CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): C.R. Rao, Xiaoping Shi and Yuehua Wu

Title of the Report: **Approximation of expected value of harmonic mean and some applications**

Research Report No.: **RR2014-17**

Date: **July 3, 2014**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

Approximation of the expected value of the harmonic mean and some applications

Calyampudi Radhakrishna Rao[a,b,1], Xiaoping Shi[c], and Yuehua Wu[c]

[a]Department of Biostatistics, University at Buffalo, The State University of New York, Buffalo, NY 14221-3000; [b]CRRAO Advanced Institute of Mathematics, Statistics And Computer Science, Hyderabad-500046, India; and [c]Department of Mathematics and Statistics, York University, Toronto, ON, Canada M3J 1P3

Contributed by Calyampudi R. Rao, July 1, 2014 (sent for review June 5, 2014)

Although the harmonic mean (HM) is mentioned in textbooks along with the arithmetic mean (AM) and the geometric mean (GM) as three possible ways of summarizing the information in a set of observations, its appropriateness in some statistical applications is not mentioned in textbooks. During the last 10 y a number of papers were published giving some statistical applications where HM is appropriate and provides a better performance than AM. In the present paper some additional applications of HM are considered. The key result is to find a good approximation to $E(H_n)$, the expectation of the harmonic mean of n observations from a probability distribution. In this paper a second-order approximation to $E(H_n)$ is derived and applied to a number of problems.

harmonic mean | second-order approximation | arithmetic mean | image denoising | marginal likelihood

The harmonic mean H_n of n observations $Z_1, \ldots, Z_n$ drawn from a population is defined by

$$H_n = \frac{n}{\sum_{i=1}^{n} \frac{1}{Z_i}}. \quad [1]$$

There have been a number of applications of the harmonic mean in recent papers. A more general version of H_n with weights $w_1, \ldots, w_n$ is

$$H_n(\boldsymbol{w}) = \frac{\sum_{i=1}^{n} w_i}{\sum_{i=1}^{n} \frac{w_i}{Z_i}}. \quad [2]$$

where $\boldsymbol{w} = (w_1, \ldots, w_n)^T$. The harmonic mean H_n is used to provide the average rate in physics and to measure the price ratio in finance as well as the program execution rate in computer engineering. Some statistical applications of the harmonic mean are given in refs. 1-4, among others. $H_n(\boldsymbol{w})$ has been used in evaluation of the portfolio price-to-earnings ratio value (ref. 5, p. 339) and the signal-to-interference-and-noise ratio (6) among others. The asymptotic properties of H_n including the asymptotic expansion of $E(H_n)$ are investigated in refs. 7 and 8 by either assuming that some moments of $1/Z_i$ are finite or that Z_i s follow the Poisson distribution. It is noted that recent papers (9, 10) enable one to use saddle-point approximation to give the asymptotic expansion of $E(H_n)$ to any given order of $1/n$ for some constants $c_0, c_1, c_2, \ldots$, i.e.,

$$E(H_n) = c_0 + \frac{c_1}{n} + \frac{c_2}{n^2} + \cdots. \quad [3]$$

However, such methods are not applicable for obtaining the asymptotic expansion of H_n when the first moment of $1/Z_i$ is infinite. In ref. 3, Z_i s are assumed to follow a uniform distribution in the interval $(0,1)$, i.e., $U(0,1)$, motivated by learning theory. Using the property that the inverse of H_n converges to the stable law, ref. 3 showed that

$$E(H_n) \sim \frac{1}{\log(n)}, \quad [4]$$

where the symbol "$\sim$" means asymptotic equivalence as $n \to \infty$. Our interest in this paper is to determine the second term in the asymptotic expansion of $E(H_n)$ or the general version $E(H_n(\boldsymbol{w}))$ under more general assumptions on distributions of Z_i s. We show that under mild assumptions,

$$E(H_n) \sim \frac{1}{\log(n)} \left\{ 1 + \frac{c_1}{\sqrt{\log(n)}} \right\}, \quad [5]$$

where the constant c_1 will be given. In addition, we use the approach for obtaining **[5]** to the case that the first moment of $1/Z_i$ is finite, motivated by evaluation of the marginal likelihood in ref. 11.

