LEAST-SQUARES ESTIMATES USING ORDERED OBSERVATIONS

By F. Downton

University of Liverpool

1. Introduction and summary. The purpose of this paper is to compare for various two-parameter distributions, of the form $f\{(x - \mu)/\sigma\}/\sigma$, the estimates of the parameters obtained by applying the method of least squares to the observations, after these have been arranged in order of magnitude. Estimates obtained by this process we shall call "ordered least-squares estimates." Such estimates are unbiased and have minimal variance among all unbiased estimates which are linear in the ordered observations.

This estimation process has been previously discussed by Godwin [1] and [2] and Lloyd [3]. In the present paper, ordered estimates are obtained explicitly for a class of two-parameter distributions having the above form. This class contains the rectangular and the right triangular distributions as special cases. It also reduces to the exponential distribution as a limiting case. Other special cases of this class of distributions have also been previously discussed by Craig [4].

Further, a general property of ordered least-squares estimates of the parameter λ in distributions of the type $f(x/\lambda)/\lambda$ is discussed. As a result it is shown that the ordered least-squares estimate of the scale parameter in the Pearson Type III distribution is identical with the maximum likelihood estimate.

2. Notation and general theory. Let $x_1, x_2, x_3, \cdots x_n$ be a sample of n independent observations on a continuous variate X whose distribution has the form $f\{(x - \mu)/\sigma\}/\sigma$. We may write

$$x_{(1)} < x_{(2)} < x_{(3)} \cdots < x_{(n)}$$

for the ordered observations.

Let $y_r = (x_r - \mu)/\sigma$ and $y_{(r)} = (x_{(r)} - \mu)/\sigma$ be the reduced observations, unordered and ordered, respectively.

For $r, s = 1, 2, 3, \cdots n$, let $\mathcal{E}(y_{(r)}) = \alpha_r$, $\text{cov}(y_{(r)}, y_{(s)}) = v_{rs}$.

Let $\boldsymbol{\alpha}$ denote the $(n \times 1)$ vector of the α_r; $\mathbf{v}$ the symmetric, positive-definite $(n \times n)$ matrix of the v_{rs}; $\mathbf{1}$ an $(n \times 1)$ vector of 1's; $\mathbf{x}$ the $(n \times 1)$ vector of the $x_{(r)}$; and $\mathbf{y}$ the $(n \times 1)$ vector of the $y_{(r)}$. The inverse of $\mathbf{v}$ is $\mathbf{v}^{-1}$ with elements v_{rs}^{-1}.

The ordered least-squares estimates of μ and σ are then

$$\hat{\mu} = \boldsymbol{\alpha}'\mathbf{v}^{-1}(\boldsymbol{\alpha}\mathbf{1}' - \mathbf{1}\boldsymbol{\alpha}')\mathbf{v}^{-1}\mathbf{x}/\Delta \tag{2.1}$$

$$\hat{\sigma} = \mathbf{1}'\mathbf{v}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{v}^{-1}\mathbf{x}/\Delta \tag{2.2}$$

where

Received 6/11/53.

$$\Delta = (\mathbf{1}'\mathbf{v}^{-1}\mathbf{1})(\mathbf{a}'\mathbf{v}^{-1}\mathbf{a}) - (\mathbf{1}'\mathbf{v}^{-1}\mathbf{a})^2 \tag{2.3}$$

Also

$$\text{var}(\hat{\mu}) = \mathbf{a}'\mathbf{v}^{-1}\mathbf{a}\sigma^2/\Delta \tag{2.4}$$

$$\text{var}(\hat{\sigma}) = \mathbf{1}'\mathbf{v}^{-1}\mathbf{1}\sigma^2/\Delta \tag{2.5}$$

$$\text{cov}(\hat{\mu}, \hat{\sigma}) = \mathbf{1}'\mathbf{v}^{-1}\mathbf{a}\sigma^2/\Delta. \tag{2.6}$$

These results were given by Lloyd [3].

3. Two-parameter distribution with explicit ordered least-squares solution. We introduce the generalized geometric variate X whose density function is

$$f(x) = \begin{cases} \dfrac{p}{\sigma b^p}\left(\dfrac{x - \mu}{\sigma} + a\right)^{p-1}, & \mu - a\sigma \leqq x < \mu - (a - b)\sigma, \\ 0, & \text{otherwise,} \end{cases} \tag{3.1}$$

where $p \geqq 1$, and $a = \sqrt{p(p + 2)}$ and $b = \sqrt{(p + 1)^3/p}$.

The expectation and variance of X are

$$\mathcal{E}(X) = \mu, \qquad \text{var}(X) = \sigma^2.$$

It will be shown that for all $p \geqq 1$ it is possible to find explicitly the expectation vector and the variance matrix of the reduced ordered observations, and hence the ordered least-squares estimates of μ and σ.

The standardised form of the variate X is $Y = (X - \mu)/\sigma$, for which

$$\mathcal{E}(Y) = 0, \qquad \text{var}(Y) = 1.$$

The density function of Y will be

$$f(y) = \begin{cases} pb^{-p}(y + a)^{p-1}, & -a \leqq y < b - a \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

In order to apply the results of Section 2 it is convenient to define the variate

$$T = (Y + a)/b = (X - \mu + a\sigma)/b. \tag{3.3}$$

Its density function is

$$f(t) = \begin{cases} pt^{p-1}, & 0 \leqq t < 1 \\ 0, & \text{otherwise.} \end{cases}$$

Its distribution function is

$$F(t) = \begin{cases} 0, & t < 0, \\ t^p, & 0 \leqq t < 1, \\ 1, & 1 \leqq t. \end{cases} \tag{3.4}$$

