ORIGINAL PAPER

On connections between skewed, weighted and distorted distributions: applications to model extreme value distributions

Jorge Navarro[1] · Jorge M Arevalillo[2,3]

Abstract

The purpose of the paper is to explore the connections between skew symmetric, weighted and distorted univariate distributions as well as how they appear related to the distributions of the extreme values in a sample of identically distributed random variables under both the independence and dependence scenarios. Some extensions of the concept of skewed distributions are proposed in order to cover the most general cases of extremes. Their natural connections to the likelihood ratio ordering and the role played by the P–P plots for handling these models are also highlighted. The results can also be applied to order statistics and coherent systems although these cases do not always lead to skewed distributions. The theoretical findings are illustrated by applied examples to model extremes as well as by several applications concerned with the analysis of artificial and real data.

Keywords Skewness · Distortions · Order statistics · Copula

Mathematics Subject Classification 60E15 · 62N05

1 Introduction

The normal (or Gaussian) model is closed under operations such as linear combinations, marginalization and conditioning. However, it is not closed under the extreme value operator: for example, if $X_1, \ldots, X_n$ is a sample of independent and identically

✉ Jorge Navarro
  jorgenav@um.es

[1] Dpto. Estadística e Investigación Operativa, Facultad de Matemáticas, Universidad de Murcia, 30100 Murcia, Spain

[2] UC3M-Santander Big Data Institute, Madrid Street 135, 28903 Getafe, Madrid, Spain

[3] Department of Statistics and Operational Research, University Nacional Educación a Distancia (UNED), Juan del Rosal 10, 28040 Madrid, Spain

distributed (IID) normal variables then $X_{n:n} = \max(X_1, \ldots, X_n)$ does not follow a normal distribution. The same thing happens if we consider $X_{1:n} = \min(X_1, \ldots, X_n)$ or if the random variables $X_1, \ldots, X_n$ are dependent, where the dependence can be modeled by copulas. This kind of data arise in areas like reliability, climatology and environmental studies (Blenkinsop et al. 2017; Buishand 1989; Ghosh and Ng 2019; Navarro 2022) just to name a few fields fertile for data analysis.

It is well-known that the models for the maximum and minimum are no longer normal since their distributions correspond to asymmetric models that were used to define the univariate *skew normal distribution*, see Roberts (1966) and Azzalini (1985). This model assumes a movement of the probability mass to the right ($X_{n:n}$) or to the left ($X_{1:n}$) which results in asymmetric departures from normality. This property was extended to the multivariate case and to other (non-normal) models by using *skew symmetric distributions* (Arellano-Valle and Genton 2008; Azzalini 2005; Azzalini and Capitanio 2003) which are useful to model the asymmetry arising for example in finance data (De Luca and Loperfido 2004). Other related distributions were studied in Ferreira and Steel (2006), Jones (2015) and Ley (2015).

On the other hand, the *weighted distributions* introduced by Fisher (1934), Rao (1965) and Patil and Rao (1978) are concerned with the modification of a given initial baseline model through a weighted function which handles unequal sampling probabilities; this is actually what happens with the extreme values selected from a set of baseline values using a specific sampling procedure. A relatively recent generalization of weighted distributions, obtained by the extension of Azzalini's skewing approach, has been addressed in Domma et al. (2015).

The third related concept is the so called *distorted distributions*. They were introduced in the theory of choice under risk (Wang 1996; Yaari 1987) to model a change (a distortion) in the baseline distribution for the risks or the claims. They have also been applied to model order statistics and coherent systems (Navarro 2022; Navarro et al. 2018).

In this paper we connect these three concepts in the univariate case showing how they can be used to model the distributions of extremes (minima and maxima). The representations based on weighted and distorted distributions always hold for extremes. In order to establish a unified framework that includes these models and skew distributions, we propose two univariate extensions that assess skewness by means of distributions defined as the product of a baseline probability density function and a distribution (survival) function. We show that both variants —distribution or survival— would cover many extreme related models, with the only limitation of disregarding a few particular cases of dependence models (copulas). The connections of the proposed representations with the likelihood ratio order and PP-plots are also discussed putting the focus on their implications for modeling extremes.

The rest of the paper is organized as follows: The notation, definitions and preliminary results are placed in Sect. 2, which also includes the new extensions of univariate skewed distributions. The main results are given in Sect. 3 which studies different representations of the distributions of extremes; some illustrative examples of these representations are presented in Sect. 4. The numerical work illustrating the results of the theory is provided in Sect. 5; it includes applications with artificial data and a real

data study about daily and monthly maximum temperatures. Finally, Sect. 6 contains the conclusions and some thoughts for future research.

2 Skewed, weighted, distorted distributions and their connections

2.1 Univariate models

Let $X_1, \ldots, X_n$ be a sample of identically distributed (ID) random variables with a common absolutely continuous cumulative distribution function (CDF) F and with probability density function (PDF) $f = F'$ (a.e.). Let $\bar{F} = 1 - F$ be the survival or reliability function and let $X_{1:n} \leq \cdots \leq X_{n:n}$ be the associated order statistics derived from the sample. In particular, the extreme values are $X_{1:n} = \min(X_1, \ldots, X_n)$ and $X_{n:n} = \max(X_1, \ldots, X_n)$. The main properties of order statistics have been studied in Arnold et al. (2008) and David and Nagaraja (2003). In many cases $X_1, \ldots, X_n$ correspond to the lifetimes (or survival times) of some items; so we only observe the first failure time $X_{1:n}$. In some other cases, they correspond to claims, environmental or climatology extreme values and we just observe $X_{1:n}$ or $X_{n:n}$.

Sometimes the assumption of independence of the sample random variables (IID case) is well suited to the sampling procedure. However, in other situations where the observed values share the same environment (or the same risks) such assumption fails; in such cases the sampling procedure responds to the dependence scheme. For both scenarios, the joint distribution function of $(X_1, \ldots, X_n)$ can be represented as

$$\Pr(X_1 \leq x_1, \ldots, X_n \leq x_n) = C(F(x_1), \ldots, F(x_n)),$$

where $C : [0, 1]^n \to [0, 1]$ is a copula function (Durante and Sempi 2016; Nelsen 2006). From Sklar's theorem (Schweizer and Sklar 1974), C is unique provided that F is continuous. Here, we will assume that both C and F are absolutely continuous with PDFs $c = \partial_{1,\ldots,n} C$ and $f = F'$ respectively, where $\partial_i C$ denotes the partial derivative of C with respect to its ith variable, $\partial_{i,j} C$ the second partial derivative of C with respect to its jth and ith variables and so on. The independence is represented by the product copula $C(u_1, \ldots, u_n) = u_1, \ldots, u_n$ with $c(u_1, \ldots, u_n) = 1$ for $u_1, \ldots, u_n \in [0, 1]$. A similar representation holds for the joint survival function

$$\Pr(X_1 > x_1, \ldots, X_n > x_n) = \widehat{C}(\bar{F}(x_1), \ldots, \bar{F}(x_n)),$$

where $\widehat{C} : [0, 1]^n \to [0, 1]$ is a copula function called *survival copula*.

The univariate skew normal (SN) distribution was introduced by Azzalini (1985) to handle asymmetry deviations from normality. Due to its simple analytical form, this distribution has become a widely used model to handle the non-normality of data; its PDF is defined by

$$f_\lambda(x) = 2\phi(x)\Phi(\lambda x) \colon x \in \mathbb{R}, \tag{2.1}$$

where ϕ and Φ are the PDF and CDF of a standard normal variable respectively and $\lambda \in \mathbb{R}$ is a shape parameter that regulates the asymmetry of the model. If $\lambda > 0$ then $\Phi(\lambda x)$ is a distribution function and the density f_λ arises as a deformation of ϕ that results by injecting probability to the right tail of the normal model (i.e. the likelihood of the greater values increases). If $\lambda < 0$ then $\Phi(\lambda x) = 1 - \Phi(-\lambda x)$ is a survival function and f_λ captures the movement of the probability mass to the left, augmenting the likelihood of lower values. In both cases, the model can be formulated to consider location and scale parameters (Azzalini 1985; Azzalini and Capitano 2014). When $\lambda = 0$ the SN reduces to the normal distribution. We will write $X \sim \mathcal{SN}(\lambda)$ to indicate that X follows a SN distribution with PDF given by expression (2.1).

The idea of perturbing the symmetry motivated the extension of the SN model to the class of *skew symmetric distribution* studied in Azzalini and Capitanio (2003), Azzalini (2005), Azzalini and Regoli (2012), see also the term 'perturbation' used in Azzalini and Capitanio (2003), Azzalini and Capitanio (2014) and the term 'modulation of distributions' in Azzalini and Capitanio (2014). The PDF of a skew symmetric scalar variable is defined by

$$f_{G,h}(x) = 2f(x)G(h(x)), \tag{2.2}$$

where f is a symmetric PDF around zero, G is a CDF such that $G(-x) = 1 - G(x)$ and h is a real function such that $h(-x) = -h(x)$ for all x. We put $X \sim \mathcal{SSD}(f, G, h)$ to denote that a random variable X has the PDF in (2.2). Clearly, the SN distribution is obtained when we take $f = \phi$, $G = \Phi$ and $h(x) = \lambda x$. Expression (2.2) can be extended naturally to the multivariate case, see Azzalini and Regoli (2018). For other extensions and their relationships see Jones (2015) and Ley (2015).

A second idea related to the perturbation of a density function is concerned with the well-known *biased or weighted distributions* which can be traced back to Fisher (1934) and Rao (1965). If $w : \mathbb{R} \to \mathbb{R}$ is a nonnegative weight function and f is a PDF then the weighted PDF associated to w and f is defined by

$$f_w(x) = c_w w(x) f(x), \tag{2.3}$$

where we assume that $0 < \mu_w = \int_{\mathbb{R}} w(x)f(x)dx < \infty$ with $c_w = 1/\mu_w$ being the normalizing constant. The weight function w appearing in (2.3) is used to modify the sampling probabilities of the density f. Of course, if w is constant then we will be sampling from the baseline variable with PDF f. However, in other situations the function w will serve to down weight or up weight the probability so that the sampling scheme is modified accordingly. Perhaps, the most typical case is the length biased (or size biased) weight function where $X \geq 0$, $w(x) = x$ for $x \geq 0$ and $\mu_w = \mu = E(X)$ (Patil and Rao 1978). In this case, the sampling probability of a data X_i is proportional to X_i and the baseline probability mass from f is moved to the right to get the PDF f_w; so it is also "right skewed". However, this is not the case for other weight functions which will not necessarily represent skewed distributions. We put $X \sim \mathcal{WD}(f, w)$ to indicate that the a random variable X follows a weighted distribution with PDF given by (2.3).

