WORKING PAPER

SCHOOL OF INDUSTRIAL MANAGEMENT

Notes on Paretian Distribution Theory

G. M. Kaufman

School of Industrial Management

Mass. Institute of Technology

November 1962 27-63

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE 39, MASSACHUSETTS

Notes on Paretian Distribution Theory

G. M. Kaufman

School of Industrial Management

Mass. Institute of Technology

November 1962 27-63

Notes on Paretian Distribution Theory

G. M. Kaufman

School of Industrial Management

Mass. Institute of Technology

Nov. 1962

Mandelbrot (see [1] for example) has suggested that many real-world random processess--first differences of stock market prices over time, the size distribution of income, etc.--are particular instances of processes characterized by "stable non-Gaussian" probability laws called "Pareto-Levy" laws (cf: M. Loève [2]).

One particular member of the class of probability laws stable in the Pareto-Levy sense is the (strong) Pareto process. It is defined as a stochastic process generating independent random variables $\tilde{x}_1, \ldots, \tilde{x}_i \ldots$ with identical Pareto density functions

$$f_\alpha(x_i|\alpha, x_o) = \begin{cases} 0 & \text{if} \quad x_i \le x_o \\ \alpha x_o^\alpha x_i^{-(\alpha+1)} & \text{if} \quad x_i > x_o \end{cases}, \quad i=1,2,\ldots \tag{1}$$

where $\alpha > 0$ and $x_o > 0$ are parameters of the above density function. The density function (1) possesses properties which make life difficult for the Bayesian: when $1 < \alpha \le 2$ the variance of f_α is infinite and when $0 < \alpha \le 1$, both the mean and variance of f_α are infinite.

Our interest here is in the following question: How do we derive distribution-theoretic results needed for a "Bayesian" analysis of decision problems where the consequences of adopting one of a set of available (terminal) acts depends on the true underlying value of α, and where information

about α can be obtained by observation of values of $\tilde{x}_1, \ldots, \tilde{x}_i, \ldots$ when the $\tilde{x}_i$'s are independent random variables, identically distributed according to (1)?

We follow the pattern of [3] (ASDT) and present the following:

1. Definition of and properties of Pareto density function.

2. Definition and characterization of the Pareto process and its properties.

3. Likelihood of a sample and the sufficient statistic.

4. Conjugate prior distribution of $\tilde{\alpha}$ and binary operation for going from the prior parameter and sample statistic to the posterior parameter.

5. Conditional distribution of the sufficient statistic for a given value of $\tilde{\alpha}$ and a fixed size sample from the process.

6. Unconditional sampling distribution of the sufficient statistic.

7. Some facts about the (prior) distribution of the mean of the process when α is not known with certainty.

A subsequent note will discuss further aspects of analysis of decision problems under uncertainty, when the underlying data generating process is of the strong Pareto type.

1. <u>The Pareto Function</u>

The Pareto density function is defined by (1). The first k moments of f_α about the origin are

$$E(\tilde{x}^k) = \begin{cases} +\infty & \text{if} \quad k \geq \alpha \\ \dfrac{\alpha x_o^k}{\alpha - k} & \text{if} \quad k < \alpha \end{cases} \qquad (1.2)$$

Proof: To prove that f_α is a density function, note that it is positive for all $x \in (x_o, \infty)$ and that

$$\int_{x_o}^{\infty} f_\alpha(z|\alpha, x_o)\,dz = \alpha x_o^\alpha \int_{x_o}^{\infty} z^{-(\alpha+1)}\,dz$$

$$= \alpha x_o^\alpha \left[-\lim_{z\to\infty} \frac{1}{\alpha z^\alpha} + \frac{1}{\alpha x_o^\alpha}\right]$$

$$= (\alpha x_o^\alpha)(\alpha x_o^\alpha)^{-1} = 1 .$$

To prove (1.2), we have

$$E(\tilde{x}^k) = \int_{x_o}^{\infty} z^k f_\alpha(z|\alpha, x_o)\,dz$$

$$= \alpha x_o^\alpha \int_{x_o}^{\infty} z^{k-\alpha-1}\,dz$$

$$= \alpha x_o^\alpha \left[\lim_{z\to\infty} \frac{z^{k-\alpha}}{k-\alpha} - \frac{x_o^{k-\alpha}}{k-\alpha}\right] .$$