Its expectation and variance are, respectively,

$$\mathcal{E}(T) = a/b = p/(p+1) \tag{3.5}$$

$$\text{var}(T) = 1/b^2 = p/(p+2)(p+1)^2.$$

For the vector $\mathbf{T}$ of ordered values of T, let

$$\mathcal{E}(\mathbf{T}) = \mathbf{a}, \qquad \text{var}(\mathbf{T}) = \mathbf{w}.$$

To obtain the relationship between these quantities and the expectation and variance of the vector of reduced ordered observations $\mathbf{y}$, we note that since (3.3) gives T as a monotonic increasing function of Y, it follows that

$$\mathbf{y} = b\mathbf{t} - a\mathbf{1} \tag{3.6}$$

where $\mathbf{t}$ is the vector of ordered observations on T. Taking expectations and variances,

$$\boldsymbol{\alpha} = b\mathbf{a} - a\mathbf{1}, \qquad \mathbf{v} = b^2\mathbf{w}.$$

In terms of p we then have

$$\boldsymbol{\alpha} = (p+1)\sqrt{(p+2)/p}\,\{\mathbf{a} - p\mathbf{1}/(p+1)\} \tag{3.7}$$

$$\mathbf{v}^{-1} = p\mathbf{w}^{-1}/(p+2)(p+1)^2. \tag{3.8}$$

We now turn to the explicit calculation of these vectors and matrices. The rth element of $\mathbf{a}$ is

$$\mathrm{a}_r = \{\Gamma(n+1)/\Gamma(r)\Gamma(n-r+1)\} \int_0^1 t\{F(t)\}^{r-1}\{1 - F(t)\}^{n-r} f(t)\,dt$$

where $f(t)$ and $F(t)$ are defined in (3.4). This reduces to

$$\mathrm{a}_r = n^{(n-r+1)} p^{n-r+1}/(np+1)^{(n-r+1)} \tag{3.9}$$

where $n^{(s)} = n(n-1) \cdots (n-s+1)$ and $(np+1)^{(s)} = (np+1)([n-1]p+1) \cdots ([n-s+1]p+1)$. Similarly if w_{rs} is the (r, s) element of $\mathbf{w}$,

$$w_{rr} = \{n^{(n-r+1)} p^{n-r+1}/(pn+2)^{(n-r+1)}\} - \mathrm{a}_r^2 \tag{3.10}$$

and for $r < s$

$$w_{sr} = w_{rs} = \{n^{(n-r+1)} p^{n-r+1}/(pn+2)^{(n-s+1)}(p[s-1]+1)^{(s-r)}\} - \mathrm{a}_r \mathrm{a}_s. \tag{3.11}$$

If w_{rs}^{-1} is the (r, s) element of a matrix $\mathbf{w}^{-1}$, where

$$w_{rs}^{-1} = 0 \qquad \text{when } |r - s| > 1, \tag{3.12}$$

$$w_{r,r-1}^{-1} = w_{r-1,r}^{-1} = -(p[r-1]+1)(pn+2)^{(n-r+2)}/p^{n-r+1}n^{(n-r+1)}, \tag{3.13}$$

$$w_{rr}^{-1} = \{p^2(2r^2 - 2r + 1) + 2p(2r-1) + 1\}(pn+2)^{(n-r+1)} / p^{n-r+1}n^{(n-r+1)}, \tag{3.14}$$

it may be shown by multiplication that $\mathbf{w}\mathbf{w}^{-1} = \mathbf{1}^{\delta}$, the unit matrix and hence $\mathbf{w}^{-1}$ is the unique inverse of $\mathbf{w}$.

We now evaluate various quantities needed for computing the estimates; we have

(3.15) $\mathbf{a}'\mathbf{w}^{-1}\mathbf{a} = pn\,(pn + 2)$

(3.16) $\mathbf{a}'\mathbf{w}^{-1}\mathbf{1} = \mathbf{1}'\mathbf{w}^{-1}\mathbf{a} = (pn + 1)(pn + 2)$

(3.17) $\mathbf{1}'\mathbf{w}^{-1}\mathbf{1} = (p - 1)S + (pn + 1)(pn + 2) + 2(pn + 2)^{(n)}/n!p^n$

where

$$S = \begin{cases} (pn + 2)/(p - 2) - 2(pn + 2)^{(n)}/n!p^n, & \text{for } p \neq 2 \quad (3.18) \\ (n + 1)\sum_{r=1}^{n} \frac{1}{r} & \text{for } p = 2. \quad (3.19) \end{cases}$$

We now convert these expressions into the form in which they are used in Section 2, using equations (3.7) and (3.8). Then

$$\Delta = (\mathbf{1}'\mathbf{v}^{-1}\mathbf{1})(\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha}) - (\mathbf{1}'\mathbf{v}^{-1}\boldsymbol{\alpha})^2$$

$$= \frac{p}{(p + 1)^2(p + 2)}\{(\mathbf{1}'\mathbf{w}^{-1}\mathbf{1})(\mathbf{a}'\mathbf{w}^{-1}\mathbf{a}) - (\mathbf{1}'\mathbf{w}^{-1}\mathbf{a})^2\}$$

$$= \begin{cases} \dfrac{(pn + 2)}{(p + 1)^2(p + 2)(p - 2)}\left\{(pn + 2)(p[n - 1] + 2) - \dfrac{2pn(pn + 2)^{(n)}}{n!p^n}\right\} & p \neq 2, \quad (3.20) \\ 2(n + 1)^2\left\{n\sum_{r=1}^{n}\dfrac{1}{r} - 1\right\} & p = 2. \quad (3.21) \end{cases}$$

Also

(3.22) $\boldsymbol{\alpha}'\mathbf{v}^{-1}(\boldsymbol{\alpha}\mathbf{1}' - \mathbf{1}\boldsymbol{\alpha}')\mathbf{v}^{-1}$

$$= p[\mathbf{a}' - p\mathbf{1}'/(p + 1)][\mathbf{w}^{-1}(\mathbf{a}'\mathbf{1}' - \mathbf{1}\mathbf{a}')\mathbf{w}^{-1}/(p + 1)^2(p + 2)]$$

and

(3.23) $\mathbf{1}'\mathbf{v}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{v}^{-1} = p^{3/2}\mathbf{1}'\mathbf{w}^{-1}(\mathbf{1}\mathbf{a}' - \mathbf{a}\mathbf{1}')/(p + 1)^3(p + 2)^{3/2}$.

