Ferdowsi University of Mashhad

Stochastic Models in Probability and Statistics

ISSN: 3060-7876

SMPS

Vol. 1, No. 1, 2024

In the Name of God

Stochastic Models in Probability and Statistics (SMPS)

Volume 1, Issue 1 July 2024

ISSN-Print: —- ISSN-Online: 3060-7876

Publisher: Faculty of Mathematical Sciences, Ferdowsi University of Mashhad
Published by: Ferdowsi University of Mashhad Press
Printing Method: Electronic

Address:
SMPS Journal, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad,
P.O. Box 1159, Mashhad 91775, Iran.

Tel.: +98-51-38806222 **Fax:** +98-51-38807358
E-mail: smps@um.ac.ir
Website: https://smps.um.ac.ir

Editorial Board

This journal is published under the auspices of Ferdowsi University of Mashhad

---

We would like to acknowledge the help of Miss Narjes Khatoon Zohorian in the preparation of this issue.

Letter from the Editor-in-Chief

It is my great pleasure to present this issue of Stochastic Models in Probability and Statistics (SMPS). As an international, peer-reviewed, open-access journal, SMPS remains committed to advancing both the theoretical foundations and practical applications of stochastic modeling, probability, and statistics—with a particular focus on reliability and related disciplines.
SMPS publishes two issues annually and is proudly supported by the Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad.

The journal's mission is to promote the development and dissemination of innovative methodologies and ideas that bridge stochastic modeling in probability and statistics with real-world applications. In this spirit, SMPS welcomes high-quality contributions across a broad range of topics, including reliability theory, lifetime data analysis, stochastic modeling in industrial and systems engineering, operations research, statistical methods in information theory, and dependence modeling.

On behalf of the editorial board, I would like to express my sincere gratitude to all authors for their valuable scholarly contributions, to the reviewers for their thoughtful and constructive evaluations, and to the editorial team for their unwavering dedication and professionalism. The continued success of SMPS is made possible by the collective efforts of this dynamic community of researchers and practitioners who share a common goal: advancing the field of stochastic modeling and statistical reliability analysis.

Finally, I warmly invite readers and contributors around the world to support SMPS by submitting original research, citing published work, and sharing new ideas that help shape and expand the journal's vision.

Jafar Ahmadi
Editor-in-Chief
Stochastic Models in Probability and Statistics (SMPS)

Contents

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 1–15. Research Paper

SMPS

A Monte Carlo simulation study of goodness-of-fit tests for Weibull distribution based on the empirical distribution function

S.Q. Alavi, H. Alizadeh Noughabi*, S. Jomhoori

Department of Statistics, University of Birjand, Birjand, Iran
Email(s): s.q.alavi@birjand.ac.ir, alizadehhadi@birjand.ac.ir, sjomhoori@birjand.ac.ir

Abstract. The Weibull distribution is widely used in reliability as a model for time to failure. In this paper, we investigate goodness-of-fit tests based on the empirical distribution function and apply them to test the validity of the Weibull model. We use the maximum likelihood estimator to estimate the scale and shape parameters of the distribution. A Monte Carlo simulation study is employed to determine the critical values and the actual size of the considered tests. The power values of the tests are computed and compared with each other. A real data example is used to illustrate the proposed tests.

Keywords: Empirical distribution function; Goodness-of-fit; Maximum likelihood estimates; Monte Carlo simulation; Test power; Weibull distribution.

1 Introduction

In the process of statistically analyzing lifetime data, the goodness-of-fit (GOF) test procedure is important to choose a distribution that adequately fits the data. The classical GOF tests are usually based on graphical analysis, moment such as skewness or kurtosis, chi-squared type, the empirical distribution function, or regression, and correlation; see Kim (2017).

Many authors have extensively studied the GOF technique, which was introduced by Karl Pearson in 1900. Those interested in a comprehensive understanding of this topic may refer to the books by D'Agostino and Stephens (1986).

There are several GOF tests based on the empirical distribution function, which are common and sophisticated, namely Cramer-von Mises (W^2), Kolmogorov-Smirnov (D), Kuiper (V),

*Corresponding author

Received: 23 October 2023 / Accepted: 9 May 2024
DOI: 10.22067/smps.2024.45177

 https://smps.um.ac.ir

Watson (U^2) and Anderson-Darling (A^2). For details, see D'Agostino and Stephens (1986). Furthermore, some different GOF tests are developed by researchers for various distributions; see for example, Chandra et al. (1981); Kim (2017); Krit (2014); Littell et al. (1979); Mann et al. (1973), and Alizadeh Noughabi and Shafaei Noughabi (2024).

Three GOF tests, which are based on the empirical distribution function, are introduced by Zhang (2002). Compared with previous competing tests to fit the normal distribution, the new tests have higher power than the classic ones. Recently, Torabi et al. (2016) investigated the new GOF test for normality based on the empirical distribution function, which proved to be effective and powerful for some alternatives. Furthermore, Torabi et al. (2018) considered a test for the exponential distribution utilizing the same test statistic in another study.

The Weibull statistical model, which was suggested by Waloddi Weibull in 1939, is one of the best statistical models in the field of reliability. Its density function is

$$f(x;\eta,\beta)=\frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1}\exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right),\quad x\geq 0,\ \eta>0,\ \beta>0.$$

It is used in many fields such as geology, chemistry, physics, medicine, environmental science, economics, geography, and engineering. See, for instance, Ghazanfari Rad and Riazi (2023); Jung and Schindler (2017) for analysis of wind speed, Garca et al. (2020) for radar square-law detection, and Klakattawi (2022) for survival analysis of cancer patients.

The aim of this paper is to compare a wide representative selection of classical and recent tests to fit the Weibull distribution. The methods, proposed by Zhang (2002) and Torabi et al. (2016) are used to construct new tests of fit for Weibull distribution.

The rest of the paper is classified as follows. In Section 2, we consider test statistics based on the empirical distribution function and apply them to the Weibull distribution. In Section 3, the critical points and the actual sizes of the tests are obtained by Monte Carlo simulations. The power values of the tests are compared with each other. Section 4 contains applications of the tests in real examples. Finally, the conclusion is discussed in section 5.

2 The test statistics

Let $X_1,\ldots,X_n$ be n independent and identically distributed random variables from a continuous distribution F, with order statistics $X_{(1)},\ldots,X_{(n)}$. The hypothesis of interest is

$$\begin{aligned} H_0: F(x) &= F_0(x;\eta,\beta), && \text{for all } x>0,\\ H_1: F(x) &\neq F_0(x;\eta,\beta), && \text{for some } x>0, \end{aligned} \tag{1}$$

where η, and β are unknown parameters and

$$F_0(x;\eta,\beta)=1-\exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right),\quad x\geq 0,\quad \eta>0,\quad \beta>0,$$

is the cumulative distribution function (CDF) of Weibull random variable, dented by $W(\eta,\beta)$. To estimate the unknown parameters, the maximum likelihood method (MLE) is used, and the

MLEs are the solution of the following equations:

$$\begin{cases} \hat{\eta}_n = \left[\frac{1}{n} \sum_{i=1}^{n} X_i^{\hat{\beta}_n} \right]^{\frac{1}{\hat{\beta}_n}}, \\ \frac{n}{\hat{\beta}_n} + \sum_{i=1}^{n} \ln X_i - \frac{n}{\sum_{i=1}^{n} X_i^{\hat{\beta}_n}} \sum_{i=1}^{n} X_i^{\hat{\beta}_n} \ln X_i = 0. \end{cases}$$

2.1 Traditional tests

The existing tests for (1) that are commonly employed in practice, are briefly explained in the following:

- Von Mises (1931) statistic (W^2)

$$\begin{aligned} W^2 &= n \int_{-\infty}^{+\infty} [F_n(x) - F_0(x)]^2 \, d\hat{F}_0(x), \\ &= \frac{1}{12n} + \sum_{i=1}^{n} \left(\frac{2i-1}{2n} - F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) \right)^2. \end{aligned}$$

- Watson (1961) statistic (U^2)

$$U^2 = W^2 - n(\bar{U} - 0.5)^2,$$

  where $\bar{U}$ is the mean of $F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right), i = 1, 2, \ldots, n.$

- The Kolmogorov (1933) statistic

$$D = \max\left(D^+, D^-\right),$$

  where

$$\begin{aligned} D^+ &= \max_{1 \le i \le n} \left\{ \frac{i}{n} - F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) \right\}, \\ D^- &= \max_{1 \le i \le n} \left\{ F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) - \frac{i-1}{n} \right\}. \end{aligned}$$

- The Kuiper (1960) statistic

$$V = D^+ + D^-.$$

- The Anderson and Darling (1954) statistic

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left\{ \ln F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) + \ln\left[1 - F_0\left(X_{(n-i+1)}; \hat{\eta}, \hat{\beta}\right)\right] \right\}.$$

2.2 Tests based on Zhang (2002)

Considering the problem of testing (1), we describe two tests based on the empirical distribution function, suggested by Zhang (2002). In short, the approach of Zhang (2002) for the Weibull distribution is described as follows. Let

$$H_t : F(t) = F_0(t;\eta,\beta),$$

and let $\bar{H}_t : F(t) \neq F_0(t;\eta,\beta)$. Hence,

$$\mathscr{H}_0 : \bigcap_{t\in(0,\infty)} H_t \quad \text{and} \quad \mathscr{H}_1 = \bigcup_{t\in(0,\infty)} \bar{H}_t.$$

According to Zhang (2002), testing $\mathscr{H}_0$ versus $\mathscr{H}_1$, is equivalent to testing H_t versus $\bar{H}_t$, for every $t \in (0,\infty)$.

Having a binary random sample, $X_{it} = I(X_i \leq t), i = 1,2,\ldots,n$, where $P(X_{it} = 1) = F(t)$ and $P(X_{it} = 0) = 1 - F(t)$, Z_t is a test statistic to testing H_t versus $\bar{H}_t$, for each fixed $t \in (0,\infty)$ and the large values of Z_t rejects H_t. In view of Zhang (2002), two test statistics for testing $\mathscr{H}_0$ verses $\mathscr{H}_1$, are given by

$$Z = \int_{-\infty}^{+\infty} Z_t dw(t) \quad \text{and} \quad Z_{\max} = \sup_{t\in(0,\infty)} [Z_t w(t)], \tag{2}$$

where $w(t)$ is some weight function and the null hypothesis will be rejected for large values of Z or $Z_{\max}$. Two well-known test statistics, namely, the Pearson's chi-squared statistic and the likelihood ratio test statistic are used as Z_t by Zhang (2002), which are, respectively,

$$\chi_t^2 = \frac{n\{F_n(t) - F_0(t)\}^2}{F_0(t)\{1 - F_0(t)\}}$$

and

$$G_t^2 = 2n\left\{F_n(t)\log\frac{F_n(t)}{F_0(t)} + [1 - F_n(t)]\log\frac{1 - F_n(t)}{1 - F_0(t)}\right\},$$

where $F_n(t)$ is the empirical distribution function of the sample.

Using χ_t^2 as Z_t in (2), with different choices of weight functions results in traditional GOF tests, for example, substituting $dw_1(t) = n^{-1}F_0(t)[1 - F_0(t)]dF_0(t)$ and $w_2(t) = F_0(t)$ in the first of equations (2) generates Cramer–von Mises and Anderson–Darling statistics, respectively, as well as $w_3(t) = n^{-1}F_0(t)[1 - F_0(t)]$, in the second which results in the famous Kolmogorov–Smirnov statistic.

Put $F_n(X_{(i)}) = \frac{i-0.5}{n}$, and select $w_4(t) = 1$, $dw_5(t) = F_0(t)^{-1}[1 - F_0(t)]^{-1}dF_0(t)$ and $dw_6(t) = F_n(t)^{-1}[1 - F_n(t)]^{-1}dF_n(t)$ as weight functions. Then G_t^2 as Z_t produces the following test statistics for the Weibull distribution, respectively,

$$
\begin{aligned}
Z_A &= -\sum_{i=1}^{n}\left(\frac{\log F_0(X_{(i)};\hat{\eta},\hat{\beta})}{n-i+0.5}+\frac{\log\left[1-F_0(X_i;\hat{\eta},\hat{\beta})\right]}{i-0.5}\right),\\
Z_c &= \sum_{i=1}^{n}\left(\log\left\{\frac{F_0(X_{(i)};\hat{\eta},\hat{\beta})^{-1}-1}{(n-0.5)/(i-0.75)-1}\right\}\right)^2,\\
Z_k &= \max_{1\le i\le n}\left((i-0.5)\log\left(\frac{i-0.5}{nF_0(X_{(i)};\hat{\eta},\hat{\beta})}\right)\right.\\
&\qquad \left. +(n-i+0.5)\log\left\{\frac{n-i+0.5}{n(1-F_0(X_{(i)};\hat{\eta},\hat{\beta}))}\right\}\right),
\end{aligned}
$$

where $\hat{\eta}$ and $\hat{\beta}$ are the MLEs. Clearly, the null hypothesis $\mathscr{H}_0$ will be rejected for large values of Z_A, Z_c, and Z_k.

2.3 Tests based on Torabi et al. (2016)

Recently, Torabi et al. (2016) defined the following discrepancy measure,

$$
D(F_0,F)=\int_{-\infty}^{+\infty} h\left(\frac{1+F_0(x)}{1+F(x)}\right)dF(x)=E_F\left[h\left(\frac{1+F_0(x)}{1+F(x)}\right)dF(x)\right],
$$

where F_0 and F are the CDF of two absolutely continuous random variables, $E_F[\cdot]$ is the expectation under F and $h:(0,\infty)\to\mathbb{R}^+$ is a continuous function, decreasing on $(0,1)$ and increasing on $(1,\infty)$ with an absolute minimum at $x=1$ such that $h(1)=0$.

Torabi et al. (2016) used this measure as a criterion of GOF, to a given distribution F_0. It is obvious that $D(F,F_0)$ can be estimated by

$$
H_n = D(F_0,F_n)=\frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_{(i)})}{1+i/n}\right).
$$

In addition, we construct a test statistic for Weibull distribution based on H_n as follows:

$$
H_n=\frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_{(i)};\hat{\eta},\hat{\beta})}{1+i/n}\right),
$$

where $F_0(x;\eta,\beta)$ is a Weibull CDF. Obviously, the null hypothesis will be rejected for large values of H_n. The following two functions suggested by Torabi et al. (2016), are considered for h,

$$
\begin{aligned}
h_1(x) &= x\log(x)-x+1,\\
h_2(x) &= \left(\frac{x-1}{x+1}\right)^2.
\end{aligned}
$$

The corresponding test statistics are

$$H_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n} h_k\left(\frac{1+F_0\left(X_{(i)};\hat{\eta},\hat{\beta}\right)}{1+i/n}\right), \quad k=1,2.$$

Not that $h_k:[0,\infty)\to \mathrm{R}^+, k=1,2$ are nonnegative functions, for which the absolute minimum is $x=1$, since $h_k(1)=0, k=1,2$. Under $\mathscr{H}_0$, we expect that $F_n(x)\approx F_0(x)$, and hence $h_k\left(\frac{1+F_0(X)}{1+F_n(x)}\right)\approx 0$. So, H_n is approximately zero as $\mathscr{H}_0$ is true. Therefore, we reject H_0 for large values of H_n.

We have the following result for all $x>0$ from (Torabi et al., 2016, Proposition 2.3):

Corollary 1.

$$0\le h_k\left(\frac{1+F_0(X)}{1+F_n(X)}\right)\le \max(h_k(1/2),h_k(2)) = \begin{cases} 0.38629 & k=1,\\ 0.1111 & k=2.\end{cases}$$

From transformed data, $H_n^{(k)}$ is the mean of $h_k(\cdot)$, which for $k=1,2$, is obtained as:

$$supp(H_n^{(1)})=[0,0.38629],\ supp(H_n^{(2)})=[0,0.11111].$$

Proposition 1. *Let F_1 be an arbitrary continuous CDF in H_1. Then under the assumption that the observed sample has CDF F_1, the test based on H_n is consistent.*

Proof. Based on the Glivenko–Cantelli theorem, for n large enough, we have $F_n(x)\approx F_1(x)$, for all $x\in R$. Therefore,

$$\begin{aligned} H_n &= \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0\left(X_{(i)};\hat{\eta},\hat{\beta}\right)}{1+F_n(X_{(i)})}\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0\left(X_i;\hat{\eta},\hat{\beta}\right)}{1+F_n(X_i)}\right)\\ &\approx \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right)\approx \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right)\\ &\to E_{F_1}\left[h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right)\right] = D(F_0,F_1), \text{as}\, n\longrightarrow\infty. \end{aligned}$$

Note that the convergence is valid according to the law of large numbers, and the divergence between F_0 and F_1 is measured by $D(F_0,F_1)$. Therefore, the test based on H_n is consistent by Torabi et al. (2016). □

Theorem 1. *The considered test statistics are invariant under the power transformations $G=\{g_c : g_c(x)=x^c\}$.*

Proof. It is enough to show that $T(X^c)=T(X)$. Let $Y=X^c$. Then

$$\begin{aligned} F_Y(y) &= P(Y\le y)=P(X^c\le y)=P\left(X\le y^{\frac{1}{c}}\right)=F_X\left(y^{\frac{1}{c}}\right),\\ &= 1-\exp\left[-\left(\frac{y^{\frac{1}{c}}}{\eta}\right)^{\beta}\right] = 1-\exp\left[-\frac{y^{\frac{\beta}{c}}}{\eta^{\beta}}\right]\sim W\left(\eta^c,\frac{\beta}{c}\right),\\ &\Rightarrow y_i=x_i^c\Rightarrow \begin{cases} \hat{\eta}_y=\hat{\eta}_x^c\\ \hat{\beta}_y=\frac{\hat{\beta}_x}{c}\end{cases} \end{aligned}$$

We have

$$F_0(x_i^c) = 1 - \exp\left[-\left(\frac{y_i}{\hat{\eta}_y}\right)^{\hat{\beta}_y}\right] = 1 - \exp\left[-\left(\frac{x_i^c}{\hat{\eta}_x^c}\right)^{\frac{\hat{\beta}_x}{c}}\right],$$

$$= 1 - \exp\left[-\left(\frac{x_i^{\hat{\beta}_x}}{\eta_x^{\hat{\beta}_x}}\right)\right] = 1 - \exp\left[-\left(\frac{x_i}{\hat{\eta}_x}\right)^{\hat{\beta}_x}\right] = F_0(x_i).$$

Consequently, $T(X^c) = T(X)$, and all test statistics based on the empirical distribution function are invariant. □

3 Simulation study

A thorough analysis of all the proposed GOF tests has been made using Monte Carlo simulation.

3.1 Critical values and type I error

Firstly, we discuss the critical values and the actual sizes of the considered tests. We can reject the null hypothesis H_0 at significance level α, if the value of the test statistic is greater than $C(\alpha)$. The critical value $C(\alpha)$ is calculated from the $(1-\alpha)$-quantile of the distribution of the test statistic under H_0.

We cannot analyze analytically the distribution of the test statistics $W^2, U^2, D, V, A^2, Z_A, Z_c, Z_k, H_n^{(1)}, H_n^{(2)}$ under the null hypothesis. Therefore, the Monte Carlo method is used to calculate the critical value of the test statistics. For each test statistic $W^2, U^2, D, V, A^2, Z_A, Z_c, Z_k, H_n^{(1)}, H_n^{(2)}$, 100,000 simulated random samples of size n are calculated from the Weibull distribution with parameters 1 and 1.
Since $\alpha = 0.05 = 5000/100000$, the 5000th order statistic is assessed to determine $C(\alpha)$. The critical values obtained for the statistics $W^2 - H_n^{(2)}$ and sample sizes $10 \le n \le 100$ are given in Table 1.

Secondly, the type I error control using the 0.05 percentile is considered in Table 2. We find the values of type I error around 5% and therefore, these values are acceptable.

3.2 Power study

For power comparison, we consider the following alternatives. These alternatives have different hazard rate: increasing hazard rate (IHR), decreasing hazard rate (DHR), bathtub shaped hazard rate (BT), and upside-down hazard rate (UBT). The usual distributions are Gamma $\mathscr{G}$, Lognormal $\mathscr{LN}$, Inverse-Gamma $\mathscr{IG}$, and Inverse-Gaussian $\mathscr{IS}$. Some distributions with CDF $F(x)$ are as follows:

- Exponentiated Weibull distribution (Mudholkar and Srivastava 1993) $\mathscr{EW}(\theta, \eta, \beta)$

$$F(x) = \left[1 - e^{-(x/\eta)^{\beta}}\right]^{\theta}, \quad \theta, \eta, \beta > 0.$$

Table 1: Critical values of the test statistics at level $\alpha = 0.05$.

n	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
10	0.1945	0.1884	0.2606	0.4373	0.7243	13.8336	7.0158	1.0934	0.003136	0.003136
20	0.2007	0.1945	0.1892	0.3181	0.7386	17.6030	9.4311	1.4670	0.001532	0.000794
30	0.2035	0.1973	0.1566	0.2641	0.7494	19.9790	10.9320	1.6911	0.001016	0.000518
40	0.2042	0.1976	0.1367	0.2301	0.7519	21.7873	12.0246	1.8413	0.000755	0.000385
50	0.2048	0.1989	0.1228	0.2073	0.7499	23.1606	12.7047	1.9436	0.000604	0.000307
60	0.2050	0.1988	0.1120	0.1894	0.7484	24.3338	13.3967	2.0456	0.000502	0.000257
70	0.2052	0.1991	0.1044	0.1761	0.7503	25.3784	13.9913	2.1279	0.000429	0.000217
80	0.2054	0.1995	0.0976	0.1649	0.7535	26.2529	14.5114	2.2006	0.000379	0.000190
90	0.2056	0.1995	0.0923	0.1558	0.7552	27.0854	14.9487	2.2666	0.000334	0.000169
100	0.2059	0.1996	0.0875	0.1480	0.7536	27.7776	15.2821	2.3113	0.000301	0.000152

Table 2: The actual size of the tests.

$W(\eta,\beta)$	n	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
$W(1,0.5)$	10	0.0504	0.0502	0.0497	0.0492	0.0500	0.0507	0.0501	0.0501	0.0494	0.0501
$W(1,0.5)$	20	0.0511	0.0514	0.0516	0.0525	0.0516	0.0501	0.0511	0.0499	0.0499	0.0498
$W(1,0.5)$	30	0.0491	0.0490	0.0499	0.0499	0.0487	0.0488	0.0502	0.0481	0.0500	0.0500
$W(1,0.5)$	40	0.0495	0.0507	0.0484	0.0500	0.0494	0.0494	0.0498	0.0485	0.0502	0.0509
$W(1,0.5)$	50	0.0490	0.0484	0.0489	0.0485	0.0499	0.0498	0.0510	0.0504	0.0498	0.0487
$W(1,1)$	10	0.0509	0.0511	0.0510	0.0516	0.0518	0.0501	0.0521	0.0528	0.0494	0.0491
$W(1,1)$	20	0.0512	0.0511	0.0513	0.0511	0.0515	0.0503	0.0508	0.0499	0.0515	0.0516
$W(1,1)$	30	0.0507	0.0505	0.0496	0.0494	0.0496	0.0494	0.0497	0.0493	0.0492	0.0486
$W(1,1)$	40	0.0495	0.0499	0.0496	0.0495	0.0498	0.0501	0.0507	0.0492	0.0503	0.0490
$W(1,1)$	50	0.0501	0.0497	0.0492	0.0503	0.0508	0.0516	0.0504	0.0513	0.0484	0.0500
$W(1,2)$	10	0.0523	0.0519	0.0521	0.0515	0.0529	0.0496	0.0519	0.0535	0.0500	0.0488
$W(1,2)$	20	0.0512	0.0514	0.0518	0.0504	0.0505	0.0489	0.0499	0.0494	0.0511	0.0501
$W(1,2)$	30	0.0485	0.0481	0.0492	0.0489	0.0491	0.0509	0.0498	0.0506	0.0494	0.0503
$W(1,2)$	40	0.0498	0.0505	0.0497	0.0499	0.0499	0.0500	0.0518	0.0498	0.0507	0.0481
$W(1,2)$	50	0.0504	0.0502	0.0497	0.0509	0.0507	0.0509	0.0506	0.0521	0.0487	0.0514
$W(1,4)$	10	0.0512	0.0503	0.0511	0.0504	0.0512	0.0506	0.0521	0.0521	0.0498	0.0502
$W(1,4)$	20	0.0503	0.0508	0.0503	0.0511	0.0499	0.0495	0.0498	0.0486	0.0513	0.0503
$W(1,4)$	30	0.0487	0.0481	0.0490	0.0491	0.0492	0.0514	0.0511	0.0506	0.0507	0.0492
$W(1,4)$	40	0.0499	0.0501	0.0498	0.0496	0.0495	0.0501	0.0508	0.0488	0.0501	0.0493
$W(1,4)$	50	0.0515	0.0511	0.0511	0.0522	0.0519	0.0507	0.0507	0.0515	0.0482	0.0499

- Generalized Gamma distribution (Stacy 1962) $\mathscr{G}\mathscr{G}(\kappa,\eta,\beta)$

$$F(x) = \frac{1}{\Gamma(k)}\gamma\left(k,(x/\eta)^{\beta}\right),\ \ k,\eta,\beta > 0,$$

where $\gamma(s,x) = \int_0^x v^{s-1}e^{-v}dv$.

- Distribution I of Dhillon (1981) $\mathscr{D}1(\beta,b)$ with CDF:

$$F(x) = 1 - e^{-\left[e^{(\beta x)^b}-1\right]},\ \ b,\beta > 0.$$

- Distribution II of Dhillon (1981) $\mathscr{D}2(\lambda,b)$ with CDF:

$$F(x) = 1 - e^{-(\ln(\lambda x+1))^{b+1}},\ \ \lambda > 0, b \geq 0.$$

- Hjorth (1980) distribution $\mathscr{H}(\beta,\delta,\theta)$ with CDF:

$$F(x) = 1 - \frac{e^{-\frac{\delta x^2}{2}}}{(1+\beta x)^{\theta/\beta}},\ \ \beta,\delta,\theta > 0.$$

- Chen (2000) distribution $\mathscr{C}(\lambda,\beta)$ with CDF:

$$F(x) = 1 - e^{\lambda\left(1-e^{x^{\beta}}\right)},\ \ \lambda,\beta > 0.$$

In short, the alternative distributions are categorized by the shape of their hazard rate in Table 3.

Table 3: Alternative distributions.

IHR	$\mathscr{G}(2)\equiv\mathscr{G}(2,1)$	$\mathscr{G}(3)\equiv\mathscr{G}(3,1)$	$\mathscr{EW}1\equiv\mathscr{EW}(6.5,20,6)$
	$\mathscr{D}2(2)\equiv\mathscr{D}2(1,2)$		
UBT	$\mathscr{LN}(0.8)\equiv\mathscr{LN}(0,0.8)$	$\mathscr{IG}(3)\equiv\mathscr{IG}(3,1)$	$\mathscr{EW}4\equiv\mathscr{EW}(4,12,0.6)$
	$\mathscr{IS}(0.25)\equiv\mathscr{IS}(1,0.25)$	$\mathscr{IS}(4)\equiv\mathscr{IS}(1,4)$	
DHR	$\mathscr{G}(0.2)\equiv\mathscr{G}(0.2,1)$	$\mathscr{EW}2\equiv\mathscr{EW}(0.1,0.01,0.095)$	$\mathscr{H}(0)\equiv\mathscr{H}(0,1,1)$
	$\mathscr{D}2(2)\equiv\mathscr{D}2(1,0)$		
BT	$\mathscr{EW}3\equiv\mathscr{EW}(0.1,100,5)$	$\mathscr{GG}1\equiv\mathscr{GG}(0.1,1,4)$	$\mathscr{GG}2\equiv\mathscr{GG}(0.2,1,3)$
	$\mathscr{C}(0.4)\equiv\mathscr{C}(2,0.4)$	$\mathscr{D}1(0.8)\equiv\mathscr{D}1(1,0.8)$	

Monte Carlo simulation is used to calculate the power of the tests W^2, D, V, U^2, A^2, Z_A, Z_c, Z_k, $H_n^{(1)}$, $H_n^{(2)}$. This process is done 100000 times for samples 10, 20, and 50. The power of the corresponding test was estimated by the frequency of the event the test statistics is larger than the critical point. The power of tests is presented in Tables 4–6. The maximal power is shown in bold type for each alternative.

Tables 4–6 show that there is no single test can be considered as the best against all alternatives. Nevertheless, when considering most alternatives, the tests U^2, Z_A, and Z_c statistics have the most power.

In addition, one notable aspect is that the effectiveness of the tests is closely connected to the pattern of the hazard rate in the simulated distribution. The patterns seem to exhibit similar behavior for the DHR and BT hazard rates for large sizes of n, as well as for the IHR

Table 4: Power results of the tests for $n = 10$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0538	**0.0554**	0.0531	0.0550	0.0493	0.0261	0.0495	0.0478	0.0503	0.0496
$\mathscr{G}(3)$	0.0599	**0.0616**	0.0575	0.0606	0.0534	0.0179	0.0546	0.0493	0.0511	0.0514
$\mathscr{EW}1$	0.0809	**0.0824**	0.0738	0.0787	0.0719	0.0085	0.0779	0.0602	0.0502	0.0496
$\mathscr{D}2(2)$	0.0664	**0.0672**	0.0620	0.0657	0.0619	0.0284	0.0664	0.0566	0.0509	0.0500
UBT										
$\mathscr{LN}(0.8)$	0.1110	0.1120	0.0959	0.1038	0.1004	0.0036	**0.1139**	0.0806	0.0502	0.0494
$\mathscr{IG}(3)$	0.2187	0.2170	0.1747	0.1998	0.2072	0.0008	**0.2383**	0.1582	0.0496	0.0506
$\mathscr{EW}4$	0.0563	0.0538	0.0571	0.0527	0.0626	**0.0812**	0.0645	0.0653	0.0515	0.0504
$\mathscr{IS}(0.25)$	0.1835	0.1826	0.1496	0.1690	0.1705	0.0005	**0.1961**	0.1289	0.0513	0.0488
$\mathscr{IS}(4)$	0.1112	0.1116	0.0961	0.1036	0.1006	0.0029	**0.1131**	0.0790	0.0504	0.0483
DHR										
$\mathscr{G}(0.2)$	0.0982	0.0882	0.0930	0.0800	0.1208	**0.1533**	0.1226	0.1196	0.0512	0.0503
$\mathscr{EW}2$	0.0516	0.0516	0.0522	0.0513	0.0524	0.0520	0.0516	**0.0529**	0.0503	0.0490
$\mathscr{H}(0)$	0.0509	0.0508	0.0511	0.0518	0.0513	0.0486	0.0511	**0.0520**	0.0503	0.0505
$\mathscr{D}2(0)$	0.1488	0.1484	0.1238	0.1370	0.1393	0.0097	**0.1562**	0.1098	0.0507	0.0511
BT										
$\mathscr{EW}3$	0.0806	**0.0812**	0.0703	0.0779	0.0720	0.0085	0.0752	0.0587	0.0500	0.0480
$\mathscr{GG}1$	0.1307	0.1161	0.1202	0.1017	0.1615	**0.1834**	0.1608	0.1539	0.0504	0.0506
$\mathscr{GG}(2)$	0.0979	0.0876	0.0926	0.0780	0.1195	**0.1530**	0.1218	0.1175	0.0503	0.0492
$\mathscr{C}(0.4)$	0.0582	0.0549	0.0594	0.0539	0.0659	**0.0840**	0.0654	0.0687	0.0517	0.0482
$\mathscr{D}1(0.8)$	0.0685	0.0621	0.0678	0.0595	0.0800	**0.1087**	0.0832	0.0841	0.0512	0.0514

and UBT hazard rates, with a few exceptions.The powerful tests for against different hazard rate are shown in Table 7. In the subsequent analysis, we evaluate the GOF tests within each specific group. The powerful tests for all shape of hazard rates are the test statistics Z_A and Z_c.

4 Application to real data sets

In this section, we present the application of GOF tests on industrial data to fit with Weibull distribution. We consider two teal data examples as follows.

Example 1. Datsiou and Overend (2018) considered the concerns glass surface strength data, and the number of data is $n = 18$. The data are
24.12, 32.98, 39.71, 24.13, 35.91, 49.10, 28.52, 35.92, 52.43, 29.18, 36.38, 52.46, 29.67, 37.60, 52.61, 30.48, 37.70, 61.72.

Example 2. The second data set is the waiting time of $n = 100$ bank customers using by Ghitany et al. (2007). The date are
0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1,

Table 5: Power results of the tests for $n = 20$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0625	**0.0634**	0.0581	0.0610	0.0590	0.0179	0.0621	0.0488	0.0487	0.0516
$\mathscr{G}(3)$	0.0770	0.0775	0.0670	0.0722	0.0729	0.0093	**0.0817**	0.0590	0.0489	0.0521
$\mathscr{EW}1$	0.1298	0.1284	0.1055	0.1143	0.1308	0.0024	**0.1582**	0.1024	0.0498	0.0499
$\mathscr{D}2(2)$	0.0895	0.0903	0.0755	0.0833	0.0893	0.0217	**0.1049**	0.0732	0.0509	0.0517
UBT										
$\mathscr{LN}(0.8)$	0.2084	0.2034	0.1610	0.1791	0.2167	0.0003	**0.2636**	0.1731	0.0512	0.0511
$\mathscr{IG}(3)$	0.4645	0.4526	0.3525	0.4142	0.4877	0.0000	**0.5623**	0.4301	0.0489	0.0509
$\mathscr{EW}4$	0.0627	0.0582	0.0614	0.0551	0.0717	**0.1028**	0.0776	0.0751	0.0499	0.0522
$\mathscr{IS}(0.25)$	0.3930	0.3787	0.2878	0.3421	0.4147	0.0000	**0.4976**	0.3698	0.0486	0.0503
$\mathscr{IS}(4)$	0.2091	0.2025	0.1604	0.1795	0.2179	0.0002	**0.2691**	0.1760	0.0510	0.0511
DHR										
$\mathscr{G}(0.2)$	0.1641	0.1412	0.1411	0.1155	0.2006	**0.2264**	0.2128	0.1824	0.0502	0.0501
$\mathscr{EW}2$	0.0499	0.0495	0.0484	0.0495	0.0500	**0.0556**	0.0510	0.0496	0.0504	0.0515
$\mathscr{H}(0)$	0.0500	0.0500	0.0495	0.0495	0.0504	0.0506	**0.0518**	0.0496	0.0501	0.0506
$\mathscr{D}2(0)$	0.2986	0.2951	0.2300	0.2642	0.3077	0.0035	**0.3446**	0.2458	0.0507	0.0497
BT										
$\mathscr{EW}3$	0.1306	0.1289	0.1055	0.1147	0.1328	0.0024	**0.1563**	0.1018	0.0484	0.0492
$\mathscr{GG}1$	0.2376	0.2064	0.1966	0.1661	0.2927	0.2755	**0.3031**	0.2448	0.0485	0.0512
$\mathscr{GG}(2)$	0.1639	0.1408	0.1406	0.1149	0.2016	**0.2270**	0.2134	0.1825	0.0489	0.0517
$\mathscr{C}(0.4)$	0.0671	0.0619	0.0635	0.0579	0.0764	**0.1068**	0.0804	0.0771	0.0503	0.0507
$\mathscr{D}1(0.8)$	0.0904	0.0797	0.0834	0.0703	0.1076	**0.1513**	0.1193	0.1083	0.0500	0.0509

7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23, 27, 31.6, 33.1, 38.5.

Figure 1 shows the empirical distribution function of the considered data set of glass and the waiting time of bank customers, respectively. We apply the GOF test for the above data examples. The values of each test statistics are computed and then compared with corresponding critical value at significance level 0.05. The results are shown in Table 8.

The results of Table 8 show that the value of each test statistic is smaller than the corresponding critical value. Therefore, the Weibull hypothesis is not rejected at the significance level of 0.05. Hence, we can infer that the probability distribution of these data sets are Weibull distribution.

5 Conclusion

In this paper, we evaluated the empirical distribution function-based GOF tests for the Weibull distribution and showed that the considered tests have a good performance. Critical points of the test statistics have been computed, and then the actual sizes of the considered test have been obtained. Through Monte Carlo simulations, we have carried out an extensive power study

Table 6: Power results of the tests for $n = 50$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0849	0.0833	0.0748	0.0764	0.0875	0.0106	**0.1014**	0.0765	0.0498	0.0487
$\mathscr{G}(3)$	0.1256	0.1208	0.1043	0.1071	0.1347	0.0035	**0.1618**	0.1177	0.0495	0.0481
$\mathscr{EW}1$	0.2858	0.2702	0.2108	0.2317	0.3225	0.0001	**0.3920**	0.2902	0.0497	0.0483
$\mathscr{D}2(2)$	0.1609	0.1600	0.1227	0.1409	0.1781	0.0125	**0.2134**	0.1484	0.0509	0.0488
UBT										
$\mathscr{LN}(0.8)$	0.4981	0.4739	0.3698	0.4240	0.5608	0.0000	**0.6539**	0.5373	0.0486	0.0487
$\mathscr{IG}(3)$	0.8822	0.8661	0.7625	0.8418	0.9184	0.0000	**0.9559**	0.9226	0.0505	0.0501
$\mathscr{EW}4$	0.0841	0.0758	0.0778	0.0676	0.0974	**0.1360**	0.1052	0.1010	0.0518	0.0482
$\mathscr{IS}(0.25)$	0.8396	0.8127	0.6825	0.7908	0.8951	0.0000	**0.9558**	0.9285	0.0511	0.0489
$\mathscr{IS}(4)$	0.5030	0.4730	0.3705	0.4258	0.5690	0.0000	**0.6766**	0.5656	0.0495	0.0488
DHR										
$\mathscr{G}(0.2)$	0.3848	0.3315	0.3097	0.2606	0.4626	0.3573	**0.4846**	0.3825	0.0493	0.0477
$\mathscr{EW}2$	0.0523	0.0521	0.0515	0.0510	0.0534	**0.0564**	0.0524	0.0518	0.0499	0.0495
$\mathscr{H}(0)$	0.0504	0.0505	0.0498	0.0496	0.0505	0.0494	**0.0515**	0.0496	0.0494	0.0488
$\mathscr{D}2(0)$	0.6509	0.6404	0.5292	0.5918	0.6910	0.0005	**0.7194**	0.6150	0.0508	0.0486
BT										
$\mathscr{EW}3$	0.2891	0.2743	0.2162	0.2360	0.3264	0.0002	**0.3944**	0.2917	0.0500	0.0486
$\mathscr{GG}1$	0.5773	0.5131	0.4608	0.4144	0.6791	0.4378	**0.7136**	0.5694	0.0512	0.0484
$\mathscr{GG}(2)$	0.3867	0.3322	0.3106	0.2618	0.4662	0.3590	**0.4876**	0.3847	0.0499	0.0482
$\mathscr{C}(0.4)$	0.1023	0.0921	0.0917	0.0809	0.1220	**0.1436**	0.1267	0.1150	0.0512	0.0470
$\mathscr{D}1(0.8)$	0.1708	0.1458	0.1449	0.1191	0.2089	0.2210	**0.2231**	0.1891	0.0490	0.0489

Table 7: The powerful test statistics in various shape of hazard rate.

IHR	UBT	DHR	BT
U^2 & Z_c	Z_c & Z_A	Z_A, Z_c & Z_k	Z_A & Z_c

on the considered tests. It is shown that some of the tests outperform in most cases all other tests. Finally, we have used a real data set and have illustrated how the considered test can be applied to test the GOF for the Weibull distribution when a random sample is available.

References

Alizadeh Noughabi, H., and Shafaei Noughabi, M. (2024). Monte Carlo comparison of goodness-of-fit tests for the Inverse Gaussian distribution based on empirical distribution function. *Journal of Mahani Mathematical Research Center*, **13** 71–84.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of American Statistical Association*, **49**, 765–769.

Table 8: Results for Examples 1 and 2.

Example 1				Example 2		
Test	Value of the test statistic	Critical value	Decision	Value of the test statistic	Critical value	Decision
W^2	0.195545	0.2001545	Not reject H_0	0.1435295	0.2059	Not reject H_0
D	0.1882295	0.1942104	Not reject H_0	0.1362945	0.1996	Not reject H_0
V	0.1958956	0.1988176	Not reject H_0	0.05779042	0.0875	Not reject H_0
U^2	0.3127805	0.3342664	Not reject H_0	0.1033005	0.1480	Not reject H_0
A^2	0.659343	0.7384829	Not reject H_0	0.4056094	0.7536	Not reject H_0
Z_A	10. 119	16.99821	Not reject H_0	18.8458	27.7776	Not reject H_0
Z_c	7.40204	9.100983	Not reject H_0	9.631016	15.2821	Not reject H_0
Z_k	1.027094	1.40989	Not reject H_0	0.9548925	2.3113	Not reject H_0
H_n^1	0.000679278	0.001706246	Not reject H_0	0.0001602	0.000301	Not reject H_0
H_n^2	0.000412255	0.000881864	Not reject H_0	0.0000710	0.000152	Not reject H_0

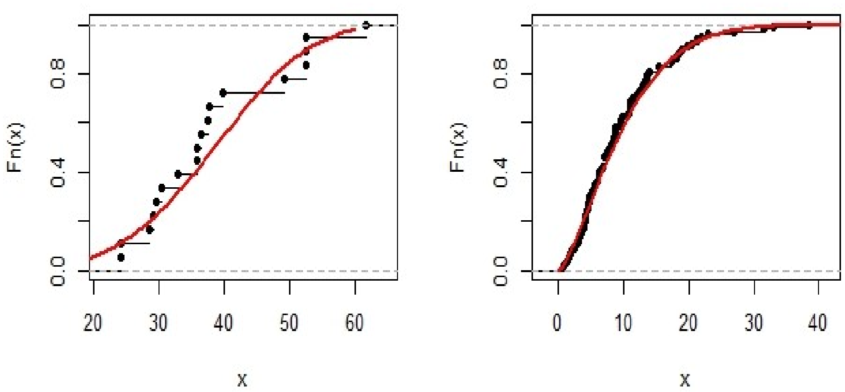


Figure 1: Empirical cumulative distribution of data in Examples 2 and 1.

Chandra, M., Singpurwalla, N. D., and Stephens, M. A. (1981). Kolmogorov statistics for tests of fit for the extreme-value and Weibull distributions. *Journal of the American Statistical Association*, **76**, 729–731.

Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, **49**, 155–161.

D'Agostino, R. B. and Stephens, M. A. (Eds.) (1986). *Goodness-of-fit Techniques.* Marcel Dekker, New York.

Datsiou, K. C., and Overend, M. (2018). Weibull parameter estimation and goodness-of-fit for glass strength data. *Structural Safety*, **73**, 29–41.

Dhillon, B. S. (1981). Life distributions. *IEEE Transactions on Reliability*, **30**, 457–460.

Garca, F. D. A., Mora, H. R. C., Fraidenraich, G., and Santos Filho, J. C. S. (2020). Square-law detection of exponential targets in Weibull-distributed ground clutter. *IEEE Geoscience and Remote Sensing Letters*, **18**, 1956–1960.

Ghazanfari-Rad, S., and Riazi, Z. (2023). Wind speed forecasting and probability distribution analysis using measured data at weather stations in Iran. *In 2023 8th International Conference on Technology and Energy Management (ICTEM)*, (pp. 1–4). IEEE.

Ghitany, M. E., Al-Awadhi, F. A., and Alkhalfan, L. (2007). Marshall-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics-Theory and Methods*, **36**, 1855–1866.

Hjorth, U. (1980). A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics*, **22**, 99–107.

Jung, C., and Schindler, D. (2017). Global comparison of the goodness-of-fit of wind speed distributions. *Energy Conversion and Management*, **133**, 216–234.

Kim, N. (2017). Goodness-of-fit tests for randomly censored Weibull distributions with estimated parameters. *Communications for Statistical Applications and Methods*, **24**, 519–531.

Klakattawi, H. S. (2022). Survival analysis of cancer patients using a new extended Weibull distribution. *Plos one*, **17**, e0264229.

Kolmogorov, A. N. (1933). Sulla Determinazione Empirica di une legge di Distribuzione. *Giornale dellIntituto Italiano degli Attuari*, **4**, 83–91.

Krit, M. (2014). Goodness-of-fit tests for the Weibull distribution based on the Laplace transform. *Journal de la Socit Franaise de Statistique*, **155**, 135–151.

Kuiper, N. H. (1960). Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, Series A, **63**, 38–47.

Littell, R. C., Mc Clave, J. T., and Offen, W. W. (1979). Goodness-of-fit tests for the two parameter Weibull distribution. *Communications in Statistics-Simulation and Computation*, **8**, 257–269.

Mann, N. R., Scneuer, E. M., and Fertig, K. W. (1973). A men goodness-of-fit test for the two-parameter Weibull or extreme-value distribution with unknown parameters. *Communications in Statistics*, **2**, 383–400.

Mudholkar, G. S., and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, **42**, 299–302.

Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 1187–1192.

