The Operational-Bayesian Approach In Reliability Theory

Richard E. Barlow[1] and Max B. Mendel[2]

Abstract: We emphasize the derivation of likelihood models starting from a well specified problem of interest and finite populations. "Parameters" are given operational meaning. In particular, parameters are specified in terms of different forms of energy. Examples relevant to reliability theory are used to illustrate ideas. Examples in engineering probability are given.

Key words: Exchangeability, isotropicity, indifference, predictive probability, symmetry.

1. Introduction.

The key ingredients in this approach are an emphasis on 1) the operational meaning of questions posed and 2) the related field of specialization within which questions are posed. Deterministic and logical relationships basic to the field of application are used in part to develop appropriate probability models.

This approach requires a deeper understanding of the questions posed than the traditional approach and is consequently less mechanical in its application. P. A. W. Bridgman (1927) was perhaps first to popularize use of the term "operational" in the context of physical phenomena. De Finetti was one of the first to consider this approach in the context of probability theory. Dawid (1982) followed de Finetti's point of view by deriving subjective probability models based on observable random quantities. Mendel (1989) used the same approach but more closely followed Savage (1954) in placing emphasis on the decision problem. Mendel advocates the "indifference principle" among acts as a means for constructing likelihood probability models.

We begin an operational-Bayesian analysis with a question or questions which require selection of an act or a decision. Like Lindley (1985 p. 4) we make no distinction between action and decision. The question or decision problem must contain terms which have operational meaning. That is, a bet based on the outcome related to our decision must have the property that it can be settled, at least in principle. We think about a value function relative to our problem. Value functions suggest "parameters" of interest. Although in practice it may be difficult to specify a value function, thinking about one helps identify "parameters"; i.e., functions of relevant but unknown observables. In general, value functions will

depend on these parameters. Parameters may be additive functions of observables. Examples of relevant "parameters" are energy, mass, force, money, etc. The value function and associated parameters together will determine the relevant sample space.

2. The Indifference Principle.

We propose that probability is essentially an analytical tool based on judgment. As an analytical tool, it plays a key role in decision making and forecasting. Its assessment, based on judgment, is the most difficult aspect. In the absence of a well defined, real problem, we can at most determine a conditional probability; e.g., a likelihood function. Even this probability function can only apply to a particular class of problems.

To illustrate the role of indifference in probability modeling, we will consider several cases where the parameter of interest is energy. Although energy can take many forms, it is conserved and it is additive.

2.1. Kinetic Energy of a Particle in Motion.

Consider a particle in an ideal gas in stable equilibrium. The gas particle is moving in space with a velocity whose coordinates in three mutually orthogonal directions can be represented by a velocity vector

$$(\dot{X}_1, \dot{X}_2, \dot{X}_3).$$

The condition of stable equilibrium implies that the probability distribution of the velocity does not depend on the choice of the coordinates and hence that this random velocity vector is rotationally symmetric. Since kinetic energy is proportional to the sum of squares of the coordinate velocities when the mass is known, the parameter of interest corresponds, in this case, to the square of the ℓ^2–norm, namely

$$\sum_{i=1}^{3} \dot{X}_i^2 = 3\theta^2.$$

Hence we would be *indifferent* between any two velocity vectors having the same sum of squares. This means that we would give the *same* predictive probability (or probability density) for any two such vectors. It can be shown that the bivariate density for two coordinate velocities, conditional on the kinetic energy, is *determined* by this indifference. The conditional bivariate density for two coordinate velocities can be shown to be

$$p(\dot{x}_1, \dot{x}_2 \mid \theta) = \frac{1}{2\pi(3\theta)^2} \left[1 - \frac{\dot{x}_1^2 + \dot{x}_2^2}{3\theta^2}\right]^{-\frac{1}{2}} \tag{1}$$

for $\dot{x}_1^2 + \dot{x}_2^2 \leq 3\theta^2$ and 0 elsewhere.

