法政大学学術機関リポジトリ
HOSEI UNIVERSITY REPOSITORY

# Statistical modeling and reliability analysis for multi-component systems with dependent failures

| | |
|---|---|
| 著者 | OTA Shuhei |
| 著者別名 | 太田 修平 |
| page range | 1-91 |
| year | 2019-03-24 |
| 学位授与番号 | 32675甲第457号 |
| 学位授与年月日 | 2019-03-24 |
| 学位名 | 博士(工学) |
| 学位授与機関 | 法政大学 (Hosei University) |
| URL | http://doi.org/10.15002/00021753 |

# Statistical modeling and reliability analysis for multi-component systems with dependent failures

Dissertation submitted in partial fulfillment for
the degree of Doctor of Engineering

**Shuhei Ota**

Under the supervision of
Professor Mitsuhiro Kimura

Major of Systems Engineering and Science,
Graduate School of Science and Engineering,
Hosei University

Tokyo, Japan

March 2019

# Abstract

Reliability analysis of systems based on component reliability models has drawn the great interest of many researchers so far, as one of the fundamental aspects of reliability assessment issues. In particular, reliability analysis considering dependent failure occurrences of system components is important because the components may fail mutually due to sharing workloads such as heat, tasks and so on. In such a situation, we are liable to incorrectly estimate the reliability of the system unless we consider the possibility of the dependent failure occurrence phenomena. Thus, there are many publications about this topic in the literature. Most of the existing studies deal with the dependent failure between any two components in a multi-component system since its mathematical formulation is comparatively easy. However, the dependent failure may occur among two or more components in actual cases.

In this thesis, we aim at developing reliability analysis techniques when several components of a system break down dependently. First, we newly formulate a reliability model of systems with the dependent failure by using a multivariate Farlie-Gumbel-Morgenstern (FGM) copula. Based on the model, we investigate the effect of the dependent failure occurrence on the system's reliability. Secondly, we deal with the parameter estimation for the model in order to evaluate the dependence among the components by using their failure times. To do so, we propose a useful estimation algorithm for the multivariate FGM copula. In addition, we theoretically reveal the asymptotic normality of the proposed estimators and numerically investigate the estimation accuracy. Finally, we present a new method for the detection of the dependent failure occurrence in an $n$-component parallel system. These results are helpful to both quantitative and qualitative reliability assessment of the system under the possibility of the dependent failure occurrences. Also, our estimation method is especially applicable not only the reliability analysis but also other research fields.

# Acknowledgments

I would like to record my deepest gratitude to my thesis supervisor, Professor Mitsuhiro Kimura, for his invaluable support and kind advice during all the time of research and writing of this thesis. I appreciate that he always gave me interesting research ideas and patiently supervised. Also, I would like to thank him for his financial support through a KAKEN grant for my study.

I would like to express my sincere gratitude to my dissertation committee: Professor Hisashi Yamamoto of Tokyo Metropolitan University and Associate Professor Nobuyuki Tamura of Hosei University for their support, valuable feedback and insightful ideas to this research.

I am grateful to the staffs at Department of Industrial and Systems Engineering in Hosei University for their kind support. Also, I would like to appreciate financial support from Hosei University for my study.

Special thanks to past and present members of Reliability Engineering Laboratory in Hosei University for providing a pleasant and stimulating research environment.

Finally, I wish to thank my parents, brothers and relatives for their continuous encouragement and support to my life over many years.

March, 2019
Shuhei Ota
太田 修平

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Reliability engineering

According to O'Connor, Newton and Bromley [42], reliability engineering originated in the United States during the 1950's as a separate engineering discipline. In its early stage, techniques of reliability engineering had been developed to improve the reliability of military systems. After that, it continuously had evolved through space exploration programs. For example, in Apollo program (1961-1972), reliability approaches had been concerned as one of the most significant factors to assure the success of the mission [55]. Today, reliability engineering has contributed to the improvement of the reliability of various engineering systems such as electronic devices, communication systems, software systems and so on. In this way, the more systems support our lives, the more important reliability techniques are.

In general, systems, components and items eventually break down if we continue to use them. Therefore, designers, manufacturers, and end users would like to often maximize system performance and efficiently use their given resources due to cost reduction, assuring safety and so on. To do so, it is important to know how often failures may occur. Then, this topic leads to predicting the frequency of the failure occurrences in the system. However, because we cannot fully analyze most failure mechanisms in a particular design, the prediction of failures is associated with uncertainty. That is, it is inherently a probabilistic problem. Therefore, mathematical and statistical methods can be used for reliability analysis.

Reliability analysis based on component reliability models has drawn the great interest of many researchers so far, as one of the fundamental aspects of reliability assessment issues. Thus, there are many publications on this topic in the literature. For instance, Carhart [3] is one of the oldest contributions which surveyed the reliability analysis for series and parallel systems with identical components. Also, Creveling [8] and Moore and Shannon [34] dealt with more complex systems like series-parallel and network systems for instance. Zelen [62] well summarized the reliability analysis for redundant systems including such systems. However, most of these studies assumed that the lifetimes of the components are statistically independent.

## 1.2 Dependent failure analysis

In actual practice, system components may fail mutually due to sharing workloads such as heat, tasks and so on [32, 56, 64]. Such a failure is called dependent failure. For example, the dependent failure was observed in the space development programs in the 1960's [56]. During the reentry of Gemini spacecraft, one of the two guidance computers failed, and a few minutes later the other one also failed. This is because the temperature inside the two computers was much higher than expected. That is, the two computers failed dependently due to sharing the heat. In general, the more complex a system is, the more likely the dependent failure occurs in the system.

The major problem of the dependent failure is that it can cause severe deterioration of the reliability of the redundant system [32]. Moreover, we are liable to incorrectly estimate the reliability of the system unless we consider the possibility of the dependent failure occurrence. Therefore, reliability analysis considering the dependence among the components is required.

The widely acceptable definition of the dependent failure is a set of events in which the probability of each failures is dependent on the occurrence of the other failures [33]. In general, the major causes of the dependent failures among systems or components can be summarized as shown in Tab. 1.1. For example, the case of Gemini spacecraft [56] is the dependent failure of intersystem due to the physical reason.

Table 1.1: Types of dependent events (cf. [33], on page 390).

| Dependent Event Type | Dependent Event Category | Subcategory | Example |
|---|---|---|---|
| Internal | Challenge (e.g., stress and damage) | – | Internal transients or deviations from the normal operating envelope introduce a challenge to a number of items. |
| | Intersystem (failure between two or more systems) | Functional | Power to several independent system is from the same source. |
| | | Shared equipment | The same equipment, for example, a valve, is shared between otherwise independent systems. |
| | | Physical | The extreme equipment, for example, high temperature, causes dependencies between independent systems. |
| | | Human | Operator error causes failure of two or more independent systems. |
| | Intercomponent (failure between two or more components) | Functional | A component in a system provides multiple functions. |
| | | Shared equipment | Two independent trains in hydraulic system share the same common header. |
| | | Physical | Same as system interdependency above. |
| | | Human | Design errors in redundant pump controls introduce a dependency in the system. |
| External | – | – | Earthquake or fire fails a number of independent systems or components. |

The effect of the dependent failure occurrence on the system reliability has been investigated by many researchers. For example, Meeker and Escobar [32] indicated that the dependence between two components deteriorates/improves the system reliability. They assumed that the distribution of logarithm of the failure times of the individual components is bivariate normal distribution with the same mean, the same deviation and a linear correlation $\rho$. They stated that when there is the positive correlation between the failure times of the individual components, the designed reliability of the series system exceeds the reliability that is predicted by the 2-independent-component series system as shown in Fig. 1.1. On the other hand, Fig. 1.2 illustrates that the reliability of the 2-component parallel system declines when the failure times of the individual components have the positive correlation. As a result, the paper suggested that the multivariate generalization of these results is important in reliability modeling application.

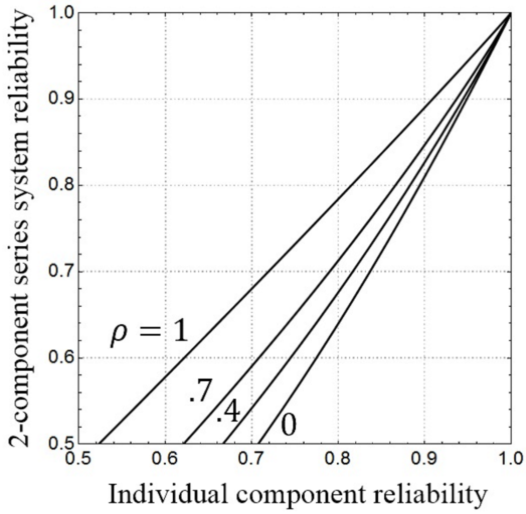


Figure 1.1: Reliability of the 2-dependent-component series system with a linear correlation $\rho$ [32].

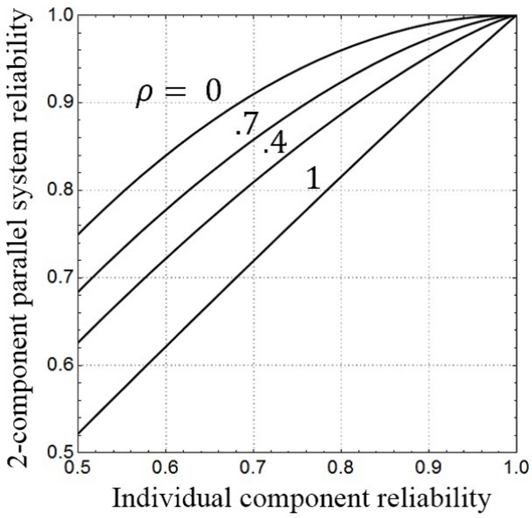


Figure 1.2: Reliability of the 2-dependent-component parallel system with a linear correlation $\rho$ [32].

An early model of the dependent failure in multi-component systems was proposed by Marshall and Olkin [30]. Their model describes common cause failure in a multi-component system by using a multivariate exponential distribution. O'Connor and Mosleh [41] applied a Bayesian network model for common cause failure in order to consider the coupling mechanism of components. Humphreys and Jenkins [20] investigated analysis methods and a cause classification scheme for the dependent failures.

In the last decade, reliability analysis for the dependent failure by using copulas is one of the challenging research topics. The copula is a function that joins several one-dimensional distribution functions to form a multivariate distribution function with dependency. Also the copula works as a multivariate distribution function by itself. For examples, Eryilmaz and Tank [11] showed that the positive dependence between the two components improves the mean time to failure (MTTF) of the 2-component series system by the FGM copula and Ali-Mikhail-Haq copula. Ota and Kimura [46] investigated the effect of the dependence into the reliability of an $n$-component parallel system and series one in the same fashion. Navarro, Ruiz and Sandoval [39] developed a reliability model for coherent systems with $n$ dependent components by many kinds of copulas.

Most of the existing studies focus on the dependent failure between any two components in order to simplify the complexity of the mathematical models. However, the dependent failure may occur among two or more components in actual cases.

Therefore, in this thesis, we aim at developing reliability analysis techniques when several components in a system break down dependently. In this thesis, we develop two statistical models of systems. The first one is based on the multivariate FGM copula. The second one is based on a switching mechanism of hazard rate functions. By using these models, we evaluate multivariate dependencies among system components. They are helpful to both quantitative and qualitative reliability assessment of the system under the possibility of the dependent failure occurrences.

## 1.3 Organization of this thesis

This thesis is divided into six chapters, Appendix and Bibliography. In detail, Chapter 2 mainly provides mathematical preliminaries and new results of a copula. In this study, we use the FGM copula as a modeling tool of a pendent failure-occurrence environment. The FGM copula is one of the copulas and a multivariate distribution with high flexibility. At first, this chapter provides the definition and some fundamental properties of the copula. Then, we introduce several unique features of the FGM copula. Finally, we newly reveal the necessary conditions of the exact ranges of the dependence parameters of the minimum and maximum value distributions which are based on the FGM copula and show their asymptotic properties. This result is given by Ota and Kimura [49].

In Chapter 3, we develop reliability models of a parallel system and series system with $n$ dependent components by using the multivariate FGM copula. Then, we investigate several reliability-related properties of the systems. As a result, we reveal that the $n$-component parallel system cannot deliver its designed reliability if the lifetimes of the individual components have positive dependence. On the other hand, we present that the $n$-component series system can exceed its designed reliability under such dependent failure-occurrence environment. This result is given by Kimura, Ota and Abe [27] and Ota and Kimura [46].

In Chapter 4, we present an estimation algorithm for the model parameter and refer to its asymptotic normality. More specifically, we deal with the parameter estimation for the $d$-variate FGM copula which consists of $2^d - d - 1$ dependence parameters to be estimated. We propose a new estimation method of the parameters of the FGM copula by using the theory of the inference function for margins [25]. Although the ordinary maximum likelihood estimation is computationally infeasible for a large number $d$, our method is feasible for the same situation. Then, we analytically show its asymptotic property. Finally, we demonstrate the performance of the proposed method through simulation studies. This result is given by Ota and Kimura [47] and Ota and Kimura [48].

In Chapter 5, we deal with detection of a dependent failure occurrence in an $n$-component parallel system. Making a system redundant by combining identical components is a useful way to ensure a highly reliable system. However, the components of such systems may fail mutually, and if the components break down dependently, the reliability of the system decreases. Therefore, reliability analysis considering the dependence among the components is important in reliability assessment. In this chapter, we propose a statistical detection method of the dependent failure occurrence in $n$-component parallel systems by using the failure occurrence times of the components. If we assume that the lifetime distribution of the components worsens if $k$ out of $n$ components failed, the dependent failure occurrence can be found by identifying the change of the distribution. Finally, the performance of the proposed method is demonstrated by simulation studies. This results is given by Ota and Kimura [43] and Ota and Kimura [45].

In Appendix, the mathematical proofs of principal theorems and corollaries derived in Chapter 2 and Chapter 4 are given.

# Chapter 2

# Copula

In Chapters 3 and 4, we deal with the FGM (Farlie-Gumbel-Morgenstern) copula as a modeling tool of a dependent failure-occurrence environment. The FGM copula is one of the copulas and a multivariate distribution with high flexibility. As the preliminaries, this chapter provides the definition and some fundamental properties of the copula at first. Then, we introduce several unique features of the FGM copula. Finally, we newly reveal the necessary conditions of the ranges of the dependence parameters of the minimum and maximum value distributions which are based on the FGM copula and show the asymptotic properties of the ranges. This result is given by Ota and Kimura [49].

## 2.1 Copula

The copula is defined as a multivariate distribution function as follows [40].

**Definition 2.1.** *For a d-variate distribution function $H$, with univariate marginal distribution functions $F_1, \ldots, F_d$, we say that $C : [0,1]^d \to [0,1]$ associated with $H$ is a copula if*

$$H(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \tag{2.1}$$

*for $x_1, \ldots, x_d \in \mathbb{R}$.*

Here, we introduce an important theorem of copulas, Sklar's theorem [40].

**Theorem 2.1.** *Let $H$ be a d-variate distribution function with univariate marginal distribution functions $F_1, \ldots, F_d$. If $F_1, \ldots, F_d$ are continuous functions, then the* *following relationship is uniquely determined.*

$$C(u_1, \ldots, u_d) = H(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)), \tag{2.2}$$

*for* $u_1, \ldots, u_d \in [0, 1]$.

Hence, for many theoretical questions about copulas, one can concentrate on that all marginal distributions are the uniform distribution in the interval $[0, 1]$.

If $C(u_1, \ldots, u_d)$ is a continuous function, then its density function is given by

$$c(u_1, \ldots, u_d) = \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \cdots \partial u_d}. \tag{2.3}$$

Next, we consider the joint density function of Eq. (2.1) with univariate marginal distribution functions $F_1, \ldots, F_d$. Let $h$ be the joint density function of Eq. (2.1). If $F_1, \ldots, F_d$ are continuous functions with respective densities $f_i(x) = \frac{d}{dx} F_i(x)$ for $i = 1, \ldots, d$, then the joint density function is given by

$$h(x_1, \ldots, x_d) = c(u_1, \ldots, u_d) \prod_{i=1}^{d} f_i(x_i). \tag{2.4}$$

Moreover, the fact that $C$ is a distribution function leads us to the following properties.

- $C(u_1, \ldots, u_d)$ is an increasing function for $u_1, \ldots, u_d \in [0, 1]$.
- If at least one of $(u_1, \ldots, u_d)$ takes the value 0, then $C(u_1, \ldots, u_d) = 0$ holds.
- The $i$-th marginal distribution function is given by setting $u_j = 1$ for $j \neq i$ as

$$C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i. \tag{2.5}$$

For further mathematical formulations and their theoretical deployment to a multivariate version, refer to [22, 25, 40] for instance.

## 2.2 FGM copula

The FGM copula is one of the copulas. In the early stage of the studies, Farlie [12], Gumbel [16] and Morgenstern [35] have discussed families of the bivariate FGM copula. Then, Johnson and Kotz [26] formulated a multivariate version of the FGM copula. In particular, the multivariate FGM copula is useful as an alternative to a multivariate normal distribution because it has a simple form and it can express mutual dependencies among two or more variables [22]. Therefore, the FGM copula has been applied to stochastic models in various research fields such as finance, economics, reliability engineering and so on (see e.g., [6, 11, 52]). For various interesting results on the FGM copula, see e.g., [26, 29, 50].

### 2.2.1 Definition

According to [26, 40], the joint distribution function of the multivariate FGM copula, written by $C$, is defined by

$$C(u_1,\dots,u_d) = \Big(\prod_{i=1}^{d} u_i\Big)\Big(1+\sum_{S\in\mathscr{S}}\theta_S\prod_{j\in S}\overline{u}_j\Big), \tag{2.6}$$

where $0 \le u_i \le 1$ for $i = 1,\dots,d$ and $\overline{u}_i = 1-u_i$. The set $\mathscr{S}$ consists of the subsets which are all combinations of at least two elements of the index number set $\{1,\dots,d\}$. The set $S$ represents an element of $\mathscr{S}$, and $j$ is an index number belongs to $S$. Equation (2.6) contains $2^d - d - 1$ parameters denoted by $\theta_S$. If all $\theta_S$'s are 0, Eq. (2.6) corresponds to $\prod_{i=1}^{d} u_i$. This means all variables are independent each other. If not, they are dependent. Therefore, we call $\theta_S$ *dependence parameter* in this thesis.

The joint density function of the $d$-variate FGM copula is given by

$$c(u_1,\dots,u_d) = 1+\sum_{S\in\mathscr{S}}\theta_S\prod_{j\in S}(1-2u_j). \tag{2.7}$$

For example, when $d = 3$, the dependence parameters are $\{\theta_{\{1,2\}},\theta_{\{1,3\}},\theta_{\{2,3\}},\theta_{\{1,2,3\}}\}$, and the joint distribution function is given by

$$C(u_1, u_2, u_3) = u_1u_2u_3(1+\theta_{\{1,2\}}\overline{u}_1\overline{u}_2+\theta_{\{1,3\}}\overline{u}_1\overline{u}_3+\theta_{\{2,3\}}\overline{u}_2\overline{u}_3+\theta_{\{1,2,3\}}\overline{u}_1\overline{u}_2\overline{u}_3). \quad (2.8)$$

There is a constraint for the dependence parameters. $\theta_S$'s must be the parameter such that the joint density function always takes non-negative values. Thus, it is easy to see that $\theta_S$'s must satisfy the following limitation (cf., [22, 26]).

$$1 + \sum_{S\in\mathscr{S}} \theta_S \prod_{j\in S}(1 - 2u_j) \geq 0, \quad (2.9)$$

where $\forall(u_1, \ldots, u_d) \in [0, 1]^d$. More specifically, in the case of $d = 2$, we have

$$1 + \theta_{\{1,2\}}(1 - 2u_1)(1 - 2u_2) \geq 0. \quad (2.10)$$

In order to find the range of $\theta_{\{1,2\}}$, we investigate the following four cases, i.e., $(u_1, u_2) = (0, 0), (0, 1), (1, 0)$ and $(1, 1)$ for the necessary and sufficient conditions for Eq. (2.10). Therefore, they yield

$$\left.\begin{array}{l} 1 + \theta_{\{1,2\}} \geq 0 \\ 1 - \theta_{\{1,2\}} \geq 0 \end{array}\right\}. \quad (2.11)$$

Hence, the appropriate range of $\theta_{\{1,2\}}$ is given by

$$-1 \leq \theta_{\{1,2\}} \leq 1. \quad (2.12)$$

For $d = 3$, since $\mathscr{S} = \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$, we have

$$\begin{aligned} C(u_1, u_2, u_3) = u_1u_2u_3\Big[&1 + \theta_{\{1,2\}}(1 - u_1)(1 - u_2) \\ &+ \theta_{\{1,3\}}(1 - u_1)(1 - u_3) + \theta_{\{2,3\}}(1 - u_2)(1 - u_3) \\ &+ \theta_{\{1,2,3\}}(1 - u_1)(1 - u_2)(1 - u_3)\Big]. \quad (2.13) \end{aligned}$$

By the same manner, the constant parameters $\theta_{\{1,2\}}$, $\theta_{\{1,3\}}$, $\theta_{\{2,3\}}$ and $\theta_{\{1,2,3\}}$ have to hold the following conditions (e.g., [26]).

$$\left.\begin{array}{l} 1+\theta_{\{1,2\}}+\theta_{\{1,3\}}+\theta_{\{2,3\}} \geq |\theta_{\{1,2,3\}}| \\ 1+\theta_{\{1,2\}}-\theta_{\{1,3\}}-\theta_{\{2,3\}} \geq |\theta_{\{1,2,3\}}| \\ 1-\theta_{\{1,2\}}+\theta_{\{1,3\}}-\theta_{\{2,3\}} \geq |\theta_{\{1,2,3\}}| \\ 1-\theta_{\{1,2\}}-\theta_{\{1,3\}}+\theta_{\{2,3\}} \geq |\theta_{\{1,2,3\}}| \end{array}\right\}. \tag{2.14}$$

Thus for example, if $\theta_{\{1,2,3\}} = 1$, the other parameters are fixed to $\theta_{\{1,2\}} = \theta_{\{1,3\}} = \theta_{\{2,3\}} = 0$. On the contrary, if $\theta_{\{1,2\}} = \theta_{\{1,3\}} = \theta_{\{2,3\}} = 0$, then $-1 \leq \theta_{\{1,2,3\}} \leq 1$. These conditions which are structured as Eq. (2.14) become complex when $n$ is large.

### 2.2.2 Relationship with correlation indices

We additionally mention three correlation indices between two variates, namely, Kendall's $\tau$, Spearman's $\rho$, and Pearson's linear correlation $r$. All of the correlation indices are known as non-parametric ones, and these can be expressed by utilizing a copula function. Considering population-based definitions of them, we obtain the following relationships between one of them and $\theta_{\{1,2\}}$ of the FGM copula [40].

Kendall's $\tau$ of the population version can be obtained by

$$\tau = 4\int_0^1 \int_0^1 C(u_1, u_2)c(u_1, u_2)du_1du_2 - 1. \tag{2.15}$$

In the case of FGM copula, we have

$$\tau = \frac{2}{9}\theta_{\{1,2\}}. \tag{2.16}$$

On the other hand, Spearman's $\rho$ of the population version is given by

$$\begin{aligned} \rho &= 12\int_0^1 \int_0^1 C(u_1, u_2)du_1du_2 - 3 \\ &= \frac{1}{3}\theta_{\{1,2\}}, \end{aligned} \tag{2.17}$$

based on the FGM copula. Finally, Pearson's linear correlation $r$ of the population version is as follows.

$$
\begin{aligned}
r &= 12\int_0^1\int_0^1 u_1u_2c(u_1,u_2)du_1du_2 - 3 \\
&= \frac{1}{3}\theta_{\{1,2\}},
\end{aligned}
\tag{2.18}
$$

Note that these indices are free from the selection of marginal distributions of the copula. This is greatly related to the fact that a copula function can separately treat the concept of 'coupling' and the selection of the marginal distributions. Also considering the range of $\theta_{\{1,2\}}$ is $-1 \leq \theta_{\{1,2\}} \leq 1$, we understand that the FGM copula cannot describe the strong dependence, that is, $-2/9 \leq \tau \leq 2/9$, $-1/3 \leq \rho \leq 1/3$ and $-1/3 \leq r \leq 1/3$, respectively.

## 2.3 Minimum and maximum value distributions based on $d$-variate FGM copula

### 2.3.1 Introduction

One unresolved problem of the FGM copula is that restrictions of its parameters have not been found as closed forms. For this reason, for example, it has been difficult to estimate the dependence parameters of the FGM copula so far. As we mentioned before, the $d$-variate FGM copula has totally $2^d - d - 1$ dependence parameters. In general, a lot of papers have referred to this restriction (e.g., [2, 17, 40]). However, there are few studies that explicitly derive the exact ranges of the parameters for $d \geq 4$ in particular because of its complexity.

