The tenets of quantile-based inference in Bayesian models

Dmytro Perepolkin* Benjamin Goodrich† Ullrika Sahlin‡

First version: 2021-06-29

This version: 2022-08-23

Abstract

This paper extends the application of Bayesian inference to probability distributions defined in terms of their quantile functions. We describe the method of *quantile-based likelihood* to be used in the Bayesian models with sampling distributions which lack an explicit probability density function. We provide examples applying the principles of *quantile-based inference* in univariate settings, as well as in *parametric quantile regression* for describing and updating the parameters of the error term. We also extend the same method to propose the definition of an *quantile-based prior* and discuss the situations where it can be useful.

You have to start looking at the world in a new way.

– Priya in TENET by Christopher Nolan

1 Introduction

Most statistics courses and textbooks introduce continuous random variables via the (cumulative) *distribution function* (CDF) and the *probability density function* (PDF). “An equally adequate representation” (Tukey, 1965) of a random variable can be made using the *inverse CDF*, known as the *quantile function* (QF), and its derivative, the *quantile density function* (QDF), but the use of such *quantile distributions* is rare. Defining a distribution via its quantile function has several advantages, including that the distributions with explicit quantile functions are easy to sample from and more complex distributions can be crafted using the simpler quantile functions as the building blocks (Gilchrist, 2000; Parzen, 1980; Hadlock, 2017; Powley, 2013).

Some of the widely used probability distributions defined in terms of the CDF and PDF (*density-defined* distributions) are not easily invertible (e.g. normal or gamma) and, therefore,

*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

the numerical approximation of their QF is used. Similarly, there are other distributions defined in terms of their QF and QDF (*quantile* distributions), that are also not invertible, and thus, the numerical approximation of their CDF can be used.

Most of the knowledge and methods for Bayesian inference have been developed for the *density-defined* distributions. While there have been several published articles where *quantile* distributions were used in the context of the likelihood-free approximate Bayesian computation (Allingham et al., 2009; Drovandi and Pettitt, 2011; Karabatsos and Leisen, 2018; Fearnhead and Prangle, 2012; Bernton et al., 2019; McVinish, 2012), the likelihood-based application of the Bayesian inference for *quantile* distributions has been limited (Rayner and MacGillivray, 2002; Haynes and Mengersen, 2005; Nair et al., 2020).

1.1 Aims of the paper

This article builds on the ideas of Gilchrist (2000), Rayner and MacGillivray (2002), Nair et al. (2020) and systematically presents and illustrates the Bayesian inference using quantile functions. We refer to this method of inference as *quantile-based* because it deals with the inverse transformation of the *intermediate cumulative probabilities* (*depths*¹) corresponding to the observations given the parametrized model. We aim to show that the *quantile-based* Bayesian inference using the intermediate depths leads to the same posterior beliefs as the conventional density-based inference. We apply the principles of *quantile-based inference* to Bayesian updating of parameters in the univariate and regression settings using the flexible and extensible quantile sampling distributions.

1.2 Paper structure

Section 2 revisits the functions and identities for characterizing the distribution of a continuous random variable. Then in Section 3 we introduce the terms of *quantile-based likelihood* and *quantile-based prior* in Bayesian inference and show that the likelihood (and prior) can be expressed without the PDF. Section 4 discusses the computational aspects of the numerical inversion of quantile functions for approximating the intermediate depths in a quantile-based likelihood expressed by a quantile distribution. Section 5 discusses the applications of quantile-based inference in univariate and regression settings. We discuss the models and provide code examples implementing quantile-based likelihood in Stan (Gabry and Češnovar, 2022) and in R (R Core Team, 2021). We conclude the paper with a discussion and summary of the results in Section 6.

¹Because they indicate how “deep” an observation is into the distribution.

Figure 1: Probability distributions, quantile distributions and parameterization.

1.3 Paper contribution

Although the description of the *quantile-based likelihood* (Rayner and MacGillivray, 2002; King, 1999; Gilchrist, 2007; Nair et al., 2020) and prior (Nair et al., 2020) appeared in the literature before, they were presented in the context of specific distributions and not as a general principle of inference. In their recent work, Nair et al. (2020) presented Bayesian inference with quantile functions, but their presentation of what we describe here as *quantile-based prior* lacked the necessary adjustment due to the nonlinear transformation of the parameters involved (see Section 3.2 below).

In this paper, we apply the principles of *quantile-based inference* to implement the Bayesian version of the *parametric quantile regression* (Gilchrist, 2008) with the error term is described by a bespoke quantile function and estimate the regression parameters using MCMC.

2 Distribution specification

To set the scene for the discussion of density-based and quantile-based Bayesian inference we briefly review the different ways of specifying a probability distribution and discuss several examples of the distributions defined by a quantile function, found in the literature (Figure 1).

2.1 Essential functions

Let Y be a continuous random variable. It can be expressed via the (cumulative) *distribution function* (CDF)

$$F_Y(y|\theta) = \Pr(Y \leq y|\theta), \quad \theta \in \mathcal{A} \subset \mathbb{R} \quad (1)$$

An alternative way of describing the random variable Y is via the *quantile function* (QF)

$$Q_Y(u|\theta) = \inf\{y : F_Y(y|\theta) \geq u\}, \quad 0 \leq u \leq 1 \quad (2)$$

The subscript $_Y$ is used to indicate the random variable that the depth u corresponds to.

New quantile functions can be easily created using Gilchrist's quantile function transformation rules (Gilchrist, 2000; Powley, 2013; Hadlock, 2017; Sharma and Chakrabarty, 2017) summarized in Table 1. We use these rules for crafting a bespoke quantile function for modeling the error term in Bayesian parametric quantile regression in Section 5.

If $F_Y(y|\theta)$ is continuous and strictly monotonically increasing over the support of Y , then $Q_Y(u|\theta)$ is simply the inverse of $F_Y(y|\theta)$. Therefore, the quantile function is often referred to as the *inverse CDF*, i.e.

$$Q_Y(u|\theta) = F_Y^{-1}(y|\theta) \quad (3)$$

Not all QFs are analytically invertible. A distribution whose quantile function $Q_Y(u|\theta)$ is not analytically invertible is called a *quantile distribution* (Gilchrist, 1997) or a *quantile-based distribution* (Sharma and Chakrabarty, 2020) (Figure 1).

The derivative of the CDF is the *probability density function* (PDF) denoted by

$$f_Y(y|\theta) = \frac{dF_Y}{dy} \quad (4)$$

Similarly, the derivative of the QF is the *quantile density function* (QDF) denoted by

$$q_Y(u|\theta) = \frac{dQ_Y}{du}, \quad 0 \leq u \leq 1 \quad (5)$$

The reciprocal of the QDF $[q_Y(u|\theta)]^{-1} = f(Q_Y(u|\theta))$ is referred to as the *density quantile function* (Parzen, 1980) or *p-pdf* (Gilchrist, 2000). Here and for the rest of the article, we will often omit the subscript $_Y$ and the conditioning on θ to simplify the notation.

Table 1: Gilchrist’s quantile function transformation rules

Original QF	Rule	Resulting QF	Resulting variable
$Q_Y(u)$	Reflection rule	$-Q(1 - u)$	QF of $-Y$
$Q_Y(u)$	Reciprocal rule	$1/Q(1 - u)$	QF of $1/Y$
$Q_1(u), Q_2(u)$	Addition rule	$Q_1(u) + Q_2(u)$	valid QF
$Q_1(u), Q_2(u)$	Linear combination rule	$aQ_1(u) + bQ_2(u)$	valid QF for $a, b > 0$
$Q_1(u), Q_2(u) > 0$	Multiplication rule	$Q_1(u)Q_2(u)$	valid QF
$Q_Y(u)$	Q-transformation	$T(Q_Y(u))$	QF of $T(Y)$, $T(Y)$ non-decreasing
$Q_Y(u)$	p-transformation	$Q_Y(H(u))$	p-transformation of $Q_Y(u)$, $H(u)$ non-decreasing

$$f(Q(u)) = \frac{dF(Q(u))}{dQ(u)} = \frac{dF(Q(u))/du}{dQ(u)/du} = \frac{dF(F^{-1}(u))/du}{q(u)} = \frac{du/du}{q(u)} = [q(u)]^{-1} \quad (6)$$

95 In Section 3 of this paper, we rely on the density quantile function (DQF) $[q(u)]^{-1}$, i.e. the
96 density of a random variable expressed in terms of the cumulative distribution function
97 (Perri and Tarsitano, 2007), to define the likelihood in a Bayesian model based on a quantile
98 sampling distribution.

