On stationary multiplier methods for the rounding of probabilities and the limiting law of the Sainte-Laguë divergence

Lothar Heinrich, Friedrich Pukelsheim, Udo Schwingenschlögl

Summary: Stationary multiplier methods are procedures for rounding real probabilities into rational proportions, while the Sainte-Laguë divergence is a reasonable measure for the cumulative error resulting from this rounding step. Assuming the given probabilities to be uniformly distributed, we show that the Sainte-Laguë divergences converge to the Lévy-stable distribution that obtains for the multiplier method with standard rounding. The norming constants to achieve convergence depend in a subtle way on the stationary method used.

1 Introduction

For a fixed number of *categories* c, let the vector $(W_1, \ldots, W_c)$ be uniformly distributed on the probability simplex $\mathcal{S}_c = \{(w_1, \ldots, w_c) \in [0, 1]^c : \sum_{j \leq c} w_j = 1\}$. We consider the problem of rounding the real *weights* W_j to rational proportions N_j/n, for some prespecified *accuracy* n, with nonnegative integer numerators $N_1, \ldots, N_c$ summing to n. There are various ways to obtain the numerators N_j. Heinrich, Pukelsheim and Schwingenschlögl (2004; henceforth quoted as HPS) treat the *multiplier method with standard rounding*. This method generates the numerators N_j by rounding the scaled weights μW_j in a standard fashion, where the multiplier $\mu > 0$ is adjusted so as to achieve $\sum_{j \leq c} N_j = n$.

In the present note we discuss the wider class of *stationary multiplier methods* depending on a *stationarity parameter* $q \in [0, 1]$, as reviewed in Section 2. In order to measure the error that comes with discretizing the continuous weights W_j into rational proportions $N_{q,j}/n$, Sainte-Laguë (1910) proposed a chi-square-type goodness-of-fit criterion,

$$S_{q,c,n} = \sum_{j \leq c} \frac{(N_{q,j} - nW_j)^2}{W_j},$$

which we call *the Sainte-Laguë divergence*. The subscript q indicates that the numerators $N_{q,j}$ are generated by the q-stationary multiplier method. The case $q = 1/2$ retrieves the multiplier method with standard rounding, considered in HPS.

We generalize the limiting results of HPS to all q-stationary multiplier methods. In Section 3 we let the accuracy n tend to infinity, and show that the Sainte-Laguë divergences $S_{q,c,n}$ converge in distribution to a limiting variable $S_{q,c}$. In Section 4 the number of categories c grows large, and we exhibit constants a_q and b_q to again achieve convergence in distribution,

$$\frac{a_q}{c^2} S_{q,c} - b_q - \log c \quad \xrightarrow[c \to \infty]{\mathrm{d}} \quad S^*,$$

where the random variable S^* has the Lévy-stable distribution (not depending on q) that appeared in HPS for the multiplier method with standard rounding.

2   Stationary multiplier methods

Let the stationarity parameter $q \in [0, 1]$ be fixed. The essential ingredient is the *q-stationary rounding function* $[x]_q$, defined through $[0]_q = 0$ and, for $x > 0$, through

$$[x]_q = \begin{cases} \text{IntegerPart}(x) + 1 & \text{for FractionalPart}(x) > q, \\ \text{IntegerPart}(x) & \text{for FractionalPart}(x) < q. \end{cases}$$

That is, a value x is rounded down when the fractional part of x is smaller than q, while it is rounded up when the fractional part is larger than q. In view of our distributional assumptions the events when $x = k + q$, where k is an integer, form Lebesgue nullsets, whence in these points we may leave the rounding function $[x]_q$ undefined.

In fact, there is no general agreement of how to extend the definition of $[x]_q$ at the jump points $k + q$, except for making sure that the ensuing function is increasing. When $q = 0$, the rounding function $[x]_q$ rounds up, which is suggestive for extending it to the (left continuous) *ceiling function* $\lceil x \rceil$, the smallest integer larger than or equal to x. When $q = 1$, the function $[x]_q$ rounds down, whence the natural extension is the (right continuous) *floor function* $\lfloor x \rfloor$, the largest integer smaller than or equal to x. When $q = 1/2$, the function $[x]_q$ yields standard rounding $\langle x \rangle$ which, at the points of jump, HPS (p. 46) define to round down. However, business people prefer to round $k + 0.5$ up to $k + 1$, while statisticians are advised to round to the nearest *even* integer, see Wallis and Roberts (1956, p. 175).

