Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Stochastics and Statistics

Maximum entropy distributions with quantile information

Amirsaman H. Bajgiran a, Mahsa Mardikoraem b, Ehsan S. Soofi b, a

a Department of Industrial and Manufacturing Engineering, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
b Sheldon B. Lubar School of Business, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

A R T I C L E   I N F O

Article history:
Received 16 February 2020
Accepted 24 July 2020
Available online 8 August 2020

Keywords:
Decision analysis
Information value
Maximum entropy prior
Newsvendor
Survey of Professional Forecasters

A B S T R A C T

Quantiles are available in various problems for developing probability distributions. In some problems quantities are elicited from experts and used for fitting parametric models, which induce non-elicited information. In some other problems comparisons are made with a quantile of an assumed model which is noncommittal to the quantile information. The maximum entropy (ME) principle provides models that avoid these issues. However, the information theory literature has been mainly concerned about models based on moment information. This paper explores the ME models that are the minimum elaborations of the uniform and moment-based ME models by quantiles. This property provides diagnostics for the utility of elaboration in terms of the information value of each type of information over the other. The ME model with quantiles and moments is represented as the mixture of truncated distributions on consecutive intervals whose shapes and existence are determined by the moments. Elaborations of several ME distributions by quantiles are presented. The ME model based only on quantiles elicited by the fixed interval method possesses a useful property for pooling information elicited from multiple experts. The elaboration of Laplace distribution is an extension of the information theory connection with minimum risk under symmetric loss functions to the asymmetric linear loss. This extension produces a new Asymmetric Laplace distribution. Application examples compare ME priors with a parametric model fitted to elicited quantiles, illustrate measuring uncertainty and disagreement of economic forecasters based on elicited probabilities, and adjust ME models for a fundamental quantile in an inventory management problem.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Quantile information is available in various decision and statistical problems. Experiments on accuracy of elicitation of probability distributions have shown that subjects can assess median reasonably more accurate than the mean (Garthwaite, Kadane, & O'Hagan, 2005). Variable interval method and fixed interval method are used for eliciting quantiles and probabilities. Central banks such as the Federal Reserve Bank of Philadelphia, European Central Bank, the Bank of England, and other institutions survey economic forecasters for eliciting subjective probabilities for a set of fixed intervals. These probabilities are used to produce economic outlook reports for policy makers. The research area of measuring the uncertainty and disagreement of economic forecasters is primarily based on using these probabilities for developing other aspects of forecast distribution of variables such as the growth of Gross Domestic Product (GDP); Lahiri and Wang (2019) and Shoja and Soofi (2017) provide reviews and latest developments on this topic. In operations and business decision problems a quantile appears as the solution to the minimum risk (expected loss) under the asymmetric linear loss function such as the newsvendor problem, where the optimal order quantity that maximizes the profit is a quantile of the demand distribution (Snyder and Shen, 2011).

Quantiles provide partial information about a probability distribution. Traditionally, this type of partial information is used to fit parametric probability distributions. The information theory literature primarily has considered developing distributions based on various types of moments (Jaynes, 1957; 1968). The use of probability allocation has been sporadic (Brockett, Charnes, & Paick, 1995; Ebrahimi, Soofi, & Soyer, 2008); quantile is only noted (Asadi, Ebrahimi, Soofi, & Zarezadeh, 2014). This paper aims to fill this void.

Two information measures are used for deriving models. The Shannon entropy of a quantity X with continuous distribution F and a probability density function (PDF) f is defined by

\[ H(f) = - \int_{\mathbb{R}} f(x) \log f(x) \, dx, \tag{1} \]
where $\mathcal{R}$ is the range of $X$, provided that the integral is finite. The Kullback-Leibler (KL) information divergence between $f$ and $f_r$ is defined by

$$K(f : f_r) = \int_{\mathcal{R}} f(x) \log \frac{f(x)}{f_r(x)} \, dx \geq 0, \quad (2)$$

provided that $f$ is absolutely continuous with respect to $f_r$ ($f(x) = 0$ whenever $f_r(x) = 0$); the inequality becomes equality if and only if $f(x) = f_r(x)$ almost everywhere. Unlike the entropy, $K(f : f_r)$ is invariant under one-to-one transformations of $X$ for the continuous case. This measure has been used in operations and decision problems (see, for example, Alwan, Ebrahimi, and Soofi (1998); Asadi, Ebrahimi, and Soofi (2018); Plischke, Borgonovo, and Smith (2013); Saghafian and Tomlin (2016)). The Minimum Discrimination Information (MDI) principle minimizes $K(f : f_r)$ where $f$ is in a class of distributions $\Omega$ and $f_r \notin \Omega$. When $f_r$ is uniform (proper PDF or improper), the MDI coincides with the Maximum Entropy (ME) principle (Jaynes, 1957). For the continuous case, the ME is defined as maximization of $-K(f : f_r)$ where $f_r$ is interpreted as the “invariance measure” function (Jaynes, 1968). In the game theory problem of selecting a distribution where the utility is defined in terms of the score function, the MDI and ME models are minimax decisions (Grunwald & Dawid, 2004; Smith, 1974). Jaynes (1968) advanced assessing the information value of additional constraints by comparing the entropies of two discrete ME models obtained from information sets of two experts. Soofi, Ebrahimi, and Habullah (1995) extended this idea through the notion of information distinguishability defined by equality between $K(f : f_r)$ and the entropy difference where $f \in \Omega$ and $f_r$ is the ME model.

We consider the following three classes of partial information about $f$ on continuous support:

$$\Omega_{\theta} = \{ f_\theta : E_f[1_j(X)] = \theta_j, \ j = 1, \ldots, J \}, \quad (3)$$

$$\Omega_{\alpha} = \{ f_\alpha : E_f[1(X \leq q_k)] = \alpha_k, \ k = 1, \ldots, m \}. \quad (4)$$

$$\Omega_{\theta, \alpha} = \Omega_{\theta} \cup \Omega_{\alpha}, \quad (5)$$

where $1_j(X)$ is integrable with respect to $f$. $1(A)$ is the indicator function of the event $A$, and $q_k = F^{-1}(\alpha_k), k = 1, \ldots, m$ are quantiles; (3) and (4) are referred to as moment information and quantile information (QI) and $\Omega_{\theta, \alpha}$ is the set of PDFs that include both types of information.

Fig. 1 gives a schematic representation of the modeling approach of this paper. Distributions in (3) and (4) are embedded in the larger family (5). Such larger families are called elaboration of the smaller models (Box & Tiao, 1973) indexed by the elaboration parameter (Carota, Parmigiani, & Nicholas, 1966). In Fig. 1 $f_{\theta, \alpha} \in \Omega_{\theta, \alpha}$ is elaboration of $f_\theta$ and $f_{\alpha}$ indexed by $\alpha$ and $\theta$, respectively. Carota et al. (1966) proposed a Bayesian KL diagnostic for measuring the utility of elaboration. Thus far, elaborations are chosen arbitrarily for model evaluation. Including additional moments to ME and MDI models is well-known, but it is not characterized as being the minimum elaboration. This research develops minimum elaborations of ME models by $\alpha$ and $\theta$ according to the MDI principle.

This research develops ME minimal elaborations of ME models and diagnostics (highlighted parts of Fig. 1). The ME models $f_\theta$ and $f^{ME}_{\theta, \alpha}$ are elaborations of the uniform distribution $f^0$ by the QI, respectively. We show that the MDI updating of $f^{MDI}_{\theta, \alpha}$ and $f_{\alpha}$ relative to each other give the equivalence of $f^{MDI}_{\theta, \alpha}$ and $f_{\alpha}$ and characterizes $f^{MDI}_{\theta, \alpha}$ as the minimum elaboration of the smaller ME models. These MDI adjustments of $f^0_{\alpha}$ and $f_{\alpha}$ for satisfying the constraints in $\Omega_{\theta, \alpha}$ provide diagnostics for assessing their information values and the utility of the minimum elaboration. A result represents $f^{MDI}_{\theta, \alpha}$ as the mixture of truncated PDFs on supports formed by consecutive quantiles whose shapes and existence are determined by the type of moments. This representation facilitates computation of the moments and entropy of $f^{MDI}_{\theta, \alpha}$ and simulations from this model. Elaborations of the classical ME models (uniform, exponential, and normal) by QI are given. A result explores a useful property of $f^{MDI}_{\theta, \alpha}$ for aggregating quantiles elicited from multiple experts by the fixed interval method. The elaboration of Laplace distribution extends the notion of minimum risk ME models under symmetric loss (Ardakani et al., 2017) to an asymmetric loss and produces a new Asymmetric Laplace (AL) distribution. The elaboration of a well-known AL is also given.

