

MINIMUM DYNAMIC DISCRIMINATION INFORMATION MODELS

MAJID ASADI,* *University of Isfahan*

NADER EBRAHIMI,** *Northern Illinois University*

G. G. HAMEDANI,*** *Marquette University*

EHSAN S. SOOFI,**** *University of Wisconsin–Milwaukee*

Abstract

In this paper, we introduce the minimum dynamic discrimination information (MDDI) approach to probability modeling. The MDDI model relative to a given distribution G is that which has least Kullback–Leibler information discrepancy relative to G , among all distributions satisfying some information constraints given in terms of residual moment inequalities, residual moment growth inequalities, or hazard rate growth inequalities. Our results lead to MDDI characterizations of many well-known lifetime models and to the development of some new models. Dynamic information constraints that characterize these models are tabulated. A result for characterizing distributions based on dynamic Rényi information divergence is also given.

Keywords: Lifetime distribution; residual life distribution; monotone density; failure rate dominance; uncertainty ordering

2000 Mathematics Subject Classification: Primary 94A15

Secondary 60E15; 60B05

1. Introduction

Laplace's principle of insufficient reason says to distribute probability uniformly in the absence of any constraint on the probabilities. The maximum entropy (ME) principle extends this to the production of probability models close to uniform, which are least sensitive to information other than that explicitly taken into account via some moment constraints (Jaynes (1957), (1982)). The minimum discrimination information (MDI) or minimum cross-entropy principle is a generalization of the ME principle for the development of models close to any given distribution, instead of the uniform distribution (Kullback (1959), Shore and Johnson (1980)). Recently, a maximum dynamic entropy (MDE) procedure for developing lifetime models has been proposed. This may be viewed as an extension of the ME principle in the case that the information is given in terms of hazard rate growth inequality constraints (Asadi *et al.* (2004)). The ME and MDI procedures (i.e. those based on the corresponding principles) use the calculus of variations, while the MDE procedure uses differential inequalities and hazard ordering. In this

Received 21 July 2004; revision received 6 January 2005.

* Postal address: Department of Statistics, University of Isfahan, Isfahan, 81744, Iran.

Email address: m.asadi@sci.ui.ac.ir

** Postal address: Division of Statistics, Northern Illinois University, DeKalb, IL 60155, USA.

Email address: nader@math.niu.edu

*** Postal address: Department of Mathematics, Statistics and Computer Science, Marquette University, PO Box 1881, Milwaukee, WI 53201-1881, USA. Email address: hossein@mscs.mu.edu

**** Postal address: School of Business Administration, University of Wisconsin–Milwaukee, PO Box 741, Milwaukee, WI 53201, USA. Email address: esoofo@uwm.edu

paper, we introduce the minimum dynamic discrimination information (MDDI), or dynamic minimum cross-entropy, approach to probability modeling. In this procedure, models are derived using simple residual moment inequalities, differential residual moment inequalities, as well as differential hazard rate inequalities. This generalizes the MDE procedure.

Closeness between two distributions F and G with nonnegative supports is measured by the Kullback–Leibler discrimination information

$$K(f : g) = \int_0^\infty f(x) \log \frac{f(x)}{g(x)} dx, \tag{1}$$

where f is the probability density function (PDF) of F , g is the PDF of G , and F is absolutely continuous with respect to G . The discrimination information $K(f : g)$ is greater than or equal to 0, but it is not symmetric, and G is referred to as the *reference distribution*.

An important application of (1) is Kullback’s MDI principle for probability modeling and statistical inference. The MDI principle of modeling considers the moment class of distributions

$$\Omega_{F_\theta} = \{F_\theta : E_f[T_j(X) | \theta] = \theta_j, j = 0, 1, \dots, J\}, \tag{2}$$

where X is a random variable with PDF f , $T_j(X)$ are functions integrable with respect to f with $T_0(X) = \theta_0 = 1$, and $\theta = (\theta_1, \dots, \theta_J)$ is a vector of moment parameters. The MDI model $F^* \in \Omega_{F_\theta}$ with reference to G has the PDF f^* that minimizes $K(f : g)$.

Recall that the MDI principle is also referred to as the minimum cross-entropy principle. The Shannon entropy of a distribution F (Shannon (1948)) is defined by

$$H(f) = - \int_0^\infty f(x) \log f(x) dx.$$

The ME model is the distribution whose PDF f^* maximizes $H(f)$ subject to the set of constraints in (2). When G is uniform, the MDI and ME procedures are equivalent.

When F is a lifetime distribution and the current age of the item must be taken into account, the set of interest for studying information is not the entire support; rather, the set of interest is the *residual lifetime* of the item, $\{x : x > t\}$. The discrimination information function of two residual life distributions $F(x; t) = P_F(X \leq x | X > t)$ and $G(x; t) = P_G(X \leq x | X > t)$, corresponding to two lifetime distributions F and G , is given by

$$K(f : g; t) \equiv K[f(x; t) : g(x; t)] = \int_t^\infty f(x; t) \log \frac{f(x; t)}{g(x; t)} dx, \tag{3}$$

where $f(x; t) = f(x)/\bar{F}(t)$ and $g(x; t) = g(x)/\bar{G}(t)$ denote the residual PDFs for $x > t$, and $\bar{F}(t) = 1 - F(t)$ and $\bar{G}(t) = 1 - G(t)$ are the survival functions.

It is clear that for $t_0 = \inf\{t : \bar{F}(t) = 1\}$, we have $K(f : g; t_0) = K(f : g)$. By (3), for each t , $K(f : g; t)$ possesses all the properties of the discrimination information function (1). If we consider $\mathcal{T} = \{t : t > 0\}$ as an index set, then $K(f : g; t)$ provides dynamic discrimination information ranging over \mathcal{T} . The Shannon entropy of the residual lifetime distribution is defined similarly:

$$H(f; t) \equiv H[f(x; t)] = - \int_t^\infty f(x; t) \log f(x; t) dx.$$

Several authors have studied information functions that take age into account (Ebrahimi (1996), Ebrahimi and Kirmani (1996a), (1996b), Di Crescenzo and Longobardi (2002),

Belzunce *et al.* (2004), Asadi *et al.* (2004)). Consideration of age has led to some important insights about lifetime models, such as an information characterization of the proportional hazards model (Ebrahimi and Kirmani (1996a)) and MDE characterizations of various lifetime models, including some mixture distributions, for which no other maximum entropy formulation is available (Asadi *et al.* (2004)). Some tests of distributional hypotheses based on dynamic information measures have been developed for reliability analysis (Ebrahimi (1998), (2001)). When the subject of study is an interval other than a lifetime (e.g. a search time or an unemployment period) the present time point plays the role of ‘age’. More generally, dynamic measures are applicable to any continuous distribution with positive support. Examples include wage distribution, where the minimum wage is the source of the dynamics in the information measures, and the distribution of the amount of a natural resource such as oil, where the remaining amount in an oil field is of primary interest and the amount extracted introduces the dynamics into the information measures.

It is natural to think of the dynamic extensions of the MDI and ME principles in terms of the residual moments. The mean residual life function is defined by

$$\mu_F(t) = E[X - t \mid X > t] = \begin{cases} \frac{1}{\bar{F}(t)} \int_t^\infty \bar{F}(x) \, dx & \text{if } \bar{F}(t) > 0, \\ 0 & \text{if } \bar{F}(t) = 0. \end{cases}$$

However, the mean residual life uniquely determines the lifetime distribution; that is,

$$\bar{F}(t) = \frac{\mu_F(0)}{\mu_F(t)} \exp\left\{-\int_0^t \frac{1}{\mu_F(x)} \, dx\right\}. \tag{4}$$

Consequently, the MDI and ME procedures for developing models subject to residual moment equations do not extend to the dynamic case. However, developing lifetime models subject to residual moment inequalities is feasible.

In Section 2, we present results for developing MDDI and MDE models based on residual moment inequality constraints. In Section 3, we give results for MDDI modeling based on hazard rate growth inequality constraints, and also present a result that extends MDDI modeling to the case in which closeness is measured by the information divergence of order α (also known as Rényi information) between two distributions (Rényi (1961)).

