TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

DISSERTATION

On the Asymptotic Behaviour of the Estimator of Kendall's Tau

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der Naturwissenschaften unter der Leitung von

Univ.-Prof. Dr. Uwe Schmock

E 105 - 1
Institut für Wirtschaftsmathematik
Forschungsgruppe Finanz- und Versicherungsmathematik

eingereicht an der Technischen Universität Wien
bei der Fakultät für Mathematik und Geoinformation

von

Dipl.-Math. Barbara Dengler
Matrikelnummer: 0626635
Krongasse 4/5, 1050 Wien

Wien, im Februar 2010

Kurzfassung der Dissertation

Diese Dissertation behandelt Abhängigkeitsmaße und ihre Schätzer. Von besonderem Interesse ist dabei das asymptotische Verhalten der Schätzer, speziell die Eigenschaft der asymptotischen Normalität und die zugehörige asymptotische Varianz. Dafür gibt es hauptsächlich zwei Gründe. Zum einen erlaubt die asymptotische Normalität die Bildung von asymptotischen Konfidenzintervallen, welche Punktschätzungen ergänzen können, beispielsweise bei Tests auf Unabhängigkeit. Die zweite Anwendung besteht darin, Schätzverfahren anhand ihrer asymptotischen Varianz zu bewerten. Dies ist vorallem von Bedeutung, wenn zwei Schätzmethoden existieren und diese miteinander verglichen werden können.

Die zwei Abhängigkeitsmaße, die wir hier betrachten, sind der lineare Korrelationskoeffizient und das auf Rängen basierende Maß Kendall's Tau. Beide Schätzer sind asymptotisch normal und in der Literatur existieren Formeln für ihre asymptotischen Varianzen, wobei diese relativ kompliziert sind. In dieser Dissertation werden verschiedene Methoden zur Vereinfachung entwickelt.

Kendall's Tau und sein Schätzer basieren auf Rängen und somit hängen beide nur von der Abhängigkeitsstruktur, also von der Copula, und nicht von den Randverteilungen ab. Daher liefert die erste Vereinfachung der asymptotischen Varianz des Tau-Schätzers eine Formel die eine Funktion der Copula ist. Analytische Lösungen werden präsentiert für verschiedene Familien von Copulas, wie beispielsweise die Clayton-, die Farlie-Gumbel-Morgenstern- und die Marshall-Olkin-Familie.

Ein zweiter Ansatz die asymptotische Varianz des Tau-Schätzers zu vereinfachen, beruht auf einer geometrischen Betrachtung und gilt für achsensymmetrische Verteilungen. Eine konkrete Anwendungsmöglichkeit bietet sich dabei durch sphärische Verteilungen. Diese bilden eine Unterklasse der elliptischen Verteilungen, welche wiederum Verallgemeinerungen der Normalverteilung sind. Elliptische Verteilungen sind in der Praxis weit verbreitet, da sie mehr Flexibilität bieten als die Normalverteilung, dabei aber einige schöne Eigenschaften behalten. Eine davon ist die Beziehung zwischen dem linearen Korrelationskoeffizienten und Kendall's Tau, die zu zwei Schätzmethoden des Abhängigkeitsmaßes führt: direkte Verwendung des Standardschätzers des linearen Korrelationskoeffizienten oder Schätzung von Kendall's Tau und anschließende Umwandlung. Beide Methoden führen zu asymptotisch normalverteilten Schätzern und können daher anhand ihrer asymptotischen Varianzen verglichen werden.

Die Vereinfachungen, die für achsensymmetrische Verteilungen entwickelt wurden, führen zu expliziten Lösungen der asymptotischen Varianz des Tau-Schätzers für ver-

schiedene Verteilungen. Eine davon ist die unkorrelierte t-Verteilung. Die Berechnungen stellten sich als schwierig heraus, aber letzendlich konnten wir analytische Lösungen der asymptotischen Varianzen beider Schätzer für alle unkorrelierten t-Verteilungen mit ganzzahligen Freiheitsgraden entwickeln. Diese zeigen, dass besonders bei kleinen Freiheitsgraden, wo die Verteilung endlastig (heavy-tailed) ist, der alternative Schätzer mittels Kendall's Tau besser funktioniert als der Standardschätzer.

Abstract

This thesis is about dependence measures and their estimators. We are interested in the asymptotic behaviour of the estimators, especially in asymptotic normality and the corresponding asymptotic variance. This is mainly due to two reasons. The first one is that asymptotic normality allows to build asymptotic confidence intervals, which can complement point estimations, e.g. for tests of independence. The second application is to rate estimating procedures by their asymptotic variance. This is especially interesting if two estimation methods exist and can be compared.

The two dependence measures we consider are the classical linear correlation coefficient and the rank-based measure Kendall's tau. Both estimators are asymptotically normal and formulas for the asymptotic variance exist within the literature, although they are quite complicated. Different methods of simplification are developed within this thesis.

Kendall's tau is based on ranks and so is its estimator, thus they both only depend on the dependence structure, i.e. on the copula, and not on the marginal distributions. Therefore the first simplification for the asymptotic variance of the tau-estimator gives a formula that is a function of the copula. Analytical solutions are given for several well-known families, like e.g. the Clayton, the Farlie–Gumbel–Morgenstern and the Marshall–Olkin family.

A second approach to simplify the asymptotic variance of the tau-estimator is based on a geometrical consideration and is valid for axially symmetric distributions. We especially apply it to spherical distributions. They are a subclass of elliptical distributions, which generalize the normal distribution. Elliptical distributions are widely used in practice, as they provide more freedom than the normal distribution, but keep some nice properties. One of them is the connection between the linear correlation coefficient and Kendall's tau, which leads to two ways of estimating the dependence measure: using the standard estimator of the linear correlation directly or estimating Kendall's tau and transforming it. Both procedures lead to asymptotically normal estimators and can therefore be compared by their asymptotic variance.

The simplifications achieved for axially symmetric distributions lead to explicit solutions of the asymptotic variance of the tau-estimator for several distributions. One of them is the uncorrelated t-distribution. Calculations were tough, but we finally derived an analytic solution for both estimators for every uncorrelated t-distribution with integer valued degrees of freedom. It shows that especially for small degrees of freedom, where the distribution is heavy-tailed, the alternative estimator performs much better than the standard estimator.

Acknowledgements

The last four years, when I worked on this thesis, have been interesting and exciting, but also very tough. There are many people who supported me and helped me to finally finish my PhD, and I want to take the opportunity to thank them.

First, I want to thank my supervisor Prof. Uwe Schmock for his support and his strong belief that I can make it. He shared his big knowledge with me and taught me a good style in mathematics.

I am grateful for the financial support I got from Bank Austria and Christian Doppler Research Association via the PRisMa laboratory. I also want to thank Margarete Sobolewski and Wolfgang Weidinger, my colleagues from Bank Austria, for the good collaboration and the interesting and fruitful discussions.

I am also very thankful to my colleagues at FAM, especially those who became good friends: Verena, Benedikt, Sandra, Beatrice, Barbara, Maria, Christina and Karin. They made my everyday life at FAM much more colourful.

The final and deepest thanks go to my family, especially to my mother, to my sisters Lisa and Vroni, and to my brother Lukas. They always encouraged me and believed in me, more then I did myself. I know that they are always there for me whenever I need them.

Contents

Chapter 1

Introduction

This thesis is about dependence measures and their estimators. We consider asymptotic properties of the estimators, and are especially interested in asymptotic normality and the corresponding asymptotic variance.

The dependence measure we are mainly concerned with is Kendall's tau. It is a rank-based measure and can be understood as a measure of concordance. It fulfils many desirable properties, like it has the range $[-1, 1]$ and attains the boarders under perfect dependence. Further it only depends on the dependence structure, i.e. on the copula, and is therefore invariant under strictly increasing transformations of the marginal distributions.

The classical estimator of Kendall's tau is a U-statistic. These statistics have been first studied by Halmos (1946) and since then many nice properties have been discovered. U-statistics provide unbiased and strongly consistent estimators. They can be understood as extensions of sums of i.i.d. random variables and therefore similar asymptotic results can be derived. Under appropriate moment conditions they are asymptotically normal, which was first shown by Hoeffding (1948). Also Berry–Esseen theorems can be proved, whereas research about sharpening the bounds goes on. Besides these characteristics of U-statistics, the estimator of Kendall's tau keeps some nice properties of the measure, like being invariant under strictly increasing transformations of the marginal distributions, which leads to a more robust behaviour concerning outliers.

Asymptotic normality of the tau-estimator is a generalization of the classical central limit theorem and can be shown easily, using basic properties of U-statistics. Nonetheless, it is by far not trivial to find explicit expressions for the corresponding asymptotic variance. So one aim of this thesis is to find such solutions for different distributions. There are two main motivations. The first one is that asymptotic normality allows to build asymptotic confidence intervals, using the quantiles of the standard normal distribution, where the knowledge of the asymptotic variance is needed to determine the correct normalization. Asymptotic confidence intervals support point estimations of Kendall's tau by rating the liability of the estimation. This can especially be applied for tests of independence. Point estimators will almost surely not hit the value zero, assuming a continuous distribution, but the decision can be based on whether the asymptotic confidence interval contains this value or not.

A second motivation is using the asymptotic variance to rate the performance of an estimator. This becomes especially interesting when there are several ways of estimating a parameter. This situation is given if the underlying distribution is elliptical. Elliptical distributions generalize the normal distribution, keep some of the nice properties but also provide more freedom like increasing weight on the tails. Within this elliptical world, the classical linear correlation coefficient is the natural measure of dependence. There is further a connection between the linear correlation coefficient ϱ and Kendall's tau τ. Lindskog et al. (2003) showed that for all two-dimensional distribution with non-degenerate marginal distributions the following equation holds true:

$$\tau = \frac{2\, a_X}{\pi} \arcsin(\varrho)\,, \tag{1.1}$$

where the constant a_X only appears in a non-continuous framework and depends on the probability of the atoms.

This connection suggests two different ways of estimating the linear correlation: using the standard estimator directly or using the estimator for Kendall's tau and transforming this value into an estimator of the linear correlation via (1.1). The nice properties of the rank-based measure and its estimator, like e.g. its robustness against outliers in the data, might lead to a more stable estimation procedure, which could especially be crucial for heavy-tailed distributions. Since both estimators are asymptotically normal, a comparison of their performances based on the asymptotic variances is possible.

On the basis of these motivations there are three main achievements of this thesis concerning the asymptotic variance of estimators for measures of dependence: investigating the situations when the asymptotic variance of the tau-estimator equals zero, calculating the asymptotic variance of the tau-estimator for several families of copulas and comparing the asymptotic variances of the standard estimator and of the transformation-estimator for elliptical distributions.

The situation, when the asymptotic variance of the tau-estimator in the central limit theorem equals zero, is called degeneracy. Examples of distributions that lead to degeneracy are difficult to find as they must have a very special form, but we were able to give assumptions that induce degeneracy. On the other hand a distribution already guarantees non-degeneracy if it possesses a density that is continuous and strictly positive in one point. This idea is generalized such that we give a criterion that can be used to exclude degeneracy. On the assumption of ellipticity of the underlying distribution we could even develop a stronger result as we give a complete characterization of degeneracy in this case.

As Kendall's tau itself, also its estimator only depends on the copula and not on the marginal distributions. Therefore we were able to develop a formula of the asymptotic variance that is a function of the copula. This led to concrete results for several well-known families of copulas like e.g. the Clayton, the Farlie–Gumbel–Morgenstern and the Marshall–Olkin family.

In the case of elliptical distributions we are given two ways of estimating the linear correlation, the standard estimator and the transformation-estimator, and we are interested in rating those methods by comparing the asymptotic variance of the estimators. For the standard estimator we derived a quite simple formula, depending just on the moments of the distribution of the norm of the random vector. The formula for the asymptotic variance of the tau-estimator is much more complicated, but we nonetheless provide analytical solutions for several distributions. In the spherical case we reduced the original representation with nested expectations to a double integral. For the uncorrelated t-distribution, the most famous spherical distribution besides the standard normal, we even got analytical results for all integer valued degrees of freedom. The t-distribution with small degrees of freedom is heavy-tailed, and the asymptotic variances of the two estimators show that in this case the alternative estimator can essentially improve the estimating procedure.

The structure of the thesis is as follows.

In *Chapter 2* we introduce Kendall's tau and its estimator and show several properties. Some of them are deduced from the fact that the estimator is a U-statistic. Therefore we also present this well-studied topic of statistics and especially mention asymptotic statements like the central limit theorem and Berry–Esseen theorems. In the last part of this chapter we give a first survey of degeneracy, i.e. the case that the asymptotic variance in the central limit theorem equals zero.

Chapter 3 is concerned with elliptical distributions. We first give the basic definitions and derive the properties we will use afterwards. Since the linear correlation coefficient is the natural measure of dependence in the elliptical world, we introduce it here. We also explain the second way of estimating the linear correlation, which is given by equation (1.1), and show that after the transformation we still have asymptotic normality. We further introduce normal variance mixture distributions, a subclass of elliptical distributions that e.g. includes the t-distribution. In the last part of this chapter we revisit the case of degeneracy. Given that the underlying distribution is elliptical we present a complete characterization of degeneracy.

In *Chapter 4* we introduce the different methods we developed to calculate the asymptotic variance of the tau-estimator. The first method is based on the fact that the asymptotic variance, like Kendall's tau itself, only depends on the copula. We shortly introduce the important facts about copulas and then show how to calculate the asymptotic variance for a given copula. The second method is based on a geometrical consideration. Under the assumption that the distribution is axially symmetric, we could find a way to come from the original representation, that consists of nested, partly conditional expectations, to an easier analytical representation. The third method is based on the second and works with a simplification that is possible if the two-dimensional, axially symmetric density is a sum of products where the factors of each summand only contain one of the variables. Also the fourth method extends the second one, this time in the case where the distribution is a standard normal variance mixture distribution. The calculation of the asymptotic variance is further simplified and an example is given where the mixing distribution is an inverse gamma distribution. This is motivated by the fact that for the right choice of parameters this gives the t-distribution.

In *Chapter 5* we give some first examples of distributions where the asymptotic variance of the tau-estimator can be analytically calculated using the methods introduced in the previous chapter. The first part of the chapter is concerned about copulas, where we calculate the asymptotic variance in the case of several well-known copula-families. The second part shows an example of a spherical decomposable density. The density is a product of a polynomial in the square of the radius and of an exponential, similar to the normal density, to get an exponential decay in the tails.

In *Chapter 6* we look at elliptical distributions and compare the asymptotic variances of the two estimators introduced in Chapter 3. We start with some easy spherical distributions and revisit again the example of the decomposable density from the previous chapter. The most popular spherical distribution is the standard normal distribution and we show that we can compute both asymptotic variances here. In the last part of this chapter we consider the t-distribution, define it as a normal variance mixture distribution, determine the asymptotic variance of the standard estimator and can even give analytical solutions of the asymptotic variance of the tau-estimator in the uncorrelated case with integer valued degrees of freedom.

In *Chapter 7* we present the proof of the asymptotic variance of the tau-estimator for the uncorrelated t-distribution. Starting with the simplifications achieved for standard normal variance mixture distributions, there is still some work to do. We first show the main ideas of the proof. In the following section we introduce some necessary tools, like e.g. the polylogarithm, and develop some useful solutions for sums over binomial coefficients and definite integrals. The details of the main proof can be found in the last part of this chapter.

Chapter 2

Kendall's tau, its estimator, and the theory of U-statistics

This work is concerned with measures that represent the strength of dependence between two random variables. The dependence measures are functions of the common probability measure μ. Although generalizations are possible, e.g. for Kendall's tau, we want μ to be defined on $\mathbb{R}^2$. Denote the set of all probability measures on $\mathbb{R}^2$ with Borel σ-algebra by $\mathcal{M}_1(\mathbb{R}^2)$. Whenever we are talking about just two random variables we want to denote them by

$$(X, Y) \sim \mu \,.$$

The dependence measure we are mainly interested in is Kendall's tau. Within this chapter we introduce this rank-based measure and present its estimator. We derive several properties of the estimator, some of them provided from the theory of U-statistics. Another well-known measure is the classical linear correlation coefficient. It will be presented and discussed in Section 3.2.

Within this work we mostly consider the two-dimensional case. Given a multivariate setting with more than two random variables, dependence measures can be applied pairwise (see e.g. McNeil et al., 2005, pp. 63–65 for the linear correlation, p. 207 for Kendall's tau). Estimation works as in the two-dimensional case, although consistency problems of the estimated dependence matrix can arise, but we do not want to go into detail here.

2.1 U-statistics

As the estimator of Kendall's tau is a U-statistic, we will use several results from this well-studied topic of statistics within our work. Therefore we give a short introduction of U-statistics before talking about concrete dependence measures. There are no proofs given in this section, instead we refer e.g. to the textbooks of Lee (1990), Koroljuk and Borovskich (1994) and Borovskikh (1996) for further details.

The history of U-statistics goes back to Halmos (1946), publishing an article about minimum-variance unbiased estimators. The name U-statistics was first introduced by Hoeffding (1948), where he also proved the central limit theorem.

Within this section we define U-statistics, explain what its rank is, and show a very useful representation, called Hoeffding decomposition. In Section 2.3 we will revisit U-statistics and show some asymptotic properties.

Definition 2.1. Fix $m \in \mathbb{N}$. For $n \geq m$ let $Z_1, \ldots, Z_n$ be random variables taking values in a measurable space $(\mathcal{Z}, \mathfrak{Z})$ and let $\kappa : \mathcal{Z}^m \to \mathbb{R}$ be a symmetric measurable function. The *U-statistic* $\hat{U}_n^m(\kappa)$ belonging to the *kernel* κ of degree m is defined as

$$\hat{U}_n^m(\kappa) := \binom{n}{m}^{-1} \sum_{1 \leq i_1 < \cdots < i_m \leq n} \kappa(Z_{i_1}, \ldots, Z_{i_m})\,. \tag{2.1}$$

Notation. To shorten notation we omit the index m or the argument κ whenever the degree or the kernel, respectively, are clear from the context.

Remark 2.2. We only talk about real-valued kernels as this will suffice for our work. More generally one could define the kernel as

$$\kappa : \mathcal{Z}^m \to B\,,$$

where B is a real separable Banach space (see e.g. Borovskikh, 1996).

2.1.1 Basic properties

From now on we assume i.i.d. random variables. We are given the probability space $(\mathcal{Z}, \mathfrak{Z}, \mu)$ and define the product space $(\Omega, \mathfrak{A}, \mathbb{P}) \equiv (\mathcal{Z}^{\mathbb{N}}, \mathfrak{Z}^{\otimes \mathbb{N}}, \mu^{\otimes \mathbb{N}})$ such that $(Z_i)_{i \in \mathbb{N}}$ are the coordinate projections from Ω to $\mathcal{Z}$, forming an i.i.d. sequence with $\mathcal{L}(Z_i) = \mu$. Assume that $\int_{\mathcal{Z}^m} |\kappa| \, d\mu^{\otimes m} < \infty$.

These assumptions suffice for unbiasedness of the U-statistic and for the strong law of large numbers. In Section 2.3.2 we will further present the central limit theorem and Berry–Esseen theorems, which hold true under appropriate moment conditions.

For i.i.d. random variables, a U-statistic with an integrable kernel κ of degree m is an unbiased estimate of

$$\theta := \int_{\mathcal{Z}^m} \kappa \, d\mu^{\otimes m} \tag{2.2}$$

(see e.g. Koroljuk and Borovskich, 1994, p. 18). This observation led to the name U-statistic.

Another desirable property that U-statistics possess is strong consistency. A series of estimators is called weakly consistent, if it converges to the real parameter in probability. If the convergence holds almost surely, then we talk about strong consistency. This is equivalent to fulfilling the strong law of large numbers:

Theorem 2.3. *Let $(Z_j)_{j\in\mathbb{N}}$ be i.i.d. random variables, $\mathcal{L}(Z_j) = \mu$, and let $\big(\hat{U}_n^m(\kappa)\big)_{n\geq m}$ be U-statistics with kernel κ of degree m such that $\int_{\mathcal{Z}^m} |\kappa| \, d\mu^{\otimes m} < \infty$. Then*

$$\hat{U}_n^m(\kappa) \overset{n\to\infty}{\longrightarrow} \theta \quad a.s.$$

Proof. See e.g. Koroljuk and Borovskich (1994, pp. 93–95). □

2.1.2 Canonical functions and the rank of a U-statistic

For every $c = 0, \dots, m$ define the functions $\kappa_c \in L^1(\mathcal{Z}^c, \mu^{\otimes c})$ as

$$\kappa_c(z_1, \dots, z_c) = \int_{\mathcal{Z}^m} \kappa \, d\big(\delta_{z_1} \otimes \cdots \otimes \delta_{z_c} \otimes \mu^{\otimes(m-c)}\big) \quad \text{for } \mu^{\otimes c}\text{-almost all } (z_1, \dots, z_c) \in \mathcal{Z}^c\,,$$

such that $\kappa_0 = \theta$ and $\kappa_m = \kappa$. For ease of notation write

$$\tilde{\kappa} = \kappa - \theta \quad \text{and} \quad \tilde{\kappa}_c = \kappa_c - \theta, \quad c = 0, \dots, m\,.$$

We can now define the functions $h_c \in L^1(\mathcal{Z}^c, \mu^{\otimes c})$ as

$$h_c(z_1, \dots, z_c) = (-1)^c \sum_{d=1}^{c} (-1)^d \sum_{1\leq j_1 < \cdots < j_d \leq c} \tilde{\kappa}_d(z_{j_1}, \dots, z_{j_d})\,, \quad c = 1, \dots, m\,, \tag{2.3}$$

called canonical functions. It can be shown (see e.g. Koroljuk and Borovskich, 1994, pp. 21–22), that

$$h_c(z_1, \dots, z_c) = \int_{\mathcal{Z}^m} \kappa \, d\big((\delta_{z_1} - \mu) \otimes \cdots \otimes (\delta_{z_c} - \mu) \otimes \mu^{\otimes(m-c)}\big)$$
$$\text{for } \mu^{\otimes c}\text{-almost all } (z_1, \dots, z_c) \in \mathcal{Z}^c\,.$$

Canonical functions will be used as kernels in the Hoeffding representation (see Section 2.1.3) and are also used in the classical definition of the rank of a U-statistic. A U-statistic $\hat{U}_n^m(\kappa)$ or its kernel κ are said to be of rank $r \in \{1, \dots, m-1\}$, if $h_c = 0$ for all $c \in \{1, \dots, r-1\}$ and $h_r \neq 0$, and to be of rank m, otherwise. If $r = 1$, then the U-statistic is called non-degenerate. In the case $r \geq 2$ the U-statistic is called degenerate and r is called the order of degeneracy. If $r = m$ we talk about complete degeneracy. The canonical functions possess the property of complete degeneracy.

In the following we will mostly be concerned with kernels of degree 2. In this case the kernel can either be non-degenerate or completely degenerate. It is degenerate if and only if

$$\int_{\mathcal{Z}^2} \kappa \, d\big(\delta_{z_1} \otimes \mu\big) = \theta \quad \text{for } \mu\text{-almost all } z_1 \in \mathcal{Z}\,.$$

2.1.3 Hoeffding representation

One of the most important properties of U-statistics is a useful decomposition first shown by Hoeffding (1961). It is called Hoeffding representation and decomposes a U-statistic of degree m into uncorrelated U-statistics of degree $1, 2, \ldots, m$ in the following way:

$$\hat{U}_n^m(\kappa) - \theta = \sum_{c=1}^{m} \binom{m}{c} \hat{U}_n^c(h_c)\,, \tag{2.4}$$

where h_c are the canonical functions as defined in (2.3).

Notation. We want to denote the U-statistics $\hat{U}_n^c(h_c)$ by $H_n^{(c)}$, i.e.

$$H_n^{(c)} = \binom{n}{c}^{-1} \sum_{1 \le i_1 < \cdots < i_c \le n} h_c(Z_{i_1}, \ldots, Z_{i_c})\,, \quad c = 1, \ldots, m\,.$$

The Hoeffding representation will be used in the next section to prove the asymptotic normality of the estimator of Kendall's tau and can also be found in the proofs of Berry–Esseen theorems for U-statistics. The following properties will be important:

- The expectation of the kernels h_c is zero,

  $$\int_{\mathcal{Z}^c} h_c \, d\mu^{\otimes c} = 0\,, \quad c = 1, \ldots, m\,,$$

  such that every U-statistic $H_n^{(c)}$, $c = 1, \ldots, m$, is an unbiased estimate for the value zero.

- The U-statistics $H_n^{(c)}$, $c = 1, \ldots, m$, are pairwise uncorrelated.

For details and proofs about the Hoeffding representation see e.g. Lee (1990, pp. 25–33) or Koroljuk and Borovskich (1994, pp. 23–24).

2.2 Kendall's tau and its estimator

2.2.1 Definitions

The dependence measure Kendall's tau is the expectation of the U-statistic that belongs to the kernel

$$\begin{aligned} &\kappa_\tau : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}\,, \\ &\kappa_\tau\big((x, y), (\tilde{x}, \tilde{y})\big) = \operatorname{sgn}(x - \tilde{x}) \operatorname{sgn}(y - \tilde{y})\,, \end{aligned} \tag{2.5}$$

where the signum function is defined as

$$\operatorname{sgn}(x) := \begin{cases} 1\,, & \text{if } x > 0\,, \\ 0\,, & \text{if } x = 0\,, \\ -1\,, & \text{if } x < 0\,. \end{cases}$$

Notation. We want to call κ_τ *tau-kernel*.

Kendall's tau is classically defined by the difference between the probability of concordance and the probability of discordance. Two points (x, y) and $(\tilde{x}, \tilde{y}) \in \mathbb{R}^2$ are called concordant if $(x - \tilde{x})(y - \tilde{y}) > 0$, and they are said to be discordant if $(x - \tilde{x})(y - \tilde{y}) < 0$. As only the ordering and not the distance of points plays a role, Kendall's tau is said to be a rank-based measure.

Definition 2.4. For a probability measure μ on $\mathbb{R}^2$ the dependence measure *Kendall's tau* τ is defined as

$$\begin{aligned} \tau = \tau(\mu) :=& (\mu \otimes \mu)\Big(\big\{\big((x,y),(\tilde{x},\tilde{y})\big) \in \mathbb{R}^2 \times \mathbb{R}^2 \;\big|\; (x-\tilde{x})(y-\tilde{y}) > 0\big\} \Big) \\ &- (\mu \otimes \mu)\Big(\big\{\big((x,y),(\tilde{x},\tilde{y})\big) \in \mathbb{R}^2 \times \mathbb{R}^2 \;\big|\; (x-\tilde{x})(y-\tilde{y}) < 0\big\} \Big) . \end{aligned} \tag{2.6}$$

Notation. For the basic definition we chose this clear notation to stress that Kendall's tau is a function of the common probability measure μ. Within the literature one can very often find the following notation for the definition of Kendall's tau:

$$\tau(X,Y) = \mathbb{P}[\,(X-\tilde{X})(Y-\tilde{Y}) > 0\,] - \mathbb{P}[\,(X-\tilde{X})(Y-\tilde{Y}) < 0\,] ,$$

with independent $(X, Y), (\tilde{X}, \tilde{Y}) \sim \mu$. For ease of notation we will also use this short form in the following whenever the underlying probability measure is clear from the context.

Remark 2.5. Since Kendall's tau is the expectation of a U-statistic, it can be equivalently defined as

$$\tau(X,Y) = \mathbb{E}[\,\operatorname{sgn}(X-\tilde{X})\,\operatorname{sgn}(Y-\tilde{Y})\,] , \tag{2.7}$$

with independent $(X, Y), (\tilde{X}, \tilde{Y}) \sim \mu$.

Remark 2.6. To define Kendall's tau we do not necessarily need the random variables to be real-valued, it is enough that they are defined on an ordered set. For ease of computation we nonetheless only look at real-valued random variables within this work.

Definition 2.7. For $n \geq 2$ pairs of real-valued random variables (X_i, Y_i), $i = 1, \ldots, n$, the estimator for Kendall's tau, short *tau-estimator*, is defined as

$$\hat{\tau}_n = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} \operatorname{sgn}(X_i - X_j)\,\operatorname{sgn}(Y_i - Y_j) . \tag{2.8}$$

Remark 2.8. As already mentioned in the beginning of this section, the estimator of Kendall's tau is a U-statistic with kernel κ_τ. This tau-kernel is bounded by 1 and therefore integrable for all probability measures on $\mathbb{R}^2 \times \mathbb{R}^2$.

Remark 2.9. From the theory of U-statistics we know that the tau-estimator is an unbiased and strongly consistent estimator (see Section 2.1.1).

Remark 2.10. If $\tau = 1$, then $\hat{\tau}_n = 1$ for every $n \geq 2$, and similarly for $\tau = -1$.

2.2.2 Variance of the tau-estimator

In this section we derive the variance of the tau-estimator for a fixed sample size using combinatorial arguments.

Lemma 2.11. *Fix $n \in \mathbb{N}$, $n \geq 2$. For n i.i.d. pairs of real-valued random variables (X_j, Y_j), $j = 1, \ldots, n$, the tau-estimator $\hat{\tau}_n$, as defined in* (2.8)*, has the variance*

$$\sigma_{\tau,n}^2 := \mathbb{V}\mathrm{ar}[\,\hat{\tau}_n\,] = \frac{4\,(n-2)}{n\,(n-1)}\,\sigma_1^2 + \frac{2}{n\,(n-1)}\,\sigma_2^2\,, \tag{2.9}$$

where

$$\sigma_1^2 := \mathbb{V}\mathrm{ar}\Big[\,\mathbb{E}\big[\,\kappa_\tau\big((X_1,Y_1),(X_2,Y_2)\big) \mid X_1, Y_1\,\big]\,\Big] \tag{2.10}$$

and

$$\sigma_2^2 := \mathbb{V}\mathrm{ar}\big[\,\kappa_\tau\big((X_1,Y_1),(X_2,Y_2)\big)\,\big]\,. \tag{2.11}$$

Proof. The variance of the tau-estimator is

$$\begin{aligned}\sigma_{\tau,n}^2 &= \mathbb{V}\mathrm{ar}[\,\hat{\tau}_n\,] = \mathbb{E}[\,\hat{\tau}_n^2\,] - \big(\mathbb{E}[\,\hat{\tau}_n\,]\big)^2 \\ &= \binom{n}{2}^{-2} \mathbb{E}\Big[\Big(\sum_{1\leq i<j\leq n} \mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)\Big)^2\Big] - \tau^2\,.\end{aligned}$$

When we expand the square we get summands of the form

$$\mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)\,\mathrm{sgn}(X_k - X_l)\,\mathrm{sgn}(Y_k - Y_l)\,]$$

with $1 \leq i < j \leq n$ and $1 \leq k < l \leq n$. The following constellations for i, j, k and l are possible:

- $i = k$ and $j = l$:
  These $\binom{n}{2} = \frac{n\,(n-1)}{2}$ summands have the value $(1 \leq i < j \leq n)$

$$\mathbb{E}\big[\,\big(\mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)\big)^2\,\big] = \sigma_2^2 + \tau^2\,. \tag{2.12}$$

- All the indices are different:
  Here we get $\binom{n}{2}\binom{n-2}{2} = \frac{n\,(n-1)\,(n-2)\,(n-3)}{4}$ summands of the form

$$\begin{aligned}&\mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)\,\mathrm{sgn}(X_k - X_l)\,\mathrm{sgn}(Y_k - Y_l)\,] \\ &\quad = \mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)\,]^2 \\ &\quad = \tau^2\,.\end{aligned}$$

- Exactly two of i, j, k, l have the same value:
  We get $\binom{n}{2}^2 - \binom{n}{2} - \binom{n}{2}\binom{n-2}{2} = n\,(n-1)\,(n-2)$ summands of this form. All the cases can be simplified similarly as we show for $i = k$ and $j \neq l$:

$$\begin{aligned}
&\mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\ \mathrm{sgn}(Y_i - Y_j)\ \mathrm{sgn}(X_i - X_l)\ \mathrm{sgn}(Y_i - Y_l)\,] \\
&\quad = \mathbb{E}\big[\,\mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\ \mathrm{sgn}(Y_i - Y_j)\ \mathrm{sgn}(X_i - X_l)\ \mathrm{sgn}(Y_i - Y_l) \mid X_i, Y_i\,]\,\big] \\
&\quad = \mathbb{E}\big[\,\mathbb{E}[\,\mathrm{sgn}(X_i - X_j)\ \mathrm{sgn}(Y_i - Y_j) \mid X_i, Y_i\,]^2\,\big] \\
&\quad = \sigma_1^2 + \tau^2\,.
\end{aligned} \tag{2.13}$$

In summary we get

$$\begin{aligned}
\sigma_{\tau,n}^2 &= \frac{2}{n\,(n-1)}\left(\sigma_2^2 + \tau^2\right) + \frac{(n-2)\,(n-3)}{n\,(n-1)}\,\tau^2 + \frac{4\,(n-2)}{n\,(n-1)}\left(\sigma_1^2 + \tau^2\right) - \tau^2 \\
&= \frac{4\,(n-2)}{n\,(n-1)}\,\sigma_1^2 + \frac{2}{n\,(n-1)}\,\sigma_2^2\,.
\end{aligned}$$

□

Remark 2.12. We decided to show this direct proof of the variance. Clearly we also could have taken the result from the theory of U-statistics, e.g. from Lee (1990, p. 14). To make the two ways comparable we chose the notation of σ_1^2 and σ_2^2 according to Lee (1990) such that it is easy to see that the results coincide.

