Lean Six Sigma Cycle Time Analysis:

Processes with a Backlog of Work.

By Rick Haynes
Master Black Belt | Austin, Texas

www.smartersolutions.com

eTech Resource
June 2007

In an improvement project targeting a reduction in process lead time, one typically faces transactions or items that spend a lot of time in queue waiting to be worked.

Then, there are questions about the "Best" method to estimate the current capability. Do you use the data from all the transaction that are in the queue in the estimate? If so, how do you use them?

This tech article shows the pro's and con's with three typical analysis methods so that you will know what is the best method for your application.

Read and learn how these quick and simple Minitab resources will help you do just that!

Software: Minitab Version 14 or 15.

Many a Lean Six Sigma practitioner crosses paths with a project to reduce the cycle time of a process when there is a significant backlog of work. Why? Typically the large backlog of work is an indicator of a problem and problems usually spawn projects.

This tech note will focus on the improvement analysis methods rather than a specific use of a tool.

Typical Problem statement: Time to finish a job is fails to meet the customer requirement of no greater than 7 days for 53% of the jobs, which results in significant customer complaints and possible loss of the contract.

OK, so what do we know about Cycle Time data.

Little's Law :

Lead time = ACR/WIP (average completion rate / work in progress)

ACR = worked hrs avail per day / avg. hrs per job

Takt time:

Customer demand rate in time between job close outs.

Cycle time data has a natural limit of zero, which leads to a non-normal distribution.
Cycle time data typically has a log-normal distribution.

If the job input rate = job output rate, the process is stable. An unstable process indicates that the input and output rates differ and a new equilibrium has not been attained. (Analogous to a mass balance analysis in related to the water level in a tank that has water going in and out)

Improvements in this process can be made by increasing the hrs worked per day (more people or longer days) or by reducing the avg. hrs per job.

Now let's look at a simulated process which has;

- 23 jobs input into the process a day.
- Today there were 32 close outs and 197 still in work that did not finish.

Note: The data for this tech note was generated with a 100 day simulation of a transactional process that had a random but stable process cycle time. Twenty three transactions entered the process each day, what was completed each day and how long each was in work is a result of the simulation parameters. Day 100 was used in this analysis to ensure the data was from a steady state condition with a true representation of what a daily production meeting would see for data if the process was reviewed.

© Smarter Solutions Inc., www.smartersolutions.com

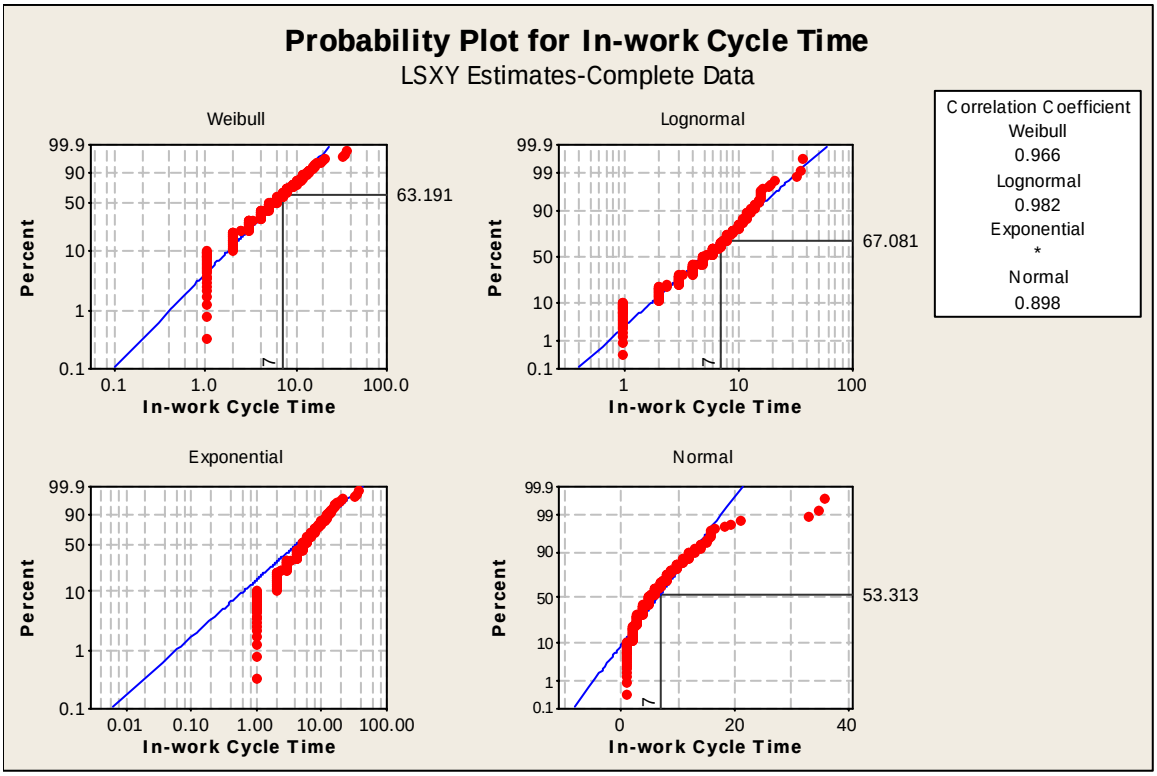
What is the capability to meet the customer requirement?

