The Australian Journal of Mathematical Analysis and Applications

AJMAA

Volume 13, Issue 1, Article 1, pp. 1-29, 2016

SOME NEW GENERALIZATIONS OF JENSEN'S INEQUALITY WITH RELATED RESULTS AND APPLICATIONS

STEVEN G. FROM

Received 13 January, 2016; accepted 20 April, 2016; published 17 May, 2016.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NEBRASKA AT OMAHA, OMAHA, NEBRASKA 68182-0243.
sfrom@unomaha.edu

ABSTRACT. In this paper, some new generalizations of Jensen's inequality are presented. In particular, upper and lower bounds for the Jensen gap are given and compared analytically and numerically to previously published bounds for both the discrete and continuous Jensen's inequality cases. The new bounds compare favorably to previously proposed bounds. A new method based on a series of locally linear interpolations is given and is the basis for most of the bounds given in this paper. The wide applicability of this method will be demonstrated. As by-products of this method, we shall obtain some new Hermite-Hadamard inequalities for functions which are 3-convex or 3-concave. The new method works to obtain bounds for the Jensen gap for non-convex functions as well, provided one or two derivatives of the nonlinear function are continuous. The mean residual life function of applied probability and reliability theory plays a prominent role in construction of bounds for the Jensen gap. We also present an exact integral representation for the Jensen gap in the continuous case. We briefly discuss some inequalities for other types of convexity, such as convexity in the geometric mean, and briefly discuss applications to reliability theory.

Key words and phrases: b-entropy, convex according to the geometric mean, convex function, decreasing mean residual life (DMRL), Hermite-Hadamard inequality, Jensen gap, Jensen's inequality, linear interpolation, moments, mean residual life function (MRL), moment-generating function, new better than used in expectation (NBUE), real analytic, Riemann-Stieltjes Sums.

2000 *Mathematics Subject Classification.* Primary 26, 28, 60. Secondary 26A, 26D..

ISSN (electronic): 1449-5910

1. INTRODUCTION

It is well-known that the discrete Jensen's inequality states that if f is a convex function on $[a, b]$, $p_i > 0$, $i = 1, \ldots, n$, $\sum_{j=1}^{n} p_j = 1$, and $x_i \in [a, b]$, $i = 1, \ldots, n$, then

$$D = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \geq 0\,. \tag{1.1}$$

The most general form of Jensen's inequality states that if $(\Omega, \mathcal{F}, \mu)$ is a measure space with $\mu(\Omega) = 1$, $f(x)$ is convex, and $g(x)$ is a μ-intergrable real-valued function, then

$$D = \int_{\Omega} (f \circ g) d\mu - f\left(\int_{\Omega} g d\mu\right) \geq 0\,. \tag{1.2}$$

In this paper, $g(x) = x$, a very important special case. We shall mainly be concerned with the discrete case and pass to the limit to the possibly continuous case.

Many papers have been written on Jensen's inequality and its many companion inequalities. See [9] and [24], for example. In this paper, we shall obtain new inequalities for the Jensen gap, D, in (1.1) and (1.2). These will allow us to obtain new companion inequalities, including new Hermite-Hadamard type inequalities for 3-convex functions.

We are concerned with obtaining bounds for the Jensen gap D in (1.1). In continuous cases where there exists a probability density function absolutely continuous with respect to Lebesgue measure $h(x)$, then (1.2) becomes, in the continuous case:

$$D = \int_{-\infty}^{\infty} f(x)h(x)dx - f\left(\int_{-\infty}^{\infty} x \cdot h(x)dx\right)\,. \tag{1.3}$$

We will obtain upper and lower bounds for D in (1.1) and (1.3) even if $f(x)$ is not convex/concave, provided f has one or two continuous derivatives. We focus first on the discrete case and present new bounds and compare them analytically and numerically to previously published bounds. Later, we obtain, as a limiting case, bounds for D in the continuous case. We shall obtain both an exact integral representation and an exact infinite series representation for the Jensen gap D in (1.3) for the special case where $h(x) = 0$, if $x \notin [a, b]$, that is, when the support of $h(x)$ is $[a, b]$, where $a < b$ are real numbers. Then it is a simple matter to extend to the cases where $b \to \infty$ or $a \to -\infty$. In applied probability and reliability theory areas, the function $g(x)$ below is of great importance:

$$\begin{aligned} g(x) &= \frac{\int_x^b (t-x)h(t)dt}{\int_x^b h(t)dt}, \quad a \leq x < b \\ g(b) &= 0\,. \end{aligned} \tag{1.4}$$

The function $g(x)$ is known as the mean residual life function. We shall see that the Jensen gap bounds are strongly related to $g(x)$ as is the exact single integral representation of D to be given later in Theorem 4.3. A new method is given to obtain the inequalities in this paper. This method is widely applicable to obtain new Jensen type inequalities.

First, let's discuss the finite n discrete case. The following bounds on the Jensen gap D have previously been published. First, let's discuss some 'global' bounds for D.

Several global bounds for the Jensen gap D have been published. Two of the most prominent are given below.

Theorem A. *(Dragomir* [12]*). Suppose* $p_i > 0$, $\sum_{j=1}^n p_j = 1$. *If* f *is a differentiable convex mapping on* $[a,b]$*, then*

$$\begin{aligned} D &= \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) \\ &\le \frac{1}{4}(b-a)(f'(b)-f'(a)) \equiv D_f(a,b)\,. \end{aligned} \tag{1.5}$$

The next theorem was given in [27].

Theorem B. Under the conditions of Theorem A above,

$$\begin{aligned} D &= \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) \\ &\le f(a)+f(b)-2f\left(\frac{a+b}{2}\right) \equiv S_f(a,b)\,. \end{aligned} \tag{1.6}$$

Theorem C. *(Theorem 1 of* [15]*.)* Let $f : [a,b] \to \mathbb{R}$ be a differentiable convex function on (a,b). Let $x_i \in (a,b)$, $i = 1,2,\ldots,n$. Let $p_i \ge 0$, $i = 1,2,\ldots,n$ and $\sum_{i=1}^n p_i = 1$. Let $\bar{x} = \sum_{i=1}^n p_i x_i$. Then

$$\begin{aligned} D &= \sum_{i=1}^n p_i f(x_i) - f(\bar{x}) \\ &\ge \left|\sum_{i=1}^n p_i |f(x_i)-f(\bar{x})| - \sum_{i=1}^n p_i |x_i-\bar{x}|\cdot|f'(\bar{x})|\right| \equiv L_{DS}\,. \end{aligned} \tag{1.7}$$

Theorem D. . *(Theorems 1 and 2 of* [14]*.)* Let $f : [a,b] \to \mathbb{R}$ be a differentiable convex function on (a,b), $x_i \in (a,b)$, $p_i \ge 0$, $i = 1,2,\ldots,n$, $\sum_{i=1}^n p_i = 1$. Let D and $\bar{x}$ be as given in Theorem C above. Then

$$\text{(a)} \qquad D \le \sum_{i=1}^n p_i x_i f'(x_i) - \bar{x}\cdot\sum_{i=1}^n p_i f'(x_i) \equiv U_{D,1} \tag{1.8}$$

and

$$\begin{aligned} \text{(b)} \qquad D &\le \sum_{i=1}^n p_i x_i f'(x_i) + f\left((f')^{-1}\left(\sum_{i=1}^n p_i f'(x_i)\right)\right) \\ &\quad -(f')^{-1}\left(\sum_{i=1}^n p_i f'(x_i)\right)\cdot\sum_{n=1}^n p_i f'(x_i) - f(\bar{x}) \equiv U_{D,2} \\ &\le U_{D,1}\,, \end{aligned} \tag{1.9}$$

where $(f')^{-1}$ is the inverse function of the derivative f'.

Next, let's present some of the best 'local' bounds for Jensen gap D. We shall compare these to the new bounds to be discussed in later sections.

Theorem E. *(Theorem 1 of* [2]*).* Let $f : I \to \mathbb{R}$ where I is an interval. Let $x_i \in I$, $i = 1, 2, \ldots, n$, $p_i \geq 0$, $i = 1, 2, \ldots, n$ with $\sum_{i=1}^{n} p_i = 1$. Let $M = \max\{x_1, x_2, \ldots, x_n\}$, $m = \min\{x_1, x_2, \ldots, x_n\}$. If f is differentiable and f' is strictly increasing, then

$$D = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \leq \lambda\,, \tag{1.10}$$

where

$$\begin{aligned} \lambda &= \left(\frac{f(M)-f(m)}{M-m}\right)(f')^{-1}\left(\frac{f(M)-f(m)}{M-m}\right) + \frac{Mf(m)-mf(M)}{M-m} \\ &\quad - f\left((f')^{-1}\left(\frac{f(M)-f(m)}{M-m}\right)\right) \equiv U_{BPP}(1) \end{aligned} \tag{1.11}$$

where $(f')^{-1}$ is the inverse function of f'.

Theorem F. *(Theorem 11 of* [2]*).* Let $f : [m, M] \to \mathbb{R}$ be a convex function on $[m, M]$, $p_i > 0$, $i = 1, 2, \ldots, n$, $\sum_{i=1}^{n} p_i = 1$, $n \geq 2$ and $x_i \in [m, M]$, $i = 1, \ldots, n$, with $m \leq x_1 < x_2 < \cdots < x_n \leq M$. Then

(1.12)

$$\begin{aligned} D &= \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &\leq \max_{1\leq k\leq n} \{p_k^* f(m) + (1-p_k^*)f(M) - f(p_k^* m + (1-p_k^*)M)\} \equiv U_{BPP}(2) \end{aligned}$$

where

$$p_k^* = \sum_{i=1}^{k} p_i, \quad k = 1, 2, \ldots, n\,.$$

Theorem G. *(Theorem 1.3 of* [17]*).* Let $f[a, b] \to \mathbb{R}$ be continuous, twice differentiable on (a, b). Let $x_i \in [a, b]$, $p_i \geq 0$, $i = 1, 2, \ldots, n$ with $\sum_{i=1}^{n} p_i = 1$. Suppose $m = \inf\{f''(x) : x \in [a, b]\}$ and $M = \sup\{f''(x) : x \in [a, b]\}$ exist. Then

$$D = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \geq \frac{m}{2}\left(\sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2\right) \equiv F_L \tag{1.13}$$

and

$$D \leq \frac{M}{2}\left(\sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2\right) \equiv F_U\,. \tag{1.14}$$

The following theorem is a special case of a more general result given in [10].

Theorem H. Let $f : [a, b] \to \mathbb{R}$ be a differentiable convex function. Let $x_i \in [a, b]$, $p_i \geq 0$, $i = 1, \ldots, n$, $\sum_{j=1}^{n} p_j = 1$. Then

(1.15)

$$\begin{aligned} D &= \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &\leq \left(\sum_{i=1}^{n} p_i (f'(x_i))^2\right)^{1/2} \cdot \left(\sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2\right)^{1/2} \equiv U_{DG} \,. \end{aligned}$$

2. NEW BOUNDS - DISCRETE CASE

To present new bounds for D, we shall first need the following lemmas.