99 2.2 Derivatives of the inverses and the numerical approximation

100 Following the inverse function theorem (Price, 1984), for a function to be invertible in the
101 neighborhood of a point, it should have a continuous non-zero derivative at that point. If the
102 function is invertible, the derivative of the inverse is reciprocal to the function’s derivative
103 (Marsden et al., 1985). Formally, if dt/dy exists and $dt/dy \neq 0$, then dy/dt also exists and
104 $dy/dt = [dt/dy]^{-1}$. Therefore, for a quantile function $Q(u) = y$, if a QDF $q(u)$ exists and
105 $q(u) \neq 0$, then PDF $f(y)$ also exists and it is equal to $f(y) = [q(u)]^{-1}$.

106 Figure 2 illustrates the relationship between the key probability functions. The distribution
107 function (CDF) and the quantile function (QF) are depicted on the opposite sides of the
108 Moebius strip. The derivatives of these functions (PDF and QDF, respectively) end up on
109 the same side, due to the geometry of the Moebius surface. It means that the derivatives are
110 no longer the inverses of each other, but rather the reciprocals, as stated in the equation at
111 the bottom. The “do-it-yourself” copy of this probability function Moebius strip is included
112 in the Supplementary Materials.

Figure 2: Moebius strip of probability functions.

Even though the quantile distributions lack the closed-form CDF $F(y) = u$, in most cases, the depths u can be approximated by numerically inverting the $Q(u)$. We denote the numerically inverted quantile function as $\widehat{Q}^{-1}(y)$ or $\widehat{F}(y)$. The inverse of a quantile function $Q(u)$ at point u , corresponding to the observation y , is obtained by minimizing the difference between the actual observation y and $Q(u)$ by iteratively refining the depth u . The details of the numerical inversion algorithm are discussed in Section 4.

2.3 Quantile distributions

Statistical methods utilizing QF and QDF were pioneered by the seminal work of Parzen (1979). Some of the quantile distributions covered in the literature are generalized g -and- h and its sibling g -and- k distribution (Haynes and Mengersen, 2005; Jacob, 2017; Prangle, 2017; Rayner and MacGillivray, 2002), Tukey Lambda Distribution and its generalizations, known as GLD (Aldeni et al., 2017; Chalabi et al., 2012; Dedduwakumara et al., 2021; Ramberg and Schmeiser, 1974; Fournier et al., 2007; Freimer et al., 1988), Wakeby distribution (Rahman et al., 2015), flattened logistic distribution (Sharma and Chakrabarty, 2019) and Govindarajulu distribution (Nair et al., 2012, 2013).

The mathematical notation for describing probability distributions has been standardized and adopted across different domains. First use of the tilde symbol $\sim$ to denote the CDF can be traced back to early 1960s: “ $X \sim F(x)$ means that x is distributed according to the distribution function $F(x)$ ” (Olkin and Tate, 1961). Today this convention is adopted by the majority of the Bayesian textbooks (Gelman et al., 2013; Johnson et al., 2022; O’Hagan et al., 2006; Lambert, 2018; Gelman et al., 2021; Koller and Friedman, 2009). For example, if a variable Y is normally distributed it is described as $Y \sim N(\mu, \sigma)$, which means that the

random variable Y has the distribution function $F_Y(y) = \Phi(y|\mu, \sigma)$, where Φ is the CDF of the normal distribution (Johnson et al., 1994).

If the distribution of a random variable Y is described by a non-invertible quantile function, such as, for example, the standard quantile function of the Skew-Logistic Distribution (SLD)² $Q_Y(u|\delta) = (1 - \delta) \ln(u) - \delta \ln(1 - u)$ stating $Y \sim \text{SLD}(\delta)$ is not strictly correct, because the SLD quantile function is not invertible and its CDF can be computed only numerically as $F_Y(y|\delta) \cong \widehat{Q_Y^{-1}}(y|\delta)$. Therefore, in this paper, we propose to denote this distribution as $u \overset{y}{\sim} \text{SLD}(\delta)$, where the *back-tilde symbol* with the variable name overscript $\overset{y}{\sim}$ should be read “inversely distributed as” to indicate that the *depth* u is fully determined given the value of the random variable Y and the parameterized inverse distribution function indicated to the right of the *back-tilde symbol*, in this case SLD³.

Although in this paper we focus on the distributions with abstract parameters, the distributions parameterized by the quantile-probability pairs (*quantile-parameterized* distributions) are also worth a mention. The most prominent examples of the *quantile-parameterized density-defined* distributions are the Myerson distribution (Myerson, 2005), Johnson QPD (J-QPD) with its generalizations (Hadlock and Bickel, 2017, 2019) and the Simple Q-Normal distribution (Keelin and Powley, 2011). The group of the *quantile-parameterized quantile distributions* is represented by the metalog distribution (Keelin, 2016). Quantile-parameterized distributions play an important role in representing expert beliefs about variables, parameters, or quantities of interest (Dion et al., 2020; Gu et al., 2018; Larrain et al., 2021; Reinhardt et al., 2016; Baey et al., 2022), although they don’t lend themselves easily as sampling distributions due to the special nature of their parameterization.

3 Bayesian inference with quantile functions

Gilchrist (2000), p. 209 used the term *quantile-based likelihood* when describing the principles of maximum likelihood applied to quantile distributions. Rayner and MacGillivray (2002) describe a three-step process of computing the log-likelihood for a quantile distribution and use it for the maximum likelihood estimation of parameters in *g-and-k* and generalized *g-and-h* distributions. Nair et al. (2020) performed quantile function substitutions both for observables $y_i = Q(u_i|\theta)$, $i = \{1, 2, \dots, n\}$ and for parameters $\theta = Q_\Theta(v)$ and computed

²The standard logistic quantile function $Q(u) = \ln(u) - \ln(1 - u)$ is a combination of the exponential $-\ln(1 - u)$ and the reflected exponential $\ln(u)$ quantile functions (Gilchrist, 2000; Lampasi, 2008). The skew-logistic quantile function introduces a weight parameter δ to scale the exponential components of the logistic function as $(1 - \delta) \ln(u) - \delta \ln(1 - u)$ (Gilchrist, 2000; van Staden and King, 2015; Balakrishnan and So, 2015) using Linear Combination Rule in Table 1.

³In situations where extra clarity is required, the *depth* variable name can also be mentioned in describing the density-defined distributions, e.g. $X \overset{p}{\sim} N(\mu, \sigma)$, where $\mu \overset{v}{\sim} N(\mu_0, \sigma_0)$

the Bayes estimator under the squared error loss for the Govindarajulu likelihood. In this section, we try to summarize this approach and introduce the terms *quantile-based* prior and *quantile-based* likelihood based on the identities and substitutions introduced in Section 2 and show the equivalence of the two ways of expressing the likelihood in Bayesian models.

