Journal of Mathematical Inequalities
Volume 17, Number 1 (2023), 403–418
doi:10.7153/jmi-2023-17-27

EXPECTATIONS OF LARGE DATA MEANS

TOMISLAV BURIĆ*, NEVEN ELEZOVIĆ AND LENKA MIHOKOVIĆ

(*Communicated by J. Pečarić*)

Abstract. In this paper we present estimation formulas for the expectations of power means of large data and associate them with means of probability distribution and means of random sample. The proposed method follows from the asymptotic expansion of power means which is applicable for sufficiently large data and it is especially useful when value of such expectation is hard to obtain. We will show the accuracy of these approximations for random samples which have uniform and normal distribution and analyse their behaviour for large sample volume.

1. Introduction

Let X be a random variable with mathematical expectation μ and standard deviation σ and consider the random sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$ of volume n from distribution X. Furthermore, define sample moments of order k about $c \in \mathbb{R}$ by

$$m_k(\mathbf{X}, c) := \frac{1}{n}\sum_{i=1}^{n}(X_i - c)^k, \quad k \in \mathbb{N},$$

and also let

$$\mu_k := \mathbb{E}[(X-\mu)^k] = \mathbb{E}[m_k(\mathbf{X}, \mu)].$$

For the n-tuple $\mathbf{a} = (a_1, \dots, a_n)$ with positive entries, the n-variable (unweighted) power mean can be defined as

$$M_r(\mathbf{a}) = \begin{cases} \left[\frac{a_1^r + a_2^r + \ldots + a_n^r}{n}\right]^{1/r}, & r \neq 0, \\ \sqrt[n]{a_1 a_2 \cdots a_n}, & r = 0. \end{cases}$$

Some well known classical means such as harmonic mean H, geometric mean G, arithmetic mean A and quadratic mean Q belong to the class of power means for $r = -1, 0, 1, 2$ respectively. Goal of this paper is to approximate the power mean of random sample and its expectation $\mathbb{E}[M_r(\mathbf{X})]$.

Mathematics subject classification (2020): 26E60, 41A60, 60E05, 62E20.
Keywords and phrases: Asymptotic expansion, power mean, expectation, moment, normal distribution.

* Corresponding author.

Paper JMI-17-27

Similary, for a positive random variable X we may also define distribution power mean of order r, if it exists, by

$$\mathfrak{m}_r = \left(\mathbb{E}[X^r]\right)^{1/r},$$

where harmonic, geometric, arithmetic, quadratic mean of a probability distribution are obtained for $r = -1, 0, 1, 2$ respectively:

$$\mathfrak{h} := \left(\mathbb{E}[\tfrac{1}{X}]\right)^{-1}, \quad \mathfrak{g} := e^{\mathbb{E}[\ln X]}, \quad \mathfrak{a} := \mathbb{E}[X] = \mu, \quad \mathfrak{q} := \left(\mathbb{E}[X^2]\right)^{\frac{1}{2}}.$$

Geometric and harmonic means have been studied within theory of financial mathematics. In establishing criteria for choosing among strategies, Latané [12] uses the fact that the final value of the investment converges in probability to the power of geometric mean. Later, Latané and Tuttle [13] established the geometric mean criterion for maximizing the income. In both papers the following approximation of geometric mean was used

$$\mathfrak{g} \simeq (\mu^2 - \sigma^2)^{\frac{1}{2}}, \tag{1}$$

which holds when the deviation is small compared to the arithmetic mean of the distribution.

Markowitz [14] presented two geometric mean approximations:

$$\ln \mathfrak{g} \simeq (\mu - 1) - \frac{1}{2}((\mu - 1)^2 + \sigma^2) \tag{2}$$

and

$$\ln \mathfrak{g} \simeq \ln \mu - \frac{\sigma^2}{2\mu^2}, \tag{3}$$

and gave the advantage to (3) when estimating the error of both approximations.

According to Renshaw [18], similar approximation to (3) was suggested by Johnson

$$\mathfrak{g} \simeq \mu - \frac{\sigma^2}{2\mu}. \tag{4}$$

Young and Trent [20] proposed the geometric mean approximation involving the first four sample moments:

$$G(\mathbf{X}) \simeq A(\mathbf{X}) - \frac{m_2(\mathbf{X}, A(\mathbf{X}))}{2A(\mathbf{X})} + \frac{m_3(\mathbf{X}, A(\mathbf{X}))}{3A(\mathbf{X})^2} - \frac{m_4(\mathbf{X}, A(\mathbf{X}))}{4A(\mathbf{X})^3}, \tag{5}$$

and examined its quality in context of convergence in mean. From here, Johnson's (4), Markowitz (3) and Latané's (1) approximations follow. Young and Trent empirically tested several geometric mean approximations including (1), (4) and (5) up to the third and fourth moment and concluded that approximations with two terms were better.

Along with the geometric mean criterion, stochastic dominance models were also developed for the same purpose. Jean [8] extended previous approximations to the infinite series representation of the geometric mean of probability distribution using Taylor series expansion:

$$\ln \mathfrak{g} = \ln \mu - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \mu_k \mu^{-k}. \tag{6}$$

Jean and Helms [10] presented series inferred from the approximations developed by Young and Trent as

$$G(\mathbf{X}) \simeq A(\mathbf{X}) + \sum_{k=1}^{\infty} (-1)^{k+1} \frac{m_k(\mathbf{X}, A(\mathbf{X}))}{kA(\mathbf{X})^k},$$

but it doesn't seem to converge to the geometric mean by adding more terms. In the same paper, the comparison of several approximations was also done and Jean series with two terms appeared to be the most accurate:

$$G(\mathbf{X}) \simeq \exp\left(\ln A(\mathbf{X}) - \frac{\sigma^2}{2A(\mathbf{X})^2}\right).$$

Jean [9] recognized ranking of harmonic means as necessary condition for stochastic dominance rankings. He found an infinite convergent series expansion of the harmonic mean

$$\frac{1}{\mathfrak{h}} = \frac{1}{b} + \sum_{k=1}^{\infty} \frac{k!}{b^{k+1}} F_{k+1}(b),$$

where b is the finite upper bound for range of X and F_k denotes the k-th integral of the probability density function.

