Limiting Behavior of High Order Correlations for Simple Random Sampling

Christopher Wayne Walker

University of Southern California and
Northrop Grumman Aerospace Systems

Abstract

For $N = 1, 2, \ldots$, let S_N be a simple random sample of size $n = n_N$ from a population A_N of size N, where $0 \leq n \leq N$. Then with $f_N = n/N$, the sampling fraction, and $\mathbf{1}_A$ the inclusion indicator that $A \in S_N$, for any $H \subset A_N$ of size $k \geq 0$, the high order correlations

$$\mathrm{Corr}(k) = E\left(\prod_{A \in H} (\mathbf{1}_A - f_N)\right)$$

depend only on k, and if the sampling fraction $f_N \to f$ as $N \to \infty$, then

$$N^{k/2}\mathrm{Corr}(k) \to [f(f-1)]^{k/2} EZ^k, \; k \textit{ even}$$

and

$$N^{(k+1)/2}\mathrm{Corr}(k) \to [f(f-1)]^{(k-1)/2}(2f-1)\frac{1}{3}(k-1)EZ^{k+1}, \; k \textit{ odd}$$

where Z is a standard normal random variable. This proves a conjecture given in [2].

[0]cwwalker@cwwphd.com

[0]AMS 2000 subject classifications. Primary 05A16 , 62D05

[0]Key words and phrases: Normal moments, Stirling numbers, Bernoulli numbers

[0]This paper appeared in Annals of Combinatorics, Springer Basel 2013.

1 Introduction

Simple random sampling is doubtless one of the most often used tools in statistics [7]. With $0 \le n \le N$, by a simple random sample of size n from a set A_N of size N we mean the random subset of S_N of A_N with distribution

$$P(S_N = r) = \begin{cases} \binom{N}{n}^{-1} & r \subset A_N, |r| = n \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

where $|r|$ denotes the size of the set r. It is easy to see that all individuals in A_N have an equal chance of being included in the sample, and that in particular, for $A \in A_N$,

$$E\mathbf{1}_A = f_N \quad \text{where} \quad f_N = \frac{n}{N}, \tag{2}$$

where the inclusion indicator $\mathbf{1}_A$ takes the value 1 when $A \in S_N$ and the value 0 otherwise. The value f_N is known as the sampling fraction. Likewise, from (1) one can show directly that the inclusion indicators $\mathbf{1}_A$ and $\mathbf{1}_B$, for distinct individuals A and B in A_N, are negatively correlated, as the inclusion of A leaves less room in the remaining sample for B. That is,

$$E\mathbf{1}_A\mathbf{1}_B = \frac{n(n-1)}{N(N-1)} < \left(\frac{n}{N}\right)^2 = E\mathbf{1}_A E\mathbf{1}_B,$$

or, considering the two-way correlation

$$\mathrm{Corr}(\mathbf{1}_A, \mathbf{1}_B) = E\left((\mathbf{1}_A - f_N)(\mathbf{1}_B - f_N)\right), \tag{3}$$

we have

$$\mathrm{Corr}(\mathbf{1}_A, \mathbf{1}_B) = \frac{-n(N-n)}{N^2(N-1)}. \tag{4}$$

However, the higher order correlations of simple random sampling, which arise in some applications [2] and exhibit rather interesting behavior, are virtually unknown.

To consider such correlations of higher order, generalizing (3), for any $H \subset A_N$ of size $|H| = k, 0 \le k \le n$, we define

$$\mathrm{Corr}(H) = E\left(\prod_{A \in H}(\mathbf{1}_A - f_N)\right). \tag{5}$$

We see from (1) that the probability that all individuals in H are included in the sample is

$$E\left(\prod_{A\in H}\mathbf{1}_A\right)=\frac{\binom{N-k}{n-k}}{\binom{N}{n}}, \tag{6}$$

which we note only depends on k, and not on which individuals comprise the set H. Hence, arguing either directly using (6), or by noting the indicators $\{\mathbf{1}_A, A\in A_N\}$ are exchangeable, we conclude that $\mathrm{Corr}(H)$ depends only on the size k of H, and hence we denote it by $\mathrm{Corr}(k)$.

In [2] the high order correlation of rejective sampling was studied in order to determine the asymptotic properties of a generalized logistic estimator. In rejective sampling [6], each individual A in A_N is associated with a non-negative weight w_A, and the probability of sampling a set $r\subset A_N$ of size n is given by

$$P(S_n=r)=\frac{w_r}{\sum_{s\subset A_N,|s|=n}w_s} \quad \text{where} \quad w_s=\prod_{j\in s}w_j.$$

We note that the high order correlations of rejective sampling may be defined exactly as in (5), and that simple random sampling is the special case of rejective sampling, taking all weights equal.

Critical in the asymptotic analysis in [2] was the fact that under some stability conditions on the weights, the second and third order correlations of rejective sampling decay at rates N^{-1} and N^{-2}, respectively, that is,

$$\lim_{N\to\infty}N\mathrm{Corr}(2)=O(1) \quad \text{and} \quad \lim_{N\to\infty}N^2\mathrm{Corr}(3)=O(1).$$

Checking for the special case of simple random sampling when $f_N\to f$ as $N\to\infty$, equality (4) implies $N\mathrm{Corr}(2)\to f(1-f)$, and further direct calculation for correlations up to order 9 obtained by expanding the expression in definition (5) and using (6) yields

$$\begin{array}{rclrcl}
N\mathrm{Corr}(2) & \to & f(f-1), & N^2\mathrm{Corr}(3) & \to & 2f(f-1)(2f-1)\\
N^2\mathrm{Corr}(4) & \to & 3f^2(f-1)^2, & N^3\mathrm{Corr}(5) & \to & 20f^2(f-1)^2(2f-1)\\
N^3\mathrm{Corr}(6) & \to & 15f^3(f-1)^3, & N^4\mathrm{Corr}(7) & \to & 210f^3(f-1)^3(2f-1)\\
N^4\mathrm{Corr}(8) & \to & 105f^4(f-1)^4, & N^5\mathrm{Corr}(9) & \to & 2520f^4(f-1)^4(2f-1).
\end{array}$$

Perhaps the most surprising feature of these correlations is that their rate of decay depends on the parity of the correlation order, in particular, one can easily conjecture that

$$N^{(k+k\,mod\,2)/2}\text{Corr}(k) = O(1) \quad k = 1, 2, \ldots. \tag{7}$$

Theorem 4.1 of [2] shows that (7) holds quite generally for rejective sampling, and therefore for simple random sampling in particular. Application of this theorem sufficed to complete the asymptotic analysis required in [2] for rejective sampling. A recent investigation of rejective sampling can be found in [3] where an approximation formula is provided for the higher order inclusion probabilities and bounds are derived on higher order correlations.

