HAL
open science

Two prior distributions for incorporating information on a parameter $\theta \in [0, 1]$

Jérôme Dupuis

To cite this version:

Jérôme Dupuis. Two prior distributions for incorporating information on a parameter $\theta \in [0, 1]$. 2017. hal-01534675

HAL Id: hal-01534675

https://hal.science/hal-01534675v1

Preprint submitted on 7 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Two prior distributions for incorporating information on a parameter $\theta \in [0, 1]$

Jérôme Dupuis

IMT, Université Paul Sabatier, Toulouse, France

In Bayesian statistics, the available prior information on a statistical parameter θ is taken into account via a prior distribution. We consider the problem of incorporating prior information on a parameter $\theta \in [0, 1]$ consisting of a 95% credible interval and either the mean or the mode. Two prior distributions are considered for incorporated these two kinds of priors: the beta distribution and the two-sided power (TSP) distribution. It appears that the beta distribution is preferable to the TSP distribution when the prior involves the mean. The preference is reversed when the prior includes the mode. We show that, in this case, there exits a unique TSP distribution which allows us to incorporate such a prior. In all the other cases, results of existence are established, but uniqueness may not hold. That is why we advocate collecting from an expert the prior mode rather the prior mean, and incorporating the resulting prior via a TSP distribution.

Key Words: Bayesian statistics; Beta distribution; Elicitation parameters; Informative prior; Prior information; Two-side power distribution.

1. Motivations

Bayesian statistics permits to take into account the information available on a parameter θ of a statistical model: it is done via the prior distribution. (Throughout this paper, θ will designate as well the random variable as the parameter, as usual in Bayesian statistics.) When $\theta \in [0, 1]$, a beta distribution is often retained as prior distribution, on account of its great flexibility via a large choice of shapes of densities. If $\theta \sim$ beta (a, b), the problem is thus, in practice, to determine the parameters a and b from the available prior information.

When the prior information consists of $\mathrm{E}(\theta)$ and $\mathrm{Var}(\theta)$, it is well known that there is a unique beta distribution which is suitable for taking into account such a prior, provided that $\mathrm{Var}(\theta) < \mathrm{E}(\theta)[1-\mathrm{E}(\theta)]$. Berger (1980) suggests the use of two quantiles (eg the median and another quantile) of the prior distribution for elicating a and b, but does not tackle the problem of existence and uniqueness of the pair (a, b). Much later, Van Dorp and Mazzuchi (2003) state a general result of existence; but the uniqueness has not been proved by these authors, but only conjectured on the basis of numerical studies. Martz and Waller (1982) specify the parameters a and b from $\mathrm{E}(\theta)$ and a prior quantile, but these authors did not tackle the questions of existence and uniqueness of (a, b).

In this paper, we first consider the problem of specifying a and b when the prior information consists of $\mu = \mathrm{E}(\theta)$ and a 95% credible interval (that is an interval $I = [\alpha, \beta]$ such that $\Pr(\theta \in I) = 0.95$). Motivations for considering such a prior are mainly practical, as collecting the prior information under this form is in general an easy task, especially when the parameter of interest

has been the subject of previous studies; see eg Dupuis (1995), or Dupuis and Joachim (2011), for different illustrations in various contexts (see also the chapter 5 of the book of Marin and Robert, 2007). In these two articles finding a and b has been solved for some particular values of μ, α and β. The theoretical problems of existence or uniqueness of the pair (a, b) have never been considered, and one of the objectives of this paper is to fill this gap. A general result of existence is first stated, and we show that existence is ensured in most situations. Nevertheless, we also exhibit realistic situations for which there is no solution to the problem (see Section 2.1.1). As for the problem studied by Van Dorp and Mazzuchi (2003), and mentioned above, obtaining uniqueness results is difficult, and, in most situations we will have to make conjectures of uniqueness based on numerical studies. It appears that uniqueness should hold in most cases of practical interest, but there exist also realistic situations for which uniqueness does not hold (see Section 2.1.2). The situation is thus not entirely satisfactory and there is a need to propose alternatives. To this end, we explore two avenues.

- A first avenue is to collect (from the expert) the prior mode rather the prior mean, while keeping the beta law as prior distribution (note that the solution will have now to be searched within the family of unimodal beta distributions). From a purely practical standpoint, collecting the prior under this form should still be an easy task, insofar as the prior mode represents the most likely value of θ. In addition, the fact that the mode is a visual element of density makes it a particularly attractive alternative to the mean. We note that Kruschke (2015) advocates the use of the mode (instead of the mean) arguing that it is more intuitive than the mean. The capacity of

the beta distribution to incorporate such a prior is studied in Section 2.2. Concerning the existence of a solution, we state a result that is more general than previously; but, concerning uniqueness, the drawbacks mentioned above remain.

- A second avenue is to try another prior distribution for incorporating the two kinds of prior information above considered. To our knowledge, the two-sided power (TSP) distribution has never been considered for incorporating prior information on a parameter $\theta \in [0,1]$, whereas it is particularly flexible, just like the beta distribution (see Kotz and Van Dorp, 2004). Indeed, the TSP distribution on $[0,1]$ has two parameters, denoted by $m \in]0,1[$ and by $r > 0$. When $r > 1$, the parameter m coincides with the mode. The parameter r is closely linked to the variance (see Section 3). These considerations suggest that the TSP distribution could constitute a valuable alternative to the standard beta distribution. This issue is adressed in Section 3.

2. The prior distribution is beta(a, b)

If $\theta \sim$ beta (a, b) where $a > 0$ and $b > 0$, the prior density is

$$\pi(\theta) = \frac{1}{B(a,b)}\, \theta^{a-1}\, (1-\theta)^{b-1}\, \mathbb{1}_{[0,1]}(\theta)$$

where $B(a, b)$ is the standard beta coefficient. The beta distribution has been studied in detail: see eg Johnson, Kotz and Balakrishnan (1994). Let us recall some results which will be used afterwards. If μ and v respectively denote the mean and variance of θ, let us recall that one has:

$$\mu = \frac{a}{a+b} \quad \text{and} \quad v = \frac{ab}{(a+b)^2(a+b+1)}\,.$$

For convenience, we will reserve the term *unimodal density* for densities which have a mode located in $]0,1[$; it is easy to check that, with this convention, a density is unimodal if and only if $a > 1$ and $b > 1$. The mode will be denoted by m and is equal to $(a-1)/(a+b-2)$. Moreover, it will be usefull to reparametrize the beta distribution via (μ, λ) where $\lambda = a + b$. It is immediate that $v = [\mu(1-\mu)]/[1+\lambda]$, providing a concrete interpretation of λ (since, for fixed μ, $1+\lambda$ and v are inversely proportional).

