A property of the order statistics of an exponential variate

by Trevor Williams,
Department of Mathematics,
Duke University,
Durham, North Carolina, U.S.A.

Kendall & Stuart (1961, Vol. II, pp. 87 et seq.) cite the work of Lloyd (1952) and Downton (1953) in estimating the scale and location of a random variable by forming the best linear unbiassed estimate based on its order statistics. The purpose of this note is to point out the rather curious consequences which ensue when this sort of technique is applied to the estimation of the scale of an exponential distribution of known location, a situation which arises not infrequently in practice. The result which emerges is that the information relevant to the scale is concentrated markedly in the upper portion of the sample, there being a logarithmic type of singularity such that the upper half of the sample contains 98.0% of the total information in the sample; the upper 10%, 68.9% of the information; and the upper 1%, 22.4%.

We assume that we have an ordered sample

$$y_{(1)} \leq y_{(2)} \leq \cdots \leq y_{(n)} \tag{1}$$

of an exponential variate

$$y \sim \mathrm{expl}(\sigma) , \tag{2}$$

by which we mean that the frequency function of y is

$$f(y) = \sigma^{-1} \exp(-y/\sigma) ; \tag{3}$$

and we write

$$\underline{y} \equiv \sigma \underline{z} \ , \quad \underline{z} \sim \text{expl}(1) \tag{4}$$

and

$$\underline{y}_{(r)} = \sigma \underline{z}_{(r)} \quad (\underline{r} = 1, 2, \ldots, \underline{n}) \ . \tag{5}$$

Then the column vector

$$\underset{\sim}{z} \equiv \{\underline{z}_{(n)}, \underline{z}_{(n-1)}, \cdots, \underline{z}_{(n-k+1)}\} \tag{6}$$

(we assume henceforth that $\underline{k}$ is a fixed integer between 1 and n) has, as Kendall & Stuart (op. cit., pp. 96, 97, problem 19.11) remark, a mean given by

$$\underset{\sim}{\mu} = \{ \sum_{1}^{n} \frac{1}{j}, \sum_{2}^{n} \frac{1}{j}, \ldots, \sum_{k}^{n} \frac{1}{j} \} \tag{7}$$

and a variance equal to

$$\underset{\sim}{V} = \begin{pmatrix} \sum_{1}^{n} \frac{1}{j^2} & \sum_{2}^{n} \frac{1}{j^2} & \cdots & \sum_{k}^{n} \frac{1}{j^2} \\ \sum_{2}^{n} \frac{1}{j^2} & \sum_{2}^{n} \frac{1}{j^2} & \cdots & \sum_{k}^{n} \frac{1}{j^2} \\ \vdots & \vdots & & \vdots \\ \sum_{k}^{n} \frac{1}{j^2} & \sum_{k}^{n} \frac{1}{j^2} & \cdots & \sum_{k}^{n} \frac{1}{j^2} \end{pmatrix} . \tag{8}$$

These results are easy consequences of the fact that the ordered sample (5) of the $\underline{z}$'s may also be generated by the equivalent sampling scheme of choosing

$$\underline{u}_1 \equiv \underline{z}_{(1)} \sim \text{expl}(1/\underline{n}) \tag{9}$$

and

$$\underline{u}_j \equiv \underline{z}_{(j)} - \underline{z}_{(j-1)} \sim \text{expl}[1/(\underline{n}-\underline{j}+1)] \quad (\underline{j} = 2, 3, \ldots, \underline{n}) , \tag{9'}$$

where the $\underline{u}$'s are all mutually independent; for, this makes the joint frequency function

$$\begin{aligned} & f(\underline{z}_{(1)}, \cdots, \underline{z}_{(n)}) \\ & \quad = \underline{n} \exp[-\underline{n}\underline{z}_{(1)}] \prod_{j=1}^{n-1} (\underline{n}-\underline{j}) \exp\{-(\underline{n}-\underline{j})[\underline{z}_{(j+1)} - \underline{z}_{(j)}]\} \\ & \quad = \underline{n}! \prod_{1}^{n} \exp[-\underline{z}_{(j)}] , \end{aligned} \tag{10}$$

which is clearly what it should be.

Corresponding to (6), we write the top $\underline{k}$ members of the ordered sample of the original variate $\underline{y}$ as the column vector

$$\underset{\sim}{\underline{Y}} \equiv \{\underline{Y}_{(n)}, \underline{Y}_{(n-1)}, \cdots, \underline{Y}_{(n-k+1)}\} \tag{11}$$

and standard results from linear estimation then state that for the least squares estimate we have

$$\hat{\sigma} = \underline{S}^{-1} \underset{\sim}{\alpha}' \underset{\sim}{\underline{V}}^{-1} \underset{\sim}{\underline{Y}} \quad \text{and} \quad \text{var}\, \hat{\sigma} = \sigma^2 \underline{S} , \tag{12}$$

where

$$\underline{S} = \underset{\sim}{\alpha}' \underset{\sim}{\underline{V}}^{-1} \underset{\sim}{\alpha} , \tag{13}$$

a scalar, since we have only a single parameter here.

