International Journal of Probability and Statistics 2022, 11(1): 1-8
DOI: 10.5923/j.ijps.20221101.01

Lebesgue−Chebyshev Synergism for the Bounds of Measurable Random Variables

Francis Egenti Nzerem*, Ukamaka Cynthia Orumie

Department of Mathematics and Statistics, University of Port Harcourt, Choba, Nigeria

Abstract The attainment of a near-faultless optimal estimate is always the chief goal in the application of measures. Boundedness is always an 'essential commodity' if optimality is desired. A function is **Lebesgue integrable** if theLebesgue integral does not 'explode'; as well the existence of Chebyshev bound assures that the estimates of certain random variables or those of certain functionals, as the case may be, do not 'explode'. The realization of the latter has the underpinning of Lebesgue measure applied on the set functions under consideration. Therefore, this work treated a veritable way of estimating the (maximal and minimal) bounds of Chebyshev-type inequalities, with the functions presupposed Lebesgue measurable.

Keywords Lebesgue−Chebyshev, Inequalities, Bounds, Maximal function, Probability

1. Introduction

The central limit theorem in mathematical probability theory reckons that, often, the addition of independent random variables to the original variables causes the normalized sum of such variables to tend toward a normal distribution even though they are not normally distributed. However, the central theorem affords simply an asymptotic distribution, and it requires several observations to stretch into the tails of a normal distribution. Chebyshev's Inequality is widely used in probability theory, and; it furnishes probability bounds on wide-ranging aspects of probability distributions. The central importance of the inequality is the use in limiting the distance between the mean and a random variable, Amaresh [1]. As precious as the inequality is in doing the said work, its drawback is the inability to employ it in setting confidence intervals in estimation problems. However, the propitiation of this Chebyshev's shortcoming is the use of concentration inequalities (see [2], [3], [4]). Besides, Roos et al. [5] provided alternative tight lower and upper bounds on the tail probability under certain conditions. Such bounds were obtained as exact solutions to semi-infinite linear programs.

Chebyshev's inequality is, fundamentally, a consequence of the *measure theory*, as any resulting bounding value is to be prescribed in a "measurable set". In the light of this, Stepaniants [6] and Billingsley [7] showed that Chebyshev's inequality can be subsumed in the theory of Lebesgue measure. Besides the qualitative use of the generic Chebyshev's inequality in probability measures, numerous applications of Chebyshev-type inequalities furnish suitable bounds for the estimation of some class of functions. One of such celebrated functions is depicted in the abiding *maximal theorem* by Hardy and Littlewood [8] - a work that has engaged the attention of Flett [9], Keith [10], among numerous others found in literature. In recent times numerous Chebyshev-type inequalities are 'crafted' to attend to varying needs (see Teimourian and Ghazanfari [11], Nisar et al. [12])).

Another salient class of inequality is the type that is obtainable from the Chebyshev *functionals* of functions of bounded variation. While numerous identities that relate to the Chebyshev functional were considered by Mitrinović *et al*. [13,14], the quest for the bounds for the functional was done by Cerone [15,16] and Dragomir [17].

This work supplied the combined outcome of (Lebesgue) measure theory and Chebyshev's inequality in furnishing the method of obtaining estimates (bounds) of functions and Chebyshev-type functionals. The linear combinations of independent random variables, their control, and the maximal averages over some parameters are desired for purposes of empirical risk minimization. In effect, this line of study is quite graceful, for purposes of optimization.

2. Preliminaries

Some basic tools for the development of the content of this paper are supplied in this section.

2.1. The Measure Space

Definition 1. Let (X, σ) be a measurable space. A set function μ is a measure on the space if:

* Corresponding author:
frankjournals@yahoo.com (Francis Egenti Nzerem)
Received: Dec. 16, 2021; Accepted: Jan. 12, 2022; Published: Jan. 21, 2022
Published online at http://journal.sapub.org/ijps

(i) $\mu(A) \in [0,\infty]$ for $A \in \sigma(X)$;

(ii) $\mu(\varnothing) =$ 0;

(iii) Given a sequence of disjoint A_1, A_2, . . . of σ -sets, and if $\bigcup_{r=1}^{\infty}\mu(A_r) \in \sigma$, then

$$\mu\left(\bigcup_{r=1}^{\infty} A_r\right) = \sum_{r=1}^{\infty}\mu(A_r) . \tag{1}$$

The triple (X,σ,μ) is called a measure space.

2.1.1. Probability Space

Definition 2. Let (Ω, F) and (Ω', F') be two measurable spaces. The function $X : \Omega \to S$ is measurable a function and

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in F \quad \forall A \in F'. \tag{2}$$

X is called a *random variable* if $\Omega' = \mathbb{R}$ and F' is the Borel σ-algebra.

If Ω is any state space that is not endowed with σ-algebra, and $X : \Omega \to \Omega'$ is a function that has value in a measurable space (Ω', F'), then the smallest σ-algebra for which X is a measurable function is

$$\sigma(X) := \sigma(\{X^{-1}(A) : A \in F' \}). \tag{3}$$

The triplet (Ω, F, P) is a probability space, in which:

(i) Ω denotes the sample space describing a non-empty set of all possible outcomes of some model being performed.

(ii) F denotes an σ-algebra on the sample space (Ω ∈F).

(iii) P: F $\to [0,1]$, $P(\Omega) = 1$ is a probability measure if it as well satisfies the countably additive property specified in 2.1 (iii).