If, in the $(1 \times n)$ vector $\boldsymbol{\alpha}'\mathbf{v}^{-1}(\boldsymbol{\alpha}\mathbf{1}' - \mathbf{1}\boldsymbol{\alpha}')\mathbf{v}^{-1} = \boldsymbol{\beta}$, the rth term is, say, β_r, then

$$\beta_r = \frac{p^2(p - 1)(n - 1)(pn + 2)(pn + 2)^{(n-r+1)}}{(p + 1)^3(p + 2)n^{(n-r+1)}p^{n-r+1}} + \eta_r. \quad (3.24)$$

where, with S as defined in (3.18) and (3.19),

$$\eta_1 = \frac{2p^2(n - 1)(pn + 2)^{(n-r+1)}}{(p + 1)^3(p + 2)n^{(n-r+1)}p^{n-r+1}}, \quad (3.25a)$$

(3.25b) $$\eta_r = 0 \qquad \text{for } 2 \leqq r \leqq n - 1,$$

(3.25c) $$\eta_n = \frac{p}{(p+1)^3(p+2)}\left\{(p-1)(pn+1)(pn+2)(pS - pn - 2) + \frac{2p(pn+1)(pn+2)^{(n)}}{n!p^n}\right\}.$$

In the $(1 \times n)$ vector $\mathbf{1}'\mathbf{v}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{v}^{-1} = \boldsymbol{\gamma}$, say, if the rth term is γ_r, then

(3.26) $$\gamma_r = \frac{-p^{3/2}(p-1)(pn+1)(pn+2)^{(n-r+1)}}{(p+1)^3(p+2)^{3/2}n^{(n-r+1)}p^{n-r+1}} + \nu_r$$

where

(3.27a) $$\nu_1 = \frac{2p^{3/2}(pn+1)(pn+2)(pn+2)^{(n)}}{(p+1)^3(p+2)^{3/2}n!p^n}$$

(3.27b) $$\nu_r = 0 \qquad \text{for } 2 \leqq r \leqq n - 1$$

(3.27c) $$\nu_n = \frac{p^{3/2}(pn+1)(pn+2)}{(p+1)^3(p+2)^{3/2}}\left\{(p-1)S + \frac{2(pn+2)^{(n)}}{n!p^n}\right\}.$$

The only other quantities necessary for substitution in the expressions of Section 2 are $\mathbf{1}'\mathbf{v}^{-1}\mathbf{1}$, $\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha}$, and $\mathbf{1}'\mathbf{v}^{-1}\boldsymbol{\alpha}$. These quantities may be shown to be given by

(3.28) $$\mathbf{1}'\mathbf{v}^{-1}\mathbf{1} = \frac{p}{(p+1)^3(p+2)}\left\{(p-1)S + (pn+1)(pn+2) + \frac{2(pn+2)^{(n)}}{n!p^n}\right\}$$

(3.29) $$\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha} = (2pn+1)(pn+2) + \frac{p}{p+1} \cdot \{p(p-1)S - 2(pn+1)(pn+2)\} + \frac{2(pn+2)^{(n)}}{n!p^n}$$

(3.30) $$\mathbf{1}'\mathbf{v}^{-1}\boldsymbol{\alpha} = \frac{p^{1/2}(pn+1)(pn+2)}{(p+1)(p+2)^{1/2}} - \frac{p^{3/2}}{(p+1)^2(p+2)^{1/2}} \cdot\left\{(p-1)S + (pn+1)(pn+2) + \frac{2(pn+2)^{(n)}}{n!p^n}\right\}.$$

As the expressions obtained in this section tend to be rather long, we do not write out the explicit formulae for the estimates. In any special case we merely have to substitute the values of (3.20) et seq. in the general formulae (2.1) to (2.6). In the following section we proceed to consider some special cases.

4. Special cases.

a) Rectangular distribution, $p = 1$. Although the solution of the rectangular distribution is well-known it is quoted here since it is a special case of the system solved in Section 3.

For the rectangular distribution, centered at μ with variance σ, the density function is

$$f(x) = \begin{cases} \frac{1}{2}\sqrt{3}\sigma, & \mu - \sqrt{3}\sigma \leqq x < \mu + \sqrt{3}\sigma \\ 0, & \text{otherwise.} \end{cases}$$

Using $p = 1$ in the expressions already derived yields the values for the expectation vector, variance matrix and inverse of the variance matrix as given by Lloyd (1952). The resulting estimates of μ and σ are:

$$\hat{\mu} = \tfrac{1}{2}(x_{(1)} + x_{(n)}), \qquad \hat{\sigma} = (n + 1)(x_{(n)} - x_{(1)})/2\sqrt{3}(n - 1)$$

$$\text{var}(\hat{\mu}) = 6\sigma^2/(n + 1)(n + 2) \qquad \text{var}(\hat{\sigma}) = 2\sigma^2/(n - 1)(n + 2)$$

$$\text{cov}(\hat{\mu}, \hat{\sigma}) = 0.$$

b) Right triangular distribution, $p = 2$. A less familiar example is the right triangular distribution, which is unsymmetrical. A convenient form for the density function is

$$f(x) = \begin{cases} \{(x - \mu)/\sigma + 2\sqrt{2}\}/9\sigma, & \mu - 2\sqrt{2}\sigma \leqq x < \mu + \sqrt{2}\sigma \\ 0, & \text{otherwise.} \end{cases}$$

