Bayesian Analysis (2007) **2**, Number 1, pp. 137–166

A Quantitative Study of Quantile Based Direct Prior Elicitation from Expert Opinion

Dipak K. Dey* and Junfeng Liu†

Abstract. Eliciting priors from expert opinion enjoys more efficiency and reliability by avoiding the statistician's potential subjectivity. Since elicitation on the predictive prior probability space requires too-simple priors and may be burdened with additional uncertainties arising from the response model, quantitative elicitation of flexible priors on the direct prior probability space deserves much attention. Motivated by precisely acquiring the shape information for the general location-scale-shape family beyond the limited and simple location-scale family, we investigate multiple numerical procedures for a broad class of priors, as well as interactive graphical protocols for more complicated priors. We highlight the quantile based approaches from several aspects, where Taylor's expansion is demonstrated to be an efficient approximate alternative to work on the regions in which the shape parameter is highly sensitive. By observing inherent associations between the scale and shape parameters, we put more weight on practical solutions under a proper sensitivity index (SI) rather than presumability. Our proposed methodology is demonstrated through skew-normal and Gamma hyperparameter elicitation where the shape parameter is numerically solved in a stable way. The performance comparisons among different elicitation approaches are also provided.

Keywords: location parameter, prior elicitation, quantile, scale parameter, shape parameter, skewness, Taylor's expansion

1 Introduction

Bayesian analysis is markedly recognized by the subjective probability belief, or quantitative *a priori* description of unknown parameter θ. Without external support statisticians can only implement the whole computational process at his/her own will, say conjugate priors, Jeffreys' priors and others. On the other hand, expert opinion may be helpful when we investigate new, rare, complex or poorly understood phenomena. Multivariate-normal related prior elicitation on the predictive prior space by requesting response summaries from experts was developed by Kadane, et al (1980), Garthwaite and Dickey (1988), Al-Awadhi and Garthwaite (1998) among others. However, those algorithms were limited to simple normal linear or AR(1) time series models. On the other hand, direct non-informative prior elicitation was discussed through piecewise conjugate priors (Meeden, 1992), entropy based priors (Jaynes, 1968, 1983), mixture of natural conjugate priors (Dalal and Hall, 1983) and others. Quantile based univari-

*Department of Statistics, University of Connecticut, Storrs, CT, http://www.stat.uconn.edu/~dey

*Department of Statistics, West Virginia University, Morgantown, WV, mailto:jfliu@stat.wvu.edu

ate prior elicitation for simple cases, say symmetric ones, was studied by Peterson and Miller (1964), Garthwaite and Dickey (1985), O'Hagan (1998) among others. A recent comprehensive review on probability elicitation was written by Garthwaite, Kadane and O'Hagan (2005). Berger (1985, Chapter 3) also discussed subjective prior determination on the direct prior space by showing that a lack of sufficient tail information in continuous parameter space causes much difficulty for most of the approaches in practice, including the "histogram" approach, the "relative likelihood" approach, the entropy based method, and even the most used "matching a given functional form" method which often needs prior moments and others. Berger (1985, Chapter 3) envisioned that a quantile based approach poses as a better method since estimation of probabilities of regions are more attractive than working on moments. However, two situations deserve caution in application of quantile approaches: disagreement among multiple quantiles and incidental matching by multiple functional forms (Berger, 1985, Chapter 3). The key point to ease these concerns is to efficiently and precisely recover flexible parametric priors in a quantitative way other than those weak symmetric ones in order to implement the "sketching" principle (Berger, 1985, Chapter 3) for the downstream graphical verification. In this paper, we propose efficient quantile based approaches for direct prior elicitation with quantitative examples, and study a broad class of parametric priors with a focus on asymmetric cases, which have yet received little attention.

The paper is organized as follows: Section 2 develops approximate and exact quantile based numerical elicitation algorithms for general symmetric and skewed priors; Section 3 introduces Taylor's expansion as an efficient tool for joint scale-shape parameter elicitation in sensitive regions and makes comparison among multiple expansion options with skew-normal and Gamma distributions as examples; Section 4 provides interactive graphical display procedures, especially for more complicated Student's t related prior elicitation; and Section 5 concludes the paper with future directions.

2 Symmetric and Asymmetric Prior Elicitation

2.1 Normal Prior

Garthwaite and Dickey (1985) evaluated double- and single-bisection methods for subjective probability assessment in a location-scale family, where single-bisection means median plus one quantile, double-bisection means 0.25, 0.50 and 0.75 quantiles: $q_{0.25}$, $q_{0.50}$ and $q_{0.75}$. They showed that, the double-bisection method is more favorable for most densities. Throughout this paper, we will use the notation IQR to represent the inter-quartile range $q_{0.75} - q_{0.25}$ and additional subscripts to specify certain parametric families. More specifically, Z_q represents the IQR of standard normal distribution. Suppose further $q_{0.75} - q_{0.50} \approx q_{0.50} - q_{0.25}$, then the prior mean and the standard deviation are estimated by

$$\hat{\mu}_N = q_{0.50} \quad \text{and} \quad \hat{\sigma}_N = \frac{\text{IQR}}{Z_q} \quad \text{respectively.} \tag{1}$$

Note that, normal distribution is symmetric and shape parameter free. If both the experimenter and the expert agree on the symmetry, then only two quantiles are necessary.

2.2 Student's t Prior

If symmetry is agreed on for Student's t prior, then three non-redundant quantiles are necessary to take into account the degrees of freedom ν. The elicitation by Kadane, et al (1980) requested 0.50, 0.75 and 0.9375 quantiles: $q_{0.50}$, $q_{0.75}$ and $q_{0.9375}$. The median $q_{0.50}$ can be taken as the location parameter μ and $a(x)$ is defined as $(q_{0.9375}$-$q_{0.50})/(q_{0.75}$-$q_{0.50})$. Since both the numerator and the denominator are independent of the center and the ratio is independent of the spread, thus $a(x)$ depends on degrees of freedom ν only (a monotonic function of ν) and ν is determined by checking the look-up table of ν vs. $a(x)$. For the spread of Student's t-distribution, after eliciting the degrees of freedom ν and obtaining the corresponding standard upper quartile $t_{\nu,.75}$, $S(q)$=$(q_{.75}-q_{.50})^2/t^2_{\nu,.75}$ is used to elicit the scale parameter σ. This idea can be applied to general location-scale-shape family (Sections 2.4, 2.5 and 2.6).

2.3 Log-normal Prior

Among the scarcity of literature on asymmetric prior elicitation, Garthwaite (1989) modeled asymmetry by a "split-normal" distribution, and O'Hagan (1998) used $\frac{1}{6}$, $\frac{3}{6}$ and $\frac{5}{6}$ quantiles to elicit a log-normal prior for positive skew distribution. For simplicity, we denote $q_{0.25}$, $q_{0.50}$ and $q_{0.75}$ to be the lower quartile, median and upper quartile of log-normal distribution, the following proposition provides parameter solution for log-normal case.

Proposition 1 If X has a log-normal distribution, i.e., $\ln X \sim N(\mu, \sigma^2)$, then the variance $D(X) = q^2_{0.50} r^2(r^2 - 1)$ and the mean $E(X) = rq_{0.50}$, where $q_{0.50} = e^{\mu}$ is the median of X, $r = \exp(\frac{\ln^2(q_{0.75}/q_{0.25})}{2Z_q^2})$, Z_q is the IQR for standard normal distribution.

The proof is done by observing $E(e^{t \ln X}) = E(X^t)$, followed by simplifications. The parameter elicitation is straightforward based on 0.25, 0.50 and 0.75 quantiles when $q_{0.75} - q_{0.50} > q_{0.50} - q_{0.25}$. A limitation of log-normal distribution is the strictly positive domain and the inherent constraint $\log q_{0.75}$-$\log q_{0.50}$=$\log q_{0.50}$-$\log q_{0.25}$, which is obviously too restrictive a rule for the expert to follow. Thus the log-normal discussion is trivial. We now study general asymmetric prior elicitation in the following sections, where special attention is paid to the shape parameter elicitation.

2.4 Skew-normal Prior

Definition 1: Skew-normal distribution We construct a skew-normal random variable X by $cZ^+ + \epsilon$ (which spans the whole real domain), where c is a constant, Z^+ is a folded normal random variable, e.g., the positive part of $N(0, \sigma_Z^2)$ random variable, ϵ is a normal random variable with mean μ and variance σ_ϵ^2, and Z^+ is independent of ϵ. The pdf of X is

$$f(x|c, \mu, \sigma_Z, \sigma_\epsilon) = \frac{2}{\sqrt{c^2\sigma_Z^2+\sigma_\epsilon^2}}\phi(\frac{x-\mu}{\sqrt{c^2\sigma_Z^2+\sigma_\epsilon^2}})\Phi(\frac{c\sigma_Z}{\sigma_\epsilon}\frac{(x-\mu)}{\sqrt{c^2\sigma_Z^2+\sigma_\epsilon^2}}),$$

where σ $(=\sqrt{c^2\sigma_Z^2+\sigma_\epsilon^2})$ is the scale parameter, λ=$(\frac{c\sigma_Z}{\sigma_\epsilon})$ is the skewness (shape) parameter, and μ is the location parameter. The derivation is given in Sahu, Dey and Branco (2003). When $c = 0$, we have an ordinary normal distribution. Under reparameterization $Y = \frac{X-\mu}{\sigma}$, we get the simplified density function of $f(y|\lambda) = 2\phi(y)\Phi(\lambda y)$ which is the skew-normal density function given by Azzalini (1985), with $E(Y) = \sqrt{\frac{2}{\pi}}\delta$ and $D(Y) = 1 - \frac{2}{\pi}\delta^2$, where $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. In the general case, $E(X) = \mu + \sigma\sqrt{\frac{2}{\pi}}\delta$ and $D(X) = \sigma^2(1 - \frac{2}{\pi}\delta^2)$.

Theorem 1 Assume $\int_{-\infty}^{Q_q(\lambda)} 2\phi(x)\Phi(\lambda x)dx = q$; where, q^{th} quantile $Q_q(\lambda)$ is determined by λ, the skewness parameter. Then (i) $Q_q(-\lambda) = -Q_{1-q}(\lambda)$; (ii) $\forall\lambda \in (-\infty, 0], q \in [0, 1]$, $Q'_q(\lambda) = \frac{\phi(\lambda Q_q(\lambda))}{(1+\lambda^2)\Phi(\lambda Q_q(\lambda))}$.

