J. Appl. Prob. **54,** 1027–1050 (2017)
*doi:*10.1017/jpr.2017.51

CONNECTIONS OF GINI, FISHER, AND SHANNON BY BAYES RISK UNDER PROPORTIONAL HAZARDS

MAJID ASADI,* *University of Isfahan and Institute of Research in Fundamental Sciences (IPM)*
NADER EBRAHIMI,** *Northern Illinois University*
EHSAN S. SOOFI,*** *University of Wisconsin-Milwaukee*

Abstract

The proportional hazards (PH) model and its associated distributions provide suitable media for exploring connections between the Gini coefficient, Fisher information, and Shannon entropy. The connecting threads are Bayes risks of the mean excess of a random variable with the PH distribution and Bayes risks of the Fisher information of the equilibrium distribution of the PH model. Under various priors, these Bayes risks are generalized entropy functionals of the survival functions of the baseline and PH models and the expected asymptotic age of the renewal process with the PH renewal time distribution. Bounds for a Bayes risk of the mean excess and the Gini's coefficient are given. The Shannon entropy integral of the equilibrium distribution of the PH model is represented in derivative forms. Several examples illustrate implementation of the results and provide insights for potential applications.

Keywords: Entropy functional; generalized entropy; equilibrium distribution; escort distribution; excess cost; mean residual life; cumulative hazard

2010 Mathematics Subject Classification: Primary 62B10; 62C05; 62N05; 94A15
Secondary 94A17

1. Introduction

Let X_α be a nonnegative random variable with the survival function $\bar{F}_\alpha(x) = \mathbb{P}(X_\alpha > x)$, $x \geq 0$. The proportional hazards (PH) model is defined by

$$\bar{F}_\alpha(x) = \bar{F}^\alpha(x), \qquad x \geq 0,\ \alpha > 0, \tag{1.1}$$

where $\bar{F}$ is the survival function of the baseline hazard model and α is the proportional hazard parameter. When X_α has a probability density function (PDF) f_α, its hazard function is defined by $\lambda_\alpha(x) = f_\alpha(x)/\bar{F}_\alpha(x)$, $\bar{F}_\alpha(x) > 0$, and the PH model can be represented as $\lambda_\alpha(x) = \alpha\lambda(x)$, $x \geq 0$, $\alpha > 0$, where $\lambda(x)$ is the baseline hazard. The cumulative hazard function of X with survival function $\bar{F}$ is defined by

$$\Lambda(x) = \int_0^x \lambda(t)\,\mathrm{d}t = -\log \bar{F}(x). \tag{1.2}$$

Received 1 November 2016; revision received 11 May 2017.
* Postal address: Department of Statistics, University of Isfahan, Isfahan, 81744, Iran.
Email address: m.asadi@stat.ui.ac.ir
** Postal address: Division of Statistics, Northern Illinois University, DeKalb, IL 60155, USA.
Email address: nebrahim@niu.edu
*** Postal address: Lubar School of Business, University of Wisconsin-Milwaukee, P.O. Box 742, Milwaukee, WI 53201, USA. Email address: esoofi@uwm.edu

The PH model and its associated distributions provide a rich medium for exploring relationships between the Gini measure of inequality, the Fisher measure of the steepness of a likelihood function over the parameter space, and the Shannon entropy.

The equilibrium distribution (ED) of a nonnegative random variable X with distribution F and a finite mean $\mathbb{E}(X) = \mu < \infty$ is defined by the following PDF:

$$p(x) = \frac{\bar{F}(x)}{\int_0^\infty \bar{F}(x)\,\mathrm{d}x} = \frac{\bar{F}(x)}{\mu}, \qquad x \geq 0. \tag{1.3}$$

The random variable A_X associated with the ED is referred to as the asymptotic age by Shaked and Shanthikumar [21] and $\mathbb{E}(A_X)$ is the expected asymptotic age of the renewal process at the age t (see [20]). The ED plays an important role in renewal processes.

The PDF of the ED of X_α with the PH model (1.1) is related to the ED of the baseline hazard model F as follows:

$$p_\alpha(x) = \frac{\bar{F}^\alpha(x)}{\mu_\alpha} = \frac{[p(x)]^\alpha}{C_\alpha}, \qquad x \geq 0,\ \alpha > 0, \tag{1.4}$$

where $C_\alpha = \mu_\alpha/\mu^\alpha$, $\alpha > 0$, and $\mu_\alpha = \mathbb{E}(X_\alpha) < \infty$.

The connections between the Gini coefficient, Fisher information, and Shannon entropy are via Bayes risks which appear in the form of a generalized entropy functional defined in the next section. In addition to establishing the connections between these measures, we derive the Shannon entropy of p_α as the derivative of its hazard function with respect to the PH parameter α. This result provides a derivative representation of the Shannon entropy which up till now has been seen as integrals of the PDF and hazard function.

The paper is organized as follows. In Section 2 we define the generalized entropy measures used in this paper. In Section 3 we present three Bayes risk measures of the mean excess of the random variable distributed as the PH model, discuss some applications, and give representations and bounds for one of the Bayes risks. The connections between the Bayes risk and Gini measures are established in Section 4. In Section 5 we present a link between Fisher information of the PH parameter and the Shannon entropy functional via the ED and the generalized entropy functional of the survival function. In Section 6 we give representations of the Shannon entropy of (1.4) in terms of a derivative. This section also contains a result for computing the entropy of the PH model and comparing it with the entropy of the baseline hazard model. In Section 7 we summarize the findings of the paper. Proofs can be found in Appendix A.

2. Generalized entropy functional

The generalized entropy (GE) of order q of a continuous nonnegative random variable with PDF f is defined by

$$H_q(X) = H_q(f) = \frac{1}{1-q}\int_0^\infty f(x)[f^{q-1}(x) - 1]\,\mathrm{d}x, \qquad q \in \Omega, \tag{2.1}$$

$$= H(f) = -\int_0^\infty f(x)\log f(x)\,\mathrm{d}x, \qquad q = 1, \tag{2.2}$$

where $\Omega = (0, 1) \cup (1, \infty)$ and $H(f) = \lim_{q\to 1} H_q(f)$ is the Shannon entropy, provided that the integrals are finite.

The GE is known as the Tsallis–Havrda–Charvat entropy or the Tsallis entropy in the statistics literature. It was introduced by Havrda and Charvát [14] and popularized by Tsallis [23]. Maasoumi [16] proposed a discrete version of (2.1) for measuring multivariate income inequality. Abe [1] provided an axiomatic derivation of the discrete version of (2.1).

The GE has been represented more compactly as

$$H_q(f) = -\int_0^\infty f^q(x) L_q(f(x))\,\mathrm{d}x, \qquad q > 0,$$

where

$$L_q(z) = \begin{cases} \dfrac{z^{1-q} - 1}{1-q}, & z > 0,\ q \in \{0\} \cup \Omega, \\ \log z, & z > 0,\ q = 1, \end{cases} \tag{2.3}$$

is known as the generalized logarithm function with $L_q(z) \to \log z$ as $q \to 1$.

The generalized logarithm (2.3) provides the following generalizations of the cumulative hazard function and the hazard function (1.2):

$$\Lambda_q^*(x) = -L_q(\bar{F}(x)) = \begin{cases} \dfrac{1}{q-1}\left[\dfrac{1 - \bar{F}^{q-1}(x)}{\bar{F}^{q-1}(x)}\right], & q \in \Omega, \\ \Lambda(x), & q = 1, \end{cases}$$

$$\lambda_q^*(x) = -\frac{\mathrm{d}L_q(\bar{F}(x))}{\mathrm{d}x} = \frac{f(x)}{[\bar{F}(x)]^q}, \qquad q > 0, \tag{2.4}$$

where $\Lambda_q^*(x)$ and $\lambda_q^*(x)$ are known as the *generalized odds ratio* and the *generalized odds rate* of the nonnegative random variable X, respectively (see [10]). The cumulative hazard function $\Lambda(x)$ is the limiting case of $\Lambda_q^*(x)$ as $q \to 1$.

The generalized logarithm function $L_q(z)$ is pseudo-additive, in that,

$$L_q(yz) = L_q(y) + L_q(z) + (1-q)L_q(y)L_q(z), \qquad y, z > 0,\ q > 0. \tag{2.5}$$

The solution to $L_q(yz) = 0$ is $y = z^{-1}$ and $L_q(z^{-1}) = -\overline{L}_q(z)$, where

$$\overline{L}_q(z) = \begin{cases} \dfrac{z^{q-1} - 1}{q-1}, & q \in \Omega, \\ \log z, & q = 1. \end{cases} \tag{2.6}$$

We will use two properties of $\overline{L}_q(z)$ given by the following lemma.

Lemma 2.1. (i) *Let $\phi_i(x) \geq 0$ such that $\int_0^\infty \phi_i(x)\,\mathrm{d}x = \mu_i < \infty$, $i = 1, 2$, and $\overline{L}_q(z)$ be as defined in (2.6). Then*

$$D_q(\phi_1 : \phi_2) = \int_0^\infty \phi_1(x)\overline{L}_q\left(\frac{\phi_1(x)}{\phi_2(x)}\right)\mathrm{d}x \geq \mu_1 \overline{L}_q\left(\frac{\mu_1}{\mu_2}\right).$$

(ii) *For all values of $q > 0$ and $z > 0$, $\overline{L}_q(z) \geq 1 - 1/z$.*

Proof. The proofs of all results can be found in Appendix A. □

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

Lemma 2.1(i) is a generalization of the log-sum inequality proved by Borland *et al.* [9]. For functions with equal measures such as PDFs, where $\mu_1 = \mu_2 = 1$, $D_q(\phi_1 : \phi_2) \geq 0$ is a divergence function and provides a generalization of the Kullback–Leibler information, which is given by $q = 1$.

The GE is suitable for the nonextensive system where the distribution of a random variable is given by the *escort distribution of order* q defined as the qth power of another distribution. Discrete escort distributions are prevalent in nonextensive statistical mechanics and source coding (see [8]) and nanothermodynamics (see [24]). The second equality in (1.4) represents p_α as the escort PDF of order α of p. This makes the GE suitable for the ED of X. For $q = \alpha$, (2.1) yields

$$H_\alpha(p) = \frac{C_\alpha - 1}{1 - \alpha}, \qquad \alpha \in \Omega.$$

In (1.1), the survival function of X_α is represented as the escort survival function of order α of $\bar{F}$. This makes the PH model similar to the nonextensive systems of statistical mechanics where the escort distributions are discrete and defined in terms of probability. We define the *GE functional of the survival function* of order α by

$$h_\alpha(\bar{F}) = -\int_0^\infty \bar{F}^\alpha(x) L_\alpha(\bar{F}(x))\,\mathrm{d}x, \qquad \alpha > 0. \tag{2.7}$$

The limiting case $h_\alpha(\bar{F}), \alpha \to 1$ yields the Shannon entropy functional of the survival function $h_1(\bar{F}) = h(\bar{F})$. Rao *et al.* [19] introduced the $\alpha = 1$ case in terms of $\mathbb{P}(|X| > x)$ and referred to it as the cumulative residual entropy. Zografos and Nadarajah [28] defined some other generalizations of $h(\bar{F})$ and referred to them as the survival exponential entropies. Asadi *et al.* [6] referred to $h(\bar{F})$ as the entropy functional of the survival function.