Komarova and Rivin [11] studied the limiting behaviour of harmonic mean of X which is uniformly distributed on $[0,1]$, and for large n we have:

$$\mathbb{E}[H(\mathbf{X})] \sim \frac{1}{\ln n}, \quad n \to \infty.$$

Shi, Wang and Reid [19] obtained an approximation for the expectation of a function of a sum of random variables, up to a given order of $1/n$:

$$\mathbb{E}[f(\sum_{i=1}^{n} X_n)] = f_0 + \frac{f_2 \kappa_2}{2n} + \frac{3f_4 \kappa_2^2 + 4f_3 \kappa_3}{24n^2} + \mathscr{O}(|f_0| n^{-3}),$$

where $p(x)$ is probability density function of independently identical distributed (i.i.d.) random variables X_i, $K(t) = \ln \int e^{tx} p(x)\,\mathrm{d}x$, $\kappa_r = K^{(r)}(0)$ and $f_r = n^r f^{(r)}(n\kappa_1)$.

This method can also be used to obtain the asymptotic expansion of the expectation of harmonic mean when $n \to \infty$, but it is not applicable when the first moment of $1/X$ is infinite. Under some conditions on distribution X, Rao, Shi, Wu [17], found the approximation even in that case when the first moment can be infinite:

$$\mathbb{E}[H(\mathbf{X})] \sim \frac{1}{\ln n}\left(1 - \frac{1}{\sqrt{\ln n}}\right), \quad n \to \infty.$$

Komarova and Rivin [11] studied random matrices, explored the behaviour of the harmonic mean of i.i.d. random variables and presented limit theorems for the expectation of the harmonic mean for large n. For other possible applications of means in statistics, see [15] and references therein.

From all these results and examples, we may see that approximation of the power mean of probability distribution, the power mean of random sample and the expectation of power mean of random sample are useful in financial mathematics and other areas.

In this paper we will obtain estimations for the expectation of power mean of random sample and associate all three kinds of power means mentioned above. The method we propose follows from the asymptotic expansion of means which is applicable for sufficiently large value of μ. Presented approximation formulas are based on large values of data, rather than on large sample volume. These methods are especially useful when exact or even approximative value of such expectation is hard to obtain. Advantage of presented approximations will be shown numerically for random samples which have uniform and normal distribution. Finally, in order to compare our results with the known results from the introduction, we will analyse behaviour of our expansions for large sample volume.

2. Preliminaries and auxiliary results

Let us denote $\mathbf{e} = (1,1,\dots,1)$, with n units. Then n-tuple $(x_1,x_2,\dots,x_n)$ can be written in the form

$$(x_1,x_2,\dots,x_n) = (x+a_1,x+a_2,\dots,x+a_n) = x\mathbf{e}+\mathbf{a},$$

where x can be taken as a mean of $x_1,x_2,\dots,x_n$. The asymptotic behaviour of numerous classical and parametric means for large value of x was recently analysed in several papers, see for example [3, 4, 5, 6], and we will now mention some of these results which will be used in the sequel.

Let us denote in this section

$$m_k = m_k(\mathbf{a},0) = \frac{a_1^k+\ldots+a_n^k}{n} = M_k(\mathbf{a})^k, \qquad m_0 := 1.$$

THEOREM 1. ([4]) *General power mean has the following asymptotic expansion*

$$M_r(x\mathbf{e}+\mathbf{a}) = x\cdot\sum_{k=0}^{\infty} c_k(r,\mathbf{a})x^{-k}, \tag{7}$$

where $c_0 = 1$ *and*

$$c_k(r,\mathbf{a}) = \frac{1}{k}\sum_{j=1}^{k}\left[j\left(1+\frac{1}{r}\right)-k\right]\binom{r}{j}m_j c_{k-j}(r,\mathbf{a}), \quad k\in\mathbb{N}.$$

The first few coefficients are

$$\begin{aligned}
c_0(r,\mathbf{a}) &= 1,\\
c_1(r,\mathbf{a}) &= m_1,\\
c_2(r,\mathbf{a}) &= -\tfrac{1}{2}(r-1)(m_1^2-m_2),\\
c_3(r,\mathbf{a}) &= \tfrac{1}{6}(r-1)\big((2r-1)m_1^3-3(r-1)m_1m_2+(r-2)m_3\big),\\
c_4(r,\mathbf{a}) &= -\tfrac{1}{24}(r-1)\big((3r-1)(2r-1)m_1^4-6(r-1)(2r-1)m_1^2m_2\\
&\qquad +3(r-1)^2m_2^2+4(r-2)(r-1)m_1m_3-(r-3)(r-2)m_4\big).
\end{aligned}$$

Asymptotic expansions of special cases of power means can be obtained by substituting parameter r with appropriate values, even in the limiting case $r=0$ for geometric mean as it was proved in [6]. Since the corresponding formulas have simpler form with special values of parameter r, we state them separately.

Since for all x it holds

$$A(x\mathbf{e}+\mathbf{a}) = x + A(\mathbf{a}),$$

the asymptotic expansion for this mean has only these two terms.

