Project Risks Analysis: Sensitivity Analysis and Correlations

Intaver Institute Inc.
303, 6707, Elbow Drive S.W.
Calgary, AB, T2V0E5, Canada
tel: +1(403)692-2252
fax: +1(403)259-4533
sales@intaver.com
www.intaver.com

Identifying which activities have the most affect on the project and understanding how activities are correlated with each other is important for establishing project priorities. Our judgment about correlation and causation is affected by a number of biases, such as illusory and invisible correlations, covariation assessment, and others. Sensitivity analysis helps to discover correlations within a project.

What Are Correlations And Why Do We Need to Analyze Them?

On January 21, 2005 the state of Maine officially launched its new Medicaid Claim System (Holmes, 2006). This web-based software system was designed to process $1.5 billon in Medicaid claims and payments per year. However, a problem surfaced within the first few days: the new system started to mistakenly reject a huge number of claims from service providers: doctors, hospitals, dentists, and others. By the end of March, the system suspended (unpaid) 300,000 claims. It soon became clear that the system was plagued by numerous software bugs. With the huge number of claims rejected, several doctors were forced to close their doors; others had to take loans to continue operation. This was only a small symptom of a larger problem. The Medicaid program accounted for one-third of the entire state budget and, as a result of the software problem, Maine's financial stability was in jeopardy. By the end of summer 2005, the system had a backlog of 647,000 unpaid claims representing $310 million in back payments. It soon became clear that the problem was due to poor project management, and in particular, communication between state's IT team, contractors and business users. One of the most critical issues was related to the prioritization of the project tasks. The IT staff had worked extremely hard trying to fix every possible bug without a clear understanding of what were the most important issues to focus their efforts on. In essence, they had given *all the issues the same priority without regard to how they affected the system. A new project manager quickly moved to prioritize the project's activities. The software developers were asked to fix the bugs that were causing the largest number of suspended claims. Other activities that were assigned a lower priority such as improvements to the system's user interface would be addressed later. By improving the lines of communication, setting up priorities, and improving other project management processes, the claim system started to get back on track.*

One of the problems that project managers and teams face is that human beings are not built to think about different things simultaneously. Project teams cannot deal with all the tasks and all the issues at the same time. Tasks that have the most effect on the project schedule and project deliverables require the most attention. Activities that require the most attention often carry the most risk and may require significant mitigation efforts; therefore, we must identify them before the project starts. Essentially, we need to be able to analyze the **correlations** between the main project parameters (duration, cost, finish time, and others) and each task's parameters.

Another important issue is to understand how correlations between different activities affect the project. You know that your supplier may be very busy. If your activity is delayed because the components did not arrive on time, other activities, which use component from the same supplier, can also be delayed. Similarly there can be synergies between different projects. These correlations will significantly affect the course of the project.

In addition, when you plan your project, you should determine which project management activities and procedures would be most effective for this particular project. For example, for some teams a ten-minute meeting every morning to coordinate team activities is very useful, but in others it can be an utter waste of time. In other words, you need to find a correlation between project management activities and project results. To do this, we need to:

1. Identify which activities have the most affect on the project to set up our priorities.
2. Understand how different projects and activities within the project are correlated with each other and analyze the effect of these correlations on the project.
3. Identify particular project management processes and procedures that address the effects of these correlations so that they can be adopted by your team or organization.

Sources of Correlation in Projects

In what manner are different tasks in a project correlated? Schuyler (Schuyler, 2001) names three different sources:

Common Drivers: Different project parameters may share a common influence. For example, changes in project scope will affect many tasks. If there is an issue with resource, all tasks in which this resource is employed will be affected.

Common Constraints: If different project activities are competing for the same resources, these activities will be correlated. For example, technological constraints will affect all of the activities using this technology.

Common Cause: Results of one activity will lead to changes in another activity. For example, design changes that occur during the course of one activity, may affect many other activities.

Psychology of Correlation and Causation

Usually project managers believe that setting up priorities is a trivial task, which they can accomplish quickly and without any tools. But if it is so easy, why do so many projects start without any prioritization when the failure to do so often results in missed deadlines, poor quality, and huge cost overruns? The answer lies, as it does in so many things, in human nature.