The GE functional (2.7) is a measure of the concentration of the distribution. That is, $h_\alpha(\bar{F}) \geq 0$ where the equality holds if and only if the distribution is degenerate. This can be seen by noting that, for $0 < z \leq 1$, $L_\alpha(z) \leq 0$ and the equality holds if and only if $z = 1$. This property implies that $h_\alpha(\bar{F}) \geq 0$ which further implies that $h_\alpha(\bar{F}) = 0$ if and only if the integrand in (2.7) is 0; i.e. $\bar{F}^\alpha(x) L_\alpha(\bar{F}(x)) = 0$, which holds if and only if $\bar{F}(x) = 0$ or 1.

The pseudo-additive property (2.5) implies that for two independent random variables X and Y, $H_q(X, Y)$ and $h_q(X, Y)$ are pseudo-additive as in (2.5). For two independent continuous nonnegative random variables with survival functions $\bar{F}_X$ and $\bar{F}_Y$, $\bar{F}_{X,Y}(x, y) = \bar{F}_X(x)\bar{F}_Y(y)$. The survival function integrates to the mean and (2.5) yields

$$h_\alpha(\bar{F}_X \bar{F}_Y) = \mathbb{E}(Y_\alpha)h_\alpha(\bar{F}_X) + \mathbb{E}(X_\alpha)h_\alpha(\bar{F}_Y) + (1-\alpha)h_\alpha(\bar{F}_X)h_\alpha(\bar{F}_Y), \qquad \alpha > 0, \tag{2.8}$$

provided that all measures exist. For $\alpha < (>)1$, the last term in (2.8) is positive (negative), thus,

$$h_\alpha(\bar{F}_X \bar{F}_Y) \geq (\leq)\mathbb{E}(Y_\alpha)h_\alpha(\bar{F}_X) + \mathbb{E}(X_\alpha)h_\alpha(\bar{F}_Y), \qquad \alpha \leq (\geq)1. \tag{2.9}$$

That is, $h_\alpha(\bar{F}_X \bar{F}_Y)$ is super-additive for $\alpha < 1$, sub-additive for $\alpha > 1$, and additive for $\alpha = 1$. For $\alpha = 1$, (2.9) yields the bivariate version of Theorem 4 of [19].

3. Bayes risk of PH mean excess

The residual or excess of a nonnegative random variable, given that it exceeds a threshold τ, denoted as $X - \tau \mid X > \tau$, is of interest in various fields. The PDF of the residual random variable is

$$f(x; \tau) = \frac{f(x)}{\bar{F}(\tau)}, \qquad x > \tau. \tag{3.1}$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

The mean residual life (MRL) function of X with a finite mean μ is defined as

$$m(\tau) = \mathbb{E}_{X>\tau}(X - \tau \mid X > \tau), \qquad \tau \geq 0, \tag{3.2}$$

where $\mathbb{E}_{X>\tau}$ denotes the expectation with respect to the residual PDF (3.1).

The MRL of X_α, denoted as $m_\alpha(\tau)$, is given by (3.2) with X_α in place of X. Under the quadratic loss function $\mathcal{L}(d, X_\alpha \mid \tau) = (X_\alpha - \tau - d)^2$, $X_\alpha > \tau$, the MRL $m_\alpha(\tau)$ is the optimal decision for prediction of the excess:

$$d^*(\tau) = \arg\min_d \mathbb{E}_{X_\alpha>\tau}[\mathcal{L}(d, X_\alpha \mid \tau)] = m_\alpha(\tau), \qquad \alpha > 0.$$

The MRL $m_\alpha(\tau)$ is a local risk measure, conditional on the threshold τ. Its global risk is the Bayes risk $\mathcal{E}(m_\alpha) = \mathbb{E}_\pi[m_\alpha(\tau)]$, where $\pi(\tau)$ is a prior distribution (weight function) for the threshold. Ardakani *et al.* [2] proposed ranking models by the Bayes risk of the mean excess of the absolute error of the forecast model $\mathcal{E}[m(\tau)]$ with $\pi(\tau) = f(\tau), \tau \geq 0$. Asadi and Zohrevand [4] have shown that, for $\pi(\tau) = f(\tau), \tau \geq 0, \mathcal{E}(m_1) = h_1(\bar{F})$. Owing to $\bar{F}(x;\tau) < 1$ for all $x > \tau$, the MRL function m_α is decreasing in $\alpha, \alpha > 0$. Thus, under any prior for the threshold, $\mathcal{E}(m_\alpha)$ is a decreasing function of α.

In the following theorem we state $\mathcal{E}(m_\alpha)$ under three priors for the threshold τ.

Theorem 3.1. *Let X_α be a nonnegative continuous random variable with survival function $\bar{F}^\alpha, \alpha > 0$, and MRL function $m_\alpha(\tau)$.*

(i) *Under the baseline prior $\pi_1(\tau) = f(\tau), \tau \geq 0$, the Bayes risk of $m_\alpha(\tau)$ is given by the GE functional of the baseline survival function, $\mathcal{E}_1(m_\alpha) = h_\alpha(\bar{F}), \alpha > 0$.*

(ii) *Under the PH prior $\pi_2(\tau \mid \alpha) = f_\alpha(\tau), \tau \geq 0$, the Bayes risk of $m_\alpha(\tau)$ is given by the Shannon entropy functional of the PH survival function, $\mathcal{E}_2(m_\alpha) = h(\bar{F}_\alpha), \alpha > 0$.*

(iii) *Under the ED of the PH prior $\pi_3(\tau \mid \alpha) = p_\alpha(\tau), \tau \geq 0$, the Bayes risk of $m_\alpha(\tau)$ is given by the expected asymptotic age of the renewal process with the arrival time distribution F_α,*

$$\mathcal{E}_3(m_\alpha) = \mathbb{E}(A_{X_\alpha}) = \frac{\mathbb{E}(X_\alpha^2)}{2\mathbb{E}(X_\alpha)}, \qquad \alpha > 0.$$

In the following corollary we provide comparisons of the three Bayes risks in Theorem 3.1 and their connections to the Shannon entropy.

Corollary 3.1. *Under the conditions of Theorem 3.1, we have the following.*

(i) *The Bayes risk of Theorem 3.1(i) is related to the Bayes risk of the baseline MRL under the PH prior for τ as follows:*

$$\mathcal{E}_1(m_\alpha) = \mathbb{E}_{\pi_1(\tau)}[m_\alpha(\tau)] = \frac{1}{\alpha}\mathbb{E}_{\pi_2(\tau\mid\alpha)}[m(\tau)] = \frac{1}{\alpha}\mathcal{E}_{2,\alpha}(m), \qquad \alpha > 0, \tag{3.3}$$

where $\mathcal{E}_{2,\alpha}(m) = \int_0^\infty f_\alpha(\tau)m(\tau)\,\mathrm{d}\tau$.

(ii) *The Bayes risks of Theorem 3.1(ii) and 3.1(iii) are related to the Shannon entropy of the ED of the PH model as follows:*

$$\mathcal{E}_2(m_\alpha) = \mu_\alpha[H(p_\alpha) - \log\mu_\alpha] \leq \mathcal{E}_3(m_\alpha), \qquad \alpha > 0. \tag{3.4}$$

Clearly, for $\alpha = 1$, $\mathcal{E}_1(m_\alpha) = \mathcal{E}_2(m_\alpha)$, which connects the Bayes risk of Theorem 3.1(i) to the Shannon entropy of the ED of the baseline model. The connection of $H(p_\alpha)$ with $\mathcal{E}_2(m_\alpha)$ is direct. But $H(p_\alpha)$ together with μ_α provide a lower bound $\mathcal{E}_3(m_\alpha)$.

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

3.1. Application examples

Application areas of Theorem 3.1 and Corollary 3.1 include wealth inequality, reliability, and actuarial science, among others. Applications to wealth inequality will be discussed in Section 4. In this section we present examples of potential applications in reliability and actuarial science. We use the following notions.

Definition 3.1. (i) A random variable X with survival function $\bar{F}_X$ is said to be stochastically smaller than or equal to random variable Y with survival function $\bar{F}_Y$, denoted by $X \leq_{\text{st}} Y$, if $\bar{F}_X(x) \leq \bar{F}_Y(x)$ for all x.

(ii) A lifetime model with survival function $\bar{F}$ and mean μ is said to be new better (worse) than used, NBUE (NWUE) if $m(\tau) \leq (\geq)\mu$ for all $\tau > 0$.

We also use the following expressions for $\mathcal{E}_1(m_\alpha)$. From (2.1), we have

$$\mathcal{E}_1(m_\alpha) = \frac{1}{\alpha - 1}\left[\int_0^\infty \bar{F}(x)\,\mathrm{d}x - \int_0^\infty \bar{F}^\alpha(x)\,\mathrm{d}x\right], \qquad \alpha \in \Omega, \tag{3.5}$$

$$= \frac{1}{\alpha - 1}(\mu - \mu_\alpha), \qquad \alpha \in \Omega. \tag{3.6}$$

3.1.1. *Weibull renewal process.* Consider the Weibull renewal processes (see [15] and [25]) where the distribution of the renewal time X_α is PH with Weibull $W(\beta, 1)$ baseline and survival function

$$\bar{F}(x) = \mathrm{e}^{-x^\beta}, \qquad x > 0,\ \beta > 0. \tag{3.7}$$

For $0 < \beta < 1$, the process is a Cox process (see [25] and [26]). In Table 1 we present the Bayes risks $\mathcal{E}_j(m_\alpha)$, $j = 1, 2, 3$, of the MRL given in Theorem 3.1 for this model.

The entries are found as follows. The PH model is $W(\beta, \lambda_\alpha)$, $\lambda_\alpha = \alpha^{-1/\beta}$. Using (3.7) and $\bar{F}_\alpha$ in (3.5), we obtain $\mathcal{E}_1(m_\alpha)$. In the left panel of Figure 1 we present the plot of $\mathcal{E}_1(m_\alpha)$ for $W(\beta, 1)$, $\beta \in (0.5, 3)$, and $\alpha \in (0.5, 3)$. The expression for $\mathcal{E}_2(m_\alpha)$ is found as follows:

$$h(\bar{F}_\alpha) = \alpha \int_0^\infty x^\beta \bar{F}_\alpha(x)\,\mathrm{d}x = \frac{\alpha}{\beta + 1}\mathbb{E}(X_\alpha^{\beta+1}),$$

where the moments of the PH models are given by

$$\mathbb{E}(X_\alpha^k) = \lambda_\alpha^k \Gamma\left(1 + \frac{k}{\beta}\right) = \alpha^{-k/\beta}\mathbb{E}(X^k), \qquad \alpha, k > 0.$$

TABLE 1: Expressions for $\mathcal{E}_j(m_\alpha)$, $j = 1, 2, 3$, of Theorem 3.1 for the Weibull baseline model.