The *q-stationary multiplier method* approximates the real weights W_j by rational proportions of the form $N_{q,j}/n$ with numerators $N_{q,j} = [\mu W_j]_q$, where $[x]_q$ is the q-stationary rounding function and n is the prespecified accuracy, while the multiplier $\mu > 0$ is adjusted so as to achieve $\sum_{j \leq c} N_{q,j} = n$. Stationary multiplier methods lend themselves to rather complete mathematical results, see Oyama (1991, p. 201), Balinski and Rachev (1993, p. 479), Oyama and Ichimori (1995, p. 305), Balinski and Rachev (1997, p. 13), Ramírez, Márquez and Pérez (1999, p. 475), Marshall, Olkin and Pukelsheim (2002, p. 892), Palomares and Ramírez (2003, p. 406), Schuster, Pukelsheim, Drton and Draper (2003, p. 669). Moreover, they permit a continous transition between the

three apportionment methods that are most widely known: the method based on rounding up ($q = 0$), the method using standard rounding ($q = 1/2$), and the method obtained from rounding down ($q = 1$).

No closed formula is known how the input weights $W_1, \ldots, W_c$ determine a "local", stochastic multiplier $\mu = \mu(W_1, \ldots, W_c)$ satisfying $\sum_{j \leq c} [\mu W_j]_q = n$. Instead, we start out with a "global", deterministic multiplier ν close to n and, in a second step, take some corrective action to arrive at the desired numerators $N_{q,1}, \ldots, N_{q,c}$. Lemma 2.1 casts the Adjustment Algorithm of HPS (p. 46) into a more generic format. We define $\mathrm{sgn}(x) = -1, 0, 1$ according to whether x is negative, zero, or positive.

Lemma 2.1 (Generic Algorithm) *Let the stationarity parameter $q \in [0, 1]$, the accuracy $n \in \{1, 2, \ldots\}$, and a global multiplier $\nu \geq 0$ be fixed. For Lebesgue almost all weight vectors $(W_1, \ldots, W_c)$ the q-stationary multiplier method yields a unique apportionment vector $(N_{q,1}, \ldots, N_{q,c})$ that is obtained as follows. Firstly, we calculate the discrepancy*

$$D = \left(\sum_{j \leq c} [\nu W_j]_q \right) - n,$$

a random variable with integer values in the interval $\big(\nu - n - cq, \nu - n + c(1 - q)\big)$. Secondly, for $j = 1, \ldots, c$, we adjust the initial assignment $[\nu W_j]_q$ to obtain the final numerators

$$N_{q,j} = [\nu W_j]_q - \mathrm{sgn}(D) m_{j,n}(D),$$

where $m_{j,n}(D)$ is the count of how often index j appears among the $|D|$-smallest quotients

$$
\begin{cases}
\dfrac{k - \nu W_i + [\nu W_i]_q + q - 1}{W_i} & \text{for } i = 1, \ldots, c \text{ and } k = 1, \ldots, -D, \quad \text{when } D < 0; \\[2ex]
\dfrac{k + \nu W_i - [\nu W_i]_q - q}{W_i} & \text{for } i = 1, \ldots, c \text{ and } k = 1, \ldots, D, \qquad \text{when } D > 0.
\end{cases}
$$

Proof: The terms $\mp \nu W_i / W_i = \mp \nu$ are included in anticipation of Theorem 3.1, the constant shift $\mp \nu$ does not affect the ordering of the quotients from smallest to largest. The assertion thus follows as in HPS (p. 47). □

The Generic Algorithm comprises a variety of algorithms, by starting out from different global multipliers ν. For instance, with $\nu = 0$ all initial assignments vanish, $[\nu W_j]_q = 0$, and the discrepancy is as negative as can be, $D = -n$. The adjustment terms $m_{j,n}(D)$ augment the initial assignment 0 until the final apportionment $N_{q,j}$ is reached; they are found by counting the appearance of the index j among the n-smallest quotients $(k - 1 + q)/W_i$, or equivalently, among the n-largest quotients $W_i/(k - 1 + q)$. This yields the recursive algorithm that is known best, see Balinski and Young (1982, p. 100).

In fact, as long as the initial multiplier ν stays below $n - c(1 - q)$ the discrepancy is negative, $D < 0$, and all adjustment terms increment the initial assignments $[\nu W_j]_q$. On the other hand, when the initial multiplier ν lies above $n + cq$ the discrepancy is positive, $D > 0$, and all adjustment terms decrement the initial assignment $[\nu W_j]_q$. In the field of political science, decrementation comes with the bad flavor that participants must give back what they have been assigned already. Mathematically, this objection is unsubstantiated because the important outcome is the final apportionment $N_{q,j}$, not a transient state during the algorithmic calculations.

The most convenient choice for the global multiplier is such that the support interval of the discrepancy is centered around zero, which is achieved by

$$\nu_n = n + c\left(q - \frac{1}{2}\right).$$

The multiplier ν_n minimizes the computational complexity of the Generic Algorithm, see Happacher and Pukelsheim (2000, p. 154). The resulting algorithm is considered in Happacher and Pukelsheim (1996, p. 378), and Dorfleitner and Klein (1999, p. 147). For the 1-stationary multiplier method, that is the multiplier method with rounding down, the multiplier $\nu_n = n + c/2$ was proposed already by Gfeller (1890, p. 130). Hagenbach-Bischoff (1905, p. 15) advocated $\nu = n + 1$ which, with the discrepancy D necessarily coming out negative, makes do with incrementation only. The distribution of the discrepancy is investigated by Happacher (1996, 2001).