Now, one may readily check that

$$\Sigma^{-1} = \begin{pmatrix} 1^2 & -1^2 & 0 & \cdots & 0 & 0 \\ -1^2 & 1^2+2^2 & -2^2 & \cdots & 0 & 0 \\ 0 & -2^2 & 2^2+3^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & (\underline{k}-2)^2+(\underline{k}-1)^2 & -(\underline{k}-1)^2 \\ 0 & 0 & 0 & \cdots & -(\underline{k}-1)^2 & (\underline{k}-1)^2+\dfrac{1}{\sum_{\underline{k}}^{n} \frac{1}{j^2}} \end{pmatrix} \tag{14}$$

and from this it follows that

$$\xi'\Sigma^{-1} = (1, 1, \ldots, 1, -\underline{k}+1+\frac{\sum_{\underline{k}}^{n} \frac{1}{j}}{\sum_{\underline{k}}^{n} \frac{1}{j^2}}) \tag{15}$$

and

$$\underline{S} = \underline{k}-1+\frac{(\sum_{\underline{k}}^{n} \frac{1}{j})^2}{\sum_{\underline{k}}^{n} \frac{1}{j^2}} . \tag{16}$$

Thus, the best linear unbiassed estimate of the scale factor, based on the top $\underline{k}$ members of the sample, is given by

$$\hat{\sigma} = \frac{1}{\underline{S}}\{\sum_{n-\underline{k}+1}^{n} Y_{(i)} + \left[\frac{\sum_{\underline{k}}^{n} \frac{1}{j}}{\sum_{\underline{k}}^{n} \frac{1}{j^2}} - \underline{k}\right] Y_{(n-\underline{k}+1)}\} , \tag{17}$$

a rather curious expression, since, for moderate values of k, the second principal term within the curly brackets completely swamps the first. When $\underline{k} = \underline{n}$, (16) yields $\underline{S} = \underline{n}$ and (17) reduces to the arithmetic mean, which is, of course, the maximum likelihood estimator based on the entire sample. Therefore, by the second half of (12), the relative efficiency of (17) is $\underline{S}/\underline{n}$; but, writing $\underline{k} = \theta\underline{n}$ $(0 < \theta < 1)$, we have

$$\sum_{k}^{n} \frac{1}{j} \sim \int_{\theta n}^{n} \frac{dx}{x} = -\log \theta \quad (n \to \infty) \tag{18}$$

and

$$\sum_{k}^{n} \frac{1}{j^2} \sim \int_{\theta n}^{n} \frac{dx}{x^2} = (\frac{1}{\theta} - 1)\frac{1}{n} \quad (n \to \infty) . \tag{19}$$

Hence, (16) yields

$$\underline{S}/\underline{n} \sim \{1 + \frac{(\log \theta)^2}{1 - \theta}\}\theta \quad (k = \theta n , \quad n \to \infty) , \tag{20}$$

the result alluded to at the outset.

An obvious application of this result is to the estimation of frequency functions which are linear combinations of distinct exponentials:

$$\underline{f}(y) = \sum_{i=1}^{\nu} \underline{A}_i \exp(-y/\sigma_i) \quad (\sigma_1 > \sigma_2 > \ldots > \sigma_\nu) . \tag{21}$$

Since

$$\underline{f}(y) \sim \underline{A}_1 \exp(-y/\sigma_1) \quad (y \to \infty) , \tag{22}$$

in a sufficiently large sample size the upper extreme values will effectively be governed by (22), and we may use (17) to estimate the dominant "eigenvalue," σ_1, provided $\underline{k}$ is not taken to be large enough to introduce contributions from the smaller σ's in (21), a delicate matter.

When it is required to estimate location as well as scale, the behaviour we have been describing no longer obtains.

ACKNOWLEDGMENT

This work was done under grant number NGR 34-001-005 from the National Aeronautics and Space Administration of the U.S. Government.

REFERENCES

Downton, F. (1953). A note on ordered least-squares estimation. Biometrika, 40, 457.

Kendall, M. G. and Stuart, A. (1961). The Advanced Theory of Statistics. Griffin, London.

Lloyd, E. H. (1952). Least-squares estimation of location and scale parameters using order statistics. Biometrika, 39, 88.