Torabi, H., Montazeri, N. H., and Gran, A. (2016). A test for normality based on the empirical distribution function. *SORT: Statistics and Operations Research Transactions*, **40**, 55–88.

Torabi, H., Montazeri, N. H., and Gran, A. (2018). A wide review on exponentiality tests and two competitive proposals with application on reliability. *Journal of Statistical Computation and Simulation*, **88**, 108–139.

Von Mises, R. (1931). Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. *Deuticke*, Leipzig and Vienna.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika*, **48**, 109-114.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **64**, 281–294.

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 17–36. Research Paper

Investigation of a variable-power FGM-type copula

Christophe Chesneau*

Department of Mathematics, LMNO, University of Caen-Normandie, 14032 Caen, France
Email(s): christophe.chesneau@gmail.com

Abstract. Copula theory is essentially based on the use of multivariate functions, commonly called copulas. These functions provide a versatile toolset for capturing a broad spectrum of dependence structures, highlighting their indispensability in a multitude of applied domains. However, in light of the evolving complexities inherent in real-world data, there is a growing demand for pioneering copula constructions. In this paper, our main goal is to increase the field of variable-power copulas by introducing an innovative Farlie–Gumbel–Morgenstern (FGM)-type power copula. It is distinguished by a unique one-parameter formulation that recovers the independence copula. In the main part, we establish its mathematical validity, which is based on differentiation techniques, appropriate factorizations, and two complementary logarithmic inequalities. Then we provide a comprehensive exploration of its modeling properties, with a focus on its negative dependence through the beta medial correlation, rho of Spearman and tau of Kendall. The corresponding copula data generation is examined with different values of the parameter. A new bivariate normal distribution is also derived, and its shapes are discussed. Finally, the minimum and maximum of two random variables connected through the proposed copula are examined from a distributional viewpoint. Our findings contribute to the advancement of copula theory, thereby enhancing its practical utility across a wide range of disciplines.

Keywords: Bivariate distributions; Bivariate plots; Copula approach; Correlation measures; Data generation; Dependence model; Inequalities.

1 Introduction

Copula theory, first introduced by Sklar (1959), has transformed into an indispensable tool for analyzing multivariate data and characterizing the interrelationships among random variables. This theory is based on copulas, which are pivotal multivariate functions enabling the separation

*Corresponding author

Received: 11 October 2023 / Accepted: 4 February 2024
DOI: 10.22067/smps.2024.45178

 https://smps.um.ac.ir

of marginal distributions from the underlying dependence structure. This separation empowers practitioners to independently model and estimate joint distributions, making copulas highly adaptable and applicable in various domains such as finance (as discussed in Joe 2015), insurance (as studied in Embrechts et al. 2001), and environmental sciences (as detailed in Becherini et al. 2013), among many others. For a comprehensive examination of the theoretical and practical aspects of copulas, we recommend consulting Nelsen (2006) and Durante and Sempi (2016). Recent advancements are documented in Susam (2020a), Susam (2020b), Chesneau (2022), Chesneau (2023), Chesneau (2021b), Michimae and Emura (2022), Shih et al. (2002), El Ktaibi et al. (2022), and Yeh et al. (2023). These references collectively provide a solid foundation for better understanding and applying copula theory.

In the current context, there is a growing demand for the development of new copula constructions, driven by several compelling factors. First, existing copula families often have limitations in capturing the complexities of dependence structures marked by asymmetric behavior of tails or outliers, which are prevalent in various domains. Second, real-world data often exhibit complex, nonlinear, or nonmonotonic dependence patterns that traditional copula models struggle to accurately represent. Finally, the rise of high-dimensional datasets requires copula models capable of efficiently adapting to large dependence structures.

The development of innovative copula families to overcome these limitations is thus of paramount importance. A recent contribution to the subject is the variable-power copula family, briefly presented below. In the bivariate setting, a variable-power copula is characterized by the following general expression:

$$G(x,y) = x^{P(x,y)} y^{Q(x,y)}, \quad (x,y) \in [0,1]^2. \tag{1}$$

Here, $P(x,y)$ and $Q(x,y)$ are bivariate functions that adhere to specific conditions while satisfying the following requirements: $G(0,0) = 0$, $G(0,1) = 0$, $G(1,0) = 0$, and $G(1,1) = 1$. It is evident that when $P(x,y) = Q(x,y) = 1$, the result is the independence copula, that is, $G(x,y) = xy = \Pi(x,y)$. The investigation of scenarios, where $P(x,y) = 1$ and $Q(x,y)$ is a function exclusively dependent on x, was explored in Chesneau (2023). This exploration led to the discovery of several new bivariate copulas with varying power, providing new insights into dependence patterns that extend beyond the conventional framework.

On the other hand, the Farlie-Gumbel-Morgenstern (FGM) copula can be considered as a cornerstone of copula theory. Indeed, it has received considerable attention in recent research due to its versatility in modeling multivariate dependence structure. This widespread interest has stimulated innovative developments in our understanding and use of copulas. From a mathematical point of view, the FGM copula is defined by the following expression:

$$G^{FGM}(x,y) = xy[1 + a(1-x)(1-y)], \quad (x,y) \in [0,1]^2,$$

with $a \in [-1,1]$. The primary advantage of the FGM copula lies in its simplicity, ease of manipulation, and its capacity to capture both negative and positive dependences. However, it does have limitations, notably its restricted ability to represent diverse copula density shapes and its moderate degree of dependence. Recent research efforts have been devoted to improving its flexibility by introducing modifications and extensions, thereby providing more robust tools for risk assessment, financial modeling, and various other applications. Notably, several articles, such as Morgenstern (1956), Celebioglu (1997), Bairamov et al. (2001), Rodríguez-Lallena and Úbeda-Flores (2004), Domma and Giordano (2012), Mirhosseini et al. (2015), Bairamov and Kotz (2002), Amini et al. (2011), and Chesneau (2021a), have contributed by introducing various modified versions of the FGM copula. These advancements are continually expanding the horizons of copula theory and its practical applications across diverse fields.

In this article, we introduce an innovative copula expression that merges the fundamental principles of variable-power and FGM copulas. Specifically, by selecting $P(x,y) = Q(x,y) = G^{FGM}(x,y)/(xy)$ in (1), our proposed copula exhibits the following distinctive expression:

$$G(x,y) = x^{G^{FGM}(x,y)/(xy)} y^{G^{FGM}(x,y)/(xy)}, \quad (x,y) \in [0,1]^2,$$

which can also be represented in a more concise manner as follows:

$$G(x,y) = (xy)^{1+a(1-x)(1-y)}, \quad (x,y) \in [0,1]^2.$$

It is worth noting that, at this stage, the parameter a remains undetermined, with the exception of the trivial case where $a = 0$ giving the independence copula. The proposed copula stands out primarily due to its novel expression, which, to the best of our knowledge, has not been examined previously. In the main result, we determine a range of values for a making it valid in the mathematical sense. The proof is based on differentiation, appropriate factorizing, and two complementary logarithmic inequalities. Then we investigate its properties through a comprehensive analysis that encompasses analytical, graphical, and numerical methods. A focus is put on diverse kinds of negative dependence. In addition, the data generation based on the proposed copula is described, and a new bivariate normal distribution is derived. Finally, a distributional perspective is used to analyze the minimum and maximum of two random variables connected by the suggested copula.

The article is structured as follows: In Section 2, we introduce the variable-power FGM-type copula. Section 3 delves into an in-depth exploration of its key characteristics. Section 4 contains complementary studies to highlight some computational or distributional aspects. Finally, in Section 5, we provide our concluding remarks.

2 A variable-power FGM-type copula

2.1 Notion of copula

Before introducing our novel variable-power copula, it is essential to provide the notion of copula in the standard bivariate absolutely continuous (SBAC) context.

Lemma 1. *Nelsen (2006) In the SBAC context, we characterize a copula as a bivariate function, denoted as $B(x,y)$, defined over $[0,1]^2$, which is twice continuously differentiable on $(0,1)^2$ and satisfies the following conditions:*

C1: $B(x,0) = 0$, $B(0,y) = 0$, $B(x,1) = x$, *and* $B(1,y) = y$,

C2: $\partial_{x,y}B(x,y) = \dfrac{\partial^2}{\partial x \partial y}B(x,y) \geq 0.$

In fact, the original definition of a bivariate copula considers the 2-increasing condition, which is equivalent to C2 in the SBAC setting. From a mathematical point of view, proving C2 is often more difficult than C1, as it may require complex differentiations, factorizations, and inequalities. We refer to Nelsen (2006) for all the technical details behind the notion of copula in the SBAC context. We will adopt this notion in this article.

2.2 Novel copula

The upcoming result emphasizes the variable-power FGM-type copula under consideration.

Proposition 1. *The bivariate function defined by*

$$G(x,y) = (xy)^{1+a(1-x)(1-y)}, \quad (x,y) \in [0,1]^2, \tag{2}$$

with $a \in [0,1]$, is a valid copula.

Proof. Let us examine C1 and C2 outlined in Lemma 1 with the proposed bivariate function, assuming that $a \in [0,1]$. Beginning with C1, for any $x \in [0,1]$, since $1 + a(1-x) \geq 0$, we can observe that

$$G(x,0) = (x \times 0)^{1+a(1-x)(1-0)} = 0.$$

Likewise, for any $y \in [0,1]$, we find that $G(0,y) = 0$. Furthermore, for any $x \in [0,1]$, we have

$$G(x,1) = (x \times 1)^{1+a(1-x)(1-1)} = x.$$

Similarly, for any $y \in [0,1]$, we can determine that $G(1,y) = y$. Thus, C1 is satisfied.

Now, let us delve into C2. By employing conventional differentiation principles and conducting a comprehensive factorization, we arrive at the following result:

$$\begin{aligned}
\partial_{x,y} G(x,y) =& (xy)^{a(1-x)(1-y)} \times \{ax(y-1) + a(x-1)y + axy\log(xy) \\
&+ [a(1-x)(1-y) + ax(y-1)\log(xy) + 1][a(1-x)(1-y) + a(x-1)y\log(xy) + 1]\} \\
=& (xy)^{a(1-x)(1-y)}\{ax(y-1) + a(x-1)y + axy\log(xy) + 2a(1-x)(1-y) + 1 \\
&+ a(x-1)y\log(xy) + a(y-1)x\log(xy) \\
&+ a^2(1-x)^2(1-y)^2 - a^2(1-x)^2(1-y)y\log(xy) - a^2(1-x)(1-y)^2x\log(xy) \\
&+ a^2(1-x)(1-y)xy[\log(xy)]^2\} \\
=& (xy)^{a(1-x)(1-y)}[R(x,y) + S(x,y)],
\end{aligned}$$

where

$$\begin{aligned}
R(x,y) =& ax(y-1) + a(x-1)y + axy\log(xy) + 2a(1-x)(1-y) + 1 \\
&+ a(1-x)y[-\log(xy)] + a(1-y)x[-\log(xy)]
\end{aligned}$$

and

$$\begin{aligned}
S(x,y) =& a^2(1-x)^2(1-y)^2 + a^2(1-x)^2(1-y)y[-\log(xy)] \\
&+ a^2(1-x)(1-y)^2x[-\log(xy)] + a^2(1-x)(1-y)xy[\log(xy)]^2.
\end{aligned}$$

Since $(x,y)\in(0,1)^2$ and $a\in[0,1]$, it is obvious that $(xy)^{a(1-x)(1-y)}\geq 0$. Let us now prove that we have $R(x,y)\geq 0$ and $S(x,y)\geq 0$.

With regard to $S(x,y)$, since $(x,y)\in(0,1)^2$, we have $x\geq 0$, $y\geq 0$, $1-x\geq 0$, $1-y\geq 0$, $-\log(xy)\geq 0$, and, obviously, $a^2\geq 0$ $(1-x)^2\geq 0$, $(1-y)^2\geq 0$, and $[\log(xy)]^2\geq 0$. So, as a direct sum of nonnegative main terms, we have $S(x,y)\geq 0$.

Concerning $R(x,y)$, the steps are more technical. The following logarithmic inequality is well known: $\log(1+u)\leq u$ for $u>-1$. As a result, since $(x,y)\in(0,1)^2$, we have $\log(xy)=\log[1+(xy-1)]\leq xy-1$, so $-\log(xy)\geq 1-xy$. On the other hand, the following complementary logarithmic inequality is well known too: $\log(1+u)\geq u/(1+u)$ for $u>-1$. As a result, since $(x,y)\in(0,1)^2$, we have $\log(xy)=\log[1+(xy-1)]\geq(xy-1)/[1+(xy-1)]=(xy-1)/(xy)$. Therefore, since $a\geq 0$, we have

$$\begin{aligned}R(x,y)&\geq ax(y-1)+a(x-1)y+a(xy-1)+2a(1-x)(1-y)+1\\&\quad+a(1-x)y(1-xy)+a(1-y)x(1-xy)\\&=2ax^2y^2-ax^2y-axy^2+3axy-2ax-2ay+a+1\\&=axy[xy+(1-x)(1-y)+1-a]+a(1-x)(1-y)+(1-ax)(1-ay).\end{aligned}$$

Since $a\in[0,1]$ specifically and $(x,y)\in(0,1)^2$, it is clear that $axy\geq 0$, $xy+(1-x)(1-y)\geq 0$, $1-a\geq 0$, $a(1-x)(1-y)\geq 0$ and $(1-ax)(1-ay)\geq(1-a)^2\geq 0$. Hence, we have $R(x,y)\geq 0$.

The above results imply that $\partial_{x,y}G(x,y)\geq 0$. Thus, C2 is satisfied. We conclude that $G(x,y)$ is a valid copula. □

We refer to the copula mentioned in (2) as the variable-power FGM (VPFGM) copula. To the best of our understanding, this is a new addition to the variable-power copula family, as discussed in Chesneau's earlier work (see Chesneau 2022, 2023).

Some related comments are formulated below. First, the independence copula is recovered by taking $a=0$, and we can write the VPFGM copula as

$$G(x,y)=[\Pi(x,y)]^{1+a(1-x)(1-y)},\quad (x,y)\in[0,1]^2.$$

We thus see how the independence structure is modified with a variable-power term.

We can also express it in an exponential-logarithmic form as follows:

$$G(x,y)=\exp\{[1+a(1-x)(1-y)]\log(xy)\}$$

or, eventually,

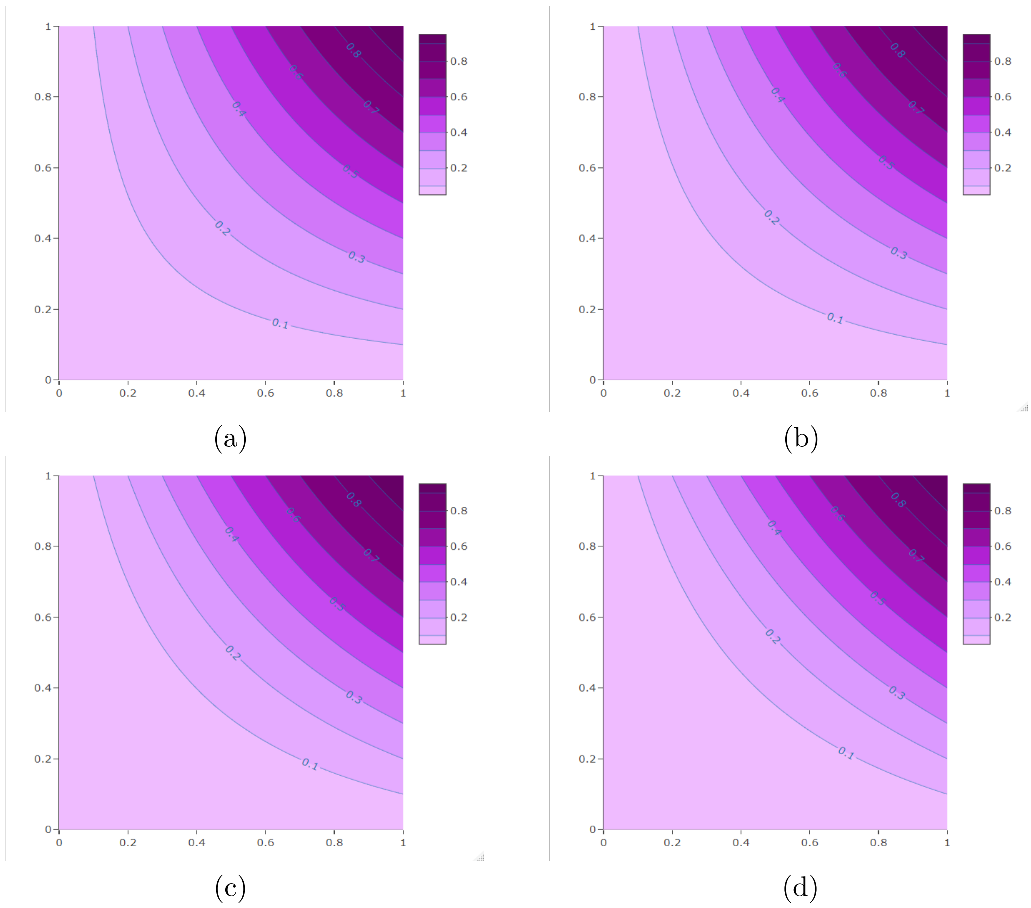
$$G(x,y)=\exp\{[1+a(1-x)(1-y)]\log(x)+[1+a(1-x)(1-y)]\log(y)\}.$$

By analyzing these expressions, the VPFGM copula is distinguished from the extreme value copula family because no function of $\log(x)/\log(xy)$ or $\log(y)/\log(xy)$ can be exhibited in the power-variable term. Furthermore, we can demonstrate that it lacks associativity, which excludes its classification in the Archimedean copula family. Another comment is that we can write it as

$$G(x,y)=xy\exp[a(1-x)(1-y)\phi(xy)],\quad (x,y)\in[0,1]^2,$$

with $\phi(t) = \log(t)$. Under this form, we remark a strong connection with the Celebioglu–Cuadras copula, which is defined with $\phi(t) = 1$ (see Celebioglu 1995). This concordance in form is intriguing and perhaps reveals an unidentified copula family.

With the use of the software R and the package `plotly` (see R Core Team 2016), Figure 1 illustrates Proposition 1 by plotting the VPFGM copula for selected values of $a \in [0,1]$.



(a) (b) (c) (d)

Figure 1: Plots of the intensity zones of the VPFGM copula for (a) $a = 0.05$, (b) $a = 0.3$, (c) $a = 0.7$, and (d) $a = 1$.

We can discern the characteristic circular intensity zones indicative of a valid copula. The remainder of the study considers the associated copula density and explores several of its pivotal characteristics.

3 Copula density and characteristics

3.1 Copula density

The copula density is a crucial component in achieving a comprehensive understanding of a copula, particularly with regard to its various shapes and their implications. In our context, the VPFGM copula density is given as

$$\begin{aligned} g(x,y) =& \partial_{x,y} G(x,y) = (xy)^{a(1-x)(1-y)} \times \{ax(y-1) + a(x-1)y + axy\log(xy) \\ & + [a(1-x)(1-y) + ax(y-1)\log(xy) + 1][a(1-x)(1-y) + a(x-1)y\log(xy) + 1]\}, \\ & (x,y) \in (0,1)^2. \end{aligned} \tag{3}$$

This function is relatively complex, primarily due to the presence of the logarithmic term $\log(xy)$. However, with the assistance of mathematical software, such as R, its implementation becomes straightforward. Figure 2 showcases the VPFGM copula density for a range of selected values of $a \in [0,1]$.

A wide panel of different colored zones is observed, highlighting the versatility of the VPFGM copula density. The most pronounced and intense zones are particularly noticeable at the coordinates $(0,1)$ and $(1,0)$. This observation becomes increasingly evident as the values of a increase.

For a deeper analysis, we can also visualize the shapes of the VPFGM copula density. Utilizing the R package `plot3D`, Figure 3 showcases some of them for selected values of $a \in [0,1]$.

By examining the relationship between Figures 2 and 3, we gain deeper insights into the behavior of the copula density, particularly its characteristics around the critical corner points.

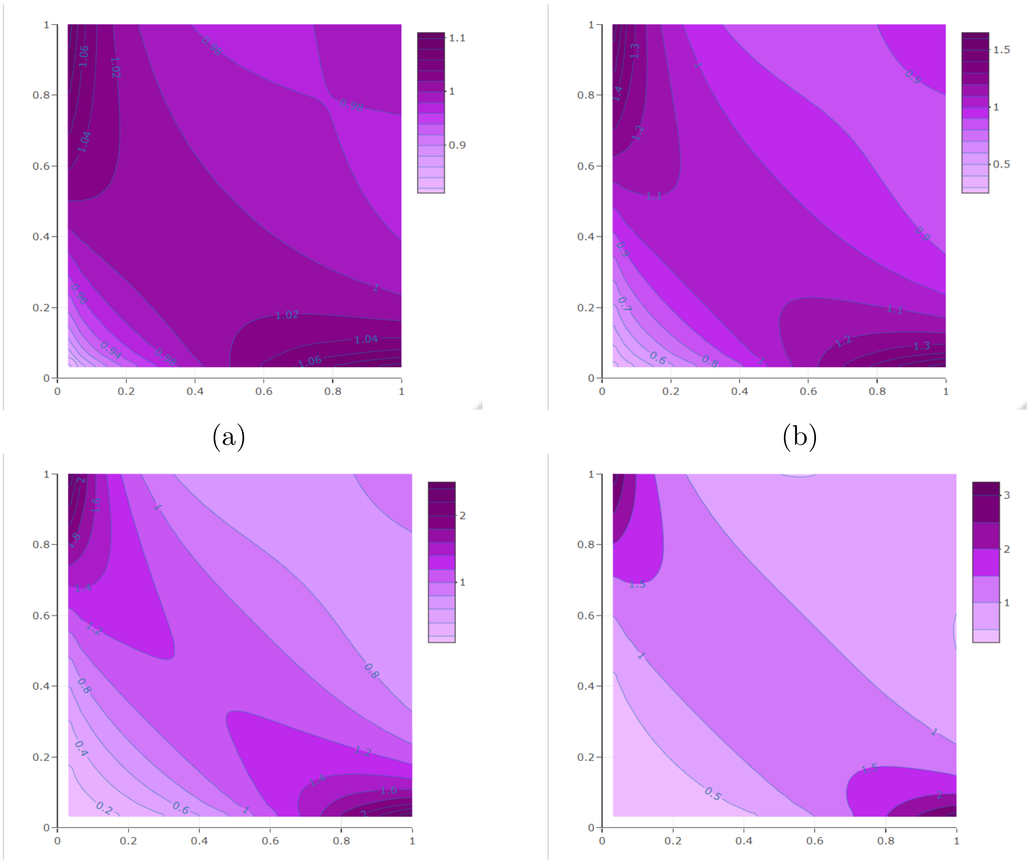
To begin a discussion involving the copula density, let us recall that Proposition 1 holds for $a \in [0,1]$. However, it is not claimed that $[0,1]$ is the optimal range of values of a. Based on Lemma 1, we can remark that C1 holds for $a > -1$. Furthermore, we have the following value for the copula density, among the rare comprehensive ones:

$$g\left(\frac{1}{2}, 1\right) = 1 - \frac{a}{2}.$$

Since a valid copula density must be nonnegative by C2, the inequality $a \leq 2$ is a necessary condition.

On the other hand, numerical tests exclude the negative values for a; in this case, we can always find $(x_\iota, y_\iota) \in (0,1)^2$ such that $g(x_\iota, y_\iota) < 0$, with x_ι approaching 1 and y_ι approaching 0; C2 is not satisfied. For example, by taking $a = -0.2$, $x_\iota = 0.999$ and $y_\iota = 0.001$, we have $g(x_\iota, y_\iota) \approx -0.178378 < 0$.

In addition, as a result of a more deep computational study, the situation is not clarified for $a \in (1,2]$; C2 is satisfied in many pointwise cases. Thus, we conjecture that the optimal range of values for a is an interval I_{op} such that $[0,1] \subseteq I_{op} \subseteq [0,2]$, but more investigations are needed for its identification.



(a) (b) (c) (d)

Figure 2: Plots of the intensity zones of the VPFGM copula density for (a) $a = 0.05$, (b) $a = 0.3$, (c) $a = 0.7$, and (d) $a = 1$.

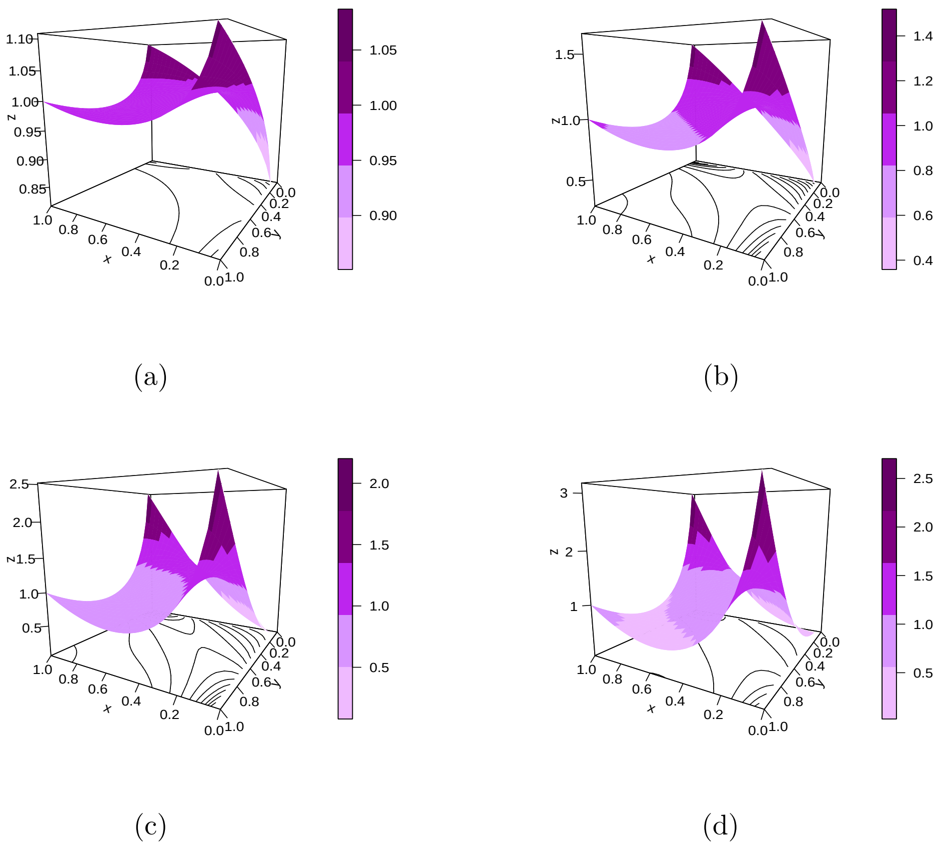


Figure 3: Plots of the shapes of the VPFGM copula density for (a) $a = 0.05$, (b) $a = 0.3$, (c) $a = 0.7$, and (d) $a = 1$.

3.2 Characteristics

We now explore the main properties of the VPFGM copula. Clearly, it is diagonally symmetric because, for any $(x,y)\in[0,1]^2$, we have $G(x,y)=G(y,x)$. Based on the well-known copula theory, the Fréchet-Hoeffding bounds hold (see Nelsen 2006). They imply the following inequalities: $\max(x+y-1,0)\leq G(x,y)\leq \min(x,y)$, that is,

$$\max(x+y-1,0)\leq (xy)^{1+a(1-x)(1-y)}\leq \min(x,y).$$

We can easily establish the upper bound by direct proof, while the lower bound seems less immediately verifiable.

Another direct and important property of a copula is the Lipschitz condition of constant 1, that is, for any $(x_1,x_2,y_1,y_2)\in[0,1]^4$, we have

$$|G(x_2,y_2)-G(x_1,y_1)|\leq |x_2-x_1|+|y_2-y_1|,$$

which, in expanded form, corresponds to

$$|(x_2y_2)^{1+a(1-x_2)(1-y_2)}-(x_1y_1)^{1+a(1-x_1)(1-y_1)}|\leq |x_2-x_1|+|y_2-y_1|.$$

On the other hand, the function $\psi(x,y)=\partial_y G(x,y)=\partial G(x,y)/(\partial y)$ is nondecreasing and satisfies $0\leq\psi(x,y)\leq 1$. Since

$$\psi(x,y)=x(xy)^{a(1-x)(1-y)}\left[1+a(1-x)(1-y)-a(1-x)y\log(xy)\right], \tag{4}$$

the following inequalities hold:

$$0\leq x(xy)^{a(1-x)(1-y)}\left[1+a(1-x)(1-y)-a(1-x)y\log(xy)\right]\leq 1.$$

Let us mention that $\psi(x,y)$ will play an important role in some coming parts, especially in the definition of integral-type correlation measures and the data generation process related to the proposed copula.

More technical properties are described below.

The VPFGM copula satisfies a notable weighted geometric result. Indeed, by setting $G(x,y;a)=G(x,y)$, for any $(x,y)\in[0,1]^2$ and $(a_1,a_2,b)\in[0,1]^3$, we have

$$\begin{aligned}[G(x,y;a_1)]^b[G(x,y;a_2)]^{1-b} &= \left[(xy)^{1+a_1(1-x)(1-y)}\right]^b\left[(xy)^{1+a_2(1-x)(1-y)}\right]^{1-b}\\ &= (xy)^{1+[ba_1+(1-b)a_2](1-x)(1-y)}=G[x,y;ba_1+(1-b)a_2].\end{aligned}$$

The tail dependence coefficients for the lower left (LL), lower right (LR), upper left (UL), and upper right (UR) directions can be calculated using the formulas provided in Nelsen (2006) and Jaworski (2023). They can be obtained through conventional limit methods in the following manner:

$$\lambda_{LL}=\lim_{x\to 0}\frac{G(x,x)}{x}=\lim_{x\to 0}x^{1+2a(1-x)^2}=\lim_{x\to 0}x^{2a+1}=0,$$

$$\lambda_{LR}=\lim_{x\to 0}\frac{x-G(1-x,x)}{x}=\lim_{x\to 0}\frac{x-[(1-x)x]^{1+ax(1-x)}}{x}=\lim_{x\to 0}x[1-a\log(x)]=0.$$

Since $G(x,y)$ is diagonally symmetric, we have

$$\lambda_{UL} = \lim_{x\to 0} \frac{x - G(x,1-x)}{x} = \lambda_{RL} = 0$$

and

$$\lambda_{UR} = \lim_{x\to 1} \frac{1-2x+G(x,x)}{1-x} = \lim_{x\to 1} \frac{1-2x+x^{2+2a(1-x)^2}}{1-x} = \lim_{x\to 1}[(1-x)-2a(1-x)^2] = 0,$$

respectively. As a result, the VPFGM copula is completely free of tail dependence, which is relatively rare for a variable-power copula (see Chesneau 2022, 2023).

For any $(x,y) \in [0,1]$ and $a \in [0,1]$, since $1+a(1-x)(1-y) \geq 1$, we have

$$G(x,y) = (xy)^{1+a(1-x)(1-y)} \leq xy,$$

which implies that the VPFGM copula is negatively quadrant dependent. Also, owing to this inequality, it is obvious that

$$G(x,y) \leq xy \leq xy[1+a(1-x)(1-y)] = G^{FGM}(x,y),$$

giving an immediate concordance ordering between the VPFGM and FGM copulas.

By using the following well-known power (Bernoulli) inequality: $(1+u)^r \geq 1+ru$ for $u \geq -1$ and $r \in \mathbb{R}/(0,1)$, then we have

$$\begin{aligned} G(x,y) = (xy)^{1+a(1-x)(1-y)} &\geq 1 + [1+a(1-x)(1-y)](xy-1) \\ &= G^{FGM}(x,y) - a(1-x)(1-y). \end{aligned}$$

We can also remark that

$$\partial_x\left(\frac{G(x,y)}{x}\right) = a\frac{1}{x}y(1-y)(xy)^{a(1-x)(1-y)}[1-x-x\log(xy)] \geq 0,$$

implying that the VPFGM copula is left tail increasing (with respect to x and y since it is diagonally symmetric). For more details on the notion of tail monotonicity, we may refer to Nelsen (2006) and Izadkhah et al. (2015).

As commented in Izadkhah et al. (2015), the FGM copula benefits from interesting likelihood ratio dependence properties. Such properties are not excluded for the VPFGM copula, but its complex copula density is an obstacle to a deep study on this aspect.

Regarding the information provided in Nelsen (2006), the expression for the beta medial correlation coefficient of the VPFGM copula is as follows:

$$\beta = 4G(0.5,0.5) - 1 = 4 \times 2^{-2-2a(1/2^2)} - 1 = 2^{-a/2} - 1.$$

It is clear that $\beta \leq 0$, implying the negative dependence feature of the VPFGM copula.

Also, based on the theory in Nelsen (2006), the basic definition of the rho of Spearman related to the VPFGM copula is

$$\begin{aligned} \rho_S &= 12\int_0^1\int_0^1 [G(x,y)-xy]\,dxdy \\ &= 12\int_0^1\int_0^1 xy\left[(xy)^{a(1-x)(1-y)}-1\right]dxdy. \end{aligned}$$

Since no direct primitive of the integrand exists and no integral technique gives a satisfying result, a numerical evaluation of ρ_S is necessary. Table 1 gives the values of ρ_S for some values of $a \in [0,1]$.

Table 1: Some values of the rho of Spearman of the VPFGM copula for $a = 0, 0.1, 0.2, \ldots, 1$.

a	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
ρ_S	0	-0.0534	-0.1028	-0.1487	-0.1917	-0.232	-0.27	-0.3059	-0.3399	-0.3722	-0.4029

From this table, we can observe that the VPFGM copula exhibits varying degrees of negative dependence, as indicated by the values of ρ_S into the interval of $[-0.4029, 0]$. Notably, this range demonstrates a larger negative dependence compared to the FGM copula, which attains values within the range of $[-0.3333, 0]$. The FGM copula permits positive dependence, whereas the VPFGM copula, for $a \in [0,1]$, lacks this property. The main distinction between the VPFGM and FGM copulas lies in their inherent characteristics and the manner in which they shape the behavior of the associated copula densities.

To complete the above study, one can also investigate the tau of Kendall of the VPFGM copula defined as

$$\tau_K = 1 - 4 \int_0^1 \int_0^1 \psi(x,y)\psi(y,x)dxdy,$$

where $\psi(x,y)$ is given in (4). A numerical exploration gives the results in Table 2.

Table 2: Some values of the tau of Kendall of the VPFGM copula for $a = 0, 0.1, 0.2, \ldots, 1$.

a	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
τ_K	0	-0.0357	-0.0691	-0.1005	-0.1302	-0.1585	-0.1854	-0.2112	-0.2359	-0.2596	-0.2825

The same conclusion drawn for the rho of Spearman holds: if we focus on the negative values, the tau of Kendall of the VPFGM copula attains the range of values $[-0.2825, 0]$ against $[-0.2222, 0]$ for the FGM copula, with still the originality arguments in terms of density copula shapes.

4 Some computational and distributional aspects

This section completes the previous study by investigating the data generation, bivariate distributions, and minimum and maximum random variables, all connected with the VPFGM copula through the subject dependence structure.

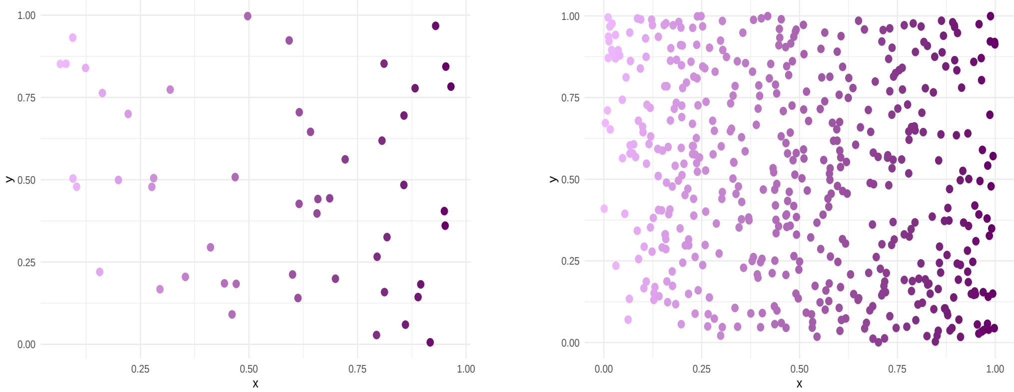
4.1 Data generation

By considering the VPFGM copula as a bivariate cumulative distribution function, we may produce random pairs of values (data) from a random vector, say (X,Y). Thus, by introducing the probability operator $\mathbb{P}$, we suppose that

$$\mathbb{P}(X \leq x, Y \leq y) = G(x,y), \quad (x,y) \in [0,1]^2.$$

Based on the data generation result in Nelsen (2006), the process to generate a single pair of values of (X,Y), say $(x_\star, y_\star)$, is described below. First, we generate a pair of independent values $(x_\star, z)$, each from random values of the uniform distribution on the interval $[0,1]$. Second, we determine $y_\star$ as the solution of the following nonlinear equation: $G_{cond}(x_\star, y_\star) = z$, where $G_{cond}(x,y) = \psi(y,x)$ as defined in (4) (with the exchange of x and y). Then the desired pair of values is $(x_\star, y_\star)$.

We can repeat this process m times to generate m such pairs of values. As an illustrative application, we generate several pairs of values for different values of m and a in Figures 4 and 5.



(a) (b)

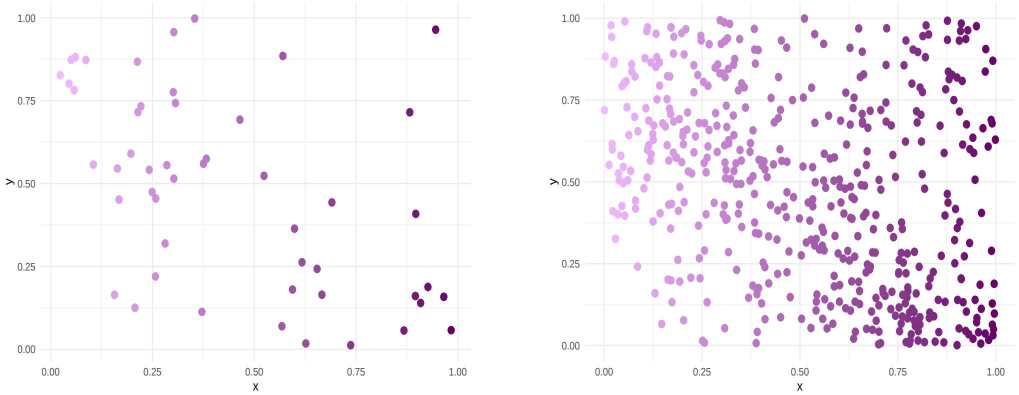
Figure 4: Examples of generated pairs of values (data) from the VPFGM copula for $a = 0.3$ and (a) $m = 50$ and (b) $m = 500$.

Various point structures, including clusters and “empty areas”, are visible in these figures, and they are all caused by the multiple dependence features of the VPFGM copula.

On the other hand, from a statistical point of view, these generated pairs of values can be used to test the effectiveness of parametric estimation methods for a. This can therefore be considered as a first step towards the statistical use of the VPFGM copula.

4.2 Bivariate distributions

The primary aim of the VPFGM copula is to create bivariate distributions. The fundamental approach is summarized below. Suppose we have two cumulative distribution functions, denoted



(a) (b)

Figure 5: Examples of generated pairs of values (data) from the VPFGM copula for $a = 0.9$ and (a) $m = 50$ and (b) $m = 500$.

as $U(x)$ and $V(y)$, representing univariate continuous distributions. We establish a novel bivariate cumulative distribution function through the following expression:

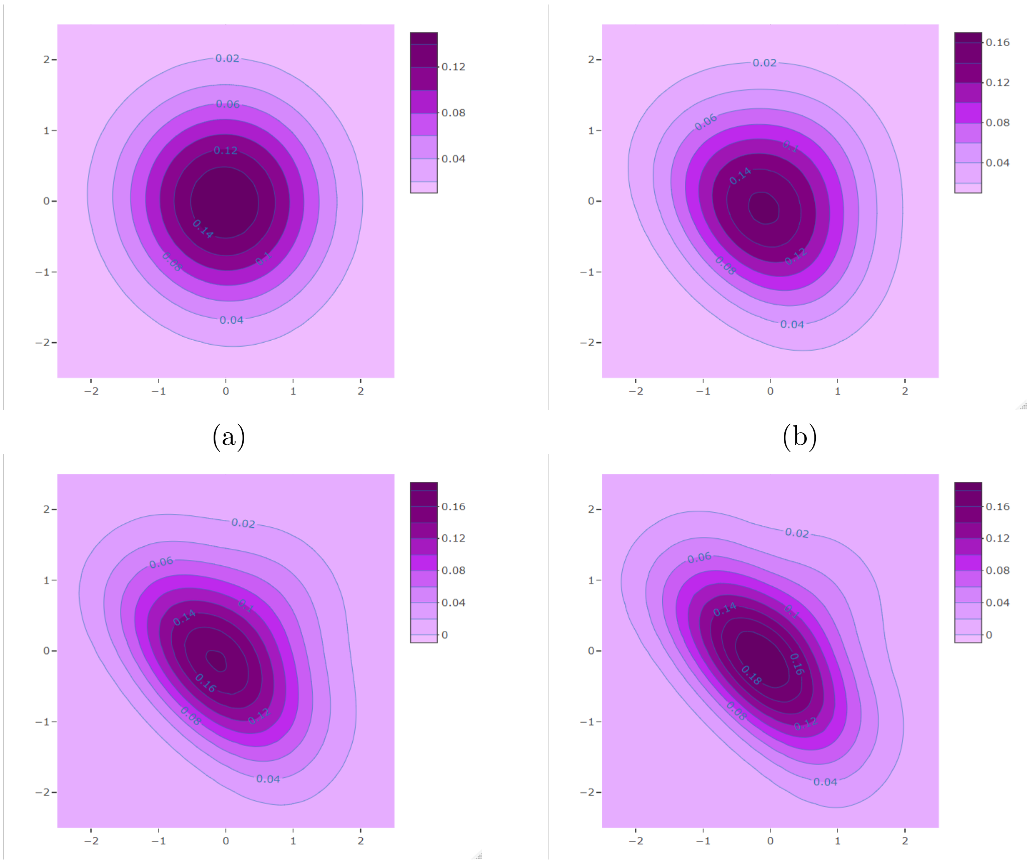
$$W(x,y) = G[U(x),V(y)] = [U(x)V(y)]^{1+a[1-U(x)][1-V(y)]}, \quad (x,y) \in \mathbb{R}^2,$$

and the corresponding probability density function is given as

$$w(x,y) = u(x)v(y)g[U(x),V(y)], \quad (x,y) \in \mathbb{R}^2,$$

where $u(x)$ and $v(y)$ are the probability density functions associated with $U(x)$ and $V(y)$, respectively, and $g(x,y)$ is the VPFGM copula density presented in (3). A novel bivariate distribution emerges from these functions. When it comes to selecting appropriate functions for $U(x)$ and $V(y)$ within the context of lifetime analysis, the comprehensive review by Taketomi et al. (2022) can be consulted. For valuable choices of $U(x)$ and $V(y)$ defined on $\mathbb{R}$, we may think to the logistic distribution (see Ali et al. 1978), or the normal distribution.

Just to give an example of a bivariate distribution defined on $\mathbb{R}^2$, let us consider $U(x)$ and $V(y)$ as both the cumulative distribution functions of the standard normal distribution, and call the related VPFGM bivariate distribution the VPFGM normal (VPFGMN) distribution. Figure 6 presents the intensity zones of the corresponding probability density function $w(x,y)$ for selected values of $a \in [0,1]$.



(a) (b) (c) (d)

Figure 6: Plots of the intensity zones of the VPFGMN probability density function for (a) $a = 0.05$, (b) $a = 0.3$, (c) $a = 0.7$, and (d) $a = 1$.

We immediately see how the circle zone of the independence bivariate standard normal distribution is deformed as the values of the parameter a increase and tends to be more dispersed at the top-right corner. This figure is the first representation of a bivariate distribution generated by the VPFGM copula.

4.3 Minimum and maximum random variables

Studying the minimum and maximum of two dependent random variables within the support $[0,1]$ is crucial to understanding extreme values in proportional type data. This analysis provides insight into the potential range of outcomes and helps identify critical scenarios. By exploring the interdependence of these variables, we can better understand the variability inherent in proportions, providing valuable information for decision-making and risk assessment in diverse fields such as finance, biology, social sciences or sport sciences. For the general context of the bivariate distributions within the support $[0,1]^2$, we may refer to Arnold and Ng (2011), Nadarajah et al. (2017), and Martínez-Flórez et al. (2022).

In the result below, we examine the probability functions associated with the minimum and maximum of two dependent random variables whose joint distribution is characterized by the VPFGM copula, which mainly governs the dependence between these random variables (the marginal distributions being in fact the uniform distribution on the interval $[0,1]$).

