EVALUATION STRATEGIES IN qs-STAT AND solara.MP

STEPHAN CONRAD | Q-DAS GMBH

Just a few clicks to reveal the absolute truth or rather “do not trust any statistics…”?

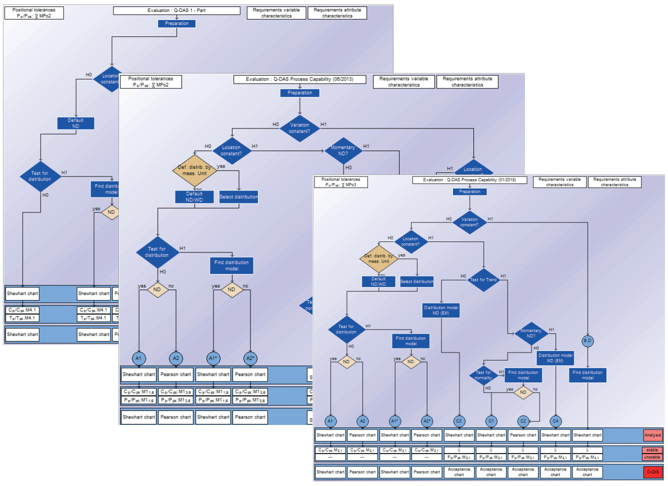
“Evaluation strategy” is the magic word in Q-DAS software products qs-STAT, solara.MP, destra and the CAMERA tools. This article tells you what an evaluation strategy actually is and how you can design it.

Before dealing with the selected topic, I first want to give a warning. Please note that I wrote parts of the text with a slight twinkle in my eye but only parts of it, not the whole text.

<mode sporadic irony on>

Be aware of the fact that any software evaluation strategy is nothing but a representation of a quality philosophy, of course, i.e. a philosopher develops it in his head, puts it on paper and calls it corporate strategy. Then he implements the strategy in the software to ensure a uniform approach throughout the company. The entire company will then apply the same correct methods.

Even though the creation of an evaluation strategy caused a lot of trouble, you will finally do everything right as long as you follow this strategy. Or more precisely: Whoever pursues the strategy of company Smith, will do everything the right way – for company Smith. However, if you work for company Taylor, you will have to use Taylor's strategy. This leads us to a decisive question. If Smith's strategy obtains results that are different from the ones calculated by Taylor, but both results are correct, which result can be considered to be more correct? We can even carry this question to the extremes: Which one of the many evaluation strategies in qs-STAT and solara.MP is the most correct?



Critical fellows only know a single answer to this question: My own strategy is the most correct.

Even in "pre"factual times, men already regarded statistics with deep suspicion. Starting a conversation with "Statistically speaking..." usually made your opponent frown. This is no surprise because statistics is a branch of mathematics. "The cognition of the reason through construction of conceptions is mathematical", said Immanuel Kant. He also stated that no natural science contained more truth than "pure mathematics". But why did M. L. H. Kessel alias Hans Bruehl dare to say statistics was "the fairy tale of rationality"? And there are even more irreverent quotes we avoid here for reasons of youth protection!

Where do these issues we have with statistics actually come from?

The first aspect you should know is that there are two different types of statistics – descriptive statistics and inferential statistics. Descriptive statistics reflects the actual situation. Any pseudo optimisation measures are often easy to identify.

However, this is not the case for inferential statistics since it draws conclusions from the actual situation. These conclusions rely on a data basis containing nothing but a fraction of reality. Otherwise, we would not need any statistics. The quality of statistics sinks or swims with the quality of data.

> Any fool can know. The point is to understand.

To reach reasonable conclusions, we must understand the structures leading to the data we collected before. We thus need a model of the very reality that generated these data.

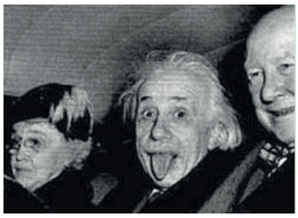


Fig. 1 Albert Einstein

Fig. 2 Niels Bohr

Models do not have to be unambiguous. They can see in the truest sense of the word. The light that has just reached your eyes is a wave and a particle at the same time. You can hear the clattering of a light particle striking a photometer and when you look at a rainbow, you will see the different colours since the refraction angle varies with wavelength. Not a single model is able to fully describe the true character. Whoever works in the field of light always applies the model best reflecting the aspect he observes. Even Albert Einstein doubted that he grasped the exact meaning of it. In 1951, he said: "Any fool can know. The point is to understand." His colleague Niels Bohr expressed the same opinion: "There are trivial truths and the great truths. The opposite of trivial truth is plainly false. The opposite of the great truth is also true."

