The Utility of Reliability and Survival

Nozer D. Singpurwalla
The George Washington University, Washington, D.C. 20052

Revised, July 2009

Abstract

Reliability (survival analysis, to biostatisticians) is a key ingredient for making decisions that mitigate the risk of failure. The other key ingredient is utility. A decision theoretic framework harnesses the two, but to invoke this framework we must distinguish between chance and probability. We describe a functional form for the utility of chance that incorporates all dispositions to risk, and propose a probability of choice model for eliciting this utility. To implement the model a subject is asked to make a series of binary choices between gambles and certainty. These choices endow a statistical character to the problem of utility elicitation. The workings of our approach are illustrated via a live example involving a military planner. The material is general because it is germane to any situation involving the valuation of chance.

Key Words: Choice Models, Decision Making, Probability, Propensity, Quality of Life, Risk Analysis.

1 Introduction and Overview

1.1 Preamble: Motivation and Objectives

Perhaps a better title for this paper could be "The Utility of Chance", but it would detract from its motivating import, which is applied and pragmatic. Indeed, the work here was suggested by a scenario wherein the author was asked to determine if the reliability of an amphibious landing tank, called the "Expeditionary Fighting Vehicle" was in excess of .999, a number sacrosanct to a commanding general of the U.S. Marine Corps. By reliability, we mean the survival function evaluated at any specified time, called the "mission time". Given the vehicle's architecture, the said number was literally impossible to achieve. Thus arose the question of why .999? It turned out, as is usual the case with reliability specifications, that such numbers are arbitrary, more a matter of decree than a consideration of need. See, for example, The Washington Post, February 2, 2008 article on "GAO Report Criticizes Defense on its Acquisition of Weapons". A more recent, albeit related, case in point is the U. S. Air Force's decision to opt for the oversized and cost-ineffective KC-30 tanker over the right-sized and less costly KC-767 tanker [cf. Washington Post, June 11, 2008, p. A16]. It is possible that similar situations may also prevail in the biomedical environment wherein choices that impact patient survival over cost and quality of life considerations, are not judiciously balanced.

There are two aims that underlie this paper. The first is to advocate the need for utility considerations in the reliability arena including a suggestion for its general shape. The second is to propose a statistical approach based on the item response theory models (also known as a ***choice models***) for eliciting an individual's utility. However, to set a formal stage for achieving the first aim, we need to distinguish between reliability (or the survival function) as an unknown ***chance*** or ***propensity***, and ***survivability*** as ones subjective (personal) probability about the unknown chance.

By propensity we mean the purported ***causes*** of observed stable relative frequencies; see Popper (1957). Propensities are invoked to ***explain why*** repeating a certain kind of experiment will generate a given outcome type at a persistent rate; they are constrained to be between 0 and 1, both inclusive. Regarding the second aim, statistical methods for eliciting utilities are virtually non-existent, and regrettably so, because such methodologies can vastly enhance the utility assessment enterprise. Some exceptions are the papers of Mosteller and Nogee (1951), Becker, De Groot and Marschak (1964), and Novick and Lindley (1979).

Stripped of the reliability centered application that has motivated our work, the underlying theme of this paper should have a wider appeal. It is germane to any situation that entails the desirability of a chance. For example, how much more desirable is a coin with a propensity of heads of say .95 to a coin whose propensity of heads is .93, given that the coins are to be used for gambling? Similarly, how much more desirable is a pill with a cure rate of .98 to one with a cure rate of .95, given that the former could cost much more than the latter? To address issues such as these we need to assess the utility of propensity. But first in order are some comments on the roles of reliability and survival analysis in risk management, the roles of probability and utility decision in making, and the structure of a decision problem. Sections 1.2 through 1.4 are devoted to these topics. Section 1.5 is a summary presentation of de Finetti's (1972) theorem on infinite exchangeable Bernoulli sequences, and how this theorem may provide a hook on which the essence of the material here can be thought of. The rest of this paper, Sections 2, 3, 4, and 5, pertains to the utility of reliability, a model for eliciting utilities, deploying the model, and a live application, respectively. Section 6 closes this paper. It may be of relevance to note that the matter of a utility of chance is to be contrasted with that of the utility of probability for which there is a precedence in the works of Lindley (1976), and Good and Card (1971).

1.2 Reliability and Survival Analysis in Risk Management

Reliability (Survival) analyses done by engineers (biostatisticians) provide yardsticks for quantifying the random nature of lifetimes. We quantify this randomness for managing the risk of failure. Managing risk means making choices that minimize the losses caused by adverse events. In the context of engineering, such choices pertain to deciding between competing designs, managing maintenance, or the acceptance/rejection of manufactured lots. In biomedicine, decision making pertains to treatment options, and other choices that impact survival and the quality of life.

1.3 Coherent Decision Making: Probability and Utility

Coherent decision making rests on two pillars, probability and utility, and the principle of maximization of expected utility (MEU). Probability quantifies uncertainty and utility quantifies preferences. Utility in statistical inference is via loss functions, such as squared error, absolute error, linex, etc.. Such loss functions are stylized. There appears to be a dearth of literature in statistical outlets about eliciting the actual loss functions of decision makers. This state of affairs is also true in the engineering sciences such as filtering, control, and information fusion, wherein a use of squared-error loss functions seems to be the norm.

1.4 The Structure of a Decision Problem

Suppose that a decision maker $\mathcal{D}$ is required to choose one among a set of n mutually exclusive and exhaustive actions, a_1, a_2,..., a_n. Associated with a_i, are k_i outcomes (states of nature) θ_{ij}, $j = 1, ..., k_i$, assumed mutually exclusive and exhaustive. When $\mathcal{D}$ chooses a_i, $\mathcal{D}$ is uncertain about the outcome. Let $\mathbf{P}(\theta_{ij})$ be $\mathcal{D}$'s probability of occurrence of θ_{ij}, $i = 1, ..., n$, $j = 1, ..., k_i$; $\mathbf{P}(\theta_{ij})$ is personal to $\mathcal{D}$. Let $U(\theta_{ij})$ be $\mathcal{D}$'s utility of θ_{ij}; $U(\theta_{ij})$ is also personal to $\mathcal{D}$. $U(\bullet)$ is a numerical quantity between 0 and

1, and if $U(\theta_{ij}) > U(\theta_{im})$, $j \neq m$, then $\mathcal{D}$ prefers θ_{ij} to θ_{im}. We assume that all the θ_{ij}'s can be preference ranked by $\mathcal{D}$ (i.e. $\mathcal{D}$ satisfies the ***axiom of completeness***). The focus of this paper is to develop a procedure for obtaining $U(\theta_{ij})$, in a manner that ensures a certain kind of consistency; this will be clarified later in Section 4.1. The expected utility of a_i is,

$$\mathbf{E}[U(a_i)] = \sum_j U(\theta_{ij})\mathbf{P}(\theta_{ij})$$

and the MEU principle prescribes that $\mathcal{D}$ choose that a_i for which $\mathbf{E}[U(a_i)]$ is a maximum. In developing an argument for this principle $\mathcal{D}$ assumes that there exists a θ_{ij}, say θ^* for which $U(\theta^*) = 1$, and some other θ_{ij}, say θ_* for which $U(\theta_*) = 0$. The θ^* and θ_* are known as ***anchor points***. In our particular application the θ_{ij} are propensities; thus $\theta_{ij} \in [0, 1]$, with anchor points $U(0) = 0$, and $U(1) = 1$.