Applications are illustrated for a wide range of problems with a common theme: ME models with QI purely carry the given information into the results because of being “maximally noncommittal with regard to missing information” (Jaynes, 1957). Three application areas illustrate cases where non-elicited information is included or existing information is excluded. The ME prior, $f^0_{\alpha}$, is piecewise uniform and proper on a bounded range, which is most directly in accord with Laplace’s principle of “insufficient reason”. The range can be set wide like fitting parametric models such as gamma priors to quantiles elicited for a finite range. However, unlike the parametric families, $f^0_{\alpha}$ does not induce non-elicited information. Model fitting also is used for measuring uncertainty and disagreement of forecasters where probabilities are elicited by the fixed interval method. The result for the ME model based on QI...
with fixed intervals is particularly suitable for this problem. This is illustrated for the Federal Reserve Bank of Philadelphia’s Survey of Professional Forecasters. The case of excluding existing information is seen in the newsvendor (NV) problem of inventory management. The ME principle has been invoked for justifying the classical ME distributions as models for the demand distribution (Andersson, Jörnsten, Nonås, Sandal, & Ubøe, 2013; Perakis & Roels, 2008), where a profit maximizing QI that appears in the analysis is not accounted for. We illustrate updating of these classical ME models in light of this information. The new minimum risk ME model and the elaboration of the known AL model are particularly useful for the NV problem.

The paper is organized as follows. Section 2 gives preliminaries and ME models with moment information. Section 3 presents the minimum elaboration of ME models by the QI. Section 4 illustrates potential applications of QI for developing Bayesian prior, measuring uncertainty and disagreement of economic forecasters, and reliability and inventory problems. Section 5 gives some concluding remarks. Computational details are available in a Supplementary Document.

2. Preliminaries

The ME model in $\Omega_\theta$, if exists, is unique and has PDF in the following form:

$$f_0(x) = C_\eta \exp \left\{ -\sum_{j=1}^{L} \eta_j T_j(x) \right\}, \quad x \in \mathbb{R},$$  \hspace{1cm} (6)

where $\eta_1, \ldots, \eta_L$ are Lagrange multipliers and

$$C_\eta = C(\eta_1, \ldots, \eta_L) = \left[ \int_{\mathbb{R}} \exp \left\{ -\sum_{j=1}^{L} \eta_j T_j(x) \right\} dx \right]^{-1} < \infty, \quad (7)$$

is the normalizing factor for $f_0$, and $\theta_j = -\partial \log C_\eta / \partial \eta_j$, see, for example, Soofi et al. (1995). The existence of the ME model is determined by the finiteness condition (7).

Well-known families of probability distributions are ME models with various types of moment information. On the finite range an ME distribution always exists. The best known examples are the uniform distribution $[a, b]$ with the range information $E_j[1(a \leq X \leq a+b)] = 1$ and the beta distribution $[1,2]$ with two geometric means: $T_1(x) = \log x$, $T_2(x) = \log (1-x)$; hereafter the range information will be omitted. On the nonnegative range, well known examples of ME models are the exponential distribution with $T_1(x) = x$, the gamma distribution with $T_1(x) = x$ and $T_2(x) = \log x$, and the log-normal distribution with $T_1(x) = \log x$, $T_2(x) = (\log x)^2$. On the unrestricted range, the ME model with $T_j(x) = x^j$, $j = 1, \ldots, J$ when $j$ is an odd number does not exist. Examples of the ME models with even $j$ are the normal with $J = 2$ and the quartic exponential with $J = 4$ (Zellner & Hidifield, 1988).

If $f_0 \in \Omega_\theta$ is the ME model, then for any $f \in \Omega_\theta$ with finite entropy,

$$K(f : f_0) = H(f_0) - H(f), \quad (8)$$

provided that $f$ is absolutely continuous with respect to $f_0$ (Soofi et al., 1995). Due to the uniqueness of $f_0 \in \Omega_\theta$, $f$ is distinguishable with $F \in \Omega$ if and only if it reduces the maximum entropy.

The information distinguishability (ID) index of $f$ is defined by normalized KL measure

$$ID_0(f) = 1 - e^{-2K(f : f_0)}. \quad (9)$$

When $f$ is also an ME in a subset of $\Omega_\theta$ with more constraints, (8) provides the information values of additional constraints. Asadi et al. (2014) have used this diagnostic for assessing the information value of additional moments in a ME problem. Mc Culloch (1989) proposed a calibration of the KL divergence based on the KL between the Bernoulli distributions for the outcome of a biased coin with $\pi > 0.5$ and a fair coin with $\pi_0 = 0.5$. This measure is represented by

$$\pi(K) = \frac{1}{2} \left( 1 + \sqrt{ID(K)} \right) \in [0.5, 1]. \quad (10)$$

2.1. Symmetric minimum risk ME models

The Laplace and normal distributions are ME models consistent with $T(X) = |X - \mu|$ for $\nu = 1, 2$, respectively. Ardakani et al. (2017) discussed the ME problem consistent with the minimum decision-theoretic risk of predicting $X$ under the general symmetric loss function $L(X, \tilde{X}) = |X - \tilde{X}|$, $\nu > 0$ on the unrestricted range. In the statistical decision theory, the consequence of a decision $d$ about an unknown quantity $X$ is a loss function $L(X, d) \geq 0$ and the risk function of a decision rule is defined by $R(X, d) = E_J[L(X, d)]$ (Berg, 1985). The optimal decision in a set of possible decisions $D$ is defined by

$$d^* = \arg \min_{d \in D} E_J[L(X, d)]. \quad (11)$$

The ME model consistent with the minimum risk of decision, $d$, is as follows:

$$\max_{f \in \Omega_\theta} \left\{ E_J[L(X, d)] \right\} = \theta. \quad$$

Let

$$\Omega_\theta = \left\{ f : E_J[L(X, d)] = \theta, E_J[T(X)] = d^* \right\},$$

where $d^* = E_J[T(X)]$ be the minimum risk of the optimal decision. The ME model in $\Omega_\theta$, if exists, is given by

$$f_\theta(x) = C_\eta \exp \left\{ -\sum_{j=1}^{L} \eta_j T_j(x) \right\}, \quad (11)$$

where the optimal decision under the quadratic loss is the mean, $d^* = E(X)$ and the risk is the variance, so the minimum risk ME problem includes the first two moments. On the unrestricted range, the minimum risk ME model is normal, but the solutions on the finite and nonnegative ranges are not so straightforward; these cases will be discussed in Section 2.2. The optimal decision under the absolute loss (symmetric linear loss function $\nu = 1$) is the median, so the minimum risk ME problem includes a moment and a quantile; this case will be extended to the asymmetric linear loss function in Section 3.3. (Fleischhacker & Folk (2015) studied the ME model with the expected loss under the loss function $L(p) = x A p$, where $p = (p_1, \ldots, p_N)$, $\sum_{j=1}^{N} p_j = 1$ is a probability vector (discrete distribution on $N$ point), $A$ is a nonnegative $N \times N$ matrix, and $x^j = (x_1, \ldots, x_N)$.)

2.2. First two moments on restricted range

The ME modeling with the first two moments, $E_J(X^j) = \theta_j$, $j = 1, 2$ on a restricted range (finite or nonnegative) is intricate. By (6), the PDF of the ME model, if exists, is in the following form:

$$f_0(x) = C_\eta e^{-\eta_1 x - \eta_2 x^2}, \quad (12)$$

where the Lagrange multipliers are given by $\theta_1 = E(X) = \mu$ and $\sigma^2 = \theta_2 - \theta_1^2$ as follows:

$$\theta_1 = \frac{\eta_1}{2\eta_2}, \quad \sigma^2 = \frac{1}{2\eta_2}. \quad (13)$$

However, unlike the unrestricted range, the existence and shape of the ME model depend on the range and the relationship between the moments. Dowson and Wragg (1973) give a rigorous treatment of the problem and provide results that are summarized in our notations as follows.
On the finite range the ME model always exists. The left panel of Fig. 2 shows various shapes of the ME PDFs with first two moments on [0,1] when \( \theta_1, \theta_2 > 0 \) (truncated normal); \( \theta_1 = \theta_2 = 0 \) (U-shaped); \( \theta_1 = 0, \theta_2 > 0 \) (truncated exponential); \( \theta_1 = \eta_1 = \eta_2 = 0 \) (Uniform); \( \theta_1, \eta_2 < 0 \) (U-shaped).

