Beta Distributions and Its Applications, 29(1), 1-37 (2003)

Parameter Specification of the Beta Distribution and its Dirichlet Extensions Utilizing Quantiles

J. René van Dorp and Thomas A. Mazzuchi

(Submitted January 2003, Revised March 2003)

To appear in Handbook of Beta Distributions and Its Applications — Publisher, Marcel Dekker, Inc.
Editors Gupta, A.K. and Nadarajah, S.

Parameter Specification of the Beta Distribution and its Dirichlet Extensions Utilizing Quantiles

J. René van Dorp and Thomas A. Mazzuchi
Department of Engineering Management and
Systems Engineering
1776 G Street, Suite 110
The George Washington University,
Washington, DC 20052

(Submitted January 2003, Revised March 2003)

I. Introduction

The time-honored 2 parameter Beta distribution

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\ x^{a-1}(1-x)^{b-1}, a, b > 0, x \in [0,1], \tag{1}$$

which is the main subject matter of this volume, is well known in Bayesian methodology as a prior distribution on the success probability p of a binomial distribution (see, e.g. Carlin & Louis (2000)). Many authors (see, e.g. Gavaskar (1988)) have quoted the suitability of a Beta random variable X in different applications due to its flexibility. The transformation $Y = -Ln(X)$ transforms the $[0,1]$ support of X into the support $[0,\infty)$ of Y, while still inheriting the flexibility of X. Hence, the use of the Beta distribution as a prior distribution is by no means restricted to a bounded domain. For example, Van Dorp (1998) utilizes the above transformation to specify a prior distribution on the positive shape parameter of a Weibull distribution.

To appear in Handbook of Beta Distributions and Its Applications
Editors Gupta, A.K. and Nadarajah, S.
Publisher, Marcel Dekker, Inc.

Amongst m-variate extensions of the Beta distribution (i.e. m - dimensional joint distributions with Beta marginals) the Dirichlet distribution (see, e.g., Kotz et al. (2000)),

$$\frac{\Gamma\left(\sum_{i=1}^{m+1}\theta_i\right)}{\prod_{i=1}^{m+1}\Gamma(\theta_i)}\left\{\prod_{i=1}^{m}x_i^{\theta_i-1}\right\}\left\{1-\sum_{i=1}^{m}x_i\right\}^{\theta_{m+1}-1}, \tag{2}$$

where $x_i \geq 0,\ i=1,\ldots,m,\ \sum_{i=1}^{m}x_i \leq 1, \theta_i > 0,\ i=1,\ldots,m+1$ has enjoyed wide popularity in Bayesian methodology (see, e.g., Cowell (1996), Johnson & Kokalis (1994) and Dennis (1998)). Application areas include Reliability Analysis (see, e.g., Kumar & Tiwari (1989), Coolen (1997) and Neath & Samaniego (1996)), Econometrics (See, e.g. Lancaster (1997), and Forensics (see, e.g., Lang (1995)). The use of the m-variate Ordered Dirichlet distribution (see, e.g., Wilks (1962))

$$\frac{\Gamma\left(\sum_{i=1}^{m+1}\theta_i\right)}{\prod_{i=1}^{m+1}\Gamma(\theta_i)}\prod_{i=1}^{m+1}(x_i-x_{i-1})^{\theta_i-1}, \tag{3}$$

where $x_0 \equiv 1-x_{m+1} \equiv 0,\ 0 < x_{i-1} < x_i < 1, i=1,\ldots,m+1,\ \theta_i > 0,\ i=1,\ldots,m+1,$ in Bayesian applications is less prevalent and the distribution is generally less well known. To the best of our knowledge, applications of (3) have been limited so far to reliability analysis problems (see, e.g., Van Dorp et al. (1996), Van Dorp et al. (1997), Erkanli et al. (1998), Van Dorp and Mazzuchi (2003)). A fundamental difference between the Dirichlet distribution (defined on a simplex) and the Ordered Dirichlet distribution (defined on the upper pyramidal cross section of the unit hyper cube) is their support. Figure 1 below illustrates this difference for the bivariate case. Both the Dirichlet and Ordered Dirichlet random vector $\underline{X}=(X_1,\ldots,X_m)$ inherit the flexibility of its Beta marginals $X_i, i=1,\ldots,m$. Transforming these m-variate extensions by means of the transformation $Y_i = -Ln(X_i), i=1,\ldots,m$ allow for flexible prior distributions not restricted to the unit hyper cube. Van Dorp and Mazzuchi (2003) used a similar transformation to define a prior distribution on a set of ordered failure rates on $[0, \infty)^m$ via an ordered Dirichlet distribution.

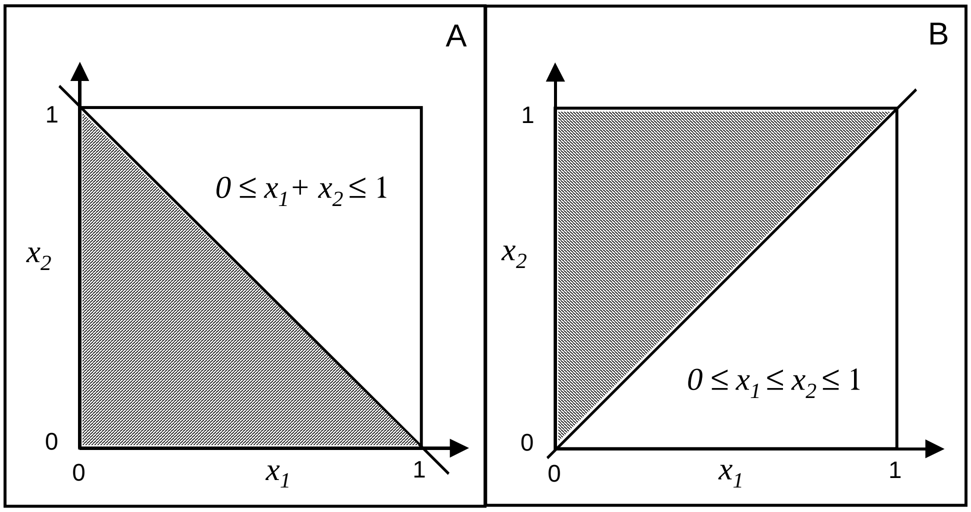

Figure 1. A: Support of a Bivariate Dirichlet Distribution

B: Support of a Bivariate Ordered Dirichlet Distribution

Practical implementation of subjective Bayesian methods involving the Beta distribution and its Dirichlet extensions evidently require the specification of their parameters. To avoid being incoherent in these Bayesian analyses, the specification of these prior parameters preferably should not rely on classical estimation techniques which use data, such as maximum likelihood estimation or the method of moments. The specification of the prior parameters ought to be based on expert judgment elicitation. To define the prior parameters, expert judgment about quantities of interest are elicited and equated to their theoretical expression for central tendency such as mean, median, or mode (see e.g. Chaloner and Duncan (1983)). In addition, some quantification of the quality of the expert judgment is often given by specifying a variance or a probability interval for the prior quantity. Solving these equations generally would lead to the required parameter estimates.

Methods for eliciting the parameters of a Beta distributions have focused on eliciting: (a) a measure of central tendency such as the mean and a measure of dispersion such as the variance (see, e.g., Press (1989)), (b) the mean and a quantile (see, e.g. Martz and Waller (1982)) or (c) equivalent observations (e.g. Cooke, 1991). Elicitation of the mean (and certainly the variance), however, requires a level of cognitive processing that elicitation procedures which demand it, may well produce little more than random noise (see, Chaloner and Duncan (1983)). Hence, it is desirable, for designing of a meaningful elicitation procedure for engineers, that elicited information can be easily related (i.e. involving little cognitive processing) to observables (see, e.g., Chaloner and Duncan (1983)). While Chaloner and Duncan (1983), (1987) elicit Beta prior parameters and Dirichlet prior parameters by relating these parameters to the modes of observable random variables and non-uniformity around their modes, they also advocate the use of quantiles, such as the median and a lower quantile, for the elicitation of prior parameters. An additional advantage of eliciting quantiles is that it allows for the use of betting strategies in an *indirect* elicitation procedure (see e.g. Cooke (1991)).

This chapter addresses the problem of specification of prior parameters of a Beta distribution and its Dirichlet extensions above via quantile estimates. It is envisioned that these quantile estimates are elicited utilizing expert judgment techniques thereby allowing coherent and practical application of the Beta distribution and its Dirichlet extensions in Bayesian Analyses. Solving for the parameters of these prior parameters via quantile estimates involves using the incomplete Beta function $B(x|a,b)$ given by

$$B(x \mid a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x p^{a-1}(1-p)^{b-1} dp, \tag{4}$$

where $a > 0,\ b > 0$. The incomplete Beta function $B(x|\,a,b)$ has no closed-form (analytic) expression. Hence, Weiler (1965) resorted to solving graphically for the two parameters of the Beta distribution given the q-th and $(1-q)$-th quantile. This graphical approach, however, is limited to the number of graphs plotted. For intermediate solutions interpolation methods must be used, which are often subject to an interpolation error.

The adaptability of the Beta distribution will be reconfirmed in Section II by proving that a solution exists for the parameters of a Beta distribution for any combination of a lower quantile and upper quantile constraint. A numerical procedure will be described which solves for parameters a and b of a Beta distribution (cf. (1)) given these constraints. The contents of Section II is based on Van Dorp and Mazzuchi (2000). The numerical procedure derived in Section II can be easily adapted to the Weiler's (1965) methodology and improves on his graphical method. In addition, the numerical procedure can be adapted to the case where the median and an another quantile are specified as measures of central tendency and dispersion. In Sections III and IV the methods of Section II will be utilized to specify the parameters of the Dirichlet and Ordered Dirichlet distributions, respectively. In addition, some properties of the Dirichlet and Ordered Dirichlet distribution will be listed in Sections III and IV.

II. Specification of Prior Beta Parameters

For reasons to become evident from the discussion below, we will reparameterize the Beta density given by (1) by setting $\beta = a + b$ and $\alpha = \frac{a}{a+b}$; This yields the following expression for the probability density function of a Beta random variable X

$$\frac{\Gamma(\beta)}{\Gamma(\beta \cdot \alpha)\Gamma(\beta \cdot (1-\alpha))} \; x^{\beta \cdot \alpha - 1}(1-x)^{\beta \cdot (1-\alpha) - 1}, x \in [0,1], \tag{5}$$

where $0 \leq \alpha \leq 1,\ \beta > 0$. The reparameterization is a one-to-one transformation from (1) to (5) and vice versa. Note that the condition $\alpha \in [0,1]$ is identical to the condition on the original random variable X. For the purpose of this chapter, a random variable X distributed following (5) will be denoted as $X \sim Beta(\alpha,\beta)$. The latter notation is somewhat unconventional as $Beta(a,b)$ usually refers to the structural form of the pdf provided by (1). Perhaps the consistent use of Greek notation α and β rather than Latin notation a and b may help alleviate this source of confusion.