Under such a situation, as a first step, we focus on the minimum and maximum value distributions which are constructed by the $d$-variate FGM copula in this paper. The former distribution can describe the lifetime (first failure occurrence time) of a series system with $d$ dependent components, and the latter one does that of a dependent parallel system. In addition, in order to reduce the complexity involved in the $d$-variate FGM copula, we assume that all dependence parameters are represented by just one parameter. As a result, the necessary conditions for the dependence parameter of minimum and maximum distributions are explicitly provided.

### 2.3.2 Main results

From Eq. (2.14), we can see that it is very hard to derive the exact ranges of the dependent parameters unless some certain conditions are made. Therefore in this section, as we mentioned in Subsection 2.3.1, we try to obtain the ranges of the dependence parameters for the minimum and maximum value distributions based on $n$-variate the FGM copula when all of the dependence parameters are identical to $\theta$. That is, we assume that $\theta \equiv \theta_S$ for all $S$ in Eq. (2.6).

We first show the minimum and maximum value distributions derived from one-parameter $d$-variate FGM copula. Suppose $\boldsymbol{U} = (U_1, \ldots, U_d)$ be a random vector that follows an FGM copula with $n$ uniform marginal distributions on the interval $[0, 1]$. Let $U_{1:d}$ and $U_{d:d}$ be the minimum and maximum values of $\boldsymbol{U}$, respectively. Let $C_{1:d}(u; \theta_{1:d})$ and $C_{d:d}(u; \theta_{d:d})$ denote the cumulative distribution functions (CDF) of $U_{1:d}$ and $U_{d:d}$, respectively. According to [39] and [27], $C_{1:d}(u; \theta_{1:d})$ is written as

$$\begin{aligned} C_{1:d}(u; \theta_{1:n}) &= \Pr[\min(U_1, \ldots, U_d) \leq u] \\ &= 1 - (1-u)^d \left(1 + \theta_{1:d} \sum_{k=2}^{d} \binom{d}{k} (-u)^k \right). \end{aligned} \tag{2.19}$$

Also $C_{d:d}(u; \theta_{d:d})$ is given by

$$\begin{aligned} C_{d:d}(u; \theta_{d:d}) &= \Pr[\max(U_1, \ldots, U_d) \leq u] \\ &= u^d \left(1 + \theta_{d:d} \sum_{k=2}^{d} \binom{d}{k} (1-u)^k \right). \end{aligned} \tag{2.20}$$

Note that we rewrite $\theta$ to $\theta_{1:d}$ and $\theta_{d:d}$ in the above equations respectively in order to distinguish these two parameters.

The following theorems present the necessary conditions for the ranges of $\theta_{1:d}$ and $\theta_{d:d}$ in Eqs. (2.19) and (2.20), respectively. In addition, these theorems yield corollaries about their asymptotic properties.

**Theorem 2.2.** *The range of* $\theta_{1:d}$ *is given by the following inequality.*

$$-\frac{1}{d-1} \leq \theta_{1:d} \leq \frac{1}{2-2(1-u_d^*)^d-(1+d)u_d^*}, \tag{2.21}$$

*where*

$$u_d^* = 1 - \left(\frac{1+d}{2d}\right)^{\frac{1}{d-1}}. \tag{2.22}$$

**Corollary 2.1.** *As* $d \to \infty$*, the range of* $\theta_{1:d}$ *is obtained as*

$$0 \leq \theta_{1:d} \leq \frac{1}{1-\log 2} \simeq 3.259. \tag{2.23}$$

**Theorem 2.3.** *The range of* $\theta_{d:d}$ *is given by the following inequality.*

$$-\frac{1}{2^d-1-d} \leq \theta_{d:d} \leq \frac{1}{(1-v_d^*)\{1+d-2(2-v_d^*)^{d-1}\}}, \tag{2.24}$$

*where* $v_d^*$ *is the solution of the following equation.*

$$d(2-v_d^*)^{d-1}-(d-1)(2-v_d^*)^{d-2}-\frac{1+d}{2}=0, \tag{2.25}$$

for $0 \leq v_d^* \leq 1$.

**Remark 2.1.** *The asymptotic property of* $v_d^*$ *is as follows.*

$$\lim_{d\to\infty} v_d^* = 1. \tag{2.26}$$

This remark offers some support to the next conjecture.

**Conjecture 2.1.** *As* $d \to \infty$*, the following equation holds.*

$$\lim_{d\to\infty} \frac{1}{(1-v_d^*)\{1+d-2(2-v_d^*)^{d-1}\}} = 0. \tag{2.27}$$

This conjecture is derived by **Remark 2.1** and the assumption that $1+d-2(2-v_d^*)^{d-1}$ diverges to positive infinity more quickly than $1-v_d^*$ converges to 0. Then we have the following corollary.

**Corollary 2.2.** *As* $d \to \infty$*, the range of* $\theta_{d:d}$ *is convergent to* 0.

Note that proofs for **Theorem 2.2** and **Corollary 2.1** are given in Appendix. Also, we would like to omit the proofs of **Theorem 2.3** and **Corollary 2.2** because they can be shown in the same way as those of **Theorem 2.2** and **Corollary 2.1**.

In summary, the above results guarantee that the ranges of $\theta_{1:d}$ and $\theta_{d:d}$ depend on only $d$, and the ranges become narrower as $d$ increases. For example, Table 2.1 shows the numerical results of the ranges of $\theta_{1:d}$ and $\theta_{d:d}$ for $d = 2, \ldots, 30$, and $\infty$. Note that every value is calculated by Eqs. (2.21) and (2.24), and rounded to the nearest thousandth. This table implies that the ranges of the negative values of $\theta_{1:d}$ and $\theta_{d:d}$ are narrower than those of the positive values, respectively. In addition, we can find that the range of $\theta_{d:d}$ is narrower than that of $\theta_{1:d}$.

Table 2.1: Numerical results of the ranges of $\theta_{1:d}$ and $\theta_{d:d}$.

| $d$ | $\theta_{1:d}$ | $\theta_{d:d}$ | $d$ | $\theta_{1:d}$ | $\theta_{d:d}$ |
|---|---|---|---|---|---|
| 2 | $[-1.000, 8.000]$ | $[-1.000, 8.000]$ | 17 | $[-0.063, 3.559]$ | $[-0.000, 1.104]$ |
| 3 | $[-0.500, 5.639]$ | $[-0.250, 4.440]$ | 18 | $[-0.059, 3.541]$ | $[-0.000, 1.070]$ |
| 4 | $[-0.333, 4.850]$ | $[-0.091, 3.229]$ | 19 | $[-0.056, 3.526]$ | $[-0.000, 1.039]$ |
| 5 | $[-0.250, 4.454]$ | $[-0.038, 2.611]$ | 20 | $[-0.053, 3.512]$ | $[-0.000, 1.011]$ |
| 6 | $[-0.200, 4.216]$ | $[-0.018, 2.232]$ | 21 | $[-0.050, 3.499]$ | $[-0.000, 0.985]$ |
| 7 | $[-0.167, 4.057]$ | $[-0.008, 1.974]$ | 22 | $[-0.048, 3.488]$ | $[-0.000, 0.961]$ |
| 8 | $[-0.143, 3.943]$ | $[-0.004, 1.786]$ | 23 | $[-0.045, 3.477]$ | $[-0.000, 0.940]$ |
| 9 | $[-0.125, 3.858]$ | $[-0.002, 1.642]$ | 24 | $[-0.043, 3.468]$ | $[-0.000, 0.920]$ |
| 10 | $[-0.111, 3.792]$ | $[-0.001, 1.528]$ | 25 | $[-0.042, 3.459]$ | $[-0.000, 0.901]$ |
| 11 | $[-0.100, 3.738]$ | $[-0.000, 1.435]$ | 26 | $[-0.040, 3.451]$ | $[-0.000, 0.884]$ |
| 12 | $[-0.091, 3.695]$ | $[-0.000, 1.357]$ | 27 | $[-0.038, 3.444]$ | $[-0.000, 0.868]$ |
| 13 | $[-0.083, 3.659]$ | $[-0.000, 1.292]$ | 28 | $[-0.037, 3.437]$ | $[-0.000, 0.852]$ |
| 14 | $[-0.077, 3.628]$ | $[-0.000, 1.235]$ | 29 | $[-0.036, 3.430]$ | $[-0.000, 0.838]$ |
| 15 | $[-0.071, 3.602]$ | $[-0.000, 1.186]$ | 30 | $[-0.034, 3.425]$ | $[-0.000, 0.825]$ |
| 16 | $[-0.067, 3.579]$ | $[-0.000, 1.143]$ | $\infty$ | $[\ 0, 1/(1-\log 2)]$ | $0$ |

### 2.3.3 Discussion and concluding remarks

The results in the previous subsection have been obtained from the necessary condition such that the density function always takes non-negative value for any $u \in [0, 1]$ and $d$ in the case of the minimum value distribution (and the maximum one as well). Therefore the ranges obtained by Eqs. (2.21) and (2.24) are both lack of sufficiency. In order to show this fact, let us go back to the FGM copula presented in Eq. (2.14) with $d = 3$. By setting all the parameters $\theta_{\{1,2\}}$, $\theta_{\{1,3\}}$, $\theta_{\{2,3\}}$ and $\theta_{\{1,2,3\}}$ be identical to $\theta$ in Eq. (2.14), we can calculate the possible range of $\theta$ as

$$-\frac{1}{4} \le \theta \le \frac{1}{2}. \tag{2.28}$$

On the other hand, from Table 2.1, we recall

$$-\frac{1}{2} \le \theta_{1:3} \le 5.639, \tag{2.29}$$

$$-\frac{1}{4} \le \theta_{3:3} \le 4.440, \tag{2.30}$$

for the two kinds of dependence parameters, respectively. It should be noted that the minimum and maximum value distributions based on the FGM copula are both exist only if the $d$-variate FGM copula is theoretically valid. In other words, for example, if we estimate that the value of $\widehat{\theta}_{1:3}$ is 2.0 from a certain data analysis concerning the minimum value distribution, this value surely satisfies Eqs. (2.21) and (2.29) but it does not satisfy the limitation denoted by Eq. (2.9) with the identical dependence parameters. Therefore in this case, we must discard the estimation result $\widehat{\theta}_{1:3} = 2.0$, and need to find the estimated value from the range $-\frac{1}{4} \le \theta \le \frac{1}{2}$ in Eq. (2.28) instead of Eq.(2.29).

In conclusion, we have revealed the necessary conditions for the dependence parameters $\theta_{1:d}$ and $\theta_{d:d}$ for general $d$ and their asymptotic properties on $d \to \infty$ in this section. However, providing their sufficient conditions and the exact ranges of the parameters has been remaining for the future work.

# Chapter 3

# Modeling of dependent failures by FGM copula

## 3.1 Introduction

In the previous chapter, we introduced the multivariate FGM copula from the viewpoint of probability theory. On the other hand, if we suppose that random variables of the FGM copula express the lifetimes of system components, we can deal with it as one of the modeling tools for dependent failure occurrences among them. Therefore, we can naturally extend the traditional reliability models in which the lifetimes are i.i.d. to d.i.d.[1] cases.

In this chapter, we investigate the effect of dependent failure occurrence on system reliability assessment. First, we propose statistical models of a parallel system and series system with $n$ dependent components by using the multivariate FGM copula. Then, we investigate several reliability-related properties of the systems. As a result, we reveal that the $n$-component parallel system cannot deliver its designed reliability if the lifetimes of the individual components have positive dependence. On the other hand, we derive that the $n$-component series system can exceed its designed reliability under such dependent failure-occurrence environment. This result is given by Kimura, Ota and Abe [27] and Ota and Kimura [46].

[1] The term "d.i.d."denotes that random variables are statistically dependent and identically distributed.

## 3.2 Reliability assessment of $n$-component systems under dependent failure-occurrence environment

### 3.2.1 Assumptions

Suppose that the system consists of $n$ components. Let $X_i$ be a random variable that represents the lifetime of the component $i$ ($i = 1, \ldots, n$). Let $F(x_1, \ldots, x_n)$ be the joint distribution function of these variables. We assume that $F(x_1, \ldots, x_n)$ follows the $n$-variate FGM copula given by Eq. (2.6) and the marginal distributions are identical exponential distributions. That is, the marginal distribution function is represented by $F_i(x_i) \equiv F(x_i) = 1 - \exp[-\lambda x_i]$ ($\lambda > 0$). In order to simplify this modeling, we introduce another dependence parameter $\theta_k$ for $k = |S|$ (e.g., $\theta_2 = \theta_{\{1,2\}} = \theta_{\{1,3\}} = \theta_{\{2,3\}}$ for $n = 3$). This assumption means that any dependence among $k$ components are equivalent to $\theta_k$. Here, we have

$$F(x_1, \ldots, x_n) = \prod_{i=1}^{n}(1 - e^{-\lambda x_i}) \left(1 + \sum_{S \in \mathscr{F}} \theta_k \prod_{j \in S} e^{-\lambda x_j}\right). \tag{3.1}$$

For those components, we analyze the reliabilities of the parallel system and series one.

### 3.2.2 Parallel system with dependent failures

Let $X_{\max}$ denote $\max(X_1, \ldots, X_n)$. Then, we can derive the CDF of the lifetime of $n$-component parallel system, $\Pr[X_{\max} \le x] \equiv F_{\max}(x)$, as

$$F_{\max}(x) = F(x)^n \Big(1 + \sum_{k=2}^{n} \binom{n}{k} \theta_k \bar{F}(x)^k\Big), \tag{3.2}$$

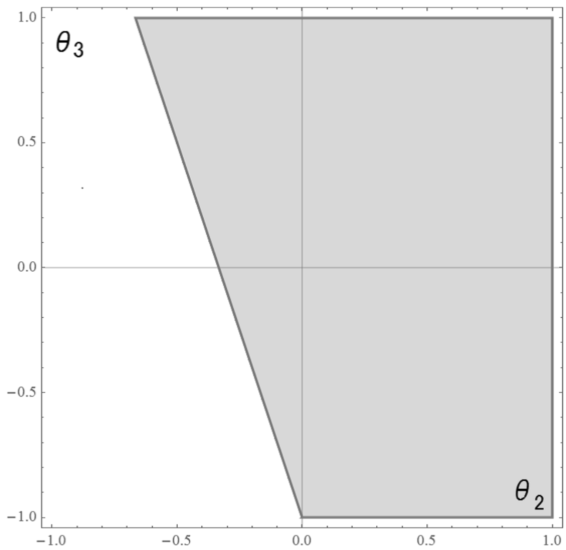


Figure 3.1: Region of $(\theta_2, \theta_3)$ for the 3-component parallel system.

from Eq. (2.6) via the simplification of the constant parameters. If $n = 3$, PDF of $X_{\max}$, namely, $f_{\max}(x)$ can be written as follows.

$$f_{\max}(x) = 3F(x)^2 f(x) \{1 + \theta_2(1 - F(x))(3 - 5F(x)) \\ +\theta_3(1 - F(x))^2(1 - 2F(x))\} . \tag{3.3}$$

The colored area in Figure 3.1 illustrates the region of $(\theta_2, \theta_3)$ satisfying $f_{\max}(x) \geq 0$ for all $x$.

From Eq. (3.2), the MTTF of an $n$-component parallel system can be obtained as

$$\mathrm{E}[X_{\max}] \equiv \mathrm{MTTF}_{\max}(n) = \frac{1}{\lambda}\Big[H(n) - \sum_{k=2}^{n} \frac{J(k)}{k}\theta_k\Big], \tag{3.4}$$

where $H(n)$ describes a partial sum of the harmonic series

$$H(n) = \sum_{i=1}^{n} \frac{1}{i}, \tag{3.5}$$

and

$$J(k) = \binom{n}{n-k} \Big/ \binom{n+k}{k}. \tag{3.6}$$

Note that if all of the components are mutually independent, the MTTF can be given by

$$\mathrm{MTTF}(n) = \frac{1}{\lambda} H(n), \tag{3.7}$$

(see, e.g., [38]). Hence we can understand that if the system operates under the positively-dependent failure-occurrence environment (i.e., $\theta_k > 0$ $(k = 2, 3, \ldots, n)$), the achieved MTTF declines from the design MTTF under the assumption of the independence.

From Eq. (3.2), the moment of second order about the origin, $\mathrm{M}^{(2)}_{\max}(n)$, can be derived as

$$\begin{aligned}\mathrm{M}^{(2)}_{\max}(n) =& -\frac{2}{\lambda^2}\left[\sum_{k=2}^{n}\frac{1}{k^2}\binom{n}{k}\theta_k + \sum_{r=1}^{n}\frac{1}{r^2}\binom{n}{r}(-1)^r \right.\\ &\left. + \sum_{k=2}^{n}\binom{n}{k}\theta_k\left\{\sum_{r=1}^{n}\frac{1}{(k+r)^2}\binom{n}{r}(-1)^r\right\}\right]\\ =&\frac{1}{\lambda^2}\left[H(n)^2 + \sum_{i=1}^{n}\frac{1}{i^2} - 2\sum_{k=2}^{n}\frac{J(k)}{k}\theta_k\sum_{r=0}^{n}\frac{1}{k+r}\right].\end{aligned} \tag{3.8}$$

Therefore the variance of $X_{\max}$ is given by

$$\mathrm{Var}[X_{\max}] \equiv \mathrm{M}^{(2)}_{\max}(n) - \mathrm{MTTF}_{\max}(n)^2. \tag{3.9}$$

### 3.2.3 Series system with dependent failures

By following the same fashion, we can also derive the CDF of the minimum of $n$ variables, i.e., $\Pr[\min(X_1, X_2, \ldots, X_n) \equiv X_{\min} \le x] = F_{\min}(x)$, as

$$F_{\min}(x) = 1 - \bar{F}(x)^n\Big(1 + \sum_{k=2}^{n}\binom{n}{k}(-1)^k\theta_k F(x)^k\Big), \tag{3.10}$$

from Eq. (2.6) with reducing the number of the dependence parameters.

If $n = 3$, PDF of $X_{\min}$, $f_{\min}(x)$ can be obtained as

$$\begin{aligned}f_{\min}(x) = 3(1-F(x))^2 f(x)\big\{&1 - \theta_2(2-5F(x))F(x)\\ &+ \theta_3(1-2F(x))F(x)^2\big\}.\end{aligned} \tag{3.11}$$

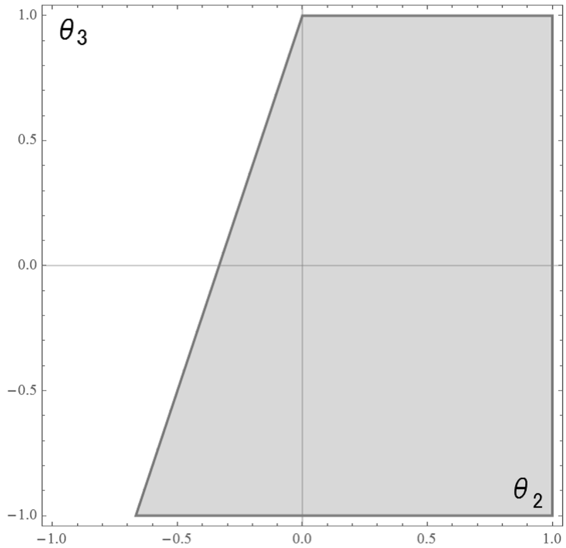


Figure 3.2: Region of $(\theta_2, \theta_3)$ for 3-component series system.

The colored area in Figure 3.2 illustrates the region of $(\theta_2, \theta_3)$ which satisfies $f_{\min}(x) \geq 0$ for all $x$.

MTTF is given by

$$\begin{aligned}
\mathrm{E}[X_{\min}] \equiv \mathrm{MTTF}_{\min}(n) &= \frac{1}{n\lambda} + \frac{1}{\lambda}\sum_{k=2}^{n}\binom{n}{k}\theta_k(-1)^k\sum_{r=0}^{k}\binom{k}{r}(-1)^r\frac{1}{n+r} \\
&= \frac{1}{n\lambda}\left\{1 + \sum_{k=2}^{n}(-1)^k J(k)\theta_k\right\}. \qquad (3.12)
\end{aligned}$$

Moment of second order about the origin, denoted by $\mathrm{M}_{\min}^{(2)}(n)$ can be obtained as

$$\mathrm{M}_{\min}^{(2)}(n) = \frac{2}{n^2\lambda^2} + 2\sum_{k=2}^{n}\binom{n}{k}\theta_k(-1)^k\sum_{r=0}^{k}\binom{k}{r}(-1)^r\frac{1}{\lambda^2(n+r)^2}. \qquad (3.13)$$

Therefore the variance of $X_{\min}$ can be calculated by

$$\mathrm{Var}[X_{\min}] \equiv \mathrm{M}_{\min}^{(2)}(n) - \mathrm{MTTF}_{\min}(n)^2. \qquad (3.14)$$

## 3.3 Numerical illustrations

### 3.3.1 Basic behavior of FGM copula with exponential marginals

At first, we are interested in the behavior of the FGM copula with exponential marginals. The CDF of the bivariate case is given by

$$\Pr[X_1 \le x_1, X_2 \le x_2] = (1 - e^{-\lambda x_1})(1 - e^{-\lambda x_2})(1 + \theta_2 e^{-\lambda x_1} e^{-\lambda x_2}). \tag{3.15}$$

Therefore, the PDF which is expressed by $f(x_1, x_2)$ is derived as

$$\begin{aligned} f(x_1, x_2) = \lambda^2 e^{-\lambda x_1} e^{-\lambda x_2} + \theta_2 \big\{ 4\lambda^2 e^{-2\lambda x_1} e^{-2\lambda x_2} \\ - 2\lambda^2 e^{-\lambda x_1} e^{-2\lambda x_2} - 2\lambda^2 e^{-2\lambda x_1} e^{-\lambda x_2} + \lambda^2 e^{-\lambda x_1} e^{-\lambda x_2} \big\}. \end{aligned} \tag{3.16}$$

We depict several typical behaviors of Eq. (3.16) in Figure 3.3 with $\lambda = 1$. If $\theta_2 = 0$, it behaves as a 2-dimensional independent bivariate exponential density. On the other hand, when $X_1$ and $X_2$ are dependent, the shape of joint density changes. The case of $\theta_2 = 1$ can be characterized by, for example, the joint density $f(x_1, x_2)$ takes larger value if both values are smaller (near by zero), and the shape of 'edge' of density can be seen like a convex curve, since $X_1$ and $X_2$ tend to take slightly closer values theoretically. On the contrary, the opposite characteristics can be seen in the case of $\theta_2 = -1$.

In general, from the viewpoint of reliability analysis taking account of the dependent failure-occurrence environment, it can be considered natural that the value of each $\theta_k$ ($k = 2, 3, \ldots, n$) is non-negative. Hence in the rest of this chapter, we perform the reliability analysis and the characteristics investigation under the assumption of $\theta_k \ge 0$ ($k = 2, 3, \ldots, n$). It is intuitively understandable that the MTTF of parallel systems gets smaller when $\theta_k$ is positive. This fact can be confirmed by Eq. (3.4), since $J(k)$ is a decreasing sequence. On the other hand, in the case of series systems, the MTTF of the system becomes larger if the term

$$\sum_{k=2}^{n} (-1)^k J(k) \theta_k \tag{3.17}$$

is positive (see, Eq. (3.12)).

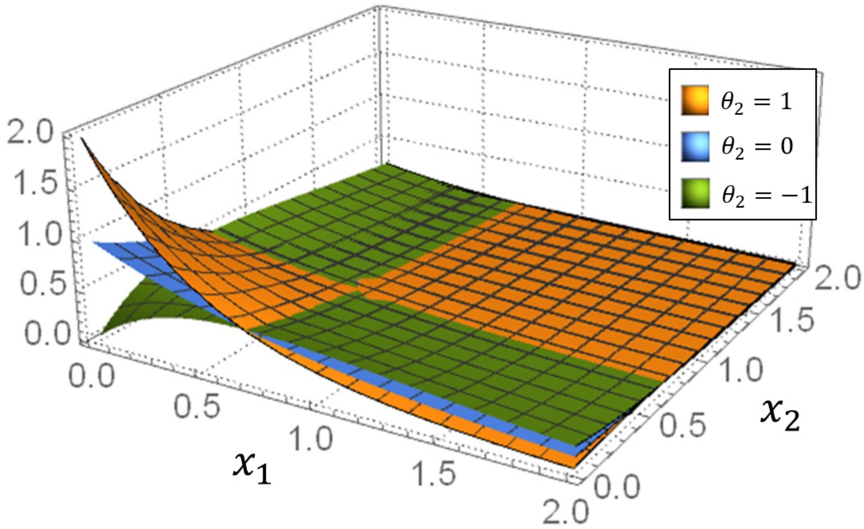


Figure 3.3: Behaviors of the bivariate joint density function ($\theta_2 = -1, 0, 1$).