Approximations

We derive the asymptotic approximation of $E(H_n)$ when the first moment of $1/Z_i$ is not finite. Let $\{Z_i\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with possible infinite first moment. Suppose that there exist constants A_n and B_n, such that the distribution $F_n(x)$ of

$$X_n = \frac{1/Z_1 + 1/Z_2 + \cdots + 1/Z_n}{B_n} - A_n \quad [6]$$

converges weakly to a nondegenerate distribution $F(x)$ such that

$$F(x) = \frac{d_1 + o(1)}{|x|^{\alpha}} \text{ as } x \to -\infty, \quad [7]$$

$$1 - F(x) = \frac{d_2 + o(1)}{|x|^{\alpha}} \text{ as } x \to \infty, \quad [8]$$

where α, d_1, and d_2 are constants with $0 \le \alpha < 2$, $d_1, d_2 \ge 0$, and $d_1 + d_2 > 0$, respectively. The set of all distributions converging to $F(x)$ is called the domain of attraction of $F(x)$. It is known that only stable laws with index $\alpha (0 \le \alpha < 2)$ have the nonempty domains of attraction as shown by refs. 12 (chap. 7) and 13 (chap. 2).

STATISTICS

Significance

The harmonic mean (HM) filter is better at removing positive outliers than the arithmetic mean (AM) filter. There are especially difficult issues when an accurate evaluation of expected HM is needed such as, for example, in image denoising and marginal likelihood evaluation. A major challenge is to develop a higher-order approximation of the expected HM when the central limit theorem is not applicable. A two-term approximation of the expected HM is derived in this paper. This approximation enables us to develop a new filtering procedure to denoise the noisy image with an improved performance, and construct a truncated HM estimator with a faster convergence rate in marginal likelihood evaluation.

Author contributions: C.R.R. designed research; C.R.R., X.S., and Y.W. performed research; X.S. and Y.W. analyzed data; and C.R.R., X.S., and Y.W. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. Email: crr1@psu.edu.

Assume that there is a positive constant d_3 which does not depend on n such that

$$X_n + A_n \geq d_3 > 0. \quad [9]$$

We further assume a uniform rate of convergence of $F_n(x)$ to $F(x)$ such that

$$\sup_x |F_n(x) - F(x)| = o(n^{-\beta}), \quad [10]$$

for some positive constant $\beta < 1$. Our assumptions are mild. Ref. 14 showed that $\sup_x|F_n(x) - F(x)|$ has the rate of $o\{n^{-1}\log(n)\}$ under some assumptions.

We have the following asymptotic approximation of $E(H_n)$:

Theorem 1. *Assume that conditions* [**7**]–[**10**] *are satisfied and* $A_n = \log(n)$, $B_n = n$, $\alpha = 1$, $d_1 = 0$, *and* $d_2 = 1$. *Then we have the following first approximation:*

$$E(H_n) = E(X_n + \log n)^{-1} = \ell_n^{-2} - \ell_n^{-3} + o(\ell_n^{-3}), \quad [11]$$

where $\ell_n = \sqrt{\log(n)}$.

The proof is given in *Appendix: Proof of Theorem 1*. Because $n^{-\beta}$ in [**10**] is smaller than the remaining terms in [**11**], the coefficients of both ℓ_n^{-2} and ℓ_n^{-3} are independent of β in [**11**].

Remark 1: For an extension of *Theorem 1* to the weighted harmonic mean in [**2**], we consider the following partial sum:

$$X_n(\mathbf{w}) = \frac{w_1/Z_1 + w_2/Z_2 + \cdots + w_n/Z_n}{W_n} - A_n, \quad [12]$$

where $W_n = (\sum_{i=1}^{n}|w_i|^\alpha)^{1/\alpha}$. Motivated by ref. 15, we may assume the following two conditions on the weights w_i s:

$$\max_{1\leq i\leq n} |w_i| = o(W_n), \quad [13]$$

and the characteristic function of $1/Z_i$ in [**6**] satisfies that

$$1 - c|t|^\alpha + o(|t|^\alpha) \text{ as } t \to 0. \quad [14]$$

Under the conditions [**13**] and [**14**], ref. 15 showed that the distribution of $X_n(\mathbf{w})$ converges to a stable distribution with characteristic function $\exp(-c|t|^\alpha)$. For example, if Z_is follow uniform distribution $U(0,1)$, the condition [**14**] is satisfied when $A_n = \log n$ and $\alpha = 1$. Following the proof of *Theorem 1*, it can be shown that

$$E\{H_n(\mathbf{w})\} = \ell_n^{-2} - \ell_n^{-3} + o(\ell_n^{-3}), \quad [15]$$

where $\ell_n = \sqrt{\log(n)}$. It is noted that the weights in [**2**] do not have to be nonnegative, but must satisfy both conditions [**9**] and [**13**].

By *Theorem 1*, c_1 in [**5**] has the value −1. It is noted that *Theorem 1* holds true if Z_i s follow a uniform distribution $U(0,1)$.

