Lehmer Transform and Its Theoretical Properties

Masoud Ataei, Shengyuan Chen and Xiaogang Wang
Department of Mathematics and Statistics
York University
Toronto, Ontario, Canada

Abstract

We propose a new class of transforms that we call Lehmer Transform which is motivated by the Lehmer mean function. The proposed Lehmer transform decomposes a function of a sample into their constituting statistical moments. Theoretical properties of the proposed transform are presented. This transform could be very useful to provide an alternative method in analyzing non-stationary signals such as brain wave EEG.

1 Introduction

In this paper, we introduce a new class of transforms referred to as the Lehmer Transform. The proposed transform decomposes a function of the data into the so-called Breve Moments, which in turn provides the means to construct some parametric families of nonparametric statistical models.

The Lehmer transform is motivated by the Lehmer mean function, named after the renowned number theorist Derrick H. Lehmer. However, the appealing properties of this function has not received much attention from statistics and machine learning communities in the past. As generalization of the power mean function and having connections to other important classes of means, the Lehmer mean function has received a lot of attentions in recent years. The function’s elementary properties like homogeneity, monotonicity and differentiability have been discussed in [2,3], whereas its more advanced properties like Schur-convexity, Schur harmonic convexity and Schur power convexity have been the focus of more recent studies [4,7,8]. Also, the inflection points of the function have been studied in [9], and the results concerning its possible connections to Gini and Toader means have been provided in [10,11,12,13,14,15]. Most application of the Lehmer mean
function rely mostly on its special cases. For instance, Terziyan [16] formulated a distance metric using the harmonic and contra-harmonic means, and evaluated its effectiveness in geographic information systems. In a similar manner, Somasundaram et al. [17] studied the disconnected graphs and their possible labeling schemes, and Sluciak [18] developed state-dependent consensus algorithms by resorting to some special cases of the Lehmer mean function. In addition, Gomes [19] successfully constructed a family of high-performance value at risk estimators through the use of this function and proved asymptotic normality of the constructed estimators.

2 Discrete Lehmer Transform

Let $X_1, X_2, \ldots$ denote a sequence of random variables defined on some given probability measure space $(\Omega, \mathcal{F}, \mathbb{P})$ such that each random variable takes its value from the same measurable space $(\mathcal{X}, \mathcal{B}_{\mathcal{X}})$. Also, let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$ denote a realization of some discrete-time stochastic process $\{X_n\}$ where no restrictions pertain to its underlying sampling procedure. For instance, $X_1, X_2, \ldots, X_n$ could be any combination of $n$ continuous- or discrete-type random variables which are not necessarily independent or identically distributed.

Let us further consider a generic framework to normalize the sample, in which every possible normalization of the vector-valued sample $X_n = (X_1, X_2, \ldots, X_n)^\top$ is viewed as a $(\mathcal{B}_{\mathcal{X}^n}, \mathcal{B}_{\mathcal{H}})$-measurable function obtained through composition of several, say $m$, maps; i.e. the function denoted by

$$
\hat{h}_m : \mathcal{X}^n \to \mathcal{H},
$$

and defined as

$$
\hat{h}_m(X_i) = (h_m \circ h_{m-1} \circ \cdots \circ h_2 \circ h_1)(X_i), \quad i = 1, 2, \ldots, n,
$$

could be considered as a normalization of $X_n$ where $\mathcal{H} \subset \mathbb{R}^n$. Throughout this work, we focus on a specific class of normalizations referred to as $m$-normalizations, which contains every possible $(\mathcal{B}_{\mathcal{X}^n}, \mathcal{B}_{\mathcal{H}})$-measurable function mapping the considered $X_n$ into a strictly positive sample.

