Available online at www.sciencedirect.com
ScienceDirect
Statistics & Probability Letters 77 (2007) 1354–1361
STATISTICS & PROBABILITY LETTERS
www.elsevier.com/locate/stapro

Betting on residual life: The caveats of conditioning

Nozer D. Singpurwalla*

The George Washington University, Washington, DC 20052, USA

Available online 24 March 2007

Abstract

Assessing conditionals based on any specified probability model is straightforward and unique when the conditioning event is in the subjunctive mood; that is, supposing that the conditioning event were to occur. The matter becomes problematic, however, when the conditioning event actually does occur as observed data, and thus becomes a reality. We illustrate this point by considering a commonly occurring scenario in the actuarial sciences, engineering reliability, survival analysis, and in general, any type of an activity that involves filtering. We argue that there could be more than one way to bet on residual life. Our message is that it is the likelihood—not Bayes' Law—which is the tail that wags the dog!

This paper should appeal to both probabilists and statisticians who are interested in foundational issues. It has been written to honour Richard Johnson whose Editorship of *Statistics and Probability Letters* has provided a platform for dialogue between probabilists, statisticians, and those who strive to be both.

Keywords: Actuarial science; Conditionalization principle; Double slit experiment; Filtering; Forecasting; Likelihood; Reliability; Survival analysis

1. Introduction

In the process of using marker data to assess the lifetime of an item experiencing failure due to ageing, we were confronted by a dilemma that sneaked upon us as a matter of course (see Singpurwalla, 2006a). It turns out that the scenario leading to the dilemma is quite common and can arise when addressing practical issues of conditioning in the actuarial, the engineering, and the biomedical sciences. Stripped to its essentials, the scenario goes as follows.

Suppose that an item's lifetime X is judged to have a distribution function $G(x) = P(X \leqslant x)$, and a survival function $\overline{G}(x) \stackrel{\text{def}}{=} 1 - G(x) = P(X > x)$. We suppose that lifetime can be continuously monitored so that $x \geqslant 0$. Were this item supposed to survive until x, its *residual* (or *remaining*) *lifetime* will be $X - x$. We are required to make statements of uncertainty about $(X - x)$, so that actuarial, engineering, or medical decisions about the item can be made. That is, we are required to specify $P(X - x > u | X > x)$, for all $u > 0$. Our interpretation of probability is de Finnetian (see de Finetti, 1937), in the sense that probability reflects one's disposition to a two-sided bet. Thus, probability assessments can be seen as a device for hedging our bets on the item's survival, or some other unknown quantity of interest, such as parameters in probability models.

*Tel.: +1 202 994 7515; fax: +1 202 994 7508.
E-mail address: nozer@gwu.edu.

0167-7152/$ - see front matter
doi:10.1016/j.spl.2007.03.021

A solution to the problem posed is elementary and unique, given a distribution function G. Specifically, for any $u>0$

$$P(X-x>u|X>x)=P(X>x+u|X>x)=\frac{P(X>x+u)}{P(X>x)}=\frac{\overline{G}(x+u)}{\overline{G}(x)}. \tag{1.1}$$

Suppose now, that instead of the subjunctive, "*were* the item to survive until x", we are told that the item actually *did* survive to x. That is, the event $(X>x)$ is no more an uncertain event; $(X>x)$ has now become observed data. What then would our assessment of the uncertainty about the residual life $(X-x)$ be? In other words, how would we bet on the event $(X-x>u)$, for $u>0$? Would it continue to be $\overline{G}(x+u)/\overline{G}(x)$, or could it be something else? If the latter, would the number to bet be unique? For a discussion of these and related questions, one may visit Freedman and Purves (1969). A more recent discourse on the different kinds of conditional beliefs is in Joyce (1999, Chapter 6).

Intuitively, it seems that there ought to be some distinction between looking at $(X>x)$ as a possibility, versus looking at it as a fact that is revealed as data. Thus, $\overline{G}(x+u)/\overline{G}(x)$ need not be the correct answer. Yet many individuals when faced with this problem would simply mimic the steps leading to Eq. (1.1) and continue to declare $\overline{G}(x+u)/\overline{G}(x)$ as their answer. In doing so they do not appear to be making a distinction between $(X>x)$ as a supposition versus a reality. Alternatively put, they may be failing to recognize the connotation that in a conditional probability statement, the word "given" does not indicate a fact; rather it indicates a supposition that the conditioning event is true. Thus, are those who declare $\overline{G}(x+u)/\overline{G}(x)$ as their answer—irrespective of the character of the conditioning event—in error, or is there a rationale for their answer?