### 3.3.2 Parallel system with dependent failures

We now focus on the values of the coefficients of dependence parameters $\theta_k$ ($k = 2, 3, \ldots, n$), denoted by $J(k)/k$ in Eq. (3.4). Table 3.1 shows several coefficients of $\theta_k$. We can see that the coefficient of $\theta_k$ with the smaller index number $k$ becomes significantly larger than other coefficients of $\theta_k$ as the number of components, $n$, becomes large. In other words, the value of $\theta_k$ with the larger index number $k$ does not contribute to reduce the MTTF by the possibility of dependent failure occurrences.

Another finding from Table 3.1 is that these values in Table 3.1 correspond to

Table 3.1: Coefficients of $\theta_k$ of Parallel Systems.

| | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|
| $n=2$ | $\frac{1}{12}$ | – | – | – | – |
| $n=3$ | $\frac{3}{20}$ | $\frac{1}{60}$ | – | – | – |
| $n=4$ | $\frac{1}{5}$ | $\frac{4}{105}$ | $\frac{1}{280}$ | – | – |
| $n=5$ | $\frac{5}{21}$ | $\frac{5}{84}$ | $\frac{5}{504}$ | $\frac{1}{1260}$ | – |
| $n=6$ | $\frac{15}{56}$ | $\frac{5}{63}$ | $\frac{1}{56}$ | $\frac{1}{385}$ | $\frac{1}{5544}$ |

the values of the coefficients which appear in a certain approximation method of the difference equation theory (see, e.g., [14, 15]). In addition, an alternative form of $J(k)/k$ can be described by

$$\frac{J(k)}{k} = \binom{n}{k} \frac{\Gamma(k)\Gamma(n+1)}{\Gamma(n+k+1)} = \binom{n}{k} B(k, n+1), \tag{3.18}$$

where $B(a, b)$ is the beta function.

Here, we assume $\theta \equiv \theta_k$ $(k = 2, 3, \ldots, n)$ for more simplicity. This assumption means that all dependence among $k$ components $(k = 1, 2, \ldots, n)$ are equivalent to $\theta$. Several behaviors of CDF and PDF are shown in Figures 3.4 and 3.5, respectively $(\lambda = 1)$.

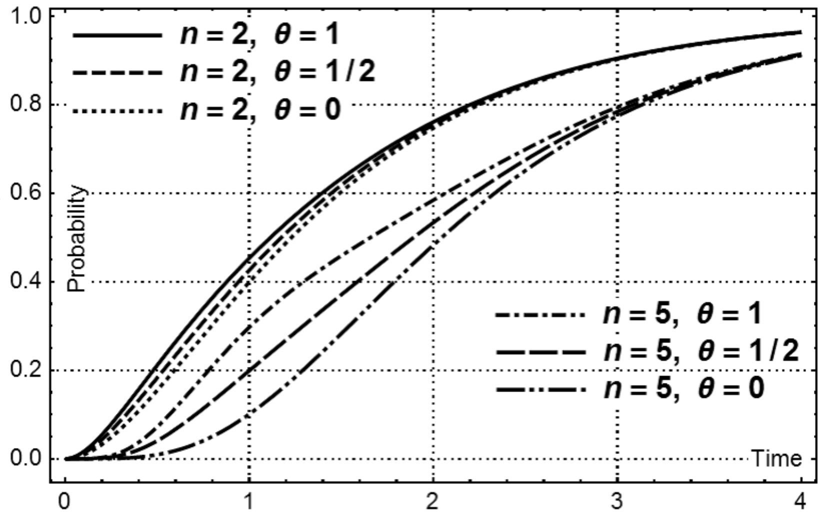


Figure 3.4: Behaviors of CDF (parallel system).

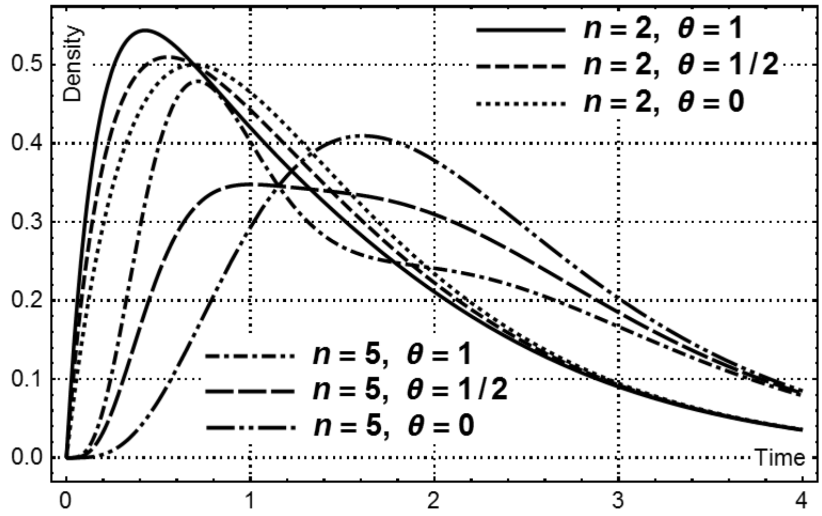


Figure 3.5: Behaviors of PDF (parallel system).

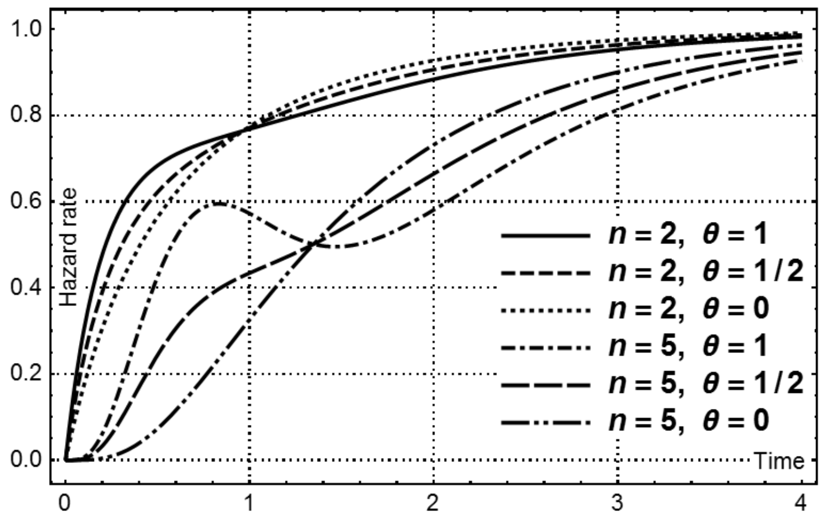


Figure 3.6: Behaviors of hazard rate (parallel system).

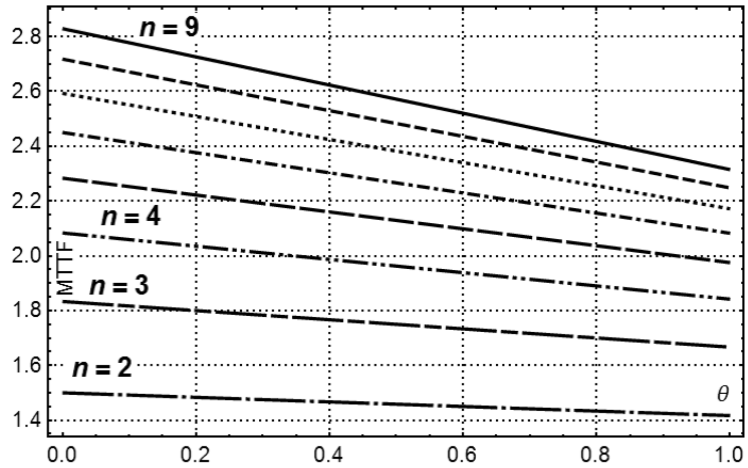


Figure 3.7: Behaviors of MTTF (parallel system).

Figure 3.5 suggests that the PDF shows a bimodal shape when both of $\theta$ and $n$ become large. Figures 3.6 and 3.7 represent the hazard rate function and MTTF, respectively. In Figure 3.6, when $\theta$ is large, the shape of the hazard rate depicts characteristic behavior as seen in Figure 3.5. It is also represented that the more stronger dependence reduces the MTTF in Figure 3.7.

Variance of the lifetime $X_{\max}$ is illustrated in Figure 3.8 ($\lambda = 1$). In the case of $\theta = 0$, it converges to $\pi^2/6$ as $n \to \infty$.

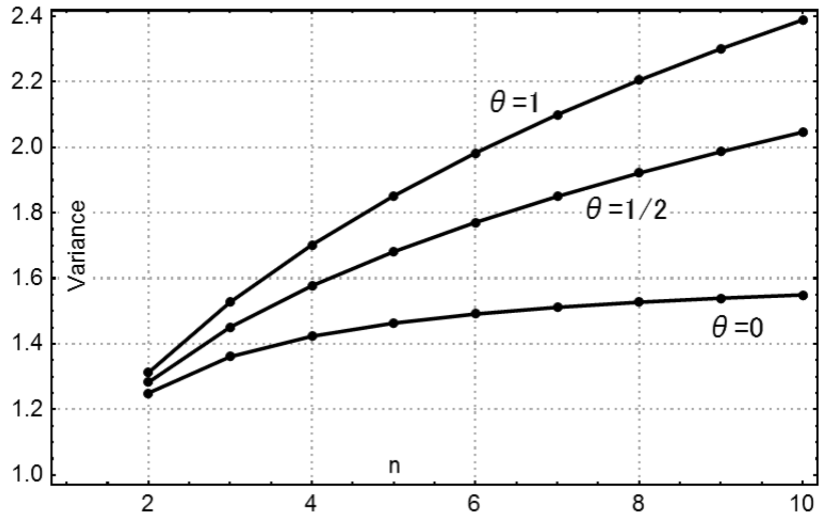


Figure 3.8: Variance of $X_{\max}$ ($\theta = 1, 1/2, 0$) (parallel system).

#### Decrement of MTTF

As shown in Figure 3.7, the dependent failure-occurrence environment has a harmful influence on the parallel system in the sense of the decrement of MTTF. For example, if one has designed an $n$-component parallel system under the expectation of the independent failure-occurrence environment, he/she would be betrayed if the actual operation environment induces the dependent failures. Therefore we calculate the value of dependence parameter which spoils the amount of MTTF of one component out of $n$-component parallel system by using the following equation.

$$\mathrm{MTTF}(n-1) = \mathrm{MTTF}_{\max}(n), \tag{3.19}$$

where $\mathrm{MTTF}(n-1)$ is derived by Eq. (3.7). Solving Eq. (3.19) with respect to $\theta$ analytically, we have

$$\theta(n) = 1/\big[n(\frac{H(n)}{2} - \frac{n}{n+1})\big]. \tag{3.20}$$

Table 3.2 is a summary of the relation between $n$ ($n \geq 5$) and $\theta$. $n \geq 5$ is needed to satisfy the condition $\theta \leq 1$. In this table, for example, it can be stated that the MTTF of the 5-component dependent parallel system cannot beyond that of 4-component independent parallel system if $\theta \geq 24/37$. Also this threshold value decreases as $n$ increases. In other words, the potentially-dependent parallel system with the large number of components can easily lose its design MTTF by the very small amount of dependency which is represented by $\theta$.

Table 3.2: Threshold value of $\theta$ which spoils MTTF of $n$-independent-component parallel system.

| $n = 5$ | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|
| $\theta = \frac{24}{37}$ | $\frac{140}{309}$ | $\frac{20}{59}$ | $\frac{630}{2369}$ | $\frac{560}{2593}$ | $\frac{5544}{30791}$ |
| =0.6487 | =0.4531 | =0.3390 | =0.2660 | =0.2160 | =0.1801 |

### 3.3.3 Series system with dependent failures

Here, we investigate several characteristics of an $n$-component series system with dependent failures by following the same manner to the previous section. Table 3.3 presents the coefficients of $\theta_k$ ($k = 2, 3, \ldots, 6$), that is, $J(k)/n$. The same behavior except for the factor of $(-1)^k$ is observed by comparing Table 3.1. Therefore, we can understand that the MTTF of a series system under the dependent failure-occurrence environment gets longer than that of the case of independent-component series system if the positive dependence [11] is possible. In addition, corresponding to Eq. (3.18), it is obvious that

$$\frac{J(k)}{n} = \binom{n-1}{k-1} \frac{\Gamma(k)\Gamma(n+1)}{\Gamma(n+k+1)} = \binom{n-1}{k-1} B(k, n+1). \tag{3.21}$$

Several behaviors of CDF and PDF based on Eq. (3.10) are shown in Figures 3.9 and 3.10, respectively ($\lambda = 1$). Different from the parallel system, these curves seem unimodal.

Table 3.3: Coefficients of $\theta_k$ of Series Systems.

| | $k=2$ | $k=3$ | $k=4$ | $k=5$ | $k=6$ |
|---|---|---|---|---|---|
| $n=2$ | $\frac{1}{12}$ | – | – | – | – |
| $n=3$ | $\frac{1}{10}$ | $-\frac{1}{60}$ | – | – | – |
| $n=4$ | $\frac{1}{10}$ | $-\frac{1}{35}$ | $\frac{1}{280}$ | – | – |
| $n=5$ | $\frac{2}{21}$ | $-\frac{1}{28}$ | $\frac{1}{126}$ | $-\frac{1}{1260}$ | – |
| $n=6$ | $\frac{5}{56}$ | $-\frac{5}{126}$ | $\frac{1}{84}$ | $-\frac{1}{462}$ | $\frac{1}{5544}$ |

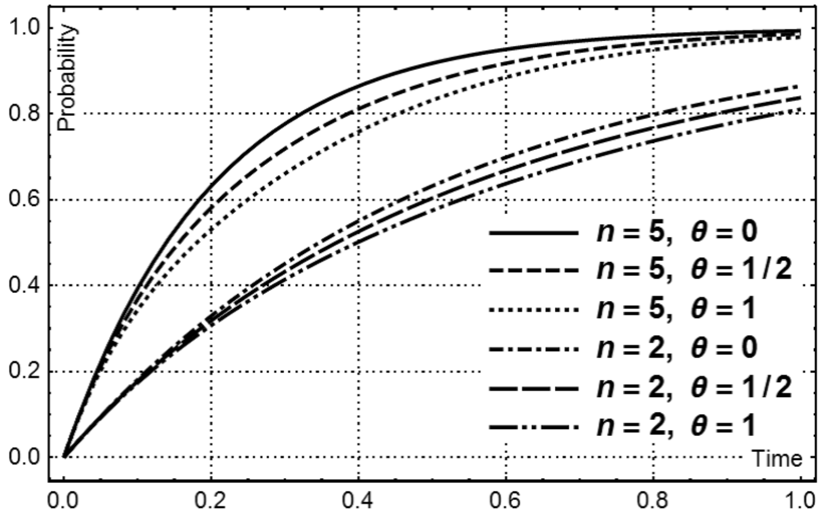


Figure 3.9: Behaviors of CDF (series system).

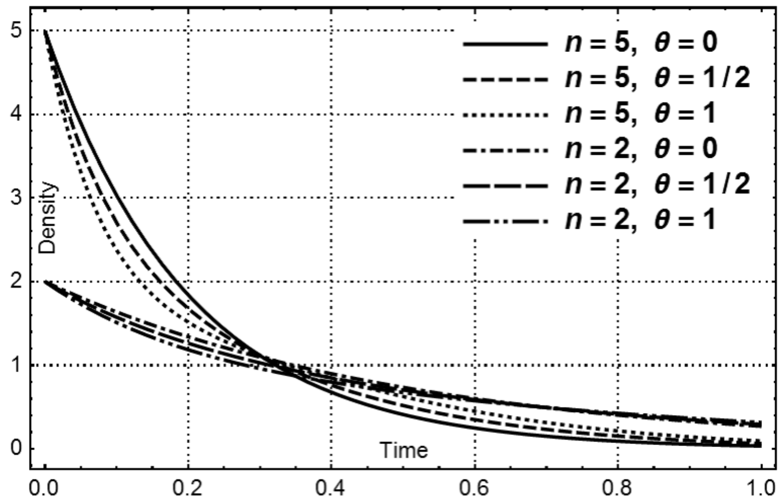


Figure 3.10: Behaviors of PDF (series system).

Figures 3.11 and 3.12 show the hazard functions and MTTF respectively ($\lambda = 1$). It is interesting that the hazard rate temporarily drops at the early stage of the lifetime. It can be confirmed that the MTTF increases when the degree of positive dependence becomes high. Variance of the lifetime $X_{\min}$ is illustrated in Figure 3.13. It converges to 0 as $n \to \infty$.

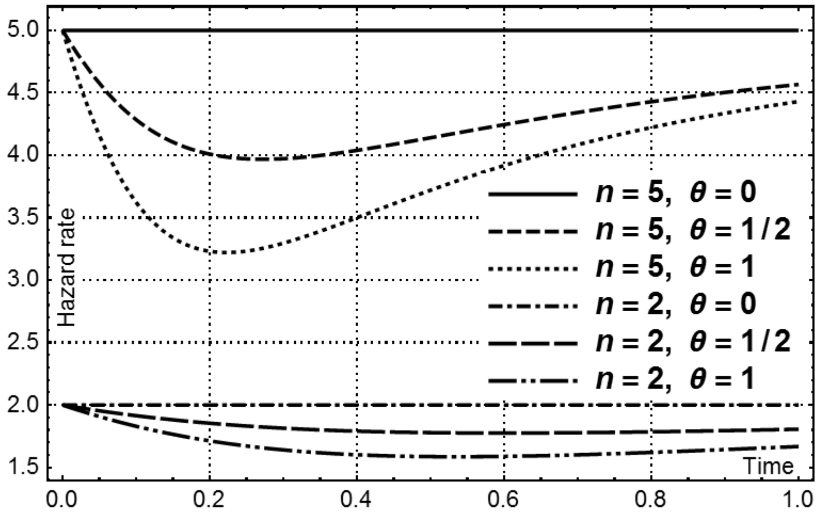


Figure 3.11: Behaviors of hazard rate (series system).

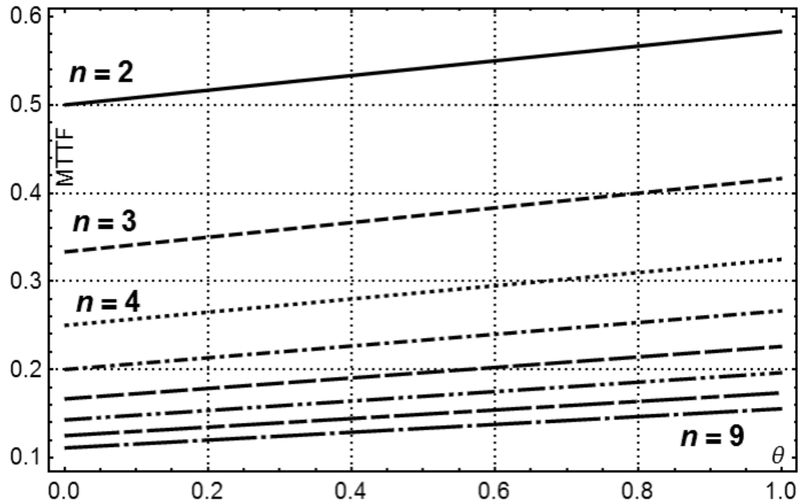


Figure 3.12: Behaviors of MTTF (series system).

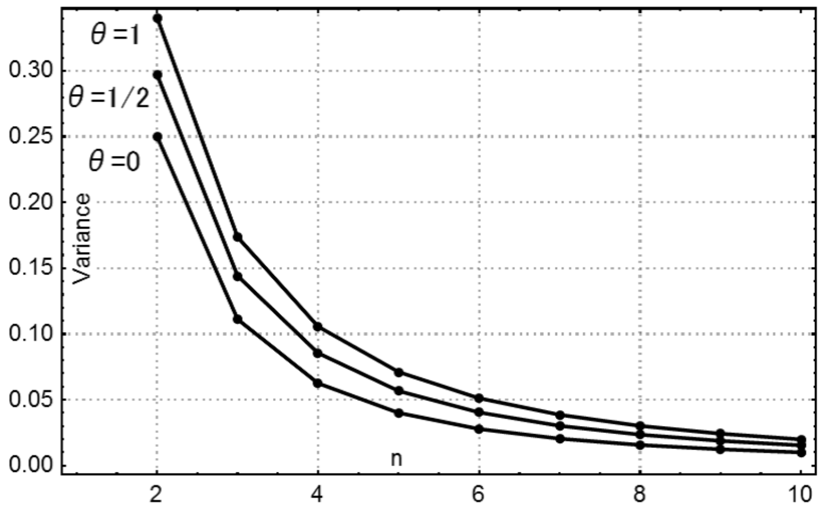


Figure 3.13: Variance of $X_{\min}$ ($\theta = 1, 1/2, 0$) (series system).

**Increment of MTTF**

In the series systems under the dependent failure-occurrence environment, the positive dependency among the components prolongs the lifetime of the system. In contradistinction to the property of $n$-component parallel systems which was discussed in Section 3.3.2, we find the threshold value of the dependence parameter $\theta$, which gains the amount of MTTF of one component out of an $n$-component series system. By solving the following equation

$$\mathrm{MTTF}(n-1) = \mathrm{MTTF}_{\min}(n), \tag{3.22}$$

with respect to $\theta$, we obtain the relation as

$$\theta(n) = \frac{2(n+1)}{(n-1)^2}. \tag{3.23}$$

Table 3.4 consists of several numerical results. By the same reason to the case of Table 3.2, $n \geq 5$ is required.

Table 3.4: Threshold value of $\theta$ which gains MTTF of $(n-1)$-independent-component series system.

| $n=5$ | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|
| $\theta=\frac{3}{4}$ | $\frac{14}{25}$ | $\frac{4}{9}$ | $\frac{18}{43}$ | $\frac{5}{16}$ | $\frac{22}{81}$ |
| =0.75 | =0.56 | =0.4444 | =0.3673 | =0.3125 | =0.2716 |

## 3.4 Conclusion

In this chapter, we have presented several reliability-related properties of the $n$-component systems under the assumption that the components may fail dependently. We have restricted the structure of $n$-component systems into the parallel and series systems, respectively. As a modeling framework, the FGM copula has attractive characteristics, and it is comparatively easy for doing mathematical manipulations. As a result, it has been confirmed that there exist several complex conditions among the dependence parameters. By using the simplification for many constant parameters, we have derived the closed forms of MTTF and variance of the systems. In order to illustrate the characteristics of two kinds of $n$-component systems, several numerical examples have been shown.

# Chapter 4

# Parameter estimation for FGM copula

In the previous chapter, we have formulated the model of the series system and parallel one with dependent components by using the multivariate FGM copula. In this chapter, we estimate the parameters of the model in order to evaluate the dependence among the components by using their failure times. To do so, we newly present an estimation algorithm for the parameters of the multivariate FGM copula by using the theory of *the inference functions for margins* (IFM, for short). Then, we theoretically analyze the asymptotic normality of the proposed estimators and numerically investigate the estimation accuracy. This result is given by Ota and Kimura [47] and Ota and Kimura [48].

## 4.1 Introduction

In general, parameter estimation for copulas is effective in analyzing dependence structures among data. However, it is computationally hard work because copulas have many parameters such as marginal parameters and dependence parameters in general. Thus, efficient/computable estimation methods have been required so far.

The problem here is that there are no practical estimation methods of the parameters of the multivariate FGM copula. The reason mainly depends on the computational complexity of estimating many parameters and dealing with the parameter space. As we mentioned in Chapter 2, the $d$-variate FGM copula consists of $2^d - d - 1$ dependence parameters. In fact, the FGM copula has relatively a lot of parameters compared with other multivariate copulas (e.g., a $d$-variate Gaussian copula [22] has $2d + \binom{d}{2}$ ($< 2^d - d - 1$) dependence parameters). It implies that the parameter estimation for the dependence parameters of the multivariate FGM copula is computationally more complicated than that for the parameters of the other copula distributions. Especially, the ordinary maximum likelihood estimation (MLE, for short) is infeasible over such high dimensional space.

One acceptable estimation method is IFM estimation (IFME, for short). IFME is known as a framework of the parameter estimation for copulas. It was proposed by Joe and Xu[23], and well explained in Xu [61]. It is based on $k$-th marginal likelihood functions. This method firstly estimates the univariate parameters from separate univariate likelihood functions of a copula. Then, it estimates bivariate, trivariate and multivariate parameters from bivariate, trivariate and multivariate likelihoods with lower order parameters fixed as estimated values. The advantages of IFME are its computational efficiency and asymptotic property that holds under several regularity conditions [24, 52, 61]. However, when we apply IFME to the multivariate FGM copula, the following problems exist: (i) it has not been known how to apply IFME to the parameter estimation of the FGM copula, (ii) it has not been found whether or not the asymptotic property holds.

