A STUDY OF A UNIVERSE OF n FINITE POPULATIONS WITH APPLICATION TO MOMENT-FUNCTION ADJUSTMENTS FOR GROUPED DATA

By Joseph A. Pierce

The object of this paper is to study the case of a universe of n finite populations, considering both the expectations of population moment-functions and the moments of sample moments, and to make applications of the results which may be of interest to mathematical statisticians. The sampling formulas which are derived reduce to the usual infinite or finite sampling formulas, under appropriate assumptions. Also a method is given whereby finite sampling formulas may be transformed into the corresponding infinite sampling formulas.

The general methods and formulas which are given in Part I for the expectations of population moment-functions are used, in Part II, to find the expectations of moments of a distribution of discrete data grouped in "k groupings of k".

I. A Study of a Universe of n Finite Populations

Let ${}_nU_N$ be a universe composed of the set of populations ${}_rX$, $(r = 1, 2, \cdots, n)$ each population ${}_rX$ consisting of a finite number of discrete variates ${}_rx_i$, $(i = 1, 2, \cdots, N)$, $(N > n)$. The tth moment of ${}_rX$ is denoted by ${}_r\mu_t$. The tth central moment of ${}_rX$ is denoted by ${}_r\bar{\mu}_t$. The tth moment and the tth central moment of ${}_nU_N$ are respectively denoted by μ_t and $\bar{\mu}_t$. The expected value of a variable y is denoted by $E(y)$. We have

$$\begin{aligned}
{}_r\mu_t &= E({}_rx_i^t) = \frac{1}{N}\sum_{i=1}^{N} {}_rx_i^t, \qquad {}_r\bar{\mu}_t = E({}_rx_i - {}_r\mu_1)^t = \frac{1}{N}\sum_{i=1}^{N} ({}_rx_i - {}_r\mu_1)^t, \\
\mu_{1:\mu_t} &= E({}_r\mu_t) = \frac{1}{n}\sum_{r=1}^{n} {}_r\mu_t, \qquad \mu_{1:\bar{\mu}_t} = E({}_r\bar{\mu}_t) = \frac{1}{n}\sum_{r=1}^{n} {}_r\bar{\mu}_t, \qquad (1.1) \\
\mu_{s_1 s_2 \cdots s_v : \mu_{t_1}\mu_{t_2}\mu_{t_v}} &= E({}_{r_1}\mu_{t_1}^{s_1}\, {}_{r_2}\mu_{t_2}^{s_2} \cdots {}_{r_v}\mu_{t_v}^{s_v}), \\
\mu_{s_1 s_2 \cdots s_v : \bar{\mu}_{t_1}\bar{\mu}_{t_2}\cdots\bar{\mu}_{t_v}} &= E({}_{r_1}\bar{\mu}_{t_1}^{s_1}\, {}_{r_2}\bar{\mu}_{t_2}^{s_2} \cdots {}_{r_v}\bar{\mu}_{t_v}^{s_v}).
\end{aligned}$$

We also note that $\mu_{s_1 s_2 \cdots s_v : \mu_{t_1}\mu_{t_2}\cdots\mu_{t_v}}$ may be written $\mu_{111\cdots 1 : \mu_{t_1}^{s_1}\mu_{t_2}^{s_2}\cdots\mu_{t_v}^{s_v}}$.

1. The expected value of moments and central moments. It follows easily from (1.1) that

$$\mu_{1:\mu_t} = \mu_t. \qquad (1.2)$$

From the usual formula for central moments in terms of moments, we get

$$\mu_{1.\bar{\mu}_t} = \sum_{i=0}^{t} (-1)^i \binom{t}{i} \mu_{i1:\mu_1\mu_{t-1}}. \tag{1.3}$$

Terms of the form $\mu_{i1:\mu_1\mu_{t-i}}$ may be evaluated by use of the well known formulas [20; p. 58] for changing from moments to central moments in the case of a multivariate distribution. Two of these formulas are given below.

$$\begin{aligned} \bar{\mu}_{11:\mu_a\mu_b} &= \mu_{11:\mu_a\mu_b} - \mu_{10:\mu_a\mu_b}\mu_{01:\mu_a\mu_b}. \\ \bar{\mu}_{111:\mu_a\mu_b\mu_c} &= \mu_{111:\mu_a\mu_b\mu_c} - \mu_{110:\mu_a\mu_b\mu_c}\mu_{001:\mu_a\mu_b\mu_c} \\ &\quad - \mu_{101:\mu_a\mu_b\mu_c}\mu_{010:\mu_a\mu_b\mu_c} - \mu_{011:\mu_a\mu_b\mu_c}\mu_{100:\mu_a\mu_b\mu_c} \\ &\quad + 2\,\mu_{100:\mu_a\mu_b\mu_c}\mu_{010:\mu_a\mu_b\mu_c}\mu_{001:\mu_a\mu_b\mu_c}. \end{aligned} \tag{1.4}$$

We find that

$$\mu_{i1:\mu_1\mu_{t-1}} = \Sigma \frac{i!}{p_1!p_2!r_1!r_2!} \bar{\mu}_{p_1r_1:\mu_1\mu_{t-i}} \mu_{1:\mu_1}^{p_2} \mu_{1:\mu_{t-i}}^{r_2}, \tag{1.5}$$

where p_1p_2 is a two-part partition of i and $r_1 + r_2 = 1$.

Using (1.3) and (1.5), we get

$$\mu_{1:\bar{\mu}_2} = \bar{\mu}_2 - \bar{\mu}_{2:\mu_1}. \tag{1.6}$$

$$\mu_{1:\bar{\mu}_3} = \bar{\mu}_3 - 3\bar{\mu}_{11:\mu_1\mu_2} + 6\mu_1\bar{\mu}_{2:\mu_1} + 2\bar{\mu}_{3:\mu_1}. \tag{1.7}$$

$$\begin{aligned} \mu_{1:\bar{\mu}_4} &= \bar{\mu}_4 + 6(\bar{\mu}_2 - 2\mu_1^2)\,\bar{\mu}_{2:\mu_1} - 12\mu_1\bar{\mu}_{3:\mu_1} + 12\mu_1\bar{\mu}_{11:\mu_1\mu_2} \\ &\quad - 4\bar{\mu}_{11:\mu_1\mu_3} + 6\bar{\mu}_{21:\mu_1\mu_2} - 3\bar{\mu}_{4:\mu_1}. \end{aligned} \tag{1.8}$$

etc.

If the n populations are identical, it is evident from the definition of $\mu_{1:\bar{\mu}_t}$ that, for all finite t,

$$\mu_{1:\bar{\mu}_t} = \bar{\mu}_t.$$

2. The expected value of Thiele seminvariants. If the tth Thiele seminvariant is denoted by λ_t, then

$$\mu_{1:\lambda_t} = \Sigma \frac{(-1)^{\rho-1} t!(\rho-1)!}{s_1!s_2!\cdots s_v!(2!)^{s_2}(3!)^{s_3}\cdots(v!)^{s_v}} \mu_{s_1s_2\cdots s_v:\mu_1\mu_2\cdots\mu_v}, \tag{1.9}$$

the summation being taken for all positive integers $s_i (i = 1, 2, \cdots v)$, for which

$$\rho = \sum_{i=1}^{v} s_i, \qquad t = \sum_{i=1}^{v} is_i.$$

Terms of the form $\mu_{s_1s_2\cdots s_v:\mu_1\mu_2\cdots\mu_v}$ are evaluated by (1.4). We have

$$\mu_{1:\lambda_2} = \lambda_2 - \bar{\mu}_{2:\mu_1}. \tag{1.10}$$

$$(1.11) \qquad \mu_{1:\lambda_3} = \lambda_3 - 3\bar{\mu}_{11:\mu_1\mu_2} + 6\lambda_1\bar{\mu}_{2:\mu_1} + 2\bar{\mu}_{3:\mu_1} .$$

$$(1.12) \qquad \mu_{1:\lambda_4} = \lambda_4 + 12[\lambda_2 - 2\lambda_1^2]\,\bar{\mu}_{2:\mu_1} - 24\lambda_1\bar{\mu}_{3:\mu_1} + 24\lambda_1\bar{\mu}_{11:\mu_1\mu_2}$$
$$- 4\bar{\mu}_{11:\mu_1\mu_3} + 12\bar{\mu}_{21:\mu_1\mu_2} - 3\bar{\mu}_{2:\mu_2} - 6\bar{\mu}_{4:\mu_1} .$$

etc.

If the n populations are identical then, for all finite t,

$$\mu_{1:\lambda_t} = \lambda_t .$$

3. Generalized sampling. It follows from definition that all rational isobaric moment-functions have the property that they may be expressed in terms of power sums and power product sums with certain coefficients. Of the power sums and power product sums which enter a sampling formula only the power product sums take different forms depending on the law of variate selection. Now, there are two possible courses which may be followed by one who wishes to derive sampling formulas for the case of a single population.

1. One may decide in advance on the law which he wishes to govern the selection of variates which enter the sample. Then he may apply this law in the evaluation, in terms of moments, of every power product term as it occurs in each formula which is derived.
2. One may derive the formulas for sampling under the condition that the law is unspecified, thereby obtaining formulas which are capable of being interpreted in terms of laws that are decided upon later.