![Fig. 2. Shapes of the maximum entropy PDFs with first two moments on [0,1] and regions of parameters for the shapes of the distributions: \( A_1, \eta_1 > 0 \) (truncated normal); \( A_2, \eta_2 = 0 \) (truncated exponential, \( \eta_1 = \eta_2 = 0 \) (Uniform); \( A_3, \eta_2 < 0 \) (U-shaped).](image)

The information measure of a mixture PDF \( f \) is the Jensen-Shannon (JS) divergence,

\[
JS(f : \{ f_i \}; W) = \sum_{i=1}^{n} w_i K(f_i : f) = H(f) - \sum_{i=1}^{n} w_i H(f_i) \geq 0,
\]

where \( W \) stands for \( \{ w_1, \ldots, w_n \} \) and \( K \) and \( H \) are the KL divergence and Shannon entropy as defined in (2) and (1). The KL divergences \( K(f_i; f) \) are well defined due to the fact that for all \( i \) the supports of \( f_i \) are subsets of the support of \( f \), but in general, \( K(f_i; f) \) and \( H(f) \) cannot be computed in closed form. A relationship between the Shannon entropy and the KL divergence implies the second equality. It is clear from (17) that the JS is equal to zero if and only if all individual models are identical. The entropy difference (18) gives the uncertainty increase (information decrease) when the pooled distribution is used for representing the set of information of individual distributions.

In the sequel, we will also encounter mixtures where the set of supports of \( f_i \) is a partition of the support of \( f \). In these cases application of the entropy decomposition formula (Di Crescenzo & Longobardi, 2002) gives the entropy of \( f \) as follows:

\[
H(f) = H(W) + \sum_{i=1}^{n} w_i H(f_i).
\]

3. ME minimum elaboration

The elaboration of \( f_{\theta, \alpha} \in \Omega_{\theta, \alpha} \) by QI is the PDF of the ME distribution in \( \Omega_{\theta, \alpha} \) given by

\[
f_{\theta, \alpha}(x) = c_{\theta, \alpha} \exp \left\{ -\sum_{j=1}^{J} \lambda_j T_j(x) - \sum_{k=1}^{m} \xi_k I(x < q_k) \right\}, \quad -\infty < x < \infty.
\]

where the normalizing factor \( c_{\theta, \alpha} \) is determined by the Lagrange multipliers \( (\lambda_1, \ldots, \lambda_J, \xi_1, \ldots, \xi_m) \) for the moment and QI constraints in \( \Omega_{\theta, \alpha} \). This ME model gives \( f_{\theta, \alpha}^{(k)}(q_k) = \alpha_k, k = 1, \ldots, m \). Because \( \lambda_j, j = 1, \ldots, J \) are determined by all constraints in \( \Omega_{\theta, \alpha} \), in general, they are different from the parameters \( \eta_j, j = 1, \ldots, J \) in (6).
The ME model (20) generalizes [6]. In the absence of the moment information, (20) gives \( f_{\theta,\alpha}^* \) as the elaboration of the uniform distribution by QI. With the moment information, \( f_{\theta,\alpha}^* \) provides the elaboration of \( f_{\theta,\alpha}^* \) by \( \alpha \) and the elaboration of \( f_{\theta,\alpha}^* \) by \( \theta \).

The KL divergence provides diagnostics for measuring the utility of elaboration (Carota et al., 1966). Because \( \Omega_{\theta,\alpha} \subseteq \Omega_{\theta,\alpha}^* \), the ME model \( f_{\theta,\alpha}^* \) does not satisfy the QI constraints and \( f_{\theta,\alpha}^* \) does not satisfy the moment constraints. The ME principle provides the minimum adjustment for these models to satisfy both types of constraints. The MDI distribution in \( \Omega_{\theta,\alpha} \), if exists, is unique and has PDF in the following form:

\[
\begin{align*}
\nu_{\theta,\alpha}(x) &= C_{\theta,\alpha} f_{\theta,\alpha}(x) \exp \left\{ \sum_{j=1}^{J} \lambda_{\text{md}_j} T_j(x) + \sum_{k=1}^{\text{md}_i} \zeta_{\text{md}_k} (x - q_k) \right\},
\end{align*}
\]

where the Lagrange multipliers are for constraints in \( \Omega_{\theta,\alpha} \), and \( C_{\theta,\alpha} \) is the normalizing factor for \( \nu_{\theta,\alpha}(x) \). Letting \( f_{\theta,\alpha}(x) = f_{\theta,\alpha}^*(x) \) and \( f_{\theta,\alpha}(x) = f_{\theta,\alpha}^*(x) \) gives the MDI updates of these ME models for satisfying both types of constraints. The following lemma gives the equivalence between \( f_{\theta,\alpha}^* \) and \( \nu_{\theta,\alpha}(x) \) which characterizes the minimum elaboration of \( f_{\theta,\alpha}^* \) and \( f_{\theta,\alpha}^* \) by \( \alpha \) and \( \theta \), respectively.

**Lemma 1.** Let \( f_{\theta,\alpha}^* \in \Omega_{\theta,\alpha} \) and \( f_{\theta,\alpha}^* \in \Omega_{\theta,\alpha} \) be the ME models and \( \nu_{\theta,\alpha}(x) \in \Omega_{\theta,\alpha} \) denote the MDI model with reference to \( f_{\theta,\alpha}^* \) or \( f_{\theta,\alpha}^* \). Then, the MDI updates of \( f_{\theta,\alpha}^* \) and \( f_{\theta,\alpha}^* \) for satisfying \( \Omega_{\theta,\alpha} \) and \( \Omega_{\theta,\alpha} \), respectively, are identical to \( f_{\theta,\alpha}^* \in \Omega_{\theta,\alpha} \), almost everywhere, where \( \lambda_j = \eta_j - \lambda_{\text{md}_j} \), \( \zeta_{\text{md}_k} = -\zeta_k \) for all \( j \) and \( k \).

Application of (8) to result of Lemma 1 establishes the following two ID relationships for updating of each type of information by the other type:

\[
\begin{align*}
K(f_{\theta,\alpha}^* : f_{\theta,\alpha}^*) &= H(f_{\theta,\alpha}^*) - H(f_{\theta,\alpha}^*) \geq 0, \quad (21) \\
K(f_{\theta,\alpha}^* : f_{\theta,\alpha}^*) &= H(f_{\theta,\alpha}^*) - H(f_{\theta,\alpha}^*) \geq 0, \quad (22)
\end{align*}
\]

where the inequalities become equalities if and only if \( f_{\theta,\alpha}^*(x) = f_{\theta,\alpha}^*(x) \) and \( f_{\theta,\alpha}^*(x) = f_{\theta,\alpha}^*(x) \), respectively, almost everywhere. By the inequalities in (21) and (22), the PDF \( f_{\theta,\alpha}^* \) is more concentrated than \( f_{\theta,\alpha}^* \) and \( f_{\theta,\alpha}^* \). Hence, the addition of one type of partial information to the other type leads to an information gain. The KL divergences in (21) and (22) give the minimum utility of the elaboration \( f_{\theta,\alpha}^* \) for \( f_{\theta,\alpha}^* \) and \( f_{\theta,\alpha}^* \). The entropy reductions in (21) and (22) quantify the information values of each type of partial information (Jaynes, 1968).

Consider the partition of the support of \( f \) formed by consecutive quantiles in (4):

\[
\mathcal{R} = \bigcup_{k=1}^{m-1} \{x: q_{k-1} < x < q_k\}, \quad (23)
\]

where \( q_m = \sup \{x: F(x) = 0\} \) and \( q_{m+1} = \inf \{x: F(x) = 1\} \). The constraints in (4) can be represented in terms of the following expectations:

\[
E[I(1(q_{k-1} < x \leq q_k))] = \alpha_k - \alpha_{k-1} = p_k, \quad k = 1, \ldots, m, \quad (24)
\]

where \( \alpha_0 = 0 \). The following proposition gives a new representation of the ME model (20) in terms of the mixture of \( m+1 \) truncated distributions on the partition of the support (23).

**Proposition 1.** Let \( \Omega_{\theta,\alpha} \) be the class of distributions defined in (5). Then the ME model, if exists, is continuous and its PDF has the following \((m+1)\)-piece truncated representation:

\[
f_{\theta,\alpha}(x) = \sum_{k=1}^{m+1} p_k f_k(x) I(q_{k-1} < x \leq q_k), \quad -\infty < x < \infty, \quad (25)
\]

where \( p_{m+1} = 1 - \alpha_m \).

\[
f_k(x) = \frac{C_k}{f_k(q_k) - f_k(q_{k-1})} \exp \left\{ - \sum_{j=1}^{J} \lambda_j T_j(x) \right\}, \quad q_{k-1} < x < q_k, \quad (26)
\]

\[C_k\] is determined by \( p_k, q_k, k = 1, \ldots, m, E[T_j(x)] = \theta_j, j = 1, \ldots, J, \]

and the Lagrange multipliers \( \lambda_j, j = 1, \ldots, J \) for the moment constraints in \( \Omega_{\theta,\alpha} \), and \( F_k \) is the CDF with PDF \( f_k(x) \).