**Definition 1 (m-normalization)** The $(\mathcal{B}_{\mathcal{X}^n}, \mathcal{B}_{\mathcal{H}})-$measurable function $\hat{h}_m : \mathcal{X}^n \to \mathcal{H}$ is said to be an $m$-normalization of $X_n$ if and only if its image satisfies $\mathcal{H} \subset \mathbb{R}_{\geq 0}$.
Definition 2 (Discrete Lehmer Transform) The discrete Lehmer transform denoted by 
$$\mathcal{L} : \mathbb{R} \to \mathbb{R}_{>0} \, ,$$

is a $$(\mathcal{B}_{\mathbb{R}}, \mathcal{B}_{\mathbb{R}_{>0}})$$–measurable function defined as

$$\mathcal{L}^n\{\hat{h}_m\}(\check{s}) = \begin{cases} \max_{i=1,...,n} \{\hat{h}_m(X_i)\} , & \text{if } \check{s} = \infty , \\
\frac{\sum_{i=1}^{n} \hat{h}_m(X_i)}{\sum_{i=1}^{n} \check{s}^{-1} \hat{h}_m(X_i)} , & \text{if } \check{s} \in (-\infty, \infty) , \\
\min_{i=1,...,n} \{\hat{h}_m(X_i)\} , & \text{if } \check{s} = -\infty , \end{cases} \quad (1)$$

where $\hat{h}_m$ represents some $m$-normalization of the sample.

According to Def. 2, the discrete Lehmer transform $\mathcal{L}^n\{\hat{h}_m\}(\check{s})$ maps every point $\check{s} \in \mathbb{R}$ into some statistic contained in the close interval

$$\left[ \min_{i=1,...,n} \{\hat{h}_m(X_i)\} , \max_{i=1,...,n} \{\hat{h}_m(X_i)\} \right] ,$$

where the points $\check{s}$ (read $s$-breve) are referred to as breve moments of $X_n$ under $m$-normalization $\hat{h}_m$. Table 1 reports some of the widely-encountered breve moments in data analysis. Furthermore, for every statistic

$$T : \mathcal{X}^n \to \left[ \min_{i=1,...,n} \{\hat{h}_m(X_i)\} , \max_{i=1,...,n} \{\hat{h}_m(X_i)\} \right] ,$$

the following inverse image

$$\mathcal{L}^{-1} : \mathbb{R}_{>0} \to \mathbb{R} ,$$

provides a set that contains the associated breve moment(s) of the sample under $m$-normalization $\hat{h}_m$ such that

$$\mathcal{L}^{-1}\{\hat{h}_m\}(T) = \{ \check{s} \in \mathbb{R} : \mathcal{L}^n\{\hat{h}_m\}(\check{s}) = T \} .$$

Lemma 1 For every injective $\hat{h}_m$, the sufficient condition for $\mathcal{L}$ to be
### Table 1: My caption

| Breve moment domain | Sample domain                  |
|---------------------|---------------------------------|
| $-\infty$           | Minimum ($n \geq 1$)            |
| 0                   | Harmonic mean ($n \geq 1$)     |
| $1/2$               | Geometric mean ($n = 2$)       |
| 1                   | Arithmetic mean ($n \geq 1$)   |
| 2                   | Contra-harmonic mean ($n \geq 1$) |
| $\infty$           | Maximum ($n \geq 1$)           |

(i) a constant function is given by
\[
\{ \exists \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies \hat{h}_m(X_i) \neq \hat{h}_m(X_j) \} \xrightarrow{a.s.} 0; \quad (2)
\]

(ii) a monotone increasing function is given by
\[
\{ \exists \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies \hat{h}_m(X_i) \neq \hat{h}_m(X_j) \} \xrightarrow{a.s.} 1; \quad (3)
\]

(iii) a strictly monotone increasing function is given by
\[
\{ \forall \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies \hat{h}_m(X_i) \neq \hat{h}_m(X_j) \} \xrightarrow{a.s.} 1. \quad (4)
\]

**Lemma 2** For every injective $\hat{h}_m$, the sufficient condition for $\hat{L}$ to be

(i) a constant function is given by
\[
P[ \exists \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies X_i \neq X_j ] = 0; \quad (5)
\]

(ii) a monotone increasing function is given by
\[
P[ \exists \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies X_i \neq X_j ] = 1; \quad (6)
\]