The proof is given in the appendix. We can see that, $\forall\lambda \in (-\infty, 0]$, $Q_{0.50}(-\lambda) = -Q_{0.50}(\lambda)$ and $Q'_{0.50}(-\lambda) = Q'_{0.50}(\lambda)$; the median of $cZ^+ + \epsilon$ is not equal to the median of cZ^+ plus the median of ϵ, since $Q_q(\lambda)$ is not a linear function of λ. The boundary condition q determines the individual solution from the nonlinear ordinary differential equation $Q'_q(\lambda) = \frac{\phi(\lambda Q_q(\lambda))}{(1+\lambda^2)\Phi(\lambda Q_q(\lambda))}$. Conditional on the skewness (shape) parameter λ, we use the notation IQRR(λ) to represent the inter-quartile range ratio of $(q_{0.75,\lambda} - q_{0.50,\lambda})$ to $(q_{0.50,\lambda} - q_{0.25,\lambda})$, a monotonic function of λ (Section 2.2). Presumably, the skewness (shape) parameter λ can be obtained by checking the look-up table of the ratio of IQRR(λ) to λ. However, the top-right panel of Figure 1 shows that, a small change of the IQRR(λ) leads to a large change of hyperparameter λ when λ is close to zero, say $|\lambda| < 1$, which is not our expectation. Thus "IQRR(λ) vs. λ" is not a good elicitation method for λ. To formalize this scenario, we define the sensitivity index (SI) as

$$\partial \text{ hyperparameter}/\partial \text{ elicitation input},$$

for which we expect a moderate magnitude ($\sim$1). For example, in the top-right panel of Figure 1, the elicitation input is IQRR and the hyperparameter is λ. The guideline is to work on "elicitation input vs. hyperparameter" curves with moderate derivative or sensitivity index (SI), say around 1. We first consider the regions in which the shape parameter is not sensitive ($|\lambda| > 1$) by the following one-way procedure, and the sensitive regions will be discussed on in Section 3. The rationale for choosing 1 as the λ critical value will be described later in Remark 1.

1. The skewness parameter λ can be obtained by the monotonic relationship between λ and IQRR(λ) (top-right, Figure 1).
2. The scale parameter σ can be obtained by $(q_{0.75,\lambda} - q_{0.50,\lambda})$ vs. λ plot (bottom-left, Figure 1): find out $(q_{0.75,\lambda} - q_{0.50,\lambda})$ at elicited λ, and the elicited σ is thus $(q_{0.75} - q_{0.50})/(q_{0.75,\lambda} - q_{0.50,\lambda})$. Or by $(q_{0.75,\lambda} - q_{0.25,\lambda})$ vs. λ plot (bottom-right, Figure 1): find out $(q_{0.75,\lambda} - q_{0.25,\lambda})$ at elicited λ, and the elicited σ is thus $(q_{0.75} - q_{0.25})/(q_{0.75,\lambda} - q_{0.25,\lambda})$.
3. The location parameter μ is simply $q_{0.25} - \sigma q_{0.25,\lambda}$, $q_{0.50} - \sigma q_{0.50,\lambda}$, or $q_{0.75} - \sigma q_{0.75,\lambda}$.

2.5 Skew Student's t Prior

Definition 2: Skew Student's *t* distribution We construct a skew Student's t random variable X by $cT^+ + \epsilon$ (which spans the whole real domain), where c is a constant, T^+ is a folded Student's t random variable, that is, the positive part of $T(0,1,\nu)$ random variable, ϵ is a Student's t random variable with mean μ and scale η, both of them have degrees of freedom ν, and T^+ is independent of ϵ. The pdf of X is

$$f(x|c,\mu,\eta,\nu) = \frac{2}{\sqrt{c^2+\eta^2}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\nu\pi)^{\frac{1}{2}}} (1 + \frac{(x-\mu)^2}{\nu(\eta^2+c^2)})^{-\frac{\nu+1}{2}} T_{1,\nu+1}[(\frac{\nu + \frac{(x-\mu)^2}{\eta^2+c^2}}{\nu+1})^{-\frac{1}{2}} \frac{c}{\eta} \frac{(x-\mu)}{\sqrt{\eta^2+c^2}}],$$

where σ $(=\sqrt{c^2+\eta^2})$ is the scale parameter, λ $(=\frac{c}{\eta})$ is the skewness (shape) parameter, μ is the location parameter, and ν is the degrees of freedom. The derivation is given in Sahu, Dey and Branco (2003). When $c = 0$, we have an ordinary Student's t-distribution. Under reparameterization of $Y = \frac{X-\mu}{\sigma}$, we get the simplified density function of

$$f(y|\lambda,\mu,\sigma,\nu) = 2\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})(\nu\pi)^{\frac{1}{2}}} (1 + \frac{y^2}{\nu})^{-\frac{\nu+1}{2}} T_{1,\nu+1}[(\frac{\nu+y^2}{\nu+1})^{-\frac{1}{2}} \lambda y].$$

A different attribute for skew Student's t-distribution is that, the nominal location-scale-shape family is conditional on the degrees of freedom ν by observing that ν cannot be combined with λ to form a composite shape parameter. From Figure 2, we note that, the inter-quartile information of skew Student's t-distribution highly depends on the degrees of freedom ν; when ν reaches a value to approach the limiting normal distribution, the sensitivity index (SI) in the top-right panel gets substantially larger than 1, as skew-normal does (Figure 1). Conditioning on ν, the elicitation procedure is the same as that in Section 2.4 when λ is not close to 0, say $|\lambda| > 1$, and the sensitive regions will be discussed on in Section 3.

2.6 Normal-exponential Prior

The following normal-exponential distribution is dated back to Aigner et al (1977).

Definition 3: Normal-exponential distribution A normal-exponential random variable X is constructed by $\lambda E_1 + \epsilon$ (which spans the whole real domain), where λ is a constant, E_1 is an exponential random variable with mean 1, ϵ is a normal random variable with mean μ and variance η^2, and E_1 is independent of ϵ. The pdf of X has closed form

$$\tfrac{\alpha}{\sigma}\exp(-\tfrac{\alpha^2}{2})\exp(-\alpha\tfrac{x-\delta}{\sigma})\Phi(\tfrac{x-\delta}{\sigma}).$$

The mean is $\delta + \sigma(\frac{1}{\alpha} - \alpha)$, where σ ($=\eta$) is the scale parameter, α ($=\frac{\eta}{\lambda}$) is the shape parameter and δ ($=\mu + \frac{\eta^2}{\lambda}$) is the location parameter. The proof is a matter of integration. The superficial skewness is produced by the normal part, other than the exponential part. As α approaches zero, the normal-exponential density becomes a flat degenerate density; as α exceeds two, the density becomes an approximate normal density (Figure 3), where leftward displacement induced by the combination of shape and scale parameters is observed. Although the curves are quite steep in the IQR vs. α plot, reasonable sensitivity index (SI) is reached after taking 5^{th} root (bottom-middle and bottom-right, Figure 4). The parameter elicitation procedure at $\alpha > 0.15$, where the IQRR is substantially away from the limit 1.7107 as α departs from 0, is similar to skew-normal case (Section 2.4). We observe that, the skewness parameter elicitation is sensitive for skew-normal and normal-exponential priors when the true value is close to zero, while it is less sensitive for skew Student's t case under small degrees of freedom.

A more rigorous mathematical statement to support the above elicitation procedures is that any three distinct quantiles uniquely determines the location-scale-shape parameter set. For two density functions belonging to the same location-scale-shape family with two sets of parameters, we have observed that one or two density function crossovers occurs if these two sets of parameters are not exactly equal to each other. If two distinct density functions have two identical sets of three quantiles, then at least four density function crossovers are required, since there is at most one common quantile between two consecutive crossovers, and there is no common quantile either less than the first crossover or greater than the last crossover. The open problem is that there are at most three density function crossovers under any two sets of different parameters.

2.7 Approximate Scale Parameter Elicitation

Now we propose an elicitation scheme based on approximation for Student's t-distribution. The expert provides three quantiles: $q_{0.25}$, $q_{0.50}$ and $q_{0.75}$, but is not sure about normality assumptions. Suppose $F_{\lambda,\mu,\sigma}$ is the cumulative density function of the distribution of expert, a member of location-scale-shape (μ-σ-λ) family, and p is the inter-quartile probability between $q_{0.25}$ and $q_{0.75}$, then by first order Taylor's expansion at $g(\star)$=μ+$\sigma g(\lambda)$, where $g(\lambda)$ is some characteristic point (median, mean, mode, etc) of the standardized density $f_{\lambda,0,1}$ with $\mu = 0$ and $\sigma = 1$. We have

$$\begin{array}{l} F_{\lambda,\mu,\sigma}(q_{0.75}) \approx F_{\lambda,\mu,\sigma}(g(\star)) + f_{\lambda,\mu,\sigma}(g(\star))(q_{0.75} - g(\star)) + \frac{1}{2} f^{(1)}_{\lambda,\mu,\sigma}(g(\star))(q_{0.75} - g(\star))^2 \cdots \\ F_{\lambda,\mu,\sigma}(q_{0.25}) \approx F_{\lambda,\mu,\sigma}(g(\star)) + f_{\lambda,\mu,\sigma}(g(\star))(q_{0.25} - g(\star)) + \frac{1}{2} f^{(1)}_{\lambda,\mu,\sigma}(g(\star))(q_{0.25} - g(\star))^2 \cdots \end{array} . \tag{2}$$

Consequently, if we take $\frac{1}{2} = F_{\lambda,\mu,\sigma}(q_{0.75}) - F_{\lambda,\mu,\sigma}(q_{0.25}) \approx f_{\lambda,\mu,\sigma}(g(\star))(q_{0.75} - q_{0.25})$ $= \frac{1}{\sigma} f_{\lambda,0,1}(g(\lambda))\mathrm{IQR}$, then a rough approximation for scale parameter elicitation is $2\mathrm{IQR} f_{\lambda,0,1}(g(\lambda))$.

Theorem 2 Under Taylor's expansion approximation (2), denote

$$\Delta = \sum_{k=1}^{\infty} \frac{1}{k+1} f^{(k)}_{\lambda,0,1}(g(\lambda))[(q_{0.75,\lambda} - g(\lambda))^{k+1} - (q_{0.25,\lambda} - g(\lambda))^{k+1}].$$

Then

(i) the relative error of the scale parameter elicitation by $2\mathrm{IQR} f_{\lambda,0,1}(g(\lambda))$ is $\frac{-2\Delta}{1-2\Delta}$, a function only associated with the shape parameter λ;

(ii) $\Delta = \frac{1}{2} - f_{\lambda,0,1}(g(\lambda))\mathrm{IQR}(\lambda)$, where $\mathrm{IQR}(\lambda)$ is the standardized IQR under $F_{\lambda,0,1}$.

The proof is given in the appendix. For $g(\lambda) =$ mode function $M(\lambda)$, mean function $E(\lambda)$ or median function $q_{0.50}(\lambda)$, we have the following observations:

Corollary 1 $\mathrm{IQR}(\lambda) = \frac{\frac{1}{2} - \Delta}{f_{\lambda,0,1}(g(\lambda))}$.

Corollary 2 Under skew-normal distribution (Section 2.4),

$$\lim_{\lambda \to 0} \sum_{k=1}^{\infty} \frac{1}{k+1} f^{(k)}_{\lambda,0,1}(g(\lambda))[(q_{0.75,\lambda} - g(\lambda))^{k+1} - (q_{0.25,\lambda} - g(\lambda))^{k+1}] = \frac{1}{2}(1 - \sqrt{\frac{2}{\pi}} Z_q).$$

Corollary 3 Under skew Student's t-distribution with degrees of freedom ν (Section 2.5), $\lim_{\lambda \to 0} \sum_{k=1}^{\infty} \frac{1}{k+1} f^{(k)}_{\lambda,0,1}(g(\lambda))[(q_{0.75,\lambda} - g(\lambda))^{k+1} - (q_{0.25,\lambda} - g(\lambda))^{k+1}] = \frac{1}{2}(1 - \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} t_{\nu,q})$, where $t_{\nu,q}$ is the standardized inter-quartile range for Student's t-distribution with ν degrees of freedom.

