Sankhyā : The Indian Journal of Statistics
2007, Volume 69, Part 2, pp. 162-189

The Discrete Mohr and Noll Inequality with Applications to Variance Bounds

G. Afendras, N. Papadatos and V. Papathanasiou
University of Athens, Greece

Abstract

In this paper, we provide Poincaré-type upper and lower variance bounds for a function $g(X)$ of a discrete integer-valued random variable (r.v.) X, in terms of the (forward) differences of g up to some order. To this end, we investigate a discrete analogue of the Mohr and Noll inequality (1952, *Math. Nachr.*, vol. 7, pp. 55–59), which may be of some independent interest in itself. It has been shown by Johnson (1993, *Statist. Decisions*, vol. 11, pp. 273–278) that for the commonly used absolutely continuous distributions that belong to the Pearson family, the somewhat complicated variance bounds take a very pleasant and simple form. We show here that this is also true for the commonly used discrete distributions. As an application of the proposed inequalities, we study the variance behaviour of the UMVU estimator of $\log p$ in Geometric distributions.

AMS (2000) *subject classification.* Primary 60E15.
Keywords and phrases. Poincaré-type variance bounds, Mohr and Noll inequality, discrete Pearson family, UMVUE of $\log p$ in geometric distribution.

1 Introduction and a Brief Review for the Continuous Case

For a n-times continuously differentiable function $g : (0,1) \longrightarrow \mathbb{R}$ with derivatives $g^{(j)}(x) = d^j g(x)/dx^j$, $j = 0, 1, \ldots, n$, Mohr and Noll (1952) obtained, as an extension of Schwarz's inequality for integrals, the identity

$$\left(\int_0^1 g(t)dt\right)^2 = \sum_{k=0}^{n-1} \frac{(-1)^k}{k!(k+1)!} \int_0^1 (g^{(k)}(t))^2 t^k (1-t)^k dt + (-1)^n R_n \quad (1.1)$$

(provided that the integrals are finite), where the remainder

$$R_n = \int \cdots \int_{0<x_1<\cdots<x_n<y_1<\cdots<y_n<1} \left(\int_{x_n}^{y_1} g^{(n)}(t)dt\right)^2 dy_n \cdots dy_1 dx_n \cdots dx_1 \quad (1.2)$$

is clearly nonnegative. The sum in the RHS of (1.1) (without the remainder term $(-1)^n R_n$) provides a lower (upper) bound for the LHS if n is even (odd), and the equality holds only for polynomials of degree at most $n-1$.

Since the median of an i.i.d. standard uniform sample of odd size $n = 2k+1$ has a Beta density, which is a multiple of $t^k(1-t)^k$, the inequality (1.1) has the following statistical representation:

PROPOSITION 1.1. *Let $U = U_{1:1}$ be a r.v. uniformly distributed over the interval $(0,1)$, and denote by $U_{1:m} < U_{2:m} < \cdots < U_{m:m}$ the corresponding order statistics from a sample of size m. Fix $n \in \{1,2,\ldots\}$, and assume that the function $g : (0,1) \longrightarrow \mathbb{R}$ satisfies the following conditions.*

(i) *For $k = 0,1,\ldots,n-2$, the function $g^{(k)}$ (where $g^{(0)} = g$) is absolutely continuous with a.s. derivative $g^{(k+1)}$.*

(ii) *The r.v.'s $g^{(k)}(U_{k+1:2k+1})$, $k = 0,1,\ldots n-1$, have finite variance, i.e.,*

$$\sum_{k=0}^{n-1} \int_0^1 x^k(1-x)^k (g^{(k)}(x))^2 dx < \infty.$$

Then,

$$\mathbb{E}^2 g(U) = \sum_{k=0}^{n-1} \frac{(-1)^k}{(k+1)(2k+1)!} \mathbb{E}(g^{(k)}(U_{k+1:2k+1}))^2 + (-1)^n R_n, \qquad (1.3)$$

where

$$R_n = \frac{1}{(2n)!} \mathbb{E}\Big(g^{(n-1)}(U_{n+1:2n}) - g^{(n-1)}(U_{n:2n})\Big)^2. \qquad (1.4)$$

Clearly $R_n = 0$ iff g is a polynomial of degree at most $n-1$. Note that (i) does not impose any condition on g when $n = 1$; on the other hand, it can be shown, with the help of the inequalities $(b-a)^2 \le 2a^2 + 2b^2$ and $\max\{x^n(1-x)^{n-1}, x^{n-1}(1-x)^n\} < x^{n-1}(1-x)^{n-1}$, $0 < x < 1$, that the condition $\mathbb{E}(g^{(n-1)}(U_{n:2n-1}))^2 < \infty$, imposed by (ii), suffices for the remainder R_n, in (1.4), to be finite.

We omit the proof of Proposition 1.1, since it is just a restatement of (1.1). However, we note that the conditions on g in Proposition 1.1 are slightly more general than those imposed by the original proof of Mohr and Noll. For instance, $g^{(n)}$ in (1.2) may not exists, and $g^{(n-1)}$ need not be continuous. In the sequel, we call (1.1) or (1.3) as the Mohr and Noll inequality.

Papathanasiou (1988), with the help of the Mohr and Noll inequality, derived a class of variance bounds for all absolutely continuous r.v.'s X with density f and finite moment of order $2n+2$, where $n \in \{0, 1, \ldots\}$ is fixed. Specifically, let g be a $(n+1)$-times continuously differentiable real function, defined on $(r, s) := (\operatorname{ess\,inf}(X), \operatorname{ess\,sup}(X))$, the minimal open interval containing the support of X, and assume that $g(X)$ has finite variance. In this case, Papathanasiou (1988) obtained the following inequality (provided that the integrals are finite):

$$(-1)^n[S_n - \operatorname{Var} g(X)] \ge 0, \text{ where } S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{(k-1)!k!} \int_r^s a_{k-1}(t)(g^{(k)}(t))^2 dt. \tag{1.5}$$

The bound (1.5) depends on the positive functions $a_k(t)$, $t \in (r, s)$, given by

$$a_k(t) = (-1)^k[\mu_{k+1}(t) I_k(t) - \mu_k(t) I_{k+1}(t)], \quad k = 0, 1, \ldots, n, \tag{1.6}$$

where

$$\mu_k(t) = \mathbb{E}(X-t)^k, \quad I_k(t) = \int_r^t (x-t)^k f(x)dx = \mathbb{E}[(X-t)^k I(X < t)]. \tag{1.7}$$

Moreover, the equality in (1.5) holds if and only if the function g is a polynomial of degree at most $n+1$.

Clearly, the explicit form of $a_k(t)$ is important for the calculation of the bound S_n in (1.5) (note that $a_0(t) = \int_{-\infty}^t (\mu - x) f(x) dx$ with $\mu = \mathbb{E}X$). Papathanasiou (1988) gave explicit forms of $a_k(t)$ only for Normal and Gamma r.v.'s, e.g., if X is $N(\mu, \sigma^2)$, then $a_k(t) = k!\,\sigma^{2k+2} f(t)$ so that

$$(-1)^n[S_n - \operatorname{Var} g(X)] \ge 0, \text{ where } S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k!} \sigma^{2k}\, \mathbb{E}(g^{(k)}(X))^2, \tag{1.8}$$

and where S_n becomes equal to $\operatorname{Var} g(X)$ only if g is a polynomial of degree at most $n+1$. It should be noted that the bound (1.8) has been independently obtained by Houdré and Kagan (1995), using a completely different method based on trigonometric polynomial approximation. Multivariate extensions, including Normal and Poisson distributions, can be found in

Houdré and Pérez-Abreu (1995), Houdré et al. (1998), Chang and Richards (1999), among others.

Johnson (1993) proved that the functions $a_k(t)$ have explicit forms when X is a Pearson variate. Specifically, for the Pearson family, it has been shown by Korwar (1991) that there exists a quadratic $q(t) = \delta t^2 + \beta t + \gamma$ such that $a_0(t) = q(t)f(t)$ (none of the parameters δ, β and γ has to be nonzero; the only restriction is that q remains strictly positive in the support (r, s)). Repeated use of this fact in combination with the covariance identity, $\mathrm{Cov}(X, g(X)) = \mathbb{E}[q(X)g'(X)]$, yields a recurrence in k so that $a_k(t) = k!c_k(\delta)q^{k+1}(t)f(t)$, where $c_k(\delta) = \prod_{j=0}^{k}(1 - j\delta)^{-1}$ depends only on k and δ. Note that the traditional way of describing the densities f of the Pearson family as those satisfying $f'(t)/f(t) = -\sigma^2 p_1(t)/q(t)$ for $t \in (r, s)$, where $p_1(t)$ is a monic linear polynomial and σ^2 the variance, limits the applicability of Johnson's result; e.g., the standard uniform distribution does not belong to this subclass, even though $q(t) = t(1 - t)/2$ for $t \in (0, 1)$. A careful reading of Johnson's derivation shows that his results continue to hold for such cases as well, since the simplification actually depends only upon the "magic relation" $\int_{-\infty}^{t}(\mu - x)f(x)dx = q(t)f(t)$. Therefore, Theorem 2 of Johnson (1993) can be restated as

$$\begin{aligned} (-1)^n[S_n - \mathrm{Var}\, g(X)] &\geq 0, \\ \text{where } S_n &= \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k! \prod_{j=0}^{k-1}(1 - j\delta)} \mathbb{E}[q^k(X)(g^{(k)}(X))^2]. \end{aligned} \tag{1.9}$$

The above bound can be applied, for instance, to Beta r.v.'s with any parameters $a > 0$ and $b > 0$ (not only with $a + b > 2$). Obviously, (1.8) is a very particular case of (1.9), since $q(t) = \sigma^2$ for the Normal r.v. Variance bounds of the form (1.9) are usually called Poincaré-type (see, e.g., Chen and Lou, 1987), and are especially useful in isoperimetric inequalities (Chernoff, 1981). Other applications concern characterizations and limit theorems (Chen, 1988; Chen and Lou, 1987), and the analysis of the behavior of future contracts (Siegel, 1993; Houdré, 1995). Similar variance inequalities for an iterated jackknife estimate of the variance have been obtained by Houdré (1997, Theorem 2.1).