2.2. ℓ^q-isotropic distributions.

If the joint distribution of $\mathbf{X}_N$ is absolutely continuous, the parameter of interest may correspond to the ℓ^q-norm of $\mathbf{X}_N$. Considering the general case, suppose that our predictive probability function for $\mathbf{X}_N$ has the property that $p(\mathbf{x}_N) = p(\mathbf{y}_N)$ when $\mathbf{x}_N$ and $\mathbf{y}_N$ have the same ℓ^q-norm. We are in effect indifferent regarding vector outcomes having the same ℓ^q-norm. Even though our value function may depend on $\mathbf{X}_N$ only through the ℓ^q-norm, it does not follow that we would always be indifferent to vector outcomes having the same ℓ^q-norm. As Lindley has pointed out, "My opinion about the data, were I to know the parameter, exists irrespective of the decision that might hinge on the data. To deny this would to fly in the face of scientific thinking which holds that pure science is meaningful."

In Barlow and Mendel (1992) we considered aging as well as the ℓ^1-norm in deriving the likelihood.

The distribution of $\mathbf{X}_N$ given $\sum_{i=1}^{N} |X_i|^q = N\theta^q$ is singular, although $p(\mathbf{x}_n \mid \theta)$ exists for $1 \leq n < N$. In Mendel (1989), $p(\mathbf{x}_n \mid \theta)$ is derived.

We have

$$p(\mathbf{x}_n | \theta) = \frac{\Gamma(\frac{N}{q})}{\Gamma^n(\frac{1}{q})\Gamma(\frac{N-n}{q})(N\theta^q)^{n/q}} \left(\frac{q}{2}\right)^n \left[1 - \frac{\sum_{i=1}^{n} |x_i|^q}{N\theta^q}\right]^{\frac{N-n}{q} - 1} \qquad (2),$$

for $\sum_{i=1}^{n} |x_i|^q \leq N\theta^q$ and 0 elsewhere.

For $N \to \infty$ and $q = 2$ we obtain the product of n independent normal densities where the parameter is the variance.

3. Strength of Materials.

The theory of static stress was, in many respects, developed in the last century. A very good reference to this subject is the book by Hoffman and Sachs (1953). However, the operational approach to the assessment of probability forms on the space of stress tensors is a very new subject still under development. More advanced results than those described below will appear in a forthcoming monograph by Mendel and Chick (1994).

3.1. Tensile Tests for Isotropic Ductile Materials.

Lindquist (1992) makes use of the ℓ^1-norm model relative to the strength of ductile materials. When the distortion energy u_1 of a material reaches a threshold

value the material will yield. The total elastic strain energy u absorbed by a volume of stressed material is the sum of a dilatational energy u_2, i.e., the elastic energy that causes a change in the material's volume, and the distortion energy, u_1, i.e., the elastic energy that results in shear deformations.

Shear deformations are those which distort the material so that initial right angles no longer remain right angles.

Using this theory, yield predictions can be made for structural members under complex states of stress based only on the uniaxial or torsional stress test data. One simply calculates the distortion energy capacity for the material using the simple yield strength test results and uses this value to make yield predictions for members under more complicated states of stress.

The formula used to calculate distortion energy for uniaxial stress is

$$u = \frac{\sigma^2}{6G} \tag{3}$$

where σ is the uniaxial stress and G is the shear modulus (a material constant). In this case, $u = u_1$.

Equation (3) follows from Hooke's law which is an empirical relationship between stress and strain.

If $p(\mathbf{u}_N) = p(\mathbf{u}'_N)$ when $\sum_{i=1}^{N} u_i = \sum_{i=1}^{N} u'_i$ and the parameter of interest is $\sum_{i=1}^{N} u_i = N\theta$, then

$$p(\mathbf{u}_n \mid \theta) = \frac{N-1}{N\theta} \cdots \frac{N-n}{N\theta} \left[1 - \frac{\sum_{i=1}^{n} u_i}{6GN\theta} \right]^{N-n-1}.$$

This is the ℓ^1–norm form of (2) for positive random quantities.