Fig. 3 Ludwig XIV

Our statistical models pose the same dilemma. If the whole problem yet seems to be too knotty for you, just select one of the pragmatic approaches mentioned above, such as "la vérité c'est moi" (in order not to get involved in current politics, we refer to the French Sun King Ludwig XIV in 60 to 75 per cent of all cases). When in doubt, just define your own truth and defend your point of view without a shadow of a doubt!

As soon as we define a model, we draw conclusions from this model, e.g. the assumed capability index of the process or how many parts that are n.o.k. we expect. At first glance, this sounds like a straightforward approach since these conclusions are based on formulas. Looking back on our data basis, however, we must admit that the truth might be a bit different. As an example, when you conduct a machine performance study and produce 51 parts, you will throw away the first one. This action slightly changes you C_m/C_{mk} value since the 51^{st} is not identical to the first part but the machine is still the same. This example shows that we get a random result having an associated confidence interval that includes the "true result" with a certain probability. We thus obtain a correct result but not the one and only true result.

You might already guess the drama. Even though all results are correct, they are not essentially true. The result could even be true by pure chance but we would not even know it. True values are by definition indeterminate.

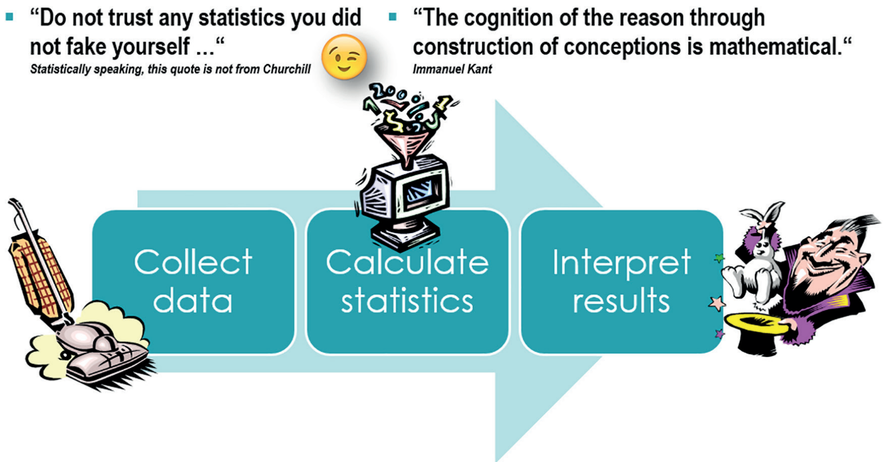
I do not want to provide you with the impression that statistics is nothing but an arbitrary conjuring of lobby-driven pseudo results, so I must admit that the knack of statistics is to express, calculate and apply these uncertainties.

The fundamental error in the comprehension of statistics is to believe that you are able to calculate the truth straight to the point. Whoever ignores confidence intervals takes the first step in reaching a false conclusion. Those who understand the meaning of statistics know that it brings us closer to the truth, much closer than any other technique.

With this knowledge in mind, we go back to the evaluation strategy. We have already discovered that there are many ways to approach reality. Different evaluation strategies, however, adopt different approaches. Many evaluation strategies even neglect specific cases since their creator considers them to be irrelevant. This is the reason why it is of utmost importance to understand the backgrounds and opportunities of evaluation strategies and to apply them in a target-oriented manner.

<mode sporadic irony reduced>

Our evaluation strategy thus has to consider the three stages of data analysis we are going to pass through.



STAGE 1 – COLLECTING DATA

Are the raw data useable?

You should find out whether the measured values represent the process to be evaluated. These values must contain any required information but nothing unnecessary. Only the evaluator is able to answer the first question ("anything required contained"). The software, however, helps supply the answer to the second part of the question. You can e.g. check whether single values keep a surprisingly large distance to the centre or even violate plausibility limits. The level of complexity rises with part anomaly analyses detecting anomalies in all characteristics of a part. You might notice that I always try to avoid the word "outlier". Not a single software product is able to `prove´ the presence of outliers. Well-known tests for outliers (Grubbs, David-Hartley-Pearson, Hampel...) only identify symptoms of an outlier under specific conditions (a normal distribution is frequently required) and provide the following indication: "Assuming the normal distribution model, this value keeps a surprisingly large distance from the tolerance centre or it is just not part of a normal distribution." Only a process expert is able to assess whether it is really an outlier or not.