Probability measures has been largely used in optimization problems as seen in Saito and Takahashi [18] and Yu [19].

If X_1, X_2,… is a sequence of real-valued random variables defined on (Ω, F, P), then the supremum

$$X : \Omega \to \mathbb{R} \cup \{\infty\},$$
$$X(\omega) = \sup_n X_n(\omega)$$

is measurable (see Lowther [20]).

2.2. Excerpts from Lebesgue Integral

Let $X \in \mathbb{R}$ be a random variable. It has a Gaussian distribution if it has a Lebesgue measurable density p on R such that

$$p(x) = \frac{1}{\left(2\pi\sigma^2\right)^{1/2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad x \in R$$

where $\mu = \mathrm{E}(X)$ and σ^2 = var(X) > 0 are the mean and variance of X.

Let (X, S, μ) be a measure space:

(i) If $f : X \to [0,\infty]$ is an S-summable function, and S_P is an S-partition B_1, B_2,…, B_m of X, then the lower Lebesgue sum L (f , S_P) is defined by

$$L \quad (f, S_P) \ = \sum_{j=1}^{m} \mu(B_j) \inf_{B_j} f.$$

(ii) If $D \in$ S and χ_D is the characteristic function of D, we get (see Axler [21])

$$\int \chi_D d\mu = \mu(D).$$

(iii) If $f : X \to [\infty,\infty]$ is S-summable, then the Lebesgue space $\mathrm{L}^1(\mu)$ reads,

$\{f : f$ is an S- measurable function from X to $\mathbb{R}$ and $\|f\|_1 = \int |f| \, d\mu < \infty\}$

(iv) **Lemma 1**. *If* $l \in L^1(\mu)$, then (*Markov's inequality*)

$$\mu(\{x \in X : |l(x)| \geq c\}) = \frac{1}{c}\|l\|_1 \quad \forall l > 0 \tag{4}$$

Proof. Assume $c > 0$. Therefore

$$\begin{aligned} \mu\left(\{x \in X : |l(x)| \geq c\}\right) &= \frac{1}{c}\int_{\mathrm{I}} c\, d\mu; \\ &\leq \frac{1}{c}\int_{\mathrm{I}} |l| \, d\mu, \ \text{since} |l(x)| \geq c \\ &\leq \frac{1}{c}\|l\|_1 , \end{aligned} \tag{5}$$

where $\mathrm{I} = \{x \in X : |l(x)| \geq c\}$.

2.3. Chebyshev's Inequality from Lebesgue Integral

The generic (probability) statement of Chebyshev's Inequality reads:

Theorem 1. *(Chebyshev's Inequality). Let $X : \Omega \to R$ be a random variable on a probability space (Ω, $\mathcal{F}$, P). Suppose X has a finite expected value m and finite nonzero variance σ^2. Then, for any real number $k > 0$*

$$P(|X - m| \geq k\sigma) \leq \frac{1}{k^2} . \tag{6}$$

Note that X above is integrable.

Theorem 2. *Suppose (X, S, μ) is a measure space with $\mu(X) = 1$ and $l \in L^1(\mu)$, then*

$$\mu\left(\left\{x \in X : \left|l(x) - \int l d\mu\right| \geq c\right\}\right) \leq \frac{1}{c^2}\left(\int l^2 d\mu - \left(\int l d\mu\right)^2\right) \quad \forall c > 0 \tag{7}$$

The proof may easily derive from the proof of Markov inequality (4) already presented here.

Compare the inequality (7) above to Markov's inequality (4). Consider the left-hand side of the inequality (7). We see that

$$\left|l(x) - \int_B l \mathrm{d}\mu\right| \leq |l(x)| - \left|\int_B l \mathrm{d}\mu\right| , \quad B \in \mathrm{S}.$$

Let χ_B be the characteristic function of *B*. Then,

$$\left|\int_B l\mathrm{d}\mu\right| = \left|\int \chi_B l\mathrm{d}\mu\right|$$
$$\leq \left|\int \chi_B l\right|\mathrm{d}\mu \leq \int \sup_B |l|\,\chi_B\mathrm{d}\mu \qquad (8)$$
$$= s\mu(B),$$

let $\left|l(x) - \int_B l\mathrm{d}\mu\right| =$ where $s = \sup_B |l|$. The above shows that the integral measure $\int l\mathrm{d}\mu$ in (7) encodes the mean value. Therefore, $|X-\mu|$; $c = k\sigma$, we recognize the integral in the rightmost expression of (7) as equal to Var(*X*) = σ^2. If *P* is the probability measure, one finds that the inequality (6) is equivalent to the inequality (7). Under some condition(s), the *tail estimate* of a given measurable function may be obtained using Chebyshev's inequality. Let *f* be a non-negative measurable function. Define a set $\{x \in E: f(x) \geq \omega \geq 0\}$. The property

$$\omega 1_{\{x \in E:\, f(x) \geq \omega\}} \leq f \qquad (9)$$

holds. On integration, one gets Chebyshev's inequality in the form

$$\omega\mu_{\{x \in E:\, f(x) \geq \omega\}} \leq \mu(f\,). \qquad (10)$$

Suppose *g* is any measurable function. The representative inequalities for *g* may be obtained by choosing some non-negative measurable function $\varphi(x)$ and applying Chebyshev's inequality to $f = \varphi \circ g$. Thus,

$$\mu\left(|g|^p \geq \omega\right) = \mu\left(|g|^p \geq \omega^p\right) \leq \omega^{-p}\mu\left(|g|^p\right) < \infty\,. \qquad (11)$$

