学位論文

Distribution theory of nonparametric test statistics and sum of generalized Lindley random variables

（ノンパラメトリック検定統計量および一般化 Lindley 確率変数の和の分布論）

2022 年 3 月

Masato Kitani
（木谷 聖人）

Acknowledgments

First and foremost, I would like to express my greatest appreciation to my thesis advisor, Associate Professor Hidetoshi Murakami. He would discuss my research with me any time, day or night, and gave me great advice. He has always been a constant source of inspiration for me. I am very thankful for his continuous guidance and encouragement throughout my studies. He has shown me how to live and think as a professional researcher.

In addition, I am deeply grateful to Professor Hiroki Hashiguchi for his insightful comments and encouragement and the question that incentivized me to widen my research from various perspectives. Without his guidance and generous support, this dissertation would not have been completed.

I express thanks to all the members of the Murakami laboratory. Through their kind help and support I had a wonderful time studying. My sincere thanks to Mr. Hikaru Yamaguchi for his constructive advice on my study. Sitting next to him, his constant enthusiasm for his own work motivated me to do my best.

I also would like to thank my family for making me what I am today. My parents have always watched over my university life from afar. I could always rely on them for great support whenever I was in pain or distress.

Lastly, I express my gratitude to all those who have helped me in my doctoral course study directly or indirectly.

Masato Kitani

Contents

Chapter 1

Introduction

One-sample and two-sample testing problems are important topics in statistics. Samples are compared using conventional tests based on the assumption that a single population or a population of differences between pairs is normal. Researchers commonly assume a normal distribution when they analyse the experimental data. However, the assumption of normality is often inappropriate in practice. As Büning (1997) and Nanna and Sawilowsky (1998) pointed out, normality is the exception rather than the rule. Micceri (1989) investigated 440 large research data sets in psychology. Regarding symmetry and tails, less than 7% of these data sets were similar to a normal distribution. As a matter of fact, every data set was non-normal at the 1% significance level. Hence, the nonparametric procedure is required when normality cannot clearly be assumed for a specific distribution. In the 1940s, the rank-based approach emerged. Initiated by Wilcoxon (1945), various nonparametric tests were subsequently developed by Mann and Whitney (1947), Mood (1954), Ansari and Bradley (1960), and many others. In the 1950s and 1960s, Pitman (1948), Hodges and Lehmann (1956), and Chernoff and Savage (1958) showed that nonparametric tests have desirable efficiency properties relative to their parametric counterparts. Their work has contributed to the use of nonparametric methods in experimental design and regression analysis. Several statistical resampling methods, such as the jackknife and the bootstrap introduced by Efron (1979), developed since the 1980s, make use of the computational power of computers to provide standard errors and confidence intervals in many applications, including complicated ones where it is difficult, if not impossible, to use a parametric approach. Nowadays, with enhanced computer performance, nonparametric methods have become the mainstream analysis method.

When inferences are related only to the population of differences between paired observations, the first step in the analysis is typically to consider the differences between the paired observations so that only a single set of observations is left. Therefore, this type of data may be legitimately classified as a one-sample problem. Conventional tests are derived based on the assumption that the single population or the population of differences within pairs is normal. In nonparametric procedures, the Wilcoxon signed-rank and the sign tests are commonly used for the location alternative, see, e.g., Gibbons and Chakraborti (2011). These two tests consist of test statistics based on the ranks of the data and associated estimates and confidence intervals for location parameters. The test statistics are distribution-free because the null distribution does not depend on the distribution of errors. Furthermore, these procedures display both robustness of validity and power. In recent years, these nonparametric methods have been extended to linear and nonlinear models by using the pseudo-norm defined by the signed-rank scores. For more details, see Hettmansperger and McKean (2010). In addition, the Wilcoxon signed-rank test has been extended and applied to various data structures. For example, Rosner et al. (2006) extended the Wilcoxon signed-rank test to clustered data. Clustered data are characterized as data that can be classified into a number of distinct groups within a particular study. Examples of clustered data include electrophysiological or optical recordings taken from synaptic terminals, repeated measurement of blood pressure from a single individual, responses of litter mates in an experiment using rodents, or body mass index of siblings. Recently, Rosenblatt and Benjamini (2018) compared the performance of the t-test and the Wilcoxon signed-rank test under mixture alternatives, and results show that the Wilcoxon signed-rank test is more useful than the t-test when normality cannot be assumed. The Bayesian analysis for the Wilcoxon signed-rank test has also been considered by Benavoli et al. (2014) and Chechile (2018).

Supposing assumptions can be made concerning the forms of the underlying populations and assuming that the difference between the two populations is with respect to the means only, then, the population can be treated as a two-sample location problem. If it is assumed that both populations follow the normal distribution, the famous and the powerful test for equality of means is the two-sample Student's t-test. As in the case of one-sample problems, if there is not enough information about the underlying distribution, the nonparametric procedure is desirable. There are many good and simple nonparametric tests for the two-sample location problem based on ranks. Since the ranks of the first sample relative to the ranks of the second sample provide information on the population medians, many researchers have proposed various score functions to reflect the difference of samples adequately. One of the most famous and powerful nonparametric two-sample tests is the Wilcoxon rank-sum test (Wilcoxon, 1945). Note that the rank-sum test was first proposed by Deuchler (1914) and discovered independently by Mann and Whitney (1947). Under the non-normal population distribution, the Wilcoxon rank-sum test is more powerful than the t-test for the location alternative as shown in Hodges and Lehmann (1956) and Neave and Granger (1968). The asymptotic relative efficiency (ARE) of the Wilcoxon rank-sum test to the t-test is about 95.5% under the assumption of normal populations differing in location. The ARE of the t-test may exceed one for non-normal cases. In addition, Hodges and Lehmann (1956) show that the ARE of the Wilcoxon rank-sum test relative to that of the t-test never falls below 0.864. For a linear rank test, Goria (1980) derived the distribution for which the test is locally most powerful. They showed that the Wilcoxon rank-sum score is the locally most powerful test when the logistic distribution is assumed. Therefore, the Wilcoxon rank-sum test is useful when the underlying distribution cannot be assumed.

In the nonparametric method, an asymptotic theory is constructed for a large sample size and the performance is compared with other nonparametric and parametric tests. Both Pitman efficiency and the ARE, which are given by Pitman (1948) and Noether (1955), are often used to compare the performance of two or more tests. In addition, it is necessary to calculate the critical value of the test statistic or the p-value when testing a problem. Therefore, deriving the asymptotic and the limiting null distributions have historically played an important role in the field of statistics. With nonparametric methods, obtaining the exact critical value is often difficult for large sample sizes due to extensive calculations. Under the circumstances, the distribution of the statistic has to be approximated. The use of the asymptotic distribution has been studied as an approximation for the statistic, and many researchers have carried out simulations to determine whether the asymptotic theory works well. Some have devised ways to improve the order of the approximation. Edgeworth (1905) provided a similar expansion as an improvement to the central limit theorem. The Edgeworth expansion can estimate the error between the exact distribution of the standardized statistic and the standard normal distribution. Cornish and Fisher (1938) proposed an alternative expansion. The Edgeworth expansion and the Cornish-Fisher expansion, which approximate the distribution by a function depending on the number of samples, have been studied and applied to complex statistics. When the first few moments are known, the Edgeworth expansion can be used and this method is often satisfactory in practice. However, the Edgeworth expansion has the drawback that the approximation can assume negative values in the tail regions of the distribution. Daniels (1954) introduced an approximation method by using the inversion of the characteristic function, so called the saddlepoint approximation. The error of the saddlepoint approximation is $O(n^{-1})$ compared to the more usual $O(n^{-1/2})$ associated with the normal approximation. Note that Monti (1993) shows the relationship between the Edgeworth expansion and saddlepoint approximations. In the literature, the saddlepoint approximation has been used with great success and discussed by many authors, e.g. Daniels (1987), Easton and Ronchetti (1986), Reid (1988), Jensen (1995), Goutis and Casella (1999), Huzurbazar (1999), Kolassa (2006), Butler (2007), and Eisinga et al. (2013). For the two-sample nonparametric statistics, Froda and van Eeden (2000) proposed a uniform saddlepoint expansion to the null distribution of the Wilcoxon rank-sum test, and Bean et al. (2004) compared the saddlepoint approximation of the Wilcoxon rank-sum test with that of Edgeworth, and determined normal and uniform approximations. In addition, Murakami (2010) and Murakami and Kamakura (2009) proposed a saddlepoint approximation to the distribution of the Bagai statistic (Bagai et al., 1989) and Jonckheere-Terpstra statistic (Jonckheere, 1954; Tertpstra, 1952), respectively.

One of the most important properties of the nonparametric test is the unbiasedness. The alternative hypothesis is typically more important to researchers than the null hypothesis as the former describes their scientific conjecture. In addition, the power of the test should be sensitive to the magnitude of the effect. Higher effects are needed if the power of the biased test reaches the same power as the unbiased test. The biased test is not sensitive to the magnitude of the effect. Thus, if the test rejects the null hypothesis with a probability of less than the significance level under the alternative of interest, it is unlikely that the test can be recommended to researchers. As Jurečková et al. (2019) indicated, the finite sample unbiasedness of some tests is still an open question. For the one-sample problem, Lehmann and Romano (2005) showed that both the Wilcoxon signed-rank and sign tests are unbiased with regard to the one-sided alternative. Amrhein (1995) provided a counter example of the Wilcoxon signed rank test, which is not unbiased with respect to the two-sided alternative. However, Lehmann and Romano (2005) mentioned that it is not known whether these results are admissible within the class of all rank tests. Therefore, the unbiasedness has to be proved for each test. Many one-sample goodness-of-fit tests are not the unbiased test. For example, Massey (1950), Thompson (1966), and Frey (2009) showed that the Kolmogorov-Smirnov test, the Cramér-von Mises test, and the weighted Kolmogorov-Smirnov test are not the unbiased test, respectively. Moreover, Ding et al. (2010) showed that neither the Berk-Jones test (Berk and Jones, 1979) nor the reversed Berk-Jones test (Jager and Wellner, 2007) are unbiased test by focusing on the structure of the confidence bands. Then, Hanyuda and Murakami (2021) composed the unbiased Berk-Jones test and reversed Berk-Jones test by applying the algorithm of Frey (2009). Since the biased test is considered undesirable, it is important to construct an unbiased test with a bias correction if the test is not unbiased.

There are different kinds of nonparametric tests. The Wilcoxon rank-sum test is powerful for symmetric distributions with medium or large tails. In particular, the Wilcoxon rank-sum test is the locally most powerful test when the underlying distribution is the logistic distribution. However, in a nonparametric model, it is not known whether the distributions are symmetric or not, whether they have short or heavy tails. Hájek et al. (1999) specified the locally optical score function which depends on the underlying distribution function and the corresponding density function. Hogg (1974) and Büning (1991) introduced a method to choose between pre-selected score functions. The statistics has to be selected before a test can be performed. Then, Büning (1996) proposed a selector for the adaptive test for two-sample problems by using measures from the order statistics. In addition, the researcher can use tests that are unbiased with respect to the two-sided alternative described above by choosing the appropriate statistic.

Sums of random variables also arise naturally in many application areas: for example, analysis of wireless communications; PERT networks; software reliability estimation; project management processes; to mention but a few. Then, many researchers derive and approximate the distribution of the sum of random variables for various distributions. For example, the sum of weighted χ^2 random variables appears in many important problems in statistics. Various statistical inferences lead to the problem of evaluating the probability of the sum of weighted χ^2 random variables. Computational methods including numerical inversion of the characteristic function of the sum of weighted χ^2 variables for various approximations were reviewed in Solomon and Stephens (1977). Furthermore, the theoretical approach of an approximation was presented along with an easily implementable algorithm in Gabler and Wolff (1987). A review of the current state for the sum of weighted non-central χ^2 random variables can be found in Duchesne and De Micheaux (2010), and methods for computing the cumulative distribution function of a single non-central χ^2 random variable are described in Ding (1992); Farebrother (1987); Penev and Raykov (2000). Recently, Miyazaki and Murakami (2020) considered the Fourier series approximation to the distribution of the sum of weighted non-central χ^2 random variables.

The distributions of sum, minima and maxima of generalized geometric random variables are considered when applying other distributions (Tank and Eryilmaz, 2015). The sum of independent, albeit not necessarily identical uniformly distributed random variables, arises naturally in the aggregation of scaled values with differing numbers of significant figures. The distribution of the sum of uniform random variables was obtained by Olds (1952) using the mathematical induction method. In addition, Bradley and Gupta (2002) derived the explicit formulae for the distribution by inverting the characteristic function. In a different approach, Sadooghi-Alvandi et al. (2009) determined this distribution by employing a Laplace transform, seemingly utilizing prior knowledge of the result. Potuschak and Müller (2009) provided a simplified derivation of the distribution of the sum of independent and non-identically distributed (inid) uniform random variables via an inverse Fourier transform. Mathai (1982) obtained the distribution of the sum of inid gamma random variables, and Moschopoulos (1985) provided an expression for the single gamma series, computing the coefficients using simple recursive relations. Since the square of a Nakagami random variable (Nakagami, 1960) follows a gamma distribution, the sum of inid gamma random variables is required in wireless communications. Alouini et al. (2001) considered applying Moschopoulos's approach to the distribution of the sum of correlated gamma random variables. In summary, determination of the distribution of the sum of inid random variables is an important topic in many scientific fields, see e.g., Nadarajah (2008).

The exponential distribution has been widely applied in many scientific fields. Kamps (1990) characterized the exponential distribution as a distribution of a weighted sum of independent, identically distributed random variables. Khuong and Kong (2006) obtained the probability density function with distinct or equal parameters using the characteristic function. Amari and Misra (1997) discussed the case of non-identically exponential random variables. Numerous authors have examined several extensions to the exponential distribution. Gómez et al. (2014) proposed an extended exponential distribution, which is useful in fitting real data. The extended exponential distribution is also considered as the extension of the Lindley distribution. The Lindley distribution was first introduced by Lindley (1958). Ghitany et al. (2008) suggested that many situations exist for which the Lindley distribution is a better model than the exponential distribution. Researchers have proposed extensions to the Lindley distribution. Zakerzadeh and Dolati (2009) introduced the generalized Lindley distribution which contains the original Lindley, the exponential, the gamma, and the extended exponential distributions. However, the exact distributions of the sum of inid generalized Lindley random variables is not clarified.

In this study, following previous studies, the approximation, the asymptotic, and the limiting null distributions of the one-sample and two-sample nonparametric statistics for location alternatives based on Kitani and Murakami (2020a) and Kitani and Murakami (2022) are discussed. In Chapter 2, the extension of the sign and the Wilcoxon signed-rank (hereinafter referred to as ESWSR) tests proposed by Policello and Hettmansperger (1976) is presented. Since the computational cost of deriving the critical value of the test by exact permutation is enormous, the saddlepoint approximation and the normal approximation to the distribution of the ESWSR statistic are applied. In addition, the unbiasedness of the test is discussed and a selector for the ESWSR statistic is proposed in terms of the asymptotic efficiency. In Chapter 3, the asymptotic and limiting null distributions of the combining t and Wilcoxon rank-sum test proposed by Neuhäuser (2015) is derived. The convergence of the maximum test to the limiting distribution is investigated for various cases via Monte Carlo simulations. In addition, it is demonstrated that calculating the limiting distribution for given data through estimators can be useful. In Chapter 4, the exact distribution of the sum of inid extended exponential random variables and the sum of inid generalized Lindley random variables are derived. These results are based on the work of Kitani and Murakami (2020b) and Kitani et al. (2021). The exact distribution contains the infinite gamma series; thus, a finite number of terms are truncated. However, since the computation of the iterations takes a lot of time, the saddlepoint and the normal approximations are also considered and the accuracy is compared with that of the numerical results. Furthermore, it is demonstrated that the distribution of the sum of inid conventional random variables is a special case of the sum of inid generalized Lindley random variables. Finally, in Chapter 5, the paper is concluded with a summary.

Chapter 2

Properties of one-sample test statistic

In one-sample testing problems, both the sign test and the Wilcoxon signed-rank test are widely used. To construct an adaptive robust procedure, Policello and Hettmansperger (1976) introduced a statistic which contains these two traditional tests. Since the calculation of the exact distribution of the statistic is not feasible, the null distribution is formulated by using a saddlepoint approximation. In addition, the unbiasedness of the test is discussed as it is an important property in analysis. A selector statistic based on the asymptotic efficiency is further proposed.

2.1 Revisiting one-sample nonparametric tests

Let $|X_1| \leq |X_2| \leq \cdots \leq |X_n|$ be the ordered observations for a random sample from $F(x-\theta)$, where F is a symmetric continuous distribution related to θ. Then, without loss of generality, the following hypothesis is tested:

$$H_{10} : \theta = 0 \quad \text{against} \quad H_{11} : \theta > 0 \quad (\text{or} \quad H_{12} : \theta \neq 0)\,.$$

Then, the Wilcoxon signed-rank test statistic is defined by $\sum_{i=1}^{n} i\mathbb{I}(X_i > 0)$, while the sign test statistic is defined by $\sum_{i=1}^{n} \mathbb{I}(X_i > 0)$, where $\mathbb{I}(\cdot)$ is the indicator function. Subsequently, Policello and Hettmansperger (1976) proposed the ESWSR test statistic, namely, T_ν, as follows:

$$T_\nu = \sum_{i=1}^{n} a_\nu(i)\mathbb{I}(X_i > 0), \quad 0 \leq \nu \leq 1, \tag{2.1}$$

where

$$a_\nu(i) = \begin{cases} i, & 1 \leq i \leq n_\nu = \lfloor (1-\nu)n \rfloor \\ 1 + n_\nu, & 1 + n_\nu \leq i \leq n \end{cases} = \min[i, 1 + n_\nu].$$

It is worth mentioning that $\nu = 1$ is equivalent to the sign test, which is asymptotically the most powerful rank test when F is a Laplace distribution. Moreover, $\nu = 0$ is equivalent to the Wilcoxon signed-rank test, which is asymptotically the most powerful rank test when F is a logistic distribution (Hájek, 1962). Therefore, Policello and Hettmansperger (1976) suggested the use of the T_ν statistic for distributions with tail weights, such as those of the Laplace and logistic families.

Policello and Hettmansperger (1976) investigated the power and robustness properties of the test through Monte Carlo simulations. It is important to calculate the critical value of the test statistic in testing problems. However, obtaining the exact critical value is often difficult when the sample size is large. Under these circumstances, the critical values have to be approximated. A saddlepoint approximation is considered for the distribution of the T_ν statistic. Furthermore, the accuracy of the saddlepoint approximation is compared with the exact probability and normal approximation. It is also important to determine the properties of the statistic under the alternative hypothesis. Therefore, the asymptotic power of the statistic is obtained by calculating the first and second moments under the alternative hypothesis H_{11}. In addition, the unbiasedness of the ESWSR test is discussed and a selector statistic, which is the rule of selecting ν, is proposed. In addition, the use of the ESWSR test is demonstrated on real data.

2.2 Approximations to the test statistic

In this section, we consider the saddlepoint and normal approximations for the test statistic T_ν. Policello and Hettmansperger (1976) derived the probability generating function of T_ν as

$$G(s) = 2^{-n}(1 + s^{1+n_\nu})^{n-n_\nu} \prod_{k=1}^{n_\nu}(1 + s^k) = 2^{-n} \prod_{i=1}^{n}(1 + s^{a_\nu(i)}). \tag{2.2}$$

Then, the moment generating function and cumulant generating function of T_ν are obtained by substituting $\exp(s)$ into Equation (2.2) as follows:

$$M(s) = 2^{-n}(1 + \exp\{s(1 + n_\nu)\})^{n-n_\nu} \prod_{k=1}^{n_\nu}(1 + \exp(sk)), \tag{2.3}$$

$$\kappa(s) = \log M(s) = -n \log 2 + (n - n_\nu) \log(1 + \exp\{s(1 + n_\nu)\}) + \sum_{k=1}^{n_\nu} \log\{1 + \exp(sk)\}, \tag{2.4}$$

respectively.

From Equations (2.3) and (2.4) and the Lugannani and Rice formula (Lugannani and Rice, 1980),

$$\Pr(T_\nu \geq t) = 1 - \Phi(\tilde{w}) + \left\{\frac{1}{\tilde{u}_1} - \frac{1}{\tilde{w}}\right\} \phi(\tilde{w}). \tag{2.5}$$

Here, $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density function and the corresponding cumulative distribution function, respectively; $\tilde{w} = \sqrt{2(\tilde{s}t - \kappa(\tilde{s}))}\text{sgn}(\tilde{s})$; $\tilde{u}_1 = \tilde{s}\sqrt{\kappa''(\tilde{s})}$; $\tilde{s}$ is a root of the saddlepoint equation $\kappa'(\tilde{s}) = t$; and κ' and κ'' are the first and second derivatives of κ, respectively. For continuity correction, we utilize $\tilde{u}_2 = (1 - e^{-\tilde{s}})\sqrt{\kappa''(\tilde{s})}$, rather than the $\tilde{u}_1$ in Equation (2.5).

In this study, we investigate the accuracy of the approximations. We calculate the upper probability $\Pr(T_\nu \geq t)$ by using Mathematica version 11. In the tables, t and "Exact" represent the exact critical value and probability of $\Pr(T_\nu \geq t)$, respectively. Additionally, NA (SA) and NA_c (SA_c) are the normal (saddlepoint) approximations with and without continuity corrections. In this study, the sample sizes are assumed to be $n = 10, 15, 20, 25, 30$ and $\nu = 0, 0.2, 0.4, 0.6, 0.8, 1.0$. The significance levels are 0.05 and 0.01.

Table 2.1: $\Pr(T_\nu \geq t)$ for the 5% significance level.

	ν	0	0.2	0.4	0.6	0.8	1.0
$n = 10$	t	45	44	40	32	22	9
	Exact	0.0420	0.0430	0.0381	0.0430	0.0273	0.0107
	NA	0.0372	0.0378	0.0336	0.0368	**0.0264**	0.0057
	SA	0.0367	0.0374	0.0331	0.0366	0.0256	0.0045
	NA_c	0.0416	0.0423	0.0383	**0.0432**	0.0341	0.0134
	SA_c	**0.0418**	**0.0427**	**0.0384**	0.0438	0.0340	**0.0116**
$n = 15$	t	90	88	78	62	40	12
	Exact	0.0473	0.0443	0.0466	0.0461	0.0394	0.0176
	NA	0.0442	0.0414	0.0432	0.0414	0.0350	0.0101
	SA	0.0443	0.0414	0.0434	0.0415	0.0348	0.0092
	NA_c	0.0469	0.0441	**0.0464**	0.0454	**0.0408**	0.0194
	SA_c	**0.0473**	**0.0443**	**0.0468**	**0.0458**	**0.0408**	**0.0179**
$n = 20$	t	150	145	129	101	63	15
	Exact	0.0487	0.0498	0.0471	0.0484	0.0457	0.0207
	NA	0.0465	0.0475	0.0447	0.0454	0.0413	0.0127
	SA	0.0467	0.0477	0.0448	0.0455	0.0413	0.0120
	NA_c	0.0483	0.0494	0.0468	0.0482	**0.0457**	0.0221
	SA_c	**0.0487**	**0.0498**	**0.0471**	**0.0485**	0.0459	**0.0209**
$n = 25$	t	225	217	192	150	91	18
	Exact	0.0479	0.0491	0.0476	0.0462	0.0491	0.0216
	NA	0.0463	0.0474	0.0458	0.0440	0.0457	0.0139
	SA	0.0465	0.0476	0.0459	0.0441	0.0458	0.0134
	NA_c	0.0476	0.0488	0.0474	0.0461	**0.0492**	0.0228
	SA_c	**0.0479**	**0.0491**	**0.0476**	**0.0462**	0.0494	**0.0218**
$n = 30$	t	314	303	267	207	125	20
	Exact	0.0481	0.0480	0.0476	0.0484	0.0462	0.0494
	NA	0.0468	0.0467	0.0463	0.0466	0.0433	0.0339
	SA	0.0470	0.0469	0.0464	0.0467	0.0433	0.0338
	NA_c	0.0479	0.0478	0.0475	0.0482	0.0459	0.0502
	SA_c	**0.0480**	**0.0480**	**0.0477**	**0.0484**	**0.0460**	**0.0496**

Table 2.2: $\Pr(T_\nu \geq t)$ for the 1% significance level.