Lemma 2.1. *Let $f(t)$ be a real-valued function of t on $[c, d]$, such that $f(t)$ and its first two derivatives $f'(t)$ and $f''(t)$ are all continuous throughout the interval $c \leq t \leq d$. Let $L(t)$ denote the interpolating polynomial of degree 1 or less passing through the points $(c, f(c))$ and $(d, f(d))$. Then for each t_0 in $[c, d]$, there is a number $\xi(t_0)$ in (c, d) such that*

$$f(t_0) = L(t_0) + \frac{f''(\xi(t_0))}{2}(t_0 - c)(t_0 - d)\,, \tag{2.1}$$

where

$$L(t_0) = f(c) + \left(\frac{f(d) - f(c)}{d - c}\right) \cdot (t_0 - c), \quad c < d\,. \tag{2.2}$$

Thus, the approximation $f(t_0) \approx L(t_0)$ has linear interpolation error

$$E(t_0) \equiv f(t_0) - L(t_0) = \frac{f''(\xi(t_0))}{2}(t_0 - c)(t_0 - d)\,. \tag{2.3}$$

Proof. See [5], p. 111-113. ∎

Lemma 2.2. *Let $f(t)$ be a real-valued function of t on $[c, d]$. Suppose $f(t)$ and its first three derivatives $f'(t)$, $f''(t)$ and $f^{(3)}(t)$ are continuous throughout $[c, d]$.*

a) *If $f''(t) \geq 0$ and $f^{(3)}(t) \geq 0$, $c \leq t \leq d$, then for each t_0 in $[c, d]$, the approximation error $E(t_0)$ in the approximation $f(t_0) \approx L(t_0)$, where $L(t_0)$ is given by (2.2), satisfies*

$$\begin{aligned} [(d-c)f'(d) - f(d) + f(c)] \cdot \frac{(t_0 - c)(t_0 - d)}{(d-c)^2} &\leq E(t_0) \\ &\leq [f(d) - f(c) - (d-c)f'(c)] \cdot \frac{(t_0 - c)(t_0 - d)}{(d-c)^2}\,, \end{aligned} \tag{2.4}$$

where $E(t_0) = f(t_0) - L(t_0)$.

b) *If $f''(t) \geq 0$ and $f^{(3)}(t) \leq 0$, then (2.4) holds with the inequality reversed, that is, for $c < d$,*

$$\begin{aligned} [f(d) - f(c) - (d-c)f'(c)]\frac{(t_0 - c)(t_0 - d)}{(d-c)^2} &\leq E(t_0) \\ &\leq [(d-c)f'(d) - f(d) + f(c)] \cdot \frac{(t_0 - c)(t_0 - d)}{(d-c)^2}\,. \end{aligned} \tag{2.5}$$

Proof. See [19], p 364–366. ∎

Now, we are ready to present some main results on bounds for the Jensen gap D in the discrete case.

Theorem 2.3. *Let $f(t)$ be a real-valued function on $[a,b]$. Suppose that f and its first two derivatives f' and f'' are continuous on $[a,b]$. Suppose that $a \le x_1 \le x_2 \le \cdots \le x_n \le b$. Suppose $p_i \ge 0$, $i = 1,2,\ldots,n$, with $\sum_{i=1}^n p_i = 1$, $p_1 > 0$, $p_n > 0$. Let*

$$R_i = p_i + p_{i+1} + \cdots + p_n = \sum_{j=i}^{n} p_j, \quad i = 1,2,\ldots,n\,, \tag{2.6}$$

$$x_i^* = \frac{p_{i+1}}{R_{i+1}} x_{i+1} + \cdots + \frac{p_n}{R_{i+1}} x_n = \frac{\sum_{L=i+1}^{n} p_L x_L}{R_{i+1}}\,, \qquad i = 1,2,\ldots,n-1\,, \tag{2.7}$$

and

$$\begin{aligned} x^{(i)} &= \left(\frac{p_i}{R_i}\right) x_i + \left(\frac{R_{i+1}}{R_i}\right) x_i^* \\ &= \left(\frac{p_i}{R_i}\right) x_i + \left(1 - \frac{p_i}{R_i}\right) x_i^*\,, \qquad i = 1,2,\ldots,n\,. \end{aligned} \tag{2.8}$$

Then there exist real numbers $\theta_1, \theta_2, \ldots, \theta_{n-1}$ with $x_i < \theta_i < x_i^*$, $i = 1,2,\ldots,n-1$ such that

$$\begin{aligned} D &= \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \\ &= \sum_{i=1}^{n-1} \frac{f''(\theta_i)}{2} (x^{(i)} - x_i)(x_i^* - x^{(i)}) \\ &= \sum_{i=1}^{n-1} \frac{f''(\theta_i)}{2} p_i \left(1 - \frac{p_i}{R_i}\right) (x_i^* - x_i)^2\,. \end{aligned} \tag{2.9}$$

Proof. Define A_i by

$$\begin{aligned} A_i &= \sum_{L=1}^{i-1} p_L f(x_L) + R_i f(x^{(i)}) \\ &= \sum_{L=1}^{i-1} p_L f(x_L) + R_i f\left(\left(\frac{p_i}{R_i}\right) x_i + \left(1 - \frac{p_i}{R_i}\right) x_i^*\right),\,,\ i = 1,2,\ldots,n\,, \end{aligned} \tag{2.10}$$

where any impossible sum is defined to be zero. Then

$$f\left(\sum_{i=1}^{n} p_i x_i\right) - \sum_{i=1}^{n} p_i f(x_i) = A_1 - A_n = \sum_{j=1}^{n-1} (A_j - A_{j+1})\,. \tag{2.11}$$

Now simple algebra gives $x^{(j+1)} = x_j^*$ and

$$A_j - A_{j+1} = R_j \left(f(x^{(j)}) - \left[\left(\frac{p_j}{R_j}\right) f(x_j) + \left(1 - \frac{p_j}{R_j}\right) f(x_j^*) \right] \right)\ j = 1,2,\ldots,n-1\,. \tag{2.12}$$

If $x_j = x_j^*$, $A_j - A_{j+1} = 0$. If $x_j < x_j^*$, then the expression in brackets of (2.12) is the linearly interpolated value of $f(x^{(j)})$ on the interval $x_j \le t \le x_j^*$. Letting $c = x_j$, $d = x_j^*$, $t_0 = x^{(j)}$ in

Lemma 2.1, we obtain this value as

$$L(x^{(j)}) = \frac{p_j}{R_j} f(x_j) + \left(1 - \frac{p_j}{R_j}\right) f(x_j^*) .$$

Thus there exists a real number θ_j in (x_j, x_j^*) such that

$$f(x^{(j)}) - L(x^{(j)}) = \frac{f''(\theta_j)}{2}(x^{(j)} - x_j)(x^{(j)} - x_j^*) . \tag{2.13}$$

Thus, (2.12) gives

$$A_j - A_{j+1} = R_j \left(\frac{f''(\theta_j)}{2}\right) \cdot (x^{(j)} - x_j)(x^{(j)} - x_j^*), \quad 1 \le j \le n-1 ,$$

which is true also when $x_j = x_j^*$, since $x^{(j)} = x_j = x_j^*$ in this case. Summing over j gives, using (2.13):

$$f\left(\sum_{i=1}^n p_i x_i\right) - \sum_{i=1}^n p_i f(x_i) = \sum_{j=1}^{n-1} \frac{f''(\theta_j)}{2}(x^{(j)} - x_j)(x^{(j)} - x_j^*) \cdot R_j . \tag{2.14}$$

Also, simple algebra gives

$$x^{(i)} - x_i = \left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)$$

and

$$x_i^* - x^{(i)} = \frac{p_i}{R_i}(x_i^* - x_i), \quad i = 1, 2, \ldots, n-1 ,$$

from which we obtain

$$(x^{(i)} - x_i)(x_i^* - x^{(i)}) = \frac{p_i}{R_i}\left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)^2 , \quad i = 1, 2, \ldots, n-1 . \tag{2.15}$$

Thus, (2.14) gives

$$\begin{aligned} D &= \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) \\ &= \sum_{i=1}^{n-1} \frac{f''(\theta_i)}{2} p_i \left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)^2 \end{aligned} \tag{2.16}$$

and Theorem 2.3 is proved. ∎

Remark 2.1. Note that $x_i \le x^{(i)} \le x_i^*$ and $(x^{(i)} - x_i)(x_i^* - x^{(i)}) \ge 0$, $i = 1, 2, \ldots, n-1$. Thus, if $f(t)$ is convex on $[a, b]$, then (2.16) above is nonnegative as required by Jensen's inequality.

From Theorem 2.3, we obtain the following corollary. The proof is omitted since it follows immediately from Theorem 2.3.

Corollary 2.4. *Under the assumptions of Theorem 2.3, we have*

$$\begin{aligned} L_J &\equiv \sum_{i=1}^{n-1} \frac{1}{2} m_i p_i \left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)^2 \le D \\ &\le \sum_{i=1}^{n-1} \frac{1}{2} M_i p_i \left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)^2 \equiv U_J \end{aligned} \tag{2.17}$$

where

$$\begin{aligned} m_i &= \inf\{f''(t) : x_i \le t \le x_i^*\} \quad \textit{and} \\ M_i &= \sup\{f''(t) : x_i \le t \le x_i^*\}, \quad i = 1, 2, \ldots, n-1\,. \end{aligned}$$

If f is 3-convex on $[a, b]$, that is, the third derivative $f^{(3)}(x) \ge 0$ on $[a, b]$, then $m_i = f''(x_i)$ and $M_i = f''(x_i^*)$. If $f^{(3)}(x) \le 0$ on $[a, b]$ instead, then $m_i = f''(x_i^*)$ and $M_i = f''(x_i)$, $i = 1, \ldots, n$.

Theorem 2.5 below improves on the bounds on D, but assumes more conditions on f and its derivatives.

Theorem 2.5. *Let $f(t)$ be a real-valued function of t on $[a, b]$. Suppose that f and its first three derivatives f', f'' and $f^{(3)}$ are continuous on $[a, b]$. Suppose that $a \le x_1 < x_2 < \cdots < x_n \le b$. Suppose $p_i \ge 0$, $i = 1, 2, \ldots, n$ with $\sum_{i=1}^n p_i = 1$, $p_1 > 0$, $p_n > 0$.*

a) *Suppose that $f''(t) \ge 0$ and $f^{(3)}(t) \ge 0$, $a \le t \le b$. Let D be the difference*

$$D = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right).$$

Then

$$h_1 \equiv \sum_{j=1}^{n-1} W_j p_j \left(1 - \frac{p_j}{R_j}\right) \le D \le \sum_{j=1}^{n-1} V_j p_j \left(1 - \frac{p_j}{R_j}\right) \equiv h_2\,, \tag{2.18}$$

where x_j^*, $x^{(j)}$ and R_j are given in Theorem 2.3 above,

$$V_j = (x_j^* - x_j) f'(x_j^*) - f(x_j^*) + f(x_j)\,,$$

and

$$W_j = f(x_j^*) - f(x_j) - (x_j^* - x_j) f'(x_j), \quad j = 1, 2, \ldots, n-1\,. \tag{2.19}$$

b) *Suppose $f''(t) \ge 0$ and $f^{(3)}(t) \le 0$ on $[a, b]$. Then the reverse inequality holds, that is*

$$h_2 \le D \le h_1\,. \tag{2.20}$$

Proof of (a). The proof proceeds in the same way as the proof of Theorem 2.3. Then, as before, using the definitions of A_j given in (2.10) above,

(2.21)

$$A_j - A_{j+1} = R_j \left(f(x^{(j)}) - \left[\left(\frac{p_j}{R_j} \right) f(x_j) + \left(1 - \frac{p_j}{R_j}\right) f(x_j^*) \right] \right),\ j = 1, 2, \ldots, n-1\,.$$

Instead of using Lemma 2.1, we use Lemma 2.2, part (a) with $c = x_j$, $d = x_j^*$. Recalling that $L(x^{(j)}) = \frac{p_j}{R_j} f(x_j) + \left(1 - \frac{p_j}{R_j}\right) f(x_j^*))$, we obtain that $A_j - A_{j+1} = 0$, if $x_j = x_j^*$. If $x_j < x_j^*$, then we obtain

$$\begin{aligned} -W_j \left(\frac{(x^{(j)} - x_j)(x_j^* - x^{(j)})}{(x_j^* - x_j)^2} \right) &\le f(x^{(j)}) - L(x^{(j)}) \\ &\le -V_j \left(\frac{(x^{(j)} - x_j)(x_j^* - x^{(j)})}{(x_j^* - x_j)^2} \right) \ 1 \le j \le n-1\,. \end{aligned}$$

Multiplying by R_j, using (2.11) and summing over j gives, as done in the proof Theorem 2.3,

$$\sum_{j=1}^{n-1} R_j W_j \frac{(x^{(j)} - x_j)(x_j^* - x^{(j)})}{(x_j^* - x_j)^2} \le D \le \sum_{j=1}^{n-1} R_j V_j \frac{(x^{(j)} - x_j)(x_j^* - x^{(j)})}{(x_j^* - x^{(j)})^2} .$$

From (1), we obtain, upon cancellation of $(x_j^* - x_j)^2$ terms

$$\sum_{j=1}^{n-1} W_j p_j \left(1 - \frac{p_j}{R_j}\right) \le D \le \sum_{j=1}^{n-1} V_j p_j \left(1 - \frac{p_j}{R_j}\right) .$$

This completes the proof of part (a). ∎

The proof of (b) uses Lemma 2.2, part (b) instead and is omitted.

