Statistical Problems Relevant to Risk Assessment

REEM SHEDAYED A ALHARBI

Statistical Problems Relevant to Risk Assessment

Reem S. Alharbi

A Thesis presented for the degree of
Doctor of Philosophy

Statistics and Probability
Department of Mathematical Sciences
University of Durham
United Kingdom

August 2021

Statistical Problems Relevant to Risk Assessment

Reem S. Alharbi

Submitted for the degree of Doctor of Philosophy
August 2021

Abstract

This thesis discusses two theoretical statistical problems which have potential uses in risk assessment. The likely application of the first problem is to ecotoxicological risk assessment, while the second problem has a wide range of potential applications in risk assessment. The two problems have important mathematical features in common.

The first problem is concerned with key dominance properties for the arithmetic mean as the sample size increases. We show mathematically that the dominance properties hold for all distributions with symmetric log-concave densities. A detailed and comprehensive analysis of what happens when the sample size increases from one to two for two-component scale and location mixtures of normal distributions is introduced.

The second problem relates to combining limited probabilistic expert judgements on multiple quantities in order to provide limited probabilistic information about a derived quantity. First, a working hypothesis that simplifies calculations for the derived quantity is developed. Second, we mathematically show that the working hypothesis holds for all distributions with symmetric log-concave densities. In addition, it holds for negatively-skewed Azzalini-style skew-symmetric distributions with log-concave kernels when two quantities are involved. Moreover, under a specific condition, the working hypothesis is valid whenever the underlying distributions have log-concave right tail probability functions or partial log-concave right tail probability functions when two quantities are involved.

Declaration

The work in this thesis is based on research carried out in the Department of Mathematical Sciences at Durham University. No part of this thesis has been submitted elsewhere for any degree or qualification.

Acknowledgements

First and foremost, I thank God for everything.

I am indebted to my supervisor, Prof. Peter Craig, for his guidance and unlimited support along with his valuable discussions in all stages of this thesis. He made the work on this research easy and smooth. I would say without his spur, this work would have never been possible.

I am grateful for my family's love, understanding, and continuing support to complete this research.

I also thank King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia, for their financial support.

Finally, I wish to acknowledge everyone who has supported me during this research.

Contents

Abstract ii

Declaration iii

Acknowledgements iv

1 Introduction and Background 1

1.1 Motivation 1

1.1.1 Problem 1: Ecotoxicological Risk Assessment 2

1.1.2 Some Concepts Related to Problem 2 7

1.1.3 Problem 2: Combining Minimal Judgements about Uncertain Quantities 11

1.1.4 Proschan's Result 14

1.2 Outline of Thesis 15

2 Log-Concave Probability Density Functions 17

2.1 Introduction 17

2.2 Log-Concave Functions and Probability Densities 19

2.3 Symmetric Log-Concave Probability Density Functions 23

2.4 Conclusions 33

3 Dominance of Arithmetic Mean for Repeated Sampling in Some Risk Problems 34

3.1 Introduction 34

3.2 Example of Application by EFSA 35

3.3 Background 36

3.4 Normal Distribution LSSDs 37
3.5 Location-Scale Family of Distributions 39
3.6 Symmetric Log-Concave Probability Density Functions 40
3.7 A Two-Component Scale Mixture of Normal Distributions 42
3.7.1 MFE_n 42
3.7.2 $\mathrm{PFE}_n(\alpha)$ 56
3.8 A Two-Component Location Mixture of Normal Distributions 64
3.8.1 MFE_n 64
3.8.2 $\mathrm{PFE}_n(\alpha)$ 83
3.9 Conclusions 89

4 Partial Probability Judgments in Risk Assessment **96**
4.1 Introduction 96
4.2 Example of Application by EFSA 97
4.3 A Simplification 102
4.4 Consequences of Proschan's Result 102
4.5 Examples 104
4.6 Skew-Symmetric Distributions with Log-Concave Kernels 106
4.7 SHE and SHEL Distribution Functions 110
4.8 Stochastic Dominance 111
4.9 Right Tail Probability Functions 113
4.10 Log-Concave Right Tail Probability Functions 119
4.10.1 Convolution of SHE Distribution Functions 119
4.10.2 Generalization for $n > 2$ 129
4.11 Partial Log-Concave Right Tail Probability Functions 131
4.11.1 Convolution of SHEL Distribution Functions 131
4.12 Re-Scaling and Shifting 145
4.13 Conclusions 146

5 Conclusions and Future Research Directions **148**

Glossary **153**

A Basic and Auxiliary Results **154**

A.1 Evaluation of Some Integrals 154

A.2 Lambert $\mathcal{W}$ Function 156

Bibliography **157**

Chapter 1

Introduction and Background

In this chapter, we will discuss the motivation for making a contribution to two topics in quantitative risk assessment. In addition, an outline of the thesis is provided.

1.1 Motivation

Risk analysis entails three phases: risk assessment, risk management, and risk communication, see WHO (1995); EFSA (2012); and references therein. Risk assessment is concerned with identifying and evaluating actions that are likely to cause harm. It entails four branches: hazard identification, hazard characterisation, exposure assessment and risk characterisation. Risk characterisation is a combination of the other three branches of risk assessment (Renwick et al., 2003). In EC (2000), risk characterisation is defined as "The quantitative or semi-quantitative estimate, including attendant uncertainties, of the probability of occurrence and severity of adverse effect(s)/event(s) in a given population under defined exposure conditions based on hazard identification, hazard characterisation and exposure assessment."

The variability in data and uncertainty due to incomplete information affect the conclusions that are made in risk characterisation (Renwick et al., 2003). Quantifying variability and uncertainty is an important part in risk characterisation since the result will be used in risk management (decision making process). Such quantification can contribute significantly to the transparency and robustness of risk assessment.

Uncertainty has various definitions. In EFSA (2018b), it is defined as
"a general term referring to all types of limitations in available knowledge that affect the range and probability of possible answers to an assessment question."
The two kinds of uncertainty are aleatory uncertainty (also called variability, objective uncertainty, dissonance, or irreducible uncertainty) and epistemic uncertainty (also called uncertainty, ignorance, incertitude, subjective uncertainty, nonspecificity, or reducible uncertainty), see Burgman (2005) and EFSA (2018a) among many in the literature. Aleatory uncertainty can not be reduced by collecting more information, whereas epistemic uncertainty can.

Uncertainty may be described qualitatively or quantitatively. Quantitative description may be deterministic or probabilistic. In this thesis, we are concerned with probabilistic methods. Frequency probability is reasonable in quantifying aleatory uncertainty, whereas epistemic uncertainty may be quantified by using subjective probability which reflects the personal degree of belief. Ramsey (1988) and Savage (1954/1972) defined subjective probabilities in terms of an individual's preferences as certain consistency assumptions are satisfied. Consequently, this probability differs from one expert to another due to the difference in their knowledge and experience.

The goal of the thesis is to make contributions to solving two mathematical problems that are directly relevant to improving the treatment of uncertainty in two related areas of practical risk assessment. These risk assessment areas and the associated mathematical problems are described in Subsections 1.1.1 and 1.1.3 and originate from the work of EFSA. Although the two problems are quite different, they have important mathematical features in common.

1.1.1 Problem 1: Ecotoxicological Risk Assessment

The first problem relates to the applied question of what to do in ecotoxicological risk assessment if more species are tested than required.

Ecotoxicology is concerned with the ecological effects of harmful (toxic) chemicals in the natural environment. An important part of the natural environment is the aquatic compartment: ditches, streams, ponds, rivers, etc.. Ecotoxicological risk assessment requires a procedure that must be imposed to ensure that a chem-

ical substance is used safely in the real environment. The risk is characterized by the risk characterisation ratio (RCR) which compares an acceptable environmental concentration (AEC), estimated under the hazard assessment, to a predicted environmental concentration (PEC), i.e. concentration of a substance expected to be found in any environment and estimated under exposure assessment. If the PEC is less than the AEC then the chemical substance is likely to pass. Otherwise further higher-tier risk assessment is required in order to release such a chemical substance (ECHA, 2012). Consequently, the decision problem for risk managers in this context is to specify the AEC. What this means is that users of a chemical in industry and agriculture are not allowed to create higher concentrations than AEC in the real environment.

In ecotoxicological risk assessment, the toxicity data which is used in the estimation procedure are species sensitivity values. These values represent the substance's concentration that causes a specific effect (endpoint) which is observable or measurable to these species. The commonly used lethal concentration (LC_{50}) represents the concentration that is responsible for the mortality of 50% of a targeted species for a specific exposure period. Other typical concentration levels are LC_{10}, LC_{25}, LC_{75}, and LC_{100}.

A simple approach, sometimes taken to this risk management problem, is to measure a single species's sensitivity in an experiment, and then divide this value by a fixed factor which is called the *assessment factor* (AF), also referred to as safety (or uncertainty) factor. The AF is expressed as a multiplication of two parts: one accounts for variability between species (AF_{spec}), while the second part (AF_{other}) allows for the other uncertainties. The magnitude of each assessment factor is chosen by experts (often working for regulatory organisations) and it is usually a power of 10. The reasoning underlying the chosen value of the assessment factor is rarely explicit and is quite opaque in many cases (EFSA, 2005). An AF is applied to allow for various uncertainties summarized in EFSA (2005) as follows:

- Intra and inter species variation.
- Short-term to long-term toxicity extrapolation.

- Laboratory data to field impact extrapolation.
- Intra and inter laboratory variation of toxicity data.

Scientifically, it is clearly beneficial to test more species as we then learn more about the inter-species variation in sensitivity. This could be captured using the so-called *Species Sensitivity Distributions* (SSDs) model, which means using probability distributions to model the inter-species variation in a particular measure of sensitivity to a chemical substance. Instead, regulation traditionally applies an AF to the minimum measured sensitivity value to obtain the AEC (EFSA, 2005). Clearly this minimum sensitivity value is no larger than the sensitivity value obtained by testing a single species. Since it is also assumed that a lower environmental concentration/dose will be better ecologically, using this minimum sensitivity value when computing AEC leads to reduced ecotoxicological risk. However, the use of minimum sensitivity value has a negative consequence as well. For a person/company seeking to license the use of a chemical, testing an additional species either leaves the minimum measured sensitivity value unchanged or lowers its value. Consequently, the value of AEC will never increase and will often decrease. Therefore, there is no incentive for the person/company to test more than the minimum number of species required by legislation. That minimum number is one in the context we are considering.

It has been realised by EFSA (2005, 2008) that it might be possible to replace the use of the minimum value by the geometric mean of available sensitivity values without actually creating a higher risk to ecosystems. Clearly the geometric mean is no less than the minimum value. However, since practice was not to test more than one species, then in fact the geometric mean, the minimum value, and the value of a single species are all equal in that practice. When multiple species are tested, they also argued that although it is quite possible that the geometric mean is larger than the single value, there will be less variability between samples for the geometric mean than for the single value. In addition, that reduced variability might actually lead to reduced risk, despite the fact that the geometric mean would often be greater than the single value, because it would reduce the probability of producing a very high value for AEC. At the same time, there would clearly be some benefit for the license seeker for the chemical because using the geometric mean would also reduce

the probability of producing a very low value of AEC.

Usually log-scale (base 10) of the original toxicity data will be used to have normal or approximately normal SSDs. The normal distribution is preferable according to its mathematical properties. For example, the empirical rule which states that 68% of the values fall within one standard deviation from the mean, 95% of the values fall within two standard deviation of the the mean, and 99.7% of the values fall within three standard deviation of the mean. In addition, it is symmetric about its mean. Therefore, dividing the geometric mean of the original data by an AF, to obtain AEC, becomes subtracting the logarithm of the AF from the arithmetic mean of the data measured on log-scale (base 10). In what follows, the distribution on log-scale will be referred to as the LSSD.

The adverse consequences, which affect ecological communities in ecosystem, are measured in EFSA (2005) as the proportion of species for which their endpoints are below the specific hazardous concentration, the so called *fraction exceeded* (FE). FE varies from one assessment to another and its value is never known. However, the assessment procedure can be designed to control the expected value of FE known as the *mean fraction exceeded* (MFE). This is called a statistical risk measure. Another measure of statistical risk, which measures the probability that the fraction exceeded (PFE) is greater than some chosen threshold level, is presented in EFSA (2008). Use of PFE is consistent with the approach in Aldenberg and Jaworska (2000).

It was found in EFSA (2005) for a normal LSSD that increasing the sample size while keeping the assessment factor fixed leads to a reduction in the MFE. Later on in EFSA (2008), these two statistical risk measures are numerically demonstrated to be lowered as the number of tested species increases for many different probability distributions.

The unanswered theoretical question is how general this conclusion is, i.e. for which distributions does risk decrease as the sample size increases. The essential point is that there is no scientific evidence yet to allow a decision in favour of any family of distributions other than the normal. Chapter 3 of this thesis studies this mathematical question (see Section 1.2 for more detail).

A Brief History of the SSD Concept

In Ecotoxicological risk assessment, toxicity data that is used in an estimation procedure are the species tolerance values (laboratory effect values or sensitivity values). The term SSD has several meanings:

- The SSD model uses probability distributions to model inter-species variability of some measure of toxicity of a chemical which was first introduced by Kooijman (1987) and Van Straalen and Denneman (1989).
- An empirical SSD distribution is an empirical distribution of measurements of toxicity of a chemical to some sample of species. This came first historically, see Posthuma et al. (2002) for a comprehensive review.
- SSD-based risk assessment uses either (a) the empirical SSD or (b) the SSD model together with measurements for a sample of species to deduce the AEC, as some statistical estimate of some percentile of the SSD, see Aldenberg et al. (2002).

In this thesis, the SSD concept used is that of SSD models.

In SSD-based risk assessment, the SSD is used to determine the AEC of a substance. SSDs are used in two ways forward method and inverse method (Van Straalen, 2002). The forward method is used to estimate the FE, which is also known as the potentially affected fraction (PAF) in Trass et al. (2002) and fraction affected (FA) in Aldenberg and Jaworska (2000). In this method a substance's concentration x is specified in advance and an empirical SSD or a SSD model is used to estimate the proportion of affected species that have an end point which is less than or equal to x, this is equivalent to obtaining cumulative distribution function of the SSD at x ($F_{\text{SSD}}(x)$). On the other hand, the inverse method is used to estimate the threshold or hazardous concentration for which a particular percentage $\alpha\%$ of species have a lower endpoint, the value of α being chosen by the risk manager. This is equivalent to finding the percentile of the SSD.

The SSD methodology was more generally criticised by Forbes and Calow (2002). We refer the reader to Posthuma et al. (2002) for more details. We do not discuss

these further because the research in this thesis relevant to ecotoxicology is not focused on the use of SSDs in risk assessment, but rather on the SSD model assumptions in relation to the geometric mean approach.

1.1.2 Some Concepts Related to Problem 2

In the following, we will introduce some background concepts that are relevant to the wider context of the second problem introduced in Subsection 1.1.3, but which are not core to the mathematical work on the problem presented in Chapter 4 of this thesis.

A Brief Introduction to Expert Knowledge Elicitation

Here, we give the reader a brief introduction to the core ideas of Expert Knowledge Elicitation (EKE). Expert judgment plays an essential role in quantifying uncertainty when available evidence is limited or costly to obtain. The EKE extracts knowledge from one or more experts. In such situations, experts may be asked to give their subjective judgments as probabilities about one or more uncertain quantities. Experts usually express their knowledge as a probability distribution, whether fully or partially, although they can use instead partial ordering, preferences, or estimates for specific values. However, expressing their knowledge, partially or fully, as a probability distribution is the basis of this research as this enables calculations based on standard probability theory to obtain distributional information for derived quantities. A full probability distribution requires specifying probabilities that are associated with all values, whereas partial probability assigns probability of a quantity being within some range(s) of values or exceeding a specific value.

In Bayesian statistics, elicitation is the foundation for constructing the prior distribution, thereby the posterior distribution is obtained via Bayes Theorem. EKE is also used in statistics, economics, engineering, different types of forecasting, and environmental risks, see for example O'Hagan (2012); EFSA (2014); Hanea et al. (2018); and the references therein.

The elicitation process involves one or more experts who have knowledge about an uncertain quantity that would be elicited. There will be also usually an elicitor

(facilitator) who conducts the elicitation exercise and who is qualified and familiar with the elicitation process and understands basic probability theory. There will be one or more experts depending on, for example, availability, resources, or the complexity of the quantities of interest. If EKE is conducted using more than one expert, then their judgements are usually combined in a single probability distribution. This process is called aggregation. Two approaches are often taken:

- Mathematical aggregation (or pooling): The judements are separately elicited, then a probability distribution is fitted to each of them. Pooling rules (mathematical formula) are applied to combine these individual probability distributions mathematically to produce the aggregate distribution. Heterogeneous beliefs due to expert variability are allowed and even preferred in order to increase the quality and credibility of the output by adding more experts especially for complex models. However, the effectiveness of such heterogeneous beliefs depends on the way of combining them, therefore it requires a specific form of pooling rule to produce the aggregate distribution such as a weighted or un-weighted average.

- Behavioral aggregation: This approach aims to reach a consensus. Compromise and persuasion are used by experts to arrive to such consensus, in which an aggregate distribution would be appropriate after a number of experts discuss their knowledge and opinions. This approach results in knowledge sharing between experts. However, some personalities might prevail in the discussion, which leads probability distributions to weigh towards a few or a single expert's judgements. The behavioral approach is also susceptible to group think.

The most well-known protocols used by practitioners of elicitation are briefly discussed in the following.

- Cooke protocol: This protocol applies a mathematical aggregation approach. It is based on weighing expert responses according to their accuracy in assessing distribution(s) for seed variable(s), which is (are) unknown to the experts. More weight is assigned to the expert whose predicted probability distribution

is more accurate. Details of the protocol and the approach to weighting may be found in Cooke (1991).

- Sheffield protocol: SHeffield ELicitation Framework (SHELF), which is used in EFSA (2014), is a behavioral aggregation approach. It is erected on two rounds. In the first round, private judgements are made by experts. In the second round, experts review their judgments. When the elicitor decides that the discussion has reached a point at which no further beneficial contributions can no longer be provided, the experts are asked to come up with joint aggregated judgements, which means they reach an agreement. Although the experts might not share the same opinions following the discussion, they ought to provide reasonable impartial judgements. The optimal number of experts is considered to be four to eight. Experts involved in the elicitation ought to be amenable to the opinions of one another.

- Delphi protocol: This protocol is a combination of behavioural and mathematical aggregation. It involves two or more rounds of judgements. Anonymity is one of its features, so that the providers of the judgements are anonymous. Like the SHELF protocol, there is interaction among experts between rounds where they can share knowledge, but this is limited. At the end individuals provide their judgments individually, and a pooling rule is required across expert final distributions where the aggregated distribution is obtained by given equal weight to individual judgements. In situations where there are strongly differing viewpoints, it may be necessary to involve a larger number of experts. Details of the Delphi approach, including selection of experts may be found in Rowe and Wright (1999).

A detailed discussion on EKE is beyond this research. For more detailed explanation we refer the reader interested in EKE, and different protocols that are used by practitioners of elicitation, to Garthwaite et al. (2005); EFSA (2014); O'Hagan (2019); and references therein.

Interval Analysis

In interval analysis, which is also called interval arithmetic, we specify a range for each input to a calculation or mathematical model and use interval arithmetic to deduce the corresponding set of output values, see Neumaier (1990) and Moore et al. (2009). One simple form is to specify an upper or lower limit for each input.

Imprecise Probability

Williams (1975/2007) and Walley (1991) consider that expressing subjective probabilities precisely as given in De Finetti (1964/1992) is unrealistic. Instead, subjective probability should be bounded by lower and upper probabilities to accommodate a range of opinions. From the standpoint of Walley (1991) these lower and upper bound of probabilities represent the maximum amount that an individual is amenable to pay to sign a contract and the minimum amount that he/she is amenable to receive to abandon signing the contract. When the maximum amount for signing the contract and the minimum amount for abandoning signing it coincide, then the common amount is called a fair amount for signing that contract, and the probability is precise. The partial probability judgements considered in Chapter 4 are a kind of imprecise probability.

When probabilities are specified as ranges rather than numbers, interval analysis can be used to obtain a range of probability as an output from a correct probability calculation. Williams (1975/2007) and Walley (1991) gave a motivation and operational definition of imprecise probability based on upper and lower prices that lead to specifying ranges rather than numbers for probabilities. They also showed that interval analysis is indeed appropriate for calculations. When applied to the distributions of random variables, we get the concept of probability bounds analysis, first through simple application of the Frechet inequalities as in EFSA (2018a) and more generally through the methodology of Tucker and Ferson (2003).

Probability Bounds Analysis

A probability box (p-box) or imprecise probability distribution for a random variable X, where its exact distribution function $F_X(x)$ is unknown, is the class of distri-

bution functions $F_X(x)$ satisfying $\underline{F}_X(x) < F_X(x) < \overline{F}_X(x)\ \forall x$ where the chosen upper and lower cumulative distribution functions $\overline{F}_X(x)$ and $\underline{F}_X(x)$ express imprecision about $F_X(x)$. The distance between these bounds represents the amount of lack of knowledge (epistemic uncertainty) about the unknown distribution function. Probability bounds analysis (PBA) is a combination of the standard interval analysis method and classical probability theory (Ferson et al., 2003). It can be used to evaluate probability boxes for mathematical expressions when there is uncertainty about the input values, their dependencies, or even the distribution shape. For example, it can be used to compute bounds on the distribution resulting from convolution (addition and other arithmetic operations for distributions) or a more complex function, given bounds on the distributions of the inputs (p-boxes). It can be used when the range or bound for probability is provided rather than an exact value. We refer the interested readers to Williamson and Downs (1990); Ferson (2001); Tucker and Ferson (2003); Ferson et al. (2004); and references therein for a detailed review.

1.1.3 Problem 2: Combining Minimal Judgements about Uncertain Quantities

The second problem is related to combining limited probabilistic expert judgements on multiple uncertain quantities in order to provide limited probabilistic information about a derived quantity.

Risk assessment is often concerned with rare events (extreme outcomes), i.e. tails of a distribution, where people have less experience and it is hard to get reliable information (Burgman, 2005). This leads to uncertainty in the analysis.

Making good use of partial probability judgments made by experts has the potential to play a pivotal role in quantifying uncertainty in risk assessment, thereby relieving experts of the need to express uncertainty using full probability distributions.

In many assessments, a model, which is a function of uncertain input quantities, is developed with the output being a proxy for an uncertain quantity of interest in the real world. This mathematical model is called *the assessment calculation*

in EFSA (2018a). This model could be composed of many uncertain quantities (complex model), or rather few uncertain quantities (simple model)(EFSA, 2014). The benefit of the model is that expressing uncertainty about the inputs may be considered easier than directly expressing uncertainty about the output and that uncertainty about the output may be deduced from uncertainty about the inputs. An example of such a model is the Rift Valley Fever Virus model discussed in Chapter 4.

One approach is to specify a full joint (multivariate) probability distribution for the uncertain input quantities. Then we will know the full probability distribution of the model output via the assessment calculation. However, in quantifying uncertainty probabilistically, the most challenging situation is using expert judgment through an EKE process, as discussed earlier, to specify a full joint (multivariate) probability distribution for uncertain quantities, for examples see Daneshkhah and Oakley (2010); O'Hagan (2012); and EFSA (2018b). In practice, it is easier if those uncertain quantities are judged to be independent. Independence substantially reduces the effort of eliciting a joint distribution to eliciting the full probability independently for each uncertain quantity. In subjective probability theory, independence (or epistemic independence) in the judgement of experts between two uncertain quantities means that they would not change their beliefs about the uncertainty of one of them, given (new) knowledge about the other one, and vice versa, see for examples Daneshkhah and Oakley (2010); O'Hagan (2012); EFSA (2018b); and references therein. However, specifying even the full probability distribution of an individual uncertain quantity is hard for experts. Therefore, they could instead provide partial probability statements for the input quantities in the form of probability bounds (Ferson et al., 2003). Probability bounds were introduced in Subsection 1.1.2. The EFSA (2018b) definition of probability bound is to specify an upper bound for the probability that an uncertain quantity exceeds some a specific threshold. However, Ferson's definition is more general.

EFSA (2018a) applies the probability bounds method in two examples: one for purely uncertain parameters and a second for uncertainty about a variable function of variable parameters. The simplest approach to combine probability bounds on

inputs in order to derive a probability bound on the output of an assessment calculation is introduced in EFSA (2018a). It applies to models which do not involve variable quantities, just uncertain parameters. Moreover, it applies when an assessment calculation is monotonic in each input. If the interest is on extreme values for the output, then a 'high' or 'extreme' value is specified for each input, which is called a threshold for the input. In addition, for each input, an upper bound for the probability that the input exceeds the threshold has to be provided by expert judgment, for example through an EKE. Then the assessment calculation is used to combine the threshold values for the inputs to obtain a threshold for the output. Finally, the probability bounds analysis approach is applied to obtain an upper bound on the probability that the output exceeds the calculated threshold. That upper bound is the sum of the individual upper bounds on the probabilities that inputs exceed their threshold values. This approach is valid under all possible assumptions about dependencies and that are consistent with the probability bounds specified for the inputs distributions. However, the upper bound on the probability for the output is intrinsically larger than the upper bounds for the inputs and is likely to be of little use in practice if there are many input quantities or the upper bounds on the probabilities for the inputs are not very small. Although assuming independence or some particular dependence would produce a tighter upper bound on the probability for the output, the probability bounds answer is not greatly improved by assuming independence. For example, for two independent uncertain quantities X_1 and X_2, the upper bound is $P(X_1 + X_2 \geq x_1 + x_2) \leq p_1 + p_2 - p_1 p_2$ where $p_i = P(X_i \geq x_i),\ i = 1, 2$. Note that this upper bound improves on the general upper bound only by the amount $p_1 p_2$ which will usually be much smaller than $p_1 + p_2$. To tighten the upper bound further, one needs additional assumptions about distribution shape as in Chapter 4 of this thesis.

The question arising here is what information we can obtain from experts about distribution shape, in addition to the probability bounds on inputs, that would lead to a useful tighter upper bound for the probability relating to the output of the assessment calculation. We started by considering the situation when X_1 and X_2 are independent and (not necessarily identically) normally distributed as will be dis-

cussed in Example 4.5.1(Section 4.5). We realised that $P(X_1 + X_2 \geq x_1 + x_2)$ is less than the maximum of p_1 and p_2. In addition, if X_1 and X_2 have standard Cauchy distributions and $p_1 = p_2 = p$, $P(X_1 + X_2 \geq x_1 + x_2) = p$, see Example 4.5.2 (Section 4.5). Moreover, if X_1 and X_2 are independent and identically distributed exponential random variables with failure rate one, $P(X_1 + X_2 \geq x_1 + x_2) \approx 0.292 > 0.29$ when $p_1 = p_2 = 0.29$, whereas $P(X_1 + X_2 \geq x_1 + x_2) \approx 0.06 < 0.1$ when $p_1 = p_2 = 0.1$. This means that if $X_i;\ i = 1, 2$ has exponential distribution with failure rate one, the upper bound of $P(X_1 + X_2 \geq x_1 + x_2)$ is in some cases less than the maximum of p_1 and p_2. These results give a motivation to propose $\max(p_1, p_2)$ as a practical upper bound for $P(X_1 + X_2 \geq x_1 + x_2)$ instead of the worst case, $(p_1 + p_2)$, in EFSA (2018a) and the challenge is to establish practical criteria for when this upper bound applies. Consequently, we formulated a working hypothesis and called it "*combined tails dominance property*". This property specifies an upper bound for the probability that the convolution of two independent random variables exceeds a threshold value. The combined tails dominance property is defined as follows

Definition 1.1.1 Two independent random variables X_1 and X_2 satisfy the combined tails dominance property for probabilities p_1 and p_2 if

$$P(X_1 + X_2 \geq x_1 + x_2) \leq \max(p_1, p_2) \text{ where } P(X_i \geq x_i) = p_i;\ i = 1, 2$$

When this property holds, the probability bounds calculation produces a tighter and easily calculated upper bound for the probability for the output than the simple completely general approach introduced in EFSA (2018a). The goal is to find useful conditions on the distributions of X_1 and X_2 (and in general more than two variables) which imply that the property holds, i.e. to understand when the working hypothesis is valid. In particular, we are interested in conditions about which experts might reasonably be able to make judgements.

1.1.4 Proschan's Result

The seminal result by Proschan (1965) will be introduced in detail in Chapter 2. It applies to all symmetric and log-concave densities and leads in Chapter 4 to the validation of the combined tails dominance property when the distributions of

$X_i;\ i = 1, 2$, in problem 2 have symmetric and log-concave densities as will be seen in Section 4.4. In addition, by using Proschan's result, we can show in Chapter 3 that the two statistical risk measures decrease as the number of tested species increases when the LSSDs have symmetric and log-concave densities as will be argued in Section 3.6. The notion of peakedness of probability distributions that underlies Proschan's result is a key mathematical connection between the two problems.

1.2 Outline of Thesis

The thesis is in four chapters. Chapter 2 firstly highlights some properties of log-concave densities that are used throughout the thesis. Secondly, An important lemma in Birnbaum (1948) that is used later by Proschan (1965) to prove his seminal result is introduced in detail. Finally, the seminal result in Proschan (1965), which is used in Sections 3.6 and 4.4, is presented with detailed proof.

Chapter 3 is allocated to show when the dominance properties for the arithmetic mean hold, i.e when the two statistical risk measures in problem 1, namely the mean fraction exceeded and the probability that the fraction exceeded is greater than some chosen threshold level α, decrease as the number of tested species increases. In this thesis, MFE_n and $\text{PFE}_n(\alpha)$ denote respectively the mean fraction exceeded and the probability that the fraction exceeded is greater than some chosen threshold level α when the sample size is n. We will examine the behaviour of MFE_n and $\text{PFE}_n(\alpha)$ when LSSDs have symmetric log-concave densities. In addition, we provide a detailed and comprehensive analysis of what happens as n changes from one to two for MFE_n and $\text{PFE}_n(\alpha)$ when we consider two-component scale and location mixtures of normal LSSDs including a complete mathematical theory of limiting behaviour as parameters tend to extreme values.

In Chapter 4, we assume that the probability bounds in the form $P(X_i \geq x_i) \leq p_i;\ i = 1, 2$ are given, which represent the partial probability judgements about inputs. Then we will explore for which distributions the combined tails dominance property in problem 2 is valid. We examine distributions with symmetric log-concave densities. In addition, we consider negatively-skewed Azzalini-style skew-symmetric

distributions with log-concave kernels when two quantities are involved. Moreover, we extend our investigation to accommodate distributions that have log-concave right tail probability functions where the symmetry of the densities is not necessary. Also, an extension to the case of having more than two uncertain quantities that have log-concave right tail probability functions is presented. At the end of this chapter, we explore the situation where the log-concavity only applies to part of the right tail probability function when two quantities are involved.

We summarise the research discussed above in Chapter 5. Furthermore, future research directions are suggested. A glossary of acronyms that will be used throughout the thesis is presented at the end of the thesis. Finally, two appendices are provided. Appendix A.1 provides an evaluation of some integrals used in Chapter 3. Appendix A.2 gives some information about Lambert function in Chapter 4.

Chapter 2

Log-Concave Probability Density Functions

2.1 Introduction

Proschan (1965) provides a seminal result on peakedness comparison for convex combinations of independent and identically distributed (iid) random variables, $X_i;\ i = 1, 2, \ldots, n$, from a symmetric log-concave density function (see Section 2.3). This result has important interpretation as it implies that $(1/n)\sum_{i=1}^{n} X_i$ is strictly increasing in peakedness as n increases. Proschan's result is sufficient to prove that the dominance properties for the arithmetic mean and the combined tails dominance property, in the two problems discussed in Chapter 1, hold when the underlying distributions have log-concave symmetric densities as we will see in Sections 3.6 and 4.4.

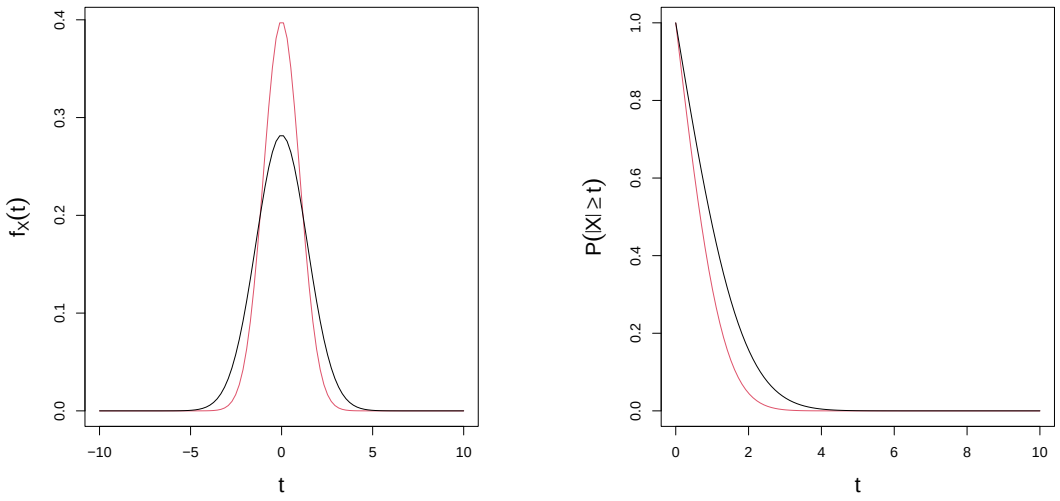
Peakedness was first introduced by Birnbaum (1948) in the following definition.

Definition 2.1.1 (Birnbaum, 1948, Definition). Let X_1 and X_2 be real random variables and x_1 and x_2 real constants. X_1 is more peaked about x_1 than X_2 about x_2 if

$$P\left(|X_1 - x_1| \geq t\right) \leq P\left(|X_2 - x_2| \geq t\right)\ \forall t \geq 0 \tag{2.1.1}$$

If the inequality in Equation 2.1.1 is strict, then X_1 is said to be strictly more peaked about x_1 than X_2 about x_2. In case $x_1 = x_2 = 0$, we simply say X_1 is (strictly)

more peaked than X_2. This definition was generalized to the multivariate case by Sherman (1955). See Figure 2.1 for illustration.



(a) (b)

Figure 2.1: (a) The probability density function of $X_1 \sim \mathrm{N}(\mu = 0, \sigma^2 = 2)$ (black) and $X_2 \sim \mathrm{N}(\mu = 0, \sigma^2 = 1)$ (red), (b) $P(|X_1| \geq t)$ (black) and $P(|X_2| \geq t)$ (red) for $t \geq 0$ to illustrate that X_1 is more peaked than X_2 (Definition 2.1.1).

Proschan's result states that if X_i are iid from symmetric log-concave density function f_X and a and b are vectors such that the elements of each vector are non-negative and sum to one and a strictly majorizes b, as in Definition 2.1.2, then $\sum_{i=1}^{n} b_i X_i$ is strictly more peaked than $\sum_{i=1}^{n} a_i X_i$. This result depends on the notion of majorization which defined as follows:

Definition 2.1.2 (Marshall et al., 2011, A.1. Definition). For any vectors $a, b \in \mathbb{R}^n$,

$$a \succ_m b \text{ if } \begin{cases} \sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i; \ k = 1, \ldots, n-1 \\ \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i \end{cases}$$

When $a \succ_m b$, b is said to be majorized by a, where

$$a_1 \geq \cdots \geq a_n$$

and

$$b_1 \geq \cdots \geq b_n$$

The term strict majorization is used when $\sum_{i=1}^{k} a_i > \sum_{i=1}^{k} b_i;\ k = 1, \ldots, n-1$.

Extensions of Proschan's result are many in the literature, among them is Chan et al. (1987), they considered that the random variables are jointly distributed from a Schur-concave density, whereas the random variables in Ma (1998) are independent and non-identically distributed from a symmetric log-concave density, see the review in Tong (1994). Later on, Ibragimov (2007) showed that the Proschan's result holds for convolutions of α symmetric distributions with $\alpha > 1$. Further studies also extend Proshan's result, see for example Xu and Hu (2011); and Zhao et al. (2011).

This chapter highlights some properties of log-concave functions in Section 2.2. Section 2.3 is devoted to reproducing two important results. The first is a lemma by Birnbaum (1948), which proved that peakedness increases under convolution when specific conditions are satisfied. This lemma is used later by Proschan (1965) in order to prove his seminal result as will be seen in Lemma 2.3.3. The second result is the seminal result by Proschan (1965) with two auxiliary lemmas. All results are reproduced with detailed proofs.

2.2 Log-Concave Functions and Probability Densities

The log-concavity property is considered throughout the following section and chapters. Hence, some log-concave properties are firstly presented.

Definition 2.2.1 A function $f : \mathbb{R} \to [0; \infty)$ is said to be log-concave if

$$f\left((1-t)x_0 + tx_1\right) \geq f(x_0)^{1-t} f(x_1)^t \ \ \forall x_0, x_1 \in \mathbb{R},\ \forall t \in [0,1]$$

or equivalently, a function $\log f$ is concave, i.e.

$$\log f\left((1-t)x_0 + tx_1\right) \geq (1-t)\log f(x_0) + t \log f(x_1) \ \ \forall x_0, x_1 \in \mathbb{R},\ \forall t \in [0,1]$$

This definition implies that any line that connects two points on the graph of $\log f$ must lie below the graph. For example $f(x) = 1 - \frac{x}{10}; 0 \leq x \leq 10$ is a log-concave

function. Therefore, the line that connects $\log f(1)$ and $\log f(8)$, and the line that connects $\log f(2)$ and $\log f(9)$ lie below the curve of $\log f(x) = \log\left(1 - \frac{x}{10}\right)$, see Figure 2.2 for illustration.

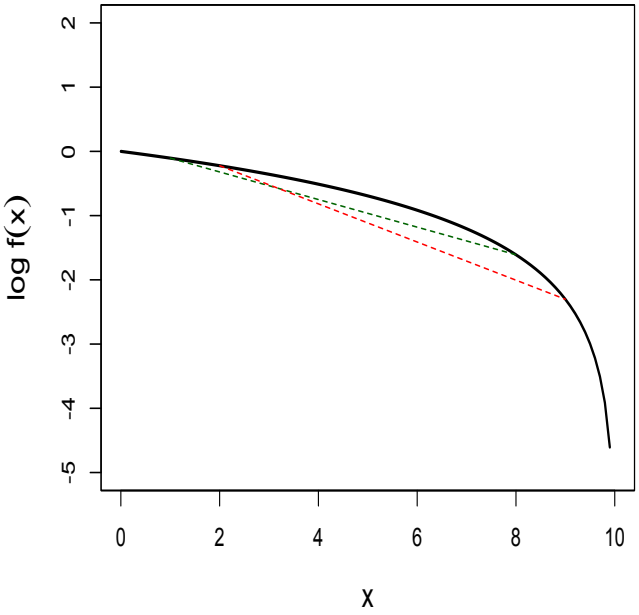


Figure 2.2: A plot of $\log f(x) = \log\left(1 - \frac{x}{10}\right)$. The red dotted line connects $\log f(1)$ and $\log f(8)$. The green dotted line connects $\log f(2)$ and $\log f(9)$.

When we have probability density functions, the following properties are true.

- Log-concave probability densities are unimodal, and may be symmetric or skewed.
- Log-concavity is preserved under convolution of log-concave probability densities.
- The tail of a log-concave density function decays exponentially or faster.
- The right tail probability functions and cumulative distribution functions of log-concave densities are both log-concave.

Examples of such densities are the normal density, gamma densities with shape parameter ≥ 1, and beta densities with both parameters ≥ 1. Many properties

of log-concave densities are discussed in Marshall et al. (1979/2011); An (1998); Saumard and Wellner (2014); and references therein.

An idea that is closely related to log-concave densities is the concept of monotone likelihood ratio that is introduced in the following definition.

Definition 2.2.2 The translation family $f_X(x-\theta);\ \theta \in \mathbb{R}$ has monotone likelihood ratio in x if

$$\frac{f_X(x_1-\theta_2)}{f_X(x_1-\theta_1)} \leq \frac{f_X(x_2-\theta_2)}{f_X(x_2-\theta_1)} \ \forall x_1 < x_2,\ \theta_1 < \theta_2 \tag{2.2.1}$$

which applies to all x in the space of X, where $f_x > 0$ (Meeden, 1971). This property will be used in proving the seminal result by Proschan (1965) since the probability density function f_X is log-concave if and only if $f_X(x-\theta);\ \theta \in \mathbb{R}$ has monotone likelihood ratio in x, as will be seen later in this section.

Another related idea is the idea of Pólya Frequency functions of order 2 (PF_2) (Schoenberg, 1951) which is introduced in the following definition.

Definition 2.2.3 A density function f_X is of Pólya Frequency function of order 2 (PF_2) if

$$\det\begin{pmatrix} f_X(x_1-y_1) & f_X(x_1-y_2) \\ f_X(x_2-y_1) & f_X(x_2-y_2) \end{pmatrix} \geq 0$$

for all $a < x_1 < x_2 < b$ and all $a < y_1 < y_2 < b$.

The term Pólya Frequency function of order 2 (PF_2) is used in Proschan (1965). As will be seen in the following, f_X is PF_2 if and only if it is log-concave.

The following connections between the preceding ideas and the log-concave densities are given in Saumard and Wellner (2014) with their proofs.

(a) f_X is log-concave density function if and only if $f_X(x-\theta), \theta \in \mathbb{R}$ has monotone likelihood ratio in x.

(b) f_X is log-concave density function if and only if f_X is a PF2.

Proof. (a) Suppose f_X is log-concave. $f_X(x-\theta)$ has monotone likelihood ratio in x if and only if

$$\frac{f_X(x_1-\theta_2)}{f_X(x_1-\theta_1)} \leq \frac{f_X(x_2-\theta_2)}{f_X(x_2-\theta_1)} \ \forall x_1 < x_2,\ \theta_1 < \theta_2 \tag{2.2.2}$$

This holds if and only if

$$\log f_X(x_1 - \theta_2) + \log f_X(x_2 - \theta_1) \leq \log f_X(x_2 - \theta_2) + \log f_X(x_1 - \theta_1) \tag{2.2.3}$$

Let $t = \frac{(x_2 - x_1)}{(x_2 - x_1 + \theta_2 - \theta_1)}$, thus

$$x_2 - x_1 = t(x_2 - \theta_1) - t(x_1 - \theta_2)$$

By adding and subtracting θ_1,

$$(x_2 - \theta_1) - (x_1 - \theta_1) = t(x_2 - \theta_1) - t(x_1 - \theta_2)$$

Hence,

$$x_1 - \theta_1 = t(x_1 - \theta_2) + (1 - t)(x_2 - \theta_1)$$

Similarly,

$$x_2 - \theta_2 = (1 - t)(x_1 - \theta_2) + t(x_2 - \theta_1)$$

Since f_X is log-concave,

$$\log f_X(x_1 - \theta_1) \geq t \log f_X(x_1 - \theta_2) + (1 - t) \log f_X(x_2 - \theta_1) \tag{2.2.4}$$

and

$$\log f_X(x_2 - \theta_2) \geq (1 - t) \log f_X(x_1 - \theta_2) + t \log f_X(x_2 - \theta_1) \tag{2.2.5}$$

Adding Equations 2.2.4 and 2.2.5 yields Equation 2.2.3. Therefore, the log-concavity of f_X implies that $f_X(x - \theta)$ has monotone likelihood ratio in x.

Now suppose that $f_X(x - \theta)$ has monotone likelihood ratio in x so that Equation 2.2.3 holds. In particular, it holds if $x_1, x_2, \theta_1, \theta_2$ satisfy

$$x_1 - \theta_2 = a < b = x_2 - \theta_1$$

and

$$t = \frac{x_2 - x_1}{(x_2 - x_1 + \theta_2 - \theta_1)} = \frac{1}{2}$$

Hence,

$$x_2 - x_1 = \theta_2 - \theta_1$$

and

$$x_2 - \theta_2 = x_1 - \theta_1 \tag{2.2.6}$$

Moreover,

$$\frac{a+b}{2} = \frac{(x_1 - \theta_1) + (x_2 - \theta_2)}{2} \tag{2.2.7}$$

By using Equation 2.2.6, thus we obtain

$$x_1 - \theta_1 = \frac{a+b}{2} = x_2 - \theta_2$$

Therefore, Equation 2.2.3 becomes

$$\log f_X(a) + \log f_X(b) \leq 2 \log f_X\Big(\frac{a+b}{2}\Big)$$

Thus f_X is a log-midpoint-concave. In addition, f_X is a Lebesgue measurable function. Therefore f_X is log-concave (Blumberg, 1919; Simon, 2011).

(b) Suppose f_X is PF_2. Thus for $x_1 < x_2, y_1 < y_2$,

$$\det \begin{pmatrix} f_X(x_1 - y_1) & f_X(x_1 - y_2) \\ f_X(x_2 - y_1) & f_X(x_2 - y_2) \end{pmatrix} \geq 0$$

$$\Longleftrightarrow$$

$$f_X(x_1 - y_2) f_X(x_2 - y_1) \leq f_X(x_1 - y_1) f_X(x_2 - y_2)$$

$$\Longleftrightarrow$$

$$\tfrac{f_X(x_1 - y_2)}{f_X(x_1 - y_1)} \leq \tfrac{f_X(x_2 - y_2)}{f_X(x_2 - y_1)}$$

That is $f_X(x-y)$ has monotone likelihood ratio in x. By using (a), this is equivalent to f_X is log-concave. □

In this thesis, we shall use the term log-concave.

2.3 Symmetric Log-Concave Probability Density Functions

This section reproduces the following:

- The lemma by Birnbaum (1948) in Lemma 2.3.1. This lemma is used later by Proschan (1965) (see Lemma 2.3.4).

- Two Auxiliary lemmas that are used by Proschan (1965) to prove his seminal result in Lemmas 2.3.3 and 2.3.4.
- The seminal result by Proschan (1965) in Theorem 2.3.5.

All the lemmas and theorem are reproduced with detailed proof.

Proshan's result is relevant to symmetric log-concave density functions. In addition, it is used later in Sections 3.6 and 4.4 when we consider situations where the LSSD and the distributions of $X_i;\ i = 1, 2$ in respectively problem 1 and problem 2 (see Chapter 1) have symmetric and log-concave density functions.

Lemma 2.3.1 (Birnbaum, 1948, Lemma). Let X_1, X_2, Y_1, Y_2 be continuous random variables with probability densities $f_{X_1}, f_{X_2}, f_{Y_1}, f_{Y_2}$ such that

(1) X_1 and X_2 are independent, Y_1 and Y_2 are independent,

(2) $f_{X_i}(x) = f_{X_i}(-x)\ \ \forall x$ and $f_{Y_i}(y) = f_{Y_i}(-y)\ \ \forall y$ for $i = 1, 2$,

(3) f_{X_2} and f_{Y_1} are not increasing functions for positive values of the variables, and

(4) X_i is more peaked about 0 than Y_i, for $i = 1, 2$.

Let $X = X_1 + X_2$ and $Y = Y_1 + Y_2$. Under these assumptions X is more peaked about 0 than Y.

