New Distributions for Modeling Subjective Lower and Upper Probabilities

Michael Smithson MICHAEL.SMITHSON@ANU.EDU.AU
The Australian National University, Canberra (Australia)

Abstract

This paper presents an investigation of approaches to modeling lower and upper subjective probabilities. A relatively unexplored approach is introduced, based on the fact that every cumulative distribution function (CDF) with support (0,1) has a "dual" CDF that obeys the conjugacy relation between coherent lower and upper probabilities. A new 2-parameter family of "CDF-Quantile" distributions with support (0,1) is extended via a third parameter for the purpose of modeling lower-upper probabilities. The extension exploits certain properties of the CDF-Quantile family, and the fact that continuous CDFs on (0,1) random variables form an algebraic group that is closed under composition. This extension also yields models for testing specific models of lower-upper probability assignments. Finally, the new models are applied to a real data-set, and compared with the alternative approaches for their relative advantages and drawbacks.

Keywords: probability judgment; distribution; quantile regression; generalized linear model.

1. Introduction

This paper presents an investigation of approaches to modeling lower and upper subjective probabilities. This investigation springs from two motivational sources. First, it is motivated by the many applications in which interval-valued probability assignments play a role in human probability judgments, whether as input into decision making and forecasting or as risk communication (e.g., Budescu et al., 2014). Second, it is motivated by recent developments for modeling random variables on the (0,1) interval, which have resulted in a new family of probability distributions with (0,1) support, described by Smithson and Merkle (2014) and elaborated in Smithson and Shou (2017).

We begin with a brief description of conventional methods for modeling lower-upper probabilities, followed by the introduction of a heretofore unexplored modeling approach. Then the new family of distributions is introduced, and extended for the purpose of modeling lower-upper probabilities via the methods described previously. Finally, the models are applied to real data-sets, and compared for their relative advantages and drawbacks.

Conventional statistical approaches to modeling lower-upper probability assignments treat them as a pair of dependent random variables. One type of method ignores the ordering and simply models the dependency either via a "subject-effect" parameter or a covariance. A somewhat more sophisticated regression-style approach uses a binary dummy predictor that takes a value of 0 for the lower probabilities and 1 for the upper probabilities and respects the ordering by restricting the coefficient to being non-negative by exponentiating it (e.g., Smithson et al., 2012).

This paper introduces another approach to modeling lower-upper probabilities, in which the probability distributions modeling the lower and upper probability assignments share parameters but take two different forms. This pair of distributions is determined by the so-called "conjugacy" relation between coherent lower and upper probabilities. Let $p_L(A) = W(p(A), \theta)$, be a lower probability with respect to probability $p(A)$ so that $0 \leq W\left(p\left(A\right), \theta\right) \leq p(A)$, for real-valued θ. The conjugate upper probability is $p_U\left(A\right) = 1 - p_L\left(\sim A\right)$, so that $p_U\left(A\right) = 1 - W\left(1 - p\left(A\right), \theta\right)$.

A version of this relationship may be identified in cumulative distribution functions (CDFs) for random variables on the (0,1) interval. Consider a CDF, $G(x, \theta)$, for $0 \leq x \leq 1$, with a location parameter, θ, so that $G(0, \theta) = 0$, $G(1, \theta) = 1$, and G is monotonically increasing in x. Define $G_D(x, \theta) = 1 - G(1 - x, \theta)$, which clearly also is a CDF. G_D is the *conjugate dual* of G, which follows by observing that

$$1 - G_D\left(1 - x, \theta\right) = 1 - \left[1 - G\left(1 - \left(1 - x\right), \theta\right)\right] = G\left(x, \theta\right) \tag{1}$$

As a simple example, consider $G(x, \theta) = x^\theta$, for $\theta > 0$. Then $G_D(x, \theta) = 1 - (1 - x)^\theta$. When $\theta < 1$ G is the upper CDF, when $\theta = 1$ we have the uniform distribution so that $G = G_D$, and when $\theta > 1$ G is the lower CDF.

A second example is the beta distribution. It is easy to show that if X is distributed $\mathrm{beta}(\omega, \tau)$ then G_D is the CDF of a random variable, X_D, say, that is distributed $\mathrm{beta}(\tau, \omega)$, i.e., the PDF of X flipped around 1/2. The absolute difference between their means, $|(\omega - \tau)/(\omega + \tau)|$, gives a convenient index of the distance between the lower and upper distributions. Reparameterizing the beta distribution so that the parameters are the mean, $\mu = \omega/(\omega + \tau)$, and precision, $\phi = \omega + \tau$, it is clear that the mean and precision of X jointly determine the magnitude of the difference between its distribution and that of and its conjugate dual X_D.

One- and two-parameter distributions of the kinds illustrated here have very limited flexibility regarding the location of G and G_D; typically the corresponding PDFs are mirror-images of one another centred on 1/2. Nevertheless, while these pairs of distributions may not be very useful for modeling real data, the concepts involved turn out to have such applications when applied to the family of distributions introduced in the next section.