We claim that the rationale cannot completely be within the calculus of probability, because the notion of probability—at least from a subjectivistic point of view—is germane only when the disposition of *all* events in question is unknown. Thus, for example, it may not make sense to say that the probability that a coin with heads on both faces when flipped will land heads, is one. This is because the disposition of the outcome is known before the flip. Consequently, a two-sided bet on the outcome heads has to be \$1, which will be exchanged for a \$1 when the coin lands heads, which it will. The two-sided bet of \$1 is thus meaningless. The rationale therefore must come from concepts in statistics wherein the notion of a likelihood plays a signal role. By all accounts the notion of a likelihood appears to be alien to probability theory.

In what follows we point out that there are both philosophical and technical arguments which support $\overline{G}(x+u)/\overline{G}(x)$ as an answer, but that this answer is one among other possible answers. This is the main point of this article. Arguments about conditioning are common among philosophers of science. That such arguments could also be relevant to reliability, survival analysis, filtering, and forecasting seems to not have been recognized.

2. Answer(s) to the question

2.1. Reassessment and the principle of conditionalization

Some individuals when faced with the matter of assessing $P(X-x>u)$ with $(X>x)$ as observed data, may chose to re-assess all probabilities treating the factual event $(X>x)$ as a part of background history; that is, they would start from ground zero, even if the observed $(X>x)$ is not a surprise. Diaconis and Zabell (1982) label a process like this, *complete reassessment*; however, the driving premise considered by the above authors is different from the one we are discussing here, in the sense that the observed event is considered to be a surprise. In a re-assessment one essentially starts all over again from scratch and possibly even rejects G as the underlying probability model. The answer that one obtains may therefore not necessarily be $\overline{G}(x+u)/\overline{G}(x)$. Reassessment is a perfectly legitimate step; its main danger is the risk of incoherence (i.e. a lack of consistency). We therefore do not pursue here this line of reasoning and do not advocate reassessment as a strategy.

To ensure coherence one may proceed formally by invoking Bayes' Law as an inferential mechanism, using $(X>x)$ as data. These are two directions from which this can be approached, one general, the other specific. These we describe in Sections 2.2 and 2.3, respectively, wherein we point out that there need not be a unique answer to the question posed, and that under a certain assumption, $\overline{G}(x+u)/\overline{G}(x)$ will indeed be one of several possible answers.

But there is another, more philosophical, argument that supports $\overline{G}(x+u)/\overline{G}(x)$ as a correct answer. This argument, known as the *Principle of conditionalization* (cf. Howson and Urbach, 1989, p. 68), proceeds as follows:

Prior to observing $(X>x)$ as factual data, we had declared that $\overline{G}(x+u)/\overline{G}(x)$ would represent our bet (or personal probability) on the event $(X-x>u)$, for some $u>0$, were the event $(X>x)$ turns out to be a fact. Now that $(X>x)$ has revealed itself as being actually true, we shall act as we had declared, and thus $\overline{G}(x+u)/\overline{G}(x)$ would continue to be our bet. As suggested by a reviewer, another way to articulate the principle of conditionalization is, to assert that "if I say I am going to do something, I will do it".

Those who subscribe to a complete reassessment by starting all over from scratch, may reject the principle of conditionalization on grounds that the *actual* occurrence of the event $(X>x)$ has changed their psychological disposition so dramatically from their disposition under the supposition that $(X>x)$, that they can no more subscribe to G as their model of uncertainty. They then seek an alternate to G, say H as a model for assessing $(X-x)$. This point was made by Ramsey (1931) (cf. Diaconis and Zabell, 1982) who stated that

> [The degree of belief in p given q] is not the same as the degree to which [a subject] would believe p, if he believed q for certain; for knowledge of q might for psychological reasons profoundly alter his whole system of beliefs.

Diaconis and Zabell (1982) also cite other, more modern, references that mention the above issue; these are Hacking (1967), de Finetti (1972, p. 150; 1975, p. 203), Teller (1976), and Freedman and Purves (1969).