2.1 The prior information consists of μ and a 95% credible interval

It is assumed that the prior information is composed of $\mu = E[\theta]$ and of an interval $I = [\alpha, \beta]$ containing μ such that $\Pr(\theta \in I) = 0.95$. The problem is to find a beta distribution which is able to incorporate such information, which reduces to finding the parameters a and b of such a beta distribution. Moreover, it is clear that, once μ is available, finding (a, b) involves finding λ such that $\Pr(\theta \in I) = 0.95$, since we return to the standard parametrization by applying the formulae $a = \lambda\mu$ and $b = \lambda(1-\mu)$. For fixed μ and I, $\Pr(\theta \in I)$ is a function of λ, which we denote by $H(\lambda)$, given by

$$H(\lambda) = \frac{\Gamma(\lambda)}{\Gamma(\lambda\mu)\Gamma(\lambda(1-\mu))} \int_\alpha^\beta \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1} \, \mathrm{d}\,\theta \tag{2.1}$$

where $\lambda > 0$. Any solution of the equation $H(\lambda) = 0.95$ is denoted by λ^*.

2.1.1. Existence

Theorem 2.1. For every μ, α, β such that $0 < \alpha < \mu < \beta < 1$, there exists a beta distribution such that $E[\theta] = \mu$ and $\Pr(\theta \in [\alpha, \beta]) = 0.95$.

• The proof of Theorem 2.1 proceeds in three steps. First, we show that $H(\lambda) \longrightarrow 1$ as $\lambda \longrightarrow \infty$ by using the Chebyshev's inequality. Second, we

show that $H(\lambda) \longrightarrow 0$ as $\lambda \longrightarrow 0$. Third, considering that the function h is continuous, we conclude by applying the intermediate value theorem on $]0, +\infty[$. Details appear in Appendix A.

• We have actually a more general result than the one stated in Theorem 2.1; it is obtained by replacing 0.95 by any $\gamma \in]0, 1[$. However, we will afterwards keep this value of 0.95 which is the one typically used in Bayesian statistics when reporting credible intervals.

• Note that the Theorem 2.1 excludes intervals of type $[0, \beta]$ (called intervals of type 1) and intervals of type $[\alpha, 1]$ (called intervals of type 2). Now, in practice, one may wish to use such intervals, especially if μ is small in the first case, and large in the second case. When the interval I is of type 1 or 2, $H(\lambda)$ no longer tends to 0 as $\lambda \longrightarrow 0$ because the density $\pi(.)$ is no longer bounded on I: we have the following result.

Proposition 2.2 Assume that I is of type 1; for fixed μ, $H(\lambda) \longrightarrow 1 - \mu$ as $\lambda \longrightarrow 0$. If I is of type 2, $H(\lambda) \longrightarrow \mu$ as $\lambda \longrightarrow 0$.

Note that, surprisingly, the limit does not depend on I; more precisely, it does not depend on β when I is of type 1, and not on α when I is of type 2. The proof of Proposition 2.2 uses inequalities of Volodine (1970) which relate the incomplete beta function: see Appendix B for details. The Theorem 2.3 below completes the Theorem of existence 2.1 when I is of type 1 or 2.

Theorem 2.3. For all $\mu > 0.05$ and all interval I of type 1, there exists a beta distribution such that $\Pr(\theta \in [\alpha, \beta]) = 0.95$ and $E[\theta] = \mu$. For all $\mu < 0.05$ and interval I of type 1 centered on μ, there exist a unimodal and a non-unimodal beta distribution such that $E[\theta]=\mu$ and $\Pr(\theta \in I) = 0.95$.

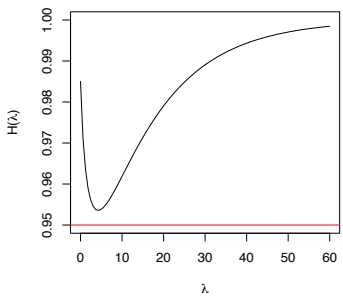

Figure 1: Graph of $H(\lambda)$ when $\mu = 0.015$ and $I = [0, 0.1]$.

The proof of Theorem 2.3 appears in Appendix C. In the particular case $\mu = 0.05$, the equation $H(\lambda) = 0.95$ has a unimodal solution if I is of type 1 and is centered on μ (see Appendix C). We have of course a similar theorem for intervals of type 2: details are omitted.

In summary, the existence of a solution is guaranteed in all configurations, except if $\mu \leq 0.05$ and I is of type 1 and is not centered on μ (or if $\mu > 0.95$ and I is of type 2 and not centered on μ). In such situations, the problem may have no solution. It is for example the case when $\mu = 0.05$ and $I = [0, 0.6]$, or when $\mu = 0.015$ and $I = [0, 0.1]$; Figure 1 provides the plot of the function H associated with the second example. A numerical study of H indeed shows that $H(\lambda) > 0.953$ (thus > 0.95) for all λ. The fact that the beta distribution may fail to incorporate certain priors constitutes of course a weakness. A possible strategy to circumvent it would be to consider intervals of type $[\epsilon, \beta]$ if $I = [0, \beta]$, and of type $[\alpha, 1-\epsilon]$ if $I = [0, \beta]$, where ϵ is chosen each time very close to 0, but > 0 so that existence Theorem 2.1 applies. Unfortunately, this strategy does not work, because the resulting prior beta distribution is very sensitive to the choice of ϵ. For example, returning to the example for

which $\mu = 0.015$ and $I = [0, 0.1]$, λ^* is equal to 60.93, 30.95, 19.00 if ϵ is respectively equal to 10^{-4}, 10^{-6}, 10^{-9}; these values have been obtained by solving the equation $H(\lambda) = 0.95$ using the command *solve* of R.

2.1.2. Uniqueness

We now examine the question of uniqueness. We have stated results only in two particular cases: first, for $\mu = 1/2$ (see below); second, for intervals I of type 1 or 2 and centered on μ. As for the problem considered by Van Dorp and Mazzuchi (2003), obtaining uniqueness results is in fact difficult. It is due to the fact the function H given by (2.1) is hard to manipulate analytically, even if its derivative can be calculated in a closed form; it is indeed easy to check that:

$$H'(\lambda) = \omega(\lambda)H(\lambda) + \int_\alpha^\beta [\mu \log(\theta) + (1-\mu)\log(1-\theta)]\, \pi(\theta) \mathrm{d}\, \theta$$

with

$$\omega(\lambda) = \Psi(\lambda) - [\mu\Psi(\lambda\mu) + (1-\mu)\Psi(\lambda(1-\mu))]$$

where $\Psi(\lambda) = \Gamma'(\lambda)/\Gamma(\lambda)$ denotes the digamma function. Interestingly, most mathematical software (in this paper we have systematically used R) allows us to visualize the graph of the function H, and then to locate the solution(s) of the equation $H(\lambda) = 0.95$, and finally to obtain them via the command *solve*. We will also use the plot of H yielded by R, either to conclude an absence of uniqueness for certain values of μ, α and β, or to form a conjecture of uniqueness.

• The only (positive) uniqueness result concerns the case $\mu = 1/2$.

Theorem 2.4. For any interval $I = [\alpha, \beta]$ containing $\mu = 1/2$, there exists a unique beta distribution such that $\Pr(\theta \in [\alpha, \beta]) = 0.95$.