2. CDF-Quantile Distributions

The family of distributions presented here is elaborated in Smithson and Shou (2017) and Shou and Smithson (2016) implement them in the R package `cdfquantile` for generalized linear modeling. This family is a special case of the T-X family presented by Aljarrah, et al. (2014), although it was independently described in Smithson and Merkle (2014). Let $G(x, \mu, \sigma)$ denote a CDF for random variable X with support (0, 1), a real-valued location parameter μ and positive scale parameter σ. We define *G* as follows:

$$G(x, \mu, \sigma) = F[U(H^{-1}(x), \mu, \sigma)] \tag{2}$$

where F is a CDF with support denoted by D_1, H is an invertible CDF with support denoted by D_2, and $U : D_2 \rightarrow D_1$ is an appropriate transform for incorporating parameters μ and σ. We limit the domains D_1 and D_2 to pairs taken from $(-\infty, \infty)$ and/or $(0, \infty)$, and the following cases of U.

For $D_1 = (-\infty, \infty)$ and $D_2 = (-\infty, \infty)$ we put

$$U(y, \mu, \sigma) = (y - \mu)/\sigma. \tag{3}$$

For $D_1 = (-\infty, \infty)$ and $D_2 = (0, \infty)$ we put

$$U\left(y, \mu, \sigma\right) = (\log\left(y\right) - \mu)/\sigma. \tag{4}$$

For $D_1 = (0, \infty)$ and $D_2 = (-\infty, \infty)$ we put

$$U(y, \mu, \sigma) = \exp(-\mu/\sigma)\exp(y/\sigma). \tag{5}$$

Finally, for $D_1 = (0, \infty)$ and $D_2 = (0, \infty)$ we put

$$U(y, \mu, \sigma) = \exp(-\mu/\sigma)\, y^{1/\sigma}. \tag{6}$$

If all the functions are differentiable then the PDF $g(x, \mu, \sigma)$ has an explicit expression. If F is invertible, then for every γ such that $G(x, \mu, \sigma) = \gamma$, the quantile functions corresponding to the cases described in equations (3) to (6) are as follows. For $D_1 = (-\infty, \infty)$ and $D_2 = (-\infty, \infty)$ we put

$$G^{-1}(\gamma, \mu, \sigma) = H\left[\sigma F^{-1}(\gamma) + \mu\right]. \tag{7}$$

For $D_1 = (-\infty, \infty)$ and $D_2 = (0, \infty)$ we put

$$G^{-1}(\gamma, \mu, \sigma) = H\left[\exp\left(\sigma F^{-1}(\gamma) + \mu\right)\right]. \tag{8}$$

For $D_1 = (0, \infty)$ and $D_2 = (-\infty, \infty)$ we put

$$G^{-1}(\gamma, \mu, \sigma) = H\left[\mu + \sigma \log\left(F^{-1}(\gamma)\right)\right]. \tag{9}$$

Finally, for $D_1 = (0, \infty)$ and $D_2 = (0, \infty)$ we put

$$G^{-1}(\gamma, \mu, \sigma) = H\left[\exp(\mu)\left(F^{-1}(\gamma)\right)^{\sigma}\right]. \tag{10}$$

Smithson and Shou (2017) present 36 members of the *CDF-Quantile* family by employing six standard distributions for *F* and *H*: The logistic, Cauchy, t with df = 2, arc-sinh, Burr VII, and Burr VIII distributions. All of these have explicit PDF, CDF, and quantile functions. Smithson and Shou observe that *F* and *H* may exchange roles. The resulting pairs of distributions are”quantile-duals” of one another in the sense that one’s CDF is the other’s quantile, with the appropriate parameterization. This duality is due to the fact that (0, 1) is both the domain and range of these functions. Smithson and Shou denote these distributions with the nomenclature *F*-*H* (e.g., Cauchit-Logistic and Logit-Cauchy).

Smithson and Shou (2017) show that the CDF-Quantile family members share the following properties:

1. The family can model a wide variety of distribution shapes, with different skew and kurtosis coverage from the beta or the Kumaraswamy.

2. (Proposition 1, from Smithson and Shou (2017)) Members are self-dual in the sense that $g(x, \mu, \sigma) = g(1 - x, -\mu, \sigma)$. Moreover, $G = G_D$, so the conjugate-CDF duals in this family consists of identical distributions.

3. (Proposition 2) The median is solely a function of μ, so that μ is genuinely a location parameter.

4. (Proposition 3) The parameter σ is a dispersion parameter.

5. (Proposition 4) Members of this family fall into four subfamilies distinguished by behavior at the boundaries of the [0, 1] interval, including a subfamily whose density is finite in the limits at 0 and at 1.