Additionally, there also happens to be empirical evidence from quantum mechanics that rejects the conditionalization principle vis-a-vis the "double slit experiment". This experiment has now become a classic thought experiment for its clarity in expressing the central puzzles of quantum mechanics. In its original version, performed by the English scientist Thomas Young sometime around 1805, the experiment consisted of letting light diffract through two slits producing fringes on a screen. The goal of the experiment was to resolve the question as to whether light is composed of particles or waves. The current versions of the experiment are performed with electrons instead of light (cf. Jonsson, 1974). Such experiments have shown that the probability (as assessed via the relative frequency) of some event, say B, when an event A always occurs is not equal to the conditional probability of B given A found from an experiment in which A occurs in some replications and the complement of A occurs in other replications. This tantamounts to a negation of the principle of conditionalization.

2.2. Using Bayes' Law, directly

The clearest, and perhaps the most natural way to address the question posed is via a use of Bayes' Law. But to better articulate the workings of this law in the present context, we introduce the convention (see Singpurwalla, 2006b) that for two events A and B, $P(A|B)$ denotes the conditioning (or supposition) that B is true, whereas $P(A;B)$ denotes the fact that B is actually true. With the above convention in place, our problem boils down to assessing $P(X>x+u;X>x)$. The answer is given by Eq. (2.2). But the arguments leading to this equation entail a transition from purely probabilistic considerations to the statistical ones, and these may be helpful to re-iterate.

To assess $P(X>x+u;X>x)$, one way to start is by considering the proposition $P(X>x+u|X>x)$, which by Bayes' Law leads us to the inverse relationship

$$P(X>x+u|X>x) \propto P(X>x|X>x+u)P(X>x+u), \tag{2.1}$$

where "$\propto$" denotes proportional to. Eq. (2.1) is an honest-to-goodness probability statement.

However, since $(X>x)$ has been observed as data, the middle term of Eq. (2.1) does not make sense as a probability. Instead, it is the *likelihood* of the event $X>x+u$ with $X>x$ fixed. We denote this likelihood by $\mathscr{L}(X>x+u;X>x)$. Similarly, $P(X>x+u|X>x)$ must now be written as $P(X>x+u;X>x)$. In writing $\mathscr{L}(X>x+u;X>x)$ we interpret $X>x+u$, $u>0$, as a hypothesis and $X>x$ as data. This interpretation is not conventional in the sense that in statistical inference likelihoods are generally functions of unknown parameters, not unknown events. However, as stated by Edwards (1992, p. 12), the likelihood can be regarded as a function of the hypotheses or of the parameters. A treatment of the question posed involving the use of a parametric model which results in the likelihood being a function of the parameter will be discussed in Section 2.3.

With the above in place, Eq. (2.1) now becomes

$$P(X > x+u; X > x) \propto \mathscr{L}(X > x+u; X > x)P(X > x+u). \tag{2.2}$$

The last term of the above expression, being an unknown quantity, is $\overline{G}(x+u)$.

According to Basu (1975, 1982), when Fisher (1912) rediscovered the Gaussian notion of likelihood, he looked upon it as "a scale of comparative support lent by the data to various possible values of θ [an unknown parameter]"; also see Edwards (1992, p. 221). This interpretation of likelihood is (symmetrically) different from the conventional interpretation in which the likelihood tells us which hypothesis better supports the data (cf. Edwards, 1992, p. 9). The point of view that we adopt here is the former. Having done so, we are—in principle—free to choose the functional form of the likelihood function as we see fit. Suppose then, that the likelihood is taken to be a constant, say 1, over all values of $x+u$, with x fixed; see Fig. 1. Note that this choice will also be in keeping with the conventional use of the likelihood. Then Eq. (2.2) would become

$$P(X > x+u; X > x) \propto 1 \cdot P(X > x+u),$$

which when normalized yields $P(X > x+u)/P(X > x) = \overline{G}(x+u)/\overline{G}(x)$ as an answer. Thus, implicit to the answer given by those who subscribe to the principle of conditionalization (i.e. those who mimic the steps to assess conditional probability) is the assumption of a constant likelihood!