Prior	Bayes risk $\mathcal{E}_j(m_\alpha)$	
$\pi_j(\tau),\ \tau \geq 0$	$W(\beta, 1)$	$W(2, 1)$
$\pi_1(\tau) = f(\tau)$	$\dfrac{\Gamma(1 + 1/\beta)(\alpha^{1/\beta} - 1)}{(\alpha - 1)\alpha^{1/\beta}}$	$\dfrac{\sqrt{\pi}}{2(\sqrt{\alpha} + 1)\sqrt{\alpha}}$
$\pi_2(\tau) = f_\alpha(\tau)$	$\dfrac{\Gamma(2 + 1/\beta)}{(\beta + 1)\alpha^{1/\beta}}$	$\dfrac{\sqrt{\pi}}{4\sqrt{\alpha}}$
$\pi_3(\tau) = p_\alpha(\tau)$	$\dfrac{\Gamma(1 + 2/\beta)}{2\Gamma(1 + 1/\beta)\alpha^{1/\beta}}$	$\dfrac{1}{\sqrt{\pi\alpha}}$

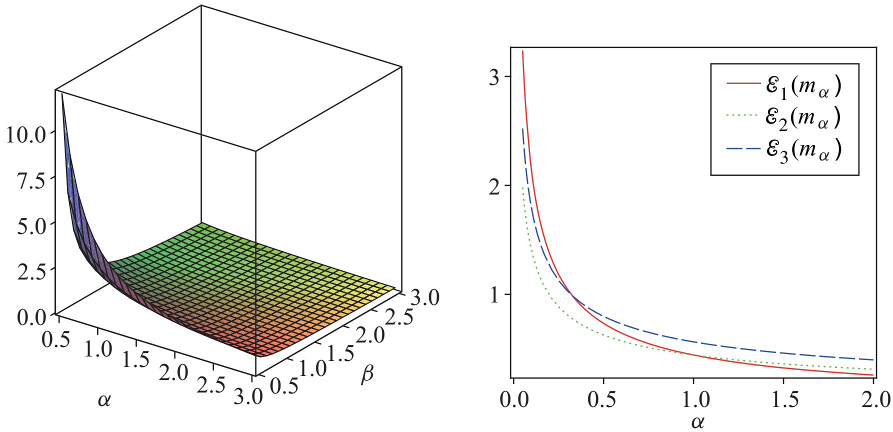

FIGURE 1: The plot of $\mathcal{E}_1(m_\alpha)$ for the Weibull distribution (*left*) and plots of $\mathcal{E}_j(m_\alpha)$ for $\beta = 2$ based on three priors $\pi_j(\tau)$, $j = 1, 2, 3$, in Theorem 3.1 (*right*).

Letting $k = \beta + 1$ yields $\mathcal{E}_2(m_\alpha)$ shown in Table 1. The expression $\mathcal{E}_3(m_\alpha)$ is found using

$$\mathbb{E}(A_{X_\alpha}) = \frac{\mathbb{E}(X_\alpha^2)}{2\mathbb{E}(X_\alpha)} = \frac{\mathbb{E}(X^2)}{2\alpha^{1/\beta}\mathbb{E}(X)}.$$

In the last column of Table 1, we have $\mathcal{E}_j(m_\alpha)$, $j = 1, 2, 3$, for $W(2, 1)$. In the right panel of Figure 1, we present the plots of these three measures. We note that $\mathcal{E}_2(m_\alpha) \le (\ge)\mathcal{E}_1(m_\alpha)$, $\alpha \le (\ge)1$. It can be shown that these inequalities hold for all distributions with decreasing $m_\alpha(\tau)$ and are reversed for distributions with increasing $m_\alpha(\tau)$.

3.1.2. *System lifetime.* Three potential reliability applications of Theorem 3.1 are as follows.

(i) For $\alpha \ge (\le)1$, $X_\alpha \le_{\text{st}} (\ge_{\text{st}})X$ implying that $\mu_\alpha \le (\ge)\mu$. It is clear that if F is NBUE (NWUE), under any prior for τ, the Bayes risk $\mathcal{E}(m_\alpha) \le (\ge)\mu$ for $\alpha \ge (\le)1$.

(ii) Consider a system with n components whose lifetimes $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) with survival function $\bar{F}$. The series system lifetime is $X_{\min} = \min\{X_1, \ldots, X_n\}$ with survival function $\bar{F}_1(t) = \bar{F}^n(t)$. The MRL of the system, given that $X_{\min} > \tau$, is

$$m_n(\tau) = \arg\min_d = \mathbb{E}_{X_{\min}>\tau}[\mathcal{L}(d, X_1, \ldots, X_n \mid \tau)], \tag{3.8}$$

where $\mathcal{L}(d, X_1, \ldots, X_n \mid \tau) = (X_{\min} - \tau - d)^2$, $X_{\min} > \tau$. The Bayes risk of (3.8) is $\mathcal{E}(m_n)$. This formulation extends to the case when the PH parameter is a rational number $\alpha = n/b$ and $X_1, \ldots, X_n$ are i.i.d. with the survival function $\bar{F}^{1/b}$; see [10]. Using (3.6), we find the Bayes risk of the series system under the prior of Theorem 3.1(i) in terms of the difference between the mean lifetimes of a component and the system as follows:

$$\mathcal{E}_1(m_n) = \frac{1}{n-1}[\mu - \mathbb{E}(X_{\min})]. \tag{3.9}$$

Thus, for large series systems the Bayes risk $\mathcal{E}_1(m_n)$ is negligible. The lifetime of the parallel system is $X_{\max}$ with the survival function $\bar{F}_n(t) = 1 - F^n(t)$. Using the binomial expansion of $(1 - \bar{F}(t))^n$ and (3.9), we find the following linear function of the Bayes

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

risks of the series systems $\mathcal{E}_1(m_k)$, $k = 2, \ldots, n$ in terms of the difference between the mean lifetimes of a component and the parallel system:

$$\sum_{k=2}^{n} c_{n,k}\mathcal{E}_1(m_k) = \mathbb{E}(X_{\max}) - \mu, \tag{3.10}$$

where $c_{n,k} = \binom{n}{k}(k-1)(-1)^k$. Equations (3.9) and (3.10) provide the range of the expected times to failure of k-out-of-n systems in terms of a linear function of the Bayes risk of the series systems $\mathcal{E}_1(m_k)$, $k = 2, \ldots, n$. For example, for $n = 3, 4$, the ranges are as follows:

$$n = 3, \qquad \mathbb{E}(X_{\max}) - \mathbb{E}(X_{\min}) = 3\mathcal{E}_1(m_2),$$
$$n = 4, \qquad \mathbb{E}(X_{\max}) - E(X_{\min}) = 6\mathcal{E}_1(m_2) - 8\mathcal{E}_1(m_3) + 6\mathcal{E}_1(m_4).$$

(iii) Consider a series system that starts operating at time 0 and fails at some point of time. It is reasonable to reuse the operating components for other systems. Under the condition that $X_{\min} = x$, the conditional distribution of the ordered lifetimes of remaining components $X_{2:n}, \ldots, X_{n:n}$, is the same as the distribution function of ordered random variables from the distribution function F which is truncated at time x (see [7]). Let $Y_i^{(1)}$, $i = 1, 2, \ldots, n-1$, denote the randomly ordered values of $X_{2:n}, \ldots, X_{n:n}$. Then Bairamov and Arnold [7] showed that the residual lifetime of $X_{2:n}, \ldots, X_{n:n}$ can be represented as

$$X_i^{(1)} = Y_i^{(1)} - X_{\min}, \qquad i = 1, \ldots, n-1,$$

and

$$\mathbb{E}(X_1^{(1)}) = \int_0^\infty \left[\int_0^\infty \frac{\bar{F}(\tau + x)}{\bar{F}(\tau)}\, \mathrm{d}\tau\right] f_{\min}(x)\, \mathrm{d}x.$$

Thus, we have $\mathcal{E}_{2,n}(m) = n\mathcal{E}_1(m_n) = \mathbb{E}(X_1^{(1)})$.

3.1.3. *Insurance deductible.* In applications to the insurer loss, the deductible amount on the policy is bounded, $\tau \le \tau_0$. A prior such as $\pi_0(\tau) = cf(\tau)$, $0 \le \tau \le \tau_0$, where $c = [F(\tau_0)]^{-1}$, serves the purpose. This yields

$$\mathcal{E}_0(m_\alpha) = \frac{\mathcal{E}(m_\alpha)}{F(\tau_0)}, \qquad \alpha > 0.$$

Note that this measure is decreasing in τ_0. As the upper bound (for the amount of deductible) increases, the risk (the insurer's expected loss) decreases. This maps the practice of insurers decreasing their average costs by increasing the maximum amount of deductible. The measure of Theorem 3.1(i) is the limit as $\tau_0 \to \infty$, hence a conservative estimate of the risk (the insurers cost). The priors in Theorem 3.1(ii) and 3.1(iii) with $0 \le \tau \le \tau_0$ yield results analogous to $\mathcal{E}_0(m_\alpha)$.

3.2. Representations and bounds.

In this section we present representations and bounds for $\mathcal{E}_1(m_\alpha)$. In the following theorem we represent $\mathcal{E}_1(m_\alpha)$ in terms of the GE of the ED of the baseline model.

Theorem 3.2. *Let X be a nonnegative continuous random variable with the survival function* $\bar{F}$*. Then*

$$\mathcal{E}_1(m_\alpha) = \mu^\alpha[H_\alpha(p) - L_\alpha(\mu)], \qquad \alpha > 0,$$

where $H_\alpha(p)$ *is the GE of the ED of the baseline model.*

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

From the following theorem we obtain the covariance representation of $\mathcal{E}_1(m_\alpha)$, which will be used to develop an upper bound.

Theorem 3.3. *Let X be a nonnegative continuous random variable with the survival function $\bar{F}$. Then under the condition that the expectations exist,*

$$\mathcal{E}_1(m_\alpha) = \frac{\alpha}{1-\alpha}\operatorname{cov}(X, \bar{F}^{\alpha-1}(X)), \qquad \alpha \in \Omega, \tag{3.11}$$

where Ω is defined in (2.1).

In the following corollary we obtain an upper bound for $\mathcal{E}_1(m_\alpha)$ in terms of the standard deviation.

Corollary 3.2. *Let X be a nonnegative continuous random variable with survival function $\bar{F}$, mean μ, and standard deviation (SD) $\sigma_x < \infty$. Then*

$$\mathcal{E}_1(m_\alpha) \le \frac{\sigma_x}{\sqrt{2\alpha - 1}}, \qquad \alpha > \frac{1}{2},\ \alpha \ne 1. \tag{3.12}$$

Remark 3.1. The bound (3.12) holds for $\alpha = 1$ and the inequality becomes an equality if and only if F is exponential (see [3, Theorem 3.1]).

The SD bound in (3.12) is decreasing in α, but it is applicable when $\alpha > \frac{1}{2}$. Next we obtain a bound for $\mathcal{E}_1(m_\alpha)$ which is defined for $0 < \alpha < 2$, $\alpha \ne 1$, and will be referred to as the GPD bound because it is based on the generalized Pareto distribution with survival function

$$\bar{F}_Y(y) = \left(1 + \frac{(\alpha - 1)y}{c}\right)^{1/(1-\alpha)}, \qquad y \in \mathcal{S}_\alpha,\ c > 0, \tag{3.13}$$

where

$$\mathcal{S}_\alpha = \begin{cases} \left\{y\colon 0 \le y \le b = \dfrac{c}{1-\alpha},\ 0 < \alpha < 1\right\}, \\ \{y\colon y \ge 0,\ 1 < \alpha < 2\}. \end{cases} \tag{3.14}$$

Theorem 3.4. *Let X be a nonnegative continuous random variable with the survival function $\bar{F}$. If the support of $\bar{F}$ is contained in $\mathcal{S}_\alpha$ defined in (3.14) and $\mathbb{E}(X)$ and $\mathbb{E}(X_\alpha^2)$ are finite, then*

$$\mathcal{E}_1(m_\alpha) \le \nu_\alpha \left(\frac{\mathbb{E}(X_\alpha^2)}{\mathbb{E}^\alpha(X)}\right)^{1/(2-\alpha)} + \omega_\alpha[\mathbb{E}^{1-\alpha}(X_\alpha^2)\mathbb{E}^\alpha(X)]^{1/(2-\alpha)} - \mathbb{E}(X), \qquad 0 < \alpha < 2,\ \alpha \ne 1, \tag{3.15}$$

where $\nu_\alpha = [2(2-\alpha)]^{1/(\alpha-2)}$ and $\omega_\alpha = [2(2-\alpha)\nu_\alpha]^{-1}$, $0 < \alpha < 2$.

