Hybrid elicitation and indirect Bayesian inference with quantile-parametrized likelihood

Dmytro Perepolkin* Benjamin Goodrich† Ullrika Sahlin‡

First version: 2021-06-29

This version: 2022-06-26

Abstract

This paper extends the application of indirect Bayesian inference to probability distributions defined in terms of quantiles of the observable quantities. Quantile-parameterized distributions are characterized by high shape flexibility and interpretability of its parameters, and are therefore useful for elicitation on observables. To encode uncertainty in the quantiles elicited from experts, we propose a Bayesian model based on the metalog distribution and a version of the Dirichlet prior. The resulting *hybrid* expert elicitation protocol for characterizing uncertainty in parameters using questions about the observable quantities is discussed and contrasted to parametric and predictive approach to elicitation.

1 Introduction

1.1 Parametric and predictive approach in elicitation

Bayesian parametric inference is about updating prior beliefs about the model parameters in light of new observations. The assumption is that the expert’s prior knowledge (or lack thereof) can be translated into the subjective probability distribution of model parameters through the process of elicitation (Winkler, 1967). The direct *elicitation of parameters* represents the *structural approach* to extracting the expert’s knowledge (Kadane, 1980). It requires that the expert understands the model and the role a particular parameter plays in it. Unfortunately, parameters may be abstract, difficult to interpret (like α and β parameters in the Gamma distribution), and sometimes not independent (like parameters in a hierarchical model).

*Centre for Environmental and Climate Science, Lund University, Sweden; dmytro.perepolkin@cec.lu.se

†Applied Statistics Center, Columbia University, New York, NY

‡Centre for Environmental and Climate Science, Lund University, Sweden; ullrika.sahlin@cec.lu.se

An alternative approach is to elicit the information about the observable quantities, possibly conditioned on the observable covariates (Kadane and Wolfson, 1998), which may be more intuitive and easy for the experts to relate to. Kadane and Wolfson (1998) advise against elicitation of moments, except, possibly, the first moment (the arithmetic average). The assessment should be done using the quantiles or probabilities of the predictive distribution. *Elicitation of predictions* aims at assessing the expert’s uncertainty about the future observation (Gelman et al., 2013). The difficulty with eliciting the predictive distribution is that it makes no distinction between the randomness explained by the model and the uncertainty about the parameters within it. Without this distinction, updating the expert predictions with the data coming from the new observations may be challenging.

Whereas the non-Bayesian elicitation literature often stops at the quantiles or probabilities related to the expert’s *predictive judgment* (Spetzler and Staël Von Holstein, 1975; Morgan, 2014; Keeney and von Winterfeldt, 1991; Hanea et al., 2021), the Bayesian literature attempts to devise a method to infer the prior distribution, which could have led to the particular predictions expressed by the expert (Akbarov, 2009; Hartmann et al., 2020; Winkler, 1980; Kadane and Wolfson, 1998). Kadane (1980) calls such elicitation of the predictive distribution, followed by the inference of the prior, the *predictive approach* since it uses the predictions to arrive at the distribution of parameters.

1.2 Aims of the paper

In this paper, we propose a **hybrid elicitation** approach, where the elicitation of observable quantities is accompanied by elicitation of the uncertainty around them. Hybrid elicitation, combining the features of the predictive and the structural approach to elicitation, can be used to define the prior distribution for a model defined by a quantile-parametrized distribution.

Quantile-parameterized distributions (QPDs) (Keelin and Powley, 2011; Hadlock, 2017) are parameterized by a set of quantile-probability pairs describing a random variable; therefore parameters in a QPD are on the same scale as the random variable they describe. These distributions can be used for characterizing either uncertainty about the future observation (predictive distribution) or a distribution of an unobservable parameter (prior distribution).

Until now there has been, to the best of our knowledge, no published research on how to update the quantile parameters of a QPD in light of the new observations. This paper extends the principles of *indirect Bayesian inference* (Rayner and MacGillivray, 2002; Nair et al., 2020; Perepolkin et al., 2021) to the models parameterized by quantiles and proposes a prior distribution, which is able to capture the uncertainty in the quantile parameters. The

proposed approach allows for elicitation and Bayesian updating of a variable quantity with minimal assumptions about the model structure.

1.3 Paper structure

Section 2 presents the method of *indirect inference* proposed by [Rayner and MacGillivray \(2002\)](#) and [Nair et al. \(2020\)](#), and summarized in [Perepolkin et al. \(2021\)](#). The core of the method is related to using one of the quantile-based distributions, which lack an explicit distribution function (CDF) and probability density function (PDF), as a prior and/or likelihood in a Bayesian model. Indirect priors and likelihoods rely on the substitutions based on the inverse distribution function, known as the quantile function (QF).

Then, in Section 3, we turn to a subclass of distributions parameterized by a set of quantile-probability pairs (see Figure 1). We review the literature on the different methods of constructing a quantile-parameterized distribution (QPD) and particularly focus on the quantile-based quantile-parameterized metalog distribution ([Keelin, 2016](#)) due to the flexibility of its parameters.

Figure 1: Probability distributions, quantile-based distributions and parameterization by quantiles.

In Section 4 we present the Bayesian model with the likelihood expressed by the metalog dis-

tribution parameterized by a set of quantile-probability pairs. We show how the uncertainty in these parameters can be specified using a variant of the Dirichlet distribution. While describing the elicitation of the QDirichlet prior in Section 5, we discuss a novel *hybrid* elicitation process for obtaining the quantile-probability pairs along with the uncertainty associated with them. We illustrate our method by the snippets of the interview and the code examples in the accompanying R package, called `qpd` (Perepolkin, 2019), which implements several quantile-parameterized distributions and contains the functionality to support the elicitation of the Dirichlet and the Connor-Mosimann distributions (Elfadaly and Garthwaite, 2013).

Section 6 discusses the MCMC-based algorithm for updating the parameters in a quantile-parameterized distribution. In this paper, we use Hamiltonian Monte Carlo MCMC in Stan interfaced by the `cmdstanr` package in R (Gabry and Češnovar, 2022). Alternative implementation using the Robust Adaptive Metropolis algorithm by Vihola (2012) interfaced by the `fmcmc` package (Vega Yon and Marjoram, 2019) in R is available in the Supplementary Materials.

We conclude the paper by discussion and summary of the results in Section 7.

2 Indirect Bayesian inference with quantile functions

The idea of using a non-invertible quantile-based distributions as a likelihood (Rayner and MacGillivray, 2002; King, 1999) or a prior (Nair et al., 2020) is not new to the scientific literature. Rayner and MacGillivray (2002) describe a three-step process for computing the log-likelihood for a quantile-based distribution and use it for the maximum likelihood estimation of the parameters in the *g-and-k* and the generalized *g-and-h* distributions. Nair et al. (2020) used the quantile function substitutions to express both prior and likelihood in a quantile form and calculated the posterior Bayes estimator of the parameters in the Govindarajulu model with uniform and generalized exponential priors. Perepolkin et al. (2021) summarized the approaches to indirect inference and provided several examples of applying the principles of indirect inference in univariate and regression settings.

Given a random sample of $\underline{x} = \{x_1, x_2, \dots, x_n\}$, the posterior distribution of θ over the parameter space Θ can be summarized by:

$$f(\theta|\underline{x}) \propto \mathcal{L}(\theta; \underline{x})f(\theta) \quad (1)$$

where $f(\theta|\underline{x})$ is the posterior distribution of θ after having observed the sample $\underline{x}$, $f(\theta)$ is the

106 prior distribution of θ , and $\mathcal{L}(\theta; \underline{x}) = \prod_{i=1}^n f(x_i|\theta)$ is the *direct* form of likelihood, because
 107 the observables $\underline{x}$ are used directly.