We illustrate the two possible courses by considering the formula,

$$(1.13) \qquad \bar{\mu}_{2:z} = \frac{r}{s}\,\Sigma\bar{x}^2 + \frac{2r(r-1)}{s(s-1)}\,\Sigma\bar{x}_i\bar{x}_j ,$$

which Carver [12; p. 102] obtains for the case of finite sampling without replacements. Here r = the number in the sample, s = the number in the parent population and z_i = the algegraic sum of the variates of ith sample. Later, by evaluating $\Sigma\bar{x}^2$ and $\Sigma\bar{x}_i\bar{x}_j$ in terms of moments, he finds

$$(1.14) \qquad \bar{\mu}_{2:z} = \frac{r(s-r)}{s-1}\,\bar{\mu}_{2:x} .$$

(It should be noted that Carver [12; p. 115] obtained the corresponding formula for infinite sampling by letting $s \to \infty$).

The preceding development is entirely in accord with the first of the courses stated above. It is also the standard procedure and is the course followed by such writers as Isserles [2], Neyman [6], Church [7], Pepper [11] and Dwyer [20], in deriving finite sampling formulas. Also, it is the course followed by such authors as "Student" [1], Tchouproff [3], Church [5], Craig [9], Fisher [10], and Georgesque [13] for the case of sampling from an infinite population.

However, in (1.13), it is possible to employ the definition,

$$\frac{2}{s(s-1)}\,\Sigma\bar{x}_i\bar{x}_j = \bar{\mu}_{1,1}\,.$$

Then (1.14) becomes

$$\bar{\mu}_{2:z} = r\bar{\mu}_2 + r(r-1)\bar{\mu}_{1,1}\,. \tag{1.15}$$

Formula (1.15) may be interpreted as holding for either finite or infinite sampling, depending on the interpretation which is given to $\bar{\mu}_{1,1}$. It may be easily shown that, if the sampling is from a limited supply, $\bar{\mu}_{1,1} = \frac{-1}{s-1}\,\bar{\mu}_2$ and (1.15) reduces to (1.14). If the sampling is from an infinite supply, $\bar{\mu}_{1,1}$ becomes $\bar{\mu}_1^2$ and therefore

$$\bar{\mu}_{2:z} = r\bar{\mu}_{2:x}\,,$$

which is the formula [12; p. 115] that corresponds, in the infinite case, to (1.14).

Thus, either of the two courses is possible in the case of sampling from a single population. However, if one wishes to get general formulas which hold for both infinite and finite sampling, he should follow the second course. Similarly, in order to obtain generalized sampling formulas where the relations between the variates are unspecified and the populations are assumed to be different, the second course should be followed.

It appears that Tchouproff [3], [4] was the first to approach the sampling problem from such a general point of view. However, his methods of derivation are quite complicated and his results, in general, are difficult to apply to a given problem [5], [8].

Samples of n are formed from ${}_nU_N$ by chosing one variate from each of the n populations. A typical sample is

$${}_1x_{i_1}\,,\ {}_2x_{i_2}\,,\ {}_3x_{i_3}\,,\ \cdots\,,\ {}_rx_{i_r}\,,\ \cdots\,,\ {}_nx_{i_n}\,.$$

We define [4; p. 472]

$$\begin{aligned} \frac{1}{k}\sum_{\substack{i_{r_1},i_{r_2},\cdots i_{r_v}=1\\ r_j\neq r_k}}^{n} {}_{r_1}x_{i_{r_1}}^{t_1}\,{}_{r_2}x_{i_{r_2}}^{t_2}\cdots{}_{r_v}x_{i_{r_v}}^{t_v} &= E({}_{r_1}x_{i_{r_1}}^{t_1}\,{}_{r_2}x_{i_{r_2}}^{t_2}\cdots{}_{r_v}x_{i_{r_v}}^{t_v}) \\ &= {}_{r_1r_2\cdots r_v}\mu_{t_1t_2\cdots t_v}\,, \\ \frac{v}{n^{(v)}}\sum_{r_i\neq r_j}{}_{r_1r_2\cdots r_v}\mu_{t_1t_2\cdots t_v} = \frac{1}{n^{(v)}}\,S_v\,{}_{r_1r_2\cdots r_v}\mu_{t_1t_2\cdots t_v} &= \mu_{t_1t_2\cdots t_v}\,, \end{aligned} \tag{1.16}$$

where k represents the number of possible terms of the given form; S_v means v times the sum for unequal values of $r_1, r_2 \cdots r_v$ and $n^{(v)} = n(n-1)\cdots(n-v+1)$.

4. Moments and product moments of sample moments. The tth moment of the jth sample is denoted by ${}_jm_t$. The sth moment of ${}_jm_t$ for all j is denoted by ${}'\mu_{s:m_t}$ where the prime indicates that the moments of the universe are measured about a fixed point. It follows that

$$(1.17) \qquad {}_jm_t = \frac{1}{n}\sum_{r=1}^{n} {}_rx^t_{i_r} \quad \text{and} \quad {}'\mu_{s:m_t} = E[{}_jm_t]^s.$$

Also, the general product moment, in which the variates of both the sample and the universe are measured about a fixed point, is defined by

$$(1.18) \qquad {}'\mu_{s_1s_2\cdots s_v:m_{t_1}m_{t_2}\cdots m_{t_v}} = E[{}_jm^{s_1}_{t_1}{}_jm^{s_2}_{t_2}\cdots {}_jm^{s_v}_{t_v}].$$

As an illustration of the methods used to derive the formulas of this section, consider a special case of (1.18) when $s_1 = 2$ and $s_i = 0$, $(i = 2, 3, \cdots, v)$. Then

$$\begin{aligned} {}'\mu_{2:m_t} &= \frac{1}{n^2}E\left[\sum_{r=1}^{n} {}_rx^t_{i_r}\right]^2 \\ &= \frac{1}{n^2}E\left[\sum_{r=1}^{n} {}_rx^{2t}_{i_r} + S_2\, {}_{r_1}x^t_{i_{r_1}}\, {}_{r_2}x^t_{i_{r_2}}\right] \\ &= \frac{1}{n^2}\left[\sum_{r=1}^{n} {}_r\mu_{2t} + S_2\, {}_{r_1r_2}\mu_{t,t}\right]. \end{aligned}$$

Therefore, by (1.1), (1.2) and (1.16), we get

$$(1.19) \qquad {}'\mu_{2:m_t} = \frac{1}{n^2}[n\mu_{2t} + n^{(2)}\mu_{t,t}].$$

Using the formulas [20; p. 34] relating products of power sums and power products to expand expressions of the type $E({}_jm^{s_1}_{t_1}{}_jm^{s_2}_{t_2}\cdots {}_jm^{s_v}_{t_v})$, we give, in the tables below, formulas for moments and product moments of sample moments through weight six. The number in a cell and the coefficient, in the same column, at the top of the table should be taken as the coefficient of the moment which is found in the same vertical division. The coefficients in the vertical division are coefficients of the entire right members of the formulas for the respective moments.

Terms of the form $\mu_{t_1t_2\cdots t_v}$, if $t_1 = t_2 = \cdots = t_r = t$, are sometimes written $\mu_{t^r, t_{r+1}\cdots t_v}$.

The numbers in the cells of the tables are identical with the numbers in the cells of the tables given by Dwyer [19; p. 30] for the expected value of partition products.

5. Moments of central moments of samples of n. The tth central moment of the jth sample is denoted by ${}_j\bar{m}_t$. Then,

$$(1.20) \qquad {}_j\bar{m}_t = \frac{1}{n}\sum_{r=1}^{n}({}_rx_{i_r} - {}_jm_1)^t$$

and

$$(1.21) \qquad {}'\mu_{s:\bar{m}_t} = E\left[\frac{1}{n}\sum_{r=1}^{n}({}_rx_{i_r} - {}_jm_1)^t\right]^s.$$

TABLE I

(1)

	Coef.	n
		μ_1
$'\mu_{1:m_1}$	n^{-1}	1

(2)

	Coef.	n	$n^{(2)}$
		μ_2	μ_{1^2}
$'\mu_{1:m_2}$	n^{-1}	1	
$'\mu_{2:m_1}$	n^{-2}	1	1

(3)

	Coef.	n	$n^{(2)}$	$n^{(3)}$
		μ_3	$\mu_{2,1}$	μ_{1^3}
$'\mu_{1:m_3}$	n^{-1}	1		
$'\mu_{11:m_1m_2}$	n^{-2}	1	1	
$'\mu_{3:m_1}$	n^{-3}	1	3	1

(5)

	Coef.	n	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(5)}$
		μ_5	$\mu_{4,1}$	$\mu_{3,2}$	$\mu_{3,1^2}$	$\mu_{2^2,1}$	$\mu_{2,1^3}$	μ_{1^5}
$'\mu_{1:m_5}$	n^{-1}	1						
$'\mu_{11:m_1m_4}$	n^{-2}	1	1					
$'\mu_{11:m_2m_3}$	n^{-2}	1		1				
$'\mu_{21:m_1m_3}$	n^{-3}	1	2	1	1			
$'\mu_{12:m_1m_2}$	n^{-3}	1	1	2		1		
$'\mu_{31:m_1m_2}$	n^{-4}	1	3	4	3	3	1	
$'\mu_{5:m_1}$	n^{-5}	1	5	10	10	15	10	1

(4)