Note that, for $0 < \alpha < 1$, the support of $\bar{F}$ (also $\bar{F}_\alpha$) is bounded as follows:

$$0 \le x \le b^* = \frac{c^*}{1-\alpha},$$

where

$$c^* = (2-\alpha)\nu_\alpha \left(\frac{\mathbb{E}(X_\alpha^2)}{\mathbb{E}^\alpha(X)}\right)^{1/(2-\alpha)},$$

and $\mathbb{E}^{1-\alpha}(X_\alpha^2)\mathbb{E}^\alpha(X)$ is the geometric mean of $\mathbb{E}(X_\alpha^2)$ and $\mathbb{E}(X)$.

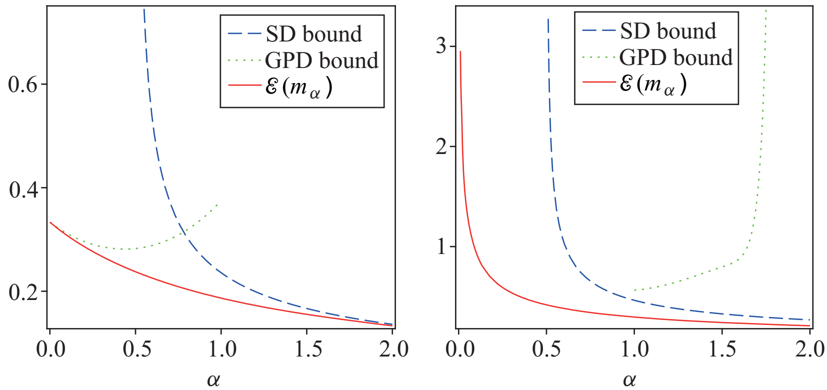

FIGURE 2: The SD and GPD bounds for the beta model (*left*) and the Weibull model (*right*).

In general, neither of the two bounds (3.12) and (3.15) uniformly dominates the other. However, as $\alpha \to 1$ from the right, the bound in (3.15) remains finite and as $\alpha \to 2$, the first term in (3.15) goes to 0 and the second term goes to ∞. This is in contrast with the fact that $\mathcal{E}_1(m_\alpha)$ is a decreasing function of α. So, the bound in (3.15) can be a useful alternative for the $\alpha < \frac{1}{2}$ case where the bound in (3.12) is not defined. In the following example we illustrate these points.

Example 3.1. (i) *Beta baseline*. Let the baseline distribution be the one-parameter beta with the survival function $\bar{F}(x) = 1 - x^2$, $0 \le x \le 1$. Then

$$\mathbb{E}(X_\alpha) = \int_0^1 \bar{F}^\alpha(x)\,\mathrm{d}x = \int_0^1 (1-x^2)^\alpha\,\mathrm{d}x = \frac{\sqrt{\pi}\,\Gamma(\alpha+1)}{2\Gamma(\alpha+3/2)}.$$

From (3.6), we obtain

$$\mathcal{E}_1(m_a) = \frac{1}{\alpha-1}\bigg(\frac{2}{3} - \frac{\sqrt{\pi}\,\Gamma(\alpha+1)}{2\Gamma(\alpha+3/2)}\bigg).$$

The variance is $\sigma_x^2 = \frac{1}{18}$ and (3.12) yields

$$\mathcal{E}_1(m_\alpha) \le \frac{1}{\sqrt{18(2\alpha-1)}}, \qquad \alpha > \frac{1}{2}.$$

The distributional ingredients of the upper bound (3.15) are as follows: $\mathbb{E}(X) = \frac{2}{3}$ and $\mathbb{E}(X_\alpha^2) = 1/(\alpha+1)$. For $0 < \alpha < 1$, $b^* > 1$, so $\bar{F}_X(x)/\bar{F}_Y(x) > 0$ for $0 \le x \le 1$ and (3.15) is well defined. In the left panel of Figure 2 we present the plots of the SD and GPD bounds along with the plot of $\mathcal{E}_1(m_\alpha)$. For this model, the bound (3.15) is preferred for $\alpha < 0.63$ and the bound (3.12) is preferred for $\alpha > 0.63$.

(ii) *Weibull renewal process*. Consider the baseline survival function (3.7) with $\beta = 2$. The variance of X is $\sigma_x^2 = 1 - \pi/4$. Thus,

$$\mathcal{E}_1(m_\alpha) \le \sqrt{\frac{1-\pi/4}{2\alpha-1}}, \qquad \alpha > \frac{1}{2}.$$

In the right panel of Figure 2 we present the plots of the SD and GPD bounds given in (3.12) and (3.15) along with the plot of $\mathcal{E}_1(m_\alpha)$ given in Table 1. For this model, the GPD bound is not defined for $\alpha < 1$ because the support of the GPD is bounded and the support of the Weibull distribution is not. For $\alpha > 1$, the SD bound is close to $\mathcal{E}_1(m_\alpha)$.

4. Gini measures

The Gini coefficient of a wealth X with distribution F, PDF f, and mean $\mu < \infty$ is defined by the expected distance between pairs of random draws X_1 and X_2 from X:

$$G(X) = \frac{1}{2\mu}\int_0^\infty \int_0^\infty f(x_1)f(x_2)|x_1 - x_2| \,\mathrm{d}x_1 \,\mathrm{d}x_2.$$

This measure is also referred to as the relative Gini index, and $0 \le G(X) \le 1$, where $G(X) = 0$ when the wealth is distributed uniformly and $G(X) = 1$ indicates full concentration of the wealth at a single point $\mathbb{P}(X = x_0) = 1$. We use the following well-known representation of $G(X)$ (see [27]):

$$G(X) = \frac{1}{\mu}\int_0^\infty F(x)\bar{F}(x)\,\mathrm{d}x.$$

It is easy to see that

$$G(X) = \frac{1}{\mu}h_2(\bar{F}) = \frac{1}{\mu}h_2(F), \tag{4.1}$$

where $h_\alpha(F)$ is the GE functional of the distribution function. The reversed PH model is defined by $F^*_\alpha(x) = F^\alpha(x)$, $x \ge 0$. It can be shown that $\mathcal{E}_1(m^*_\alpha) = h_\alpha(F)$, where

$$m^*_\alpha(\tau) = \mathbb{E}(\tau - X_\alpha \mid X_\alpha \le \tau) = \frac{\int_0^\tau F^\alpha(x)\,\mathrm{d}x}{F^\alpha(\tau)}$$

is the mean inactivity time of X_α; see [11] and [12] for results on the past lifetime. From (4.1) and Theorem 3.1(i) with $\alpha = 2$, we have

$$G(X) = \frac{1}{\mu}\mathcal{E}_1(m_2) \tag{4.2}$$

$$= \frac{1}{2\mu}[\mathcal{E}_1(m_2) + \mathcal{E}_1(m^*_2)]. \tag{4.3}$$

Equation (4.2) represents the Gini index in terms of the Bayes risk of the minimum of two wealths $X_{\min} = \min\{X_1, X_2\}$ that exceeds a threshold τ:

$$G(X) = \frac{1}{\mu}\mathbb{E}_\pi\{\mathbb{E}_{X_{\min}>\tau}[X_{\min} - \tau \mid X_{\min} > \tau]\}.$$

Equation (4.3) represents the Gini index in terms of the Bayes risks of the MRL function of the PH model and the mean inactivity time of the reversed PH model with $\alpha = 2$.

From (4.2), (3.3), and Theorem 3.2, we obtain the following representations of $G(X)$.

Corollary 4.1. *The Gini coefficient of the wealth X with a mean $\mu < \infty$ is related to the GE of the ED of X_2 as follows:*

$$G(X) = \frac{1}{2\mu}\mathcal{E}_{2,2}(m) = \frac{\mu}{2}[H_2(p) - \mu + 1],$$

where $\mathcal{E}_{2,\alpha}(m)$ is defined in Corollary 3.1.

Remark 4.1. (i) Corollary 4.1 establishes a relationship between the Gini coefficient and the GE used by Maasoumi [16] for measuring multivariate wealth inequality. Letting $\alpha = \gamma + 1$ in the result of Corollary 4.1 and multiplying by $k = -\gamma$ yields the result for Maasoumi's measure.

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

(ii) The absolute Gini index, also known as Gini's mean difference, is defined without the scaling by μ, which has various representations including in terms of $\mathrm{cov}(X, F(X))$ and the expectation of the difference between the maximum and minimum of two random draws X; see [22] and [27] and the references therein. Shalit and Yitzhaki defined the extended Gini as $G_\nu(X) = -\upsilon\,\mathrm{cov}\,(X, \bar{F}^{\nu-1})$, $\nu > 1$ (see [22, Equation (4)]. Clearly, $G_\nu(X) = \mathcal{E}_1(m_{\nu/(\nu-1)})$, $\nu > 1$.

In the next theorem we state a class of bounds for the Gini coefficient.

Theorem 4.1. *Let X and Y be continuous random variables with finite means μ_x, μ_y and survival functions $\bar{F}_X$ and $\bar{F}_Y$ such that $X \le_{\mathrm{st}} Y$.*

(i) *The GE functionals of $\bar{F}_X$ is bounded as*

$$h_\alpha(\bar{F}_X) \le h_\alpha(\bar{F}_Y) - \mu_x \overline{L}_\alpha\left(\frac{\mu_x}{\mu_y}\right), \tag{4.4}$$

where $\overline{L}_\alpha$ is defined in (2.6).

(ii) *The Gini coefficient of X is bounded as*

$$G(X) \le \frac{1}{\rho} G(Y) - \rho + 1, \qquad \rho \le 1, \tag{4.5}$$

where $\rho = \mu_x/\mu_y$.

Note that $X \le_{\mathrm{st}} Y$ implies $\mu_x \le \mu_y$ which implies that $\overline{L}_\alpha(\mu_x/\mu_y) < 0$. Hence, the bound in (4.4) is greater than $h_\alpha(\bar{F}_Y)$. In particular, if we let $\alpha \to 1$ then the inequality in (4.4) yields the result of Navarro *et al.* [17, Proposition 2.1].

In the following example we present some applications of Theorem 4.1.

Example 4.1. (*Bounds for wealth inequality.*) In Table 2 we present the ingredients for computing the upper bound (4.5) for the Gini coefficients of all distributions stochastically dominated by each distribution shown in the table. The Gini coefficient for these distributions is known and for given μ_x, the mean ratio ρ can easily be computed. For example, for any

TABLE 2: Ingredients for computing the upper bound (4.5) for the Gini index $G(X)$. Note that $B(\cdot, \cdot)$ is the beta function and $\mathrm{erf}(\cdot)$ is the error function.

