Statistical Science
2016, Vol. 31, No. 4, 555–557
DOI: 10.1214/16-STS593
Main article DOI: 10.1214/16-STS565

Rejoinder: Concert Unlikely, "Jugalbandi" Perhaps

Nozer D. Singpurwalla

Abstract. This rejoinder to the discussants of Filtering and Tracking Survival Propensity begins with a brief history of the statistical aspects of reliability and its impact on survival analysis and responds to the several issues raised by the discussants, some of which are conceptual and some pragmatic.

Key words and phrases: Military standards, Weibull distribution, Wash 1400.

The current version of this paper would not have come into being without the endless patience and selfless encouragement of the Editor and the Associate Editor. They orchestrated a series of demanding, but insightful, referee reports followed up by reactions from four discussants, each adding a new dimension to the paper. The author is indebted to each discussant for raising several issues, almost all of them brutally piercing, and most of them germane. To ensure the paper's completeness, I am obliged to respond to these, and will do so, but in the context of a historical development. The hope here is that a historical perspective could help diffuse some of the concerns expressed. But before doing this, it is appropriate to start with the cautionary comment that history is written from the vantage point of the writer, and thus what follows is a manifestation of this thesis.

Engineering reliability, and its offshoot survival analysis, have garnered a track record of success starting with the work of von Braun and his rocket scientists. But the beginning of its statistical dimension can be attributed to Epstein and Sobel's (1953, 1954) papers on life testing for the exponential under different censoring protocols. Whereas relevance of the Weibull for describing lifetime data was advocated by Weibull (1951), its endorsement as a general distribution for use in reliability came about because of a series of papers by John H. K. Kao, two of which are (Kao, 1956, 1959). The U.S. Department of Defense found value in all of the above works, and codified them as Military Standards. At about this time, Kaplan and Meier (1958) published their paper on survival function estimation, with a focus on medical applications. This paper was a landmark event in survival analysis (followed by the more encompassing paper by David Cox, 1972). Whereas Kaplan–Meier make reference to Epstein and Sobel, they also cite Davis (1952) on the analysis of failure data, suggesting a time frame which precedes 1953.

It therefore seems appropriate to claim that a paradigm for statistical inference in reliability and survival analysis was formulated by the above individuals, and was based on the state of the art of their times. Specifically, reliability was conceptualized as an objective, frequency based probability, that is fixed at the time of assessment, over all future time. This viewpoint was also embraced in the books by Barlow and Proschan (1965, 1975), on the mathematical theory of reliability, and by Gnedenko, Belyaev and Soloview (1965), in their book on reliability; this book was awarded the Soviet State Prize.

Whereas excursions from this paradigm have appeared in the literature, its essential dominance prevails. Regarding excursions, the notion that reliability could be dynamic was articulated by Arjas (1981). Also, papers invoking the Bayesian argument for one of a kind entity, like nuclear reactors, began to appear in the mid 1970s (cf. Wash 1400, 1975). However, these Bayesian works lacked a philosophical compass, and the goal of this paper is to propose one. The genesis of the compass can be traced to Dennis Lindley's visit to Richard Barlow at Berkley. Lindley emphasized the importance of thinking like a Bayesian, over the

Nozer D. Singpurwalla is Chair Professor, Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong, China (e-mail: nsingpur@um.cityu.edu.hk).

mechanical process of assigning arbitrary priors and turning the crank. In the sequel, it became transparent that a defensible foundation for reliability (and survival analysis) can be had if reliability were interpreted as a chance, not a probability. This viewpoint would enable one to introduce a dynamic into the process and also consider one of a kind scenarios. Furthermore, it would enable one to endow a personal probability to an unknown chance, and avoid the conundrum of endowing a probability to a probability.

Hopefully, the above preamble provides a context for a reply to the discussants who fall in two categories. The first are those who seem to be in overall agreement with the thesis that reliability is a dynamic entity, or are at least partial to it, but nonetheless seem to say "what is new here? Laplace has said it a long time ago, and Dawid has said it more recently". First, an apology to Phil for not mentioning Dawid (2004), though I had seen his paper, albeit in a different context. The second category of discussants wants to know how a change in the current view will help their day to day activities like tracking software, administering drugs or changing a bicycle tyre.

Frank Coolen's Section 2, captures the essence of my paper, and adds to it, by highlighting the work done by him and his co-workers. In the sequel, he alludes to the matter of partial exchangeability, raises the spectre of multiple propensities, and flirts with the notion of a probabilistic structure function. All these are ideas worth pursuing, but his kiss of death appears in his last sentence, questioning benefit. My response to him is the same I give to Professor Jane Hutton's discussion which begins with the salvo "What does 'propensity' add; as yet I have not found benefit". My answer is simple: clarity of thought! To her question "does it matter in practice", my answer is yes, especially when it comes to changing bicycle tyres. This is because each rider has a unique riding style, and a unique route. Furthermore, the bicycle experiences deterioration due to wear and tear. The tyre replacement phenomenon therefore needs to be modelled dynamically, and via a Bayesian approach. Were the bicycle's reliability be interpreted as a probability, then a prior probability on it would be tantamount to placing a bet on a bet. There are other matters in Professor Hutton's discussion that warrant comment. First, to Kolmogorov, it is the very notion of probability that is an undefined primitive, not merely the axioms of probability. Second, propensity has not been taken as an interpretation of probability; much verbiage is devoted to make the case that it is not. Propensity is seen as a primitive which manifests as a frequency, or put differently, the cause of an observed frequency is propensity. Finally, there is a difference between notional infinite repetitions used to define probability, versus a theorem being valid for any finite sequence that is a sub-set of an infinite sequence. Regarding the motto of remaining silent in the context of entirely unique entities, whereas silence is golden, it may be necessary to yell when a gun is held to your head, and you are forced to make a decision.