A higher-order approximation may be similarly obtained but extra conditions on $F_n(x)$ in [**7**] and [**8**] may be needed. In view of the proof of *Theorem 2.1* given in *Appendix: Proof of Theorem 1*, the higher-order term should be $\ell_n^{-4}\log(\ell_n)$. Because it is difficult to obtain the coefficient of this term theoretically, it may be constructed empirically. As a demonstration, we consider the case where Z_i s follow a uniform distribution $U(0,1)$. We perform Monte Carlo simulation with 1,000,000 replications of n independent observations from standard uniform distribution $U(0,1)$ for different values of $n = 10, 15, 20, \ldots, 200$. The coefficient of $\ell_n^{-4}\log(\ell_n)$ is estimated to be 0.5673 by fitting the simulated data to the following model by least squares:

$$\log(n)H_n - 1 + \frac{1}{\sqrt{\log(n)}} = \beta\frac{\log[\log(n)]}{\log(n)}.$$

Thus, we obtain the following approximation:

$$E(H_n) \sim \ell_n^{-2} - \ell_n^{-3} + 0.5673\log(\log(n))\ell_n^{-4}. \quad [16]$$

As in ref. 3, suppose that Z_i s follow a uniform distribution $U(0,1)$. The distribution of $Y_i = 1/Z_i$ is easily seen to be given by

$$P(Y \leq t) = (1 - 1/t)I(t \geq 1),$$

where $I(\cdot)$ is an indicator function. It is well known that the mean of Y_i is infinite but $EY_i^r < \infty$ for $r < 1$. By considering the limit stable distribution with index $\alpha = 1$ of the distribution of X_n for $A_n = \log(n)$ and $B_n = n$, ref. 3 obtained the result [**4**], which is

$$E\{[\log(n)]H_n\} \sim 1. \quad [17]$$

According to our *Theorem 1* and the approximation [**16**],

$$E\{[\log(n)]H_n\} \sim 1 - \frac{1}{\sqrt{\log(n)}}, \quad [18]$$

$$E\{[\log(n)]H_n\} \sim 1 - \frac{1}{\sqrt{\log(n)}} + \frac{0.5673\log[\log(n)]}{\log(n)}. \quad [19]$$

Fig. 1 displays the approximations given in [**17**]–[**19**] compared with the sample mean of 1,000,000 replications of n independent observations from the uniform distribution $U(0,1)$ that serves as a proxy for the exact value of $E(H_n)$. Here n takes values 10, 15, 20, …, and 200. From Fig. 1, it can be seen that the approximation [**18**] is better than the approximation [**17**]. Although the approximation [**19**] is purely empirical, this empirical exercise basically achieves the desired result as shown in Fig. 1; it clearly gives much better approximation of $E\{[\log(n)]H_n\}$ than its other two counterparts.

We now consider the case that $\alpha > 1$. In this case, $B_n = n^{1/\alpha}$ and $A_n = E(1/Z_1)n^{1-1/\alpha}$. Thus, we have

$$H_n = \frac{n^{1-1/\alpha}}{X_n + n^{1-1/\alpha}E(1/Z_1)}. \quad [20]$$

In light of the proof of *Theorem 1*, we have the following asymptotic approximation of $E(H_n)$:

Theorem 2. *Assume that conditions in* [**7**]–[**10**] *are satisfied and* $A_n = E(1/Z_1)n^{1-1/\alpha}$, $B_n = n^{1/\alpha}$, $\alpha > 1$, $d_1 = 0$, and $d_2 = 1$; *then we have the following approximation:*

$$E(H_n) = \ell_n^{-2} + \ell_n^{-3} + o(\ell_n^{-3}), \quad [21]$$

where $\ell_n = \sqrt{A_n}$.

Remark 2: A similar result as in *Theorem 2* can be obtained for the weighted harmonic mean in [**2**] by assuming that conditions [**13**] and [**14**] are satisfied with $\alpha > 1$ and $A_n = E(1/Z_1)\sum_{i=1}^{n} w_i/W_n$. It can be shown that

$$E(H_n(\mathbf{w}) = \{\ell_n^{-2} + \ell_n^{-3} + o(\ell_n^{-3})\}\sum_{i=1}^{n} w_i/W_n, \quad [22]$$

where $\ell_n = \sqrt{A_n}$.