(iii) a strictly monotone increasing function is given by
\[
P[ \forall \, i, j \in \{1, 2, \ldots, n\} : i \neq j \implies X_i \neq X_j ] = 1. \quad (7)
\]

**Lemma 3** The discrete Lehmer transform $\hat{L}$ is differentiable at $\hat{s}_0$ in $P$-probability, and its $n$th order derivative is given by
\[
\frac{\partial^n}{\partial \hat{s}^n} \hat{L} = \hat{L} \left[ \sum_{k=0}^{n} \frac{1}{k!} \left( \sum_{j=0}^{k} (-1)^j \binom{k}{j} \Lambda^j \frac{\partial^n}{\partial \hat{s}^n} \Lambda^{k-j} \right) \right],
\]
where

\[ \Lambda := \log \hat{L}. \]

**Theorem 4 (Inversion Theorem)** For a sample \( X_n \) having a (strictly) monotone increasing \( \hat{L} \), its discrete inverse Lehmer transform is a \((\mathcal{B}_{\mathbb{R}^+}, \mathcal{B}_\mathbb{R})\)–measurable function defined as

\[
\hat{L}_n^{-1}\{\hat{h}_m\}(T) = \begin{cases} 
\infty, & \text{if } T = \max_{i=1,\ldots,n} \left\{ \hat{h}_m(X_i) \right\}, \\
g(s), & \text{if } T \in \left( \min_{i=1,\ldots,n} \left\{ \hat{h}_m(X_i) \right\}, \max_{i=1,\ldots,n} \left\{ \hat{h}_m(X_i) \right\} \right), \\
-\infty, & \text{if } T = \min_{i=1,\ldots,n} \left\{ \hat{h}_m(X_i) \right\},
\end{cases}
\]

where

\[
g(s) = s_0 + \sum_{k=1}^{\infty} \left( T - \hat{L}_n\{\hat{h}_m\}(s_0) \right)^k \frac{k!}{k} \left[ \lim_{s \to s_0} \frac{\partial^{k-1}}{\partial s^{k-1}} \left( \frac{s - s_0}{\hat{L}_n\{\hat{h}_m\}(s) - \hat{L}_n\{\hat{h}_m\}(s_0)} \right) \right].
\]

for some \( s_0 \in \mathbb{R} \).

**Theorem 5** In a family \( \mathcal{F} \) of linear transformations of some (strictly) monotone increasing \( \hat{L} \) where

\[
\mathcal{F} = \left\{ a + b\hat{L} : a \in \mathbb{R}, b \in \mathbb{R}^+ \right\},
\]

there exists one and only one \((\mathcal{B}_{\mathbb{R}}, \mathcal{B}_\mathbb{R})\)–measurable function \( F(s) \in \mathcal{F} \) that satisfies the condition for being a distribution.

**Corollary 6** Let \( \{A_k\}_{k \geq 1} \) and \( \{B_k\}_{k \geq 1} \) be sequences of \((\mathcal{B}_\mathbb{X}, \mathcal{B}_\mathbb{R})\)– and \((\mathcal{B}_\mathbb{X}, \mathcal{B}_{\mathbb{R}^+})\)–measurable functions such that \( A_k \overset{P}{\to} A \) and \( B_k \overset{P}{\to} B \), respectively. Then, for every (strictly) monotone increasing \( \hat{L} \), the following \((\mathcal{B}_\mathbb{R}, \mathcal{B}_\mathbb{R})\)–measurable function

\[ F(s) = A + B\hat{L}(s), \]

is a distribution if

\[ A \overset{P}{\to} \frac{\hat{L}(-\infty)}{\hat{L}(\infty) - \hat{L}(-\infty)}, \]

and

\[ B \overset{P}{\to} \frac{1}{\hat{L}(\infty) - \hat{L}(-\infty)}. \]
Corollary 7 For every (strictly) monotone increasing $\mathcal{L}$, the following $(\mathcal{B}_{\mathbb{R}},\mathcal{B}_{\mathbb{R}})$–measurable function

