

Stationary Proportional Hazard Processes via Complementary Power Function Distribution Processes

B. G. Manjunath ^a, Barry C. Arnold ^b, and S. Sachdeva^c

^a*Department of Statistics, Rajiv Gandhi University, Arunachal Pradesh, India.*

^b*Department of Statistics, University of California, Riverside, USA.*

^c*School of Mathematics and Statistics, University of Hyderabad, Hyderabad, India.*

ARTICLE HISTORY

Compiled September 8, 2024

Received 02 December 2023; Accepted 14 August 2024

ABSTRACT

In the following, we introduce new proportional hazard (P.H.) processes, which are derived by a marginal transformation applied to complementary power function distribution (CPFD) processes. Also, we introduce two new Pareto processes, which are derived from the proportional hazard family. We discuss the distributional features of such processes, explore inferential aspects, and include an example of applications of the new processes to real-life data.

KEYWORDS

proportional hazard process; power function distribution process; complementary power function distribution process; pareto process, stationarity; markovian property

1. Introduction

An important tool that will be used in constructing proportional hazard processes is a distribution that we will call the complementary power function distribution (CPFD). The CPFD processes are the perfect tools for developing stationary proportional hazard processes, which will be of interest since an assumption of proportional hazards is frequently deemed to be plausible in many applications. Another important specific case of a proportional hazard family of distributions is given by the classical Pareto distribution, which is used in the study of inequality in income distribution. In the following sections, we discuss the distributional features of CPFD processes and the Pareto processes, explore inferential aspects, and include an example of the application of such processes to real-life data.

2. Proportional reversed hazard processes

It is well known that proportional reversed hazard (PRH) processes can be viewed as being obtained by marginal transformations applied to a particular PFD process. In the following, we define two PFD processes introduced in Kundu [4] (called Kundu process) and Arnold et al. [2] (called A-M process).

2.1. PFD-processes

We begin by recalling the definition of the power function distribution (PFD) and two PFD processes introduced in Arnold et al. [2]. We say that X has a power function distribution (PFD) with parameter α , if its distribution is of the form

$$F_X(x) = x^\alpha, \quad 0 < x < 1, \quad \alpha > 0. \quad (1)$$

If X has distribution (1) then we will write $X \sim PFD(\alpha)$. The corresponding density is

$$f_X(x) = \alpha x^{\alpha-1}, \quad 0 < x < 1, \quad \alpha > 0. \quad (2)$$

This can be recognized as a $Beta(\alpha, 1)$ density. There are two special properties of power function distributions that will be useful for defining stationary PFD processes.

- (A) **Closure under maximization:** If X and Y are independent with $X \sim PFD(\alpha)$ and $Y \sim PFD(\beta)$, then $Z = \max\{X, Y\} \sim PFD(\alpha + \beta)$.
- (B) **Closure under raising to a power:** If $X \sim PFD(\alpha)$ and if $\delta > 0$, then $V = X^\delta \sim PFD(\alpha/\delta)$.

The following two stationary power function distribution processes (PFD-processes) with PFD marginals proved to be useful for generating stationary proportional reversed hazard processes.

- (I) **A Kundu PFD process:** Consider an i.i.d. sequence, $\{U_n\}$, of *uniform*(0, 1) random variables. The PFD process $\{X_n^*\}$ will be defined by

$$X_n^* = \max \left\{ U_n^{1/\alpha}, U_{n-1}^{1/\beta} \right\}, \quad (3)$$

where $\alpha, \beta > 0$, see Kundu [4].

- (II) **An A-M PFD process:** Define $Y_0^* = U_0^{1/\alpha} \sim PF(\alpha)$ and for $n = 1, 2, \dots$ define the PFD process

$$Y_n^* = \max \left\{ (Y_{n-1}^*)^{\alpha/(\alpha-\delta)}, U_n^{1/\delta} \right\}, \quad (4)$$

where $\delta \in (0, \alpha)$ and the U_n 's are i.i.d. *uniform*(0, 1) random variables, see Arnold et al. [2].

Now, if $\{X_n\}$ is either one of the above PFD processes, we can construct a stationary proportional reverse hazard process of the form

$$Y_n = F_0^{-1}(X_n), \quad n = 0, 1, 2, \dots \quad (5)$$

where F_0 is an arbitrary distribution function with support $(0, \infty)$.

We refer to Arnold et al. [2] for stochastic properties, inferential aspects, and also an example of applications of the above two new PFD processes.

3. Proportional hazard processes

An important tool that will be used in constructing proportional hazard processes is a distribution that we will call the complementary power function distribution (CPFD). It might also be called a power function survival distribution. The construction of proportional hazard processes, using the CPFD will be a close parallel to that used in developing the proportional reverse hazard processes above.

3.1. The complementary power function distribution

We say that X has a complementary power function distribution (CPFD) with parameter α , if its distribution is of the form

$$F_X(x) = 1 - (1 - x)^\alpha, \quad 0 < x < 1, \quad \alpha > 0, \quad (6)$$

i.e., if $\bar{F}_X(x) = (1 - x)^\alpha$, $0 < x < 1$. If X has distribution (6) then we will write $X \sim \text{CPFD}(\alpha)$. The corresponding density is

$$f_X(x) = \alpha(1 - x)^{\alpha-1}, \quad 0 < x < 1, \quad \alpha > 0. \quad (7)$$

This can be recognized as a $Beta(1, \alpha)$ density. There are two special properties of complementary power function distributions that will be useful for defining stationary CPFD processes.

(C) Closure under minimization: If X and Y are independent with $X \sim \text{CPFD}(\alpha)$ and $Y \sim \text{CPFD}(\beta)$, then $Z = \min\{X, Y\} \sim \text{CPFD}(\alpha + \beta)$.

(D) Closure under complementary raising to a power: If $X \sim \text{CPFD}(\alpha)$ and if $\delta > 0$, then $V = 1 - (1 - X)^\delta \sim \text{CPFD}(\alpha/\delta)$. Of course, this is just a rewritten version of statement (B) for PFD variables.

3.2. Two CPFD-processes

Note that, if $U \sim \text{Uniform}(0, 1)$ and $X = 1 - U^{1/\alpha}$ then $P(X > x) = (1 - x)^\alpha$.

Now, analogous to the two processes defined in Section 2.1, we have the following CPFD processes.

(I) A Kundu CPFDP process: Consider an i.i.d. sequence, $\{U_n\}$, of *uniform*(0, 1) random variables. The Kundu CPFDP-process $\{V_n^*\}$ will be defined by

$$V_n^* = \min\{1 - U_{n-1}^{1/\alpha}, 1 - U_n^{1/\beta}\} \quad (8)$$

where $\alpha, \beta > 0$.

Note that $V_n^* = 1 - X_n^*, \forall n$ where $\{X_n^*\}$ is the Kundu PFD process .