Since one is free to choose the functional form of the likelihood, what if the likelihood was chosen by us, see Fig. 1, to be some other function of u, say $\exp(-u)$, for $u > 0$? Our assessment of $P(X > x+u; X > x)$ would be different; namely, it would be $\exp(-u)\overline{G}(x+u)/\overline{G}(x)$. This means that it is the form of the likelihood that dictates how we would bet on residual life. The standard answer $\overline{G}(x+u)/\overline{G}(x)$ arises only under the special case of a constant likelihood.

The constant likelihood encapsulates a user's disposition of indifference with respect to the observed $X > x$. A decreasing likelihood one of conservatism. The form of likelihood can therefore be given a behaviouristic justification.

2.3. Using Bayes' Law, conventionally

By a conventional use of the Bayes Law we mean the introduction of a parametric model into the analysis followed by a prior to posterior transformation of our uncertainty about the parameters. When we do so, an argument similar to the one of Section 2.2 can be made, and possibly with more transparency, because of the concrete nature of the set-up. Suppose then, that $P(X \leqslant x|\theta) = G(x|\theta)$, where $\theta > 0$ is some unknown

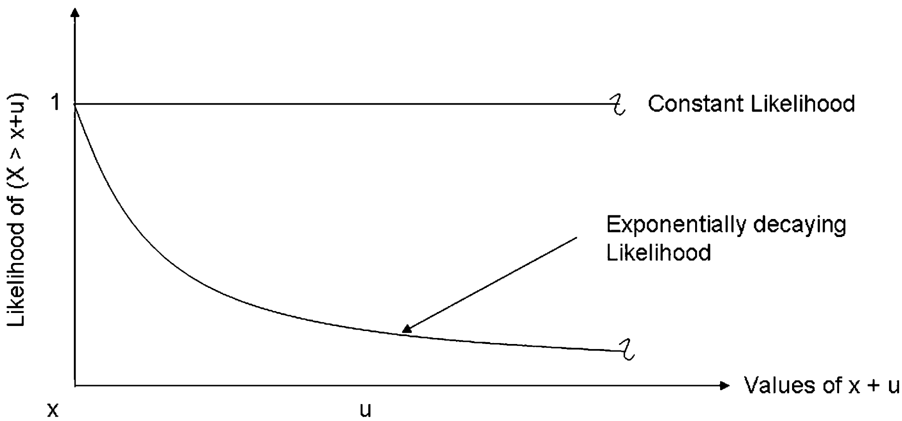

Fig. 1. Likelihood of event $(X > x+u)$ with $(X > x)$ fixed.

parameter. Using standard arguments involving the law of total probability, we may write

$$P(X \geqslant x+u|X \geqslant x) = \int_\theta P(X \geqslant x+u|X \geqslant x, \theta)\pi(\theta|X \geqslant x)\,\mathrm{d}\theta,$$

where by Bayes' Law

$$\pi(\theta|X \geqslant x) \propto P(X \geqslant x|\theta)\pi(\theta);$$

$\pi(\theta)$ is our prior distribution of $\theta > 0$.

With the event $(X \geqslant x)$ as data, the above relationship can be written as

$$P(X \geqslant x+u; X \geqslant x) = \int_\theta P(X \geqslant x+u|\theta; X \geqslant x)\pi(\theta; X \geqslant x)\,\mathrm{d}\theta \tag{2.3}$$

with

$$\pi(\theta; X \geqslant x) \propto \mathscr{L}(\theta; X \geqslant x)\pi(\theta); \tag{2.4}$$

$\mathscr{L}(\theta; X \geqslant x)$ is the likelihood of θ, with $X \geqslant x$ taken to be fixed, known, and also assumed to be credible.

Were we to subscribe to the principle of conditionalization, then $\mathscr{L}(\theta; X \geqslant x)$ will be prescribed by our chosen model $G(x|\theta)$. If otherwise, we are free to choose any other meaningful form for $\mathscr{L}(\theta; X \geqslant x)$, and thus our answers to $P(X \geqslant x+u; X \geqslant x)$ could be different. The example below illustrates this point.