Thus if $k \geq \alpha$, $E(\tilde{x}^k) = +\infty$, and if $k < \alpha$,

$$E(\tilde{x}^k) = \frac{\alpha x_o^k}{\alpha - k} .$$

The incomplete kth moment about the origin follows directly and is, for $z_o \geq x_o$,

$$E_{z_o}^{\infty}(\tilde{x}^k) = \begin{cases} +\infty & \text{if} \quad k \geq \alpha \\ \dfrac{\alpha x_o^{\alpha} z_o^{k-\alpha}}{\alpha - k} & \text{if} \quad k < \alpha \end{cases} \quad . \tag{1.3}$$

The cumulative distribution function is

$$F_{\alpha}(x|\alpha, x_o) = 1 - \left(\frac{x}{x_o}\right)^{-\alpha} \quad . \tag{1.4}$$

2. The Pareto Process

The Pareto process is defined as a stochastic process generating independent random variables $\tilde{x}_1, \ldots, \tilde{x}_i, \ldots$ with identical Pareto density functions

$$f_{\alpha}(x_i|\alpha, x_o) = \alpha x_o^{\alpha} x_i^{-(\alpha+1)} , \qquad \begin{array}{l} \alpha > 0 , \\ \cdot > 0 , \\ x_i > x_o , \\ i=1,2,\ldots . \end{array} \tag{2.1}$$

From (1.2) we may obtain the mean and variance of the process:

$$E(\tilde{x}|\alpha, x_o) = \begin{cases} +\infty & \text{if} \quad \alpha \leq 1 \\ \dfrac{\alpha x_o}{\alpha - 1} & \text{if} \quad \alpha > 1 \end{cases} \tag{2.2}$$

and

$$V(\tilde{x}|\alpha, x_o) = \begin{cases} +\infty & \text{if} \quad \alpha \leq 2 \\ \dfrac{\alpha x_o^2}{(\alpha-2)(\alpha-1)^2} & \text{if} \quad \alpha > 2 \end{cases} \tag{2.3}$$

Proof: Formula (2.2) follows from setting k=1 in (1.2). Formula 2.3 is determined by use of the fact that

$$V(\tilde{x}|\alpha,x_o) = E(\tilde{x}^2|\alpha,x_o) - E^2(\tilde{x}|\alpha,x_o) \text{ , and}$$

that

$$V(\tilde{x}|\alpha,x_o) \equiv \int_{x_o}^{\infty} [z - E(\tilde{x}|\alpha,x_o)]^2 f_\alpha(z|\alpha,x_o)dz \; .$$

By (1.2) if $0 \leq \alpha \leq 1$ then $E(\tilde{x}|\alpha,x_o) = +\infty$. Thus for all values of the integrand z, $[z - E(\tilde{x}|\alpha,x_o)]^2$ is unbounded and hence $V(\tilde{x}|\alpha,x_o) = +\infty$. If $1 \leq \alpha \leq 2$ then $V(\tilde{x}|\alpha,x_o) = +\infty$ for

$$V(\tilde{x}|\alpha,x_o) = E(\tilde{x}^2|\alpha,x_o) - E^2(\tilde{x}|\alpha,x_o) \text{ ,}$$

and

$$E(\tilde{x}^2|\alpha,x_o) = +\infty \quad \text{and} \quad E^2(\tilde{x}|\alpha,x_o) < \infty$$

by (1.2). When $\alpha > 2$, by (1.2) again

$$= \frac{\alpha x_o^2}{\alpha-2} - \left(\frac{\alpha x_o}{\alpha-1}\right)^2$$

$$= \frac{\alpha x_o^2}{(\alpha-2)(\alpha-1)^2} \quad .$$

A "unique and important" property of a Poisson process is that if the independent random variables

$$\tilde{u}_1, \ldots, \tilde{u}_i, \ldots$$

are generated by a Poisson process, then

$$P(\tilde{u}_i > \xi + u | \tilde{u}_i > \xi) = P(\tilde{u}_i > u), \quad \begin{matrix} i=1,2,\ldots \\ \xi > 0 \end{matrix} \tag{2.4}$$