A. Basic Properties of the Beta Distribution

It easily follows from (5) that

$$E[X|\alpha,\beta] = \alpha \tag{6}$$

$$Var[X|\alpha,\beta] = \frac{\alpha\cdot(1-\alpha)}{(\beta+1)}. \tag{7}$$

Hence, the reparameterization provided in (6) allows one to interpret α as a location parameter and β as a shape parameter that determines the uncertainty in X. The n-th moment of X around zero in terms of α and β can be expressed utilizing (5) as

$$E[X^n|\alpha,\beta] = \frac{\prod_{i=1}^{n}(\beta\cdot\alpha+n-i)}{\prod_{i=1}^{n}(\beta+n-i)} = \alpha\cdot\frac{\prod_{i=1}^{n-1}(\beta\cdot\alpha+n-i)}{\prod_{i=1}^{n-1}(\beta+n-i)},\ n=1,2,3,\ldots \tag{8}$$

with the usual convention that $\prod_{i=1}^{0}\{\,\cdot\,\}=1$. Using the structure of (5), (6), (7) and (8) we can readily draw conclusions regarding the limiting distributions of a Beta random variable by letting $\beta\to\infty$ and $\beta\downarrow 0$ (for any fixed value of α). Consider the two different classes of degenerate distributions presented in Figure 2. It follows from (6) and (7) that the degenerate distribution in Class 1 of Figure 2 is the limiting distribution obtained by letting $\beta\to\infty$. From (8) it follows that the moments of the limiting distribution when letting $\beta\downarrow 0$ coincide with the moments of the degenerate distribution in Class 2 of Figure 2 (i.e. of a Bernoulli variable with a point mass of α at 1). As both the limiting distribution of X by letting $\beta\downarrow 0$ and the degenerate distribution of Class 2 have a bounded support, it follows from the agreement of their moments that the degenerate distribution in Class 2 <u>is</u> the limiting distribution by letting $\beta\downarrow 0$ (see e.g. Harris (1966), p. 103). The limiting distributions of Class 1 and Class 2 (and how they arise from the limiting behavior of the parameter β) play a central role in deriving the theoretical result in the next section.

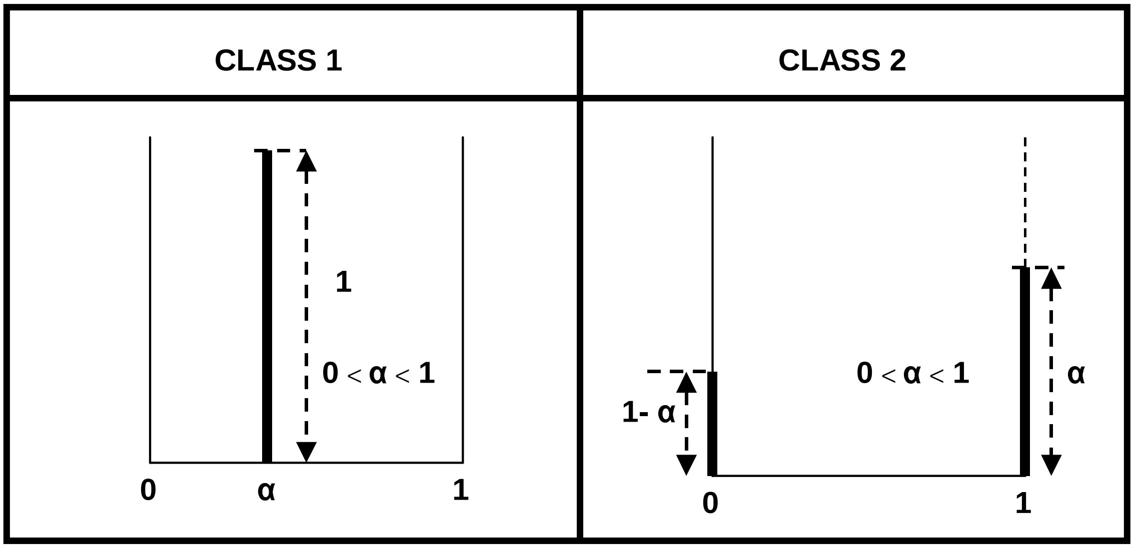

Figure 2. Two classes of degenerate Beta distributions.

An additional property of the Beta distribution utilized in this derivation is that for $b > 0$ (using the notation of (1))

$$0 < a_1 < a_2 \Rightarrow B(x \mid a_1, b) > B(x \mid a_2, b), \forall \, \mathrm{x} \in (0,1), \tag{9}$$

and for $a > 0$

$$b_1 > b_2 > 0 \Rightarrow B(x \mid a, b_1) > B(x \mid a, b_2), \forall \, \mathrm{x} \in (0,1), \tag{10}$$

(see e.g. Proschan and Singpurwalla (1979)). From (9) and (10) it follows (using the notation of (5)) that

$$\alpha_2 > \alpha_1 > 0, \, \beta > 0 \Rightarrow Pr(X \leq x|\alpha_2, \beta) < Pr(X \leq x|\alpha_1, \beta). \tag{11}$$

Finally, the *quantile constraint* concept defined in Definition 1 below will be used as well.

Definition 1: *Let* $0 < x_q < 1$, $0 < q < 1$. *A random variable* X *with support* $[0,1]$ *satisfies quantile constraint* (x_q, q) *if and only if* $Pr\{X \leq x_q\} = q$.

B. Solving for the Prior Beta Parameters

To specify the prior Beta parameters based on quantile estimates we need to solve problem $\mathcal{P}_1$ below. Solving problem $\mathcal{P}_1$ involves the use of the incomplete Beta function given by (4) and therefore has no closed form (analytic) solution. Also, the quantile constraints in problem $\mathcal{P}_1$ can be considered a set of two nonlinear constraints in two unknowns, i.e. α and β, and may not necessarily have a feasible solution. To construct a numerical procedure with solves problem $\mathcal{P}_1$ in a finite number of iterations, it is necessary to verify that problem $\mathcal{P}_1$ has a solution for any combination of the two quantile constraints. This assertion will be proved in Theorem 1 by means of limiting arguments.

***Problem $\mathcal{P}_1$* :** *Solve α and β for $X \sim Beta(\alpha, \beta)$ (cf. (5)) under the two quantile constraints (x_{q_L}, q_L) and (x_{q_U}, q_U), where $q_L < q_U$.*

Theorem 1: *There exists a solution (α^*, β^*) of problem $\mathcal{P}_1$.*

***Proof* :** The proof involves four steps. In the first step it will be proved, using the notation in (5), that for a given $\beta > 0$ and a quantile constraint (x_q, q), a unique α° exists such that $X \sim Beta(\alpha^\circ, \beta)$ (cf. (5)) satisfies this quantile constraint. In the second step it will be shown that for $\beta \downarrow 0$ the parameter $\alpha^\circ \to (1 - q)$. The third step validates that for $\beta \to \infty$ the parameter $\alpha^\circ \to x_q$. Finally, in the fourth step, the statement of this theorem will be verified.

Step 1: Let a quantile constraint (x_q, q) be specified for X. Assume that $\beta > 0$ is given and introduce the function $\xi(\alpha, \beta)$ such that

$$\xi(\alpha, \beta) = Pr\{X \leq x_q | \alpha, \beta\} - q, \; 0 < \alpha < 1, \beta > 0. \tag{12}$$

From the structure of (5) it follows that $\xi(\alpha, \beta)$ is a continuous differentiable function for $0 < \alpha < 1, \beta > 0.$ Consider $\xi(\alpha, \beta)$ when $\alpha \downarrow 0$ and $\beta > 0$ fixed. From (6) and (7) it follows, respectively, that

$$\underset{\alpha \downarrow 0}{Lim} E[X|\alpha, \beta] = 0,$$
$$\underset{\alpha \downarrow 0}{Lim} Var[X|\alpha, \beta] = 0,$$

respectively, for any fixed $\beta > 0$. Hence, when $\alpha \downarrow 0$, the distribution of X converges to a degenerate distribution with a single point mass concentrated at 0. With $0 < x_q < 1,\ 0 < q < 1$ it thus follows from (12) that

$$\underset{\alpha \downarrow 0}{Lim} \xi(\alpha, \beta) = 1 - q > 0, \tag{13}$$

for any fixed $\beta > 0$. Similarly, using (6), (7) and using the fact that the distribution of X converges to a degenerate distribution with a single point mass concentrated at 1 as $\alpha \uparrow 1$, we obtain

$$\underset{\alpha \uparrow 1}{Lim} \xi(\alpha, \beta) = -q < 0 \tag{14}$$

for any fixed $\beta > 0$. From (13), (14) with $\xi(\alpha, \beta)$ being a continuous function, it follows that

$$\exists\, \alpha^\circ \in (0,1): \xi(\alpha^\circ, \beta) = 0, \forall\, \beta > 0. \tag{15}$$

Utilizing expression (11), it follows that $\xi(\alpha, \beta)$ is a strictly decreasing function in α for any fixed $\beta > 0$. Thus, given fixed $\beta > 0$, α° is the unique solution to $\xi(\alpha, \beta) = 0$ and $X \sim Beta(\alpha^\circ, \beta)$ (cf. (5)) satisfies the quantile constraint (x_q, q) given fixed $\beta > 0$.

Before proceeding to Step 2, note that the solution α° depends on q, x_q (cf. (12)) and β (cf. (15)) motivating the following notation

$$\alpha^\circ = \mathcal{G}_{x_q}(\beta), \tag{16}$$

where $\mathcal{G}_{x_q}(\,\cdot\,)$: $(0, \infty) \rightarrow (0, 1)$, $0 < x_q < 1$, $0 < q < 1$, such that

$$\xi(\mathcal{G}_{x_q}(\beta), \beta) = 0, \forall\, \beta > 0. \tag{17}$$

From the structure of (5), (17), and the implicit function theorem, it follows that $\mathcal{G}_{x_q}(\beta)$ is also a continuous function for $\beta > 0$. Using the definition of $\xi(\alpha, \beta)$ given by (12) and (17), it follows that

$$Pr\{X \leq x_q | \mathcal{G}_{x_q}(\beta), \beta\} = q, \forall \beta > 0. \tag{18}$$

Step 2: Consider $X \sim Beta(\alpha^\circ, \beta)$ (cf. (5)), where $\alpha^\circ = \mathcal{G}_{x_q}(\beta)$, and let $\beta \downarrow 0$. From continuity of $Pr\{X \leq x_q | \mathcal{G}_{x_q}(\beta), \beta\}$ in β for fixed x_q it follows from (18) that

$$\lim_{\beta \downarrow 0} Pr\{X \leq x_q | \mathcal{G}_{x_q}(\beta), \beta\} = q. \tag{19}$$

For the structure of the density (6) it has been shown above that as $\beta \downarrow 0$, the distribution of X converges to a degenerate distribution of Class 2 in Figure 2. The limiting expectation of X as $\beta \downarrow 0$ thus becomes the expectation of a Bernoulli random variable and from (19) it follows that

$$\lim_{\beta \downarrow 0} E[\, X | \, \mathcal{G}_{x_q}(\beta), \beta \,] = 1 - q. \tag{20}$$

However, from (6) we have

$$E[X | \, \mathcal{G}_{x_q}(\beta), \beta] = \mathcal{G}_{x_q}(\beta), \tag{21}$$

for any $\beta > 0$ and using (20) and (21) one concludes

$$\lim_{\beta \downarrow 0} \mathcal{G}_{x_q}(\beta) = 1 - q. \tag{22}$$

In other words, the parameter $\alpha^\circ \to (1 - q)$ as $\beta \downarrow 0$.

Step 3: Consider $X \sim Beta(\alpha^\circ, \beta)$ (cf. (5)), where $\alpha^\circ = \mathcal{G}_{x_q}(\beta)$, and let $\beta \to \infty$. From (7) it follows that as $\beta \to \infty$ the distribution of X converges to a degenerate distribution of Class 1 in Figure 2 with a single point mass concentrated at some $x^\bullet \in [0, 1]$. From continuity of $Pr\{X \leq x_q | \mathcal{G}_{x_q}(\beta), \beta\}$ in β for fixed x_q it follows from (18) that $x^\bullet = x_q$. This means that,

$$\lim_{\beta \to \infty} E[\, X | \, \mathcal{G}_{x_q}(\beta), \beta \,] = x_q.$$

Hence, from (6) we have

$$\lim_{\beta \to \infty} \mathcal{G}_{x_q}(\beta) = x_q. \tag{23}$$

In other words, the parameter $\alpha^\circ \to x_q$ as $\beta \to \infty$.