The required tail estimate is [21]

$$\mu\left(|g| \geq \omega\right) = O(\omega^{-1}), \quad as\, \omega \to \infty\,. \qquad (12)$$

Still from the measure-theoretic point, let (X, Ω, μ) be a measure space and let *f* be an extended real-valued measurable function defined on *X*. Then for any real number $t > 0$ and $0 < p < \infty$ [22],

$$\mu(\{x \in X : |f(x)| \geq t\}) \leq \frac{1}{t^p}\int_{|f|\geq t} |f|^p d\mu.$$

In general, if *g* is an extended real-valued measurable function, nonnegative and nondecreasing, with $g(t) \neq 0$ then:

$$\mu(\{x \in X : f(x) \geq t\}) \leq \frac{1}{g(t)}\int_X g \circ f\, d\mu.$$

Chebyshev's inequality may be expressed in the sense of Lebesgue integration (see Stepaniants [5]) in the theorem that follows:

Theorem 3. *(Generalized Chebyshev's Inequality). Let (Y, σ, μ) be a measure space and let f be a real-valued measurable function defined on Y. Suppose μ is the Lebesgue measure and let h be a non-negative and non-decreasing real-valued measurable function on the range of f. Then, for any positive real number n and 0 < t < ∞,*

$$\mu(\{x \in Y : f(x) \geq n\}) \leq \frac{1}{h(n)}\int_Y h(f(x))d\mu(x) \qquad (13)$$

Proof: For a fixed number, *n*, Let G_n be a set and let $G_n = \{x \in Y : f(x) \geq n\}$. Denote the characteristic function of the set G_n by χG_n. By the theorem, *h* is non-decreasing and it is non-negative on the range of *f*, therefore

$$0 \leq h(n)\chi G_n \leq h(f(x))\chi G_n \qquad (14)$$

Apply Lebesgue integration and integrate over *Y*, to get

$$h(n)\mu(G_n) = h(n)\int_Y \chi G_n d\mu = \int_Y h(n)\chi G_n d\mu \qquad (15)$$

Now,

$$\int_Y h(n)\chi G_n d\mu \leq \int_Y h(f(x))\chi G_n d\mu$$
$$= \int_{G_n} h(f(x))d\mu \leq \int_Y h(f(x))d\mu. \qquad (16)$$

In (13) above the last inequality holds since *h* is nonnegative everywhere. In conclusion, we have

$$\mu(G_n) \leq \frac{1}{h(n)}\int_Y h(f(x))d\mu \qquad (17)$$

2.4. Equi-measurable Functions

So far, the emphasis had been on non-decreasing functions. What happens if a function is decreasing in an interval? The relationship between the above two types of functions is well articulated in the seminal and yet abiding work by Hardy and Littlewood [8].

Assume that $f(x)$ is a positive, bounded, and measurable in (a_0, a_1); let $\beta(y)$ be a measure such that $f(x) \geq y$, therefore $\beta(y)$ is a decreasing function of *y* which vanishes for sufficiently large *y*. The function $f^*(x)$ is defined for $x \in [a_0, a_1]$ by

$$f^*\{\,\beta\,(y)\} = y\,\beta(y) \in [a_0, a_1]. \qquad (18)$$

It is constant in some interval (i_1, i_2) if $\beta(y)$ has a discontinuity, with a jump from i_1 to i_2. It is known as the rearrangement of $f(x)$ in decreasing order. The *effective upper bound* of $f(x)$ corresponds to the upper bound of $f^*(x)$ excluding in a null set, as sets of zero measure do not count in the definition of $f^*(x)$. The functions $f(x)$ and $f^*(x)$ are *equimeasurable* when the measures of the sets wherein they take up values lying in a prescribed interval are equal. In the interval (a_0, a_1) if $\varphi(x)$ is a function, then

$$\int_0^a \varphi(f)dx = \int_0^a \varphi(f*)dx\,, \qquad (19)$$

whenever each integral exists. If $\lambda \in [a_0, x)$ is a function, the functional *F* is such that

$$F(x,\lambda) = F(x,\lambda, f) = \frac{1}{x-\lambda}\int_\lambda^x f(t)dt, \quad F(x,x) = f(x). \qquad (20)$$

Define M(*x*, *f*), the maximum average of *f*(*x*) about a point *x* by

$$(x, f) = \max_{\lambda \in [a_0, x]} F(x, \lambda). \tag{21}$$

Lemma 2. *If a function φ(x) is positive, continuous, and non-decreasing with x, and λ = λ (x) is any measurable function such that λ ϵ (a₀, x), then* [8]

$$\int_{a_0}^{a_1} \varphi\{F(x, \lambda, f)\} dx \le \int_{a_0}^{a_1} \varphi\{F(x, a_0, f*)\} dx. \tag{22}$$

Note that we may, in most cases, take $a_0 = 0$, $a_1 = 1$.