We point out that (2) seems to work for many flexible location-scale-shape classes (details later, in Propositions 3, 4). Suppose f is the probability density function of general Student's t-distribution with ν degrees of freedom (shape parameter) and inter-quantile range IQR, then we get the following approximation (after neglecting the error)

$$\frac{p}{\text{IQR}} = \frac{1}{\sqrt{\nu\pi}\sigma}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}(1+\frac{(q_{0.50}-\mu)^2}{\nu\sigma^2})^{-\frac{\nu+1}{2}}, \tag{3}$$

where μ is the location parameter and σ is the scale parameter. Then $\hat{\mu} = q_{0.50}$ is the estimate of the prior mean (location parameter), and the approximate estimate of σ can be obtained as

$$\hat{\sigma} = \frac{\text{IQR}}{p\sqrt{\nu\pi}}\frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})}. \tag{4}$$

Proposition 2 Scale-parameter elicitation is implemented by scaling the requested IQR by certain value, or in accordance with the underlying distribution assumption. Denote the inter-quantile range $T_{q,\nu}$ of Student's t-distribution with degrees of freedom ν to be $T_{0.75,\nu}$-$T_{0.25,\nu}$, then the ratio of the approximate scale parameter by (4) to the exact scale parameter from Student's t version of (1) is a monotonically decreasing function of the degrees of freedom ν with limit $Z_q\sqrt{\frac{2}{\pi}}$=1.0763 (Z_q=1.349).

We now describe an exploratory scale parameter elicitation for Student's t-distribution with degrees of freedom ν (shape parameter), the exact scale parameter is $\frac{\text{IQR}}{T_{q,\nu}}$, which goes to $\frac{\text{IQR}}{Z_q}$ as ν increases. By Stirling's formula, we can show that the approximately elicited scale parameter (4) goes to $\frac{\text{IQR}}{p\sqrt{2\pi}}$, where the inter-quartile probability p=0.50. Although (4) seems to be a rough approximation, from Figure 6 it turns out to be an efficient estimator retaining proportional high-fidelity (stable relative error) as degrees of freedom ν increases. Of the two relative error curves, one is between the approximate and the exact scale parameters both under Student's t-distribution (solid curve), the other is between the approximate scale parameter under normal distribution and the exact scale parameter under Student's t-distribution (dotted curve), both approach the limit ratio and the former one approaches the limit more quickly. Consequently, when the degrees of freedom ν is greater than 5, the approximation (4) under Student's t is realistic after scaling by 1.0763. However, the crude approximation under the normal assumption is worse for most values of degrees of freedom, say less than 20. Taylor's expansion based approximate scale parameter elicitation may be quite efficient in the sense of a proportional high-fidelity when the shape parameter (say degrees of freedom ν for Student's t-distribution) approaches certain limit (say positive infinity). This observation will guide us to elicit scale and shape parameters jointly in certain regions by means of flexible Taylor's expansion based iterations (Section 3). From exploratory point of view, we may potentially elicit the degrees of freedom reversely conditional on $\hat{\sigma}$, although intuitively they are independent parameters without any inherent association. The $\nu \sim \frac{p\hat{\sigma}}{\text{IQR}}$ plot derived from (4) is given in Figure 7, where $\frac{p\hat{\sigma}}{\text{IQR}}$ range of $\sim$(0.2, 0.396) gives sensitive ν (shape parameter) estimates in (0,10). In other words, other values will lead to the normal distribution. Theoretically, If we assume symmetry *a priori*, then eliciting degrees of freedom ν and the scale parameter σ concurrently is not possible, since these two parameters are mutually determined within a certain domain for σ (Figure 7).

For only scale-parameter model, no approximation is needed. For example, if f is the pdf of an exponential distribution with only scale parameter σ, then $\hat{\sigma}$ should be obtained by solving the equation

$$q_{0.50}\sigma - \ln\sigma = \ln(\frac{\mathrm{IQR}}{p}). \tag{5}$$

Although approximation (4) is shown to be an efficient scale elicitation for symmetric Student's t priors, a similar procedure may not work for some skewed cases. For the log-normal case in Proposition 1, assume we are given $q_{0.25}$, $q_{0.50}$ and $q_{0.75}$ for X which forms a symmetric normal distribution for $\ln(X)$ requiring $q_{0.50}{=}\sqrt{q_{0.25} \times q_{0.75}}$. We use first order Taylor's expansion to make a rough approximation: $p = \mathrm{IQR}f(q_{0.50})$, where f is the pdf of a log-normal distribution, then $\hat{\mu}{=}\ln(q_{0.50})$ and

$$p = \frac{\mathrm{IQR}}{q_{0.50}\sqrt{2\pi}\sigma}, \qquad \text{thus} \qquad \hat{\sigma} = \frac{\mathrm{IQR}}{q_{0.50}\sqrt{2\pi}p}. \tag{6}$$

The exact σ elicitation under log-transformation is $\frac{\Delta}{Z_q}$, where Δ is $\ln(\frac{q_{0.75}}{q_{0.25}})$, and (6) becomes $\sqrt{\frac{2}{\pi}}(e^{\Delta} - e^{-\Delta})$. The ratio of the approximate and exact scale elicitation is thus approximately $\sqrt{\frac{2}{\pi}}Z_q(1+\frac{\Delta^2}{6})$ with quadratic increment of Δ. The weakness of Taylor's expansion may be due to the fact that, log-normal distribution is not a location-scale family member.

For the skew-normal case (Section 2.4), the exact scale elicitation is $\frac{\mathrm{IQR}}{\mathrm{IQR}(\lambda)}$, where $\mathrm{IQR}(\lambda)$ is the IQR for the standard skew-normal distribution with location parameter 0, scale parameter 1 and skewness (shape) parameter λ.

- ▷ If we apply Taylor's expansion at the location parameter μ, then $\hat{\sigma} = \sqrt{\frac{2}{\pi}}\mathrm{IQR}$, thus location based expansion makes use of no skewness information. The ratio of the approximate to the exact scale elicitation is $\sqrt{\frac{2}{\pi}}\mathrm{IQR}(\lambda)$.

- ▷ If we apply Taylor's expansion at mean $E_\lambda = \sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$, then $\hat{\sigma}$ is $[2\phi(E_\lambda)\Phi(\lambda E_\lambda)]$ IQR/p, where p=0.50 and E_λ is the mean of the standardized skew-normal distribution with location parameter 0, scale parameter 1 and skewness parameter λ. The ratio of the approximate to the exact scale elicitation is $\mathrm{IQR}(\lambda)[2\phi(E_\lambda)\Phi(\lambda E_\lambda)]/p$.

- ▷ If we apply Taylor's expansion at mode M_λ, then $\hat{\sigma}$ is IQR$[2\phi(M_\lambda)\Phi(\lambda M_\lambda)]/p$, where p=0.50 and M_λ is the mode of the standardized skew-normal distribution with location parameter 0, scale parameter 1 and skewness parameter λ. The ratio of the approximate to the exact scale elicitation is $[2\phi(M_\lambda)\Phi(\lambda M_\lambda)]$ IQR(λ)/p.

- ▷ If we apply Taylor's expansion at median $q_{0.50,\lambda}$, then $\hat{\sigma}$ is $[2\phi(q_{0.50,\lambda})\Phi(\lambda q_{0.50,\lambda})]$ IQR/p, where p=0.50 and $q_{0.50,\lambda}$ is the median of the standardized skew-normal distribution with location parameter 0, scale parameter 1 and skewness parameter λ. The ratio of the approximate to the exact scale elicitation is $[2\phi(q_{0.50,\lambda})\Phi(\lambda q_{0.50,\lambda})]$ IQR(λ)/p.

Remark 1 The scale parameter σ elicitation based on (2) expanded at the location parameter μ leads to identical results under naive normal or skew-normal distributions. If we ignore the skewness completely by assuming naive normal, then the elicited scale parameter σ is $\frac{\text{IQR}}{Z_q}$ accordingly. All of these options are compared by a "whiskers" plot (Figure 8). Except for the native normal based approximation, all skew-normal based approximations are visually undistinguishable as horizontal segments expanding from ∼-1 to ∼1, which is exactly where λ is highly sensitive to the IQRR(λ) (the top-right panel of Figure 1), also represented by a horizontal segment expanding from ∼-1 to ∼1. Thus the sensitive region and the critical value of λ are to be identified visually. Overall, the performance ordering may be Taylor's expansion at: median, mean, mode, and location. We envision that this observation also works for skew Student's t conditional on the degrees of freedom ν. Thus quantile (probability) based prior elicitation could increase the potential flexibility and efficiency.

Proposition 3 For the skew-normal distribution, the ratio of the approximate scale parameter to the exact scale parameter goes to $\sqrt{\frac{2}{\pi}}Z_q$=1.0763 as λ approaches 0.

For the normal-exponential case (Section 2.6), the exact scale elicitation is $\frac{\text{IQR}}{\text{IQR}_{(\alpha)}}$, where α is the shape parameter, and IQR(α) is the IQR for the standard normal-exponential distribution with location parameter 0 and scale parameter 1. Again, p=0.50.

- ▷ If we apply Taylor's expansion at the location parameter δ, then $\hat{\sigma}=\frac{\alpha}{2p}\exp(-\frac{\alpha^2}{2})$IQR, in contrast to skew-normal case, location based expansion incorporates the shape information. The ratio of the approximate to the exact scale elicitation is $\frac{2p}{\alpha}\exp(\frac{\alpha^2}{2})$ IQR(α).

- ▷ If we apply Taylor's expansion at mean $E_\alpha=\sigma(\frac{1}{\alpha}-\alpha)$, then $\hat{\sigma} = \frac{\alpha}{p}\exp(\frac{\alpha^2}{2} - 1)\Phi(\frac{1}{\alpha}-\alpha)$IQR. The ratio of the approximate to the exact scale elicitation is $\frac{\alpha}{p}\exp(\frac{\alpha^2}{2}-1)\Phi(\frac{1}{\alpha}-\alpha)IQR(\alpha)$.

- ▷ If we apply Taylor's expansion at mode M_α which does not have an explicit formula, then $\hat{\sigma} = \frac{\alpha}{p}\exp[-(\frac{\alpha^2}{2}+\alpha M_\alpha)]\Phi(M_\alpha)$IQR. The ratio of the approximate to the exact scale elicitation is $\frac{\alpha}{p}\exp[-(\frac{\alpha^2}{2}+\alpha M_\alpha)]\Phi(M_\alpha)\text{IQR}(\alpha)$.

▷ If we apply Taylor's expansion at median $q_{0.50,\alpha}$, then $\hat{\sigma}$ is $\frac{\alpha}{p}\exp[-(\frac{\alpha^2}{2}+\alpha q_{0.50,\alpha})]$ $\Phi(q_{0.50,\alpha})$IQR. The ratio of the approximate to the exact scale elicitation is $\frac{\alpha}{p}\exp[-(\frac{\alpha^2}{2}+\alpha q_{0.50,\alpha})]\Phi(q_{0.50,\alpha})\text{IQR}(\alpha)$.

Remark 2 The comparisons are given by a "whiskers" plot (Figure 9), where the location or median based approximation performs more stably and retains more proportional high-fidelity (≈1.09) than others at small α. Note that the location based approximation takes into account the shape information, which is a different attribute from skew-normal and skew Student's t. The performance ordering may be Taylor's expansion at: median, location, mean and mode at small α; at large α, the location based approximation cannot be used since the shape and scale parameters jointly induces large location displacement, and all other approximations have visually undistinguishable performances.