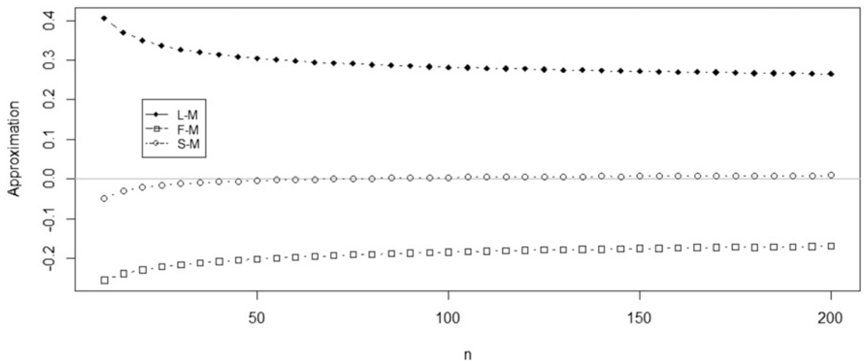

Fig. 1. Comparisons of three approximations of $E\{[\log(n)]H_n\}$ with respect to the sample mean (denoted by $\mathcal{M}$) of $[\log(n)]H_n$ with 1,000,000 replications of n independent observations from $U(0,1)$ for $n = 10, 15, 20, \ldots, 200$. (*i*) "**L-M**" denotes the approximations of $E\{[\log(n)]H_n\}$ by [**17**] less $\mathcal{M}$. (*ii*) "**F-M**" stands for the approximations of $E\{[\log(n)]H_n\}$ by [**18**] less $\mathcal{M}$. (*iii*) "**S-M**" represents the approximations of $E\{[\log(n)]H_n\}$ by [**19**] less $\mathcal{M}$.

Some Applications

We present two applications which involve the use of the approximation of $E(H_n)$.

Image Denoising. Image denoising is very important in image processing. There are many methods for image denoising in the literature of image process. We are interested in the local filters such as the arithmetic and harmonic mean filters that have been used in image denoising. The harmonic mean filter is better at removing positive outliers and preserving edge features than the arithmetic mean filter. However, both of them fail when the image is contaminated by a uniform noise. Comparing the difference between the two means on different segments, we use the ratio of the harmonic mean and the arithmetic mean (defined in [**23**]) as a local filter and select the corresponding threshold of the ratio using the improved approximation [**16**] plus a saddle-point approximation. This application shows how the local filter can improve the performance of image denoising. The details are given below.

For demonstration, we consider a test image with dimension 250 × 250 (Fig. 2*A*) including disk, hand, human body, ring, sunflower, and triangle as shown in figure 2 of ref. 16. We contaminate the image with uniform noise, which is displayed in Fig. 2*B*. The usual harmonic mean filter method in image denoising is to replace the value of each pixel with the harmonic mean of values of the pixels in a surrounding region. We consider a square containing 9 pixels for each pixel such that this pixel is located at the center. Here the variable Z_i represents the value of the pixel taking values 0 (black), 1/255,. . .,255/255 (white) in this 256 grayscale image and the sample size is 9. For the edge of an image with dimension 250 × 250 such as the first or last row and column, where the pixels are not surrounded by a square, we copy them to the neighbor areas in the original image and the new image becomes 252 × 252. Note that this handling is only for convenience of filtering and the added pixels will not be analyzed. From Fig. 2 *C* and *D*, it can be seen that even though the harmonic mean filter outperforms the arithmetic mean filter, both arithmetic mean filter and the harmonic mean filter fail to denoise the noisy image given in Fig. 2*B*. However, we can first

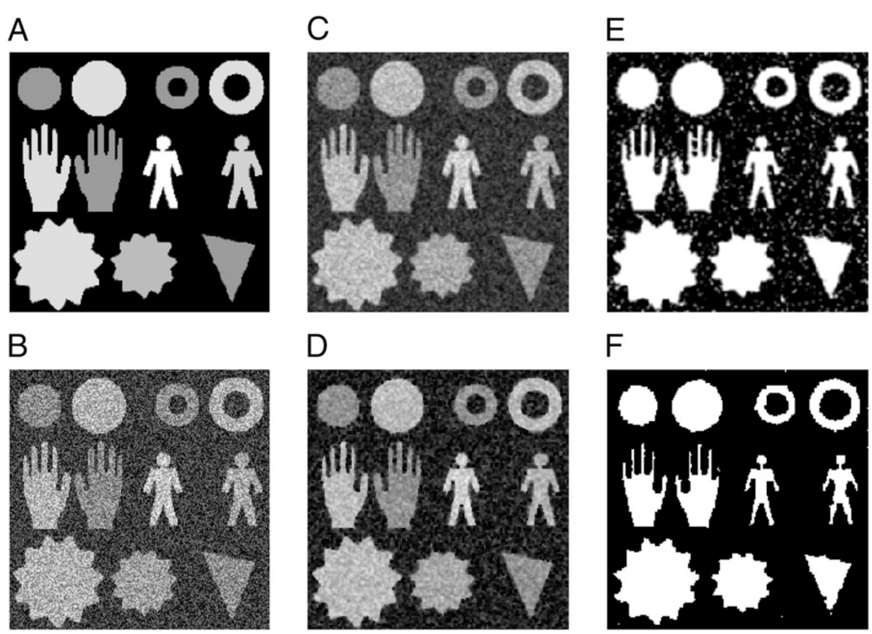

Fig. 2. (*A*) Initial unnoisy image. (*B*) Image that is noised by adding $U(0,1)$ noise to each pixel value of the image *A*. (*C*) Image obtained by denoising the noisy image *B* using the arithmetic mean filter. (*D*) Image obtained denoising the noisy image *B* using the harmonic mean filter. (*E*) The arithmetic mean filtered image of $\{\tilde{Z}_{i,j}\}$ (see [**23**]). (*F*) The harmonic mean filtered image of $\{\tilde{Z}_{i,j}\}$.