Let $G(x|\theta) = 1 - \exp(-\theta x)$, an exponential distribution with mean $1/\theta$, $\theta > 0$, and let our prior on θ be a gamma distribution with scale (shape) parameter 1 (k). This is a natural conjugate prior for θ, though any other prior will also do. Then

$$\begin{aligned} P(X \geqslant x+u; X \geqslant x) &= \int_0^\infty P(X \geqslant x+u|\theta; X \geqslant x)\pi(\theta; X \geqslant x)\,\mathrm{d}\theta \\ &= \int_0^\infty \mathrm{e}^{-u\theta}\pi(\theta; X \geqslant x)\,\mathrm{d}\theta, \end{aligned}$$

and

$$\pi(\theta; X \geqslant x) \propto \mathscr{L}(\theta; X \geqslant x)\mathrm{e}^{-\theta}\theta^{k-1}/\Gamma(k).$$

When $\mathscr{L}(\theta; X \geqslant x) = \mathrm{e}^{-\theta x}$—which is what the principle of conditionality would mandate, and which is what is conventionally done—then it can be verified that the posterior distribution of θ is also a gamma with scale [shape] $(x+1)[k]$; i.e.

$$\pi(\theta; X \geqslant x) = \mathrm{e}^{-\theta(x+1)}\theta^{k-1}(x+1)^k/\Gamma(k).$$

It now follows that

$$\begin{aligned} P(X \geqslant x+u; X \geqslant x) &= \int_0^\infty \mathrm{e}^{-u\theta}\mathrm{e}^{-\theta(x+1)}\frac{\theta^{k-1}}{\Gamma(k)}(x+1)^k\,\mathrm{d}\theta \\ &= \left(\frac{x+1}{x+u+1}\right)^k. \end{aligned} \tag{2.5}$$

As an aside if the prior on θ were taken to be an *improper prior*, $\pi(\theta) = 1$, $\theta > 0$, then $P(X \geqslant x+u; X \geqslant x) = (x/(x+u))$. This assessment of residual life is similar, but not identical, to that of Eq. (2.5) with $k = 1$.

Suppose now that one were to not subscribe to the principle of conditionality and chose $\mathscr{L}(\theta; X \geqslant x) = c$; i.e. the likelihood is a constant $c > 0$. Then the posterior of θ would equal its prior, and Eq. (2.5) would become $(u+1)^{-k}$. Here the effect of x vanishes, because in choosing a flat likelihood one essentially says that irrespective of what x is, an equal weight is given to all values of θ. Clearly, this choice for a likelihood is not appealing. However, the following choice for $\mathscr{L}(\theta; X \geqslant x)$ appears to be a more sensible alternative.

Suppose that instead of choosing $\mathscr{L}(\theta; X \geqslant x) = \exp(-\theta x)$—a decreasing function of θ—one were to choose $\mathscr{L}(\theta; X \geqslant x) = \exp(-\theta\beta x)$, for some $\beta > 0$. The likelihood would still be a decreasing function of θ, but the rate of decrease would vary, depending on the value of β; see Fig. 2.

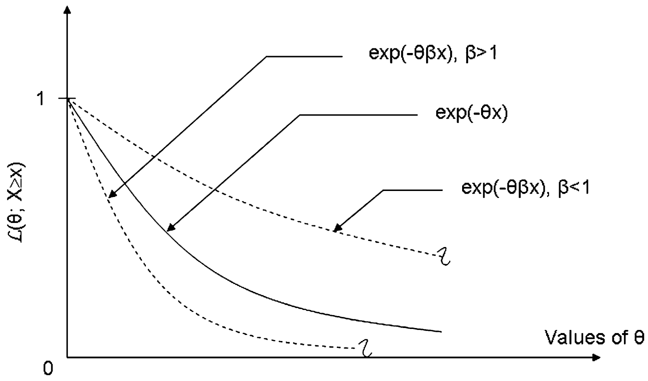

Fig. 2. Likelihood of θ with $X > x$ fixed.

For $\beta > 0$, Eq. (2.5) would become

$$P(X \geqslant x + u; X \geqslant x) = \left(\frac{\beta x + 1}{\beta x + u + 1}\right)^k, \tag{2.6}$$

so that the introduction of a β in the likelihood tantamounts to assigning a weight β to the observed value of x. This in some scenarios could be a desirable feature to have, say when the accuracy (i.e. the credibility) of the observed x is suspect. The choice $\beta > (<)1$ would inflate (deflate) x, and this in turn would cause the likelihood to decay faster (slower) than the conventional $\exp(-\theta x)$. Since θ is the reciprocal of the mean time to failure, accentuating large values of θ, as the choice $\beta < 1$ would tend to do, boils down to accentuating small values of the mean time to failure and thence small values of the residual life. Similarly with $\beta > 1$. The choice $\beta = 1$ encapsulates full faith in the observed x and also an adherence to the principle of conditionality. Eqs. (2.5) and (2.6) support our claim that the introduction of a parametric model increases the transparency of the point we are trying to make.