Proposition 4 For normal-exponential distribution, the ratio of the approximate scale to the exact scale approaches 1.0763 as α approaches positive infinity.

3 Elicitation on Shape-Parameter-Sensitive Regions

3.1 Skew-normal case

Complementary to the discussion on skew-normal in Section 2.4, we propose several iterative algorithms for $|\lambda| \leq 1$ under alternative moderate sensitivity index (∼1). If the original procedure in Section 2.4 is followed, then λ will be highly sensitive to minor changes of IQRR(λ); while the approximate σ elicitation based on first order Taylor's expansion at mean (with a favorable explicit expression, Section 2.4) retains proportional ($\sqrt{\frac{2}{\pi}}Z_q$) high-fidelity (the middle panel of Figure 8) for that very λ range.

◇ **Iteration Based on Taylor's Expansion at Median**

1. Start with current λ, apply the first order Taylor's expansion at median $q_{0.50,\lambda}$ with location parameter 0 and scale parameter 1. Then $\hat{\sigma}$ is $\text{IQR}[2\phi(q_{0.50,\lambda})\Phi(\lambda q_{0.50,\lambda})]/p$, the elicited scale parameter is $\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q}$, and the standardized ($q_{0.75,\lambda}$-$q_{0.50,\lambda}$) is ($q_{0.75}$-$q_{0.50}$)$/(\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q})$.

2. The skewness parameter λ can be obtained by $(q_{0.75,\lambda}-q_{0.50,\lambda})$ vs. λ plot (bottom-left, Figure 1) under moderate sensitivity index (∼1). Note that, one-to-two matching around small λ is not an issue, since the sign of λ is correctly kept.

3. Go to 1 until convergence (complete λ and σ).

4. The location parameter μ is simply $q_{0.25}-\sigma q_{0.25,\lambda}$, $q_{0.50}-\sigma q_{0.50,\lambda}$ or $q_{0.75}-\sigma q_{0.75,\lambda}$.

◇ **Iteration Based on Taylor's Expansion at Mean**

1. Start with current λ, apply the first order Taylor's expansion at mean $E_\lambda=\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$ with location parameter 0 and scale parameter 1. Then $\hat{\sigma}$ is IQR$[2\phi(E_\lambda)\Phi(\lambda E_\lambda)]/p$, the elicited scale parameter is $\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q}$, and the standardized ($q_{0.75,\lambda}$-$q_{0.50,\lambda}$) is ($q_{0.75}$-$q_{0.50}$)/($\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q}$).
2. The skewness parameter λ can be obtained by $(q_{0.75,\lambda}-q_{0.50,\lambda})$ vs. λ plot (bottom-left, Figure 1) under moderate sensitivity index (∼1). Note that, one-to-two matching around small λ is not an issue, since the sign of λ is correctly kept.
3. Go to 1 until convergence (complete λ and σ).
4. The location parameter μ is simply $q_{0.25}-\sigma q_{0.25,\lambda}$, $q_{0.50}-\sigma q_{0.50,\lambda}$ or $q_{0.75}-\sigma q_{0.75,\lambda}$.

◇ **Iteration Based on Taylor's Expansion at Mode**

1. Start with current λ, apply the first order Taylor's expansion at mode $M_\lambda=\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$ with location parameter 0 and scale parameter 1. Then $\hat{\sigma}$ is IQR$[2\phi(E_\lambda)\Phi(\lambda E_\lambda)]/p$, the elicited scale parameter is $\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q}$, and the standardized ($q_{0.75,\lambda}$-$q_{0.50,\lambda}$) is ($q_{0.75}$-$q_{0.50}$)/($\frac{\hat{\sigma}}{\sqrt{\frac{2}{\pi}}Z_q}$).
2. The skewness parameter λ can be obtained by $(q_{0.75,\lambda}-q_{0.50,\lambda})$ vs. λ plot (bottom-left, Figure 1) under moderate sensitivity index (∼1). Note that, one-to-two matching around small λ is not an issue, since the sign of λ is correctly kept.
3. Go to 1 until convergence (complete λ and σ).
4. The location parameter μ is simply $q_{0.25}-\sigma q_{0.25,\lambda}$, $q_{0.50}-\sigma q_{0.50,\lambda}$ or $q_{0.75}-\sigma q_{0.75,\lambda}$.

◇ **Iteration Based on the Inter-quartile Ranges**

1. Start with current λ, $(q_{0.75,\lambda}-q_{0.25,\lambda})$ vs. λ plot (Figure 1) with moderate sensitivity index (∼1). The elicited σ is $(q_{0.75}-q_{0.25})/(q_{0.75,\lambda}-q_{0.25,\lambda})$.
2. The skewness parameter λ can be obtained by $(q_{0.75,\lambda}-q_{0.50,\lambda})$ vs. λ plot (bottom-left, Figure 1) under moderate sensitivity index (∼1).
3. Go to 1 until convergence (complete λ and σ).
4. The location parameter μ is simply $q_{0.25}-\sigma q_{0.25,\lambda}$, $q_{0.50}-\sigma q_{0.50,\lambda}$ or $q_{0.75}-\sigma q_{0.75,\lambda}$.

Remark 3 We create the ($q_{0.25,\lambda}$, $q_{0.50,\lambda}$, $q_{0.75,\lambda}$) vs. λ look-up table (the top-left panel of Figure 1), where the λ step is 0.01, the iteration involves obtaining the optimal λ given $(q_{0.75,\lambda} - q_{0.50,\lambda})$ and the optimal $(q_{0.75,\lambda} - q_{0.25,\lambda})$ given λ, till convergence. The one-way elicitation of λ followed by σ (Section 2.4) is replaced by a two-way joint λ and σ elicitation. The elicitation comparison is given in Table 1, and Figures 11 through 14, where ratios are calculated at all declared 50% quantiles (from 0.01 to 0.20 with step 0.01). The IQR, mean and median based Taylor's expansion iterations are consistent, while mode based Taylor's expansion iteration is likely to be less satisfactory. Compared to the IQR based iteration, we recommend the explicit mean based Taylor's expansion in view of numerical stability. In the top-right panel of Figure 1, when the IQRR varies within a very short range around 1.0, the λ ranges among [-1,1], leading to a sensitivity index with a huge magnitude. Under iterative elicitation procedures, highly sensitive regions are avoided, and the applied regions in the IQR(λ) vs. λ plots (the bottom-left and bottom-right panels of Figure 1) show a workable sensitivity index ($\sim$1). By observing that, the standard mean $\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$ has a derivative $\sqrt{\frac{2}{\pi}}(1+\lambda^2)^{-\frac{3}{2}}$ ($\sim$1 at small λ) and conclusion (ii) in Theorem 2, we conclude that, iteration by Taylor's expansion at mean will work well under moderate sensitivity index. Generally, ∂shape parameter/∂standardized IQR needs to be moderate in order to apply the iterative algorithms, similar procedures will apply to the normal-exponential prior and skew Student's t prior conditional on the degrees of freedom ν. Because the standard IQRs are tabulated by numerical integration which is subject to accuracy evaluation for irregular skewed classes, we prefer an explicit expression involved in the algorithm, such as mean based Taylor's expansion. One piece of evidence is seen in Table 1, where there is a switching between two elicited λ's at $q_{0.50}$=0.07 and $q_{0.50}$=0.08.

3.2 Gamma Case

Gamma distribution can be generalized into a location-scale-shape family member as $f_{\alpha,\beta,\mu} = \frac{1}{\beta}\frac{1}{\Gamma(\alpha)}(\frac{x-\mu}{\beta})^{\alpha-1}e^{-\frac{x-\mu}{\beta}}1_{[x\geq\mu]}$, whose quantile plots are given in Figure 5. We can find that the sensitivity index is around 100 by using $(q_{0.75,\alpha} - q_{0.50,\alpha})/(q_{0.50,\alpha} - q_{0.25,\alpha})$ vs. α plot when $\alpha \geq$1.5, and the sensitivity index is decreased to less than 10 by using $(q_{0.75,\alpha} - q_{0.50,\alpha})$ or $(q_{0.50,\alpha} - q_{0.25,\alpha})$ vs. α plot. For $\alpha \geq$1.5, we propose the following iteration procedure.

Iteration Based on the Inter-quartile Ranges

1. Start with current α, $(q_{0.75,\alpha} - q_{0.50,\alpha})$ vs. α plot (Figure 5) with sensitivity index $\sim$10, the elicited β is $(q_{0.75} - q_{0.50})/(q_{0.75,\alpha} - q_{0.50,\alpha})$.
2. The shape parameter α can be obtained by $(q_{0.50,\alpha} - q_{0.25,\alpha})$ vs. α plot, where $(q_{0.50,\alpha} - q_{0.25,\alpha})$=$(q_{0.50} - q_{0.25})/\beta$.
3. Repeat 1-2 until convergence (complete α and β).
4. The location parameter μ is simply $q_{0.25} - \beta q_{0.25,\alpha}$, $q_{0.50} - \beta q_{0.50,\alpha}$ or $q_{0.75} - \beta q_{0.75,\alpha}$.

Remark 4 IMSL library is used to call the *gamin* function to return the standard quantiles for arbitrary α (Table 2). We create the look-up table of ($q_{0.25,\alpha}$, $q_{0.50,\alpha}$, $q_{0.75,\alpha}$) vs. α, where the α step is 0.001, the iteration alternates between finding the optimal α given $(q_{0.50,\alpha} - q_{0.25,\alpha})$ and finding the optimal $(q_{0.75,\alpha} - q_{0.50,\alpha})$ given α from the look-up table (the lower panels of Figure 5), till convergence. The one-way elicitation of α followed by β (the top-right panel of Figure 5) is replaced by a two-way joint elicitation of α and β. The comparison is given in Table 2. where the numerical difference is substantially greater than 0.001 (α step). One-way elicitation introduced in Section 2.4 only works well for those regions where the shape parameter is not sensitive, only two-way iterating elicitation works on regions where the shape parameter is highly sensitive, including the IQR based iteration and Taylor's expansion based iterations. Unfortunately, we also observe those regions where both elicitations encounter high sensitivity index (SI), such as the tail regions (λ approaches 10) in the lower panels of Figure 1.

4 Interactive Graphical Display of Prior Probability Distribution

In this section, we propose visual aids to help elicit skewed priors and behavior of the tail shape with an emphasis on the degrees of freedom ν of skew Student's t priors. This section is a follow-up of preceding sections to streamline the "sketching" principle recommended by Berger (1985, Chapter 3).

Trial on Degrees of Freedom

A. We consider $\nu = 1, 5, 10, 15, 30, \ldots$ to apply the general procedure in Section 2.5, with an exact or an approximate scale elicitation, to select the most agreeable ν for skewed Student's t prior by repeatedly drawing distribution shapes.

B. We start with a tentative ν out of (1, 5, 10, 15, 30,...) to obtain $\hat{\sigma}$ in an exact way, then immediately update the degrees of freedom for Student's t-distribution by reverse of (4) and show the distribution shape to the expert in order to make an exploratory effort to reach an agreement on ν.

Other Cases

C. The expert provides his/her opinion in the form of a graph, then we estimate $q_{0.25}$, $q_{0.50}$ and $q_{0.75}$ directly from the graph. If the graph is close to symmetric normal we use (1); if it is close to log-normal with a positive domain, we apply Proposition 1; if it is close to symmetric heavier tail, we go to option [B].