Lemma 3. *If* M(x) is *the upper bound of*

$$F(x, \lambda, f) = \frac{1}{x - \lambda} \int_{\lambda}^{x} f(t) dt \quad \lambda \in [x_1, x_2], \tag{23}$$

Then

$$\int_{x_1}^{x_2} \varphi\{M(x)\} dx \le 2 \int_{x_1}^{x_2} \varphi\left(\frac{1}{x - x_1} \int_{x_1}^{x} f*(t) dt\right) dx. \tag{24}$$

The proofs of *Lemma* 2 and *Lemma* 3 are available in Hardy and Littlewood [8]). The consequence of Hardy-Littlewood maximal theorem is that for a decreasing function, $f^*(x)$, that is equi-measurable with f on $(0, \infty)$, and for a measurable set, U, with a measure, μ, if t is a positive real number, then

$$\int_{\mathbb{R}} f d\mu = \int_0^{\infty} f * d\mu = \int_0^{\infty} \mu(U_t[f]) dt. \tag{25}$$

It was shown in Phillips [9] that inequality of the form

$$\int_{\Xi} f d\mu \le \int_0^{\mu(\Xi)} f * d\mu \tag{26}$$

holds well for all Lebesgue measurable set Ξ.

3. Estimation of Bounds

The estimation of the bounds of measurable functions is of much importance in applicable phenomena. Inequality measures contribute much to near guileless estimates. Extensive energies that have so far been expended on the *theory of maximal inequalities* for several decades have produced a sumptuous result. As expected, such inequalities must be rigged in the Lebesgue ($\mathcal{L}^p$ -) function space if the bounds are to be sought. This section is devoted to inequality bounds that are consequences of Chebyshev's inequality and with the underpinning of $\mathcal{L}^p$ measures.

3.1. Bounds in Function ($\mathcal{L}^p$) Spaces

3.1.1. Probability Bound

Many a time the linear combinations of independent random variables, their control, and the maximal averages over some parameters are desired for purposes of empirical risk minimization.

Let P be a probability distribution over some set X. An ε-**net** for a class $H \subseteq 2^X$ of subsets X is any subset G ⊆ X such that for any $q \epsilon H$

$$P(q) \ge \varepsilon \Rightarrow G \cap q \ne \phi.$$

The subset G approximates the probability distribution. An **ε**-approximation for class H is such that [23]

$$\left| P(q) - \frac{|G \cap h|}{|G|} \right| < \varepsilon.$$

The following argument derives from [24].

Let $X_1, \ldots, X_n$ be n independent and random variables such that $\mathbf{E}[X_i] = \mu$ and $\mathrm{var}(X_i) \le \sigma^2$. Let $\delta \in (0, 1)$ and assume that n can be factored into $n = K \cdot W$ where $W = 8 \log(1/\delta)$ is a positive integer. For w = 1, ..., W, let $\bar{X}_w$ denote the average over the w-th group of k variables. Then, this average takes the form

$$\bar{X}_w = \frac{1}{k} \sum_{i=(w-1)k+1}^{wk} X_i.$$

For any w = 1,...,W, it can be shown [24] that

$$\mathrm{P}\left[\bar{X}_w - \mu \ge \frac{2\sigma}{\sqrt{k}}\right] \le \frac{1}{4}. \tag{27}$$

Moreover, define $\hat{\mu}$ as the median of {$X1$,..., XW}. Then

$$\mathrm{P}\left[\hat{\mu} - \mu \ge \frac{2\sigma}{\sqrt{k}}\right] \le \mathrm{P}\left[\mathcal{B} \ge \frac{W}{2}\right].$$

where B ~ Bin(W, 1/4). Evidently,

$$\mathrm{P}\left[\hat{\mu} - \mu \ge 4\sigma\left(\frac{2\log(1/\delta)}{n}\right)^{1/2}\right] \le \delta. \tag{28}$$

Definition 3. The space of functions $\boldsymbol{L}^p(\Omega)$, $p \in [1, \infty)$, for which the p-th power is Lebesgue integrable over Ω is defined by

$$L^p(\Omega) = \left\{ w : \|w\|_{L^p(\Omega)} = \left(\int_{\Omega} |w|^p \, d\mu\right)^{1/p} < \infty \right\}; \tag{29}$$

and the space of functions L^{∞} with finite essential supremum is

$$L^p(\Omega) = \left\{ w : \|w\|_{L^{\infty}(\Omega)} = \text{ess sup}_{\Omega} |w| < \infty \right\}, \tag{30}$$

where

$$\text{ess sup}_{\Omega} |w| = \inf\{C \ge 0 : |w(x)| \le C \text{ for a.e. } x \in \Omega\}.$$

Consider the L^2 ball. Let $\mathcal{H} \subset R^d$ be fixed and let $\varepsilon > 0$. A set G is called an ε-net of $\mathcal{H}$ with respect to a distance $d(\cdot, \cdot)$ on R^d, if $G \subset \mathcal{H}$ and for any $z \in \mathcal{H}$, there exists $x \in G$ such that $d(x, z) \le \varepsilon$. The unit L^2 ball of R^d is the set of vectors u that have Euclidean norm $|u|_2$ at most 1, usually defined by

$$B_2 = \left\{ x \in R^d : \sum_{i=1}^{d} x_i^2 \le 1 \right\}.$$

The upper bound on the magnitude of the smallest ε-net of $\mathcal{B}_2$ is given below.

Lemma 4. [24] Let ε ∈ (0, 1). Then the unit Euclidean ball $\mathcal{B}_2$ has an *ε-net* $\mathcal{H}$ with respect to the Euclidean distance of cardinality $|\mathcal{H}| \le (3/\varepsilon)^d$

Proof. Going by [24], construct an iteration of the *ε-net*. Choose $x_1 = 0$. For a given $i \ge 2$, let x_i be any $x \in \mathcal{B}_2$ such that $|x-x_j|_2 > \varepsilon$ for all $j < i$. If there is no such x, stop the process. Thus, an ε-net is generated. Then, the size of the ε-net has to be controlled.