(II) An A-M CPFDP process: We define the A-M CPFDP process $\{W_n^*\}$ as follows. Define $W_0^* = 1 - U_0^{1/\alpha}$ and for $n = 1, 2, \dots$ define the CPFDP process

$$W_n^* = \min\{1 - (1 - W_{n-1}^*)^{\alpha/(\alpha-\delta)}, 1 - U_n^{1/\delta}\}, \quad (9)$$

where $\delta \in (0, \alpha)$ and the U_n 's are i.i.d. *uniform*(0, 1) random variables.

Note that $W_n^* = 1 - Y_n^* \forall n$ where $\{Y_n^*\}$ is the A-M PFD process .

3.3. Two proportional hazard processes

The motivation for introducing complementary power function distribution processes (which are so simply related to power function distribution processes) is that the CPFDP processes are the perfect tools for developing stationary proportional hazard processes, which will be of interest since an assumption of proportional hazards is frequently deemed to be plausible in many applications. A simple marginal transformation applied to a CPFDP process will be all that is required.

We can then define a Kundu PH process $\{S_n^*\}$ with PH marginals using a Kundu CPFDP process $\{V_n^*\}$ as follows

$$S_n^* = F_0^{-1}(V_n^*), \quad n = 0, 1, 2, \dots \quad (10)$$

where it can be verified that

$$P(S_n^* > s) = [\overline{F}_0(s)]^{\alpha+\beta}.$$

In parallel fashion, we can define an A-M PH process as follows

$$T_n^* = F_0^{-1}(W_n^*) = F_0^{-1}(1 - Y_n^*) \quad (11)$$

where W_n^* is the A-M CPFDP process.

We may readily verify that $\{T_n^*\}$ has proportional hazard marginal distributions as follows:

$$\begin{aligned} P(T_n^* > t) &= P(F_0^{-1}(1 - Y_n^*) > t) = P(1 - Y_n^* > F_0(t)) \\ &= P(Y_n^* < 1 - F_0(t)) = P(Y_n^* < \overline{F}_0(t)) \\ &= [\overline{F}_0(t)]^\alpha. \end{aligned}$$

3.4. New stationary Pareto processes

A specific case of a proportional hazard family of distributions is provided by the classical Pareto distribution, which is much used in the study of inequality in economic variables such as income [1]. Recall that a random variable X has a classical Pareto distribution or Pareto(I) distribution if its survival function is of the form

$$\overline{F}_X(x) = (x/\sigma)^{-\alpha}, \quad x > \sigma.$$

To indicate this, we write $X \sim P(I)(\sigma, \alpha)$.

Note that if $X_1 \sim P(I)(\sigma, \alpha_1)$ and $X_2 \sim P(I)(\sigma, \alpha_2)$ are independent and if we define $Z = \min\{X_1, X_2\}$ then $Z \sim P(I)(\sigma, \alpha_1 + \alpha_2)$.

Note that the family of $P(I)(\sigma, \alpha)$ distributions, for a fixed value of σ is a proportional-hazard family with its basic survival function of the form $\overline{F}_0(x) = (x/\sigma)^{-1}$, with corresponding quantile function

$$F_0^{-1}(y) = \sigma(1 - y)^{-1}.$$

It will thus be possible to define two stationary Pareto processes as in Subsection 3.3 using this particular form for $\overline{F}_0(x)$. Instead, we will equivalently develop the processes by utilizing a simple relationship between Pareto and uniform distributions.

Recall that the basic proportional hazard distribution (or CPGD) family has a density of the form:

$$f_0(x : \alpha) = \alpha(1 - x)^{\alpha-1}, \quad 0 < x < 1,$$

Which is the density of a random variable of the form $Y = 1 - U^{1/\alpha}$

where $U \sim \text{uniform}(0, 1)$.

Note that a $Pareto(I)(\sigma, \alpha)$ variable X can be represented in the form

$$X = \sigma U^{-1/\alpha},$$

or equivalently

$$X = \sigma(1 - U)^{-1/\alpha}.$$

Now, we may define two new Pareto processes as follows:

(I) **A Kundu Pareto process:** $S_n^* = \sigma/X_n^*$ where $\{X_n^*\}$ is a Kundu PFD process. i.e.,

$$S_n^* = \sigma \min\{U_{n-1}^{-1/\alpha}, U_n^{-1/\beta}\}, \quad (12)$$

where $\alpha, \beta > 0$. For this process we have $S_n^* \sim P(\sigma, \alpha + \beta) \forall n$. We refer to Figures 1 for examples of simulated sample paths of steps $n = 10$ and $n = 50$.

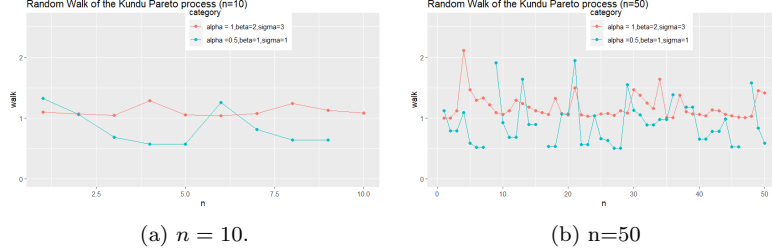


Figure 1. Sample Path for the Kundu Pareto process

(II) **An A-M Pareto process:** define

$$T_{-1}^* \sim \text{Pareto}(I)(\sigma, \alpha)$$

and

$$T_n^* = \sigma \min\{(T_{n-1}^*/\sigma)^{-\alpha/(\alpha-\delta)}, U_n^{-1/\delta}\}. \quad (13)$$

where $0 < \delta < \alpha$. For this process we have $T_n^* \sim P(\sigma, \alpha)$.

We refer to Figures 2 for examples of simulated sample paths of steps $n = 10$ and $n = 50$.

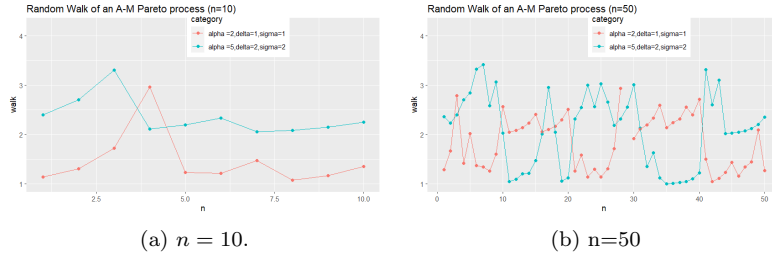


Figure 2. Sample Path for an A-M Pareto process

4. Stochastic properties of the CFPD and Pareto processes

4.1. Kundu CFPD process

Theorem 4.1. *If the sequence of random variables $\{V_n^*\}$ is as defined in equation (8), then the following statements are true:*

- (a) *For each n , $\{V_n^*\}$ has a CFPD($\alpha + \beta$) distribution.*

(b) The joint distribution function of (V_{n-1}^*, V_n^*) is

$$1 - (1 - v_{n-1})^{\alpha+\beta} - (1 - v_n)^{\alpha+\beta} + (1 - v_{n-1})^\alpha (1 - v_n)^\beta \min\{(1 - v_n)^\alpha, (1 - v_{n-1})^\beta\}.$$

(c) $\{V_n^*\}$ is a stationary non-Markovian process.

