

相依随机变量一般有限混合的随机比较

王艺婷

西北师范大学数学与统计学院, 甘肃 兰州

收稿日期: 2023年2月15日; 录用日期: 2023年3月14日; 发布日期: 2023年3月22日

摘要

本文讨论了相依随机变量一般有限混合的随机比较, 运用随机序和超优序的理论, 给出了有限相依混合之间的普通随机序和似然比序, 并通过数值例子直观说明了主要结论。

关键词

一般有限混合模型, 相依性, 随机比较

Stochastic Comparisons of Generalized Finite Mixture with Dependent Random Variables

Yiting Wang

School of Mathematics and Statistics, Northwest Normal University, Lanzhou Gansu

Received: Feb. 15th, 2023; accepted: Mar. 14th, 2023; published: Mar. 22nd, 2023

Abstract

In this paper, we carry out stochastic comparisons of generalized finite mixtures with dependent random variables using stochastic orders and majorization order in the sense of the usual stochastic and likelihood ratio ordering and illustrate the main conclusions visually with some numerical examples.

Keywords

Generalized Finite Mixture Model, Dependence, Stochastic Comparison

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

设 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 是一个由 n 个非负随机变量组成的随机向量。 X_1, X_2, \dots, X_n 的有限混合的分布函数为

$$F_{X,\alpha}(x) = \sum_{i=1}^n \alpha_i F_i(x),$$

其中 F_i 是随机变量 X_i 的分布函数, α_i 是混合比例(权重), 对于 $i \in \{1, 2, \dots, n\}$, 有 $\sum_{i=1}^n \alpha_i = 1$ 且 $\alpha_i \geq 0$ 。 同样地, X_1, X_2, \dots, X_n 的有限混合的生存函数和密度函数分别被定义为

$$\bar{F}_{X,\alpha}(x) = \sum_{i=1}^n \alpha_i \bar{F}_i(x), \quad f_{X,\alpha}(x) = \sum_{i=1}^n \alpha_i f_i(x),$$

其中 $\bar{F}_i = 1 - F$ 和 f_i 分别是随机变量 X_i 的生存函数和密度函数。

有限混合模型在可靠性理论、精算科学、医学和经济等多领域有着重要应用。因为在实际生产过程中使用的资源和部件的质量、操作历史和人为失误等各种的原因, 制造出来的元件往往是异质的(见 Finkelstein [1] 和 Cha 和 Finkelstein [2])。因此, 有限混合模型是建模异质随机变量分布的有效工具。比如, 在工业工程中, 导致项目样本失败的原因不止一个, 每个原因的失效分布都可以用一个简单的密度函数来充分近似, 如负指数; 从而总体分布可以建模为负指数随机变量的有限混合。

在过去的几年里, 一些学者致力于研究有限混合模型的随机特性, 但大部分研究是基于所有随机变量之间是独立的假设下。例如, Hazra 和 Finkelstein [3] 讨论了当两个有限混合模型的随机变量的分布都服从比例失效率(HR)模型、比例反失效率(PHR)模型和加速寿命(AL)模型其中的之一的时候, 系统的一些随机比较结果。Amini-Seresht 和 Zhang [4] 针对两个有限混合模型分别在不同基分布和不同混合比例下研究了在失效率序、反失效率序、似然比序下的随机比较结果。其他更多相关结果可参考 Nadeb 和 Torabi [5]、Navarro [6]、Franco 和 Balakrishnan [3] 的研究。

然而, 在大多数实际情况下, 一个系统的组件通常共享相同的生存环境, 这导致组件之间的具有相依性。Amini-Seresht 和 Balakrishnan [7] 考虑了由相依同分布(d.i.d.)随机变量构成的一般有限混合模型, 并给出了该模型在反失效率和失效率序方面的随机结果。但对于其他随机序的结果并没有给出, 这值得我们进一步探讨。

受 Hazra 和 Misra [8]、Amini-Seresht 和 Balakrishnan [7] 和 Shojaee 等人 [9] 工作的影响, 本文将重点关注相依随机变量的一般有限混合模型, 研究该模型在普通随机序和似然比序下的随机比较结果。这些结果不仅补充了 Amini-Seresht 和 Balakrishnan [7] 的文献结果, 弥补了现阶段文章仅研究随机变量间为独立情况, 而且在可靠性理论中也有潜在的应用。本文的主要内容安排如下: 第一节中, 给出有限混合模型的定义以及国内外研究现状。第二节中, 介绍本文研究中需要的一些基本理论知识。第三节给出由不同权重的随机向量组成的两个有限混合模型, 在普通随机序和似然比序下的随机比较结果, 并给出数值例子。