2. Mean residual life constraints

Consider a set of distributions $\Omega_F = \{F\}$, where F has PDF f and is absolutely continuous with respect to a reference distribution G that has PDF g . The MDDI model in Ω_F relative to G is F^* , with PDF f^* such that $K(f^* : g; t) \leq K(f : g; t)$ for all $t \geq 0$. That is, among all the residual PDFs $f(x; t)$ induced by all members of Ω_F , the MDDI model $F^* \in \Omega_F$ is that whose residual PDF $f^*(x; t)$ retains its MDI property for all $t \geq 0$.

The following theorem gives the properties of the MDDI distributions in classes of distributions with mean residual life inequality constraints.

Theorem 1. *Let $\Omega_F = \{F : \mu_F(t) \leq q(t)\}$ be a compact set of distributions, where F is absolutely continuous with respect to a reference distribution G . Let $F^* \in \Omega_F$ be such that $\mu_{F^*}(t) = q(t)$. If $\log(f^*(x)/g(x))$ is decreasing and concave then F^* is the MDDI distribution relative to G . The same result holds, with $\Omega_F = \{F : \mu_F(t) \geq q(t)\}$, if $\log(f^*(x)/g(x))$ is increasing and convex.*

Proof. We will prove the case in which $\mu_F(t) \leq q(t)$. First, note that

$$\begin{aligned} K(f : g; t) &= \int_t^\infty f(x; t) \log \frac{f(x; t)}{g(x; t)} dx \\ &= \int_t^\infty f(x; t) \log \frac{f^*(x; t)}{g(x; t)} dx + \int_t^\infty f(x; t) \log \frac{f(x; t)}{f^*(x; t)} dx \quad (5) \\ &\geq \int_t^\infty f(x; t) \log \frac{f^*(x; t)}{g(x; t)} dx, \end{aligned}$$

where the inequality is due to the fact that the second integral in (5) equals $K(f : f^*; t) \geq 0$. This gives

$$\begin{aligned} K(f : g; t) - K(f^* : g; t) &\geq \int_t^\infty f(x; t) \log \frac{f^*(x; t)}{g(x; t)} dx - \int_t^\infty f^*(x; t) \log \frac{f^*(x; t)}{g(x; t)} dx \\ &= \int_t^\infty f(x; t) \log \frac{f^*(x)}{g(x)} dx - \int_t^\infty f^*(x; t) \log \frac{f^*(x)}{g(x)} dx \\ &\geq 0. \end{aligned}$$

Since $\log(f^*(x)/g(x))$ is decreasing and concave and $\mu_F(t) \leq q(t) = \mu_{F^*}(t)$, the last inequality follows from Theorem 3.A.13 of Shaked and Shantikumar (1994). The proof for $\mu_F(t) \geq q(t)$ is similar.

When G has a uniform PDF over $\{x : 0 < x < b\}$, the residual life PDF $g(x; t)$ is also uniform over $\{x : t < x < b\}$ and the MDDI model reduces to the MDE model. The MDE model in a set of distributions $\Omega_F = \{F\}$ is the distribution with PDF $f^*(x)$ such that $H(f; t) \leq H(f^*; t)$ for all $t \geq 0$.

Corollary 1. *Let $\Omega_F = \{F : \mu_F(t) \leq q(t)\}$ be a compact set of absolutely continuous distributions. Let $F^* \in \Omega_F$ be such that $\mu_{F^*}(t) = q(t)$. If $f^*(x)$ is increasing and log-concave then F^* is the MDE distribution. The same results holds, with*

$$\Omega_F = \{F : \mu_F(t) \geq q(t)\},$$

if $f^(x)$ is decreasing and log-concave.*

Applications of the above results include the identification of classes of distributions based on mean residual life inequality constraints, where the well-known models are MDDI and/or MDE, and the development of MDDI and MDE models for classes of distributions whose mean residual life functions are bounded (above or below) by a given function $q(t)$. Any continuous function $q(t)$ with the following properties is the mean residual life function of a distribution function F :

- $0 \leq q(t) < \infty;$ (6)

- $q(0) > 0;$ (7)

- $q'(t) + 1 > 0;$ (8)

- $\begin{cases} \text{if } q(t_0) = 0 \text{ for some } t_0 \text{ then } q(t) = 0 \text{ for all } t \geq t_0; \\ \text{if } q(t) > 0 \text{ for all } t > 0 \text{ then } \int_0^\infty (1/q(t)) dt = \infty. \end{cases}$ (9)

For any function satisfying (6)–(9), we can obtain a unique distribution given by (4). Then, for any reference distribution, the conditions of Theorem 1 can easily be verified by comparing the first and second derivatives of the two densities: $f(x)/g(x)$ is increasing or decreasing if

$$\frac{d \log f(x)}{dx} \geq \frac{d \log g(x)}{dx} \quad \text{or} \quad \frac{d \log f(x)}{dx} \leq \frac{d \log g(x)}{dx}, \tag{10}$$

respectively, and $\log(f(x)/g(x))$ is convex or concave if

$$\frac{d^2 \log f(x)}{dx^2} \geq \frac{d^2 \log g(x)}{dx^2} \quad \text{or} \quad \frac{d^2 \log f(x)}{dx^2} \leq \frac{d^2 \log g(x)}{dx^2}, \tag{11}$$

respectively. Under the conditions of Corollary 1, the right-hand sides of both (10) and (11) equal 0.

The natural choice of reference distribution in lifetime analysis is the exponential distribution with PDF

$$g_e(x; \lambda) = \lambda e^{-\lambda x}, \quad \lambda > 0, x \geq 0. \tag{12}$$

For the exponential reference distribution, the right-hand side of (10) equals $-\lambda$ and the right-hand side of (11) equals 0. That is, $\log(f(x)/g_e(x))$ is concave or convex if $f(x)$ is log-concave or log-convex, respectively. Furthermore, $K(f : g_e)$ is well defined if F has positive support, finite entropy, and finite mean. These conditions hold for all distributions presented in this paper.

Example 1. This example illustrates several applications of the above results, summarized in Table 1.

(a) Let $q(t) = \mu$ for all t . It is clear that $q(t)$ satisfies (6)–(9). Then (4) gives the exponential distribution with PDF $f(x) = (1/\mu)e^{-x/\mu}$. The exponential distribution PDF is decreasing and log-concave. By Corollary 1, this is the MDE model in the class of distributions with mean residual life $\mu_F(t) \leq \mu$.

(b) Let $q(t) = \beta + \alpha t$ for $t < \infty$, where $\alpha, \beta > 0$ and $q(\infty) = 0$. It is clear that $q(t)$ satisfies (6)–(9). Then (4) gives the generalized Pareto distribution with PDF shown in Table 1. This PDF is decreasing and log-convex, so Corollary 1 is not applicable. However, for the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is increasing and convex when $\lambda \geq (2\alpha + 1)/\beta$. By Theorem 1, in the class of distributions with mean residual life $\mu_F(t) \geq \beta + \alpha t$, the generalized Pareto distribution is therefore the MDDI model relative to the exponential distribution with $\lambda \geq (2\alpha + 1)/\beta$.

(c) Let $q(t) = \alpha/(\beta + t)$. It is clear that (6), (7), and (9) are satisfied when $\alpha, \beta > 0$, and (8) is satisfied when $\beta^2 \geq \alpha$. Then (4) gives a distribution with PDF shown in Table 1. This PDF is a combination of the PDFs of the normal and generalized normal distributions. It is log-concave when $\beta^2 \geq 3\alpha$. It is well known that log-concavity is equivalent to strong unimodality and implies that the distribution has increasing failure rate.

(i) This PDF is decreasing when $\beta^2 \geq 3\alpha$. By Corollary 1, it is the PDF of the MDE model in the class of distributions with mean residual life $\mu_F(t) \leq \alpha/(\beta + t)$.

(ii) For the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is decreasing and concave when $\lambda \leq \beta(\beta^2 - 3\alpha)/[\alpha(\beta^2 - \alpha)]$. By Theorem 1, in the class of distributions with mean residual life $\mu_F(t) \leq \alpha/(\beta + t)$, this is the PDF of the MDDI distribution relative to the exponential distribution with $\lambda \leq \beta(\beta^2 - 3\alpha)/[\alpha(\beta^2 - \alpha)]$.

