Print

What Is Statistical Thinking?

In the opening video, Roger Hoerl described statistical thinking, emphasizing that all processes can produce information that enables improvement, learning occurs when you have an informative event and a perceptive observer, and basic tools are very important.

For this course, we define statistical thinking as a philosophy of learning, and taking action, based on three fundamental principles: All work occurs in a system of interconnected processes, variation exists in all processes, and understanding and reducing variation are keys to success. At the heart of statistical thinking is making decisions based on data.

This means that you need to be able to quantify and interpret the variation that you observe in your data, and determine which statistical methods to use, and when to use them. These are all integral questions, which serve as the foundation of this course.

Consider this scenario. You are studying the yield of a process. You want to understand whether the yield is improving over time. You collect yield data for two weeks and graph the data using a bar chart.

To better see the difference between the yields for the two weeks, you add a line. The endpoints of the line are the yields for the two weeks. What do you learn? Is the yield getting worse over time? Should you panic? What if you add data for a third week?

Now it looks like the yield is improving! Should you celebrate? What if you have more data? Suppose that you have yield data for 18 weeks.

You plot these data using a run (or line) chart. The yield is bouncing around from one week to the next. Most of the values are between 90% and 95%. You see that the yield was particularly good in week 12, and then it was low the next week.

What caused these results? You're excited about the spike in week 12. Can you improve the overall yield if you can identify and implement the change that caused this?

The yield in week 13 is disappointing. Should you reprimand the production team for the poor performance? This is where statistical thinking can help. When you think statistically, you might ask, "Do the observed values for weeks 12 and 13 indicate real changes in the process, or are they simply the result of random variation?"

You can't answer this question using a simple run chart. When you have time-ordered data, you can plot the data using a control chart instead.

This graph shows the overall mean and control limits. The control limits represent the range of random variation that you can expect from this process.

This is called common cause variation. It is variation that is inherent to the process. It turns out that the yield for week 12 is fairly typical, given the random variation in the process.

That is, there probably isn't anything uniquely good about week 12 causing the yield to be high, or anything bad in week 13 causing the yield to be low.

In any given week, you can expect the yield to be between 86% and 98%, just by random chance. This gets at the heart of statistical thinking. To improve any process, you need to understand the process, you need to understand the variation in the process, and you need to understand the causes of this variation.

Throughout this module, you learn a number of tools for understanding the process. You also learn a variety of tools for identifying potential sources of variation within the process.

In the remaining modules, you learn tools for exploring and characterizing process variation, and you learn tools for understanding and reducing this variation.

Close