Proof. Let $F_{X_i}(x) = P(X_i \leq x)$, $F_{Y_i}(y) = P(Y_i \leq y)$, for $i = 1, 2$, be the cumulative distribution functions. For any random variables X_1, X_2, Y_1, Y_2 (not necessarily continuous) where assumption (1) is satisfied, and any t,

$$P(X \leq t) - P(Y \leq t) \ = \ \int_{-\infty}^{\infty} \big[F_{X_1}(t-s) f_{X_2}(s)\, ds - F_{Y_1}(t-s) f_{Y_2}(s)\, ds\big]$$

By adding and subtracting $\int_{-\infty}^{\infty} F_{Y_1}(t-s) f_{X_2}(s)\, ds$, thus

$$\begin{aligned} P(X \leq t) - P(Y \leq t) \ &= \ \int_{-\infty}^{\infty} \left[F_{X_1}(t-s) - F_{Y_1}(t-s)\right] f_{X_2}(s)\, ds \\ &+ \ \int_{-\infty}^{\infty} F_{Y_1}(t-s)\big[f_{X_2}(s)\, ds - f_{Y_2}(s)\, ds\big] \end{aligned} \qquad (2.3.1)$$

Using integration by parts,

$$\int_{-\infty}^{\infty} F_{Y_1}(t-s)\big[f_{X_2}(s)\, ds - f_{Y_2}(s)\, ds\big] = \int_{-\infty}^{\infty} \big[F_{X_2}(s) - F_{Y_2}(s)\big] f_{Y_1}(t-s)\, ds$$

Thus Equation 2.3.1 becomes

$$\begin{aligned} P(X \le t) - P(Y \le t) &= \int_{-\infty}^{\infty} \big[F_{X_1}(t-s) - F_{Y_1}(t-s)\big] f_{X_2}(s)\, ds \\ &\quad + \int_{-\infty}^{\infty} \big[F_{X_2}(s) - F_{Y_2}(s)\big] f_{Y_1}(t-s)\, ds \end{aligned}$$

which is equivalent to

$$\begin{aligned} P(X \le t) - P(Y \le t) &= \int_{-\infty}^{\infty} \big[F_{X_1}(t-s) - F_{Y_1}(t-s)\big] f_{X_2}(s)\, ds \\ &\quad + \int_{-\infty}^{\infty} \big[F_{X_2}(t-s) - F_{Y_2}(t-s)\big] f_{Y_1}(s)\, ds \\ &= I_1(t) + I_2(t) \end{aligned}$$

where

$$\begin{aligned} I_1(T) &= \int_{-\infty}^{\infty} \big[F_{X_1}(t-s) - F_{Y_1}(t-s)\big] f_{X_2}(s)\, ds \\ &= \int_{-\infty}^{\infty} \big[F_{X_1}(-s) - F_{Y_1}(-s)\big] f_{X_2}(t+s)\, ds \\ &= \left(\int_{-\infty}^{0} + \int_{0}^{\infty}\right) \big[F_{X_1}(-s) - F_{Y_1}(-s)\big] f_{X_2}(t+s)\, ds \\ &= A + B \end{aligned} \tag{2.3.2}$$

where

$$A = \int_{-\infty}^{0} \big[F_{X_1}(-s) - F_{Y_1}(-s)\big] f_{X_2}(t+s)\, ds$$

which is equivalent to

$$A = \int_{0}^{\infty} \big[F_{X_1}(s) - F_{Y_1}(s)\big] f_{X_2}(t-s)\, ds \tag{2.3.3}$$

Using Equation 2.3.3, Equation 2.3.2 becomes

$$\begin{aligned} I_1 &= \int_{0}^{\infty} \big[F_{X_1}(s) - F_{Y_1}(s)\big] f_{X_2}(t-s)\, ds \\ &\quad + \big[F_{X_1}(-s) - F_{Y_1}(-s)\big] f_{X_2}(t+s)\, ds \\ &= \int_{0}^{\infty} \big[P(X_1 \le s) - P(Y_1 \le s)\big] f_{X_2}(t-s)\, ds \\ &\quad + \big[P(X_1 \le -s) - P(Y_1 \le -s)\big] f_{X_2}(t+s)\, ds \end{aligned}$$

Using assumption (2),

$$\begin{aligned} I_1 &= \int_0^\infty \Bigg\{ \Big[(1 - P(X_1 > s)) - (1 - P(Y_1 > s)) \Big] f_{X_2}(t-s)\, ds \\ &\qquad + \big[P(X_1 \geq s) - P(Y_1 \geq s)\big] f_{X_2}(t+s)\, ds \Bigg\} \\ &= \int_0^\infty \Bigg\{ \big[P(Y_1 > s) - P(X_1 > s)\big] f_{X_2}(t-s)\, ds \\ &\qquad + \big[P(X_1 \geq s) - P(Y_1 \geq s)\big] f_{X_2}(t+s)\, ds \Bigg\} \end{aligned}$$

By adding and subtracting $\big[P(X_1 = s) - P(Y_1 = s)\big] f_{X_2}(t-s)\, ds$, we obtain

$$\begin{aligned} I_1 &= \int_0^\infty \Bigg\{ \big[P(X_1 \geq s) - P(Y_1 \geq s)\big] \big[f_{X_2}(t+s)\, ds - f_{X_2}(t-s)\, ds\big] \\ &\qquad - \big[P(X_1 = s) - P(Y_1 = s)\big] f_{X_2}(t-s)\, ds \Bigg\} \\ &= \int_0^\infty \big[P(X_1 \geq s) - P(Y_1 \geq s)\big] \big[f_{X_2}(t+s)\, ds - f_{X_2}(t-s)\, ds\big] \\ &\qquad - \int_0^\infty \big[P(X_1 = s) - P(Y_1 = s)\big] f_{X_2}(t-s)\, ds \end{aligned} \tag{2.3.4}$$

Similarly,

$$\begin{aligned} I_2 &= \int_0^\infty \big[P(X_2 \geq s) - P(Y_2 \geq s)\big] \big[f_{Y_1}(t+s)\, ds - f_{Y_1}(t-s)\, ds\big] \\ &- \int_0^\infty \big[P(X_2 = s) - P(Y_2 = s)\big] f_{Y_1}(t-s)\, ds \end{aligned} \tag{2.3.5}$$

Since X_1, X_2, Y_1, Y_2 are continuous random variables, the second integrals in Equations 2.3.4 and 2.3.5 are zero, and

$$I_1 = \int_0^\infty \big[P(X_1 \geq s) - P(Y_1 \geq s)\big] \big[f_{X_2}(t+s) - f_{X_2}(t-s)\big]\, ds \tag{2.3.6}$$

and

$$I_2 = \int_0^\infty \big[P(X_2 \geq s) - P(Y_2 \geq s)\big] \big[f_{Y_1}(t+s) - f_{Y_1}(t-s)\big]\, ds \tag{2.3.7}$$

Using assumption (3), for $t \geq 0$,

$$f_{X_2}(t+s) - f_{X_2}(t-s) \leq 0 \text{ if } 0 \leq s \leq t$$

$$f_{X_2}(t+s) - f_{X_2}(t-s) = f_{X_2}(s+t) - f_{X_2}(s-t) \leq 0 \text{ if } 0 \leq t \leq s$$

Similarly,

$$f_{Y_1}(t+s) - f_{Y_1}(t-s) \leq 0 \text{ for all } t \geq 0 \text{ and } s \geq 0 \tag{2.3.8}$$

Since X_i are more peaked than Y_i, thus for $s \geq 0$,

$$P(X_1 \geq s) - P(Y_1 \geq s) \leq 0$$
$$P(X_2 \geq s) - P(Y_2 \geq s) \leq 0$$

Hence, both integrands in Equations 2.3.6 and 2.3.7 are non negative $\forall s$, thus

$$P(X \leq t) - P(Y \leq t) = I_1(t) + I_2(t) \geq 0$$

Hence,

$$P(X \geq t) - P(Y \geq t) \leq 0 \text{ for } t \geq 0 \tag{2.3.9}$$

and

$$2P(X \geq t) - 2P(Y \geq t) \leq 0 \text{ for } t \geq 0 \tag{2.3.10}$$

Using assumption (2) and Equation 2.3.10,

$$P(X \geq t) + P(X \leq -t) - \Big(P(Y \geq t) + P(Y \leq -t)\Big) \leq 0$$

Hence,

$$P(|X| \geq t) - P(|Y| \geq t) \leq 0 \text{ for } t \geq 0 \tag{2.3.11}$$

Therefore, X is more peaked than Y. □

Lemma 2.3.2 shows that symmetric log-concave functions are non-increasing functions. In addition, it will be used in Lemma 2.3.3.

Lemma 2.3.2 Let f_X be symmetric around zero, log-concave and $0 \leq a \leq b$. Then $f_X(0) \geq f_X(a) \geq f_X(b)$.

Proof. If f_X is log-concave,

$$1/2 \log f_X(-a) + 1/2 \log f_X(a) \leq \log f_X(0)$$

Since f_X is symmetric around zero, then $f_X(-a) = f_X(a)$, thus

$$f_X(a) \leq f_X(0) \tag{2.3.12}$$

Similarly,

$$f_X(b) \leq f_X(0) \tag{2.3.13}$$

Since f_X is log-concave, and for $\alpha \in [0,1]$ such that $\alpha(0)+(1-\alpha)b = a$, the following inequality should be true

$$\log f_X(a) \geq \alpha \log f_X(0) + (1-\alpha)\log f_X(b)$$

Using Equation 2.3.13,

$$\begin{aligned} \alpha \log f_X(0) + (1-\alpha)\log f_X(b) &\geq \alpha \log f_X(b) + (1-\alpha)\log f_X(b) \\ &= \log f_X(b) \end{aligned}$$

Thus,

$$f_X(a) \geq f_X(b) \tag{2.3.14}$$

Using Equations 2.3.12 and 2.3.14,

$$f_X(0) \geq f_X(a) \geq f_X(b)$$

□

The seminal result obtained by Proschan (1965) is presented in Theorem 2.3.5. In addition, he presented two lemmas in order to prove his theorem. However, Proschan's proof for the first lemma is concise. Therefore, we will reintroduce the two lemmas in Lemmas 2.3.3 and 2.3.4 with detailed proof for the first lemma.

Lemma 2.3.3 (Proschan, 1965, Lemma 2.1). Let f_X be a symmetric log-concave density function, X_1 and X_2 independently distributed with density f_X. Then $aX_1 + (1-a)X_2$ is strictly increasing in peakedness as a increases from 0 to $\frac{1}{2}$

Proof. Fix $t > 0$ and define

$$H(a,t) = P\left(aX_1 + (1-a)X_2 \leq t\right) = \int_{-\infty}^{\infty} F_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)\, du; \;\; 0 < a < \frac{1}{2}$$

Differentiation under the integral sign is permissible since the derivative of $F_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)$ is bounded by a Lebesgue integrable function: $\left|f_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)(u-t)\right| \leq M f_X(u)\,|u-t|$ where M is the mode of f_X, and

$\int_{-\infty}^{\infty} M f_X(u)(u-t)\,du < \infty$ because all moments exist for a log-concave density (Chen and Samworth, 2013). Then

$$a^2\left(\frac{\partial H(a,t)}{\partial a}\right) = \int_{-\infty}^{\infty} f_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)(u-t)\,du \tag{2.3.15}$$

Rewrite Equation 2.3.15,

$$\begin{aligned} a^2\left(\frac{\partial H(a,t)}{\partial a}\right) &= \int_{-\infty}^{t} f_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)(u-t)\,du \\ &+ \int_{t}^{\infty} f_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)(u-t)\,du \end{aligned}$$

Set $v = t-u$ in the first integral and $v = u-t$ in the second integral. We get

$$\begin{aligned} \frac{a^2\partial H(a,t)}{\partial a} &= \int_0^{\infty} v\Bigg\{ f_X\big(t+v\big) f_X\left(t-\frac{(1-a)}{a}v\right) \\ &\qquad -f_X\left(t-v\right) f_X\left(t+\frac{(1-a)}{a}v\right)\Bigg\}\,dv \end{aligned} \tag{2.3.16}$$

By symmetry of f_X,

$$\begin{gathered} f_X\left(t+v\right) f_X\left(t-\tfrac{(1-a)}{a}v\right) - f_X\left(t-v\right) f_X\left(t+\tfrac{(1-a)}{a}v\right) \\ = \\ f_X\left(t+v\right) f_X\left(\tfrac{(1-a)}{a}v-t\right) - f_X\left(v-t\right) f_X\left(t+\tfrac{(1-a)}{a}v\right) \end{gathered}$$

Since f_X is log-concave,

$$\frac{f_X\left(v-t\right)}{f_X\left(v+t\right)} \leq \frac{f_X\left(\frac{(1-a)}{a}v-t\right)}{f_X\left(\frac{(1-a)}{a}v+t\right)} \tag{2.3.17}$$

for $1-a > a,\ v > 0$ because $v < \frac{(1-a)}{a}v$ and $-t < t$, and

$$f_X\left(v-t\right) f_X\left(\frac{(1-a)}{a}v+t\right) \leq f_X\left(\frac{(1-a)}{a}v-t\right) f_X\left(v+t\right) \tag{2.3.18}$$

So that the integrand in Equation 2.3.16 is not negative. Hence, $\frac{a^2\partial H(a,t)}{\partial a} \geq 0$, so that $\frac{\partial H(a,t)}{\partial a} \geq 0$.

Now suppose $\frac{\partial H(a,t)}{\partial a} = 0$. Since f_X is log-concave, the integrand has at most a finite number of points of discontinuity because the density is continuous except at the boundary of its support. Therefore, the integrand must be zero almost everywhere.

We will consider two cases for the support of f_X as follows:

Case (1) Finite support as given in Proschan's proof:

Consider the support of f_X is an interval $(-l, l)$. Then we need only consider $t < l$ to make Proschan's proof work.

Define $k = (1-a)/a$, and consider $v \in ((l-t)/k, \min((l+t)/k, l-t))$. Then we will examine if $t - kv,\ t - v,\ t + v,\ t + kv$ are within the interval $(-l, l)$ for the specified range of values of v as follows:

$$\begin{aligned} l > t - kv > t - k(l+t)/k = -l &\implies f_X(t-kv) > 0 \\ l > t - v > t - (l-t) = 2t - l > -l &\implies f_X(t-v) > 0 \\ 0 < t + v < t + (l-t) = l &\implies f_X(t+v) > 0 \\ t + kv > t + k(l-t)/k = l &\implies f_X(t+kv) = 0 \end{aligned}$$

So that the integrand is not zero for such v.

We have shown that there is an interval where the integrand is non-zero and this contradicts the requirement that the integrand is zero almost everywhere if $\frac{\partial H(a,t)}{\partial a} = 0$. Therefore $\frac{\partial H(a,t)}{\partial a} > 0$ when f_X has finite support.

Case (2) Support is real line:

In this case, we make the support of f_X wider than the one proposed by Proschan's proof. let $v \in (0, t/k)$. Then $0 < t - kv$ and the following relationships are true

$$0 < t - kv < t - v < t + v < t + kv$$

Using Lemma 2.3.2, when $0 < t - kv < t - v < t + v < t + kv$ we have

$$f_X(0) \geq A \geq B \geq C \geq D$$

where $A = f_X(t - kv),\ B = f_X(t - v),\ C = f_X(t + v)$ and $D = f_X(t + kv)$. If $C > D$, then $CA > BD$ and the integrand $CA - BD$ in Equation 2.3.16 is positive. Whereas if $C = D$ implies that $A = B$ since the integrand is supposed to be zero. Moreover, $f_X(0) = A = B = C = D$ since f_X is log-concave, and for $\alpha \in (0, 1)$ such that $\alpha(0) + (1-\alpha)(t + kv) = t + v$, the following inequality should be true

$$\log C \geq \alpha \log f_X(0) + (1-\alpha) \log D$$

Since $C = D$,

$$\alpha \log D \geq \alpha \log f_X(0)$$

However, $0 < t + kv$, thus $f_X(0)$ must be equal to D.

Hence, for the integrand to be zero almost everywhere, f_X must be constant for the specified range of values of v. Therefore, we obtain the support of f_X as follows: the maximum value for f_X is corresponding to the maximum value of v, i.e

$$t + kv = t + k(t/k) = 2t$$

and the minimum value of f_X is $t - k(t/k) = 0$. Consequently, f_X must be constant on the interval $(0, 2t)$.

Now suppose that we know that f_X is constant on some interval $(-\gamma, \gamma)$ where $\gamma \geq 2t$. Consider $v \in (0, \min((\gamma + t)/k, \gamma - t))$. Then

$$\begin{aligned}
0 < v < \gamma - t \rightarrow 0 < t + v < \gamma \Longrightarrow \quad & f_X(t + v) > 0 \\
0 < v < \gamma - t \rightarrow \gamma > t - v > t - (\gamma - t) = 2t - \gamma > -\gamma \Longrightarrow \quad & f_X(t - v) > 0 \\
\gamma > t - kv > t - k(\gamma + t)/k = -\gamma \Longrightarrow \quad & f_X(t - kv) > 0
\end{aligned}$$

So that $f_X(0) = A = B = C > 0$ and so $D = f_X(0) = A = B = C$ for the integrand to be zero. The maximum value of $t + kv$ for the specified range of values of v is

$$\begin{aligned}
\min\left(t + k(\gamma + t)/k, t + k(\gamma - t)\right) &= \min\left(2t + \gamma, k(\gamma - t) + t\right) \\
&= \gamma + \min\left(2t, (k - 1)(\gamma - t)\right)
\end{aligned}$$

Since $\gamma - t \geq t$,

$$\begin{aligned}
\gamma + \min\left(2t, (k - 1)(\gamma - t)\right) &\geq \gamma + \min\left(2t, (k - 1)t\right) \\
&= \gamma + \min\left(2, k - 1\right) t \\
&= \gamma + \Delta t
\end{aligned}$$

Hence, $\min\left(t + k(\gamma + t)/k, t + k(\gamma - t)\right) \geq \gamma + \Delta t$ where $\Delta = \min\big(2, k - 1\big) > 0$. Therefore, f_X should be constant on the interval $(-(\gamma + \Delta t), (\gamma + \Delta t))$.

We conclude that, when f_X is supposed to be constant on some interval, f_X should be constant on wider than that interval by $2\Delta t$.

By repeating this process, we conclude that f_X is constant on the interval $(-(\gamma + n\Delta t), (\gamma + n\Delta t))$ for any positive integer n, and so f_X is constant on the whole real line which is impossible. This contradiction means that the integrand cannot be zero almost everywhere and therefore that $\frac{\partial H(a,t)}{\partial a} > 0$. □

Lemma 2.3.4 (Proschan, 1965, Lemma 2.2). Let f_X be a symmetric log-concave density function, $X_1, \ldots, X_n$ independently distributed with density f_X. Then $\sum_{i=1}^{n} a_i X_i$ is strictly increasing in peakedness as a_1 increases from 0 to $\frac{1}{2}c$, with $a_1 + a_2 = c$, $0 < c \leq 1$, $a_i \geq 0, i = 1, \ldots, n$ and $\sum_{i=1}^{n} a_i = 1$

Proof. First note that $\sum_{i=1}^{2} a_i X_i$ and $\sum_{i=1}^{n} a_i X_i$ are each symmetric unimodal random variables since each X_i is, see Kanter (1977). Suppose that $a_1 < b_1$, $a_1 < a_2$,$b_1 < b_2$, $a_1 + a_2 = c = b_1 + b_2$. Then by Lemma 2.3.3, $a_1X_1 + a_2X_2$ is less peaked than $b_1X_1 + b_2X_2$. By Lemma 2.3.1, it follows that $\sum_{i=1}^{n} a_i X_i$ is less peaked than $\sum_{i=1}^{2} b_i X_i + \sum_{i=3}^{n} a_i X_i$. Finally, the strictness in Lemma 2.3.4 is because of the strictness in Lemma 2.3.3. □

To state Theorem 2.3.5, if a vector $b = (b_1, \ldots, b_n)$ is majorized by vector $a = (a_1, \ldots, a_n)$, then b can be derived from a by a finite number of transformations T of the form

$$T(a) = \alpha(a_1, \ldots, a_n) + (1 - \alpha)(a_1, \ldots, a_{j-1}, a_k, a_{j+1}, \ldots, a_{k-1}, a_j, a_{k+1}, \ldots, a_n)$$

where $0 \leq \alpha \leq 1$, see Hardy et al. (1934).

Theorem 2.3.5 (Proschan, 1965, Theorem 2.3). Let f_X be a symmetric log-concave density function, $X_1, \ldots, X_n$ independently distributed with density f_X, $a \succ_m b$, a, b not identical,$\sum_{i=1}^{n} a_i = 1 = \sum_{i=1}^{n} b_i$. Then $\sum_{i=1}^{n} b_i X_i$ is strictly more peaked than $\sum_{i=1}^{n} a_i X_i$. i.e

$$P\left(\left|\sum_{i=1}^{n} b_i X_i\right| \geq t\right) < P\left(\left|\sum_{i=1}^{n} a_i X_i\right| \geq t\right);\ t \geq 0 \tag{2.3.19}$$

Proof. b can be obtained from a by a finite number of T transformations. Applying Lemma 2.3.4 in each case, therefore the theorem is proved. □

2.4 Conclusions

We have highlighted the lemma by Birnbaum (1948), which proved that peakedness increases under convolution when specific conditions are satisfied, used later by Proschan (1965) in proving his seminal result on peakedness comparison for convex combinations of independent and identically distributed random variables from a symmetric log-concave density function. The Birnbaum's lemma and Proschan's seminal result are reproduced with detailed proofs.

Chapter 3

Dominance of Arithmetic Mean for Repeated Sampling in Some Risk Problems

3.1 Introduction

This chapter is devoted to theoretically generalize, to a wider range of distribution shapes, the dominance properties, first introduced in EFSA (2005, 2008) and described in Subsection 1.1.1, for the arithmetic mean for LSSDs as the number of tested species increases. The chapter opens by introducing an example of practical application by EFSA in Section 3.2. Section 3.3 defines the two statistical risk measures used, including methods of calculation. In Section 3.4, we consider the case when the LSSD is normal, and in Section 3.5 discuss location-scale families of distributions. In Section 3.6, the generalization to all LSSDs with symmetric log-concave densities is presented. Subsequently, LSSDs in the form of mixtures of symmetric log-concave densities acquired our attention. Sections 3.7 and 3.8 are respectively devoted to two-component scale and location mixtures of normal LSSDs. They include a comprehensive analysis of what happens to the dominance properties as parameters tend to extreme values when the number of tested species increases from one to two. The chapter closes with some conclusions in Section 3.9. A glossary of acronyms, including those used in this chapter, is presented at the end of the thesis.

3.2 Example of Application by EFSA

The fungicide thiophanate-methyl and its metabolite carbendazim are used on many crops. However, their residues could have harmful effects on the environment. Therefore, risk assessments are essential for the safety of various groups of organisms. In EFSA (2018c), the risk assessment was made for the effects of the metabolite carbendazim on fish. Five species of fish were tested: ictalurus punctatus, oncorhynchus mykiss, cyprinodon variegatus, lepomis macrochirus, and cyprinus carpio. They were exposed to carbendazim for 96 hours and the concentration, in milligrams per liter of water (mg/L), that leads to the mortality of 50% of each of the five species was determined as indicated in Table 3.1

Fish	Toxicity (mg/L)
Ictalurus punctatus	0.019
Oncorhynchus mykiss	0.54
Cyprinodon variegatus	> 1.158
Lepomis macrochirus	> 3.2
Cyprinus carpio	0.44

Table 3.1: Toxicity data, in (mg/L), for five tested species of fish that were exposed to carbendazim for 96 hours. Reproduced from EFSA (2018d, Toxicity data for all aquatic tested species).

Following standard first tier regulatory practice, the minimum value of the toxicity data in Table 3.1 is 19 μg/L (micrograms per liter of water) and, applying the standard assessment factor of 100, the AEC is 0.19 μg/L. However, for the higher tier risk assessment, the geometric mean for the toxicity data of the above five fish is calculated to be 441 μg/L and, by applying the assessment factor of 100, the AEC is 4.41 μg/L. As discussed in Subsection 1.1.1, the basis for using the higher tier AEC as the final value in the risk assessment was established in EFSA (2005) assuming that the LSSD was normal and considering the two statistical risk measures: the mean fraction exceeded and the probability that the fraction exceeded exceeds a

specified threshold α. The claim is not that replacing the minimum by the geometric mean leads to reduced risk when more than one species is tested; clearly it does not since the geometric mean is always greater than or equal to the minimum. Rather, the justification is that the regulations require only that a single species be tested, that the minimum is equal to the geometric mean when a single species is tested, and that the risk using the geometric mean decreases with increasing sample size for both statistical risk measures.

Note that, as described in Subsection 1.1.1, using log-scale base 10 for the original toxicity data means that an equivalent higher tier calculation to obtain the AEC is to subtract $\log_{10} 100 = 2$ from the arithmetic mean of the log base 10 of the above five toxicity data which is $2.64 \log_{10} \mu$g/L. Hence, the AEC is $0.64 \log_{10} \mu$g/L. This quantity will be denoted later by Y_n where here $n = 5$.

3.3 Background

Denote by $S_1, \ldots, S_n$ the results of testing a substance on n species, where each S_i is the concentration of the substance that causes a specified effect (endpoint). As discussed in Subsection 1.1.1, a common statistical model in ecotoxicology is that $S_1, \ldots, S_n$ are a random sample from the SSD. On taking the logarithm (base 10), the geometric mean of $S_1, \ldots, S_n$ becomes the arithmetic mean, $\bar{X}_n$, of $X_1 = \log_{10} S_1, \ldots, X_n = \log_{10} S_n$, where the values X_i are randomly sampled from the LSSD. Moreover, dividing the geometric mean by an AF becomes subtracting $c = \log(\text{AF})$ from $\bar{X}_n$. Hence, $Y_n = \bar{X}_n - c$ is the statistic that addresses inter-species variability. The fraction exceeded $\text{FE}_n = F_X(Y_n)$ is the fraction of species having an endpoint that is less than Y_n, and it is a measure of the fraction of species at risk. The FE_n is in general not observed. However, it is possible to make mathematical statements about its behaviour under random sampling. We shall use the notation MFE_n, as in EFSA (2008), to stand for the mean fraction exceeded when n species are tested. Moreover, we will use $\text{PFE}_n(\alpha)$ to stand for the probability that FE_n is greater than some chosen level α.

The mathematical formulas of MFE_n and $\text{PFE}_n(\alpha)$ are given below:

MFE_n

The Mean Fraction Exceeded MFE_n is the expected value of $\text{FE}_n = F_X(Y_n)$, i.e

$$\begin{aligned} \text{MFE}_n &= E\big(F_X(Y_n)\big) \\ &= \int_{-\infty}^{\infty} F_X(y) f_{Y_n}(y)\, dy \end{aligned} \tag{3.3.1}$$

$\text{PFE}_n(\alpha)$

The probability that the fraction exceeded FE_n exceeds some chosen level α is

$$\text{PFE}_n(\alpha) = P\big(F_X(Y_n) > \alpha\big) \tag{3.3.2}$$

$$= P\big(\bar{X}_n > F_X^{-1}(\alpha) + c\big) \tag{3.3.3}$$

In addition, EFSA (2008) requires that c is large enough to guarantee that

$$F_X(\mu - c) < \alpha \tag{3.3.4}$$

where μ, σ, and F_X are the mean, standard deviation, and the cumulative distribution function of X respectively. This is not an unreasonable requirement because $c = \mu - F_X^{-1}(\alpha)$ is what is required to move the expected value of $\bar{X}_n$ to the α quantile of the LSSD, the quantile that is being targetted.

In the rest of this chapter we are looking at which conditions on the distributions lead to MFE_n or $\text{PFE}_n(\alpha)$ decreasing as we increase n. In EFSA (2005), a normal distribution is assumed for the LSSD. Therefore, we first consider the case when X is normally distributed in Section 3.4.

3.4 Normal Distribution LSSDs

The normal distribution is parameterized by the mean μ and variance σ^2. If $X \sim N(\mu, \sigma^2);\ \mu \in \mathbb{R}, \sigma > 0$, then the probability density and the cumulative distribution functions are respectively given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right);\ -\infty < x < \infty$$

and

$$F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right);\ -\infty < x < \infty \tag{3.4.1}$$

where Φ is the cumulative distribution function of a standard normal random variable, whose density function is

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right);\ -\infty < x < \infty$$

It is straightforward to show that $Y_n = \bar{X}_n - c$ has a normal distribution with mean $\mu - c$ and variance σ^2/n, as it is a linear function of the sample mean drawn from the normal distribution with mean μ and variance σ^2. Y_n has probability density function

$$f_{Y_n}(y) = \frac{\sqrt{n/2\pi}}{\sigma} \exp\left(\frac{-n\big(y-(\mu-c)\big)^2}{2\sigma^2}\right);\ -\infty < y < \infty$$

The mean fraction exceeded when the distribution on the log-scale is normal is evaluated as follows

Using Equation 3.3.1,

$$\begin{aligned}
\text{MFE}_n &= \frac{\sqrt{n/2\pi}}{\sigma}\int_{-\infty}^{\infty} \Phi\left(\frac{y-\mu}{\sigma}\right) \exp\left(\frac{-n(y-(\mu-c))^2}{2\sigma^2}\right) dy \\
&= \frac{\sqrt{n/2\pi}}{2\sigma}\int_{-\infty}^{\infty} \left(1+\text{erf}\left(\frac{y-\mu}{\sqrt{2}\sigma}\right)\right) \exp\Big(\frac{-n(y-(\mu-c))^2}{2\sigma^2}\Big) dy \\
&= \frac{1}{2} + \frac{\sqrt{n/2\pi}}{2\sigma}\int_{-\infty}^{\infty} \text{erf}\Big(\frac{y-\mu}{\sqrt{2}\sigma}\Big) \exp\Big(\frac{-n(y-(\mu-c))^2}{2\sigma^2}\Big) dy
\end{aligned}$$

Using the result from Appendix A.1,

$$\begin{aligned}
\text{MFE}_n &= \frac{1}{2}\left(1+\text{erf}\left(\frac{-c}{\sigma\sqrt{2(1+\frac{1}{n})}}\right)\right) \\
&= \Phi\left(\frac{-c}{\sigma\sqrt{1+\frac{1}{n}}}\right)
\end{aligned}$$

It is clear for $c > 0$ that MFE_n decreases as n increases.

In the following, we will examine the behaviour of the probability that the fraction exceeded exceeds a specific α when the number of tested species increases in the case that LSSD is normal, provided that the condition in Equation 3.3.4 is satisfied. Recall Equation 3.3.2

$$\text{PFE}_n(\alpha) = P\big(F_X(Y_n) > \alpha\big)$$

Since $F_X(x) = \Phi(\frac{x-\mu}{\sigma})$,

$$\begin{aligned} \text{PFE}_n(\alpha) &= P\left(\Phi\left(\frac{Y_n - \mu}{\sigma}\right) > \alpha\right) \\ &= P\left(Y_n > \mu + \sigma\Phi^{-1}(\alpha)\right) \\ &= 1 - \Phi\left(\frac{\sqrt{n}(\sigma\Phi^{-1}(\alpha) + c)}{\sigma}\right) \end{aligned} \tag{3.4.2}$$

By using Equations 3.3.4 and 3.4.1, thus we obtain

$$F_X(\mu - c) = \Phi\left(\frac{-c}{\sigma}\right) < \alpha$$

Consequently, $\sigma\Phi^{-1}(\alpha) + c > 0$. In addition, $\Phi\left(\frac{\sqrt{n}\left(\sigma\Phi^{-1}(\alpha)+c\right)}{\sigma}\right)$ in Equation 3.4.2 is increasing as n increases. As a result, $\text{PFE}_n(\alpha)$ decreases as n increases provided that $c > \mu - F_X^{-1}(\alpha)$.

3.5 Location-Scale Family of Distributions

We now show the stability of the domainance properties of the two statistical measures under changes to the location and/or scale of a distribution.

A location-scale family of distributions is formed by starting with a standard probability density function f_Z and considering the linear transformation $X = \mu + \sigma Z$, where $Z \sim f_Z$, $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$. μ is the location parameter which is responsible for shifting the graph on the horizontal line to the right or left. The scale parameter σ is responsible for stretching or compressing the probability density function. The probability density function of X and its cumulative distribution function are then

$$\begin{aligned} f_X(x) &= \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right) \\ F_X(x) &= F_Z\left(\frac{x-\mu}{\sigma}\right) \end{aligned} \tag{3.5.1}$$

Proposition 3.5.1 $E\left(F_Z\left(\bar{Z}_n - c^*\right)\right)$ and $P\left(F_Z\left(\bar{Z}_n - c^*\right) > \alpha\right)$ decrease, as n increases, if and only if $E\left(F_X\left(\bar{X}_n - c^*\sigma\right)\right)$ and $P\left(F_X\left(\bar{X}_n - c^*\sigma\right) > \alpha\right)$ respectively decrease.

Proof: It is sufficient to prove that $F_Z\left(\bar{Z}_n - c^*\right) = F_X\left(\bar{X}_n - c^*\sigma\right)$ in order to prove this proposition. By using Equation 3.5.1,

$$\begin{aligned} F_Z\left(\bar{Z}_n - c^*\right) &= F_X\left(\left(\bar{Z}_n - c^*\right)\sigma + \mu\right) \\ &= F_X\left(\left(\frac{\sum_{i=1}^n X_i - n\mu}{n\sigma} - c^*\right)\sigma + \mu\right) \\ &= F_X\left(\bar{X}_n - c^*\sigma\right) \end{aligned}$$

Therefore, we conclude that if either dominance property holds for F_Z, then it holds for any distribution in the same location-scale family. This happens by re-scaling c^* using the scale parameter of that distribution. For example, if we had proved only that the dominance properties of the arithmetic mean hold for the standard normal distribution, then they hold for any normal distribution with mean $\mu \in \mathbb{R}$ and standard deviation $\sigma > 0$. □

3.6 Symmetric Log-Concave Probability Density Functions

In this section, we consider distributions with symmetric and log-concave densities. We shall exploit the seminal result obtained by Proschan (1965), reproduced in Section 2.3, to prove the following theorem.

Theorem 3.6.1 Let $X_1, \ldots, X_n$ be independently distributed with symmetric log-concave density f_X. As n increases, MFE_n decreases and $\text{PFE}_n(\alpha)$ decreases provided that $c > -F_X^{-1}(\alpha)$.

Proof. First: To prove MFE_n decreases as n increases, we argue as follows:
Recall first that $Y_n = \bar{X}_n - c$. Now

$$\begin{aligned} \text{MFE}_n = E\left(F_X\left(Y_n\right)\right) &= E\left(P\left(X \leq Y_n | Y_n\right)\right) \\ &= P\left(X \leq Y_n\right) \end{aligned}$$

where X is an independently sampled value from f_X. Let $a_1 = \cdots = a_{n-1} = 1/(n-1), a_n = 0$ and $b_1 = \cdots = b_n = 1/n$ in Theorem 2.3.5. Hence,

$$P\left(\left|\bar{X}_n\right| \geq t\right) < P\left(\left|\bar{X}_{n-1}\right| \geq t\right) \quad \forall\, t \geq 0 \tag{3.6.1}$$

Moreover, log-concavity is preserved under convolution, therefore $\sum_{i=1}^{k} X_i$ is log-concave. In addition, the linear transformation of log-concave random variables is log-concave. Consequently, $-\bar{X}_k$ is log-concave. Furthermore, the symmetry of $-\bar{X}_k$ results from the symmetry of $\bar{X}_k$. Therefore, $X - \bar{X}_n$ has the same distribution as $X + \bar{X}_n$. Let $a = \left(\frac{1}{2}, \frac{1}{2(n-1)}, \ldots, \frac{1}{2(n-1)}, 0\right)$ and $b = \left(\frac{1}{2}, \frac{1}{2n}, \ldots, \frac{1}{2n}\right)$ in Theorem 2.3.5,

$$P\left(\frac{\left|X + \bar{X}_n\right|}{2} \geq t\right) < P\left(\frac{\left|X + \bar{X}_{n-1}\right|}{2} \geq t\right) \quad \forall\, t \geq 0$$

Hence,

$$P\left(\left|X + \bar{X}_n\right| \geq 2t\right) < P\left(\left|X + \bar{X}_{n-1}\right| \geq 2t\right) \quad \forall\, t \geq 0$$

which is equivalent to

$$P\left(\left|X - \bar{X}_n\right| \geq 2t\right) < P\left(\left|X - \bar{X}_{n-1}\right| \geq 2t\right) \quad \forall\, t \geq 0 \tag{3.6.2}$$

By symmetry of $X - \bar{X}_k$, Equation 3.6.2 becomes

$$P\left(X - \bar{X}_n \leq -2t\right) < P\left(X - \bar{X}_{n-1} \leq -2t\right)$$

Set $t = c/2$, thus we obtain

$$P\left(X - \bar{X}_n + c \leq 0\right) < P\left(X - \bar{X}_{n-1} + c \leq 0\right)$$

and

$$P\left(X \leq Y_n\right) < P\left(X \leq Y_{n-1}\right)$$

Therefore, MFE_n decreases as n increases.

Second: To prove $\mathrm{PFE}_n(\alpha)$ decreases as n increases, set $t = F_X^{-1}(\alpha) + c > 0$ in Equation 3.6.1, and note symmetry of $\bar{X}_k$. Thus

$$P\left(\bar{X}_n \geq F_X^{-1}(\alpha) + c\right) < P\left(\bar{X}_{n-1} \geq F_X^{-1}(\alpha) + c\right)$$

which is equivalent to

$$P\left(F_X(Y_n) \geq \alpha\right) < P\left(F_X(Y_{n-1}) \geq \alpha\right)$$

Therefore, $\mathrm{PFE}_n(\alpha)$ decreases as n increases. □

In Sections 3.7 and 3.8, we consider the two-component scale mixture of normal

distributions and the two-component location mixture of normal distributions respectively. We consider the two-component scale mixture of normal distributions as a way of broadening to families of distributions that are heavy tailed distributions. The two-component location mixture of normal distributions is considered as away of addressing skewness. In addition, bimodality of the empirical SSD is common (Fox et al., 2021). Moreover, we chose the mixture of normals due to the possibility of analytical calculations.

3.7 A Two-Component Scale Mixture of Normal Distributions

In this section, we consider that the distribution of X is a scale mixture of normal (SMN) distribution having two components. Thus
$X|\sigma^2 \sim \mathrm{N}(0, \sigma^2)$ and $\sigma^2 \sim 1 + \mathrm{Bern}(1-w)\Delta$, where $\Delta = \psi - 1 > 0$, and the distribution of $\mathrm{Bern}(1-w)$ is the Bernoulli distribution that has probability mass function

$$\begin{cases} 1 & \text{with probability } 1-w;\ 0 \leq w \leq 1 \\ 0 & \text{with probability } w \end{cases}$$

The mixture components both have zero means and the variances are 1 and $\psi > 1$. The probability density function of $X \sim \mathrm{SMN}(w, 0, 1, \psi)$ is

$$f_X(x) = w\phi(x) + \frac{(1-w)}{\sqrt{\psi}}\phi\left(\frac{x}{\sqrt{\psi}}\right);\ -\infty < x < \infty$$

3.7.1 MFE_n

In this subsection, we first obtain MFE_n when the distribution of X is a mixture of two normal distributions on both the location and scale parameters, i.e. $X \sim \mathrm{MN}(w, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ where $\mu_1, \mu_2 \in \mathbb{R}$; $\sigma_1^2, \sigma_2^2 > 0$; and $0 \leq w \leq 1$ are location, scale, and weight parameters respectively. Second, we obtain MFE_n when $X \sim \mathrm{SMN}(w, 0, 1, \psi)$ as a special case. The probability density and characteristic functions of $X \sim \mathrm{MN}(w, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$ are given respectively as

$$f_X(x) = \frac{w}{\sigma_1}\phi\left(\frac{x-\mu_1}{\sigma_1}\right) + \frac{1-w}{\sigma_2}\phi\left(\frac{x-\mu_2}{\sigma_2}\right) \tag{3.7.1}$$

and

$$\Psi_X(t) = w\Psi_1(t) + (1-w)\Psi_2(t);\ t \in \mathbb{R} \tag{3.7.2}$$

where

$$\Psi_j(t) = \exp\left(it\mu_j - \frac{t^2\sigma_j^2}{2}\right);\ j = 1, 2 \tag{3.7.3}$$

Let $X_1, \ldots, X_n$ be independent and identically distributed from the distribution with density function in Equation 3.7.1. By using Equation 3.7.2, the characteristic function of $\bar{X}_n$ is

$$\begin{aligned}\Psi_{\bar{X}_n}(t) &= \left(\Psi_X\left(\frac{t}{n}\right)\right)^n \\ &= \left(w\Psi_1\left(\frac{t}{n}\right) + (1-w)\Psi_2\left(\frac{t}{n}\right)\right)^n\end{aligned}$$

By using Binomial expansion, the previous equation becomes

$$\Psi_{\bar{X}_n}(t) = \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\Psi_1^k\left(\frac{t}{n}\right)\Psi_2^{(n-k)}\left(\frac{t}{n}\right)$$

Using Equation 3.7.3, we obtain

$$\begin{aligned}\Psi_{\bar{X}_n}(t) &= \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\left(\exp\left(\frac{it\mu_1}{n} - \frac{t^2\sigma_1^2}{2n^2}\right)\right)^k\left(\exp\left(\frac{it\mu_2}{n} - \frac{t^2\sigma_2^2}{2n^2}\right)\right)^{(n-k)} \\ &= \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\exp\left(\frac{it\left(k\mu_1 + (n-k)\mu_2\right)}{n} - \frac{t^2\left(k\sigma_1^2 + (n-k)\sigma_2^2\right)}{2n^2}\right)\end{aligned}$$

Consequently, $\bar{X}_n$ is a mixture of (n+1) normal distributions. Moreover, its probability density function is

$$f_{\bar{X}_n}(x) = \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\frac{n\exp\left(\frac{-(nx-(k\mu_1+(n-k)\mu_2))^2}{2\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}\right)}{\sqrt{2\pi\left(k\sigma_1^2 + (n-k)\sigma_2^2\right)}} \tag{3.7.4}$$

and the characteristic function of Y_n is

$$\begin{aligned}\Psi_{Y_n}(t) &= e^{-itc}\Psi_{\bar{X}_n}(t) \\ &= \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\exp\left(it\left(\frac{k\mu_1 + (n-k)\mu_2}{n} - c\right) - \frac{t^2\left(k\sigma_1^2 + (n-k)\sigma_2^2\right)}{2n^2}\right)\end{aligned}$$

Hence, Y_n is a mixture of (n+1) normal distributions. In addition, its probability density function is

$$f_{Y_n}(y) = \sum_{k=0}^{n} \binom{n}{k} w^k (1-w)^{(n-k)} \frac{n \ \exp\left(\frac{-(ny-(k\mu_1+(n-k)\mu_2-nc))^2}{2\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}\right)}{\sqrt{2\pi\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}} \tag{3.7.5}$$

Using Equations 3.3.1 and 3.7.5, thus

$$\begin{aligned}
\text{MFE}_n &= \frac{w}{2}\int_{-\infty}^{\infty}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\frac{n}{\sqrt{2\pi\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}} \\
&\times \exp\left(\frac{-\left(ny-(k\mu_1+(n-k)\mu_2-nc)\right)^2}{2\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}\right)\left(1+\text{erf}\left(\frac{y-\mu_1}{\sqrt{2}\sigma_1}\right)\right)dy \\
&+ \frac{(1-w)}{2}\int_{-\infty}^{\infty}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\frac{n}{\sqrt{2\pi\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}} \\
&\times \exp\left(\frac{-\left(ny-(k\mu_1+(n-k)\mu_2-nc)\right)^2}{2\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}\right)\left(1+\text{erf}\left(\frac{y-\mu_2}{\sqrt{2}\sigma_2}\right)\right)dy \\
&= \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)} \\
&\quad\times\int_{-\infty}^{\infty}\frac{n}{\sqrt{2\pi\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}}\exp\left(\frac{-\left(ny-(k\mu_1+(n-k)\mu_2-nc)\right)^2}{2\left(k\sigma_1^2+(n-k)\sigma_2^2\right)}\right) \\
&\qquad\times\left\{w\ \text{erf}\left(\frac{y-\mu_1}{\sqrt{2}\sigma_1}\right)+(1-w)\text{erf}\left(\frac{y-\mu_2}{\sqrt{2}\sigma_2}\right)\right\}dy
\end{aligned}$$

Using the result from Appendix A.1,

$$\begin{aligned}
\text{MFE}_n &= \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\left\{w\ \text{erf}\left(\frac{(n-k)(\mu_2-\mu_1)-nc}{\sqrt{2\left((n^2+k)\sigma_1^2+(n-k)\sigma_2^2\right)}}\right)\right. \\
&\qquad\left.+(1-w)\text{erf}\left(\frac{k(\mu_1-\mu_2)-nc}{\sqrt{2\left(k\sigma_1^2+(n^2+n-k)\sigma_2^2\right)}}\right)\right\}
\end{aligned}$$

When $X \sim \text{SMN}(w, 0, 1, \psi)$, the MFE_n will be

$$\begin{aligned}
\text{MFE}_n &= \frac{1}{2}+\frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\left\{w\ \text{erf}\left(\frac{-nc}{\sqrt{2\left((n^2+k)+(n-k)(1+\Delta)\right)}}\right)\right. \\
&\qquad\left.+(1-w)\text{erf}\left(\frac{-nc}{\sqrt{2\left(k+(n^2+n-k)(1+\Delta)\right)}}\right)\right\}
\end{aligned} \tag{3.7.6}$$

Which is equivalent to

$$\begin{aligned}\text{MFE}_n &= \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\left(w\Phi\left(\frac{-nc}{\sqrt{n^2+n+(n-k)\Delta}}\right)\right.\\ &\left.+(1-w)\Phi\left(\frac{-nc}{\sqrt{n^2+n+(n^2+n-k)\Delta}}\right)\right) \qquad (3.7.7)\end{aligned}$$

The previous equation can be rewritten in terms of ψ, thus we obtain

$$\begin{aligned}\text{MFE}_n &= \sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)}\left(w\Phi\left(\frac{-nc}{\sqrt{n^2+k+(n-k)\psi}}\right)\right.\\ &\left.+(1-w)\Phi\left(\frac{-nc}{\sqrt{k+(n^2+n-k)\psi}}\right)\right) \qquad (3.7.8)\end{aligned}$$

In this section, MFE_n are expressed in terms of ψ or $\Delta = \psi - 1$ according to the case that we would like to prove.

The MFE_n in Equation 3.7.8 is plotted in Figure 3.1 for sample sizes $n = 1, \ldots, 50$ and different values of w in the case when $c = 1$ and $\psi = 43$. In some of these plots $\text{MFE}_1 \leq \text{MFE}_2$, which means that the increase in sample size from one observation to two observations does not guarantee the decreasing of the mean fraction exceeded.

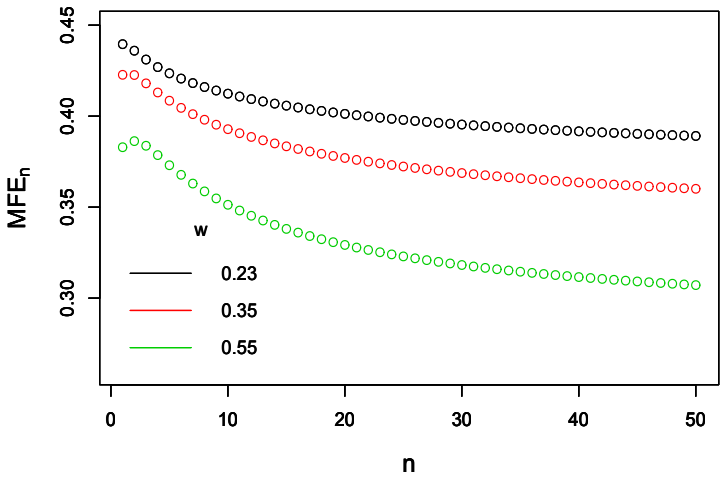


Figure 3.1: A plot of MFE_n versus n for the SMN$(w, 0, 1, 43)$ distribution and selected values of w.

Based on Figure 3.1, for $c = 1$ and $\psi = 43$, when $w = 0.23$ we found $\text{MFE}_1 > \text{MFE}_2$. Whereas, when $w = 0.35$, $\text{MFE}_1 = \text{MFE}_2$. In addition, when $w = 0.55$, $\text{MFE}_1 < \text{MFE}_2$. Hence, we became interested in exploring the case when MFE_1 and MFE_2 are equal, since we will know when $\text{MFE}_1 > \text{MFE}_2$, i.e. when MFE_n decreases as n increases from one to two. Using Equation 3.7.8, MFE_1 and MFE_2 are respectively given by

$$\text{MFE}_1 = (1-w)^2\Phi\left(\frac{-c}{\sqrt{2\psi}}\right) + 2w(1-w)\Phi\left(\frac{-c}{\sqrt{\psi+1}}\right) + w^2\Phi\left(\frac{-c}{\sqrt{2}}\right) \quad (3.7.9)$$

and

$$\begin{aligned}\text{MFE}_2 = {} & (1-w)^2 w\Phi\left(\frac{-2c}{\sqrt{2\psi+4}}\right) + (1-w)^3\Phi\left(\frac{-2c}{\sqrt{6\psi}}\right) \\ & +2w^2(1-w)\Phi\left(\frac{-2c}{\sqrt{\psi+5}}\right) + 2w(1-w)^2\Phi\left(\frac{-2c}{\sqrt{5\psi+1}}\right) \\ & +w^3\Phi\left(\frac{-2c}{\sqrt{6}}\right) + w^2(1-w)\Phi\left(\frac{-2c}{\sqrt{4\psi+2}}\right) \end{aligned} \quad (3.7.10)$$

We shall use the term ΔMFE_{12} to refer to the difference between MFE_1 and MFE_2. Thus

$$\Delta\text{MFE}_{12} = \text{MFE}_1 - \text{MFE}_2 \quad (3.7.11)$$

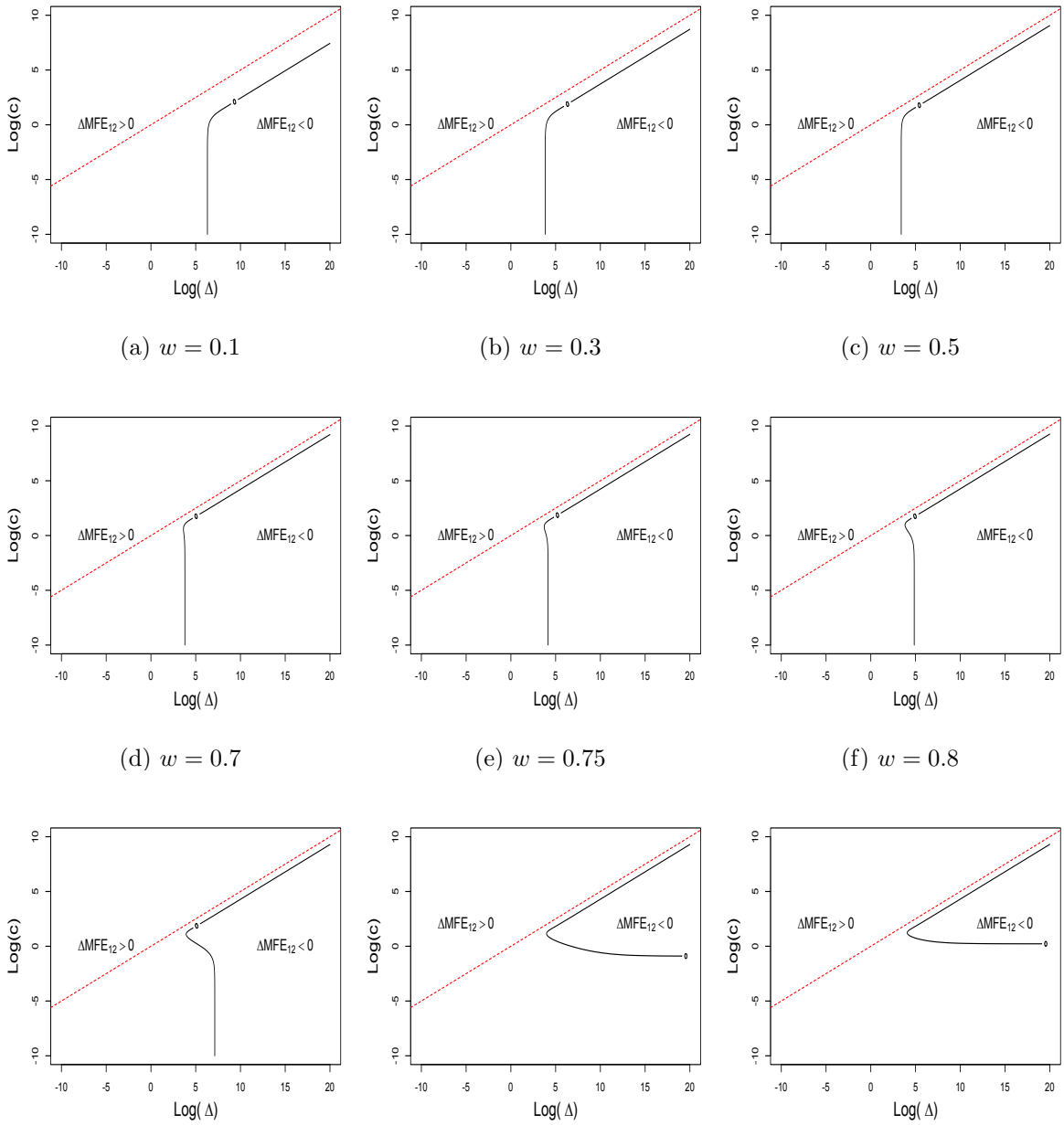
By using Equations 3.7.9 and 3.7.10,

$$\begin{aligned}\Delta\text{MFE}_{12} = {} & w^3\left[\Phi\left(\frac{-2c}{\sqrt{4\psi+2}}\right) - \Phi\left(\frac{-2c}{\sqrt{6}}\right) + 2\Phi\left(\frac{-2c}{\sqrt{\psi+5}}\right) - 2\Phi\left(\frac{-2c}{\sqrt{5\psi+1}}\right)\right. \\ & \left. -\Phi\left(\frac{-2c}{\sqrt{2\psi+4}}\right) + \Phi\left(\frac{-2c}{\sqrt{6\psi}}\right)\right] \\ & + w^2\left[\Phi\left(\frac{-c}{\sqrt{2}}\right) - \Phi\left(\frac{-2c}{\sqrt{4\psi+2}}\right) - 2\Phi\left(\frac{-2c}{\sqrt{\psi+5}}\right) - 2\Phi\left(\frac{-c}{\sqrt{\psi+1}}\right)\right. \\ & \left. +4\Phi\left(\frac{-2c}{\sqrt{5\psi+1}}\right) + \Phi\left(\frac{-c}{\sqrt{2\psi}}\right) + 2\Phi\left(\frac{-2c}{\sqrt{2\psi+4}}\right) - 3\Phi\left(\frac{-2c}{\sqrt{6\psi}}\right)\right] \\ & + w\left[2\Phi\left(\frac{-c}{\sqrt{\psi+1}}\right) - 2\Phi\left(\frac{-2c}{\sqrt{5\psi+1}}\right) - 2\Phi\left(\frac{-c}{\sqrt{2\psi}}\right) - \Phi\left(\frac{-2c}{\sqrt{2\psi+4}}\right)\right. \\ & \left. +3\Phi\left(\frac{-2c}{\sqrt{6\psi}}\right)\right] + \left[\Phi\left(\frac{-c}{\sqrt{2\psi}}\right) - \Phi\left(\frac{-2c}{\sqrt{6\psi}}\right)\right] \end{aligned}$$

With respect to w, ΔMFE_{12} is a cubic function with coefficients that depend on c and ψ. Furthermore, ΔMFE_{12} has a value of zero if and only if the real root(s) of

this cubic function is (are) restricted to be in the interval $(0, 1)$.