TABLE 1: MDE and MDDI models relative to the exponential distribution, with mean residual life inequality constraints, for the densities shown.

Mean residual inequality constraint	Parameter restrictions	
	MDE	MDDI: $g_e(x; \lambda)$
<p>Exponential: $f^*(x) = \frac{1}{\mu}e^{-x/\mu}$ with $\mu > 0$</p> <p>$\mu_F(t) \leq \mu$</p>	None	None
<p>Generalized Pareto: $f^*(x) = \frac{\alpha + 1}{\beta} \left(1 + \frac{\alpha}{\beta}x\right)^{-1/\alpha - 2}$ with $\alpha, \beta > 0$</p> <p>$\mu_F(t) \geq \beta + \alpha t$</p>	Not applicable	$\lambda \geq \frac{2\alpha + 1}{\beta}$
<p>Generalized normal combination: $f^*(x) = \frac{e^{\beta^2/2\alpha}}{\beta} \left[\frac{(x + \beta)^2}{\alpha} - 1\right] e^{-(1/2\alpha)(x + \beta)^2}$ with $0 < \alpha < \beta^2$</p> <p>$\mu_F(t) \leq \frac{\alpha}{t + \beta}$</p>	$\beta^2 > 3\alpha$	$\beta^2 \geq 3\alpha, \lambda \leq \frac{\beta(\beta^2 - 3\alpha)}{\alpha(\beta^2 - \alpha)}$
<p>Extreme-value-type combination I: $f^*(x) = \left(\frac{1}{\alpha}e^{\beta x} - \beta\right)e^{\beta x}e^{(1/\alpha\beta)(1 - e^{\beta x})}$ with $\alpha\beta < 1$</p> <p>$\mu_F(t) \leq \alpha e^{-\beta t}$</p>	$\alpha\beta \leq \frac{3 - \sqrt{5}}{2}$	$\lambda \leq \frac{1 - \alpha\beta}{\alpha} - \frac{\beta}{1 - \alpha\beta}$
<p>Extreme-value-type combination II: $f^*(x) = \frac{1}{\gamma + 1} \left[\frac{1}{\alpha}(\gamma + e^{\beta x})^2 - \beta e^{\beta x}\right] e^{(1/\alpha\beta)(1 - \gamma\beta x - e^{\beta x})}$ with $\gamma > 0, \alpha\beta < 2\gamma$</p> <p>$\mu_F(t) \leq \frac{\alpha}{\gamma + e^{\beta t}}$</p>	$\alpha\beta < \gamma$	$\lambda \leq \min\left\{\frac{3(\gamma - \alpha\beta)}{\alpha}, \frac{\sqrt{3}\beta}{2}\right\}$
<p>Half-logistic: $f^*(x) = \frac{\beta(\alpha + 1)e^{\beta x}}{(\alpha + e^{\beta x})^2}$ with $\alpha, \beta > 0$</p> <p>$\mu_F(t) \leq \frac{1}{\alpha\beta}(\alpha + e^{\beta t}) \log(1 + \alpha e^{-\beta t})$</p>	$\alpha \leq 1$	$\alpha < 1, \lambda \leq \left(\frac{1 - \alpha}{1 + \alpha}\right)\beta$
<p>Mixture of exponentials: $f^*(x) = \alpha\beta_1 e^{-\beta_1 x} + (1 - \alpha)\beta_2 e^{-\beta_2 x}$ with $0 < \alpha < 1, \beta_1, \beta_2 > 0$</p> <p>$\mu_F(t) \geq \frac{\alpha\beta_1^{-1}e^{-\beta_1 t} + (1 - \alpha)\beta_2^{-1}e^{-\beta_2 t}}{\alpha e^{-\beta_1 t} + (1 - \alpha)e^{-\beta_2 t}}$</p>	Not applicable	$\lambda \geq \max\{\beta_1, \beta_2\}$

The extension to the case $q(t) = \alpha/(\beta_0 + \beta_1 t)$ is straightforward, and other models may be developed similarly. For example, let $q(t) = r(t)/s(t) > 0$, where $r(t) = \sum_{i=0}^{n_1} \alpha_i t^i$ and $s(t) = \sum_{j=0}^{n_2} \beta_j t^j$ are polynomials of degrees n_1 and n_2 , respectively, such that $n_1 < n_2, \alpha_0\beta_0 > 0, \alpha_{n_1}\beta_{n_2} > 0$, and $(d/dt)(r(t)/s(t)) > -1$. It can be shown that $q(t)$ satisfies (6)–(9). Then (4) gives a distribution with PDF

$$f(x) = \frac{\alpha_0}{\beta_0} \left[\left(\frac{s(x)}{r(x)} \right)^2 - \frac{d}{dt} \left(\frac{s(x)}{r(x)} \right) \right] \exp \left\{ - \int_0^x \frac{s(v)}{r(v)} dv \right\}.$$

For suitably chosen $r(\cdot)$ and $s(\cdot)$, this PDF satisfies the conditions of Corollary 1 and Theorem 1 for a given reference distribution G , and, thus, new MDE and MDDI models can be developed.

(d) Let $q(t) = \alpha e^{-\beta t}$. It is clear that (6), (7), and (9) are satisfied when $\alpha, \beta > 0$, and (8) is satisfied when $\alpha\beta < 1$. Then (4) gives a distribution with the extreme-value-type combination I PDF shown in Table 1. It is easy to show that this PDF is log-concave.

(i) It can be shown that this PDF is decreasing when $\alpha\beta \leq (3 - \sqrt{5})/2$. By Corollary 1, it is the PDF of the MDE model in the class of distributions with mean residual life $\mu_F(t) \leq \alpha e^{-\beta t}$.

(ii) For the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is decreasing when $\lambda \leq (1 - \alpha\beta)/\alpha - \beta/(1 - \alpha\beta)$. By Theorem 1, in the class of distributions with mean residual life $\mu_F(t) \leq \alpha e^{-\beta t}$, this is the PDF of the MDDI distribution relative to the exponential distribution with $\lambda \leq (1 - \alpha\beta)/\alpha - \beta/(1 - \alpha\beta)$.

(e) Let $q(t) = \alpha/(\gamma + e^{\beta t})$. It is clear that (6), (7), and (9) are satisfied when $\alpha, \beta > 0$, and it can be shown that (8) is satisfied when $\alpha\beta \leq 2\gamma$. Then (4) gives a distribution with the extreme-value-type combination II PDF shown in Table 1.

(i) It can be shown that this PDF is decreasing and log-concave when $\alpha\beta \leq \gamma$. By Corollary 1, it is the PDF of the MDE model in the class of distributions with mean residual life $\mu_F(t) \leq \alpha/(\gamma + e^{\beta t})$.

(ii) For the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is decreasing and concave when $\lambda \leq \min\{3(\gamma - \alpha\beta)/\alpha, \sqrt{3}\beta/2\}$. By Theorem 1, in the class of distributions with mean residual life $\mu_F(t) \leq \alpha/(\gamma + e^{\beta t})$, this is the PDF of the MDDI distribution relative to the exponential distribution with $\lambda \leq \min\{3(\gamma - \alpha\beta)/\alpha, \sqrt{3}\beta/2\}$.

(f) Let $q(t) = (1/\alpha\beta)(\alpha + e^{\beta t}) \log(1 + \alpha e^{-\beta t})$. It is clear that (6)–(9) are satisfied when $\alpha\beta > 0$. Then (4) gives the half-logistic distribution with PDF shown in Table 1. The PDF is log-concave.

(i) It can be shown that this PDF is decreasing when $\alpha \leq 1$. By Corollary 1, it is the PDF of the MDE model in the class of distributions with mean residual life $\mu_F(t) \leq (1/\alpha\beta)(\alpha + e^{\beta t}) \log(1 + \alpha e^{-\beta t})$.

(ii) For the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is decreasing when $\lambda \leq \beta(1 - \alpha)/(1 + \alpha)$. By Theorem 1, in the class of distributions with mean residual life $\mu_F(t) \leq (1/\alpha\beta)(\alpha + e^{\beta t}) \log(1 + \alpha e^{-\beta t})$, the half-logistic distribution is the MDDI distribution relative to the exponential distribution with $\lambda \leq \beta(1 - \alpha)/(1 + \alpha)$.