D. For generalized symmetric Student's t-distribution with two degrees of freedom, the graphical approach seems more necessary and similar procedures apply.

E. For skew-normal and normal-exponential priors, the aforementioned general procedures apply (Sections 2.4, 2.6 and 3), we need to redraw distribution shape for expert's evaluation.

Now we briefly explore the shapes of generalized Student's t-distribution. Generalized Student's t-distribution is a special case of a scale mixture of normal distribution. It is obtained as $X|\lambda \sim N(\mu, \lambda/\sigma^2)$ and $\lambda \sim \text{Gamma}(\nu_1/2, \nu_2/2)$, the marginal distribution produces the generalized t, which has the form $f(x) = \frac{1}{\sqrt{\pi\nu_2}} \frac{\Gamma(\frac{\nu_1+1}{2})}{\sigma\Gamma(\frac{\nu_1}{2})}(1 + \frac{(x-\mu)^2}{\nu_2\sigma^2})^{-\frac{\nu_1+1}{2}}$. When $\nu_1 = \nu_2$ we get back to the usual Student's t-distribution. The two degrees of freedom parameters make the distribution extremely flexible keeping the symmetry intact. For example, t(30,1) and t(1,30), t(10,1) and t(1,10) densities are shown in Figure 10. The tail could be even heavier than Cauchy. For prior elicitation, this distribution can be used for any symmetric specification, similar to Section 2, the realistic approximate algorithm will be to set $\frac{p}{\text{IQR}}$ as $\frac{1}{\sqrt{\pi\nu_2}} \frac{\Gamma(\frac{\nu_1+1}{2})}{\sigma\Gamma(\frac{\nu_1}{2})}(1 + \frac{(q_{0.50}-\mu)^2}{\nu_2\sigma^2})^{-\frac{\nu_1+1}{2}}$, thus

$$\hat{\mu} = q_{0.50} \quad \text{and} \quad \hat{\sigma} = \frac{\text{IQR}}{p} \frac{1}{\sqrt{\pi\nu_2}} \frac{\Gamma(\frac{\nu_1+1}{2})}{\Gamma(\frac{\nu_1}{2})}. \tag{7}$$

Note that the only difference between (4) and (7) is the presence of ν_2 in the denominator, a very minor change in the code.

5 Concluding Remarks

We demonstrate that, for those univariate skewed priors belonging to location-scale-shape families, the elicitation procedure on the non-sensitive regions is: shape parameter followed by the scale and location parameters. Moreover, we take the initiatives to quantitatively study technical details for flexible direct prior elicitation, which will streamline the induced interactive graphical procedures. The quantile based Taylor's expansion approximation was interestingly found to enjoy sufficient efficiency to pinpoint the scale parameter on shape-parameter sensitive regions, when the shape information is incorporated for approximation. The exact procedure is in the sense of numerical computation, and the iterative elicitation is numerically tested to be satisfactorily stable and efficient. Prior elicitation from multiple expert opinions could be done by averaging these summaries (Kadane, et al, 1980) or applying linear regression to take into account multiple experts (Gill and Walker, 2005). Throughout we assumed that experts provide a unimodal distribution, which is realistic in practice. However, if the experts believe that the distribution could be multimodal then we have to consider discrete mixtures of unimodal distributions. If the prior predictive space is necessary and a well-established response model is available, then expert opinions on the response distribution may be more practical. But we can envision that the calculation load will be much more intensive under skewed priors. How to develop efficient prior elicitation algorithm under that circumstance is another area of research.

Supplemental Material: Proofs, Figures and Results

Proof of Theorem 1

Proof of part 1. From $2\int_{-\infty}^{Q_q(\lambda)} \phi(x)\Phi(\lambda x)dx = q$, we have $2\int_{Q_q(\lambda)}^{+\infty} \phi(x)\Phi(\lambda x)dx = 1-q$ and $2\int_{+\infty}^{Q_q(\lambda)} -\phi(-x)\Phi(\lambda x)dx = 1-q$. Substituting $y=-x$, it follows that $2\int_{-\infty}^{-Q_q(\lambda)} \phi(y)\Phi(-\lambda y)dy = 1-q$, which proves part 1.

Proof of part 2.

$$\begin{aligned}
& \int_{-\infty}^{Q_q(\lambda+\Delta_\lambda)} 2\phi(x)\Phi((\lambda+\Delta_\lambda)x)dx \\
= & \int_{-\infty}^{Q_q(\lambda)} 2\phi(x)\Phi((\lambda+\Delta_\lambda)x)dx + \int_{Q_q(\lambda)}^{Q_q(\lambda+\Delta_\lambda)} 2\phi(x)\Phi((\lambda+\Delta_\lambda)x)dx \\
= & \int_{-\infty}^{Q_q(\lambda)} 2\phi(x)\Phi((\lambda+\Delta_\lambda)x)dx \\
& +(Q_q(\lambda+\Delta_\lambda)-Q_q(\lambda))2\phi(pQ_q(\lambda)+(1-p)Q_q(\lambda+\Delta_\lambda)) \\
& \times\Phi((\lambda+\Delta_\lambda)(pQ_q(\lambda)+(1-p)Q_q(\lambda+\Delta_\lambda))), (\exists p\in[0,1]) \\
= & \int_{-\infty}^{Q_q(\lambda)} 2\phi(x)\Phi(\lambda x)dx = q.
\end{aligned} \tag{8}$$

Thus,

$$\begin{aligned}
& \int_{-\infty}^{Q_q(\lambda)} 2x\phi(x)\left(\tfrac{\Phi((\lambda+\Delta_\lambda)x)-\Phi(\lambda x)}{x\Delta_\lambda}\right)dx + \tfrac{Q_q(\lambda+\Delta_\lambda)-Q_q(\lambda)}{\Delta_\lambda} \\
& \times 2\phi(pQ_q(\lambda)+(1-p)Q_q(\lambda+\Delta_\lambda)) \\
& \times\Phi((\lambda+\Delta_\lambda)(pQ_q(\lambda)+(1-p)Q_q(\lambda+\Delta_\lambda))) \\
& = 0.
\end{aligned} \tag{9}$$

Letting Δ_λ approach 0, we get $\int_{-\infty}^{Q_q(\lambda)} 2x\phi(x)\phi(\lambda x)dx + Q'_q(\lambda)2\phi(Q_q(\lambda))\Phi(\lambda Q_q(\lambda)) = 0$. The final result follows after some simplifications. Our result shows that $Q_q(\lambda)$ is differentiable. The proof ends.

Proof of Theorem 2

Obviously, $g(\star)$ retains linearity of the standardized characteristic point by $g(\star) = \mu + \sigma g(\lambda)$, where $g(\star)$ is the characteristic point of general location-scale-shape family and $g(\lambda)$ is the characteristic point of the standard location-scale-shape family with location 0 and scale 1. Thus, (2) becomes

$$\begin{aligned}
F_{\lambda,\mu,\sigma}(q_{0.75}) &= F_{\lambda,0,1}(g(\lambda)) + \tfrac{1}{\sigma}f_{\lambda,0,1}(g(\lambda))(q_{0.75}-g(\star)) + \tfrac{1}{2}f^{(1)}_{\lambda,0,1}(g(\lambda))(q_{0.75,\lambda}-g(\lambda))^2\cdots \\
F_{\lambda,\mu,\sigma}(q_{0.25}) &= F_{\lambda,0,1}(g(\lambda)) + \tfrac{1}{\sigma}f_{\lambda,0,1}(g(\lambda))(q_{0.25}-g(\star)) + \tfrac{1}{2}f^{(1)}_{\lambda,0,1}(g(\lambda))(q_{0.25,\lambda}-g(\lambda))^2\cdots
\end{aligned}. \tag{10}$$

Consequently,

$$\tfrac{1}{2} = \tfrac{1}{\sigma} f_{\lambda,0,1}(g(\lambda))\mathrm{IQR} + \textstyle\sum_{k=1}^{\infty} \tfrac{1}{k+1} f^{(k)}_{\lambda,0,1}(g(\lambda))[(q_{0.75,\lambda} - g(\lambda))^2 - (q_{0.25,\lambda} - g(\lambda))^2].$$

Denote $\Delta = \sum_{k=1}^{\infty} \frac{1}{k+1} f^{(k)}_{\lambda,0,1}(g(\lambda))[(q_{0.75,\lambda} - g(\lambda))^2 - (q_{0.25,\lambda} - g(\lambda))^2]$, the relative error for scale parameter elicitation is $(\frac{f^{(k)}_{\lambda,0,1}(g(\lambda))}{0.50} - \frac{f^{(k)}_{\lambda,0,1}(g(\lambda))}{0.50-\Delta})/(\frac{f^{(k)}_{\lambda,0,1}(g(\lambda))}{0.50-\Delta}) = \frac{-2\Delta}{1-2\Delta}$. The second statement in Theorem 2 comes from the following fact: the true scale parameter σ is equal to $\dfrac{\mathrm{IQR}}{\mathrm{IQR}_{(\lambda)}}$ by direct numerical calculation from the look-up table $\mathrm{IQR}(\lambda)$ vs. λ, or $\frac{\mathrm{IQR} f_{\lambda,0,1}(g(\lambda))}{0.50-\Delta}$ by keeping the first term from the fully Taylor's expansion. The proof ends.

Table 1: Iteration Comparison: Inter-quartile range (IQR), Taylor's expansion at median, mean and mode (skew-normal: $q_{0.25} = -10$, $q_{0.75} = 10$)

$q_{0.50}$	λ				σ				μ			
	IQR	median	mean	mode	IQR	median	mean	mode	IQR	median	mean	mode
0.01	-0.23	-0.23	-0.23	-0.23	15.06	15.07	15.07	15.07	2.69	2.69	2.69	2.69
0.02	-0.37	-0.32	-0.32	-0.32	15.42	15.29	15.29	15.29	4.26	3.70	3.70	3.70
0.03	-0.41	-0.41	-0.41	-0.41	15.56	15.56	15.56	15.56	4.69	4.69	4.69	4.69
0.04	-0.53	-0.54	-0.54	-0.54	15.99	16.04	16.04	16.04	5.93	6.04	6.04	6.04
0.05	-0.54	-0.55	-0.55	-0.55	16.03	16.08	16.08	16.08	6.04	6.15	6.15	6.15
0.06	-0.64	-0.56	-0.56	-0.56	16.44	16.12	16.12	16.12	7.00	6.24	6.23	6.24
0.07	-0.70	-0.67	-0.69	-0.67	16.70	16.58	16.66	16.58	7.56	7.30	7.47	7.30
0.08	-0.68	-0.68	-0.68	-0.71	16.62	16.62	16.62	16.76	7.39	7.39	7.39	7.67
0.09	-0.71	-0.72	-0.72	-0.72	16.74	16.80	16.80	16.80	7.65	7.75	7.74	7.75
0.10	-0.77	-0.82	-0.77	-0.85	17.01	17.25	17.02	17.40	8.17	8.62	8.18	8.88
0.11	-0.82	-0.85	-0.85	-0.86	17.23	17.39	17.38	17.44	8.60	8.87	8.86	8.95
0.12	-0.85	-0.86	-0.86	-0.89	17.37	17.43	17.42	17.58	8.85	8.94	8.93	9.21
0.13	-0.88	-0.89	-0.89	-0.92	17.50	17.57	17.56	17.72	9.08	9.19	9.18	9.43
0.14	-0.89	-0.92	-0.91	-0.95	17.55	17.71	17.65	17.86	9.17	9.41	9.33	9.67
0.15	-0.94	-0.95	-0.92	-1.01	17.78	17.85	17.69	18.14	9.55	9.65	9.40	10.12
0.16	-0.95	-1.01	-0.95	-1.03	17.83	18.12	17.83	18.23	9.64	10.10	9.63	10.26
0.17	-1.01	-1.03	-1.01	-1.04	18.09	18.21	18.09	18.28	10.07	10.24	10.07	10.33
0.18	-1.01	-1.04	-1.01	-1.11	18.09	18.25	18.09	18.60	10.07	10.30	10.07	10.82
0.19	-1.04	-1.09	-1.04	-1.24	18.24	18.48	18.22	19.18	10.29	10.65	10.27	11.66
0.20	-1.09	-1.11	-1.04	-1.26	18.46	18.57	18.22	19.27	10.62	10.78	10.27	11.80