By substituting $p = 2$ in the general expressions of Section 3, the elements of the expectation vector $\boldsymbol{\alpha}$ and the variance matrix $\mathbf{v}$ are found to be

$$\alpha_r = \{6n^{(n-r+1)}2^{n-r+1}/(2n + 1)^{(n-r+1)} - 4\}/\sqrt{2},$$

$$v_{rr} = 18\{r/(n + 1) - [n^{(n-r+1)}2^{n-r+1}/(2n + 1)^{(n-r+1)}]^2\},$$

$$v_{rs} = (s - 1)^{(s-r)}2^{s-r}v_{ss}/(2s - 1)^{(s-r)}, \qquad r < s.$$

Also, for $\mathbf{v}^{-1}$, the inverse of $\mathbf{v}$, we have

$$\mathbf{v}^{-1} = \frac{n+1}{18}\begin{bmatrix} 9 & -6 & 0 & & & 0 \\ -6 & 33/2 & -10 & 0 & & \\ 0 & -10 & 73/3 & -14 & 0 & \\ & 0 & -14 & \cdot & \cdot & 0 \\ & & 0 & \cdot & \cdot & -2(2n-1) \\ 0 & & & 0 & -2(2n-1) & (8n^2+1)/n \end{bmatrix}.$$

Thus the estimates of μ and σ, and their variances, are

$$\hat{\mu} = \frac{n-1}{3}\left[\sum \frac{x_{(r)}}{r} + 2x_{(1)} + \frac{2n+1}{n-1} x_{(n)} \sum_{r=2}^{n} \frac{1}{r}\right] \Big/ \left[n \sum \frac{1}{r} - 1\right]$$

$$\hat{\sigma} = \frac{2n-1}{6\sqrt{2}}\left[\left(\sum \frac{1}{r} + 2\right) x_{(n)} - 2x_{(1)} - \sum \frac{x_{(r)}}{r}\right] \Big/ \left[n \sum \frac{1}{r} - 1\right]$$

$$\operatorname{var}(\hat{\mu}) = \frac{2\sigma^2}{n+1}\left[\sum \frac{1}{r} + n - 2\right] \Big/ \left[n \sum \frac{1}{r} - 1\right]$$

$$\operatorname{var}(\hat{\sigma}) = \frac{\sigma^2}{4(n+1)}\left[\sum \frac{1}{r} + 4(n+1)\right] \Big/ \left[n \sum \frac{1}{r} - 1\right]$$

$$\operatorname{cov}(\hat{\mu}, \hat{\sigma}) = \frac{\sigma^2}{\sqrt{2}\,(n+1)}\left[2n - 1 - \sum \frac{1}{r}\right] \Big/ \left[n \sum \frac{1}{r} - 1\right]$$

where all summations are for r from 1 to n unless otherwise indicated.

To facilitate estimation of the parameters of a distribution of this type, the elements of the expectation vector $\boldsymbol{\alpha}$, variance matrix $\mathbf{v}$, and the coefficients of the ordered observations in the estimates $\hat{\mu}$ and $\hat{\sigma}$ have been computed for samples of size $n \leqq 10$.

Tables I and II give the values of α_r and v_{rs} (for $r \leqq s$). Values of $v_{sr} (r < s)$ may be determined by considerations of symmetry. Tables III and IV give the values of b_r and c_r, the elements of $\mathbf{b}$ and $\mathbf{c}$, respectively, where $\hat{\mu} = \mathbf{b}'\mathbf{x}$ and $\hat{\sigma} = \mathbf{c}'\mathbf{x}$. Table V gives the values of var $(\hat{\mu})/\sigma^2$, var $(\hat{\sigma})/\sigma^2$ and cov $(\hat{\mu}, \hat{\sigma})/\sigma^2$.

Perhaps in some cases it would be more natural to estimate the extremities of the distribution. Since these extremities are linear functions of the parameters μ and σ, their least square estimates will be the same linear functions of $\hat{\mu}$ and $\hat{\sigma}$.

5. Single parameter system. When a distribution is defined by a single parameter, λ, say, which is a measure of dispersion, then the density function will be of the form $f(x/\lambda)/\lambda$. If we take $Y = X/\lambda$, then Y will have a nonparametric distribution, and we may assume that the expectation vector, $\mathcal{E}(Y) = \boldsymbol{\alpha}$, and the variance matrix, var $(\mathbf{Y}) = \mathbf{v}$, are known. Then the ordered least-squares estimate, $\hat{\lambda}$, of λ is

$$\hat{\lambda} = \boldsymbol{\alpha}'\mathbf{v}^{-1}\mathbf{x}/\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha} \tag{5.1}$$

and

$$\operatorname{var}(\hat{\lambda}) = \lambda^2/\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha}. \tag{5.2}$$

The distributions of Section 3 may clearly be reduced to a system of single parameter distributions to give density functions of the type (for $p \geqq 1$):

$$f(x) = p(b - x/\lambda)^{p-1}/\lambda b^p; \qquad 0 \leqq x < b\lambda \tag{5.3}$$

where $b = (p+1)\sqrt{(p+2)/p}$.