To harness the notions of reliability and survival analysis for risk management, we need to cast them in the decision making framework described above. One approach to doing so is motivated by de Finetti's Theorem (1972) on exchangeable sequences. The theorem leads to the view that the reliability and the survival function are a chance (or a propensity), and not a probability. By contrast, Kolmogorov (1969) does not distinguish between chance and probability, so that to him reliability is a probability. It is not clear as to how under this conventional view of Kolmogorov, reliability and survival functions can be formally cast in a decision theoretic framework.

1.5 de Finetti's Theorem: Infinite Exchangeable Sequences

Let $X_1, X_2, ...,$ be an infinite sequence of (non-negative) continuous random variables with the property that for some $\boldsymbol{\theta}$ and all i

$$\mathbf{P}(X_i \geq x|\boldsymbol{\theta}) = \overline{F}(x|\boldsymbol{\theta}),\ x \geq 0.$$

Then for every finite $n \geq 1$,

$$\mathbf{P}(X_1 \geq x_1, ..., X_n \geq x_n) = \int_0^\infty \prod_1^n \overline{F}(x_i|\boldsymbol{\theta})\Pi(\boldsymbol{\theta})d\boldsymbol{\theta},$$

where $\Pi(\boldsymbol{\theta})$ encapsulates $\mathcal{D}$'s uncertainty about $\boldsymbol{\theta}$. Thus for $n = 1$

$$\mathbf{P}(X \geq x) = \int_0^\infty \overline{F}(x|\boldsymbol{\theta})\Pi(\boldsymbol{\theta})d\boldsymbol{\theta}; \tag{1.1}$$

$\overline{F}(x|\boldsymbol{\theta})$ is the reliability (or the survival function) of X, and is likened to chance; see Lindley and Phillips (1976), or Lindley and Singpurwalla (2002). The quantity $\mathbf{P}(X \geq x)$ is $\mathcal{D}$'s uncertainty about the event $(X \geq x)$ described via a personal probability. We label it as the ***survivability*** of X, and is distinguished from the ***survival function*** $\mathbf{P}(X \geq x|\boldsymbol{\theta})$. In a decision theoretic set-up, $\mathbf{P}(X \geq x|\boldsymbol{\theta}) = \overline{F}(x|\boldsymbol{\theta}) \in [0, 1]$ is the unknown state of nature, and $\Pi(\boldsymbol{\theta})$ encapsulates $\mathcal{D}$'s uncertainty about it.

The goal of this paper is to assess the utility of $\overline{F}(x|\boldsymbol{\theta})$, for fixed $x \geq 0$.

2 The Utility of Reliability / The Survival Function

Let $U[\overline{F}(x|\boldsymbol{\theta})]$ denote $\mathcal{D}$'s utility of $\overline{F}(x|\boldsymbol{\theta})$. With $\overline{F}(x|\boldsymbol{\theta}) \in [0, 1]$, we anchor on two points $\overline{F}(x|\boldsymbol{\theta}) = 1$ and $\overline{F}(x|\boldsymbol{\theta}) = 0$, setting $U(1) = 1$ and $U(0) = 0$, and ask what is $U(\overline{F}(x|\boldsymbol{\theta}))$ for any $\overline{F}(x|\boldsymbol{\theta}) \in (0, 1)$?

To address this question, we offer $\mathcal{D}$ the following two choices:

i) Receive $\overline{F}(x|\boldsymbol{\theta})$ for sure, or

ii) Agree to a gamble wherein $\mathcal{D}$ receives $\overline{F}(x|\boldsymbol{\theta}) = 1$ with chance p, and $\overline{F}(x|\boldsymbol{\theta}) = 0$ with chance $(1 - p)$.

Then, $\mathcal{D}$'s utility of $\overline{F}(x|\boldsymbol{\theta})$ is that value of p for which D is indifferent between the two choices of certainty versus uncertainty -henceforth a **p-*gamble***. With $U[\overline{F}(x|\boldsymbol{\theta})]$ so anchored archetypal utility functions can be prescribed by the relationship $U[\overline{F}(x|\boldsymbol{\theta})] =$

$(\overline{F}(x|\boldsymbol{\theta}))^{\frac{\beta}{x}}$, for $\beta > 0$; see Figure 2.1. Note that $\mathcal{D}$ is risk neutral for $\beta = x$, and risk prone (averse) for $\beta > (<)\ x$.

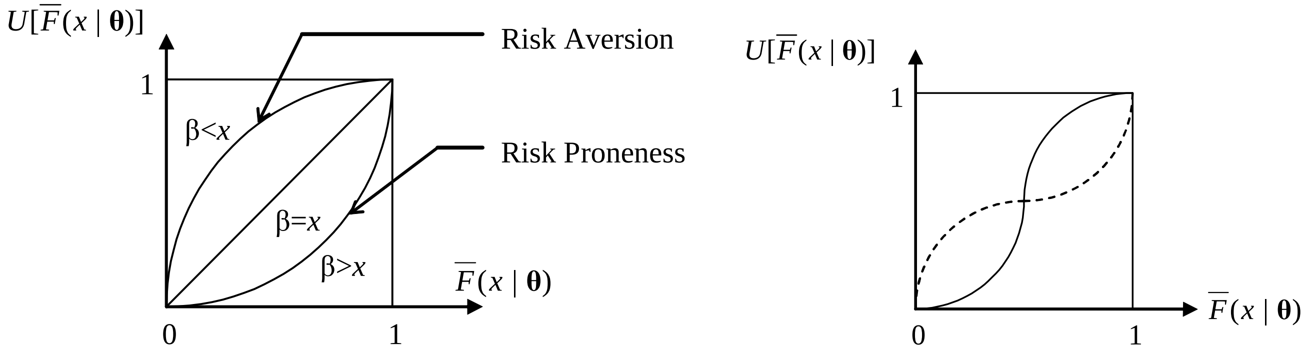

Figure 2.1: Archetypal Forms of $\mathcal{D}$'s Utility of $\overline{F}(x|\boldsymbol{\theta})$.

It can be so that $\mathcal{D}$ is risk prone for small values of $\overline{F}(x|\boldsymbol{\theta})$ and risk averse for large values of $\overline{F}(x|\boldsymbol{\theta})$ making the utility function S-shaped; similarly, a reverse S-shape, mutatis-mutandis.