Proof. If V_n^* is a CPDF process, then $V_n^* = 1 - X_n^*$, where X_n^* is a stationary non-Markovian PFD process (see Arnold et al. [2]). Hence, the stationarity and non-Markovian nature of V_n^* follows. To prove part (a) we use the following notation

$$\begin{aligned} F_{V_n^*}(v_n) = P(V_n^* \leq v) &= P(1 - X_n^* \leq v_n) \\ &= P(X_n^* \geq 1 - v_n) \\ &= 1 - P(X_n^* \leq 1 - v_n) \\ &= 1 - (1 - v_n)^{\alpha+\beta} \end{aligned} \quad (14)$$

For the proof of Part (b), we refer to $V_n^* = 1 - X_n^*$ and the analogous result for X_n^* in Arnold et al. [2]. Now, the joint distribution of V_{n-1}^* and V_n^* ,

$$\begin{aligned} F_{V_n^*, V_{n-1}^*}(v_n, v_{n-1}) &= P(V_n^* \leq v_n, V_{n-1}^* \leq v_{n-1}) \\ &= P(1 - X_n^* \leq v_n, 1 - X_{n-1}^* \leq v_{n-1}) \\ &= P(X_n^* \geq 1 - v_n, X_{n-1}^* \geq 1 - v_{n-1}) \\ &= 1 - F_{X_{n-1}^*}(1 - v_{n-1}) - F_{X_n^*}(1 - v_n) + F_{X_n^*, X_{n-1}^*}(1 - v_n, 1 - v_{n-1}) \\ &= 1 - (1 - v_{n-1})^{\alpha+\beta} - (1 - v_n)^{\alpha+\beta} + (1 - v_n)^\alpha \\ &\quad (1 - v_{n-1})^\beta \min\{(1 - v_n)^\alpha, (1 - v_{n-1})^\beta\}. \end{aligned} \quad (15)$$

□

Corollary 4.2. If the sequence of random variables $\{V_n^*\}$ is as defined in equation (8), then the mean and variance are

$$E(V_n^*) = \frac{1}{\alpha + \beta + 1} \quad (16)$$

$$Var(V_n^*) = \frac{\alpha + \beta}{(\alpha + \beta + 1)^2(\alpha + \beta + 2)} \quad (17)$$

Proof. We know that

$$E(V_n^*) = E(1 - X_n^*) = 1 - E(X_n^*) = 1 - \frac{\alpha + \beta}{\alpha + \beta + 1} = \frac{1}{\alpha + \beta + 1} \quad (18)$$

$$Var(V_n^*) = Var(1 - X_n^*) = Var(X_n^*) = \frac{\alpha + \beta}{(\alpha + \beta + 1)^2(\alpha + \beta + 2)}. \quad (19)$$

□

If the sequence of random variables $\{V_n^*\}$ is as defined in equation (8), the auto-

correlation function is

$$\begin{aligned} \text{Corr}(V_n^*, V_{n-1}^*) &= \frac{\text{Cov}(V_n^*, V_{n-1}^*)}{\sqrt{\text{Var}(V_n^*)\text{Var}(V_{n-1}^*)}} \\ &= \frac{E(V_n^*V_{n-1}^*) - \left(\frac{1}{\alpha+\beta+1}\right)^2}{\frac{\alpha+\beta}{(\alpha+\beta+1)^2(\alpha+\beta+2)}} \end{aligned} \quad (20)$$

From the usual expression for the auto-correlation in (20), it is evident that the detail lacking is an expression for $E(V_n^*V_{n-1}^*)$. To evaluate this, we argue as follows

$$\begin{aligned} E(V_n^*V_{n-1}^*) &= E((1 - X_n^*)(1 - X_{n-1}^*)) \\ &= E(1 - X_{n-1}^* - X_n^* + X_n^*X_{n-1}^*) \\ &= 1 - 2\frac{1}{\alpha + \beta + 1} + E(X_n^*X_{n-1}^*). \end{aligned} \quad (21)$$

For the computation of $E(X_n^*X_{n-1}^*)$ we refer to page 8 of Arnold et al. [2], hence

$$E(X_n^*X_{n-1}^*) = A + B + C + D \quad (22)$$

where

$$\begin{aligned} A &= \frac{\beta^2}{\beta + 1} \left[\frac{1}{\alpha + \beta + 1} - \frac{\alpha}{\beta^2 + \alpha^2 + \alpha\beta + \alpha + \beta} \right], \\ B &= \left[\frac{\alpha + \beta + \alpha^2 + \beta^2 + \alpha\beta}{\alpha\beta} \right]^{-1}, \\ C &= \frac{\beta\alpha}{(\beta + 1)(\alpha + 1)} \left[1 - \frac{\beta}{\alpha + \beta + 1} - \frac{\alpha}{\alpha + \beta + 1} + \frac{\alpha\beta}{\alpha^2 + \beta^2 + \alpha\beta + \alpha + \beta} \right], \\ D &= \frac{\alpha^2\beta}{\alpha + 1} \left[\frac{1}{\beta^2 + \alpha + \alpha\beta} - \frac{1}{\alpha^2 + \beta^2 + 2\alpha + \alpha\beta} \right]. \end{aligned}$$

Theorem 4.3. *If the sequence of random variable $\{V_n^*\}$ is as defined in equation (8) then*

$$P(V_1^* < V_2^*) = \begin{cases} \frac{\alpha+\beta}{2\alpha+\beta} & \text{if } \alpha > \beta \\ \frac{\beta}{2\beta+\alpha} & \text{if } \alpha \leq \beta. \end{cases} \quad (23)$$

Proof. By definition of the process, we have $V_1^* = 1 - X_1^*$ and $V_2^* = 1 - X_2^*$.

$$P(V_1^* < V_2^*) = P(1 - X_1^* < 1 - X_2^*) = P(X_1^* > X_2^*) \quad (24)$$

So, if $\alpha > \beta$ we have

$$P(V_1^* < V_2^*) = \frac{\alpha + \beta}{2\alpha + \beta} \quad (25)$$

Now, for $\alpha < \beta$

$$P(V_1^* > V_2^*) = \frac{\alpha + \beta}{2\beta + \alpha} \quad (26)$$

□

4.2. A-M CPF D process

Theorem 4.4. *If the sequence of random variables $\{W_n^*\}$ is as defined in equation (9), then the following statements are true:*

- (a) *For each n , $\{W_n^*\}$ has a CPF D(α) distribution.*
- (b) *The joint distribution function of (W_{n-1}^*, W_n^*) is $1 - (1 - w_{n-1})^\alpha - (1 - w_n)^\alpha + (1 - w_n)^\delta \min\{(1 - v_n)^{\alpha-\delta}, (1 - v_{n-1})^\alpha\}$.*
- (c) *$\{W_n^*\}$ is a stationary Markovian process.*