Remark 2.13. As the signum function is bounded by 1, equations (2.12) and (2.13) suggest the bounds

$$\sigma_1^2 \leq 1 - \tau^2 \quad \text{and} \quad \sigma_2^2 \leq 1 - \tau^2\,.$$

So we know for every $n \geq 2$

$$\sigma_{\tau,n}^2 \leq 4\,\frac{n - \frac{3}{2}}{n\,(n-1)}\,(1 - \tau^2)\,.$$

Remark 2.14. Using again formula (2.12) we can write the variance σ_2^2 as

$$\begin{aligned}
\sigma_2^2 &= \mathbb{E}\big[\,\big(\mathrm{sgn}(X_1 - X_2)\big)^2\,\big(\mathrm{sgn}(Y_1 - Y_2)\big)^2\,\big] - \tau^2 \\
&= \mathbb{P}[\,X_1 \neq X_2, Y_1 \neq Y_2\,] - \tau^2\,.
\end{aligned}$$

For continuous marginal distributions this reduces to

$$\sigma_2^2 = 1 - \tau^2$$

and formula (2.9) for the variance simplifies accordingly.

2.2.3 Asymptotic normality

U-statistics that are based on i.i.d. random variables can be understood as extensions of sums of i.i.d. random variables. Hence it is no surprise that there exists a central limit theorem for U-statistics, too. Under the assumption of square-integrability of the kernel the proof is quite short and we show it here for the tau-kernel, which is clearly square-integrable. In Section 2.3.2 we will present a central limit theorem with weaker conditions like it is proved e.g. in Koroljuk and Borovskich (1994, pp. 129–131).

The proof is based on the Hoeffding representation from the theory of U-statistics and on the result of the variance of the tau-estimator as proved in Lemma 2.11.

Theorem 2.15. *For i.i.d. pairs of real-valued random variables (X_j, Y_j), $j \in \mathbb{N}$, the tau-estimators $(\hat{\tau}_n)_{n\geq 2}$, as defined in (2.8), normalized with $\sqrt{n}$, are asymptotically normal,*

$$\sqrt{n}\,\big(\hat{\tau}_n - \tau\big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, 4\,\sigma_1^2\big), \quad n \to \infty\,, \tag{2.14}$$

with σ_1^2 as defined in (2.10):

$$\sigma_1^2 = \mathbb{E}\big[\,\mathbb{E}[\,\operatorname{sgn}(X_1 - X_2)\,\operatorname{sgn}(Y_1 - Y_2) \mid X_1, Y_1\,]^2\,\big] - \tau^2\,.$$

Notation. For simplicity we use a special notation for the asymptotic variance, namely

$$\sigma_\tau^2 := 4\,\sigma_1^2\,. \tag{2.15}$$

Proof. We first want to ensure that the asymptotic variance is correct. Knowing the variance for fixed n from Lemma 2.11, we get

$$\begin{aligned}\mathbb{V}\mathrm{ar}\big[\,\sqrt{n}\,\big(\hat{\tau}_n - \tau\big)\,\big] = n\,\sigma_{\tau,n}^2 &= \frac{4\,(n-2)}{n-1}\,\sigma_1^2 + \frac{2}{n-1}\,\sigma_2^2 \\ &\overset{n\to\infty}{\longrightarrow} 4\,\sigma_1^2\,.\end{aligned}$$

To show asymptotic normality we use the Hoeffding representation (2.4). In the case of the tau-estimator we have a kernel of degree 2 and therefore the representation

$$\begin{aligned}\sqrt{n}\,\big(\hat{\tau}_n - \tau\big) &= 2\,\sqrt{n}\,H_n^{(1)} + \sqrt{n}\,H_n^{(2)} \\ &= \frac{2}{\sqrt{n}}\sum_{1\leq i\leq n} h_1\big((X_i, Y_i)\big) + \sqrt{n}\,H_n^{(2)}\,.\end{aligned}$$

Since $h_1 = \tilde{\kappa}_1$ by (2.3), the variance of the components of the first summand is the following:

$$\mathbb{V}\mathrm{ar}\big[\,h_1\big((X_i, Y_i)\big)\,\big] = \mathbb{V}\mathrm{ar}\big[\,\mathbb{E}[\,\operatorname{sgn}(X_1 - X_2)\,\operatorname{sgn}(Y_1 - Y_2) \mid X_1, Y_1\,]\,\big] = \sigma_1^2\,.$$

As the pairs (X_i, Y_i), $i \in \mathbb{N}$, are assumed to be i.i.d., we know from the central limit theorem that

$$\frac{2}{\sqrt{n}}\sum_{1\leq i\leq n} h_1\big((X_i, Y_i)\big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, 4\,\sigma_1^2\big), \quad n \to \infty\,.$$

This asymptotic variance equals the asymptotic variance of the tau-estimator and the components of the Hoeffding representation are uncorrelated, so we know that the variance of the second component of the Hoeffding representation must vanish for $n \to \infty$:

$$\mathbb{V}\mathrm{ar}[\, \sqrt{n}\, H_n^{(2)} \,] \overset{n\to\infty}{\longrightarrow} 0\,.$$

As the U-statistics $H_n^{(c)}$ are estimates of the value zero this component must converge to zero in probability for $n \to \infty$:

$$\sqrt{n}\, H_n^{(2)} \overset{\mathbb{P}}{\longrightarrow} 0, \quad n \to \infty\,.$$

Slutsky's theorem (see e.g. Lehmann, 1999, p. 70) gives then the asymptotic behaviour claimed in the theorem. □

Remark 2.16. For some special cases the asymptotic distribution of $\sqrt{n}\,(\hat{\tau}_n - \tau)$ can be degenerate, i.e. $\sigma_\tau^2 = 0$. We will have a closer look at these degenerate cases in Section 2.4 and, for elliptical distributions, in Section 3.4.

Remark 2.17. We know from Remark 2.13 that the variance of the tau-estimator is bounded for every $n \geq 2$. This leads to the following upper bound of the asymptotic variance:

$$\sigma_\tau^2 \leq 4\,(1 - \tau^2)\,.$$

2.2.4 Asymptotic confidence intervals

As we know the asymptotic distribution of the tau-estimator we can determine asymptotic confidence intervals.

Theorem 2.15 shows that in the non-degenerate case, i.e. if $\sigma_\tau^2 > 0$, the normalized tau-estimator behaves asymptotically like a standard normally distributed random variable. So we have the following convergence:

$$\mathbb{P}\Big[\, \frac{\sqrt{n}}{\sigma_\tau}\,\big(\hat{\tau}_n - \tau\big) \leq z \,\Big] \overset{n\to\infty}{\longrightarrow} \mathfrak{N}(z)\,, \quad z \in \mathbb{R}\,.$$

Using the theorem of Pólya as given in Witting and Müller-Funk (1995, p. 71), we know that this convergence holds even uniformly in z, i.e.

$$\sup_{z\in\mathbb{R}} \Big| \mathbb{P}\Big[\, \frac{\sqrt{n}}{\sigma_\tau}\,\big(\hat{\tau}_n - \tau\big) \leq z \,\Big] - \mathfrak{N}(z) \Big| \overset{n\to\infty}{\longrightarrow} 0\,.$$

So for every fixed confidence level $(1-\alpha)$, $\alpha \in (0,1)$, we can use the α-quantile u_α of the standard normal distribution to formulate an asymptotic confidence interval as

$$\Big[\hat{\tau}_n - \frac{\sigma_\tau}{\sqrt{n}}\, u_{1-\frac{\alpha}{2}}\,,\, \hat{\tau}_n + \frac{\sigma_\tau}{\sqrt{n}}\, u_{1-\frac{\alpha}{2}}\Big]\,. \tag{2.16}$$

The probability, that this method produces an interval which contains the true parameter τ, converges to $(1-\alpha)$ for a growing sample size,

$$\mathbb{P}\Big[\,\hat{\tau}_n - \frac{\sigma_\tau}{\sqrt{n}}\,u_{1-\frac{\alpha}{2}} \le \tau \le \hat{\tau}_n + \frac{\sigma_\tau}{\sqrt{n}}\,u_{1-\frac{\alpha}{2}}\,\Big] \overset{n\to\infty}{\longrightarrow} 1-\alpha$$

uniformly in $\alpha \in (0,1)$. Nonetheless, the exact confidence level for a fixed n is not known (see Hartung et al., 2005, p. 131). For more details about asymptotic confidence intervals see e.g. Lehmann (1999).

2.2.5 Behaviour under weak convergence

We are interested in the convergence of Kendall's tau and the asymptotic variance of its estimator under weak convergence of the underlying probability measure. Therefore we want to rewrite formula (2.7) of Kendall's tau as

$$\tau(\mu) = \int_{\mathbb{R}^2\times\mathbb{R}^2} \kappa_\tau(z,\tilde{z})\,(\mu\otimes\mu)(dz,d\tilde{z}) \tag{2.17}$$

and formula (2.15) of the asymptotic variance of its estimator as

$$\sigma_\tau^2(\mu) = 4\int_{(\mathbb{R}^2)^3} \kappa_\tau(z,\tilde{z})\,\kappa_\tau(z,\bar{z})\,\mu^{\otimes 3}(dz,d\tilde{z},d\bar{z}) - 4\,\tau^2(\mu)\,. \tag{2.18}$$

Lemma 2.18. *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathbb{R}^2$, which converges weakly to μ. If*

$$\mu\big(\{\tilde{z}\in\mathbb{R}^2 \mid \kappa_\tau(z,\tilde{z})=0\}\big) = 0 \tag{2.19}$$

for μ-almost all $z\in\mathbb{R}^2$, then

(i) $\lim_{n\to\infty}\tau(\mu_n) = \tau(\mu)$ *and*

(ii) $\lim_{n\to\infty}\sigma_\tau^2(\mu_n) = \sigma_\tau^2(\mu)$.

Proof. From Billingsley (1968, Thm 3.2, p. 21) we know that also the product measures $\mu_n\otimes\mu_n$ and $\mu_n^{\otimes 3}$ converge weakly to $\mu\otimes\mu$ and $\mu^{\otimes 3}$, respectively.

(i) Define

$$D_\tau = \{(z,\tilde{z})\in\mathbb{R}^2\times\mathbb{R}^2 \mid \kappa_\tau(z,\tilde{z})=0\}\,.$$

κ_τ is continuous on $(\mathbb{R}^2)^2\setminus D_\tau$ since D_τ contains the set of discontinuities of κ_τ. By Fubini's theorem and the assumptions, $(\mu\otimes\mu)(D_\tau)=0$. Since κ_τ is bounded, part (i) follows from Billingsley (1968, Thm 5.2(iii), p. 31).

(ii) Define $s(z,\tilde{z},\bar{z}) = \kappa_\tau(z,\tilde{z})\,\kappa_\tau(z,\bar{z})$ for all $z,\tilde{z},\bar{z}\in\mathbb{R}^2$ and the set

$$D_\sigma = \{(z,\tilde{z},\bar{z})\in(\mathbb{R}^2)^3 \mid s(z,\tilde{z},\bar{z})=0\}\,,$$

which contains the set of discontinuities of s. As in part (i), $\mu^{\otimes 3}(D_\sigma)=0$. Since s is bounded, part (ii) also follows from Billingsley (1968, Thm 5.2(iii), p. 31).

□

2.3 Asymptotic statements for U-statistics

Additionally to showing the asymptotic normality of the tau-estimator directly, we also give a summary of asymptotic statements for U-statistics in general. We start again with asymptotic normality, but the main part is devoted to the rate of convergence, showing asymptotic expansions and Berry–Esseen theorems. As the theorems for U-statistics are based on the ones for sums of i.i.d. random variables, we start with this classical part. Proofs are mostly omitted or just sketched, but we give references.

2.3.1 Asymptotics for sums of i.i.d. random variables

For the classical case of sums of i.i.d. random variables we follow the book of Feller (1971, pp. 531–544). We first introduce some basic definitions and results we will need, and show then expansions for densities, for distributions and finally the Berry–Esseen theorem. The proof of the last one is given, for the other results we just give the main ideas of the proofs.

Fourier transform and characteristic function

Within the following subsections we will need the Fourier transform, the characteristic function and some of its properties.

Definition 2.19. The *Fourier transform* $\hat{u}$ of an integrable function u is defined as

$$\hat{u}(\zeta) := \int_{\mathbb{R}} e^{i\zeta x}\, u(x)\, dx, \quad \zeta \in \mathbb{R}\,.$$

The Fourier transform of an integrable function is continuous and vanishes at infinity. If it is further integrable, then

$$u(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\zeta x}\, \hat{u}(\zeta)\, d\zeta, \quad \text{for almost all } x \in \mathbb{R}\,,$$

(see Rudin, 1987, pp. 180–185), which suggests the estimate

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}} |u(x)| \le \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{u}(\zeta)|\, d\zeta\,, \tag{2.20}$$

called Fourier norm of u.

If u is a density, then the Fourier transform equals the characteristic function.

Definition 2.20. The *characteristic function* of a distribution function F is defined as

$$\varphi(\zeta) = \int_{\mathbb{R}} e^{i\zeta x}\, F(dx), \quad \zeta \in \mathbb{R}\,. \tag{2.21}$$

Lemma 2.21. *Let φ be the characteristic function of a real-valued random variable X and let $a, b \in \mathbb{R}$. Then*

(i) *φ is continuous;*

(ii) *$\varphi(0) = 1$ and $|\varphi(\zeta)| \leq 1$ for all $\zeta \in \mathbb{R}$;*

(iii) *$aX + b$ has the characteristic function $\zeta \mapsto e^{ib\zeta}\varphi(a\zeta)$. In particular, the characteristic function of $-X$ is the complex conjugate of φ.*

Proof. See Feller (1971, pp. 499–500). □

Theorem 2.22 (Uniqueness and continuity, Lévy)**.** *Let G and G_n, $n \in \mathbb{N}$, be distribution functions with characteristic functions γ and γ_n, respectively. Then the sequence $(G_n)_{n\in\mathbb{N}}$ converges weakly to G if and only if $\gamma_n(\zeta) \to \gamma(\zeta)$ for every $\zeta \in \mathbb{R}$, and then $\gamma_n \to \gamma$ uniformly on every bounded set.*

Proof. See Kallenberg (2002, p. 86). □

Definition 2.23. Let G be a distribution function and h an integrable or a bounded, measurable function. Then the *convolution* $G \star h$ is defined by

$$(G \star h)(x) := \int_{\mathbb{R}} h(x-y)\, G(dy), \quad \text{for almost all } x \in \mathbb{R}\,.$$

If G possesses a density g, this simplifies to

$$(g * h)(x) := \int_{\mathbb{R}} h(x-y)\, g(y)\, dy, \quad \text{for almost all } x \in \mathbb{R}\,. \tag{2.22}$$

Notation. The convolution of a function h with a distribution function G will be denoted by $G \star h$, the convolution with a density g will be denoted by $g * h$.

Lemma 2.24. *If H is a distribution function so is $G \star H$.*

Proof. See Feller (1971, p. 144). □

Lemma 2.25. *Let G be a distribution function with density g and with characteristic function γ and let h be an integrable function with Fourier transform η. Then the convolution $G \star h$ has the Fourier transform $\gamma \cdot \eta$.*

Proof. As G has a density we can use representation (2.22) and get the result from Rudin (1987, Thm 9.2(c), p. 179). □

Lemma 2.26. *Let G_1 and G_2 be two distribution functions with characteristic functions γ_1 and γ_2, respectively. Then the convolution $G_1 \star G_2$ has the characteristic function $\gamma_1 \cdot \gamma_2$.*

Proof. See Feller (1971, p. 500). □

Theorem 2.27 (Fourier inversion)**.** *Let φ be the characteristic function of a distribution function F and assume that $\int_{\mathbb{R}} |\varphi(\zeta)|\, d\zeta < \infty$. Then F has a bounded continuous density f given by*

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\zeta x}\, \varphi(\zeta)\, d\zeta, \quad x \in \mathbb{R}\,.$$

Proof. See Feller (1971, pp. 509–510). □

Lemma 2.28 (Riemann–Lebesgue)**.** *If g is integrable and*

$$\gamma(\zeta) = \int_{\mathbb{R}} e^{i\zeta x}\, g(x)\, dx, \quad \zeta \in \mathbb{R}\,,$$

then $\gamma(\zeta) \to 0$ as $\zeta \to \pm\infty$.

Proof. See Feller (1971, pp. 513–514). □

Notation

Throughout the next subsections we assume to be given i.i.d. random variables $(X_j)_{j\in\mathbb{N}}$, following a distribution F with characteristic function φ. If F has a density we denote it by f. If the kth moment exists we denote it by $\mu_k = \mathbb{E}[X_1^k]$, $k \in \mathbb{N}_0$. We always assume that $\mu_1 = 0$ and $\sigma^2 = \mu_2 > 0$. For every $n \in \mathbb{N}$ define the normalized sum

$$S_n := \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} X_j\,. \tag{2.23}$$

Every S_n has the distribution function $F_n(x) = F^{n\star}(x\sigma\sqrt{n})$ and we can conclude from Lemma 2.21 and Lemma 2.26 that its characteristic function is $\zeta \mapsto \varphi_n(\zeta) = \varphi^n(\frac{\zeta}{\sigma\sqrt{n}})$. If F_n has a density we denoted it by f_n.

Like always we denote a standard normal density by $\mathfrak{n}$ and a standard normal distribution function by $\mathfrak{N}$. The Fourier transform of $\mathfrak{n}$ is the characteristic function $\zeta \mapsto e^{-\frac{1}{2}\zeta^2}$.

Central limit theorem

The theorems in the following are based on the classical central limit theorem, which we want to present here.

Theorem 2.29. *Let X_j, $j \in \mathbb{N}$, be i.i.d. random variables. Assume that $\mu = \mathbb{E}[X_1]$ and $\sigma^2 = \mathbb{V}\mathrm{ar}[X_1]$ exist. Then*

$$\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n} X_j - \mu\right) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, \sigma^2\big), \quad n \to \infty\,.$$

Proof. We sketch the proof from Williams (1991, p. 189) as it gives a nice intuition why we get exactly the normal distribution in the limit.

We start with Taylor expansion of the function $g(x) = e^{ix}$. For every $n \in \mathbb{N}$ the remainder term is

$$R_n(x) = e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!}, \quad x \in \mathbb{R}\,.$$

Using induction it can be shown that (see Williams, 1991, p. 188)

$$|R_n(x)| \le \min\Big(\frac{2\,|x|^n}{n!}, \frac{|x|^{n+1}}{(n+1)!}\Big), \quad x \in \mathbb{R}\,. \tag{2.24}$$

Assume that Z is a random variable with zero expectation, variance $s^2 > 0$ and with characteristic function γ. Then we get for $\theta \in \mathbb{R}$

$$\begin{aligned}\mathbb{E}[\,R_2(\theta Z)\,] &= \mathbb{E}[\,e^{i\theta Z}\,] - 1 - \mathbb{E}[\,i\theta Z\,] + \mathbb{E}\Big[\,\frac{\theta^2 Z^2}{2}\,\Big]\\ &= \gamma(\theta) - 1 + \frac{\theta^2 s^2}{2}\,.\end{aligned}$$

The remainder term can be estimated using (2.24):

$$\big|\mathbb{E}[\,R_2(\theta Z)\,]\big| \le \mathbb{E}\big[\,|R_2(\theta Z)|\,\big] \le \theta^2\,\mathbb{E}\Big[\,|X|^2 \wedge \frac{|\theta|\,|X|^3}{6}\,\Big]\,,$$

where the argument of the last expectation is dominated by the integrable random variable $|X|^2$ and tends to zero for $\theta \to 0$. Hence by dominated convergence

$$\gamma(\theta) = 1 - \frac{\theta^2 s^2}{2} + o(\theta^2) \quad \text{as } \theta \to 0\,. \tag{2.25}$$

Assume now that the variance σ^2 of our random variables X_j is greater than zero, as for $\sigma^2 = 0$ the assertion is trivially true. We look at the random variables

$$G_n = \frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} (X_j - \mu)$$

and want to show that their distributions converge weakly to the standard normal distribution. If we denote the characteristic function of $X_1 - \mu$ by $\bar{\varphi}$, we know that $\sum_{j=1}^{n}(X_j - \mu)$ has the characteristic function $\bar{\varphi}^n$, as the X_j are assumed to be i.i.d. The characteristic function φ_{G_n} of G_n becomes

$$\varphi_{G_n}(\theta) = \bar{\varphi}^n\Big(\frac{\theta}{\sigma\sqrt{n}}\Big), \quad \theta \in \mathbb{R}\,.$$

Using (2.25) we can say for every $\theta \in \mathbb{R}$

$$\varphi_{G_n}(\theta) = \bigg(1 - \frac{\theta^2}{2n} + o\Big(\frac{\theta^2}{\sigma^2\,n}\Big)\bigg)^n \quad \text{as } n \to \infty\,.$$

Now we use that (Williams, 1991, p. 188)

$$|\log(1+z) - z| \leq |z|^2, \quad |z| \leq \frac{1}{2}\,,$$

such that we know for every $\theta \in \mathbb{R}$

$$\begin{aligned}\log \varphi_{G_n}(\theta) &= n\,\log\Big(1 - \frac{\theta^2}{2n} + o\Big(\frac{\theta^2}{n}\Big)\Big)\\ &= n\left(-\frac{\theta^2}{2n} + o\Big(\frac{\theta^2}{n}\Big)\right) \to -\frac{\theta^2}{2} \quad \text{as } n \to \infty\,.\end{aligned}$$

So the characteristic functions of G_n converge to the characteristic function of the standard normal distribution and therefore, by uniqueness of the characteristic function (Theorem 2.22), we get the result. □

Expansions for densities

We want to strengthen the central limit theorem, which is possible if higher moments exist. With a further assumption on the integrability of the characteristic function φ of F we get the following expansion for densities:

Theorem 2.30. *Assume that $\mu_3 < \infty$ and that there exists an $\eta_0 \in \mathbb{N}$ such that*

$$\int_{\mathbb{R}} |\varphi(\zeta)|^{\eta_0}\, d\zeta < \infty\,. \tag{2.26}$$

Then f_n exists for $n \geq \eta_0$ and as $n \to \infty$

$$f_n(x) - \mathfrak{n}(x) - \frac{\mu_3}{6\sigma^3\sqrt{n}}\,(x^3 - 3x)\,\mathfrak{n}(x) = o\Big(\frac{1}{\sqrt{n}}\Big) \tag{2.27}$$

uniformly in $x \in \mathbb{R}$.

Proof. A proof can be found in Feller (1971, pp. 533–534). We just want to stress again the importance of assumption (2.26). As characteristic functions fulfil $|\varphi(\zeta)| \leq 1$ (see Lemma 2.21) it implies

$$\int_{\mathbb{R}} |\varphi^n(\zeta)|\, d\zeta < \infty \quad \text{for all } n \in \mathbb{N},\, n \geq \eta_0\,.$$

So knowing that F_n has the characteristic function $\varphi^n(\frac{\zeta}{\sigma\sqrt{n}})$, we can apply the Fourier inversion theorem 2.27 and get the existence of f_n for every $n \geq \eta_0$. □

Smoothing

The integrability condition (2.26) is essential for Theorem 2.30. To get a result for a more general framework we proceed indirectly.

For $T > 0$ we introduce the distribution function V_T with density

$$v_T(x) = \frac{1}{\pi} \frac{1 - \cos(Tx)}{Tx^2}, \quad x \in \mathbb{R},$$

and characteristic function w_T. This function will help us satisfying the integrability condition as we have

$$w_T(\zeta) = \begin{cases} 1 - \frac{|\zeta|}{T}, & \text{if } |\zeta| < T, \\ 0, & \text{if } |\zeta| \geq T. \end{cases} \tag{2.28}$$

We want to approximate functions by their convolutions with V_T. To shorten notation we denote such a convolution of an integrable or a bounded, measurable function g by

$${}^T g(t) = (V_T \star g)(t) = \int_{\mathbb{R}} g(t-x)\, v_T(x)\, dx, \quad t \in \mathbb{R}.$$

The next lemma looks at the difference of two functions, $\Delta = G - H$, and tells how the supremum of ${}^T\Delta$ limits the supremum of Δ.

Lemma 2.31. *Let G be a distribution function and H a differentiable function with bounded derivative h such that $\lim_{x\to-\infty} H(x) = 0$ and $\lim_{x\to\infty} H(x) = 1$. Take $T > 0$. Define $\Delta(x) = G(x) - H(x)$ and*

$$\xi = \sup_{x\in\mathbb{R}} |\Delta(x)|, \quad \xi_T = \sup_{x\in\mathbb{R}} |{}^T\Delta(x)|.$$

Then

$$\frac{\xi}{2} \leq \xi_T + \frac{12\, \|h\|_\infty}{\pi T}. \tag{2.29}$$

Proof. See Feller (1971, p. 537). □

So we are looking for an upper bound for the supremum of ${}^T\Delta$. Our argumentation works with the Fourier transforms. We assume that the derivative h of the function H has a Fourier transform η with $\eta(0) = 1$ and $\eta'(0) = 0$. This is clearly true for $H = \mathfrak{N}$. We then know from Lemma 2.25 that the convolution ${}^T h$ has the Fourier transform $w_T \cdot \eta$. We know from Lemma 2.24 that $V_T \star G$ is a distribution function and Lemma 2.26 tells that its characteristic function is $w_T \cdot \gamma$, where γ denotes the characteristic function of G. As the multiplication with w_T ensures the integrability of the characteristic function, we further get the existence of a density ${}^T g$ by the Fourier inversion theorem 2.27. So we have

$${}^T g(x) - {}^T h(x) = \frac{1}{2\pi} \int_{-T}^{T} e^{-i\zeta x} \big(\gamma(\zeta) - \eta(\zeta)\big)\, w_T(\zeta)\, d\zeta, \quad x \in \mathbb{R}.$$

Integrating with respect to x gives the anti-derivative

$$^T\Delta(x) = \frac{1}{2\pi}\int_{-T}^{T} e^{-i\zeta x}\,\frac{\gamma(\zeta)-\eta(\zeta)}{-i\zeta}\,w_T(\zeta)\,d\zeta, \quad x \in \mathbb{R}\,.$$

No integration constant appears because both sides tend to zero as $x \to \pm\infty$, the left because $\Delta(x) = G(x) - H(x) \to 0$, $x \to \pm\infty$, the right by the Riemann–Lebesgue lemma 2.28. Using the Fourier norm and applying (2.29) we come to the following smoothing inequality:

Lemma 2.32. *Let G be a distribution function with vanishing expectation and characteristic function γ. Let H be a function such that $G - H$ vanishes at $\pm\infty$. Assume further that H has a bounded derivative h. Finally, suppose that h has a continuously differentiable Fourier transform η such that $\eta(0) = 1$ and $\eta'(0) = 0$. Then for all $T > 0$*

$$\sup_{x\in\mathbb{R}} |G(x) - H(x)| \le \frac{1}{\pi}\int_{-T}^{T}\Big|\frac{\gamma(\zeta)-\eta(\zeta)}{\zeta}\Big|\,d\zeta + \frac{24\,\|h\|_\infty}{\pi T}\,. \tag{2.30}$$

This inequality is called smoothing inequality and will be used in the proofs of the expansion theorem and of the Berry–Esseen theorem in the next two subsections.

Expansions for distributions

Theorem 2.33. *Let F be a distribution function that is not concentrated on a set $\{b + zh \mid z \in \mathbb{Z}\}$ with $b, h \in \mathbb{R}$ (i.e. F is not a lattice distribution) such that the third moment μ_3 exists. Then as $n \to \infty$*

$$F_n(x) - \mathfrak{N}(x) - \frac{\mu_3}{6\sigma^3\sqrt{n}}\,(1-x^2)\,\mathfrak{n}(x) = o\Big(\frac{1}{\sqrt{n}}\Big) \tag{2.31}$$

uniformly in $x \in \mathbb{R}$.

Proof. The proof can be found in Feller (1971, p. 539). If F_n has a density, the assertion is clear from the expansion for densities (Theorem 2.27). In the case that no densities exist, the proof is based on the smoothing inequality (2.30), setting

$$H_n(x) = \mathfrak{N}(x) + \frac{\mu_3}{6\sigma^3\sqrt{n}}\,(1-x^2)\,\mathfrak{n}(x), \quad x \in \mathbb{R}\,.$$

□

Berry–Esseen theorems

Berry–Esseen theorems specify the rate of convergence to the normal distribution. A first version was shown independently by Berry (1941) and by Esseen (1942). Our presentation follows Feller (1971, pp. 543–544) and is slightly sharper than the original one, in the sense of having a smaller constant. In Remark 2.35 right after the theorem and its proof we will add a brief overview concerning the improvements of this constant.