	ν	0	0.2	0.4	0.6	0.8	1.0
$n = 10$	t	50	49	44	36	25	10
	Exact	0.0098	0.0098	0.0098	0.0068	0.0029	0.0010
	NA	0.0109	0.0107	0.0107	0.0085	0.0044	**0.0008**
	SA	0.0082	0.0082	0.0086	**0.0068**	**0.0030**	-
	NA_c	0.0125	0.0123	0.0124	0.0104	0.0061	0.0022
	SA_c	**0.0100**	**0.0100**	**0.0106**	0.0088	0.0048	-
$n = 15$	t	101	98	87	69	44	13
	Exact	0.0090	0.0094	0.0097	0.0099	0.0092	0.0037
	NA	0.0099	0.0101	0.0102	0.0096	**0.0089**	0.0023
	SA	0.0082	0.0085	0.0089	0.0085	0.0079	0.0017
	NA_c	0.0107	0.0109	0.0111	0.0108	0.0107	0.0049
	SA_c	**0.0090**	**0.0094**	**0.0098**	**0.0097**	0.0098	**0.0038**
$n = 20$	t	167	162	143	112	70	16
	Exact	0.0096	0.0093	0.0097	0.0098	0.0086	0.0059
	NA	0.0103	0.0099	0.0101	0.0098	0.0079	0.0036
	SA	0.0091	0.0088	0.0091	0.0090	0.0072	0.0031
	NA_c	0.0108	0.0104	0.0107	0.0106	0.0091	0.0070
	SA_c	**0.0096**	**0.0093**	**0.0097**	**0.0098**	**0.0083**	**0.0060**
$n = 25$	t	249	240	212	165	100	19
	Exact	0.0094	0.0098	0.0094	0.0094	0.0098	0.0073
	NA	0.0100	0.0103	0.0097	0.0095	**0.0098**	0.0047
	SA	0.0090	0.0094	0.0090	0.0088	0.0092	0.0042
	NA_c	0.0103	0.0107	0.0102	0.0101	0.0108	0.0082
	SA_c	**0.0094**	**0.0098**	**0.0094**	**0.0094**	0.0102	**0.0074**
$n = 30$	t	345	333	293	227	137	22
	Exact	0.0098	0.0097	0.0096	0.0096	0.0087	0.0081
	NA	0.0103	0.0101	0.0099	0.0097	0.0085	0.0053
	SA	0.0095	0.0094	0.0093	0.0092	0.0080	0.0049
	NA_c	0.0106	0.0104	0.0102	0.0102	0.0091	0.0088
	SA_c	**0.0098**	**0.0097**	**0.0096**	**0.0096**	**0.0087**	**0.0081**

Tables 2.1 to 2.2 show that the saddlepoint approximation with continuity correction is the closest to the exact probability for $n = 30$. Although no significant differences are observed between the saddlepoint approximation and the normal approximation, the results revealed that the saddlepoint approximation of the statistic is useful when the sample size increases. Note that the exact method requires extensive calculations to obtain the critical values, while the saddlepoint approximation generates the critical values instantly. Therefore, the saddlepoint approximation is appropriate for the ESWSR statistic.

2.3 Asymptotic power

In this section, we derive the asymptotic power of the statistic T_ν by calculating the first and second moments of T_ν under H_{11}. Herein, the statistic T_ν in Equation (2.1) is rewritten as follows:

$$T_\nu = \sum_{i=1}^{n} \min[i, 1 + n_\nu]\mathbb{I}(X_i > 0). \tag{2.6}$$

Defining $\mathrm{E}_A[T]$ as the expectation of T_ν under H_{11}, possible partitions of the indicator function in Equation (2.6) are considered to express the first moment of T_ν as,

$$\begin{aligned}
\mathrm{E}_A[T_\nu] &= \sum_{i=1}^{n}\sum_{k=0}^{n-1}\binom{n-1}{k}\min[k+1, 1+n_\nu] \\
&\quad\times \mathrm{E}_A[\mathbb{I}(X_i > 0)\underbrace{\mathbb{I}(|X_i| > |X'_1|)\cdots\mathbb{I}(|X_i| > |X'_k|)}_{k}\underbrace{\mathbb{I}(|X_i| < |X'_{k+1}|)\cdots\mathbb{I}(|X_i| < |X'_{n-1}|)}_{n-k-1}] \\
&= n\sum_{k=0}^{n-1}\binom{n-1}{k}\min[k+1, 1+n_\nu] \\
&\quad\times \int_0^\infty \{F(x) - F(-x)\}^k \{1 - F(x) + F(-x)\}^{n-k-1}\, dF(x),
\end{aligned} \tag{2.7}$$

where X' is an unordered sample of X. In addition, the second moment is

$$\begin{aligned}
\mathrm{E}_A[T_\nu^2] &= \sum_{i=1}^{n}\mathrm{E}_A\left[\min[i, 1+n_\nu]^2\mathbb{I}(X_i > 0)\right] \\
&\quad + \mathrm{E}_A\left[\sum\sum_{i\neq j}\min[i, 1+n_\nu]\min[j, 1+n_\nu]\mathbb{I}(X_i > 0)\mathbb{I}(X_j > 0)\right].
\end{aligned} \tag{2.8}$$

Now focusing on the second term of the right-hand side of Equation (2.8), the following is obtained:

$$\begin{aligned}
&\mathrm{E}_A\left[\sum\sum_{i\neq j}\min[i, 1+n_\nu]\min[j, 1+n_\nu]\mathbb{I}(X_i > 0)\mathbb{I}(X_j > 0)\right] \\
&= 2\sum\sum_{i\neq j}\sum_{a=0}^{n-2}\sum_{b=0}^{n-2-a}\frac{(n-2)!}{a!b!(n-2-a-b)!}\min[a+1, 1+n_\nu]\min[a+b+2, 1+n_\nu]
\end{aligned}$$

$$
\begin{aligned}
&\times \mathrm{E}_A[\mathbb{I}(X_i > 0)\mathbb{I}(X_j > 0)\mathbb{I}(|X_i| < |X_j|) \\
&\times \underbrace{\mathbb{I}(|X_1'| < |X_i| < |X_j|)\cdots}_{a}\underbrace{\mathbb{I}(|X_i| < |X_{a+1}'| < |X_j|)\cdots}_{b}\underbrace{\mathbb{I}(|X_i| < |X_j| < |X_{a+b+1}'|)\cdots}_{n-2-a-b}] \\
= &\ 2n(n-1)\sum_{a=0}^{n-2}\sum_{b=0}^{n-2-a}\frac{(n-2)!}{a!b!(n-2-a-b)!}\min[a+1, 1+n_\nu]\min[a+b+2, 1+n_\nu] \\
&\times \int_0^\infty\int_0^{x_2}\{F(x_1) - F(-x_1)\}^a\{F(x_2) - F(x_1) + F(-x_1) - F(-x_2)\}^b \\
&\times \{1 - F(x_2) + F(-x_2)\}^{n-2-a-b}\, dF(x_1)dF(x_2). \qquad (2.9)
\end{aligned}
$$

Thus, with the exact mean and variance of T_ν under H_{11} from Equations (2.7) and (2.8), the asymptotic power under H_{11} is obtained as,

$$
\frac{1}{\sqrt{2\pi}}\int_c^\infty \exp\left(-\frac{1}{2}x^2\right)dx, \quad c = \frac{z_\alpha \operatorname{Var}_0[T_\nu] + \mathrm{E}_0[T_\nu] - E_A[T_\nu]}{\sqrt{\operatorname{Var}_A[T_\nu]}}, \qquad (2.10)
$$

where z_α is α percentile for the standard normal distribution,

$$
\mathrm{E}_0[T_\nu] = \frac{1}{4}(n_\nu + 1)(2n - n_\nu), \quad \operatorname{Var}_0[T_\nu] = \frac{1}{24}(n_\nu + 1)(-4n_\nu^2 - 5n_\nu + 6n + 6nn_\nu)
$$

are the expectation and the variance of T_ν under H_{10}, and $\operatorname{Var}_A[T_\nu]$ is the variance of T_ν under H_{11}.

Table 2.3: Numerical comparisons for the normal distribution.

		ν								
Δ		0	0.2	0.35	0.4	0.55	0.6	0.75	0.8	1.0
0	Sim	0.054	0.050	0.054	0.054	0.055	0.055	0.053	0.053	0.059
	AP	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
0.2	Sim	0.187	0.178	0.182	0.182	0.179	0.178	0.167	0.167	0.173
	AP	0.173	0.172	0.169	0.169	0.163	0.161	0.157	0.157	0.148
0.4	Sim	0.434	0.418	0.418	0.418	0.403	0.403	0.375	0.376	0.368
	AP	0.396	0.393	0.383	0.383	0.368	0.361	0.347	0.347	0.321
0.6	Sim	0.711	0.695	0.687	0.688	0.665	0.664	0.628	0.628	0.604
	AP	0.675	0.670	0.654	0.654	0.628	0.617	0.592	0.592	0.546
0.8	Sim	0.898	0.888	0.882	0.882	0.862	0.863	0.831	0.831	0.804
	AP	0.904	0.899	0.883	0.883	0.856	0.844	0.818	0.818	0.764
1.0	Sim	0.976	0.973	0.970	0.970	0.960	0.960	0.944	0.944	0.924
	AP	0.992	0.991	0.986	0.986	0.975	0.969	0.953	0.953	0.914
1.2	Sim	0.997	0.996	0.995	0.995	0.992	0.992	0.986	0.986	0.978
	AP	1.000	1.000	1.000	1.000	0.999	0.999	0.996	0.996	0.982

Table 2.4: Numerical comparisons for the Laplace distribution.

		ν								
Δ		0	0.2	0.35	0.4	0.55	0.6	0.75	0.8	1.0
0	Sim	0.054	0.050	0.054	0.054	0.055	0.055	0.053	0.053	0.059
	AP	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
0.2	Sim	0.170	0.163	0.175	0.175	0.183	0.184	0.181	0.181	0.196
	AP	0.158	0.159	0.163	0.163	0.168	0.169	0.171	0.171	0.169
0.4	Sim	0.363	0.356	0.377	0.376	0.393	0.392	0.387	0.388	0.398
	AP	0.334	0.338	0.349	0.349	0.359	0.362	0.363	0.363	0.348
0.6	Sim	0.581	0.574	0.599	0.599	0.616	0.617	0.606	0.607	0.605
	AP	0.543	0.550	0.565	0.565	0.578	0.580	0.576	0.576	0.546
0.8	Sim	0.761	0.756	0.778	0.779	0.790	0.791	0.779	0.779	0.767
	AP	0.739	0.746	0.761	0.761	0.770	0.769	0.760	0.760	0.720
1.0	Sim	0.880	0.878	0.893	0.893	0.900	0.900	0.889	0.890	0.874
	AP	0.883	0.888	0.898	0.898	0.902	0.900	0.889	0.889	0.850
1.2	Sim	0.947	0.945	0.954	0.953	0.957	0.957	0.950	0.949	0.937
	AP	0.963	0.965	0.970	0.970	0.970	0.970	0.960	0.960	0.931

Table 2.5: Numerical comparisons for the logistic distribution.

		ν								
Δ		0	0.2	0.35	0.4	0.55	0.6	0.75	0.8	1.0
0	Sim	0.053	0.050	0.054	0.054	0.055	0.055	0.053	0.052	0.059
	AP	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
0.2	Sim	0.118	0.113	0.119	0.119	0.119	0.118	0.114	0.113	0.120
	AP	0.111	0.111	0.111	0.111	0.109	0.109	0.107	0.107	0.103
0.4	Sim	0.226	0.216	0.224	0.224	0.222	0.222	0.210	0.210	0.214
	AP	0.208	0.208	0.207	0.207	0.203	0.201	0.195	0.195	0.185
0.6	Sim	0.370	0.358	0.367	0.367	0.362	0.361	0.341	0.341	0.339
	AP	0.338	0.338	0.337	0.337	0.330	0.326	0.316	0.316	0.295
0.8	Sim	0.532	0.520	0.529	0.528	0.520	0.519	0.493	0.493	0.481
	AP	0.492	0.493	0.490	0.490	0.480	0.474	0.459	0.459	0.426
1.0	Sim	0.687	0.676	0.682	0.682	0.671	0.671	0.642	0.641	0.623
	AP	0.652	0.653	0.650	0.650	0.635	0.628	0.608	0.608	0.565
1.2	Sim	0.811	0.802	0.808	0.807	0.797	0.797	0.769	0.769	0.746
	AP	0.797	0.797	0.792	0.792	0.776	0.776	0.746	0.746	0.697

Table 2.6: Numerical comparisons for the Student's t-distribution with two degrees of freedom.

		ν								
Δ		0	0.2	0.35	0.4	0.55	0.6	0.75	0.8	1.0
0	Sim	0.053	0.050	0.053	0.054	0.055	0.055	0.053	0.053	0.059
	AP	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
0.2	Sim	0.144	0.138	0.148	0.149	0.153	0.152	0.146	0.146	0.155
	AP	0.134	0.135	0.138	0.138	0.139	0.139	0.138	0.138	0.133
0.4	Sim	0.293	0.288	0.308	0.308	0.316	0.316	0.305	0.306	0.311
	AP	0.270	0.275	0.284	0.284	0.288	0.288	0.284	0.284	0.270
0.6	Sim	0.477	0.473	0.502	0.502	0.514	0.516	0.501	0.501	0.499
	AP	0.396	0.414	0.442	0.442	0.450	0.450	0.412	0.412	0.381
0.8	Sim	0.647	0.649	0.679	0.679	0.695	0.696	0.683	0.683	0.674
	AP	0.614	0.627	0.650	0.650	0.662	0.662	0.654	0.654	0.618
1.0	Sim	0.779	0.782	0.811	0.810	0.827	0.827	0.817	0.817	0.806
	AP	0.761	0.775	0.800	0.800	0.812	0.812	0.804	0.804	0.765
1.2	Sim	0.868	0.872	0.895	0.895	0.908	0.908	0.901	0.902	0.891
	AP	0.868	0.880	0.901	0.901	0.911	0.911	0.904	0.904	0.871

Table 2.7: Numerical comparisons for the Cauchy distribution.

		ν								
Δ		0	0.2	0.35	0.4	0.55	0.6	0.75	0.8	1.0
0	Sim	0.054	0.050	0.054	0.054	0.055	0.055	0.053	0.053	0.059
	AP	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050	0.050
0.2	Sim	0.119	0.114	0.126	0.126	0.133	0.132	0.131	0.130	0.142
	AP	0.110	0.112	0.117	0.117	0.122	0.123	0.124	0.124	0.121
0.4	Sim	0.216	0.214	0.237	0.238	0.255	0.256	0.256	0.255	0.270
	AP	0.200	0.206	0.220	0.220	0.234	0.237	0.240	0.240	0.233
0.6	Sim	0.336	0.335	0.371	0.372	0.405	0.404	0.404	0.406	0.421
	AP	0.309	0.319	0.345	0.345	0.370	0.376	0.382	0.382	0.370
0.8	Sim	0.455	0.459	0.504	0.504	0.549	0.549	0.552	0.553	0.566
	AP	0.422	0.436	0.473	0.473	0.509	0.517	0.525	0.525	0.508
1.0	Sim	0.563	0.569	0.617	0.618	0.669	0.669	0.674	0.674	0.687
	AP	0.526	0.545	0.590	0.590	0.632	0.642	0.652	0.652	0.632
1.2	Sim	0.650	0.660	0.708	0.707	0.762	0.762	0.766	0.768	0.777
	AP	0.617	0.638	0.689	0.689	0.734	0.734	0.754	0.754	0.733

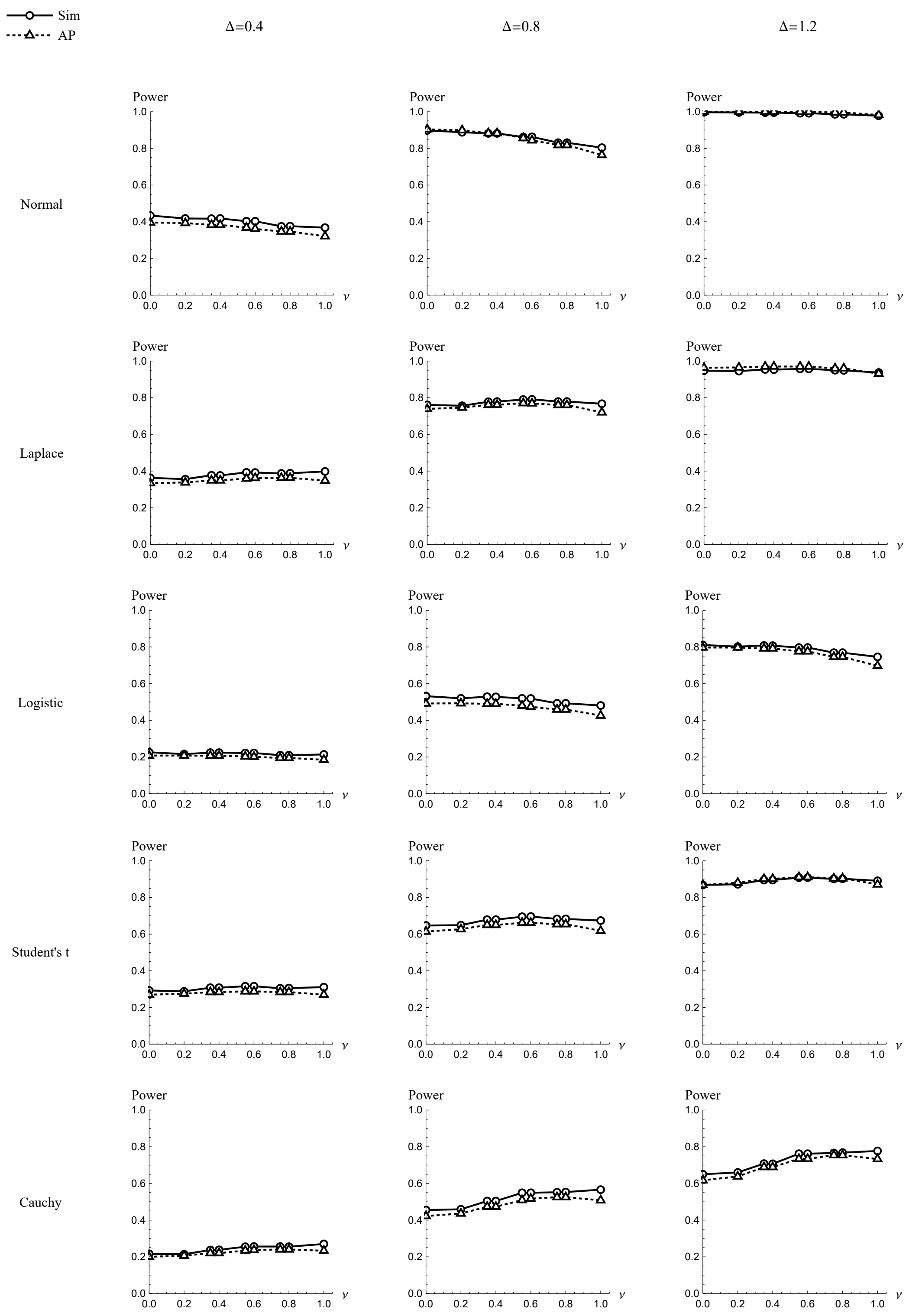

Figure 2.1: Simulated powers and asymptotic powers for various distributions for $\Delta = 0.4, 0.8, 1.2$.

Herein, Tables 2.3 to 2.7 illustrate the numerical comparison between the asymptotic and simulated power using Mathematica and R software. Mathematica was used for both the saddlepoint and normal approximation. In addition, Mathematica was used to calculate the asymptotic power as it contains multiple integrals. Meanwhile, R software was used for the Monte Carlo simulations of the power. The distributions considered were: the normal distributions $N(0,1)$ and $N(\theta,1)$ in Table 2.3, the Laplace distributions Laplace$(0,1)$ and Laplace$(\theta,1)$ in Table 2.4, the logistic distributions Logis$(0,1)$ and Logis$(\theta,1)$ in Table 2.5, the Student's t-distributions with two degrees of freedom $t(2)$ and $t(2)+\theta$ in Table 2.6, and the Cauchy distributions Cauchy$(0,1)$ and Cauchy$(\theta,1)$ in Table 2.7, where $\theta = 0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.2$. In addition, it was assumed that $n = 15$, $\nu = 0, 0.2, 0.35, 0.4, 0.55, 0.6, 0.75, 0.8, 1.0$, while the significance level was 0.05. In the tables, Sim. represents the simulated power based on 1,000,000 Monte Carlo simulations, and AP represents the asymptotic power when Equation (2.10) is used.

Tables 2.3 to 2.7 and Figure 2.1 show that the asymptotic power is roughly the same as the simulated power. Although the asymptotic power is discussed for a large sample size, even a small sample size of $n = 10$ and 15 yielded values close to those of the simulation. Therefore, it is guaranteed that the calculation of the asymptotic power is correct. Note that calculating the asymptotic power is time consuming; therefore, to overcome the computational complexity, the algorithm of the numerical calculation has to be improved. For $n = 20$, the calculations for the summations in Equation (2.9) are extensive, and hence, performing the calculation becomes impossible. However, the derivation of the exact mean and variance under the alternative hypothesis is meaningful. Although the asymptotic power can be calculated as an application, exact calculation is impossible when the sample size is large. Addressing this problem would enable the accurate calculation of the asymptotic power.

2.4 Unbiasedness and selector

In this section, the focus is on the unbiasedness of the ESWSR test. Let Ω be a parameter space and let $\Omega_H, \Omega_K \subset \Omega$, $\Omega_K = \Omega \setminus \Omega_H$. The test Ψ of the null hypothesis $H : \theta \in \Omega_H$ against the alternative $K : \theta \in \Omega_K$ is said to be unbiased if

$$\sup_{\theta \in \Omega_H} \beta_\Psi(\theta) \leq \alpha \quad \text{and} \quad \inf_{\theta \in \Omega_K} \beta_\Psi(\theta) \geq \alpha,$$

where $\beta_\Psi(\theta)$ is the power function and α the chosen significance level. By applying results from Lehmann and Romano (2005, p.324), it is proved that the ESWSR test is unbiased for the one-sided location alternative H_{11}. Thus, we obtain Theorem 2.1.

Theorem 2.1. *The ESWSR test is unbiased against the one-sided location alternative* H_{11}.