Proposition 2. *Let us consider a random vector (X,Y) that has the VPFGM copula as the cumulative distribution function. Then we define the minimum and maximum of X and Y by $M_{\min} = \min(X,Y)$ and $M_{\max} = \max(X,Y)$, respectively. Then the results below hold.*

- *The cumulative distribution function of $M_{\max}$ is given by*

$$F_{\max}(t) = t^{2+2a(1-t)^2}, \quad t \in [0,1],$$

(with the standard complementary values for $t \notin [0,1]$).

Furthermore, the corresponding probability density function is indicated as

$$f_{\max}(t) = 2t^{1+2a(1-t)^2}[a(1-t)^2 - 2at(1-t)\log(t) + 1], \quad t \in (0,1],$$

(with the standard complementary value for $t \notin (0,1]$).

- *The cumulative distribution function of $M_{\min}$ is given by*

$$F_{\min}(t) = 2t - t^{2+2a(1-t)^2}, \quad t \in [0,1].$$

Furthermore, the corresponding probability density function is indicated as

$$f_{\min}(t) = 2\left\{1 - t^{1+2a(1-t)^2}[a(1-t)^2 - 2at(1-t)\log(t) + 1]\right\}, \quad t \in (0,1].$$

Proof. Let us prove each item with the use of the probability operator $\mathbb{P}$.

- For any $t \in [0,1]$, we have

$$F_{\max}(t) = \mathbb{P}(M_{\max} \leq t) = \mathbb{P}(X \leq t, Y \leq t) = G(t,t) = t^{2+2a(1-t)^2}.$$

The corresponding probability density function is obtained upon differentiation with respect to t; we have

$$f_{\max}(t) = F'_{\max}(t) = 2t^{1+2a(1-t)^2}[a(1-t)^2 - 2at(1-t)\log(t) + 1].$$

- Concerning $M_{\min}$, for any $t \in [0,1]$, since $\mathbb{P}(X > t) = \mathbb{P}(Y > t) = 1 - t$, by using the previous result, we have

$$\begin{aligned} F_{\min}(t) &= \mathbb{P}(M_{\min} \leq t) = 1 - \mathbb{P}(M_{\min} > t) = 1 - \mathbb{P}(X > t, Y > t) \\ &= 1 - \mathbb{P}(X > t) - \mathbb{P}(Y > t) + \mathbb{P}(M_{\max} > t) \\ &= 1 - \mathbb{P}(X > t) - \mathbb{P}(Y > t) + 1 - F_{\max}(t) \\ &= 2t - F_{\max}(t) = 2t - t^{2+2a(1-t)^2}. \end{aligned}$$

The corresponding probability density function is obtained upon differentiation with respect to t; we have

$$f_{\min}(t) = F'_{\min}(t) = 2\left\{1 - t^{1+2a(1-t)^2}[a(1-t)^2 - 2at(1-t)\log(t) + 1]\right\}.$$

The desired functions are obtained. □

Beyond the minimum and maximum paradigms, an interest in the two exhibited distributions in Proposition 2 is that they are new in the literature and can be considered one-parameter statistical models in the field of proportional or rate data analysis. Thus, in some senses, the nature of the VPFGM copula is detoured for such univariate perspectives. Further details and applications of the univariate distributions with support $[0,1]$ can be found in Mazucheli et al. (2019), Korkmaz (2020), and Korkmaz and Chesneau (2021).

5 Conclusion

In this article, we introduced and comprehensively analyzed a novel copula that amalgamates the structural elements of variable-power and FGM-type copulas, giving rise to an innovative dependence model. Our examination demonstrated that it has a diagonal symmetry and manifests negative correlation characteristics, including negatively quadrant dependence and negative-valued beta medial correlation, among others.

Moreover, through a numerical investigation, it was determined that the rho of Spearman is confined to the interval $[-0.4029, 0]$, while the tau of Kendall is found within the range of $[-0.2825, 0]$. In addition, an adapted data generation process is described, a new bivariate normal distribution is derived, and a distributional analysis is conducted on the minimum and maximum of two random variables that are related by the suggested copula. These findings represent a slight enhancement over the corresponding negative range associated with the FGM copula. Consequently, our work makes a contribution to the advancement of variable-power-type copulas, establishing a theoretical foundation for applied research using them.

Some potential avenues of work are described below.

- The study of the following more general and possibly diagonally asymmetric version of the proposed copula:

$$G_{\dagger}(x,y) = (xy)^{1+a\theta(x)\xi(y)}, \quad (x,y) \in [0,1]^2,$$

  where $\theta(x)$ and $\xi(y)$ are functions that satisfy $\theta(1) = 0$ and $\xi(1) = 0$, and other conditions that need to be determined to make it valid in the mathematical sense. Like this, we transpose the idea in Rodríguez-Lallena and Úbeda-Flores (2004) for the FGM copula to the VPFGM copula.

- The development of a multivariate extension of the proposed copula in the following form:

$$G(x_1,x_2,\ldots,x_n) = \left(\prod_{i=1}^{n} x_i\right)^{1+a\prod_{i=1}^{n}(1-x_i)}, \quad (x_1,x_2,\ldots,x_n) \in [0,1]^n,$$

  where n denotes a positive integer. In this setting, the possible range of values on a making $G(x_1,x_2,\ldots,x_n)$ a valid copula is a challenge that we postpone for the future.

Acknowledgments

The author would like to thank the two reviewers for their thorough and constructive comments on the former version of the paper, which have significantly helped to improve it.

References

Ali, M. M., Mikhail, N. N. and Haq, M. S. (1978). A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis*, **8**, 405–412.

Amini, M., Jabbari, H. and Borzadaran, G. R. M. (2011). Aspects of dependence in generalized Farlie-Gumbel-Morgenstern distributions. *Communication in Statistics - Simulation and Computations*, **40**, 1192–1205.

Arnold, B.C. and Ng, H. K. T. (2011). Flexible bivariate beta distributions. *Journal of Multivariate Analysis*, **102**, 1194–1202.

Bairamov, I. and Kotz, S. (2002). Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions. *Metrika*, **56**, 55–72.

Bairamov, I. G., Kotz, S. and Bekci, M. (2001). New generalized Farlie-Gumbel-Morgenstern distributions and concomitants of order statistics. *Journal of Applied Statistics*, **28**, 521–536.

Becherini, A., Martina, M. L. and Marchi, M. (2013). Copula-based statistical analysis of environmental data. *Environmentalmetrics*, **24**, 1–11.

Celebioglu, S. (1995). Yeni bir copulalar ailesi, Arastirma Sempozyumu. 95, Program ve Bildiriler, Ankara, 152-153, 27–29 Kasim.

Celebioglu, S. (1997). A way of generating comprehensive copulas. *Journal of the Institute of Science and Technology*, **10**, 57–61.

Chesneau, C. (2021a). A new two-dimensional relation copula inspiring a generalized version of the Farlie-Gumbel-Morgenstern copula. *Research and Communications in Mathematics and Mathematical Sciences*, **13**, 99–128.

Chesneau, C. (2021b). On new types of multivariate trigonometric copulas. *AppliedMath*, **1**, 3–17.

Chesneau, C. (2022). On a weighted version of the Gumbel-Barnett copula. *Innovative Journal of Mathematics (IJM)*, **1(2)**, 1–13.

Chesneau, C. (2023). Theoretical validation of new two-dimensional one-variable-power copulas. *Axioms*, **12**, 392.

Domma, F. and Giordano, S. (2012). A copula-based approach to account for dependence in stress-strength models. *Statistical Papers*, **54**, 807–826.

Durante, F. and Sempi, C. (2016). *Principles of Copula Theory.* CRS Press, Boca Raton.

El Ktaibi, F., Bentoumi, R., Sottocornola, N. and Mesfioui, M. (2022). Bivariate copulas based on counter-monotonic shock method. *Risks*, **10**, 202.

Embrechts, P., Klüppelberg, C. and Mikosch, T. (2001). *Modelling extremal events: for insurance and finance* (Vol. 33). Springer Science & Business Media.

Izadkhah, S., Ahmadzade, H. and Amini, M. (2015). Further results for a general family of bivariate copulas. *Communications in Statistics-Theory and Methods*, **44**, 3146–157.

Jaworski, P. (2023). On copulas with a trapezoid support. *Dependence Modeling*, **11**, 20230101.

Joe, H. (2015) *Dependence Modeling with Copulas.* CRS Press, Boca Raton.

Korkmaz, M. Ç. (2020). The unit generalized half normal distribution: A new bounded distribution with inference and application. *UPB Scientific Bulletin, Series A: Applied Mathematics and Physics*, **82**, 133–140.

Korkmaz, M. Ç. and Chesneau, C. (2021). On the unit Burr-XII distribution with the quantile regression modeling and applications. *Computational and Applied Mathematics*, **40**, 1–26.

Martínez-Flórez, G., Lemonte, A. J., Moreno-Arenas, G. and Tovar-Falón, R. (2022). The bivariate unit-sinhnormal distribution and its related regression model. *Mathematics*, **10**, 3125.

Mazucheli, J., Menezes, A. F. B. and Dey, S. (2019). Unit-Gompertz distribution with applications. *Statistica*, **79**, 25–43.

Michimae, H. and Emura, T. (2022). Likelihood inference for copula models based on left-truncated and competing risks data from field studies. *Mathematics*, **10**, 2163.

Mirhosseini, S. M., Amini, M. and Dolati, A. (2015). On a general structure of the bivariate FGM type distributions. *Applications of Mathematics*, **60**, 91–108.

Morgenstern, D. (1956). Einfache Beispiele zweidimensionaler Verteilungen. *Mitteilungsblatt für Mathematische Statistik*, **8**, 234–235.

Nadarajah, S., Shih, S.H. and Nagar, D. K. (2017). A new bivariate beta distribution. *Statistics*, **51**, 455–474.

Nelsen, R. B. (2006). *An Introduction to Copulas.* 2nd ed. Springer Science+Business Media, Inc.: Berlin, Germany.

R Core Team (2016). R: A Language and Environment for Statistical Computing. Vienna, Austria. Available at: https://www.R-project.org/.

Rodríguez-Lallena, J. A. and Úbeda-Flores, M. (2004). A new class of bivariate copulas. *Statistics and Probability Letters*, **66**, 315–325.

Shih, J.-H., Konno, Y., Chang, Y.-T. and Emura, T. (2022). Copula-based estimation methods for a common mean vector for bivariate meta-analyses. *Symmetry*, **14**, 186.

Sklar, A. (1959). Fonctions de repartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

Susam, S. O. (2020a). Parameter estimation of some Archimedean copulas based on minimum Cramér-von-Mises distance. *Journal of The Iranian Statistical Society*, **19**, 163–183.

Susam, S. O. (2020b). A new family of Archimedean copula via trigonometric generator function. *Gazi University Journal of Science*, **33**, 806–813.

Taketomi, N., Yamamoto, K., Chesneau, C. and Emura, T. (2022). Parametric distributions for survival and reliability analyses, a review and historical sketch. *Mathematics*, **10**, 3907.

Yeh, C.-T., Liao, G.-Y. and Emura, T. (2023). Sensitivity analysis for survival prognostic prediction with gene selection: A copula method for dependent censoring. *Biomedicines*, **11**, 797.

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 37–45. Research Paper

Bayes factor: hypothesis testing in restricted Stein-rule regression analysis with autocorrelated error

Oloyede Isiaka*

Department of Statistics, University of Ilorin, Ilorin, Nigeria
Email(s): oloyede.i@unilorin.edu.ng

Abstract. Regression is a generic tool used to make inferences so as to draw logical conclusions and judgments about a particular problem. It has been widely used in engineering, sciences, education, technology, and social sciences for a long time. Breaking down of the underlying estimation mar the judgment and decision of the study. Chiefly among the error structure that can lead to the drawback of the inferences of regression is the autocorrelation of error term of both AR and MA processes. Restricted Stein-rule regression analysis was used with data injected with autocorrelated error; H_1 was modeled with autocorrelated error whereas H_0 was modeled without, alternative approach in Bayes factor of $AR(1)$ and $MA(1)$ processes were introduced and compared with Bayesian information criterion approach in both negative and positive ρ (symmetrical). The choice of the Bayes factor (Bf) is due to the probabilistic nature of Bayesian inference, which over the years has been performed better than the classical approach. The datasets with five covariates was set at 25 to capture the error structure and project the property of a small sample. The study, therefore, concluded that Bayesian inference being probabilistic about the uncertainty of the parameters should be adopted to verify the presence or absence of autocorrelated error in the data before estimation.

Keywords: Autocorrelated Error; Bayesian; Bayes factor; Hypothesis; Restricted; Stein-Rule.

1 Introduction

Han and Carlin (2001) attempted to compare several methods of estimations under the categories of simple regression, hierarchical longitudinal model, and binary data latent. They thought that the joint model parameters search space technique performs well but is only complicated in terms

*Corresponding author
Received: 31 October 2023 / Accepted: 7 March 2024
DOI: 10.22067/smps.2024.45179

 https://smps.um.ac.ir

of computing algorithm whereas marginal likelihood approach is of less difficult. They therefore suggested from their study that the marginal likelihood approach is appropriate for researchers with model choice setting while the joint space method can be used for comparing models of varying dimensions. It was learned from the literature that the origin of the Bayesian hypothesis testing was credited to Jeffreys (1961), and ordinarily it was referred to as a significant test.

Kass and Raftery (1995) reviewed the work of Jeffreys (1961) and added that the Bayesian technique to hypothesis testing meets the test of time with moderate computation technique. They found out that the Bayes factor is sensitive to the assumption of the parametric model and choice of prior. Giampaoli et al. (2015) examined the Bayes factor in a restricted simple linear regression and pointed out that their approach considered restricted parameter space to be more informative than unrestricted parameter space, which permits evidence for null hypothesis. Bayes factor provided a means through which two competing hypothesis may be compared; see Johnson et al. (2023). The Bayes factor $B10$ is used to compute the probability of the observed data under the alternative hypothesis H_1 versus the null hypothesis H_0, which is expressed as

$$B_{10} = \frac{H_1}{H_0} = \frac{\int P(Y|\beta)\,\pi(\beta|H_1)\,d\beta}{\int P(Y|\beta)\,\pi(\beta|H_0)\,d\beta}. \tag{1}$$

Jarosz and Wileyr (2014) in their presentation, showcased the similarities between the Bayes factor and p-value. They believe that bf should have similar information as p-value. The only difference is that bf allowed the researcher to give evidence in both the alternative and null hypothesis unlike p-value that will solely give evidence about null hypothesis. Roozbeh and Hamzah (2020) pointed out that the restriction of parameter in partial linear regression can overcome the problem of multicollinearity. They believed that restricted estimator outperformed traditional estimator. They added that restriction can be hypothesis that may have to be tested.

Over the years, many studies have been carried out in Bayesian inference for linear and nonlinear models on the premise of model averaging, parameter estimation, and hypothesis testing, but most importantly restricted stem-rule are not well studied in the Bayesian framework. Against this backdrop, this paper seems to examine nonspherical disturbance (autocorrelated error) in the Bayesian restricted stem-rule paradigm.

2 Materials and methods

Let $y = X\beta + u$ be the linear regression model, where y is an $n \times 1$ set of observations on the regressand, X is a set of $n \times p$ full column rank of regressors, and β is $p \times 1$ vectors of unknown parameters while u is an $n \times 1$ vectors of disturbance error not necessarily well behave, which is characterized by autocorrelated errors. Let there be m linearly independent restriction that constrains the regression coefficients such that

$$r = R\beta, \tag{2}$$

where r is an $m \times 1$ vector and R is an $m \times p$ matrix of rank $m < p$ Chaturvedi et al. (2001). The $AR(1)$ and $MA(1)$ processes are expressed as

$$\widehat{\Omega}_{AR} = \frac{\sigma_u^2}{1-\rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

The $MA(1)$ with autocorrelated error is expressed as

$$\widehat{\Omega}_{MA} = \sigma_u^2 \begin{bmatrix} (1+\phi^2) & \phi & \cdot & \cdot & \cdots \\ \phi & (1+\phi^2) & \phi & \cdot & \cdots \\ \cdot & \phi & (1+\phi^2) & \phi & \cdots \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \cdot & \cdot & \cdot & \cdots & (1+\phi^2) \end{bmatrix},$$

where ϕ ranges between -1 and 1. The generalized restricted least squares estimates are obtained as follows: Adopting the criterion of minimizing the sum of squares $(y - X\beta)'\widehat{\Omega}(y - X\beta)$ subject to the condition that $R\beta = r$. This leads to the Lagrangian function

$$\widehat{\beta}_R = \widehat{\beta} + \left(X\widehat{\Omega}X\right)^{-1} \left[R\left(X\widehat{\Omega}X\right)^{-1} R\right]^{-1} R'\left(r - R\widehat{\beta}\right), \tag{3}$$

$$\widehat{\beta}_R = \widehat{\beta} + \left(X'\widehat{\Omega}X\right)^{-1} R' \left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} \left(r - R\widehat{\beta}\right). \tag{4}$$

Thus β_R is a constrained estimates, following Chaturvedi et al. (2001), the restricted stein-rule version of intertwined disturbance errors can be expressed as

$$\widehat{\beta}_S = \left[1 - \frac{a}{n} \frac{\left(y - X\widehat{\beta}\right)' \widehat{\Omega} \left(y - X\widehat{\beta}\right)}{\widehat{\beta} X \widehat{\Omega} X \widehat{\beta}}\right] \widehat{\beta}. \tag{5}$$

We have

$$\widehat{\beta}_{RS} = \widehat{\beta}_S + \left(X'\widehat{\Omega}X\right)^{-1} R' \left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} \left(r - R\widehat{\beta}_S\right), \tag{6}$$

where $\widehat{\Omega} = \Omega(b)$, in which b is a consistent and efficient estimator of β; thus $b = \left(X'\widehat{\Omega}X\right)^{-1}\left(X'\widehat{\Omega}y\right)$. Following Chaturvedi and Shukla (1990), we have modified the Stein rule estimator of beta.

3 Bayesian restricted least squares estimator

The restricted posterior density of restricted β and σ is obtained by marginalizing the conjugate of normal-inverse gamma and restricted likelihood since both are of the same family of distribution Oloyede (2023). The linear model is expressed as

$$y = X\beta_R + u. \tag{7}$$

The likelihood function of β, X, and y, where $\theta = (\beta, \lambda)$ given sample vectors X_1, $X_2 = (1, 2, \ldots, n)'$ and $y = (y_1, y_2, \ldots, y_n)'$ and incorporating restricted β_R, is expressed as

$$L\left(\beta_R, \sigma^2 | X, y\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(y - X\beta_R\right)' \widehat{\Omega}^{-1}\left(y - X\beta_R\right)\right] \to H_1, \tag{8}$$

$$L\left(\beta_R, \sigma^2 | X, y\right) = \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\left(y - X\beta_R\right)'\left(y - X\beta_R\right)\right] \to H_0. \tag{9}$$

Note that normal-inverse gamma priors are conjugate priors and selected because the prior and posterior densities are of the same family of distributions; see Oloyede (2023). Moreover,

$$p\left(\beta_R | \sigma^2\right) p\left(\sigma^2\right) = (2\pi)^{-\frac{k}{2}} \left|\widehat{\Omega}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\beta_R - \mathbb{B}\right)' \widehat{\Omega}^{-1}\left(\beta - \mathbb{B}\right)\right] \times \sigma^{-(a-k)} \exp\left[-\frac{b}{\sigma^2}\right], \tag{10}$$

$$p\left(\beta_R | \sigma^2\right) p\left(\sigma^2\right) = \sigma^{-(a-k)} \exp\left[-\frac{1}{2\sigma^2}\left(\beta_R - \mathbb{B}\right)' \widehat{\Omega}^{-1}\left(\beta_R - \mathbb{B}\right) + 2b\right], \tag{11}$$

where $X$is an $n \times k$ matrix,
β_R is unknown parameter,
B is a prior mean vector of β (true value),
σ^2 is a prior variance for β,
$\widehat{\sigma}^2 = \frac{(y - X\beta_R)\widehat{\Omega}(y - X\beta_R)}{n-k}$,
$a - k$ is first hyper-parameter,
b is second hyper-parameter.

Since σ^2 is known, normal-inverse gamma conjugate prior is adopted. For details of the posterior, the reader is advised to see Oloyede (2023).

Posterior density function is

$$\widehat{\beta}_R \sim MVN\left(\widehat{\beta}_R \ , \ \widehat{\sigma}^2\left(X'\widehat{\Omega}X\right)^{-1}\left[1 - \left(X'\widehat{\Omega}X\right)^{-1} R'\left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} R\right]\right), \tag{12}$$

$$\breve{\beta}_{RS} \sim MVN\left(\widehat{\beta}_{RS} \ , \ \widehat{\sigma}^2\left(X'\widehat{\Omega}X\right)^{-1}\left[1 - \left(X'\widehat{\Omega}X\right)^{-1} R'\left[R\left(X'\widehat{\Omega}X\right)^{-1} R'\right]^{-1} R\right]\right), \tag{13}$$

for restricted the Stein-rule estimator.

4 Bayesian hypothesis testing

Let $M_i = X\beta + u*$ be the regression model with autocorrelated error, setting up dual distinct model M_1 and M_0, where 1 and 0 represent the model with autocorrelated error and without autocorrelated error, respectively. The prior belief of parameter estimates of both models is the same. Let D be the set of observed datasets corrupted with autocorrelated error of $AR(1)$ and $MA(1)$ independently. Thus, the posterior of M_1 and M_0 can be expressed as follows:

$$P(\beta|\mathscr{D}, M_0) \propto P(\beta|\, M_0) * P(D|\beta, M_0) \rightarrow H_0, \tag{14}$$

$$P(\beta|\mathscr{D}, M_1) \propto P(\beta|\, M_1) * P(D|\beta, M_1) \rightarrow H_1. \tag{15}$$

Bayes factor is therefore expressed as

$$BF_{10} = \frac{Posterior\ model\ odds}{prior\ model\ odds}, \tag{16}$$

$$\frac{P(D|\beta, M_1)}{P(D|\beta, M_0)} = \frac{P(M_1|D)}{P(M_0|D)} \div \frac{P(M_1)}{P(M_0)}, \tag{17}$$

$$\frac{P(M_1|D)}{P(M_0|D)} = \frac{P(M_1)}{P(M_0)} \times \frac{P(D|\beta, M_1)}{P(D|\beta, M_0)}. \tag{18}$$

5 Data generation processes and simulation experiment

The Markov Chain Monte Carlo simulation algorithm was adopted to examine the small sample properties of the family of restricted least squares estimators with autocorrelated error in Bayesian frameworks. Data were generated based on the following parameters of the model: $P = 6, n = 25$, $y_t = X_t\beta + u_t$, $t = 1, \ldots, 25$, where ε_t assumed to be generated by the $AR(1)$ process $u_t = \rho u_{t-1} + \varepsilon_t$ or the $MA(1)$ process $u_t = \varepsilon_t - \rho\varepsilon_{t-1}$, $_t$ $N(0,1)$. Moreover, $\hat{\beta}$ was set as $(1.2, 2, 0.8, 0.3, 2.1, 1.1)$ while seed was set at 1234. Also, 10000 iterations were set for Bayesian Monte Carlo simulation, and ρ was set at $-0.8, -0.5, -0.3, 0, 0.3, 0.5, 0.8$ for both $AR(1)$ process and $MA(1)$ process. The restriction of parameters was set as

$$R = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{19}$$

$$r = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \tag{20}$$

where $\beta_1 - \beta_3 = 0$, $\beta_2 + \beta_4 = 1$ and $\beta_5 = 0$. All computations were carried out using Statistical software R-Core (2022). The datasets class contained the posterior sample for the model parameters.

6 Hypothesis metrics

Bayesian Information Criterion (BIC) and Quadratic weight loss and risk function approaches that incorporated autocorrelated error structures were used to evaluate the Bayes factor of both the Bayes estimate and posterior mean. Let $\left(\widehat{\beta}_R - \beta\right) = \left(\widehat{\beta}_R - \beta\right) Q \left(\widehat{\beta}_R - \beta\right)$ be a quadratic or square error loss function, where $\hat{\beta}_R$ is an estimator of β and Q is the $\sum_{i=1}^{\beta_R} \widehat{\beta}_R$ weight of loss function. For obtaining the Bayes factor for $\hat{\beta}_R$ and $hat\beta_{RS}$, ratio of quadratic loss function of both H_1 and H_0 was computed as well as ratio of BIC for both H_1 and H_0.

7 Data analysis and discussion

Both null and alternative hypotheses were subjected to the same datasets having autocorrelation error. Two types of autocorrelation error $AR(1)$ and $MA(1)$ processes were examined. The null hypothesis was modeled in a natural setting without being generalized, but the alternative hypothesis was modeled in a generalized pattern by incorporating autocorrelation error structure into the model. Bayesian restricted least squares and Bayesian restricted stem-rule were the predominant estimators adopted in the study. The study looked into the evidence that alternative hypothesis have over null hypothesis due to the presence of autocorrelated error in the datasets in Bayesian paradigm. The threshold of 1 is the decision value.

Table 1: Bayes factor with loss and risk function approach with $AR(1)$ process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.001761	1.001954	1.000197	0.999665
-0.6	0.947783	0.947851	0.94699	0.946795
-0.3	0.964658	0.964679	0.964214	0.964158
0	1.000008	1.000014	0.999745	0.999731
0.3	1.039884	1.039891	1.03971	1.039685
0.6	1.073562	1.07359	1.073367	1.073277
0.8	1.094908	1.094954	1.094644	1.094493

Table 2: Bayes factor with BIC approach with $AR(1)$ process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.172955243	1.188872	1.010712	1.010932
-0.6	2.273047888	2.286148	1.021469	1.021607
-0.3	2.005146615	2.008686	1.013486	1.01364
0	0.998911484	0.999406	1.000062	1.000056
0.3	0.380557759	0.380816	1.035379	1.035065
0.6	0.18084545	0.181293	1.253861	1.253553
0.8	0.33050558	0.332143	2.335213	2.336675

Tables 1 and 2 provide information about the outcome of study about the Bayes factor with autocorrelated error of autoregressive process of order 1. Table 1 shows that the autocorrelated error of magnitude -0.8, 0 to 0.8 provided evidence in favour of the alternative hypothesis H_1 for both the Bayes estimates and posterior mean considering Bayesian restricted least squares whereas the magnitude of -0.6 and -0.3 provided evidence in support of accepting null hypothesis H_0. In the same vein, the Bayes factor for Bayesian restricted Stein-rule provided information in favor of alternative hypothesis considering the magnitude of 0.3 to 0.8 autocorrelated errors. In Table 2 where BIC approach was adopted to compute the Bayes factor, the magnitudes of -0.8 through -0.3 portend evidences in favour of alternative hypothesis whereas other magnitudes were otherwise. This is in Bayesian restricted least squares for both Bayes estimates and posterior mean. Interestingly, Bayesian restricted stein rule has Bayes factor that portends evidence in favour of alternative hypothesis for both Bayes estimates and posterior mean for all the magnitudes.

Table 3: Bayes factor with Loss and risk function with MA(1) Process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	1.052104	1.052086	1.047152	1.047216
-0.6	1.002316	1.002319	1.000276	1.000283
-0.3	0.98147	0.981478	0.980911	0.980893
0	1.000008	1.000014	0.999745	0.999731
0.3	1.058752	1.05874	1.058577	1.058607
0.6	1.069787	1.069688	1.069916	1.07014
0.8	0.799912	0.799774	0.799965	0.800259

Table 4: Bayes factor with Bayesian Information Criteria with MA(1) Process

$n = 25$	Bayesian Restricted Least Squares		Bayesian Restricted Stein-Rule	
ρ	Bayes Estimates	Posterior Mean	Bayes Estimates	Posterior Mean
-0.8	0.428503544	0.428626	1.294917	1.294175
-0.6	1.728149432	1.729083	1.137145	1.137505
-0.3	1.770699531	1.771931	1.022389	1.022589
0	0.998911484	0.999406	1.000062	1.000056
0.3	0.373687701	0.373339	1.023952	1.023549
0.6	0.792633093	0.786735	1.050708	1.051685
0.8	41.83140879	41.63028	1.511668	1.523933

Tables 3 and 4 show Bayes factor Loss and risk function with $MA(1)$ Process . Tables 3 and 4 provide information about the outcome of the study about the Bayes factor with autocorrelated error of moving average process of order 1. Table 3 shows that autocorrelated error of magnitude -0.8,-0.6, 0, 0.3, and 0.6 provided evidence in favour of alternative hypothesis H_1 for both Bayes estimates and posterior mean considering Bayesian restricted least squares whereas other magnitudes provided evidence in support of accepting null hypothesis H_0. In the same vein, the

Bayes factor for Bayesian restricted the Stein-rule provided information in favour of alternative hypothesis considering the magnitude of -0.8, -0.6, 0.3, and 0.6 autocorrelated errors. In Table 4 where the BIC approach was adopted to compute the Bayes factor, the magnitudes of -0.6, -0.3, and 0.8 portend evidences in favour of alternative hypothesis whereas other magnitudes were otherwise. This is in Bayesian restricted least squares for both Bayes estimates and posterior mean. Interestingly, the Bayesian restricted Stein rule has Bayes factors that portend evidence in favour of alternative hypothesis for both Bayes estimates and posterior mean for all the magnitudes.

8 Conclusion

This study introduced Bayes factor as an alternative to probability value, which has an age long drawback in science, technology, education, and so on, with the view of having probability, not only in the value of data, but also in the uncertainty of unknown parameters. The study established and compared the BIC approach and loss/risk function approach in determining the Bayes factor. The small sample behaviour of regression parameter with uncertainty was studied via the violation of non-serial correlation of error terms in a restricted least squares and restricted stein-rule least squares. The study allays fears in the minds of researchers in engineering, science, technology, education, and so on, who have over the years unsatisfied with p-value or using it wrongly. The Bayesian test hypothesis now provides alternative means through which they can verify the presence/absence of autocorrelated errors in their data.

References

Chaturvedi, A. and Shukla, G. (1990). Stein rule estimation in linear model with non-scalar error covariance matrix. *Sankhyä: The Indian Journal of Statistics*, **Series B**, 293–304.

Chaturvedi, A., Wan, A. T. K. and Singh, S. P. (2001). Stein-rule restricted estimator in a linear regression model with nonspherical disturbances. *Communications in Statistics - Theory and Methods*, **30**, 55–68.

Giampaoli, V., Pereira, C., Bolfarine, H., Singer, J. (2015). *Bayes Factors for Comparison of Restricted Simple Linear Regression Coefficients.* In: Polpo, A., Louzada, F., Rifo, L., Stern, J., Lauretto, M. (eds) Interdisciplinary Bayesian Statistics. Springer Proceedings in Mathematics & Statistics, 118. Springer, Cham.

Han, C. and Carlin, B. P. (2001). Markov chain Monte Carlo methods for computing Bayes factors: A comparative review. *Journal of the American Statistical Association,* **96**, 1122–1132.

Jarosz, A. F. and Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving,* **7**, 2.

Jeffreys, H. (1961). *Theory of Probability.* (3rd ed) Oxford University Press, Oxford.

Johnson, V. E., Pramanik, S. and Shudde, R., (2023). Bayes factor functions for reporting outcomes of hypothesis tests. *Statistics Psychological And Cognitive Sciences*, **120**, e2217331120.

Kass, R. E. and Raftery, A. E., (1995). Bayes factors. *Journal of the American Statistical Association,* **90**, 773–795.

Oloyede, I. (2023). Bayesian linear restricted least squares estimators. *Journal of Indian Society Probability and Statistics*, **24** 65–74.

R-Core Team (2022). R: A language and environment for statistical computing, *R Foundation for Statistical Computing*, Vienna, Austria.

Roozbeh, M., and Hamzah, N. A. (2020). Uncertain stochastic ridge estimation in partially linear regression models with elliptically distributed errors. *Statistics*, **54**, 494–523.

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 47–73. Research Paper

SMPS

Ordering results of extreme order statistics from multiple-outlier scale models with dependence

Sangita Das[†], **Suchandan Kayal**[‡*]

[†] *Theoretical Statistics and Mathematics Unit, Indian Statistical Institute Bangalore, Bangalore-560059, India*

[‡] *Department of Mathematics, National Institute of Technology Rourkela, Rourkela-769008, India*

Email(s): sangitadas118@gmail.com, suchandan.kayal@gmail.com

Abstract. Many authors have studied ordering results between extreme order statistics from multiple-outlier models when the observations are independent. However, the independence assumption is not very attractive in many situations due to the complexity of the problems. This paper focuses on stochastic comparisons of extreme order statistics stemming from multiple-outlier scale models with dependence. Archimedean copula is used to model dependence structure among nonnegative random variables. Sufficient conditions are obtained to compare the largest order statistics in the sense of the usual stochastic, reversed hazard rate, likelihood ratio, dispersive, star, and Lorenz orders. The smallest order statistics are also compared with respect to the usual stochastic, hazard rate, star, and Lorenz orders. Here, the sufficient conditions are based on the weak-super majorization, weak-sub majorization, and p-larger orders between the model parameters. To illustrate the theoretical establishments, some examples are provided. Furthermore, some counterexamples are provided to establish that ignorance of sufficient conditions may not lead to the established ordering results between the order statistics.

Keywords: Archimedean copula; Dispersive order; Hazard rate order; Majorization; Multiple-outlier model; Reversed hazard rate order.

1 Introduction

Order statistics play a vital role in many fields such as statistical inference, economics, reliability theory and operations research. Consider a random sample $X_1,\ldots,X_n$ from a population. Then, the ith order statistic is denoted by $X_{i:n}$, where $i=1,\ldots,n$. In reliability theory, the ith order

*Corresponding author

Received: 9 October 2023 / Accepted: 8 March 2024
DOI: 10.22067/smps.2024.45213

 https://smps.um.ac.ir

statistic represents the lifetime of an $(n-i+1)$-out-of-n system, which functions if at least $n-i+1$ of n components work. In particular, the order statistics $X_{1:n}$ and $X_{n:n}$ represent the lifetimes of series and parallel systems, respectively. Due to the correspondence between the order statistics and the systems' reliability, a lot of effort has been put to study ordering results between order statistics in terms of many well-known stochastic orders. In this paper, we deal with the comparison of extreme order statistics arising from dependent multiple-outlier scale models in the sense of the usual stochastic, reversed hazard rate, hazard rate, star, and Lorenz orders.

Due to the robustness of different estimators of model parameters, multiple-outlier models have been widely used by many researchers. Now, we present some developments on stochastic comparisons between order statistics arising from multiple-outlier models.

Kochar and Xu (2011) considered multiple-outlier exponential models. They showed that more heterogeneity among the scale parameters of the model results more skewed order statistics. Zhao and Balakrishnan (2012) took similar model and obtained ordering results between the largest order statistics with respect to the likelihood ratio, reversed hazard rate, hazard rate, and usual stochastic orderings. Zhao and Balakrishnan (2015) discussed stochastic comparisons of the largest order statistics from multiple-outlier gamma models in terms of various stochastic orderings such as the likelihood ratio, hazard rate, star, and dispersive orders. Kochar and Torrado (2015) established likelihood ratio ordering between the largest order statistics arising from independent multiple-outlier scale models. Sufficient conditions for the comparison of the lifetimes of series systems with respect to dispersive order have been obtained by Fang et al. (2016). They considered that the components of the series systems follow multiple-outlier Weibull models. Amini Seresht et al. (2016) studied multiple-outlier proportional hazard rate models and developed ordering results with respect to the star, Lorenz, and dispersive orders. Furthermore, they proved that more heterogeneity among the multiple-outlier components led to a more skewed lifetime of a k-out-of-n system consisting of these components. Balakrishnan and Torrado (2016) obtained conditions under which the likelihood ratio order holds between largest order statistics under the set-up of multiple-outlier exponential model. Torrado (2017) developed the comparison result similar to Balakrishnan and Torrado (2016) for the multiple-outlier scale models when the random variables are independent. Wang and Cheng (2017) studied an open problem on mean residual life ordering between two parallel systems under multiple-outlier exponential models, which was proposed by Balakrishnan and Zhao (2013).

It is noted that almost all concerned research in this area has been developed under the assumption of statistically independent component lifetimes. However, there are some practical situations, where the condition of statistically mutual independence among the component lifetimes is evidently unsuitable. For an example, let us consider a mechanical system. The components of the system are suffering a common stress. Then, it is of huge interest to include statistical dependence among component lifetimes into the study of stochastic comparison of the lifetimes of the series and parallel systems. Furthermore, note that due to the complexity of working with the dependent random variables, marginal effort was put to the study of dependent multiple-outlier models by the researchers (see Navarro et al. 2018). These are the main motivations to investigate ordering properties of the extreme order statistics arising from multiple-outlier dependent scale components. The dependency structure among the random variables is modeled by the concept of Archimedean copulas. We recall that a nonnegative random variable X with distribution function F_X is said to follow the scale model if there exists $\lambda > 0$ such that $F_X(x) = F(\lambda x)$, where F is the baseline distribution function and λ is the scale parameter.

In this paper, we will develop different ordering results between the largest as well as the smallest order statistics stemming from multiple-outlier dependent scale models with respect to several stochastic orderings such as the usual stochastic, hazard rate, reversed hazard rate, star, and Lorenz orders. Let $\{X_1,\ldots,X_{n_1^*},X_{n_1^*+1},\ldots,X_{n^*}\}$ be a set of dependent and heterogeneous random observations. The observations are sharing a common Archimedean copula with generator ψ_1 and are taken from the multiple-outlier scale model, where for $i = 1,\ldots,n_1^*$, $X_i \sim F_1(\lambda_1 x)$ and for $j = n_1^*+1,\ldots,n^*$, $X_j \sim F_2(\lambda_2 x)$, where $\lambda_1,\ \lambda_2 > 0$. Note that $F_1(\cdot)$ and $F_2(\cdot)$ are two different baseline distribution functions. Also, let $\{Y_1,\ldots,Y_{n_1^*},Y_{n_1^*+1},\ldots,Y_{n^*}\}$ be another set of dependent and heterogeneous random observations sharing a common Archimedean copula with generator ψ_2, drawn from the multiple-outlier scale model, where for $i = 1,\ldots,n_1^*$, $Y_i \sim F_1(\mu_1 x)$ and for $j = n_1^*+1,\ldots,n^*$, $Y_j \sim F_2(\mu_2 x)$, where $\mu_1,\ \mu_2 > 0$.

Denote by $r_1,\ \tilde{r}_1$ and $r_2,\ \tilde{r}_2$ the hazard rate and reversed hazard rate functions for F_1 and F_2, respectively. Furthermore, $X_{n:n}(n_1,n_2)$, $Y_{n^*:n^*}(n_1^*,n_2^*)$ and $X_{1:n}(n_1,n_2)$, $Y_{1:n^*}(n_1^*,n_2^*)$ denote the largest and the smallest order statistics, respectively, arising from $\{X_1,\ldots,X_{n_1},X_{n_1+1},\ldots,X_n\}$ and $\{Y_1,\ldots,Y_{n_1^*},Y_{n_1^*+1},\ldots,Y_{n^*}\}$, where $1 \le n_1 \le n_1^* \le n_2^* \le n_2$, $n = n_1+n_2$ and $n^* = n_1^*+n_2^*$. We aim to establish sufficient conditions, under which the following implications hold:

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \quad \Rightarrow \quad Y_{n^*:n^*}(n_1^*,n_2^*) \le_{st} [\le_{rh}] X_{n:n}(n_1,n_2),$$

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \quad \Rightarrow \quad X_{1:n}(n_1,n_2) \le_{st} Y_{1:n^*}(n_1^*,n_2^*)$$

and

$$(\underbrace{u_1,\ldots,u_1}_{n_1^*},\underbrace{u_2,\ldots,u_2}_{n_2^*}) \succeq_w (\underbrace{v_1,\ldots,v_1}_{n_1^*},\underbrace{v_2,\ldots,v_2}_{n_2^*}) \Rightarrow X_{1:n}(n_1,n_2) \le_{hr} Y_{1:n^*}(n_1^*,n_2^*),$$

where $u_i = \log\lambda_i$ and $v_i = \log\mu_i$, $i = 1,\ 2$.

The remainder of the paper is rolled out as follows. Some basic definitions and important lemmas are provided in Section 2. Section 3 consists of two subsections. In Subsection 3.1, we obtain sufficient conditions, under which two largest order statistics are comparable according to the usual stochastic order, reversed hazard rate order, likelihood ratio, dispersive order, star order, and Lorenz order, whereas in Subsection 3.2, we study the usual stochastic order, hazard rate order, star order, and Lorenz order between two smallest order statistics. We also present some examples to illustrate the established results. Finally, we conclude the paper in Section 4.

Throughout the paper, we only concern about nonnegative random variables. Increasing and decreasing mean nondecreasing and nonincreasing, respectively. Also, the prime "$\prime$" stands for the first order derivative.

2 Basic notions

In this section, we recall some basic definitions and well-known concepts of stochastic orders and majorization. Let $\mathbf{x} = (x_1,\ldots,x_n)$ and $\mathbf{y} = (y_1,\ldots,y_n)$ be two n-dimensional vectors such that

$\mathbf{x}$,$\mathbf{y} \in \mathbb{A}$, where $\mathbb{A} \subset \mathbb{R}^n$ and $\mathbb{R}^n$ be an n-dimensional Euclidean space. Also, consider the order coordinates of the vectors $\mathbf{x}$ and $\mathbf{y}$ as $x_{1:n} \le \cdots \le x_{n:n}$ and $y_{1:n} \le \cdots \le y_{n:n}$, respectively.

Definition 1. *A vector* $\mathbf{x}$ *is said to be*

- *majorized by another vector* $\mathbf{y}$, *(denoted by* $\mathbf{x} \preceq^m \mathbf{y}$*), if for each* $l = 1,\ldots,n-1$, *we have* $\sum_{i=1}^{l} x_{i:n} \ge \sum_{i=1}^{l} y_{i:n}$ *and* $\sum_{i=1}^{n} x_{i:n} = \sum_{i=1}^{n} y_{i:n}$;
- *weakly submajorized by another vector* $\mathbf{y}$, *denoted by* $\mathbf{x} \preceq_w \mathbf{y}$, *if for each* $l = 1,\ldots,n$, *we have* $\sum_{i=l}^{n} x_{i:n} \le \sum_{i=l}^{n} y_{i:n}$;
- *weakly supermajorized by another vector* $\mathbf{y}$, *denoted by* $\mathbf{x} \preceq^w \mathbf{y}$, *if for each* $l = 1,\ldots,n$, *we have* $\sum_{i=1}^{l} x_{i:n} \ge \sum_{i=1}^{l} y_{i:n}$.

Note that $\mathbf{x} \preceq^m \mathbf{y}$ implies both $\mathbf{x} \preceq_w \mathbf{y}$ and $\mathbf{x} \preceq^w \mathbf{y}$. For a brief introduction of majorization orders and their applications, we refer to Marshall et al. (2011). Now, we present notions of stochastic orderings. Let X_1 and X_2 be two nonnegative random variables with probability density functions f_{X_1} and f_{X_2}, cumulative density functions F_{X_1} and F_{X_2}, survival functions $\bar{F}_{X_1} = 1 - F_{X_1}$ and $\bar{F}_{X_2} = 1 - F_{X_2}$, hazard rate functions $r_{X_1} = f_{X_1}/\bar{F}_{X_1}$ and $r_{X_2} = f_{X_2}/\bar{F}_{X_2}$, and reversed hazard rate functions $\tilde{r}_{X_1} = f_{X_1}/F_{X_1}$ and $\tilde{r}_{X_2} = f_{X_1}/F_{X_2}$, respectively.

Definition 2. *A random variable* X_1 *is said to be smaller than* X_2 *in the*

- *likelihood ratio order (denoted by* $X_1 \le_{lr} X_2$*) if* $f_{X_2}(x)/f_{X_1}(x)$ *is increasing in* $x > 0$,
- *hazard rate order (denoted by* $X_1 \le_{hr} X_2$*) if* $r_{X_1}(x) \ge r_{X_2}(x)$, *for all* $x > 0$;
- *reversed hazard rate order (denoted by* $X_1 \le_{rh} X_2$*) if* $\tilde{r}_{X_1}(x) \le \tilde{r}_{X_2}(x)$, *for all* $x > 0$;
- *usual stochastic order (denoted by* $X_1 \le_{st} X_2$*) if* $\bar{F}_{X_1}(x) \le \bar{F}_{X_2}(x)$, *for all* x;
- *star order (denoted by* $X_1 \le_* X_2$ *or* $F_{X_1}(x) \le_* F_{X_2}(x)$*) if* $F_{X_2}^{-1} F_{X_1}(x)$ *is star shaped in the sense that* $\frac{F_{X_2}^{-1} F_{X_1}(x)}{x}$ *is increasing in* x *on the support of* X_1;
- *Lorenz order (denoted by* $X_1 \le_{Lorenz} X_2$*) if* $\frac{1}{E(X_1)} \int_0^{F_{X_1}^{-1}(u)} x dF_{X_1}(x) \ge \frac{1}{E(X_2)} \int_0^{F_{X_2}^{-1}(u)} x dF_{X_2}(x)$, *for all* $u \in (0,1]$;
- *dispersive order, (denoted by* $X_1 \le_{disp} X_2$*) if* $F_{X_1}^{-1}(\beta) - F_{X_1}^{-1}(\alpha) \le F_{X_2}^{-1}(\beta) - F_{X_2}^{-1}(\alpha)$ *for all* $0 < \alpha \le \beta < 1$.