Step 4: Let $X \sim Beta(\alpha, \beta)$ (cf. (6)). Let (x_{q_L}, q_L) and (x_{q_U}, q_U) be two quantile constraints specified for X, such that $q_L < q_U$. Consider the associated functions $\mathcal{G}_{x_{q_L}}(\beta)$ and $\mathcal{G}_{x_{q_U}}(\beta)$ each defined implicitly by (12) and (17), respectively. Introducing the function

$$H(\beta) = \mathcal{G}_{x_{q_L}}(\beta) - \mathcal{G}_{x_{q_U}}(\beta)$$

it follows from (22) that

$$\underset{\beta \downarrow 0}{Lim} H(\beta) = (1 - q_L) - (1 - q_U) = q_U - q_L > 0. \tag{24}$$

Similarly, from (23) it follows that

$$\underset{\beta \to \infty}{Lim} H(\beta) = x_{q_L} - x_{q_U} < 0. \tag{25}$$

From the continuity of $\mathcal{G}_{x_{q_L}}(\beta)$ and $\mathcal{G}_{x_{q_U}}(\beta)$, (24) and (25) it follows that

$$\exists \beta^* > 0 : H(\beta^*) = 0. \tag{26}$$

Denoting $\alpha^* = \mathcal{G}_{x_{q_L}}(\beta^*)$ (cf. (16)), it follows from (26) that

$$\alpha^* = \mathcal{G}_{x_{q_L}}(\beta^*) = \mathcal{G}_{x_{q_U}}(\beta^*).$$

In other words, $X \sim Beta(\alpha^*, \beta^*)$ (cf. (5)) satisfies both quantile constraints (x_{q_L}, q_L) and (x_{q_U}, q_U) and thus (α^*,β^*) is a solution to problem $\mathcal{P}_1$. □

Theorem 1 proves the existence of a solution to problem $\mathcal{P}_1$. The uniqueness of the solution (α^*, β^*) to $\mathcal{P}_1$ would follow by showing that; (i) $H(\beta)$ has 0 or 1 stationary points for $\beta > 0$; (ii) if $H(\beta)$ has a stationary point for $\beta > 0$ this stationary point coincides with a global maximum. It is conjectured that the above assertions hold. Numerical analyses in the examples below support this conjecture (see Van Dorp and Mazzuchi (2000)). In case multiple solutions exist to problem $\mathcal{P}_1$, the numerical algorithm below is designed so that the selected solution coincides with the solution with the lowest value for β^*, and thus the highest level of uncertainty. The latter solution would be a preferred solution, given that x_{q_L} and x_{q_U} ought to be elicited through expert judgment in Bayesian Analysis.

C. Design of a Numerical Procedure

Since problem $\mathcal{P}_1$ cannot be solved in a closed form, a numerical procedure that determines a solution to problem $\mathcal{P}_1$ with a prescribed level of accuracy, in a finite number of iterations, is desirable. Below, such a numerical procedure will be informally described. The numerical method uses a procedure for solving for the q-th quantile of a Beta distribution. Such a procedure is described in the Appendix in Pseudo Pascal (denoted $BISECT\,1$).

From (6) and (7) it follows that α is a *location parameter* and β is an *uncertainty parameter* given the value of α and higher values of β coincide with lower uncertainty levels. These interpretations of the parameters α and β are used in the design of the numerical procedure to obtain a solution to $\mathcal{P}_1$. Assume for now that an interval $[a_1, b_1]$ is obtained containing β^* which yields a solution (α^*, β^*) of $\mathcal{P}$, where $\alpha^* = \mathcal{G}_{x_q}(\beta^*)$. Let β_1 be the midpoint of this interval. The k-th iteration of the numerical procedure will be described below.

To solve $(\alpha^\circ)_k$ satisfying the quantile constraint (x_{q_U}, q_U) of $\mathcal{P}_1$ given a value for β_k, successive shrinking intervals $[d_n, e_n]$ are calculated containing the solution $(\alpha^\circ)_k$. From (5) follows that $(\alpha^\circ)_k \in [0,1]$. Hence, $[d_1, e_1] = [0,1]$. Next, α_n is set to the midpoint of $[d_n, e_n]$ and the probability mass $(q_U)_n = Pr\{X \leq x_{q_U} | \alpha_n, \beta_k\}$ is calculated. In case $(q_U)_n \leq q_U$, the Beta distribution is skewed excessively towards 1. Therefore, it follows from (6) that the value of the location parameter α_n is too high. Hence, the next interval containing $(\alpha^\circ)_k$ is $[d_{n+1}, e_{n+1}] = [d_n, \alpha_n]$. On the other hand, when $(q_U)_n > q_U$, the Beta distribution is skewed excessively towards 0. Therefore it follows from (6) that the value of the location parameter α_n is too small. Hence, the next interval containing $(\alpha^\circ)_k$ can be set to $[d_{n+1}, e_{n+1}] = [\alpha_n, e_n]$. Finally, the next estimate α_{n+1} is set to be the midpoint of the interval $[d_{n+1}, e_{n+1}]$. The above procedure is repeated until $(q_U)_n$ is close to q_U with a pre-assigned level of accuracy. The quantile constraint (x_{q_U}, q_U) of $\mathcal{P}_1$ is satisfied once this accuracy has been reached and $(\alpha^\circ)_k$ is set equal to the α_n. The algorithm above is a bisection method (See, for example, Press et al., 1989) and is provided in the Appendix in Pseudo Pascal (denoted $BISECT\,2$). A specific example of the algorithm in $BISECT\,2$ is presented in Figure 3, where $(x_{q_U}, q_U) = (0.80, 0.70)$

and $\beta_1 = 3$. The starting interval for α_1 equals $[d_1, e_1] = [0, 1]$, hence $\alpha_1 = 0.5$. Thus it follows that $(q_U)_1 > 0.7$, hence $[d_2, e_2] = [0.5, 1]$ and $\alpha_2 = 0.75$. Now we have $(q_U)_2 < 0.7$, hence $[d_3, e_3] = [0.5, 0.75]$ and $\alpha_3 = 0.625$. Consequently we have $(q_U)_3 \approx 0.7 = q_U$, $(\alpha^\circ)_1$ is set to $\alpha_3 = 0.625$, the algorithm terminates and $X \sim Beta((\alpha^\circ)_1, \beta_1)$ (cf. (5)) satisfies the quantile constraint $(x_{q_U}, q_u) = (0.80, 0.70)$.

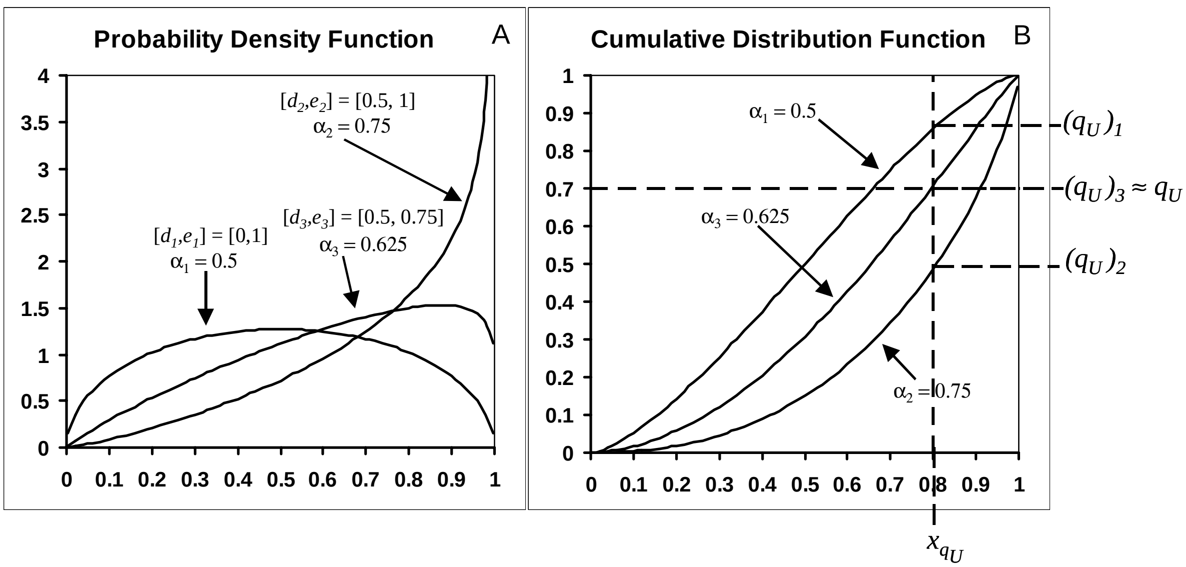

Figure 3. An example of bisection method $BISECT\,2$. A: Beta PDF's and shrinking bisection intervals $[d_n, e_n]$, B: Beta CDF's and sequence of $(q_U)_n$, $n = 1, \ldots, 3$

After solving for $(\alpha^\circ)_k$ (utilizing $BISECT2$) the procedure calculates the q_L-th quantile $(x_{q_L})_k$ (utilizing $BISECT1$) of $Beta((\alpha^\circ)_k, \beta_k)$. When $(x_{q_L})_k < x_{q_L}$ the uncertainty in $Beta((\alpha^\circ)_k,\ \beta_k)$ (cf. (5)) is too high. Therefore, the current estimate of the uncertainty parameter β_k should be too low. Hence, the next interval which contains β^* can be set to $[a_{k+1}, b_{k+1}] = [\beta_k, b_k]$. On the other hand, if $(x_{q_L})_k > x_{q_L}$ the uncertainty in $Beta((\alpha^\circ)_k,\ \beta_k)$ is too low. Therefore, the current estimate of the uncertainty parameter β_k is too high. Hence, the next interval which contains β^* can be set to $[a_{k+1}, b_{k+1}] = [a_k, \beta_k]$. Finally, the next estimate β_{k+1} is taken to be the midpoint of the interval $[a_{k+1}, b_{k+1}]$. The above procedure is then repeated until the current estimate $(x_{q_L})_k$ is close to x_{q_L} with the pre-assigned desired level of accuracy.

The quantile constraint (x_{q_L}, q_L) of $\mathcal{P}_1$ is met once this accuracy has been reached. The parameters (α^*, β^*) that solve $\mathcal{P}_1$ are set equal to the pair $((\alpha^\circ)_k,\ \beta_k)$. The algorithm above is a bisection method and is provided in the Appendix in Pseudo Pascal (denoted $BISECT\,3$). A specific example of the algorithm $BISECT\,3$ is presented in Figure 4, where $(x_{q_L}, q_L) = (0.20, 0.10)$, $(x_{q_U}, q_U) = (0.80, 0.70)$, $\beta_1 = 3$ and $(\alpha^\circ)_1 = 0.625$.