STAGE 2 – CALCULATING STATISTICS

What are the statistics you need to calculate and which methods do you apply?

This is supposed to be the most fortunate part of the evaluation strategy since this is what "computers" are made for. In addition, there are various documents specifying possible calculation methods. As soon as you go for a specific guideline, you adjust the settings accordingly. Since these guidelines, however, are often not consistent, you also have the option to adapt to an international standard. We like to recommend the ISO 22514 series `Statistical methods in process management - Capability and performance´. But why the hell are there companies that do not stick to international standards? Well, from the Beatles to Lady Gaga, from van Gogh to Andy Warhol – is there any better way to express one's individuality than by deviating from the norm? And on the other hand, "the norm is just" the current state of the art – why not try to make it even better?

STAGE 3 – EVALUATING THE RESULT

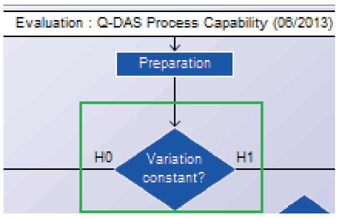
How do you interpret statistics?

The automated computer-assisted evaluation strategy supports you as long as you have to make objective decisions. It helps you find out whether a value reached a limit or is still within a confidence interval. The subsequent part of the interpretation is once again rather subjective and based on pure experience. The evaluator thus carries sole responsibility.

Let's assume that we want to create an evaluation strategy for the qs-STAT Process Capability Analysis module. To comply with a validated standard, we apply the ISO 22514-2 standard. Since I have no intention to write a never-ending article, I will just focus on the key aspects.

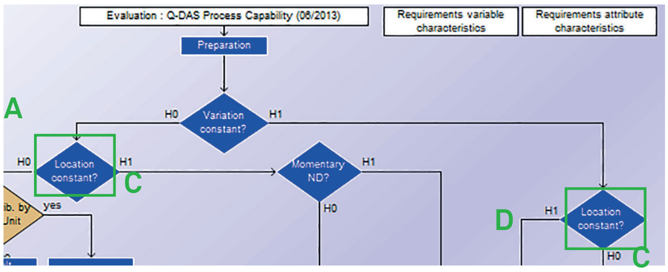
The following specifications help us determine the suitable distribution time model.

The classification into distribution time models offers the advantage to draw conclusions about instabilities over time, possibly required process optimisation and control mechanisms to be applied.



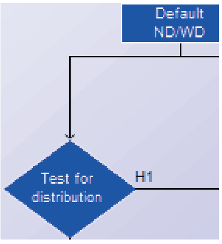
The first step is to find out whether the variation of the process is stable over time. After confirming H_0 (null hypothesis), you continue with distribution time models A and C. When you confirm H_1 (alternative hypothesis), you either apply distribution time model B or D.

In both cases we must check whether the location is constant over time. The left-hand branch leads from H_0 to distribution time model A whereas H_1 results in distribution time model C. The right-hand branch leads from H_0 to model B and from H_1 to model D. The first simple classification is done now.



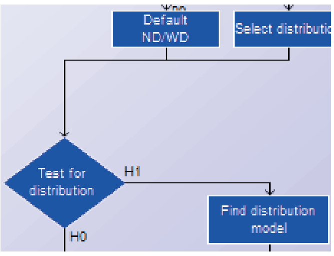
DISTRIBUTION TIME MODELS ACCORDING TO ISO 22514-2

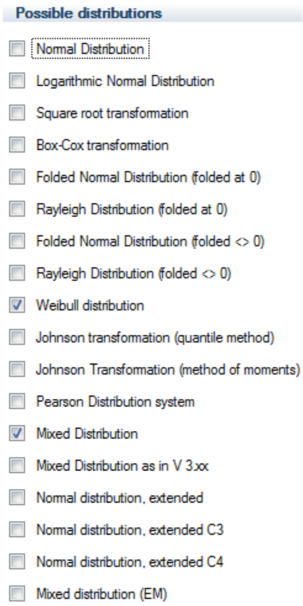
Process standard deviation	Process average							
	constant			Not constant				
		A			C			
		A1	A2		C1	C2	C3	C4
constant				**Location**	Random	Random	Systematic (e.g. trend)	Systematic and random (e.g. lot to lot)
	Short-time distribution	Normally distributed	Not normally distributed - unimodal	**Short-time distribution**	Normally distributed	Normally distributed	Normally distributed	Normally distributed
				Resulting distribution	Normally distributed	Not normally distributed - unimodal	Any shape	Any shape (e.g. multimodal)
not constant	**Resulting distribution**	B		Resulting distribution	D			
		Any shape			Any shape			