Elja Arjas, who has made substantial, but subtle, contributions to reliability and survival analysis, introduces the metaphor of wearing two hats, presumably, one in spring and the other in summer. In so doing, he encapsulates the essence of the idea of sequentially linking propensity and probability more transparently than my paper does. I am in agreement with much of the rest of what he has to say, including the feature that θ remains in the intellectual world, though I may have slipped by calling θ objective, and becoming a victim of *mind projection fallacy*. Arjas' crucially important remark pertaining to his *duck test*, was also raised by Jim Berger as an off-hand cocktail hour conversation in Varanasi, India. The response here is that one may use any approach one feels appropriate to reflect ones sense of the strength of propensity, be it a scientific theory, or a probability based system of rules; I have chosen the latter.

Glen Shafer's encompassing knowledge on the foundations of probability, as evidenced by his several publications on the topic, Shafer (1990) being an example I neglected to cite and for which I owe him an apology, helps cast the paper in a broader philosophical and historical perspective. Shafer's insertion of his recent works on defensive forecasting, game-theoretic generalizations of Kolmogorov's framework and his discourse on the works of Ville (1939) and Bienvenu, Shafer and Shen (2009), endow a uniquely different dimension to the subject of this paper that could have an impact on its future development. Whereas Shafer is eminently known for his work on belief functions, my first encounter with him was in the context of reliability when along with Melvin Springer he wrote a paper alluding to a use of belief functions in reliability. Currently, there is a plethora of papers using belief functions in the context of prognosis, a latest one being in the context of rechargeable batteries (He et al., 2011).

Whereas Shafer endorses a main thesis of the paper that an objective probability is, and has been dynamic, he does not see a need for reliance on exchangeability. I am tempted to agree, but find exchangeability a convenient stepping stone. Shafer also claims that efforts at reconciliation of the different interpretations of probability have not gained much traction. This sentiment is also apparent in Arjas' kiss of death, which arrives in the guise of a dissonance in western classical music. It is unlikely that Popper and de Finetti would have appeared on the same concert stage. But if Popper were a guitar, and de Finetti a sitar, then there would be an uplifting *jugalbandi* (entwined twins) of *saval–javab* (question–answer) in a concert of Hindustani classical music with the guitar playing the saval and the sitar playing the javab. I rest my case!

REFERENCES

ARJAS, E. (1981). The failure and hazard processes in multivariate reliability systems. *Math. Oper. Res.* **6** 551–562. MR0703096

BARLOW, R. E. and PROSCHAN, F. (1965). *Mathematical Theory of Reliability*. Wiley, New York. MR0195566

BARLOW, R. E. and PROSCHAN, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt, Rinehart and Winston, New York. MR0438625

BIENVENU, L., SHAFER, G. and SHEN, A. (2009). On the history of martingales in the study of randomness. *J. Électron. Hist. Probab. Stat.* **5** 1–40. MR2520666

COX, D. R. (1972). Regression models and life-tables. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **34** 187–220. MR0341758

DAVIS, D. J. (1952). An analysis of some failure data. *J. Amer. Statist. Assoc.* **47** 113–150.

DAWID, A. P. (2004). Probability, causality and the empirical world: A Bayes–de Finetti–Popper–Borel synthesis. *Statist. Sci.* **19** 44–57. MR2082146

EPSTEIN, B. and SOBEL, M. (1953). Life testing. *J. Amer. Statist. Assoc.* **48** 486–502. MR0056898

EPSTEIN, B. and SOBEL, M. (1954). Some theorems relevant to life testing from an exponential distribution. *Ann. Math. Stat.* **25** 373–381. MR0061339

GNEDENKO, B. V., BELYAEV, Y. K. and SOLOVIEW, A. D. (1965) *Mathematical Methods in Reliability Theory*. Nauka, Moscow.

HE, W., WILLIARD, N., OSTERMAN, M. and PECHT, M. (2011). Pronostics of lithium-ion batteries on Dempster–Shafer theory and the Bayesian Monte Carlo method. *J. Power Sources* **196** 10314–10321.

KAO, J. H. K. (1956). A new life quality measure for electron tubes. *IRE Trans. Reliab. Qual. Control* **7** 1–11.

KAO, J. H. K. (1959). A graphical estimation of mixed Weibull parameters in life testing of electronic tubes. *Technometrics* **1** 389–407.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481. MR0093867

SHAFER, G. (1990). The unity and diversity of probability. *Statist. Sci.* **5** 435–462. With comments and a rejoinder by the author. MR1092984

VILLE, J. (1939). *Étude Critique de la Notion de Collectif*. Gauthier-Villars, Paris.

WASH 1400 (1975). Reactor Safety Study. NUREG-75/014. U.S. Nuclear Regulatory Commission, Washinton, D.C.

WEIBULL, W. (1951). A statistical distribution functions of wide applicabiity. *J. Appl. Mech.* 293–297.