The purpose of the present article is in deriving similar Poincaré-type bounds for the discrete case. Specifically, in Section 2, we state and prove a novel discrete Mohr and Noll inequality (Lemma 2.1; see also Corollary 3.1), which may be of some independent interest in itself. We also give its statistical interpretation, which, as in the continuous case, is connected with some properties of the sample medians from a uniform model of various sample sizes. In Section 3, we use this inequality to derive a discrete analogue of the bound (1.5), satisfied by any integer-valued r.v. with finite moments of the appropriate order.

In Section 4, we apply the discrete Mohr and Noll inequality to the Pearson family of integer-valued r.v.'s, and we obtain the discrete analogue of the Poincaré-type variance bound (1.9). Although the new bounds of Theorems 3.1 and 4.1 formally coincide with their continuous analogues (1.5) and (1.9), the discrete bounds are quite different in nature, and the appropriate inequality needed for their proof (see (3.1), below) is not very easy to recognize from its continuous counter-part. In fact, it can be shown, using Riemann integral, that the inequality of Corollary 3.1(a), applied to the function $g_N(j) = g(j/N)$, $j = 1, 2, \ldots, N$, implies the inequality in (1.1), as $N \to \infty$, while the converse is not true (see Remark 3.1). Perhaps this provides a reason for the fact that the only known results of this kind for the discrete case are those obtained for Poisson processes by Houdré and Pérez-Abreu (1995), using Malliavin calculus, and the corresponding correlation identities and inequalities for infinitely divisible variables, established by Houdré et al. (1998).

Finally, in Section 5, we present an application of the bounds in a particular situation concerning statistical inference. Specifically, the main result of Theorem 4.1 is applied in order to approximate the variance of the UMVU estimator of $\log p$ in Geometric distributions. We show that for any fixed n, the corresponding bounds, $S_n = S_n(\nu, p)$, behave well as the sample size, ν, becomes large. Also, we show that for any fixed sample size ν, the bounds $S_n(\nu, p)$ approximate the true variance as $n \to \infty$, if and only if $p \geq 1/2$. Some numerical values of the first few bounds are also given. The order of approximation of the bounds to the true variance seems to be quite satisfactory in this example, for small and moderate values of ν, and for various values of the parameter p.

2 The Discrete Mohr and Noll Inequality

Turning to the discrete case, consider a "real-valued function" (finite sequence) g defined on $\{1, 2, \ldots, N\}$, and denote its n-th forward difference as $g^{(n)}$ or $\Delta^n g$, i.e.,

$$g^{(n)}(x) = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} g(x+j), \quad n = 0, 1, \ldots, \quad x = 1, 2, \ldots, N-n.$$

For convenience, set $g^{(n)}(x) = 0$ if $x > \max\{0, N-n\}$. This definition is compatible with the fact that $g^{(n)}(x) = 0$ for all $x = 1, 2, \ldots$ if and only if $g(x)$ is a polynomial of degree at most $n-1$ (it is well-known that any function g defined on the set $\{1, 2, \ldots, N\}$ is the restriction of a polynomial of degree at most $N-1$). Denoting by $S(g^{(n)})$ the "true support" of $g^{(n)}$, i.e., the maximal set of x's where $g^{(n)}(x)$ is completely defined in terms of the values $g(1), \ldots, g(N)$, we observe that $S(g^{(n)}) = \{1, 2, \ldots, N-n\}$, if $n < N$, and $S(g^{(n)}) = \emptyset$ otherwise.

The appropriate uniform model is provided by the order statistics $U_{1:m}^{(N)} < U_{2:m}^{(N)} < \cdots < U_{m:m}^{(N)}$, obtained from a simple random sample of size m, taken without replacement from the discrete uniform r.v. $U_N = U_{1:1}^{(N)}$, assigning probability $1/N$ to each point $1, 2, \ldots, N$. This simple model, which clearly is not well-defined if $m > N$, has been recently studied by some authors, due to its useful connection with general order statistics from finite populations; see, e.g., Arnold et al. (1992), Kochar and Korwar (1997), (2001), Takahasi and Futatsuya (1998), Balakrishnan et al. (2003), Papadatos and Rychlik (2004), López-Blázquez and Castaño-Martínez (2006), among others.

In the sequel, we shall make use of the following properties:

$$I\!P[U_{n:2n}^{(N)}{=}k_1, U_{n+1:2n}^{(N)}{=}k_2] = \binom{k_1-1}{n-1}\binom{N-k_2}{n-1}\Big/\binom{N}{2n}, \quad 1 \le k_1 < k_2 \le N, \tag{2.1}$$

(provided $2n \le N$)

and

$$I\!P[U_{n-1:2n-1}^{(N)}{=}k] = \binom{k-1}{n-1}\binom{N-k}{n-1}\Big/\binom{N}{2n-1}, \qquad 1 \le k \le N, \tag{2.2}$$

(provided $2n-1 \le N$)

where the binomial coefficient $\binom{a}{b}$ should be treated as 0 whenever $a < b$.

Assume that $2k+1 \le N$ so that the sample median $U_{k+1:2k+1}^{(N)}$ is well-defined. Clearly, $k+1 \le U_{k+1:2k+1}^{(N)} \le N-k$, and thus, the support of $U_{k+1:2k+1}^{(N)}$ is strictly smaller than the "true support" $S(g^{(k)})$ of $g^{(k)}$ when $k > 0$. If we require sample medians to be supported in the set $S(g^{(k)})$, as in the continuous Mohr and Noll inequality (1.3), then a kind of standardization should be done. To this end, we define the "sample medians"

$$V_k = V_k^{(N)} := U_{k+1:2k+1}^{(N+k)} - k, \quad k = 0, 1, \ldots, N-1. \tag{2.3}$$

The standardized "sample medians" V_k can now assume any value in $S(g^{(k)}) = \{1, 2, \ldots, N-k\}$, with positive probability, for all $k \le N-1$ (i.e., for those values of k for which $S(g^{(k)})$ is nonempty). Similar arguments apply to the case of even sample sizes $2k+2$, $k = 0, 1, 2, \ldots$, where the standardized "sample medians" $(Z_k, W_k) = (Z_k^{(N)}, W_k^{(N)})$ are now defined as the random pairs

$$(Z_k, W_k) := \left(U_{k+1:2k+2}^{(N+k)} - k, \quad U_{k+2:2k+2}^{(N+k)} - k\right), \quad k = 0, 1, \ldots, N-2, \tag{2.4}$$

so that the support for each (Z_k, W_k) is the set $\{(i_1, i_2) : 1 \le i_1 < i_2 \le N-k\}$, and thus, the r.v. $g^{(k)}(W_k) - g^{(k)}(Z_k)$ is well-defined for $k \le N-2$.

We are now in a position to state and prove the discrete Mohr and Noll inequality in its probabilistic form.

Lemma 2.1. *Under the above notations, the following identity holds for* $n = 1, 2, \ldots$ *and for any function* g *defined on* $\{1, 2, \ldots, N\}$:

$$\mathbb{E}^2 g(U_N) = \frac{1}{N} \sum_{k=0}^{n-1} \frac{(-1)^k}{k+1} \binom{N+k}{2k+1} \mathbb{E}\left[g^{(k)}(V_k)\right]^2 + (-1)^n \frac{1}{N^2} R_n, \tag{2.5}$$

where $\binom{N+k}{2k+1} \mathbb{E}\left[g^{(k)}(V_k)\right]^2$ *should be treated as* 0 *if* $k > N-1$. *The remainder in* (2.5) *is*

$$R_n = \binom{N+n-1}{2n} \mathbb{E}\left[g^{(n-1)}(W_{n-1}) - g^{(n-1)}(Z_{n-1})\right]^2, \tag{2.6}$$

and it should be treated as 0 *if* $n > N-1$. *Moreover,* $R_n = 0$ *if and only if* g *is the restriction of a polynomial of degree at most* $n-1$.