We have a customer requirement of less than or equal to 7 days. We know it is cycle time data, but might not recall this is best fit with a lognormal distribution. Since we only measure capability on the actual data, we need to find the distribution.

But what actual data do we use?

Use all the data, finished and unfinished work.

How about all the data for jobs in work today. This will include the jobs that finished today and the ones that are still in work but have not finished.

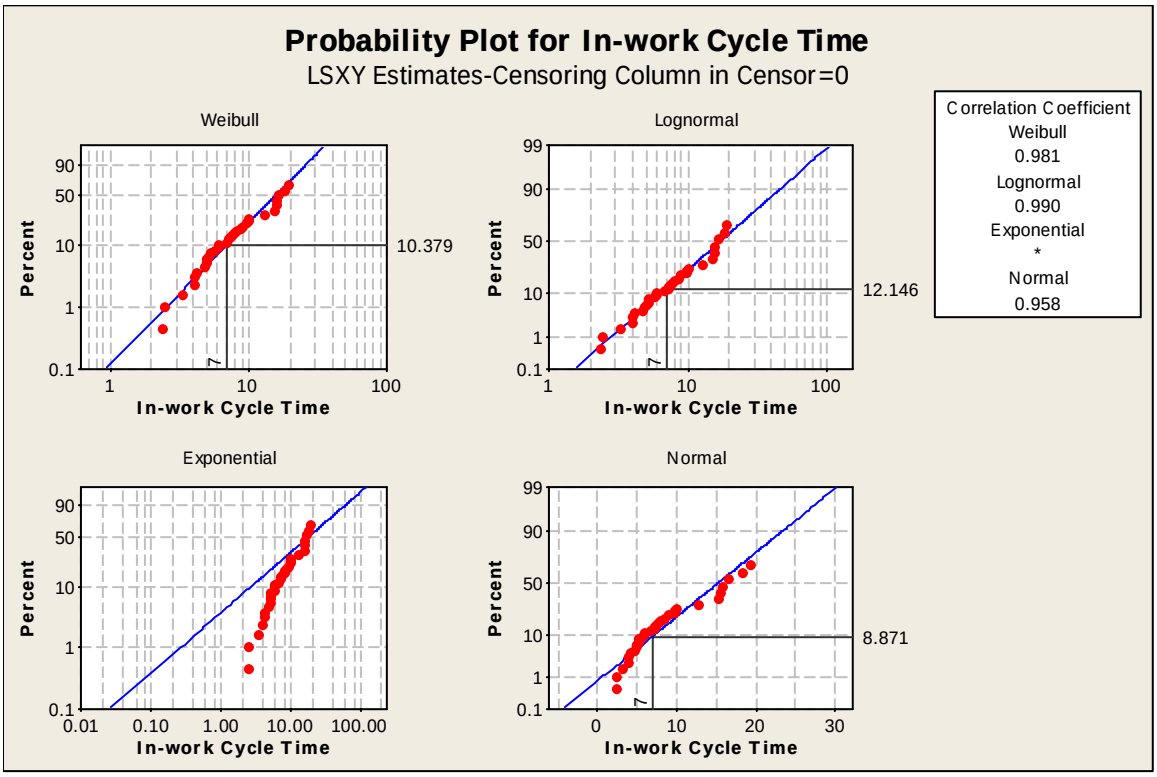


It shows that the lognormal distribution fit is the best choice. If we use a target of seven days, the lognormal distribution predicts that we will have 67% in tolerance or 33% of the work will exceed the customer requirement. This is different from our historical average of 53% failing to meet the requirements. The process must have improved, or did it?

© Smarter Solutions Inc., www.smartersolutions.com

Using all the data, but considering the unfinished work as censored.

Now let's consider a more statistical type of analysis. We realize that 32 jobs finished today and we know the exact cycle time for them. The remaining 197 jobs have yet to finish, so the cycle time value reported for those jobs is really how long they have been worked. The final cycle time values will be more than these values, but how much more is not known. Sounds a lot like "Right-Censored Data" for a reliability test where 129 items are being evaluated and 32 have shown failures with the remainder still in testing. So maybe a censored data analysis would be appropriate.