3.1 Density-based and quantile-based likelihood

The traditional Bayesian inference formula can be restated using the substitutions involving quantile functions (Nair et al., 2020). Assume that the prior information about the scalar parameter θ can be summarized by the prior distribution over the parameter space Θ . Then, given a random sample of $\underline{y} = \{y_1, y_2, \dots, y_n\}$, the posterior distribution of θ can be expressed as:

$$f(\theta|\underline{y}) \propto \mathcal{L}(\theta; \underline{y})f(\theta) \quad (7)$$

where $f(\theta|\underline{y})$ is the posterior distribution of θ after having observed the sample $\underline{y}$, $f(\theta)$ is the prior distribution of θ , and $\mathcal{L}(\theta; \underline{y}) = \prod_{i=1}^n f(y_i|\theta)$ is the likelihood, which is a function of θ . We refer to this form of likelihood as *density-based*, because it is expressed using the density function (PDF) of the observable $\underline{y}$.

Given the random sample $\underline{y}$, we can use the quantile function $Q_Y(u|\theta)$ to compute $\underline{Q} = \{Q_1(u_1), \dots, Q_n(u_n)|\theta\}$, such that $u_i = F(y_i|\theta)$, $i = \{1, 2, \dots, n\}$. The depths u_i are degenerate random variables because they are fully determined given the values of $\underline{y}$ and the parameter θ . Since $Q_Y(u_i|\theta) = y_i$ we can substitute $\underline{Q}$ for $\underline{y}$. Then the Bayesian inference formula (7) becomes

$$f(\theta|\underline{Q}) \propto \mathcal{L}(\theta; \underline{Q})f(\theta) \quad (8)$$

We refer to the likelihood $\mathcal{L}(\theta; \underline{Q}) = \prod_{i=1}^n f(Q(u_i|\theta)) = \prod_{i=1}^n [q(u_i|\theta)]^{-1}$ as *quantile-based*, because it relies on computing the intermediate depths $u_i = F(y_i|\theta)$ corresponding to the observations y_i , $i = \{1, 2, \dots, n\}$. The two forms of likelihood $\mathcal{L}(\theta; \underline{Q})$ and $\mathcal{L}(\theta; \underline{y})$ are equivalent to each other. Therefore, following the likelihood principle, the posterior beliefs about θ in both cases are identical.

Since the likelihood in the Equation (8) is expressed in terms of $\underline{Q}$, an additional transformation is required to arrive at $\underline{u} = F(\underline{y}|\theta)$. In case the CDF $F(\underline{y}|\theta)$ is not available, the numeric approximation of $\widehat{Q}^{-1}(\underline{y}|\theta)$ may be used. We discuss the details of the numerical approximation of the inverse quantile function in Section 4 of this paper.

3.2 Density-based and quantile-based prior

It is possible to extend the same logic of quantile function substitution to define the *density-based* and the *quantile-based* prior. In this section, we discuss the parameter transformation required for defining a quantile-based prior and show its connection to the inverse transform used for non-uniform sampling.

The Bayesian inference formula can be restated using the quantile form of the prior (Nair et al., 2020). Assume that the prior distribution of θ can be described using the invertible CDF $F_{\Theta}(\theta|\psi) = v$ with hyperparameter ψ , so that $Q_{\Theta}(v|\psi) = \theta$. Substituting the quantile values $Q_{\Theta}(v|\psi)$ for values of θ , prior beliefs about the parameter(s) of the distribution of θ can be expressed using the distribution of the *quantile values* corresponding to the random variate v , given hyperparameter ψ of the prior distribution as $f(Q_{\Theta}(v|\psi)) = [q_{\Theta}(v|\psi)]^{-1}$. We refer to the such formulation of the prior as *quantile-based* because it describes the prior distribution of the random variate v given hyperparameter ψ corresponding to the parameter $\theta = Q_{\Theta}(v|\psi)$ and not the distribution of the parameter θ itself.

Likewise, the likelihood $\mathcal{L}(Q_{\Theta}(v|\psi); \underline{y})$ will also rely on the parameter transformation $\theta = Q_{\Theta}(v|\psi)$. However, such non-linear parameter transformation requires a Jacobian adjustment (Andrilli and Hecker, 2010), which is equal to the absolute derivative of the transform, i.e. $J(Q_{\Theta}(v|\psi)) = |dQ_{\Theta}(v|\psi)/dv| = |q_{\Theta}(v|\psi)|$. Provided that the $Q_{\Theta}(v)$ is a valid (non-decreasing) quantile function, meaning that $q_{\Theta}(v|\psi)$ is non-negative on $v \in [0, 1]$, the density quantile term $[q_{\Theta}(v|\psi)]^{-1}$ representing the prior and the Jacobian adjustment $|q_{\Theta}(v|\psi)|$ can be dropped as they are reciprocal to each other. Therefore the indirect posterior of the random variate v after observing the sample $\underline{y}$ can be expressed as

$$\begin{aligned} [q_{\Theta}(v|\underline{y})]^{-1} &\propto \mathcal{L}(Q_{\Theta}(v|\psi); \underline{y}) [q_{\Theta}(v|\psi)]^{-1} |q_{\Theta}(v|\psi)| \implies \\ [q_{\Theta}(v|\underline{y})]^{-1} &\propto \mathcal{L}(Q_{\Theta}(v|\psi); \underline{y}) \end{aligned} \tag{9}$$

where $[q_{\Theta}(v|\underline{y})]^{-1}$ is the quantile form of the posterior, and the (quantile) prior density $[q_{\Theta}(v|\psi)]^{-1}$ is implied by QF transform $Q_{\Theta}(v|\psi) = \theta$, $v \in [0, 1]$.

The quantile-based prior can also be used in combination with a quantile-based likelihood, since, as we showed previously, regardless of the form of the likelihood used, the posterior beliefs about the parameter θ will be the same. In such a case, neither prior nor likelihood would require the PDF, and, therefore, both of them can be represented by quantile distributions.

4 Numerical inversion of quantile functions

The core element of the *quantile-based* likelihood method is the use of the intermediate depths $\underline{u}$, corresponding to the observables $\underline{y}$ given the parameter θ . These values can either be found analytically, as $F(y)$ for distributions with CDF, or numerically via root-finding algorithm, as $\widehat{F}_Y(y) = \widehat{Q}_Y^{-1}(y)$ e.g. in case of quantile distributions (Figure 1).

The problem of inverting a quantile function is tantamount to finding the root of a target function

$$\Omega(u; y, \theta) = [y - Q_Y(u|\theta)] \quad (10)$$

where y is a known observation, θ is the parameter value, and u is the depth. Provided that the $Q_Y(u|\theta)$ is a non-decreasing function and y is a fixed observable value, the target function $\Omega(u; y, \theta)$ is non-increasing. The root-finding algorithm uses the target function to take an observable y and “pull in” its inverted equivalent $Q_Y(u|\theta)$ until the two values exactly meet by iteratively adjusting u . Since the target function $\Omega(u; y, \theta)$ has a range $u \in [0, 1]$, the *bracketing* root-finding algorithms, such as bisection or *regula falsi* (Atkinson, 2008; Burden and Faires, 2011) may be used, although depending on the shape of the quantile function their convergence can be slow. The convergence may be accelerated with the help of the *non-bracketing* root-finding algorithms e.g. Newton-Raphson method (Householder, 1970), which rely on computing the derivatives of the target function Ω . The first derivative of the target function is simply the negative QDF $\Omega'(u; y, \theta) = \frac{d[y - Q_Y(u|\theta)]}{du} = -q_Y(u|\theta)$. Unfortunately, the derivative-based algorithms do not guarantee that the root will be found and may end up in infinite loops and divergences. Modern root-finding algorithms such as TOMS748 (Alefeld et al., 1995), Chandrupatla (Chandrupatla, 1997), NewtSafe (Acton, 1990; Press, 2007), Ridders’ (Ridders, 1979) and Zhang (Zhang, 2011; Stage, 2013), combine advantages of bracketing and non-bracketing methods to ensure robust and efficient convergence. Many of these algorithms have been implemented in high-performance math libraries, such as Boost (Schäling, 2011).

