MITOCW | Lec 8 | MIT 2.830J Control of Manufacturing Processes, S08

DUANE BONING: So let me begin by just giving you a little bit of a preview of what's coming up over the next couple of weeks in terms of the calendar. The problem set that's due today you'll notice is perhaps a little bit long. But the good news is the next problem set you'll look at and go, is that all? Because it will be a very short. Part of the reason is the problem set that goes out today will be due a week from today, next Tuesday. But a week from Thursday will be our first quiz. And so the upcoming problem set is meant to give you some time to go back and re-integrate and re-digest the material that we've covered so far.

The material that is on this problem set that's going out, it's going to be a couple more, a few more problems still related to control charting, some of the things we'll talk about today. And that will be covered on the quiz. So the quiz will be statistical basics up through statistical process control and control charting. That'll be a week from Thursday. This Thursday I'll be giving a lecture that shifts a little bit to yield modeling. And that will not be on either this problem set or this quiz.

So that's the plan. We'll have more details on the quiz and I think even some example quizzes that you can look at from the past. But other quick warnings or previews on the quiz itself is the plan is to make that closed book. But you will be able to bring a page of notes in. So you can have a single page of cheat sheet or notes, formulas, examples, whatever you can fit for quiz one on the one side of one page. And also bring a calculator. We will provide for you any statistics tables that might be needed, typically out of the back of Montgomery or [INAUDIBLE]. So we'll have those if you need them. Yes?

AUDIENCE: So [INAUDIBLE]

DUANE BONING: We're not giving you those. So those should go on your one page of notes that you can bring in. Or you can just remember them if you're really good.

AUDIENCE: But you'll give us like a Z table?

DUANE BONING: We'll give you a Z table. We'll give you, if you need it, t, Chi squared. We'll give you those tables.

AUDIENCE: Is there a gamma table?

DUANE BONING: I'm sure there is a gamma table. I'm not sure whether or not it's in the--

AUDIENCE: [INAUDIBLE] gamma of something--

DUANE BONING: It arises in--

AUDIENCE: [INAUDIBLE]

DUANE BONING: And that might be a clue. If you don't have a gamma table, don't try to use the gamma.

AUDIENCE: [INAUDIBLE]

Can you put a sample exam so that we know--

DUANE BONING: Yeah. We'll try to dig up an example exam so you can take a look. OK. So any other questions on the quiz? Any questions from Singapore on the upcoming quiz? You guys look so eager for it.

OK. So today what I want to do is pick up a little bit where Dave Hardt left off. So the plan here today will be a very brief review on control charting. The basic idea I want to be able to hit, now that you've seen what he described and what you've been doing hopefully on the problem set that's due at the end of the day, is reinforce the relationship here between hypothesis tests and control chart.

And so I'm going to just remind us a little bit about hypothesis tests and these different types of errors. The probability of false alarm, probability of a misalarm, missed alarm or beta, and how the sample size plays with that. Because when you design control charts, alpha, recall, basically sets your upper control limits and lower control limits. Beta falls out, beta being related to essentially the probability, again, of missing an alarm when the process has, in fact, shifted.

And how do you change or affect beta? By modifying the sample size. So I want to talk a little bit about those. I think you might have touched on them briefly. But I want to touch on those and also remind us of the definition or give you the definition of average run length. And I think if you've gotten to that point in the problem set, you've already read about that as well.

So let me start, again, with the basic hypothesis test. Because remember, all that a control chart is is a running hypothesis test where every new sample comes in, you're doing a hypothesis or you're asking the test on the hypothesis, has the process shifted or is it still my nominal distribution?

So the basic situation with the hypothesis test, if you recall, what we have is our assumed distribution. I'm going to assume here that it's a Gaussian. We'll call that h0. It's got some mu, some true process mean. And what we said is we pick control limits, or in the hypothesis test case acceptance regions. But I'm going to use the control chart terminology, since that's what we're dealing with here. We pick where these limits are based on a probability of false alarm.

Seems a little weird, right? But basically what we do is we assign if alpha is our false alarm rate, we put alpha over 2 in each of the two tails. Now, that's based on this key idea that if we're in control, most of the time the data will be in here. Every now and then, just by random chance alone, even when nothing has changed, we're still in h0, I do get an unusual point out in the tails.

And if we do 3 sigma, we saw that's about 3 in 1,000 points will fall way outside in the tails. So when we get a point out in the tail, we say something might be unusual going on. Either something has changed in the process or I got unlucky and I just pulled a sample from the tails. So in fact, there is, in a process that's always in control, there is this false alarm rate.

Now, another way of talking about this false alarm rate is also the inverse of the false alarm rate, which is how many runs does it take on average before you get a false alarm? And this is referred to as the average run length. So here sliding back to the slides. Can we put the PowerPoint slides on both screens? Great. Thank you very much.

So looking back here, this is our nominal probability. Again, 3 out of 1,000. But now if I turn it around and ask the question, how often or what is the average number of runs that we have to go? Well, it's just 1 over that probability on any one run of having the point fall in the tails. So that's the definition of our average run length. And actually, a notation that I like in Montgomery is we'll define this as average run length sub-zero. What that does is remind me that that's the average run length associated with the h0, the default probability distribution.

So that's the alpha part. Where does beta come in? Beta comes in when we actually have a shift. So the whole idea is now imagine that I've got a shift in the distribution. I've had a shift from mu to mu plus some delta. I'm going to label that the alternative hypothesis. So the idea is for a given delta, I'm now actually drawing parts from a different distribution, my h1 distribution. I've already defined the acceptance region based on my acceptable average run length or my acceptable false alarm rate.