It is clear that the $\mathcal{SN}(\lambda)$ distribution is a weighted distribution with weight function $w(x) = \Phi(\lambda x)$ and $c_w = 2$. Hence, it can be seen as a model having a probability assignment scheme proportional to $\Phi(\lambda x) = \Pr(X \leq \lambda x)$. If $\lambda > 0$ the probability increases in x; meanwhile, if $\lambda < 0$ then it is a decreasing function of x. In the first case, the higher values are more likely than the smaller ones, and vice versa for the second case. In particular, if $\lambda = 1$ then we get $w(x) = \Phi(x) = \Pr(X \leq x)$ (the standard normal distribution function); on the other hand, when $\lambda = -1$ it is obtained that $w(x) = \Phi(-x) = 1 - \Phi(x) = \Pr(X > x)$ (the standard normal survival function). The same observation applies to the $\mathcal{SSD}(f, G, h)$ family which is a subclass of the class of weighted distributions with weight function $w(x) = G(h(x))$ and $c_w = 2$. The weighted distributions can also be extended to the multivariate case, see e.g. Navarro et al. (2006). For the connections between (2.2) and (2.3) in the multivariate case see Section 7 of Azzalini and Regoli (2018).

The third related concept is known as the *distorted distribution*. It was introduced in the context of the theory of choice under risk in order to allow a change (distortion) of the initial (or past) baseline distribution (Wang 1996; Yaari 1987). The *distorted distribution function* associated to a CDF F and a distortion function $q : [0, 1] \to [0, 1]$ is defined by

$$F_q(x) = q(F(x)), \tag{2.4}$$

where q is a continuous increasing function such that $q(0) = 0$ and $q(1) = 1$. Under these assumptions for q, F_q is a proper CDF and the respective survival functions $\bar{F}_q = 1 - F_q$ and $\bar{F} = 1 - F$ satisfy a similar relationship $\bar{F}_q(x) = \bar{q}(\bar{F}(x))$, where $\bar{q}(u) = 1 - q(1 - u)$ is another distortion function called *dual distortion*. Note that $\bar{q}$ is a CDF and that it is not the survival function associated to q. The distribution F_q has PDF given by

$$f_q(x) = q'(F(x))f(x) = \bar{q}'(\bar{F}(x))f(x), \tag{2.5}$$

which is also a weighted model with w depending on both q and F. This representation can be applied to extreme data, order statistics and coherent systems with ID components (Navarro 2022); a recent multivariate extension can be seen in Navarro et al. (2022). We put $X \sim \mathcal{DD}(q, F)$ to denote that a variable X follows a distribution with PDF given by expression (2.5).

A typical case arises when $q(u) = u^\alpha$ for $\alpha > 0$. This distortion leads to $F_q = F^\alpha$, known as Lehmann's alternative in hypothesis testing, also equivalent to the Proportional Reversed Hazard Rate (PRHR) model. Analogously, the dual distortion $\bar{q}(u) = u^\alpha$ leads to the well-known Proportional Hazard Rate (PHR) Cox model with $\bar{F}_q = \bar{F}^\alpha$ (Cox 1972). Here, the index α represents a risk parameter that can be related to the characteristics of each individual.

Note that if $\alpha = 2$, the PRHR and PHR models lead to the PDFs $f_q(x) = 2f(x)F(x)$ and $f_q(x) = 2f(x)\bar{F}(x)$, respectively giving the PDFs of $X_{2:2}$ and $X_{1:2}$ in the IID case. The first one is a skew symmetric distribution $\mathcal{SSD}(f, G, h)$ from f, with $G = F$ and $h(x) = x$ if we assume that f is symmetric around zero. Under this assumption, the second one f_q is also a skew symmetric distribution from f, with

$G = F$ and $h(x) = -x$ since $\bar{F}(x) = F(-x)$. The survival function is $\bar{F}^2$ and it represents an accelerated life testing with double hazard rate $h_q = f_q/\bar{F}_q = 2f/\bar{F} = 2h$; hence, the probability mass of X is moved to the left (i.e. the sample items fail before).

Ferreira and Steel (2006) (see also Jones 2015) defined a class of skewed distributions by 'perturbing' a reference baseline density f to get the PDF $f^*(x) = f(x)h(F(x))$, where f is a symmetric PDF, F is its CDF and h is a PDF with support on the unit interval [0, 1]. The CDF of the model is $F^*(x) = H(F(x))$ where H is the CDF of h; thus, it is equivalent to the distortion model (applied to a symmetric PDF). It is worthwhile noting that this model does not always lead to a more skewed version of f. Actually, in some cases, the PDF f^* may be symmetric (see Remark 3.8).

2.2 Connections with stochastic orderings

Weighted, skew symmetric and distorted distributions have close connections with the convex ordering defined in Chan et al. (1990), also known as the likelihood ratio ordering (Shaked and Shanthikumar 2007). We say that the random variable X is less than Y in the likelihood ratio order, shortly written as $X \leq_{lr} Y$ or as $F_X \leq_{lr} F_Y$, if f_Y/f_X is increasing in the union of their supports, where f_X, F_X and f_Y, F_Y represent the PDFs and CDFs of the variables X and Y. This ordering implies other popular orders as the usual stochastic and hazard rate orders; its main properties can be seen in Shaked and Shanthikumar (2007), pp. 42–70.

The weighted distribution is more (less) skewed than the baseline distribution with respect to the likelihood ratio ordering (Chan et al. 1990) when w is increasing (decreasing) since $f_w(x)/f(x) = c_w w(x)$. Thus we have

$$F \leq_{lr} F_w \ (\geq_{lr}) \Leftrightarrow w \text{ increases (decreases).}$$

As an immediate consequence we get

$$\Phi \leq_{lr} F_\lambda \ (\geq_{lr}) \Leftrightarrow \lambda \geq 0 \ (\lambda \leq 0),$$

where F_λ is the skew normal distribution having the PDF f_λ in equation (2.1). A similar statement can be obtained for the $\mathcal{SSD}(f, G, h)$ family:

$$F \leq_{lr} F_{G,h} \ (\geq_{lr}) \Leftrightarrow h \text{ increases (decreases),}$$

where $F_{G,h}$ is the distribution function of the skew symmetric model, with PDF defined by equation (2.2), and F is the distribution function associated to the PDF f.

Analogously, the distorted distribution is more (less) skewed than the baseline distribution with respect this convex ordering when q is convex (concave). This is due to the facts that the distortion function q from equation (2.4) can be written as $q(u) = F_q(F^{-1}(u))$ for $u \in (0, 1)$ and that the ordering $F \leq_{lr} G$ holds if and only if $G(F^{-1})$ is convex, see Shaked and Shanthikumar (2007), p. 43. Thus we have

$$F \leq_{lr} F_q \ (\geq_{lr}) \Leftrightarrow q \text{ convex (concave).}$$

It is worthwhile noting that the convexity of $F_q(F^{-1})$ implies the convexity of the P–P plot since $F_q(x) = F_q(F^{-1}(F(x)))$ so that the CDF F_q is obtained through a convex transformation of the CDF F. Note that a P–P plot can be obtained by plotting $(u, F_q(F^{-1}(u)))$ for $u \in (0, 1)$, that is, the distortion function q, or equivalently by plotting $(F(t), F_q(t))$ for $t \in \mathbb{R}$, see Thas (2010), Chapter 8.

Interestingly, the P–P plot and the Q–Q plot have emerged as useful graphical tools for comparing probabilities and quantiles among distributions, see e.g. Thas (2010), Chapter 8. P–P plots are handy for the visual assessment of the likelihood ratio ordering since the convexity of $G(F^{-1}(u))$ is equivalent to $F \leq_{lr} G$, see Shaked and Shanthikumar (2007), p. 43. Moreover, it depicts the distortion function $q(u) = G(F^{-1}(u))$; such distortion is actually the mapping that transforms the less skewed distribution to the distribution with the higher skewness: $G(x) = q(F(x))$. Unlike P–P plots, the Q–Q plot depicts the function $G^{-1}(F(x))$ which in turn defines the mapping between quantile functions: $G^{-1}(u) = G^{-1}(F(F^{-1}(u)))$; its convexity ensures the convex transform ordering between distributions, denoted by $F \leq_c G$, as defined by Van Zwet (1964). This ordering has been used to describe the concept of skewness in a better way so that skewness measures have been revisited in connection with the ordering (Groeneveld and Meeden 1984); moreover, the Q–Q plot has arisen as a useful tool for the visual assessment of this ordering (Arriaza et al. 2019). We are not aware of similar implications, as the aforementioned ones for the P–P plot and the likelihood ratio ordering, that allow to connect the convex transform ordering with weighted distributions, skew symmetric distributions and distortions, with the exception of a work concerned with the skew normal distribution (Arevalillo and Navarro 2019). In fact, the likelihood ratio ordering provides a general framework that comprises the concept of being more (less) skewed for a gamut of distribution families like the previous ones. This fact motivates the following definition to capture the idea of being more right (less left) skewed with respect to a given baseline PDF f.

Definition 2.1 If f is a univariate PDF and G is a univariate CDF, we define the *right skewed distribution* associated to f and G (shortly written as $\mathcal{RS}(f, G)$) as the model with PDF

$$f_R(x) = c_R G(x) f(x), \tag{2.6}$$

where $c_R = 1/\int_{\mathbb{R}} G(x) f(x) dx$. Analogously, If f is a PDF and $\bar{G}$ is a survival function, we define the *left skewed distribution* associated to f and $\bar{G}$ (shortly written as $\mathcal{LS}(f, \bar{G})$) as the model with PDF

$$f_L(x) = c_L \bar{G}(x) f(x), \tag{2.7}$$

where $c_L = 1/\int_{\mathbb{R}} \bar{G}(x) f(x) dx$.

It can be noted that

$$0 \leq \int_{\mathbb{R}} G(x) f(x) dx \leq \int_{\mathbb{R}} f(x) dx = 1.$$

Hence, the function f_R is a proper PDF iff $\int_{\mathbb{R}} G(x) f(x) dx > 0$; moreover $c_R \geq 1$. Also note that $c_R = 1/\Pr(U < X)$, where U and X are two independent random variables with CDFs G and F, respectively. Similar properties hold for the left skewed PDF defined in (2.7) with $c_L \geq 1$ and $c_L = 1/\Pr(U > X)$.