108 Consider a set of conditional probabilities, called *depths*¹, $\underline{u} = \{u_i|\theta\} = \{F(x_i|\theta)\}$,
 109 $i = \{1, 2, \dots n\}$, corresponding to the sample $\underline{x}$ of the observable x with distri-
 110 bution function F given the parameter θ . Using the depths $\underline{u}$ we can compute
 111 $\underline{Q} = \{Q_1(u_1), Q_2(u_2), \dots Q_n(u_n)|\theta\}$, where $Q(u|\theta) = F^{-1}(u|\theta)$ is the quantile func-
 112 tion or inverse CDF. Since $Q(u_i|\theta) = x_i$ we can substitute $\underline{Q}$ for $\underline{x}$, so the Bayesian inference
 113 formula (1) becomes:

$$f(\theta|\underline{Q}) \propto \mathcal{L}(\theta; \underline{Q})f(\theta) \quad (2)$$

114 We call this form of the likelihood $\mathcal{L}(\theta; \underline{Q}) = \prod_{i=1}^n f(Q(u_i|\theta)) = \prod_{i=1}^n [q(u_i|\theta)]^{-1}$ *indirect*,
 115 because it relies on computing of intermediate depths $u_i = F(x_i|\theta)$, $i = \{1, 2, \dots n\}$.

116 The two forms of likelihood $\mathcal{L}(\theta; \underline{Q})$ and $\mathcal{L}(\theta; \underline{x})$ are equivalent and produce the same poste-
 117 rior beliefs about the parameter θ (Perepolkin et al., 2021).

118 3 Quantile parameterization of distributions

119 In this section, we consider a special class of distributions, whose parameters are specified by
 120 quantile-probability pairs (Figure 1) and see how the concept of *indirect inference* (Perepolkin
 121 et al., 2021) can be applied to these distributions, as well.

122 A set of n quantile-probability pairs $\mathbf{S} = \{(p_i, q_i)\}, i = \{1, 2, \dots n\}$ can be viewed as consisting
 123 of a tuple-pair of ordered vectors: a vector of probabilities p and a vector of quantiles q ,
 124 $p = \{p_1, \dots p_n\}$, $p_i \in [0; 1]$, $q = \{q_1, \dots q_n\}$, $i = \{1, 2, \dots n\}$. Since CDF $F(x) = p$ is a non-
 125 decreasing function, vectors p and q are considered properly ordered iff $q_i \leq q_{i+1}, \forall q_i \in q$
 126 and $p_i \leq p_{i+1}, \forall p_i \in p$. Moreover, the quantile-probability pairs in the set $\mathbf{S}$ are considered
 127 distinct iff $\forall \{(p_i, q_i)\} \in \mathbf{S}, \exists! \{(p_j, q_j)\} = \{(p_i, q_i)\}, j \neq i, i = \{1, 2, \dots n\}, j = \{1, 2, \dots n\}$. In
 128 this paper we refer to the set of n distinct properly ordered quantile-probability pairs as a
 129 size- n *quantile-probability tuple* (QPT) denoted by $\{p, q\}_n$.

130 3.1 Parameterization using implicit functions

131 Keelin and Powley (2011) and Powley (2013) present the method of parametrizing a dis-

¹The conditional probabilities $\underline{u} = \{u_i|\theta\}$, $i = \{1, 2, \dots n\}$ are degenerate random variables fully deter-
 mined given the observations $\underline{x}$ and the value of the parameter θ . They are called *depths*, because they
 indicate how “deep” is certain observation into the distribution.

tribution by a set of quantile-probability pairs. The method is based on the finite Taylor series expansion of the parameters in a known inverse distribution function with the linear functions of the cumulative probability p .

The authors created Simple Q-Normal (SQN) distribution by taking the quantile function of a normally distributed random variable $x \equiv \mu + \sigma\Phi^{-1}(p)$ and making parameters μ and σ be the functions of p , namely $\mu(p) = a_1 + a_4p$ and $\sigma(p) = a_2 + a_3p$. Keelin (2016) proposed meta-logistic (*metalog*) distribution by making the parameters μ and s in the logistic quantile function $x \equiv \mu + s\text{logit}(p)$ be the functions of p , i.e. $\mu = a_1 + a_4(p - 0.5) + a_5(p - 0.5)^2 + \dots$ and $s = a_2 + a_3(p - 0.5) + a_6(p - 0.5)^2 + \dots$, where μ is the mean, s is proportional to the standard deviation such that $\sigma = s\pi/\sqrt{3}$, $\text{logit}(p) = \ln(p/(1 - p))$ is the log-odds of probability $p \in [0, 1]$ and a_i , $i = \{1, 2, \dots, n\}$ are real constants.

In both cases, the quantile function $Q(p)$ whose parameters also depend on p becomes an *implicit function*. This means that such a quantile function can not be simply computed for an arbitrary value of p . However, having a set of n quantile-probability pairs, it is possible to determine the constants a_i , $i = 1, 2, \dots, n$ by solving a system of n linear equations (Keelin and Powley, 2011; Powley, 2013).

Keelin and Powley (2011) show the conditions under which size- n QPT $\{p, q\}_n$ can uniquely determine the constants $a = \{a_1, \dots, a_n\}$ using the matrix equation

$$a = \mathbb{P}^{-1}q \quad (3)$$

The details of the metalog distribution (including the composition of the matrix $\mathbb{P}$) are provided in Appendix A.

The shape flexibility of the QPD increases with the number of terms added to the finite Taylor expansion of the parameters in the parent distribution. In order to obtain the estimates of the coefficients for the n -term quantile-parameterized distribution $a = \{a_1, \dots, a_n\}$ a minimum of n quantile-probability pairs are required. The order of the terms n is limited by the size of the parameterizing QPT m , $n \leq m$ and concerns about overfitting. QPT used for parameterization of a QPD can be obtained by the expert elicitation or from the empirical CDF (ECDF) constructed from a sample of observations. ECDF starts at zero and increases by $1/m$ at each of the m data points in the sample, representing the fraction of observations that are less than or equal to the specified value (Wasserman, 2006)

Depending on the relationship between the size of the parameterizing QPT m and the number of terms n in the QPD QF we use the following terminology:

- If the size of the parameterizing QPT m is equal to the number of terms n in the QPD QF, i.e. $m = n$, we refer to the process of estimating the vector of coefficients $a = \{a_1 \dots a_n\}$ as “fitting” and the resulting QPD as “properly parameterized”. For the properly parameterized QPD, the QF curve is guaranteed to go through every QPT point. We refer to the n -term metalog parameterized by the n -size QPT $\{p, q\}_n$ as the **proper n -metalog**.
- If the size of the parameterizing QPT m exceeds the number of terms n in the QPD QF, e.g. when QPT comes from the sample ECDF and $m > n$, we refer to the process of estimating the vector of coefficients $a = \{a_1 \dots a_n\}$ as “approximating” and the resulting QPD as “approximated”. Such approximation is usually done by optimization or regression, and the resulting QF curve is no longer guaranteed to go through every QPT point. We refer to the n -term metalog parameterized by the m -size QPT $\{p, q\}_m$, $m > n$ as the **approximate n -metalog**.

Given the matrix Equation (3), we can use two alternative parameterizations of the proper n -metalog: it can either be parameterized directly by the coefficients $a = \{a_1, \dots a_n\}$ (the A-parameterization) or indirectly by a QPT $\{p, q\}_n$ (the QPT-parameterization). Therefore, in this paper, when referring to the proper n -metalog, the notations $Q_{M_n}(u|a)$ and $Q_{M_n}(u|p, q)$ (where $u|a$ or $u|p, q$ represents the *depths* corresponding to the observation x) are used interchangeably. In case the metalog is approximated, only the A-parameterization is appropriate, as the number of metalog terms n is not implied by the m data points from the ECDF, used for estimating the parameter vector a using the Equation (9).

More details about the parameterization using the implicit functions, as a method of creating the quantile-parameterized distributions, can be found in [Powley \(2013\)](#) and [Hadlock \(2017\)](#).