After several major delays in your recent projects, you decided to analyze how your subcontractor affected the project deadline. As part of this investigation, you created 1 (adopted from Plous, 1993):

	Project deadline is missed	Project was on schedule
Subcontractor is involved in a project	8 times	2 times
Subcontractor is not involved in a project	4 times	1 time

Table 1 Example of covariation assessment

A short glance of this table you may lead you to believe that your subcontractor is a major source of your problems; the largest number of failures (eight) occurred when subcontractor was involved in the project. Moreover, you may be misled by the higher numbers in the first row (with subcontractor) than the second row (without subcontractor). In reality the project success ratio (four failures versus one success) is the same, regardless of whether the subcontractor was involved or not. In psychology, this phenomenon is referred to as "**covariation assessment**" or the analysis of whether two parameters are related to each other. If we only use a small data set, as in the example above, it is easy to draw misleading conclusions. Therefore, to understand the correlation between two parameters, you need to consider all of the available information.

Another interesting observation is **people pay more attention to events that have negative effects than events that have positive effects;** although, events that have positive effects may yield as much information. For example, we pay more attention to events that caused a delay, than to events that caused the project to be on time. Moreover, a clear understanding of these positive events (opportunities) may be very important for the planning of future projects (Nisbett and Ross, 1980).

Sometimes a small number of samples can significantly skew our judgment about correlations. This situation is very common in project management because usually the total number of projects performed by the same organization is limited. In addition, people tend to forget what happened in previous projects and the farther back they occurred in the past, the more likely they will be overlooked, even if they are more relevant to the current project. For example, if you have problems with a supplier twice in the past three projects, this does not mean that you need to switch to another supplier. The information you have may not be enough to analyze this correlation. First, you need to understand the underlying reasons that are causing the problems with the supplier. It might be the supplier, but perhaps your organization's procurement system has been tardy in sending out its purchase orders causing late delivery of supplies.

Sometimes people tend to find correlations where they do not exist. Psychologists refer to this phenomenon as "**illusory correlations**" (Chapman and Chapman, 1971). For example, a programmer is always late for work and the sales of the software that he is working on are not doing well. Is there any correlation between these two events? Most likely not; however, if the manager focuses a lot of attention on team discipline and sales numbers, he or she will **expect** the two numbers to be related and perceive that the slow sales are due to the inattentiveness of the team members.

The opposite phenomenon is "**invisible correlations**", or correlations that go unnoticed because people do not expect them to be related. For example, team members have been complaining to senior management about a certain project manager's lack of people skills. The results of this lack of people skills in unclear, as some of the projects managed by this individual have problems, but some do not. In addition, the project manager in question has worked in the organization for a long time and is viewed by senior management as a valuable contributor to the business. Is there a relationship between the manager's perceived inability to work with people and problems with projects? Perhaps, but because not all of the project manager's projects have problems and the project manager's high status in the organization, this correlation may become invisible. Essentially, as senior management has not expected to find a correlation between the project manager's people skills and poor results, it has not found it. Unfortunately, invisible correlations often go on to cause even larger problems. In this case, the effects of the project manager's poor people skills are cumulative and, over time, the project team's effectiveness may diminish as members leave to other teams or organizations in search of a less stressful environment.

If there is a correlation between two variables, this does not mean that one variable has caused another one. In other words, correlation does not necessarily equal **causation.** Remember, a common cause is only one possible source of correlation. Every day, we are constantly bombarded with all sorts of dubious claims of causation. The media is constantly announces the research findings that show the health benefit of different types of foods: eat this and you will live 20% longer or fail to include this in your diet and expect to die 20% younger (which by the way, is often the same product). Although for some groups of products, these claims may be true, in many cases, other factors, which can have a positive or negative effect on our health, are not taken into account. For example, research into the health benefits of red wine may not include an analysis of situations when people do not drink wine because they are already sick. Similar situations occur in project management. We may think the project succeeded because we created and managed a risk list. Correlations between project success rate and the presence of a risk list is not enough to make a judgment that a risk list led to a positive result.