Proof. We use induction on n. For $n = 1$, the identity (2.5) reduces to

$$\mathbb{E}^2 g(U_N) = \mathbb{E}[g(U_N)]^2 - \frac{N-1}{2N} \mathbb{E}\left[g(U_{2:2}^{(N)}) - g(U_{1:2}^{(N)})\right]^2,$$

which can be easily verified. (Clearly, $R_1 = 0$ if and only if g is a constant function.) Assuming that the assertion is true for some n, the induction step for $n+1$ will be deduced if it can be shown that

$$R_{n+1} + R_n = \frac{N}{n+1} \binom{N+n}{2n+1} \mathbb{E}\left[g^{(n)}(V_n)\right]^2. \tag{2.7}$$

Thus, we have nothing to show if $n \geq N$ (everything vanishes in (2.7)), while for $n \leq N-1$, using (2.1)–(2.4), we have

$$\begin{aligned}
R_n &= \sum_{n\leq k_1<k_2\leq N} \binom{k_1-1}{n-1}\binom{N+n-k_2-1}{n-1}\left[\sum_{j=k_1-n+1}^{k_2-n} g^{(n)}(j)\right]^2 \\
&= \sum_{j_1=1}^{N-n}\sum_{j_2=1}^{N-n} g^{(n)}(j_1)g^{(n)}(j_2)\binom{\min\{j_1,j_2\}+n-1}{n}\binom{N-\max\{j_1,j_2\}}{n} \\
&= \sum_{k=n+1}^{N} (g^{(n)}(k-n))^2\binom{k-1}{n}\binom{N+n-k}{n} \\
&\qquad +2\sum_{k_1<k_2} g^{(n)}(k_1-n)g^{(n)}(k_2-n)\binom{k_1-1}{n}\binom{N+n-k_2}{n} \\
&= \binom{N+n}{2n+1}\mathbb{E}\left[g^{(n)}(V_n)\right]^2 + 2\binom{N+n}{2n+2}\mathbb{E}\left[g^{(n)}(Z_n)g^{(n)}(W_n)\right].
\end{aligned}$$

Using the above expression for R_n and (2.6) with $n+1$ in place of n, we get

$$\begin{aligned}
R_{n+1}+R_n &= \binom{N+n}{2n+2}\left(\mathbb{E}\left[g^{(n)}(Z_n)\right]^2 + \mathbb{E}\left[g^{(n)}(W_n)\right]^2\right) \\
&\qquad + \binom{N+n}{2n+1}\mathbb{E}\left[g^{(n)}(V_n)\right]^2. \qquad (2.8)
\end{aligned}$$

Observe that the first term in the RHS of (2.8) vanishes if $n=N-1$, and thus, (2.7) follows immediately from (2.8). Otherwise, if $n \leq N-2$, the desired equality (2.7) follows again by (2.8), if we apply for the function $h=(g^{(n)})^2$ the identity

$$\mathbb{E}h(Z_n)+\mathbb{E}h(W_n)=2\mathbb{E}h(V_n), \quad n=0,1,\ldots,N-2,$$

which can be obtained from (2.1) and (2.2), since $\mathbb{P}[Z_n=k]+\mathbb{P}[W_n=k]=2\mathbb{P}[V_n=k]$ for all k. This proves the identity (2.5) for all n.

It is now obvious from the form of the remainder that $R_n=0$ if and only if either $n\geq N$ and g is completely arbitrary, or $n\leq N-1$ and $g^{(n-1)}(y)-g^{(n-1)}(x)$ vanishes for all $(x,y)\in\{(i_1,i_2):1\leq i_1<i_2\leq N-n+1\}$, the support of (Z_{n-1},W_{n-1}). This can happen if and only if $g^{(n)}$ identically vanishes in its "true support", i.e., if and only if g is the restriction of a polynomial of degree at most $n-1$, and the proof is complete. $\square$

3 Application to Variance Bounds

For any $x \in \mathbb{R}$ and $k = 1, 2, \ldots$, we define $(x)_k = x(x-1)\cdots(x-k+1)$, $[x]_k = x(x+1)\cdots(x+k-1) = (x+k-1)_k$, and, as usual, we set $(x)_0 = [x]_0 = 1$ for all x. The next result is evident from Lemma 2.1.

Corollary 3.1. (a) *For $n = 0, 1, \ldots$ and any function g defined in $\{1, 2, \ldots, N\}$,*

$$\left(\sum_{j=1}^{N} g(j)\right)^2 = N \sum_{k=0}^{n} \frac{(-1)^k}{k!(k+1)!} \sum_{j=1}^{N-k} (g^{(k)}(j))^2 [j]_k (N-j)_k + (-1)^{n+1} R_{n+1},$$

where

$$R_{n+1} = \frac{1}{(n!)^2} \sum_{1 \le j_1 < j_2 \le N-n} (g^{(n)}(j_2) - g^{(n)}(j_1))^2 [j_1]_n (N-j_2)_n,$$

and where empty sums are treated as 0.
(b) *For $n = 0, 1, \ldots$ and for any function g defined on $\{x, x+1, \ldots, y\}$, where $x < y$ are assumed to be integers,*

$$(g(y)-g(x))^2 = (y-x) \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{(k-1)!k!} \sum_{j=x}^{y-k} (g^{(k)}(j))^2 [j-x+1]_{k-1} (y-j-1)_{k-1} + (-1)^{n+1} R_{n+1}(x,y), \qquad (3.1)$$

where

$$R_{n+1}(x,y) = \frac{1}{(n!)^2} \sum_{x \le j_1 < j_2 \le y-1-n} (g^{(n+1)}(j_2) - g^{(n+1)}(j_1))^2 [j_1-x+1]_n (y-j_2-1)_n,$$

and where empty sums are treated as 0.

Remark 3.1. (a) The inequality suggested by Corollary 3.1(a) can be written as

$$(-1)^n \left[\left(\frac{1}{N} \sum_{j=1}^{N} h(j) \right)^2 - \sum_{k=0}^{n} \frac{(-1)^k}{k!(k+1)!} \left\{ \frac{1}{N} \sum_{j=1}^{N-k} \frac{[j]_k}{N^k} \frac{(N-j)_k}{N^k} (N^k \Delta^k h(j))^2 \right\} \right]$$

$\le 0,$

for any $h : \{1, 2, \ldots, N\} \longrightarrow \mathbb{R}$, while the original Mohr and Noll inequality suggests that

$$(-1)^n \left[\left(\int_0^1 g(t)dt \right)^2 - \sum_{k=0}^{n} \frac{(-1)^k}{k!(k+1)!} \left\{ \int_0^1 t^k (1-t)^k (g^{(k)}(t))^2 dt \right\} \right] \leq 0,$$

for all n-times continuously differentiable $g : (0, 1) \longrightarrow \mathbb{R}$. It can be shown that the discrete inequality is stronger than the continuous one. To see this, assume first that the functions $g, g', \ldots, g^{(n)}$ are defined in the compact interval $[0, 1]$, and are (uniformly) continuous on $[0, 1]$. Then apply the discrete inequality to the function $h_N(j) = g(j/N)$, $j = 1, 2 \ldots, N$, observing that for any $j = 1, 2, \ldots, N - k$ and any $\epsilon > 0$,

$$\begin{aligned} & |g^{(k)}(j/N) - N^k \Delta^k h_N(j)| \\ = \ & N^k \left| \int_{j/N}^{(j+1)/N} \int_{t_1}^{t_1+1/N} \cdots \int_{t_{k-1}}^{t_{k-1}+1/N} [g^{(k)}(j/N) - g^{(k)}(t_k)] dt_k \ldots dt_2 dt_1 \right| \\ \leq \ & N^k \int_{j/N}^{(j+1)/N} \int_{t_1}^{t_1+1/N} \cdots \int_{t_{k-1}}^{t_{k-1}+1/N} |g^{(k)}(j/N) - g^{(k)}(t_k)| dt_k \ldots dt_2 dt_1 \\ \leq \ & \epsilon, \end{aligned}$$

if N is large enough, because $|t_k - j/N| < k/N$ will be arbitrarily small as $N \to \infty$ (k remains fixed). This implies that $g^{(k)}(j/N) - N^k \Delta^k h_N(j) \to 0$ uniformly in $j \in \{1, 2, \ldots, N - k\}$, as $N \to \infty$. Define now

$$\begin{aligned} \Sigma_N^{(k)} &= \frac{1}{N} \sum_{j=1}^{N-k} \frac{[j]_k}{N^k} \frac{(N-j)_k}{N^k} (N^k \Delta^k h_N(j))^2, \\ \widetilde{\Sigma}_N^{(k)} &= \frac{1}{N} \sum_{j=1}^{N-k} \frac{[j]_k}{N^k} \frac{(N-j)_k}{N^k} (g^{(k)}(j/N))^2. \end{aligned}$$

The above argument shows that $\Sigma_N^{(k)} - \widetilde{\Sigma}_N^{(k)} \to 0$. Obviously, from the theory of Riemann integration, $\widetilde{\Sigma}_N^{(k)} \to \int_0^1 t^k (1-t)^k (g^{(k)}(t))^2 dt$ and $N^{-1} \sum_{j=1}^{N} h_N(j) \to \int_0^1 g(t)dt$. Hence, $\Sigma_N^{(k)} \to \int_0^1 t^k (1-t)^k (g^{(k)}(t))^2 dt$ for all $k = 0, 1, \ldots, n$, and the continuous bound follows as the limiting case of the discrete one. For the general case (where $g, g', \ldots, g^{(n)}$ are merely assumed to be continuous in the *open* interval $(0, 1)$), it suffices to apply the previous continuous bound to the function $g_\epsilon(t) = g(\epsilon + (1 - 2\epsilon)t)$, $t \in [0, 1]$, and its derivatives. The general form of the inequality now follows by monotone convergence theorem, since

$$(t-\epsilon)^k (1-\epsilon-t)^k (g^{(k)}(t))^2 I_{[\epsilon,1-\epsilon]}(t) \nearrow t^k (1-t)^k (g^{(k)}(t))^2 I_{(0,1)}(t), \text{ as } \epsilon \searrow 0.$$

(b) Although the continuous inequality arises as a limiting case of the discrete one, it seems that no simple way exists to get the discrete form of the inequality from the continuous one. We shall try to provide a convincing explanation of this fact. First observe that the discrete inequality of Corollary 3.1(a) is equivalent to the variance bound (see also Theorem 4.1 and Table 1)

$$(-1)^n[S_n - \mathrm{Var}\, g(U_N)] \geq 0,$$

$$\text{where } \; S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k!(k+1)!} \frac{1}{N} \sum_{j=1}^{N-k} (\Delta^k g(j))^2 [j]_k (N-j)_k.$$