By a change of variable we obtain the joint probability density for an exchangeable sample of items resulting in stress test data, $\Sigma_n = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, namely

$$p(\Sigma_n | \theta) = \left(\prod_{i=1}^{n} \frac{2\sigma_i(N-i)}{6GN\theta} \right) \left[1 - \frac{\sum_{i=1}^{n} \sigma_i^2}{6GN\theta} \right]^{N-n-1} \tag{4.}$$

When $N \to \infty$ this is the product of n Weibull densities with shape parameter 2.

If we had judged $p(\Sigma_n) = p(\Sigma'_n)$ when $\sum_{i=1}^{N} u_i = \sum_{i=1}^{N} u'_i$ then the likelihood for test stresses would have been based on the ℓ^2–norm. In this case the likelihood would not be given by (4). The first indifference assumption seems more reasonable to us unless additional information is available.

3.2. Elastic Energy of Isotropic Materials: The State of Plane Stress.

In section 3.1 we noted that we need only use the distortion energy capacity, u_1, of a ductile material to make yield predictions based on tensile tests. Total elastic strain energy, however, is the sum of the distortion energy, u_1, and the dilatational energy, u_2, of a material. We will now make use of ℓ^2−norm isotropicity to analyze multivariate data concerning the state of plane stress relative to both u_1 and u_2. In this example, the population size $N = 1$.
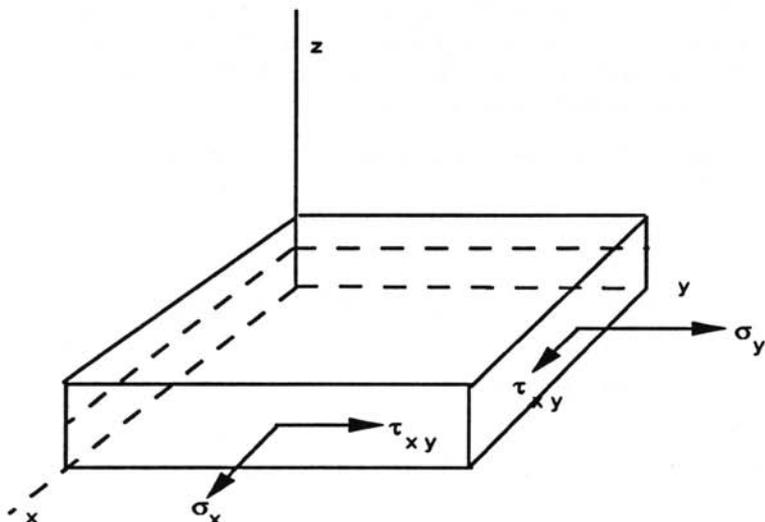

Figure 1. Plane stress

In Figure 1 we illustrate the case of plane stress at a "point" in a material relative to an arbitrarily chosen coordinate system. In the case of plane stress we assume $\sigma_z = \tau_{zx} = \tau_{zy} = 0$ so that the stress tensor can be represented by the 2×2 matrix

$$S = \begin{pmatrix} \sigma_x & \tau_{yx} \\ \tau_{xy} & \sigma_y \end{pmatrix}$$

where σ_x, $\tau_{xy} = \tau_{yx}$ and σ_y are dependent random quantities. Material isotropicity implies $\tau_{xy} = \tau_{yx}$. Since our coordinate representation for the stress tensor is symmetric and non-singular, it can be diagonalized to

$$S' = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$$

where $\sigma_1 \geq \sigma_2$ are the principal stresses. The principal stresses, but not the principal axes of the so-called stress-quadric, are invariant with respect to the coordinate system chosen. The distortion energy in terms of the principal stresses is

$$u_1 = \frac{1 + 2G}{6E} [\sigma_1^2 + \sigma_2^2 - \sigma_1 \sigma_2] \tag{5}$$

where E is Young's modulus and G is the shear modulus.