Thus, the CDF-Quantile family enables a wide variety of quantile regression models for random variables on the (0, 1) interval with predictors for both location and dispersion parameters, and simple interpretations of those parameters. Smithson and Shou demonstrate that members of the family can out-perform the beta and other two-parameter distributions in fitting real data. Because they have explicit CDFs and quantile functions, the CDF-Quantile family is well-suited for multivariate models using copulas, and an example of this application will be presented later in this paper. Shou and Smithson (2017) fit a trivariate copula model to real data as a demonstration of how this may be done using their `cdfquantreg` package in conjunction with the R package `copula`.

3. Introducing a Third Parameter to the CDF-Quantile Family

The fact that $G = G_D$ for the entire CDF-Quantile family implies that they may be well-suited to testing the conjugate-CDF model of lower and upper probabilities via the introduction of a third parameter. Unlike two-parameter distributions such as the beta distribution, for a three-parameter distribution the third parameter can determine the difference between a CDF and its conjugate dual CDF.

There are several ways to introduce a third parameter, but we will focus on doing so through a composition operator. Marshall and Olkin (2007, pp. 494-495) state that the class **G** of CDFs G whose support is (0,1) form an algebraic group. This is true of continuous CDFs. The class of continuous CDFs is closed under the composition operation $G_1 \bullet G_2 = G_1(G_2)$, and this operation also is associative. The uniform distribution is the identity. Likewise, for any G in **G**, the quantile function G^{-1} also is in **G**. The quantile-dual relation described in the preceding section is a special case of this type of closure.

A straightforward way to introduce a third parameter is via an invertible monotonic function applied either at the outermost or innermost level of the CDF or the quantile function. Applying an invertible $(0,1) \to (0,1)$ transformation W to the innermost level of the CDF, for instance, we have

$$G(x,\mu,\sigma,\theta) = F\left[U\left(H^{-1}\left(W(x,\theta)\right),\mu,\sigma\right)\right] \tag{11}$$

and

$$G^{-1}(\gamma,\mu,\sigma,\theta) = W^{-1}\left[H\left(U^{-1}\left(F^{-1}(\gamma),\mu,\sigma\right)\right),\theta\right] \tag{12}$$

If we additionally require that $W(0,\theta) = 0$, $W(1,\theta) = 1$ and W monotonically increasing in x, then W behaves as a CDF. The conjugate dual CDF therefore is

$$G_D(x,\mu,\sigma,\theta) = F\left[U\left(H^{-1}\left(1 - W(1-x,\theta)\right),\mu,\sigma\right)\right]. \tag{13}$$

Several kinds of CDFs for W and application of the CDF-composition operator are available from the literature on lifetime distributions. A power (resilience) parameter or a frailty parameter can be introduced in this way, by applying the CDF-composition operator. The relevant CDF is x^{θ}, for some $\theta > 0$. Slightly less obviously, introducing a tilt parameter also involves a CDF-composition, because, for $\theta > 0$, it is a composition of the CDF $x/(x + \theta(1-x))$ with $G(x,\mu,\sigma)$. Likewise, a hazard parameter can be introduced via composition using the CDF

$1-\exp\left[-(-\log(1-x))^{\theta}\right]$, for $\theta > 0$; and a Laplace transform parameter with the CDF $\left(1-e^{-\theta x}\right)/\left(1-e^{-\theta}\right)$, for real θ.

In the cases where the composition is $G \bullet W$, the introduction of the third parameter yields a three-parameter CDF-Quantile family with distinct CDFs and conjugate dual CDFs (i.e., $G \neq G_D$) and possessing certain properties paralleling those derived by Smithson and Shou (2017) for the two-parameter family. The following Proposition is an extension of Proposition 1 (the self-dual property) from Smithson and Shou (2017).

Proposition 5.1: Let $W(x,\theta)$ be defined as earlier, so that it behaves as a CDF. Let

$$G\left(W\left(x,\theta\right),\mu,\sigma\right)=F\left[U\left(H^{-1}\left(W\left(x,\theta\right)\right),\mu,\sigma\right)\right].$$

Then if the CDFs F and H satisfy certain symmetry conditions (in the 4 cases detailed below),

$$1-G\left(W\left(1-x,\theta\right),-\mu,\sigma\right)=G\left(1-W\left(1-x,\theta\right),\mu,\sigma\right). \tag{14}$$

Now define

$$G^{-1}\left(Z_1\left(\gamma,\mu,\sigma\right),\theta\right)=W^{-1}\left[H\left(U^{-1}\left(F^{-1}\left(\gamma\right),\mu,\sigma\right)\right),\theta\right],$$

and

$$G^{-1}\left(Z_2\left(\gamma,\mu,\sigma\right),\theta\right)=1-W^{-1}\left[1-H\left(U^{-1}\left(F^{-1}\left(\gamma\right),\mu,\sigma\right)\right),\theta\right].$$

These are the quantile functions corresponding to the conjugate dual CDFs $G\left(W\left(x,\theta\right),\mu,\sigma\right)$ and $G\left(1-W\left(1-x,\theta\right),\mu,\sigma\right)$, respectively. Then $G^{-1}\left(Z_1\left(\gamma,\mu,\sigma\right),\theta\right)$ and $G^{-1}\left(Z_2\left(\gamma,\mu,\sigma\right),\theta\right)$ behave as conjugate lower-upper probabilities.