Wealth distribution	Gini coefficient		
	$f_Y(y)$	$G(Y)$	$\rho = \mu_x/\mu_y$
Weibull	$\frac{\beta}{\lambda^\beta} y^{\beta-1} \mathrm{e}^{-(y/\lambda)^\beta}$	$1 - 2^{-1/\beta}$	$\frac{\mu_x}{\lambda\Gamma(1+1/\beta)}$
gamma	$\frac{1}{\lambda^\beta \Gamma(\beta)} y^{\beta-1} \mathrm{e}^{-y/\lambda}$	$\frac{1}{\beta B(\beta, 1/2)}$	$\frac{\mu_x}{\lambda\beta}$
Pareto	$\frac{\beta y_0^\beta}{y^{\beta+1}},\ y \ge y_0 > 0$	$\frac{y_0}{2\beta - 1},\ \beta > 1$	$\frac{(\beta-1)\mu_x}{y_0},\ \beta > 1$
lognormal	$\frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log(y) - \mu)^2\right)$	$\mathrm{erf}\left(\frac{\sigma}{2}\right)$	$\mu_x \mathrm{e}^{-(\mu+\sigma/2)}$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

distribution F_X stochastically dominated by the exponential distribution with $\lambda = 1$, using the entries in the third and fourth columns of the first row of Table 2 with $\beta = 1$ in (4.5) yields

$$G(X) \leq \frac{1}{\mu_x} - \mu_x + 1, \qquad 0 \leq \mu_x \leq 1.$$

5. Bayes risk of Fisher information

In the following theorem we state the Fisher information of the ED (1.4) about α.

Theorem 5.1. *The Fisher information provided by the PDF of the ED of the PH model about α is given by*

$$I_{p_\alpha}(\alpha) = \mathbb{E}_{p_\alpha}\left[\frac{\partial \log p_\alpha(X)}{\partial \alpha}\right]^2 = \text{var}_{p_\alpha}[\Lambda(X)], \tag{5.1}$$

where $\mathbb{E}_{p_\alpha}$ and var_{p_α} denote the expectation and variance with respect to the ED (1.4) and $\Lambda(x)$ is the baseline cumulative hazard function (1.2).

In (5.1), $I_{p_\alpha}(\alpha)$ is represented as the minimum of the quadratic loss function for estimating $\Lambda(X)$ by $\mathbb{E}_{p_\alpha}[\Lambda(X)]$. In the following theorem, we state the Bayes risk of $I_{p_\alpha}(\alpha)$ under two proper prior distributions and an improper uniform prior. Without loss of generality, we consider the $\alpha \geq 1$ case. For the $\alpha \leq 1$ case the results apply to $\alpha' = 1/\alpha$ in place of α.

Theorem 5.2. *The Bayes risks of Fisher information (5.1) under three prior distributions for α are as follows.*

(i) *Under the Pareto prior $\pi_\theta(\alpha) = \theta\alpha^{-(\theta+1)}$, $\alpha \geq 1, \theta > 0$, if, for all x,*

$$\lim_{\alpha\to\infty} \alpha^{\theta+1} p_\alpha(x) \to 0,$$

then the Bayes risk is

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \theta\{\mathbb{E}_{p^*}[\Lambda(X)] + \mathbb{E}_p[\Lambda(X)]\},$$

where $\mathbb{E}_{p^*}$ denotes the expectation with respect to the PDF of the prior predictive distribution given by $p^*(x) = \int p_\alpha(x)\pi_{\theta+1}(\alpha)\,\mathrm{d}\alpha$.

(ii) *Under the proper uniform prior $\pi(\alpha) = 1/(a-1)$, $1 \leq \alpha \leq a$, the Bayes risk is*

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \frac{1}{a-1}\{\mathbb{E}_{p_a}[\Lambda(X)] + \mathbb{E}_p[\Lambda(X)]\}.$$

(iii) *If $\lim_{\alpha\to\infty} p_\alpha(x) \to 0$ for all x then under the improper uniform prior $\pi(\alpha) \propto 1$, $\alpha \geq 1$, the Bayes risk is*

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] \propto \mathbb{E}_p[\Lambda(X)].$$

In the following corollary we connect the Shannon entropy of the ED of the baseline hazard model, p, to the Bayes risk of Fisher information of its escort distribution p_α about the PH parameter.

Corollary 5.1. *In Theorem 5.2,*

$$\mathbb{E}_p[\Lambda(X)] = \frac{1}{\mu}h_1(\bar{F}) = \frac{1}{\mu}\mathcal{E}_1(m),$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

and under the condition of part (iii), the Bayes risk of Fisher information is related to the Shannon entropy of the ED of the baseline model as follows:

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] \propto \frac{1}{\mu} h_1(\bar{F}) = H(p) - \log \mu. \tag{5.2}$$

In the following example we give some applications of Theorem 5.2 for a renewal process.

Example 5.1. (*Weibull renewal process.*) Consider the baseline survival function (3.7). The ED of X_α is the generalized gamma $\mathrm{GG}(1/\beta, \beta, \lambda_\alpha)$, $\lambda_\alpha = \alpha^{-1/\beta}$ with PDF

$$p_\alpha(x) = \frac{1}{\lambda_\alpha \Gamma(1/\beta)} \mathrm{e}^{-(x/\lambda_\alpha)^\beta}, \qquad x > 0,\ \alpha, \beta, \lambda_\alpha > 0.$$

Using the expression for the moments of this distribution,

$$\mathbb{E}_{p_\alpha}(X^k) = \frac{\Gamma(1/\beta + k/\beta)}{\alpha^{k/\beta}\Gamma(1/\beta)},$$

we find that

$$I_{p_\alpha}(\alpha) = \mathrm{var}_{p_\alpha}[\log \bar{F}(X)] = \frac{1}{\alpha^2 \beta}.$$

Under the Pareto and uniform priors given in Theorem 5.2(i)–(iii), we obtain, respectively,

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \frac{\theta}{(\theta+2)\beta}, \qquad \mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \frac{1}{\beta}, \qquad \mathbb{E}_\pi[I_{p_\alpha}(\alpha)] \propto \frac{1}{\beta}.$$

The assumption of Theorem 5.2(iii) is met, hence, the relationship (5.2) holds.

6. Shannon measures

In the following theorem we state the Shannon entropy of the residual distribution of the ED of X_α, denoted by $H(p_\alpha; \tau)$, and a representation of (5.2).

Theorem 6.1. *Let $\lambda_{p_\alpha}(\tau)$ denote the hazard rate function of the ED (1.4). Then*

$$H(p_\alpha; \tau) = \alpha^2 \frac{\partial}{\partial \alpha}\bigg(\frac{\log \lambda_{p_\alpha}(\tau)}{\alpha}\bigg) \tag{6.1}$$

$$= -\alpha^2 \frac{\partial}{\partial \alpha}\bigg(\frac{\log m_\alpha(\tau)}{\alpha}\bigg); \tag{6.2}$$

and under the improper uniform prior $\pi(\alpha) \propto 1$, $\alpha \geq 1$,

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] \propto H(p, 0) - \log \mu.$$

Thus far, the Shannon entropy and Fisher information of a continuous PDF are defined and computed by integration. Representations of these measures in terms of a derivative is novel.

The final issue that we address is the comparison of the entropies of the PH and baseline models given by the following theorem.

Theorem 6.2. (i) *The Shannon entropy of PH model is given by*

$$H(f_\alpha) = H(f) - K(f_\alpha : f) - \alpha \operatorname{cov}(\log f(X), \bar{F}^{\alpha-1}(X)), \tag{6.3}$$

where $K(f_\alpha : f)$ *is the Kullback–Leibler divergence between the PH and baseline models,*

$$K(f_\alpha : f) = \int_0^\infty f_\alpha(x) \log \frac{f_\alpha(x)}{f(x)} \, \mathrm{d}x = \log \alpha + \frac{1}{\alpha} - 1. \tag{6.4}$$

(ii) *It holds that* $H(f_\alpha) \le (\ge) H(f)$ *if and only if*

$$\operatorname{cov}(\log f(X), \bar{F}^{\alpha-1}(X)) \ge (\le) \frac{1}{\alpha}\left(1 - \log \alpha - \frac{1}{\alpha}\right). \tag{6.5}$$

(iii) *It holds that* $\log f(X)$ *and* $\bar{F}^{\alpha-1}(X)$ *are uncorrelated if and only if*

$$K(f_\alpha : f) = H(f) - H(f_\alpha). \tag{6.6}$$

Some important consequences of Theorem 6.2 are as follows. Equation (6.6) provides a new condition for the Kullback–Leibler divergence between two distributions to be equal to the difference between the respective entropies. Thus far, it is known that such an equality holds for $K(f : f^*)$ when f is in a class of distributions with given moments, where f^* is the maximum entropy model. Theorem 6.2(ii) provides a new condition. For example, (6.6) holds for $\alpha = 2$ when f is in a class of models where $\log f(x)$ is a symmetric function of x.

For some baseline models the entropies of the baseline and the PH models are available in closed form. But such cases are exceptional rather than being general. For many models the entropy expression for many baseline models are available, but the entropies for the corresponding PH models are not available in closed form. Yet, for a third class of models neither the entropy of the baseline model nor the entropy of the PH model is available in closed form. The order between the entropies of the baseline and PH models may be obtained using known results such as the dispersive order (see [18]) and the hazard rate and stochastic orders with an additional monotone PDF condition (see [5] and [13]). The covariance condition (6.5) applies to any baseline hazard model. When the entropy of the baseline model is available in closed form, and computation of the PH entropy requires a numerical integration procedure, (6.3) offers a simple procedure in order to compute the PH entropy. When the baseline entropy cannot be computed so easily, such as the case of a mixture model, (6.5) provides a comparison between the entropies of the baseline and PH models. In the following example we illustrate some applications of Theorem 6.2.

Example 6.1. (*Uncertainties of the PH model.*) Consider baseline models shown in Table 3. The entropies for the PH models with the Weibull, Pareto, and uniform baseline distributions are computed using the entropy expressions for the Weibull, Pareto, and symmetric beta distributions beta(β, β). The expressions for the entropy of the gamma, lognormal, and beta baseline models are available, but not for the corresponding PH models. The entropies of these PH models can be easily computed using (6.3) with the entropy expressions for these models, $H(f)$, and computing the covariance term by simulating large samples from the baseline model.

In the upper panels of Figure 3 we present the plots of the entropies of the PH models listed in Table 3, with the exception of the mixture model for which the entropy of the baseline model is not available in closed form. For the PH models whose entropies are not available in closed form the covariance term is computed using $n = 100,000$ data points simulated from the baseline model.

TABLE 3: Comparison of Shannon entropies of the baseline and PH models of Example 6.1. Note that $B(\cdot,\cdot)$ is the beta function, $\psi(\cdot)$ is the digamma function, and $\gamma = \psi(1)$ is the Euler number.