$$F(\hat{s}) = \mathcal{L}(\hat{s}),$$

is a distribution if and only if

$$\mathcal{L}(\infty) \xrightarrow{P} W_0(e),$$

and

$$\mathcal{L}(-\infty) \xrightarrow{P} W_0(0).$$

Theorem 8 Assume the discrete Lehmer transform of a sample $X_n$ satisfies the conditions $\mathcal{L}(\infty) \xrightarrow{P} W_0(e)$ and $\mathcal{L}(-\infty) \xrightarrow{P} W_0(0)$. Then, a probability density function for the extended real-valued random variable $\hat{S}$ could be derived using

$$f_{\hat{S}}(\hat{s}; \alpha, \beta) = \left( \frac{\partial}{\partial \hat{s}} \mathcal{L}(\hat{s}) \right) I_{(-\infty, \infty)}(\hat{s})$$

$$= \left[ \sum_{i=1}^{n} \sum_{k=i+1}^{n} \left( \hat{h}_m(X_i) - \hat{h}_m(X_k) \right) \left( \log \hat{h}_m(X_i) - \log \hat{h}_m(X_k) \right) \left( \hat{h}_m(X_i) \hat{h}_m(X_k) \right)^{\hat{s}-1} \right]$$

$$\times \left( \sum_{i=1}^{n} \hat{s}_m^{\hat{s}-1}(X_i) \right)^{-2} I_{(-\infty, \infty)}(\hat{s}),$$

for every $\alpha \in (0, 1]$ and $\beta \in \mathbb{R}_{>0}$. 
Theorem 9 In a family $F$ of non-linear transformations of some (strictly) monotone increasing $\hat{L}$ where

$$F = \left\{ a + b \hat{L}^{1/\alpha} e^{\beta \hat{L}} : a \in \mathbb{R}, b \in \mathbb{R}_{>0}, \alpha \in (0, 1], \beta \in \mathbb{R}_{>0} \right\}, \quad (12)$$

there exists one and only one $(\mathcal{B}_{X}, \mathcal{B}_{Z})$—measurable function $F(\hat{s}) \in F$ that satisfies the condition for being a distribution.

Corollary 10 Let $\{A_k\}_{k \geq 1}, \{B_k\}_{k \geq 1}$ be sequences of $(\mathcal{B}_{X}, \mathcal{B}_{Z})$— and $(\mathcal{B}_{X}, \mathcal{B}_{Z_{>0}})$—measurable functions such that $A_k \overset{P}{\to} A$ and $B_k \overset{P}{\to} B$, respectively. Then, for every (strictly) monotone increasing $\hat{L}$, the following $(\mathcal{B}_{X}, \mathcal{B}_{Z})$—measurable function

$$F(\hat{s}) = A + B \left( \hat{L}(\hat{s}) \right)^{1/\alpha} e^{\beta \hat{L}(\hat{s})},$$

is a distribution if

$$A \overset{P}{\to} - \frac{\exp \left\{ \frac{1}{\alpha} \log \hat{L}(\infty) + \beta \hat{L}(\infty) \right\}}{\exp \left\{ \frac{1}{\alpha} \log \hat{L}(\infty) + \beta \hat{L}(\infty) \right\} - \exp \left\{ \frac{1}{\alpha} \log \hat{L}(\infty) + \beta \hat{L}(\infty) \right\}},$$

and

$$B \overset{P}{\to} \frac{1}{\exp \left\{ \frac{1}{\alpha} \log \hat{L}(\infty) + \beta \hat{L}(\infty) \right\} - \exp \left\{ \frac{1}{\alpha} \log \hat{L}(\infty) + \beta \hat{L}(\infty) \right\}},$$

where $\alpha \in (0, 1]$ and $\beta \in \mathbb{R}_{>0}$ denote some given parameters.