Proof: The identity in equation (14) has four cases, corresponding to the four combinations of domains in the CDF-Quantile family.

Case 1: For $D_1=(-\infty,\infty)$ and $D_2=(-\infty,\infty)$ when $-H^{-1}(x)=H^{-1}(1-x)$ and $f(x)=f(-x)$, $1-G\left(W\left(1-x,\theta\right),-\mu,\sigma,\theta\right)=1-F\left[\left(H^{-1}\left(W\left(1-x,\theta\right)\right)+\mu\right)/\sigma\right]$
$=1-F\left[\left(-H^{-1}\left(1-W\left(1-x,\theta\right)\right)+\mu\right)/\sigma\right]=F\left[\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)-\mu\right)/\sigma\right]$
$=G\left(1-W\left(1-x,\theta\right),\mu,\sigma,\theta\right).$

Case 2: For $D_1=(-\infty,\infty)$ and $D_2=(0,\infty)$ when $H^{-1}(x)=1/H^{-1}(1-x)$ and $f(x)=f(-x)$, $1-G\left(W^{-1}\left(1-x,\theta\right),-\mu,\sigma,\theta\right)=1-F\left[\left(\log\left(H^{-1}\left(W\left(1-x,\theta\right)\right)\right)+\mu\right)/\sigma\right]$
$=1-F\left[\left(-\log\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)\right)+\mu\right)/\sigma\right]=F\left[\left(\log\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)\right)-\mu\right)/\sigma\right]$
$=G\left(1-W\left(1-x,\theta\right),\mu,\sigma,\theta\right).$

Case 3: For $D_1=(0,\infty)$ and $D_2=(-\infty,\infty)$ when $H^{-1}(x)=1/H^{-1}(1-x)$ and $F(x)=1-F(1/x)$, $1-G\left(1-x,-\mu,\sigma\right)=1-F\left[\left(H^{-1}\left(W\left(1-x,\theta\right)\right)\exp\left(\mu\right)\right)^{1/\sigma}\right]$
$=1-F\left[\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)\right)^{\sigma}\left(\exp\left(\mu\right)\right)^{1/\sigma}\right]=F\left[\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)\exp\left(-\mu\right)\right)^{1/\sigma}\right]$
$=G\left(1-W\left(1-x,\theta\right),\mu,\sigma,\theta\right).$

Case 4. For $D_1=(0,\infty)$ and $D_2=(0,\infty)$ when $-H^{-1}(x)=H^{-1}(1-x)$ and $F(x)=1-F(1/x)$, $1-G\left(1-x,-\mu,\sigma\right)=1-F\left[\exp\left(\left(-H^{-1}\left(W\left(1-x,\theta\right)\right)+\mu\right)/\sigma\right)\right]$
$=1-F\left[\exp\left(\left(-H^{-1}\left(1-W\left(1-x,\theta\right)\right)+\mu\right)/\sigma\right)\right]=F\left[\exp\left(\left(H^{-1}\left(1-W\left(1-x,\theta\right)\right)-\mu\right)/\sigma\right)\right]$
$=G\left(1-W\left(1-x,\theta\right),\mu,\sigma,\theta\right).$

The conjugacy relationship immediately follows immediately by observing that, in the definition of the quantile functions, $H\left(U^{-1}\left(F^{-1}(\gamma), \mu, \sigma\right)\right)$ fulfills the role of x in the function W^{-1}. *End of proof*.

The conjugate dual CDFs straddle the CDF $G(x, \mu, \sigma)$ and the resultant lower and upper quantile functions straddle the quantile function $G^{-1}(\gamma, \mu, \sigma)$. That is, the location of the conjugate-dual pair is determined by μ, which makes them flexible enough to be worthy candidates for modeling real data. Propositions 2-4 in Smithson and Shou (2017) also hold for these three-parameter CDF-Quantile distributions because W is monotonically increasing in x and we can write the quantile function as $W^{-1}\left[H\left(U^{-1}\left(F^{-1}(\gamma), \mu, \sigma\right)\right), \theta\right]$. Thus, the median is solely a function of μ and θ, and σ still is a dispersion parameter. Moreover, the θ parameter has an interpretation as a risk-attitude parameter, because it determines the difference between the lower and upper CDFs (and likewise the difference between the corresponding quantile functions). This three-parameter family therefore is suited to ascertaining whether samples of lower and upper probability assignments behave as though they come from populations with conjugate dual distributions.