By the way, there's nothing magic about plus minus 3 sigma. If you don't want to respond once out of three times out of every 1,000 where about once every 333 runs, you don't want to respond to a false alarm rate, it doesn't cost you that much to have a false alarm or it costs a lot to have a false alarm, but it doesn't cost you that much to have a bad part, then maybe you only want to respond 1/10 as much. In which case, you would change your control limits, move them out a little bit so that you set it based on the average-- set based on the average run length or the probability that you can accept out of the tail.

So that's set. But now we said there is also the type two error or the beta error, which is the probability of a missed alarm. And that's set by those control limits. So now the probability that I would have thought it was a good part even though an alarm occurred is just the other side of that same control limit but now drawn from the h1 distribution. I really have to shift, but the point fell in this part of the tail of the h1 distribution. And I declared it good. I said there's no alarm.

And we saw last time basically, or a couple of classes ago, this is also it's just the tail there in that distribution. It's UCL minus delta over sigma over square root of n. By the way, this is just this part of the tail. To be really careful, I should also subtract off this tail, this part right there, because I got something way out there, I would call it a failed part. And so my beta is slightly smaller than this. But as I've drawn it here, I'm really only worried about this part of the tail. But there is the other part of a lower control limit plus delta minus. Excuse me. It's still a lower control limit still minus delta over single route in.

OK. Great. There is also another definition of average run length associated with this. And the question you would pose here is what is the average run length to detect shift when it actually occurred? It's a little tricky here, because there's actually two steps. You can imagine this would be the probability of not missing an alarm and then inverting it, just like we did here.

So this little tail here is the probability of missing alarm. There is also a definition referred to as the power of a test, which is 1 minus beta. And that's basically the probability of catching an alarm on any one run. And that's just, it's everything to the right of this. And now if I draw from h1 but not in the region where I would have declared it h0, that's 1 minus beta. An average run length sub 1 is defined as 1 over 1 minus beta. So this is the average number of runs it takes to detect the shift assuming it actually occurred.

OK? Everybody happy with this so far? Is it starting to fit together? Now how do I change-- well, why would I change alpha? I change alpha based on a decision outside on acceptance of my acceptability of making a false alarm. How do I change beta? What if I say it's taking me too long to detect a shift? I want it to take less time.

Change n. So what changes when I change n? Does alpha change when I change n? No. Alpha is set a priori based on my false alarm rate. When I change beta-- or when I change n, I do change beta. Do I change the control limits? I do change the control limits. I change the control limits because my sample size changed and I move them to maintain the same alpha. So let me draw the picture. Let's see. If I draw down here, can people in Singapore still see this? OK, great.

So if I increase n, what's going to happen is I have a sample size that tightens by my sampling distribution for x bar. Remember, h0 is really an x bar based on. And so that distribution will have the same mean, but it will get tighter. So now I have an h0. But I'm still setting my control limits based on an acceptable alpha.

But the control limits will get tighter, because I've got a tighter distribution. I still have the same alpha in the tail. And what's cool here is I'm still detecting the same mean shift. Trying to draw down here. Boom, boom, boom. So now I've got a tighter distribution here as well. Here's my h1. I've got a tighter distribution on h1 and I've still got just a little bit of tiny stuff over here on the tail. And that's now my beta, and it's much smaller if I've increased n.

So I can basically pick my sample size to give me the acceptable false-- or the acceptable missed alarm rate or the acceptable average run length to detect the shift. And there are tables in both Montgomery and [INAUDIBLE] that are referred to as operating curves or OC's that basically give you the relationship between sample size, size of shift, and the beta or average run length one that you can accept. You could go in and calculate them.

Turns out this is actually a really nasty calculation to do, because there's no close-- if you need to find the n to give you a certain beta, you have to solve it iteratively. There's no closed form for that. Q sum are the cumulative probability that's associated with beta. So there's no formula for it. And these tables, basically, or charts have basically calculated that for you. You can set up Excel or whatever to do that, to solve it iteratively if you want. Yeah?

AUDIENCE: [INAUDIBLE]

DUANE BONING: So the question was, we're changing alpha or changing sample size. The distribution gets tighter. Why doesn't--

AUDIENCE: Yeah, why doesn't alpha?

DUANE BONING: Why doesn't alpha change? If we kept the same lower control limit and upper control limit and I tightened this. So if I kept these the same but the distribution tightened, then alpha would change. What's interesting about classic control chart design is the upper control limit and lower control limit are based on the statistics. So they are based on alpha. We nailed down alpha first and then set the upper and lower control limits.

And there's a really important thing that we'll get to, actually, in just a second about process capability where then we start talking about upper and lower spec limits, which is related to the notion of where the parts property really needs to be for it to function correctly. And those are very different than control limits but often confused. And so nominally, your spec limits wouldn't actually change if you're changing sample size. The part still needs to function correctly within a certain band.

So in that case, you change sample size. Your spec limits would not. But what people often do is erroneously set the control limits to equal the spec limits. And that's wrong. What you want to do is set the control limits based purely on the statistics based on your false alarm rate. Chance of observing by chance or by random probability, an unusual event.

AUDIENCE: [INAUDIBLE]

For example, in the problem set, there was a question [INAUDIBLE].

DUANE BONING: I guess I don't want to say too much about the problem set right now. But you can often relate things back to the underlying probabilities. And what you should usually do is relate them back to the underlying probabilities and set up the hypothesis test based on shifts, questions about there can be two sided shifts, one sided shifts.