COROLLARY 1. ([4]) *Geometric mean has the following asymptotic expansion*

$$G(x\mathbf{e}+\mathbf{a}) = x\cdot\sum_{k=0}^{\infty} c_k x^{-k},$$

where $c_0=1$ *and*

$$c_k = \frac{1}{k}\sum_{j=1}^{k}(-1)^{j-1}m_jc_{k-j}, \quad k\geqslant 1.$$

COROLLARY 2. ([4]) *Harmonic mean has the following expansion*

$$H(x\mathbf{e}+\mathbf{a}) = x\sum_{k=0}^{\infty} c_k x^{-k+1},$$

where coefficients are given by $c_0=1$ *and*

$$c_k = \sum_{j=1}^{k}(-1)^{j-1}m_jc_{k-j}, \quad k\geqslant 1.$$

COROLLARY 3. ([4]) *Quadratic mean has the following asymptotic expansion*

$$Q(x\mathbf{e}+\mathbf{a}) = x\cdot\sum_{k=0}^{\infty} c_k x^{-k},$$

where $c_0=1$, $c_1=m_1$ *and*

$$c_k = \left(\frac{3}{k}-2\right)m_1c_{k-1} + \left(\frac{3}{k}-1\right)m_2c_{k-2}, \qquad k\geqslant 2.$$

In particular, the first few coefficients are

$$G(x\mathbf{e}+\mathbf{a}) = x + m_1 + \tfrac{1}{2}(m_1^2 - m_2)x^{-1} + \tfrac{1}{6}(m_1^3 - 3m_1m_2 + 2m_3)x^{-2} + \tfrac{1}{24}(m_1^4 - 6m_1^2m_2 + 3m_2^2 + 8m_1m_3 - 6m_4)x^{-3} + \ldots \tag{8}$$

$$H(x\mathbf{e}+\mathbf{a}) = x + m_1 + (m_1^2 - m_2)x^{-1} + (m_1^3 - 2m_1m_2 + m_3)x^{-2} + (m_1^4 - 3m_1^2m_2 + m_2^2 + 2m_1m_3 - m_4)x^{-3} + \ldots \tag{9}$$

$$Q(x\mathbf{e}+\mathbf{a}) = x + m_1 + \tfrac{1}{2}(-m_1^2 + m_2)x^{-1} + \tfrac{1}{2}m_1(m_1^2 - m_2)x^{-2} + \tfrac{1}{8}(-5m_1^4 + 6m_1^2m_2 - m_2^2)x^{-3} + \ldots \tag{10}$$

In the sequel, we will also use the following property of the coefficients which was proved in the same paper.

THEOREM 2. ([4]) *Coefficients c_k, $k \in \mathbb{N}_0$, are homogeneous polynomials in variables $(a_1,\ldots,a_n)$ and have the following form:*

$$c_k(r,\mathbf{a}) = \sum_{\substack{\alpha_1,\alpha_2,\ldots,\alpha_k \geqslant 0 \\ \alpha_1+2\alpha_2+\cdots+k\alpha_k = k}} q_{\alpha_1,\ldots,\alpha_k}(r)m_1(\mathbf{a})^{\alpha_1}\cdots m_k(\mathbf{a})^{\alpha_k},$$

where

$$\sum_{\substack{\alpha_1,\alpha_2,\ldots,\alpha_k \geqslant 0 \\ \alpha_1+2\alpha_2+\cdots+k\alpha_k = k}} q_{\alpha_1,\ldots,\alpha_k}(r) = 0, \quad k \geqslant 2.$$

3. Expectation of power mean

Let X_1, $X_2,\ldots,$ X_n be independent copies of random variable X which satisfy following properties:

1) the distribution of X is known and it is symmetric with respect to some point μ, obviously the expectation of X,

2) μ is sufficiently large with respect to the essential support of X, in a sense that $\mathbb{P}(X<0) \to 0$ as $\mu \to \infty$,

3) $\mathbb{E}[(X-\mu)^k]$ is finite for $k \in \mathbb{N}$.

Let us denote centered random variables

$$A_i := X_i - \mu,$$

and random sample

$$\mathbf{A} = (A_1, A_2, \ldots, A_n).$$

In order to estimate the expectation $\mathbb{E}[M_r(\mathbf{X})]$, we shall use the asymptotic expansion (7):

$$\mathbb{E}[M_r(\mathbf{X})] = \mathbb{E}[M_r(x\mathbf{e}+\mathbf{A})] = x \cdot \sum_{k=0}^{\infty} \mathbb{E}[c_k(r,\mathbf{A})]x^{-k}, \quad x \to \infty. \tag{11}$$

Recall the moments

$$m_k = m_k(\mathbf{X}, \mu) = \frac{A_1^k + \ldots + A_n^k}{n}.$$

Then, for each i we have k-th central moment of variable X_i:

$$\mu_k = \mathbb{E}[m_k] = \mathbb{E}[A_i^k].$$

It holds $\mathbb{E}[A_i] = 0$ for all i and we also have $\mu_k = 0$ for all odd k, since by assumption the distribution of each A_i is symmetric.