Baseline model	$f(x)$	$H(f)$	$H(f_\alpha)$
Weibull(2,1)	$2x\mathrm{e}^{-x^2},\ x \ge 0$	$1 - \log 2 + \dfrac{\gamma}{2}$	$H(f) + \dfrac{\log\alpha}{2}$
Pareto(2)	$\dfrac{2}{(1+x)^3},\ x \ge 0$	$\frac{3}{2} - \log 2$	$1 - \log(2\alpha) + \dfrac{1}{2\alpha}$
Uniform	$1,\ 0 \le x \le 1$	0	$\dfrac{1}{\alpha} - \log\alpha + 1$
gamma(2,1)	$x\mathrm{e}^{-x},\ x \ge 0$	γ	Not available
Lognormal(0,1)	$\dfrac{1}{x\sqrt{2\pi}}\mathrm{e}^{-(\log x)^2/2},\ x \ge 0$	$\frac{1}{2} + \frac{1}{2}\log(2\pi)$	Not available
beta(β,β), $\beta = 0.5, 2, 6$	$\dfrac{1}{B(\beta,\beta)}x^{\beta-1}(1-x)^{\beta-1},\ 0 \le x \le 1$	$\log B(\beta,\beta) -2(\beta-1)[\psi(\beta)-\psi(2\beta)]$	Not available
Mixture	$\frac{1}{2}\mathrm{e}^{-x} + \frac{3}{2}x^2\mathrm{e}^{-x^3},\ x \ge 0$	Not available	Not available

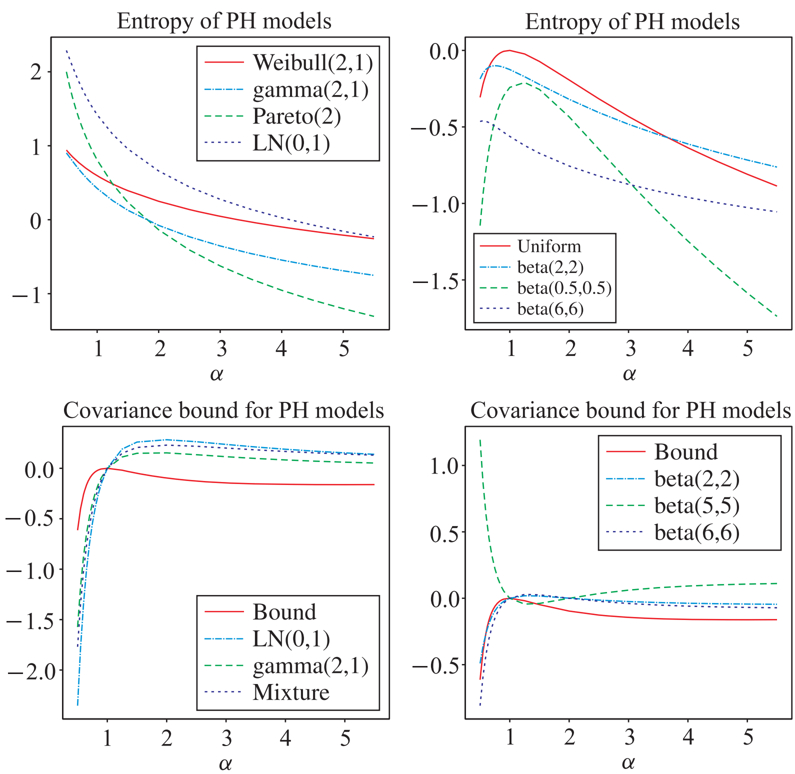

FIGURE 3: Plots of the covariance condition (6.5) for several baseline models.

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

In the lower panels of Figure 3 we present the plots of the bound and the covariance term in (6.5) against α for the baseline models listed in Table 3 for which the entropy of the PH model is not available in closed form. The left panel pertains to PH models with the gamma, lognormal, and mixture baseline models. All plots intersect at (1, 0), which is the trivial case for the covariance being 0. For these models, condition (6.5) corresponds to the order of $H(f)$ and $H(f_\alpha)$ within each family by α. But across the families and the bound, the order is reversed for $\alpha \le (\ge) 1$. The right panel pertains to PH models with three symmetric beta baseline distributions. These models are examples of distributions where $Y = \log f(X)$ is a symmetric function of X which is uncorrelated with the monotone function $U = \bar{F}(X)$. For these models, the plots of the covariance intersect for $\alpha = 1, 2$, hence, (6.6) is applicable for $\alpha = 2$.

7. Conclusions

The PH model is used in various applications in many fields. We showed that the PH model and the associated distributions also provide a rich medium for studying some important measures used in statistics, reliability and survival analysis, information theory, and related fields. The MRL function of the ED of the PH model is the optimal prediction of the excess of a random variable over a given threshold τ under the quadratic loss. We derived and compared the Bayes risks of the MRL under three priors (weight functions) for τ, showed their relationships with Shannon entropy of the ED of the PH model, and presented some potential applications in reliability and insurance.

When the prior for τ is the density of the baseline distribution, the Bayes risk of the MRL function is the GE functional of the baseline survival function. We derived two upper bounds for this Bayes risk. For $\alpha = 2$, this Bayes risk is related to the Gini coefficient of the wealth inequality and, in turn, gives the connection between the Gini coefficient and the GE. A bound for this Bayes risk provides upper bounds for the Gini coefficients of wealth distributions stochastically dominated by some known families of distributions.

We computed the Bayes risks of the Fisher information of the ED of the PH model about the PH parameter α under three priors for α. These measures are linear functions of the Shannon entropy functional of the baseline survival function. Under an improper uniform prior, the Bayes risk is proportional to a linear function of the Shannon entropy of the ED of the baseline model.

Thus far, the Shannon entropy of a random variable with continuous density is seen in the integral forms of the density and hazard rate functions. We presented two derivative representations for the Shannon entropy of the residual distribution of the ED of X_α in terms of the hazard rate and MRL functions. We also provided a covariance condition for comparison of the entropies of the PH and baseline models. This result gives the necessary and sufficient condition for ordering of the entropies of the baseline and PH models.

Appendix A. Proofs

Proof of Lemma 2.1. (i) This result is a continuous analog of the following generalized log-sum inequality:

$$\sum_{i=1}^{n} a_i \overline{L}_\alpha\left(\frac{a_i}{b_i}\right) \ge \left(\sum_{i=1}^{n} a_i\right) \overline{L}_\alpha\left(\frac{\sum_{i=1}^{n} a_i}{\sum_{i=1}^{n} b_i}\right).$$

This inequality was obtained by Borland *et al.* [9] using Jensen's inequality

$$\sum_{i=1}^{n} \pi_i \varphi(z_i) \ge \varphi\left(\sum_{i=1}^{n} \pi_i z_i\right),$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

where $\varphi(z) = z\overline{L}_\alpha(z)$ is a convex function, $z_i = a_i/b_i$, and $\pi_i = b_i/\sum_{j=1}^n b_j$, $i = 1, \ldots, n$. The continuous version is obtained by letting $z = \phi_1(x)/\phi_2(x)$ and $\pi(x) = \phi_2(x)/\mu_2$, and using $\int_0^\infty \phi_i(x)\,\mathrm{d}x = \mu_i$, $i = 1, 2$.

(ii) We show that $L_\alpha(z)$ is an increasing function of α for all values of z. Differentiating in terms of α yields

$$\frac{\mathrm{d}\overline{L}_\alpha(z)}{\mathrm{d}\alpha} = \frac{z^{\alpha-1}\log z^{\alpha-1} - z^{\alpha-1} + 1}{(\alpha-1)^2} \geq 0;$$

the inequality is implied by $\log u \geq 1 - 1/u$ for all $u = z^{\alpha-1} > 0$. □

Proof of Theorem 3.1. For any $\pi(\tau)$, $\tau \geq 0$, and $\alpha > 0$, we have

$$\begin{aligned}\mathcal{E}_\pi(m_\alpha) &= \int_0^\infty m_\alpha(\tau)\pi(\tau)\,\mathrm{d}\tau \\ &= \int_0^\infty \left(\frac{\int_\tau^\infty \bar{F}^\alpha(x)\,\mathrm{d}x}{\bar{F}^\alpha(\tau)}\right)\pi(\tau)\,\mathrm{d}\tau \\ &= \int_0^\infty \bar{F}^\alpha(x)\left(\int_0^x \frac{\pi(\tau)}{\bar{F}^\alpha(\tau)}\,\mathrm{d}\tau\right)\mathrm{d}x. \end{aligned} \tag{A.1}$$

The results are obtained by letting $\pi(\tau) = \pi_j(\tau)$, $j = 1, 2, 3$, in (A.1) as follows.

(i) With $\pi_1(\tau) = f(\tau)$, $\tau \geq 0$, in (A.1), for $\alpha \in \Omega$ as defined in (2.1), we obtain

$$\mathcal{E}_1(m_\alpha) = \int_0^\infty \bar{F}^\alpha(x)\left(\int_0^x \frac{f(\tau)}{\bar{F}^\alpha(\tau)}\,\mathrm{d}\tau\right)\mathrm{d}x = -\int_0^\infty \bar{F}^\alpha(x)L_\alpha(\bar{F}(x))\,\mathrm{d}x = h_\alpha(\bar{F}),$$

where the second equality is obtained by noting that the inner integrand in the double integral is in the form of (2.4). See [4] for the proof for $\alpha = 1$.

(ii) With $\pi_2(\tau) = f_\alpha(\tau) = \alpha f(\tau)\bar{F}^{\alpha-1}(\tau)$, $\tau \geq 0$, in (A.1), we obtain

$$\begin{aligned}\mathcal{E}_2(m_\alpha) &= \int_0^\infty \alpha\bar{F}^\alpha(x)\left(\int_0^x \frac{f(\tau)}{\bar{F}(\tau)}\,\mathrm{d}\tau\right)\mathrm{d}x \\ &= -\int_0^\infty \bar{F}^\alpha(x)\log(\bar{F}^\alpha(x))\,\mathrm{d}x \\ &= h(\bar{F}_\alpha), \qquad \alpha > 0.\end{aligned}$$

(iii) With $\pi_3(\tau) = p_\alpha(\tau)$, $\tau \geq 0$, in (A.1), we obtain

$$\begin{aligned}\mathcal{E}_3(m_\alpha) &= \int_0^\infty \bar{F}^\alpha(x)\left(\int_0^x \frac{\bar{F}^\alpha(\tau)}{\mu_\alpha\bar{F}^\alpha(\tau)}\,\mathrm{d}\tau\right)\mathrm{d}x \\ &= \frac{1}{\mu_\alpha}\int_0^\infty x\bar{F}^\alpha(x)\,\mathrm{d}x \\ &= \frac{\mathbb{E}(X_\alpha^2)}{2\mathbb{E}(X_\alpha)} \\ &= \mathbb{E}(A_{X_\alpha}), \qquad \alpha > 0;\end{aligned}$$

the last equality is well known (see [20]). □

Proof of Corollary 3.1. (i) With the PH prior $\pi_2(\tau) = \alpha f(\tau)\bar{F}^{\alpha-1}(\tau)$, we can write

$$\begin{aligned}\mathcal{E}_{2,\alpha}(m) &= \int_0^\infty m(\tau)\alpha f(\tau)\bar{F}^{\alpha-1}(\tau)\,\mathrm{d}\tau \\ &= \int_0^\infty \biggl(\int_\tau^\infty \bar{F}(x)\,\mathrm{d}x\biggr)\alpha f(\tau)\bar{F}^{\alpha-2}(\tau)\,\mathrm{d}\tau \\ &= \frac{\alpha}{\alpha-1}\int_0^\infty \bar{F}(x)\int_0^x (\alpha-1)f(\tau)\bar{F}^{\alpha-2}(\tau)\,\mathrm{d}\tau\,\mathrm{d}x \\ &= \frac{\alpha}{\alpha-1}\int_0^\infty \bar{F}(x)(1-\bar{F}^{\alpha-1}(x))\,\mathrm{d}x \\ &= \alpha\mathcal{E}_1(m_\alpha),\end{aligned}$$

where the last equality is obtained from (3.5).