The minimum complexity multiplier ν_n is also instrumental for carrying out the asymptotical analysis, as is the accompanying discrepancy

$$D_{q,c,n} = \left(\sum_{j \leq c} [\nu_n W_j]_q\right) - n,$$

attaining the integer values $-\left\lfloor \frac{c-1}{2} \right\rfloor, \ldots, \left\lfloor \frac{c-1}{2} \right\rfloor$.

This discrepancy version permits various reformulations, emphasizing its distinct aspects. While the discrepancy identities are implicit in HPS, it is worth the space to single them out more explicitly. Insertion of $n = \nu_n - c(q - 1/2)$ and $\nu_n = \sum_{j \leq c} \nu_n W_j$ yields

$$D_{q,c,n} = c\left(q - \frac{1}{2}\right) - \sum_{j \leq c}\left(\nu_n W_j - [\nu_n W_j]_q\right),$$

exhibiting the discrepancy as a function of the *q-stationary residuals* $\nu_n W_j - [\nu_n W_j]_q$.

For a study of the rounding residuals a passage to standard rounding transpires to be helpful, via $[x]_q = \langle x - q + 1/2 \rangle$, which follows instantly from the equivalences

$$[x]_q = k \iff x \in (k + q - 1, k + q)$$

$$\iff x - q + \frac{1}{2} \in \left(k - \frac{1}{2}, k + \frac{1}{2}\right) \iff \left\langle x - q + \frac{1}{2} \right\rangle = k.$$

Upon introducing the *centering residual function* $u_q(x) = x - q + 1/2 - \langle x - q + 1/2 \rangle \in [-1/2, 1/2]$, the q-stationary residuals become $x - [x]_q = u_q(x) + q - 1/2$. This

establishes the first of the two identities

$$D_{q,c,n} = -\sum_{j \leq c} u_q(\nu_n W_j) = -\left\langle \sum_{j<c} u_q(\nu_n W_j) \right\rangle.$$

The second identity follows from $\sum_{j<c} u_q(\nu_n W_j) = -D_{q,c,n} - u_q(\nu_n W_c) \in (-D_{q,c,n} - 1/2, -D_{q,c,n} + 1/2)$. We are now prepared to turn to asymptotic considerations.

3 Infinite rounding accuracy

In this section we assume that the truncated vector $(W_1, \ldots, W_{c-1})$ has a Riemann integrable Lebesgue density on its domain $\mathcal{T}_c = \{(w_1, \ldots, w_{c-1}) \in [0, 1]^{c-1} : \sum_{j<c} w_j < 1\}$, as in HPS (p. 50). We consider the q-stationary multiplier method, for some fixed $q \in [0, 1]$. As the accuracy n tends to infinity, so does the multiplier $\nu_n = n + c(q - 1/2)$.

Theorem 3.1 *(i) For $j < c$, the q-stationary residuals $\nu_n W_j - [\nu_n W_j]_q$ converge in distribution to $U_j + q - 1/2$, where the random variables $U_1, \ldots, U_{c-1}$ are uniformly distributed on $(-1/2, 1/2)$, and are independent of each other as well as of $W_1, \ldots, W_{c-1}$.*

(ii) Moreover, upon setting $D_c = -\left\langle \sum_{j<c} U_j \right\rangle$ and $U_c = -D_c - \sum_{j<c} U_j$, the Sainte-Laguë divergences $S_{q,c,n}$ converge in distribution to

$$S_{q,c} = -c^2 \left(q - \frac{1}{2} \right)^2 + \sum_{j \leq c} \frac{\left(U_j + q - \frac{1}{2} + \text{sgn}(D_c) m_j(D_c) \right)^2}{W_j},$$

where $m_j(D_c)$ is the count of how often index j appears among the $|D_c|$-smallest quotients

$$\frac{k + \text{sgn}(D_c) U_i - \frac{1}{2}}{W_i} \qquad \text{for } i = 1, \ldots, c \text{ and } k = 1, \ldots, |D_c|.$$

Proof:

(i) Fix $j < c$. We start with the representation $\nu_n W_j - [\nu_n W_j]_q = u_q(\nu_n W_j) + q - 1/2$ introduced above. With some constants $\alpha, \beta \in \mathbb{R}$, the random variables $u_q(\nu_n W_j)$ are of the form $(n + \alpha) W_j + \beta - \langle (n + \alpha) W_j + \beta \rangle$ and hence converge to U_j. In case $\alpha = \beta = 0$ this is shown in HPS (p. 48), but their Fourier transform argument is easily seen to extend to arbitrary values $\alpha, \beta \in \mathbb{R}$.