3.2 SPT-parameterization

One of the earliest examples of a probability distribution parameterized by quantiles was proposed by [Myerson \(2005\)](#), who developed the generalization of the lognormal distribution, known as the Myerson distribution. The distribution is constructed from the lognormal density function, using the *symmetric percentile triplet* (SPT), a QPT of size 3 where the middle cumulative probability p_2 is equal to 0.5 (median) and $p_1 = 1 - p_3 = \alpha$, $\alpha \in (0, 0.5)$ (e.g. $\{0.25, 0.50, 0.75\}$ or $\{0.10, 0.50, 0.90\}$). If the quantile values in the parameterizing SPT triplet are equally spaced (e.g. 4, 7, 10), the resulting Myerson distribution is equivalent to the normal distribution. If the median quantile value is equal to the geometric mean of the top and bottom quantiles (e.g. 1, 4, 16), the resulting distribution is lognormal. The

Myerson distribution is bounded from the side of the narrower quantile range, similar to its parent, the lognormal distribution. The distribution has both CDF and QF and, therefore, it is ideally suited for capturing expert knowledge, be it the predictive judgement or the prior distribution of an uncertain parameter.

Hadlock and Bickel (2017) proposed an alternative family of the SPT-parameterized distributions based on the distribution system proposed by Johnson et al. (1994). When creating the new distribution the authors formulated five desiderata, for the properties of the new distribution: it should be quantile-parameterized, fully invertible, maximally feasible, highly flexible and be able to closely approximate the shapes of the lognormal, gamma, inverse-gamma and beta-prime distributions (Hadlock and Bickel, 2017). Johnson Quantile-Parameterized Distribution (J-QPD) and its generalization (Hadlock and Bickel, 2017, 2019) represent a flexible family of the quantile-parameterized distributions, which can accommodate a wide range of bounded and semi-bounded distribution shapes and can be conveniently parameterized by the three symmetrical points off of the CDF.

A special case of the SPT-parameterization exists also for the metalog distribution (Keelin, 2016). For the purpose of this paper we are not considering the SPT-parameterized QPDs: neither the SPT-metalog, nor Myerson or J-QPD due to their strict requirement for the symmetric probability parameterization. Since the cumulative probability parameters in these distributions are restricted by the requirement of symmetry around the median, uncertainty in the parameters would have to be admitted to the quantile values. This would require defining the covariance matrix to model the dependency between the quantile values. Eliciting the covariance matrix is, in our view, more difficult, than using the *hybrid elicitation*, proposed in this paper. Therefore, the hybrid elicitation approach amounts to eliciting the uncertainty in the cumulative probabilities corresponding to the fixed quantile values.

4 QDirichlet-Metalog model

In this section we propose a Bayesian model with the likelihood defined by the proper n -term metalog. Since metalog is a quantile-based distribution (Figure 1), we use the *indirect likelihood* following Equation (2). The indirect likelihood relies on the intermediate depths $\underline{u}|\theta$ corresponding to the sample of observations $\underline{x}$ as input. Since the closed-form CDF for the metalog distribution is not available, we use the numerical approximation of the $\widehat{Q}_X^{-1}(x) = \widehat{F}_X(x)$.

In order to consider the uncertainty in the QPT parameters of a quantile-parameterized distribution, such as metalog, the uncertainty needs to be admitted either to cumulative probabilities or the corresponding quantile values (or both). Coles and Tawn (1996) specified

the prior for an extreme value model in terms of the quantile values $(q_{p_1}, q_{p_2}, q_{p_3})$ for certain fixed cumulative probability values $p_1 > p_2 > p_3$. Crowder (1992) suggested that the priors can be constructed on the space of probabilities, for fixed quantiles. In the recent paper on predictive elicitation Hartmann et al. (2020) divided the observable space into several exhaustive and mutually exclusive categories and asked the experts to assign probabilities that the next observation falls into each of the categories. Then they treated the probabilities provided by the expert as uncertain and assigned a Dirichlet prior to these probability judgments.

We follow a similar route and use the quantile values provided by the expert as partitions of the outcome space. Then we characterize the uncertainty in the corresponding cumulative probabilities by the Dirichlet distribution (together referred to as the *QDirichlet* prior). Our approach to constructing a prior distribution for the simplex Δ is similar to the method adopted by Bürkner and Charpentier (2020) for modeling the monotonic effects in ordinal regression. The parameter vector of the Dirichlet distribution combined with the vector of elicited quantiles act as hyper-parameters of the proposed QDirichlet prior, which describes the uncertainty in the parameters of the quantile-parameterized model.

Consider a size- n QPT $\{p, q\}_n$, consisting of a vector of probabilities p and a vector of quantile values q . Now consider an extended vector of probabilities $b = \{0, p, 1\}$ of size $n + 2$ containing the vector p . Also consider a forward difference $\Delta = \{\Delta_1 \dots \Delta_{n+1}\}$, $\Delta_i = b_{i+1} - b_i$, $i = 1, 2, \dots (n + 1)$, which is a simplex of size $n + 1$. The simplex Δ is properly ordered iff it is based on the properly ordered vector b , and, therefore, also p .

The simplex Δ is converted back to the vector p using the cumulative sum operator $\Xi_1^n()$, so that $p = \Xi_1^n(\Delta) : p_j = \sum_{i=1}^j \Delta_i$, $j \in (1 \dots n)$, assuming the simplex Δ is properly ordered².

We can express the prior uncertainty in the simplex Δ by a Dirichlet distribution (Johnson et al., 1997) with the hyperparameter vector α of size $n + 1$ conditional upon the specified quantile values q . We refer to this variant of the Dirichlet prior as the *QDirichlet* prior, as its parameter vector α is specified in relation to the fixed quantile values q . Therefore our model becomes:

²If the simplex Δ for whatever reason can no longer be called “properly ordered” the index vector of distinct values $I = \{I_1 \dots I_n\} : I_j = \{1, 2, \dots n\}, j = \{1, 2, \dots n\}$, $\exists! I_j = I_i, j \neq i$ might be required to restore the proper order before the simplex Δ can be accumulated into the vector of probabilities p again.

$$\begin{aligned}
\Delta &\sim \text{Dirichlet}(\alpha|q); \\
p &= \Xi_1^n(\Delta); \\
Q(u|p, q) &\sim \text{Metalog}_{M_n}(p, q) \\
u &= \widehat{F}_{M_n}(x|p, q);
\end{aligned} \tag{4}$$

259 where Δ is a simplex of size $n + 1$, $\Xi_1^n()$ is the cumulative sum operator, p is a size- n tuple of
 260 cumulative probabilities, q is the size- n vector of quantiles, corresponding to the cumulative
 261 probabilities p , u is the depth corresponding to the observable x given the parameterizing
 262 QPT $\{p, q\}$, $\widehat{F}_{M_n}(x|p, q)$ is the approximated CDF of the metalog distribution, represented
 263 by the numerically inverted $\widehat{Q}_{M_n}^{-1}(x|p, q)$, indirectly parameterized by $\{p, q\}_n$ via the vec-
 264 tor of metalog coefficients a determined by the matrix Equation (3). The likelihood is rep-
 265 resented by the metalog density quantile function (DQF) $f(Q_{M_n}(u|p, q)) = [q_{M_n}(u|p, q)]^{-1}$.
 266 The model (4) can be visualized as a graph (Figure 2)

Figure 2: QDirichlet-Metalog model.

267 [Elfadaly and Garthwaite \(2013\)](#) describe a method for assessing the parameter vectors of
 268 Dirichlet and Generalized Dirichlet (Connor-Mosimann) distribution using the conditional
 269 univariate beta distributions. In this method the expert assesses the quartiles of the probabil-
 270 ity for each category using the symmetric percentile triplet (SPT) elicitation. After assessing
 271 the probability for the first category p_1 , the expert is asked to assume that the median value
 272 they provided in the assessment of p_1 , is in fact the correct probability (true value) for
 273 the first category. Then the expert is invited to proceed to assessing the SPT for the next
 274 category, conditional upon the previous assessment $p_2|p_1$. The three quartiles of $p_2|p_1$ will
 275 be divided through by $1 - p_1$ to arrive at the normalized quartiles of $p_2^*|p_1$ from which the
 276 hyperparameters of the beta distribution representing $p_2^*|p_1$ are obtained. The process is
 277 repeated for all categories except the last one. [Elfadaly and Garthwaite \(2013\)](#) propose an
 278 improvement to [Pratt et al. \(1995\)](#)'s method of fitting the beta distribution to the elicited
 279 conditional symmetric percentile triplets based on the normal approximation described in
 280 [Patel and Read \(1996\)](#). The α and β parameters of the conditional beta-distributions are
 281 then normalized and the hyperparameter vector $\alpha = \{\alpha_1 \dots \alpha_{n+1}\}$ is estimated.