This inequality holds for all $n = 0, 1, \ldots$, and for any $g : \{1, 2, \ldots, N\} \longrightarrow \mathbb{R}$, where U_N is the discrete uniform r.v. on $\{1, 2, \ldots, N\}$. On the other hand, one can consider a large class of variance bounds, as in (1.5), for a class of continuous r.v.'s that are close to U_N. Clearly, for the bound (1.5) to be applicable, the function g should be extended to a function $\widetilde{g}$ defined at least in the interval $[1, N]$, in such a way that $\widetilde{g}', \widetilde{g}'', \ldots, \widetilde{g}^{(n+1)}$ exist in $(1, N)$, and the graph of $\widetilde{g}$ is passing through the points $\{(j, g(j)), j = 1, 2, \ldots, N\}$, i.e., $\widetilde{g}(1) = g(1), \ldots, \widetilde{g}(N) = g(N)$. (In the sequel we write g instead of $\widetilde{g}$.) It seems that the simplest way to approximate U_N by a continuous r.v. is to form the convolution $X_\epsilon = U_N + \epsilon V$, where V is a uniform r.v. in $(-1, 1)$ and $\epsilon > 0$ is small. Letting $\epsilon \searrow 0$, we see that $\mathrm{Var}\, g(X_\epsilon) \to \mathrm{Var}\, g(U_N)$, since g is assumed to be (absolutely) continuous in $[1, N]$. If $S_n^\epsilon(g)$ is the bound (1.5) for $\mathrm{Var}\, g(X_\epsilon)$, then its limit, $\widetilde{S}_n(g) = \lim_{\epsilon \searrow 0} S_n^\epsilon(g)$, provides a "differential" bound for $\mathrm{Var}\, g(U_N)$, and a natural question is if this procedure can obtain the discrete variance bound of Corollary 3.1(a), for a suitable choice of g. Surprisingly enough, the answer is negative. To see this, observe that the variance bound $S_n^\epsilon(g)$ depends on the functions $a_k^\epsilon(t)$, $t \in (1-\epsilon, N+\epsilon)$ corresponding to the density $f_\epsilon(t) = (2N\epsilon)^{-1} \sum_{j=1}^N I_{(j-\epsilon, j+\epsilon)}(t)$ of X_ϵ (see (1.6)). It is easy to see that $a_k^\epsilon(t) \to a_k(t)$ as $\epsilon \searrow 0$, where

$$a_k(t) = \frac{1}{N^2} \sum_{j_1 \leq [t]} \sum_{j_2 > [t]} (j_2 - j_1)(t - j_1)^k (j_2 - t)^k, \quad 1 < t < N.$$

Thus, taking limits as $\epsilon \searrow 0$ in the inequality $(-1)^n[S_n^\epsilon(g) - \mathrm{Var}\, g(X_\epsilon)] \geq 0$, and using dominated convergence, we get the bound

$$(-1)^n[\widetilde{S}_n(g) - \mathrm{Var}\, g(U_N)] \geq 0, \quad \text{where}$$

$$\widetilde{S}_n(g) = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{(k-1)!k!} \int_1^N a_{k-1}(t)(g^{(k)}(t))^2 dt,$$

and where $a_k(t)$ is as above. [Note that under mild moment conditions on the absolutely continuous r.v. V involved in the convolution $U_N + \epsilon V$, the form of $a_k(t)$ does not depend on the distribution of V, so that the specific choice of a uniform V does not limit the applicability of the procedure.] After some algebra, it follows that for $j < t < j+1$ ($j = 1, 2, \ldots, N-1$), $a_0(t) = j(N-j)/(2N)$, while

$$a_1(t) = \frac{j(N-j)}{2N}(t-j)(j+1-t) + \frac{[j]_2(N-j)_2}{6N}(t-j) + \frac{(j)_2[N-j]_2}{6N}(j+1-t).$$

The above calculations suffice for obtaining the first two bounds $\widetilde{S}_0(g)$ and $\widetilde{S}_1(g)$. The first one is an upper bound and the second one a lower bound for $\mathrm{Var}\, g(U_N)$. Regarding the first discrete and continuous bounds, S_0 and $\widetilde{S}_0(g)$, resp., it can be seen that $\mathrm{Var}\, g(U_N) \le S_0 \le \widetilde{S}_0(g)$, and that $\inf_g \widetilde{S}_0(g) = S_0$, where the infimum is taken over all absolutely continuous g passing through the points. (The infimum in this case is minimum, and it is attained by the broken line passing through the points.) Thus, S_0 can be obtained from $\widetilde{S}_0(g)$, for a suitable choice of g. However, this is not the case for the second discrete bound S_1, which cannot be recovered from its continuous counterpart $\widetilde{S}_1(g)$. Indeed, we have

$$S_1 - \widetilde{S}_1(g) = \frac{1}{2N}\sum_{j=1}^{N-1} j(N-j)\theta_j(g) + \frac{1}{12N}\sum_{j=1}^{N-2} [j]_2(N-j)_2\phi_j(g),$$

where

$$\begin{aligned}\theta_j(g) &= \left(\int_j^{j+1} g'(t)dt\right)^2 - \int_j^{j+1} (g'(t))^2 dt \\ &\quad + \frac{1}{2}\int_j^{j+1} (t-j)(j+1-t)(g''(t))^2 dt, \\ \phi_j(g) &= \int_j^{j+1}\int_{t_1}^{t_1+1} (g''(t_2))^2 dt_2 dt_1 - \left(\int_j^{j+1}\int_{t_1}^{t_1+1} g''(t_2)dt_2 dt_1\right)^2.\end{aligned}$$

Observe that $\theta_j(g) \ge 0$ (as follows from the inequality (1.1)) and, obviously, $\phi_j(g) \ge 0$, showing that $\widetilde{S}_1(g) \le S_1 \le \mathrm{Var}\, g(U_N)$ for all g passing through the given points. Moreover, if $N \ge 4$, it can be shown that $\phi_j(g) + \phi_{j+1}(g) \ge [\Delta^2 g(j+1) - \Delta^2 g(j)]^2/4$ for all $j = 1, \ldots, N-3$ (we omit the tedious details), implying that

$$\sup_g \widetilde{S}_1(g) \le S_1 - \frac{(N-1)(N-2)}{48N}\sum_{j=1}^{N-3} [\Delta^2 g(j+1) - \Delta^2 g(j)]^2,$$

where the supremum is taken over all g passing through the given points and having an absolutely continuous derivative g' in $(1,N)$. Thus, S_1 cannot be recovered by $\widetilde{S}_1(g)$, unless the $N \ge 4$ points belong to a polynomial curve of degree at most 2. The same is true for all bounds $\widetilde{S}_n(g)$, i.e., $\inf_g (-1)^n[\widetilde{S}_n(g) - S_n] \ge 0$, and in the non-trivial case, where $n \le N-3$ (otherwise S_n is equal to $\mathrm{Var}\, g(U_N)$) and $n \ge 1$, the equality is attained if and only if the N points $\{(j, g(j)),\ j = 1, 2, \ldots, N\}$ lie on a polynomial curve of degree at most $n+1$.

The above analysis shows that the discrete inequality is strictly better than the continuous one. The interested reader can find the details of the present Remark in Afendras, Papadatos and Papathanasiou (2007).

The main result of the present section is stated in the following theorem, which is the discrete analogue of Papathanasiou's (1988) main result (cf. (1.5)–(1.7)). Here we use the terminology "integer interval" in order to denote any subset J of integers with the property: if $j_1 \in J$ and $j_2 \in J$ then all integers j between j_1 and j_2 belong to J.

THEOREM 3.1. *Let X be an integer-valued r.v. with probability mass function* (p.m.f.) *$p(x)$ and finite moment of order $2n+2$, and assume that $g(x)$ is an arbitrary function defined in the smallest integer interval $J = J(X)$, that contains the support of X. Then, for $k = 0, 1, \ldots, n$, the nonnegative functions defined by*

$$a_k(j) = \sum_{x \le j} \sum_{y > j} (y-x) p(x) p(y) (j+k-x)_k (y-j-1)_k, \quad j \in J, \tag{3.2}$$

can be rewritten as

$$a_k(j) = (-1)^k [\mu_{k+1}(j) s_k(j+1) - \mu_k(j+1) s_{k+1}(j)], \quad j \in J, \tag{3.3}$$

where

$$\mu_k(j) = \mathbb{E}(X-j)_k, \quad s_k(j) = \sum_{x<j} p(x)(x-j)_k = \mathbb{E}[(X-j)_k I(X<j)]. \tag{3.4}$$

If $g(X)$ has finite second moment and, furthermore, the function g satisfies the conditions

$$\sum_{j \in J} a_{k-1}(j) (g^{(k)}(j))^2 < \infty, \quad k = 1, 2, \ldots, n+1, \tag{3.5}$$

then

$$(-1)^n[S_n - \operatorname{Var} g(X)] \geq 0, \quad \textit{where} \;\; S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{(k-1)!k!} \sum_{j \in J} a_{k-1}(j)(g^{(k)}(j))^2, \tag{3.6}$$

and the equality in (3.6) *holds if and only if* $g(x)$, $x \in J$, *is the restriction of a polynomial of degree at most* $n+1$.