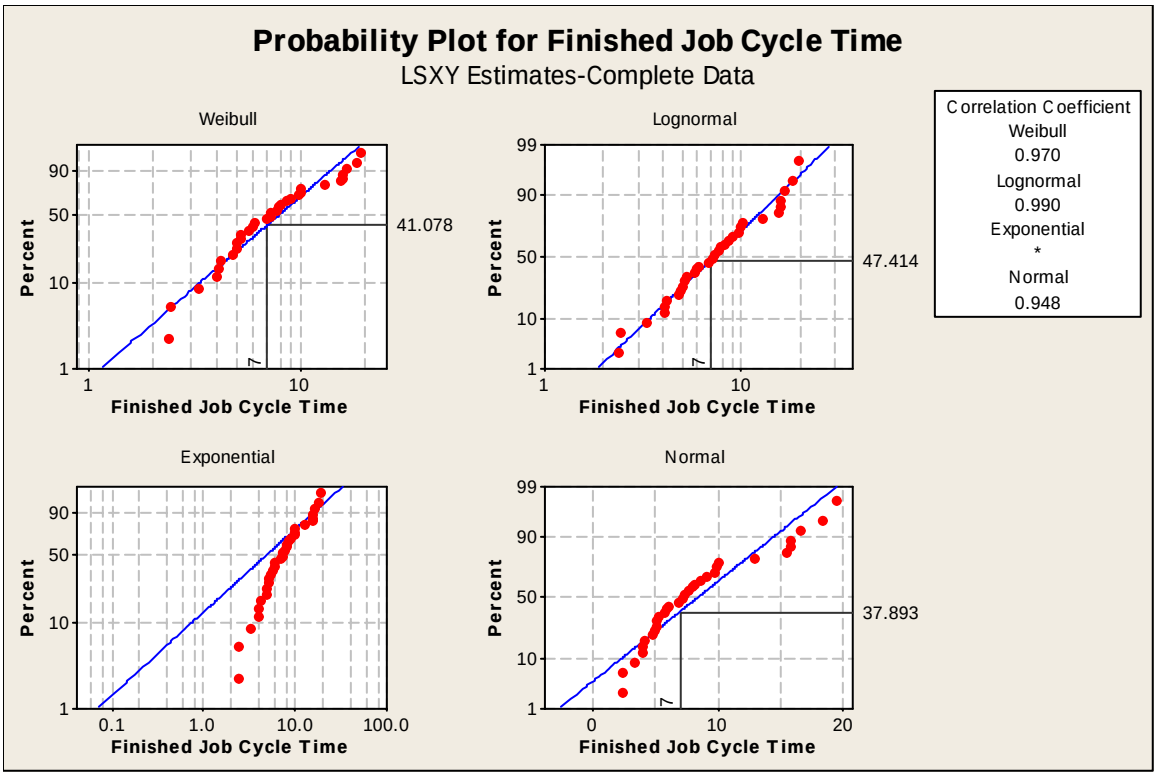


With this analysis it again shows that the lognormal distribution best fits the data, but the prediction is only that 12% of the work will complete within 7 days and 88% will fail to meet the customer demand. Has the process taken a turn for the worse?

Using only the jobs that finished today, ignore the unfinished work.

Now we will try the third way to look at the data, using only the jobs that have finished today. We will ignore all the in-process work that has not finished today.

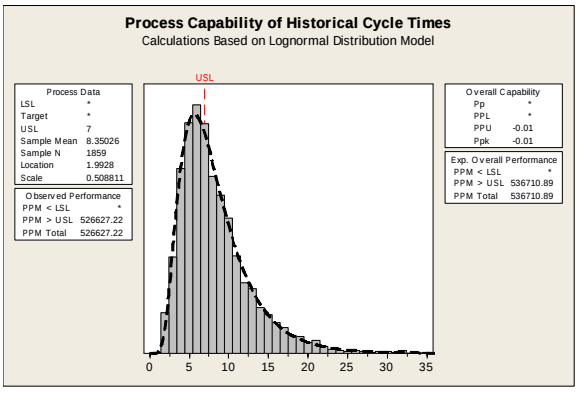
© Smarter Solutions Inc., www.smartersolutions.com

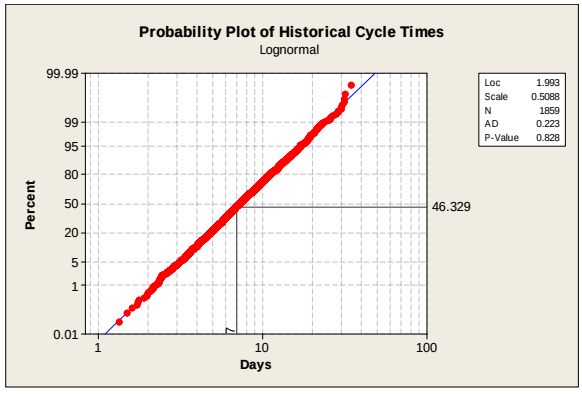


Again, the lognormal is the best choice for a distribution, but this time we are predicting that 47% of the jobs will complete within 7 days and 53% will exceed the customer requirement. This result matches the historical performance.

This leads into the message of this tech note. Why are the results so different for three rational methods of estimating the capability?

First of all, let's look at the baseline process for the simulation. Here is a capability analysis of the true process.