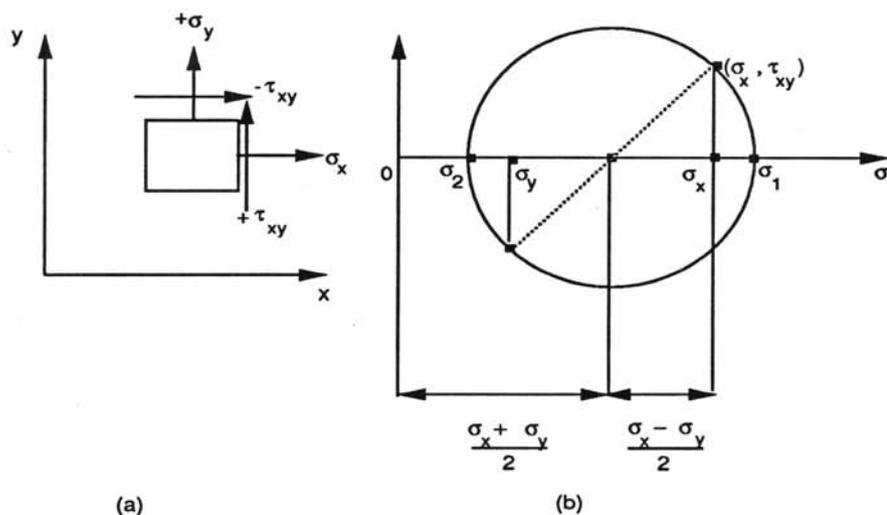<div align="center">(a)   (b)</div>

<div align="center">Figure 2. Mohr's Circle of Stress</div>

The dilatational energy is

$$u_2 = \frac{1 + 2G}{6E}\left[\sigma_1 + \sigma_2\right]^2 \tag{6}$$
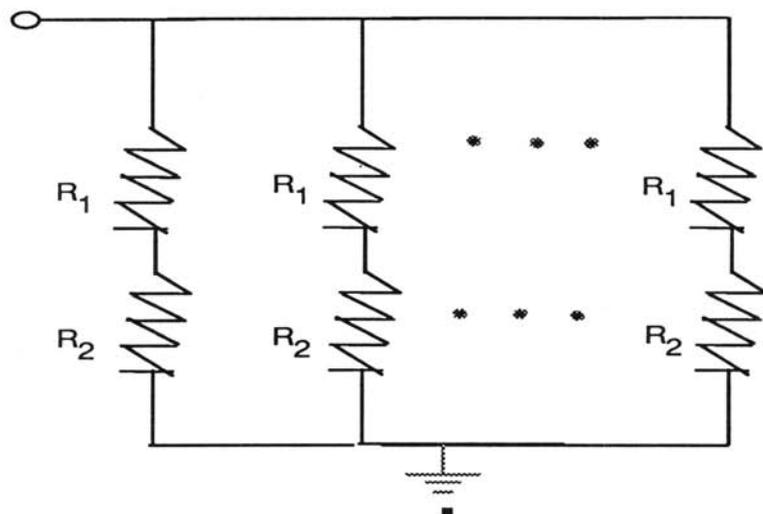

<div align="center">Figure 3. Resistors in Series</div>

It is easy to see that u_1, u_2 determine $\sigma_1 \geq \sigma_2$ and vice versa and that u_1, u_2 are invariants.

Conditional on u_1 and u_2 (or σ_1 and σ_2) we draw the so-called "Mohr circle" passing through $(\sigma_1, 0)$ and $(\sigma_2, 0)$ with center

$$\frac{\sigma_1 + \sigma_2}{2}.$$

Mohr's circle of stress (see Figure 2b) is widely used in practice for stress transformations.