The existence and shape of the ME model \( f_{\theta,\alpha}^* \) is determined by the existence and shapes of \( f_k, k = 1, \ldots, m+1 \). The kernels of the truncated PDFs \( f_k \) are in the form of the ME model \( f_{\theta,\alpha}^* \), but the Lagrange multipliers of \( f_k \) and \( f_k^* \) are different due to the additional QI constraints. **Proposition 1** is insightful and facilitates computations of the moments of \( f_{\theta,\alpha}^* \) via (15). Application of the entropy decomposition (19) facilitates computation of diagnostics (21) and (22) via \( H(f_k) \) given by (5.10) of Supplementary Document. Representation (25) facilitates simulations via mixing outcomes generated from \( f_k, k = 1, \ldots, m+1 \).

### 3.1. Elaboration of uniform with QI

The ME model with only QI constraints is in the following mixture of uniform PDFs:

\[
f_u(x) = \sum_{k=1}^{m+1} \frac{p_k}{B_k} I(q_{k-1} \leq x \leq q_k), \quad a < x < b, \quad (27)
\]

where \( B_k = q_k - q_{k-1} \), \( q_0 = 0 \), \( q_{m+1} = b \), \( q_0 = 0 \), and \( \alpha_{m+1} = 1 \). This piecewise uniform PDF is a density histogram with unequal bins. On an unbounded range, additional information such as moments to supplement (24) is needed for the existence of an ME model. In the absence of moment information lower and upper limits for the unbounded bins, \( X \leq q_1 \) and \( X > q_m \), should be set (like constructing histograms).

The entropy of (27) is given by (19) with \( H(f_k) = \log B_k \) and \( H(W) = \log P_k \), where \( P_k \) is distribution with probabilities \( p_1, \ldots, p_{m+1} \) of the partition sets \( B_k \). This measure is sensitive to the choice of end points for unbounded bins. Its sensitivity can be easily assessed using \( p_k \log B_k \) and \( p_{m+1} \log B_{m+1} \). The diagnostics (21) and (22) are invariant for the choice of endpoints.

When partial information is elicited from multiple experts pooling the data is in order. The arithmetic pool is commonplace. For developing ME models two options may be followed: (a) develop the ME model that is consistent with arithmetic pool of partial information; or (b) pool the set of ME models that are consistent with the partial information provided by individual respondents. In general, these options provide two different forecast models. For example, the ME model consistent with the arithmetic pools of moments is in the same family as the individual ME models, but is not the mixture of the individual ME models (Shoja & Soofi, 2017). However, the ME model (27) produces the same model for the two options when probabilities are elicited from multiple experts using fixed interval method. In this approach, the range of \( X \) is partitioned into intervals and the expert provides the probability that \( X \) will fall in each interval.

The following proposition highlights this interesting mixture property of ME model (27).

**Proposition 2.** Let \( p_k = P_k[q_{k-1} \leq X \leq q_k] \) denote probabilities assigned by an individual to given bins \( B_k = q_k - q_{k-1}, k = 1, \ldots, m+1 \) and \( \{f_{\theta,\alpha}^*, i = 1, \ldots, n\} \) be the set of individual ME models (27). Then,

\[
f^*(x) = \sum_{i=1}^{n} W_i f_{\theta,\alpha}^*(x), \quad (28)
\]
where $f^*_M$ is the ME model consistent with the average bin probabilities $\tilde{w}_k = \sum_{i=1}^{m} w_i p_{ik}$.

In pooling applications, the subscript of $f^*_M$ stands for its designation as the consensus distribution and the weights are uniform. The divergence and information measures (17) and (18) are invariant for the choice of endpoints for the unbounded bins (Supplementary Document, Section S.4). In the context of pooling, (18) gives the information increment of the set of partial information provided by individuals over their pool.

3.2. Elaboration of exponential and normal with Ql

The main moment information sets used in ME modeling are the mean and the first two moments. The following corollary gives the ME model with the QI and mean constraints.

**Corollary 1.** Let $\Omega_{\theta,\alpha}$ be the class of distributions on support $0 \leq x < b \leq \infty$, with $E_J(X) = \mu$. Then the ME model is the $(m+1)$-piece truncated exponential with the following PDF:

$$f^*_{\mu,\alpha}(x) = \sum_{k=1}^{m+1} p_k f_k(x) 1(q_{k-1} < x \leq q_k), \quad 0 \leq b < \infty,$$

where $q_0 = 0$, $q_{m+1} = b$, $f_k(x) = \frac{\lambda e^{-\lambda x}}{e^{-\lambda q_k} - e^{-\lambda q_{k-1}}}$, $q_{k-1} < x < q_k$, and $\lambda$ is given by

$$\lambda = \frac{1}{\mu} \sum_{k=1}^{m+1} p_k \left[1 - e^{-\lambda q_k} + e^{-\lambda q_{k-1}}\right].$$

When QI constraint is included in addition to the first two moments, the ME model is given by Proposition 1, where the shapes and existences of the truncated distributions in (26) are given according to results of Dowson and Wragg (1973), summarized in Section 2.2. Let $A\alpha = A(\alpha)$ where $A: [0, 1] \times [0, 1] \rightarrow [a, b] \times [a, b]$ and the regions $A\alpha, \alpha = 1, 2, 3$ are shown in Fig. 2. This is an affine transformation which affects the location and scale of a PDF, but does not affect its shape. The following corollary presents the shape and existence of the ME model with QI in addition to the first two moments for the finite, nonnegative, and unrestricted ranges.

**Corollary 2.** Let $f_k(x)$ be as defined in (26) and denote their first two moments by $\theta_j = E_k(X)$, $j = 1, 2$. Then the shape and existence of the ME model $f^*_{\theta,\alpha}$ is determined by $f_k, k = 1, \ldots, m + 1$ as follows.

(a) On the finite range $a \leq x \leq b$, $f^*_{\theta,\alpha}$ always exists and its shape is given by

$$f^*_k(x) = \begin{cases} \frac{TN(\mu_k, \sigma_k)}{\tilde{w}_k} \text{if } (\theta_1k, \theta_2k) \in A_1, \\ \frac{TE(\theta_1)k}{\tilde{w}_k} \text{if } (\theta_1k, \theta_2k) \in A_2, \theta_1k \neq \frac{1}{q_k - q_{k-1}}, \\ \text{Uniform} \text{if } (\theta_1k, \theta_2k) \in A_2, \theta_1k = \frac{1}{q_k - q_{k-1}}, \\ \text{U-shaped} \text{if } (\theta_1k, \theta_2k) \in A_3, \end{cases}$$

where $TN(\mu_k, \sigma_k)$ and $TE(\theta_1k)$ are truncated normal and exponential with parent distributions $N(\mu_k, \sigma_k^2)$ and $Exp(\theta_1k)$ on the unrestricted and nonnegative range, respectively.

(b) On the nonnegative range, the shapes of $f^*_k$ on $q_{k-1} \leq x \leq q_k, k = 1, \ldots, m$ are as in (a) and its shape and existence on $x > q_m$ is given by

$$f^*_m(x) = \begin{cases} \frac{TN(\mu_m, \sigma_m)}{\tilde{w}_m} \text{if } c \nu_m < 1, \\ \text{Exp}(\theta_m) \text{if } c \nu_m = 1, \\ \text{Does not exist} \text{if } c \nu_m > 1. \end{cases}$$

(c) On the unrestricted range, the shapes of $f^*_M$ on $q_{k-1} \leq x \leq q_k, k = 2, \ldots, m$ are as in (a) and shapes and existence on $k = 1, m + 1$ are as in (b), where $f_1(x) = f^*_0(x) = (q_1 - x) / c_{\alpha}, x \leq q_1$.

3.3. Elaboration of Laplace models with QI

The ME characterization of the Laplace distribution extends to the family of ME models with expected minimum risk under the following asymmetric linear loss function:

$$L(\xi, X) = \begin{cases} c_0(\xi - X), & X \leq \xi, \\ c_0(X - \xi), & X > \xi. \end{cases}$$

where $c_0$ and $c_0$ are factors for losses of overestimation and underestimation, respectively. The loss function (33) is used in various problems such as quantile regression estimation and inventory problems. The optimal decision that minimizes $E_I[L(\xi, X)]$ is given by

$$q = \arg \min_{\xi} \left\{ c_0E_I([\xi - X]) + c_0E_I(X - \xi) \right\} + c_0E_I([\xi + X]) = \theta = \frac{\alpha}{\mu}.$$
metric quadratic and linear loss functions, respectively. The ME model (36) is the PDF of a new AL distribution. It generalizes the symmetric Laplace distribution with \( c_0 = c_0 \).