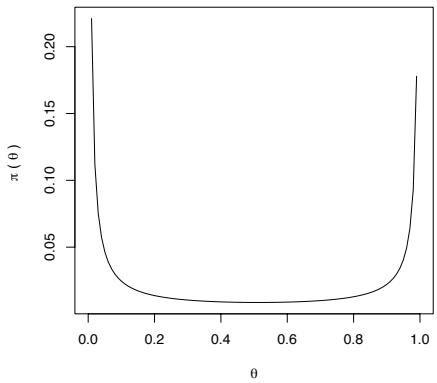

Figure 2: Non-unimodal solution

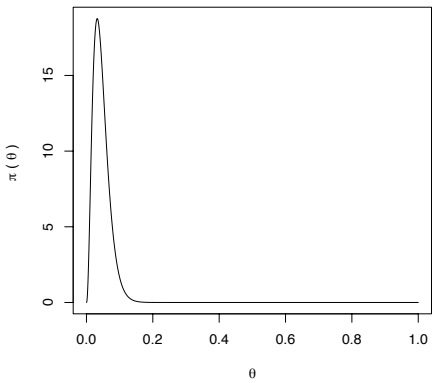

Figure 3: Unimodal solution

The proof appears in Appendix D.

• Conversely, Theorem 2.3 states that uniqueness does not hold when I is of type 1 and centered on $\mu < 0.05$ (or when I is of type 2 and centered on $\mu > 0.95$). We illustrate this situation by an example. When $\mu = 0.045$ and $I = [0, 0.09]$ the software R yields the solutions $\lambda^* = 0.052$ and $\lambda^* = 74.44$. The beta distribution associated with the smaller value of λ^* appears in Figure 2, and the one associated with the higher value in Figure 3. It is expected that, in practice, only the unimodal solution is of interest, insofar as the non-unimodal beta distribution solution of the problem gives a non negligible probability to the interval [0.5,1], which should constitute an undesirable characterictic in most situations; in our example, this probability is equal to 0.045 (while it is almost null with the unimodal solution).

• We now assume that I is of type 1 and is not centered on $\mu < 0.05$. In this configuration, uniqueness may not hold. For example, if $I = [0, 0.08]$ and $\mu = 0.02$ we observe that the problem has two solutions: $\lambda_1^* = 1.01$ and

$\lambda_2^* = 21.79$. It is easy to check that these two solutions do not correspond to unimodal beta distributions. Computing, for example, the probability that θ exceeds 0.02 should allow the expert to choose between the two solutions, since they strongly differ. Indeed, the probability is equal to 0.076 if the retained solution is 1.01 and to 0.311 if one retains the other solution.

• In all other configurations, we conjecture that uniqueness holds: that is when I is not of type 1 or 2 (this framework is the one of Theorem 2.1), or when I is either of type 1 with $\mu > 0.05$ or of type 2 with $\mu < 0.95$.

2.1.3. A first assessment

When $\mu \in]0.05, 0.95[$ the situation is quite satisfactory since the existence of a solution is ensured and uniqueness should hold. Concerning uniqueness, we miss theoretical results; however, in practice, the plot of the function H yielded by R consitutes a simple way to check that the conjecture of uniqueness effectively holds for the prior at hand.

When $\mu \leq 0.05$ or ≥ 0.95 some drawback and deficiencies occur. The fact that the problem may have no solution when I is type 1 or 2 and is not centered on μ, constitutes a weakness of the beta distribution. The fact that uniqueness may not hold when I is of type 1 with $\mu < 0.05$ (or when I is of type 2 with $\mu > 0.95$) should not constitute a difficulty in practice. However, such a situation will require some additional prior information to be collected from the expert for discriminating between the different solutions (as illustrated in Section 2.1.2).

2.2 The prior information consists of m and a 95% credible interval

Considering that the prior mode m (assumed to belong to $]0, 1[$) is a part

of the prior information, only unimodal beta distributions have to be considered, which implies that $a > 1$ and $b > 1$. We reparametrize the beta distribution via (m, λ). With this new parametrization, one has:

$$a = 1 + m(\lambda - 2) \quad \text{and} \quad b = 1 + (1 - m)(\lambda - 2)\,.$$

For fixed m and $I = [\alpha, \beta]$, $\Pr(\theta \in I)$ is a function of λ, where $\lambda > 2$; it is now denoted by $G(\lambda)$. We have:

$$G(\lambda) = \int_\alpha^\beta \frac{\theta^{m(\lambda-2)}(1-\theta)^{(1-m)(\lambda-2)}}{B\Big(1 + m[\lambda - 2], 1 + [1 - m][\lambda - 2]\Big)}\ \mathrm{d}\,\theta\,. \tag{2}$$

For fixed m and I, the problem is to solve the equation $G(\lambda) = 0.95$; any solution will be afterwards denoted by λ^*.

2.2.1. Existence

Theorem 2.5. For every m, α, β such that $0 \leq \alpha < m < \beta \leq 1$ and $\beta - \alpha < 0.95$, there exists a prior beta distribution having m as mode and such that $\Pr(\theta \in [\alpha, \beta]) = 0.95$.

• The proof proceeds as for the Theorem 2.1. We first show that $H(\lambda)$ tends to 1 as $\lambda \longrightarrow \infty$. Then, we show that $G(\lambda) \longrightarrow \beta - \alpha$ as $\lambda \longrightarrow 2$. As the function g is clearly continuous, we conclude by applying the intermediate value theorem on $]2, +\infty[$. Details appear in Appendix E.

• Interestingly, Theorem 2.5 is more general than Theorem 2.1, insofar as it does not exclude intervals of type 1 or 2; the price to pay is that the length of $I = [\alpha, \beta]$ should be < 0.95. But it is not really constraining in practice, because in an informative situation (which is the one considered in this paper) it will typically be satisfied.

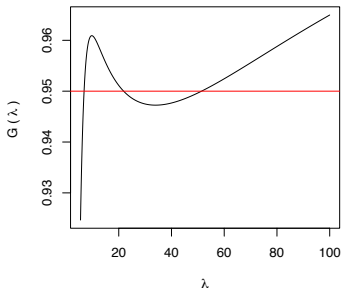

Figure 4: Graph of $G(\lambda)$ when $m = 0.02$ and $I = [0.007, 0.5]$.

2.2.2. Uniqueness

First, note that uniqueness Theorem 1.3 obviously applies to the particular case $m = 1/2$ (since the mean and the mode coincide when $m = 1/2$). Apart from this particular case, obtaining general uniqueness results is difficult in view of the expression of $G(\lambda)$. Consequently, as previously, we will have to form conjectures based on extensive numerical studies, and in particular on the plot of the function G yielded by R.

• When I is of type 1 and 2, we conjecture that uniqueness holds. This constitutes a major difference with the previous study (in which μ was part of the prior information).