Both things could shift by a certain amount, and it might be just worried about what's happening in the tail. There's other hypotheses. These are all for changes in the mean. If I'm looking at some other distribution, like changes in the variance, I could set up a similar hypothesis test and perhaps other probability distributions apply.

Turns out with the S chart, we're really tracking the mean of standard deviation or the mean of the sample standard deviation. And in those cases, we typically still are assuming that the distribution of standard deviation for large samples is Gaussian within correction factors to make it come close in setting the control limits. Things like C4's and those sorts of things.

But going back, it's not always exactly the same set up. But you can almost always reason it out by going back and setting up what is my normal situation, what is the thing I'm trying to detect, and then look at the probabilities associated with those cases.

AUDIENCE: I have a question, Professor.

DUANE BONING: Yes, question.

AUDIENCE: If I give you a control chart, you can't really tell whether it's a false alarm or it's a true alarm, can you?

DUANE BONING: That's correct.

AUDIENCE: Like how do you use a control chart then? Because if I gave a control chart, it could be with alpha probability of a false alarm or with 1 minus beta probability of a true alarm. [INAUDIBLE]

DUANE BONING: This is a fundamental important point about how control charting works, which goes back to the idea that if nothing changes, you do have false alarms. And if something changes, you've got a real alarm. All you've got on the control chart is a point that's lying outside of your normal bounds. All by itself from the control chart, you can't tell. You are exactly right. So in SPC methodology, what you are supposed to do any time you see an alarm is go and look at the equipment, investigate the problem, and decide was this a false alarm? In other words, is the process still operating normally? Or has something important changed?

Now, it turns out that in an awful lot of factories, people have adapted to this notion of false alarm in a very qualitative way, which is they see one alarm point. They look at it. They don't investigate the process. They ignore it. And then they wait. They do the next point. If it's also an alarm, then they say, ah, I think something may really be wrong and then they investigate. What are they doing there? How could you change your control chart to actually do what they are doing?

What essentially they're doing is saying, I don't believe the evidence. The 3 in 1,000 chance of seeing one point outside of a control limit is not a strong enough evidence for me. What they're really doing is looking for a smaller probability of two points in a row being outside of the control chart as much stronger evidence. So in this case, roughly it's twice as infrequent. Or not twice, but the product.

The probability of seeing both of those points outside of the control limit, and again, not alarming on the first one is the square of that. So that's a very infrequent thing. Then they say I think there's really something wrong. Now, you could actually have adapted your control chart either by having it only signal alarms once you've got two alarm points or changing the control limits so that you basically move the control limits out even further. So you've got to really have an alpha over 2 that's that small. So maybe you go to 4, plus or minus 4 sigma or 4.5 sigma, something like that, in order to have that much evidence.

AUDIENCE: What steps are efficient to find [INAUDIBLE]?

DUANE BONING: I'm sorry? What steps?

AUDIENCE: Yeah. There is a huge machine, and everything can be somewhat changed. I don't know where I should start.

DUANE BONING: Well, I think the steps that you use to actually investigate depend completely on the process, but there's a few rules of thumb. Many of you have probably had actual process experience on different tools, different manufacturing lines. What are some things that you would do? I'll tell you the number one thing I would do.

AUDIENCE: Check the screen?

DUANE BONING: Look at the screen. Pull the power plug! Plug it back in! No. Yes?

AUDIENCE: Check the measurement?

DUANE BONING: Yes. That's exactly what I would do. I would always doubt my measurement equipment first. They also have failures. Rather than pull the process, turn the process off and panic with the actual equipment, I'd want to verify the measurement first. So I take it off, do an independent measurement of those parts, and try to understand it. So let's say that indeed I see, yes, that part looks like it's outside of the control limits. What's another thing it might do? Do you look just at that one single point? What other data do you have?

AUDIENCE: You can look at the [INAUDIBLE].

DUANE BONING: Exactly. So now you can look back at the control chart data that has been plotted either manually in old systems or in the computer system. Track back historically and start to look, is there some other evidence that wasn't triggering an alarm but that might give you a sense? And a telltale one might be, in fact, a slow drift in the points.

Now, some people might set up the WECO rules, the Western Electric rules, to also watch for drifts. But that's pretty rare, frankly. People tend to just use the upper and lower control limits in a single point outside of them. So you might look back at the rest of the data and try to see is there any additional evidence.

AUDIENCE: The trends can be incorporated into [INAUDIBLE].

DUANE BONING: Exactly. So we could also have-- the point was trends could be incorporated in the control chart as well, either with this and additional rules or, in fact, other control charting approaches we'll talk about later that are designed to track or detect very subtle trends. But again, the point would be often you are trying to debug the process and look back. So you are a good pattern detector, and you might detect a pattern.

Failing that, I think then you start looking around to see if maybe there's not a systematic trend. What you're hypothesizing, in fact, is that there was a shift, not a trend. And you start talking to people, looking at the equipment, talking with the operators, and ask, did something unusual happen? The whole idea is you've detected an unusual event. Might that be correlated with something else in the operation of the machine changing? So you start investigating and looking for things that might explain a shift.

So you go to talk to the operator, and it's a new operator. You've never seen that operator before. Say when did you start? They say, an hour ago. Aha. And then it gets progressively more complicated. You're going to have to dive into more details on either the process, the feedstock, all of those sources of variability that we talked about earlier. Good questions.