Fig. 3 depicts the PDFs and CDFs of three examples of the ME model given by Proposition 3 with \( \theta = 5 \) and \( q = 20 \) chosen for the purpose of illustration, but different under and over estimation costs: \((c_0, c_0) = (9, 1)\) (solid blue), \((c_0, c_0) = (9.5, 3.3)\) (dashed blue), and \( c_0 = c_0 \) = 2.78 (dotted red); the expression for the CDF is given in the Supplementary Document.

The following corollary gives the minimum risk ME model for nonnegative range \( \mathcal{R} = [0, \infty) \).

**Corollary 3.** The ME model in (35) on nonnegative range is the truncated continuous AL distribution with the following piecewise PDF:

\[
 f_{\theta, \eta_1}^{\mathcal{R}}(x) = \begin{cases} 
 \frac{\alpha \lambda c_0}{1 - e^{-\lambda c_0 \eta}} e^{-\lambda c_0 (x - \eta)} & (0 \leq x \leq q) \\
 (1 - \alpha) \lambda c_0 e^{-\lambda c_0 (x - q)} & (x > q),
\end{cases}
\]

where the Lagrange multipliers for the risk and Ql constraints are given, respectively, by

\[
 \theta = 1 - \frac{\alpha c_0 \eta_1 e^{\lambda c_0 q}}{1 - e^{-\lambda c_0 \eta}}, \quad \xi = \log \left( 1 - e^{-\lambda c_0 \eta} \right).
\]

Due to the truncation of the PDF in Proposition 3 for the nonnegative range the normalizing factor for the first branch has been adjusted which makes the PDF (38) discontinuous at \( x = q \). Also the second term in the expression for risk and the negative expression for \( \xi \) have been adjusted. The model for the finite range can be adjusted similarly.

For \( \lambda = 1/(c_0 + c_0) \), (36) gives an AL distribution used by Poiraud-Casanova and Thomas-Agnan (2000) to prove a property of a quantile regression estimator. For \( \lambda c_0 = \eta_1 \) and \( \lambda c_0 = \eta_2 \), the PDF (36) gives the known AL distribution (41). Kotz, Kozubowski, and Podgórski (2001) have extensively studied this model, including an ME characterization for a version with \( \tau = 0 \) in terms of the following constraints: \( E_f(X) = \mu, \ E_f[X] = \theta, \ |\mu| \leq \theta \). We give the elaboration of this AL distribution by QI as follows.

We first give a new ME characterization of this well-known AL model. Consider the class of PDFs with the following equivalent representations:

\[
\Omega_{\mu, \mu_1, \tau} = \left\{ f : E_f(X) = \mu, \ E_f[X \leq \tau] = \mu_1 \right\},
\]

\[
\begin{align*}
\Omega_{\theta_1, \theta_2, \tau} &= \left\{ f : E_f[\tau - X] 1(X \leq \tau) = \theta_1, \ E_f[(X - \tau) 1(X > \tau)] = \theta_2 \right\},
\end{align*}
\]

where \( \theta_1 = F(\tau) (\mu - 1) \), \( \theta_2 = F(\tau) (\mu_2 - \tau) \), and \( \mu = F(\tau) \mu_1 + F(\tau) \mu_2 \).

**Lemma 2.** The ME model in (40) is an AL distribution with the following absolutely continuous PDF

\[
 f_{\theta_1, \theta_2, \tau}^{\mathcal{R}}(x) = \frac{1}{1/\eta_1 + 1/\eta_2} \left[ e^{-\eta_1 (x - \tau)} 1(x \leq \tau) + e^{-\eta_2 (x - \tau)} 1(x > \tau) \right],
\]

where the Lagrange multipliers for the constraints in (40) are given by \( \eta_1 = 1/(\theta_1 + \sqrt{\theta_1 \theta_2}) \) and \( \eta_2 = 1/(\theta_1 + \sqrt{\theta_1 \theta_2}) \). The constraints in \( \Omega_{\theta_1, \theta_2, \tau} \) can be written as \( E_f[\mathcal{X} X \leq \tau] = \tau - \theta_1 \), \( E_f[(\mathcal{X} X > \tau)] = \theta_2 - \tau \). Asadi et al. (2014) referred to these as “local moments” and showed that the ME model that satisfies these constraints on the nonnegative range is piecewise exponential. The truncated exponential on \( 0 \leq x \leq \tau \) can be increasing, decreasing, or uniform (see Fig. 2), but on the unrestricted range \( f_{\theta_1, \theta_2, \tau}^{\mathcal{R}} \) must be increasing on \( 0 \leq x \leq \tau \) which gives an AL distribution.

Letting \( \eta_1 = 1/(\sigma \kappa) \) and \( \eta_2 = \kappa / \sigma \) in (41) gives a well-known AL distribution, where \( \sigma \) is the scale parameter and \( \kappa = (\theta_1 / \theta_2)^{1/2} \) is the asymmetry parameter (Kotz et al., 2001). Bera, Galvao, Mobte-Rojas, & Park (2016) related this formulation to the quantile regression estimation. In general, (41) does not satisfy a given Ql, since

\[
F(\tau) = \frac{\sqrt{\theta_1}}{\sqrt{\theta_1} + \sqrt{\theta_2}}.
\]

Lemma 2 provides an ME formulation suitable for the elaboration of (41) by the QI constraint in the following class of distributions:

---

*Fig. 3. Plots of the PDFs and CDFs of three ME models given by Proposition 3 (\( \theta = 5, q = 20 \)). (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)*
\[
\Omega_{\theta_1, \theta_2, \alpha} \\
= \left\{ f : \int f(q-x) \mathbf{1}(X \leq q) \, dq = \theta_1, \int f(q) \mathbf{1}(X > q) \, dq = \theta_2, \int f(X \leq q) \, dq = \alpha \right\} 
\]

(43)

The first two constraints in (43) can be represented by the risk constraint in (35) where \( \theta = c_d \theta_1 + c_u \theta_2 \), and one of the local moments, for example, \( \int f(q-x) \mathbf{1}(X \leq q) \, dq = \theta_1 \). \( \int f(q) \mathbf{1}(X > q) \, dq = \theta_2 \), \( \int f(X \leq q) \, dq = \alpha \).

The following corollary gives the ME model in this class.

**Corollary 4.** The ME model in \( \Omega_{\theta_1, \theta_2, \alpha} \) is a continuous AL distribution with the following piecewise PDF:

\[
f_{\theta_1, \theta_2, \alpha}(x) = \alpha \lambda_1 e^{-\lambda_1 q - \lambda_2} \mathbf{1}(x \leq q) + (1 - \alpha) \lambda e^{-\lambda x - \lambda_2} \mathbf{1}(x > q), \quad \lambda, \zeta > 0, 
\]

(44)

where the Lagrange multipliers for the risk and QI constraints in (43) are given by

\[
\lambda_1 = \frac{\alpha}{\theta_1}, \quad \lambda_2 = \frac{1 - \alpha}{\theta_2}, \quad \zeta = \log \left( \frac{1 - \alpha}{\alpha} \right) \lambda_2.
\]

Corollary 4 gives the ME model for unrestricted range. The ME model in (43) for nonnegative demand can be obtained similarly.

### 4. Application examples

#### 4.1. ME priors

Eliciting quantiles is known to be an effective approach for developing Bayesian priors. The QI information is usually used to fit a convenient parametric distribution. We illustrate ME priors with QI. A commonly used variable interval method is bisection elicitation. Faucett [Lecture Notes http://www.mas.ncl.ac.uk/~nfl8/teaching/mas2317/notes/chapter3.pdf] uses a scenario to illustrate the bisection elicitation method for eliciting a prior distribution for the rate, \( \beta \), of retreat (per foot) of the Zachariae Isstrøm glacier in Greenland. The range [0, 70] is bisected into two intervals of equal probability for eliciting the expert’s median. Then the expert is asked to further bisect each of these intervals for eliciting the lower and upper quartiles. The process resulted in \( q_{.25} = 19, q_{.50} = 24 \), and \( q_{.75} = 30 \). Faucett used these quartiles for computing the parameters of the gamma conjugate prior (shape \( r = 9 \) and scale \( s = 36 \)) for a Pareto likelihood model. Fig. 4 depicts the PDF and CDF of this gamma and the PDF of the ME piecewise uniform priors (27) for the median and for all elicited information. The uniform ME reference prior is also included.