Table 2: Shape parameter elicitation comparison: iterative inter-quartile range algorithm (Gamma: $q_{0.25} = -2.0$, $q_{0.75} = 3.0$)

$q_{0.50}$	α (iterative)	α (IQR ratio)	$q_{0.50}$	α (iterative)	α (IQR ratio)
0.02	1.655	1.652	0.22	4.329	4.324
0.04	1.778	1.775	0.24	4.975	4.969
0.06	1.919	1.916	0.26	5.775	5.782
0.08	2.080	2.076	0.28	6.818	6.825
0.10	2.266	2.261	0.30	8.189	8.197
0.12	2.481	2.476	0.32	10.040	10.049
0.14	2.711	2.728	0.34	12.631	12.640
0.16	3.029	3.024	0.36	16.408	16.418
0.18	3.382	3.378	0.38	22.226	22.244
0.20	3.809	3.804	0.40	31.864	31.899

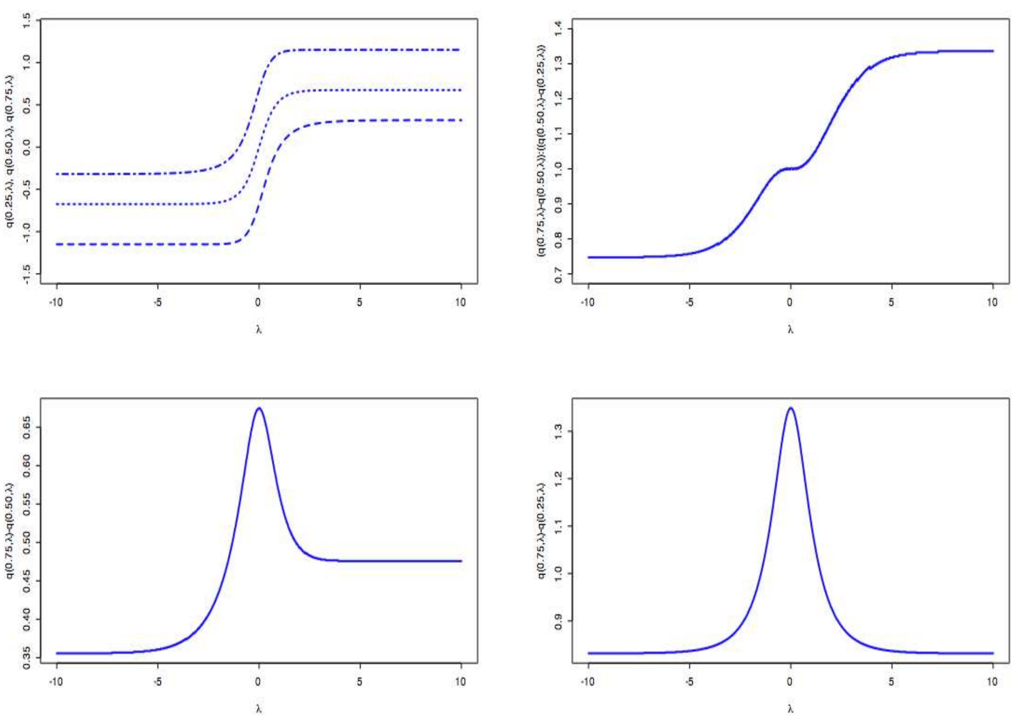

Figure 1: Relationship between quantiles and skewness parameter λ for skew-normal distribution. (Top-left: $(q_{0.75,\lambda},\ q_{0.50,\lambda},\ q_{0.25,\lambda})$ vs. λ; top-right: $(\frac{q_{0.75,\lambda}-q_{0.50,\lambda}}{q_{0.50,\lambda}-q_{0.25,\lambda}})$ vs. λ; bottom-left: $(q_{0.75,\lambda}-q_{0.50,\lambda})$ vs. λ; bottom-right: $(q_{0.75,\lambda}-q_{0.25,\lambda})$ vs. λ)

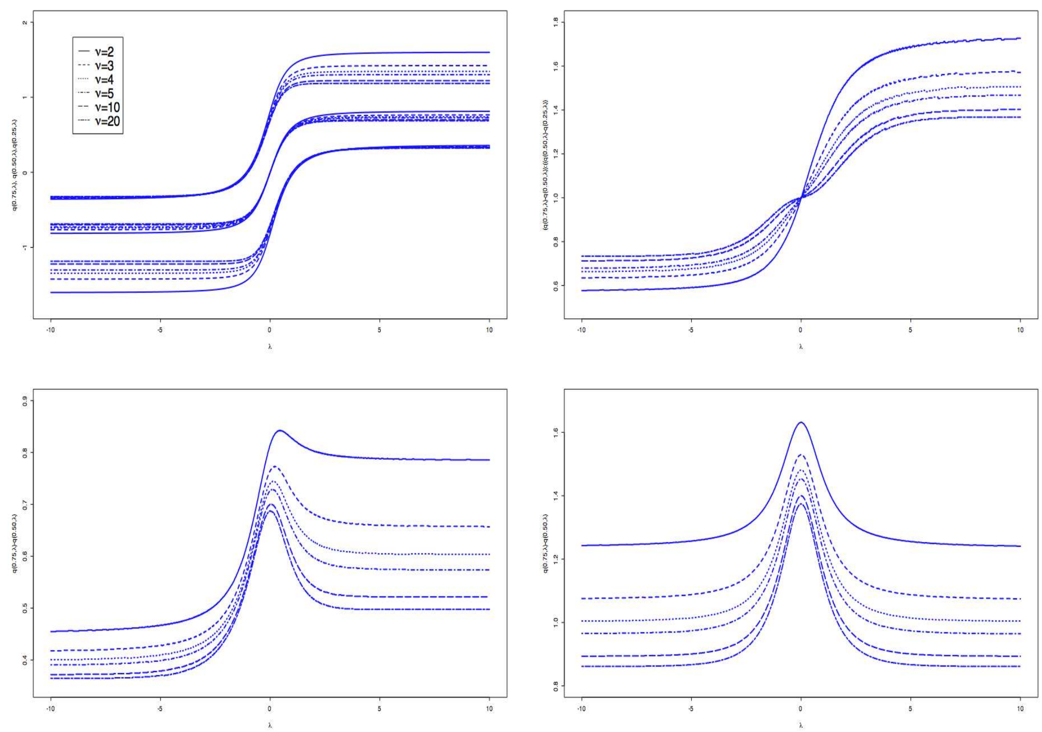

Figure 2: Relationship between quantiles and skewness parameter λ for skew Student's t distribution. (Top-left: $(q_{0.75,\lambda},\, q_{0.50,\lambda},\, q_{0.25,\lambda})$ vs. λ; top-right: $(\frac{q_{0.75,\lambda}-q_{0.50,\lambda}}{q_{0.50,\lambda}-q_{0.25,\lambda}})$ vs. λ; bottom-left: $(q_{0.75,\lambda} - q_{0.50,\lambda})$ vs. λ; bottom-right: $(q_{0.75,\lambda} - q_{0.25,\lambda})$ vs. λ)

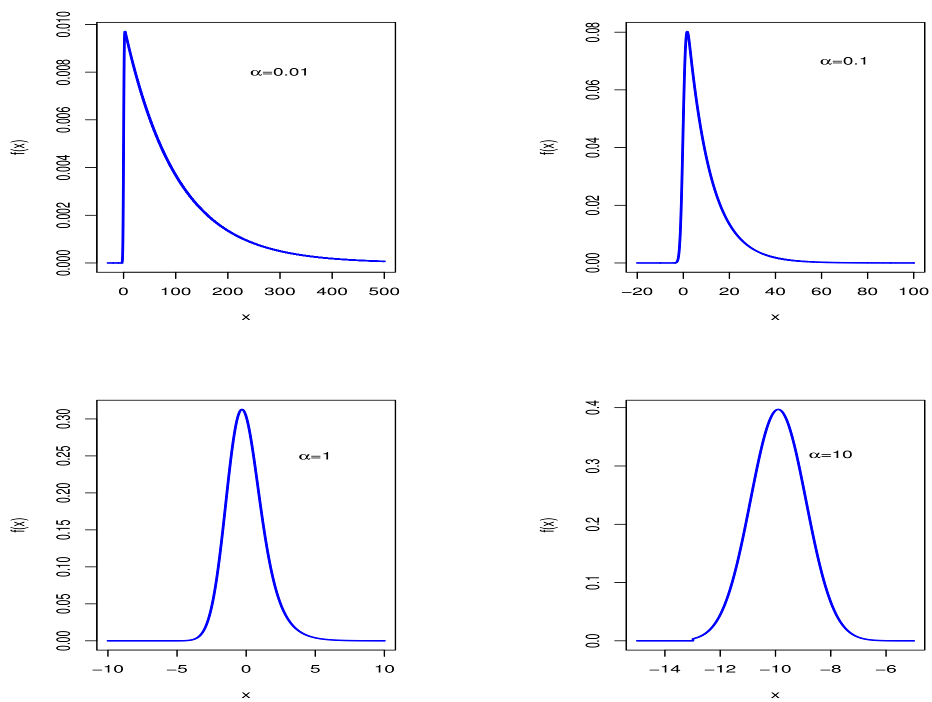

Figure 3: Relationship between density function and shape parameter α for normal-exponential distribution.