Remark 2.2. If $f''(t) \le 0$ on $[a,b]$ and either $f^{(3)}(t) \ge 0$ on $[a,b]$ or $f^{(3)}(t) \le 0$ on $[a,b]$, then we may apply Theorem 2.5 to $-f(t)$ instead to get bounds on $-D$, hence on D.

Remark 2.3. The quantities h_1 and h_2 are nearly Riemann sums for various integrals when passing from the finite n case to the continuous case as $n \to \infty$ and will be discussed later. At least one of these bounds for large n is usually very good. This will be seen in Theorem 4.3 later.

Remark 2.4. The bounds given in Theorems 2.3 and 2.5 use a series of local linear interpolations 'moving forward' from $x = x_1$ to $x = x_n$ from left to right. However, a moment's reflection reveals that we could also perform the sequence of local linear interpolations in 'reverse' order. Let $f^*(x) = f(a+b-x)$, where $a = x_1$, $b = x_n$, $a \le x \le b$. Then $f^*(a+b-x) = f(x)$, $a \le x \le b$ and $\sum_{i=1}^n p_i f(x_i) = \sum_{i=1}^n p_{n+1-i} f^*(a+b-x_{n+1-i})$. Replacing p_i by p_{n+1-i}, x_i by $a+b-x_{n+1-i}$ and f by f^* in Theorems 2.3 and 2.5 allows us to obtain 'reverse order' versions of L_J, J_J, h_1 and h_2. Call these L_J^*, U_J^*, h_1^* and h_2^*, respectively. These new 'reverse order' bounds sometimes do improve on L_J, U_J, h_1, h_2 bounds.

Next, we state the famous Hermite-Hadamard inequality. This will be needed in the proofs of several new results to be discussed in this paper. In addition, we shall present a refinement of this inequality later for convex functions which are also 3-convex on $[a,b]$.

Hermite-Hadamard inequality. Let f be a convex function on $[a,b]$, where $a < b$. Then

$$f\left(\frac{a+b}{2}\right) \le \frac{1}{b-a}\int_a^b f(x)dx \le \frac{f(a)+f(b)}{2} . \tag{2.22}$$

Next, we show that the bounds of Theorem 2.5 are better than Corollary 2.4 bounds, if we have the assumptions of Corollary 2.4 met.

Theorem 2.6. *Suppose that the assumptions of Theorem 2.5 hold.*

a) *If $f''(t) \ge 0$ and $f^{(3)}(t) \ge 0$ on $[a,b]$, then*

$$L_J \le h_1 \le D \le h_2 \le U_J .$$

b) *If $f''(t) \ge 0$ and $f^{(3)}(t) \le 0$ on $[a,b]$, then*

$$L_J \le h_2 \le D \le h_1 \le U_J .$$

Proof. We shall prove a) only. The proof of b) is very similar and is omitted. From Theorem 2.5, it remains only to prove $L_J \le h_1$ and $h_2 \le J$. We shall prove $h_2 \le U_J$ only; the proof that

$L_J \leq h_1$ is very similar and is omitted. Then

$$\begin{aligned} h_2 &= \sum_{j=1}^{n-1} V_j p_j \left(1 - \frac{p_j}{R_j}\right) \\ &= \sum_{j=1}^{n-1} [(x_j^* - x_j) f'(x_j^*) - f(x_j^*) + f(x_j)] \cdot p_j \left(1 - \frac{p_j}{R_j}\right) . \end{aligned} \tag{2.23}$$

By the Mean Value Theorem, there is a number $\lambda_j \in (0, 1)$ such that

$$f'(x_j + \lambda_j (x_j^* - x_j)) = \frac{f(x_j^*) - f(x_j)}{x_j^* - x_j} = \left(\int_{x_j}^{x_j^*} f'(t)dt\right) \left(\frac{1}{x_j^* - x_j}\right) . \tag{2.24}$$

But $f^{(3)}(t) \geq 0$, so f' is convex on $[a, b]$. By the Hermite-Hadamard inequality applied to f',

$$\frac{1}{x_j^* - x_j} \int_{x_j}^{x_j^*} f''(t)dt \geq f'\left(\frac{x_j + x_j^*}{2}\right) = f'\left(x_j + \frac{1}{2}(x_j^* - x_j)\right) . \tag{2.25}$$

From (2.24) and (2.25), we get $f'(x_j + \lambda_j (x_j^* - x_j)) \geq f'\left(x_j + \frac{1}{2}(x_j^* - x_j)\right)$. Since f' is increasing, $\lambda_j \geq \frac{1}{2}$ must hold. Thus,

$$f(x_j^*) - f(x_j) \geq f'\left(x_j + \frac{1}{2}(x_j^* - x_j)\right) \cdot (x_j^* - x_j)$$

and (2.23) gives

$$h_2 \leq \sum_{j=1}^{n-1} \left[f'(x_j^*) - f'\left(x_j + \frac{1}{2}(x_j^* - x_j)\right)\right] \cdot (x_j^* - x_j) p_j \left(1 - \frac{p_j}{R_j}\right) .$$

Applying the Mean Value Theorem again to the expression in brackets, we can find a number λ_j^* in $(0, 1/2)$ such that

$$f'(x_j^*) - f'\left(x_j + \frac{1}{2}(x_j^* - x_j)\right) = f''(\lambda_j^*) \cdot \frac{1}{2}(x_j^* - x_j) .$$

Thus,

$$\begin{aligned} h_2 &\leq \sum_{j=1}^{n-1} \frac{1}{2} f''(\lambda_j^*) p_j \left(1 - \frac{p_j}{R_j}\right) (x_j^* - x_j)^2 \\ &\leq \sum_{j=1}^{n-1} \frac{1}{2} M_j \cdot p_j \left(1 - \frac{p_j}{R_j}\right) (x_j^* - x_j)^2 = U_J , \end{aligned}$$

where $M_j = \sup\{f''(t) : x_j \leq t \leq x_j^*\}$. This completes the proof of a). ∎

How do the new bounds given in this paper compare to previously published bounds? For large n, the new bounds are quite good, based upon many numerical comparisons done and not reported here

Next, let's compare some of the new bounds to the previously proposed bounds discussed in previous works, both analytically, when it is possible to do so, and numerically. In most cases, comparisons are somewhat delicate since different bounds have slightly different assumptions, and have greater degrees of computational complexities so that we are comparing 'apples and oranges'. However, such comparisons give some idea about utility and relative merits of bounds. From the results of numerical comparisons, it is the case that most of the bounds, both old and new, are the best for some choices of p_i and x_i and n, but the new bounds compare very favorably to previously published bounds. Moreover, some of the new bounds are still valid for non-convex/non-concave choices of $f(x)$ and lead to some nice applications.

The following theorem states that bounds L_J and U_J are at least as good as the bounds F_L and F_U discussed [17] and given in Theorem G earlier. It should be mentioned, however, that the F_L and F_U bounds do not require continuity of f'', whereas the new bounds L_J and U_J of Theorem 2.3 do require it. Hence, the bounds of [17] are slightly more generally applicable.

Theorem 2.7. *Suppose that Theorem 2.3 assumptions hold. Let L_J and U_J be the new bounds of Corollary 2.4. Let F_L and F_U be the bounds discussed in Theorem G. Let*

$$D = \sum_{i=1}^{n} p_i f(x_i) - f\left(\sum_{I=1}^{n} p_i x_i\right).$$

Then

a) $F_L \le L_J \le D$
b) $D \le U_J \le F_L$.

Proof. We shall prove part a). The proof of part b) is similar and is omitted. The inequality $L_J \le D$ was proven in Corollary 2.4. First, replacing $f(x)$ by $f_A(x) = x^2$ in Theorem 2.3, we obtain

$$\sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2 = \sum_{j=1}^{n-1} p_j \left(1 - \frac{p_j}{R_j}\right)(x_j^* - x_j)^2.$$

Then, with $m = \inf\{f''(t) : a \le t \le b\}$, $m_i = \inf\{f''(t) : x_i \le t \le x_i^*\}$, we get

$$\begin{aligned} F_L &= \frac{m}{2}\left(\sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{I=1}^{n} p_i x_i\right)^2\right) = \frac{m}{2}\left(\sum_{j=1}^{n-1} p_j \left(1 - \frac{p_j}{R_j}\right)(x_j^* - x_j)^2\right) \\ &\le \sum_{j=1}^{n-1} \frac{m_j}{2} p_j \left(1 - \frac{p_j}{R_j}\right)(x_j^* - x_j)^2 = L_J. \end{aligned}$$

This completes the proof of part a). ∎

From numerical comparisons done, it appears that the following general conclusions are valid.

(1) The new bounds based on h_1 and h_2 of Theorem 2.5 are often the best, but these are not as widely applicable as the other bounds, both old and new.
(2) The new lower bounds L_J and L_{J^*} are not as good as L_{DS} when the variance of x_i is large, unless n is large. Then they are significantly better.
(3) Among the upper bounds, U_J and U_{J^*} are usually better than $U_{D,1}$, $U_{D,2}$, $U_{BPP}(1)$, $U_{BPP}(2)$, U_{DG}, if the variance of the x_i values is not too large. But for large n, U_J and U_{J^*} are substantially better regardless of variance.
(4) For non-convex/non-concave choices for $f(x)$ with $f''(x)$ continuous, only the new bounds given in Theorem 2.3 and Corollary 2.4 and the F_L, F_U bounds of [17] are applicable. Of course, in this case, Theorem 2.7 stated that the new bounds are at least as good.
(5) Most of the bounds discussed in this paper are the best or nearly so, for some choices of p_i and x_i. Given this fact, how amenable are the bounds, both old and new, for convenient applications? It will be seen that the new bounds can be used to refine various

inequalities, such as the Hermite-Hadamard inequality and to obtain brand new inequalities in probability and reliability theory. We briefly discuss some other applications first, which are more of a classical nature.

Remark 2.5. If $f(x)$ is log-convex on $[a, b]$, then we may apply Theorems 2.3 and 2.5 to $Lnf(x)$ instead of $f(x)$ to obtain bounds on $D^* = \sum_{i=1}^n p_i Lnf(x_i) - Lnf\left(\sum_{i=1}^n p_i x_i\right)$. Then exponentiation will provide bounds on the ratio

$$e^{D^*} = \frac{\prod_{i=1}^n (f(x_i))^{p_i}}{f\left(\sum_{I=1}^n p_i x_i\right)} .$$

If $f(x) = Ln\left[\left(\frac{1-x}{x}\right)^r\right]$, $r > 0$, $0 < x < 1/2$, then we obtain new and improved inequalities of Ky-Fan type. In [11], upper bounds are given for e^{D^*}. See their Theorem 1, p. 52 and Proposition 5.6, p. 61. Numerical comparisons of these bounds to the new bounds suggest that the new bounds derived from Theorem 2.5 are always better and that the Corollary 2.4 bounds for e^{D^*} are better if either the variance of the x_i values is not too great, or if n is large. Theorems 2.3 and 2.5 can be used to obtain new refinements of the classical arithmetic mean-geometric mean (AM-GM) inequality. These new refinements compare favorably to many previously discussed bounds for the difference and ratio of arithmetic mean and geometric means. Theorems 2.3 and 2.5 have been applied to obtain many new bounds in information theory such as b-entropy bounds discussed in Theorem 4.3 of [11]. These new bounds, in numerical comparisons done have been found to compare very favorably to other bounds in this case as well.