Theorem 2.34. *For $n \in \mathbb{N}$ let X_i, $i = 1, \ldots, n$, be i.i.d. random variables with distribution function F and with*

$$\mathbb{E}[\, X_1 \,] = 0, \quad \sigma^2 = \mathbb{E}[\, X_1^2 \,] > 0, \quad \nu_3 = \mathbb{E}[\, |X_1|^3 \,] < \infty\,.$$

Denote the distribution of the normalized sum

$$\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i$$

by F_n. Then

$$\sup_{x\in\mathbb{R}} |F_n(x) - \mathfrak{N}(x)| \le \frac{3\,\nu_3}{\sigma^3\,\sqrt{n}}\,. \tag{2.32}$$

Proof. We want to apply the smoothing inequality (2.30) with $G = F_n$ and $H = \mathfrak{N}$. We choose

$$T = \frac{4}{3}\,\frac{\sigma^3}{\nu_3}\,\sqrt{n} \le \frac{4}{3}\,\sqrt{n}\,, \tag{2.33}$$

where the last inequality is due to Jensen. Since we know that $\mathfrak{n}(x) \le \frac{1}{\sqrt{2\pi}} < \frac{2}{5}$ for all $x \in \mathbb{R}$, the smoothing inequality (2.30) becomes

$$\pi \sup_{x\in\mathbb{R}} |F_n(x) - \mathfrak{N}(x)| \le \int_{-T}^{T} \Big|\varphi^n\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big) - e^{-\frac{1}{2}\zeta^2}\Big|\,\frac{d\zeta}{|\zeta|} + \frac{48}{5\,T}\,. \tag{2.34}$$

To find a bound for the integrand, we use the inequality

$$|\alpha^n - \beta^n| \le n\,|\alpha - \beta|\,\max\big(|\alpha|^{n-1}, |\beta|^{n-1}\big), \quad \alpha, \beta \in \mathbb{C}\,, \tag{2.35}$$

with $\alpha = \varphi(\frac{\zeta}{\sigma\sqrt{n}})$ and $\beta = e^{-\frac{\zeta^2}{2n}}$. We now find a bound for $\max(|\alpha|^{n-1}, |\beta|^{n-1})$. We apply the inequality

$$\Big|e^{it} - 1 - it + \frac{t^2}{2}\Big| \le \frac{|t|^3}{6}, \quad t \in \mathbb{R}\,,$$

(for a proof see e.g. Feller, 1971, p. 512) to get

$$\begin{aligned} |\varphi(t)| - |1 - \frac{1}{2}\,\sigma^2\,t^2| \le |\varphi(t) - 1 + \frac{1}{2}\,\sigma^2\,t^2| &= \Big|\int_{\mathbb{R}} \big(e^{itx} - 1 - itx + \frac{1}{2}\,t^2\,x^2\big)\,F(dx)\Big| \\ &\le \int_{\mathbb{R}} \frac{|tx|^3}{6}\,F(dx) = \frac{1}{6}\,\nu_3\,|t|^3, \quad t \in \mathbb{R}\,, \end{aligned} \tag{2.36}$$

and hence

$$|\varphi(t)| \le 1 - \frac{1}{2}\,\sigma^2 t^2 + \frac{1}{6}\,\nu_3\,|t|^3, \quad \text{if } \frac{1}{2}\,\sigma^2 t^2 \le 1\,.$$

For $t = \frac{\zeta}{\sigma\sqrt{n}}$ with $|\zeta| \le T$ this assumption is fulfilled, since by the choice of T in (2.33)

$$\frac{1}{2}\,\sigma^2 t^2 = \frac{\zeta^2}{2n} \le \frac{T^2}{2n} \le \frac{8}{9} < 1\,,$$

and so we can conclude that for all $|\zeta| \le T$

$$\Big|\varphi\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big)\Big| \le 1 - \frac{\zeta^2}{2n} + \frac{\nu_3}{6\sigma^3\sqrt{n^3}}\,|\zeta|^3 \le 1 - \frac{\zeta^2}{2n} + \frac{\nu_3}{6\sigma^3\sqrt{n^3}}\,\zeta^2\,T = 1 - \frac{5}{18n}\,\zeta^2 \le e^{-\frac{5}{18n}\,\zeta^2}\,.$$

Since $\sigma^3 \le \nu_3$ the assertion of the theorem is trivially true for $\sqrt{n} \le 3$ so we can assume $n \ge 10$ and get

$$\Big|\varphi\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big)\Big|^{n-1} \le e^{-\frac{1}{4}\,\zeta^2}, \quad |\zeta| \le T\,.$$

Hence the value $e^{-\frac{1}{4}\,\zeta^2}$ can be used as estimate for the maximum in (2.35). To get an estimate for $n\,|\alpha - \beta|$ in (2.35) we write

$$n\,\Big|\varphi\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big) - e^{-\frac{\zeta^2}{2n}}\Big| \le n\,\Big|\varphi\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big) - 1 + \frac{\zeta^2}{2n}\Big| + n\,\Big|1 - \frac{\zeta^2}{2n} - e^{-\frac{\zeta^2}{2n}}\Big|, \quad \zeta \in \mathbb{R}\,.$$

The first summand can be estimated using (2.36), for the second we use that

$$e^{-x} - 1 + x \le \frac{1}{2}\,x^2 \quad \text{for } x > 0$$

and get

$$n\,\Big|\varphi\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big) - e^{-\frac{\zeta^2}{2n}}\Big| \le \frac{\nu_3}{6\sigma^3\sqrt{n}}\,|\zeta|^3 + \frac{1}{8n^2}\,\zeta^4, \quad \zeta \in \mathbb{R}\,.$$

Using again the facts that $\sigma^3 \le \nu_3$, $n \ge 10$ and $|\zeta| \le T$ we get for the integrand in (2.34) that

$$\frac{1}{|\zeta|}\,\Big|\varphi^n\Big(\frac{\zeta}{\sigma\sqrt{n}}\Big) - e^{-\frac{1}{2}\zeta^2}\Big| \le \frac{1}{T}\,\Big(\frac{2}{9}\,\zeta^2 + \frac{1}{18}\,|\zeta|^3\Big)\,e^{-\frac{1}{4}\,\zeta^2}\,.$$

This can be integrated over $\zeta \in \mathbb{R}$. We get from (2.34)

$$\begin{aligned}
\pi\,\sup_{x\in\mathbb{R}}|F_n(x) - \mathfrak{N}(x)| &\le \int_{-T}^{T} \frac{1}{T}\,\Big(\frac{2}{9}\,\zeta^2 + \frac{1}{18}\,|\zeta|^3\Big)\,e^{-\frac{1}{4}\,\zeta^2}\,d\zeta + \frac{48}{5\,T} \\
&\le \frac{2}{T}\int_0^{\infty}\Big(\frac{2}{9}\,\zeta^2 + \frac{1}{18}\,|\zeta|^3\Big)\,e^{-\frac{1}{4}\,\zeta^2}\,d\zeta + \frac{10}{T} \\
&= \frac{8}{9\,T}\,(\sqrt{\pi} + 1) + \frac{10}{T} \\
&< \frac{4\pi}{T}
\end{aligned}$$

and as we chose $T = \frac{4}{3}\,\frac{\sigma^3}{\nu_3}\,\sqrt{n}$ this finishes the proof. □

Remark 2.35. In a more general way we can write inequality (2.32) as

$$\sup_{x\in\mathbb{R}}|F_n(x) - \mathfrak{N}(x)| \le \frac{C\,\nu_3}{\sigma^3\,\sqrt{n}}\,,$$

where $C > 0$ is a constant which is independent of F and n. The value of the smallest possible constant C_0 is not known. Esseen (1945, p. 43) stated $C_0 \leq 7.5$ as a first upper bound. In a later work (see Esseen, 1956) he showed the lower bound

$$C_0 \geq \frac{3+\sqrt{10}}{6\sqrt{2\pi}} \approx 0.4097\,.$$

There have been many improvements concerning the upper bound, including publications by van Beek (1972) showing $C_0 \leq 0.7975$ and Shiganov (1986) proving $C_0 \leq 0.7655$. Up to our knowledge the best current bound is $C_0 \leq 0.7056$, shown by Shevtsova (2007).

2.3.2 Asymptotics for U-statistics

As U-statistics are closely related to sums of i.i.d. random variables, it is possible to state similar results also for U-statistics. We will first present the central limit theorem and then show two Berry–Esseen theorems.

Recall that a U-statistic $\hat{U}_n^m(\kappa)$ with kernel κ of degree m can be decomposed into a sum over U-statistics of degree $1, 2, \ldots, m$, called Hoeffding representation (see Section 2.1.3). The kernels of the U-statistics in the Hoeffding representations, as defined in (2.3), are called canonical functions and we denote them by h_c.

Central limit theorem

A central limit theorem for U-statistics was first proved by Hoeffding (1948). He showed the asymptotic normality under the condition

$$\mathbb{E}[\,\kappa^2(Z_1, \ldots, Z_m)\,] < \infty\,.$$

This assumption can be relaxed as shown in the following theorem.

Theorem 2.36. *Assume that $\hat{U}_n^m(\kappa)$ is a non-degenerate U-statistic, based on i.i.d. random variables, with a kernel κ of degree m. Assume that the canonical functions h_c, as defined in* (2.3)*, satisfy the moment conditions*

$$\mathbb{E}[\,|h_c(Z_1, \ldots, Z_c)|^{\gamma_c}\,] < \infty, \quad c = 1, \ldots, m\,,$$

where $\gamma_c = \frac{2c}{2c-1}$. Then

$$\sqrt{n}\,\big(\hat{U}_n^m(\kappa) - \theta\big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, m^2\,\sigma_1^2(\kappa)\big), \quad n \to \infty\,,$$

with

$$\sigma_1^2(\kappa) = \mathbb{E}[\,h_1^2(Z_1)\,]\,. \tag{2.37}$$

Proof. See Koroljuk and Borovskich (1994, pp. 129–131). □

Berry–Esseen theorems

The first general Berry–Esseen theorem stating the convergence rate $n^{-1/2}$ for U-statistics of degree 2, not just for selected examples, was published by Bickel (1974), assuming that the kernel is bounded. Over the years the assumptions have been successively reduced. Bentkus et al. (1994) showed that the weakest possible conditions are

$$\mathbb{E}\big[\,|h_1(Z_1)|^3\,\big] < \infty \quad \text{and} \quad \mathbb{E}\big[\,|\kappa(Z_1, Z_2)|^{\frac{5}{3}}\,\big] < \infty\,.$$

Berry–Esseen theorems under these conditions are e.g. stated by Friedrich (1989) (for degree 2) and Korolyuk and Borovskikh (1985) (for general degree $m \in \mathbb{N}$).

As the degree of the U-statistic does not play an important role, most publications assume $m = 2$. A good overview over the existing literature and an extension of the optimal bounds to higher-degree U-statistics is given by Bentkus et al. (2009).

We want to show two examples of Berry–Esseen theorems for U-statistics. Throughout this section assume that the asymptotic variance of the U-statistic, as defined in (2.37), is greater than zero, $\sigma_1^2(\kappa) > 0$. To formulate the other assumptions we will use the notation from the Hoeffding representation (Section 2.1.3) and further denote

$$\lambda^{(1)} = \frac{1}{\sigma_1^3(\kappa)}\,\mathbb{E}[\,|h_1(Z_1)|^3\,] \tag{2.38}$$

and

$$\lambda_p^{(2)} = \frac{1}{\sigma_1^p(\kappa)}\,\mathbb{E}[\,|h_2(Z_1, Z_2)|^p\,], \quad 0 \le p \le \frac{5}{3}\,. \tag{2.39}$$

In most of the proofs the Hoeffding representation plays an important role. As its first term is a normalized sum of the i.i.d. random variables $h_1(X_i)$, we can apply the classical Berry–Esseen theorem 2.34, and so we expect a summand of the form $\frac{1}{\sqrt{n}}\,C\,\lambda^{(1)}$ with some constant $C > 0$ in the right-hand side of the Berry–Esseen inequality.

We start with a theorem given in Korolyuk and Borovskikh (1985). It states a Berry–Esseen theorem for U-statistics of general degree m and gives a concrete bound in the case $m = 2$.

Theorem 2.37. *Let $\hat{U}_n^m(\kappa)$ be a non-degenerate U-statistic with kernel κ of degree m, based on i.i.d. random variables Z_i, $i \in \mathbb{N}$. Assume $\sigma_1^2(\kappa) > 0$ and further*

$$\mathbb{E}\big[\,|h_1(Z_1)|^3\,\big] < \infty \quad \textit{and} \quad \mathbb{E}\big[\,|\kappa(Z_1, \ldots, Z_m)|^{\frac{5}{3}}\,\big] < \infty\,.$$

Then there exists a positive constant C_κ such that for all $n \ge m+1$

$$\sup_{x \in \mathbb{R}} \Big|\mathbb{P}\Big[\,\frac{\sqrt{n}}{2\,\sigma_1(\kappa)}\,\big(\hat{U}_n^m(\kappa) - \theta\big) \le x\,\Big] - \mathfrak{N}(x)\Big| \le \frac{C_\kappa}{\sqrt{n}}\,.$$

In the case $m = 2$ we have

$$C_\kappa \le 12!\,\big(\lambda^{(1)} + \lambda_{5/3}^{(2)}\big)\,.$$

The second version of a Berry–Esseen theorem for U-statistics we want to present is from Lee (1990, pp. 97–105). It gives a more refined version and is an adaption of the proof in Friedrich (1989).

Theorem 2.38. *Let $\hat{U}_n^2(\kappa)$ be a non-degenerate U-statistic with kernel κ of degree 2, based on i.i.d. random variables Z_i, $i \in \mathbb{N}$. Assume $\sigma_1^2(\kappa) > 0$ and further*

$$\mathbb{E}\big[\, |h_1(Z_1)|^3 \,\big] < \infty \quad \textit{and} \quad \mathbb{E}\big[\, |h_2(Z_1, Z_2)|^{\frac{5}{3}} \,\big] < \infty\,.$$

Then there exist constants C_1, C_2 and C_3 depending neither on n, κ nor on the distribution of Z_1 such that for all $n \geq 2$

$$\sup_{x\in\mathbb{R}} \left| \mathbb{P}\Big[\frac{\sqrt{n}}{2\,\sigma_1(\kappa)} \big(\hat{U}_n^2(\kappa) - \theta \big) \leq x \Big] - \mathfrak{N}(x) \right| \leq \frac{1}{\sqrt{n}} \Big(C_1\, \lambda^{(1)} + C_2\, \lambda^{(2)}_{5/3} + C_3 \left(\lambda^{(1)}\, \lambda^{(2)}_{3/2} \right)^{\frac{2}{3}} \Big)\,.$$

2.4 Degenerate tau-kernel

Within this section we have a closer look at the cases where the tau-kernel is degenerate, as defined in Section 2.1.2. Recall that for a given probability measure $\mu \in \mathcal{M}_1(\mathbb{R}^2)$ the tau-kernel κ_τ, as defined in (2.5), is called μ-degenerate, if for independent $Z, \tilde{Z} \sim \mu$

$$\mathbb{E}[\, \kappa_\tau(Z, \tilde{Z}) \mid Z \,] = \mathbb{E}[\, \kappa_\tau(Z, \tilde{Z}) \,] = \tau \quad \mu\text{-a.s.} \tag{2.40}$$

Note that in the case of degeneracy this equation is not only true for $Z \sim \mu$, but also for all Z that are independent of $\tilde{Z}$ and whose distribution is absolutely continuous w.r.t. μ, $\mathcal{L}(Z) \ll \mu$.

Within our proofs we will also use an equivalent formulation to assure degeneracy. The tau-kernel is μ-degenerate if and only if the following property is fulfilled for μ-almost all $(x, y) \in \mathbb{R}^2$:

$$\mathbb{E}\big[\, \kappa_\tau\big((x, y), (\tilde{X}, \tilde{Y})\big) \,\big] = \mathbb{E}\big[\, \operatorname{sgn}(x - \tilde{X})\, \operatorname{sgn}(y - \tilde{Y}) \,\big] = \tau\,, \tag{2.41}$$

where $(\tilde{X}, \tilde{Y}) \sim \mu$.

The next lemma shows that the degeneracy of the tau-kernel is equivalent to a degenerate asymptotic distribution, i.e. that the asymptotic variance σ_τ^2, as defined in (2.15), equals zero.

Lemma 2.39. *The tau-kernel κ_τ is μ-degenerate if and only if the corresponding asymptotic distribution is degenerate, i.e. $\sigma_\tau^2 = 0$.*

Proof. As we have

$$\sigma_\tau^2 = 4\ \mathbb{V}\mathrm{ar}\Big[\mathbb{E}\big[\, \kappa_\tau\big((X, Y), (\tilde{X}, \tilde{Y})\big) \mid X, Y \,\big] \Big]$$

for independent $(X, Y), (\tilde{X}, \tilde{Y}) \sim \mu$, we know that

$$\sigma_\tau^2 = 0 \quad \Leftrightarrow \quad \mathbb{E}\big[\, \kappa_\tau\big((X, Y), (\tilde{X}, \tilde{Y})\big) \mid X, Y \,\big] = \text{const.} \quad \mu\text{-a.s.}$$

□

In the following we will provide some statements how the property of degeneracy or non-degeneracy changes under transformations of the two-dimensional distribution. Therefore we need the following lemma that enables us to change the sigma-algebra on which we condition.

Lemma 2.40. *Let $(\Omega, \mathfrak{F}, \mathbb{P})$ be a probability space and let $(\mathcal{Z}_1, \mathfrak{Z}_1)$, $(\mathcal{Z}_2, \mathfrak{Z}_2)$, and $(\mathcal{Z}_3, \mathfrak{Z}_3)$ be measurable spaces. Let $g : \mathcal{Z}_1 \to \mathcal{Z}_2$ and $h : \mathcal{Z}_2 \times \mathcal{Z}_3 \to \mathbb{R}$ be measurable functions, h bounded, and let $X : \Omega \to \mathcal{Z}_1$ and $Y : \Omega \to \mathcal{Z}_3$ be independent random variables. Then, for every sigma-algebra $\mathfrak{A} \subset \mathfrak{F}$ that is independent of Y and fulfils $\sigma(X) \subset \mathfrak{A}$,*

$$\mathbb{E}\big[\, h\big(g(X), Y\big) \mid \mathfrak{A} \,\big] \overset{\text{a.s.}}{=} \mathbb{E}\big[\, h\big(g(X), Y\big) \mid \sigma\big(g(X)\big) \,\big] \,.$$

Proof. The proof is based on the monotone-class theorem, see e.g. Williams (1991, pp. 205–206).

Denote the class of functions

$$\mathcal{H} = \Big\{ h : \mathcal{Z}_2 \times \mathcal{Z}_3 \to \mathbb{R} \text{ measurable and bounded } \Big| \\ \mathbb{E}\big[\, h\big(g(X), Y\big) \mid \mathfrak{A} \,\big] \overset{\text{a.s.}}{=} \mathbb{E}\big[\, h\big(g(X), Y\big) \mid \sigma\big(g(X)\big) \,\big] \Big\} \,.$$

Then $\mathcal{H}$ is a vector space due to linearity of the conditional expectation. It also contains the constant function 1 and further contains the limit of a sequence $(h_n)_{n \in \mathbb{N}}$ of non-negative functions in $\mathcal{H}$ with $h_n \nearrow h$, h bounded, due to monotone convergence for conditional expectations.

Now consider functions $k : \mathcal{Z}_2 \times \mathcal{Z}_3 \to \mathbb{R}$ that are measurable and bounded, and that can be written as $k(x, y) = u(x)\, v(y)$ with $u : \mathcal{Z}_2 \to \mathbb{R}$ and $v : \mathcal{Z}_3 \to \mathbb{R}$ measurable and bounded. Note that Y is independent of X and therefore also of $g(X)$, and by definition it is independent of $\mathfrak{A}$. Further we know that $u(g(X))$ is $\mathfrak{A}$-measurable and $\sigma(g(X))$-measurable. So we have

$$\begin{aligned} \mathbb{E}\big[\, k\big(g(X), Y\big) \mid \mathfrak{A} \,\big] &= \mathbb{E}\big[\, u\big(g(X)\big)\, v(Y) \mid \mathfrak{A} \,\big] = u\big(g(X)\big)\, \mathbb{E}\big[\, v(Y) \mid \mathfrak{A} \,\big] \\ &= u\big(g(X)\big)\, \mathbb{E}\big[\, v(Y) \mid \sigma\big(g(X)\big) \,\big] \\ &= \mathbb{E}\big[\, u\big(g(X)\big)\, v(Y) \mid \sigma\big(g(X)\big) \,\big] \\ &= \mathbb{E}\big[\, k\big(g(X), Y\big) \mid \sigma\big(g(X)\big) \,\big] \quad \text{a.s.}, \end{aligned}$$

and therefore we know that $k \in \mathcal{H}$. This holds especially true for u and v being indicator functions 1_A, $A \in \mathfrak{Z}_2$, and 1_B, $B \in \mathfrak{Z}_3$, respectively. As the family of such subsets $A \times B$ builds a π-system on the product space $\mathcal{Z}_2 \times \mathcal{Z}_3$ which generates its sigma-algebra, we get the result. $\square$

Remark 2.41. We will use this lemma twice in the next subsection and also in Section 4.1. The function h will be the tau-kernel κ_τ, which is measurable and bounded. The random variable X will take values in $\mathbb{R}$ or $\mathbb{R}^2$, the function g will map to $\mathbb{R}^2$, Y will be an independent copy of $g(X)$, and we take $\mathfrak{A} = \sigma(X)$.

2.4.1 Conditions for degeneracy

Lemma 2.42. *If $|\tau| = 1$, then the tau-kernel κ_τ is degenerate.*

Proof. If $\tau = \pm 1$, then we know from Remark 2.10 that $\hat{\tau}_n = \pm 1$ for every $n \geq 2$. So the variance of the estimator for every $n \geq 2$ and also the asymptotic variance equal zero, which implies a degenerate tau-kernel by Lemma 2.39. □

As the absolute value of Kendall's tau itself remains unchanged under strictly monotone transformations of the marginal distributions so does the property of degeneracy of the tau-kernel:

Lemma 2.43. *Let the tau-kernel κ_τ be μ-degenerate and let f and g be monotone functions such that for independent $(X, Y), (\tilde{X}, \tilde{Y}) \sim \mu$*

$$\mathbb{P}[\, f(X) = f(\tilde{X}),\, X \neq \tilde{X}\,] = 0 \quad \textit{and} \quad \mathbb{P}[\, g(Y) = g(\tilde{Y}),\, Y \neq \tilde{Y}\,] = 0\,. \tag{2.42}$$

Let μ' denote the distribution of the vector $(f(X), g(Y))$. Then the tau-kernel is also μ'-degenerate.

Proof. As f is monotone and fulfils (2.42), we have for independent $(X, Y), (\tilde{X}, \tilde{Y}) \sim \mu$

$$\operatorname{sgn}\big(f(X) - f(\tilde{X})\big) \overset{\text{a.s.}}{=} \begin{cases} \operatorname{sgn}(X - \tilde{X})\,, & \text{if } f \text{ is increasing,} \\ -\operatorname{sgn}(X - \tilde{X})\,, & \text{if } f \text{ is decreasing.} \end{cases}$$

As this holds similarly for g we get

$$\begin{aligned} &\kappa_\tau\Big(\big(f(X), g(Y)\big), \big(f(\tilde{X}), g(\tilde{Y})\big)\Big) \\ &\quad \overset{\text{a.s.}}{=} \begin{cases} \kappa_\tau\big((X, Y), (\tilde{X}, \tilde{Y})\big)\,, & \text{if } f \text{ and } g \text{ are both increasing or both decreasing,} \\ -\kappa_\tau\big((X, Y), (\tilde{X}, \tilde{Y})\big)\,, & \text{otherwise.} \end{cases} \end{aligned}$$

From Lemma 2.40 we know that we can equivalently condition on $(f(X), g(Y))$ or on (X, Y), so equation (2.40) also holds for the transformed random variables. □

Lemma 2.44. *Let X be a real-valued random variable and let f and g be monotone functions such that*

$$\mathbb{P}[\, f(X) = f(\tilde{X}),\, X \neq \tilde{X}\,] = 0 \quad \textit{and} \quad \mathbb{P}[\, g(X) = g(\tilde{X}),\, X \neq \tilde{X}\,] = 0\,, \tag{2.43}$$

where $\tilde{X}$ is an independent copy of X. Let μ denote the distribution of $(f(X), g(X))$. Then the tau-kernel κ_τ is μ-degenerate if and only if one of the following properties is satisfied:

(i) *X is continuously distributed;*

(ii) *X is uniformly distributed on a finite, non-empty set.*

Proof. As in the previous proof we know that, since f and g are monotone functions and fulfil (2.43), we have for X and an independent copy $\tilde{X}$

$$\operatorname{sgn}\big(f(X) - f(\tilde{X})\big) \operatorname{sgn}\big(g(X) - g(\tilde{X})\big) \overset{\text{a.s.}}{=} s_{fg} \operatorname{sgn}(X - \tilde{X}) \operatorname{sgn}(X - \tilde{X}),$$

where $s_{fg} = 1$ if f and g are both increasing or both decreasing and $s_{fg} = -1$ otherwise. Using Lemma 2.40 we can equivalently condition on $(f(X), g(X))$ or on X, and hence

$$\begin{aligned}
&\mathbb{E}\big[\operatorname{sgn}\big(f(X) - f(\tilde{X})\big) \operatorname{sgn}\big(g(X) - g(\tilde{X})\big) \mid f(X), g(X)\big] \\
&\quad = s_{fg}\, \mathbb{E}[\operatorname{sgn}(X - \tilde{X}) \operatorname{sgn}(X - \tilde{X}) \mid X] \\
&\quad = s_{fg}\, \mathbb{P}[X \neq \tilde{X} \mid X] \quad \text{a.s.}
\end{aligned}$$

So the tau-kernel is μ-degenerate if and only if

$$\mathbb{P}[X \neq \tilde{X} \mid X] = c \quad \text{a.s.} \tag{2.44}$$

for a constant $c \in [0, 1]$. We have to distinguish between two cases:

- If $c = 1$, then the random variable X fulfils (2.44) if and only if it is continuously distributed.
- If $c \in [0, 1)$, then the distribution of X has atoms. The distribution can even not have any continuous parts, as this would imply $c = 1$ for the continuous areas. Since $\mathbb{P}[X \neq \tilde{X} \mid X = x] = c$ must hold for all atoms x, we further know that they must have the same probability, so the set of all atoms must be finite. In total this means that the random variable X fulfils (2.44) with $c \in [0, 1)$ if and only if it is uniformly distributed on a finite, non-empty set.

□

Remark 2.45. Assumptions (2.42) and (2.43) are especially fulfilled if f and g are strictly monotone functions.

We now give two further examples of distributions that lead to a degenerate tau-kernel. The first example tells that the tau-kernel is degenerate if one variable has a degenerate distribution.

Example 2.46. Consider the random vector (X, c) where $c \in \mathbb{R}$. Then the tau-kernel is degenerate for every distribution of X as

$$\mathbb{E}[\operatorname{sgn}(X - \tilde{X}) \operatorname{sgn}(c - c) \mid X] = 0 \quad \text{a.s.}$$

The second example considers a distribution where the random tuple (X, Y) can just take values on a cross, and two opposite branches have each the same probability.

Example 2.47. Consider the random vector (X, Y) where there exist $x, y \in \mathbb{R}$ such that $\mathbb{P}[\,(X-x)\,(Y-y) = 0\,] = 1$ and additionally $\mathbb{P}[\,X > x\,] = \mathbb{P}[\,X < x\,]$ and $\mathbb{P}[\,Y > y\,] = \mathbb{P}[\,Y < y\,]$. For this random vector the tau-kernel is degenerate as

$$\begin{aligned}
&\mathbb{E}[\,\operatorname{sgn}(X-\tilde{X})\,\operatorname{sgn}(Y-\tilde{Y}) \mid X, Y\,] \\
&\quad = \mathbb{E}\big[\,\operatorname{sgn}\big((X-x)-(\tilde{X}-x)\big)\,\operatorname{sgn}\big((Y-y)-(\tilde{Y}-y)\big) \mid X, Y\,\big] \\
&\quad = \mathbb{E}\big[\,\operatorname{sgn}\big(-(\tilde{X}-x)\big)\,\operatorname{sgn}\big((Y-y)-(\tilde{Y}-y)\big) \mid Y\,\big]\,1_{\{X=x, Y\neq y\}} \\
&\qquad + \mathbb{E}\big[\,\operatorname{sgn}\big((X-x)-(\tilde{X}-x)\big)\,\operatorname{sgn}\big(-(\tilde{Y}-y)\big) \mid X\,\big]\,1_{\{X\neq x, Y=y\}} \\
&\qquad + \mathbb{E}[\,\operatorname{sgn}(\tilde{X}-x)\,\operatorname{sgn}(\tilde{Y}-y)\,]\,1_{\{X=x, Y=y\}} \\
&\quad = \mathbb{E}\big[\,1_{\{\tilde{X}\neq x\}}\,\operatorname{sgn}\big(-(\tilde{X}-x)\big)\,\operatorname{sgn}(Y-y) \mid Y\,\big]\,1_{\{X=x, Y\neq y\}} \\
&\qquad + \mathbb{E}\big[\,\operatorname{sgn}(X-x)\,1_{\{\tilde{Y}\neq y\}}\,\operatorname{sgn}\big(-(\tilde{Y}-y)\big) \mid X\,\big]\,1_{\{X\neq x, Y=y\}} \\
&\quad = 0 \quad \text{a.s.}\,,
\end{aligned}$$

where the last equation holds as $\mathbb{E}[\,\operatorname{sgn}(\tilde{X}-x)\,] = \mathbb{E}[\,\operatorname{sgn}(\tilde{Y}-y)\,] = 0$.

2.4.2 Conditions for non-degeneracy

In the previous subsection we saw examples of distributions that imply a degenerate tau-kernel. Now we look for assumptions that assure a non-degenerate tau-kernel. We will show in Corollary 2.51 that, loosely speaking, if there exists a rectangle in $\mathbb{R}^2$ where the measure is strictly positive, then it is already clear that the tau-kernel is not degenerate. There is even a more general statement, proved in Lemma 2.49. For the proof we need the following technical lemma:

Lemma 2.48. *For $u_0, u_1, v_0, v_1 \in \mathbb{R}$ with $u_0 < u_1$ and $v_0 < v_1$ we have*

$$\begin{aligned}
&\sum_{i,j=0}^{1} (-1)^{i+j}\,\kappa_\tau\big((u_i, v_j), \cdot\big) \\
&\quad = 4 \cdot 1_{(u_0,u_1)\times(v_0,v_1)} + 2 \cdot 1_{\{u_0,u_1\}\times(v_0,v_1)} + 2 \cdot 1_{(u_0,u_1)\times\{v_0,v_1\}} + 1_{\{u_0,u_1\}\times\{v_0,v_1\}} \\
&\quad \geq 1_{[u_0,u_1]\times[v_0,v_1]}\,,
\end{aligned} \tag{2.45}$$

with κ_τ given by (2.5).