Figure 3.2 displays the zero contour of ΔMFE_{12} for different values of w. The vertical axis represents a wide range of c, and the horizontal axis represents $\Delta = \psi - 1$. Both axes are taken on the log-scale.



(a) $w = 0.1$ (b) $w = 0.3$ (c) $w = 0.5$

(d) $w = 0.7$ (e) $w = 0.75$ (f) $w = 0.8$

(g) $w = 0.85$ (h) $w = 0.87$ (i) $w = 0.9$

Figure 3.2: Plots of the zero contour of ΔMFE_{12} with respect to $\log(\Delta)$ (x-axis) and $\log(c)$ (y-axis) for the $\text{SMN}(w, 0, 1, \Delta)$ distribution. Each of the nine panels is related to a specific value of w. The red dashed lines equation is $\log(c) = 0.5\log(\Delta)$.

From Figure 3.2, the zero contour of ΔMFE_{12} behaves differently in three parts. The first part which is formed of the vertical line that appears when c is small. The second part which is about the horizontal line that is formed when ψ has a large value. In the third part, when both c and ψ have large values, the zero contour of ΔMFE_{12} appears as a line that is parallel to the line $\log(c) = 0.5\log(\Delta)$. The aim of the rest of this section and Section 3.8, as we will see later, is to bound the area where the dominance properties for the arithmetic mean do not hold. Therefore, we are interested in exploring the limiting behaviour as parameters tend to extreme values. The purpose of those limiting calculations is to be sure that the behaviour seen in the figures is correct for more extreme values of parameters where the numerical methods used to produce the figures might become unstable. Consequently, we shall consider three cases: c is small, ψ approaches ∞, and both c and ψ approach ∞. In addition, we will add lines, the equations for which are obtained by assuming these three cases, to copies of those figures in order to verify that the figures were showing the correct limiting behaviours. The summary of findings for these three cases is presented in Table 3.2 followed by detailed proofs.

Case A: Small c
$\Delta\text{MFE}_{12} \approx \frac{c}{\sqrt{\pi}}\left[\sum_{k=0}^{2}\binom{2}{k}w^{(k+1)}(1-w)^{(2-k)}\sqrt{\frac{2}{4+k+(2-k)\psi}}\right.$
$+\sum_{k=0}^{2}\binom{2}{k}w^{k}(1-w)^{(3-k)}\sqrt{\frac{2}{k+(6-k)\psi}}$
$-\sum_{k=0}^{1}\binom{1}{k}\frac{w^{(k+1)}(1-w)^{(1-k)}}{\sqrt{2\{1+k+(1-k)\psi\}}}$
$\left.-\sum_{k=0}^{1}\binom{1}{k}\frac{w^{k}(1-w)^{(2-k)}}{\sqrt{2\{k+(2-k)\psi\}}}\right]$

Table 3.2: Summary of the behaviour limiting of the zero contour of ΔMFE_{12} for the $\text{SMN}(w, 0, 1, \psi)$ distribution when c is small, ψ approaches ∞, and c and ψ both approach ∞. Table continues on the next page.

Case B: ψ approaches ∞
$\Delta\text{MFE}_{12} \to \Big(2w(1-w) - w^2(1-w)\Big)\Phi\Big(\frac{-1}{\sqrt{\phi}}\Big) + (1-w)^2\Phi\Big(\frac{-1}{\sqrt{2\phi}}\Big)$ $- w(1-w)^2\Phi\Big(-\sqrt{\frac{2}{\phi}}\Big) - (1-w)^3\Phi\Big(-\sqrt{\frac{2}{3\phi}}\Big)$ $- 2w^2(1-w)\Phi\Big(\frac{-2}{\sqrt{\phi}}\Big) - 2w(1-w)^2\Phi\Big(\frac{-2}{\sqrt{5\phi}}\Big)$
Case C: c and ψ both approach ∞
$\Delta\text{MFE}_{12} \to \Big(2w(1-w) - w^2(1-w)\Big)\Phi\Big(\frac{-1}{\sqrt{\phi}}\Big) + (1-w)^2\Phi\Big(\frac{-1}{\sqrt{2\phi}}\Big)$ $- w(1-w)^2\Phi\Big(-\sqrt{\frac{2}{\phi}}\Big) - (1-w)^3\Phi\Big(-\sqrt{\frac{2}{3\phi}}\Big)$ $- 2w^2(1-w)\Phi\Big(\frac{-2}{\sqrt{\phi}}\Big) - 2w(1-w)^2\Phi\Big(\frac{-2}{\sqrt{5\phi}}\Big)$

Table 3.2: Continued from previous page.

The detailed proof of the summary results in Table 3.2 is as follows:

Case A: Small c

This case corresponds to the vertical lines in Figure 3.2. In the following, we will explore the equation of the vertical line given a specific value of w.

Equation 3.7.6 can be rewritten in terms of ψ, thus we obtain

$$\begin{aligned}\text{MFE}_n = \frac{1}{2} \quad &+ \quad \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^{(k+1)}(1-w)^{(n-k)} \\ &\qquad \times\text{erf}\left(\frac{-nc}{\sqrt{2\{n^2+k+(n-k)\psi\}}}\right) \\ &+ \quad \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^{k}(1-w)^{(n-k+1)} \\ &\qquad \times\text{erf}\left(\frac{-nc}{\sqrt{2\{k+(n^2+n-k)\psi\}}}\right) \qquad (3.7.12)\end{aligned}$$

For small c, The first-order Maclaurin series expansion of MFE_n in Equation 3.7.12

is

$$\begin{aligned}\text{MFE}_n &\approx \frac{1}{2} - \frac{c}{\sqrt{\pi}}\left[\sum_{k=0}^{n}\binom{n}{k}\frac{nw^{(k+1)}(1-w)^{(n-k)}}{\sqrt{2\{n^2+k+(n-k)\psi\}}}\right.\\ &\qquad\left.+\sum_{k=0}^{n}\binom{n}{k}\frac{nw^{k}(1-w)^{(n-k+1)}}{\sqrt{2\{k+(n^2+n-k)\psi\}}}\right]\end{aligned}$$

Consequently,

$$\begin{aligned}\Delta\text{MFE}_{12} &\approx \frac{c}{\sqrt{\pi}}\left[\sum_{k=0}^{2}\binom{2}{k}w^{(k+1)}(1-w)^{(2-k)}\sqrt{\frac{2}{4+k+(2-k)\psi}}\right.\\ &\qquad+\sum_{k=0}^{2}\binom{2}{k}w^{k}(1-w)^{(3-k)}\sqrt{\frac{2}{k+(6-k)\psi}}\\ &\qquad-\sum_{k=0}^{1}\binom{1}{k}\frac{w^{(k+1)}(1-w)^{(1-k)}}{\sqrt{2\{1+k+(1-k)\psi\}}}\\ &\qquad\left.-\sum_{k=0}^{1}\binom{1}{k}\frac{w^{k}(1-w)^{(2-k)}}{\sqrt{2\{k+(2-k)\psi\}}}\right]\end{aligned}$$

The previous equation is the equation of the vertical line at any value of ψ when a specific value of w is given.

Case B: ψ approaches ∞

When $\psi \to \infty$, the zero contour of ΔMFE_{12} forms a horizontal line. The equations of the horizontal lines in Figure 3.2 are obtained as follows:

By assuming $k = n$ in the first summation of Equation 3.7.12, the term that does not involve ψ is extracted. Hence,

$$\begin{aligned}\text{MFE}_n = \frac{1}{2} &+ \frac{1}{2}\sum_{k=0}^{n-1}\binom{n}{k}w^{(k+1)}(1-w)^{(n-k)}\text{erf}\left(\frac{-nc}{\sqrt{2\{n^2+k+(n-k)\psi\}}}\right)\\ &+ \frac{1}{2}w^{(n+1)}\text{erf}\left(\frac{-nc}{\sqrt{2(n^2+n)}}\right)\\ &+ \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^{k}(1-w)^{(n-k+1)}\text{erf}\left(\frac{-nc}{\sqrt{2\{k+(n^2+n-k)\psi\}}}\right)\end{aligned}$$

Set $\psi \to \infty$ in the previous equation, thus

$$\begin{aligned}\text{MFE}_n &\to \frac{1}{2} + \frac{w^{(n+1)}}{2}\text{erf}\left(\frac{-nc}{\sqrt{2(n^2+n)}}\right)\\ &= \frac{1}{2} + w^{(n+1)}\left\{\Phi\left(\frac{-nc}{\sqrt{n^2+n}}\right) - \frac{1}{2}\right\}\end{aligned}$$

Consequently,

$$\text{MFE}_1 \to \frac{1}{2} + w^2 \left\{ \Phi\left(\frac{-c}{\sqrt{2}}\right) - \frac{1}{2} \right\}$$

and

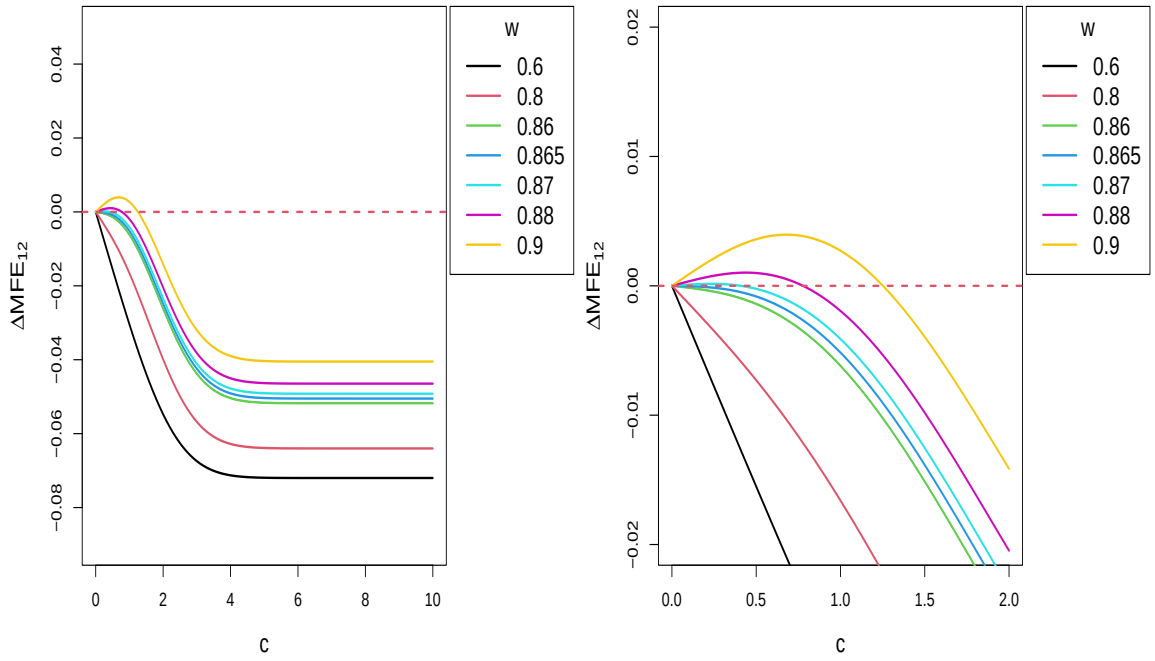
$$\text{MFE}_2 \to \frac{1}{2} + w^3 \left\{ \Phi\left(\frac{-2c}{\sqrt{6}}\right) - \frac{1}{2} \right\}$$

Therefore,

$$\Delta\text{MFE}_{12} \to w^2 \left\{ \Phi\left(\frac{-c}{\sqrt{2}}\right) - w\Phi\left(\frac{-2c}{\sqrt{6}}\right) - \frac{(1-w)}{2} \right\} \tag{3.7.13}$$

Equation 3.7.13 is used to obtain c that makes $\Delta\text{MFE}_{12} \to 0$ for a specific value of w, this represents the equation of the horizontal line when $\psi \to \infty$.

Figure 3.3 presents ΔMFE_{12} in Equation 3.7.13 as a function of c for selected values of w. The left panel demonstrates this argument for a wide range of c. It is clear that ΔMFE_{12} could be zero for c between 0 and 2. However, it is not clear for which value of w. To clarify this argument visually, we set the ranges of c and ΔMFE_{12} to be respectively $[0, 2]$ and $[-0.02, 0.02]$ in the right panel.



(a) $c \in (0, 10)$ and $\Delta\text{MFE}_{12} \in (-0.09, 0.05)$ (b) $c \in (0, 2)$ and $\Delta\text{MFE}_{12} \in (-0.02, 0.02)$

Figure 3.3: ΔMFE_{12} versus c for the $\text{SMN}(w, 0, 1, \psi)$ distribution and selected values of w when $\psi \to \infty$ (Equation 3.7.13).

Figure 3.3 suggests that ΔMFE_{12} in Equation 3.7.13 could tend to zero when w becomes larger.

Case C: c and ψ both approach ∞

As illustrated in Figure 3.2, when c and ψ approach ∞, the zero contour of ΔMFE_{12} is parallel to $\log(c) = 0.5\log(\Delta)$. The equation of such a line is

$$\log(c) = 0.5(\log(\Delta) - \log(\phi)) \tag{3.7.14}$$

Therefore, we will obtain $\phi = \Delta/c^2$ as follows:

Equation 3.7.7 can be written as

$$\begin{aligned}\text{MFE}_n &= \sum_{k=0}^{n} \binom{n}{k} w^k (1-w)^{(n-k)} \\ &\quad \times \left(w\Phi\left(\frac{-n}{\sqrt{(n^2+n)/c^2 + (n-k)\Delta/c^2}} \right) \right. \\ &\quad \left. + (1-w)\Phi\left(\frac{-n}{\sqrt{(n^2+n)/c^2 + (n^2+n-k)\Delta/c^2}} \right) \right)\end{aligned}$$

Let c and Δ approach infinity, and notice that $\phi = \Delta/c^2$ becomes constant. Therefore,

$$\begin{aligned}\text{MFE}_n &\to \sum_{k=0}^{n} \binom{n}{k} w^k (1-w)^{(n-k)} \left(w\Phi\left(\frac{-n}{\sqrt{(n-k)\phi}} \right) \right. \\ &\quad \left. + (1-w)\Phi\left(\frac{-n}{\sqrt{(n^2+n-k)\phi}} \right) \right)\end{aligned}$$

Hence,

$$\begin{aligned}\Delta\text{MFE}_{12} &\to \left(2w(1-w) - w^2(1-w)\right)\Phi\left(\frac{-1}{\sqrt{\phi}}\right) + (1-w)^2\Phi\left(\frac{-1}{\sqrt{2\phi}}\right) \\ &- w(1-w)^2\Phi\left(-\sqrt{\frac{2}{\phi}}\right) - (1-w)^3\Phi\left(-\sqrt{\frac{2}{3\phi}}\right) \\ &- 2w^2(1-w)\Phi\left(\frac{-2}{\sqrt{\phi}}\right) - 2w(1-w)^2\Phi\left(\frac{-2}{\sqrt{5\phi}}\right)\end{aligned} \tag{3.7.15}$$

Equation 3.7.15 is used to obtain ϕ numerically for a specific value of w.

The findings in Case A, Case B, and Case C for selected values of w are presented in Table 3.3.

w	Small c	ψ approaches ∞	c and ψ both approach ∞
0.1	$\log(\Delta) = 6.31$	–	$\log(\Delta/c^2) = 5.13$
0.3	$\log(\Delta) = 3.82$	–	$\log(\Delta/c^2) = 2.56$
0.5	$\log(\Delta) = 3.36$	–	$\log(\Delta/c^2) = 1.88$
0.7	$\log(\Delta) = 3.77$	–	$\log(\Delta/c^2) = 1.57$
0.75	$\log(\Delta) = 4.14$	–	$\log(\Delta/c^2) = 1.52$
0.8	$\log(\Delta) = 4.85$	–	$\log(\Delta/c^2) = 1.47$
0.85	$\log(\Delta) = 7.13$	–	$\log(\Delta/c^2) = 1.43$
0.87	–	$\log(c) = -0.89$	$\log(\Delta/c^2) = 1.41$
0.9	–	$\log(c) = 0.22$	$\log(\Delta/c^2) = 1.39$

Table 3.3: The zero contour equations of ΔMFE_{12} for the $\text{SMN}(w, 0, 1, \Delta)$ distribution, and selected values of w, when c is small, ψ approaches ∞, and c and ψ both approach ∞. The numbers in the second, third, and fourth column are rounded to two decimal places.

In Table 3.3, ψ is calculated for different values of w. Clearly ψ decreases to a certain value, and then it increases. The minimum value of ψ (ψ_{min}) is calculated numerically to be 29.8187 at $w = 0.515$. Furthermore, ϕ decreases in w. Hence, the minimum value of ϕ (ϕ_{min}) is numerically found to be 3.777 at $w = 0.9999$.
In addition, ϕ_{min} and ψ_{min} determine respectively the boundary lines

$$\log(c) = 0.5(\log(\Delta) - 1.329)$$

and

$$\log(\Delta) = 3.361024$$

These lines are used to determine the area where $\Delta\text{MFE}_{12} \lessapprox 0$, see Figure 3.4 for illustration.

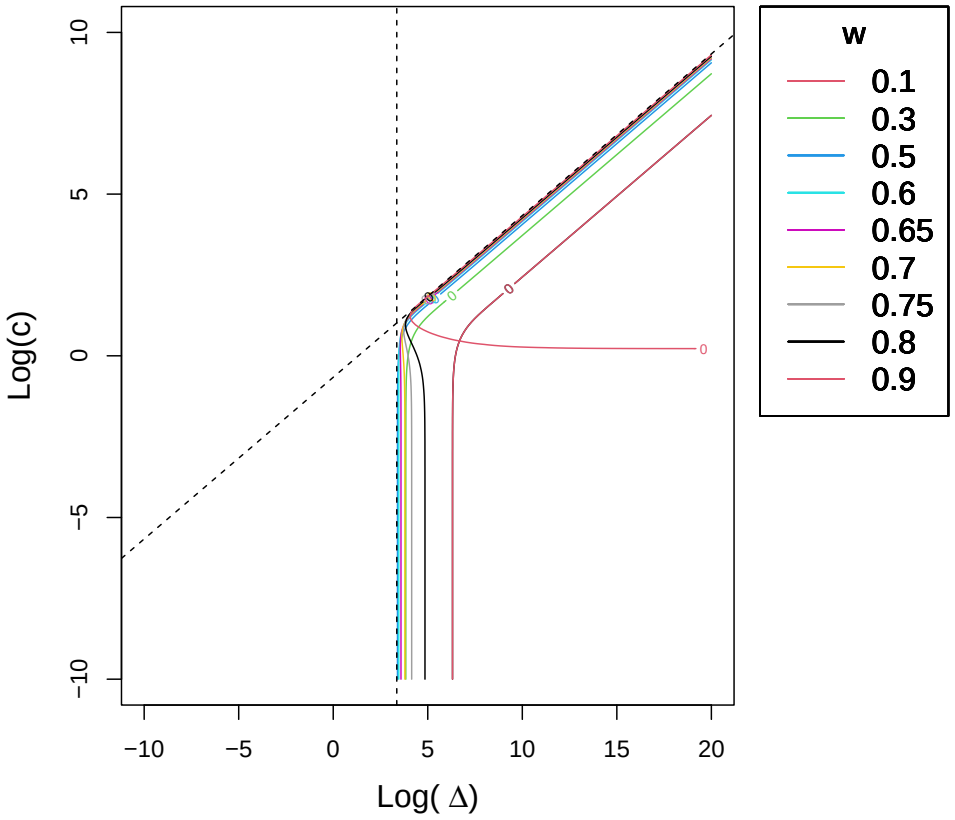
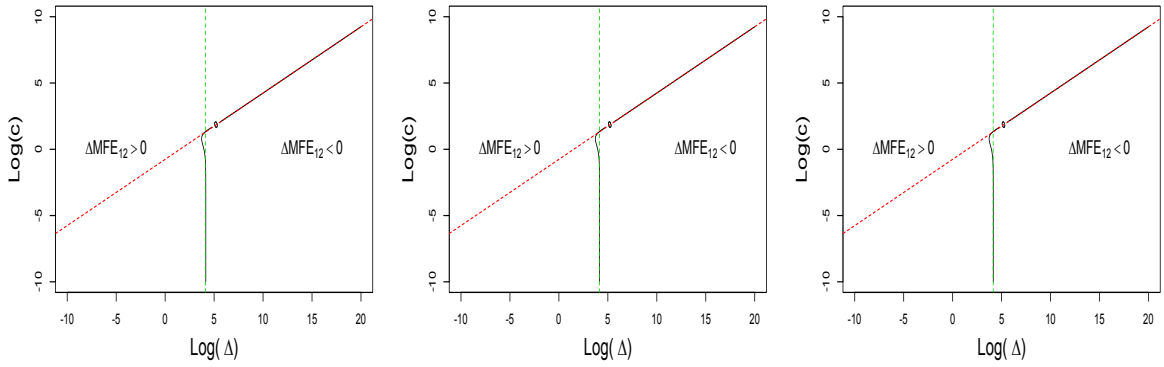
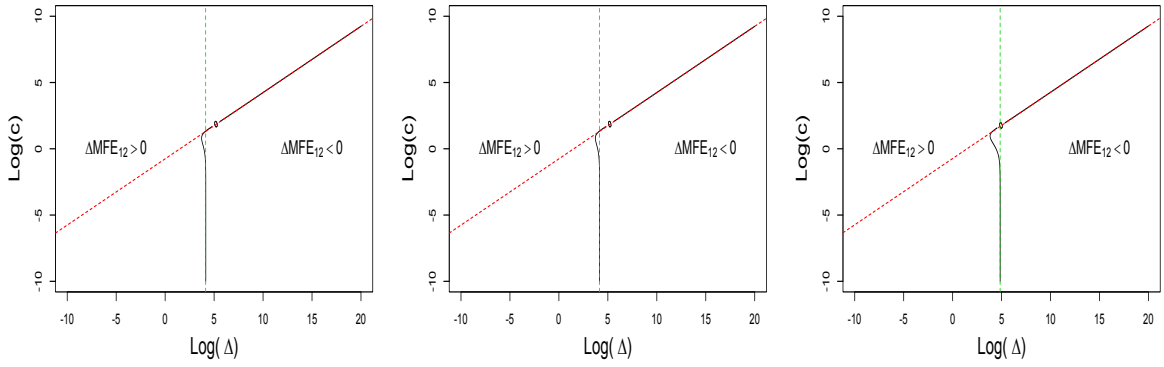


Figure 3.4: A plot of the zero contours of ΔMFE_{12} with respect to $\log\Delta$ (x-axis) and $\log(c)$ (y-axis) for the $\text{SMN}(w, 0, 1, \Delta)$ distribution, and selected values of w. The black dashed lines equations are $\log(\Delta) = 3.361024$ and $\log(c) = 0.5(\log(\Delta) - 1.329)$

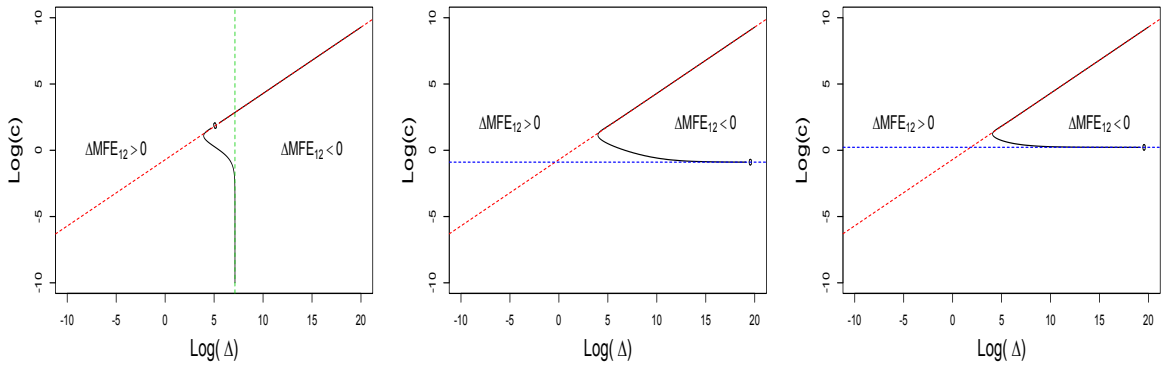
In order to make comparative plots between the approximations of ΔMFE_{12} in the previous cases and the real contours, the zero contour of ΔMFE_{12} for different values of w is illustrated with the lines, the equations for which are presented in Table 3.3, in Figure 3.5.



(a) $w = 0.1$ (b) $w = 0.3$ (c) $w = 0.5$



(d) $w = 0.7$ (e) $w = 0.75$ (f) $w = 0.8$



(g) $w = 0.85$ (h) $w = 0.87$ (i) $w = 0.9$

Figure 3.5: Plots of the zero contour of ΔMFE_{12} with respect to $\log(\Delta)$ (x-axis) and $\log(c)$ (y-axis) for the $\text{SMN}(w, 0, 1, \Delta)$ distribution. Each of the nine panels related to a specific value of w. The equations of dashed lines are presented in Table 3.3.

□

3.7.2 $\mathbf{PFE}_n(\alpha)$

In this subsection, we first obtain $\mathrm{PFE}_n(\alpha)$ when $X \sim \mathrm{MN}(w, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$. Second, we obtain $\mathrm{PFE}_n(\alpha)$ when $X \sim \mathrm{SMN}(w, 0, 1, \psi)$ as a special case.

Define $b = F_X^{-1}(\alpha) + c$, the condition $c > \mu - F_X^{-1}(\alpha)$ introduced earlier in Equation 3.3.4 is equivalent in this part to $b > 0$. Hence Equation 3.3.3 becomes

$$\mathrm{PFE}_n(\alpha) = P\big(\bar{X}_n > b\big) \tag{3.7.16}$$

Using Equations 3.7.4 and 3.7.16, the $\mathrm{PFE}_n(\alpha)$ is

$$\mathrm{PFE}_n(\alpha) = \sum_{k=0}^{n} \binom{n}{k} w^k (1-w)^{(n-k)} \left(1 - \Phi\left(\frac{nb - (k\mu_1 + (n-k)\mu_2)}{(k\sigma_1^2 + (n-k)\sigma_2^2)/n}\right)\right) \tag{3.7.17}$$

Hence, for $\mathrm{SMN}(w, 0, 1, \psi)$ distribution and $n = 1, 2$; $\mathrm{PFE}_1(\alpha)$ and $\mathrm{PFE}_2(\alpha)$ respectively are

$$\mathrm{PFE}_1(\alpha) = 1 - (1-w)\Phi\left(\frac{b-\mu_2}{\sigma_2}\right) - w\Phi\left(\frac{b-\mu_1}{\sigma_1}\right) \tag{3.7.18}$$

and

$$\mathrm{PFE}_2(\alpha) = 1 - \left[(1-w)^2\Phi\left(\frac{b-\mu_2}{\sqrt{\sigma_2^2/2}}\right) + 2w(1-w)\Phi\left(\frac{b-(\mu_1+\mu_2)/2}{\sqrt{(\sigma_1^2+\sigma_2^2)/4}}\right) + w^2\Phi\left(\frac{b-\mu_1}{\sqrt{\sigma_1^2/2}}\right)\right] \tag{3.7.19}$$

Consequently, when $X \sim \mathrm{SMN}(w, 0, 1, \psi)$, Equations 3.7.18 and 3.7.19 respectively become

$$\mathrm{PFE}_1(\alpha) = 1 - (1-w)\Phi\left(\frac{b}{\sqrt{\psi}}\right) - w\Phi(b)$$

and

$$\mathrm{PFE}_2(\alpha) = 1 - (1-w)^2\Phi\left(\frac{b}{\sqrt{\psi/2}}\right) - 2w(1-w)\Phi\left(\frac{b}{\sqrt{(1+\psi)/2}}\right) - w^2\Phi\left(\frac{b}{\sqrt{1/2}}\right)$$

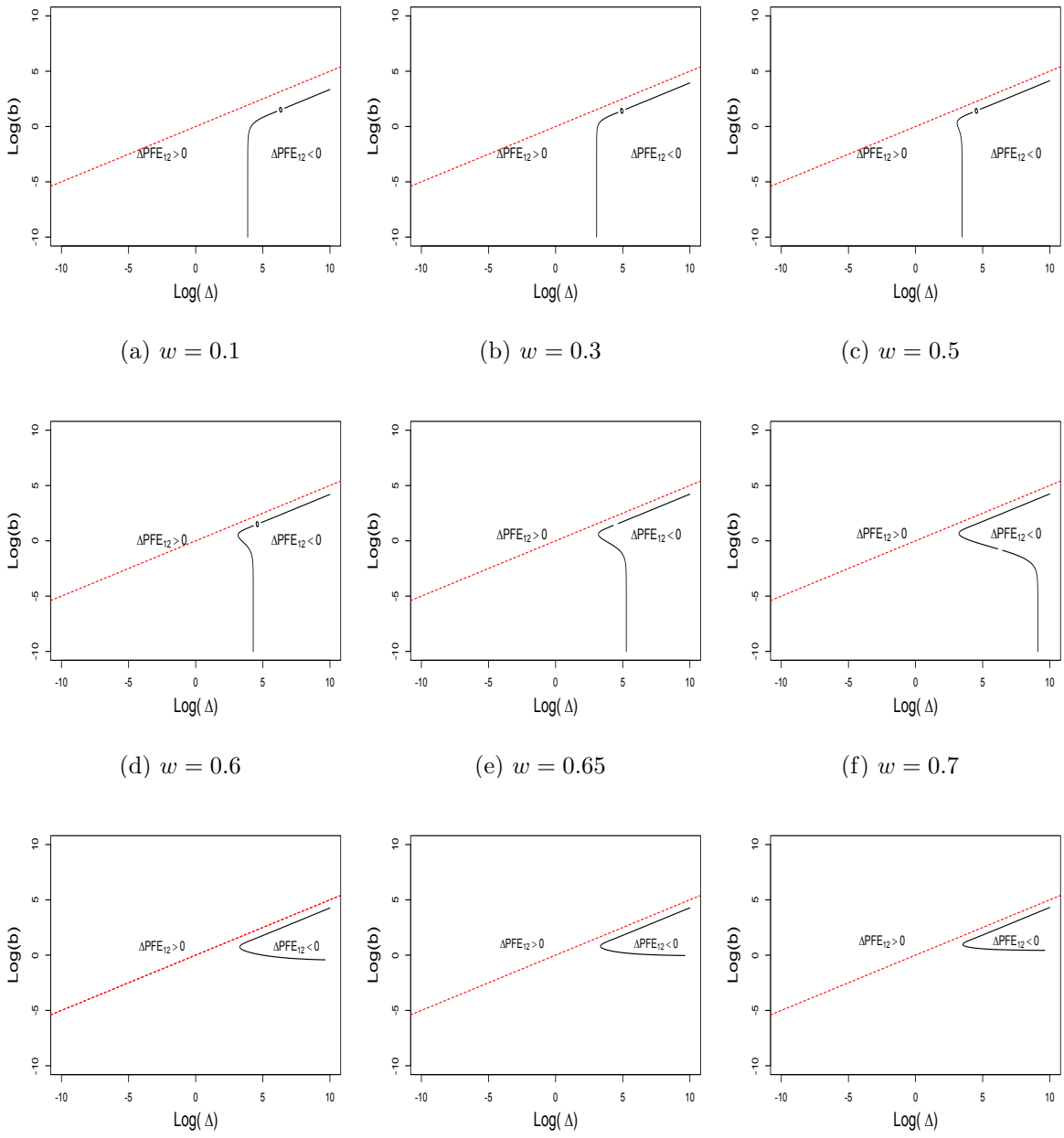
The difference between $\mathrm{PFE}_1(\alpha)$ and $\mathrm{PFE}_2(\alpha)$ is defined as

$$\Delta\mathrm{PFE}_{12}(\alpha) = \mathrm{PFE}_1(\alpha) - \mathrm{PFE}_2(\alpha) \tag{3.7.20}$$

Therefore,

$$\begin{aligned}\Delta\mathrm{PFE}_{12}(\alpha) &= (1-w)^2\Phi\left(\frac{b}{\sqrt{\psi/2}}\right) + 2w(1-w)\Phi\left(\frac{b}{\sqrt{(1+\psi)/2}}\right) + w^2\Phi\left(\frac{b}{\sqrt{1/2}}\right) \\ &- (1-w)\Phi\left(\frac{b}{\sqrt{\psi}}\right) - w\Phi(b)\end{aligned} \tag{3.7.21}$$

In Figure 3.6, we display the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ for different values of w. The vertical axis represents a wide range of b, while the horizontal axis represents $\Delta = \psi - 1$. Both axes are taken on the log-scale.



(a) $w = 0.1$ (b) $w = 0.3$ (c) $w = 0.5$

(d) $w = 0.6$ (e) $w = 0.65$ (f) $w = 0.7$

(g) $w = 0.75$ (h) $w = 0.8$ (i) $w = 0.9$

Figure 3.6: Plots of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log(\Delta)$ (x-axis) and $\log(b)$ (y-axis) for the $\text{SMN}(w, 0, 1, \Delta)$ distribution. Each of the nine panels is related to a specific value of w. The red dashed lines equation is $\log(b) = 0.5\log(\Delta)$.

Equation 3.7.21 can be rewritten as

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= w^2\left\{\Phi\left(\frac{b}{\sqrt{1/2}}\right) - 2\Phi\left(\frac{2b}{\sqrt{(1+\psi)}}\right) + \Phi\left(\frac{b}{\sqrt{\psi/2}}\right)\right\} \\ &+ w\left\{2\Phi\left(\frac{2b}{\sqrt{(1+\psi)}}\right) - 2\Phi\left(\frac{b}{\sqrt{\psi/2}}\right) + \Phi\left(\frac{b}{\sqrt{\psi}}\right)\right. \\ &\left. -\Phi\left(b\right)\right\} - \Phi\left(\frac{b}{\sqrt{\psi}}\right) + \Phi\left(\frac{b}{\sqrt{\psi/2}}\right)\end{aligned}$$

With respect to w, $\Delta\text{PFE}_{12}(\alpha)$ is a quadratic function with coefficients depending on b and ψ. It equals zero if and only if the real root(s) of this quadratic function is (are) restricted to be in the interval $(0, 1)$. As illustrated in Figure 3.6, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ behaves differently in three different ways depending on the values of b, ψ, and such root(s).

In this section, we will analyse the behaviour of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ as parameters tend to extreme values. Consequently, we will explore three cases: b is small, ψ approaches ∞, and b and ψ both approach ∞. The summary of findings will be presented in Table 3.4 followed by detailed proof.

Case A: Small b
$\Delta\text{PFE}_{12}(\alpha) \approx \frac{b}{\sqrt{\pi}}\Big(\frac{(1-w)^2}{\sqrt{\psi}} + 2w(1-w)\sqrt{\frac{2}{1+\psi}} + w^2 - \frac{(1-w)}{\sqrt{2\psi}} - \frac{w}{\sqrt{2}}\Big)$
Case B: ψ approaches ∞
$\Delta\text{PFE}_{12}(\alpha) \to w^2\Phi(\sqrt{2}b) - w\Phi(b) + \frac{w(1-w)}{2}$
Case C: b and ψ both approach ∞
$\Delta\text{PFE}_{12}(\alpha) \to (1-w)^2\Phi\big(\sqrt{\frac{2}{\phi}}\big) + 2w(1-w)\Phi\big(\frac{2}{\sqrt{\phi}}\big) - w(1-w) - (1-w)\Phi\big(\frac{1}{\sqrt{\phi}}\big)$

Table 3.4: Summary of the behaviour limiting of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ for the $\text{SMN}(w, 0, 1, \psi)$ distribution when b is small, ψ approaches ∞, and both b and ψ approach ∞.

The detailed proof of the summary results in Table 3.4 is as follows:

Case A: Small b

In this case, we will explore the equations of the vertical lines in Figure 3.6 that are shown when b is small as follows:

Equation 3.7.21 is expressed in terms of erf (.), thus

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= \frac{1}{2}\Bigg[(1-w)^2\text{erf}\left(\frac{b}{\sqrt{\psi}}\right)+2w(1-w)\text{erf}\left(\frac{b}{\sqrt{(1+\psi)/2}}\right)+w^2\text{erf}\,(b)\\ &\qquad -(1-w)\text{erf}\left(\frac{b}{\sqrt{2\psi}}\right)-w\ \text{erf}\left(\frac{b}{\sqrt{2}}\right)\Bigg]\end{aligned}$$

For small b,

$$\Delta\text{PFE}_{12}(\alpha)\approx\frac{b}{\sqrt{\pi}}\left(\frac{(1-w)^2}{\sqrt{\psi}}+2w(1-w)\sqrt{\frac{2}{1+\psi}}+w^2-\frac{(1-w)}{\sqrt{2\psi}}-\frac{w}{\sqrt{2}}\right)$$

Based on a given value of w, the previous equation is used to calculate ψ that makes $\Delta\text{PFE}_{12}(\alpha)\approx 0$.

Case B: ψ approaches ∞

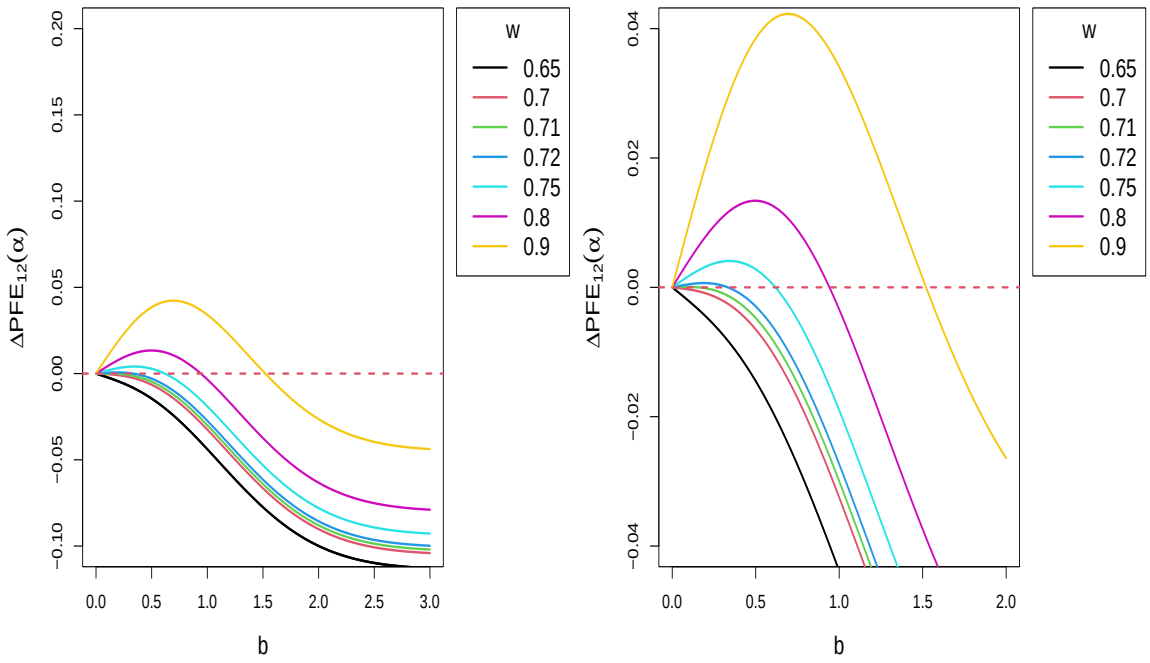
When $\psi\to\infty$, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ appears as a horizontal line. The equations of the horizontal lines in Figure 3.6 are obtained as follows:

Let $\psi\to\infty$ in Equation 3.7.21. Hence,

$$\Delta\text{PFE}_{12}(\alpha)\to w^2\Phi\left(\sqrt{2}b\right)-w\Phi(b)+\frac{w(1-w)}{2} \tag{3.7.22}$$

In order to get the equation of the horizontal line for a specific value of w, Equation 3.7.22 is used to obtain b that makes $\Delta\text{PFE}_{12}(\alpha)\to 0$

Figure 3.7 presents $\Delta\text{PFE}_{12}(\alpha)$ in Equation 3.7.22 as a function of b for selected values of w. The left panel demonstrates this argument for a specific range of b. As shown in the figure, $\Delta\text{PFE}_{12}(\alpha)$ could be zero for $b\in(0,2)$ and $\Delta\text{PFE}_{12}(\alpha)\in(-0.1,0.2)$. However, it is not clear for which value of w. To clarify this argument visually, we set the ranges of b and $\Delta\text{PFE}_{12}(\alpha)$ to be respectively $[0,2]$ and $[-0.04,0.04]$ in the right panel.



(a) $b \in (0, 3)$ and $\Delta\text{PFE}_{12}(\alpha) \in (-0.1, 0.2)$ (b) $b \in (0, 2)$ and $\Delta\text{PFE}_{12}(\alpha) \in (-0.04, 0.04)$

Figure 3.7: $\Delta\text{PFE}_{12}(\alpha)$ versus b for the SMN$(w, 0, 1, \psi)$ distribution and selected values of w when $\psi \to \infty$ (Equation 3.7.22).

Figure 3.7 suggests that $\Delta\text{PFE}_{12}(\alpha)$ in Equation 3.7.22 could tend to zero when w becomes larger.

Case C: b and ψ both approach ∞

As illustrated in Figure 3.6, when b and Δ approach infinity, the zero contour of ΔPFE12 forms the line that is parallel to $\log(b) = 0.5\log(\Delta)$. The equation of such a line is

$$\log(b) = 0.5(\log(\Delta) - \log(\phi)) \tag{3.7.23}$$

where $\phi = \Delta/b^2$ is obtained as follows:

Set $\psi = \Delta + 1$ in Equation 3.7.21,

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= (1-w)^2\Phi\left(\frac{\sqrt{2}b}{\sqrt{\Delta+1}}\right) + 2w(1-w)\Phi\left(\frac{2b}{\sqrt{(\Delta+2)}}\right) \\ &+ w^2\Phi\left(\frac{b}{\sqrt{1/2}}\right) - (1-w)\Phi\left(\frac{b}{\sqrt{\Delta+1}}\right) - w\Phi(b)\end{aligned}$$

The previous equation can be rewritten as

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= (1-w)^2\Phi\left(\frac{\sqrt{2}}{\sqrt{\Delta/b^2+1/b^2}}\right)+2w(1-w)\Phi\left(\frac{2}{\sqrt{\Delta/b^2+2/b^2}}\right)\\ &+ w^2\Phi\left(\frac{b}{\sqrt{1/2}}\right)-(1-w)\Phi\left(\frac{1}{\sqrt{\Delta/b^2+1/b^2}}\right)-w\Phi\left(b\right)\end{aligned}$$

Let b and Δ approach infinity, and notice that Δ/b^2 becomes constant. Define $\phi=\Delta/b^2$, therefore

$$\Delta\text{PFE}_{12}(\alpha)\to(1-w)^2\Phi\left(\sqrt{\frac{2}{\phi}}\right)+2w(1-w)\Phi\left(\frac{2}{\sqrt{\phi}}\right)-w(1-w)-(1-w)\Phi\left(\frac{1}{\sqrt{\phi}}\right) \tag{3.7.24}$$

Equation 3.7.24 is used to obtain ϕ numerically for a specific value of w.
The findings in Case A, Case B, and Case C for different values of w are presented in Table 3.5.

w	Small b	ψ approaches ∞	b and ψ both approach ∞
0.1	$\log(\Delta)=3.87$	–	$\log(\Delta/b^2)=3.30$
0.3	$\log(\Delta)=3.06$	–	$\log(\Delta/b^2)=2.11$
0.5	$\log(\Delta)=3.46$	–	$\log(\Delta/b^2)=1.72$
0.6	$\log(\Delta)=4.29$	–	$\log(\Delta/b^2)=1.60$
0.65	$\log(\Delta)=5.26$	–	$\log(\Delta/b^2)=1.55$
0.7	$\log(\Delta)=9.10$	–	$\log(\Delta/b^2)=1.51$
0.75	–	$\log(b)=-0.48$	$\log(\Delta/b^2)=1.47$
0.8	–	$\log(b)=-0.06$	$\log(\Delta/b^2)=1.44$
0.9	–	$\log(b)=0.42$	$\log(\Delta/b^2)=1.38$

Table 3.5: The zero contour equations of $\Delta\text{PFE}_{12}(\alpha)$ for the $\text{SMN}(w,0,1,\Delta)$ distribution, and selected values of w, when b is small, ψ approaches ∞, and both b and ψ approach ∞. The numbers in the second, third, and fourth column are rounded to two decimal places.

Table 3.5 shows that as w increases, ψ decreases until it reaches the minimum value of ψ, after that ψ continues to increase. The minimum value of ψ (ψ_{min}) is calculated numerically to be 22.1924 at $w=0.3154$. Moreover, ϕ decreases in w.

Hence, the minimum value of ϕ ($\phi_{\min}$) where $\Delta\text{PFE}_{12} \to 0$ is numerically found to be 3.777 at $w = 0.9999$. Furthermore, $\phi_{\min}$ and $\psi_{\min}$ respectively determine the lines

$$\log(\Delta) = 3.053643$$

and

$$\log(b) = 0.5(\log(\Delta) - 1.329)$$

These lines bound the area where $\Delta\text{PFE}_{12}(\alpha) \lessapprox 0$. See Figure 3.8 for illustration.

