LEARNING TO APPLY STATISTICS USING A VIRTUAL ENVIRONMENT

Jock MacKay and Stefan Steiner
Business and Industrial Statistics Research Group, Department of Statistics and Actuarial Science,
University of Waterloo, Canada
rjmackay@uwaterloo.ca

We have designed a virtual environment, a manufacturing process that needs improvement, where we challenge students to come up with ideas for, design, execute and analyze a wide variety of studies. The motivation for performing these studies is easily understood. And it is all done on hugely abbreviated time scales with little cost. The purpose of this presentation is to demonstrate the environment. We also discuss some of our experiences that we hope apply to other virtual environments. One observation is that students need a framework to help to formulate, plan and execute statistical investigations A second is that students view the simulation as a game, enjoying the freedom and challenge of working in an environment where they have a well-defined goal and control over the entire project.

Learning to apply Statistics, like playing the piano, is difficult without practice. Understanding the theory and watching others perform are not sufficient. To provide the opportunity to apply a wide range of statistical plans and methods, we designed an easy-to-use virtual manufacturing process, called Watfactory. A process map is given in Figure 1. One part per minute is produced over three eight hour shifts per day, five days a week.

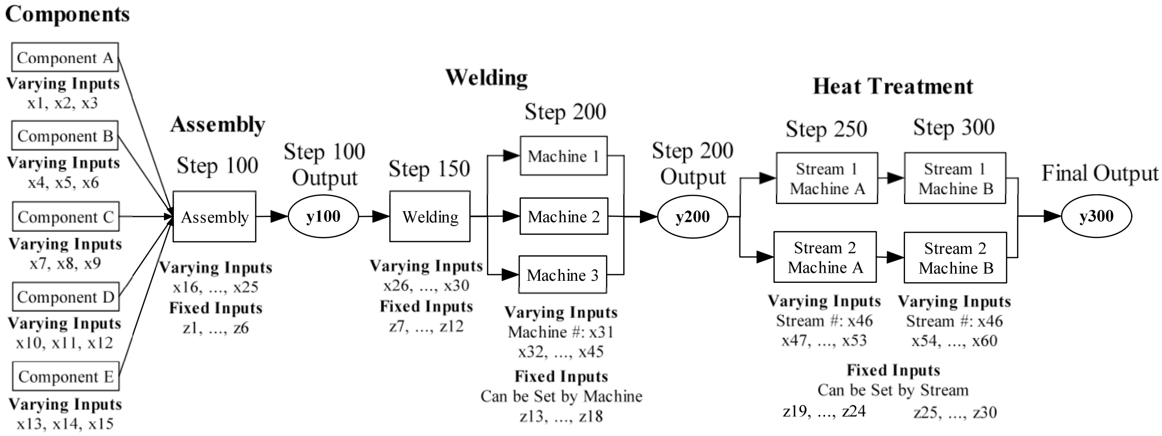

Figure 1. Watfactory Process map

The final output, denoted by y300, is a key process characteristic with given specification limits. Under the current operating conditions, some parts have values for y300 outside of these limits. The process starts with five components that are assembled at Step 100. It is possible to measure the unfinished output characteristic at this point and again after Step 200 where the components are welded together. Not surprisingly, the values change over the process steps. Within the process, there are varying inputs, denoted by x1,…, x60, that may change from one part to the next. The changes in one or more of these varying inputs cause the variation in the final output. These variable inputs can be measured for each part if we choose to do so. As well, there are fixed inputs, denoted by z1, …, z30 which the user can decide to change and which might then change the behaviour of y300.

Watfactory is available at http://www.stats.uwaterloo.ca/Faculty/VirtualProcess.shtml. More details about the virtual environment and some suggestions on how to proceed to reduce the output variation are available at this website and in the paper by Steiner and MacKay (2009). Needless to say, it involves a number of empirical studies.

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

We have used Watfactory as a teaching tool in a number of settings. In one course for senior undergraduate and graduate students, student teams are assigned the task of reducing variation in y300 as a semester-long project. In another first course in Statistics for engineering students, we use Watfactory to generate assignment questions and data. In both cases, each team or student has their own version of Watfactory with different underlying structure so the nature of the data generated from the process is different from team to team. The data produced by Watfactory are presented in a row-column format that can easily be cut and pasted into an analysis program such as Minitab or Excel.

Users can carry out a wide variety of empirical studies in Watfactory. For example, we might be interested in the relationship between y200 and y300 in the current process. We can determine values of these variates for as many parts as we like with a wide choice of sampling protocols. Or we may want to conduct an experimental study where we change the levels of one of the fixed inputs, say z1, to understand its effects on y100. Again, we can design this simple experiment in a wide variety of ways.

As can be seen from these examples, one of the strengths of Watfactory is that it forces students to make decisions about the plan of the study. In the first example, an observational study, students must decide on sample size and which parts should be included. Is a sample of 30 consecutive parts appropriate? Should we use a random sample from the next five days of production? Should we measure any other characteristics (e.g. y100 or one or more of the varying inputs x1, …, x60)? There is a cost to every investigation. Typically larger and more complicated studies cost more. For any planned investigation, Watfactory will calculate the cost before the study is executed. We give students a budget so that they can factor study costs into their decisions about the plan.

We believe that Watfactory helps students learn to apply Statistics in ways that standard textbook problems cannot. One observation is that students need a framework to help to formulate, plan and execute statistical investigations. We use Problem/Plan/Data/Analysis/Conclusions (PPDAC) as described by MacKay and Oldford (2000) and Steiner and MacKay (2005). A second observation is that students view the simulation as a game, enjoying the freedom and challenge of working in an environment where they have a well-defined goal, many important choices to make with incomplete information, and control over both the plan and analysis of one or more investigations.

REFERENCES

MacKay R. J., & Oldford R. W. (2000). Scientific Method, Statistical Method, and the Speed of Light, *Statistical Science*, *15*(3), 254-278.

Steiner S. H., & MacKay R. J. (2005). *Statistical Engineering: An Algorithm for Reducing Variation in Manufacturing Processes*. Milwaukee WI: ASQ Quality Press.

Steiner S. H., & MacKay, R. J. (2009). Teaching Process Improvement Using a Virtual Manufacturing Environment. *American Statistician*, *63*(4), 361-365.