Proof. The tau-kernel can be written for all $u, v, x, y \in \mathbb{R}$ as

$$\begin{aligned}
\kappa_\tau\big((u,v),(x,y)\big) &= \operatorname{sgn}(u-x)\,\operatorname{sgn}(v-y) \\
&= 1_{(u,\infty)\times(v,\infty)}(x,y) + 1_{(-\infty,u)\times(-\infty,v)}(x,y) \\
&\qquad - 1_{(-\infty,u)\times(v,\infty)}(x,y) - 1_{(u,\infty)\times(-\infty,v)}(x,y)\,.
\end{aligned} \tag{2.46}$$

Taking the first summand in (2.46) and building the alternating sum by inserting the different values of (u_i, v_j), as needed in the left-hand side of (2.45), we get

$$\begin{aligned}
&1_{(u_0,\infty)\times(v_0,\infty)} + 1_{(u_1,\infty)\times(v_1,\infty)} - 1_{(u_0,\infty)\times(v_1,\infty)} - 1_{(u_1,\infty)\times(v_0,\infty)} \\
&\quad = 1_{(u_0,u_1)\times(v_0,v_1)} + 1_{(u_0,u_1)\times\{v_1\}} + 1_{\{u_1\}\times(v_0,v_1)} + 1_{\{u_1\}\times\{v_1\}}\,.
\end{aligned}$$

If we do this for all the four summands and sum up, we get four times the interior $1_{(u_0,u_1)\times(v_0,v_1)}$, twice the edges $1_{(u_0,u_1)\times\{v_i\}}$ and $1_{\{u_i\}\times(v_0,v_1)}$, $i \in \{0,1\}$, and once every corner $1_{\{u_i\}\times\{v_j\}}$, $i,j \in \{0,1\}$, and therefore the result. $\square$

Lemma 2.49. *Assume that κ_τ is μ-degenerate. Consider $u_0, u_1, v_0, v_1 \in \mathbb{R}$ with $u_0 < u_1$ and $v_0 < v_1$. If there exists a random vector (U_0, U_1, V_0, V_1) such that the distribution of (U_i, V_j) is absolutely continuous w.r.t. μ for all $i, j \in \{0,1\}$, $\mathcal{L}(U_i, V_j) \ll \mu$, as well as*

$$\mathbb{P}[\, U_0 \le u_0, U_1 \ge u_1, V_0 \le v_0, V_1 \ge v_1 \,] > 0\,, \tag{2.47}$$

then

$$\mu\big([u_0, u_1] \times [v_0, v_1]\big) = 0\,. \tag{2.48}$$

Proof. We want to denote $A := \{U_0 \le u_0, U_1 \ge u_1, V_0 \le v_0, V_1 \ge v_1\}$. Since $\mathcal{L}(U_i, V_j) \ll \mu$ for all $i, j \in \{0,1\}$ we know from equation (2.40) and the comment afterwards that

$$\mathbb{E}\big[\, \kappa_\tau\big((U_i, V_j), (\tilde{X}, \tilde{Y})\big) \,\big] = \tau \quad \text{for all } i, j \in \{0,1\}$$

with $(\tilde{X}, \tilde{Y}) \sim \mu$ independent of (U_i, V_j). Using (2.45) we get

$$\begin{aligned} 0 \;&=\; \sum_{i,j=0}^{1} (-1)^{i+j}\, \mathbb{E}\big[\, \kappa_\tau\big((U_i, V_j), (\tilde{X}, \tilde{Y})\big) \,\big] = \mathbb{E}\Big[\, \sum_{i,j=0}^{1} (-1)^{i+j}\, \kappa_\tau\big((U_i, V_j), (\tilde{X}, \tilde{Y})\big) \,\Big] \\ &\overset{(2.45)}{\ge}\; \mathbb{E}[\, 1_{[U_0,U_1]\times[V_0,V_1]}(\tilde{X}, \tilde{Y}) \,] \ge \mathbb{E}[\, 1_A\, 1_{[U_0,U_1]\times[V_0,V_1]}(\tilde{X}, \tilde{Y}) \,] \\ &\ge\; \mathbb{E}[\, 1_A\, 1_{[u_0,u_1]\times[v_0,v_1]}(\tilde{X}, \tilde{Y}) \,] = \mathbb{P}[A]\, \mu\big([u_0, u_1] \times [v_0, v_1]\big)\,. \end{aligned}$$

As we assumed $\mathbb{P}[A] > 0$ we know that $\mu\big([u_0, u_1] \times [v_0, v_1]\big) = 0$. $\square$

Example 2.50. Assume there exist $u_0, u_1, v_0, v_1 \in \mathbb{R}$ with $u_0 < u_1$ and $v_0 < v_1$ such that $\mu\big(\{(u_i, v_j)\}\big) > 0$ for all $i, j \in \{0,1\}$. Then κ_τ is not μ-degenerate, because $(U_0, U_1, V_0, V_1) \equiv (u_0, u_1, v_0, v_1)$ satisfies the assumptions of Lemma 2.49, but

$$\mu\big([u_0, u_1] \times [v_0, v_1]\big) \ge \sum_{i,j=0}^{1} \mu\big(\{(u_i, v_j)\}\big) > 0\,.$$

Corollary 2.51. *Assume there exists a closed rectangle $I \times J \subset \mathbb{R}^2$ with non-empty interior and $\delta > 0$ such that*

$$\mu(B) \ge \delta\, \lambda_2(B)$$

for all Borel subsets B of $I \times J$, where λ_2 denotes the Lebesgue measure on $\mathbb{R}^2$. Then the tau-kernel κ_τ is not μ-degenerate.

Proof. There exist $u_0, u_1 \in I$, $v_0, v_1 \in J$ and $\varepsilon > 0$ such that $u_0 < u_1$, $v_0 < v_1$ and $I = [u_0 - \varepsilon, u_1 + \varepsilon]$, $J = [v_0 - \varepsilon, v_1 + \varepsilon]$.

Apply Lemma 2.49 with (U_0, U_1, V_0, V_1) uniformly distributed on $[u_0 - \varepsilon, u_0] \times [v_0 - \varepsilon, v_0] \times [u_1, u_1 + \varepsilon] \times [v_1, v_1 + \varepsilon]$ to obtain the contradiction $\mu([u_0, u_1] \times [v_0, v_1]) = 0$. $\square$

Remark 2.52. The corollary applies in particular when μ has a density f which is continuous at a point $x \in \mathbb{R}^2$ with $f(x) > 0$.

2.4.3 Invariance of degeneracy under linear transformations

The property of degeneracy or non-degeneracy does not necessarily stay unchanged under linear transformation. It does, if the transformation matrix is diagonal or anti-diagonal with non-zero diagonal or anti-diagonal entries, respectively. In the case of degeneracy this is known from Lemma 2.43, and for non-degeneracy the arguments of the proof can be similarly used. If the matrix does not have this special shape, then we can find examples where degeneracy or non-degeneracy is destroyed.

Example 2.53. Assume that the transformation matrix $\boldsymbol{A}$ has a column where both entries are non-zero and assume w.l.o.g. that it is the first one. Consider the random vector $(X, 0)^{\mathrm{t}}$ where X is any real-valued random variable. We know degeneracy of the tau-kernel in this situation from Example 2.46. The transformed random vector is $(a_{11}X, a_{21}X)^{\mathrm{t}}$, which is only degenerate if X is continuously distributed or uniformly distributed on a finite non-empty set (see Lemma 2.44).

Example 2.54. Assume that in one row of the transformation matrix $\boldsymbol{A}$ both entries are zero. Then the transformed random tuple has a degenerate tau kernel for every original distribution, as one component equals zero.

A last example shows how rotation can change the property of degeneracy.

Example 2.55. Example 2.50 shows that the distribution that is uniformly distributed on the corners of a square has no degenerate kernel. But if we rotate the square by $\frac{\pi}{4}$, then we get a distribution that is uniformly distributed on $\{(c, 0), (0, c), (-c, 0), (0, -c)\}$ for a $c > 0$, which is a distribution as described in Example 2.47 and therefore implies a degenerate kernel.

Chapter 3

Measures of dependence for elliptical distributions

Dealing with multi-dimensional distributions still holds problems. Therefore a widely used approach is to assume to have a normal distribution. The normal distribution is well studied and possesses a lot of nice properties, but it is very restrictive and often not suitable. So we follow the well-known approach of generalizing the normal distribution by introducing a class of distributions called elliptical distributions. The symmetric form of the distribution stays, but one is free e.g. to put more weight on the tails.

We first give the basic definitions and theorems. We then introduce the linear correlation coefficient as the natural measure of dependence in the case of elliptical distributions. The relationship between Kendall's tau and the linear correlation coefficient, which is one of the nice properties provided by elliptical distributions, leads to two different ways of estimating the dependence measure. We finally revisit the case of degenerate tau-kernels on the background of elliptical distributions.

3.1 Definitions and basic properties of elliptical distributions

We give a short introduction to elliptical distributions. The main definitions and theorems are given. For details and proofs see e.g. Cambanis et al. (1981), Fang et al. (1990) or McNeil et al. (2005, pp. 89–94).

Elliptical distributions are a generalization of the normal distribution. The name comes from the fact that, in case of existing continuous densities, the level curves of these densities are ellipses. A special class of elliptical distributions are spherical distributions where the level curves are circles. They are defined by their invariance under orthogonal transformations.

Notation. For $d \in \mathbb{N}$ let $\mathcal{O}(d)$ denote the set of all orthogonal matrices $\boldsymbol{U} \in \mathbb{R}^{d\times d}$, i.e. $\boldsymbol{U}\boldsymbol{U}^{\mathrm{t}} = \boldsymbol{U}^{\mathrm{t}}\boldsymbol{U} = \boldsymbol{I}_d$, where $\boldsymbol{I}_d$ denotes the $(d \times d)$-dimensional identity matrix.

Definition 3.1. An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^{\mathrm{t}}$ has a *spherical distribution* if for all $\boldsymbol{U} \in \mathcal{O}(d)$

$$\boldsymbol{U}\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{X} \,. \tag{3.1}$$

The characteristic function $\varphi_{\boldsymbol{X}}(\boldsymbol{t}) = \mathbb{E}[\, e^{i\boldsymbol{t}^{\mathrm{t}}\boldsymbol{X}}\,]$, $\boldsymbol{t} \in \mathbb{R}^d$, of a spherically distributed random vector has special properties. Note that if $\boldsymbol{X}$ is spherically distributed, then $\boldsymbol{X} \stackrel{\mathrm{d}}{=} -\boldsymbol{X}$, so both random vectors must have the same characteristic function. Since we further know that the characteristic function of $-\boldsymbol{X}$ is the complex conjugate of the characteristic function of $\boldsymbol{X}$ (see Lemma 2.21(iii) which similarly holds in higher dimensions), the characteristic function of spherically distributed random vectors must be real-valued. There is even a stronger property which is often used in the literature as definition of spherical distributions:

Theorem 3.2. *An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^{\mathrm{t}}$ has a spherical distribution if and only if its characteristic function φ satisfies one of the following equivalent conditions:*

(i) $\varphi(\boldsymbol{U}^{\mathrm{t}}\boldsymbol{t}) = \varphi(\boldsymbol{t})$ *for all* $\boldsymbol{U} \in \mathcal{O}(d)$ *and* $\boldsymbol{t} \in \mathbb{R}^d$*;*

(ii) *There exists a function* $\psi : [0, \infty) \to \mathbb{R}$ *such that* $\varphi(\boldsymbol{t}) = \psi(\boldsymbol{t}^{\mathrm{t}}\boldsymbol{t})$ *for all* $\boldsymbol{t} \in \mathbb{R}^d$.

Proof. See e.g. McNeil et al. (2005, pp. 89–90). □

Notation. The function ψ is called the *characteristic generator* of the spherical distribution and the resulting distribution is denoted by $S_d(\psi)$.

Within some of our proofs we will need another characterization of a spherical distribution which is given by the following theorem:

Theorem 3.3. *An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^{\mathrm{t}}$ has a spherical distribution if and only if it has the stochastic representation*

$$\boldsymbol{X} \stackrel{\mathrm{d}}{=} R\boldsymbol{S} \,, \tag{3.2}$$

where $\boldsymbol{S}$ is uniformly distributed on the unit sphere $\mathcal{S}^{d-1} = \big\{\boldsymbol{s} \in \mathbb{R}^d : \boldsymbol{s}^{\mathrm{t}}\boldsymbol{s} = 1\big\}$ and $R \geq 0$ is a radial random variable, independent of $\boldsymbol{S}$.

Proof. See e.g. McNeil et al. (2005, pp. 90–91). □

Remark 3.4. For every spherically distributed $\boldsymbol{X}$ with $\mathbb{P}[\boldsymbol{X} = \boldsymbol{0}] = 0$ we have

$$\Big(\|\boldsymbol{X}\|, \frac{\boldsymbol{X}}{\|\boldsymbol{X}\|}\Big) \stackrel{\mathrm{d}}{=} (R, \boldsymbol{S}) \tag{3.3}$$

(see e.g. McNeil et al. (2005, p. 91)), so the distribution of the radial random variable R is uniquely determined.

Example 3.5. Let $\boldsymbol{X}$ be a d-dimensional, standard normally distributed random vector, $\boldsymbol{X} \sim \mathcal{N}_d(\boldsymbol{0}, \boldsymbol{I}_d)$. Since $\boldsymbol{X}^{\mathrm{t}}\boldsymbol{X} \sim \chi^2_d$, where χ^2_d denotes a chi-squared distribution with d degrees of freedom, it follows from (3.3) that $R^2 \sim \chi^2_d$.

The components of spherically distributed random variables are pairwise uncorrelated, which implies that the dependence measures linear correlation coefficient and Kendall's tau equal zero, $\varrho = 0$ and $\tau = 0$. To introduce correlation we have to generalize the spherical distributions to elliptical ones. This generalization is realized by an affine transformation:

Definition 3.6. An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \dots, X_d)^{\mathrm{t}}$ has an *elliptical distribution* with parameters $\boldsymbol{\mu} \in \mathbb{R}^d$ and positive semi-definite $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ if there exist $k \in \mathbb{N}$, $\boldsymbol{Y} \in \mathbb{R}^k$ and $\boldsymbol{A} \in \mathbb{R}^{d\times k}$ such that

$$\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Y}, \tag{3.4}$$

where $\boldsymbol{Y}$ is spherically distributed and $\boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \boldsymbol{\Sigma}$.

Notation. We refer to $\boldsymbol{\mu}$ as the *location vector* and $\boldsymbol{\Sigma}$ as the *dispersion matrix.*

Again it is common to characterize the elliptical distribution by the characteristic generator. If we assume $\boldsymbol{Y} \sim S_k(\psi)$ in Definition 3.6, then we get for the characteristic function of $\boldsymbol{X}$

$$\begin{aligned}\varphi_{\boldsymbol{X}}(\boldsymbol{t}) &= \mathbb{E}[\, e^{i\boldsymbol{t}^{\mathrm{t}}\boldsymbol{X}}\,] = \mathbb{E}[\, e^{i\boldsymbol{t}^{\mathrm{t}}(\boldsymbol{\mu}+\boldsymbol{A}\boldsymbol{Y})}\,] = e^{i\boldsymbol{t}^{\mathrm{t}}\boldsymbol{\mu}}\,\mathbb{E}[\, e^{i(\boldsymbol{A}^{\mathrm{t}}\boldsymbol{t})^{\mathrm{t}}\boldsymbol{Y}}\,]\\ &= e^{i\boldsymbol{t}^{\mathrm{t}}\boldsymbol{\mu}}\,\psi(\boldsymbol{t}^{\mathrm{t}}\boldsymbol{\Sigma}\,\boldsymbol{t})\end{aligned}$$

and denote the elliptical distribution by $E_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$.

Remark 3.7. The representation of an elliptical distribution by formula (3.4) or by $E_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$ is not unique:

(a) As spherical distributions are invariant under orthogonal transformations we have for every orthogonal matrix $\boldsymbol{U} \in \mathcal{O}(k)$ that

$$\boldsymbol{\mu} + (\boldsymbol{A}\boldsymbol{U})\boldsymbol{Y} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{Y}\,.$$

(b) If $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$ and also $\boldsymbol{X} \sim E_d(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*, \psi^*)$, then there exists a constant $c > 0$ such that $\boldsymbol{\mu} = \boldsymbol{\mu}^*$, $\boldsymbol{\Sigma} = c\,\boldsymbol{\Sigma}^*$ and $\psi(\cdot) = \psi^*(c^{-1}\,\cdot)$ (see e.g. Cambanis et al., 1981, Thm 3(i), pp. 372–373).

As for spherical distributions there is also another way of characterizing an elliptical distribution:

Theorem 3.8. *An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \dots, X_d)^{\mathrm{t}}$ has an elliptical distribution with location vector $\boldsymbol{\mu}$ and dispersion matrix $\boldsymbol{\Sigma}$ if and only if there exist $k \in \mathbb{N}$, $\boldsymbol{S}$, R and $\boldsymbol{A}$ satisfying*

$$\boldsymbol{X} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + R\boldsymbol{A}\boldsymbol{S}, \tag{3.5}$$

with

(i) $\boldsymbol{S}$ *uniformly distributed on the unit sphere* $\mathcal{S}^{k-1} = \big\{\boldsymbol{s} \in \mathbb{R}^k : \boldsymbol{s}^{\mathrm{t}}\boldsymbol{s} = 1\big\}$,

(ii) $R \geq 0$, *a radial random variable, independent of* $\boldsymbol{S}$, *and*

(iii) $\boldsymbol{A} \in \mathbb{R}^{d\times k}$ *with* $\boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \boldsymbol{\Sigma}$.

Proof. It follows directly from Definition 3.6 and Theorem 3.3. □

Remark 3.9. In the following it will be helpful to specify $k = \operatorname{rank}(\boldsymbol{\Sigma})$, which is always possible if $\operatorname{rank}(\boldsymbol{\Sigma}) > 0$ (see Cambanis et al., 1981, pp. 369–370). We refer to (3.5) with this special choice of k as full rank representation. As we further required $\boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \boldsymbol{\Sigma}$ we always have $\operatorname{rank}(\boldsymbol{A}) = \operatorname{rank}(\boldsymbol{\Sigma}) = k$ in this case.

Remark 3.10. In the following we are mainly interested in the two-dimensional case. We will need the distribution and the moments of the components of $\boldsymbol{S}$. We have to distinguish between three cases.

(i) $\operatorname{rank}(\boldsymbol{\Sigma}) = 0$. In this case both components of the elliptically distributed random vector are degenerate.

(ii) $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$. Here we have $\boldsymbol{S} = S$ which is uniformly distributed on $\{-1, 1\}$. So it is easy to see that for all $n \in \mathbb{N}$

$$\mathbb{E}[\,S^{2n-1}\,] = 0 \quad \text{and} \quad \mathbb{E}[\,S^{2n}\,] = 1\,.$$

(iii) $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$. Here we have $\boldsymbol{S} = (S_1, S_2)^{\mathrm{t}}$ which is uniformly distributed on the unit circle. So we can say that

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \overset{\mathrm{d}}{=} \begin{pmatrix} \cos\chi \\ \sin\chi \end{pmatrix},$$

where χ is uniformly distributed on $[0, 2\pi]$. Denoting $\arccos : [-1,1] \to [0,\pi]$ the principle branch of the arc cosine, the distribution of $S = S_1 \overset{\mathrm{d}}{=} S_2$ becomes

$$\begin{aligned} \mathbb{P}[\,S \leq x\,] &= \mathbb{P}[\,\cos\chi \leq x\,] = \mathbb{P}[\,\arccos x \leq \chi \leq 2\pi - \arccos x\,] \\ &= \int_{\arccos x}^{2\pi - \arccos x} \frac{1}{2\pi}\,d\chi \\ &= 1 - \frac{1}{\pi}\arccos x\,, \quad x \in [-1,1]\,, \end{aligned}$$

and its density

$$f_S(x) = \frac{1}{\pi\sqrt{1-x^2}}\,, \quad x \in (-1,1)\,.$$

Due to its symmetry and boundedness we know that for all $n \in \mathbb{N}$

$$\mathbb{E}[\,S^{2n-1}\,] = 0$$

and can further calculate

$$\mathbb{E}[\, S^2 \,] = \int_0^{2\pi} \frac{\cos^2 \chi}{2\pi} \, d\chi = \frac{1}{2} \quad \text{and} \quad \mathbb{E}[\, S^4 \,] = \int_0^{2\pi} \frac{\cos^4 \chi}{2\pi} \, d\chi = \frac{3}{8} \,.$$

We already know that the two components are uncorrelated:

$$\mathbb{E}[\, S_1 \, S_2 \,] = \int_0^{2\pi} \frac{\cos \chi \, \sin \chi}{2\pi} \, d\chi = 0 \,.$$

Later we will further need

$$\mathbb{E}[\, S_1^3 \, S_2 \,] = \mathbb{E}[\, S_1 \, S_2^3 \,] = 0 \quad \text{and} \quad \mathbb{E}[\, S_1^2 \, S_2^2 \,] = \frac{1}{8} \,.$$

3.2 Measures of dependence for elliptical distributions

For elliptical distributions the linear correlation coefficient is the natural measure of dependence (see e.g. Embrechts et al., 1999). So we present now the definitions of the measure and its estimator. Under appropriate moment conditions the estimator is asymptotically normal and the asymptotic variance can be calculated. In the case of elliptical distributions it even gets a very simple form.

In the elliptical world, using the linear correlation coefficient is equivalent to using Kendall's tau as there is a unique connection between them, which was shown in Lindskog et al. (2003). This connection leads to a second way of estimating correlation in the case of elliptical distributions. Again we can show asymptotic normality of this alternative estimator such that we can compare the classical and the new estimator by their asymptotic variance.

3.2.1 Linear correlation for elliptical distributions

Definition 3.11. For two non-degenerate, real-valued random variables X and Y with $\mathbb{E}[X^2] < \infty$ and $\mathbb{E}[Y^2] < \infty$ the *linear correlation coefficient* is defined as

$$\varrho = \varrho(X, Y) := \frac{\mathbb{C}\mathrm{ov}[X, Y]}{\sqrt{\mathbb{V}\mathrm{ar}[X]} \, \sqrt{\mathbb{V}\mathrm{ar}[Y]}} \,, \tag{3.6}$$

where $\mathbb{C}\mathrm{ov}[X, Y] = \mathbb{E}[\, (X - \mathbb{E}[X]) \, (Y - \mathbb{E}[Y]) \,]$, $\mathbb{V}\mathrm{ar}[X] = \mathbb{E}[\, (X - \mathbb{E}[X])^2 \,]$ and $\mathbb{V}\mathrm{ar}[Y] = \mathbb{E}[\, (Y - \mathbb{E}[Y])^2 \,]$.

In the case of elliptical distributions the linear correlation coefficient gets an easy form:

Lemma 3.12. *For a two-dimensional elliptical distribution with non-degenerate, square-integrable components and with dispersion matrix*

$$\boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

the linear correlation coefficient equals

$$\varrho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\,\Sigma_{22}}}.$$

Proof. As the components are non-degenerate we know that $\operatorname{rank}(\boldsymbol{\Sigma}) > 0$ and can use the full rank representation (3.5). Assume first $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$ and write $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ such that (3.5) becomes

$$\begin{pmatrix} X \\ Y \end{pmatrix} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + R \begin{pmatrix} a_{11}S_1 + a_{12}S_2 \\ a_{21}S_1 + a_{22}S_2 \end{pmatrix}.$$

As R and $\boldsymbol{S}$ are independent and as $\mathbb{E}[R^2] < \infty$ by the assumption of square-integrable components, we get

$$\varrho = \frac{\mathbb{E}[\,(a_{11}S_1 + a_{12}S_2)\,(a_{21}S_1 + a_{22}S_2)\,]}{\mathbb{E}[\,(a_{11}S_1 + a_{12}S_2)^2\,]^{\frac{1}{2}}\,\mathbb{E}[\,(a_{21}S_1 + a_{22}S_2)^2\,]^{\frac{1}{2}}},$$

which is well-defined since $a_{i1} = a_{i2} = 0$, $i \in \{1, 2\}$, would violate our assumption $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$. Using the moments of S_1 and S_2 from Remark 3.10 we can calculate

$$\varrho = \frac{a_{11}\,a_{21} + a_{12}\,a_{22}}{\sqrt{(a_{11}^2 + a_{12}^2)\,(a_{21}^2 + a_{22}^2)}}. \tag{3.7}$$

Knowing that

$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \begin{pmatrix} a_{11}^2 + a_{12}^2 & a_{11}\,a_{21} + a_{12}\,a_{22} \\ a_{11}\,a_{21} + a_{12}\,a_{22} & a_{21}^2 + a_{22}^2 \end{pmatrix}$$

we get the result.

Now consider $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$. For representation (3.5) we have $\boldsymbol{S} = S$, uniformly distributed on $\{-1, 1\}$, $\boldsymbol{A} = (a_1, a_2)^{\mathrm{t}}$ and $(X, Y)^{\mathrm{t}} \stackrel{\mathrm{d}}{=} \boldsymbol{\mu} + RS(a_1, a_2)^{\mathrm{t}}$. The linear correlation coefficient becomes (again R and S are independent and $\mathbb{E}[R^2] < \infty$)

$$\varrho = \frac{\mathbb{E}[\,a_1\,a_2\,S^2\,]}{\mathbb{E}[\,a_1^2\,S^2\,]^{\frac{1}{2}}\,\mathbb{E}[\,a_2^2\,S^2\,]^{\frac{1}{2}}} = \frac{a_1\,a_2}{\sqrt{a_1^2\,a_2^2}}. \tag{3.8}$$

Again this is well-defined as assuming a distribution with non-degenerate components assures that $a_1 \neq 0$ and $a_2 \neq 0$. The result follows with the observation

$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \begin{pmatrix} a_1^2 & a_1\,a_2 \\ a_1\,a_2 & a_2^2 \end{pmatrix}.$$

□

Definition 3.13. For a two-dimensional elliptical distribution with non-degenerate components and with dispersion matrix $\boldsymbol{\Sigma}$ define the *linear correlation coefficient* as

$$\varrho = \frac{\Sigma_{12}}{\sqrt{\Sigma_{11}\,\Sigma_{22}}}\,. \tag{3.9}$$

Remark 3.14. This new definition is an extension of the classical one, including also elliptical distributions with infinite variances, since for elliptical distributions with square-integrable components the two definitions coincide, as shown in Lemma 3.12. The extended definition leads to a uniquely defined ϱ despite the fact that for an elliptical distribution the dispersion matrix is not uniquely defined (see Remark 3.7), because $\boldsymbol{\Sigma}$ can only differ by a multiplicative constant and this cancels out.

Lemma 3.15. *A two-dimensional elliptical distribution with non-degenerate components has perfect linear correlation, $|\varrho| = 1$, if and only if* $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$.

Proof. Let $|\varrho| = 1$. $\operatorname{rank}(\boldsymbol{\Sigma}) = 0$ is not possible as we assumed non-degeneracy of the components. So assume $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$. Then we know from (3.7) that

$$\begin{aligned} |\varrho| = 1 &\Leftrightarrow |a_{11}\,a_{21} + a_{12}\,a_{22}| = \sqrt{(a_{11}^2 + a_{12}^2)\,(a_{21}^2 + a_{22}^2)} \\ &\Leftrightarrow (a_{11}\,a_{21} + a_{12}\,a_{22})^2 - (a_{11}^2 + a_{12}^2)\,(a_{21}^2 + a_{22}^2) = 0 \\ &\Leftrightarrow a_{11}\,a_{22} = a_{12}\,a_{21}\,. \end{aligned}$$

But in this case we have $\operatorname{rank}(\boldsymbol{A}) = 1$ and therefore also $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$ which is a contradiction.

Let $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$. Then it follows from (3.8) that

$$\varrho = \frac{a_1\,a_2}{|a_1|\,|a_2|} = \operatorname{sgn}(a_1\,a_2) \in \{-1, 1\}$$

and therefore $|\varrho| = 1$. ☐

3.2.2 Standard estimator for the linear correlation coefficient

Definition 3.16. For n i.i.d. pairs of non-degenerate, real-valued random variables (X_j, Y_j), $j = 1, \ldots, n$, the *standard estimator* for the linear correlation coefficient is defined as

$$\hat{\varrho}_n = \frac{\sum_{j=1}^n (X_j - \bar{X}_n)(Y_j - \bar{Y}_n)}{\sqrt{\sum_{j=1}^n (X_j - \bar{X}_n)^2 \, \sum_{j=1}^n (Y_j - \bar{Y}_n)^2}}\,, \tag{3.10}$$

where $\bar{X}_n = \frac{1}{n}\sum_{j=1}^n X_j$ and $\bar{Y}_n = \frac{1}{n}\sum_{j=1}^n Y_j$ are the sample averages. (On the event where the denominator vanishes define $\hat{\varrho}_n = 0$.)

Also the standard estimator is asymptotically normal. The asymptotic variance has a nice form in the case of elliptical distributions.

Theorem 3.17. *For i.i.d. pairs of non-degenerate, real-valued random variables (X_j, Y_j), $j \in \mathbb{N}$, with $\mathbb{E}[X_1^4] < \infty$ and $\mathbb{E}[Y_1^4] < \infty$ the standard estimators $(\hat{\varrho}_n)_{n\geq 2}$, as defined in* (3.10), *normalized with $\sqrt{n}$, are asymptotically normal,*

$$\sqrt{n}\,\big(\hat{\varrho}_n - \varrho\big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, \sigma_\varrho^2\big), \quad n \to \infty\,. \tag{3.11}$$

The asymptotic variance is

$$\sigma_\varrho^2 = \Big(1 + \frac{\varrho^2}{2}\Big)\frac{\sigma_{22}}{\sigma_{20}\sigma_{02}} + \frac{\varrho^2}{4}\Big(\frac{\sigma_{40}}{\sigma_{20}^2} + \frac{\sigma_{04}}{\sigma_{02}^2} - 4\frac{\sigma_{31}}{\sigma_{11}\sigma_{20}} - 4\frac{\sigma_{13}}{\sigma_{11}\sigma_{02}}\Big)\,, \tag{3.12}$$

where $\sigma_{kl} := \mathbb{E}[\,(X - \mathbb{E}[X])^k\,(Y - \mathbb{E}[Y])^l\,]$, $k, l \in \mathbb{N}_0$ with $k + l \leq 4$.

If in addition the distribution of (X_1, Y_1) is elliptical, then the asymptotic variance can be simplified to

$$\sigma_\varrho^2 = \frac{\mathbb{E}[R^4]}{2\,\mathbb{E}[R^2]^2}\,\big(1 - \varrho^2\big)^2\,, \tag{3.13}$$

where R is the radial random variable as in representation (3.5).

Proof. The asymptotic normality is proved in Witting and Müller-Funk (1995, pp. 108–109).