Therefore, in this chapter, we solve the above problems. We present how to apply IFME to the multivariate FGM copula. Then, we theoretically analyze the asymptotic normality of the proposed estimators. Moreover, we investigate its computation time and estimation accuracy via numerical studies.

This result facilitates evaluating the dependence structure among system components. If we have lifetime data of the components, we can find the dependencies by estimating the dependence parameters of the multivariate FGM copula. Thus, this result contributes to the dependent failure analysis in reliability engineering.

The remainder of this chapter is organized as follows. In Section 4.2, we present algorithms of MLE and IFME for the multivariate FGM copula. Their asymptotic properties are analytically revealed in Section 4.3. The performance of IFME is discussed through simulation in Section 4.4. Finally, we conclude our study with a summary in Section 4.5.

## 4.2 Estimation

First of all, we start with redefining the distribution function of the $d$-variate FGM copula. Suppose $\boldsymbol{X} = (X_1, \ldots, X_d)$ be a random vector that follows the $d$-variate FGM copula with arbitrary continuous marginal distributions. For $i = 1, \ldots, d$, let $F_i(x_i; \boldsymbol{\delta_i})$ and $f_i(x_i; \boldsymbol{\delta_i})$ be the $i$-th marginal distribution function and the density function with a parameter vector $\boldsymbol{\delta_i}$, respectively. Let us define $\boldsymbol{\delta}$ as $\boldsymbol{\delta} = (\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_d})$. Then, the joint distribution function of $\boldsymbol{X}$, denoted by $H_d$, can be written as

$$\begin{aligned} H_d(x_1, \ldots, x_d; \boldsymbol{\delta}, \boldsymbol{\theta}) &= \Pr[X_1 \leq x_1, \ldots, X_d \leq x_d], \\ &= \prod_{i=1}^{d} F_i(x_i; \boldsymbol{\delta_i}) \left(1 + \sum_{k=2}^{d} \sum_{1 \leq j_1 < \cdots < j_k \leq d} \theta_{j_1 \cdots j_k} \overline{F}_{j_1} \cdots \overline{F}_{j_k}\right), \end{aligned} \tag{4.1}$$

where $\overline{F}_i \equiv 1 - F_i(x_i; \cdot)$ for $i = 1, \ldots, d$, $\boldsymbol{\delta} = (\boldsymbol{\delta_1}, \ldots, \boldsymbol{\delta_d})$, and $\boldsymbol{\theta}$ is the parameter vector of the FGM copula (e.g., for $d = 3$, $\boldsymbol{\theta} = (\theta_{12}, \theta_{13}, \theta_{23}, \theta_{123})$). Also, the joint density function of of $\boldsymbol{X}$ is given by

$$\begin{aligned} &h_d(x_1, \ldots, x_d; \boldsymbol{\delta}, \boldsymbol{\theta}) \\ &\quad = \prod_{i=1}^{d} f_i(x_i; \boldsymbol{\delta_i}) \left(1 + \sum_{k=2}^{d} \sum_{1 \leq j_1 < \cdots < j_k \leq d} \theta_{j_1 \cdots j_k} (2\overline{F}_{j_1} - 1) \cdots (2\overline{F}_{j_k} - 1)\right). \end{aligned} \tag{4.2}$$

Note that, $\boldsymbol{\theta}$ must satisfy the following limitation.

$$\Theta = \left\{ \boldsymbol{\theta} \;\middle|\; 1 + \sum_{k=2}^{d} \sum_{1 \leq j_1 < \cdots < j_k \leq d} \theta_{j_1 \cdots j_k} (1 - 2u_{j_1}) \cdots (1 - 2u_{j_k}) \geq 0 \right\}, \tag{4.3}$$

where $\forall (u_1, \ldots, u_d) \in [0, 1]^d$ and $\Theta$ denotes the parameter space. As we mentioned in Chapter 2, the parameter space $\Theta$ becomes more and more complex as $d$ becomes large. When we estimate the dependence parameters, we need to consider this parameter space.

### 4.2.1 Ordinary maximum likelihood estimation

The ordinary MLE is one of the natural choices to estimate parameters of random distributions because its estimators satisfy the asymptotic property under several regularity conditions (see e.g., [25]). In this method, we estimate all of the parameters $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ simultaneously. For a sample size $n$, with observed independent random vectors $\boldsymbol{x_i} = (x_{i1}, \ldots, x_{id})$ for $i = 1, \ldots, n$, the full-dimensional log-likelihood function of $\boldsymbol{\delta}$ and $\boldsymbol{\theta}$ can be written by

$$\ell \equiv \ell(\boldsymbol{\delta}, \boldsymbol{\theta}; \boldsymbol{x_1}, \ldots, \boldsymbol{x_n}) = \sum_{i=1}^{n} \log h_d(x_{i1}, \ldots, x_{id}; \boldsymbol{\delta}, \boldsymbol{\theta}). \tag{4.4}$$

Thus, the ordinary maximum likelihood estimators for the parameters, denoted by $\widehat{\boldsymbol{\delta}}$ and $\widehat{\boldsymbol{\theta}}$, are given below.

$$(\widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\theta}}) = \underset{\boldsymbol{\delta}, \boldsymbol{\theta} \in \Theta}{\arg\max}\, \ell. \tag{4.5}$$

More specifically, $(\widehat{\boldsymbol{\delta}}, \widehat{\boldsymbol{\theta}})$ is the root of the following non-linear simultaneous equations.

$$\begin{bmatrix} \frac{\partial}{\partial \boldsymbol{\delta}^T} \sum \log h_d(\cdot; \boldsymbol{\delta}, \boldsymbol{\theta}) \\ \frac{\partial}{\partial \boldsymbol{\theta}^T} \sum \log h_d(\cdot; \boldsymbol{\delta}, \boldsymbol{\theta}) \end{bmatrix} = \mathbf{0}. \tag{4.6}$$

To find the above root, we need to use some numerical optimization technique. However, a numerical optimization in such a high dimensional space needs a huge amount of computation resources. This issue relates to computational feasibility. In addition, its estimation accuracy behaves worse because MLE may not find the global optimum but a local one. Even if all of the marginal distributions are uni-parameter distributions (i.e., $\boldsymbol{\delta_i} = \delta_i$), Eq. (4.6) consists of the $2^d - 1$ unknown parameters. For example, the 4-variate FGM copula has 4 marginal parameters and 11 dependence parameters. Thus, one should find the global maximum over the 15-dimensional space to obtain the maximum likelihood estimators. Its computation is not practical. Therefore, other simple estimation methods have been required for the parameters of the FGM copula.

### 4.2.2 Inference functions for margins estimation

To reduce the computational difficulty of the ordinary MLE, we use the IFME. In this method, we estimate the dependence parameters one by one so that each estimation result satisfies the parameter space $\Theta$ given by Eq. (4.3).

In the previous subsection, we defined the full-dimensional likelihood. Now, let us define $k$-dimensional marginal likelihood functions for $k = 1, 2, \ldots, d-1$. It is easy to see that the $k$-dimensional marginal distribution function is as follows.

$$H_k(x_1, \ldots, x_k; \cdot) = \prod_{i=1}^{k} F_i(x_i; \boldsymbol{\delta_i}) \left(1 + \sum_{p=2}^{k} \sum_{1 \le j_1 < \cdots < j_p \le k} \theta_{j_1 \cdots j_p} \overline{F}_{j_1} \cdots \overline{F}_{j_p}\right). \tag{4.7}$$

Thus, for example, we have

$$H_1(x_{j_1}; \boldsymbol{\delta_{j_1}}) = F_{j_1}, \tag{4.8}$$

$$H_2(x_{j_1}, x_{j_2}; \boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \theta_{j_1 j_2}) = F_{j_1} F_{j_2} (1 + \theta_{j_1 j_2} \overline{F}_{j_1} \overline{F}_{j_2}), \tag{4.9}$$

$$\begin{aligned} &H_3(x_{j_1}, x_{j_2}, x_{j_3}; \boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \boldsymbol{\delta_{j_3}}, \theta_{j_1 j_2}, \theta_{j_1 j_3}, \theta_{j_2 j_3}, \theta_{j_1 j_2 j_3}) \\ &= F_{j_1} F_{j_2} F_{j_3} (1 + \theta_{j_1 j_2} \overline{F}_{j_1} \overline{F}_{j_2} + \theta_{j_1 j_3} \overline{F}_{j_1} \overline{F}_{j_3} + \theta_{j_2 j_3} \overline{F}_{j_2} \overline{F}_{j_3} + \theta_{j_1 j_2 j_3} \overline{F}_{j_1} \overline{F}_{j_2} \overline{F}_{j_3}). \end{aligned} \tag{4.10}$$

The key point of these functions is that their variables are controlled by the only parameters with the same subscript, respectively. Then, we obtain the following log-likelihood functions of the $k$-dimensional marginal distribution for $k < d$ and $1 \le j_1 < \cdots < j_k \le d$.

$$\ell_{j_1} \equiv \ell_{j_1}(\boldsymbol{\delta_{j_1}}; \boldsymbol{x_{j_1}}) = \sum \log f_{j_1}(\cdot; \boldsymbol{\delta_{j_1}}), \tag{4.11}$$

$$\ell_{j_1, j_2} \equiv \ell_{j_1, j_2}(\boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \theta_{j_1 j_2}; \boldsymbol{x_{j_1}}, \boldsymbol{x_{j_2}}) = \sum \log h_2(\cdot; \boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \theta_{j_1 j_2}), \tag{4.12}$$

$$\begin{aligned} \ell_{j_1, j_2, j_3} &\equiv \ell_{j_1, j_2, j_3}(\boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \boldsymbol{\delta_{j_3}}, \theta_{j_1 j_2}, \theta_{j_1 j_3}, \theta_{j_2 j_3}, \theta_{j_1 j_2 j_3}; \boldsymbol{x_{j_1}}, \boldsymbol{x_{j_2}}, \boldsymbol{x_{j_3}}) \\ &= \sum \log h_3(\cdot; \boldsymbol{\delta_{j_1}}, \boldsymbol{\delta_{j_2}}, \boldsymbol{\delta_{j_3}}, \theta_{j_1 j_2}, \theta_{j_1 j_3}, \theta_{j_2 j_3}, \theta_{j_1 j_2 j_3}), \end{aligned} \tag{4.13}$$

$$\vdots$$

$$
\begin{aligned}
\ell_{j_1,\cdots,j_{d-1}} &\equiv \ell_{j_1,\cdots,j_{d-1}}(\boldsymbol{\delta_{j_1}},\ldots,\boldsymbol{\delta_{j_{d-1}}},\theta_{j_1j_2},\ldots,\theta_{j_1\cdots j_{d-1}};\boldsymbol{x_{j_1}},\ldots,\boldsymbol{x_{j_{d-1}}}) \\
&= \sum_{i=1}^{n} \log h_{d-1}(\cdot;\boldsymbol{\delta_{j_1}},\ldots,\boldsymbol{\delta_{j_{d-1}}},\theta_{j_1j_2},\ldots,\theta_{j_1\cdots j_{d-1}}). \qquad (4.14)
\end{aligned}
$$

In the case of IFME, we firstly estimate parameters of marginal distributions, denoted by $\widetilde{\boldsymbol{\delta}}$, as follows.

$$
\widetilde{\boldsymbol{\delta}}_{j_1} = \arg\max_{\boldsymbol{\delta_{j_1}}} \ell_{j_1}(\boldsymbol{\delta_{j_1}};\boldsymbol{x_{j_1}}), \qquad (4.15)
$$

for $j_1 = 1, 2, \ldots, d$. Next, we estimate the dependence parameters of 2-dimensional marginal distributions as follows.

$$
\widetilde{\theta}_{j_1j_2} = \arg\max_{\theta_{j_1j_2}\in\{\Theta\cap\mathscr{F}\}} \ell_{j_1,j_2}(\widetilde{\boldsymbol{\delta}}_{\boldsymbol{j_1}},\widetilde{\boldsymbol{\delta}}_{\boldsymbol{j_2}},\theta_{j_1j_2};\boldsymbol{x_{j_1}},\boldsymbol{x_{j_2}}), \qquad (4.16)
$$

where $\mathscr{F}$ is the set consisting of all estimates of the dependence parameters that are already obtained. Since $\widetilde{\boldsymbol{\theta}} \in \Theta$ must hold, we need to find the estimate on $\mathscr{F}$. As a result, $\Theta \cap \mathscr{F}$ may not contain the true value of $\theta_{j_1j_2}$. However, we can omit this issue if we have a large sample from the viewpoint of the asymptotic unbiasedness which is mentioned in the next section. Then, we have the estimates of the trivariate dependence parameters as follows.

$$
\widetilde{\theta}_{j_1j_2j_3} = \arg\max_{\theta_{j_1j_2j_3}\in\{\Theta\cap\mathscr{F}\}} \ell_{j_1,j_2,j_3}(\widetilde{\boldsymbol{\delta}}_{\boldsymbol{j_1}},\widetilde{\boldsymbol{\delta}}_{\boldsymbol{j_2}},\widetilde{\boldsymbol{\delta}}_{\boldsymbol{j_3}},\widetilde{\theta}_{j_1j_2},\widetilde{\theta}_{j_1j_3},\widetilde{\theta}_{j_2j_3},\theta_{j_1j_2j_3};\boldsymbol{x_{j_1}},\boldsymbol{x_{j_2}},\boldsymbol{x_{j_3}}). \qquad (4.17)
$$

In the same manner, we estimate all dependence parameters of higher dimensions in order. Finally, $\theta_{12\cdots d}$ is estimated in a bottom-up fashion. Thus, we can estimate all dependence parameters by repeatedly maximizing one-unknown parameter functions. In section 4, we demonstrate the estimation procedure with an example. Therefore, IFME can be used to estimate the dependence parameters of the $d$-variate FGM copula no matter how $d$ is a large number (even if MLE is infeasible). In addition, if the ordinary MLE is feasible, the estimators of IFME can be dealt with a good starting point for the numerical maximization of the full log-likelihood function [25].

## 4.3 Asymptotic normality

It is well known that the estimator of the ordinary MLE satisfies the asymptotic normality under the several conditions for the distribution function (see e.g., [7]). Because the asymptotic property means that the estimator is convergent in probability to the actual value, MLE has been frequently used as a method of the parameter estimation. As for the case of IFME, it is also known that its estimator supports the asymptotic property if the copula distribution satisfies the several conditions [25]. However, no studies assure that these properties hold in the case of the multivariate FGM copula.

In this section, therefore, we newly present that the asymptotic normalities of MLE and IFME hold for the FGM copula. Then, we compare their asymptotic efficiencies for a specific case of the parameters.

Without loss of generality, let us consider the asymptotic properties of each estimator in the case of $d = 3$ and $\boldsymbol{\delta_j} = \delta_j$. Let $\boldsymbol{\eta}$ be $(\boldsymbol{\delta}, \boldsymbol{\theta})$ (i.e., $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_7) = (\delta_1, \delta_2, \delta_3, \theta_{12}, \theta_{13}, \theta_{23}, \theta_{123})$), and a parameter with the subscript 0 denotes the true value of the parameter (e.g., $\boldsymbol{\theta_0}$ and $\boldsymbol{\delta_0}$). Then, for a random vector of the FGM copula $\boldsymbol{X} = (X_1, X_2, X_3)$, the following theorems hold.

**Theorem 4.1.** *The ordinary maximum likelihood estimators, $\widehat{\boldsymbol{\eta}}$, satisfy the following equation for $n \to \infty$.*

$$\sqrt{n}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{\mathcal{D}} N(\mathbf{0}, I^{-1}), \tag{4.18}$$

*where $\xrightarrow{\mathcal{D}}$ denotes the convergence in distribution, and $I$ is the Fisher information matrix which is given by*

$$I = -E\left[ \left. \frac{\partial^2 \log h_3(X_1, X_2, X_3; \cdot)}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^T} \right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}} \right]. \tag{4.19}$$

**Theorem 4.2.** *The estimators of IFME, $\widetilde{\boldsymbol{\eta}}$, satisfy the following equation for $n \to \infty$.*

$$\sqrt{n}(\widetilde{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{\mathcal{D}} N(\mathbf{0}, V), \tag{4.20}$$

*where*

$$V = D^{-1}M(D^{-1})^T, \tag{4.21}$$

$$D = E\left[\left.\frac{\partial \boldsymbol{s}(\boldsymbol{X};\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}}\right]$$

$$= E\left[\left.\begin{pmatrix} \frac{\partial s_1}{\partial \delta_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\partial s_2}{\partial \delta_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial s_3}{\partial \delta_3} & 0 & 0 & 0 & 0 \\ \frac{\partial s_4}{\partial \delta_1} & \frac{\partial s_4}{\partial \delta_2} & 0 & \frac{\partial s_4}{\partial \theta_{12}} & 0 & 0 & 0 \\ \frac{\partial s_5}{\partial \delta_1} & 0 & \frac{\partial s_5}{\partial \delta_3} & 0 & \frac{\partial s_5}{\partial \theta_{13}} & 0 & 0 \\ 0 & \frac{\partial s_6}{\partial \delta_2} & \frac{\partial s_6}{\partial \delta_3} & 0 & 0 & \frac{\partial s_6}{\partial \theta_{23}} & 0 \\ \frac{\partial s_7}{\partial \delta_1} & \frac{\partial s_7}{\partial \delta_2} & \frac{\partial s_7}{\partial \delta_3} & \frac{\partial s_7}{\partial \theta_{12}} & \frac{\partial s_7}{\partial \theta_{13}} & \frac{\partial s_7}{\partial \theta_{23}} & \frac{\partial s_7}{\partial \theta_{123}} \end{pmatrix}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}}\right], \tag{4.22}$$

$$M = E[\boldsymbol{s}(\boldsymbol{X};\boldsymbol{\eta_0})\boldsymbol{s}(\boldsymbol{X};\boldsymbol{\eta_0})^T], \tag{4.23}$$

$$\boldsymbol{s}(\boldsymbol{X};\boldsymbol{\eta}) = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial \delta_1} \log f_1(X_1;\cdot) \\ \frac{\partial}{\partial \delta_2} \log f_2(X_2;\cdot) \\ \frac{\partial}{\partial \delta_3} \log f_3(X_3;\cdot) \\ \frac{\partial}{\partial \theta_{12}} \log h_2(X_1, X_2;\cdot) \\ \frac{\partial}{\partial \theta_{13}} \log h_2(X_1, X_3;\cdot) \\ \frac{\partial}{\partial \theta_{23}} \log h_2(X_2, X_3;\cdot) \\ \frac{\partial}{\partial \theta_{123}} \log h_3(X_1, X_2, X_3;\cdot) \end{bmatrix}. \tag{4.24}$$

Note that we cannot obtain the explicit forms of $D$ and $V$ but can compute them for a specific $\boldsymbol{\eta_0}$. Proofs for **Theorem 4.1** and **Theorem 4.2** are given in Appendix.

Now, a natural question arises: how much is $\widehat{\boldsymbol{\eta}}$ relatively efficient compared with $\widetilde{\boldsymbol{\eta}}$? To take account of this question, we investigate one analytical example. To simplify the asymptotic covariance matrices, let us consider the case that $\boldsymbol{\delta}$ is given. That is, the parameters of interest are only $\boldsymbol{\theta}$. Suppose $(\theta_{12}, \theta_{13}, \theta_{23}, \theta_{123}) =$

$(0.3, 0.3, 0.3, 0.3)$. Then, each asymptotic covariance matrix is given by

$$I^{-1} = \begin{pmatrix} 8.48 & 0.725 & 0.725 & -0.462 \\ 0.725 & 8.48 & 0.725 & -0.462 \\ 0.725 & 0.725 & 8.48 & -0.462 \\ -0.462 & -0.462 & -0.462 & 24.18 \end{pmatrix}, \quad (4.25)$$

$$V = \begin{pmatrix} 8.70 & 0.736 & 0.736 & -0.474 \\ 0.736 & 8.70 & 0.736 & -0.474 \\ 0.736 & 0.736 & 8.70 & -0.474 \\ -0.474 & -0.474 & -0.474 & 24.18 \end{pmatrix}. \quad (4.26)$$

Since the $i$-th diagonal element of Eq. (4.25) is less than that of Eq. (4.26) for $i = 1, \ldots, 4$, we can find that MLE is more effective than IFME. For example, the variance of $\widehat{\theta}_{12}$ is around 97% ($= 8.48/8.70$) of the variance of $\widetilde{\theta}_{12}$. In the next section, we verify these results through simulation.

## 4.4 Numerical study

### 4.4.1 Random variate

To perform a simulation, we introduce a method that generates random variate from the $d$-variate FGM copula with arbitrary marginal distributions. Here, *the conditional method* described in [25] allows us to generate a random sequence from the $d$-variate FGM copula. Let $(X_1, \ldots, X_d)$ be a random vector of the $d$-variate FGM copula. For $i = 2, 3, \ldots, d$, we define the conditional distribution functions as follows.

$$H_{i|12\cdots i-1}(x_i|x_1, \ldots, x_{i-1}) = \Pr[X_i \leq x_i | X_1 = x_1, \ldots, X_{i-1} = x_{i-1}]. \quad (4.27)$$

Let $v_1, \ldots, v_d$ be samples from the i.i.d. uniform distribution in the interval $(0, 1)$. Here, we construct a random sequence through the following processes.

- Let $u_1 = H_1^{-1}(v_1)$.
- Let $u_2 = H_{2|1}^{-1}(v_2|u_1)$,$\ldots$,

- Let $u_d = H^{-1}_{d|1\cdots d-1}(v_d|u_1, \ldots, u_{d-1})$.
- Return $(x_1, x_2, \ldots, x_d) = (F_1^{-1}(u_1; \boldsymbol{\delta_1}), F_2^{-1}(u_2; \boldsymbol{\delta_2}), \ldots, F_d^{-1}(u_d; \boldsymbol{\delta_d}))$.

Consequently, such a $(x_1, \ldots, x_d)$ follows the $d$-variate FGM copula. Thus, we need to know the closed form of $H^{-1}_{i|1\cdots i-1}(\cdot|\cdot)$ to perform the above process. Here, we present it as follows. First, the conditional distribution function is given by

$$H_{i|1\cdots i-1}(u_i|u_1, \ldots, u_{i-1}) = \int_0^{u_i} \frac{h_i(u_1, \ldots, u_{i-1}, u)}{h_{i-1}(u_1, \ldots, u_{i-1})} du. \tag{4.28}$$

With the FGM copula, since

$$h_i(u_1, \ldots, u_i) = h_{i-1}(u_1, \ldots, u_{i-1}) + \sum_{p=2}^{i} \sum_{1 \le j_1 < \cdots < j_p = i} \theta_{j_1 \cdots j_p}(1 - 2u_{j_1}) \cdots (1 - 2u_{j_p}), \tag{4.29}$$

it leads that

$$\frac{h_i(u_1, \ldots, u_i)}{h_{i-1}(u_1, \ldots, u_{i-1})} = 1 + \frac{\sum_{p=2}^{i} \sum_{1 \le j_1 < \cdots < j_p = i} \theta_{j_1 \cdots j_p}(1 - 2u_{j_1}) \cdots (1 - 2u_{j_p})}{h_{i-1}(u_1, \ldots, u_{i-1})}. \tag{4.30}$$

Let us define $D$ as

$$D = \frac{\sum_{p=2}^{i} \sum_{1 \le j_1 < \cdots < j_p = i} \theta_{j_1 \cdots j_p}(1 - 2u_{j_1}) \cdots (1 - 2u_{j_{p-1}})}{h_{i-1}(u_1, \ldots, u_{i-1})}. \tag{4.31}$$

Thus, by integrating both sides of Eq. (4.30), we obtain

$$H_{i|1\cdots i-1}(u_i|u_1, \ldots, u_{i-1}) = (1 + D)u_i - Du_i^2, \tag{4.32}$$

where

$$D = \frac{\sum_{p=2}^{i} \sum_{1 \le j_1 < \cdots < j_p = i} \theta_{j_1 \cdots j_p}(1 - 2u_{j_1}) \cdots (1 - 2u_{j_{p-1}})}{h_{i-1}(u_1, \ldots, u_{i-1})}. \tag{4.33}$$

From $Du_i^2 - (1 + D)u_i + H_{i|1\cdots i-1}(u_i|u_1, \ldots, u_{i-1}) = 0$, we are able to represent $u_i$ by $D$ and $H_{i|1\cdots i-1}(u_i|u_1, \ldots, u_{i-1})$ based on the quadratic formula. Hence,

$H^{-1}_{i|1\cdots i-1}(v_i|u_1,\dots,u_{i-1})$ takes one of the values of

$$\frac{1+D\pm\sqrt{(1+D)^2-4Dv_i}}{2D}, \tag{4.34}$$

for which it is greater than 0 and less than 1. For example, for $i=2$, we have

$$H^{-1}_{2|1}(v_2|u_1)=\frac{1+\theta_{12}(1-2u_1)+\sqrt{(1+\theta_{12}(1-2u_1))^2-4\theta_{12}(1-2u_1)v_2}}{2\theta_{12}(1-2u_1)}. \tag{4.35}$$

As far as we have found, the signs of Eq. (4.34) for $i=2,3$ and 4 are $+,-$, and $-$, respectively. Consequently, we can generate samples of the multivariate FGM copula by the conditional method with Eq. (4.34).