2.1 Eliciting Utility: Some Caveats.

For convenience set $\overline{F}(x|\boldsymbol{\theta}) = c$, for $c \in (0, 1)$. Recall that $U(c)$ is that chance, say p^*, at which $\mathcal{D}$ is indifferent between receiving a c for sure, versus a p^*-gamble.

The conventional approach for eliciting p^*, revolves around two methods, ***fixed probability***, and ***fixed state*** [cf. Hull, Moore, and Thomas (1973), or Farquhar (1984)]. In the former, $\mathcal{D}$ is presented with a p^*-gamble, and is asked to choose a $c \in (0, 1)$ for which $\mathcal{D}$ is indifferent between c and the p^*-gamble. The c so chosen is called the ***certainty equivalent*** of the gamble. This exercise is repeated for a range of values of $p^* \in (0,1)$. In the fixed state method, c is fixed and $\mathcal{D}$ is interrogated over a range of values of p until $\mathcal{D}$ converges on a p^* for which $\mathcal{D}$ is indifferent between receiving the fixed c and a p^*-gamble. Either method is cumbersome because it is difficult to iterate around an indifference value of c or p. However, for any fixed p, it may be easier for $\mathcal{D}$ to make a binary choice between receiving a sure c versus a p-gamble. Indeed, this is the essence of our proposed choice model based approach for eliciting utility; this approach is discussed in Sections 3.

In addition to the elicitation difficulty mentioned above, there are two other issues associated with the conventional approach. For one, there is no assurance that $\mathcal{D}$ will be consistent in the declared values of p. Specifically, the p's should be non-decreasing in $\overline{F}(x|\boldsymbol{\theta})$ -the ***monotonicity requirement***- and they must be invariant with respect to the anchor points used to elicit them -the ***invariance requirement***. Attempts at resolving this latter type of inconsistency have entailed a use of linear programming and regression based approaches [cf. Meyer and Pratt (1968), and Novick and Lindley (1979), respectively].

Another feature of the conventional approach is that it being fundamentally deterministic and iterative, there is no provision for accommodating $\mathcal{D}$'s lack of definitiveness about the declared values. By contrast, a Bayesian procedure based on binary choices will have a built in mechanism for the treatment of $\mathcal{D}$'s uncertainties.

Finally, there is a price to be paid for achieving a high reliability and the survival probability. This tantamounts to a ***disutility***. The matter of disutilities associated with high survival probabilities, has spawned the topic of "***quality of life***" in the health sciences, [cf. Mesbah and Singpurwalla (2008)].

2.2 Incorporating Disutility: An Omnibus Utility.

For purposes of illustration, an archetypal function that is able to reflect the feature that an increase in reliability should be accompanied by an increase in cost, so that the disutility due to cost is an increasing function of reliability, can be of the form

$$1 - \exp\left(-\frac{\delta \overline{F}(x|\boldsymbol{\theta})}{1 - \overline{F}(x|\boldsymbol{\theta})}\right), \tag{2.1}$$

for some $\delta > 0$.

Combining this disutility with the utility, yields an ***omnibus utility*** for reliability (or survival) as:

$$(\overline{F}(x|\boldsymbol{\theta}))^{\frac{\beta}{x}} - \left[1 - \exp\left(-\frac{\delta \overline{F}(x|\boldsymbol{\theta})}{1 - \overline{F}(x|\boldsymbol{\theta})}\right)\right], \tag{2.2}$$

for $x \geq 0$, β, $\delta > 0$, and $\overline{F}(x|\boldsymbol{\theta}) \in [0, 1]$. Like Equation (2.1), Equation (2.2) is also illustrative.

3 A Probability of Choice Model for Utility Elicitation

The thesis that it is easier for $\mathcal{D}$ to make a binary choice between the options of receiving an $\overline{F}(x|\boldsymbol{\theta}) \in (0, 1)$ for sure, or receiving $\overline{F}(x|\boldsymbol{\theta}) = 1(0)$ with chance $p(1-p)$, versus arriving upon a p by iteration motivates us to consider ***Probability of Choice Models*** as a possible mechanism for eliciting utilities. In Section 3.1 we motivate and introduce our model. A use of this model entails fixing c at some c_i, and p at p_{ij}, and then asking $\mathcal{D}$ to make a binary choice between a sure c_i versus a p_{ij}-gamble, for $j = 1, 2, ..., n_i$. We set $Y_{ij} = 1(0)$ if $\mathcal{D}$ opts (does not opt) for the gamble. We repeat this exercise for different values of c_i, $i = 1, 2, ..., n_i$. Using the Y_{ij} as data, we estimate the model parameters, and for every c, find that p for which the probability of choosing the p-gamble is 0.5. This binary choice strategy resonates with the method used by Mosteller and Nogee (1951) who consider the proportion of times a subject chooses the various gambles.

3.1 The Proposed Model for Utility Elicitation

The proposed model is based on how hard or easy it is for $\mathcal{D}$ to make the binary choices. Given below are the boundary conditions for a rational $\mathcal{D}$:

- When $p = 1$, $\mathcal{D}$ will choose $Y = 1$ for all $c < 1$;
- When $p = 0$, $\mathcal{D}$ will choose $Y = 0$ for all $c > 0$;
- When $p = c$, a ***risk neutral*** $\mathcal{D}$ will choose $Y = 1$ or $Y = 0$, equally often; by contrast, a ***risk prone*** (***averse***) $\mathcal{D}$ will choose $Y = 1$ more (less) often than $Y = 0$.

Let $\mathbf{P}(Y = 1) = \Pi$ denote an elicitor $\mathcal{E}$'s personal probability of $\mathcal{D}$ choosing a p-gamble over the certain c. Note that there are two kinds of entities that come into play; $\mathcal{E}$'s personal probability Π, and a chance p. $\mathcal{D}$'s indifference between the two choices tantamounts to $\Pi = \frac{1}{2}$. $\mathcal{D}$'s choices are the easiest with $\mathcal{E}$'s $\Pi = 1$, when c and p lie on the lines joining $(c = 0, p = 0)$, $(0, 1]$, and $(1, 1)$ and with $\Pi = 0$, when they lie on the lines joining $(0, 0)$, $(1, 0]$ and $(1, 1)$. These are the boundaries of Figure 3.1. $\mathcal{D}$'s choices become difficult as p and c get close to each other, becoming the most difficult when $c = p$. The roles of p and c are analogous to those of the ability and the difficulty parameters in the Rasch Model used in education testing and quality of life studies [cf. Mesbah, Cole and Lee (2002)].