Another feature of the simple random sampling correlations is also easy to conjecture, that their scaled limits are equal to a constant depending on k, times the factor $f(f-1)$ raised to $(k - k\,mod\,2)/2$, and if k is odd, times the additional factor $(2f-1)$. Hence, one need only determine the constants to completely specify their limiting behavior. On the basis of the above observations and the constants corresponding to the even and odd values of k up to 9, that is, the sequences $1, 3, 15, 105$ and $2, 20, 210, 2520$, respectively, a conjecture was put forth in [2], which is now validated by the following theorem which is proven in this analysis.

Theorem 1.1. *For $N = 1, 2, \ldots$ let S_N be a sequence of simple random samples from populations A_N of size N, whose sampling fractions f_N obey*

$$f_N \to f \in (0,1) \quad \textit{as } N \to \infty.$$

Then

$$\lim_{N\to\infty} N^{k/2}\,Corr(k) = [f(f-1)]^{k/2} EZ^k, \; k \textit{ even}$$

and

$$\lim_{N\to\infty} N^{(k+1)/2}\,Corr(k) = [f(f-1)]^{(k-1)/2}(2f-1)\frac{1}{3}(k-1)EZ^{k+1}, \; k \textit{ odd}$$

where Z is a standard normal variate.

2 Main Result

The proof of Theorem 1.1 requires a few identities for the Stirling and Bernoulli numbers, which can be found in [5]. Letting $(x)_j$ denote the falling factorial, or Pochhammer symbol, and expanding $(x)_j$ as a polynomial in x

we have $(x)_j = x(x-1)\cdots(x-j+1) = \sum_{v=0}^{j} \left[{j \atop v}\right] (-1)^{j-v} x^v$ where $\left[{j \atop v}\right]$ is the (unsigned) Stirling numbers of the first kind; for instance,

$$\begin{bmatrix} j \\ j \end{bmatrix} = 1, \quad \begin{bmatrix} j \\ j-1 \end{bmatrix} = \frac{1}{2} j(j-1), \quad \begin{bmatrix} j \\ 1 \end{bmatrix} = (j-1)! \quad \text{and} \quad \begin{bmatrix} j \\ 0 \end{bmatrix} = 0.$$

Regarding Stirling numbers of the second kind, denoted by $\left\{{m \atop k}\right\}$, we make use of the identity

$$\sum_{j=0}^{k} \binom{k}{j} (-1)^j j^m = (-1)^k k! \left\{ {m \atop k} \right\}. \tag{8}$$

In particular we note that

$$\left\{ {m \atop k} \right\} = 0 \text{ for } m < k, \quad \left\{ {k \atop k} \right\} = 1, \quad \text{and} \quad \left\{ {k+1 \atop k} \right\} = \binom{k+1}{2}. \tag{9}$$

Furthermore,

$$\sum_{p=0}^{k-1} p^m = \frac{1}{m+1} \sum_{p=0}^{m} \binom{m+1}{p} B_p k^{m+1-p} \tag{10}$$

where B_p denotes the p^{th} Bernoulli number defined by the explicit recurrence

$$\sum_{j=0}^{m} \binom{m+1}{j} B_j = \delta(m), \tag{11}$$

where $\delta(m)$ is the Kronecker delta function, taking the value 1 for $m = 0$ and zero otherwise.

For any integer m let $r_{\boldsymbol{\alpha},m}(x)$ denote a polynomial in x of degree no more than $\max(m, 0)$ whose coefficients are a function of $\boldsymbol{\alpha}$, possibly a vector; the polynomial $r_{\boldsymbol{\alpha},m}(x)$ is not necessarily the same at each occurrence. Note with this notation that $r_{\boldsymbol{\alpha},m}(x) + r_{\boldsymbol{\alpha},m}(x) = r_{\boldsymbol{\alpha},m}(x)$ and if $f(\boldsymbol{\alpha})$ is a function of $\boldsymbol{\alpha}$ not involving x then $f(\boldsymbol{\alpha}) r_{\boldsymbol{\alpha},m}(x) = r_{\boldsymbol{\alpha},m}(x)$. From (11) one easily finds that $B_0 = 1$ and $B_1 = -1/2$, and hence (10) yields

$$\sum_{p=0}^{k-1} p^m = \frac{1}{m+1} k^{m+1} - \frac{1}{2} k^m + r_{m,m-1}(k). \tag{12}$$

Other examples that illustrate this notation are

$$k^4 - 2km + m^2 + j^2/m - 2j = r_{(k,m),2}(j)$$

and

$$k^4 - 2km + m = r_{(k,m),0}(j).$$

Similarly, let $r_{\boldsymbol{\alpha},m,n,p}(x,y)$ denote a polynomial in x and y whose coefficients are a function of $\boldsymbol{\alpha}$ where the highest power of x that occurs is at most $\max(m,0)$, the highest power of y that occurs is at most $\max(n,0)$, and the highest sum $a+b$ in terms of the form $x^a y^b$ is at most $\max(p,0)$; again the polynomial $r_{\boldsymbol{\alpha},m,n,p}(x,y)$ is not necessarily the same at each occurrence. Examples that illustrate this notation are

$$k^4 + k^3 + k^2 j^2 + j^3 + j^4 = r_{1,4,4,4}(j,k)$$

and

$$k^4 + mk^4 j^2 + k^2 j^2 + j^6 = r_{m,6,4,6}(j,k) = r_{(k,m),6}(j).$$

An identity that can be obtained directly from Equation 5.114 in [5] is

$$\sum_{n=0}^{m-1} (-1)^n \binom{m}{n} = (-1)^{m-1}$$

for $m \geq 1$. We can manipulate this last expression to deduce

$$\sum_{n=0}^{m-1} (-1)^n \binom{m}{n+1} = u(m-1) \tag{13}$$

where $u(n)$ is the unit step function defined to be 1 when $n \geq 0$ and is zero otherwise.

Using the well-known identity $\binom{m}{n+1} = \binom{m-1}{n+1} + \binom{m-1}{n}$ and (13) we easily find

$$\sum_{n=0}^{m-1} (-1)^n \binom{m-1}{n} = \delta(m-1). \tag{14}$$

The following can be found in [4]:

$$\sum_{n=0}^{m-1} (-1)^n \binom{m-1}{n} \frac{1}{n+\beta} = \frac{\Gamma(m)\Gamma(\beta)}{\Gamma(m+\beta)} u(m-1) \tag{15}$$

where, $\Gamma(m)$ denotes the gamma function. We note from [1] that for positive integers m,

$$\Gamma(m) = (m-1)!, \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \quad \Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}, \tag{16}$$

and we also find

$$\Gamma\left(m + \frac{1}{2}\right) = \frac{(2m-1)!}{2^{2m-1}(m-1)!}\sqrt{\pi}. \tag{17}$$

Using (14) and (15) we get

$$\sum_{n=0}^{m-1} (-1)^n \binom{m-1}{n} \frac{n}{n+\beta} = \delta(m-1) - \frac{\beta\Gamma(m)\Gamma(\beta)}{\Gamma(m+\beta)} u(m-1). \tag{18}$$