Proof. We use the property that , if $\{W_n^*\}$ is an A-M CPF D process $W_n^* = 1 - Y_n^*$, where $\{Y_n^*\}$ is an A-M PFD process. To prove part (a) we use the following notation

$$\begin{aligned} F_{W_n^*}(w) &= P(W_n^* \leq w) = P(1 - Y_n^* \leq w) = P(Y_n^* \geq 1 - w) = 1 - P(Y_n^* \leq 1 - w) \\ &= 1 - (1 - w)^\delta P(Y_{n-1}^* \leq (1 - w)^{\frac{\alpha-\delta}{\alpha}}) \end{aligned} \quad (27)$$

Hence, using mathematical induction, we have

$$F_{W_n^*}(w) = w^\alpha. \quad (28)$$

Now, for Part (b) of the theorem, the joint distribution of W_n^* and W_{n+1}^* is, by stationarity of the process, the same as the joint distribution of W_0^* and W_1^* which may be computed as follows:

$$\begin{aligned} F_{W_0^*, W_1^*}(w_0, w_1) &= P(W_0^* \leq w_0, W_1^* \leq w_1) \\ &= P(1 - Y_0^* \leq w_0, 1 - Y_1^* \leq w_1) \\ &= P(Y_0^* \geq 1 - w_0, Y_1^* \geq 1 - w_1) \\ &= 1 - F_{Y_0^*}(1 - w_0) - F_{Y_1^*}(1 - w_1) + F_{Y_0^*, Y_1^*}(1 - w_0, 1 - w_1) \\ &= 1 - (1 - w_0)^\alpha - (1 - w_1)^\alpha + (1 - w_1)^\delta \min\{w_0^\alpha, w_1^{\alpha-\delta}\}. \end{aligned} \quad (29)$$

It remains to prove Part (c). Stationarity is trivial from the definition of the process. For the Markovian property, recall that $W_n^* = 1 - Y_n^*$, we know that Y_n^* is Markovian, and W_n^* is an invertible transformation of a Markovian process, hence W_n^* is also Markovian. □

Corollary 4.5. *If the sequence of random variable $\{W_n^*\}$ is as defined in equation (9), then the mean and variance are*

$$E(W_n^*) = \frac{1}{\alpha + 1} \quad (30)$$

$$Var(W_n^*) = \frac{\alpha}{(\alpha + 1)^2(\alpha + 2)}. \quad (31)$$

Proof. We know that

$$E(W_n^*) = E(1 - Y_n^*) = 1 - E(Y_n^*) = 1 - \frac{\alpha}{\alpha + 1} = \frac{1}{\alpha + 1} \quad (32)$$

$$Var(W_n^*) = Var(1 - Y_n^*) = Var(Y_n^*) = \frac{\alpha}{(\alpha + 1)^2(\alpha + 2)}. \quad (33)$$

□

If the sequence of random variables $\{W_n^*\}$ is as defined in equation (9) then the auto-correlation function will be

$$Corr(W_0^*, W_1^*) = \frac{E(W_0^*W_1^*) - \left(\frac{1}{\alpha+1}\right)^2}{\frac{\alpha}{(\alpha+1)^2(\alpha+2)}}. \quad (34)$$

which is monotone in δ .

The only thing needed in order to get the auto-correlation is $E(W_0^*W_1^*)$. Now,

$$\begin{aligned} E(W_0^*W_1^*) &= E((1 - Y_0^*)(1 - Y_1^*)) = 1 - E(Y_0^*) - E(Y_1^*) + E(Y_0^*Y_1^*) \\ &= 1 - 2\frac{1}{\alpha+1} + \frac{\delta}{\delta+1} \left[\frac{\alpha}{\alpha+1} - \left(\frac{1}{\alpha} + \frac{\delta+1}{\alpha-\delta} + 1 \right)^{-1} \right] + \left[\frac{1}{\alpha} + \frac{1+\delta}{\alpha-\delta} + 1 \right]^{-1}. \end{aligned} \quad (35)$$

For the computation of $E(Y_0^*Y_1^*)$, we refer to page 18 of Arnold et al. [2].

Theorem 4.6. *If the sequence of random variables $\{W_n^*\}$ is as defined in equation (9) then*

$$P(W_1^* < W_0^*) = \frac{\delta}{\alpha + \delta}. \quad (36)$$

Proof. By the definition of the process we have $W_1^* = 1 - Y_1^*$, consequently $P(W_1^* < W_0^*)$ is given by

$$P(W_1^* < W_0^*) = P(1 - Y_1^* < 1 - Y_0^*) = P(Y_1^* > Y_0^*) = \frac{\delta}{\alpha + \delta} \quad (37)$$

For the computation of $P(Y_1^* > Y_0^*)$, we refer to page 19 of Arnold et al. [2]. □

4.3. Kundu Pareto process

Theorem 4.7. *If the sequence of random variables $\{S_n^*\}$ is as defined in equation (12), then the following statements are true:*

- (a) *For each n , $\{S_n^*\}$ has a Pareto($\alpha + \beta$) distribution.*
- (b) *The joint distribution function of (S_{n-1}^*, S_n^*) is*

$$1 - \left(\frac{s_{n-1}}{\sigma}\right)^{-(\alpha+\beta)} - \left(\frac{s_n}{\sigma}\right)^{-(\alpha+\beta)} + \left(\frac{s_n}{\sigma}\right)^{-\alpha} \left(\frac{s_{n-1}}{\sigma}\right)^{-\beta} \min \left\{ \left(\frac{s_n}{\sigma}\right)^{-\alpha}, \left(\frac{s_{n-1}}{\sigma}\right)^{-\beta} \right\}.$$

(c) $\{S_n^*\}$ is a stationary non-Markovian process.

Proof. If S_n^* is a Pareto process, then $S_n^* = \sigma/X_n^*$, where X_n^* is a stationary non-Markovian PFD process (see Arnold et al. [2]). Hence, the stationarity and non-Markovian nature of S_n^* follows. To prove part (a) we use the following notation

$$\begin{aligned}
F_{S_n^*}(s_n) = P(S_n^* \leq s_n) &= P(\sigma/X_n^* \leq s_n) \\
&= P(X_n^* \geq \sigma/s_n) \\
&= 1 - P(X_n^* \leq \sigma/s_n) \\
&= 1 - \left(\frac{s_n}{\sigma}\right)^{-(\alpha+\beta)}
\end{aligned} \tag{38}$$