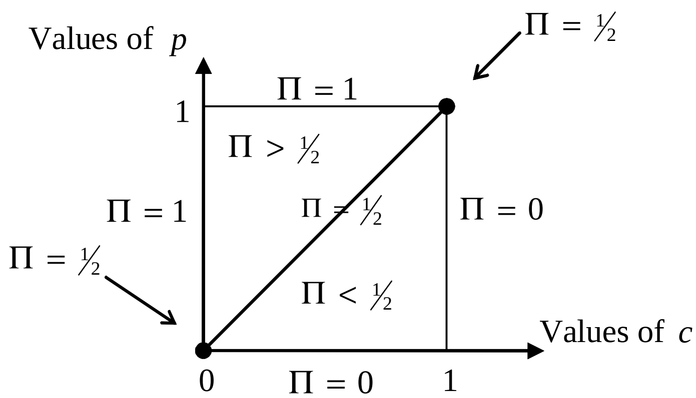

Figure 3.1: Boundary Conditions for a Risk Neutral $\mathcal{D}$

The boundary conditions of Figure 3.1 motivate the general forms of Figure 3.2 as $\mathcal{E}$'s model for Π, seen as a function of $(p - c)$.

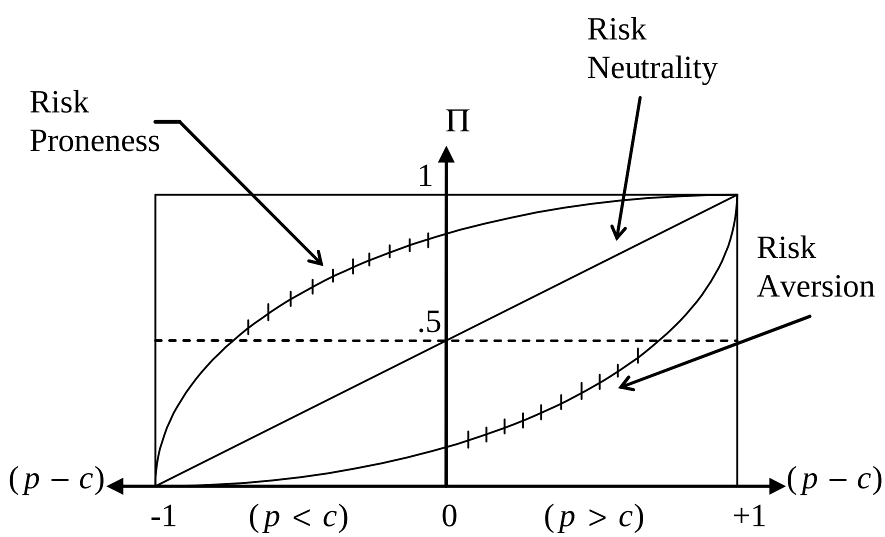

Figure 3.2: General Forms for $\mathcal{E}$'s Model for Π.

For a risk neutral $\mathcal{D}$, the right diagonal line of Figure 3.2 would encapsulate $\mathcal{E}$'s choice probabilities, whereas the concave (convex) curve would encapsulate these probabilities for a risk prone (averse) $\mathcal{D}$. The above relationships between $(p-c)$ and Π can, for $p, c \neq 0, 1$, be encapsulated via the equation

$$\Pi = \mathbf{P}(Y = 1|\beta; c, p) = \left(\frac{(p-c)+1}{2}\right)^{\beta}. \tag{3.1}$$

The inclusion of an additional parameter $\alpha > 0$ enables us to incorporate varying degrees of risk aversion and proneness; see Figures 3.3 a) and b). Thus a penultimate version of the model is

$$\Pi = \mathbf{P}(Y = 1|\alpha, \beta; c, p) = \left(\frac{(p-c)^{\alpha}+1}{2}\right)^{\beta}. \tag{3.2}$$

Figure 3.3 a) pertains to the case of $\alpha = \frac{1}{3}$ with $\beta = \frac{1}{2}$, 1, and 2 respectively, and Figure 3.3 b) is for $\alpha = 3$ with similar values for β.

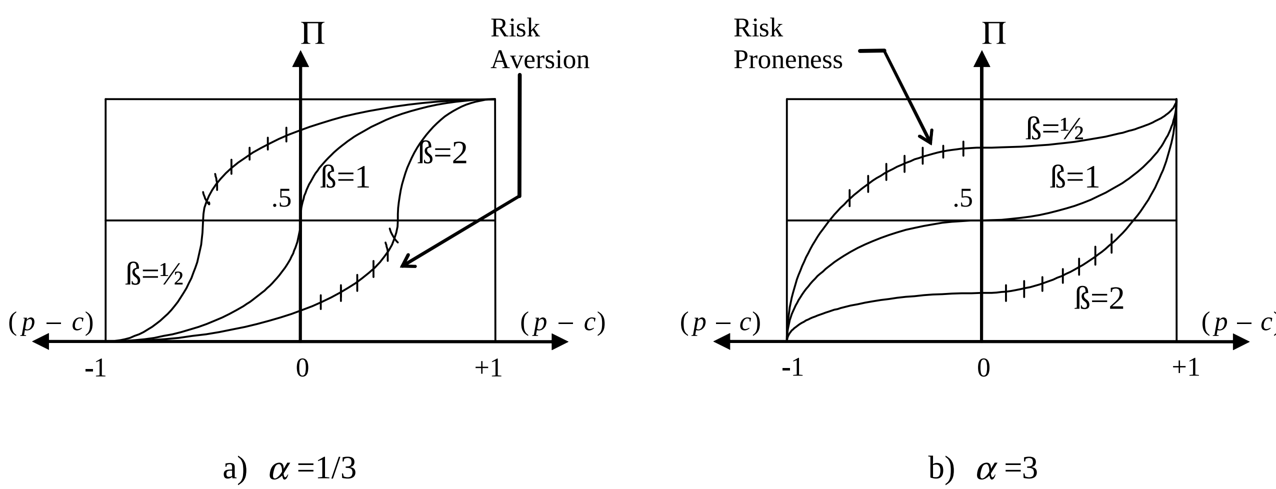

Figure 3.3: Illustration of $\mathcal{D}$'s Probability of Choice.

Whereas Equation (3.2) is intuitively appealing, it suffers from a technical deficiency which arises when we wish to solve for $(p-c)$ with Π set at 0.5. Specifically, to avoid complex roots, only certain combinations of values of α and β are admissible. Thus as a refinement of Equation (3.2), the final version of our model for $\mathcal{E}$'s choice probability is:

$$\mathbf{P}(Y=1|\alpha,\beta;c,p) = \begin{cases} 0, & \text{if } p=1 \text{ and } c<1, \text{ or } p>1 \text{ and } c=0; \\ \frac{1}{2}, & \text{if } p=0 \text{ and } c=0, \text{ or } p=1 \text{ and } c=1; \\ 1, & \text{if } p=0 \text{ and } c>0, \text{ or } p<1 \text{ and } c=1; \\ \frac{1}{2}\left[1+\operatorname{sgn}(p-c)|p-c|^{\alpha}\right]^{\beta}, & \text{otherwise,} \end{cases} \tag{3.3}$$

where $\operatorname{sgn}(\mathfrak{z}) = -1(+1)[0]$, when $\mathfrak{z} < (>)[=]0$, and $\alpha, \beta > 0$.