This is interpretable as "independence of past history," and is a <u>unique</u> property of the Poisson process among all processes generating continuous random variables $\tilde{x} \geq 0$. (See [1] for discussion of this property). By analogy, it is easy to show that for independent random variables $\tilde{x}_1, \ldots, \tilde{x}_i, \ldots$ generated according to (1) and $\xi' \equiv e^{\xi} > 1$,

$$P(\tilde{x}_i > \xi' x | \tilde{x}_i > \xi') = P(\tilde{x}_i > x) = (x/x_o)^{-\alpha}, \quad \begin{matrix} i=1,2,\ldots \\ x > x_o \end{matrix}, \tag{2.5}$$

and that the (strong) Pareto process is the only process among all processes generating continuous random variables $\tilde{x}_i > x_o > 0$ to possess this property.

<u>Proof</u>: Make the integrand transform $u_i = \log (x_i/x_o)$ in (1) and observe that (2.5) implies (2.4) and conversely since this transform is one to one from x_i to u_i, and from α_i to x_i, since $x_o e^{u_i} = x_i$.

3. Independent Pareto Process When x_o is Known

3.1 Likelihood of a Sample When x_o is Known

The likelihood that an independent Pareto process will generate r successive values $x_1, x_2 \ldots x_r$ is the product of their individual likelihoods (1):

$$(\alpha x_o^{\alpha})^r \left[\prod_{i=1}^{r} x_i\right]^{-(\alpha+1)} \quad . \tag{3.1}$$

If we define the statistic

$$t = [\prod_{i=1}^{r} x_i] \quad , \tag{3.2}$$

we may write (3.1) as

$$(\alpha x_o^{\alpha})^r \; t^{-(\alpha+1)} \quad . \tag{3.1)'$$

Alternatively, (3.1) may be written as

$$\xi \alpha^r \; e^{-\lambda\alpha} \tag{3.3)''$$

where

$$\xi \equiv 1/t \qquad \text{and} \qquad \lambda \equiv \log_e(t/x_o^{\,r}) \; . \tag{3.4}$$

Since ξ in no way depends on the unknown parameter α, it is clear that

$$\alpha^r e^{-\lambda\alpha} \tag{3.5}$$

is the kernel of the likelihood. If the sampling process is non-informative then clearly (r,t) is a sufficient statistic when x_o is known.

3.2. Conjugate Distribution of $\tilde{\alpha}$

When x_o is known but $\tilde{\alpha}$ is regarded as a random variable, the most convenient distribution of $\tilde{\alpha}$ is a gamma-1 distribution which may be written as

$$f_{\gamma 1}(\alpha | r,t) = \frac{\lambda}{\Gamma(r)} (\lambda\alpha)^{r-1} e^{-\lambda\alpha}, \qquad \lambda > 0 \; , \; \alpha > 0 \; , \; r > 0 \; , \tag{3.6}$$

where $\lambda = \log_e(t/x_o^{\,r})$. If such a distribution with parameter (r',t') has been assigned to $\tilde{\alpha}$ and if then a sample from the process yields a

sufficient statistic (r,t) the posterior distribution of $\tilde{\alpha}$ will be gamma-1 with parameter (r'',t'') where

$$r'' \equiv r' + r \quad , \quad t'' = t't \quad , \tag{3.7a}$$

or alternatively (r'',λ'') where

$$\lambda'' = \lambda' + \lambda \quad . \tag{3.7b}$$

<u>Proof</u>: Formula (3.6) follows directly from the discussion of sufficiency. Formulas (3.7) follow from Bayes Theorem; i.e. the posterior density of $\tilde{\alpha}$ is proportional to the product of the kernel of (3.6), the prior density of $\tilde{\alpha}$ and the kernel (3.5) of the sample likelihood:

$$D''(\alpha|r',t';r,t) \sim \alpha^{r} e^{-\lambda\alpha} \cdot \alpha^{r'-1} e^{-\lambda\alpha}$$

$$= \alpha^{r+r'-1} e^{-\alpha(\lambda+\lambda')}$$

where

$$\lambda = \log_e(t/x_o^{\,r}) \quad \text{and} \quad \lambda' = \log_e(t'/x_o^{\,r'})$$

so that if we define $r'' = r' + r$ and

$$\lambda'' = \lambda + \lambda' = \log_e \left[\left(\frac{t}{x_o^{\,r}}\right)\left(\frac{t'}{x_o^{\,r'}}\right)\right]$$

$$= \log_e (t''/x_o^{\,r''}) \quad ,$$

the posterior density has a kernel of the form (3.5).

The mean, variance, and partial moments of the distribution of $\tilde{\alpha}$ are from ASDT:

$$E(\tilde{\alpha}|r,t) \equiv \bar{\alpha} = \frac{r}{t} \; , \tag{3.7}$$

$$E_o^{\alpha}(\tilde{\alpha}|r,t) \equiv \bar{\alpha}\, F_{\gamma*}(\alpha t|r+1) \quad , \tag{3.8}$$

$$V(\tilde{\alpha}|r,t) = \frac{r}{t^2} \quad . \tag{3.9}$$

In ASDT it is shown that the linear (in α) loss integrals are

$$L_\ell(\alpha) \equiv \int_0^{\alpha} (\alpha - z)\, f_{\gamma 1}(z|r,t)dz =$$

$$= \alpha\, F_{\gamma*}(\alpha t|r) - \bar{\alpha}\, F_{\gamma*}(\alpha t|r+1) \tag{3.10a}$$

and

$$L_r(\alpha) \equiv \int_{\alpha}^{\infty} (z - \alpha)\, f_{\gamma 1}(z|r,t)dz$$

$$= \bar{\alpha}\, G_{\gamma*}(\alpha t|r+1) - \alpha\, G_{\gamma*}(\alpha t|r) \tag{3.10b}$$

although these are <u>not</u> loss integrals in which we will be interested subsequently.

3.3 Distribution of the Mean $\tilde{\bar{a}}$

If $\alpha > 1$ then the mean

$$\bar{a} \equiv \frac{\alpha x_o}{\alpha - 1} \quad , \qquad \text{for } \alpha > 1 \quad ,$$

of

$$\tilde{x}|\alpha \quad \sim \quad f_\alpha(x|\alpha, x_o)$$

exists by (1.2). If α is not known with certainty then provided that we assume that

$$\tilde{\alpha} \sim f_{\gamma 1}(\alpha|r,t) \quad ,$$

it follows that for $\tilde{\alpha} > 1$, (conditional on $\tilde{\alpha} > 1$),

$$\tilde{\bar{a}} \sim p^{-1} f_{\gamma 1} (h(\bar{a})|r,t) \frac{x_o^2}{(\bar{a}-x_o)^2} \quad , \quad \bar{a} > x_o \quad , \qquad (3.11)$$

where

$$h(\bar{a}) \equiv \frac{\bar{a}}{\bar{a}-x_o} \quad \text{and} \quad p \equiv G_{\gamma 1}(1|r,t) \quad .$$

<u>Proof</u>: Since $\bar{a} = \alpha x_o/(\alpha-1)$ when $\alpha > 1$,

$$\alpha \equiv h(\bar{a}) = \bar{a}/(\bar{a}-x_o) \quad .$$

The function $h(\bar{a})$ is continuous and monotonic decreasing for $\bar{a} > \max\{1,x_o\}$, so that

$$\tilde{\bar{a}}|\tilde{\alpha} > 1 \sim p^{-1} f_{\gamma 1} (h(\bar{a})|r,t) \left|\frac{dh(\bar{a})}{d\bar{a}}\right| \quad , \bar{a} > x_o \ ;$$

as

$$\frac{dh(\bar{a})}{d\bar{a}} = \frac{x_o^2}{(\bar{a}-x_o)^2} \quad ,$$

(3.11) follows directly.