4. Examples and Applications

4.1 $G \bullet W$ Conjugate Duals

In this subsection we will survey two examples of three-parameter CDF-Quantile distributions of the $G \bullet W$ type, each one corresponding to a well-known kind of parameterization borrowed from the life distributions literature. These include the power parameter (which in this case corresponds to a frailty parameter) and the tilt parameter. The Cauchit-Cauchy distribution will be used throughout this subsection for illustrative purposes (it also is employed in the data-fitting example in the next subsection).

Starting with the power parameter, $W(x, \theta) = x^\theta$ and so $1 - W(1 - x, \theta) = 1 - (1 - x)^\theta$. Applied to the Cauchit-Cauchy distribution, we have the conjugate CDF duals As its name suggests, both F and H are Cauchy CDFs, the power parameter (exponentiated) model simply replaces x with x^θ, and the conjugate-dual CDF pair is

$$G(x, \mu, \sigma) = \frac{1}{2} + \frac{\arctan\left(\left(\tan\left((2\pi x^\theta - \pi)/2\right) - \mu\right)/\sigma\right)}{\pi} \tag{15}$$

and

$$G_D(x, \mu, \sigma) = \frac{1}{2} + \frac{\arctan\left(\left(\tan\left((2\pi\left(1 - (1 - x)^\theta\right) - \pi)/2\right) - \mu\right)/\sigma\right)}{\pi} \tag{16}$$

When $\theta < 1$ then $G > G_D$, and when $\theta > 1$ then $G < G_D$.

The tilt parameter, as mentioned earlier, uses the CDF $W(x, \theta) = x/(x + \theta(1 - x))$. Applying it to the Cauchit-Cauchy distribution yields the conjugate CDF duals

$$G(x, \mu, \sigma) = \frac{1}{2} + \frac{\arctan\left(\left(\tan\left((2\pi x/(x + \theta(1 - x)) - \pi)/2\right) - \mu\right)/\sigma\right)}{\pi} \tag{17}$$

and

$$G_D(x, \mu, \sigma) = \frac{1}{2} + \frac{\arctan\left(\left(\tan\left((2\pi\theta x/(1 + x(\theta - 1)) - \pi)/2\right) - \mu\right)/\sigma\right)}{\pi} \tag{18}$$

This model behaves as a rescaled version of the constant-odds-ratio imprecise probability model described in Walley (1991) and elsewhere. When $\theta < 1$ then $G > G_D$, and when $\theta > 1$ then

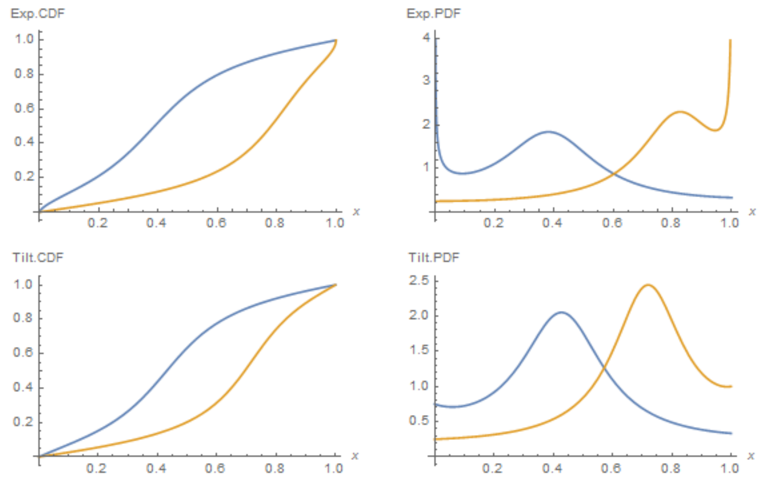

Figure 1: Power- and Tilt-Parameter Conjugate Dual Distributions

$G < G_D$. Figure 1 displays the pairs of CDFs and PDFs for the exponentiated and tilt parameter models when $\mu = 0.1, \sigma = 0.5$, and $\theta = 1.5$.

Finally, it is worth mentioning that because any CDF whose support is (0,1) can play the role of W, a one-parameter version of any member of the CDF-Quantile family may be used in that capacity, with θ as the location parameter. These alternatives would seem to present a forbiddingly large variety of models for analysts to consider. However, it turns out that under some conditions all of them can be very similar to one another with appropriate choices of θ. For many practical modeling purposes we may restrict attention to a subset of such models, such as the power and tilt parameter (constant odds-ratio) models, but at this stage of research on these models the best procedure for selecting among them remains an open topic for further investigation. The next section presents examples of model-fitting with a real data-set, demonstrating that conjugate dual lower-upper CDF models can fit lower-upper probability assignments quite well.