### 4.4.2 Short example

In this subsection, we present the detailed procedure of IFME for the FGM copula and the asymptotic normality of its estimators. The several settings of this example are as follows.

1. The sample size is 30000.
2. Assume that all of the marginal distributions are given (i.e., the target parameters are the only dependence parameters).
3. For $d=3$, $(\theta_{12},\theta_{13},\theta_{23},\theta_{123})=(0.3,0.3,0.3,0.3)$.

We generate the random samples under the situation, and then we estimate the parameters from the samples. The procedure of generating the samples from the FGM copula is discussed in the previous subsection. Here, we compute the estimates through the following steps.

**Step 1** $\widetilde{\theta}_{12}$ is given by $\underset{-1\le\theta 12\le 1}{\arg\max}\,\ell_{12}$. $\widetilde{\theta}_{12}=0.304$.

**Step 2** $\widetilde{\theta}_{13}$ is given by $\underset{\theta_{13}\in\{\Theta\cap\theta_{12}=0.304\}}{\arg\max}\,\ell_{13}$. $\widetilde{\theta}_{13}=0.304$.

**Step 3** $\widetilde{\theta}_{23}$ is given by $\underset{\theta_{23}\in\{\Theta\cap\theta_{12}=0.304\cap\theta_{13}=0.304\}}{\arg\max}\,\ell_{23}$. $\widetilde{\theta}_{23}=0.310$.

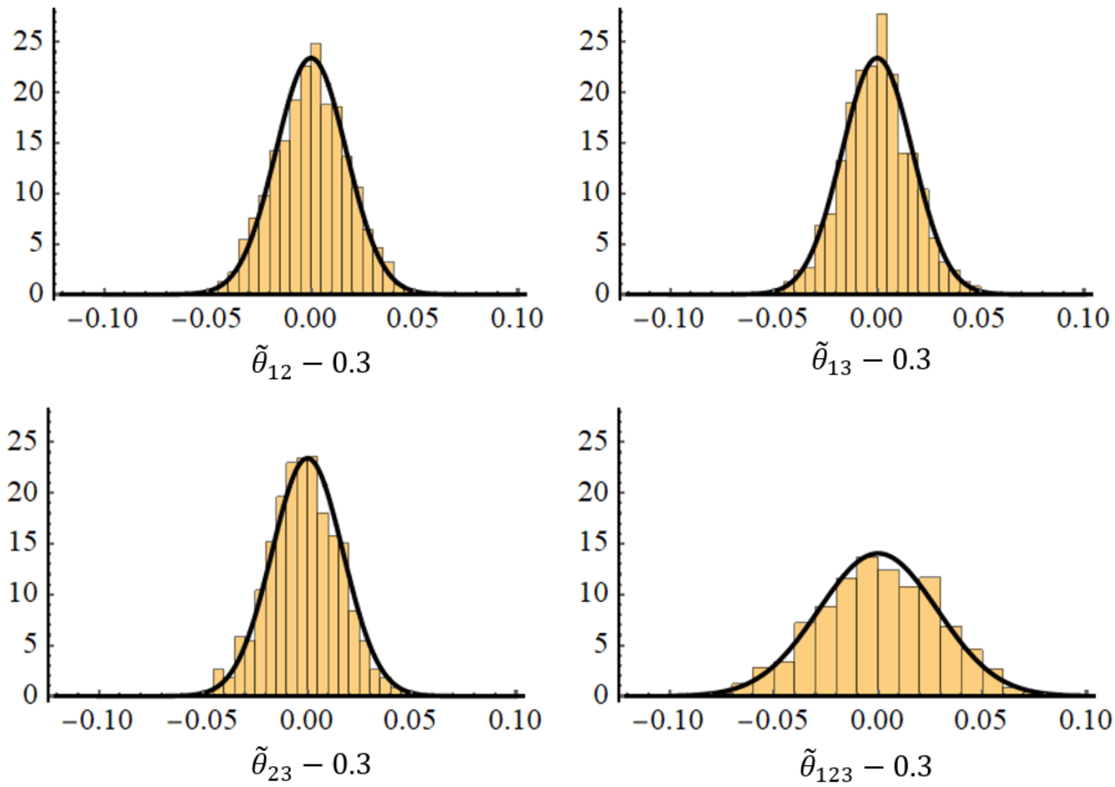


Figure 4.1: Histogram for each estimator and the theoretically derived PDF of its asymptotic distribution whose variance is a diagonal element of Eq. (4.26).

**Step 4** $\widetilde{\theta}_{123}$ is given by $\underset{\theta_{123}\in\{\Theta\cap\theta_{12}=0.304\cap\theta_{13}=0.304\cap\theta_{23}=0.310\}}{\arg\max}\ \ell$. $\widetilde{\theta}_{123} = 0.310$.

Note that the steps 1, 2 and 3 are exchangeable although we estimate the dependence parameters in the order of $\theta_{12}, \theta_{13}$ and $\theta_{23}$ in this example. If the order of the steps is changed, the estimation result is different from the above result.

Next, we also investigate how well these estimates fit to the asymptotic distribution. We iterate above simulation 1000 times, and then obtain 1000 samples of each estimate. Figure 4.1 depicts the histogram of these samples with the PDF of the asymptotic distribution whose covariance matrix is given by Eq. (4.26). Since the histograms well fit the asymptotic distribution, the figure implies that each estimator holds the asymptotic normality. Actually, this result satisfies the Kolmogorov-Smirnov test for multivariate normality at with the 5% significance level.

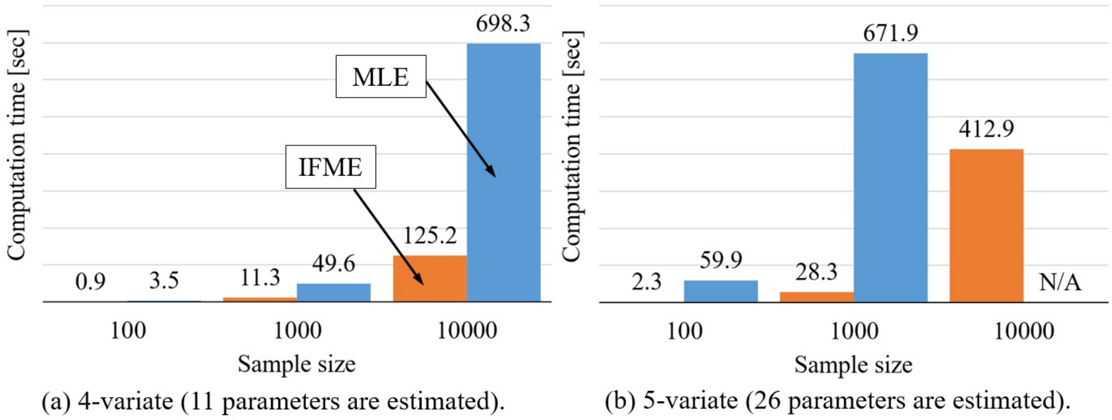


Figure 4.2: Average computation times of estimating all dependence parameters of the $d$-variate FGM copula for $d = 4, 5$ (number of iterations: 5).

### 4.4.3 Computation time

In this subsection, we investigate the computation times of MLE and IFME. We use a PC equipped with Intel(R)Core(TM) i7-6700k 4.00GHz, 32GB RAM, running on Windows 10 professional 64bit. We perform simulation under the following settings.

1. The dimensions $d$ are 4 and 5.
2. The sample sizes, $n$, are $100, 1000, 10000$.
3. Assume that all of the marginal distributions are given (i.e., the target parameters are the only dependence parameters).
4. In the case of $d = 4$, the values of the dependence parameters are as follows: $\theta_{j_1j_2} = 0, \theta_{j_1j_2j_3} = 0.2, \theta_{1234} = 0.2$ for $1 \leq j_1 < j_2 < j_3 \leq 4$.
5. In the case of $d = 5$, the values of the dependence parameters are as follows: $\theta_{j_1j_2} = 0, \theta_{j_1j_2j_3} = 0, \theta_{j_1j_2j_3j_4} = 0.2, \theta_{12345} = 0.2$ for $1 \leq j_1 < j_2 < j_3 < j_4 \leq 5$.

Figure 4.2 depicts the computation times to estimate the dependence parameters under the above situation. We can find that IFME works faster than MLE in all of the cases, and both of MLE and IFME spend more computation time as $d$ and $n$ increase. In particular, for $d = 5$ and $n = 10000$, MLE cannot work due to out of memory (it is denoted by the symbol "N/A"). Therefore, estimating the parameters of the multivariate FGM copula, we should not use MLE but IFME from the viewpoint of the computation costs. For future works, a theoretical analysis of the computational complexity of IFME is required.

### 4.4.4 Performance evaluation

Now, we investigate the estimation accuracy of IFME through Monte Carlo simulation. In this subsection, we again consider that all of the marginal distributions are given. For $d = 4$, we perform the simulation under the following settings.

1. The sample sizes, $n$, are $100, 1000, 10000$, and the number of iteration times, $m$, is 100.
2. All of the marginal distributions are the i.i.d. uniform distribution in the interval $[0, 1]$. It means that the target parameters are only the dependence parameters.
3. We deal with the following four situations of the dependence parameters.
   (a) $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234})$
   $= (0, 0, 0, 0, 0, 0, 0.2, 0.2, 0.2, 0.2, 0.2)$.
   (b) $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234})$
   $= (0, 0, 0, 0, 0, 0, 0.2, 0.2, 0.2, 0.2, 0.5)$.
   (c) $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234})$
   $= (0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2)$.
   (d) $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234})$
   $= (0, 0, 0, 0, 0, 0, -0.2, -0.2, -0.2, -0.2, -0.2)$.

For the settings, we calculate the following *Bias* and *RMSE* (Root-Mean-Squared error) as the indicators of the estimation accuracy for each estimator.

$$Bias(\widetilde{\boldsymbol{\theta}}) = \begin{bmatrix} \frac{1}{m}\sum_{i=1}^{m}(\widetilde{\theta}_{12,i} - \theta_{12(0)}) \\ \vdots \\ \frac{1}{m}\sum_{i=1}^{m}(\widetilde{\theta}_{1234,i} - \theta_{1234(0)}) \end{bmatrix},$$

$$RMSE(\widetilde{\boldsymbol{\theta}}) = \begin{bmatrix} \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\widetilde{\theta}_{12,i} - \theta_{12(0)})^2} \\ \vdots \\ \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\widetilde{\theta}_{1234,i} - \theta_{1234(0)})^2} \end{bmatrix},$$

where $\widetilde{\theta}_{\cdot,i}$ is the $\widetilde{\theta}_{\cdot}$'s estimate obtained in the $i$-th iteration step and $\widetilde{\theta}_{\cdot,(0)}$ is the true value of $\widetilde{\theta}_{\cdot}$. *Bias* and *RMSE* mean the average and dispersion of the estimation errors, respectively.

The numerical results are presented in Tables 4.1, 4.2, 4.3 and 4.4. These tables lead to the following remarks. The estimation accuracy is improved with a power of 10 samples. In order to obtain estimates with Bias of less than 0.1, we need 100, 1000 and 10000 samples for $\theta_{j_1j_2}, \theta_{j_1j_2j_3}$, and $\theta_{1234}$, respectively. On the other hand, *Bias* tends to be a negative value for small samples. This means that the proposed method estimates lower than the true values of the dependence parameters (e.g., in Tab. 4.2 for $\widetilde{\theta}_{1234}$ and $n = 100$, the mean value of the estimate is 0.051 even though its true value is 0.5).

## 4.5 Conclusion

In this chapter, we have presented how to estimate the parameters of the multivariate FGM copula in practical computation time. Consequently, the computational difficulties of the ordinary MLE have been reduced by using the IFM framework. Because of this, IFME can estimate all of the dependence parameters of the FGM copula no matter how the dimension $d$ is high. Then, we have revealed that the estimators of IFME satisfy the asymptotic normality in the case of the FGM copula. Finally, we have numerically shown the computation time and estimation accuracy of IFME. In conclusion, IFME is computationally easier than MLE, and the estimation accuracies are not much difference between them. Therefore, IFME is a useful approach to estimate the parameters of the FGM copula.

From the viewpoint of reliability engineering, this result can be applied to a quantitative evaluation of the dependence among system components. If we have the lifetime data of the components, we can find the latent dependence structure among them by estimating the parameters of the multivariate FGM copula with arbitrary marginal distributions. Moreover, our estimation method is especially applicable not only the reliability analysis but also other research fields. Therefore, this result also contributes to the evolution of statistical theory.

Table 4.1: Estimation result $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234}) = (0, 0, 0, 0, 0, 0, 0.2, 0.2, 0.2, 0.2, 0.2)$.

| $Bias$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | 0.017 | -0.036 | 0.009 | -0.019 | 0.009 | 0.026 | -0.133 | -0.103 | -0.133 | -0.169 | -0.157 |
| $n = 1000$ | 0.015 | -0.002 | -0.016 | -0.013 | -0.002 | 0.011 | -0.009 | -0.002 | -0.025 | 0.002 | -0.106 |
| $n = 10000$ | 0.005 | -0.002 | 0.001 | -0.004 | 0.001 | 0.001 | 0.001 | 0.002 | 0.001 | 0.001 | 0.003 |
| $RMSE$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
| $n = 100$ | 0.325 | 0.304 | 0.297 | 0.294 | 0.307 | 0.273 | 0.333 | 0.328 | 0.307 | 0.325 | 0.300 |
| $n = 1000$ | 0.093 | 0.095 | 0.092 | 0.091 | 0.088 | 0.095 | 0.163 | 0.158 | 0.173 | 0.132 | 0.201 |
| $n = 10000$ | 0.032 | 0.029 | 0.029 | 0.029 | 0.033 | 0.029 | 0.057 | 0.048 | 0.053 | 0.050 | 0.079 |

Table 4.2: Estimation result $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234}) = (0, 0, 0, 0, 0, 0, 0.2, 0.2, 0.2, 0.2, 0.5)$.

| $Bias$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | 0.017 | -0.036 | 0.008 | -0.019 | 0.009 | 0.029 | -0.133 | -0.102 | -0.138 | -0.167 | -0.449 |
| $n = 1000$ | 0.015 | -0.002 | -0.016 | 0.013 | -0.002 | 0.011 | -0.009 | -0.002 | -0.024 | 0.003 | -0.323 |
| $n = 10000$ | 0.005 | -0.002 | 0.001 | -0.004 | 0.001 | 0.001 | 0.001 | 0.003 | 0.001 | 0.001 | -0.065 |
| $RMSE$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
| $n = 100$ | 0.325 | 0.304 | 0.297 | 0.294 | 0.307 | 0.272 | 0.334 | 0.334 | 0.311 | 0.325 | 0.516 |
| $n = 1000$ | 0.093 | 0.095 | 0.093 | 0.091 | 0.089 | 0.095 | 0.163 | 0.158 | 0.173 | 0.134 | 0.371 |
| $n = 10000$ | 0.032 | 0.029 | 0.029 | 0.029 | 0.033 | 0.029 | 0.057 | 0.048 | 0.053 | 0.050 | 0.098 |

Table 4.3: Estimation result $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234}) = (0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2)$.

| *Bias* | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | 0.014 | -0.034 | 0.012 | -0.039 | 0.000 | 0.003 | -0.143 | -0.121 | -0.144 | -0.165 | -0.052 |
| $n = 1000$ | 0.015 | -0.001 | -0.016 | -0.011 | -0.004 | 0.011 | -0.013 | -0.006 | -0.019 | 0.011 | -0.100 |
| $n = 10000$ | 0.005 | -0.002 | 0.001 | -0.003 | 0.001 | 0.000 | 0.000 | 0.003 | 0.001 | 0.002 | 0.001 |
| *RMSE* | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
| $n = 100$ | 0.324 | 0.298 | 0.298 | 0.284 | 0.273 | 0.263 | 0.313 | 0.280 | 0.296 | 0.290 | 0.244 |
| $n = 1000$ | 0.090 | 0.096 | 0.089 | 0.089 | 0.089 | 0.097 | 0.157 | 0.150 | 0.171 | 0.137 | 0.196 |
| $n = 10000$ | 0.032 | 0.028 | 0.029 | 0.029 | 0.032 | 0.030 | 0.055 | 0.048 | 0.051 | 0.050 | 0.081 |

Table 4.4: Estimation result $(\theta_{12}, \theta_{13}, \theta_{14}, \theta_{23}, \theta_{24}, \theta_{34}, \theta_{123}, \theta_{124}, \theta_{134}, \theta_{234}, \theta_{1234}) = (0, 0, 0, 0, 0, 0, -0.2, -0.2, -0.2, -0.2, -0.2)$.

| $Bias$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | 0.017 | -0.034 | 0.010 | -0.023 | 0.002 | 0.022 | 0.079 | 0.145 | 0.148 | 0.170 | 0.226 |
| $n = 1000$ | 0.015 | -0.002 | -0.016 | -0.013 | -0.003 | 0.009 | -0.004 | 0.002 | 0.003 | 0.044 | 0.148 |
| $n = 10000$ | 0.005 | -0.002 | 0.001 | -0.004 | 0.001 | 0.000 | 0.001 | 0.003 | 0.002 | 0.002 | 0.055 |
| $RMSE$ | $\widetilde{\theta}_{12}$ | $\widetilde{\theta}_{13}$ | $\widetilde{\theta}_{14}$ | $\widetilde{\theta}_{23}$ | $\widetilde{\theta}_{24}$ | $\widetilde{\theta}_{34}$ | $\widetilde{\theta}_{123}$ | $\widetilde{\theta}_{124}$ | $\widetilde{\theta}_{134}$ | $\widetilde{\theta}_{234}$ | $\widetilde{\theta}_{1234}$ |
| $n = 100$ | 0.325 | 0.302 | 0.298 | 0.298 | 0.302 | 0.270 | 0.309 | 0.332 | 0.304 | 0.306 | 0.335 |
| $n = 1000$ | 0.093 | 0.093 | 0.093 | 0.089 | 0.087 | 0.097 | 0.165 | 0.159 | 0.162 | 0.146 | 0.244 |
| $n = 10000$ | 0.032 | 0.028 | 0.029 | 0.029 | 0.033 | 0.029 | 0.057 | 0.047 | 0.053 | 0.051 | 0.101 |

# Chapter 5

# Dependent failure detection

## 5.1 Introduction

In this chapter, we propose a statistical detection method of the dependent failure occurrence in $n$-component parallel systems that utilizes the failure occurrence times of the components. If we assume that the lifetime distribution of the components worsens if $k$ out of $n$ components failed, the dependent failure occurrence can be found by identifying the change of the distribution. The performance of the proposed method is demonstrated by simulation studies. This results is given by Ota and Kimura [43] and Ota and Kimura [45].

Cascading failure, one kind of dependent failures, is a phenomenon in which the failure occurrence of one component triggers other failures [1, 9, 18]. In this case, the lifetimes of the triggered components depend on the lifetime of the trigger component. For example, this phenomenon is observed in complex network systems such as blackouts of power transmission systems. On 30th and 31st July, 2012, the northern region of India experienced large blackouts due to the late arrival of monsoons that caused excessive demands for the power [54]. Indeed, a disturbance in a part of the power grid triggered the cascading failure of the power transmission systems. In this way, the cascading failure tends to occur in network systems whose components cannot be immediately repaired/replaced after their failures. Dobson, Carreras and Newman [9] have also reported that the trigger component can cause a huge number of failures after it fails if other components strongly rely on the trigger component. Consequently, if a cascading failure occurs in a redundant system, the system cannot deliver the designed performance [9, 56].

The related works of the cascading failure have been mainly focused on the modeling and reliability assessment for the actual cases. Dobson, Carreras and Newman [9] proposed a stochastic model of load-dependent cascading failure in electric power transmission systems. Watts [60] studied the conditions that the cascading failure occurs in large-scale network systems by developing a stochastic model. Bialek et al. [1] surveyed the state of the art in the cascading failure with a lot of the actual incidents. These important studies are helpful in understanding the mechanisms of the cascading failure and suggesting the ways to reduce particular factors of cascading failure occurrences.

From the viewpoint of reliability management, the factors of cascading failure should be eliminated at the design or test stages. At the design stage, which is located before the system implementation, Failure Mode and Effect Analysis and Fault Tree Analysis [28, 36, 57] are effective methods in identifying the factors behind dependent failure occurrences in a redundant system. This is because these two methods can explain the logical structure of the components. However, system designers cannot always be aware of all the likely causes of the dependent failure (cf. Section 11 in [42]). This fact implies the importance of the test stage, which is performed after the implementation.

As for the test stage, detection methods of the dependent failure occurrence are useful. In general, it is assumed that anomaly of the system tend to cause the dependent failure occurrence, and there are a lot of statistical detection methods of such anomaly phenomena in the literature. For example, McCool [31] proposed a detection method for a parallel system with 2 components whose lifetimes follow identical Weibull distributions. the anomaly detection methods for the monitored signal data were well studied by [58, 59]. Rocco and Zio [53] used an SVM (Support Vector Machine) in order to detect a specific anomaly in a nuclear power plant. Zhang, Lin and Karim [63] studied the anomaly detection in a hydroelectric plant with the signals captured from several parts of the units. However, the detection of the anomaly events and locating them are difficult as the system structure becomes more complex despite that these are essential topics in reliability engineering [10, 21].

Ota and Kimura [44] have recently developed and proposed a statistical test method that detects the dependent failure occurrence in a redundant system by analyzing the failure occurrence times of the components. They dealt with a parallel system having two identical components as a simple network model. Thus, this study aims at evaluating whether or not the first failure occurrence worsens the reliability of the surviving component. The result implies the possibility that the dependent failure occurrence in other systems can be also detected.

Hence, we expand the modeling framework to $n$-component parallel redundant systems in this chapter. That is, we propose the detection method of the dependent failure occurrence for the system. Note that, we regard the cascading failure as the dependent failure in this chapter. The novelty of our research is that this method can specify the trigger component of the cascading failure and the time point at which a cascading failure occurred. Therefore, this result contributes to the cause analysis of the dependent failure occurrence in the $n$-component parallel system.

In the next section, we explain a stochastic model that represents both of independent and dependent failure occurrences in a parallel system. In Section 5.3, we present an algorithm for the dependent failure detection method with failure occurrence times. The accuracy of the detection method is demonstrated by numerical studies in Section 5.4. We conclude our study in Section 5.5 with a summary.

## 5.2 Failure model

This model expresses both independent and dependent failure occurrences by the switching mechanism of the hazard rate functions. Let $T_i$ and $t_i$ be the random variable and the observed value of the $i$-th failure occurrence time for $i = 1, 2, \ldots, n$, respectively. Suppose the lifetimes of the components follow the $i$-th conditional hazard rate function $h_i(t|t_1, t_2, \ldots, t_{i-1})$ in the time interval between $t_{i-1}$ and $t_i$. Then, we consider the $n$-component parallel system and the dependent failure occurrence with the following assumptions.

**Assumption I :** The parallel system is constructed by $n$ components whose lifetimes are independent and identically distributed.

**Assumption II :** The hazard rate function of each component is one of the fol-

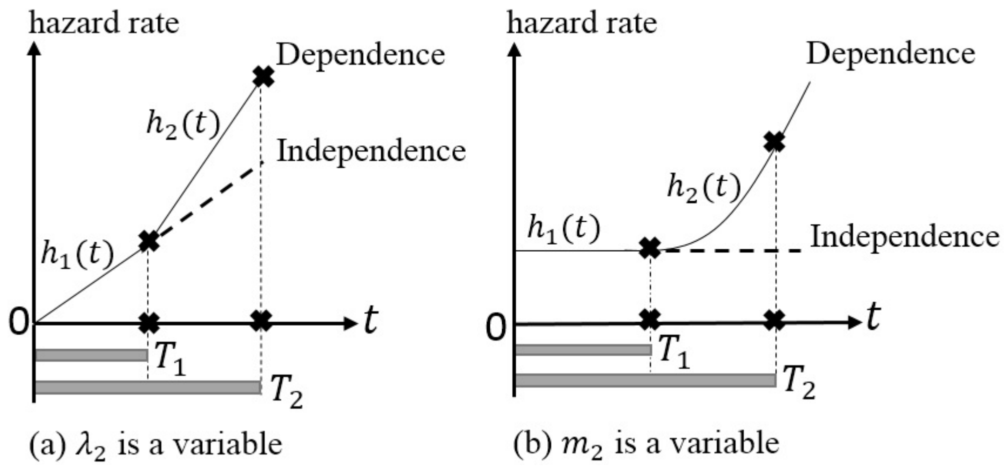


Figure 5.1: Behavior of the hazard rate function [43].

lowing: DFR (decreasing failure rate), CFR (constant failure rate) or IFR (increasing failure rate).