Table 1 presents the elicited information provided by the median and the three quartiles, and by the additional non-elicited information induced by the gamma prior. The information value, \( K_{\theta} \), is computed by (8) with \( f^* \) being the ME prior with less constraint. The information value of the median relative to the uniform reference corresponds to the information provided by experimental results of a biased coin with probability.66 against the fair coin, which is rather informative. The information value of the three quartiles, given by \( K_{\theta_0} \), is comparable with a biased coin with probability.86, which is substantial. The non-elicited information, \( K_{\theta_2} \), is computed by (8) with \( f^*(x) \) being the fitted gamma prior is in \( \Omega_{\theta} \), in an approximate sense (its support is \( x > 0 \) but probability over the range [0,70] is greater than.9999 and it approximately satisfies the quartile constraints). This is comparable with the information divergence of a biased coin with probabilities.87 from a fair coin. The entropy reduction of the gamma prior relative to the ME QI prior \( K_{\theta_2} \) is about the same as the information provided by the three quartiles. That is, this fitted prior induces non-elicited information as much as the elicited quartiles \( K_{\theta_0} \). This is particularly important as it underscores the amount of extraneous information induced by the choice of this prior.

Suppose that one also elicits the mean, \( E(\beta) = 25 \), in the above problem; (this is the mean of the fitted gamma prior). The upper

![Fig. 4. ME prior consistent with the range, the median, three quartiles (QI), and the gamma prior with parameters chosen to satisfy the quartiles.](image-url)
row of Fig. 5 shows the PDFs and CDFs of the ME priors with the mean, the median, and both. The CDFs of these priors are continuous. The ME prior with the mean is the truncated exponential and the prior with the mean and median is the two-piece truncated exponential. Table 2 presents the information analysis of these priors (rows $i = 0, 1, 2, 3$). The mean relative to the uniform reference is more informative than the median. The information value of the mean relative to the ME with the median corresponds to a biased coin with probability $71\%$ against the fair coin, which is not negligible and the information value of the median relative to the ME with the mean corresponds to a biased coin with probability $67\%$.

The lower row of Fig. 5 shows the PDFs and CDFs of the ME priors with the mean, the three quartiles, and both. Plots for the gamma prior fitted to the quartiles are also shown. Table 2 presents the information analysis of these priors. Comparison of $K_{30}$ and $K_{20}$ indicates that the mean is more informative than the median. The indices $K_{31}$ and $K_{32}$ indicate that the mean provides more information over the median than the median over the mean. Comparison of $K_{50}$ and $K_{40}$ indicates that the three quartiles are substantially more informative than the mean and median jointly. The indices $K_{53}$ and $K_{54}$ indicate that the first and third quartiles provide more information over the mean and median than the mean over the three quartiles. Information analysis of the fitted gamma indicates that it induces substantial non-elicited information over the elicited information, which corresponds to a biased coin with probability $85\%$; $K_{65}$ is calculated using (8).

**Table 2**

Information analysis of distributions shown in Fig. 5.

| $i$ | Prior (Fig. 5 row) | Constraints | Information relative to uniform | Information relative to other ME |
|----|-------------------|-------------|---------------------------------|--------------------------------|
| 0  | Uniform (both)    | QI          | $E(X)$                          | $K_{ij}$ $ID$ $\pi$             |
| 1  | Median (upper)    | 24          | 4.20                            | $K_{30}$ $0.10$ $0.18$ $0.71$   |
| 2  | Mean (both)       | 24          | 4.16                            | $K_{32}$ $0.06$ $0.11$ $0.67$   |
| 3  | Mean & median (upper) | 24 25  | 4.10                           | $K_{31}$ $0.10$ $0.18$ $0.71$   |
| 4  | Three quartiles (lower) | 19 24 30 | 3.89                           | $K_{51}$ $0.25$ $0.40$ $0.81$   |
| 5  | Mean & three quartiles (lower) | 19 24 30 25 | 3.85                        | $K_{54}$ $0.05$ $0.10$ $0.66$   |
| 6  | gamma (lower)     | 19 24 30 25 | 3.50                         | $K_{65}$ $0.35$ $0.50$ $0.85$   |

![Fig. 5](image-url) ME models consistent with the range, only QI (median, upper panels, three quartiles (QI), lower panels), only the mean, both types of constraints, and the gamma model with parameters chosen to satisfy the quartiles (lower panels).
Table 3
The subjective probabilities (in percentage) assigned by four SPF respondents in the first and second quarters (Q1 and Q2) of 2020 to intervals of the U.S. GDP in 2020.

| Forecaster | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 |
|------------|----|----|----|----|----|----|----|----|----|
| 421        | 2  | 5  | 45 | 45 | 3  |     |    |    |    |
| 504        | 1  | 3  | 5  | 27 | 50 | 10 | 1  |     |    |
| 563        | 10 | 50 | 35 | 24.40 | 46.60 | 20.92 | 3.44 | .52 | .12 |
| 588        | 18 | 55 | 17 | 2  |     |    |    |    |    |
| Pooled     | 0.25 | 1.00 | 1.82 | 4.68 | 28.60 | 49.15 | 12.73 | 1.61 | 0.13 | 0.03 |

4.2. Survey of Professional Forecasters

In the middle of each quarter, the Federal Reserve Bank of Philadelphia surveys experts to solicit subjective forecast probabilities for the GDP growth during the current year and in the following year. The Survey of Professional Forecasters (SPF) offers a set of m fixed intervals (bins) for the experts to assign probabilities. Since 1992, SPF offers ten bins, where the lower and upper bins are unbounded. The SPF also includes a question asking for the point forecast of the respondent. The respondents’ point forecasts cannot be simply interpreted as the means of their forecast distributions (Engelberg, Manski, & Williams, 2009).

The subjective probabilities are used to estimate the individual forecast distributions \( f_i, i = 1, \ldots, n \) and the arithmetic pool of the individual probabilities (consensus distribution with PDF \( f_p \)) as defined in (14), with \( w_i = 1/n \). A considerable amount of attention is given to the variance decomposition (16) for \( \sigma^2_p \). Both, \( \sigma^2_p \) and the average of \( \sigma^2_i \), are used as measures of uncertainty and variance of the means (last term in (16)) is used as a measure of disagreement. Various variance estimates are used, including the variance of the midpoints of the bins \( x_k = (x_k + x_{k+1})/2, k = 1, \ldots, m \), this with the Sheppard correction 1/12, the variance of the normal fit to histogram, and the variance of beta distribution fitted to the histogram.

Information measures also have been used for measuring uncertainty and disagreement. Rich and Tracy (2010) used the discrete entropy of the probabilities \( H(P_{i}) > 0 \), but did not use any information divergence for disagreement. Shoja and Soofi (2017) proposed an information framework for measuring uncertainty and disagreement. This framework requires a set of individual forecast distributions, specifying weights for their arithmetic pooling, and measures of uncertainty and divergence. These authors used the discrete entropy, \( H(P_{i}) \), and the ME models based on the variance of the midpoints to illustrate their proposed framework. Lahiri and Wang (2019) avoided using discrete entropy and applied the information framework of Shoja and Soofi (2017) using the entropies of beta and triangular distributions fitted to the histograms of forecasters’ subjective probabilities and their average; they estimated bounds for the unbounded intervals according to the fit. The use of fitted parametric distributions, such as the normal and beta, induces non-elicited information, so does the use of their variance. In addition, a parametric model fitted to the pooled quantities is not a mixture of the models fitted to the elicited quantities, which is a pool of the models fitted for the individuals.

We use the ME model (27) to measure uncertainty, disagreement and information of economic forecasters within the framework of Shoja and Soofi (2017). The subjective probabilities of each forecaster provide an ME model \( f_{i} = f_{i}^{*} \), \( i = 1, \ldots, n \). By Proposition 2, \( f_{i}^{*} \) is the ME forecast model based on the pool of subjective probabilities of forecasters. Then \( H(f_{i}^{*}) \) and (18) with \( f_{i} = f_{i}^{*} \) provide two uncertainty measures for the set of forecasters and (17) gives the disagreement measure for the set of forecasters. Following the literature, we use \( w_i = 1/n \).

We first illustrate the implementation of the information framework with the ME model (27) using data given by four respondents of the SPF in the first and second quarters of 2020 for the U.S. GDP growth in 2020 and then apply to all SPF respondents (2019-2020Q2) who provided probabilities for the 2020 U.S. GDP growth. Table 3 gives the probabilities assigned by the four forecasters selected to represent typical forecasters of the first two quarters of 2020 (before and during the coronavirus pandemic). The last row gives the pooled elicited information.