Causation is often used by people who are trying to sell you something. Salesmen want you to believe that their technology or services are the cause of successful projects. Project management consultants will claim that twenty-three projects with an overall budget of $3 billon succeeded because of their involvement. It may be true, but these projects may have succeeded without the consultants. It is very difficult to prove or disprove. As you can see, there are many biases related to correlations and causation. Is any way we can overcome these mental traps?

How to Improve Your Judgment

From this discussion, it must be obvious that having a proper understanding of correlations is not easy. People have a lot of difficulty accurately judging project correlations. For example, you have just implemented a new tool and you must decide whether it is important to provide formal training for your team or count on them to learn it themselves as they go. Basically, you want to determine the correlation between training and project success rates. Training could delay the project by a few weeks, but without training the project may completely fail. So it is important to take the time to research this issue. Crocker has proposed a six-step process on how to perform this analysis (Crocker, 1982).

1. **Decide what information is relevant**. At this stage you need to discover which data should be used for the analysis. Do you have enough data within the organization? What happens with the project if training is not provided? What about the people who will provide the training: experience, education, etc.?
2. **Get samples or observations (randomly, if possible) related to your problem**. This is the data collection phase. You need to collect as much data as you can from all previous training sessions and how they affected project results. At this stage you are filling your table with data.
3. **Interpret and classify these observations**. Should data be placed in separate categories: long/short training, training for recent grads and experienced engineers? What counts as a project success: are there categories of project success? At this stage you design a table similar to Table 1.
4. **Estimate frequencies: how often a positive correlation occurred or did not occur**. At this stage, you fill out the table.
5. **Integrate your estimates to come up with a measure that can be used to make a judgment**. For example, you can come up with some rough averages. With training there is a 60% chance that project is on budget. Without training there is a 30% chance that project is on budget.
6. **Finally, use the integrated estimate to make a judgment**.

As you can see, this is a complex mental process, which cannot be done intuitively. Unfortunately, wrong judgments related to correlation and causation are very common even with training. Also, as this is such a complex process, can anyone perform this type of analysis? The answer is definitely yes, but only as long as the following conditions are met:

- Reliable data is available.
- The problem is serious enough to justify spending scarce resources on this analysis. The reality is if you performed an advanced analysis on everything, you

would never start a project. The project would end up being the analysis of how to do the project.

Fortunately, there are a number of tools that you can use to improve your judgment.

Sensitivity Analysis

> Sensitivity analysis determines how sensitive results of the analysis are to uncertainties in input variables.

If you have a valuation model related to your project, you can determine which parameters in your model have the most potential to affect your project. Simply change one parameter, for example a task duration, while keeping all of the other parameters the same and see how this change affects your results. Those inputs parameters that cause the largest fluctuations in the results will be most important.

For example, you have a model that calculates revenue based on unit price, cost of fuel, and cost labor. For these three parameters, you have three estimates: low, base, and high. To start, calculate your revenue using the base values of all the parameters. In this trial, revenue equals $400,000. Then recalculate the revenue multiple times by changing one parameter, for example unit price, and keep the other parameters constant. When we change the unit price, low unit price revenue is $180,000; high unit price revenue is $650,000.

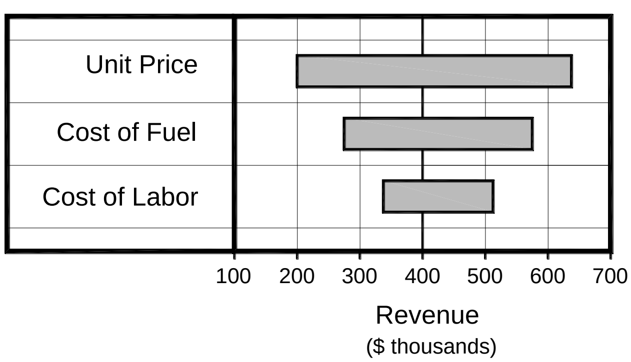

Figure 1. Tornado Diagram

These results can be displayed on a diagram (Figure 1). The ranges of revenue are displayed as a bar associated with each parameter. Parameters are sorted in such a way that parameters with a higher range of revenue will be displayed on top. This diagram looks like a “tornado”; therefore, it is called a **tornado diagram**.