	Coef.	n	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(4)}$
		μ_4	$\mu_{3,1}$	$\mu_{2,2}$	$\mu_{2,1^2}$	μ_{1^4}
$'\mu_{1:m_4}$	n^{-1}	1				
$'\mu_{11:m_1m_3}$	n^{-2}	1	1			
$'\mu_{2:m_2}$	n^{-2}	1		1		
$'\mu_{21:m_1m_2}$	n^{-3}	1	2	1	1	
$'\mu_{4:m_1}$	n^{-4}	1	4	3	6	1

(6)

	Coef.	n	$n^{(2)}$	$n^{(2)}$	$n^{(2)}$	$n^{(3)}$	$n^{(3)}$	$n^{(3)}$	$n^{(4)}$	$n^{(4)}$	$n^{(5)}$	$n^{(6)}$
		μ_6	$\mu_{5,1}$	$\mu_{4,2}$	$\mu_{3,3}$	$\mu_{4,1^2}$	$\mu_{3,2,1}$	μ_{2^3}	$\mu_{3,1^3}$	$\mu_{2^2,1^2}$	$\mu_{2,1^4}$	μ_{1^6}
$'\mu_{1:m_6}$	n^{-1}	1										
$'\mu_{11:m_1m_5}$	n^{-2}	1	1									
$'\mu_{11:m_2m_4}$	n^{-2}	1		1								
$'\mu_{2:m_3}$	n^{-2}	1			1							
$'\mu_{21:m_1m_4}$	n^{-3}	1	2	1		1						
$'\mu_{111:m_1m_2m_3}$	n^{-3}	1	1	1	1		1					
$'\mu_{3:m_2}$	n^{-3}	1		3				1				
$'\mu_{31:m_1m_3}$	n^{-4}	1	3	3	1	3	3		1			
$'\mu_{22:m_1m_2}$	n^{-4}	1	2	3	2	1	4	1		1		
$'\mu_{41:m_1m_2}$	n^{-5}	1	4	7	4	6	16	3	4	6	1	
$'\mu_{6:m_1}$	n^{-6}	1	6	15	10	15	60	15	20	45	15	1

After writing $({}_r x_{i_r} - {}_j m_1)^t$ as the sum of the general term of a binomial series and then expanding the resulting right member of (1.21) as a product of power sums [20; p. 19], we get

$$(1.22)\qquad {}'\mu_{s:\bar{m}_t} = \sum \frac{s!}{r_1!r_2!\cdots r_v!\pi_1!\pi_2!\cdots} \sum_{\substack{i_1,i_2,\cdots i_v=0 \\ i_1\neq i_2\neq\cdots\neq i_v}}^{t} (-1)^{\rho} \binom{t}{i_1}^{r_1}\binom{t}{i_2}^{r_2} \cdots \binom{t}{i_v}^{r_v} {}'\mu_{r_1r_2\cdots r_v\rho:m_t-i_1m_t-i_2\cdots m_t-i_vm_1}$$

where $\sum_{j=1}^{v} r_j = s$, $\sum_{j=1}^{v} i_j r_j = \rho$ and $\pi_1, \pi_2, \cdots$ are the numbers of the repeated parts of s.

The mean of the tth central moment takes the following simple form,

$$(1.23)\qquad {}'\mu_{1:\bar{m}_t} = \sum_{i=0}^{t} (-1)^i \binom{t}{i} {}'\mu_{1i:m_t-im_1},$$

where the moments in the right member of (1.23) through weight six are given in the tables of section four. Also,

$$(1.24)\qquad {}'\mu_{2:\bar{m}_2} = {}'\mu_{2:m_2} - 2'\mu_{21:m_1m_2} + {}'\mu_{4:m_1}.$$

$$(1.25)\qquad {}'\mu_{3:\bar{m}_2} = {}'\mu_{3:m_2} - 3'\mu_{22:m_1m_2} + 3'\mu_{41:m_1m_2} - {}'\mu_{6:m_1}.$$

$$(1.26)\qquad {}'\mu_{2:\bar{m}_3} = {}'\mu_{2:m_3} + 9'\mu_{22:m_1m_2} + 4'\mu_{6:m_1} - 6'\mu_{111:m_1m_2m_3} + 4'\mu_{31:m_1m_3} - 12'\mu_{41:m_1m_2}.$$

After substituting from the tables of section four, (1.23) through (1.26) become

$$(1.27)\qquad {}'\mu_{1:\bar{m}_2} = \frac{n^{(2)}}{n^2}[\mu_2 - \mu_{1,1}].$$

$$(1.28)\qquad {}'\mu_{1:\bar{m}_3} = \frac{n^{(3)}}{n^3}[\mu_3 - 3\mu_{2,1} + 2\mu_{1^3}].$$

$$(1.29)\qquad {}'\mu_{1:\bar{m}_4} = \frac{1}{n^4}[n^{(2)}(n^2 - 3n + 3)(\mu_4 - 4\mu_{3,1}) + 3n^{(2)}(2n - 3)\mu_{2,2} + 3n^{(4)}(2\mu_{2,1^2} - \mu_{1^4})].$$

$$(1.30)\qquad {}'\mu_{1:\bar{m}_5} = \frac{1}{n^5}[n^{(3)}(n^2 - 2n + 2)(\mu_5 - 5\mu_{4,1}) + 10n^{(3)}(n - 2)\mu_{3,2} + 10n^{(3)}(n + 1)(n - 4)\mu_{3,1^2} - 30n^{(3)}(n - 2)\mu_{2^2,1} - 10n^{(4)}(3n - 4)\mu_{2,1^3} + 4n^{(5)}\mu_{1^5}].$$

$$
\begin{aligned}
{}'\mu_{1:\bar{m}_6} = \frac{1}{n^6}[&n^2(n^4 - 5n^3 + 10n^2 - 10n + 5)(\mu_6 - \mu_{5,1}) \\
&+ 15n^{(2)}(n^3 - 4n^2 + 7n - 5)\mu_{4,2} - 10n^{(3)}(2n^2 - 6n + 5)\mu_{3^2} \\
&+ 15n^{(3)}(n^3 - 4n^2 + 6n - 5)\mu_{4,1^2} - 60n^{(3)}(n^2 - 4n + 5)\mu_{3,2,1} \\
&+ 15n^{(3)}(3n - 5)\mu_{2^3} - 20n^{(4)}(n^2 - 3n + 5)\mu_{3,1^3} \\
&+ 45n^{(4)}(2n - 5)\mu_{2^2,1^2} + 15n^{(5)}(n - 5)\mu_{2,1^4} - 5n^{(6)}\mu_{1^6}].
\end{aligned} \tag{1.31}
$$

$$
{}'\mu_{2:\bar{m}_2} = \frac{1}{n^4}[n^{(2)}(n - 1)(\mu_4 - 4\mu_{3,1}) + n^{(3)}(n + 1)\mu_{2,2} - n^{(4)}(2\mu_{2,1^2} - \mu_{1^4})]. \tag{1.32}
$$

$$
\begin{aligned}
{}'\mu_{3:\bar{m}_2} = \frac{1}{n^6}[&n^{(2)}(n - 1)^2(\mu_6 - 6\mu_{5,1}) + 3n^{(2)}(n - 1)(n^2 - 2n + 5)\mu_{4,2} \\
&- 2n^{(2)}(3n^2 - 6n + 5)\mu_{3,3} + n^{(3)}(n^3 - 3n^2 + 9n - 15)\mu_{2^3} \\
&- 3n^{(3)}(n - 1)(n - 5)\mu_{4,1^2} - 12n^{(3)}(n^2 - 4n + 5)\mu_{3,2,1} \\
&+ 4n^{(4)}(3n - 5)\mu_{3,1^3} - 3n^{(4)}(n^2 - 6n + 15)\mu_{2^2,1^2} \\
&+ n^{(6)}(3\mu_{2,1^4} - \mu_{1^6})].
\end{aligned} \tag{1.33}
$$

$$
\begin{aligned}
{}'\mu_{2:\bar{m}_3} = \frac{1}{n^6}[&n^{(2)}(n - 1)^2(n - 2)(\mu_6 - 6\mu_{5,1}) - 3n^{(2)}(n - 2)^2(2n - 5)\mu_{4,2} \\
&+ n^{(2)}(n - 2)^2(n^2 - 2n + 10)\mu_{3,3} \\
&- 6n^{(3)}(n - 2)(n^2 - 6n + 20)\mu_{3,2,1} + 3n^{(3)}(n - 2)(7n - 10)\mu_{4,1^2} \\
&+ 3n^{(3)}(3n^2 - 12n + 20)\mu_{2^3} + 4n^{(4)}(n - 2)(n - 10)\mu_{3,1^3} \\
&+ 9n^{(4)}(n^2 - 8n + 20)\mu_{2^2,1^2} - 4n^{(6)}(3\mu_{2,1^4} - \mu_{1^6})].
\end{aligned} \tag{1.34}
$$

6. The variance of the variance of samples of n. The variance of the variance of samples of n, when the moments of the universe are measured about a fixed point, is defined as

$$
{}'\bar{\mu}_{2:\bar{m}_2} = {}'\mu_{2:\bar{m}_2} - [{}'\mu_{1:\bar{m}_2}]^2. \tag{1.35}
$$

Therefore, from (1.27) and (1.32),

$$
\begin{aligned}
{}'\mu_{2:\bar{m}_2} = \frac{1}{n^4}[&n^{(2)}(n - 1)(\mu_4 - 4\mu_{3,1}) + n^{(3)}(n - 1)\mu_{2,2} - n^{(4)}(2\mu_{2,1^2} - \mu_{1^4})] \\
&- \left(\frac{n - 1}{n}\right)^2 (\mu_2 - \mu_{1,1})^2.
\end{aligned} \tag{1.36}
$$

Tchouproff [4; p. 492] gave a formula (8) for the variance of the sample variance but his result is unwieldy due to the fact that moments of the universe are measured about the mean.