OK. So I've sort of flown by these. These are actually slides, again, that were pulled out of last Thursday's lecture but just illustrate this notion of and give some example numbers of exactly what we're talking about here with average run length. The example plotted here is the same chart but drawn a little bit differently for kind of smaller shifts. Here we have a delta u that moves the probability under the assumption of a shift. And this pe down here is basically 1 minus beta. So this is the probability of detecting the true shift under the control limits again that were set based on h0.

And so here is an example where, again, the false alarm rate was 3 out of 1,000 with this small one sigma shift. Now your tail probability there only becomes 24 out of 1,000 or an average run length, if I invert that, of about 42 samples are required before you can detect that small one sigma shift. Now, one sigma shift doesn't seem all that small to me.

If that's too long to detect a shift, then what do you do? Increase sample size. That's exactly right. And here is simply the definition. What's going on there if we increase the sample size is the standard deviation, again, of our charts, our x bar chart. h0 and h1 are shrinking. That moves our three sigma limits closer together so that you get smaller probabilities. Yeah?

AUDIENCE: Even if we increase the n, still we have to get more samples [INAUDIBLE].

DUANE BONING: Yes. So the question was to increase, then, you got to take more samples. So there is a little bit of this two level hierarchy in sampling assumed here when I'm talking about this. We'll also talk about ways to deal with more continuously sampled, even driving in smaller, not larger, so that you can improve detectability. But the two level hierarchy that I'm talking about here is basically this notion that in many lines, you've got lots and lots of parts coming off very quickly.

So in time, you've got parts coming very rapidly. And the basic picture here is you're grabbing a sample of size n at some point in time. Maybe it's a four point sample. And then much later, I'm grabbing another four. So I am sampling from a what's pictured as a much larger group. And so the idea here is I actually have more than one degree of freedom. We've been talking about I can change just the sample size, but I might still in time only be sampling--

I could have actually the same time period. So I'm not actually changing if I just modify my sample size here. I might actually not change the frequency at which I draw my size and sample. I might do this once an hour. And once every hour, I draw a sample of size four. So there's an additional degree of freedom that's not in here, which is I could also shorten my sample time, if you will. Because then it takes fewer minutes of clock time to get to my 42 runs that I sample. So I haven't really talked about sort of the trade off between these two. But this is the assumed model.

And you could also do things, and I think what was kind of implicit in your question, is I could continuously sample everything. And these days of automated inspection equipment on the equipment, on the manufacturing equipment, you're very often sampling and measuring virtually every part. In which case, it's sort of like every part gets measured, in which case I have a very interesting trade off between n, which is exactly the point you were making. If I increase n, then it takes me longer to get more runs that would form each point on the control chart. So there is a little bit of a trade off there. Is that clear? Almost?

AUDIENCE: Almost. Because why we don't simple increase-- because [INAUDIBLE]

So why we just don't increase the coefficient? So it's more sensitive to any change. Why we should increase the x?

DUANE BONING: Well, let's see. If you increase, go from 3 sigma to 5 sigma, you increase the control limit and you make it less sensitive. You have to have even more extreme low points. So in that case, you decrease the sensitivity of the chart to false alarms. So in that case, you would want to-- so alpha and beta are different. When you play with them a little bit on the problem set, it'll reinforce, I think, a little bit of this. So again, this is basically the example we just talked about.

OK. So I want to shift to kind of this point that I was making earlier about the difference between control limits and spec limits. But I'm going to get to it through this notion of process capability. And CP and CPK, if some of you have run into that or played with that yet on the problem set.

The point here is that once we've got a distribution like this, based on historical data that we use to build our control chart, we basically have a model of our process. It's a very simple statistical model that says it has a particular mean and it has a particular variance in the normal distribution case. And an important thing that we can do with that model is answer a couple of key questions that really come up a lot. How good is the process? Is the mean where we want it? Is the variance where we want it?

And in particular, how good is it with respect to our requirements for manufacturing a particular part or product? This is the notion of process capability. And what this also pulls in, then, for the first time, unlike all of this, which is just probabilities on a normally operating or a not normally operating process, process capability pulls in for the first time the notion of design specifications. Different than control limits.

So the basic way to characterize your process is you have to assure yourself that the process is indeed in control. It's operating normally. No common cause events. When that's the case for a normally distributed process, it's characterized entirely by mean and variance or on our control chart by x bar and s. So indeed, if the control limits have been set correctly in relation to known sample sizes and known probabilities of error and so on, then x bar and s tells you everything about the underlying process mean and variance.

And now we can start to use those to look at where the process sits and those statistics sit with respect to tolerances or quality loss functions. Quality loss functions may be a little less familiar to you. Everybody has an intuitive idea of tolerance. So let's talk about that first. Basic idea of a tolerance on, say, a part measurement is that there is an upper and lower limit. So let's say it's some characteristic dimension. We have a nominal dimension or a target dimension. I'll label that x star.

And then we have an upper spec limit and a lower spec limit that we simply say, as long as the part dimension is within that range, we're going to call it a good part. And it meets maybe the design requirements. It meets our consumer's requirements, what have you. And what we're saying is in essence, anything, anywhere within those spec limits is equally good and equally shippable. So there's no notion of better or worse. It's simply on off decision, it's acceptable or not. So that's a spec limit or not.

Now, the idea of a quality loss function is a beautiful idea that actually conforms a little bit better to what we really care about. And one could think of a quality loss function as this penalty function for any deviation away from our target x star. Now graphically, a typical quality loss function that is used is a quadratic cost for these deviations. Where the further I get away from x bar, I penalize more and more. Not just even in a linear fashion but in a square fashion. And then I have some normalized or some constant out front.