Note that the symmetry of f is not necessarily assumed in Definition 2.1; in fact, if the baseline distribution with PDF f is already skewed then the derived distribution with PDF f_R (f_L) results in a more (less) skewed distribution in accordance to the convex ordering. Also note that the normalizing constants c_R and c_L can be different from 2. Therefore, the models in (2.6) and (2.7) are not included in the skew symmetric distribution model, the model in Proposition 1.1 of Azzalini and Capitanio (2014) or in the model called 'family 1' in Jones (2015), that is, they are more flexible models than skew symmetric distributions and the family 1 class. In fact, they naturally respond to the intuitive idea of injecting right or left asymmetry into a given not necessarily symmetric PDF.

Clearly, both models in Definition 2.1 are weighted models with sampling probabilities proportional to $G(x)$ and $\bar{G}(x)$, respectively. Note that the skew normal model $\mathcal{SN}(\lambda)$ is a right skewed distribution of the normal model when $\lambda > 0$ and it is a left skewed distribution when $\lambda < 0$. However, the skew symmetric distribution cannot be represented in this way since we do not know if the function h is monotone. Conversely, the $\mathcal{RS}$ and $\mathcal{LS}$ models do not admit the skew symmetric formulation (2.2) with $h(x) = \pm x$ since it is not necessarily assumed that f and $g = G'$ are symmetric around zero. Some extensions to the multivariate case can be seen in Jupp et al. (2016).

Another two particular cases of interest are obtained when we take $G(x) = F(\lambda x)$ for $\lambda > 0$ or $\bar{G}(x) = F(\lambda x)$ for $\lambda < 0$, where F is the distribution function associated to the baseline PDF f.

A desired and relevant property which follows immediately from the previous observations is stated by the next proposition.

Proposition 2.2 *If F is the CDF of f, $F_R \sim \mathcal{RS}(f, G)$ and $F_L \sim \mathcal{LS}(f, \bar{G})$, then*

$$F_L \leq_{lr} F \leq_{lr} F_R$$

for any f, G.

Note that this property tells us that the new distributions are more (less) skewed than the baseline distribution. Also note that F_R satisfies this property as long as G is an increasing function and the expression for f_R in (2.6) defines a PDF. Hence, this family could be extended by replacing the condition "G is a CDF" with both requirements. As we will see in the following section, this extension will allow us to cover additional models for extreme observations under dependency (see also Example 4.5). A similar extension can be obtained for the left skewed distribution by relaxing the condition "$\bar{G}$ is a survival function".

The chain of stochastic inequalities obtained in Proposition 2.2 implies that the theoretical P–P plot $F_R(F^{-1}(u))$ is a convex function mapping $F(x)$ in $F_R(x)$; meanwhile the P–P plot $F_L(F^{-1}(u))$ is a concave function that maps $F(x)$ in $F_L(x)$. On

the other hand, the P–P plot can be put as follows:

$$
\begin{aligned}
F_R(F^{-1}(u)) &= \int_{-\infty}^{F^{-1}(u)} c_R G(x) f(x) dx \\
&= c_R \int_0^u G(F^{-1}(v)) dv = \frac{\int_0^u G(F^{-1}(v)) dv}{\int_0^1 G(F^{-1}(v)) dv}
\end{aligned} \tag{2.8}
$$

which allows to characterize it as the proportional cumulative area under the P–P plot given by $G(F^{-1}(u))$, which in turn is the theoretical P–P plot between the baseline CDF F and the CDF G, which controls the right skewed perturbation in Definition 2.1. If we assume that the CDF G is a distortion of F, that is, $G = q(F)$, then so is F_R with $F_R = q^*(F)$ for

$$
q^*(u) = \frac{\int_0^u q(v) dv}{\int_0^1 q(v) dv}.
$$

Also note that the PDF associated to the CDF q^* is proportional to q.

Analogously, for the left skewed distribution we get

$$
\bar{F}_L(\bar{F}^{-1}(u)) = \int_{\bar{F}^{-1}(u)}^{\infty} c_L \bar{G}(x) f(x) dx = c_L \int_0^u \bar{G}(\bar{F}^{-1}(v)) dv = \frac{\int_0^u \bar{G}(\bar{F}^{-1}(v)) dv}{\int_0^1 \bar{G}(\bar{F}^{-1}(v)) dv}.
$$

If $\bar{G}$ admits a distortion representation from $\bar{F} = 1 - F$ with dual distortion $\bar{q}$, that is, $\bar{G} = \bar{q}(\bar{F})$, then we get a similar representation for $\bar{F}_L$ with dual distortion

$$
\bar{q}^*(u) = \frac{\int_0^u \bar{q}(v) dv}{\int_0^1 \bar{q}(v) dv}
$$

for $u \in [0, 1]$.

Finally, we include some preservation properties for the unimodality of models (2.6) and (2.7). In this sense a PDF f is said to be increasing likelihood ratio (ILR) or strongly unimodal if f is log-concave (i.e. $\log f$ is concave). This is an aging class that implies the popular increasing hazard rate (IHR) and decreasing reversed hazard rate (DRHR) aging classes, where the hazard rate and reversed hazard rate functions are defined by $f/\bar{F}$ and f/F, respectively. The ILR property is equivalent to the increasing property for the eta Glaser function $\eta = -f'/f$. Analogously, the IHR and DRHR classes can be characterized by the log-concavity of functions $\bar{F}$ and F, respectively.

Thus we can obtain the following results for the unimodality property for f_R and f_L.

Proposition 2.3 *Let f_R and f_L be the PDFs in (2.6) and (2.7).*

(i) If f is ILR and G is DRHR, then f_R is ILR and, in particular, it is unimodal.
(ii) If f is ILR and G is IHR, then f_L is ILR and, in particular, it is unimodal.
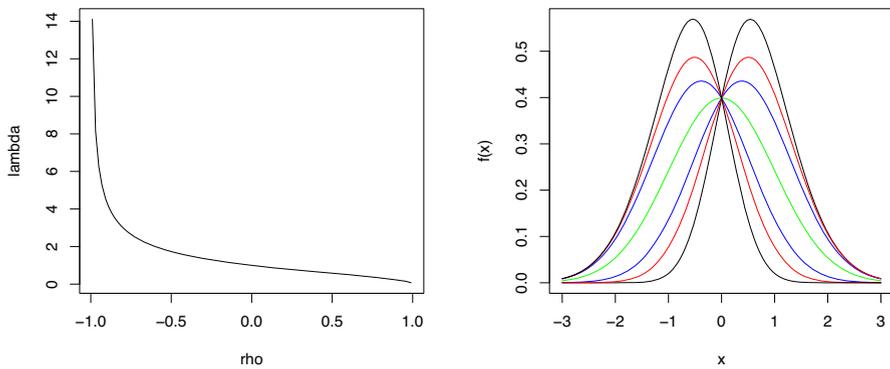
Fig. 1 Value of λ (left) in a bivariate normal distribution with correlation ρ. PDFs (right) for $X_{2:2}$ and $X_{1:2}$ for $\rho = -0.5$ (black), 0 (red) and 0.5 (blue). The green line represents the standard normal PDF ($\rho = 1$, $\lambda = 0$) (color figure online)

The proof is easy. In particular, note that if both f and G are ILR, then both models are unimodal. For the skew normal model we have $G(x) = F(\lambda x)$ when $\lambda > 0$ or $\bar{G}(x) = F(\lambda x)$ when $\lambda < 0$. Moreover, the normal PDF f is ILR and then so is G (in both cases). Therefore, we get that the skew normal model is ILR and, in particular, it is unimodal (a well-known property). The particular cases in which $G = q(F)$ or $\bar{G} = \bar{q}(\bar{F})$ (they are distorted distribution of F) can be studied with the preservation properties for these classes given in Navarro (2022), p. 120.

3 Distributions of extremes

Once we have explored the relationships among weighted distributions, skew symmetric distributions and distortions along with their connections with the likelihood ratio ordering, in this section we study how they are related to the distributions of extremes. Our purpose is to show that the distributions of the sample extremes $X_{1:n}$ and $X_{n:n}$ belong to the family of skewed distributions introduced in Definition 2.1.

For the sake of motivation, let us start with the case $n = 2$ under normality: we will assume that (X_1, X_2) has a bivariate normal distribution with standardized marginal CDFs Φ and correlation ρ; then it can be shown that $X_{2:2}$ has a skew normal distribution $\mathcal{SN}(\lambda)$ with $\lambda = \sqrt{(1-\rho)/(1+\rho)} > 0$. This result dates back to an early work by Roberts (1966) and it is a precursor of the skew normal model introduced by Azzalini's seminal paper (Azzalini 1985); later on it has been revisited and generalized under different settings (Azzalini and Capitanio 2003; Loperfido et al. 2007) including its extension to elliptical distributions (Loperfido 2008); an analogous argument would prove that $X_{1:2}$ has a skew normal distribution with $\mathcal{SN}(-\lambda)$. Both cases can be interpreted as sampling procedures for which the higher (lower) values have the largest sampling probabilities. It is easy to see that λ is a decreasing function of ρ with $\lambda \to \infty$ as $\rho \to -1$ and $\lambda \to 0$ as $\rho \to 1$ (see the left plot of Fig. 1). If $\rho = 0$ (IID case) then $\lambda = 1$ and the distribution function of the maximum becomes $F_{2:2}(x) = \Phi^2(x)$.

The skewed PDFs of the extremes $X_{2:2}$ and $X_{1:2}$ for the values $\rho = -0.5, 0, 0.5$ can be seen in the right plot of Fig. 1. Note that the most skewed distributions correspond to $\rho = -0.5$ (black lines), that is, the most extreme values are obtained with negative dependence as expected. Of course, when $\rho \to 1$ we get the normal distribution, that is, the singular case with $X_1 = X_2 = X_{1:2} = X_{2:2}$.

The results for the bivariate normal distribution can be extended as follows for exchangeable (EXC) random vectors: We say that a random vector $(X_1, \ldots, X_n)$ is EXC if it has the same distribution as $(X_{\sigma(1)}, \ldots, X_{\sigma(n)})$ for any permutation σ. The random vector $(X_1, \ldots, X_n)$ is EXC if and only if $X_1, \ldots, X_n$ are ID and its copula (or its survival copula) is invariant under permutations.

The following proposition, partially borrowed from Corollary 1 in Arellano-Valle and Genton (2008), gives the extension to EXC bivariate vectors. For completeness, we provide a proof here since the statement establishes a new copula formulation for the resulting skewed distribution.

Proposition 3.1 *Let (X_1, X_2) be an EXC random vector with absolutely continuous copula C and common marginal CDF F and PDF f. Then the PDF of $X_{2:2}$ can be written as*

$$f_{2:2}(x) = c_R f(x) G(x), \tag{3.1}$$

with $G(x) = \partial_1 C(F(x), F(x))$ and $c_R = 2$.