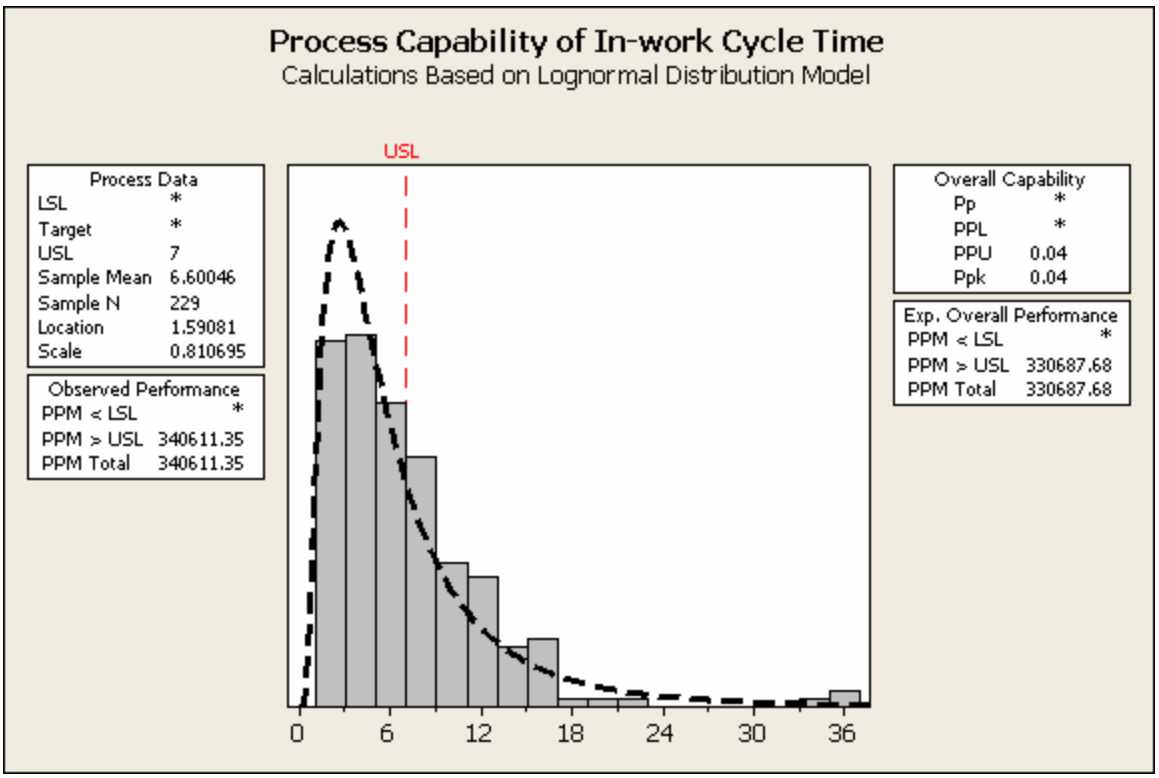


The process mean is 8.35 days and the capability is 53.7% exceeding the customer specification of 7 days.

© Smarter Solutions Inc., www.smartersolutions.com

Using all of the existing data, finished and unfinished.

In this method all jobs are considered equal and will be reporting the average close out cycle time using all of the data existing in a single day. This would mean that you would average all the times for the jobs closed out today and all of the cycle times accumulated for the jobs not yet closed. Here is what that would look like.