Fig. 6 shows the plots of the ME forecast distributions for these forecasters, superimposed by the pooled ME forecast model which is derived using the pooled QI constraints according to Proposition 2. The unbounded bins are assigned \( B_{1} = B_{10} = (q_{1} + q_{9})/2 = 4\% \). The plots in the first row are for the first quarter of 2020 and the second row shows the corresponding plots for the second quarter. Shifts of all distributions toward the lower side are apparent.

The uncertainty of individual forecasters and disagreement among them are tabulated below the plots of the ME models. These measures are computed using (19) for the individual \( i \) and pooled distributions within the information framework of Shoja and Soofi (2017). For each forecaster, \( H_i = H(f_{i}^{*}) \), maps the extent of its concentration and quantifies the uncertainty of the ME model (solid blue). The KL divergence \( K_{i} = K(f_{i}^{*} : f_{i}^{*}) \) maps the extent to which the individual ME model disagrees with the pooled ME model for the quarter (dashed red); \( \pi(K_{i}) \) index gives the probability according to the biased coin calibration for the KL divergences \( K_{i} \). These measures indicate that in most cases the uncertainty is decreased in the second quarter and in all cases the disagreement is sharply increased. The pooled uncertainty \( H_{c} \) and disagreement \( J_{SC} \), are averages of \( H(f_{i}^{*}) \) and \( K(f_{i}^{*} : f_{i}^{*}) \). For the quarter pool, \( H(f_{p}^{*}) \) and the average of four individual entropies, \( H_{c} \), are measures of overall uncertainty. The disagreement information, \( J_{SC} \), is the average of four individual divergences, which is also the difference between the two measures of uncertainty for the pool. These measures confirm that, overall, the uncertainty of these four forecasters is decreased in the second quarter and their disagreement is sharply increased. The coronavirus pandemic might have played a role in these changes.

Next we present application to all 2019 and 2020 quarterly forecasters of the 2020 U.S. GDP in the SPF database. Table 4 summarizes the information analysis of quarterly data for all 2019 and 2020 forecasters. For the unbounded bins we set \( B_{1} = B_{10} = 4\% \). The sensitivity of entropies to the choice of the lower and upper bins, \( B_{1} \) and \( B_{10} \), can be easily assessed using \( \tilde{p}_{1} \log(B_{1}/4) + \tilde{p}_{10} \log(B_{10}/4) \)
referred

Table

| Data         | Forecast | Qf measures | H(\hat{f}_c) | H_c | JS_c | \pi(JS_c) |
|--------------|----------|-------------|-------------|-----|------|----------|
| Quarter      | horizon  | No.         | H(f_c)      | H_c | JS_c | \pi(JS_c) |
| 2019Q1       | 7.5      | 31          | 1.615       | 1.355 | .260  | .82      |
| 2019Q2       | 6.5      | 34          | 1.578       | 1.321 | .257  | .82      |
| 2019Q3       | 5.5      | 30          | 1.583       | 1.369 | .214  | .80      |
| 2019Q4       | 4.5      | 30          | 1.571       | 1.343 | .228  | .80      |
| 2020Q1       | 3.5      | 30          | 1.313       | 1.143 | .170  | .77      |
| 2020Q2       | 2.5      | 34          | 1.619       | 1.266 | .353  | .86      |

\( \hat{p}_{10} \log(B_{10}/4) \). The average probabilities of the unbounded bins are minute which make the entropy practically insensitive to the choice of the bins. The following points are evident from Table 4.

(a) During 2019 quarters the uncertainty of the pooled model \( H(f_c) \) remains stable while the pooled uncertainty \( H(f_c) \) varies a little bit more, resulting in varying the information/disagreement \( JS \) through the quarters. The disagreement is stronger in the first half of 2019, where the forecast horizon is longer, as compared with the second half (shorter forecast horizon).

(b) In the first quarter of 2020, the uncertainty and disagreement of the forecaster decrease; a shorter horizon might have played a role. But in the second quarter of 2020, the uncertainty sharply and disagreement modestly increase. In spite of a shorter horizon these changes can be attributed to the coronavirus pandemic.

The SPF provides information about the forecaster’s industry. Table 5 reports the uncertainty and disagreement according to the forecaster’s firm. The uncertainty of the forecaster from the financial service providers is higher and their disagreement is lower than respective measures for forecasters from other industries. The time patterns of the measures are similar to Table 5. (R codes for computing Tables 4 and 5 are available upon request from the authors.)

4.3 Reliability and inventory management

Two truncated distributions below and above a threshold, \( \tau \), on the support of a distribution are prominent in reliability modeling, where \( \tau \) represents the current age of an item and in inventory management where \( \tau \) represents the inventory. We consider the case when \( \tau = q \) is a quantile of the ME PDF \( f_{q,\alpha}^{*} \), on the entire support. Then the truncated distributions in (26) are:

\[
\begin{align*}
    f_1(x) &= \frac{f_{q,\alpha}^{*}(x)}{F_{q,\alpha}^{*}(q)} \mathbf{1}(x \leq q), \\
    f_2(x) &= \frac{f_{q,\alpha}^{*}(x)}{F_{q,\alpha}^{*}(q)} \mathbf{1}(x > q).
\end{align*}
\]

where \( F_{q,\alpha}^{*}(q) = 1 - \alpha \) is the survival function at the optimal order. In reliability modeling, (45) and (46) correspond to the past-life or down-time and the residual distributions at a given time \( q \). The “mean residual” is also used in inventory problems. In inventory analysis, (45) is the distribution of sales (the conditional distribution of the demand, given that demand does not exceed the inventory), and (46) is the distribution of stockout (the conditional distribution of the demand, given that demand exceeds the inventory). We present ME models useful for inventory problems.

A widely studied topic in inventory management is the so-called newsvendor (NV) problem, which is defined by an
optimal order quantity for a product with an uncertain demand $X$ and fixed prices. Let $c$ be the cost per unit, $r$ the selling price per unit, and $s$ the salvage value per unit, then $c_0 = c - s$ and $c_v = r - c$ are the overage and underage costs, respectively. Assuming, $s < c < r$, the expected profit maximizing (loss minimization) solution for the optimal quantity order is the $\alpha$th quantile of the demand distribution where $\alpha = c_v/(c_0 + c_v)$. This solution is traditionally represented in terms of the $\beta$th upper quantile $\beta = 1 - \alpha$. This optimization, however, does not offer a model for the unknown demand distribution, $F$, and without a distribution $q$ remains unknown.

The NV optimal decision theoretic decision under the loss function (33) coincides with NV decision based on the expected profit maximization; see for example, Snyder & Shen (2011, p. 78). Thus, the new ME Asymmetric Laplace model (36) provides a demand distribution consistent with the optimal order for the basic NV problem. Corollary 4 also provides a model for the NV problem. In (39), $\mu_1$ is the average sales which can be assessed based on the sales data. This and the usual assumption of given mean demand, $\mu$, provide the partial information about the demand distribution. In (40), the constraints represent the mean inventory, the mean stock-out, and the profit maximizing quantile.

The NV literature offers moment-based rules for computing optimal order, such as the Scarf (1958)’s rule and the minimax regret rule (Perakis & Roels, 2008). These rules include the cost ratio, $\alpha$, but do not offer models for the demand distribution. Instead, well-known probability models are then chosen for the demand distribution as a separate task. The classical ME models (uniform, exponential, and normal) are used for the distribution of the demand and justified based on the ME principle (Andersson et al., 2013; Perakis & Roels, 2008). The moment based ME models are said to be non-robust to various degrees because their $\alpha$th quantiles are different from the optimal solutions of the moment-based rules. Such discrepancies between the optimal order quantities and the $\alpha$th quantiles of moment-based ME models are due to the lack of inclusion of the QI in the ME calculations. The ME models with moments and quantiles resolve this discrepancy.