Combining (15) and (18) we obtain for real numbers $\alpha, \delta, \gamma, \beta$ and integer m such that $\gamma \neq 0$ and $\frac{\beta}{\gamma} \neq 0, -1, -2, \ldots, -(m-1)$,

$$\begin{aligned} &\sum_{n=0}^{m-1} (-1)^n \binom{m-1}{n} \frac{\alpha n + \delta}{\gamma n + \beta} \\ &= \frac{\alpha}{\gamma}\delta(m-1) + \frac{\Gamma(m)\Gamma\left(\frac{\beta}{\gamma}\right)}{\Gamma\left(m + \frac{\beta}{\gamma}\right)} \frac{1}{\gamma} \left(\delta - \frac{\alpha\beta}{\gamma}\right) u(m-1). \end{aligned} \tag{19}$$

We now derive another result that will be useful. Using (12) and the binomial theorem we can write for integers $n, p, s \geq 0$,

$$\begin{aligned} \sum_{q=1}^{k-m} q(q+1)^{pn+s} &= \sum_{q=1}^{k-m} \sum_{j=0}^{pn+s} \binom{pn+s}{j} q^{pn+s+1-j} \\ &= \sum_{q=1}^{k-m} \left(q^{pn+s+1} + (pn+s)q^{pn+s} + r_{(n,p,s),pn+s-1}(q) \right) \\ &= \frac{k^{pn+s+2}}{pn+s+2} \\ &\quad + \frac{((pn+s)(3-2m) + 1 - 2m)\, k^{pn+s+1}}{2(pn+s+1)} \\ &\quad + r_{(n,m,p,s),pn+s}(k). \end{aligned} \tag{20}$$

If we let $k = j$ in this last identity and set $m = 1$ we get

$$\sum_{q=1}^{j-1} q(q+1)^{pn+s} = \frac{j^{pn+s+2}}{pn+s+2} + \frac{(pn+s-1)j^{pn+s+1}}{2(pn+s+1)} + r_{(n,p,s),pn+s}(j). \tag{21}$$

For the following definition we adopt the empty sum convention that

$$\sum_{q=a}^{b} x_q = 0 \quad \text{when } b < a.$$

Definition 2.1. *For nonnegative integers k, j and m, let*

$$P_{k,0}(j) = 1,$$
$$P_{k,m}(j) = \sum_{q=1}^{k-m} qP_{k,m-1}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1}(q+1), \ m \geq 1 \tag{22}$$

and

$$P^0_{k,0}(j) = 1 \quad \textit{and} \quad P^0_{k,m}(j) = \sum_{q=j}^{k-m} qP^0_{k,m-1}(q+1), \ m \geq 1.$$

The following lemma shows that the identity $P_{k,1}(k) = 0$ is a specific instance of a more general fact.

Lemma 2.1. *For all nonnegative integers k, j and m,*

$$P_{k,m}(j) = 0 \quad \textit{for } 0 \leq k - m + 1 \leq j \leq k, \tag{23}$$

and

$$P_{k,m}(j) = P^0_{k,m}(j) \quad \textit{for} \quad 0 \leq j \leq k. \tag{24}$$

Furthermore, for all integers k, m and j satisfying $0 \leq m, j \leq k$,

$$P_{k,m}(j) = (-1)^m \left(\frac{j^{2m} + \frac{1}{3}m(2m-5)j^{2m-1}}{2^m m!} \right) + r_{(k,m),2m-2}(j), \tag{25}$$

and for integers j and v satisfying $0 \leq v \leq j$,

$$P_{j,v}(1) = \frac{j^{2v} - \frac{1}{3}v(2v+1)j^{2v-1}}{2^v v!} + r_{v,2v-2}(j). \tag{26}$$

Proof. First note that (23) implies (24), as the equality holds for $0 \le j \le k-m$ by construction. To argue by induction, we have that (23) holds when $m = 0$, as in this case the premise $0 \le k - m + 1 \le j \le k$ is vacuous. Now assume (23) holds for $m \ge 0$. From Definition 2.1, for $k - m \le j \le k$,

$$\begin{aligned} P_{k,m+1}(j) &= \sum_{q=1}^{k-m-1} qP_{k,m}(q+1) - \sum_{q=1}^{j-1} qP_{k,m}(q+1) \\ &= -\sum_{q=k-m}^{j-1} qP_{k,m}(q+1) = -\sum_{q=k-m+1}^{j} (q-1)P_{k,m}(q). \end{aligned}$$

For $k - m + 1 \le q \le j \le k$ we have $P_{k,m}(q) = 0$ by (23). Hence,

$$P_{k,m+1}(j) = 0 \quad \text{for } k - m \le j \le k,$$

which is (23) with $m + 1$ replacing m, completing the proof of the first two claims.

To prove the rest of the lemma we will show that

$$\begin{aligned} P_{k,m}(j) &= \sum_{n=0}^{m} \frac{(-1)^n j^{2n} k^{2m-2n}}{2^m (m-n)!n!} \\ &\quad + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)!n!} \\ &\quad + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m (m-1-n)!n!} \\ &\quad + r_{m,2m-2,2m-2,2m-2}(j,k) \end{aligned} \tag{27}$$

from which both (25) and (26) will easily follow. Let

$$D_{n,m} = 2^m (m-n)!n!.$$

We will use induction on m in our proof. Clearly (27) is true for $m = 0$. Now assume (27) is true for some $m-1 \geq 0$. Then, to evaluate $P_{k,m}(j)$ let

$$\begin{aligned}
P_{k,m;1}(j) &= \sum_{n=0}^{m} \frac{(-1)^n j^{2n} k^{2m-2n}}{2^m (m-n)! n!} \\
P_{k,m;2}(j) &= \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)! n!} \\
P_{k,m;3}(j) &= \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m (m-1-n)! n!} \\
P_{k,m;4}(j) &= r_{m,2m-2,2m-2,2m-2}(j,k)
\end{aligned}$$

so that (22) becomes

$$P_{k,m}(j) = \sum_{q=1}^{k-m} q \sum_{t=1}^{4} P_{k,m-1;t}(q+1) - \sum_{q=1}^{j-1} q \sum_{t=1}^{4} P_{k,m-1;t}(q+1). \tag{28}$$

We will evaluate the terms in (28) separately. We find for $t=1$

$$\begin{aligned}
&\sum_{q=1}^{k-m} q P_{k,m-1;1}(q+1) - \sum_{q=1}^{j-1} q P_{k,m-1;1}(q+1) \\
&= \sum_{n=0}^{m-1} \frac{(-1)^n k^{2m-2-2n}}{D_{n,m-1}} \sum_{q=1}^{k-m} q(q+1)^{2n} \\
&\quad - \sum_{n=0}^{m-1} \frac{(-1)^n k^{2m-2-2n}}{D_{n,m-1}} \sum_{q=1}^{j-1} q(q+1)^{2n}.
\end{aligned} \tag{29}$$