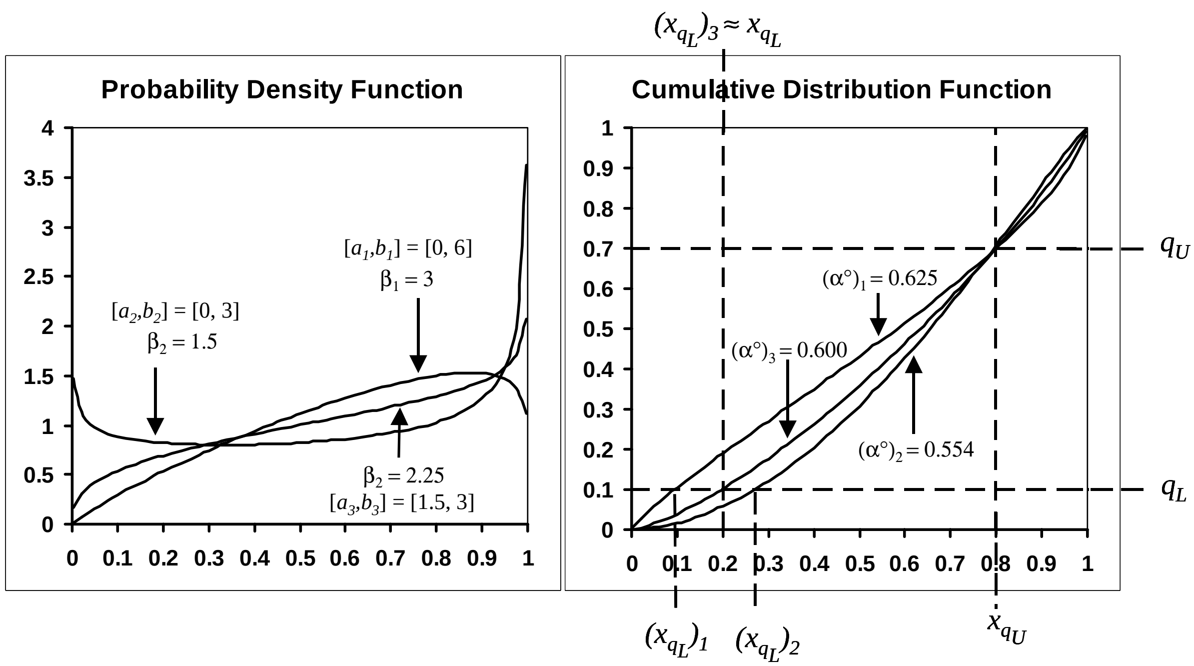

Figure 4. An example of bisection method $BISECT\,3$. A: Beta PDF's and shrinking bisection intervals $[a_k, b_k]$, B: Beta CDF's and a sequence of $(x_{q_L})_k,\ k = 1, \ldots, 3$

The starting interval for $\beta_1$equals $[a_1, b_1] = [0, 6]$. It follows that $(x_{q_L})_1 < 0.2$, hence $[a_2, b_2] = [0, 3]$ and $\beta_2 = 1.5$, $(\alpha^\circ)_2 = 0.554$ (determined using $BISECT2$). It follows that $(x_{q_L})_1 > 0.2$, hence $[a_3, b_3] = [1.5, 3]$ and $\beta_3 = 2.25$, $(\alpha^\circ)_2 = 0.600$ (determined using $BISECT2$). It now follows that $(x_{q_L})_1 \approx 0.2 = q_L$, the algorithm terminates, β^* is set $\beta_3 = 2.25$, α^* is set to $(\alpha^\circ)_3 = 0.600$ and we have $X \sim Beta(\alpha^*, \beta^*)$ which satisfies the quantile constraints $(x_{q_L}, q_L) = \ (0.20, 0.10)$, $(x_{q_U}, q_U) = \ (0.80, 0.70)$.

To determine a starting interval $[a_1, b_1]$ containing β^* the following steps could be adopted in the procedure. Set the lower bound $a_1 = 0$. To obtain the upper bound b_1, set $\beta_{1,k} = 1$, where $k = 1$, and solve for $(\alpha^\circ)_{1,k}$ satisfying the quantile constraint x_{q_U} of problem $\mathcal{P}_1$ utilizing $BISECT2$. Next, solve for the q_L-th quantile $(x_{q_L})_{1,k}$ of $Beta((\alpha^\circ)_{1,k}, \beta_{1,k})$ (utilizing $BISECT1$). In case $(x_{q_L})_{1,k} < x_{q_L}$ the uncertainty in $Beta((\alpha^\circ)_{1,k}, \beta_{1,k})$ is too high. Therefore, $\beta_{1,k} < \beta^*$. In that case, set $\beta_{1,k+1} = 2\beta_{1,k}$ and repeat the above procedure. Conversely, in the case $(x_{q_L})_{1,k} > x_{q_L}$ the uncertainty in $Beta((\alpha^\circ)_{1,k}, \beta_{1,k})$ is too low. Therefore, $\beta_{1,k} > \beta^*$. In that case, set $b_1 = \beta_{1,k}$ and the starting interval $[a_1, b_1]$ has been determined. Note that if multiple solutions exist to problem $\mathcal{P}_1$, the starting interval is chosen in such a manner that the selected solution for $\mathcal{P}_1$ by means of the algorithm coincides with the solution with the lowest value for β^*, and consequently the highest level of uncertainty.

The three different bisection methods $BISECT1$, $BISECT2$ and $BISECT3$ were implemented in a PC-based program BETA-CALCULATOR. Figure 5 displays a screen capture of BETA-CALCULATOR.

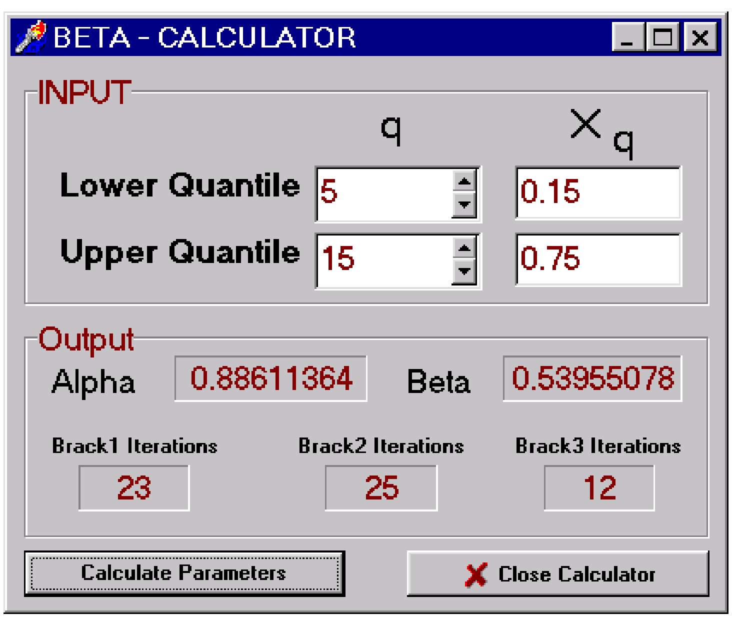

Figure 5. Screen Capture of BETA-CALCULATOR with calculation results for the first row in Table 1.

The accuracy for δ in the bisection methods $BISECT1$ and $BISECT2$ was set to be 10^{-8}. The accuracy in the bisection method $BISECT3$ was set to be 10^{-4}. Table 1 contains solutions to problem $\mathcal{P}$ for 4 different combinations of a lower quantile and upper quantile constraint calculated using BETA-CALCULATOR. In addition, Table 1 provides the maximum number of iterations in each bisection method to yield the solutions with the above settings of error-tolerances. Figure 5 contains the results for Example 1 in Table 1. Example 4 in Table 1 coincides with the setup of the Weiler's (1965) graphical method. Finally, Figure 6 depicts the probability density functions and cumulative distribution functions associated with the examples in Table 1. Note that the U-Shaped, J-Shaped and Unimodal forms of the Beta distribution are represented in Figure 6.

Table 1. Some Calculation Examples of Beta Parameters given an upper and a lower quantile constraints

		q	x_q	α^*	β^*	# 1	# 2	# 3
$Example\,1$	L	0.05	0.15	0.8861	0.5396	23	25	12
	U	0.15	0.75					
$Example\,2$	L	0.49	0.25	0.2672	11.3906	71	24	11
	U	0.99	0.60					
$Example\,3$	L	0.20	0.10	0.3437	2.6328	26	26	10
	U	0.50	0.30					
$Example\,4$	L	0.05	0.45	0.6330	19.5625	44	25	10
	U	0.95	0.80					

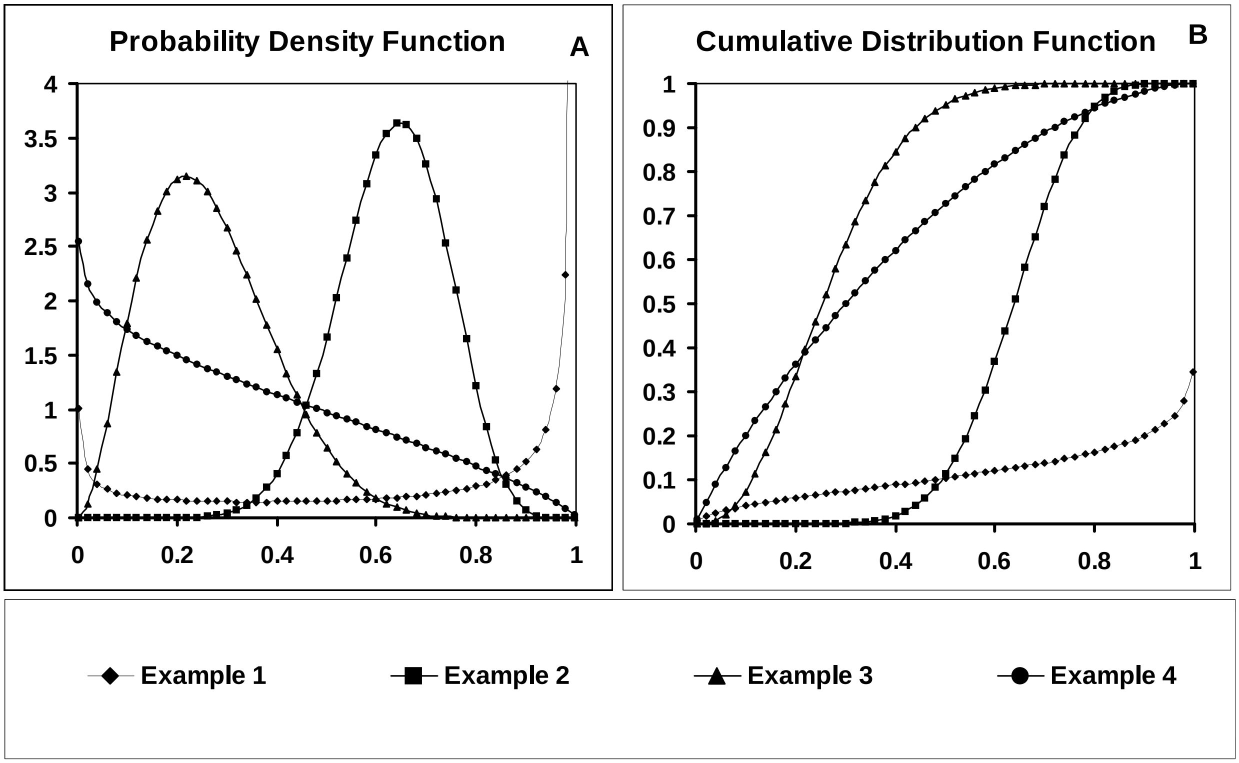

Figure 6. A: Beta Probability Density Functions for the Examples in Table 1,

B: Beta Cumulative Distribution Functions for the Examples in Table 1.