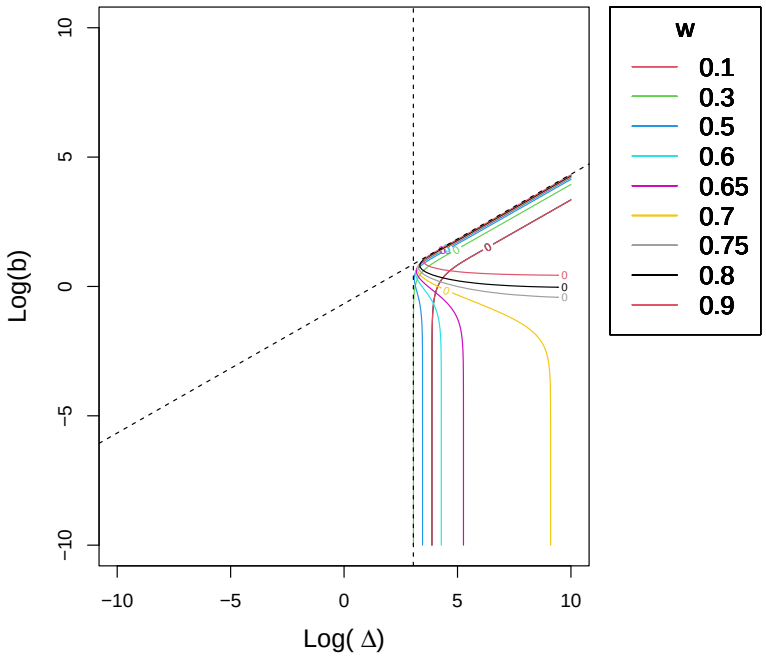
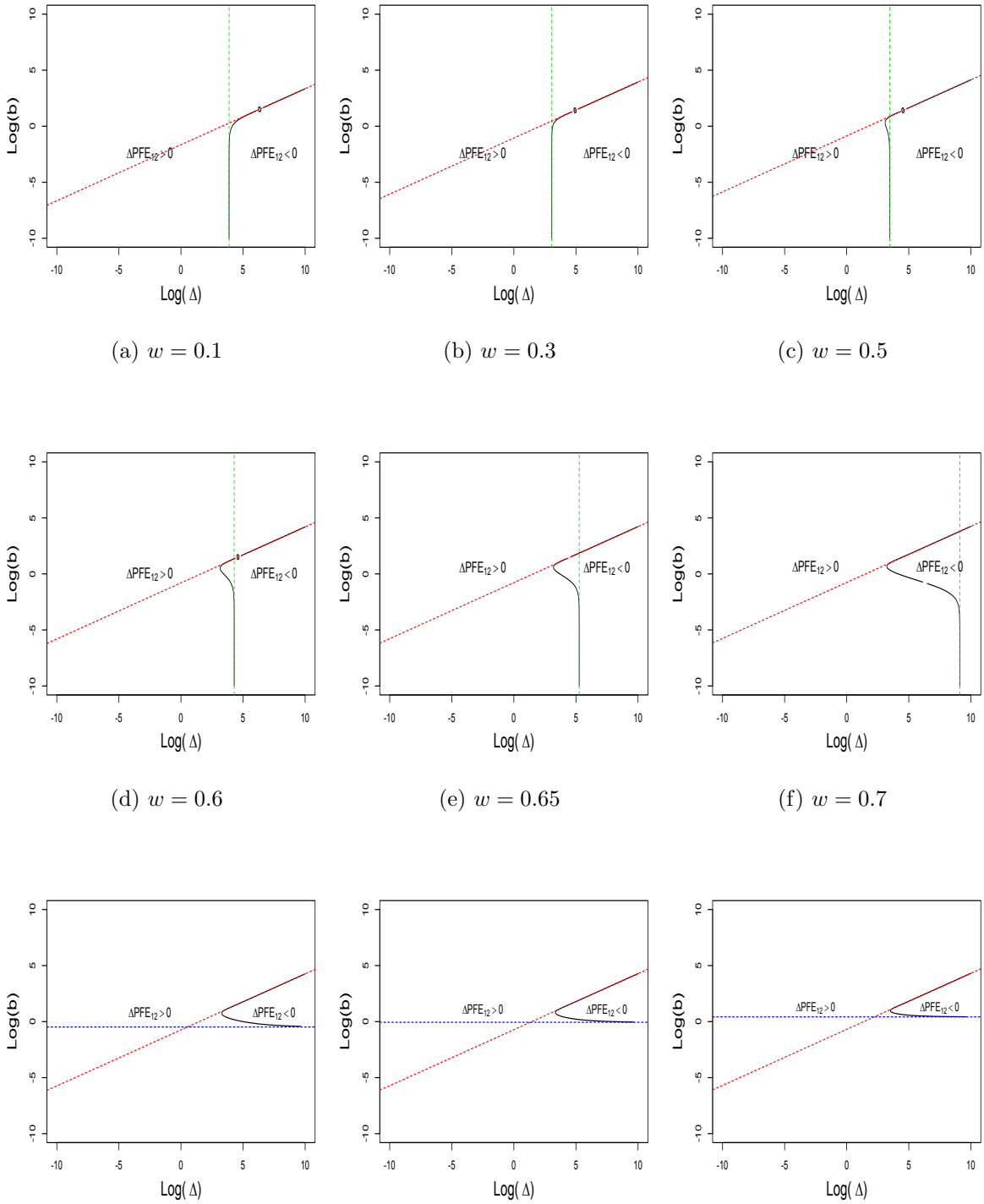


Figure 3.8: A Plot of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log(\Delta)$ (x-axis) and $\log(b)$ (y-axis) for the $\text{SMN}(w, 0, 1, \Delta)$ distribution, and selected values of w. The black dashed lines equations are $\log(\Delta) = 3.053643$ and $\log(c) = 0.5(\log(\Delta) - 1.32896)$.

In order to make comparative plots between the approximations of $\Delta\text{PFE}_{12}(\alpha)$ in the previous cases and the real contours, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ for different values of w is illustrated with the lines, the equations for which are presented in Table 3.5, in Figure 3.9.



(a) $w = 0.1$ (b) $w = 0.3$ (c) $w = 0.5$

(d) $w = 0.6$ (e) $w = 0.65$ (f) $w = 0.7$

(g) $w = 0.75$ (h) $w = 0.8$ (i) $w = 0.9$

Figure 3.9: Plots of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log\Delta$ (x-axis) and $\log(b)$ (y-axis)for the $\text{SMN}(w, 0, 1, \Delta)$ distribution. Each of the nine panels is related to a specific value of w. The equations of dashed lines are presented in Table 3.5.

3.8 A Two-Component Location Mixture of Normal Distributions

In this section, we consider that the distribution of X is a location mixture of normal (LMN) distribution having two components. The mixture components have means $-a$ and a and both have variances 1. The probability density function of $X \sim \text{LMN}(w, -a, a, 1)$ is

$$f_X(x) = w\phi(x+a) + (1-w)\phi(x-a);\ a > 0$$

3.8.1 MFE_n

If $X \sim \text{LMN}(w, -a, a, 1)$, MFE_n is obtained directly from Equation 3.7.6 as a special case to be

$$\begin{aligned}\text{MFE}_n &= \frac{1}{2} + \frac{1}{2}\sum_{k=0}^{n}\binom{n}{k}w^k(1-w)^{(n-k)} \\ &\times \left\{w\text{erf}\left(\frac{2a(n-k)-nc}{\sqrt{2(n^2+n)}}\right) + (1-w)\text{erf}\left(\frac{-2ak-nc}{\sqrt{2(n^2+n)}}\right)\right\}\end{aligned}$$

Therefore, MFE_1 and MFE_2 are respectively obtained to be

$$\begin{aligned}\text{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\Bigg\{w(1-w)\left[\text{erf}\left(\frac{2a-c}{2}\right) + \text{erf}\left(\frac{-2a-c}{2}\right)\right] \\ &\qquad +(w^2+(1-w)^2)\text{erf}\left(\frac{-c}{2}\right)\Bigg\} \qquad (3.8.1)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &= \frac{1}{2} + \frac{1}{2}\Bigg\{w(1-w)^2\left[\text{erf}\left(\frac{2a-c}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{-a-c}{\sqrt{3}}\right)\right] \\ &\qquad +w^2(1-w)\left[\text{erf}\left(\frac{-2a-c}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{a-c}{\sqrt{3}}\right)\right] \\ &\qquad +(w^3+(1-w)^3)\text{erf}\left(\frac{-c}{\sqrt{3}}\right)\Bigg\} \qquad (3.8.2)\end{aligned}$$

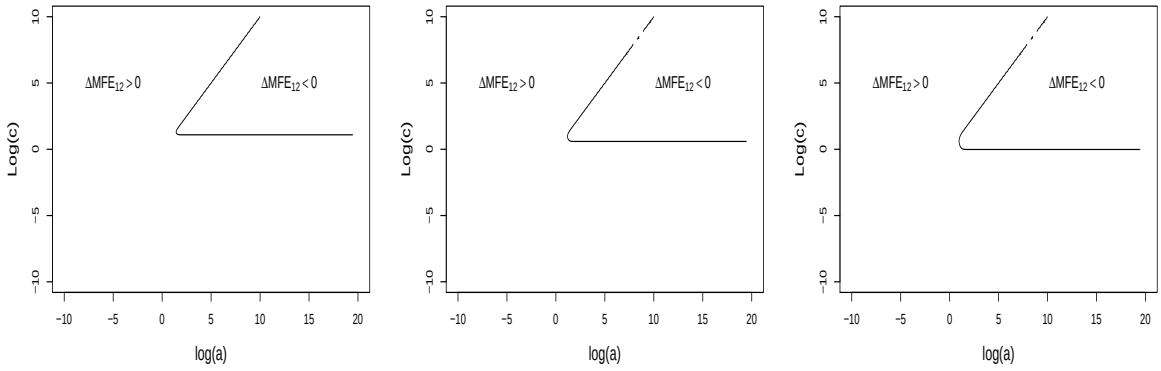
Recall Equation 3.7.11,

$$\Delta\text{MFE}_{12} = \text{MFE}_1 - \text{MFE}_2$$

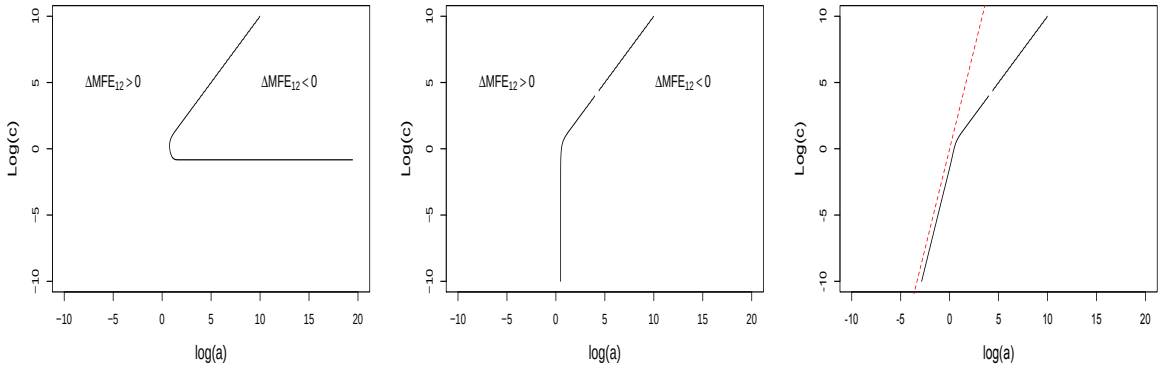
By using Equations 3.8.1 and 3.8.2, ΔMFE_{12} will be

$$\begin{aligned}
\Delta\text{MFE}_{12} &= \frac{w^3}{2}\Big[\text{erf}\,(\frac{-2a-c}{\sqrt{3}}) + 2\text{erf}\,(\frac{a-c}{\sqrt{3}}) - \text{erf}\,(\frac{2a-c}{\sqrt{3}}) - 2\text{erf}\,(\frac{-a-c}{\sqrt{3}})\Big] \\
&+ \frac{w^2}{2}\Big[2\text{erf}\,(\frac{-c}{2}) - 3\text{erf}\,(\frac{-c}{\sqrt{3}}) - \text{erf}\,(\frac{2a-c}{2}) - \text{erf}\,(\frac{-2a-c}{2}) \\
&\qquad +2\text{erf}\,(\frac{2a-c}{\sqrt{3}}) + 4\text{erf}\,(\frac{-a-c}{\sqrt{3}}) - 2\text{erf}\,(\frac{a-c}{\sqrt{3}}) - \text{erf}\,(\frac{-2a-c}{\sqrt{3}})\Big] \\
&+ \frac{w}{2}\Big[\text{erf}\,(\frac{2a-c}{2}) + \text{erf}\,(\frac{-2a-c}{2}) - 2\text{erf}\,(\frac{-c}{2}) + 3\text{erf}\,(\frac{-c}{\sqrt{3}}) \\
&\qquad -\text{erf}\,(\frac{2a-c}{\sqrt{3}}) - 2\text{erf}\,(\frac{-a-c}{\sqrt{3}})\Big] + \frac{1}{2}\Big[\text{erf}\,(\frac{-c}{2}) - \text{erf}\,(\frac{-c}{\sqrt{3}})\Big]
\end{aligned}$$

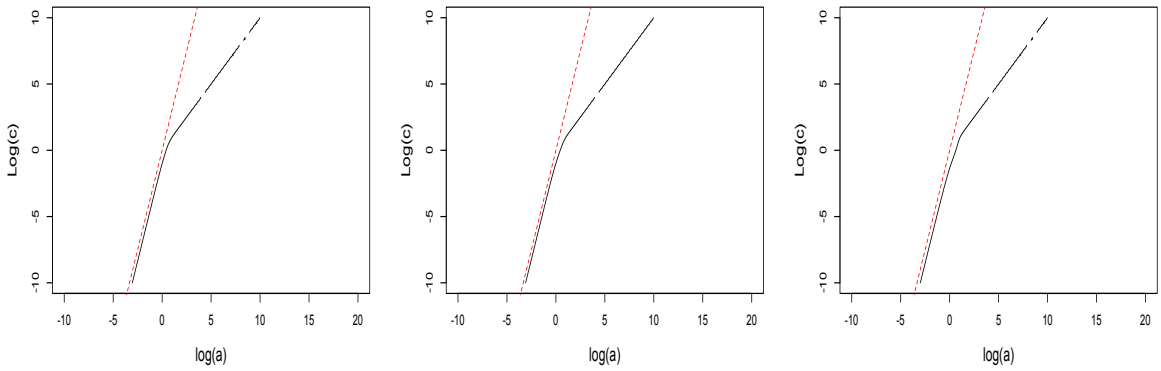
With respect to w, ΔMFE_{12} is a cubic function with coefficients that depend on c and a. Furthermore, ΔMFE_{12} has a value of zero if and only if the real root(s) of this cubic function is(are) restricted to be in the interval $(0, 1)$. Figure 3.10 illustrates the zero contour of ΔMFE_{12} for different values of w. The vertical axis represents a wide range of c, and the horizontal axis represents a. Both axes are taken on the log-scale.



(a) $w = 0.1$ (b) $w = 0.2$ (c) $w = 0.3$



(d) $w = 0.4$ (e) $w = 0.5$ (f) $w = 0.6$



(g) $w = 0.7$ (h) $w = 0.8$ (i) $w = 0.9$

Figure 3.10: Plots of the zero contour of ΔMFE_{12} with respect to $\log(a)$ (x-axis) and $\log(c)$ (y-axis) for the $\text{LMN}(w, -a, a, 1)$ distribution. Each of the nine panels is related to a specific value of w. The red dashed line's equation is $\log(c) = 3\log(a)$

As shown in Figure 3.10, the zero contour of ΔMFE_{12} has four distinct behaviours. Therefore, to bound the area where $\Delta\text{MFE}_{12} < 0$, we become interested in exploring the limiting behaviour of the zero contour of ΔMFE_{12} for the extreme values of parameters that are summarized in Table 3.6 and discussed afterwards in detail.

Case A: a and c both approach ∞ and $c = ra^m$		
m	r	ΔMFE_{12}
m<1	r>0	No zero contour
m=1	r <1	$-w^2(1-w)$
	r=1	0
	1<r<2	$w^2(1-w)$
	r=2	$\frac{w^2(1-w)}{2}$
	r>2	No zero contour
m>1	r>0	No zero contour
Case B: a approaches ∞		
$\Delta\text{MFE}_{12} \to (w^3+(1-w)^3)\Phi(\frac{2c}{\sqrt{6}}) - (w^2+(1-w)^2)\Phi(\frac{c}{\sqrt{2}}) + w(1-w)^2$		
Case C: Small c and $w = 0.5$		
$\Delta\text{MFE}_{12} \approx 0 \iff a = 1.63044$		
Case D: Small a and c approaches ∞		
No zero contour		
Case E: Both a and c small		
$\Delta\text{MFE}_{12} \approx 0 \iff \eta = \frac{3(\sqrt{3}-2)}{4(1-2w)w(1-w)};\ \eta = a^3/c$		

Table 3.6: Summary of the behaviour limiting of the zero contour of ΔMFE_{12} for the $\text{LMN}(w, -a, a, 1)$ distribution in 5 cases when: a and c both approach ∞, $a \to \infty$, small c and $w = 0.5$, small a and $c \to \infty$, both a and c small.

The detailed proof of the summary results in Table 3.6 is as follows:

Case A: a and c both approach ∞

This case corresponds to the straight line that appears in Figure 3.10 when a and c approach infinity. Since there are different possible lines, we will examine the case

when $c \propto a^m$ for $m < 1, m = 1$ and $m > 1$.

(1) $m < 1$

When $a \to \infty$, $a > a^m$ which means that a approaches infinity faster than a^m. Replace c by ra^m in Equations 3.8.1 and 3.8.2. Thus

$$\begin{aligned} \text{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\left[\text{erf}\left(\frac{2a - ra^m}{2}\right) + \text{erf}\left(\frac{-2a - ra^m}{2}\right)\right] \\ &\qquad + (w^2 + (1-w)^2)\text{erf}\left(\frac{-ra^m}{2}\right)\Bigg\} \end{aligned} \tag{3.8.3}$$

and

$$\begin{aligned} \text{MFE}_2 &= \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)^2\left[\text{erf}\left(\frac{2a - ra^m}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{-a - ra^m}{\sqrt{3}}\right)\right] \\ &\qquad + w^2(1-w)\left[\text{erf}\left(\frac{-2a - ra^m}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{a - ra^m}{\sqrt{3}}\right)\right] \\ &\qquad + (w^3 + (1-w)^3)\text{erf}\left(\frac{-ra^m}{\sqrt{3}}\right)\Bigg\} \end{aligned} \tag{3.8.4}$$

Let $a \to \infty$, Equations 3.8.3 and 3.8.4 become

$$\begin{aligned} \text{MFE}_1 &\to \frac{1}{2} + \frac{1}{2}\Big\{ w(1-w)\Big(\text{erf}\,(\infty) - \text{erf}\,(\infty)\Big) + (w^2 + (1-w)^2)\text{erf}\,(-\infty)\Big\} \\ &= \frac{1}{2} - \frac{1}{2}(w^2 + (1-w)^2) \end{aligned} \tag{3.8.5}$$

and

$$\begin{aligned} \text{MFE}_2 &\to \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)^2\Big(\text{erf}\,(\infty) - 2\text{erf}\,(\infty)\Big) \\ &\qquad + w^2(1-w)\Big(-\text{erf}\,(\infty) + 2\text{erf}\,(\infty)\Big) \\ &\qquad - (w^3 + (1-w)^3)\text{erf}\,(\infty)\Bigg\} \\ &= \frac{1}{2} + \frac{1}{2}\Big(w^2(1-w) - w(1-w)^2 - (w^3 + (1-w)^3)\Big) \end{aligned} \tag{3.8.6}$$

By using Equations 3.8.5 and 3.8.6, we obtain

$$\begin{aligned} \Delta\text{MFE}_{12} &\to \frac{1}{2}\Big(-w^2 - (1-w)^2 - w^2(1-w) + w(1-w)^2 + w^3 + (1-w)^3\Big) \\ &= \frac{1}{2}\Big(w^2(-1-(1-w)+w) + (1-w)^2(-1+w+(1-w))\Big) \\ &= -w^2(1-w) \end{aligned}$$

Since ΔMFE_{12} tends to negative value, there is no solution for this case. Consequently, no contours.

(2) $m = 1$

When $m = 1$, $c = ra$. Substitute c with ra in Equations 3.8.1 and 3.8.2, thus

$$\begin{aligned}\text{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\Bigg\{w(1-w)\left[\text{erf}\left(\frac{a(2-r)}{2}\right) + \text{erf}\left(\frac{-a(2+r)}{2}\right)\right] \\ &\quad +(w^2 + (1-w)^2)\text{erf}\left(\frac{-ra}{2}\right)\Bigg\} \qquad (3.8.7)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &= \frac{1}{2} + \frac{1}{2}\Bigg\{w(1-w)^2\left[\text{erf}\left(\frac{a(2-r)}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{-a(1+r)}{\sqrt{3}}\right)\right] \\ &\quad +w^2(1-w)\left[\text{erf}\left(\frac{-a(2+r)}{\sqrt{3}}\right) + 2\text{erf}\left(\frac{a(1-r)}{\sqrt{3}}\right)\right] \\ &\quad +(w^3 + (1-w)^3)\text{erf}\left(\frac{-ra}{\sqrt{3}}\right)\Bigg\} \qquad (3.8.8)\end{aligned}$$

As $a \to \infty$, $\text{erf}\left(\frac{a(2-r)}{2}\right)$ in Equations 3.8.7 and 3.8.8 tends to different values. These values are $-1, 0$ or 1 when $r < 2, r = 2$ or $r > 2$, respectively. Similarly, $\text{erf}\left(\frac{a(1-r)}{\sqrt{3}}\right)$ in Equation 3.8.8 tends to different values. These values are $-1, 0$ or 1 when $r < 1, r = 1$ or $r > 1$, respectively. Therefore, we shall study the behavior of ΔMFE_{12} when $r < 1, r = 1, 1 < r < 2, r = 2$, and $r > 2$.

Table 3.7 summarizes the above cases when $m = 1$.

r	ΔMFE_{12}
$r < 1$	$-w^2(1-w)$
$r = 1$	0
$1 < r < 2$	$w^2(1-w)$
$r = 2$	$\frac{w^2(1-w)}{2}$
$r > 2$	There is no solution.

Table 3.7: ΔMFE_{12} for the $\text{LMN}(w, -a, a, 1)$ distribution when a and c approach ∞, $c = ra$ and r has different values.

From the table shown above, $\Delta\text{MFE}_{12} > 0$ when $c = ra;\ r > 1$.

(2.1) $r < 1$

Let $r < 1$ in Equations 3.8.7 and 3.8.8. Moreover, let a approach ∞, hence

$$\begin{aligned}\text{MFE}_1 &\rightarrow \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\text{erf}\,(\infty) + \big(w(1-w) + w^2 + (1-w)^2\big)\text{erf}\,(-\infty)\Bigg\} \\ &= \frac{1}{2} - \frac{1}{2}(w^2 + (1-w)^2) \qquad (3.8.9)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &\rightarrow \frac{1}{2} + \frac{1}{2}\Bigg\{ \big(w(1-w)^2 + 2w^2(1-w)\big)\text{erf}\,(\infty) \\ &\qquad + \big(2w(1-w)^2 + w^2(1-w) + w^3 + (1-w)^3\big)\text{erf}\,(-\infty)\Bigg\} \\ &= \frac{1}{2} + \frac{1}{2}\Big(- w(1-w)^2 + w^2(1-w) - (w^3 + (1-w)^3)\Big) \qquad (3.8.10)\end{aligned}$$

Using Equations 3.8.9 and 3.8.10,

$$\begin{aligned}\Delta\text{MFE}_{12} &\rightarrow \frac{1}{2}\Big(- w^2 - (1-w)^2 + w(1-w)^2 - w^2(1-w) + w^3 + (1-w)^3\Big) \\ &= \frac{1}{2}\Big(w^2(-1-(1-w)+w) + (1-w)^2(-1+w+(1-w))\Big) \\ &= -w^2(1-w)\end{aligned}$$

Since ΔMFE_{12} tends to negative value, there is no solution for this case. Consequently, no contours.

(2.2) $r = 1$

Let $r = 1$ in Equations 3.8.7 and 3.8.8,

$$\begin{aligned}\text{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\Bigg[\text{erf}\,\big(\frac{a}{2}\big) + \text{erf}\,\big(\frac{-3a}{2}\big)\Bigg] \\ &\qquad + (w^2 + (1-w)^2)\text{erf}\,\big(\frac{-a}{2}\big)\Bigg\}\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &= \frac{1}{2}+\frac{1}{2}\Bigg\{w(1-w)^2\left[\text{erf}\left(\frac{a}{\sqrt{3}}\right)+2\text{erf}\left(\frac{-2a}{\sqrt{3}}\right)\right]\\ &\qquad +w^2(1-w)\left[\text{erf}\left(\frac{-3a}{\sqrt{3}}\right)+2\text{erf}\,(0)\right]\\ &\qquad +(w^3+(1-w)^3)\text{erf}\left(\frac{-a}{\sqrt{3}}\right)\Bigg\}\end{aligned}$$

Set $a \to \infty$ in the previous equations, thus

$$\text{MFE}_1 \;\to\; \frac{1}{2}-\frac{1}{2}(w^2+(1-w)^2) \tag{3.8.11}$$

and

$$\begin{aligned}\text{MFE}_2 &\to \frac{1}{2}-\frac{1}{2}\Big(w(1-w)^2+w^2(1-w)+w^3+(1-w)^3\Big)\\ &= \frac{1}{2}-\frac{1}{2}\Big(w^2+(1-w)^2\Big)\end{aligned} \tag{3.8.12}$$

Using Equations 3.8.11 and 3.8.12. Hence,

$$\Delta\text{MFE}_{12} \to 0$$

(2.3) $1 < r < 2$

Let $1 < r < 2$ in Equations 3.8.7 and 3.8.8. Moreover, let a tend to ∞, so we obtain

$$\begin{aligned}\text{MFE}_1 &\to \frac{1}{2}+\frac{1}{2}\Bigg\{w(1-w)\left[\text{erf}\,(\infty)+\text{erf}\,(-\infty)\right]\\ &\qquad +(w^2+(1-w)^2)\text{erf}\,(-\infty)\Bigg\}\\ &= \frac{1}{2}-\frac{1}{2}(w^2+(1-w)^2)\end{aligned} \tag{3.8.13}$$

and

$$\begin{aligned}\text{MFE}_2 &\to \frac{1}{2}+\frac{1}{2}\Bigg\{w(1-w)^2\left[\text{erf}\,(\infty)+2\text{erf}\,(-\infty)\right]\\ &\qquad +w^2(1-w)\left[\text{erf}\,(-\infty)+2\text{erf}\,(-\infty)\right]\\ &\qquad +(w^3+(1-w)^3)\text{erf}\,(-\infty)\Bigg\}\\ &= \frac{1}{2}+\frac{1}{2}\Big(-w(1-w)^2-3w^2(1-w)-(w^3+(1-w)^3)\Big)\end{aligned} \tag{3.8.14}$$

Using Equations 3.8.13 and 3.8.14. Thus

$$\begin{aligned}\Delta\text{MFE}_{12} &\rightarrow \frac{1}{2}\Big(-w^2-(1-w)^2+w(1-w)^2+3w^2(1-w)+w^3+(1-w)^3\Big)\\ &= \frac{1}{2}\Big(w^2\Big(-1+3(1-w)+w\Big)+(1-w)^2\Big(-1+w+(1-w)\Big)\Big)\\ &= \frac{1}{2}\Big(2w^2(1-w)\Big)\\ &= w^2(1-w)\end{aligned}$$

(2.4) $r = 2$

Set $r = 2$ in Equations 3.8.7 and 3.8.8. Then let a tend to ∞, hence

$$\begin{aligned}\text{MFE}_1 &\rightarrow \frac{1}{2}+\frac{1}{2}\Bigg\{w(1-w)\Big[\text{erf}\,(0)+\text{erf}\,(-\infty)\Big]\\ &\qquad\qquad +(w^2+(1-w)^2)\text{erf}\,(-\infty)\Bigg\}\\ &= \frac{1}{2}-\frac{1}{2}(w(1-w)+w^2+(1-w)^2)\\ &= \frac{1}{2}-\frac{1}{2}(1-w(1-w)) \qquad (3.8.15)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &\rightarrow \frac{1}{2}+\frac{1}{2}\Bigg\{w(1-w)^2\Big[\text{erf}\,(0)+2\text{erf}\,(-\infty)\Big]\\ &\qquad +w^2(1-w)\Big[\text{erf}\,(-\infty)+2\text{erf}\,(-\infty)\Big]\\ &\qquad\qquad +(w^3+(1-w)^3)\text{erf}\,(-\infty)\Bigg\}\\ &= \frac{1}{2}-\frac{1}{2}\Big(2w(1-w)^2+3w^2(1-w)+w^3+(1-w)^3\Big)\\ &= \frac{1}{2}-\frac{1}{2}(1-w(1-w)^2) \qquad (3.8.16)\end{aligned}$$

Using Equations 3.8.15 and 3.8.16. Thus

$$\begin{aligned}\Delta\text{MFE}_{12} &\rightarrow \frac{1}{2}\Big(w(1-w)-w(1-w)^2\Big)\\ &= \frac{1}{2}\Big(w(1-w)\big(1-(1-w)\big)\Big)\\ &= \frac{1}{2}w^2(1-w)\end{aligned}$$

(2.5) $r > 2$

Equations 3.8.7 and 3.8.8 are respectively equivalent to

$$\begin{aligned} \mathrm{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\left[\mathrm{erf}\left(\frac{-a(r-2)}{2}\right) + \mathrm{erf}\left(\frac{-a(r+2)}{2}\right)\right] \\ &\qquad + (w^2 + (1-w)^2)\mathrm{erf}\left(\frac{-ra}{2}\right)\Bigg\} \end{aligned} \tag{3.8.17}$$

and

$$\begin{aligned} \mathrm{MFE}_2 &= \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)^2\left[\mathrm{erf}\left(\frac{-a(r-2)}{\sqrt{3}}\right) + 2\mathrm{erf}\left(\frac{-a(r+1)}{\sqrt{3}}\right)\right] \\ &\qquad + w^2(1-w)\left[\mathrm{erf}\left(\frac{-a(r+2)}{\sqrt{3}}\right) + 2\mathrm{erf}\left(\frac{-a(r-1)}{\sqrt{3}}\right)\right] \\ &\qquad + (w^3 + (1-w)^3)\mathrm{erf}\left(\frac{-ra}{\sqrt{3}}\right)\Bigg\} \end{aligned} \tag{3.8.18}$$

Let $r > 2$ and a tend to ∞ in Equations 3.8.17 and 3.8.18, thus we obtain

$$\begin{aligned} \mathrm{MFE}_1 &\to \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\Big[\mathrm{erf}\,\big(-\infty\big) + \mathrm{erf}\,\big(-\infty\big)\Big] \\ &\qquad + \big(w^2 + (1-w)^2\big)\mathrm{erf}\,\big(-\infty\big)\Bigg\} \\ &= \frac{1}{2} + \frac{1}{2}\Big(-2w(1-w) - (w^2 + (1-w)^2)\Big) \\ &= 0 \end{aligned} \tag{3.8.19}$$

and

$$\begin{aligned} \mathrm{MFE}_2 &\to \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)^2\Big[\mathrm{erf}\,\big(-\infty\big) + 2\mathrm{erf}\,\big(-\infty\big)\Big] \\ &\qquad + w^2(1-w)\Big[\mathrm{erf}\,\big(-\infty\big) + 2\mathrm{erf}\,\big(-\infty\big)\Big] \\ &\qquad + (w^3 + (1-w)^3)\mathrm{erf}\,\big(-\infty\big)\Bigg\} \\ &= \frac{1}{2} + \frac{1}{2}\Big(-3w(1-w)^2 - 3w^2(1-w) - (w^3 + (1-w)^3)\Big) \\ &= 0 \end{aligned} \tag{3.8.20}$$

From Equations 3.8.19 and 3.8.20, MFE_1 and MFE_2 are zeros as a approaches ∞. Therefore, we shall use L'Hôspital's rule as follows:

$$\lim_{a\to\infty} \frac{\mathrm{MFE}_1}{\mathrm{MFE}_2} = \lim_{a\to\infty} \frac{\frac{\partial}{\partial a}\mathrm{MFE}_1}{\frac{\partial}{\partial a}\mathrm{MFE}_2}$$

Differentiate Equations 3.8.17 and 3.8.18 with respect to a, hence

$$\begin{aligned}\frac{\partial}{\partial a}\text{MFE}_1 &= \frac{w(1-w)}{\sqrt{\pi}}\Bigg[-\frac{(r-2)}{2}\exp\Big(-\big(\frac{a(r-2)}{2}\big)^2\Big) \\ &\qquad -\frac{(r+2)}{2}\exp\Big(-\big(\frac{a(r+2)}{2}\big)^2\Big)\Bigg] \\ &- \frac{r(w^2+(1-w)^2)}{2\sqrt{\pi}}\exp\Big(-\big(\frac{ra}{2}\big)^2\Big) \qquad (3.8.21)\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial a}\text{MFE}_2 &= \frac{w(1-w)^2}{\sqrt{\pi}}\Bigg[-\frac{(r-2)}{\sqrt{3}}\exp\Big(-\big(\frac{a(r-2)}{\sqrt{3}}\big)^2\Big) \\ &\qquad -\frac{2(r+1)}{\sqrt{3}}\exp\Big(-\big(\frac{a(r+1)}{\sqrt{3}}\big)^2\Big)\Bigg] \\ &+ \frac{w^2(1-w)}{\sqrt{\pi}}\Bigg[-\frac{(r+2)}{\sqrt{3}}\exp\Big(-\big(\frac{a(r+2)}{\sqrt{3}}\big)^2\Big) \\ &\qquad -\frac{2(r-1)}{\sqrt{3}}\exp\Big(-\big(\frac{a(r-1)}{\sqrt{3}}\big)^2\Big)\Bigg] \\ &- \frac{r(w^3+(1-w)^3)}{\sqrt{3\pi}}\exp\Big(-\big(\frac{ra}{\sqrt{3}}\big)^2\Big) \qquad (3.8.22)\end{aligned}$$

Divide Equations 3.8.21 and 3.8.22 by $\exp\Big(-a^2(\frac{r-2}{2})^2\Big)$. Notice that as $a \to \infty$,
$\exp\left(-a^2\Big(\big(\frac{r+2}{2}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$, $\exp\left(-a^2\Big(\big(\frac{r}{2}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$,
$\exp\left(-a^2\Big(\big(\frac{r-2}{\sqrt{3}}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$, $\exp\left(-a^2\Big(\big(\frac{r+1}{\sqrt{3}}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$,
$\exp\left(-a^2\Big(\big(\frac{r+2}{\sqrt{3}}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$, $\exp\left(-a^2\Big(\big(\frac{r-1}{\sqrt{3}}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$,
and $\exp\left(-a^2\Big(\big(\frac{r}{\sqrt{3}}\big)^2-\big(\frac{r-2}{2}\big)^2\Big)\right)$ tend to zero. Thus

$$\begin{aligned}\lim_{a\to\infty}\frac{\frac{\partial}{\partial a}\text{MFE}_1}{\frac{\partial}{\partial a}\text{MFE}_2} &\to \frac{-w(1-w)\frac{(r-2)}{2\sqrt{\pi}}}{0^-} \\ &= \infty\end{aligned}$$

Hence, there is no solution for this case. Furthermore, there are no contours. □

(3) $m > 1$

When $a \to \infty$, $ra^m > 2a$. Let $a \to \infty$ in Equations 3.8.3 and 3.8.4, thus

$$\begin{aligned}
\text{MFE}_1 &\to \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)\Big[\text{erf}\big(-\infty\big) + \text{erf}\big(-\infty\big)\Big] \\
&\qquad + (w^2 + (1-w)^2)\text{erf}\big(-\infty\big)\Bigg\} \\
&= \frac{1}{2} + \frac{1}{2}\Big(-2w(1-w) - (w^2 + (1-w)^2)\Big) \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\text{MFE}_2 &\to \frac{1}{2} + \frac{1}{2}\Bigg\{ w(1-w)^2\Big[\text{erf}\big(-\infty\big) + 2\text{erf}\big(-\infty\big)\Big] \\
&\qquad + w^2(1-w)\Big[\text{erf}\big(-\infty\big) + 2\text{erf}\big(-\infty\big)\Big] \\
&\qquad + (w^3 + (1-w)^3)\text{erf}\big(-\infty\big)\Bigg\} \\
&= \frac{1}{2} + \frac{1}{2}\Big(-3w(1-w)^2 - 3w^2(1-w) - (w^3 + (1-w)^3)\Big) \\
&= 0
\end{aligned}$$

MFE_1 and MFE_2 are zeros when a tends to ∞. Therefore, we shall use L'Hôspital's rule as follows:

$$\lim_{a\to\infty} \frac{\text{MFE}_1}{\text{MFE}_2} = \lim_{a\to\infty} \frac{\frac{\partial}{\partial a}\text{MFE}_1}{\frac{\partial}{\partial a}\text{MFE}_2}$$

Differentiate Equations 3.8.3 and 3.8.4 with respect to a, hence

$$\begin{aligned}
\frac{\partial}{\partial a}\text{MFE}_1 &= \frac{w(1-w)}{\sqrt{\pi}}\Bigg[\frac{(2 - rma^{m-1})}{2}\exp\Big(-\big(\frac{ra^m - 2a}{2}\big)^2\Big) \\
&\qquad - \frac{(2 + rma^{m-1})}{2}\exp\Big(-\big(\frac{ra^m + 2a}{2}\big)^2\Big)\Bigg] \\
&- \frac{rma^{m-1}(w^2 + (1-w)^2)}{2\sqrt{\pi}}\exp\Big(-\big(\frac{ra^m}{2}\big)^2\Big) \qquad (3.8.23)
\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial a}\text{MFE}_2 &= \frac{w(1-w)^2}{\sqrt{\pi}}\left[\frac{(2-rma^{m-1})}{\sqrt{3}}\exp\Big(-\big(\frac{ra^m-2a}{\sqrt{3}}\big)^2\Big)\right.\\ &\qquad\left.-\frac{2(rma^{m-1}+1)}{\sqrt{3}}\exp\Big(-\big(\frac{ra^m+a}{\sqrt{3}}\big)^2\Big)\right]\\ &+ \frac{w^2(1-w)}{\sqrt{\pi}}\left[-\frac{(2+rma^{m-1})}{\sqrt{3}}\exp\Big(-\big(\frac{ra^m+2a}{\sqrt{3}}\big)^2\Big)\right.\\ &\qquad\left.+\frac{2(1-rma^{m-1})}{\sqrt{3}}\exp\Big(-\big(\frac{ra^m-a}{\sqrt{3}}\big)^2\Big)\right]\\ &- \frac{rma^{m-1}(w^3+(1-w)^3)}{\sqrt{3\pi}}\exp\Big(-\big(\frac{ra^m}{\sqrt{3}}\big)^2\Big) \qquad (3.8.24)\end{aligned}$$

Divide Equations 3.8.23 and 3.8.24 by $\frac{\exp\left(-\left(\frac{ra^m-2a}{2}\right)^2\right)}{a^{m-1}}$. Notice that as $a\to\infty$, $\exp\left(\left(\frac{ra^m+2a}{2}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$, $\exp\left(\left(\frac{ra^m}{2}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$, $\exp\left(\left(\frac{ra^m-2a}{\sqrt{3}}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$, $\exp\left(\left(\frac{ra^m+a}{\sqrt{3}}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$, $\exp\left(\left(\frac{ra^m+2a}{\sqrt{3}}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)\exp\left(\left(\frac{ra^m-a}{\sqrt{3}}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$, and $\exp\left(\left(\frac{ra^m}{\sqrt{3}}\right)^2-\left(\frac{ra^m-2a}{2}\right)^2\right)$ tend to ∞. Thus

$$\begin{aligned}\lim_{a\to\infty}\frac{\frac{\partial}{\partial a}\text{MFE}_1}{\frac{\partial}{\partial a}\text{MFE}_2} &\to \frac{\frac{-rmw(1-w)}{2\sqrt{\pi}}}{0^-}\\ &= \infty\end{aligned}$$

Therefore, there is no solution for this case. Furthermore, there are no contours.

Case B: a approaches ∞

When $a\to\infty$, the zero contour of ΔMFE_{12} is a horizontal line. The equations of the horizontal lines in Figure 3.10 are obtained as follows:

Let a tend to ∞ in Equations 3.8.1 and 3.8.2, thus

$$\begin{aligned}\text{MFE}_1 &\to \frac{1}{2}-\frac{1}{2}(w^2+(1-w)^2)\text{erf}\left(\frac{c}{2}\right)\\ &= 1-w(1-w)-(w^2+(1-w)^2)\Phi(\frac{c}{\sqrt{2}}) \qquad (3.8.25)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &\to \frac{1}{2}+\frac{1}{2}\left\{w(1-w)(2w-1)-(w^3+(1-w)^3)\text{erf}\left(\frac{c}{\sqrt{3}}\right)\right\}\\ &= \frac{1}{2}-(w^3+(1-w)^3)\Phi\left(\frac{2c}{\sqrt{6}}\right)+\frac{w(1-w)(2w-1)+(w^3+(1-w)^3)}{2}\end{aligned}$$

Consequently,

$$\Delta\text{MFE}_{12} \to (w^3 + (1-w)^3)\Phi(\frac{2c}{\sqrt{6}}) - (w^2 + (1-w)^2)\Phi(\frac{c}{\sqrt{2}}) + w(1-w)^2 \quad (3.8.26)$$

Equation 3.8.26 is used to obtain c that makes $\Delta\text{MFE}_{12} \to 0$ for a specific value of w, which represents the equation of the horizontal line.

Figure 3.11 presents ΔMFE_{12} in Equation 3.8.26 as a function of c for selected values of w.

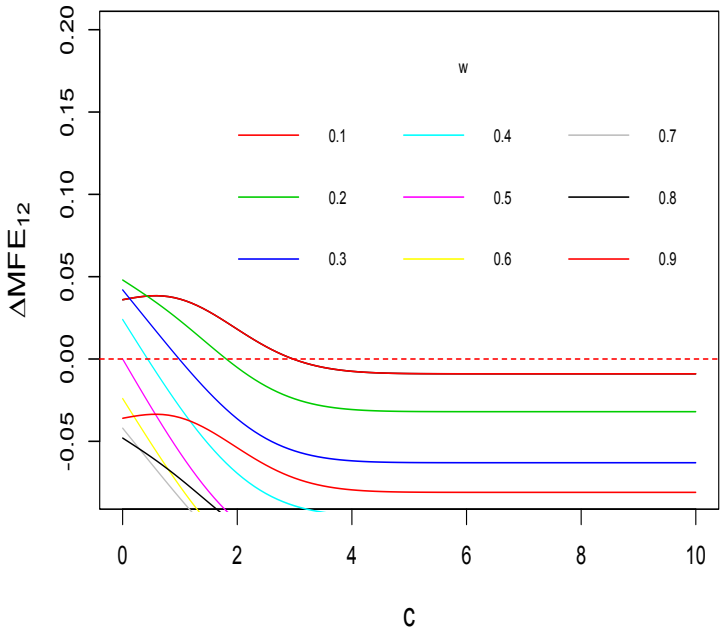


Figure 3.11: ΔMFE_{12} versus c for the LMN$(w, -a, a, 1)$ distribution and selected values of w when $a \to \infty$ (Equation 3.8.26).

Figure 3.11 shows that ΔMFE_{12} could tend to zero when $w < 0.5$.

Case C: Small c and $w = 0.5$

This case represents the vertical line in Figure 3.10. To obtain its equation, set $w = (1-w)$ in Equations 3.8.1 and 3.8.2, thus

$$\begin{aligned}\text{MFE}_1 &= \frac{1}{2} + \frac{1}{2}\left\{\frac{1}{4}\left[\text{erf}\left(\frac{2a-c}{2}\right) + \text{erf}\left(\frac{-2a-c}{2}\right)\right]\right. \\ &\qquad \left. + \frac{1}{2}\text{erf}\left(\frac{-c}{2}\right)\right\} \qquad (3.8.27)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &= \frac{1}{2}+\frac{1}{2}\Bigg\{\frac{1}{8}\text{erf}\left(\frac{2a-c}{\sqrt{3}}\right)+\frac{1}{4}\text{erf}\left(\frac{-a-c}{\sqrt{3}}\right)+\frac{1}{8}\text{erf}\left(\frac{-2a-c}{\sqrt{3}}\right)\\ &\qquad +\frac{1}{4}\text{erf}\left(\frac{a-c}{\sqrt{3}}\right)+\frac{1}{4}\text{erf}\left(\frac{-c}{\sqrt{3}}\right)\Bigg\} \qquad (3.8.28)\end{aligned}$$

For small c in Equations 3.8.27 and 3.8.28, we obtain

$$\text{MFE}_1 \approx \frac{1}{2}-\frac{c}{2\sqrt{\pi}}\left(\frac{e^{-a^2}}{2}+\frac{1}{2}\right)$$

and

$$\text{MFE}_2 \approx \frac{1}{2}-\frac{c}{2\sqrt{\pi}}\left(\frac{e^{\frac{-4a^2}{3}}}{2\sqrt{3}}+\frac{e^{-\frac{a^2}{3}}}{\sqrt{3}}+\frac{1}{2\sqrt{3}}\right)$$

Consequently,

$$\Delta\text{MFE}_{12} \approx \frac{c}{2\sqrt{\pi}}\left(\frac{e^{-(4a^2/3)}+2e^{-(a^2/3)}-\sqrt{3}e^{-a^2}+1-\sqrt{3}}{2\sqrt{3}}\right)$$

As a result,

$$\Delta\text{MFE}_{12} \approx 0 \iff a = 1.63044$$

by numerical solution.

Case D: Small a and c approaches ∞

By using the first-order Maclaurin series expansion of each error function with respect to a in Equations 3.8.1 and 3.8.2, we obtain

$$\begin{aligned}\text{MFE}_1 &\approx \frac{1}{2}+\frac{1}{2}\Bigg(w(1-w)\Big(\text{erf}\,(\frac{-c}{2})+\frac{2}{\sqrt{\pi}}e^{-c^2/4}a\Big)\\ &\qquad -(w^2+(1-w)^2)\text{erf}\,(\frac{c}{2})-w(1-w)\Big(\text{erf}\,(\frac{c}{2})+\frac{2}{\sqrt{\pi}}e^{-c^2/4}a\Big)\Bigg)\\ &= \frac{1}{2}-\frac{1}{2}\text{erf}\,(\frac{c}{2}) \qquad (3.8.29)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &\approx \frac{1}{2}+\frac{1}{2}\Bigg(w(1-w)^2\Big(\text{erf}\,(\frac{-c}{\sqrt{3}})+\frac{4}{\sqrt{3\pi}}e^{-c^2/3}a\Big)\\ &\qquad +2w^2(1-w)\Big(\text{erf}\,(\frac{-c}{\sqrt{3}})+\frac{2}{\sqrt{3\pi}}e^{-c^2/3}a\Big)\\ &\qquad -2w(1-w)^2\Big(\text{erf}\,(\frac{c}{\sqrt{3}})+\frac{2}{\sqrt{3\pi}}e^{-c^2/3}a\Big)\\ &\qquad -w^2(1-w)\Big(\text{erf}\,(\frac{c}{\sqrt{3}})+\frac{4}{\sqrt{3\pi}}e^{-c^2/3}a\Big)\\ &\qquad -(w^3+(1-w)^3)\text{erf}\,(\frac{c}{\sqrt{3}})\Bigg)\end{aligned}$$

Hence,

$$\text{MFE}_2 \approx \frac{1}{2} - \frac{1}{2}\text{erf}\,(\frac{c}{\sqrt{3}}) \tag{3.8.30}$$

As $c \to \infty$, both MFE_1 and MFE_2 are approximately zero. Therefore, we shall use L'Hôspital's rule as follows:

$$\lim_{c\to\infty} \frac{\text{MFE}_1}{\text{MFE}_2} = \lim_{c\to\infty} \frac{\frac{\partial}{\partial c}\text{MFE}_1}{\frac{\partial}{\partial c}\text{MFE}_2}$$

Differentiate Equations 3.8.29 and 3.8.30 with respect to c,

$$\frac{\partial}{\partial c}\text{MFE}_1 \approx \frac{-e^{-\frac{c^2}{4}}}{2\sqrt{\pi}}$$

and

$$\frac{\partial}{\partial c}\text{MFE}_2 \approx \frac{-e^{-\frac{c^2}{3}}}{\sqrt{3\pi}}$$

Hence,

$$\begin{aligned}
\lim_{c\to\infty} \frac{\frac{\partial}{\partial c}\text{MFE}_1}{\frac{\partial}{\partial c}\text{MFE}_2} &\approx \lim_{c\to\infty} \frac{\sqrt{3}e^{-\frac{c^2}{4}}}{2e^{-\frac{c^2}{3}}} \\
&= \lim_{c\to\infty} \frac{\sqrt{3}}{2e^{-\frac{c^2}{12}}} \\
&\to \frac{\sqrt{3}}{0} \\
&= \infty
\end{aligned}$$

Hence, there is no solution for this case. Furthermore, there are no contours.

Case E: Both a and c small

In this case, the zero contour of ΔMFE_{12} forms the line that is parallel to

$$\log(c) = 3\log(a)$$

The equation of such a line is

$$\log(c) = 3\log(a) - \log(\eta)$$

Hence, our objective is to obtain $\eta = a^3/c$.