(ii) It follows from the Continuous Mapping Theorem that the discrepancies $D_{q,c,n} = -\left\langle \sum_{j<c} u_q(\nu_n W_j) \right\rangle$ converge to $D_c = -\left\langle \sum_{j<c} U_j \right\rangle$. The terminal residuals $u_q(\nu_n W_c) = -D_{q,c,n} - \sum_{j<c} u_q(\nu_n W_j)$ then converge to $U_c = -D_c - \sum_{j<c} U_j$.

As for the Sainte-Laguë divergences, we insert $nW_j = \nu_n W_j - c(q - 1/2)W_j$ to obtain

$$S_{q,c,n} = \sum_{j \leq c} \frac{\left(c\left(q - \frac{1}{2}\right)W_j + N_{q,j} - \nu_n W_j\right)^2}{W_j}$$

$$= -c^2 \left(q - \frac{1}{2}\right)^2 + \sum_{j \leq c} \frac{\left(\nu_n W_j - N_{q,j}\right)^2}{W_j},$$

since $\sum_{j \leq c} c^2 (q - 1/2)^2 W_j = c^2 (q - 1/2)^2$, and since $2c(q-1/2)\sum_{j \leq c}(N_{q,j} - \nu_n W_j) = -2c^2(q-1/2)^2$. In the sum of squares we write $\nu_n W_j - N_{q,j} = \nu_n W_j - [\nu_n W_j]_q + \text{sgn}(D_{q,c,n})m_{j,n}(D_{q,c,n})$. These terms converge to $U_j + q - 1/2 + \text{sgn}(D_c)m_j(D_c)$, provided the counts $m_{j,n}(D_{q,c,n})$ behave appropriately. Indeed, in the augmentation case $D_c < 0$, we have $(k - \nu_n W_i + [\nu_n W_i]_q + q - 1)/W_i$ converging to $(k - U_i - 1/2)/W_i$. In the reduction case $(k + \nu_n W_i - [\nu_n W_i]_q - q)/W_i$ converge to $(k + U_i - 1/2)/W_i$. Either way, the limiting quotients to be sorted from smallest to largest are $\left(k + \text{sgn}(D_c)U_i - 1/2\right)/W_i$ for $i = 1, \dots, c$ and $k = 1, \dots, |D_c|$. The proof is complete. $\square$

As remarked by a referee, convergence in distribution in Theorem 3.1.i may be sharpened to convergence in total variation. To this end let f denote the probability density of $(W_1, ..., W_{c-1})$, which vanishes outside the domain $\mathcal{T}_c$. Then the vector of residuals $(\nu_n W_1 - [\nu_n W_1]_q], \dots, \nu_n W_{c-1} - [\nu_n W_{c-1}]_q)$ has the density function

$$p_{q,c,n}(x_1, ..., x_{c-1}) = \frac{1}{\nu_n^{c-1}} \sum_{k_1=0}^{\lfloor \nu_n \rfloor + 1} \cdots \sum_{k_{c-1}=0}^{\lfloor \nu_n \rfloor + 1} f\left(\frac{k_1 + x_1}{\nu_n}, \dots, \frac{k_{c-1} + x_{c-1}}{\nu_n}\right).$$

With $(x_1, ..., x_{c-1}) \in (q - 1, q)^{c-1}$ fixed, this is a Riemann sum of the density f over $[0, 1]^{c-1}$, whence Riemann integrability of f entails

$$\lim_{n \to \infty} p_{q,c,n}(x_1, ..., x_{c-1}) = \int_{\mathcal{T}_c} f(w_1, ..., w_{c-1})\, d(w_1, ..., w_{c-1}) = 1.$$

But then the densities $p_{q,c,n}$ are uniformly bounded, since f is Riemann integrable and hence bounded. Finally the Dominated Convergence Theorem is seen to yield convergence in variation,

$$\lim_{n \to \infty} \int_{(q-1,q)^{c-1}} \left| p_{q,c,n}(x_1, ..., x_{c-1}) - 1 \right| d(x_1, ..., x_{c-1}) = 0.$$

These arguments are closely related to the concept of asymptotically uniformly distributed sequences of probability measures on locally compact Abelian groups, as developed in Kerstan and Matthes (1968). By exploiting the depth of this more general approach the smoothness assumptions underlying Theorem 3.1 may be weakened.

Sainte-Laguë (1910) showed that, for a fixed accuracy n, the divergence $S_{q,c,n}$ is pointwise minimized at $q = 1/2$. Hence the limiting variables $S_{q,c}$, for $q \neq 1/2$, are

stochastically larger than $S_{1/2,c}$. Furthermore the limiting Sainte-Laguë divergences $S_{q,c}$ and $S_{1-q,c}$ have the same distribution, that is, the parametrization by the stationarity parameter q is symmetric about the point $1/2$.

Theorem 3.1 entails some general formulas for biases, that is, for the expected absolute error, or for the expected relative error. Specifically, the *seat bias of the k-th largest party* is defined to be

$$B_k^{(q)}(n, t) = \text{E}\left(N_{q,k} - nW_k \,\middle|\, W_1 \geq \cdots \geq W_c \geq t\right).$$

In other words, the seat bias is the expected difference between the actually allocated numerators $N_{q,k}$ and the corresponding ideal shares nW_k, under the condition that the weights are ordered from largest to smallest and that the smallest weight lies above a prespecified threshold $t \in [0, 1/c]$. Note that all seat biases together always sum to zero, $\sum_{k \leq c} B_k^{(q)}(n, t) = 0$.