Proof The distribution function of $X_{2:2}$ is

$$F_{2:2}(x) = \Pr(\max(X_1, X_2) \leq x) = \Pr(X_1 \leq x, X_2 \leq x) = C(F(x), F(x))$$

for all x. Hence its PDF is

$$f_{2:2}(x) = F'_{2:2}(x) = f(x)\partial_1 C(F(x), F(x)) + f(x)\partial_2 C(F(x), F(x)).$$

Taking into account that C is permutation symmetric, we obtain (3.1). □

Remark 3.2 From (3.1), $X_{2:2}$ has a right skewed representation if and only if $G(x) = \partial_1 C(F(x), F(x))$ is a CDF. Note that equation (3.1) is useful to get the explicit expression of the skewed distribution of $X_{2:2}$ for different copulas as illustrated in the examples of Sect. 4. In the independent case $\partial_1 C(u, v) = v$ so that $f_{2:2}(x) = 2f(x)F(x)$ and the distribution of $X_{2:2}$ has a right skewed representation. In order to check if $G(x) = \partial_1 C(F(x), F(x))$ is a distribution function, we argue as follows: from the copula representation, the joint PDF of the bivariate vector (X_1, X_2) can be written as

$$\mathbf{f}(x_1, x_2) = f(x_1) f(x_2) \partial_{1,2} C(F(x_1), F(x_2)).$$

Hence the PDF of the conditional random variable $(X_2 | X_1 = x)$ is

$$f_{2|1}(x_2 | x_1) = \frac{\mathbf{f}(x_1, x_2)}{f(x_1)} = f(x_2) \partial_{1,2} C(F(x_1), F(x_2))$$

for x_1 such that $f(x_1) > 0$. Then its distribution function is

$$F_{2|1}(x_2|x_1) = \int_{-\infty}^{x_2} f_{2|1}(z|x_1)dz = \int_{-\infty}^{x_2} f(z)\partial_{1,2}C(F(x_1), F(z))dz$$
$$= [\partial_1 C(F(x_1), F(z))]_{z=-\infty}^{x_2} = \partial_1 C(F(x_1), F(x_2))$$

whenever $\lim_{z \to -\infty} \partial_1 C(F(x_1), F(z)) = 0$ holds. Therefore we get

$$G(x) = \partial_1 C(F(x), F(x)) = F_{2|1}(x|x)$$

provided that $\lim_{v \to 0^+} \partial_1 C(F(x), v) = 0$ holds for all x. The expression above can also be obtained from Proposition 1 in Arnold et al. (2008). However, note that $G(x) = F_{2|1}(x|x)$ is not necessarily a CDF in x. For many copulas, it can be shown that G is a CDF, although this is not always the case (see Sect. 4).

Using an analogous argument, we could prove a similar result for the minimum. The result is stated by the following proposition.

Proposition 3.3 *Let (X_1, X_2) be an EXC random vector with absolutely continuous survival copula $\widehat{C}$ and common marginal CDF F and PDF f. Then the PDF of $X_{1:2}$ can be written as*

$$f_{1:2}(x) = c_L f(x)\bar{G}(x), \tag{3.2}$$

with $\bar{G}(x) = \partial_1 \widehat{C}(\bar{F}(x), \bar{F}(x))$ and $c_L = 2$.

Remark 3.4 Expression (3.2) shows that $X_{1:2}$ has a $\mathcal{LS}(f, \bar{G})$ distribution if and only if $\bar{G}$ is a survival function. The requirements resulting from Propositions 3.1 and 3.3, that allow to derive skewed distributions, hold for many copulas. However, they are not satisfied in all the cases; see Navarro and Sordo (2018) or the Examples 4.4 and 4.5 of the next section for counterexamples. An alternative direct proof without copulas can be obtained by using that

$$f_{2:2}(x) = f_1(x)\Pr(X_2 \le x|X_1 = x) + f_2(x)\Pr(X_1 \le x|X_2 = x)$$

which under the EXC assumption leads to

$$f_{2:2}(x) = 2f(x)\Pr(X_2 \le x|X_1 = x)$$

as stated in the first proposition. Note that here we just use the conditional distribution function of $(X_2 \le x|X_1 = x)$. This representation can be applied to models with known conditionals. Several models with known conditionals were obtained in Arnold et al. (1999). If we relax the EXC condition and we just assume that X_1 and X_2 are ID with a common PDF f, then

$$f_{2:2}(x) = 2f(x)G(x)$$

where

$$G(x) = \frac{1}{2}\Pr(X_2 \leq x|X_1 = x) + \frac{1}{2}\Pr(X_1 \leq x|X_2 = x).$$

Provided that we can ensure that $\Pr(X_2 \leq x|X_1 = x)$ and $\Pr(X_1 \leq x|X_2 = x)$ are CDFs, we found that G is a mixture of two CDFs which in turn implies that the maximum $X_{2:2}$ has a right skewed distribution. Similar results hold for the minimum $X_{1:2}$. These findings are in agreement with the results in Arellano-Valle and Genton (2008), showing that $X_{2:2}$ has a $\mathcal{RS}(f, G)$ under the elliptically contoured distribution.

The results for the bivariate case can be extended to the n dimensional case, that is, for $X_{1:n}$ and $X_{n:n}$. The results for the n dimensional normal distribution were obtained in Loperfido et al. (2007). In order to establish the result for the general n dimensional ID case, the following definition is needed: The *diagonal section* of a copula C is $\delta_C(u) = C(u, \ldots, u)$ for $u \in [0, 1]$. It can be shown that $\delta_C(u)$ is a distortion function which can be extended to a continuous CDF with support included in [0, 1]. For additional properties and details about diagonal sections of copulas see Durante and Sempi (2016) and Nelsen (2006).

Proposition 3.5 *Let $(X_1, \ldots, X_n)$ be a random vector with copula C and common marginal CDF and PDF given by F and f, respectively. Then the PDF of $X_{n:n}$ can be written as*

$$f_{n:n}(x) = f(x)\delta'_C(F(x)), \tag{3.3}$$

where δ'_C is the first derivative of the diagonal section of C.

Proof The CDF of $X_{n:n}$ is

$$\begin{aligned} F_{n:n}(x) &= \Pr(\max(X_1, \ldots, X_n) \leq x) = \Pr(X_1 \leq x, \ldots, X_n \leq x) \\ &= C(F(x), \ldots, F(x)) = \delta_C(F(x)) \end{aligned}$$

for all x. Hence its PDF is $f_{n:n}(x) = F'_{n:n}(x)$ and (3.3) holds. $\square$

We have a similar result for the minimum.

Proposition 3.6 *Let $(X_1, \ldots, X_n)$ be a random vector with survival copula $\widehat{C}$ and common marginal CDF and PDF given by F and f, respectively. Then the PDF of $X_{1:n}$ can be written as*

$$f_{1:n}(x) = f(x)\delta'_{\widehat{C}}(\bar{F}(x)), \tag{3.4}$$

where $\delta'_{\widehat{C}}$ is the first derivative of the diagonal section of $\widehat{C}$.

Proof The survival function of $X_{n:n}$ can be written as

$$\bar{F}_{1:n}(x) = \Pr(\min(X_1, \ldots, X_n) > x) = \Pr(X_1 > x, \ldots, X_n > x)$$
$$= \widehat{C}(\bar{F}(x), \ldots, \bar{F}(x)) = \delta_{\widehat{C}}(\bar{F}(x))$$

for all x. Hence its PDF is

$$f_{1:n}(x) = F'_{1:n}(x) = f(x)\delta'_{\widehat{C}}(\bar{F}(x))$$

and (3.4) holds.   □

Remark 3.7 As in the bivariate case, (3.3) gives a weighted PDF for $X_{n:n}$. This implies a movement of the probability mass of f to the right when δ'_C is increasing (i.e. δ_C is convex). However, in order to get a proper right skewed distribution, as in Definition 2.1, the function $\delta'_C(F(x))$ needs to be proportional to a distribution function $G(x)$. In such a case, $f_{n:n}(x) = c_R f(x) G(x)$, admitting a right skewed representation, and the respective P–P plots can be obtained as $G(F^{-1}(u)) = (1/c_R)\delta'_C(u)$ and $F_{n:n}(F^{-1}(u)) = \delta_C(u)$ which is a convex function. For example, in the IID n-dimensional case, we get $\delta_C(u) = u^n$ and $\delta'_C(u) = nu^{n-1}$. Therefore, the distribution of $X_{n:n}$ gets more right skewed than f with skewness mechanism given by $G(x) = F^{n-1}(x)$ and $c_R = n$. It can also be interpreted as a sampling procedure over f where the sampling probability for a value $X = x$ is proportional to $w(x) = F^{n-1}(x) = \Pr(X_1 \leq x, \ldots, X_{n-1} \leq x)$, that is, $X = x$ is bigger than $n-1$ random independent values $X_1, \ldots, X_{n-1}$ from X. Other examples with different copulas are given in Sect. 4. The results are similar for the minimum: for example, in the independent case, we get that $X_{1:n}$ has more left skewed distribution than f with $\bar{G}(x) = \bar{F}^{n-1}(x)$ and $c_L = n$. The interpretation as a weighted sampling is analogous: a value $X = x$ is available in the sample if it is less than $n-1$ independent copies of X.

Remark 3.8 Similar results can be obtained for the other order statistics $X_{i:n}$ for $i = 2, \ldots, n-1$. If we assume that $X_1, \ldots, X_n$ are ID (dependent or independent), their PDFs can also be written as

$$f_{i:n}(x) = f(x)q'_{i:n}(F(x))$$

for $i = 2, \ldots, n-1$, where $q_{i:n}$ is a distortion function which depends on C, i and n (see e.g. Navarro 2022, p. 61). In particular, in the IID case $q_{i:n}$ is a polynomial and

$$q'_{i:n}(u) = i\binom{n}{i}u^{i-1}(1-u)^{n-i}$$

for $u \in [0, 1]$. In this case, $q'_{i:n}(F(x))$ is not proportional to a distribution function so we cannot obtain a skewed representation for $i = 2, \ldots, n-1$. However, note that $X_{i:n}$ has both a weighted representation with $w(x) = F^{i-1}(x)\bar{F}^{n-i}(x)$ and a distortion representation with distortion $q_{i:n}$. For example, for $X_{2:3}$ we get $w(x) = F(x)\bar{F}(x)$

and $q_{2:3}(u) = 3u^2 - 2u^3$ for $u \in [0, 1]$. These transformations do not provide a skewing mechanism; it can be noted that if f is symmetric then $f_{2:3}$ is also symmetric.