III. Specification of Prior Dirichlet Parameters

Analogously to (5) and for a straightforward application of the numerical procedures derived in Section 2 (and provided in the Appendix), we reparameterize the Dirichlet distribution given by (2) by introducing the new parameters $\beta = \sum_{i=1}^{m+1} \theta_i$ and $\alpha_i = \frac{\theta_i}{\beta},\, i = 1, \ldots, m,$ yielding the probability density

$$\frac{\Gamma(\beta)}{\left\{\prod_{i=1}^{m}\Gamma(\beta\cdot\alpha_i)\right\}\Gamma\Big(\beta\cdot(1-\sum_{i=1}^{m}\alpha_i)\Big)}\left\{\prod_{i=1}^{m} x_i^{\beta\cdot\alpha_i-1}\right\}\left\{1-\sum_{i=1}^{m}x_i\right\}^{\beta\cdot(1-\sum_{i=1}^{m}\alpha_i)-1}, \qquad (27)$$

where $x_i \geq 0,\ \sum_{i=1}^{m} x_i \leq 1,\ \alpha_i \geq 0, i = 1, \ldots, m,\ \sum_{i=1}^{m}\alpha_i \leq 1$ and $\beta > 0$. A random vector $\underline{X} = (X_1, \ldots, X_m)$ distributed according to the reparameterized Dirichlet distribution (27) will be denoted by $Dirichlet(\underline{\alpha}, \beta)$. Note that, as in the case of (5), the condition on the parameters $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$ is identical to the conditions on the variables $X_i, i = 1, \ldots, m$.

A. Basic Properties of the Dirichlet Distribution

Let a random vector $\underline{X} = (X_1, \ldots, X_m) \sim Dirichlet(\underline{\alpha}, \beta)$. It may be derived from (27) that marginals distribution of X_i are given by $X_i \sim Beta(\alpha_i, \beta)$ (in parameterization of (5)), $i = 1, \ldots, m$. The moments $E[X_i^n | \underline{\alpha}, \beta]$ follow by substituting α_i for α in (8). Analogously, the mean and the variance of X_i follow by substituting α_i for α in (6) and (7), respectively. Hence, the parameter β of the $Dirichlet(\underline{\alpha}, \beta)$ distribution may be interpreted as the common shape parameter amongst $\underline{X} = (X_1, \ldots, X_m)$, whereas the vector $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$ may be interpreted as a location parameter of $\underline{X}$. Such an interpretation was not valid for the original parameterization given by (2) involving the parameters $\theta_i,\ i = 1, \ldots, m + 1$. Similar to the analysis in Section 2.1 it follows that we may draw conclusions regarding the limiting distributions of the $Dirichlet(\underline{\alpha}, \beta)$ based solely on the limiting behavior of the parameter β. Letting $\beta \rightarrow \infty$ we observe that a $Dirichlet(\underline{\alpha}, \beta)$ distribution converges to a degenerate distribution with a single point mass concentrated at $\underline{\alpha}$. Letting $\beta \downarrow 0$, we deduce that the $Dirichlet(\underline{\alpha}, \beta)$ distribution converges to an m-variate Bernoulli distribution with marginal parameters α_i in Class 2 of Figure 2, $i = 1, \ldots, m$. The dependence structure in the limiting m-variate Bernoulli distribution is obtained by studying the limiting behavior of the pairwise correlation coefficients in a $Dirichlet(\underline{\alpha}, \beta)$ distribution as $\beta \downarrow 0$. Utilizing the reparameterization in (27) it follows that

$$Cov(X_i, X_j) = -\frac{\alpha_i \alpha_j}{\beta + 1} \tag{28}$$

and with (28) and (7)

$$Cor(X_i, X_j) = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i) Var(X_j)}} = -\sqrt{\frac{\alpha_i \alpha_j}{(1 - \alpha_i)(1 - \alpha_i)}}. \tag{29}$$

Apparently, the correlation structure in a $Dirichlet(\underline{\alpha}, \beta)$ distribution does not depend on the common scale parameter β, and (29) describes the dependence structure of the limiting m-variate Bernoulli distribution when $\beta \downarrow 0$. Relation (29) is consistent with the well-known result (see Kotz et al. 2000) that the correlations in a "classical" Dirichlet distribution are negative.

We now present several basic properties of the $Dirichlet(\underline{\alpha}, \beta)$ distribution below utilizing reparameterization (27). It would appear (similar to the result in (29)) that some further transparency may be achieved by expressing these properties in terms of $(\underline{\alpha}, \beta)$. Firstly, for any index set $A \subset \{1, \ldots, m\}$,

$$\underline{X}^A \sim Dirichlet(\underline{\alpha}^A, \beta), \qquad (30)$$

where $\underline{X}^A = \{X_i | i \in A\}$ and $\underline{\alpha}^A = \{\alpha_i | i \in A\}$. Next,

$$\sum_{i \in A} X_i \sim Beta(\sum_{i \in A} \alpha_i, \beta)$$

and

$$\frac{X_j}{\sum_{i \in A} X_i} \sim Beta(\frac{\alpha_j}{\sum_{i \in A} \alpha_i}, \beta \sum_{i \in A} \alpha_i).$$

Finally, utilizing (30) we may derive the conditional probability density function of $(\underline{X}^A | \underline{X}^{A^c})$, where A^c denotes the complement of A, i.e. $A^c = \{1, \ldots, m\} \setminus A$, yielding

$$\left\{\frac{1}{\xi}\right\}^{\beta^\bullet - 1} \frac{\Gamma(\beta^\bullet)}{\left\{\prod_{i \in A} \Gamma(\beta^\bullet \cdot \alpha_i^\bullet)\right\} \Gamma\Big(\beta \cdot (1 - \sum_{i \in A} \alpha_i^\bullet)\Big)} \times \qquad (31)$$
$$\left\{\prod_{i \in A} y_i^{\beta^\bullet \cdot \alpha_i^\bullet - 1}\right\} \left\{\xi - \sum_{i \in A} y_i\right\}^{\beta^\bullet \cdot (1 - \sum_{i \in A} \alpha_i^\bullet) - 1}$$

where

$$\xi = 1 - \sum_{j \in A^c} x_j, \quad \beta^\bullet = \beta \cdot (1 - \sum_{i \in A^c} \alpha_i), \quad \alpha_i^\bullet = \frac{\alpha_i}{1 - \sum_{i \in A^c} \alpha_i}. \tag{32}$$

The distribution in (31) may be recognized as that of an $|A|$- dimensional vector $\underline{Y}$, where $|A|$ indicates the cardinality of the index set A, and

$$\underline{Y} = \xi\, \underline{Z}, \underline{Z} \sim Dirichlet(\underline{\alpha}^\bullet, \beta^\bullet),$$

where $\underline{\alpha}^\bullet = \{\alpha_i^\bullet | i \in A\}$ and $\alpha_i^\bullet$ and $\beta^\bullet$ are given by (32). Setting $A = \{i\}$ in (31) and (32) yields what is called the *full conditional distribution* of X_i as a transformed $Beta(\alpha_i^\bullet, \beta^\bullet)$ with the support

$$[0, 1 - \sum_{j=1, j \neq i} x_j].$$

The latter result is relevant to the application of the Markov Chain Monte Carlo (MCMC) methods utilizing a $Dirichlet(\underline{\alpha}, \beta)$ (see, e.g., Casella & George (1992)). The MCMC methods have spurted an emergence of numerous Bayesian applications (see, e.g. Gilks et al. (1995)) as these methods allow for sampling from a posterior distribution by successively sampling from posterior full conditional distributions, without having a closed form of the posterior distribution.

B. Solving for the Dirichlet Prior Parameters

In order to solve for the common shape parameter β and location parameter $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$ of an m dimensional random vector $\underline{X} \sim Dirichlet(\underline{\alpha}, \beta)$ distribution using quantile estimates, it is required to solve problem $\mathcal{P}_2$ below.

Problem $\mathcal{P}_2$: *Solve $\underline{\alpha}$ and β for $\underline{X} \sim Dirichlet(\underline{\alpha}, \beta)$, $\underline{X} = (X_1, \ldots, X_m)$ under the two quantile constraints $(x^i_{q_L}, q^i_L)$ and $(x^i_{q_U}, q^i_U)$ for X_i , where $q^i_L < q^i_U,$ and $m-1$ single quantile constraints (x^j_q, q^j) for $X_j,\ j = 1, \ldots, m, j \neq i$.*

Note that the quantile levels q^j and q_U^i (or q_L^i) may differ amongst X_j, $j = 1, \ldots, m.$ Since $X_i \sim Beta(\alpha_i, \beta)$, it follows immediately from Theorem 1 that a solution to problem $\mathcal{P}_2$ exists. When multiple solutions are available to problem $\mathcal{P}_2$, the numerical algorithms in the Appendix are designed such that the selected solution coincides with the solution with the lowest value for β, and thus the highest level of uncertainty. The latter solution would be a preferred solution given that the quantile constraints in $\mathcal{P}_2$ ought to be elicited through expert judgment in Bayesian Analysis. The algorithm to solve $\mathcal{P}_2$ is provided below in Pseudo Pascal using the bisection methods described in the Appendix.

$STEP\,1:\quad BISECT3(\boldsymbol{\alpha_i}, \boldsymbol{\beta}, x^i_{q_L}, x^i_{q_U}, q^i_L, q^i_U): j = 1;$

$STEP\,2:\quad If\; j \neq i\; then\, BISECT2(\boldsymbol{\alpha_j}, x^j_q, \beta, q^j);$

$STEP\,3:\quad If\; j < m\, then\, j: \; = j + 1;\; Goto\, STEP\,2; \qquad Else\, Stop;$

Table 2 below describes two instances of problem $\mathcal{P}_2$ and their solutions using the algorithm above for $\underline{X} = (X_1, X_2)$, where $\underline{X} \sim Dirichlet(\underline{\alpha}\,, \beta)$ and $\underline{\alpha} = (\alpha_1, \alpha_2)$. Note that, $(x^1_{q_L}, q^1_L)$ and $(x^1_{q_U}, q^1_U)$ in Table 2 coincide with the third row of Table 1. Hence, α_1 and β also coincide in Tables 1 and 2 resulting in a J-shaped marginal form for the pdf of X_1 given in Figure 6A. The resulting marginal form in Example 1(2) of Table 2 for X_2 is J-shaped (uni-modal). Figure 7 displays the resulting Dirichlet densities for the examples in Table 2. Note that, the marginal density of X_1 in both Figures 7A and 7B is identical and J-Shaped, whereas the marginal form of X_2 in Figure 7B is uni-modal. As a result, the joint pdf in Figure 7B has a single mode at $(x_1, x_2) = (0, 1)$. Figure 7A displays three modes at each corner of the unit simplex.