TABLE I

Expectation vector, $\boldsymbol{\alpha}$

n = 2	−0.56569	+0.56569								
3	−0.88893	+0.08081	+0.80812							
4	−1.10443	−0.24244	+0.40406	+0.94281						
5	−1.26116	−0.47753	+0.11020	+0.59997	+1.02852					
6	−1.38172	−0.65837	−0.11585	+0.33625	+0.73183	+1.08786				
7	−1.47817	−0.80304	−0.29669	+0.12527	+0.49448	+0.82677	+1.13137			
8	−1.55759	−0.92218	−0.44561	−0.04848	+0.29902	+0.61176	+0.89844	+1.16465		
9	−1.62448	−1.02251	−0.57103	−0.19479	+0.13441	+0.43070	+0.70229	+0.95448	+1.19092	
10	−1.68181	−1.10850	−0.67852	−0.32020	−0.00667	+0.27550	+0.53416	+0.77435	+0.99952	+1.21218

TABLE II

Variance matrix, **v**

$n = 2$	0.88000	0.32000								
		0.48000								
$n = 3$	0.73837	0.35755	0.14694							
		0.53633	0.22041							
			0.27551							
$n = 4$	0.62785	0.34177	0.18721	0.08127						
		0.51265	0.28082	0.12190						
			0.35102	0.15238						
				0.17778						
$n = 5$	0.54367	0.31551	0.19439	0.11250	0.05037					
		0.47327	0.29158	0.16875	0.07556					
			0.36448	0.21094	0.09445					
				0.24610	0.11019					
					0.12397					
$n = 6$	0.47846	0.28913	0.18998	0.12368	0.07384	0.03372				
		0.43369	0.28497	0.18552	0.11075	0.05059				
			0.35621	0.23190	0.13844	0.06323				
				0.27055	0.16152	0.07377				
					0.18171	0.08299				
						0.09129				
$n = 7$	0.42680	0.26519	0.18149	0.12603	0.08464	0.05154	0.02387			
		0.39779	0.27224	0.18904	0.12696	0.07731	0.03580			
			0.34030	0.23630	0.15869	0.09664	0.04476			
				0.27568	0.18514	0.11275	0.05221			
					0.20829	0.12684	0.05874			
						0.13953	0.06462			
							0.07000			
$n = 8$	0.38498	0.24414	0.17184	0.12429	0.08903	0.06100	0.03766	0.01762		
		0.36621	0.25776	0.18644	0.13355	0.09149	0.05649	0.02643		
			0.32220	0.23304	0.16694	0.11437	0.07062	0.03304		
				0.27189	0.19476	0.13343	0.08239	0.03854		
					0.21910	0.15011	0.09269	0.04336		
						0.16512	0.10195	0.04770		
							0.11045	0.05167		
								0.05536		
$n = 9$	0.35051	0.22576	0.16221	0.12067	0.09004	0.06580	0.04570	0.02851	0.01344	
		0.33865	0.24331	0.18100	0.13506	0.09869	0.06856	0.04276	0.02016	
			0.30414	0.22625	0.16882	0.12337	0.08570	0.05345	0.02520	
				0.26396	0.19696	0.14393	0.09998	0.06236	0.02940	
					0.22158	0.16192	0.11247	0.07016	0.03308	
						0.17811	0.12372	0.07718	0.03639	
							0.13403	0.08361	0.03942	
								0.08958	0.04224	
									0.04488	
$n = 10$	0.32163	0.20972	0.15306	0.11624	0.08921	0.06790	0.05031	0.03531	0.02220	0.01053
		0.31458	0.22959	0.17435	0.13381	0.10185	0.07547	0.05296	0.03329	0.01579
			0.28699	0.21794	0.16726	0.12732	0.09433	0.06620	0.04162	0.01974
				0.25427	0.19514	0.14854	0.11006	0.07723	0.04855	0.02303
					0.21953	0.16710	0.12381	0.08689	0.05462	0.02591
						0.18381	0.13620	0.09557	0.06008	0.02850
							0.14755	0.10354	0.06509	0.03088
								0.11093	0.06974	0.03309
									0.07410	0.03515
										0.03711

TABLE III

Coefficients for estimating $\hat{\mu}$. (Vector **b***)*

n = 2	0.5000	0.5000								
3	0.4444	0.0741	0.4815							
4	0.4091	0.0682	0.0455	0.4773						
5	0.3840	0.0640	0.0427	0.0320	0.4773					
6	0.3650	0.0608	0.0406	0.0304	0.0243	0.4789				
7	0.3499	0.0583	0.0389	0.0292	0.0233	0.0194	0.4810			
8	0.3375	0.0562	0.0375	0.0281	0.0225	0.0187	0.0161	0.4834		
9	0.3271	0.0545	0.0363	0.0273	0.0218	0.0182	0.0156	0.0136	0.4857	
10	0.3181	0.0530	0.0353	0.0265	0.0212	0.0177	0.0151	0.0133	0.0118	0.4879

TABLE IV

Coefficients for estimating $\hat{\sigma}$. (Vector **c***)*

n = 2	−0.8839	+0.8839								
3	−0.5500	−0.0917	+0.6416							
4	−0.4339	−0.0723	−0.0482	+0.5544						
5	−0.3734	−0.0622	−0.0415	−0.0311	+0.5082					
6	−0.3355	−0.0559	−0.0373	−0.0280	−0.0224	+0.4790				
7	−0.3030	−0.0505	−0.0337	−0.0253	−0.0202	−0.0168	+0.4495			
8	−0.2898	−0.0483	−0.0322	−0.0241	−0.0193	−0.0161	−0.0138	+0.4436		
9	−0.2746	−0.0458	−0.0305	−0.0229	−0.0183	−0.0153	−0.0131	−0.0114	+0.4319	
10	−0.2624	−0.0437	−0.0292	−0.0219	−0.0175	−0.0146	−0.0125	−0.0109	−0.0097	+0.4225