The new bounds can also be widely applied to obtain bounds for quantities of interest in information theory. We consider here one such application. Suppose X is a discrete random variable with support $\{x_1, x_2, \ldots, x_n\}$ where $0 < x_1 < \cdots < x_n$. Let $p_i = \text{Prob}(X = x_i)$, $i = 1, 2, \ldots, n$. Suppose $p_i > 0$, $i = 1, 2, \ldots, n$. Then the b-entropy of X is, using the notation of Dragomir and Goh (1996):

$$H_b(x) = \sum_{I=1}^n p_i \text{Log}_b\left(\frac{1}{p_i}\right), \quad b > 1 .$$

In Dragomir and Goh (1996), the following theorem is given.

Theorem I. *(Theorem 4.3 of Dragomir and Goh (1991).)*

$$0 \le \text{Log}_b n - H_b(x) \le \frac{1}{Ln\, b}\left[n \sum_{i=1}^n p_i^2 - 1\right] \equiv U_H . \tag{2.26}$$

Equality holds if and only if $p_i = \frac{1}{n}$, $i = 1, 2, \ldots, n$.

If we take $f(x) = -\text{Log}_b x$ in Theorem 2.5, part (b), we immediately obtain the following corollary, which gives both an upper and lower bound for $\text{Log}_b n - H_b(x)$.

Corollary 2.8. *Suppose that $p_j > 0$, $j = 1, 2, \ldots, n$, $\sum_{j=1}^n p_j = 1$. Let $P_{(1)} \ge P_{(2)} \ge \cdots \ge P_{(n)}$ denote the p_j values ordered from largest to smallest. Let $x_j = \frac{1}{P_{(j)}}$, $j = 1, 2, \ldots, n$. Let $f(x) = -Log_b x$. Then*

$$I_L \le Log_b n - H_b(x) \le I_U , \tag{2.27}$$

where I_L is the value of h_2 in Theorem 2.5 with $P_{(j)}$ replacing p_j everywhere in Theorems 2.3 and 2.5 and I_U is the value of h_1 with these same replacements.

Proof. The results is immediate, except we note here that since it was assumed that $x_1 \leq x_2 \leq \cdots \leq x_n$, and we are choosing $x_i = \frac{1}{p_i}$, we must order the p_j values before we can apply Theorem 2.5, part (b). This is no problem since entropies are invariant to permutations. Of course, if $p_j = \frac{1}{n}$ for all j, then we obtain $I_L = J_U = 0$. ∎

Numerical comparisons of the upper bound U_H given in Theorem 4.3 of [10] with the new bound I_U of Corollary 2.4 found that the new upper bound I_U is always at least as good as U_H, but no proof could be found. However, Corollary 2.4 also provides a very good lower bound I_L. Here we present a small numerical comparison.

Suppose $n = 3$, $P_1 = .5$, $P_2 = .3$, $P_3 = .2$ Then we take $b = e$ here, but the choice of the base $b > 1$ is irrelevant for purposes of comparison. We obtain

$$\text{Log}_e n - H_e(X) = 0.0690, \quad U_H = 0.1400\,,$$

$$I_L = 0.0569,\ I_U = 0.0881,\ L_1 = 0.0379,\ U_1 = 0.1400\,.$$

The bound I_U improves on U_H especially if the p_i values are not extremely diverse. The lower bound I_L is also quite good. A theoretical reason for this will be given later. The U_H bound is, however, easier to compute than the new bounds. U_1 is better than U_H when the p_i values are not too diverse. Otherwise, U_H is better than U_1. (Here, both U_H and U_1 are the same.)

3. INTEGRAL REPRESENTATIONS AND BOUNDS

Theorem 3.1 below extends the bounds to the continuous case. We present bounds for the continuous Jensen gap D.

In the continuous case, we assume that $H(x)$ in Theorem 3.1 below is continuous on $[a, b]$, with $H(a) = 0$, $H(b) = 1$. Thus, $H(x)$ is a continous distribution function on $[a, b]$. In most, but not all cases, there will exist a probability density function absolutely continuous with respect to Lebesgue measure, $h(x)$ with $H'(x) = h(x)$ on $[a, b]$, but we do not have to make this assumption in Theorem 3.1. For this reason, we shall use Riemann-Stieltjes integrals in Theorem 3.1 to obtain a slightly more general result. All integrals below are of Lebesgue-Stieltjes type. Since we are intergrating with respect to a bounded continuous monotonic function, they can be considered Riemann-Stieltjes integrals as well. The function $g(x)$ given in (3.1) in Theorem 3.1 is called the mean residual life function and plays a prominent role in applied probability, reliability theory and statistics. See [18], for example. First, we consider the bounded support case.

Next, we consider the continuous version of Jensen's gap utilizing Lebesgue/Riemann Stieltjes integrals.

Theorem 3.1. *Suppose that f'' is continuous on $[a, b]$. Let $H(x)$ be a bounded, continuous nondecreasing function on $[a, b]$ with $H(a) = 0$ and $H(b) = 1$. Suppose $0 < H(x) < 1$ on (a, b). Let*

$$g(x) = \begin{cases} \dfrac{\int_x^b (t - x)\,dH(t)}{1 - H(x)}, & a \leq x < b, \\ \lim_{x \to b^-} g(x) = 0, & x = b. \end{cases} \tag{3.1}$$

Let

$$q_1(x) = \inf\{f''(t) : x \leq t \leq x + g(x)\}\,, \tag{3.2}$$

and

$$q_2(x) = \sup\{f''(t) : x \leq t \leq x + g(x)\}\,. \tag{3.3}$$

Let

$$L_1 = \frac{1}{2}\int_a^b q_1(x)(g(x))^2 dH(x)\,, \tag{3.4}$$

and

$$U_1 = \frac{1}{2}\int_a^b q_2(x)(g(x))^2 dH(x)\,. \tag{3.5}$$

Then

$$L_1 \le \int_a^b f(x)dH(x) - f\left(\int_a^b x dH(x)\right) \le U_1\,. \tag{3.6}$$

The proof of Theorem 3.1 will utilize Lebesgue-Stieltjes sum approximations. Since $H(t)$ is a bounded nondecreasing function, this is equivalent to Riemann-Stieltjes sum approximations.

Given any partition $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ with $a = x_0 < x_1 < \cdots < x_n = b$, we shall let $\|\mathcal{P}\|$ denote the norm or mesh of the partition. Then $\|\mathcal{P}\| = \max\{x_{i+1} - x_i : 0 \le i \le n-1\}$. We shall apply Theorems 2.3 and 2.5 to a partition $\mathcal{P}$ with suitably small enough $\|\mathcal{P}\|$. We shall assume $x_{i+1} - x_i = \|\mathcal{P}\| = \Delta x$, $i = 0, 1, \ldots, n-1$ throughout. Let $p_j = H(x_j) - H(x_{j-1})$, $j = 1, 2, \ldots, n$. Then $p_j \ge 0$, $j = 2, \ldots, n-1$, $p_1 > 0$, $p_n > 0$ and $\sum_{j=1}^n p_j = 1$. The Riemann-Stieltjes sums involved with utilize integrand values at $x = x_j$ $j = 1, 2, \ldots, n$. Let R_i, x_i^* and $x^{(i)}$ be given by (2.7)–(2.8) of Theorem 2.3 for these choices of x_j and p_j, $j = 1, 2, \ldots, n$. Then $R_i = 1 - H(x_i)$, $i = 1, 2, \ldots, n-1$.

To prove Theorem 3.1, we need the following lemmas.

Lemma 3.2. *There exists a constant $K_1 > 0$ such that*

$$|g^2(x_i^*) - (x_i^* - x_i)^2| \le K_1 \Delta x,\ i = 1, 2, \ldots, n\,. \tag{3.7}$$

Proof. Let

$$\begin{aligned} B_i &= 1 - H(x_i), \quad C_i = \int_{x_i}^b (t - x_i) dH(t)\,, \\ \widehat{C}_i &= \sum_{L=i}^{n-1} (x_{L+1} - x_i)\cdot (H(x_{L+1}) - H(x_i))\,. \end{aligned}$$

Then

$$g(x_i) - (x_i^* - x_i) = \frac{C_i - \widehat{C}_i}{B_i} = \frac{\sum_{L=i}^{n-1}\int_{x_L}^{x_{L+1}} (t - x_{L+1}) dH(t)}{\sum_{L=i}^{n-1}\int_{x_L}^{x_{L+1}} dH(t)}\,.$$

Now $|t - x_{L+1}| \le \Delta x$, $L = i, \ldots, n-1$. Thus,

$$|g(x_i) - (x_i^* - x_i)| \le \Delta x,\ i = 1, 2, \ldots, n-1\,. \tag{3.8}$$

Also, clearly we have

$$\begin{aligned} |g(x_i) + (x_i^* - x_i)| &\le \sup\{g(t) : a \le t \le b\} + \max\{x_i^* - x_i,\ i = 1, 2, \ldots, n-1\} \\ &\le (b-a) + (b-a) = 2(b-a)\,. \end{aligned} \tag{3.9}$$

Multiplication of (3.8) and (3.9) produces

$$|g^2(x_i) - (x_i^* - x_i)^2| \le 2(b-a)\Delta x,\ i = 1, 2, \ldots, n\,.$$

Letting $K_1 = 2(b-a)$, the proof of Lemma 3.2 is complete. ∎

Lemma 3.3. *Under the conditions of Lemma 3.2, we have the following. Let*

$$\begin{aligned} Z_i &= \sup\left\{|f''(t) - f''(x_i^*)| : x_i^* \le t \le x_i + g(x_i)\right\}, \text{ if } x_i^* \le x_i + g(x_i), \\ Z_i &= \sup\left\{|f''(t) - f''(x_i + g(x_i)| : x_i + g(x_i) \le t \le x_i^*\right\}, \\ &\qquad \text{if } x_i + g(x_i) \le x_i^*,\ i = 1, 2, \ldots, n-1. \end{aligned}$$

Then

$$|\sup\{f''(t) : x_i \le t \le x_i^*\} - \sup\{f''(t) : x_i \le t \le x_i + g(x_i)\}| \le Z_i,\ i = 1, 2, \ldots, n-1.$$

Proof. Consider the case $x_i^* \le x_i + g(x_i)$. Then for $x_i \le t \le x_i + g(x_i)$, we have

$$\begin{aligned} f''(t) &= f''(x_i^*) + (f''(t) - f''(x_i^*)) \\ &\le \sup\{f''(t) : x_i \le t \le x_i^*\} + (f''(t) - f''(x_i^*)) \\ &\le \sup\{f''(t) : x_i \le t \le x_i^*\} + Z_i, \end{aligned}$$

since $x_i \le x_i^* \le x_i + g(x_i)$. Since this holds for all t with $x_i \le t \le x_i + g(x_i)$, we obtain

$$\sup\{f''(t) : x_i \le t \le x_i + g(x_i)\} \le \sup\{f''(t) : x_i \le t \le x_i^*\} + Z_i.$$

Since $[x_i, x_i^*] \subseteq [x_i, x_i + g(x_i)]$, subtraction proves the lemma. The proof of the other case is similar and is omitted. ∎

Proof of Theorem 3.1. First, assume that $H(x)$ satisfies $H(x) \le \frac{k+x-a}{k+b-a}$, $a \le x \le b$ or some integer k. We shall then prove the general case.