The terminology alludes to the situation where in a political body n designates the total number of seats to be apportioned, $N_{q,k}$ is the number of seats that the q-stationary multiplier method allocates to the k-th largest party, and t is the threshold that parties must pass in order to be eligible to participate in the apportionment process.

Corollary 3.2 *Let $k = 1, \ldots, c$ be fixed, and let $t \in [0, 1/c]$ be a prespecified threshold.*

(i) With the assumptions and notions of Theorem 3.1 the asymptotic seat bias of the k-th largest party is

$$\lim_{n \to \infty} B_k^{(q)}(n, t) = \left(q - \frac{1}{2}\right)\left\{c\,\text{E}\left(W_k \,\middle|\, W_1 \geq \cdots \geq W_c \geq t\right) - 1\right\}.$$

(ii) In particular, if the weights $W_1, \ldots, W_c$ are uniformly distributed on the probability simplex $\mathcal{S}_c$ then

$$\lim_{n \to \infty} B_k^{(q)}(n, t) = \left(q - \frac{1}{2}\right)\left\{\left(\sum_{j=k}^{c} \frac{1}{j}\right) - 1\right\}(1 - ct).$$

Proof:

(i) Denoting the conditional expectation by $\text{E}^{(t)}(\cdot) = \text{E}\left(\cdot \,\middle|\, W_1 \geq \cdots \geq W_c \geq t\right)$, we need to determine the limit of $B_k^{(q)}(n, t) = \text{E}^{(t)}\left(N_{q,k} - nW_k\right)$. Insertion of $N_{q,k} = [\nu_n W_k]_q - \text{sgn}\left(D_{q,c,n}\right)m_{k,n}\left(D_{q,c,n}\right)$ and $nW_k = \nu_n W_k - c\left(q - 1/2\right)W_k$, and substitution of $[\nu_n W_k]_q - \nu_n W_k = -u_q(\nu_n W_k) - (q - 1/2)$ yield

$$\text{E}^{(t)}\left(N_{q,k} - nW_k\right) = \left(q - \frac{1}{2}\right)\left\{c\,\text{E}^{(t)}(W_k) - 1\right\}$$
$$- \text{E}^{(t)}\left(u_q(\nu_n W_k) + \text{sgn}(D_{q,c,n})m_{k,n}(D_{q,c,n})\right).$$

The integrands in the last expectation are bounded. Hence convergence in distribution implies that they converge to $-\text{E}^{(t)}\left(U_k + \text{sgn}(D_c)m_k(D_c)\right)$. Since the limiting

variables U_k and D_c are independent of the conditioning variables $W_1, \dots, W_c$, the expectation is actually unconditional. Clearly we have $E(U_k) = 0$. As $D_c = \sum_{j<c} U_j$ is distributed symmetrically around zero, so is $\mathrm{sgn}(D_c)m_k(D_c)$, which yields $E\big(\mathrm{sgn}(D_c)m_k(D_c)\big) = 0$. In summary we obtain $\lim_{n\to\infty} B_k^{(q)}(n,t) = (q - 1/2)\big\{c\,E^{(t)}(W_k) - 1\big\}$.

(ii) On the conditioning event $\{W_1 \geq \cdots \geq W_c \geq t\}$, the transformed variables $W_j^* = (W_j - t)/(1 - ct)$, for $j \leq c$, are nonnegative and sum to unity, and hence inherit the uniform distribution from $W_1, \dots, W_c$. Thus we get $E^{(t)}(W_k) = E\big(W_k^* | W_1^* \geq \cdots \geq W_c^*\big)(1 - ct) + t = E^{(0)}(W_k)(1 - ct) + t$. From Drton and Schwingenschlögl (2004, p. 90) or Johnson, Kotz and Balakrishnan (1994, p. 500) we know that $E^{(0)}(W_k) = c^{-1}\sum_{j=k}^{c} j^{-1}$, leading to $c\,E^{(t)}(W_k) - 1 = \{(\sum_{j=k}^{c} j^{-1}) - 1\}(1 - ct)$. The proof is complete.   □

An alternate proof of Corollary 3.2.ii is given by Schwingenschlögl and Pukelsheim (2005), based on the geometric combinatorial results of Schwingenschlögl and Drton (2004). Comparisons of empirical data with the abstract formulas are rather encouraging in that, practically, the asymptotic formulas are perfectly acceptable for finite accuracy n provided the accuracy is at least twice as large as the number of categories, $n \geq 2c$, see Schuster, Pukelsheim, Drton and Draper (2003, p. 668) who first conjectured Corollary 3.2.ii when $t = 0$, which then was rigorously proved by Drton and Schwingenschlögl (2005, Theorem 1). Corollary 3.2.i indicates what happens under general distributional assumptions.

