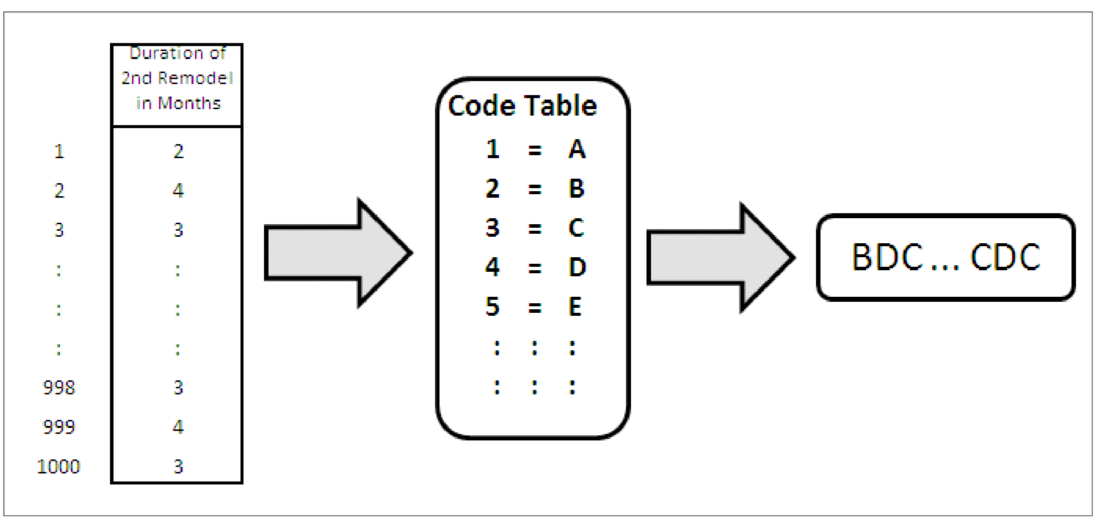

Figure 11: Translating numbers into characters when creating a DIST.

[1] The actual encoding is more complex. Visit ProbabilityManagement.org for more information.

Figure 12: The DIST representing the duration of the house construction.

References

Kwaka, Y.H., & Ingall, L. (2007). Exploring Monte Carlo simulation applications for project management. *Risk Management* (9) pp., 44–57. Retrieved from http://www.palgrave-journals.com/rm/journal/v9/n1/full/8250017a.html#bib29D

http://www.oracle.com/us/products/applications/042528.pdf

http://www.palisade.com/riskproject/

Savage, S. L. (2009). *The flaw of averages: Why we underestimate risk in the face of uncertainty*. Hoboken: John Wiley & Sons.

Savage, S.L., Scholtes, S., & Zweidler, D. (2006). Probability management. *OR/MS Today*, February, 33(1). See also ProbabilityManagement.org

About the Authors

Philip Fahringer is an operations analyst with Lockheed Martin Corporation and has been there since 2006. He is retired from the Navy, having served the last ten years as a logistics analyst, determining logistics requirements for Naval and Joint Forces contingency scenarios, and developing budget priorities. Currently, Philip is the lead developer for the Affordable Readiness Model for Lockheed Martin, and focuses on balancing costs and outcomes under conditions of uncertainty. He holds master's degrees in operations research and strategic planning.

John Hinton is vice president and chief financial officer of De Mattei Construction, a large residential contractor in San Jose, California, USA. At De Mattei, John has managed and tracked the financials of over 1000 projects, has worked at several start-up companies, and has held positions at Coopers & Lybrand and Bechtel, where he worked in the consulting, information management, and project management areas. He has served in the U.S. Army in military intelligence in Germany and Vietnam, retiring as a captain with a bronze star. John holds an MBA degree from Stanford University, Palo Alto, California, USA.

Marc Thibault is an independent consultant with a twenty-year practice that focuses on technical analysis, design, and planning. His clients have included a large fraction of the Canadian federal government's departments and a variety of high-tech companies. Marc's earlier experience includes over a decade of marketing and technology roles at Xerox and senior management in two high-tech start-up companies. He has a physics degree from Loyola College in Montreal, Canada. Marc is the author of the "Art of the Plan" blog at goodplan.ca, and is on a mission to fix project planning and the systemic errors that result in high-risk plans and unattainable targets.

Sam L. Savage is chairman of Vector Economics, a consulting professor at Stanford University's School of Engineering, and a fellow at Cambridge University's Judge Business School. His recent book, *The Flaw of Averages*, describes a class of systematic errors we make in the face of uncertainty, and explores significant new technologies for illuminating and managing risk. In 2008, in collaboration with Oracle Corp. SAS Institute, and Frontline Systems, he led the development of the distribution string (DIST), a revolutionary open standard for describing probability distributions. Sam holds a PhD in the area of computational complexity from Yale University.