Remark 3.9 In order to write the distribution of $X_{n:n}$ as a right skewed distribution, Definition 2.1 could be extended by considering

$$f_R(x) = c_R f(x) \mathbf{G}_x(h(x))$$

for all $x \in \mathbb{R}$, where f is a PDF, $\mathbf{G}_x$ is a k-dimensional distribution function and $h : \mathbb{R} \to \mathbb{R}^k$. By doing so, from Corollary 1 in Arellano-Valle and Genton (2008), we get that if $(X_1, \ldots, X_n)$ is EXC, then $X_{n:n}$ has a right skewed distribution of order $k = n - 1$ derived from the common marginal PDF f,

$$\mathbf{G}_x(x_1, \ldots, x_{n-1}) = \Pr(X_1 \le x_1, \ldots, X_{n-1} \le x_{n-1} | X_n = x),$$

$h(x) = (x, \ldots, x)$ and $c_R = n$. The same thing happens for elliptically contoured distributions (see Corollary 2 in Arellano-Valle and Genton 2008). In this case $\mathbf{G}_x$ is a fixed distribution, that is, it does not depend on x. Actually, as in this case $G(x) = \mathbf{G}_x(h(x))$ is a univariate CDF, then $X_{n:n}$ follows a $\mathcal{RS}(f, G)$ distribution. This is not always the case for other EXC distributions.

Finally, we provide a characterization of the right skewed representation for $X_{n:n}$. A similar characterization can be obtained for $X_{1:n}$.

Proposition 3.10 *Let $(X_1, \ldots, X_n)$ be a random vector with copula C and common marginal CDF and PDF given by F and f. Then $X_{n:n}$ follows a distribution having a right skewed representation from f if and only if δ_C is convex and $\delta_C'(0) = 0$. Moreover, in this case, $c_R \in (0, n]$.*

Proof If $X_{n:n}$ can be written as $\mathcal{RS}(f, G)$, then $f_{n:n}(x) = c_R f(x) G(x)$, where G is a CDF. Taking into account (3.3), we get $\delta_C'(F(x)) = c_R G(x) \ge 0$ and

$$\delta_C'(0) = \delta_C'(F(-\infty)) = c_R G(-\infty) = 0.$$

Conversely, if we assume that δ_C is convex and $\delta_C'(0) = 0$ then δ_C' is an increasing function in $(0, 1)$. Moreover, from $\delta_C(u) = C(u, \ldots, u)$, we get

$$\delta_C'(u) = \sum_{i=1}^{n} \partial_i C(u, \ldots, u)$$

for $u \in [0, 1]$. Then, from Theorem 1.6.1 in Durante and Sempi (2016), p. 21, we obtain $\delta_C'(u) \le n$ for all $u \in [0, 1]$. Thus we define $c_R = \delta_C'(1) \in (0, n]$ so that we get

$$f_{n:n}(x) = f(x) \delta_C'(F(x)) = c_R f(x) G(x),$$

where $G(x) = (1/c_R)\delta'_C(F(x))$ is an increasing function satisfying

$$G(-\infty) = \frac{1}{c_R}\delta'_C(F(-\infty)) = \frac{1}{c_R}\delta'_C(0) = 0$$

and

$$G(\infty) = \frac{1}{c_R}\delta'_C(F(\infty)) = \frac{1}{c_R}\delta'_C(1) = 1.$$

Hence, G is a CDF and $X_{n:n} \sim \mathcal{RS}(f, G)$. $\square$

Note that the condition $\delta'_C(0) = 0$ can be removed from the assumptions in the previous proposition provided that we extend the right skewed class by allowing G to be a pseudo-distribution function with a possible mass at $-\infty$. In fact the case $\delta'_C(0) > 0$ may come up for some copulas as the Example 4.5 in the next section will show; meanwhile, the condition about the convexity of the diagonal section cannot be dropped out from the assumptions (see Example 4.4).

4 Examples

In the first example we consider a dependent case (copula) where $X_{2:2}$ admits a right skewed representation.

Example 4.1 Let us consider a bivariate random vector (X_1, X_2) with a common marginal CDF F, PDF $f = F'$ and the following EXC copula, known as Farlie-Gumbel-Morgenstern (FGM) copula:

$$C(u, v) = uv + \theta uv(1 - u)(1 - v)$$

for $u, v \in [0, 1]$, where $\theta \in [-1, 1]$ is a dependence parameter (see Nelsen 2006, p. 77). Its first partial derivative is

$$\partial_1 C(u, v) = v + \theta(1 - 2u)(v - v^2)$$

for $u, v \in [0, 1]$. A straightforward calculation shows that

$$q(u) = \partial_1 C(u, u) = u + \theta(1 - 2u)(u - u^2)$$

for all $u \in [0, 1]$ is a distortion function for any $\theta \in [-1, 1]$. Hence, from (3.1), $X_{2:2}$ has the following right skewed representation

$$f_{2:2}(x) = 2f(x)G(x),$$

where

$$G(x) = q(F(x)) = F(x) + \theta(1 - 2F(x))F(x)\bar{F}(x)$$
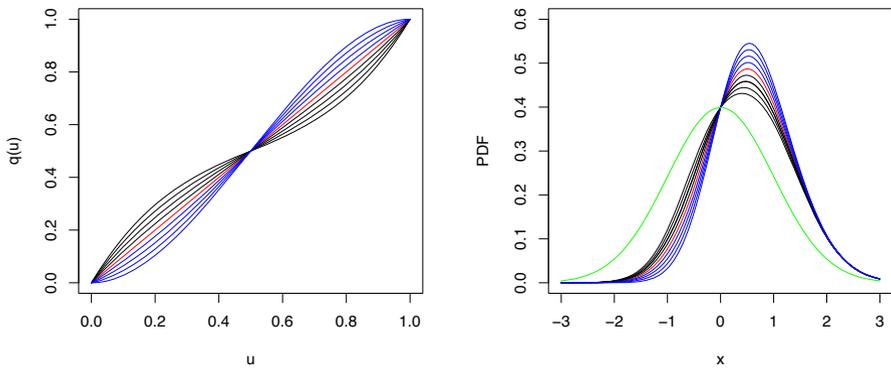
Fig. 2 Distortion function q (left) in Example 4.1 for $\theta = 0$ (red), $-1, -0.75, -0.5, -0.25$ (blue) and $0.25, 0.5, 0.75, 1$ (black). PDF $f_{2:2}$ (right) of the maximum for the same parameter values and a standard normal PDF (green line) (color figure online)

is a genuine distribution function for any $\theta \in [-1, 1]$. The distortion functions q obtained for different values of θ are plotted in Fig. 2, left. The respective PDFs $f_{2:2}$ for the standard normal model are plotted in Fig. 2, right. All of them are right skewed versions of the standard normal PDF (green line). As we can see, all of them are unimodal and very similar in skewness. This is due to the weak dependence induced by the FGM copula. We can get more skewed models by using extreme copulas with negative dependency, whereas copulas with positive dependency will lead to models similar to the parent distribution. For example, with the counter-monotonic copula (i.e. the lower Fréchet-Hoeffding bound), the measure of skewness of Arnold and Groeneveld (1995) of $X_{2:2}$ reaches the maximum value 1 when f is symmetric with respect to its mode.

The diagonal section of C is

$$\delta_C(u) = u^2 + \theta(u - u^2)^2$$

with first derivative

$$\delta'_C(u) = 2u + 2\theta(1 - 2u)(u - u^2).$$

Therefore, by using (3.3), we get the same representation for $f_{2:2}$ (as expected). The representations for $X_{1:2}$ are obtained in a similar way due to the symmetry of the model; they are left skewed distributions.

On the other hand, from expression (2.8) and Remark 3.7, the respective theoretical P–P plots are $G(F^{-1}(u)) = q(u)$ and $F_R(F^{-1}(u)) = \delta_C(u)$ for $u \in [0, 1]$. As expected, δ_C is a convex function for any $\theta \in [-1, 1]$; hence, $F \leq_{lr} F_{2:2}$ holds. However, $G(F^{-1})$ is not a convex function (see Fig. 2, left).

The previous example can be generalized to the n dimensional case as follows.

Example 4.2 Let us consider a random vector $(X_1, \ldots, X_n)$ with a common marginal CDF F, PDF $f = F'$ and the following FGM EXC copula:

$$C(u_1, \ldots, u_n) = u_1 \ldots u_n + \theta u_1 \ldots u_n (1 - u_1) \ldots (1 - u_n)$$

for $u_1, \ldots, u_n \in [0, 1]$, where $\theta \in [-1, 1]$ is a dependence parameter. Its diagonal section is

$$\delta_C(u) = u^n + \theta(u - u^2)^n$$

and its derivative

$$\delta_C'(u) = nu^{n-1} + n\theta(1 - 2u)(u - u^2)^{n-1}.$$

Then, by using (3.3), we get

$$f_{n:n}(x) = f(x)\delta_C'(F(x)) = nf(x)G(x),$$

where

$$G(x) = F^{n-1}(x) + \theta(1 - 2F(x))(F(x) - F^2(x))^{n-1} = q(F(x))$$

and $q(u) = u^{n-1} + \theta(1 - 2u)(u - u^2)^{n-1}$. In order to prove that G is a distribution function for any $n = 3, 4, \ldots$ (the case $n = 2$ was addressed in the previous example) and $\theta \in [-1, 1]$, we consider

$$\begin{aligned}
q'(u) &= (n-1)u^{n-2} - 2\theta(u - u^2)^{n-1} + \theta(n-1)(1 - 2u)^2(u - u^2)^{n-2} \\
&= (n-1)u^{n-2} - 2\theta(u - u^2)^{n-1} + \theta(n-1)(1 - 4u + 4u^2)(u - u^2)^{n-2} \\
&= (n-1)u^{n-2} - 2\theta(u - u^2)^{n-1} + \theta(n-1)(u - u^2)^{n-2} \\
&\quad - 4\theta(n-1)(u - u^2)^{n-1}.
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{q'(u)}{(n-1)u^{n-2}} &= 1 - \frac{2}{n-1}\theta(1 - u)^{n-2}(u - u^2) + \theta(1 - u)^{n-2} \\
&\quad - 4\theta(1 - u)^{n-2}(u - u^2) \\
&= 1 + \theta(1 - u)^{n-2}\left[1 - \frac{4n - 2}{n - 1}(u - u^2)\right].
\end{aligned}$$