Provided the initial stresses are given, the Mohr circle method is applicable whether the material behaves elastically or plastically.

Figure 2a shows the states of stress on a plane (in 3 dimensions) that would be determined by the x-axis. σ_x is the stress normal to this plane while τ_{xy} is the stress parallel to this plane (the shearing stress). The dilatational energy, u_2, determines the center of the circle, namely

$$\frac{\sigma_1 + \sigma_2}{2} = \sigma_m.$$

u_1 and u_2 together determine $\sigma_1 \geq \sigma_2$ which in turn determines the radius of the circle, namely

$$\tau = \sigma_1 - \frac{\sigma_1 + \sigma_2}{2} = \frac{\sigma_1 - \sigma_2}{2}.$$

The radius of the circle corresponds to the maximum possible shear stress at the given point for planes orthogonal to the x-y plane. The point on the circle (σ_x, τ_{xy}) corresponds to the state of stress on the plane determined by the x-axis which intersects the right hand face of the square in Figure 2a. In general, points on the circle correspond to the normal stress and the shearing stress on planes orthogonal to the x-y plane. We are indifferent to normal and shearing stresses corresponding to points on the circle by the laws governing the equilibrium of forces.

Since

$$\tau^2 = \tau_{xy}^2 + (\sigma_x - \sigma_m)^2$$

$p(\sigma_x|\sigma_1, \sigma_2)$ can be calculated using ℓ^2–norm isotropicity to give

$$p(\sigma_x|\sigma_1, \sigma_2) = \frac{1}{\pi\tau}\left[1 - \left(\frac{\sigma_x - \sigma_m}{\tau}\right)^2\right]^{-1/2} \tag{7}$$

for $\sigma_2 \leq \sigma_x \leq \sigma_1$.

4. Lifetime Models.

The previous probability models were for multivariate observations on a single particle or a state of stress at a specified point in a body of isotropic material. The following example pertains to exchangeable observations on similar systems or units.

4.1. Lifetimes of Resistors in Series.

This section proposes an energy approach using an "energy potential", rather than the classical failure rate, to specify classes of lifetime models. First, we consider the univariate case and the operating energy usage of N exchangeable resistors on the same voltage source. Such resistors could represent heating elements, light bulbs on a voltage source, turbines on a water reservoir of constant height; etc.

Let a_i be the i-th unit's effective age which may be measured in units other than calendar time; e. g. cycles per second, etc. We call $\dot{a}_i$ the aging rate for the i-th unit. This could, for example, be a derivative with respect to calendar time.

The energy usage of a single resistor with resistance R_i ohms is V^2/R_i watts and its energy usage potential function becomes:

$$c(a_i) = V^2/R_i$$

watts (or joules per second) if this is constant in the unit's age, a_i. We call $\dot{a}_i$ the aging rate for the i-th unit. This could, for example, depend on a specified maintenance policy. In any event it is specified although the unit's lifetime, x_i, is unknown. The energy potential evaluated at a unit's lifetime, x_i is

$$\lambda_i = c(x_i).$$

The lifetime energy usage for unit i is

$$\Lambda_i = \int_0^{x_i} c(a_i)da_i.$$

If a unit's energy potential does not depend on the life of other units, then

$$\sum_{i=1}^{N} \Lambda_i$$

is the total lifetime energy used by a population of N units. If the parameter of interest is the average, namely

$$\frac{\sum_{i=1}^{N} \Lambda_i}{N} = \theta$$

and we are indifferent to vectors of N lifetime energies having the same sum, then for

$1 \leq n < N$ and by ℓ^1 isotropicity, we have

$$p(\mathbf{x}_n|\theta) \propto \left(\frac{\prod_{i=1}^{n} \lambda_i}{\theta}\right) \left[1 - \frac{\sum_{i=1}^{n} \Lambda_i}{N\theta}\right]^{N-n-1}. \tag{8}$$