To prove the formula for the elliptical case we use the full rank representation (3.5), which is possible as the assumption of non-degenerate components implies $\mathrm{rank}(\boldsymbol{\Sigma}) > 0$. Assume first that $\mathrm{rank}(\boldsymbol{\Sigma}) = 2$ and denote $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Using the independence of R and $\boldsymbol{S}$ we get for every $k, l \in \mathbb{N}_0$, $k + l \leq 4$,

$$\begin{aligned} \sigma_{kl} &= \mathbb{E}[\,(X - \mu_X)^k\,(Y - \mu_Y)^l\,] \\ &= \mathbb{E}[\,R^{k+l}\,]\,\mathbb{E}[\,(a_{11}\,S_1 + a_{12}\,S_2)^k\,(a_{21}\,S_1 + a_{22}\,S_2)^l\,]\,. \end{aligned}$$

To shorten notation define $\tilde{\sigma}_{kl} := \mathbb{E}[\,(a_{11}\,S_1 + a_{12}\,S_2)^k\,(a_{21}\,S_1 + a_{22}\,S_2)^l\,]$. Knowing the needed moments of S_1 and S_2 (see Remark 3.10) we can calculate

$$\begin{aligned} \tilde{\sigma}_{11} &= \frac{1}{2}\,\big(a_{11}\,a_{21} + a_{12}\,a_{22}\big)\,, \\ \tilde{\sigma}_{20} &= \frac{1}{2}\,\big(a_{11}^2 + a_{12}^2\big)\,, \\ \tilde{\sigma}_{02} &= \frac{1}{2}\,\big(a_{21}^2 + a_{22}^2\big)\,, \\ \tilde{\sigma}_{22} &= \frac{1}{8}\,\big(3\,a_{11}^2\,a_{21}^2 + a_{11}^2\,a_{22}^2 + 4\,a_{11}\,a_{12}\,a_{21}\,a_{22} + a_{12}^2\,a_{21}^2 + 3\,a_{12}^2\,a_{22}^2\big)\,, \\ \tilde{\sigma}_{31} &= \frac{3}{8}\,\big(a_{11}^3\,a_{21} + a_{11}^2\,a_{12}\,a_{22} + a_{11}\,a_{12}^2\,a_{21} + a_{12}^3\,a_{22}\big)\,, \\ \tilde{\sigma}_{13} &= \frac{3}{8}\,\big(a_{11}\,a_{21}^3 + a_{11}\,a_{21}\,a_{22}^2 + a_{12}\,a_{21}^2\,a_{22} + a_{12}\,a_{22}^3\big)\,, \\ \tilde{\sigma}_{40} &= \frac{3}{8}\,\big(a_{11}^2 + a_{12}^2\big)^2\,, \text{ and} \\ \tilde{\sigma}_{04} &= \frac{3}{8}\,\big(a_{21}^2 + a_{22}^2\big)^2\,. \end{aligned}$$

Note that

$$\frac{\tilde\sigma_{11}\,\tilde\sigma_{40}}{\tilde\sigma_{20}} = \tilde\sigma_{31} \quad \text{and} \quad \frac{\tilde\sigma_{11}\,\tilde\sigma_{04}}{\tilde\sigma_{02}} = \tilde\sigma_{13}\,, \tag{3.14}$$

and further

$$\tilde\sigma_{31}\,\tilde\sigma_{02} = \tilde\sigma_{13}\,\tilde\sigma_{20}\,. \tag{3.15}$$

The square of the linear correlation coefficient can be written as

$$\varrho^2 = \frac{\sigma_{11}^2}{\sigma_{20}\,\sigma_{02}} = \frac{\tilde\sigma_{11}^2}{\tilde\sigma_{20}\,\tilde\sigma_{02}}\,. \tag{3.16}$$

So we can rewrite equation (3.12) as

$$\begin{aligned}
\frac{\mathbb{E}[R^2]^2}{\mathbb{E}[R^4]}\,\sigma_\varrho^2 &= \frac{\tilde\sigma_{22}}{\tilde\sigma_{20}\,\tilde\sigma_{02}} + \frac{\tilde\sigma_{11}^2\,\tilde\sigma_{22}}{2\,\tilde\sigma_{20}^2\,\tilde\sigma_{02}^2} + \frac{\tilde\sigma_{11}^2\,\tilde\sigma_{40}}{4\,\tilde\sigma_{20}^3\,\tilde\sigma_{02}} + \frac{\tilde\sigma_{11}^2\,\tilde\sigma_{04}}{4\,\tilde\sigma_{20}\,\tilde\sigma_{02}^3} - \frac{\tilde\sigma_{11}\,\tilde\sigma_{31}}{\tilde\sigma_{20}^2\,\tilde\sigma_{02}} - \frac{\tilde\sigma_{11}\,\tilde\sigma_{13}}{\tilde\sigma_{20}\,\tilde\sigma_{02}^2} \\
&= \frac{1}{\tilde\sigma_{20}^2\,\tilde\sigma_{02}^2}\Big(\tilde\sigma_{22}\,\tilde\sigma_{20}\,\tilde\sigma_{02} + \frac{1}{2}\,\tilde\sigma_{11}^2\,\tilde\sigma_{22} + \frac{\tilde\sigma_{11}^2\,\tilde\sigma_{40}\,\tilde\sigma_{02}}{4\,\tilde\sigma_{20}} + \frac{\tilde\sigma_{11}^2\,\tilde\sigma_{04}\,\tilde\sigma_{20}}{4\,\tilde\sigma_{02}} \\
&\qquad\qquad - \tilde\sigma_{11}\,\tilde\sigma_{31}\,\tilde\sigma_{02} - \tilde\sigma_{11}\,\tilde\sigma_{13}\,\tilde\sigma_{20}\Big) \qquad\qquad (3.17) \\
&\overset{(3.14)}{=} \frac{1}{\tilde\sigma_{20}^2\,\tilde\sigma_{02}^2}\Big(\tilde\sigma_{22}\,\tilde\sigma_{20}\,\tilde\sigma_{02} + \frac{1}{2}\,\tilde\sigma_{11}^2\,\tilde\sigma_{22} - \frac{3}{4}\,\tilde\sigma_{11}\,\tilde\sigma_{31}\,\tilde\sigma_{02} - \frac{3}{4}\,\tilde\sigma_{11}\,\tilde\sigma_{13}\,\tilde\sigma_{20}\Big) \\
&\overset{(3.15)}{=} \frac{1}{\tilde\sigma_{20}^2\,\tilde\sigma_{02}^2}\Big(\tilde\sigma_{22}\,\tilde\sigma_{20}\,\tilde\sigma_{02} + \frac{1}{2}\,\tilde\sigma_{11}^2\,\tilde\sigma_{22} - \frac{3}{2}\,\tilde\sigma_{11}\,\tilde\sigma_{31}\,\tilde\sigma_{02}\Big)\,.
\end{aligned}$$

Inserting the values of the remaining $\tilde\sigma$ we can calculate that

$$\tilde\sigma_{22}\,\tilde\sigma_{20}\,\tilde\sigma_{02} + \frac{1}{2}\,\tilde\sigma_{11}^2\,\tilde\sigma_{22} - \frac{3}{2}\,\tilde\sigma_{11}\,\tilde\sigma_{31}\,\tilde\sigma_{02} = \frac{1}{32}\,\big(a_{11}\,a_{22} - a_{12}\,a_{21}\big)^4\,.$$

As we can also calculate that

$$\tilde\sigma_{20}^2\,\tilde\sigma_{02}^2\,(1-\varrho^2)^2 = \big(\tilde\sigma_{20}\,\tilde\sigma_{02} - \tilde\sigma_{11}^2\big)^2 = \frac{1}{16}\,\big(a_{11}\,a_{22} - a_{12}\,a_{21}\big)^4$$

we get the result.

In the case of a non-degenerate elliptical distribution with $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$ we have $(X,Y)^{\mathrm{t}} \overset{\mathrm{d}}{=} RS(a_1,a_2)^{\mathrm{t}}$. For every $k,l \in \mathbb{N}_0$, $k+l \le 4$, we get

$$\begin{aligned}
\sigma_{kl} &= \mathbb{E}[\,(X-\mu_X)^k\,(Y-\mu_Y)^l\,] = \mathbb{E}[\,R^{k+l}\,]\,\mathbb{E}[\,(a_1\,S_1)^k\,(a_2\,S_2)^l\,] \\
&= \begin{cases} \mathbb{E}[\,R^{k+l}\,]\,a_1^k\,a_2^l\,, & \text{if } k+l \in 2\mathbb{N}\,, \\ 0\,, & \text{otherwise}\,. \end{cases}
\end{aligned}$$

Defining $\tilde\sigma_{kl} = a_1^k\,a_2^l$ for even $k+l$, equations (3.14)–(3.17) hold, and it is easy to see that the last line in (3.17) equals zero, i.e. $\sigma_\varrho^2 = 0$. As further $\varrho \in \{-1,1\}$ (see Lemma 3.15), it follows that also in this case equation (3.13) holds. □

3.2.3 Alternative estimator via transformation of Kendall's tau

We want to compare this standard estimator to an alternative estimating procedure, which is based on the following relation between the two measures of dependence in the case of elliptical distributions.

Theorem 3.18. *Let* $(X,Y)^{\mathrm{t}}$ *be elliptically distributed with non-degenerate components and with location vector* $(\mu_X,\mu_Y)^{\mathrm{t}}$ *and dispersion matrix* $\mathbf{\Sigma}$. *Denote*

$$a_X := 1 - \sum_{x\in\mathbb{R}} \big(\mathbb{P}[\,X = x\,]\big)^2 ,$$

where the sum extends over all atoms of the distribution of X. *Then*

$$\tau(X,Y) = \frac{2\,a_X}{\pi}\,\arcsin\varrho(X,Y)\,. \tag{3.18}$$

If in addition $\operatorname{rank}(\mathbf{\Sigma}) = 2$, *then* $a_X = 1 - \big(\mathbb{P}[\,X = \mu_X\,]\big)^2$.

Proof. See Lindskog et al. (2003). For elliptical distributions having a density (which we discuss in the corollary) this has been proved by Fang et al. (2002, Thm 3.1). □

Remark 3.19. Since we can use the extended definition of ϱ (see Definition 3.13 and the remark afterwards), we do not need to assume finite variances in this theorem and its corollary.

Remark 3.20. Assuming non-degenerate components we always have $a_X > 0$.

To further simplify this equation, we can either take the general case (3.18) and additionally assume that the distribution of X has no atoms, i.e. that X is continuously distributed, or we take the assumption $\operatorname{rank}(\mathbf{\Sigma}) = 2$ and further restrict ourselves to distributions with $\mathbb{P}[\,X = \mu_X\,] = 0$.

Corollary 3.21. *Let* $(X,Y)^{\mathrm{t}}$ *be elliptically distributed with location vector* $(\mu_X,\mu_Y)^{\mathrm{t}}$ *and dispersion matrix* $\mathbf{\Sigma}$. *If further* $\operatorname{rank}(\mathbf{\Sigma}) = 2$ *and* $\mathbb{P}[\,X = \mu_X\,] = 0$ *or if* X *is continuously distributed, then* $a_X = 1$ *and* (3.18) *simplifies to*

$$\tau = \frac{2}{\pi}\,\arcsin\varrho\,. \tag{3.19}$$

This relationship suggests to estimate first Kendall's tau and to transform the obtained value into an estimate of the linear correlation in the following way:

Definition 3.22. For elliptical distributions with non-degenerate components the *transformation-estimator* for the linear correlation ϱ is defined for every $n \in \mathbb{N}$, $n \geq 2$, as

$$\hat{\varrho}_{\tau,n} = \sin\Big(\frac{\pi}{2\,a_X}\,\hat{\tau}_n\Big)\,, \tag{3.20}$$

where $\hat{\tau}_n$ is the tau-estimator as defined in (2.8).

Remark 3.23. This estimator has implicitly been used in the simulation study in Lindskog et al. (2003) and has also been used in the case $a_X = 1$ by Kuhn (2006, pp. 89–90) and Klüppelberg and Kuhn (2009, pp. 745–746).

Remark 3.24. We require non-degenerate components as this is needed in the definition of the linear correlation coefficient and therefore also in Theorem 3.18, although the tau-estimator is well-defined even if one component is degenerate.

To compare the two estimators we want to look at their asymptotic variances. A comparison is possible because also the transformation-estimator is asymptotically normal:

Theorem 3.25. *For i.i.d. pairs of non-degenerate random variables* (X_j, Y_j), $j \in \mathbb{N}$, *following an elliptical distribution, the transformation-estimators* $(\hat{\varrho}_{\tau,n})_{n\geq 2}$, *as defined in* (3.20), *normalized with* $\sqrt{n}$, *are asymptotically normal,*

$$\sqrt{n}\,\big(\hat{\varrho}_{\tau,n} - \varrho\big) \xrightarrow{\mathrm{d}} \mathcal{N}\big(0, \sigma^2_{\varrho(\tau)}\big), \quad n \to \infty\,. \tag{3.21}$$

The asymptotic variance is

$$\sigma^2_{\varrho(\tau)} = \frac{\pi^2}{4\,a_X^2}\,\sigma_\tau^2\,\cos^2\!\Big(\frac{\pi}{2\,a_X}\,\tau\Big) = \frac{\pi^2}{4\,a_X^2}\,\sigma_\tau^2\,(1-\varrho^2)\,, \tag{3.22}$$

where σ_τ^2 *is the asymptotic variance of the tau-estimator as defined in* (2.15).

Proof. If $\tau = \pm 1$, then $\hat{\tau}_n = \pm 1$ for all $n \geq 2$ (see Remark 2.10) and $\sigma_\tau^2 = 0$. It further implies that the distribution can not have any atoms, so $a_X = 1$ and therefore, by Definition 3.22, $\hat{\varrho}_{\tau,n} = \pm 1$ for all $n \geq 2$. So the variance of the transformation-estimator and especially its asymptotic variance equal zero and equations (3.21) and (3.22) hold true with $\sigma^2_{\varrho(\tau)} = 0$.

For $|\tau| < 1$ we use a method called delta method, as proved in Lehmann and Casella (1998, p. 58). The transforming function is

$$h(x) = \sin\!\Big(\frac{\pi}{2\,a_X}\,x\Big)\,, \quad x \in [-1, 1]\,,$$

with an existing first derivative

$$h'(x) = \frac{\pi}{2\,a_X}\,\cos\!\Big(\frac{\pi}{2\,a_X}\,x\Big)\,, \quad x \in [-1, 1]\,,$$

which is non-zero on $|x| < 1$. Knowing the asymptotic normality of the tau-estimators $(\hat{\tau}_n)_{n\geq 2}$ from Theorem 2.15, we get

$$\sqrt{n}\,\big(h(\hat{\tau}_n) - h(\tau)\big) \xrightarrow{\mathrm{d}} \mathcal{N}\Big(0, \sigma_\tau^2\,\big(h'(\tau)\big)^2\Big), \quad n \to \infty\,.$$

Since $h(\hat{\tau}_n) = \hat{\varrho}_{\tau,n}$ by definition and $h(\tau) = \varrho$ by Theorem 3.18, we get the result. □

Remark 3.26. Note that we only assume that the elliptically distributed random vector has non-degenerate components, whereas the standard estimator additionally requires finite fourth moments for asymptotic normality. We will notice this difference again when we consider the t-distribution in Section 6.4.

As the tau-estimator also the transformation-estimator possesses the property of strong consistency:

Lemma 3.27. *For i.i.d. pairs of non-degenerate random variables* (X_j, Y_j), $j \in \mathbb{N}$, *following an elliptical distribution, the transformation-estimators* $(\hat{\varrho}_{\tau,n})_{n\geq 2}$, *as defined in* (3.20), *are strongly consistent,*

$$\hat{\varrho}_{\tau,n} \overset{n\to\infty}{\longrightarrow} \varrho \quad a.s.$$

Proof. Since the tau-estimators are strongly consistent (see Remark 2.9) and h is continuous, we also know that $h(\hat{\tau}_n) = \hat{\varrho}_{\tau,n}$ converge almost surely (see e.g. Gänssler and Stute, 1977, p. 59). □

3.3 Normal variance mixture distributions

3.3.1 Definition and basic properties

This section gives an introduction to a subset of elliptical distributions, called normal variance mixture distributions. There are many definitions within the literature. We chose a definition that makes clear that a normal variance mixture distribution is elliptical.

Definition 3.28. An $\mathbb{R}^d$-valued random vector $\boldsymbol{X} = (X_1, \ldots, X_d)^{\mathrm{t}}$ has a *normal variance mixture distribution* with parameters $\boldsymbol{\mu} \in \mathbb{R}^d$ and $\boldsymbol{\Sigma} \in \mathbb{R}^{d\times d}$ if it has the stochastic representation

$$\boldsymbol{X} \overset{\mathrm{d}}{=} \boldsymbol{\mu} + \sqrt{W}\boldsymbol{A}\boldsymbol{Z}, \tag{3.23}$$

with

(i) $\boldsymbol{Z}$ a k-dimensional standard normally distributed random vector,

(ii) $W \geq 0$, a radial random variable, independent of $\boldsymbol{Z}$, and

(iii) $\boldsymbol{A} \in \mathbb{R}^{d\times k}$ with $\boldsymbol{A}\boldsymbol{A}^{\mathrm{t}} = \boldsymbol{\Sigma}$.

Notation. If $\boldsymbol{\mu} = \boldsymbol{0}$, $k = d$ and $\boldsymbol{\Sigma} = \boldsymbol{I}_d$, then the distribution is spherical and we want to call it *standard normal variance mixture distribution.*

Notation. The distribution of W is called *mixing distribution function* and will be denoted by G.

Remark 3.29. Early references on mixture distributions are Robbins (1948) and Beale and Mallows (1959). Andrews and Mallows (1974) introduced normal variance mixture distributions for one-dimensional spherical distributions, but they used the ratio Z/W. Our discussion is mainly inspired by McNeil et al. (2005). The same definition can also be found in Gupta and Varga (1993).

Remark 3.30. Like for general elliptical distributions also every random vector with a non-degenerate normal variance mixture distribution has a representation like (3.23) with $k = \operatorname{rank}(\boldsymbol{\Sigma})$.

The name normal variance mixture distribution comes from the observation that, conditioned on $W = w$, $w \geq 0$, the random variable $\boldsymbol{X}$ is normally distributed with covariance matrix $w\boldsymbol{\Sigma}$, $\mathcal{L}(\boldsymbol{X} \mid W = w) = \mathcal{N}_d(\boldsymbol{\mu}, w\boldsymbol{\Sigma})$. This observation helps specifying a density, which exists if $\boldsymbol{\Sigma}$ has full rank and $\mathbb{P}[\,W = 0\,] = G(0) = 0$. Since the multivariate normal distribution has a density

$$f_{\mathcal{N}_d(\boldsymbol{\mu},\boldsymbol{\Sigma})}(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \exp\Big(-\frac{1}{2}\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big), \quad \boldsymbol{x} \in \mathbb{R}^d\,,$$

(see McNeil et al., 2005, p. 66), the normal variance mixture distribution has a density

$$\begin{aligned} f(\boldsymbol{x}) &= \int_0^\infty f_{\boldsymbol{X}|W}(\boldsymbol{x} \mid w)\, G(dw) \\ &= \int_0^\infty \frac{1}{(2\pi w)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \exp\Big(-\frac{1}{2w}\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big)\, G(dw), \quad \boldsymbol{x} \in \mathbb{R}^d\,, \end{aligned} \tag{3.24}$$

(see McNeil et al., 2005, pp. 74–75).

Remark 3.31. It is not only possible to mix normal distributions with different variances, but also with different means. The resulting distributions are called normal mean-variance mixtures. They provide more flexibility concerning asymmetry, but in general they are not elliptical. For details see e.g. McNeil et al. (2005, pp. 77–78).

3.3.2 Asymptotic variance of the standard estimator for normal variance mixture distributions

Every normal variance mixture distribution can be seen as an elliptical distribution with radial random variable $\sqrt{W}R_{\mathcal{N}}$, where W is the mixing variable as in Definition 3.28 and $R_{\mathcal{N}}$ is the radial random variable of the normal distribution, i.e. $R_{\mathcal{N}}^2 \sim \chi_d^2$ (see Example 3.5). Since those two random variables are independent, we can simplify formula (3.13) of the asymptotic variance of the standard estimator as

$$\begin{aligned} \sigma_\varrho^2 &= \frac{\mathbb{E}[R^4]}{2\,\mathbb{E}[R^2]^2}\,\big(1-\varrho^2\big)^2 = \frac{\mathbb{E}[W^2]\,\mathbb{E}[R_{\mathcal{N}}^4]}{2\,\mathbb{E}[W]^2\,\mathbb{E}[R_{\mathcal{N}}^2]^2}\,\big(1-\varrho^2\big)^2 \\ &= \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]^2}\,\big(1-\varrho^2\big)^2\,. \end{aligned} \tag{3.25}$$

The last step is due to the fact that $\mathbb{E}[R_{\mathcal{N}}^2] = 2$ and $\mathbb{E}[R_{\mathcal{N}}^4] = \mathbb{V}\mathrm{ar}[R_{\mathcal{N}}^2] + \mathbb{E}[R_{\mathcal{N}}^2]^2 = 8$.

3.4 Characterization of elliptical distributions with degenerate tau-kernel

We again look at distributions on $\mathbb{R}^2$ with degenerate tau-kernel, as defined in Section 2.1.2. We are interested in the question whether there exist elliptical distributions with degenerate tau-kernel and if yes, how they look. The answer to the first question is indeed yes. Depending on the rank of the dispersion matrix we can even give the necessary and sufficient conditions for an elliptical distribution to have a degenerate tau-kernel.

Theorem 3.32. *A two-dimensional elliptical distribution with dispersion matrix $\boldsymbol{\Sigma}$ has a degenerate tau-kernel if and only if one of the following properties is satisfied:*

(i) $\operatorname{rank}(\boldsymbol{\Sigma}) = 0$*;*

(ii) $\operatorname{rank}(\boldsymbol{\Sigma}) = 1$ *and additionally one of the following properties is satisfied:*

  (a) *one of the components has a degenerate distribution;*

  (b) *the radial random variable R is continuously distributed;*

  (c) *the random variable RS, where S is uniformly distributed on $\{-1, 1\}$, is uniformly distributed on a finite non-empty set;*

(iii) $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$ *and the radial random variable R has a degenerate distribution.*

Proof. (i) If $\operatorname{rank}(\boldsymbol{\Sigma}) = 0$, then the distribution is degenerate which clearly leads to a degenerate tau-kernel.

(ii) We use the full rank representation, i.e. (3.5) with $k = 1$, and assume w.l.o.g. $\boldsymbol{\mu} = (0, 0)^{\mathrm{t}}$ (as the location vector does not influence Kendall's tau and its estimator). So we have $\boldsymbol{A} = (a_1, a_2)^{\mathrm{t}} \in \mathbb{R}^2$ and $\boldsymbol{S} = S$, uniformly distributed on $\{-1, 1\}$, and therefore $(X, Y)^{\mathrm{t}} \stackrel{\mathrm{d}}{=} RS\,(a_1, a_2)^{\mathrm{t}}$.

If $a_1 = 0$ or $a_2 = 0$ (i.e. case (a)), then the situation is the same as in Example 2.46 with $c = 0$, which means degeneracy of the tau-kernel.

If both $a_1 \neq 0$ and $a_2 \neq 0$, then we have $Y = f(X)$ with $f(x) = \frac{a_2}{a_1}\, x$ being a strictly increasing (if $\operatorname{sgn}(a_1\, a_2) = 1$) or strictly decreasing (if $\operatorname{sgn}(a_1\, a_2) = -1$) function. Then we know from Lemma 2.44 that the tau-kernel is degenerate if and only if one of the following holds:

- X is continuously distributed: this is true if and only if R is continuously distributed, i.e. case (b);
- X is uniformly distributed on a finite, non-empty set: as we have $X \stackrel{\mathrm{d}}{=} a_1 RS$, $a_1 \neq 0$, this is exactly case (c).

(iii) "if": see Lemma 3.33;

"only if": see Lemma 3.36.

□

Lemma 3.33. *A two-dimensional elliptical distribution with a degenerate radial variable R has a degenerate tau-kernel.*

Proof. If $R = 0$, then the assertion is trivially true. So lets assume $R > 0$.

The cases $\operatorname{rank}(\boldsymbol{\Sigma}) < 2$ are already proved in Theorem 3.32.

So assume $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$. To prove that the tau-kernel is degenerate we have to show property (2.40), i.e. that the probability

$$\mathbb{P}[\,(X - \tilde{X})(Y - \tilde{Y}) > 0 \mid X, Y\,] \tag{3.26}$$

is almost surely constant and therefore independent of $(X, Y)^{\mathrm{t}}$ (this is sufficient as the distribution is continuous). So for a given $(X, Y)^{\mathrm{t}}$ we are interested in the areas $\{(\tilde{x}, \tilde{y})^{\mathrm{t}} \in \mathbb{R}^2 \mid \tilde{x} > X,\, \tilde{y} > Y\}$ and $\{(\tilde{x}, \tilde{y})^{\mathrm{t}} \in \mathbb{R}^2 \mid \tilde{x} < X,\, \tilde{y} < Y\}$. In the elliptical setting the areas are simple to describe, as they are separated by horizontal and vertical lines, but the probability that a point $(\tilde{X}, \tilde{Y})^{\mathrm{t}}$ lies in these areas is not easy to see. So we change to the spherical setting, using the full rank representation (3.5) with $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\boldsymbol{S} = (S_1, S_2)^{\mathrm{t}}$ and, w.l.o.g., $\boldsymbol{\mu} = (0, 0)^{\mathrm{t}}$. Remember that the radial random variable is degenerate. We first look at the case $a_{12} > 0$ and $a_{22} > 0$ such that we know

$$\begin{aligned} & X > \tilde{X} \quad \text{and} \quad Y > \tilde{Y} \\ \iff\; & a_{11}S_1 + a_{12}S_2 > a_{11}\tilde{S}_1 + a_{12}\tilde{S}_2 \quad \text{and} \quad a_{21}S_1 + a_{22}S_2 > a_{21}\tilde{S}_1 + a_{22}\tilde{S}_2 \\ \iff\; & \text{(I)} \quad \tilde{S}_2 < \Big(\frac{a_{11}}{a_{12}}\, S_1 + S_2\Big) - \frac{a_{11}}{a_{12}}\, \tilde{S}_1 \quad \text{and} \quad \text{(II)} \quad \tilde{S}_2 < \Big(\frac{a_{21}}{a_{22}}\, S_1 + S_2\Big) - \frac{a_{21}}{a_{22}}\, \tilde{S}_1 \end{aligned} \tag{3.27}$$

In this setting the random vector $(\tilde{S}_1, \tilde{S}_2)^{\mathrm{t}}$ is uniformly distributed on the unit circle. The lines that border the set of points $(\tilde{S}_1, \tilde{S}_2)^{\mathrm{t}}$ where the last expression of (3.27) holds true, are no longer horizontal and vertical, but they stay lines, they cross at $(S_1, S_2)^{\mathrm{t}}$, i.e. on the circle, and their gradient and therefore also the angle between them is fixed independently of $(S_1, S_2)^{\mathrm{t}}$. So to get the probability $\mathbb{P}[\,X > \tilde{X},\, Y > \tilde{Y} \mid X, Y\,]$ we have to integrate over all points on the unit circle that lie under these two lines. The same argumentation tells that the probability $\mathbb{P}[\,X < \tilde{X},\, Y < \tilde{Y} \mid X, Y\,]$ corresponds to the part of the circle which is over the two lines. So we have two lines with a fixed enclosed angel crossing on the circle and we are interested in the part of the circle that lies between the lines. But this is fixed independently of the crossing point (inscribed angle theorem) and therefore independently of $(S_1, S_2)^{\mathrm{t}}$. This means that the probability (3.26) is independent of $(X, Y)^{\mathrm{t}}$ which is the condition of having a degenerate tau-kernel.

For $a_{12} < 0$ or $a_{22} < 0$ the situation stays the same, simply unequal signs can change and so it can change whether the areas contribute to probability (3.26) or to the complementary probability . But as the angle still stays the same also the probabilities are independent of $(S_1, S_2)^{\mathrm{t}}$.

In the case $a_{12} = 0$, $a_{22} \neq 0$ the argumentation stays the same, only inequality (I) of (3.27) changes to $\tilde{S}_1 < S_1$, but the angle stays fixed independently of $(S_1, S_2)^{\mathrm{t}}$.

In the case $a_{12} > 0$, $a_{22} \neq 0$ inequality (II) changes to $\tilde{S}_1 < S_1$. Again the angle stays fixed independently of $(S_1, S_2)^{\mathrm{t}}$.

As we assumed $\operatorname{rank}(\mathbf{\Sigma}) = 2$ we know that $a_{12} = a_{22} = 0$ is not possible. □

Remark 3.34. If the degenerate radial variable has a value greater than zero, then the assumptions of Theorem 3.18 are fulfilled and we can compare the two ways of estimating the linear correlation. As the tau-kernel is degenerate, the tau-estimator for an elliptically distributed random vector with a degenerate radial variable has an asymptotic variance of value zero, $\sigma_\tau^2 = 0$ (see Lemma 2.39). The asymptotic variance stays the same after the transformation into an estimator of linear correlation, $\sigma^2_{\varrho(\tau)} = 0$.

In this setting also the standard estimator is asymptotically normal, but its asymptotic variance does not necessarily equal zero:

$$\sqrt{n}\,\big(\hat{\varrho}_n - \varrho\big) \xrightarrow{\mathrm{d}} \mathcal{N}\Big(0, \frac{1}{2}\,(1-\varrho^2)^2\Big), \quad n \to \infty\,. \tag{3.28}$$

The variance is strictly greater than zero if $\operatorname{rank}(\mathbf{\Sigma}) = 2$ as this implies $|\varrho| < 1$.

It still remains to prove the “only if”-part in Theorem 3.32(iii). We need it for elliptical distributions, which is done in Lemma 3.36. Nonetheless we first prove it for spherical distributions, as this shows better the idea of the proof.

Lemma 3.35. *A two-dimensional spherical distribution can only have a degenerate tau-kernel if the distribution of the radial variable R is degenerate.*

Proof. For the spherically distributed random vector we use the notation

$$(X, Y)^{\mathrm{t}} \overset{\mathrm{d}}{=} R\,(\cos\chi, \sin\chi)^{\mathrm{t}}$$

where $R \sim F_R$ and $\chi \sim \mathcal{U}[0, 2\,\pi]$.

Assume that the distribution of R is non-degenerate. We will show the contradiction by showing that Lemma 2.49 is violated. If F_R is non-degenerate, then there exists $r_0 \in [0, \infty)$ with $F_R(r_0) \in (0, 1)$. By right-continuity of the distribution function we can assume that $r_0 > 0$. Right-continuity also ensures that there exists $r_2 \in (r_0, \infty)$ with $F_R(r_2) \in (0, 1)$. Fix any $r_1 \in (r_0, r_2)$.

Define $u_1 = r_0$, $v_1 = \sqrt{r_1^2 - r_0^2}$, $u_0 = -u_1$, $v_0 = -v_1$ and $\chi_0 = \arcsin(\frac{v_1}{r_1})$. So all the points (u_i, v_j), $i, j \in \{0, 1\}$, are on the circle with radius r_1 and $\chi_0 \in (0, \frac{\pi}{2})$. We have

$$\begin{aligned} \mu\big([u_0, u_1] \times [v_0, v_1]\big) &\geq \mathbb{P}[\, R \leq r_0\,,\ |\sin\chi| \leq \chi_0\,] \\ &= F_R(r_0)\,\frac{4\chi_0}{2\pi} \\ &> 0\,. \end{aligned}$$

Consider the random vector $(U_0, U_1, V_0, V_1) = (-\tilde{R}\cos\tilde{\chi}, \tilde{R}\cos\tilde{\chi}, -\tilde{R}\sin\tilde{\chi}, \tilde{R}\sin\tilde{\chi})$ with $\tilde{R} \sim F_R$ and $\tilde{\chi} \sim \mathcal{U}[0, 2\pi]$. We need to show that assumption (2.47) is fulfilled:

$$\begin{aligned}
\mathbb{P}\big[\,U_0 \le u_0,\, U_1 \ge u_1,\, V_0 \le v_0,\, V_1 \ge v_1\,\big] &= \mathbb{P}\big[\cos\tilde{\chi} \ge \frac{u_1}{\tilde{R}}\,, \sin\tilde{\chi} \ge \frac{v_1}{\tilde{R}}\,\big] \\
&\ge \mathbb{P}\big[\,\tilde{R} \ge r_2\,, \cos\tilde{\chi} \ge \frac{u_1}{r_2}\,, \sin\tilde{\chi} \ge \frac{v_1}{r_2}\,\big] \\
&\ge \big(1 - F_R(r_2)\big)\,\mathbb{P}\big[\cos\tilde{\chi} \ge \frac{r_0}{r_2}\,, \sin\tilde{\chi} \ge \frac{v_1}{r_2}\,\big]\,.
\end{aligned}$$

We chose r_2 such that $\big(1 - F_R(r_2)\big) > 0$. Further we have

$$\begin{aligned}
&\Big\{\tilde{\chi} \mid \sin\tilde{\chi} \ge \frac{v_1}{r_2}\,, \cos\tilde{\chi} \ge \frac{r_0}{r_2}\Big\} \\
&= \Big\{\tilde{\chi} \mid \sin\tilde{\chi} \ge \frac{v_1}{r_2}\,, \sin\tilde{\chi} = \sqrt{1 - \cos^2\tilde{\chi}} \le \sqrt{1 - \frac{r_0^2}{r_2^2}} = \frac{\sqrt{r_2^2 - r_0^2}}{r_2}\Big\} \\
&= \Big\{\tilde{\chi} \mid \frac{\sqrt{r_1^2 - r_0^2}}{r_2} \le \sin\tilde{\chi} \le \frac{\sqrt{r_2^2 - r_0^2}}{r_2}\Big\} \\
&\ne \{\}
\end{aligned}$$

as $r_2 > r_1$, which is a contradiction to (2.48). □

Lemma 3.36. *A two-dimensional elliptical distribution with* $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$ *can only have a degenerate tau-kernel if the distribution of the radial variable* R *is degenerate.*

Proof. For the elliptically distributed random vector we use the notation

$$(X, Y)^{\mathrm{t}} \overset{\mathrm{d}}{=} R\,\boldsymbol{A}\,(\cos\chi, \sin\chi)^{\mathrm{t}}$$

where $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $R \sim F_R$ and $\chi \sim \mathcal{U}[0, 2\pi]$ (w.l.o.g. $\boldsymbol{\mu} = (0,0)^{\mathrm{t}}$). We can assume that $\varrho \ne 0$ as the case $\varrho = 0$ is treated in Lemma 3.35. We further know that $|\varrho| < 1$ as we assumed $\operatorname{rank}(\boldsymbol{\Sigma}) = 2$ (see Lemma 3.15). Lets first look at the case of positive correlation, $\varrho \in (0,1)$, which means that $a_{11}\,a_{21} + a_{12}\,a_{22} > 0$.