(g) Let

$$q(t) = \frac{\alpha\beta_1^{-1}e^{-\beta_1 t} + (1 - \alpha)\beta_2^{-1}e^{-\beta_2 t}}{\alpha e^{-\beta_1 t} + (1 - \alpha)e^{-\beta_2 t}}$$

It can be shown that (6)–(9) are satisfied. Then (4) gives the mixture of two exponential distributions with PDF shown in Table 1. This PDF is decreasing and log-convex, so Corollary 1 is not applicable. However, for the exponential reference distribution (12), $\log(f(x)/g_e(x))$ is increasing and convex when $\lambda \geq \max\{\beta_1, \beta_2\}$. By Theorem 1, in the class of distributions with mean residual life

$$\mu_F(t) \geq \frac{\alpha\beta_1^{-1}e^{-\beta_1 t} + (1 - \alpha)\beta_2^{-1}e^{-\beta_2 t}}{\alpha e^{-\beta_1 t} + (1 - \alpha)e^{-\beta_2 t}},$$

the mixture of two exponential distributions is the MDDI distribution relative to the exponential distribution with $\lambda \geq \max\{\beta_1, \beta_2\}$.

(h) The exponential distribution is the MDDI distribution relative to all distributions shown in Table 1. If $\log(f(x)/g_e(x))$ is decreasing and concave or increasing and convex, then $\log(g_e(x)/f(x))$ is increasing and convex or decreasing and concave, respectively. We therefore have the following results.

- (i) In the class of distributions $\Omega_F = \{F : \mu_F(t) \geq \mu\}$, the exponential distribution is the MDDI distribution relative to the generalized Pareto distribution and the mixture of two exponential distributions shown in Table 1.
- (ii) In the class of distributions $\Omega_F = \{F : \mu_F(t) \leq \mu\}$, the exponential distribution is the MDDI distribution relative to the generalized normal combination, the extreme-value-type combinations I and II, and the half-logistic distributions shown in Table 1.

The convexity or concavity condition on $\log(f^*(x)/g(x))$ in Theorem 1 may be difficult to meet when G is a lifetime distribution such as the gamma distribution. Our next theorem provides a solution to this problem by using the constraints on the mean residual life and its growth.

Theorem 2. *Let*

$$\Omega_F = \left\{ F : \mu_F(t) \leq q(t), \frac{\mu'_F(t)}{\mu_F(t)} \geq \frac{q'(t)}{q(t)} \right\}$$

be a compact set of distributions, where F is absolutely continuous with respect to a reference distribution G and a prime denotes differentiation. Let $F^ \in \Omega_F$ be such that $\mu_{F^*}(t) = q(t)$. If $f^*(x)/g(x)$ is decreasing then F^* is the MDDI distribution relative to G . The same result holds, with*

$$\Omega_F = \left\{ F : \mu_F(t) \geq q(t), \frac{\mu'_F(t)}{\mu_F(t)} \leq \frac{q'(t)}{q(t)} \right\},$$

if $f^(x)/g(x)$ is increasing.*

Proof. The proof follows the same steps as did the proof of Theorem 1. Here, however, the last inequality is obtained by noting that the conditions of Theorem 2 imply hazard ordering (see Theorem 3).

Corollary 2. *Let*

$$\Omega_F = \left\{ F : \mu_F(t) \leq q(t), \frac{\mu'_F(t)}{\mu_F(t)} \geq \frac{q'(t)}{q(t)} \right\}$$

be a compact set of absolutely continuous distributions. Let $F^ \in \Omega_F$ be such that $\mu_{F^*}(t) = q(t)$. If $f^*(x)$ is increasing then F^* is the MDE distribution. The same result holds if $f^*(x)$ is decreasing and*

$$\Omega_F = \left\{ F : \mu_F(t) \geq q(t), \frac{\mu'_F(t)}{\mu_F(t)} \leq \frac{q'(t)}{q(t)} \right\}.$$

Example 2. Consider the gamma distribution with PDF

$$g_g(x; \lambda, \nu) = \frac{\lambda^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\lambda x}, \quad \lambda, \nu > 0, x \geq 0. \tag{13}$$

Here, $K(f : g_g)$ is well defined if F has positive support, finite entropy, finite mean, and if $E_f[\log X] < \infty$. These conditions hold for all distributions presented in this paper.

TABLE 2: MDDI models relative to the gamma $g_g(x; \lambda)$ and exponential ($\nu = 1$) distributions, with mean residual life and growth inequality constraints, for the densities shown.

Mean residual inequality and growth constraints	Parameter restrictions on MDDI: $g_g(x; \lambda, \nu)$
<i>Exponential:</i> $f^*(x) = \frac{1}{\mu}e^{-x/\mu}$ with $\mu > 0$	
$\mu_F(t) \leq \mu, \quad \mu'_F(t) \geq 0$ or as above with the inequalities reversed	$\lambda \leq \beta, \quad \nu \geq 1$ or as above with the inequalities reversed
<i>Generalized Pareto:</i> $f^*(x) = \frac{\alpha + 1}{\beta} \left(1 + \frac{\alpha}{\beta}x\right)^{-1/\alpha - 2}$ with $\alpha, \beta > 0$	
$\mu_F(t) \geq \beta + \alpha t, \quad \mu'_F(t) \leq \alpha$	$\frac{(\lambda\beta - \alpha\nu - \alpha - 1)^2}{4\alpha\beta\lambda} + \nu \leq 1, \quad \nu < 1$
<i>Generalized normal combination:</i> $f^*(x) = \frac{e^{\beta^2/2\alpha}}{\beta} \left[\frac{(x + \beta)^2}{\alpha} - 1\right] e^{-(1/2\alpha)(x + \beta)^2}$ with $0 < \alpha < \beta^2$	
$\mu_F(t) \leq \frac{\alpha}{t + \beta}, \quad \frac{\mu'_F(t)}{\mu_F(t)} \geq -\frac{1}{\alpha}\mu_F(t)$ or as above with the inequalities reversed	$\lambda \leq \frac{\beta(\beta^2 - 3\alpha)}{\alpha(\beta^2 - \alpha)}, \quad \beta^2 > 3\alpha, \quad \nu \geq 1$ or as above with the inequalities reversed
<i>Extreme-value-type combination I:</i> $f^*(x) = \left(\frac{1}{\alpha}e^{\beta x} - \beta\right)e^{\beta x}e^{(1/\alpha\beta)(1 - e^{\beta x})}$ with $\alpha, \beta > 0, \alpha\beta < 1$	
$\mu_F(t) \leq \alpha e^{-\beta t}, \quad \frac{\mu'_F(t)}{\mu_F(t)} \geq -\beta$	$\lambda \leq \frac{(1 - \alpha\beta)^2 - \alpha\beta}{\alpha(1 - \alpha\beta)}, \quad \nu \geq 1$
<i>Extreme-value-type combination II:</i> $f^*(x) = \frac{1}{\gamma + 1} \left[\frac{1}{\alpha}(\gamma + e^{\beta x})^2 - \beta e^{\beta x}\right] e^{(1/\alpha\beta)(1 - \gamma\beta x - e^{\beta x})}$ with $\gamma > 0, \alpha\beta < 2\gamma$	
$\mu_F(t) \leq \frac{\alpha}{\gamma + e^{\beta t}}, \quad \frac{\mu'_F(t)}{\mu_F(t)} \geq -\frac{\beta}{\alpha}e^{\beta t}\mu_F(t)$	$\lambda \leq \min\left\{\frac{3(\gamma - \alpha\beta)}{\alpha}, \frac{\sqrt{3}\beta}{2}\right\}, \quad \nu \geq 1$
<i>Half-logistic:</i> $f^*(x) = \frac{\beta(\alpha + 1)e^{\beta x}}{(\alpha + e^{\beta x})^2}$ with $\alpha, \beta > 0$	
$\mu_F(t) \leq \frac{1}{\alpha\beta}(\alpha + e^{\beta t}) \log(1 + \alpha e^{-\beta t}),$ $\frac{\mu'_F(t)}{\mu_F(t)} \leq \frac{\beta}{1 + \alpha e^{-\beta t}} - \frac{1}{\mu_F(t)}$ or as above with the inequalities reversed	$\lambda \leq \left(\frac{1 - \alpha}{1 + \alpha}\right)\beta, \quad \alpha < 1, \quad \nu > -\lambda + \beta + 1$ $\lambda \geq \beta, \quad \nu \leq 1$
<i>Mixture of exponentials:</i> $f^*(x) = \alpha\beta_1 e^{-\beta_1 x} + (1 - \alpha)\beta_2 e^{-\beta_2 x}$ with $0 < \alpha < 1, \beta_1, \beta_2 > 0$	
$\mu_F(t) \leq q(t), \quad \frac{\mu'_F(t)}{\mu_F(t)} \leq \frac{q'(t)}{q(t)}$ or as above with the inequalities reversed	$\lambda \geq \beta_1, \quad \lambda \geq \beta_2, \quad \nu \geq 1$ or as above with the inequalities reversed