Of more interest than the loss integrals (3.10) are loss integrals linear in $\bar{a}$ -- which, as we would expect, have certain undesirable properties if $\tilde{\alpha} < 1$ with positive probability. However, we can make a step in the direction of doing (prior) terminal analysis of a two action linear loss decision problem if we assume that $\tilde{\alpha} > 1$ with probability 1. In order to proceed in this direction we will need to know the conditional expectations $E(\tilde{\bar{a}}|\tilde{\alpha} \leq \epsilon)$ and $E(\tilde{\bar{a}}|\tilde{\alpha} \geq \epsilon)$ for values of ϵ

between 0 and $+\infty$. We show that

$$\bar{m}_\epsilon \equiv E(\tilde{\bar{a}}|\tilde{\alpha} \geq \epsilon) = \begin{cases} +\infty & \text{for } \epsilon \leq 1 \\ \dfrac{x_o e^{-\lambda}}{p\Gamma(r)} \left[\left\{\sum_{i=1}^{r} \Gamma(i)\binom{r}{i}\lambda^{r-i} G_{\gamma *}(\epsilon-1|i,1)\right\} + 0_o(\epsilon)\right] & \text{for } \epsilon > 1 , \end{cases}$$

where

$$0_o(\epsilon) \equiv \int^{\infty} y^{-1}e^{-y}dy \quad , \quad \lambda \equiv \log_e(t\, x_o^{\,r}) \text{ and } p \equiv G_{\gamma 1}(\epsilon|r,t) .$$

Proof: Formula (3.10) follows directly from (1.2). Now suppose that

$$\tilde{\alpha} \sim f_{\gamma 1}(\alpha|r,t) = \frac{\lambda}{\Gamma(r)} (\lambda\alpha)^{r-1} e^{-\lambda\alpha} , \qquad \begin{matrix} \lambda > 0 \\ r > 0 \end{matrix}$$

Then for $\epsilon > 1$,

$$E(\tilde{\bar{a}}|\epsilon > 1) = \frac{x_o}{p} \int_{\epsilon}^{\infty} \left(\frac{z}{z-1}\right) f_{\gamma 1}(z|r,t)\, dz$$

$$= \frac{x_o}{p\lambda\Gamma(r)} \int_{\epsilon}^{\infty} (z-1)^{-1} e^{-\lambda z}(\lambda z)^r d(\lambda z) .$$

Letting $u = z-1$, we may write the above as

$$\frac{x_o e^{-\lambda}}{p\lambda\Gamma(r)} \int_{\epsilon-1}^{\infty} y^{-1}e^{-y} [\lambda u + \lambda]^r \, d(\lambda u) ,$$

or letting $y = \lambda u$, as

$$\frac{x_o e^{-\lambda}}{p\Gamma(r)} \int_{\epsilon=1}^{\infty} y^{-1} e^{-y} [\sum_{i=0}^{r} \binom{r}{i} y^i \lambda^{r-i}] \, dy$$

$$= \frac{x_o e^{-\lambda}}{p\Gamma(r)} \sum_{i=0}^{r} \binom{r}{i} \lambda^{r-i} \int_{\epsilon-1}^{\infty} y^{i-1} e^{-y} dy \quad .$$

Since

$$G_{\gamma*}(\epsilon-1|i,1) = \frac{1}{\Gamma(i)} \int_{\epsilon-1}^{\infty} y^{i-1} e^{-y} dy \quad , \quad i > 0 \quad ,$$

we may write

$$E(\tilde{\bar{a}}|\tilde{\alpha} \geq \epsilon) = \frac{x_o e^{-\lambda}}{p\Gamma(r)} \{\sum_{i=1}^{r} \Gamma(i) \binom{r}{i} \lambda^{r-i} G_{\gamma*}(\epsilon|i,1) + O_o(\epsilon)\}$$

where

$$O_o(\epsilon) \equiv \int_{\epsilon-1}^{\infty} y^{-1} e^{-y} dy \quad .$$

For any $\epsilon > 1$, the integral immediately above is bounded, and is unbounded for $\epsilon = 1$.