PROOF. First note that the finiteness of the $(k+1)$-th moment is sufficient for the finiteness of $a_k(j)$ for $j \in J$. Using $(j+k-x)_k = (-1)^k(x-j-1)_k$ and $y - x = (y-j) - (x-j)$, we get

$$\begin{aligned}(y-x)(j+k-x)_k(y-j-1)_k &= (-1)^k(x-j-1)_k(y-j)_{k+1} \\ &\quad +(-1)^{k+1}(x-j)_{k+1}(y-j-1)_k,\end{aligned}$$

so that

$$\begin{aligned} a_k(j) &= (-1)^k \sum_{x<j+1} p(x)(x-(j+1))_k \sum_{y \geq j} p(y)(y-j)_{k+1} \\ &\quad +(-1)^{k+1} \sum_{x<j} p(x)(x-j)_{k+1} \sum_{y \geq j+k+1} p(y)(y-(j+1))_k \\ &= (-1)^k s_k(j+1) \sum_{y \geq j} p(y)(y-j)_{k+1} \\ &\quad +(-1)^{k+1} s_{k+1}(j) \sum_{y \geq j+k+1} p(y)(y-(j+1))_k \\ &= (-1)^k[s_k(j+1)(\mu_{k+1}(j)-s_{k+1}(j))-s_{k+1}(j)(\mu_k(j+1)-s_k(j+1))], \end{aligned}$$

and thus, (3.3) follows. If Y is an independent copy of X, we can write (cf. Cacoullos and Papathanasiou (1985))

$$\operatorname{Var} g(X) = \frac{1}{2}\mathbb{E}(g(Y)-g(X))^2 = \sum_{x<y} p(x)p(y)(g(y)-g(x))^2,$$

and the upper/lower bound S_n in (3.6) is immediately obtained if we estimate $\{g(y)-g(x)\}^2$ with the sum in the RHS of (3.1). Moreover, $S_n = \operatorname{Var} g(X)$ if and only if

$$\sum_{x<y} p(x)p(y)R_{n+1}(x,y) = 0,$$

which implies that $g^{(n+1)}$ is constant, i.e., $g(j)$, $j \in J$, is the restriction of a polynomial $g(x)$ of degree at most $n+1$. This completes the proof. □

It should be noted that $a_0(j) = \sum_{x\le j}(\mu - x)p(x)$, where $\mu = \mathbb{E}X$, and Theorem 3.1 with $n = 0$ yields the upper bound of Cacoullos and Papathanasiou (1985) for the discrete case; see also Cacoullos (1982) and Klaassen (1985). Also, using (3.3) and the identity $(x-j)_{k+1} = (x-j-1)_{k+1}+(k+1)(x-j-1)_k$, it follows that $s_{k+1}(j) = s_{k+1}(j + 1) + (k + 1)s_k(j + 1)$ and similarly for $\mu_{k+1}(j)$, yielding

$$a_k(j) = (-1)^k[\mu_{k+1}(j + 1)s_k(j + 1) - \mu_k(j + 1)s_{k+1}(j + 1)]. \qquad (3.7)$$

4 Variance Bounds for the Discrete Pearson System

The functions $a_k(j)$, given in (3.2), (3.3) or (3.7), are too involved to be useful. Johnson (1993) simplified considerably the expression for $a_k(t)$ for the continuous case (given by (1.6), (1.7)), and obtained the general form (1.9) of the bound (1.5), which is satisfied by the most commonly used continuous r.v.'s. The same simplification for the discrete case is the purpose of the present section. To this end, we have to impose on the p.m.f. p of the integer-valued r.v. X the condition

$$\sum_{x\le j}(\mu - x)p(x) = q(j)p(j), \quad j \in J, \qquad (4.1)$$

with $\mu = \mathbb{E}X$, J the smallest integer interval containing the support of X, and $q(j) = \delta j^2 + \beta j + \gamma$ an arbitrary quadratic. By Korwar's (1991) characterization, relation (4.1) entails that X belongs to the Pearson system of discrete distributions. Independently of Korwar's result, however, the following lemma is needed for our applications.

Lemma 4.1. *Let X be an integer-valued r.v. with p.m.f. $p(x)$ and finite mean μ, and let J be the smallest integer interval containing the support of X. Assume that there exists a quadratic $q(j) = \delta j^2 + \beta j + \gamma$, such that* (4.1) *holds. Then, the support of X equals J and, moreover, if $\delta \le 0$ then X has finite moments of any order. Furthermore, in the case where J is infinite and $\delta > 0$, X has finite moments of any order $a \in [1, 1 + 1/\delta)$, while $\mathbb{E}|X|^{1+1/\delta} = \infty$.*

Proof. Obviously, the function defined in the LHS of (4.1) is unimodal for $j \in J$. Moreover, it is strictly positive for $j \in J$, except if J has a finite upper endpoint, say N, in which case it has a zero at $j = N$. This shows that the support of X is an integer interval, and thus, it coincides with J.

Assume first that $\delta \neq 0$. If $\delta < 0$ then J is finite, because $q(j)$ has to be nonnegative for all $j \in J$. If $\delta > 0$, we provide a detailed proof only for the case where $J = \mathbb{Z}$, since the other two cases $J = \{\ldots, N-1, N\}$ and $J = \{m, m+1, \ldots\}$ can be treated similarly. It can be easily seen that $J = \mathbb{Z}$ implies that $q(j) > 0$ for all $j \in \mathbb{Z}$, $\mu < \min_{j\in\mathbb{Z}}\{j + q(j)\}$, and that the r.v. $Y = -X$ has quadratic q_Y given by $q_Y(j) = q(-j) - \mu - j$, $j \in \mathbb{Z}$. Define the sequences $a_n = p(n) > 0$ and $b_n = p(-n) > 0$, $n = 0, 1, \ldots$, with $a_0 = b_0 = p(0)$. Then, (4.1) obtains the recurrence

$$a_{n+1} = \frac{q(n)}{q(n+1)+n+1-\mu}\, a_n, \quad b_{n+1} = \frac{q(-n)-n-\mu}{q(-n-1)}\, b_n, \quad n = 0, 1, \ldots.$$

It is easy to see that for any fixed $a \geq 1$,

$$\frac{(n+1)^a a_{n+1}}{n^a a_n} = \frac{(n+1)^a q(n)}{n^a(q(n+1)+n+1-\mu)} = \frac{n^2 + (a+\beta/\delta)n + O(1)}{n^2 + (2+(\beta+1)/\delta)n + O(1)},$$

as $n \to \infty$, and the Gauss-Bertrand-Raabe summation criterion yields

$$\sum_{n=0}^{\infty} n^a a_n < \infty \ \text{ if and only if } \ \left(2 + \frac{\beta+1}{\delta}\right) - \left(a + \frac{\beta}{\delta}\right) > 1;$$

that is, if and only if $a < 1 + 1/\delta$. Similar arguments, applied to b_n, yield

$$\sum_{n=0}^{\infty} n^a b_n < \infty \ \text{ if and only if } \ \left(2 - \frac{\beta}{\delta}\right) - \left(a - \frac{\beta+1}{\delta}\right) > 1,$$

completing the proof in the case where $J = \mathbb{Z}$.

Assume next that $\delta = \beta = 0$ and, in order to avoid trivialities, suppose that J contains at least two integers. It then follows that $\gamma > 0$ and that J is bounded below; otherwise, (4.1) leads to the contradiction that $p(j-1)/p(j) = 1 + (j-\mu)/\gamma$ is negative for sufficiently small j. Thus, J is of the form $\{m, m+1, \ldots\}$, and (4.1) implies that for all large enough j, $p(j)/p(j-1) = \gamma/(\gamma + j - \mu)$, so that $\lim_{j\to+\infty} j^a p(j)/((j-1)^a p(j-1)) = 0$ for any fixed $a \geq 1$. Therefore, X has finite moments of any order.

Assume next that $\delta = 0$ and $\beta > 0$. In this case, it is obvious that $q(j) \geq 0$ implies that $j \geq -\gamma/\beta$, so that J is bounded below. Since J cannot be bounded above (otherwise, the linear polynomial q should have 2 real roots), it is again immediate by (4.1) that $\lim_{j\to+\infty} j^a p(j)/((j-1)^a p(j-1)) = \beta/(\beta+1) < 1$, so that X has finite moments of any order.

Finally, in the remaining case $\delta = 0$, $\beta < 0$, it is again obvious that J is bounded above by $-\gamma/\beta$, and also that if J is infinite then $\lim_{j\to-\infty} |j-1|^a p(j-1)/(|j|^a p(j)) = 1+1/\beta < 1$ (this incidentally implies that $-1 \le \beta < 0$ in this case) so that X has finite moments of any order. This completes the proof of lemma. □

It can be seen that there exist discrete r.v.'s satisfying (4.1) with $\delta > 0$ and finite support J with cardinality $|J| \ge 3$; however, the inequality $\delta < (2(|J|-2))^{-1}$ should be necessarily satisfied in this case.

Using a special case of the discrete covariance identity given in Cacoullos and Papathanasiou (1989), namely,

$$\mathrm{Cov}(X, h(X)) = \mathbb{E}[q(X)\Delta h(X)], \tag{4.2}$$

for all functions h for which $\mathbb{E}[q(X)|\Delta h(X)|] < \infty$, we can prove the following lemma.

Lemma 4.2. *Let $k \ge 1$. If X satisfies* (4.1) *and has $k+1$ finite moments, then, under the notation* (3.4), *we have that for all $j \in J$,*

$$\mu_{k+1}(j+1) = b_k(j)\mu_k(j+1) + d_k(j)\mu_{k-1}(j+1) \tag{4.3}$$

and

$$s_{k+1}(j+1) = b_k(j)s_k(j+1) + d_k(j)s_{k-1}(j+1), \tag{4.4}$$

where

$$b_k(j) = \frac{\mu - j - k - 1 + k((2j+2k+1)\delta + \beta)}{1-k\delta}, \quad d_k(j) = \frac{k}{1-k\delta}\, q(k+j), \tag{4.5}$$

provided that $\delta \ne 1/k$.