Assume that the distribution of R is non-degenerate. Again we want to show the contradiction using Lemma 2.49. With the same arguments as in the proof of Lemma 3.35 we can find $0 < r_0 < r_1 < r_2 < \infty$ such that $F_R(r_i) \in (0,1)$, $i \in \{0,1,2\}$.

For every $r > 0$ the points $(x, y)^{\mathrm{t}} = r\,\boldsymbol{A}\,(\cos\chi, \sin\chi)^{\mathrm{t}}$, $\chi \in [0, 2\pi]$, have a maximal y-value if $\chi_{\max} = \arctan(\frac{a_{22}}{a_{21}})$ as we have

$$\begin{aligned}
&\frac{\partial y}{\partial \chi} = \frac{\partial}{\partial \chi}\big(r\,(a_{21}\,\cos\chi + a_{22}\,\sin\chi)\big) = r\,(-a_{21}\,\sin\chi + a_{22}\,\cos\chi) \\
\Longrightarrow\quad &\frac{\partial y}{\partial \chi}\,\big(\chi_{\max}\big) = -r\,a_{21}\,\sin\Big(\arctan\Big(\frac{a_{22}}{a_{21}}\Big)\Big) + r\,a_{22}\,\cos\Big(\arctan\Big(\frac{a_{22}}{a_{21}}\Big)\Big) \\
&\qquad = -r\,a_{21}\,\frac{\frac{a_{22}}{a_{21}}}{\sqrt{1 + \frac{a_{22}^2}{a_{21}^2}}} + r\,a_{22}\,\frac{1}{\sqrt{1 + \frac{a_{22}^2}{a_{21}^2}}} = 0\,,
\end{aligned}$$

and the corresponding $y_{\max}(r)$ is positive for positive r:

$$y_{\max}(r) = r\,\big(a_{21}\,\cos\chi_{\max} + a_{22}\,\sin\chi_{\max}\big) = r\,\sqrt{a_{21}^2 + a_{22}^2} > 0\,.$$

We further get

$$x_{\max}(r) = r\,\big(a_{11}\,\cos\chi_{\max} + a_{12}\,\sin\chi_{\max}\big) = r\,\frac{a_{11}\,a_{21} + a_{12}\,a_{22}}{\sqrt{a_{21}^2 + a_{22}^2}} = \frac{r\,\varrho}{\sqrt{a_{21}^2 + a_{22}^2}}\,.$$

Note that $\chi_{\max}$ is independent of the radial variable r. The coordinates of the maximal point are both positive for $r > 0$ as we assumed $\varrho \in (0,1)$.

Define $u_1 = x_{\max}(r_1)$, $v_1 = y_{\max}(r_1)$, $u_0 = -u_1$ and $v_0 = -v_1$. The points with maximal y-values for different radial values lie on a line with gradient

$$\frac{y_{\max}(r)}{x_{\max}(r)} = \frac{a_{21}^2 + a_{22}^2}{\varrho} < \infty\,.$$

Therefore we know that $(x_{\max}(r), y_{\max}(r))^{\mathrm{t}} \in [u_0, u_1] \times [v_0, v_1]$ for $r \le r_1$. As $r_0 < r_1$ we even know that there exists an $\varepsilon_0 > 0$ such that

$$r\,\boldsymbol{A}\,(\cos\chi_0, \sin\chi_0)^{\mathrm{t}} \in [u_0, u_1] \times [v_0, v_1] \tag{3.29}$$

for $\chi_0 \in [\chi_{\max} - \varepsilon_0, \chi_{\max} + \varepsilon_0]$ and $r \le r_0$. We can use the same arguments to show that there exists an $\varepsilon_2 > 0$ such that

$$r\,\boldsymbol{A}\,(\cos\chi_2, \sin\chi_2)^{\mathrm{t}} \in [u_1, \infty] \times [v_1, \infty] \tag{3.30}$$

for $\chi_2 \in [\chi_{\max} - \varepsilon_2, \chi_{\max} + \varepsilon_2]$ and $r \ge r_2$. From the choice of u_0, u_1, v_0 and v_1 we get

$$\begin{aligned}
\mu\big([u_0, u_1] \times [v_0, v_1]\big) &\overset{(3.29)}{\ge} \mathbb{P}\big[\,R \le r_0, \chi \in [\chi_{\max} - \varepsilon_0, \chi_{\max} + \varepsilon_0]\,\big] \\
&= F_R(r_0)\,\frac{2\,\varepsilon_0}{2\,\pi} \\
&> 0\,.
\end{aligned}$$

Consider the random vector (U_0, U_1, V_0, V_1) with

$$\begin{aligned}
U_0 &= -\tilde{R}\,(a_{11}, a_{12})\,(\cos\tilde{\chi}, \sin\tilde{\chi})^{\mathrm{t}}\,, \\
U_1 &= \tilde{R}\,(a_{11}, a_{12})\,(\cos\tilde{\chi}, \sin\tilde{\chi})^{\mathrm{t}}\,, \\
V_0 &= -\tilde{R}\,(a_{21}, a_{22})\,(\cos\tilde{\chi}, \sin\tilde{\chi})^{\mathrm{t}}\,, \\
V_1 &= \tilde{R}\,(a_{21}, a_{22})\,(\cos\tilde{\chi}, \sin\tilde{\chi})^{\mathrm{t}}\,,
\end{aligned}$$

where $\tilde{R} \sim F_R$ and $\tilde{\chi} \sim \mathcal{U}[0, 2\,\pi]$. To observe the contradiction we need to show that assumption (2.47) is fulfilled:

$$\begin{aligned}
&\mathbb{P}\big[\,U_0 \le u_0, U_1 \ge u_1, V_0 \le v_0, V_1 \ge v_1\,\big] \\
&\overset{(3.30)}{\ge} \mathbb{P}\big[\,\tilde{R} \ge r_2, \tilde{\chi} \in [\chi_{\max} - \varepsilon_2, \chi_{\max} + \varepsilon_2]\,\big] \\
&\ge \big(1 - F_R(r_2)\big)\,\frac{2\,\varepsilon_2}{2\,\pi} \\
&> 0\,.
\end{aligned}$$

If $\varrho \in (-1, 0)$ we take $\bar{\boldsymbol{A}} = \begin{pmatrix} a_{11} & a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}$ and get a random vector

$$(\bar{X}, \bar{Y})^{\mathrm{t}} \stackrel{\mathrm{d}}{=} R\, \bar{\boldsymbol{A}}\, (\cos\chi, \sin\chi)^{\mathrm{t}} \stackrel{\mathrm{d}}{=} (X, -Y)^{\mathrm{t}}$$

with linear correlation $\bar{\varrho} = -\varrho \in (0, 1)$. We know from above that the distribution of $(\bar{X}, \bar{Y})^{\mathrm{t}}$ has a degenerate tau-kernel. As $g(y) = -y$ is a strictly monotone function, we know from Lemma 2.43 that also the distribution of $(X, Y)^{\mathrm{t}} \stackrel{\mathrm{d}}{=} (\bar{X}, -\bar{Y})^{\mathrm{t}}$ has a degenerate tau-kernel. $\square$

Chapter 4

Methods to calculate the asymptotic variance of the tau-estimator

In the previous chapters we introduced the estimator of the dependence measure Kendall's tau and showed that it is asymptotically normal. This property is useful to determine asymptotic confidence intervals and later we will also use it to compare two estimating procedures in the case of elliptical distributions. In all these cases we need the value of the asymptotic variance, which is defined as (see (2.15))

$$\sigma_\tau^2 = 4\,\mathbb{E}\big[\,\mathbb{E}[\,\operatorname{sgn}(X_1 - X_2)\,\operatorname{sgn}(Y_1 - Y_2) \mid X_1, Y_1\,]^2\,\big] - 4\,\tau^2\,.$$

So within this chapter we develop simplifications of this formula, assuming different properties of the two-dimensional distribution function.

4.1 Calculation of the asymptotic variance via the copula

Copulas play an important role within the discussion of dependence as they represent the dependence structure between two random variables. We just give the definition and some basic properties here, for more details see e.g. Nelsen (2006).

Definition 4.1. A *two-dimensional copula* C is a function $C : [0,1]^2 \to [0,1]$ with the following properties:

(i) For every $u, v \in [0,1]$:

$$C(u,0) = C(0,v) = 0$$

and

$$C(u,1) = u \quad \text{and} \quad C(1,v) = v\,.$$

(ii) For every $u_1, u_2, v_1, v_2 \in [0, 1]$ such that $u_1 \leq u_2$ and $v_1 \leq v_2$:

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0\,.$$

Equivalently one could simply say that a two-dimensional copula is a distribution function on $[0, 1]^2$ with uniform marginal distributions.

Sklar's theorem ensures that for every given marginal distributions F_X and F_Y and every two-dimensional distribution function G there exists a copula C such that for all $x, y \in \mathbb{R}$

$$G(x, y) = C\big(F_X(x), F_Y(y)\big)\,.$$

If the marginal distributions are continuous, then the copula is unique. Note that in this case we further know that the random variables $F_X(X)$ and $F_Y(Y)$ are uniformly distributed on $[0, 1]$.

One important property of copulas are the Fréchet–Hoeffding bounds. For every copula C we have

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v)\,, \quad u, v \in [0, 1]\,. \tag{4.1}$$

The bounds themselves are copulas, the lower bound indicates perfect negative dependence, the upper bound perfect positive dependence.

As the copula describes the dependence structure entirely and is not influenced by the marginal distributions, it is a desirable property of dependence measures to depend only on the copula. Kendall's tau fulfils this property and for continuous marginal distributions we have the following representation:

Lemma 4.2. *Let X and Y be two continuous random variables whose copula is C. Kendall's tau for X and Y is then given by*

$$\tau = 4 \int_{[0,1]^2} C(u, v)\, dC(u, v) - 1\,. \tag{4.2}$$

Proof. See e.g. McNeil et al. (2005, pp. 207–208). □

As Kendall's tau itself is independent of the marginal distributions, also the asymptotic variance of its estimator can be expressed only by the copula:

Theorem 4.3. *Let X and Y be two random variables with continuous marginal distributions F_X and F_Y and copula C. Defining*

$$U := F_X(X) \quad \textit{and} \quad V := F_Y(Y)\,, \tag{4.3}$$

we can calculate the asymptotic variance for the tau-estimator as

$$\begin{aligned} \sigma_\tau^2 = {} & 64\,\mathbb{E}[\,C^2(U, V)\,] - 64\,\mathbb{E}[\,U\,C(U, V)\,] - 64\,\mathbb{E}[\,V\,C(U, V)\,] + 32\,\mathbb{E}[\,UV\,] \\ & + \frac{20}{3} + 8\,\tau - 4\,\tau^2\,. \end{aligned} \tag{4.4}$$

Proof. When calculating the asymptotic variance (2.15) we are not interested in the exact values of the random variables but only in the signs of the differences $X - \tilde{X}$ and $Y - \tilde{Y}$. Strictly monotone transformations do not change the order of numbers. Marginal distributions are monotone, but not necessarily strictly monotone functions. But the constant areas do not matter as they are just attained on a null set. So we have

$$\operatorname{sgn}(X - \tilde{X}) = \operatorname{sgn}\big(F_X(X) - F_X(\tilde{X})\big) \quad \text{a.s.}$$

and

$$\operatorname{sgn}(Y - \tilde{Y}) = \operatorname{sgn}\big(F_Y(Y) - F_Y(\tilde{Y})\big) \quad \text{a.s.}$$

The random variables $F_X(X)$, $F_X(\tilde{X})$, $F_Y(Y)$ and $F_Y(\tilde{Y})$ are uniformly distributed on the unit interval. Similar to (4.3) we use the notation

$$\tilde{U} := F_X(\tilde{X}) \quad \text{and} \quad \tilde{V} := F_Y(\tilde{Y})\,.$$

To calculate the asymptotic variance we are interested in the conditional expectation

$$\begin{aligned}
&\mathbb{E}[\operatorname{sgn}(X - \tilde{X}) \operatorname{sgn}(Y - \tilde{Y}) \mid X, Y\,] \\
&\quad = \mathbb{E}\big[\operatorname{sgn}\big(F_X(X) - F_X(\tilde{X})\big) \operatorname{sgn}\big(F_Y(Y) - F_Y(\tilde{Y})\big) \mid X, Y\,\big] \quad \text{a.s.}
\end{aligned}$$

From Lemma 2.40 we know that we can also change the conditioning from $\sigma(X, Y)$ to $\sigma(U, V)$, so we have

$$\mathbb{E}[\operatorname{sgn}(X - \tilde{X}) \operatorname{sgn}(Y - \tilde{Y}) \mid X, Y\,] = \mathbb{E}[\operatorname{sgn}(U - \tilde{U}) \operatorname{sgn}(V - \tilde{V}) \mid U, V\,] \quad \text{a.s.}$$

Hence we can rewrite σ_τ^2 as

$$\sigma_\tau^2 = 4\,\mathbb{E}[\,\mathbb{E}[\operatorname{sgn}(U - \tilde{U}) \operatorname{sgn}(V - \tilde{V}) \mid U, V\,]^2\,] - 4\,\tau^2.$$

Using the definition of the copula and again the continuity of the distribution, we can further simplify this formula:

$$\begin{aligned}
\frac{1}{4}\,\sigma_\tau^2 &= \mathbb{E}\big[\,\big(2\,\mathbb{P}[\,(U - \tilde{U})(V - \tilde{V}) > 0 \mid U, V\,] - 1\big)^2\,\big] - \tau^2 \\
&= \mathbb{E}\Big[\Big(2\,\big(\mathbb{P}[\tilde{U} \le U, \tilde{V} \le V \mid U, V\,] + \mathbb{P}[\tilde{U} > U, \tilde{V} > V \mid U, V\,]\big) - 1\Big)^2\Big] - \tau^2 \\
&= \mathbb{E}\Big[\Big(2\,\big(C(U, V) + 1 - U - V + C(U, V)\big) - 1\Big)^2\Big] - \tau^2 \\
&= \mathbb{E}\big[\,\big(4\,C(U, V) - 2\,U - 2\,V + 1\big)^2\,\big] - \tau^2 \\
&= 16\,\mathbb{E}[\,C^2(U, V)\,] + 4\,\mathbb{E}[\,U^2\,] + 4\,\mathbb{E}[\,V^2\,] + 1 - 16\,\mathbb{E}[\,U\,C(U, V)\,] \\
&\quad - 16\,\mathbb{E}[\,V\,C(U, V)\,] + 8\,\mathbb{E}[\,C(U, V)\,] + 8\,\mathbb{E}[\,UV\,] - 4\,\mathbb{E}[\,U\,] - 4\,\mathbb{E}[\,V\,] - \tau^2\,.
\end{aligned}$$

Knowing the moments of random variables that are uniformly distributed on the unit interval and using the copula-representation (4.2) of Kendall's tau, we get the result. □

Remark 4.4. If we assume that the copula C is symmetric, then we can further simplify this formula:

$$\sigma_\tau^2 = 64\,\mathbb{E}[\,C^2(U,V)\,] - 128\,\mathbb{E}[\,U\,C(U,V)\,] + 32\,\mathbb{E}[\,U\,V\,] + \frac{20}{3} + 8\,\tau - 4\,\tau^2\,. \tag{4.5}$$

In Section 5.1 we will use these formulas to specify the asymptotic variance of the tau-estimator for several copulas.

4.2 Axially symmetric distributions

Another approach to simplify the calculation of the asymptotic variance of the tau-estimator is to work with the common distribution directly, assuming it fulfils some helpful properties. Within this section we look for simplifications under the assumption of axial symmetry. We first show a very technical lemma:

Lemma 4.5. *For every $x, y > 0$ we have for the tau-kernel κ_τ, as defined in* (2.5)*,*

$$\begin{aligned}
&\kappa_\tau\big((x,y),\cdot\big) + \kappa_\tau\big((-x,-y),\cdot\big)\\
&\quad = 2\cdot 1_{(-x,x)\times(-y,y)} + 2\cdot 1_{(-\infty,-x)\times(-\infty,-y)} + 2\cdot 1_{(x,\infty)\times(y,\infty)}\\
&\qquad - 2\cdot 1_{(-\infty,-x)\times(y,\infty)} - 2\cdot 1_{(x,\infty)\times(-\infty,-y)}\\
&\qquad + 1_{\{-x\}\times(-\infty,-y)} + 1_{\{-x\}\times(-y,y)} - 1_{\{-x\}\times(y,\infty)}\\
&\qquad + 1_{(-\infty,-x)\times\{-y\}} + 1_{(-x,x)\times\{-y\}} - 1_{(x,\infty)\times\{-y\}}\\
&\qquad + 1_{\{x\}\times(y,\infty)} + 1_{\{x\}\times(-y,y)} - 1_{\{x\}\times(-\infty,-y)}\\
&\qquad + 1_{(x,\infty)\times\{y\}} + 1_{(-x,x)\times\{y\}} - 1_{(-\infty,-x)\times\{y\}}\\
&\qquad + 1_{\{-x\}\times\{-y\}} + 1_{\{x\}\times\{y\}}\,.
\end{aligned} \tag{4.6}$$

Proof. From the definition of the tau-kernel we get

$$\begin{aligned}
&\kappa_\tau\big((x,y),\cdot\big) + \kappa_\tau\big((-x,-y),\cdot\big)\\
&\quad = 1_{(-\infty,x)\times(-\infty,y)} + 1_{(x,\infty)\times(y,\infty)} - 1_{(-\infty,x)\times(y,\infty)} - 1_{(x,\infty)\times(-\infty,y)}\\
&\qquad + 1_{(-\infty,-x)\times(-\infty,-y)} + 1_{(-x,\infty)\times(-y,\infty)} - 1_{(-\infty,-x)\times(-y,\infty)} - 1_{(-x,\infty)\times(-\infty,-y)}\,.
\end{aligned}$$

After decomposing the intervals in the way ($x > 0$)

$$(-\infty,x) = (-\infty,-x) \cup \{-x\} \cup (-x,x)$$

severals terms cancel out and reordering the summands gives the result. □

With the help of this lemma we can simplify the formula of the asymptotic variance for axially symmetric distributions:

Lemma 4.6. *If a distribution μ on $\mathbb{R}^2$ is axially symmetric, then the asymptotic variance of the tau-estimator is given by*

$$\sigma_\tau^2 = 16\int_{(0,\infty)^2}\Big(\mu\big((-x,x]\times(-y,y]\big)\Big)^2\,\mu(dx,dy)\,. \tag{4.7}$$

Proof. By the axial symmetry it is easy to see that for $x, y \geq 0$ one has

$$\mathbb{E}\big[\,\kappa_\tau\big((x,y),(X,Y)\big)\,\big] = \mathbb{E}\big[\,\kappa_\tau\big((-x,-y),(X,Y)\big)\,\big]\,,$$

where $(X,Y) \sim \mu$. Assuming $x, y > 0$ we can apply Lemma 4.5. Additionally using the symmetry we can calculate that

$$\begin{aligned}
\mathbb{E}\big[\,\kappa_\tau\big((x,y),(X,Y)\big)\,\big] &= \frac{1}{2}\,\mathbb{E}\big[\,\kappa_\tau\big((x,y),(X,Y)\big) + \kappa_\tau\big((-x,-y),(X,Y)\big)\,\big] \\
&= \mu\big((-x,x)\times(-y,y)\big) \\
&\quad + \frac{1}{2}\,\mu\big(\{-x,x\}\times(-y,y)\big) + \frac{1}{2}\,\mu\big((-x,x)\times\{-y,y\}\big) \\
&\quad + \frac{1}{2}\,\mu\big(\{-x\}\times\{-y\}\big) + \frac{1}{2}\,\mu\big(\{x\}\times\{y\}\big) \qquad (4.8) \\
&= \mu\big((-x,x)\times(-y,y)\big) \\
&\quad + \mu\big(\{x\}\times(-y,y)\big) + \mu\big((-x,x)\times\{y\}\big) \\
&\quad + \mu\big(\{x\}\times\{y\}\big) \\
&= \mu\big((-x,x]\times(-y,y]\big)\,.
\end{aligned}$$

If we take $x = 0$ it is easy to see that due to the symmetry we have

$$\mathbb{E}\big[\,\kappa_\tau\big((0,y),(X,Y)\big) + \kappa_\tau\big((0,-y),(X,Y)\big)\,\big] = 0\,.$$

As further

$$\mu\big((0,0]\times(-y,y]\big) = \mu(\varnothing) = 0\,,$$

the result of equation (4.8) holds true for all $x \geq 0$ and $y > 0$ or, if we argue analogously for $y = 0$, for all $x, y \geq 0$. If we change the sign of one variable, then axial symmetry implies that the sign of the expectation changes:

$$\begin{aligned}
&\mathbb{E}\big[\,\kappa_\tau\big((x,y),(X,Y)\big)\,\big] = \mathbb{E}\big[\,\kappa_\tau\big((-x,-y),(X,Y)\big)\,\big] \\
&\quad = -\mathbb{E}\big[\,\kappa_\tau\big((-x,y),(X,Y)\big)\,\big] = -\mathbb{E}\big[\,\kappa_\tau\big((x,-y),(X,Y)\big)\,\big]\,.
\end{aligned}$$

For the asymptotic variance this value gets squared and therefore the sign does not matter and we have for all $x, y \in \mathbb{R}$

$$\Big(\mathbb{E}\big[\,\kappa_\tau\big((x,y),(X,Y)\big)\,\big]\Big)^2 = \Big(\mu\big((-|x|,|x|]\times(-|y|,|y|]\big)\Big)^2\,.$$

But for the calculation of Kendall's tau itself the change of sign effects that by integrating over $\mathbb{R}^2$ the terms cancel out and we get $\tau = 0$. Therefore the asymptotic variance equals

$$\begin{aligned}
\sigma_\tau^2 &= 4\int_{\mathbb{R}^2}\Big(\mu\big((-|x|,|x|]\times(-|y|,|y|]\big)\Big)^2\,\mu(dx,dy) \\
&= 16\int_{(0,\infty)^2}\Big(\mu\big((-x,x]\times(-y,y]\big)\Big)^2\,\mu(dx,dy)\,.
\end{aligned}$$

□

If there exists a density, then this formula can be written as follows.

Corollary 4.7. *If a distribution μ on $\mathbb{R}^2$ is axially symmetric and possesses a density f, then the asymptotic variance of the tau-estimator is given by*

$$\sigma_\tau^2 = 4 \int_{\mathbb{R}^2} \left(4 \int_0^y \int_0^x f(u,v)\, du\, dv \right)^2 f(x,y)\, d(x,y)\,. \tag{4.9}$$

Remark 4.8. The symmetry within Lemma 4.6 and Corollary 4.7 can also be defined with respect to the lines $x = \mu_X$ and $y = \mu_Y$ where the conclusions, formulated in an analogue way, hold true.

Within the next two sections we look at special axially symmetric distributions where we have further simplifications beyond formula (4.9). In Section 6.1 the formula will again be used to calculate the asymptotic variance of the tau-estimator for special spherical distributions.

4.3 Symmetric decomposable densities

Formula (4.9) for the asymptotic variance can further be simplified if the two-dimensional density can be decomposed into a product where the factors only contain one of the variables. This is e.g. true for the standard normal distribution. The simplification still works when the density is a sum of summands that possess this decomposition property. So we now look at densities of the form

$$f(x,y) = \sum_{i \in \mathcal{I}} c_i\, g_i(x)\, h_i(y)\,, \tag{4.10}$$

where $\mathcal{I}$ is a finite index set and where the constants $c_i \in \mathbb{R}$ and the integrable functions $g_i, h_i : \mathbb{R} \to [0, \infty)$ are chosen such that f is an axially symmetric density. Using formula (4.9) and writing $G_i(x) := \int_0^x g_i(u)\, du$ and $H_i(x) := \int_0^y h_i(u)\, du$ we get

$$\begin{aligned}
\sigma_\tau^2 &= 64 \int_{\mathbb{R}^2} \left(\int_0^y \int_0^x f(u,v)\, du\, dv \right)^2 f(x,y)\, d(x,y) \\
&= 64 \int_{\mathbb{R}^2} \left(\sum_{i \in \mathcal{I}} c_i\, G_i(x)\, H_i(y) \right)^2 \left(\sum_{l \in \mathcal{I}} c_l\, g_l(x)\, h_l(y) \right) d(x,y) \\
&= 64 \int_{\mathbb{R}^2} \left(\sum_{i,j,l \in \mathcal{I}} c_i\, c_j\, c_l\, G_i(x)\, G_j(x)\, g_l(x)\, H_i(y)\, H_j(y)\, h_l(y) \right) d(x,y) \\
&= 64 \sum_{i,j,l \in \mathcal{I}} c_i\, c_j\, c_l \left(\int_{\mathbb{R}} G_i(x)\, G_j(x)\, g_l(x)\, dx \int_{\mathbb{R}} H_i(y)\, H_j(y)\, h_l(y)\, dy \right).
\end{aligned} \tag{4.11}$$

An example of a decomposable density can be found in Section 5.2.

4.4 Normal variance mixture distributions

4.4.1 Asymptotic variance of the tau-estimator

This section revisits normal variance mixture distributions, as defined in Section 3.3. Since we are interested in simplifying the calculation of the asymptotic variance of the tau-estimator, we only consider the two-dimensional case. Moreover we only look at standard normal variance mixture distributions to ensure axial symmetry. To use formula (4.9) we further need the existence of a density, which is given if the mixing variable W has no point mass at zero, $G(0) = 0$. In this case representation (3.24) of a density f simplifies to

$$f(u,v) = \int_0^\infty \psi_\zeta(u)\,\psi_\zeta(v)\,G(d\zeta), \quad u,v \in \mathbb{R}\,, \tag{4.12}$$

where ψ_ζ denotes the continuous density of the $\mathcal{N}(0,\zeta)$-distribution, given by

$$\psi_\zeta(u) = \frac{1}{\sqrt{2\pi\zeta}}\exp\Big(-\frac{u^2}{2\zeta}\Big), \quad u \in \mathbb{R},\ \zeta > 0\,.$$

Theorem 4.9. *For a two-dimensional standard normal variance mixture distribution with mixing distribution function G that fulfils $G(0) = 0$, the asymptotic variance of the tau-estimator simplifies to*

$$\sigma_\tau^2 = \frac{16}{\pi^2}\int_0^\infty\int_0^\infty\int_0^\infty \arctan^2\Big(\frac{\zeta}{\sqrt{\upsilon\,\xi + \upsilon\,\zeta + \xi\,\zeta}}\Big)\,G(d\upsilon)\,G(d\xi)\,G(d\zeta)\,. \tag{4.13}$$

Proof. As we are given a spherical distribution with existing density, we can start with formula (4.9) for the asymptotic variance of the tau-estimator. Writing

$$\Psi_\zeta(x) = \int_0^x \psi_\zeta(u)\,du, \quad x \in \mathbb{R},\ \zeta > 0\,,$$

we can use representation (4.12) of the two-dimensional density and get

$$\begin{aligned}
\sigma_\tau^2 &= 4\int_{\mathbb{R}^2}\Big(4\int_0^y\int_0^x f(u,v)\,du\,dv\Big)^2 f(x,y)\,d(x,y) \\
&= 4^3\int_{\mathbb{R}^2}\Big(\int_0^\infty \Psi_\upsilon(x)\,\Psi_\upsilon(y)\,G(d\upsilon)\Big)\Big(\int_0^\infty \Psi_\xi(x)\,\Psi_\xi(y)\,G(d\xi)\Big) \\
&\qquad\qquad \Big(\int_0^\infty \psi_\zeta(x)\,\psi_\zeta(y)\,G(d\zeta)\Big)\,d(x,y) \\
&= 4\int_0^\infty\int_0^\infty\int_0^\infty \big(H(\upsilon,\xi,\zeta)\big)^2\,G(d\upsilon)\,G(d\xi)\,G(d\zeta)
\end{aligned}$$

where

$$H(\upsilon,\xi,\zeta) := 4\int_{\mathbb{R}} \Psi_\upsilon(x)\,\Psi_\xi(x)\,\psi_\zeta(x)\,dx\,, \quad \upsilon,\xi,\zeta > 0\,.$$

So we only have to show that

$$H(\upsilon,\xi,\zeta) = \frac{2}{\pi}\,\arctan\Big(\frac{\zeta}{\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}}\Big) =: G(\upsilon,\xi,\zeta)\,.$$

We first want to fix the third argument and treat G_ζ and H_ζ as functions on $(0,\infty)^2$ with fixed parameter ζ. We want to show that

$$G_\zeta(\upsilon,\xi) = H_\zeta(\upsilon,\xi) \quad \text{for all } \zeta > 0\,. \tag{4.14}$$

Therefor we claim that both functions are a potential of the same gradient field, i.e. that $\nabla G_\zeta = \nabla H_\zeta$. As both functions are symmetric in the first two arguments it is enough to show that

$$\frac{\partial}{\partial\upsilon}\,G_\zeta(\upsilon,\xi) = \frac{\partial}{\partial\upsilon}\,H_\zeta(\upsilon,\xi)\,.$$

We first simplify the derivative $\frac{\partial}{\partial\upsilon}\,\Psi_\upsilon(x)$. Using the representation $\Psi_\upsilon(x) = \Psi_1(\frac{x}{\sqrt{\upsilon}})$, where the variable υ is only in the upper bound and no longer in the integrand of the integral, it is easy to see that

$$\begin{aligned}\frac{\partial}{\partial\upsilon}\,\Psi_\upsilon(x) &= -\frac{x}{2\sqrt{\upsilon^3}}\,\psi_1\big(\frac{x}{\sqrt{\upsilon}}\big)\\ &= -\frac{x}{2\upsilon}\,\psi_\upsilon(x)\,.\end{aligned}$$

Hence we know that

$$\begin{aligned}\frac{\partial}{\partial\upsilon}\,H_\zeta(\upsilon,\xi) &= 4\int_{\mathbb{R}}\Big(\frac{\partial}{\partial\upsilon}\,\Psi_\upsilon(x)\Big)\,\Psi_\xi(x)\,\psi_\zeta(x)\,dx\\ &= -\frac{2}{\upsilon}\int_{\mathbb{R}} x\,\psi_\upsilon(x)\,\Psi_\xi(x)\,\psi_\zeta(x)\,dx\,.\end{aligned}$$

To solve this integral we want to use integration by parts. Therefor we need

$$\begin{aligned}\int x\,\psi_\upsilon(x)\,\psi_\zeta(x)\,dx &= \frac{1}{2\pi\,\sqrt{\upsilon\,\zeta}}\int x\,\exp\Big(-\frac{(\upsilon+\zeta)\,x^2}{2\,\upsilon\,\zeta}\Big)\,dx\\ &= \frac{-\sqrt{\upsilon\,\zeta}}{2\pi\,(\upsilon+\zeta)}\,\exp\Big(-\frac{(\upsilon+\zeta)\,x^2}{2\,\upsilon\,\zeta}\Big)\,.\end{aligned}$$