Proof. Let R_i, $i = 1, 2, \ldots, n^*$, denote the ranks of the absolute values of observation $X_1, X_2, \ldots, X_n$ whose signs are positive. The adaptive Wilcoxon singed rank test is given by

$$T_\nu = a(R_1) + a(R_2) + \cdots + a(R_{n^*}),$$

and the alternative hypothesis H_{11} is rejected when T_ν is sufficiently large, say $T_\nu \geq c$. We now consider the power function $\beta(\theta) = \Pr(T_\nu \geq c)$ of the test against the location sift alternative. Let $\theta_0 < \theta_1$, and $X_1, \ldots, X_n$ be independently distributed with distribution $F(x - \theta_0)$. If $U_i = X_i + (\theta_1 - \theta_0)$, the distribution of the random variables $U_1, \ldots, U_n$ is given by

$$\Pr(U_i \leq x) = \Pr(X_i + \theta_1 - \theta_0 \leq x) = \Pr(X_i \leq x - (\theta_1 - \theta_0)) = F(x - \theta_1).$$

Hence, $\beta(\theta_0) = \Pr(T_X \geq c)$, $\beta(\theta_1) = \Pr(T_U \geq c)$, where T_X (T_U) is the statistic calculated from X (Y). From the fact that $X_i < U_i$ for all i, it is seen that $T_X \leq T_U$ and $\beta(\mu_0) \leq \beta(\theta_1)$. Therefore, since the power function $\beta(\theta)$ is a nondecreasing function of θ, we have the result. □

Furthermore, the biasedness of the ESWSR test for the two-sided alternative is demonstrated. A similar process as that of Amrhein (1995) is followed to provide a counterexample. Suppose $F(x)$ has a density function,

$$f(x) = \begin{cases} \frac{1}{2}, & \frac{1}{2} < |x| < \frac{3}{2}, \\ 0, & \text{otherwise}, \end{cases} \tag{2.11}$$

and let $n = 3$. Then we consider the distribution of $S = (S_1, S_2, S_3)$ for $\theta = 0$ and $\theta = \frac{1}{2}$, where $S_i = \mathbb{I}(X_i > 0)$. For $\theta = 0$, it is generally known that each observation $s = (s_1, s_2, s_3) \in \{0, 1\}^3$ appears with the same probability

$$\Pr_{\theta=0}\{S = s\} = \frac{1}{8}.$$

When $\theta = \frac{1}{2}$, the density function (2.11) is replaced with $f(x - \frac{1}{2})$. Figure 2.2 displays the density function $f(x - \frac{1}{2})$.

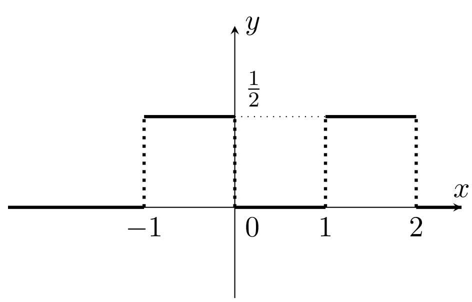

Figure 2.2: Density function under the alternative.

Subsequently, based on the assumption that $|X_1| \leq |X_2| \leq |X_3|$, the distribution of S is provided in Table 2.8.

Table 2.8: Distributions of $S = (s_1, s_2, s_3)$ for some ν.

s_1	s_2	s_3	$\Pr_{\theta=0}$	$\Pr_{\theta=1/2}$	T_0	$T_{1/3}$	$T_{2/3}$	T_1
0	0	0	1/8	1/8	0	0	0	0
1	0	0	1/8	0	1	1	1	1
0	1	0	1/8	0	2	2	2	1
0	0	1	1/8	3/8	3	3	2	1
1	1	0	1/8	0	3	3	3	2
1	0	1	1/8	0	4	4	3	2
0	1	1	1/8	3/8	5	5	4	2
1	1	1	1/8	1/8	6	6	5	3

Therefore, for $\nu = 0$ and $\nu = 1/3$, which correspond to the Wilcoxon signed rank test, the test is not unbiased in this example. Hence, the ESWSR test is not unbiased against the two-sided location alternative H_{12}. However, when $\nu = 2/3$ or $\nu = 1$ is selected, the ESWSR test is unbiased. Therefore, the appropriate ν in terms of bias correction should be selected.

We are also interested in determining ν for the specific distribution. To perform the ESWSR test, ν has to be determined in advance. However, a statistician cannot assume the underlying distribution in practice; therefore, a selector is required to determine the appropriate ν for the given dataset. Thus, we propose utilizing a selector to maximize the asymptotic efficiency of the test. The asymptotic efficiency was obtained in Policello and Hettmansperger (1976). Thereafter, the type of the underlying unknown distribution function is classified with respect to two measures of kurtosis, as discussed in Hogg (1974), Randles and Hogg (1973), and Büning (2001). We implement the following measures:

$$\hat{M}_1 = \frac{\hat{z}_{0.975} - \hat{z}_{0.025}}{\hat{z}_{0.875} - \hat{z}_{0.125}}, \quad \hat{M}_2 = \frac{\hat{z}_{0.65} - \hat{z}_{0.35}}{\hat{z}_{0.55} - \hat{z}_{0.45}}$$

with the empirical p-quantile $\hat{z}_p$ given by

$$\hat{z}_p = \begin{cases} Z_{(1)}, & p \le 0.5/n \\ (1-\lambda)Z_{(j)} + \lambda Z_{(j+1)}, & 0.5/n < p \le 1 - 0.5/n \\ Z_{(n)}, & p > 1 - 0.5/n \end{cases},$$

where $Z_{(1)}, \ldots, Z_{(n)}$ are the order of the samples and $j = \lfloor np + 0.5 \rfloor$, $\lambda = np + 0.5 - j$. Table 2.9 presents the theoretical values of $\hat{M}_1$ and $\hat{M}_2$ for various distributions.

As the asymptotic efficiency of the ESWSR test depends on the weight of the tail of the distribution, the value of ν is selected to obtain a higher power. For heavy-tailed distributions, a large value of M_1 and a value of M_2 around 3.3 are appropriate. For light-tailed distributions, a small value of M_1 and a value of M_2 around 3.0 are appropriate. Through this rule of selecting ν, a test, that can easily reject the null hypothesis for the given data, is adopted. Let $M = (\hat{M}_1, \hat{M}_2)$, which is called a selector. Hence, four categories can be defined based on the measures $\hat{M}_1$ and $\hat{M}_2$:

$$D_1 = \{M;\ 0 \le \hat{M}_1 \le 1,\ 0 \le \hat{M}_2 \le 3.3\},$$

Table 2.9: Theoretical $\hat{M}_1$ and $\hat{M}_2$ values for some distributions.

Distribution	$\hat{M}_1$	$\hat{M}_2$
Uniform	1.267	3.000
Normal	1.704	3.066
Logistic	1.883	3.085
Laplace	2.161	3.385
Student's t ($df = 2$)	2.683	3.129
Cauchy	5.263	3.217
Hyperbolic sec	2.004	3.105

$$D_2 = \{M;\ 1 < \hat{M}_1 \leq 2.5,\ 0 \leq \hat{M}_2 \leq 3.3\},$$
$$D_3 = \{M;\ 2.5 < \hat{M}_1 \leq 5.0,\ 0 \leq \hat{M}_2 \leq 3.3\},$$
$$D_4 = \{M;\ 5.0 < \hat{M}_1,\ 0 \leq \hat{M}_2 \leq 3.3\},$$
$$D_5 = \{M;\ 3.3 < \hat{M}_2\}.$$

Hence, we propose the following rule for selecting ν based on asymptotic efficiency:

$$\nu = \begin{cases} 0, & M \in D_1 \\ 0.35, & M \in D_2 \\ 0.55, & M \in D_3 \\ 0.75, & M \in D_4 \\ 1, & M \in D_5, \end{cases} \tag{2.12}$$

and Figure 2.3 illustrates the rule of selecting ν.

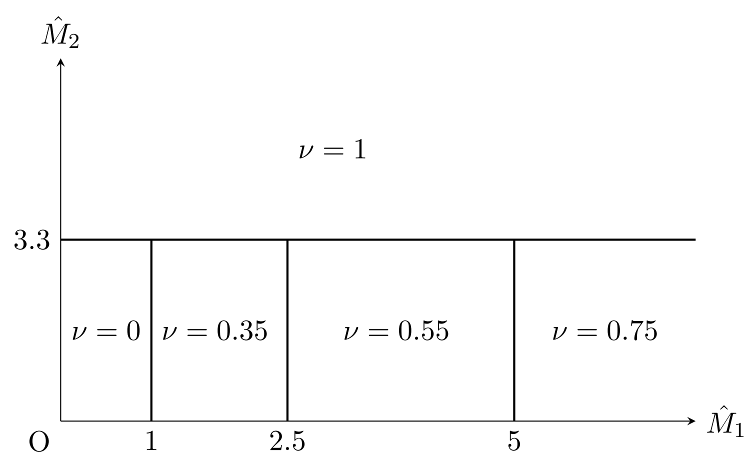

Figure 2.3: Illustration of the adaptive scheme.

Tables 2.3 to 2.7 and Figure 2.1 indicate that the selector proposed in this section is advantageous, in general. For example, when the underlying distribution is normal, $\nu = 0$ is selected.

In contrast, the case $\nu = 0.55$ is relatively powerful when a Student's t-distribution with two degrees of freedom is assumed. Meanwhile, the case for $\nu = 0.75$ is more powerful when the Cauchy distribution is presumed.

2.5 Real data analysis

For clinical testing, the method is applied to cosmetics data. The dataset was obtained from Gibbons and Chakraborti (2011, p.224). A manufacturer of suntan lotion is testing a new formula (Y) to determine whether it provides more protection against sunburn than the old formula (X). Specifically, the higher numbers represent more severe sunburn. Then, the null hypothesis is to verify the difference ($X - Y$) of the degree of sunburn, which has median zero against the positive one-sided alternative.

Old Formula (X)	41	42	48	38	38	45	21	28	29	14
New Formula (Y)	37	39	31	39	34	47	19	30	25	8

Initially, we obtain $\hat{M}_1 = 1.76$ and $\hat{M}_2 = 2.00$, and then $\nu = 0.35$ is selected using Equation (2.12). Therefore, for the ESWSR test, $T_{0.35} = 43$, the exact p-value is 0.0098. For the Wilcoxon singed-rank and sign tests, $T_0 = 48$ and $T_1 = 7$, and the exact p-values are 0.0137 and 0.0547, respectively. These results show that the null hypothesis can be rejected at the 1% significance level, only with the ESWSR test.

Chapter 3

Distribution of two-sample test statistic

In two-sample testing problems, Student's t-test is widely used under the assumption of normality. Alternatively, the Wilcoxon rank-sum test is used in nonparametric two-sample testing problems. Neuhäuser (2015) proposed a maximum test combining the t-test and the Wilcoxon rank-sum test and showed that the power of the maximum test is similar to the more powerful of the two tests. However, the limiting distribution of the test has not been derived. Therefore, we derive the limiting distribution and investigate the convergence of the null distribution of the maximum test to the limiting null distribution, via Monte Carlo simulations, for various cases. The usefulness of the maximum test is also demonstrated by applying the test to reaction time data.

3.1 Review of two-sample tests

Let $\boldsymbol{X} = (X_1, \ldots, X_{n_1})$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_{n_2})$ be two independent samples of size n_1 and n_2 $(N = n_1 + n_2)$ with a common variance σ^2, from a continuous cumulative distribution function F and G, respectively. In addition, let f and g be the probability density function corresponding to F and G, respectively. In the location-shift model, $F(x) = G(x - \theta)$ for every x. Note that we the focus on the maximum test for one-sided alternatives because the two-sided Wilcoxon rank-sum test is not unbiased with respect to location parameters (Sugiura, 1965; Sugiura et al., 2006). Because of the unbiasedness, testing the following hypothesis is of interest:

$$H_{20} : \theta = 0 \quad \text{against} \quad H_{21} : \theta > 0.$$

Student's t-test is defined by

$$t = \sqrt{\frac{n_1 n_2}{N}} \frac{\bar{X} - \bar{Y}}{S_t}, \quad \text{where} \quad S_t^2 = \frac{(n_1 - 1)S_X^2 + (n_2 - 1)S_Y^2}{N - 2},$$

and $\bar{X}$ $(\bar{Y})$ is the sample mean and S_X^2 (S_Y^2) is the unbiased sample variance of $\boldsymbol{X}$ $(\boldsymbol{Y})$ as follows:

$$S_X^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} \left(X_i - \bar{X}\right)^2, \quad S_Y^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} \left(Y_j - \bar{Y}\right)^2.$$

Furthermore, R_i denotes the rank of X_i with combined samples $(\boldsymbol{X}, \boldsymbol{Y})$ and $W = \sum R_i$ is the sum of the ranks of $\boldsymbol{X}$'s, where $i = 1, \ldots, n_1$. In Conover and Iman (1981), a rank transformation procedure for computing the t-test is introduced based on the ranks R_i as follows:

$$t_R = \left[\frac{1}{n_1}W - \frac{1}{n_2}\left(\frac{N(N+1)}{2} - W\right)\right]$$
$$\div \left[\left(\frac{N(N+1)(2N+1)}{6} - \frac{1}{n_1}W^2 - \frac{1}{n_2}\left(\frac{N(N+1)}{2} - W\right)^2\right)\frac{N}{n_1 n_2(N-2)}\right]^{1/2}.$$

This studentized test is widely applied in many fields. For example, Chung and Romano (2016) discuss the two-sample Wilcoxon rank-sum test using this statistic. Recently, Neuhäuser (2015) proposed a maximum test combining the t-test and the Wilcoxon rank-sum test $\max(|t|, |t_R|)$. In Neuhäuser (2015), it is shown that the power of the maximum test is close to that of the better of the t-test or the Wilcoxon rank-sum test. They conclude that the maximum test is useful when there is no special reason to use either the t-test or the Wilcoxon rank-sum test. In Welz et al. (2018), the nonparametric maximum test is suggested for the Behrens-Fisher problem. Here, we focus on the maximum test without the absolute value $\max(t, t_R)$ because the test is one-sided.

Since the critical value of the maximum test proposed by Neuhäuser (2015) depends on the permutation method, the calculation is extensive when sample sizes increase. Moreover, the asymptotic and the limiting distributions have historically played an important role in all statistical fields. Here, the limiting distribution of the maximum test is discussed under the hypotheses of this chapter.

3.2 Limiting distribution of the maximum test statistic

In this section, the null and non-null limiting distributions of the maximum test are derived. Let H_j be the rank of Y_j with combined samples $(\boldsymbol{X}, \boldsymbol{Y})$ and $V = \sum H_j$ be the sum of the ranks of $\boldsymbol{Y}$'s, where $j = 1, 2, \ldots, n_2$. Moreover, define $\bar{R} = W/n_1$ and $\bar{H} = V/n_2$. Then, the following proposition is proved to obtain the theorem of the limiting distribution.

Proposition 3.1. *Let*

$$S_R^2 = \frac{1}{N-2}\left\{\sum_{i=1}^{n_1}(R_i - \bar{R})^2 + \sum_{j=1}^{n_2}(H_j - \bar{H})^2\right\}, \quad \sigma_R^2 = \frac{1}{12}N(N+1).$$

Then, the expectation of S_R^2 is equal to σ_R^2.

Proof. By simple calculation,

$$\begin{aligned} \mathrm{E}[S_R^2] &= \frac{1}{N-2}\,\mathrm{E}\left[\sum_{i=1}^{n_1} R_i^2 + \sum_{j=1}^{n_2} H_j^2 - \frac{1}{n_1}W^2 - \frac{1}{n_2}V^2\right] \\ &= \frac{1}{N-2}\left\{\frac{1}{6}N(N+1)(2N+1) - \frac{1}{n_1}\,\mathrm{E}[W^2] - \frac{1}{n_2}\,\mathrm{E}[V^2]\right\}. \end{aligned}$$

Now focusing on the term $\mathrm{E}[W^2]$, we obtain

$$\begin{aligned}
\mathrm{E}[W^2] &= \mathrm{E}\left[\left(\sum_{i=1}^{n_1} R_i\right)^2\right] \\
&= \mathrm{E}\left[\sum_{i=1}^{n_1} R_i^2 + \sum_{i\neq k} R_i R_k\right] \\
&= \frac{1}{12} n_1(n_1+n_2+1)(3n_1^2+3n_1n_2+3n_1+n_2).
\end{aligned}$$

Similarly,

$$\mathrm{E}[V^2] = \frac{1}{12} n_2(n_1+n_2+1)(3n_2^2+3n_1n_2+3n_2+n_1).$$

Hence,

$$\mathrm{E}[S_R^2] = \frac{1}{12} N(N+1) = \sigma_R^2.$$

Then, the result holds. □

Nádas (1971) derived the exact distribution of the maximum of two normal random variables. By applying Nádas (1971)'s result, the limiting distribution of $\max(t, t_R)$ is stated in Theorem 3.2.

Theorem 3.2. *Let X and Y be independent random variables with distribution functions F and G, respectively. Then, the limiting distribution of* $\max(t, t_R)$ *is*

$$L_1(x) = \int_{-\infty}^{x} \{\ell_{11}(-s) + \ell_{21}(-s)\}\, ds,$$

where

$$\begin{aligned}
\ell_{11}(-s) &= \phi\left(\mathrm{E}[t]-s\right)\Phi\left(\frac{\rho(\mathrm{E}[t]-s)}{\sqrt{1-\rho^2}} - \frac{\mathrm{E}[t_R]-s}{\sqrt{1-\rho^2}}\right), \\
\ell_{21}(-s) &= \phi\left(\mathrm{E}[t_R]-s\right)\Phi\left(\frac{\rho(\mathrm{E}[t_R]-s)}{\sqrt{1-\rho^2}} - \frac{\mathrm{E}[t]-s}{\sqrt{1-\rho^2}}\right),
\end{aligned}$$

$\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density function and the corresponding cumulative distribution function, respectively, and

$$\begin{aligned}
\mathrm{E}[t] &= \frac{\theta}{\sigma}\sqrt{\frac{n_1 n_2}{N}}, \\
\mathrm{E}[t_R] &= \left(\int_{-\infty}^{\infty} G(x) f(x) dx - \frac{1}{2}\right)\sqrt{\frac{12 n_1 n_2}{N^2(N+1)}}, \\
\rho &= \mathrm{E}[t \cdot t_R] - \frac{n_1 n_2 \theta}{N\sigma}\sqrt{\frac{12}{N(N+1)}}\left(\int_{-\infty}^{\infty} G(x) f(x) dx - \frac{1}{2}\right),
\end{aligned}$$

$$
\begin{aligned}
\mathrm{E}[t\cdot t_R] = \frac{1}{\sigma}\sqrt{\frac{12}{N(N+1)}}\Bigg[&\frac{n_1(n_1+1)}{2}\,\mathrm{E}[X] + n_2\,\mathrm{E}[XG(X)] + n_2(n_1-1)\,\mathrm{E}[X]\,\mathrm{E}[G(X)] \\
&+\frac{n_2(n_2+1)}{2}\,\mathrm{E}[Y] + n_1\,\mathrm{E}[YF(Y)] + n_1(n_2-1)\,\mathrm{E}[Y]\,\mathrm{E}[F(Y)] \\
&-\frac{n_1n_2(N+1)}{2}\left(\frac{1}{n_2}\,\mathrm{E}[X] + \frac{1}{n_1}\,\mathrm{E}[Y]\right)\Bigg].
\end{aligned}
$$

Proof. At first, $\mathrm{E}[t]$ and $\mathrm{E}[t_R]$ are introduced under H_{21}. It is well known that the t-statistic for populations with a common variance σ^2 is

$$
t = \sqrt{\frac{n_1n_2}{N}}\left(\frac{\bar{X}-\bar{Y}}{S_t}\right) = \sqrt{\frac{n_1n_2}{N}}\left(\frac{\bar{X}-\bar{Y}-\theta}{\sigma} + \frac{\theta}{\sigma}\right)\frac{\sigma}{S_t},
$$

where

$$
S_t^2 = \frac{1}{N-2}\left\{\sum_{i=1}^{n_1}(X_i-\bar{X})^2 + \sum_{j=1}^{n_2}(Y_j-\bar{Y})^2\right\}.
$$

Since $S_t/\sigma \to 1$ as $N\to\infty$, $n_1/N \in (0,1)$, then,

$$
\mathrm{E}[t] = \frac{\theta}{\sigma}\sqrt{\frac{n_1n_2}{N}},\quad \mathrm{Var}[t] = \frac{n_1n_2}{N}\frac{\sigma^2/n_1+\sigma^2/n_2}{\sigma^2} = 1.
$$

Following a similar procedure for Student's t-statistic, by using Proposition 3.1, the statistic t_R can be expressed as

$$
t_R = \sqrt{\frac{n_1n_2}{N}}\left(\frac{\bar{R}-\bar{H}}{S_R}\right) = \sqrt{\frac{n_1n_2}{N}}\left(\frac{\bar{R}-\bar{H}-\theta_R}{\sigma_R} + \frac{\theta_R}{\sigma_R}\right)\frac{\sigma_R}{S_R},
$$

$S_R/\sigma_R \to 1$ as $N\to\infty$, $n_1/N\in(0,1)$. Then,

$$
\mathrm{E}[t_R] = \frac{\theta_R}{\sigma_R}\sqrt{\frac{n_1n_2}{n_1+n_2}},\quad \mathrm{Var}[t_R] = \frac{n_1n_2}{n_1+n_2}\frac{\sigma_R^2/n_1+\sigma_R^2/n_2}{\sigma_R^2} = 1.
$$

Herein, θ_R is determined. Define

$$
U(i,j) = \begin{cases} 1, & X_i < Y_j, \\ 0, & X_i > Y_j, \quad i=1,\dots,n_1,\ j=1,\dots,n_2. \end{cases}
$$

If $F\le G$, then

$$
\begin{aligned}
\theta_R &= \mathrm{E}[\bar{R}-\bar{H}] \\
&= \mathrm{E}\left[\frac{1}{n_1}\sum_{i=1}^{n_1}R_i\right] - \mathrm{E}\left[\frac{1}{n_2}\sum_{j=1}^{n_2}H_j\right] \\
&= \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\mathrm{E}\left[\sum_{i=1}^{n_1}R_i\right] - \frac{N(N+1)}{2n_2}
\end{aligned}
$$

$$
\begin{aligned}
&= \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \mathrm{E}\left[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} U(i,j) + \frac{n_1(n_1+1)}{2}\right] - \frac{N(N+1)}{2n_2} \\
&= \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \mathrm{E}\left[U(i,j)\right] + \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \frac{n_1(n_1+1)}{2} - \frac{N(N+1)}{2n_2} \\
&= \frac{n_1+n_2}{n_1 n_2} n_1 n_2 \int_{-\infty}^{\infty} F(x)g(x)dx + \frac{(n_1+n_2)(n_1+1)}{2n_2} - \frac{N(N+1)}{2n_2} \\
&= N \int_{-\infty}^{\infty} F(x)g(x)dx - \frac{N}{2}.
\end{aligned}
$$

Next, the correlation of t and t_R, that is ρ, is given by

$$
\rho = \frac{\mathrm{Cov}(t, t_R)}{S_t S_R} = \frac{\mathrm{E}[t \cdot t_R] - \mathrm{E}[t] \cdot \mathrm{E}[t_R]}{S_t S_R}.
$$