For the proof of Part (b), noting that $S_n^* = \sigma/X_n^*$, one can refer to the analogous result for X_n^* in Arnold et al. [2]. Now, the joint distribution of S_{n-1}^* and S_n^* ,

$$\begin{aligned}
F_{S_n^*, S_{n-1}^*}(s_n, s_{n-1}) &= P(S_n^* \leq s_n, S_{n-1}^* \leq s_{n-1}) \\
&= P(\sigma/X_n^* \leq s_n, \sigma/X_{n-1}^* \leq s_{n-1}) \\
&= P(X_n^* \geq \sigma/s_n, X_{n-1}^* \geq \sigma/s_{n-1}) \\
&= 1 - F_{X_{n-1}^*}(\sigma/s_{n-1}) - F_{X_n^*}(\sigma/s_n) + F_{X_n^*, X_{n-1}^*}(\sigma/s_n, \sigma/s_{n-1}) \\
&= 1 - \left(\frac{s_{n-1}}{\sigma}\right)^{-(\alpha+\beta)} - \left(\frac{s_n}{\sigma}\right)^{-(\alpha+\beta)} + \\
&\quad \left(\frac{s_n}{\sigma}\right)^{-\alpha} \left(\frac{s_{n-1}}{\sigma}\right)^{-\beta} \min \left\{ \left(\frac{s_n}{\sigma}\right)^{-\alpha}, \left(\frac{s_{n-1}}{\sigma}\right)^{-\beta} \right\}.
\end{aligned} \tag{39}$$

□

Corollary 4.8. If the sequence of random variables $\{S_n^*\}$ is as defined in equation (12), then the mean (for $\alpha + \beta > 1$) and variance (for $\alpha + \beta > 2$) are

$$E(S_n^*) = \frac{\sigma(\alpha + \beta)}{\alpha + \beta - 1} \tag{40}$$

$$Var(S_n^*) = \frac{\sigma^2(\alpha + \beta)}{(\alpha + \beta - 1)^2(\alpha + \beta - 2)} \tag{41}$$

Proof. We know that

$$E(S_n^*) = E(\sigma/X_n^*) = \sigma E(1/X_n^*) = \sigma \frac{\alpha + \beta}{\alpha + \beta - 1} \tag{42}$$

$$Var(S_n^*) = Var(\sigma/X_n^*) = \sigma^2 Var(1/X_n^*) = \sigma^2 \frac{\alpha + \beta}{(\alpha + \beta - 1)^2(\alpha + \beta - 2)}. \tag{43}$$

□

If the sequence of random variables $\{S_n^*\}$ is as defined in equation (12), the auto-

correlation function is

$$\begin{aligned}
Corr(S_n^*, S_{n-1}^*) &= \frac{Cov(S_n^*, S_{n-1}^*)}{\sqrt{Var(S_n^*)Var(S_{n-1}^*)}} \\
&= \frac{E(S_n^* S_{n-1}^*) - (\sigma \frac{\alpha+\beta}{\alpha+\beta-1})^2}{\sigma^2 \frac{\alpha+\beta}{(\alpha+\beta-1)^2(\alpha+\beta-2)}}
\end{aligned} \tag{44}$$

From the usual expression for the auto-correlation in (4.7), it is evident that the detail lacking is an expression for $E(S_n^* S_{n-1}^*)$. To evaluate this, we argue as follows

$$\begin{aligned}
E(S_n^* S_{n-1}^*) &= E((\sigma/X_n^*)(\sigma/X_{n-1}^*)) \\
&= \sigma^2 E(1/X_{n-1}^* 1/X_n^*)
\end{aligned} \tag{45}$$

For the computation of $E(1/X_{n-1}^* 1/X_n^*)$

$$\begin{aligned}
E(1/X_{n-1}^* 1/X_n^*) &= E(V^{-1/\beta} W^{-1/\beta} I(U < V^{\alpha/\beta}, W > V^{\beta/\alpha})) \\
&\quad + E(V^{-1/\beta} V^{-1/\alpha} I(U < V^{\alpha/\beta}, W < V^{\beta/\alpha})) \\
&\quad + E(U^{-1/\alpha} W^{-1/\beta} I(U > V^{\alpha/\beta}, W > V^{\beta/\alpha})) \\
&\quad + E(U^{-1/\alpha} V^{-1/\beta} I(U > V^{\alpha/\beta}, W < V^{\beta/\alpha})) \\
&= A + B + C + D
\end{aligned} \tag{46}$$

where

$$\begin{aligned}
X_n^* &= \max\{U^{1/\alpha}, V^{1/\beta}\} \\
X_{n-1}^* &= \max\{V^{1/\alpha}, W^{1/\beta}\}
\end{aligned}$$

and expressions for A, B, C, and D are as follows.

$$\begin{aligned}
A &= \int_0^1 dv \int_0^{v^{\alpha/\beta}} du \int_{v^{\beta/\alpha}}^1 dw [v^{-1/\beta} w^{-1/\beta}] \\
&= \frac{\beta^2}{\beta-1} \left[\frac{1}{\alpha+\beta-1} - \frac{\alpha}{\alpha^2 + \beta^2 + \alpha\beta - \alpha - \beta} \right],
\end{aligned}$$

$$\begin{aligned}
B &= \int_0^1 dv \int_0^{v^{\alpha/\beta}} du \int_0^{v^{\beta/\alpha}} dw [v^{-1/\alpha} v^{-1/\beta}] \\
&= \left[\frac{\alpha^2 + \beta^2 + \alpha\beta - \alpha - \beta}{\alpha\beta} \right]^{-1},
\end{aligned}$$

$$\begin{aligned}
C &= \int_0^1 dv \int_{v^{\alpha/\beta}}^1 du \int_{v^{\beta/\alpha}}^1 dw [u^{-1/\alpha} w^{-1/\beta}] \\
&= \frac{\alpha\beta}{(\beta-1)(\alpha-1)} \left[1 - \frac{\beta}{\alpha+\beta-1} - \frac{\alpha}{\alpha+\beta-1} + \frac{\alpha\beta}{1+2\alpha\beta-\alpha-\beta} \right],
\end{aligned}$$

$$\begin{aligned}
D &= \int_0^1 dv \int_{v^{\alpha/\beta}}^1 du \int_0^{v^{\beta/\alpha}} dw [v^{-1/\alpha} u^{-1/\beta}] \\
&= \frac{\alpha^2}{\alpha-1} \left[\frac{1}{\alpha+\beta-1} - \frac{\beta}{\alpha^2+\beta^2-\alpha-\beta} \right].
\end{aligned}$$

Theorem 4.9. *If the sequence of random variable $\{S_n^*\}$ is as defined in equation (12) then*

$$P(S_1^* < S_2^*) = \begin{cases} \frac{\alpha+\beta}{2\alpha+\beta} & \text{if } \alpha > \beta \\ \frac{\beta}{2\beta+\alpha} & \text{if } \alpha \leq \beta. \end{cases} \quad (47)$$

Proof. By definition of the process, we have $S_1^* = \sigma/X_1^*$ and $S_2^* = \sigma/X_2^*$.

$$P(S_1^* < S_2^*) = P(\sigma/X_1^* < \sigma/X_2^*) = P(X_1^* > X_2^*) \quad (48)$$

So, if $\alpha > \beta$ we have

$$P(S_1^* < S_2^*) = \frac{\alpha+\beta}{2\alpha+\beta} \quad (49)$$

Now, for $\alpha < \beta$

$$P(S_1^* > S_2^*) = \frac{\alpha+\beta}{2\beta+\alpha} \quad (50)$$

□

4.4. A-M Pareto process

Theorem 4.10. *If the sequence of random variables $\{T_n^*\}$ is as defined in equation (13), then the following statements are true:*

- (a) (a) For each n , $\{T_n^*\}$ has a Pareto(σ, α) distribution.
- (b) The joint distribution function of (T_{n-1}^*, T_n^*) is

$$1 - (w_0/\sigma)^{-\alpha} - (w_1/\sigma)^{-\alpha} + (w_1/\sigma)^{-\delta} \min\{(w_0/\sigma)^{-\alpha}, (w_1/\sigma)^{-(\alpha/\delta)}\}.$$

- (c) $\{T_n^*\}$ is a stationary Markovian process.