• When $I = [\alpha, \beta]$ is not of type 1 or of type 2, we conjecture that uniqueness holds in general, but it may fail when $m < 0.05$ (or > 0.95). Let us give one example. Figure 4 provides the plot of G when $I = [0.007, 0.5]$ and $m = 0.02$. We observe that the equation $G(\lambda) = 0.95$ has three solutions. The software R gives $\lambda_1^* = 6.85$ and $\lambda_2^* = 21.85$ and $\lambda_3^* = 51.39$; Figure 5 provides the corresponding densities. A possible strategy for discriminating between the

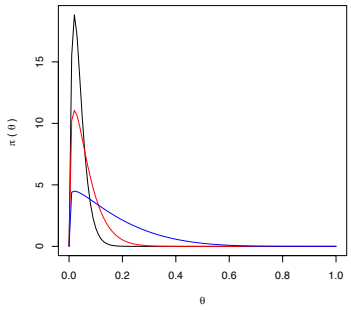

Figure 5: The three beta solutions when $m = 0.02$ and $I = [0.007, 0.5]$.

different solutions is to compare the probabilities of wisely chosen intervals; for example, computing $\Pr(\theta \in J)$ where $J = [0.2, 1]$ for each λ^* should allow us to retain (or to remove) the solution λ_1^*; indeed, this probability is equal to 0.31 when $\lambda = 6.85$, to 0.02 when $\lambda = 21.85$, and to 0.0002 when $\lambda = 51.39$. If one removes the solution λ_1^*, computing the probability that θ exceeds 0.1 should allow us to chose between λ_2^* and λ_3^*, since this probability is respectively equal to 0.20 and to 0.03.

2.2.3. A second assessment

As regards the capacity of the beta distribution to incorporate the prior considered in Section 2.2, the assessment is more positive than previously, the fact that existence is guaranted being particularly attractive. As for the previous study, uniqueness will have to be appreciated from the plot of the function G and, if uniqueness does not hold, additional prior information will have to be collected from the expert for choosing one solution.

3. The prior distribution is a two-sided power distribution

Recall that a random variable θ which takes its values in $[0,1]$ follows a standard two-sided power distribution with parameters (m,r), where $m \in]0,1[$ and $r>0$, if its density is given by:

$$\pi(\theta) = r\left(\frac{\theta}{m}\right)^{r-1} \mathbb{1}_{[0,m[}(\theta) + r\left(\frac{1-\theta}{1-m}\right)^{r-1} \mathbb{1}_{[m,1]}(\theta).$$

One can find a detailed study of the standard two-sided power distribution (in short TSP, omitting thus the term standard for convenience) in the book of Kotz and Van Dorp (2004). These authors provide, in particular, all the differents shapes of density we can obtain according to the values of m and r. Let us give some useful results. One has:

$$\mathrm{E}\,[\,\theta\,] = \frac{1+m(r-1)}{1+r} \quad \text{and} \quad \mathrm{Var}\,[\,\theta\,] = \frac{r-2m(m-1)(r-1)}{(r+1)^2(r+2)}\,.$$

Moreover it is easy (though tedious) to check that, for any fixed m, $\mathrm{Var}\,[\,\theta\,]$ is a strictly decreasing function of r; hence a concrete interpretation of r exists (as for the parameter λ of the beta distribution).

When $r=1$, the TSP distribution coincides with the uniform distribution on $[0,1]$ (whatever the value of m). When $r<1$ the density is convex and reaches its minimum when $\theta = m$; the density has no upper bound on $[0,1]$ (therefore it has no mode). The TSP is unimodal if and only if $r>1$; when $r>1$, the parameter m coincides with the mode. We provide, as illustration, the densities of the TSP distribution for $(m,r) = (0.3,10)$ in Figure 6, and for $(m,r) = (0.3,10)$ in Figure 7.

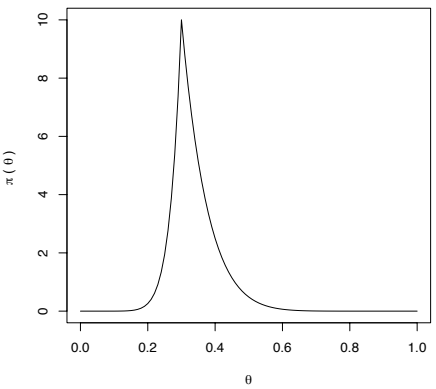

Figure 6: TSP (0.3,10)

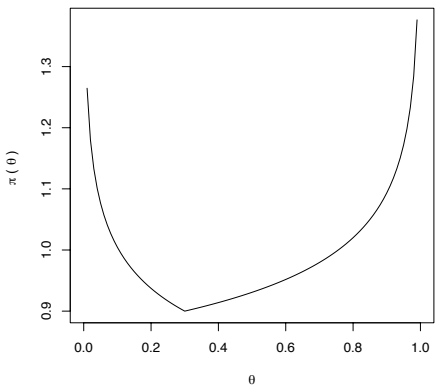

Figure 7: TSP(0.3,0.9)

3.1. The prior information consists of μ and a 95% credible interval

As for the beta distribution, it is possible to reparametrize the TSP distibution via its means μ and another parameter linked to the variance, here r. But, contrary to the beta distribution, working with the mean is much less natural, due to the presence of numerous constraints (see below). Without loss of generality, we can assume that $r \neq 1$ (since the uniform distribution on $[0, 1]$ is typically chosen in the absence of prior information).

The problem is to find (m, r) such that $E(\theta) = \mu$ and $\Pr(\theta \in I) = 0.95$, where $I = [\alpha, \beta]$ and μ are given and satisfy $0 \leq \alpha < \mu < \beta \leq 1$. The mean μ being given, and the fact that m should belong to $]0, 1[$ induces some constraints on r. It is indeed easy to check that one should have $r > \max\{\frac{1-\mu}{\mu}, \frac{\mu}{1-\mu}\}$ when $r > 1$, and $r < \min\{\frac{1-\mu}{\mu}, \frac{\mu}{1-\mu}\}$ when $r < 1$ (to ensure that $m \in]0, 1[$ when μ is given).

We will assume afterwards that $\mu \leq 1/2$; the developments for the case $\mu > 1/2$ are very similar and have been omitted for brevity.

We provide below a result of existence and uniqueness when I is of type 1 and when the solution is searched within the class of unimodal TSP distributions (which corresponds to TSP distributions for which $r > 1$).

Theorem 3.1. We assume that $\mu \in]0, \frac{1}{2}]$ and that I is of type 1; thus, in the class of unimodal TSP distributions there exists a distribution such that $\mathrm{E}(\theta) = \mu$ and $\Pr(\theta \in I) = 0.95$ if and only if $1-(1-\beta)^{\frac{1-\mu}{\mu}} < 0.95$; moreover, when such a TSP distribution exists, it is unique.

The proof appears in Appendix F. The condition $1 - (1 - \beta)^{\frac{1-\mu}{\mu}} < 0.95$ is called afterwards the condition (C).

For fixed μ and $I = [0, \beta]$, $\Pr(\theta \in I)$ is a function of r, denoted by $L(r)$; its expression is given in Appendix F.

