How Much Confidence Should You Have in Your Style Analysis?

Thomas Becker

Zephyr Associates, Inc.

Zephyr Cove, NV

thomas@styleadvisor.com

Question Addressed by This Talk:

How much confidence should you have in your returns-based style analysis (RBSA)?

First, we must ask: when and under what circumstances is it necessary to discuss *confidence* in the result of a calculation?

It is necessary whenever there is:

- Uncertainty
- The possibility of error
- Approximation instead of exactness

What kind of uncertainty could there possibly be in a calculation such as RBSA?

To answer this question, we must distinguish between two different kinds, or sources, of uncertainty.

First Source of Uncertainty: *Calculation Error*

Engineers often use calculations or algorithms that are inherently error-prone.

For example, a probabilistic algorithm is an algorithm that gives the right answer most of the time and a randomly bad answer in the remaining cases.

In our context, the only kind of calculation error that is relevant is one that is exemplified by the square root.

Suppose you ask me to calculate for you the square root of 2.

My answer could be one of the following:

- 1.41 (off the top of my head)
- 1.414213562 (Excel with default settings)
- 1.4142135623730950488016887242097 (MS Windows Calculator)

Each of these contains a *calculation error*.

The decimal representation of the true result has infinitely many digits (with no periodic repetition).

Any communicable result in decimal form is an approximation.

However, this is not really a problem, for two reasons:

1. It is always known what the error is.

   A universally accepted convention is that a given result in decimal form is accurate up to the number of displayed digits.

   For example, 1.41 really means 1.41 $\pm$0.005.

2. The square of a number can be calculated to any desired precision.

A calculation with property 2. above is called an *arbitrary precision calculation*.

For all practical purposes, it is as good as an exact calculation.

Remark:

The vast majority of all scientific software programs, including MS Excel and StyleADVISOR – as well as all of its competitors – limit the precision of their calculations for the sake of better performance. However, that is a mere technicality.

What about calculation error in returns-based style analysis?

Returns-based style analysis as set forth by William F. Sharpe is an arbitrary precision calculation.

It can be shown that for any given (non-degenerate) input, RBSA has a unique solution, and that solution can be calculated with arbitrary precision.

Second Source of Uncertainty: ***Error Propagation (Propagation of Uncertainty)***

Suppose you ask me to calculate for you the square of the number 2.00.

You are indicating that this number comes from a source where it has been subject to rounding: 2.00 really means 2.00 ±0.005.

Now you can really ask two questions:

- What is the square of my number?
- What is the uncertainty in the result, based on the uncertainty in the input? In other words, how did the input error propagate through the calculation?

One way to give a combined answer is: The square is 4 +0.020025 –0.019975.

Relative error of input and output:

Input: .01 / 2 = .5%

Output: .04 / 4 = 1%

The relative error has doubled!

Does this issue of magnifying the error by error propagation potentially affect RBSA?

It potentially affects *any* calculation:

Whenever there is uncertainty in the input of a calculation (and there almost always is), the question arises, what is the resulting uncertainty in the output?

Remark on Terminology:

Investigating error propagation can be viewed as a form of statistical significance test:

In the presence of input error, the output of the calculation is a point estimator of the true output.

Understanding the error propagation gives us the distribution of the point estimates, which allows us to determine the significance of the estimates.

When preparing this talk, I felt that emphasizing the error propagation point of view helped clarifying some issues.

How To Measure Uncertainty

This is a complex issue ([1]). A standard way of dealing with it is as follows.

1. One assumes that the uncertainty in the input numbers is *Gaussian* ([2]):

   Input is given by a mean and a standard deviation.

   The probability of the true input being in any given interval is given by a normal distribution with that mean and standard deviation.

2. The propagation of the uncertainty is measured by *confidence intervals*.

   For example, suppose your input is “mean = 2, stddev = .1”

   and you want the square and the 95% confidence interval.

   The square is 4, of course, and the confidence interval is determined by the following steps.

   a) The interval [1.804, 2.196] captures the true input with a 95% probability.

   b) When calculating the square, this interval gets mapped exactly (1-1) to the interval [3.254416, 4.822416].

   c) Hence, this interval has a 95% probability of capturing the true output. *It is the 95% confidence interval.*

Confidence Intervals for William F. Sharpe's RBSA Calculation

There is no known method to calculate them.

The literature is not entirely silent on the subject (more about that later), but the fact remains: We cannot really, exactly calculate confidence intervals for RBSA.

The fact that we do not have confidence intervals does not mean that we are left in the dark when it comes to RBSA error propagation.

When there is no closed-form calculation for a statistical result such as confidence intervals, one can always resort to *Monte Carlo simulation*.

Monte Carlo simulation turns out to be an excellent tool for our purpose.

To measure RBSA error propagation via Monte Carlo simulation:

1. Choose a set of inputs for a style analysis.
2. Repeat the style analysis a large number of times. In the process, add to the input random noise of the same nature and magnitude as the error whose propagation you are analyzing ([3]).
3. Observe and statistically analyze the set of outputs thus obtained.

What Monte Carlo Simulation Can Teach Us about Confidence in RBSA

We saw: The only source of uncertainty in RBSA is error propagation.

Therefore, any investigation of uncertainty and confidence must first determine and clearly state what the perceived nature and magnitude of the error, or uncertainty, in the input is.

The most obvious and undeniable kind of uncertainty is *rounding of the input numbers*.

Data providers typically provide returns in fractional form with six digits after the decimal point. Hence, there is an error in the form of a discrete interval of width 10^{-6}.

Running Monte Carlo simulations of RBSA with added random noise sampled from a discrete interval of width 10^{-6} shows: the effect of this error on the output is so small that it is almost always immeasurable with the precision afforded by MS Excel or Zephyr StyleADVISOR.