And what this basically says is you really want to be on target. So as many of your parts are right close to the target, the lower your loss function is. And then the further away you get, the worse and worse things are. So this really drives you to do two things. To minimize quality loss function, what do you need to do? Want to target your process, get it means centered. And second, you want to reduce the deviation, reduce the variance in your process. So with one quality loss function, you're really driving the process towards both goals.

Now, there are questions about how do you calibrate this? And there's some rules of thumb, for example, people take in upper and lower spec limits, you can start to sort of relate these to some cost associated with if a part actually goes out of spec and either you shift it or not, what is the cost of throwing away that part? That might give you a point on the curve and establish an actual dollar value, for example, right at those spec limits. But it actually isn't that important what the number is. It's really that you've got this quadratic dependence which drives you towards minimizing both being off target and the variance.

So here's an example on the use of tolerances. And now pulling in both our distributions and control charts. So here we're assuming the process is normal. And what I've done here is drawn, first off, our in control process with its mean and its plus minus 3 sigma points. Now, notice here I've superimposed x star. So this is the desired target. And notice we're off target a little bit. There's a deviation between where our process mean really sits and our target. And I've also drawn these spec limits.

So now you can start to evaluate various questions like, well, what's the probability not of a false alarm, but the probability of a nonconforming part? And that's easy. You're just looking at the tails. So I can actually go in and say, OK, the probability of a nonconforming part up on this side is just that tail to the right based on the actual performance I need. You can also start to ask questions of, well, how does my probability of getting nonconforming parts change as a function of how far off target I am? That's just shifting the thing and it's moving those tails around. So you can evaluate both of these.

And essentially, a very nice, canonical way of talking about how far apart your spec limits are compared to your distribution, how good your underlying process is compared to your spec limits. Is this notion of process capability, or CP, and then CPK, as we'll see in a minute, also pulls in the idea of the mean difference. So first off, definition of CP or process capability ratio is simply the distance between your upper and spec limits and six sigma of your underlying process. Now, you start to hear the six sigma kinds of terminology. Yeah?

AUDIENCE: [INAUDIBLE]. For example, if you have 30 or 40 steps, and each unit [INAUDIBLE] has its own [INAUDIBLE], is there a system [INAUDIBLE]? And the second question was, for [INAUDIBLE], for example, or for [INAUDIBLE], you have thousands of parameters, each of which has a spec. So what's the quality loss function? Each of them is at some point [INAUDIBLE] all of them as something [INAUDIBLE]. Otherwise, you're not s

DUANE BONING: So you're asking, really, I think two questions. I'll give you a really quick answer right now and then we'll defer some of those. So your first question is, for example, in semiconductor manufacturing or in many processes, we have multiple process steps. And what you can often do is have control charts for each individual one of those. You can talk about the process capability of each one of those. And yes, indeed, there is also very often the question about sort of quality roll up of how quality loss or deviations either add up or compound themselves across multiple steps. So we'll talk about that.

But the first quick answer there is if each of those process steps are independent, then it's simply an additive function. Where it gets interesting is if those two process steps are not independent. So for example, the second process step may magnify or shrink or, in fact, even be designed to compensate for deviations in the first process step, in which case there is intentional correlation between the two. So it's not always a trivial question of how multiple steps interact to aggregate.

The second question is a little different, which is especially in today's highly instrumented processes, there are a zillion parameters that you can measure. There's a zillion control charts. And what we've been talking about so far is a single parameter control chart. And there's some very important corrections one has to make when you've got many, many control charts and many, many parameters. And just to give you the intuition, if the probability for a single control chart with plus minus 3 sigma control limits of a false alarm is 3 out of 1,000 or an average run length of 333 and you have 1,000 independent control charts, how often do you think you're going to see a false alarm? All the time.

So the quick answer there is you actually have to think about an aggregate false alarm rate, not an individual alpha on every chart. And you really should make an adjustment. Basically spread your control limits on every one of your independent individual charts much wider so that you don't drive the engineer nuts chasing false alarms. So we'll, I think, talk about multivariate situations also a little bit later. Very good questions.

So the key idea here on the process capability is simply talking about the spread in your process compared to your spec limits and how good that is. Notice that if I'm off target, this definition CP doesn't change. This is purely talking about the inherent variance, the inherent spread in your process compared to your spec limits. Your inherent six sigma spread compared to your spec limits. If you also want to worry about mean, and you often do, another measure is the CPK.

And what that looks at is the shortest distance either on the right or the left of the chart between your spec limit and your mean. How many sigma apart is your mean from one of the spec limits? And divide that by 3 sigma, 3 sigma on either side. And so now you're penalized for being offset from your target. So let's look at this graphically and get a little bit of a feel here. And here I'm just using tolerance limits. I'm not talking quality loss at this point. CP, CPK are actually just defined for pure plus minus tolerance limits.

So what's drawn here is a set of spec limits that go-- these are sigma markings. So if I look at, first off, CP and CPK, my lower spec limit, upper spec limit, those are six sigma apart compared to the variance of my underlying process. That's just a CP of 1. So qualitatively, CP of 1 you should mesh in your mind just says my spread is such that my plus minus 3 sigma tails are right at the spec limits if I'm perfectly centered. And as pictured here, I am perfectly centered. So my CPK is also 1. In other words, my minimum distance between mu and my upper spec limit or mu and my lower spec limit is both three sigma. So the minimum of those two is still 1. So my CPK is 1.