LEMMA 1. *For all odd k we have $\mathbb{E}[c_k(r, \mathbf{A})] = 0$.*

Proof. Because of Theorem 2, it is sufficient to prove that

$$\mathbb{E}[m_1^{\alpha_1} \cdots m_k^{\alpha_k}] = 0.$$

After expanding the polynomial in the brackets into sum of monomials, each term will be of the form

$$A_1^{\beta_1} A_2^{\beta_2} \cdots A_n^{\beta_n},$$

where $\beta_1, \ldots, \beta_n \geqslant 0$ and $\beta_1 + \ldots + \beta_n = k$. Therefore, at least one of exponents is odd. Now,

$$\mathbb{E}[A_1^{\beta_1} A_2^{\beta_2} \cdots A_n^{\beta_n}] = \mathbb{E}[A_1^{\beta_1}]\mathbb{E}[A_2^{\beta_2}] \cdots \mathbb{E}[A_n^{\beta_n}] = 0. \quad \square$$

LEMMA 2. *It holds*

$$\begin{aligned}
\mathbb{E}[m_1^2] &= \frac{\mu_2}{n}, \\
\mathbb{E}[m_1^4] &= \frac{\mu_4}{n^3} + \frac{3(n-1)\mu_2^2}{n^3}, \\
\mathbb{E}[m_1^2 m_2] &= \frac{\mu_4}{n^2} + \frac{(n-1)\mu_2^2}{n^2}, \\
\mathbb{E}[m_2^2] &= \frac{\mu_4}{n} + \frac{(n-1)\mu_2^2}{n}, \\
\mathbb{E}[m_1 m_3] &= \frac{\mu_4}{n}.
\end{aligned}$$

Proof. We have

$$\mathbb{E}[m_1^2] = \frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^n A_i\right]^2 = \frac{1}{n^2}\left(\sum_{i=1}^n \mathbb{E}[A_i^2] + \sum_{i \neq j} \mathbb{E}[A_i]\mathbb{E}[A_j]\right) = \frac{\mu_2}{n}.$$

$$\mathbb{E}[m_1^4] = \frac{1}{n^4}\left(\sum_{i=1}^n \mathbb{E}[A_i^4] + 3\sum_{i \neq j} \mathbb{E}[A_i^2]\mathbb{E}[A_j^2]\right) = \frac{\mu_4}{n^3} + \frac{3(n-1)\mu_2^2}{n^3}.$$

The rest follows in a similar way. $\square$

Finally, using previous two lemmas, we can estimate the expectation of power mean with the first few terms from (11) in a following way.

THEOREM 3. *Expectation of the power mean has the following expansion:*

$$\mathbb{E}[M_r(\mathbf{X})] = \mu + \frac{d_2}{\mu} + \frac{d_4}{\mu^3} + \mathscr{O}(\mu^{-5}), \quad \mu \to \infty,$$

where μ *is the expectation of* X *and*

$$d_2 = \frac{r-1}{2} \cdot \frac{n-1}{n} \mu_2,$$

$$d_4 = -\frac{r-1}{24} \cdot \frac{n-1}{n} \left[3\mu_2^2 \left(\frac{(3r-1)(2r-1)}{n^2} - \frac{2(r-1)(2r-1)}{n} + (r-1)^2 \right) - \mu_4 \left(\frac{(3r-1)(2r-1)}{n^2} - \frac{(2r-1)(3r-5)}{n} + (r-2)(r-3) \right) \right].$$

In the sequel, we will show application to the geometric, harmonic and quadratic mean which follows from this theorem for $r = 0, -1, 2$, respectively. For the convenience of the reader, let us state them in the following corollaries. Note that they can also be deduced from (8), (9) and (10).

COROLLARY 4. *For the expectation of the geometric mean it holds:*

$$\mathbb{E}[G(\mathbf{X})] = \mu - \frac{\mu_2}{2}\left(1 - \frac{1}{n}\right)\mu^{-1} + \frac{1}{24}\left[3\mu_2^2 \frac{(n-1)^3}{n^3} + \mu_4 \frac{(1-n)(1-2n)(1-3n)}{n^3}\right]\mu^{-3} + \mathscr{O}(\mu^{-5}). \tag{12}$$

COROLLARY 5. *For the expectation of the harmonic mean it holds:*

$$\mathbb{E}[H(\mathbf{X})] = \mu - \left(1 - \frac{1}{n}\right)\mu_2\mu^{-1} + \left(1 - \frac{1}{n}\right)\left[\mu_2^2\left(1 - \frac{3}{n} + \frac{3}{n^2}\right) - \mu_4\left(1 - \frac{1}{n}\right)^2\right]\mu^{-3} + \mathscr{O}(\mu^{-5}).$$

COROLLARY 6. *For the expectation of the quadratic mean it holds:*

$$\mathbb{E}[Q(\mathbf{X})] = \mu + \frac{1}{2}\left(1 - \frac{1}{n}\right)\mu_2\mu^{-1} - \frac{1}{8}\left(1 - \frac{1}{n}\right)\left[\mu_2^2\left(1 - \frac{6}{n} + \frac{15}{n^2}\right) + \mu_4\left(\frac{1}{n} - \frac{5}{n^2}\right)\right]\mu^{-3} + \mathscr{O}(\mu^{-5}).$$

Note that the case for geometric mean is important since we can establish the quality of the approximation (12). Namely, for the geometric mean we have the exact formula for its expectation:

$$\mathbb{E}[G(\mathbf{X})] = \prod_{i=1}^{n} \mathbb{E}[X_i^{1/n}] = \left(\mathbb{E}[X^{1/n}]\right)^n. \tag{13}$$

Hence, for the specific distributions, we will be able to numerically check precision of our approximation formula which will be done in the next sections.