By using the third-order Maclaurin series expansion of each error function with

respect to a in Equations 3.8.1 and 3.8.2, we obtain

$$\begin{aligned}\text{MFE}_1 &\approx \frac{1}{2}+\frac{1}{2}\Bigg(w(1-w)\Big(\text{erf}\,(\frac{-c}{2})+\frac{2}{\sqrt{\pi}}e^{-c^2/4}\big(a+c\frac{a^2}{2}+(c^2-2)\frac{a^3}{6}\big)\Big)\\ &\qquad -(w^2+(1-w)^2)\text{erf}\,(\frac{c}{2})\\ &\qquad -w(1-w)\Big(\text{erf}\,(\frac{c}{2})+\frac{2}{\sqrt{\pi}}e^{-c^2/4}\big(a-c\frac{a^2}{2}+(c^2-2)\frac{a^3}{6}\big)\Big)\Bigg)\\ &= \frac{1}{2}+\frac{1}{2}\Bigg(w(1-w)\frac{2ca^2}{\sqrt{\pi}}e^{-c^2/4}-\text{erf}\,(\frac{c}{2})\Bigg) \qquad (3.8.31)\end{aligned}$$

and

$$\begin{aligned}\text{MFE}_2 &\approx \frac{1}{2}+\frac{1}{2}\Bigg(w(1-w)^2\Big(\text{erf}\,(\frac{-c}{\sqrt{3}})+\frac{4}{\sqrt{3\pi}}e^{-c^2/3}\big(a+\frac{2}{3}ca^2+\frac{4}{9}(\frac{2}{3}c^2-1)a^3\big)\Big)\\ &\qquad +2w^2(1-w)\Big(\text{erf}\,(\frac{-c}{\sqrt{3}})+\frac{2}{\sqrt{3\pi}}e^{-c^2/3}\big(a+\frac{1}{3}ca^2+\frac{1}{9}(\frac{2}{3}c^2-1)a^3\big)\Big)\\ &\qquad -2w(1-w)^2\Big(\text{erf}\,(\frac{c}{\sqrt{3}})+\frac{2}{\sqrt{3\pi}}e^{-c^2/3}\big(a-\frac{1}{3}ca^2+\frac{1}{9}(\frac{2}{3}c^2-1)a^3\big)\Big)\\ &\qquad -w^2(1-w)\Big(\text{erf}\,(\frac{c}{\sqrt{3}})+\frac{4}{\sqrt{3\pi}}e^{-c^2/3}\big(a-\frac{2}{3}ca^2+\frac{4}{9}(\frac{2}{3}c^2-1)a^3\big)\Big)\\ &\qquad -(w^3+(1-w)^3)\text{erf}\,(\frac{c}{\sqrt{3}})\Bigg)\\ &= \frac{1}{2}+\frac{1}{2}\Bigg(-\text{erf}\,(\frac{c}{\sqrt{3}})+\frac{4a^2w(1-w)}{\sqrt{3\pi}}e^{-c^2/3}\Big(c+\frac{a}{3}(1-2w)\big(\frac{2}{3}c^2-1\big)\Big)\Bigg)\end{aligned} \qquad (3.8.32)$$

For small c, Equations 3.8.31 and 3.8.32 respectively become

$$\text{MFE}_1 \approx \frac{1}{2}+\frac{c}{2\sqrt{\pi}}\Big(2w(1-w)a^2e^{-c^2/4}-1\Big) \qquad (3.8.33)$$

and

$$\begin{aligned}\text{MFE}_2 &\approx \frac{1}{2}+\frac{1}{2}\Bigg(-\frac{2c}{\sqrt{3\pi}}+\frac{4a^2w(1-w)}{\sqrt{3\pi}}e^{-c^2/3}\Big(c+\frac{a}{3}(1-2w)\big(\frac{2}{3}c^2-1\big)\Big)\Bigg)\\ &= \frac{1}{2}+\frac{c}{2\sqrt{\pi}}\Bigg(-\frac{2}{\sqrt{3}}+\frac{4w(1-w)}{\sqrt{3}}e^{-c^2/3}\Big(a^2+\frac{(1-2w)}{3}\big(\frac{2}{3}ca^3-a^3/c\big)\Big)\Bigg)\end{aligned}$$

Set $\eta = a^3/c$,

$$\text{MFE}_2 \approx \frac{1}{2}+\frac{c}{2\sqrt{\pi}}\Bigg(-\frac{2}{\sqrt{3}}+\frac{4w(1-w)}{\sqrt{3}}e^{-c^2/3}\Big(a^2+\frac{(1-2w)}{3}\big(\frac{2}{3}ca^3-\eta\big)\Big)\Bigg) \qquad (3.8.34)$$

By using Equations 3.8.33 and 3.8.34, we obtain

$$\begin{aligned}\Delta\text{MFE}_{12} &\approx \frac{c}{2\sqrt{\pi}}\Bigg(2w(1-w)a^2e^{-c^2/4}-1+\frac{2}{\sqrt{3}}\\ &- \frac{4w(1-w)}{\sqrt{3}}e^{-c^2/3}\Big(a^2+\frac{(1-2w)}{3}\big(\frac{2}{3}ca^3-\eta\big)\Big)\Bigg)\end{aligned}$$

$\Delta\text{MFE}_{12} \approx 0$, if and only if

$$2w(1-w)a^2e^{-c^2/4}-1+\frac{2}{\sqrt{3}}-\frac{4w(1-w)}{\sqrt{3}}e^{-c^2/3}\Big(a^2+\frac{(1-2w)}{3}\big(\frac{2}{3}ca^3-\eta\big)\Big) = 0$$

Let a and c tend to zero in the previous equation, thus

$$\Delta\text{MFE}_{12} \approx 0 \iff \eta = \frac{3(\sqrt{3}-2)}{4(1-2w)w(1-w)} \tag{3.8.35}$$

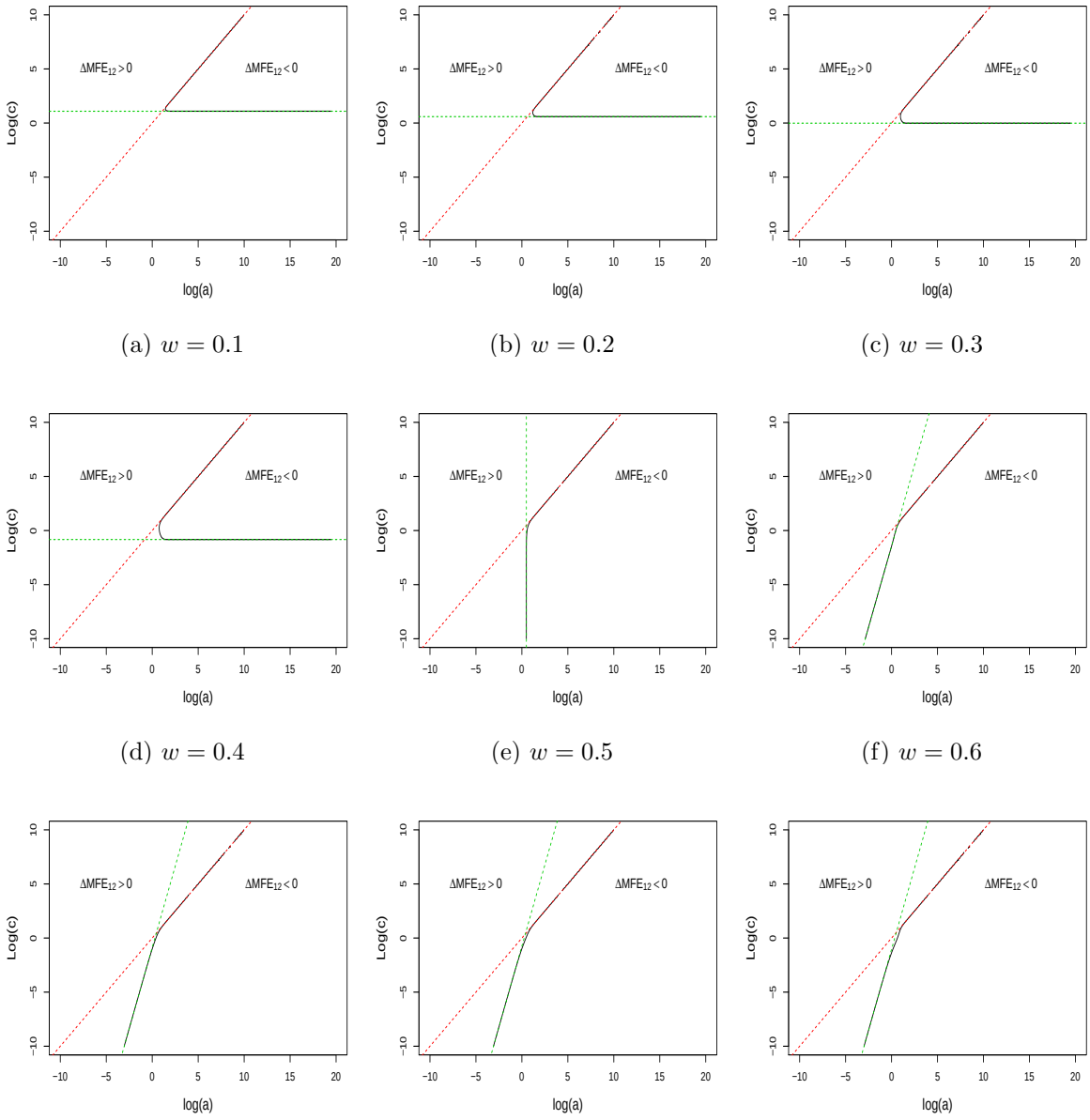
Since η is positive, w should be greater than (1-w).

In Table 3.8, we provide the lines equations that are obtained in Case A, Case B, Case C, and Case E for different values of w. The diagonal line $\log(c) = \log(a)$ is common for all values of w, when both a and c approaches ∞ (Case A).

w	Small c	a approaches ∞	Both a and c small
0.1	–	$\log(c) = 1.09$	–
0.2	–	$\log(c) = 0.59$	–
0.3	–	$\log(c) = -0.01$	–
0.4	–	$\log(c) = -0.83$	–
0.5	$\log(a) = 0.49$	–	–
0.6	–	–	$\log(a^3/c) = 1.43$
0.7	–	–	$\log(a^3/c) = 0.87$
0.8	–	–	$\log(a^3/c) = 0.74$
0.9	–	–	$\log(a^3/c) = 1.03$

Table 3.8: The zero contour equations of ΔMFE_{12} for the LMN$(w, -a, a, 1)$ distribution, and selected values of w, when c is small, a approaches ∞, and both a and c are small. The numbers in the second, third, and fourth column are rounded to two decimal places.

In order to make comparative plots between the approximations of ΔMFE_{12} in the previous cases and the real contours, the zero contour of ΔMFE_{12} for different values of w is illustrated with the lines, the equations for which are presented in Table 3.8, in Figure 3.12.



(a) $w = 0.1$ (b) $w = 0.2$ (c) $w = 0.3$

(d) $w = 0.4$ (e) $w = 0.5$ (f) $w = 0.6$

(g) $w = 0.7$ (h) $w = 0.8$ (i) $w = 0.9$

Figure 3.12: Plots of the zero contour of ΔMFE_{12} with respect to $\log(a)$ (x-axis), $\log(c)$ (y-axis) for the $\text{LMN}(w, -a, a, 1)$ distribution. Each of the nine panels is related to a specific value of w. The equations of dashed lines are presented in Table 3.8.

3.8.2 $\mathbf{PFE}_n(\alpha)$

If $X \sim \text{LMN}(w, -a, a, 1)$, $\text{PFE}_n(\alpha)$ is obtained by making $\mu_1 = -a,\ \mu_2 = a$ and $\sigma_1 = \sigma_2 = 1$ in Equation 3.7.17. In addition, for $n = 1, 2$; $\text{PFE}_1(\alpha)$ and $\text{PFE}_2(\alpha)$ respectively become

$$\text{PFE}_1(\alpha) = 1 - (1-w)\Phi(b-a) - w\Phi(b+a) \tag{3.8.36}$$

and

$$\begin{aligned}\text{PFE}_2(\alpha) = 1 - \Bigg\{&(1-w)^2\Phi\left(\frac{b-a}{\sqrt{1/2}}\right) + 2w(1-w)\Phi\left(\sqrt{2}b\right) \\ &+w^2\Phi\left(\frac{b+a}{\sqrt{1/2}}\right)\Bigg\}\end{aligned} \tag{3.8.37}$$

Recall Equation 3.7.20,

$$\Delta\text{PFE}_{12}(\alpha) = \text{PFE}_1(\alpha) - \text{PFE}_2(\alpha)$$

By using Equations 3.8.36 and 3.8.37

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= (1-w)^2\Phi\left(\frac{b-a}{\sqrt{1/2}}\right) + 2w(1-w)\Phi\left(\sqrt{2}b\right) \\ &+ w^2\Phi\left(\frac{b+a}{\sqrt{1/2}}\right) - (1-w)\Phi(b-a) - w\Phi(b+a)\end{aligned}$$

which is can be rewritten as

$$\begin{aligned}\Delta\text{PFE}_{12}(\alpha) &= w^2\left(\Phi\left(\sqrt{2}(b+a)\right) - 2\Phi\left(\sqrt{2}b\right) + \Phi\left(\sqrt{2}(b-a)\right)\right) \\ &+ w\left(2\Phi\left(\sqrt{2}b\right) - 2\Phi\left(\sqrt{2}(b-a)\right) + \Phi(b-a) - \Phi(b+a)\right) \\ &+ \Phi\left(\sqrt{2}(b-a)\right) - \Phi(b-a)\end{aligned} \tag{3.8.38}$$

With respect to w, Equation 3.8.38 represents a quadratic function with coefficients depending on b and a. Moreover, it equals zero if and only if the real root(s) of this quadratic function is(are) restricted to be in the interval $(0, 1)$.

The zero contour of $\Delta\text{PFE}_{12}(\alpha)$ is presented in Figure 3.13 for different values of w. The vertical axis represents a wide range of b, while the horizontal axis represents a. Both axes are taken on the log-scale.

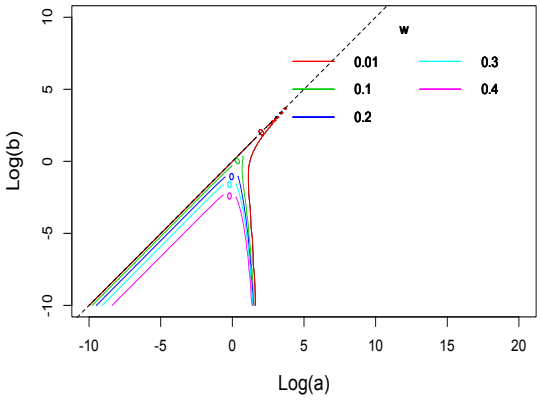
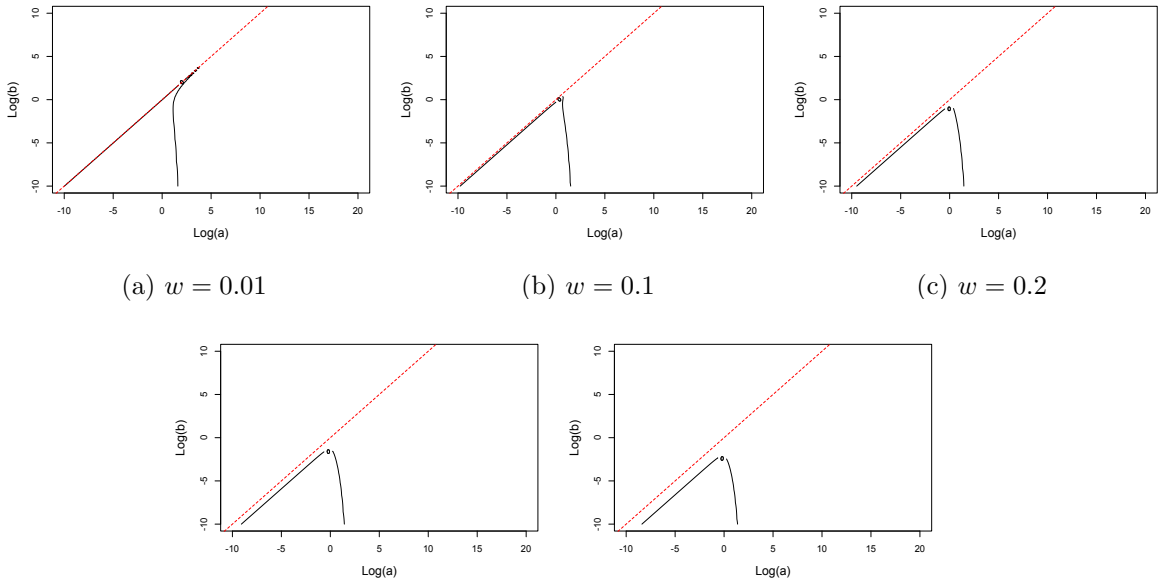


Figure 3.13: A Plot of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log(a)$ (x-axis) and $\log(b)$ (y-axis) for the $\text{LMN}(w, -a, a, 1)$ distribution, and selected values of w. The dashed line is the diagonal line $\log(b) = \log(a)$.

The contours in Figure 3.13 are plotted individually in Figure 3.14.



(a) $w = 0.01$ (b) $w = 0.1$ (c) $w = 0.2$

(d) $w = 0.3$ (e) $w = 0.4$

Figure 3.14: Plots of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log(a)$ (x-axis) and $\log(b)$ (y-axis) for the $\text{LMN}(w, -a, a, 1)$ distribution. Each of the five panels is related to a specific value of w. The red dashed line is the diagonal line $\log(b) = \log(a)$.

We do not include any value of $w \geq 0.5$ in Figures 3.13 and 3.14 because there are no zero contours if we do so, which means $\text{PFE}_n(\alpha)$ decreases as n increases

from one observation to two for all $w \geq 0.5$. The behaviour of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ for extreme parameters is summarized in Table 3.9 and discussed afterwards in detail.

Case A: Both a and b small
$\Delta\text{PFE}_{12}(\alpha) \approx 0 \iff \eta = 1/(1-2w)$
Case B: a approaches ∞
$\Delta\text{PFE}_{12}(\alpha) \to 0 \iff b = 0$

Table 3.9: Summary of the behaviour limiting of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ for the $\text{LMN}(w, -a, a, 1)$ distribution when, both a and b small, and a approaches ∞.

The detailed proof of the summary results in Table 3.9 is as follows:

Case A: Both a and b small

In this case, where both a and b are small, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ in Figures 3.13 and 3.14 forms the line that is parallel to the diagonal line $\log(b) = \log(a)$. The equation of such a line is

$$\log(b) = -\log(\eta) + \log(a)$$

Therefore, we will obtain $\eta = a/b$ as follows:

Rewrite Equations 3.8.36 and 3.8.37 in terms of erf (.). Hence,

$$\text{PFE}_1(\alpha) = 1 - \frac{1}{2}\Big\{(1-w)\text{erf}\left(\frac{b-a}{\sqrt{2}}\right) + w\ \text{erf}\left(\frac{b+a}{\sqrt{2}}\right) + 1\Big\} \tag{3.8.39}$$

and

$$\begin{aligned}\text{PFE}_2(\alpha) \quad = \quad & 1 - \frac{1}{2}\Bigg\{(1-w)^2\text{erf}\,(b-a) + 2w(1-w)\text{erf}\,(b) \\ & + w^2\text{erf}\,(b+a) + 1\Bigg\}\end{aligned} \tag{3.8.40}$$

By using the first-order Maclaurin series expansion of each error function with re-

spect to a in Equations 3.8.39 and 3.8.40, thus we obtain

$$\begin{aligned}\mathrm{PFE}_1(\alpha) &\approx 1-\frac{1}{2}\Bigg\{(1-w)\left(\mathrm{erf}\left(\frac{b}{\sqrt{2}}\right)-\sqrt{\frac{2}{\pi}}e^{-b^2/2}a\right)\\ &\qquad +w\left(\mathrm{erf}\left(\frac{b}{\sqrt{2}}\right)+\sqrt{\frac{2}{\pi}}e^{-b^2/2}a\right)+1\Bigg\}\\ &= 1-\frac{1}{2}\Bigg\{\mathrm{erf}\left(\frac{b}{\sqrt{2}}\right)+(2w-1)\sqrt{\frac{2}{\pi}}e^{-b^2/2}a+1\Bigg\} \qquad (3.8.41)\end{aligned}$$

and

$$\begin{aligned}\mathrm{PFE}_2(\alpha) &\approx 1-\frac{1}{2}\Bigg\{(1-w)^2\left(\mathrm{erf}\,(b)-\frac{2}{\sqrt{\pi}}e^{-b^2}a\right)+2w(1-w)\mathrm{erf}\,(b)\\ &\qquad +w^2\left(\mathrm{erf}\,(b)+\frac{2}{\sqrt{\pi}}e^{-b^2}a\right)+1\Bigg\}\\ &= 1-\frac{1}{2}\Bigg\{\mathrm{erf}\,(b)+\left(w^2-(1-w)^2\right)\frac{2}{\sqrt{\pi}}e^{-b^2}a+1\Bigg\} \qquad (3.8.42)\end{aligned}$$

For small b, Equations 3.8.41 and 3.8.42 respectively become

$$\mathrm{PFE}_1(\alpha)\approx 1-\left(\frac{b}{\sqrt{2\pi}}+\frac{(2w-1)}{\sqrt{2\pi}}a+\frac{1}{2}\right) \qquad (3.8.43)$$

and

$$\begin{aligned}\mathrm{PFE}_2(\alpha) &\approx 1-\left\{\frac{b}{\sqrt{\pi}}+\left(w^2-(1-w)^2\right)\frac{a}{\sqrt{\pi}}+\frac{1}{2}\right\}\\ &= 1-\left\{\frac{b}{\sqrt{\pi}}+(2w-1)\frac{a}{\sqrt{\pi}}+\frac{1}{2}\right\} \qquad (3.8.44)\end{aligned}$$

Using Equations 3.8.43 and 3.8.44. Hence,

$$\Delta\mathrm{PFE}_{12}(\alpha)\approx\frac{b\left(1-1/\sqrt{2}\right)}{\sqrt{\pi}}\left(1+(2w-1)a/b\right)$$

Define $\eta=a/b$. Thus

$$\Delta\mathrm{PFE}_{12}(\alpha)\approx\frac{b\left(1-1/\sqrt{2}\right)}{\sqrt{\pi}}\left(1+(2w-1)\eta\right)$$

Moreover,

$$\Delta\mathrm{PFE}_{12}(\alpha)\approx 0 \iff \eta=1/(1-2w)$$

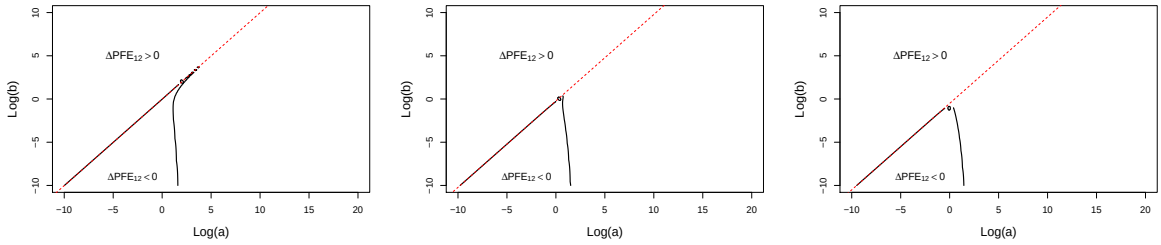
since $\eta>0$, w should be less than $(1-w)$.

Table 3.10 lists different values of $\log(a/b)$, for selected values of w.

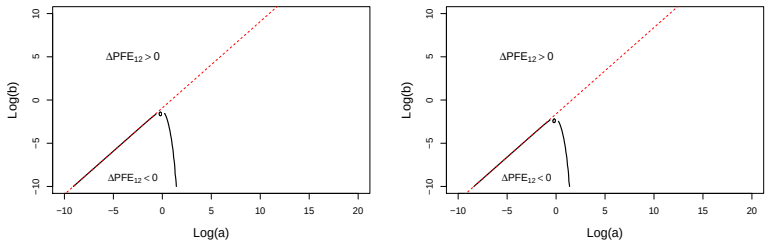
w	Both a and b small
0.01	$\log(a/b) = 0.02$
0.1	$\log(a/b) = 0.22$
0.2	$\log(a/b) = 0.51$
0.3	$\log(a/b) = 0.92$
0.4	$\log(a/b) = 1.61$

Table 3.10: The zero contours equations of $\Delta\text{PFE}_{12}(\alpha)$ for the $\text{LMN}(w, -a, a, 1)$ distribution and selected values of w when both a and b small. The numbers in the second column are rounded to two decimal places.

In order to make a comparative plot between the approximation of $\Delta\text{PFE}_{12}(\alpha)$ in Case A and the real contour, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ is illustrated in Figure 3.15 for different values of w with the line, the equation for which is presented in Table 3.10.



(a) $w = 0.01$ (b) $w = 0.1$ (c) $w = 0.2$



(d) $w = 0.3$ (e) $w = 0.4$

Figure 3.15: Plots of the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with respect to $\log(a)$ (x-axis) and $\log(b)$ (y-axis) for the $\text{LMN}(w, -a, a, 1)$ distribution. Each of the five panels is related to a specific value of w. The equation of the black dashed lines are presented in Table 3.10

.

Case B: a approaches ∞

In this case, the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ does not appear when both axes in Figure 3.13 are taken on the log-scale. Hence, it is re-presented in Figure 3.16 where both axes are taken on the original scale.

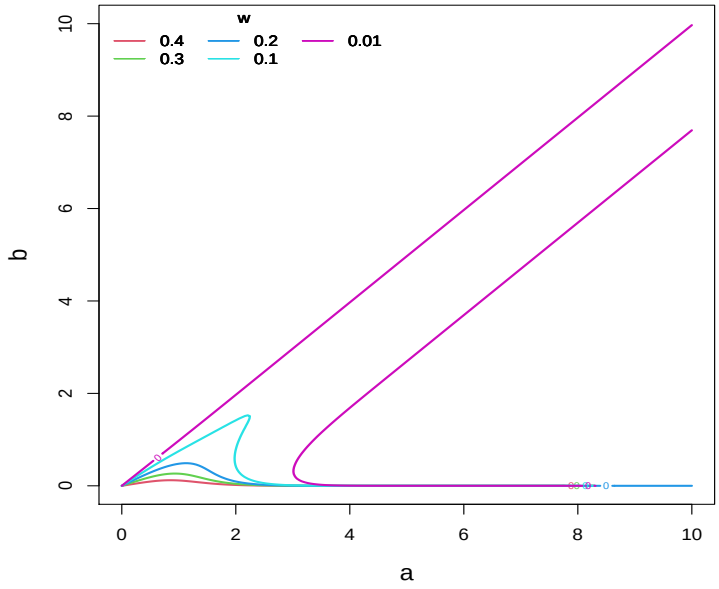


Figure 3.16: A Plot of the zero contours of ΔPFE_{12} with respect to a (x-axis), b (y-axis) for the LMN$(w, -a, a, 1)$ distribution and selected values of w.

The zero contour of $\Delta\text{PFE}_{12}(\alpha)$ when $a \to \infty$ forms the horizontal line. To obtain the equation of this line, let a tend to ∞ in Equations 3.8.39 and 3.8.40, thus

$$\begin{aligned} \text{PFE}_1(\alpha) &\to 1 - \frac{1}{2}\left(-(1-w) + w + 1\right) \\ &= 1 - w \end{aligned}$$

and

$$\begin{aligned} \text{PFE}_2(\alpha) &\to 1 - \frac{1}{2}\left(-(1-w)^2 + 2w(1-w)\text{erf}\,(b) + w^2 + 1\right) \\ &= 1 - \frac{1}{2}\left(2w + 2w(1-w)\text{erf}\,(b)\right) \end{aligned}$$

Therefore,

$$\Delta\text{PFE}_{12}(\alpha) \to w(1-w)\text{erf}\,(b)$$

and

$$\Delta\text{PFE}_{12}(\alpha) \to 0 \iff b = 0$$

Figure 3.17 illustrates the zero contour of $\Delta\text{PFE}_{12}(\alpha)$ with the line $b = 0$.

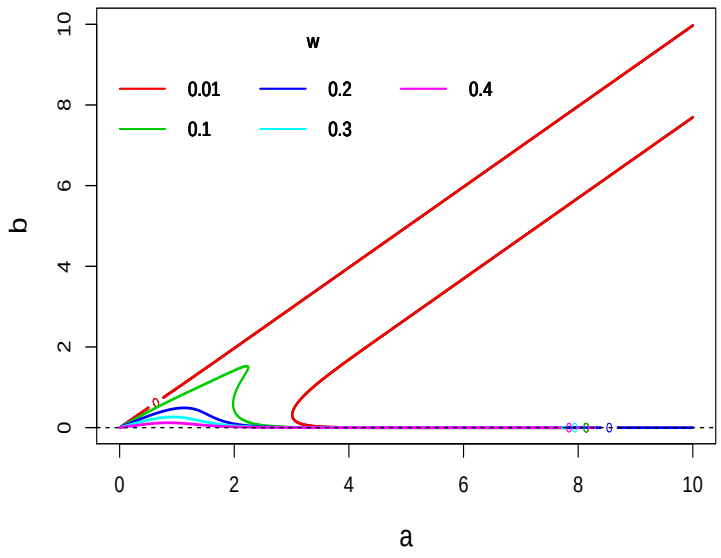


Figure 3.17: A Plot of the zero contours of $\Delta\text{PFE}_{12}(\alpha)$ with respect to a (x-axis) and b (y-axis) for the LMN$(w, -a, a, 1)$ distribution, and selected values of w. The black dashed line is $b = 0$.

3.9 Conclusions

In this chapter, we generalized the dominance properties for the arithmetic mean, known to be true for normal LSSDs, to all LSSDs with symmetric log-concave densities. Furthermore, we examined two-component scale and location mixtures of normal LSSDs in the case where the sample size increases from one to two, including a complete mathematical theory of limiting behaviour as parameters tend to extreme values. The study shows that the dominance properties do not always hold and sheds some light on factors that lead to them not holding. For the first part of the discussion we shall focus on the change from MFE_1 to MFE_2 and will briefly consider higher n at the end.

The scale and location mixture of normals are considered since some empirical LSSDs are heavy-tailed or bimodal and analytic mathematical calculations are tractable. When a LSSD is the two-component scale mixture of normal distributions, where ψ is the ratio of the variances of the two components and w is the weight given to the component with smaller variance, we see from Figure 3.2 that MFE_n decreases whenever $\psi < 30$ for all values of w and all $c > 0$. Considering the boundary situation where $\psi = 30$, Figure 3.18 compares the probability density functions of SMN(0.5, 0, 1, 30) and the normal distribution having the same variance and mean, showing the heavy-tailed nature of the scale mixture. For heavier tails, i.e. $\psi > 30$, MFE_n decreases or increases depending on the values of w and c. However, applying a sufficiently large c always leads to decreasing MFE_n. Similarly, Figure 3.6 shows that $\text{PFE}_n(\alpha)$ decreases when $\psi < 22$ for any value of w and any $b > 0$. Whereas the value of w affects the decision of choosing a value for b (and hence for c) only when $\psi \geq 22$. Figure 3.19 compares the probability density functions of the boundary case SMN(0.5, 0, 1, 22) and the normal distribution having the same variance and mean, showing the heavy-tailed nature of the scale mixture. However, a sufficiently large value of c guarantees always that $\text{PFE}_n(\alpha)$ will decrease.

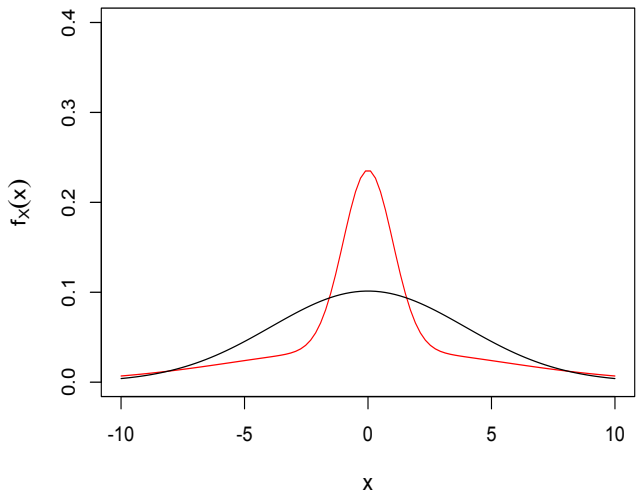


Figure 3.18: The probability density function of $\text{N}(0, 15.5)$ (black) and SMN $(0.5, 0, 1, 30)$ distributions (red), both of which have zero mean and variance 15.5.

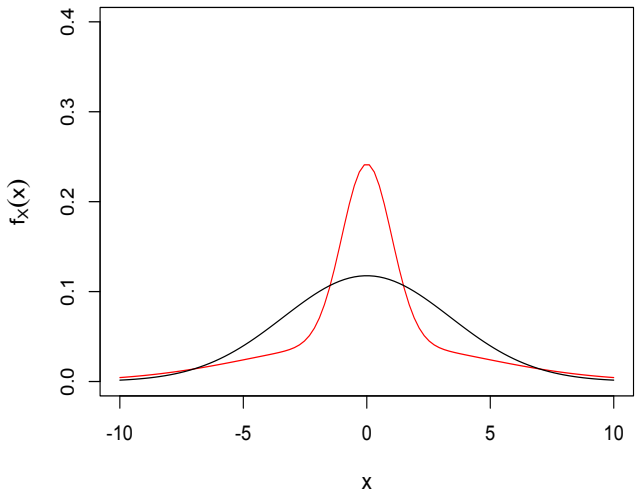
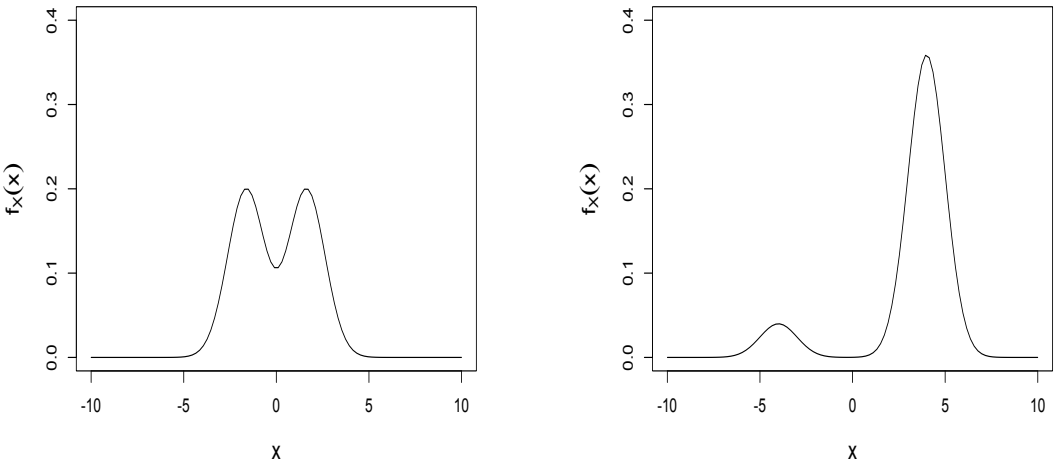


Figure 3.19: The probability density function of $N(0, 11.5)$ (black) and SMN $(0.5, 0, 1, 22)$ distributions (red), both of which have zero mean and variance 11.5.

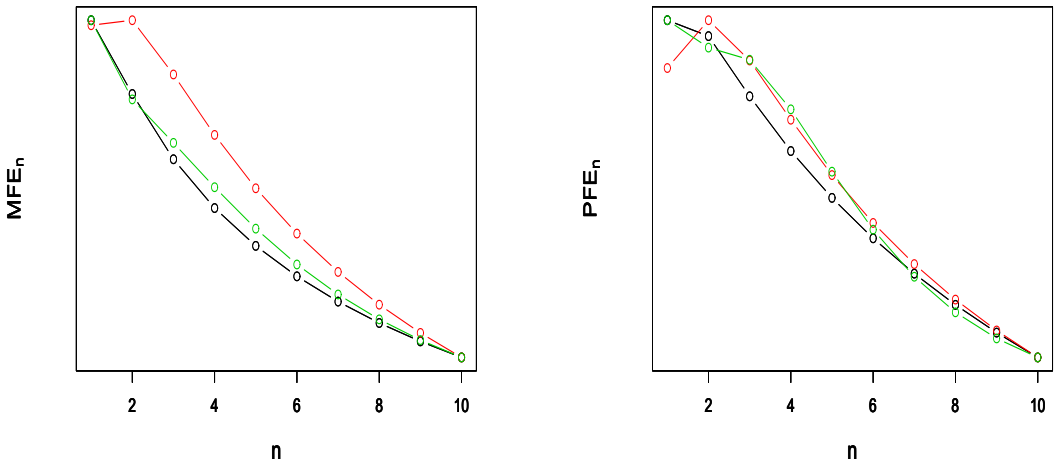
When the LSSD is a two-component location mixture of normal distributions located at $-a$ and a, both with unit variance and with weight w for the component located at $-a$, we see from Figure 3.10 that MFE_n decreases as n increases when applying small c for $a < 1.63$ and $w = 0.5$, i.e. the LSSD is symmetric. In general, choosing c to be greater than a guarantees that MFE_n decreases for all values of w. Moreover, when $w > 0.5$, i.e. the LSSD is positively skewed, MFE_n decreases if $c > ka^3$ for $a < 1$ where $k < 1$ depends on w. However, when $w < 0.5$, i.e. the LSSD is negatively skewed, MFE_n decreases for any value of c for small a and for small c for larger values of a. Regarding $\text{PFE}_n(\alpha)$, the property holds for $w \geq 0.5$ without any condition on the value of $b > 0$. Additionally, Figures 3.13 and 3.14 show that $\text{PFE}_n(\alpha)$ decreases as n increases when $w < 0.5$ and $b > a$. Taken together, these results suggest $b > a$ is sufficient to make MFE_n and $\text{PFE}_n(\alpha)$ decrease as n increases from one to two. Figure 3.20 illustrates two probability density functions of $\text{LMN}(0.5, -1.63, 1.63, 1)$ and $\text{LMN}(0.1, -4, 4, 1)$ distribution where respectively MFE_n and $\text{PFE}_n(\alpha)$ are stable when two species are tested instead of one and applying small c.



(a) LMN(0.5,-1.63,1.63,1) (b) LMN(0.1,-4,4,1)

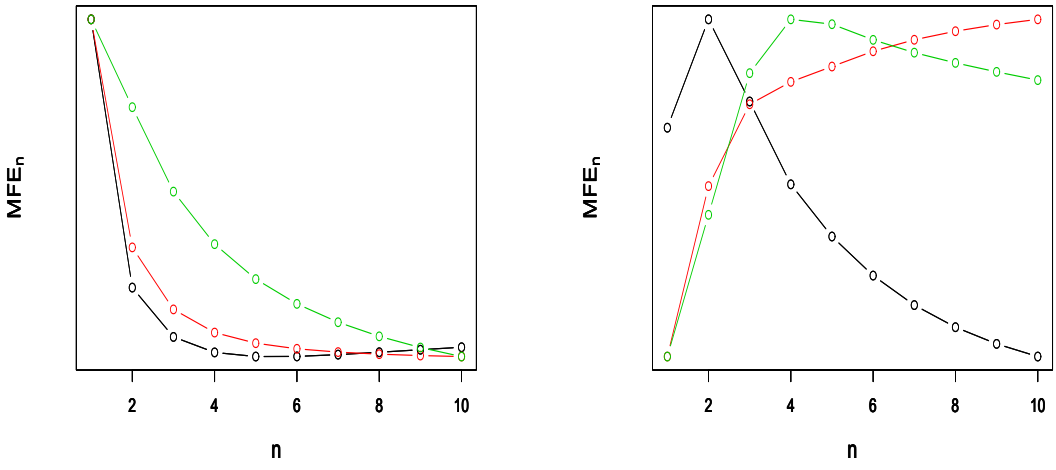
Figure 3.20: The probability density function in (a) makes $\text{MFE}_1 = \text{MFE}_2$ if small c is applied. While, the probability density function in (b) makes $\text{PFE}_1(\alpha) = \text{PFE}_2(\alpha)$ for small b.

However, assuming the LSSD is two-component scale mixture of normal distributions or two-component location mixture of normal distributions, our findings in Sections 3.7 and 3.8 are based on increasing the tested species from one to two observations. In Figures 3.21, 3.22, and 3.23, we will illustrate the behaviour of MFE_n and $\text{PFE}_n(\alpha)$ as we increase the tested species from one to ten. The values of distributions parameters, c, and b are chosen in the manner that different outcomes can be shown, and then the dominance properties for the arithmetic mean sometimes hold. Figure 3.21 shows how MFE_n and $\text{PFE}_n(\alpha)$ behave if $X \sim \text{SMN}(w, 0, 1, 43)$, Figure 3.22 (a) and (b) show how MFE_n behaves if $X \sim \text{LMN}(w, -2, 2, 1)$ and $X \sim \text{LMN}(w, -6, 6, 1)$ respectively when $w = 0.2, 0.5, \text{or } 0.9$, and Figure 3.23 shows how $\text{PFE}_n(\alpha)$ behaves if $X \sim \text{LMN}(0.1, -3, 3, 1)$; $b = 2, 3, \text{or } 4$.



(a) (b)

Figure 3.21: MFE_n in Equation 3.7.8 with $c = 2$ (a) and $\text{PFE}_n(\alpha)$ in Equation 3.7.17 with $b = 2$ (b) for $X \sim \text{SMN}(w, 0, 1, 43)$ and $w = 0.2$ (black), 0.5 (red), 0.9 (green).



(a) $\text{LMN}(w, -2, 2, 1)$ (b) $\text{LMN}(w, -6, 6, 1)$

Figure 3.22: MFE_n for two LMN distributions with $c = 2$ and $w = 0.2$ (black), 0.5 (red), 0.9 (green).

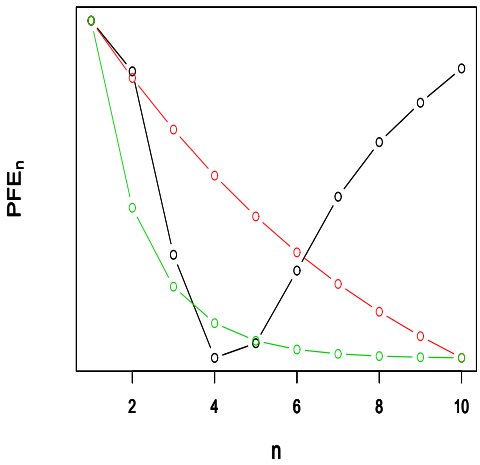


Figure 3.23: Plot of $\text{PFE}_n(\alpha)$ in Equation 3.7.17 for LMN$(0.1, -3, 3, 1)$ distribution and $b = 2$ (black), 3 (red), 4 (green).

For SMN distribution, Figure 3.21 illustrates that if MFE_n or $\text{PFE}_n(\alpha)$ decreases as n changes from 1 to 2, then it will continue decreasing. Although we have no proof of this, we have not seen any examples where this is not true.

For LMN distributions with location parameters 2 and 6, and for $c = 2$, we obtain different outcomes of MFE_n for the same value of w. Figure 3.22 shows that MFE_n is monotonically decreasing as n changes from 1 to 5 when the location parameter is 2 and $w = 0.2, 0.5$, or 0.9. Whereas when the location parameter is 6 and $w = 0.2$, MFE_n increases as n changes from 1 to 2, then it starts to decrease. In addition, when $w = 0.9$, MFE_n increases as n increases until it reaches 4 and then it starts to decrease. However, when $w = 0.5$, MFE_n is monotonically increasing. Moreover, Figure 3.23 shows that for LMN$(0.1, -3, 3, 1)$, $\text{PFE}_n(\alpha)$ is monotonically decreasing when $b = 3$ or 4. Whereas it decreases when $b = 2$ as n increases until 4, then it begins to increase. Basically, this shows that there are no safe general conclusions in the LMN distributions.

In summary, based on the selected values of c, b, and the distribution parameters, Figures 3.21, 3.22, and 3.23 show that MFE_n and $\text{PFE}_n(\alpha)$ decrease if we test five species instead of one. For other choices of c, b, and the distribution parameters leading to MFE_n or $\text{PFE}_n(\alpha)$ increasing, or both increasing, we could prevent

this outcome by applying a large value of c which is the parameter under the risk manager's control.

Returning to the application to ecotoxicological risk assessment discussed in Subsection 1.1.1 and Section 3.2, we have learned that if the density of LSSD is symmetric and log-concave, there is support for the argument in EFSA (2005) and EFSA (2008) that using the geometric mean for $n > 1$ leads to reduced risk relative to the use of a single test result, i.e. $n = 1$. However, for skewed or heavy-tailed LSSDs as exemplified by two component mixtures of normals, the situation is more complex and depends on the nature and extent of departure from univariate normality.

The two statistical risk measures, MFE_n and $\text{PFE}_n(\alpha)$, focus on the fraction of species FE_n that are affected, in the sense that their toxicity endpoint is exceeded, when they are exposed at the level of the AEC. This does match use of the criterion that the RCR=PEC/AEC should be less than one. If the PEC>AEC, the pesticide is not allowed to be used. If PEC<AEC, it may be allowed if passes the assessments of risk to other species groups and to humans. The focus in this chapter has been on dominance properties for the two risk measures as the sample size changes and has not involved specifying an acceptable value for MFE_n or a value of α and acceptable value for $\text{PFE}_n(\alpha)$. However, if specifying such acceptable values, MFE_n has the advantage that is the expected value of the FE_n and that, for using it, no further decision by risk managers is required other than to specify an acceptable value for MFE_n. Use of $\text{PFE}_n(\alpha)$ on the other hand requires risk managers to specify the value of α as well as an acceptable level of probability that FE_n exceeds α. However, there is evidence (Aldenberg and Jaworska, 2000) that $\alpha = 0.05$ is considered suitable in some situations. Then, the focus is on what size of assessment factor should be applied to obtain acceptable MFE_n or $\text{PFE}_n(\alpha)$. Hickey et al. (2009) followed a more sophisticated approach and introduced an asymmetric LINEX(linear-exponential) loss function from which to derive the AEC rather than just using the geometric mean. However, it requires sophisticated judgements from risk managers and we have not considered it here.

Chapter 4

Partial Probability Judgments in Risk Assessment

4.1 Introduction

As discussed in Subsection 1.1.3, for the second problem we are interested in conditions on the shape of distributions that lead to

$$P\left(X_1 + X_2 \geq x_1 + x_2\right) \leq \max\left(P\left(X_1 \geq x_1\right), P\left(X_2 \geq x_2\right)\right)$$

as a substitute for the worst case

$$P\left(X_1 \geq x_1\right) + P\left(X_2 \geq x_2\right)$$

where X_1 and X_2 are two independent uncertain quantities. This we call the combined tails dominance property. In this context, we expect that $P(X_1 \geq x_1)$ and $P\left(X_2 \geq x_2\right)$ are given as imprecise probabilities which represent partial probability judgements.

The convenience of the combined tail dominance property is particularly clear when more quantities are involved. For example, if $P(X_i \geq x_i) = 0.1$ for $i = 1, 2, \ldots, 10$, then the worst case outcome is $P(X_1 + \cdots + X_{10} \geq x_1 + \cdots + x_{10}) \leq 0.1 + \cdots + 0.1 = 1$ which provides no information, whereas if the combined tails dominance property holds, the bound is 0.1 not 1. In what follows, we will explore situations where we can show that this property does always hold, and others where

we can see that it may not hold.

In the rest of this chapter, an illustrative example of application is presented in Section 4.2. The seminal result of Proschan (1965) in Section 2.3 is exploited to explore the validation of the combined tails dominance property for distributions with symmetric and log-concave density functions in Section 4.4. The normal and Cauchy distributions are considered in Section 4.5 as illustrative examples. The peakedness comparison for convex combinations of independent and identically distributed two random variables, $X_i;\ i = 1, 2$, from skew-symmetric distributions with log-concave kernels is discussed in Section 4.6. The reason for this discussion is to achieve our goal in discovering the validation of the combined tails dominance property for such distributions when two uncertain quantities are involved. Two important distributions that will be used later in this chapter are the shifted exponential (SHE) and the mixture of shifted exponential and lump (SHEL). These two distributions are presented in Section 4.7. The stochastic dominance property with some results needed for the rest of this chapter are introduced in Section 4.8. In Section 4.9, we focus on the log-concavity of the right tail probability function which may be fully log-concave or partially log-concave. We present an important theorem that is used later in Sections 4.10 and 4.11. A novel approach is presented in Sections 4.10 and 4.11 to explore the validation of the combined tails dominance property for distributions with log-concave right tail probability and distributions with partial log-concave right tail probability respectively. Section 4.13 presents some conclusions. A glossary of acronyms, including those used in this chapter, is presented at the end of the thesis.

4.2 Example of Application by EFSA

Rift Valley Fever Virus (RVFV) is a virus that is transmitted by mosquitoes and affects buffalo, camels, cattle, goats, and sheep. It causes severe illness and could transfer to people. The endemic areas are West and East Africa. EFSA (2013) discussed the uncertainty about the number of infected animals that are exported in a single year from West Africa to specific countries of North Africa and the Near

East. The model for the number of infected animals exported to North Africa from West Africa is

$$N_W = v_W \times p_W \times t_W$$

where these uncertain quantities, as presented in EFSA (2013), are:

- N_W is the quantity of interest which represents the number of infected animals arriving in North Africa from the West source in one year.
- v_W is the volume of trade from the West source: the number of animals transported in one year to North Africa from the West source.
- p_W is the prevalence of RVFV in animals in the West source: approximates the proportion of transported animals which are infected.
- t_W is the proportion of infected animals which remain infected after transport.

Uncertainties about the quantities v_W, p_W and t_W were judged to be independent, and their distributions were elicited from experts, using the SHELF protocol in an EKE process, to be:

- The distribution of v_W is scaled beta (3.77, 5.99) on range 10 to 700000.
  The beta distribution is used to model continuous random variables $X \in [0, 1]$. Its probability density function is

  $$f_X(x) = \frac{x^{p-1}(1-x)^{q-1}}{\text{Beta(p,q)}};\ 0 \leq x \leq 1;\ p > 0,\ q > 0$$

  where $\text{Beta}(\alpha, \beta)$ is the beta function and its formula is

  $$\text{Beta}(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

  However, X can be scaled and shifted to obtain a scaled beta random variable on $[a, b]$ by using the transformation $Y = a+(b-a)X$. The probability density function of the scaled beta distribution on $[a, b]$ is

  $$f_Y(x) = \frac{(x-a)^{p-1}(b-x)^{q-1}}{\text{Beta}(p,q)(b-a)^{p+q-1}};\ a \leq x \leq b,\ p > 0,\ q > 0$$

  Hence the probability density function of v_W is

  $$f_{v_W}(x) = \frac{(x-10)^{2.77}(700000-x)^{4.99}}{\text{Beta}(3.77, 5.99)(700000-10)^{3.77+5.99-1}};\ 10 \leq x \leq 700000$$

- The distribution of p_W is log-normal with parameters $\mu = -6.7$ and $\sigma = 1.11$. A continuous random variable X has a log-normal distribution $X \sim \text{LN}(\mu, \sigma)$ if its probability density function is

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right);\ 0 < x < \infty,\ \mu \in \mathbb{R}, \sigma > 0$$

  If $X \sim \text{LN}(\mu, \sigma)$, then $Y = \log X \sim \text{N}(\mu, \sigma)$, the normal distribution with mean μ and standard deviation σ. Hence the probability density function of p_W is

$$f_{p_W}(x) = \frac{1}{(1.11)x\sqrt{2\pi}} \exp\left(-\frac{(\ln x + 6.7)^2}{2(1.11)^2}\right);\ 0 < x < \infty$$

- In addition, $t_W \sim \text{LN}(-1.5, 0.758)$. Its probability density function is

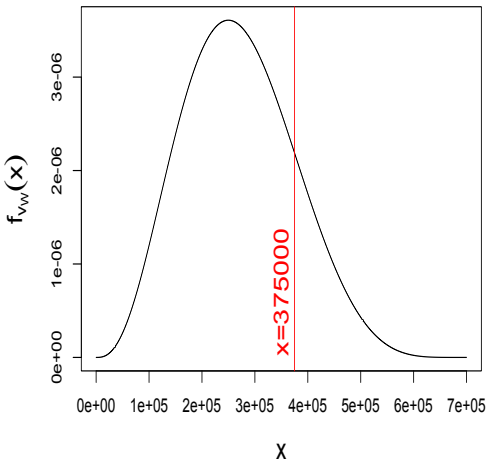
$$f_{t_W}(x) = \frac{1}{(0.758)x\sqrt{2\pi}} \exp\left(-\frac{(\ln x + 1.5)^2}{2(0.758)^2}\right);\ 0 < x < \infty$$

The three probability density functions are shown in Figure 4.1. Following the approach in (EFSA, 2013), we would then compute the distribution of N_W. This could easily be done using Monte Carlo or by numerical calculation based on characteristic functions, noting that the distribution of $\log N_W = \log v_W + \log p_W + \log t_W$ is the convolution of the distributions of $\log v_W$, $\log p_W$ and $\log t_W$. The result would be a full probability distribution for N_W based on the three full probability distributions for the individual quantities. However, we should note that the experts did not directly provide these three distributions. Rather, they were derived by the EKE facilitator from the partial probability judgements made by the experts and presented to the experts for them to approve. Therefore, it is interesting to explore alternative approaches based directly on judgements made by experts. One such approach is probability bounds analysis, as recommended in EFSA (2018a).