Proof. Write $q(x)$ as

$$q(x) = \delta(x-j-k)(x-j-k-1) + ((2j+2k+1)\delta+\beta)(x-j-k) + q(j+k) \tag{4.6}$$

and observe that

$$\begin{aligned}
\mu_{k+1}(j+1) &= \mathbb{E}[(X-j-k-1)(X-j-1)_k] \\
&= \mathbb{E}[((X-\mu)+(\mu-j-k-1))(X-j-1)_k] \\
&= \mathrm{Cov}[X,(X-j-1)_k] + (\mu-j-k-1)\mu_k(j+1) \\
&= k\,\mathbb{E}[q(X)(X-j-1)_{k-1}] + (\mu-j-k-1)\mu_k(j+1),
\end{aligned}$$

where we applied (4.2) to the function $h(x) = (x - j - 1)_k$. With the help of (4.6), the expectation above is

$$\mathbb{E}[q(X)(X-j-1)_{k-1}] = \delta\mu_{k+1}(j+1) + ((2j+2k+1)\delta+\beta)\mu_k(j+1) + q(j+k)\mu_{k-1}(j+1),$$

and (4.3) follows. Since $s_{k+1}(j+1) = \mathbb{E}[(X-j-1)_{k+1}I(X < j+1)]$, we can apply the same procedure to $s_{k+1}(j+1)$. Now, identity (4.2) will operate on the function $h(x) = (x-j-1)_k I(x < j+1)$. Taking into account the formula $\Delta[h_1(x)h_2(x)] = \Delta[h_1(x)]h_2(x+1) + h_1(x)\Delta[h_2(x)]$, we have

$$\begin{aligned}\Delta[h(x)] &= \Delta[(x-j-1)_k]I(x+1 < j+1) + (x-j-1)_k\Delta[I(x < j+1)] \\ &= k(x-j-1)_{k-1}I(x < j) - (x-j-1)_k I(x = j) \\ &= k(x-j-1)_{k-1}I(x < j+1).\end{aligned}$$

[The last equality holds because $-(x-j-1)_k I(x=j) = k(x-j-1)_{k-1}I(x = j) = (-1)^{k+1}k!I(x=j)$.] Thus, using exactly the same steps, (4.4) follows. □

COROLLARY 4.1. *Assume that X has $k+1$ finite moments and satisfies* (4.1) *with $\delta \notin \{1, 1/2, \ldots, 1/k\}$. Then*

$$a_k(j) = \frac{k!}{\prod_{s=0}^{k}(1-s\delta)}\, q^{[k+1]}(j)p(j), \quad j \in J, \tag{4.7}$$

where $q^{[n]}(j) = q(j)q(j+1)\cdots q(j+n-1)$.

PROOF. For $k = 0$ the assertion is (4.1) itself, while for $k \geq 1$, (3.7), (4.4) and (4.5) yield the recurrence $a_k(j) = d_k(j)a_{k-1}(j)$, with $d_k(j)$ given by (4.5). Thus, (4.7) is proved. □

Our main result follows immediately from Theorem 3.1 and Corollary 4.1.

THEOREM 4.1. *Let X be an integer-valued r.v. with finite mean μ, support J and p.m.f. $p(x)$ satisfying* (4.1). *Moreover, fix $n \in \{0, 1, \ldots\}$, assume that X has $2n+2$ finite moments and, if $n \geq 1$, suppose further that $\delta \notin \{1, 1/2, \ldots, 1/n\}$. Then, for any function $g : J \longrightarrow \mathbb{R}$ satisfying*

$$\mathbb{E}[q^{[k]}(X)(g^{(k)}(X))^2] < \infty, \quad k = 0, 1, \ldots, n+1,$$

where $q^{[0]}(X) = 1$*, the following inequality holds:*

$$(-1)^n[S_n - \mathrm{Var}\, g(X)] \ge 0, \tag{4.8}$$

$$\textit{where} \quad S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k! \prod_{j=0}^{k-1}(1 - j\delta)} \mathbb{E}[q^{[k]}(X)(g^{(k)}(X))^2].$$

Equality in (4.8) *holds if and only if* $g(x)$, $x \in J$, *is the restriction of a polynomial* $g : \mathbb{R} \longrightarrow \mathbb{R}$, *of degree at most* $n + 1$.

Variance bounds for the most commonly used discrete Pearson variates are given in Table 1.

Observe that if X is Poisson(λ), then the upper/lower variance bound

$$S_n = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k!} \lambda^k \mathbb{E}(g^{(k)}(X))^2$$

is very similar to the bound (1.8) for the Normal (cf. Houdré and Pérez-Abreu, 1995). It should also be mentioned that since the Negative Binomial is an example of a discrete infinite divisible r.v., similar inequalities in this case can be obtained by an application of Corollary 2 in Houdré et al. (1998).

5 Estimation of logθ in Geometric Distributions

Let $X_1, X_2, \ldots, X_\nu$ be a random sample of size ν from Geometric(θ), so that

$$I\!P(X_i = j) = \theta(1-\theta)^j, \quad j = 0, 1, \ldots, \quad i = 1, 2, \ldots, \nu.$$

Let $X = X_1 + X_2 + \cdots + X_\nu$ be the complete sufficient statistic, and define

$$U(X_1) = \begin{cases} 0, & \text{if } X_1 = 0, \\ 1 + \frac{1}{2} + \cdots + \frac{1}{X_1}, & \text{if } X_1 \in \{1, 2, \ldots\}. \end{cases}$$

It is clear that $0 \le U(X_1) \le X_1$, showing that U has finite moments of any order. Observing that $\Delta U(j) = 1/(1 + j)$, $j = 0, 1, \ldots$, and applying the identity

$$\mathbb{E}[U(X_1)] - U(0) = \sum_{j=0}^{\infty} (\Delta U(j)) I\!P(X_1 > j), \tag{5.1}$$

Table 1. Specific form of $\alpha_k(X)$, needed for computation of $\sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k!} \mathbb{E}[\alpha_k(X)(g^{(k)}(X))^2]$, for some discrete Pearson variates.

name parameters	p.m.f. $p(j)$ support	quadratic $q(j)$ δ, β, γ	$\alpha_k(X) = \dfrac{q^{[k]}(X)}{\prod_{j=0}^{k-1}(1-j\delta)}$
Poisson(λ) $\lambda > 0$	$e^{-\lambda}\dfrac{\lambda^j}{j!}$ $j = 0, 1, \ldots$	λ $0, 0, \lambda$	λ^k
Binomial(N, p) $0 < p < 1$, $N = 1, 2, \ldots$	$\binom{N}{j} p^j (1-p)^{N-j}$ $j = 0, 1, \ldots, N$	$(N-j)p$ $0, -p, Np$	$p^k (N-X)_k$ (exact if $n \ge N-1$)
Negative Binomial(r, p) $0 < p < 1$, $r > 0$	$\binom{r+j-1}{j} p^r (1-p)^j$ $j = 0, 1, \ldots$	$\dfrac{1-p}{p}(r+j)$ $0, \frac{1-p}{p}, \frac{r(1-p)}{p}$	$\left(\dfrac{1-p}{p}\right)^k [r+X]_k$
Hyper-geometric($r, s; N$) $r \ge 1$, $s \ge 1$, $N = 1, 2, \ldots$, $N \le \min\{r, s\}$	$\binom{N}{j}\dfrac{(r)_j (s)_{N-j}}{(r+s)_N}$ $j = 0, 1, \ldots, N$	$\dfrac{(r-j)(N-j)}{r+s}$ $\frac{1}{r+s}, -\frac{r+N}{r+s}, \frac{Nr}{r+s}$	$\dfrac{(r-X)_k (N-X)_k}{(r+s)_k}$ (exact if $n \ge N-1$)
Negative Hyper-geometric($r, s; N$) $r > 0$, $s > 0$, $N = 1, 2, \ldots$,	$\binom{N}{j}\dfrac{(-r)_j (-s)_{N-j}}{(-r-s)_N}$ $j = 0, 1, \ldots, N$	$\dfrac{(r+j)(N-j)}{r+s}$ $-\frac{1}{r+s}, \frac{N-r}{r+s}, \frac{Nr}{r+s}$	$\dfrac{[r+X]_k (N-X)_k}{[r+s]_k}$ (exact if $n \ge N-1$)

one finds that $\mathbb{E}(U) = -\log\theta$. According to the Rao-Blackwell/Lehmann-Scheffé Theorem, the uniformly minimum variance unbiased (UMVU) estimator of $-\log\theta$ is given by $T_\nu = \mathbb{E}(U|X)$, and the following Lemma shows that it has an explicit form.

Lemma 5.1. *The UMVU estimator of* $-\log\theta$ *is given by*

$$T_\nu = T_\nu(X) = \sum_{j=0}^{X-1} \frac{1}{\nu+j}, \tag{5.2}$$

where $X = X_1 + X_2 + \cdots + X_\nu$, *and where* T_ν *is assumed to be* 0 *if* $X = 0$.