Now let's consider a slightly different picture. Here I have my spec limits different from my-- or basically my x bar or x star here and my mu are very different. My process spread is exactly the same as before. I still have an upper spec limit minus lower spec limit equal to 6 sigma. So my CP, in some sense, my inherent capability of the process in terms of spread is still exactly the same. But now because I'm off centered, the distance from mu to the lower spec limit is 0. So my CPK is 0 in this case. So qualitatively, these are very different.

Here's another situation. What's changed here? Now in this case, my spread is much smaller. I've got a tighter distribution. And in fact, my upper spec limit to lower spec limit is basically 12 sigma. That's great, right? So that's a CP of 2. Now, I'm still off centered. I'm mistargeted. And so in this case my minimum distance from mean to the lower spec limit, that is the smallest distance. And that's still three sigma. So when I divide that distance there, I get CPK of 1.

And here's one last example. In this case, now I've re-centered, got my CPK up to my full theoretical possible CP of 2 by putting it back on the target. Putting my mean back on the target. So one of the things you use these CPs and CPKs for is in manufacturing. And you also can use it in design to get a quick feel for the relationship between your manufacturing capability, the inherent spread and mean to target centering, compared to your design specs.

So you can characterize your CP and CPK and then evaluate, well, if I make some change in my design spec, what probabilities would that do? If I tightened up my spec limits and start to say, if you just even know just a CP and CPK, you can back out, because that's all in terms of normalized standard deviation, you can back probabilities associated with the number of parts that are not going to conform to those specifications. So if I purely tell you CP and CPK, you know probabilities, fractions of nonconforming parts.

And then you can start to also compare how much would it buy me to move my mean in terms of fraction of nonconforming parts. How much would it buy me? It might cost a lot of energy and time to move the mean a little bit closer to the center of my spec limits. Is it worth it or not? Or would my time be better spent shrinking my variance if I were looking for possible process improvements? You can start to make those comparisons.

And the last thing I want to do with the CP and CPK is give you a little bit of qualitative feel for these. Because what is a good value for CP and CPK? Well, the larger the better. A CP or CPK of 0, that's not good. When you start getting in the ranges of two, that's often what is talked about in manufacturing as being a goal. But it really depends on your process and what the costs of nonconforming parts are.

But just to give you a feel, then, for these probabilities of nonconforming parts and what kind of CPs or CPKs those correspond to, just recognize that a CPK of 1 corresponds to 3 sigma tail right at the spec limit. And being a 3 sigma tail off of whatever my mean is. So I might be off target. But 3 sigma out from the process mean, that's 3 sigma. Think of this is the number of sigmas. And that corresponds to about 1 in 1,000 parts falling outside of our tail.

In fact, you can ask it upper spec limit, lower spec limit. And we actually know the number much more carefully. It's really 3 out of 1,000, those probabilities. If I go to a CPK of about 1.33, that corresponds roughly to a Z of 4. 4 sigmas. I'm four sigmas away from the spec limit. You can see I get about 100x reduction in the number of nonconforming parts. If I can go another standard deviation, or a CPK of about 1.67, I get another factor of 100. And similarly, I get about another factor of 100 if I can get to a CPK of 2. CPK of 2 means my mean is 6 sigma away from either my minimum distance from my upper or lower spec limit. And that's a very small number. That's really 6 sigma away or about one part per billion. 10 to the minus 9.

So why do you want something like a CP of 2 or a CPK of 2? Because one part per billion seems awfully small. And the reason, really, is so that you are robust to small mean shifts. To get down to even if you've got that small mistargeting that you've still got enough room in an inherent process with, say, a CP of 2 or CPK of 2 to be able to have a very small number of nonconforming parts. So whoops, what happened here?

So here's the picture where my CPK and CP are both 2. Perfectly mean centered. Again, that's the 6 sigma distance here between the tail. And we said the probability of having way out here in the tail, that's my fraction of nonconforming. And it's about one part per billion, maybe two parts per billion.

But now let's say that I actually have a mean shift. That was a little weird. Where I've actually shifted by 2 sigma. And recall we said back with control charting, detecting small mean shifts at 1 or 2 sigma are actually kind of hard. You actually have to derive sample size way high to have a very rapid detection. And what we'd like to do in this case now, I've got still the same inherent variability or a CP of 2, but that 2 sigma shift still leaves me 4 sigma before my tail starts impinging very much on my spec limit.

And so now I'm here at about 31 or 32 parts per million nonconforming. So that's kind of the drive towards six sigma is that it allows you to accommodate or have a process that is capable enough that it can accommodate both the variance and target deviations. Both of them are hard to maintain as small as possible.

Now, one can also go back and, if you're given the quality loss function, you can also ask if you have the probability. So think of this as the PDF. If I have the probability distribution function, I've got my underlying process with some mean and variance. You can also do that integral and ask questions about what is my expected quality loss and do the same kinds of things for mean deviations or variance changes. You just have to go in and essentially do the calculation.

One thing that's nice is if you look at the expectation of that quality loss function and just go through and do a little bit of the basic expectation math on that where you split out, factor out x minus x bar, do that squaring and then relate components of the expansion, what you get is one component that is clearly related to your variance in the process and another component that is clearly related to your mistargeting or deviation from the mean. So in one quality loss function, again, you can see and relate in the normal distribution sense how both variance and how target deviation contributes to quality loss.

OK, so just to recap a little bit on process capability, what you've got is some characterization. You've got a model of your real process. Where it's really operating when in control in the best situation. You know its mean. You know its variance. You've got the requirements, the spec limits. You don't use the spec limits to set your control limits. Control limits are based on detection of unusual events. And that's based purely on your model.