The elicitation method proposed by [Elfadaly and Garthwaite \(2013\)](#) can be used to elicit the uncertainty in the cumulative probabilities parameterizing the metalog likelihood. If elicitation starts with the tail of the distribution, the first category will coincide with the first cumulative probability p_1 in the parameter vector p . Subsequent (higher) cumulative probabilities would always be conditional upon and include the median value of the lower probability³.

Note that the parameter vector p in the model (4) is not independent: it is tightly related to the vector of q denoting the fixed quantile values. There could be several approaches to specifying the hyperparameter vector q .

- **Predictive distribution.** The vector q could be coming from the characterization of the (prior) predictive distribution. In this case, we could ask the expert to specify his/her uncertainty in the next observation using the standard predictive elicitation techniques described in [Morgan et al. \(1990\)](#) or [Spetzler and Staël Von Holstein \(1975\)](#). Predictive elicitation results in the QPT $\{p^*, q^*\}_n$, of which vector q^* can be adopted as the true values of q . The vector of cumulative probabilities p^* can be used as the initialization values for the MCMC/HMC algorithm, as shown later.
- **Hypothetical sample.** The vector q could be viewed as a *representative sample* from the predictive distribution. The advantage of randomly sampling the predictive distribution is that the values closer to the mode of the distribution are more likely. The danger is that if the values of q are too close to each other, it might be difficult to fit the metalog to the QPT $\{p, q\}_n$ using the matrix Equation (3) within the MCMC loop.

In any case, the primary goal of eliciting the vector q is to *position* the prior on the data (x) scale and provide a reasonable baseline for the follow up elicitation. In fact, the hyperparameter vector q specifies the *location* of the QDirichlet prior, while the hyperparameter vector α is responsible for defining its *shape*.

In Equation (4) the prior is represented by the Dirichlet distribution with hyperparameter α specifying the uncertainty in the cumulative probabilities and a vector q representing the quantile values corresponding to the sampled cumulative probabilities (QDirichlet prior). The metalog (indirect) likelihood parameterized by the QPT $\{p, q\}_n$ relies on depths u which can be estimated using the numerical inverse of the metalog quantile function ([Perepolkin et al., 2021](#)).

³If the elicitation is performed out of order, which might be expedient to avoid anchoring ([Spetzler and Staël Von Holstein, 1975](#); [Abbas et al., 2008](#)), the integer index vector I of the same size can be provided along with the results of the assessment to restore the proper ordering of the simplex Δ after sampling and before its accumulation into the vector of probabilities p

5 Applications

In this section we provide an example of the *hybrid* elicitation for parameterizing the QDirichlet prior to describe the uncertainty in the $\{p, q\}_n$ parameters of the proper metalog.

5.1 Steelhead trout weights

We take a sample of 100 observations from the records of the fish weights of the steelhead trout caught and released in Babine River, Canada during 2006-2014 (Figure 3). The data was published by Keelin (2016), as well as included in the `rmetalog` (Faber and Jung, 2021) package available on CRAN.

Figure 3: Summary of the sample from the steelhead trout dataset

Our objective is to elicit the prior beliefs about the distribution of the fish weights from a hypothetical expert and then update those beliefs in light of the sampled data. Sections 5.2 and 5.3 document the elicitation process, as well as describe the necessary diagnostics for specification of the prior and the inference of the posterior.

5.2 Specifying location of the prior

The *hybrid* elicitation consists of two phases: elicitation of the quantile values q and elicitation of uncertainty in the cumulative probabilities associated with them (i.e. possible vectors of p which could correspond to the specified vector q).

Let's imagine that we conduct an interview with an experienced fly-fisherman to elicit the weights of the steel-head trout in the Canadian rivers. We begin with the elicitation of the predictive QPT following the probability-value (PV) method (Spetzler and Staël Von Holstein, 1975). After walking the expert through the important preparatory steps (motivating,

structuring, conditioning, encoding and verifying), we elicit the following predictive QPT
 $\{p^*, q^*\}_3$ (Table 1).

Table 1: Quantile-probability pairs for the predictive distribution of fish weights

Cumulative probability, p^*	Weight in lbs, q^*
0.1	4
0.5	9
0.9	17

Figure 4: Probability bands corresponding to the four fish size categories

The physical weight can be expressed using the non-negative values, so it should be modeled using the distribution which is bounded on the left. We can fit a quantile-parameterized distribution, such as the semi-bounded log-metalog to these QPT values (Figure 4). Note, that the three cumulative probabilities p^* provided by the expert divide the y-axis of the CDF into the four bands (highlighted by different colors in Figure 4). We can think of those as categories into which a randomly drawn fish could fall on the empirical CDF curve. Similar to the method adopted in Hartmann et al. (2020), the elicited quantile values q^* partition the outcome space into the exhaustive and mutually exclusive categories. The widths of the

bands are the increments in the cumulative probabilities provided by the expert, adding up to 1. They are represented by the simplex Δ in the model (Figure 2).

In the second step of the *hybrid* elicitation we elicit the uncertainty in these probability band widths Δ using the values q as anchors for delineating the categories to which the random variate $p_i \in p$, $i = \{1, 2, \dots n\}$ would correspond to.

5.3 Hybrid elicitation of the QDirichlet prior

It is worth noting that, in the *predictive elicitation*, the elicited QPT characterizing the prior predictive distribution $\{p^*, q^*\}_n$ do not have any uncertainty in them (except for the imprecision/inconsistencies in the expression of the expert’s own belief). In this phase of the elicitation, we ask the expert to consider uncertainty around his/her quantile assessments, and therefore, we attempt to separate the aleatory from the epistemic uncertainty (Knight, 1991). One possible drawback of starting with the predictive elicitation, as we did in our example, is that the provided vector p^* might be anchoring the expert on the specified values, therefore affecting the spread of probabilities that the expert could reasonably associate with the vector q . The expert might become tempted to allow only a small (and, likely, symmetrical) variation around the previously specified values of p^* and be more hesitant to revise their judgments. On the other hand, the specified p^* values can be used as a starting point for conversation with the expert, confirming and validating their belief and deliberately challenging the expert to revise them.

Whenever we use the word “aleatory” we should be thinking of sampling and therefore, we ought to change the assessment frame from assessing the properties of the population to assessing the properties of an imaginary sample. Elicitation in this phase is performed converting the cumulative probability scale to natural frequencies (Gigerenzer, 2011) by treating the values of $p_i \in p$, $i = \{1, 2, \dots n\}$ as proportions of a large hypothetical sample.

Interviewer:

Consider a large sample of steel-head trout caught in British Columbia over the last few years, say 1000 fish. According to your assessment there should be only 100 fish that would weigh below 4 lbs.

In this frame, the particular curve that we have drawn through the three points provided by the expert (Figure 4) is just one of many empirical CDF curves that can be drawn given the sampling uncertainty. Admitting the uncertainty in the QPT will allow us to update our beliefs about the distribution of observables, in this case, the fish weights.

We shall proceed with the elicitation, asking the expert to consider the fish weight cutoff of 4 lbs.

Interviewer:

Let’s talk about this hypothetical sample of 1000 fish. You said there will be 100 fish that weigh less than 4 lbs. We will interpret this assessment as you believing that there’s about equal chance that the actual number of “small” fish (smaller than 4 lbs) in this sample will be above or below 100, i.e we will interpret it as the median assessment. Would you like to reconsider this value?