2. 预备知识

2.1. 随机序

随机序理论是用于两个随机变量进行随机比较的有效工具。由于其在科学和工程的不同分支中的应

用, Shaked 和 Shanthikumar [10]对这一理论进行了详细的概述。首先给出随机序的基本概念。

对于任意两个非负绝对连续的随机变量 X 和 Y , 分布函数分别记为 F 和 G , 生存函数分别记为 $\bar{F}=1-F$ 和 $\bar{G}=1-G$, 密度函数分别记为 f 和 g , 失效率函数分别记为 $r_X = f/F$ 和 $r_Y = g/G$, 反失效率函数分别记为 $\tilde{r}_X = f/F$ 和 $\tilde{r}_Y = g/G$, 则

- (i) 若对于任意的 $t \in R$, 都有 $\bar{F}(t) \leq \bar{G}(t)$, 则称 X 在普通随机序下小于 Y , 记作 $X \leq_{st} Y$;
- (ii) 若对于任意的 $t \in R$, $\bar{G}(t)/\bar{F}(t)$ 关于 t 是递增的, 则称 X 在失效率序下小于 Y , 记作 $X \leq_{hr} Y$;
- (iii) 若对于任意的 $t \in R$, $G(t)/F(t)$ 关于 t 是递增的, 则称 X 在反失效率序下小于 Y , 记作 $X \leq_{rh} Y$;
- (iv) 若对于任意的 $t \in R$, $g(t)/f(t)$ 关于 t 是递增的, 则称 X 在似然比序下小于 Y , 记作 $X \leq_r Y$ 。

失效率序与反失效率序都蕴含了普通随机序; 而失效率序与反失效率序之间并没有必然的联系, 由普通随机序既不能得失效率序也不能得反失效率序。

2.2. 超优序

在统计学中, 超优序已经被广泛的应用于建立各种不等式并且起着关键作用, 下面给出超优序的定义。

设 $\mathbf{a} = (a_1, a_2, \dots, a_n)$ 和 $\mathbf{b} = (b_1, b_2, \dots, b_n)$ 是两个 n 维实值向量, 记 $a_{(1)} \leq a_{(2)} \leq \dots \leq a_{(n)}$ 和 $b_{(1)} \leq b_{(2)} \leq \dots \leq b_{(n)}$ 分别为 \mathbf{a} 和 \mathbf{b} 的递增排列,

(i) 若对于任意的 $i = 1, 2, \dots, n-1$, 都有 $\sum_{j=1}^i a_{(j)} \geq \sum_{j=1}^i b_{(j)}$ 并且 $\sum_{j=1}^n a_{(j)} = \sum_{j=1}^n b_{(j)}$, 则称向量 \mathbf{b} 超优于向量 \mathbf{a} , 记作 $\mathbf{a} \preceq^m \mathbf{b}$ 。

(ii) 若对于任意的 $i = 1, 2, \dots, n$, 都有 $\sum_{j=1}^i a_{(j)} \geq \sum_{j=1}^i b_{(j)}$, 则称向量 \mathbf{b} 弱上优于向量 \mathbf{a} , 记作 $\mathbf{a} \preceq^w \mathbf{b}$ 。

更多详细有关超优序的性质及应用, 请参阅[11]。

2.3. Copula 函数

Copula 函数是描述随机变量之间相依的数学概念。自 Nelsen [12]引入以来, 其广泛应用于可靠性和保险理论等领域, 从而受到广泛关注。Copula 理论的优势在于将多元随机向量的边际分布和联合分布函数联接起来。因此十分适合描述随机变量之间的相依关系。下面给出 Copula 函数的定义。