Proof. We use the property that , if $\{T_n^*\}$ is an A-M Pareto process $T_n^* = \sigma/Y_n^*$, where $\{Y_n^*\}$ is an A-M PFD process. To prove part (a) we use the following notation

$$\begin{aligned} F_{T_n^*}(t) &= P(T_n^* \leq t) = P(\sigma/Y_n^* \leq t) = P(Y_n^* \geq \sigma/t) = 1 - P(Y_n^* \leq \sigma/t) \\ &= 1 - (\sigma/t)^\delta P\left(Y_{n-1}^* \leq (\sigma/t)^{\frac{\alpha-\delta}{\alpha}}\right) \end{aligned} \quad (51)$$

Hence, using mathematical induction, we have

$$F_{T_n^*}(t) = \left(\frac{t}{\sigma}\right)^{-\alpha}. \quad (52)$$

Now, for Part (b) of the theorem, the joint distribution of T_n^* and T_{n+1}^* is, by stationarity of the process, the same as the joint distribution of T_0^* and T_1^* which may be computed as follows:

$$\begin{aligned} F_{T_0^*, T_1^*}(t_0, t_1) &= P(T_0^* \leq t_0, T_1^* \leq t_1) \\ &= P(\sigma/Y_0^* \leq w_0, \sigma/Y_1^* \leq w_1) \\ &= P(Y_0^* \geq \sigma/w_0, Y_1^* \geq \sigma/w_1) \\ &= 1 - F_{Y_0^*}(\sigma/w_0) - F_{Y_1^*}(\sigma/w_1) + F_{Y_0^*, Y_1^*}(\sigma/w_0, \sigma/w_1) \\ &= 1 - (w_0/\sigma)^{-\alpha} - (w_1/\sigma)^{-\alpha} + (w_1/\sigma)^{-\delta} \min\{(w_0/\sigma)^{-\alpha}, (w_1/\sigma)^{-(\alpha-\delta)}\} \end{aligned} \quad (53)$$

It remains to prove Part (c). Stationarity is trivial from the definition of the process. For the Markovian property, recall that $T_n^* = \sigma/Y_n^*$, we know that Y_n^* is Markovian, and T_n^* is an invertible transformation of a Markovian process, hence T_n^* is also Markovian. \square

Corollary 4.11. *If the sequence of random variable $\{T_n^*\}$ is as defined in equation (13), then the mean (for $\alpha > 1$) and variance (for $\alpha > 2$) are*

$$E(T_n^*) = \frac{\sigma\alpha}{\alpha - 1} \quad (54)$$

$$Var(T_n^*) = \frac{\sigma^2\alpha}{(\alpha - 1)^2(\alpha - 2)}. \quad (55)$$

Proof. We know that

$$E(T_n^*) = E(\sigma/Y_n^*) = \sigma E(1/Y_n^*) = \sigma \frac{1}{\alpha - 1} \quad (56)$$

$$Var(T_n^*) = Var(\sigma/Y_n^*) = \sigma^2 Var(1/Y_n^*) = \sigma^2 \frac{\alpha}{(\alpha - 1)^2(\alpha - 2)}. \quad (57)$$

\square

If the sequence of random variables $\{T_n^*\}$ is as defined in equation (13) then the

auto-correlation function will be

$$\text{Corr}(T_0^*, T_1^*) = \frac{E(T_0^* T_1^*) - \left(\frac{\sigma\alpha}{\alpha-1}\right)^2}{\frac{\sigma^2\alpha}{(\alpha-1)^2(\alpha-2)}}. \quad (58)$$

which is monotone in δ .

The only thing needed in order to get the auto-correlation is $E(T_0^* T_1^*)$. Now,

$$\begin{aligned} E(T_0^* T_1^*) &= E((\sigma/Y_0^*)(\sigma/Y_1^*)) = \sigma^2 E((1/Y_0^*)(1/Y_1^*)) \\ &= \sigma^2 E(U_0^{-1/\alpha} \min\{(T_0^*)^{-\frac{\alpha}{\alpha-\delta}}, U_1^{-1/\delta}\}) \\ &= \sigma^2 E(U_0^{-1/\alpha} \min\{(U_0)^{-\frac{1}{\alpha-\delta}}, U_1^{-1/\delta}\}) \\ &= \sigma^2 \int_0^1 u_0^{-\frac{1}{\alpha}} \left[\int_0^{u_0^{\frac{\delta}{\alpha-\delta}}} u_1^{-\frac{1}{\alpha-\delta}} du_1 \right] du_0 + \sigma^2 \int_0^1 u_0^{-\frac{1}{\alpha}} \left[\int_{u_0^{\frac{\delta}{\alpha-\delta}}}^1 u_0^{-\frac{1}{\delta}} du_1 \right] du_0 \\ &= \sigma^2 \frac{\alpha\delta}{\delta-1} \left[\frac{1}{\alpha-1} - \frac{\alpha-\delta}{\alpha^2-2\alpha+\delta} \right] + \sigma^2 \frac{\alpha(\alpha-\delta)}{\alpha^2-2\alpha+\delta}. \end{aligned}$$

Theorem 4.12. *If the sequence of random variables $\{T_n^*\}$ is as defined in equation (13) then*

$$P(T_1^* < T_0^*) = \frac{\delta}{\alpha + \delta}. \quad (59)$$

Proof. By the definition of the process we have $T_1^* = \sigma/Y_1^*$, consequently $P(T_1^* < T_0^*)$ is given by

$$P(T_1^* < T_0^*) = P(\sigma/Y_1^* < \sigma/Y_0^*) = P(Y_1^* > Y_0^*) = \frac{\delta}{\alpha + \delta} \quad (60)$$

For the computation of $P(Y_1^* > Y_0^*)$, we refer to page 19 of Arnold et al. [2]. \square

5. Statistical Inference

5.1. Estimation

In the following, we use the method of moments to obtain consistent estimators for the parameters for both of the CPFDF processes.