4.2 Fitting Models to Data

We now present an example of model-fitting that compares the conjugate lower-upper distributions with appropriate alternatives for modeling lower-upper probability assignments. The fourth Intergovernmental Panel on Climate Change (IPCC) report utilizes verbal phrases such as "likely" and "unlikely" to describe the uncertainties in climate science. Budescu et al. (2009) conducted an experimental study of lay interpretations of these phrases, using 13 sentences from the IPCC report, in which they asked 223 participants to provide lower, "best", and upper numerical estimates of the probabilities to which they believed each sentence referred. For example, participants were presented with the sentence "The Greenland ice sheet and other Arctic ice fields likely contributed no more than 4 m of the observed sea level rise.", and asked tp consider the probability they thought the report authors may have had in mind for the term "likely" in this sentence. Participants were required to provide their lowest, highest, and their best numerical estimates of this probability. Budescu et al. found that participants' "best" estimates were more regressive (toward the middle of the [0, 1] interval) than the IPCC stipulations, but they did not report systematic analyses of the lower and upper estimates.

I present 11 models fitted to the lower and upper probability estimates in the Budescu et al. data. The first three models are based on the two-parameter CDF-Quantile distribution. Model 1 is just the two-parameter distribution, as defined in equation (3), with intercept-only submodels $\hat{\mu} = \beta_0$ and $\hat{\sigma} = \exp(\delta_0)$. Model 2 has conditional parameter estimates, with submodels $\hat{\mu} = \beta_0 + \beta_1 x$ and $\hat{\sigma} = \exp(\delta_0 + \delta_1 x)$, where $x = 0$ for lower probabilities and $x = 1$ for upper probabilities. Model 3, in addition to the submodels from Model 2, also estimates the dependency between the lower and upper estimates via a t-copula with CDF-Quantile margins. This model therefore also includes estimates of the t-copula dependency parameter, ρ, and degrees of freedom parameter, ϕ.

Models 4-7 are based on the 3-parameter power (exponentiated) CDF-Quantile distribution, as in the CDF defined in equation (11) with $W(x, \theta) = x^{\theta}$. Model 4 has intercept-only submodels $\hat{\mu} = \beta_0$, $\hat{\sigma} = \exp(\delta_0)$, and $\hat{\theta} = \exp(\gamma_0)$. Model 5 is the conjugate-dual model, as defined in equations (11) and (13). This has the same intercept-only submodels as Model 4 but is a two-component distribution mixture model with a fixed mixture parameter, so that the first CDF, G, is weighted 1 and the second, G_D, is weighted 0 for the upper probabilities and the reverse weighting is applied to the lower probabilities. Technically, it is a four-parameter model although the mixture parameter is not being estimated. Model 6 has conditional parameter estimates, $\hat{\mu} = \beta_0 + \beta_1 x$ and $\hat{\sigma} = \exp(\delta_0 + \delta_1 x)$ with $x = 0$ and 1 for lower and upper probabilities, but an intercept-only submodel $\hat{\theta} = \exp(\gamma_0)$. Model 7 has the conditional μ and σ submodels in Model 6 plus $\hat{\theta} = \exp(\gamma_0 + \gamma_1 x)$. Finally, models 8-11 are based on the tilt-parameter CDF-Quantile distribution, as in the CDF defined in equation (11) with $W(x, \theta) = x/(x + \theta(1 - x))$. These models have the same variants as Models 4-7.

The best-fitting models from the CDF-Quantile family are from the "finite-tailed" subfamily, whose members have defined, finite densities at 0 and 1 (Smithson and Shou, 2017). The best-fitting distribution from this subfamily is the Cauchit-Cauchy, so the models considered here are mainly limited to that distribution. Table 1 displays goodness-of-fit statistics for the 11 models. The top section of the table presents these results for the three models using the two-parameter Cauchit-Cauchy. The middle section contains the power-parameter (exponentiated) models, and the lower section contains the tilted-parameter models. The "Params" column displays the number of parameters in each model, the "2LL" column shows twice the log-likelihood of the fitted models, and the "AIC" column is the Akaike Information Criterion, $AIC = -2LL + 2p$, where p is the number of parameters in the Params column.

Remarkably, the 4-parameter conjugate-dual models fit the data better than most of the 5- and 6-parameter conditional models and better than the 6-parameter copula model. The conjugate-dual power-parameter model is superior to the conjugate-dual tilted-parameter model, and is out-performed only by the 6-parameter conditional tilted-parameter model. Likewise, the conjugate-dual tilted-parameter model is out-performed only by the 5- and 6-parameter conditional tilted-parameter models and the 6-parameter conditional power-parameter model.

These results are not due to some kind of fluke in the Cauchit-Cauchy distribution. Other members of the finite-tailed subfamily have similar fits for their conjugate-dual models. For instance, the T2-T2 and the Cauchit-ArcSinh conjugate-dual power-parameter models have AIC's of -2159 and -2062, respectively, and both of these out-perform their respective 5- and 6-parameter conditional power-parameter counterparts.