We shall prove only the upper bound right half of (3.6) in Theorem. 3.1 The proof of the other half is very similar and is omitted. Let's show that there exists a positive number M such that for every $\epsilon > 0$, we have

$$\int_a^b f(x)dH(x) - f\left(\int_a^b xdH(x)\right) \le U_1 + \epsilon M, \tag{3.10}$$

where $U_1 = \frac{1}{2}\int_a^b q_2(x)(g(x))^2 dH(x)$. Since f is uniformly continuous on $[a, b]$, there is a $\delta_1 > 0$ such that if $\mathcal{P} = \{x_1, x_2, \ldots, x_n\}$ is a partition of $[a, b]$ with $\left|\int_a^b xdH(x) - \sum_{i=1}^n p_i x_i\right| < \delta_1$, then

$$\left|f\left(\int_a^b xdH(x)\right) - f\left(\sum_{i=1}^n p_i x_i\right)\right| < \epsilon.$$

Since $\sum_{i=1}^n p_i x_i$ is a Riemann-Stieltjes sum for $\int_a^b xdH(x)$, there is a $\delta_2 > 0$ such that $\|\mathcal{P}\| < \delta_2$ implies $\left|\int_a^b xdH(x) - \sum_{i=1}^n p_i x_i\right| < \delta_1$ and thus

$$\left|f\left(\int_a^b xdH(x)\right) - f\left(\sum_{i=1}^n p_i x_i\right)\right| < \epsilon. \tag{3.11}$$

Similarly, there is a $\delta_3 > 0$ such that $\|\mathcal{P}\| < \delta_3$ implies

$$\left|\int_a^b f(x)dH(x) - \sum_{i=1}^n p_i f(x_i)\right| < \epsilon. \tag{3.12}$$

Taking $\delta_4 = \min(\delta_1, \delta_3)$, we see that if $\|\mathcal{P}\| < \delta_4$, then

$$\int_a^b f(x)dH(x) - f\left(\int_a^b x dH(x)\right)$$

$$\leq \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) + 2\epsilon\,. \tag{3.13}$$

Now apply Theorem 2.3 to the right-hand side of (3.13) utilizing Lemmas 3.2 and 3.3 to show that (3.10) holds. To this end, define

$$\begin{aligned} Y_{1,i} &= \sup\{f''(t) : x_i \leq t \leq x_i^*\} \\ Y_{2,i} &= \sup\{f''(t) : x_i \leq t \leq x_i + g(x_i)\},\ i = 1, 2, \ldots, n-1 \\ &= q_2(x_i)\,. \end{aligned}$$

Let

$$\begin{aligned} Q_1 &= \frac{1}{2}\sum_{i=1}^{n-1} Y_{1,i} p_i \left(1 - \frac{p_i}{R_i}\right)(x_i^* - x_i)^2\,, \\ Q_2 &= \frac{1}{2}\sum_{i=1}^{n-1} Y_{2,i} p_i \left(1 - \frac{p_i}{R_i}\right)(g(x_i))^2\,, \\ Q_3 &= \frac{1}{2}\sum_{i=1}^{n-1} Y_{1,i} p_i \left(1 - \frac{p_i}{R_i}\right)(g(x_i))^2\,, \\ Q_4 &= \frac{1}{2}\sum_{i=1}^{n-1} Y_{2,i} p_i (g(x_i))^2\,. \end{aligned}$$

Note that $Y_{1,i} = M_i$ in Corollary 2.4 and Q_1 is the upper bound on

$$D = \sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right)$$

given there. Then

$$\sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right) \leq Q_1\,. \tag{3.14}$$

Now, clearly, we have

$$|Q_1 - U_1| \leq |Q_1 - Q_3| + |Q_3 - Q_2| + |Q_2 - Q_4| + |Q_4 - U_1|\,. \tag{3.15}$$

Then

$$|Q_1 - Q_3| \leq \frac{1}{2}\sum_{i=1}^{n-1} |Y_{1,i}| \cdot p_i \left(1 - \frac{p_i}{R_i}\right) \cdot |(x_i^* - x_i)^2 - (g(x_i))^2|\,. \tag{3.16}$$

Since $|f''|$ is bounded, Lemma 3.2 shows that $|Q_1 - Q_3|$ is $O(\Delta x)$, that is, there is a constant $K_2 > 0$ such that

$$|Q_1 - Q_3| \leq K_2 \Delta x\,.$$

By Lemmas 3.2 and 3.3, from the fact that $|Y_{1,i} - Y_{2,i}| \leq Z_i$ and the uniform continuity of f'' on $[a, b]$, given $\epsilon > 0$, we may take Δx small enough so that

$$|Q_3 - Q_2| \leq \frac{\epsilon}{2}\sup\{(g(t))^2 : a \leq t \leq b\} \leq \frac{\epsilon}{2}(b-a)^2\,. \tag{3.17}$$

Since Q_4 is a Riemann sum approximation to U_1, given $\epsilon > 0$, we may choose Δx small enough so that

$$|Q_4 - U_1| \le \epsilon\,. \tag{3.18}$$

It remains to show that $|Q_2 - Q_4|$ can be made arbitrarily small as $\|\mathcal{P}\| = \Delta x \to 0$. By the boundedness of $|f''|$ on $[a, b]$, there exists a constant K_2 such that $|f''(x)| \le K_2$, $a \le x \le b$. Then

$$|Q_2 - Q_4| \le \frac{1}{2} K_2 \sum_{i=1}^{n-1} \frac{p_i^2}{R_i} (g(x_i))^2\,.$$

Clearly, $g(x_i) \le b - x_i$, $i = 1, \dots, n$. Also, $R_i = 1 - H(x_i)$, $i = 1, 2, \dots, n-1$. Let $S(x) = \frac{(b-x)^2}{1-H(x)}$, $a \le x < b$. Since $H(x) \le \frac{k+x-a}{k+b-a}$, $S(x) \le (k + b - a)(b - x)$, $a \le x < b$. Let $S(b) \equiv \lim_{x\to b^-} S(x) = 0$. Then $S(x)$ is bounded on $[a, b]$. Since $0 \le \frac{(g(x_i))^2}{R_i} \le S(x_i)$, there is a constant K_3 such that $0 \le \frac{(g(x_i))^2}{R_i} \le K_3$. Thus,

$$|Q_2 - Q_4| \le \frac{1}{2} K_2 K_3 \sum_{i=1}^{n-1} p_i^2\,.$$

Now

$$\sum_{i=1}^{n-1} p_i^2 \le \sum_{i=1}^{n} p_i^2 \le \sum_{i=1}^{n} p_i \cdot \sup_{1\le i\le n} (H(x_i) - H(x_{i-1}))$$

which can be made arbitrarily small as $n \to \infty$, by the uniform continuity of $H(\cdot)$ on $[a, b]$, and since $\sum_{i=1}^{n} p_i = 1$. Thus, we may take $\|\mathcal{P}\| = \Delta x$ small enough so that

$$|Q_2 - Q_4| \le \epsilon\,. \tag{3.19}$$

Since all four absolute differences in (3.15) can be made arbitrarily small as $\Delta x = \|\mathcal{P}\| \to 0$, we can make $|Q_1 - U_1|$ arbitrarily small. So we can choose Δx small enough so that, from (3.15)–(3.19), we obtain

$$|Q_1 - U_1| \le K_2 \Delta x + \frac{\epsilon}{2} (b - a)^2 + 2\epsilon\,.$$

Choosing $\Delta x \le \epsilon$, then

$$|Q_1 - U_1| \le \left(K_2 + \frac{1}{2}(b - a)^2 + 2\right) \cdot \epsilon\,.$$

In particular,

$$Q_1 \le U_1 + \left(K_2 + \frac{1}{2}(b - a)_2^2\right) \cdot \epsilon\,. \tag{3.20}$$

From the definition of Q_1 and Corollary 2.4, and (3.20), we obtain for Δx small enough,

$$\begin{aligned} &\int_a^b f(x)dh(x) - f\left(\int_a^b x dH(x)\right) \\ &\le \sum_{i-1}^{n} p_i f(x_i) - f\left(\sum_{i=1}^{n} p_i x_i\right) \le Q_1 \le U_1 + M\epsilon\,, \end{aligned}$$

where $M = K_2 + \frac{1}{2}(b - a)^2 + 2\,.$

Thus, (3.10) holds for all $\epsilon > 0$ and $H(x)$ with $H(x) \le \frac{k+x-a}{k+b-a}$ for some integer k. Now suppose the general case for $H(x)$. Let

$$H_k(x) = \frac{H(x) + \frac{x-a}{k}}{1 + \frac{b-a}{k}}, \quad a \le x \le b, \quad k = 1, 2, 3, \dots .$$

Then

$$H_k(x) \le \frac{k+x-a}{k+b-a}, \quad a \le x \le b, \quad k = 1, 2, 3, \dots .$$

Thus, the conclusion of Theorem 3.1 is valid for $H_k(x)$ in place of $H(x)$. Now $\lim_{k\to\infty} H_k(x) = H(x)$, $a \le x \le b$, $|H_k(x)| \le 1$, $a \le x \le b$, $k \ge 1$. Let $g_k(x)$, $q_{2,k}(x)$, $L_{1,k}$ and $U_{1,k}$ be the values of $g(x)$, $q_2(x)$, L_1 and U_1 given in Theorem 3.1 with $H_k(x)$ replacing $H(x)$ throughout, respectively. Then

$$\int_a^b f(x)dH_k(x) - f\left(\int_a^b x \cdot dH_k(x)\right) \le U_{1,k}, \tag{3.21}$$

where

$$\begin{aligned} U_{1,k} &= \frac{1}{2}\int_a^b q_{2,k}(x)(g_k(x))^2 dH_k(x), \\ g_k(x) &= \frac{\int_x^b (t-x)dH_k(t)}{1-H_k(x)}, \quad a \le x < b, \quad g_k(b) = 0, \text{ and} \\ q_{2,k}(x) &= \sup\{f''(t) : x \le t \le x + g_k(x)\}, \end{aligned}$$

Clearly, $|q_{2,k}(x)| \le \sup\{|f''(t)| : a \le t \le b\}$. Also, $|g_k(x)| \le b-a$, $a \le x \le b$, $k \ge 1$. Also, $\lim_{k\to\infty} q_{2,k}(x) = q_2(x)$. In addition,

$$\begin{aligned} \lim_{k\to\infty}\int_a^b f(x)dH_k(x) &= \int_a^b f(x)dH(x) + \lim_{k\to\infty}\frac{-(b-a)}{(k+b-a)}\int_a^b f(x)dx \\ &= \int_a^b f(x)dH(x). \end{aligned} \tag{3.22}$$

Similarly, by the continuity of f,

$$\begin{aligned} \lim_{k\to\infty} f\left(\int_a^b xdH_k(x)\right) &= \lim_{k\to\infty} f\left(\int_a^b xdH(x) - \frac{\frac{1}{2}(b-a)(b^2-a^2)}{k+b-a}\right) \\ &= f\left(\int_a^b xdH(x)\right), \end{aligned} \tag{3.23}$$

and

$$\begin{aligned} \lim_{k\to\infty} g_k(x) &= \frac{\lim_{k\to\infty}\left(\int_x^b (t-x)dH(t) + \int_x^b (t-x)\cdot\left(\frac{-(b-a)}{k+b-a}\right)dt\right)}{1-H(x)} \\ &= g(x), \quad a \le x < b. \end{aligned}$$

This holds for $x = b$, also. Also, by the Dominated Convergence Theorem,

$$\lim_{k\to\infty} U_{1,k} = \lim_{k\to\infty} \frac{1}{2}\left(\int_a^b q_{2,k}(x)(g_k(x))^2 dH(x) + \int_a^b q_{2,k}(x)(g_k(x))^2 \cdot \frac{-(b-a)}{k+b-a}dx\right)$$
$$= \frac{1}{2}\int_a^b q_2(x)(g(x))^2 dH(x) = U_1 . \tag{3.24}$$

and the proof of Theorem 3.1 is complete. ∎

Remark 3.1. In a few results to come later, we shall need to assume the existence of a density function $h(x) = H'(x)$ on $[a, b]$. This will be needed in some applications to be discussed next. In this case, the mean residual life function is

$$g(x) = \frac{\int_x^b (t-x)h(t)dt}{\int_x^b h(t)dt}, \quad a \le x < b$$
$$g(b) = 0 . \tag{3.25}$$

4. EXACT INTEGRAL REPRESENTATIONS AND APPLICATIONS

Next, we present various applications of some theorems in Section 3. They are improvements on the classical Hermite-Hadamard inequality. Many such papers have been published and they are too numerous to cite. However, see [1], [3], [22] and [23].

Theorem 4.1 below is a refinement of the Hermite-Hadamard inequality for convex functions which are also 3-convex on $[a, b]$, and is an application of Theorem 3.1.