5.1.1. Method of moments

We define the following statistics for the observed sample paths $\{V_1^*, \dots, V_m^*\}$ from the Kundu CPFDF process given in (8).

$$\bar{V}^* = \frac{1}{m} \sum_{i=1}^m V_i^* \quad (61)$$

$$P = \frac{1}{m} \sum_{i=1}^m I(V_i^* < V_{i-1}^*). \quad (62)$$

where $I(\cdot)$ is the indicator function.

For the Kundu CPFD process in (8) and the expression for $P(V_1^* < V_2^*)$ for the case $\alpha > \beta$, the following two moment equations can be solved for estimates of α and β .

$$\overline{V^*} = \frac{1}{\alpha + \beta + 1} \quad (63)$$

$$P = \frac{\alpha + \beta}{2\alpha + \beta}. \quad (64)$$

We obtain the following consistent estimates.

$$\hat{\alpha} = \frac{1 - \overline{V^*}(1 + \hat{\beta})}{\overline{V^*}} \quad (65)$$

$$\hat{\beta} = \frac{(2P - 1)(1 - \overline{V^*})}{\overline{V^*}P}. \quad (66)$$

This solution does not always satisfy the condition $\alpha > \beta$. Now, for the case $\alpha < \beta$ we have

$$P(V_1^* > V_2^*) = \frac{\alpha}{2\alpha + \beta}.$$

The following two moment equations can be solved for estimates of α and β .

$$\overline{V^*} = \frac{1}{\alpha + \beta + 1} \quad (67)$$

$$P = \frac{\beta}{2\beta + \alpha}. \quad (68)$$

We can then obtain the following consistent estimates.

$$\hat{\beta} = \frac{(1 - \overline{V^*})P}{(1 - P)\overline{V^*}} \quad (69)$$

$$\hat{\alpha} = \frac{\hat{\beta}(1 - 2P)}{P} \quad (70)$$

Now, for the A-M CPFD process in (9) and the expression for $P(W_1^* < W_2^*)$, the following two moment equations can be solved for estimates of α and β .

$$\overline{W^*} = \frac{1}{\alpha + 1} \quad (71)$$

$$P = \frac{\delta}{\alpha + \delta}. \quad (72)$$

We obtain the following consistent estimates.

$$\hat{\alpha} = \frac{1 - \overline{W}^*}{\overline{W}^*} \quad (73)$$

$$\hat{\delta} = \frac{P(1 - \overline{W}^*)}{\overline{W}^*(1 - P)}. \quad (74)$$

For the Kundu Pareto process in (12) and the expression for $P(S_1^* < S_2^*)$, for the case, $\alpha > \beta$, the following two moment equations can be solved for estimates of α and β . We also know that a consistent estimator for σ is

$$\hat{\sigma} = S_{(1)} = \min\{S_1, \dots, S_m\}.$$

Now, using the following two moment equations

$$\overline{S}^* = \sigma \frac{\alpha + \beta}{\alpha + \beta - 1} \quad (75)$$

$$P = \frac{\alpha + \beta}{2\alpha + \beta}. \quad (76)$$

We obtain the following consistent estimates.

$$\hat{\alpha} = \frac{(1 - P)\overline{S}^*}{P(\overline{S}^* - \hat{\sigma})} \quad (77)$$

$$\hat{\beta} = \frac{\hat{\alpha}(2P - 1)}{1 - P}. \quad (78)$$

Now, for the case $\alpha < \beta$, the two moment equations are

$$\overline{S}^* = \sigma \frac{\alpha + \beta}{\alpha + \beta - 1} \quad (79)$$

$$P = \frac{\beta}{2\beta + \alpha}. \quad (80)$$

We obtain the following consistent estimates.

$$\hat{\alpha} = \frac{\overline{S}^*(1 - 2P)}{(1 - P)(\overline{S}^* - \hat{\sigma})} \quad (81)$$

$$\hat{\beta} = \frac{\hat{\alpha}P}{(1 - 2P)}. \quad (82)$$

For an A-M Pareto process in (13) and the expression for $P(T_1^* < T_2^*)$, the following two moment equations can be solved for estimates of α and δ . We also know that consistent estimator for σ is

$$\hat{\sigma} = S_{(1)} = \min\{S_1, \dots, S_m\}$$

Now,

$$\overline{T^*} = \sigma \frac{\alpha}{\alpha - 1} \quad (83)$$

$$P = \frac{\delta}{\alpha + \delta}. \quad (84)$$

We obtain the following consistent estimates.

$$\hat{\delta} = \frac{\overline{T^*} P}{(\overline{T^*} - \hat{\sigma})(1 - P)} \quad (85)$$

$$\hat{\alpha} = \frac{\hat{\delta}(1 - P)}{P}. \quad (86)$$

6. Applications

In the following two subsections, we illustrate a simulation study and give examples of real-life applications of the two CPFD-processes given in (8) and (9) and two Pareto processes given in (12) and (13).

6.1. Simulation study

For the processes given in equation (8) & (9) and (12)& (13) for the same moment estimators given in Section 5, we illustrate a simulation procedure on the behavior of estimators by varying the parameter values with increasing sample path sizes.

We have simulated 2000 data sets of sample path size $m = 20, 30, 50, 100, 200, 500$ for the parameter vectors $(\alpha = 0.5, \beta = 0.1)$ for the process in (8) and $(\alpha = 1, \delta = 0.1)$ for the process in (9). Similarly, for the same number of data sets with the same varying sample path sizes, we have simulated observations from the process in (12) for the parameter vectors $(\sigma = 1, \alpha = 1, \beta = 2)$ and $(\sigma = 1, \alpha = 4, \delta = 2)$ from the process in (13).

We refer to the following Figures 3–6 for the bootstrapped distribution of each parameter estimate for the processes defined above. The numerical evidence suggests that as sample size increases, the moment estimates approach the true parameter values with standard errors that decrease as the sample size increases.

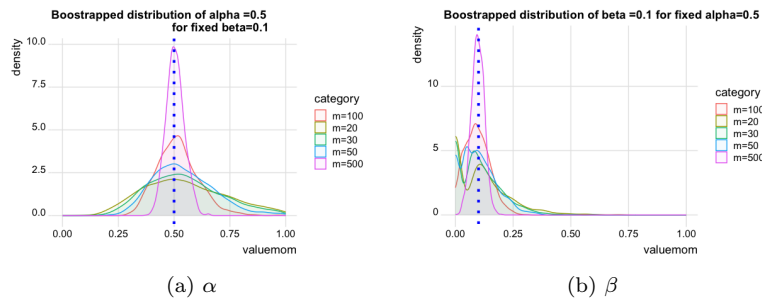


Figure 3. Kundu CPFD process.