Figure 2 shows the fitted distributions from the conjugate-dual model (top half of the figure) and the 6-parameter conditional exponentiated model. The two pairs of fitted distributions are strikingly similar and the conjugate-dual AIC is the better of the two. The facts that the 4-parameter

Table 1: Cauchit-Cauchy Models and Fits

Model	Description	Params.	2LL	AIC
1	2-parameter	2	595	-591
2	2-parameter condit. μ, σ	4	1378	-1370
3	2-parameter condit. t-copula	6	1584	-1572
4	exponentiated 3-param.	3	616	-609
5	conjugate-dual exponentiated	4	2378	-2372
6	exponentiated condit. μ, σ	5	1392	-1382
7	exponentiated condit. μ, σ, θ	6	1967	-1955
8	tilted 3-param.	3	880	-874
9	conjugate-dual tilted	4	1736	-1730
10	tilted condit. μ, σ	5	2152	-2142
11	tilted condit. μ, σ, θ	6	3118	-3106

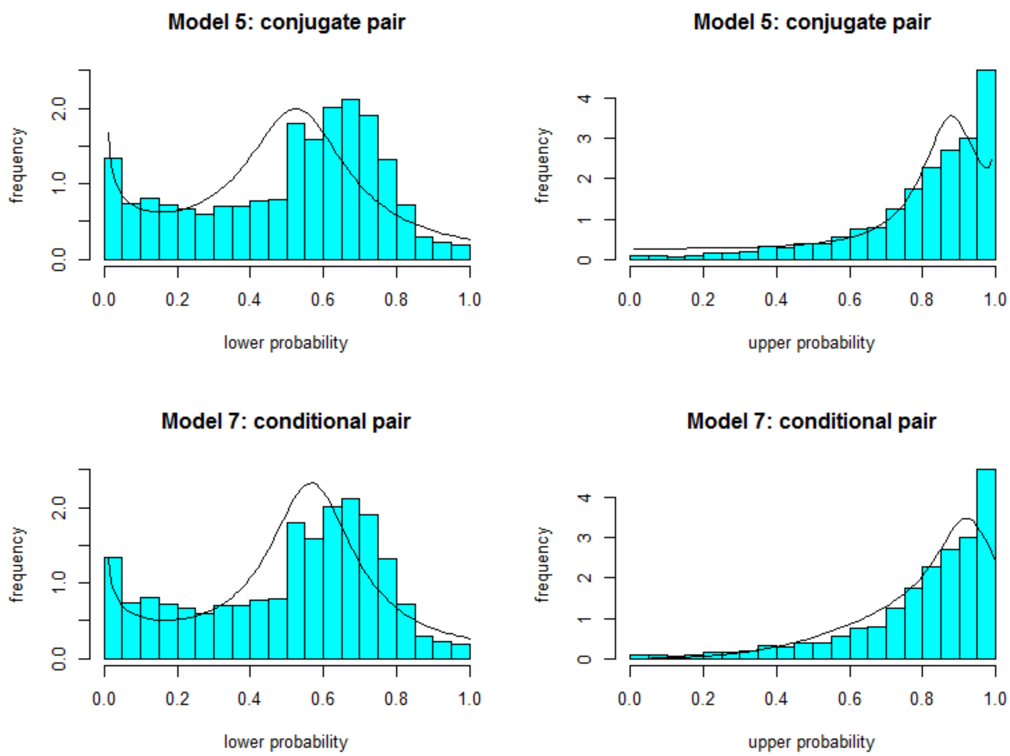

Figure 2: IPCC Data and Fitted Distributions

conjugate-dual model fits the data better than a regression model with 6 parameters and that the fitted distribution shapes are reasonably similar to the empirical distributions lend plausibility to the seemingly unlikely hypothesis that human lower-upper probability judgments are distributed approximately as conjugate-dual distributions.

The exponentiated Cauchit-Cauchy 4-parameter conjugate-dual and 6-parameter conditional regression models may be compared further via the 5-number summaries in Table 2. When compared with their empirical counterparts (rows 1 and 4 in the table), the conditional model is more accurate than the conjugate-dual model at the 10^{th} quantile, but the reverse is the case for most of the other quantiles. Both models appear to be fairly accurate in the middle 50% of the distributions. Again, this is an intriguing outcome for the conjugate-dual model, given that only three of its four parameters are being estimated from the data.