Theorem 4.1. *Let f be a convex function on $[a, b]$. Suppose that f is also 3-convex on $[a, b]$, that is, the third derivative $f^{(3)}(x) \ge 0$ on $[a, b]$. Then*

a)

$$0 \le N_1 \le \frac{1}{b-a}\int_a^b f(x)dx - f\left(\frac{a+b}{2}\right) \le N_2 \tag{4.1}$$

where

$$N_1 = \frac{1}{2}\left[\frac{-(b-a)}{4}f'(a) + \frac{1}{2}\left(f\left(\frac{a+b}{2}\right) - f(a)\right)\right] \tag{4.2}$$

and

$$N_2 = \frac{-(b-a)}{4}f'\left(\frac{a+b}{2}\right) + \frac{1}{2}\left(f(b) - f\left(\frac{a+b}{2}\right)\right) \tag{4.3}$$

and

b)

$$\frac{1}{b-a}\int_a^b f(x)dx \le N_2 + f\left(\frac{a+b}{2}\right) \le \frac{f(a)+f(b)}{2}, \tag{4.4}$$

that is, $N_2 + f\left(\frac{a+b}{2}\right)$ is at least as good an upper-bound as $\frac{f(a)+f(b)}{2}$ is for $\frac{1}{b-a}\int_a^b f(x)dx$.

Proof of (a). We apply Theorem 3.1 to obtain N_1 from L_1 in Theorem 3.1. Then $h(x) = \frac{1}{b-a}$, $a \le x \le b$, $g(x) = \frac{b-x}{2}$, and $q_1(x) = f''(x)$, since $f^{(3)} \ge 0$. Then

$$\begin{aligned} &\frac{1}{b-a}\int_a^b f(x)dx - f\left(\frac{a+b}{2}\right) \\ &\ge L_1 = \frac{1}{2}\int_a^b q_1(x)h(x)(g(x))^2 dx = \frac{1}{2}\int_a^b f''(x)\cdot\frac{1}{b-a}\cdot\left(\frac{b-x}{2}\right)^2 dx\,. \end{aligned}$$

Integrating by parts twice, we obtain

$$\begin{aligned} L_1 &= \frac{-(b-a)}{8}f'(a) + \frac{1}{2(b-a)}\int_a^b \left(\frac{b-x}{2}\right) f'(x)dx \\ &= \frac{1}{2(b-a)}\left[-\frac{(b-a)^2}{4}f'(a) - \left(\frac{b-a}{2}\right)f(a) + \frac{1}{2}\int_a^b f(x)dx\right]. \end{aligned} \tag{4.5}$$

Applying the Hermite-Hadamard inequality to the integral in (4.5), we obtain:

$$\begin{aligned} L_1 &\ge \frac{1}{2(b-a)}\left[-\frac{(b-a)^2}{4}f'(a) - \left(\frac{b-a}{2}\right)f(a) + \frac{1}{2}(b-a)\cdot f\left(\frac{a+b}{2}\right)\right] \\ &= \frac{1}{2}\left[-\frac{(b-a)}{4}f'(a) - \frac{1}{2}f(a) + \frac{1}{2}f\left(\frac{a+b}{2}\right)\right] \\ &= N_1\,. \end{aligned}$$

This proves the first non-trivial left half of (4.1). To prove the other half involving N_2, we proceed similarly with a few modifications. Using $q_2(x) = f''(x+g(x)) = f''\left(\frac{x+b}{2}\right)$, Theorem 3.1 gives

$$\begin{aligned} U_1 &= \frac{1}{2}\int_a^b q_2(x)\cdot h(x)\cdot(g(x))^2 dx \\ &= \frac{1}{2}\int_a^b f''\left(\frac{x+b}{2}\right)\cdot\frac{1}{b-a}\cdot\left(\frac{b-x}{2}\right)^2 dx\,. \end{aligned}$$

Making the substitution $u = \frac{x+b}{2}$ and again integrating by parts twice, we obtain

$$\begin{aligned} U_1 &= \int_{\frac{a+b}{2}}^b \frac{1}{b-a}f''(u)\cdot(b-u)^2 du \\ &= \frac{-(b-a)}{4}f'\left(\frac{a+b}{2}\right) - \frac{1}{b-a}\int_{\frac{a+b}{2}}^b 2(u-b)f'(u)du \\ &= \frac{-(b-a)}{4}f'\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2}\right) + 2\int_{\frac{a+b}{2}}^b f(u)du \end{aligned} \tag{4.6}$$

Applying the Hermite-Hadamard inequality on $[\frac{a+b}{2}, b]$ to the integral in (3.6), we obtain:

$$\begin{aligned} U_1 &\le \frac{-(b-a)}{4}f'\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2}\right) + \frac{1}{2}\left(f\left(\frac{a+b}{2}\right) + f(b)\right) \\ &= \frac{-(b-a)}{4}f'\left(\frac{a+b}{2}\right) - \frac{1}{2}f\left(\frac{a+b}{2}\right) + \frac{1}{2}f(b) \\ &= N_2\,. \end{aligned}$$

This complets the proof. ∎

Proof of (b). Now simple algebra gives

$$N_2 + f\left(\frac{a+b}{2}\right) \le \frac{f(a)+f(b)}{2}$$

if and only if

$$f\left(\frac{a+b}{2}\right) - f(a) \le \left(\frac{b-a}{2}\right) f'\left(\frac{a+b}{2}\right),$$

which is equivalent to

$$\int_a^{\frac{a+b}{2}} f'(t)dt \le \frac{(b-a)}{2} f'\left(\frac{a+b}{2}\right).$$

But f is convex, so $f'(t) \le f'\left(\frac{a+b}{2}\right)$, $a \le t \le \frac{a+b}{2}$. So

$$\int_a^{\frac{a+b}{2}} f'(t)dt \le \left(\frac{b-a}{2}\right) f'\left(\frac{a+b}{2}\right)$$

holds and the proof of (b) is complete. ∎

Remark 4.1. Upper bound (4.6) is a better upper bound than N_2 and is exact for quadratic choices of $f(x)$. Similarly, (4.5) is a better lower bound than N_1 and is also exact for quadratic choices for $f(x)$. Using (4.5) and (4.6), we could, if desired, get even better refinements by applying the Theorem 4.3 to the integrals in (4.5) and (4.6) and 'iterating' to the limit on successive intervals half as large at each iteration. Of course, this would give much longer expressions for bounds and involve evaluation of f and f' at more and more points in $[a, b]$.

Theorem 4.1 required convexity of f in the form of $f''(x) \ge 0$ on $[a, b]$. But, as pointed out in Fink and Páles (2007), we may often replace this condition by less restrictive assumptions on $f(x)$. In any case, the next theorem, Theorem 4.2, gives a Hermite-Hadamard type of bound requiring only 3-convexity or 3-concavity on $[a, b]$, and not convexity of f itself.

Theorem 4.2. *Suppose $f^{(3)}(x)$ is continuous on $[a, b]$.*

a) *If $f(x)$ is 3-convex on $[a, b]$, that is, $f^{(3)}(x) \ge 0$ on $[a, b]$, then*

$$\frac{\int_a^b f(x)dx}{b-a} \le f\left(\frac{a+b}{2}\right) + \left(\frac{b-a}{12}\right) \cdot \left(f'(b) - f'\left(\frac{a+b}{2}\right)\right) \tag{4.7}$$

and

$$\frac{\int_a^b f(x)dx}{b-a} \ge f\left(\frac{a+b}{2}\right) + \left(\frac{b-a}{12}\right) \cdot \left(f'\left(\frac{a+b}{2}\right) - f'(a)\right). \tag{4.8}$$

b) *If $f(x)$ is 3-concave on $[a, b]$, then inequalities (4.7) and (4.8) hold with the inequality signs reversed.*

Proof. We shall prove only (4.7) part (a). The proof of (4.8) of part (a) follows upon considering the function $f^*(x) = f(a + b - x)$ and noting that $f^*(x)$ is 3-convex, if $f(x)$ is 3-concave and vice-versa. Also, $\int_a^b f(x)dx = \int_a^b f^*(x)dx$ and $f\left(\frac{a+b}{2}\right) = f^*\left(\frac{a+b}{2}\right)$. Part (b) follows from part (a) using $-f(x)$ instead of $f(x)$ in part (a). To prove (a), apply Theorem 3.1 with $H(x) = \frac{x-a}{b-a}$, $a \le x \le b$. Then $g(x) = \frac{b-x}{2}$, $a \le x \le b$. Since f is 3-convex, f'' is nondecreasing in its argument. Clearly $(g(x))^2 = \left(\frac{b-x}{2}\right)^2$ is nonincreasing in x. Then Theorem 3.1 gives

$$\frac{\int_a^b f(x)dx}{b-a} \le f\left(\frac{a+b}{2}\right) + \frac{1}{2}\int_a^b f''(x + g(x)) \cdot g^2(x) \cdot \frac{1}{b-a}dx.$$

An application of the Chebychev-Gruss inequality gives:

$$\begin{aligned}\frac{\int_a^b f(x)dx}{b-a} &\leq f\left(\frac{a+b}{2}\right)+\frac{1}{2}\int_a^b f''\left(\frac{x+b}{2}\right)\cdot\left(\frac{b-x}{2}\right)^2\cdot\frac{1}{b-a}dx \\ &\leq f\left(\frac{a+b}{2}\right)+\frac{1}{2}\cdot\left(\int_a^b f''\left(\frac{x+b}{2}\right)\cdot\frac{1}{b-a}dx\right) \\ &\quad\cdot\left(\int_a^b\left(\frac{b-x}{2}\right)^2\cdot\frac{1}{b-a}dx\right) \\ &= f\left(\frac{a+b}{2}\right)+\left(\frac{b-a}{12}\right)\cdot\left(f'(b)-f'\left(\frac{a+b}{2}\right)\right).\end{aligned}$$

The proof of (4.7) in part (a) is complete. ∎

Next, let's consider bounds for $\int_a^b f(x)h(x)dx - f\left(\int_a^b x\cdot h(x)dx\right) = D$. From Theorem 2.5, we obtained in (2.18)

$$\begin{aligned}h_1=\sum_{j=1}^{n-1}W_jp_j\left(1-\frac{p_j}{R_j}\right) &\leq \sum_{i=1}^{n}p_if(x_i)-f\left(\sum_{i=1}^{n}p_ix_i\right) \\ &\leq \sum_{j=1}^{n-1}V_jp_j\left(1-\frac{p_j}{R_j}\right)=h_2\end{aligned}$$

or its reversal (2.20). We shall now show that the bound h_2 is a 'best possible' bound as $n\to\infty$ when considering $p_i = H(x) - H(x_{i-1})$. Even if f'' and $f^{(3)}$ are not of one sign, it will be proven that h_2 approaches D as $n\to\infty$, so that the integral or continuous analogue for h_2 will be an exact representation. Thus h_2 behaves like a Riemann sum for D. We have the following theorem.