• When I is of type 1 and centered on $\mu \leq 0.5$, the condition (C) becomes $1 - (1 - 2\mu)^{\frac{1-\mu}{\mu}} < 0.95$. It is straightforward to check that the function $\mu \longmapsto 1 - (1 - 2\mu)^{\frac{1-\mu}{\mu}}$ is strictly increasing on $[0, 1/2]$. The condition (C) is thus satisfied for every $\mu < \mu^*$, where μ^* denotes the solution of the equation $1 - (1 - 2\mu)^{\frac{1-\mu}{\mu}} = 0.95$; the software R yields $\mu^* \approx 0.4617$. Consequently, the condition (C) will, in particular, be satisfied for every interval I of length < 0.90, which is typically the case in informative situations.

• When I is of type 1 and not centered on $\mu \leq 0.5$, it is easy to exhibit quite commonplace priors where the condition (C) fails (which implies that the equation $L(r) = 0.95$ has no solution in $]1, +\infty[$). For example, the condition (C) is not fullfilled in the following cases: $\mu = 0.2$ and $I = [0, 0.6]$; $\mu = 0.1$ and $I = [0, 0.3]$; $\mu = 0.05$ and $I = [0, 0.15]$; $\mu = 0.04$ and $I = [0, 0.12]$.

• When, for fixed μ and $I = [\alpha, \beta]$, the problem has no solution in the class of unimodal TSP distributions (that is when $r > 1$), it remains to examine

if it has a solution when $0 < r < 1$. In fact, when $0 < r < 1$, the function L has a behavior which strongly depends on the values of μ, α and β; it may be strictly monotone (increasing or decreasing) or not. That means that a study of the function L has to be undertaken, case by case. However, we have observed that, in most situations, the equation $L(r) = 0.95$ has no solution in $]0, 1[$ when I is type 1 and is not centered on $\mu \leq 0.5$; it is, in particular, the case in all the four situations considered above. Therefore, for these four examples, there does not exist a TSP distribution able to incorporating the prior information.

The capacity of the TSP distribution to incorporate the prior information considered in this Section is thus seriously limited since it is not able to deal with quite commonplace priors (see above). That is why we consider that it is not useful to pursue this study further, and we recommend the use of the beta distribution for incorporating prior information consisting of μ and I.

3.2. The prior information consists of m and a 95% credible interval

In this Section the prior information consists of the prior mode m and of $I =]\alpha, \beta[$ such that $0 \leq \alpha < m < \beta \leq 1$ and $\Pr(\theta \in [\alpha, \beta]) = 0.95$. For fixed $m \in]0, 1[$ and $I = [\alpha, \beta]$, $\Pr(\theta \in I)$ is a function of r, denoted by $K(r)$. Note that r is necessarily > 1 since one assumes that the density has a mode. Contrary to the situations examined in Section 2, $\Pr(\theta \in I)$ can be written in a closed form. Indeed, it is easy to check that

$$K(r) = 1 - \left[m \left(\frac{\alpha}{m} \right)^r + (1 - m) \left(\frac{1 - \beta}{1 - m} \right)^r \right].$$

The theorem below provides a general result of existence and uniqueness.

Theorem 3.2 For every m, α, β such that $0 \leq \alpha < m < \beta \leq 1$ and $\beta - \alpha < 0.95$, there exists a unique unimodal two-sided power distribution having m as mode and such that $\Pr(\theta \in [\alpha, \beta]) = 0.95$.

The proof appears in Appendix G.

• As explained in Section 2.2.1, assuming that $\beta - \alpha < 0.95$ is actually not constraining, in informative situations. Note that, interestingly, Theorem 3.2 applies to intervals of types $[0, \beta]$ or $[\alpha, 1]$.

• We now briefly describe an algorithm which allows us to obtain, with the desired accuracy ϵ, the parameters of the TSP distribution solution of the problem. The mode m and interval $I = [\alpha, \beta]$ being fixed, Theorem 3.2 implies that there exists a unique $r_0 \in I\!N$ such that $[r_0, r_0 + 1[$ contains the solution (called r^*) of the equation $K(r) = 0.95$. For such a r_0, note that one has $U(r_0)U(r_0 + 1) \leq 0$ where $U(r_0) = K(r_0) - 0.95$. Two cases may occur: either $U(r_0) = 0$ (a rare occurence in practice) and then $r^* = r_0$, or $U(r_0)U(r_0+1) < 0$. In this second case, one finds r such that $|K(r)-0.95| < \epsilon$ by implementing a dichotomy on $]r_0, r_0+1[$. The code in R is available from the author on request.

5. Conclusion and extensions

We have showed that, in most cases, the beta distribution is able to incorporate prior information on $\theta \in [0, 1]$ composed of $\mu = E(\theta)$ and of $I = [\alpha, \beta]$ containing μ such that $\Pr(\theta \in I) = 0.95$ (it is not the case of the TSP distribution, as illustrated in Section 3.1). The fact that theoretical results of uniqueness are lacking should not prevent us from adopting the beta distribution for incorporing such a prior, insofar as one has the possibility

(via the plot of the function H yielded by R) to know if, for the prior at hand, uniqueness holds or not. If it fails, discriminating between the different solutions will require additional prior information to be collected from the expert. An appealing alternative to the prior mean is the prior mode. When the latter is a part of the prior information, the TSP distribution is clearly preferable to the beta distribution on account of the result of existence and uniqueness stated in this paper.

We believe that the procedures designed for incorporating the prior will be effectively used by non-statisticians, if certain conditions are met: first, the prior information has to be easy to collect (by involving meaningful and intuitive quantities); second, the existence and the uniqueness of the distribution chosen for incorporating the prior information have to be ensured; third, an algorithm which yields the parameters of the prior distribution solution of the problem has to be available. For a parameter $\theta \in [0, 1]$, the paper provides a contribution in this direction (via the TSP distribution and the use of the mode combined with a 95% credible interval). A natural extension of this work would be to examine other situations, such as: $\theta \in I\!R$, $\theta \in I\!R^+$ and $\theta \in I\!N$, while keeping the prior information available under the forms so far considered. For a parameter, $\theta \in I\!R$, the normal distribution is a natural choice. If $\theta \in I\!R^+$, the gamma distribution is a possible candidat because it can be re-parametrized via its mean and its variance. For a parameter $\theta \in I\!N$, we advocate the use of a negative binomial distribution, for the same reason. Studying the questions of existence and uniqueness is straightforward when $\theta \in I\!R$; on the other hand, the two other cases raise difficulties which are the subject of current research.

References

Alzer, H. (1997). On some inequalities for the gamma and psi functions. *Mathematics of Computation*, **66**, No 217, 373-389.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Second Edition. New York, Spinger-Verlag.

Dupuis, J. A. (1995). Bayesian estimation of movement and survival probabilities from capture-recapture data. *Biometrika* **82**, 761-772.

Dupuis, J. A., Bled, F., and Joachim, J. (2010). Estimating occupancy rate of spatially rare or hard to detect species: a conditional approach. *Biometrics* **67**, 290-298

Johnson, N. L., and Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions.* Vol. 2 (2nd ed.). New York, Wiley.

Kotz S. and Van Dorp J. R. (2004). *Beyond beta: other continuous Families of Distributions with Bounded Support and Applications.* World Scientific.

