

**Arizona State University**

---

**From the Selected Works of Joseph M Hilbe**

---

August, 2007

# The Black Swan: Praise and Criticism

Peter H Westfall

Joseph M Hilbe, *Arizona State University*



Available at: [https://works.bepress.com/joseph\\_hilbe/9/](https://works.bepress.com/joseph_hilbe/9/)

# The Black Swan: Praise and Criticism

Peter H. WESTFALL and Joseph M. HILBE

It is unusual for the Editors of *The American Statistician* to contribute regular articles to the journal. However, this is a special case. If you are a statistician, or statistically minded, Nassim Nicholas Taleb really wants to get under your skin. In his book, *The Black Swan: The Impact of the Highly Improbable*, (henceforth *TBS*), he says:

“‘social scientists’ . . . for over a century have operated under the false belief that their tools could measure uncertainty.”

“ . . . certain professionals [including especially statisticians] . . . do not know more about their subject matter than the general population, but they are much better at . . . smoking you with mathematical models.”

“ . . . the bell curve . . . is [the] . . . Great Intellectual Fraud.”

“The beast in this book is not just the bell curve and the self-deceiving statistician . . . ”

He also makes several other comments disparaging of statisticians and statistically minded people he calls “quants.” His often-unfounded and sometimes outrageous criticisms require a response from the statistical community.

Despite the numerous irritating comments peppered throughout, the book is quite engaging, well-written, and tells an interesting story. Like Taleb’s first book, *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets* (2001), *TBS* is likely to be popular with the general public. As the quotes demonstrate, however, *TBS* is also likely to influence the way that statistics and statisticians are perceived by the general public.

First some praise: We highly recommend *TBS* for statisticians, students of randomness and statistics, and “quants” in general. Our recommendation is not necessarily for its correctness, accuracy, and appropriate placement of statistics within the scientific enterprise, but instead for its stimulating ideas. By analogy, W. Edwards Deming often irritated statisticians with admonishments not to use probability models for anything but finite population sampling (Deming 1986). Even if you considered Deming’s advice misplaced, you have to acknowledge that he made you think more carefully about the use of randomness in models for process data (“analytical studies” in Deming’s terminology); coming to grips with such issues is always preferable to rote memorization and teaching of statistical formulas.

Taleb’s *The Black Swan* will also make you consider fundamental questions about the nature of randomness, and about the models we use for randomness. The central thesis of Taleb’s book

is that outliers, not part of the bell curve, are too often ignored by social scientists, statisticians in particular, and that these outliers have enormous importance for all matters of scientific, worldly, economic, and societal importance.

Paradoxically, for all Taleb’s disparaging comments about statisticians, he will find few arguments among statisticians about this central thesis. Of course outliers are important and should not be ignored or “swept under the rug.” Finney (2006) provided a recent discussion of the problem, but the statistical literature is rife with tools for outlier diagnosis and modeling. Taleb reminds us we can perhaps do more with outliers, a message that perhaps is under-emphasized in our statistics classes, or perhaps under-appreciated by our students. But again, there will be few arguments from statisticians. In addition to Taleb’s admonishments, we add the following anecdotes:

- First, a pharmaceutical company statistician (working in toxicology) once remarked, “All this discussion of deleting the outliers is completely backwards. In my work, I usually throw away all the good data, and just analyze the outliers.”

- Second, the famous statistician Fred Wood [of Daniel and Wood (1971) fame] reportedly discussed outliers at length in his presentations, claiming to have developed a number of patents resulting from outliers.

Statisticians will also find agreement with Taleb’s admonishment to scientists concerning post-hoc rationalization: one of “Young’s False Positive Rules” is that well-intentioned (but perhaps misguided) scientists can “explain” any anomalous result (Westfall and Young 1993, p. 7). Also, the central tenet of Taleb’s previous book admonishing people not to be “Fooled by Randomness” is also the central concern of statistical inference, since margins of error accounting for randomness are mandatory. Taleb would question the assumptions used to derive such margins of error, but it is standard practice for most statisticians to question these assumptions as well.

It seems, therefore, that Taleb has fallen victim to his own curse: having observed a few statisticians with an outlier-avoidance mindset (let us call these statisticians “White Swans”), he then violates Hume’s anti-inductive admonishment and assumes that all statisticians are “White Swans.” What makes Taleb’s error particularly egregious is that his sample of statisticians from which he makes such generalizations is both small and systematically biased!

There is much to criticize in Taleb’s book besides the disparaging (and unfounded) comments about statisticians. For example:

- He devotes much space to less-important topics (e.g., Mandelbrot’s special distribution) and he devotes little space to more important ones (e.g., heavy-tailed distributions in general).

- He disparages many deductive models proposed by various scientists (especially Nobel Prize winners), while espousing

---

Peter H. Westfall is Editor, *The American Statistician*, and Professor of Statistics, Area of Information Systems and Quantitative Sciences, Texas Tech University, Lubbock, TX 79409 (E-mail: [peter.westfall@ttu.edu](mailto:peter.westfall@ttu.edu)). Joseph M. Hilbe is Software Reviews Editor, *The American Statistician*, and has the dual position of Emeritus Professor, University of Hawaii, and Adjunct Professor of Statistics, School of Social and Family Dynamics, Arizona State University, Tempe, AZ (E-mail: [Hilbe@asu.edu](mailto:Hilbe@asu.edu) and [jhilbe@aol.com](mailto:jhilbe@aol.com)).

some less well-known contributions of Benoit Mandelbrot (also deductive).