First, consider the term $\mathrm{E}[t \cdot t_R]$. Since $S_t/\sigma \to 1$ and $S_R/\sigma_R \to 1$ as $N \to \infty$, then

$$
\begin{aligned}
\mathrm{E}[t \cdot t_R] &= \frac{n_1 n_2}{N} \mathrm{E}\left[\frac{\bar{X}-\bar{Y}}{S_t} \cdot \frac{\bar{R}-\bar{H}}{S_R}\right] \\
&= \frac{n_1 n_2}{N} \mathrm{E}\left[\frac{\bar{X}-\bar{Y}}{\sigma} \cdot \frac{\sigma}{S_t} \cdot \frac{\bar{R}-\bar{H}}{\sigma_R} \cdot \frac{\sigma_R}{S_R}\right] \\
&= \frac{n_1 n_2}{N} \frac{1}{\sigma\sigma_R} \mathrm{E}\left[(\bar{X}-\bar{Y})\left(\frac{1}{n_1}W - \frac{1}{n_2}V\right)\right] \\
&= \frac{n_1 n_2}{N} \frac{1}{\sigma\sigma_R} \mathrm{E}\left[\frac{1}{n_1}\bar{X}W + \frac{1}{n_2}\bar{Y}V - \frac{1}{n_2}\bar{X}\left(\frac{N(N+1)}{2} - W\right) - \frac{1}{n_1}\bar{Y}\left(\frac{N(N+1)}{2} - V\right)\right] \\
&= \frac{1}{\sigma\sigma_R} \mathrm{E}\left[\bar{X}W + \bar{Y}V - \frac{n_1(N+1)}{2}\bar{X} - \frac{n_2(N+1)}{2}\bar{Y}\right].
\end{aligned}
$$

Second, consider the term $\mathrm{E}[\bar{X}W]$. Then,

$$
\begin{aligned}
\mathrm{E}[\bar{X}W] &= \mathrm{E}\left[\frac{1}{n_1}\sum_{k=1}^{n_1} X_k \left\{\sum_{i=1}^{n_1}\left(i + \sum_{j=1}^{n_2} \mathbb{I}(Y_j < X_{(i)})\right)\right\}\right] \\
&= \frac{1}{n_1}\sum_{k=1}^{n_1} \mathrm{E}\left[X_k \sum_{i=1}^{n_1} i\right] + \frac{1}{n_1}\mathrm{E}\left[\sum_{k=1}^{n_1} X_k \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \mathbb{I}(Y_j < X_{(i)})\right] \\
&= \frac{n_1(n_1+1)}{2}\mathrm{E}[X] + \frac{n_2}{n_1}\sum_{k=i}\sum \mathrm{E}[X_k \mathbb{I}(Y < X_i)] + \frac{n_2}{n_1}\sum_{k\neq i}\sum \mathrm{E}[X_k \mathbb{I}(Y < X_i)] \\
&= \frac{n_1(n_1+1)}{2}\mathrm{E}[X] + n_2\,\mathrm{E}[XG(X)]] + n_2(n_1-1)\,\mathrm{E}[X]\,\mathrm{E}[G(X)]].
\end{aligned}
$$

Similarly,

$$
\mathrm{E}[\bar{Y}V] = \frac{n_2(n_2+1)}{2}\mathrm{E}[Y] + n_1\,\mathrm{E}[YF(Y)]] + n_1(n_2-1)\,\mathrm{E}[Y]\,\mathrm{E}[F(Y)]].
$$

When $\mathrm{E}[\bar{X}W]$ and $\mathrm{E}[\bar{Y}V]$ are substituted into $\mathrm{E}[t \cdot t_R]$, the results hold. □

Corollary 3.3. *Under H_{20}, the null distribution is obtained by substituting $\theta = 0$ into Theorem 3.2. In this case,* $\mathrm{E}[t] = \mathrm{E}[t_R] = 0$, $\mathrm{E}[F(X)] = 1/2$, *and the correlation coefficient is given by*

$$\rho_0 = \frac{1}{\sigma}\sqrt{\frac{12N}{N+1}}\left(\mathrm{E}[XF(X)] - \frac{1}{2}\,\mathrm{E}[X]\right). \tag{3.1}$$

Corollary 3.4. *For the limiting null distribution $n_1, n_2 \to \infty$, the correlation coefficient is given by*

$$\rho_* = \frac{\sqrt{3}}{\sigma}\left(2\,\mathrm{E}[XF(X)] - \mathrm{E}[X]\right). \tag{3.2}$$

Note that the one-sided hypothesis $H_{22} : \theta < 0$ considers the statistic $\min(t, t_R)$, and the corresponding limiting distribution is obtained as follows:

$$L_2(x) = \int_{-\infty}^{x} \left\{\ell_{12}(-s) + \ell_{22}(-s)\right\} ds,$$

where

$$\ell_{12}(-s) = \phi\left(-s - \mathrm{E}[t]\right)\Phi\left(\frac{s + \mathrm{E}[t_R]}{\sqrt{1-\rho^2}} - \frac{\rho(s + \mathrm{E}[t])}{\sqrt{1-\rho^2}}\right),$$
$$\ell_{22}(-s) = \phi\left(-s - \mathrm{E}[t_R]\right)\Phi\left(\frac{s + \mathrm{E}[t]}{\sqrt{1-\rho^2}} - \frac{\rho(s + \mathrm{E}[t_R])}{\sqrt{1-\rho^2}}\right),$$

$\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density function and the corresponding cumulative distribution function, respectively. The limiting null distribution of $\max(|t|, |t_R|)$ is further obtained for the two-sided hypothesis $H_{23} : \theta \neq 0$ by applying Philonenko et al. (2016)'s result. The cumulative distribution function is given by

$$L_3(x) = \int_0^x 4\phi(s)\left\{\Phi_0\left(s\sqrt{\frac{1-\rho}{1+\rho}}\right) + \Phi_0\left(s\sqrt{\frac{1+\rho}{1-\rho}}\right)\right\} ds,$$

where

$$\Phi_0(x) = \int_0^x \phi(s)ds.$$

3.3 Numerical results

In the previous section, the limiting distribution of the $\max(t, t_R)$ test was derived under the null and the alternative hypotheses. In this section, we verify the convergence to each of the critical points of the asymptotic and limiting distributions for various cases. Mathematica version 11 was employed to investigate the behavior of the $\max(t, t_R)$ test as it approached the limiting distribution through simulation studies. Here, the null distribution was estimated with $1,000,000$ iterations of the Monte Carlo simulations. In this study, equal sample sizes $n_1 = n_2$ as well as unequal sample sizes $2n_1 = n_2$, $n_1 = 2n_2$ were considered. Because of the focus on convergence of the $\max(t, t_R)$, set $N = 40$, 100, 200, 1,000 and 2,000 for equal cases, and $N = 60$, 120, 300, 1,200 and 3,000 for unequal cases. The results for the standard normal distribution are presented in Table 3.1, the standard exponential distribution in Table 3.2, and the standard logistic distribution in Table 3.3. In the tables, Sim. and Perm. represent the critical point, based on Monte Carlo simulations and permutations, respectively. Note that 1,000 permutations and 1,000 iterations are run to obtain the permutation critical values. In addition, CV_A and CV_L represent the asymptotic critical points which are calculated by correlation (3.1) and (3.2), respectively. In our study, it is assumed that populations have the same variance. In practice, however, it is necessary to estimate the population variance. Hence, the population variance σ was estimated by the sample variance and the sample mean $\mathrm{E}[X]$. In addition,

$$\frac{1}{n_1(n_1-1)}\sum_{i=1}^{n_1}(i-1)x_{(i)}$$

was utilized as an unbiased estimator of $\mathrm{E}[XF(X)]]$, where $x_{(i)}$ is the i-th order statistic of the sample (Landwehr et al., 1979). In the tables, CV_E represents the mean of the estimated critical points based on 1,000,000 iterations. Furthermore, RMSE represents the root mean squared error between the estimated critical points CV_E and the asymptotic critical point CV_A.

Table 3.1: Critical value of $\max(t, t_R)$ for the standard normal distribution.

	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990
CV_L	1.362	1.724	2.038	2.403	1.362	1.724	2.038	2.403	1.362	1.724	2.038	2.403
	$(n_1, n_2) = (20, 20)$				$(20, 40)$				$(40, 20)$			
Sim.	1.396	1.779	2.117	2.527	1.375	1.755	2.088	2.486	1.384	1.755	2.091	2.480
Perm.	1.395	1.779	2.117	2.520	1.377	1.755	2.084	2.485	1.382	1.755	2.092	2.486
CV_A	1.380	1.741	2.055	2.419	1.374	1.736	2.050	2.414	1.374	1.736	2.050	2.414
CV_E	1.349	1.711	2.026	2.391	1.341	1.703	2.017	2.383	1.360	1.722	2.036	2.401
RMSE	0.050	0.048	0.047	0.046	0.053	0.052	0.051	0.050	0.030	0.029	0.028	0.027
	$(50, 50)$				$(40, 80)$				$(80, 40)$			
Sim.	1.378	1.748	2.067	2.452	1.369	1.739	2.059	2.437	1.371	1.744	2.065	2.439
Perm.	1.371	1.741	2.067	2.452	1.371	1.740	2.059	2.437	1.370	1.737	2.059	2.437
CV_A	1.370	1.732	2.045	2.410	1.369	1.730	2.044	2.409	1.369	1.730	2.044	2.409
CV_E	1.357	1.719	2.033	2.398	1.351	1.713	2.028	2.393	1.361	1.723	2.037	2.402
RMSE	0.028	0.027	0.026	0.026	0.035	0.034	0.033	0.032	0.020	0.019	0.019	0.018
	$(100, 100)$				$(100, 200)$				$(200, 100)$			
Sim.	1.369	1.734	2.050	2.424	1.366	1.730	2.048	2.417	1.362	1.731	2.048	2.420
Perm.	1.366	1.731	2.052	2.422	1.366	1.727	2.047	2.420	1.368	1.732	2.050	2.421
CV_A	1.366	1.728	2.042	2.407	1.365	1.727	2.041	2.406	1.365	1.727	2.041	2.406
CV_E	1.360	1.722	2.036	2.401	1.358	1.720	2.034	2.400	1.362	1.724	2.038	2.403
RMSE	0.018	0.018	0.017	0.017	0.019	0.018	0.018	0.017	0.013	0.012	0.012	0.011
	$(500, 500)$				$(400, 800)$				$(800, 400)$			
Sim.	1.359	1.723	2.041	2.409	1.362	1.725	2.038	2.406	1.363	1.723	2.034	2.400
Perm.	1.363	1.727	2.043	2.408	1.362	1.722	2.035	2.401	1.363	1.729	2.042	2.407
CV_A	1.363	1.725	2.039	2.404	1.363	1.725	2.039	2.404	1.363	1.725	2.039	2.404
CV_E	1.362	1.724	2.038	2.403	1.361	1.723	2.037	2.402	1.362	1.724	2.038	2.403
RMSE	0.008	0.008	0.007	0.007	0.009	0.009	0.008	0.008	0.006	0.006	0.006	0.006
	$(1000, 1000)$				$(1000, 2000)$				$(2000, 1000)$			
Sim.	1.360	1.723	2.039	2.402	1.362	1.724	2.037	2.408	1.365	1.726	2.040	2.408
Perm.	1.362	1.726	2.040	2.409	1.364	1.724	2.040	2.401	1.362	1.725	2.040	2.403
CV_A	1.363	1.725	2.039	2.404	1.362	1.724	2.038	2.404	1.362	1.724	2.038	2.404
CV_E	1.362	1.724	2.038	2.403	1.362	1.724	2.038	2.403	1.362	1.724	2.038	2.403
RMSE	0.006	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.004	0.004	0.004	0.004

Table 3.2: Critical value of $\max(t, t_R)$ for the standard exponential distribution.

	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990
CV_L	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486
	$(n_1, n_2) = (20, 20)$				$(20, 40)$				$(40, 20)$			
Sim.	1.469	1.839	2.149	2.542	1.439	1.776	2.079	2.448	1.502	1.879	2.217	2.614
Perm.	1.478	1.839	2.149	2.542	1.435	1.772	2.079	2.450	1.500	1.881	2.219	2.609
CV_A	1.467	1.822	2.131	2.490	1.465	1.821	2.129	2.489	1.465	1.821	2.129	2.489
CV_E	1.425	1.783	2.094	2.456	1.421	1.780	2.091	2.452	1.443	1.800	2.110	2.471
RMSE	0.065	0.061	0.057	0.054	0.068	0.064	0.060	0.057	0.043	0.039	0.037	0.034
	$(50, 50)$				$(40, 80)$				$(80, 40)$			
Sim.	1.467	1.822	2.138	2.505	1.443	1.787	2.083	2.437	1.485	1.858	2.183	2.568
Perm.	1.465	1.821	2.136	2.508	1.443	1.787	2.086	2.431	1.477	1.843	2.164	2.534
CV_A	1.463	1.819	2.128	2.488	1.463	1.819	2.128	2.487	1.463	1.819	2.128	2.487
CV_E	1.445	1.802	2.112	2.472	1.440	1.797	2.108	2.468	1.451	1.808	2.117	2.478
RMSE	0.038	0.035	0.033	0.031	0.044	0.041	0.038	0.035	0.030	0.028	0.026	0.023
	$(100, 100)$				$(100, 200)$				$(200, 100)$			
Sim.	1.465	1.824	2.134	2.498	1.453	1.800	2.104	2.447	1.475	1.842	2.163	2.539
Perm.	1.463	1.819	2.131	2.490	1.450	1.798	2.101	2.450	1.477	1.843	2.164	2.534
CV_A	1.462	1.818	2.127	2.487	1.462	1.818	2.127	2.486	1.462	1.818	2.127	2.486
CV_E	1.452	1.809	2.119	2.479	1.452	1.809	2.118	2.478	1.457	1.813	2.122	2.482
RMSE	0.027	0.025	0.023	0.021	0.027	0.025	0.023	0.021	0.019	0.018	0.016	0.015
	$(500, 500)$				$(400, 800)$				$(800, 400)$			
Sim.	1.459	1.819	2.127	2.492	1.454	1.806	2.111	2.462	1.469	1.827	2.140	2.501
Perm.	1.464	1.821	2.128	2.492	1.456	1.806	2.109	2.465	1.466	1.828	2.144	2.507
CV_A	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486
CV_E	1.459	1.815	2.124	2.484	1.458	1.815	2.124	2.484	1.460	1.816	2.125	2.485
RMSE	0.012	0.011	0.010	0.009	0.014	0.013	0.012	0.011	0.010	0.076	0.072	0.069
	$(1000, 1000)$				$(1000, 2000)$				$(2000, 1000)$			
Sim.	1.458	1.815	2.126	2.483	1.458	1.814	2.121	2.478	1.464	1.821	2.131	2.490
Perm.	1.463	1.819	2.132	2.490	1.458	1.809	2.116	2.472	1.464	1.825	2.138	2.502
CV_A	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486	1.461	1.817	2.126	2.486
CV_E	1.460	1.816	2.125	2.485	1.460	1.816	2.125	2.485	1.460	1.817	2.126	2.485
RMSE	0.009	0.008	0.007	0.007	0.009	0.008	0.007	0.007	0.006	0.006	0.005	0.005

Table 3.3: Critical value of $\max(t, t_R)$ for the standard logistic distribution.

	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990	0.900	0.950	0.975	0.990
CV_L	1.392	1.753	2.066	2.430	1.392	1.753	2.066	2.430	1.392	1.753	2.066	2.430
	$(n_1, n_2) = (20, 20)$				$(20, 40)$				$(40, 20)$			
Sim.	1.416	1.792	2.120	2.539	1.407	1.781	2.111	2.503	1.407	1.775	2.107	2.489
Perm.	1.407	1.781	2.117	2.516	1.407	1.779	2.105	2.493	1.407	1.779	2.107	2.497
CV_A	1.404	1.765	2.077	2.440	1.401	1.761	2.074	2.437	1.401	1.761	2.074	2.437
CV_E	1.369	1.731	2.044	2.408	1.362	1.724	2.037	2.402	1.383	1.744	2.057	2.421
RMSE	0.056	0.055	0.053	0.051	0.061	0.059	0.057	0.055	0.036	0.034	0.033	0.031
	$(50, 50)$				$(40, 80)$				$(80, 40)$			
Sim.	1.401	1.770	2.096	2.474	1.403	1.768	2.089	2.460	1.397	1.763	2.081	2.455
Perm.	1.399	1.767	2.082	2.456	1.397	1.763	2.080	2.461	1.399	1.763	2.083	2.455
CV_A	1.397	1.758	2.071	2.434	1.397	1.757	2.070	2.434	1.397	1.757	2.070	2.434
CV_E	1.382	1.743	2.057	2.421	1.377	1.738	2.051	2.416	1.387	1.748	2.061	2.425
RMSE	0.033	0.032	0.031	0.029	0.040	0.038	0.037	0.035	0.026	0.024	0.023	0.022
	$(100, 100)$				$(100, 200)$				$(200, 100)$			
Sim.	1.398	1.762	2.074	2.446	1.395	1.757	2.071	2.445	1.396	1.758	2.077	2.446
Perm.	1.396	1.763	2.079	2.449	1.394	1.757	2.074	2.440	1.396	1.757	2.069	2.438
CV_A	1.395	1.756	2.068	2.432	1.394	1.755	2.068	2.431	1.394	1.755	2.068	2.431
CV_E	1.387	1.748	2.061	2.425	1.386	1.747	2.060	2.424	1.390	1.751	2.064	2.428
RMSE	0.023	0.022	0.021	0.020	0.023	0.023	0.022	0.020	0.016	0.016	0.015	0.014
	$(500, 500)$				$(400, 800)$				$(800, 400)$			
Sim.	1.392	1.756	2.067	2.433	1.393	1.754	2.068	2.431	1.394	1.754	2.069	2.435
Perm.	1.394	1.755	2.069	2.435	1.394	1.755	2.067	2.431	1.392	1.753	2.066	2.428
CV_A	1.393	1.754	2.066	2.430	1.393	1.753	2.066	2.430	1.393	1.753	2.066	2.430
CV_E	1.391	1.752	2.065	2.429	1.391	1.751	2.064	2.428	1.392	1.752	2.065	2.429
RMSE	0.011	0.010	0.010	0.009	0.012	0.011	0.011	0.010	0.008	0.008	0.008	0.007
	$(1000, 1000)$				$(1000, 2000)$				$(2000, 1000)$			
Sim.	1.391	1.752	2.066	2.425	1.391	1.753	2.064	2.423	1.393	1.752	2.068	2.426
Perm.	1.391	1.755	2.067	2.436	1.393	1.755	2.068	2.434	1.393	1.754	2.069	2.435
CV_A	1.392	1.753	2.066	2.430	1.392	1.753	2.066	2.430	1.392	1.753	2.066	2.430
CV_E	1.392	1.752	2.065	2.429	1.392	1.752	2.065	2.429	1.392	1.753	2.066	2.430
RMSE	0.007	0.007	0.007	0.006	0.007	0.007	0.007	0.006	0.005	0.005	0.005	0.005

Tables 3.1 to 3.3 show that the convergence of the maximum test $\max(t, t_R)$ with equal sample sizes is faster than that of unequal sample sizes. In addition, convergence to the nominal significance level is slow for the tail probability. The simulations reveal that the limiting distribution is necessary when the sample size is larger than 200 for equal sample sizes and 300 for unequal sample sizes.

Moreover, the asymptotic power is compared with the simulated power. The results are shown in Tables 3.4 to 3.6. The simulated powers are obtained by $1,000,000$ iterations of Monte Carlo simulations. In this study, a significance level of 0.05 is assumed and $(n_1, n_2) = (20, 20)$, $(50, 50)$, $(40, 80)$, $(80, 40)$. Now consider the normal distribution $N(0, 1)$ and $N(\theta, 1)$ in Table 3.4, the exponential distribution Exp(1) and Exp(1)$+\theta$ in Table 3.5, and the logistic distribution Logis$(0, 1)$ and Logis$(\theta, 1)$ in Table 3.6, where $\theta = 0, 0.2, 0.4, 0.6, 0.8, 1.0$. In addition, use the 0.95 percentile of the asymptotic null distribution as the critical value. In the tables, Sim.(ρ_0) and Sim.$(\hat{\rho})$ represent the simulated power and the mean of the estimated p-value based on $1,000,000$ iterations of Monte Carlo simulations. In addition, AP is the asymptotic power calculated by $L_1(x)$.

Table 3.4: Asymptotic power of $\max(t, t_R)$ under $N(0, 1)$ and $N(\theta, 1)$.

		θ					
(n_1, n_2)		0	0.2	0.4	0.6	0.8	1.0
$(20, 20)$	Sim.(ρ_0)	0.053	0.159	0.352	0.593	0.803	0.929
	Sim.$(\hat{\rho})$	0.056	0.165	0.361	0.604	0.810	0.932
	AP	0.050	0.157	0.361	0.618	0.832	0.948
$(50, 50)$	Sim.(ρ_0)	0.051	0.259	0.632	0.908	0.990	1.000
	Sim.$(\hat{\rho})$	0.052	0.262	0.637	0.909	0.990	0.999
	AP	0.050	0.261	0.648	0.921	0.993	1.000
$(40, 80)$	Sim.(ρ_0)	0.051	0.268	0.657	0.923	0.993	1.000
	Sim.$(\hat{\rho})$	0.052	0.274	0.662	0.925	0.993	1.000
	AP	0.050	0.272	0.672	0.934	0.995	1.000
$(80, 40)$	Sim.(ρ_0)	0.051	0.269	0.656	0.923	0.993	1.000
	Sim.$(\hat{\rho})$	0.052	0.271	0.659	0.924	0.993	1.000
	AP	0.050	0.272	0.672	0.934	0.995	1.000

Table 3.5: Asymptotic power of $\max(t, t_R)$ under Exp(1) and Exp(1) + θ.

(n_1, n_2)		θ = 0	0.2	0.4	0.6	0.8	1.0
$(20, 20)$	Sim.(ρ_0)	0.051	0.229	0.530	0.785	0.922	0.976
	Sim.$(\hat{\rho})$	0.057	0.242	0.544	0.794	0.926	0.978
	AP	0.050	0.222	0.513	0.767	0.912	0.974
$(50, 50)$	Sim.(ρ_0)	0.050	0.422	0.867	0.987	0.999	1.000
	Sim.$(\hat{\rho})$	0.053	0.427	0.869	0.988	0.999	1.000
	AP	0.050	0.414	0.855	0.984	0.999	1.000
$(40, 80)$	Sim.(ρ_0)	0.064	0.474	0.919	0.997	1.000	1.000
	Sim.$(\hat{\rho})$	0.057	0.448	0.907	0.996	1.000	1.000
	AP	0.050	0.427	0.870	0.987	0.999	1.000
$(80, 40)$	Sim.(ρ_0)	0.046	0.441	0.868	0.985	0.999	1.000
	Sim.$(\hat{\rho})$	0.049	0.445	0.869	0.985	0.999	1.000
	AP	0.050	0.441	0.882	0.990	1.000	1.000

Table 3.6: Asymptotic power of $\max(t, t_R)$ under Logis$(0, 1)$ and Logis$(\theta, 1)$.