设随机变量 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 具有联合分布函数 F , 联合生存函数 \bar{F} , 边际分布函数 F_1, F_2, \dots, F_n 和边际生存函数 $\bar{F}_1, \bar{F}_2, \dots, \bar{F}_n$ 。若存在函数 $\mathbf{C} : [0, 1]^n \rightarrow [0, 1]$ 和 $\hat{\mathbf{C}} : [0, 1]^n \rightarrow [0, 1]$ 的联合分布函数使得, 对所有的 $x_i, i = 1, 2, \dots, n$, 有

$$F_{\mathbf{X}}(\mathbf{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n) = \mathbf{C}(F(x_1), F(x_2), \dots, F(x_n)),$$

$$\bar{F}_{\mathbf{X}}(\mathbf{x}) = P(X_1 > x_1, X_2 > x_2, \dots, X_n > x_n) = \hat{\mathbf{C}}(\bar{F}(x_1), \bar{F}(x_2), \dots, \bar{F}(x_n)),$$

则将 \mathbf{C} 和 $\hat{\mathbf{C}}$ 分别称为随机变量 \mathbf{X} 的 Copula 和生存 Copula。

令 $\bar{K}_i(F(\mathbf{x})) = \hat{\mathbf{C}}(\bar{F}(x)\mathbf{1}_i, \mathbf{1}_{n-i})$ 表示串联系统 $X_{1:i} = \min(X_1, X_2, \dots, X_i)$ 的生存函数, 其中 $\mathbf{1}_i$ 和 $\mathbf{1}_{n-i}$ 均为 1, 并且 $\bar{K}_1(\bar{F}(x)) = \bar{F}(x)$ 和 $\bar{K}_n(\bar{F}(x)) = \hat{\mathbf{C}}(\bar{F}(x), \dots, \bar{F}(x))$ 。

由此, 相依随机变量构成的一般有限混合模型的生存函数可以写成(混合比例可能取负值),

$$\bar{H}_{\mathbf{X}, \mathbf{a}}(\bar{F}(x)) = \sum_{i=1}^n a_i \bar{K}_i(\bar{F}(x)),$$

其中 $\mathbf{a} = (a_1, a_2, \dots, a_n)$, 并且 a_i 是实数使得 $\sum_{i=1}^n a_i = 1$ 。

3. 主要结果

本节主要研究由不同权重的随机向量组成的两个有限混合模型，在普通随机序和似然比序下的随机比较结果，并给出数值例子。

定理 1 设 $H_{X,a}$ 和 $H_{X,b}$ 是两个权重分别为 \mathbf{a} 和 \mathbf{b} 且相依随机变量为 X 构成的一般有限混合模型。若

(i) 对于任意的 $1 \leq i \leq j \leq n$, $a_i \geq a_j$,

(ii) $\mathbf{a} \stackrel{m}{\preceq} \mathbf{b}$,

(iii) 对于任意的 $1 \leq i \leq n$, $\bar{K}_i(\bar{F}(x))$ 关于 i 是递减的, 那么

$$H_{X,a} \leq_{st} H_{X,b}$$

证明: 用 $\bar{F}_{X,a}(t)$ 和 $\bar{F}_{X,b}(t)$ 分别表示 $H_{X,a}$ 和 $H_{X,b}$ 的生存函数, 那么

$$\phi(t, \mathbf{a}) = \bar{F}_{X,a}(t) = \sum_{i=1}^n a_i \bar{K}_i(\bar{F}(t)),$$

$$\phi(t, \mathbf{b}) = \bar{F}_{X,b}(t) = \sum_{i=1}^n b_i \bar{K}_i(\bar{F}(t)).$$

现只需证明 $\phi(t, \mathbf{a})$ 关于 \mathbf{a} 是舒尔凸的。

$$\frac{\partial \phi(t, \mathbf{a})}{\partial a_i} = \bar{K}_i(\bar{F}(t)),$$

$$\frac{\partial \phi(t, \mathbf{a})}{\partial a_j} = \bar{K}_j(\bar{F}(t)),$$

$$(a_i - a_j) \left[\frac{\partial \phi(t, \mathbf{a})}{\partial a_i} - \frac{\partial \phi(t, \mathbf{a})}{\partial a_j} \right] = (a_i - a_j) (\bar{K}_i(\bar{F}(t)) - \bar{K}_j(\bar{F}(t))) \geq 0,$$