Table 2. Calculation Examples

		X_1	X_2			
		(q, x_q)	(q, x_q)	α_1	α_2	β
$Example\,1$	L	$(0.20, 0.10)$	$(0.20, 0.10)$	0.3437	0.3437	2.6328
	U	$(0.50, 0.30)$				
$Example\,2$	L	$(0.20, 0.10)$	$(0.40, 0.50)$	0.3437	0.5665	2.6328
	U	$(0.50, 0.30)$				

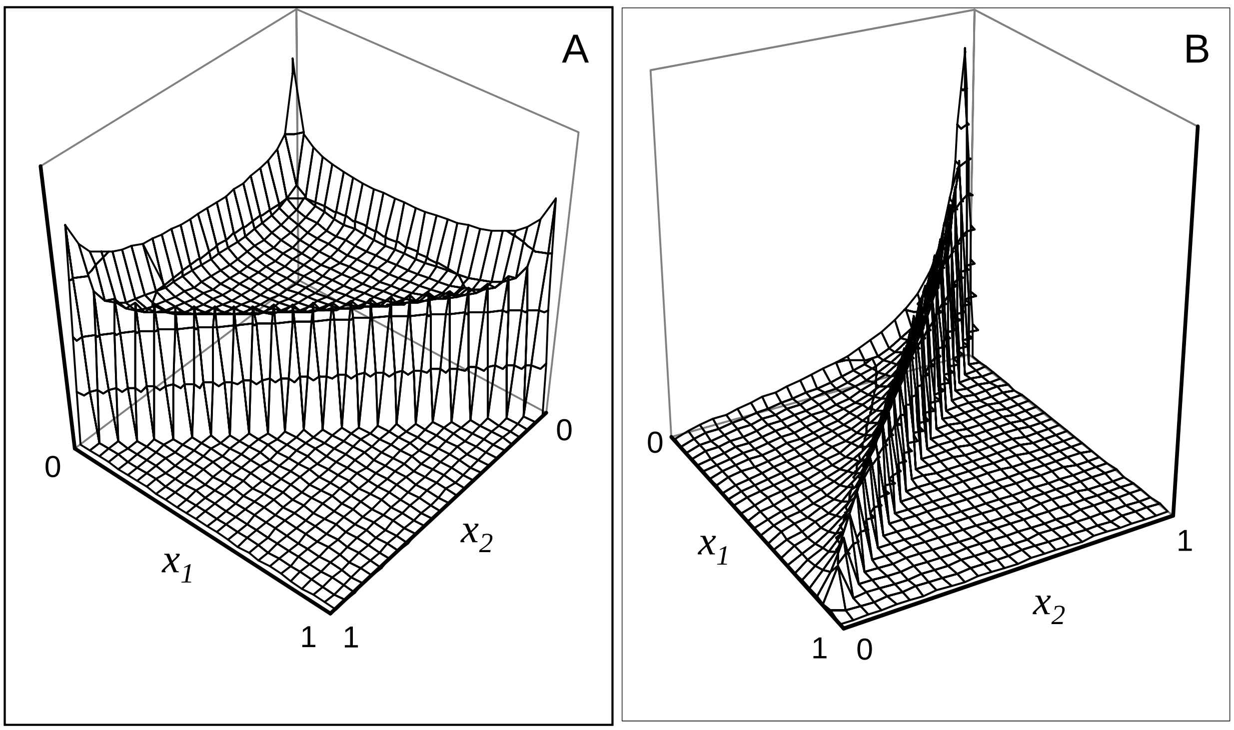

Figure 7. A: Dirichlet PDF of Example 1 in Table 2,

B: Dirichlet PDF of Example 2 in Table 2.

IV. Specification of Ordered Dirichlet Parameters

Analogously to (5) and (27) we reparameterize the Ordered Dirichlet distribution given by (3) by introducing $\beta = \sum_{i=1}^{m+1} \theta_i$ and $\alpha_i = \frac{\theta_i}{\beta}$, $i = 1, \ldots, m,$ yielding the probability density

To appear in Handbook of Beta Distributions and Its Applications Publisher, Marcel Dekker, Inc.
Editors Gupta, A.K. and Nadarajah, S.

$$\frac{\Gamma(\beta)}{\left\{\prod_{i=1}^{m}\Gamma(\beta\cdot\alpha_i)\right\}\Gamma\Big(\beta\cdot(1-\sum_{i=1}^{m}\alpha_i)\Big)}\times \tag{33}$$
$$x_1^{\beta\cdot\alpha_i-1}\left\{\prod_{i=2}^{m}(x_i-x_{i-1})^{\beta\cdot\alpha_i-1}\right\}(1-x_m)^{\beta\cdot(1-\sum_{i=1}^{m}\alpha_i)-1}$$

where $0\leq x_{i-1}<x_i\leq 1,\ \alpha_i\geq 0, i=1,\ldots,m,\ \sum_{i=1}^{m}\alpha_i\leq 1$ and $\beta>0$. The distribution given by (33) will be denoted by $OD(\underline{\alpha},\beta)$, where $\underline{\alpha}=(\alpha_1,\ldots,\alpha_m)$.

A. Some Properties of the Ordered Dirichlet Distribution

Let a random vector $\underline{X}=(X_1,\ldots,X_m)\sim OD(\underline{\alpha},\beta)$. It easily follows from (33) that marginals distribution of X_i are given by

$$X_i\sim Beta(\alpha_i^+,\beta),\ \ \alpha_i^+=\sum_{k=1}^{i}\alpha_k \tag{34}$$

(in parameterization (5)), $i=1,\ldots,m$. The moments $E[X_i^n|\underline{\alpha},\beta]$ follow by substituting α_i^+ for α in (8). Analogously, the mean and the variance of X_i follow by substituting α_i^+ for α in (6) and (7), respectively. As above, the parameter β of the $OD(\underline{\alpha},\beta)$ may be interpreted as the common scale parameter amongst $\underline{X}=(X_1,\ldots,X_m)$, whereas the vector $\underline{\alpha}^+=(\alpha_1^+,\ldots,\alpha_m^+)$ may be interpreted as a location parameter of $\underline{X}$. Similarly to the analysis in Section 3.1 it follows that that the degenerate distribution with a point mass concentrated at parameter $\underline{\alpha}^+$ (cf. (34)) is the degenerate distribution of an $OD(\underline{\alpha},\beta)$ distribution by letting $\beta\rightarrow\infty$. Letting $\beta\downarrow 0$ we deduce that the $OD(\underline{\alpha},\beta)$ distribution converges to an ordered m-variate Bernoulli distribution with marginal parameters α_i^+ (cf. (34)), $i=1,\ldots,m$. The dependence structure in the limiting ordered m-variate Bernoulli distribution is obtained by studying the limiting behavior of the pairwise correlation coefficients in a $OD(\underline{\alpha},\beta)$ distribution as $\beta\downarrow 0$. Utilizing the reparameterization in (33) it follows that

$$Cov(X_i,X_j) = \frac{Cov(X_i, X_j)}{\sqrt{Var(X_i)Var(X_j)}} = \frac{\alpha_i^+(1-\alpha_j^+)}{\beta+1} \tag{35}$$

and with (35) and (7) we have

$$Cor(X_i,X_j) = \sqrt{\frac{\alpha_i^+(1-\alpha_j^+)}{(1-\alpha_i^+)\,\alpha_j^+}}. \tag{36}$$

Utilizing the pdf reparameterization in (33), it follows from (36) that the correlation structure in a $OD(\underline{\alpha}, \beta)$ does not depend on the common scale parameter β as it is in the case of the $Dirichlet(\underline{\alpha}, \beta)$ distribution. Note that, unlike the case of $Dirichlet(\underline{\alpha}\,, \beta)$ distribution, the correlations are positive. The difference between the signs of the correlations in the $OD(\underline{\alpha}, \beta)$ and $Dirichlet(\underline{\alpha}, \beta)$ distributions may in part be explained by the differences in their support (see, Figure 1). An additional useful property for the $OD(\underline{\alpha}, \beta)$ is that for any index set $A \subset \{1, \ldots, m\}$,

$$\underline{X}^A \sim OD(\underline{\alpha}^A, \beta), \tag{37}$$

where $\underline{X}^A = \{X_i | i \in A\}$, $\underline{\alpha}^A = \{\alpha_l^A | l = 1, \ldots, A\}$ and

$$\alpha_l^A = \begin{cases} \alpha_{A_{(l)}}^+ & l = 1 \\ \alpha_{A_{(l)}}^+ - \alpha_{A_{(l-1)}}^+ & l = 2, \ldots, |A| \end{cases}$$

where, as above, $|A|$ indicates the number of elements in the index set A and $A_{(l)}$ indicates the l-th element in A, such that $A_{(l)} > A_{(l-1)}, k = 2, \ldots, |A|$. Furthermore,

$$(X_j - X_i) \sim Beta(\alpha_j^+ - \alpha_i^+, \beta)$$

and

$$\frac{X_i}{X_j} \sim Beta(\frac{\alpha_i^+}{\alpha_j^+}, \beta\alpha_j^+).$$

where $j > i$ and α_i^+ are defined in (34). Finally, utilizing (37) we may derive the conditional probability density function of $(\underline{X}^A | \underline{X}^{A^c})$, where A^c denotes the complement of A, i.e. $A^c = \{1, \ldots, m\} \setminus A$ and $A = \{i\}, i \in \{1, \ldots, m\}$ yielding

$$\frac{\Gamma\Big(\beta(\alpha_i + \alpha_{+1})\Big)}{\Gamma(\beta\alpha_i)\Gamma(\beta\alpha_{+1}))} \frac{(x_i - x_{i-1})^{\beta\alpha_i - 1}(x_{i+1} - x_i)^{\beta\alpha_{i+1} - 1}}{(x_{i+1} - x_{i-1})^{\beta(\alpha_{i+1} + \alpha_i) - 1}} \tag{38}$$

The distribution in (38) may be recognized as that of a random variable Y where

$$Y = (x_{i+1} - x_{i-1})Z + x_{i-1},\, Z \sim Beta(\frac{\alpha_i}{\alpha^+_{i+1} - \alpha^+_{i-1}}, \beta(\alpha^+_{i+1} - \alpha^+_{i-1})). \tag{39}$$

where α^+_{i+1} is defined by (34). Hence, the distribution of $(\underline{X}^A | \underline{X}^{A^c})$ where $A = \{i\}$ (referred to as the *full conditional distribution* of X_i) is a transformed Beta distribution with support

$$\begin{cases} [0, x_{i+1}] & i = 1 \\ [x_{i-1}, x_{i+1}] & i = 1, \ldots, m-1 \\ [x_{i-1}, 1] & i = m \end{cases}$$

As above, the result in (38) and (39) is relevant to the application of Markov Chain Monte Carlo (MCMC) methods (see, e.g. Casella & George (1992)) utilizing an $OD(\underline{\alpha}, \beta)$ distribution.

B. Solving for the Ordered Dirichlet prior parameters

To solve for the common shape parameter β and location parameter $\underline{\alpha} = (\alpha_1, \ldots, \alpha_m)$ of an m dimensional random vector $\underline{X} \sim OD(\underline{\alpha}, \beta)$ distribution using quantile estimates, we are required to solve problem $\mathcal{P}_3$ below.

Problem $\mathcal{P}_3$: *Solve $\underline{\alpha}$ and β for $\underline{X} \sim OD(\underline{\alpha}, \beta)$, $\underline{X} = (X_1, \ldots, X_m)$ under the two quantile constraints $(x^i_{q_L}, q^i_L)$ and $(x^i_{q_U}, q^i_U)$ for X_i , where $q^i_L < q^i_U$, and $m-1$ single quantile constraints (x^j_q, q^j) for X_j, $j = 1, \ldots, m, j \neq i$ such that*

$$x^1_q < x^2_q \ldots < x^{i-1}_q < x^i_{q_L} < x^i_{q_U} < x^{i+1}_q < \ldots < x^m_q \tag{40}$$

Note that the quantile levels q^j and q^i_U (or q^i_L) may differ amongst X_j, $j = 1, \ldots, m.$ Since $X_i \sim Beta(\alpha^+_i, \beta)$, it follows immediately from Theorem 1 that a solution to problem $\mathcal{P}_3$ exists. In case multiple solutions exist to problem $\mathcal{P}_3$, the numerical algorithms in the Appendix are designed in such a manner that the selected solution coincides with the solution with the lowest value for β, and thus the highest level of obtained uncertainty. As mentioned above, the latter solution would be a preferred solution. The algorithm for solving $\mathcal{P}_3$ is provided below in Pseudo Pascal using the bisection methods described in the appendix.