TABLE V

Variance and covariance of estimates

n	var $(\hat{\mu})/\sigma^2$	var $(\hat{\sigma})/\sigma^2$	cov $(\hat{\mu}, \hat{\sigma})/\sigma^2$
2	0.5000	0.5625	0.1768
3	0.3148	0.2477	0.1244
4	0.2227	0.1506	0.09482
5	0.1691	0.1051	0.07599
6	0.1345	0.07938	0.06304
7	0.1107	0.06303	0.05364
8	0.09340	0.04946	0.04652
9	0.08037	0.04205	0.04097
10	0.07024	0.03641	0.03652

It is of interest to study this single parameter system, chiefly because the limiting case, as $p \to \infty$, is the exponential distribution $f(x) = e^{-x/\lambda}/\lambda$. Before considering this system, however, it is useful to consider one general property of estimates of this type. Lloyd derived necessary and sufficient conditions for the variance of the ordered least-squares estimate of the mean to attain its upper bound σ^2/n for symmetric distributions. The author of the present paper has since extended these conditions to include the unsymmetric case [5]. This would suggest that in the single parameter case it may be possible to derive an upper bound for var $(\hat{\lambda})$, using a similar approach.

We first note that the reduced ordered observations $y_{(r)}(= x_{(r)}/\lambda$, here) are a permutation of the unordered observations $y_{(r)}$, and hence

$$\mathbf{1}'\boldsymbol{\alpha} = \sum \mathcal{E}(y_{(r)}) = \mathcal{E}(\sum y_{(r)}) = \mathcal{E}\sum(y_{(r)}) = n\mathcal{E}(Y), \tag{5.4}$$

where $\mathcal{E}(Y)$ is the mean of the nonparametric parent population. A similar argument shows that

$$\mathbf{1}'\mathbf{v}\mathbf{1} = n \text{ var } (Y). \tag{5.5}$$

We also note that both $\mathbf{v}$ and $\mathbf{v}^{-1}$ are symmetric and positive definite and hence may be expressed as $\mathbf{v} = \mathbf{t}\mathbf{t}'$ and $\mathbf{v}^{-1} = (\mathbf{t}^{-1})'\mathbf{t}^{-1}$, where $\mathbf{t}$ is a lower triangular matrix. It then follows that

$$\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha} = \boldsymbol{\alpha}'(\mathbf{t}^{-1})'\mathbf{t}^{-1}\boldsymbol{\alpha} = \mathbf{h}'\mathbf{h} = \sum h_i^2,$$

say, where $\mathbf{h} = \mathbf{t}^{-1}\boldsymbol{\alpha}$. Similarly we find that

$$\mathbf{1}'\mathbf{v}\mathbf{1} = \mathbf{1}'\mathbf{t}\mathbf{t}'\mathbf{1} = \mathbf{k}'\mathbf{k} = \sum k_i^2,$$

say, where $\mathbf{k} = \mathbf{t}'\mathbf{1}$.

Now, the Cauchy-Schwarz inequality $(\sum h_i^2)(\sum k_i^2) \geqq (\sum h_i k_i)^2$ in matrix form becomes $(\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{v}\mathbf{1}) \geqq (\mathbf{1}'\boldsymbol{\alpha})^2$, or

$$\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha} \geqq n[\mathcal{E}(Y)]^2/\text{var } (Y). \tag{5.6}$$

The necessary and sufficient condition that the equality shall hold is that $k_i = qh_i$ for some constant q and for all i. In matrix form this becomes

$$\mathbf{v}\mathbf{1} = q\boldsymbol{\alpha}. \tag{5.7}$$

Premultiplying by $\mathbf{1}'$, it follows that, necessarily

$$q = \text{var } (Y)/\mathcal{E}(Y). \tag{5.8}$$

Since var $(\hat{\lambda}) = \lambda^2/\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha}$, the variance of the ordered least-squares estimate of λ has an upper bound such that

$$\text{var } (\hat{\lambda}) \leqq \lambda^2 \text{ var } (Y)/n[\mathcal{E}(Y)]^2. \tag{5.9}$$

This upper bound is attained if, and only if,

$$\mathcal{E}(Y)\cdot\mathbf{v}\mathbf{1} = \text{var } (Y)\cdot\boldsymbol{\alpha}. \tag{5.10}$$

It may also be shown by substituting (5.10) in (5.1) that if this upper bound is attained the estimate of λ is, necessarily, given by

$$\hat{\lambda} = \mathbf{1}'\mathbf{x}/n\mathcal{E}(Y), \tag{5.11}$$

which is proportional to the arithmetic mean.

We now proceed to examine the ordered least-squares estimates of λ for distributions of the type given by equation (5.3). In the notation of Section 3, we consider a variate $T = Y/b = X/\lambda b$, for which $\mathcal{E}(\mathbf{T}) = \mathbf{a}$ and var $(\mathbf{T}) = \mathbf{w}$. Then for the distribution (5.3),

$$\alpha = (p+1)\sqrt{p+2}(1 - \mathbf{J}\mathbf{a})/\sqrt{p} \tag{5.12}$$

$$\mathbf{v} = (p+1)^2(p+2)\mathbf{J}\mathbf{w}\mathbf{J}/p \tag{5.13}$$

$$\mathbf{v}^{-1} = p\mathbf{J}\mathbf{w}^{-1}\mathbf{J}/(p+1)^2(p+2) \tag{5.14}$$

where $\mathbf{J}$ is the permutation matrix

$$\mathbf{J} = \begin{bmatrix} 0 & & 1 \\ & \cdot^{\cdot^{\cdot}} & \\ 1 & & 0 \end{bmatrix}$$