The Euclidean balls centred at $x \in \mathcal{H}$ and with radius ε/2 are disjoint as $|x-y|_2 > \varepsilon$ for all $x, y \in \mathcal{H}$. Besides,

$$\bigcup_{z\in\mathcal{H}} \{z + \varepsilon x \mathcal{B}_2 \ \} \subset \left(1+\frac{\varepsilon}{2}\right)\mathcal{B}_2 \tag{31}$$

where $\{z + \varepsilon B_2\} = \{z + \varepsilon x,\ x \in \mathcal{B}_2\}$. Thus, measuring volumes, one finds

$$\sum_{z\in\mathcal{H}} \mathrm{vol}\left(z+\frac{\varepsilon}{2}\mathcal{B}_2\right) = \mathrm{vol}\left(\bigcup_{z\in\mathcal{H}}\{z+\frac{\varepsilon}{2}\mathcal{B}_2\}\right) \le \mathrm{vol}\left(\left(1+\frac{\varepsilon}{2}\right)\mathcal{B}_2\right). \tag{32}$$

Equivalently,

$$|\mathcal{H}|\left(\frac{\varepsilon}{2}\right)^d \le \left(1+\frac{\varepsilon}{2}\right)^d. \tag{33}$$

Thus, the following bound holds

$$|\mathcal{H}| \le \left(1+\frac{\varepsilon}{2}\right)^d \le \left(\frac{3}{\varepsilon}\right)^d \tag{34}$$

Let $B(x, r) := \{y \in R^d: |x - y| < r\}$, $r > 0$, denote the open ball of radius r centred at x for any $x \in R^d$. The averaging operators ψ_r on R^d for a locally integrable function f read:

$$\Psi_r f := \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy. \tag{35}$$

On every $\mathcal{L}^p(R^d)$, $p \in [1,\infty]$, the operators ψ_r are contractions for which Young's inequality

$$\|\Psi_r\|_{L^p(\boldsymbol{R}^d)} \le \|f\|_{L^p(\boldsymbol{R}^d)} \tag{36}$$

holds.

For an $\mathcal{L}^p$-measurable function, f, the maximal function is of the form

$$(\Psi_m f)(x) = \sup_{B\in B_x} \frac{1}{L^p(B)}\int_B |f| dL^p, \tag{37}$$

with the weak $\mathcal{L}^1$-estimate

$$\mu\{\Psi_m f > t\} \le \frac{C}{t}\|f\| L^1, \forall t > 0 \tag{38}$$

The averages $\psi_r f$ are uniformly bounded in magnitude and in shape as r varies. From the foregoing, we state:

Proposition 1. (Hardy-Littlewood maximal inequality)

$$\left\|\sup_{r>0}|\Psi_r f|\right\|_{L^p(\boldsymbol{R}^d)} \le \|f\|_{L^p(\boldsymbol{R}^d)} \quad \forall p \in (1,\infty], \tag{39a}$$

for any $f \in L^p(\boldsymbol{R}^d)$

and

$$\left\|\sup_{r>0}|\Psi_r f|\right\|_{L^{1,\infty}(\boldsymbol{R}^d)} \le \|f\|_{L^1(\boldsymbol{R}^d)} \quad \text{for any } f \in L^1(\boldsymbol{R}^d). \tag{39b}$$

Consider the equations describing the Borel-measurable *maximal averages* of f in some prescribed interval:

$$\begin{aligned} \Psi_{r_0} f(x) &= \sup\left\{\frac{1}{\mu(I)}\int_{I=[x,z]} f d\mu : z \in (x,\infty)\right\} \\ \Psi_{r_1} f(x) &= \sup\left\{\frac{1}{\mu(I)}\int_{I=[z,x]} f d\mu : z \in (-\infty,x)\right\} \\ \Psi f(x) &= \sup\left\{\frac{1}{\mu(I)}\int_I f d\mu : I \text{ is a closed interval that contains } x\right\} \end{aligned} \tag{40}$$

In equations (40) above,

$$\Psi f = \mathrm{Max}(\Psi_{r_0} f, \Psi_{r_1} f) \tag{41}$$

is the sublinear operator encoding the Hardy-Littlewood maximal operator. The proof of (41) above may be found in Keith [10].

Suppose t is a positive real number and w is an extended real-valued positive function. Let

$$B_t[w] = \{x \in R : w(x) > t\} \tag{42}$$

Lemma 5. *For every $k \in (0, 1)$ and every $t > 0$:*

$$(i) \qquad \mu\left(B_t[\Psi_i f]\right) \le \frac{1}{(1-k)t}\int_{B_{kt}[f]} f d\mu \quad (i = r_0, r_1); \tag{43}$$

$$(ii) \qquad \mu\left(B_t[\Psi f]\right) \le \frac{2}{(1-k)t}\int_{B_{kt}[f]} f d\mu.$$

Proof. Following Philip [9], define the distribution function $w(x)$ by

$$w(x) = \begin{cases} f(x) & \text{if } f(x) > kt \\ 0 & \text{otherwise} \end{cases}. \tag{44}$$

So $\Psi_{r_0} f \le \Psi_{r_0} w + k\,t$ and we have $B_t\left[\Psi_{r_0} f\right] \subset B_{(1-k)t}\left[\Psi_{r_0} w\right]$. Since the equality

$$\mu\left(B_t[\Psi_i f]\right) = \frac{1}{t}\int_{B_t[\Psi_i f]} f d\mu \tag{45}$$

holds for every $t > 0$, we have

$$\mu\left(B_t\left[\Psi_{r_o} f\right]\right) \le \frac{1}{(1-k)t}\int_{B_{(1-k)t}[\Psi_{r_o} f]} w d\mu. \tag{46}$$

The function w(x) was assumed to be nonnegative so

$$\mu(w) = 0 \text{ on } B_{kt}[\Psi]'$$

since $w = 0$ there.