$STEP\,1:$ $\quad BISECT3(\boldsymbol{\alpha_i^+}, \boldsymbol{\beta}, , x^i_{q_L}, x^i_{q_U}, q^i_L, q^i_U): j = 1;$

$STEP\,2:$ $\quad If\ j \neq i\ then\ BISECT2(\boldsymbol{\alpha_j^+}, x^j_q, \beta, q^j);$

$STEP\,3:$ $\quad If\ j < m\, then\, j:\ = j+1;\ Goto\, STEP\,2;$ $\qquad Else\, Goto\, STEP\,4;$

$STEP\,4:$ $\quad \alpha_1:\ = \alpha_1^+$

$STEP\,5:$ $\quad For\, i:\ = 2\ to\ m\ do\ \alpha_i:\ = \alpha_i^+ - \alpha_{i-1}^+$

Note that Steps 1 to 3 are identical to those in the algorithm to solve problem $\mathcal{P}_2$ associated with a $Dirichlet(\underline{\alpha}, \beta)$ distribution. As an example of the procedure above, note that the Example 2 in the second row of Table 2 satisfies the order restriction (52) since $x^1_{q_L} = 0.10,\ x^1_{q_U} = 0.30$ and $x^2_q = 0.50$. Note that the corresponding quantile levels $q^1_L = 0.20,\ q^1_U = 0.50$ and $q^2 = 0.40$ differ. From Table 2 it follows that

$$\alpha_1^+ = 0.3437; \alpha_2^+ = 0.5665; \beta = 2.6328$$

Executing Steps 4 and 5 in the algorithm above it follows that for this example

$$\beta = 2.6328; \alpha_1 = \alpha_1^+ = 0.3437;\ \alpha_2 = \alpha_2^+ - \alpha_1^+ = 0.5665 - 0.3437 = 0.2228. \qquad (41)$$

The probability density function of the Ordered Dirichlet distribution associated with (41) is presented in Figure 8.

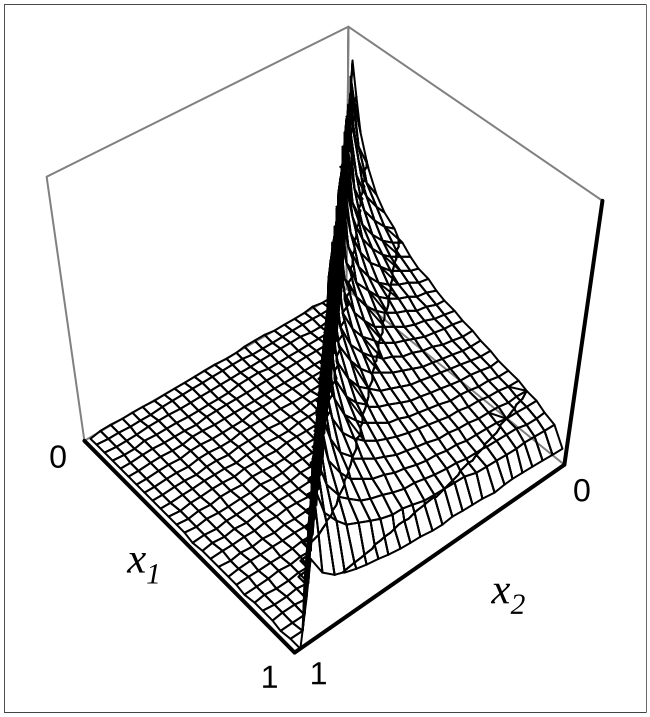

Figure 8. Ordered Dirichlet PDF associated with (41).

C. Transforming the Ordered Dirichlet Distribution and Numerical Stability

The use of an m-variate Ordered Dirichlet distribution to specify a prior distribution may be extended to the unbounded domain $(\mathbb{R}^{+})^{m}$ by transforming to $Y_i = -Ln(X_i)$ or $X_i = e^{-Y_i}$, where $\underline{X} = OD(\underline{\alpha}, \beta)$, $\underline{X} = (X_1, \ldots, X_m)$. Although the quantile levels q^j and q_U^i (or q_L^i), $j = 1, \ldots, m,$ may differ in the specification of problem $\mathcal{P}_3$ and in its solution, it is perhaps more practical from an expert judgment elicitation point of view to set e.g. $q^j = q_L^i = 0.50$ and $q_U^i = 0.95$ (which are quantile levels widely used in practice). Hence, measures of location for every Y_i are established by eliciting their median values and one could utilize the median (50% quantile) of Y_i and the 95% quantile for Y_i to determine the common shape parameter β. Table 1 below contains such median estimates $y_1^{0.50}$for failure rates of an exponential life time distribution in different stress environments (see, Van Dorp and Mazzuchi (2003)). In addition, the 95% quantile $y_1^{0.95}$ of the failure rate at Environment 1 (which is typically the use-stress environment in an accelerated life testing set-up) is provided as well.

Table 3. Environments, Prior Failure Rates

and Transformations

$Environment$	$Temp$ $(°F)$	$Volt$ (VDC)	$Prior\, y_i^{0.50}$ $(hours)^{-1}$	$Prior\, x_i^{0.50}$ $c = 1$	$Prior\, x_i^{0.50}$ $c = 841.61$
1	100	10.0	$5.036 \cdot 10^{-5}$	0.99995	0.9585
2	125	13.0	$1.100 \cdot 10^{-4}$	0.99998	0.9116
3	160	15.0	$5.732 \cdot 10^{-4}$	0.99943	0.6173
4	200	17.0	$1.429 \cdot 10^{-3}$	0.99857	0.3004
5	250	19.0	$3.781 \cdot 10^{-3}$	0.99623	0.0415
$y_1^{0.95}\, at\, use\, stress$			$1.315 \cdot 10^{-3}$	0.99869	0.3306

Note that the median failure rates $y_i^{0.50}, i = 1, \ldots, 5$, are very small (common to the accelerated life testing area) and increase rapidly when the stress in an environment increases (see, e.g., Van Dorp and Mazzuchi (2003)). The resulting transformed medians $x_i^{0.50}$ follow utilizing the transformation $x_i^{0.50} = e^{-y_i^{0.50}}$ and are provided in fourth column in Table 3. Note that these values for $x_i^{0.50}$, $i = 1, \ldots, 5$ are very close to 1.

Instead of the transformation $Y_i = -Ln(X_i)$ we may utilize a more general transformation

$$Y_i = \frac{-Ln(X_i)}{c} \Leftrightarrow X_i = e^{-cY_i} \tag{42}$$

to define a prior distribution on the domain $[0, \infty)^m$ via the Ordered Dirichlet distribution, where c is a preset transformation factor. As an example, the fifth column in Table 3 contains the transformed median values $x_i^{0.50} = e^{-cy_i^{0.50}}$ $i = 1, \ldots, 5$, where $c = 841.61$. Note that, these median values are now spread over the entire support $[0, 1]$ rather than being in the vicinity at 1. The motivation for utilizing (42) follows from (i) the value of the medians $y_i^{0.50}$, (ii) the Beta marginal distributions of X_i and (iii) the fact that no closed-form expression is available for the incomplete Beta function given $B(x \mid a, b)$ given by (4).

Several numerical algorithms exist to approximate the incomplete Beta function given by (4) (see, e.g., Press et al. (1989)). These approximations are well-behaved for parameter values $a \geq 1$, $b \geq 1$. However, in case $a < 1$, the Beta density *explodes* at $x = 0$, resulting in numerical instability for the approximation of $B(u \mid a, b)$ for the values close to $u = 0$. Vice versa, in case $b < 1$, the Beta density *explodes* at $x = 1$, resulting in numerical instability of $B(u \mid a, b)$ for values in the vicinity of $u = 1$. Figure 9A depicts the situation in Table 1 with $c = 1$ and medians $x_i^{0.50}$ close to 1. As the probability mass in the interval $[x_i^{0.50}, 1]$ has to account for 50% of the total probability mass, the associated Beta marginal densities will have an infinite mode at 1, resulting in numerically unstable behavior of evaluation of $B(x \mid a, b)$ (cf. (4)) when setting $c = 1$ into (42). The closer these median values $x_i^{0.50}$ are to the boundaries 1 (or 0), the larger will the numerical instability be in evaluating $B(x \mid a, b)$ (cf. (4)).

To reduce instability in numerical evaluation of $B(x \mid a, b)$ (cf. (4)) one may select a preset transformation factor c in (42) such that distance between median estimates $x_i^{0.50}, i = 1, \ldots, 5$, and the boundaries of the support $[0, 1]$ be as large as possible. Figure 9B depicts the motivation behind using (42) with a value of $c > 1$. The idea in Figure 9B is to choose a preset transformation parameter c such that the incomplete Beta function $B(x \mid a, b)$ corresponding to $x_5^{0.50}$ (the median at the highest stress level) is as well-behaved at $x = 0$ as $B(x|a, b)$ corresponding to $x_1^{0.50}$ (the median at the lowest stress level) at $x = 1$. The suggestion in Figure 7B is to select c such that $x_5^{0.50} = 1 - x_1^{0.50}$, or

$$e^{-c \cdot y_5^{0.50}} = 1 - e^{-c \cdot y_1^{0.50}}. \tag{43}$$

Unfortunately, (43) cannot be solved in closed form. General root finding algorithms may be used to solve for the transformation constant c numerically up to a pre-assigned level of accuracy. The resulting value for the transformation factor c utilizing $y_1^{0.50}$ and $y_5^{0.50}$ in Table 3 and (43) equals 841.61 as indicated in Table 3. Note that, indeed in the fifth column in Table 3, $x_5^{0.50} = 1 - x_1^{0.50}$. Utilizing the median $x_i^{0.50}$ in the fifth column of Table 3 the parameters of the $OD(\underline{\alpha}, \beta)$ may be solved by means of the algorithm described in the previous section.

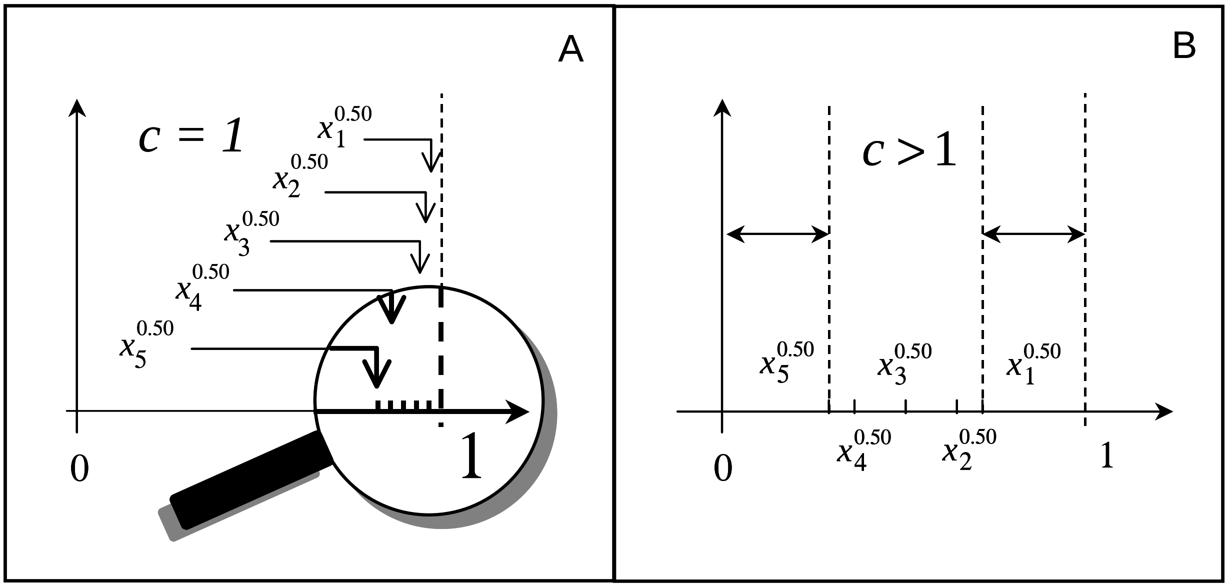

Figure 9. Transformation to medians $x_i^{0.50}$ $i = 1, \ldots, 5$ with Transformation Factor c

The resulting parameter values are:

$$\beta = 1.6656;\ \alpha_1 = 0.1522; \alpha_2 = 0.0481; \tag{44}$$

$$\alpha_3 = 0.2198; \alpha_4 = 0.2168; \alpha_5 = 0.2109;$$

It follows from (34) and (44) that

$$1 - \sum_{i=1}^{5} \alpha_i = 1 - \alpha_5^+ = 0.1552 = \alpha_1 = \alpha_1^+$$

indicating an identical numerical behavior when evaluating the cdf of X_5 and X_1.