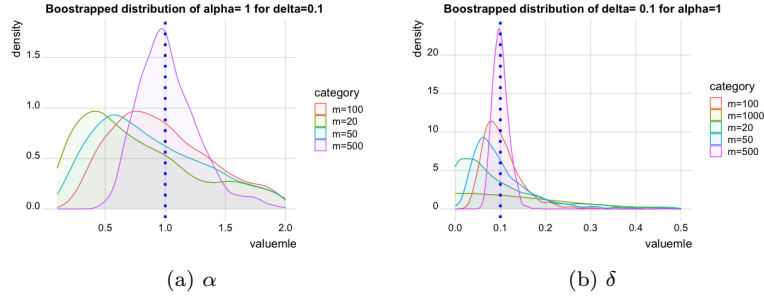


Figure 4. A-M CPFD process.

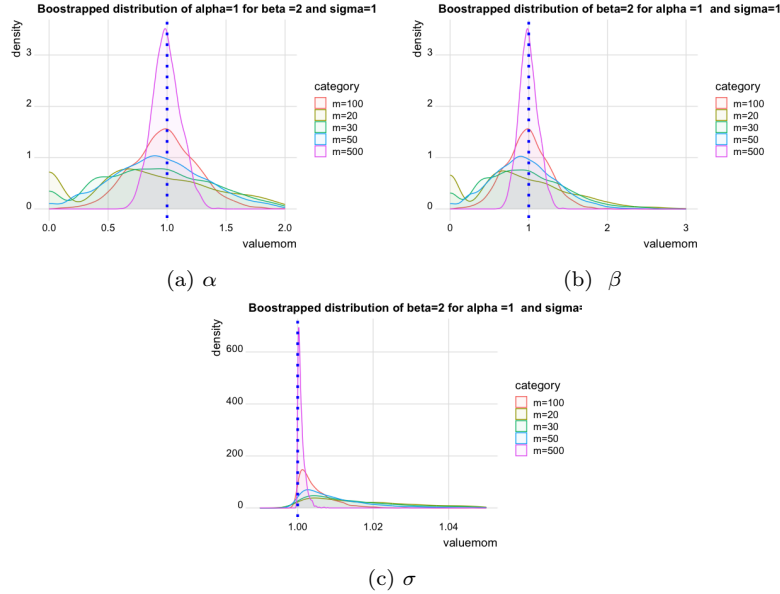


Figure 5. Kundu Pareto Process.

6.2. Real-life data

In the following, we consider the S&P data set from 3 January 1928 to 20 February 1928 of size consisting of 34 observations. The data set can be obtained from the website <https://www.kaggle.com/datasets/camnugent/sandp500>.

Here, we assume that the data generating process for the S&P data is a proportional hazard process with unknown marginal distribution F^* and underlying CPFD process of the kind in either (8) or (9). We used a non-parametric approach to estimate $1 - F^*$ by using $1 - F_m$, the complementary empirical distribution function of the $m = 34$ available data points of S&P. Now, the transformation $1 - F_m$ applied to the available S&P data will be approximately a sample from a CPFD process.

We refer to Table 1 and 2 for the fitted parameter values for the two CPFD processes and the mean square error(MSE) between the complementary empirical distribution function with the estimated CPFD processes. We can conclude that the AM CPFD Process (9) fits reasonably well for the data.

For the new Pareto processes, we make similar assumptions but without applying a distributional transformation to the available data. We refer to Table 3 and 4 for the fitted parameter values for the two new Pareto processes and the mean square error(MSE) between the actual S&P data with the estimated Pareto processes. We

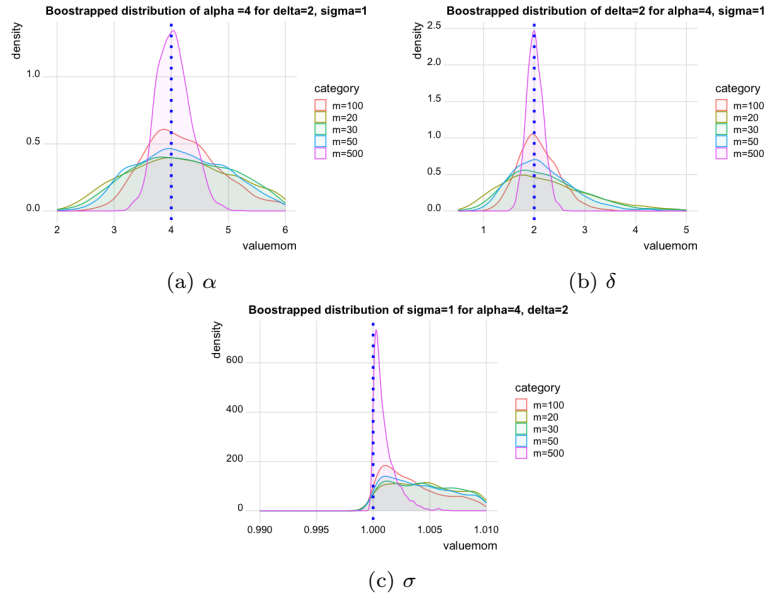


Figure 6. AM Pareto Process.

can conclude that a Kundu Pareto process (12) provides a reasonable fit for the data.

Unfortunately, there is no proper goodness of fit measure available to specifically quantify how well the processes fit the data. The problem of assessing the suitability of either of the power function process models to a real-world data set will require further investigation.

Table 1. S&P data fitted to (8) process

Parameters	Estimates
α	0.963
β	0.1203
MSE	0.212

Table 2. S&P data fitted to (9) process

Parameters	Estimates
α	1.083
δ	0.855
MSE	0.155

Table 3. S&P data fitted to (12) process

Parameters	Estimates
σ	16.950
α	7.015
β	26.307
MSE	0.257

Table 4. S&P data fitted to (13) process

Parameters	Estimates
σ	16.950
α	33.321
δ	37.487
MSE	9.636

7. Conclusion

We have considered two complementary power function distribution processes and two new Pareto processes, all stationary while some are Markovian. We have derived the marginal and bivariate distributional properties and included moment computation up to the auto-correlation function of all of the processes. Since the derived bivariate distributions do not have densities, likelihood-based inferences such as maximum likelihood and the Bayesian approach cannot be used to estimate the process parameters. However, we have derived one more variation property of each of the processes and have used it to obtain simple expressions for consistent estimators of the parameters of the processes. We also included a simulation study of the behavior of the estimators with varying sample sizes. Finally, we included an example of an application to the S&P data and illustrated the general procedure to fit such processes. The investigation of such processes is still at the beginning and will eventually merit a place in the toolkit of modelers dealing with proportional hazard processes.

Acknowledgment

The second author, S. Sachdeva's research, is Sponsored by the DST INSPIRE.

References

- [1] Arnold, B.C. (2015). Pareto Distributions (2nd ed.). Chapman and Hall/CRC.
- [2] Arnold, B.C., Sachin, S. and Manjunath, B.G. (2023), Some power function distribution processes, (Submitted).
- [3] Gupta, R.C., Gupta, P.L. and Gupta, R.D. (1998), Modeling failure time data by Lehmann alternatives, Comm. in Stat.- Theory and Methods, 27, 887–904
- [4] Kundu, Debasis. (2022), Stationary GE-Process in its applications in analyzing gold price data, Sankhya, 84-B, 575-595.