4 Implementing the Model

For any fixed c, our aim is to find that $(p-c)$ for which $\mathbf{P}(Y=1|\alpha,\beta;c,p)=\frac{1}{2}$; we can then solve for p, $\mathcal{D}$'s p-gamble, for the fixed c. To do the above we need estimates of $\mathcal{D}$'s values of α and β for the chosen c. This can be done using the Y_{ij}'s declared by $\mathcal{D}$, for c fixed at c_i, and a range of values of p, say p_{ij}, $j=1,...,n_i$.

In actual practice the values chosen for c_i and p_{ij} will not involve the boundary conditions since there is no need to elicit preferences for such clear cut situations. For purposes of discussion, suppose that the data Y_{ij}, $j=1,...,n_i$, yield $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ as the maximum likelihood estimators of α and β, respectively; see Appendix A. Then, the desired $(p_i - c_i) \stackrel{def}{=} \omega_i$, which is a solution to the relationship $\mathbf{P}(Y_i=1|\widehat{\alpha}_i,\widehat{\beta}_i;c_i,p_i)=\frac{1}{2}$, will be of the form:

$$\widehat{\omega}_i = \operatorname{sgn}(\widehat{\beta}_i - 1)\left[\operatorname{sgn}(\widehat{\beta}_i - 1)(2^{1-\frac{1}{\widehat{\beta}_i}} - 1)\right]^{\frac{1}{\widehat{\alpha}_i}}; \tag{4.1}$$

also, $\widehat{\omega}_i \in [-1,+1]$.

Consequently, with utilities being constrained to lie between 0 and 1, $\mathcal{D}$'s utility for c_i will be:

$$U(c_i) = \begin{cases} \min(1, c_i + \widehat{\omega}_i), & \text{if } \widehat{\omega}_i > 0, \\ \max(0, c_i + \widehat{\omega}_i), & \text{if } \widehat{\omega}_i < 0, \\ c_i, & \text{if } \widehat{\omega}_i = 0. \end{cases} \tag{4.2}$$

When $\widehat{\omega}_i < (>)[=]0$, $\mathcal{D}$ is risk prone (averse) [neutral] for c_i. The above exercise is repeated for a range of values of c, say $c_1, ..., c_i, ..., c_n$, to yield $\mathcal{D}$'s utility for survival as $U(c_1), ..., U(c_n)$, as perceived by $\mathcal{E}$.

4.1 Adjacent Point Gambles and Coherence

The elicitation of $\mathcal{D}$'s Y_{ij}'s discussed in Section 3.2 presumes ***end point gambles***; i.e. $\mathcal{D}$ receives either a sure c_i, or a 1 with chance p_{ij} and a 0 with chance $(1 - p_{ij})$. A problematic feature of these end point gambles is that the resulting utilities may not be increasing (non-decreasing) in c_i. This obstacle can be tempered by considering ***adjacent point gambles***. That is, $\mathcal{D}$ is offered a choice between a sure c_i and a p_{ij} gamble involving $U(c_{i-1})$ and $U(c_{i+1})$, with $U(c_0) = U(0) = 1$, and $U(c_{n+1}) = U(1) = 1$, $i = 1, ..., n$.

The adjacent gamble process can start-off with some c_i and an end point gamble to obtain a $U(c_i)$. This is followed by picking a c_k between c_i and 1 (or between 0 and c_i) to obtain a $U(c_k)$ via a p_{kj}-gamble involving $U(c_i)$ and 1 (or involving 0 and $U(c_i)$). Once $U(c_i)$ and $U(c_k)$ are at hand, these can be used to obtain $U(c_m)$ for $c_i < c_m < c_k$ via a sure c_m versus a p_{mj}-gamble.

The adjacent gamble approach will help achieve the monotonicity requirement of the utility function but will not guarantee it. This is because of the inherent randomness in the proposed approach which is based on estimating α and β. For a $c_m \in (c_i, c_k)$, should $U(c_m)$ be greater (less) than c_k (c_i), then a way to achieve monotonicity would be to let $U(c_m)$ lie on the line segment joining $U(c_i)$ and $U(c_k)$. A strategy such as this is used in ***isotonic regression*** [see, for example, Barlow, Bartholomew, Bremner, and Brunk (1972)].

The matter of resolving inconsistencies caused by the failure of invariance is more difficult to deal with. Novick and Lindley (1979) propose elicitation based on <u>all</u>

possible triplets of the form $0 \leq i < m < k \leq 1$, each triplet entailing a gamble, which we denote as a p_{imk}-gamble. A least-squares analysis involving the minimization of the $U(\bullet)$'s with respect to log-odds of the type

$$\sum_{i,m,k} [\log(p_{imk}/1 - p_{imk}) - \log(U(c_m) - U(c_i)/U(c_k) - U(c_m))]^2$$

is then performed; this results in a utility function $U(c_m)$. Some discussion on an approach such as this is in Becker, De Groot and Marschak (1963).

4.2 The Incorporation of Uncertainties: A Bayesian Approach

Philosophical considerations aside, an approach based on the maximum likelihood estimators $\widehat{\alpha}$ and $\widehat{\beta}$ suffers from the drawback that exact measures of uncertainty about the inferred utility function $U(\bullet)$ are difficult to obtain. This is so even at the level of pointwise confidence limits for the $U(c_i)$'s, $i = 1, ..., n$. On the other hand, a parametric Bayesian approach is able to account for the underlying uncertainties by the process of averaging out with respect to the posterior distributions of α and β. To see how, suppose that $\Pi(\alpha_i, \beta_i)$ denotes the joint prior distribution on α_i and β_i, for $i = 1, ..., n$; a specific case is described in Section A.1 of Appendix A. For any fixed c_i and data Y_{ij}, $j = 1, ..., n_i$, the above prior will lead to a posterior distribution of α_i and β_i, say $\Pi(\alpha_i, \beta_i; \bullet)$. Then $\widetilde{\omega}_i$, a Bayes assessment of $\omega_i = (p_i - c_i)$, would be the solution of

$$\int_{\alpha_i}\int_{\beta_i} \left[\left(\frac{1 + \mathrm{sgn}(\omega_i)|\omega_i|^{\alpha_i}}{2} \right)^{\beta_i} - \frac{1}{2} \right] \Pi(\alpha_i, \beta_i; \bullet) d\alpha_i d\beta_i = 0. \tag{4.3}$$

With $\widetilde{\omega}_i$, $i = 1, ..., n$ at hand, we may obtain the $U(c_i)$, $i = 1, ..., n$, by mimicking the procedure outlined before the $\widehat{\omega}_i$ obtained via the method of maximum likelihood.