It should be noted that (42) may be useful to increase numerical stability when transforming a $Dirichlet(\underline{\alpha}, \beta)$ distribution. In the latter case, the indices 1 and 5 in (56) would have to be replaced by those indices representing the largest and smallest median values in the $Dirichlet(\underline{\alpha}, \beta)$ distribution.

V. Conclusions

Algorithms have been developed for solving the prior parameters of the Beta, Dirichlet and Ordered Dirichlet distributions based on quantile constraints. The development of these algorithms involved a reparameterization which allows interpretation of these prior parameters in terms of a location parameter and a shape parameter $\beta > 0$. Limiting distributions of the Beta, Dirichlet and Ordered Dirichlet distributions follow in a transparent manner from the limiting behavior of the common shape parameter β. Existence of parameter solutions for the Beta, Dirichlet and Ordered Dirichlet distributions was proved utilizing their limiting distributions. We note in passing that the reparameterization advocated in this paper may be related to the orthoganality of parameters. We have not checked as yet whether indeed the condition for orthoganality as described for example in Cox and Reid (1987), is valid in our case.

Appendix

Let $X \sim Beta(\alpha,\, \beta)$ using the reparameterization given by (5). The bisection methods below use the numerical algorithm given in Press et al. (1989) to evaluate the incomplete Beta function $B(\,\cdot\,|\,a, b)$ given by (5). $BISECT1(\boldsymbol{x_q}, \alpha, \beta, q)$ solves for the q-th quantile x_q of X. Output parameters are indicated in bold. $BISECT2(\boldsymbol{\alpha^\circ}, x_q, \beta, q)$ solves for the parameter α° satisfying the quantile constraint (x_q, q). $BISECT3(\boldsymbol{\alpha^*, \beta^*}, x_{q_L}, x_{q_U}, q_L, q_U)$ solve for the required solution (α^*, β^*). A method to determine a starting interval $[a_1, b_1]$ containing β^* is given in $STEP\ 1$, $STEP\ 2$, and $STEP\,3$ of $BISECT\,3$.

$BISECT1(\boldsymbol{x_q}, \alpha, \beta, q)$:

$STEP\,1$ $\quad m :\, = 1;\ Set\,[d_1, e_1] = [0, 1]\,;$

$STEP\,2$ $\quad x_{q,m} :\, = \frac{d_m + e_m}{2}\ ;\ q_m :\, = B(x_{q,m}|\,\alpha, \beta);$

$STEP\,3$ $\quad If\ q_m \leq q\ then\ d_{m+1} :\, = x_{q,m}\,; e_{m+1} :\, = e_m;$

$\quad Else\ e_{m+1} :\, = x_{q,m}\,;\ d_{m+1} :\, = d_m;$

$STEP\,4$ $If\,|\,q_m - q| < \delta\,then\,x_q :\,= x_{q,m};\ Stop;$

$Else\ \ m :\,= m + 1; Goto\,STEP\,2;$

$BISECT2(\boldsymbol{\alpha}^{\circ}, x_q, \beta, q):$

$STEP\,1$ $n :\,= 1;\ Set\,[d_1, e_1] = [0, 1]\,;$

$STEP\,2$ $\alpha_{n+1} :\,= \frac{d_n + e_n}{2};\ q_n :\,= B(x_q \mid \alpha_{n+1}, \beta);$

$STEP\,3$ $If\ q_n \leq q\ then\,e_{n+1} :\,= \alpha_{n+1};\ d_{n+1} :\,= d_n;$

$Else\,d_{n+1} = \alpha_{n+1}\,; e_{n+1} :\,= e_n;$

$STEP\,4$ $If\,|\,q_n - q| < \delta\ then\ \ Stop;$

$Else\,n :\,= n + 1;\ Goto\,STEP\,2;$

$BISECT3(\boldsymbol{\alpha^*}, \boldsymbol{\beta^*}, x_{q_L}, x_{q_U}, q_L, q_U):$

$STEP\,1$ $k :\,= 1;\ \beta_{1,k} :\,= 1;$

$STEP\ 2$ $(\alpha^{\circ})_{\,1,k} :\,= \ BISECT2(x_{q_U}, \beta_{1,k}, q_U);$

$(x_{q_L})_{1,k} :\,= BISECT1(\,(\alpha^{\circ})_{\,1,k}, \beta_{1,k}, q_L);$

$STEP\,3$ $If\,(x_{q_L})_{1,k} < x_{q_L}\,then\,\beta_{1,k+1} :\,= 2 \cdot \beta_{1,k}; Goto\ STEP\,2;$

$Else\,[a_1, b_1] :\,= [\,0, \beta_{1,k}];$

$STEP\,4$ $k :\,= 1;$

$STEP\,5$ $\beta_k :\,= \frac{a_k + b_k}{2}; (\alpha^{\circ})_k :\,= BISECT2(x_{q_U}, \beta_k, q_U);$

$(x_{q_L})_k :\,= BISECT1(\,(\alpha^{\circ})_k, \beta_k, q_L);$

$STEP\,6$ $If\,(x_{q_L})_k < x_{q_L}\,then\,a_{k+1} :\,= \beta_k; b_{k+1} :\,= b_k;$

$Else\,a_{k+1} :\,= a_k; b_{k+1} :\,= \beta_k;$

$STEP\,7$ $If\ |\,(x_{q_L})_k - x_{q_L}| < \delta\,then\,\alpha^* :\,= (\alpha^{\circ})_k;\ \beta^* :\,= \beta_k; Stop;$

$Else\,k :\,= k + 1;\ Goto\,STEP\,5;$

Acknowledgment

The authors would like to extend our gratitude to Samuel Kotz for his most valuable comments in developing this chapter. We are indebted to Saralees Nadarajah for his editorial support.

References

1. Carlin, B.P. and Louis, T.A. (2000). Bayes and Empirical Bayes Methods for Data Analysis. Chapman & Hall/CRC, New York.
2. Chaloner, K.M and Duncan, G.T. (1983). Assessment of a Beta Prior Distribution: PM Elicitation. *The Statistician*, Vol. 32, pp. 174 - 181.
3. Chaloner, K.M and Duncan, G.T. (1987). Some Properties of the Dirichlet-Multinomial Distribution and Its use in Prior Elicitation. *Communications in Statistics, Theory and Methods,* Vol. 16., No.2, pp. 511 - 523.
4. Casella, G. and E.I. George (1992). Explaining the Gibbs Sampler. *The American Statistician*, Vol. 46, No. 3, pp. 167-174.
5. Cooke, R.M. (1991). Experts in Uncertainty: Opinions and Subjective Probability in Science. Oxford University Press, Oxford.
6. Coolen, F.P.A. (1997). Am Imprecise Dirichlet Model for Bayesian Analysis of Failure Data including Right-Censored Observations. *Reliability Engineering and Systems Safety*, Vol. 56, pp. 61-68.
7. Cowell, R.G. (1996). On Compatible Priors for Bayesian Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 9, pp. 901-911.
8. Cox, D.R. and Reid, N. (1987). Parameter Orthoganality and Approximate Conditional Inference. Journal of Royal Statistical Society, Series B, Vol. 49, No.1, pp.1-39.
9. Dennis III, S.Y. (1998). A Bayesian Analysis of Ratios of Proportions in Tree-Structured Experiments. *Journal of Statistical Planning and Inference*, Vol. 74, pp. 293-322.
10. Erkanli, A., Mazzuchi A. and Soyer, R. (1998). Bayesian Computations for a Class of Reliability Growth Models. *Technometrics*, Vol. 40., No.1, pp. 14-23.
11. Gavaskar, U. (1988). A Comparison of Two Elicitation Methods for a Prior Distribution for a Binomial Parameter. *Management Science*, Vol. 34, pp. 784 - 790.
12. Gilks, W.R., Richardson S. & Spiegelhalter, D. J. (1995). Markov Chain Monte Carlo in Practice, CRC Press; ISBN: 0412055511.
13. Harris, B. (1966). Theory of Probability. Addision-Wesley, Reading.

14. Johnson, W.O. and Kokolasis, G.E. (1994). Bayesian Classification Based on Multivariate Binary Data. *Journal of Statistical Planning and Inference*, Vol. 41, pp. 21-35.

15. Lancaster, T. (1997). Bayes WESML Posterior Inference from Choice-Based Samples. *Journal of Econometrics,* Vol. 79, pp. 291-303.

16. Lange, K. (1995). Application of the Dirichlet distribution to forensic match probabilities. *Genetica,* Vol. 96, pp. 107-117.

17. Kotz, S., Balakhrishnan, N and Johnson, N.L. (2000). Continuous Multivariate Distributions, Volume 1: Models and Applications (2nd Edition). Wiley, New York.

18. Kumar, S. and Tiwari, R.C. (1989). Bayesian Estimation of Reliability under a Random Environment Governed by a Dirichlet Prior. *IEEE Transactions of Reliability,* Vol. 38, No.2, pp. 218-223.

19. Martz, H.F. and Waller, R.A. (1982). Bayesian Reliability Analysis. Wiley, New York.

20. Neath A.N. and Samaniego, F.J. (1996). On Bayesian Estimation of the Multiple Decrement Function in the Competing Risks problem. *Statistics and Probability Letters*, Vol. 31, pp. 75-83.

21. Press, S.J. (1989). Bayesian Statistics, Principles, Models and Applications. Wiley, New York.

22. Press, W.H., Flannery, B.P., Teukolsky S.A. and Vettering W.T. (1989). Numerical Recipes in Pascal, Cambridge University Press, Cambridge.

23. Proschan, F. and Singpurwalla, N.D. (1979). Accelerated Life Testing - A Pragmatic Bayesian Approach. In: *J. Rustagi ed., Optimizing Methods in Statistics,* Academic Press, New York, pp. 385 -401.

24. Van Dorp, J.R., Mazzuchi T.A., Fornell, G.E. and Pollock L.R. (1996). A Bayes Approach to Step-Stress Accelerated Life Testing. *IEEE Transactions on Reliability*, Vol. 45, No.3, pp. 491-498.

25. Van Dorp, Mazzuchi, T.A. and Soyer, R. (1997). Sequential inference and decision making for single mission systems development. *Journal of Statistical Planning and Inference*, Vol. 62, No. 2, pp. 207-218.

26. Van Dorp, J.R. (1998) "A General Bayesian Inference Model for Accelerated Life Testing", Dissertation, The George Washington University, Washington, D.C..

27. Van Dorp, J.R. and Mazzuchi, T.A. (2000). Solving for the parameters of a Beta Distribution under two quantile constraints. *Journal of Statistical Computation and Simulation*, Vol. 67, pp. 189-201.

28. Van Dorp, J.R. and Mazzuchi, T.A. (2003). A General Bayes Exponential Inference Model for Accelerated Life Testing, To appear in *Journal of Statistical Planning and Inference.*
29. Weiler, H. (1965). The Use of Incomplete Beta Functions for Prior Distributions in Binomial Sampling. *Technometrics*, Vol. 7, pp. 335 - 346.
30. Wilks, S. S.(1962). Mathematical Statistics, Wiley, New York.