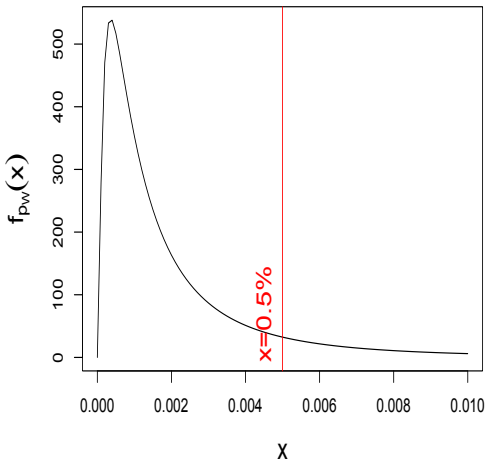
We do not have probability bounds provided by the experts for this example. Instead, we will work with illustrative probability bounds for v_W, p_W and t_W shown in red in Figure 4.1: $P(v_w \geq 375000) \leq 0.1$, $P(p_W \geq 0.5\%) \leq 0.1$ and $P(t_W \geq 60\%) \leq 0.1$. These illustrative values are consistent with the distributions obtained by the EKE. We then proceed with the probability bounds analysis. First we obtain the threshold $375000 \times 0.5\% \times 60\% = 1125$ and then calculate the upper bound probability of $P(v_W \times p_W \times t_W \geq 1125)$ to be $0.1 + 0.1 + 0.1 = 0.3$. The result of the

probability bounds analysis is a probability bound for N_W: $P(N_W \geq 1125) \leq 0.3$.

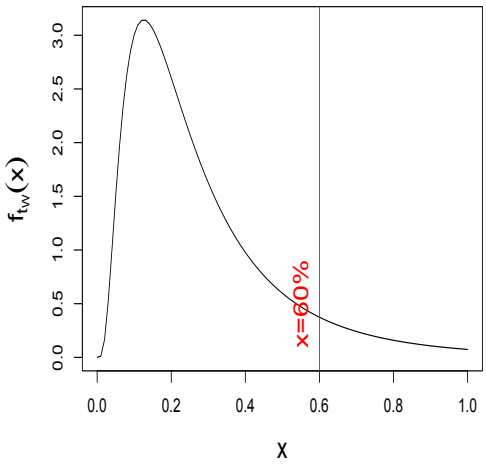
If the combined tails dominance property were to apply, the benefit would be that the probability bound would be tighter, namely $P(N_W \geq 1125) \leq 0.1$. This would follow from the facts that $\log N_W = \log v_W + \log p_W + \log t_W$ and that the logarithm function is strictly monotonic increasing so that the individual probability bounds become $P(\log v_w \geq \log 375000) \leq 0.1$, $P(\log p_W \geq \log 0.5\%) \leq 0.1$ and $P(t_W \geq \log 60\%) \leq 0.1$. The combined tails dominance property would then yield $P(\log v_W + \log p_W + \log t_W \geq \log 375000 + \log 0.5\% + \log 60\%) \leq \max(0.1, 0.1, 0.1) = 0.1$ and inverting the logarithm would give $P(N_W > 1125) \leq 0.1$. An interesting question is whether or not it is reasonable to assume the combined tails dominance property in this example. In Section 4.10, we shall see that the property holds for random variables having log-concave right tail probability functions if the individual probabilities in the probability bounds are less than 0.28 as they are here. Moreover, $\log X$ has log-concave right tail probability function if X has log-concave right tail probability function (Bagnoli and Bergstrom, 2005) and X has log-concave right tail probability function if it has log-concave probability density function (ibid, 2005). All normal distributions have log-concave densities as do beta distributions where the shape parameters exceed 1. Therefore the distributions from the EKE do all have log-concave right tail probability functions. This is of course not the same as experts making direct judgements about log-concavity of the right tail probability function and this issue is considered briefly in Section 4.9.



(a) $v_W \sim$ Scaled beta$(3.77, 5.99)$



(b) $p_W \sim$ LN$(-6.7, 1.11)$



(c) $t_W \sim$ LN$(-1.5, 0.758)$

Figure 4.1: Probability density functions of uncertain quantities (a) v_W , (b) p_W, and (c) t_W. The vertical red lines show the thresholds for the hypothetical probability bounds used to illustrate the application of probability bounds analysis and the potential benefit of the combined tail dominance property.

4.3 A Simplification

The following theorem shows that we need only look in the rest of the chapter at the case where the tail probabilities for the individual quantities are equal.

Theorem 4.3.1 Let the distribution functions of independent and continuous random variables $X_i;\ i = 1, 2, \ldots, n$ be fixed. Moreover, let p_i and x_i be such that $p_i = P(X_i \geq x_i)$. In addition, suppose that when $p_1 = p_2 = \cdots = p_n = p$,

$$P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) \leq p \tag{4.3.1}$$

Consequently, whenever $\max(p_1, \ldots, p_n) = p$,

$$P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) \leq p$$

Proof. Define x_i' so that $P(X_i \geq x_i') = p$. Then Equation 4.3.1 is equivalent to

$$P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i'\right) \leq p \tag{4.3.2}$$

If $\max(p_1, \ldots, p_n) = p$, each $p_i \leq p$ and so

$$x_i \geq x_i'$$

Using Equation 4.3.2, we find that

$$\begin{aligned}\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) &\leq P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i'\right) \\ &\leq p\end{aligned}$$

The inequality is strict when $\min(p_1, \ldots, p_n) < p$. □

4.4 Consequences of Proschan's Result

In Theorem 4.4.1 we exploit the result by Proschan (1965), which is reintroduced in Theorem 2.3.5, to show that the combined tail dominance property holds for n independent and identically random variables having symmetric log-concave densities.

Theorem 4.4.1 Let $X_i;\ i = 1, 2, \ldots, n$ be independent and identically random variables having symmetric log-concave densities. Moreover, let x_i be such that $P(X_i \geq x_i) = p$. Then

$$P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) < p$$

Proof. Set $a = (1, 0, 0, \ldots, 0)$ and $b = (1/n, 1/n, \ldots, 1/n)$ in Theorem 2.3.5. Thus

$$P\left(\frac{\sum_{i=1}^{n} X_i}{n} \geq t\right) < P(X_n \geq t);\ t \geq 0$$

Let $t = \frac{\sum_{i=1}^{n} x_i}{n}$ in the previous inequality, hence

$$P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) < P\left(X_n \geq \frac{\sum_{i=1}^{n} x_i}{n}\right)$$

Since X_i are iid, the x_i are equal for all i. Consequently,

$$\frac{\sum_{i=1}^{n} x_i}{n} = x_i$$

Therefore,

$$\begin{aligned} P\left(\sum_{i=1}^{n} X_i \geq \sum_{i=1}^{n} x_i\right) &< P(X_n \geq x_n) \\ &= p \end{aligned}$$

Hence, theorem is proved. □

Moreover, by using Theorem 2.3.5, Theorem 4.4.2 shows that the combined tail dominance property is stable under re-scaling the distributions in Theorem 4.4.1.

Theorem 4.4.2 Let $X_i;\ i = 1, 2, \ldots, n$ be independent and identically random variables having symmetric log-concave densities. Moreover, let x_i and $l_i > 0$ be such that $P(l_i X_i \geq x_i) = p$. Then

$$P\left(\sum_{i=1}^{n} l_i X_i \geq \sum_{i=1}^{n} x_i\right) < p$$

Proof. Set $a = (1, 0, 0, \ldots, 0)$ and $b = (l_1/\sum_{i=1}^{n} l_i, l_2/\sum_{i=1}^{n} l_i, \ldots, l_n/\sum_{i=1}^{n} l_i)$ in Theorem 2.3.5. This yields

$$P\left(\sum_{i=1}^{n} l_i X_i \geq t \sum_{i=1}^{n} l_i\right) < P(X_n \geq t);\ t \geq 0$$

Set $t = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} l_i}$ in the last inequality, thus

$$\begin{aligned} P\left(\sum_{i=1}^{n} l_i X_i \geq \sum_{i=1}^{n} x_i\right) &< P\left(X_n \geq \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} l_i}\right) \\ &= P\left(l_n X_n \geq \frac{\sum_{i=1}^{n} l_n x_i}{\sum_{i=1}^{n} l_i}\right) \end{aligned} \tag{4.4.1}$$

Since $P\left(l_i X_i \geq x_i\right) = p$, then we can assume that $\frac{x_i}{l_i} = x$. Moreover,

$$\begin{aligned} P\left(l_n X_n \geq \frac{\sum_{i=1}^{n} l_n x_i}{\sum_{i=1}^{n} l_i}\right) &= P\left(l_n X_n \geq \frac{\sum_{i=1}^{n} l_n l_i x}{\sum_{i=1}^{n} l_i}\right) \\ &= P\left(l_n X_n \geq l_n x\right) \\ &= P\left(l_n X_n \geq x_n\right) \\ &= p \end{aligned}$$

Therefore, Equation 4.4.1 becomes

$$P\left(\sum_{i=1}^{n} l_i X_i \geq \sum_{i=1}^{n} x_i\right) < p$$

□

4.5 Examples

In this section, we show directly that the combined tail dominance property holds for normal and Cauchy distributions without using Proschan's result. The normal distribution is the first choice due to its popularity and mathematical simplicity. The Cauchy distribution is of interest since if X_1 and X_2 are independent and identically distributed random variables, each with a standard Cauchy distribution, then the sample mean $(X_1 + X_2)/2$ has the same standard Cauchy distribution. The combined tail dominance property holds exactly for Cauchy which is symmetric but does not have log-concave density function. This result coincides with findings in Ibragimov (2007), who proved that the seminal result of Proschan (1965) extends to hold for symmetric distributions with not extremely heavy-tailed densities and does not hold for distributions with extremely heavy-tailed densities. In addition the Cauchy distribution is on the boundary between these two cases.

Example 4.5.1 Let $X_1 \sim N(0, \sigma_1^2)$ and $X_2 \sim N(0, \sigma_2^2)$ be independent. Moreover, let σ_1^2, σ_2^2, and $0 < \rho \leq 1$ be such that $\sigma_1^2 = \rho\sigma_2^2$. In addition, let p_i be such that $p_i = P(X_i \geq x_i) = p$ for $i = 1$ and 2. Then

$$P\left(X_1 + X_2 \geq x_1 + x_2\right) < p$$

Proof. The distribution of X_2 is used to find p_1, p_2, and $P\big(X_1 + X_2 \geq x_1 + x_2\big)$ as follows:

$$\begin{aligned} p_1 &= P\left(X_2 \geq \frac{x_1}{\sqrt{\rho}}\right) \\ p_2 &= P\left(X_2 \geq x_2\right) \\ P(X_1 + X_2 \geq x_1 + x_2) &= P\left(Z \geq \frac{x_1 + x_2}{\sqrt{1+\rho}\sigma_2}\right) \\ &= P\left(X_2 \geq \frac{x_1 + x_2}{\sqrt{1+\rho}}\right) \end{aligned} \tag{4.5.1}$$

(1) Assume $\rho = 1$. If $p_1 = p_2 = p$, then $x_1 = x_2$, and Equation 4.5.1 becomes

$$\begin{aligned} P(X_1 + X_2 \geq x_1 + x_2) &= P\left(X_2 \geq \sqrt{2}x_2\right) \\ &< P\left(X_2 \geq x_2\right) \\ &= p \end{aligned}$$

(2) Assume $0 < \rho < 1$. If $p_1 = p_2$, then $\frac{x_1}{\sqrt{\rho}} = x_2$ and Equation 4.5.1 becomes

$$P(X_1 + X_2 \geq x_1 + x_2) = P\left(X_2 \geq \frac{\sqrt{\rho}+1}{\sqrt{1+\rho}}x_2\right) \tag{4.5.2}$$

Since $(\sqrt{\rho}+1)^2 > \rho + 1$, thus $\sqrt{\rho}+1 > \sqrt{\rho+1}$. Therefore, Equation 4.5.2 becomes

$$\begin{aligned} P(X_1 + X_2 \geq x_1 + x_2) &< P\left(X_2 \geq x_2\right) \\ &= p \end{aligned}$$

□

Example 4.5.2 Let $X_i;\ i = 1, 2$ be independent and have a Cauchy distribution with probability density function

$$f_{X_i}(x; x_0, \sigma_i) = \frac{1}{\pi\sigma_i}\left(1 + \left(\frac{x - x_0}{\sigma_i}\right)^2\right) + \frac{1}{2};\ x \in \mathbb{R}$$

where $x_0 \in \mathbb{R}; \sigma_i > 0$ are the location and scale parameter respectively. In addition, let p_i be such that $p_i = P(X_i \geq x_i)$. If $p_1 = p_2 = p$, then $P\big(X_1 + X_2 \geq x_1 + x_2\big) = p$.

Proof. For no loss of generality, assume $x_0 = 0$. The cumulative distribution function of $X_i \sim \text{Cauchy}(0, \sigma_i);\ i = 1, 2$ is

$$F_{X_i}(x) = \frac{1}{\pi}\text{arctan}\left(\frac{x}{\sigma_i}\right) + \frac{1}{2};\ x \in \mathbb{R}$$

where 0 and $\sigma > 0$ are the location and scale parameter respectively. Thus the sum of two independent Cauchy random variables is again a Cauchy, with the scale parameters summed.

Let $p_1 = p_2 = p$, then $x_1 = \frac{\sigma_1}{\sigma_2}x_2$, and

$$\begin{aligned} P\big(X_1 + X_2 \geq x_1 + x_2\big) &= \frac{1}{2} - \frac{1}{\pi}\text{arctan}\left(\frac{\big(\frac{\sigma_1}{\sigma_2} + 1\big)x_2}{\sigma_1 + \sigma_2}\right) \\ &= \frac{1}{2} - \frac{1}{\pi}\text{arctan}\left(\frac{x_2}{\sigma_2}\right) \\ &= p \end{aligned}$$

This means that Cauchy distribution is on the boundary between the situation where the combined tail property holds and situation where it does not. □

4.6 Skew-Symmetric Distributions with Log-Concave Kernels

Azzalini and Regoli (2012) define the probability density function of a skew-symmetric distribution as given in Lemma 4.6.1, which is presented originally in Azzalini and Capitanio (2003).

Lemma 4.6.1 (Azzalini and Regoli, 2012, Lemma 1). Denote a D-dimensional probability density function centrally symmetric about zero by f_X^*, a continuous distribution function on the real line by $G*_X$, with symmetric density function $g*_X$, so that $g*_X$ exists almost everywhere and is an even function, and an odd function by $\gamma : \mathbb{R}^D \to \mathbb{R}$, so that $\gamma(-x) = -\gamma(x)\ \forall x \in \mathbb{R}^D$. Thus

$$f_X(x) = 2f_X^*(x)G_X^*(\gamma(x));\ x \in \mathbb{R}^D \tag{4.6.1}$$

is a density function. The perturbation function $G_X(x) = G_X^*\big(\gamma(x)\big)$ satisfies

$$G_X(x) \geq 0, \;\; G_X(x) + G_X(-x) = 1; \; x \in \mathbb{R}^D$$

In Lemma 4.6.2, we present the behaviour of the right tail probability of the convex combination of two iid random variables with density function in Equation 4.6.1 under specific conditions. Subsequently, the obtained result will be used to prove Theorem 4.6.3 which states that the combined tail dominance property holds for distributions having such densities when two components are involved.

Lemma 4.6.2 Let $f_X(x) = 2f_X^*(x)G_X^*(\alpha x); \; x \in \mathbb{R}, \; \alpha < 0$, where f_X^* and G_X^* are defined in Lemma 4.6.1 such that f_X^* is log-concave and $D = 1$. Moreover, let X_1 and X_2 be independently distributed with f_X. Then $P\big(aX_1 + (1-a)X_2 \leq t\big)$ is strictly increasing as a increases from 0 to $\frac{1}{2}$

Proof. Fix $t > 0$ and define

$$H(a,t) = P\left(aX_1 + (1-a)X_2 \leq t\right) = \int_{-\infty}^{\infty} F_X\left(\frac{t-(1-a)u}{a}\right) f_X(u)\, du; \;\; 0 < a < \frac{1}{2}$$

Follow the same steps in Lemma 2.3.3 until reaching Equation 2.3.16. Hence,

$$\begin{aligned}
\frac{a^2 \partial H(a,t)}{\partial a} &= \int_0^{\infty} v \Bigg\{ f_X\left(t+v\right) f_X\left(t - \frac{(1-a)}{a} v\right) \\
&\qquad\qquad - f_X\left(t-v\right) f_X\left(t + \frac{(1-a)}{a} v\right) \Bigg\} dv \\
&= \int_0^{\infty} v J(a,v)\, dv
\end{aligned}$$

Since $f_X(x) = 2f_X^*(x)G_X^*(\alpha x)$,

$$\begin{aligned}
J(a,v) &= 4 f_X^*\left(t+v\right) G_X^*\left(\alpha(t+v)\right) f_X^*\left(t - \frac{(1-a)}{a} v\right) G_X^*\left(\alpha\left(t - \frac{(1-a)}{a} v\right)\right) \\
&- 4 f_X^*\left(t-v\right) G_X^*\left(\alpha(t-v)\right) f_X^*\left(t + \frac{(1-a)}{a} v\right) G_X^*\left(\alpha\left(t + \frac{(1-a)}{a} v\right)\right)
\end{aligned}$$

Since f_X^* is symmetric around zero, thus we get

$$\begin{aligned}
J(a,v) &= 4 f_X^*\left(t+v\right) f_X^*\left(\frac{(1-a)}{a} v - t\right) G_X^*\left(\alpha(t+v)\right) G_X^*\left(\alpha\left(t - \frac{(1-a)}{a} v\right)\right) \\
&- 4 f_X^*\left(v-t\right) f_X^*\left(\frac{(1-a)}{a} v + t\right) G_X^*\left(\alpha\left(t-v\right)\right) G_X^*\left(\alpha\left(t + \frac{(1-a)}{a} v\right)\right)
\end{aligned}$$

Furthermore, f_X^* is a log-concave function. Thus for $v \geq 0;\ t > 0;\ (1-a)/a > 1$,

$$\frac{f_X^*\big(v-t\big)}{f_X^*\big(v+t\big)} \leq \frac{f_X^*\big(\frac{(1-a)}{a}v-t\big)}{f_X^*\big(\frac{(1-a)}{a}v+t\big)} \tag{4.6.2}$$

Moreover,

$G_X^*\left(\alpha\left(t-\frac{(1-a)}{a}v\right)\right) > G_X^*\left(\alpha(t-v)\right)$ and $G_X^*\left(\alpha\left(t+\frac{(1-a)}{a}v\right)\right) < G_X^*\left(\alpha\left(t+v\right)\right)$

Hence,

$$G_X^*\left(\alpha(t-v)\right)G_X^*\left(\alpha\left(t+\frac{(1-a)}{a}v\right)\right) < G_X^*\left(\alpha(t+v)\right)G_X^*\left(\alpha\left(t-\frac{(1-a)}{a}v\right)\right) \tag{4.6.3}$$

By using Equations 4.6.2 and 4.6.3, we get

$$f_X^*\left(v-t\right)f_X^*\left(\tfrac{(1-a)}{a}v+t\right)G_X^*\left(\alpha(t-v)\right)G_X^*\left(\alpha\left(t+\tfrac{(1-a)}{a}v\right)\right)$$
$$<$$
$$f_X^*\left(v+t\right)f_X^*\left(\tfrac{(1-a)}{a}v-t\right)G_X^*\left(\alpha(t+v)\right)G_X^*\left(\alpha\left(t-\tfrac{(1-a)}{a}v\right)\right)$$

Hence, $J(a,v) > 0$, which implies $\frac{\partial H(a,t)}{\partial a} > 0$ as a consequence. □

Remark. when f_X^* is the probability density of the standard normal distribution, and G_X^* is the cumulative distribution function of the standard normal distribution, this is the case of having negatively skewed-normal distribution (Azzalini, 1985).

Theorem 4.6.3 Let X_1 and X_2 be independent and identically distributed with

$$f_X(x) = 2f_X^*(x)G_X^*(\alpha x);\ x \in \mathbb{R},\ \alpha < 0$$

where f_X^* and G_X^* are defined in Lemma 4.6.1 such that f_X^* is log-concave and $D = 1$. Moreover, let x_i be such that $p_i = P\big(X_i \geq x_i\big) = p$. Then

$$P\big(X_1 + X_2 \geq x_1 + x_2\big) < p$$

Proof. Using Lemma 4.6.2, for $a < b, a < (1-a), b < (1-b)$, then

$$P(aX_1 + (1-a)X_2 \leq t) < P(bX_1 + (1-b)X_2 \leq t)$$

which is equivalent to

$$P(bX_1 + (1-b)X_2 \geq t) < P(aX_1 + (1-a)X_2 \geq t) \tag{4.6.4}$$

Let $a = (1, 0)$ and $b = (1/2, 1/2)$ in Equation 4.6.4. Thus

$$P\big(X_1 + X_2 \geq 2t\big) < P\big(X_2 \geq t\big);\ t \geq 0$$

Let $t = \frac{x_1+x_2}{2}$ in the previous inequality, hence

$$P\big(X_1 + X_2 \geq x_1 + x_2\big) \ < \ P\big(X_2 \geq \frac{x_1 + x_2}{2}\big)$$

Since $X_i;\ i = 1, 2$ are iid, $x_1 = x_2$. Therefore,

$$\begin{aligned} P\big(X_1 + X_2 \geq x_1 + x_2\big) \ &< \ P\big(X_2 \geq x_2\big) \\ &= \ p \end{aligned}$$

□

Corollary 4.6.4 shows that the combined tail dominance property is stable under re-scaling either of the distributions in Theorem 4.6.3.

Corollary 4.6.4 Let X_1 and X_2 be independent and identically distributed with

$$f_X(x) = 2f_X^*(x)G_X^*(\alpha x);\ x \in \mathbb{R},\ \alpha < 0$$

where f_X^* and G_X^* are defined in Lemma 4.6.1 such that f_X^* is log-concave and $D = 1$. Moreover, let x_1, x_2 and $k > 0$ be such that $P(kX_1 \geq x_1) = P(X_2 \geq x_2) = p$. Then

$$P\big(kX_1 + X_2 \geq x_1 + x_2\big) < p$$

Proof. Set $a = (0, 1)$ and $b = (k/(k + 1), 1/(k + 1))$ in Equation 4.6.4. Thus

$$P(kX_1 + X_2 \geq (k + 1)t) < P(X_2 \geq t)$$

Set $t = \frac{x_1+x_2}{k+1}$ in the previous inequality,

$$P(kX_1 + X_2 \geq x_1 + x_2) \ < \ P(X_2 \geq \frac{x_1 + x_2}{k + 1}) \tag{4.6.5}$$

Since $P(kX_1 \geq x_1) = P(X_2 \geq x_2)$, thus $x_1 = kx_2$ and Equation 4.6.5 becomes

$$\begin{aligned} P(kX_1 + X_2 \geq x_1 + x_2) \ &< \ P(X_2 \geq x_2) \\ &= \ p \end{aligned}$$

□

4.7 SHE and SHEL Distribution Functions

The distribution of the shifted exponential distribution (SHE) is the distribution of X where $X-c$ is exponentially distributed for some real c. Therefore, the probability density function, right tail probability function, and characteristic function of $X \sim \text{SHE}(c, \lambda)$ are given respectively by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda(x-c)} & , \ x \geq c \\ 0 & , \ x < c \end{cases} \tag{4.7.1}$$

$$\bar{F}_X(x) = \begin{cases} e^{-\lambda(x-c)} & , \ x \geq c \\ 1 & , \ x < c \end{cases} \tag{4.7.2}$$

$$\psi_X(t) = e^{it(c)}(1 - it\lambda^{-1})^{-1} \tag{4.7.3}$$

The discrete random variable Y has lump distribution $L(c)$ if its probability mass function, right tail probability function and characteristic function are given respectively by

$$f_Y(x) = \begin{cases} 1 & , \ x = c \\ 0 & , \ x \neq c \end{cases}$$

$$\bar{F}_Y(x) = \begin{cases} 0 & , \ x \geq c \\ 1 & , \ x < c \end{cases} \tag{4.7.4}$$

$$\psi_Y(t) = e^{itc} \tag{4.7.5}$$

The variable Z has a mixture distribution of shifted exponential and lump (SHEL) with parameters $0 < w \leq 1, c \in \mathbb{R}, \ \lambda > 0$, simply represented by $Z \sim \text{SHEL}(w, c, \lambda)$, if it is distributed like X for $x > c$ with probability w, but takes on the value c with probability $1 - w$. Using Equations 4.7.3 and 4.7.5, the characteristic function of Z is

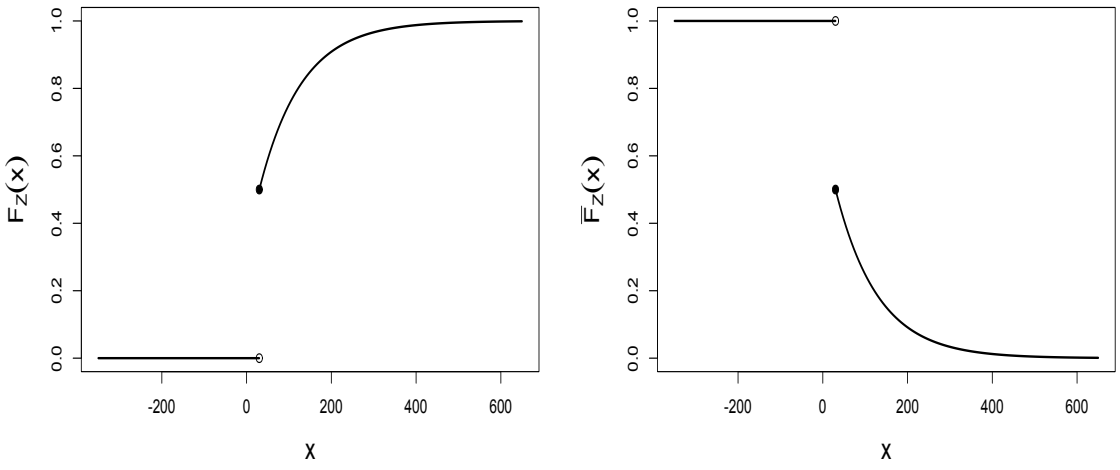
$$\begin{aligned} \psi_Z(t) &= w\psi_X(t) + (1-w)\psi_Y(t) \\ &= we^{it(c)}(1 - it\lambda^{-1})^{-1} + (1-w)e^{it(c)} \end{aligned} \tag{4.7.6}$$

The cumulative distribution function and right tail probability function of Z are respectively as follows:

$$\begin{aligned} F_Z(x) &= w \begin{cases} 0, & x < c, \\ 1 - e^{-\lambda(x-c)}, & x \geq c \end{cases} + (1-w) \begin{cases} 0, & x < c, \\ 1, & x \geq c \end{cases} \\ &= \begin{cases} 0, & x < c, \\ 1 - we^{-\lambda(x-c)}, & x \geq c \end{cases} \end{aligned} \tag{4.7.7}$$

$$\bar{F}_Z(x) = \begin{cases} 1, & x < c, \\ we^{-\lambda(x-c)}, & x \geq c \end{cases} \tag{4.7.8}$$

where w is the probability of being in the continuous state. Graphs of $F_Z(x)$ and $\bar{F}_Z(x)$ in Figure 4.2 are continuous except at the point c, where they have a jump discontinuity of height $(1-w)f_Y(c)$ and $wf_Y(c)$ respectively.



(a) $F_Z(x)$ (b) $\bar{F}_Z(x)$

Figure 4.2: The cumulative distribution (a) and right tail probability function (b) of Z in Equations 4.7.7 and 4.7.8 respectively.

4.8 Stochastic Dominance

The stochastic dominance (SD) is an approach used for risk assessment in decision making. It is commonly used in economics to rank different actions by comparing the probability distributions over possible outcomes, taking into account the

preferences of decision makers (Hanoch and Levy, 1969). There are infinite degree stochastic dominance (Thistle, 1993). In this thesis, we shall extensively use first-order stochastic dominance.

The first-order stochastic dominance is defined in Hadar and Russell (1969) as follows

Definition 4.8.1 The distribution G_Y is said to first-order stochastically dominate F_X (alternatively, Y first-order stochastically dominate X, and use the term $Y \succcurlyeq_{FOSD} X$ to stand for it) if

$$F_X(x) \geq G_Y(x) \; \forall x \in \mathbb{R}$$

which is equivalent to

$$\bar{G}_Y(x) \geq \bar{F}_X(x) \; \forall x \in \mathbb{R} \tag{4.8.1}$$

For convenience and to be consistent with our interest in $\bar{F}_X$, we work usually with Equation 4.8.1.

Some properties of first-order stochastic dominance are presented in Propositions 4.8.1 and 4.8.2.

Proposition 4.8.1 If X_1 first-order stochastically dominates X_2, and Y is independent of both X_1 and X_2, then $X_1 + Y$ first-order stochastically dominates $X_2 + Y$.

Proof.

$$P(X_1 + Y \leq t) = \int_{-\infty}^{\infty} F_{X_1}(t - y)\, dF_Y(y)$$

Since $X_1 \succcurlyeq_{FOSD} X_2$,

$$\begin{aligned} \int_{-\infty}^{\infty} F_{X_1}(t - y)\, dF_Y(y) &\leq \int_{-\infty}^{\infty} F_{X_2}(t - y)\, dF_Y(y) \\ &= P(X_2 + Y \leq t) \end{aligned}$$

Hence,

$$P(X_1 + Y \leq t) \leq P(X_2 + Y \leq t)$$

i.e. $X_1 + Y \succcurlyeq_{FOSD} X_2 + Y$ □

Proposition 4.8.2 If X_i first-order stochastically dominates Y_i; $i = 1, 2$, and X_1, X_2 are independent and Y_1, Y_2 are independent, then $X_1 + X_2$ first-order stochastically dominates $Y_1 + Y_2$.

Proof. By using Proposition 4.8.1,

$$X_1 + X_2 \succcurlyeq_{FOSD} Y_1 + X_2 \tag{4.8.2}$$

Moreover,

$$\begin{aligned} Y_1 + X_2 = X_2 + Y_1 \quad &\succcurlyeq_{FOSD} \quad Y_2 + Y_1 \\ &= \quad Y_1 + Y_2 \end{aligned} \tag{4.8.3}$$

Using Equations 4.8.2 and 4.8.3,

$$X_1 + X_2 \succcurlyeq_{FOSD} Y_1 + Y_2$$

□

4.9 Right Tail Probability Functions

The right tail probability function (complementary cumulative distribution function or tail distribution) is defined as

$$\bar{F}_X(x) = 1 - F_X(x)$$

It has different names such as the survival function in survival analysis, or reliability function in engineering. We will use right tail probability function as our term for $\bar{F}_X(x)$ unless redefined elsewhere in this thesis. The right tail probability and cumulative distribution functions inherit log-concavity from its corresponding log-concave density function $f_X(x)$ (Bagnoli and Bergstrom, 2005). However the probability density function of the mirror-image of Pareto distribution is

$$f_X(x) = \beta(-x)^{-\beta-1};\ x \in (-\infty, -1],\ \beta > 1$$

which is log-convex, whereas its right tail probability function $\bar{F}_X(x) = 1 - (-x)^{-\beta}$ is log-concave. This means that the class of distributions with log-concave right tail

probability functions is larger than the class of distributions that have log-concave densities.

The assumption of log-concavity property of right tail probability functions could be supported by using Proposition 4.9.1, it shows that the distance is decreasing between successive quantiles with the tail probability halving each time for a selection of such quantiles. Therefore there is some possibility that experts would be able to make such a judgement.

Proposition 4.9.1 If $\bar{F}_X$ is log-concave and $q > 0$ is fixed, then

$$\bar{F}_X^{-1}(q/4) - \bar{F}_X^{-1}(q/2) \leq \bar{F}_X^{-1}(q/2) - \bar{F}_X^{-1}(q)$$

See Figure 4.3 for an illustration of the proposition.

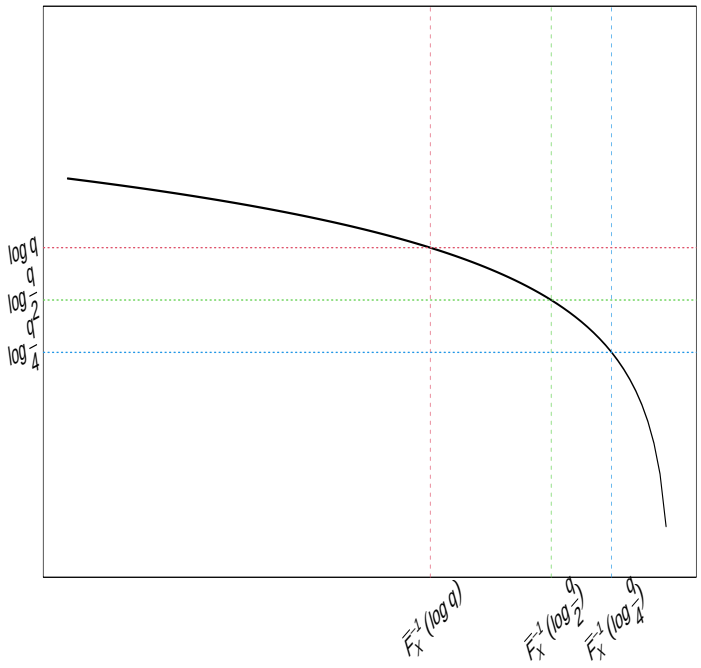


Figure 4.3: An example of log-concave right tail probability function, $\bar{F}_X$ in Proposition 4.9.1, to illustrate the relationship between the quantiles $\bar{F}_X^{-1}(q), \bar{F}_X^{-1}(q/2)$ and $\bar{F}_X^{-1}(q/4)$ of three successive right tail probabilities $q/4 < q/2 < q$.

Proof. If $q > 0$ is fixed, then

$$q/4 < q/2 < q$$

Hence,

$$\bar{F}_X^{-1}(q) \leq \bar{F}_X^{-1}(q/2) \leq \bar{F}_X^{-1}(q/4)$$

Using Definition 2.2.1 for $t = 1/2$, thus

$$\begin{aligned} \bar{F}_X\left(\frac{\bar{F}_X^{-1}(q) + \bar{F}_X^{-1}(q/4)}{2}\right) &\geq \left(\bar{F}_X\left(\bar{F}_X^{-1}(q)\right)\right)^{\frac{1}{2}} \left(\bar{F}_X\left(\bar{F}_X^{-1}(q/4)\right)\right)^{\frac{1}{2}} \\ &= q/2 \\ &= \bar{F}_X\left(\bar{F}_X^{-1}(q/2)\right) \end{aligned}$$

Since $\bar{F}_X$ is a decreasing function in its argument,

$$\begin{aligned} \frac{\bar{F}_X^{-1}(q) + \bar{F}_X^{-1}(q/4)}{2} &\leq \bar{F}_X^{-1}(q/2) \\ \bar{F}_X^{-1}(q) + \bar{F}_X^{-1}(q/4) &\leq 2\bar{F}_X^{-1}(q/2) \end{aligned}$$

Consequently,

$$\bar{F}_X^{-1}(q/4) - \bar{F}_X^{-1}(q/2) \leq \bar{F}_X^{-1}(q/2) - \bar{F}_X^{-1}(q)$$

□

In Sections 4.10 and 4.11, we develop a novel approach to explore the validity of the combined tails dominance property for distributions that have log-concave or partial log-concave right tail probability functions. In the latter case the log-concavity only applies to the upper part of the right tail probability function. This approach passes through two phases. In the first phase, which is presented in Theorem 4.9.1 and the corollary that follows after, we show that the distribution that has log-concave or partial log-concave right tail probability function is first-order stochastically dominated by SHE and SHEL distribution respectively. The second phase shows when the combined tail dominance property holds for SHE and SHEL distributions. Consequently, we will know when the combined tail dominance property holds for the distributions that have log-concave or partial log-concave right tail probability functions.

Theorem 4.9.1 Let X_1 have partial log-concave right tail probability function, i.e. log-concave right tail probability function for $x > c'$, and define $0 < w = \bar{F}_{X_1}(c') \leq 1$. Moreover, let $0 < p < w$ and x' be such that $\bar{F}_{X_1}(x') = p$. Then there exists

$X_2 \sim \text{SHEL}(w, c, \lambda)$, with $c \geq c'$ such that X_2 first-order stochastically dominates X_1.

Proof. Since $\log \bar{F}_{X_1}$ is concave for $x > c'$, it is absolutely continuous on each closed subinterval of (c', M) where the supremum M of the support of X_1 may be ∞. Left and right derivatives of $\log \bar{F}_{X_1}$ with respect to x exist at every point in (c', M) and are monotonically decreasing. The derivative exists at almost all points in (c', M) and is Lebesgue integrable.

Moreover, the left-hand derivative of $\log \bar{F}_{X_1}$ at x' is $-\lambda_1$ which is greater than or equal to the right-hand derivative at x' which is equal to $-\lambda_2$. Let λ satisfy $\lambda_1 \leq \lambda \leq \lambda_2$. Note that $\lambda_1 > 0$, as $p < w$ implies $c' < x'$ and the derivative of $\log \bar{F}_{X_1}$ must be negative somewhere in (c', x'). Furthermore, since the left-hand derivative is decreasing between c' and x', $\lambda_1 > 0$. Therefore, $\lambda > 0$ as a consequence.

In addition, c is chosen so that $\bar{F}_{X_2}(x') = p$, i.e. to solve

$$we^{-\lambda(x'-c)} = p$$

After taking the logarithm of both sides and simple algebraic,

$$c \;=\; x' + \frac{\log \frac{p}{w}}{\lambda}$$

Furthermore, If $G_X(x) = we^{-\lambda(x-c)};\ 0 < w < 1$, then $\big(\log G_X(x)\big)' = -\lambda$ which is constant for all x. Moreover, for $x \in (c', x')$,

$$\big(\log \bar{F}_{X_1}(x)\big)' \geq -\lambda_1 \geq -\lambda = \big(\log G_X(x)\big)'$$

Hence,

$$\int_x^{x'} \big(\log \bar{F}_{X_1}(t)\big)'\, dt \geq \int_x^{x'} \big(\log G_X(t)\big)'\, dt \tag{4.9.1}$$

$$\log \bar{F}_{X_1}(x') - \log \bar{F}_{X_1}(x) \geq \log G_X(x') - \log G_X(x)$$

Since $\Big(\log \bar{F}_{X_1}(x')\Big)' = \Big(\log \bar{G}_X(x')\Big)'$, $\log \bar{F}_{X_1}(x') = \log G_X(x')$ and

$$G_X(x) \geq \bar{F}_{X_1}(x) \;\; \forall x \in (c', x')$$

Therefore,

$$G_X(c') \;\; \geq \;\; \bar{F}_{X_1}(c') = w = G_X(c)$$

Hence, $c \geq c'$. See Figure 4.4 for illustration.

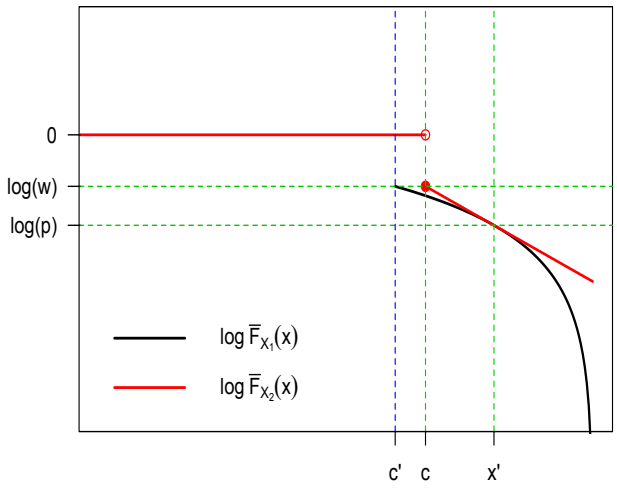


Figure 4.4: log right tail probability functions of X_1 and $X_2 \sim \text{SHEL}(w, c, \lambda)$ in Theorem 4.9.1.

In Figure 4.4 there are three areas that should be examined.

(1) For $x \geq x'$, the first derivative of $\log \bar{F}_{X_1}(x)$ is less than or equal to $-\lambda_2 \leq -\lambda$ which in turn means that

$$\int_{x'}^{x} \left(\log \bar{F}_{X_1}(t)\right)' dt \leq \int_{x'}^{x} \left(\log \bar{F}_{X_2}(t)\right)' dt$$

Hence,

$$\log \bar{F}_{X_1}(x) - \log \bar{F}_{X_1}(x') \leq \log \bar{F}_{X_2}(x) - \log \bar{F}_{X_2}(x') \tag{4.9.2}$$

Since $\log \bar{F}_{X_1}(x') = \log \bar{F}_{X_2}(x')$, Equation 4.9.2 becomes

$$\log \bar{F}_{X_1}(x) \leq \log \bar{F}_{X_2}(x)$$

Therefore,

$$F_{X_2}(x) \leq F_{X_1}(x) \; \forall x \geq x' \tag{4.9.3}$$

(2) For $x \in [c, x')$, the first derivative of $\log \bar{F}_{X_1}(x)$ is greater than or equal to $-\lambda_1 \geq -\lambda$ and this leads to

$$\int_{x}^{x'} \left(\log \bar{F}_{X_1}(t)\right)' dt \geq \int_{x}^{x'} \left(\log \bar{F}_{X_2}(t)\right)' dt$$

By applying the same method to solve Equation 4.9.1,

$$F_{X_2}(x) \leq F_{X_1}(x) \; \forall x \in [c, x') \tag{4.9.4}$$

(3) For $x < c$

Since $F_{X_2}(x) = 0$ for this range of x, so it must be the case that

$$F_{X_2}(x) \leq F_{X_1}(x) \; \forall x < c \tag{4.9.5}$$

Hence, from Equations 4.9.3, 4.9.4 and 4.9.5, we obtain

$$F_{X_2}(x) \leq F_{X_1}(x) \; \forall x \in \mathbb{R}$$

□

Corollary 4.9.2 Let X_1 have log-concave right tail probability function. Moreover, let x' be such that $\bar{F}_{X_1}(x') = p$. Then there exists $X_2 \sim \text{SHE}(c, \lambda)$ such that $\bar{F}_{X_2}(x') = p$, and X_2 first-order stochastically dominates X_1.

Proof. Put $w = 1$ in Proposition 4.9.1, then the proof follows. This problem is presented in Figure 4.5.

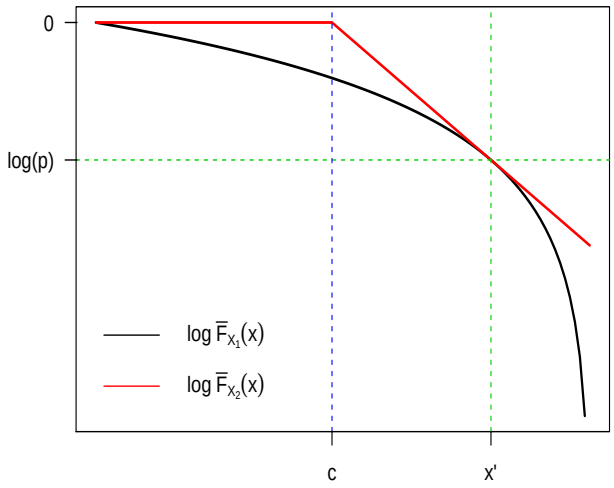


Figure 4.5: log right tail probability functions of X_1 and $X_2 \sim \text{SHE}(c, \lambda)$ in Corollary 4.9.2.

□

4.10 Log-Concave Right Tail Probability Functions

In the present section, we consider distributions that have log-concave right tail probability functions.

4.10.1 Convolution of SHE Distribution Functions

In Lemma 4.10.1 we shall obtain $P(X_1, X_2 \geq x_1' + x_2')$ when $X_i \sim \text{SHE}(c_i, \lambda_i);\ i = 1, 2$. Therefore, we can explore when $P(X_1, X_2 \geq x_1' + x_2') \leq p$.

Lemma 4.10.1 Let $X_i \sim \text{SHE}(c_i, \lambda_i);\ i = 1, 2$ be independent. Moreover, let $\lambda_1 \leq \lambda_2$ and x_i' be such that $\bar{F}_{X_i}(x_i') = p$. If $d = -\ln p$ and $s = \frac{\lambda_1}{\lambda_2}$, then

$$\frac{\bar{F}_{X_1+X_2}(x_1' + x_2')}{p} = \begin{cases} \frac{e^{-ds} - se^{-ds^{-1}}}{1-s} & ;\ \lambda_1 < \lambda_2 \\ \big(1 + 2d\big)e^{-d} & ;\ \lambda_1 = \lambda_2 \end{cases}$$

Proof. Let $X_i \sim \text{SHE}(c_i, \lambda_i);\ i = 1, 2$. Recall Equation 4.7.1

$$f_{X_i}(x_i) = \lambda_i e^{-\lambda_i(x_i - c_i)};\ x_i \geq c_i$$

To convolve X_1 and X_2, let $Z = X_1 + X_2$. Therefore, $X_2 = Z - X_1$. The probability density function of Z will be found as follows:

$$\begin{aligned} f_{X_1,Z}(x_1, z) &= f_{X_1,X_2}(x_1, z - x_1) \left| \frac{\partial(x_1, x_2)}{\partial(x_1, z)} \right| \\ &= \lambda_1\lambda_2 \exp\left(-\lambda_1(x_1 - c_1) - \lambda_2(z - x_1 - c_2)\right) \end{aligned} \tag{4.10.1}$$

where $c_1 \leq x_1 \leq z - c_2$.

We have two possibilities for the values of λ_1 and λ_2, either $\lambda_1 < \lambda_2$ or $\lambda_1 = \lambda_2$.

(1) If $\lambda_1 < \lambda_2$, then the probability density function of Z is

$$\begin{aligned} f_Z(z) &= \int_{c_1}^{z-c_2} f_{X_1,Z}(x_1, z)\, dx_1 \\ &= \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1}\left(e^{-\lambda_1(z-(c_1+c_2))} - e^{-\lambda_2(z-(c_1+c_2))}\right);\ z \geq c_1 + c_2 \end{aligned}$$

By integrating the last equation, thus

$$\begin{aligned} F_Z(z) &= \frac{\lambda_1\lambda_2}{\lambda_2-\lambda_1}\int_{c_1+c_2}^{z}\left(e^{-\lambda_1(t-(c_1+c_2))}-e^{-\lambda_2(t-(c_1+c_2))}\right)dt \\ &= 1+\frac{\lambda_1 e^{-\lambda_2(z-(c_1+c_2))}}{\lambda_2-\lambda_1}-\frac{\lambda_2 e^{-\lambda_1(z-(c_1+c_2))}}{\lambda_2-\lambda_1},\ z\geq c_1+c_2 \end{aligned}$$

Hence,

$$\begin{aligned} \bar{F}_{X_1+X_2}(x_1'+x_2') &= \frac{\lambda_2 e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}}{\lambda_2-\lambda_1}-\frac{\lambda_1 e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}}{\lambda_2-\lambda_1} \\ &= \frac{e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}}{1-\frac{\lambda_1}{\lambda_2}}-\frac{\frac{\lambda_1}{\lambda_2}e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}}{1-\frac{\lambda_1}{\lambda_2}} \end{aligned} \tag{4.10.2}$$

Since

$$\begin{aligned} \bar{F}_{X_i}(x_i') &= e^{-\lambda_i(x_i'-c_i)} \\ &= p \end{aligned}$$

Thus

$$\bar{F}_{X_1+X_2}(x_1'+x_2') = p\left(\frac{e^{-\lambda_1(x_2'-c_2)}-\frac{\lambda_1}{\lambda_2}e^{-\lambda_2(x_1'-c_1)}}{1-\frac{\lambda_1}{\lambda_2}}\right)$$

Moreover, $e^{-\lambda_i(x_{3-i}'-c_{3-i})}=p^{\frac{\lambda_i}{\lambda_{3-i}}}$. Hence,

$$\bar{F}_{X_1+X_2}(x_1'+x_2') = p\left(\frac{p^{\frac{\lambda_1}{\lambda_2}}-\frac{\lambda_1}{\lambda_2}p^{\frac{\lambda_2}{\lambda_1}}}{1-\frac{\lambda_1}{\lambda_2}}\right)$$

Since $s=\frac{\lambda_1}{\lambda_2}$ and $p=e^{-d}$, the last equation becomes

$$\bar{F}_{X_1+X_2}(x_1'+x_2') = p\left(\frac{e^{-ds}-se^{-ds^{-1}}}{1-s}\right)$$

Hence,

$$\frac{\bar{F}_{X_1+X_2}(x_1'+x_2')}{p} = \left(\frac{e^{-ds}-se^{-ds^{-1}}}{1-s}\right) \tag{4.10.3}$$

(2) If $\lambda_1=\lambda_2=\lambda$, then Equation 4.10.1 becomes

$$f_{X_1,Z}(x_1,z) = \lambda^2\exp\left(-\lambda(z-c_1-c_2)\right)$$

Therefore,

$$\begin{aligned} f_Z(z) &= \int_{c_1}^{z-c_2} f_{X_1,Z}(x_1,z)\,dx_1 \\ &= \lambda^2(z-c_1-c_2)\exp\left(-\lambda(z-c_1-c_2)\right);\ z\geq c_1+c_2 \end{aligned}$$

By using integration by parts, the last equation becomes

$$F_Z(z) = -(\lambda(z - c_1 - c_2) + 1)\exp(-\lambda(z - c_1 - c_2)) + 1$$

Therefore,

$$\bar{F}_Z(z) = (\lambda(z - c_1 - c_2) + 1)\exp(-\lambda(z - c_1 - c_2));\ z \geq c_1 + c_2$$

and

$$\begin{aligned} \bar{F}_{X_1+X_2}(x_1' + x_2') &= (\lambda(x_1' + x_2' - c_1 - c_2) + 1) \\ &\times \exp(-\lambda(x_1' + x_2' - c_1 - c_2)) \end{aligned} \tag{4.10.4}$$

Since $e^{-\lambda(x_i' - c_i)} = p$,

$$\bar{F}_{X_1+X_2}(x_1' + x_2') = (1 - 2\ln p)\, p^2$$

Moreover, since $p = e^{-d}$,

$$\bar{F}_{X_1+X_2}(x_1' + x_2') = p(1 + 2d)\, e^{-d}$$

and

$$\frac{\bar{F}_{X_1+X_2}(x_1' + x_2')}{p} = (1 + 2d)\, e^{-d} \tag{4.10.5}$$

Therefore, the result follows from Equations 4.10.3 and 4.10.5. □

Define

$$\begin{aligned} W(d, s) &= \frac{\bar{F}_{X_1+X_2}(x_1' + x_2')}{p} \\ &= \frac{e^{-ds} - se^{-ds^{-1}}}{1 - s} \end{aligned} \tag{4.10.6}$$

which represents the ratio of $\bar{F}_{X_1+X_2}(x_1' + x_2')$ to p.