Theorem 4.3. *Suppose f' and h are continuous on $[a,b]$. Let*

$$a_1(x) = f(x+g(x)) - f(x) - g(x)f'(x) \tag{4.9}$$

and

$$a_2(x) = g(x)f'(x+g(x)) - f(x+g(x)) + f(x)\,. \tag{4.10}$$

Then

a)

$$D \equiv \int_a^b f(x)h(x)dx - f\left(\int_a^b xh(x)dx\right) = \int_a^b a_2(x)h(x)dx \equiv H_2\,. \tag{4.11}$$

b) *If $f^{(3)}(x)$ is continuous on $[a,b]$, $f''(x)\geq 0$ and $f^{(3)}(x)\geq 0$ on $[a,b]$, then*

$$H_1 \equiv \int_a^b a_1(x)h(x)dx \leq D = \int_a^b a_2(x)h(x)dx \equiv H_2\,. \tag{4.12}$$

c) *If $f^{(3)}(x)$ is continuous on $[a,b]$, $f''(x)\geq 0$ on $[a,b]$ and $f^{(3)}(x)\leq 0$ on $[a,b]$, then*

$$H_2 = D \leq H_1\,. \tag{4.13}$$

d) *Under the assumptions of part b), H_1 is a better lower bound for D than is L_1 given in Theorem 3.1, that is, $L_1 \leq H_1$.*

e) *Under the assumptions of part c), H_1 is a better upper bound for D than is U_1 given in Theorem 3.1, that is, $U_1 \geq H_1$.*

Proof of (a). Recall that $g(x) = \frac{\int_x^b (t-x)h(t)dt}{\int_x^b h(t)dt}$, upon differentiation of $g(x)$ using Leibnitz's rule for differentiating an integral, we obtain

$$g'(x) = \frac{h(x)g(x)}{\bar{H}(x)} - 1\,, \tag{4.14}$$

where $\bar{H}(x) = \int_x^b h(t)dt$. Thus

$$\begin{aligned} h_2 &= \int_a^b a_2(x)h(x)dx = \int_a^b (g(x)f'(x+g(x)) - f(x+g(x)) + f(x))h(x)dx \\ &= \int_a^b g(x)f'(x+g(x))h(x)dx - \int_a^b f(x+g(x))h(x)dx + \int_a^b f(x)h(x)dx\,. \end{aligned}$$

Making the substitution $w = w(x) = x + g(x)$ and using (4.14) to obtain $w'(x)$, we obtain

$$h_2 = \int_a^b f'(w(x))w'(x)\bar{H}(x)dx - \int_a^b f(x+g(x))h(x)dx + \int_a^b f(x)h(x)dx\,.$$

Integrating by parts, and using $\bar{H}(b) = 0$, $\bar{H}(a) = 1$, we obtain

$$\begin{aligned} h_2 = -\bar{H}(a)f(w(a)) + \left(\int_a^b f(w(x))h(x)dx - \int_a^b f(w(x)h(x)dx)\right) \\ + \int_a^b f(x)h(x)dx\,, \end{aligned} \tag{4.15}$$

upon application of (4.14) above. Since $w(a) = a + g(a) = \int_a^b xh(x)dx$, we obtain

$$h_2 = -f\left(\int_a^b x\cdot h(x)dx\right) + \int_a^b f(x)h(x)dx\,, \tag{4.16}$$

since the expression in parentheses in (4.15) equals zero. Thus, $h_2 = D$ and part (a) is proven.

Parts (b) and (c) explain why the bound h_2 is so good for larger n when bounding $\sum_{i=1}^n p_i f(x_i) - f\left(\sum_{i=1}^n p_i x_i\right)$ and has been verified in many numerical comparisons done. ∎

Proof of (b). Since the proof of the left half of (4.12) is very similar to the proof of Theorem 3.1 given earlier, we merely sketch its proof. Now choosing the same partition of $[a, b]$ as used for Theorem 3.1, we have

$$h_1 = \sum_{j=1}^{n-1} W_j p_j - \sum_{j=1}^{n-1} W_j \frac{p_j^2}{R_j}\,,$$

where from (2.19).

$$W_j = f(x_j^*) - f(x_j) - (x_j^* - x_j)f'(x_j) \approx f(x_j + g(x_j)) - f(x_j) - f(x_j)f'(x_j)\,,$$

since $x_j^* \approx x_j + g(x_j)$ was shown in Lemma 3.2 earlier. Also, $\sum_{j=1}^{n-1} W_j \frac{p_j^2}{R_j}$ can be shown to be $O(\Delta x)$. Now $\sum_{j=1}^{n-1} W_j p_j$ is a Riemann sum for $D = \int_a^b a_2(x)h(x)dx$. Similarly, for a sketch of a proof for part (c).

The proofs of parts (d) and (e) are similar to the proof of Theorem 3.1, part (a) given earlier and are omitted. ∎

Remark 4.2. The results given on bounds for the Jensen's gap in the continuous case on an interval $[a, b]$ are easily extended to certain types of improper integrals. For many applications in applied probability, the choice $a = 0$, $b = \infty$ is very important. For example, if $a = 0$ and $b = \infty$, Theorems 3.1 and 4.3 are valid in this case provided the improper integrals exist, and f'' is continuous on $[0, \infty)$.

For purposes of comparison, let's determine the values of L_1, U_1 in Theorem 3.1 and H_1 and H_2 in Theorem 3.1 and compare them to the lower bound given in Walker [28] for the case where $f(x) = e^{x/2}$ and $h(x) = e^{-x}$, $x > 0$. As discussed in Remark 4.2, Theorem 3.1 is valid for improper integrals as well since convergence of all integrals holds. These choices for $f(x)$ and $h(x)$ were also used in [28]. There, the author obtained $D = 2.00$ and Walker's lower bound on D was given as 1.920. The new lower bounds on D (using $a = 0$, $b = \infty$ in Theorems 3.1 and 4.3, per Remark 4.2), are $L_1 = 1.899$ and $H_1 = 1.946$. Thus, H_1 improves on the lower bound of Walker [28], but L_1 does not. However, Theorems 3.1 and 4.3 provide upper bounds as well. They are $U_1 = 2.062$ and $H_2 = 2.000$. Per Theorem 4.3, part (b), H_2 is exact in this case. The computer algebra package 'MAPLE' easily computed all necessary integrals. The bounds of Walker [28] assumes $f(x)$ has a power series representation of the form $f(S) = \sum_{n=0}^{\infty} \omega_n S^n$ for $0 < S < R$, where $R > 0$, and assumes $f(x)$ is convex. The new bounds L_1 and U_1 make no convexity assumption and require only continuity of $f''(x)$. The H_1 and H_2 bounds do, however, require convexity and either 3-convexity or 3-concavity. The bounds of Walker [28] are given in terms of an arbitrary probability measure on $(0, \infty)$, so it is valid, in particular, in the discrete and continuous cases; one does not need separate formulas for these two cases, unlike the bounds of Theorem 4.3 given in this paper. However, the new bounds presented in this paper can handle cases that previously proposed bounds can not. Moreover, they lead to extensions of various well-known results, such as Levinson's inequalities and other types of convexity, such as convexity in the geometric and harmonic mean. They also have applications to applied probability and reliability theory some of which was discussed in [21].

Next, we demonstrate how Theorem 4.2 can be used, along with Theorem 4.3 to get more inequalities of Hermite-Hadamard type for 3-convex/concave functions. Note that convexity of f is not required, just convexity of f'.

Theorem 4.4. *Let $f^{(3)}$ be continuous on $[a, b]$. If $f^{(3)}(x) \geq 0$ on $[a, b]$, then*

$$\frac{\int_a^b f(x)dx}{b-a} - f\left(\frac{a+b}{2}\right) \leq -f\left(\frac{a+b}{2}\right) + f\left(\frac{a+3b}{4}\right) + \frac{b-a}{24}\left(f'(b) - f'\left(\frac{a+3b}{4}\right)\right)$$

$$+\frac{2}{3}f\left(\frac{3a+b}{4}\right) - \frac{2}{3}f\left(\frac{3a+5b}{8}\right) - \frac{b-a}{24}\left(f'\left(\frac{3a+5b}{8}\right) - f'\left(\frac{3a+b}{4}\right)\right). \tag{4.17}$$

If $f^{(3)}(x) \leq 0$ on $[a, b]$, then the reverse inequality holds.

Proof. We shall prove only the 3-convexity case. Theorem 4.3, part (a) gives, with $h(x) = \frac{1}{b-a}$ and $g(x) = \frac{b-x}{2}$, $a \leq x \leq b$

$$\frac{\int_a^b f(x)dx}{b-a} - f\left(\frac{a+b}{2}\right) = \int_a^b \left[\left(\frac{b-x}{2}\right) f'\left(\frac{b+x}{2}\right) - f\left(\frac{b+x}{2}\right) + f(x)\right] \cdot \frac{1}{b-a} dx. \tag{4.18}$$

Now

$$f\left(\frac{b+x}{2}\right) - f(x) = \int_x^{\frac{b+x}{2}} f'(t)dt.$$

Since f' is convex, the Hermite-Hadamard inequality applied to f' gives

$$f\left(\frac{b+x}{2}\right) - f(x) \geq f'\left(\frac{b+3x}{4}\right) \cdot \left(\frac{b-x}{2}\right) . \tag{4.19}$$

Then (4.18) and (4.19) and integration by parts gives

$$\frac{\int_a^b f(x)dx}{b-a} - f\left(\frac{a+b}{2}\right) \leq -f\left(\frac{a+b}{2}\right) + \frac{\int_a^b f\left(\frac{b+x}{2}\right)dx}{b-a} - \frac{\int_a^b f'\left(\frac{b+3x}{4}\right)\cdot\left(\frac{b-x}{2}\right)dx}{b-x} . \tag{4.20}$$

Also,

$$\frac{\int_a^b f'\left(\frac{b+3x}{4}\right)\cdot\left(\frac{b-x}{2}\right)dx}{b-a} = -\frac{2}{3}f\left(\frac{3a+b}{4}\right) + \frac{2}{3}\int_a^b f\left(\frac{b+3x}{4}\right)dx . \tag{4.21}$$

Applying Theorem 4.2, part (a) to the functions $f_1(x) = f\left(\frac{b+x}{2}\right)$ and $f_2(x) = f\left(\frac{b+3x}{4}\right)$ in (4.20), we obtain, since f_1 and f_2 are 3-convex,

$$\frac{\int_a^b f\left(\frac{b+x}{2}\right)dx}{b-a} \leq f\left(\frac{a+3b}{4}\right) + \left(\frac{b-a}{24}\right)\left(f'(b) - f'\left(\frac{a+3b}{4}\right)\right) \tag{4.22}$$

and

$$\int_a^b f\left(\frac{b+3x}{4}\right)dx \geq f\left(\frac{3a+5b}{8}\right) + \left(\frac{b-a}{16}\right)\left(f'\left(\frac{3a+5b}{8}\right) - f'\left(\frac{3a+b}{4}\right)\right) . \tag{4.23}$$

From (4.21)–(4.23), we obtain upon addition and subtraction, the desired result. ∎

Remark 4.3. Numerical investigations suggest (4.17) improves on (4.7) in Theorem 4.2, but no proof has been found. Also, the bounds given in Theorem 4.4 are quite good. If $f(x)$ is 3-convex, then we may apply the above theorem to $f^*(x) = f(a+b-x)$ instead in the obvious way to obtain a lower bound for $\frac{\int_a^b f(x)dx}{b-a} - f\left(\frac{a+b}{2}\right)$, (or an upper bound for this difference, if f is 3-concave). Many more such inequalities of Hermite-Hadamard type can be obtained using the methods given in this paper and will be discussed in a forthcoming paper.

Next, let's consider some applications to reliability theory and applied probaiblity. First, let's obtain more convenient representations of the Jensen gap $D = \int_a^b f(x)h(x)dx - f\left(\int_a^b xh(x)dx\right)$ in the form of (possibly) infinite series. Assume for the moment that $f(x)$ is real analytic on some open set containing $[a, b)$. In reliability theory, $0 \leq a < b \leq \infty$, so we assume this here also. Then by Theorem 4.3, part (a), with the extension to a possibly improper integral with $g(x) < \infty$ on $[a, b)$, we obtain

$$D = \int_a^b [g(x)f'(x+g(x)) - f(x+g(x)) + f(x)] \cdot h(x)dx . \tag{4.24}$$

Since f has a Taylor series expansion at x, we may write

$$f(x+g(x)) = f(x) + \sum_{j=1}^{\infty} \frac{f^{(j)}(x)}{j!}(g(x))^j , \tag{4.25}$$

assuming that the interval of convergence centered at x has radius of convergence greater than $g(x)$. (This would be true, for example, if f is a polynomial or exponential function.) Similarly,

$$f'(x+g(x)) = \sum_{j=1}^{\infty} \frac{f^{(j)}(x)}{(j-1)!}(g(x))^{j-1} . \tag{4.26}$$

From (4.24)–(4.26) above, we obtain

$$D = \sum_{j=2}^{\infty} \int_a^b \left(\frac{j-1}{j!} \right) \cdot f^{(j)}(x) \cdot (g(x))^j h(x) dx \,. \tag{4.27}$$

Equation (4.27) gives a nice and convenient representation of the Jensen gap D in terms of the mean residual life function $g(x)$ for purposes of obtaining numerous new inequalities in reliability theory. It can also be used to obtain many more new inequalities of Hermite-Hadamard type. Note that (4.27) is a finite series for polynomial choices of $f(x)$ and is an easily manageable infinite series for exponential functions, two very important special choices in applied probability and reliability theory, since moments about the origin and moment probability generating functions are important in these areas.