**Assumption III :** At $t = 0$, the hazard rate function of each component is given by

$$h_1(t) = \lambda_1 m_1 t^{m_1 - 1} \quad (\lambda_1, m_1 > 0). \tag{5.1}$$

This assumption describes that the lifetime distribution of the $n$ components is an identical Weibull distribution, which is frequently used for the modeling of failure data [37].

**Assumption IV :** If the $i$-th failure occurs at $t_i$, the $i$-th hazard rate function $h_i(t|t_1, \ldots, t_{i-1})$ switches to $h_{i+1}(t|t_1, \ldots, t_i)$. This conditional hazard rate function $h_{i+1}(t|t_1, \ldots, t_i)$ is given by

$$\begin{aligned} &h_{i+1}(t|t_1, \ldots, t_i) \\ &\quad = g_{i+1}(t) - g_{i+1}(t_i) + h_i(t_i|t_1, \ldots, t_{i-1}) \quad (t > t_i), \end{aligned} \tag{5.2}$$

where

$$g_i(t) = \lambda_i m_i t^{m_i - 1} \quad (\lambda_i, m_i > 0, i = 1, 2, \ldots, n). \tag{5.3}$$

From the above, this model has $2n$ parameters denoted by $\lambda_i$ and $m_i$ for $i = 1, 2, \ldots, n$. Figure 5.1 shows two examples of the hazard rate functions in the case of two components. The lifetimes of the two components obey the distribution function with the hazard rate function $h_1(t)$ until the first failure occurs. After the first failure occurrence, the lifetime of the surviving component obeys $h_2(t)$. In the case of the graph on the left (a), since the hazard rate function switches, the surviving component breaks down sooner than originally expected. This is because the residual lifetime of the surviving component, $T_2 - T_1$, is statistically shortened in the sense that $\lambda_2$ is greater than $\lambda_1$. The graph on the right (b) demonstrates the switching from CFR to IFR. In Section 5.4, we generate the lifetimes of the components by Monte Carlo simulation [19] based on this model in order to evaluate the accuracy of the proposed method.

This model describes one failure occurrence phenomenon by setting the model parameters. We call such a parameter setting *"failure case"*. It is the aim of our research to investigate how the detection method accurately detects the dependent failure occurrences for the various failure cases. Moreover, the failure cases discussed in this chapter are the following cases: (i) no changes of the hazard rate functions occur, and (ii) the hazard rate functions do not change twice but only once at any of $T_1, T_2, \ldots,$ or $T_{n-1}$.

We investigate behaviors of CDF of $T_2$ in the case of the independent and dependent failures. Figure 5.2 illustrates the CDFs of $T_2$ when only $\lambda_2$ is a variable. In this figure, the functions in the case of the dependent failures are apart from the independent one as $\lambda_2$ increases. Figure 5.3 also indicates that the same fact holds between $m_2$ and the CDFs of $T_2$. Note that $F_2(t)$ is written by

$$F_2(t) = \int_0^t G_2(t - t_1|t_1) f_1(t_1) dt_1, \tag{5.4}$$

where $G_2(t|t_1)$ denotes the conditional CDF of $h_2(t|t_1)$ (i.e., $G_2(t|t_1) = 1 - e^{-\int_0^t h_2(s+t_1|t_1)ds}$ for $t \geq 0$), and $f_1(t)$ denotes the PDF of $T_1$. Because of its complexity, $F_2(t)$ is not an elementary function except for some of the failure cases.

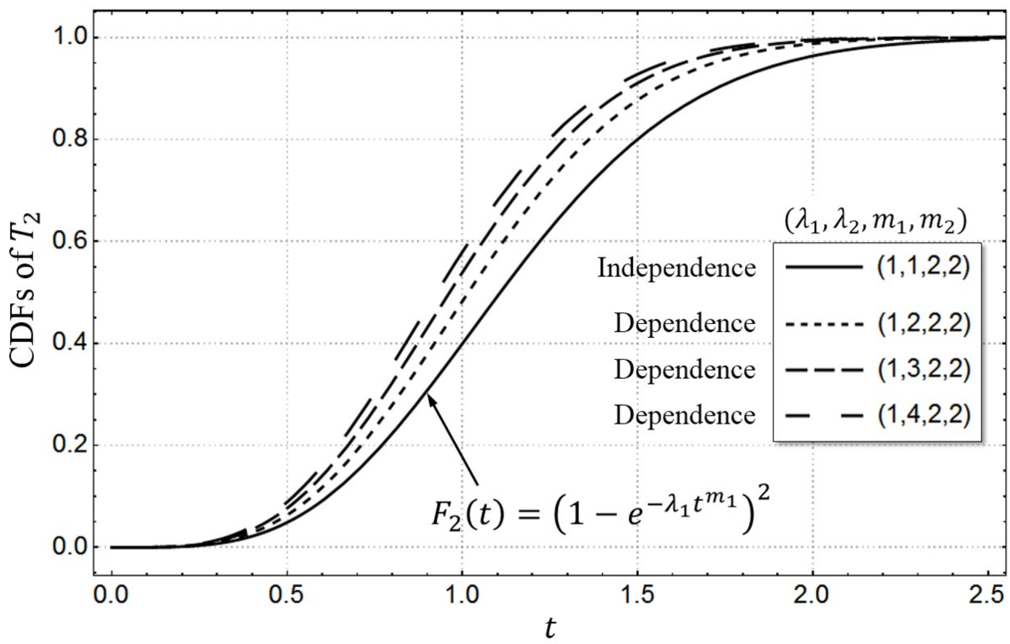


Figure 5.2: Cumulative distribution functions of $T_2$ when only $\lambda_2$ is a variable.

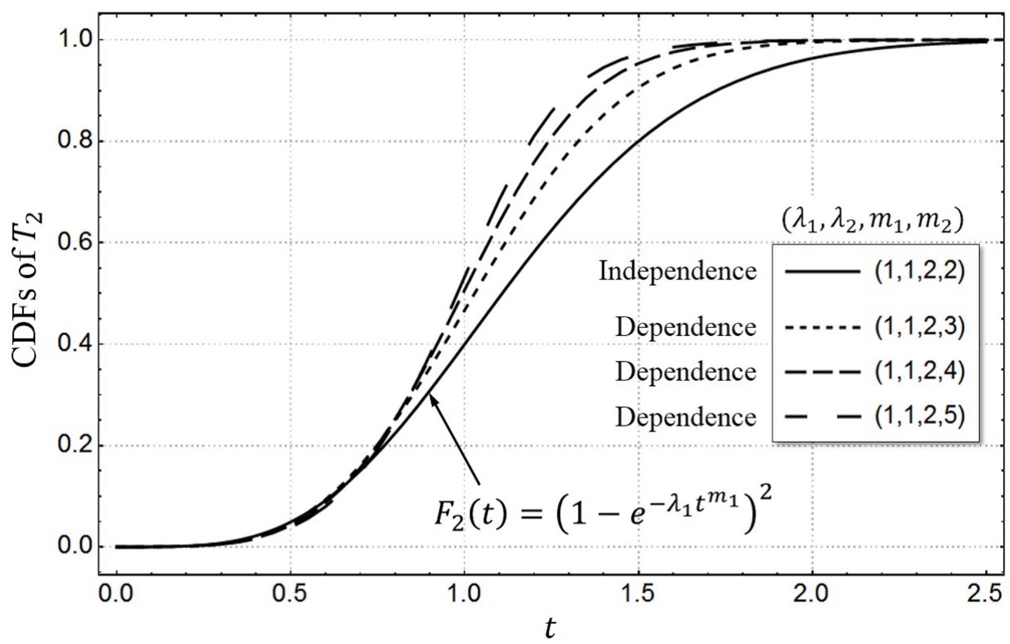


Figure 5.3: Cumulative distribution functions of $T_2$ when only $m_2$ is a variable.

## 5.3 Method

In this section, we propose a statistical detection method of the dependent failure occurrence for the $n$-component parallel system based on the samples of failure occurrence times. First, we discuss the detection method of the dependent failure occurrence for the 2-component parallel system. As Ota and Kimura [44] mentioned, the key point of the detection method of the dependent failure occurrence is to statistically identify the change of the hazard rate function. After that, we expand the method to the case of the $n$-component parallel system.

### 5.3.1 Algorithm for 2-component parallel system

As preliminaries, let $X_1$ and $X_2$ be random variables of the lifetimes of two components, which follow an identical Weibull distribution function $F(t) = 1 - e^{-\lambda_1 t^{m_1}}$. If $X_1$ and $X_2$ are independent, we obtain the following relationships. CDF of $T_1$ is given by $F_1(t) = 1 - e^{-2\lambda_1 t^{m_1}}$, and CDF of $T_2$ is equal to $F_2(t) = (1 - e^{-\lambda_1 t^{m_1}})^2$, since $T_1$ corresponds to $\mathrm{Min}(X_1, X_2)$ and $T_2$ corresponds to $\mathrm{Max}(X_1, X_2)$. Note that these functions have common parameters $\lambda_1$ and $m_1$. Hence, we can estimate $F_2(t)$ with the sample of $T_1$ if the independent failure occurs [44].

In order to detect the dependent failure occurrence, let us consider the following value.

$$\sup_{0 \leq t < \infty} \left| (1 - e^{\lambda_1 t^{m_1}})^2 - F_2(t) \right|. \tag{5.5}$$

In Eq. (5.5), the first term corresponds to $F_2(t)$ that holds under the condition of the independent failure, and the second one is given by Eq. (5.4). That is, the value of Eq. (5.5) is theoretically 0 if and only if the independent failure occurs, otherwise greater than 0 (i.e., in the case that the change of the hazard rate function occurs). Furthermore, we can infer from Figures 5.2 and 5.3 that the value of Eq. (5.5) increases as $\lambda_2 (\geq \lambda_1)$ or $m_2 (\geq m_1)$ increases. Accordingly, these facts indicate that the dependent failure occurrence can be detected by performing a goodness-of-fit-test between the empirical distribution of $T_2$ and the estimated $F_2(t)$ by the sample of $T_1$.

The detection method judges by performing a goodness-of-fit test between the sample of $T_2$ and $\widehat{F}_2(t)$ estimated with the sample of $T_1$. First, we estimate $\widehat{F}_1(t)$

with the sample of $T_1$ by the method of maximum likelihood (i.e., we obtain $\widehat{\lambda_1}$ and $\widehat{m_1}$). Next, we can obtain $\widehat{F}_2(t)$ by using $\widehat{\lambda_1}$ and $\widehat{m_1}$ with the assumption that these components failed independently. Finally, we perform a goodness-of-fit test between the sample of $T_2$ and $\widehat{F}_2(t)$. That is, we investigate whether or not the estimated value of Eq. (5.5) is statistically significant. If $\widehat{F}_2(t)$ fits the data, it implies that the independent failure occurred in the system. If not, the dependent failure is considered to have occurred. Thus, the accuracy of the test is affected by the parameters $\lambda_i$ or $m_i$ $(i = 1, 2)$ because Eq. (5.5) contains them. From these points of view, we use the Kolmogorov-Smirnov (K-S) test [13] as a goodness-of-fit test.

The algorithm to detect the dependent failure occurrence is as follows.

**Algorithm 1: 2-component parallel system**

(i) Estimate $\widehat{F_1}(t) = 1 - e^{-2\widehat{\lambda}_1 t^{\widehat{m}_1}}$ by the method of maximal likelihood with the sample of $T_1$.

(ii) Set $\widehat{F_2}(t) = (1 - e^{-\widehat{\lambda}_1 t^{\widehat{m}_1}})^2$ with the assumption that the independent failure occurs.

(iii) Calculate K-S statistics $D$ as

$$D = \sup_{0 \le t < \infty} |\widehat{F_2}(t) - S_2(t)|,$$

where $S_2(t)$ denotes the empirical distribution function [13] of the sample of $T_2$. Let $(t_{21}, t_{22}, \ldots, t_{2N})$ be $N$ random samples of $T_2$. Then,

$$S_2(t) = \begin{cases} 0 \quad (0 \le t < t_{21}) \\ k/N \quad (t_{2k} <\le t < t_{2k+1}) \\ 1 \quad (t \ge t_{2N}) \end{cases} \tag{5.6}$$

(iv) Set the null hypothesis $H_0$ and the alternative hypothesis $H_1$ as follows.

$H_0$: The sample of $T_2$ obeys $\widehat{F_2}(t)$.

$H_1$: The sample of $T_2$ does not obey $\widehat{F_2}(t)$.

(v) Perform K-S test with the hypotheses.

If $\sqrt{N}D > 1.36$, reject $H_0$ with the significance level of 5% and accept $H_1$.

If not, accept $H_0$.

(vi) If $H_0$ is accepted, the independent failure occurred. If $H_1$ is accepted, the dependent failure occurred.

---

In this algorithm, $\widehat{F_2}(t)$ is the substitution of the first term of Eq. (5.5) and $S_2(t)$ is the substitution of $F_2(t)$. Consequently, if the null hypothesis $H_0$ is accepted on the test, we can state that an independent failure occurred in the 2-component parallel system under the significance level of 5% (N.B., $\Pr[\sqrt{N}D > 1.36] = 0.05$ under the null hypothesis $H_0$ as $N \to \infty$). If not, the dependent failure occurred.

### 5.3.2 Algorithm for $n$-component parallel system

In this subsection, we propose the detection method of the dependent failure occurrence for the $n$-component parallel system by analogy with the 2-component parallel system. To do so, let us consider the distribution of the following order statistics. Suppose $X_1, X_2, \ldots, X_n$ be the components' lifetimes that initially follow the hazard rate function $h_1(t)$. Let $F_k(t)$ be the CDF of the time to the $k$-th failure occurrence $T_k$. If there is no dependent failure occurrence, $F_k(t)$ can be obtained as

$$F_k(t) = \sum_{i=k}^{n} \binom{n}{i} (1 - e^{-\lambda_1 t^{m_1}})^i e^{-\lambda_1 (n-i) t^{m_1}} \quad (k = 1, 2, \ldots, n), \tag{5.7}$$

by the theory of order statistics [42, 56]. That is, $F_k(t)$ can be estimated with the sample of $T_1$. Therefore, we can detect the dependent failure occurrence by performing K-S test between the sample of $T_k$ and the estimated $F_k(t)$ with the sample of $T_1$ for $k = 2, 3, \ldots, n$. Let $\{t_{k1}, t_{k2}, \ldots, t_{kN}\}$ be the $N$ samples of $T_k$. The algorithm of the detection method is shown below, and its schematic chart is given in Figure 5.4.

---

**Algorithm 2: $n$-component parallel system**

---

(i) Estimate $\widehat{F_1}(t) = 1 - e^{-n\widehat{\lambda}_1 t^{\widehat{m}_1}}$ by the method of maximal likelihood with the sample of $T_1$.

(ii) Set $k = 2$, and iterate the following steps.

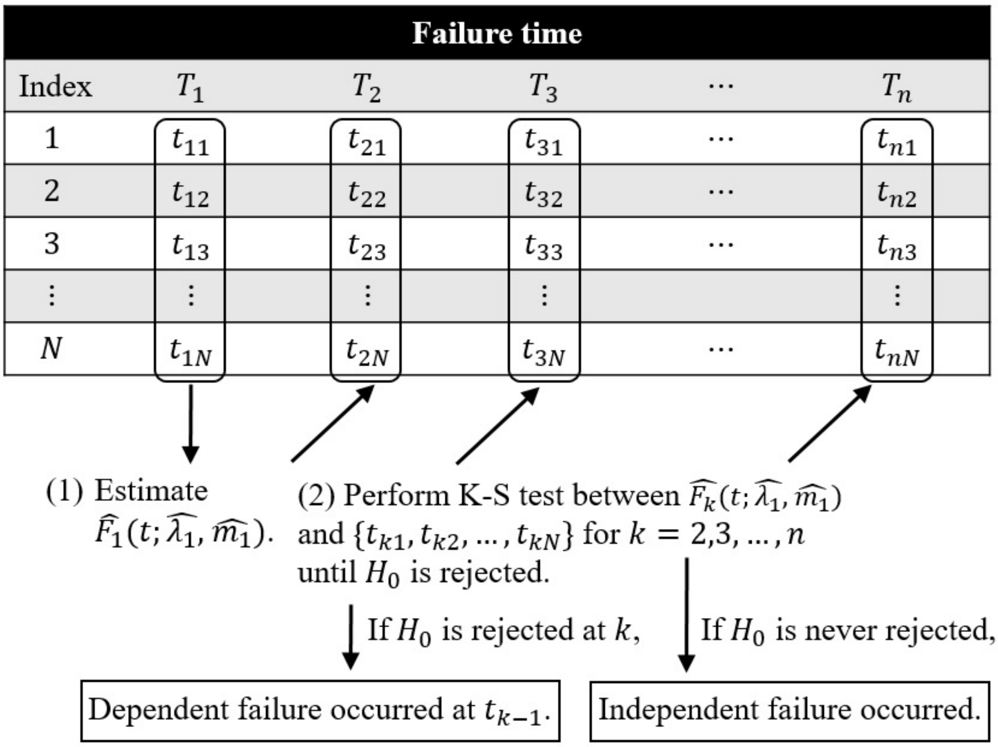


Figure 5.4: Schematic chart of the detection method.

(iii) Set $\widehat{F_k}(t) = \sum_{i=k}^{n} \binom{n}{i}(1 - e^{-\widehat{\lambda}_1 t^{\widehat{m}_1}})^i e^{-\widehat{\lambda}_1 (n-i) t^{\widehat{m}_1}}$ with the assumption that the independent failure occurs at $T_{k-1}$.

(iv) Perform K-S test with a significance level of 5% between $\widehat{F_k}(t)$ and the $k$-th sample $\{t_{k1}, t_{k2}, \ldots, t_{kN}\}$.

(v) If $H_0$ is accepted, let $k$ be $k+1$ and go back to (iii). If $H_0$ is rejected or $k = n+1$, go to the next step.

(vi) If $H_0$ is never rejected, the dependent failure never occurred. If $H_0$ is rejected, the dependent failure occurred at time $T_{k-1}$.

---

Throughout these straightforward steps, we can detect the time point of the first dependent failure occurrence with the failure data under the significance level of 5%. Note that the accuracy of the proposed method depends on $\lambda_i$ or $m_i$ ($i = 1, 2, \ldots, n$) as the case of the 2-component parallel system.

## 5.4 Numerical studies

In this section, we perform simulation studies to evaluate how the proposed method can accurately detect the dependent failure occurrence. After a demonstration of the detection scheme with a short example, we investigate the performance of the proposed method.

### 5.4.1 Short example

In this subsection, we explain how the proposed method detects the dependent failure occurrence by using one example. Suppose a RAID 1 system (Redundant Arrays of Inexpensive Disks: 1st level [51]) constructed by two hard disks as shown in Figure 5.5. This system consists of an exact copy of a set of data on the two disks. Although the disks are typically designed to work independently, the dependent failure may occur if they share heat (i.e., workload [4]). For this system, we try to evaluate whether or not the dependent failure occurs after one hard disk fails. Note that, we assume that these hard disks are used until both of them fail (i.e., we suppose that the failed disk are not repaired/replaced due to some unavoidable reasons).

To model the RAID 1 system, we deal with it as a 2-component parallel system. Several assumptions for the model are as follows. Let the values of the parameters be $(\lambda_1, \lambda_2, m_1, m_2) = (1, 3, 2, 2)$. In this case, the setting of the parameters expresses a dependent failure because $T_2$ sooner occurs than the originally expected one in the sense that $\lambda_2$ is $3 \times \lambda_1$. Thus, the correct evaluation is that a dependent failure occurs at $T_1$. We obtain $N$ $(= 100)$ samples of $(T_1, T_2)$ by Monte Carlo simulation [19] based on the failure model, and show the data set in Table 5.1.

Here, the estimation result is listed in Table 5.2. We obtain the estimates $\widehat{\lambda_1} = 0.846$ and $\widehat{m_1} = 1.85$ by the method of maximal likelihood with the sample of $T_1$. As the result of K-S test, the test statistics $\sqrt{N}D$ equals to 2.22, and its $p$-value equals to 0.0001. In this case, it is statistically natural to consider that the failure at $T_1$ caused the dependent failure rather than the independent failure of the small significance level. Accordingly, the null hypothesis $H_0$ is rejected with the significance level of 5%. That is, the proposed method identifies this failure case as the dependent failure case. Therefore, the dependent failure occurrence

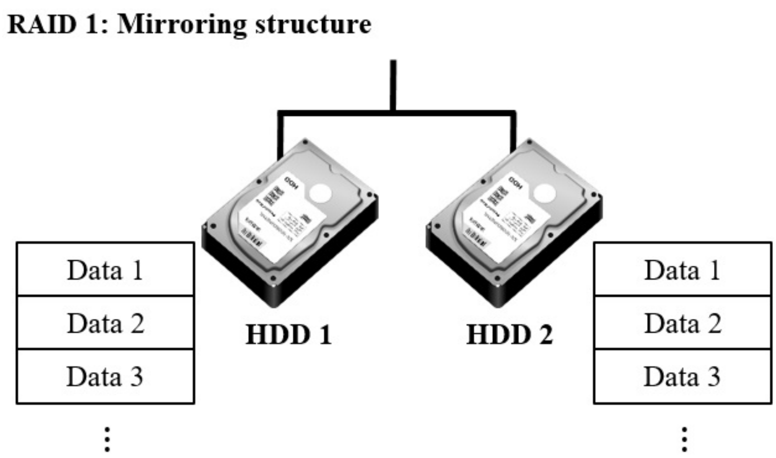


Figure 5.5: Example of a 2-component parallel system.

can be correctly detected in this example. In the next subsection, we investigate the various failure cases to evaluate the performance of the proposed method.

### 5.4.2 Performance evaluation

We analyze the 3-component parallel system as examples of the $n$-component parallel system. In this numerical studies, we assume that the dependent failure occurs at most once. Thus, in this case, the failure cases are divided into the following three types:

- Type 0: no dependent failure occurs.
- Type 1: dependent failure that occurs at $T_1$.
- Type 2: dependent failure that occurs at $T_2$.

For example, the failure case of $(\lambda_1, \lambda_2, \lambda_3, m_1, m_2, m_3) = (1, 1, 1, 2, 2, 2)$ is Type 0 because the hazard rate function does not change. The accuracy of the proposed method is defined as the probability of a correct judgment that specifies the true failure case of the system.