4.2. Multivariate Lifetimes of Resistors in Series.

A unit's energy potential may depend on the lifetimes of other units, thus introducing an energy dependence of one unit on other units. For simplicity, consider N exchangeable systems each with 2 resistors in series. We generalize the notation by letting a_i denote the ages of the components of the ith series system and,

$$\mathbf{a} = (a_1, a_2, \ldots, a_N) = (\{ \begin{matrix} a_{11} \\ a_{12} \end{matrix} \} \cdots \{ \begin{matrix} a_{N1} \\ a_{N2} \end{matrix} \})$$

and similarly for the lifetimes $\mathbf{x}$, the aging rate $\dot{\mathbf{a}}$ and energy potential $\mathbf{c}$.

The energy usage of a resistor now also depends on the lifetime of the other in the pair. If resistors fail by shorting, then the power dissipated in the first resistor of a pair is either

$$V^2/R_1 \quad \text{or} \quad [VR_1/(R_1 + R_2)]^2/R_1$$

watts, depending on whether the other of the pair has failed or not, respectively. The power dissipated in the second resistor depends on the lifetime of the first in a similar way. The energy potential functions of the components of the ith system become:

$$c_{i1}(a_{i1}) = \frac{V^2}{R_1} \Big[1 - \frac{R_2^2 + 2R_1 R_2}{(R_1 + R_2)^2} \phi(a_{i1}, x_{i2}) \Big] = \kappa_{11} + \kappa_{12} \phi(a_{i1}, x_{i2})$$

$$c_{i2}(a_{i2}) = \frac{V^2}{R_2} \Big[1 - \frac{R_1^2 + 2R_1 R_2}{(R_1 + R_2)^2} \phi(x_{i1}, a_{i2}) \Big] = \kappa_{22} + \kappa_{21} \phi(x_{i1}, a_{i2}).$$

Here ϕ is the structure function of a series system[1]

Both these potentials depend on the lifetime of the other unit. Although their components are energy dependent, the systems are, however, energy independent and even identical with respect to energy.

The lifetime energy potential of the ith system is obtained by evaluating the given energy potential at the lifetimes x_i of its components:

$$\{ \begin{matrix} \lambda_{i1} \\ \lambda_{i2} \end{matrix} \} = \{ \begin{matrix} k_{11} + k_{12} \phi(x_{i1}, x_{i2}) \\ k_{22} + k_{21} \phi(x_{i1}, x_{i2}) \end{matrix} \}$$

---

[1] In particular,

$$\phi(a_1, x_2) = \begin{cases} 1 & \text{if } a_1 < x_2, \\ 0 & \text{otherwise.} \end{cases}$$

and the Jacobian $|\lambda_i| = \lambda_{i1}\lambda_{i2}$ is found by multiplying these potentials. The absolute value of the lifetime energy, Λ_i, is obtained by taking the dot product of the energy potential with the aging rates $\dot{a}_i$ and integrating up to the lifetimes:

$$||\Lambda_i|| = \Lambda_{i1} + \Lambda_{i2} = \int_0^{x_{i1}} c(a_{i1})da_{i1} + \int_0^{x_{i2}} c(a_{i2})da_{i2}$$

$$= k_{11}x_{i1} + k_{22}x_{i2} + (k_{12} + k_{21})\min(x_{i1}, x_{i2}).$$

Consider the $2N$ random quantities

$$\{\Lambda_{11}, \Lambda_{12}, \cdots, \Lambda_{N1}, \Lambda_{N2}\}.$$

Suppose we are indifferent to all such sequences having the same lifetime energy average, θ, where

$$\theta = \frac{\sum_1^N ||\Lambda_i||/2}{N}.$$

The determination of the likelihood proceeds in much the same way as in the univariate case, the only difference being that the change of variable involves the determinant

$$|\lambda_i| = \lambda_{i1} \, \lambda_{i2}$$

of the Jacobian λ_i of the coordinate-transformation map $\Lambda_i = \Lambda_i(x_i)$ taking 2-vectors x_i of component lifetimes into 2-vectors Λ_i of component lifetime energies. Using ℓ^1 isotropicty we obtain:

$$p(\mathbf{x}_n|\theta) \propto \left(\prod_{i=1}^n \frac{|\lambda_i|}{\theta^2}\right)\left[1 - \frac{\sum_{i=1}^n ||\Lambda_i||/2}{N\theta}\right]^{2(N-n)-1}. \qquad (9)$$