7. Conventional infinite sampling formulas derived from generalized sampling formulas. The term "infinite sampling" is to be interpreted as meaning: *sampling from an unlimited supply* or *sampling from a limited supply with repetitions permitted.* In each of these situations the variates are independent [5; p. 79].

First, it is assured that the n populations are identical, that is, ${}_1X = {}_2X = \cdots = {}_nX$. This assumption results in the fact that, for a fixed t, ${}_1\mu_t = {}_2\mu_t = \cdots = {}_n\mu_t$ and ${}_1\bar{\mu}_t = {}_2\bar{\mu}_t = \cdots = {}_n\bar{\mu}_t$. Therefore, under the assumption of identical populations, every moment may be interpreted as either the moment of n identical populations or as the moment of a single population. The only other assumption is that the sampling is "infinite".

From the condition of independence [3; p. 141], we have

$$E({}_{r_1}x_{i_{r_1}}^{t_1}\,{}_{r_2}x_{i_{r_2}}^{t_2}\cdots{}_{r_v}x_{i_{r_v}}^{t_v}) = (E\,{}_{r_1}x_{i_{r_1}}^{t_1})(E\,{}_{r_2}x_{i_{r_2}}^{t_2})\cdots(E\,{}_{r_v}x_{i_{r_v}}^{t_v}).$$

Therefore,

$${}_{r_1r_2\cdots r_v}\mu_{t_1t_2\cdots t_v} = {}_{r_1}\mu_{t_1}\,{}_{r_2}\mu_{t_2}\cdots{}_{r_v}\mu_{t_v}.$$

Combining the condition of independence with that of identical populations, we have

$$(1.37)\qquad \frac{1}{n^{(v)}}S_v\,{}_{r_1r_2\cdots r_v}\mu_{t_1t_2\cdots t_v} = \frac{1}{n^{(v)}}S_v\,{}_{r_1}\mu_{t_1}\,{}_{r_2}\mu_{t_2}\cdots{}_{r_v}\mu_{t_v} = \mu_{t_1}\,\mu_{t_2}\cdots\mu_{t_v}.$$

By (1.16) and (1.37), we may write

$$(1.38)\qquad \mu_{t_1t_2\cdots t_v} = \mu_{t_1}\mu_{t_2}\cdots\mu_{t_v}.$$

Since the only terms of the generalized sampling formulas [illegible] affected by the assumption of "infinite sampling" are those of the form $\mu_{t_1t_2\cdots t_v}$, the problem of obtaining conventional infinite sampling formulas from generalized sampling formulas is, in practice, a mechanical one. Simply write terms of the form $\mu_{t_1t_2\cdots t_v}$ which appear in a generalized sampling formula, as $\mu_{t_1}\mu_{t_2}\cdots\mu_{t_v}$ and one automatically obtains the corresponding infinite sampling formula.

As an illustration of the method, consider the generalized sampling formula (1.36) for the variance of the sample variance. When (1.38) is utilized to change it into the corresponding infinite sampling formula, (1.36) becomes

$$(1.39)\qquad {}'\bar{\mu}_{2:\bar{m}_2} = \frac{n^{(2)}}{n^4}[(n-1)(\mu_4 - 4\mu_3\mu_1) - (n-3)\mu_2^2 + 2(2n-3)(2\mu_2\mu_1^2 - \mu_1^4)],$$

which is the usual formula [20; p. 75] for the variance of the sample variance when the moments of the universe are measured about a fixed point. If it is assumed that the moments of ${}_nU_N$ are measured about the mean, formula (1.39) becomes

$$(1.40)\qquad \bar{\mu}_{2:\bar{m}_2} = \frac{n^{(2)}}{n^4}[(n-1)\bar{\mu}_4 - (n-3)\bar{\mu}_2^2],$$

which was published by "Student" [1; p. 3] in 1908.

8. Conventional finite sampling formulas derived from generalized sampling formulas. The term "finite sampling" is to be interpreted as meaning: *sampling from a limited supply when repetitions are not permitted.*

In order to reduce generalized sampling formulas to the corresponding formulas for finite sampling, the assumptions are made that the n populations are identical and that N and n are finite, $N > n$. The selection of variates which enter each sample is restricted in the following manner. If a variate having a given post-subscript is chosen, then no other variate having the same post-subscript may be chosen for the same sample.

Now it is evident that terms of the form $\mu_{t_1 t_2 \dots t_v}$ must be redefined on the basis of the preceding assumptions. From the expansions [20; p. 32] of power product sums in terms of products of power sums, we get the formulas for $\mu_{t_1 t_2 \dots t_v}$ which are given in the following tables.

The formulas in the tables of this section are called *transformation formulas for finite sampling* or more briefly *transformation formulas.*

The transformation of generalized sampling formulas into corresponding finite sampling formulas is illustrated by the substitution of $\frac{N^2\mu_1^2 - N\mu_2}{N^{(2)}}$ for $\mu_{1,1}$ in (1.27). We get

$$(1.41) \qquad {}'\mu_{1:\bar{m}_2} = \frac{N(n-1)}{n(N-1)} [\mu_2 - \mu_1^2],$$

which is the well-known finite sampling formula for the mean of the variance of samples of n.

From this and the preceding section it is evident that the generalized sampling formulas may be considered as formulas for either infinite or finite sampling depending upon the interpretation given to terms of the form $\mu_{t_1 t_2 \dots t_v}$.

9. Transformation of infinite sampling formulas into corresponding finite sampling formulas. It is a well-known fact that infinite sampling formulas may be obtained from those for finite sampling by letting the size of the parent population become infinite. But, prior to this paper, apparently no one has presented a method of obtaining finite sampling formulas from infinite sampling formulas. However, by making use of the relations between finite, infinite, and generalized sampling, we shall demonstrate that it is possible to transform any infinite sampling formula into the corresponding finite sampling formula.

Since the infinite sampling formulas are obtained from the generalized sampling formulas by replacing

$$\mu_{t_1 t_2 \dots t_v} \quad \text{by} \quad \mu_{t_1}\mu_{t_2} \cdots \mu_{t_v}$$

it follows that generalized sampling formulas may be obtained from the infinite

TABLE II

(1)

	Coef.	N
		μ_1
μ_1	$N^{(-1)}$	1

(2)

	Coef.	N	N^2
		μ_2	μ_1^2
μ_2	$N^{(-1)}$	1	
$\mu_{1,1}$	$N^{(-2)}$	−1	1

(3)

	Coef.	N	N^2	N^3
		μ_3	$\mu_2\mu_1$	μ_1^3
μ_3	$N^{(-1)}$	1		
$\mu_{2,1}$	$N^{(-2)}$	−1	1	
μ_{1^3}	$N^{(-3)}$	2	−3	1

(4)

	Coef.	N	N^2	N^2	N^3	N^4
		μ_4	$\mu_{3,1}$	μ_2^2	$\mu_2\mu_1^2$	μ_1^4
μ_4	$N^{(-1)}$	1				
$\mu_{3,1}$	$N^{(-2)}$	−1	1			
$\mu_{2,2}$	$N^{(-2)}$	−1		1		
$\mu_{2,1^2}$	$N^{(-3)}$	2	−2	−1	1	
μ_{1^4}	$N^{(-4)}$	−6	8	3	−6	1

(5)

	Coef.	N	N^2	N^2	N^3	N^3	N^4	N^5
		μ_5	$\mu_4\mu_1$	$\mu_3\mu_2$	$\mu_3\mu_1^2$	$\mu_2^2\mu_1$	$\mu_2\mu_1^3$	μ_1^5
μ_5	$N^{(-1)}$	1						
$\mu_{4,1}$	$N^{(-2)}$	−1	1					
$\mu_{3,2}$	$N^{(-2)}$	−1		1				
$\mu_{3,1^2}$	$N^{(-3)}$	2	−2	−1	1			
$\mu_{2^2,1}$	$N^{(-3)}$	2	−1	−2		1		
$\mu_{2,1^3}$	$N^{(-4)}$	−6	6	5	−3	−3	1	
μ_{1^5}	$N^{(-5)}$	24	−30	−20	20	15	−10	1

(6)

	Coef.	N	N^2	N^2	N^2	N^3	N^3	N^3	N^4	N^4	N^5	N^6
		μ_6	$\mu_5\mu_1$	$\mu_4\mu_2$	μ_3^2	$\mu_4\mu_1^2$	$\mu_3\mu_2\mu_1$	μ_2^3	$\mu_3\mu_1^3$	$\mu_2^2\mu_1^2$	$\mu_2\mu_1^4$	μ_1^6
μ_6	$N^{(-1)}$	1										
$\mu_{5,1}$	$N^{(-2)}$	−1	1									
$\mu_{4,2}$	$N^{(-2)}$	−1		1								
$\mu_{3,3}$	$N^{(-2)}$	−1			1							
$\mu_{4,1^2}$	$N^{(-3)}$	2	−2	−1		1						
$\mu_{3,2,1}$	$N^{(-3)}$	2	−1	−1	−1		1					
μ_{2^3}	$N^{(-3)}$	2		−3				1				
$\mu_{3,1^3}$	$N^{(-4)}$	−6	6	3	2	−3	−3		1			
$\mu_{2^2,1^2}$	$N^{(-4)}$	−6	4	5	2	−1	−4	−1		1		
$\mu_{2,1^4}$	$N^{(-5)}$	24	−24	−18	−8	12	20	3	−4	−6	1	
μ_{1^6}	$N^{(-6)}$	−120	144	90	40	−90	−120	−15	40	45	−15	1

formulas by replacing

$$\mu_{t_1}\mu_{t_2} \cdots \mu_{t_v} \quad \text{by} \quad \mu_{t_1 t_2 \cdots t_v}. \tag{1.42}$$

However, it must be emphasized that the application of (1.42) demands formulae which are expressed in terms of moments of sample moments rather than central moments of sample moments (although the sample moments may be measured about a fixed point or about the mean) and the moments of the universe must be measured about a fixed point The reason for these restrictions is to insure that each term is accounted for individually.