Note that both the hazard rate and reversed hazard rate orderings imply the usual stochastic ordering. Also, star order implies Lorenz order (see Marshall and Olkin 2007). One may refer to Shaked and Shanthikumar (2007) for a detailed discussion on various stochastic orderings. The next definition is for the Schur-convex and Schur-concave functions.

Definition 3. *A function* $\Psi : \mathbb{R}^n \to \mathbb{R}$ *is said to be Schur-convex (Schur-concave) in* $\mathbb{R}^n$ *if*

$$\mathbf{x} \overset{m}{\succeq} \mathbf{y} \Rightarrow \Psi(\mathbf{x}) \ge (\le) \Psi(\mathbf{y}), \quad \textit{for all} \ \ \mathbf{x}, \ \mathbf{y} \in \mathbb{R}^n.$$

Throughout the article, we will use the notations. (i) $\mathscr{D}_+ = \{(x_1,\ldots,x_n) : x_1 \geq x_2 \geq \cdots \geq x_n > 0\}$ and (ii) $\mathscr{E}_+ = \{(x_1,\ldots,x_n) : 0 < x_1 \leq x_2 \leq \cdots \leq x_n\}$. Set $h'(z) = \frac{dh(z)}{dz}$. The following consecutive lemmas due to Kundu et al. (2016) are useful to prove the results in the subsequent sections. The partial derivative of h with respect to its kth argument is denoted by $h_{(k)}(\mathbf{z}) = \partial h(\mathbf{z})/\partial z_k$, for $k = 1,\ldots,n$, where $\mathbf{z} = (z_1,\ldots,z_n)$.

Lemma 1. *Let $h : \mathscr{D}_+ \to \mathbb{R}$ be a function, continuously differentiable on the interior of $\mathscr{D}_+$. Then, for $\mathbf{x},\ \mathbf{y} \in \mathscr{D}_+$,*

$$\mathbf{x} \succeq^m \mathbf{y} \text{ implies } h(\mathbf{x}) \geq (\leq)\ h(\mathbf{y}),$$

if and only if $h_{(k)}(\mathbf{z})$ is decreasing (increasing) in $k = 1,\ldots,n$.

Lemma 2. *Let $h : \mathscr{E}_+ \to \mathbb{R}$ be a function, continuously differentiable on the interior of $\mathscr{E}_+$. Then, for $\mathbf{x},\ \mathbf{y} \in \mathscr{E}_+$,*

$$\mathbf{x} \succeq^m \mathbf{y} \text{ implies } h(\mathbf{x}) \geq (\leq)\ h(\mathbf{y}),$$

if and only if $h_{(k)}(\mathbf{z})$ is increasing (decreasing) in $k = 1,\ldots,n$.

The following lemma due to Saunders and Moran (1978) is useful to establish star order between the order statistics.

Lemma 3. *Let $\{F_\lambda | \lambda \in \mathbb{R}\}$ be a class of distribution functions, such that F_λ is supported on some interval $(a,b) \subseteq (0,\infty)$ and has density f_λ that does not vanish on any subinterval of (a,b). Then,*

$$F_\lambda \leq_* F_{\lambda^*}, \quad \lambda \leq \lambda^*$$

if and only if

$$\frac{F'_\lambda(x)}{x f_\lambda(x)} \text{ is decreasing in } x,$$

where F'_λ is the derivative of F_λ with respect to λ.

To model the dependency structure among the random variables, the concept of copulas plays a vital role. One of the important characteristics of the copula is that it involves the information of the dependencies between the random variables apart from the behavior of the marginal distributions. Archimedean copulas are important class of copulas. These are used widely because of its simplicity. Let F and $\bar{F}$ be the joint distribution function and the joint survival function of the random vector $\mathbf{X} = (X_1,\ldots,X_n)$. Suppose there exist functions $C(\mathbf{z}) : [0,1]^n \to [0,1]$ and $\hat{C}(\mathbf{z}) : [0,1]^n \to [0,1]$ such that for all $x_i,\ i \in \mathscr{I}_n$, where $\mathscr{I}_n$ is the index set

$$F(x_1,\ldots,x_n) = C(F_1(x_1),\ldots,F_n(x_n))$$

and

$$\bar{F}(x_1,\ldots,x_n) = \hat{C}(\bar{F}_1(x_1),\ldots,\bar{F}_n(x_n))$$

hold. Then, $C(\mathbf{z})$ and $\hat{C}(\mathbf{z})$ are said to be the copula and survival copula of $\mathbf{X}$, respectively. Here, $F_1,\ldots,F_n$ and $\bar{F}_1,\ldots,\bar{F}_n$ are the univariate marginal distribution functions and survival functions of the random variables $X_1,\ldots,X_n$, respectively.

Now, let $\psi : [0,\infty) \to [0,1]$ be a nonincreasing and continuous function, satisfying $\psi(0) = 1$ and $\psi(\infty) = 0$. Also, let $\psi = \phi^{-1} = \sup\{x \in \mathscr{R} : \phi(x) > v\}$ be the right continuous inverse. Furthermore, suppose ψ satisfies the conditions (i) $(-1)^i \psi^i(x) \geq 0,\ i = 0,\ 1,\dots,d-2$ and (ii) $(-1)^{d-2}\psi^{d-2}$ is nonincreasing and convex. That implies the generator ψ is d-monotone. Then, a copula C_ψ is said to be an Archimedean copula if it can be written as the following form

$$C_\psi(v_1,\dots,v_n) = \psi(\phi(v_1)+\cdots+\phi(v_n)), \quad \text{for all } v_i \in [0,1],\ i \in \mathscr{I}_n.$$

For further discussion on Archimedean copulas, one may refer to McNeil and Nešlehová (2009); Nelsen (2006). A few well known Archimedan copulas including different properties of the generators are enlisted in Table 1. Next lemma is taken from Li and Fang (2015), which has been used to prove the results in Theorems 1, 7 and 9.

Lemma 4. *For two n-dimensional Archimedean copulas C_{ψ_1} and C_{ψ_2}, if $\phi_2 \circ \psi_1$ is super-additive, then $C_{\psi_1}(\mathbf{z}) \leq C_{\psi_2}(\mathbf{z})$, for all $\mathbf{z} \in [0,1]^n$. A function f is said to be super-additive, if $f(x)+f(y) \leq f(x+y)$, for all x and y in the domain of f.*

3 Main results

This section is completely devoted to establishing sufficient conditions, under which the extreme order statistics arising from multiple outlier dependent scale models are comparable in different stochastic senses. The usual stochastic, hazard rate, reversed hazard rate, star, and Lorenz orders are used in what follows. Throughout this section, we denote two-dimensional vectors by bold symbols. For example, $\boldsymbol{\lambda} = (\boldsymbol{\lambda_1},\boldsymbol{\lambda_2})$ and $\boldsymbol{\mu} = (\boldsymbol{\mu_1},\boldsymbol{\mu_2})$.

3.1 Orderings between the largest order statistics

This subsection addresses ordering results between the largest order statistics arising from multiple-outlier models. The following three consecutive theorems present different conditions, for which the usual stochastic order between the largest order statistics holds.

The distribution functions of $X_{n^*:n^*}(n_1^*,n_2^*)$ and $Y_{n^*:n^*}(n_1^*,n_2^*)$ are, respectively, given by

$$F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) = \psi_1\left[\sum_{i=1}^{n^*}\phi_1(F_i(x\lambda_i))\right] = \psi_1\left[n_1^*\phi_1\left(F_1\left(x\lambda_1\right)\right)+n_2^*\phi_1\left(F_2\left(x\lambda_2\right)\right)\right] \tag{1}$$

and

$$F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) = \psi_2\left[\sum_{i=1}^{n^*}\phi_2(F_i(x\mu_i))\right] = \psi_2\left[n_1^*\phi_2\left(F_1\left(x\mu_1\right)\right)+n_2^*\phi_2\left(F_2\left(x\mu_2\right)\right)\right]. \tag{2}$$

Li et al. (2016) established comparison results between two largest order statistics in the sense of the usual stochastic order when the underlying random variables follow the dependent scale models. The authors have used same baseline distribution.

In this section, we consider two sets of dependent observations where the random variables follow multiple-outliers dependent scale models sharing Archimedean copula with same/different generators. Here, the first n_1 observations of $\{X_1,\dots,X_{n_1},X_{n_1+1},\dots,X_n\}$ have baseline distribution

Table 1: List of several Archimedean copulas and generators

Archimedean copula	$\psi(x) \equiv \psi(x;\sigma)$	$\sigma \in$	ψ is log-convex	ψ is log-concave	$\frac{1-\psi}{\psi'}$ is decreasing	$[\frac{1-\psi}{\psi'}][\frac{1-\psi}{\psi'}]'$ is increasing	$[\frac{1-\psi}{\psi'}]'[\frac{\psi}{\psi'}]$ is increasing	$\psi_1(x)$	$\psi_2(x)$	$\phi_2 \circ \psi_1$ is super-additive
Independence copula	e^{-x}	-	Yes	Yes	Yes	Yes	Yes	e^{-x}	e^{-x}	Yes
Ali–Mikhail–Haq copula	$\frac{1-\sigma}{e^x-\sigma}$	$[-1,1)$	Yes $\sigma \in (0,1)$	Yes $\sigma \in [-1,0)$	Yes $\sigma \in [-1,1)$	Yes $\sigma \in [-1,1)$	-	$\frac{1-\sigma_1}{e^x-\sigma_1}$	$\frac{1-\sigma_2}{e^x-\sigma_2}$	Yes $0 \leq \sigma_1 < \sigma_2 < 1$
Gumbel–Barnett copula	$e^{\frac{1}{\sigma}(1-e^x)}$	$(0,1]$	No	Yes $\sigma \in (0,1]$	Yes $\sigma \in (0,1]$	-	$\sigma = 0.99$	$e^{\frac{1}{\sigma_1}(1-e^x)}$	$e^{\frac{1}{\sigma_2}(1-e^x)}$	Yes $0 < \sigma_2 \leq \sigma_1 \leq 1$
Gumbel–Hougaard copula	$e^{1-(1+x)^{\sigma}}$	$[1,\infty)$	No	Yes $\sigma \in [1,\infty)$	Yes $\sigma \in [1,\infty)$	-	-	$e^{1-(1+x)^{\sigma_1}}$	$e^{1-(1+x)^{\sigma_2}}$	Yes $\sigma_1 \leq \sigma_2$
Gumbel copula	$e^{-x^{\frac{1}{\sigma}}}$	$[1,\infty)$	Yes $\sigma \in [1,\infty)$	No	Yes $\sigma \in [1,\infty)$	-	-	$e^{-x^{\frac{1}{\sigma_1}}}$	$e^{-x^{\frac{1}{\sigma_2}}}$	Yes $\sigma_1 \leq \sigma_2$
Clayton copula	$(1+\sigma t)^{-\frac{1}{\sigma}}$	$[1,\infty)$	Yes $\sigma \in [1,\infty)$	No	Yes $\sigma \in [1,\infty)$	-	-	$(1+\sigma_1 t)^{-\frac{1}{\sigma_1}}$	$(1+\sigma_2 t)^{-\frac{1}{\sigma_2}}$	Yes $\sigma_1 \leq \sigma_2$
-	$(1-t)^{\sigma}$	$[1,\infty)$	No	Yes $\sigma \in [1,\infty)$	Yes $\sigma \in [1,\infty)$	Yes $\sigma \in [1,\infty)$	-			

function F_1 and remaining observations have baseline distribution function F_2. Furthermore, assume that first n_1^* observations of $\{Y_1,\ldots,Y_{n_1^*},Y_{n_1^*+1},\ldots,Y_{n^*}\}$ have a baseline distribution function F_1, and remaining $(n^*-n_1^*)$ observations have baseline distribution function F_2. It is also assumed that the number of observations $n\ (=n_1+n_2)$ and $n^*\ (=n_1^*+n_2^*)$ may be different. Now, we investigate whether the usual stochastic order between $Y_{n^*:n^*}(n_1^*,n_2^*)$ and $X_{n:n}(n_1,n_2)$ holds under this set-up. To obtain this result, we first prove following two theorems. In the following theorem, it is assumed that the dependence structures of two sets of samples having multiple-outliers are different. Also, first n_1^* observations of $\{X_1,\ldots,X_{n_1^*},X_{n_1^*+1},\ldots,X_{n^*}\}$ have baseline distribution function F_1 and remaining observations have baseline distribution function F_2. Before presenting the first result, we state the following assumption.

Assumption 1. *Let $X_1,\ldots,X_{n^*}$ be n^* dependent nonnegative random variables sharing Archimedean (survival) copula with generator ψ_1, with $X_i\sim F_1(x\lambda_1)$, for $i=1,\ldots,n_1^*$ and $X_j\sim F_2(x\lambda_2)$, for $j=n_1^*+1,\ldots,n^*$. Also, let $Y_1,\ldots,Y_{n^*}$ be n^* dependent nonnegative random variables sharing Archimedean copula with generator ψ_2, with $Y_i\sim F_1(x\mu_1)$, for $i=1,\ldots,n_1^*$ and $Y_j\sim F_2(x\mu_2)$, for $j=n_1^*+1,\ldots,n^*$. Here, $n_1^*+n_2^*=n^*$, $\psi_1=\phi_1^{-1}$ and $\psi_2=\phi_2^{-1}$.*

Theorem 1. *Under the set-up as in Assumption 1, let $\tilde r_1(x)\ge(\le)\tilde r_2(x)$ and let $n_1^*\ge(\le)n_2^*$. Also, let $\boldsymbol\lambda$, $\boldsymbol\mu\in\mathscr{E}_+\ (\mathscr{D}_+)$, $\phi_2\circ\psi_1$ be super-additive and let ψ_1 or ψ_2 be log-convex. Then,*

(i) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*})\succeq^w(\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*})\Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*)\le_{st}X_{n^*:n^*}(n_1^*,n_2^*)$, *provided $\tilde r_1(x)$ or $\tilde r_2(x)$ is decreasing;*

(ii) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*})\succeq^p(\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*})\Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*)\le_{st}X_{n^*:n^*}(n_1^*,n_2^*)$, *provided $x\tilde r_1(x)$ or $x\tilde r_2(x)$ is decreasing.*

Proof. Here, we prove the first part. Second part can be proved by similar argument. Denote $A(\boldsymbol\lambda,\psi_1,x)=F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$ and $B(\boldsymbol\mu,\psi_2,x)=F_{Y_{n^*:n^*}(n_1^*,n_2^*)}(x)$ (given in (1) and (2)). Using the fact that $\phi_2\circ\psi_1$ is super-additive, one can easily obtain $A(\boldsymbol\mu,\psi_1,x)\le B(\boldsymbol\mu,\psi_2,x)$. Therefore, to prove the desired result, we have to show that $A(\boldsymbol\lambda,\psi_1,x)\le A(\boldsymbol\mu,\psi_1,x)$. This is equivalent to establish that the function $A(\boldsymbol\lambda,\psi_1,x)$ is increasing and Schur-concave with respect to $\boldsymbol\lambda$ (see Marshall et al. 2011, Theorem $A.8$). Furthermore, on differentiating $A(\boldsymbol\lambda,\boldsymbol\psi_1,\mathbf{x})$ with respect to λ_i partially, we get

$$\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_i}=xn_i^*\tilde r_i(x\lambda_i)\frac{\psi_1\left[\phi_1\left[F_i\left(x\lambda_i\right)\right]\right]}{\psi_1'\left[\phi_1\left[F_i\left(x\lambda_i\right)\right]\right]}\psi_1'\left[\sum_{m=1}^{n^*}\phi_1(F_m(x\lambda_m))\right],\tag{3}$$

where $i=1,2$. From (3), it is not difficult to check that $\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_i}\ge0$ for $i=1,2$. Thus, $A(\boldsymbol\lambda,\psi_1,x)$ is increasing in λ_i, for $i=1,2$. To establish Schur-concavity of $A(\boldsymbol\lambda,\psi_1,x)$, in view of Lemma 2 (Lemma 1), we only need to show that for $1\le i\le j\le n^*$, the following inequality holds:

$$\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_i}-\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_j}\ge(\le)0,\text{ for }\boldsymbol\lambda\in\mathscr{E}_+\ (\mathscr{D}_+).\tag{4}$$

Next, consider three cases.
Case I: For $1\le i\le j\le n_1^*$, $\lambda_i=\lambda_j=\lambda_1$. In this case, $\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_i}-\frac{\partial A(\boldsymbol\lambda,\psi_1,x)}{\partial\lambda_j}=0$.

Case II: For $n_1^*+1 \le i \le j \le n^*$, $\lambda_i = \lambda_j = \lambda_2$. Here, $\frac{\partial A(\lambda,\psi_1,x)}{\partial \lambda_i} - \frac{\partial A(\lambda,\psi_1,x)}{\partial \lambda_j} = 0$.
Case III: For $1 \le i \le n_1^*$ and $n_1^*+1 \le j \le n^*$, $\lambda_i = \lambda_1$ and $\lambda_j = \lambda_2$. For this case, consider $\lambda_1 \le (\ge)\lambda_2$, which implies $\phi_1(F_1(x\lambda_1)) \ge (\le)\phi_1(F_1(x\lambda_2))$. Furthermore, under the given assumption, we get $\phi_1(F_1(x\lambda_2)) \ge (\le)\phi_1(F_2(x\lambda_2))$. Hence, $\phi_1(F_1(x\lambda_1)) \ge (\le)\phi_1(F_2(x\lambda_2))$. Again, ψ_1 is log-convex. Therefore, we have

$$-\frac{\psi_1(w)}{\psi_1'(w)}\Big|_{w=\phi_1[F_1(x\lambda_1)]} \ge (\le) -\frac{\psi_1(w)}{\psi_1'(w)}\Big|_{w=\phi_1[F_2(x\lambda_2)]}. \tag{5}$$

Moreover, $\tilde{r}_1(w)$ is decreasing in $w>0$; hence

$$\tilde{r}_1(x\lambda_1) \ge (\le)\tilde{r}_1(x\lambda_2). \tag{6}$$

Also, $\tilde{r}_1(x) \ge (\le)\tilde{r}_2(x)$ gives

$$\tilde{r}_1(x\lambda_2) \ge (\le)\tilde{r}_2(x\lambda_2). \tag{7}$$

Equations (6), (7) and $n_1^* \ge (\le)n_2^*$, together imply

$$n_1^*\tilde{r}_1(x\lambda_1) \ge (\le)n_2^*\tilde{r}_2(x\lambda_2). \tag{8}$$

Finally, combining (5) and (8), we obtain (4). $\square$

The following corollary, which is a direct consequence of Theorem 1 presents some special cases.

Corollary 1. *In addition to Assumption 1, let $\psi_1 = \psi_2 = \psi$, let $n_1^* \ge (\le)n_2^*$, and let ψ be log-convex. Furthermore, let λ, $\mu \in \mathscr{E}_+$ $(\mathscr{D}_+)$. Then,*

(i) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*) \le_{st} X_{n^*:n^*}(n_1^*,n_2^*)$, *provided* $\tilde{r}_1(x)$ *or* $\tilde{r}_2(x)$ *is decreasing and* $\tilde{r}_1(x) \ge (\le)\tilde{r}_2(x)$;

(ii) for $\tilde{r}_1 = \tilde{r}_2 = \tilde{r}$, *we have*

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*) \le_{st} X_{n^*:n^*}(n_1^*,n_2^*),$$

provided $\tilde{r}(x)$ *is decreasing.*

The next theorem states that the ordering result holds between the largest order statistics $X_{n:n}(n_1,n_2)$ and $X_{n^*:n^*}(n_1^*,n_2^*)$ according to the usual stochastic ordering. Here, the samples are collected from multiple-outlier dependent scale models. Also, it is assumed that the samples are sharing Archimedean copula with a common generator.

Assumption 2. *Let $X_1,\ldots,X_{n^*}$ be n^* dependent nonnegative random variables sharing Archimedean (survival) copula with generator ψ_1, such that $X_i \sim F_1(x\lambda_1)$, for $i = 1,\ldots,n_1^*$ and $X_j \sim F_2(x\lambda_2)$, for $j = n_1^*+1,\ldots,n^*$. We assume that there exist two natural numbers n_1 and n_2 such that $1 \le n_1 \le n_1^* \le n_2^* \le n_2$. Also, $n = n_1+n_2$, $n^* = n_1^*+n_2^*$ and $\psi_1 = \phi_1^{-1}$.*

Theorem 2. *Let Assumption 2 hold with $F_1 \geq F_2$. Then, for $\lambda \in \mathscr{D}_+$, we have*

$$(n_1, n_2) \succeq_w (n_1^*, n_2^*) \Rightarrow X_{n^*:n^*}(n_1^*, n_2^*) \leq_{st} X_{n:n}(n_1, n_2).$$

Proof. The distribution functions of $X_{n:n}(n_1, n_2)$ and $X_{n^*:n^*}(n_1^*, n_2^*)$ can be, respectively, written as

$$F_{X_{n:n}(n_1,n_2)}(x) = \psi_1\left(\sum_{i=1}^{n} \phi_1\left(F_i\left(x\lambda_i\right)\right)\right) = \psi_1[n_1\phi_1\left(F_1\left(x\lambda_1\right)\right) + n_2\phi_1\left(F_2\left(x\lambda_2\right)\right)]$$

and

$$F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) = \psi_1\left(\sum_{i=1}^{n^*} \phi_1\left(F_i\left(x\lambda_i\right)\right)\right) = \psi_1[n_1^*\phi_1\left(F_1\left(x\lambda_1\right)\right) + n_2^*\phi_1\left(F_2\left(x\lambda_2\right)\right)].$$

To obtain the desired result, one needs to show $F_{X_{n:n}(n_1,n_2)}(x) \leq F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$, which is equivalent to establish

$$\begin{aligned}
&\psi_1\left(\sum_{i=1}^{n} \phi_1\left(F_i\left(x\lambda_i\right)\right)\right) \leq \psi_1\left(\sum_{i=1}^{n^*} \phi_1\left(F_i\left(x\lambda_i\right)\right)\right)\\
&\Rightarrow n_1\phi_1\left(F_1\left(x\lambda_1\right)\right) + n_2\phi_1\left(F_2\left(x\lambda_2\right)\right) \geq n_1^*\phi_1\left(F_1\left(x\lambda_1\right)\right) + n_2^*\phi_1\left(F_2\left(x\lambda_2\right)\right)\\
&\Rightarrow (n_1^* - n_1)\phi_1\left(F_1\left(x\lambda_1\right)\right) \leq (n_2 - n_2^*)\phi_1\left(F_2\left(x\lambda_2\right)\right). \qquad (9)
\end{aligned}$$

Now, $(n_1, n_2) \succeq_w (n_1^*, n_2^*) \Rightarrow (n_1 + n_2) \geq (n_1^* + n_2^*) \Rightarrow (n_2 - n_2^*) \geq (n_1^* - n_1) \geq 0$. Also, $\lambda_1 \geq \lambda_2 \Rightarrow \phi_1\left(F_2\left(x\lambda_2\right)\right) \geq \phi_1\left(F_1\left(x\lambda_1\right)\right) \geq 0$. Combining above two inequalities, we have the inequality given in (9). □

Next, we observe that two largest order statistics $X_{n:n}(n_1, n_2)$ and $Y_{n^*:n^*}(n_1^*, n_2^*)$ are comparable with respect to the usual stochastic order under the presence of Archimedean copula. It is worth mentioning that the order statistics are constructed from two multiple-outlier dependent samples having sample sizes n and n^*. The pairs of the sizes of both the outliers (n_1, n_2) and (n_1^*, n_2^*) are assumed to be connected according to the weakly submajorization order and p-larger order. The following assumption is useful for the next theorem.

Assumption 3. *Let $X_1, \ldots, X_n$ be n nonnegative dependent random variables sharing Archimedean (survival) copula with generator ψ_1, such that $X_i \sim F_1(x\lambda_1)$, for $i = 1, \ldots, n_1$ and $X_j \sim F_2(x\lambda_2)$, for $j = n_1 + 1, \ldots, n$. Also, let $Y_1, \ldots, Y_{n^*}$ be n^* dependent nonnegative random variables sharing Archimedean copula with generator ψ_2, such that $Y_i \sim F_1(x\mu_1)$, for $i = 1, \ldots, n_1^*$ and $Y_j \sim F_2(x\mu_2)$, for $j = n_1^* + 1, \ldots, n^*$. Here, $1 \leq n_1 \leq n_1^* \leq n_2^* \leq n_2$, $n = n_1 + n_2$ and $n^* = n_1^* + n_2^*$.*

The following result states as remark immediately follows from Theorems 1 and 2.

Remark 1. *Assume that Assumption 3 holds with $\tilde{r}_1(x) \leq \tilde{r}_2(x)$. Let $(n_1, n_2) \succeq_w (n_1^*, n_2^*)$, let λ, $\mu \in \mathscr{D}_+$, let $\phi_2 \circ \psi_1$ be super-additive, and let ψ_1 or ψ_2 be log-convex.*

(i) $(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1^*}, \underbrace{\lambda_2, \ldots, \lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1, \ldots, \mu_1}_{n_1^*}, \underbrace{\mu_2, \ldots, \mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*, n_2^*) \leq_{st} X_{n:n}(n_1, n_2)$, *provided $\tilde{r}_1(x)$ or $\tilde{r}_2(x)$ is decreasing;*

(ii) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^p (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*) \leq_{st} X_{n:n}(n_1,n_2)$, *provided* $x\tilde{r}_1(x)$ *or* $x\tilde{r}_2(x)$ *is decreasing.*

The condition stated in Remark 1 "$\phi_2 \circ \psi_1$ is super-additive" demonstrates that the copula having generator ψ_2 is more positively dependent than that having generator ψ_1. Therefore, from Remark 1(i), one can conclude that in the presence of the Archimedean copula for more positively dependent components having multiple-outlier dependent scale models, less heterogeneous scale parameter vector (with respect to the weakly supermajorized order) leads to less reliable parallel system in the sense of the stochastically smaller lifetime.

Remark 2. *From Table 1, one can see that in literature, there are several Archimedean copulas whose generators are satisfied all the conditions of Theorems 1 and 2 and Remark 1. For example, we can consider Independence copula, Clayton copula, Ali-Mikhail-Haq copula, and Gumbel copula.*

We now present a numerical example, which provides an illustration of the result in Remark 1.

Example 1. Set $\lambda = (\lambda_1,\lambda_2) = (5,2)$, $\mu = (\mu_1,\mu_2) = (6,3)$, $(n_1,n_2) = (1,11)$, $(n_1^*,n_2^*) = (5,6)$, $\psi_1(x) = e^{-x^{\frac{1}{9}}}$, $\psi_2(x) = e^{-x^{\frac{1}{10}}}$, $x > 0$. Consider the baseline distribution functions as $F_2(x) = 1 - e^{1-(1+x^2)^{\frac{1}{5}}}$ and $F_1(x) = 1 - e^{-x}$, $x > 0$. Here, both the reversed hazard rate functions $\tilde{r}_1$ and $\tilde{r}_2$ are decreasing and satisfy $\tilde{r}_1(x) \leq \tilde{r}_2(x)$, for $x > 0$. Furthermore, ψ_1 and ψ_2 are log-convex and $\phi_2 \circ \psi_1$ is super-additive (see Table 1). Thus, all the conditions of Remark 1 are satisfied. Now, we plot the graphs of $F_{X_{12:12}(1,11)}(x)$ and $F_{Y_{11:11}(5,6)}(x)$ in Figure 1, which shows that $Y_{11:11}(5,6) \leq_{st} X_{12:12}(1,11)$ holds.

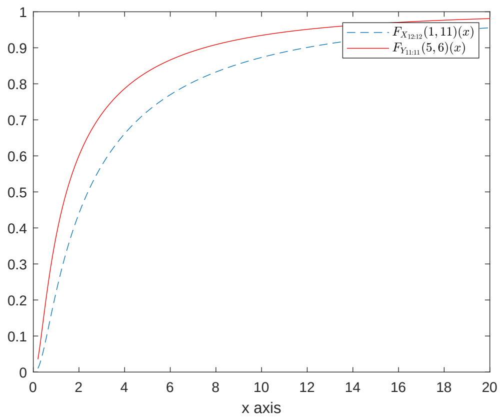


Figure 1: Plots of the distribution functions $F_{X_{12:12}(1,11)}(x)$ and $F_{Y_{11:11}(5,6)}(x)$ as in Example 1.

Next, we present a counterexample to illustrate that the result does not hold if $\tilde{r}_1(x) \geq \tilde{r}_2(x)$ and $\lambda \in \mathscr{E}_+$ in Remark 1.

Counterexample 1. Consider $\lambda = (\lambda_1, \lambda_2) = (2,6)$, $\mu = (\mu_1, \mu_2) = (8,2)$, $(n_1, n_2) = (1,8)$, $(n_1^*, n_2^*) = (3,4)$, $\psi_1(x) = e^{-x^{\frac{1}{3}}}$, $\psi_2(x) = e^{-x^{\frac{1}{10}}}$, $x > 0$. Baseline distribution functions are taken as $F_1(x) = 1 - e^{-x}$ and $F_2(x) = 1 - (1+2x)^{-0.5}$, $x > 0$. It can be seen that all the conditions of Remark 1 are satisfied except $\lambda \in \mathscr{D}_+$ and $\tilde{r}_1(x) \le \tilde{r}_2(x)$. Now, we plot the graph of $F_{X_{9:9}(1,8)}(x) - F_{Y_{7:7}(3,4)}(x)$ in Figure 2, which reveals that $Y_{7:7}(3,4) \nleq_{st} X_{9:9}(1,8)$.

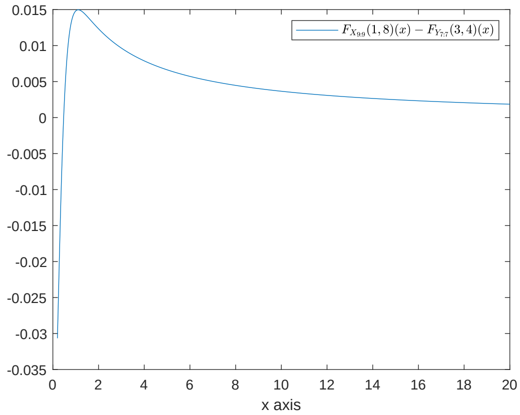


Figure 2: Plot of $F_{X_{9:9}(1,8)}(x) - F_{Y_{7:7}(3,4)}(x)$ as in Counterexample 1

In the preceding results, we have derived sufficient conditions, under which the largest order statistics from multiple-outlier dependent scale models obey the usual stochastic order. However, naturally, it is of interest to extend the ordering results to some other stronger concepts of the stochastic orders. In this part of the subsection, we establish sufficient conditions, under which the reversed hazard rate order holds between the largest order statistics. The following theorem shows that the largest order statistics $X_{n^*:n^*}(n_1^*, n_2^*)$ and $Y_{n^*:n^*}(n_1^*, n_2^*)$ have the reversed hazard rate ordering when the scale parameters are associated with the weakly supermajorization order. The samples are heterogeneous and follow multiple-outlier dependent scale models.

Theorem 3. *Let Assumption 1 hold with $r_1 = r_2 = r$, $n_1^* \ge (\le)\ n_2^*$ and $\psi_1 = \psi_2 = \psi$. Also, suppose that ψ is log-concave, that $\frac{1-\psi}{\psi'}$ is decreasing, and that $\frac{1-\psi}{\psi'}[\frac{1-\psi}{\psi'}]'$ is increasing. Then,*

$$(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1^*}, \underbrace{\lambda_2, \ldots, \lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1, \ldots, \mu_1}_{n_1^*}, \underbrace{\mu_2, \ldots, \mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*, n_2^*) \le_{rh} X_{n^*:n^*}(n_1^*, n_2^*),$$

provided λ, $\mu \in \mathscr{E}_+\ (\mathscr{D}_+)$, $r(x)$ is decreasing and $xr(x)$ is decreasing and convex.

Proof. Under the given assumption, $r_1 = r_2 = r$ implies $F_1 = F_2 = F$. The reversed hazard rate function of $X_{n^*:n^*}(n_1^*,n_2^*)$ is

$$\begin{aligned}\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) &= \frac{\psi'\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}\left[\sum_{i=1}^{n^*}\frac{n_1^*\lambda_i f(x\lambda_i)}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}\right] \\ &= \frac{\psi'\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}\left[\sum_{i=1}^{n^*}\frac{n_1^*\lambda_i r(x\lambda_i)[1-\psi[\phi\left(F\left(x\lambda_i\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}\right],\end{aligned}$$

where f is the probability density function corresponding to F. Denote $z = n_1^*\phi\left(F\left(x\lambda_1\right)\right) + n_2^*\phi\left(F\left(x\lambda_2\right)\right)$. The partial derivative of $\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$ with respect to λ_i is obtained as

$$\begin{aligned}\frac{\partial[\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)]}{\partial\lambda_i} &= n_i^* x r(x\lambda_i)\frac{d}{dz}\left[\frac{\psi'(z)}{\psi(z)}\right]\left[\frac{1-\psi\left[\phi\left[F\left(x\lambda_i\right)\right]\right]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}\right]\left[\sum_{i=1}^{n^*}\frac{\lambda_i f\left(x\lambda_i\right)}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}\right] \\ &\quad + n_i^* x\lambda_i\left[r(x\lambda_i)\right]^2\frac{\psi'(z)}{\psi(z)}\left[\frac{1-\psi(v)}{\psi'(v)}\frac{d}{dv}\left[\frac{1-\psi(v)}{\psi'(v)}\right]\right]_{v=\phi(F(x\lambda_i))} \\ &\quad + n_i^*\frac{d}{dw}\left[wr(w)\right]_{w=x\lambda_i}\left[\frac{1-\psi\left[\phi\left[F\left(x\lambda_i\right)\right]\right]}{\psi'\left[\phi\left[F\left(x\lambda_i\right)\right]\right]}\right]\frac{\psi'(z)}{\psi(z)}, \qquad (10)\end{aligned}$$

for $i = 1,2$. Utilizing (Marshall et al., 2011, Theorem A.8), to obtain the desired result, we need to prove that $\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$ is decreasing and Schur-convex with respect to λ. Using the given assumptions and (10), the decreasing property of $\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$ with respect to λ is obvious. Furthermore, according to Lemma 2 (Lemma 1), to show the Schur-convexity of $\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$, we have to establish for $1 \le i \le j \le n^*$,

$$\left[\frac{\partial[\tilde{r}_{X_{n^*:n^*}}(n_1^*,n_2^*)(x)]}{\partial\lambda_i} - \frac{\partial[\tilde{r}_{X_{n^*:n^*}}(n_1^*,n_2^*)(x)]}{\partial\lambda_j}\right] \le (\ge) 0, \text{ for } \lambda \in \mathscr{E}_+\ (\mathscr{D}_+). \qquad (11)$$

Now, consider the following three cases.

Case I: For $1 \le i \le j \le n_1^*$, $\lambda_i = \lambda_j = \lambda_1$. Here, $\frac{\partial[\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)]}{\partial\lambda_i} - \frac{\partial[\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)]}{\partial\lambda_j} = 0$.

Case II: For $n_1^*+1 \le i \le j \le n^*$, $\lambda_i = \lambda_j = \lambda_2$. Hence, $\frac{\partial[\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)]}{\partial\lambda_i} - \frac{\partial[\tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)]}{\partial\lambda_j} = 0$.

Case III: For $1 \le i \le n_1^*$ and $n_1^*+1 \le j \le n^*$, $\lambda_i = \lambda_1$ and $\lambda_j = \lambda_2$. Consider $\lambda_1 \le \lambda_2$, which gives $\phi(F(x\lambda_1)) \ge \phi(F(x\lambda_2))$. Here, we only discuss the proof when $\lambda_1 \le \lambda_2$. The other case when $\lambda_1 \ge \lambda_2$ can be proved in the similar way. The concavity property of $\ln\psi$ provides $\frac{d}{dz}\left[\frac{\psi'(z)}{\psi(z)}\right] \le 0$. Again, using decreasing property of $\frac{1-\psi}{\psi'}$, we have

$$\left.\frac{1-\psi(w)}{\psi'(w)}\right|_{w=\phi[F(x\lambda_1)]} \le \left.\frac{1-\psi(w)}{\psi'(w)}\right|_{w=\phi[F(x\lambda_2)]} \le 0. \qquad (12)$$

Furthermore, it has been assumed that $r(x)$ is decreasing and that $xr(x)$ is decreasing and convex. Therefore, using $n_1^* \ge n_2^*$, we have

$$r(x\lambda_1) \ge r(x\lambda_2), \qquad (13)$$

$$n_1^* x\lambda_1 r(x\lambda_1) \ge n_2^* x\lambda_2 r(x\lambda_2), \qquad \text{and} \qquad (14)$$

$$n_1^* \frac{d}{dw}[wr(w)]_{w=x\lambda_1} \leq n_2^* \frac{d}{dw}[wr(w)]_{w=x\lambda_2} \leq 0. \tag{15}$$

Moreover, $\frac{1-\psi(w)}{\psi'(w)} \frac{d}{dw}[\frac{1-\psi(w)}{\psi'(w)}]$ is increasing. Therefore, we obtain the following inequality:

$$\left[\frac{1-\psi(w)}{\psi'(w)} \frac{d}{dw}\left[\frac{1-\psi(w)}{\psi'(w)}\right]\right]_{w=\phi[F(x\lambda_1)]} \geq \left[\frac{1-\psi(w)}{\psi'(w)} \frac{d}{dw}\left[\frac{1-\psi(w)}{\psi'(w)}\right]\right]_{w=\phi[F(x\lambda_2)]} \geq 0. \tag{16}$$

Now, combining (12)–(16) and the given assumptions, we obtain that the inequality given by (11) holds. □

In the next theorem, we show that the largest order statistics $X_{n:n}(n_1,n_2)$ and $X_{n^*:n^*}(n_1^*,n_2^*)$ are comparable according to the reversed hazard rate order.

Theorem 4. *Let Assumption 2 hold with $\psi_1 = \psi$ and $r_1 = r_2 = r$. Then, for $\lambda \in \mathscr{D}_+$, we have*

$$(n_1,n_2) \succeq_w (n_1^*,n_2^*) \Rightarrow X_{n^*:n^*}(n_1^*,n_2^*) \leq_{rh} X_{n:n}(n_1,n_2),$$

provided $\ln\psi$ is concave and $\frac{1-\psi}{\psi'}$ and $xr(x)$ are decreasing.

Proof. The stated result will be proved, if we show that $\tilde{r}_{X_{n:n}(n_1,n_2)}(x) \geq \tilde{r}_{X_{n^*:n^*}(n_1^*,n_2^*)}(x)$. Equivalently,

$$\frac{\psi'\left[\sum_{i=1}^{n}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi\left[\sum_{i=1}^{n}\phi\left(F\left(x\lambda_i\right)\right)\right]} \times \frac{\psi\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi'\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]} \geq \frac{\sum_{i=1}^{n^*}\frac{\lambda_i r(x\lambda_i)[1-\psi[\phi\left(F\left(x\lambda_i\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}}{\sum_{i=1}^{n}\frac{\lambda_i r(x\lambda_i)[1-\psi[\phi\left(F\left(x\lambda_i\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]}}. \tag{17}$$

The preceding inequality holds if the following two inequalities are satisfied:

$$\frac{\psi'\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi\left[\sum_{i=1}^{n^*}\phi\left(F\left(x\lambda_i\right)\right)\right]} \geq \frac{\psi'\left[\sum_{i=1}^{n}\phi\left(F\left(x\lambda_i\right)\right)\right]}{\psi\left[\sum_{i=1}^{n}\phi\left(F\left(x\lambda_i\right)\right)\right]} \Leftrightarrow (n_1^*-n_1)\phi(F(x\lambda_1)) \leq (n_2-n_2^*)\phi(F(x\lambda_2)) \tag{18}$$

and

$$\begin{aligned} \sum_{i=1}^{n^*}\frac{\lambda_i r(x\lambda_i)[1-\psi[\phi\left(F\left(x\lambda_i\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]} &\geq \sum_{i=1}^{n}\frac{\lambda_i r(x\lambda_i)[1-\psi[\phi\left(F\left(x\lambda_i\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_i\right)\right)]} \\ \Leftrightarrow (n_1^*-n_1)\frac{\lambda_1 r(x\lambda_1)[1-\psi[\phi\left(F\left(x\lambda_1\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_1\right)\right)]} &\geq (n_2-n_2^*)\frac{\lambda_2 r(x\lambda_2)[1-\psi[\phi\left(F\left(x\lambda_2\right)\right)]]}{\psi'[\phi\left(F\left(x\lambda_2\right)\right)]}. \end{aligned} \tag{19}$$

Furthermore,

$$(n_1,n_2) \succeq_w (n_1^*,n_2^*) \Rightarrow (n_1+n_2) \geq (n_1^*+n_2^*) \Rightarrow (n_2-n_2^*) \geq (n_1^*-n_1) \geq 0. \tag{20}$$

Also,

$$\lambda_1 \geq \lambda_2 \Rightarrow \phi\left(F\left(x\lambda_2\right)\right) \geq \phi\left(F\left(x\lambda_1\right)\right) \geq 0.$$

Moreover, $\frac{1-\psi}{\psi'}$ is decreasing. Thus,

$$\left.\frac{1-\psi(w)}{\psi'(w)}\right|_{w=\phi[F(x\lambda_2)]} \leq \left.\frac{1-\psi(w)}{\psi'(w)}\right|_{w=\phi[F(x\lambda_1)]} \leq 0. \tag{21}$$

Using the decreasing property of $xr(x)$, we have

$$x\lambda_1 r(x\lambda_1) \leq x\lambda_2 r(x\lambda_2). \tag{22}$$

Combining (20), (21), and (22), the inequality (19) can be obtained. Using (20) and the assumption that ψ is log-concave, we get the inequality (18). $\square$

Now, we are ready to state a result which shows that the largest order statistics $X_{n:n}(n_1,n_2)$ and $Y_{n^*:n^*}(n_1^*,n_2^*)$ can be compared with respect to the reversed hazard rate order. Here, n and n^* may be different.

Theorem 5. *Let the set-up in Assumption 3 hold with $\psi_1 = \psi_2 = \psi$ and $r_1 = r_2 = r$. Also, assume $\lambda,\ \mu \in \mathscr{D}_+$ and $(n_1,n_2) \succeq_w (n_1^*,n_2^*)$. Then,*

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*) \leq_{rh} X_{n:n}(n_1,n_2),$$

provided ψ is log-concave, $\frac{1-\psi}{\psi'}$ is decreasing, $\frac{1-\psi}{\psi'}[\frac{1-\psi}{\psi'}]'$ is increasing, $xr(x)$ is decreasing and convex, and $r(x)$ is decreasing.

Proof. According to Theorem 3, we have

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*,n_2^*) \leq_{rh} X_{n^*:n^*}(n_1^*,n_2^*). \tag{23}$$

Also, from Theorem 4, we get

$$(n_1,n_2) \succeq_w (n_1^*,n_2^*) \Rightarrow X_{n^*:n^*}(n_1^*,n_2^*) \leq_{rh} X_{n:n}(n_1,n_2). \tag{24}$$

Thus, the proof of the theorem follows after combining the inequalities given by (23) and (24). $\square$

Below, we consider an example to illustrate Theorem 5.

Example 2. Consider $\lambda = (3,2)$, $\mu = (6,5)$, $(n_1,n_2) = (2,10)$, $(n_1^*,n_2^*) = (3,4)$, $\psi(x) = e^{\frac{1}{d}(1-e^x)}$, $x > 0$. Also, let the baseline distribution be $F(x) = 1 - \left(\frac{x}{b}\right)^{-a}$, $x \geq b > 0$, $a > 0$. It is not hard to see that, for $d = 0.2$, $a = 5$ and $b = 1$, all the conditions of Theorem 5 are satisfied (see Table 1). Furthermore, we plot the graph of $F_{X_{12:12}(2,10)}(x)/F_{Y_{7:7}(3,4)}(x)$ in Figure 3. This shows that the result in Theorem 5 holds.

Remark 3. *Let us consider (i) Independence copula with generator $\psi(x) = e^{-x}$, $x > 0$ and (ii) Ali-Mikhail-Haq copula with generator $\psi(x) = \frac{1-\sigma}{e^x-\sigma}$, $\sigma \in [-1,0)$, $x > 0$. One can easily check that the above two copulas satisfies all the conditions of Theorems 3– 5 (see Table 1).*

Now, we derive conditions such that the star order holds between $X_{n^*:n^*}(n_1^*,n_2^*)$ and $Y_{n^*:n^*}(n_1^*,n_2^*)$. Denote $\lambda_{2:2} = \max\{\lambda_1,\lambda_2\}$, $\lambda_{1:2} = \min\{\lambda_1,\lambda_2\}$, $\mu_{2:2} = \max\{\mu_1,\mu_2\}$, and $\mu_{1:2} = \min\{\mu_1,\mu_2\}$.