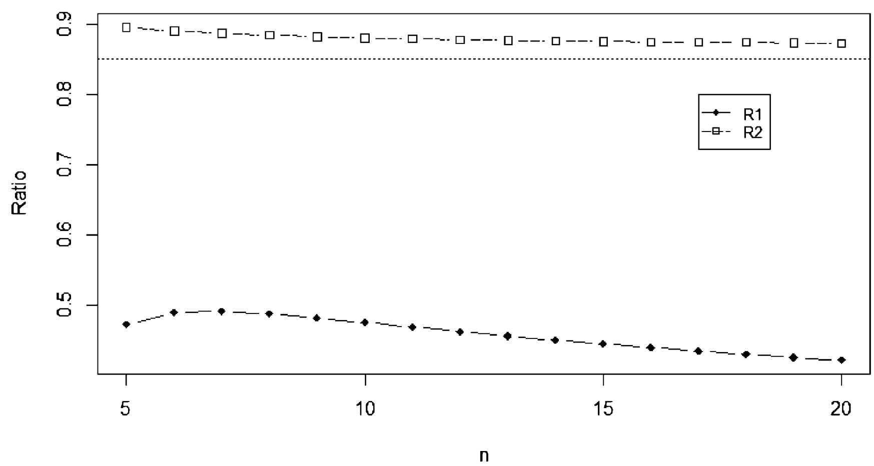

Fig. 3. Ratios $E(H_n)/E(\mathcal{A}_n)$ with $n=5,6,\ldots,20$ for both cases. "R1" denotes the ratio for case (*i*), whereas "R2" stands for the ratio for case (*ii*). The dotted line is 0.85.

use the ratio of the harmonic mean and the arithmetic mean jointly with a threshold θ to transform the pixel $Z_{i,j}$ at the pixel location (i,j) as follows:

$$\tilde{Z}_{i,j}=\begin{cases}1, & \text{if } H_{i,j}/\mathcal{A}_{i,j}\geq\theta,\\ 0, & \text{otherwise,}\end{cases} \qquad [23]$$

where $H_{i,j}$ and $\mathcal{A}_{i,j}$ are, respectively, the harmonic mean and the arithmetic mean of 9 pixels centering at $Z_{i,j}$. We then apply the arithmetic or harmonic mean filter to the pixels $\{\tilde{Z}_{i,j}\}$ to denoise the image of pixels $\{\tilde{Z}_{i,j}\}$. By Fig. 2 *E* and *F*, it can be seen that both images look much better than the images in Fig. 2 *C* and *D*. The image in Fig. 2*F* (by the harmonic mean filter) looks almost the same as the initial unnoisy image.

We note that only when using the ratio of the harmonic mean and the arithmetic mean, we assign 1 or 0 according to a threshold θ in [**23**], which is determined by the asymptotic behavior of the ratio of their expected values. How to select the threshold θ is important in practice. To demonstrate how to select θ, we consider two cases of uniform distributions with sample size n: (*i*) $Z_i\sim U(0,1)$; (*ii*) $Z_i\sim U(0.2,0.8)$. Let H_n and $\mathcal{A}_n$ be, respectively, the harmonic mean and the arithmetic mean of this sample. An approximation to $H_n/\mathcal{A}_n$ would be the ratio of their means, $E(H_n)/E(\mathcal{A}_n)$ as in ref. 9. For case (*i*), $E(H_n)$ can be approximated by [**16**], an improved approximation compared with the result of *Theorem 1*. For case (*ii*), $1/Z_{i,j}$ has moment of any order. Hence the saddle-point approximation [3.12] in ref. 10 can be applied, and $E(H_n)$ can thus be approximated by the three terms in that expansion. Fig. 3 displays the approximations of ratios of $E(H_n)/E(\mathcal{A}_n)$ with $n=5,6,\ldots,20$ for both cases. It can be seen that the approximation for case (*ii*) is larger than the one for case (*i*). By this figure, a practical recommendation of the threshold θ may be 0.85, which has been used for obtaining images displayed in Fig. 2 *E* and *F*.