You might superimpose your spec limits to be able to quickly calculate and think about what the probability of nonconforming parts are. But those are different things. To relate the two, a good way of thinking about it is the CP and CPK or an expected loss function. What's nice about CP and CPK is we basically normalized away so that if I came up to you and told you my process has a CP of 3 and a CPK of 2.5, you'd go, wow, I'm impressed. That's a pretty good process or awfully loose spec limits. One of the two. You don't really know.

AUDIENCE: If each step is six sigma [INAUDIBLE] but there is no-- and let's say, the real determinant [INAUDIBLE] you could have six sigma for everything and you could still have zero, because [INAUDIBLE]. Is there a CP and CPK for [INAUDIBLE]?

DUANE BONING: I'm not quite sure what you mean by tracking. If by that you mean being on target.

AUDIENCE: Yeah, let's say [INAUDIBLE] horizontal and vertical [INAUDIBLE]. And each step by itself would to be six sigma. And if the tracking is off, what would be within that [INAUDIBLE]?

DUANE BONING: So that's, again, another multivariate question, looking at how important it is for two parameters to be close to each other, even if one or the other could deviate. It's OK if they both.

AUDIENCE: [INAUDIBLE]

DUANE BONING: And so what I would do is if the importance is the difference between those two dimensions, that's what I would characterize and that's what I would chart, put a plot on, and that's what I would set my spec limits on. I might also still care about, say, channel length. And so I might have another chart on that too. I have to be careful, because they're not completely independent. But what you really want to do is get your statistical process control apparatus targeted at or applied to the things that are most critical to you. And so in your scenario, the tracking between those two parameters. That's what I would monitor.

OK. So this is just a recap of the x bar chart with a little-- that's kind of ugly. A little example. I wonder if I can just get rid of all those. Let me just try to go forward here. I think this is trying to reveal a data point one at a time, which we don't really want to do.

So here's an example. Based on what you've seen about an x bar chart, and assuming here we've really set plus minus 3 sigma control limits on this chart, if you were wandering, moving along, and just triggering just on, say, extremal points, what do you think? Was this an alarm? Well, sure. How about that one? That was an alarm. Do you think they were real? Was it a false alarm? A real alarm? Why do you think it was real?

AUDIENCE: Because [INAUDIBLE]. And those two are [INAUDIBLE] as well. So I expect to have more points [INAUDIBLE].

DUANE BONING: So the real answer is we don't know, but you're doing exactly the right thing going back to some of the discussion we had earlier of looking for additional evidence. So your evidence here was saying if I got this alarm, I might look back and say one of the previous points, hmm, there's an awful lot of them to one side of the mean. And in fact, maybe adrift, although at this point sort of looks like it's counter to the drift. So there's a little counter evidence.

By the time I get to this point, now I'm starting to get really convinced that something is weird. Because now look at this. What's the probability of having all of those points? Almost all except for that one or two to one side of the mean. Well, you could actually go and calculate that probability. Sounds like a problem set problem. But qualitatively, you know something looks like it's off target. So this is the kind of thinking that you would want to do.

Now, this is a little tricky. If you were just eyeballing it, what do you think a CP or a CPK would be for this process? I have no idea. I haven't done the calculation. You could go and do the calculation based on what? What would you do to actually formally try to calculate a CP or a CPK for this process?

AUDIENCE: [INAUDIBLE]

DUANE BONING: Yeah. You'd calculate the mean and the standard deviation. First off, you have the true mean and you have the true control limits. Could you calculate a CP or CPK based on mean and standard deviation? I can give you those numbers.

AUDIENCE: Because you need upper and lower spec limits.

DUANE BONING: You need spec limits. Control chart by itself never gives you spec limits. So you don't know. It may well be I've got really tight spec limits. My parts really need to operate down in here. By the way, you also have to be careful because if these are based on n samples, the sample distribution is tighter than the underlying process distribution. So if you see spec limits also on your control chart, it's kind of a warning. Because usually a control chart is based on sampling.

So by itself, you can't really know what your CP and CPK are. And I would be cautious about assuming your spec limits are anywhere near your control limit. Don't assume your spec limits are plus minus 3 sigma. You're really falling in the trap of sort of basically assuming a CP or a CPK of 1 process, which is not often the case.

AUDIENCE: [INAUDIBLE] outside your control limits?

DUANE BONING: Absolutely. Why? Here you're having a huge fraction of nonconforming. Yeah. So that's my point. Your spec limits may be very tight and your process would be horrible. This is not a capable process. So normally yes. What you normally want to do is trigger on alarms much before you have parts that are falling outside of spec so that you can react to drifts in the process before they cause you disaster. Absolutely. That was implicit. I'm glad you made it explicit. Yes, yes.

OK. So I think we've talked about a lot of these. What I'd like to do is give you a little glimpse of some additional control chart ideas that are beyond the x bar s or x bar range chart. There's some very good things about the x bar chart. It's relatively simple, fairly transparent, meaning you sort of know what each data point means. You know what an alarm means. It relates very closely to assumptions of normality.

And in fact, by the act of sampling, you're enforcing some of these assumptions. By sampling within larger than 1, generally, you've got the central limit theorem on your side. So x bar really trends towards a normal distribution. And we'll talk about this a little bit. If I've got long times between samples, if there's very, very short time scale correlation between sample points, these samples are truly independent measures of the process. So as long as I've got reasonably long sampling times, each of my samples are an independent draw from the process.