We illustrate applications of Corollaries 1 and 2 for two moment-based optimal order rules, one is based on the mean and one is based on the mean and variance. The minimax regret rule, $y^*_\alpha$, depends on the partial information. With a given mean on the nonnegative range $y^*_\alpha = \mu(x, \alpha = \xi, \alpha \leq 1)$, $y^*_\alpha = \mu(1 - \alpha, \alpha \geq 1)$. For the case of the first two moments on the nonnegative range, an approximate minimax regret rule (Roels, 2006) is

$$y^*_\alpha \approx \max \left\{0, \mu + \frac{2(2\alpha - 1)\sigma}{5\sqrt{\alpha(1 - \alpha)}}, \frac{\sqrt{x} - \beta}{\sqrt{\alpha(1 - \alpha)}} \right\}, \quad \text{for} \quad cv \leq \sqrt{\alpha}.$$

The Scarf’s exact rule for the optimal order quantity based on the mean $\mu$ and the variance $\sigma^2$ is

$$y^*_\alpha = \left\{ \begin{array}{ll}
\mu + \frac{(2\alpha - 1)\sigma}{2\sqrt{\alpha(1 - \alpha)}}, & \text{if } cv \leq \sqrt{\alpha} \\
0, & \text{otherwise}
\end{array} \right.$$
These rules give $\psi^* = 3.49$ and $\eta^* = 3.42$. With $\psi^* = 3.49$, the ME model is $TN(3.77, 1.11)$, where $\mu$ and $\sigma$ are the mean and standard deviation of its parent normal distribution. The first two moments of $f_{\alpha \psi^* \eta^*}(k) = 1.2, \theta_1, b = 1.2, \theta_2 = (1.65, 3.54) \in A_1$ and $(\theta_2, \theta_2) = (4.91, 25.7)$ which give $c_{uv} = 0.26$; (the moments are computed using $(\mu_4, \sigma_4) = (3.49, 1.16)$ in $(S.1)$ and $(S.2)$ in the Supplementary Document). Part (b) of Corollary 2 with $c_v < 1$ gives a continuous distribution with a two-piece PDF: a truncated normal on [0, $q$] and a truncated normal on $x \geq q$. The CDF plot illustrates the lack of robustness of the truncated normal model for this rule.

Table 6 gives the information values of QI when the uniform, exponential, and truncated normal distributions are adjusted to satisfy the maximizing profit quantiles as well as the minimax regret or Scarf’s rules. The Scarf’s rule for the uniform ME model on a finite range, $[a, b]$, gives $\psi^* = 0$ for $\alpha \leq 0.25$ and $\psi^* < F^{-1}(\alpha)$ for $\alpha < 0.5$ (for $\alpha < 0.5$, $\psi^* > F^{-1}(\alpha)$). The minimax regret rule when the partial information is a finite range $[a, b]$ gives $\psi^* = (1 - \alpha)a + \alpha b$ which coincides with $F^{-1}(\alpha)$ (Perakis & Roels, 2008).

In Table 6 the uniform distribution is on [0, 70] and $\alpha = 0.75$. The minimax regret rule gives $\psi^* = 75$ and there is no information value for the QI of this rule. The Scarf’s rule gives $\psi^* = 46.67$ and the information gain of QI over the uniform model is comparable to the prediction of outcomes of a coin with probability,83, indicating that the QI of this rule is very informative. Notice that, here we have a single QI constraint and QI is nearly as informative as the four quantiles information in Table 1. The information values of QI constraint with the minimax regret rule $q = \psi^*$ and Scarf’s rule $q = \psi^*$ in addition to the mean (Fig. 7), are comparable to the prediction of outcomes of a coin with probability,59 and 69, respectively. Accordingly, QI is rather informative indicating that these rules are not robust for the exponential model, however, the minimax regret improves over the Scarf’s rule. The information values of QI constraint with $q = \psi^*$ and Scarf’s rule $q = \psi^*$, in addition to the first two moments (Fig. 7), are comparable to the prediction of outcomes of a coin with probability,62 and 53, respectively. Accordingly, QI with the minimax regret is rather informative and the Scarf’s rule is rather robust under the truncated normal for the demand.

5. Summary and discussion

This paper presented several results and an assortment of ME models with given quantiles with or without moments and illustrated their potential applications in various problems. The results explored properties, existence, and shapes of models for the minimum information elaborations of the uniform and moment-based ME models by quantiles, and provided diagnostics for assessment of the utility of the minimal elaborations. The application examples illustrated the merits of the results by explicating problems where non-elicited information is induced in the assumed models and information that appears in the analysis is not included in the assumed ME models. The results for the ME model with quantile and moments information enabled us to determine the shape and existence of the ME model, compute its entropy rather easily, and assess the information utility of the elaboration. Specific models provided elaborations of the uniform, exponential, normal, Laplace, and a known asymmetric Laplace distribution. The elaboration of the Laplace distribution is along the lines of Ardakani, Ebrahimí, and Soofi (2018) that explored the link between the information theory and Laplace’s first and second laws of error in terms of the minimum risk ME models under the absolute error and quadratic loss functions (Laplace and normal). This extension provides a gateway to further research on developing ME models with minimum risks of other asymmetric loss functions such as Linex loss which is used in various problems.

The ME distributions with quantile information are continuous, but their PDFs are not, so they do not provide the convenience of parametric priors. However, ME models enable assessing the non-elicited information that is induced by assuming parametric models that fit the elicited quantiles. From an example we learned that a gamma distribution fitted to quantiles elicited for a Bayesian prior induces non-elicited information as much as the elicited quantiles. The example also compared information values of eliciting median, quartiles, and the mean.

Application of the ME model with QI elicited by fixed interval method to the Federal Reserve Bank of Philadelphia’s SPF revealed the effect of the coronavirus outbreak on the uncertainty and disagreement of the economic forecasters. The results informed that disagreement among forecasters about the U.S. 2020 GDP growth increased as time evolved until the coronavirus outbreak. After the outbreak, the uncertainty and disagreement of forecasters increased sharply above the level of the first quarter of 2020. An industry level analysis indicated that disagreement among forecasters in financial services is lower than those who are in other industries.

Our results shed light on an issue in the inventory management literature. A demand distribution is said to be (perfectly) robust if its profit maximizing quantile is consistent with a moment-based rule for an optimal order quantity. Achieving this property requires inclusion of this important quantile within the partial information set used for developing the rule. This is illustrated using elaborations of the uniform, exponential, and normal distributions by the profit maximizing quantile.

This paper considered only the case of continuous random variables. The discrete case can be illustrated similarly. The biased coin calibration of the KL divergence has been supplied so that the tabulated disagreement measures can be judged sizeable or not. The Bayesian inferences about the entropy and KL divergence can be used for more formal inferences.
Acknowledgments

The authors are thankful to two reviewers for their comments that led to substantial improvement of the exposition of this paper. Soofi’s research was supported by the 2019 Business Advisory Council Faculty Scholar Award.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2020.07.052.

References

Alwan, L. C., Ebrahimi, N., & Soofi, E. S. (1998). Information theoretic framework for statistical process control. European Journal of Operational Research, 111(3), 526–542.

Andersson, J., Jönsten, K., Nonså, S. L., Sandal, L., & Ubej, J. (2013). A maximum entropy approach to the newsvendor problem with partial information. European Journal of Operational Research, 228(1), 190–200.

Ardakan, O. M., Ebrahimi, N., & Soofi, E. S. (2018). Ranking forecasts by stochastic error distance, information, and reliability measures. International Statistical Review, 83, 442–468.

Asadi, M., Ebrahimi, N., & Soofi, E. S. (2018). Optimal hazard models based on partial information. European Journal of Operational Research, 270(2), 723–733.

Asadi, M., Ebrahimi, N., Soofi, E. S., & Zarezadeh, S. (2014). New maximum entropy methods for modeling lifetime distributions. In Naval research logistics: 61 (pp. 427–434).

Bera, A. K., Galvao Jr., A. F., Montes-Rojas, G. V., & Park, S. Y. (2016). Asymmetric Laplace regression: Maximum likelihood, maximum entropy and quantile regression. Journal of Econometric Methods, 5(1), 79–101.

Berger, J. O. (1985). Statistical decision theory and Bayesian analysis (2nd ed.). New York: Springer-Verlag.

Box, G. E. P., & Tiao, G. C. (1973). Bayesian inference in statistical analysis. New York: John Wiley.

Brockett, P. L., Charnes, A., & Paick, K. H. (1995). Information-theoretic approach to unimodal density estimation. IEEE Transactions on Information Theory, 41(3), 624–629.

Carota, C., Parmigiani, G., & Nicholas, G. P. (1966). Diagnostic measures for model criticism. Journal of the American Statistical Association, 61(345), 753–762.

Di Crescenzo, A., & Longobardi, M. (2002). Entropy-based measure of uncertainty in joint lifetime distributions. Journal Applied Probability, 39, 434–440.

Dowson, D. C., & Wragg, A. (1973). Maximum-entropy distributions having prescribed first and second moments. IEEE Transactions of Information Theory, 19(5), 689–693.

Ebrahimi, N., Soofi, E. S., & Soyer, R. (2008). Multivariate maximum entropy identification, transformation, and dependence. Journal of Multivariate Analysis, 99, 1217–1231.

Engelberg, J., Manski, C. F., & Williams, J. (2009). Comparing the point predictions and subjective probability distribution of professional forecasters. Journal of Business and Economic Statistics, 27, 30–41.