For the derivative of H_ζ we get

$$\begin{aligned}\frac{\partial}{\partial\upsilon}\,H_\zeta(\upsilon,\xi) &= -\frac{2}{\upsilon}\int_{\mathbb{R}} x\,\psi_\upsilon(x)\,\Psi_\xi(x)\,\psi_\zeta(x)\,dx\\ &= -\frac{2}{\upsilon}\left(\frac{-\sqrt{\upsilon\,\zeta}}{2\pi\,(\upsilon+\zeta)}\,\exp\Big(-\frac{(\upsilon+\zeta)\,x^2}{2\,\upsilon\,\zeta}\Big)\,\Psi_\xi(x)\right)\Bigg|_{x=-\infty}^{\infty}\\ &\quad+\frac{2}{\upsilon}\int_{\mathbb{R}}\Big(\frac{\partial}{\partial x}\,\Psi_\xi(x)\Big)\,\frac{-\sqrt{\upsilon\,\zeta}}{2\pi\,(\upsilon+\zeta)}\,\exp\Big(-\frac{(\upsilon+\zeta)\,x^2}{2\,\upsilon\,\zeta}\Big)\,dx\\ &= -\frac{\sqrt{\zeta}}{\pi\,\sqrt{2\pi}\,\sqrt{\upsilon\,\xi}\,(\upsilon+\zeta)}\int_{\mathbb{R}}\exp\Big(-\frac{x^2}{2}\,\Big(\frac{1}{\upsilon}+\frac{1}{\xi}+\frac{1}{\zeta}\Big)\Big)\,dx\\ &= -\frac{\sqrt{\zeta}}{\pi\,\sqrt{\upsilon\,\xi}\,(\upsilon+\zeta)}\,\Big(\frac{1}{\upsilon}+\frac{1}{\xi}+\frac{1}{\zeta}\Big)^{-\frac{1}{2}} = -\frac{\zeta}{\pi\,(\upsilon+\zeta)\,\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}}\,.\end{aligned}$$

One can easily verify that G_ζ has exactly the same derivative:

$$\begin{aligned}
\frac{\partial}{\partial \upsilon} G_\zeta(\upsilon,\xi) &= \frac{\partial}{\partial \upsilon}\Big(\frac{2}{\pi}\arctan\Big(\frac{\zeta}{\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}}\Big)\Big) \\
&= \frac{2}{\pi}\Big(1+\Big(\frac{\zeta}{\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}}\Big)^2\Big)^{-1}\Big(-\frac{\zeta\,(\xi+\zeta)}{2(\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta})^3}\Big) \\
&= -\frac{2\,\zeta}{\pi\,(\upsilon+\zeta)\,\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}} \\
&= \frac{\partial}{\partial \upsilon} H_\zeta(\upsilon,\xi)\,.
\end{aligned}$$

As $(0,\infty)^2$ is an open and connected set and therefore a domain and as the two functions G_ζ and H_ζ are potentials of the same gradient field, the functions can only differ in an additive constant (see e.g. Amann and Escher, 1999, p. 321). So it just remains to show that they are equal in one point. For every $\zeta > 0$ we have

$$\begin{aligned}
H_\zeta(\zeta,\zeta) &= 4\int_{\mathbb{R}}\big(\Psi_\zeta(x)\big)^2\,\psi_\zeta(x)\,dx = 4\int_{\mathbb{R}}\frac{\partial}{\partial x}\Big(\frac{1}{3}\big(\Psi_\zeta(x)\big)^3\Big)\,dx \\
&= \frac{4}{3}\Big(\big(\Psi_\zeta(\infty)\big)^3-\big(\Psi_\zeta(-\infty)\big)^3\Big) = \frac{4}{3}\Big(\frac{1}{2^3}+\frac{1}{2^3}\Big) \\
&= \frac{1}{3}
\end{aligned}$$

and therefore

$$G_\zeta(\zeta,\zeta) = \frac{2}{\pi}\arctan\Big(\frac{\zeta}{\sqrt{3\,\zeta^2}}\Big) = \frac{1}{3} = H_\zeta(\zeta,\zeta)\,.$$

As the resulting equality

$$G_\zeta(\upsilon,\xi) = H_\zeta(\upsilon,\xi)$$

is true for all $\zeta \in (0,\infty)$ we also have

$$G(\upsilon,\xi,\zeta) = H(\upsilon,\xi,\zeta)$$

which we wanted to prove. $\square$

Formula (4.13) consists of three integrals. We can reduce this to two integrals, but we loose the independence of the components. This can be seen by rewriting the argument of the arc tangent in the following way:

$$\frac{\zeta}{\sqrt{\upsilon\,\xi+\upsilon\,\zeta+\xi\,\zeta}} = \Big(\frac{\upsilon}{\zeta}\,\frac{\xi}{\zeta}+\frac{\upsilon}{\zeta}+\frac{\xi}{\zeta}\Big)^{-\frac{1}{2}}\,.$$

We want to define the two-dimensional distribution function H as

$$H(x,y) := \mathbb{P}\Big[\,\frac{X}{Z}\le x,\,\frac{Y}{Z}\le y\,\Big],\quad x,y\in[0,\infty)\,, \tag{4.15}$$

where X, Y and Z are i.i.d. random variables with distribution function G. This is well-defined as we require that $G(0) = 0$. So we have

$$H(x, y) = \int_0^\infty G(xz)\, G(yz)\, G(dz)\,, \tag{4.16}$$

and the asymptotic variance of the tau-estimator can be simplified as follows.

Corollary 4.10. *For a two-dimensional standard normal variance mixture distribution with mixing distribution function G that fulfils $G(0) = 0$, the asymptotic variance of the tau-estimator simplifies to*

$$\sigma_\tau^2 = \frac{16}{\pi^2} \int_{(0,\infty)^2} \arctan^2\Big(\frac{1}{\sqrt{x + y + xy}}\Big)\, H(dx, dy)\,, \tag{4.17}$$

where H is the distribution function as defined in (4.15)*.*

4.4.2 An example where the mixing distribution is inverse gamma

There are examples where H can be calculated explicitly, e.g. when G is the inverse gamma distribution. We first show the definition and basic properties of this distribution before we simplify the formula for the asymptotic variance of the tau-estimator.

The inverse gamma distribution

Definition 4.11. A non-negative random variable X follows a *gamma distribution* with parameters $\alpha, \beta > 0$, $X \sim \mathrm{Ga}(\alpha, \beta)$, if

$$\mathbb{P}[X \le x] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^x z^{\alpha-1}\, e^{-\beta z}\, dz\,, \quad x \ge 0\,.$$

Definition 4.12. A non-negative random variable X follows an *inverse gamma distribution* with parameters $\alpha, \beta > 0$, $X \sim \mathrm{Ig}(\alpha, \beta)$, if $\frac{1}{X} \sim \mathrm{Ga}(\alpha, \beta)$.

We can calculate the distribution function for the inverse gamma distribution in the following way:

$$\begin{aligned}
\mathbb{P}[\, X \le x\,] = \mathbb{P}\Big[\, \frac{1}{X} \ge \frac{1}{x}\, \Big] &= 1 - \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{\frac{1}{x}} z^{\alpha-1}\, e^{-\beta z}\, dz \\
&= 1 - \frac{\beta^\alpha}{\Gamma(\alpha)} \int_x^\infty u^{-\alpha-1}\, e^{-\frac{\beta}{u}}\, du \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^x u^{-\alpha-1}\, e^{-\frac{\beta}{u}}\, du\,, \quad x \ge 0\,.
\end{aligned}$$

In Section 6.4.3 we will further need the first and the second moment of the inverse gamma distribution to calculate the asymptotic variance of the standard estimator for the t-distribution.

Lemma 4.13. *Let $X \sim \mathrm{Ig}(\alpha, \beta)$. Then, if $\alpha > 1$,*

$$\mathbb{E}[X] = \frac{\beta}{\alpha - 1}$$

and, if $\alpha > 2$,

$$\mathbb{V}\mathrm{ar}[X] = \frac{\beta^2}{(\alpha-1)^2\,(\alpha-2)}\,.$$

Proof. We know that for all $0 \le \delta < \alpha$

$$\mathbb{E}[X^\delta] = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{-\alpha-1+\delta}\, e^{-\frac{\beta}{x}}\, dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty z^{\alpha-1-\delta}\, e^{-\beta z}\, dz\,.$$

The integrand is, up to a constant, a density of the gamma distribution $\mathrm{Ga}(\alpha - \delta, \beta)$, so we know

$$\mathbb{E}[X^\delta] = \frac{\beta^\delta\, \Gamma(\alpha - \delta)}{\Gamma(\alpha)}$$

and for $\delta \in \mathbb{N}$

$$\mathbb{E}[X^\delta] = \frac{\beta^\delta}{(\alpha-1)\,(\alpha-2)\,\ldots\,(\alpha-\delta)}\,.$$

□

Lemma 4.14. *The sequence of inverse gamma distributions $\big(\mathrm{Ig}(\alpha,\alpha)\big)_{\alpha>0}$ converges weakly to a Dirac measure in 1 as $\alpha \to \infty$.*

Proof. If $X \sim \mathrm{Ig}(\alpha, \alpha)$, then

$$\mathbb{E}[X] = \frac{\alpha}{\alpha - 1} \overset{\alpha\to\infty}{\longrightarrow} 1 \quad \text{and} \quad \mathbb{V}\mathrm{ar}[X] = \frac{\alpha^2}{(\alpha-1)^2\,(\alpha-2)} \overset{\alpha\to\infty}{\longrightarrow} 0\,.$$

So X converges in probability to a degenerate random variable of value 1 and therefore its distribution converges weakly to the Dirac measure. □

Formula for the asymptotic variance of the tau-estimator

The choice, that G is inverse gamma, is motivated by the fact, that the inverse gamma distribution with parameters $\alpha = \beta$ as mixing distribution leads to a t-distribution, a well-known family of elliptical distributions. In Section 6.4 we will use the derived simplifications to calculate the asymptotic variance of the tau-estimator for uncorrelated t-distributions.

Lemma 4.15. *For a two-dimensional standard normal variance mixture distribution where the mixing distribution function G is an inverse gamma distribution $\mathrm{Ig}(\alpha, \beta)$, $\alpha, \beta > 0$, the asymptotic variance of the tau-estimator simplifies to*

$$\sigma_\tau^2 = \frac{16\,\Gamma(3\alpha)}{\pi^2\,\Gamma^3(\alpha)} \int_0^\infty \int_0^\infty \frac{(xy)^{2\alpha-1}}{(x+y+xy)^{3\alpha}}\, \arctan^2\!\Big(\frac{1}{\sqrt{x+y+xy}}\Big)\, dx\, dy\,. \tag{4.18}$$

Proof. We use representation (4.16) to determine the two-dimensional distribution function H. As the inverse gamma distribution has a density so does H, and we can derive such a density $h(x,y)$, $x,y \in (0,\infty)$, as

$$\begin{aligned}
h(x,y) &= \int_0^\infty z^2\, g(xz)\, g(yz)\, g(z)\, dz \\
&= \frac{\beta^{3\alpha}}{\Gamma^3(\alpha)}\, (xy)^{-\alpha-1} \int_0^\infty z^{-3\alpha-1}\, \exp\Big(-\frac{\beta}{z}\,\big(1+\frac{1}{x}+\frac{1}{y}\big)\Big)\, dz \\
&= \frac{\beta^{3\alpha}}{\Gamma^3(\alpha)}\, (xy)^{-\alpha-1} \int_0^\infty u^{3\alpha-1}\, \exp\Big(-\beta u\,\big(1+\frac{1}{x}+\frac{1}{y}\big)\Big)\, du \\
&= \frac{\beta^{3\alpha}}{\Gamma^3(\alpha)}\, (xy)^{-\alpha-1}\, \Gamma(3\alpha)\, \beta^{-3\alpha} \Big(1+\frac{1}{x}+\frac{1}{y}\Big)^{-3\alpha} \\
&= \frac{\Gamma(3\alpha)}{\Gamma^3(\alpha)}\, (xy)^{2\alpha-1}\, (x+y+xy)^{-3\alpha}\,,
\end{aligned}$$

where the integral is solved by realizing that it is, up to a constant, an integral over a density of the gamma distribution. Inserting this density in (4.17) gives the formula claimed in the lemma. □

Remark 4.16. Note that the scale parameter β vanishes, as was expected due to the definition in (4.15).

Chapter 5

Examples of the asymptotic variance of the tau-estimator for several distributions

5.1 Examples for several copulas

Within this chapter we want to apply the formulas that we derived in Chapter 4. We start with examples of different copulas where we can use the results from Section 4.1. As those results require continuity of the random variables we also assume it within this section.

5.1.1 Product copula

Two independent random variables possess the product copula, which is defined as follows.

Definition 5.1. The *product copula* is the function

$$C^{\perp}(u,v) = u\,v, \quad u,v \in [0,1]\,.$$

Lemma 5.2. *For two continuous random variables linked by the product copula $C^{\perp}$, Kendall's tau has the value*

$$\tau^{\perp} = 0$$

and the asymptotic variance of the tau-estimator, as defined in (2.15)*, equals*

$$\left(\sigma_{\tau}^{\perp}\right)^2 = \frac{4}{9}\,.$$

Proof. Since U and V are independent we know that $\tau^{\perp} = 0$. Starting with formula (4.5) we can again use the independence and the moments of $\mathcal{U}[0,1]$-distributed random variables to get

$$\begin{aligned}(\sigma_\tau^\perp)^2 &= 4\Big(16\,\mathbb{E}[\,U^2\,V^2\,] - 32\,\mathbb{E}[\,U^2\,V\,] + 8\,\mathbb{E}[\,U\,V\,] + \frac{5}{3}\Big)\\ &= 4\Big(16\,\mathbb{E}[\,U^2\,]\,\mathbb{E}[\,V^2\,] - 32\,\mathbb{E}[\,U^2\,]\,\mathbb{E}[\,V\,] + 8\,\mathbb{E}[\,U\,]\,\mathbb{E}[\,V\,] + \frac{5}{3}\Big)\\ &= \frac{4}{9}\,.\end{aligned}$$

□

5.1.2 Archimedean copulas

Archimedean copulas are widely used due to their nice construction and the resulting properties. They are constructed with the help of a function φ, called generator:

Definition 5.3.

(a) A *generator* of an Archimedean copula is a continuous, strictly decreasing and convex function $\varphi : [0,1] \to [0,\infty]$ such that $\varphi(1) = 0$.

(b) The *pseudo-inverse* of φ is the function $\varphi^{[-1]} : [0,\infty] \to [0,1]$ given by

$$\varphi^{[-1]}(t) = \begin{cases}\varphi^{-1}(t)\,, & \text{if } t \in [0,\varphi(0)]\,,\\ 0\,, & \text{if } t \in (\varphi(0),\infty]\,.\end{cases}$$

(c) If $\varphi(0) = \infty$, then the pseudo-inverse is the normal inverse function, $\varphi^{[-1]} = \varphi^{-1}$. In this case φ is called a *strict* generator.

Definition 5.4.

(a) An *Archimedean copula* with generator φ is a function

$$C(u,v) = \varphi^{[-1]}\big(\varphi(u) + \varphi(v)\big), \quad u, v \in [0,1]\,.$$

(b) If φ is a strict generator, then C is called a *strict* Archimedean copula.

The proof, that the function C, as defined in Definition 5.4, is indeed a copula can e.g. be found in Nelsen (2006, pp. 111–112).

Example 5.5. The product copula is a strict Archimedean copula with generator $\varphi^\perp(t) = -\log(t)$.

One of the nice properties that copulas of the Archimedean family possess is an easy formula for Kendall's tau:

Lemma 5.6. *Let X and Y be continuous random variables with an Archimedean copula C generated by φ. Then Kendall's tau is given by*

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)}\, dt\,. \tag{5.1}$$

Proof. See e.g. Nelsen (2006, p. 163). □

Remark 5.7. For formula (5.1) we do not have to assume that φ is differentiable. As φ is convex on $(0,1)$ we know that the set where φ' fails to exist is countable (see e.g. Roberts and Varberg, 1973, p. 7).

Clayton copula

The Clayton copula is a one-parameter Archimedean copula with generator

$$\varphi^{\mathrm{Cl},\theta}(t) = \frac{1}{\theta}\,(t^{-\theta} - 1), \quad t \in (0,1]\,,$$

and with parameter $\theta \in [-1,\infty) \setminus \{0\}$.

Definition 5.8. A *Clayton copula* with parameter $\theta \in [-1,\infty) \setminus \{0\}$ is a function

$$C^{\mathrm{Cl},\theta}(u,v) = \begin{cases} \big(\max(u^{-\theta} + v^{-\theta} - 1, 0)\big)^{-\frac{1}{\theta}}\,, & \text{if } u,v > 0\,,\\ 0\,, & \text{otherwise}\,, \end{cases} \qquad u,v \in [0,1]\,.$$

Remark 5.9.

(a) All the three special copulas, the product copula and the two Fréchet–Hoeffding bounds, as introduced in Section 4.1, can be reached by a Clayton copula. We get the product copula in the limit $\theta \to 0$, the lower Fréchet–Hoeffding bound for $\theta = -1$ and the upper Fréchet–Hoeffding bound for $\theta \to \infty$.

(b) On the full range of the parameter the generator and the copula are not strict. If we only allow $\theta \in (0,\infty)$, then we get a strict Archimedean copula and can simplify the copula to

$$C^{\mathrm{Cl},\theta}(u,v) = \begin{cases} \big(u^{-\theta} + v^{-\theta} - 1\big)^{-\frac{1}{\theta}}\,, & \text{if } u,v > 0\,,\\ 0\,, & \text{otherwise}\,, \end{cases} \qquad u,v \in [0,1]\,.$$

In this case only positive dependence can be modelled.

Lemma 5.10. *For two continuous random variables linked by a Clayton copula $C^{\mathrm{Cl},\theta}$ with parameter $\theta \in [-1,\infty) \setminus \{0\}$, Kendall's tau has the value*

$$\tau^{\mathrm{Cl},\theta} = \frac{\theta}{\theta + 2}\,.$$

For $\theta \in (0, \infty)$ the asymptotic variance of the tau-estimator, as defined in (2.15), equals

$$\begin{aligned}\left(\sigma_\tau^{\mathrm{Cl},\theta}\right)^2 &= \frac{16}{3}\left(\frac{1+3\,\theta}{3+\theta} - \frac{(1+\theta)\,(7+5\,\theta)}{(2+\theta)^2}\right) \\ &\quad + 32\,(1+\theta)\int_{(0,1]^2} u^{-\theta}\,v^{-\theta}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}-2}\,d(u,v)\,.\end{aligned} \tag{5.2}$$

In the special cases $\theta = 1$ and $\theta = 2$ we get the values

$$\left(\sigma_\tau^{\mathrm{Cl},1}\right)^2 = \frac{16}{9}\left(6\,\pi^2 - 59\right) \approx 0.387$$

and

$$\left(\sigma_\tau^{\mathrm{Cl},2}\right)^2 = \frac{337}{15} - 32\,\log(2) \approx 0.286\,.$$

Proof. We use formula (5.1) to calculate Kendall's tau. The derivative of the generator is

$$\left(\varphi^{\mathrm{Cl},\theta}\right)'(t) = -t^{-\theta-1}\,,$$

such that we get

$$\begin{aligned}\tau^{\mathrm{Cl},\theta} &= 1 + \frac{4}{\theta}\int_0^1 \left(t^{\theta+1} - t\right)dt = 1 + \frac{4}{\theta}\left(\frac{1}{\theta+2} - \frac{1}{2}\right) \\ &= \frac{\theta}{\theta+2}\,.\end{aligned}$$

To calculate the asymptotic variance of the tau-estimator we use formula (4.5). As we set $\theta \in (0, \infty)$ we can use the simple form of the copula and can determine a density for $u, v > 0$ as

$$\begin{aligned}c^{\mathrm{Cl},\theta}(u,v) &= \frac{\partial^2}{\partial u\,\partial v}\,C^{\mathrm{Cl},\theta}(u,v) \\ &= \frac{\partial^2}{\partial u\,\partial v}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}} \\ &= (\theta+1)\,u^{-\theta-1}\,v^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}-2}, \quad u,v \in (0,1]\,.\end{aligned}$$

We need to solve the integrals

$$\begin{aligned}&\int_{(0,1]^2}\left(C^{\mathrm{Cl},\theta}(u,v)\right)^2 c^{\mathrm{Cl},\theta}(u,v)\,d(u,v) \\ &\quad = (1+\theta)\int_{(0,1]^2} u^{-\theta-1}\,v^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{3}{\theta}-2}\,d(u,v)\end{aligned}$$

and

$$\begin{aligned}&\int_{(0,1]^2} u\,C^{\mathrm{Cl},\theta}(u,v)\,c^{\mathrm{Cl},\theta}(u,v)\,d(u,v) \\ &\quad = (1+\theta)\int_{(0,1]^2} u^{-\theta}\,v^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{2}{\theta}-2}\,d(u,v)\,.\end{aligned}$$

They can be treated equally. We start with the integral over v by substituting $z := v^{-\theta}$ (for the first integral set $\alpha = 3$, for the second $\alpha = 2$):

$$\begin{aligned}\int_0^1 v^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{\alpha}{\theta}-2} dv &= \frac{1}{\theta}\int_1^\infty \left(u^{-\theta}+z-1\right)^{-\frac{\alpha}{\theta}-2} dz \\ &= \frac{1}{\theta\left(-\frac{\alpha}{\theta}-1\right)}\left(u^{-\theta}+z-1\right)^{-\frac{\alpha}{\theta}-1}\Big|_{z=1}^{\infty} \\ &= \frac{1}{\alpha+\theta}\, u^{\alpha+\theta}\,.\end{aligned}$$

So the whole integral equals

$$\int_{(0,1]^2} u^{-\theta-\alpha+2}\, v^{-\theta-1}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{\alpha}{\theta}-2} d(u,v) = \frac{1}{\alpha+\theta}\int_0^1 u^2\, du = \frac{1}{3\,(\alpha+\theta)}\,.$$

The third term in formula (4.5) can be written as

$$\int_{(0,1]^2} u\, v\, c^{\mathrm{Cl},\theta}(u,v)\, d(u,v) = (1+\theta)\int_{(0,1]^2} u^{-\theta}\, v^{-\theta}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}-2} d(u,v)\,.$$

Inserting the calculated values as well as the value of $\tau^{\mathrm{Cl},\theta}$ we get

$$\begin{aligned}\left(\sigma_\tau^{\mathrm{Cl},\theta}\right)^2 &= \frac{64\,(1+\theta)}{3\,(3+\theta)} - \frac{128\,(1+\theta)}{3\,(2+\theta)} + \frac{20}{3} + \frac{8\,\theta}{2+\theta} - \frac{4\,\theta^2}{(2+\theta)^2} \\ &\quad + 32\,(1+\theta)\int_{(0,1]^2} u^{-\theta}\, v^{-\theta}\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}-2} d(u,v)\,.\end{aligned}$$

Simplifications give the general formula claimed in the lemma. To solve the last remaining integral for $\theta = 1$ we use the substitution $z := (1-v)\,u + v$ to get

$$\begin{aligned}&\int_{(0,1]^2} u^{-1}\, v^{-1}\left(u^{-1}+v^{-1}-1\right)^{-3} d(u,v) \\ &= \int_0^1 v^2 \int_0^1 \frac{u^2}{\left((1-v)\,u+v\right)^3}\, du\, dv = \int_0^1 \frac{v^2}{(1-v)^3}\int_v^1 \frac{(z-v)^2}{z^3}\, dz\, dv \\ &= \int_0^1 \frac{v^2}{(1-v)^3}\left(-\frac{v^2}{2z^2}+\frac{2v}{z}+\log(z)\right)\Big|_{z=v}^{1} dv \\ &= \int_0^1 \frac{v^2}{(1-v)^3}\left(-\frac{v^2}{2}+2v-\frac{3}{2}-\log(v)\right) dv \\ &= \int_0^1 \frac{(1-z)^2}{z^3}\left(-\frac{z^2}{2}-z-\log(1-z)\right) dz \\ &=: (*)\end{aligned}$$

where we used the substitution $z := 1-v$ in the last step. Expanding the sums we can calculate the integral over several summands directly (Li_2 denotes the dilogarithm, see

Section 7.4.1 for details):

$$
\begin{aligned}
(*) &= \int_0^1 \left(\frac{3}{2z} - \frac{1}{z^2} - \left(\frac{1}{z^3} - \frac{2}{z^2}\right) \log(1-z)\right) dz - \int_0^1 \frac{z}{2}\, dz - \int_0^1 \frac{1}{z} \log(1-z)\, dz \\
&= \left(\frac{1}{2z} + \left(\frac{1}{2z^2} - \frac{2}{z} + \frac{3}{2}\right) \log(1-z)\right) \Big|_{z=0}^{1} - \frac{z^2}{4} \Big|_{z=0}^{1} + \mathrm{Li}_2(1-z) \Big|_{z=0}^{1} \\
&= \frac{1}{2} - \lim_{z \searrow 0} \left(\frac{1}{2z} + \left(\frac{1}{2z^2} - \frac{2}{z}\right) \log(1-z)\right) - \frac{1}{4} + \frac{\pi^2}{6}\,.
\end{aligned}
$$

The remaining limit can be determined by l'Hôpital's rule to be

$$
\lim_{z \searrow 0} \left(\frac{1}{2z} + \left(\frac{1}{2z^2} - \frac{2}{z}\right) \log(1-z)\right) = \frac{7}{4}\,.
$$

Putting everything together we get the value for the asymptotic variance in the case $\theta = 1$ claimed in the lemma.

We can also find a solution for $\theta = 2$. We use the substitution $z := (1 - v^2)\, u^2 + v^2$, similar to the previous case:

$$
\begin{aligned}
&\int_{(0,1]^2} u^{-2}\, v^{-2} \left(u^{-2} + v^{-2} - 1\right)^{-\frac{5}{2}} d(u,v) \\
&\quad = \int_0^1 v^3 \int_0^1 \frac{u^3}{((1-v^2)\, u^2 + v^2)^{\frac{5}{2}}}\, du\, dv = \frac{1}{2} \int_0^1 \frac{v^3}{(1-v^2)^2} \int_{v^2}^1 \frac{z - v^2}{z^{\frac{5}{2}}}\, dz\, dv \\
&\quad = \frac{1}{3} \int_0^1 \frac{v^3}{(1-v^2)^2} \left(\frac{v^2 - 3z}{z^{\frac{3}{2}}}\right) \Big|_{z=v^2}^{1} dv = \frac{1}{3} \int_0^1 \frac{v^3}{(1-v^2)^2}\, \frac{v^3 - 3v + 2}{v}\, dv \\
&\quad = \frac{1}{3} \int_0^1 \frac{(2+v)\, v^2}{(1+v)^2}\, dv = \frac{1}{6} \left(v^2 - \frac{2}{(1+v)} - 2 \log(1+v)\right) \Big|_{v=0}^{1} \\
&\quad = \frac{1}{3} \left(1 - \log(2)\right).
\end{aligned}
$$

Again, we get the result claimed in the lemma. □

Remark 5.11. The last remaining integral in (5.2) can be solved numerically, Mathematica even gives analytic solutions for $\theta = \frac{1}{n}$, $n \in \mathbb{N}$.

Ali–Mikhail–Haq copula

Another famous family of Archimedean copulas is the Ali–Mikhail–Haq family, introduced in Ali et al. (1978). With parameter $\theta \in [-1, 1)$ its strict generator is

$$
\varphi^{\mathrm{AMH},\theta}(t) = \log\left(\frac{1 - \theta\,(1-t)}{t}\right), \quad t \in (0, 1]\,.
$$

Definition 5.12. An *Ali–Mikhail–Haq copula*, short AMH-copula, with parameter $\theta \in [-1, 1)$ is a function

$$
C^{\mathrm{AMH},\theta}(u,v) = \frac{u\, v}{1 - \theta\,(1-u)\,(1-v)}\,, \quad u, v \in [0,1]\,.
$$

Remark 5.13. For $\theta = 0$ we get the product copula. The restriction $\theta < 1$ is needed to let $\varphi^{\mathrm{AMH},\theta}$ be a generator. Having a non-negative density requires $\theta \geq -1$.

The AMH-copula leads to the following value of Kendall's tau:

$$\tau^{\mathrm{AMH},\theta} = \begin{cases} \frac{3\theta-2}{3\theta} - \frac{2\,(1-\theta)^2}{3\theta^2}\,\log(1-\theta)\,, & \text{if } \theta \in [-1,1) \setminus \{0\}\,, \\ 0\,, & \text{if } \theta = 0\,. \end{cases}$$

Also the asymptotic variance of the tau-estimator can be calculated analytically, but the calculations are lengthy. Nonetheless e.g. Mathematica gives the following solution for every $\theta \in [-1,1) \setminus \{0\}$:

$$\begin{aligned} \left(\sigma_\tau^{\mathrm{AMH},\theta}\right)^2 = &-\frac{16\,\big(7+13\,\theta\,(\theta+3)\big)}{9\,\theta^2} + \frac{32\,(\theta-1)\,\big(4+\theta\,(3\,\theta+11)\big)}{9\,\theta^3}\,\log(1-\theta) \\ &- \frac{16\,(\theta-1)^4}{9\,\theta^4}\,\log^2(1-\theta) + \frac{32\,(\theta+1)}{\theta^2}\,\mathrm{Li}_2(\theta)\,, \end{aligned}$$

where Li_2 denotes the dilogarithm (see Section 7.4.1). As we get the product copula for $\theta = 0$ we know that $\left(\sigma_\tau^{\mathrm{AMH},0}\right)^2 = \frac{4}{9}$.

5.1.3 Farlie–Gumbel–Morgenstern copula

The idea that led to the Farlie–Gumbel–Morgenstern copula was to construct a copula with a mathematically easy representation. This explains why it is possible to calculate the asymptotic variance for all possible values of the parameter in a very simple way. The definition of the copula is the following:

Definition 5.14. A *Farlie–Gumbel–Morgenstern copula*, short FGM-copula, with parameter $\theta \in [-1,1]$ is a function

$$C^{\mathrm{FGM},\theta}(u,v) = u\,v + \theta\,u\,v\,(1-u)\,(1-v), \quad u,v \in [0,1]\,.$$

This form of the copula leads to the drawback that it is impossible to model complete dependence. The interval of possible values for Kendall's tau is small, more precisely we will see that $\tau^{\mathrm{FGM}} \in [-\frac{2}{9}, \frac{2}{9}]$. Nonetheless, this copula has an easy form and it allows analytical computations, among others of the asymptotic variance.

Lemma 5.15. *For two continuous random variables linked by a FGM-copula $C^{\mathrm{FGM},\theta}$ with parameter $\theta \in [-1,1]$, Kendall's tau has the value*

$$\tau^{\mathrm{FGM},\theta} = \frac{2\theta}{9}$$

and the asymptotic variance of the tau-estimator, as defined in (2.15), *equals*

$$\left(\sigma_\tau^{\mathrm{FGM},\theta}\right)^2 = \frac{4}{9} - \frac{184\,\theta^2}{2025}\,.$$

Proof. Knowing that a density of the copula is

$$c^{\mathrm{FGM},\theta}(u,v) = 1 + \theta\,(1-2u)\,(1-2v), \quad u,v \in [0,1]\,,$$

the calculation of the integrals in formula (4.2) for Kendall's tau and in formula (4.5) for the asymptotic variance gets lengthy but not complicated and gives the results claimed in the lemma. □

5.1.4 Marshall–Olkin copula

Marshall and Olkin (1967) want to give a meaningful derivation of a multivariate exponential distribution. One of the explanations works with a system of two components where shocks occur that affect either one or both components. These "fatal shocks" are modelled by three independent Poisson processes with parameters λ_1, λ_2 and λ_{12}, such that the survival probability becomes

$$\mathbb{P}[\,X > s, Y > t\,] = \exp\bigl(-\lambda_1 s - \lambda_2 t - \lambda_{12}\max(s,t)\bigr)\,, \quad s,t > 0\,,$$

which they call bivariate exponential, $\mathrm{BVE}(\lambda_1, \lambda_2, \lambda_{12})$. The copula that belongs to this survival probability is called Marshall–Olkin copula and can be derived by changing the parameters to $\alpha = \frac{\lambda_{12}}{\lambda_1+\lambda_{12}}$ and $\beta = \frac{\lambda_{12}}{\lambda_2+\lambda_{12}}$ (for details see e.g. Embrechts et al., 2003).