Here we know that $(u - u^2) \in [0, 1/4]$ for $u \in [0, 1]$ and that $(4n - 2)/(n - 1) \in [4, 5]$ for $n = 3, 4, \ldots$. Therefore

$$1 - \frac{4n - 2}{n - 1}(u - u^2) \in [-1/4, 1]$$
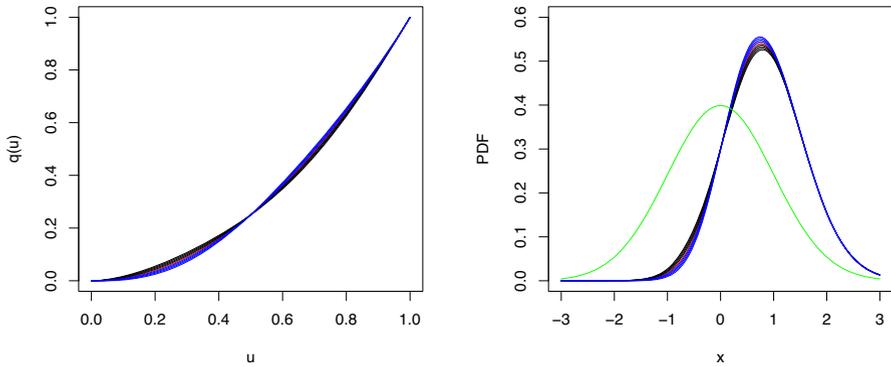
Fig. 3 Distortion function q (left) in Example 4.2 for $\theta = 0$ (red), $-1, -0.75, -0.5, -0.25$ (blue) and $0.25, 0.5, 0.75, 1$ (black). PDF $f_{3:3}$ (right) of the maximum for the same parameter values and a standard normal PDF (green line) (color figure online)

and so, for $u \in [0, 1]$, we have

$$\frac{q'(u)}{(n-1)u^{n-2}} \geq 1 - 0.25\theta(1-u)^{n-2} \geq 0$$

for $\theta \geq 0$ and

$$\frac{q'(u)}{(n-1)u^{n-2}} \geq 1 + \theta(1-u)^{n-2} \geq 0$$

for $\theta \leq 0$. Hence q is an increasing function in the interval $u \in [0, 1]$ for which $q(0) = 0$ and $q(1) = 1$; consequently, it is a distortion function and G is a distribution function for any CDF F, dependency parameter $\theta \in [-1, 1]$ and $n = 3, 4, \ldots$. Thus the distribution of the maximum $X_{n:n}$ admits a right skewed representation.

The different distortion functions q obtained for $n = 3$ and $\theta = 0$ (red, IID case), $-1, -0.75, -0.5, -0.25$ (blue) and $0.25, 0.5, 0.75, 1$ (black) are displayed in Fig. 3, left. The respective PDFs $f_{2:2}$ obtained with a standard normal PDF are shown in the right plot. Note that they are more right skewed than the distribution function obtained with $n = 2$ (see the plot of Fig. 2, right).

As in the previous example, the theoretical P–P plots can be obtained as $G(F^{-1}(u)) = q(u)$ and $F_R(F^{-1}(u)) = \delta_C(u)$. We have proved above that the last one is a convex function and so $F \leq_{lr} F_{n:n}$ holds. The first one is also a convex function when $n = 3$ (see Fig. 2, left).

In the next example we consider a non-exchangeable tridimensional copula that also leads to a right skewed representation for $X_{3:3}$.

Example 4.3 To get a non-exchangeable copula we consider a random vector (X_1, X_2, X_3) with copula $C(u, v, w) = uD(v, w)$, where

$$D(v, w) = \frac{vw}{v + w - vw}$$
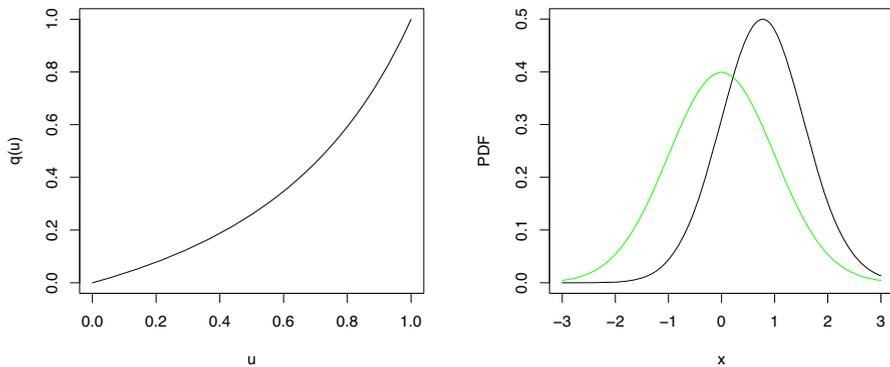
Fig. 4 Distortion function q (left) in Example 4.3 and PDF $f_{3:3}$ (right, black line) for a standard normal PDF (right, green line) (color figure online)

is a Clayton copula with $\theta = 1$ (see Nelsen 2006, p. 116). Under the copula C, X_1 is independent of X_2 and X_3 but X_2 and X_3 are dependent with copula D. The diagonal section of C has derivative

$$\delta'_C(u) = \frac{4u - u^2}{(2 - u)^2}.$$

Hence, from (3.3), $X_{3:3}$ has a right skewed PDF given by

$$f_{3:3}(x) = f(x)\delta'_C(F(x)) = 3f(x)G(x)$$

with

$$G(x) = q(F(x)) = \frac{1}{3}\frac{4F(x) - F^2(x)}{(2 - F(x))^2},$$

and

$$q(u) = \frac{1}{3}\frac{4u - u^2}{(2 - u)^2}.$$

A straightforward calculation shows that G is a genuine CDF for any distribution function F. In Fig. 4, we plot the distortion function q (left) and the PDF $f_{3:3}$ (right) for a standard normal distribution.

From expression (2.8) and Remark 3.7, the theoretical P–P plots are $G(F^{-1}(u)) = q(u)$ (plotted in Fig. 4, left) and $F_R(F^{-1}(u)) = \delta_C(u)$ which are convex functions.

We must say that with the majority of the copulas we get right and left skewed representations for the distributions of the extremes $X_{n:n}$ and $X_{1:n}$, respectively. However, for some specific copulas, these representations fail. Let us consider an example based on a copula defined in Example 4.1 of Navarro et al. (2018).

Example 4.4 Let (X_1, X_2) be two ID random variables having an EXC copula C with diagonal section

$$\delta_C(u) = \begin{cases} u, & \text{for } 0 \le u \le 1/3; \\ 1/3, & \text{for } 1/3 < u \le 2/3; \\ 2u - 1, & \text{for } 2/3 < u \le 1. \end{cases}$$

This function satisfies the properties of a proper diagonal section of an EXC copula, see Example 4.1 of Navarro et al. (2018). Hence, from (3.3), the PDF of $X_{2:2}$ is

$$f_{2:2}(x) = f(x)\delta'_C(F(x)),$$

where

$$\delta'_C(u) = \begin{cases} 1, & \text{for } 0 \le u \le 1/3; \\ 0, & \text{for } 1/3 < u \le 2/3; \\ 2, & \text{for } 2/3 < u \le 1. \end{cases}$$

Therefore $\delta'_C(F(x))$ is not proportional to a CDF so that $X_{2:2}$ PDF does not admit a right skewed representation. However, note that $F_{2:2}$ can be written as a distortion distribution since

$$F_{2:2}(x) = \delta_C(F(x)) = \begin{cases} F(x), & \text{for } x: 0 \le F(x) \le 1/3; \\ 1/3, & \text{for } x: 1/3 < F(x) \le 2/3; \\ 2F(x) - 1, & \text{for } x: 2/3 < F(x) \le 1. \end{cases}$$

Its PDF can be written as a weighted distribution since $f_{2:2}(x) = f(x)w(x)$ with

$$w(x) = \begin{cases} 1, & \text{for } x: 0 \le F(x) \le 1/3; \\ 0, & \text{for } x: 1/3 < F(x) \le 2/3; \\ 2, & \text{for } x: 2/3 < F(x) \le 1. \end{cases}$$

As w is not increasing, the "natural" likelihood ratio ordering $F \le_{lr} F_{2:2}$ does not hold. The same thing happens for the hazard rate order (see Example 4.1 of Navarro et al. 2018). The PDFs $f_{2:2}$ for baseline uniform and normal distributions are displayed in Fig. 5.

In the last example we show a case where G is increasing but it is not a genuine CDF since it assigns probability mass at $-\infty$. As we have already mentioned, this case could also fall within the right skewed representation scheme as long as the conditions in Definition 2.1 is slightly relaxed, because we can still ensure the ordering $F \le_{lr} F_R$ due to the increasing behavior of G.

Example 4.5 Let us consider the maximum of two ID random variables with a standard normal distribution, that is, $X_{2:2} = \max(X_1, X_2)$, where $X_i \sim N(0, 1)$. If they are dependent with the following Clayton copula:

$$C(u, v) = \frac{uv}{u + v - uv}$$
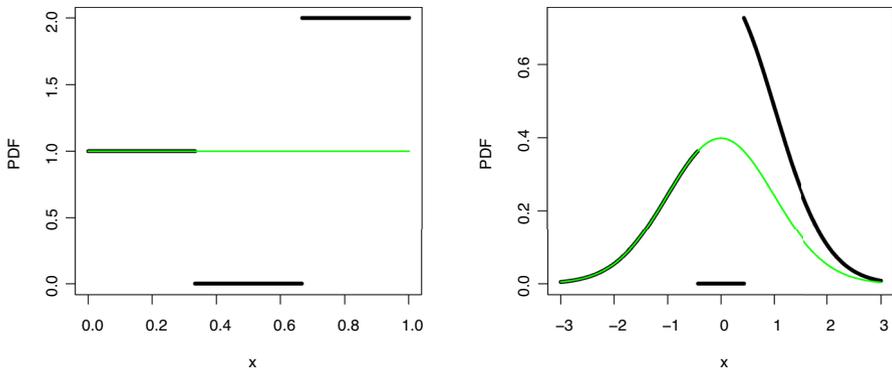
Fig. 5 PDF $f_{2:2}$ for the representation of $X_{2:2}$ in Example 4.4 for a standard uniform distribution (left) and a standard normal distribution (right). The green lines represent the baseline (marginal) PDF (color figure online)

for $u, v \in (0, 1)$ thence, the CDF of $X_{2:2}$ can be obtained as

$$F_{2:2}(x) = \Pr(X_{2:2} \leq x) = C(\Phi(x), \Phi(x)) = \delta_C(\Phi(x)),$$

where the diagonal section of C is

$$\delta_C(u) = C(u, u) = \frac{u}{2 - u}.$$

Therefore $\delta'_C(u) = \dfrac{2}{(2 - u)^2}$ for $u \in [0, 1]$. This is a convex function satisfying that $\delta'_C(1) = 2$ whereas $\delta'_C(0) = 1/2$. Hence, the PDF $f_{2:2}$ admits the following representation:

$$f_{2:2}(x) = \phi(x)\delta'_C(\Phi(x)) = 2\phi(x)G(x),$$

where ϕ is the PDF of a standard normal and $G(x) = 0.5\delta'_C(\Phi(x))$. Note that G is an increasing and continuous function satisfying $G(\infty) = 1$ so the ordering $F \leq_{lr} F_{2:2}$ follows; however, G is not a CDF because $G(-\infty) = 1/4$. Even so, we could get a right skewed formulation of the distribution by allowing G to be a pseudo-CDF having positive mass at $-\infty$. If we replace the normal distribution with a lifetime distribution with support $(0, \infty)$ (e.g. an exponential distribution), we would overcome this problem since G is a CDF with a positive mass probability at $x = 0$.