(ii) The equality in (3.4) is seen by using the PDF (1.4) in (2.2) and Theorem 3.1(ii). The inequality in (3.4) is obtained by applying the bound given by Rao *et al*. [19, Theorem 10] to $h_1(\bar{F}_\alpha)$. □

Proof of Theorem 3.2. Using (3.6), for $\alpha \in \Omega$, we obtain

$$\begin{aligned}\mathcal{E}_1(m_\alpha) &= \frac{\mu^\alpha}{\alpha-1}\biggl[\mu^{1-\alpha} - \frac{\mu_\alpha}{\mu^\alpha}\biggr] \\ &= \frac{\mu^\alpha}{\alpha-1}[\mu^{1-\alpha} - C_\alpha] \\ &= \mu^\alpha\biggl[\frac{\mu^{1-\alpha}-1}{\alpha-1} - \frac{C_\alpha-1}{1-\alpha}\biggr] \\ &= -\mu^\alpha L_\alpha(\mu) + \mu^\alpha H_\alpha(p).\end{aligned}$$

The $\alpha = 1$ case is easily seen by taking the log of the two sides of (1.4). □

Proof of Theorem 3.3. Noting that the first integral in (3.5) yields $\mathbb{E}(X)$ and the second integral yields $\alpha\mathbb{E}(X\bar{F}^{\alpha-1}(X))$, we have

$$\begin{aligned}\mathcal{E}_1(m_\alpha) &= \frac{1}{\alpha-1}(\mathbb{E}(X) - \alpha\mathbb{E}(X\bar{F}^{\alpha-1}(X))) \\ &= \frac{1}{\alpha-1}(\mathbb{E}(X)\mathbb{E}(\alpha\bar{F}^{\alpha-1}(X)) - \alpha\mathbb{E}(X\bar{F}^{\alpha-1}(X))) \\ &= \frac{-\alpha}{\alpha-1}(\mathbb{E}(X\bar{F}^{\alpha-1}(X)) - \mathbb{E}(X)\mathbb{E}(\bar{F}^{\alpha-1}(X))),\end{aligned}$$

where the second equality follows from the fact that $\mathbb{E}(\alpha\bar{F}^{\alpha-1}(X)) = 1$. □

Proof of Corollary 3.2. Let $U = \bar{F}^{\alpha-1}(X)$. Then $\mathbb{E}(U^k) = 1/(k(\alpha-1)+1)$. For $k = 1, 2$, this yields

$$\sigma_u^2 = \frac{(\alpha-1)^2}{\alpha^2(2\alpha-1)}, \qquad \alpha > \frac{1}{2}.$$

Applying the Cauchy–Schwarz inequality to (3.11) yields

$$\mathcal{E}_1(m_\alpha) = \frac{\alpha}{1-\alpha}\operatorname{cov}(X, U), \qquad \alpha \neq 1,$$
$$\leq \frac{\alpha}{1-\alpha}\sigma_x\sigma_u, \qquad \alpha > \frac{1}{2},\ \alpha \neq 1. \qquad \square$$

Proof of Theorem 3.4. For two survival functions with a common support, $z = \bar{F}_X(x)/\bar{F}_Y(x)$ is well defined except for sets of measure 0. Then

$$\int_{\mathcal{S}_\alpha} \bar{F}_X(x)\overline{L}_\alpha\bigg(\frac{\bar{F}_X(x)}{\bar{F}_Y(x)}\bigg)\,\mathrm{d}x$$
$$= \int_{\mathcal{S}_\alpha} \bar{F}_X(x)\bigg(\frac{\bar{F}_X^{\alpha-1}(x)-1}{\alpha-1}\bigg)\,\mathrm{d}x + \int_{\mathcal{S}_\alpha} \bar{F}_X^{\alpha}(x)\bigg(\frac{\bar{F}_Y^{1-\alpha}(x)-1}{\alpha-1}\bigg)\,\mathrm{d}x$$
$$= -h_\alpha(\bar{F}_X) + \int_{\mathcal{S}_\alpha} \bar{F}_X^{\alpha}(x)\bigg(\frac{\bar{F}_Y^{1-\alpha}(x)-1}{\alpha-1}\bigg)\,\mathrm{d}x.$$

That is,

$$h_\alpha(\bar{F}_X) = -\int_{\mathcal{S}_\alpha} \bar{F}_X^{\alpha}(x)\bigg(\frac{1-\bar{F}_Y(x)^{1-\alpha}}{\alpha-1}\bigg)\,\mathrm{d}x - \int_{\mathcal{S}_\alpha} \bar{F}_X(x)\overline{L}_\alpha\bigg(\frac{\bar{F}_X(x)}{\bar{F}_Y(x)}\bigg)\,\mathrm{d}x.$$

Using Lemma 2.1(i) with $\phi_1(x) = \bar{F}_X(x)$ and $\phi_2(x) = \bar{F}_Y(x)$,

$$h_\alpha(\bar{F}_X) \leq -\int_{\mathcal{S}_\alpha} \bar{F}_X^{\alpha}(x)\bigg(\frac{1-\bar{F}_Y(x)^{1-\alpha}}{\alpha-1}\bigg)\,\mathrm{d}x - \mu_x\overline{L}_\alpha\bigg(\frac{\mu_x}{\mu_y}\bigg). \tag{A.2}$$

With $\bar{F}_Y$ as in (3.13), $\mu_y = c/(2-\alpha)$. Then (A.2) implies that

$$-h_\alpha(\bar{F}) \geq -\frac{1}{c}\int_{\mathcal{S}_\alpha} x\bar{F}_X^{\alpha}(x)\,\mathrm{d}x + \mu_x\overline{L}_\alpha\bigg(\frac{(2-\alpha)\mu_x}{c}\bigg).$$

The integral yields $\mathbb{E}(X_\alpha^2)/2$. Thus, for all values of $c > 0$,

$$-h_\alpha(\bar{F}) \geq -\frac{1}{2c}\mathbb{E}(X_\alpha^2) + \mu_x\overline{L}_\alpha\bigg(\frac{(2-\alpha)\mu_x}{c}\bigg) = \varphi(c), \qquad c > 0. \tag{A.3}$$

That is,

$$-h_\alpha(\bar{F}) \geq \max_{c>0}\phi(c) = \varphi(c^*).$$

Setting $\varphi'(c) = 0$ and noting that $\varphi''(c) = [(2-\alpha)/c]^{\alpha+1} > 0$, we find that $c^* = (2-\alpha)\nu_\alpha Q_\alpha(X)$, where

$$Q_\alpha(X) = \bigg(\frac{\mathbb{E}(X_\alpha^2)}{\mathbb{E}^\alpha(X)}\bigg)^{1/(2-\alpha)}, \qquad 0 < \alpha < 2.$$

Substituting the value of c^* in (A.3) and then applying Lemma 2.1(ii), we obtain

$$-h_\alpha(\bar{F}) \geq -\frac{\mathbb{E}(X_\alpha^2)}{2(2-\alpha)\nu_\alpha Q_\alpha(X)} + \mu_x\overline{L}_\alpha\bigg(\frac{\mu_x}{\nu_\alpha Q_\alpha(X)}\bigg)$$
$$\geq -\omega_\alpha S_\alpha(X) + \mu_x\bigg(1 - \frac{\nu_\alpha Q_\alpha(X)}{\mu_x}\bigg)$$
$$= -\omega_\alpha S_\alpha(X) - \nu_\alpha Q_\alpha(X) + \mu_x,$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

where

$$S_\alpha(X) = \frac{\mathbb{E}(X_\alpha^2)}{Q_\alpha(X)} = [\mathbb{E}^{1-\alpha}(X_\alpha^2)\mathbb{E}^\alpha(X)]^{1/(2-\alpha)}, \qquad 0 < \alpha < 2.$$

Hence, we obtain (3.15). □

Proof of Theorem 4.1. (i) Applying the stochastic assumption to the integral in (A.2), we obtain

$$\begin{aligned} h_\alpha(\bar{F}_X) &\le -\int_0^\infty \bar{F}_Y^\alpha(x)\bigg(\frac{1-\bar{F}_Y(x)^{1-\alpha}}{\alpha-1}\bigg)\,\mathrm{d}x - \mu_x\overline{L}_\alpha\bigg(\frac{\mu_x}{\mu_y}\bigg) \\ &= -\int_0^\infty \bar{F}_Y^\alpha(x)\bigg(\frac{\bar{F}_Y(x)^{1-\alpha}-1}{1-\alpha}\bigg)\,\mathrm{d}x - \mu_x\overline{L}_\alpha\bigg(\frac{\mu_x}{\mu_y}\bigg) \\ &= h_\alpha(\bar{F}_Y) - \mu_x\overline{L}_\alpha\bigg(\frac{\mu_x}{\mu_y}\bigg). \end{aligned}$$

(ii) From (4.2), we have $h_2(\bar{F}_X) = \mu_x G(X)$ and $h_2(\bar{F}_Y) = \mu_y G(Y)$. Substituting into (4.4) and noting that $\overline{L}_2(\rho) = \rho - 1$ yields the result. The restriction $\rho \le 1$ is implied by the stochastic order assumption. □

Proof of Theorem 5.1. This result can be obtained as a corollary of [8, Theorem 9]. In order to avoid additional definitions and notation we state the following proof. The score function of p_α is as follows:

$$\begin{aligned} \frac{\partial \log p_\alpha(X)}{\partial\alpha} &= \frac{\partial(\log \bar{F}_\alpha(x) - \log\int_0^\infty \bar{F}_\alpha(x)\,\mathrm{d}x)}{\partial\alpha} \\ &= \log\bar{F}(x) - \frac{\int_0^\infty \bar{F}^\alpha(x)\log\bar{F}(x)\,\mathrm{d}x}{\int_0^\infty \bar{F}^\alpha(x)\,\mathrm{d}x} \\ &= \log\bar{F}(x) - \int_0^\infty \log\bar{F}(x)p_\alpha(x)\,\mathrm{d}x \\ &= \log\bar{F}(x) - \mathbb{E}_{p_\alpha}[\log\bar{F}(X)]. \end{aligned}$$

The expectation of the square of the partial derivative yields the result. □

Proof of Theorem 5.2. First note that

$$\begin{aligned} I_{p_\alpha}(\alpha) &= \int_0^\infty \bigg(\frac{\partial\log p_\alpha(x)}{\partial\alpha}\bigg)^2 p_\alpha(x)\,\mathrm{d}x \\ &= \int_0^\infty \bigg(\frac{\partial p_\alpha(x)}{\partial\alpha}\bigg)(\log\bar{F}(x) - \mathbb{E}_{p_\alpha}[\log\bar{F}(X)])\,\mathrm{d}x \\ &= \int_0^\infty \bigg(\frac{\partial p_\alpha(x)}{\partial\alpha}\bigg)\log\bar{F}(x)\,\mathrm{d}x - \mathbb{E}_{p_\alpha}[\log\bar{F}(X)]\int_0^\infty \frac{\partial p_\alpha(x)}{\partial\alpha}\,\mathrm{d}x \\ &= \int_0^\infty \bigg(\frac{\partial p_\alpha(x)}{\partial\alpha}\bigg)\log\bar{F}(x)\,\mathrm{d}x, \end{aligned}$$

where the last equality follows from the regularity conditions, implying that

$$\int_0^\infty \frac{\partial p_\alpha(x)}{\partial\alpha}\,\mathrm{d}x = 0.$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

Thus,

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \int_1^\infty \biggl[\int_0^\infty \biggl(\frac{\partial p_\alpha(x)}{\partial \alpha}\biggr) \log \bar{F}(x)\,\mathrm{d}x\biggr]\pi(\alpha)\,\mathrm{d}\alpha$$
$$= \int_0^\infty \biggl[\int_1^\infty \biggl(\frac{\partial p_\alpha(x)}{\partial \alpha}\biggr)\pi(\alpha)\,\mathrm{d}\alpha\biggr] \log \bar{F}(x)\,\mathrm{d}x. \qquad \text{(A.4)}$$