5 Applications

In this section, we illustrate the application of the *quantile-based inference* to univariate and regression models and provide code examples for models based on the quantile sampling distributions in Stan (Gabry and Češnovar, 2022) and in R (R Core Team, 2021). For the univariate model, we update the parameters of a bathtub-shaped Govindarajulu distribution and for the regression model, we pick the flattened skew-logistic distribution to model the

error term.

5.1 Univariate model

We take the dataset provided in Aarset (1987) on the time-to-failure of 50 devices (Figure 3). Lifetime reliability data are often modeled using specialized distributions (Nadarajah, 2009) or 2(3)-component mixtures. Nair et al. (2020) obtained the Bayes estimator under the squared error loss function for the posterior mean of the parameter γ in the Govindarajulu distribution (Nair et al., 2012), under the generalized exponential prior (Gupta and Kundu, 2007). We extend their approach to estimating the full posterior distribution of both parameters in the Govindarajulu likelihood implementing the model in Stan (Gabry and Češnovar, 2022).

The QF and the DQF of the Govindarajulu distribution (Nair et al., 2012) are given by:

$$\begin{aligned} Q(u|\sigma, \gamma) &= \sigma\gamma u^\gamma(1 + \gamma^{-1} - u) \\ [q(u|\sigma, \gamma)]^{-1} &= [\sigma\gamma(\gamma + 1)u^{\gamma-1}(1 - u)]^{-1} \end{aligned} \quad (11)$$

where $\gamma > 0$. The distribution has support on $Q(u|\sigma, \gamma) \in [0, \sigma]$. For a definition of the Govindarajulu distribution with shifted support see Nair et al. (2012).

Figure 3: Histogram of time-to-failure for 50 devices

We adopt the generalized exponential prior (Gupta and Kundu, 2007) for the parameter γ of the Govindarajulu distribution with hyperparameters $\alpha = 5$ and $\lambda = 1$. The CDF and PDF of the generalized exponential distribution are given by

Table 2: Summary of the posterior samples from the GenExp-Govindarajulu model

parameter	mean	median	q5	q95	rhat
gamma	2.0355	2.0067	1.5749	2.5978	1.001
sigma	86.0225	86.0004	86.0000	86.1240	1.001

$$\begin{aligned}
 F(x|\lambda, \alpha) &= (1 - e^{-\lambda x})^\alpha \\
 f(x|\lambda, \alpha) &= \alpha \lambda (1 - e^{-\lambda x})^{\alpha-1} e^{-\lambda x}
 \end{aligned}
 \tag{12}$$

where $\alpha, \lambda > 0$. The support of the distribution is $x \in [0, \infty]$.

The parameter σ of the Govindarajulu distribution can not be lower than the maximum of the observed times-to-failure. Therefore, we defined a shifted exponential prior with PDF $f_X(x|\lambda) = \lambda e^{-\lambda(x-l)}$ with the rate $\lambda = 0.5$ and the varying lower bound l corresponding to the highest time to failure (in our example equal to 86). The Jacobian for adding a constant to the sampled parameter σ is equal to one, so there is no impact on the log density, as the shifting transform produces a Jacobian derivative matrix with a unit determinant.

The resulting model takes the following form

$$\begin{aligned}
 u &\overset{y}{\sim} \text{Govindarajulu}(\sigma, \gamma) \\
 \gamma &\sim \text{GenExp}(5, 1) \\
 \sigma - \max(y) &\sim \text{Exp}(1)
 \end{aligned}
 \tag{13}$$

We ran 2500 post-warmup iterations and 4 chains in Stan (Gabry and Češnovar, 2022). Table 2 summarizes the posterior distribution of the parameters γ and σ of Govindarajulu distribution. The code is provided in the Supplementary Materials.

5.2 Parametric Quantile Regression

Quantile functions are useful not only for modeling the observables, but they can also be used to represent unobserved quantities of interest, such as the error term in a parametric quantile regression.

Using Gilchrist’s Linear Combination rule in Table 1 any quantile function can be represented as

$$Q(u|\mu, \sigma, \theta) = \mu + \sigma Q_s(u|\theta) \tag{14}$$

where $Q_s(u)$ is a “basic” quantile function (Gilchrist, 2000), μ and σ are location and scale parameters, respectively, and θ is an optional shape parameter. Many quantile functions, such as logistic $Q(u) = \mu + \sigma \text{logit}(u)$ or normal $Q(u) = \mu + \sigma \Phi^{-1}(u)$, are already in this form. Others, such as the SLD discussed in Section 2, can have location and scale parameters added to them to enable shifted and scaled support, e.g. $Q(u|\mu, \sigma, \delta) = \mu + \sigma [(1 - \delta) \ln(u) - \delta \ln(1 - u)]$. The basic quantile functions (i.e. $\text{logit}(u)$ for the logistic and $\Phi^{-1}(u)$ for the normal) can be useful as the building blocks for constructing more complex distributions (Table 1). Basic quantile functions with the median centered at zero are called *standard* quantile functions (Gilchrist, 2000) denoted here as $S(u|\theta) = Q_s(u|\theta) - Q_s(0.5|\theta)$.

A simple linear regression of a random variable Y given the covariate X can be written as

$$y_i = \underbrace{\alpha + \beta x_i}_{\text{deterministic term}} + \underbrace{\varepsilon_i}_{\text{stochastic term}} \quad (15)$$

where y_i is the i -th observation of Y , x_i is i -th observation of covariate X , α and β are unknown intercept and slope, respectively. The ε_i represents the error, which in ordinary least squares (OLS) regression is assumed (or forced through the link function) to be normally distributed with the mean of zero. An alternative way of representing the error term is through a *standard quantile function* $\varepsilon_i = S_\varepsilon(u_i|\theta)$, where u_i is the depth corresponding to the error ε_i in the regression model with the intercept α , slope β and the shape parameter θ (which are assumed to be independent).

$$y_i = \alpha + \beta x_i + S_\varepsilon(u_i|\theta) \quad (16)$$

We want to emphasize that the *traditional* quantile regression introduced by Koenker (2005) is in essence semi-parametric, because it does not require the user “to specify the distribution of the error term as it is allowed to take any form” (Yu and Moyeed, 2001). The regression Equation (16) represents the *parametric quantile regression* (PQR), because in this type of regression the error term is modeled explicitly (Gilchrist, 2008; Sharma and Chakrabarty, 2020; Su, 2015; Dean and King, 2009; Muraleedharan et al., 2016; Perri and Tarsitano, 2007, 2008).

Note that the deterministic term in (16) can be viewed as a location parameter in the quantile function

$$Q_Y(u_i|\mu_i, \theta) = \mu_i + S_\varepsilon(u_i|\theta) \quad (17)$$

where $\mu_i = \alpha + \beta x_i$. Likewise, if the stochastic component $S_\varepsilon(u_i|\theta)$ is made dependent on the covariate x_i , the resulting PQR QF can capture the heteroscedasticity of the error term.

The depth u_i can be found by inverting the quantile function $u_i = \widehat{Q_Y^{-1}}(y_i|\mu_i, \theta)$. In cases where inverting the PQR QF may be analytically difficult (e.g. when the $S_\varepsilon(u_i, \theta)$ is not invertible), the numerical approximation can be used (see Section 4 above). Once the depths $\underline{u} = \{u_1, u_2, \dots, u_n\}$ are found, the likelihood of N observations $\underline{y} = \{y_1, y_2, \dots, y_n\}$ given parameter θ can be calculated using the density quantile function corresponding to the PQR QF.

Because the deterministic term μ_i in PQR QF $Q_Y(u_i|\mu_i, \theta)$ is additive and does not depend on the depth u_i it can be dropped from the derivative.

$$[q_Y(u_i|\mu_i, \theta)]^{-1} = \left[\frac{dQ_Y(u_i|\mu_i, \theta)}{du} \right]^{-1} = \left[\frac{dS_\varepsilon(u_i|\theta)}{du} \right]^{-1} = [q_\varepsilon(u_i|\theta)]^{-1} \quad (18)$$

where $[q_\varepsilon(u_i|\theta)]^{-1}$ is the *density quantile function* of the error term.