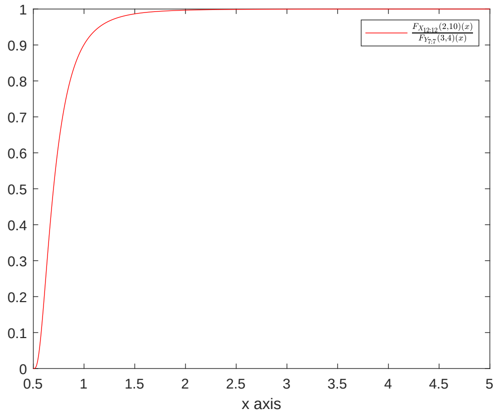


Figure 3: Plot of the ratio of two distribution functions $F_{X_{12:12}(2,10)}(x)/F_{Y_{7:7}(3,4)}(x)$ in Example 2.

Theorem 6. *Under the set-up as in Assumption 1, with $\tilde{r}_1 = \tilde{r}_2 = \tilde{r}$ and $\psi_1 = \psi_2 = \psi$, we have*

$$\frac{\lambda_{2:2}}{\lambda_{1:2}} \geq \frac{\mu_{2:2}}{\mu_{1:2}} \Rightarrow Y_{n^*:n^*}(n_1^*, n_2^*) \leq_* X_{n^*:n^*}(n_1^*, n_2^*),$$

provided $\frac{\psi}{\psi'}$ is decreasing and convex, $\frac{x\tilde{r}'(x)}{\tilde{r}(x)}$ is decreasing, and $x\tilde{r}(x)$ is increasing.

Proof. Under the assumption, $\tilde{r}_1 = \tilde{r}_2 = \tilde{r}$ gives $F_1 = F_2 = F$. The distribution functions of $X_{n^*:n^*}(n_1^*, n_2^*)$ and $Y_{n^*:n^*}(n_1^*, n_2^*)$ are, respectively, given by

$$F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) = \psi\left[n_1^*\phi\left(F\left(x\lambda_1\right)\right) + n_2^*\phi\left(F\left(x\lambda_2\right)\right)\right]$$

and

$$F_{Y_{n^*:n^*}(n_1^*,n_2^*)}(x) = \psi\left[n_1^*\phi\left(F\left(x\mu_1\right)\right) + n_2^*\phi\left(F\left(x\mu_2\right)\right)\right].$$

To obtain the required result, we consider two cases.
Case I: $\lambda_1 + \lambda_2 = \mu_1 + \mu_2$.

For convenience, we assume $\lambda_1 + \lambda_2 = \mu_1 + \mu_2 = 1$. For this case, it is clear that $(\lambda_1, \lambda_2) \succeq^m (\mu_1, \mu_2)$. Now, take $\lambda_2 = \lambda \geq \lambda_1$, $\mu_2 = \mu \geq \mu_1$. Hence, $\lambda_1 = 1 - \lambda$ and $\mu_1 = 1 - \mu$. Based on this, the distribution functions of $X_{n^*:n^*}(n_1^*, n_2^*)$ and $Y_{n^*:n^*}(n_1^*, n_2^*)$ can be written in the following form

$$F_{X_{n^*:n^*}(n_1^*,n_2^*)}(x) \stackrel{def}{=} F_\lambda(x) = \psi\left[n_1^*\phi\left(F\left(x(1-\lambda)\right)\right) + n_2^*\phi\left(F\left(x\lambda\right)\right)\right]$$

and

$$F_{Y_{n^*:n^*}(n_1^*,n_2^*)}(x) \stackrel{def}{=} F_\mu(x) = \psi\left[n_1^*\phi\left(F\left(x(1-\mu)\right)\right) + n_2^*\phi\left(F\left(x\mu\right)\right)\right].$$

Now, according to Lemma 3, we have to show that $\frac{F_\lambda'(x)}{xf_\lambda(x)}$ is decreasing in $x \in \mathbb{R}^+$, for $\lambda \in (1/2, 1]$. The derivative of F_λ, with respect to λ is given by

$$\begin{aligned} F_\lambda'(x) &= \left[-xn_1^*\tilde{r}(x(1-\lambda))\frac{\psi[\phi\left(F\left(x(1-\lambda)\right)\right)]}{\psi'[\phi\left(F\left(x(1-\lambda)\right)\right)]} + xn_2^*\tilde{r}(x\lambda)\frac{\psi[\phi\left(F\left(x\lambda\right)\right)]}{\psi'[\phi\left(F\left(x\lambda\right)\right)]}\right] \\ &\quad \times \psi'\left[n_1^*\phi\left(F\left(x(1-\lambda)\right)\right) + n_2^*\phi\left(F\left(x\lambda\right)\right)\right]. \end{aligned} \tag{25}$$

Also, the probability density function corresponding to F_λ is

$$\begin{aligned} f_\lambda(x) &= \left[(1-\lambda)n_1^*\tilde{r}(x(1-\lambda))\frac{\psi[\phi(F(x(1-\lambda)))]}{\psi'[\phi(F(x(1-\lambda)))]} + \lambda n_2^*\tilde{r}(x\lambda)\frac{\psi[\phi(F(x\lambda))]}{\psi'[\phi(F(x\lambda))]}\right] \\ &\quad \times \psi'\left[n_1^*\phi(F(x(1-\lambda))) + n_2^*\phi(F(x\lambda))\right]. \end{aligned} \tag{26}$$

Therefore,

$$\frac{F'_\lambda(x)}{xf_\lambda(x)} = \left(\lambda + \left[\frac{n_2^*\tilde{r}(x\lambda)\dfrac{\psi[\phi(F(x\lambda))]}{\psi'[\phi(F(x\lambda))]}}{n_1^*\tilde{r}(x(1-\lambda))\dfrac{\psi[\phi(F(x(1-\lambda)))]}{\psi'[\phi(F(x(1-\lambda)))]}} - 1\right]^{-1}\right)^{-1}.$$

Thus, it suffices to show that $L(x) = \left(\tilde{r}(x\lambda)\frac{\psi[\phi(F(x\lambda))]}{\psi'[\phi(F(x\lambda))]}\right) / \left(\tilde{r}(x(1-\lambda))\frac{\psi[\phi(F(x(1-\lambda)))]}{\psi'[\phi(F(x(1-\lambda)))]}\right)$ is decreasing in $x \in \mathbb{R}^+$, for $\lambda \in (1/2, 1]$. The derivative of $L(x)$ with respect to x is obtained as

$$\begin{aligned} L'(x) &\overset{sign}{=} \frac{\lambda\tilde{r}'(x\lambda)}{\tilde{r}(x\lambda)} + \lambda\tilde{r}(x\lambda)\left[\frac{\psi[\phi(F(x\lambda))]}{\psi'[\phi(F(x\lambda))]}\right]' - \frac{(1-\lambda)\tilde{r}'(x(1-\lambda))}{\tilde{r}((1-\lambda)x)} \\ &\quad -(1-\lambda)\tilde{r}(x(1-\lambda))\left[\frac{\psi[\phi(F(x(1-\lambda)))]}{\psi'[\phi(F(x(1-\lambda)))]}\right]'. \end{aligned}$$

Under the assumptions made, $\frac{x\tilde{r}'(x)}{\tilde{r}(x)}$ is decreasing and $x\tilde{r}(x)$ is increasing. Therefore, for $\lambda \in (1/2, 1]$,

$$\frac{x\lambda\tilde{r}'(x\lambda)}{\tilde{r}(x\lambda)} \le \frac{x(1-\lambda)\tilde{r}'(x(1-\lambda))}{\tilde{r}(x(1-\lambda))} \le 0 \text{ and } x\lambda\tilde{r}(x\lambda) \ge x(1-\lambda)\tilde{r}(x(1-\lambda)) \ge 0. \tag{27}$$

Also, since $\frac{\psi}{\psi'}$ is decreasing and convex, we have

$$\left[\frac{\psi[\phi(F(x\lambda))]}{\psi'[\phi(F(x\lambda))]}\right]' \le \left[\frac{\psi[\phi(F(x(1-\lambda)))]}{\psi'[\phi(F(x(1-\lambda)))]}\right]' \le 0. \tag{28}$$

Now, combining (27) and (28), we get $L'(x) \le 0$, for $x \in \mathbb{R}^+$.
Case II. $\lambda_1 + \lambda_2 \ne \mu_1 + \mu_2$.

In this case, we can take $\lambda_1 + \lambda_2 = k(\mu_1 + \mu_2)$, where k is a scalar. Hence, $(k\mu_1, k\mu_2) \preceq^m (\lambda_1, \lambda_2)$. Let us consider n^* dependent nonnegative random variables sharing Archimedean copula with generator ψ, such that $Z_i \sim F(k\mu_1 x)$, for $i = 1, \ldots, n_1^*$ and $Z_j \sim F(k\mu_2 x)$, for $j = n_1^* + 1, \ldots, n^*$. Here, $n_1^* + n_2^* = n^*$. Then, from Case I, we have $Z_{n^*:n^*}(n_1^*, n_2^*) \le_* X_{n^*:n^*}(n_1^*, n_2^*)$. Furthermore, star order is scale invariant, and hence we obtain $Y_{n^*:n^*}(n_1^*, n_2^*) \le_* X_{n^*:n^*}(n_1^*, n_2^*)$. □

Remark 4. *Let us consider* $\psi(x) = e^{-x}$, $x > 0$ *in Theorem 6. Then, Theorem 6 reduces to (Zhang et al., 2019, Theorem 3.14) when* $\alpha_1 = \alpha_2 = 1$.

Using the fact that the star order implies the Lorenz order, the following result is a direct consequence of Theorem 6. Furthermore, since the Lorenz order is mainly used to compare the income distributions, the following corollary is more interesting from the point of its applications in the study of incomes.

Corollary 2. *Under the set-up as in Theorem 6,*

$$\frac{\lambda_{2:2}}{\lambda_{1:2}} \geq \frac{\mu_{2:2}}{\mu_{1:2}} \Rightarrow Y_{n^*:n^*}(n_1^*, n_2^*) \leq_{Lorenz} X_{n^*:n^*}(n_1^*, n_2^*),$$

provided $\frac{\psi}{\psi'}$ *is decreasing and convex,* $\frac{x\tilde{r}'(x)}{\tilde{r}(x)}$ *is decreasing, and* $x\tilde{r}(x)$ *is increasing.*

In the following theorem, an interesting result has been developed to compare $Y_{n^*:n^*}(n_1^*, n_2^*)$ and $X_{n^*:n^*}(n_1^*, n_2^*)$ according to the dispersive order.

Theorem 7. *Under the same assumptions of Theorem 6, let us consider* $\frac{\psi}{\psi'}$ *be decreasing and convex,* $\tilde{r}(x)$, $\frac{x\tilde{r}'(x)}{\tilde{r}(x)}$ *be decreasing, and* $x\tilde{r}(x)$ *be increasing. Also, let* $\lambda_2 \leq \mu_2 \leq \mu_1 \leq \lambda_1$*. Then,* $(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1^*}, \underbrace{\lambda_2, \ldots, \lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1, \ldots, \mu_1}_{n_1^*}, \underbrace{\mu_2, \ldots, \mu_2}_{n_2^*}) \Rightarrow Y_{n^*:n^*}(n_1^*, n_2^*) \leq_{disp} X_{n^*:n^*}(n_1^*, n_2^*)$.

Proof. Ahmed et al. (1986) established that, for two continuous random variables X and Y, if $X \leq_* Y$ holds, then $X \leq_{st} Y \Rightarrow X \leq_{disp} Y$. Therefore, from Corollary 1(ii), we have $Y_{n^*:n^*}(n^*, n^*) \leq_{st} X_{n^*:n^*}(n^*, n^*)$. Hence, considering Theorem 6, we can conclude that the required result holds. □

Remark 5. *It is notable that, in Theorem 7, if we take* $\psi(x) = e^{-x}$, $x > 0$, *then Theorem 7 reduces to (Zhang et al., 2019, Theorem 3.15(*i*)) when* $\alpha_1 = \alpha_2 = 1$.

Remark 6. *Let* $F(x) = (\frac{x}{a})^l$, $0 < x \leq a$*. For this baseline distribution, it can be easily checked that* $\tilde{r}(x)$, $\frac{x\tilde{r}'(x)}{\tilde{r}(x)}$ *are decreasing and* $x\tilde{r}(x)$ *is increasing. Therefore, we can consider the power distribution as the baseline distribution function in Theorems 6 and 7.*

The following theorem provides some conditions for comparing two largest order statistics $X_{n:n}(n_1, n_2)$ and $Y_{n^*:n^*}(n_1^*, n_2^*)$ according to the likelihood ratio order. This can be proved by the similar approach of Theorem 3.12 of Torrado (2017). Here, we consider $\psi(x) = e^{-x}$, $x > 0$, as the generator of the Independence copula. Also, denote $u(x) = x\tilde{r}(x)$, $\eta(x) = -\frac{xf'(x)}{f(x)}$ and $v(x) = \frac{x\tilde{r}'(x)}{\tilde{r}(x)}$.

Theorem 8. *Let* $X_1, \ldots, X_n$ *be* n *nonnegative dependent random variables sharing independence copula with generator* ψ*, such that* $X_i \sim F(x\lambda_1)$*, for* $i = 1, \ldots, n_1^*$ *and* $X_j \sim F(x\lambda_2)$*, for* $j = n_1^* + 1, \ldots, n^*$*. Also, let* $Y_1, \ldots, Y_n$ *be* n *dependent nonnegative random variables sharing Independence copula with generator* ψ*, such that* $Y_i \sim F(x\mu_1)$*, for* $i = 1, \ldots, n_1$ *and* $Y_j \sim F(x\mu_2)$*, for* $j = n_1 + 1, \ldots, n$*. Assume that* $v(x)$ *is decreasing and that* $\eta(x)$*,* $x^2\tilde{r}(x)$*, and* $u(x)v'(x)$ *are increasing. Suppose* $\mu_1 \leq \lambda_1 \leq \lambda_2 \leq \mu_2$ *and* $1 \leq n_1 \leq n_1^* \leq n_2^* \leq n_2$*,* $n = n_1 + n_2$*,* $n^* = n_1^* + n_2^*$*. Then,* $(n_1, n_2) \succeq^w (n_1^*, n_2^*)$ *and*

$$(\underbrace{\lambda_1, \ldots, \lambda_1}_{n_1^*}, \underbrace{\lambda_2, \ldots, \lambda_2}_{n_2^*}) \succeq^w (\underbrace{\mu_1, \ldots, \mu_1}_{n_1^*}, \underbrace{\mu_2, \ldots, \mu_2}_{n_2^*}) \Rightarrow X_{n:n}(n_1, n_2) \leq_{lr} Y_{n^*:n^*}(n_1^*, n_2^*).$$

3.2 Orderings between the smallest order statistics

In the previous subsection, we focus on the conditions, under which the largest order statistics are comparable according to various stochastic orders. Here, we develop conditions such that the usual stochastic, hazard rate, star, and Lorenz orders hold between the smallest order statistics. The reliability functions of $X_{1:n^*}(n_1^*,n_2^*)$ and $Y_{1:n^*}(n_1^*,n_2^*)$ are, respectively, given by

$$\bar{F}_{X_{1:n^*}(n_1^*,n_2^*)}(x) = \psi_1\left[\sum_{i=1}^{n^*}\phi_1\left(\bar{F}_i\left(x\lambda_i\right)\right)\right] = \psi_1\left[n_1^*\phi_1\left(\bar{F}_1\left(x\lambda_1\right)\right)+n_2^*\phi_1\left(\bar{F}_2\left(x\lambda_2\right)\right)\right] \tag{29}$$

and

$$\bar{F}_{Y_{1:n^*}(n_1^*,n_2^*)}(x) = \psi_2\left[\sum_{i=1}^{n^*}\phi_2\left(\bar{F}_i\left(x\mu_i\right)\right)\right] = \psi_2\left[n_1^*\phi_2\left(\bar{F}_1\left(x\mu_1\right)\right)+n_2^*\phi_2\left(\bar{F}_2\left(x\mu_2\right)\right)\right]. \tag{30}$$

Li et al. (2016) compared two series systems with same number of components in the sense of the usual stochastic order when the corresponding random variables follow the dependent scale model with same baseline distribution function.

Here, we consider multiple-outlier models with different baseline distribution functions and prove the usual stochastic order between $X_{1:n}(n_1,n_2)$ and $Y_{1:n^*}(n_1^*,n_2^*)$. This is presented in Theorem 7. To obtain this result, let us first prove the following two theorems. In the following theorems, it is assumed that the samples are heterogeneous and taken from the multiple-outlier dependent scale models with different baseline F_1 and F_2. We now consider the following assumption.

Theorem 9. *Under the set-up as in Assumption 1, with $r_1(x)\le(\ge)r_2(x)$ and $n_1^*\le(\ge)n_2^*$,*

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*})\succeq_w(\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*})\Rightarrow X_{1:n^*}(n_1^*,n_2^*)\le_{st}Y_{1:n^*}(n_1^*,n_2^*),$$

provided $\lambda,\ \mu\in\mathscr{E}_+\ (\mathscr{D}_+)$, $\phi_2\circ\psi_1$ is super-additive, ψ_1 or ψ_2 is log-convex, and $r_1(x)$ or $r_2(x)$ is increasing.

Proof. Let us denote $C(\lambda,\psi_1,x)=\bar{F}_{X_{1:n^*}(n_1^*,n_2^*)}(x)$ and $D(\mu,\psi_2,x)=\bar{F}_{Y_{1:n^*}(n_1^*,n_2^*)}(x)$ (given in (29) and (30)). According to Lemma 4, the super-additivity property of $\phi_2\circ\psi_1$ provides $C(\mu,\psi_1,x)\le D(\mu,\psi_2,x)$. In order to prove the desired result, we need to show that $C(\lambda,\psi_1,x)\le C(\mu,\psi_1,x)$. This is equivalent to show that $C(\lambda,\psi_1,x)$ is decreasing and Schur-concave with respect to λ. Taking derivative of $C(\lambda,\psi_1,x)$ with respect to λ_i, for $i=1,2$, we get

$$\frac{\partial C(\lambda,\psi_1,x)}{\partial\lambda_i} = -n_i^*xr_i(x\lambda_i)\frac{\psi_1\left[\phi_1\left(\bar{F}_i\left(x\lambda_i\right)\right)\right]}{\psi_1'\left[\phi_1\left(\bar{F}_i\left(x\lambda_i\right)\right)\right]}\psi_1'\left[\sum_{i=1}^{n^*}\phi_1\left(\bar{F}_i\left(x\lambda_i\right)\right)\right]. \tag{31}$$

Equation (31) shows that $C(\lambda,\psi_1,x)$ is decreasing in λ_1. The other part can be proved in the similar argument as of Theorem 1. $\square$

The following corollary is immediate from Theorem 9.

Corollary 3. *Let the set-up as in Assumption 1 hold with $\psi_1=\psi_2=\psi$ and $n_1^*\le(\ge)n_2^*$. Also, let ψ be log-convex and let $\lambda,\ \mu\in\mathscr{E}_+\ (\mathscr{D}_+)$. Then,* *(i)* $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow X_{1:n^*}(n_1^*,n_2^*) \leq_{st} Y_{1:n^*}(n_1^*,n_2^*)$, *provided* $r_1(x)$ *or* $r_2(x)$ *is increasing with* $r_1(x) \leq (\geq) r_2(x)$.

(ii) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow X_{1:n^*}(n_1^*,n_2^*) \leq_{st} Y_{1:n^*}(n_1^*,n_2^*)$, *provided* $r(x)$ *is increasing, where* $r_1(x) = r_2(x) = r(x)$.

The next result reveals that the smallest order statistics $X_{1:n^*}(n_1^*,n_2^*)$ dominates $X_{1:n}(n_1,n_2)$ in the sense of the usual stochastic order under the condition that (n_1^*,n_2^*) is weakly submajorized by (n_1,n_2). The following assumption will be helpful to prove the next two results.

Theorem 10. *Let Assumption 2 hold. Then, for* $\lambda = (\lambda_1,\lambda_2) \in \mathscr{E}_+$ *and* $F_1 \leq F_2$, *we have*

$$(n_1,n_2) \succeq_w (n_1^*,n_2^*) \Rightarrow X_{1:n}(n_1,n_2) \leq_{st} X_{1:n^*}(n_1^*,n_2^*).$$

Proof. Using the similar approach of Theorem 2, one can obtain the required result. □

The following remark follows from Theorems 9 and 10.

Remark 7. *Let Assumption 3 hold, with* $r_1(x) \leq r_2(x)$. *Also, let* $\lambda,\ \mu \in \mathscr{E}_+$. *Then, for* $(n_1,n_2) \succeq_w (n_1^*,n_2^*)$,

$$(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow X_{1:n}(n_1,n_2) \leq_{st} Y_{1:n^*}(n_1^*,n_2^*),$$

provided $\phi_2 \circ \psi_1$ *is super-additive,* ψ_1 *or* ψ_2 *is log-convex and* $r_1(x)$ *or* $r_2(x)$ *is increasing.*

As an illustration of Remark 7, we present the following example.

Example 3. Take $\lambda = (2,6)$, $\mu = (1,3)$, $(n_1,n_2) = (4,8)$, $(n_1^*,n_2^*) = (6,7)$, $\psi_1(x) = e^{-x^{\frac{1}{d_1}}}$ and $\psi_2(x) = e^{-x^{\frac{1}{d_2}}}$, $x > 0$, $d_1,\ d_2 \in [1,\infty)$. Also, let $F_1(x) = \left(\frac{x}{a}\right)^l$, $0 < x \leq a$ and $F_2(x) = 1 - e^{-x}$, $x > 0$. It can be seen that, for $d_1 = 9$, $d_2 = 10$ $a = 400$ and $l = 2$, all the conditions of Remark 7 are satisfied (see Table 1). Now, we plot the graph of $\bar{F}_{X_{1:12}(4,8)}(x) - \bar{F}_{Y_{1:13}(6,7)}(x)$, given in Figure 4. The figure suggests that $X_{1:12}(4,8) \leq_{st} Y_{1:13}(6,7)$ holds.

Next, we present a counterexample, which shows that the stated usual stochastic order in Remark 7 does not hold if the conditions $r_1(x) \leq r_2(x)$ and r_2 is increasing, are dropped out.

Counterexample 2. Take $\lambda = (1.2,3.6)$, $\mu = (1.4,3)$, $(n_1,n_2) = (2,11)$, $(n_1^*,n_2^*) = (3,9)$, $\psi_1(x) = e^{-x^{\frac{1}{4.5}}}$, and $\psi_2(x) = e^{-x^{\frac{1}{5}}}$, $x > 0$. Also, suppose $F_1(x) = 1 - e^{-x}$ and $F_2(x) = 1 - (1+2x)^{-0.5}$, $x > 0$. Clearly, all the conditions of Remark 7 are satisfied except $r_1 \leq r_2$ and r_2 is increasing. Now, the graphs of $\bar{F}_{X_{1:13}}(2,11)(x)$ and $\bar{F}_{Y_{1:12}}(3,9)(x)$ are depicted in Figure 5. It reveals that the usual stochastic order in Remark 7 does not hold.

Upon using Remark 7, one can easily conclude the following corollary.

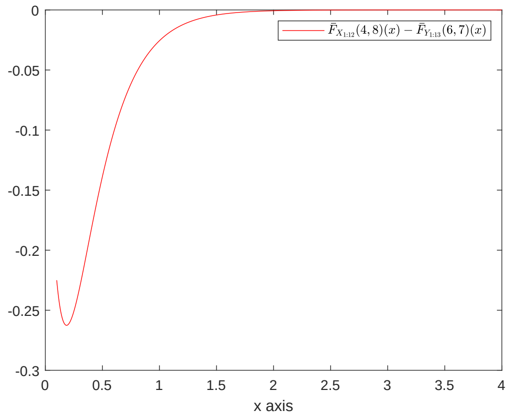


Figure 4: Plot of $\bar{F}_{X_{1:12}(4,8)}(x) - \bar{F}_{Y_{1:13}(6,7)}(x)$ as in Example 3.

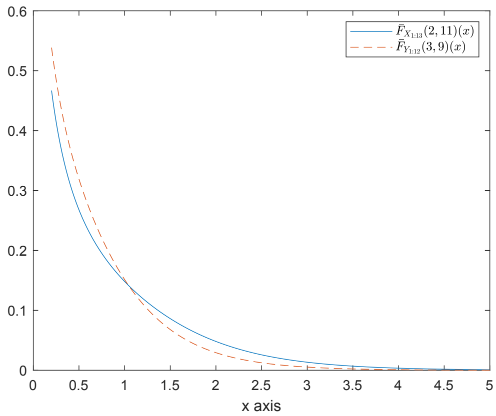


Figure 5: Plots of $\bar{F}_{X_{1:13}(2,11)}(x)$ and $\bar{F}_{Y_{1:12}(3,9)}(x)$ as in Counterexample 2

Corollary 4. *Let Assumption 3 hold with $\psi_1 = \psi_2 = \psi$. Also, let λ, $\mu \in \mathscr{E}_+$, ψ is log-convex and $(n_1,n_2) \succeq_w (n_1^*,n_2^*)$. Then,*

(i) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow X_{1:n}(n_1,n_2) \leq_{st} Y_{1:n^*}(n_1^*,n_2^*)$, *provided* $r_1(x)$ *or* $r_2(x)$ *is increasing and* $r_1(x) \leq r_2(x)$.

(ii) $(\underbrace{\lambda_1,\ldots,\lambda_1}_{n_1^*},\underbrace{\lambda_2,\ldots,\lambda_2}_{n_2^*}) \succeq_w (\underbrace{\mu_1,\ldots,\mu_1}_{n_1^*},\underbrace{\mu_2,\ldots,\mu_2}_{n_2^*}) \Rightarrow X_{1:n}(n_1,n_2) \leq_{st} Y_{1:n^*}(n_1^*,n_2^*)$, *provided* $r(x)$ *is increasing, where* $r_1 = r_2 = r$.

Next, we provide three consecutive theorems, which deal with the hazard rate ordering between the smallest order statistics.

Theorem 11. *Let Assumption 1 hold with $n_1^* \leq (\geq) n_2^*$, $\psi_1 = \psi_2 = \psi$, $F_1 = F_2 = F$, $\tilde{r}_1 = \tilde{r}_2 = \tilde{r}$ and $r_1 = r_2 = r$. Also, suppose that ψ is log-concave, that $\frac{1-\psi}{\psi'}$ is decreasing, that $[\frac{1-\psi}{\psi'}]' \frac{\psi}{\psi'}$ is increasing, and that λ, $\mu \in \mathscr{E}_+$ $(\mathscr{D}_+)$. Then,*

$$(\underbrace{m_1,\ldots,m_1}_{n_1^*},\underbrace{m_2,\ldots,m_2}_{n_2^*}) \succeq_w (\underbrace{v_1,\ldots,v_1}_{n_1^*},\underbrace{v_2,\ldots,v_2}_{n_2^*}) \Rightarrow X_{1:n^*}(n_1^*,n_2^*) \leq_{hr} Y_{1:n^*}(n_1^*,n_2^*),$$

where $m_i = \log\lambda_i$ and $v_i = \log\mu_i$, $i = 1,\ 2$, provided $r(x)$ is increasing, $x\tilde{r}(x)$ is increasing and convex.

Proof. Denote by f the probability density function corresponding to the distribution function F. The hazard rate function of $X_{1:n^*}(n_1^*,n_2^*)$ is given by

$$\begin{aligned} r_{X_{1:n^*}(n_1^*,n_2^*)}(x) &\overset{def}{=} \mathscr{E}(\mathbf{m}) = \frac{\psi'[z]}{\psi[z]}\left[\sum_{i=1}^{n^*}\frac{e^{m_1} f(xe^{m_1})}{\psi'[\phi(\bar{F}(xe^{m_1}))]}\right] \\ &= \frac{\psi'[z]}{\psi[z]}\left[\sum_{i=1}^{n^*}\frac{e^{m_i}\tilde{r}(xe^{m_i})[1-\psi[\phi(\bar{F}(xe^{m_i}))]]}{\psi'[\phi(\bar{F}(xe^{m_i}))]}\right], \end{aligned}$$

where $z = n_1^*\phi(\bar{F}(xe^{m_1})) + n_2^*\phi(\bar{F}(xe^{m_2}))$, $m_i = \log\lambda_i$, for $i = 1,2$ and $\mathbf{m} = (m_1,m_2)$. Also, f is the probability density function of F. The partial derivative of $\mathscr{E}(\mathbf{m})$ with respect to m_i for $i = 1,2$ is given by

$$\begin{aligned} \frac{\partial\mathscr{E}(\mathbf{m})}{\partial m_i} &= -n_i^* xe^{m_i}\tilde{r}(xe^{m_i})\frac{d}{dz}\left[\frac{\psi'(z)}{\psi(z)}\right]\left[\frac{1-\psi[\phi[\bar{F}(xe^{m_i})]]}{\psi'[\phi(\bar{F}(xe^{m_i}))]}\right]\left[\sum_{i=1}^{n^*}\frac{e^{m_i} f(xe^{m_i})}{\psi'[\phi(\bar{F}(xe^{m_i}))]}\right] \\ &\quad -n_i^* r(xe^{m_i})\left[x[e^{m_i}]^2\tilde{r}(xe^{m_i})\right]\frac{\psi'(z)}{\psi(z)}\left[\frac{\psi(v)}{\psi'(v)}\left[\frac{d}{dv}\left[\frac{1-\psi(v)}{\psi'(v)}\right]\right]\right]_{v=\phi(\bar{F}(xe^{m_i}))} \\ &\quad +n_i^*\frac{d}{dw}[w\tilde{r}(w)]_{w=xe^{m_i}}\frac{1-\psi[\phi[\bar{F}(xe^{m_i})]]}{\psi'[\phi[\bar{F}(xe^{m_i})]]}\frac{\psi'(z)}{\psi(z)}. \end{aligned} \tag{32}$$

From (32), it is easy to see that $\mathscr{E}(\mathbf{m})$ is increasing with respect to $\mathbf{m}$. Now, we only need to show the Schur-convexity of $\mathscr{E}(\mathbf{m})$ with respect to $\mathbf{m}$. This is equivalent to show that for $1 \leq i \leq j \leq n^*$,

$$\left[\frac{\partial\mathscr{E}(\mathbf{m})}{\partial m_i} - \frac{\partial\mathscr{E}(\mathbf{m})}{\partial m_j}\right] \leq (\geq) 0, \text{ for } \mathbf{m} \in \mathscr{E}_+\ (\mathscr{D}_+). \tag{33}$$

Utilizing the assumptions made, the rest of the proof follows from the similar arguments of Theorem 9. Thus, it is omitted for the sake of conciseness.

□

The following theorem demonstrates that under some conditions, the hazard rate ordering between $X_{1:n}(n_1,n_2)$ and $X_{1:n^*}(n_1^*,n_2^*)$ exists.

Theorem 12. *Let Assumption 2 hold with $\psi_1=\psi$ and $\tilde{r}_1=\tilde{r}_2=\tilde{r}$. Then, for $\lambda\in\mathscr{E}_+$, we have*

$$(n_1,n_2)\succeq_w (n_1^*,n_2^*)\Rightarrow X_{1:n}(n_1,n_2)\leq_{hr} X_{1:n^*}(n_1^*,n_2^*),$$

provided $x\tilde{r}(x)$ is increasing, ψ'/ψ and $\frac{1-\psi}{\psi'}$ are decreasing.

Proof. The required result can be proved if we show that $r_{X_{1:n}(n_1,n_2)}(x)\geq r_{X_{1:n^*}(n_1^*,n_2^*)}(x)$ and equivalently,

$$\frac{\psi'[\sum_{i=1}^{n}\phi(\bar{F}(x\lambda_i))]}{\psi[\sum_{i=1}^{n}\phi(\bar{F}(x\lambda_i))]}\left[\sum_{i=1}^{n}\frac{\lambda_i\tilde{r}(x\lambda_i)[1-\psi[\phi(\bar{F}(x\lambda_i))]]}{\psi'[\phi(\bar{F}(x\lambda_i))]}\right]\geq\frac{\psi'\left[\sum_{i=1}^{n^*}\phi(\bar{F}(x\lambda_i))\right]}{\psi\left[\sum_{i=1}^{n^*}\phi(\bar{F}(x\lambda_i))\right]}\left[\sum_{i=1}^{n^*}\frac{\lambda_i\tilde{r}(x\lambda_i)[1-\psi[\phi(\bar{F}(x\lambda_i))]]}{\psi'[\phi(\bar{F}(x\lambda_i))]}\right].\tag{34}$$

To prove inequality (34), it is sufficient to show that the following two inequalities hold:

$$(n_1^*-n_1)\phi(\bar{F}(x\lambda_1))\leq(n_2-n_2^*)\phi(\bar{F}(x\lambda_2))\tag{35}$$

and

$$(n_1^*-n_1)\frac{\lambda_1\tilde{r}(x\lambda_1)[1-\psi[\phi(\bar{F}(x\lambda_1))]]}{\psi'[\phi(\bar{F}(x\lambda_1))]}\geq(n_2-n_2^*)\frac{x\lambda_2\tilde{r}(x\lambda_2)[1-\psi[\phi(\bar{F}(x\lambda_2))]]}{\psi'[\phi(\bar{F}(x\lambda_2))]}.\tag{36}$$

Furthermore, $(n_1,n_2)\succeq_w(n_1^*,n_2^*)\Rightarrow(n_1+n_2)\geq(n_1^*+n_2^*)\Rightarrow(n_2-n_2^*)\geq(n_1^*-n_1)\geq 0$. Also, $\lambda_1\leq\lambda_2\Rightarrow\phi(\bar{F}(x\lambda_2))\geq\phi(\bar{F}(x\lambda_1))\geq 0$. With the help of decreasing property of $\frac{1-\psi}{\psi'}$, we obtain

$$\frac{1-\psi(w)}{\psi'(w)}|_{w=\phi[\bar{F}(x\lambda_2)]}\leq\frac{1-\psi(w)}{\psi'(w)}|_{w=\phi[\bar{F}(x\lambda_1)]}\leq 0.\tag{37}$$

Since $x\tilde{r}(x)$ is increasing,

$$x\lambda_1\tilde{r}(x\lambda_1)\leq x\lambda_2\tilde{r}(x\lambda_2).\tag{38}$$

Thus, the proof is completed from(37), (38), and the given assumptions. □

The next theorem states that if the scale parameters are connected with the weakly submajorized order and the sample size pairs (n_1,n_2) and (n_1^*,n_2^*) have weakly submajorized order, then the smallest order statistics of $X_{1:n}(n_1,n_2)$ is dominated by $Y_{1:n^*}(n_1^*,n_2^*)$ according to the hazard rate order.

Theorem 13. *Let Assumption 3 hold with $\psi_1=\psi_2=\psi$, $r_1=r_2=r$ and $\tilde{r}_1=\tilde{r}_2=\tilde{r}$. Then, for λ, $\mu\in\mathscr{E}_+$ and $(n_1,n_2)\succeq_w(n_1^*,n_2^*)$,*

$$(\underbrace{m_1,\dots,m_1}_{n_1^*},\underbrace{m_2,\dots,m_2}_{n_2^*})\succeq_w(\underbrace{v_1,\dots,v_1}_{n_1^*},\underbrace{v_2,\dots,v_2}_{n_2^*})\Rightarrow X_{1:n}(n_1,n_2)\leq_{hr}Y_{1:n^*}(n_1^*,n_2^*),$$

provided ψ is log-concave, $\frac{1-\psi}{\psi'}$ is decreasing, $[\frac{1-\psi}{\psi'}]'\frac{\psi}{\psi'}$ and $r(x)$ are increasing, $x\tilde{r}(x)$ is increasing and convex, where $m_i=\log\lambda_i$ and $v_i=\log\mu_i$, $i=1,\ 2$.

Proof. The proof of the theorem follows from Theorems 11 and 12. Thus, it is omitted. □

To illustrate Theorem 13, we now consider the following example.

Example 4. Set $\lambda = (e^{0.5}, e^{0.6})$, $\mu = (e^{0.2}, e^{0.3})$, $(n_1, n_2) = (2, 11)$, $(n_1^*, n_2^*) = (3, 7)$, and $\psi(x) = e^{\frac{1}{b}(1-e^x)}$, $x > 0$, $b \in (0, 1]$. Furthermore, let $F(x) = \left(\frac{x}{a}\right)^l$, $0 < x \leq a$. It can be easily shown that for $b = 0.99$, $a = 1000$, and $l = 2$, all the conditions of Theorem 13 are satisfied (see Table 1). Now, we plot the ratio $\frac{\bar{F}_{Y_{1:10}}(3,7)(x)}{\bar{F}_{X_{1:13}}(2,11)(x)}$ in Figure 6, which is consistent with the result in Theorem 13.

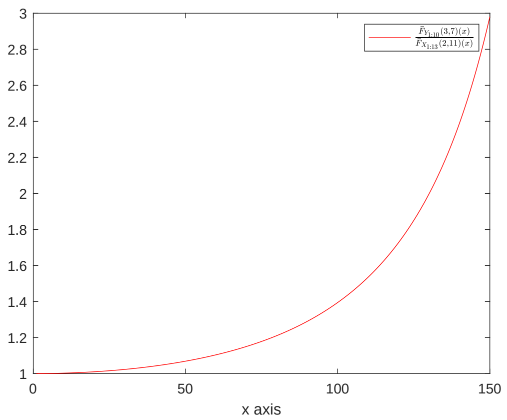


Figure 6: Plot of $\bar{F}_{Y_{1:10}(3,7)}(x)/\bar{F}_{X_{1:13}(2,11)}(x)$ as in Example 4.

Remark 8. *Consider the Independence copula with generator $\psi(x) = e^{-x}$, $x > 0$, which satisfies all the conditions of Theorems 11–13 (see Table 1).*

In the next theorem, we develop some conditions under which two smallest order statistics are comparable according to the star order.

Theorem 14. *Under the set-up as in Assumption 1, with $\tilde{r}_1 = \tilde{r}_2 = \tilde{r}$ and $\psi_1 = \psi_2 = \psi$,*

$$\frac{\lambda_{2:2}}{\lambda_{1:2}} \geq \frac{\mu_{2:2}}{\mu_{1:2}} \Rightarrow Y_{1:n^*}(n_1^*, n_2^*) \leq_* X_{1:n^*}(n_1^*, n_2^*),$$

provided $\frac{\psi}{\psi'}$ is decreasing and convex and $\frac{xr'(x)}{r(x)}$ and $xr(x)$ are decreasing.

Proof. The distribution functions of $X_{1:n^*}(n_1^*, n_2^*)$ and $Y_{1:n^*}(n_1^*, n_2^*)$ are, respectively, given by

$$F_{X_{1:n^*}}(n_1^*, n_2^*)(x) = 1 - \psi\left[n_1^* \phi\left(\bar{F}(x\lambda_1)\right) + n_2^* \phi\left(\bar{F}(x\lambda_2)\right)\right]$$

and

$$F_{Y_{1:n^*}}(n_1^*, n_2^*)(x) = 1 - \psi\left[n_1^* \phi\left(\bar{F}(x\mu_1)\right) + n_2^* \phi\left(\bar{F}(x\mu_2)\right)\right].$$

Now, the rest of the proof follows using similar arguments as in Theorem 6. Thus, it is omitted. □

The following result is a direct consequence of Theorem 14.

Corollary 5. *Under the assumptions as in Theorem 14,*

$$\frac{\lambda_{2:2}}{\lambda_{1:1}} \geq \frac{\mu_{2:2}}{\mu_{1:2}} \Rightarrow Y_{1:n^*}(n_1^*, n_2^*) \leq_{Lorenz} X_{1:n^*}(n_1^*, n_2^*),$$

provided $\frac{\psi}{\psi'}$ *is decreasing and convex and* $\frac{xr'(x)}{r(x)}$ *and* $xr(x)$ *are decreasing.*

Remark 9. *Suppose* $F(x) = 1 - (\frac{x}{b})^a$, $0 < x \leq b$, $a > 0$. *For this baseline distribution, it is easy to check that* $\frac{xr'(x)}{r(x)}$ *and* $xr(x)$ *are decreasing. Therefore, one can consider this distribution function as the baseline distribution of Theorem 14.*

4 Concluding remarks

In this paper, we discussed some comparison results between the lifetimes of both parallel and series systems consisting of multiple-outlier dependent scale components in the sense of the usual stochastic, reversed hazard rate, dispersive order, hazard rate, likelihood ratio, star, and Lorenz orders. The dependence structure has been modeled by Archimedean copulas. Sufficient conditions have been established for the purpose of the comparisons of the order statistics. Several examples and counterexamples are presented to illustrate the established results.

Acknowledgments

Both the authors thank the Editor and two anonymous referess for their comments, which have led to this improved version. Sangita Das wishes to thank the Ministry of Education (formerly known as MHRD), Government of India for the financial support to carry out this research work. Also, Sangita Das gratefully acknowledges the financial support for this research work under the NPDF, grant No: PDF/2022/00471, SERB, Government of India. Suchandan Kayal gratefully acknowledges the partial financial support for this research work under a grant MTR/2018/000350, SERB, India.

Disclosure statement

Both the authors states that there is no conflict of interest.

References

Ahmed, A. N., Alzaid, A., Bartoszewicz, J. and Kochar, S. C. (1986). Dispersive and superadditive ordering. *Advances in Applied Probability*, **18**, 1019–1022.

Amini-Seresht, E., Qiao, J., Zhang, Y. and Zhao, P. (2016). On the skewness of order statistics in multiple-outlier PHR models. *Metrika*, **79**, 817–836.

Balakrishnan, N. and Torrado, N. (2016). Comparisons between largest order statistics from multiple-outlier models. *Statistics*, **50**, 176–189.

Balakrishnan, N. and Zhao, P. (2013). Ordering properties of order statistics from heterogeneous populations: a review with an emphasis on some recent developments. *Probability in the Engineering and Informational Sciences*, **27**, 403–443.

Fang, L., Barmalzan, G. and Ling, J. (2016). Dispersive order of lifetimes of series systems in multiple-outlier Weibull models. *Journal of Systems Science & Complexity*, **29**, 1693–1702.

Kochar, S. C. and Torrado, N. (2015). On stochastic comparisons of largest order statistics in the scale model. *Communications in Statistics-Theory and Methods*, **44**, 4132–4143.

Kochar, S. and Xu, M. (2011). On the skewness of order statistics in multiple-outlier models. *Journal of Applied Probability*, **48**, 271–284.

Kundu, A., Chowdhury, S., Nanda, A. K. and Hazra, N. K. (2016). Some results on majorization and their applications. *Journal of Computational and Applied Mathematics*, **301**, 161–177.

Li, C., Fang, R. and Li, X. (2016). Stochastic comparisons of order statistics from scaled and interdependent random variables. *Metrika*, **79**, 553–578.

Li, X. and Fang, R. (2015). Ordering properties of order statistics from random variables of Archimedean copulas with applications. *Journal of Multivariate Analysis*, **133**, 304–320.

Marshall, A. W. and Olkin, I. (2007). *Life Distributions.* Springer, New York.

Marshall, A.W., Olkin, I. and Arnold, B.C. (2011). *Inequalities: Theory of Majorization and its Applications.* Second edn, Springer, New York.

McNeil, A. J. and Nešlehová, J. (2009). Multivariate archimedean copulas, d-monotone functions and l^1-norm symmetric distributions. *The Annals of Statistics*, **37**, 3059–3097.

Navarro, J., Torrado, N. and Águila, Y.d. (2018). Comparisons between largest order statistics from multiple-outlier models with dependence. *Methodology and Computing in Applied Probability*, **20**, 411–433.

Nelsen, R. B. (2006) *An Introduction to Copulas.* Second edn. Springer, New York.

Saunders, I. W. and Moran, P. A. P. (1978). On the quantiles of the gamma and F distributions. *Journal of Applied Probability*, **15**, 426–432.

Shaked, M. and Shanthikumar, J.G. (2007) *Stochastic Orders.* Springer, New York.

Torrado, N. (2017). Stochastic comparisons between extreme order statistics from scale models. *Statistics*, **51**, 1359–1376.

Wang, J. and Cheng, B. (2017). Answer to an open problem on mean residual life ordering of parallel systems under multiple-outlier exponential models. *Statistics & Probability Letters*, **130**, 80–84.

Zhang, Y., Cai, X., Zhao, P. and Wang, H. (2019). Stochastic comparisons of parallel and series systems with heterogeneous resilience-scaled components. *Statistics*, **53**, 126–147.

Zhao, P. and Balakrishnan, N. (2012). Stochastic comparisons of largest order statistics from multiple-outlier exponential models. *Probability in the Engineering and Informational Sciences*, **26**, 159–182.

Zhao, P. and Balakrishnan, N. (2015). Comparisons of largest order statistics from multiple-outlier gamma models. *Methodology and Computing in Applied Probability*, **17**, 617–645.