5 Numerical work

This section illustrates the previous theoretical findings with several numerical examples that involve both artificial and real data. In the first example we analyze a simulated data set from IID random variables whereas the second example considers a data
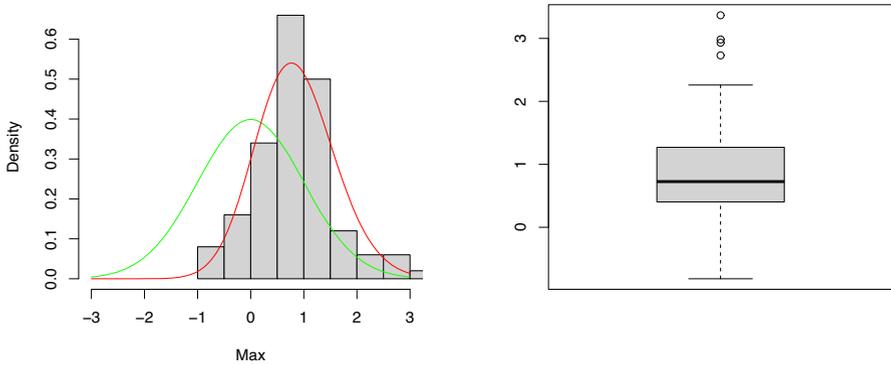
Fig. 6 Histogram (left) made with 100 data from $X_{3:3}$ in Example 5.1 for a standard normal distribution. The green line represent the baseline (marginal) normal PDF and the red line the real skewed PDF $f_{3:3}$. The box-plot (right) shows the asymmetry of the data (color figure online)

experiment that involves two dependent random variables with the copula considered in Example 4.5. Finally, a real data set about daily maximum temperatures in the Spanish Iberian Peninsula throughout a thirty year period is analyzed.

Example 5.1 Let us consider the maximum of three IID random variables with a standard normal distribution, that is, $X_{3:3} = \max(X_1, X_2, X_3)$, where $X_i \sim N(0, 1)$. We simulate 100 observations from this model; the histogram of the sample data is displayed in Fig. 6, left. The maximum $X_{3:3}$ follows a right skewed distribution with PDF

$$f_{3:3}(x) = 3\phi(x)G(x),$$

where $G(x) = \Phi^2(x)$ is a distribution function and Φ is the standard normal distribution. The histogram shows a clear skewed model derived from the baseline normal distribution. In Fig. 6 (right) we also depict the box-plot that shows the asymmetry of the data; it is worthwhile noting that four sample observations are highlighted as possible outliers in a wrong way as they do not come from a normal model. In this case the P–P plot derived from (2.8) is given by the convex function $F_R(F^{-1}(u)) = u^3$ for $u \in [0, 1]$. Note that it is the diagonal section of the product copula of dimension 3.

In the second example we simulate data from the maximum of the dependent random variables of Example 4.5.

Example 5.2 Let us consider the maximum of two ID random variables with a standard normal CDF Φ, that is, $X_{2:2} = \max(X_1, X_2)$, where $X_i \sim N(0, 1)$. Let us assume that they are dependent variables whose dependency is assessed by the Clayton copula of Example 4.5. In order to simulate data from this model, we use the inverse transform method based on

$$F_{2:2}^{-1}(u) = \Phi^{-1}(\delta_C^{-1}(u)),$$
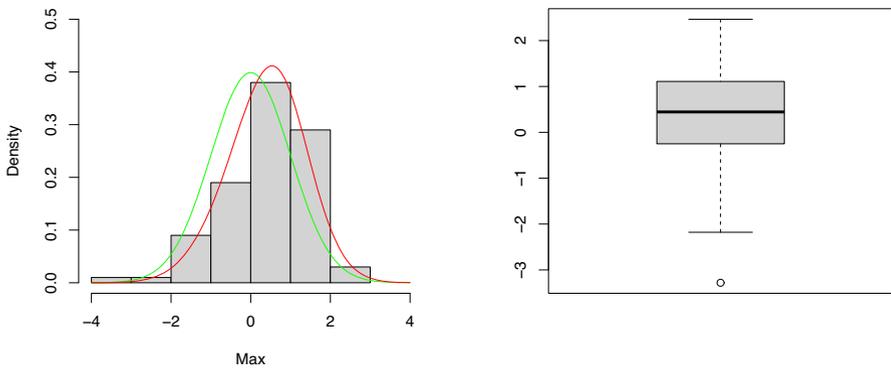
Fig. 7 Histogram (left) made with 100 data from $X_{2:2}$ in Example 5.2 for a standard normal distribution. The green line represent the baseline (marginal) normal PDF and the red line the real skewed PDF $f_{2:2}$. The box-plot (right) shows the weak asymmetry of the data (color figure online)

for $0 < u < 1$, where Φ^{-1} is the quantile function of a standard normal distribution and δ_C^{-1} is the inverse of the diagonal section obtained in Example 4.5 given by

$$\delta_C^{-1}(u) = \frac{2u}{1+u}.$$

Using this expression we draw 100 observations from $X_{2:2}$; the histogram of the sample data is shown in Fig. 7 (left) together with the baseline normal PDF and the theoretical skewed PDF which is given by

$$f_{2:2}(x) = \delta_C'(\Phi(x))\phi(x) = 2\phi(x)G(x),$$

where $G(x) = 1/(2 - \Phi(x))^2$ is a pseudo-distribution function having mass 1/4 at $-\infty$. Figure 7 (right) displays the sample data box-plot which highlights the asymmetry of the underlying model. As in the previous example, one observation is wrongly identified as a potential outlier since the data do not come from a normal model. However, in this case, the asymmetry is weak due to the positive dependence represented by this Clayton copula.

Example 5.3 This example analyzes data about maximum daily temperatures collected during a thirty year period from 1981 to 2010 in the Spanish Iberian Peninsula. The dataset is available from the agroclim R package (Serrano-Notivoli 2022).

We will focus on daily temperatures for July and August summer months; then the maximum monthly temperature values are calculated from the daily records. The histograms for daily and monthly maximum temperatures are shown in Fig. 8, left; it can be noted that daily temperatures exhibit a skewed shape. On the other hand, the monthly values exhibit an apparent change in location as well as a more right skewed behavior than daily ones. The convexity of the P–P plot at the right side of Fig. 8 gives empirical support to assert that the distribution of maximum monthly temperatures is more skewed to the right than the distribution of daily ones, which is indeed a natural fact.
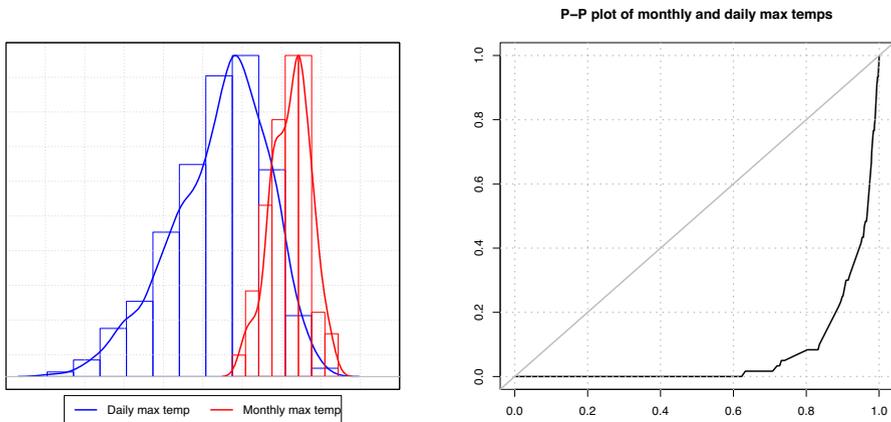
Fig. 8 Histograms of daily and monthly maximum temperatures observed for July and August summer months (left). Empirical P–P plot between daily and monthly maximum temperatures (right)

We now address the comparison of summer temperatures by decades. Our purpose is to elucidate whether there has been a decade with hotter summers on the basis of their comparisons in skewness. To this aim, we consider daily temperatures of both summer months for the following three decades: $1981 - 1990$ (decade 1), $1991 - 2000$ (decade 2) and $2001 - 2010$ (decade 3). The estimated PDFs in Fig. 9 highlight a slightly more right-skewed distribution for decade 2; this is corroborated by the nearly convex shape of P–P plots between decades (1, 2) and decades (3, 2) —see the right panels. Therefore, the analysis with P–P plots highlights hotter summers in the nineties than in the eighties and the first decade of XXI century, with the latter two having nearly equal temperatures (see the left bottom P–P plot of Fig. 9).