Then the rth element of α is given by

$$\alpha_r = (p+1)\sqrt{(p+2}\,(1 - n^{(r)}p^r/(pn+1)^{(r)})/\sqrt{p}. \tag{5.15}$$

Also, for $r < s$, when a_r satisfies (3.9),

$$v_{sr} = v_{rs} = (p+1)^2(p+2)\{n^{(r)}p^r/(pn+2)^{(s)}(p[n-s]+1)^{r-s} - a_{n-r+1}a_{n-s+1}\}/p \tag{5.16}$$

$$v_{rr} = (p+1)^2(p+2)\{n^{(r)}p^r/(pn+2)^{(r)} - a^2_{n-r+1}\}/p. \tag{5.17}$$

Also

$$v^{-1}_{rs} = 0, \qquad |r-s| \geqq 2 \tag{5.18}$$

$$v^{-1}_{r+1,r} = v_{r,r+1} = -p(p[n-r]+1(pn+2)^{(r+1)}/p^r n^{(r)}(p+1)^2(p+2), \tag{5.19}$$

$$v^{-1}_{rr} = \frac{p\left\{\begin{matrix} p^2(2[n-r+1]^2 - 2[n-r+1]+1) \\ +\,2p(2[n-r+1]-1)+1\end{matrix}\right\}(pn+2)^{(r)}}{p^r n^{(r)}(p+1)^2(p+2)} \tag{5.20}$$

Also, with S satisfying (3.18) and (3.19),

$$\alpha'\mathbf{v}^{-1}\alpha = (p-1)S + 2(pn+2)^{(n)}/p^n n! - (pn+2) \tag{5.21}$$

$$= (pn+2)/(p-2) - 2(pn+2)^{(n)}/p^n n!(p-2), \qquad p \neq 2. \tag{5.22}$$

The rth element of the vector $\alpha'\mathbf{v}^{-1}$ is given by

$$\{\alpha'\mathbf{v}^{-1}\}_r = (p-1)(pn+2)^{(r)}\sqrt{p}/(p+1)n^{(r)}p^r\,\sqrt{p+2} + \eta_r \tag{5.23}$$

where $\eta_r = 0$ for $1 \leqq r \leqq n-1$, and

$$\eta_n = 2(pn+2)^{(n)}\,\sqrt{p}/(p+1)n!p^n\,\sqrt{p+2}. \tag{5.24}$$

From these expressions the ordered least-squares estimate of λ and its variance may be calculated.

6. Exponential distribution. Taking the limit of (5.3) as $p \to \infty$, we find $f(x) = e^{-x/\lambda}/\lambda$. The expressions given for α and $\mathbf{v}$ in Section 5 now become indeterminate. However, taking limits in (5.19) and (5.20), we obtain

$$\mathbf{v}^{-1} = \frac{1}{2}\begin{bmatrix} (2n-1)^2+1 & -2(n-1)^2 & 0 & & & & 0 \\ -2(n-1)^2 & (2n-3)^2+1 & -2(n-2)^2 & 0 & & & \\ 0 & -2(n-2)^2 & (2n-5)^2+1 & \cdot & 0 & & \\ & 0 & \cdot & \cdot & \cdot & 0 & \\ & & 0 & \cdot & 26 & -8 & 0 \\ & & & 0 & -8 & 10 & -2 \\ 0 & & & & 0 & -2 & 2 \end{bmatrix}$$

Also, $\boldsymbol{\alpha}'\mathbf{v}^{-1}\boldsymbol{\alpha} = n$, from (5.21), and $\boldsymbol{\alpha}'\mathbf{v}^{-1} = \mathbf{1}'$, from (5.23). Thus the ordered least-squares estimate of the parameter λ is the sample mean, which is also the maximum likelihood estimate.

To determine the variance matrix $\mathbf{v}$ in any particular case, the author has found it simpler to invert the matrix $\mathbf{v}^{-1}$ by triangular resolution [6] than to evaluate the integrals which give the individual elements of $\mathbf{v}$. It is also possible to use this matrix $\mathbf{v}$, obtained by inversion of $\mathbf{v}^{-1}$, to compute the elements of the expectation matrix $\boldsymbol{\alpha}$.

We note that for the reduced exponential distribution

$$f(y) = e^{-y}, \qquad \mathcal{E}(Y) = 1, \quad \text{and} \quad \text{var}\,(Y) = 1,$$

and therefore the upper bound for var $(\hat{\lambda})$ is λ^2/n. But var $(\hat{\lambda}) = \lambda^2/n$, and hence condition (5.10) operates. We have therefore that $\mathbf{v1} = \boldsymbol{\alpha}$, or that the elements of $\boldsymbol{\alpha}$ are the row sums of $\mathbf{v}$.

In this case an alternative method may be used for deriving the variance matrix. It may be noted that since increasing the sample size by unity involves adding a row and a column to the leading edges of the matrix $\mathbf{v}^{-1}$, such an increase in sample size will affect $\mathbf{v}$ in a similar way. Thus, calculating the variance matrix for any specific sample size n enables us to determine the variance matrix for a sample size $n + 1$, and successively for all sample sizes greater than or, conversely, smaller than n.