4. Sampling Distribution and Preposterior Analysis With Fixed Sample Size r

We assume that a sample of size r is taken from an independent Pareto process, and that the statistic $\tilde{t}$ is left to chance. It is also assumed that x_o is known and fixed and that the parameter $\tilde{\alpha}$ has a gamma-1 distribution of the type (3.6). We need the following theorem in the sequel:

4.4 Convolution of r Pareto Density Functions

We define the (multiplicative)convolution g* of any r density functions,

$$g^* = f_1 * f_2 *\ldots* f_r \quad , \tag{4.1}$$

by

$$g^*(t) = \int_{R_t} f_1(z_1)\ f_2(z_2)\ldots f(z_r)dA \tag{4.2}$$

where R_t is the r-1 dimensional hypersurface

$$\prod_{i=1}^{r} z_i = t \quad . \tag{4.3}$$

Theorem: The convolution g* of $r \geq 1$ Pareto densities, each with parameter $(\alpha,1)$ is

$$g^*(t|\alpha,1) = \frac{1}{\Gamma(r)}\ \alpha^r t^{-(\alpha+1)}(\log_e t)^{r-1} \quad , \quad \begin{array}{l} r \geq 1 \ , \\ t > 1 \ , \\ \alpha > 0 \ . \end{array} \tag{4.4}$$

Proof: We prove this by showing that if

$$\tilde{x}_i|\alpha \sim f_\alpha(x|\alpha,x_o) \quad , \quad i=1,2,\ldots,r$$

and we scale all $\tilde{x}_i$ into units of x_o, then g* may be represented as an (additive) convolution g of gamma-1 densities. We then show that the fact that g is a gamma-1 density(ASDT, p. 224) implies that g* is as stated above.

Define $\tilde{z}_i = \tilde{x}_i/x_o$ and $e^{\tilde{y}_i} = \tilde{z}_i$, so that

$$\tilde{z}_i \sim f_\alpha(z_i|\alpha,1)dz_i = az_i^{-(a+1)}dz_i \quad , \quad \begin{array}{l} z_i > 1 \ , \\ \alpha > 0 \ , \\ i=1,2,\ldots r \ , \end{array} \tag{4.5}$$

or

$$\tilde{y}_i \sim f_e(y_i|\alpha)dy_i = \alpha e^{-\alpha y_i} dy_i \quad , \quad \begin{array}{l} \alpha > 0 \;, \\ y_i > 0 \;, \\ i=1,2,\ldots r \;, \end{array} \tag{4.6}$$

or letting $u_i = \alpha y_i$,

$$u_i \sim f_e(u_i|1)du_i = e^{-u_i} du_i \quad , \quad \begin{array}{l} u_i > 0 \;, \\ i=1,2,\ldots r \;. \end{array} \tag{4.7}$$

By the Theorem on page 224 of ASDT, the convolution g of the r density functions (4.7) is a standardized gamma-1 density with parameter r:

$$\tilde{v} \equiv \sum_{i=1}^{r} \tilde{u}_i \sim f_{\gamma *}(v|r) = \frac{1}{\Gamma(r)} e^{-v} v^{r-1} \;. \tag{4.8}$$

If we define

$$\tilde{v} \equiv \sum_{i=1}^{r} \tilde{u}_i = \sum_{i=1}^{r} \log_e \tilde{z}_i^{\alpha} = \log_e \prod_{i=1}^{r} \tilde{z}_i^{\alpha} = \log_e \tilde{t}^{\alpha} \;,$$

since $\alpha > 0$, $\log t^{\alpha}$ is a monotonic increasing function of t. Hence (4.8) implies that

$$g^*(t) = f_{\gamma *}(\log_e t^{\alpha}) \left| \frac{d}{dt} \log_e t^{\alpha} \right|$$

$$= \frac{1}{\Gamma(r)} e^{-[\log t^{\alpha}]} [\log t^{\alpha}]^{r-1} \cdot (\alpha t^{-1})$$

$$= \frac{1}{\Gamma(r)} \alpha^r t^{-(\alpha+1)} [\log_e t]^{r-1} \quad , \quad \begin{array}{l} r > 0 \;, \\ t > 1 \;. \end{array}$$