Table 2 shows the classes of distributions with mean residual life inequality and mean residual life growth rate constraints, where the distributions shown in Table 1 are the MDDI distributions relative to the gamma distribution with PDF $g_g(x; \lambda, \nu)$. Note that, in Table 2,

$$q(t) = \frac{\alpha\beta_1^{-1}e^{-\beta_1 t} + (1 - \alpha)\beta_2e^{-\beta_2 t}}{\alpha e^{-\beta_1 t} + (1 - \alpha)e^{-\beta_2 t}}.$$

The $g_g(x; \lambda, \nu)$ column in Table 2 shows the parameter restrictions for $f^*(x)/g_g(x)$ to be decreasing or increasing. MDDI models relative to the exponential distribution with PDF $g_e(x; \lambda)$ are recovered by letting $\nu = 1$ in the parameter restrictions of $g_g(x; \lambda, \nu)$. Simpler restrictions on the exponential parameters, $\lambda \geq (2\alpha + 1)/\beta$ and $\alpha\beta < \gamma$, can be obtained for the generalized Pareto and for the extreme value type-II distributions, respectively. We note that the lack of a convexity or a concavity restriction in Theorem 2 also allows us to reverse the mean residual life inequalities for the exponential, the generalized normal combination, the half-logistic, and the mixture of exponential distributions. Reversing the inequalities gives new classes of distribution in which these models are the MDDI distributions relative to the exponential distribution. The new classes are subsets of the complements of the classes of distribution shown in Table 1.

3. Hazard growth constraints

In this section, we give a result that generalizes the MDE procedure using differential inequality constraints for the hazard rate developed in Asadi *et al.* (2004). In general, for any distribution F with PDF f and corresponding hazard function $\lambda_F(t) = f(t)/(1 - F(t))$, the relative growth of the hazard function is given by

$$\frac{\lambda'_F(t)}{\lambda_F(t)} = \frac{f'(t)}{f(t)} + \lambda_F(t). \tag{14}$$

For any distribution with PDF $f^*(x)$, we can identify a set of distributions $\Omega_F = \{F\}$ such that

$$\frac{\lambda'_F(t)}{\lambda_F(t)} \leq \frac{\lambda'_{F^*}(t)}{\lambda_{F^*}(t)} \quad \text{or} \quad \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \frac{\lambda'_{F^*}(t)}{\lambda_{F^*}(t)}, \tag{15}$$

where the right-hand side of each inequality is given by (14). Solution of the differential inequalities (15) with appropriate initial conditions gives a $\lambda_{F^*}(t)$ that dominates or, respectively, is dominated by $\lambda_F(t)$ for all distributions in Ω_F . Thus, by relating the dynamic entropy and the hazard rate of distributions in Ω_F , $H(f^*; t)$ dominates $H(f, t)$ and one can use it to identify Ω_F for which a given $f^*(x)$ is the MDE model.

Note that the hazard function is the rate of growth of the mean residual life, i.e.

$$\lambda_F(t) = \frac{1 + \mu'_F(t)}{\mu_F(t)}, \quad t \geq 0.$$

Thus, a set of differential inequality constraints may also be expressed in terms of a set of mean residual life constraints. However, the constraints can be in more complicated forms, including second-order differential equations.

Asadi *et al.* (2004) gave a result that relates hazard rate ordering with the dynamic entropy ordering for distributions with monotone densities. Our next theorem provides a generalization of Theorem 1 of Asadi *et al.* (2004) that facilitates the derivation of various MDDI models.

Theorem 3. Let $\Omega_F = \{F : \lambda_F(t) \leq r(t)\}$ be a compact set of distributions, where F is absolutely continuous with respect to a reference distribution G . Let $F^* \in \Omega_F$ be such that $\lambda_{F^*}(t) = r(t)$. If $f^*(x)/g(x)$ is decreasing then F^* is the MDDI distribution relative to G . The same result holds, with $\Omega_F = \{F : \lambda_F(t) \geq r(t)\}$, if $f^*(x)/g(x)$ is increasing.

Proof. The proof for $\lambda_F(t) \leq r(t)$ follows the same steps as did the proof of Theorem 1. Here, the last inequality is obtained by noting that hazard rate ordering implies that the residual lifetimes are stochastically ordered and that $f^*(x)/g(x)$ is decreasing in x . The proof in the increasing case is similar and is given in Ebrahimi and Kirmani (1996b).

When the reference distribution is uniform, Theorem 3 reduces to the result of Asadi *et al.* (2004) for dynamic entropy ordering of distributions with monotone densities. This result is applicable for many suitable choices of reference distribution, including well-known lifetime models such as the exponential, gamma, Weibull, and Pareto distributions.

Since, for $\nu = 1$, the gamma distribution PDF (13) reduces to the exponential distribution PDF (12), it suffices to identify the MDDI models relative to the gamma distribution in the classes of distributions with hazard rate growth constraints. MDDI models relative to the Weibull and Pareto distributions can be obtained via transformation of the results for exponential distributions, as will be shown in the sequel.

Our next example derives numerous well-known lifetime distributions as MDDI models relative to the gamma distribution. Most of them are also MDDI models relative to the exponential distribution.

Example 3. Tables 3 and 4 show hazard rate constraints that define classes of distribution where well-known distributions are MDDI models relative to the gamma distribution with PDF $g_g(x; \lambda, \nu)$. The differential inequality constraints shown in these tables are from Asadi *et al.* (2004), with two exceptions. For Pareto type III, the right-hand side of the differential inequality for the evolution of the hazard function is adopted from Hamedani (2005) and the differential inequality for the log-normal distribution is a new result. Since the log-normal distribution is not monotone, it was not included in Asadi *et al.* (2004); however, the proof for the log-normal distribution is similar to the proof of Proposition 3.3 of Asadi *et al.* (2004). The types and complexities of the differential inequality constraints were discussed in Asadi *et al.* (2004). Since the hazard functions of the log-normal and half-normal distributions are not available in closed form, the right-hand sides of the differential inequality constraints for these distributions include their hazard functions $\lambda_{F^*}(t)$. For all other distributions listed in Tables 3 and 4, the differential inequality constraints have closed forms. Note that, in Table 4,

$$A_{\alpha,\beta,\lambda} = \max \left\{ 1, 2 - \frac{\alpha(\alpha - \lambda)}{\beta} \right\},$$

$$B_{\alpha,\beta,\lambda,\nu} = \frac{(\alpha + \lambda - \beta\nu)^2}{4\alpha\nu(1 - \lambda)},$$

$$C_{\alpha,\beta,\lambda,\nu} = \frac{(\lambda\beta - \alpha\nu - \alpha - 1)^2}{4\alpha\beta\lambda} + \nu.$$

Table 3 shows the restrictions on the parameters required in order to guarantee the monotonicity of the likelihood ratio $f^*(x)/g_g(x)$. In this table, setting $\nu = 1$ in the parameter restrictions gives the parameter restrictions for guaranteeing monotonicity of the likelihood ratio with exponential reference distribution $f^*(x)/g_e(x)$.

TABLE 3: MDDI models relative to the gamma $g_g(x; \lambda)$ and exponential ($\nu = 1$) distributions, with hazard rate growth inequality constraints, for the densities shown.