Corollary 11 For every (strictly) monotone increasing $\hat{L}$, the following $(\mathcal{B}_{X}, \mathcal{B}_{Z})$—measurable function

$$F(\hat{s}) = \left( \hat{L}(\hat{s}) \right)^{1/\alpha} e^{\beta \hat{L}(\hat{s})},$$

is a distribution if and only if

$$\hat{L}(\infty) \overset{P}{\to} \frac{W_0(\alpha \beta)}{\alpha \beta},$$

and

$$\hat{L}(\infty) \overset{P}{\to} W_0(0),$$

where $\alpha \in (0, 1]$ and $\beta \in \mathbb{R}_{>0}$ denote some given parameters.
Theorem 12 (Breve Distribution) Assume the discrete Lehmer transform of a sample $X_n$ satisfies the conditions $\hat{L}(\infty) \overset{P}{\to} W_0(\alpha\beta)$ and $\hat{L}(-\infty) \overset{P}{\to} W_0(0)$. Then, an extended real-valued random variable $\hat{S}$ is said to follow a Breve distribution if its probability density function is given by

$$f_{\hat{S}}(\hat{s}; \alpha, \beta) = \left( \frac{\partial}{\partial \hat{s}} \hat{L}(\hat{s}) \right) \frac{1}{\alpha} \exp \{ \beta \hat{L}(\hat{s}) \} \left( 1 + \alpha \beta \hat{L}(\hat{s}) \right)^{\frac{1}{\alpha} - 1} \mathbb{I}_{(-\infty, \infty)}(\hat{s})$$

for every $\alpha \in (0, 1]$ and $\beta \in \mathbb{R}_{>0}$. A Breve random variable having the probability density function (10) is denoted by $\hat{S} \sim \text{Breve}(\hat{s}; \alpha, \beta)$.

Theorem 13 The extreme points of (15) satisfy the following relation

$$\hat{L}(\hat{s}) = \frac{1}{\alpha\beta} W_0 \left( \alpha\beta \left( \frac{C_1}{\alpha} (\hat{s} + C_2) \right)^{\alpha} \right), \quad (14)$$

for some constants $C_1$ and $C_2$ such that

$$\hat{L}(\infty) \overset{P}{\to} \frac{W_0(\alpha\beta)}{\alpha\beta},$$

and

$$\hat{L}(-\infty) \overset{P}{\to} W_0(0).$$

Theorem 14 (Log-Breve Distribution) Assume the discrete Lehmer transform of a sample $X_n$ satisfies the conditions $\hat{L}(\infty) \overset{P}{\to} \exp\{\frac{W_0(\alpha\beta)}{\alpha\beta}\}$ and $\hat{L}(-\infty) \overset{P}{\to} \exp\{W_0(0)\}$. Then, an extended real-valued random variable $\hat{S}$
is said to follow a Log-Breve distribution if its probability density function is given by

\[
f_{\tilde{S}}(\tilde{s}; \alpha, \beta) = \left[ \sum_{i=1}^{n} \sum_{k=i+1}^{n} \left( \hat{h}_m(X_i) - \hat{h}_m(X_k) \right) \left( \log \hat{h}_m(X_i) - \log \hat{h}_m(X_k) \right) \left( \hat{h}_m(X_i) \hat{h}_m(X_k) \right)^{\tilde{s} - 1} \right] \times \left( \sum_{i=1}^{n} \hat{h}_{m}^{\tilde{s}-1}(X_i) \right)^{-2} \frac{1}{\alpha} \exp \left\{ \frac{\beta \sum_{i=1}^{n} \hat{h}_m^{\tilde{s}}(X_i)}{\sum_{i=1}^{n} \hat{h}_m^{\tilde{s}-1}(X_i)} \right\} \left( 1 + \alpha \beta \frac{\sum_{i=1}^{n} \hat{h}_m^{\tilde{s}}(X_i)}{\sum_{i=1}^{n} \hat{h}_m^{\tilde{s}-1}(X_i)} \right) \left( \sum_{i=1}^{n} \hat{h}_m^{\tilde{s}}(X_i) \right) \times \left( \exp \left\{ \frac{W_0(\alpha \beta)}{\alpha^2 \beta} + \beta \exp \left\{ \frac{W_0(\alpha \beta)}{\alpha \beta} \right\} \right\} - \exp \{ \beta \} \right)^{-1} \mathbb{I}_{(-\infty, \infty)}(\tilde{s}) \]