Kruschke, J., K. (2015). *Doing Bayesian Data Analysis.* Second Edition. Accademic Press.

Marin, J., M. and Robert C., P. (2007). *Bayesian Core: A practical Approach To Computational Bayesian Statistics.* Springer, New York.

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

Van Dorp, J.R. and Mazzuchi, T.A. (2003). Solving for the parameters of a beta distribution under two quantile constraints. *Journal of Statistical Computation and Simulation* **67**, 189-201.

Volodine, I., (1970). Beta-distribution with small parameters. *Theory of Probability and its Applications.* **15**, 546-549.

Appendix A

In this Appendix, μ, α, and β are fixed such that $0 < \alpha < \mu < \beta < 1$.

• Step 1.

It is clear that $\Pr(\theta \in I) > \Pr(|\theta - \mu| < a)$ where $a = \min(|\mu - \alpha|, |\mu - \beta|)$. Moreover the Chebyshev inequality implies that $\Pr(|\theta - \mu| > a) < \frac{Var(\theta)}{a^2}$. We thus have:

$$1 - \frac{\mu(1-\mu)}{a^2(1+\lambda)} \leq \Pr(\theta \in I) \leq 1$$

from which we deduce that $H(\lambda) = \Pr(\theta \in I) \longrightarrow 1$ as $\lambda \longrightarrow \infty$.

• Step 2.

To find the limit of $H(\lambda)$ as $\lambda \longrightarrow 0$ we write $H(\lambda)$ as the product of $\frac{\Gamma(\lambda)}{\Gamma(\lambda\mu)\Gamma(\lambda(1-\mu))}$ and $\int_\alpha^\beta \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1} \, \mathrm{d}\,\theta$.

a) Concerning the first term, it is well known that $\Gamma(\lambda) \sim 1/\lambda$ as $\lambda \longrightarrow 0$. Thus one has:

$$\frac{\Gamma(\lambda)}{\Gamma(\lambda\mu)\Gamma(\lambda(1-\mu))} \sim \lambda\mu(1-\mu);$$

from which we deduce that $\frac{\Gamma(\lambda)}{\Gamma(\lambda\mu)\Gamma(\lambda(1-\mu))} \longrightarrow 0$ as $\lambda \longrightarrow 0$.

b) Concerning the second term, we can write:

$$\lim_{\lambda\to 0} \int_\alpha^\beta \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1} \, \mathrm{d}\,\theta = \int_\alpha^\beta \lim_{\lambda\to 0} \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1} \, \mathrm{d}\,\theta\,.$$

The above permutation is justified as follows. First, it is easy to check that $[\theta^\mu(1-\theta)^{(1-\mu)}]^\lambda < 1$ for all $\mu \in]0,1[$, $\theta \in [\alpha,\beta]$ and $\lambda \in]0,+\infty[$. Therefore, the function $\theta \mapsto \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1}$ is uniformly dominated by the function $\theta \mapsto \frac{1}{\theta(1-\theta)}$ which is integrable on $[\alpha,\beta]$, which justifies the permutation. Thus, one has:

$$\lim_{\lambda\to 0} \int_\alpha^\beta \theta^{\lambda\mu-1}(1-\theta)^{\lambda(1-\mu)-1} \, \mathrm{d}\,\theta = \log\frac{\beta(1-\alpha)}{\alpha(1-\beta)}\,.$$

From a) and b) we deduce that $H(\lambda) \longrightarrow 0$ as $\lambda \longrightarrow 0$.

• Step 3.

It is clear that the function H is continuous on $]0, +\infty[$. Considering that $\Pr(\theta \in I) \longrightarrow 0$ as $\lambda \longrightarrow 0$ and that $\Pr(\theta \in I) \longrightarrow 1$ as $\lambda \longrightarrow +\infty$, we deduce, that the equation $H(\lambda) = 0.95$ has at least one solution in $]0, +\infty[$ (by appling the intermediate value theorem).

Appendix B

We only consider the case where I is of type 1, therfore $I = [0, \beta]$ where $0 < \beta < 1$. In what follows μ and β are fixed. The proof for the case I is of type 2 is very similar and is thus omitted.

For finding the limit of $H(\lambda)$ as $\lambda \longrightarrow 0$ we cannot proceed as in Appendix A, because, $\pi(\theta)$ is no longer bounded on $]0, \beta[$ when $a < 1$ (since θ^{a-1} tend to ∞ as $\theta \longrightarrow 0$). For obtaining this limit, we use inequalities stated by Volodine (1970). As in Volodine (1970) we denote by $I_x(a, b)$ the incomplete beta function which is defined as follows:

$$I_x(a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{t-1}(1-t)^{t-1} \, \mathrm{d}\, t \, .$$

Volodine (1970) showed that, for all a and b such $a+b < 1$, and all $x \in]0, 1/2]$, one has:

$$J_x(a, b) \leq I(x, a, b) < J_x(a, b) + \lambda^2$$

where

$$J_x(a, b) = \left(1 - \frac{a}{a+b}\right)\left(\frac{x}{1-x}\right)^a \, .$$

In our context, the inequalities of Volodine (1970) are writen as:

$$(1-\mu)\left(\frac{\beta}{1-\beta}\right)^{\lambda\mu} \leq \Pr(0 \leq \theta \leq \beta) < (1-\mu)\left(\frac{\beta}{1-\beta}\right)^{\lambda\mu} + \lambda^2 \, ,$$

where $0 < \beta \leq 1/2$. Insofar as we will make tend λ to 0, we can assume (without loss of generality) that $\lambda < 1$. Using the above inequalities, we deduce that, for every fixed μ and $\beta \in]0, 1/2]$, $H(\lambda) \longrightarrow 1-\mu$ as $\lambda \longrightarrow 0$ since $\left(\frac{\beta}{1-\beta}\right)^{\lambda\mu} \longrightarrow 1$ as $\lambda \longrightarrow 0$. If $\beta > 1/2$, the result still holds. Indeed, on can write $\Pr(0 \leq \theta \leq \beta)$ as the sum of $\Pr(0 \leq \theta \leq 1/2)$ and of $\Pr(1/2 \leq \theta \leq \beta)$, and the first term tends $1 - \mu$ while the second term tends to 0 (from the step 2 of Appendix A).

Appendix C

• The first part of the Theorem is an immediate consequence of Proposition 2.2. Indeed, assume that $\mu > 0.05$, then Proposition 2.2 implies that the limit of $H(\lambda)$ as $\lambda \longrightarrow 0$ is < 0.95. The existence of a solution to the equation $H(\lambda) = 0.95$ is thus an immediate consequence of the intermediate values theorem (recall that $H(\lambda) \longrightarrow 1$ as $\lambda \longrightarrow +\infty$).