5 Application: The Utility of a Vehicle's Reliability

The following real life example pertains to assessing the utility of reliability of a yet to be designed manned ground combat vehicle, as perceived by $\mathcal{D}$, a military analyst. $\mathcal{D}$ is also an officer in uniform (retired) and is thus knowledgeable about strategic needs. Furthermore, $\mathcal{D}$ is a well trained operations research analyst exposed to analytical thinking and decision making. Knowing the utility of reliability (for a mission time that is specified by $\mathcal{D}$) will help the government procurers specify the vehicle's ***design reliability***. The vehicle in question belongs to the family of systems called "Future Combat Systems". There are several hundred such new vehicles that are to be commissioned in a brigade, each costing several hundred thousand U.S. dollars. We assume exchangeability of all the new vehicles in the brigade. The reliabilities of interest to $\mathcal{D}$ are in the range of 0.5 to 0.9, the lower reliabilities of no value, and the higher reliabilities deemed unnecessary. Because of security classification, no additional details can be made available.

Elicitations from $\mathcal{D}$ entail both end point gambles and adjacent point gambles. Here c_i takes values .5, .6, .7, .8 and .9, for $i = 1, ..., 5$ respectively. The values chosen for the p_{ij}'s are .3, .4, .5, .6, .7, .8, .9 and .95, for $j = 1, ..., 8$, respectively, in the case of the end point gambles, and .3, .4, .45, .5, .55, .6, and .7, for the adjacent point gambles. The gamble probabilities p_{ij} were chosen to ensure that neither the gamble nor the sure thing would always be preferred for any of the reliabilities.

Y_{ij}		p_{ij}							
		.3	.4	.5	.6	.7	.8	.9	.95
c_i	.5	0	0	1	1	1	1	1	1
	.6	0	0	0	0	1	1	1	1
	.7	0	0	0	0	0	1	1	1
	.8	0	0	0	0	0	0	0	1
	.9	0	0	0	0	0	0	0	1

Table 5.1: $\mathcal{D}$'s Choices Under End Point Gambles.

Table 5.1 shows the results of the elicitation for the end point gambles, and Table 5.2 for the adjacent point gambles. The entries in these tables gives the values Y_{ij}, with $Y_{ij} = 1$, whenever $\mathcal{D}$ opts for the gamble. $\mathcal{D}$'s choices are solely based on strategic needs, not costs.

An inspection of the entries in Table 5.1 suggests that $\mathcal{D}$ tends to be risk averse, because $\mathcal{D}$ opts for the p_{ij} - gamble only when $p_{ij} \geq c_i$. The entries of Table 5.1 when used to obtain the $U(c_i)$'s via the method of maximum likelihood and also the Bayesian approach (using the independent gamma priors described in Section A.1 of Appendix) yields results that are almost identical; specifically $U(.5) = .5$, $U(.6) = .6$, $U(.7) = .7$, $U(.8) = .93$, $U(.9) = .92$. Figure 5.1 shows a plot of the $U(c_i)$'s versus c_i, $i = 1, ..., 5$.

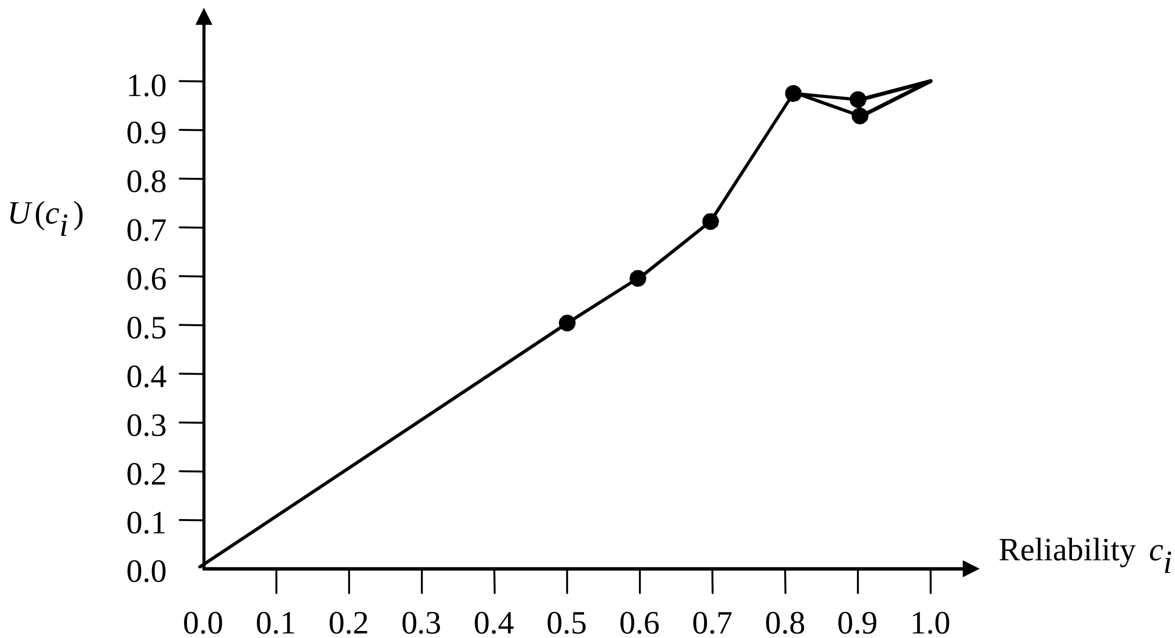

Figure 5.1: $\mathcal{D}$'s Utility of Reliability Based on End Point Gambles (MLE and Bayes).

Figure 5.1 suggests that $\mathcal{D}$'s utility, based on end point gambles, is linear in c_i, save for an upward jump at .8 followed by a slight drop of .01 at .9. This suggests that $\mathcal{D}$ is risk neutral for values of reliability up to .7 and is risk averse for values of reliability greater than .7. The drop in utility at .9 is an aberration that, hopefully, can be avoided by using adjacent point gambles. The essence of the message of Figure 5.1 is that there does not appear to be any gain in utility in going from a reliability of .8 to a reliability of .9. Thus to this $\mathcal{D}$, the strategic worthiness of the vehicle matures at a reliability of about .8; higher reliabilities are of little strategic consequence. A similar conclusion seems to be true with a consideration of adjacent point gambles, the data for which are given in Table 5.2 below.

Y_{ij}		p_{ij}						
		.3	.4	.45	.5	.55	.6	.7
c_i	.5	0	0	0	1	1	1	1
	.6	0	0	0	1	1	1	1
	.7	0	0	0	1	1	1	1
	.8	0	0	1	1	1	1	1
	.9	1	1	1	1	1	1	1

Table 5.2: $\mathcal{D}$'s Choices Under Adjacent Point Gambles.

An examination of the entries in Table 5.2 suggests that under adjacent point gambles, $\mathcal{D}$ tends to be risk prone in disposition towards the vehicle reliability for the values c_i considered here. $\mathcal{D}$'s shift from the risk aversion phenomenon of the entries of table 5.1 to the proneness phenomenon of Table 5.2 is intriguing. It could be attributed to the feature that it may be easier for $\mathcal{D}$ to contemplate end point gambles than adjacent point gambles.

c_i	$U(c_i)$	
	MLE	Bayes
.5	.330	.275
.6	.660	.551
.7	.880	.734
.8	.974	.825
.9	.997	.825

Table 5.3: $\mathcal{D}$'s Utility of Reliability Based on Adjacent Point Gambles.