- He apparently favors empirical induction in his personal endeavors while simultaneously espousing the anti-inductive stances of Hume and Popper.

- He seems to have missed the entire point of Bayesian statistics, whose aim is to integrate empiricism and deduction seamlessly [Jeffreys (1961) would be a good place to start].

- He claims the normal curve is a fiction produced by statisticians' minds, then he proceeds to describe a naturalistic process, independent of the mind of any statistician, that gives rise to the normal curve.

Taleb promotes “Black Swan” thinking for decision-making, which necessarily entails forecasting states of nature, while at the same time stating that nature cannot be forecast. (If concern for Black Swans is so great as to outweigh the mundane, a silly corollary is that we should avoid plane flights!) We are reminded of former United States Secretary of Defense Donald Rumsfeld's quotes on forecasting and uncertainty:

- “I would not say that the future is necessarily less predictable than the past. I think the past was not predictable when it started.”

- “There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. These are things we don't know we don't know.”

Like Rumsfeld, it seems that Taleb wants it both ways: to call himself an expert in the field, but also to leave himself a clear escape route if pinned down. Taleb calls himself a probabilist, but wants to avoid using formal probabilistic models, instead invoking the inductive nihilism of Hume and Popper. Paradoxically, the anti-inductive stance taken by Hume that justifies The Black Swan also justifies The White Swan: if forecasting is impossible, with no rational basis, then we may as well assume that the future will be populated with only beautiful white swans!

To make his Black Swan philosophy rational, certain probabilistic assumptions are unavoidable. If Taleb is truly a probabilist and can make logical arguments concerning his philosophy, any number of journals, including this one, would be happy to publish his work. Here is a start: suppose that a relevant time series is

$$\mathbf{X} = \{\dots, X_{-1}, X_0, X_1, X_2, \dots\},$$

any measurements (possibly vector-valued) of interest regarding finance, science, or life in general. Taleb seems to eschew model assumptions so, at the outset, let us assume nothing about the dependence structure or about the distributional forms, assuming only

$$X_t \sim F_t,$$

with  $F_t$  an arbitrary distribution, and with arbitrary dependence among the  $\{X_t\}$ . Since virtually nothing is assumed here, one really can't argue with the model. Taleb likes this model! However, he also prefers a “Black Swan strategy,” where the optimal stance is to anticipate the occurrence of Black Swans. If this

strategy is somehow preferred, what is the logic? Two questions arise immediately:

1. What regularity conditions on  $\mathbf{X}$  are sufficient to ensure that the Black Swan strategy is the preferred strategy? Clearly, if the outlier probability is sufficiently small, then there is no point in worrying about it. How large does the outlier probability have to be before we can exploit it?

2. What assurance can Taleb provide to readers (and possibly to investors) that such regularity conditions are satisfied? Black Swans are justified empirically (they have occurred), yet a strict interpretation of Hume would force us to acknowledge that there might be no future Black Swans. Here is where the art and science of forecasting comes most directly to the fore: Taleb is right in saying that we cannot predict Black Swans; however, sensible judgmental predictions (perhaps Bayesian?), aided by empirical data, are essential here. So, what are those predictions? Please be specific! After all, without such predictions, there can be no formal justification for any strategy based on Black Swan thinking.

It is doubtful that we will hear specific answers to these questions, as Taleb apparently prefers not to be pinned down. It is too bad, because answers to these questions would make for very interesting reading in *The American Statistician* or similar journals.

Despite our criticisms, we hope *The Black Swan* is widely read and successful in the popular press. At the foundation, Taleb is on our side. We statisticians have fought an uphill battle—forever—to inject probabilistic thinking into daily enterprise, whether science, business, government, or driving one's car to work. The forces we counter are deeply ingrained deterministic mindsets, for example, from the biologist who says “If you have to use statistics, then you don't really know what you are doing,” to the computer scientist who thinks probabilistic data-generating processes have no relevance to the algorithmic analysis of datasets. Students and teachers alike who read *The Black Swan* will pay more attention to both the normal distribution and to outliers in their statistics classes. If the end result of the publication of *The Black Swan* is that the general public becomes more aware of the importance of randomness in all areas of human, societal, government, and scientific enterprise, then Taleb will have done the statistical profession a great service.

## REFERENCES

- Daniel, C., and Wood, F.S. (1980), *Fitting Equations to Data: Computer Analysis of Multifactor Data* (2nd ed.), New York: Wiley.
- Deming, W. E. (1986), *Out of the Crisis*, Cambridge, MA: MIT Press.
- Finney, D. J. (2006), “Calibration Guidelines Challenge Outlier Practices,” *The American Statistician*, 60, 309–314.
- Jeffreys, H. (1961), *Theory of Probability*, Oxford: Oxford University Press.
- Taleb, N. N. (2001), *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets*, London: TEXRE Publishing Limited.
- (2007), *The Black Swan: The Impact of the Highly Improbable*, New York: Random House.
- Westfall, P. H., and Young, S.S. (1993), *Resampling-Based Multiple Testing: Examples and Methods for P Value Adjustment*, New York: Wiley.