Unbiased estimates for products of moments and cumulants for finite populations

Kit Withers (Kit.Withers@gmail.com)

formerly Industrial Research, Lower Hutt & DSIR

2015

Abstract

Let $F = F_N$ be the distribution of a finite real population of size N. Let $\hat{F} = F_n$ be the empirical distribution of a sample of size n drawn from the population without replacement. We prove the following remarkable **inversion principle** for obtaining unbiased estimates. Let $T(F_N)$ be any product of the moments or cumulants of F_N.

$$\text{Set } T_{nN}(F_N) = E\ T(F_n). \text{ Then } E\ T_{Nn}(F_n) = T(F_N).$$

Introduction and Summary

Consider a random sample of size n without replacement, say $X_1, ..., X_n$, from a finite real population $x_1, ..., x_N$ with mean $\mu = m_1 = N^{-1} \sum_{i=1}^N x_i$, rth moment, central moment and cumulant $m_r = N^{-1} \sum_{i=1}^N x_i^r$, $\mu_r = \mu_r$ and κ_r. We obtain unbiased estimates (**UEs**) for products of them

$$m(1^{r_1} 2^{r_2} \ldots) = m_1^{r_1} m_2^{r_2} \ldots,$$
$$mu(1^{r_1} 2^{r_2} ...) = \mu^{r_1} \mu_2^{r_2} \mu_3^{r_3} \ldots,$$
$$\kappa(1^{r_1} 2^{r_2} \ldots) = \kappa_1^{r_1} \kappa_2^{r_2} \ldots$$

and for products of the joint central moments

$$\mu(1^{r_1}2^{r_2}...) = E(\hat{\mu} - \mu)^{r_1}(\hat{\mu}_2 - E\ \hat{\mu}_2)^{r_2}...$$

and for products of the corresponding joint cumulants $\kappa(1^{r_1}2^{r_2}...)$, for **weight** ≤ 6 where the weight is

$$r(1^{r_1}2^{r_2}...) = 1.r_1 + 2.r_2 + ... \text{ that is } r(\pi) = a + b + \cdots$$

for $\pi = (a, b, \cdots)$ a partition of r.
We assume that all partitions π are put into ascending order $(1^{r_1}2^{r_2}...)$. For example we write $(1^2 2)$ or (112) rather than (21^2) or (121).

Our UEs are given in terms of

$$\hat{m}(1^{r_1}2^{r_2}...) = \hat{m}_1^{r_1}\hat{m}_2^{r_2}\cdots, \text{ and } \hat{m}u(1^{r_1}2^{r_2}\cdots) = \hat{\mu}^{r_1}\hat{\mu}_2^{r_2}\cdots \text{ where}$$

$$\hat{m}_1 = \overline{X} = n^{-1}\sum_{i=1}^{n} X_i,\ \hat{m}_r = n^{-1}\sum_{i=1}^{n} X_i^r,\ \hat{\mu}_r = \mu_r(\hat{F}) = n^{-1}\sum_{i=1}^{n}(X_i - \overline{X})^r.$$

For π a partition of r, 2 gives $E\ \hat{m}(\pi)$ and an UE of $m(\pi)$ for $r \leq 6$. 3 gives $E\ \hat{m}u(\pi)$ and an UE of $mu(\pi)$ for $r \leq 6$.

We discover a remarkable **inversion principle**. A result of this is that these UEs do not require having to invert any matrices or solve any sets of linear equations. Set

$$m_{(r)} = \{m(\pi) : \ \pi \text{ a partition of } r\},$$
$$mu_{(r)} = \{mu(\pi) : \pi \text{ a partition of } r\},$$
$$\text{and } K_{(r)} = \{K(\pi) : \ \pi \text{ a partition of } r\}.$$

These are vectors of length n_r, the number of partitions of r, tabled on p307 of Comtet (1974). We shall use the ordering $(1), (2.1^2), (3.12.1^3), (4.13.2^2.1^2 2.1^4)$, $(5.14.23.1^2 3.12^2.1^3 2.1^5)$ and $(6.15.24.3^2.1^2 4.123.2^3.1^3 3.1^2 2^2.1^4 2.1^6)$. For example $m_{(3)} = (m_2,\ m_1 m_2,\ m_1^3)$.