(n_1, n_2)		θ = 0	0.2	0.4	0.6	0.8	1.0
$(20, 20)$	Sim.(ρ_0)	0.052	0.103	0.182	0.290	0.421	0.559
	Sim.$(\hat{\rho})$	0.056	0.109	0.190	0.301	0.433	0.572
	AP	0.050	0.100	0.179	0.291	0.428	0.574
$(50, 50)$	Sim.(ρ_0)	0.051	0.143	0.306	0.525	0.733	0.881
	Sim.$(\hat{\rho})$	0.053	0.146	0.312	0.530	0.739	0.885
	AP	0.050	0.141	0.309	0.532	0.745	0.892
$(40, 80)$	Sim.(ρ_0)	0.051	0.147	0.321	0.547	0.758	0.900
	Sim.$(\hat{\rho})$	0.053	0.150	0.327	0.554	0.764	0.903
	AP	0.050	0.146	0.323	0.554	0.768	0.908
$(80, 40)$	Sim.(ρ_0)	0.051	0.146	0.321	0.548	0.759	0.899
	Sim.$(\hat{\rho})$	0.052	0.149	0.323	0.551	0.761	0.901
	AP	0.050	0.146	0.323	0.554	0.768	0.908

Tables 3.4 to 3.6 show that the asymptotic power is roughly the same as the simulated power. In addition, the results demonstrate the consistency of the maximum test. Furthermore, Sim.($\hat{\rho}$) is close to AP and Sim(ρ_0). Thus, the simulations reveal that the maximum statistic can be utilized when the underlying distribution cannot be assumed.

3.4 Real data analysis

We apply our method to reaction time data Sedlmeier and Renkewitz (2008) in a clinical trial similar to Neuhäuser (2015).

Active drug (X):	171	172	178	179	184	185	186	194	196	223
Placebo (Y):	154	155	158	159	161	163	177	183	192	219

At first, the Mood test is used (see, e.g., Gibbons and Chakraborti (2011)) by adjusting the median of each sample, then the p-value of 0.906 is obtained for the homogeneity of variance. It is assumed that the two samples are from two distributions with the same variance. Next, $t = 1.800$ is obtained and the one-sided p-value is 0.045. For the Wilcoxon rank-sum test, $t_R = 2.356$, and the p-value is 0.014, as determined by the exact permutation test. Furthermore, $\bar{X} = 186.8$, $\hat{\sigma} = 15.15$, $\mathrm{E}[\widehat{XF(X)}] = 97.5$, and $\hat{\rho} = 0.915$. The estimated p-value is 0.013 for the maximum test. Thus, the decision is to reject the null hypothesis. Therefore, the p-value can be estimated by using unbiased estimators even when the underlying distribution cannot be assumed. This result is close to the Wilcoxon rank-sum test. The limiting distribution is sometimes not suitable, as in the case of small sample sizes. Then, the exact permutation test is used to compare with the asymptotic result and the p-value of 0.018 is obtained for the data discussed. Although the sample sizes are small, the p-value of the limiting distribution is close to that of the exact permutation.

Chapter 4

Sum of random variables

Determination of the distribution of the sum of independent random variables is one of the most important topics in real data analysis. Gómez et al. (2014) introduced the extended exponential distribution which is useful in fitting real data. In this work, the distribution of the sum of n inid extended exponential random variables are derived. In addition, the extended exponential distribution is interpreted as a special case of the generalized Lindley distribution, which was introduced by Zakerzadeh and Dolati (2009). Therefore, the exact probability density function and cumulative distribution function of the sum of inid generalized Lindley random variables is also derived, and the sum of some inid conventional random variables is shown to be a special case of the generalized Lindley distribution.

4.1 Extended exponential random variables

The exponential distribution is widely used and is the most significant distribution in statistical analysis. For instance, the exponential distribution is used to model waiting times of services, the time interval between phone calls, and the lifetime of a machine. Many researchers have discussed extensions of the exponential distribution. For example, Gupta and Kundu (2001) introduced an extended exponential distribution, such as

$$f_{\mathrm{GK}}(x;\alpha,\lambda)=\alpha\lambda(1-e^{-\lambda x})^{\alpha-1}e^{-\lambda x},\quad x>0,$$

for $\alpha>0,\lambda>0$. In addition, Nadarajah and Haghighi (2011) introduced another extension of the exponential distribution with a density function given by,

$$f_{\mathrm{NH}}(x;\alpha,\lambda)=\alpha\lambda(1+\lambda x)^{\alpha-1}\exp\{1-(1+\lambda x)^{\alpha}\},\quad x>0,$$

where $\alpha>0$ and $\lambda>0$. Recently, Lemonte and Moreno-Arenas (2016) proposed a three parameter extension of the exponential distribution. More recently, Almarashi et al. (2019) extended the exponential distribution by using the type I half-logistic family of distributions.

An extension of the exponential distribution based on mixtures of positive distributions is introduced by Gómez et al. (2014). A random variable X is distributed according to the extended exponential distribution with parameters α and β with the probability density function and cumulative distribution function given by

$$f_{\mathrm{EE}}(x;\alpha,\beta) = \frac{\alpha^2(1+\beta x)e^{-\alpha x}}{\alpha+\beta},$$
$$F_{\mathrm{EE}}(x;\alpha,\beta) = \frac{\alpha+\beta-(\beta+\alpha+\alpha\beta x)e^{-\alpha x}}{\alpha+\beta},$$

respectively, where $x > 0$, $\alpha > 0$, and $\beta \geq 0$. Then, the moment generating function is given by

$$M_{\mathrm{EE}}(t) = \frac{\alpha^2(\alpha+\beta-t)}{(\alpha+\beta)(t-\alpha)^2}, \quad t < \alpha. \tag{4.1}$$

Gómez et al. (2014) showed that the extended exponential distribution is more useful in fitting real data than other extensions for the life of the fatigue fracture of Kevlar 49/epoxy. In this study, the exact probability density function and cumulative distribution function of the sum of inid extended exponential random variables are derived. However, the probability density function and cumulative distribution function contain an infinite sum of gamma series. Then, the saddlepoint and the normal approximations are applied to overcome the computational complexity. The accuracy of the saddlepoint approximation to the sum of inid extended exponential random variables is further compared with the exact distribution function through numerical studies. In addition, parameter estimation by the maximum likelihood method is discussed for the case of $n = 2$ and real data analysis.

4.2 Sum of inid extended exponential random variables

In this section, the exact distribution of the sum of the inid extended exponential random variables is derived. Let $X_1, \ldots, X_n$ be independent, extended exponential random variables with parameters $\alpha_i > 0$, $\beta_i \geq 0$ for $i = 1, \ldots, n$. By Equation (4.1), the moment generating function of $S = X_1 + X_2 + \cdots + X_n$ is

$$\begin{aligned} M_{\mathrm{SEE}}(t) &= \prod_{i=1}^{n} \frac{\alpha_i^2(\alpha_i+\beta_i-t)}{(\alpha_i+\beta_i)(t-\alpha_i)^2} \\ &= \left(\prod_{i=1}^{n} \frac{\alpha_i^2}{\alpha_i+\beta_i}\right) \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{(\alpha_i-t)^{(\boldsymbol{\tau})_i+1}} \\ &= \left(\prod_{i=1}^{n} \frac{\alpha_i^2}{\alpha_i+\beta_i}\right) \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} \left(\prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_i^{(\boldsymbol{\tau})_i+1}}\right) \prod_{i=1}^{n} \left(1-\frac{t}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i-1}, \end{aligned} \tag{4.2}$$

where $\mathbb{B} = \{0, 1\}$, and $(\boldsymbol{\tau})_i$ is the ith component of $\boldsymbol{\tau}$.

Herein, let

$$h(t) = \prod_{i=1}^{n} \left(1-\frac{t}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i-1}.$$

The inverse transformation of the moment generating function is applicable to $h(t)$. Then we obtain Theorem 4.1.

Theorem 4.1. *The probability density function of S is expressed as*

$$f_{\mathrm{SEE}}(x) = C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}+k)} x^{\rho_{\boldsymbol{\tau}}+k-1} e^{-\alpha_1 x}, \quad x > 0, \tag{4.3}$$

where

$$C = \prod_{i=1}^{n} \frac{\alpha_i^2}{\alpha_i + \beta_i}, \quad D_{\boldsymbol{\tau}} = \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_i^{(\boldsymbol{\tau})_i+1}}, \quad \rho_{\boldsymbol{\tau}} = \sum_{i=1}^{n} \{(\boldsymbol{\tau})_i + 1\},$$

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j \eta_j \delta_{k+1-j}, \ k = 0, 1, 2, \ldots, \quad \delta_0 = 1,$$

$$\eta_j = \sum_{s=1}^{n} \{(\boldsymbol{\tau})_s + 1\} \left(1 - \frac{\alpha_s}{\alpha_1}\right)^j / j, \quad \alpha_1 = \max_i(\alpha_i).$$

Proof. A procedure similar to Moschopoulos (1985) is followed to prove the theorem. Without loss of generality, assume $\alpha_1 = \max_i(\alpha_i)$. Then apply the identity

$$1 - \frac{t}{\alpha_i} = \left(1 - \frac{t}{\alpha_1}\right) \frac{\alpha_1}{\alpha_i} \left[1 - \frac{1 - \frac{\alpha_i}{\alpha_1}}{1 - \frac{t}{\alpha_1}}\right]$$

to $h(t)$, and then

$$\begin{aligned}
\log h(t) &= \log \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i - 1} \\
&= \log \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_1}\right)^{-(\boldsymbol{\tau})_i - 1} \left(\frac{\alpha_1}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i - 1} \left[1 - \frac{1 - \frac{\alpha_i}{\alpha_1}}{1 - \frac{t}{\alpha_1}}\right]^{-(\boldsymbol{\tau})_i - 1} \\
&= \log \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_1}\right)^{-(\boldsymbol{\tau})_i - 1} \left(\frac{\alpha_1}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i - 1} + \sum_{i=1}^{n} \log \left[1 - \frac{1 - \frac{\alpha_i}{\alpha_1}}{1 - \frac{t}{\alpha_1}}\right]^{-(\boldsymbol{\tau})_i - 1}.
\end{aligned}$$

By using the Maclaurin expansion $\log(1-x) = -\sum_{j=1}^{\infty} \frac{x^j}{j}$, we have

$$\begin{aligned}
\log h(t) &= \log \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_1}\right)^{-(\boldsymbol{\tau})_i - 1} \left(\frac{\alpha_1}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i - 1} + \sum_{i=1}^{n} (-(\boldsymbol{\tau})_i - 1) \left[-\sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{1 - \frac{\alpha_i}{\alpha_1}}{1 - \frac{t}{\alpha_1}}\right)^j\right] \\
&= \log \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_1}\right)^{-(\boldsymbol{\tau})_i - 1} \left(\frac{\alpha_1}{\alpha_i}\right)^{-(\boldsymbol{\tau})_i - 1} + \sum_{j=1}^{\infty} \left\{\frac{1}{j} \sum_{i=1}^{n} ((\boldsymbol{\tau})_i + 1) \left(1 - \frac{\alpha_i}{\alpha_1}\right)^j\right\} \left(1 - \frac{t}{\alpha_1}\right)^{-j} \\
&= \log \left[\prod_{i=1}^{n} \left(\frac{\alpha_i}{\alpha_1}\right)^{(\boldsymbol{\tau})_i + 1} \left(1 - \frac{t}{\alpha_1}\right)^{-\rho_{\boldsymbol{\tau}}}\right] + \sum_{j=1}^{\infty} \eta_j \left(1 - \frac{t}{\alpha_1}\right)^{-j},
\end{aligned}$$

where

$$\rho_{\boldsymbol{\tau}} = \sum_{i=1}^{n} \{(\boldsymbol{\tau})_i + 1\}, \quad \eta_j = \frac{1}{j} \sum_{i=1}^{n} \{(\boldsymbol{\tau})_i + 1\} \left(1 - \frac{\alpha_i}{\alpha_1}\right)^j.$$

This expression is defined by t such that $\max_i |(1 - \alpha_i/\alpha_1)/(1 - t/\alpha_1)| < 1$. Therefore,

$$h(t) = \prod_{i=1}^{n} \left(\frac{\alpha_i}{\alpha_1}\right)^{(\boldsymbol{\tau})_i + 1} \left(1 - \frac{t}{\alpha_1}\right)^{-\rho_{\boldsymbol{\tau}}} \exp\left(\sum_{j=1}^{\infty} \eta_j \left(1 - \frac{t}{\alpha_1}\right)^{-j}\right).$$

Herein, the terms of the same order in the Taylor series are calculated together to obtain,

$$\exp\left(\sum_{j=1}^{\infty} \eta_j \left(1 - \frac{t}{\alpha_1}\right)^{-j}\right) = \sum_{k=0}^{\infty} \delta_k \left(1 - \frac{t}{\alpha_1}\right)^{-k}.$$

The coefficient δ_k is obtained by the recursive formula,

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j \eta_j \delta_{k+1-j}, \ k = 0, 1, 2, \ldots,$$

with $\delta_0 = 1$. Thus, the moment generating function of Y is

$$M_{\mathrm{SEE}}(t) = \left(\prod_{i=1}^{n} \frac{\alpha_i^2}{\alpha_i + \beta_i}\right) \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_1^{(\boldsymbol{\tau})_i + 1}} \sum_{k=0}^{\infty} \delta_k \left(1 - \frac{t}{\alpha_1}\right)^{-(\rho_{\boldsymbol{\tau}} + k)}.$$

Remark that $\left(1 - \frac{t}{\alpha_1}\right)^{-(\rho_{\boldsymbol{\tau}} + k)}$ is the same as the moment generating function of the gamma distribution. Then, we apply the inverse transformation of the moment generating function term-by-term. Therefore, the theorem is completely proved. □

Theorem 4.2. *The exact cumulative distribution function $F_{\mathrm{SEE}}(y) = P(S \leq y)$ is derived by term-by-term integration of* (4.3)*, that is,*

$$F_{\mathrm{SEE}}(x) = C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \delta_k \int_0^x \frac{\alpha_1^{\rho_{\boldsymbol{\tau}} + k}}{\Gamma(\rho_{\boldsymbol{\tau}} + k)} w^{\rho_{\boldsymbol{\tau}} + k - 1} e^{-\alpha_1 w} dw \tag{4.4}$$

$$= C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho_{\boldsymbol{\tau}} + k}}{\Gamma(\rho_{\boldsymbol{\tau}} + k + 1)} e^{-\alpha_1 x} {}_1F_1(1; \rho_{\boldsymbol{\tau}} + k + 1; \alpha_1 x), \tag{4.5}$$

where ${}_1F_1(a; b; z)$ is the confluent hypergeometric function with the integral formula of

$${}_1F_1(a; b; z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{zw} w^{a-1} (1-w)^{b-a-1} dw, \quad 0 < a < b.$$

In addition, the truncation error is obtained by

$$E_m(x) = C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} \int_0^x w^{\rho_{\boldsymbol{\tau}} - 1} \exp\{-(1-a)\alpha_1 w\} dw - F_m(x),$$

where $F_m(x)$ is the sum of the first $m + 1$ terms of (4.5) *for $k = 0, 1, \ldots, m$.*

Proof. The interchange of the integration and summation in $F_{\mathrm{SEE}}(y)$ can be justified from the uniform convergence. For $i = 1, 2, \dots$ and $a = \max_{2\le\ell\le n}(1 - \alpha_\ell/\alpha_1)$, to obain,

$$|\eta_j| = \sum_{s=1}^{n} \frac{((\boldsymbol{\tau})_s + 1)(1 - \alpha_s/\alpha_1)^j}{j} \le \frac{\rho_{\boldsymbol{\tau}} a^j}{j},\ j = 1, 2, \dots, k+1.$$

From the definition of δ,

$$|\delta_{k+1}| \le \frac{\rho_{\boldsymbol{\tau}}}{k+1} \sum_{j=1}^{k+1} a^j |\delta_{k+1-j}|, \quad k = 0, 1, 2, \dots,$$

from the recursive equation,

$$|\delta_{k+1}| \le \frac{\rho_{\boldsymbol{\tau}}(\rho_{\boldsymbol{\tau}} + 1)\cdots(\rho_{\boldsymbol{\tau}} + k)}{(k+1)!} a^{k+1}.$$

Therefore,

$$\begin{aligned} f_{\mathrm{SEE}}(x) &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} \exp(-\alpha_1 x) \sum_{k=0}^{\infty} \frac{\delta_k}{\rho_{\boldsymbol{\tau}}(\rho_{\boldsymbol{\tau}} + 1)\cdots(\rho_{\boldsymbol{\tau}} + k - 1)} (\alpha_1 x)^k \\ &\le C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} \exp(-\alpha_1 x) \sum_{k=0}^{\infty} \frac{(\alpha_1 a x)^k}{k!} \\ &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} \exp\{-(1-a)\alpha_1 x\}. \end{aligned} \tag{4.6}$$

Here, (4.6) shows the uniform convergence of (4.3), and then we have (4.5). □

4.3 Numerical results

In the previous section, we derived the exact distribution of the sum of inid extended exponential random variables. However, it is difficult to calculate the exact probability when the number of random variables increases. Hence, a more accurate approximation of the distribution is needed. Under these circumstances, the saddlepoint approximation is commonly used. In literature, many researchers have considered applying the saddlepoint approximation to the distribution of the sum of inid random variables. For example, Eisinga et al. (2013) discussed the use of the saddlepoint approximation for the sum of the inid binomial random variables; Murakami (2014) and Nadarajah et al. (2015) applied the saddlepoint approximation to the sum of the inid uniform and beta random variables, respectively, while Murakami (2015) gave the approximation for the sum of the inid gamma random variables. In this study, the evaluation of the tail probability is discussed using the saddlepoint approximation. Furthermore, the parameters are estimated by the maximum likelihood method and applied to real data analysis.

4.3.1 Saddlepoint approximation

Herein, we consider an approximation for the distribution of S. The cumulant generating function of S is given by

$$k(t) = \sum_{i=1}^{n} \log\left(\frac{\alpha_i^2(\alpha_i+\beta_i-t)}{(\alpha_i+\beta_i)(t-\alpha_i)^2}\right).$$

Lugannani and Rice (1980) provided the formula for approximating the distribution function as follows:

$$F_{\text{SA}}(x) = \Phi(\hat{w}) + \phi(\hat{w})\left(\frac{1}{\hat{w}} - \frac{1}{\hat{u}}\right) + O(n^{-\frac{3}{2}}), \tag{4.7}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal probability density function and the corresponding cumulative distribution function, respectively. In addition, we denote

$$\hat{w} = \text{sgn}(\hat{s})\sqrt{2\{\hat{s}x - k(\hat{s})\}}, \quad \hat{u} = \hat{s}\sqrt{k''(\hat{s})},$$

where $\hat{s}$ is the root of $k'(s) = x$, which is solved numerically by the Newton-Raphson algorithm; $\text{sgn}(\hat{s}) = \pm 1, 0$ if $\hat{s}$ is positive, negative, or zero; and

$$k'(t) = \sum_{i=1}^{n} \frac{t-\alpha_i-2\beta_1}{(t-\alpha_i)(\alpha_i+\beta_i-t)},$$
$$k''(t) = \sum_{i=1}^{n} \left(\frac{2}{(t-\alpha_i)^2} - \frac{1}{(\alpha_i+\beta_i-t)^2}\right).$$

Herein, the accuracy of approximation is compared with the following distributions by calculating the probability. $\hat{s}$ is a percentile derived from $100,000,000$ random numbers generated by S, and p is the exact probability. Note that X is generated as a mixture distribution of two random variables (Gómez et al., 2014). More precisely, they are the exponential and gamma distributions. Then, a random number S is obtained by the sum of n random numbers X.

- F_m: The approximate cumulative distribution function, which is truncated in the infinite series in (4.5) after $m+1$ terms.
- F_{NA}: The normal approximation.
- F_{SA}: The saddlepoint cumulative distribution function from (4.7).

In the tables, Conv. represents the exact probability calculated by convolution, r.e. is the relative error between the approximation and p, and MCT is the mean calculating time. In our study, Mathematica version 11 (CPU 2.80 GHz and 32.0 GB RAM) is used. The parameters $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_n)$ are generated from the uniform distribution $\text{Unif}(0, 3)$, as in Murakami (2014) and Nadarajah et al. (2015):

Case 1: $n = 2$

Case 1-A: $\boldsymbol{\alpha} = (2.307, 1.438)$, $\boldsymbol{\beta} = (2.138, 2.694)$.

Case 1-B: $\boldsymbol{\alpha} = (2.767, 0.301)$, $\boldsymbol{\beta} = (0.659, 1.709)$.

Case 2: $n = 5$

Case 2-A: $\boldsymbol{\alpha} = (1.792, 1.522, 0.828, 1.690, 1.189)$,
$\boldsymbol{\beta} = (0.628, 2.730, 0.796, 1.556, 1.226)$.

Case 2-B: $\boldsymbol{\alpha} = (0.628, 2.730, 07.96, 1.556, 1.226)$,
$\boldsymbol{\beta} = (0.942, 0.867, 1.438, 0.631, 2.014)$.

Case 3: $n = 10$

Case 3-A: $\boldsymbol{\alpha} = (1.820, 1.537, 0.887, 1.707, 0.867, 1.857, 1.308, 0.165, 1.405, 1.120)$,
$\boldsymbol{\beta} = (1.172, 0.331, 0.812, 1.107, 0.575, 1.634, 0.412, 0.151, 1.834, 1.377)$.

Case 3-B: $\boldsymbol{\alpha} = (1.366, 2.631, 2.944, 0.106, 1.769, 2.205, 2.227, 2.347, 1.911, 1.761)$,
$\boldsymbol{\beta} = (2.620, 0.210, 1.000, 0.473, 2.635, 1.267, 1.255, 0.7433, 0.103, 1.609)$.

Table 4.1: Numerical results for $n = 2$.

$\hat{s}$	Conv.	p	F_m	F_{SA}	F_{NA}	r.e. F_{SA}	r.e. F_{NA}
Case 1-A							
1.8547	0.6000	0.6000	0.6000	0.6004	0.5229	0.0007	0.1286
2.1821	0.7000	0.7000	0.7000	0.7005	0.6377	0.0007	0.0891
2.6082	0.8000	0.8000	0.8000	0.8005	0.7691	0.0006	0.0386
3.2795	0.9000	0.9000	0.9000	0.9003	0.9099	0.0003	0.0110
3.9072	0.9500	0.9500	0.9500	0.9502	0.9716	0.0002	0.0228
4.5079	0.9750	0.9750	0.9750	0.9751	0.9928	0.0001	0.0182
5.2739	0.9900	0.9900	0.9900	0.9900	0.9991	0.0000	0.0092
MCT(sec.)		64.1	30.9	0.11	0.0		
Case 1-B							
6.6464	0.6000	0.6000	0.6000	0.6000	0.5058	0.0001	0.1570
8.0298	0.7000	0.7000	0.7000	0.6999	0.6215	0.0002	0.1121
9.8729	0.8000	0.8000	0.8000	0.7998	0.7588	0.0002	0.0515
12.8465	0.9000	0.9000	0.9000	0.8998	0.9093	0.0002	0.0103
15.6850	0.9500	0.9500	0.9500	0.9499	0.9739	0.0001	0.0252
18.4362	0.9750	0.9750	0.9750	0.9749	0.9943	0.0001	0.0198
21.9809	0.9900	0.9900	0.9900	0.9900	0.9995	0.0000	0.0096
MCT(sec.)		63.0	28.6	0.11	0.00		

Determining the appropriate value of m is difficult. Therefore, percentiles of F_m were compared for various m with the exact probability calculated by convolution. Then, it was determined that

Table 4.2: Numerical results for $n = 5$.