The simulation is performed as follows. First, the failure model described in Section 5.2 generates $N = 100$ samples of $(T_1, T_2, T_3)$ under the several failure case environments by Monte Carlo simulation [19]. Then, the detection method

Table 5.1: Simulation data ($t_{1i}$ and $t_{2i}$ are the pair of the failure times for $i = 1, 2, \ldots, 100$). Failure case: $(\lambda_1, \lambda_2, m_1, m_2) = (1, 3, 2, 2)$.

| |
|---|
| $t_{1i} = \{0.52, 1.04, 0.47, 0.59, 0.89, 0.30, 1.09, 0.56, 0.36, 0.67, 0.15, 0.21, 0.52, 1.43, 0.59, 0.52, 0.29, 0.85, 0.20, 0.41, 1.16, 1.18, 0.87, 0.69, 0.63, 1.13, 1.29, 0.24, 0.47, 0.29, 0.28, 0.28, 0.94, 0.53, 0.18, 0.79, 0.60, 0.58, 1.10, 0.16, 0.50, 0.54, 1.10, 0.76, 0.14, 0.98, 1.97, 0.33, 1.11, 1.15, 1.18, 0.29, 0.43, 0.70, 0.38, 1.71, 0.59, 0.98, 0.79, 0.85, 0.49, 0.31, 0.63, 0.06, 0.68, 0.40, 1.15, 0.87, 0.42, 0.12, 0.33, 0.73, 0.75, 0.32, 0.75, 1.09, 1.00, 0.62, 0.33, 0.83, 0.79, 0.75, 0.03, 1.09, 0.93, 0.50, 0.69, 0.65, 1.36, 0.67, 0.30, 0.23, 0.73, 0.35, 1.65, 0.51, 0.57, 0.57, 0.51, 0.56\}$ |
| $t_{2i} = \{0.92, 1.23, 0.54, 0.98, 1.04, 0.86, 1.29, 0.65, 0.95, 1.11, 0.17, 0.42, 0.99, 1.52, 1.00, 0.69, 0.51, 1.47, 1.04, 0.84, 1.39, 1.44, 1.38, 0.89, 1.12, 1.39, 1.47, 0.45, 0.54, 1.10, 1.38, 0.73, 1.89, 1.13, 0.50, 1.19, 0.84, 0.98, 1.50, 0.24, 0.81, 0.85, 1.15, 1.88, 0.72, 1.34, 2.00, 0.68, 1.14, 1.20, 1.20, 0.45, 0.74, 1.31, 0.81, 2.00, 1.22, 1.26, 1.65, 0.89, 0.51, 0.76, 0.92, 0.64, 0.68, 0.89, 1.45, 0.96, 0.97, 0.52, 1.16, 0.94, 1.28, 0.49, 1.16, 1.87, 1.06, 0.78, 0.93, 0.98, 1.66, 0.76, 0.72, 1.33, 1.44, 1.14, 1.48, 1.00, 1.51, 0.98, 1.11, 0.42, 1.26, 1.22, 2.48, 0.63, 0.82, 0.95, 0.98, 1.50\}$ |

Table 5.2: Estimation result.

| $\widehat{\lambda_1}$ | $\widehat{m_1}$ | $\sqrt{N}D$ | $p$-value |
|---|---|---|---|
| 0.846 | 1.85 | 2.22 | 0.0001 |

identifies the failure case as one of Type 0, 1 and 2 based on the data set. Finally, we iterate this evaluation 500 times and calculate the accuracy as follows

$$\text{Accuracy (\%)} = \frac{\text{\# of times of the correct judgment given}}{500} \times 100. \tag{5.8}$$

### 5.4.3 Accuracy for Type 1

We calculate the accuracy of the test for the following 20 failure cases of Type 1. In the first 10 cases, $\lambda_2$ $(= \lambda_3)$ is a variable that takes the values $1, 2, \ldots, 10$. The other parameters are fixed as $(\lambda_1, m_1, m_2, m_3) = (1, 2, 2, 2)$. Figure 5.6 shows these behaviors of the hazard rate functions. If $\lambda_2 = \lambda_3 = 1$, the failure case is an independent failure case because there is no change of the hazard rate function.

Table 5.3 shows the accuracy of the test in these 10 cases. As shown in the table, the detection method correctly judges with the probabilities of 43.0(%) and

83.0(%) if $\lambda_2 = \lambda_3 = 2$ and $\lambda_2 = \lambda_3 = 3$, respectively. This result demonstrates that the accuracy gets higher as $\lambda_2$ and $\lambda_3$ increase.

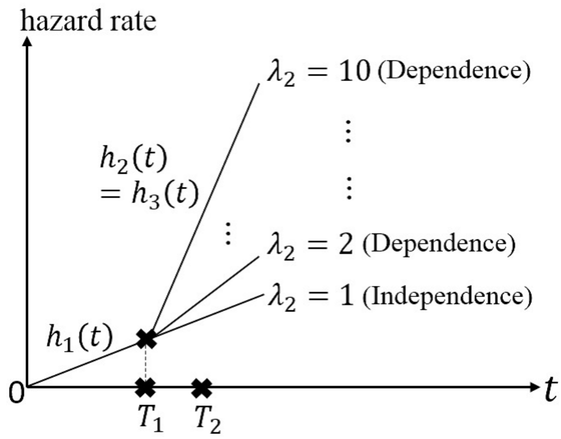


Figure 5.6: Behavior of the hazard rate functions (Type 1).

Table 5.3: Type 1 accuracy for the 3-component parallel system.

| | Accuracy (%) | | |
|---|---|---|---|
| $\lambda_2(=\lambda_3)$ | Type 0 | Type 1 | Type 2 |
| 1 | *65.8 | 9.6 | 24.6 |
| 2 | 16.0 | *43.0 | 41.0 |
| 3 | 1.8 | *83.0 | 15.2 |
| 4 | 0.0 | *97.4 | 2.6 |
| 5 | 0.0 | *99.6 | 0.4 |
| 6 | 0.0 | *100.0 | 0.0 |
| 7 | 0.0 | *100.0 | 0.0 |
| 8 | 0.0 | *100.0 | 0.0 |
| 9 | 0.0 | *100.0 | 0.0 |
| 10 | 0.0 | *100.0 | 0.0 |

1 Fixed parameters: $(\lambda_1, m_1, m_2, m_3) = (1, 2, 2, 2)$.

2 * symbol denotes the probability of a correct judgment.

Table 5.4: Type 1 accuracy for the 3-component parallel system.

| | Accuracy (%) | | |
|---|---|---|---|
| $m_2(= m_3)$ | Type 0 | Type 1 | Type 2 |
| 2 | *65.8 | 9.6 | 24.6 |
| 3 | 11.8 | *30.0 | 58.2 |
| 4 | 1.0 | *55.8 | 43.2 |
| 5 | 0.0 | *72.0 | 28.0 |
| 6 | 0.0 | *82.4 | 17.6 |
| 7 | 0.0 | *89.0 | 11.0 |
| 8 | 0.0 | *91.4 | 8.6 |
| 9 | 0.0 | *97.0 | 3.0 |
| 10 | 0.0 | *98.4 | 1.6 |

[1] Fixed parameters: $(\lambda_1, \lambda_2, \lambda_3, m_1) = (1, 1, 1, 2)$.

[2] * symbol denotes the probability of a correct judgment.

In the remaining 10 cases, $m_2$ $(= m_3)$ is a variable that takes the values $2, 3, \ldots, 10$. The other parameters are fixed as $(\lambda_1, \lambda_2, \lambda_3, m_1) = (1, 1, 1, 2)$. The accuracy shown in Table 5.4 suggests that $m_2$ also affects the accuracy.

### 5.4.4 Accuracy for Type 2

We also investigate how the proposed method can accurately detect the dependent failure that occurred at time $T_2$. Here, the hazard rate function does not change at $T_1$ but $T_2$. Hence, for the dependent failure cases of Type 2, we deal with a correct judgment as a judgment for the dependent failure occurrence at $T_2$ without any incorrect judgments (i.e., judgments that the dependent failure occurred at $T_1$ or that no dependent failure occurred).

First, we calculate the accuracy for Type 2 while $\lambda_3$ is a variable and other parameters are fixed as $(\lambda_1, \lambda_2, m_1, m_2, m_3) = (1, 1, 2, 2, 2)$. Figure 5.7 shows these behaviors of the hazard rate functions. The case of $\lambda_3 = 1$, of course, describes the independent failure. These results are listed in Table 5.5. For example, we can find that the detection method correctly judges with the probabilities of 37.4(%) and 87.8(%) if $\lambda_3 = 2$ and $\lambda_3 = 10$, respectively. That is, the accuracy for Type 2 is

Table 5.5: Type 2 accuracy for the 3-component parallel system.

| | Accuracy (%) | | |
|---|---|---|---|
| $\lambda_3$ | Type 0 | Type 1 | Type 2 |
| 1 | *65.8 | 9.6 | 24.6 |
| 2 | 52.8 | 9.8 | *37.4 |
| 3 | 39.4 | 8.4 | *52.2 |
| 4 | 21.0 | 8.8 | *70.2 |
| 5 | 15.0 | 9.8 | *75.2 |
| 6 | 10.8 | 8.0 | *81.2 |
| 7 | 6.8 | 12.6 | *80.6 |
| 8 | 4.4 | 9.2 | *86.4 |
| 9 | 2.8 | 8.4 | *88.8 |
| 10 | 2.4 | 9.8 | *87.8 |

[1] Fixed parameters: $(\lambda_1, \lambda_2, m_1, m_2, m_3) = (1, 1, 2, 2, 2)$.
[2] * symbol denotes the probability of a correct judgment.

getting better as $\lambda_3$ increases. Note that although the dependent failure occurred at $T_2$, the method gives an approximately 10% wrong judgment for Type 1 among the entire failure cases. The reason is the significance level of the test.

Next, we derive the accuracy for Type 2 while $m_3$ is a variable and other parameters are given by $(\lambda_1, \lambda_2, \lambda_3, m_1, m_2) = (1, 1, 1, 2, 2)$. Table 5.6 shows the accuracy in these failure cases. These results demonstrate that $m_3$ affects the test performance similarly to $\lambda_3$.

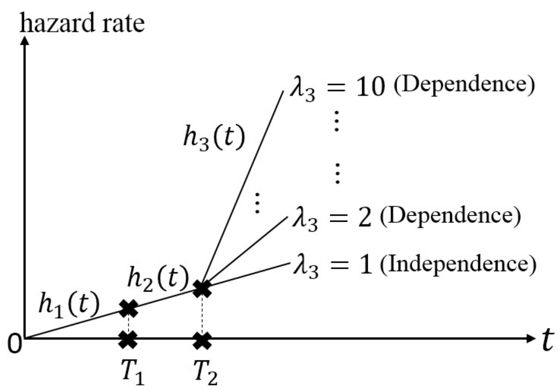


Figure 5.7: Behavior of the hazard rate functions (Type 2).

Table 5.6: Type 2 accuracy for the 3-component parallel system.

| | Accuracy (%) | | |
|---|---|---|---|
| $m_3$ | Type 0 | Type 1 | Type 2 |
| 2 | *65.8 | 9.6 | 24.6 |
| 3 | 40.0 | 7.6 | *52.4 |
| 4 | 9.2 | 11.0 | *79.8 |
| 5 | 3.0 | 9.0 | *88.0 |
| 6 | 0.4 | 8.4 | *91.2 |
| 7 | 0.2 | 9.0 | *90.8 |
| 8 | 0.0 | 11.2 | *88.8 |
| 9 | 0.0 | 9.4 | *90.6 |
| 10 | 0.0 | 8.6 | *91.4 |

[1] Fixed parameters: $(\lambda_1, \lambda_2, \lambda_3, m_1, m_2) = (1, 1, 1, 2, 2)$.
[2] * symbol denotes the probability of a correct judgment.

### 5.4.5 Discussion

We found that, for $i < j$, the proposed method can detect the dependent failure that occurred at $T_i$ with a higher probability than that occurred at $T_j$. For example, the probability of a correct judgment for Type 1 is higher than for Type 2. On the other hand, as $i$ increases, it is hard for the proposed method to precisely identify $T_i$ due to the significance level of the test. However, this is not a serious problem because of the following reasons. We are primarily interested in detecting the dependent failure that occurred at $T_i$ with a small number $i$ rather than a large number $j$. This is because the dependent failure that occurred at $T_i$ shortens a system's reliability more strongly than the case of $T_j$. Moreover, in actual cases, a large number of components $n$ is not applied because the effectiveness of redundancy declines as $n$ increases [5]. That is, $i$ is a small number in the actual cases. Therefore, the proposed method makes sense for use in practical cases.

## 5.5 Conclusion

In this chapter, we have developed a method of detecting the dependent failure occurrence for the $n$-component parallel system. This method has succeeded in detecting the dependent failure occurrence by testing whether or not the lifetimes of the components follow an expected Weibull distribution between $t_{k-1}$ and $t_k$ (for $k = 2, 3, \ldots, n$) with the sample of $T_1$. In other words, we can discriminate whether the components of the parallel system failed independently or dependently by obtaining the failure occurrence times of each component. It helps in making the objective judgments in order to identify the first dependent failure occurrence out of $n$ failures in the parallel system. In Subsection 5.4.1, we have illustrated how the proposed method can be applied to the RAID 1 system with two hard disks. In the same fashion, the proposed method can also detect the dependent failure occurrence in a RAID 1 system with $n$ hard disks. Therefore, we recommend using the statistical detection method of the dependent failure occurrence for quality control at the stage of system testing.

A limitation of this work is that the assumption for the hazard rate of the components. In Section 5.2, we assumed that the hazard rate function of the components is one of the following: DFR, CFR, or IFR (Assumption II). However, this is not always realistic in practical engineering. For example, the typical bathtub curve is a combination of these three [42]. In that case, the proposed method cannot work correctly because the switching of the hazard rate function is not the only evidence for the dependent failure occurrence. This problem can be ignored only if we use the proposed method for a particular period of the life-cycle of the item (i.e., one of the DFR, CFR and IFR periods). Thus, we need to pay attention to the appropriateness of the Assumption II when we apply the proposed method to real data.

For the future work, we would like to improve the detection method which can also find two or more dependent failure occurrences. Moreover, an application of this method to real data is required for its verification.

# Chapter 6

# Conclusion

## 6.1 Novelties and contributions of this thesis

This thesis has the following novelties and contributions in the research fields of reliability engineering and statistics.

- Necessary conditions of the ranges of the dependence parameters for the minimum and maximum value distributions based on the FGM copula are shown (Chapter 2).
- MTTFs of the parallel system and series one with dependent components are analytically found (Chapter 3).
- The result enables us to estimate dependencies among several components by using their lifetime data in an acceptable computation time (Chapter 4).
- The result enables us to detect the dependent failure occurrence in an $n$-component parallel system (Chapter 5).

## 6.2 Summary

In this thesis, we have considered reliability analysis techniques when several components in a multi-component system break down dependently. First, we have formulated the reliability model for the dependent failure by using the multivariate FGM copula. Based on the model, we have investigated the effect of the dependent failure occurrence on the system's reliability. Secondly, we have proposed the useful estimation algorithm for the multivariate FGM copula. In addition, we have theoretically revealed the asymptotic normality of the proposed estimators. Finally, we have presented the new method for the detection of the dependent failure occurrence in the $n$-component parallel system. These results are helpful to both quantitative and qualitative reliability assessment of the system under the possibility of the dependent failure occurrences. Our estimation method is applicable not only the reliability analysis but also other research fields in particular.

In Chapter 2, we have discussed mathematical preliminaries for our study and the new results of the FGM copula in order to use them as the modeling tool of the interdependent failure-occurrence environment in Chapters 3 and 4. Chapter 2 has been devoted to the definition and some fundamental properties of the copula. Then, several unique features of the FGM copula have been introduced. Finally, we have discovered the necessary conditions of the ranges of the dependence parameters for the minimum and maximum value distributions which are based on the FGM copula and presented the asymptotic properties of the ranges.

In Chapter 3, we have developed the reliability models of a parallel system and series system with $n$ dependent components by using the multivariate FGM copula. Then, we have investigated the several reliability-related properties of the systems. As a result, we have analytically shown that the $n$-component parallel system cannot deliver its designed reliability if the lifetimes of the individual components have positive dependence. On the other hand, we have derived that the $n$-component series system can exceed its designed reliability under such dependent failure-occurrence environment.

In Chapter 4, we have presented the estimation algorithm for the model parameter and referred to its asymptotic normality. More specifically, we have dealt with the parameter estimation for the $d$-variate FGM copula which consists of $2^d - d - 1$ dependence parameters to be estimated. We have proposed the new estimation method for the FGM copula by using the theory of the inference function for margins. Although the ordinary maximum likelihood estimation is computationally infeasible for a large number $d$, our method is feasible for the same situation.

Then, we have presented its asymptotic property. Finally, we have demonstrated the performance of the proposed method through the simulation studies.

In Chapter 5, we have studied the detection method of dependent failure occurrence in the $n$-component parallel system. Making a system redundant by combining identical components is a useful way to ensure a highly reliable system. However, the components of such systems may fail mutually, and if the components break down dependently, the reliability of the system decreases. Therefore, reliability analysis considering the dependence among the components is important in reliability assessment. In this chapter, we have proposed the statistical detection method of the dependent failure occurrence in $n$-component parallel systems by using the failure occurrence times of the components. If we assume that the lifetime distribution of the components worsens if $k$ out of $n$ components failed, the dependent failure occurrence can be found by identifying the change of the distribution. Finally, the performance of the proposed method has been demonstrated by the simulation studies.

# Appendix A

# Proofs

In this appendix, we present the mathematical proofs of principal theorems and corollaries derived in Chapter 2 and Chapter 4.

**Proof of Theorem 2.2:** We need to derive the closed form of the limitation $\theta_{1:d} \in \{\theta \mid \frac{d}{du}C_{1:d}(u;\theta) \geq 0, 0 \leq u \leq 1\}$. First we define the probability density function (PDF) of $U_{1:d}$ by the following $c_{1:d}(u;\theta_{1:d})$.

$$\begin{aligned} c_{1:d}(u;\theta_{1:d}) &= \frac{d}{du}C_{1:d}(u;\theta_{1:d}) \\ &= d(1-u)^{d-1}\left\{1+\theta_{1:d}(-2+2(1-u)^d+(1+d)u)\right\}. \end{aligned} \tag{A.1}$$

Here, $\theta_{1:d}$ has the following relationship.

$$\begin{aligned} &\theta_{1:d} \in \{\theta \mid c_{1:d}(u;\theta) \geq 0, 0 \leq u \leq 1\} \\ &\Leftrightarrow \theta_{1:d} \in \{\theta \mid \min_{0\leq u\leq 1}[c_{1:d}(u;\theta)] \geq 0\}. \end{aligned} \tag{A.2}$$

That is, the problem is equivalent to finding $\theta_{1:d}$ such that the minimum value of $c_{1:d}(u;\theta_{1:d})$ is non-negative. Since $d(1-u)^{d-1} \geq 0$ for $0 \leq u \leq 1$ and $d \geq 2$, we have

$$\theta_{1:d} \in \{\theta \mid \min_{0\leq u\leq 1}[\tilde{c}_{1:d}(u;\theta)] \geq 0\}, \tag{A.3}$$

where

$$\tilde{c}_{1:d}(u;\theta_{1:d}) = \frac{1}{d(1-u)^{d-1}} c_{1:d}(u;\theta_{1:d}). \tag{A.4}$$

Hence, we can verify the **Theorem 2.2** by solving the minimization problem of Eq. (A.3). In order to solve this problem, we define the first and second derivatives with respect to $u$ of $\tilde{c}_{1:d}(u;\theta_{1:d})$ as follows.

$$\begin{aligned} \tilde{c}'_{1:d}(u;\theta_{1:d}) &\stackrel{\text{def}}{=} \frac{d}{du}\tilde{c}_{1:d}(u;\theta_{1:d}) \\ &= \theta_{1:d}(-2d(1-u)^{d-1}+1+d), &(\text{A.5}) \\ \tilde{c}''_{1:d}(u;\theta_{1:d}) &\stackrel{\text{def}}{=} \frac{d^2}{du^2}\tilde{c}_{1:d}(u;\theta_{1:d}) \\ &= \theta_{1:d}(d-1)d(1-u)^{d-2}. &(\text{A.6}) \end{aligned}$$

Let $u_d^*$ be an critical point of $\tilde{c}_{1:d}$ (i.e., $\tilde{c}'_{1:d}(u_d^*;\theta_{1:d}) = 0$). Then, $u_d^*$ uniquely exists on the interval $[0,1]$, and it is easy to see that $u_d^* = 1 - \left(\frac{1+d}{2d}\right)^{\frac{1}{d-1}}$.

Consider $\theta_{1:d} \geq 0$. In this case, for $0 \leq u \leq 1$, $\tilde{c}_{1:d}(u;\theta_{1:d})$ is a convex function because $\tilde{c}''_{1:d}(u;\theta_{1:d}) \geq 0$. Thus, $\tilde{c}_{1:d}(u_d^*;\theta_{1:d})$ is the absolute minimum value. Hence, we have $\theta_{1:d} \in \{\theta \mid \tilde{c}_{1:d}(u_d^*;\theta) \geq 0\}$. This yields

$$\theta_{1:d} \leq \frac{1}{2-2(1-u_d^*)^d-(1+d)u_d^*}, \tag{A.7}$$

where $1/(2-2(1-u_d^*)^d-(1+d)u_d^*)$ gives the upper bound of $\theta_{1:d}$.

Consider $\theta_{1:d} < 0$. In this case, $\tilde{c}_{1:d}(u;\theta_{1:d})$ is a concave function that takes the minimum value if and only if $u = 1$. Thus, we have $\theta_{1:d} \in \{\theta \mid \tilde{c}_{1:d}(1;\theta) \geq 0\}$. This implies

$$-\frac{1}{d-1} \leq \theta_{1:d}, \tag{A.8}$$

where $-1/(d-1)$ gives the lower bound of $\theta_{1:d}$. Hence, the proof is complete. □

**Proof of Corollary 2.1:** By **Theorem 2.2**, the upper bound of $\theta_{1:d}$ is given by $1/(2-2(1-u_d^*)^d-(1+d)u_d^*)$. Thus as $d \to \infty$, the upper bound of $\theta_{1:d}$ can be written by

$$\lim_{d\to\infty} \frac{1}{2-2(\frac{1+d}{2d})^{\frac{d}{d-1}}-(1+d)(1-\frac{1+d}{2d})^{\frac{1}{d-1}}}. \tag{A.9}$$

Note that $u_d^*$ is replaced by Eq. (2.22). By considering its Taylor series, Eq. (A.9) equals to

$$\lim_{d\to\infty} \frac{1}{2-(1+O(\frac{1}{d}))-(\log 2+O(\frac{1}{d}))} = \frac{1}{1-\log 2}, \tag{A.10}$$

where $O(\cdot)$ denotes Landau's symbol.

Moreover, for the lower bound, we have

$$\lim_{d\to\infty} -\frac{1}{d-1} = 0. \tag{A.11}$$

Hence, the proof is complete. □

We would like to omit the proofs of **Theorem 2.3** and **Corollary 2.2** because they can be shown in the same way as those of **Theorem 2.2** and **Corollary 2.1**.

# Proofs in Chapter 4

Before we prove **Theorem 4.1** and **Theorem 4.2**, we introduce an important lemma and theorem as follows.

**Lemma A.1.** *Let* $(X_1, X_2, X_3)$ *be a random vector of the trivariate FGM copula. Then,* $E\left[\boldsymbol{s}(X_1, X_2, X_3; \boldsymbol{\eta_0})\right] = \mathbf{0}$ *and* $E[\frac{\partial}{\partial \boldsymbol{\eta}} \log h_3(X_1, X_2, X_3; \cdot) = \mathbf{0}$.