Hazard rate growth constraints	Parameter restrictions on MDDI: $g_g(x; \lambda, \nu)$
<i>Exponential:</i> $f^*(x) = \beta e^{-\beta x}$ with $\beta > 0$	
$\lambda_F(0) = \beta, \quad \lambda'_F(t) \leq 0$ or as above with the inequality reversed	$\lambda \leq \beta, \quad \nu \geq 1$ or as above with the inequalities reversed
<i>Mixture of exponentials:</i> $f^*(x) = \alpha\beta_1 e^{-\beta_1 x} + (1 - \alpha)\beta_2 e^{-\beta_2 x}$ with $0 < \alpha < 1, \beta_1, \beta_2 > 0$	
$\lambda_F(0) = \alpha\beta_1 + (1 - \alpha)\beta_2,$ $\beta_2 \leq \lambda_F(t) \leq \beta_1, \quad \frac{\lambda'_F(t)}{\lambda_F(t) - \beta_2} \geq \lambda_F(t) - \beta_1$ or as above with the inequalities reversed	$\lambda \leq \beta_2, \quad \nu \geq 1$ $\lambda \geq \beta_1, \quad \nu \geq 1$
<i>Half-logistic:</i> $f^*(x) = \frac{\beta(\alpha + 1)e^{\beta x}}{(\alpha + e^{\beta x})^2}$ with $\alpha, \beta > 0$	
$\lambda_F(0) = \lambda_0, \quad \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \beta - \lambda_F(t) \geq 0$ or reverse the first inequality and drop the last constraint	$\lambda \leq \left(\frac{1 - \alpha}{1 + \alpha}\right)\beta, \alpha < 1, \quad \nu > \lambda + \beta + 1$ $\lambda \geq \beta, \quad \nu \leq 1$
<i>Minimum of exponential and Pareto (type III):</i> $f^*(x) = (\alpha + \beta(x + 1))e^{-\beta x}(x + 1)^{-(\alpha+1)}$ with $\alpha, \beta > 0$	
$\lambda_F(0) = \alpha + \beta, \quad \lambda_F(t) \geq \beta,$ $\frac{\lambda'_F(t)}{\lambda_F(t) - \beta} \leq -\frac{1}{\alpha}(\lambda_F(t) - \beta)$ or as above with the last inequality reversed	$\lambda \leq \beta, \quad \nu \geq 1$ $\lambda - \nu \geq \alpha + 2\beta, \quad \nu < \min\{1, \lambda\}$
<i>Extreme value (Gumbel):</i> $f^*(x) = \alpha e^{(\beta/\alpha)x} \exp\left\{\frac{\alpha^2}{\beta}(1 - e^{(\beta/\alpha)x})\right\}$ with $0 < \beta < \alpha^2$	
$\lambda_F(0) = \alpha, \quad \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \frac{\beta}{\alpha}$	$\lambda \leq \alpha - \frac{\beta}{\alpha}, \quad \nu \geq 1$
<i>Minimum of exponential and extreme value:</i> $f^*(x) = (\beta_1 e^x + \beta_2) \exp\{-\beta_2 x - \beta_1(e^x - 1)\}$ with $\beta_1 > 0, \beta_2 \geq 1$	
$\lambda_F(0) = \beta_1 + \beta_2, \quad \frac{\lambda'_F(t)}{\lambda_F(t)} \geq 1 - \frac{\beta_2}{\lambda_F(t)}$	$\lambda \leq \beta_2 - 1, \quad \nu \geq 1$
<i>Log-normal:</i> $f^*(x) = \frac{1}{\sqrt{2\pi x}} e^{-(\log x)^2/2}$	
$\lambda_F(0) = 0, \quad \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \lambda_{F^*}(t) - \frac{\log t + 1}{t}$	$\lambda \geq \nu e^\nu, \quad \nu \geq 1$

TABLE 4: MDDI models relative to the exponential $g_e(x; \lambda)$ and gamma $g_g(x; \lambda)$ distributions, with hazard rate growth inequality constraints, for the densities shown.

Hazard rate growth constraints	Parameter restrictions	
	MDDI: $g_e(x; \lambda)$	MDDI: $g_g(x; \lambda, \nu)$
<i>Weibull:</i> $f^*(x) = \alpha\beta x^{\alpha-1}e^{-\beta x^\alpha}$ with $\alpha \geq 1, \beta > 0$		
$\lambda_F(1) = \alpha\beta, \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \frac{\alpha-1}{t}$	Not applicable	$\lambda \leq \alpha\beta, \nu \geq \alpha + \lambda$
or reverse the first inequality and replace the constraint in the PDF with $\alpha, \beta \leq 1$	Not applicable	$\lambda \geq \alpha\beta, \nu \leq (1-\beta)\alpha$
<i>Linear failure rate:</i> $f^*(x) = (\alpha + \beta x) \exp\{-\alpha x + \frac{1}{2}\beta x^2\}$ with $0 < \beta \leq \alpha^2$		
$\lambda_F(0) = \alpha, \lambda'_F(t) \geq \beta$	$\lambda \leq \alpha - \frac{\beta}{\alpha}$	$\lambda \leq \alpha, A_{\lambda, \beta, \alpha} \leq \nu \leq 2$
<i>Pareto type II:</i> $f^*(x) = \alpha\beta^\alpha(x + \beta)^{-(\alpha+1)}$ with $\alpha > 1, \beta > 0$		
$\lambda_F(0) = \alpha, \frac{\lambda'_F(t)}{\lambda_F(t)} \leq -\frac{\lambda_F(t)}{\alpha}$	$\lambda \geq \frac{\alpha+1}{\beta}$	$B_{\alpha, \beta, \nu, \lambda} \leq 1, \nu \leq 1$
<i>Generalized Pareto:</i> $f^*(x) = \frac{\alpha+1}{\beta} \left(1 + \frac{\alpha}{\beta}x\right)^{-1/\alpha-2}$ with $\alpha, \beta > 0$		
$\lambda_F(0) = \frac{\alpha+1}{\beta}, \frac{\lambda'_F(t)}{\lambda_F(t)} \leq -\frac{\alpha}{\alpha+1}\lambda_F(t)$	$\lambda \geq \frac{2\alpha+1}{\beta}$	$C_{\alpha, \beta, \lambda, \nu} \leq 1$
<i>Half-normal:</i> $f^*(x) = \sqrt{\frac{2}{\pi}}e^{-x^2/2}$		
$\lambda_F(0) = \sqrt{\frac{2}{\pi}}, \frac{\lambda'_F(t)}{\lambda_F(t)} \geq \lambda_{F^*}(t) - \frac{1}{2}t$	Not applicable	$\nu \geq 1 + \frac{1}{4}\lambda^2$

Table 4 shows more complicated cases. For the Weibull and half-normal models, the monotonicity condition can not be satisfied by $g_e(x; \lambda)$. For other models, the parameter restrictions for a gamma reference distribution are complicated and simpler restrictions, for an exponential reference distribution, can be obtained as shown in the table.

Table 5 shows several other distributions that, by Theorem 3, can be obtained as MDDI distributions relative to the exponential and gamma distributions. The constraints for these distributions have very complicated forms and, thus, are not listed.

Finally, noting that if $\log f(x)/g_g(x)$ is decreasing or increasing then $\log g_g(x)/f(x)$ is respectively increasing or decreasing, we can apply Theorem 3 to derive the gamma model as the MDDI distribution relative to those listed in Tables 3–5 for which $K(g_e : f)$ is well defined. Consider the class of distributions with the following hazard rate growth constraints, where $\lambda_{F^*}(t)$ is the hazard rate of the gamma distribution, i.e. $\lambda_F(1) = \lambda_0$,

$$\frac{\lambda'_F(t)}{\lambda_F(t)} \begin{cases} \geq \lambda_{F^*}(t) + \frac{(\alpha-1)}{t} - 1 & \text{for } \alpha \geq 1, \\ \leq \lambda_{F^*}(t) + \frac{(\alpha-1)}{t} - 1 & \text{for } \alpha \leq 1. \end{cases}$$

TABLE 5: MDDI models relative to the exponential $g_e(x; \lambda)$ and gamma $g_g(x; \lambda, \nu)$ distributions, with hazard rate growth inequality constraints on the PDF f .