In 2 we derive a matrix $B_r = B_r(N, n)$ from Skellam (1949) such that

$$E\ \hat{m}_{(r)} = B_r m_{(r)} \text{ so } B_r(N, n)^{-1}\hat{m}_{(r)} \text{ is an UE of } m_{(r)}.$$

In 3 we derive a matrix $C_r = C_r(N, n)$ from Sukhatme (1944) such that

$$E\ \hat{mu}_{(r)} = C_r mu_{(r)}, \text{ so } C_r(N, n)^{-1}\hat{mu}_{(r)} \text{ is an UE of } mu_{(r)}.$$

So expressing cumulants in terms of moments as

$$K_{(r)} = G_r m_{(r)} \text{ where } G_r \text{ is a matrix of constants,}$$

it follows that

$$E\ \hat{K}_{(r)} = D_r(N, n)K_{(r)} \text{ where } D_r(N, n) = G_r C_r(N, n)G_r^{-1}.$$

The inversion principle (proved in 7) states the following beautiful result.

$$B_r(N, n)^{-1} = B_r(n, N),\ C_r(N, n)^{-1} = C_r(n, N)),\ D_r(N, n)^{-1} = D_r(n, N).$$

This implies that for $T(F)$ a product of moments or cumulants,

$$\text{if } T_{nN}(F) = E\ T(\hat{F}) \text{ then } E\ T_{Nn}(\hat{F}) = T(F).$$

This can be extended to general smooth functionals.
4 gives equivalent multivariate results to those of 3.
For partitions $\pi_1, \pi_2, \cdots$ set

$$\mu(\pi_1, \pi_2, \cdots) = \mu(\pi_1)\mu(\pi_2)\cdots$$

and

$$\kappa(\pi_1, \pi_2, \cdots) = \kappa(\pi_1)\kappa(\pi_2)\cdots$$

5 gives $E\ \hat{\mu}(\pi_1, \pi_2, \cdots)$ and an UE of $\mu(\pi_1, \pi_2, \cdots)$ up to total order 6, in particular UEs for $\mu(1^r)$. MAPLE was used to simplify the UE of $\mu(1^r)$ given by Dwyer and Tracy (1980) for $r \leq 5$ and to confirm they agreed with our results. Earlier, Nath (1968,1969) gave $\mu(1^r)$ for $r = 3, 4$ and UEs for them.

6 gives $E\ \hat{\kappa}(\pi_1, \pi_2, \cdots)$ and an UE of $\kappa(\pi_1, \pi_2, \cdots)$ up to total order 6.

7 proves the inversion principle and a multivariate version: in this case the dimension $n_r \times n_r$ in (1.7) jumps to $N_r \times N_r$ where

$$N_r = \sum_{\pi}^{r} P(\pi), \text{ is the partition function,}$$

$$P(1^{r_1} 2^{r_2} \cdots) = r!/\Pi_{i=1}(i!^{r_i} r_i!) \text{ for } r = \sum_{i=1} i r_i.$$

So N_r is the sum of the partitions of r. Pierce (1940) gives transformations that allow one to obtain noncentral moments up to order 6 for finite populations from those for infinite populations. He also covers the case where the sample values $X_1, \cdots, X_n$ may have different moments.

We shall use π for a partition of r, π_- for a partition of r excluding 1's, say $\pi_- = (2^{r_2}3^{r_3}\cdots)$ with $2r_2 + 3r_3 + \cdots = r$, and π_+ for a partition of r including at least one 1. Let n_r be the number of partitions of r, and N_r the sum of the partitions of r. We also partition vectors and matrices using subscript $+$ and $-$. For example $mu_{-(r)} = \{mu(\pi_-)\}$, $mu_{+(r)} = \{mu(\pi_+)\}$, and

$B_r = B_r(N,n) = \begin{pmatrix} B_{--r} & 0 \\ B_{+-r} & B_{++r} \end{pmatrix}$ since $B_{-+r} = 0$.