Thus

$$\int_{B_{(1-k)t}[\Psi_r f]} w d\mu < \int_{B_{kt}[f]} f d\mu. \tag{47}$$

A similar approach may be used to prove (i) for $i = r_1$, and (ii) may be proved by using the property that

$$\mu\left(B_t\left[\Psi f\right]\right) \le \frac{2}{t}\int_{B_t[\Psi f]} f d\mu \quad \forall t > 0 \tag{48}$$

The following relations on the maximal operator are often used in estimations

Lemma 6. *If $f \in L^p$, then for $p > 1$ the following inequalities*

$$\text{(i)} \quad \left\|\Psi_i f\right\|_p \le \frac{p}{p-1}\left\|f\right\|_p, \ (i = r_0, r_1) \tag{49a}$$

and

$$\text{(ii)} \quad \left\|\Psi f\right\|_p \le \frac{2^{1/p} p}{p-1}\left\|f\right\|_p \tag{49b}$$

apply.

The proof to the above theorem is readily found in Flett [9] and Keith [10].

3.2. Chebyshev Functional

The bounding of the Chebyshev functional has extensive applicability in probability problems and the bounding of special functions in applied mathematics. Some theories and identities that relate to the Chebyshev functional are treated in Mitrinovi´c *et al*. [13,14] while Fink [25] looked at an applicable aspect of such inequalities. Elegant and enduring work in *the quest to bound the Chebyshev functional* was done by Cerone [15].

As usual, let f, g: $[a,b] \to \mathbb{R}$ be two measurable functions. The weighted Chebyshev functional $W(f, g)$ is defined by

$$W(f, g; q) := \mathrm{M}(f g; q) - \mathrm{M}(f; q)\mathrm{M}(g; q), \tag{50}$$

with the weighted integral mean given by

$$Q.\ \mathrm{M}(f;q) = \int_a^b q(x) f(x) dx,$$

where $0 < Q = \int_a^b q(x)dx < \infty$ and the above integral is assumed to exist. The following equivalences hold [22],

$$Q(f;g;1) \equiv Q(f, g),\ M(f;1) \equiv \mathrm{M}(f). \tag{51}$$

It is conventional to present the Chebyshev functional $W(f, g)$ as Korkine's identity. If f, g: $[a,b] \to \mathbb{R}$ are two equi-monotonic functions, then

$$W(f, g)$$

$$\equiv \frac{1}{b-a}\int_a^b f(x)g(x)dx \ge \frac{1}{b-a}\int_a^b f(x)dx . \frac{1}{b-a}\int_a^b g(x)dx \tag{52}$$

$$= \frac{1}{2(b-a)^2}\int_a^b\int_a^b \left(f(x) - f(y))(g(x) - g(y)\right)dxdy,$$

provided the integrals exist. If $f(x)$ and $g(y)$ are synchronous functions *i.e.* $(f(x) - f(y))(g(x) - g(y)) \ge 0$, a.e. $x, y \in [a, b]$, the classical Chebyshev functional inequality is satisfied by

$$W(f, g) \ge 0. \tag{53}$$

As usual, let the triple (X, σ, μ) be a measure space, with μ as a countably additive and positive measure on σ with values in R ∪ {∞}.

Given a μ-measurable function v: $X \to \mathbb{R}$, with $v(x) \ge 0$ for μ, a.e. $x \in X$, define a Lebesgue space $L_v(X, \sigma, \mu) := \{f : X \to \mathbb{R}$, f is μ-measurable and $\int_X v(x)\left|f(x)\right| d\mu(x) < \infty\}$. Assume $\int_X v(x)d\mu(x) > 0$. If f, g: $X \to \mathbb{R}$ are μ-measurable functions and $f, g, fg \in L_v(X, \sigma, \mu)$, then the Chebyshev functional that admits $v(x)$ may be written as

$$\mathrm{W}_v(f, g) = \mathrm{W}_v(f, g; X)$$

$$:= \frac{1}{\int_X v(x)d\mu(x)}\int_X v(x) f(x) g(x) d\mu(x) \text{-} I_{(v(x), f(x), g(x))} \tag{54}$$

where

$$I_{(v(x), f(x), g(x))} =$$

$$\frac{1}{\int_X v(x)d\mu(x)}\int_X v(x) f(x) d\mu(x) \frac{1}{\int_X v(x)d\mu(x)}\int_X v(x) g(x) d\mu(x) \cdot$$

In Cerone [15], remarkable work was done in using the above to furnish some bounds. As well, the bounds for the modulus of complex Chebyshev functionals engaged the attention of Dragomir [17]. In Grüss [26], the Chebyshev functional was evaluated as

$$\left|W(f, g)\right| \le \frac{1}{4}\left(\tau - \tau_1\right)\left(\upsilon - \upsilon_1\right), \tag{55}$$

provided τ, τ_1, υ, υ_1 are real numbers such that

$$-\infty < \tau \le f \le \tau_1 < \infty, \quad -\infty < \upsilon \le g \le \upsilon_1 < \infty \ \text{ a.e.on } [a,b]$$

The constant ¼ is assumed sharp, as there would be no smaller number in its stead. The proof of this inequality is available in Mitrinovi´c *et al.* [13].