Using (20) and (21) with $p=2$ and $s=0$, (29) becomes

$$\begin{aligned}
&\sum_{q=1}^{k-m} q P_{k,m-1;1}(q+1) - \sum_{q=1}^{j-1} q P_{k,m-1;1}(q+1) \\
&= \frac{k^{2m}}{2^m} \sum_{n=0}^{m-1} \frac{(-1)^n}{(m-1-n)!(n+1)!} + \frac{k^{2m-1}}{2^m} \sum_{n=0}^{m-1} \frac{(-1)^n (2n(3-2m)+1-2m)}{(2n+1)(m-1-n)! n!} \\
&\quad - \sum_{n=0}^{m-1} \frac{(-1)^n j^{2n+2} k^{2m-2-2n}}{2^m (m-1-n)!(n+1)!} - \sum_{n=0}^{m-1} \frac{(-1)^n (2n-1) j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)! n! (2n+1)} \\
&\quad + r_{m,2m-2,2m-2,2m-2}(j,k)
\end{aligned}$$

which becomes

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;1}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;1}(q+1) \\
&= \frac{k^{2m}}{2^m m!}\sum_{n=0}^{m-1}(-1)^n\binom{m}{n+1} \\
&\quad +\frac{k^{2m-1}}{2^m(m-1)!}\sum_{n=0}^{m-1}(-1)^n\binom{m-1}{n}\frac{2(3-2m)n+1-2m}{2n+1} \\
&\quad +\sum_{n=1}^{m}\frac{(-1)^n j^{2n}k^{2m-2n}}{2^m(m-n)!n!} - \sum_{n=0}^{m-1}\frac{(-1)^n j^{2n+1}k^{2m-2-2n}}{2^m(m-1-n)!n!}\frac{(2n-1)}{(2n+1)} \\
&\quad +r_{m,2m-2,2m-2,2m-2}(j,k).
\end{aligned}
$$

Using (13) and (19) with $\alpha = 2(3-2m),\ \delta = 1-2m,\ \gamma = 2$, and $\beta = 1$ this becomes

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;1}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;1}(q+1) \\
&= \frac{k^{2m}}{2^m m!}u(m-1) - \frac{k^{2m-1}}{2^m(m-1)!}\left(\frac{\Gamma(m)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(m+\frac{1}{2}\right)}u(m-1) - (3-2m)\delta(m-1)\right) \\
&\quad +\sum_{n=1}^{m}\frac{(-1)^n j^{2n}k^{2m-2n}}{2^m(m-n)!n!} - \sum_{n=0}^{m-1}\frac{(-1)^n j^{2n+1}k^{2m-2-2n}}{2^m(m-1-n)!n!}\frac{(2n-1)}{(2n+1)} \\
&\quad +r_{m,2m-2,2m-2,2m-2}(j,k).
\end{aligned}
$$

Using (16) and (17) we get

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;1}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;1}(q+1) \\
&= \sum_{n=0}^{m}\frac{(-1)^n j^{2n}k^{2m-2n}}{2^m(m-n)!n!}u(m-1) \qquad (30) \\
&\quad -\frac{2^{m-1}(m-1)!k^{2m-1}}{(2m-1)!}u(m-1) + \frac{k^{2m-1}}{2^m(m-1)!}(3-2m)\delta(m-1) \\
&\quad -\sum_{n=0}^{m-1}\frac{(-1)^n j^{2n+1}k^{2m-2-2n}}{2^m(m-1-n)!n!}\frac{(2n-1)}{(2n+1)} + r_{m,2m-2,2m-2,2m-2}(j,k).
\end{aligned}
$$

We next find from (28) with $t=2$

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;2}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;2}(q+1)\\
&= \sum_{n=0}^{m-2}\sum_{q=1}^{k-m} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}q(q+1)^{2n+1}k^{2m-4-2n}}{2^{m-1}(m-2-n)!n!}\\
&\quad -\sum_{n=0}^{m-2}\sum_{q=1}^{j-1} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}q(q+1)^{2n+1}k^{2m-4-2n}}{2^{m-1}(m-2-n)!n!}.
\end{aligned}
$$

Using (20) and (21) with $p=2$ and $s=1$ we get

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;2}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;2}(q+1)\\
&= \sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}k^{2m-1}}{2^{m-1}(m-2-n)!n!(2n+3)} + r_{m,2m-2}(k)\\
&\quad -\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}j^{2n+3}k^{2m-4-2n}}{2^{m-1}(m-2-n)!n!(2n+3)}\\
&\quad +r_{m,2m-2,2m-4,2m-2}(j,k)\\
&= -\frac{\frac{1}{3}k^{2m-1}}{2^{m-1}(m-2)!}\sum_{n=0}^{m-2}(-1)^n\binom{m-2}{n}\frac{2n-3}{2n+3}\\
&\quad -\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}j^{2n+3}k^{2m-4-2n}}{2^{m-1}(m-2-n)!n!(2n+3)}\\
&\quad +r_{m,2m-2,2m-2,2m-2}(j,k).
\end{aligned}
$$

Using (19) with $m=m-1,\ \alpha=2,\ \delta=-3,\ \gamma=2,\ \beta=3$ and then using (16) and (17) we obtain

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;2}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;2}(q+1)\\
&= \frac{(m-1)!2^{m-1}k^{2m-1}}{(2m-1)!}u(m-2) - \frac{\frac{1}{3}k^{2m-1}}{2^{m-1}(m-2)!}\delta(m-2) \qquad (31)\\
&\quad -\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1}j^{2n+3}k^{2m-4-2n}}{2^{m-1}(m-2-n)!n!(2n+3)} + r_{m,2m-2,2m-2,2m-2}(j,k).
\end{aligned}
$$

Next we find from (28) with $t = 3$

$$\sum_{q=1}^{k-m} qP_{k,m-1;3}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;3}(q+1)$$
$$= \sum_{q=1}^{k-m}\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-2m+1)(-1)^n q(q+1)^{2n}k^{2m-3-2n}}{2^{m-1}(m-2-n)!n!}$$
$$-\sum_{q=1}^{j-1}\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-2m+1)(-1)^n q(q+1)^{2n}k^{2m-3-2n}}{2^{m-1}(m-2-n)!n!}.$$

Using (20) and (21) with $p = 2$ and $s = 0$ this becomes

$$\sum_{q=1}^{k-m} qP_{k,m-1;3}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;3}(q+1)$$
$$= \sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-2m+1)(-1)^n k^{2m-1}}{2^{m-1}(m-2-n)!n!(2n+2)} + r_{m,2m-2}(k)$$
$$-\sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-2m+1)(-1)^n j^{2n+2}k^{2m-3-2n}}{2^{m-1}(m-2-n)!n!(2n+2)} + r_{m,2m-3,2m-3,2m-2}(j,k)$$
$$= \frac{\frac{1}{3}k^{2m-1}}{2^m(m-2)!}\sum_{n=0}^{m-2}(-1)^n\binom{m-2}{n}\frac{2n+1-2m}{n+1}$$
$$+\sum_{n=1}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n}k^{2m-1-2n}}{2^m(m-1-n)!n!} + r_{m,2m-3,2m-2,2m-2}(j,k).$$