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 75–104. Research Paper

SMPS

On sequential growth of trees subject to various labeling constraints: from enumeration to probability theory

Thierry E. Huillet*

Laboratoire de Physique Théorique et Modélisation (CNRS, UMR 8089)
CY Cergy Paris Université, 95302 Cergy-Pontoise, France
Email(s): thierry.huillet@cyu.fr

Abstract. Trees, as loop-free graphs, are fundamental hierarchical structures of Nature. Depending on the way their constitutive atoms are labeled, their growth obeys different sequential dynamics when a new atom is being appended to a current tree, possibly forming a new tree. Randomized versions of the underlying counting problems are shown to lead, in general, to Markovian triangular sequences.

Keywords: Combinatorial probability; Distinguishability; Forests; Markovian triangular sequences; Randomization; Simply generated and increasing trees.

1 Introduction

The purpose of this note is to present explicit and asymptotic methods to count various kinds of topological trees (for which there is no data giving the distance between any two of their nodes). In all cases, the use of generating functions (g.f.'s) is an essential ingredient. Explicit formulas are derived with the help of Lagrange's inversion formula. On the other hand, singularity analysis of g.f.'s leads to asymptotic formulas aiming at describing large trees.

The analysis concerns counting of the following labeled tree structures (with no restriction on the degree of their nodes):

- Simply generated labeled trees, either ordered (plane) or not (Cayley trees).
- Recursive increasing labeled trees, either ordered (plane) or not.
- The height, depth, and number of leaves of such trees.
- A weighted version of such trees.

*Corresponding author

Received: 24 December 2023 / Accepted: 22 April 2024
DOI: 10.22067/smps.2024.45214

 https://smps.um.ac.ir

- Forests of such trees are also considered.

The various labeling conditions translate various distinguishability possibilities of the nodes (somewhere called here, alternatively vertices or atoms) of the tree (or forest). We will herewith be also concerned by the stochastic randomization of the underlying strictly combinatorial aspects of the counting issues. Relation of enumeration problems of trees and forests to random walks can be found in Pitman (1998).

Focus will be on sequential growth of trees and forests as given, after randomization, by Markov triangular probability sequences. By that, we mean the update of their shapes when the number of its atoms goes from n to $n+1$, in an aggregation process of individuals. Whenever the additional atom connects to one of the internal nodes of the size-n tree, the number of external nodes (leaves) grows by one unit, whereas if the additional atom connects to one of its leaves, the number of its leaves remains constant. Leaves of a tree consist of its boundary, connecting it to the external world. When dealing with forests, the additional atom can either connect to one of the trees of the current forest (aggregation) or generate a new tree (nucleation). Trees are efficient graph structures to fully connect people as the number of edges required to do so is minimum but these are fragile in that the deletion of an edge disconnects the subtree below it from the main tree.

If, by recursion from the root, $\Phi(z)$ solves the functional equation

$$\Phi(z) = ze^{\sum_{k\geq 1}\frac{1}{k}\Phi\left(z^k\right)}, \tag{1}$$

then

$$c_n = [z^n]\,\Phi(z),$$

(the z^n-coefficient in the power series expansion of $\Phi(z)$) is the number of rooted unlabeled size-n trees (see (Wolfram.com) and the references therein). The number of unrooted such trees is given by $c_n^* = [z^n]\,\Phi^*(z) = c_n/n$, where

$$\Phi^*(z) = \Phi(z) - \frac{1}{2}\left[\Phi^2(z) - \Phi\left(z^2\right)\right].$$

The enumeration of such trees with n up to 6 is given in (Wolfram.com). There are no known expressions for c_n (or c_n^*) but, from local singularity analysis, $c_n \sim \beta n^{-3/2} z_c^{-n}$ (respectively, $c_n^* \sim \beta n^{-5/2} z_c^{-n}$) for large n, where $\beta = 0.5349\ldots$ and $1/z_c = 2.955765\ldots$ is defined from $\Phi(z_c) = 1$. Note that the functional equation (1) is also,

$$\Phi(z) = z\prod_{n\geq 1}(1-z^n)^{-c_n},$$

with $\prod_{n\geq 1}(1-z^n)^{-c_n} = e^{\sum_{k\geq 1}\frac{1}{k}\Phi\left(z^k\right)}$ the unordered "exponential" of $\Phi(z)$. Moreover,

$$\Phi(z,u) = ze^{\sum_{k\geq 1}k^{-1}\left[(u-1)z^k + \frac{1}{k}\Phi\left(z^k,u\right)\right]} = z(1-z)^{u-1}\,e^{\sum_{k\geq 1}\frac{1}{k}\Phi\left(z^k,u\right)}$$

is the joint g.f. for the nodes (vertices) and the dangling nodes (culs de sac or leaves) of such trees, with

$$c_{n,k} := \left[z^n u^k\right]\Phi(z,u)$$

the number of unlabeled rooted trees with size-n having k dangling nodes, $k = 2, \ldots, n-1$. From Wolfram, $c^*_{6,1} = 1$, $c^*_{6,2} = 1$, $c^*_{6,3} = 2$, $c^*_{6,4} = 2$, and $c^*_{6,5} = 1$. For each such tree, the question of how many ways they can be labeled to yield simple or recursive unrooted trees arises.

Let B_l be a 0-diagonal $n \times n$ matrix with entries in $\{0,1\}$ whose the first row is $(0;1,\ldots,1;0,\ldots,0)$ (with $n_1 \geq 1$ ones), the second row is $(0;0,\ldots,0;1,\ldots,1;0,\ldots,0)$ (with $n_2 \geq 1$ following 1 's not overlapping the n_1 first ones), ..., whose the lth row is $(0;0,\ldots,0;0,\ldots,0;\ldots;1,\ldots,1)$ (with $n_l \geq 1$ final 1's not overlapping the $n_1 + \ldots + n_{l-1}$ previous ones). The remaining rows are zero-sum rows, and $n_1 + \ldots + n_l = n-1$ is the overall number of edges in the size-n unrooted tree; the nonoverlapping condition in the formation of B_l guarantees that there are no cycles in the tree graph.

If $n_1 \geq n_2 \geq \ldots \geq n_l \geq 1$, then there are $p_{n-1,l}$ such matrices B_l, where $p_{n,l}$ is the number of unordered partitions of n into l positive summands, with $p_{n,l}$ obeying the three-term recursion

$$p_{n,l} = p_{n-1,l-1} + p_{n-l,l},$$

with boundary conditions $p_{n,l} = 0$ if $l \leq 0$ or $l > n$. Upon transposition of B_l, the symmetrized matrix

$$A_l = B_l + B_l'$$

is thus, up to permutation of its rows and columns, the incidence matrix of unrooted unlabeled trees with $k = n-l$ dangling nodes. Both the row and column sums vectors of A_l are the same. We conclude that $c^*_{n,k}$, $k = 2, \ldots, n-1$, is the number of matrices obtained as PA_lP', where P runs over all permutation $n \times n$ matrices (the equivalence class of A_l).

2 Number of atoms and leaves in a size-n simple tree

We shall distinguish two main types of simply generated (or simple) trees, namely, as follows:

- Ordered (or plane) trees: The reason is that one can draw the tree in the half-plane so that the children of every parent are ordered from left to right, say from the youngest child to the oldest one. Embeddings obtained from cyclic rotations of the sub-trees around the root are not allowed.

Such trees are amenable to the Ulam–Harris–Neveu ordering of their nodes (horizontal ordering holds), and they can be represented as strata with the founder on top and the successive layers below; see Neveu (1986). Given that an individual of the population at generation k has been labeled by vertex $\mathbf{v} = v_1 \ldots v_k$ (as a concatenation of k positive integers) and gives birth to $K_{\mathbf{v}} \geq 1$ children, its offspring is labeled by $\mathbf{v}1, \ldots, \mathbf{v}K_{\mathbf{v}}$. Each individual at generation k thus gets a concatenated label $\mathbf{v} = v_1 \ldots v_k$ for which label $v_1 \ldots v_{k-1}$ is one of its mothers, $v_1 \ldots v_{k-2}$, is one of its grandmothers, ..., up to $\emptyset$, is the conventional label of the root. Such ordered trees are the combinatorial versions of Galton–Watson trees in the theory of branching processes.

- Unordered (nonplane) trees: If the above condition is not imposed, then such trees are called Cayley trees.

2.1 Nonincreasing (simply generated) trees

By recursion from the root, the number of labeled simply generated (or simple) trees of size n generated by the local g.f. $\phi(z)$ (with nonnegative $\phi_m := [z^m]\,\phi(z)$, $m \geq 0$, $\phi(0) = 1$) is obtained as

$$C_n = n!\,[z^n]\,\Phi(z), \tag{2}$$

where $\Phi(z)$ solves the functional equation $\Phi(z) = z\phi(\Phi(z))$, $\Phi(0) = 0$. Lagrange inversion formula states that for all $n \geq 1$

$$[z^n]\,\Phi(z) = \frac{1}{n}\left[z^{n-1}\right]\phi(z)^n. \tag{3}$$

In enumeration problems from combinatorics, $k!\phi_k$ are assumed to be nonnegative integers.

A more general form of Lagrange inversion formula states that (with $'$ denoting derivative)

$$[z^n]\,h(\Phi(z)) = \frac{1}{n}\left[z^{n-1}\right]\left(h'(z)\,\phi(z)^n\right), \tag{4}$$

for any arbitrary analytic output function h; see Surya and Warnke (2023).

2.1.1 Cayley (unordered) trees Rényi (1959)

For rooted and labeled Cayley trees, $\phi(z) = e^z$. By Lagrange inversion formula,

$$\Phi(z) = \sum_{n\geq 1} \frac{n^{n-1}}{n!} z^n$$

solves $\Phi(z) = ze^{\Phi(z)}$ (with $\phi(z) = e^z$ and $C_n = n^{n-1}$).

Note $z_c = \sup(z > 0 : \Phi(z) < \infty) = 1/e$ with $\Phi(z_c) = 1$, $\Phi'(z_c) = \infty$. The probability to observe a particular size-n tree τ_n among all such size-n trees is $1/c_n$. The tilted probability to observe a tree of size n among all possible trees is

$$\frac{z^n C_n/n!}{\Phi(z)},$$

for those $z \in (0, z_c]$, where $z_c = 1/e$. Tilting is necessary because the overall number of Cayley trees, regardless of their sizes, is infinite.

The joint g.f. of their nodes and leaves reads:

$$\Phi(z,u) = z(u - 1 + e^{\Phi(z,u)}),$$

hence with

$$\begin{aligned}
[z^n]\,\Phi(z,u) &= \frac{1}{n}\left[z^{n-1}\right](u-1+e^z)^n \\
&= \frac{1}{n}\left[z^{n-1}\right]\sum_{k=0}^{n}\binom{n}{k}u^k(e^z-1)^{n-k}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
C_{n,k} &:= n!\left[z^n u^k\right]\Phi(z,u) = (n-1)!\binom{n}{k}\left[z^{n-1}\right](e^z-1)^{n-k} \\
&= \frac{n!}{k!} S_{n-1,n-k}, \qquad k = 1, \ldots, n-1,
\end{aligned} \tag{5}$$

due to the vertical "g.f. of Stirling numbers of the second kind" $S_{n,k}$:

$$\sum_{n\geq k} S_{n,k}\frac{z^n}{n!} = \frac{(e^z-1)^k}{k!} \Rightarrow [z^n](e^z-1)^k = \frac{k!}{n!}S_{n,k}.$$

We get

$$kC_{n+1,k} = k(n+1)C_{n,k} + (n+1)(n-k+1)C_{n,k-1}, \tag{6}$$

with boundary conditions $C_{n,0} = C_{n,n} = 0$, $C_{n,1} = n!$ and $C_{n,n-1} = n$. In addition,

$$C_n := [z^n]\Phi(z,1) = n^{n-1} = \sum_{k=1}^{n-1} C_{n,k}.$$

Assuming uniform sampling, we have $\mathbf{P}(L_n = k) = \frac{C_{n,k}}{C_n} > 0$, $k = 1,\ldots,n-1$ (otherwise 0): The law of L_n has finite support, varying with n. From (6), with $\mathbf{P}(L_n = k) = \frac{C_{n,k}}{C_n}$, we get

$$k\mathbf{P}(L_{n+1} = k) = \left(\frac{n}{n+1}\right)^{n-1}[k\mathbf{P}(L_n = k) + (n-(k-1))\mathbf{P}(L_n = k-1)]. \tag{7}$$

The latter recursion (6) may be written as

$$\mathbf{P}(L_{n+1} = k) = q_{k,k}^{(n)}\mathbf{P}(L_n = k) + q_{k-1,k}^{(n)}\mathbf{P}(L_n = k-1),$$

defining the (positive) transition coefficients $q_{k,k}^{(n)}$ and $q_{k-1,k}^{(n)}$, which are not transition probabilities. This three-term ("space-time" inhomogeneous) recurrence is therefore not one of standard Markov chains with a usual probability transition matrix. However, it is one of triangular Markov probability sequences whose support varies with n linearly. Note that the ratio $\rho_{k,n} := q_{k-1,k}^{(n)}/q_{k,k}^{(n)} = (n-(k-1))/k$ obeys

$$\begin{aligned} q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &< 1 \text{ if } k > (n+1)/2, \\ q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &> 1 \text{ if } k < (n+1)/2, \end{aligned}$$

translating that L_n is attracted to the central part of its support, whose two endpoints $\{1, n-1\}$ are (asymmetrically) weakly repelling. Introduce the superdiagonal transition matrix

$$Q_{n,n} := \begin{pmatrix} q_{1,1}^{(n)} & q_{1,2}^{(n)} & 0 & \cdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & q_{n-1,n-1}^{(n)} & q_{n-1,n}^{(n)} \\ \vdots & \vdots & 0 & q_{n,n}^{(n)} \end{pmatrix},$$

and let $Q_{n-1,n}$ be its $(n-1)\times n$ truncated version. With $\pi_2^L := 1$, taking indeed into account the boundary conditions, the distributions $\pi_n^L := (\mathbf{P}(L_n = k), k \in \{1,\ldots,n-1\})$, $n \geq 2$, satisfy the recursion

$$\pi_{n+1}^L = (\pi_n^L, 0)\, Q_{n,n} = \pi_n^L Q_{n-1,n}, \qquad n \geq 2.$$

Thus, $\pi_n^L = \overleftarrow{\prod}_{m=2,\ldots,n-1} Q_{m-1,m}$, $n \geq 3$ is an integrated form solution of the recursion, as a left product of nested rectangular matrices (starting from $Q_{n-2,n-1}$ to the right, ending up with $Q_{1,2}$ to the left). Because for each $n \geq 1$, π_n^L is a probability vector, we get that π_n^L is orthogonal to the 1-shifted column sum vector $\mathbf{q}_{n-1} - \mathbf{1}$ of $Q_{n-1,n}$ with kth entry $q_{k,k}^{(n)} + q_{k,k+1}^{(n)} - 1$, $k = 1, \ldots, n-1$. This can easily be seen while post-multiplying π_{n+1}^L by the column vector $\mathbf{1}$ and observing $\pi_{n+1}^L \mathbf{1} = \pi_n^L \mathbf{1} = 1$.

Next, the identity (Rényi 1959)

$$\sum_{k=1}^{n-1} C_{n,k} \frac{\binom{x}{n-k}}{\binom{n}{n-k}} = x^{n-1}$$

yields (with $x = n-1$): $\sum_{k=1}^{n-1} k C_{n,k} = n(n-1)^{n-1}$. Also,

$$\partial_u \Phi(z,1) = \frac{z}{1 - \Phi(z,1)},$$

with

$$\begin{aligned} [z^n] \partial_u \Phi(z,1) &= \left[z^{n-1}\right] \frac{1}{1-\Phi(z,1)} = \frac{1}{n-1} \left[z^{n-2}\right] \left(\left(\frac{1}{1-z} \right)' e^z \right) \\ &= n(n-1)^{n-1}, \end{aligned}$$

so that

$$\begin{aligned} \mathbf{E}(L_n) &= \frac{[z^n] \partial_u \Phi(z,1)}{[z^n] \Phi(z,1)} = n \left(1 - \frac{1}{n}\right)^{n-1} \sim n/e, \\ \sigma^2(L_n) &= n(n-1)\left(1 + \frac{2}{n}\right)^{n-1} + n\left(1 - \frac{1}{n}\right)^{n-1} - n^2 \left(1 - \frac{1}{n}\right)^{2(n-1)} \\ &\sim n \frac{e-2}{e^2}. \end{aligned}$$

The variance term is obtained while plugging $x = n-2$ in the identity. The Central Limit Theorem (CLT) therefore holds:

$$\frac{L_n - \mathbf{E}(L_n)}{\sigma(L_n)} \xrightarrow{d} \mathcal{N}(0,1). \tag{8}$$

Rényi (1959) and Steele (1987) rather studies unrooted Cayley trees (there are n^{n-2} such trees) with no major difference with the rooted case here. Trees, in Rényi's writing, consist of the optimal graphs to connect n cities with just one path joining any pair of cities, passing through their most recent common ancestor.

2.1.2 Ordered (plane) trees

For rooted ordered trees, $\phi(z) = 1/(1-z)$. The g.f.

$$\Phi(z) = \frac{1-\sqrt{1-4z}}{2}$$

solves $\Phi(z) = z/(1-\Phi(z))$ (with $\phi(z) = 1/(1-z)$).

Note $z_c = \sup(z > 0 : \Phi(z) < \infty) = 1/4$ with $\Phi(z_c) = 1/2$ and $\Phi'(z_c) = \infty$. $\Phi(z)$ has an algebraic singularity at $z_c = 1/2$.

Here, $C_n = n!\,[z^n]\,\Phi(z) = \frac{(2n-2)!}{(n-1)!}$ and $c_n = [z^n]\,\Phi(z) \sim \frac{1}{\sqrt{\pi}} n^{-3/2} 4^{n-1}$ (by singularity analysis of $\Phi(z)$). Note also $C_{n+1} = 2(2n-1)\,C_n$.

Here, C_n counts the number of labeled size-n trees respecting the Ulam-Harris-Neveu ordering of the children, while c_n counts the number of such size-n trees, dropping the global labeling condition. There are thus $c_n = \frac{(2n-2)!}{n!(n-1)!}$ rooted ordered trees with n atoms.

Let us now come to leaves of such trees:

With $\Phi(z) = \Phi(z,1)$, the joint g.f. of nodes and leaves solves

$$\Phi(z,u) = z\,(u-1+1/(1-\Phi(z,u)))\,.$$

Hence,

$$\Phi(z,u) = \frac{1+(u-1)\,z-\sqrt{(1+(u-1)\,z)^2-4zu}}{2},$$

with

$$\begin{aligned}
[z^n]\,\Phi(z,u) &= \frac{1}{n}\left[z^{n-1}\right](u+z/(1-z))^n \\
&= \frac{1}{n}\left[z^{n-1}\right]\sum_{k=0}^{n}\binom{n}{k}u^k\,(z/(1-z))^{n-k} \\
&= \frac{1}{n}\sum_{k=0}^{n}\binom{n}{k}u^k\left[z^{k-1}\right](1-z)^{-(n-k)} \\
&= \frac{1}{n}\sum_{k=0}^{n}\binom{n}{k}\frac{[n-k]_{k-1}}{(k-1)!}u^k \\
&= \frac{1}{n}\sum_{k=0}^{n}\binom{n}{k}\binom{n-2}{k-1}u^k.
\end{aligned}$$

Therefore,

$$C_{n,k} := n!\left[z^n u^k\right]\Phi(z,u) = (n-1)!\binom{n}{k}\binom{n-2}{k-1}, \qquad k = 1,\ldots,n-1. \tag{9}$$

Using (Pascal triangle)

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$

and

$$\binom{n}{k} = \frac{n}{k}\binom{n-1}{k-1}$$

entails (with $C_{n,1}=n!$ and $C_{n,n}=0$)

$$C_{n+1,k}=\frac{n\left(n-1\right)}{n-k}C_{n,k}+\frac{n\left(n-1\right)}{k-1}C_{n,k-1}, \tag{10}$$

from which the following three-term recurrence for $\mathbf{P}\left(L_n=k\right)=\frac{C_{n,k}}{C_n}$ can be deduced (with $\mathbf{P}\left(L_n=n\right)=\mathbf{P}\left(L_n=0\right)=0$):

$$\mathbf{P}\left(L_{n+1}=k\right)=\frac{C_{n+1,k}}{C_{n+1}}=\frac{n\left(n-1\right)}{2\left(2n-1\right)}\left[\frac{1}{n-k}\mathbf{P}\left(L_n=k\right)+\frac{1}{k-1}\mathbf{P}\left(L_n=k-1\right)\right]. \tag{11}$$

For such random walks, the next step is either a no-move step or a move-up one. What (11) says is that whenever L_n approaches its extremal values, either 1 (or $n-1$), there is a smaller chance (inversely proportional to this value) that the next connection will result in a no-move step (a move-up step); anti-preferential attachment rule holds. Here, the ratio $\rho_{k,n}:=q_{k-1,k}^{(n)}/q_{k,k}^{(n)}=\left(n-k\right)/\left(k-1\right)$ obeys

$$\begin{aligned} q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &< 1 \text{ if } k>\left(n+1\right)/2,\\ q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &> 1 \text{ if } k<\left(n+1\right)/2, \end{aligned}$$

translating that L_n is again attracted to the central part of its support, whose two endpoints $\{1,n-1\}$ are (asymmetrically) repelling. Here,

$$n!\left[z^n\right]\Phi\left(z,1\right)=\frac{\left(2n-2\right)!}{\left(n-1\right)!}=\sum_{k=1}^{n-1}C_{n,k}=C_n.$$

Next,

$$\begin{aligned} \partial_u\Phi\left(z,1\right)&=\frac{z}{2}+\frac{z}{2\sqrt{1-4z}}=z\frac{1-\Phi\left(z\right)}{1-2\Phi\left(z\right)},\\ n!\left[z^n\right]\partial_u\Phi\left(z,1\right)&=n!\left[z^{n-1}\right]\left(\frac{1}{2}+\frac{1}{2\sqrt{1-4z}}\right)=\frac{n}{2}C_n \text{ for } n\geq 2. \end{aligned}$$

$$\begin{aligned} \mathbf{E}\left(L_n\right) &= \frac{\left[z^n\right]\partial_u\Phi\left(z,1\right)}{\left[z^n\right]\Phi\left(z,1\right)}=\frac{n}{2},\\ \sigma^2\left(L_n\right) &\sim n/8. \end{aligned}$$

The variance term can be computed from (see Drmota 2009, p. 83)

$$\sigma^2\left(L_n\right)=n!\left[z^n\right]\left[\partial_u^2\Phi\left(z,1\right)+\partial_u\Phi\left(z,1\right)\right]/C_n-\left[n!\left[z^n\right]\partial_u\Phi\left(z,1\right)/C_n\right]^2.$$

A CLT holds.

2.2 Increasing (or recursive) trees

A size-n rooted and increasing Cayley tree has vertices with indices or labels $\{1,\ldots,n\}$ increasing for any path from the root to its leaves. Wherever a new connection is created in this tree, the adjunction of a new node with index $n+1$ will result in a size-$(n+1)$ rooted increasing tree. Increasing trees can in addition be unordered (Cayley) or ordered. Such trees were studied by Bergeron et al. (1992).

2.2.1 Cayley (unordered) increasing trees

The g.f. of unordered, rooted increasing trees solves the ordinary differential equation $\Phi'(z) = e^{\Phi(z)}$, $\Phi(0)=0$, hence $\Phi(z) = -\log(1-z)$ with $z_c = \sup(z>0 : \Phi(z) < \infty) = 1$ with $\Phi(z_c) = \infty$. Hence, $\Phi(z)$ has a logarithmic singularity at $z_c = 1$. We get $C_n = (n-1)!$ and $c_n = 1/n$.

The joint generating of nodes and leaves solves:

$$\partial_z \Phi(z,u) = u - 1 + e^{\Phi(z,u)}.$$

Hence,

$$z = \int_0^{\Phi(z,u)} \frac{dz'}{u-1+e^{z'}},$$

with solution

$$\Phi(z,u) = \log\left(\frac{1-u}{1-ue^{z(1-u)}}\right).$$

With $k=1,\ldots,n-1$, we have

$$\begin{aligned} C_{n,k} &:= n!\left[z^n u^k\right]\Phi(z,u) = (n-1)!\left[z^{n-1}u^k\right]\left(u-1+\frac{1-u}{1-ue^{z(1-u)}}\right) \\ &=: E_{n-1,k}, \qquad k=1,\ldots,n-1, \end{aligned} \tag{12}$$

a shifted version of the first kind Eulerian triangle, for which

$$C_{n+1,k} = kC_{n,k} + (n-k+1)C_{n,k-1}. \tag{13}$$

Also, from

$$\Phi(z,1) = -\log(1-z),$$

$$C_n := [z^n]\Phi(z,1) = (n-1)! = \sum_{k=1}^{n-1} C_{n,k}.$$

With $\mathbf{P}(L_n = k) = \frac{C_{n,k}}{C_n}$, $k=1,\ldots,n-1$, therefore,

$$\mathbf{P}(L_{n+1} = k) = \frac{C_{n+1,k}}{C_{n+1}} = \frac{k}{n}\mathbf{P}(L_n = k) + \left(1-\frac{k-1}{n}\right)\mathbf{P}(L_n = k-1), \tag{14}$$

the dynamics of a conventional inhomogeneous Markov chain with transition probabilities. Multiplying (14) by k (respectively, k^2) and summing over the admissible range of k yields recurrences for $\mathbf{E}(L_n)$ (respectively, $\mathbf{E}\left(L_n^2\right)$) from which

$$\begin{aligned}\mathbf{E}(L_n) &= \frac{[z^n]\,\partial_u\Phi(z,1)}{[z^n]\,\Phi(z,1)} = n/2,\\ \sigma^2(L_n) &= 7n/12 + 1/3 \sim 7n/12.\end{aligned}$$

A CLT holds.

What (14) says is that, whenever L_n approaches its extremal values, either 1 (or $n-1$), there is a large probability (proportional to this value) that the next connection will propel the walker back inside its support. Else, the ratio $\rho_{k,n} := q_{k-1,k}^{(n)}/q_{k,k}^{(n)} = (n-k+1)/k$ also obeys:

$$\begin{aligned} q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &< 1 \text{ if } k > (n+1)/2,\\ q_{k-1,k}^{(n)}/q_{k,k}^{(n)} &> 1 \text{ if } k < (n+1)/2,\end{aligned}$$

translating that the dynamics of L_n is attracted by the central body of its support, whose endpoints remain repelling.

2.2.2 Increasing ordered (plane) trees

The exponential generating function (e.g.f.) $\Phi(z)$ of such ordered, rooted increasing trees solves: $\Phi'(z) = 1/(1-\Phi(z))$, $\Phi(0) = 0$. Hence, $\Phi(z) = 1-\sqrt{1-2z}$ with

$$C_n = n!\,[z^n]\,\Phi(z) = (2n-3)!! = 2^{-(n-1)}(2n-2)!/(n-1)!, \tag{15}$$

and $C_{n+1} = (2n-1)\,C_n$. There are thus 2^{n-1} more of simple ordered n-trees than increasing ordered n-trees. The factor 2^{n-1} is the number of ordered partitions of n into positive summands.

Here, $z_c = \sup(z > 0 : \Phi(z) < \infty) = 1/2$ with $\Phi(z_c) = 1$, $\Phi'(z_c) = \infty$.

The joint e.g.f. of atoms and leaves solves:

$$\partial_z\Phi(z,u) = u - 1 + 1/(1-\Phi(z,u))$$

or

$$z = \int_0^{\Phi(z,u)} \frac{dz'}{u-1+1/(1-z')}.$$

With $C(z) = \sum_{n\geq 1}\frac{n^{n-1}}{n!}z^n$ the Cayley function solving $C(z)\,e^{C(z)} = z$,

$$\Phi(z,u) = \frac{u - C\left(ue^{-u}e^{z(1-u)^2}\right)}{1-u}$$

as the inverse function of

$$\int_0^z \frac{dz'}{u-1+1/(1-z')} = \frac{\log(1+(1-u)\,z/u) - (1-u)\,z}{(1-u)^2},$$

(see Bergeron et al. (1992)). The triangular array

$$C_{n,k} := n!\left[z^n u^k\right]\Phi(z,u), \qquad k = 1,\ldots,n-1,$$

constitutes a shifted version of the second kind Eulerian triangle, with

$$C_{n+1,k} = kC_{n,k} + (2n-k)\,C_{n,k-1}. \tag{16}$$

Furthermore,

$$C_n := n!\,[z^n]\,\Phi(z,1) = (2n-3)!! = \sum_{k=1}^{n-1} C_{n,k}.$$

With $\mathbf{P}(L_n = k) = \frac{C_{n,k}}{C_n}$, $k = 1,\ldots,n-1$, therefore,

$$\mathbf{P}(L_{n+1} = k) = \frac{C_{n+1,k}}{C_{n+1}} = \frac{k}{2n-1}\mathbf{P}(L_n = k) + \left(1 - \frac{k-1}{2n-1}\right)\mathbf{P}(L_n = k-1). \tag{17}$$

Multiplying (17) by k and summing over the admissible range of k yields recurrences for $\mathbf{E}(L_n)$ from which

$$\mathbf{E}(L_n) := \frac{[z^n]\,\partial_u \Phi(z,1)}{C_n} = \frac{2n-1}{3}.$$

A CLT holds.

Remark: When conditioning, as above, on the number n of nodes of a tree with g.f. $\Phi(z)$, the law of L_n has bounded support $1,\ldots,n-1$. One may wish to consider the distribution of say N_k, the number of nodes of the tree given its number of leaves is k; but then, the law of N_k has unbounded support $k, k+1, \ldots$. To this end, we first observe that the probability law of N, the number of nodes in a tree with g.f. $\Phi(z)$, is given by the tilting

$$\mathbf{P}(N = n) = \frac{z^n C_n / n!}{\Phi(z)},$$

for any $z < z_c$ ($z \leq z_c$), depending on $\Phi(z_c) = \infty$ ($\Phi(z_c) < \infty$). Tilting is necessary in the randomization process because of the divergence of the series c_n. Then, with $n \geq k$,

$$\begin{aligned}\mathbf{P}(N_k = n) &= \mathbf{P}(L_n = k)\,\frac{z^n C_n/n!}{\Phi(z)} \Big/ \sum_{n \geq k} \text{Num} \\ &= \frac{z^n C_{n,k}/n!}{[u^k]\Phi(z,u)},\end{aligned}$$

where $[u^k]\Phi(z,u) = \sum_{n \geq k} z^n C_{n,k}/n!$ is the horizontal g.f. of $\Phi(z,u)$. The law of N_k necessarily depends on z.

3 Root degree

With $\Phi(z,1) = \Phi(z)$ solving $\Phi(z) = z\phi(\Phi(z))$ or $\Phi'(z) = \phi(\Phi(z))$, define

$$R(z,u) = z\phi(u\Phi(z)) \text{ or } R(z,u) = \int_0^z \phi\left(u\Phi\left(z'\right)\right) dz', \tag{18}$$

where u marks the root-degree of the tree in that

$$R_{n,k} := n!\left[u^k z^n\right] R(z,u)$$

is the number of trees with n nodes and root-degree k.

3.1 Nonincreasing trees

3.1.1 Cayley (unordered) trees

- If $R(z,u) = ze^{u\Phi(z)}$, $\partial_u R(z,1) = z\Phi(z) e^{\Phi(z)} = \Phi(z)^2$. By Lagrange inversion theorem, $\mathbf{E}(R_n) = 2\left(1-\frac{1}{n}\right) \sim 2$. More generally, we have

$$
\begin{aligned}
\left[z^n\right] R(z,u) &= \left[z^{n-1}\right] 1/\left(1-u\Phi(z)\right) \\
&= \frac{1}{n-1}\left[z^{n-2}\right] \phi(uz)' \phi(z)^{n-1} \\
&= \frac{u}{n-1}\left[z^{n-2}\right] \phi'(uz) \phi(z)^{n-1} \\
&= \frac{u}{n-1}\left[z^{n-2}\right] e^{z(u+n-1)} \\
&= \frac{u(u+n-1)^{n-2}}{(n-1)!}.
\end{aligned}
$$

Hence,

$$R_{n,k} := n!\left[z^n u^k\right] R(z,u) = n\binom{n-2}{k-1}(n-1)^{n-1-k} \text{ and}$$

$$\mathbf{E}\left(u^{R_n}\right) = \frac{n!\left[z^n\right] R(z,u)}{C_n} = \frac{nu(u+n-1)^{n-2}}{n^{n-1}} = u\left(1-\frac{1}{n}+\frac{u}{n}\right)^{n-2}, \tag{19}$$

a shifted binomial distribution for which $\mathbf{E}\left(u^{R_n}\right) \underset{n\to\infty}{\to} ue^{-(1-u)}$, the probability generating function (p.g.f.) of a shifted Poisson (1) random variable (r.v.).

3.1.2 Ordered (plane) trees

- If, as for this case, $R(z,u) = z/\left(1-u\Phi(z)\right)$,

$$
\begin{aligned}
\left[z^n\right] R(z,u) &= \left[z^{n-1}\right] 1/\left(1-u\Phi(z)\right) \\
&= \frac{1}{n-1}\left[z^{n-2}\right] \phi(uz)' \phi(z)^{n-1} \\
&= \frac{u}{n-1}\left[z^{n-2}\right] \phi'(uz) \phi(z)^{n-1} \\
&= \frac{u}{n-1}\left[z^{n-2}\right] \frac{1}{(1-uz)^2(1-z)^{n-1}} \\
&= \frac{1}{n-1}\sum_{k=1}^{n-1} k\binom{2n-k-1}{n-2} u^k,
\end{aligned}
$$

$$r_{n,k} := \left[z^n u^k\right] R(z,u) = \frac{k}{n-1}\binom{2n-k-1}{n-2},$$

$$\begin{aligned}\mathbf{E}\left(u^{R_n}\right) &= \frac{n!\left[z^n\right]R(z,u)}{C_n} = \frac{u}{\binom{2n-2}{n-1}}\left[z^{n-2}\right]\frac{1}{(1-uz)^2(1-z)^{n-1}} \\ &= \frac{1}{\binom{2n-2}{n-1}}\sum_{k=1}^{n-1}k\binom{2n-k-1}{n-2}u^k.\end{aligned}$$

We have

$$\partial_u R(z,1) = \frac{\Phi(z)^3}{z} \text{ and}$$

$$\begin{aligned}\left[z^n\right]\partial_u R(z,1) &= \left[z^{n+1}\right]\Phi(z)^3 = \frac{3}{n+1}\left[z^n\right]\frac{z^2}{(1-z)^n} \\ &= \frac{3}{n+1}\left[z^{n-2}\right](1-z)^{-n} = \frac{3}{n+1}\binom{2n-3}{n-2},\end{aligned}$$

$$\mathbf{E}(R_n) = \frac{\left[z^n\right]\partial_u R(z,1)}{C_n} = \frac{3n!}{n+1}\binom{2n-3}{n-2}\frac{(n-1)!}{(2n-2)!} = \frac{3n}{2(n+1)} \sim \frac{3}{2}.$$

For simple trees, there is a small (finite) number of root sub-trees (at least one of which must be very large!).

3.2 Increasing (recursive) trees

3.2.1 Cayley (unordered) trees

- For nonplane increasing trees (see Bergeron et al. 1992, p. 40 and Mahmoud et al. 1993), with $\Phi(z) = -\log(1-z)$, we have

$$\begin{aligned}\partial_z R(z,u) &= e^{u\Phi(z)} = (1-z)^{-u}, \\ R(z,u) &= \frac{1-(1-z)^{1-u}}{1-u}, \\ n!\left[z^n\right]R(z,u) &= [u]_{n-1}, \\ n!\left[z^n u^k\right]R(z,u) &= \left|s_{n-1,k}\right| =: R_{n,k}, \\ \mathbf{E}\left(u^{R_n}\right) &= \frac{[u]_{n-1}}{(n-1)!},\end{aligned}$$

$$\mathbf{E}(R_n) = \frac{\left[z^n\right]\partial_u R(z,1)}{C_n} = H_{n-1} \sim \log n, \tag{20}$$

$$\sigma^2(R_n) \sim \log n.$$

A CLT holds.

3.2.2 Ordered (plane) trees

For plane increasing trees, with $\Phi(z) = 1 - \sqrt{1-2z}$, we have

$$\begin{aligned}\partial_z R(z,u) &= \frac{1}{1-u\Phi(z)} = \frac{1}{1-u+u\sqrt{1-2z}}, \\ R(z,u) &= \frac{1-\sqrt{1-2z}}{u} + \frac{1-u}{u^2}\log\left(1-u+u\sqrt{1-2z}\right).\end{aligned}$$

Hence,

$$\begin{aligned}n!\left[z^n u^k\right] R(z,u) &= \frac{(2n-k-3)!}{2^{n-k-1}(n-k-1)!} \\ &\sim \sqrt{\frac{2}{\pi n}} e^{-k^2/(4n)},\end{aligned}$$

and

$$\mathbf{P}(R_n = k) = \frac{n!\left[z^n u^k\right] R(z,u)}{(2n-3)!!} = \frac{1}{(2n-3)!!}\frac{(2n-k-3)!}{2^{n-k-1}(n-k-1)!},$$

with

$$\mathbf{E}(R_n) = \frac{[z^n]\,\partial_u R(z,1)}{C_n} \sim \sqrt{\pi n}. \tag{21}$$

4 Height of a size-n tree (Flajolet and Sedgewick, 2009, pp. 216–217)

The height parameter is an extremal one. The height $H_n := h_{\tau_n}$ of a size-n tree τ_n is the height of one of its nodes at largest distance to the root. It obeys

$$h_{\tau_n} = 1 + \max_{\tau \prec \tau_n} h_\tau,$$

where the maximum is over all root sub-trees τ of the full size-n tree τ_n. Taking into account that the number of trees with n nodes and root-degree k is $R_{n,k} = n! r_{n,k}$, with $C_n(h) = n! c_n(h)$ the number of labeled trees' configurations with size n having height less than h, the following bivariate recurrence holds:

$$c_{n+1}(h+1) = \sum_{k=1}^{n} r_{n,k} \sum_{n_1+\ldots+n_k=n} \prod_{r=1}^{k} c_{n_r}(h).$$

There are $\binom{n-1}{k-1}$ terms in the sum over $n_l \geq 1$, $l = 1,\ldots,k$ (the number of ordered partitions of n into k parts). We have $\mathbf{P}(H_n \leq h) = c_n(h)/c_n$.

Alternatively, let $\Phi_h(z)$ be the g.f. of those simple trees generated by ϕ with height $\leq h$. We have

$$\Phi_{h+1}(z) = z\phi(\Phi_h(z)), \qquad \Phi_0(z) = z,$$

where $n!\left[z^n\right]\Phi_h(z) =: C_h(n)$ is the number of size-n labeled trees whose height is $\leq h$. With $c_h(n) := C_h(n)/n!$ and $\phi_k = \left[z^k\right]\phi(z)$, it holds that

$$\begin{aligned} c_{h+1}(n+1) &= \sum_{k=1}^{n} \phi_k \left[z^n\right] \Phi_h(z)^k \\ &= \sum_{k=1}^{n} \phi_k \sum_{n_1+\ldots+n_k=n} \prod_{l=1}^{k} c_h(n_l). \end{aligned} \tag{22}$$

Therefore, with $\overline{N}_h$ the number of trees whose height is less than or equal to h,

$$\mathbf{P}\left(\overline{N}_h = n\right) = \frac{n!\left[z^n\right]\Phi_h(z)}{n!\left[z^n\right]\Phi(z)} = \frac{c_h(n)}{c_n},$$

where $\Phi(z) = \Phi_\infty(z)$. We have $\overline{N}_h = n$ if and only if $H_n \leq n$. From implied linear recurrences and a complex analysis based on Mellin transforms:

- Cayley trees: Rényi and Szekeres (1967) have shown that $(2\pi n)^{-1/2} H_n$ has a nontrivial weak limit. Hence $\mathbf{E}(H_n) \sim \sqrt{2\pi n}$.

- For ordered trees (see De Bruijn et al. 1972): $\mathbf{E}(H_n) \sim \sqrt{\pi n}$.

In both cases, $\mathbf{E}(H_n) \sim \lambda\sqrt{n}$, where $\lambda = (\pi\zeta)^{1/2}$ with $\zeta = \frac{2\phi'(\tau)^2}{\phi(\tau)\phi''(\tau)}$ and $\phi(\tau) = \tau\phi'(\tau)$. Furthermore, $H_n/\mathbf{E}(H_n) \xrightarrow{d} X$, where X has theta density, with q-moments $(q > 1)$: $\mathbf{E}(X^q) = q(q-1)\Gamma(q/2)\zeta^q$. See Proposition $VII.16$ of Flajolet and Sedgewick (2009).

The e.g.f. $\Phi_h(z)$ of the number of increasing trees whose height is $\leq h$ obeys

$$\Phi_{h+1}(z) = \int_0^z \phi\left(\Phi_h\left(z'\right)\right) dz', \qquad \Phi_0(z) = z$$

with $C_h(n) = n!\left[z^n\right]\Phi_h(z)$ and $\mathbf{P}\left(\overline{N}_h = h\right) = \frac{C_h(n)}{C_n}$. A bivariate recurrence similar to (22) holds with $(n+1)c_{h+1}(n+1)$ instead of $c_{h+1}(n+1)$ at the left hand-side.

It holds:

- Cayley increasing trees Pittel (1994): $\mathbf{E}(H_n) \sim e\log n$, $H_n/\log n \to e$ (a.s.).

- Ordered increasing trees Pittel (1994): $\mathbf{E}(H_n) \sim \frac{1}{2s}\log n$, $H_n/\log n \to 1/(2s)$ (a.s.), $se^s = 1$ $(s = 0.27846...)$.

Increasing recursive trees' height is much smaller than the one of simple trees; concomitantly, their root degrees are much larger than the ones of simple trees.

5 Depth of a size-n tree

The depth D_n of a size-n tree is the height of a randomly chosen node in this tree. So,

$$D_n = h \text{ w.p. } \frac{N_{h,n}}{n}, \qquad h = 0, \ldots, H_n,$$

where $N_{h,n}$ is the number of nodes at height h in the size-n tree ($N_{0,n}=1$), the profile of trees. It requires the joint law of $(N_{h,n};h=0,\ldots,H_n)$ obeying $\sum_{h=0}^{H_n} N_{h,n}=n$. With $N_{h,\nu}^{(r)} \overset{d}{=} N_{h,\nu}$ i.i.d. copies,

$$N_{h+1,n}=\sum_{r=1}^{R_n} N_{h,\nu_{r,n}}^{(r)}, \qquad h=0,\ldots,H_n.$$

- For Cayley increasing trees, Devroye (1998) and Mahmoud et al. (1993) have shown that $\mathbf{E}(D_n)\sim\log n$, $\sigma^2(D_n)\sim\log n$. A CLT holds.

- For ordered increasing trees, Mahmoud et al. (1993) have shown that $\mathbf{E}(D_n)\sim\frac{1}{2}\log n$, $\sigma^2(D_n)\sim\frac{1}{2}\log n$. A CLT holds.

Equivalently, the mean depth can be obtained as the mean path-length$/n$, (see Bergeron et al. 1992, pp. 36–37), where the path-length is $\sum_{h=0}^{H_n} N_{h,n}$.

Define some local additive parameters g.f. generated by a sequence a_n's as

$$a(z,u)=\sum_{n\geq 1} C_n u^{a_n}\frac{z^n}{n!}.$$

The global additive parameters generated by the a_n's will be

$$A_{\tau_n}=a_n+\sum_{\tau\prec\tau_n} A_\tau,$$

where the sum is over all root sub-trees τ of the full size-n tree τ_n.

For tree size: $a_n=1$, $a(z,u)=u\Phi(z)$; For path length: $a_n=n$, $a(z,u)=\Phi(zu)$; For leaves: $a_n=\delta_{n,1}$, $a(z,u)=zu$.

With $A(z,u)=\sum_{n\geq 1} C_n u^{A_n}\frac{z^n}{n!}$ the additive parameters global g.f., we get

$$\begin{aligned}
\text{Simple} &: \quad A(z,u)=a(z,u)-a(z,1)+z\phi(A(z,u)),\\
\text{Recursive} &: \quad A(z,u)=a(z,u)-a(z,1)+\int_0^z \phi\left(A\left(z',u\right)\right)dz'.
\end{aligned}$$

Concerning the mean, with $A_u(z,1)$ denoting the derivative with respect to u evaluated at $u=1$,

$$\text{Simple trees: } A_u(z,1)=a_u(z,1)+z\phi'(A(z,1))A_u(z,1),$$

so (owing to: $A(z,1)=\Phi(z)$) $\Phi'(z)=\phi(\Phi(z))/(1-z\phi'(\Phi(z)))$,

$$A_u(z,1)=\frac{a_u(z,1)}{1-z\phi'(\Phi(z))},$$

where $a_u(z,1)=\sum_{n\geq 1} C_n a_n\frac{z^n}{n!}$ is the e.g.f. of the Hadamard product sequence $C_n a_n$.

$$\text{Recursive trees: } A_u(z,1)=a_u(z,1)+\int_0^z \phi'\left(A\left(z',1\right)\right)A_u\left(z',1\right)dz',$$

whose integrated form is ($a_u(z,1)=\sum_{n\geq 1} C_n a_n\frac{z^n}{n!}$, $a_{u,z}(z,1)=\sum_{n\geq 1} C_n a_n\frac{z^{n-1}}{(n-1)!}$)

$$A_u(z,1)=\Phi'(z)\int_0^z \frac{a_{z,u}\left(z',1\right)}{\Phi'\left(z'\right)}dz'.$$

As a result, we get for the mean path lengths: (simple trees) $a_u(z,1) = z\Phi'(z)$ and (recursive trees) $a_{u,z}(z,1) = \Phi'(z) + z\Phi''(z)$.

