**NCC_BSL_DavisBalestracci_2_10292015_v**

Hello, and welcome to Part 2 of Plot the Dots. Today's topic is Control Charts, also known as Process Behavior Charts.

The term "control" is so misunderstood and it actually isn't a very good term, so that's why we can call them Process Behavior Charts. Control is in the concept of statistical control, which as you will see, just means the process is stable. So it assesses the stability of the process, which will make a lot more sense when I'm through with this presentation.

Last lesson was about Run Charts. I call that filter number one for a time sequence of data. All it answers is did this process have any significant shifts during this time. That's it. But Control Chart now, and once we kind of see, do we have boiling water, ice water, take the average or correct averages, Control Chart now becomes filter number two. It answers the question well how much variation in my process is natural, and are any individual observations possible outliers. So those are two very important questions as well and, as I hope I will show you, the math required to convert a Run Chart to a Control Chart is elegantly simple. As a result you will have a chart where it will motivate conversations you've never had before, and I'm hoping more productive conversations.

So I'm hoping that as a result of this session you will be able to see how easy it is to calculate how much difference between two consecutive points is too much; you know, as you look at the performance this month versus last month, or the performance this week versus last week. Of course they're not the same number, but how much of a difference is too much? Or where is the threshold between common cause and cause? The other thing you can calculate is the expected range of your process' output so you don't overreact to common cause. In other words, what you will calculate is what your process is perfectly designed to get, even if you don't like it.

So then my other objectives are: can we quickly facilitate agreement on a strategy to improve a process even if the variation to you seems out of control. Part of my job with this video is to get you very comfortable with the amount of variation, and hopefully show you that it really doesn't matter in terms of improving the process. Let's stop the tendency to jump right to a process redesign or a known solution in reaction to common cause. There are some trainers, many of them as a matter of fact, who are teaching you the wrong thing in that they say if a process is common cause you have to accept it; and the only way to do something about it is a process redesign. That is not true at all, and one of the upcoming lessons will show you common cause strategies. I'm going to demonstrate to you today an introduction to one of the more important ones, which is how do you take a vague situation that is in control and zero in on the important part of it. That gets to my last objective: see the need to use stratification to focus a vital opportunity within a vague problem.

Okay, quick review. Last lesson was about a Run Chart, which starts off as our bread and butter tool. Always start by plotting your data in a Run Chart. It's a time-ordered plot with a median drawn in as the reference.

And I taught you two special cause tests. One is the Trend Test, which is six successive increases, or six successive decreases. That will be rare, and I think in the example I showed you all it indicates is a process in transition. Another way that this test becomes very helpful is if you're in a meeting with tables of numbers, and everybody is shooting from the hip. If that's all you're going to do then you have to use six. You will see once we have the common cause and do the Control Chart, we don't have to be as conservative.

Then the other rule is testing for a shift where you have eight data points in a row, either all above the median or all below the median. This is indicative of a shift during the time in which you studied the process. I just want to reinforce what I said last lesson was that I would prefer when you do improvement work to collect more frequent samples over time.

Okay, do you remember this? This is real data on bacteraemias in a hospital, and if all people are doing is shooting from the hip, the tendency is to say well look right here. We trended down and things got better. And then we went back up and all of a sudden everybody has to drop everything and find out why we went up. Which as I showed you with the Run Chart, over these 19 quarters nothing had changed despite the fact that they did 150 Root Cause Analyses. Every month they had a quarterly meeting – every three months. Every quarter they had a meeting where they pulled the charts of the bacteraemias and said what should we have done differently, which is a special cause strategy; and this plot shows it to be common cause and it also shows that all that work didn't really yield any change.

Then I showed you that this process is perfectly designed in any one quarter, even though we're averaging eight. You're not going to get exactly eight every quarter, but you will get a number between zero and twenty; and that one quarter can differ from its previous quarter by as many as, yes, fifteen. You're all saying well, that's out of control, we don't like that. Well I don't like it either, but I'm going to get you used to the fact that despite how much variation you see or how out of control it seems, it could very well be in control and we can deal with it.

Okay. The math is very simple, as I hope you'll agree. Here are the 19 data points in their time order, and that's important; time order. I determined what are called the moving ranges, which simply means I take each data point and subtract its predecessor; and it's the absolute value. In other words, its variation; I don't care whether it's positive or negative.

It also turns out this number has very nice statistical properties. The one time where – how do I put it? Easiness to calculate and statistics go hand in hand, and it is useful. So the difference between seven and ten is three; the difference between three and seven is four; and the next one would be the difference between ten and ten, which would be zero. Yes, a moving range can be zero. And we go all the way down the numbers like this until we get to the 18th where the difference between five and twelve is seven. So 19 data points produce 18 moving ranges because the first data point doesn't have a predecessor. So then we sort those 18 numbers to find the median; yes, the median. So in this case with 18 data points it's going to be the average of the 9th and 10th.