At this point the expert might adjust the assessment of the median. Once the value of the median is confirmed, we can proceed with the elicitation of the range around it. From this point, elicitation takes the conventional *fixed probability* encoding method (Abbas et al., 2008; Spetzler and Staël Von Holstein, 1975): we ask the expert about the range of fish counts in the sample (which are in fact representing the cumulative probabilities) corresponding to the predetermined SPT (say, the quartiles or 10th/90th percentile).

Let’s say the expert provides the following uncertainty for the “small” fish count from the imaginary sample of 1000 fish, as summarized in Table 2. In this case the expert has provided us the revised median and the 50% IQR around the initially assessed probability $p_1^* = 0.1$.

Table 2: Symmetric percentile triplet for the count of the small fish (less than 4 lbs) in the 1000 fish sample

	Cumulative probability	Fish count
Small fish	0.25	70
	0.50	90
	0.75	120

Note:

All fish counts are out of the sample of 1000

From here we can immediately derive the uncertainty in the “width” of our first bin: it is now characterized by a symmetric percentile triplet $\{0.07, 0.09, \text{ and } 0.12\}$ with $\alpha = p_1 = 1 - p_3 = 0.25$. Using this SPT, we can fit the beta distribution (Figure 5) following the method proposed in Elfadaly and Garthwaite (2013).

We proceed to make the conditional elicitation of the probabilities for the other fish weight categories and uncertainties around them following method of Elfadaly and Garthwaite (2013), conditional on the median values of the previously elicited categories. We ask the expert:

Interviewer:

Let’s assume that in the 1000 fish sample the number of small fish (those that weigh under 4 lbs) was, in fact, found to be exactly 90. What would be your

Figure 5: Conditional beta distribution fitted to the quartiles provided by the expert

assessment of the number of fish that would measure between 4 and 9 lbs in such a sample?

Because in this question we are asking the expert for the conditional probability distribution, we don't have to keep the expert accountable for the previously made assessment, where they implied that there would be roughly half of the population weighing 9 lbs or less (as suggested by the cumulative probability of 0.5). We expect the median number to be close to 500 fish, but not necessarily precisely so. Our objective is to elicit the number of fish that weigh between 4 and 9 lbs, but if the expert prefers to provide us with the count corresponding to the “exceedance probability” (i.e. for 2 categories combined), we should deduct the median of the first category, i.e. 90.

For N groups we should only need to make $N - 1$ elicitation (3 elicitation of triplets total in our case). It might make sense to elicit the top quantile (the upper tail of the CDF) as $1 - \sum_{i=1}^3 p_i$ and leave the quantiles for the “Large fish” category (fish that weigh between 9 and 17 lbs) to be calculated from the rest. [Elfadaly and Garthwaite \(2013\)](#) provide detailed walk-through for eliciting the conditional SPTs and fitting beta distributions to them.

Let's assume that after the elicitation of the three conditional SPTs and converting the hypothetical sample counts to probabilities we got the following assessments (Table 3).

5.4 Fitting Dirichlet and Connor-Mosimann distributions

We can use the elicited conditional SPTs to arrive at the parameter vectors for the Dirichlet (Table 4) or Connor-Mosimann (Table 5) distributions. The algorithm of converting the conditional SPTs into the parameter vector(s) ([Elfadaly and Garthwaite, 2013](#)) is implemented in the `qpd` R package ([Perepolkin, 2019](#)).

Table 3: Conditional SPTs for the counts of the fish in the 1000 fish sample

	Category*	P25	P50	P75
1	Small fish	0.07	0.090	0.12
2	Medium fish	0.34	0.410	0.50
4	Huge fish	0.05	0.075	0.15

* Small fish is under 4 lbs, Medium fish is 4-9 lbs, Huge fish is over 17 lbs

Table 4: Dirichlet parameter vector

Category	a
1	3.77
2	12.86
4	2.70
3	10.72

Table 5: Connor-Mosimann parameter vectors

Category	a	b
1	5.87	55.24
2	6.77	8.01
4	1.32	5.25

5.5 Prior predictive check

Prior predictive checks are important for providing the expert with the feedback about the elicited values and for diagnosing the potential problems (Gabry et al., 2019). Since uncertainties in the quantile probabilities were elicited as conditional probabilities, it is important to show the expert the impact made by the provided probability ranges on the overall multivariate distribution. This can be done, for example, using the marginal plots. We can sample from the Dirichlet distribution (Figure 6) or the Connor-Mosimann distribution (Figure 7) and provide the expert with the overview of the distribution of parameters in the same format, as will later be used for the posterior predictive check (Figure 11 in Section 6).

The Dirichlet distribution defines a rigid negative dependence between the elements of the simplex Δ , and an increase in the probability of one element necessarily decreases the probability of every other element (Balakrishnan, 2014). The Connor-Mosimann distribution

relaxes this assumption of a strong negative correlation between the categories and allows for more flexible encoding of dependence between the quantiles (Wilson, 2017).

Figure 6: Prior predictive check for Dirichlet distribution

6 MCMC implementation

In order to sample from the posterior distribution of Δ , we used HMC algorithm in Stan interfaced via the `cmdstanr` package (Gabry and Češnovar, 2022) in R. Alternative implementation using Vihola (2012)’s Robust Adaptive Metropolis algorithm implemented in `fmcmc` package (Vega Yon and Marjoram, 2019) is provided in the Supplementary Materials.

We used 2500 post-warmup iterations and 4 chains (see Figure 10 for NUTS parameter plots). The posterior distribution of parameter Δ , presented in Table 6 and Figures 8 and 9 show the reduced uncertainty in the cumulative probabilities corresponding to the quantile values of 4, 9, and 17 lbs. The lowest value 4 lbs corresponds to the cumulative probability 0.03-0.09, while the upper value 17 lbs corresponds to 0.89-0.96 in cumulative probability, both 90% credible interval.

The posterior predictive check also shows reduction of uncertainty about the distribution of parameter Δ (compare prior predictive checks in Figures 6, 7 with posterior predictive check in Figure 11).

Figure 7: Prior predictive check for Connor-Mosimann distribution

Figure 8: Summary of the posterior draws for Δ simplex

Figure 9: Pairs plot for the posterior draws from Δ simplex

Figure 10: NUTS parameters for QDirichlet-Metalog model with indirect likelihood

Table 6: Posterior sample summary for the simplex Δ

category	mean	median	q5	q95	rhat
Small	0.0552	0.0536	0.0306	0.0862	1.000
Medium	0.4328	0.4318	0.3763	0.4927	1.000
Huge	0.0724	0.0703	0.0427	0.1093	1.001
Large	0.4396	0.4398	0.3803	0.4986	1.000

Figure 11: Posterior predictive check for the QDirichlet-Metalog model

7 Discussion

7.1 Uses of quantile-parameterized distributions in elicitation

In the last 20 years *quantile-based distributions* have transitioned from being a relatively exotic area of scientific interest to being one of the heavily researched and developed topics in statistics and reliability engineering. The interest in the scientific community was fueled by the development of the highly flexible and versatile distributions, such as the Generalized Lambda Distribution (Freimer et al., 1988), the g -and- k distribution (Rayner and MacGillivray, 2002), the Govindarajulu distribution (Nair et al., 2012) and their various modifications and generalizations (Aldeni et al., 2017) (Figure 1). At the same time the research area related to the *quantile-parameterized distributions*, remain largely unexplored.

Describing a probability distribution by the quantiles is not a new concept in the scientific literature. [Jeffreys \(1939\)](#) was perhaps the first to mention the idea of the *substitution likelihood* in a Bayesian context. He proposed to use a set of sample quantiles for making the inference about the population median. Alternative bootstrap-based formulation of the *substitution likelihood* was proposed by [Boos and Monahan \(1986\)](#). Modern developments of [Jeffreys \(1939\)](#) ideas can be found in [Lavine \(1995\)](#) and [Dunson and Taylor \(2005\)](#).