- For simple Cayley trees: $\phi(z) = e^z$, $\Phi(z) = ze^{\Phi(z)}$, $z\Phi'(z) = \Phi(z)/(1-\Phi(z))$,

$$A_u(z,1) = \frac{a_u(z,1)}{1 - z\phi'(\Phi(z))} = \frac{\Phi(z)}{(1-\Phi(z))^2},$$

$$[z^n]A_u(z,1) = \frac{1}{n}\left[z^{n-1}\right]\frac{1+z}{(1-z)^3}e^{zn},$$

taking the explicit convolution form:

$$[z^n]A_u(z,1) = \frac{1}{n}(a_\cdot * b_\cdot)_{n-1},$$

with $a_m = m(m+1)^2$ and $b_m = n^m/m!$. Next, by the Stirling formula,

$$c_n := [z^n]\Phi(z) = \frac{n^{n-1}}{n!} \sim \frac{1}{\sqrt{2\pi}}n^{-3/2}e^n,$$

translating that $\Phi(z)$ shows up an algebraic singularity of order $-1/2$ at $z_c = 1/e$. Equivalently, observing $\Phi(z_c) = 1$,

$$\Phi(z) \sim 1 - \sqrt{2}(1 - z/z_c)^{1/2} \text{ as } z \to z_c,$$

and so

$$A_u(z,1) = \frac{\Phi(z)}{(1-\Phi(z))^2} \sim \frac{1}{2}(1 - z/z_c)^{-1} \text{ as } z \to z_c,$$

so with $[z^n]A_u(z,1) \sim 1/2 \cdot e^n$. We conclude that

$$\frac{n!\,[z^n]A_u(z,1)}{C_n} = \frac{[z^n]A_u(z,1)}{c_n} \sim \sqrt{\frac{\pi}{2}}n^{3/2}.$$

The mean depth therefore is $\sqrt{\frac{\pi}{2}}n^{1/2}$.

- For simple ordered trees: $\phi(z) = (1-z)^{-1}$, $\Phi(z) = \left(1 - \sqrt{1-4z}\right)/2$, $\Phi'(z) = (1-4z)^{-1/2}$,

$$\begin{aligned} A_u(z,1) &= \frac{a_u(z,1)}{1 - z\phi'(\Phi(z))} = \frac{z(1-4z)^{-1/2}}{\left(1 - \frac{4z}{\left(1+\sqrt{1-4z}\right)^2}\right)} \\ &= \frac{z(1-4z)^{-1}}{2}\left(1 + \sqrt{1-4z}\right). \end{aligned}$$

Hence,

$$[z^n]A_u(z,1) \sim \left[z^{n-1}\right]\frac{(1-4z)^{-1}}{2} = \frac{4^{n-1}}{2}.$$

Owing to $c_n = [z^n]\Phi(z) = \frac{(2n-2)!}{n!(n-1)!} \sim \frac{1}{\sqrt{\pi}}n^{-3/2}4^{n-1}$,

$$\frac{[z^n]A_u(z,1)}{c_n} \sim \frac{1}{2}\sqrt{\pi}n^{3/2}.$$

The mean depth therefore is $\frac{1}{2}\sqrt{\pi}n^{3/2}/n = \frac{1}{2}\sqrt{\pi}n^{1/2}$.

- Concerning recursive Cayley trees (with $\Phi(z) = -\log(1-z)$ and $\Phi'(z) = 1/(1-z)$),

$$\begin{aligned} A_u(z,1) &= \Phi'(z)\int_0^z \frac{\Phi'(z') + z'\Phi''(z')}{\Phi'(z')}dz' \\ &= \frac{1}{1-z}\int_0^z \left(1 + z'\left[\log \Phi'(z')\right]'\right) \\ &= \frac{-\log(1-z)}{1-z}. \end{aligned}$$

Therefore, $[z^n]A_u(z,1) \sim \log n \Rightarrow n!\,[z^n]A_u(z,1)/(n-1)! = n\log n$. The mean depth therefore is of order $\log n$.

- For recursive ordered trees,

$$A_u(z,1) = \frac{2z - \log(1-2z)}{4(1-2z)^{1/2}}.$$

Owing to $c_n = [z^n]\Phi(z) = [z^n]\left(1 - \sqrt{1-2z}\right) \sim -2^n n^{-3/2}/\Gamma(-1/2)$,

$$[z^n]A_u(z,1) \sim 2^n/\Gamma(1/2)\,n^{-1/2}\log n \Rightarrow [z^n]A_u(z,1)/c_n = \frac{n}{2}\log n.$$

The mean depth therefore is of order $\frac{1}{2}\log n$.

6 Forests (nucleation/aggregation process)

Besides aggregation of atoms to a preexisting tree, growth of trees can also result from the creation of new trees forming a forest. Whenever an incoming atom launches on a new tree, we speak of a nucleation event.

With $\Phi(z,1) = \Phi(z)$ solving $\Phi(z) = z\phi(\Phi(z))$ or $\Phi'(z) = \phi(\Phi(z))$, define

$$K(z,u) = e^{u\Phi(z)} \text{ (or } K(z,u) = 1/(1-u\Phi(z))\text{)}, \tag{23}$$

where u marks the number of trees (connected components) in a forest in that

$$\overline{C}_{n,k} := n!\left[u^k z^n\right]K(z,u) = \frac{n!}{k!}[z^n]\Phi(z)^k \text{ or} \tag{24}$$

$$\overline{C}_{n,k} := n!\left[u^k z^n\right]K(z,u) = n!\,[z^n]\Phi(z)^k$$

is the number of unordered (or ordered) forests with n labeled nodes and k trees, $k = 1,\dots,n$. Furthermore,

$$\begin{aligned} \overline{C}_n &:= \sum_{k=1}^n \overline{C}_{n,k} = [z^n]K(z,1) = [z^n]e^{\Phi(z)} \text{ (unordered)}, \\ \overline{C}_n &:= \sum_{k=1}^n \overline{C}_{n,k} = [z^n]K(z,1) = [z^n]\frac{1}{1-\Phi(z)} \text{ (ordered)}, \end{aligned}$$

is the number of forests with n atoms, regardless of its number of (in) distinguishable trees.

6.1 Nonincreasing (simple) trees

6.1.1 Cayley (unordered) trees Rényi (1959)

- If $\Phi(z) = ze^{\Phi(z)}$, with $K(z,u) = e^{u\Phi(z)}$, the by Lagrange inversion theorem,

$$\begin{aligned}
[z^n] K(z,u) &= \frac{u}{n}\left[z^{n-1}\right] e^{z(u+n)} \\
&= \frac{u}{n!}(u+n)^{n-1}, \\
n!\left[z^n u^k\right] K(z,u) &= :\overline{C}_{n,k} = \binom{n-1}{k-1} n^{n-k}, \qquad k = 1,\dots,n, \\
\mathbf{P}(K_n = k) &= \frac{n!\left[z^n u^k\right] K(z,u)}{n!\left[z^n\right] K(z,1)} = \frac{1}{(1+n)^{n-1}}\binom{n-1}{k-1} n^{n-k}, \\
\binom{n-1}{k-1} n^{n-k} &= \binom{n}{k} kn^{n-k-1}.
\end{aligned}$$

Takács (1990) rather gave $\overline{C}_{n,k} = kn^{n-k-1}$ as the number of unordered forests with k Cayley trees while fixing the k different founders of the distinct trees out of $\binom{n}{k}$ different ways. See also Rényi (1959). Also,

$$\mathbf{E}\left(u^{K_n}\right) = \frac{n!\left[z^n\right] K(z,u)}{n!\left[z^n\right] K(z,1)} = u\left(1 - \frac{1}{1+n} + \frac{u}{1+n}\right)^{n-1}, \tag{25}$$

a shifted binomial distribution. In particular, $\mathbf{E}(K_n) = \frac{2n}{1+n} \sim 2$. A remarkable fact is that $R_n \overset{d}{=} K_{n-1}$. Note

$$\mathbf{E}\left(u^{K_n}\right) \sim ue^{-(1-u)} \text{ as } n \text{ is large}, \tag{26}$$

the p.g.f. of a shifted mean 1 Poisson r.v..

We finally observe as in Clarke (1958) that the triangular array $\overline{C}_{n,k}$ obeys the backward recursion

$$(n-k)\overline{C}_{n,k} = nk\overline{C}_{n,k+1}.$$

Hence, $\mathbf{P}(K_n = k) = \overline{C}_{n,k}/\overline{C}_n$, $k = 1,\dots,n$ obeys

$$\mathbf{P}(K_n = k) = \frac{nk}{n-k}\mathbf{P}(K_n = k+1), \qquad k = n-1,\dots,1,$$

with terminal condition $\mathbf{P}(K_n = n) = (n+1)^{-(n-1)}$.

6.1.2 Ordered and unordered forests

- (ordered forests of ordered trees) For such models of forests of trees, the trees themselves are assumed distinguishable, resulting in ordered forests. If $\Phi(z) = \left(1 - \sqrt{1-4z}\right)/2$, solving

$\Phi(z) = z/(1-\Phi(z))$, with $K(z,u) = 1/(1-u\Phi(z))$, then

$$\begin{aligned}
[z^n]K(z,u) &= \frac{u}{n}\left[z^{n-1}\right]\frac{1}{(1-zu)^2}(1-z)^{-n} \\
&= \frac{1}{n}\sum_{k=1}^{n} k\frac{[n]_{n-k}}{(n-k)!}u^k \\
&= \sum_{k=1}^{n}\frac{k}{n}\binom{2n-k-1}{n-k}u^k, \\
[z^n]K(z,1) &= e[z^n]e^{-\sqrt{1-2z}} = \frac{1}{n+1}\binom{2n}{n},
\end{aligned}$$

where we used the identity

$$[z^n]K(z,1) = \sum_{k=1}^{n}\frac{k}{n}\binom{2n-k-1}{n-k} = \frac{1}{n+1}\binom{2n}{n}.$$

Hence

$$\mathbf{P}(K_n = k) = \frac{n!\left[z^n u^k\right]K(z,u)}{n!\,[z^n]K(z,1)} = \frac{n+1}{\binom{2n}{n}}\frac{k}{n}\binom{2n-k-1}{n-k}. \tag{27}$$

Observing

$$\partial_u K(z,1) = \Phi(z)/(1-\Phi(z))^2,$$

$$\begin{aligned}
[z^n]\partial_u K(z,1) &= \frac{1}{n}\left[z^{n-1}\right]\frac{1+z}{(1-z)^{n+3}} \\
&= \frac{1}{n}\left(\left[z^{n-1}\right](1-z)^{-(n+3)} + \left[z^{n-2}\right](1-z)^{-(n+3)}\right) \\
&= \frac{1}{n}\binom{2n}{n-2}\frac{3n+1}{n-1},
\end{aligned}$$

so

$$\begin{aligned}
\mathbf{E}(K_n) &= \frac{[z^n]\partial_u K(z,1)}{[z^n]K(z,1)} = \frac{n+1}{n}\frac{3n-1}{n-1}\frac{\binom{2n}{n-2}}{\binom{2n}{n}} \\
&= \frac{3n-1}{n}\frac{n-1}{n+2} \sim 3.
\end{aligned}$$

The variance can be computed using

$$\partial_u^2 K(z,1) = 2\Phi(z)^2/(1-\Phi(z))^3.$$

We get $\mathbf{E}(K_n(K_n-1)) = 10\frac{n-1}{n+2} \sim 10$, so that $\sigma^2(K_n) \sim 4$ and showing that the limit law is not Poisson.

For such tree models, there is a small number of connected components (at least one of which must be very large).

- (Unordered forests of ordered trees): If $K(z,u) = e^{u\Phi(z)}$, for $k = 1, \ldots, n$, then

$$\begin{aligned}
\overline{C}_{n,k} : &= n! \left[z^n u^k \right] K(z,u) = \frac{n!}{k!} [z^n] \Phi(z)^k \\
&= \frac{(n-1)!}{(k-1)!} \left[z^{n-1} \right] z^{k-1} (1-z)^{-n} \\
&= \frac{(n-1)!}{(k-1)!} \left[z^{n-k} \right] (1-z)^{-n} \\
&= \frac{(n-1)!}{(k-1)!} \frac{[n]_{n-k}}{(n-k)!} = \frac{(2n-k-1)!}{(k-1)!\,(n-k)!}.
\end{aligned}$$

These $\overline{C}_{n,k}$'s are readily checked to obey the three term recurrence

$$\overline{C}_{n+1,k} = (2n-k)\,\overline{C}_{n,k} + \frac{n}{k-1}\overline{C}_{n,k-1}.$$

Furthermore,

$$\begin{aligned}
\overline{C}_n &= n!\,[z^n]\, e^{\Phi(z)} = (n-1)! \left[z^{n-1} \right] e^z (1-z)^{-n} \\
&= (n-1)! \sum_{m=1}^{n} \frac{1}{(m-1)!} \left[z^{n-m} \right] (1-z)^{-n} = \sum_{k=1}^{n} \frac{(2n-k-1)!}{k!\,(n-k-1)!} \\
&= \sum_{l=1}^{n} \frac{(n+l-1)!}{(l-1)!\,(n-l)!}.
\end{aligned} \tag{28}$$

6.2 Increasing trees

6.2.1 Cayley (unordered) trees

- In the case of nonplane increasing trees for which $\Phi(z) = -\log(1-z)$,

$$\begin{aligned}
K(z,u) &= e^{u\Phi(z)} = (1-z)^{-u}, \\
n!\,[z^n]\, K(z,u) &= [u]_n, \\
n! \left[z^n u^k \right] K(z,u) &= |s_{n,k}|, \\
\mathbf{E}\left(u^{K_n} \right) &= \frac{[u]_n}{n!}.
\end{aligned}$$

Here $\overline{C}_{n,k} = |s_{n,k}| \equiv S_{n,k}(-1,0,0)$ are the signless Stirling numbers of the first kind, obeying

$$\overline{C}_{n+1,k} = n\overline{C}_{n,k} + \overline{C}_{n,k-1}, \tag{29}$$

$$\mathbf{E}(K_n) = \frac{[z^n]\, \partial_u K(z,1)}{C_n} = H_n \sim \log n,$$

$$\sigma^2(K_n) \sim \log n.$$

A CLT holds.

6.2.2 Ordered (plane) trees

- In the case of ordered increasing trees for which $\Phi(z) = 1-\sqrt{1-2z}$ solving $\Phi(z) = 1/(1-\Phi(z))$,

$$K(z,u) = e^{u\left(1-\sqrt{1-2z}\right)}. \tag{30}$$

Here, with $S_{n,k}(-\alpha_2,-\alpha_1;w_2)$ the generalized Stirling triangle defined in Hsu and Shiue (1998),

$$\begin{aligned}
n!\left[z^n u^k\right] K(z,u) &= \frac{n!}{k!}[z^n]\Phi(z)^k = \overline{C}_{n,k} := S_{n,k}(-2,-1;0),\\
\overline{C}_{n,k} &\equiv S_{n,k}(-2,-1,0) = (2(n-k)-1)!!\binom{2n-k-1}{2(n-k)}\\
&= \frac{(2(n-k)-1)!}{2^{n-k-1}(n-k-1)!}\binom{2n-k-1}{2(n-k)} = \frac{1}{2^{n-k}}\frac{(2n-k-1)!}{(n-k)!(k-1)!},\\
\overline{C}_{n+1,k} &= (2n-k)\overline{C}_{n,k} + \overline{C}_{n,k-1}.
\end{aligned}$$

Hence,

$$\mathbf{P}(K_{n+1}=k) = \frac{\overline{C}_n}{\overline{C}_{n+1}}\left((2n-k)\mathbf{P}(K_n=k) + \mathbf{P}(K_n=k-1)\right). \tag{31}$$

With $\overline{C}_n := \sum_{k=1}^{n}\overline{C}_{n,k} = \sum_{l=0}^{n-1}2^{-l}\frac{(n+l-1)!}{l!(n-l-1)!}$,

$$\mathbf{P}(K_n=k) = \frac{n!\left[z^n u^k\right]K(z,u)}{n![z^n]K(z,1)} = \frac{\overline{C}_{n,k}}{\overline{C}_n},$$

$$\overline{C}_{n+1} = (2n+1)\overline{C}_n - \left\langle \overline{C}_n\right\rangle,$$

$$\frac{\overline{C}_{n+1}}{\overline{C}_n} \sim 2n-1 \Rightarrow \mathbf{E}(K_n) = \frac{n![z^n]\partial_u K(z,1)}{n![z^n]K(z,1)} = \frac{\left\langle \overline{C}_n\right\rangle}{\overline{C}_n} \sim 2,$$

K_n has a shifted Poisson(1) limit law. We have

$$K(z,u) = e^u\left(1+\sum_{k\geq 1}\frac{(-u)^k}{k!}(1-2z)^{k/2}\right),$$

with singularity at $z=1/2$. Singularity analysis of the singular expansion yields

$$\begin{aligned}
n![z^n]K(z,u) &= n!e^u\left(\sum_{k\geq 1}\frac{(-u)^k}{k!}\frac{n^{-k/2}2^n}{n}\frac{1}{\Gamma(-k/2)}\right)\\
&\sim (n-1)!2^n e^u\left(\frac{un^{-1/2}}{2\sqrt{\pi}} + O\left(n^{-3/2}\right)\right).
\end{aligned}$$

Therefore ,

$$\begin{aligned}
\overline{C}_n &= n![z^n]K(z,1) \sim (n-1)!2^n e\left(\frac{n^{-1/2}}{2\sqrt{\pi}} + O\left(n^{-3/2}\right)\right),\\
\frac{\overline{C}_{n+1}}{\overline{C}_n} &\sim 2n\left(1-\frac{1}{2n}+O\left(n^{-2}\right)\right) \sim 2n-1.
\end{aligned}$$

Therefore,

$$\mathbf{E}\left(u^{K_n}\right)=\frac{n!\left[z^n\right]K\left(z,u\right)}{n!\left[z^n\right]K\left(z,1\right)}\sim ue^{-(1-u)}, \tag{32}$$

the p.g.f. of a shifted Poisson(1) r.v..

This three-term recurrence, giving $\mathbf{P}\left(K_n=k\right)$ in all cases, is the one of triangular Markov probability sequences. The latter recursion (31), for example, may be written as

$$\mathbf{P}\left(K_{n+1}=k\right)=q_{k,k}^{(n)}\mathbf{P}\left(K_n=k\right)+q_{k-1,k}^{(n)}\mathbf{P}\left(K_n=k-1\right),$$

defining the transition coefficients $q_{k,k}^{(n)}$ and $q_{k-1,k}^{(n)}$ (not probabilities). Introduce the superdiagonal transition matrix

$$Q_{n,n}:=\begin{pmatrix} q_{1,1}^{(n)} & q_{1,2}^{(n)} & 0 & \cdots \\ 0 & \ddots & \ddots & 0 \\ 0 & 0 & q_{n-1,n-1}^{(n)} & q_{n-1,n}^{(n)} \\ \vdots & \vdots & 0 & q_{n,n}^{(n)} \end{pmatrix},$$

and let $Q_{n-1,n}$ be its $(n-1)\times n$ truncated version. With $\pi_1^K:=1$, taking into account the boundary conditions, the distributions $\pi_n^K:=\left(\mathbf{P}\left(K_n=k\right),\ k\in\{1,\ldots,n\}\right)$, $n\geq 1$, satisfy the recursion $\pi_n^K=\left(\pi_{n-1}^K,0\right)Q_{n,n}=\pi_{n-1}^K Q_{n-1,n}$, $n\geq 2$. Thus, $\pi_n^K=\overleftarrow{\prod}_{m=2,\ldots,n}Q_{m-1,m}$, $n\geq 2$ is an integrated form solution of the recursion, as a left product of nested rectangular matrices. Because for each n, π_n^L is a probability vector, we get that π_n^L is orthogonal to the 1-shifted column sum vector $\mathbf{q}_n-\mathbf{1}$ of $Q_{n,n+1}$ with kth entry $q_{k,k}^{(n+1)}+q_{k,k+1}^{(n+1)}-1$, $k=1,\ldots,n$.

7 Weighted trees

Simply generated weighted trees are weighted versions of rooted trees and have been introduced by Meir and Moon (1978). They are obtained while assigning to each node x of a size-n tree τ_n a weight $w_{b_n(x)}$, where $b_n(x)$ is the outdegree of x. The weight of a particular tree τ_n is then the product $\prod_{x\in\tau_n}w_{b_n(x)}$, and while summing over all τ_n, we get the weight of all size-n trees.

7.1 Weighted simple trees

Let $w\left(\tau_n\right)=\prod_{b=0}^{n-1}w_b^{n_b(\tau_n)}$ be the (multiplicative) weight of an unlabeled rooted tree τ_n with n nodes ($|\tau_n|=n$) having $n_b\left(\tau_n\right)$ nodes with outdegree (branching number) b. The weight $w\left(\tau_n\right)$ is the product over the n nodes x of τ_n of the $w_{b_n(x)}$'s, where $b_n(x)$ is the outdegree of x. Then $W_n=\sum_{\tau_n}w\left(\tau_n\right)$ is the weight of all size-n such trees associated to the weight sequence $\mathbf{w}:=\left(w_b\geq 0,\ b\geq 0\right)$. The number of these trees is $c_n:=C_n/n!=\left[z^n\right]\Phi\left(z\right)$. Let $\Phi_{\mathbf{w}}\left(z\right)=\sum_{n\geq 1}z^nW_n$. Then $\Phi_{\mathbf{w}}\left(z\right)$ solves $\Phi_{\mathbf{w}}\left(z\right)=zg\left(\Phi_{\mathbf{w}}\left(z\right)\right)$, where $g\left(z\right)=\sum_{b\geq 0}w_bz^b$.

By Lagrange inversion formula, we get the identity

$$W_n=\left[z^n\right]\Phi_{\mathbf{w}}\left(z\right)=\frac{1}{n}\left[z^{n-1}\right]g\left(z\right)^n=\sum_{\tau_n}\prod_{x\in\tau_n}w_{b_n(x)}=\sum_{\tau_n}\prod_{b=0}^{n-1}w_b^{n_b(\tau_n)}.$$

Examples of $\mathbf{w}$ are as follows:

- $\mathbf{w}=(\varepsilon_b;\ b\geq 0)$ with $\varepsilon_b\in\{0,1\}$. Only the branches of the tree with $\varepsilon_b=1$ contribute to its weight.

- $w_b=a_1a_2^b$, $a_1,a_2>0$ with $g(z)=a_1\sum_{b\geq 0}(a_2z)^b=a_1/(1-a_2z)$. Note $w(\tau_n)=a_1^na_2^{n-1}$ in view of $\sum_{b=0}^{n-1}n_b(\tau_n)=n$ and $\sum_{b=1}^{n-1}bn_b(\tau_n)=n-1$ (the total tree length). As a result, $W_n=a_1^na_2^{n-1}c_n$, where $c_n=C_n/n!=\frac{(2n-2)!}{n!(n-1)!}$. In that separable case, each tree τ_n has equal weight $a_1^na_2^{n-1}$ and $W_n=a_1^na_2^{n-1}c_n$. Note $\Phi_{\mathbf{w}}(z)=\left(1-\sqrt{1-4a_1a_2z}\right)/(2a_2)=\Phi(a_1a_2z)/a_2$.

- If the sequence of weights w_b is summable, then $g(z)$'s can be normalized to yield the p.g.f.'s $g(z)/g(1)$; the corresponding integrated solutions change from $\Phi_{\mathbf{w}}(z)$ to $\Phi_{\mathbf{w}}(z/g(1))$.

- If $\mathbf{w}:=(w_b\geq 0,\ b\geq 0)$ is directly a probability sequence with $\sum_{b\geq 0}w_b=1$, then $\Phi_{\mathbf{w}}(z)$ is the g.f. of the total progeny of a Galton–Watson tree with branching mechanism g, solving $\Phi_{\mathbf{w}}(z)=zg(\Phi_{\mathbf{w}}(z))$, $\Phi_{\mathbf{w}}(0)=0$. The numbers $[z^n]\Phi_{\mathbf{w}}(z)=W_n$ are the (sub-)probabilities of a progeny with size n.

Examples:

• (Poisson offspring) If $w_b=e^{-\mu}\mu^b/b!$, $b\geq 0$ with $\mu>0$, then, with $g(z)=e^{\mu(z-1)}$ and $\Phi(z)$ solving $\Phi(z)=ze^{\Phi(z)}$, the Cayley e.g.f., we have

$$\Phi_{\mathbf{w}}(z)=\frac{1}{\mu}\Phi\left(\mu e^{-\mu}z\right)=\frac{1}{\mu}\sum_{n\geq 1}\frac{n^{n-1}}{n!}\left(\mu e^{-\mu}z\right)^n.$$

The Borel e.g.f. $\Phi_{\mathbf{w}}(z)$ has thus now a displaced algebraic singularity of order $-1/2$ at $z_c=\frac{1}{\mu}e^{\mu-1}>1$ with $\Phi_{\mathbf{w}}(z_c)=1/\mu$ and $\Phi_{\mathbf{w}}'(z_c)=\infty$. Note

$$\begin{aligned}\mu &\leq 1:\ \Phi_{\mathbf{w}}(1)=1,\\ \mu &> 1:\ \Phi_{\mathbf{w}}(1)=\rho<1/\mu,\end{aligned}$$

where ρ is the extinction probability of the supercritical Poisson GW process solving $g(\rho)=e^{\mu(\rho-1)}=\rho$.

• (geometric offspring) If $w_b=\overline{a}a^b$, $b\geq 0$ with $a\in(0,1)$ and $\overline{a}=1-a$, then, with $g(z)=\overline{a}/(1-az)$ and $\Phi(z)$ solving $\Phi(z)=z/(1-\Phi(z))$, the e.g.f. of ordered trees, we have

$$\Phi_{\mathbf{w}}(z)=\frac{1}{a}\Phi(\overline{a}az)=\frac{1-\sqrt{1-4\overline{a}az}}{2a}.$$

Also, $\Phi_{\mathbf{w}}(z)$ has thus now a displaced algebraic singularity of order $-1/2$ at $z_c=1/(4\overline{a}a)>1$ with $\Phi_{\mathbf{w}}(z_c)=1/(2a)$ and $\Phi_{\mathbf{w}}'(z_c)=\infty$. Note, with $\mu=a/\overline{a}$, the mean offspring number is

$$\begin{aligned}\mu &\leq 1:\ \Phi_{\mathbf{w}}(1)=1,\\ \mu &> 1:\ \Phi_{\mathbf{w}}(1)=\rho<1/(2a),\end{aligned}$$

where $\rho=\overline{a}/a$ is the extinction probability of the supercritical geometric GW process solving $g(\rho)=\overline{a}/(1-a\rho)=\rho$.

In both cases, there is a possibility for such weighted trees that they are finite with some nonzero probability.

7.2 Weighted increasing trees

Consider the increasing ordered tree g.f. $\Phi(z)$ solving $\Phi'(z) = 1/(1-(\Phi(z)))$, $\Phi(0) = 0$, with $\Phi(z) = 1 - \sqrt{1-2z}$ and

$$C_n = n![z^n]\Phi(z) = (2n-3)!! = 2^{-(n-1)}(2n-2)!/(n-1)!.$$

Let $W_n = \sum_{\tau_n} w(\tau_n)$ (where $w(\tau_n) = \prod_{b=0}^{n-1} w_b^{n_b(\tau_n)}$) be the weight of all size n increasing labeled rooted trees ($|\tau_n| = n$) associated to the weight sequence $\mathbf{w} := (w_b \geq 0,\ b \geq 0)$, with $w_b = [z^b]\, g(z)$ the weight of an atom with out-degree b. Let then $\Phi_{\mathbf{w}}(z) = \sum_{n\geq 1} z^n W_n/n!$, with $W_n = n![z^n]\Phi_{\mathbf{w}}(z)$. Then $\Phi_{\mathbf{w}}(z)$ solves $\Phi'_{\mathbf{w}}(z) = g(\Phi_{\mathbf{w}}(z))$, where $g(z) = \sum_{b\geq 0} w_b z^b$. Note $\Phi(z)$ is obtained when $w_b = 1$, $b \geq 0$: the number of such n-trees is $C_n = (2n-3)!!$.

The probability to observe a particular size-n tree τ_n among all size-n trees is $w(\tau_n)/W_n$. The tilted probability to observe a tree of size n among all possible trees is

$$\frac{z^n W_n/n!}{\Phi_{\mathbf{w}}(z)},$$

for those $z \in (0, z_c)$, where $z_c = \inf(z > 0 : \Phi_{\mathbf{w}}(z) = \infty) = \int_0^{z_*} \frac{dz'}{g(z')}$ with $z_* = \inf(z > 0 : g(z) = \infty)$.

Examples of $\mathbf{w}$ are as follows:

- $w_b = a_1 a_2^b [c]_b/b!$, $a_1, a_2, c > 0$ with $g(z) = a_1 \sum_{b\geq 0} \frac{[c]_b}{b!}(a_2 z)^b = a_1(1-a_2 z)^{-c}$.

- $w_b = a_1 a_2^b$, $a_1, a_2 > 0$ with $g(z) = a_1/(1-a_2 z)$. In that separable case, each tree τ_n has equal weight $a_1^n a_2^{n-1}$ and $W_n = a_1^n a_2^{n-1} C_n$ with $W_{n+1}/W_n = a_1 a_2(2n-1)$.

For the three special g.f.'s,

$$\begin{aligned} g(z) &= (1-\alpha_1 z)^{-(\alpha_2/\alpha_1 - 1)} \ (\alpha_2 > \alpha_1 > 0, 0 < \alpha := \alpha_1/\alpha_2 < 1), \\ g(z) &= e^{\alpha_2 z} \text{ (obtained from the latter } g(z) \text{ when } \alpha_1 \to 0\text{), or} \\ g(z) &= (1+\alpha_1 z)^d, \qquad \alpha_1 > 0, d \in \{2,3,\ldots\}, \end{aligned}$$

the weighted e.g.f.'s $\Phi_{\mathbf{w}}(z)$ solving $\Phi'_{\mathbf{w}}(z) = g(\Phi_{\mathbf{w}}(z))$, $\Phi_{\mathbf{w}}(0) = 0$ are, respectively,

$$\begin{aligned} \Phi_{\mathbf{w}}(z) &= \frac{1}{\alpha_1}\left(1-(1-\alpha_2 z)^{\alpha_1/\alpha_2}\right) \\ &= -\log(1-\alpha_2 z) \text{ as } \alpha_1 \to 0 \\ &= \frac{1}{\alpha_1}\left(-1+[1-\alpha_1(d-1)z]^{-1/(d-1)}\right). \end{aligned}$$

Note that these special $g(z)$'s can be normalized to the p.g.f.'s $g(z)/g(1)$ with the corresponding integrated solutions changing from $\Phi_{\mathbf{w}}(z)$ to $\Phi_{\mathbf{w}}(z/g(1))$. The normalized p.g.f.'s $g(z)/g(1)$ are then, respectively, negative-binomial, Poisson, and binomial p.g.f.'s.

The formation of such increasing trees admits the following recursive tree evolution scheme (label 1 is assigned to the root):

With probability $p_b(n) := K_n^{-1}(b+1)w_{b+1}/w_b$, attach uniformly node $n+1$ to any of the $n_b(\tau_n)$ nodes with out-degree $b \in \{0,\ldots,b^*\}$ of a previous size-n increasing tree τ_n ($b^* \leq n-1$).

The normalization constant is $K_n = \sum_{b=0}^{n-1} n_b(\tau_n)(b+1)w_{b+1}/w_b$, representing the "number" of ways the new atom with label $n+1$ can be inserted in τ_n. This preferential (or uniform) attachment procedure results in a realization of τ_{n+1}; see Panholzer and Prodinger (2007). With $(B_b(n_b(\tau_n)), b \in \{0,\ldots,b^*\})$ mutually exclusive Bernoulli r.v.'s (summing to 1), each with success probability $n_b(\tau_n)\, p_b(n)$, for each $n \geq 1$, we have

$$\begin{aligned} n_0(\tau_{n+1}) &= n_0(\tau_n) + 1 - B_0(n_0(\tau_n)), \\ n_b(\tau_{n+1}) &= n_b(\tau_n) + B_{b-1}(n_{b-1}(\tau_n)) - B_b(n_b(\tau_n)), \qquad b \in \{1,\ldots,b^*\}, \\ n_{b^*+1}(\tau_{n+1}) &= 0 + B_{b^*}(n_{b^*}(\tau_n)). \end{aligned}$$

Whenever a connection to a node with outdegree b occurs, the number of nodes with out-degree b (respectively, $b+1$) decreases (increases) by one unit. In addition, a new node with outdegree-0 is always created whatever the degree of the node to which the new incoming atom connects to τ_n . Here $n_0(\tau_n)$ is the number L_n of leaves in a size-n tree.

From the first equation above giving the evolution of the number of **leaves**,

$$\begin{cases} L_{n+1} = L_n + 1 \text{ with probability } 1 - L_n p_0(n), \\ L_{n+1} = L_n \text{ with probability } L_n p_0(n). \end{cases} \tag{33}$$

This is one of standard space-time inhomogeneous Markov chains.

For the three particular $\Phi_{\mathbf{w}}$-models generated by the special g's above, using $\sum_{b=0}^{n-1} n_b(\tau_n) = n$ and $\sum_{b=1}^{n-1} b n_b(\tau_n) = n-1$ for any τ_n, we get

$$p_b(n) = \frac{1-\alpha+\alpha b}{n-\alpha}, \ \frac{1}{n}, \ \frac{d-b}{1+n(d-1)},$$

respectively, depending only on (b,n) and not on the full sequence of weights $(w_b; b = 0,\ldots,b^*)$. In the first two examples, $b \in \{0,\ldots,b^* = n-1\}$ while $b \in \{0,\ldots,b^* = d-1\}$ in the third d-ary labeled plane trees case. The number C_n of such d-ary plane trees is obtained while integrating $\Phi'_{\mathbf{w}}(z) = g_d(\Phi_{\mathbf{w}}(z))$, where $g_d(z) := (1+z)^d$. We get

$$\Phi_{\mathbf{w}}(z) = -1 + [1-(d-1)z]^{-1/(d-1)}, \text{ with}$$

$$C_n = n!\,[z^n]\,\Phi_{\mathbf{w}}(z) = (d-1)^n\,[1/(d-1)]_n = [1:d-1]_n,$$

where $[a:b]_n := a(a+b)\ldots(a+(n-1)b)$.

Note that for $g(z) = (1-\alpha_1 z)^{-(\alpha_2/\alpha_1 - 1)}$, $z_* = 1/\alpha_1$, $z_c = 1/\alpha_2$ and $\Phi_{\mathbf{w}}(z_c) = z_*$.

Such increasing trees may serve as models for phylogenetic trees in which nodes represent species and labels encode their order of appearance, and thus the chronology of evolution. The leaves of the tree are the currently living species; the different trees consist of genera.

Remark:(33) is the weighted tree extension of (17) and (14) obtained, respectively, when $\alpha_2 = 1 > \alpha_1 = 0$ and $\alpha_2 = 2 > \alpha_1 = 1$.

The joint e.g.f. of the number of atoms and the number of trees of the **forests** is

$$K_{\mathbf{w}}(z,u) = e^{u\Phi_{\mathbf{w}}(z)},$$

with

$$\overline{C}_{n,k} = n! \left[z^n u^k \right] K_{\mathbf{w}}(z,u).$$

The generalized Stirling numbers

$$\begin{array}{rcll} \overline{C}_{n,k} & \equiv & S_{n,k}(-\alpha_2, -\alpha_1; 0), & k = 0, \ldots, n, \\ \overline{C}_{n,k} & \equiv & 0 \text{ if } k > n, & \end{array}$$

are defined by the identity Hsu and Shiue (1998)

$$[w : \alpha_2]_n = \sum_{k=0}^{n} \overline{C}_{n,k} [w : \alpha_1]_k.$$

The $\overline{C}_{n,k}$ obey the (Stirling's triangle) recurrence

$$\overline{C}_{n+1,k} = (n\alpha_2 - k\alpha_1)\overline{C}_{n,k} + \overline{C}_{n,k-1} \tag{34}$$

with boundary conditions $\overline{C}_{n,0} = \delta_{n,0}$ and $\overline{C}_{n,n} = 1$ for all $n \geq 0$. With $\mathbf{P}(K_n = k) = \frac{n!\left[z^n u^k\right] K_{\mathbf{w}}(z,u)}{n!\left[z^n\right] K_{\mathbf{w}}(z,1)} = \frac{\overline{C}_{n,k}}{\overline{C}_n}$, therefore

$$\mathbf{P}(K_{n+1} = k) = \frac{\overline{C}_n}{\overline{C}_{n+1}} \left[(n\alpha_2 - k\alpha_1) \mathbf{P}(K_n = k) + \mathbf{P}(K_n = k-1) \right], \tag{35}$$

where $\overline{C}_n := \sum_{k=0}^{n} \overline{C}_{n,k} = n! [z^n] K_{\mathbf{w}}(z,1) = n! [z^n] e^{\Phi_{\mathbf{w}}(z)}$. The e.g.f. $K_{\mathbf{w}}(z,1) = e^{\Phi_{\mathbf{w}}(z)}$ has an algebraic singularity of order $-\alpha$ at $z_c = 1/\alpha_2$; by singularity analysis therefore

$$\begin{array}{rcll} \overline{C}_n & \sim & -(n-1)! \dfrac{1}{\alpha_1 \Gamma(-\alpha)} \alpha_2^n n^{-\alpha}, & \\ \dfrac{\overline{C}_n}{\overline{C}_{n+1}} & \sim & \dfrac{1}{n\alpha_2} \left(1 + O\left(n^{-1}\right)\right), & \text{for large } n. \end{array}$$

Summing (34) over k,

$$\overline{C}_{n+1} = (n\alpha_2 + 1)\overline{C}_n - \alpha_1 \left\langle \overline{C}_n \right\rangle,$$

$$\frac{\overline{C}_n}{\overline{C}_{n+1}} \sim \frac{1}{n\alpha_2} \Rightarrow \mathbf{E}(K_n) = \frac{n! [z^n] \partial_u K_{\mathbf{w}}(z,1)}{n! [z^n] K_{\mathbf{w}}(z,1)} = \frac{\left\langle \overline{C}_n \right\rangle}{\overline{C}_n} \sim 1/\alpha_1.$$

Multiplying (35) by k and summing over k yield $\mathbf{E}\left(K_n^2\right) \sim (1 + 1/\alpha_1)/\alpha_1$; hence $\sigma^2(K_n) \sim 1/\alpha_1 \sim \mathbf{E}(K_n)$.

By the Lagrange inversion formula, we have

$$\begin{array}{rcl} [z^n] K_{\mathbf{w}}(z,1) & = & \dfrac{1}{n} \left[z^{n-1}\right] e^z g(z)^n \\ & = & \dfrac{1}{n} (a_{\cdot} * b_{\cdot})_{n-1}, \end{array}$$

where $a_m = 1/m!$ and
$b_m = [z^m] g(z)^n = [z^m] (1 - \alpha_1 z)^{-n(1/\alpha - 1)} = \alpha_2^m [n(1-\alpha)]_m / m!$.

Proposition 1 (condensed phase). *If* $g(z) = (1-\alpha_1 z)^{-(\alpha_2/\alpha_1 - 1)}$ *with* $\alpha_2 > \alpha_1 > 0,\ 0 < \alpha := \alpha_1/\alpha_2 < 1$, *then* K_n *has a Poisson limit-law with mean* $1/\alpha_1$ *as* $n \to \infty$. *It corresponds to the distribution*

$$\mathbf{P}(K_n = k) = \frac{\overline{C}_{n,k}}{\overline{C}_n}, \qquad k = 0, \ldots, n,$$

of the normalized weight of unordered forests with n *labeled atoms and* k *increasing ordered trees, resulting from a random uniform choice of a configuration proportional to its weight.*

Remark: (35) is the weighted tree extension of (29) and (31) obtained, respectively, when $\alpha_2 = 1 > \alpha_1 = 0$ and $\alpha_2 = 2 > \alpha_1 = 1$. When $\alpha_1 \to 0$, K_n diverges logarithmically, with both $\mathbf{E}(K_n)$ and $\sigma^2(K_n) \sim \alpha_2 \log n$.

- Finally, we mention a random version of increasing trees. With $m > 0$, $\mathbf{w} := (w_b = m\pi_b \geq 0,\ b \geq 0)$, where (π_b) is a probability sequence ($\sum_{b\geq 0} \pi_b = 1$), $\Phi_{\mathbf{w}}(z)$ is the e.g.f. of the total progeny of a Galton–Watson increasing tree with branching mechanism $g(z) = \sum_{b\geq 0} \pi_b z^b$ and parameter m. If the total progeny is finite with probability (w.p.) 1, $\Phi_{\mathbf{w}}(1) = 1$ and $\Phi'_{\mathbf{w}}(1) = m$ is the mean number of its nodes, fixing $m = \int_0^1 \frac{dz'}{g(z')}$.

The numbers $W_n/n! = [z^n]\Phi_{\mathbf{w}}(z) = m^n \sum_{\tau_n} \prod_{b=0}^{n-1} \pi_b^{n_b(\tau_n)}/n!$ are then the probabilities of a progeny with size n. The e.g.f. $\Phi_{\mathbf{w}}(z)$ solves

$$z = \frac{1}{m}\int_0^{\Phi_{\mathbf{w}}(z)} \frac{dz'}{g(z')} =: \frac{1}{m} P(\Phi_{\mathbf{w}}(z)).$$

By Lagrange inversion formula,

$$n!\,[z^n]\Phi_{\mathbf{w}}(z) = m^n (n-1)!\,\left[z^{n-1}\right]\left(\frac{z}{P(z)}\right)^n = W_n.$$

With $z_* = \sup(z > 0 : g(z) < \infty) \in [1, \infty]$, the convergence radius of $\Phi_{\mathbf{w}}(z)$ is

$$z_c = \sup(z > 0 : \Phi_{\mathbf{w}}(z) < \infty) = \frac{1}{m}\int_0^{z_*} \frac{dz'}{g(z')},$$

with $\Phi_{\mathbf{w}}(z_c) = z_*$ owing to

$$z_c - z = \frac{1}{m}\int_{\Phi_{\mathbf{w}}(z)}^{z_*} \frac{dz'}{g(z')}.$$

Assuming $m < \int_0^{z_*} \frac{dz'}{g(z')}$ then $z_c > 1$. If so, $\Phi_{\mathbf{w}}(1) < \infty$ and $\Phi_{\mathbf{w}}(z)$ is a g.f. candidate. It is a defective p.g.f. only if $m < \int_0^1 \frac{dz'}{g(z')}$ because then $\rho := \Phi_{\mathbf{w}}(1) < 1$ (in view of $1 = \frac{1}{m}\int_0^{\Phi_{\mathbf{w}}(1)} \frac{dz'}{g(z')}$): the model is supercritical, having ρ as its extinction probability. The total progeny is finite only w.p. ρ and $\Phi'_{\mathbf{w}}(1) = mg(\rho)$ is the mean number of its nodes on the extinction set. Fixing the parameter m to its critical upper value $m_c = \int_0^1 \frac{dz'}{g(z')}$ entails $\Phi_{\mathbf{w}}(1) = 1$, $\Phi'_{\mathbf{w}}(1) = m_c > 1$. The size of the corresponding tree is then finite w.p. 1, with mean value m_c. Note that if $m = 1 < m_c = \int_0^1 \frac{dz'}{g(z')}$, then $\Phi_{\mathbf{w}}(1) < 1$ for all p.g.f.'s g.