4. Uniform distribution

First, we shall apply the results from the previous section to the case of the uniformly distributed random variable. Here we assume that X has an uniform distribution on $[a,b]$, where $0<a<b$. Then

$$x=\mu=\mathbb{E}(X)=(a+b)/2,$$

random variables $A_1,\ldots,A_n$ are uniform on $\left[-\frac{b-a}{2},\frac{b-a}{2}\right]=:[-t,t]$ and we have

$$\mu_2=\mathbb{E}[A_1^2]=\frac{(b-a)^2}{12}=\frac{1}{3}t^2,\qquad \mu_4=\mathbb{E}[A_1^4]=\frac{(b-a)^4}{80}=\frac{1}{5}t^4.$$

With these central moments, we have the following consequence of Theorem 3.

COROLLARY 7. *Let X be a uniform variable on interval $[a,b]\subset\mathbb{R}^+$. Then the expectation of the power mean has the following expansion:*

$$\mathbb{E}[M_r(\mathbf{X})]=\mu+\frac{d_2}{\mu}+\frac{d_4}{\mu^3}+\mathscr{O}(\mu^{-5}),\quad \mu\to\infty,$$

where

$$d_2=\frac{r-1}{6}\cdot\frac{n-1}{n}t^2,$$
$$d_4=-\frac{r-1}{360}\cdot\frac{n-1}{n}\left[\frac{2(3r-1)(2r-1)}{n^2}-\frac{(2r-1)(r+5)}{n}+(2r^2+5r-13)\right]t^4.$$

As mentioned before, we will discuss the results for the geometric, harmonic and quadratic mean. In particular, for $r=0,-1,2$ we get the expectations for the geometric, harmonic and quadratic mean, when $\mu\to\infty$:

$$\mathbb{E}[G(\mathbf{X})]=\mu-\frac{1}{6}\cdot\frac{n-1}{n}t^2\mu^{-1}+\frac{1}{360}\cdot\frac{n-1}{n}\left(\frac{2}{n^2}+\frac{5}{n}-13\right)t^4\mu^{-3}+\mathscr{O}(\mu^{-5}),\quad(14)$$

$$\mathbb{E}[H(\mathbf{X})]=\mu-\frac{1}{3}\cdot\frac{n-1}{n}t^2\mu^{-1}+\frac{1}{90}\cdot\frac{n-1}{n}\left(\frac{12}{n^2}+\frac{6}{n}-8\right)t^4\mu^{-3}+\mathscr{O}(\mu^{-5}),\quad(15)$$

$$\mathbb{E}[Q(\mathbf{X})]=\mu+\frac{1}{8}\cdot\frac{n-1}{n}t^2\mu^{-1}-\frac{1}{360}\cdot\frac{n-1}{n}\left(\frac{30}{n^2}-\frac{21}{n}+5\right)t^4\mu^{-3}+\mathscr{O}(\mu^{-5}).\quad(16)$$

For the geometric mean we also have the exact formula for the expectation which follows from (13):

$$\mathbb{E}[G(\mathbf{X})]=\left[\frac{n}{n+1}\cdot\frac{b^{1+1/n}-a^{1+1/n}}{b-a}\right]^n.\quad(17)$$

Numerical results will be shown in the next three tables for sample size of $n=10,20,100$ on intervals $[100,110]$ and $[1000,1010]$. In Table 1 we have the following data: the exact value of the mean (17), asymptotic approximation (14) up to order

Table 1: *Expectation of the geometric mean for uniform distribution*

		$n=10$	$n=20$	$n=100$
$a=100$ $b=110$	exact	104.964268853	104.962283411	104.960695039
	asym	104.964268869	104.962283429	104.960695059
	MC4	104.973283147	104.964084374	104.955314488
	MC5	104.961915767	104.963161009	104.960514008
$a=1000$ $b=1010$	exact	1004.9962686375058	1004.9960613391589	1004.9958955004609
	asym	1004.9962686375060	1004.9960613391591	1004.9958955004611
	MC4	1004.9795366485	1005.0035736128	1004.9924337216
	MC5	1004.9975345659	1004.9976453115	1004.9960081227

μ^{-3} and Monte-Carlo simulation with 10^4 and 10^5 repetitions (MC4 and MC5 respectively). As we can see, for larger data we obviously obtain better results where the error of asymptotic expansion is less than 10^{-12}.

Similar numerical calculations can be made for harmonic and quadratic mean using (15) and (16) (up to order μ^{-3}), which is presented in the Tables 2 and 3. Note that here we don't have the exact value for the approximation so we compare it with Monte-Carlo simulations.

Table 2: *Expectation of the harmonic mean for uniform distribution*

		$n=10$	$n=20$	$n=100$
$a=100$ $b=110$	asym	104.928532124	104.924559464	104.921381424
	MC4	104.939539156	104.934724836	104.923835838
	MC5	104.923631934	104.924903744	104.921067563
$a=1000$ $b=1010$	asym	1004.99253727	1004.99212267	1004.99179099
	MC4	1004.99600812	1004.97657863	1005.00017099
	MC5	1004.99473527	1004.98943082	1004.99329475

5. Normal distribution

Let us now show application to the random variable X which has normal distribution with mean $\mu > 0$ and standard deviation σ. Then it holds

$$\mu_2 = \sigma^2, \qquad \mu_4 = 3\sigma^4,$$

and the result of the Theorem 3 reads as:

Table 3: *Expectation of the quadratic mean for uniform distribution*

		$n=10$	$n=20$	$n=100$
$a=100$ $b=110$	asym	105.035709967	105.037692678	105.039278598
	MC4	105.019672348	105.033824351	105.036214548
	MC5	105.031103452	105.038155429	105.038242199
$a=1000$ $b=1010$	asym	1005.00373134	1005.00393863	1005.00410447
	MC4	1005.01124638	1005.00192709	1005.00280217
	MC5	1005.00191186	1005.00619483	1005.00485681