We illustrate the application of PQR using the car stopping distance data from [Gilchrist \(2000\)](#), sec. 12.4. The dataset (Figure 4) contains 30 observations of the car speed and the corresponding stopping distances. As suggested by the physics' kinetic energy equation ([Lutus, 2021](#)) the speed of the car is proportional to the square root of the braking distance. We can draw a *mean regression line* through the observations, as shown in Figure 4 relating the car to the square root of the stopping distance using ordinary least squares (OLS). In the rest of this section we will estimate the *quantile regression lines* for the *median*, the 5th, and the 95th quantile using the PQR.

Figure 4: Car stopping distance dataset

One of the simplest quantile functions which could be used to model the error in PQR is the logistic quantile function $Q(u) = \ln(u) - \ln(1 - u)$. The distribution of the errors in the stopping distance model might be less “peaked” than the standard logistic distribution due to various factors (about vehicles or the drivers) not included in the model. Therefore, adding some effect of the standard uniform quantile function u (Lampasi, 2008) might be reasonable. Flattened Logistic Distribution (FLD) described by Sharma and Chakrabarty (2019) combines the standard QFs of logistic and uniform distributions by applying positive affine transformation for scale and shape parameters (ref. Addition and Linear Combination rules in Table 1).

$$Q_\varepsilon(u|\chi, \eta, \kappa) = \chi + \eta \left[\underbrace{\ln(u) - \ln(1 - u)}_{\text{logistic}} + \kappa \times \underbrace{u}_{\text{uniform}} \right] \quad (19)$$

where χ is the location parameter, $\eta, \kappa > 0$ are scale and shape parameters, respectively. For the standard quantile function $S_\varepsilon(u|\kappa) = Q_\varepsilon(u|\kappa) - Q_\varepsilon(0.5|\kappa)$, the location should be set to 0 and the scale set to 1.

The Flattened Logistic Distribution is symmetrical. This assumption might be too restrictive for modeling the residuals in the car stopping distance model (e.g. because of inertia). Sharma and Chakrabarty (2020) replaced the logistic quantile function in the FLD with the skew-logistic quantile function; the resulting QF can be referred to as the flattened skew-logistic distribution (FSLD).

The FSLD QF and the DQF are

$$Q_\varepsilon(u|\chi, \eta, \delta, \kappa) = \chi + \eta \left(\underbrace{(1 - \delta) \ln(u) - \delta \ln(1 - u)}_{\text{skew-logistic}} + \kappa \times \underbrace{u}_{\text{uniform}} \right) \quad (20)$$

$$[q_\varepsilon(u|\chi, \eta, \delta, \kappa)]^{-1} = \left[\eta \left(\frac{1 - \delta}{u} + \frac{\delta}{1 - u} + \kappa \right) \right]^{-1}$$

Since the variance in the speed Y increases with the car stopping distance X , a heteroscedastic model can be used to describe the error term in the PQR for the stopping distances. The resulting PQR QF and the corresponding DQF can be expressed as

$$Q_Y(u|\alpha, \beta, \theta; x) = \alpha + \beta\sqrt{x} + S_\varepsilon(u; \theta)\sqrt{x}$$

$$[q_Y(u|\theta; x)]^{-1} = \left[\frac{dQ_Y(u|\alpha, \beta, \theta; x)}{du} \right]^{-1} = \frac{1}{\sqrt{x}} [q_\varepsilon(u; \theta)]^{-1} \quad (21)$$

where α, β are intercept and slope, $\theta = \{\eta, \delta, \kappa\}$ represent the parameters of the standard flattened logistic distribution $S_\varepsilon(u; \theta)$ with the density quantile function $[q_\varepsilon(u|\theta)]^{-1}$, u is the *depth* corresponding to the error in the model for the speed y given the stopping distance x and the regression parameters $\{\alpha, \beta, \theta\}$. The depth u can be computed by inverting the PQR QF $u = \widehat{Q_Y^{-1}}(y|\alpha, \beta, \theta; x)$ (21).

For each of the n observations of speed in the sample $\underline{Y} = \{y_1, y_2, \dots, y_n\}$ we can compute $\underline{Q_Y} = \{Q_Y(u_1|\alpha, \beta, \theta, x_1), \dots, Q_Y(u_n|\alpha, \beta, \theta, x_n)\}$, such that $u_i = \widehat{Q_Y^{-1}}(y_i|\alpha, \beta, \theta, x_i)$, $i = \{1, 2, \dots, n\}$.

Let's further assume that the expert's prior belief about the intercept was elicited using a set of quantile-probability pairs and the best fit was achieved using the FLD quantile function with hyperparameters $\chi = 1, \eta = 1$, and $\kappa = 10$. Similarly, the expert belief about the slope is described by the FSLD with hyperparameters $\chi = 2, \eta = 2, \delta = 0.8$, and $\kappa = 2$. The density quantile functions of the priors for the slope and intercept are shown in Figure 5.

Figure 5: Prior density (DQF) for slope and intercept

Since FLD and FSLD are quantile distributions, the prior for the parameters α and β of PQR must be defined in the quantile form. This means that the density quantile functions $f(Q_\alpha(v)), f(Q_\beta(w))$ and the Jacobian adjustments $|q_\alpha(v)|, |q_\beta(w)|$ can be dropped, as explained in Section 3.2 above.

$$\begin{aligned} f(Q_\alpha(v), Q_\beta(w), \theta | \underline{Q_Y}, x) &\propto \mathcal{L}(\theta; \underline{Q_Y}, x) f(Q_\alpha(v)) |q_\alpha(v)| f(Q_\beta(w)) |q_\beta(w)| f(\theta) \implies \\ f(Q_\alpha(v), Q_\beta(w), \theta | \underline{Q_Y}, x) &\propto \mathcal{L}(\theta; \underline{Q_Y}, x) f(\theta) \end{aligned} \quad (22)$$

Therefore, the posterior distribution of the PQR parameters $\alpha = Q_\alpha(v), \beta = Q_\beta(w)$, and θ can

Table 3: Summary of the posterior samples from the FSLD PQR model

parameter	mean	median	q5	q95	rhat
Q(v)	3.8107	3.8201	3.4817	4.0702	1.012
Q(w)	4.5148	4.5133	4.4254	4.6059	1.013
eta	0.2691	0.2625	0.1937	0.3655	1.014
k	0.1142	0.0804	0.0063	0.3171	1.057
dlt	0.8065	0.8329	0.5766	0.9426	1.084

be expressed using the *quantile-based likelihood* (and the *quantile-based prior* for parameters α and β).

$$\begin{aligned}
f(\alpha, \beta, \theta | \underline{Q}_Y, \underline{x}) &\propto \mathcal{L}(\theta; \underline{Q}_Y, \alpha, \beta, \theta, \underline{x}) f(\theta) \\
\mathcal{L}(\theta; \underline{Q}_Y, \alpha, \beta, \theta, \underline{x}) &= \prod_{i=1}^n f(Q_Y(u_i | \alpha, \beta, \theta; x_i)) = \prod_{i=1}^n [q_\varepsilon(u_i | \theta) \sqrt{x_i}]^{-1} \\
\underline{Q}_Y(\underline{u} | \alpha, \beta, \theta; \underline{x}) &= \alpha + \beta \sqrt{\underline{x}} + S_\varepsilon(\underline{u}; \theta) \sqrt{\underline{x}} \\
u &\overset{y}{\sim} Q_Y(\alpha, \beta, \theta; d) \\
v &\overset{\alpha}{\sim} \text{FLD}(1, 1, 10) \\
w &\overset{\beta}{\sim} \text{FSLD}(2, 2, 0.8, 2) \\
\eta &\sim \text{Exp}(1/2) \\
\delta &\sim \text{Beta}(2, 1) \\
\kappa &\sim \text{Exp}(1/0.1)
\end{aligned} \tag{23}$$

where $f(\theta) = f(\eta)f(\delta)f(\kappa)$.