for every \( \alpha \in (0, 1] \) and \( \beta \in \mathbb{R}_{>0} \). A Log-Breve random variable having the probability density function \( f_{\tilde{S}}(\tilde{s}; \alpha, \beta) \) is denoted by \( \tilde{S} \sim \text{LogBreve}(\tilde{s}; \alpha, \beta) \).

References

[1] J Havil. Gamma. exploring euler’s constant. 2003.

[2] Peter S Bullen. Handbook of means and their inequalities, volume 560. Springer Science & Business Media, 2013.

[3] Gleb Beliakov, Humberto Bustince Sola, and Tomasa Calvo Sánchez. A practical guide to averaging functions, volume 329. Springer, 2016.

[4] Vera Čuljak. Schur-convexity of the weighted čebyšev functional ii. Journal of mathematical inequalities, 6(1):141–147, 2012.

[5] Chun-Ru Fub, Dongsheng Wanga, and Huan-Nan Shic. Schur-convexity for lehmer mean of n variables. Journal of Nonlinear Sciences & Applications (JNSA), 9(10), 2016.

[6] WF Xia and YM Chu. The schur harmonic convexity of lehmer means. In Int. Math. Forum, volume 4, pages 2009–2015. Citeseer, 2009.

[7] Zhen-Hang Yang. Schur power convexity of gini means. Bulletin of the Korean Mathematical Society, 50(2):485–498, 2013.

[8] Yu-Ming Chu and Tie-Hong Zhao. Convexity and concavity of the complete elliptic integrals with respect to lehmer mean. Journal of Inequalities and Applications, 2015(1):396, 2015.
[9] Ondrej Sluciak. On inflection points of the lehmer mean function. *arXiv preprint arXiv:1509.09277*, 2015.

[10] Yu-Ming Chu and WEIMAO QIAN. Sharp lehmer mean bounds for neuman means with applications. *JOURNAL OF MATHEMATICAL INEQUALITIES*, 10(2):583–596, 2016.

[11] Yu-Ming Chu and Miao-Kun Wang. Optimal lehmer mean bounds for the toader mean. *Results in Mathematics*, 61(3-4):223–229, 2012.

[12] Tiberiu Trif. Sharp inequalities involving the symmetric mean. *Miskolc Math. Notes*, 3:157–164, 2002.

[13] Tie-Hong Zhao, Yu-Ming Chu, and Wen Zhang. Optimal inequalities for bounding toader mean by arithmetic and quadratic means. *Journal of Inequalities and Applications*, 2017(1):26, 2017.

[14] Yu-Ming Chu and Miao-Kun Wang. Inequalities between arithmetic-geometric, gini, and toader means. In *Abstract and Applied Analysis*, volume 2012. Hindawi, 2012.

[15] Yun Hua and Feng Qi. A double inequality for bounding toader mean by the centroidal mean. *Proceedings-Mathematical Sciences*, 124(4): 527–531, 2014.

[16] Vagan Terziyan. Social distance metric: from coordinates to neighborhoods. *International Journal of Geographical Information Science*, 31 (12):2401–2426, 2017.

[17] S Somasundaram, SS Sandhya, and TS Pavithra. Lehmer-3 mean labeling of some new disconnected graphs. *International Journal of Mathematics Trends and Technology*, 35(1), 2016.

[18] Ondrej Sluciak and Markus Rupp. Consensus algorithms with state-dependent weights. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 5462–5466. IEEE, 2014.

[19] M Ivette Gomes, Fernanda Figueiredo, and Lígia Henriques-Rodrigues. Port value-at-risk estimation through generalized means. In *Proceedings 2nd ISI Regional Statistics Conference*, 2017.