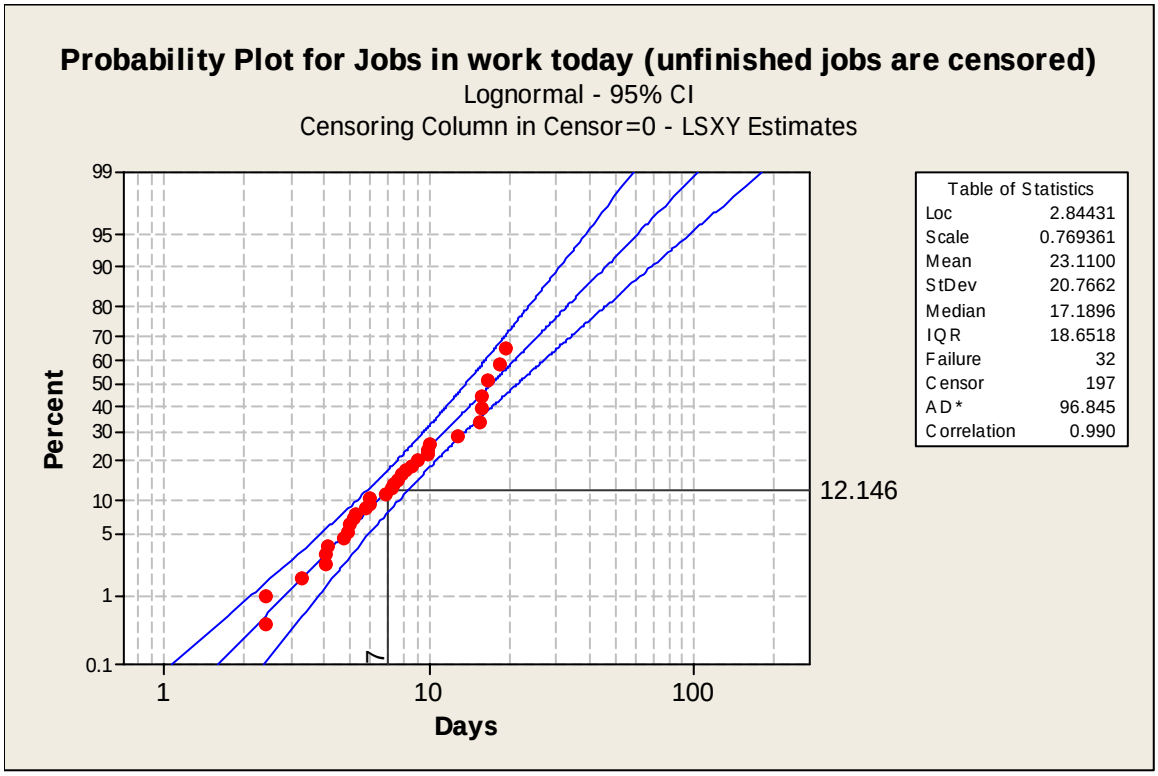
Notice that this method dramatically reduces the estimate of the average time to complete a job.(6.6 days vs. true 8.35 day average) The decrease is because the unfinished jobs are being averaged with the time spent to date instead of their actual completion cycle time. For example, there are 23 jobs that started today that are included into the average that have a cycle time value of 1.

Conclusion: This method underestimates the average cycle time.

© Smarter Solutions Inc., www.smartersolutions.com

Using all the data, but considering the unfinished work as censored.

Using a Censored data analysis on today's data seems like it should be a better method. This method would recognize that all the unfinished work would be considered as right censored data, where the true cycle time is greater than the current accumulated cycle time. This analysis is shown here.