Evaluating Marginal Likelihood. It is of importance to calculate the marginal likelihood in the process of likelihood maximization. Let $\pi(\theta|x)=f(x|\theta)\pi_0(\theta)/f_m(x)$ be the posterior density for prior $\pi_0(\theta)$, which implies that $[f_m(x)]^{-1}=E_\pi\{[f(x|\theta)]^{-1}\}$. Ref. 11 proposed the harmonic mean estimator for the marginal likelihood $f_m(x)$ by letting $Z_i=f(x|\theta_i)$ in [**1**], where θ_i s are i.i.d. draws from the posterior distribution. Ref. 11 noted that $1/Z_i$ can have infinite variance, in which case the central limit theorem is not applicable to the partial sums. Later, ref. 17 showed that in typical applications $[f(x|\theta_i)]^{-1}$ may lie in the domain of attraction

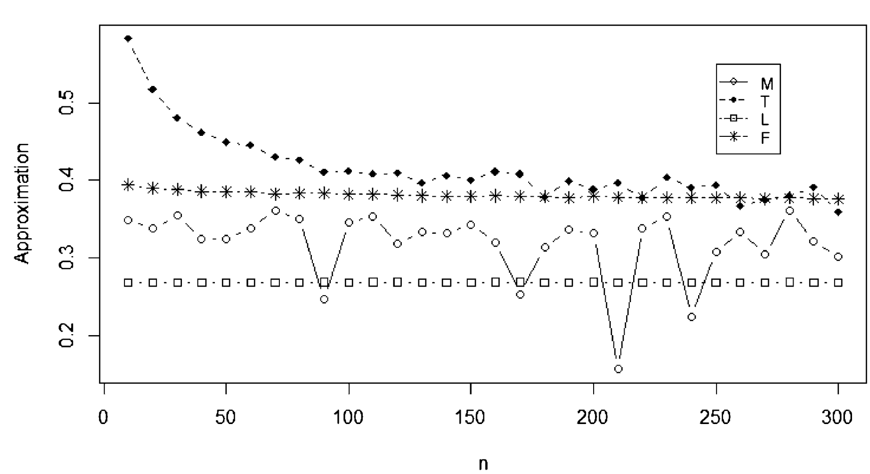

Fig. 4. Comparison of four approximations of the marginal likelihood with $n=10,20,30,\ldots,300$. (*i*) "**M**" denotes the sample mean of H_n in [**1**] with 100,000 replications of n independent observations from the posterior distribution. (*ii*) "**T**" stands for the sample mean of $\tilde{H}_n$ in [**24**] ($\delta=1.5$ is used) with 100,000 replications of n independent observations from the posterior distribution. (*iii*) "**L**" represents the sample mean of $f(\overline{X})$ in [**25**]. (*iv*) "**F**" denotes the sample mean of $f(\overline{X})+f^{3/2}(\overline{X})/n^{(1-1/\alpha)/2}$ in [**25**].

of a one-sided α-stable law with index $\alpha \in (1, 2]$. If the sample information exceeds the prior information in an application, the limit law for a harmonic mean estimator is stable with index α close to 1, and the convergence is very slow at rate $n^{1-\alpha^{-1}}$. In the following we demonstrate via one of their examples that if $\{1/Z_i\}$ are properly right truncated, a good approximation can be constructed so that it converges to the expected harmonic mean of the right truncated $\{1/Z_i\}$, which converges to the marginal likelihood.

Suppose we want to evaluate the marginal likelihood $f(\overline{X})$ based on $X_1, \ldots, X_r$ independently normally distributed $N(\theta, 1)$ variables with mean θ and variance 1 for a sample $\{X_i\}$ of size $r = 10$ with sample mean $\overline{X}$. Set the prior distribution $\theta \sim N(0, 1)$. The exact marginal likelihood for $r = 10$ is available analytically, $f(\overline{X}) = (2.2\pi)^{-1/2} e^{-\overline{X}^2/2.2}$. Our aim is to estimate the marginal likelihood $f(\overline{X})$, where $P(\overline{X} = 0) = 0$. The harmonic mean estimate of the marginal likelihood is $H_n = n/[\sum_{i=1}^n 1/Z_i]$, where $1/Z_i = \sqrt{\pi/5} e^{5(\theta_i - \overline{X})^2}$ for independent and identical draws θ_i from the posterior distribution $N(10\overline{X}/11, 1/11)$. Ref. 17 showed that the convergence rate of H_n to the marginal likelihood $f(\bar{x})$ is slow because of $\alpha = 1.1$, and the harmonic mean estimator behaves badly (Fig. 4). As described above, in light of the truncation method used in refs. 18 and 19, we consider the right truncated variable $1/Z_i I(1/Z_i < n^\delta)$, where $I(\cdot)$ is an indicator function and δ is a positive constant. Let

$$\tilde{H}_n = \frac{n}{\sum_{i=1}^n [1/Z_i I(1/Z_i < n^\delta)]}. \qquad [24]$$