$\hat{s}$	p	F_m	F_{SA}	F_{NA}	r.e. F_{SA}	r.e. F_{NA}
Case 2-A						
6.0019	0.6000	0.6000	0.6007	0.5464	0.0011	0.0893
6.6914	0.7000	0.7000	0.7007	0.6579	0.0010	0.0601
7.5633	0.8000	0.8000	0.8007	0.7804	0.0008	0.0245
8.8929	0.9000	0.9000	0.9005	0.9088	0.0005	0.0097
10.1016	0.9500	0.9500	0.9503	0.9673	0.0003	0.0182
11.2365	0.9750	0.9750	0.9752	0.9898	0.0002	0.0152
12.6623	0.9900	0.9900	0.9901	0.9982	0.0001	0.0083
MCT(sec.)	106	511	0.16	0.00		
Case 2-B						
5.4987	0.6000	0.6001	0.6014	0.5308	0.0023	0.1153
6.2522	0.7000	0.7000	0.7012	0.6435	0.0017	0.0807
7.2267	0.8000	0.8000	0.8010	0.7714	0.0012	0.0357
8.7527	0.9000	0.9000	0.9005	0.9085	0.0005	0.0095
10.1783	0.9500	0.9500	0.9502	0.9700	0.0002	0.0211
11.5417	0.9750	0.9750	0.9751	0.9920	0.0001	0.0174
13.2824	0.9900	0.9900	0.9900	0.9990	0.0000	0.0090
MCT(sec.)	102	506	0.17	0.00		

$m = 1000$ for simulating F_m. There is no difference between the Conv. and p, as shown in Table 4.1. Hence, the simulated p is used as the exact probability for Case 2 ($n = 5$) and Case 3 ($n = 10$).

Tables 4.1 to 4.3 show that F_m is closer to p than any other approximation. In Kitani and Murakami (2020b) it took more than 100 hours to calculate F_m using Equation (4.4) for Case 3. However, the value of F_m can be obtained using Equation (4.5) in numerical calculations. Nonetheless, it is difficult to apply real data analysis because it takes a long time to calculate the probability and F_m does not work well in Case 3-B. On the other hand, F_{NA} and F_{sa} overcome the problem of calculation time; in particular, F_S gives better accuracy than F_N.

4.3.2 Parameter estimation

In this section, parameter estimation for the sum of extended exponential distribution for the case of $n = 2$ is discussed. After generating r random numbers which follow the extended exponential distribution, the parameters were estimated using the maximum likelihood estimator. In Table 4.4, the mean and the variance of estimated parameters are shown, along with the first four moments of $S = X_1 + X_2$ based on 1,000 simulations.

Table 4.3: Numerical results for $n = 10$.

$\hat{s}$	p	F_m	F_{SA}	F_{NA}	r.e. F_{SA}	r.e. F_{NA}
Case 3-A						
19.0493	0.6000	0.6001	0.6004	0.5106	0.0006	0.1489
21.5079	0.7000	0.7001	0.6996	0.6233	0.0005	0.1095
24.7999	0.8000	0.8001	0.7994	0.7578	0.0008	0.0528
30.1437	0.9000	0.9000	0.8995	0.9073	0.0006	0.0081
35.2711	0.9500	0.9500	0.9497	0.9728	0.0003	0.0240
40.2577	0.9750	0.9750	0.9749	0.9939	0.0001	0.0194
46.6919	0.9900	0.9900	0.9899	0.9994	0.0001	0.0095
MCT(sec.)	177	29446	0.25	0.02		
Case 3-B						
23.4716	0.6000	0.6000	0.5999	0.5069	0.0001	0.1551
27.4031	0.7000	0.7000	0.6998	0.6223	0.0003	0.1110
32.6411	0.8000	0.8000	0.7998	0.7591	0.0002	0.0512
41.0841	0.9000	0.9000	0.8998	0.9090	0.0002	0.0100
49.1511	0.9500	0.9471	0.9499	0.9737	0.0001	0.0250
56.9604	0.9750	0.9517	0.9749	0.9942	0.0001	0.0197
67.0207	0.9900	0.9425	0.9900	0.9995	0.0000	0.0096
MCT(sec.)	171	28280	0.25	0.02		

Table 4.4: The mean (the variance) of numerical simulation based on 1000 iterations and the first four moments.

		r					r		
	True	100	500	2000		True	100	500	2000
$\hat{\alpha}_1$	2.2	1.99 (2.21)	2.07 (0.76)	2.12 (0.64)	E$[S]$	1.67	1.74	1.70	1.68
$\hat{\alpha}_2$	1.4	1.45 (0.07)	1.39 (0.04)	1.38 (0.03)	E$[S^2]$	3.99	4.21	4.06	4.01
$\hat{\beta}_1$	2.6	4.28 (18.4)	4.28 (13.9)	3.60 (5.34)	E$[S^3]$	12.27	13.04	12.48	12.34
$\hat{\beta}_2$	0.8	0.60 (3.07)	0.43 (1.62)	0.49 (0.89)	E$[S^4]$	46.26	49.08	46.98	46.57

In Table 4.4, the estimated parameters get closer to the true parameters as sample size r increases. Parameters $\hat{\beta}_1$ and $\hat{\beta}_2$ are different from true parameters β_1 and β_2; however, there is almost no difference in the moments. The simulation results show that parameter estimation works well. Nevertheless, since the variance is large when the sample size is small, the initial value problem of parameter estimation is considered. It will be necessary to discuss the identifiability of the parameters in the future.

4.3.3 Real data analysis

The Akaike information criterion (AIC) is compared with the extended exponential distribution of Gómez et al. (2014) (i.e., the case of $n = 1$, EE1) and f_Y with $n = 2, 3$ (EE2, EE3), Gamma distribution Gamma(α_1, α_2), and Weibull distribution Weibull(α_1, α_2) based on the maximum likelihood approach.

Two data sets are considered for the life of the fatigue fracture of Kevlar 49/epoxy, which is a widely used synthetic fiber, given in Glaser (1983) as follows:

Dataset 1:
0.0251, 0.0886, 0.0891, 0.2501, 0.3113, 0.3451, 0.4763, 0.5650, 0.5671, 0.6566, 0.6748, 0.6751, 0.6753, 0.7696, 0.8375, 0.8391, 0.8425, 0.8645, 0.8851, 0.9113, 0.9120, 0.9836, 1.0483, 1.0596, 1.0773, 1.1733, 1.2570, 1.2766, 1.2985, 1.3211, 1.3503, 1.3551, 1.4595, 1.4880, 1.5728, 1.5733, 1.7083, 1.7263, 1.7460, 1.7630, 1.7746, 1.8275, 1.8375, 1.8503, 1.8808, 1.8878, 1.8881, 1.9316, 1.9558, 2.0048, 2.0408, 2.0903, 2.1093, 2.1330, 2.2100, 2.2460, 2.2878, 2.3203, 2.3470, 2.3513, 2.4951, 2.5260, 2.9911, 3.0256, 3.2678, 3.4045, 3.4846, 3.7433, 3.7455, 3.9143, 4.8073, 5.4005, 5.4435, 5.5295, 6.5541, and 9.0960.

Dataset 2:
0.7367, 1.1627, 1.8945, 1.9340, 2.3180, 2.6483, 2.8573, 2.9918, 3.0797, 3.1152, 3.1335, 3.2647, 3.4873, 3.5390, 3.6335, 3.6541, 3.7645, 3.8196, 3.8520, 3.9653, 4.2488, 4.3017, 4.3942, 4.6416, 4.7070, 4.8885, 5.1746, 5.4962, 5.5310, 5.5588, 5.6333, 5.7006, 5.8730, 5.8737, 5.9378, 6.1960, 6.2217, 6.2630, 6.3163, 6.4513, 6.8320, 6.9447, 7.2595, 7.3183, 7.3313, 7.7587, 8.0393, 8.0693, 8.1928, 8.4166, 8.7558, 8.8398, 9.2497, 9.2563, 9.5418, 9.6472, 9.6902, 9.9316, 10.018, 10.4028 , 10.4188, 10.7250, 10.9411, 11.7962, 12.075, 12.6933, 13.5307, 13.8105, 14.5067, 15.3013, 16.2742, 18.2682, and 19.2033.

In addition, we consider another type of data set consisting of the waiting times between 65 consecutive eruptions of the Kiama Blowhole. These data are available at `http://www.statsci.org/data/oz/kiama.html`.

Dataset 3:
83, 51, 87, 60, 28, 95, 8, 27, 15, 10, 18, 16, 29, 54, 91, 8, 17, 55, 10, 35,47, 77, 36, 17, 21, 36, 18, 40, 10, 7, 34, 27, 28, 56, 8, 25, 68, 146, 89, 18, 73, 69, 9, 37, 10, 82, 29, 8, 60, 61, 61, 18, 169, 25, 8, 26, 11, 83, 11, 42, 17, 14, 9 and 12.

Results of the parameter estimation of various models are shown in Table 4.5.

Table 4.5: Parameter estimates for various models.

Distribution	$\hat{\alpha_1}$	$\hat{\beta_1}$	$\hat{\alpha_2}$	$\hat{\beta_2}$	$\hat{\alpha_3}$	$\hat{\beta_3}$	AIC
Dataset 1							
EE1	0.954	6.365	-	-	-	-	247.3
EE2	0.709	0.000	2.296	0.817	-	-	253.2
EE3	156.3	1.320	99.54	118.6	1.007	15.91	256.8
Gamma	1.641	0.838	-	-	-	-	248.5
Weibull	1.326	2.133	-	-	-	-	249.0
Dataset 2							
EE1	0.281	835.7	-	-	-	-	405.1
EE2	0.474	10.77	0.476	0.348	-	-	402.5
EE3	1.898	0.498	0.462	1.064	0.465	0.201	406.5
Gamma	3.071	0.432	-	-	-	-	398.5
Weibull	1.877	8.039	-	-	-	-	400.0
Dataset 3							
EE1	0.050	7.038	-	-	-	-	597.6
EE2	0.335	2.5×10^6	0.032	0.003	-	-	594.8
EE3	0.715	4.125	0.030	0.001	0.519	4.823	596.9
Gamma	1.621	0.041	-	-	-	-	595.8
Weibull	1.274	43.21	-	-	-	-	597.8

Figures 4.1 to 4.3 display the data fitting for various models.

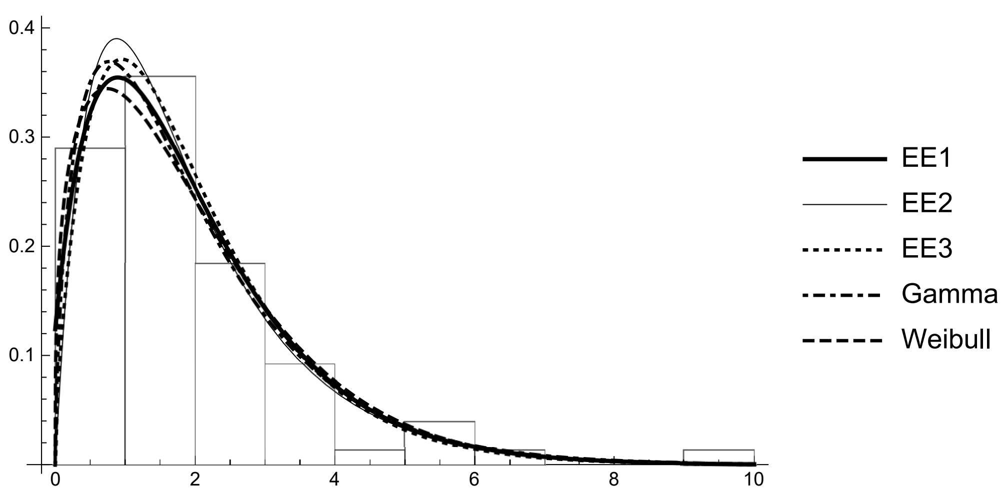

Figure 4.1: Fitted probability density functions of the distribution for Dataset 1.

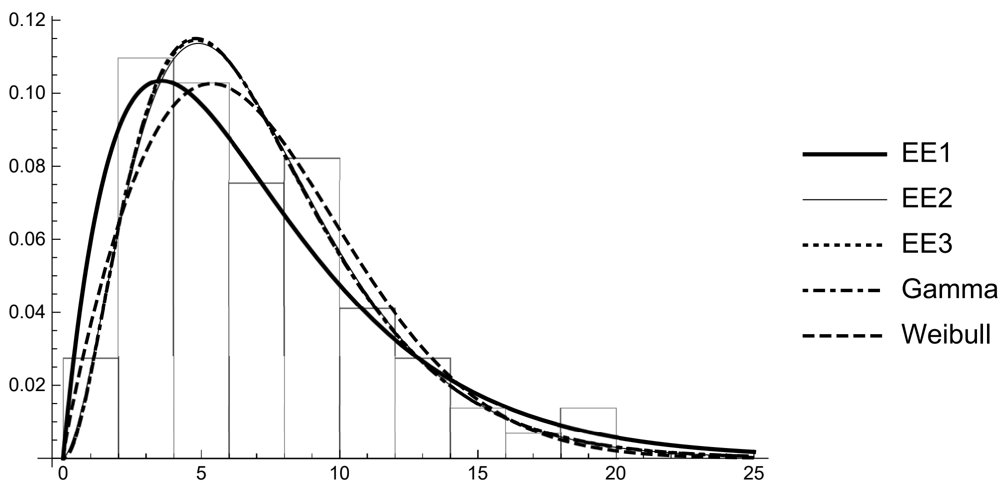

Figure 4.2: Fitted probability density functions of the distribution for Dataset 2.

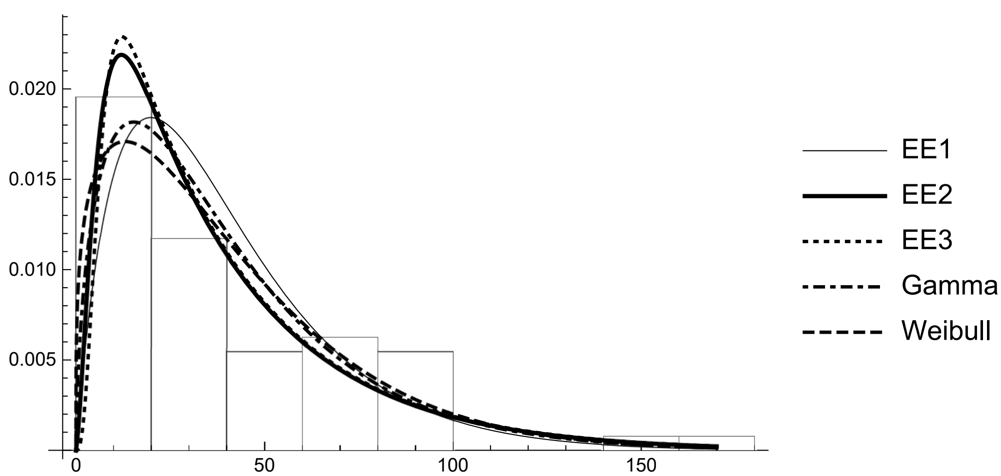

Figure 4.3: Fitted probability density functions of the distribution for Dataset 3.

In Dataset 1, as in Gómez et al. (2014), AIC indicates that EE1 is the most suitable model. In another example, EE2 is more suitable than EE1 while the gamma model is a better fit in Dataset2. However, EE2 is the best model for Dataset3. As a whole, the sum of the extended exponential models have versatility for these data sets.

4.4 Generalized Lindley random variables

The extended exponential distribution can be expressed as a mixture of exponential and gamma distributions as follows:

$$f_{\mathrm{EE}}(x;\alpha,\beta)=\frac{\alpha}{\alpha+\beta}f_{\mathrm{E}}(x;\alpha)+\frac{\beta}{\alpha+\beta}f_{\mathrm{G}}(x;2,\alpha),\quad x>0,$$

where $\alpha>0$, $\beta\geq 0$, and $f_{\mathrm{E}}(x;\alpha)$ and $f_{\mathrm{G}}(x;2,\alpha)$ are the probability density function of the exponential distribution $\mathrm{Exp}(\alpha)$ and gamma distribution $\mathrm{Gamma}(2,\alpha)$, respectively. Then, the extended exponential distribution is also the extension of the Lindley distribution, which is proposed by Lindley (1958). The probability density function of the Lindley distribution is given by

$$\begin{aligned} f_{\mathrm{L}}(x;\alpha)&=\frac{\alpha}{1+\alpha}f_{\mathrm{E}}(x;\alpha)+\frac{1}{1+\alpha}f_{\mathrm{G}}(x;2,\alpha)\\ &=\frac{\alpha^2}{1+\alpha}(x+1)e^{-\alpha x},\quad x>0, \end{aligned}$$

where $\alpha>0$. Ghitany et al. (2008) discussed various properties of this distribution and demonstrated its importance for modeling various sets of lifetime data and reliability modeling. Ghitany et al. (2008) also suggested that many situations exist in which the Lindley distribution is a better model than the exponential distribution. In the literature, many authors have extended the Lindley distribution. For example, Shanker and Mishra (2013) proposed a quasi-Lindley distribution with probability density function

$$f_{\mathrm{QL}}(x;\alpha,\lambda)=\frac{\alpha(\lambda+\alpha x)}{1+\lambda}e^{-\alpha x},\quad x>0,$$

where $\alpha>0$, $\lambda>-1$. The quasi-Lindley distribution is defined by a mixture of $\mathrm{Exp}(\alpha)$ and $\mathrm{Gamma}(2,\alpha)$, where the mixing probabilities are $\lambda/(1+\lambda)$ and $1/(1+\lambda)$, respectively. Another two-parameter Lindley distribution was introduced by Shanker et al. (2013), which is the same as the extended exponential distribution proposed by Gómez et al. (2014). Note that the extended exponential distribution is reduced to the original Lindley distribution when $\beta=1$.

For three-parameter extension, Abd El-Monsef (2016) proposed the three-parameter Lindley distribution by adding the location parameter with probability density function

$$f_{3\mathrm{L}}(x;\alpha,\beta,\theta)=\frac{\alpha^2}{\alpha+\beta}\{1+\beta(x-\theta)\}e^{-\alpha(x-\theta)},\quad x>\theta\geq 0,$$

where $\alpha > 0$, $\beta \geq 0$. Moreover, Zakerzadeh and Dolati (2009) introduced a further generalization, which is the generalized Lindley distribution, with probability density function

$$f_{\mathrm{GL}}(x; \alpha, \beta, \gamma) = \frac{\alpha^2(\alpha x)^{\gamma-1}(\gamma + \beta x)e^{-\alpha x}}{(\alpha+\beta)\Gamma(\gamma+1)}, \quad x > 0,$$

where $\alpha, \gamma > 0$, and $\beta \geq 0$. Zakerzadeh and Dolati (2009) showed that the generalized Lindley distribution is flexible and a better fitting model than the gamma or Weibull distributions for the failure time of electronic components and the number of cycles until the specimen breaks. The generalized Lindley distribution is a mixture of Gamma(γ, α) and Gamma$(\gamma+1, \alpha)$ with probabilities $\alpha/(\alpha+\beta)$ and $\beta/(\alpha+\beta)$. Note that the generalized Lindley distribution is reduced to the gamma distribution with parameter (γ, α) and the exponential distribution with parameter α when $\beta = 0$ and $\beta = 0$, $\gamma = 1$, respectively. Also note that this distribution is reduced to the extended exponential distribution when $\gamma = 1$ and the original Lindley distribution when $\beta = 1$ and $\gamma = 1$. Furthermore, when $\beta = 1/\alpha$ and $\gamma = 1$, the distribution is reduced to the Shanker distributuon introduced by Shanker (2015). Thus, the generalized Lindley distribution is an extension of the exponential and gamma distributions. Zakerzadeh and Dolati (2009) derived the distribution of the sum of independent random variables. Let $X_1, \ldots, X_n$ denote the independent variables for the generalized Lindley distribution with parameters $(\alpha, \beta, \gamma_i)$, where $\alpha, \gamma_i > 0$, $\beta \geq 0$ for $i = 1, \ldots, n$. Then, the probability density function of $S = X_1 + \cdots + X_n$ is given by

$$f_S(x) = \sum_{k=0}^{n} p_k f_{\mathrm{G}}(x; \gamma^* + k, \beta),$$

where

$$\gamma^* = \sum_{i=1}^{n} \gamma_i, \quad p_k = \binom{n}{k} \frac{\beta^{n-k}\alpha^k}{(\alpha+\beta)^n}, \quad k = 0, 1, \ldots, n.$$

However, they considered the sum of generalized Lindley random variables only in the case $\alpha_1 = \alpha_2 = \cdots = \alpha_n = \alpha$, $\beta_1 = \beta_2 = \cdots = \beta_n = \beta$, and γ_i are not equal. Therefore, in this study, a more general case is discussed to obtain the distribution of the sum of n independent generalized Lindley random variables with parameters $(\alpha_i, \beta_i, \gamma_i)$ using a simple gamma series and recurrence relation similar to Section 4.2.

4.5 Sum of inid generalized Lindley random variables

In this section, the probability density function and cumulative distribution function of the sum of inid random variables is considered. In addition, the distributions of the sum of inid extended exponential are derived: Lindley, three-parameter Lindley, and gamma random variables as special cases.

4.5.1 The exact probability density function

Let X_i be independent generalized Lindley random variables with parameters α_i, $\gamma_i > 0$, $\beta_i \geq 0$ for $i = 1, \ldots, n$. Then, using the direct product of $\mathbb{B} = \{0, 1\}$, we obtain the exact probability density function of $S = X_1 + \cdots + X_n$ as follows:

Theorem 4.3. *The exact probability density function of S is expressed as*

$$f_{\mathrm{SGL}}(x) = C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}} + k)} x^{\rho_{\boldsymbol{\tau}}+k-1} e^{-\alpha_1 x}, \quad x > 0, \tag{4.8}$$

where

$$C = \prod_{i=1}^{n} \frac{\alpha_i^{\gamma_i+1}}{\alpha_i + \beta_i}, \quad D_{\boldsymbol{\tau}} = \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_1^{\gamma_i+(\boldsymbol{\tau})_i}}, \quad \rho_{\boldsymbol{\tau}} = \sum_{i=1}^{n} \left\{\gamma_i + (\boldsymbol{\tau})_i\right\},$$

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j \eta_j \delta_{k+1-j}, \quad k = 0, 1, 2, \ldots, \quad \delta_0 = 1,$$

$$\eta_j = \sum_{s=1}^{n} \frac{\gamma_s + (\boldsymbol{\tau})_s}{j} \left(1 - \frac{\alpha_s}{\alpha_1}\right)^j, \quad \alpha_1 = \max_i(\alpha_i),$$

and $(\boldsymbol{\tau})_i$ is i-th element of $\boldsymbol{\tau} \in \mathbb{B}^n$. If $\beta_i = 0$ and $(\boldsymbol{\tau})_i = 0$, define $\beta_i^{(\boldsymbol{\tau})_i} = 1$.