2.3.1. *Discussion: the advantage of parametric models*

Parametric models are used because they facilitate a coherent updating of the assessed uncertainties via a mechanistic application of Bayes' formula. The example of Section 2.3.2 underscores this point. By contrast, the direct approach of Section 2.2 requires of the user a fresh specification of the likelihood every time new evidence becomes available. This process, besides being cumbersome, has the danger of leading one to incoherence should one not be thoughtful about one's specifications. The disadvantage of parametric models is that the chosen model may not be an accurate reflection of reality. All the same the computational advantage offered by parametric models outweighs the disadvantage of misspecification, and thus their common use.

2.3.2. *Application to survival time data on winding life*

To illustrate the workings of the material of this section we consider here some service life data on "field windings" of generators given by Nelson (2000). The data below, abstracted from Nelson (2000, Table 1), consists of months in service of failed and unfailed windings. The 16 ranked failures and survival times—with the former tagged by an asterisk—in months are

$$31.7^*, 39.2^*, 57.5^*, 65.0, 65.8^*, 70.0^*, 75.0, 75.0, 87.5, 88.3, 94.2, 101.7, 105.8^*, 109.2, 110.0^*, \text{ and } 130.0.$$

Observe that seven out of the 16 field windings have experienced failures and of the nine that have not the largest (smallest) service life is 130 (65) months. Suppose that for the purposes of planning for maintenance, we are interested in the probability of any one of the surviving units not failing for an additional $u > 0$ months. For the sake of discussion let us pick the unit with the largest accumulated life. That is, we need to assess $P(X \geqslant 130 + u; \mathbf{d})$, where $\mathbf{d}$ denotes the life history data given above.

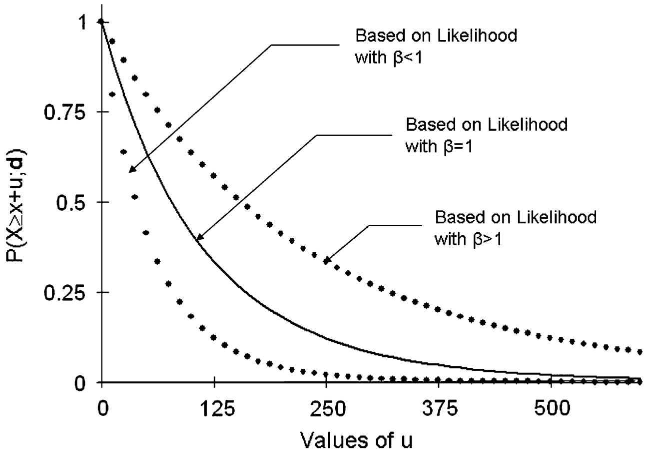

Fig. 3. Probability of the longest surviving unit surviving an additional u months.

Assuming that $P(X \geqslant x|\theta) = \exp(-\theta x)$, with a gamma prior for θ with scale (shape) parameter $1(k)$, it can be verified that under an adherence to the principle of conditionality, the posterior distribution of θ is also a gamma with scale $(\sum_1^m x_i + \sum_1^n t_i + 1)$, and shape $k+n$, where $\sum_1^m x_i$ is the sum of the m survival times and $\sum_1^n t_i$ is the sum of the n failure times. When such is the case, we have—as an analogue to Eq. (2.5)—the result that for any unfailed unit, that has experienced a service life of x,

$$P(X \geqslant x+u; \bullet) = \left(\frac{\sum_1^m x_i + \sum_1^n t_i + 1}{\sum_1^m x_i + \sum_1^n t_i + u + 1}\right)^{k+n}. \tag{2.7}$$

Eq. (2.7) when invoked—for $k=5$—in the context of the surviving unit with an accumulated service life of 130 months and the life history data given above yields, for $u \geqslant 0$,

$$P(X \geqslant 130 + u; \mathbf{d}) = \left(\frac{1306.9}{1306.9 + u}\right)^{12}. \tag{2.8}$$

A plot of $P(X \geqslant 130 + u; \mathbf{d})$ versus u, for $u \geqslant 0$, is shown as the bold faced curve of Fig. 3.