Using (19) with m replaced by $m-1$, $\alpha = 2$, $\delta = 1-2m$, $\gamma = 1$ and $\beta = 1$

this becomes

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;3}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;3}(q+1) \\
&= \frac{\frac{1}{3}k^{2m-1}}{2^m(m-2)!}\left[\frac{\Gamma(m-1)\Gamma(1)}{\Gamma(m-1+1)}(1-2m-2)u(m-2)+2\delta(m-2)\right] \\
&\quad + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m(m-1-n)!n!} \\
&\quad - \frac{\frac{1}{3}(-2m-1)k^{2m-1}}{2^m(m-1)!}u(m-1) + r_{m,2m-3,2m-2,2m-2}(j,k)
\end{aligned}
$$

or

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;3}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;3}(q+1) \\
&= \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m(m-1-n)!n!} \qquad (32) \\
&\quad + \frac{\frac{2}{3}k^{2m-1}}{2^m(m-2)!}\delta(m-2) + \frac{\frac{1}{3}(2m+1)k^{2m-1}}{2^m(m-1)!}\delta(m-1) \\
&\quad + r_{m,2m-3,2m-2,2m-2}(j,k).
\end{aligned}
$$

Next we find from (28) with $t = 4$

$$
\begin{aligned}
&\sum_{q=1}^{k-m} qP_{k,m-1;4}(q+1) - \sum_{q=1}^{j-1} qP_{k,m-1;4}(q+1) \\
&= \sum_{q=1}^{k-m} q\left(r_{m,2m-4,2m-4,2m-4}(q+1,k)\right) \\
&\quad - \sum_{q=1}^{j-1} q\left(r_{m,2m-4,2m-4,2m-4}(q+1,k)\right) \\
&= r_{m,2m-2}(k) + r_{m,2m-2,2m-4,2m-2}(j,k) \\
&= r_{m,2m-2,2m-2,2m-2}(j,k). \qquad (33)
\end{aligned}
$$

Combining (30), (31), (32) and (33) and neglecting terms in the resulting expression that sum to zero and observing

$$
\begin{aligned}
& - \sum_{n=0}^{m-1} \frac{(-1)^n j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)! n!} \frac{(2n-1)}{(2n+1)} - \sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} j^{2n+3} k^{2m-4-2n}}{2^{m-1}(m-2-n)! n! (2n+3)} \\
& = \sum_{n=0}^{m-1} \frac{(-1)^{n+1} \frac{1}{3}(2n-3) j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)! n!}
\end{aligned}
$$

and noting that $m-1 \geq 0$ in this induction proof we get

$$
\begin{aligned}
& \sum_{q=1}^{k-m} q P_{k,m-1}(q+1) - \sum_{q=1}^{j-1} q P_{k,m-1}(q+1) \\
& \quad = \sum_{n=0}^{m} \frac{(-1)^n j^{2n} k^{2m-2n}}{2^m (m-n)! n!} + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} j^{2n+1} k^{2m-2-2n}}{2^{m-1}(m-1-n)! n!} \\
& \qquad + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m (m-1-n)! n!} + r_{m,2m-2,2m-2,2m-2}(j,k)
\end{aligned}
$$

which proves (27).

Now we may write (27) as

$$
\begin{aligned}
& P_{k,m}(j) \\
& \quad = \frac{(-1)^m j^{2m}}{2^m m!} + \sum_{n=0}^{m-1} \frac{(-1)^n j^{2n} k^{2m-2n}}{2^m (m-n)! n!} \\
& \qquad + \frac{\frac{1}{3}(2m-5)(-1)^m j^{2m-1}}{2^m (m-1)!} + \sum_{n=0}^{m-2} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} j^{2n+1} k^{2m-2-2n}}{2^m (m-1-n)! n!} \\
& \qquad + \sum_{n=0}^{m-1} \frac{\frac{1}{3}(2n-2m-1)(-1)^n j^{2n} k^{2m-1-2n}}{2^m (m-1-n)! n!} + r_{m,2m-2,2m-2,2m-2}(j,k) \\
& \quad = (-1)^m \left(\frac{j^{2m} + \frac{1}{3} m (2m-5) j^{2m-1}}{2^m m!} \right) + r_{(k,m),2m-2}(j)
\end{aligned}
$$

which proves (25). Now interchanging k and j and letting $m = v$ (27)

becomes

$$
\begin{aligned}
&P_{j,v}(k) \\
&= \sum_{n=0}^{v} \frac{(-1)^n k^{2n} j^{2v-2n}}{2^v (v-n)! n!} + \sum_{n=0}^{v-1} \frac{\frac{1}{3}(2n-3)(-1)^{n+1} k^{2n+1} j^{2v-2-2n}}{2^v (v-1-n)! n!} \\
&\quad + \sum_{n=0}^{v-1} \frac{\frac{1}{3}(2n-2v-1)(-1)^n k^{2n} j^{2v-1-2n}}{2^v (v-1-n)! n!} + r_{v,2v-2,2v-2,2v-2}(k,j)
\end{aligned}
$$

so

$$P_{j,v}(1) = \frac{j^{2v} - \frac{1}{3} v(2v+1) j^{2v-1}}{2^v v!} + r_{v,2v-2}(j)$$

which proves (26). The proof of the lemma is now complete. □

Lemma 2.2. *With $P^0_{j,v}$ given in Definition 2.1,*

$$(f_N N)_j = \sum_{v=0}^{j} (-1)^v P^0_{j,v}(1) (f_N N)^{j-v} \quad \text{for } j \geq 0, \tag{34}$$

and

$$(N-k)_{j-k} = \sum_{v=0}^{j-k} (-1)^v P^0_{j,v}(k) N^{j-k-v} \quad \text{for } j-k \geq 0. \tag{35}$$

Proof. First note that the validity of (35) establishes (34) since with $k=0$, $P^0_{j,v}(0) = P^0_{j,v}(1)$. Expanding $(x-k)_j$ in $x \in \mathbb{R}$ for $k \in \mathbb{N}$ and $j \geq 0$ we obtain $(x-k)_j = \sum_{v=0}^{j} (-1)^v a_{k+j-1,v}(k) x^{j-v}$, where

$$a_{j,v}(u) = \sum_{u \leq l_1 < l_2 < \cdots < l_v \leq j} \prod_{m=1}^{v} l_m$$

with any empty product set to 1. Comparison with (34) shows that it suffices to prove

$$a_{k+j-1,v}(k) = P^0_{j,v}(k). \tag{36}$$

Equality holds for $v = 0$. Assuming (36) holds for some $v - 1 \geq 0$, we have

$$\begin{aligned}
a_{k+j-1,v}(k) &= \sum_{k \leq l_1 < l_2 < \cdots < l_v \leq k+j-1} \prod_{m=1}^{v} l_m \\
&= \sum_{k \leq q \leq k+j-v} q \sum_{l_2,\ldots,l_v : q < l_2 < \cdots < l_v \leq k+j-1} \prod_{m=2}^{v} l_m \\
&= \sum_{k \leq q \leq k+j-v} q \sum_{l_1,\ldots,l_{v-1} : q+1 \leq l_1 < l_2 < \cdots < l_{v-1} \leq k+j-1} \prod_{m=1}^{v-1} l_m \\
&= \sum_{q=k}^{k+j-v} q a_{k+j-1,v-1}(q+1).
\end{aligned}$$