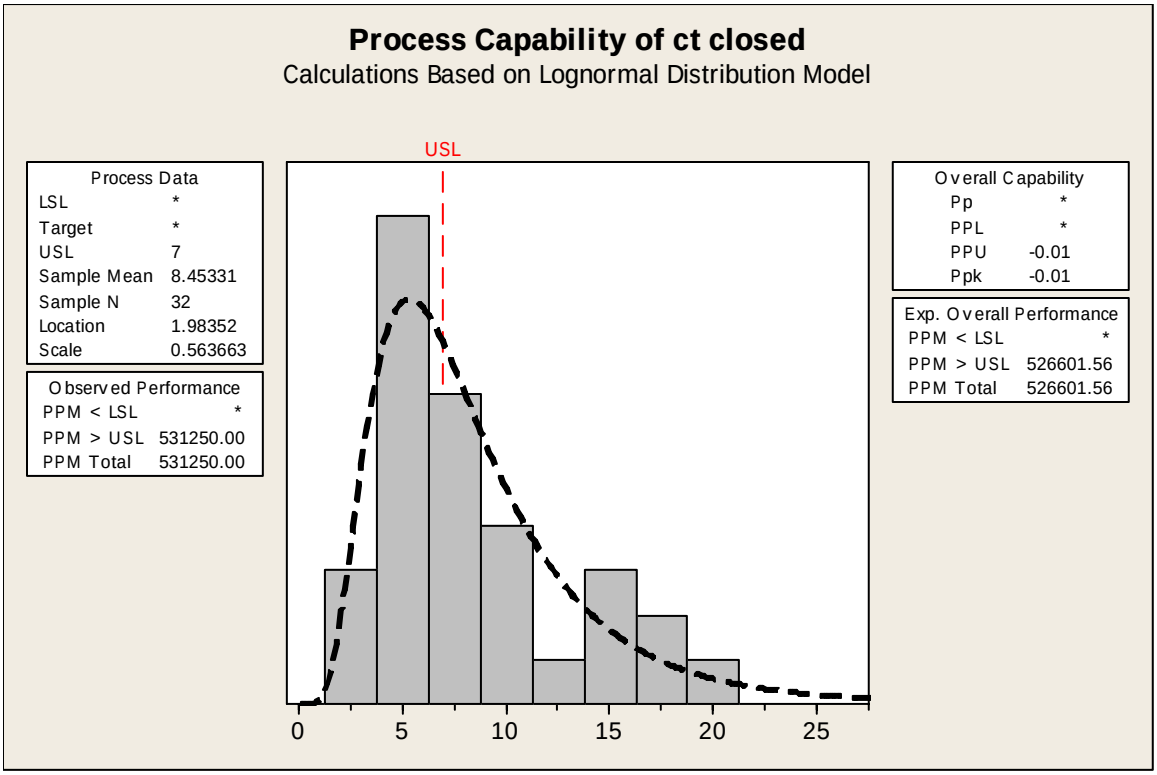
Notice that this method increases the estimated average close out cycle time (23 days vs. 8.35 day) since it is considering the daily close out along with the unfinished jobs as if they closed today. Why it over estimates the average is because those unfinished jobs will be closed out as part of a series of quicker closing jobs that are not represented in this data set resulting in a higher proportion of long cycle time jobs in this analysis.

Conclusion: This method over estimates the average cycle time of the process.

© Smarter Solutions Inc., www.smartersolutions.com

Using only the work completed today

In this case we only use the 19 jobs that finished today. For this analysis the result is as follows;



For this case there is significantly less data used, about 90% less than both of the prior methods. We are told that bigger samples are better in our classes, but in this case the analysis provides a good estimate of the mean (8.45 vs. 8.35 and the capability (53% vs. an actual 54%).

Why is less better than more data? It is because of one of the other frequently forgotten assumptions in data analysis. Only combine like data into an analysis. In this case the only difference in the 19 points was the random nature of the process. In the prior two analysis we had two or more groups of data that had different expected averages and variances, far more difference than just normal variation.

© Smarter Solutions Inc., www.smartersolutions.com

To be complete, let us revisit the lean tools when applied to this question.