We ran 2500 post-warmup iterations and 4 chains using the Robust Adaptive Metropolis algorithm by [Vihola \(2012\)](#) implemented in `fmcmc` package ([Vega Yon and Marjoram, 2019](#)) in R ([R Core Team, 2021](#)). The code is provided in the Supplementary Materials.

Table 3 summarizes the posterior distribution of the parameters in the parametric quantile regression model for the car stopping distances.

Posterior predictive check ([Gabry et al., 2019](#)) can be done by generating a grid of values for the car stopping distances x and using randomly sampled parameters from the posterior distribution to compute the value of the response y using the PQR QF. Since in the PQR the regression equation is expressed in terms of the depth u we can extract the coherent (non-crossing) quantile regression lines for any set of fractiles. Figure 6 illustrates hypothetical outcome plots for the 5th, 50th, and 95th quantile regression lines.

Figure 6: Posterior predictive quantiles (0.05, 0.5, 0.95) for stopping distances

6 Discussion and conclusion

In the past 20 years, many examples of using quantile distributions for the approximate Bayesian computation (ABC) appeared in the literature (Allingham et al., 2009; Drovandi and Pettitt, 2011; Dunson and Taylor, 2005; McVinish, 2012; Smithson and Shou, 2017). ABC methods normally do not require computation of the likelihood, which, in case of the quantile distributions, is convenient, as these distributions lack an explicit CDF and PDF.

Regardless of whether the distribution is defined by the CDF of the QF, the defining function sometimes needs to be inverted. If the inverse does not exist in closed form, the function has to be inverted numerically. In the case of the *density-based likelihood*, the inverse distribution function may be needed for sampling from the posterior (e.g. for the posterior predictive check). In the case of the *quantile-based likelihood*, the inverse is needed for computing the intermediate depth values, corresponding to observations (conditional on covariates) for every draw of the parameters. No numerical inversion of the quantile function is needed for defining the *quantile-based prior*. A wide selection of efficient root-finding algorithm implementations in the popular statistical software makes the inversion of custom quantile functions accessible. We provide a generic wrapper for inverting arbitrary quantile functions using Brent method in the accompanying R package (Perepolkin, 2019). Further research of custom root-finding algorithms for non-decreasing functions on unit-interval can make inverting of quantile function even more computationally efficient.

The *quantile-based inference* opens up a wide set of new distributions to serve as likelihood and/or prior in Bayesian models. Although many flexible *density-defined* distributions have been proposed in recent decades (Jones, 2015; Steel and Rubio, 2015), *quantile* distributions

play an important role in certain field applications (Nair et al., 2013; Chalabi et al., 2012), as well as in expert knowledge elicitation and decision analysis (Mikkola et al., 2021; Hadlock, 2017; Powley, 2013). Besides, the flexibility offered by the distributions defined in terms of the quantile function (Gilchrist, 2007), and in particular their easily extensible nature (Table 1), allows ultimate freedom in expressing the expert-informed priors. In this paper we showed the connection of quantile parameter transformation to inverse transform sampling and used quantile distribution as a prior for regression parameters.

Multivariate versions of quantile distributions have been explored in the past (Field and Genton, 2006; Vineshkumar and Nair, 2019), but their adoption in the scientific literature remains low. One possibility of utilizing the flexibility of the quantile distributions in a multivariate setting is to employ them as marginal distributions for bivariate copulas, which can be assembled into higher-dimensional structures using vines (Czado, 2019; Kurowicka and Joe, 2011). When used as priors (Wilson, 2018), the copula structure can be elicited from the experts (Elfadaly and Garthwaite, 2017) along with the marginal quantile-probability pairs for fitting the quantile distribution (O’Hagan et al., 2006; Mikkola et al., 2021). Versatile and user-friendly multivariate quantile distributions represent an opportunity for further research.

Gilchrist (2007) provides a review of the traditional approach to quantile regression, as proposed by Koenker and Bassett (1978) and contrasts it with the *fully parametric approach* taken by PQR (Gilchrist, 2000, 2008; Su, 2015). The parametric approach to regression provides coherent (non-crossing) estimates of posterior quantiles, allowing the scientists to model the distribution of the error term explicitly (instead of making assumptions). Note that the parametric quantile regression may also be used with invertible distributions (logistic, normal, etc), as long they have computable QF and QDF ($\Phi^{-1}(u)$ and $\Phi^{-1}(u)/du$, for normal distribution).

Traditionally, the fitting of parameters in quantile distributions was performed using the matching of moments or L-moments (Gilchrist, 2008; Asquith, 2007; Karvanen and Nuutinen, 2008), matching of percentiles (Karian and Dudewicz, 2011), location and scale-free shape functionals (King and MacGillivray, 2007), distributional least squares/absolutes (Gilchrist, 2007; Sharma and Chakrabarty, 2020), and maximum likelihood (Rayner and MacGillivray, 2002; Su, 2007; Tarsitano, 2005). The various methods of obtaining parameter estimates for the quantile distributions have been extensively studied and compared, primarily in application to GAD (King and MacGillivray, 1999; Karian and Dudewicz, 2011; Fournier et al., 2007; Tarsitano, 2010), but also to some other distributions (Rayner and MacGillivray, 2002; Jeong-Soo, 2005). This paper generalizes the approach to *quantile-based likelihood* (Gilchrist, 2000; Rayner and MacGillivray, 2002; Haynes and Mengersen, 2005; Nair et al., 2020) connecting the previous research on parametric quantile regression (Gilchrist, 2008; Su,

2015; Sharma and Chakrabarty, 2020) with more recently introduced work on *quantile-based priors* (Nair et al., 2020) and implementing both of these concepts in Stan (Gabry and Češnovar, 2022) and R (Vega Yon and Marjoram, 2019).

Embracing and expanding the use of quantile distributions in Bayesian analysis can enable new solutions for old problems and enrich the toolkit available to scientists for performing hard inference tasks. We hope that the *quantile-based inference* methods presented in this paper can contribute to expanding the body of knowledge about the use of quantile functions in Bayesian statistics and fuel further research in the area of quantile distributions.

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.

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>

Supplementary materials

Stan and R code

Target markdown (Landau, 2021) with corresponding targets defined in R and Stan for executing all examples in the paper.

Moebius strip of probability functions

DIY souvenir illustrating the relationship between the probability functions.

References

- Aarset, M. V. (1987). How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, R-36(1):106–108.
- Acton, F. S. (1990). *Numerical Methods That Work*. Mathematical Association of America, Washington, D.C.
- Aldeni, M., Lee, C., and Famoye, F. (2017). Families of distributions arising from the quantile of generalized lambda distribution. *Journal of Statistical Distributions and Applications*, 4(1):25.
- Alefeld, G. E., Potra, F. A., and Shi, Y. (1995). Algorithm 748: Enclosing zeros of continuous functions. *ACM Transactions on Mathematical Software*, 21(3):327–344.
- Allingham, D., King, R. A. R., and Mengersen, K. L. (2009). Bayesian estimation of quantile distributions. *Statistics and Computing*, 19(2):189–201.
- Andrilli, S. F. and Hecker, D. (2010). *Elementary Linear Algebra*. Elsevier Academic Press, Amsterdam ; Boston, 4th ed edition.
- Asquith, W. H. (2007). L-moments and tl-moments of the generalized lambda distribution. *Computational Statistics & Data Analysis*, 51(9):4484–4496.
- Atkinson, K. E. (2008). *An Introduction to Numerical Analysis*. John wiley & sons.
- Baey, C., Smith, H. G., Rundlöf, M., Olsson, O., Clough, Y., and Sahlin, U. (2022). Calibration of a bumble bee foraging model using approximate bayesian computation.
- Balakrishnan, N. and So, H. Y. (2015). A generalization of quantile-based skew logistic distribution of van staden and king. *Statistics & Probability Letters*, 107:44–51.
- Bernton, E., Jacob, P. E., Gerber, M., and Robert, C. P. (2019). Approximate bayesian computation with the wasserstein distance. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(2):235–269.
- Burden, R. L. and Faires, J. D. (2011). *Numerical Analysis*. Brooks/Cole, Cengage Learning, Boston, MA, 9th ed edition.
- Chalabi, Y., Scott, D. J., and Wuertz, D. (2012). Flexible distribution modeling with the generalized lambda distribution. Working paper, ETH, Zurich, Switzerland.
- Chandrupatla, T. R. (1997). A new hybrid quadratic/bisection algorithm for finding the zero of a nonlinear function without using derivatives. *Advances in Engineering Software*, 28(3):145–149.