Were the principle of conditionality not adhered to and the likelihood function be modulated by the constant $\beta > 0$, then our analogue to Eq. (2.6) would be

$$P(X \geqslant x+u; \bullet) = \left(\frac{\beta(\sum_1^m x_i + \sum_1^n t_i) + 1}{\beta(\sum_1^m x_i + \sum_1^n t_i) + u + 1}\right)^{k+n}. \tag{2.9}$$

Eq. (2.9) when invoked in the context of the scenario leading up to Eq. (2.8) for $\beta = \frac{1}{2}$ and 2 would result in the dotted curves of Fig. 3. Our assessed survival probability depends on the form chosen for the likelihood. In principle, likelihood plays a more crucial role than the prior, because whereas the prior gets updated with new evidence, the likelihood stays put from the start.

3. Conclusion

The innocuously simple problem of assessing conditional probabilities can get riddled with issues, both philosophical and technical, when the conditioning event becomes a reality. The cleanest way to approach it is through Bayes' Law. When this is done it can be seen that the standard answer arises as a special case under the assumption of a constant likelihood. Other forms of the likelihood will lead to other answers. Since the choice of a likelihood is an assessors prerogative—just like the choice of a probability model—there is no unique and correct way to bet on residual life. However, the traditional answer (presumably the one that will be subscribed to by *card carrying probabilists*) will be the correct and unique answer, but only when its argument is sheltered under the philosophical (or behaviouristic) principle of conditionalization.

Acknowledgements

The comments of a referee are acknowledged. The referee drew attention to Joyce (1999) and made us aware that matters discussed here have been considered by philosophers of science. Supported by Grants N00014-06-1-037 by Office of Naval Research and W911NF-05-1-2009 by the US Army Research Office.

References

Basu, D., 1975. Statistical information and likelihood. Sankhya Ser. A 37A (Pt. 1), 1–71.

Basu, D., 1982. A note on likelihood. In: Symposio Nacional de Probabilidade e Estatisticia, Universidade De Sao Paolo, pp. 1–5.

de Finetti, 1937. La Prevision: ses Lois Logiques, ses Sources Subjectives. Ann. Inst. H. Poincare (Paris) 7, 1–68.

de Finetti, B., 1972. Probability, Induction and Statistics. Wiley, New York.

de Finetti, B., 1975. Theory of Probability, vol. 2. Wiley, New York.

Diaconis, P., Zabell, S.L., 1982. Updating subjective probabilities. J. Amer. Statist. Assoc. 77 (380), 822–830.

Edwards, A.W.F., 1992. Likelihood. The Johns Hopkins Press, Baltimore, MD.

Fisher, R.A., 1912. On an absolute criterion for fitting frequency curves. Mess. Math. 41, 155–160.

Freedman, D., Purves, R., 1969. Bayes method for bookies. Ann. Math. Statist. 40, 1177–1186.

Hacking, I., 1967. Slightly more realistic personal probability. Philos. Sci. 34, 311–325.

Howson, C., Urbach, P., 1989. Scientific Reasoning: The Bayesian Approach. Open Co., La Salle, IL.

Joyce, J., 1999. The Foundations of Causal Decision Theory. Cambridge University Press, Cambridge.

Jonsson, C., 1974. Electron diffraction at multiple slits. Amer. J. Phys. 42, 4–11.

Nelson, W., 2000. Theory and applications of hazard plotting for censored failure data. Technometrics 42 (1), 12–25.

Ramsey, F.P., 1931. Truth and probability. In: Braithwaite, R.G. (Ed.), The Foundations of Mathematics and Other Logical Essays. Routledge and Kegan Paul, London, pp. 156–198.

Singpurwalla, N.D., 2006a. On competing risk and degradation processes. In: Rojo, J. (Ed.), Optimality—The Second Erich Lehmann Symposium. IMS Lecture Notes—Monograph Series, vol. 49, pp. 289–304.

Singpurwalla, N.D., 2006b. Reliability and Risk: A Bayesian Perspective. Wiley, New York.

Teller, P., 1976. Conditionalization, observation, and change of preference. In: Harper, W.L., Hooker, C.A. (Eds.), Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science, vol. 1. D. Deidel, Dordrecht, pp. 205–253.