In [18], the concept of mean residual life (MRL) function and decreasing mean residual life (DMRL) is discussed. The theoretical aspects, importance, and wide range of applications of these are surveyed and discussed. In [4], the increasing failure rate (IFR), increasing failure rate average (IFRA) and new better than used in expectation (NBUE) nonparametric classes of life distributions are discussed. See the above two references for definitions of these subclasses, the IFR is the smallest, with IFR → IFRA → NBUE. Also, IFR → DMRL → NBUE. It is also well known that if X is a lifetime random variable and is a member of any of the above four subclasses, then the n^{th} moments about the origin of X, $\mu_n' = E(X^n) = \int_0^\infty x^n h(x) dx$, $n = 0, 1, 2, \ldots$, satisfy the inequality

$$\mu_n' \le (n!) \cdot \mu^n \,, \tag{4.28}$$

where $\mu = \mu_1' = E(X)$. This bound is sharp, since the exponential distribution with $h(x) = \lambda e^{-\lambda x}$, $x > 0$, $\lambda > 0$, is a member of all four subclasses. See [4], p. 116, for example.

First, let's give another proof of (4.28) for the NBUE class. In the sequel, let $f(x) = x^n$, $n = 1, 2, 3, \ldots$. We assume without loss of generality, that $a = 0$, $b = \infty$ below, since we may take $g(x) = 0$ for $x > b$, if b is finite.

Theorem 4.5. *Suppose a random variable X is NBUE, that is, $g(x) \le \mu$ on $[0, \infty)$. Then (4.28) holds, that is, $\mu_n' \le (n!)\mu^n$.*

Proof. The result is trivially true for $n = 1$. We shall use mathematical induction. Supposing that $\mu_n' \le n!\mu^n$ holds for some positive integer n, (4.27) gives

$$\begin{aligned} \mu_{n+1}' &= E(X^{n+1}) = \sum_{j=2}^{n+1} \int_0^\infty \left(\frac{j-1}{j!} \right) f^{(j)}(x) (g(x))^j h(x) dx + \mu^{n+1} \\ &\le \sum_{j=2}^{n+1} \int_0^\infty \frac{j-1}{j!} \frac{(n+1)!}{(n+1-j)!} x^{n+1-j} \mu^j h(x) dx + \mu^{n+1} \,. \end{aligned}$$

But by the induction hypothesis, $\mu_{n+1-j}' \le (n+1-j)!\mu^{n+1-j}$, so

$$\begin{aligned} \mu_{n+1}' &\le \sum_{j=2}^{n+1} \left(\frac{j-1}{j!} \right) \mu^j \frac{(n+1)!}{(n+1-j)!} \cdot (n+1-j)!\mu^{n+1-j} + \mu^{n+1} \\ &= (n+1)!\mu^{n+1} \sum_{j=2}^{n+1} \frac{j-1}{j!} = (n+1)!\mu^{n+1} \left(1 - \frac{1}{(n+1)!} \right) + \mu^{n+1} \\ &= (n+1)!\mu^{n+1} \,. \end{aligned}$$

This completes the proof. ∎

Inequality (4.28) gives bounds on μ_n' in terms of only μ. If the variance of X or, equivalently, μ_2' is available as well, then it is clear that we can improve on these bounds, using representation (4.27) in a recursive fashion. In a future paper, we shall discuss the details. Also, lower bounds for D can be obtained as well.

The above theorem can be generalized to a more general class than the NBUE. If there exists a constant M with $g(x) \leq M$ on $[a, b)$, where $M \geq \mu$, then it is easily shown that $\mu_n' \leq n!M^{n+1}$.

Remark 4.4. Either representation (4.27) or Theorem 3.1 given earlier quickly yields the result

$$\sigma^2 = \text{Var}(X) = \int_a^b (g(x))^2 dH(x) = \int_a^b (g(x))^2 h(x)dx\,, \tag{4.29}$$

if $h(x)$ exists, where σ^2 denotes the variance of X. If this representation of the variance as the mean of the squared MRL function is already known, it is unknown by this author. In any case, representation (4.27) is of much more general use. Using (4.27), it can be shown, for example, that if X is NBUE, then $E(X^3) \leq 6\mu\sigma^2$, which is an improvement on (4.28) for $n = 3$, since $\sigma \leq \mu$ holds in the NBUE case. Bounds for $E(X^n)$ for $n \geq 4$ can also be obtained in a recursive manner using (4.27) in conjunction with the Chebychev-Gruss inequality which also improve on (4.28). We omit the details here.

More inequalities of type similar to those in [6], [7] and [8] can be obtained, if the equation for $h(x)$ is completely known, to obtain bounds on moment-generating functions and moments, in particular.

Finally, we discuss the extensions of results given in this paper to other kinds of convexity. In Niculescu [21], the following definition is given.

Definition. Let I, J be subintervals of $(0, \infty)$. Suppose that $t_1, t_2 \in I$ and $p \in (0, 1)$. Let $f : I \to J$. Then f is multiplicatively convex on I, if

$$f(t_1^p t_2^{1-p}) \leq f(t_1)^p f(t_2)^{1-p}\,. \tag{4.30}$$

This is a type of convexity according to the geometric mean, instead of the arithmetic mean. From (4.29), it follows that if $t_i \in I$, $p_i \in (0, 1)$ with $\sum_{i=1}^n p_i = 1$, then, if f is multiplicatively convex on I, then

$$f\left(\prod_{i=1}^n t_i^{p_i}\right) \leq \prod_{i=1}^n f(t_i)^{p_i}\,. \tag{4.31}$$

As discussed in [21], if $f : I \to (0, \infty)$ is a multiplicatively convex function and if we define F by $F = \log \circ f \circ \exp : \log(I) \to \mathbb{R}$, where '$\circ$' denotes functional composition, then F is a convex function. We can rewrite (4.31) in terms of F as:

$$F\left(\sum_{i=1}^n p_i \text{Log} t_i\right) \leq \sum_{i=1}^n p_i F(\text{Log} t_i)\,.,$$

Let D_{GM} denote the 'Jensen gap'

$$\begin{aligned} D_{GM} &= \prod_{i=1}^n f(t_i)^{p_i} - f\left(\prod_{i=1}^n t_i^{p_i}\right) \\ &= \sum_{i=1}^n p_i F(x_i) - F\left(\sum_{i=1}^n p_i x_i\right) \end{aligned}$$

where $x_i = \text{Log} t_i$, $i = 1, 2, \ldots, n$. Assuming $x_1 \leq x_2 \leq \cdots \leq x_n$, we can apply Theorems 2.3 and 2.5 and corollaries to obtain bounds on D_{GM}. We may also apply Theorems 3.1 and 4.3 to the continuous analogue of D_{GM} as well, except we would substitute $F(x)$ for $f(x)$ throughout.

REFERENCES

[1] A. ABRAMOVICH, J. BARIĆ, and J. PEČARIĆ, Fejer and Hermite-Hadamard type inequalities for superquadratic functions, *J. Math. Anal. Appl.*, **344** (2008), pp.1048–1056.

[2] M. K. BAKULA, J. PEČARIĆ, and J. PERIĆ, On the converse Jensen inequality, *Appl. Math. Comp.*, **218** (2012), pp. 6566–6575.

[3] A. BARANI, S. BARANI and S. S. DRAGOMIR, Refinements of Hermite-Hadamard inequalities for functions when a power of the absolute value of the second derivative is P-convex, *J. Appl. Math.*, **2012**, Art. ID 615737, p. 10.

[4] R. E. BARLOW and F. PROSCHAN, *Statistical Theory of Reliability and Life Testing*, 1981.

[5] R. L. BURDEN and J. D. FAIRES, *Numerical Analysis*, Sixth Ed., Brooks/Cole, 1997.

[6] P. CERONE, Special functions: Approximations and Bounds, *Appl. Anal. and Discr. Math.*, **1** (2007), pp.72–91.

[7] P. CERONE, On an identity for the Chebychev functional and some ramifications, *J. Ineq. Pure and Appl. Math.*, **3** (2002), pp. 1–15.

[8] P. CERONE and S. S. DRAGOMIR, On some inequalities arising from Montegormer's identity, *J. Cmput. Anal. Appl.*, **5** (2003), No. 4, pp. 341–367.

[9] D. CASTORELLI and R. SPIGLER, How sharp is the Jensen inequality? *J. Inequal. Appl.*, **69** (2015) , p. 10.

[10] S. S. DRAGOMIR and C. J. GOH, A counterpart of Jensen's discrete inequality for differentiable convex mappings and applications in information theory, *Mathl. Cmput. Modeling*, **24** (1996), No. 2, pp.1–11.

[11] S. S. DRAGOMIR and B. MOND, A refinement of Jensen's inequality for log-convex mappings and applications, *Bull. Allahabad Math. Soc.*, **10/11** (1995/1996), pp. 37–67.

[12] S. S. DRAGOMIR, A converse result for Jensen's inequality via Gruss inequality and applications in information theory, *Analele Univ. Oradea. Fasc. Math.*, **7** (1999-2000), p. 178–179.

[13] S. S. DRAGOMIR and C. E. M. PEARCE, *Selected Topics on Hermite-Hadamard Inequalities*, RGMIA Monographs, Victoria University, 2000.

[14] S. S. DRAGOMIR, On a converse of Jensen's inequality, *Univ. Beograd. Publ. Elektrotehn. Fak.*, **12** (2001), pp. 48–51.

[15] S. S. DRAGOMIR and F. P. SCARMOZZINO, A refinement of Jensen's discrete inequality for differentiable convex functions, *East Asian Math. J.*, **1** (2004), pp. 1–6.

[16] A. M. FINK and Z. PÁLES, What is Hadamard's inequality? *Appl. Anal. and Disc. Math.*, **1** (2007), pp 29–35.

[17] A. E. FRISSI and Z. LAREUCH, Jensen type inequalities for twice differentiable functions, *J. Nonlinear Sci. Appl.*, **5** (2012), pp. 350–356.

[18] F. GUESS and F. PROSCHAN *Mean Residual Life: Theory and Applications*, Handbook of Statistics, 05/1985.

[19] P. M. HUMMEL, The accuracy of linear interpolation, *Amer. Math. Monthly*, **53** (1946), pp. 364–366.

[20] D. S. MITRINOVIC, J. PEČARIĆ, A. M. FINK, *Classical and New Inequalities in Analysis (Mathematics in its Applications)*, Springer, 1992.

[21] C. P. NICULESCU, Convexity according to the geometric mean, *Math. Inequal. and Appl.*, **3** (2000), No. 2, pp. 155–167.

[22] C. P. NICULESCU, Old and new on the Hermite-Hadamard inequality, *Real Anal. Exchange*, **29** (2003/2004), No. 2, pp. 663–685.

[23] C. P. NICULESCU, The Hermite-Hadamard inequality for log-convex functions, *Nonlinear Anal.*, **75** (2012), pp. 662-669.

[24] J. PARK, Hermite-Hadamard-like type inequalities for twice differentiable convex functions, *Int. J. Math. Anal.*, **8** (2014), No. 59, pp. 2909–2926.

[25] J. E. PEČARIĆ, F. PROSCHAN, Y. L. TONG, *Convex Functions, Partial Orderings and Statistical Applications*, **Vol. 187**, Mathematics in Science and Engineering, Academic Press, 1992.

[26] F. QI, Z. L. WEI and Q. YANG, Generalizations and refinements of Hermite-Hadamard inequality, *Rocky Mountain J. Math.*, **35** (2005), No. 1, pp. 235–251.

[27] S. SIMIC, On an upper bound for Jensen's inequality, *JIPAM, J. Inequal. Pure and Appl. Math.*, **10** (2009), No. 2, Article 60, p. 5.

[28] S. G. WALKER, On a lower bound for the Jensen inequality, *SIAM J. Math. Anal.*, **46** (2014), No. 5, pp. 3151–3157.

[29] S. ZLOBEC, Jensen's inequality for nonconvex functions, *Math. Commun.*, **9** (2004), No. 2, pp. 119–124.