There is another notion of bias that is more relevant from the viewpoint of constitutional law. This is the average *relative* deviation between realized and ideal allocations. We call

$$A_k^{(q)}(n,t) = E\left(\frac{N_{q,k} - nW_k}{W_k} \,\bigg|\, W_1 \geq \cdots \geq W_c \geq t\right)$$

the *success-value bias of the voters of the k-th largest party*. In fact, the decisions of the German Federal Constitutional Court center around the *success value of a voter's ballot* (Erfolgswert einer Wählerstimme), which numerically corresponds to the likelihood ratio $\big(N_{q,k}/n\big)/W_k$, see Pukelsheim (2000, p. 450). The *ideal success value* equals unity. Hence $n^{-1}A_k^{(q)}(n,t)$ measures the expected difference between realized and ideal success values.

Corollary 3.3 *Let $k = 1, \dots, c$ be fixed, and let the threshold $t \in (0, 1/c]$ be nonzero. With the assumptions and notions of Theorem 3.1 the asymptotic success-value bias of the voters of the k-th largest party is*

$$\lim_{n\to\infty} A_k^{(q)}(n,t) = \left(q - \frac{1}{2}\right)\left\{c - E\left(\frac{1}{W_k} \,\bigg|\, W_1 \geq \cdots \geq W_c \geq t\right)\right\}.$$

Proof: With the threshold assumed positive all inverse weights remain bounded, $1/W_k \leq 1/t < \infty$. Hence the proof of Corollary 3.2.i carries over.   □

A neat formula paralleling Corollary 3.2.ii would be desirable, but evaded our assaults. We next approach the question what happens for a large number of categories, $c \to \infty$. For approximating the limiting Sainte-Laguë divergences $S_{q,c}$ through terms of which the distribution is handled more easily we provide the following bounds.

Corollary 3.4 *With the assumptions and notions of Theorem 3.1 we have*

$$0 \leq S_{q,c} + c^2 \left(q - \frac{1}{2}\right)^2 - \sum_{j \leq c} \left(\frac{\left(U_j + q - \frac{1}{2}\right)^2}{W_j} + 2\left(q - \frac{1}{2}\right) \frac{\text{sgn}(D_c) m_j(D_c)}{W_j} \right) \leq c|D_c|.$$

Proof: Expanding the squares in the sum of Theorem 3.1.ii, the center expression is seen to be the sum, over $j \leq c$, of the terms $\left(2\text{sgn}(D_c)U_j + m_j(D_c)\right)m_j(D_c)/W_j$, which HPS (p. 52) show to lie between 0 and $c\, m_j(D_c)$. Now $c \sum_{j \leq c} m_j(D_c) = c|D_c|$ completes the proof. $\qquad\square$

The law of large numbers makes $c^{-1}|D_c|$ converge to zero as c tends to infinity. Hence $c^{-2}S_{q,c}$ is approximated by the other terms that are appearing in Corollary 3.4. For determining their limiting distribution the rather generous distributional assumptions of Theorem 3.1 are narrowed down.

4 Infinitely many categories

The behavior of the limiting Sainte-Laguë divergences $S_{q,c}$ for a growing number of categories c is studied under the assumption that, for c fixed, the weight vector $(W_1, \ldots, W_c)$ is uniformly distributed on the probability simplex $\mathcal{S}_c$. Then the weights may be taken to be generated from exponentially distributed random variables E_j, via $W_j = E_j/\sum_{i \leq c} E_i$, and the terminal variable U_c – which depends on $U_1, \ldots, U_{c-1}$ – may be replaced by an independent copy, see HPS (p. 53).

Lemma 4.1 *Let the random variables $V_1, E_1, V_2, E_2, \ldots$ be independent such that every V_j is uniformly distributed on $(-1/2, 1/2)$ and every E_j is exponentially distributed with mean one. Then the vector $(W_1, \ldots, W_c)$ with components defined by $W_j = E_j/\sum_{i \leq c} E_i$ $(j = 1, \ldots, c)$ is uniformly distributed on $\mathcal{S}_c$, for every c, and*

$$\frac{1}{c^2} S_{q,c} + \left(q - \frac{1}{2}\right)^2 - \frac{1}{c} \sum_{j \leq c} \frac{\left(V_j + q - \frac{1}{2}\right)^2}{E_j}$$

converge in probability to zero, as c tends to infinity.