从而定理成立。

下面给出一个数值例子来验证定理 1。

例 1 考虑存在两个系统分别为 $T_1 = \min(X_1, X_2, X_3)$ 和 $T_2 = \max(X_1, X_2, X_3)$, 其中随机向量 $\mathbf{X} = (X_1, X_2, X_3)$ 服从 FGM copula 相依,

$$\hat{C}(u_1, u_2) = \prod_{i=1}^n u_i \left[1 + \theta \prod_{i=1}^n (1 - u_i) \right], \quad -1 \leq \theta \leq 1.$$

当 $\mathbf{a} = (0.7, 0.2, 0.1)$ 和 $\mathbf{b} = (0.5, 0.3, 0.2)$ 的时候, 可以验证, 对于任意的 $1 \leq i \leq j \leq 3$, $\mathbf{a} \stackrel{m}{\preceq} \mathbf{b}$ 且 $a_i \geq a_j$, 得到定理 1 的假设(i)和(ii)均被满足。因此, 只需检查定理 1 的假设(iii)成立即可。

$$\bar{K}_1(u) = u,$$

$$\bar{K}_2(u) = u^2,$$

$$\bar{K}_3(u) = u^3 [1 + \theta(1-u)^3]$$

$$\bar{K}_1(u) - \bar{K}_2(u) = u - u^2 = u(1-u) > 0,$$

$$\bar{K}_2(u) - \bar{K}_3(u) = u^2 \left[1 - u [1 + \theta(1-u)^3] \right] \stackrel{\text{sgn}}{\cong} 1 - u [1 + \theta(1-u)^3] \triangleq \mu(u; \theta).$$

下面检验 $\mu(u; \theta)$ 的非负性。

情况 1: $u \in [0, 1]$, $\theta \in [-1, 0]$ 。可以验证对于任意的 $\theta \in [-1, 0]$, $\mu(u; \theta)$ 关于 u 是递减的, 从而

$$\mu(u; \theta) \geq \mu(1; \theta) = 0.$$

情况 2: $u \in [0,1]$, $\theta \in [0,1]$ 。注意到对于任意的 $\theta \in [0,1]$, $\mu(u;\theta)$ 关于 u 是递增的, 进一步地

$$\mu(u;\theta) \geq \mu(0;\theta) = 1.$$

综合上述关系, 对于任意的 $u \in [0,1]$, $\theta \in [-1,1]$, 可以得到 $\bar{K}_2(u) - \bar{K}_3(u)$ 关于 u 是非负的。因此, 假设(iii)条件满足。

定理 2 设 $H_{X,a}$ 和 $H_{X,b}$ 是两个权重分别为 \mathbf{a} 和 \mathbf{b} 且相依随机变量为 \mathbf{X} 构成的一般有限混合模型。若

(i) 对于任意的 $1 \leq j \leq n$, $\frac{(1-u)K_j''(u)}{K_j'(u)}$ 关于 j 是递增的,

(ii) 对于任意的 $1 \leq i \leq j \leq n$, $a_i b_j \leq a_j b_i$, 那么

$$H_{X,a} \leq_{lr} H_{X,b}.$$

证明: 分别用 $f_{X,a}(x)$ 和 $f_{X,b}(x)$ 表示 $H_{X,a}$ 和 $H_{X,b}$ 的密度函数, 从而得到

$$f_{X,a}(x) = \left[\sum_{i=1}^n a_i K_i(\bar{F}(x)) \right]' = - \sum_{i=1}^n a_i K_i'(\bar{F}(x)) f(x),$$

$$f_{X,b}(x) = \left[\sum_{j=1}^n b_j K_j(\bar{F}(x)) \right]' = - \sum_{j=1}^n b_j K_j'(\bar{F}(x)) f(x).$$

当且仅当

$$\psi(x) = \frac{f_{X,b}(x)}{f_{X,a}(x)} = \frac{\sum_{j=1}^n b_j K_j'(\bar{F}(x))}{\sum_{i=1}^n a_i K_i'(\bar{F}(x))}$$