As $P^0_{j,v}(k)$ is characterized by the same recursion, the inductive step is complete, proving (35). □

Proof of Theorem 1.1. Expanding out the product in (5), by (6) and (2) we have

$$\begin{aligned}
\mathrm{Corr}(k) &= E\left(\prod_{A \in H}(I_A - f_N)\right) = E \sum_{G \subset H} \left(\prod_{A \in G} I_A\right)(-f_N)^{|H \setminus G|} \\
&= \sum_{G \subset H} \frac{\binom{N-|G|}{n-|G|}}{\binom{N}{n}} (-f_N)^{|H|-|G|} = \sum_{j=0}^{k} \sum_{G \subset H, |G|=j} \frac{\binom{N-j}{n-j}}{\binom{N}{n}} (-f_N)^{k-j} \\
&= \sum_{j=0}^{k} \frac{\binom{k}{j}\binom{N-j}{n-j}}{\binom{N}{n}} (-f_N)^{k-j} = \sum_{j=0}^{k} \binom{k}{j} \frac{(n)_j}{(N)_j} (-f_N)^{k-j} \\
&= \frac{\sum_{j=0}^{k} \binom{k}{j} (n)_j (N-j)_{k-j} (-f_N)^{k-j}}{(N)_k} = \frac{\alpha(k, f_N)}{(N)_k},
\end{aligned}$$

where

$$\alpha(k, f_N) = (-f_N)^k \sum_{j=0}^{k} \binom{k}{j} (-1)^j \lambda(k, j, f_N), \quad \text{for } f_N \in (0,1) \tag{37}$$

and

$$\lambda(k, j, f_N) = f_N^{-j} (f_N N)_j (N-j)_{k-j}, \quad \text{for } j = 0, 1, 2, \ldots, k. \tag{38}$$

We also write

$$\mathrm{Corr}(k, f_N) = \frac{\alpha(k, f_N)}{(N)_k} \quad \text{for } f_N \in (0,1). \tag{39}$$

Using (34) and (35), and recalling that $P^0_{j,v}(1) = 0$ when $v < 0$ then (38) becomes

$$\begin{aligned}
\lambda(k, j, f_N) &= f_N^{-j} \left(\sum_{v=0}^{j} (-1)^v P^0_{j,v}(1)(f_N N)^{j-v} \right) \left(\sum_{i=0}^{k-j} (-1)^i P^0_{k,i}(j) N^{k-j-i} \right) \\
&= f_N^{-j} \sum_{r=0}^{k} \left(\sum_{v=j-r}^{k-r} (-1)^v P^0_{j,v}(1) f_N^{j-v} (-1)^{k-v-r} P^0_{k,k-v-r}(j) \right) N^r \\
&= \sum_{r=0}^{k} \sum_{v=j-r}^{k-r} f_N^{-v} P^0_{j,v}(1)(-1)^{k-r} P^0_{k,k-v-r}(j) N^r \\
&= \sum_{v=j-k}^{k} f_N^{-v} P^0_{j,v}(1) \sum_{r=j-v}^{k-v} (-1)^{k-r} P^0_{k,k-v-r}(j) N^r \\
&= \sum_{v=0}^{k} f_N^{-v} P^0_{j,v}(1) \sum_{r=j-v}^{k-v} (-1)^{k-r} P^0_{k,k-v-r}(j) N^r,
\end{aligned}$$

since $P^0_{j,v}(1) = 0$ for $v < 0$.

For a function represented as the power series $\mu(f) = \sum_{j=-\infty}^{\infty} c_j f^j$, for $k \in \mathbb{Z}$ let $\mu(f; k)$ denote the coefficient of f^k in $\mu(f)$, that is, $\mu(f; k) = c_k$.

Making the change of variable $m = k - r$ in the first sum below, and again using that $P^0_{k,m}(j) = 0$ for $m < 0$, for $0 \le v \le k$ we obtain

$$\begin{aligned}
\lambda(k, j, f_N; -v) &= P^0_{j,v}(1) \sum_{r=j-v}^{k-v} (-1)^{k-r} P^0_{k,k-v-r}(j) N^r \\
&= P^0_{j,v}(1) \sum_{m=k-j+v}^{v} (-1)^m P^0_{k,m-v}(j) N^{k-m} \\
&= P^0_{j,v}(1) \sum_{m=v}^{k-j+v} (-1)^m P^0_{k,m-v}(j) N^{k-m} \\
&= \sum_{m=v}^{k-j+v} \lambda(k, j, f_N; -v : m)
\end{aligned} \tag{40}$$

where $\lambda(k, j, f_N; -v : m) = (-1)^m P^0_{j,v}(1) P^0_{k,m-v}(j) N^{k-m}$.

Now, by (37) and (40),

$$\begin{aligned}
&\alpha(k, f_N; k-v) \\
&= (-1)^k \sum_{j=0}^{k} \binom{k}{j} (-1)^j \lambda(k, j, f_N; -v) \\
&= (-1)^k \sum_{j=0}^{k} \sum_{m=v}^{k-j+v} \binom{k}{j} (-1)^j \lambda(k, j, f_N; -v : m) \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k-m+v} \binom{k}{j} (-1)^j \lambda(k, j, f_N; -v : m) \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k-m+v} \binom{k}{j} (-1)^{j+m} P^0_{j,v}(1) P^0_{k,m-v}(j) N^{k-m} \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+m} P^0_{j,v}(1) P^0_{k,m-v}(j) N^{k-m} \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+m} P_{j,v}(1) P_{k,m-v}(j) N^{k-m}, \qquad (41)
\end{aligned}$$

where in the last two equalities we invoke Lemma 2.1 to apply $P^0_{k,m-v}(j) = 0$ for $j > k-m+v$, and $P^0_{j,v}(1) = P_{j,v}(1)$, $P^0_{k,m-v}(j) = P_{k,m-v}(j)$ for $0 \le j \le k$.