So there's the data in its time order, there are the moving ranges, and here they are sorted from smallest to largest.  And since there are 18, if you start counting from either end they kind of meet in the middle here; I have to average the 9th and 10th, both of which happen to be 4, so the median moving range is 4.

So now that is the key.  The median moving range is the number from which all the estimates of variation come.  That is a very important number.  And it doesn't matter whether you  have 10 data points, 20, 30, 50, 100, 500, 1000; if you do what I just did, and that is take the successive differences, absolute value of the successive differences, and take the median, all the information about the variation is contained in that median moving range.

So taking that number; as you see I have it here.  This 3.865 is a constant from theory.  It will never change and you use it with the median moving range.  Now I can hear some of you saying well Davis, can't I take the average of the moving ranges?  Yes you can.  Some of you may have even been taught that, and that's just as valid except you have to use different constants.  Now here's why I teach the median moving range.  As you will see, it doesn't make a whole lot of difference.  And the reason I teach it is you won't always have your computer; this is easy to do by hand in a boring meeting.  Outside of that meeting with your computer there's plenty of good software that will do this.  So the reason I teach the median is to appeal to your intuition and that you can do it by hand in a boring meeting.

So this constant is from theory to be used with the median moving range.  And if I multiply the median moving range times 3.865 you get approximately 15; which once again, that is how much two consecutive – and let me emphasize – consecutive months can differ; just due two common cause.  So in other words I'm looking for a big difference here.  Here it looks like it went from three to ten.  Now some people might get upset about that.  Well it's less than 15; I really can't call it a special cause.

Then the other thing we calculate is because the Run Chart showed no special causes I can calculate the average, and it's about eight.  Now I need to say well what is the – what I like to call the dead band of common cause around that average.  And the way we get that is, once again, the median moving range times in this case 3.14, which is a number from statistical theory.  And it has nothing to do with pi, but if you multiply that times the median moving range that gives you the dead band of common cause on either side of the average.  So in this case it would be approximate; approximately 8 +/- 12.

Now here's something that's very important if you're going to present charts like this to doctors.  You will see that 8 minus 12 is a negative number.  Don't ever put a chart with a negative number in front of docs, if the negative number doesn't make sense.  I mean you can't have a negative number of bacteraemias.  Set it equal to zero or they will hoot you out of the room.  So in this case then, in any one quarter averaging eight, given this pattern of variation, we will observe between zero and twenty bacteraemias.  Hard to believe; you might not like it.  But that's what you are perfectly designed to get with this process.

But here's the good news.  Despite that wide variation I'm going to introduce you to the first common cause strategy, which is called Stratification.  What this means is the same process produced it.  In essence every infection was produced by the same process.  So what that means is here's the good news. So the bad news is you're perfectly designed to get this horrible result; but the good news is it's stable.  So what does that mean?  I can put all 150 together.

Now here's a case where I say I'm the statistician; I know nothing.  You're the healthcare people; you know too much.  That makes us a good team.  I've done my job.  I've kept you out of the data swamp.  What I want you to put your brains to work on is what are some ways we could slice and dice these 150 bacteraemias?

Joseph Teran, one of the quality giants of the 20th century, firmly believed in what he named the Pareto Principle in that the cause of a problem is never vague.  That usually in a situation you can find 20% of the process causing 80% of the problem.  So in other words, if you were to slice and dice these you could find the 20% of reasons that account for 80% of these, and that allows you to focus because obviously focusing on each one individually didn't work.

And it's so easy.  Once you learn the quality tools we can say well, this is common cause; oh my goodness, we've got to do something, we've got to redesign.  And you can go into a whole process of saying what causes blood infections and use that cause and effect or Ishikawa Diagram to categorize all the causes.  You've probably all done it, I've done it, where we go into a situation, say we need to redesign, and we come up with what I call the Ishikawa Diagram from Hell that takes up a whole wall.  Whereas what we can do with stratification is what if we find out the problem is here and then start brainstorming?  Because sometimes what happens too is we brainstorm all these reasons and we collect data on all these reasons, and then quickly find that the problem was here.

Now what has happened?  You've made all these people mad and you've lost some credibility for the next time you ask them to collect data.  They have the memory of hey, I did that once and they never used it.  So this quote by Brian Joiner is very telling: "You've got a big vague problem, you're going to have big vague solutions that aren't going to work very well.  You've got to focus, focus, focus so you get the effective solutions."

So yes; we try to use stratification to say what is the 20% of the process causing 80% of the problem?  And as I said, the stability is the bad news but it's also the good news because we can aggregate any stable period.  And I think rather than look at the quarterly four infections, five infections, three infections, two infections; if you look at 150 you're going to get the bigger picture.  So once again I want to stop the misconception that if it's common cause I need to redesign the process.  There's going to be a lot more about this in the lesson on common cause strategy.  So all common cause means is I can't look at infections individually and I can't treat the data points individually.  That's all it means.  But I can group things.