By *Theorem 2.2*, it follows that

$$E(H_n) \sim E(\tilde{H}_n) \sim f(\overline{X}) + f^{3/2}(\overline{X})/n^{(1-1/\alpha)/2}. \qquad [25]$$

As displayed in Fig. 4, the convergence rate of H_n is very slow as described in ref. 17. The main reason is that the value of α is close to 1. From Fig. 4, it can be seen that $\tilde{H}_n$ given in [24] has a faster convergence rate to the two-term approximation in [25]. It is noted that this two-term approximation converges to the marginal likelihood $f(\overline{X})$. Thus, $\tilde{H}_n$ may be used as its approximation.

Similar results are obtained for different values of δ, although rate increases with less accuracy or decreases when δ is larger or smaller than 1.5, e.g., $\delta = 2$ or $\delta = 1$.

Appendix: Proof of Theorem 1

First we prove the case $\alpha = 1$, which implies that $A_n = \log(n)$, $B_n = n$ in [6], and the distribution of X_n converges to the stable distribution $F(x)$ with index $\alpha = 1$ satisfying [7] and [8] where $d_1 = 0$ and $d_2 = 1$. Denote $\ell_n = \sqrt{\log(n)}$.

$$\begin{aligned} E(H_n) &= \int_{-\infty}^{-\ell_n} \frac{1}{x + \log(n)} dF_n(x) + \int_{-\ell_n}^{\ell_n} \frac{1}{x + \log(n)} dF_n(x) \\ &\quad + \int_{\ell_n}^{\infty} \frac{1}{x + \log(n)} dF_n(x) \\ &= I_{1n} + I_{2n} + I_{3n}. \end{aligned}$$

Integrating by parts, we have

$$I_{1n} = \frac{F_n(-\ell_n)}{-\ell_n + \log(n)} + \int_{-\infty}^{-\ell_n} \frac{F_n(x)}{\{x + \log(n)\}^2} dx = I_{1n,1} + I_{1n,2}.$$

By [7] and [10], $I_{1n,1} = o(\ell_n^{-3})$. We now show that $I_{1n,2} = o(\ell_n^{-3})$. By [9]

$$\begin{aligned} I_{1n,2} &= \int_{d_3}^{\ell_n^2/2} \frac{F_n(x - \log(n))}{x^2} dx + \int_{-\ell_n^2/2}^{-\ell_n} \frac{F_n(x)}{\{x + \log(n)\}^2} dx \\ &= I_{1n,2,1} + I_{1n,2,2}. \end{aligned}$$

Because $I_{1n,2,1} \le F_n(-\log(n)/2) \to 0$, by applying l'Hôpital's rule,

$$\lim_{n\to\infty} \frac{I_{1n,2,1}}{1/\log^2(n)} = \lim_{n\to\infty} -\log(n) F_n(-\log(n)/2) = o(1),$$

then $I_{1n,2,1} = o(\ell_n^{-4})$. For the other part, $I_{1n,2,2} \le F_n(-\ell_n) O(\log^{-1}(n)) = o(\ell_n^{-3})$. So, $I_{1n} = o(\ell_n^{-3})$.

Using Taylor expansion, we have

$$\begin{aligned} I_{2n} &= \frac{1}{\log(n)} \int_{-\ell_n}^{\ell_n} dF_n(x) + \frac{1}{\log(n)} \int_{-\ell_n}^{\ell_n} \frac{-x}{\log(n)} dF_n(x) \\ &\quad + \frac{1}{\log(n)} \int_{-\ell_n}^{\ell_n} O\left(\frac{1}{\log(n)}\right) dF_n(x) \\ &= I_{2n,1} + I_{2n,2} + o(\ell_n^{-4}). \end{aligned}$$

By [8] and [10]

$$\begin{aligned} I_{2n,1} &= \frac{F_n(\ell_n) - F_n(-\ell_n)}{\log(n)} = \frac{1 - \{1 - F_n(\ell_n)\} - F_n(-\ell_n)}{\log(n)} \\ &= \frac{1 - \ell_n^{-1} + o(\ell_n^{-1})}{\log(n)}, \end{aligned}$$

$$\begin{aligned} I_{2n,2} &= \frac{-\ell_n F_n(\ell_n) - \ell_n F_n(-\ell_n)}{\log^2(n)} + \int_{-\ell_n}^{\ell_n} \frac{F(x)}{\log^2(n)} dx, \\ &= \frac{-\ell_n^{-1}}{\log(n)} + \int_{-\ell_n}^{\ell_n} \frac{F(x)}{\log^2(n)} dx + o(\ell_n^{-4}) \\ &= \frac{-\ell_n^{-1}}{\log(n)} + o(\ell_n^{-4}). \end{aligned}$$