We began by exploring the behaviour of $W(d, s)$ for different values of d as presented in Figure 4.6.

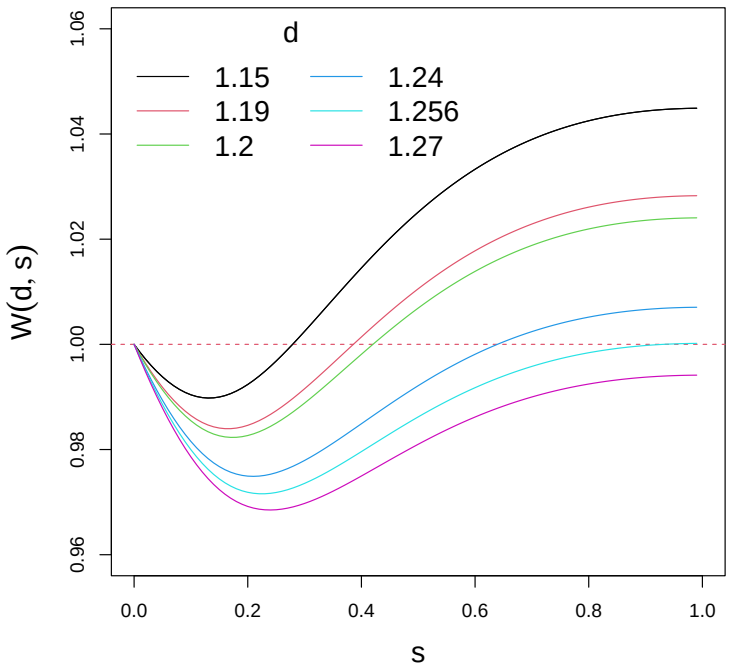


Figure 4.6: Plot of $W(d,s)$ for different values of d. The horizontal dashed line highlights when $W(d,s) \leq 1$.

From Figure 4.6, $W(d,s) \leq 1$ when s approaches one for a specific value of d. Thus define

$$\tilde{W}(d) = \lim_{s \to 1} W(d,s)$$

Using L'Hôspital's rule,

$$\tilde{W}(d) = (2d+1)e^{-d} \tag{4.10.7}$$

The previous equation is the same as Equation 4.10.5.

Differentiate Equation 4.10.7 with respect to d. Thus

$$\tilde{W}'(d) = e^{-d}\Big(1 - 2d\Big)$$

Therefore, $\tilde{W}(d)$ increases for $d < 0.5$ and then decreases for $d > 0.5$. The maximum value of $\tilde{W}(d)$ is 1.213061 at $d = 0.5$, see Figure 4.7 for illustration.

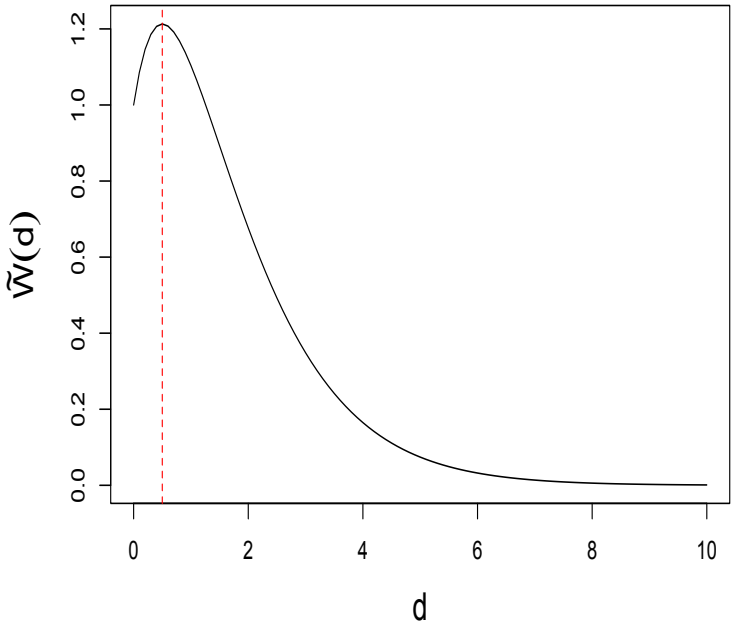


Figure 4.7: Plot of $\tilde{W}(d)$. The red dashed line is at $\arg\max_d \tilde{W}(d) = 0.5$.

Proposition 4.10.1 explores the behaviour of $W(d, s)$ as d increases. In addition, it will be used later in Theorem 4.10.2.

Proposition 4.10.1 $W(d, s)$, which represents the ratio of $\bar{F}_{X_1+X_2}(x_1' + x_2')$ to p, is decreasing in d for $d > \frac{-\ln(s)}{2\sinh(-\ln(s))}$ and $s \in (0, 1)$.

Proof. Differentiating Equation 4.10.6 with respect to d, thus

$$\frac{\partial}{\partial d} W(d, s) = \frac{e^{-ds^{-1}} - se^{-ds}}{1 - s} \tag{4.10.8}$$

Define $a = -\ln(s);\ a > 0$. Thus Equations 4.10.6 and 4.10.8 respectively become

$$W(d, e^{-a}) = \frac{e^{-de^{-a}} - e^{-a}e^{-de^{a}}}{1 - e^{-a}}$$

and

$$\frac{\partial}{\partial d} W(d, e^{-a}) = \frac{e^{-de^{a}} - e^{-a}e^{-de^{-a}}}{1 - e^{-a}} \tag{4.10.9}$$

For any $a > 0,\ d = 0$,

$$W(0, e^{-a}) = 1 \text{ and } \frac{\partial}{\partial d} W(0, e^{-a}) = 1$$

Moreover, the denominator of Equation 4.10.9 is always positive and for sufficiently large d the numerator $e^{-de^a}\left(1 - e^{-a}e^{d(e^a - e^{-a})}\right)$ is negative. Therefore, $\frac{\partial}{\partial d}W(d, e^{-a})$ is negative for sufficiently large d.

In addition, if there is a single $d^* > 0$ such that $\frac{\partial}{\partial d}W(d^*, e^{-a}) = 0$, by continuity of $\frac{\partial}{\partial d}W(d, e^{-a})$, $W(d, e^{-a})$ increases for $d < d^*$ and decreases for $d > d^*$.
Equate Equation 4.10.9 to zero,

$$\begin{aligned}
\frac{\partial}{\partial d}W(d, e^{-a}) = 0 \quad &\iff \quad e^{-de^a} - e^{-a}e^{-de^{-a}} = 0 \\
&\iff \quad e^{-de^a} = e^{-a}e^{-de^{-a}} \\
&\iff \quad \ln\left(e^{-de^a}\right) = \ln\left(e^{-a}e^{-de^{-a}}\right) \\
&\iff \quad -de^a = -a - de^{-a} \\
&\iff \quad a = d(e^a - e^{-a}) \\
&\iff \quad d = a/2\sinh(a)
\end{aligned}$$

thus $d^*(a) = a/2\sinh(a)$ is unique. The Taylor expansion of $\sinh(a)/a$ is

$$\sinh(a)/a = 1 + \frac{a^2}{3!} + \frac{a^4}{5!} + \frac{a^6}{7!} + \ldots$$

which is increasing for $a > 0$, therefore $a/\sinh(a)$ is decreasing for $a > 0$, and its largest value is 1 at $a = 0$. Thus the largest value of $d^*(a)$ is $1/2$. □

In Theorem 4.10.2, we will prove that the combined tail dominance property holds (under a specific condition) when considering two random variables that have shifted exponential distributions.

Theorem 4.10.2 Let $X_i \sim \text{SHE}(c_i, \lambda_i);\ i = 1, 2$ be independent. Moreover, let $\lambda_1 \leq \lambda_2$ and x_i' be such that $\bar{F}_{X_i}(x_i') = p$. If $p \leq \left(2\mathcal{W}_{-1}\left(\frac{e^{-\frac{1}{2}}}{2}\right)\right)^{-1} \approx 0.2846681$, then

$$P(X_1 + X_2 \geq x_1' + x_2') \leq p$$

Proof. Notice that $\bar{F}_{X_1+X_2}(x_1' + x_2') \leq p \iff W(d, s) \leq 1$. From Figure 4.6, it appears that $W(d, s) \leq 1$ when $\tilde{W}(d)$ equals to one and this will be proved formally later. Hence, we will first equate $\tilde{W}(d)$ to one in order to obtain the value of d that

makes $W(d,s) \leq 1$, for $s \in (0,1)$, as follows:

$$(2d+1)e^{-d} = 1 \rightarrow d + \frac{1}{2} = \frac{e^d}{2} \rightarrow -\ln(p) + \frac{1}{2} = \frac{1}{2p}$$

The last equation is equivalent to

$$-\frac{1}{2p}e^{-\frac{1}{2p}} = -\frac{1}{2}e^{-\frac{1}{2}} \rightarrow \frac{-1}{2p} = \mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)$$

Hence,

$$p = \left(-2\mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)\right)^{-1} \approx 0.2846681$$

Thus

$$d = \ln\left(-2\mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)\right) \approx 1.256431$$

where $\mathcal{W}_{-1}$ is the bottom branch of the Lambert $\mathcal{W}$ function, see Figure 4.8 for illustration and Appendix A.2 for more detail. Notice that the upper branch $\mathcal{W}_0(\frac{-e^{-\frac{1}{2}}}{2})$ is not of our interest since d will be 0 which is not acceptable in this context.

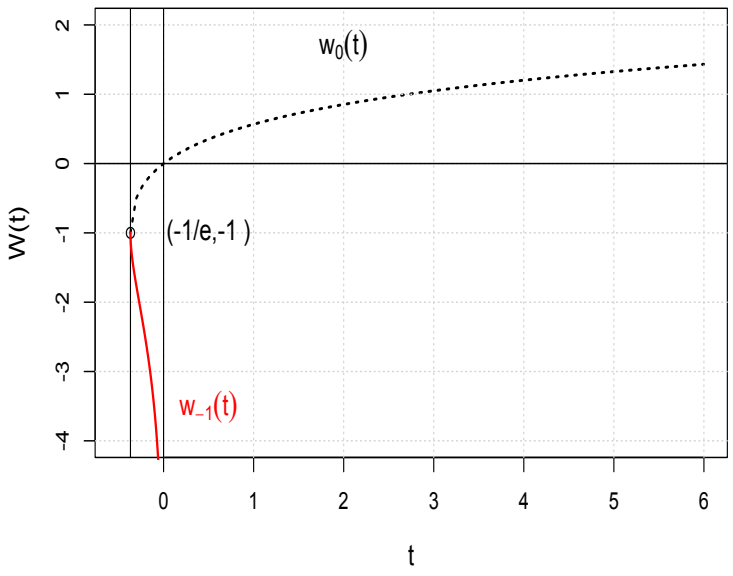


Figure 4.8: Two real branches of the Lambert $\mathcal{W}$ function.

Now we shall study the behaviour of $W(d,s)$ when s approaches zero and one as follows:

(1) s approaches 0

By differentiating Equation 4.10.6 with respect to s. Thus

$$\begin{aligned}\frac{\partial}{\partial s}W(d,s) &= \frac{\big(1-(1-s)d\big)e^{-ds}-\big((1-s)(1+\frac{d}{s})+s\big)e^{-\frac{d}{s}}}{(1-s)^2}\\ &= \frac{\big(1-(1-s)d\big)e^{-ds}-\big(1+(1-s)\frac{d}{s}\big)e^{-\frac{d}{s}}}{(1-s)^2}\end{aligned} \tag{4.10.10}$$

For small s, the first-order Maclaurin series expansion of $W(d,s)$ is

$$W(d,s)\approx 1+(1-d)s$$

Therefore, $W(d,s)<1$ for $d>1$ and small s.

(2) s approaches 1

When s approaches one, $W(d,s)$ in Equation 4.10.6 can be rewritten as

$$W(d,1-x)=\frac{e^{-d(1-x)}-(1-x)e^{\frac{-d}{(1-x)}}}{x} \tag{4.10.11}$$

Define $f(x)=e^{-d(1-x)}-(1-x)e^{\frac{-d}{(1-x)}}$, hence $f(x)$ is infinitely differentiable for $x>0$. The Maclaurin series expansion of $f(x)$ is obtained as follows:

$$\begin{aligned}f'(x) &= de^{-d(1-x)}+\left(\frac{d}{1-x}+1\right)e^{-\frac{d}{(1-x)}}\\ f''(x) &= d^2e^{-d(1-x)}-\frac{d^2}{(1-x)^3}e^{-\frac{d}{(1-x)}}\\ f'''(x) &= d^3e^{-d(1-x)}-\left(\frac{3d^2}{(1-x)^4}-\frac{d^3}{(1-x)^5}\right)e^{-\frac{d}{(1-x)}}\end{aligned}$$

Hence,

$$f(0)=0,\ f'(0)=(1+2d)e^{-d},\ f''(0)=0,\ \text{and } f'''(0)=\big(2d^3-3d^2\big)e^{-d}$$

and

$$f(x)\approx(1+2d)e^{-d}x+d^2\big(2d-3\big)e^{-d}x^3/6 \tag{4.10.12}$$

Therefore, the Maclaurin series expansion of $W(d,1-x)$ for small x is

$$\begin{aligned}W(d,1-x) &\approx (1+2d)e^{-d}+d^2\big(2d-3\big)e^{-d}x^2/6\\ &= e^{-d}\Big(1+2d+\frac{d^2\big(2d-3\big)}{6}x^2\Big)\end{aligned}$$

Hence, for small x, $W(d,1-x)<e^{-d}(1+2d)$ if and only if the coefficient of x^2 is negative, and the coefficient of x^2 is negative if and only if $d<1.5$. Therefore, for

$1 < d < 1.5$, $W(d,s)$ decreases for $s = 0^+$ and increases for $s = 1^-$.

To know the behaviour of $W(d,s)$ for s between 0 and 1, we shall argue as follows:

Equation 4.10.10 can be rewritten as

$$\frac{\partial}{\partial s}W(d,s) = (1-s)^{-2}\big(g(ds) - g(ds^{-1})\big)$$

where

$$g(x) = (x + 1 - d)e^{-x}$$

Notice that, for $d > 1$,

$$\frac{\partial}{\partial s}W(d,s) = 0 \iff g(ds) = g(ds^{-1})$$

Since $g(ds^{-1})$ is positive for $s \in (0,1)$, the only possibility for $\frac{\partial}{\partial s}W(d,s)$ to be zero is when $g(ds) > 0$, and this happens when $\frac{d-1}{d} < s < 1$.

Write $s = e^{-a}$; $0 < a < -\log(\frac{d-1}{d})$, and set

$$h(a) = \log g(de^{-a}) - \log g(de^{a})$$

Therefore,

$$h(a) = 0 \iff \frac{\partial}{\partial s}W(d,s)\big|_{s=e^{-a}} = 0$$

Differentiating $h(a)$ with respect to a,

$$\begin{aligned} h'(a) &= \frac{d^2e^{-a}\big(e^{-a}-1\big)}{de^{-a}+1-d} + \frac{d^2e^{a}\big(e^{a}-1\big)}{de^{a}+1-d} \\ &= \frac{(d^3 - d^2(1-d))(e^{-a}+e^{a}) + (1-d)d^2(e^{-2a}+e^{2a}) - 2d^3}{(de^{-a}+1-d)(de^{a}+1-d)} \\ &= \frac{2d^2\Big((2d-1)\cosh(a) + (1-d)\cosh(2a) - d\Big)}{(de^{-a}+1-d)(de^{a}+1-d)} \end{aligned}$$

Since $\cosh(2a) = 2\cosh^2(a) - 1$,

$$\begin{aligned} h'(a) &= \frac{-2d^2\Big(2(d-1)\big(\cosh^2(a) - \cosh(a)\big) - \cosh(a) + 1\Big)}{(de^{-a}+1-d)(de^{a}+1-d)} \\ &= \frac{-2d^2\Big(2(d-1)\cosh(a) - 1\Big)\Big(\cosh(a) - 1\Big)}{(de^{-a}+1-d)(de^{a}+1-d)} \end{aligned}$$

The numerator is a quadratic equation in $\cosh(a)$ and the denominator of $h'(a)$ is positive. One root is $\cosh(a) = 1$ and $\cosh(a) = 1 \iff a = 0$. The other root is

$\cosh(a) = \frac{1}{2(d-1)}$.

The function $h(a)$ is continuous over the interval $\left[0, -\log\left(\frac{d-1}{d}\right)\right]$ and differentiable over the interval $\left(0, -\log\left(\frac{d-1}{d}\right)\right)$. Since $h(0) = 0$, and there is at most one positive a which equals to $\cosh^{-1}\left(\frac{1}{2(d-1)}\right)$ such that $h'\left(\cosh^{-1}\left(\frac{1}{2(d-1)}\right)\right) = 0$, there is at most one positive $a_o > \cosh^{-1}\left(\frac{1}{2(d-1)}\right)$ such that $h(a_o) = 0$ and satisfies the Rolle's theorem (extension of mean value theorem)

$$h'\left(\cosh^{-1}\left(\frac{1}{2(d-1)}\right)\right) = \frac{h(a_o) - h(0)}{a_o} = 0$$

This means that, there is at most one turning point of $W(d, s)$ at $s = e^{-a_o}$ provided that $d > 1$.

Define $d_{\text{SHE}} = \ln\left(-2\mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)\right)$. Therefore, $d_{\text{SHE}} \approx 1.256431$ which is between 1 and 1.5.

Moreover, $W(d, s)$ decreases when $s = 0^+$ and increases when $s = 1^-$, for $1 < d < 1.5$. Hence, the value of $W(d_{\text{SHE}}, s)$ should be less than one at that turning point. Therefore, $W(d_{\text{SHE}}, s) \leq 1$. See Figure 4.9 for illustration.

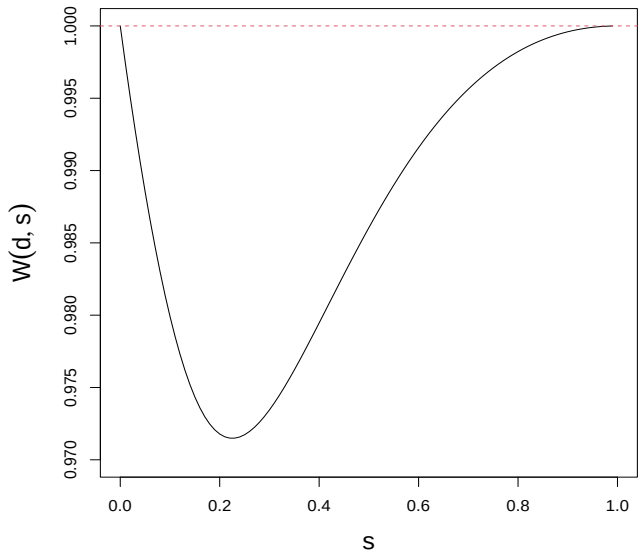


Figure 4.9: Plot of $W(d_{\text{SHE}}, s)$. The horizontal dashed line highlights that $W(d_{\text{SHEL}}, s) \leq 1$.

Using Proposition 4.10.1, $W(d, s) \leq 1$ for $d \geq d_{\text{SHE}}$. Therefore, if $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)\right)^{-1}$,

then $P(X_1 + X_2 \geq x_1' + x_2') \leq p$ □

For numerical illustration, we shall consider $X_i \sim \text{SHE}(1,1);\ i = 1,2$. Using Equations 4.7.2 and 4.10.4, the right tail probability functions of X_i at x_i' and $X_1 + X_2$ at $x_1' + x_2'$ are respectively given by

$$\bar{F}_{X_i}(x_i') = e^{-(x_i'-1)},\ x_i' \geq 1$$

and

$$\begin{aligned} \bar{F}_{X_1+X_2}(x_1' + x_2') &= ((x_1' + x_2' - 2) + 1) \\ &\times \exp\left(-\left(x_1' + x_2' - 2\right)\right);\ x_1' + x_2' \geq 2 \end{aligned}$$

Suppose that we take

$$\begin{aligned} p_i = p &= e^{-d_{\text{SHE}}} \\ &\approx 0.2846681 \end{aligned}$$

Then $x_i' = 2.256431$ and $\bar{F}_{X_1+X_2}(2\,(2.256431)) = 0.2846681 = p$. If we take $p_i = p = 0.3$, then $x_i' = 2.203973$ and $\bar{F}_{X_1+X_2}\,(2(2.203973)) = 0.3067151 > 0.3 = p$. Whereas, if $p_i = p = 0.2$, then $x_i' = 2.609438$ and $\bar{F}_{X_1+X_2}\,(2(2.609438)) = 0.168755 < 0.2 = p$. This illustrates the boundary between situations where the combined tails dominance property holds and where it does not, in the case when the distributions of uncertain quantities $X_i;\ i = 1,2$ are $\text{SHE}(c_i, \lambda_i)$.

4.10.2 Generalization for $n > 2$

Theorem 4.10.3 generalises the result obtained in Theorem 4.10.2 for $n > 2$. In addition, it will be used later in Theorem 4.10.4.

Theorem 4.10.3 Let $X_i \sim \text{SHE}(c_i, \lambda_i);\ i = 1, 2, \ldots, n$ be independent. Moreover, let $\lambda_1 \leq \lambda_2$ and x_i' be such that $\bar{F}_{X_i}(x_i') = p$. If $p \leq \left(2\mathcal{W}_{-1}\left(\frac{e^{-\frac{1}{2}}}{2}\right)\right)^{-1} \approx 0.2846681$, then

$$P(X_1 + X_2 + \cdots + X_n \geq x_1' + x_2' + \cdots + x_n') \leq p$$

Proof. We will use mathematical induction to prove this theorem.

(1) Consider $n = 2$, the statement is true by using Theorem 4.10.2.

(2) Assume the statement is true for n, hence

$$P(X_1 + X_2 + \cdots + X_n \geq x_1' + x_2' + \cdots + x_n') \leq p$$

(3) We shall prove that it is true for $n+1$ as follows:

The log-concavity of probability density function is preserved under convolution, hence $X_1 + X_2 + \cdots + X_n$ has log-concave probability density function, and then has log-concave right tail probability function, and so by Corollary 4.9.2, there exists $X^* \sim \text{SHE}(c^*, \lambda^*)$ such that

$$X^* \succcurlyeq_{FOSD} X_1 + X_2 + \cdots + X_n \tag{4.10.13}$$

and

$$\begin{aligned} P(X^* \geq x_1' + x_2' + \cdots + x_n') &= P(X_1 + X_2 + \cdots + X_n \geq x_1' + x_2' + \cdots + x_n') \\ &\leq p \end{aligned}$$

Therefore, there exists $x^* \leq x_1' + x_2' + \cdots + x_n'$ such that

$$P(X^* \geq x^*) = p$$

By applying Theorem 4.10.2 for X^*, X_{n+1} and tail probability p. Thus

$$P(X^* + X_{n+1} \geq x^* + x_{n+1}') \leq p \tag{4.10.14}$$

Moreover, using Equation 4.10.13 and Proposition 4.8.1, we obtain

$$X^* + X_{n+1} \succcurlyeq_{FOSD} X_1 + X_2 + \cdots + X_n + X_{n+1}$$

Therefore,

$$\begin{gathered} P\big(X_1 + X_2 + \cdots + X_n + X_{n+1} \geq x_1' + x_2' + \cdots + x_n' + x_{n+1}'\big) \\ \leq \\ P(X^* + X_{n+1} \geq x_1' + x_2' + \cdots + x_n' + x_{n+1}') \\ \leq \\ P(X^* + X_{n+1} \geq x^* + x_{n+1}') \leq p \end{gathered}$$

□

Theorem 4.10.4 shows that the combined tails dominance property holds for distributions that have log-concave right tail probability functions under a specific condition.

Theorem 4.10.4 Let $Y_i;\ i = 1, 2, \ldots, n$ be independent and have log-concave right tail probability function. Moreover, let x_i be such that $\bar{F}_{Y_i}(x_i') = p$.
If $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-e^{-\frac{1}{2}}}{2}\right)\right)^{-1}$, then

$$P\big(Y_1 + Y_2 + \cdots + Y_n \geq x_1' + x_2' + \cdots + x_n'\big) \leq p$$

Proof. Firstly, using Corollary 4.9.2, Y_i are first-order stochastically dominated by $X_i \sim \text{SHE}(c_i, \lambda_i)$. Secondly, applying Proposition 4.8.2 $(n-1)$ times implies that $Y_1 + Y_2 + \cdots + Y_n$ is first-order stochastically dominated by $X_1 + X_2 + \cdots + X_n$. Thus

$$P\big(Y_1 + Y_2 + \cdots + Y_n \geq x_1' + x_2' + \cdots + x_n'\big)$$
$$\leq$$
$$P(X_1 + X_2 + \cdots + X_n \geq x_1' + x_2' + \cdots + x_n')$$

Finally, using Theorem 4.10.3, then

$$P\big(Y_1 + Y_2 + \cdots + Y_n \geq x_1' + x_2' + \cdots + x_n'\big) \leq p$$

□

4.11 Partial Log-Concave Right Tail Probability Functions

In this section, we consider distributions that have partial log-concave right tail probability functions. This is a generalization of the case in Section 4.10 where the considered distributions have log-concave right tail probability functions.

4.11.1 Convolution of SHEL Distribution Functions

In Theorem 4.11.1 we shall obtain $P(Z_1, Z_2 \geq x_1' + x_2')$ when $Z_i \sim \text{SHEl}(w, c_i, \lambda_i);\ i = 1, 2$. Therefore, we can explore when $P(Z_1, Z_2 \geq x_1' + x_2') \leq p$.

Theorem 4.11.1 Let $Z_i \sim \text{SHEL}(w, c_i, \lambda_i);\ i = 1, 2$ be independent. Moreover, let $\lambda_1 \leq \lambda_2$ and x_i' be such that $\bar{F}_{Z_i}(x_i') = p < w$. If $d = -\ln(p/w) > 0$ and $s = \frac{\lambda_1}{\lambda_2}$, then

$$\frac{\bar{F}_{Z_1+Z_2}(x_1' + x_2')}{p} = \begin{cases} wW^*(d,s) + (1-w)W^{**}(d,s) & ;\ \lambda_1 < \lambda_2 \\ we^{-d}\Big(2\big(d + \frac{1}{w}\big) - 1\Big) & ;\ \lambda_1 = \lambda_2 \end{cases}$$

where

$W^*(d,s) = \frac{e^{-ds} - se^{-\frac{d}{s}}}{1-s}$ and $W^{**}(d,s) = e^{-ds} + e^{-\frac{d}{s}}$.

Proof. Using Equation 4.7.6, then the characteristic function of $Z_1 + Z_2$ is

$$\begin{aligned} \psi_{Z_1+Z_2}(t) &= \psi_{Z_1}(t)\psi_{Z_2}(t) \\ &= w^2\psi_{X_1}(t)\psi_{X_2}(t) + w(1-w)\left(\psi_{X_1}(t)\psi_{Y_2}(t) + \psi_{Y_1}(t)\psi_{X_2}(t)\right) \\ &+ \psi_{Y_1}(t)\psi_{Y_2}(t) \\ &= w^2 e^{it(c_1+c_2)}\left(1 - it\lambda_1^{-1}\right)^{-1}\left(1 - it\lambda_2^{-1}\right)^{-1} \\ &+ w(1-w)e^{it(c_1+c_2)}\left(\left(1 - it\lambda_1^{-1}\right)^{-1} + \left(1 - it\lambda_2^{-1}\right)^{-1}\right) \\ &+ (1-w)^2 e^{it(c_1+c_2)} \end{aligned} \tag{4.11.1}$$

We have two possibilities for the values of λ_1 and λ_2, either $\lambda_1 < \lambda_2$ or $\lambda_1 = \lambda_2$.

(1) Assume $\lambda_1 < \lambda_2$. By using Equation 4.11.1, the right tail probability function of $Z_1 + Z_2$ is

$$\begin{aligned} \bar{F}_{Z_1+Z_2}(x) &= w^2\bar{F}_{X_1+X_2}(x) + w(1-w)\left(\bar{F}_{X_1^*}(x) + \bar{F}_{X_2^*}(x)\right) \\ &+ (1-w)^2\bar{F}_{Y^*}(x) \end{aligned} \tag{4.11.2}$$

where $X_i^* \sim \text{SHE}(c_1 + c_2, \lambda_i)$, $Y^* \sim L(c_1 + c_2)$. Using Equations 4.7.2 and 4.7.4,

$$\bar{F}_{X_i^*}(x) = e^{-\lambda_i(x-(c_1+c_2))}\begin{cases} 0, & x < c_1 + c_2, \\ 1, & x \geq c_1 + c_2 \end{cases} \tag{4.11.3}$$

$$\bar{F}_{Y^*}(x) = \begin{cases} 1, & x < c_1 + c_2, \\ 0, & x \geq c_1 + c_2 \end{cases} \tag{4.11.4}$$

For $x \geq c_1 + c_2$, and by using Equations 4.10.2, 4.11.3 and 4.11.4 , Equation

4.11.2 becomes

$$\begin{aligned}\bar{F}_{Z_1+Z_2}(x_1'+x_2') &= \frac{w^2}{\lambda_2-\lambda_1}\left(\lambda_2 e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}-\lambda_1 e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}\right)\\ &+ w(1-w)\left(e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}+e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}\right)\end{aligned}$$

which is equivalent to

$$\begin{aligned}\bar{F}_{Z_1+Z_2}(x_1'+x_2') &= w^2\left(\frac{e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}-\frac{\lambda_1}{\lambda_2}e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}}{1-\frac{\lambda_1}{\lambda_2}}\right)\\ &+ w(1-w)\left(e^{-\lambda_1\left(x_1'+x_2'-(c_1+c_2)\right)}+e^{-\lambda_2\left(x_1'+x_2'-(c_1+c_2)\right)}\right)\end{aligned}$$

Since

$$\begin{aligned}\bar{F}_{Z_i}(x_i') &= we^{-\lambda_i(x_i'-c_i)}\\ &= p\end{aligned}$$

Thus

$$\begin{aligned}\bar{F}_{Z_1+Z_2}(x_1'+x_2') &= wp\left(\frac{e^{-\lambda_1\left(x_2'-c_2\right)}-\frac{\lambda_1}{\lambda_2}e^{-\lambda_2\left(x_1'-c_1\right)}}{1-\frac{\lambda_1}{\lambda_2}}\right)\\ &+ (1-w)p\left(e^{-\lambda_1\left(x_2'-c_2\right)}+e^{-\lambda_2\left(x_1'-c_1\right)}\right)\end{aligned}$$

Moreover, $e^{-\lambda_i(x_{3-i}'-c_{3-i})}=\left(\frac{p}{w}\right)^{\frac{\lambda_i}{\lambda_{3-i}}}$. Hence,

$$\bar{F}_{Z_1+Z_2}(x_1'+x_2') = wp\left(\frac{\left(\frac{p}{w}\right)^{\frac{\lambda_1}{\lambda_2}}-\frac{\lambda_1}{\lambda_2}\left(\frac{p}{w}\right)^{\frac{\lambda_2}{\lambda_1}}}{1-\frac{\lambda_1}{\lambda_2}}\right)+(1-w)p\left(\left(\frac{p}{w}\right)^{\frac{\lambda_1}{\lambda_2}}+\left(\frac{p}{w}\right)^{\frac{\lambda_2}{\lambda_1}}\right)$$

Since $s=\frac{\lambda_1}{\lambda_2}$ and $\frac{p}{w}=e^{-d}$,

$$\bar{F}_{Z_1+Z_2}(x_1'+x_2') = wp\left(\frac{e^{-ds}-se^{-\frac{d}{s}}}{1-s}\right)+(1-w)p\left(e^{-ds}+e^{-\frac{d}{s}}\right)$$

$$\frac{\bar{F}_{Z_1+Z_2}(x_1'+x_2')}{p} = wW^*(d,s)+(1-w)W^{**}(d,s) \tag{4.11.5}$$

where $W^*(d,s)=\frac{e^{-ds}-se^{-\frac{d}{s}}}{1-s}$ and $W^{**}(d,s)=e^{-ds}+e^{-\frac{d}{s}}$

(2) If $\lambda_1 = \lambda_2 = \lambda$, then Equation 4.11.1 becomes

$$\begin{aligned}\psi_{Z_1+Z_2}(t) &= w^2 e^{it(c_1+c_2)}\left(1-it\lambda^{-1}\right)^{-2} \\ &+ 2w(1-w)e^{it(c_1+c_2)}\left(1-it\lambda^{-1}\right)^{-1} \\ &+ (1-w)^2 e^{it(c_1+c_2)}\end{aligned}$$

Hence, the right tail probability function of $Z_1 + Z_2$ is

$$\begin{aligned}\bar{F}_{Z_1+Z_2}(x) &= w^2\bar{F}_{X_1+X_2}(x) + 2w(1-w)\bar{F}_{X^*}(x) \\ &+ (1-w)^2\bar{F}_{Y^*}(x) \qquad (4.11.6)\end{aligned}$$

where $X^* \sim \text{SHE}(c_1 + c_2, \lambda),\ Y^* \sim L(c_1 + c_2)$.

Using Equation 4.7.2,

$$\bar{F}_{X^*}(x) = e^{-\lambda(x-c_1-c_2)} \qquad (4.11.7)$$

For $x_1' + x_2' \geq c_1 + c_2$, using Equations 4.10.4, 4.11.4 and 4.11.7, Equation 4.11.6 becomes

$$\begin{aligned}\bar{F}_{Z_1+Z_2}(x_1'+x_2') &= w^2\left(\lambda\left(x_1'+x_2'-c_1-c_2\right)+1\right) \\ &\times \exp\left(-\lambda\left(x_1'+x_2'-c_1-c_2\right)\right) \\ &+ 2w(1-w)\exp\left(-\lambda\left(x_1'+x_2'-c_1-c_2\right)\right)\end{aligned}$$

Since $we^{-\lambda(x_i'-c_i)} = p$,

$$\bar{F}_{Z_1+Z_2}(x_1'+x_2') = p^2\left(1-2\ln(p/w)\right) + 2p^2\left(\frac{1-w}{w}\right)$$

and

$$\frac{\bar{F}_{Z_1+Z_2}(x_1'+x_2')}{p} = we^{-d}\left(2\left(d+\frac{1}{w}\right)-1\right) \qquad (4.11.8)$$

Therefore, the result follows from Equations 4.11.5 and 4.11.8. □

Define $W(w,d,s) = \frac{\bar{F}_{Z_1+Z_2}(x_1'+x_2')}{p}$, which represents the ratio of $\bar{F}_{Z_1+Z_2}(x_1'+x_2')$ to p. Hence,

$$\begin{aligned}W(w,d,s) &= wW^*(d,s) + (1-w)W^{**}(d,s) \\ &= w\left(\frac{e^{-ds} - se^{-\frac{d}{s}}}{1-s}\right) + (1-w)\left(e^{-ds} + e^{-\frac{d}{s}}\right)\end{aligned} \qquad (4.11.9)$$

We begin by exploring the behaviour of $W(w,d,s)$, $s \in (0,1)$, for given w and different values of d as presented in Figure 4.10.

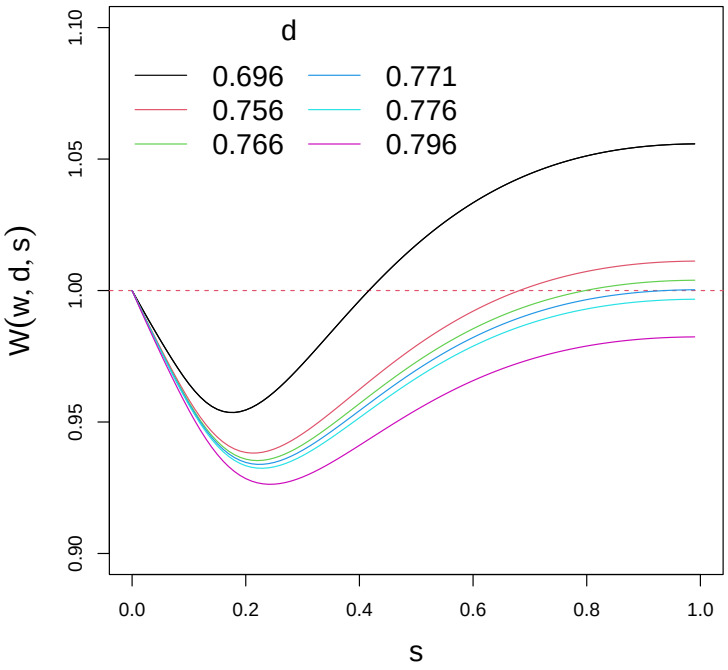


Figure 4.10: Plot of $W(w,d,s)$ for $w = 0.3$ and different values of d. The horizontal dashed line highlights when $W(w,d,s) \leq 1$.

Figure 4.10 shows that for $w = 0.3$, there is a value of d for which $W(w,d,1) = 1$, and such that $W(w,d,s) < 1$ for all $s \in (0,1)$. Indeed, this is what we are looking for. Hence, we shall try to find d depending on w that satisfies these results as follows:

First: We will examine the behaviour of $W(w,d,s)$ when s approaches zero as follows:

By differentiating Equation 4.11.9 with respect to s. Thus

$$\begin{aligned}\frac{\partial}{\partial s}W(w,d,s) &= w\left(\frac{(1-(1-s)d)\,e^{-ds} - \left(1+(1-s)\frac{d}{s}\right)e^{-\frac{d}{s}}}{(1-s)^2}\right) \\ &+ (1-w)\left(-de^{-ds} + \frac{d}{s^2}e^{-\frac{d}{s}}\right)\end{aligned}$$

For small s, the first-order Maclaurin series expansion of $W(w,d,s)$ is

$$\begin{aligned} W(w,d,s) &\approx 1 + w(1-d)s - (1-w)ds \\ &= 1 + (w-d)s \end{aligned}$$

Thus

$$W(w,d,s) \approx 1 + (w-d)s$$

Therefore, $W(w,d,s) < 1$ for $d > w$ and small s.

Second: We will obtain $\lim_{s\to 1} W(w,d,s)$ as follows:

Using L'Hospital's rule. Thus

$$\begin{aligned} \lim_{s\to 1} W^*(d,s) &= \lim_{s\to 1} \frac{e^{-ds} - se^{-\frac{d}{s}}}{1-s} \\ &= e^{-d}\big(1+2d\big) \end{aligned} \tag{4.11.10}$$

Using Equation 4.11.10, and $\lim_{s\to 1} W^{**}(d,s) = 2e^{-d}$. Hence,

$$\lim_{s\to 1} W(w,d,s) = we^{-d}\left(2\left(d+\frac{1}{w}\right)-1\right)$$

Define $\tilde{W}(w,d) = \lim_{s\to 1} W(w,d,s)$. Hence,

$$\tilde{W}(w,d) = we^{-d}\left(2\left(d+\frac{1}{w}\right)-1\right) \tag{4.11.11}$$

The previous equation is the same as Equation 4.11.8.

Third: We shall find d that makes $\tilde{W}(w,d)$ equal to one as follows:

Using Equation 4.11.11, and recall $p/w = e^{-d}$. Thus

$$\begin{aligned} we^{-d}\left(2\left(d+\frac{1}{w}\right)-1\right) &= 1 \\ d + \frac{1}{w} - \frac{1}{2} &= \frac{e^d}{2w} \\ -\frac{e^d}{2w} + d + \frac{1}{w} - \frac{1}{2} &= 0 \\ -\frac{1}{2p} - \ln p + \ln w + \frac{1}{w} - \frac{1}{2} &= 0 \end{aligned}$$

Take the exponential of the latter equation and re-arrange it, thus

$$\frac{-1}{w} e^{-(\frac{1}{w}-\frac{1}{2})} = \frac{-1}{p} e^{-\frac{1}{2p}}$$

which is equivalent to

$$\begin{aligned} \frac{-1}{2p}e^{-\frac{1}{2p}} &= \frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)} \\ \frac{-1}{2p} &= \mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right) \end{aligned}$$

Hence,

$$p = \left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$$

and

$$d = \ln\left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right) + \ln(w)$$

Define

$$d_{\text{SHEL}}(w) = \ln\left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right) + \ln(w)$$

Table 4.1 presents the values of $d_{\text{SHEL}}(w)$ and the corresponding $p = \exp\left(-d_{\text{SHEL}}(w)\right)$ for selected values of w.

w	$d_{\text{SHEL}}(w)$	p
0.1	0.71	0.49
0.2	0.74	0.48
0.3	0.77	0.46
0.5	0.86	0.42

Table 4.1: $d_{\text{SHEL}}(w)$ and its corresponding $p = \exp\left(-d_{\text{SHEL}}(w)\right)$ for selected values of w. The numbers in the second and third column are rounded to two decimal places.

Figure 4.11 illustrates $d_{\text{SHEL}}(w)$ for $0 < w \leq 1$. In addition, we added a red line showing where $d_{\text{SHEL}}(w)$ would equal w in order to highlight the conclusion that $d_{\text{SHEL}}(w) > w$ for all values of w.

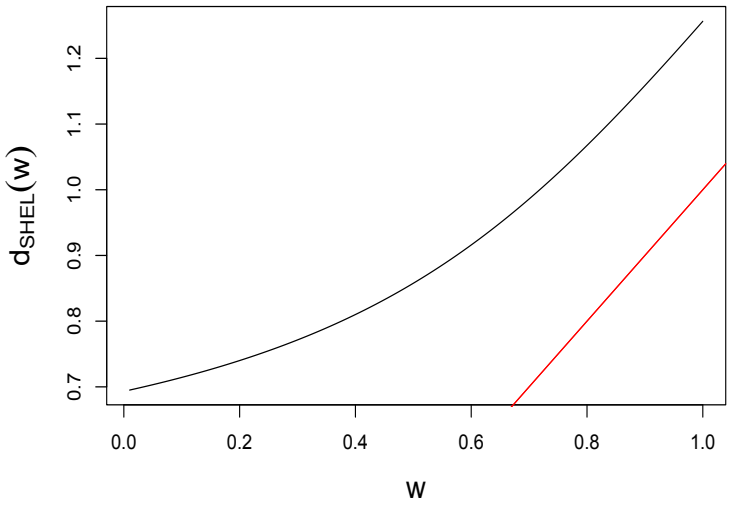


Figure 4.11: Plot of $d_{\text{SHEL}}(w)$. The red line is the diagonal line highlighting where $d_{\text{SHEL}}(w) = w$.

As shown in Figure 4.11, $d_{\text{SHEL}}(w) > w$. Thus $W(w, d, s) < 1$ for $d \geq d_{\text{SHEL}}(w)$ and small s.

$W\left(w, d_{\text{SHEL}}(w), s\right)$ is plotted in Figure 4.12 for different values of w.

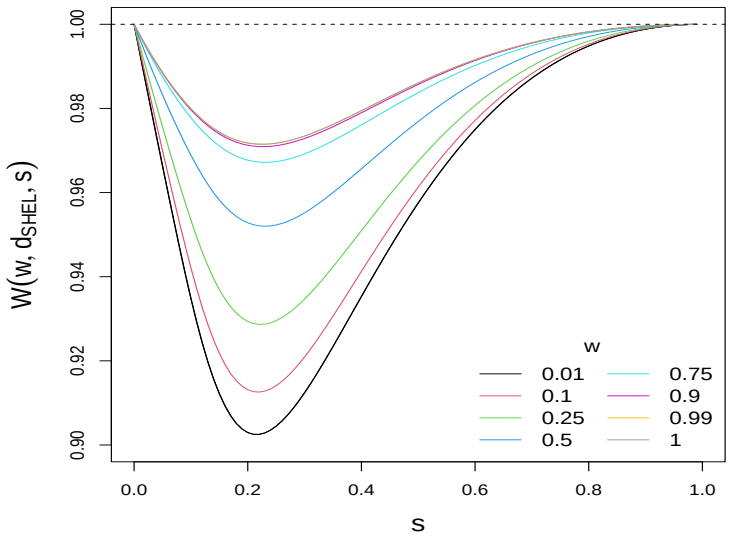


Figure 4.12: Plot of $W(w, d_{\text{SHEL}}(w), s)$ for different values of w. The horizontal dashed line highlights that $W(w, d_{\text{SHEL}}(w), s) \leq 1$.

Figure 4.12 suggests that $d = d_{\text{SHEL}}(w)$ is sufficient to make $W(w, d, s) \leq 1$ for $s \in (0, 1)$.

Now we will show how $W(w, d_{\text{SHEL}}(w), s)$ behaves when s approaches one as follows:

When s approaches one, $W(w, d, s)$ in Equation 4.11.9 can be rewritten as

$$W(w, d, 1-x) = w\left(\frac{e^{-d(1-x)} - (1-x)e^{\frac{-d}{(1-x)}}}{x}\right) + (1-w)\left(e^{-d(1-x)} + e^{-\frac{d}{(1-x)}}\right) \tag{4.11.12}$$

Define

$$f(x) = e^{-d(1-x)} - (1-x)e^{\frac{-d}{(1-x)}}$$

By using Equation 4.10.12, the Maclaurin series expansion of $f(x)$ is

$$f(x) \approx (1+2d)e^{-d}x + d^2\left(2d-3\right)e^{-d}x^3/6$$

The Maclaurin series expansion of $(1-w)\left(e^{-d(1-x)} + e^{-\frac{d}{(1-x)}}\right)$ in Equation 4.11.12 is

$$(1-w)e^{-d}\left(2 + d(d-1)x^2\right)$$

Therefore, the Maclaurin series expansion of $W(w, d, 1-x)$ for small x is

$$\begin{aligned} W(w, d, 1-x) &\approx e^{-d}\left(w\left((1+2d) + d^2\left(2d-3\right)x^2/6\right) + (1-w)\left(2 + d(d-1)x^2\right)\right) \\ &= e^{-d}\left(w + 2wd + 2(1-w) + w\frac{d^2\left(2d-3\right)}{6}x^2 + (1-w)d(d-1)x^2\right) \end{aligned}$$

Hence,

$$W(w, d, 1-x) \approx we^{-d}\left(\frac{2}{w} - 1 + 2d\right) + e^{-d}\left(w\frac{d^2\left(2d-3\right)}{6} + (1-w)d(d-1)\right)x^2 \tag{4.11.13}$$

Equation 4.11.13 is less than $we^{-d}\left(\frac{2}{w} - 1 + 2d\right)$ if and only if the coefficient of x^2 is negative. Define

$$d_o(w) = w\frac{d_{\text{SHEL}}^2(w)\left(2d_{\text{SHEL}}(w) - 3\right)}{6} + (1-w)d_{\text{SHEL}}(w)(d_{\text{SHEL}}(w) - 1)$$

It was difficult to see $d_o(w)$ is always negative when it is plotted. Therefore, $-d_o(w)$ will be plotted instead using a logarithmic axis in the following figure.

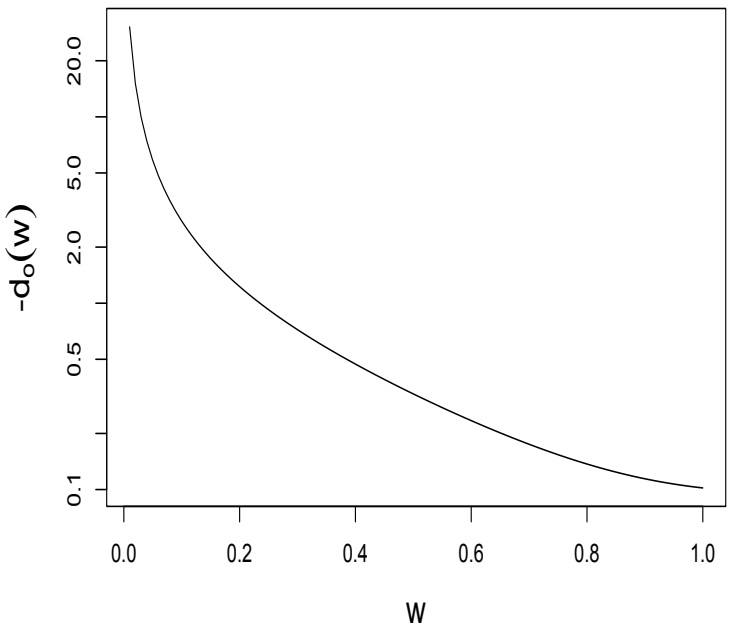


Figure 4.13: Plot of $-d_o(w)$.

Figure 4.13 shows that $-d_0(w)$ is always positive, thus $d_0(w)$ is always negative for $w \in (0,1)$. Hence, $W(w, d_{\text{SHEL}}(w), 1-x)$ is less than $we^{-d}\left(\frac{2}{w} - 1 + 2d\right)$ for s approaches 1.