Proof: The crucial point is that the sums $c^{-2} \sum_{j \leq c} \text{sgn}(D_c) m_j(D_c)/W_j$ from Corollary 3.4 converge in probability to zero, which is proved in Heinrich and Schwingenschlögl (2006). The remaining steps are as in HPS (p. 53). $\qquad\square$

Upon setting $p = \min\{q, 1-q\}$, the terms $(V_j + q - 1/2)^2$ turn out to have density $g(x) = \frac{1}{\sqrt{x}}\mathbf{1}_{(0,p^2]}(x) + \frac{1}{2\sqrt{x}}\mathbf{1}_{(p^2,(1-p)^2]}(x)$. The common density f_q of the random variables $Z_j = (V_j + q - 1/2)^2/E_j$ is now found to be

$$f_q(z) = \int_0^{(1-p)^2/z} g(zx)e^{-x}|x|\,dx$$

$$= \frac{1}{2\sqrt{z}}\int_0^{p^2/z} e^{-x}\sqrt{x}\,dx + \frac{1}{2\sqrt{z}}\int_0^{(1-p)^2/z} e^{-x}\sqrt{x}\,dx$$

$$= \frac{q^3 + (1-q)^3}{3z^2} - \frac{q^5 + (1-q)^5}{5z^3} + \frac{q^7 + (1-q)^7}{14z^4} \mp \cdots,$$

for $z > 0$. We have $f_q(z) = f_{1-q}(z)$, for $q \in [0, 1]$.

Let S^* be a random variable with characteristic function

$$\mathrm{E}e^{itS^*} = \exp\left(-\frac{\pi}{2}|t| - it\log|t|\right) \quad \text{for} \quad t \in \mathbb{R},$$

that is, S^* has a Lévy-stable distribution. As compared to the random variable S of HPS (p. 53) we have $S^* = (\pi/2)S + \log(\pi/2)$; we find S^* slightly more convenient to work with, for our present purposes. Furthermore let $\gamma = \lim_{k\to\infty}(1 + 1/2 + \cdots + 1/k - \log k) \simeq 0.5772$ denote the Euler–Mascheroni constant. Now we are in a position to formulate the asymptotic behavior of the limiting Sainte-Laguë divergences $S_{q,c}$ when they are suitably scaled and shifted.

Theorem 4.2 *Suppose that the weight vector* $(W_1, \ldots, W_c)$ *is uniformly distributed on the probability simplex* $\mathcal{S}_c$, *and let* $q \in [0, 1]$ *be fixed. Then we have*

$$\frac{a_q}{c^2} S_{q,c} - b_q - \log c \xrightarrow[c\to\infty]{\mathrm{d}} S^*,$$

with constants a_q *and* b_q *given by*

$$a_q = \frac{3}{q^3 + (1-q)^3},$$

$$b_q = \frac{2}{3} - 2\gamma + \log\frac{q^3 + (1-q)^3}{3q(1-q)} + \frac{1 + 4(q^3 - (1-q)^3)\log\frac{1-q}{q}}{4(q^3 + (1-q)^3)}.$$

The constants a_q *and* b_q *are symmetric about* $1/2$, *and we have* $b_0 = b_1 = 11/12 - 2\gamma - \log 3$.

Proof: In view of Lemma 4.1 it suffices to show that the random variables

$$\widetilde{S}_{q,c} = \frac{a_q}{c}\left(\sum_{j\le c} Z_j\right) - \log c - b_q - a_q\left(q - \frac{1}{2}\right)^2$$

converge in distribution to S^* as $c \to \infty$. To this end we show that for any fixed $t \in \mathbb{R}$ the characteristic functions

$$\mathrm{E}e^{it\widetilde{S}_{q,c}} = \left(\mathrm{E} \exp\left\{ \frac{ia_q t}{c} Z_1 - \frac{it}{c} \left(\log c + b_q + a_q \big(q - \tfrac{1}{2}\big)^2 \right) \right\} \right)^c$$

converge to the characteristic function of S^*. This follows from the approximation

$$\mathrm{E}e^{isZ_1} = 1 - \frac{\pi|s|}{2a_q} + \frac{is}{a_q} \left(\frac{5}{3} - 2\gamma - \log|s| - \frac{2\big(q^3 \log q + (1-q)^3 \log(1-q)\big)}{q^3 + (1-q)^3} \right) + o(|s|)$$

$$= 1 - \frac{\pi|s|}{2a_q} + \frac{is}{a_q} \left(- \log|s| + \log a_q + b_q + a_q \big(q - \tfrac{1}{2}\big)^2 \right) + o(|s|),$$

for $s = a_q t/c \to 0$, which we obtain from a careful analysis of the Fourier transforms $\int_0^\infty e^{isz} f_q(z) \, dz$, quite similar to the case of standard rounding in HPS (p. 54–55). □

References

[1] M.L. Balinski and S.T. Rachev. Rounding proportions: Rules of rounding. *Numerical Functional Analysis and Optimization*, 14:475–501, 1993.

[2] M.L. Balinski and S.T. Rachev. Rounding proportions: Methods of rounding. *Mathematical Scientist*, 22:1–26, 1997.

[3] M.L. Balinski and H.P. Young. *Fair Representation – Meeting the Ideal of One Man, One Vote.* New Haven CT: Yale University Press, 1982. [Second Edition (with identical pagination): Washington DC: Brookings Institution Press, 2001.]