After replacements (1.42) are made in the formula for sampling from an infinite population, the resulting formula is the corresponding generalized one. The step to the corresponding finite sampling formulas is simply the one outlined in section eight, namely, the use of the transformation formulas.

We shall consider, as the first illustration, the infinite sampling formula for the mean of the sample variance when the moments of the parent population are measured about the mean. The formula is

$$\mu_{1:\bar{m}_2} = \frac{n-1}{n} \bar{\mu}_2. \tag{1.43}$$

When (1.43) is expressed in terms of moments of the parent population about a fixed point, we have

$$'\mu_{1:\bar{m}_2} = \frac{n-1}{n} [\mu_2 - \mu_1^2]. \tag{1.44}$$

Following (1.42), μ_1^2 is replaced by $\mu_{1,1}$ and (1.44) becomes (1.27). The use of the transformation formula for $\mu_{1,1}$ gives (1.41) which, when the moments of the parent population are measured about the mean, becomes

$$\mu_{1:\bar{m}_2} = \frac{N(n-1)}{n(N-1)} \bar{\mu}_2. \tag{1.45}$$

Infinite sampling formulas expressed in terms of moment-function, may be similarly transformed into the corresponding finite sampling formulas. For example, Craig [9; p. 57] gives the second Thiele seminvariant of the variance of samples as

$$\lambda_{2:\bar{m}_2} = \frac{(n-1)^2}{n^3} \lambda_4 + 2 \frac{(n-1)}{n^2} \lambda_2^2. \tag{1.46}$$

First, we express (1.46) in terms of moments about a fixed point by use of the formulas relating Thiele seminvariants and moments [9; p. 12]. We also recall that the resulting formula should be expressed in terms of moments of sample moments rather than in terms of central moments of sample moments. We obtain

$$\begin{aligned} '\mu_{2:\bar{m}_2} = \frac{(n-1)}{n^3} [(n-1)\mu_4 - 4(n-1)\mu_3\mu_1 + (n^2 - 2n + 3)\mu_2^2 \\ - 2(n-2)(n-3)\mu_2\mu_1^2 + (n-2)(n-3)\mu_1^4]. \end{aligned} \tag{1.47}$$

The next step is to transform (1.47) into the corresponding generalized sampling formula by use of (1.42). We obtain (1.32). Since we desire to obtain the finite sampling formula which exactly corresponds to (1.46), it is necessary to transform (1.32) from the second moment of $\bar{m}_2$ to the variance of $\bar{m}_2$ and we get (1.36). Next the transformation formulas are applied to (1.36). When the moments of the parent population are measured about the mean and are replaced by Thiele seminvariants, (1.36) becomes

$$\lambda_{2:\bar{m}_2} = \frac{N(N-n)(n-1)}{n^3(N-1)^2(N-2)(N-3)}[(N-1)(Nn-N-n-1)\lambda_4 + 2(N^2n - 3Nn - 3N + 3n + 3)\lambda_2^2]. \tag{1.48}$$

Formula (1.48) gives the second Thiele seminvariant of the variance of samples of n drawn from a finite parent population of N. When $N \to \infty$, in (1.48), we obtain immediately (1.46).

It is generally true that infinite sampling formulas are more easily derived than are the corresponding finite sampling formulas. The methods of this section make it possible to derive the desired sampling formulas for the infinite parent population and then transform these infinite sampling formulas into the corresponding finite sampling formulas.

II. Moment Function Adjustments for Grouped Data

A given distribution of discrete variates may be grouped in "k groupings of k". We desire to find the correction which eliminates the error made in replacing a given moment of the original distribution by the average of the corresponding moments of the k grouped-distributions.

Formulas for the adjustments for moments of a grouped-distribution of discrete variates were first given (without proof) in the Editorial of Vol. I, No. 1 of the *Annals of Mathematical Statistics*. Later, more satisfactory derivations of adjustment formulas were given by Abernethy [24] Craig [25] and Carver [26]. However, it was observed by Carver [26; p. 162] that the developments of Abernethy and Craig are adjustments about a fixed point and that they fail to hold for the case of expectations of central moments if we accept the definition

$$\mu_{1:\bar{\mu}_t} = \frac{1}{k}\sum_{r=1}^{k} {}_r\bar{\mu}_t, \qquad (t = 2, 3, \cdots).$$

Here ${}_r\bar{\mu}_t$ represents the tth central moment of the rth grouped-distribution. The formula for the true value of $\mu_{1:\bar{\mu}_2}$ was supplied by Carver [26; p. 162] but he did not indicate a general method which might be used for the derivation of $\mu_{1:\bar{\mu}_t}$, $(t > 2)$.

A distribution of discrete variates grouped in "k groupings of k" is a special case of a universe of n finite populations and hence the methods and formulas for the expectations of population moments are applicable to our present problem.

It is found that the adjustment formulas for moment-functions of grouped data involve central moments of a rectangular distribution. It will be convenient for our present purposes to give a brief treatment of the moment-functions of a rectangular distribution.

1. Moment-functions of a rectangular distribution. Consider the rectangular distribution of discrete variates,

$$h, 2h, 3h, \cdots, kh. \tag{2.1}$$

It is readily shown that the moment generating function of (2.1),

$$G_x(\theta) = \mu_0 + \mu_1\theta + \mu_2\frac{\theta^2}{2!} + \cdots + \mu_n\frac{\theta^n}{n!} + \cdots \tag{2.2}$$

may be written

$$G_x(\theta) = \frac{e^{\frac{1}{2}(k+1)h\theta}\sinh\frac{1}{2}kh\theta}{k\sinh\frac{1}{2}h\theta}. \tag{2.3}$$

Setting the expansion of the right member of (2.3) equal to the right member of (2.2) and equating coefficients of like powers of θ, we obtain the following recursion formula for the moments of (1.1)

$$\frac{(n+1)^{(1)}}{1!}\mu_{n:R} - \frac{(n+1)^{(2)}}{2!}h\mu_{n-1:R} + \cdots + (-1)^{r-1}\frac{(n+1)^{(r)}}{r!}h^{r-1}\mu_{n-r+1:R} + \cdots = k^n h^n, \tag{2.4}$$

where $\mu_{n:R}$ represents the nth moment of a rectangular distribution. Formulas for $\mu_{n:R}$, $(n = 0, 1, \cdots, 10)$ are given below. See Sasuly [27; p. 27].

$$\begin{aligned}
\mu_{0:R} &= 1.\\
\mu_{1:R} &= \tfrac{1}{2}(k+1)h.\\
\mu_{2:R} &= \tfrac{1}{6}(k+1)(2k+1)h^2 = \tfrac{1}{3}(2k+1)h\mu_{1:R}.\\
\mu_{3:R} &= \tfrac{1}{4}(k+1)^2 kh^3 = kh\,\mu_{1:R}^2.\\
\mu_{4:R} &= \tfrac{1}{5}(3k^2+3k-1)h^2\,\mu_{2:R}.\\
\mu_{5:R} &= \tfrac{1}{3}(2k^2+2k+1)h^2\,\mu_{3:R}.\\
\mu_{6:R} &= \tfrac{1}{7}(3k^4+6k^3-3k+1)h^4\,\mu_{2:R}.\\
\mu_{7:R} &= \tfrac{1}{6}(3k^4+6k^3-k^2-4k+2)h^4\,\mu_{3:R}.\\
\mu_{8:R} &= \tfrac{1}{15}(5k^6+15k^5+5k^4-15k^3-k^2+9k-3)h^6\,\mu_{2:R}.\\
\mu_{9:R} &= \tfrac{1}{5}(2k^6+6k^5+k^4-8k^3+k^2+6k-3)h^6\,\mu_{3:R}.\\
\mu_{10:R} &= \tfrac{1}{11}(3k^8+12k^7+8k^6-18k^5-10k^4+24k^3+2k^2-15k+5)h^8\,\mu_{2:R}.
\end{aligned} \tag{2.5}$$

The deviations about the mean of (2.1) are

$$(2.6)\qquad -\tfrac{1}{2}(k-1)h, \qquad -\tfrac{1}{2}(k-3)h, \cdots, \tfrac{1}{2}(k-3)h, \qquad \tfrac{1}{2}(k-1)h.$$