The same results can be presented in another format (Figure 2). Because it can resemble a spider web it is called a **spider diagram**. One problem with tornado diagrams, it is difficult to visualize how an increase in one parameter leads to an increase or decrease in another parameter. Spider diagrams address this issue. If you calculate results with the low, high and an additional intermediate estimate, it is possible to grasp the non-linear relationship between inputs and outputs. In our example, the relationship between the cost of fuel and revenue is non-linear, as increase in fuel costs above a certain level does not lead to further reduction in revenue.

Using sensitivity analysis, it is possible to determine which parameters are the most important. They are:

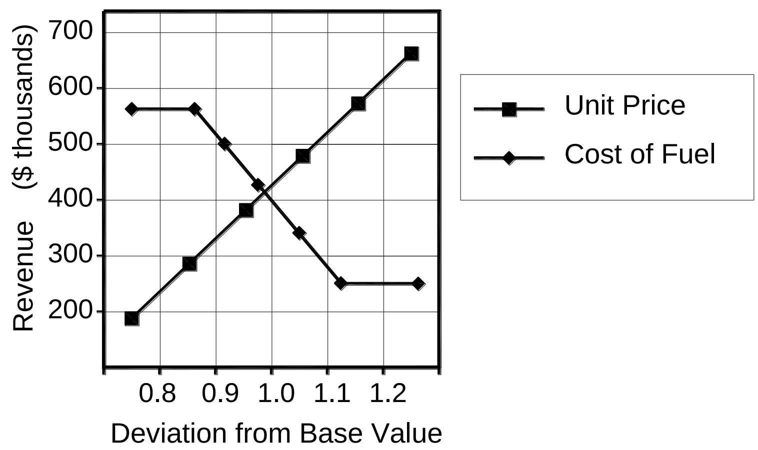

Figure 2. Spider diagram

- The upper bars of a tornado diagram
- The line with the steepest slope on a spider diagram

This type of sensitivity analysis is widely used in economic evaluations, where we have a quantitative economic model of the project. Because these models may be very complex, tornado and spider diagrams have proven to be very useful tools.

Please note the following issue regarding this type of sensitivity analysis. As with any quantitative analysis, the results of sensitivity analysis are "as good as model is". If the model does not take into account an important parameter, you will not be able to see it using a tornado or spider diagram.

How do you develop low, base, and high estimates for input variables? Why, for example, did you select a low and high cost estimates as 80% and 120% of the base value? In theory, these estimates should come from an analysis of the underling nature of the parameter. In reality, people often do not bother to learn about the nature of the variable or the data is simply not available. To get their low and high estimates, they just multiply their base estimate by 0.8 and 1.2 respectively. This is a classic case of **anchoring**: and almost always leads to misleading results.

Sensitivity analysis as one of quantitative risk analysis techniques is recommended in the PMBOK ® Guide (Chapter 11).

Quantitative Analysis of Correlations

Let us assume that you spent a great deal of time and effort collecting data on how the average experience of team members is related to their hourly rate, the cost of materials, and the cost of the project. These results can be presented on scatter diagrams (Figure 3):

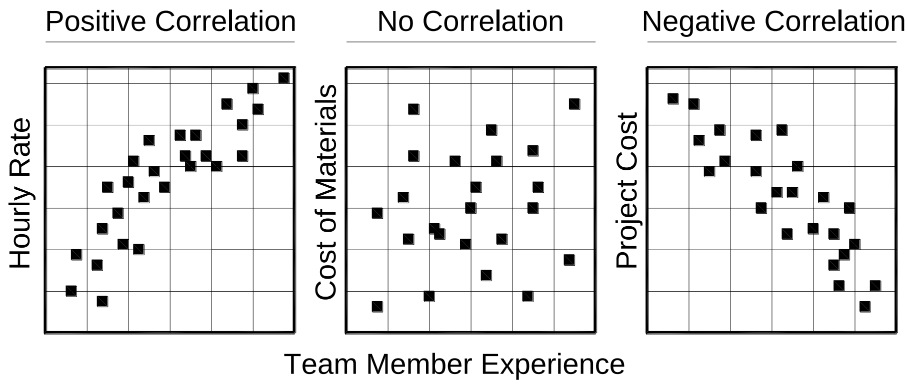

Figure 3. Scatter diagrams showing different types of correlation

Each point on these diagrams represents an actual data sample. Here is what we discover:

1. In most cases you have to pay more (higher hourly rate) for more experienced team members. There is a **positive correlation** between the two variables.
2. Cost of materials has **no correlation** with team member experience.
3. Overall project costs may be reduced with more experienced team members. Here is an example of a **negative correlation**. By the way, this is also a classic example of an invisible correlation in project management. Managers know that experienced workers will cost more, but they don't want to admit that a more experienced team will most likely lower overall project costs as they tend to focus on a short term goals (reduced monthly payroll) rather than long-term results.

It is possible to use a number referred to as the **correlation coefficient** to define these correlations. The coefficient for a strong positive correlation (higher values of one parameter are always associated with higher values of another parameter) is 1. The coefficient for strong negative correlation is –1. If there is no correlation, the coefficient is 0. There are many statistical formulas used to calculate the correlation coefficient, among which is the **Pearson and** **Spearman Rank Order** correlation. You can find the actual formula for these correlations in statistical textbooks (Freedman, Pisani and Purves, 1998). It should be noted that these formulas allow you to calculate correlation coefficients for different types of data, such as experience and cost in our example. Experience would be defined in years and cost in dollars.

Now that we know how to calculate correlations between two different sets of data, we can use it to answers our main questions: what is most important and how are different project variables related to each other?

Crucial Tasks

> Uncertainties associated with crucial tasks should be analyzed first.

Crucial tasks are those tasks that can have the most effect on your project parameters (duration, cost, success rates, and so on). As we have been describing correlations, crucial tasks represent a positive correlation between the task parameters and project parameters.

To explain crucial tasks correlation to project duration, we use the "Spring analogy" (Figure 4). Let's assume that each task within a Gantt chart is a spring in a big system of springs. When we start moving springs back and forth, we find that some of them significantly affect the movement of the full spring system, while others do not. The amount of movement depends on how springs are connected to each other (links between tasks) and how flexible they are (the type of risks and uncertainties assigned to them).

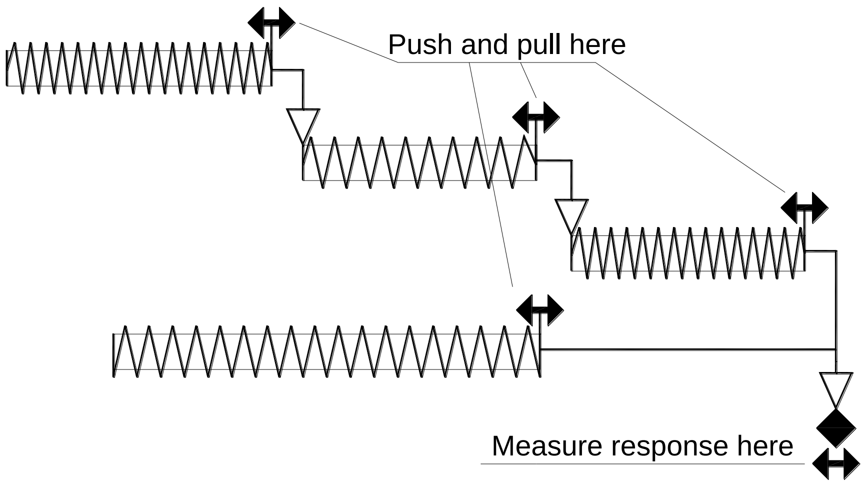

Figure 4. Spring analogy for crucial tasks

Here are a few questions commonly asked about crucial tasks:

Question: How do you determine crucial tasks?