• In what follows μ is fixed and < 0.05. Recall now that the density is unimodal (in the sense given in Section 2) if and only if $a > 1$ and $b > 1$. This condition is equivalent to $\lambda > \max\{\frac{1}{\mu}, \frac{1}{1-\mu}\}$. Consequently, when $\mu < 0.05$, the presence of a mode is equivalent to $\lambda > \frac{1}{\mu}$ (and its absence to $\lambda \leq \frac{1}{\mu}$) since $\frac{1}{\mu} > \frac{1}{1-\mu}$ is equivalent to $\mu < 0.5$. Considering this remark, for proving that there exist two beta distributions (one unimodal and the other non-unimodal) which are solutions of the problem, it is sufficient to show that $H(1/\mu) < 0.95$. Indeed, applying the intermediate value theorem first on $]0, \frac{1}{\mu}[$, and then on $]\frac{1}{\mu}, +\infty[$ will allow us to conclude, since $H(\lambda) \longrightarrow 1 - \mu > 0.95$ as $\lambda \longrightarrow 0$, and $H(\lambda) \longrightarrow 1$ as $\lambda \longrightarrow +\infty$. Now, one has:

$$H(1/\mu) = \frac{\Gamma(\frac{1}{\mu})}{\Gamma(\frac{1}{\mu} - 1)} \int_0^{2\mu} (1-\theta)^{\frac{1}{\mu}-2} \, \mathrm{d}\,\theta\,;$$

and it is easy to check that $H(1/\mu) = 1 - (1-2\mu)^{\frac{1}{\mu}-1}$. A study of this function shows that effectively $H(1/\mu) < 0.865 < 0.95$ when $\mu \in]0, 0.05]$, hence the result. In the particular case $\mu = 0.05$, we can only assert that there exits an unimodal solution (by applying the theorem of intermediate values on $]\frac{1}{\mu}, +\infty[$).

Appendix D

Thoughout this Appendix, $\mu = 1/2$; the mode and the median are thus also equal to $1/2$, since the density of such a beta distribution is symetric around $\theta = 1/2$.

To prove Proposition 2.3, we will show that the function H is a one-to-one function from $]0, \infty[$ to $]0, 1[$. Considering that the function H is continuous, $H(\lambda) \longrightarrow 1$ as $\lambda \longrightarrow +\infty$, and $H(\lambda) \longrightarrow 0$ as $\lambda \longrightarrow 0$ (see Appendix A), it remains to show that the function H is strictly increasing on $]0, \infty[$. To do that, we introduce functions u and v defined on $]0, \infty[$ by $u(\lambda) = \Pr(\theta \in [\alpha, 1/2])$, and $v(\lambda) = \Pr(\theta \in [1/2, \beta])$. To show that H is increasing, it is sufficient to show that u and v are both increasing (note that $H(\lambda) = u(\lambda) + v(\lambda)$ for all $\lambda \in]0, \infty[$). We limit ourselves to show that u is increasing (the proof for v being similar due to the symetry of the density of a beta $(\lambda/2, \lambda/2)$). We proceed in three steps.

Step 1. For $j = 1, 2$, we denote by $\pi_j(\theta)$ the density of the beta$(\lambda_j/2, \lambda_j/2)$, and by C_j the plot of $\pi_j(\theta)$ on $]0, 1/2]$. Note that C_1 and C_2 do not exhibit points of discontinuity because the functions π_1 and π_2 are both continuous on $]0, 1/2]$. In this first step, we show that, on $]0, 1/2[$, $C_1 \bigcap C_2$ is reduced to one point (whatever λ_1 and λ_2). To prove it, we assume that $C_1 \bigcap C_2 = \emptyset$,

and we show that this leads to a contradiction. Assuming that $C_1 \bigcap C_2 = \emptyset$ implies that either C_2 is located below C_1, or the reverse (recall that C_1 and C_2 do not exhibit point of discontinuity). Both configurations lead to a contradiction. Indeed, assume for instance that C_2 is above C_1; it will imply that $\pi_2(\theta) > \pi_1(\theta)$ on $]0, 1/2[$ and we will have $\int_0^{1/2} \pi_2(\theta)\, d\,\theta > \int_0^{1/2} \pi_1(\theta)\, d\,\theta$; this leads to contradiction since $\int_0^{1/2} \pi_2(\theta)\, d\,\theta = \int_0^{1/2} \pi_1(\theta)\, d\,\theta = 1/2$. Of course, assuming the reverse leads to the same contradiction. Therefore, C_1 and C_2 have at least one point of intersection. If we denote by i the abcissa of such a point, it is easy, starting from $\pi_1(i) = \pi_2(i)$, to check that i should satisfy:

$$i(i-1) = \left(\frac{\Gamma(\lambda_2)}{\Gamma(\lambda_1)}\right)^{\frac{\lambda_2-\lambda_1}{2}} \times \left(\frac{\Gamma(\lambda_1/2)}{\Gamma(\lambda_2/2)}\right)^{\lambda_2-\lambda_1} . \qquad (E)$$

Considering that the function $i \longmapsto i(i-1)$ is strictly monotone on $]0, 1/2[$, we deduce that there exists a unique $i \in]0, 1/2[$ which satisfies (E), and thus $C_1 \bigcap C_2$ is reduced to one point. Therefore, we have only two possibilities, either C_2 is located below C_1 on $]0, i[$ and above on $]i, 1/2[$, or the roles are reversed.

Step 2. To know which configuration is the good one, we compare $\pi_2(1/2)$ and $\pi_1(1/2)$. We introduce the function $S(\lambda)$ defined on $]0, \infty[$ by $S(\lambda) = \pi(1/2)$ where π represents the density of a beta $(\lambda\mu, \lambda(1-\mu))$ with $\mu = 1/2$. Thus one has:

$$S(\lambda) = \frac{1}{B(\lambda/2, \lambda/2)} \left(\frac{1}{2}\right)^{\lambda-2} .$$

It is easy to check that

$$S'(\lambda) = \frac{1}{B(\lambda/2, \lambda/2)} \left(\frac{1}{2}\right)^{\lambda-2} \left(\Psi(\lambda) - \Psi\left(\frac{\lambda}{2}\right) - \log(2)\right) ,$$

where Ψ denotes the digamma function. For studing the sign of $S'(\lambda)$, we use the inequality

$$\log(\lambda) - \frac{1}{\lambda} < \Psi(\lambda) < \log(\lambda) - \frac{1}{2\lambda} \tag{1}$$

where $\lambda \in]0, \infty[$ (see Alzer, 1997), from which we deduce that

$$\log\left(\frac{\lambda}{2}\right) - \frac{1}{2\lambda} < \Psi\left(\frac{\lambda}{2}\right) < \log\left(\frac{\lambda}{2}\right) - \frac{1}{\lambda}\,. \tag{2}$$

From (1) and (2) it is easy to check that $\Psi(\lambda) - \Psi(\lambda/2) > \log(2)$. Consequently, the function S is increasing on $]0, \infty[$ and $\pi_2(1/2) > \pi_1(1/2)$ if $\lambda_2 > \lambda_1$; in other words, at the point $\theta = 1/2$, C_2 is above C_1. Therefore, C_2 is necessarily located above C_1 on $]i, 1/2]$, and below on $]0, i[$. Consequently, $\pi_2(\theta) < \pi_1(\theta)$ on $]0, i[$ and $\pi_2(\theta) > \pi_1(\theta)$ on $]i, 1/2]$.