Density	Direction of inequality	
	$g_e(x; \lambda)$	$g_g(x; \lambda, \nu)$
Beta prime: $f(x) \propto x^{\alpha-1}(1+x)^{-(\alpha+\beta)}$	\geq	\geq
F: $f(x) \propto x^{\alpha/2-1}(\beta + \alpha x)^{-(\alpha+\beta)/2}$	\geq	\geq
Generalized inverse Gaussian: $f(x) \propto x^{\xi-1}e^{-(\alpha x + \beta/x)/2}$	\geq	\geq
Inverse Gaussian: $\xi = -\frac{1}{2}, \alpha = \beta/\gamma^2$	\geq	\geq
Levy-Smirnov: $\xi = -\frac{1}{2}, \alpha = 0, \beta = 1$	\geq	\geq
Generalized normal: $f(x) \propto x^{\alpha-1}e^{-(x/\beta)^2}$	\leq	\leq
Inverse gamma: $f(x) \propto x^{-\alpha-1}e^{-\beta/x}$	\geq	\geq
Inverse generalized normal: $f(x) \propto x^{-\alpha-1}e^{-(\beta/x)^2}$	\geq	\geq
Toranzos: $f(x) \propto x^{\xi}e^{-(1/2)(\alpha+\beta x)^2}, \xi > 0$	Not applicable	\leq

As in the log-normal and half-normal cases, the right-hand sides of the differential inequality constraints include the gamma hazard rate. Under the parameter restrictions shown in Tables 3 and 4, the gamma distribution is the MDDI model relative to the respective distributions shown in these tables.

It can be shown that many of the distributions in Tables 3–5 have the MDDI property relative to the Weibull and Pareto distributions. In addition to the extreme value distribution, the minimum of the exponential and extreme value distributions, and the log-normal distribution, all distributions listed in Tables 4 and 5 (except the beta prime distribution, in the latter case) are MDDI models relative to the Weibull distribution, subject to some parameter restrictions. It can also be shown that, except for the log-normal distribution, all distributions listed in Table 3, the linear failure rate and Pareto type-II distributions shown in Table 4, and all distributions listed in Table 5 are MDDI models relative to the generalized Pareto distribution, subject to some parameter restrictions.

Theorem 4. *Let $Y = \phi(X)$ be a one-to-one transformation. Then*

$$K(f_\phi^* : g_\phi; \phi(t)) = K(f^* : g; t) \leq K(f : g; t) = K(f_\phi : g_\phi; \phi(t)), \tag{16}$$

where

$$f_\phi^*(y) = f^*(\phi^{-1}(y)) \left| \frac{d\phi^{-1}(y)}{dy} \right|,$$

$$g_\phi(y) = g(\phi^{-1}(y)) \left| \frac{d\phi^{-1}(y)}{dy} \right|.$$

Proof. The proof follows from the fact that $K(f : g)$ is invariant under nonsingular transformations (Kullback (1959)).

This result allows us to produce new MDDI models F_{ϕ}^* , with reference G_{ϕ} , in the new classes of distribution obtained by the following transformations of the hazard constraints:

$$\begin{aligned} \lambda_{F_{\phi}}(t) &= \lambda_F(\phi^{-1}(t)) \left| \frac{d\phi^{-1}(t)}{dt} \right|, \\ \frac{\lambda'_{F_{\phi}}(t)}{\lambda_{F_{\phi}}(t)} &= \frac{\lambda'_F(\phi^{-1}(t))}{\lambda_F(\phi^{-1}(t))} \left| \frac{d\phi^{-1}(t)}{dt} \right| + \left| \frac{d^2\phi^{-1}(t)/dt^2}{d\phi^{-1}(t)/dt} \right|. \end{aligned} \tag{17}$$

The parameter restrictions for the transformed reference and MDDI models are found from the fact that $f_{\phi}(y)/g_{\phi}(y)$ is decreasing or increasing in y if and only if

$$\frac{f'_{\phi}(y)}{f_{\phi}(y)} \leq \frac{g'(\phi^{-1}(y))}{g(\phi^{-1}(y))} + \left| \frac{d^2\phi^{-1}(y)/dy^2}{d\phi^{-1}(y)/dy} \right| \quad \text{or} \quad \frac{f'_{\phi}(y)}{f_{\phi}(y)} \geq \frac{g'(\phi^{-1}(y))}{g(\phi^{-1}(y))} + \left| \frac{d^2\phi^{-1}(y)/dy^2}{d\phi^{-1}(y)/dy} \right|, \tag{18}$$

respectively.

The following example discusses some important applications of the properties of the transformation outlined above.

Example 4. Weibull, Pareto, and linear failure rate distributions can be obtained by one-to-one transformations of the exponential distribution. Figure 1 illustrates transformations from an exponential random variable X to random variables that have these distributions. Their PDFs are

$$\begin{aligned} g_W(y; \lambda, \tau) &= \lambda \tau y^{\tau-1} e^{-\lambda y^{\tau}}, & y \geq 0, \lambda, \tau > 0, \\ g_P(y; \lambda, \tau) &= \lambda \tau^{\lambda} (\tau + y)^{-(\lambda+1)}, & y \geq 0, \lambda, \tau > 0, \\ g_{lfr}(y; \eta, \gamma) &= (\gamma + \eta y) e^{-(\gamma y + (\eta/2)y^2)}, & y \geq 0, \gamma, \eta > 0, \end{aligned}$$

where $\eta = 2\lambda/b$ and $\gamma = \sqrt{a\eta}$, $a, b > 0$.

The MDDI models relative to these distributions can be obtained from the MDDI models, relative to the gamma distribution, listed in Tables 3 and 4 via transformations of the exponential random variable shown in Figure 1. An important implication of the transformation invariance in this case is that the constant hazard property of the exponential distribution may be too restrictive, and we may seek to develop models that are as close to a distribution with a more flexible hazard function as a given model is to the exponential distribution. For example, we may develop a model that is as close to the Weibull distribution (with a power hazard function) as the extreme value distribution is to the exponential distribution. The solution is $F_{\phi_W}^*$ with PDF

$$f_{\phi_W}^*(y) = \alpha \tau y^{\tau-1} e^{(\beta/\alpha)y^{\tau}} \exp \left\{ \frac{\alpha^2}{\beta} (1 - e^{(\beta/\alpha)y^{\tau}}) \right\},$$

obtained by applying ϕ_W to the extreme value PDF shown in Table 3. By the invariance property (16), we have $K(f_{\phi_W}^* : g_W; t) = K(f^* : g_e; t)$ for all $t \geq 0$. Furthermore, the new distribution $F_{\phi_W}^*$ is the MDDI model relative to the Weibull distribution in a new class of distributions that can be obtained using (17), an appropriate initial condition, and the parameter restrictions imposed by (18).

Next, we develop a result that generalizes the MDDI approach to the information divergence of order α (Rényi (1961)) between two distributions, defined by

$$K_{\alpha}(f : g) = \frac{1}{\alpha - 1} \log \int_0^{\infty} f^{\alpha}(x) g^{1-\alpha}(x) dx,$$

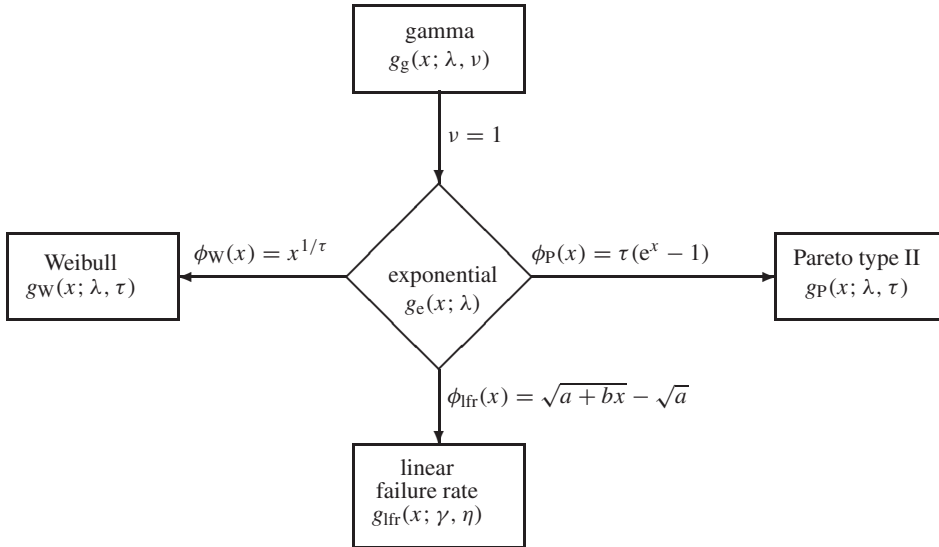


FIGURE 1: Derivation of reference distributions via transformations of the exponential distribution.

where $\alpha \neq 1$. It is well known that $\lim_{\alpha \rightarrow 1} K_\alpha(f : g) =: K_1(f : g) = K(f : g)$. Also, $K_{1/2}(f : g)$ is symmetric in f and g .