The data of Table 5.2 was used to obtain $\mathcal{D}$'s utility of reliability via both the method of maximum likelihood and the Bayesian approach (described in Section A.1 (involving independent gamma priors). Table 5.3 shows the results. Plots of the $U(c_i)$ versus c_i obtained via the two methods are shown in Figure 5.2.

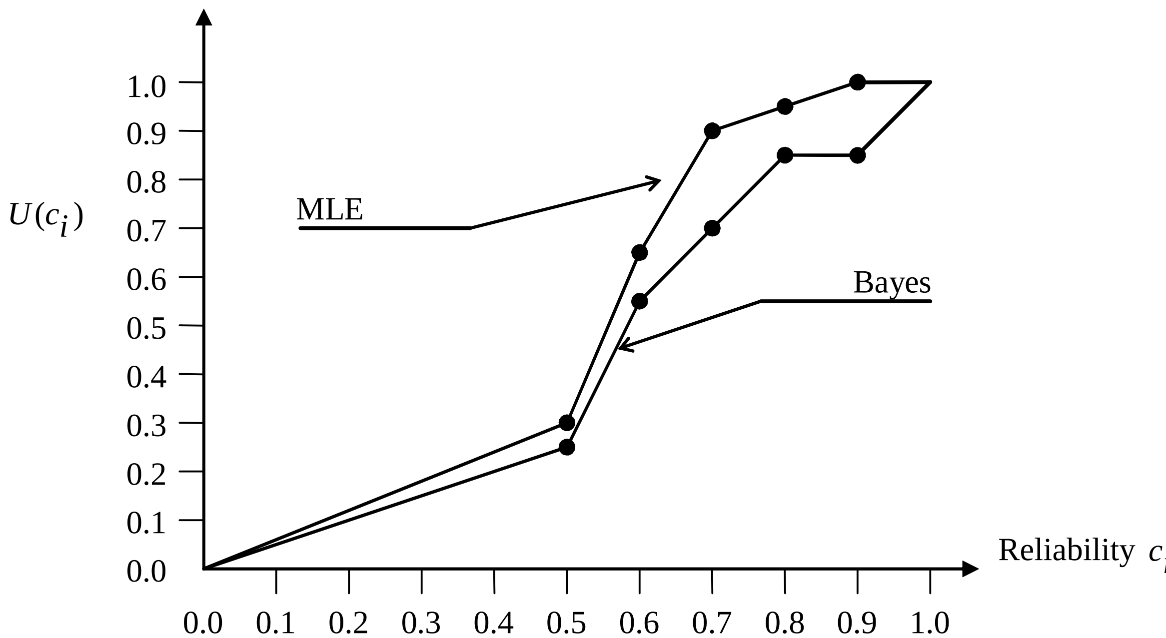

Figure 5.2: A Plot of $\mathcal{D}$'s Utility of Reliability Based on Adjacent Point Gambles.

The plots of Figure 5.2 suggests that for the priors chosen, the method of maximum likelihood yields uniformly higher values for the utility function than those yielded by the Bayesian approach, at least for the chosen priors. The differences however, are not substantial, and especially so, considering the fact that the Bayesian approach has a built in mechanism for incorporating the underlying uncertainties. Even though the raw entries of Table 5.2 suggest that $\mathcal{D}$ appears to be risk prone, the S-shaped nature of the plots of Figure 5.2 suggests that $\mathcal{D}$ tends to be risk neutral. The disparity between the intuitive conclusions that are formed by an inspection of the raw data, and those revealed by a formal analysis of these data is due to the smoothing of the data and the consistency that is enforced by the model of Equations (3.3) and (4.2).

Finally, the likes of Figure 5.2 suggest that there is little by the way of utility for reliabilities .5 or below, and that there is not much, if any, to be gained in going from a reliability of .8 to .9, at least under a Bayesian approach. Thus the need to push for reliability numbers such as .999 that General Officers usually tend to demand is unwarranted. Whereas we have not performed a reliability analysis of the vehicle in question, we have been told by $\mathcal{D}$ that the new vehicle reliability specifications tend to be in the form of 0 failures in 1000 operational hours, a requirement that is "fanciful and without connection to the real world".

6 Summary and Conclusions.

In this paper, we have advocated the point of view that a purpose for assessing the reliability and the survival functions is to help make decisions that mitigate risk. We have then cast the matter in a decision theoretic framework by leaning on a distinction between chance and probability. This has then been followed by proposing a general architecture for the utility of chance or reliability, including some boundary conditions. We have then proposed a statistical approach for assessing utilities based on binary choices between gambles and certainties. To facilitate this approach we have proposed a new choice (or item response theory) model that has features which parallel the Rasch Model in the sense that both models are indexed by a difference $-$ in our case $(p - c)$. Here, this difference is germane because the ease with which $\mathcal{D}$ can make a choice depends on the closeness of c and p. The *Grade of Membership Model* (GOM), discussed by Ershova, Fienberg and Joutard (2007), encompass aspects of the Rasch Model [cf. Ershova (2005)], and offers prospects for developing utility elicitation models more sophisticated than ours. The GOM model is difficult to appreciate and will require much thought to put it to work. Even though the material of Section 3 on using a Choice Model for utility elicitation is focused on the utility of reliability, the ideas therein are general enough for eliciting utilities in contexts that go beyond the special case of reliability.

There is much literature on utility theory and utility elicitation by economists, decision theorists, game theorists, mathematicians and statisticians. The names associated with these literatures are distinguished. Whereas we have endeavoured to gain an appreciation of as much of these works as is possible, it is likely that we may have missed some contributions that make our approach and our model for utility elicitation not new. But assuming that the above is not true, the work described here may open up avenues for future investigations.

Acknowledgements

Philip Wilson contributed to the development of the final elicitation model, and the numerical work in connection with the model. Professor Mounir Mesbah of University of Paris VI suggested the possibility of using the Rasch Model for eliciting utilities. The military problem was brought to our attention by Arthur Fries of the Institute for Defense Analysis (IDA). Robert Holcomb of the IDA served as the military planner whose utilities were assessed in Section 5. Stephen Fienberg drew our attention to the GOM model and in so doing has opened our eyes to things that we were unaware of. The comments of an insightful referee helped clarify some muddled thinking of the author; thank you! The author's research has been supported by the Office of Naval Research Contract N00014-06-1-0037 and The Army Research Office Grants W911NF-05-1-0209 and W911NF-09-1-0039 with The George Washington University.