- 497 Czado, C. (2019). *Analyzing Dependent Data with Vine Copulas*. Springer Berlin Heidelberg,
498 New York, NY.
- 499 Dean, B. and King, A. (2009). Versatile regression: Simple regression with a non-normal
500 error distribution. In *Third Annual Applied Statistics Education and Research Collaboration*
501 *(ASEARC) Conference*, pages 7–8.
- 502 Dedduwakumara, D. S., Prendergast, L. A., and Staudte, R. G. (2021). An efficient estimator
503 of the parameters of the generalized lambda distribution. *Journal of Statistical Computation*
504 *and Simulation*, 91(1):197–215.
- 505 Dion, P., Galbraith, N., and Sirag, E. (2020). Using expert elicitation to build long-term
506 projection assumptions. In Mazzucco, S. and Keilman, N., editors, *Developments in*
507 *Demographic Forecasting*, volume 49, pages 43–62. Springer International Publishing,
508 Cham.
- 509 Drovandi, C. C. and Pettitt, A. N. (2011). Likelihood-free bayesian estimation of multivariate
510 quantile distributions. *Computational Statistics & Data Analysis*, 55(9):2541–2556.
- 511 Dunson, D. B. and Taylor, J. A. (2005). Approximate bayesian inference for quantiles. *Journal*
512 *of Nonparametric Statistics*, 17(3):385–400.
- 513 Elfadaly, F. G. and Garthwaite, P. H. (2017). Eliciting dirichlet and gaussian copula prior
514 distributions for multinomial models. *Statistics and Computing*, 27(2):449–467.
- 515 Fearnhead, P. and Prangle, D. (2012). Constructing summary statistics for approximate
516 bayesian computation: Semi-automatic approximate bayesian computation. *Journal of the*
517 *Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474.
- 518 Field, C. and Genton, M. G. (2006). The multivariate g-and-h distribution. *Technometrics*,
519 48(1):104–111.
- 520 Fournier, B., Rupin, N., Bigerelle, M., Najjar, D., Iost, A., and Wilcox, R. (2007). Estimating
521 the parameters of a generalized lambda distribution. *Computational Statistics & Data*
522 *Analysis*, 51(6):2813–2835.
- 523 Freimer, M., Kollia, G., Mudholkar, G. S., and Lin, C. T. (1988). A study of the generalized
524 tukey lambda family. *Communications in Statistics-Theory and Methods*, 17(10):3547–3567.
- 525 Gabry, J. and Češnovar, R. (2022). *Cmdstanr: R Interface to 'CmdStan'*.
- 526 Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., and Gelman, A. (2019). Visualization in
527 bayesian workflow. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*,
528 182(2):389–402.

- 529 Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013).
530 *Bayesian Data Analysis*. CRC press.
- 531 Gelman, A., Hill, J., and Vehtari, A. (2021). *Regression and Other Stories*. Analytical
532 Methods for Social Research. Cambridge University Press, Cambridge New York, NY Port
533 Melbourne, VIC New Delhi Singapore.
- 534 Gilchrist, W. (1997). Modelling with quantile distribution functions. *Journal of Applied*
535 *Statistics*, 24(1):113–122.
- 536 Gilchrist, W. (2000). *Statistical Modelling with Quantile Functions*. Chapman & Hall/CRC,
537 Boca Raton.
- 538 Gilchrist, W. (2008). Regression revisited. *International Statistical Review*, 76(3):401–418.
- 539 Gilchrist, W. G. (2007). Modeling and fitting quantile distributions and regressions. *American*
540 *Journal of Mathematical and Management Sciences*, 27(3-4):401–439.
- 541 Gu, M., Bhattacharjya, D., and Subramanian, D. (2018). Nonparametric estimation of utility
542 functions. *arXiv:1807.10840 [stat]*.
- 543 Gupta, R. D. and Kundu, D. (2007). Generalized exponential distribution: Existing results
544 and some recent developments. *Journal of Statistical Planning and Inference*, 137(11):3537–
545 3547.
- 546 Hadlock, C. C. (2017). *Quantile-Parameterized Methods for Quantifying Uncertainty in*
547 *Decision Analysis*. PhD thesis, University of Texas, Austin, TX.
- 548 Hadlock, C. C. and Bickel, J. E. (2017). Johnson quantile-parameterized distributions.
549 *Decision Analysis*, 14(1):35–64.
- 550 Hadlock, C. C. and Bickel, J. E. (2019). The generalized johnson quantile-parameterized
551 distribution system. *Decision Analysis*, 16(1):67–85.
- 552 Haynes, M. and Mengersen, K. (2005). Bayesian estimation of g-and-k distributions using
553 mcmc. *Computational Statistics*, 20(1):7–30.
- 554 Householder, A. S. (1970). *The Numerical Treatment of a Single Nonlinear Equation*.
555 McGraw-Hill, London.
- 556 Jacob, P. (2017). Likelihood calculation for the g-and-k distribution.
- 557 Jeong-Soo, P. (2005). Wakeby distribution and the maximum likelihood estimation algorithm
558 in which probability density function is not explicitly expressed. *Communications for*
559 *Statistical Applications and Methods*, 12(2):443–451.

- Johnson, A. A., Ott, M. Q., and Dogucu, M. (2022). *Bayes Rules! An Introduction to Bayesian Modeling with R*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, Boca Raton.
- 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.
- Jones, M. C. (2015). On families of distributions with shape parameters. *International Statistical Review / Revue Internationale de Statistique*, 83(2):175–192.
- Karabatsos, G. and Leisen, F. (2018). An approximate likelihood perspective on abc methods. *Statistics Surveys*, 12(none):66–104.
- Karian, Z. A. and Dudewicz, E. J. (2011). *Handbook of Fitting Statistical Distributions with R*. CRC Press, Boca Raton, FL.
- Karvanen, J. and Nuutinen, A. (2008). Characterizing the generalized lambda distribution by l-moments. *Computational Statistics & Data Analysis*, 52(4):1971–1983.
- 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.
- King, R. A. and MacGillivray, H. L. (2007). Fitting the generalized lambda distribution with location and scale-free shape functionals. *American Journal of Mathematical and Management Sciences*, 27(3-4):441–460.
- King, R. A. R. (1999). *New Distributional Fitting Methods Applied to the Generalised [Lambda] Distribution*. PhD thesis, Queensland University of Technology, Australia.
- King, R. A. R. and MacGillivray, H. L. (1999). A starship estimation method for the generalized lambda distributions. *Australian & New Zealand Journal of Statistics*, 41(3):353–374.
- Koenker, R. (2005). *Quantile Regression*. Number no. 38 in Econometric Society Monographs. Cambridge University Press, Cambridge ; New York.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.
- Koller, D. and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. Adaptive Computation and Machine Learning. MIT Press, Cambridge, MA.
- Kurowicka, D. and Joe, H., editors (2011). *Dependence Modeling: Vine Copula Handbook*. World Scientific, Singapore.