What would the Lean tools tell us?

Lead time = ACR/WIP. Today that would be 23/197 = 8.57 days.
This is pretty close to the true average of 8.35 days.

Takt time= customer demand rate. For this example it is 23 per day, assuming an 8*60 minutes per day the takt time would be 21 minutes or a job should finish every 21 minutes for every process step.

The lean tools are quick and easy, but also allow the practitioner to check on their true understanding. If the lean tools provided significantly different answers, it would lead to to investigate xxxxx.

What would xxxxx be, well if you figure it out, you do have a true understanding of the tools.

You can find references to these methods in Implementing Six Sigma, 2nd ed. in the Lean chapter 44, Table 7.1 “Distribution/Process Application Overview”, Chapter 43 examples 43.15 and 43.16, along with the censored data analysis in Chapter 41.

Rick Haynes
Master Black Belt
Smarter Solutions, Inc. | **Rick.h@smartersolutions.com**

Rick Haynes is a Master Black Belt at Smarter Solutions. Haynes holds a Bachelor of Science in Chemistry from the University of Idaho in Moscow, Idaho, and a Master of Science in Statistics from the University of Texas in Austin, Texas. With more than 20 years’ experience in project management, training programs, applications, systems, engineering, and customer support, Haynes is responsible from providing consulting, coaching, and training to Smarter Solutions’ clients.

© Smarter Solutions Inc., www.smartersolutions.com