By applying l'Hôpital's rule, we obtain

$$\lim_{n\to\infty} \frac{\int_{-\ell_n}^{\ell_n} \frac{F(x)}{\log(n)} dx}{\ell_n^{-1}} = \lim_{n\to\infty} \frac{\int_{-\ell_n^{-1}}^{\ell_n^{-1}} F(t\log(n)) dt}{\ell_n^{-1}} = 1,$$

which implies that $\int_{-\ell_n}^{\ell_n} [F(x)/\log(n)] dx = \ell_n^{-1} + o(\ell_n^{-1})$, and hence $I_{2n,2} = [1/\log(n)] - 2\ell_n^{-3} + o(\ell_n^{-3})$.

Because $\int_{\ell_n}^{\infty} dx/(x^2(x + \log(n))) = \ell_n^{-3} - \ell_n^{-4} \log(1 + \ell_n)$, we have $I_{3n} = \ell_n^{-3} + o(\ell_n^{-3})$ by [8] and [10]. In sum, we have

$$E(H_n) = \ell_n^{-2} - \ell_n^{-3} + o(\ell_n^{-3}).$$

ACKNOWLEDGMENTS. This work was partially supported by the Natural Sciences and Engineering Research Council of Canada.

1. Hamerly G, Elkan C (2002) Alternatives to k-means algorithm that find better clusterings. *Proceedings of the 11th International Conference on Information and Knowledge Management* (ACM, New York), pp 600–607.
2. Iman RL (1983) Harmonic mean. *Encyclopedia of Statistical Sciences*, eds Kotz S, Johnson NL (Wiley, New York), Vol 3, pp 575–576.
3. Komarova NL, Rivin I (2003) Harmonic mean, random polynomials and stochastic matrices. *Adv Appl Math* 31(2):501–526.
4. Zhang B, Hsu M, Dayal U (1999) K-harmonic means a data clustering algorithm. (Hewlett-Packard Labs), Technical Report HPL-1999-124. Available at www.hpl.hp.com/techreports/1999/HPL-1999-124.html. Accessed October 10, 2014.
5. Pinto JE, Henry E, Robinson TR, Stowe JD (2010) *Equity Asset Valuation* (John Wiley & Sons, Hoboken, NJ), 2nd Ed.
6. Lim MCH, McLernon DC, Ghogho M (2009) Weighted harmonic mean SINR maximization for the MIMO downlink. *IEEE International Conference on Acoustics, Speech, and Signal Processing* (Institute of Electrical and Electronics Engineers, Taipei, Taiwan) pp 2381–2384.
7. Jones CM (2003) Approximating negative and harmonic mean moments for the Poisson distribution. *Math Commun* 8(2):157–172.
8. Pakes AG (1999) On the convergence of moments of geometric and harmonic means. *Stat Neerl* 53(1):96–110.
9. Shi X, Reid N, Wu Y (2014) Approximation to the moments of ratios of cumulative sums. *Can J Stat* 42(2):325–336.
10. Shi X, Wang X-S, Reid N (2014) Saddlepoint approximation of nonlinear moments. *Stat Sin*.
11. Newton MA, Raftery AE (1994) Approximate Bayesian inference with the weighted likelihood bootstrap. *J R Stat Soc, B* 56(1):3–48.
12. Gnedenko BV, Kolmogorov AN (1954) *Limit Distributions for Sums of Independent Random Variables* (Addison-Wesley, Cambridge, MA).
13. Ibragimov IA, Linnik YV (1971) *Independent and Stationary Sequences of Random Variables* (Wolters-Noordhoff, Groningen, The Netherlands).
14. Hall P (1981) On the rate of convergence to a stable law. *J Lond Math Soc* 23:179–192.
15. Berkes I, Tichy R (2014) Lacunary series and stable distributions. Available at www.math.tugraz.at/discrete/publications/projects/files/berkes_tichy_lac.pdf. Accessed March 10, 2014.
16. Aue A, Lee TCM (2011) On image segmentation using information theoretic criteria. *Ann Stat* 39(6):2912–2935.
17. Wolpert RL, Schmidler SC (2012) α-Stable limit laws for harmonic mean estimators of marginal likelihoods. *Stat Sin* 22(3):1233–1251.
18. Shi X, Wu Y, Liu Y (2010) A note on asymptotic approximations of inverse moments of nonnegative random variables. *Stat Probab Lett* 80(15-16):1260–1264.
19. Wu T-J, Shi X, Miao B (2009) Asymptotic approximation of inverse moments of nonnegative random variables. *Stat Probab Lett* 79(11):1366–1371.