(i) In (A.4), let $u = \pi(\alpha) = \theta\alpha^{-(\theta+1)}$, $\alpha \geq 1$, and $\mathrm{d}v = (\partial p_\alpha(x)/\partial\alpha)\,\mathrm{d}\alpha$. Then $\mathrm{d}u = -\theta\pi_{\theta+1}(\alpha)\,\mathrm{d}\theta$ and $v = p_\alpha(x)$. Integration by parts yields

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \int_0^\infty \biggl(\int_1^\infty \biggl(\frac{\partial p_\alpha(x)}{\partial \alpha}\biggr)\pi(\alpha)\,\mathrm{d}\alpha\biggr) \log \bar{F}(x)\,\mathrm{d}x$$
$$= \int_0^\infty [p_\alpha(x)\theta\alpha^{-(\theta+1)}]_1^\infty \log \bar{F}(x)\,\mathrm{d}x$$
$$+ \int_0^\infty \biggl[\int_1^\infty p_\alpha(x)\theta\pi_{\theta+1}(\alpha)\,\mathrm{d}\alpha\biggr] \log \bar{F}(x)\,\mathrm{d}x$$
$$= -\theta \int_0^\infty p_1(x)\log \bar{F}(x)\,\mathrm{d}x + \theta\int_0^\infty p^*(x)\log\bar{F}(x)\,\mathrm{d}x$$
$$= \theta\{\mathbb{E}_{p^*}[\log\bar{F}(X)] - \mathbb{E}_{p_1}[\log\bar{F}(X)]\}.$$

(ii) In (A.4), let $\pi(\alpha) = 1/(a-1)$, $1 \leq \alpha \leq a$. Then

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] = \frac{1}{a-1}\int_0^\infty \biggl[\int_1^a \biggl(\frac{\partial p_\alpha(x)}{\partial \alpha}\biggr)\mathrm{d}\alpha\biggr] \log \bar{F}(x)\,\mathrm{d}x$$
$$= \frac{1}{a-1}\int_0^\infty [p_a(x) - p_1(x)]\log\bar{F}(x)\,\mathrm{d}x$$
$$= \frac{1}{a-1}\{\mathbb{E}_{p_a}[\log\bar{F}(X)] - \mathbb{E}_{p_1}[\log\bar{F}(X)]\}.$$

(iii) In (A.4), let $\pi(\alpha) \propto 1$, $\alpha \geq 1$. Then

$$\mathbb{E}_\pi[I_{p_\alpha}(\alpha)] \propto \int_0^\infty \biggl[\int_1^\infty \biggl(\frac{\partial p_\alpha(x)}{\partial \alpha}\biggr)\mathrm{d}\alpha\biggr] \log \bar{F}(x)\,\mathrm{d}x$$
$$\propto \int_0^\infty [p_\infty(x) - p_1(x)]\log\bar{F}(x)\,\mathrm{d}x$$
$$\propto -\mathbb{E}_{p_1}[\log\bar{F}(X)],$$

where the last expression is obtained from the assumption that $\lim_{\alpha\to\infty} p_\alpha(x) \to 0$ for all x.

The result follows by noting that $p_1(x) = p(x)$ given in (1.3). □

Proof of Theorem 6.1. We have

$$\frac{\partial}{\partial\alpha}\lambda_{p_\alpha}(\tau) = \frac{\bar{F}^\alpha(\tau)\log\bar{F}(\tau)\int_\tau^\infty \bar{F}^\alpha(u)\,\mathrm{d}u - \bar{F}^\alpha(\tau)\int_\tau^\infty \log\bar{F}(\tau)\bar{F}^\alpha(u)\,\mathrm{d}u}{(\int_\tau^\infty \bar{F}^\alpha(u)\,\mathrm{d}u)^2}$$
$$= \lambda_{p_\alpha}(\tau)\biggl(\log\bar{F}(\tau) - \frac{\int_\tau^\infty \log\bar{F}(\tau)\bar{F}^\alpha(u)\,\mathrm{d}u}{\int_\tau^\infty \bar{F}^\alpha(u)\,\mathrm{d}u}\biggr)$$

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press

$$= \frac{\lambda_{p_\alpha}(\tau)}{\alpha}\left(\log p_\alpha(\tau) - \frac{\int_\tau^\infty \log p_\alpha(\tau) p_\alpha(u)\,\mathrm{d}u}{\int_\tau^\infty p_\alpha(u)\,\mathrm{d}u}\right)$$
$$= \frac{\lambda_{p_\alpha}(\tau)}{\alpha}\left(\log p_\alpha(\tau) + H(p_\alpha;\ \tau) - \log \int_\tau^\infty p_\alpha(u)\,\mathrm{d}u\right)$$
$$= \frac{\lambda_{p_\alpha}(\tau)}{\alpha}(\log \lambda_{p_\alpha}(\tau) + H(p_\alpha;\ \tau)).$$

That is,

$$H(p_\alpha;\ \tau) = \alpha \frac{\partial}{\partial \alpha} \log \lambda_{p_\alpha}(\tau) - \log \lambda_{p_\alpha}(\tau),$$

which is equivalent to (6.1). Representation (6.2) is obtained using the following reciprocal relationship between the hazard function of the ED of the PH model and $\lambda_{p_\alpha}(\tau) = 1/m_\alpha(\tau)$. This completes the proof. □

Proof of Theorem 6.2. Using the PDF of the PH $f_\alpha(x) = \alpha f(x)[\bar{F}(x)]^{\alpha-1}$, we obtain

$$H(f_\alpha) = -\int_0^\infty \alpha f(x)[\bar{F}(x)]^{\alpha-1} \log(\alpha f(x)[\bar{F}(x)]^{\alpha-1})\,\mathrm{d}x$$
$$= -\alpha \int_0^\infty f(x)[\bar{F}(x)]^{\alpha-1} \log f(x)\,\mathrm{d}x - K(f_\alpha : f)$$
$$= -K(f_\alpha : f) - \alpha \mathbb{E}_f([\bar{F}(X)]^{\alpha-1} \log f(X))$$
$$= -K(f_\alpha : f) - \alpha \operatorname{cov}(\log f(X), [\bar{F}(X)]^{\alpha-1}) - \alpha \mathbb{E}_f[\log f(X)] \mathbb{E}_f([F(X)]^{\alpha-1})$$
$$= -K(f_\alpha : f) - \alpha \operatorname{cov}(\log f(X), [\bar{F}(X)]^{\alpha-1}) + \alpha H(f) \int_0^1 u^{\alpha-1}\,\mathrm{d}u$$
$$= -K(f_\alpha : f) - \alpha \operatorname{cov}(\log f(X), [\bar{F}(X)]^{\alpha-1}) + H(f).$$

This yields (6.3). Letting $u = \bar{F}(x)$, we obtain

$$K(f_\alpha : f) = \int_0^1 \alpha u^{\alpha-1} \log(\alpha u^{\alpha-1})\,\mathrm{d}u = \log \alpha + \frac{1}{\alpha} - 1.$$

This yields (6.4) and (6.5). □

Acknowledgements

The authors thank the anonymous reviewers for their constructive comments which led to an improvement in the exposition of this paper. Asadi's research was carried out in IPM Isfahan branch and was in part supported by a grant from IPM (grant number 95620411). Ebrahimi's research was partially supported by a grant from the NSF (grant number DMS 1208273). Soofi's research was supported by a Roger L. Fitzsimonds Distinguished Scholar Award.

References

[1] Abe, S. (2000). Axioms and uniqueness theorem for Tsallis entropy. *Phys. Lett. A* **271,** 74–79.

[2] Ardakani, O. M., Ebrahimi, N. and Soofi, E. S. (2018). Ranking forecasts by stochastic error distance, information, and reliability measures. To appear in *Internat. Statist. Rev.*

[3] Ardakani, O. M., Asadi, M., Ebrahimi N. and Soofi, E. S. (2017). On the Bayes risk of mean residual with system reliability and empirical applications. Submitted.

[4] Asadi, M. and Zohrevand, Y. (2007). On the dynamic cumulative residual entropy. *J. Statist. Planning Infer.* **137,** 1931–1941.
[5] Asadi, M., Ebrahimi, N., Hamedani, G. G. and Soofi, E. S. (2004). Maximum dynamic entropy models. *J. Appl. Prob.* **41,** 379–390.
[6] Asadi, M., Ebrahimi, N., Soofi, E. S. and Zarezadeh, S. (2014). New maximum entropy methods for modeling lifetime distributions. *Naval Res. Logistics* **61,** 427–434.
[7] Bairamov, I. and Arnold, B. C. (2008). On the residual lifelengths of the remaining components in an $n-k+1$ out of n system. *Statist. Prob. Lett.* **78,** 945–952.
[8] Bercher, J.-F. (2012). A simple probabilistic construction yielding generalized entropies and divergences, escort distributions and q-Gaussians. *Physica A* **391,** 4460–4469.
[9] Borland, L., Plastino, A. R. and Tsallis, C. (1998). Information gain within nonextensive thermostatistics. *J. Math. Phys.* **39,** 6490–6501. (Erratum: **40** (1999), 2196.)
[10] Dabrowska, D. M. and Doksum, K. A. (1988). Estimation and testing in a two-sample generalized odds-rate model. *J. Amer. Statist. Assoc.* **83,** 744–749.
[11] Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions. *J. Appl. Prob.* **39,** 434–440.
[12] Di Crescenzo, A. and Toomaj, A. (2015). Extension of the past lifetime and its connection to the cumulative entropy. *J. Appl. Prob.* **52,** 1156–1174.
[13] Ebrahimi, N., Soofi, E. S. and Soyer, R. (2013). When are observed failures more informative than observed survivals? *Naval Res. Logistics* **60,** 102–110.
[14] Havrda, J. and Charvát, F. (1967). Quantification method of classification processes: Concept of structural a-entropy. *Kybernetika* **3,** 30–35.
[15] Lomnicki, Z. A. (1966). A Note on the Weibull renewal process. *Biometrika* **53,** 375–381.
[16] Maasoumi, E. (1986). The measurement and decomposition of multi-dimensional inequality. *Econometrica* **54,** 991–998.
[17] Navarro, J., del Aguila, Y. and Asadi, M. (2010). Some new results on the cumulative residual entropy. *J. Statist. Planning Infer.* **140,** 310–322.
[18] Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. Statist.* **8,** 154–168.
[19] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F. (2004). Cumulative residual entropy: a new measure of information. *IEEE Trans. Inf. Theory* **50,** 1220–1228.
[20] Ross, S. M. (1983). *Stochastic Processes*. John Wiley, New York.
[21] Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer, New York.
[22] Shalit, H. and Yitzhaki, S. (2005). The Mean-Gini efficient portfolio frontier. *J. Financial Res.* **XXVIII,** 59–75.
[23] Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *J. Statist. Phys.* **52,** 479–487.
[24] Vakili-Nezhaad, G. R. and Mansoori, G. A. (2004). An application of non-extensive statistical mechanics to nanosystems. *J. Comput. Theort. Nanosci.* **1,** 227–229.
[25] Yannaros, N. (1994). Weibull renewal processes. *Ann. Inst. Statist. Math.* **46,** 641–648.
[26] Yannaros, N. (1988). On Cox processes and gamma renewal processes. *J. Appl. Prob.* **25,** 423–427.
[27] Yitzhaki, S. (1998). More than a dozen alternative ways of spelling Gini. *Res. Econom. Ineq.* **8,** 13–30.
[28] Zografos, K. and Nadarajah, S. (2005). Survival exponential entropies. *IEEE Trans. Inf. Theory* **51,** 1239–1246.

https://doi.org/10.1017/jpr.2017.51 Published online by Cambridge University Press