6 Concluding remarks

We have studied the relationships between three stochastic strategies that allow to change the shape of an initial baseline distribution; these shape perturbation schemes lead to skewed, weighted and distorted distributions. For the first one, the probability mass of the baseline PDF gets moved to the right or to the left; such a movement is assessed by an order called the likelihood ratio (lr) order which is quite related to the idea of injecting skewness into the baseline model. This is not the case for the other two shape deformation schemes which just represent some "changes" in the baseline initial model, not necessarily related to the idea of skewing a baseline model to get a more right (left) skewed distribution. In this sense we propose two extensions of the classical univariate skewed models, called right and left skewed distributions, which are compatible with the lr stochastic ordering and the underlying idea of comparing distributions in skewness. These models do not assume a symmetric baseline PDF function and can also be seen as weighted or distorted distributions.
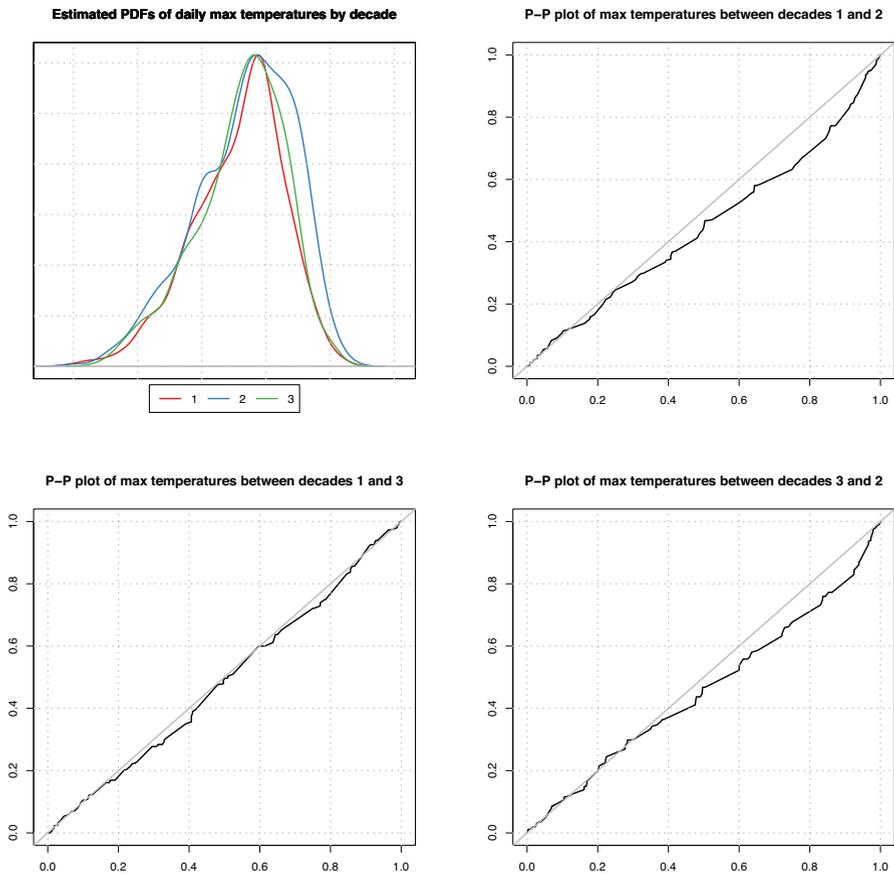
Fig. 9 Histograms of the maximum daily temperature by decade (top left). Empirical P–P plot between decades 1 and 2 (top right), 1 and 3 (bottom left) and 3 and 2 (bottom right) for summer daily temperatures

In the second part of the paper we have proved that the distributions of extremes (maxima or minima) of independent or dependent identically distributed random variables can be represented by means of these models; the weighted and distortion models always can be used for this purpose. The right and left skewed distributions come up very often when modeling extremes, although they may also fail to model a few dependence scenarios as shown in some counterexamples. When they appear as valid models, the lr-ordering between the original data and the associated extreme values will necessarily hold, a fact that can be corroborated graphically by means of P–P plots. Several theoretical and numerical examples have been presented throughout the paper to illustrate these findings.

There are several tasks for future research projects. The main one is concerned with the development of parameter estimation procedures and fit-tests for specific models and extreme data. To this end, we must fix an initial baseline model (e.g. normal or exponential) and then introduce skewness parameters under some operations for

extreme observations; such parameters are also related to the copula used for modeling the data dependence.

On the other hand, some of the univariate distributions of the paper have multivariate extensions which are closed under affine transformations, as for example the skew-normal and the extended skew-normal (Arnold and Beaver 2000) which is a weighted extension of the normal distribution. This fact encourages to look for projections of multivariate skewed vectors in the context of projection pursuit (Huber 1985) and study if there might be a connection between skewness maximization and the likelihood ratio ordering; some ideas on projection pursuit by skewness maximization (Loperfido 2018; Arevalillo and Navarro 2020) may serve to start this project. Finally, the monograph work on skewness and kurtosis orderings (Arnold and Groeneveld 1993) also hints the exploration of kurtosis orderings and multivariate extensions, which may be addressed by kurtosis maximization along the lines of previous work (Loperfido 2020) and seemingly non related results on convex transform orderings (Wang 2009; Arevalillo and Navarro 2012, 2023).

References

Arellano-Valle RB, Genton MG (2008) On the exact distribution of the maximum of absolutely continuous dependent random variables. Stat Probab Lett 78:27–35

Arevalillo JM, Navarro H (2012) A study of the effect of kurtosis on discriminant analysis under elliptical populations. J Multivar Anal 107:53–63

Arevalillo JM, Navarro H (2020) Data projections by skewness maximization under scale mixtures of skew-normal vectors. Adv Data Anal Classif 14:435–461

Arevalillo JM, Navarro H (2023) New insights on the multivariate skew exponential power distribution. Math Slovaca 73(2):529–544

Arevalillo JM, Navarro H (2019) A stochastic ordering based on the canonical transformation of skew-normal vectors. TEST 28:475–498

Arnold BC, Groeneveld RA (1993) Skewness and kurtosis orderings: an introduction. Stochastic Inequalities IMS Lecture Notes - Monograph Series Volume 22

Arnold BC, Groeneveld RA (1995) Measuring skewness with respect to the mode. Am Stat 49:34–38

Arnold BC, Castillo E, Sarabia JM (1999) Conditional specification of statistical models. Springer Series in Statistics. Springer-Verlag, New York

Arnold BC, Beaver RJ (2000) Hidden truncation models. Sankhya Ser A 62:22–35

Arnold BC, Balakrishnan N, Nagaraja HN (2008) A first course in order statistics. SIAM, Philadelphia

Arriaza A, Di Crescenzo A, Sordo MA, Suáirez-Llorens A (2019) Shape measures based on the convex transform order. Metrika 82:99–124

Azzalini A (1985) A class of distributions which includes the normal ones. Scand J Stat 12:171–178

Azzalini A (2005) The skew-normal distribution and related multivariate families. Scand J Stat 32:159–188

Azzalini A, Capitano A (2003) Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. J R Stat Soc B 65:367–389

Azzalini A, Capitanio A (2014) The skew-normal and related families. Cambridge University Press, Cambridge

Azzalini A, Regoli G (2012) Some properties of skew-symmetric distributions. Ann Inst Stat Math 64:857–879

Azzalini A, Regoli G (2018) On symmetry-modulated distributions: revisiting an old result and a step further. Stat 7:e171

Blenkinsop S, Lewis E, Chan SC, Fowler HJ (2017) Quality-control of an hourly rainfall dataset and climatology of extremes for the UK. Int J Climatol 37:722–740

Buishand TA (1989) Statistics of extremes in climatology. Stat Neerl 43:1–30

Chan W, Proschan F, Sethuraman J (1990) Convex-ordering among functions,with applications to reliability and mathematical statistics. In Block HW, Sampson AR, Savits TH (eds) Topics in statistical dependence. IMS Lecture Notes-Monograph Series 16. Hayward, California, pp 121–134

Cox DR (1972) Regression models and life-tables. J R Stat Soc B 34:187–220

David HA, Nagaraja HN (2003) Order statistics, 3rd edn. Wiley, Hoboken

Domma F, Popoviś BV, Nadarajah S (2015) An extension of Azzalini's method. J Comput Appl Math 278:37–47

Durante F, Sempi C (2016) Principles of copula theory. CRC/Chapman & Hall, London

De Luca G, Loperfido N (2004) A skew-in-mean GARCH model for financial returns. In: Skew-elliptical distributions and their applications: a journey beyond normality. CRC/Chapman & Hall, pp 205–222

Fisher RA (1934) The effect of methods of ascertainment upon the estimation of frequencies. Ann Eugen 6:13–25

Ferreira JTAS, Steel MFJ (2006) A constructive representation of univariate skewed distributions. J Am Stat Assoc 101:823–829

Ghosh I, Ng HKT (2019) A class of skewed distributions with applications in environmental data. Commun Stat Case Stud Data Anal Appl 5:346–365

Groeneveld RA, Meeden G (1984) Measuring skewness and kurtosis. The Statistician 33:391–399

Huber PJ (1985) Projection pursuit (with discussion). Ann Stat 13:435–475

Jones MC (2015) On families of distributions with shape parameters. Int Stat Rev 83:175–192

Jupp PE, Regoli G, Azzalini A (2016) A general setting for symmetric distributions and their relationship to general distributions. J Multivar Anal 148:107–119

Ley C (2015) Flexible modelling in statistics: past, present and future. Journal de la Société Française de Statistique 156:76–96

Loperfido N (2008) A note on skew-elliptical distributions and linear functions of order statistics. Stat Probab Lett 78:3184–3186

Loperfido N, Navarro J, Ruiz JM, Sandoval CJ (2007) Some relationships between skew-normal distributions and order statistics from exchangeable normal random vectors. Commun Stat Theory Methods 36:1719–1733

Loperfido N (2018) Skewness-based projection pursuit: a computational approach. Comput Stat Data Anal 120:42–57

Loperfido N (2020) Kurtosis-based projection pursuit for outlier detection in financial time series. Eur J Finance 26:142–164

Navarro J (2022) Introduction to system reliability theory. Springer, Berlin

Navarro J, Calì C, Longobardi M, Durante F (2022) Distortion representations of multivariate distributions. Stat Methods Appl 31:925–954

Navarro J, Ruiz JM, del Aguila Y (2006) Multivariate weighted distributions: a review and some extensions. Statistics 40:51–64

Navarro J, Sordo MA (2018) Stochastic comparisons and bounds for conditional distributions by using copula properties. Depend Model 6:156–177. https://doi.org/10.1515/demo-2018-0010

Navarro J, Torrado N, del Águila Y (2018) Comparisons between largest order statistics from multiple-outlier models with dependence. Methodol Comput Appl Probab 20:411–433

Nelsen RB (2006) An introduction to copulas. Springer, New York

Patil GP, Rao CR (1978) Weighted distributions and size biased sampling with applications to wildlife populations and human families. Biometrics 34:179–189

Rao CR (1965) On discrete distributions arising out of methods of ascertainment. Sankhya Ser A 27:311–324

Roberts C (1966) A correlation model useful in the study of twins. J Am Stat Assoc 61:1184–1190

Schweizer B, Sklar A (1974) Operations on distribution functions not derivable from operations on random variables. Stud Math 52:43–52

Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, New York

Serrano-Notivoli R (2022) agroclim: climatic indices for agriculture. R package version 0.2.0. https://CRAN.R-project.org/package=agroclim

Thas O (2010) Comparing distributions. Springer, New York

Van Zwet WR (1964) Convex transformations of random variables. Mathematish Centrum, Amsterdan

Wang S (1996) Premium calculation by transforming the layer premium density. Astin Bull 26:71–92

Wang J (2009) A family of kurtosis orderings for multivariate distributions. J Multivar Anal 100:509–517

Yaari ME (1987) The dual theory of choice under risk. Econometrica 55:95–115