COROLLARY 8. *Let X be a normal variable with mean $\mu>0$ and standard deviation σ. Then the expectation of the power mean has the following expansion:*

$$\mathbb{E}[M_r(\mathbf{X})]=\mu+\frac{d_2}{\mu}+\frac{d_4}{\mu^3}+\mathscr{O}(\mu^{-5}),\quad \mu\to\infty,$$

where

$$d_2=\frac{r-1}{2}\cdot\frac{n-1}{n}\sigma^2,$$
$$d_4=-\frac{r-1}{8}\cdot\frac{n-1}{n}\left[\frac{(2r-1)(r-3)}{n}+(3r-5)\right]\sigma^4.$$

Specifically, for $r=0,-1,2$ we get the expectations for the geometric, harmonic and quadratic mean:

$$\mathbb{E}[G(\mathbf{X})]=\mu-\frac{n-1}{2n}\sigma^2\mu^{-1}-\frac{(n-1)(5n-3)}{8n^2}\sigma^4\mu^{-3}+\mathscr{O}(\mu^{-5}), \tag{18}$$

$$\mathbb{E}[H(\mathbf{X})]=\mu-\frac{n-1}{n}\sigma^2\mu^{-1}-\frac{(n-1)(2n-3)}{n^2}\sigma^4\mu^{-3}+\mathscr{O}(\mu^{-5}), \tag{19}$$

$$\mathbb{E}[Q(\mathbf{X})]=\mu+\frac{n-1}{2n}\sigma^2\mu^{-1}-\frac{(n-1)(n-3)}{8n^2}\sigma^4\mu^{-3}+\mathscr{O}(\mu^{-5}). \tag{20}$$

Here we have to note that $M_r(\mathbf{X})$ is defined only for the positive entries, but normal variable is defined on $\mathbb{R}$ and can take negative vaules. However, since we are studying asymptotic behaviour for large values of μ, probability $\mathbb{P}(X<0)$ does not have an effect on the results. Namely, let $\phi(x)$ be probability density function of the standard (unit) normal variable Z:

$$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2},\quad x\in\mathbb{R}.$$

Then it holds:

$$\mathbb{P}(Z>x)=\int_x^\infty\phi(u)\,\mathrm{d}u\leqslant\int_x^\infty\frac{u}{x}\phi(u)\,\mathrm{d}u=\frac{1}{x}\phi(x).$$

Now for general variable X we have

$$\mathbb{P}(X<0)=\mathbb{P}(Z<-\frac{\mu}{\sigma})=\mathbb{P}(Z>\frac{\mu}{\sigma})\leqslant\frac{1}{\sqrt{2\pi}}\frac{\sigma}{\mu}e^{-\frac{\mu^2}{2\sigma^2}}\to 0,\quad \mu\to\infty. \tag{21}$$

The geometric mean can again serve as a control point for our approximations since (13) holds. Explicit formulas for the p-th order moments, $\mathbb{E}[X^p]$, of the normal distribution are known and they involve special functions on $\mathbb{C}$. But because of (21), we have

$$\mathbb{E}[X^p]=\mathbb{E}[|X|^p]+\mathscr{O}(\tfrac{\sigma}{\mu}e^{-\frac{\mu^2}{2\sigma^2}}),$$

and therefore we can use a known formula for absolute moments involving special functions on $\mathbb{R}$, see [7]:

$$\mathbb{E}[|X|^p]=\tfrac{1}{\sqrt{\pi}}\sigma^p 2^{p/2}\Gamma(\tfrac{p+1}{2})\,{}_1F_1(-\tfrac{p}{2},\tfrac{1}{2},-\tfrac{1}{2}(\tfrac{\mu}{\sigma})^2),$$

where ${}_1F_1$ is confluent hypergeometric function. Finally, for large μ we have:

$$\mathbb{E}[G(\mathbf{X})]\approx\sigma\sqrt{\tfrac{2}{\pi^n}}\left(\Gamma(\tfrac{n+1}{2n})\,{}_1F_1(-\tfrac{1}{2n},\tfrac{1}{2},-\tfrac{1}{2}(\tfrac{\mu}{\sigma})^2)\right)^n. \tag{22}$$

In the next three tables we will show numerical results in a same way as in previous section. We use sample sizes $n=10,20,100$ and normal distribution with $\mu=105$ and $\mu=1005$ with $\sigma=5/\sqrt{3}$ which corresponds with mean and standard deviation of uniform distribution from before. Note that in this case formula (22) has error less than 10^{-289} so in the Table 4 we call it exact. Asymptotic formula (18) is up to order μ^{-3} and as we can see, the error of asymptotic expansion is again less than 10^{-10} and 10^{-12} respectively, and of course it is better for larger data.

Table 4: *Expectation of the geometric mean for normal distribution*

		$n=10$	$n=20$	$n=100$
$\mu=105$, $\sigma=\frac{5}{\sqrt{3}}$	exact	104.9642539163	104.9622669495	104.9606772947
	asym	104.9642539952	104.9622670375	104.9606773904
	MC4	104.9765259849	104.9675671710	104.9610958531
	MC5	104.9668869488	104.9632025986	104.9606851131
$\mu=1005$, $\sigma=\frac{5}{\sqrt{3}}$	exact	1004.996268620542	1004.996061320465	1004.995895480310
	asym	1004.996268620543	1004.996061320467	1004.995895480311
	MC4	1004.988085274467	1004.989778581215	1004.998768562838
	MC5	1004.995026427079	1004.998085981116	1004.996087905275

In Tables 5 and 6 we show numerical results of asymptotic formulas up to order μ^{-3} for harmonic (19) and quadratic mean (20). Since we do not have the exact value, we compare with Monte-Carlo simulations MC4 and MC5 same as before.