Proof. For $k \in \{0, 1, \ldots\}$ and $x \in \mathbb{R}$ we write $\begin{bmatrix} x \\ k \end{bmatrix} = \frac{[x]_k}{k!}$. Since X follows the Negative Binomial(ν, θ) distribution, we have that for any fixed $k \in \{0, 1, \ldots\}$,

$$I\!P(X_1 = j | X = k) = \frac{\begin{bmatrix} \nu-1 \\ k-j \end{bmatrix}}{\begin{bmatrix} \nu \\ k \end{bmatrix}}, \quad j = 0, 1, \ldots, k,$$

and thus,

$$I\!P(X_1 > j | X = k) = \begin{cases} \dfrac{\begin{bmatrix} \nu \\ k-j-1 \end{bmatrix}}{\begin{bmatrix} \nu \\ k \end{bmatrix}}, & j = 0, 1, \ldots, k-1, \\ 0, & j = k, k+1, \ldots . \end{cases}$$

Therefore, $T_\nu(0) = 0$. Also, using (5.1) we find that for any $k \geq 1$,

$$\begin{aligned} T_\nu(k) &= \mathbb{E}(U(X_1)|X = k) \\ &= \sum_{j=0}^{\infty} (\Delta U(j)) I\!P(X_1 > j | X = k) \\ &= \frac{\sum_{j=0}^{k-1} \frac{1}{k-j} \begin{bmatrix} \nu \\ j \end{bmatrix}}{\begin{bmatrix} \nu \\ k \end{bmatrix}}. \end{aligned} \tag{5.3}$$

Clearly, $\Delta T_\nu(0) = \frac{1}{\nu}$. Also, for $k \ge 1$, we have

$$\begin{aligned}
\Delta T_\nu(k) &= \frac{(k+1)!}{[\nu]_{k+1}} \sum_{j=0}^{k} \frac{[\nu]_j}{(k+1-j)j!} - \frac{k!}{[\nu]_k} \sum_{j=0}^{k-1} \frac{[\nu]_j}{(k-j)j!} \\
&= \frac{k!}{[\nu]_{k+1}} \Bigg\{ 1 + (k+1) \sum_{j=0}^{k-1} \frac{[\nu]_{j+1}}{(k-j)(j+1)!} \\
&\qquad - \sum_{j=0}^{k-1} ((\nu+j)+(k-j)) \frac{[\nu]_j}{(k-j)j!} \Bigg\} \\
&= \frac{k!}{[\nu]_{k+1}} \left\{ 1 + \sum_{j=0}^{k-1} ((k+1)-(j+1)) \frac{[\nu]_{j+1}}{(k-j)(j+1)!} - \sum_{j=0}^{k-1} \frac{[\nu]_j}{j!} \right\} \\
&= \frac{k!}{[\nu]_{k+1}} \left\{ 1 + \sum_{j=1}^{k} \frac{[\nu]_j}{j!} - \sum_{j=0}^{k-1} \frac{[\nu]_j}{j!} \right\} = \frac{k!}{[\nu]_{k+1}} \frac{[\nu]_k}{k!} = \frac{1}{\nu+k}.
\end{aligned}$$

Hence,

$$T_\nu(k) = T_\nu(k) - T_\nu(0) = \sum_{j=0}^{k-1} \Delta T_\nu(j) = \sum_{j=0}^{k-1} \frac{1}{\nu+j},$$

completing the proof. □

The variance of T_ν is, clearly, quite complicated. Also, it can be shown that it does not attain the Cramér-Rao bound,

$$\frac{(\frac{\partial}{\partial\theta}(-\log\theta))^2}{\nu I_{X_1}(\theta)} = \frac{1-\theta}{\nu},$$

where $I_{X_1}(\theta) = \theta^{-2}(1-\theta)^{-1}$ is the Fisher information of X_1. However, since X belongs to the Pearson family, the bounds of Theorem 4.1, namely (see Table 1)

$$S_n(\nu,\theta) = \sum_{k=1}^{n+1} \frac{(-1)^{k-1}}{k!} \left(\frac{1-\theta}{\theta}\right)^k \mathbb{E}[[X+\nu]_k (T_\nu^{(k)}(X))^2], \quad n = 0,1,\ldots \tag{5.4}$$

may provide some useful information for $\mathrm{Var} T_\nu$, due to the simple form of the forward differences:

$$T_\nu^{(k)}(X) = \frac{(-1)^{k-1}(k-1)!}{[\nu+X]_k}, \quad k = 1,2,\ldots\ .$$

Indeed, in order to calculate the bound $S_n(\nu,\theta)$ in (5.4), it suffices to obtain explicit forms for the expectations

$$\begin{aligned}\mathbb{E}\left[\frac{1}{[X+\nu]_k}\right] &= \theta^{\nu}\sum_{j=0}^{\infty}\begin{bmatrix}\nu\\ j\end{bmatrix}\frac{(1-\theta)^j}{[\nu+j]_k}\\ &= \frac{\theta^{\nu}}{x^{\nu+k-1}}\sum_{j=0}^{\infty}\begin{bmatrix}\nu\\ j\end{bmatrix}\frac{x^{\nu+j+k-1}}{[\nu+j]_k}, \quad k=1,2,\ldots,\end{aligned}$$

where $x=1-\theta$. Using the identity

$$\frac{x^{\nu+j+k-1}}{[\nu+j]_k}=\frac{1}{(k-1)!}\int_0^x (x-y)^{k-1}y^{\nu+j-1}dy, \quad k=1,2,\ldots,$$

and interchanging the order of summation and integration, one finds that

$$\mathbb{E}\left[\frac{1}{[X+\nu]_k}\right]=\frac{\theta^{\nu+k}}{(k-1)!(1-\theta)^{\nu+k}}\int_{\theta}^{1} t^{-\nu}(1-t)^{\nu-1}\left(\frac{t}{\theta}-1\right)^{k-1}dt. \quad (5.5)$$

Therefore, expanding $(1-t)^{\nu-1}$ and $(t/\theta-1)^{k-1}$ according to binomial formula, it is possible to express $S_n(\nu,\theta)$ in terms of finite sums. Specifically, we have the following.

LEMMA 5.2. *Let $S_n(\nu,\theta)$ be the variance bound in* (5.4)*, where T_ν is the UMVU estimator of $-\log\theta$, given in* (5.2)*. Then, with the notation $\mu=\mu(\theta)=(1-\theta)/\theta$, the following expressions are valid:*

(i) $S_n(\nu,\theta)=(-1/\mu)^{\nu-1}\sum_{k=0}^{n}\frac{\beta_k(\nu,\theta)}{k+1}$,

where

$$\begin{aligned}\beta_k(\nu,\theta) &= (-\log\theta)\sum_{j=0}^{k}\binom{\nu-1}{j}\binom{k}{j}\theta^{-j}\\ &\quad+\sum_{j=0}^{k}(-1)^j\binom{k}{j}\sum_{s=0,\ s\neq j}^{\nu-1}(-1)^s\binom{\nu-1}{s}\frac{\theta^{-j}-\theta^{-s}}{j-s}\ .\end{aligned}$$

In particular,

$$\beta_0(\nu,\theta)=(-\log\theta)+\sum_{j=1}^{\nu-1}\frac{(-\mu)^j}{j}\ .$$

(ii) $S_n(\nu,\theta) = (1/\mu)^{\nu-1} \sum_{k=0}^{n} \frac{(-1)^k}{k+1} \int_\theta^1 t^{-\nu}(1-t)^{\nu-1} \left(\frac{t}{\theta}-1\right)^k dt.$

(iii) $S_n(\nu,\theta) = \sum_{k=0}^{n} \nu^{-k-1} \frac{(-1)^k}{k+1} \mu^{k+1} \int_0^\nu y^k \left(1-\frac{y}{\nu}\right)^{\nu-1} \left(1+\frac{\mu y}{\nu}\right)^{-\nu} dy.$

PROOF. (i) and (ii) follow from (5.5) using straightforward manipulations. For $k=0$, the form of $\beta_0(\nu,\theta)$ follows from the identity

$$\sum_{j=1}^{\nu-1} \frac{(-1)^j}{j} \binom{\nu-1}{j} (\theta^{-j}-1) = \sum_{j=1}^{\nu-1} \frac{(-1)^j}{j} \left(\frac{1-\theta}{\theta}\right)^j, \quad 0<\theta<1,$$

which can be verified on taking forward differences with respect to ν. Also, (iii) follows from (ii), using the substitution $y = \nu(t-\theta)/(1-\theta)$ in the corresponding integrals. □

The form of $\beta_k(\nu,\theta)$ is quite complicated. However, with the help of Lemma 5.2 it is possible to investigate the asymptotic behaviour of the bounds $S_n(\nu,\theta)$, as n or ν becomes large.

LEMMA 5.3.

(i) *For any fixed sample size* ν, $\lim_n S_n(\nu,\theta) = \mathrm{Var}(T_\nu)$ *if and only if* $\theta \ge 1/2$.

(ii) *For large* ν *and for any fixed* n,

$$S_n(\nu,\theta) \approx \sum_{k=0}^{n} \frac{(-1)^k}{k+1} k! \left(\frac{1-\theta}{\nu}\right)^{k+1},$$

in the sense that there exist functions $\gamma_k(\nu,\theta)$, $k=0,1,\ldots$, *such that*

$$S_n(\nu,\theta) = \sum_{k=0}^{n} \frac{(-1)^k}{k+1} \frac{\gamma_k(\nu,\theta)}{\nu^{k+1}},$$

where

$$\lim_{\nu\to\infty} \gamma_k(\nu,\theta) = k!(1-\theta)^{k+1}, \quad k=0,1,\ldots .$$

In particular, $\lim_\nu[\nu S_n(\nu,\theta)] = 1-\theta$ *for all* $n=0,1,\ldots$ *and all* $\theta \in (0,1)$, *and thus,* $\lim_\nu[\nu \mathrm{Var} T_\nu] = 1-\theta$.

Proof. (i) By Lemma 5.3(ii),

$$S_n(\nu,\theta) = \sum_{k=0}^{n} (-1)^k x_k,$$

$$\text{where } \ x_k = x_k(\nu,\theta) = \frac{(1/\mu)^{\nu-1}}{k+1} \int_\theta^1 t^{-\nu}(1-t)^{\nu-1} \left(\frac{t}{\theta}-1\right)^k dt.$$

If $\theta \ge 1/2$, then the sequence x_k is positive, decreasing in k, and tends to 0 as $k \to \infty$, because $0 < t/\theta - 1 < 1$ for all $t \in (\theta, 1)$. Therefore, by Leibnitz criterion, $S_n(\nu,\theta) \to s(\nu,\theta) \in \mathbb{R}$, as $n \to \infty$, and the inequalities $S_{2n+1} \le \mathrm{Var}T_\nu \le S_{2n}$ show that $s(\nu,\theta) = \mathrm{Var}T_\nu$. Also, when $\theta < 1/2$ it is easy to see that there exists a constant $c(\nu,\theta) > 0$ such that for all k, $x_k(\nu,\theta) \ge c(\nu,\theta)(2\theta)^{-k}/(k+1) \to \infty$, as $k \to \infty$, so that $\lim_n S_n(\nu,\theta)$ does not exist.