Proof. Following a similar procedure to that of (4.2). The moment generating function of the generalized Lindley distribution is given by

$$M_{\mathrm{GL}}(t) = \left(\frac{\alpha}{\alpha - t}\right)^{\gamma+1} \frac{\alpha + \beta - t}{\alpha + \beta}, \quad t < \alpha.$$

Therefore, the moment generating function of $S = X_1 + \cdots + X_n$ is

$$\begin{aligned} M_{\mathrm{SGL}}(t) &= \left(\prod_{i=1}^{n} \frac{\alpha_i^{\gamma_i+1}}{\alpha_i + \beta_i}\right) \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{(\alpha_i - t)^{\gamma_i + (\boldsymbol{\tau})_i}} \\ &= \left(\prod_{i=1}^{n} \frac{\alpha_i^{\gamma_i+1}}{\alpha_i + \beta_i}\right) \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} \prod_{i=1}^{n} \frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_i^{\gamma_i + (\boldsymbol{\tau})_i}} \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_i}\right)^{-\gamma_i - (\boldsymbol{\tau})_i}. \end{aligned}$$

Let $\alpha_1 = \max_i(\alpha_i)$. Herein, denote

$$h(t) = \prod_{i=1}^{n} \left(1 - \frac{t}{\alpha_i}\right)^{-\gamma_i - (\boldsymbol{\tau})_i}.$$

Using the identity

$$1 - \frac{t}{\alpha_i} = \left(1 - \frac{t}{\alpha_1}\right) \frac{\alpha_1}{\alpha_i} \left[1 - \frac{1 - \frac{\alpha_i}{\alpha_1}}{1 - \frac{t}{\alpha_1}}\right],$$

then,

$$\log h(t) = \log\left[\prod_{i=1}^{n}\left\{\left(1-\frac{t}{\alpha_1}\right)\left(\frac{\alpha_1}{\alpha_i}\right)\right\}^{-\gamma_i-(\boldsymbol{\tau})_i}\right] + \sum_{i=1}^{n}\log\left[1-\frac{1-\frac{\alpha_i}{\alpha_1}}{1-\frac{t}{\alpha_1}}\right]^{-\gamma_i-(\boldsymbol{\tau})_i}.$$

By applying the Maclaurin expansion of $\log(1-x)$,

$$\log h(t) = \log\left[\prod_{i=1}^{n}\left(\frac{\alpha_i}{\alpha_1}\right)^{\gamma_i+(\boldsymbol{\tau})_i}\left(1-\frac{t}{\alpha_1}\right)^{-\rho_{\boldsymbol{\tau}}}\right] + \sum_{j=1}^{\infty}\eta_j\left(1-\frac{t}{\alpha_1}\right)^{-j},$$

where

$$\rho_{\boldsymbol{\tau}} = \sum_{i=1}^{n}\{\gamma_i+(\boldsymbol{\tau})_i\}, \quad \eta_j = \frac{1}{j}\sum_{i=1}^{n}\{\gamma_i+(\boldsymbol{\tau})_i\}\left(1-\frac{\alpha_i}{\alpha_1}\right)^{j}.$$

Therefore,

$$h(t) = \prod_{i=1}^{n}\left(\frac{\alpha_i}{\alpha_1}\right)^{\gamma_i+(\boldsymbol{\tau})_i}\left(1-\frac{t}{\alpha_1}\right)^{-\rho_{\boldsymbol{\tau}}}\exp\left(\sum_{j=1}^{\infty}\eta_j\left(1-\frac{t}{\alpha_1}\right)^{-j}\right).$$

The terms of the same order in the Taylor series are calculated together to obtain,

$$\exp\left(\sum_{j=1}^{\infty}\eta_j\left(1-\frac{t}{\alpha_1}\right)^{-j}\right) = \sum_{k=0}^{\infty}\delta_k\left(1-\frac{t}{\alpha_1}\right)^{-k}.$$

The coefficient δ_k is obtained by the following recursive formula:

$$\delta_{k+1} = \frac{1}{k+1}\sum_{i=1}^{k+1} i\eta_i\delta_{k+1-i}, \quad k=0,1,2,\ldots,$$

where $\delta_0 = 1$. Hence, the moment generating function of S is

$$\begin{aligned} M_{\mathrm{SGL}}(t) &= \left(\prod_{i=1}^{n}\frac{\alpha_i^{\gamma_i+1}}{\alpha_i+\beta_i}\right)\sum_{\boldsymbol{\tau}\in\mathbb{B}^n}\prod_{i=1}^{n}\frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_i^{\gamma_i+(\boldsymbol{\tau})_i}}\prod_{i=1}^{n}\left(1-\frac{t}{\alpha_i}\right)^{-\gamma_i-(\boldsymbol{\tau})_i} \\ &= \left(\prod_{i=1}^{n}\frac{\alpha_i^{\gamma_i+1}}{\alpha_i+\beta_i}\right)\sum_{\boldsymbol{\tau}\in\mathbb{B}^n}\prod_{i=1}^{n}\frac{\beta_i^{(\boldsymbol{\tau})_i}}{\alpha_1^{\gamma_i+(\boldsymbol{\tau})_i}}\sum_{k=0}^{\infty}\delta_k\left(1-\frac{t}{\alpha_1}\right)^{-(\rho_{\boldsymbol{\tau}}+k)}. \end{aligned}$$

Note that $\sum_{k=0}^{\infty}\delta_k\left(1-\frac{t}{\alpha_1}\right)^{-(\rho_{\boldsymbol{\tau}}+k)}$ is the same as the moment generating function of the mixture of gamma distributions. Then, we apply the inverse transformation of the moment generating function term-by-term. Therefore, the theorem is completely proved. □

4.5.2 The exact cumulative distribution function

We derive the cumulative distribution function by the term-by-term integration of $f_S(x)$ in (4.8).

Theorem 4.4. *The exact cumulative distribution function $F_{\text{SGL}}(x) = P(S \le x)$ is*

$$F_{\text{SGL}}(x) = C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}+k+1)} e^{-\alpha_1 x} {}_1F_1(1; \rho_{\boldsymbol{\tau}}+k+1; \alpha_1 x), \tag{4.9}$$

where ${}_1F_1(a;b;z)$ is the confluent hypergeometric function with the integral formula of

$${}_1F_1(a;b;z) = \frac{\Gamma(b)}{\Gamma(a)\Gamma(b-a)} \int_0^1 e^{zw} w^{a-1}(1-w)^{b-a-1} dw, \quad 0 < a < b.$$

Proof. Let $a = \max_{2\le \ell \le n}(1-\alpha_\ell/\alpha_1)$, then

$$|\eta_j| = \sum_{s=1}^{n} \frac{(\gamma_s + (\boldsymbol{\tau})_s)(1-\alpha_j/\alpha_1)^j}{j} \le \frac{\rho_{\boldsymbol{\tau}} a^j}{j}, \quad j = 1, 2, \ldots, k+1.$$

From the definition of δ, we obtain

$$|\delta_{k+1}| \le \frac{\rho_{\boldsymbol{\tau}}}{k+1} \sum_{j=1}^{k+1} a^j |\delta_{k+1-j}|, \quad k = 0, 1, 2, \ldots$$

from the recursive equation as

$$|\delta_{k+1}| \le \frac{\rho_{\boldsymbol{\tau}}(\rho_{\boldsymbol{\tau}}+1)\cdots(\rho_{\boldsymbol{\tau}}+k)}{(k+1)!} a^{k+1}.$$

Therefore,

$$\begin{aligned} f_{\text{SGL}}(x) &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} e^{-\alpha_1 x} \sum_{k=0}^{\infty} \frac{\delta_k}{\rho_{\boldsymbol{\tau}} \cdots (\rho_{\boldsymbol{\tau}}+k-1)} (\alpha_1 x)^k \\ &\le C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} e^{-\alpha_1 x} \sum_{k=0}^{\infty} \frac{(\alpha_1 a x)^k}{k!} \\ &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{\alpha_1^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}})} x^{\rho_{\boldsymbol{\tau}}-1} e^{-(1-a)\alpha_1 x}. \end{aligned} \tag{4.10}$$

Here, (4.10) shows that the uniform convergence of (4.8). Then,

$$\begin{aligned} F_{\text{SGL}}(x) &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k}{\Gamma(\rho_{\boldsymbol{\tau}}+k)\alpha_1^{-(\rho_{\boldsymbol{\tau}}+k)}} \int_0^x w^{\rho_{\boldsymbol{\tau}}+k-1} e^{-\alpha_1 w} dw \\ &= C \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k}{\Gamma(\rho_{\boldsymbol{\tau}}+k)\alpha_1^{-(\rho_{\boldsymbol{\tau}}+k)}} \frac{1}{\alpha_1^{\rho_{\boldsymbol{\tau}}+k}} \int_0^{\alpha_1 x} e^{-t} t^{\rho_{\boldsymbol{\tau}}+k-1} dt \end{aligned}$$

$$= C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}+k+1)} {}_1F_1(\rho_{\boldsymbol{\tau}}+k; \rho_{\boldsymbol{\tau}}+k+1; -\alpha_1 x)$$

$$= C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}+k+1)} e^{-\alpha_1 x} {}_1F_1(1; \rho_{\boldsymbol{\tau}}+k+1; \alpha_1 x).$$

□

The truncation of Equation (4.9) up to $k \leq m$ is defined by

$$F_m(x) = C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \sum_{k=0}^{m} \frac{\delta_k (\alpha_1 x)^{\rho_{\boldsymbol{\tau}}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}+k+1)} e^{-\alpha_1 x} {}_1F_1(1; \rho_{\boldsymbol{\tau}}+k+1; \alpha_1 x), \tag{4.11}$$

and the truncation error bound is obtained, using Equation (4.10), by

$$C \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}} \frac{(\alpha_1 x)^{\rho_{\boldsymbol{\tau}}}}{\Gamma(\rho_{\boldsymbol{\tau}}+1)} e^{-(1-a)\alpha_1 x} {}_1F_1(1; \rho_{\boldsymbol{\tau}}+1; (1-a)\alpha_1 x) - F_m.$$

4.5.3 Special cases

Here, we derive the exact distributions of the sum of conventional random variables by substituting specific parameters into Theorems 4.3 and 4.4. First, it is shown that the sum of inid extended exponential random variables is the special case of the sum of inid generalized Lindley random variables in Corollary 4.5.

Corollary 4.5. *The exact probability density function and the exact cumulative distribution function of the sum of inid extended exponential random variables are expressed as f_{SEE} and F_{SEE} in Theorems 4.1 and 4.2, respectively, by substituting $\gamma_i = 1$ into Theorems 4.3 and 4.4*

In the location shift case, Abd El-Monsef (2016) proposed the three-parameter Lindley distribution. When location parameters θ_i are equal to zero, the three-parameter Lindley distribution is reduced to the extended exponential distribution. Thus, the sum of the three-parameter Lindley random variables is expressed as Corollary 4.6.

Corollary 4.6. *Let X_i be the three-parameter Lindley distribution and $Y_i = X_i - \theta_i$. Then, Y_i is distributed as the extended exponential distribution with parameters (α_i, β_i). Therefore, the sum of inid three-parameter Lindley random variables is obtained as $\sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \theta_i$.*

Additionally, the distribution of the sum of inid Lindley random variables f_{SL} and F_{SL} are obtained by substituting $\beta_i = 1$ and $\gamma_i = 1$ as shown in Corollary 4.7.

Corollary 4.7. *The exact probability density function and the exact cumulative distribution function of the sum of inid Lindley random variables are respectively expressed as*

$$f_{\mathrm{SL}}(x) = C^{\dagger} \sum_{\boldsymbol{\tau} \in \mathbb{B}^n} D_{\boldsymbol{\tau}}^{\dagger} \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{\rho_{\boldsymbol{\tau}}^{\dagger}+k}}{\Gamma(\rho_{\boldsymbol{\tau}}^{\dagger}+k)} x^{\rho_{\boldsymbol{\tau}}^{\dagger}+k-1} e^{-\alpha_1 x}, \quad x > 0,$$

$$F_{\mathrm{SL}}(x) = C^{\dagger} \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D^{\dagger}_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho^{\dagger}_{\boldsymbol{\tau}}+k}}{\Gamma(\rho^{\dagger}_{\boldsymbol{\tau}}+k+1)} e^{-\alpha_1 x} {}_1F_1(1;\rho^{\dagger}_{\boldsymbol{\tau}}+k+1;\alpha_1 x),$$

where

$$C^{\dagger} = \prod_{i=1}^{n} \frac{\alpha_i^2}{\alpha_i+1}, \quad D^{\dagger}_{\boldsymbol{\tau}} = \prod_{i=1}^{n} \frac{1}{\alpha_1^{1+(\boldsymbol{\tau})_i}}, \quad \rho^{\dagger}_{\boldsymbol{\tau}} = \sum_{i=1}^{n} \left\{1+(\boldsymbol{\tau})_i\right\},$$

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j\eta^{\dagger}_j \delta_{k+1-j}, \quad k=0,1,2,\ldots, \quad \delta_0 = 1,$$

$$\eta^{\dagger}_j = \sum_{s=1}^{n} \frac{1+(\boldsymbol{\tau})_s}{j} \left(1-\frac{\alpha_s}{\alpha_1}\right)^j, \quad \alpha_1 = \max_i(\alpha_i), \quad \mathbb{B} = \{0,1\},$$

and $(\boldsymbol{\tau})_i$ *is i-th element of* $\boldsymbol{\tau} \in \mathbb{B}^n$.

Similarly, the distribution of the sum of inid Shanker random variables f_{SS} and F_{SS} are obtained by substituting $\beta_i = 1/\alpha_i$ and $\gamma_i = 1$ as shown in Corollary 4.8.

Corollary 4.8. *The exact probability density function and the exact cumulative distribution function of the sum of inid Shanker random variables are respectively expressed as*

$$f_{\mathrm{SS}}(x) = C^{\star} \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D^{\star}_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{\rho^{\star}_{\boldsymbol{\tau}}+k}}{\Gamma(\rho^{\star}_{\boldsymbol{\tau}}+k)} x^{\rho^{\star}_{\boldsymbol{\tau}}+k-1} e^{-\alpha_1 x}, \quad x>0,$$

$$F_{\mathrm{SS}}(x) = C^{\star} \sum_{\boldsymbol{\tau}\in\mathbb{B}^n} D^{\star}_{\boldsymbol{\tau}} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho^{\star}_{\boldsymbol{\tau}}+k}}{\Gamma(\rho^{\star}_{\boldsymbol{\tau}}+k+1)} e^{-\alpha_1 x} {}_1F_1(1;\rho^{\star}_{\boldsymbol{\tau}}+k+1;\alpha_1 x)$$

where

$$C^{\star} = \prod_{i=1}^{n} \frac{\alpha_i^3}{\alpha_i^2+1}, \quad D^{\star}_{\boldsymbol{\tau}} = \prod_{i=1}^{n} \frac{1}{\alpha_1^{\gamma_i+(\boldsymbol{\tau})_i} \alpha_i^{(\boldsymbol{\tau})_i}}, \quad \rho^{\star}_{\boldsymbol{\tau}} = \sum_{i=1}^{n} \left\{1+(\boldsymbol{\tau})_i\right\},$$

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j\eta^{\star}_j \delta_{k+1-j}, \quad k=0,1,2,\ldots, \quad \delta_0 = 1,$$

$$\eta^{\star}_j = \sum_{s=1}^{n} \frac{\gamma_s+(\boldsymbol{\tau})_s}{j} \left(1-\frac{\alpha_s}{\alpha_1}\right)^j, \quad \alpha_1 = \max_i(\alpha_i), \quad \mathbb{B} = \{0,1\},$$

and $(\boldsymbol{\tau})_i$ *is i-th element of* $\boldsymbol{\tau} \in \mathbb{B}^n$.

Moreover, when $\beta_i = 0$, the probability density function (4.8) is reduced to the distribution of the sum of inid gamma random variables with parameters (γ_i, α_i), which is equivalent to Moschopoulos (1985) as shown in Corollary 4.9.

Corollary 4.9. *The exact probability density function and the exact cumulative distribution function of the sum of inid gamma random variables are respectively expressed as:*

$$f_{\mathrm{SG}}(x) = C^{\ddagger} \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{\rho^{\ddagger}+k}}{\Gamma(\rho^{\ddagger}+k)} x^{\rho^{\ddagger}+k-1} e^{-\alpha_1 x}, \quad x>0,$$

$$F_{\mathrm{SG}}(x) = C^{\ddagger} \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{\rho^{\ddagger}+k}}{\Gamma(\rho^{\ddagger}+k+1)} e^{-\alpha_1 x} {}_1F_1(1; \rho^{\ddagger}+k+1; \alpha_1 x),$$

where

$$C^{\ddagger} = \prod_{i=1}^{n} \left(\frac{\alpha_i}{\alpha_1}\right)^{\gamma_i}, \quad \rho^{\ddagger} = \sum_{i=1}^{n} \gamma_i,$$

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j \eta_j^{\ddagger} \delta_{k+1-j}, \quad k = 0, 1, 2, \ldots, \quad \delta_0 = 1,$$

$$\eta_j^{\ddagger} = \sum_{s=1}^{n} \frac{\gamma_s}{j} \left(1 - \frac{\alpha_s}{\alpha_1}\right)^j, \quad \alpha_1 = \max_i(\alpha_i).$$

Furthermore, when $\beta_i = 0$ and $\gamma_i = 1$, the probability density function (4.8) is reduced to the distribution of the sum of inid exponential random variables with parameters α_i as shown in Corollary 4.10.

Corollary 4.10.

$$f_{\mathrm{SE}}(x) = \prod_{i=1}^{n} \left(\frac{\alpha_i}{\alpha_1}\right) \sum_{k=0}^{\infty} \frac{\delta_k \alpha_1^{n+k}}{\Gamma(n+k)} x^{n+k-1} e^{-\alpha_1 x}, \quad x > 0,$$

$$F_{\mathrm{SE}}(x) = \prod_{i=1}^{n} \left(\frac{\alpha_i}{\alpha_1}\right) \sum_{k=0}^{\infty} \frac{\delta_k (\alpha_1 x)^{n+k}}{\Gamma(n+k+1)} e^{-\alpha_1 x} {}_1F_1(1; n+k+1; \alpha_1 x),$$

where

$$\delta_{k+1} = \frac{1}{k+1} \sum_{j=1}^{k+1} j \eta^{*} \delta_{k+1-j}, \quad k = 0, 1, 2, \ldots, \quad \delta_0 = 1,$$

$$\eta^{*} = \sum_{s=1}^{n} \frac{1}{j} \left(1 - \frac{\alpha_s}{\alpha_1}\right)^j, \quad \alpha_1 = \max_i(\alpha_i).$$

4.5.4 Comparison of approximations

In the previous section, the exact cumulative distribution function of S was obtained. However, it is difficult to calculate the exact probability when the number of random variables increases. Since the summation of overall elements $\boldsymbol{\tau}$ in $\mathbb{B}^n$ increases, it takes a long time to obtain the probability using F_m for $n = 10$. For example, when $n = 10$, $2^{10} = 1024$ summations have to be calculated. Hence, we need to approximate the distribution function. The accuracy of the approximation is compared by calculating the percentile.

We consider the use of the saddlepoint approximation for the distribution of the sum of inid generalized Lindley random variables. The saddlepoint approximation is given by (4.7), where the cumulant generating function and its derivatives of the sum of inid generalized Lindley random variables are given by

$$k(t) = \sum_{i=1}^{n} \log\left\{ \left(\frac{\alpha_i}{\alpha_i - t}\right)^{\gamma_i+1} \frac{\beta_i + \alpha_i - t}{\alpha_i + \beta_i} \right\},$$
$$k'(t) = \sum_{i=1}^{n} \left(\frac{\gamma_i + 1}{\alpha_i - t} - \frac{1}{\alpha_i + \beta_i - t} \right),$$
$$k''(t) = \sum_{i=1}^{n} \left(\frac{\gamma_i + 1}{(\alpha_i - t)^2} - \frac{1}{(\alpha_i + \beta_i - t)^2} \right).$$

Here, the following parameters that are generated from the uniform distribution $\mathrm{Unif}(0, 3)$ in a similar way to Murakami (2014); Nadarajah et al. (2015) are considered for Case 1:

- $n = 2$:

$$\boldsymbol{\alpha} = (2.680, 1.640),$$
$$\boldsymbol{\beta} = (0.208, 2.852),$$
$$\boldsymbol{\gamma} = (1.903, 0.552).$$

- $n = 5$:

$$\boldsymbol{\alpha} = (2.984, 0.510, 2.241, 1.286, 1.641),$$
$$\boldsymbol{\beta} = (1.705, 2.727, 1.569, 1.419, 2.292),$$
$$\boldsymbol{\gamma} = (1.548, 0.735, 2.754, 0.975, 2.651).$$

- $n = 10$:

$$\boldsymbol{\alpha} = (0.171, 1.765, 0.595, 0.504, 0.481, 1.564, 1.527, 1.821, 1.678, 1.276),$$
$$\boldsymbol{\beta} = (2.373, 2.037, 0.190, 1.052, 2.938, 0.656, 1.647, 1.221, 2.511, 2.752),$$
$$\boldsymbol{\gamma} = (2.198, 2.728, 2.991, 1.985, 0.503, 2.778, 1.123, 0.265, 1.851, 2.154).$$

In addition, the following parameters that are generated from the log-normal distribution $\mathrm{LogNorm}(0, 1)$ are considered for Case 2:

- $n = 2$:

$$\boldsymbol{\alpha} = (1.322, 1.042),$$
$$\boldsymbol{\beta} = (0.406, 0.310),$$
$$\boldsymbol{\gamma} = (1.185, 0.994).$$

- $n = 5$:

$$\boldsymbol{\alpha} = (0.655, 1.608, 0.322, 0.798, 1.507),$$
$$\boldsymbol{\beta} = (1.359, 0.273, 1.873, 0.870, 0.291),$$
$$\boldsymbol{\gamma} = (0.448, 3.524, 0.605, 1.564, 0.925).$$

- $n = 10$:

$$\boldsymbol{\alpha} = (1.398, 4.367, 1.030, 0.916, 1.347, 2.019, 0.596, 0.279, 1.602, 0.569),$$
$$\boldsymbol{\beta} = (0.331, 0.418, 0.885, 1.088, 1.767, 1.675, 0.722, 0.550, 1.938, 0.668),$$
$$\boldsymbol{\gamma} = (2.758, 0.215, 1.252, 0.505, 2.212, 4.133, 0.370, 1.022, 2.270, 1.243).$$

Tables 4.6 and 4.7 present the approximations to the distribution function (4.9). In the tables, $\tilde{s}$ is a percentile derived from 100, 000, 000 random numbers generated by S, and p is the corresponding probability. Here, the generalized Lindley distribution is a mixture of gamma distributions. Then, S is obtained by generating two gamma random numbers with parameter (γ_i, β_i) and $(\gamma_i + 1, \beta_i)$. In addition, F_m^{-1} represents the p-th percentile derived by solving $F_m(s) = p$ numerically, where F_m is defined in Equation (4.11). In numerical experiments, the accuracy and calculation time are compared for different values of m, namely, $m = 50$, 150, 500, 1000. Moreover, SA and NA are the percentile for the saddlepoint approximation and the normal approximation, respectively. Additionally, CT denotes the calculation time in seconds. We use Mathematica version 11 (CPU 2.80 GHz and 32.0 GB RAM).

Tables 4.6 and 4.7 reveal that F_{1000}^{-1} is enough to obtain the percentiles, except for $p = 0.9900$, $n = 10$ in Case 2. To obtain the value of $\tilde{s}$ more accurately, larger m have to be selected. However, SA and NA overcome such computational difficulty and SA is more accurate than NA.