Step 3. We have to distinguish two cases, according to $\alpha \geq i$ or $\alpha < i$.

- If $\alpha \geq i$, it is clear that $u(\lambda_2) > u(\lambda_1)$ since $\pi_2(\theta) > \pi_1(\theta)$ on $]i, 1/2]$.

- If $0 < \alpha < i$, it is clear that $\int_0^\alpha \pi_2(\theta)\, d\,\theta < \int_0^\alpha \pi_1(\theta)\, d\,\theta$ since $\pi_2(\theta) < \pi_1(\theta)$ on $]0, i[$. Now, one has $\mathrm{u}(\lambda_j) = \frac{1}{2} - \int_0^\alpha \pi_j(\theta)\, d\,\theta$ for both $j = 1, 2$; hence $\pi_2(\theta) > \pi_1(\theta)$ on $]0, i[$. Therefore, $u(\lambda_2) > u(\lambda_1)$ when $\alpha \in]0, \frac{1}{2}[$.

Consequently, for all $\alpha \in]0, \frac{1}{2}[$, the function u is increasing on $]0, +\infty[$.

Appendix E

In this Appendix, m, α, and β are fixed such that $0 \leq \alpha < m < \beta \leq 1$. Recall that $\Pr(\theta \in I)$ is thus a function of λ we denote by $G(\lambda)$, and that we seek a solution in the class of unimodal beta distribution; consequently a and b are both assumed > 1.

• Step 1.

It is clear that $\Pr(\theta \in I) > \Pr(|\theta - m| < a)$ where $a = \min(|m - \alpha|, |m - \beta|)$. The Markov inequality implies that $\Pr(|\theta - m| > a) < \frac{E(\theta - m)^2}{a^2}$, from which

we deduce that

$$1 - \frac{E(\theta - m)^2}{a^2} \leq \Pr(\theta \in I) \leq 1\,.$$

Now, it is easy to check that:

$$E(\theta - m)^2 = \frac{\mu(1-\mu)}{1+\lambda} + \left(\frac{1-2\mu}{\lambda - 2}\right)^2$$

which tends to 0 as $\lambda \longrightarrow +\infty$; consequently, $G(\lambda) \longrightarrow 1$ as $\lambda \longrightarrow \infty$.

• Step 2.

Recall that $a = 1 + m(\lambda - 2)$ and $b = 1 + (1 - m)(\lambda - 2)$ (see Section 2.2). Consequently, $a \longrightarrow 1$ and $b \longrightarrow 1$ as $\lambda \longrightarrow 2$, from which we deduce that $G(\lambda) \longrightarrow \beta - \alpha$ as $\lambda \longrightarrow 2$.

• Step 3.

Applying the intermediate value theorem on $]2, \infty[$ to the continuous function G we deduce (from the steps 2 and 3) that the equation $G(\lambda) = 0.95$ has at least one solution if $\beta - \alpha < 0.95$; hence the result.

Appendix F

In this Appendix, I is of type 1. Moreover, β and μ are fixed such that $0 < \mu < \beta < 1$, and $\mu \leq 0.5$. Since we seek a solution of the problem within the class of unimodal TSP distributions, r belong to $]1, \infty[$; in fact the existence of a mode implies that $r > \frac{1-\mu}{\mu}$ (see Section 3). Note that, for fixed μ, the mode m is a function of r:

$$m(r) = \frac{(r+1)\mu - 1}{r - 1}\,,$$

this expression being deduced from the one for μ given in Section 3. The problem is thus to solve the equation $L(r) = 0.95$ on $]\frac{1-\mu}{\mu}, +\infty[$, where

$$L(r) = 1 - [1 - m(r)]^{1-r}[1 - \beta]^r\,.$$

This expression is directly deduced from the expression for $K(r)$ appearing in Section 3.2; indeed, $r > 1$ and $\mu < 0.5$ implies that $m < \mu$, and we are thus in the framework for using $K(r)$ which requires that $\alpha < m < \beta$. It is now easy to check that:

$$L'(r) = -\Big[(1-m(r)\Big]^{1-r} \Big[1-\beta\Big]^{r} \left[\frac{1-2\mu}{r(1-\mu)-\mu} + \log\frac{1-\beta}{1-m(r)}\right].$$

Now the sign of $\xi(r) = \frac{1-2\mu}{r(1-\mu)-\mu} + \log\frac{1-\beta}{1-m(r)}$ is not constant; we thus have to study the function $\xi(r)$. One finds:

$$\xi'(r) = \frac{\mu r - (1-\mu)^2}{(r-1)(r(1-\mu)-\mu)^2}.$$

Considering that $r > \frac{1-\mu}{\mu}$ implies $r > \frac{(1-\mu)^2}{\mu}$ (since $0 < 1-\mu < 1$), we deduce that $\xi'(r) > 0$ for every $r \in]\frac{1-\mu}{\mu}, \infty[$; moreover, $\xi(r) \longrightarrow \mu + \log(1-\beta)$ as $r \longrightarrow \frac{1-\mu}{\mu}$ and $\xi(r) \longrightarrow \log\frac{1-\beta}{1-\mu}$ as $r \longrightarrow +\infty$. Now, it is clear that $\log\frac{1-\beta}{1-\mu} < 0$ (since $\mu < \beta$). Consequently, $\xi(r) < 0$ for all $r \in]\frac{1-\mu}{\mu}, \infty[$, and the function L is thus strictly increasing on $]\frac{1-\mu}{\mu}, \infty[$. Considering that $L(\frac{1-\mu}{\mu}) = 1 - (1-\beta)^{\frac{1-\mu}{\mu}}$ we deduce that the function L is a one-to-one function from $]\frac{1-\mu}{\mu}, \infty[$ to $]1-(1-\beta)^{\frac{1-\mu}{\mu}}, 1[$, which concludes the proof.

Appendix G

1) It is easy to check that, for all $m \in]\alpha, \beta[$, $K(r) \longrightarrow 1$ as $r \longrightarrow \infty$.

2) It is clear that $K(r) \longrightarrow \beta - \alpha$ as $r \longrightarrow 1$.

3) The function K is increasing on $]1, +\infty[$; indeed it is easy to check

$$K'(r) = (1-m)^{r-1}(1-\beta)^{r} \log\frac{1}{(1-m)(1-\beta)} + m^{r-1}\alpha^{r} \log\frac{1}{m\alpha}$$

is > 0 on $]1, +\infty[$, since $\frac{1}{m\alpha} > 1$ and $\frac{1}{(1-m)(1-\beta)} > 1$ (recall m, $1-m$, α, and β are all in $]0, 1[$).

From 1), 2) and 3) we deduce that the function K is a one-to-one function from $]1, +\infty[$ to $]\beta - \alpha, 1[$. Therefore, for any fixed value of α, β, $m \in]\alpha, \beta[$ such that $0 \leq \alpha < m < \beta \leq 1$ and $\beta - \alpha < 0.95$, there exists a unique solution in $]1, +\infty[$ to the equation $K(r) = 0.95$; the result follows.