In the last two decades several researchers proposed new probability distributions with interpretable parameters defined on the same scale as observable quantities ([Myerson, 2005](#); [Keelin and Powley, 2011](#); [Hadlock and Bickel, 2017](#)). The aim behind the quantile-parameterized distribution research is to simplify the elicitation process and make the job of the experts easier. In our proposed *hybrid* elicitation framework tailored specifically for the models with the QPD likelihood, experts are invited to adopt a sampling frame and turn the difficult task of expressing the uncertainty about the cumulative probability into an easier task of expressing the uncertainty about the natural frequency in a hypothetical sample ([Gigerenzer, 2011](#); [Hoffrage et al., 2002, 2015](#)). Original elicitation method proposed by [Elfadaly and Garthwaite \(2013\)](#) asks the expert to assume that the median of the previously assessed category p_{i-1} is, in fact, the true value of the probability for the category $i - 1$, and elicits the value p_i for the next category i , conditional on it. We believe that such conditioning can be made simpler if one adopts the natural frequency framework. The inputs elicited from the expert in the natural frequency frame can be easily validated by simulation, providing the expert with the instant feedback on what their judgement means for the model.

Many of the popular elicitation protocols used for both *predictive* and *parametric* elicitation rely on the assessment of QPTs. Instead of fitting a parametric distribution to the elicited QPT ([O’Hagan, 2019](#)), assessors could use the elicited QPT directly as an input into one of the QPD QFs, which, as we have shown, can be easily used in both quantile-parameterized and parametric models. Quantile-parameterized distributions are guaranteed to fit the provided points off of the CDF curve exactly. Therefore, QPDs are able to represent expert’s belief with *high fidelity* making no trade-offs between convenience and precision. There can be minor differences between the shapes of the various quantile-parameterized distributions fitted to the same QPT. However, given the variety of QPDs proposed in the literature, a well-equipped assessor should be able to select a suitable distribution and validate the choice with the expert, accounting for the width of the distribution tails.

7.2 Quantile-parameterized vs parametric distributions in Bayesian analysis

Quantile-parameterized distributions have a unique type of parameterization: they are parameterized by a set of quantile-probability pairs, or as we referred to in this paper, by a QPT $\{p, q\}_n$. For metalog, such parameterization is made possible by the virtue of the matrix Equation (3). When referring to the metalog’s flexibility in the parameters, we are highlighting the fact that there’s no requirement for the p vector in the parameterizing QPT to be necessarily symmetric, as the case is with SPT-parameterized QPDs (e.g. J-QPD, and Myerson). In order to characterize the uncertainty in the parameterizing QPT with a Dirichlet distribution, we need full flexibility in the probability vector p (which is assembled from the *varying* simplex Δ), barring the feasibility conditions. The quantiles vector q used as a hyperparameter in the QDirichlet-metalog model, contains the quantiles of a random variable the distribution describes.

The quantile-parameterized and the parametric likelihoods in Bayesian analysis are, in our view, complimentary. Defining a model with the likelihood expressed by a quantile-parameterized distribution could be a starting point for understanding the phenomenon, when the only input available is the QPT judgement from the expert, there is no strong choice for a specific parametric distribution to use and the data is scarce. As the data become more abundant and understanding of the data-generating process is gained, a switch to a parametric likelihood can be justified.

All of the indirect likelihood models (i.e. models with the likelihood represented by a quantile-based distribution) rely on the numerically inverted CDF for estimating the *depths* corresponding to the sample (given parameters). [Perepolkin et al. \(2021\)](#) discussed various algorithms which can be used for numerical inversion of QFs, which is also relevant for the metalog likelihood.

7.3 Less informative priors

The QDirichlet-Metalog model described in this paper can be applied in combination with the predictive approach to elicitation ([Kadane, 1980](#)). Assuming that the only information elicited from the expert is the predictive QPT $\{p, q\}_n$, the quantiles q vector can be combined with a *uniform* Dirichlet prior, allowing data alone to define the posterior for the simplex Δ . Since the Dirichlet distribution is a generalization of the Beta distribution to higher dimensions, a weakly informative prior can be specified with a unit vector i.e. $Dirichlet(1, 1, \dots, 1)$. We discuss inference using weakly informative priors in Appendix B.

7.4 Hybrid elicitation

Parametric elicitation aims to describe the epistemic uncertainty contained in the *parameters* of the model with the help of the experts. Predictive elicitation describes the uncertainty in the next *observation* without distinguishing between the randomness in the model and the lack of knowledge about the model parameters. Mikkola et al. (2021) proposed the prior elicitation hypercube with 7 dimensions related to the elicitation of prior distributions (Figure 12).

PRIOR AND MODEL	D3. Elicitation space - parameter vs observable space	EXPERT
	D4. Elicitation model - fitting vs supra-Bayesian	
D1. Properties of the distribution - univariate vs multivariate - parametric vs nonparametric - prior-specific vs prior-agnostic		D6. Interaction with experts - assessment tasks - one-shot vs iterative - active elicitation - multiple experts
D2. Model family - model-specific vs model-agnostic	D5. Computation	D7. Expert capabilities - domain knowledge - statistical understanding - heuristics and biases

Figure 12: Prior elicitation hypercube

Following this classification, the proposed *hybrid elicitation* is a univariate, parametric, prior-specific (D1), model-specific (D2) method, performed in the observable space (D3). *Hybrid elicitation* uses the fitting approach proposed by Elfadaly and Garthwaite (2013) for arriving at the parameter vector(s) of the (Generalized) Dirichlet distribution (D4), which relies on the simple arithmetical computation (D5) to transform the parameters of the conditional marginal beta distributions into the (Generalized) Dirichlet parameter vector(s). Finally, *hybrid elicitation* relies on the active, iterative elicitation (D6), making minimum assumptions regarding the expert’s familiarity with the statistical concepts (e.g. subjective probability) and does not require any understanding of the statistical model (D7).

Hybrid elicitation starts off by describing the next observation using a QPT, but then pivots to characterization of uncertainty contained in the assessment of the QPT itself. This is accomplished by describing a hypothetical sample from the target population, which corresponds to the cumulative probabilities. These probabilities, along with a set of quantile values, can serve as the parameters in the quantile-parameterized model. Hybrid elicitation, like predictive elicitation, describes only observable quantities. At the same time, like parametric elicitation, the *hybrid elicitation* results in the characterization of uncertainty in the model parameters. Hybrid elicitation, therefore, can be viewed as the observations-level parametric elicitation for quantile-parameterized models.

Miscellaneous

Data Availability Statement

The `qpd` R package used in this paper is available on Github at <https://github.com/dmi3kno/qpd>. Contact corresponding author Dmytro Perepolkin (Dmytro.Perepolkin@cec.lu.se) for requests for data.

Supplemental materials

Supplemental materials contain the R and Stan code for all examples used in the article.

Acknowledgments

The authors have no conflict of interest to declare. D Perepolkin and U Sahlin were funded by the strategic research environment Biodiversity and Ecosystem Services in Changing Climate (BECC). U Sahlin was also funded by the Crafoord Foundation (ref 20200626). B Goodrich is supported by U.S. National Science Foundations grants 2051246 and 2153019.

ORCID

Dmytro Perepolkin <https://orcid.org/0000-0001-8558-6183>

Ullrika Sahlin <http://orcid.org/0000-0002-2932-6253>

References

Appendices

Appendix A. Metalog distribution

Fitting and parameterization

This section recapitulates ideas and formulas provided in Keelin (2016) with our own notation and minor reinterpretations.