When $m = m_c$, the e.g.f. $\Phi_{\mathbf{w}}(z)$ is proper. It admits a closed form integrable expression in the following cases (respectively, binomial, Poisson, and negative-binomial):

$$g(z) = (\pi_0 + \pi_1 z)^d \; ; \; \Phi_{\mathbf{w}}(z) = \frac{\pi_0}{\pi_1}\left((1-z/z_c)^{-1/(d-1)} - 1\right), \; z_c = 1/\left(1 - \pi_0^{d-1}\right),$$

$$g(z) = e^{-\mu(1-z)} \; ; \; \Phi_{\mathbf{w}}(z) = -\frac{1}{\mu}\log(1-z/z_c), \; z_c = 1/\left(1-e^{-\mu}\right),$$

$$g(z) = \left(\frac{q}{1-pz}\right)^{\theta} \; ; \; \Phi_{\mathbf{w}}(z) = \frac{1}{p}\left(1-(1-z/z_c)^{1/(\theta+1)}\right), \qquad z_c = 1/\left(1-q^{\theta+1}\right).$$

With $\alpha, \lambda \in (0,1)$, suppose $g(z) = 1 - \lambda (1-z)^{\alpha}$ with $w_b = \alpha\lambda [1-\alpha]_{b-1}/b!$, $b \geq 1$ and $z_* = 1$ (a Sibuya branching mechanism with infinite mean μ). In that case, $z_c = \frac{1}{m}\int_0^1 \frac{dz'}{g(z')} = \frac{1}{m}\sum_{k\geq 0}\frac{\lambda^k}{1+k\alpha} < \infty$, with $z_c > 1$ if and only if $m < \int_0^1 \frac{dz'}{g(z')} = m_c$. When $m = m_c > 1$, $z_c = 1$ and $\Phi_{\mathbf{w}}(1) = 1$. There is no major impact of the mean value μ of the branching mechanism on the extinction possibility of the increasing tree.

Acknowledgements

The author is indebted to his collaborators Francois Dunlop and Arif Mardin for many fruitful discussions during the elaboration of this manuscript.

References

Bergeron, F., Flajolet, P. and Salvy, B. (1992). Varieties of increasing trees. In: CAAP'92: 17th Colloquium on Trees in Algebra and Programming Rennes, **17** (pp. 24–48). Springer Berlin Heidelberg.

Clarke, L. E. (1958). On Cayley's formula for counting trees. *Journal of the London Mathematical Society*, **33**, 471–474.

De Bruijn, N. G., Knuth, D. E., and Rice, S. O. (1972). The average height of planted plane trees. In: *Graph Theory and Computing*, R. C. Read, Ed., Academic Press, pp. 15–22.

Devroye, L. (1998). Universal limit laws for depth in random trees. *SIAM Journal on Computing*, **28**, 409–432.

Drmota, M. (2009) *Random Trees. An Interplay Between Combinatorics and Probability.* Springer, New York.

Flajolet, P. and Sedgewick, R. (2009). *Analytic Combinatorics.* Illustrated Edition. Cambridge University Press.

Hsu, L. C. and Shiue, P. J.-S. (1998). A unified approach to generalized Stirling numbers. *Advances in Applied Mathematics*, **20**, 366–384.

Mahmoud, H., Smythe, R. T. and Szymanski, J. (1993). On the structure of random plane-oriented recursive trees and their branches. *Random Structures and Algorithms*, **4**, 151–176.

Meir, A. and Moon, J. W. (1978). On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, **30**, 997–1015.

Neveu, J. (1986) Arbres et processus de Galton-Watson. *Annales Institue Henri Poincaré Probabilités et Statistiques* **22**, 199–207.

Pitman, J. (1998). Enumerations of trees and forests related to branching processes and random walks. In: *Microsurveys in Discrete Probability* Edited by: D Aldous, J Propp. 163–180 Providence RI: American Mathematical Society.

Pittel, B. (1994). Note on the heights of random recursive trees and random m-array search trees. *Random Structures and Algorithms*, **5**, 337–347.

Rényi, A. (1959). Some remarks on the theory of trees. *Magyar Tudományos Akadémia Matematikai Kutató Intézetének Közleményei*, **4**, 73–85.

Rényi, A. and Szekeres, G. (1967). On the height of trees. *Australian Journal of Mathematics*, **7**, 497–507.

Steele, J. M. (1987). Gibbs'measures on combinatorial objects and the central limit theorem for an exponential family. *Probability in the Engineering and Informational Sciences*, **1**, 47–59.

Surya, E. and Warnke, L. (2023). Lagrange inversion formula by induction. *The American Mathematical Monthly*, **130**, 944–948.

Takács, L. (1990). On Cayley's formula for counting forests. *Journal of Combinatorial Theory*, Series A, **53**, 321–323.

Panholzer, A. and Prodinger, H. (2007). Level of nodes in increasing trees revisited. *Random Structures and Algorithms*, **31**, 203–226.

Wolfram: https://mathworld.wolfram.com/Tree.html

Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 105–118. Research Paper

On the remaining lifespan of devices

Mario Lefebvre*

Department of Mathematics and Industrial Engineering, Polytechnique Montréal, C.P. 6079, Succursale Centre-ville, Montréal, Québec, Canada H3C 3A7
Email(s): mlefebvre@polymtl.ca

Abstract. Let $X(t)$ denote the remaining lifespan of a device. Two-dimensional degenerate diffusion processes $(X(t),Y(t))$, where $Y(t)$ is a variable that influences the remaining lifespan, are proposed to model the evolution of $X(t)$ over time. These processes are defined in such a way that $X(t)$ will be strictly decreasing as time increases: $dX(t) = \rho[X(t),Y(t)]dt$, where ρ is a strictly negative function and $\{Y(t), t \geq 0\}$ is a diffusion process. Next, optimal control problems for such diffusion processes, in which the final time is the random time when the device is considered to be worn out, are considered. This type of problems is known as homing problems. The dynamic programming equation satisfied by the value function is derived and particular problems are solved explicitly for diffusion processes $\{Y(t), t \geq 0\}$ that are important for applications. To do so, we must solve non-linear partial differential equations, subject to the appropriate boundary conditions.

Keywords: Brownian motion; Degradation; Diffusion processes; First-passage time; Optimal control.

1 Introduction

Let $D(t)$ be the amount of degradation, or wear, of a device at time t. Numerous papers have been written on models for the evolution of degradation over time. Shahraki et al. (2017) wrote a review paper in which they gave 126 relatively recent references on degradation modelling. The models proposed by various authors include gamma, Wiener and inverse Gaussian processes.

Whitmore and Schenkelberg (1997), Nicolai and Dekker (2007), and Ye et al. (2013), in particular, used one-dimensional Wiener processes as models for the stochastic process $\{D(t), t \geq 0\}$. More recently, Zhang et al. (2017) as well as Zhai et al. (2018) proposed random-effects Wiener processes to model degradation. In Zhou et al. (2021), a generalized Wiener process is used. Ghamlouch et al. (2015) considered a jump-diffusion process for the evolution of the *health indicator* of a system.

*Corresponding author
Received: 15 February 2024 / Accepted: 17 June 2024
DOI: 10.22067/smps.2024.45416

 https://smps.um.ac.ir

All these papers are of good quality. Indeed, depending on the choice for their infinitesimal parameters, one-dimensional diffusion (or jump-diffusion) processes can produce good results. However, in general degradation $D(t)$ should be a strictly increasing function of t, while the *remaining useful lifetime* (RUL) should decrease as time t increases. As is well known, any diffusion process both increases and decreases in *any* time interval.

Rishel (1991) has pointed out that in order to obtain a strictly increasing degradation with models based on diffusion processes, one must consider $(n+1)$-dimensional degenerate diffusion processes defined by a system of stochastic differential equations of the following form (see Øksendal (2003)):

$$\mathrm{d}X(t) = \rho[X(t),Y_1(t),\ldots,Y_n(t)]\,\mathrm{d}t, \tag{1}$$

$$\mathrm{d}Y_i(t) = f[X(t),Y_1(t),\ldots,Y_n(t)]\,\mathrm{d}t + \sigma[X(t),Y_1(t),\ldots,Y_n(t)]\,\mathrm{d}B_i(t) \tag{2}$$

for $i=1,2,\ldots,n$, where $Y_1(t),\ldots,Y_n(t)$ are variables that influence the wear or degradation $X(t)$, ρ is a function that is always positive, and $B_1(t),\ldots,B_n(t)$ are independent standard Brownian motions. If $X(t)$ represents the remaining lifespan, rather than the degradation, of the device at time t, then the function ρ should be always negative.

The author has written a number of papers on the optimal control of *wear processes*; see, in particular, Lefebvre and Gaspo (1996a) and Lefebvre and Gaspo (1996b), and Lefebvre (2000). He also computed in Lefebvre (2010) the mean first-passage time to zero for wear processes.

In this paper, we will propose in Section 2 two degenerate two-dimensional diffusion processes that could be used to model the remaining lifespan (or degradation) of various devices. Then, in Section 3, optimal control problems for such processes will be set up and solved explicitly. In these problems, the optimizer tries to maximize a certain performance criterion from the initial time $t_0 = 0$ until the time τ when the device is considered to be worn out. This time τ is a random variable.

Remark. The function ρ must be strictly increasing (or decreasing) in the interval $[0,\tau]$, not necessarily for any t. Moreover, if the probability that the remaining lifespan $X(t)$ will increase (or the wear will decrease) in $[0,\tau]$ is negligible, then the proposed model is considered to be acceptable in practice.

Optimal control problems for which the final time is a random variable were called *homing problems* by Whittle (1982). He also considered the case when the cost criterion is risk-sensitive; Whittle (1990).

Finally, we will make a few concluding remarks in Section 4.

2 Integrated diffusion processes as models for the remaining lifespan

First, we recall an important result on n-dimensional stochastic processes; see Lefebvre (2007).

Proposition 1. *Let* $\{\mathbf{X}(t), t \geq 0\}$ *be an* n*-dimensional stochastic process defined by*

$$\mathrm{d}\mathbf{X}(t) = (\mathbf{A}\mathbf{X}(t) + \mathbf{a})\,\mathrm{d}t + \mathbf{N}^{1/2}\,\mathrm{d}\mathbf{B}(t),$$

where $\{\mathbf{B}(t), t \geq 0\}$ *is an* n*-dimensional standard Brownian motion,* $\mathbf{A}$ *is a square matrix of order* n*,* $\mathbf{a}$ *is an* n*-dimensional vector, and* $\mathbf{N}^{1/2}$ *is a positive definite square matrix of order* n*. Then, given that* $\mathbf{X}(t_0) = \mathbf{x}$*, we may write that*

$$\mathbf{X}(t) \sim \mathrm{N}\left(\mathbf{m}(t), \mathbf{K}(t)\right) \quad \textit{for } t \geq t_0,$$

where

$$\mathbf{m}(t) := \Phi(t) \left(\mathbf{x} + \int_{t_0}^{t} \Phi^{-1}(u)\, \mathbf{a}\, du \right)$$

and

$$\mathbf{K}(t) := \Phi(t) \left(\int_{t_0}^{t} \Phi^{-1}(u) \mathbf{N} [\Phi^{-1}(u)]' \, du \right) \Phi'(t),$$

where the symbol prime *denotes the* transpose *of the matrix, and the function* $\Phi(t)$ *is given by*

$$\Phi(t) := e^{\mathbf{A}(t-t_0)} = \sum_{n=0}^{\infty} \mathbf{A}^n \frac{(t-t_0)^n}{n!}.$$

In two dimensions, the system (1), (2) can be written as follows:

$$\begin{aligned} \mathrm{d}X(t) &= \rho[X(t), Y(t)]\,\mathrm{d}t, & (3) \\ \mathrm{d}Y(t) &= f[X(t), Y(t)]\,\mathrm{d}t + \sigma[X(t), Y(t)]\,\mathrm{d}B(t). & (4) \end{aligned}$$

We will now give two examples of degenerate two-dimensional diffusion processes that could serve as models for the remaining lifespan $X(t)$ of devices, and for which Proposition 1 can be used to obtain the distribution of $X(t)$.

Example 1. Consider first the two-dimensional degenerate diffusion process $(X(t), Y(t))$ defined by

$$\begin{aligned} \mathrm{d}X(t) &= cY(t)\,\mathrm{d}t, \\ \mathrm{d}Y(t) &= \mu\,\mathrm{d}t + \sigma\,\mathrm{d}B(t), \end{aligned} \tag{5}$$

where c is a negative constant and $\{B(t), t \geq 0\}$ is a standard Brownian motion, so that $\{Y(t), t \geq 0\}$ is a Wiener process with drift coefficient $\mu \in \mathbb{R}$ and diffusion coefficient $\sigma > 0$. Moreover $\{X(t), t \geq 0\}$ is an integrated Brownian motion multiplied by c. Assuming that $(X(0), Y(0)) = (x, y)$, we find (see Lefebvre (2007), p. 212) that

$$\mathbf{m}(t) = \begin{bmatrix} x + cyt + \frac{1}{2}c\mu t^2 \\ y + \mu t \end{bmatrix}$$

and

$$\mathbf{K}(t) = \begin{bmatrix} c^2\sigma^2 t^3/3 & c\sigma^2 t^2/2 \\ c\sigma^2 t^2/2 & \sigma^2 t \end{bmatrix}.$$

Now, in theory, the Wiener process $\{Y(t), t \geq 0\}$ could become negative before $X(t) = 0$, so that the remaining lifespan would start to increase. Since $Y(t) \sim \mathrm{N}(y + \mu t, \sigma^2 t)$, we have

$$p_1(t) := P[Y(t) < 0] = \Phi\left(\frac{-y - \mu t}{\sigma \sqrt{t}} \right),$$

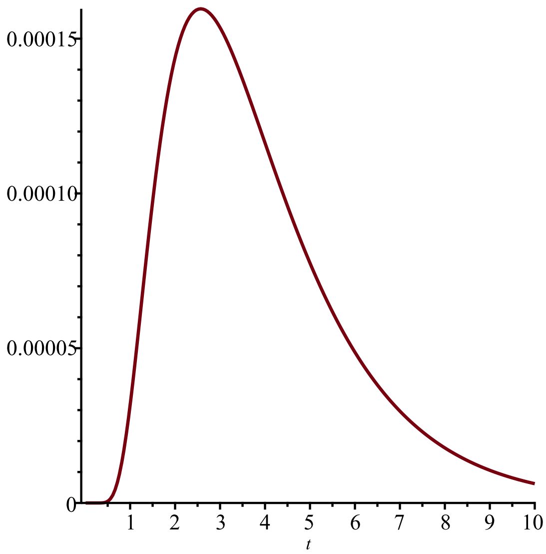


Figure 1: Value of $p_1(t) := \Phi\left(\frac{-3-t}{\sqrt{t}}\right)$ in the interval [0,10].

where here $\Phi(\cdot)$ denotes the distribution function of a $\mathrm{N}(0,1)$ random variable. Hence, if y is large enough and $\mu > 0$, the above probability becomes rapidly negligible. For instance, if $y = 3$ and $\mu = \sigma = 1$, we see in Figure 1 that $p_1(t) := \Phi\left(\frac{-3-t}{\sqrt{t}}\right)$ is smaller than approximately 0.00016. Moreover, since $Y(t)$ has a Gaussian distribution, the parameters μ and σ are easy to estimate.

Remark. To obtain an appropriate model for the degradation of the device, one has simply to assume that the constant c in Eq. (5) is positive instead of negative.

Example 2. Next, we consider the degenerate two-dimensional diffusion process $\mathbf{Z}(t) = (X(t), Y(t))$ defined by the system of stochastic differential equations

$$\begin{aligned} \mathrm{d}X(t) &= cY(t)\,\mathrm{d}t, \\ \mathrm{d}Y(t) &= -\alpha Y(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t), \end{aligned}$$

where $c < 0$ and $\sigma > 0$ are constants, $\alpha \neq 0$ and $\{B(t), t \geq 0\}$ is a standard Brownian motion. Hence, if $\alpha > 0$, $\{Y(t), t \geq 0\}$ is an Ornstein-Uhlenbeck process and $\{X(t), t \geq 0\}$ is its integral times the constant c. We suppose that $(X(t_0), Y(t_0)) = (x, y)$, with $x \geq 0$ and $y > 0$.

We have

$$\mathbf{A} = \begin{pmatrix} 0 & c \\ 0 & -\alpha \end{pmatrix}, \quad \mathbf{a} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \mathbf{N}^{1/2} = \begin{pmatrix} 0 & 0 \\ 0 & \sigma \end{pmatrix}.$$

We calculate

$$\mathbf{A}^2 = \begin{pmatrix} 0 & c \\ 0 & -\alpha \end{pmatrix}\begin{pmatrix} 0 & c \\ 0 & -\alpha \end{pmatrix} = \begin{pmatrix} 0 & -c\alpha \\ 0 & \alpha^2 \end{pmatrix},$$

$$\mathbf{A}^3 = \begin{pmatrix} 0 & c \\ 0 & -\alpha \end{pmatrix}\begin{pmatrix} 0 & -c\alpha \\ 0 & \alpha^2 \end{pmatrix} = \begin{pmatrix} 0 & c\alpha^2 \\ 0 & -\alpha^3 \end{pmatrix}$$

and

$$\mathbf{A}^4 = \begin{pmatrix} 0 & c \\ 0 & -\alpha \end{pmatrix} \begin{pmatrix} 0 & c\,\alpha^2 \\ 0 & -\alpha^3 \end{pmatrix} = \begin{pmatrix} 0 & -c\,\alpha^3 \\ 0 & \alpha^4 \end{pmatrix}.$$

Hence, we deduce that we can write that

$$\mathbf{A}^n = \begin{pmatrix} 0 & (-1)^{n-1}\,c\,\alpha^{n-1} \\ 0 & (-1)^n\,\alpha^n \end{pmatrix} \quad \text{for } n \in \mathbb{N}.$$

It follows that

$$\begin{aligned}
\Phi(t) &= \sum_{n=0}^{\infty} \mathbf{A}^n \frac{(t-t_0)^n}{n!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sum_{n=1}^{\infty} \begin{pmatrix} 0 & (-1)^{n-1}\,c\,\alpha^{n-1} \\ 0 & (-1)^n\,\alpha^n \end{pmatrix} \frac{(t-t_0)^n}{n!} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & c\,(t-t_0)\,e^{-\alpha(t-t_0)} \\ 0 & e^{-\alpha(t-t_0)} - 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & c\,(t-t_0)\,e^{-\alpha(t-t_0)} \\ 0 & e^{-\alpha(t-t_0)} \end{pmatrix}.
\end{aligned}$$

Next,

$$\begin{aligned}
\Phi^{-1}(t) &= e^{\alpha(t-t_0)} \begin{pmatrix} e^{-\alpha(t-t_0)} & -c\,(t-t_0)\,e^{-\alpha(t-t_0)} \\ 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -c\,(t-t_0) \\ 0 & e^{\alpha(t-t_0)} \end{pmatrix}.
\end{aligned}$$

We have

$$\mathbf{m}(t) = \Phi(t) \left[\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right] = \begin{pmatrix} x + c\,(t-t_0)\,e^{-\alpha(t-t_0)}\,y \\ e^{-\alpha(t-t_0)}\,y \end{pmatrix}$$

and

$$\begin{aligned}
&\Phi^{-1}(u)\,\mathbf{N}\,[\Phi^{-1}(u)]' \\
&= \begin{pmatrix} 1 & -c\,(u-t_0) \\ 0 & e^{\alpha(u-t_0)} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & \sigma^2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c\,(u-t_0) & e^{\alpha(u-t_0)} \end{pmatrix} \\
&= \begin{pmatrix} 0 & -c\,\sigma^2\,(u-t_0) \\ 0 & \sigma^2\,e^{\alpha(u-t_0)} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -c\,(u-t_0) & e^{\alpha(u-t_0)} \end{pmatrix} \\
&= \begin{pmatrix} c^2\,\sigma^2\,(u-t_0)^2 & -c\,\sigma^2\,(u-t_0)\,e^{\alpha(u-t_0)} \\ -c\,\sigma^2\,(u-t_0)\,e^{\alpha\,(u-t_0)} & \sigma^2\,e^{2\alpha(u-t_0)} \end{pmatrix},
\end{aligned}$$

so that we obtain the symmetric matrix

$$\begin{aligned}
\mathbf{M}(t) &:= \int_{t_0}^{t} \Phi^{-1}(u)\,\mathbf{N}\,[\Phi^{-1}(u)]'\,\mathrm{d}u \\
&= \begin{pmatrix} c^2\,\sigma^2\,\dfrac{(t-t_0)^3}{3} & -c\,\sigma^2\,\dfrac{1 + [\alpha\,(t-t_0) - 1]\,e^{\alpha(t-t_0)}}{\alpha^2} \\ — & \sigma^2\,\dfrac{e^{2\alpha(t-t_0)} - 1}{2\,\alpha} \end{pmatrix}.
\end{aligned}$$

Finally, we calculate

$$\begin{aligned}\mathbf{K}(t) &:= \Phi(t)\,\mathbf{M}(t)\,\Phi'(t)\\ &= \begin{pmatrix} 1 & c\,(t-t_0)\,e^{-\alpha(t-t_0)} \\ 0 & e^{-\alpha(t-t_0)} \end{pmatrix} \mathbf{M}(t) \begin{pmatrix} 1 & 0 \\ c\,(t-t_0)\,e^{-\alpha(t-t_0)} & e^{-\alpha(t-t_0)} \end{pmatrix}\\ &= \begin{pmatrix} K_{11} & K_{12} \\ K_{12} & K_{22} \end{pmatrix},\end{aligned}$$

where

$$\begin{aligned}K_{11} &:= \frac{c^2\,\sigma^2\,(t-t_0)}{3\,\alpha^2}\Big[6-(9/2)\,\alpha\,(t-t_0)+\alpha^2\,(t-t_0)^2\\ &\qquad -6\,e^{-\alpha(t-t_0)}-(3/2)\,(t-t_0)\,\alpha\,e^{-2\alpha(t-t_0)}\Big],\end{aligned}$$

$$K_{12} := \frac{c\,\sigma^2}{2\,\alpha^2}\left[2-\alpha\,(t-t_0)-2\,e^{-\alpha(t-t_0)}-\alpha\,(t-t_0)\,e^{-2\alpha(t-t_0)}\right]$$

and

$$K_{22} := \frac{\sigma^2}{2\,\alpha}\left[1-e^{-2\alpha(t-t_0)}\right].$$

The probability

$$p_2(t) := \Phi\left(\frac{-e^{-\alpha t}\,y}{\frac{\sigma}{\sqrt{2\,\alpha}}\sqrt{1-e^{-2\alpha t}}}\right) \tag{6}$$

that $Y(t)<0$ is negligible if $\alpha<0$, as can be seen in Figure 2. However, if $\alpha>0$, $Y(0)=y$ must be large for $p_2(t)$ to be very small. Finally, since $Y(t)$ has again a Gaussian distribution, the parameters α and σ can be easily estimated.

In the next section, optimal control problems for degenerate two-dimensional diffusion processes that can be used to model the evolution of the remaining lifespan of devices over time will be set up and solved explicitly.

3 Optimal control

In this section, we consider controlled versions of the two-dimensional diffusion processes defined in the previous section (see Rishel (1991)):

$$\begin{aligned}\mathrm{d}X_u(t) &= \rho\,[X_u(t),Y_u(t),u(t)]\,\mathrm{d}t,\\ \mathrm{d}Y_u(t) &= f\,[X_u(t),Y_u(t),u(t)]\,\mathrm{d}t+\sigma[X_u(t),Y_u(t),u(t)]\,\mathrm{d}B(t),\end{aligned}$$

where $u(t)\,(=u[X_u(t),Y_u(t)])$ is the control variable, which is assumed to be a continuous function.

We define the *first-passage time*

$$\tau(x,y)=\inf\{t>0 : X(t)=d \mid X(0)=x, Y(0)=y\},$$

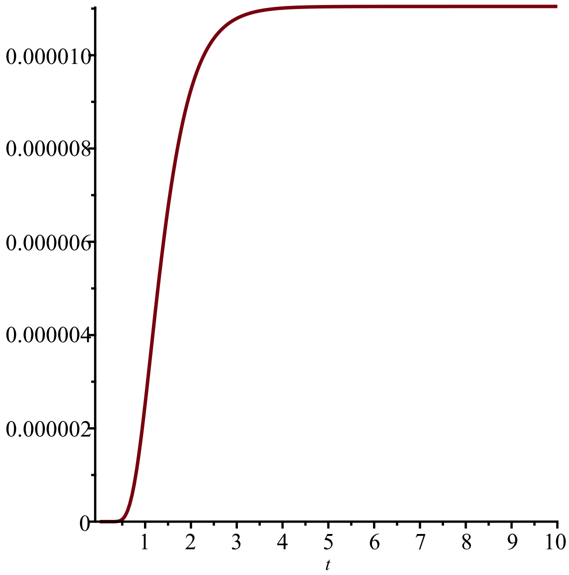


Figure 2: Value of the function $p_2(t)$ defined in Eq. (6) in the interval [0,10], when $\alpha = -1$, $\sigma = 1$ and $y = 3$.

where $x > d$ and $y > 0$, and $d \geq 0$ is a value for which the device of interest is considered to be worn out. The random variable $\tau(x,y)$ represents the remaining useful lifetime of the device, denoted by RUL.

We look for the value $u^*(t)$ of the control variable that maximizes the expected value of the performance criterion

$$J(x,y) := \int_0^{\tau(x,y)} \{r[u(t)] - c[X_u(t), Y_u(t)]\} \, \mathrm{d}t,$$

where $r[u(t)] > 0$ and $c[X_u(t), Y_u(t)] > 0$ are respectively the return and the cost per unit time.

To find the optimal control $u^*(t)$, we will use *dynamic programming*. First, we define the *value function*:

$$F(x,y) = \sup_{u(t),\, 0 \leq t < \tau(x,y)} E[J(x,y)]. \tag{7}$$

That is, $F(x,y)$ is the expected reward (or expected cost, if it is negative) obtained when the optimizer chooses the optimal value of $u(t)$ in the interval $[0, \tau(x,y))$.

Let $u(0) = u_0$. Making use of Bellman's principle of optimality, we can write that

$$\begin{aligned} F(x,y) &= \sup_{u(t),0\le t<\tau(x,y)} E\left[\int_0^{\Delta t} \{r[u(t)] - c[X_u(t),Y_u(t)]\}\,\mathrm{d}t \right. \\ &\quad \left. + F\left[x+\rho(x,y,u_0)\Delta t, y+f(x,y,u_0)\Delta t+\sigma(x,y,u_0)B(\Delta t)\right]\right] \\ &\quad + o(\Delta t) \\ &= \sup_{u(t),0\le t\le\Delta t} E\left[[r(u_0)-c(x,y)]\,\Delta t \right. \\ &\quad \left. + F\left[x+\rho(x,y,u_0)\Delta t, y+f(x,y,u_0)\Delta t+\sigma(x,y,u_0)B(\Delta t)\right]\right] \\ &\quad + o(\Delta t). \end{aligned} \tag{8}$$

Next, assuming that the value function $F(x,y)$ is differentiable with respect to x and twice differentiable with respect to y, we deduce from Taylor's formula for functions of two variables that

$$\begin{aligned} &F\left[x+\rho(x,y,u_0)\Delta t, y+f(x,y,u_0)\Delta t+\sigma(x,y,u_0)B(\Delta t)\right] \\ &= F(x,y)+\rho(x,y,u_0)\Delta t\,F_x(x,y) \\ &\quad + [f(x,y,u_0)\Delta t+\sigma(x,y,u_0)B(\Delta t)]\,F_y(x,y) \\ &\quad + \tfrac{1}{2}\,[f(x,y,u_0)\Delta t+\sigma(x,y,u_0)B(\Delta t)]^2\,F_{yy}(x,y)+o(\Delta t). \end{aligned}$$

Since, as is well known, $E[B(\Delta t)] = 0$ and $E[B^2(\Delta t)] = V[B(\Delta t)] = \Delta t$, Eq. (8) implies that

$$\begin{aligned} 0 &= \sup_{u(t),0\le t\le\Delta t} \left\{[r(u_0)-c(x,y)]\Delta t+\rho(x,y,u_0)\Delta t\,F_x(x,y) \right. \\ &\qquad \left. + f(x,y,u_0)\Delta t\,F_y(x,y)+\tfrac{1}{2}\sigma^2(x,y,u_0)\Delta t\,F_{yy}(x,y)\right\} \\ &\qquad + o(\Delta t). \end{aligned} \tag{9}$$

Finally, dividing both sides of Eq. (9) by Δt and letting Δt decrease to zero, we obtain the following proposition.

Proposition 2. *The value function $F(x,y)$ defined in Eq. (7) satisfies the* dynamic programming equation *(DPE)*

$$\begin{aligned} 0 &= \sup_{u_0} \left\{ r(u_0)-c(x,y)+\rho(x,y,u_0)F_x(x,y)+f(x,y,u_0)F_y(x,y) \right. \\ &\qquad \left. + \tfrac{1}{2}\sigma^2(x,y,u_0)F_{yy}(x,y)\right\}, \end{aligned} \tag{10}$$

subject to the boundary condition

$$F(d,y) = 0 \quad \forall y > 0. \tag{11}$$

Rishel (1991) has considered the following particular cases for the optimal control of wear processes: first, he assumed that

$$\left.\begin{array}{rcl} \rho[X_u(t),Y_u(t),u(t)] &=& \rho_0[X_u(t),Y_u(t)]\,u^2(t), \\ f[X_u(t),Y_u(t),u(t)] &=& f_0[X_u(t),Y_u(t)]\,u^2(t), \\ \sigma[X_u(t),Y_u(t),u(t)] &=& \sigma_0[X_u(t),Y_u(t)]\,u(t), \\ r[u(t)] &=& r_0\,u(t), \end{array}\right\}$$

where r_0 is positive constant. He proved that the optimal control can then be expressed as follows:

$$u_0^* = u^*(x,y) = \frac{2\,c(x,y)}{r}.$$

Next, he chose

$$\left.\begin{array}{rcl} \rho[X_u(t),Y_u(t),u(t)] &=& \rho_0[X_u(t),Y_u(t)]\,u(t), \\ f[X_u(t),Y_u(t),u(t)] &=& f_0[X_u(t),Y_u(t)]\,u(t), \\ \sigma[X_u(t),Y_u(t),u(t)] &=& \sigma_0[X_u(t),Y_u(t)]\,u(t), \\ r[u(t)] &=& r_0\,u(t). \end{array}\right\}$$

Substituting these functions into the DPE (10), we obtain (after differentiating with respect to u_0) that the optimal control is given, in terms of the value function, by

$$u^*(x,y) = -\frac{\rho_0(x,y)\,F_x + f_0(x,y)\,F_y + r_0}{\sigma_0^2(x,y)F_{yy}}.$$

Remark. There is a minus sign missing in Eq. (26) of Rishel's paper.

It follows that the value function satisfies the partial differential equation (PDE)

$$[\rho_0(x,y)\,F_x + f_0(x,y)\,F_y + r_0]^2 + 2\,c(x,y)\,\sigma_0^2(x,y)\,F_{yy} = 0, \tag{12}$$

subject to the boundary condition (11). This time, Rishel did not find an explicit expression for $u^*(x,y)$, but Lefebvre and Gaspo (1996a) first generalized Rishel's result and then solved three particular problems.

In this paper, we will solve three new problems for important diffusion processes $\{Y(t), t \geq 0\}$ when $X(t)$ represents the remaining lifespan of a device.

Problem I. Suppose that $f_0(x,y) \equiv f_{00}$ and $\sigma_0^2(x,y) = \sigma_{00}^2\,y^{\nu}$, where $\sigma_{00} > 0$ and $\nu \in \{0,1,2\}$. Then, Eq. (4) becomes

$$\mathrm{d}Y(t) = f_{00}\,\mathrm{d}t + \sigma_{00}\,Y^{\nu/2}(t)\,\mathrm{d}B(t).$$

If $\nu = 0$, $\{Y(t), t \geq 0\}$ is a Wiener process with drift f_{00} and dispersion parameter σ_{00}. As we have seen in Section 2, although this diffusion process is Gaussian, the probability that it will become negative is negligible if f_{00} is large enough and $Y(0) = y$ is not too small.

When $\nu = 1$ (and $\sigma_{00}^2 = 2$), $\{Y(t), t \geq 0\}$ is a *squared Bessel process of dimension* $\delta = f_{00}$. It can be shown that if $\delta \geq 2$ and $y > 0$, then the process cannot attain the origin. Therefore, it is a good model for the remaining lifespan (or the wear) of a device.

In the case when $\nu = 2$, we have the following proposition.

Proposition 3. *The diffusion process $\{Y(t), t \geq 0\}$ with infinitesimal mean $f_{00} > 0$ and infinitesimal variance $\sigma_{00}^2 y^2$ has an* exit boundary *at the origin. That is, starting from $Y(0) = y > 0$, the process can reach the origin in finite time and, if it does, it will remain equal to zero indefinitely.*

Proof. We define (see Eq. (9.7.19) in Kannan (1979), p. 279, but there is a mistake in the equation)

$$h(y) = \exp\left\{\int_a^y 2\frac{f_{00}}{\sigma_{00}^2 u^2}\,\mathrm{d}u\right\}$$

and

$$H(y) = \frac{2}{\sigma_{00}^2 y^2 h(y)}.$$

We calculate

$$\mu_1 := \int_0^a \int_z^a h(y) H(z)\,\mathrm{d}y\,\mathrm{d}z$$

and

$$\sigma_1 := \int_0^a \int_z^a H(y) h(z)\,\mathrm{d}y\,\mathrm{d}z.$$

We find that $\mu_1 = \infty$ and $\sigma_1 < \infty$. Then, we can state that the origin is indeed an exit boundary. ■

Assume that

$$\rho_0(x,y) = -\rho_{00}\frac{(x-d)}{y},$$

where $\rho_{00} > 0$. We must then solve the PDE

$$\left[-\rho_{00}\frac{(x-d)}{y}F_x + f_{00}F_y + r_0\right]^2 + 2c(x,y)\,\sigma_{00}^2 y^{\nu} F_{yy} = 0, \tag{13}$$

subject to the boundary condition $F(d,y) = 0$ for any $y > 0$.

Remark. We will see that with the function $\rho_0(x,y)$ defined above, the optimally controlled process $X_u^*(t)$ will actually decrease immediately to d if $Y_u^*(t)$ decreases to zero. Hence, here the type of boundary at the origin for the diffusion process $\{Y(t), t \geq 0\}$ does not really matter.

Let us look for a solution of the form

$$F(x,y) = F_0\left[e^{(x-d)y} - 1\right] \quad \text{for } x \geq d,\ y > 0, \tag{14}$$

which satisfies the boundary condition $F(d,y) = 0$. Substituting $F(x,y)$ into Eq. (13), we obtain the following proposition.

Proposition 4. *If $\rho_{00} = f_{00}$ and the function $c(x,y)$ is given by*

$$c(x,y) = -\frac{r_0^2 e^{-(x-d)y}}{2F_0\,\sigma_{00}^2 y^{\nu}(x-d)^2}, \tag{15}$$

where $F_0 < 0$, then the function $F(x,y)$ defined in Eq. (14) is the value function in Problem I.

Remark. Notice that the value function is negative. Thus, with the cost $c(x,y)$ defined in Eq. (15), the best that the optimizer can do is to minimize the losses.

We deduce from Eq. (3) the value of the optimal control.

Corollary 1. *The optimal control in Problem I is*

$$u^*(x,y) = -\frac{r_0\, e^{-(x-d)y}}{F_0\, \sigma_{00}^2\, y^{\nu}\,(x-d)^2}. \tag{16}$$

Remarks. (i) The optimal control is positive, so that the optimizer does not try to make $\rho[X_u(t),Y_u(t)]$ become positive.

(ii) We have

$$\begin{aligned} \mathrm{d}X_u^*(t) &= \rho[X_u^*(t),Y_u^*(t),u^*(t)]\,\mathrm{d}t \\ &= -\rho_{00}\,\frac{[X_u^*(t)-d]}{Y_u^*(t)}\,u^*[X_u^*(t),Y_u^*(t)]\,\mathrm{d}t \\ &= \rho_{00}\,\frac{r_0\,\exp\{[d-X_u^*(t)]\,Y_u^*(t)\}}{F_0\,\sigma_{00}^2\,[Y_u^*(t)]^{\nu+1}\,[X_u^*(t)-d]}\,\mathrm{d}t. \end{aligned}$$

Therefore, as mentioned above, the optimally controlled process $X_u^*(t)$ will decrease at once to d if $Y_u^*(t)$ decreases to zero.

Problem II. Assume now that $f_0(x,y) = f_{00}\,(\kappa - y)$, where f_{00} and κ are real parameters, and $\sigma_0^2(x,y) = \sigma_{00}^2\, y$. With these choices, Eq. (4) becomes

$$\mathrm{d}Y(t) = f_{00}\,[\kappa - Y(t)]\,\mathrm{d}t + \sigma_{00}\,Y^{1/2}(t)\,\mathrm{d}B(t).$$

Therefore, $\{Y(t), t \geq 0\}$ is a (generalized) Cox-Ingersoll-Ross (CIR) process. This process is used in financial mathematics as a model for the evolution of interest rates. When

$$2\,f_{00}\,\kappa \geq \sigma_{00}^2, \tag{17}$$

the process, starting from $Y(0) = y > 0$, will always remain positive.

The transition density function of $\{Y(t), t \geq 0\}$ is known to be

$$p(y,t;y_0,t_0) = \gamma e^{-u-v}\left(\frac{v}{u}\right)^{q/2} I_q(2\sqrt{uv}),$$

where

$$\gamma := \frac{2\,f_{00}}{\left(1-e^{-f_{00}(t-t_0)}\right)\sigma_{00}^2}, \quad u := \gamma y_0\, e^{-f_{00}(t-t_0)}, \quad v := \gamma y, \quad q := \frac{2\,f_{00}\,\kappa}{\sigma_{00}^2} - 1$$

and $I_q(\cdot)$ is a modified Bessel function of the first kind of order q (see Abramowitz and Stegun (1965)).

Assume that both f_{00} and κ are negative, and that the condition in Eq. (17) holds. We try a solution of the PDE (12) of the same form as in Problem I: $F(x,y) = F_0\left[e^{(x-d)y} - 1\right]$. One can check that this function is indeed a solution of Eq. (12) that satisfies the boundary condition $F(d,y) = 0$ if and only if

$$\rho_0(x,y) = -f_{00}\,(\kappa - y)\frac{(x-d)}{y} \quad (<0)$$

and $c(x,y)$ is the same as in Eq. (15), with again $F_0 < 0$. Moreover, the optimal control is given by the function defined in Eq. (16), with $\nu = 1$.

Problem III. In this last problem, we now assume that $X(t)$ represents the wear of the device, rather than its remaining lifespan. Therefore, the function $\rho_0(x,y)$ should be positive. Moreover, we suppose that $X(0) = x \in (0,d)$. The differential equation that we must solve is the same PDE as above, and the boundary condition is still $F(d,y) = 0$.

Let us choose

$$\rho_0(x,y) = \frac{d-x}{xy}, \quad f_0(x,y) = \frac{1}{x} \quad \text{and} \quad \sigma_0(x,y) \equiv 1.$$

We then find that the function

$$F(x,y) = F_0\left[e^{(x-d)y} - 1\right] \quad \text{for } 0 < x \le d,\ y > 0$$

is a solution of Eq. (12) that satisfies $F(d,y) = 0$ if and only if

$$c(x,y) = -\frac{r_0^2\, e^{(d-x)y}}{2F_0\,(d-x)^2},$$

where F_0 is negative. Furthermore, the optimal control is

$$u^*(x,y) = -\frac{r_0\, e^{(d-x)y}}{F_0\,(d-x)^2}.$$

Remarks. (i) In this problem, the value function is positive. Therefore, it is a reward, rather than a cost as in the previous problems.

(ii) The diffusion process $\{Y(t), t \ge 0\}$ is a Bessel process of *dimension* or *parameter* $\alpha = 3$. For Bessel processes of dimension $\alpha \ge 2$, the origin is an *entrance boundary*, which means that if $X(0) = x > 0$, the process will never hit the origin. The transition density function of $\{Y(t), t \ge 0\}$ is (see Karlin and Taylor (1981), p. 368)

$$p(y,t;y_0,0) = \sqrt{\frac{2}{\pi t}}\exp\left\{-\frac{y_0^2+y^2}{2t}\right\}\frac{y}{y_0}\sinh\left(\frac{y_0 y}{t}\right)$$

for $t > 0$, $y_0 > 0$ and $y > 0$.

4 Concluding remarks

In this paper, we proposed models for the evolution of the remaining lifespan $X(t)$ of devices that are more realistic than one-dimensional diffusion processes. Indeed, the degenerate two-dimensional processes defined in the paper were such that $X(t)$ was a strictly decreasing function of time, as should be. Actually, in Section 2, we mentioned the fact that there was a positive probability that $X(t)$ would increase before reaching the value $d \geq 0$ for which the device is worn out. However, we saw that in practice this probability can be really negligible. Moreover, the aim in Section 2 was to give examples for which it is possible to compute the exact distribution of $X(t)$. This is possible, thanks to Proposition 1, when the diffusion process $\{Y(t), t \geq 0\}$ is Gaussian and the function $\rho[X(t), Y(t)]$ in Eq. (3) is affine.

In Section 3, we presented three optimal control problems for degenerate two-dimensional diffusion processes that could serve as models for the remaining lifespan of devices. These problems are particular homing problems. Obtaining exact and explicit solutions to such problems in two or more dimensions is generally a very difficult task. In Lefebvre and Gaspo (1996a), the authors used the method of similarity solutions to reduce the PDE that must be solved to an *ordinary* differential equation. Here, we proposed a simple function for the value function $F(x, y)$, and we saw that this simple function was indeed the exact solution to the problem considered, if certain conditions hold. Furthermore, the three problems solved in Section 3 were for diffusion processes $\{Y(t), t \geq 0\}$ that are very important for the applications, whereas the ones solved in Lefebvre and Gaspo (1996a) were purely mathematical problems.

When it is not possible to find the exact solution to a particular optimal control problem, one may try to make use of numerical methods to obtain the value function and the optimal control in any special case. It is also sometimes possible to determine upper bounds for both the value function and the optimal control, as Makasu (2022) did for certain homing problems.

Acknowledgments
This work was supported by the Natural Sciences and Engineering Research Council of Canada. The authors also wish to thank the anonymous reviewers of this paper for their constructive comments.

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.

Ghamlouch, H., Grall, A. and Fouladirad, M. (2015). On the use of jump-diffusion process for maintenance decision-making: A first step. In *2015 Annual Reliability and Maintainability Symposium (RAMS)*, 1–6. DOI: 10.1109/RAMS.2015.7105099.

Kannan, D. (1979). *An Introduction to Stochastic Processes*. North Holland, New York.

Karlin, S. and Taylor, H. M. (1981). *A Second Course in Stochastic Processes*. Academic Press, New York.

Lefebvre, M. and Gaspo, J. (1996a). Optimal control of wear processes. *IEEE Transactions on Automatic Control*, **41**, 112–115.

Lefebvre, M. and Gaspo, J. (1996b). Controlled wear processes: a risk-sensitive formulation. *Engineering Optimization*, **26**, 187–194.

Lefebvre, M. (2000). Risk-averse control of a three-dimensional wear process. *International Journal of Control*, **73**, 1307–1311.

Lefebvre, M. (2007). *Applied Stochastic Processes.* Springer, New York.

Lefebvre, M. (2010). Mean first-passage time to zero for wear processes. *Stochastic Models*, **26**, 46–53.

Makasu, C. (2022). Homing problems with control in the diffusion coefficient. *IEEE Transactions on Automatic Control*, **67**, 3770–3772.

Nicolai, R. P. and Dekker, R. (2007). A comparison of two non-stationary degradation processes. In: Aven, T. and Vinnem, J. E. (eds.) *Risk, Reliability and Societal Safety: Proceedings ESREL 2007, 25-27 June, Stavanger*, Taylor & Francis, London, 1109–1114.

Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications.* 6th Ed., Springer-Verlag, Berlin.

Rishel, R. (1991). Controlled wear processes: modeling optimal control. *IEEE Transactions on Automatic Control*, **36**, 1100–1102.

Shahraki, A. F., Yadav, O. P. and Liao, H. (2017). A review on degradation modelling and its engineering applications. *International Journal of Performability Engineering*, **13**, 299–314.

Whittle, P. (1982). *Optimization Over Time I.* Wiley, Chichester.

Whittle, P, (1990) *Risk-Sensitive Optimal Control.* Wiley, Chichester.

Whitmore, G. A. and Schenkelberg, F, (1997). Modelling accelerated degradation data using Wiener diffusion with a time scale transformation. *Lifetime Data Analysis*, **3**, 27–45.

Ye, Z.-S., Wang, Y., Tsui, K.-L. and Pecht, M. (2013). Degradation data analysis using Wiener processes with measurement errors. *IEEE Transactions on Reliability*, **62**, 772–780. DOI: 10.1109/TR.2013.2284733

Zhai, Q., Chen, P., Hong, L. and Shen, L. (2018) A random-effects Wiener degradation model based on accelerated failure time. *Reliability Engineering & System Safety*, **180**, 94–103.

Zhang, Z., Hu, C., Si, X., Zhang, J. and Zheng, J. (2017) Stochastic degradation process modeling and remaining useful life estimation with flexible random-effects. *Journal of the Franklin Institute*, **354**, 2477–2499.

Zhou, S., Tang, Y. and Xu, A. (2021). A generalized Wiener process with dependent degradation rate and volatility and time-varying mean-to-variance ratio. *Reliability Engineering & System Safety*, **216**, 107895.

End of SMPS Journal – Volume 1, Issue 1

For manuscript submission and future issues, visit:
www.smps.um.ac.ir