Table 2: Quantiles and Exponentiated Model Quantile Estimates

Model	Estimate	.1	.25	.5	.75	.9
	empirical lower	0.092	0.301	0.570	0.699	0.779
5	conjugate-dual lower	0.059	0.303	0.535	0.688	0.825
7	conditional lower	0.091	0.378	0.584	0.713	0.834
	empirical upper	0.540	0.729	0.858	0.948	0.998
5	conjugate-dual upper	0.298	0.684	0.863	0.935	0.977
7	conditional upper	0.495	0.672	0.846	0.935	0.975

That said, there are practical and technical issues in estimating both conjugate-dual and regression models for the 3-parameter CDF-Quantile distributions. For several of these distributions, maximum-likelihood estimations of conjugate-dual models of the IPCC data failed to converge, and regression models yielded high correlations between the parameter estimates for μ and θ (although the latter problem did not occur for any of the successful conjugate-dual models). Moreover, as Smithson and Shou (2017) observe, model diagnostics and related aspects of model evaluation for the 2-parameter CDF-Quantile family have yet to be completely thought through. Thus, the questions of effective estimation procedures and diagnostics for these models are active topics of research. Nonetheless, the evidence from the example in this section suggests that a sufficiently well-specified conjugate-dual model using 3-parameter CDF-Quantile distributions can be used to test a specific type of coherent lower-upper probability relationship.

5. Conclusions and Future Directions

A new family of probability distributions, the CDF-Quantile family, shows promise in modeling probability judgments. The two-parameter version of the family has been sufficiently well-explored by Smithson and Shou (2017) to have been made available for generalized linear modeling via the `cdfquantreg` package in R and a SAS macro, as presented by Shou and Smithson (2016, 2017), and those authors also have demonstrated that these distributions can model probabilities better than other two-parameter distributions such as the beta. This paper has presented an investigation of the application of the CDF-Quantile family to modeling imprecise distributions of probabilities, by extending it to incorporate a third parameter.

Because CDFs whose support is the (0,1) interval are closed under composition, and due to the properties of the CDF-Quantile distributions, three-parameter extensions via the composition of CDF functions yield conjugate dual pairs of CDFs. This result may hold some theoretical interest. A future line of research may elaborate the connections between these conjugate duals and imprecise probability frameworks. There is a natural link with probability boxes (p-boxes, as coined by Ferson et al. (2003)), given that the conjugate-dual CDFs form a p-box. Conjugate duals are noteworthy cases of p-boxes because the "width" of the gap between them is determined in a different way from the data-driven methods to which Ferson et al. (2003) refer. To my awareness, p-boxes have not been systematically studied regarding methods of fitting them to lower-upper probability data.

Some conjugate-dual models, in turn, have been found to fit a data-set reasonably well, raising the possibility that human lower-upper probability assignments may approximate a conjugacy relationship in their CDFs. Further research will determine whether these findings generalize to other such data-sets, if elicitation methods influence the results, and what judgment mechanisms or heuristics account for the phenomenon. However, perhaps the first priority is to ascertain the connections between the θ parameter, measurement error, and sampling error.

Finally, the three-parameter CDF-Quantile distributions also beg for further investigation. The overview in this paper only skims their characteristics, and little is known about the advantages and drawbacks of alternative parameterization methods for θ (e.g., power versus tilt parameters). Preliminary investigations suggest that the high correlations between parameter estimates may be a pervasive problem for three-parameter distributions on the unit interval (including three-parameter generalizations of the beta distribution). Likewise, as mentioned earlier, much remains to be developed and explored regarding parameter estimation methods and model diagnostics, even for the two-parameter CDF-Quantile family. The primary goals here have been to introduce this extension of the CDF-Quantile family and to make a case that it holds some promise for modeling distributions of lower-upper probability assignments. Accordingly, this paper may be regarded as a preliminary exploration of three-parameter CDF-Quantile distributions, with the unexpected finding that conjugate-dual distributions may be useful for modeling lower-upper probability assignments.

References

M. A. Aljarrah, C. Lee, and F. Famoye. On generating t-x family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, 1(1):2, 2014.

D. V. Budescu, S. Broomell, and H.-H. Por. Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, 20(3):299–308, 2009.

D. V. Budescu, H.-H. Por, S. B. Broomell, and M. Smithson. The interpretation of ipcc probabilistic statements around the world. *Nature Climate Change*, 4(6):508–512, 2014.

S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers, and K. Sentz. Constructing probability boxes and dempster-shafer structures. *Sandia National Laboratories*, pages 143–180, 2003.

A. W. Marshall and I. Olkin. *Life Distributions*. Springer, 2007.

Y. Shou and M. Smithson. *cdfquantreg: Quantile Regression for Random Variables on the Unit Interval*, 2016. R package version 1.1.0.

Y. Shou and M. Smithson. *cdfquantreg: An R package for CDF-Quantile Regression*, 2017. The Australian National University, Canberra, Australia.

M. Smithson and E. C. Merkle. *Generalized Linear Models for Categorical and Continuous Limited Dependent Variables*. Chapman and Hall/CRC Press, 2014.

M. Smithson and Y. Shou. Cdf-quantile distributions for modeling random variables on the unit interval. *British Journal of Mathematical and Statistical Psychology*, 2017. doi:10.1111/bmsp.12091.

M. Smithson, D. V. Budescu, S. B. Broomell, and H.-H. Por. Never say not: Impact of negative wording in probability phrases on imprecise probability judgments. *International Journal of Approximate Reasoning*, 53(8):1262–1270, 2012.

P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1991.