Metalog distribution is created from the logistic quantile function $Q(p) = \mu + s \text{slogit}(p)$, where μ is the mean, s is proportional to the standard deviation such that $\sigma = s\pi/\sqrt{3}$, p is the probability $p \in [0, 1]$. The metalog quantile function is built by substitution and series expansion of its parameters μ and s with the polynomial of the form:

$$\begin{aligned}\mu &= a_1 + a_4(p - 0.5) + a_5(p - 0.5)^2 + a_7(p - 0.5)^3 + a_9(p - 0.5)^4 + \dots, \\ s &= a_2 + a_3(p - 0.5) + a_6(p - 0.5)^2 + a_8(p - 0.5)^3 + a_{10}(p - 0.5)^4 + \dots,\end{aligned}\tag{5}$$

where a_i , $i \in (1 \dots n)$ are real constants. Given a size- m QPT $\{p, q\}_m$, where $p = \{p_1 \dots p_m\}$ and $q = \{q_1 \dots q_m\}$ the vector of coefficients $a = \{a_1 \dots a_m\}$ can be determined through the set of linear equations.

$$\begin{aligned}q_1 &= a_1 + a_2 \text{logit}(p_1) + a_3(p_1 - 0.5) \text{logit}(p_1) + a_4(p_1 - 0.5) + \dots, \\ q_2 &= a_1 + a_2 \text{logit}(p_2) + a_3(p_2 - 0.5) \text{logit}(p_2) + a_4(p_2 - 0.5) + \dots, \\ &\vdots \\ q_m &= a_1 + a_2 \text{logit}(p_m) + a_3(p_m - 0.5) \text{logit}(p_m) + a_4(p_m - 0.5) + \dots.\end{aligned}\tag{6}$$

In the matrix form, this system of equations is equivalent to $q = \mathbb{P}a$, where q and a are column vectors and $\mathbb{P}$ is a $m \times n$ matrix:

$$\mathbb{P} = \begin{bmatrix} 1 & \text{logit}(p_1) & (p_1 - 0.5) \text{logit}(p_1) & (p_1 - 0.5) & \dots \\ 1 & \text{logit}(p_2) & (p_2 - 0.5) \text{logit}(p_2) & (p_2 - 0.5) & \dots \\ & & \vdots & & \\ 1 & \text{logit}(p_m) & (p_m - 0.5) \text{logit}(p_m) & (p_m - 0.5) & \dots \end{bmatrix}\tag{7}$$

If $m = n$ and $\mathbb{P}$ is invertible, then the vector of coefficients a of this *properly parameterized* metalog QPD can be uniquely determined by

$$a = \mathbb{P}^{-1}q\tag{8}$$

If $m > n$ and $\mathbb{P}$ has a rank of at least n , then the vector of coefficients a of the *approximated* metalog QPD, can be estimated using

$$a = [\mathbb{P}^T \mathbb{P}]^{-1} \mathbb{P}^T q\tag{9}$$

The matrix to be inverted in the Equation (9) is always $n \times n$ regardless of the size m of QPT used.

7.4.1 Metalog functions

Metalog **quantile function** (QF) with n terms $Q_{M_n}(u|a)$ can be expressed as

$$Q_{M_n}(u|a) = \begin{cases} a_1 + a_2 \text{logit}(u), & \text{for } n = 2, \\ a_1 + a_2 \text{logit}(u) + a_3(u - 0.5) \text{logit}(u), & \text{for } n = 3, \\ a_1 + a_2 \text{logit}(u) + a_3(u - 0.5) \text{logit}(u) + a_4(u - 0.5), & \text{for } n = 4, \\ Q_{M_{n-1}} + a_n(u - 0.5)^{(n-1)/2}, & \text{for odd } n \geq 5, \\ Q_{M_{n-1}} + a_n(u - 0.5)^{n/2-1} \text{logit}(u), & \text{for even } n \geq 6, \end{cases} \quad (10)$$

where $u \in [0, 1]$ is the cumulative probability and a is the size- n parameter vector of real constants $a = \{a_1 \dots a_n\}$.

The metalog **quantile density function** (QDF) can be found by differentiating the equations (10) with respect to u :

$$q_{M_n}(u|a) = \begin{cases} a_2 \mathcal{J}(u), & \text{for } n = 2, \\ a_2 \mathcal{J}(u) + a_3 ((u - 0.5) \mathcal{J}(u) + \text{logit}(u)), & \text{for } n = 3, \\ a_2 \mathcal{J}(u) + a_3 ((u - 0.5) \mathcal{J}(u) + \text{logit}(u)) + a_4, & \text{for } n = 4, \\ q_{M_{n-1}} + 0.5 a_n (n - 1) (u - 0.5)^{(n-3)/2}, & \text{for odd } n \geq 5, \\ q_{M_{n-1}} + a_n ((u - 0.5)^{n/2-1} \mathcal{J}(u) + (0.5n - 1)(u - 0.5)^{n/2-2} \text{logit}(u)), & \text{for even } n \geq 6, \end{cases} \quad (11)$$

where $\mathcal{J}(u) = [u(1 - u)]^{-1}$. Keelin (2016) showed that the constants a are feasible iff $q_{M_n}(u|a) > 0, \forall u \in [0, 1]$.

Metalog **density quantile function** (DQF), referred to as the “metalog pdf” in Keelin (2016) can be obtained by $f(Q_{M_n}(u|a)) = [q_{M_n}(u|a)]^{-1}$.

Metalog **cumulative distribution function** (CDF) $F_{M_n}(x|a)$ does not have an explicit form because $Q_{M_n}(u|a)$ is not invertible (Keelin, 2016). It is, however, possible to approximate $\widehat{Q}_{M_n}^{-1}(x|a)$ using the method outlined in Perepolkin et al. (2021).

604 Metalog and boundedness

Metalog distribution is defined for all $x \in \mathbb{R}$. Keelin (2016) provides semi-bounded *log-metalog*, and the bounded *logit-metalog* variations of the metalog distribution. As the names suggest, this is achieved through the variable substitution with $z = \ln(x - b_l)$ or $z = -\ln(b_u - x)$ for the semi-bounded case, and $z = \ln((x - b_l)/(b_u - x))$ for the bounded case, where z is metalog-distributed and b_l, b_u are the lower and upper limits, respectively.

Substituting one of the transformations into the QF and QDF functions above, yields semi-bounded or bounded metalog distribution. For the exact formulae of the log-metalog and logit-metalog refer to Keelin (2016).

Appendix B. Weakly informative priors with metalog likelihood

We could also adopt a weakly informative “flat” QDirichlet prior and associate it with the vector of elicited quantile values q^* . Here we use $Dirichlet(1, 1, 1, 1)$ as the prior for the Δ parameter vector of the QDirichlet-metalog model and associate it with the vector of quantile values q^* elicited from the expert in Section 5.2 above.

Figure 13: Prior predictive check for a metalog model with a weakly informative prior

The advantage of the “flat” prior is that we don’t need to ask the expert anything beyond the predictive QPT $\{p^*, q^*\}_n$. The limitation is, of course, that only 36% of the fitted metalogs in the prior predictive check (Figure 13) were valid. This is due to the fact that many simplices Δ drawn from the “flat” Dirichlet distribution, when combined with the quantile values vector q^* , do not produce a valid metalog.

The “flat” prior can also be updated using HMC. We use a vector of quantile values q^* for parameterization and a vector p^* as initial values for the HMC algorithm. The summary of the posterior is presented in Table 7. The posterior predictive check is provided in Figure 14. As we can see the Bayesian inference is also possible with weakly informative priors. The

influence of the “flat” prior on the posterior distribution will vary with the number of data points available for updating.

Table 7: Posterior sample summary for the simplex Δ

category	mean	median	q5	q95	rhat
Small	0.0463	0.0441	0.0236	0.0770	1.000
Medium	0.4252	0.4239	0.3634	0.4895	1.001
Huge	0.0750	0.0725	0.0422	0.1156	1.001
Large	0.4535	0.4540	0.3878	0.5172	1.000

Figure 14: Posterior predictive check for a metalog model with a weakly informative prior

References

- Abbas, A. E., Budescu, D. V., Yu, H.-T., and Haggerty, R. (2008). A comparison of two probability encoding methods: Fixed probability vs. fixed variable values. *Decision Analysis*, 5(4):190–202.
- Akbarov, A. (2009). *Probability Elicitation: Predictive Approach*. PhD thesis, University of Salford.