So here's an example, and this is real data; and it might be a meeting you typically go to or have even experienced where you're looking at a year over year comparison of say falls. And I know falls are very important to you folks. So let's say this year through November we have 64, but the same period last year we had 51. So you might say we're up 25%. Now 25% to you seems like a big number, and we're going to treat it like a special cause because it's too big, we don't like it, we've got to do something about it. But we don't know if it's common cause or special cause, yet. Not only that, they finally got a zero; wonderful. But then what happened after the zero? They went up to nine. Horrible. What's wrong? So now fear starts creeping in; is the reporting process going to be jeopardized? So let's plot the dots.

Now I notice what people do a lot too is you'll have this year over year plot, and they always plot the last 12 months with the excuse being we need to see the trend. Well you know how I feel about the word "trend". But okay, here are the last 12 months; there's a trend line put in, and you know how I feel about trend lines. And the CEO talks about how during these first four months we were above average. So now he insisted on root cause analyses for every fall. Well look what happened; it dropped and that's good work.

But look what happened after that. Look what happened; do you see that trend? It isn't a trend. But a lot of people would call that a trend, get worried, and say we need to start the root cause analyses again. Then you look and you say good, it worked but it's still at an unacceptable level, but at least we're now below level, whatever that means.

So he gave you the funds to have a safety fair so you could have key chains made up that say "Fall in love with Quality". Give everybody a key chain, give them education, show little videos, and oh my goodness, look what happened; it went to zero. Good work! In fact, buy everybody pizza. But then, good gracious, look what happened. You know what happens when you reward people, you know.

So does this seem like déjà vu for a lot of you? I'll bet it does. Now here's the beauty the next time you're in a meeting like this. First of all we don't do this: we don't panic because we saw this and said all right, time for the cause and effect diagram. Remember, vague solutions to vague problems get you vague results. Let's understand the process.

So do you realize by having this data, the year over year for 11 months, coupled with the last 12 months, you can reconstruct the last 23 months of performance? So there it is; the 23 months of performance. I sort it to find the median; so 23 numbers, which means it's going to be the 12th in the sorted sequence, 11 smaller, 11 bigger. Well let's plot the dots. So over these 23 months do you see? Using our rules you see six successive increases, six successive decreases. Do you see any clumps of eight above the median, below the median? If you had improved over this time you might see six successive decreases; or you might see a clump eight above the median early in the data, and/or a clump of eight below the median later in the data. You see none of that. And all those root cause analyses, what has the net effect been? No change. So the last 23 months your process has been behaving like it is perfectly designed to behave.

So now, since I don't have boiling water and ice water, I can calculate the average of 5.4. Now we want to know in any one month what can I expect? And we all panicked as it went from zero to nine – is that a special cause?

So there are the 23 numbers in their time order, there are the moving ranges, which I now sort. So 23 numbers, 22 moving ranges, which means I have to average the 11th and 12th, both of which happen to be 3.

So first of all, since the median moving range is 3, I multiply it by the 3.865 and it's approximately 11. So in other words, two months can differ; consecutive months can differ by as many as 11 falls. This jump from zero to nine is common cause; there was no need to overreact, even though you felt it was big. Some people might feel ten is big, but this says no; you have to wait until it's greater than 11.

Okay. So now let's see what we're designed to get in any one month. So once again the median moving range is 3; so to get the common cause there's our average, and I take the median moving range times the 3.14. And if I do that it turns out to be a range, and we can't have negative falls, zero to 15; believe it or not. There have been no special causes over this two-week period, but as you remember from that graph, there's been a lot of effort going on.

So I hope what you've learned from this lesson is if you have a vague problem but you get a plot of data over time and a data sanity "aha", you'll be on your way to solving the right problem. Once again I teach you to use the median moving range because that's easy to do by hand in a boring meeting. There are plenty of good software packages. And the one thing I want to emphasize is do not – do not use the standard deviation of all the data. You will never use the standard deviation calculation as you were taught in your basic statistics course again, ever. You have to get it by the moving ranges.

So to sum up: after you do the run chart, a control chart analysis helps to understand that natural inherent variation that's present in your situation that you may or may not like. But regardless – see – regardless, it enables you to act appropriately in response to avoid the human tendency to treat all variations as special causes. Even when it seems out of control it may not be. But the good news is we can deal with it if it's common cause. And the best thing to do if you can is to somehow aggregate the data and come up with ways to slice and dice it to see if you can find the 20% of the process causing 80% of the problem.

Like I say, I'm the statistician, I know nothing; you're the healthcare people, you know too much. That makes us a good team. You're smart people, and that's the problem. Because if in that falls example I said well, why did we go from three to ten? You're smart people, you go, you find a reason. Except each one of you would find a different reason. Whereas I say listen; we've got X number of falls here, I think it was 191 falls. What are some ways we can slice and dice those to get some insight as to where the biggest problem in your process is? You're going to come up with wonderful theories that I couldn't even think of. That's where I want your brains to go. I want to focus your wonderful brainpower.

So thank you.  Here's how to contact me.  You can tell I like what I do a little bit; I welcome contact.