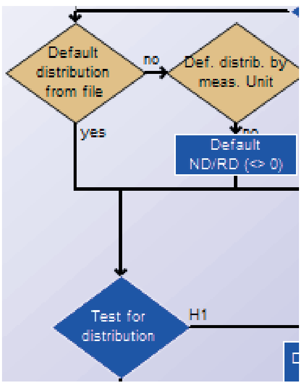


Now let's have a closer look at branch A. We must distinguish distribution time models A1 and A2 in detail. It would be enough to test for normality.

But what will we do if it is a non-normal distribution? The standard does not provide much information about it. We need a unimodal distribution model now describing the data sufficiently well. However, there are plenty of it. Specify the pool qs-STAT may choose from.

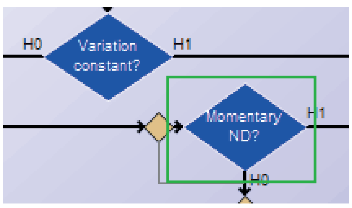


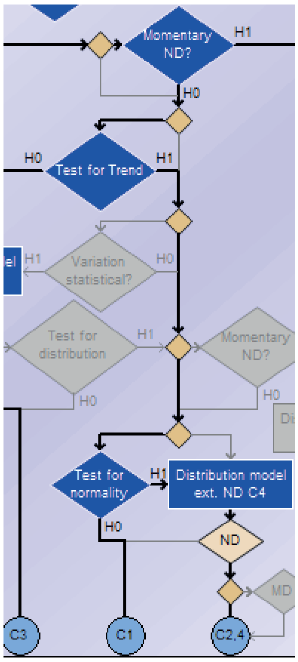




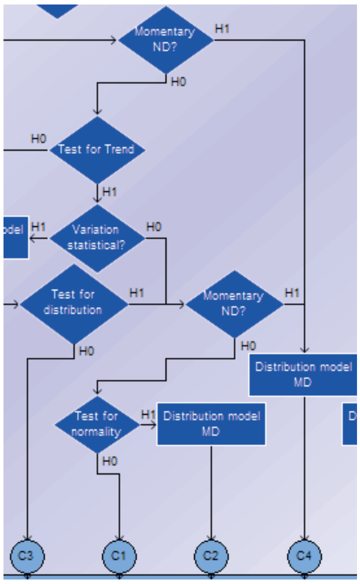
Do you even possess the technical knowledge helping you draw conclusions about probable distribution models? Apply this knowledge to make a preselection.

In the C-branch, however, we must distinguish models C1 to C4. So we need to find out whether it is a momentary normal distribution.

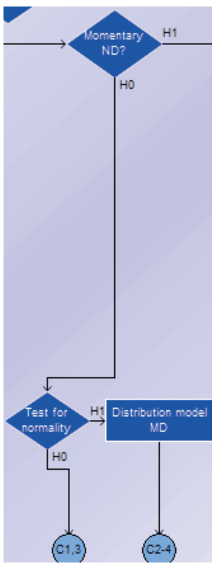




In case the momentary normal distribution does not apply, the data do not meet the C-model requirements. We thus accept H_1 which leads us to model. However, if all the facts indicate a momentary normal distribution, it is important to distinguish between different types of C-models. C1 results from a normal distribution – despite the trend. C2 is unimodal but not normally distributed. The special feature about C3 is a linear trend. C4 is a kind of rag-and-bone man collecting everything that does not fit otherwise. This basically concerns mixed distributions with a clearly constant momentary variation resulting from a variable location. Therefore we have to check this once again.