- 635 Aldeni, M., Lee, C., and Famoye, F. (2017). Families of distributions arising from the quantile
636 of generalized lambda distribution. *Journal of Statistical Distributions and Applications*,
637 4(1):25.
- 638 Balakrishnan, N. (2014). Continuous multivariate distributions. In *Wiley StatsRef: Statistics*
639 *Reference Online*. American Cancer Society.
- 640 Boos, D. D. and Monahan, J. F. (1986). Bootstrap methods using prior information.
641 *Biometrika*, 73(1):77–83.
- 642 Bürkner, P.-C. and Charpentier, E. (2020). Modelling monotonic effects of ordinal predictors
643 in bayesian regression models. *British Journal of Mathematical and Statistical Psychology*,
644 73(3):420–451.
- 645 Coles, S. G. and Tawn, J. A. (1996). A bayesian analysis of extreme rainfall data. *Journal*
646 *of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4):463–478.
- 647 Crowder, M. (1992). Bayesian priors based on a parameter transformation using the distri-
648 bution function. *Annals of the Institute of Statistical Mathematics*, 44(3):405–416.
- 649 Dunson, D. B. and Taylor, J. A. (2005). Approximate bayesian inference for quantiles.
650 *Journal of Nonparametric Statistics*, 17(3):385–400.
- 651 Elfadaly, F. G. and Garthwaite, P. H. (2013). Eliciting dirichlet and connor–mosimann prior
652 distributions for multinomial models. *TEST*, 22(4):628–646.
- 653 Faber, I. and Jung, J. (2021). *Rmetallog: The Metalog Distribution*.
- 654 Freimer, M., Kollia, G., Mudholkar, G. S., and Lin, C. T. (1988). A study of the generalized
655 tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–
656 3567.
- 657 Gabry, J. and Češnovar, R. (2022). *Cmdstanr: R Interface to 'CmdStan'*.
- 658 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization
659 in bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in*
660 *Society)*, 182(2):389–402.
- 661 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013).
662 *Bayesian Data Analysis*. CRC press.
- 663 Gigerenzer, G. (2011). What are natural frequencies? *BMJ*, 343(oct17 2):d6386–d6386.
- 664 Hadlock, C. C. (2017). *Quantile-Parameterized Methods for Quantifying Uncertainty in*
665 *Decision Analysis*. PhD thesis, University of Texas, Austin, TX.

- Hadlock, C. C. and Bickel, J. E. (2017). Johnson quantile-parameterized distributions. *Decision Analysis*, 14(1):35–64.
- Hadlock, C. C. and Bickel, J. E. (2019). The generalized johnson quantile-parameterized distribution system. *Decision Analysis*, 16(1):67–85.
- Hanea, A. M., Hemming, V., and Nane, G. F. (2021). Uncertainty quantification with experts: Present status and research needs. *Risk Analysis*, page risa.13718.
- Hartmann, M., Agiashvili, G., Bürkner, P., and Klami, A. (2020). Flexible prior elicitation via the prior predictive distribution. *arXiv:2002.09868 [stat]*.
- Hoffrage, U., Gigerenzer, G., Krauss, S., and Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, 84(3):343–352.
- Hoffrage, U., Krauss, S., Martignon, L., and Gigerenzer, G. (2015). Natural frequencies improve bayesian reasoning in simple and complex inference tasks. *Frontiers in Psychology*, 6.
- Jeffreys, H. (1939). *The Theory of Probability*. OUP Oxford.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994). *Continuous Univariate Distributions*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, 2nd ed edition.
- Johnson, N. L., Kotz, S., and Balakrishnan, N. (1997). *Discrete Multivariate Distributions*. Wiley Series in Probability and Statistics. Wiley, New York.
- Kadane, J. and Wolfson, L. J. (1998). Experiences in elicitation. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(1):3–19.
- Kadane, J. B. (1980). Predictive and structural methods for eliciting prior distributions. *Bayesian analysis in econometrics and statistics*, 18.
- Keelin, T. W. (2016). The metalog distributions. *Decision Analysis*, 13(4):243–277.
- Keelin, T. W. and Powley, B. W. (2011). Quantile-parameterized distributions. *Decision Analysis*, 8(3):206–219.
- Keeney, R. and von Winterfeldt, D. (1991). Eliciting probabilities from experts in complex technical problems. *IEEE Transactions on Engineering Management*, 38(3):191–201.
- King, R. A. R. (1999). *New Distributional Fitting Methods Applied to the Generalised [Lambda] Distribution*. PhD thesis, Queensland University of Technology, Australia.
- Knight, F. H. (1921). *Risk, Uncertainty and Profit*, volume 31. Houghton Mifflin.
- Lavine, M. (1995). On an approximate likelihood for quantiles. *Biometrika*, 82(1):220–222.

697 Mikkola, P., Martin, O. A., Chandramouli, S., Hartmann, M., Pla, O. A., Thomas, O.,
698 Pesonen, H., Corander, J., Vehtari, A., Kaski, S., Bürkner, P.-C., and Klami, A. (2021).
699 Prior knowledge elicitation: The past, present, and future. *arXiv:2112.01380 [stat]*.

700 Morgan, M. G. (2014). Use (and abuse) of expert elicitation in support of decision making
701 for public policy. *Proceedings of the National Academy of Sciences*, 111(20):7176–7184.

702 Morgan, M. G., Henrion, M., and Small, M. (1990). *Uncertainty: A Guide to Dealing with*
703 *Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge university press.

704 Myerson, R. B. (2005). *Probability Models for Economic Decisions*. Duxbury Applied Series.
705 Thomson/Brooke/Cole, Belmont, CA.

706 Nair, N. U., Sankaran, P. G., and Dileepkumar, M. (2020). Bayesian inference in quantile
707 functions. *Communications in Statistics - Theory and Methods*, 0(0):1–13.

708 Nair, N. U., Sankaran, P. G., and Vineshkumar, B. (2012). The govindarajulu distribution:
709 Some properties and applications. *Communications in Statistics - Theory and Methods*.

710 O’Hagan, A. (2019). Expert knowledge elicitation: Subjective but scientific. *The American*
711 *Statistician*, 73(sup1):69–81.

712 Patel, J. K. and Read, C. B. (1996). *Handbook of the Normal Distribution*, volume 150. CRC
713 Press.

714 Perepolkin, D. (2019). *Qpd: Tools for Quantile-Parameterized Distributions*.

715 Perepolkin, D., Goodrich, B., and Sahlin, U. (2021). The tenets of indirect inference in
716 bayesian models. Preprint, Open Science Framework, <https://osf.io/enzgs>.

717 Powley, B. W. (2013). *Quantile Function Methods for Decision Analysis*. Stanford University.

718 Pratt, J. W., Raiffa, H., and Schlaifer, R. (1995). *Introduction to Statistical Decision Theory*.
719 MIT press.

720 Rayner, G. D. and MacGillivray, H. L. (2002). Numerical maximum likelihood estimation for
721 the g-and-k and generalized g-and-h distributions. *Statistics and Computing*, 12(1):57–75.

722 Spetzler, C. S. and Staël Von Holstein, C.-A. S. (1975). Probability encoding in decision
723 analysis. *Management Science*, 22(3):340–358.

724 Vega Yon, G. and Marjoram, P. (2019). Fmcmc: A friendly mcmc framework. *The Journal*
725 *of Open Source Software*, 4(39).

726 Vihola, M. (2012). Robust adaptive metropolis algorithm with coerced acceptance rate.
727 *Statistics and Computing*, 22(5):997–1008.

- 728 Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer Texts in Statistics. Springer,
729 New York.
- 730 Wilson, E. (2017). Fitting a modified connor-mosimann distribution to elicited quantiles of
731 multinomial probabilities.
- 732 Winkler, R. L. (1967). The assessment of prior distributions in bayesian analysis. *Journal*
733 *of the American Statistical Association*, 62(319):776–800.
- 734 Winkler, R. L. (1980). Prior information, predictive distributions, and bayesian model-
735 building. *Bayesian Analysis in Econometrics and Statistics*. North-Holland Publishing
736 Company, pages 95–109.