Therefore,

$$(2.7)\qquad \bar{\mu}_{2n+1:R} = 0.$$

If we denote (2.6) by $\bar{x}$, we have

$$(2.8)\qquad G_{\bar{x}}(\theta) = \frac{\sinh \frac{1}{2}(kh\theta)}{k \sinh \frac{1}{2}(h\theta)}.$$

The recursion formula for central moments of (2.1) is

$$(2.9)\qquad \frac{(2n+1)^{(1)}}{1!}\bar{\mu}_{2n:R} + \frac{h^2}{2^2}\frac{(2n+1)^{(3)}}{3!}\bar{\mu}_{2n-2:R} + \cdots + \frac{h^r}{2^r}\frac{(2n+1)^{(r+1)}}{(r+1)!}\bar{\mu}_{2n-r:R} + \cdots = \frac{k^{2n}h^{2n}}{2^{2n}}.$$

Formulas for $\bar{\mu}_{2n:R}$, $(n = 0, 1, \cdots, 5)$ are given below. See [27; p. 27].

$$(2.10)\qquad \begin{aligned} \bar{\mu}_{0:R} &= 1,\\ \bar{\mu}_{2:R} &= \tfrac{1}{12}(k^2-1)h^2,\\ \bar{\mu}_{4:R} &= \tfrac{1}{20}(3k^2-7)h^2\bar{\mu}_{2:R},\\ \bar{\mu}_{6:R} &= \tfrac{1}{112}(3k^4-18k^2+31)h^4\bar{\mu}_{2:R},\\ \bar{\mu}_{8:R} &= \tfrac{1}{960}(5k^6-55k^4+239k^2-381)h^6\bar{\mu}_{2:R},\\ \bar{\mu}_{10:R} &= \tfrac{1}{2816}(3k^8-52k^6+410k^4-1636k^2+2555)h^8\mu_{2:R}. \end{aligned}$$

From the relation which connects Thiele seminvariants and the moment generating function, we get, see [25; p. 57],

$$(2.11)\qquad \begin{aligned} &\lambda_{0:R} = 0, \qquad \lambda_{1:R} = \frac{(k+1)h}{2}, \qquad \lambda_{2n+1:R} = 0.\\ &\lambda_{2n:R} = (-1)^{n+1}\frac{B_n h^{2n}(k^{2n}-1)}{2n}, \qquad n = 1, 2, 3, \cdots \end{aligned}$$

where $\lambda_{n:R}$ represents the nth Thiele seminvariant of a rectangular distribution of discrete variates and B_n, $(n = 1, 2, \cdots)$, the Bernoulli numbers: $\frac{1}{6}, \frac{1}{30}, \cdots$.

In each of the cases considered in this section, corresponding formulas may be found for a rectangular distribution of continuous variates by setting $h = m/k$ (which makes the range m with k subdivisions) and then letting $k \to \infty$

2. Adjustments for moments. As our basic distribution we consider the set of discrete variates, x_i, $(i = 1, 2, \cdots, N)$, where some of the x_i's may not be distinct. We assume that the given distribution is grouped in "k groupings of k".

When x_i is placed in the rth position of a class, the limits of the class are $x_i - (r - 1)h$ and $x_i + (k - r)h$ and the class mark is $x_i + \left[\frac{k - (2r - 1)}{2}\right]h$. Thus, when the class mark is used as the value of x_i, the quantity $\left[\frac{k-(2r-1)}{2}\right]h$ is added to the true value of x_i. Therefore, when the expected value of a particular moment for "k groupings of k" is found, each variate has made a definite contribution as it was placed in each of the k positions of a class.

For convenience, we define

$$e_r = \left[\frac{k - (2r - 1)}{2}\right]h, \qquad (r = 1, 2, \cdots, k). \tag{2.12}$$

As was previously indicated, the expected value of a given moment involves the contribution of each variate as it occupies the k class positions. A convenient method of finding these contributions is by means of a universe ${}_kU_n$ which is composed of the populations ${}_rX$, $(r = 1, 2, \cdots, k)$. The rth population consists of the values of the variates when they occupy the rth position of the class. Hence ${}_rX$ consists of ${}_rx_i = x_i + e_r$, $(i = 1, 2, \cdots, N)$.

The notation for moments is the same as that of Part I. Since ${}_kU_N$ is of the same form as the universe studied in Part I, we use the definitions (1.1) of that part.

The expected value of the tth moment is

$$\begin{aligned} \mu_{1:\mu_t} &= \frac{1}{k}\sum_{r=1}^{k} E(x_i + e_r)^t \\ &= \sum_{s=0}^{t} \binom{t}{s}\left[\frac{1}{k}\sum_{r=1}^{k} e_r^s\right]\mu_{t-s}. \end{aligned}$$

Many devices have been used by previous writers [24; p. 269], [25; p. 57], [26; p. 157], to evaluate terms of the form $\frac{1}{k}\sum_{r=1}^{k} e_r^s$. However, it should be noticed that the quantities e_r, $(i = 1, 2, \cdots, k)$, are respectively identical with the deviations (2.6) about the mean of a rectangular distribution of discrete variates. It follows that

$$\bar{\mu}_{s:R} = \frac{1}{k}\sum_{r=1}^{k} e_r^s.$$

And since $\bar{\mu}_{2s+1:R} = 0$, we have

$$\mu_{1:\mu_t} = \sum_{s=0}^{[t/2]} \binom{t}{2s} \mu_{t-2s}\bar{\mu}_{2s:R}. \tag{2.13}$$

Formulas for $\bar{\mu}_{2s:R}$, $(s = 0, 1, \cdots, 5)$ are given by (2.10).

If the class marks are selected as the unit of x, we set $h = 1$ in (2.10). If the class interval is chosen as the unit of x, we set $h = 1/k$ in (2.10). If k consecutive values of the discrete variable are grouped in a frequency class of width m, we put $h = m/k$ in (2.10).

Usually we desire to estimate the value of the moments that would have been obtained if we had not grouped the data. Therefore (2.13) is solved for the moments of the ungrouped data. We have

$$\mu_t = \sum_{s=0}^{[t/2]} \binom{t}{2s} P_{2s} \mu_{1:\mu_{t-2s}} \tag{2.14}$$

wherein

$$P_{2s} = \Sigma \frac{(-1)^{\rho}(2s)!\,\rho!\,\bar{\mu}_{2p_1:R}^{\pi_1}\bar{\mu}_{2p_2:R}^{\pi_2}\cdots\bar{\mu}_{2p_v:R}^{\pi_v}}{[(2p_1)!]^{\pi_1}[(2p_2)!]^{\pi_2}\cdots[(2p_v)!]^{\pi_v}\pi_1!\,\pi_2!\cdots\pi_v!},$$

the summation being taken for every possible product of moments for which

$$\sum_{i=1}^{v} p_i = s, \qquad \sum_{i=1}^{v} \pi_i = \rho.$$

Formulas, corresponding to (2.13) and (2.14), for a distribution of continuous variates are written by replacing the moment symbols for discrete variates by those for continuous variates.

3. Adjustments for central moments. Consider the universe U which consists of the population ${}_rX$, $(r = 1, 2, \cdots, k)$, where ${}_rX$ is the rth grouped-distribution.

The expected value of the tth central moment of the k grouped-distribution is given by (1.3), (1.4) and (1.5) of Part I, where now $\mu_{1:\mu_{t-i}}$ is given by (2.13) of the preceding section. Thus, the development of this section is identical with that of section one of Part I with the single exception that $\mu_{1:\mu_t} = \mu_t$ no longer holds but is replaced by $\mu_{1:\mu_t} = \mu_t +$ *a correction*. Therefore, the formulas for the adjustments for central moments may be obtained immediately from the formulas derived in section one, Part I, if the corrections of the preceding section are inserted. We have

$$\mu_{1:\bar{\mu}_2} = \bar{\mu}_2 + \bar{\mu}_{2:R} - \bar{\mu}_{2:\mu_1} \tag{2.15}$$

$$\mu_{1:\bar{\mu}_3} = \bar{\mu}_3 + 6\mu_1\bar{\mu}_{2:\mu_1} - 3\bar{\mu}_{11:\mu_1\mu_2} + 2\bar{\mu}_{3:\mu_1} \tag{2.16}$$

$$\begin{aligned}\mu_{1:\bar{\mu}_4} = {} & \bar{\mu}_4 + 6\bar{\mu}_2\bar{\mu}_{2:R} + \bar{\mu}_{4:R} + 6(\bar{\mu}_2 - 2\mu_1^2 + \bar{\mu}_{2:R})\bar{\mu}_{2:\mu_1} \\ & + 12\mu_1\bar{\mu}_{11:\mu_1\mu_2} - 12\mu_1\bar{\mu}_{3:\mu_1} - 4\bar{\mu}_{11:\mu_1\mu_3} \\ & + 6\bar{\mu}_{21:\mu_1\mu_2} - 3\bar{\mu}_{4:\mu_1}\end{aligned} \tag{2.17}$$

The moments of the ungrouped data can be obtained readily from formulas (2.15) through (2.17).

Adjustment formulas for central moments of a distribution of continuous variates may be obtained from (2.13) by replacing the moment symbols for discrete variates by those for continuous variates and taking the moments about the mean. Also, it may be observed that adjustment formulas for central moments of a distribution of continuous variates may be obtained from formulas (1.3), (1.4) and (1.5) of Part I, provided the moment symbols are exchanged as indicated above and terms of the form $\bar{\nu}_{s_1 s_2 \cdots s_v : \nu_{t_1} \nu_{t_2} \cdots \nu_{t_v}}$ are set equal to zero.