It follows from the aboved stated theorem that the convolution g* of r Pareto densities each with parameter (α, x_o) is

$$g^*(t) = \frac{1}{\Gamma(r)} \alpha^r (x_o^{\,r})^{\alpha+1} t^{-(\alpha+1)} [\log_e t - \tfrac{k}{\alpha}]^{r-1} \quad , \quad \begin{array}{l} x_o > 0 \;, \\ t > x_o^{\,r} \;, \\ r > 0 \;. \end{array} \tag{4.9}$$

where $t \equiv \prod_{i=1}^{r} x_i$ and $k \equiv r \log_e x_o$.

4.2 Conditional Distribution of $\tilde{t}|\alpha$

Using the results of section 4.1, the conditional probability given α and x_o that the product of r identically distributed Pareto random variables will have a value t may be written as

$$D(t|\alpha,x_o;r) \equiv f_t(t|\alpha,x_o;r)$$

$$= g^*(t) \quad . \tag{4.9}$$

4.3 Unconditional Distribution of $\tilde{t}$

If a sample of size r is drawn from a Pareto process with unknown parameter $\tilde{\alpha}$ regarded as a gamma-1 random variable distributed with parameter (r',t') as shown in (3.6) and all sample observations are scaled into units of x_o, then

$$D(t|x_o=1,r',t';r) = \frac{(\lambda')^{r'}}{B(r,r')} \cdot \frac{(\log_e t)^{r-1}}{t[(\log_e t)+\lambda']^{r''}} \quad , \tag{4.11a}$$

$$t > 1 \quad ,$$
$$r,r',\lambda > 0 \quad ,$$

where

$$\lambda' \equiv \log_e t' \quad , \quad B(r,r') \equiv \frac{\Gamma(r+r')}{\Gamma(r)\,\Gamma(r')} \quad . \tag{4.11b}$$

Alternatively, the unconditional distribution of the sufficient statistic $\tilde{y} \equiv \log_e \tilde{t}$ is inverted beta-2:

$$D(y|x_o=1,r',t';r) = f_{i\beta2}(y|r,r',\lambda')$$

$$= \frac{(\lambda')^{r'}}{B(r,r')} \cdot \frac{y^{r-1}}{(y+\lambda')^{r''}} \quad , \tag{4.12}$$

$$y > 0 \quad ,$$
$$r,r',\lambda' > 0 \quad .$$

Proof: To prove (4.11a) note that from (4.4) and (3.6)

$$D(t|x_o=1,r',t';r) = \int_0^\infty g*(t|1,\alpha)\ f_{\gamma 1}(\alpha|r',t')d\alpha$$

$$= \frac{(\log_e t)^{r-1}(\lambda')^{r'}}{t\Gamma(r)\Gamma(r')} \int_0^\infty \alpha^{r''-1} e^{-\lambda''\alpha} d\alpha$$

$$= \frac{(\lambda')^{r'}}{B(r,r')} \cdot \frac{(\log_e t)^{r-1}}{t[(\log_e t) + \lambda']^{r''}} \quad , \quad t > 1 \ .$$

Formula (4.12) follows by making the integrand substitution $y = \log_e t$ in (4.11a).

---

[1] Mandelbrot, B., "The Pareto-Levy" Law and the Distribution of Income," International Economic Review (May 1960, Vol. I, No. 2).

[2] Loève, M. Probability Theory (Van Nostrand, 1960, 2nd ed.) pp. 326-331.

[3] H. Raiffa and R. O. Schlaifer, Applied Statistical Decision Theory (Division of Research, Harvard Graduate School of Business Administration, 1961)