Suppose $\mathbf{v}_n$ and $\mathbf{v}_{n+1}$ are the variance matrices, with inverses $\mathbf{v}_n^{-1}$ and $\mathbf{v}_{n+1}^{-1}$, of samples of n and $n + 1$ ordered observations, respectively, and suppose $\mathbf{v}_n$ is known (by inverting $\mathbf{v}_n^{-1}$, or otherwise). It may be easily shown that $\mathbf{v}_{n+1}$ may be represented in partitioned form by

$$\mathbf{v}_{n+1} = a\left[\begin{array}{c|c} 1 & -\mathbf{h}'\mathbf{v}_n \\ \hline -\mathbf{v}_n\mathbf{h} & \mathbf{v}_n/\mathbf{a} + \mathbf{v}_n\mathbf{h}\mathbf{h}'\mathbf{v}_n \end{array}\right]$$

where $a = \{\frac{1}{2}[(2n+1)^2 + 1] - \mathbf{h}'\mathbf{v}_n\mathbf{h}\}^{-1}$ and $\mathbf{h}' = (-n^2, 0, 0, \cdots, 0)$. This follows directly from the fact that in partitioned form

$$\mathbf{v}_{n+1}^{-1} = \left(\begin{array}{c|c} \frac{1}{2}[(2n+1)^2+1] & \mathbf{h}' \\ \hline \mathbf{h} & \mathbf{v}_n^{-1} \end{array}\right).$$

From the explicit expression for the matrix of $\mathbf{v}^{-1}$, above, it may be shown that $\mathbf{1}'\mathbf{v}_n^{-1} = (n^2, 0, 0, \cdots, 0) = -\mathbf{h}'$. Post-multiplying by $\mathbf{v}_n$ gives $-\mathbf{h}'\mathbf{v}_n = \mathbf{1}'$ and hence $\mathbf{h}'\mathbf{v}_n\mathbf{h} = n^2$, giving $a = (n + 1)^{-2}$. Thus $\mathbf{v}_{n+1}$ and $\mathbf{v}_n$ are connected by the difference equation

$$\mathbf{v}_{n+1} = \frac{1}{(n+1)^2}\left[\begin{array}{c|c} 1 & \mathbf{1}' \\ \hline \mathbf{1} & (n+1)^2\mathbf{v}_n + \mathbf{11}' \end{array}\right] = \frac{\mathbf{11}'}{(n+1)^2} + \left[\begin{array}{c|c} 0 & \mathbf{0}' \\ \hline \mathbf{0} & \mathbf{v}_n \end{array}\right].$$

A similar relation connecting $\boldsymbol{\alpha}_n$ and $\boldsymbol{\alpha}_{n+1}$, the respective expectation vectors, is obtained from the fact that $\mathbf{1}'\mathbf{v} = \boldsymbol{\alpha}'$. Thus $\boldsymbol{\alpha}_{n+1}$ may be partitioned, such that

$$\boldsymbol{\alpha}_{n+1}' = \{1 \mid (n+1)\boldsymbol{\alpha}_n' + \mathbf{1}'\}/(n+1) = \mathbf{1}'/(n+1) + (0 \mid \boldsymbol{\alpha}_n),$$

which is the necessary difference equation. In the foregoing, the vectors **1** and **0** have the number of elements suitable to their context.

7. Pearson Type III distribution. Consider a variate X, whose density function is given by

$$f(x) = x^{p-1}e^{-x/\lambda}/\Gamma(p)\lambda^p \qquad 0 \leqq x < \infty. \tag{7.1}$$

If we assume p is known, this is the Pearson Type III distribution depending on a single dispersion parameter λ. Since it has the functional form discussed in Section 5, the relationships (5.1), (5.2), and (5.4) to (5.11) hold. If $Y = X/\lambda$ we have for distribution (7.1) $\mathcal{E}(Y) = p$ and var $(Y) = p$ so that from (5.9)

$$\text{var}(\hat{\lambda}) \leqq \lambda^2/np. \tag{7.2}$$

It is known from general theory that for an unbiased estimate λ^* of λ, the variance of λ^* has a lower bound given by

$$\text{var}(\lambda^*) \geqq \left[n \int (\partial \ln f/\partial\lambda)^2 f\, dx\right]^{-1}.$$

Thus for any unbiased estimate λ^* of λ,

$$\text{var}(\lambda^*) \geqq \lambda^2/np. \tag{7.3}$$

Since $\hat{\lambda}$ is an unbiased estimate of λ, both (7.2) and (7.3) hold. This can be so only if

$$\text{var}(\hat{\lambda}) = \lambda^2/np. \tag{7.4}$$

This means that the variance of the ordered least-squares estimate attains its upper bound, so that $\hat{\lambda}$ is, necessarily, given by (5.11), that is, by $\hat{\lambda} = \mathbf{1}'\mathbf{x}/np$, which is also the maximum likelihood estimate. This proves also, for the Pearson Type III distribution, that $\mathbf{v1} = \boldsymbol{\alpha}$, since (7.4) can be true only under this condition.

I am indebted to Dr. E. H. Lloyd for suggesting many improvements to the original draft of this paper.

REFERENCES

[1] H. J. GODWIN, "On the estimation of dispersion by linear systematic statistics," *Biometrika*, Vol. 36 (1949), pp. 92–100.

[2] H. J. GODWIN, "Some low moments of order statistics," *Ann. Math. Stat.*, Vol. 20 (1949), pp. 279–285.

[3] E. H. LLOYD, "Least-squares estimation of location and scale parameters using order statistics," *Biometrika*, Vol. 39 (1952), pp. 88–95.

[4] A. T. CRAIG, "A note on the best linear estimate," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 88–90.

[5] F. DOWNTON, "A note on ordered least-squares estimation," *Biometrika*, Vol. 40 (1953), pp. 457–458.

[6] L. FOX AND J. G. HAYES, "More practical methods for the inversion of matrices," *J. Roy. Stat. Soc. Ser. B*, Vol. 13 (1951), pp. 83–91.