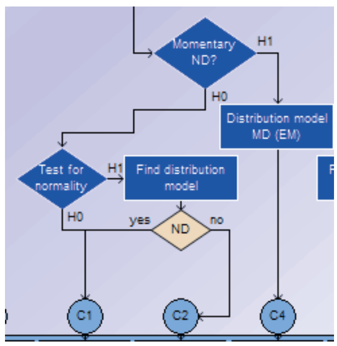


Whoever pays attention to the deactivated options will quickly realise that this is not the only possible way. The graphic on the left shows the alternative.

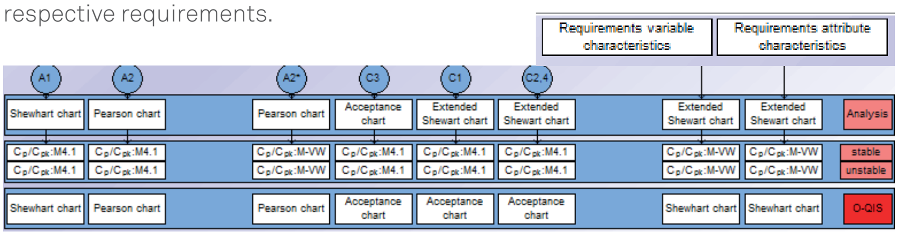


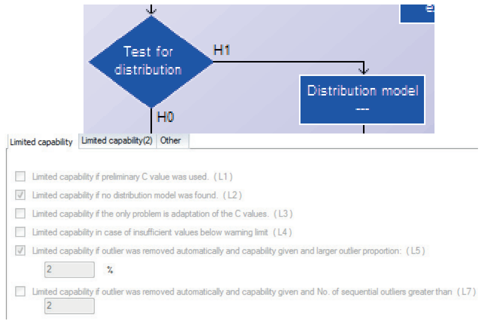
Many companies do not assign any importance to the differentiation between all these C-models. In this case, the result might look as follows:

or



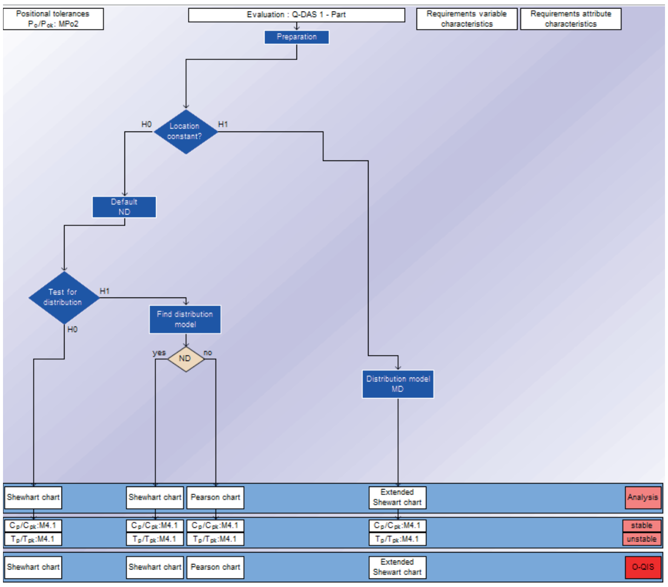
After having identified these models, we continue by evaluating the previous process stability and by selecting future control criteria based on quality control charts. Now we calculate capability indices and compare them to the respective requirements.





On the other hand, you can extend these methods, e.g. to prompt a manual intervention in case of deviations from (theoretically) expected distribution models.

We do not want to go into further details. So let's give critics a chance to speak. A typical plea they raise on the one hand is that even though the analysis of distribution time models is quite interesting, it is complex and incompatible with the "low level" statistics of tier (n+1). If you want to reduce the methods to a "best fit" distribution model, the strategy might e.g. look as follows.



We thus gained a first insight into the meaning, purpose and structure of the evaluation strategy in qs-STAT. The same applies to solara.MP. The only difference is that the procedure in solara.MP is rather a kind of flow chart; the methods, however, are defined as individual study types. Even solara.MP provides a multitude of different versions that are good entertainment value.

It is time to come to an end... Our seminars 017-SW on machine performance studies and process capability analyses in qs-STAT and 018-SW on the capability of measurement systems and measurement processes provide you with many more details.

<mode sporadic irony off>

Interested in this topic?

Q-DAS GmbH
Eisleber Str. 2
69469 Weinheim
HexagonMi.com | q-das.de | teq.de
stephan.conrad@hexagon.com