关于非负 x 单调递增时, 定理成立。

$$\begin{aligned} \psi'(x) &\stackrel{\text{sgn}}{=} \sum_{i=1}^n \sum_{j=1}^n a_i b_j ij \left[K_j'(\bar{F}(x)) K_i''(\bar{F}(x)) i f'(x) - K_i'(\bar{F}(x)) K_j''(\bar{F}(x)) j f'(x) \right] \\ &= \tilde{r}_X(x) \sum_{i=1}^n \sum_{j=1}^n a_i b_j ij \left[i \frac{F(x) K_i''(\bar{F}(x))}{K_i'(\bar{F}(x))} - j \frac{F(x) K_j''(\bar{F}(x))}{K_j'(\bar{F}(x))} \right] K_i'(\bar{F}(x)) K_j'(\bar{F}(x)) \\ &= \tilde{r}_X(x) \sum_{i=1}^n \sum_{j=1}^n a_i b_j ij \left[i \frac{F(x) K_i''(\bar{F}(x))}{K_i'(\bar{F}(x))} - j \frac{F(x) K_j''(\bar{F}(x))}{K_j'(\bar{F}(x))} \right] K_i'(\bar{F}(x)) K_j'(\bar{F}(x)) \\ &\quad + \tilde{r}_X(x) \sum_{i=1}^n \sum_{j=1}^n a_j b_i ij \left[j \frac{F(x) K_j''(\bar{F}(x))}{K_j'(\bar{F}(x))} - i \frac{F(x) K_i''(\bar{F}(x))}{K_i'(\bar{F}(x))} \right] K_j'(\bar{F}(x)) K_i'(\bar{F}(x)) \\ &= \tilde{r}_X(x) \sum_{i=1}^n \sum_{j=1}^n [a_j b_i - a_i b_j] ij \left[j \frac{F(x) K_j''(\bar{F}(x))}{K_j'(\bar{F}(x))} - i \frac{F(x) K_i''(\bar{F}(x))}{K_i'(\bar{F}(x))} \right] K_i'(\bar{F}(x)) K_j'(\bar{F}(x)) \\ &\triangleq \tilde{r}_X(x) \sum_{i=1}^n \sum_{j=1}^n [a_j b_i - a_i b_j] ij \cdot \Lambda^* \cdot K_i'(\bar{F}(x)) K_j'(\bar{F}(x)) \end{aligned}$$

其中 $\Lambda^* = j \frac{F(x) K_j''(\bar{F}(x))}{K_j'(\bar{F}(x))} - i \frac{F(x) K_i''(\bar{F}(x))}{K_i'(\bar{F}(x))}$ 。

现只需证明 Λ^* 是非负的。令 $\bar{F} = u$, 从而 Λ^* 可以表示为

$$\Lambda^* = j \frac{(1-u) K_j''(u)}{K_j'(u)} - i \frac{(1-u) K_i''(u)}{K_i'(u)} \geq i \left[\frac{(1-u) K_j''(u)}{K_j'(u)} - \frac{(1-u) K_i''(u)}{K_i'(u)} \right] \geq 0,$$

从而定理成立。

为了验证定理 2 的有效性，下面给出一个数值例子。为了使结果更加直观、清晰，减轻计算量，用图像来辅助验证。

例 2 考虑存在两个系统分别为 $T_1 = \min(X_1, X_2)$ 和 $T_1 = \max(X_1, X_2)$ ，其中随机向量 $X = (X_1, X_2)$ 服从 FGM copula 相依，并且 $\mathbf{a} = (0, 1)$ 和 $\mathbf{b} = (3, -2)$ ，经过计算得到，对于任意的 $1 \leq i \leq j \leq 2$ ， $a_i b_j \leq a_j b_i$ 。从而定理 2 的假设(ii)成立，因此，只需验证假设(i)是否成立。

通过计算，可得

$$\begin{aligned}
 K_1(u) &= 1-u, \quad K_1'(u) = -1, \quad K_1''(u) = 0, \\
 K_2(u) &= 1-u^2 [1+\theta(1-u)^2], \\
 K_2'(u) &= 2u^2\theta(1-u) - 2u [1+\theta(1-u)^2], \\
 K_2''(u) &= 2\theta u(2-3u) - 2 [1+\theta(1-u)^2] + 4u\theta(1-u), \\
 \frac{(1-u)K_1''(u)}{K_1'(u)} &= 0, \\
 \frac{(1-u)K_2''(u)}{K_2'(u)} &= \frac{(1-u)[2\theta(4-5u) - 2[1+\theta(1-u)^2]]}{2u^2\theta(1-u) - 2u [1+\theta(1-u)^2]},
 \end{aligned}$$