At this point in our proof we will consider k even and k odd separately. For k even it will be convenient to write (25) and (26), respectively, as

$$P_{k,m}(j) = (-1)^m \frac{j^{2m}}{2^m m!} + r_{(k,m),2m-1}(j),$$

and

$$P_{j,v}(1) = \frac{j^{2v}}{2^v v!} + r_{v,2v-1}(j).$$

Substituting these last two expressions into (41) we get

$$\begin{aligned}\alpha(k, f_N; k-v) &= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \left[\binom{k}{j} (-1)^{j+m} \left(\frac{j^{2v}}{2^v v!} + r_{v,2v-1}(j) \right) \right. \\ &\quad \left. \times \left(\frac{(-1)^{m-v} j^{2(m-v)}}{2^{m-v}(m-v)!} + r_{(k,m,v),2(m-v)-1}(j) \right) \right] N^{k-m} \\ &= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \left[\binom{k}{j} (-1)^{j+m} \right. \\ &\quad \left. \times \left(\frac{(-1)^{m-v} j^{2m}}{2^m v!(m-v)!} + r_{(k,m,v),2m-1}(j) \right) \right] N^{k-m}. \qquad (42)\end{aligned}$$

From (39) and (42) we obtain

$$\begin{aligned}&N^{k/2}\mathrm{Corr}(k, f_N; k-v) \\ &= N^{k/2} \frac{\alpha(k, f_N; k-v)}{(N)_k} \\ &= \frac{N^{k/2}}{(N)_k} (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \left[\binom{k}{j} (-1)^{j+m} \right. \\ &\quad \left. \times \left(\frac{(-1)^{m-v} j^{2m}}{2^m v!(m-v)!} + r_{(k,m,v),2m-1}(j) \right) \right] N^{k-m} \\ &= \frac{N^{k/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \sum_{m=v}^{k+v} \left[\frac{1}{2^m (m-v)!} \sum_{j=0}^{k} \binom{k}{j} (-1)^j j^{2m} \right. \qquad (43) \\ &\quad \left. + \sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-1}(j) \right] N^{k-m}.\end{aligned}$$

Using (8), (43) becomes

$$\begin{aligned}&N^{k/2}\mathrm{Corr}(k, f_N; k-v) \\ &= \frac{N^{k/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \left(\sum_{m=v}^{k+v} \frac{1}{2^m (m-v)!} (-1)^k k! \left\{ \begin{matrix} 2m \\ k \end{matrix} \right\} \right. \qquad (44) \\ &\quad \left. + \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-1}(j) \right) N^{k-m}.\end{aligned}$$

Since $(N)_k$ is of degree k, then when $m > k/2$,

$$\lim_{N\to\infty} \frac{N^{k/2}}{(N)_k} N^{k-m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{k/2}}{N^m} = 0$$

and when $m \le k/2$, from (8) and (9) we have that

$$\sum_{j=0}^{k} \binom{k}{j} (-1)^j j^p = (-1)^k k! \left\{ {p \atop k} \right\} = 0 \quad \text{for all } p \le 2m-1.$$

Therefore, by linearity,

$$\sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-1}(j) = 0 \quad \text{for } m \le k/2,$$

and since for k even, $\left\{ {2m \atop k} \right\} = 0$ for $m < k/2$, then upon letting $N \to \infty$, (44) is zero for $v > k/2$ and for $v \le k/2$ (44) becomes

$$\lim_{N\to\infty} N^{k/2} \mathrm{Corr}(k, f_N; k-v) \tag{45}$$

$$= \lim_{N\to\infty} \frac{N^{k/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \left(\sum_{m=k/2}^{k/2} \frac{1}{2^m (m-v)!} (-1)^k k! \left\{ {2m \atop k} \right\} \right) N^{k-m}.$$

When $m = k/2$,

$$\lim_{N\to\infty} \frac{N^{k/2}}{(N)_k} N^{k-m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{k/2}}{N^m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{k/2}}{N^{k/2}} = 1.$$

Therefore, using (9), (45) becomes

$$\lim_{N\to\infty} N^{k/2} \cdot \mathrm{Corr}(k, f_N; k-v)$$

$$= (-1)^v \frac{k!}{v! 2^{k/2} \left(\frac{k}{2} - v\right)!} = (-1)^v \frac{k!}{2^{k/2} \left(\frac{k}{2}\right)!} \binom{k/2}{v} = (-1)^v E Z^k \binom{k/2}{v}.$$

Now, by our previous notation, $\mathrm{Corr}(k) = \mathrm{Corr}(k, f_N)$, so

$$\lim_{N\to\infty} N^{k/2} \cdot \mathrm{Corr}(k)$$

$$= \lim_{N\to\infty} N^{k/2} \cdot \mathrm{Corr}(k, f_N) = \lim_{N\to\infty} N^{k/2} \sum_{v=0}^{k} \mathrm{Corr}(k, f_N; k-v) f_N^{k-v}$$

$$= \sum_{v=0}^{k/2} \binom{k/2}{v} f^{k-v} (-1)^v E Z^k = [f(f-1)]^{k/2} E Z^k$$

which proves the theorem for k even.

For k odd we have from (25), (26) and (41) that

$$\begin{aligned}
&\alpha(k, f_N; k-v) \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+m} \left(\frac{j^{2v} - \frac{1}{3}v(2v+1)j^{2v-1}}{2^v v!} + r_{v,2v-2}(j) \right) \\
&\quad \times \left[(-1)^{m-v} \left(\frac{j^{2(m-v)} + \frac{1}{3}(m-v)(2(m-v)-5)j^{2(m-v)-1}}{2^{m-v}(m-v)!} \right) \right. \\
&\quad \left. + r_{(k,m,v),2(m-v)-2}(j) \right] N^{k-m}
\end{aligned}$$

which becomes

$$\begin{aligned}
&\alpha(k, f_N; k-v) \\
&= (-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+m} \\
&\quad \times \left[(-1)^{m-v} \left(\frac{j^{2m} + \frac{1}{3}\left[(m-v)(2m-2v-5) - v(2v+1)\right] j^{2m-1}}{2^m v!(m-v)!} \right) \right. \\
&\quad \left. + r_{(k,m,v),2m-2}(j) \right] N^{k-m}.
\end{aligned}$$

From (39) and (42) we obtain

$$N^{(k+1)/2}\text{Corr}(k, f_N; k-v) = N^{(k+1)/2}\frac{\alpha(k, f_N; k-v)}{(N)_k}$$

$$= \frac{N^{(k+1)/2}}{(N)_k}(-1)^k \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^{j+m}$$

$$\times \left[(-1)^{m-v} \left(\frac{j^{2m} + \frac{1}{3}\left[(m-v)(2m-2v-5) - v(2v+1)\right] j^{2m-1}}{2^m v!(m-v)!} \right) \right.$$

$$\left. + r_{(k,m,v),2m-2}(j) \right] N^{k-m}$$

$$= \frac{N^{(k+1)/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \sum_{m=v}^{k+v} \left[\frac{1}{2^m (m-v)!} \sum_{j=0}^{k} \binom{k}{j} (-1)^j \right. \tag{46}$$

$$\times \left(j^{2m} + \frac{1}{3}\left[(m-v)(2m-2v-5) - v(2v+1)\right] j^{2m-1} \right)$$

$$\left. + \sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-2}(j) \right] N^{k-m}.$$

Using (8), (46) becomes

$$N^{(k+1)/2}\text{Corr}(k, f_N; k-v)$$

$$= \frac{N^{(k+1)/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \left[\sum_{m=v}^{k+v} \frac{1}{2^m (m-v)!} (-1)^k k! \left\{ \begin{matrix} 2m \\ k \end{matrix} \right\} \right. \tag{47}$$

$$+ \sum_{m=v}^{k+v} \frac{\frac{1}{3}\left[(m-v)(2m-2v-5) - v(2v+1)\right]}{2^m (m-v)!} (-1)^k k! \left\{ \begin{matrix} 2m-1 \\ k \end{matrix} \right\}$$

$$\left. + \sum_{m=v}^{k+v} \sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-2}(j) \right] N^{k-m}.$$