The Rényi divergence between two residual distributions with densities $f(x; t)$ and $g(x; t)$ is given by

$$K_\alpha(f : g; t) = \frac{1}{\alpha - 1} \log \int_t^\infty f^\alpha(x; t) g^{1-\alpha}(x; t) dx.$$

It is clear that $K_\alpha(f : g; 0) = K_\alpha(f : g)$ and $K_1(f : g; t) = K(f : g; t)$.

Theorem 5. Let $\Omega_F = \{F : \lambda_F(t) \leq r(t)\}$ be a compact set of distributions, where F is absolutely continuous with respect to a reference distribution G . Let $F^* \in \Omega_F$ be such that $\lambda_{F^*}(t) = r(t)$. If $\lambda'_F(x)/\lambda_F(x) \geq g'(x)/g(x)$, where g is a PDF, then, for $\alpha > 1$, F^* is the MDDI α model in Ω_F relative to a distribution G with PDF g . The same result holds, with $\Omega_F = \{F : \lambda_F(t) \geq r(t)\}$, if $\lambda'_F(x)/\lambda_F(x) \leq g'(x)/g(x)$.

Proof. We give the proof for the case of $\lambda_F(t) \geq r(t)$. A similar argument can be used in the case of $\lambda_F(t) \leq r(t)$. Since $\lambda_F(t) \geq r(t)$ for all $t \geq 0$ and $\lambda_{F^*}(t) = r(t)$, by Theorem 1.B.2 of Shaked and Shantikumar (1994), we have

$$\left[\frac{\bar{F}(x)}{\bar{F}(t)} \right]^\alpha \leq \left[\frac{\bar{F}^*(x)}{\bar{F}^*(t)} \right]^\alpha.$$

The condition $\lambda'_F(x)/\lambda_F(x) \leq g'(x)/g(x)$ implies that

$$\begin{aligned} \mathbb{E} \left[\frac{\lambda_F^{\alpha-1}(X^*)}{g^{\alpha-1}(X^*)} \mid X^* \geq t \right] &\geq \mathbb{E} \left[\frac{\lambda_F^{\alpha-1}(Y^*)}{g^{\alpha-1}(Y^*)} \mid Y^* \geq t \right] \\ &\geq \mathbb{E} \left[\frac{\lambda_{F^*}^{\alpha-1}(Y^*)}{g^{\alpha-1}(Y^*)} \mid Y^* \geq t \right], \end{aligned}$$

where X^* and Y^* have survival functions \bar{F}^α and $\bar{F}^{*\alpha}$, respectively. From the last inequality, we obtain

$$\frac{\int_t^\infty f^{*\alpha}(x)g^{1-\alpha}(x) dx}{\bar{F}^{*\alpha}(t)\bar{G}^{\alpha-1}(t)} \leq \frac{\int_t^\infty f^\alpha(x)g^{1-\alpha}(x) dx}{\bar{F}^\alpha(t)\bar{G}^{\alpha-1}(t)}.$$

That is,

$$K_\alpha(f^* : g; t) \leq K_\alpha(f : g; t).$$

Remark 1. For $0 < \alpha < 1$, we can use the same argument to show similar results for $\Omega_F = \{F : \lambda_F(t) \leq r(t)\}$, where $\lambda'_F(x)/\lambda_F(x) \leq g'(x)/g(x)$, and $\Omega_F = \{F : \lambda_F(t) \geq r(t)\}$, where $\lambda'_F(x)/\lambda_F(x) \geq g'(x)/g(x)$.

Example 5. Let G be the exponential distribution with PDF (12). Then the condition of Theorem 5 is $\lambda'_F(x)/\lambda_F(x) \geq -\lambda$. Some examples of MDDI α models listed in Tables 3 and 4 are as follows. Note that θ is used here in place of the distributional parameter α .

(a) The Weibull distribution is the MDDI α model in

$$\Omega_F = \begin{cases} \{F : \lambda_F(t) \leq \theta t^{\theta-1}\}, & \theta > 1, \\ \{F : \lambda_F(t) \geq \theta t^{\theta-1}\}, & 0 < \theta < 1. \end{cases}$$

(b) The linear failure rate distribution is the MDDI α model in

$$\Omega_F = \begin{cases} \{F : \lambda_F(t) \leq \theta + \beta t\}, & \theta > 1, \\ \{F : \lambda_F(t) \geq \theta + \beta t\}, & 0 < \theta < 1. \end{cases}$$

(c) The extreme value distribution is the MDDI α model in

$$\Omega_F = \begin{cases} \{F : \lambda_F(t) \leq \theta e^{(\beta/\theta)t}\}, & \theta > 1, \\ \{F : \lambda_F(t) \geq \theta e^{(\beta/\theta)t}\}, & 0 < \theta < 1. \end{cases}$$

(d) The half-logistic distribution is the MDDI α model in

$$\Omega_F = \begin{cases} \left\{ F : \lambda_F(t) \leq \frac{\beta e^{\beta t}}{\theta + e^{\beta t}} \right\}, & \theta > 1, \\ \left\{ F : \lambda_F(t) \geq \frac{\beta e^{\beta t}}{\theta + e^{\beta t}} \right\}, & 0 < \theta < 1. \end{cases}$$

References

ASADI, M., EBRAHIMI, N., HAMEDANI, G. G. AND SOOFI, E. S. (2004). Maximum dynamic entropy models. *J. Appl. Prob.* **41**, 379–390.

BELZUNCE, F., NAVARRO, J., RUIZ, J. M. AND DEL AGUILA, Y. (2004). Some results on residual entropy functions. *Metrika* **59**, 147–161.

DI CRESCENZO, A. AND LONGOBARDI, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *J. Appl. Prob.* **39**, 434–440.

EBRAHIMI, N. (1996). How to measure uncertainty in the residual lifetime distributions. *Sankhyā A* **58**, 48–57.

EBRAHIMI, N. (1998). Testing for exponentiality of the residual lifetime based on dynamic Kullback–Leibler information. *IEEE Trans. Reliab.* **47**, 197–201.

EBRAHIMI, N. (2001). Testing for uniformity of the residual lifetime based on dynamic Kullback–Leibler information. *Ann. Inst. Statist. Math.* **53**, 325–337.

EBRAHIMI, N. AND KIRMANI, S. N. U. A. (1996a). A characterization of the proportional hazards model through a measure of discrimination between two residual life distributions. *Biometrika* **83**, 233–235.

- EBRAHIMI, N. AND KIRMANI, S. N. U. A. (1996b). Some results on ordering of survival functions through uncertainty. *Statist. Prob. Lett.* **29**, 167–176.
- HAMEDANI, G. G. (2005). Characterizations of univariate continuous distributions based on hazard functions. To appear in *J. Appl. Statist. Sci.*
- JAYNES, E. T. (1957). Information theory and statistical mechanics. *Physics Rev.* **106**, 620–630.
- JAYNES, E. T. (1982). On the rationale of maximum entropy methods. *Proc. IEEE* **70**, 939–952.
- KULLBACK, S. (1959). *Information Theory and Statistics*. John Wiley, New York.
- RÉNYI, A. (1961). On measures of entropy and information. In *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, Vol. 1, University of California Press, Berkeley, CA, pp. 547–561.
- SHAKED, M. AND SHANTIKUMAR, J. G. (1994). *Stochastic Orders and Their Applications*. Academic Press, Boston, MA.
- SHANNON, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J.* **27**, 379–423, 623–656.
- SHORE, J. E. AND JOHNSON, R. W. (1980). Axiomatic derivation of the principle of maximum entropy and principle of minimum cross-entropy. *IEEE Trans. Inf. Theory* **26**, 26–37.