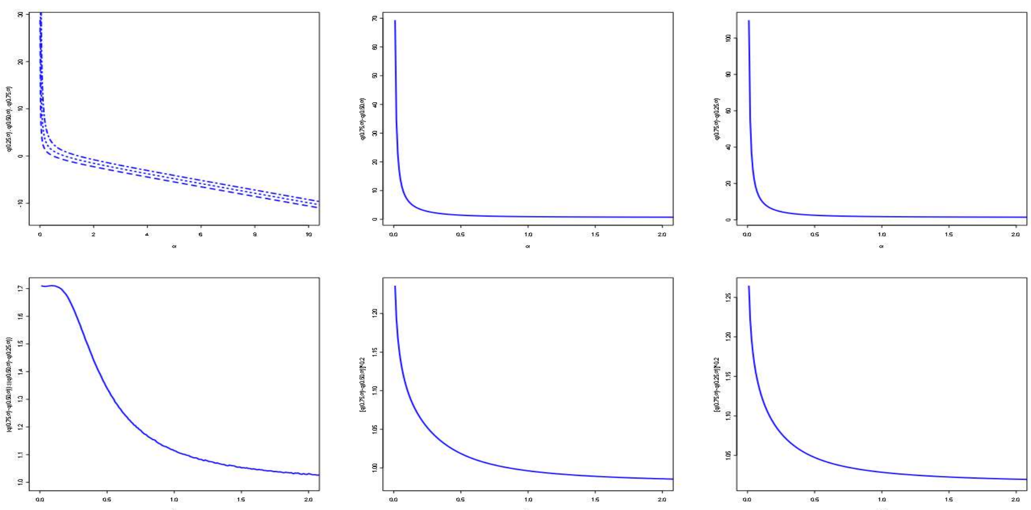

Figure 4: Relationship between quantiles and shape parameter α for normal-exponential distribution. (Top-left: $(q_{0.75,\alpha},\, q_{0.50,\alpha},\, q_{0.25,\alpha})$ vs. α; top-middle: $(q_{0.75,\alpha} - q_{0.50,\alpha})$ vs. α; top-right: $(q_{0.75,\alpha} - q_{0.25,\alpha})$ vs. α; bottom-left: $(\frac{q_{0.75,\alpha} - q_{0.50,\alpha}}{q_{0.50,\alpha} - q_{0.25,\alpha}})$ vs. α; bottom-middle: $(q_{0.75,\alpha} - q_{0.50,\alpha})^{\frac{1}{5}}$ vs. α; bottom-right: $(q_{0.75,\alpha} - q_{0.25,\alpha})^{\frac{1}{5}}$ vs. α)

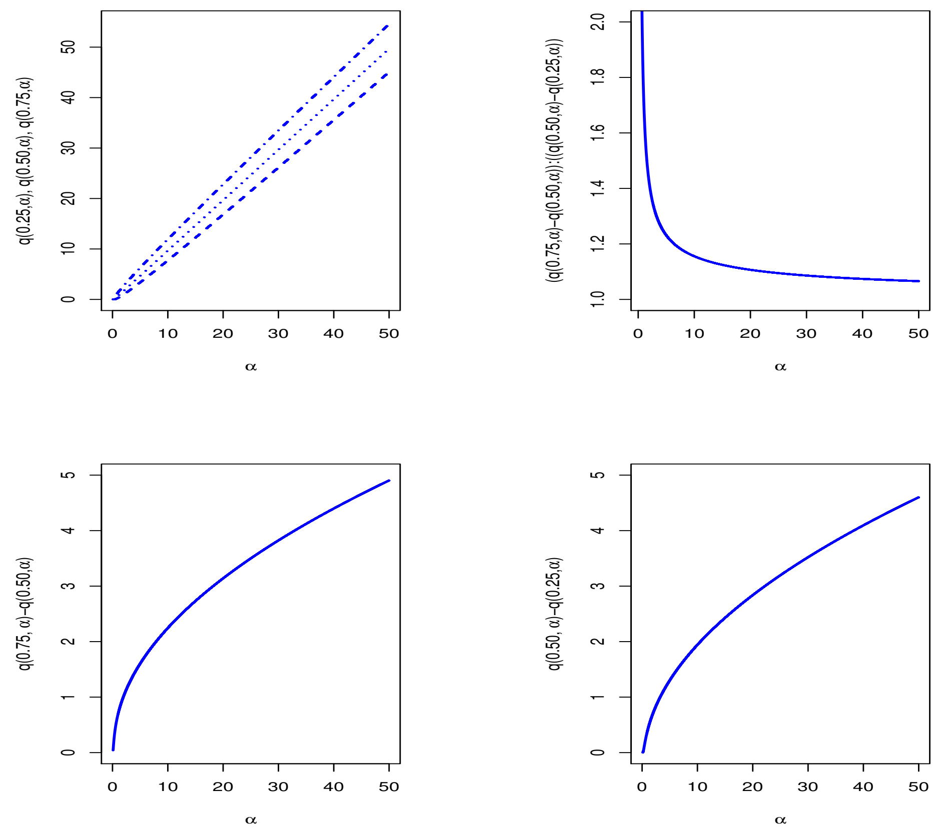

Figure 5: Relationship between quantiles and shape parameter α for Gamma distribution. (Top-left: $(q_{0.75,\alpha},\ q_{0.50,\alpha},\ q_{0.25,\alpha})$ vs. α; top-right: $(q_{0.75,\alpha} - q_{0.50,\alpha})/(q_{0.50,\alpha} - q_{0.25,\alpha})$ vs. α; bottom-left: $(q_{0.75,\alpha} - q_{0.50,\alpha})$ vs. α; bottom-right: $(q_{0.50,\alpha} - q_{0.25,\alpha})$ vs. α)

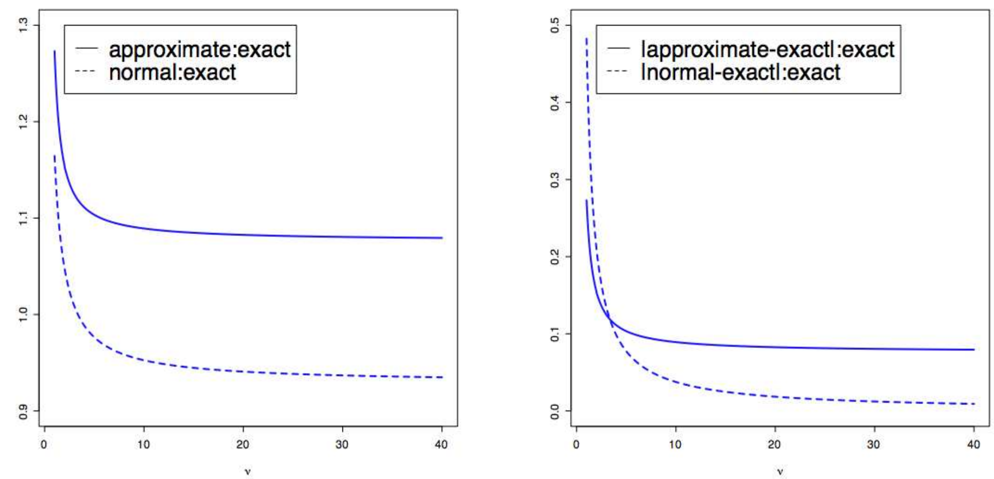

Figure 6: Scale parameter elicitation comparison vs. degrees of freedom ν in terms of relative error for Student's t prior.

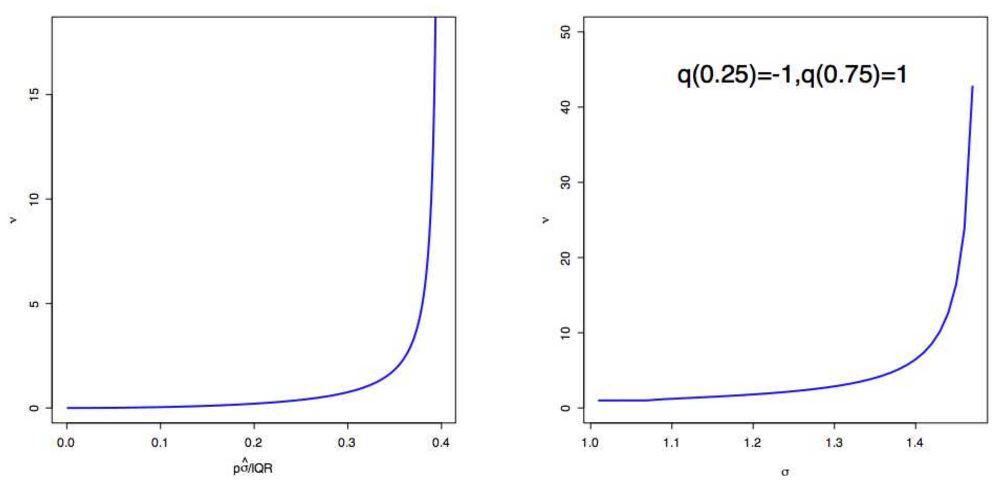

Figure 7: Estimated degrees of freedom ν from $\hat{\sigma}$.

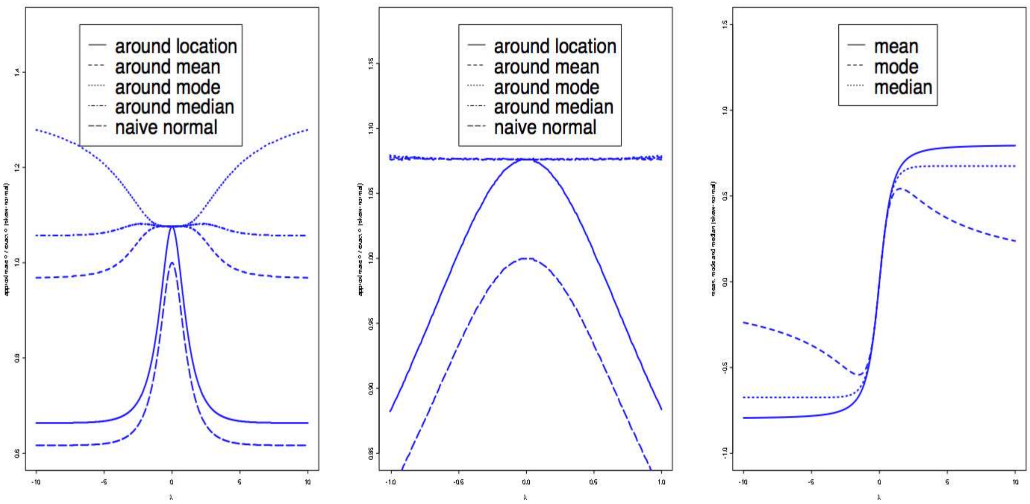

Figure 8: Scale parameter elicitation comparison for skew-normal. (Left and middle: approximate σ/exact σ vs. λ, Right: mean, mode and median vs. λ.)

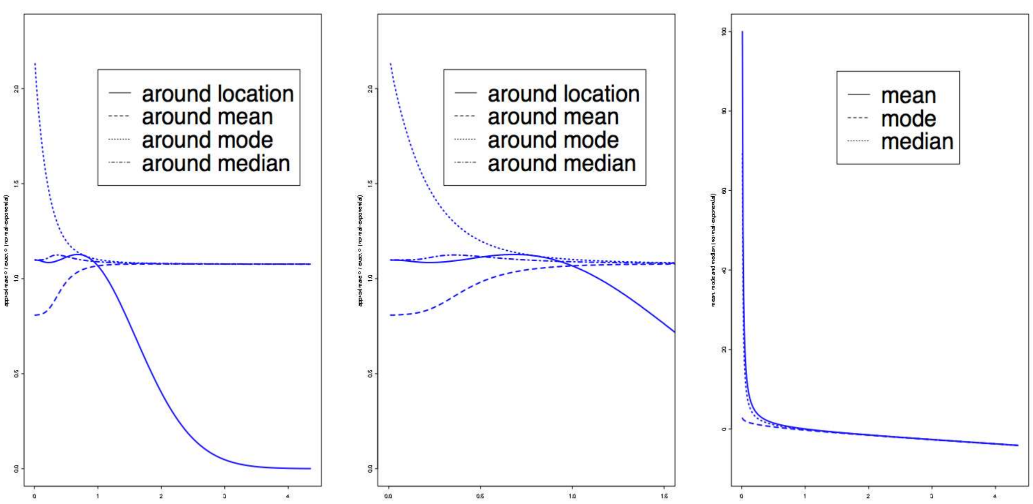

Figure 9: Scale parameter elicitation comparison for normal-exponential. (Left and middle: approximate σ/exact σ vs. α, Right: mean, mode and median vs. α.)