Since $(N)_k$ is of degree k, then when $m > (k+1)/2$,

$$\lim_{N\to\infty} \frac{N^{(k+1)/2}}{(N)_k} N^{k-m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{(k+1)/2}}{N^m} = 0$$

and when $m \le (k+1)/2$, from (8) and (9) we have that

$$\sum_{j=0}^{k} \binom{k}{j} (-1)^j j^p = (-1)^k k! \left\{ {p \atop k} \right\} = 0 \quad \text{for all } p \le 2m-2.$$

Therefore, by linearity,

$$\sum_{j=0}^{k} \binom{k}{j} (-1)^j r_{(k,m,v),2m-2}(j) = 0 \quad \text{for } m \le (k+1)/2,$$

and since for k odd, $\left\{ {2m \atop k} \right\} = 0$ for $m < (k+1)/2$ and $\left\{ {2m-1 \atop k} \right\} = 0$ for $m < (k+1)/2$, then upon letting $N \to \infty$ (47) is zero for $v > (k+1)/2$ and for $v \le (k+1)/2$ (47) becomes

$$\begin{aligned} &\lim_{N\to\infty} N^{(k+1)/2} \mathrm{Corr}(k, f_N; k-v) \\ &= \lim_{N\to\infty} \frac{N^{(k+1)/2}}{(N)_k} \frac{(-1)^{k-v}}{v!} \left[\sum_{m=(k+1)/2}^{(k+1)/2} \frac{1}{2^m (m-v)!} (-1)^k k! \left\{ {2m \atop k} \right\} \right. \\ &\quad \left. + \sum_{m=(k+1)/2}^{(k+1)/2} \frac{\frac{1}{3}\left[(m-v)(2m-2v-5) - v(2v+1)\right]}{2^m (m-v)!} (-1)^k k! \left\{ {2m-1 \atop k} \right\} \right] N^{k-m}. \end{aligned} \tag{48}$$

When $m = (k+1)/2$,

$$\lim_{N\to\infty} \frac{N^{(k+1)/2}}{(N)_k} N^{k-m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{(k+1)/2}}{N^m} = \lim_{N\to\infty} \frac{N^k}{(N)_k} \frac{N^{(k+1)/2}}{N^{(k+1)/2}} = 1.$$

Therefore,(48) becomes

$$\begin{aligned} &\lim_{N\to\infty} N^{(k+1)/2} \cdot \mathrm{Corr}(k, f_N; k-v) \\ &= \frac{(-1)^{k-v}}{v!} \left[\frac{1}{2^{(k+1)/2} \left(\frac{k+1}{2} - v\right)!} (-1)^k k! \left\{ {k+1 \atop k} \right\} \right. \\ &\quad \left. + \frac{\frac{1}{3}\left[\left(\frac{k+1}{2} - v\right)(k+1-2v-5) - v(2v+1)\right]}{2^{(k+1)/2} \left(\frac{k+1}{2} - v\right)!} (-1)^k k! \left\{ {k \atop k} \right\} \right]. \end{aligned}$$

Using (9) this last result becomes

$$\begin{aligned}
&\lim_{N\to\infty} N^{(k+1)/2}\cdot \mathrm{Corr}(k, f_N; k-v) \\
&= \frac{\frac{1}{3}(-1)^v k!}{v!2^{(k+1)/2}\left(\frac{k+1}{2}-v\right)!}\left[3\binom{k+1}{2} + \left(\frac{k+1}{2}-v\right)(k-2v-4) - v(2v+1)\right] \\
&= \frac{\frac{2}{3}(-1)^v (k-1)(k+1-v)k!}{v!2^{(k+1)/2}\left(\frac{k+1}{2}-v\right)!}.
\end{aligned}$$

We may write this last result as

$$\begin{aligned}
&\lim_{N\to\infty} N^{(k+1)/2}\cdot \mathrm{Corr}(k, f_N; k-v) \\
&= \frac{\frac{2}{3}(-1)^v (k-1)(k+1-v)k!\binom{\frac{k+1}{2}}{v}}{2^{(k+1)/2}\left(\frac{k+1}{2}\right)!} \\
&= \frac{\frac{2}{3}(-1)^v (k-1)(k+1-v)EZ^{k+1}}{k+1}\binom{\frac{k+1}{2}}{v}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{N\to\infty} N^{(k+1)/2}\cdot \mathrm{Corr}(k) &= \lim_{N\to\infty} N^{(k+1)/2}\cdot \mathrm{Corr}(k, f_N) \\
&= \lim_{N\to\infty} N^{(k+1)/2}\sum_{v=0}^{k} \mathrm{Corr}(k, f_N; k-v) f_N^{k-v} \\
&= \sum_{v=0}^{k} \frac{\frac{2}{3}(-1)^v (k-1)(k+1-v)}{k+1}\binom{\frac{k+1}{2}}{v} f^{k-v} EZ^{k+1} \\
&= \frac{\frac{2}{3}(k-1)}{k+1} EZ^{k+1}\frac{d}{df}\left(\sum_{v=0}^{(k+1)/2} (-1)^v \binom{\frac{k+1}{2}}{v} f^{k+1-v}\right) \\
&= \frac{\frac{2}{3}(k-1)}{k+1} EZ^{k+1}\frac{d}{df}\left([f(f-1)]^{(k+1)/2}\right) \\
&= [f(f-1)]^{(k-1)/2}(2f-1)\frac{1}{3}(k-1)EZ^{k+1}
\end{aligned}$$

which proves the theorem for k odd.

The proof of the theorem is now complete. □

Acknowledgments. I would like to express my sincere gratitude to Professor Larry Goldstein of the Department of Mathematics at the University of Southern California for suggesting this topic and for his many valuable comments that greatly improved this manuscript.

References

1. Abramowitz, M., and Stegun, I.: Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables. Dover Publications, New York (1966)

2. Arraita, R., Goldstein, L., and Langholz, B. (2005). Local central limit theorems, the high-order correlations of rejective sampling and logistic likelihood asymptotics. *Ann. Statist.*, **33**, pp. 871–914.

3. Boistard, H., Lopuhaä, H.P., and Ruiz-Gazen, A.: Approximation of rejective sampling inclusion probabilities and application to high order correlations. Electron. J. Stat. 6, 1967-1983 (2012)

4. Garrappa, R. (2007) Some Formulas for Sums of Binomial Coefficients and Gamma Functions. *International Mathematical Forum*, **2**, pp. 725-733.

5. Graham, R., Knuth, D., and Patashnik, O. (1994). Concrete Mathematics: A Foundation for Computer Science. Second edition. Addison-Wesley Publishing Company, Reading, MA.

6. Hájek, J. (1964) Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Statist.*, **35**, pp. 1491–1523.

7. Hájek, J. (1981) Sampling from a finite population. Edited by Vclav Dupač. With a foreword by P. K. Sen. Statistics: Textbooks and Monographs, 37. Marcel Dekker, Inc., New York.