Table 5: *Expectation of the harmonic mean for normal distribution*

		$n = 10$	$n = 20$	$n = 100$
$\mu = 105$	asym	104.9284796458	104.9244977444	104.9213115754
$\sigma = \frac{5}{\sqrt{3}}$	MC4	104.9515265759	104.9377540489	104.9234181353
	MC5	104.9301124738	104.9237058313	104.9206201061
$\mu = 1005$	asym	1004.9925372088	1004.9921225995	1004.9917909114
$\sigma = \frac{5}{\sqrt{3}}$	MC4	1004.9961238421	1004.9906718465	1004.9880447393
	MC5	1004.9939451195	1004.9937613795	1004.9914757101

Table 6: *Expectation of the quadratic mean for normal distribution*

		$n = 10$	$n = 20$	$n = 100$
$\mu = 105$	asym	105.0357095616	105.0376923576	105.0392785134
$\sigma = \frac{5}{\sqrt{3}}$	MC4	105.0429017594	105.0406447484	105.0413994623
	MC5	105.0309596004	105.0368516623	105.0386420319
$\mu = 1005$	asym	1005.0037313379	1005.0039386332	1005.0041044694
$\sigma = \frac{5}{\sqrt{3}}$	MC4	1004.9989585130	1004.9985805634	1005.0025865899
	MC5	1005.0029674739	1005.0013258903	1005.0053547868

REMARK 1. As a consequence of symmetry of distribution and Lemma 1, coefficients in Lemma 2 and subsequent Theorems have simpler form. Similar analysis could be done without the assumption of symmetry.

REMARK 2. In the case of unknown distribution of X, similar results for power mean of random sample could be obtained by approximating distribution mean μ and higher order distribution moments with a sample mean $A(\mathbf{X})$ and corresponding sample moments.

6. Large sample volume

Altough this is not the main goal of our paper, in order to compare our expansions with known results from the introduction we shall make analysis for large n.

We may apply the strong law of large numbers on i.i.d. variables X_i^r, $r \neq 0$, to see that

$$M_r(\mathbf{X})^r = \frac{1}{n}\sum_{i=1}^{n} X_i^r \xrightarrow{\text{a.s.}} \mathbb{E}[X^r] = \mathfrak{m}_r^r, \quad n \to \infty,$$

hence

$$M_r(\mathbf{X}) \xrightarrow{\text{a.s.}} \mathfrak{m}_r, \quad n \to \infty,$$

and

$$\mathbb{E}[M_r(\mathbf{X})] \to \mathfrak{m}_r, \quad n \to \infty.$$

Same reasoning can be done for i.i.d. variables $\ln X_i$ (case $r = 0$), where we have

$$\mathbb{E}[G(\mathbf{X})] \to \mathfrak{g}, \quad n \to \infty.$$

Therefore, for all r, the distribution power mean $\mathfrak{m}_r$ may be approximated with the expectation of the power mean of random sample for large n. Then from Theorem 3, it follows, as $n \to \infty$:

$$\mathfrak{m}_r \approx \mu + \frac{r-1}{2}\mu_2\mu^{-1} - \frac{r-1}{24}\left[3\mu_2^2(r-1)^2 - \mu_4(r-2)(r-3)\right]\mu^{-3}. \tag{23}$$

Note that for $r = 0$, we have:

$$\mathfrak{g} \approx \mu - \frac{1}{2}\mu_2\mu^{-1} + \frac{1}{8}\left(\mu_2^2 - 2\mu_4\right)\mu^{-3}, \tag{24}$$

which corresponds with the terms obtained from the formula (6) and it has better numerical precision than other formulas from the introduction.

Let us now show examples of these approximations for the uniform and normal distribution.

6.1. Uniform distribution

Notice that for X uniformly distributed, we may easily calculate the distribution geometric mean in the following way:

$$\ln\mathfrak{g} = \mathbb{E}[\ln X] = \int_a^b \frac{\ln x}{b-a}\,\mathrm{d}x = \frac{b\ln b - a\ln a}{b-a} - 1,$$

wherefrom it follows

$$\mathfrak{g} = \frac{1}{e}\left(\frac{a^a}{b^b}\right)^{\frac{1}{a-b}}.$$

We may conclude that $\mathfrak{g}$ equals the identric mean (the special case of Stolarsky mean) whose asymptotic expansion for large values of a and b has been derived in [5]. With $a = x - t$ and $b = x + t$, we have

$$\mathfrak{g} = I(a,b) = x - \frac{1}{6}t^2x^{-1} - \frac{13}{360}t^4x^{-3} + \mathscr{O}(x^{-5}),$$

which is exactly the same as (24) with $x = \mu$, and μ_2 and μ_4 were obtained in Section 4.

Analogously as with geometric mean, we have

$$\mathfrak{h} = \left(\frac{1}{b-a}\int_a^b \frac{1}{x}\,\mathrm{d}x\right)^{-1} = \frac{b-a}{\ln b - \ln a}.$$

Now $\mathfrak{h}$ equals the logarithmic mean $L(a,b)$ whose asymptotic expansion for large values of a and b has also been derived in [5]. With $a = x - t$ and $b = x + t$, we have

$$\mathfrak{h} = L(a,b) = x - \frac{1}{3}t^2x^{-1} - \frac{4}{45}t^4x^{-3} + \mathscr{O}(x^{-5}),$$

which is again same as (23) for $r = -1$.