3.3. Bound for the Generalized Prime Numbers

Beurling [27] conducted an extensive and enduring study on the asymptotic distribution of prime numbers. The distribution function $N(x)$ for the generalized prime numbers satisfies the inequality

$$\int_1^x x^{-1}\left\{\sup_{x \le y} \frac{\left|N(y) - C(y)\right|}{y}\right\} dx < \infty, \tag{56}$$

with $C > 0$.

More general is the case in which the counting function of the integer satisfies

$$N(x) = C(x) + o(x \log^{-\gamma} x) \tag{57}$$

for some positive C and γ. The prime number theorem holds if γ > 3/2, else it may fail to hold. The Chebyshev prime counting estimate

$$\liminf_{x\to\infty} \frac{\psi(x)\log x}{x} \ge a > 0, \quad \limsup_{x\to\infty} \frac{\psi(x)\log x}{x} \le A < \infty \tag{58}$$

is known to hold if $\gamma > 1$ but it may fail if $\gamma < 1$ (see Hall [28]). It was shown (see Hall [29]) that if the distribution function $N(x)$ is of the form

$$N(x) = C(x) + o(x), \tag{59}$$

then the Chebyshev's lower and upper bounds are satisfied by

$$\liminf_{x\to\infty} \frac{\psi(x)\log x}{x} \le 1, \quad \limsup_{x\to\infty} \frac{\psi(x)\log x}{x} \ge 1. \tag{60}$$

Following Diamond [30,31]

Let some function F: $[1, \infty) \to \boldsymbol{R}$ be of bounded variation on each interval $[1, x]$, with an accompanying Borel-Stieltjes measure, dF, on $[1, \infty)$. Define an operator T on the measures in the form

$$\int_E TdF = \int_{t\in E} \log t dF(t). \tag{61}$$

The convolution

$$d\xi * dN = TdN, \tag{62}$$

was seen to hold well for Beurling generalized numbers (see Diamond [31] and Zhang [32]).

Let

$$\eta_\alpha(x) := \int_1^x dN \otimes (\delta - \alpha t^{-\alpha} dt) \otimes (\log et)^{-\gamma} dt \tag{63}$$

be the convoluted sides of (59) by $(\delta - \alpha t^{-\varepsilon} dt)(\log et)^{-\gamma}$, where δ encodes the Dirac measure at 1, dt is the Borel-Lebesgue measure on $[1,\infty)$ and α is a positive number. Then we state, without proof, the following lemma whose proof furnishes Chebyshev's upper and lower bounds.

Lemma 6. Suppose that (36) holds with C > 0 and $\gamma > 1$. Then there exists a positive number α depending on N(x) that satisfies the conditions: (i) $\eta_\alpha(x) \ge 0 \ \forall x \ge 1$ (ii) $\eta_\alpha(x) \to \infty$ as $x \to \infty$

The proof of the above *lemma* is available in Diamond [31]. As a result, the upper bound is satisfied by

$$\psi(x/D) \le \int_1^x \eta_\alpha(x/t) d\psi(t) = o(x) \Rightarrow \limsup_{x\to\infty} \psi(x)/x < \infty, \tag{64}$$

where D is a number for which given a number n such that $\eta_\alpha(n) \ge 0 \ \forall n \ge 1$, $\eta_\alpha(n) \ge 1 \ \forall n \ge D$. The lower bound is satisfied by

$$\liminf_{x\to\infty} \psi(x)/x \ge \frac{C}{3KQ} > 0, \quad K > 0. \tag{65}$$

4. Summary and Conclusions

The achievement of a near-flawless optimal estimate is always a desirable end. Boundedness is always a *sine qua non* to the achievement optimality. The linear combinations of independent random variables, their control, and the maximal averages over some parameters are desired for purposes of empirical risk minimization. This work supplied the (Lebesgue) measure theory and Chebyshev's inequality synergism in providing the method of obtaining estimates (bounds) of functions and Chebyshev-type functionals which may be applicable in such risk minimizations.

Chebyshev's inequality in the main is a consequence of the *measure theory*, as any resulting bounding value is embedded in a "measurable set". It was shown in this work that Chebyshev's inequality and Chebyshev-type inequalities can be subsumed in the theory of (Lebesgue) measures. Besides the qualitative use of the generic Chebyshev's inequality in probability measures, numerous applications of Chebyshev-type inequalities provide appropriate bounds for the estimation of some class of functions. This reasoning calls for an in-depth study of measure theory as it acts as a catalyst to the resolution of optimization problems.

ACKNOWLEDGEMENTS

The authors are grateful to the reviewers for their useful suggestions that led to the revision of this work.

REFRENCES

[1] A. Das, "A Note on Chebyshev Inequality: To Explain or to Predict," Global Journal of Science Frontier Research: F Mathematics and Decision Sciences, Volume 17, Issue 5 Version 1.0, pp. 25-29, 2017.

[2] P-M. Samson, "Concentration of measure inequalities for Markov chains and φ-mixing processes," The Annals of Probability, Vol. 28, No. 1, pp. 416–461, 2000.

[3] M. Talagrand, "Concentration of measure and isoperimetric inequalities in product spaces," Publ. Math. I.H.E.S. 81, pp.73–205, 1995.

[4] M. Talagrand, "New concentration inequalities in product spaces," Invent. Math. 126, 505–563, 1996.