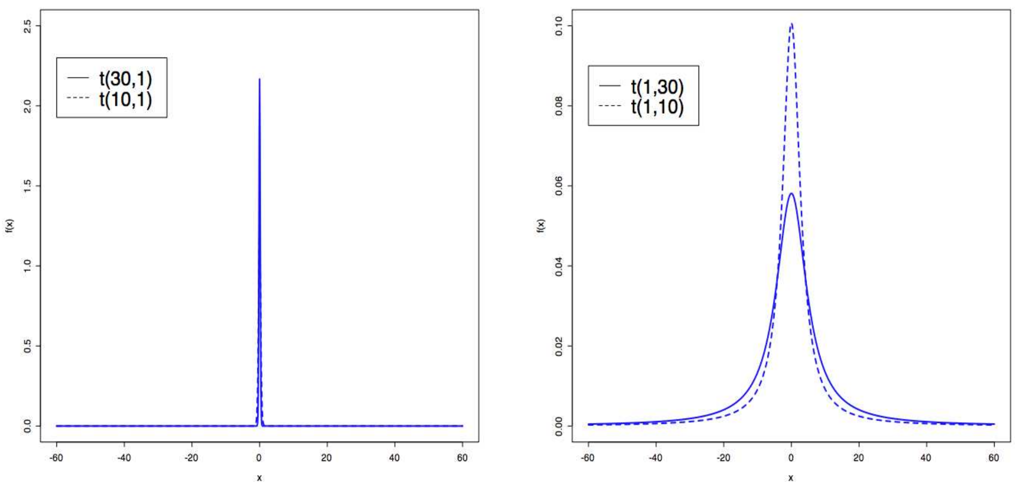

Figure 10: Generalized Student's t probability density function with different degrees of freedom.

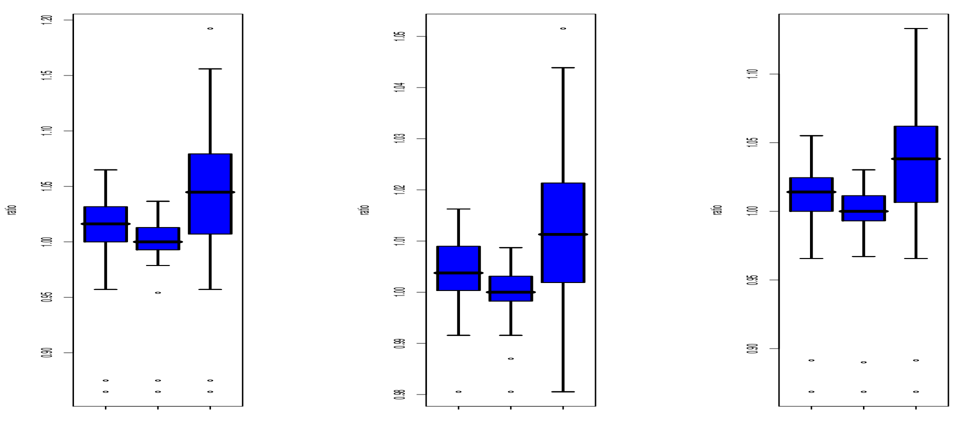

Figure 11: Skew-normal prior elicitation comparison (relative to IQR based iteration algorithm) among Taylor's expansions at median, mean and mode: the left panel is for λ, the middle panel is for σ, the right panel is for μ; from left to right in each panel, the ratios are for median, mean and mode

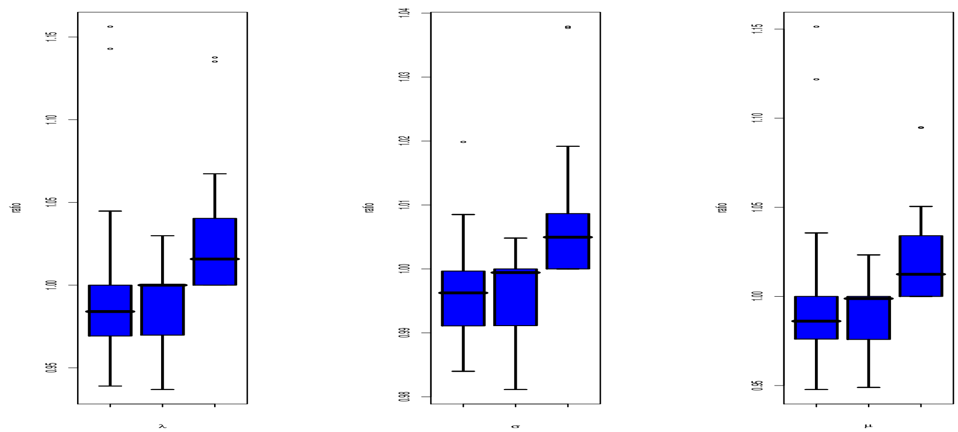

Figure 12: Skew-normal prior elicitation comparison (relative to median based iteration algorithm) among IQR based iteration, Taylor's expansions at mean and mode: the left panel is for λ, the middle panel is for σ, the right panel is for μ; from left to right in each panel, the ratios are for IQR, mean and mode

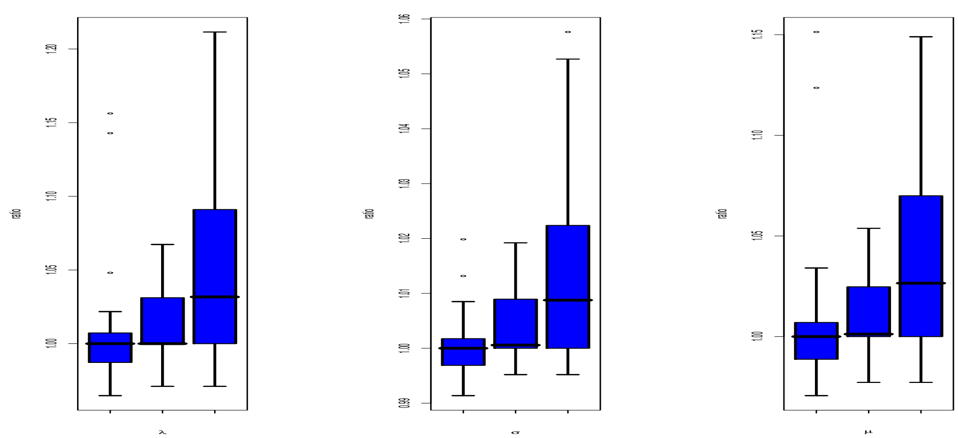

Figure 13: Skew-normal prior elicitation comparison (relative to mean based iteration algorithm) among IQR based iteration, Taylor's expansions at median and mode: the left panel is for λ, the middle panel is for σ, the right panel is for μ; from left to right in each panel, the ratios are for IQR, median and mode

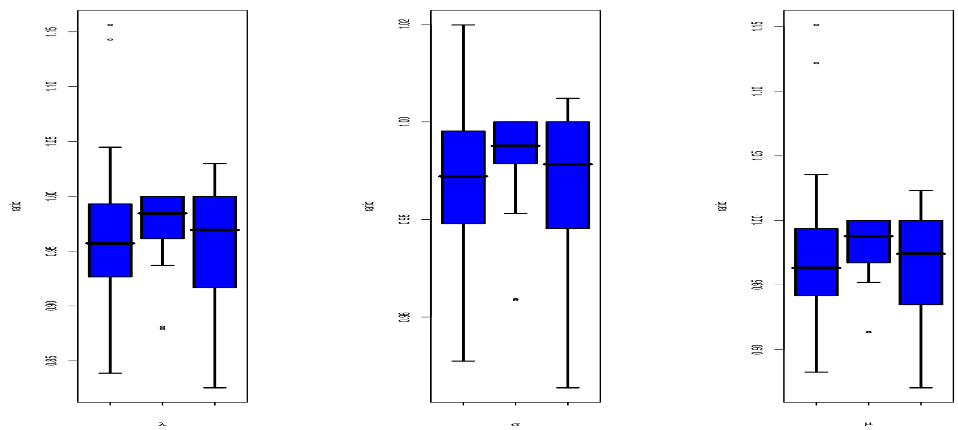

Figure 14: Skew-normal prior elicitation comparison (relative to mode based iteration algorithm) among IQR based iteration, Taylor's expansions at median and mean: the left panel is for λ, the middle panel is for σ, the right panel is for μ; from left to right in each panel, the ratios are for IQR, median and mean

References

Aigner, D. J., Lovell, C. K., and Schmidt, P. (1977). "Formulation and estimation of stochastic frontier production function model." *Journal of Econometrics*, 6(1): 21–37.

Al-Awadhi, S. A. and Garthwaite, P. H. (1998). "An elicitation method for multivariate normal distributions." *Communications in Statistics: Theory and Methods*, 27(5): 1123–1142.

Azzalini, A. (1985). "A class of distributions which includes the normal ones." *Scandinavian Journal of Statistics*, 12(2): 171–178.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 2 edition.

Dalal, S. R. and Hall, W. J. (1983). "Approximating priors by mixtures of natural conjugate priors." *Journal of the Royal Statistical Society - Series B*, 45(2): 278–286.

Garthwaite, P. H. (1989). "Fractile assessment for a linear regression model: an experimental study." *Organizational Behavior and Human Decision Processes*, 43: 188–206.

Garthwaite, P. H. and Dickey, J. M. (1985). "Double- and single-bisection methods for subjective probability assessment in a location-scale family." *Special Bayesian issue of J. Econometrics*, 29: 149–163.

— (1988). "Quantifying expert opinion in linear regression problems." *Journal of the Royal Statistical Society - Series B*, 50(3): 462–474.

Garthwaite, P. H., Kadane, J. B., and O'Hagan, A. (2005). "Statistical methods for elicitating probability distributions." *Journal of the American Statistical Association*, 100(470): 680–701.

Gill, J. and Walker, L. D. (2005). "Elicited priors for Bayesian model specifications in political science research." *The Journal of Politics*, 67(3): 841–872.

Jaynes, E. T. (1968). "Prior probabilities." *IEEE Transactions on Systems Science and Cybernetics*, SSC-4: 227–241.

Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S., and Peters, S. C. (1980). "Interactive elicitation of opinion for a normal linear model." *Journal of the American Statistical Association*, 75(372): 845–854.

Meeden, G. (1992). "An elicitation procedure using piecewise conjugate priors." In Goel, P. K. and Iyengar, N. S. (eds.), *Bayesian analysis in statistics and econometrics*. Springer-Verlag.

O'Hagan, A. (1998). "Eliciting expert beliefs in substantial practical applications." *The Statistician*, 47: 21–35.

Peterson, C. R. and Miller, A. (1964). "Mode, median and mean as optimal strategies." *Journal of Experimental Psychology*, 68: 363–367.

Sahu, S., Dey, D. K., and Branco, M. (2003). "A new class of multivariate skew distributions with application to Bayesian regression models." *The Canadian Journal of Statistics*, 31: 129–150.

Acknowledgments

The authors are grateful to Joseph Kadane for his insightful suggestions which have lead to significant improvement of the presentation, we also thank the editor and two referees for their constructive comments, Anthony Billings for his dedicated proof reading of the manuscript and Jun Tan for helping us with graphical outputs.