Since also $C_{-+r} = D_{-+r} = 0$, the inversion principle (1.8) implies

$$B_{--r}(N,n)^{-1} = B_{--r}(n,N),$$

$$C_{--r}(N,n)^{-1} = C_{--r}(n,N),$$

$$D_{--r}(N,n)^{-1} = D_{--r}(n,N),$$

so that $B_{--r}(n,N)\hat{m}_{-(r)}$, $C_{--r}(n,N)\hat{mu}_{-(r)}$, $D_{--r}(n,N)\hat{K}_{-(r)}$ are UEs of $m_{-(r)}$, $mu_{-(r)}$, $K_{-(r)}$.

The number of parts in π is denoted by $q(\pi)$:
$q(1^{r_1}2^{r_2}\cdots) = r_1 + r_2 + \cdots$. Set
$(n)_i = n(n-1)\cdots(n-i+1) = n!/(n-i)!$.

Products of non-central moments

Here we derive the result

$$E\ \hat{m}_{(r)} = B_r m_{(r)}. \tag{1}$$

Skellam (1949) showed that for

$$S_r = \sum_{i=1}^{n} X_i^r = n\hat{m}_r \text{ and } s_r = \sum_{i=1}^{N} x_i^r = Nm_r,$$

$$E\ S_{a_1}\cdots S_{a_s} = \sum_{\pi}^{s} \lambda_\pi \sum^{P(\pi)} s_{R_{\pi_1}}\ s_{R_{\pi_2}} \cdots \text{ where} \tag{2}$$

$$R_{\pi_1} = a_1 + \cdots + a_{\pi_1}, R_{\pi_2} = a_{\pi_1+1} + \cdots + a_{\pi_1+\pi_2}, \cdots, \tag{3}$$

$\sum^{P(\pi)}$ sums over all $P(\pi)$ such terms and λ_π is the **Carver function**, a function of (N, n) defined as follows:

$$\lambda_\pi = \Psi_{\pi_1}(e)\Psi_{\pi_2}(e)\ldots \text{ evaluated at } e^j = e_j = (n)_j/(N)_j.$$

where $\Psi_i = \Psi_i(u)$ are the polynomials

$$\Psi_1 = u, \Psi_2 = u - u^2, \Psi_3 = u - 3u^2 + 2u^3, \Psi_4 = u - 7u^2 + 12u^3 - 6u^4,$$
$$\Psi_5 = u - 15u^2 + 50u^3 - 60u^4 + 24u^5,$$
$$\Psi_6 = u - 31u^2 + 180u^3 - 390u^4 + 360u^5 - 120u^6, \tag{4}$$

and so on. (He actually replaces $m(\pi)$ by a more complicated function.) For example

$$E\ S_{a_1}S_{a_2}S_{a_3} = \lambda_3 s_{a_1+a_2+a_3} + \lambda_1 2\sum^3 s_{a_1+a_2}s_{a_3} + \lambda_{1^3}s_{a_1}s_{a_2}s_{a_3},$$
$$\text{where } \lambda_3 = e_1 - 3e_2 + 2e_3, \lambda_{12} = e_2 - e_3, \lambda_{1^3} = e_3,$$

and

$$
\begin{aligned}
&E\ S_{a_1}\cdots S_{a_5} = \lambda(5)s_{a_1+\ldots+a_5} + \lambda(14)\sum^{3} s_{a_1+\ldots+a_4}s_{a_5} \\
&+\lambda_{23}\sum^{10} s_{a_1+a_2+a_3}s_{a_4+a_5} + \lambda_{1^2 3}\sum^{10} s_{a_1+a_2+a_3}s_{a_4}s_{a_5} \\
&+\lambda_{12^2}\sum^{15} s_{a_1+a_2}s_{a_3+a_4}s_{a_5} + \lambda_{1^3 2}\sum^{10} s_{a_1+a_2}s_{a_3}s_{a_4}s_{a_5} \\
&+\lambda_{1^5}s_{a_1}\cdots s_{a_5},
\end{aligned}
$$