4. Usual adjustments for Thiele seminvariants. The usual adjustments for Thiele seminvariants, for the univariate discrete population, may be developed directly by use of one of the fundamental properties of Thiele seminvariants.

It is assumed (see [25; p. 55]) that k consecutive values of the discrete variable are grouped in a frequency class of width m. The k smaller intervals of width $m/k = h$ go to make up the class width m, the actual points representing the k values of the variable being plotted at the centers of the sub-intervals. Now, let us suppose that each of the k consecutive boundary points of the subintervals is as likely to be chosen as a boundary point of the larger intervals as any other. Then, if x_i is the class mark of the ith frequency class, for any true value, x, of the discrete variable included in this frequency class, we have

$$x_i = x + e_r$$

in which x and e_r are independent variables and e_r takes on the k values (2.12) with equal relative frequencies $1/k$.

Since we have noted that the equally likely values which e_r may take on are deviations about the mean of a rectangular distribution of discrete variates, we employ the cumulative property of Thiele seminvariants [9; p. 4] and obtain directly

$$\lambda'_{t:x} = \lambda_{t:x} + \lambda_{t:R}, \qquad (t = 1, 2, \cdots), \tag{2.18}$$

where $\lambda'_{t:x}$ is the tth seminvariant computed from the grouped data, $\lambda_{t:x}$ is the tth seminvariant computed from the ungrouped data and $\lambda_{t:R}$ is defined by (2.11).

Formulas corresponding to (2.18), for special values of t, are given by Craig [25; p. 57]. However, the present development indicates the dependence of adjustment formulas on central moments of a rectangular distribution and provides a general formula for these adjustments which is expressed completely in terms of Thiele seminvariants.

5. New adjustments for Thiele seminvariants. If we accept the definition

$$\mu_{1:\bar{\mu}_t} = \frac{1}{k} \sum_{r=1}^{k} {}_r\bar{\mu}_t, \qquad (t = 2, 3, \cdots),$$

then (2.18) is at best only an approximation formula. We now desire exact formulas for $\mu_{1:\lambda_t}$ for the case of a grouped-distribution of discrete variates.

First (1.9) is used and terms of the form $\mu_{s_1 s_2 \cdots s_v : \mu_1 \mu_2 \cdots \mu_v}$ are evaluated in terms of central moments by (1.3). Then terms of the form $\mu_{1:\mu_t}$ are evaluated by (2.13) and finally the relations between moments and Thiele seminvariants are employed. Exact formulas for the expected values of the second, third, and fourth Thiele seminvariants for grouped-distributions of discrete variables are given below.

(2.19) $$\mu_{1:\lambda_2} = \lambda_2 + \lambda_{2:R} - \bar{\mu}_{2:\mu_1} .$$

(2.20) $$\mu_{1:\lambda_3} = \lambda_3 + 6\lambda_1\bar{\mu}_{2:\mu_1} - 3\bar{\mu}_{11:\mu_1\mu_2} + 2\bar{\mu}_{3:\mu_1} .$$

(2.21) $$\mu_{1:\lambda_4} = \lambda_4 + \lambda_{4:R} + 12[\lambda_2 - 2\lambda_1^2 + \lambda_{2:R}]\bar{\mu}_{2:\mu_1} + 24[\bar{\mu}_{11:\mu_1\mu_2} - \bar{\mu}_{3:\mu_1}]\lambda_1 - 4\bar{\mu}_{11:\mu_1\mu_3} + 12\bar{\mu}_{21:\mu_1\mu_2} - 6\bar{\mu}_{4:\mu_1} - 3\bar{\mu}_{2:\mu_2} .$$

Formulas for Thiele seminvariants of ungrouped data in terms of expectations may be obtained from (2.19) through (2.21).

Adjustment formulas for Thiele seminvariants of a distribution of continuous variates are given by Langdon and Ore [23; p. 231] and Craig [25; p. 57]. If we denote the tth Thiele seminvariant of a distribution of continuous variates by L_t, then

(2.22) $$\nu_{1:L_t} = L_t + L_{t:R} ,$$

where

(2.23) $$L_{2t+1:R} = 0, \qquad L_{2t:R} = \frac{(-1)^{t-1}B_t m^{2t}}{2t}, \qquad t = 1, 2, \cdots .$$

Formulas (2.19) through (2.21) may be used for continuous variates by changing the moment symbols and setting terms of the form $\bar{\mu}_{s_1 s_2 \cdots s_v : \mu_{t_1}\mu_{t_2}\cdots\mu_{t_v}}$ equal to zero.

6. Adjustment formulas applied to a numerical problem. We consider the arbitrary distribution given in Table III.

TABLE III

An Arbitrary Distribution of Discrete Variates

v	f	v	f	v	f	F
1	2	4	30	7	1	2 + 30 + 1 = 33
2	8	5	4	8	1	8 + 4 + 1 = 13
3	10	6	3	9	1	10 + 3 + 1 = 14

The three grouped distributions, when the variates are grouped in "groupings of three," appear in Table IV.

TABLE IV

Distributions Derived from Data of Table III by Making the Three Possible Groupings of Three

(1) Class	(1) f	(2) Class	(2) f	(3) Class	(3) f
1-3	20	0-2	10	−1 to 1	2
4-6	37	3-5	44	2-4	48
7-9	3	6-8	5	5-7	8
10-12	0	9-11	1	8-10	2

Using the fixed point 4, moment-functions are computed for the distribution of Table III and for each of the distributions of Table IV. These quantities along with the average of each moment function appear in Table V.

TABLE V

Moment-Functions of the Distributions of Table III and Table IV. Averages of Moment-Functions of Distributions of Table IV

Dist.	μ_1	μ_2	μ_3	μ_4	$\bar{\mu}_2 = \lambda_2$	$\bar{\mu}_3 = \lambda_3$	$\bar{\mu}_4$	λ_4
(1)	$\frac{9}{60}$	$\frac{165}{60}$	$\frac{69}{60}$	$\frac{1125}{60}$	$\frac{9819}{(60)^2}$	$\frac{-17442}{(60)^3}$	$\frac{238,849,317}{(60)^4}$	$\frac{-50,388,966}{(60)^4}$
(2)	$\frac{-9}{60}$	$\frac{171}{60}$	$\frac{81}{60}$	$\frac{2511}{60}$	$\frac{10179}{(60)^2}$	$\frac{567162}{(60)^3}$	$\frac{557,840,277}{(60)^4}$	$\frac{247,004,154}{(60)^4}$
(3)	$\frac{-30}{60}$	$\frac{162}{60}$	$\frac{138}{60}$	$\frac{1938}{60}$	$\frac{8820}{(60)^2}$	$\frac{1317600}{(60)^3}$	$\frac{528,282,000}{(60)^4}$	$\frac{294,904,800}{(60)^4}$
Ave.	$\frac{-10}{60}$	$\frac{166}{60}$	$\frac{96}{60}$	$\frac{1858}{60}$	$\frac{9606}{(60)^2}$	$\frac{622440}{(60)^3}$	$\frac{441,657,198}{(60)^4}$	$\frac{163,839,996}{(60)^4}$
Orig. Dist.	$\frac{-10}{60}$	$\frac{126}{60}$	$\frac{116}{60}$	$\frac{1314}{60}$	$\frac{7460}{(60)^2}$	$\frac{642400}{(60)^3}$	$\frac{305,034,000}{(60)^4}$	$\frac{138,079,200}{(60)^4}$

Table VI gives the expected values of the moment-functions as obtained by substituting from Table V into the formulas of sections two, three, and five. Also the expected values, computed from the usual formulas, are given and the errors which would be made, if the usual formulas were used, are indicated.

TABLE VI

Expected Values of Moment-Functions Computed by Formulas

Expectations by	$\mu_{1:\mu_1}$	$\mu_{1:\mu_2}$	$\mu_{1:\mu_3}$	$\mu_{1:\mu_4}$	$\mu_{1:\bar{\mu}_2} = \mu_{1:\lambda_2}$	$\mu_{1:\bar{\mu}_3} = \mu_{1:\lambda_3}$	$\mu_{1:\bar{\mu}_4}$	$\mu_{1:\lambda_4}$
New Formulas	$\frac{-10}{60}$	$\frac{166}{60}$	$\frac{96}{60}$	$\frac{1858}{60}$	$\frac{9606}{(60)^2}$	$\frac{622440}{(60)^3}$	$\frac{441,657,198}{(60)^4}$	$\frac{163,839,996}{(60)^4}$
Usual Formulas	$\frac{-10}{60}$	$\frac{166}{60}$	$\frac{96}{60}$	$\frac{1858}{60}$	$\frac{9860}{(60)^2}$	$\frac{642400}{(60)^3}$	$\frac{416,778,000}{(60)^4}$	$\frac{133,795,200}{(60)^4}$
Error	—	—	—	—	$\frac{254}{(60)^2}$	$\frac{19960}{(60)^3}$	$\frac{-24,879,198}{(60)^4}$	$\frac{-30,060,796}{(60)^4}$

7. Evaluation of $\bar{\mu}_{2:\mu_1}$. It appears at first that it is necessary to form the "k groupings of k" in order to evaluate the term $\bar{\mu}_{2:\mu_1}$ which enters the precise formula for the expected value of the variance. That was the procedure followed by Carver [26; p. 161]. However, it is possible to evaluate $\bar{\mu}_{2:\mu_1}$ from the ungrouped data without forming a single grouped-distribution.