(ii) By Lemma 5.3(iii) it suffices to show that

$$\lim_{\nu\to\infty} \int_0^\nu y^k \left(1-\frac{y}{\nu}\right)^{\nu-1} \left(1+\frac{\mu y}{\nu}\right)^{-\nu} dy = \frac{k!}{(\mu+1)^{k+1}}, \quad k = 0,1,\ldots,$$

where $\mu = (1-\theta)/\theta > 0$. For $y > 0$ define the functions

$$f_\nu(y) = y^k \left(1-\frac{y}{\nu}\right)^{\nu-1} \left(1+\frac{\mu y}{\nu}\right)^{-\nu} I_{(0,\nu)}(y), \quad f(y) = y^k \exp(-(\mu+1)y),$$

so that $\lim_\nu f_\nu(y) = f(y)$. Using the inequality

$$\left(1-\frac{y}{\nu}\right)^{\nu-1} \left(1+\frac{\mu y}{\nu}\right)^{-\nu} I_{(0,\nu)}(y) \le \exp\left(-\frac{\mu}{\mu+1}y\right), \quad y > 0, \ \nu = 1,2,\ldots,$$

and the facts that

$$\int_0^\infty y^k \exp\left(-\frac{\mu}{\mu+1}y\right) dy < \infty, \quad \int_0^\infty f(y)dy = \frac{k!}{(\mu+1)^{k+1}},$$

the assertion follows from dominated convergence. □

Table 2 gives an idea on how the first few bounds behave, for small and moderate values of ν, and for various θ-values.

Table 2. Numerical values of the first six upper/lower variance bounds $S_n(\nu,\theta)$, $n=0,1,\ldots,5$, given in (5.4), for sample sizes $\nu=2,5,10,30$, and parametric values $\theta=0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$ (Best recorded bounds are indicated in bold).

		Lower bounds			Upper bounds		
ν	θ	$S_1(\nu,\theta)$	$S_3(\nu,\theta)$	$S_5(\nu,\theta)$	$S_0(\nu,\theta)$	$S_2(\nu,\theta)$	$S_4(\nu,\theta)$
2	0.1	**0.337022**	-1.06166	-32.0592	**0.744157**	1.04612	7.29043
2	0.3	0.366037	**0.373802**	0.354041	0.484012	**0.436289**	0.443401
2	0.5	0.267132	0.273879	**0.274889**	0.306853	0.278553	**0.276214**
2	0.7	0.157179	0.158325	**0.158368**	0.167758	0.158583	**0.158382**
2	0.9	0.0508305	0.0508623	**0.0508624**	0.0517554	0.0508640	**0.0508624**
5	0.1	0.189473	**0.191325**	0.177515	0.218115	**0.201580**	0.205491
5	0.3	0.147183	0.149112	**0.149319**	0.160880	0.150395	**0.149834**
5	0.5	0.103976	0.104595	**0.104627**	0.109814	0.104765	**0.104644**
5	0.7	0.0614806	0.0615923	**0.0615938**	0.0632906	0.0616055	**0.0615941**
5	0.9	0.0201665	0.0201699	**0.0201699**	0.0203432	0.0201700	**0.0201699**
10	0.1	0.0934586	0.0940524	**0.0940911**	0.0987841	0.0942976	**0.0941646**
10	0.3	0.0721594	0.0724140	**0.0724241**	0.0750804	0.0724719	**0.0724296**
10	0.5	0.0511213	0.0512043	**0.0512056**	0.0524877	0.0512149	**0.0512060**
10	0.7	0.0304074	0.0304234	**0.0304235**	0.0308622	0.0304244	**0.0304235**
10	0.9	0.0100454	0.0100460	**0.0100460**	0.0100925	0.0100460	**0.0100460**
30	0.1	0.0304330	0.0304528	**0.0304529**	0.0309245	0.0304544	**0.0304529**
30	0.3	0.0235959	0.0236048	**0.0236049**	0.0238849	0.0236053	**0.0236049**
30	0.5	0.0168009	0.0168040	**0.0168040**	0.0169443	0.0168041	**0.0168040**
30	0.7	0.0100484	0.0100490	**0.0100490**	0.0100986	0.0100490	**0.0100490**
30	0.9	0.00333871	0.00333873	**0.00333873**	0.00334415	0.00333873	**0.00333873**

Acknowledgements. This work was partially supported by the University of Athens' research fund under Grant 70/4/5637. We would like to thank Professor Arup Bose, Editor, for careful reading of the paper and the preceding Technical Report. We also thank an anonymous Co-Editor and a referee for a number of helpful suggestions and comments which led to the results of Section 5 and Remark 3.1. Thanks are also due to Professor Roger W. Johnson for providing us with a copy of his paper.

References

Afendras, G., Papadatos, N. and Papathanasiou, V. (2007). A Remark on the paper: The discrete Mohr and Noll inequality with applications to variance bounds. Technical Report TR\11\2007, Dept. of Math. and Stat., University of Cyprus (`http://www.mas.ucy.ac.cy/english/technical2007/TR_11_2007.pdf`).

Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1992). *A First Course in Order Statistics.* Wiley, New York.

Balakrishnan, N., Charalambides, C. and Papadatos, N. (2003). Bounds on expectation of order statistics from a finite population. *J. Statist. Plann. Inference* **113**, 569–588.

Cacoullos, T. (1982). On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.* **10**, 799–809.

Cacoullos, T. and Papathanasiou, V. (1985). On upper bounds for the variance of functions of random variables. *Statist. Probab. Lett.* **3**, 175–184.

Cacoullos, T. and Papathanasiou, V. (1989). Characterizations of distributions by variance bounds. *Statist. Probab. Lett.* **7**, 351–356.

Chang, W. and Richards, D.S.P. (1999). Variance inequalities for functions of multivariate random variables. In *Advances in Stochastic Inequalities* (Atlanta, GA, 1997), Theodore P. Hill and Christian Houdré, eds., Contemporary Mathematics **234**, American Mathematical Society, Providence, 43–67.

Chen, Louis H.Y. (1988). The central limit theorem and Poincaré-type inequalities. *Ann. Probab.* **16** 300–304.

Chen, Louis H.Y. and Lou, J.H. (1987). Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. H. Poincaré Probab. Statist.* **23**, 91–110.

Chernoff, H. (1981). A note on an inequality involving the normal distribution. *Ann. Probab.* **9**, 533–535.

Houdré, C. (1995). Some applications of covariance identities and inequalities to functions of multivariate normal variables. *J. Amer. Statist. Assoc.* **90**, 965–968.

Houdré, C. (1997). The iterated jackknife estimate of variance. *Statist. Probab. Lett.* **35**, 197–201.

Houdré, C. and Kagan, A. (1995). Variance inequalities for functions of Gaussian variables. *J. Theoret. Probab.* **8**, 23–30.

Houdré, C. and Pérez-Abreu, V. (1995). Covariance identities and inequalities for functionals on Wiener and Poisson spaces. *Ann. Probab.* **23**, 400–419.

Houdré, C., Pérez-Abreu, V. and Surgailis, D. (1998). Interpolation, correlation identities and inequalities for infinitely divisible variables. *J. Fourier Anal. Appl.* **4**, 651–668.

Johnson, R.W. (1993). A note on variance bounds for a function of a Pearson variate. *Statist. Decisions* **11**, 273–278.

Klaassen, C.A.J. (1985). On an inequality of Chernoff. *Ann. Probab.* **3**, 966–974.

Kochar, S.C. and Korwar, R.M. (1997). Stochastic properties of order statistics from finite populations. *Sankhyā Ser. A* **59**, 102–116.

Kochar, S.C. and Korwar, R.M. (2001). On random sampling without replacement from a finite population. *Ann. Inst. Statist. Math.* **53**, 631–646.

Korwar, R.M. (1991). On characterizations of distributions by mean absolute deviation and variance bounds. *Ann. Inst. Statist. Math.* **43**, 287–295.

López-Blázquez, F. and Castaño-Martínez, A. (2006). Upper and lower bounds for the correlation ratio of order statistics from a sample without replacement. *J. Statist. Plann. Inference*, **136**, 43–52.

Mohr, E. and Noll, W. (1952). Eine Bemerkung zur Schwarzschen Ungleichheit (in German). *Math. Nachr.* **7**, 55–59.

Papadatos, N. and Rychlik, T. (2004). Bounds on expectations of L-statistics from without replacement samples. *J. Statist. Plann. Inference*, **124**, 317–336.

PAPATHANASIOU, V. (1988). Variance bounds by a generalization of the Cauchy-Schwarz inequality. *Statist. Probab. Lett.* **7**, 29–33.

SIEGEL, A.F. (1993). A surprising covariance involving the minimum of multivariate normal variables. *J. Amer. Statist. Assoc.*, **88**, 77–80.

TAKAHASI, K. and FUTATSUYA, M. (1998). Dependence between order statistics in samples from finite population and its application to ranked set sampling. *Ann. Inst. Statist. Math.*, **50**, 49–70.

G. AFENDRAS, N. PAPADATOS AND V. PAPATHANASIOU
DEPARTMENT OF MATHEMATICS
SECTION OF STATISTICS AND OPERATIONS RESEARCH
UNIVERSITY OF ATHENS
PANEPISTEMIOPOLIS, 157 84 ATHENS, GREECE
E-mail: g_afendras@math.uoa.gr
npapadat@math.uoa.gr
bpapath@math.uoa.gr

Paper received September 2005; revised December 2006.