where $\lambda_5 = e_1 - 15e_2 + 50e_3 - 60e_4 + 24e_5,\ \lambda_{14} = e_2 - 7e_3 + 12e_4 - 6e_5,$
$\lambda_{23} = e_2 - 4e_3 + 5e_4 - 2e_5,\ \lambda_{1^2 3} = e_3 - 3e_4 + 2e_5,$
$\lambda_{12^2} = e_3 - 2e_4 + e_5,\ \lambda_{1^3 2} = e_4 - e_5,\ \lambda_{1^5} = e_5.$

He wrote (2.1) out in full for $s \leq 4$. We extend this to $s \leq 6$ in Appendix E.

$$\text{Set } S_{(r)} = \{S_{\pi_1} S_{\pi_2} \cdots : (\pi_1 \pi_2 \cdots) \text{ a partition of } r\},$$

and similarly for $s_{(r)}$. It follows from (2.1) that

$$E\ S_{(r)} = A_r s_{(r)} \text{ where } A_r = \{A_{\pi.\pi'} : \pi, \pi' \text{ partitions of } r\} \tag{5}$$

and $A_1, \cdots, A_6$ are as follows.

$$A_1 = \lambda_1, A_2 = \begin{pmatrix} \lambda_1 & \\ \lambda_2 & \lambda_{11} \end{pmatrix}, A_3 = \begin{pmatrix} \lambda_1 & & \\ \lambda_2 & \lambda_{11} & \\ \lambda_3 & 3\lambda_{12} & \lambda_{1^3} \end{pmatrix}.$$

Since $S_r = n\hat{m}_r$ and $s_r = Nm_r$, (2.5) implies (2.1) with

$$B_r = D_{rn}^{-1} A_r D_{rN} \text{ where } D_{rN} = diag\{N^{q(\pi)} : r(\pi) = r\}.$$

For example

$$D_{4N} = diag(n, n^2, n^2, n^3, n^4).$$

$$\text{Writing } A_r(N,n) = A_r = \{A_{\pi.\pi'}(N,n) : \pi, \pi' \text{ partition of } r\}$$
$$\text{and } B_r(N,n) = B_r = \{B_{\pi.\pi'}(N,n) : \pi, \pi' \text{ partitions of } r\}$$

it follows from (2.1) that

$$E\ \hat{m}_{\pi_1}\hat{m}_{\pi_2}\cdots = \sum_{\pi'}^{r} B_{\pi.\pi'}(N,n) m_{\pi_1'} m_{\pi_2'} \cdots = b_\pi(N,n,m) \text{ say.}$$

By the inversion principle

$$B_r(N,n)^{-1} = B_r(n,N) \tag{6}$$

$$\text{so } b_\pi(n,N,\hat{m}) \text{ is an UE of } m_{\pi_1} m_{\pi_2} \cdots . \tag{7}$$

For example

$$E\ \hat{m}_1^3 = n^{-3}(\lambda_3 N m_3 + 3\lambda_{21} N^2 m_2 m_1 + \lambda_{1^3} N^3 m_1^3)$$

so $N^{-3}(\overline{\lambda}_3 n \hat{m}_3 + 3\overline{\lambda}_{21} n^2 \hat{m}_2 \hat{m}_1 + \overline{\lambda}_{1^3} n^3 \hat{m}_1^3)$ is an UE of m_1^3

where $\overline{\lambda}_\pi$ is λ_π with n and N reversed;
that is, with $e_j = (n)_j/(N)_j$ replaced by $1/e_j = (N)_j/(n)_j$.

The reason for the inversion principle

Recall the symmetric functions

$$[\pi_1\pi_2\cdots]_n = \sum_n{}' X_{i_1}^{\pi_1} X_{i_2}^{\pi_2}\cdots, \quad [\pi_1\pi_2\cdots]_N = \sum_N{}' x_{i_1}^{\pi_1} x_{i_2}^{\pi_2}\cdots$$

where $\sum_n{}'$ sums over $i_1, i_2, \cdots$ distinct in $1, \cdots, n$.