By definition,

$$\bar{\mu}_{2:\mu_1} = \frac{1}{k}\sum_{r=1}^{k} [{}_r\mu_1 - \mu_1]^2,$$

where ${}_r\mu_1$ is the mean of the rth grouped-distribution and μ_1 is the mean of the ungrouped distribution. We wish to study the terms ${}_r\mu_1$ and μ_1. Consider a set of variates x_i, $(i = 1, 2, \cdots, s)$, with corresponding frequencies f_i, $(i = 1, 2, \cdots, s)$. The x's are subject to the condition, $x_i - x_{i-1} = 1$, and consequently some of the f's may be zero. The mean of this distribution is $\frac{\Sigma xf}{\Sigma f}$.

We define

$$F_i = f_i + f_{k+i} + f_{2k+i} + \cdots, \qquad (i = 1, 2, \cdots, k)$$

Then, if a grouped-distribution is formed with x_i in the ith $(i = 1, 2, \cdots, k)$ position of a class, the mean of this grouped-distribution is

$$\frac{\sum xf + \sum_{j=1}^{k} F_j e_{i+j-1}}{\sum f}$$

where $e_{i-1} = e_k$ if $e_i = 1$ and $e_{i+1} = e_1$ if $e_i = e_k$. Similarly if a grouped-distribution is formed with x_i in the $(i + 1)$st position of a class, the mean is

$$\frac{\sum xf + \sum_{j=1}^{k} F_j e_{i+j}}{\sum f}.$$

Thus, it is evident that, given the expression for the mean of any grouped-distribution in which x_i is in the ith position of a class, we may form the expression for the mean of the grouped-distribution in which x_i is in the $(i + 1)$st position of a class by a cyclic permutation of the e_i's of the given expression.

Therefore, it follows that if we call ${}_r\mu_1$ the mean of the grouped-distribution in which x_i is in the rth $(r = 1, 2, \cdots, k)$ position of a class, then

$$ {}_r\mu_1 - \mu_1 = \frac{\sum_{j=1}^{k} F_j e_{r+j-1}}{\sum f}, \qquad (r = 1, 2, \cdots, k). $$

If we define

$$ N = \Sigma f \quad \text{and} \quad \phi_r = \sum_{j=1}^{k} F_j e_{r+j-1} $$

then,

$$ \bar{\mu}_{2:\mu_1} = \frac{1}{kN^2} \sum_{r=1}^{k} \phi_r^2 . $$

Thus, it is evident that $\bar{\mu}_{2:\mu_1}$ is a function of the frequencies of the variates and of the e_i's. The fact that the values of the variates do not enter $\bar{\mu}_{2:\mu_1}$ permits one to quickly calculate its value.

Consider $\bar{\mu}_{2:\mu_1}$ for the distribution of Table III. We find

$$ \phi_1 = 33e_1 + 13e_2 + 14e_3 . $$

Then, by successive cyclic permutations of the e_i's,

$$ \phi_2 = 33e_2 + 13e_3 + 14e_1 , $$

$$ \phi_3 = 33e_3 + 13e_1 + 14e_2 . $$

Substituting the values $e_1 = 1$, $e_2 = 0$, $e_3 = -1$ we have $\phi_1 = 19$, $\phi_2 = 1$ and $\phi_3 = -20$. Therefore,

$$ \bar{\mu}_{2:\mu_1} = \frac{254}{(60)^2} $$

which is identical with the value which was found when Table V was used.

It follows from the preceding development that

$$ \bar{\mu}_{t:\mu_1} = \frac{1}{kN^t} \sum_{r=1}^{k} \phi_r^t $$

and if $F_1 = F_2 = \cdots = F_k$ then $\bar{\mu}_{t:\mu_1}$ is zero.

8. Conclusion. The results of this paper include:

1. The derivation of general and specific formulas for the expected values of population moment-functions.

2. The derivation of generalized sampling formulas under the condition that samples of n are formed by selecting one variate from each population.

3. Methods for the transformation of generalized sampling formulas into the corresponding infinite and finite sampling formulas.

4. A method for the transformation of infinite sampling formulas into the corresponding finite sampling formulas.

5. A demonstration of the fact that adjustment formulas for moment-function of grouped data involve central moments of a rectangular distribution.

6. A general formula for the expected value of the tth moment of grouped data.

7. New adjustment formulas for central moments of grouped data.

8. New adjustment formulas for Thiele seminvariants of grouped data.

9. A method for the evaluation of the term $\bar{\mu}_{2:\mu_1}$ which appears in the precise adjustment formula for the variance.

Many thanks are due Prof. P. S. Dwyer, to whom the writer is greatly indebted for advice and encouragement.

REFERENCES

[1] "Student", "The Probable Error of the Mean," *Biometrika*, Vol. 6 (1908), pp. 1–25.
[2] L. Isserlis, "On the Value of a Mean as Calculated from a Sample," *Jour. of Roy. Stat. Soc.*, Vol. 81 (1918), pp. 75–81.
[3] A. Tchouproff, "On the Mathematical Expectation of the Moments of Frequency Distributions," Part I, *Biometrika*, Vol. 12 (1918), pp. 140-169, 185–210; Part II, *Biometrika*, Vol. 13 (1920-21), pp. 283-295.
[4] A. Tchouproff, "On the Mathematical Expectation of the Moments of Frequency Distributions in the Case of Correlated Observations," *Metron*, Vol. 2 (1923), pp. 461–493; 646-683.
[5] A. Church, "On the Moments of the Distributions of Squared Deviations for Samples of N Drawn from an Indefinitely Large Population," *Biometrika*, Vol. 17 (1925), pp. 79-83.
[6] J. Neyman, "Contributions to the Theory of Small Samples Drawn from a Finite Population," *Biometrika*, Vol. 17 (1925), pp. 472-479.
[7] A. Church, "On the Means and Squared Deviations of Small Samples from any Population," *Biometrika*, Vol. 18 (1926), pp. 321-394.
[8] K. Pearson, "Another Historical Note on the Problem of Small Samples," *Biometrika* Vol. 19 (1927), pp. 207–210.
[9] C. C. Craig, "An Application of Thiele's Seminvariants to the Sampling Problem," *Metron*, Vol. 7 (1928-29), pp. 3-74.
[10] R. A. Fisher, "Moments and Product Moments of Sampling Distributions," *Proc. London Math. Soc.*, Vol. 2 (30) (1929), pp. 199-238.
[11] J. Pepper, "Studies in the Theory of Sampling," *Biometrika*, Vol. 11 (1929), pp. 231–258
[12] H. C. Carver, "Fundamentals of Sampling," *Annals of Math. Stat.*, Vol. 1 (1930), pp. 101-121, 260–274.
[13] N. St. Georgesque, "Further Contributions to the Sampling Problem," *Biometrika*, Vol. 24 (1932), pp. 65–107.
[14] G. M. Brown, "On Sampling from Compound Populations," *Annals of Math. Stat.*, Vol. 4 (1933), pp. 288–342.
[15] J. Neyman, "On the Two Different Aspects of the Representative Method," *Jour. Roy. Stat. Soc.*, Vol. 97 (1934), pp. 558-625.

[16] A. L. Bowley, "Measurement of the Precision Attained in Sampling," *Bull. International Stat. Inst.*, Vol. 22.

[17] H. Feldman, "Mathematical Expectations of Product Moments of Samples Drawn from a Set of Infinite Populations," *Annals of Math. Stat.*, Vol. 6 (1935), pp. 30-52.

[18] H. C. Carver, "Sheppard's Corrections for a Discrete Variable," *Annals of Math. Stat.*, Vol. 76 (1936), pp. 154-163.

[19] P. S. Dwyer, "Moments of Any Rational, Integral Isobaric Sample Moment Function," *Annals of Math. Stat.*, Vol. 8 (1937), pp. 21-65.

[20] P. S. Dwyer, "Combined Expansions of Products of Symmetric Power Sums and of Sums of Symmetric Power Products with Application to Sampling," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 1-84.

[21] A. T. Craig, "On the Mathematics of the Representative Method of Sampling," *Annals of Math. Stat.*, Vol. 10 (1939), pp. 26-34.

[22] A. Fisher, *The Mathematical Theory of Probabilities*, New York, The Macmillan Company. 1928.

[23] W. H. Langdon, and Ore, O., "Seminvariants and Sheppard's Corrections," *Annals of Math.*, Vol. 31 (1930), pp. 230-232.

[24] J. R. Abernethy, "On the Elimination of Systematic Errors Due to Grouping," *Annals of Math. Stat.*, Vol. 4 (1933), pp. 263-277.

[25] C. C. Craig, "Sheppard's Corrections for a Discrete Variable," *Annals of Math. Stat.*, Vol. 7 (1936), pp. 55-61.

[26] H. C. Carver, "The Fundamental Nature and Proof of Sheppard's Adjustments," *Annals of Math. Stat.*, Vol. 9 (1936), pp. 154-163.

[27] M. Sasuly, *Trend Analysis of Statistics*, Washington, The Brookings Institution (1934).

The University of Michigan
Ann Arbor, Michigan