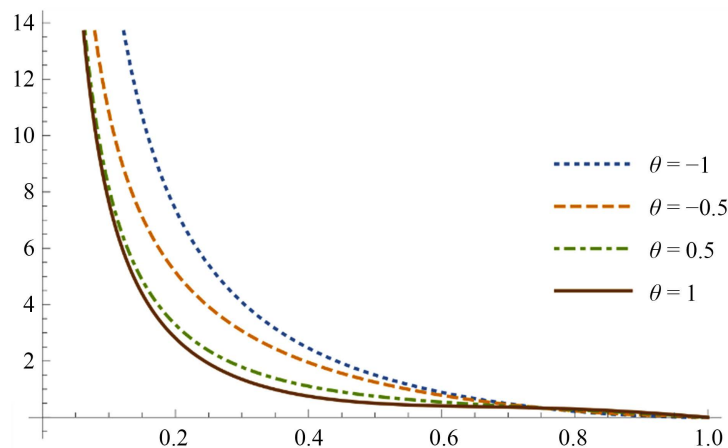


Figure 1. Plot of functions $\frac{(1-u)K_2''(u)}{K_2'(u)}$, for $\theta = -1, -0.5, 0.5, 1$

图 1. 当 $\theta = -1, -0.5, 0.5, 1$ 时， $\frac{(1-u)K_2''(u)}{K_2'(u)}$ 的函数图像

由图 1 所示，函数 $\frac{(1-u)K_2''(u)}{K_2'(u)}$ 在 $u \in [0, 1]$ 时，函数值均为正。由此可得，该例子满足定理 2。

参考文献

[1] Finkelstein, M. (2008) Failure Rate Modelling for Reliability and Risk. Springer Science & Business Media, Berlin.

-
- [2] Cha, J.H. and Finkelstein, M. (2013) The Failure Rate Dynamics in Heterogeneous Populations. *Reliability Engineering & System Safety*, **112**, 120-128. <https://doi.org/10.1016/j.res.2012.11.012>
- [3] Franco, M., Balakrishnan, N., Kundu, D. and Vivo, J.-M. (2014) Generalized Mixtures of Weibull Components. *Test*, **23**, 515-535. <https://doi.org/10.1007/s11749-014-0362-x>
- [4] Amini-Seresht, E. and Zhang, Y.Y. (2017) Stochastic Comparisons on Two Finite Mixture Models. *Operations Research Letters*, **45**, 475-480. <https://doi.org/10.1016/j.orl.2017.07.009>
- [5] Nadeb, H. and Torabi, H. (2022) New Results on Stochastic Comparisons of Finite Mixtures for Some Families of Distributions. *Communications in Statistics-Theory and Methods*, **51**, 3104-3119. <https://doi.org/10.1080/03610926.2020.1788082>
- [6] Navarro, J. (2016) Stochastic Comparisons of Generalized Mixtures and Coherent Systems. *Test*, **25**, 150-169. <https://doi.org/10.1007/s11749-015-0443-5>
- [7] Amini-Seresht, E. and Balakrishnan, N. (2021) Stochastic Properties of Generalized Finite Mixture Models with Dependent Components. *Journal of Applied Probability*, **58**, 794-804. <https://doi.org/10.1017/jpr.2021.4>
- [8] Hazra, N.K. and Misra, N. (2020) On Relative Ageing of Coherent Systems with Dependent Identically Distributed Components. *Advances in Applied Probability*, **52**, 348-376. <https://doi.org/10.1017/apr.2019.63>
- [9] Shojaei, O., Asadi, M. and Finkelstein, M. (2022) Stochastic Properties of Generalized Finite α -Mixtures. *Probability in the Engineering and Informational Sciences*, **36**, 1055-1079. <https://doi.org/10.1017/S0269964821000243>
- [10] Shaked, M. and George Shanthikumar, J. (2007) *Stochastic Orders*. Springer, Berlin. <https://doi.org/10.1007/978-0-387-34675-5>
- [11] Marshall, A.W., Olkin, I. and Arnold, B.C. (1979) *Inequalities: Theory of Majorization and Its Applications*. Springer, Berlin, Vol. 143.
- [12] Nelsen, R.B. (2007) *An Introduction to Copulas*. Springer Science & Business Media, Berlin.