[4] G. Dorfleitner and T. Klein. Rounding with multiplier methods: An efficient algorithm and applications in statistics. *Statistical Papers*, 40:143–157, 1999.

[5] M. Drton and U. Schwingenschlögl. Surface volumes of rounding polytopes. *Linear Algebra and Its Applications*, 378:71–91, 2004.

[6] M. Drton and U. Schwingenschlögl. Asymptotic seat bias formulas. *Metrika*, forthcoming, 2005.

[7] J. Gfeller. Du transfer des suffrages et de la réparition des sièges complémentaires. *Répresentation proportionelle*, 9:120–131, 1890.

[8] E. Hagenbach-Bischoff. *Die Verteilungsrechnung beim Basler Gesetz nach dem Grundsatz der Verhältniswahl.* Basel: Buchdruckerei zum Basler Berichthaus, 1905.

[9] M. Happacher. Die Verteilung der Diskrepanz bei stationären Multiplikatorverfahren zur Rundung von Wahrscheinlichkeiten. *Augsburger Mathematisch–Naturwissenschaftliche Schriften* 9. Augsburg: Wißner, 1996.

[10] M. Happacher. The discrepancy distribution of stationary multiplier rules for rounding probabilities. *Metrika*, 53:171–181, 2001.

[11] M. Happacher and F. Pukelsheim. Rounding probabilities: Unbiased multipliers. *Statistics & Decisions*, 14:373–382, 1996.

[12] M. Happacher and F. Pukelsheim. Rounding probabilities: Maximum probability and minimum complexity multipliers. *Journal of Statistical Planning and Inference*, 85:145–158, 2000.

[13] L. Heinrich, F. Pukelsheim, and U. Schwingenschlögl. Sainte-Laguë's chi-square divergence for the rounding of probabilities and its convergence to a stable law. *Statistics & Decisions*, 22:43–59, 2004.

[14] L. Heinrich and U. Schwingenschlögl. Goodness-of-fit criteria for the Adams and Jefferson rounding methods and their limiting laws. *Metrika*, forthcoming, 2006.

[15] N.L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous Univariate Distributions, Volume 1*. New York: Wiley, 1994.

[16] J. Kerstan and K. Matthes. Gleichverteilungseigenschaften von Faltungen von Verteilungsgesetzen auf lokal-kompakten abelschen Gruppen I. *Mathematische Nachrichten*, 37:267–312, 1968.

[17] A.W. Marshall, I. Olkin, and F. Pukelsheim. A majorization comparison of apportionment methods in proportional representation. *Social Choice and Welfare*, 19:885–900, 2002.

[18] T. Oyama. On the parametric divisor method for the apportionment problem. *Journal of the Operations Research Society of Japan*, 34:187–221, 1991.

[19] T. Oyama and T. Ichimori. On the unbiasedness of the parametric divisor method for the apportionment problem. *Journal of the Operations Research Society of Japan*, 38:301–321, 1995.

[20] A. Palomarez and V. Ramírez. Thresholds of the divisor methods. *Numerical Algorithms*, 34:405–415, 2003.

[21] F. Pukelsheim. Mandatszuteilungen bei Verhältniswahlen: Erfolgswertgleichheit der Wählerstimmen. *Allgemeines Statistisches Archiv – Journal of the German Statistical Society*, 84:447–459, 2000.

[22] V. Ramírez, M.L. Márquez, and R. Pérez. Parametric subfamilies of apportionment methods. *Advances in Computational Mathematics – Proceedings of the Guangzhou International Symposium on Computational Mathematics*, Lecture Notes in Pure and Applied Mathematics Series, 202:471–479, New York, 1999.

[23] A. Sainte-Laguë. La représentation proportionnelle et la méthode des moindres carrés. *Annales Scientifiques de l'École Normale Supérieure, Troisième Série*, 27: 529–542, 1910.

[24] K. Schuster, F. Pukelsheim, M. Drton, and N.R. Draper. Seat biases of apportionment methods for proportional representation. *Electoral Studies*, 22:651–676, 2003.

[25] U. Schwingenschlögl and M. Drton. Seat allocation distributions and seat biases of stationary apportionment methods for proportional representation. *Metrika*, 60:191–202, 2004.

[26] U. Schwingenschlögl and F. Pukelsheim. Seat biases in proportional representation systems with thresholds. *Social Choice and Welfare*, forthcoming, 2005.

[27] W.A. Wallis and H.V. Roberts. *Statistics – A New Approach.* London: Methuen, 1956.

Lothar Heinrich
Institut für Mathematik
Universität Augsburg
86135 Augsburg
Germany
heinrich@math.uni-augsburg.de

Friedrich Pukelsheim
Institut für Mathematik
Universität Augsburg
86135 Augsburg
Germany
pukelsheim@math.uni-augsburg.de

Udo Schwingenschlögl
Institut für Physik
Universität Augsburg
86135 Augsburg
Germany
schwingenschloegl@physik.uni-augsburg.de