The propagation of the rounding error in the input data is not a problem when performing RBSA.

There is another kind of potential uncertainty in the input data to RBSA, namely, noise caused by active management decisions.

Ideally, active management just adds alpha (which is not a source of error for RBSA).

In practice, active management may well add noise (unsystematic variance) to the manager's return series.

From the point of view of style analysis, this noise constitutes input error.

As for the nature of this error, it is a reasonable assumption that it is Gaussian.

But what about its magnitude?

The only place to look really is the variance of the excess return (tracking error) of the manager vs. the style benchmark.

Recall: The style benchmark is the passive investment that represents the manager's style as determined by the RBSA.

The tracking error of the manager vs. the style benchmark is – in William F. Sharpe's words – the portion of the manager's variance that RBSA could not explain.

There are two extreme schools of thought regarding the proper interpretation of the tracking error.

School 1 says: So, the tracking error is the part of the manager's variance that RBSA could not explain? Then clearly, it must be our input error ([4])!

School 2 says: So, the tracking error is the part of the manager's variance that RBSA could not explain? Then clearly, it must represent some portion of the manager's style that was not captured by our choice of asset classes ([5])!

My personal belief – shored up by Monte Carlo simulation experiments – is that the truth is in the middle.

Monte Carlo simulation demonstrates very convincingly that random noise added to the manager series does show up in the tracking error.

Therefore, if you believe in the presence of noise other than rounding error in your manager return series, then that will show up in the tracking error (School 1).

On the other hand, everyone who has used StyleADVISOR more than just casually will agree that a high tracking error (that is, a low R^2) can also be the result of a style component that is not captured by the asset classes (School 2).

School 1: Tracking error = input noise (probably overestimates the input noise)

School 2: Tracking error = unexplained style (probably underestimates the input noise)

More likely: Tracking error = unexplained style + input noise

How do we separate the two?

I don't think it can be done mathematically.

If you are looking at a style analysis with a high tracking error (low R^2), then it is up to you as an investment professional to find out why that is the case and how you can possibly improve the situation (change your asset classes, look at rolling windows to see how stable the analysis is over time, …)

Generally speaking, the best advice you can get is, bring your tracking error down. When the tracking error is low, then regardless of whether it's input noise or unexplained style, you won't have much of either ([6]).

In addition to this general goal of achieving a high R^2, William F. Sharpe has recommended to focus on getting a high rolling-window out-of-sample R^2.

This amounts to maximizing the predictive power of your style analysis as opposed to the explanatory power.

The Work of Lobosco and DiBartolomeo

There is a 1997 paper entitled “Approximating the Confidence Intervals for Sharpe Style Weights” ([7]) by A. Lobosco and D. DiBartolomeo. This is a brilliant paper that is full of important insights. It should be required reading for everyone interested in the mathematics of RBSA.

Nonetheless, I will argue that the paper’s main result, interesting as it is in terms of research, is of limited value to the practitioner ([8]).

1. The main result of the paper is based on the “tracking error = input error” school of thought (see [4]). Some people do not agree with this.

2. The main result of the paper provides an approximation to the standard deviation of the style weights (that is, the output of RBSA).
In order to draw conclusions on confidence intervals from this, one would have to assume that the style weights are normally – or at least symmetrically – distributed. Monte Carlo simulation shows that the style weights can be asymmetrically distributed.

3. As indicated by the title of the paper, its main result is an approximation. Although the authors discuss the origin and the nature of the error extensively, there is no error estimate. In other words, the confidence calculation has a confidence issue.

Where Does All This Leave Us?

If you really always insist on exact knowledge of your input error and its propagation through any calculation that you use, then you must dismiss RBSA.

However, that may not be a very wise engineering decision: you are dismissing a useful tool because its numbers don’t meet the highest possible standards.

You are probably be putting yourself at a competitive disadvantage.

(Getting on my soapbox…)

> Good engineering is not all about numbers. Good engineering is about making good decisions based on good judgment. Numbers are just one of many things that go into that decision making process. Numbers receive their meaning in the context of their use and by the judgment of the person who uses them.

(Getting off my soapbox…)

Notes and References

[1] This is in fact an issue that is addressed by the National Institute of Standards and Technology (NIST). See http://physics.nist.gov/cuu/Uncertainty

[2] This assumption can of course be a rather gross simplification. However, since *some* assumption must be made, the Gaussian assumption is clearly the best.

[3] Depending on your point of view, it may be more appropriate to start with the style benchmark of the analysis and then perform RBSA on that, applying noise in the process. See e.g. the Monte Carlo simulation on p. 82 of [6].

[4] This point of view is taken e.g. by [6]. To see how there, the tracking error is interpreted as input error, you can look for example at the way the Monte Carlo simulation is set up on p. 82. This is also evident from the final result (4) stated on p. 82: The tracking error appears in the numerator of the formula for the style weight's standard deviation; hence, the latter grows proportionally with the tracking error.

[5] This point of view is expressed on p. 3 and on p. 7/8 of

T. Idzorek and F. Bertsch, The Uncertainty of Style Weights, preliminary manuscript

[6] Another thing to watch out for is colinearity of the asset classes. High colinearity of the asset classes – although not itself a source of error – will exacerbate the error propagation.

[7] A. Lobosco and D. DiBartolomeo, Approximating the Confidence Intervals for Sharpe Style Weights, Financial Analysts Journal, July/August 1997, p 80–85

[8] Please note that my arguments do not constitute criticism of Lobosco's and DiBartolomeo's work. Their research is brilliant and perfectly sound.