The standardised function $\langle\pi_1\cdots\pi_j\rangle_n = [\pi_1\cdots\pi_j]_n/(n)_j$ satisfies the invariance principle

$$E\langle\pi_1\cdots\pi_j\rangle_n = \langle\pi_1\cdots\pi_j\rangle_N.$$

So for every partition π of r, $E\langle\pi\rangle_n = \langle\pi\rangle_N$.

Examples

The UE of $\mu(1^2) = var(\overline{X})$ is $(1-n/N)\hat{\mu}_2/(n-1)$.
The UE of $\mu(1^3) = mu_3(\overline{X})$ is
$(1-n/N)(1-2n/N)\hat{\mu}_3/(n-1)(n-2)$.
The UE of $\mu(2^2) = var(\hat{\mu}_2)$ is $\hat{\mu}_2^2$.
These quickly get complicated. Set $f_j = \lambda_j/n^j$. Then
$\mu_2(\hat{\mu}_2) = C_{2^2.4}\mu_4 + (C_{2^2.2^2} - C_{2.2}^2)\mu_2^2$ for
$C_{2.2} = N(f_1 - f_2) = (1-n^{-1})/(1-N^{-1}),\ C_{2^2.4} =$
$N(f_2 - 2f_3 + f_4),\ C_{2^2.2^2} = N^2(\lambda_{11}/n^2 - 2\lambda_{12}/n^3 + 3\lambda_{22}/n^4)$.

References

Carver, H. C. (1930). Fundamentals of the theory of sampling. *Ann. Math. Stat.*, **1**, 101-121.

Carver, H. C. (1930). Fundamentals of the theory of sampling. *Ann. Math. Stat.*, **1**, 260-274.

Dwyer, P. S. and Tracy, D. S. (1980). Expectation and estimation of product moments in sampling from a finite population. *J. Amer. Statist. Ass.*, **75**, 431 - 437.

Dwyer, P. S. (1938). On combined expansions of products of symmetric power sums and of sums of symmetric power products with applications to sampling (continued). *Ann. Math. Stat.*, **9**, 97-132.

Fisher, R.A. (1929). Moments and product moments of sampling distributions. *Proc.London Math. Soc.* Series 2, **30**, 199-238.

James, G.S. (1958). On moments and cumulants of systems of statistics. *Sankhya*, **20**, 1-30.

Nath, S. N. (1968). On product moments from a finite universe. *J. Amer. Statist. Ass.*, **63**, 535-541.
Pierce, J.A. (1940). A study of a universe of n finite populations with application to moment-function adjustments for grouped data. *Annals Math. Statist.*, **11**, 311-334.
Raghunandanan, K. and Srinivasan, R. (1973). Some product moments useful in sampling theory. *J. Amer. Stat. Assoc.* **68**, 409-413.
Skellam, J. G. (1949). The distribution of moment statistics of samples drawn without replacement from a finite population. *J. Royal Statistical Society*, Ser. B, **11**, 291-296.
Stuart, A. and Ord, J.K. (1987). Kendall's Advanced Theory of Statistics, 1, Griffin, London. 5th Edition.
Sukhatme, P.V. (1944). Moments and product moments of moment-statistics for samples of the finite and infinit populations. *Sankhya*, **6**, 363-382.

Tiit, E. (1988). Unbiased and n^{-k}-biased estimations of entire rational functions of moments. *Statistical and Probabilistic Models, Report 798, Tartu State University, Tartu, Estonia*, 3-17.

Wishart, J. (1952). Moment coefficients of the k-statistics in samples from a finite population. *Biometrika*, **39**, 1-13.

Withers, C.S. and Nadarajah, S. (2008) *Sankhya: The Indian Journal of Statistics, Series A*, **70** (2), 186–222.
http://sankhya.isical.ac.in/search/70a2/08006SB/08006SB.pdf

Withers, C.S. and Nadarajah, S. (2012) Nonparametric estimates of low bias. *REVSTAT Statistical Journal*, **10** (2), 229–283.