The problems are, yes, sample size, I need n more than one. And very often I may be measuring every one. So I'd like to get to some ideas or control charts that might be able to either react faster, a shorter average run length, or be able to utilize the fact that you've got more data and I can have this faster time response.

So in fact, let's think real briefly about the limit. What if your sample size is just one? Could you go to that limit? There's a couple of tricky points here. The good thing is you have a lot of data. I can look at every part and respond quickly. But now a lot of these assumptions on the statistics or the calculation of these basic statistics don't have meaning anymore.

For a sample of size one, what is the average? Well, OK, it's just that value. What's the standard deviation? No, it's undefined. 1 over n minus 1 in the standard. You don't have enough data. You have no estimate of a variance from that sample size of one.

So an approach here is essentially use kind of a variant referred to as a running control chart or a running average. So you look back not just on the new sample but some number of previous data points and kind of have a window that moves along or some calculated average of past history where I might weight my most recent measurement either equally or more importantly than my past measurements and use that as an estimate for the process average and the process variance.

AUDIENCE: Does that violate the independent assumption?

DUANE BONING: Absolutely. It does violate the independent assumption. If, however, they truly are independent, you can still formulate statistics and probabilities associated with that. In other words, if your process data was independent. What you've got is from any one point in time, I can calculate the probabilities associated with that. But if I then calculated that for my next point in time using some of the same data, I would never use some of the WECO rules like the probability of two points in a row falling above, because they're no longer independent. I can't just multiply those two probabilities.

So it would get really messy trying to calculate sequence probabilities using a running average. Now, that's different for we'll see this notion of a cumulative sum chart where actually you want to use the fact that there is dependence to aggregate deviations and detect them even faster. So you use that very notion of lack of independence.

So let's talk about these sort of running average or an n equals 1 kind of design. What you'd like to do is have increased sensitivity. You want to detect small changes, especially small delta changes. So small shifts are one class. There's also cases where you want to detect a small shift. We'll come to that chart a little bit later. You could also have the desire to not have too many false alarms. So you still want noise rejection. And one of the problems with some of these are you might, in fact, actually have higher variance estimators. And you've got to be very careful, because you can get larger false alarm rates unless you compensate for that.

But the basic idea, and this almost goes back to your earlier question, imagine, first off, what we're doing with just a regular x bar chart but where we're measuring every part and then just aggregating in samples of different size. So what I've got here is the run data. And then in the little purple pictures is an x bar calculated for each group of four sample points. So every four I calculate the average, then I move on to the next four and calculate the average.

And so in a typical sort of x bar chart with sample size n equals 4 where my sample time is also n equals 4, I've got a smoother curve here with the x bar. And we didn't explicitly talk about it, but one of the advantages of the x bar chart, what we talked about was the squeezing of the distribution down. Another way of looking at it is we're filtering high frequency deviations down to try to detect a smoother average that's more indicative of the true mean of the process. I'm really not trying to detect variances. I'd just as soon have one signal that just told me how good my mean is. And so an x bar really does filtering of your data.

So if you think about it from the filtering perspective, what does the filtering do? It reduces some of these sharp peaks. It aggregates or hides some of this intermediate data, reduces this high frequency content to try to get to the low frequency, either drift or a mean shift a single mean shift that's applying to many data points together. And so you can also think about these as filtering operations.

And any time you've got filtering, you do get at some of these questions we were kind of asking about about how related or how independent are samples at different points in time. So one can actually go in and have a measure of this independence. And if the process itself within a sample, the points are highly correlated, you actually have to explicitly think about this and for control charting spread those data points out.

There is an autocorrelation function that's rxx as a function of the time separation or the number of sample points between them that has this definition here. If you think about it in terms of the time for an autocorrelated function, that is in fact inherent in a linear first order system, the autocorrelation function drops off fairly rapidly from 0 time gap. So if I sample at this point in time and I go some time away, the similarity of those sample points decays away.

If it were truly uncorrelated, each and every new data point is completely independent of the other. And the point is simply here that if you're sampling from some physical processes, there may be an inherent correlation that isn't the one you're trying to get at. You're trying to detect long term trends to not detect very small shifts. And so you can play around with the sample time compared to the inherent correlation of the process.

All this is saying is you want to pick your sample time ultimately such that if I pick this sample here and then my sample time some point later, this underlying short term correlation is not what I'm detecting. I'm detecting again a big mean shift. So we've been talking about this. Again, what you're looking for here is maybe pulling your small sample out, using, in fact, some of the correlation there to be able to say what the best estimate of the mean is right at that point in time and then have uncorrelated samples around that.

OK, well what I think I'll do is probably stop now. I think I'll pick up next time, defer or wait a little bit on diving into the yield until maybe mid time Thursday. Because I do want you to get a feel for these alternative charts. So the charts we'll talk about next time include a moving average chart. Essentially you aggregate your data and use that to form estimates of local means and variances. And then we'll look at other clever ways of using more recent data and weight towards those. So we'll talk about an exponentially weighted moving average control chart and what that idea is.

And then finally, we'll also give you a little bit of a feel for a cumulative sum chart where we aggregate deviations from the mean, because we're really looking for was there a small change or is there a small drift going on in the process? And how do we increase the sensitivity to detect that small shift as rapidly as possible? So we'll talk about charts that are really meant to squeeze that average run length for detection down.

So we'll see you on Thursday. And again, while you're working on today's problem set, if you want to feel good, take a glance as soon as it's released at next week's problem set. It will make you feel much better.