6.2. Normal distribution

When X is normally distributed, the distribution power means are not representable in terms of elementary functions and are not easy to compute. In this case the discussion from the beginning of this section justifies to approximate $\mathfrak{g}$, $\mathfrak{h}$, $\mathfrak{q}$ and other distribution power means, using the expressions (18), (19) and (20) when $n \to \infty$, that is with negative powers of n neglected:

$$\begin{aligned}\mathfrak{g} &\approx \mu - \frac{1}{2}\sigma^2\mu^{-1} - \frac{5}{8}\sigma^4\mu^{-3},\\ \mathfrak{h} &\approx \mu - \sigma^2\mu^{-1} - 2\sigma^4\mu^{-3},\\ \mathfrak{q} &\approx \mu + \frac{1}{2}\sigma^2\mu^{-1} - \frac{1}{8}\sigma^4\mu^{-3}.\end{aligned}$$

In general, either from (23) or from Corollary 8, for normal distribution we have

$$\mathfrak{m}_r \approx \mu + \frac{r-1}{2}\sigma^2\mu^{-1} - \frac{(r-1)(3r-5)}{8}\sigma^4\mu^{-3}.$$

REFERENCES

[1] M. BARCZY, P. BURAI, *Limit theorems for Bajraktarevic and Cauchy quotient means of independent identically distributed random variables*, Aequationes Math. **96** (2022), 279–305.

[2] R. L. BERGER, G. CASELLA, *Deriving generalized means as least squares and maximum likelihood estimates*, Amer. Statist. **46** (1992), 279–282.

[3] T. BURIĆ, N. ELEZOVIĆ, *Asymptotic expansion of the arithmetic-geometric mean and related inequalities*, J. Math. Inequal. **9** (4) (2015), 1181–1190, dx.doi.org/10.7153/jmi-09-90.

[4] N. ELEZOVIĆ, L. MIHOKOVIĆ, *Asymptotic behaviour of power means*, Math. Inequal. Appl. **19** (4) (2016), 1399–1412, doi:10.7153/mia-19-103.

[5] N. ELEZOVIĆ, L. VUKŠIĆ, *Asymptotic expansions of bivariate classical means and related inequalities*, J. Math. Inequal. **8** (4) (2014), 707–724, https://doi.org/10.1016/j.amc.2014.02.026.

[6] N. ELEZOVIĆ, L. VUKŠIĆ, *Asymptotic expansions and comparison of bivariate parameter means*, Math. Inequal. Appl. **17** (4) (2014), 1225–1244, dx.doi.org/10.7153/mia-17-91

[7] I. GRADSHTEYN, I. RYZHIK, *Table of integrals, series, and products*, 6th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press (2000).

[8] W. H. JEAN, *The Geometric Mean and Stochastic Dominance*, Journal of Finance, **35** (1980), 151–158, https://doi.org/10.2307/2327187.

[9] W. H. JEAN, *The Harmonic Mean and Other Necessary Conditions for Stochastic Dominance*, The Journal of Finance, **39** (2) (1984), 527–534, https://doi.org/10.2307/2327876.

[10] W. H. JEAN, B. P. HELMS, *Geometric Mean Approximations*, The Journal of Financial and Quantitative Analysis, **18** (3) (1983), 287–293, https://doi.org/10.2307/2330720.

[11] N. L. KOMAROVA, I. RIVIN, *Harmonic mean, random polynomials and stochastic matrices*, Adv. Appl. Math **31** (2) (2003), 501–526, https://doi.org/10.1016/S0196-8858(03)00023-X.

[12] H. A. LATANÉ, *Criteria for Choice among Risky Ventures*, Journal of Political Economy, **67** (1959), 144–155, https://doi.org/10.1142/9789814293501_0004.

[13] H. A. LATANÉ AND D. L. TUTTLE, *Criteria for Portfolio Building* The Journal of Finance, **22**(3)(1967), 359–373, https://doi.org/10.2307/2978890.

[14] H. M. MARKOWITZ, *Portfolio Selection: Efficient Diversification of Investments*, Chapter VI Return in the long run, Yale University Press (1959).

[15] N. NORRIS, *General Means and Statistical Theory*, Amer. Statist. **30** (1) (1976), 8–12.

[16] A. G. PAKES, *On the convergence of moments of geometric and harmonic means*, Stat. Neerl. **53** (1999), 96–110.

[17] C. R. RAO, X. SHI, Y. WU, *Approximation of the expected value of the harmonic mean and some applications*, Proceedings of the National Academy of Sciences **111** (44) (2014), 15681–15686, https://doi.org/10.1073/pnas.1412216111.

[18] E. F. RENSHAW, *Portfolio Balance Models in Perspective: Some Generalizations That Can be Derived From the Two Asset Case*, Journal of Financial and Quantitative Analysis, **2** (2) (1967), 138–139, https://doi.org/10.2307/2329898.

[19] X. SHI, X. S. WANG, N. REID, *Saddle point approximation of nonlinear moments*, Statistica Sinica, **24** (4) (2014), 1597–1611, http://dx.doi.org/10.5705/ss.2012.333.

[20] W. E. YOUNG, R. H. TRENT, *Geometric Mean Approximations of Individual Security and Portfolio Performance*, Journal of Financial and Quantitative Analysis, **4** (1996), 179–199, https://doi.org/10.2307/2329839.

(Received January 5, 2023)

Tomislav Burić
University of Zagreb
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
e-mail: tomislav.buric@fer.hr

Neven Elezović
University of Zagreb
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
e-mail: neven@element.hr

Lenka Mihoković
University of Zagreb
Faculty of Electrical Engineering and Computing
Unska 3, 10000 Zagreb, Croatia
e-mail: lenka.mihokovic@fer.hr