[5] E. Roos, R. Brekelmans, W. vanEekelen, D. denHertog, J. S.H. vanLeeuwaarden "Tight tail probability bounds for distribution-free decision making" European Journal of Operational Research, [ONLINE, 25 December 2021]. Available: https://www.sciencedirect.com/science/article/pii/S0377221721010171.

[6] G. Stepaniants, The Lebesgue Integral, Chebyshev's Inequality, and the Weierstrass Approximation Theorem. [ONLINE]. Available: https://sites.math.washington.edu.

[7] P. Billingsley, Probability and Measure 3rd ed. John Wiley & Sons, 1995.

[8] G. H. Hardy, J. E. Littlewood "A maximal theorem with function-theoretic applications," Acta Mathematica, Acta Math. 54, pp. 81-116, 1930.

[9] T. M. Flett, "Some remarks on a maximal theorem of Hardy and Littlewood," The Quarterly Journal of Mathematics, Volume 6, Issue 1, pp. 275–282, 1955, https://doi.org/10.1093/qmath/6.1.275.

[10] P. Keith. “The Maximal Theorems of Hardy and Littlewood.” The American Mathematical Monthly Vol. 74, no.6, pp. 648–60, 1967, https://doi.org/10.2307/2314249.

[11] R. Teimourian and A. G. Ghazanfari, “Some Chebyshev type inequalities involving the Hadamard product of Hilbert space operators,” MATEMATI ˇCKI VESNIK, 72, (4), 303–313, 2020.

[12] K.S. Nisar, G. Rahman and K. Mehrez, “Chebyshev type inequalities via generalized fractional conformable integrals,” *J Inequal Appl*, 245, 2019. https://doi.org/10.1186/s13660-019-2197-1.

[13] D.S. Mitrinovi´c, J.E. Peˇcari´c, and A.M. Fink, Classical and New Inequalities in Analysis, Mathematics and Its Applications (East European Series), Kluwer Academic, Dordrecht, 1993, vol. 61.

[14] D.S. Mitrinovi´c, J.E. Pe ˇcari´c and A .M. Fink, Inequalities for Functions and their Integrals and Derivatives, Kluwer Academic Publishers, 1994.

[15] P. Cerone, On Chebyshev functional bounds, Hindawi Publishing Corporation Proceedings of the Conference on Differential & Difference Equations and Applications, pp. 267—277.

[16] P. Cerone, A refinement of the Gr ¨uss inequality and applications, RGMIA Research Report Collection 5, no. 2, article 14, 2002.

[17] S.S. Dragomir, Some Bounds for the Complex Cebyšev Functional of Functions of Bounded Variation, Symmetry, 13, 990, pp.1-15, 2021, https://doi.org/10.3390/sym13060990.

[18] T. Saito and A. Takahashi, “Portfolio optimization with choice of a probability measure,” [ONLINE, December 14, 2021]. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3816173.

[19] H. Yu, “Average-Cost Optimality Results for Borel-space Markov Decision Processes with universally measurable policies,” [ONLINE]. Available: https://arxiv.org/pdf/2104.00181.pdf.

[20] G. Lowther, “Essential suprema”, [ONLINE]. Available: https://almostsuremath.com/2019/01/06/essential-suprema/.

[21] S. Axler, Measure, Integration & Real Analysis, Springer Open, [ONLINE], Available: https://link.springer.com/book/10.1007/978-3-030-33143-6.

[22] J. R. Norris, Probability and Measure, Cambridge University. [ONLINE]. Available: http://www.statslab.cam.ac.uk/james/Lectures/pm.pdf.

[23] ε-net (computational geometry). [ONLINE]. Available: https://en.wikipedia.org/wiki/%CE%95-net.

[24] Sub-Gaussian Random Variables. [ONLINE]. Available: https://ocw.mit.edu/courses/mathematics/18-s997-high-dimensional-statistics-spring-2015/lecture-notes/MIT18_S997S15_Chapter1.pdf.

[25] A. M. Fink, “A treatise on Gr¨uss’ inequality, Analytic and Geometric Inequalities and Applications,” Math. Appl., vol. 478, pp. 93—113, 1999.

[26] G. Grüss, Math. Z. 1934, 39, 215–226. [CrossRef].

[27] A. Beurling, “Analyse de la loi asymptotique de la distribution des numbers premiers generalises I,” Acta Math. 68, pp.255-291, 1937.

[28] R.S. Hall, “Beurling’s generalized prime number systems in which the Chebyshev inequalities fail,” Proceedings of the American Mathematical society, Vol. 40, Number 1, 1973.

[29] R.S Hall, Theorems about Beurling’s generalized primes and the associated zeta function, Ph.D. Thesis, University of Illinois, Urbana. III., 1967.

[30] H.G. Diamond, “Asymptotic distribution of Beurling’s generalized integers,” Illinois J. Math. 14, pp.12-28, 1970.

[31] H.G. Diamond, “Chebyshev estimates for Beurling’s generalized prime numbers,” Proceedings of the American Mathematical Society, 39 (3), 1973.

[32] W-B. Zhang, “Chebyshev Type Estimates for Beurling Generalized Prime Numbers,” Proceedings of the American Mathematical Society, vol. 101, no. 2, October, 1987.

Copyright © 2022 The Author(s). Published by Scientific & Academic Publishing
This work is licensed under the Creative Commons Attribution International License (CC BY). http://creativecommons.org/licenses/by/4.0/