With this notation, the class of densities p for the lifetimes of the first n systems can be represented as a mixture over known likelihoods $p(\mathbf{x}_n|\theta)$:

$$p(\mathbf{x}_n) = \int_0^\infty p(\mathbf{x}_n|\theta)\pi(d\theta).$$

However, now $\mathbf{x}_n$ consists of n 2-dimensional vectors and, in this sense, the predictive density p and the likelihood $p(\mathbf{x}_n|\theta)$ are multivariate distributions. Notice that although the model is multivariate, the parameter is still a one-dimensional quantity.

Infinite batches of such units follow a Bayesian version of the so-called Freund model, after Freund [1961] who proposed the classical version for two-component systems. Although his model has four unknown classical parameters, the only unknown in the present model is the average lifetime energy θ.

5. Conclusions.

The likelihood model used to analyze data is surely as subjective as any "prior" probabilities. However, to make this model assessment we need to first concentrate on the question of interest and its relevant field of application. From the operational viewpoint statistical inference should be more closely associated with particular fields of application. This must occur if statistics is to be more than a cosmetic veneer attached to scientific papers.

6. References.

1 Barlow, R. E. and M. B. Mendel, "De Finetti-Type Representations for Life Distributions," *Journal of the American Statistical Association*, Vol. **87**, Number 420, pp. 1116-1122, (1992).

2 Bridgman, P., *The Logic of Modern Physics*, New York, The Macmillan Co, (1927).

3 Dawid, A. P., "Intersubjective Statistical Models." *Exchangeability in Probability and Statistics*, G. Koch and F. Spizzichino (editors), North-Holland Publishing Co, (1982).

4 De Finetti, B., "Foresight: Its Logical Laws, Its Subjective Sources." *Annales de l'Institute Henri Poincaré* **7**, 1-68. English translation in *Studies in Subjective Probability*. Kyburg, H.E. Jr., and Smokler, H.E., eds., 1980, (2nd ed.), Huntington, New York: Robert E. Krieger Pub. Co., pp. 53-118, (1937).

5 Freund, J. E., "A Bivariate Extension of the Exponential Distribution." JASA 56: 971-977. Hoffman, O. and G. Sachs (1953). *Introduction to the Theory of Plasticity for Engineers*, New York, McGraw-Hill, (1961).

6 Lindley, D. V., *Making Decisions*, J. Wiley & Sons, New York, 2nd edition,(1985).

7 Lindquist, E., "Strength of Materials and the Weibull Distribution." UC Berkeley Technical Report ESRC 92-32, Engineering Systems Research Center. To appear in the journal *Probabilistic Engineering Mechanics*, (1992).

8 Mendel, M. B. and S. Chick, *The Geometry and Calculus of Engineering Probability*, Springer-Verlag. Mendel, M. B. (1991). "An Economic Approach to Lifetime Modelling." UC Berkeley Technical Report ESRC 91-6, Engineering Systems Research Center, (1994)

9 Mendel, M. B., *Development of Bayesian Parametric Theory with Applications to Control*, MIT Ph.D. thesis, (1989).

10 Mendel, M. B., "Bayesian Parametric Models for Lifetimes." *Bayesian Statistics 4*. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds., Claredon Press, Oxford, 697-705, (1992).

11 Savage, L. J., *The Foundations of Statistics*, New York: J. Wiley & Sons, (1954).