Answer: Crucial tasks are by-products of Monte Carlo simulations. A Monte Carlo simulation produces an array of project durations, cost, success rates, and other parameters. You will also have an array of task durations, cost, and other parameters as a result of sampling. Now input these arrays into a correlation coefficient formula (if you are really ambitious, you could do it manually, but it is better to use one of the many applications on the market that can perform this for you) and it will return a correlation coefficient. Those tasks with the highest coefficients are crucial tasks.

Question: What is the relationship between crucial tasks and critical tasks?

Answer: Critical tasks lie on the critical path. Crucial tasks do not necessarily lie on the critical path. They are called crucial to distinguish them from tasks that have been identified using critical path method (CPM).

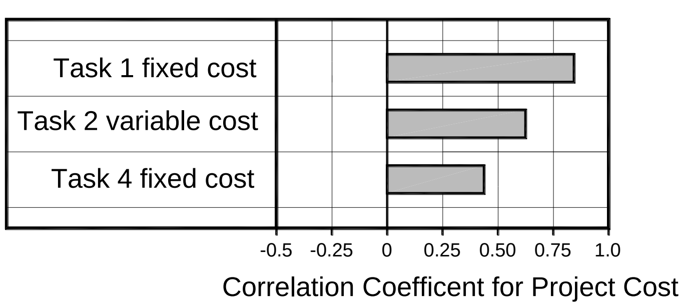

Figure 5. Sensitivity chart

Question: Can a task be crucial based on cost sensitivity and not crucial based on duration?

Answer: Yes. This is a fairly common phenomenon that depends on the combination of two factors: uncertainties in cost and duration associated with the task and the task's position in the project schedule.

Though you can use sensitivity analysis in the form of tornado and spider diagrams, the results of your analysis can be shown using a **sensitivity chart** (Figure 5). Input parameters are sorted in such a way that variables with higher correlation coefficients will be shown on the top of the chart (similar to tornado diagram). In our example, uncertainties in Task 1's fixed costs could have the largest affect on the project. Therefore, Task 1 should be the first candidate for analysis and risk mitigation.

Correlations between Tasks

Now, we have analyzed correlations between project inputs and outputs or the main project parameters. This allowed us to determine which parameters are the most important to the project. But we also need to analyze how correlations between different task parameters can affect the project schedule.

We discussed how to calculate the correlation coefficient between two data arrays. We can also perform an opposite operation where we **predefine correlation coefficients** (from –1 to 1) for certain variables and then **perform a Monte Carlo simulation**. The mathematics of the analysis is quite complex, but again, leave it to the software. The result of this analysis will illustrate what might happen to a project if a correlation existed between certain task parameters. There are also some ways to define correlations between different inputs.

For example, we have two activities, which involve the same subcontractor. If one activity is delayed because of the subcontractor, there is a 90% chance that another activity will be delayed as well. When you run the Monte Carlo simulation, assign a correlation coefficient of 0.9 for the duration of these tasks to calculate the effect of this correlation.

We strongly recommend you use correlation tools when you use Monte Carlo simulations to analyze projects because without this analysis it is impossible to tell how correlations could affect the project schedule.

Summary

- Many projects including the Maine Medicaid Claim System had problems because project managers were unable to identify what were the most important issues (tasks, risks). You must identify and address the important issues first.
- Different project parameters can be correlated with each other. These correlations can significantly affect a project schedule. Three sources of these correlations are: common drivers, common constraints, and common causes.
- Sensitivity analysis is a quantitative method, which helps to identify which project input parameters are the most important.

- You can make the probabilistic analysis of your project schedule much more accurate if you define correlations between different input variables.

References

Freedman, D., Pisani R., Purves R.1998. Statistics, 3rd edition, W. W. Norton & Company; New York

Holmes, A.,Maine’s medicaid mistakes, *CIO Magazine*, April 15, 2006

Nisbett, R.E. and Ross, L., Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall, 1980

Plous, S. 1993. The Psychology of Judgment and Decision Making, McGraw-Hill

Schuyler, J. 2001. Risk and Decision Analysis in Projects, 2nd Edition, Project Management Institute, Newton Square, PA