**Proof of Lemma A.1:** With Eq. (4.24), we prove that each element of $E\left[\boldsymbol{s}(\cdot; \boldsymbol{\eta_0})\right]$ is 0, that is,

$$\begin{aligned} &E\left[\frac{\partial}{\partial \delta_j} \log f_j(X_j; \cdot)\right] \\ &= E\left[\frac{\partial}{\partial \theta_{jk}} \log h_2(X_j, X_k; \cdot)\right] \\ &= E\left[\frac{\partial}{\partial \theta_{123}} \log h_3(X_1, X_2, X_3; \cdot)\right] = 0, \end{aligned} \tag{A.12}$$

for $1 \leq j < k \leq 3$. First, we have

$$
\begin{aligned}
E\left[\frac{\partial}{\partial \delta_j}\log f_j(X_j;\cdot)\right] &= \int_{-\infty}^{\infty} f_j(x_j;\cdot)\frac{\partial}{\partial \delta_j}\log f_j(x_j;\cdot)dx_j \\
&= \int_{-\infty}^{\infty} \frac{\partial}{\partial \delta_j} f_j(x_j;\cdot)dx_j \\
&= \frac{\partial}{\partial \delta_j}\int_{-\infty}^{\infty} f_j(x_j;\cdot)dx_j = 0, \qquad \text{(A.13)}
\end{aligned}
$$

where we change the order of the differentiation and integration in the third step. In the same fashion,

$$
\begin{aligned}
E\left[\frac{\partial}{\partial \theta_{jk}}\log h_2(X_j, X_k;\cdot)\right] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_{jk}(x_j, x_k;\cdot)\frac{\partial}{\partial \theta_{jk}}\log h_2(x_j, x_k;\cdot)dx_j dx_k \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta_{jk}} h_2(x_j, x_k;\cdot)dx_j dx_k \\
&= \frac{\partial}{\partial \theta_{jk}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_2(x_j, x_k;\cdot)dx_j dx_k = 0, \qquad \text{(A.14)}
\end{aligned}
$$

and

$$
\begin{aligned}
&E\left[\frac{\partial}{\partial \theta_{123}}\log h_3(X_1, X_2, X_3;\cdot)\right] \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(x_1, x_2, x_3;\cdot)\frac{\partial}{\partial \theta_{123}}\log h_3(x_1, x_2, x_3;\cdot)dx_1 dx_2 dx_3 \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta_{123}} h_3(x_1, x_2, x_3;\cdot)dx_1 dx_2 dx_3 \\
&= \frac{\partial}{\partial \theta_{123}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_3(x_1, x_2, x_3;\cdot)dx_1 dx_2 dx_3 = 0. \qquad \text{(A.15)}
\end{aligned}
$$

Therefore, $E\left[\boldsymbol{s}(\cdot;\boldsymbol{\eta_0})\right] = \mathbf{0}$. Also, $E[\frac{\partial}{\partial \boldsymbol{\eta}}\log h_3(X_1, X_2, X_3;\cdot) = \mathbf{0}$ holds in the same way. Hence, the proof is complete. □

**Definition A.1.** *Let $\{Y_n\}$ be a sequence of random variables. Let $c$ be a real number. We say that $\{Y_n\}$ is convergent in probability to $c$, shown by $Y_n \xrightarrow{p} c$, if*

$$
\lim_{n\to\infty} \Pr[|Y_n - c| < \varepsilon] = 1, \qquad \text{(A.16)}
$$

*for any $\varepsilon > 0$.*

**Definition A.2.** *Let $\{Y_n\}$ be a sequence of random variables, and $\{F_n\}$ be their CDFs. Let $Y$ be a random variable, and $F$ be the CDF of $Y$. We say that $\{Y_n\}$ is convergent in distribution to $Y$, shown by $Y_n \overset{\mathcal{D}}{\to} Y$, if*

$$\lim_{n\to\infty} F_n(y) = F(y). \tag{A.17}$$

**Theorem A.1.** *(Slutsky's theorem). This theorem is about the convergence in probability and distribution (see e.g., [7], p. 254). Let $\{Y_n\}, \{Z_n\}$ be sequences of random variables. Let $Z$ be a random variable. If $Y_n \overset{p}{\to} c \in \mathbb{R}$ and $Z_n \overset{\mathcal{D}}{\to} Z$, then $Y_nZ_n \overset{\mathcal{D}}{\to} cZ$.*

**Proof of Theorem 4.1:** We provide the sketch of the proof of **Theorem 4.1**. Suppose $\boldsymbol{Y}, \boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$ be i.i.d. random vectors of the $d$-variate FGM copula. Applying the mean value theorem for the maximum log likelihood function at $\boldsymbol{\eta_0}$, we have

$$\mathbf{0} = \left.\frac{\partial \ell(\boldsymbol{\eta}; \mathbf{Y_1}, \ldots, \mathbf{Y_n})}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}=\widehat{\boldsymbol{\eta}}} \tag{A.18}$$

$$= \left.\frac{\partial \ell(\boldsymbol{\eta}; \mathbf{Y_1}, \ldots, \mathbf{Y_n})}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}} + K(\overline{\boldsymbol{\eta}})(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta_0}), \tag{A.19}$$

where $K(\overline{\boldsymbol{\eta}})$ is a matrix as

$$K(\overline{\boldsymbol{\eta}}) = \left.\frac{\partial^2 \ell(\boldsymbol{\eta}; \mathbf{Y_1}, \ldots, \mathbf{Y_n})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}^T}\right|_{\boldsymbol{\eta}=\overline{\boldsymbol{\eta}}}, \tag{A.20}$$

and $\overline{\boldsymbol{\eta}} \equiv diag(\alpha_1, \ldots, \alpha_7)\widehat{\boldsymbol{\eta}} + diag(1-\alpha_1, \ldots, 1-\alpha_7)\boldsymbol{\eta_0}$ for $(\alpha_1, \ldots, \alpha_7) \in [0,1]^7$ (N.B., $diag(a_1, \ldots, a_7)$ is a diagonal matrix whose elements starting in the upper left corner are $a_1, \ldots, a_7$). From Eq. (A.19), we have

$$\sqrt{n}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta_0}) = -\left[\frac{1}{n}K(\overline{\boldsymbol{\eta}})\right]^{-1} \frac{1}{\sqrt{n}} \left.\frac{\partial \ell(\boldsymbol{\eta}; \mathbf{Y_1}, \ldots, \mathbf{Y_n})}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}} \tag{A.21}$$

$$= -\left[\frac{1}{n}K(\overline{\boldsymbol{\eta}})\right]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left.\frac{\partial}{\partial \boldsymbol{\eta}} \log h_3(\mathbf{Y_i}; \boldsymbol{\eta})\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}} \tag{A.22}$$

Here, the first factor of the right-hand side of Eq. (A.22) can be applied to the low of large numbers because it is divided by the sample size $n$. Since $\widetilde{\boldsymbol{\eta}} \xrightarrow{p} \boldsymbol{\eta_0}$ for $n \to \infty$, $\overline{\boldsymbol{\eta}} \xrightarrow{p} \boldsymbol{\eta_0}$ also holds for $n \to \infty$ . Recalling Eq. (4.19), we obtain

$$\frac{1}{n}K(\overline{\boldsymbol{\eta}}) \xrightarrow{p} -E\left[\left.\frac{\partial^2 \log h_3(\boldsymbol{Y};\cdot)}{\partial\boldsymbol{\eta}\partial\boldsymbol{\eta}^T}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}}\right] = I. \tag{A.23}$$

The second factor can be applied to the multivariate central limit theorem because it is divided by $\sqrt{n}$. From **Lemma A.1** and Eq. (A.23), We obtain

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{\partial}{\partial\boldsymbol{\eta}}\log h_3(\boldsymbol{Y_i};\boldsymbol{\eta})\bigg|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}} \xrightarrow{\mathcal{D}} N(\mathbf{0}, I). \tag{A.24}$$

Therefore, by applying **Theorem A.1** to Eqs. (A.23) and (A.24), the following equation holds.

$$\sqrt{n}(\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{\mathcal{D}} N(\mathbf{0}, I^{-1}). \tag{A.25}$$

Hence, the proof is complete. □

**Proof of Theorem 4.2:** By the same analogy with the proof of Theorem 4.1, we provide the sketch of the proof of **Theorem 4.2**. Suppose $\boldsymbol{Y}, \boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}$ be i.i.d. random vectors of the $d$-variate FGM copula. Let us define $\boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta})$ as follows.

$$\boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta}) = (\frac{\partial\ell_1}{\partial\delta_1}, \frac{\partial\ell_2}{\partial\delta_2}, \frac{\partial\ell_3}{\partial\delta_3}, \frac{\partial\ell_{12}}{\partial\theta_{12}}, \frac{\partial\ell_{13}}{\partial\theta_{13}}, \frac{\partial\ell_{23}}{\partial\theta_{23}}, \frac{\partial\ell}{\partial\theta_{123}})^T. \tag{A.26}$$

Note that $\boldsymbol{g}(\cdot;\widetilde{\boldsymbol{\eta}}) = \mathbf{0}$ holds because $\widetilde{\boldsymbol{\eta}}$ is defined as the solution of $\boldsymbol{g}(\cdot;\boldsymbol{\eta}) = \mathbf{0}$. By using the mean value theorem for $\boldsymbol{g}(\cdot;\widetilde{\boldsymbol{\eta}})$ at $\boldsymbol{\eta_0}$, we have

$$\mathbf{0} = \boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \widetilde{\boldsymbol{\eta}}) = \boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta_0}) + J(\overline{\boldsymbol{\eta}})(\widetilde{\boldsymbol{\eta}} - \boldsymbol{\eta_0}), \tag{A.27}$$

where

$$J(\overline{\boldsymbol{\eta}}) = \frac{\partial\boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta})}{\partial\boldsymbol{\eta}}\bigg|_{\boldsymbol{\eta}=\overline{\boldsymbol{\eta}}}, \tag{A.28}$$

and $\overline{\boldsymbol{\eta}} \equiv diag(\alpha_1, \ldots, \alpha_7)\widetilde{\boldsymbol{\eta}} + diag(1-\alpha_1, \ldots, 1-\alpha_7)\boldsymbol{\eta_0}$ for $(\alpha_1, \ldots, \alpha_7) \in [0, 1]^7$.

From Eq. (A.27), we have $J(\overline{\boldsymbol{\eta}})(\widetilde{\boldsymbol{\eta}} - \boldsymbol{\eta_0}) = -\boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta_0})$. Thus,

$$\sqrt{n}(\widetilde{\boldsymbol{\eta}} - \boldsymbol{\eta_0}) = -\left[\frac{1}{n}J(\overline{\boldsymbol{\eta}})\right]^{-1} \frac{1}{\sqrt{n}}\boldsymbol{g}(\cdot; \boldsymbol{\eta_0}). \tag{A.29}$$

Here, the first factor of the right-hand side of Eq. (A.29) can be applied to the low of large numbers because it is divided by the sample size $n$. Since $\widetilde{\boldsymbol{\eta}} \xrightarrow{p} \boldsymbol{\eta_0}$ for $n \to \infty$, $\overline{\boldsymbol{\eta}} \xrightarrow{p} \boldsymbol{\eta_0}$ also holds for $n \to \infty$ . Recalling Eq. (4.22), we obtain

$$\frac{1}{n}J(\overline{\boldsymbol{\eta}}) \xrightarrow{p} E\left[\left.\frac{\partial \boldsymbol{s}(\boldsymbol{Y}; \boldsymbol{\eta})}{\partial \boldsymbol{\eta}}\right|_{\boldsymbol{\eta}=\boldsymbol{\eta_0}}\right] = D. \tag{A.30}$$

The second factor can be applied to the multivariate central limit theorem because it is divided by $\sqrt{n}$. From **Lemma A.1** and Eq. (4.23), we obtain

$$\frac{1}{\sqrt{n}}\boldsymbol{g}(\boldsymbol{Y_1}, \ldots, \boldsymbol{Y_n}; \boldsymbol{\eta_0}) \xrightarrow{\mathcal{D}} N(\boldsymbol{0}, M). \tag{A.31}$$

Therefore, by applying **Theorem A.1** to Eqs. (A.30) and (A.31), the following equation holds.

$$\sqrt{n}(\widetilde{\boldsymbol{\eta}} - \boldsymbol{\eta}_0) \xrightarrow{\mathcal{D}} N(\boldsymbol{0}, D^{-1}M(D^{-1})^T). \tag{A.32}$$

Hence, the proof is complete. □

# References


[1] J. Bialek, et al, "Benchmarking and validation of cascading failure analysis tools", *IEEE Transactions on Power Systems*, Vol. 31, Issue 6, pp. 4887-4900, 2016.

[2] S. Cambanis, "Some Properties and Generalization of Multivariate Eyraud-Gumbel-Morgenstern Distributions", *Journal of Multivariate Analysis*, Vol. 7, Issue 4, pp. 551-559, 1977.

[3] R. R. Carhart, "A survey of the current status of the electronic reliability problem", *Rand Corporation Research Memorandum*, RM-1131, 1953.

[4] J. A. Chandy and P. Banerjee. "Reliability evaluation of disk array architectures", *Proceedings of the International Conference on Parallel Processing*, Vol. 1, pp. 263-267, 1993.

[5] E. Chiodo and D. Lauria, "Some basic properties of the failure rate of redundant reliability systems in industrial electronics applications", *IEEE Transactions on Industrial Electronics*, Vol. 62, Issue 8, pp. 5055-5062, 2015.

[6] H. Cossette, M. P. Cote, E. Marceau and K. Moutanabbir, "Multivariate distribution defined with Farlie-Gumbel-Morgenstern copula and mixed Erlang marginals: Aggregation and capital allocation", *Insurance: Mathematics and Economics*, Vol. 52, Issue 3, pp. 560-572, May 2013.

[7] H. Cramer, *Mathematical methods of statistics*, Princeton University Press, Princeton, 1945.

[8] C. J. Creveling, "Increasing the reliability of electronic equipment by the use of redundant circuits", *Proceedings of the IRE*, Vol. 44, Issue 4, pp. 509-515, 1956.

[9] I. Dobson, B. A. Carreras and D.E. Newman, "A loading–dependent model of probabilistic cascading failure", *Probability in the Engineering and Informational Sciences*, Vol. 19, Issue 1, pp. 15-32, 2005.

[10] I. Dobson, "Where is the edge for cascading failure?: challenges and opportunities for quantifying blackout risk", *IEEE Power Engineering Society General Meeting 2007*, pp. 1-8, 2007.

[11] S. Eryilmaz and F. Tank, "On reliability analysis of a two-dependent-unit series system with a standby unit", *Applied Mathematics and Computation*, Vol. 218, Issue 15, pp. 7792-7797, 2012.

[12] D. J. Farlie, "The performance of some correlation coefficients for a general bivariate distribution", *Biometrika*, Vol. 47, Issues 3 and 4, pp. 307-323, 1960.

[13] W. Feller, "On the Kolmogorov-Smirnov limit theorem for empirical distribution", *The Annals of Mathematical Statistics*, Vol. 19, Issue 2, pp. 177-189, 1948.

[14] B. Fornberg, "Generation of finite difference formulas on arbitrarily spaced grids", *Journal of Mathematics of Computation*, Vol. 51, No. 184, pp. 699-706, 1988.

[15] A. Z. Górski and J. Szmigielski, "On pairs of difference operators satisfying: [D,X]=Id", *Journal of Mathematical Physics*, Vol. 39, pp. 545-568, 1998.

[16] E. J. Gumbel, "Bivariate exponential distributions", *Journal of the American Statistical Association*, Vol. 55, No. 292, pp. 698-707, 1960.

[17] E. Hashorva, "Asymptotic results for FGM random sequences", *Statistics & Probability Letters*, Vol. 54, Issue 4, pp. 417-425, 2001.

[18] I. Hernandez-Fajardo and L. Dueñas-Osorio, "Probabilistic study of cascading failures in complex interdependent lifeline systems", *Reliability Engineering and System Safety*, Vol. 111, pp. 260-272, 2013.

[19] W. Hörmann, J. Leydold and G. Derflinger, *Automatic nonuniform random variate generation*, Springer-Verlag, Berlin, 2004.

[20] P. Humphreys and A. M. Jenkins, "Dependent failures developments", *Reliability Engineering and System Safety*, Vol. 34, Issue 3, pp. 417-427, 1991.

[21] A. A. Ismaeel and R. Bhatnagar, "Test for detection and location of intermittent faults in combinational circuits", *IEEE Transactions on Reliability*, Vol. 46, Issue 2, pp. 269-274, 1997.

[22] P. Jaworski, F. Durante, W. Härdle and T. Rychlik (eds.), *Copula Theory and Its Applications*, Springer-Verlag, Berlin, 2010.

[23] H. Joe and J. J. Xu, "The estimation method of inference functions for margins for multivariate models", *Technical Report, Department of Statistics*, University of British Columbia, Vol. 166, pp. 1-21, 1996.

[24] H. Joe, "Asymptotic efficiency of the two-stage estimation method for copula-based models", *J. Multivariate Anal.*, Vol. 94, Issue 2, pp. 401-419, 2005.

[25] H. Joe, *Dependence Modeling with Copulas*, Chapman & Hall, London, 2014.

[26] N. L. Johnson and S. Kotz, "On some generalized Farlie-Gumbel-Morgenstern distributions", *Communications in Statistics - Theory and Methods*, Vol. 4, Issue 5, pp. 415-427, 1975.

[27] M. Kimura, S. Ota and S. Abe, "Some reliability properties of $n$-component parallel/series systems with interdependent failures", in *Reliability Modeling with Computer and Maintenance Applications*, eds. S. Nakamura, C.H. Qian and T. Nakagawa, World Scientific, Singapore, pp. 155-173, 2017.

[28] W. Long, Y. Sato and M. Horigome, "Quantification of sequential failure logic for fault tree analysis", *Reliability Engineering and System Safety*, Vol. 67, Issue 3, pp. 269-274, 2000.

[29] D. D. Mari and S. Kotz, *Correlation and Dependence*, Imperial College Press, London, 2001.

[30] A. W. Marshall and I. Olkin, “A multivariate exponential distribution”, *Journal of the American Statistical Association*, Vol. 62, No. 317, pp. 30-44, 1967.

[31] J. I. McCool, “Testing for dependency of failure times in life testing”, *Technometrics*, Vol. 48, No. 1, pp. 41-48, 2006.

[32] W. Q. Meeker and L. A. Escobar, *Statistical methods for reliability data*, John Wiley & Sons, Toronto, 1998.

[33] M. Modarres, K. P. Kaminskiy and V. Krivtsov, *Reliability engineering and risk analysis*, CRC Press, Boca Raton, 2017.

[34] E. F. Moore and C. E. Shannon, “Reliable circuits using less reliable relays”, *Journal of the Franklin Institute*, Vol. 262, Issue 3, pp. 191-208, 1956.

[35] D. Morgenstern, “Einfache Beispiele zweidimensionaler Verteilungen”, *Mitteislingsblatt für Mathematische Statistik*, Vol. 8, pp. 234-235, 1956.

[36] A. Mosleh, “Dependent failure analysis”, *Reliability Engineering and System Safety*, Vol. 34, Issue 3, pp. 243-248, 1991.

[37] G. S. Mudholkar, D. K. Srivastava and G. D. Kollia, “A generalization of the Weibull distribution with application to the analysis of survival data”, *Journal of the American Statistical Association*, Vol. 91, No. 436, pp. 1575-1583, 1996.

[38] T. Nakagawa and W. Y. Yun, “Note on MTTF of a parallel system”, *International Journal of Reliability, Quality and Safety Engineering*, Vol. 18, Issue 1, pp. 43-50, 2011.

[39] J. Navarro, J. M. Ruiz and C. J. Sandoval, “Properties of coherent systems with dependent components”, *Communications in Statistics - Theory and Methods*, Vol. 36, Issue 1, pp. 175-191, 2007.

[40] R. B. Nelsen, *An Introduction to Copulas*, 2nd edition, Springer-Verlag, New York, 2010.

[41] A. O'Connor and A. Mosleh, "A general cause based methodology for analysis of common cause and dependent failures in system risk and reliability assessments", *Reliability Engineering and System Safety*, Vol. 145, pp. 341-350, 2016.

[42] P. D. T. O'Connor, D. W. Newton and R. C. Bromley, *Practical Reliability Engineering*, 4th edition, Jon Wiley & Sons, Chichester, 2002.

[43] S. Ota, T. Kageyama and M. Kimura, "Improvement of reliability evaluation for 2-unit parallel system with cascading failures by using maximal copula", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E98-A, No. 10, pp. 2096-2100, 2015.

[44] S. Ota and M. Kimura, "A method of dependent failure–occurrence detection for parallel systems", *Proceedings of the 9th IMA International Conference on Modelling in Industrial Maintenance and Reliability*, pp. 174-178, 2016.

[45] S. Ota and M. Kimura, "A statistical dependent failure detection method for n-component parallel systems", *Reliability Engineering and System Safety*, Vol. 167, pp. 376-382, 2017.

[46] S. Ota and M. Kimura, "A study on reliability deterioration and improvement of coherent systems under dependent failure-occurrence environment", *Proceedings of the 10th International Conference on Mathematical Methods in Reliability*, 8 pages (USB), 2017.

[47] S. Ota and M. Kimura, "Parameter estimation for an n-variate FGM copula and its application to dependent failure analysis", *Proceedings of the 24th ISSAT International Conference on Reliability & Quality in Design*, pp. 77-81, Toronto, Canada, August, 2018.

[48] S. Ota and M. Kimura, "A note on statistical inference of dependence parameters of an n-variate FGM copula", *Proceedings of the 8th Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling*, pp. 367-372, Qingdao, China, August, 2018.

[49] S. Ota and M. Kimura, "On necessary conditions for dependence parameters of minimum and maximum value distributions based on n-variate FGM copula", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E102-A, No. 3, pp. 586-589, 2019.

[50] A. K. Pathak and P. Vellaisamy, "A note on generalized Farlie-Gumbel-Morgenstern copulas", *Journal of Statistical Theory and Practice*, Vol. 10, No. 1, pp. 40-58, 2016.

[51] D. A. Patterson, G. Gibson and R. H. Katz, "A case for redundant arrays of inexpensive disks (RAID)", *Proceedings of the 1988 ACM SIGMOD International Conference on Management of Data*, pp. 109-116, 1988.

[52] A. J. Patton, "Estimation of multivariate models for time series of possibly different lengths", *Journal of Applied Econometrics*, Vol. 21, pp. 147-173, 2006.

[53] C. M. S. Rocco and E. Zio, "A support vector machine integrated system for the classification of operation anomalies in nuclear components and systems", *Reliability Engineering and System Safety*, Vol. 92, Issue 5, pp. 593-600, 2007.

[54] J. J. Romero, "Blackouts illuminate India's power problems", *IEEE Spectrum*, Vol. 49, Issue 10, pp. 11-12, 2012.

[55] Y. Sato, "Reliability in the Apollo program, a balanced approach behind the success", *Quest: the History of Spaceflight Quarterly*, Vol. 13, No. 1, pp. 22-29, 2006. (This article can also be accessed from *https://kscddms.ksc.nasa.gov/Reliability/Documents/Reliability_in_the_Apollo_Program.pdf*).

[56] M. L. Shooman, *Reliability of computer systems and networks: fault tolerance, analysis, and design*, John Wiley & Sons, Toronto, 2002.

[57] S. H. Teng and S. Y. Ho, "Failure mode and effects analysis: An integrated approach for product design and process control", *International Journal of Quality and Reliability Management*, Vol. 13, Issue 5, pp. 8-26, 1996.

[58] M. Thottan and C. Ji, "Anomaly detection in IP networks", *IEEE Transactions on Signal Processing*, Vol. 51, Issue 8, pp. 2191-2204, 2003.

[59] H. L. V. Trees, *Detection, estimation, and modulation theory Part I*, Jon Wiley & Sons, New York, 1968.

[60] D. J. Watts, "A simple model of global cascades on random networks", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 99, No. 9, pp. 5766-5771, 2002.

[61] J. J. Xu, "Statistical modeling and inference for multivariate and longitudinal discrete response data", *Ph.D. thesis, Department of Statistics, University of British Columbia*, 1996.

[62] M. Zelen ed., *Statistical Theory of Reliability*, University of Wisconsin Press, Madison, 1964.

[63] L. Zhang, J. Lin, and R. Karim, "An angle-based subspace anomaly detection approach to high-dimensional data: With an application to industrial fault detection", *Reliability Engineering and System Safety*, Vol. 142, pp. 482-497, 2015.

[64] N. Zhang, M. Fouladirad, and A. Barros, "Maintenance analysis of a two-component load-sharing system", *Reliability Engineering and System Safety*, Vol. 167, pp. 67-74, 2017.

# Publication list of the author

### Refereed journal papers


[1] S. Ota, T. Kageyama and M. Kimura, "Improvement of reliability evaluation for 2-unit parallel system with cascading failures by using Maximal copula", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E98-A, No. 10, pp. 2096-2100, 2015.

[2] S. Ota and M. Kimura, "A study on regression analysis by expanded RBF network based on copula with linear correlation and rank correlation", *International Journal of Reliability, Quality and Safety Engineering*, Vol. 22, No. 5, pp. 1550022-1-1550022-14, 2015.

[3] S. Ota and M. Kimura, "A statistical dependent failure detection method for n-component parallel systems", *Reliability Engineering and System Safety*, Vol. 167, pp. 376-382, 2017.

[4] S. Ota and M. Kimura, "On necessary conditions for dependence parameters of minimum and maximum value distributions based on n-variate FGM copula", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E102-A, No. 3, pp. 586-589, 2019.


### Refereed proceedings of international conferences


[1] S. Ohta, T. Kageyama and M. Kimura, "On a reliability assessment by copula functions for a 2-unit parallel system with cascading failures", *Proceedings of the 8th IMA International Conference on Modelling in*

*Industrial Maintenance and Reliability*, pp. 229-233, Oxford, U.K., July, 2014.

[2] S. Ota and M. Kimura, "A dependent-multivariate data analysis method by skewed-RBF network based on FGM copula", *Proceedings of the 21st ISSAT International Conference on Reliability & Quality in Design*, pp. 256-260, Philadelphia, U.S.A., August, 2015.

[3] S. Ota and M. Kimura, "A method of dependent failure-occurrence detection for parallel systems", *Proceedings of the 9th IMA International Conference on Modelling in Industrial Maintenance and Reliability*, pp. 174-178, London, U.K., July, 2016.

[4] S. Ota and M. Kimura, "A study on the MTTF of parallel systems under dependent failure-occurrence environment", *Proceedings of the 7th Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling*, pp. 423-429, Seoul, Korea, August, 2016.

[5] S. Ota and M. Kimura, "A study on reliability deterioration and improvement of coherent systems under dependent failure-occurrence environment", *Proceedings of the 10th International Conference on Mathematical Methods in Reliability*, 8 pages (USB), Grenoble, France, July, 2017.

[6] S. Ota and M. Kimura, "Parameter estimation for an n-variate FGM copula and its application to dependent failure analysis", *Proceedings of the 24th ISSAT International Conference on Reliability & Quality in Design*, pp. 77-81, Toronto, Canada, August, 2018.

[7] S. Ota and M. Kimura, "A note on statistical inference of dependence parameters of an n-variate FGM copula", *Proceedings of the 8th Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling*, pp. 367-372, Qingdao, China, August, 2018.

**Books**

[1] M. Kimura, S. Ota and S. Abe, "Some reliability properties of n-component parallel/series systems with interdependent failures", in:

Reliability modeling with computer and maintenance applications, World Scientific, Singapore, Chapter 8, pp. 155-173, August, 2017.