- Lambert, B. (2018). *A Student's Guide to Bayesian Statistics*. SAGE, Los Angeles.
- Lampasi, D. (2008). An alternative approach to measurement based on quantile functions. *Measurement*, 41(9):994–1013.
- Landau, W. M. (2021). The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software*, 6(57):2959.
- Larrain, M., Van Passel, S., Thomassen, G., Van Gorp, B., Nhu, T. T., Huysveld, S., Van Geem, K. M., De Meester, S., and Billen, P. (2021). Techno-economic assessment of mechanical recycling of challenging post-consumer plastic packaging waste. *Resources, Conservation and Recycling*, 170:105607.
- Lutus, P. (2021). The physics behind stopping a car. https://arachnoid.com/braking_physics/index.html.
- Marsden, J. E., Weinstein, A., and Marsden, J. E. (1985). *Calculus I*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 2nd ed edition.
- McVinish, R. (2012). Improving abc for quantile distributions. *Statistics and Computing*, 22(6):1199–1207.
- Mikkola, P., Martin, O. A., Chandramouli, S., Hartmann, M., Pla, O. A., Thomas, O., Pesonen, H., Corander, J., Vehtari, A., Kaski, S., Bürkner, P.-C., and Klami, A. (2021). Prior knowledge elicitation: The past, present, and future. *arXiv:2112.01380 [stat]*.
- Muraleedharan, G., Lucas, C., and Guedes Soares, C. (2016). Regression quantile models for estimating trends in extreme significant wave heights. *Ocean Engineering*, 118:204–215.
- Myerson, R. B. (2005). *Probability Models for Economic Decisions*. Duxbury Applied Series. Thomson/Brooke/Cole, Belmont, CA.
- Nadarajah, S. (2009). Bathtub-shaped failure rate functions. *Quality & Quantity*, 43(5):855–863.
- Nair, N. U., Sankaran, P., and Balakrishnan, N. (2013). *Quantile-Based Reliability Analysis*. Springer New York, New York, NY.
- Nair, N. U., Sankaran, P. G., and Dileepkumar, M. (2020). Bayesian inference in quantile functions. *Communications in Statistics - Theory and Methods*, 0(0):1–13.
- Nair, N. U., Sankaran, P. G., and Vineshkumar, B. (2012). The govindarajulu distribution: Some properties and applications. *Communications in Statistics - Theory and Methods*.

- O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities: O'Hagan/Uncertain Judgements: Eliciting Experts' Probabilities*. John Wiley & Sons, Ltd, Chichester, UK.
- Olkin, I. and Tate, R. F. (1961). Multivariate correlation models with mixed discrete and continuous variables. *The Annals of Mathematical Statistics*, 32(2):448–465.
- Parzen, E. (1979). Nonparametric statistical data modeling. *Journal of the American Statistical Association*, 74(365):105–121.
- Parzen, E. (1980). Data modeling using quantile and density-quantile functions. Technical report, Texas A & M Univ College Station Inst Of Statistics.
- Perepolkin, D. (2019). *Qpd: Tools for Quantile-Parameterized Distributions*.
- Perri, P. F. and Tarsitano, A. (2007). Partially adaptive estimation via quantile functions. *Communications in Statistics - Simulation and Computation*, 36(2):277–296.
- Perri, P. F. and Tarsitano, A. (2008). Distributional least squares based on the generalized lambda distribution. In *International Conference on Computational Statistics. COMPSTAT 2008*, volume 400, pages 341–348. Physica-Verlag, A Springer Company.
- Powley, B. W. (2013). *Quantile Function Methods for Decision Analysis*. Stanford University.
- Prangle, D. (2017). Gk: An r package for the g-and-k and generalised g-and-h distributions. *arXiv:1706.06889 [stat]*.
- Press, W. H., editor (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK ; New York, 3rd ed edition.
- Price, G. B. (1984). *The Inverse-Function Theorem*, pages 237–262. Springer New York, New York, NY.
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rahman, A., Zaman, M. A., Haddad, K., El Adlouni, S., and Zhang, C. (2015). Applicability of wakeby distribution in flood frequency analysis: A case study for eastern australia. *Hydrological Processes*, 29(4):602–614.
- Ramberg, J. S. and Schmeiser, B. W. (1974). An approximate method for generating asymmetric random variables. *Communications of the ACM*, 17(2):78–82.

- Rayner, G. D. and MacGillivray, H. L. (2002). Numerical maximum likelihood estimation for the g-and-k and generalized g-and-h distributions. *Statistics and Computing*, 12(1):57–75.
- Reinhardt, J. C., Chen, X., Liu, W., Manchev, P., and Paté-Cornell, M. E. (2016). Asteroid risk assessment: A probabilistic approach. *Risk Analysis*, 36(2):244–261.
- Ridders, C. (1979). A new algorithm for computing a single root of a real continuous function. *IEEE Transactions on circuits and systems*, 26(11):979–980.
- Schäling, B. (2011). *The Boost C++ Libraries*. Boris Schäling.
- Sharma, D. and Chakrabarty, T. K. (2017). Some general results on quantile functions for the generalized beta family. *Statistics, Optimization & Information Computing*, 5(4):360–377.
- Sharma, D. and Chakrabarty, T. K. (2019). The quantile-based flattened logistic distribution: Some properties and applications. *Communications in Statistics - Theory and Methods*, 48(14):3643–3662.
- Sharma, D. and Chakrabarty, T. K. (2020). A quantile-based approach to supervised learning. In Johri, P., Verma, J. K., and Paul, S., editors, *Applications of Machine Learning*, pages 321–340. Springer Singapore, Singapore.
- Smithson, M. and Shou, Y. (2017). Cdf-quantile distributions for modelling random variables on the unit interval. *British Journal of Mathematical and Statistical Psychology*, 70(3):412–438.
- Stage, S. A. (2013). Comments on an improvement to the brent’s method. *International Journal of Experimental Algorithms*, 4(1):1–16.
- Steel, M. F. J. and Rubio, F. J. (2015). On families of distributions with shape parameters: Discussion. *International Statistical Review*, 83(2):218–222.
- Su, S. (2007). Numerical maximum log likelihood estimation for generalized lambda distributions. *Computational Statistics & Data Analysis*, 51(8):3983–3998.
- Su, S. (2015). Flexible parametric quantile regression model. *Statistics and Computing*, 25(3):635–650.
- Tarsitano, A. (2005). Fitting wakeby model using maximum likelihood. In *Statistica e Ambiente*, volume 1, pages 253–256.
- Tarsitano, A. (2010). Comparing estimation methods for the fpld. *Journal of Probability and Statistics*, 2010:1–16.

- 682 Tukey, J. W. (1965). Which part of the sample contains the information? *Proceedings of the*
683 *National Academy of Sciences*, 53(1):127–134.
- 684 van Staden, P. J. and King, R. A. R. (2015). The quantile-based skew logistic distribution.
685 *Statistics & Probability Letters*, 96:109–116.
- 686 Vega Yon, G. and Marjoram, P. (2019). Fmcmc: A friendly mcmc framework. *The Journal*
687 *of Open Source Software*, 4(39).
- 688 Vihola, M. (2012). Robust adaptive metropolis algorithm with coerced acceptance rate.
689 *Statistics and Computing*, 22(5):997–1008.
- 690 Vineshkumar, B. and Nair, N. U. (2019). Bivariate quantile functions and their applications
691 to reliability modelling. *Statistica*, 79(1):3–21.
- 692 Wilson, K. J. (2018). Specification of informative prior distributions for multinomial models
693 using vine copulas. *Bayesian Analysis*, 13(3):749–766.
- 694 Yu, K. and Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics & Probability*
695 *Letters*, 54(4):437–447.
- 696 Zhang, Z. (2011). An improvement to the brent’s method. *International Journal of Experi-*
697 *mental Algorithms*, 2(1):21–26.