Proposition 4.11.1 explores the behaviour of $W(w,d,s)$ as d increases. In addition, it will be used in Theorem 4.11.2.

Proposition 4.11.1 $W(w,d,s)$, which represents the ratio of $\bar{F}_{Z_1+Z_2}(x_1'+x_2')$ to p, is a decreasing function in d for

$$\begin{cases} d > 0 & ;\text{ if } \ s+w \leq 1 \\ d > \dfrac{-\ln(s)+\ln\left(\frac{1-s^{-1}(1-w)}{1-s(1-w)}\right)}{2\sinh(-\ln(s))} & ;\text{ if } \ s+w > 1 \end{cases}$$

Proof. Recall Equation 4.11.9

$$W(w,d,s) \ = \ w\left(\frac{e^{-ds} - se^{-\frac{d}{s}}}{1-s}\right) + (1-w)\left(e^{-ds} + e^{-\frac{d}{s}}\right)$$

By differentiating with respect to d, thus

$$\begin{aligned}\frac{\partial}{\partial d}W(w,d,s) &= w\left(\frac{e^{-\frac{d}{s}}-se^{-ds}}{1-s}\right)+(1-w)\left(-se^{-ds}-\frac{1}{s}e^{-\frac{d}{s}}\right)\\ &= -s^2e^{-ds}\left(\frac{w+(1-s)(1-w)}{s(1-s)}\right)\\ &+ e^{-\frac{d}{s}}\left(\frac{sw-(1-s)(1-w)}{s(1-s)}\right)\end{aligned} \tag{4.11.14}$$

The sign of Equation 4.11.14 depends on the second part of the right hand side of the equation, specifically on the quantity $sw-(1-s)(1-w)=s+w-1$. Hence, two cases should be considered:

Case A: If $s+w\leq 1$, then $e^{-\frac{d}{s}}\left(\frac{sw-(1-s)(1-w)}{s(1-s)}\right)\leq 0$ and $\frac{\partial}{\partial d}W(w,d,s)<0$. Therefore, $W(w,d,s)$ is a monotonically decreasing function in $d>0$ and $s+w\leq 1$.

Case B: If $s+w>1$, then we will argue as follows:

Set $s=e^{-a};\ a>0$, Equations 4.11.9 and 4.11.14 respectively become

$$W(w,d,e^{-a}) = w\left(\frac{e^{-de^{-a}}-e^{-a}e^{-de^{a}}}{1-e^{-a}}\right)+(1-w)\left(e^{-de^{-a}}+e^{-de^{a}}\right)$$

and

$$\begin{aligned}\frac{\partial}{\partial d}W(w,d,e^{-a}) &= w\left(\frac{e^{-de^{a}}-e^{-a}e^{-de^{-a}}}{1-e^{-a}}\right)+(1-w)\left(-e^{-a}e^{-de^{-a}}-e^{a}e^{-de^{a}}\right)\\ &= \frac{w\left(e^{-de^{a}}-e^{-a}e^{-de^{-a}}\right)-(1-w)(1-e^{-a})\left(e^{-a}e^{-de^{-a}}+e^{a}e^{-de^{a}}\right)}{1-e^{-a}}\end{aligned} \tag{4.11.15}$$

For any $a>0$ and $d=0$,

$$W(w,0,e^{-a})=2-w$$

and

$$\begin{aligned}\frac{\partial}{\partial d}W(w,0,e^{-a}) &= w-(1-w)\left(e^{-a}+e^{a}\right)\\ &= w-2(1-w)\cosh(a)\end{aligned}$$

Hence, $\frac{\partial}{\partial d}W(w,0,e^{-a})>0 \iff w>2(1-w)\cosh(a)$.

Moreover, the denominator of Equation 4.11.15 is always positive, and for sufficiently large d the numerator is negative. Therefore, $\frac{\partial}{\partial d}W(w,d,e^{-a})$ is negative for

sufficiently large d. In addition, by continuity of $\frac{\partial}{\partial d}W(w,d,e^{-a})$, if there is a single $d^* > 0$ such that $\frac{\partial}{\partial d}W(w,d^*,e^{-a}) = 0$, then $W(w,d,e^{-a})$ increases for $d < d^*$ and decreases for $d > d^*$. Now, for $e^{-a} > 1-w$, Equate Equation 4.11.15 to zero,

$$\frac{\partial}{\partial d}W(w,d,e^{-a}) = 0$$

$$\iff$$

$$w\left(e^{-de^a} - e^{-a}e^{-de^{-a}}\right) - (1-w)(1-e^{-a})\left(e^{-a}e^{-de^{-a}} + e^{a}e^{-de^{a}}\right) = 0$$

$$\iff$$

$$w\left(e^{-de^a} + (1-e^{-a})\left(e^{-a}e^{-de^{-a}} + e^{a}e^{-de^{a}}\right)\right)$$

$$=$$

$$\left(we^{-a}e^{-de^{-a}} + (1-e^{-a})\left(e^{-a}e^{-de^{-a}} + e^{a}e^{-de^{a}}\right)\right)$$

$$\iff$$

$$(1-e^a(1-w))\,e^{-de^a} = e^{-a}\,(1-e^{-a}(1-w))\,e^{-de^{-a}}$$

$$\iff$$

$$\ln\left((1-e^a(1-w))\,e^{-de^a}\right) = \ln\left(e^{-a}\,(1-e^{-a}(1-w))\,e^{-de^{-a}}\right)$$

$$\iff$$

$$\ln\left(1-e^a(1-w)\right) - de^a + a - \ln\left(1-e^{-a}(1-w)\right) + de^{-a} = 0$$

$$\iff$$

$$d = \frac{a+\ln\left(\frac{1-e^a(1-w)}{1-e^{-a}(1-w)}\right)}{e^a-e^{-a}}$$

$$\iff$$

$$d = \frac{a+\ln\left(\frac{1-e^a(1-w)}{1-e^{-a}(1-w)}\right)}{2\sinh(a)}$$

Define

$$d^*(w,a) = \frac{a+\ln\left(\frac{1-e^a(1-w)}{1-e^{-a}(1-w)}\right)}{2\sinh(a)} \tag{4.11.16}$$

Thus $d^*(w,a)$ is unique. Hence, $W(w,d,s)$ is a decreasing function in d for

$$\begin{cases} d > 0 & ;\text{ if } s+w \le 1 \\ d > \frac{-\ln(s)+\ln\left(\frac{1-s^{-1}(1-w)}{1-s(1-w)}\right)}{2\sinh(-\ln(s))} & ;\text{ if } s+w > 1 \end{cases}$$

□

Moreover, $a/\sinh(a)$ and $\frac{\ln\left(\frac{1-e^a(1-w)}{1-e^{-a}(1-w)}\right)}{\sinh(a)}$, in Equation 4.11.16, are both decreasing

functions in $a > 0$, therefore, the largest value of $d^*(w, a)$ is at $a = 0$. Define

$$d_U^*(w) = \lim_{a \to 0} d^*(w, a)$$

Thus

$$d_U^*(w) = \lim_{a \to 0} \frac{a + \ln\left(\frac{1-e^a(1-w)}{1-e^{-a}(1-w)}\right)}{2\sinh(a)} = \frac{0}{0}$$

By using L'Hospital's rule,

$$\begin{aligned} d_U^*(w) &= \lim_{a \to 0} \frac{1 + \left(\frac{1-e^{-a}(1-w)}{1-e^a(1-w)}\right) \frac{-(1-w)e^a\left(1-e^{-a}(1-w)\right)-(1-w)e^{-a}(1-e^a(1-w))}{(1-e^{-a}(1-w))^2}}{2\cosh(a)} \\ &= 3/2 - 1/w \end{aligned}$$

We plot $d_{\text{SHEL}}(w) - d_U^*(w)$ in Figure 4.14 using a logarithmic axis.

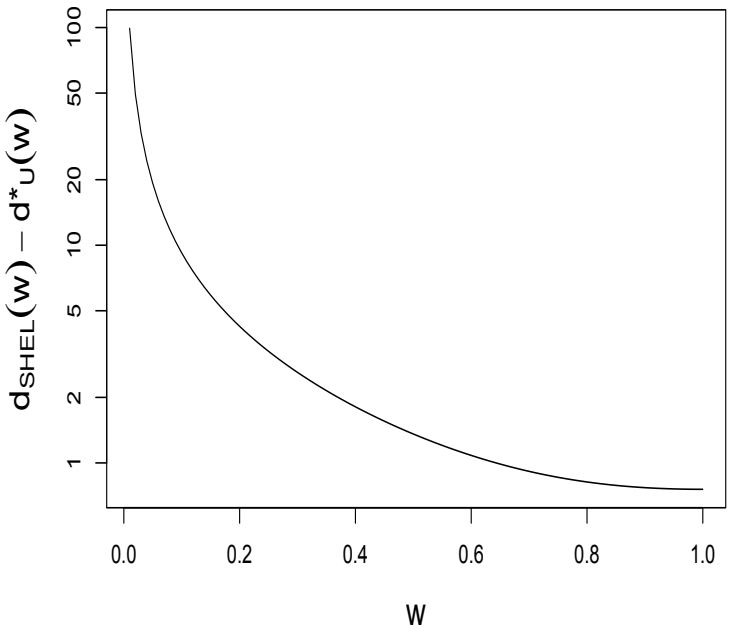


Figure 4.14: Plot of $d_{\text{SHEL}}(w) - d_U^*(w)$.

Figure 4.14 shows that $d_{\text{SHEL}}(w) > d_U^*(w)$. As a result, the derivative of $W(w, d, s)$ with respect to d for $d > d_{\text{SHEL}}(w)$ is negative for all s. □

In principle, based on the examples shown in Figure 4.12 for $d = d_{\text{SHEL}}(w)$ and Proposition 4.11.1 which shows what happens as d increases for fixed w, we believe

that the following theorem is true. However, there is currently a gap in completing the proof which is to show that $W(w, d_{\text{SHEL}}(w), s) \leq 1$ for all w and s.

Theorem 4.11.2 Let $Z_i \sim \text{SHEL}(w, c_i, \lambda_i);\ i = 1, 2$ be independent. Moreover, let $\lambda_1 \leq \lambda_2$ and x_i' be such that $\bar{F}_{Z_i}(x_i') = p < w$. If $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$, then

$$P(Z_1 + Z_2 \geq x_1' + x_2') \leq p$$

The idea of proof. $\bar{F}_{Z_1+Z_2}(x_1' + x_2') \leq p \iff W(w, d, s) \leq 1$. From the discussion above, in particular Figure 4.12, when $d = d_{\text{SHEL}}(w)$, it looks like $W(w, d_{\text{SHEL}}(w), s) \leq 1$. Moreover, by using Proposition 4.11.1, if $d \geq d_{\text{SHEL}}(w)$, i.e. $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$, then

$$P\left(Z_1 + Z_2 \geq x_1' + x_2'\right) \leq p$$

□

Theorem 4.11.2 is interesting in its own right since SHEL distributions are commonly encountered in the modeling of property/casualty claim processes (Bahnemann, 2015). Moreover, assuming that Theorem 4.11.2 is correct, we can use stochastic dominance properties to extend easily the validation of the combined tails dominance property to distributions that have partial log-concave right tail probability functions, as stated in Theorem 4.11.3.

Theorem 4.11.3 Let $U_i;\ i = 1, 2$ be independent and have log-concave right tail probability function for $x_i > c_i'$, and define $w = \bar{F}_{U_i}(c_i')$. Moreover, let $p < w$ and x_i' be such that $\bar{F}_{U_i}(x_i') = p$. If $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$. Then

$$P\big(U_1 + U_2 \geq x_1' + x_2'\big) \leq p$$

Proof. Firstly, from Theorem 4.9.1, there exists $Z_i \sim \text{SHEL}(w, c_i, \lambda_1)$, c_i be such that $p = \bar{F}_{Z_i}(x_i')$. Moreover, Z_i are first-order stochastically dominating U_i. Secondly, Proposition 4.8.2 implies that $U_1 + U_2$ is first-order stochastically dominated by $Z_1 + Z_2$, i.e.

$$P\left(U_1 + U_2 \geq x_1' + x_2'\right) \leq P\left(Z_1 + Z_2 \geq x_1' + x_2'\right)$$

Finally by using Theorem 4.11.2, thus

$$\begin{aligned} P\big(U_1 + U_2 \geq x_1' + x_2'\big) &\leq P\big(Z_1 + Z_2 \geq x_1' + x_2'\big) \\ &\leq p \end{aligned}$$

□

For the RVFV example in Section 4.2, if the experts are interested in the upper parts of the right tail probability functions of those uncertain quantities, then Theorem 4.11.3 can be applied since their distributions have partial log-concave right tail probability functions.

4.12 Re-Scaling and Shifting

The stability of the combined tails dominance property is examined under changing the location or re-scaling either of the distributions involved in the following propositions.

Proposition 4.12.1 If the property holds for F and G and tail probability p, then it holds for F^* and G, where $F^*(x) = F(x + k)\ \forall x$ and $\forall k \in \mathbb{R}$.

Proof. $P(X + k + Y \geq x + k + y) = P(X + Y \geq x + y) < p$ □

Proposition 4.12.2 If the property holds for F and G and tail probability p, then it does not necessarily hold for F^* and G, where $F^*(x) = F(kx)\ \forall k > 1$.

Proof. Assume $X_1 \sim F \equiv \text{SHE}(0, \lambda_1)$ and $X_2 \sim G \equiv \text{SHE}(c_2, \lambda_2)$ such that $s = \frac{\lambda_1}{\lambda_2} = 0.2789$

Figure 4.15 illustrates that when $d = 1.15$ which is equivalent to $p = 0.3166$, the property holds for $s \leq 0.2789$, whereas does not hold for $X_1^* \sim F^* \equiv \text{SHE}(0, k\lambda_1)$ for all $k > 1$ and G, where $s > 0.2789$ with the same value of p.

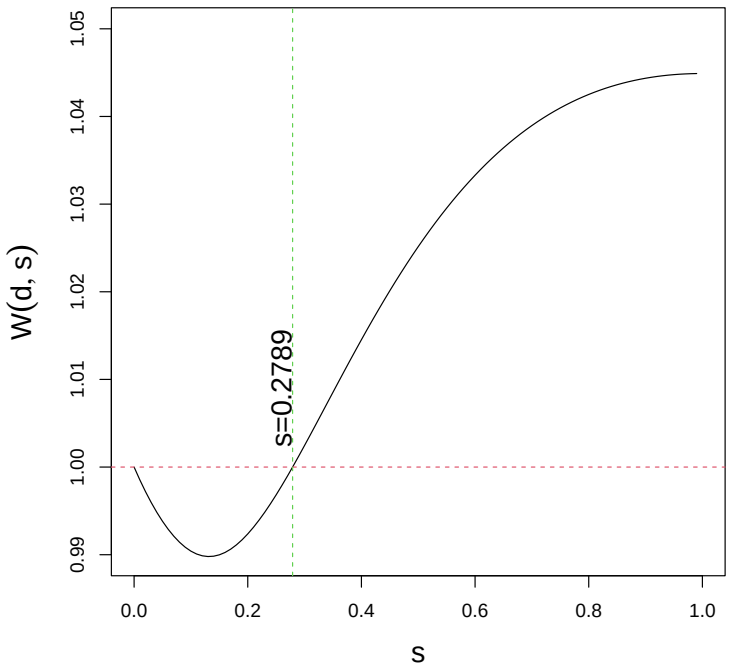


Figure 4.15: Plot of $W(1.15, s)$. The horizontal dashed line highlights where $W(d, s) = 1$. The vertical dashed line is at $s = 0.2789$, where $W(1.15, s) \leq 1$ for $0 \leq s \leq 0.2789$.

4.13 Conclusions

The validation of the combined tails dominance property has been investigated by assuming different distributions. We found that the property holds for all distributions with log-concave symmetric densities. Moreover, it holds for negatively-skewed Azzalini-style skew-symmetric distributions with log-concave kernels when two quantities are involved. In addition, we considered distributions that have log-concave right tail probability functions. In such cases, p, which represents the individual tail probabilities, needs to be less than or equal to approximately 0.2846681 to lead to the validation of the combined tails dominance property. Furthermore, distributions that have partial log-concave right tail probability functions are considered, and we found that $p \leq \left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$ is needed when two quantities are involved, where w is the probability of being in the continuous state for the

mixture of shifted exponential and lump used to dominate distribution having partial log-concave right tail probability function. The stability of the combined tails dominance property is preserved under changing the location of either of the distributions involved. However, re-scaling one of them does not guarantee preserving the property.

Chapter 5

Conclusions and Future Research Directions

In this thesis, we have presented two statistical problems in risk assessment. These problems emerge because of different sources of uncertainty that are quantified probabilistically.

The first problem relates to the applied question of what to do in ecotoxicological risk assessment if more species are tested than required. Specifically, we aimed to show theoretically certain key dominance properties for the arithmetic mean as the number of tested species increases. The dominance properties for the arithmetic mean hold when the mean fraction exceeded and the probability that the fraction exceeded exceeds a specific proportion decrease as the number of tested species increases.

We have shown mathematically that the dominance properties for the arithmetic mean hold for all distributions with symmetric log-concave densities as the number of tested species increases. This result motivated us to examine the mixture of such distributions. We have presented results for two-component scale and location mixtures of normal distributions. The findings show that the dominance properties sometimes hold. However, our investigation was based primarily on increasing the sample size from one observation to two observations.

The second problem relates to combining limited probabilistic expert judgements on multiple quantities in order to provide limited probabilistic information about a

derived quantity. In this study, we first proposed a working hypothesis, which we called the combined tails dominance property, that simplifies calculations for the derived quantity. The goal of the research is to understand when the combined tails dominance property is valid.

We have shown mathematically that the combined tails dominance property is valid for all distributions with symmetric log-concave densities. In addition, it holds for negatively-skewed Azzalini-style skew-symmetric distributions with log-concave kernels when two quantities are involved. Furthermore, we enlarged our zone of investigation by considering distributions that have log-concave right tail probability functions. This investigation has been extended to the situation where the log-concavity only applies to part of the right tail probability function of the distribution. We found that the combined tails dominance property holds for all distributions that have log-concave right tail probability function when the individual tail probabilities are less than or equal to 0.2846681. Moreover, it is valid when the distributions have partial log-concave right tail probability functions, when two quantities are involved and when the individual tail probabilities are less than or equal to $\left(-2\mathcal{W}_{-1}\left(\frac{-1}{2w}e^{-\left(\frac{1}{w}-\frac{1}{2}\right)}\right)\right)^{-1}$, where w is the probability of being in the continuous state for the mixture of shifted exponential and lump used to dominate distribution having partial log-concave right tail probability function. The combined tails dominance property is stable under changing the location of either of the two distributions, whereas re-scaling either of them does not guarantee the stability of the property.

These findings suggest that further research could be undertaken in this area since there is a connection between the two problems, for example what would happen for the dominance of the arithmetic mean properties if the LSSD has log-concave right tail probability function or has log-concave cumulative distribution function. We have presented a unified approach for the second problem if the distribution has log-concave right tail probability function, and so the questions arises whether we can exploit those results in order to extend our findings in Chapter 3 to cover all distributions with log-concave right tail probability functions or log-concave cumulative distribution functions. The reason for this suggestion is that when the

combined tails dominance property holds for any distribution, it might be possible to prove that the dominance of the arithmetic mean properties holds as well. For example, for distributions that have symmetric log-concave densities, we proved that the combined tails dominance property and the dominance properties for the arithmetic mean hold in Sections 3.6 and 4.4 respectively. Moreover, in Section 4.5, the combined tails dominance property holds exactly for Cauchy distribution. In addition, MFE_n and $\text{PFE}_n(\alpha)$ are constant if the LSSD is Cauchy, since its characteristic function is $\psi_X(t) = \exp(-|t|), t \in \mathbb{R}$, therefore the characteristic function of $Y_n = \bar{X} - c$ is

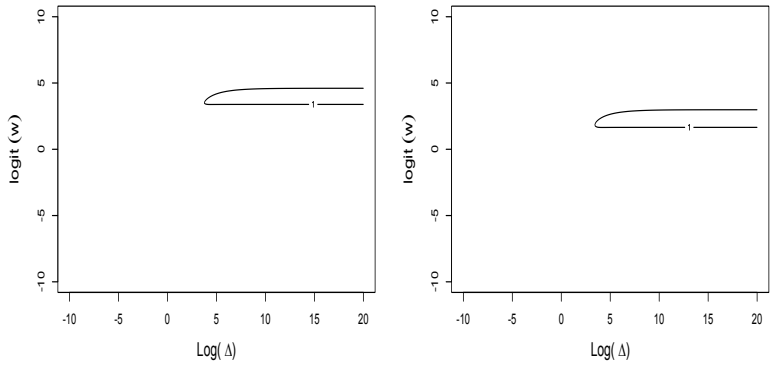
$$\begin{aligned} \psi_{Y_n}(t) &= \exp(-cit)\left(\exp\left(-\left|\frac{t}{n}\right|\right)\right)^n \\ &= \exp(-cit - |t|) \end{aligned}$$

Hence, Y_n has Cauchy distribution with probability density and cumulative distribution functions as follows:

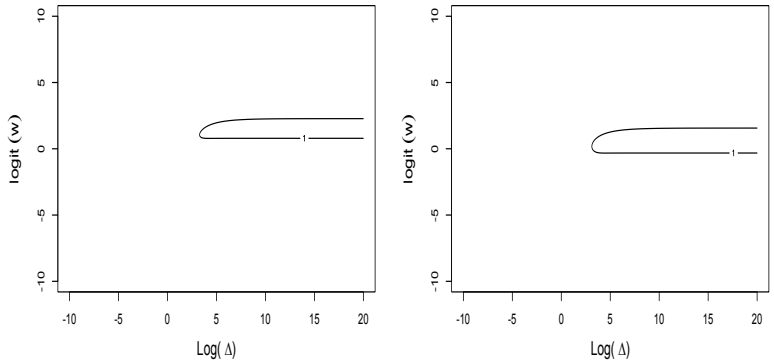
$$f_{Y_n}(y) = \frac{1}{\pi\left(1 + (y+c)^2\right)} \text{ and } F_{Y_n}(y) = \frac{1}{2} + \frac{1}{\pi}\arctan(y+c);\ y \in \mathbb{R}$$

It is clear that the distribution of Y_n does not depend on n, as a result MFE_n and $\text{PFE}_n(\alpha)$ do not change as the sample size, n, increases.

Another suggestion for further research is that the Cauchy distribution is conventionally considered to be longer-tailed than any mixture of normals, and yet MFE_n and $\text{PFE}_n(\alpha)$ do not always decrease as n increases from one to two observations for extreme mixtures of normals. Figure 5.1 shows exactly where the combined tails dominance property holds for SMN distributions for selected values of the individual tail probabilities $p_1 = p_2 = p$.



(a) $p = 0.01$ (b) $p = 0.05$



(c) $p = 0.1$ (d) $p = 0.2$

Figure 5.1: Plots of the unit contour of the ratio of $P\left(X_1 + X_2 \geq x_1 + x_2\right)$ to p with respect to $\log(\Delta)$ (x-axis) and $\log\left(\frac{w}{1-w}\right)$ (y-axis) for the SMN$(w, 0, 1, \Delta)$ distribution. Each of the four panels is related to a specific value of p.

From Figure 5.1, it is clear that the combined tails dominance property does not always hold for extreme mixtures of normals. So the question is what precludes the dominance properties for the arithmetic mean for mixtures of normals even though they have short tails and allows them for Cauchy distributions even though they have long tails?. We suspect that the answer has something to do with sharp locally convex bends in the log of the right tail probability function, at least for the second problem.

It is not so clear that experts will be able to really judge (partial) log-concavity of the right tail probability function. There are possibly other criteria that might be

easier for them to judge that would lead to the combined tails dominance property.

Finally, in the work in this thesis on the second problem, we are basically looking for $P(X_1 + X_2 \geq x_1 + x_2) \leq \max(P(X_1 \geq x_1), P(X_2 \geq x_2))$ as a substitute for the worst case $P(X_1 \geq x_1) + P(X_2 \geq x_2)$. It might be easier if we relaxed our criterion a little. There might be some criterion in between our current working hypothesis and the worst case that would still be useful and which would follow from something quite easy for experts to judge. Another area worth investigating is what happens for functions of X_1 and X_2 other than $X_1 + X_2$.

Acronyms

AEC	Acceptable Environmental Concentration
AF	Assessment Factor
ATS	Adjusted Toxicity Statistic
EKE	Expert Knowledge Elicitation
erf	Error function
f	probability density function of the random variables in the subscript
F	cumulative distribution function of the random variables in the subscript
FE_n	Fraction Exceeded, the subscript n is the number of tested species
FOSD	First Order Stochastic Dominance
L	Lump
LMN	Location Mixture of Normals
LSSD	SSD on log-scale (base 10)
MFE/MFE_n	Mean Fraction Exceeded, the subscript n is the number of tested species
MN	Mixture of Normals
P(E)	denotes the probability of the event E
PEC	Predicted Environmental Concentration
PFE/$\mathrm{PFE}_n(\alpha)$	Probability Fraction Exceeded exceeds a chosen α, the subscript n is the number of tested species
PNEC	Predicted No Effect Concentration
$\mathbb{R}$	denotes the space of real numbers
SHE	Shifted Exponential
SHEL	Mixture of shifted exponential and Lump
SMN	Scale Mixture of Normals
SSD	Species Sensitivity Distribution

Appendix A

Basic and Auxiliary Results

A.1 Evaluation of Some Integrals

In this appendix, we give the technique to solve the following integral

$$I(\alpha, \beta, a, b) = \int_{-\infty}^{\infty} e^{-(\alpha t+\beta)^2} \text{erf } (at + b)\, dt;\ \alpha \neq 0,\ a \neq 0 \tag{A.1.1}$$

Proof. Let

$$z = \alpha t + \beta \tag{A.1.2}$$

Then

$$t = \frac{z - \beta}{\alpha};\ \alpha \neq 0 \tag{A.1.3}$$

and

$$dt = \frac{1}{\alpha}\, dz \tag{A.1.4}$$

Using Equation A.1.3, the quantity $at + b$ becomes

$$at + b = \frac{a}{\alpha} z - \frac{a\beta}{\alpha} + b \tag{A.1.5}$$

Use Equations A.1.2, A.1.4, and A.1.5, Equation A.1.1 becomes

$$I(\alpha, \beta, a, b) = \frac{1}{\alpha} \int_{-\infty}^{\infty} e^{-z^2} \text{erf } \left(\frac{a}{\alpha} z - \frac{a\beta}{\alpha} + b\right) dz;\ \alpha \neq 0,\ a \neq 0 \tag{A.1.6}$$

If $b = \frac{a\beta}{\alpha}$, then Equation A.1.6 becomes

$$I\left(\alpha, \beta, a, \frac{a\beta}{\alpha}\right) = \frac{1}{\alpha} \int_{-\infty}^{\infty} e^{-z^2} \text{erf } \left(\frac{a}{\alpha} z\right) dz;\ \alpha \neq 0,\ a \neq 0 \tag{A.1.7}$$

Since e^{-z^2} is even function and erf $\left(\frac{a}{\alpha}z\right)$ is odd function, $e^{-z^2}\text{erf}\,(\frac{a}{\alpha}z)$ is odd function. Thus Equation A.1.7 becomes

$$I\left(\alpha,\beta,a,\frac{a\beta}{\alpha}\right)=0 \tag{A.1.8}$$

To calculate $I(\alpha,\beta,a,b)$, the technique of taking the derivative under the integral sign is used.

Note the error function is defined as

$$\begin{aligned}\text{erf}\,(x) &= \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt\\ &= 2\Phi(x\sqrt{2})-1\end{aligned}$$

The derivative of erf (z) with respect to x is

$$\frac{\partial}{\partial x}\text{erf}\,(z)=\frac{2}{\sqrt{\pi}}e^{-z^2}\frac{\partial}{\partial x}z$$

Differentiating Equation A.1.1 with respect to b, we get

$$\begin{aligned}\frac{\partial I(\alpha,\beta,a,b)}{\partial b} &= \frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty}e^{-(\alpha t+\beta)^2}e^{-(at+b)^2}\,dt\\ &= \frac{2}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp\left(-(\alpha^2+a^2)t^2-2(\alpha\beta+ab)t-\left(\beta^2+b^2\right)\right)\,dt\end{aligned}$$

Substituting $u=\sqrt{\alpha^2+a^2}t$, thus we obtain

$$\begin{aligned}\frac{\partial I(\alpha,\beta,a,b)}{\partial b} &= \frac{2}{\sqrt{\pi}}\frac{\exp\left(-\left(\beta^2+b^2\right)+\frac{(\alpha\beta+ab)^2}{\alpha^2+a^2}\right)}{\sqrt{\alpha^2+a^2}}\int_{-\infty}^{\infty}\exp\left(-\left(u+\frac{\alpha\beta+ab}{\sqrt{\alpha^2+a^2}}\right)^2\right)\,du\\ &= 2\frac{\exp\left(-\left(\beta^2+b^2\right)+\frac{(\alpha\beta+ab)^2}{\alpha^2+a^2}\right)}{\sqrt{\alpha^2+a^2}}\\ &= 2\frac{\exp\left(\frac{-(\alpha b-\beta a)^2}{\alpha^2+a^2}\right)}{\sqrt{\alpha^2+a^2}}\end{aligned}$$

since the integral is Gaussian.

Integrate the latter equation from $u=\frac{a\beta}{\alpha}$ to $u=b$. Thus we get

$$\int_{\frac{a\beta}{\alpha}}^{b}\frac{\partial I(\alpha,\beta,a,u)}{\partial u}\,du=I(\alpha,\beta,a,b)-I\left(\alpha,\beta,a,\frac{a\beta}{\alpha}\right)$$

Using Equation A.1.8. Thus

$$I(\alpha,\beta,a,b)=\frac{2}{\sqrt{\alpha^2+a^2}}\int_{\frac{a\beta}{\alpha}}^{b}\exp\left(\frac{-\left(\alpha u-\beta a\right)^2}{\alpha^2+a^2}\right)\,du$$

Hence,

$$I(\alpha,\beta,a,b)=\frac{2}{\sqrt{\alpha^2+a^2}}\int_0^{\frac{\alpha b-\beta a}{\sqrt{\alpha^2+a^2}}}\frac{\sqrt{\alpha^2+a^2}}{\alpha}e^{-w^2}\,dw=\frac{\sqrt{\pi}}{\alpha}\text{erf}\left(\frac{\alpha b-\beta a}{\sqrt{\alpha^2+a^2}}\right)$$

□

A.2 Lambert $\mathcal{W}$ Function

The Lambert $\mathcal{W}$ function is the inverse of the function

$$f(x)=xe^x$$

Satisfying

$$\mathcal{W}(x)e^{\mathcal{W}(x)}=x \tag{A.2.9}$$

The Lambert $\mathcal{W}$ function is a multivalued function denoted by $\mathcal{W}_k(x)$. If $x<e^{-1}$ is real, then $\mathcal{W}(x)$ is complex. If $-e^{-1}\leq x<0$ is real, then there are two real branches $\mathcal{W}_0(x)$ and $\mathcal{W}_{-1}(x)$. $\mathcal{W}_{-1}(x)$ is the bottom branch, and it is equivalent to $\mathcal{W}(x)\leq -1;\ x\in[-e^{-1},0)$. $\mathcal{W}_0(x)$ is the upper branch (*principal branch*), witch is equivalent to $\mathcal{W}(x)\geq -1;\ x\in[-e^{-1},\infty)$, where a branching point located at $(-e^{-1},-1)$, If $x\geq 0$ is real, then there is a single real value which belongs to $\mathcal{W}_0(x)$, see Corless et al. (1996).

The Lambert $\mathcal{W}$ function solves equations that have the form

$$ay+b\ln(y)+c=0; a\neq 0,\ b\neq 0 \tag{A.2.10}$$

Define $r=\ln(y)$, Equation A.2.10 becomes

$$ae^r+br+c=0 \tag{A.2.11}$$

Bibliography

Abramowitz, M., and Stegun, I. A. Eds. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. (9th ed.). New York: Dover.

Aldenberg, T., and Jaworska, J. S. (2000). Uncertainty of the Hazardous Concentration and Fraction Affected for Normal Species Sensitivity Distributions. *Ecotoxicology and Environmental Safety*, *46*: 1-18.

Aldenberg, T., Jaworska, J. S., and Traas, T. P. (2002). Normal Species Sensitivity Distributions and Probabilistic Ecological Risk Assessment. In: Posthuma, L., II Suter, G. W. and Traas, T. P. Eds. *Species Sensitivity Distributions in Ecotoxicology*. Boca Raton: Lewis Publishers, 49-102.

An, M. Y. (1998). Log-Concavity Versus Log-Convexity: A Complete Characterization. *Journal of Economic Theory*, *80*: 350-369. DOI: `https://doi.org/10.1006/jeth.1998.2400`

Azzalini, A. (1985). A Class of Distributions Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, *12*: 171-178. `http://www.jstor.org/stable/4615982`

Azzalini, A., and Capitanio, A. (2003). Distributions Generated by Perturbation of Symmetry with Emphasis on a Multivariate Skew t Distribution. *Journal of the Royal Statistical Society*, *65*(2): 367-389. `http://www.jstor.org/stable/3647510`

Azzalini, A., and Regoli, G. (2012). Some Properties of Skew-Symmetric Distributions. *Annals of the Institute of Statistical Mathematics*, *64*: 857-879. DOI: `https://doi-org.ezphost.dur.ac.uk/10.1007/s10463-011-0338-5`

Bagnoli, M., and Bergstrom, T. (2005). Log-concave Probability and Its Applications. *Economic Theory*, *26*: 445-469. DOI: https://doi-org.ezphost.dur.ac.uk/10.1007/s00199-004-0514-4

Bahnemann, D. (2015). Distributions for Actuaries. CAS Monograph Series, Casualty Actuarial Society.

Birnbaum, Z. W. (1948). On Random Variables with Comparable Peakedness. *The Annals of Mathematical Statistics*, *19*: 76-81. DOI: http://www.jstor.org/stable/2236059

Blumberg, H. (1919). On convex functions. *Transactions of the American Mathematical Society*, *20*(1): 40-44. https://doi.org/10.2307/1988982

Brito, P. B., Fabiao, F. and Staubyn, A. (2008). Euler, Lambert, and the Lambert W-function Today. *Mathematical Scientist*, *33*: 127-133.

Burgman, M. (2005). Experts, Stakeholders and Elicitation. In: *Risks and Decisions for Conservation and Environmental Management* (Ecology, Biodiversity and Conservation, pp. 62-126). Cambridge, United Kingdom: Cambridge University Press.

Chan, W., Park, D. H., and Proschan, F. (1987). *Peakedness of Weighted Averages of Jointly Distributed Random Variables*. Report No. M712R. The Florida State University. Department of Statistics. https://apps.dtic.mil/dtic/tr/fulltext/u2/a162876.pdf

Chen, Y., and Samworth, R. J. (2013). Smoothed Log-concave Maximum Likelihood Estimation with Applications. *Statistica Sinica*, *23*(3): 1373-1398.

Cooke, R. (1991). *Experts in Uncertainty : Opinion and Subjective Probability in Science* (Environmental ethics and science policy series). New York, Oxford: Oxford University Press.

Corless, R. M., Gonnet, G. H., Hare, D. E., Jeffrey, D. J., and Knuth, D. E. (1996). On the Lambert W Function. *Advances in Computational Mathematics*, *5*: 329-359.

Daneshkhah, A., and Oakley, J. (2010). Eliciting Multivariate Probability Distributions. In: Böcker, K. Ed. *Rethinking Risk Measurement and Reporting: Volume I.* London, United Kingdom: Risk Books.

De Finetti, B. (1992). Foresight: Its Logical Laws, Its Subjective Sources. In: Kotz S., Johnson N.L. Eds. *Breakthroughs in Statistics*. Springer Series in Statistics (Perspectives in Statistics). New York, NY: Springer, 134-174. (Original work published 1964).

EC (European Commission) (2000). First Report of the EC Scientific Steering Committee's Working Group on Harmonisation of Risk Assessment Procedures in the Scientific Committees advising the European Commission in the area of human and environmental health. Part 2, Appendix 1- Glossary of Terms. 26-27 October 2000.

ECHA (European Chemicals Agency) (2012). Guidance for the Implementation of REACH: Guidance on Information Requirements and Chemical Safety Assessment Chapter R.19: Uncertainty Analysis. November 2012.

EFSA (European Food Safety Authority) (2005). Opinion of the Scientific Panel on Plant Health, Plant Protection Products and their Residues on a Request from EFSA Related to the Assessment of the Acute and Chronic Risk to Aquatic Organisms with Regard to the Possibility of Lowering the Uncertainty Factor if Additional Species were Tested. *The EFSA Journal*, *4*(1), 301: 1-45.

EFSA (European Food Safety Authority) (2008). Risk Assessment for Birds and Mammals-Revision of Guidance Document under Council Directive 91/414/EEC(SANCO/4145/2000-final of 25 September 2002)-Scientific Opinion of the Panel on Plant protection products and their Residues (PPR) on the Science behind the Guidance Document on Risk Assessment for birds and mammals. *The EFSA Journal*, *6*(7), 734: 1-181, Appendix 7. DOI: `https://doi.org/10.2903/j.efsa.2008.734`

EFSA (European Food Safety Authority) Scientific Committee (2012). Scientific

Opinion on Risk Assessment Terminology. *The EFSA Journal*, *10*(5), 2664: 1-43. DOI: https://doi.org/10.2903/j.efsa.2012.2664

EFSA (European Food Safety Authority)(2013). Technical meeting of the EFSA Scientific Network on EFSA Scientific Network for Risk Assessment in Animal Health and Welfare. *Risk of Introduction of Rift Valley Fever into the Southern Mediterranean Area through Undocumented Movement of Infected Animals*, 2013:EN-416: 1-24. DOI: https://doi.org/10.2903/sp.efsa.2013.EN-416

EFSA (European Food Safety Authority) (2014). Guidance on Expert Knowledge Elicitation in Food and Feed Safety Risk Assessment. *The EFSA Journal*, *12* (6), 3734: 1-278. DOI: https://doi.org/10.2903/j.efsa.2014.3734

EFSA (European Food Safety Authority) Scientific Committee (2018a). The principles and Methods Behind EFSA's Guidance on Uncertainty Analysis in Scientific Assessment. *The EFSA Journal*, *16*(1), 5122: 1-235. DOI: https://doi.org/10.2903/j.efsa.2018.5122

EFSA (European Food Safety Authority) Scientific Committee (2018b). Guidance on Uncertainty Analysis in Scientific Assessments. *The EFSA Journal*, *16*(1), 5123: 1-39. DOI: https://doi.org/10.2903/j.efsa.2018.5123

EFSA (European Food Safety Authority)(2018c). Conclusion on the Peer Review of the Pesticide Risk Assessment of the Active Substance Thiophanate-methyl. *The EFSA Journal*, *16*(1), 5133: 1-31. DOI: https://doi.org/10.2903/j.efsa.2018.5133

EFSA (European Food Safety Authority)(2018d). Conclusion on the Peer Review of the Pesticide Risk Assessment of the Active Substance Thiophanate-methyl. *The EFSA Journal*, *16*(1), 5133: 1-123, Appendix. DOI: https://doi.org/10.2903/j.efsa.2018.5133

Ferson, S. (2001). Probability Bounds Analysis Solves the Problem of Incomplete Specification in Probabilistic Risk and Safety Assessments. In: Haimes, Y. Y.,

Moser, D. A. and Stakhiv, E.Z. Eds. *Risk-Based Decision Making in Water Resources IX*. Reston, Virginia: American Society of Civil Engineers, 173-188. DOI: https://doi.org/10.1061/40577(306)16

Ferson, S., Kreinovich, V., Ginzburg, L. R., Sentz, K. and Myers, D. S. (2003). *Constructing Probability Boxes and Dempster-Shafer Structures*. Report No. SAND2002-4015. Albuquerque, New Mexico: Sandia National Laboratories.

Ferson, S., Nelsen R.B., Hajagos, J., Berleant, D.J., Zhang, J., Tucker, W.T., Ginzburg, L.R. and Oberkampf, W.L. (2004). *Dependence in Probabilistic Modeling, Dempster-Shafer Theory, and Probability Bounds Analysis*. Report No. SAND2004-3072. Albuquerque, New Mexico: Sandia National Laboratories.

Forbes, V. E. and Calow, P. (2002). Species Sensitivity Distributions Revisited: A Critical Appraisal. *Human and Ecological Risk Assessment*, *8*(3): 473-492. DOI: https://doi.org/10.1080/10807030290879781

Fox, D., Dam, R., Fisher, R., Batley, G., Tillmanns, A., Thorley, J., Schwarz, C., Spry, D. and McTavish, K. (2021). Recent Developments in Species Sensitivity Distribution Modeling. *Environmental Toxicology and Chemistry*, *40*(2): 293-308. C.J. Schwarz,h D.J. Spry,i and K. McTavishi

Garthwaite, P. H., Kadane, J. B.,and O'Hagan, A. (2005). Statistical Methods for Eliciting Probability Distributions. *Journal of the American Statistical Association*, *100*(470): 680-700. DOI: https://doi-org.ezphost.dur.ac.uk/10.1198/016214505000000105

Hadar, J., and Russell, W. (1969). Rules for Ordering Uncertain Prospects. *The American Economic Review*, *59*(1): 25-34. https://www.jstor.org/stable/1811090

Hanea, A. M., McBride, M. F., Burgman, M. A., and Wintle, B. C. (2018). Classical Meets Modern in the IDEA Protocol for Structured Expert Judgement. *Journal of Risk Research*, *21*(4): 417-433. DOI: 10.1080/13669877.2016.1215346

Hanoch, G., and Levy, H. (1969). The Efficiency Analysis of Choices Involving Risk. *The Review of Economic Studies*, *36*(3): 335-346. DOI: https://doi-org.ezphost.dur.ac.uk/10.2307/2296431

Hardy, G. H., Littlewood, J. E., and Pólya, G. (1934). *Inequalities*. Cambridge: Cambridge University Press.

Hickey, Graeme L., Craig, Peter S. and Hart, Andy. (2009). On the Application of Loss Functions in Determining Assessment Factors for Ecological Risk. *Ecotoxicology and Environmental Safety*, *72*(2): 293-300.

Ibragimov, R. (2007). Efficiency of Linear Estimators Under Heavy-tailedness: Convolutions of [alpha]-symmetric Distributions. *Econometric Theory*, *23*(3): 501-517. DOI: 10.1017/S0266466607070223

Kanter, M. (1977). Unimodality and Dominance for Symmetric Random Vectors. *Transactions of the American Mathematical Society*, *229*: 65-85. DOI: https://doi.org/10.1090/S0002-9947-1977-0445580-7

Kooijman, S. A. L. M. (1987). A Safety Factor for LC_{50} Values Allowing for Differences in Sensitivity Among Species. *Water Research*, *21*(3): 269-276. DOI: https://doi.org/10.1016/0043-1354(87)90205-3

Ma, C. (1998). On Peakedness of Distributions of Convex Combinations. *Journal of Statistical Planning and Inference*, *70*(1): 51-56. DOI: https://doi.org/10.1016/S0378-3758(97)00178-X

Marshall, A. W., Olkin, I., and Arnold, B. C. (2001). *Inequalities: Theory of Majorization and its Applications*. 2nd ed. New York: Springer Series in Statistics. (Original work published 1979)

Meeden, G. (1971). Finding Best Tests Approximately for Testing Hypotheses about a Random Parameter. *The Annals of Mathematical Statistics*, *42*(4): 1452-1454.

Mood, A. M., Graybill, F. A., and Boes, D.C. (1974). *Introduction to the Theory of Statistics*. (3rd ed.). McGraw-Hill.

Moore, R. E., Kearfott, R. B., and Cloud, M. J. (2009). *Introduction to Interval Analysis*. Philadelphia: Society for Industrial and Applied Mathematics. DOI 10.1137/1.9780898717716

Neumaier, A. (1990). *Interval Methods for Systems of Equations*. Cambridge: Cambridge University Press. DOI: https://doi-org.ezphost.dur.ac.uk/10.1017/CBO9780511526473

O'Hagan, A. (2012). Probabilistic Uncertainty Specification: Overview, Elaboration Techniques and their Application to a Mechanistic Model of Carbon Flux. *Environmental Modelling and Software*, *36*(C): 35-48. DOI: https://doi.org/10.1016/j.envsoft.2011.03.003

O'Hagan, A. (2019). Expert Knowledge Elicitation: Subjective but Scientific. *The American Statistician*, *73*(S 1): 69-81. DOI: https://doi.org/10.1080/00031305.2018.1518265

Posthuma, L., II Suter, G. W., and Traas, T. P. Eds. (2002). *Species Sensitivity Distributions in Ecotoxicology*. Boca Raton: Lewis Publishers.

Proschan, F. (1965). Peakedness of Distributions of Convex Combinations. *Annals of Mathematical Statistics*, *36*(5): 1703-1706. DOI: 10.1214/aoms/1177699798. https://projecteuclid.org/euclid.aoms/1177699798

R Development Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org

Ramsey, F. (1988). Truth and probability. In: Gärdenfors, P. and Sahlin, N. Eds. *Decision, Probability and Utility: Selected Readings*. Cambridge: Cambridge University Press, 19-47. DOI: https://doi-org.ezphost.dur.ac.uk/10.1017/CBO9780511609220

Renwick, A. G., Barlow, S. M., Hertz-Picciotto, I., Boobis, A. R., Dybing, E., Edler, L., Eisenbrandg,G., Greigh, J.B., Kleineri, J., Lambej, J., Müllerk, D.J.G., Smithl, M.R., Tritscherm, A., Tuijtelaarsi, S., van den Brandtn, P.A., Walker, R.,

and Kroes, R. (2003). Risk Characterisation of Chemicals in Food and Diet. *Food and Chemical Toxicology*, *41*(9): 1211-1271. DOI: https://doi.org/10.1016/S0278-6915(03)00064-4

Rowe G and Wright G, 1999. The Delphi Technique as a Forecasting Tool: Issues and Analysis. *International Journal of Forecasting*, *15*(4): 353-375.

Rice, J. A. (1995). *Mathematical Statistics and Data Analysis*. (2nd ed.). Belmont, CA: Duxbury Press.

Saumard, A., and Wellner, J. A. (2014). Log-concavity and Strong Log-concavity: A Review. *Statistics Surveys*, *8*: 45-114. DOI: 10.1214/14-SS107

Savage, L. J. (1972). *The Foundations of Statistics*. (2nd ed.). New York: Dover. (Original work published 1954)

Schoenberg, I. J. (1951). On Pólya Frequency Functions, I. The Totally Positive Functions and their Laplace Transforms. *Journal d'Analyse Mathématique*, *1*(1): 331-374. DOI: https://doi-org.ezphost.dur.ac.uk/10.1007/BF02790092

Sherman, S. (1955). A Theorem on Convex Sets with Applications. *Annals of Mathematical Statistics*, *26*: 763-766.

Simon, B. (2011). *Convexity: An Analytic Viewpoint*. (Vol. 187). Cambridge, United Kingdom: Cambridge University Press. DOI: https://doi-org.ezphost.dur.ac.uk/10.1017/CBO9780511910135

Thistle, P. (1993). Negative Moments, Risk Aversion, and Stochastic Dominance. *Journal of Financial and Quantitative Analysis*, *28*(2): 301-311. DOI: 10.2307/2331292

Tong, Y. L. (1994). Some Recent Developments on Majorization Inequalities in Probability and Statistics. *Linear Algebra and its Applications*. *199*(S 1): 69-90. DOI: https://doi.org/10.1016/0024-3795(94)90342-5

Trass, T. P., Van De Meent, D., Posthuma, L., Hamers, T., Kater, B. J., De Zwart, D., and Aldenberg, T. (2002). The Potentially Afected Fraction as a Measure of

Ecological Risk. In: Posthuma, L., II Suter, G. W. and Traas, T. P. Eds. *Species Sensitivity Distributions in Ecotoxicology*. Boca Raton: Lewis Publishers, 315-344.

Tucker, W. T., and Ferson, S. (2003). *Probability Bounds Analysis in Environmental Risk Assessment*. Setauket, New York Applied: Applied Biomathematics.

Van Straalen, N. M. (2002). Theory of Ecological Risk Assessment Based on Species Sensitivity Distributions. In: Posthuma, L., II Suter, G. W. and Traas, T. P. Eds. *Species Sensitivity Distributions in Ecotoxicology*. Boca Raton: Lewis Publishers, 37-48.

Van Straalen, N. M., and Denneman, C. A. J. (1989). Ecotoxicological Evaluation of Soil Quality Criteria. *Ecotoxicology and Environmental Safety*, *18*(3): 241-251. DOI: https://doi.org/10.1016/0147-6513(89)90018-3

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

WHO (World Health Organization)(1995). *Application of Risk Analysis to Food Standards Issues*. Report of the Joint FAO/WHO Expert Consultation. (Report No. WHO/FNU/FOS/95.3). Geneva, Switzerland: World Health Organisation. Retrieved from: https://apps.who.int/iris/bitstream/handle/10665/58913/WHO_FNU_FOS_95.3.pdf

Williams, P. (2007). Notes on Conditional Previsions. *The International Journal of Approximate Reasoning*, *44*(3): 366-383. (Original work published 1975). DOI: 10.1016/j.ijar.2006.07.019

Williamson, R. C., and Downs, T. (1990). Probabilistic Arithmetic I: Numerical Methods for Calculating Convolutions and Dependency Bounds. *International Journal of Approximate Reasoning*, *4*(2): 89-158. DOI: https://doi.org/10.1016/0888-613X(90)90022-T

Xu, M., and Hu, T. (2011). Some Inequalities of Linear Combinations of Independent Random Variables. I. *Journal of Applied Probability*, *48*(4): 1179-1188. DOI: https://doi.org/10.1239/jap/1324046026

Zhao, P., Chan, P. S., and Ng, H. K. T. (2011). Peakedness for Weighted Sums of Symmetric Random Variables. *Journal of Statistical Planning and Inference*, *141*(5): 1737-1743. DOI: https://doi.org/10.1016/j.jspi.2010.11.023