Appendix A

A.1 *The Likelihood Function and Maximum Likelihood*

With respect to the notation of Section 3.1, with c fixed at c_i, and p at p_{ij}, the likelihood of α_i and β_i given the data y_{ij}, $j = 1, ..., n_i$, with $y_{ij} = 1$ or 0, is of the form:

$$\prod_{j=1}^{n_i} \left(\frac{1}{2} \left[1 + \text{sgn}(p_{ij} - c_i)|p_{ij} - c_i|^{\alpha_i} \right] \right)^{\beta_i y_{ij}} \left(1 - \left(\frac{1}{2} \left[1 + \text{sgn}(p_{ij} - c_i)|p_{ij} - c_i|^{\alpha_i} \right] \right)^{\beta_i} \right)^{1-y_{ij}} . \tag{A.1}$$

In writing out the above, we assume that given α_i and β_i, D's choices for the Y_{ij}'s, $j = 1, ..., n_i$ are independent over the p_{ij}'s. This tantamounts to assuming that in making a choice y_{ij}, $\mathcal{D}$ forgets his (her) previous choices y_{ik}, $k = j-1, j-2, ..., 2, 1$. One way to achieve this independence would be to select the p_{ij} in a random order with respect to the j's. By avoiding choosing $c_i = 1$ or 0, and $p_{ij} = 1$ or 0, in the elicitation process, we can ensure that the likelihood will not involve the boundary conditions of Figure 3.2. Equation (A.1) can be maximized numerically to yield $\widehat{\alpha}_i$ and $\widehat{\beta}_i$ as the maximum likelihood estimators of α_i and β_i, respectively.

A.2 *Bayes Inference for* α *and* β

An examination of Figure 3.4 shows that a large value of α causes Π to be steeper for $(p - c)$ in the vicinity of -1 or $+1$, and flatter for $(p - c)$ close to zero, than a small value of α. This type of change characterizes a $\mathcal{D}$ who switches slowly from preferring a certain outcome to the gamble.

Such a $\mathcal{D}$ exhibits a poorer ability to discriminate between gambles that $\mathcal{D}$ considers to be worse than the certain outcome than those gambles that are better. Thus α may be viewed as a parameter that encapsulates $\mathcal{D}$'s ***ability to discriminate between gambles***, with higher values of α representing a lower ability to discriminate.

Similarly, an examination of Figure 3.4 a) or b) shows that the parameter β encapsulates $\mathcal{D}$'s attitude to risk, $\beta < (=) > 1$ representing a risk prone (neutral) averse $\mathcal{D}$.

It is reasonable to suppose that $\mathcal{D}$'s ability to discriminate between gambles is independent of $\mathcal{D}$'s disposition towards risk. Thus α and β can be treated as being independent. With α, $\beta > 0$, it is reasonable to suppose that $\alpha[\beta]$ has a gamma distribution with scale $l[s]$ and shape $k[r]$. Choosing $k = r = 2$ and $l = s = .5$, $\Pi(\alpha, \beta)$ the joint prior on α, β is of the form $\Pi(\alpha, \beta) \propto \beta\alpha \exp(-2(\alpha + \beta))$. Consequently, $\Pi(\alpha_i, \beta_i) \propto \beta_i\alpha_i \exp(-2(\alpha_i + \beta_i))$.

Combining $\Pi(\alpha_i, \beta_i)$ with the likelihood (A.1) gives us the posterior distribution of α_i, β_i, denoted $\Pi(\alpha_i, \beta_i | \mathbf{y}, \mathbf{p}_i, c_i)$, where $\mathbf{y} = (y_{i1}, ..., y_{i,n_i})$ and $\mathbf{p}_i = (p_{i1}, ..., p_{i,n_i})$. The ingredients necessary to solve Equation (4.4) are at hand, with $(0, \infty)$ as the limits of integration. The integration has been done numerically using the bisection method.

References

Barlow, R. E., D. J. Bartholomew, J. M. Bremner, and H. D. Brunk (1972). *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression.* Wiley, New York.

Becker, G. M., M. H. DeGroot, and J. Marshak (1963). 'Stochastic Models of Choice Behavior.' *Behavioral Science*, 9, 41-55.

Becker, G. M., M. H. DeGroot, and J. Marshak (1964). 'Measuring Utility by a Single-Response Sequential Method.' *Behavioral Science*, 9, 226-232.

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

Ershova, E. (2005). 'Comparing Latent Structures of the Grade of Membership, Rasch, and Latent Class Models.' *Psychometrica,* 70, 4, 619-628.

Ershova, E., Fienberg, S., and Joutard, C. (2007). 'Describing Disability Through Individual-Level Mixture Models for Multivariate Binary Data.' *The Annals of Applied Statistics*, 1, 2, 502-537.

Farquhar, P. H. (1984). 'Utility Assessment Methods.' *Management Science*, 30 (11): 1283-1300.

Good, I. J., and W. I. Card (1971). 'The Diagnostic Process with Special Reference to Errors.' *Methods of Information in Medicine,* 10, 3, 176-188.

Hull, J., P. G. Moore and H. Thomas (1973). 'Utility and its Measurement.' *Journal of the Royal Statistical Society A* 136 (2): 226-247.

Kolmogorov, A. N. (1969). 'The Theory of Probability.' *Mathematics, Its Content, Methods and Meaning* (ed. Aleksandrov, A. D., Kolmogorov, A. N., Lavrent'ev, M. A.), 2, 3, 229-264. MIT Press, Cambridge, MA.

Lindley, D. V. (1976). 'A Class of Utility Functions.' *The Annals of Statistics*, 4, 1, 1-10.

Lindley, D. V., and M. R. Novick (1979). 'Fixed-State Assessment of Utility Functions.' *Journal of the American Statistical Association*, 306-311.

Lindley, D. V., and L. D. Phillips (1976). 'Inference for a Bernoulli Process (A Bayesian View).' *The American Statistician*, 30, 112-119.

Lindley, D. V., and N. D. Singpurwalla (2002). 'On Exchangeable, Causal and Cascading Failures.' *Statistical Science*, 17, 2, 209-219.

Mesbah, M., B. F. Cole and M. T. Lee (2002), Editors. *Statistical Methods for Quality of Life Studies.* Kluwer Academic Publishers, Boston.

Mesbah, M., and N. D. Singpurwalla (2008). 'A Bayesian Ponders the Quality of Life.' *Statistical Models and Methods for Biomedical and Technical Systems*, 373-384 (F. Vonta, Editor). To appear.

Meyer, R. F., and J. W. Pratt (1968). 'The Consistent Assessment and Fairing of Preference Functions.' *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, 3, 270-278.

Mosteller, F., and P. Nogee (1951). 'An Experimental Measurement of Utility.' *Journal of Political Economy*, 59: 371-404.

Popper, K. (1957). 'The Propensity Interpretation of the Calculus of Probability and of the Quantum Theory.' *In Observation and Interpretation* (Korner and Price, eds.), Bultersworth Scientific Publications, 65-70.