Table 4.6: Approximations to the distribution for Case 1.

p	$\tilde{s}$	F_{50}^{-1}	F_{150}^{-1}	F_{500}^{-1}	F_{1000}^{-1}	SA	NA
			$n = 2$				
0.6000	1.505	1.505	1.505	1.505	1.505	1.503	1.689
0.7000	1.770	1.770	1.770	1.770	1.770	1.768	1.933
0.8000	2.117	2.117	2.117	2.117	2.117	2.113	2.219
0.9000	2.666	2.665	2.665	2.665	2.665	2.661	2.615
0.9500	3.183	3.182	3.182	3.182	3.182	3.177	2.943
0.9750	3.679	3.678	3.678	3.678	3.678	3.674	3.227
0.9900	4.316	4.315	4.315	4.315	4.315	4.311	3.557
CT		0.1	0.9	8.2	35.4	15.6	0.0
			$n = 5$				
0.6000	8.583	8.583	8.583	8.583	8.583	8.561	9.078
0.7000	9.478	9.478	9.478	9.478	9.478	9.455	9.928
0.8000	10.628	10.628	10.628	10.628	10.628	10.605	10.924
0.9000	12.426	12.425	12.425	12.425	12.425	12.406	12.304
0.9500	14.106	14.104	14.104	14.104	14.104	14.090	13.444
0.9750	15.717	15.716	15.715	15.715	15.715	15.706	14.432
0.9900	17.784	17.794	17.781	17.781	17.781	17.778	15.582
CT		1.7	14.0	145.5	637.8	41.1	0.0
			$n = 10$				
0.6000	42.228	*	42.228	42.228	42.228	42.199	43.824
0.7000	45.658	*	45.658	45.658	45.658	45.630	47.065
0.8000	50.001	*	50.000	50.000	50.000	49.975	50.858
0.9000	56.648	*	56.648	56.648	56.648	56.629	56.119
0.9500	62.728	*	62.724	62.724	62.724	62.712	60.463
0.9750	68.451	*	68.448	68.448	68.448	68.442	64.231
0.9900	75.658	*	75.657	75.657	75.657	75.656	68.613
CT		106.2	894.5	8314.9	37371.5	64.8	0.0

*There does not exist a root of $F_m(s) = p$.

Table 4.7: Approximations to the distribution for Case 2.

p	$\tilde{s}$	F_{50}^{-1}	F_{150}^{-1}	F_{500}^{-1}	F_{1000}^{-1}	SA	NA
				$n = 2$			
0.6000	2.303	2.303	2.303	2.303	2.303	2.304	2.624
0.7000	2.743	2.742	2.742	2.742	2.742	2.744	3.027
0.8000	3.321	3.321	3.321	3.321	3.321	3.322	3.498
0.9000	4.241	4.240	4.240	4.240	4.240	4.242	4.151
0.9500	5.108	5.107	5.107	5.107	5.107	5.108	4.690
0.9750	5.940	5.939	5.939	5.939	5.939	5.940	5.158
0.9900	7.006	7.003	7.003	7.003	7.003	7.004	5.702
CT		0.2	1.0	10.0	44.3	18.1	0.0
				$n = 5$			
0.6000	12.310	12.309	12.309	12.309	12.309	12.278	13.103
0.7000	13.719	13.718	13.718	13.718	13.718	13.683	14.433
0.8000	15.529	15.528	15.528	15.528	15.528	38.334	15.989
0.9000	18.357	18.355	18.355	18.355	18.355	38.334	18.147
0.9500	20.995	20.992	20.992	20.992	20.992	38.334	19.930
0.9750	23.523	23.520	23.520	23.520	23.520	23.494	21.476
0.9900	26.757	26.757	26.757	26.757	26.757	26.741	23.274
CT		2.0	13.3	134.3	678.1	41.2	0.0
				$n = 10$			
0.6000	22.585	*	22.584	22.584	22.584	22.554	23.459
0.7000	24.460	*	24.460	24.460	24.460	24.426	25.239
0.8000	26.835	*	26.834	26.834	26.834	26.799	27.323
0.9000	30.479	*	30.477	30.477	30.477	30.441	30.212
0.9500	33.822	*	33.827	33.821	33.821	33.786	32.598
0.9750	36.984	*	37.143	37.032	37.032	36.954	34.667
0.9900	40.987	*	*	42.739	42.739	40.968	37.073
CT		93.8	884.1	9078.9	29247.3	83.6	0.0

*There does not exist a root of $F_m(s) = p$.

Chapter 5

Conclusions and Discussions

This paper discusses the approximate, asymptotic, and limiting distribution of the one-sample and two-sample nonparametric statistics. The approximate distribution overcomes the problem of computation time to derive a null distribution by the exact permutation when the sample size is large. The asymptotic and the limiting distributions make it easier to obtain the critical values for a discussion of the asymptotic power of the test. In Chapter 2, the saddlepoint approximation and the normal approximation for the distribution of the ESWSR test were considered, and the accuracy of the approximations was investigated. Numerical calculations indicated that the saddlepoint approximation with a continuity correction is the most suitable. Furthermore, we obtained the asymptotic power and discussed the unbiasedness and biasedness of the test. Notably, the exact mean and variance were derived under the alternative hypothesis. However, calculating the second moment of T_ν under H_{11} can be challenging because of computational complexity. Once this problem is addressed, the accurate asymptotic power can be calculated and sample size determination can be discussed. Utilizing the selector based on the asymptotic efficiency of the test was also suggested and the use of the ESWSR test was demonstrated with real data. In Chapter 3, we derived the asymptotic non-null and limiting null distribution of the maximum test $\max(t, t_R)$ to overcome the disadvantage of computation time. A simulation study demonstrated the convergence of the test to the limiting distribution. In addition, indicators, such as, sample size, the data size ratio of samples, were presented for the rejection point of the limiting distribution. Although the limiting distribution of the maximum test requires probability weighted moments of the underlying distribution, the simulation results revealed that the limiting distribution computed from the sample with unbiased estimators could also be used. In practice, there are many cases where homoscedasticity cannot be assumed. Thus, for future work, it is worth considering the test for the location parameter with different variances such as in Welz et al. (2018) for a maximum test based on the Brunner-Munzel test (Brunner and Munzel, 2000).

In addition, this paper discusses the distribution of the sum of inid extended exponential random variables and the sum of inid generalized Lindley random variables. In chapter 4, we obtained the exact distribution of the sum of the inid extended exponential n random variables using the simple gamma series and recursive formulas. Numerical simulation showed that the saddlepoint approximation is the most appropriate for the cumulative distribution function in terms of calculation time. The distribution of the sum of the extended exponential random variables was a suitable model for real data based on AIC. Since the generalized Lindley distribution contains the extended exponential distribution as a special case, we also obtained the exact distribution of the sum of inid generalized Lindley n random variables. It was demonstrated that the sum of inid generalized Lindley random variables includes the sum of inid exponential, gamma, Lindley, Shanker, and extended exponential random variables as special cases. Numerical simulations revealed that the saddlepoint approximation is suitable for the cumulative distribution function in terms of calculation time. However, the calculation of the exact cumulative distribution function took much longer when the number of variables and m increased. Therefore, an algorithm to reduce the calculation time for F_m, would benefit from the exact distribution of the sum of inid random variables in modelling real data.

Bibliography

Abd El-Monsef, M. M. E. (2016). A new Lindley distribution with location parameter. *Communications in Statistics - Theory and Methods*, 45:5204–5219.

Almarashi, A. M., Elgarhy, M., Elsehetry, M. M., Golam Kibria, B. M., and Algarni, A. (2019). A new extension of exponential distribution with statistical properties and applications. *Journal of Nonlinear Sciences and Applications*, 12:135–145.

Alouini, M. S., Abdi, A., and Kaveh, M. (2001). Sum of gamma variates and performance of wireless communication systems over Nakagami-fading channels. *IEEE Transactions on Vehicular Technology*, 50:1471–1480.

Amari, S. V. and Misra, R. B. (1997). Closed-form expressions for distribution of sum of exponential random variables. *IEEE Transactions on Reliability*, 46:519–522.

Amrhein, P. (1995). An example of a two-sided Wilcoxon signed rank test which is not unbiased. *Annals of the Institute of Statistical Mathematics*, 47:167–170.

Ansari, A. R. and Bradley, R. A. (1960). Rank-sum tests for dispersions. *The Annals of Mathematical Statistics*, 31:1174–1189.

Bagai, I., Deshpandé, J. V., and Kochar, S. C. (1989). Distribution free tests for stochastic ordering in the competing risks model. *Biometrika*, 76:775–781.

Bean, R., Froda, S., and van Eeden, C. (2004). The normal, Edgeworth, saddlepoint and uniform approximations to the Wilcoxon-Mann-Whitney null-distribution: a numerical comparison. *Journal of Nonparametric Statistics*, 16:279–288.

Benavoli, A., Corani, G., Mangili, F., Zaffalon, M., and Ruggeri, F. (2014). A Bayesian Wilcoxon signed-rank test based on the Dirichlet process. *In Proceedings of the 30th International Conference on Machine Learning*, pages 1–9.

Berk, R. H. and Jones, D. H. (1979). Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Zeitschrift für Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, 47:47–59.

Bradley, D. M. and Gupta, C. R. (2002). On the distribution of the sum of n non-identically distributed uniform random variables. *Annals of the Institute of Statistical Mathematics*, 54:689–700.

Brunner, E. and Munzel, U. (2000). The nonparametric Behrens-Fisher problem: Asymptotic theory and a small - sample approximation. *Biometrical Journal*, 42:17–25.

Büning, H. (1991). *Robuste und adaptive Tests*. De Gruyter.

Büning, H. (1996). Adaptive tests for the c-sample location problem - The case of two-sided alternatives. *Communications in Statistics - Theory and Methods*, 25:1569–1582.

Büning, H. (1997). Robust analysis of variance. *Journal of Applied Statistics*, 24:319–332.

Büning, H. (2001). Kolmogorov-Smirnov- and Cramér-von Mises type two-sample tests with various weight function. *Communications in Statistics - Simulation and Computation*, 30:847–865.

Butler, R. W. (2007). *Saddlepoint Approximations with Applications*. Cambridge University Press.

Chechile, R. A. (2018). A Bayesian analysis for the Wilcoxon signed-rank statistic. *Communications in Statistics - Theory and Methods*, 47:5241–5254.

Chernoff, H. and Savage, R. I. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *The Annals of Mathematical Statistics*, 29:972–994.

Chung, E. and Romano, J. P. (2016). Asymptotically valid and exact permutation tests based on two-sample U-statistics. *Journal of Statistical Planning and Inference*, 168:97–105.

Conover, W. J. and Iman, R. L. (1981). Rank transformation as a bridge between parametric and nonparametric statistics. *The American Statician*, 35:124–129.

Cornish, E. A. and Fisher, R. A. (1938). Moments and cumulants in the specification of distributions. *Review of the International Statistical Institute*, 5:307–320.

Daniels, H. E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, 25:631–650.

Daniels, H. E. (1987). Tail probability approximations. *International Statistical Review*, 55:37–48.

Deuchler, G. (1914). Über die methoden der korrelationsrechnung in der pädagogik und psychologie. *Zeitschrift für Pädagogische Psychologie und Experimentelle Pädagogik*, 15:114–31, 145–59, 229–42.

Ding, C. G. (1992). Algorithm AS 275: computing the non-central χ^2 distribution function. *Journal of the Royal Statistical Society, Series C*, 41:478–482.

Ding, X., Xu, X., and Zhao, S. (2010). On exactness and unbiasedness of confidence bands for a continuous distribution function. *Science China Mathematics*, 53:2665–78.

Duchesne, P. and De Micheaux, P. L. (2010). Computing the distribution of quadratic forms: Further comparisons between the Liu-Tang-Zhang approximation and exact methods. *Computational Statistics & Data Analysis*, 54:858–862.

Easton, G. S. and Ronchetti, E. (1986). General saddlepoint approximations with applications to L statistics. *Journal of the American Statistical Association*, 81:420–430.

Edgeworth, F. Y. (1905). The low of error. *Transactions of the Cambridge Philosophical Society*, 20:36–66.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7:1–26.

Eisinga, R., Grotenhuis, M. T., and Pelzer, B. (2013). Saddlepoint approximation for the sum of independent non-identically distributed binomial random variables. *Statistica Neerlandica*, 67:190–201.

Farebrother, R. W. (1987). Algorithm AS 231: the distribution of a noncentral χ^2 variable with nonnegative degrees of freedom. *Journal of the Royal Statistical Society, Series C*, 17:402–405.

Frey, J. (2009). Unbiased goodness-of-fit. *Journal of Statistical Planning and Inference*, 139:3690–3697.

Froda, S. and van Eeden, C. (2000). A uniform saddlepoint expansion for the null-distribution of the Wilcoxon-Mann-Whitney statistic. *Canadian Journal of Statistics*, 28:137–149.

Gabler, S. and Wolff, C. (1987). A quick and easy approximation to the distribution of a sum of weighted chi-square variables. *Statistische Hefte*, 28:317–325.

Ghitany, M. E., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, 78:493–506.

Gibbons, J. and Chakraborti, S. (2011). *Nonparametric statistical inference*. Chapman & Hall/CRC, 5th edition.

Glaser, R. E. (1983). *Statistical analysis of Kevlar 49/epoxy composite stress-rupture data*. Livermore, CA: Report UCID-19849, Lawrence Livermore National Laboratory.

Gómez, Y. M., Bolfarine, H., and Gómez, H. W. (2014). A new extension of the exponential distribution. *Revista Colombiana de Estadistica*, 37:25–34.

Goria, M. N. (1980). Some locally most powerful generalized rank tests. *Biometrika*, 67:497–500.

Goutis, C. and Casella, G. (1999). Explaining the saddlepoint approximation. *The American Statistician*, 53:216–224.

Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: An alternative to gamma and Weibull distributions. *Biometrical Journal*, 43:117–130.

Hájek, J. (1962). Asymptotically most powerful rank-order tests. *The Annals of Mathematical Statistics*, 33:1124–1147.

Hájek, K., Šidák, L., and Sen, P. K. (1999). *Theory of Rank Tests*. Academic Press, 2nd edition.

Hanyuda, B. and Murakami, H. (2021). Power comparisons of the unbiased Berk-Jones test and the unbiased reversed Berk-Jones test. *Communications in Statistics - Simulation and Computation*, 50:1009–1024.

Hettmansperger, T. P. and McKean, J. W. (2010). *Robust Nonparametric Statistical Methods*. CRC Press.

Hodges, J. L. and Lehmann, E. L. (1956). The efficiency of some nonparametric competitors of the t-test. *The Annals of Mathematical Statistics*, 27:324–335.

Hogg, R. V. (1974). Adaptive robust procedures: A partial review and some suggestions for future applications and theory. *Journal of the American Statistical Association*, 69:909–927.

Huzurbazar, S. (1999). Practical saddlepoint approximations. *The American Statistician*, 53:225–232.

Jager, L. and Wellner, J. A. (2007). Goodness-of-fit tests via phi-divergences. *The Annals of Statistics*, 35:2018–2053.

Jensen, J. L. (1995). *Saddlepoint Approximations*. Oxford University Press.

Jonckheere, A. R. (1954). A distribution-free k-sample test agsinst ordered alternatives. *Biometrika*, 41:133–145.

Jurečková, J., Picek, J., and Schindler, M. (2019). *Robust Statistical Methods with R*. Chapman & Hall/CRC, second edition.

Kamps, U. (1990). Characterizations of the exponential distribution by weighted sums of iid random variables. *Statistical Papers*, 31:233–237.

Khuong, H. V. and Kong, H.-Y. (2006). General expression for pdf of a sum of independent exponential random variables. *IEEE Communications Letters*, 10:159–161.

Kitani, M. and Murakami, H. (2020a). The limiting distribution of combining the t and Wilcoxon rank sum tests. *Statistics*, 54:871–884.

Kitani, M. and Murakami, H. (2020b). On the distribution of the sum of the independent and non-identically extended exponential random variables. *Japanese Journal of Statistics and Data Science*, 3:23–37.

Kitani, M. and Murakami, H. (2022). One-sample location test based on the sign and Wilcoxon signed-rank tests. *Journal of Statistical Computation and Simulation*, 92:610–622.

Kitani, M., Murakami, H., and Hashiguchi, H. (2021). The distribution of the sum of independent and non-identically generalized Lindley random variables. *Communications in Statistics - Theory and Methods*. DOI: 10.1080/03610926.2021.1955387.

Kolassa, J. E. (2006). *Series Approximation Methods in Statistics*. Springer-Verlag.

Landwehr, J. M., Matalas, N. C., and Wallis, J. R. (1979). Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, 15:1055–1064.

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*. Springer, 3rd edition.

Lemonte, A. J. ans Cordeiro, G. M. and Moreno-Arenas, G. (2016). A new useful three-parameter extension of the exponential distribution. *Statistics*, 50:312–337.

Lindley, D. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society. Series B*, 20:102–107.

Lugannani, R. and Rice, S. O. (1980). Saddlepoint approximation for the distribution of the sum of independent random variables. *Advances in Applied Probability*, 12:475–490.

Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18:50–60.

Massey, F. J. (1950). A note on the power of a non-parametric test. *The Annals of Mathematical Statistics*, 21:440–443.

Mathai, A. M. (1982). Storage capacity of a dam with gamma type inputs. *Annals of the Institute of Statistical Mathematics*, 34:591–597.

Micceri, T. (1989). The unicorn, the normal curve, and other improbable creatures. *Psychological Bulletin*, 105:156–166.

Miyazaki, R. and Murakami, H. (2020). The non-null limiting distribution of the generalized Baumgartner statistic based on the Fourier series approximation. *Statistical Papers*, 61:1893–1909.

Monti, A. C. (1993). A new look at the relationship between Edgeworth expansion and saddlepoint approximation. *Statistics & Probability Letters*, 17:49–52.

Mood, A. M. (1954). On the asymptotic efficiency of certain nonparametric two-sample tests. *The Annals of Mathematical Statistics*, 25:514–522.

Moschopoulos, P. G. (1985). The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics*, 37:541–544.

Murakami, H. (2010). A numerical comparison of the normal and some saddlepoint approximations to a distribution-free test for stochastic ordering in the competing risks model. *Computational Statistics*, 25:633–643.

Murakami, H. (2014). A saddlepoint approximation to the distribution of the sum of independent non-identically uniform random variables. *Statistica Neerlandica*, 68:267–275.

Murakami, H. (2015). Approximations to the distribution of sum of independent non-identically gamma random variables. *Mathematical Sciences*, 9:205–213.

Murakami, H. and Kamakura, T. (2009). A saddlepoint approximation to a nonparametric test for ordered alternatives. *Journal of the Japan Statistical Society*, 39:143–153.

Nadarajah, S. (2008). A review of results on sums of random variables. *Acta Applicandae Mathematicae*, 103:131–140.

Nadarajah, S. and Haghighi, F. (2011). An extension of the exponential distribution. *Statistics*, 45:543–558.

Nadarajah, S., Jiang, X., and Chu, J. (2015). A saddlepoint approximation to the distribution of the sum of independent non-identically beta random variables. *Statistica Neerlandica*, 69:102–114.

Nádas, A. (1971). The distribution of the identified minimum of a normal pair determines' the distribution of the pair. *Technometrics*, 13:201–202.

Nakagami, M. (1960). The m-distribution - A general formula of intensity distribution of rapid fading. *Statistical Methods in Radio Wave Propagation*, pages 3–36.

Nanna, M. J. and Sawilowsky, S. S. (1998). Analysis of Likert scale data in disability and medical rehabilitation research. *Psychological Methods*, 3:55–67.

Neave, H. R. and Granger, C. W. J. (1968). A Monte Carlo study comparing various two-sample tests for differences in means. *Technometrics*, 10:509–522.

Neuhäuser, M. (2015). Combining the t test and Wilcoxon's rank-sum test. *Journal of Applied Statistics*, 42:2760–2775.

Noether, G. E. (1955). On a theorem of Pitman. *The Annals of Mathematical Statistics*, 26:64–68.

Olds, E. G. (1952). A note on the convolution of uniform distributions. *The Annals of Mathematical Statistics*, 23:282–285.

Penev, S. and Raykov, T. (2000). A wiener germ approximation of the noncentral chi square distribution and of its quantiles. *Computational Statistics*, 15:219–228.

Philonenko, P., Postovalov, A., and Kovalevskii, A. (2016). The limit distribution of the maximum value test statistic in the general case. *11th International Forum on Strategic Technology*, pages 428–430.

Pitman, E. J. G. (1948). *Notes on Non-parametric Statistical Inference*. Columbia University.

Policello, G, E. and Hettmansperger, T. P. (1976). Adaptive robust procedures for the one-sample location problem. *Journal of the American Statistical Association*, 71:624–633.

Potuschak, H. and Müller, W. G. (2009). More on the distribution of the sum of uniform random variables. *Statistical Papers*, 50:177–183.

Randles, R. H. and Hogg, R. V. (1973). Adaptive distribution-free tests. *Communications in Statistics*, 2:337–356.

Reid, N. (1988). Saddlepoint methods and statistical inference (with discussion). *Statistical Science*, 3:213–238.

Rosenblatt, J. D. and Benjamini, Y. (2018). On mixture alternatives and Wilcoxon's signed rank test. *The American Statistician*, 72:344–347.

Rosner, B., Glynn, R. J., and Lee, M. L. T. (2006). The Wilcoxon signed rank test for paired comparisons of clustered data. *Biometrics*, 62:185–192.

Sadooghi-Alvandi, S. M., Nematollahi, A., and Habibi, R. (2009). On the distribution of the sum of independent uniform random variables. *Statistical Papers*, 50:171–175.

Sedlmeier, P. and Renkewitz, F. (2008). *Forschungsmethoden und Statistik in der Psychologie*. Pearson.

Shanker, R. (2015). Shanker distribution and its apprications. *International Journal of Statistics and Applications*, 5:338–348.

Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*, 6:64–71.

Shanker, R., Sharma, S., and Shanker, R. (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*, 4:363–368.

Solomon, H. and Stephens, M. A. (1977). Distribution of a sum of weighted chi-square variables. *Journal of the American Statistical Association*, 72:881–885.

Sugiura, N. (1965). An example of the two-sided Wilcoxon test which is not unbiased. *Annals of the Institute of Statistical Mathematics*, 17:261–263.

Sugiura, N., Murakami, H., and Lee, S.-K. Maeda, Y. (2006). Biased and unbiased two-sided Wilcoxon tests for equal sample sizes. *Annals of the Institute of Statistical Mathematics*, 58:93–100.

Tank, F. and Eryilmaz, S. (2015). The distributions of sum, minima and maxima of generalized geometric random variables. *Statistical Papers*, 56:1191–1203.

Tertpstra, T. J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Proceedings if the Kononklijke Nederlandse Akademie von Wetenschappen*, 55:327–333.

Thompson, R. (1966). Bias of the one-sample Cramér-von Mises test. *Journal of the American Statistical Association*, 61:246–247.

Welz, A., Ruxton, G. D., and Neuhäuser, M. (2018). A non-parametric maximum test for the Behrens-Fisher problem. *Journal of Statistical Computation and Simulation*, 88:1336–1347.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics*, 1:80–83.

Zakerzadeh, H. and Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical Extension*, 3:13–25.