Definition 5.16. A *Marshall–Olkin copula* with parameters $\alpha, \beta \in (0,1)$ is a function

$$\begin{aligned} C^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \min\bigl(u^{1-\alpha}\,v, u\,v^{1-\beta}\bigr) \\ &= \begin{cases} u^{1-\alpha}\,v\,, & \text{if } u^\alpha \ge v^\beta\,, \\ u\,v^{1-\beta}\,, & \text{if } u^\alpha < v^\beta\,, \end{cases} \quad u,v \in [0,1]\,. \end{aligned} \tag{5.3}$$

Remark 5.17. In the symmetric case $\alpha = \beta$ the copulas can be seen as a weighted geometric mean of the product copula and the upper Fréchet–Hoeffding bound:

$$C^{\mathrm{MO}}_{\alpha} = \bigl(\min(u,v)\bigr)^\alpha\,\bigl(u\,v\bigr)^{1-\alpha}\,, \quad 0 \le \alpha \le 1\,.$$

Copulas of this form are called *Cuadras–Augé copulas* and were first introduced by Cuadras and Augé (1981).

The Marshall–Olkin copula is absolutely continuous on

$$D_{\mathrm{MO}} := \{(u,v) \in [0,1]^2 \mid u^\alpha \ne v^\beta\}\,,$$

so on this area it has a density

$$c^{\mathrm{MO}}_{\alpha,\beta}(u,v) = \frac{\partial^2}{\partial u\,\partial v}\,C^{\mathrm{MO}}_{\alpha,\beta}(u,v) = \begin{cases} (1-\alpha)\,u^{-\alpha}\,, & \text{if } u^\alpha > v^\beta\,, \\ (1-\beta)\,v^{-\beta}\,, & \text{if } u^\alpha < v^\beta\,, \end{cases} \quad (u,v) \in D_{\mathrm{MO}}\,. \tag{5.4}$$

Integration on this area is no problem. But the integral over the density does not equal one, so the curve $u^\alpha = v^\beta$ has mass greater than zero:

$$\mathbb{P}[\, U^\alpha = V^\beta \,] = \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} \,. \tag{5.5}$$

Lemma 5.18. *For two continuous random variables linked by a Marshall–Olkin copula $C^{\mathrm{MO}}_{\alpha,\beta}$ with parameters $\alpha, \beta \in (0,1)$, Kendall's tau has the value*

$$\tau^{\mathrm{MO}}_{\alpha,\beta} = \frac{\alpha\beta}{\alpha+\beta-\alpha\beta}$$

and the asymptotic variance of the tau-estimator, as defined in (2.15)*, equals*

$$\begin{aligned}\left(\sigma_\tau^{\mathrm{MO},\alpha,\beta}\right)^2 = {} & \frac{64\,(\alpha+\beta+\alpha\beta)}{9\,(\alpha+\beta-\alpha\beta)} - \frac{32\,(2\alpha+3\beta+\alpha\beta)}{3\,(2\alpha+3\beta-2\alpha\beta)} - \frac{32\,(3\alpha+2\beta+\alpha\beta)}{3\,(3\alpha+2\beta-2\alpha\beta)} \\ & + \frac{16\,(\alpha+\beta)}{(2\alpha+2\beta-\alpha\beta)} + \frac{20}{3} + \frac{8\,\alpha\beta}{\alpha+\beta-\alpha\beta} - \frac{4\,\alpha^2\beta^2}{(\alpha+\beta-\alpha\beta)^2} \,.\end{aligned}$$

Proof. To determine the value of Kendall's tau we want to use formula (4.2) and have to calculate

$$\int_{[0,1]^2} C^{\mathrm{MO}}_{\alpha,\beta}(u,v)\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) \,.$$

We first integrate over D_{MO} where we can use the density and get

$$\begin{aligned}&\int_{D_{\mathrm{MO}}} C^{\mathrm{MO}}_{\alpha,\beta}(u,v)\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) \\ &= \int_0^1 \int_0^{u^{\frac{\alpha}{\beta}}} u^{1-\alpha}\, v\,(1-\alpha)\, u^{-\alpha}\, dv\, du + \int_0^1 \int_0^{v^{\frac{\beta}{\alpha}}} u\, v^{1-\beta}\,(1-\beta)\, v^{-\beta}\, du\, dv \\ &= \frac{1}{2}\,(1-\alpha) \int_0^1 u^{1-2\alpha}\, u^{2\frac{\alpha}{\beta}}\, du + \frac{1}{2}\,(1-\beta) \int_0^1 v^{1-2\beta}\, v^{2\frac{\beta}{\alpha}}\, dv \\ &= \frac{\alpha+\beta-2\alpha\beta}{4\,(\alpha+\beta-\alpha\beta)} \,.\end{aligned}$$

The curve $u^\alpha = v^\beta$ can be parametrized by a single variable. We get

$$\begin{aligned}\int_{[0,1]^2 \setminus D_{\mathrm{MO}}} C^{\mathrm{MO}}_{\alpha,\beta}(u,v)\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} \int_0^1 C^{\mathrm{MO}}_{\alpha,\beta}(u, u^{\frac{\alpha}{\beta}})\, dC^{\mathrm{MO}}_{\alpha,\beta}(u, u^{\frac{\alpha}{\beta}}) \\ &= \frac{\alpha\beta}{\alpha+\beta-\alpha\beta} \left(1-\alpha+\frac{\alpha}{\beta}\right) \int_0^1 u^{1-2\alpha+2\frac{\alpha}{\beta}}\, du \\ &= \frac{\alpha\beta}{2\,(\alpha+\beta-\alpha\beta)} \,.\end{aligned}$$

Adding those two integrals gives the result for Kendall's tau.

To calculate the asymptotic variance of the tau-estimator we use formula (4.4) as the copula is not symmetric. The integrals in this formula can be solved analogously to the integrals for Kendall's tau. We get the following results:

$$\begin{aligned}\int_{[0,1]^2} C^{\mathrm{MO}}_{\alpha,\beta}(u,v)^2\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \frac{\alpha+\beta-2\alpha\beta}{9\,(\alpha+\beta-\alpha\beta)} + \frac{\alpha\beta}{3\,(\alpha+\beta-\alpha\beta)} \\ &= \frac{\alpha+\beta+\alpha\beta}{9\,(\alpha+\beta-\alpha\beta)}\end{aligned}$$

$$\begin{aligned}\int_{[0,1]^2} u\, C^{\mathrm{MO}}_{\alpha,\beta}(u,v)\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \frac{2\alpha+3\beta-5\alpha\beta}{6\,(2\alpha+3\beta-2\alpha\beta)} + \frac{\alpha\beta}{2\alpha+3\beta-2\alpha\beta} \\ &= \frac{2\alpha+3\beta+\alpha\beta}{6\,(2\alpha+3\beta-2\alpha\beta)}\end{aligned}$$

$$\begin{aligned}\int_{[0,1]^2} v\, C^{\mathrm{MO}}_{\alpha,\beta}(u,v)\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \frac{3\alpha+2\beta-5\alpha\beta}{6\,(3\alpha+2\beta-2\alpha\beta)} + \frac{\alpha\beta}{3\alpha+2\beta-2\alpha\beta} \\ &= \frac{3\alpha+2\beta+\alpha\beta}{6\,(3\alpha+2\beta-2\alpha\beta)}\end{aligned}$$

$$\begin{aligned}\int_{[0,1]^2} u\, v\, dC^{\mathrm{MO}}_{\alpha,\beta}(u,v) &= \frac{\alpha+\beta-2\alpha\beta}{2\,(2\alpha+2\beta-\alpha\beta)} + \frac{\alpha\beta}{2\alpha+2\beta-\alpha\beta} \\ &= \frac{\alpha+\beta}{2\,(2\alpha+2\beta-\alpha\beta)}\end{aligned}$$

Inserting these integrals in formula (4.4) gives the result. □

5.2 An example for a spherical decomposable density

In Section 4.3 we looked at axially symmetric densities of the form

$$f(x,y) = \sum_{i\in\mathcal{I}} c_i\, g_i(x)\, h_i(y), \quad x,y\in\mathbb{R}\,,$$

and derived the formula

$$\sigma_\tau^2 = 64 \sum_{i,j,l\in\mathcal{I}} c_i\, c_j\, c_l \left(\int_{\mathbb{R}} G_i(x)\, G_j(x)\, g_l(x)\, dx \int_{\mathbb{R}} H_i(y)\, H_j(y)\, h_l(y)\, dy \right).$$

As an example for a spherical density that can be decomposed in this way we now take the product of a polynomial in (x^2+y^2) and $e^{-c(x^2+y^2)}$,

$$f(x,y) = \Big(a_n(x^2+y^2)^n + a_{n-1}(x^2+y^2)^{n-1} + \ldots + a_1(x^2+y^2) + a_0 \Big)\, e^{-c(x^2+y^2)}, \quad x,y\in\mathbb{R}\,, \tag{5.6}$$

with $c > 0$ and $a_0 \geq 0$, $a_1, a_2 \ldots, a_{n-1} \in \mathbb{R}$ and $a_n > 0$ such that $f(x,y) \geq 0$ for all $x, y \in \mathbb{R}$ and $\int_{\mathbb{R}^2} f(x,y)\, d(x,y) = 1$ (see Remark 5.21 if the density is not normalized).

To calculate the asymptotic variance we use the following notations ($i = 0, \ldots, n$ and $j = 0, \ldots, i$):

$$c_{ij} := \binom{i}{j} \frac{a_i}{(2c)^{i+1}}\,, \tag{5.7}$$

$$p_i := \frac{(2i)!}{2^i\, i!} \quad \text{and} \quad s_i := \sum_{u=0}^{i} \frac{p_u}{3^u\, u!} = \sum_{u=0}^{i} \frac{(2u)!}{6^u\,(u!)^2}\,. \tag{5.8}$$

We further denote ($i, k, p = 0, \ldots, n$)

$$\begin{aligned} b\big(i,k,p\big) = p_i\ p_k \Bigg(& \frac{2}{3^{i+k+p+1}} \sum_{j=1}^{i} \sum_{l=1}^{k} \Big(3^{j+l}\ \frac{p_{i+k+p-j-l+1}}{p_{i-j+1}\ p_{k-l+1}} \Big) \\ & - \sum_{j=1}^{i} \Big((p+i-j)!\, \frac{s_{p+i-j}}{p_{i-j+1}} \Big) - \sum_{l=1}^{k} \Big((p+k-l)!\, \frac{s_{p+k-l}}{p_{k-l+1}} \Big) \\ & + 2\ p_p \sum_{q=1}^{p} \Big((p-q)!\, \frac{s_{p-q}}{p_{p-q+1}} \Big) + \frac{\pi}{\sqrt{3}}\ p_p \Bigg)\,, \end{aligned} \tag{5.9}$$

which is symmetric in i and k.

Lemma 5.19. *For two random variables with a common density of the form* (5.6) *the asymptotic variance of the tau-estimator, as defined in* (2.15)*, equals*

$$\sigma_\tau^2 = \frac{32\,\pi}{3} \sum_{i,k,p=0}^{n} \sum_{j=0}^{i} \sum_{l=0}^{k} \sum_{q=0}^{p} c_{ij}\, c_{kl}\, c_{pq}\, b\big(j,l,q\big)\, b\big(i-j, k-l, p-q\big)\,, \tag{5.10}$$

with the notations as in (5.7)–(5.9).

Proof. We can write the density as

$$f(x,y) = \sum_{i=0}^{n} a_i \sum_{j=0}^{i} \binom{i}{j} \Big(x^{2j}\, e^{-cx^2} \Big) \Big(y^{2(i-j)}\, e^{-cy^2} \Big)\,, \quad x, y \in \mathbb{R}\,.$$

The asymptotic variance of the tau-estimator for a continuous distribution only depends on the copula and not on the marginal distributions, see Theorem 4.3. Therefore it is invariant under strictly monotone transformations of the marginals. We want to apply here the following transformation:

$$x \mapsto \frac{x}{\sqrt{2c}} \quad \text{and} \quad y \mapsto \frac{y}{\sqrt{2c}}\,.$$

After this transformation we get the following density:

$$\begin{aligned}
\tilde{f}(x,y) &= \frac{1}{2c} \sum_{i=0}^{n} \frac{a_i}{(2c)^i} \sum_{j=0}^{i} \binom{i}{j} \left(x^{2j} \, e^{-\frac{1}{2}x^2} \right) \left(y^{2(i-j)} \, e^{-\frac{1}{2}y^2} \right) \\
&= \sum_{(i,j) \in \mathcal{I}} c_{ij} \, g_j(x) \, g_{i-j}(y)
\end{aligned}$$

with $g_i(x) = x^{2i} \, e^{-\frac{1}{2}x^2}$, c_{ij} as defined in (5.7) and $\mathcal{I} = \{(i,j) \in \{0,1,\dots,n\}^2 \mid i \geq j\}$. With this representation we can apply formula (4.11), which shows that the asymptotic variance can be computed by $(G_i(z) := \int_0^z g_i(x) \, dx)$

$$\begin{aligned}
\sigma_\tau^2 &= 256 \int_0^\infty \int_0^\infty \left(\int_0^y \int_0^x \tilde{f}(u,v) \, du \, dv \right)^2 \tilde{f}(x,y) \, dx \, dy \\
&= 256 \sum_{(i,j) \in \mathcal{I}} \sum_{(k,l) \in \mathcal{I}} \sum_{(p,q) \in \mathcal{I}} c_{ij} \, c_{kl} \, c_{pq} \\
&\qquad \times \left(\int_0^\infty G_j(x) \, G_l(x) \, g_q(x) \, dx \int_0^\infty G_{i-j}(y) \, G_{k-l}(y) \, g_{p-q}(y) \, dy \right).
\end{aligned}$$

So we have to show that

$$\int_0^\infty G_i(z) \, G_k(z) \, g_p(z) \, dz = \frac{\sqrt{6}\,\pi}{12} \; b\big(i,k,p\big) \tag{5.11}$$

with $b\big(i,k,p\big)$ as defined in (5.9).
Most of the integrals in the following are solved by recursion. For ease of notation we want to introduce

$$p_{i,j} := \prod_{h=0}^{j-1} \big(2i - 2h - 1\big) \, .$$

Note that for $i = j$ we get p_i as defined in (5.8). We further have

$$p_{i,j} = \frac{p_i}{p_{i-j}} \, .$$

To show equation (5.11) we start by integrating the functions g_i using the following recursion:

$$\begin{aligned}
G_i(z) &= \int_0^z g_i(x) \, dx = \int_0^z x^{2i} \, e^{-\frac{1}{2}x^2} \, dx = - \int_0^z x^{2i-1} \left(e^{-\frac{1}{2}x^2} \right)' dx \\
&= -z^{2i-1} \, e^{-\frac{1}{2}z^2} + (2i-1) \int_0^z x^{2i-2} \, e^{-\frac{1}{2}x^2} \, dx \\
&= \sum_{k=1}^{i} \Big(-\, p_{i,k-1} \; z^{2i-2k+1} \, e^{-\frac{1}{2}z^2} \Big) + p_i \int_0^z e^{-\frac{1}{2}x^2} \, dx \\
&= -z^{2i+1} \, e^{-\frac{1}{2}z^2} \sum_{k=1}^{i} \Big(p_{i,k-1} \; z^{-2k} \Big) + p_i \, G_0(z) \, .
\end{aligned}$$

Inserting this last expression for G_i and G_k and expanding the sum gives

$$\begin{aligned}\int_0^\infty G_i(z)\,G_k(z)\,g_p(z)\,dz &= \sum_{j=1}^{i}\sum_{l=1}^{k} p_{i,j-1}\;p_{k,l-1}\int_0^\infty z^{2(i+k+p-j-l+1)}\,e^{-\frac{3}{2}z^2}\,dz\\ &\quad-\sum_{j=1}^{i} p_{i,j-1}\;p_k\int_0^\infty z^{2(i+p-j)+1}\,e^{-z^2}\,G_0(z)\,dz\\ &\quad-\sum_{l=1}^{k} p_i\;p_{k,l-1}\int_0^\infty z^{2(k+p-l)+1}\,e^{-z^2}\,G_0(z)\,dz\\ &\quad+p_i\;p_k\int_0^\infty z^{2p}\,e^{-\frac{1}{2}z^2}\,G_0^2(z)\,dz\,.\end{aligned}\tag{5.12}$$

To simplify the integrals we will need the following solution several times ($\mu \in \mathbb{N}$)

$$\begin{aligned}\int_0^\infty z^{2\mu}\,e^{-\frac{3}{2}z^2}\,dz &= -\frac{1}{3}\int_0^\infty z^{2\mu-1}\left(e^{-\frac{3}{2}z^2}\right)'\,dz\\ &= -\frac{1}{3}\left(z^{2\mu-1}\,e^{-\frac{3}{2}z^2}\Big|_{z=0}^{\infty}-(2\mu-1)\int_0^\infty z^{2\mu-2}\,e^{-\frac{3}{2}z^2}\,dz\right)\\ &= \frac{2\mu-1}{3}\int_0^\infty z^{2\mu-2}\,e^{-\frac{3}{2}z^2}\,dz = \frac{1}{3^\mu}\;p_\mu\int_0^\infty e^{-\frac{3}{2}z^2}\,dz\\ &= \frac{1}{2\cdot 3^\mu}\sqrt{\frac{2\pi}{3}}\;p_\mu\;.\end{aligned}\tag{5.13}$$

This result can also be derived by knowing the moments of a normal distribution with variance $\frac{1}{3}$. Now we show the solutions for the summands in (5.12):

- The first summand in (5.12) is exactly integral (5.13) with $\mu = i+k+p-j-l+1$, so we have

$$\begin{aligned}&p_{i,j-1}\;p_{k,l-1}\int_0^\infty z^{2(i+k+p-j-l+1)}\,e^{-\frac{3}{2}z^2}\,dz\\ &\quad= p_{i,j-1}\;p_{k,l-1}\;\frac{1}{2\cdot 3^{p+i+k-j-l+1}}\sqrt{\frac{2\pi}{3}}\;p_{p+i+k-j-l+1}\\ &\quad= \frac{\sqrt{6\,\pi}\cdot 3^{j+l}}{2\cdot 3^{i+k+p+2}}\;p_{i,j-1}\;p_{k,l-1}\;p_{p+i+k-j-l+1}\;.\end{aligned}$$

- The integral for the two summands with negative sign can be solved recursively, also using solution (5.13) ($\mu \in \mathbb{N}$)

$$
\begin{aligned}
\int_0^\infty z^{2\mu+1}\, e^{-z^2}\, G_0(z)\, dz &= -\frac{1}{2}\int_0^\infty z^{2\mu}\, \big(e^{-z^2}\big)'\, G_0(z)\, dz \\
&= \frac{1}{2}\int_0^\infty z^{2\mu}\, e^{-\frac{3}{2}z^2}\, dz + \mu \int_0^\infty z^{2\mu-1}\, e^{-z^2}\, G_0(z)\, dz \\
&= \frac{1}{4\cdot 3^\mu}\sqrt{\frac{2\pi}{3}}\; p_\mu + \mu \int_0^\infty z^{2\mu-1}\, e^{-z^2}\, G_0(z)\, dz \\
&= \sum_{s=1}^{\mu}\Big(\frac{1}{4\cdot 3^s}\sqrt{\frac{2\pi}{3}}\; p_s\; \frac{\mu!}{s!}\Big) + \frac{\mu!}{4}\sqrt{\frac{2\pi}{3}} \\
&= \sum_{s=1}^{\mu}\Big(\frac{\sqrt{6\pi}}{4\cdot 3^{s+1}}\,\frac{\mu!}{s!}\; p_s\Big) + \frac{\mu!\,\sqrt{6\pi}}{12} = \frac{\mu!\,\sqrt{6\pi}}{12}\left(\sum_{s=0}^{\mu}\frac{p_s}{3^s\, s!}\right).
\end{aligned}
$$

It follows that

$$
p_{i,j-1}\; p_k \int_0^\infty z^{2(i+p-j)+1}\, e^{-z^2}\, G_0(z)\, dz = \frac{\sqrt{6\pi}\,(p+i-j)!}{12}\; p_{i,j-1}\; p_k \left(\sum_{s=0}^{p+i-j}\frac{p_s}{3^s\, s!}\right)
$$

and

$$
p_i\; p_{k,l-1} \int_0^\infty z^{2(k+p-l)+1}\, e^{-z^2}\, G_0(z)\, dz = \frac{\sqrt{6\pi}\,(p+k-l)!}{12}\; p_i\; p_{k,l-1} \left(\sum_{s=0}^{p+k-l}\frac{p_s}{3^s\, s!}\right).
$$

- The calculation of the integral within the fourth summand of (5.12) is a little bit more complicated. To get the recursion we first change the order of integration:

$$
\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz = \int_0^\infty e^{-\frac{1}{2}t^2}\int_0^\infty e^{-\frac{1}{2}s^2}\int_{\max(s,t)}^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, dz\, ds\, dt\,.
$$

The inner integral can now be calculated and equals

$$
\begin{aligned}
\int_{\max(s,t)}^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, dz &= -\int_{\max(s,t)}^\infty z^{2p-1}\, \big(e^{-\frac{1}{2}z^2}\big)'\, dz \\
&= \max(s,t)^{2p-1}\, e^{-\frac{1}{2}\max(s,t)^2} + (2p-1)\int_{\max(s,t)}^\infty z^{2p-2}\, e^{-\frac{1}{2}z^2}\, dz\,.
\end{aligned}
$$

By changing the order of integration back we get

$$
\begin{aligned}
&\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz \\
&= \int_0^\infty\int_0^\infty \max(s,t)^{2p-1}\, e^{-\frac{1}{2}(\max(s,t)^2+s^2+t^2)}\, ds\, dt \\
&\quad + (2p-1)\int_0^\infty z^{2p-2}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz \\
&= 2\int_0^\infty e^{-\frac{1}{2}s^2}\int_s^\infty t^{2p-1}\, e^{-t^2}\, dt\, ds + (2p-1)\int_0^\infty z^{2p-2}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz\,.
\end{aligned}
$$

The integration over t can again be calculated recursively by

$$\begin{aligned}\int_s^\infty t^{2p-1}\, e^{-t^2}\, dt &= -\frac{1}{2}\int_s^\infty t^{2p-2}\left(e^{-t^2}\right)' dt \\ &= \frac{1}{2}\, s^{2p-2}\, e^{-s^2} + \frac{2p-2}{2}\int_s^\infty t^{2p-3}\, e^{-t^2}\, dt \\ &= \sum_{u=1}^{p}\left(\frac{(p-1)!}{2\,(p-u)!}\, s^{2(p-u)}\, e^{-s^2}\right).\end{aligned}$$

The first step of the recursion is therefore

$$\begin{aligned}&\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz \\ &\quad= \sum_{u=1}^{p}\left(\frac{(p-1)!}{(p-u)!}\,\frac{\sqrt{6\pi}}{2\cdot 3^{p-u+1}}\; p_{p-u}\right) + (2p-1)\int_0^\infty z^{2p-2}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz\,.\end{aligned}$$

The whole recursion is then

$$\begin{aligned}&\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz \\ &\quad= \sum_{v=1}^{p}\left(\sum_{u=1}^{p-v+1}\left(\frac{(p-v)!}{(p-u-v+1)!}\,\frac{\sqrt{6\pi}}{2\cdot 3^{p-u-v+2}}\; p_{p-u-v+1}\right)\, p_{p,v-1}\right) + \frac{\sqrt{2\,\pi^3}}{6}\; p_p \\ &\quad= \sum_{v=1}^{p}\left((p-v)!\,\sqrt{\frac{\pi}{6}}\; p_{p,v-1}\sum_{u=1}^{p-v+1}\left(\frac{1}{(p-u-v+1)!}\,\frac{1}{3^{p-u-v+1}}\; p_{p-u-v+1}\right)\right) \\ &\qquad+ \frac{\sqrt{2\,\pi^3}}{6}\; p_p\;.\end{aligned}$$

We can further simplify the sum over u by substituting $z := p-u-v+1$:

$$\sum_{u=1}^{p-v+1}\frac{p_{p-u-v+1}}{3^{p-u-v+1}\,(p-u-v+1)!} = \sum_{z=0}^{p-v}\frac{p_z}{3^z\, z!}\,.$$

Therefore we have

$$\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz = \sum_{v=1}^{p}\left((p-v)!\,\sqrt{\frac{\pi}{6}}\; p_{p,v-1}\left(\sum_{u=0}^{p-v}\frac{p_u}{3^u\, u!}\right)\right) + \frac{\sqrt{2\,\pi^3}}{6}\; p_p\;.$$

Finally we get

$$\begin{aligned}&p_i\; p_k\int_0^\infty z^{2p}\, e^{-\frac{1}{2}z^2}\, G_0^2(z)\, dz \\ &\quad= \sum_{v=1}^{p}\left((p-v)!\,\sqrt{\frac{\pi}{6}}\; p_i\; p_k\; p_{p,v-1}\left(\sum_{u=0}^{p-v}\frac{p_u}{3^u\, u!}\right)\right) + \frac{\sqrt{2\,\pi^3}}{6}\; p_i\; p_k\; p_p\;.\end{aligned}$$

Putting everything together gives the result:

$$\begin{aligned}
\int_0^\infty G_i(z)\,G_k(z)\,g_p(z)\,dz &= \frac{\sqrt{6}\,\pi}{2\cdot 3^{i+k+p+2}} \sum_{j=1}^{i}\sum_{l=1}^{k}\Big(3^{j+l}\ p_{i,j-1}\ p_{k,l-1}\ p_{i+k+p-j-l+1}\Big) \\
&\quad - \frac{\sqrt{6\pi}}{12}\ p_k \sum_{j=1}^{i}\left((p+i-j)!\ p_{i,j-1}\left(\sum_{s=0}^{p+i-j}\frac{p_s}{3^s\, s!}\right)\right) \\
&\quad - \frac{\sqrt{6\pi}}{12}\ p_i \sum_{l=1}^{k}\left((p+k-l)!\ p_{k,l-1}\left(\sum_{s=0}^{p+k-l}\frac{p_s}{3^s\, s!}\right)\right) \\
&\quad + \frac{\sqrt{6\pi}}{6}\ p_i\ p_k \sum_{v=1}^{p}\left((p-v)!\ p_{p,v-1}\left(\sum_{u=0}^{p-v}\frac{p_u}{3^u\, u!}\right)\right) \\
&\quad + \frac{\pi\,\sqrt{2\pi}}{12}\ p_i\ p_k\ p_p \\
&= \frac{\sqrt{6}\,\pi}{12}\, b\big(i,k,p\big)\,.
\end{aligned}$$

□

Remark 5.20. Note that the coefficients a_i, $i = 0,\dots,n$, and c only enter in c_{ij}. The values of $b(i,k,p)$ are the same for every density of the form (5.6). The degree n of the polynomial then tells how many of the $b(i,k,p)$ are needed.

Remark 5.21. If the coefficients a_i, $i = 0,\dots,n$, in (5.6) are not chosen such that $\int_{\mathbb{R}^2} f(x,y)\,d(x,y) = 1$, then we have to normalize the density by dividing by

$$\begin{aligned}
C_f := \int_{\mathbb{R}^2} f(x,y)\,d(x,y) &= \sum_{i=0}^{n} a_i \sum_{j=0}^{i}\binom{i}{j}\left(\int_{\mathbb{R}} x^{2j}\,e^{-cx^2}\,dx\right)\left(\int_{\mathbb{R}} y^{2(i-j)}\,e^{-cy^2}\,dy\right) \\
&= \sum_{i=0}^{n} a_i \sum_{j=0}^{i}\binom{i}{j}\frac{\pi}{c\,(2c)^i}\ p_j\ p_{i-j}\,,
\end{aligned}$$

since we know

$$\begin{aligned}
\int_{\mathbb{R}} x^{2i}\,e^{-cx^2}\,dx &= -\frac{1}{c}\int_0^\infty x^{2i-1}\left(e^{-cx^2}\right)'dx = \frac{(2i-1)}{c}\int_0^\infty x^{2i-2}\,e^{-cx^2}\,dx \\
&= \frac{2\,p_i}{(2c)^i}\int_0^\infty e^{-cx^2}\,dx = \frac{p_i}{(2c)^i}\,\frac{\sqrt{\pi}}{\sqrt{c}}\,.
\end{aligned}$$

We can further simplify

$$\begin{aligned}
C_f &= \sum_{i=0}^{n} a_i \sum_{j=0}^{i}\binom{i}{j}\frac{\pi}{c\,(2c)^i}\ p_j\ p_{i-j} = \frac{\pi}{c}\sum_{i=0}^{n}\frac{a_i}{(4c)^i}\sum_{j=0}^{i}\binom{i}{j}\frac{(2j)!\,(2i-2j)!}{j!\,(i-j)!} \\
&= \frac{\pi}{c}\sum_{i=0}^{n}\frac{i!\,a_i}{(4c)^i}\sum_{j=0}^{i}\binom{2j}{j}\binom{2\,(i-j)}{i-j} = \frac{\pi}{c}\sum_{i=0}^{n}\frac{i!\,a_i}{c^i}\,,
\end{aligned} \tag{5.14}$$

where we used

$$\sum_{j=0}^{i} \binom{2j}{j} \binom{2\,(i-j)}{i-j} = 4^i \,,$$

which is shown in Lemma 7.1. To derive the formula for the asymptotic variance for the modified density we notice that in order to normalize the density function we just have to divide the vector $(a_0, a_1, \ldots, a_n)$ by the normalizing factor. So this factor enters the formula of the asymptotic variance just at the coefficients c_{ij}. The asymptotic variance has then the form

$$\begin{aligned}
\sigma_\tau^2 &= \frac{32\,\pi}{3} \Big(\frac{\pi}{c} \sum_{i=0}^{n} \frac{i!\, a_i}{c^i} \Big)^{-3} \sum_{i,k,p=0}^{n} \sum_{j=0}^{i} \sum_{l=0}^{k} \sum_{q=0}^{p} c_{ij}\, c_{kl}\, c_{pq}\, b\big(j, l, q\big)\, b\big(i-j, k-l, p-q\big) \\
&= \frac{4}{3\,\pi^2} \Big(\sum_{i=0}^{n} \frac{i!\, a_i}{c^i} \Big)^{-3} \Bigg(\sum_{i,k,p=0}^{n} \frac{a_i\, a_k\, a_p}{(2c)^{i+k+p}} \\
&\qquad \times \sum_{j=0}^{i} \sum_{l=0}^{k} \sum_{q=0}^{p} \binom{i}{j} \binom{k}{l} \binom{p}{q}\, b\big(j, l, q\big)\, b\big(i-j, k-l, p-q\big) \Bigg) \,.
\end{aligned} \tag{5.15}$$

Chapter 6

Comparison of the estimators for elliptical distributions

In the case of elliptical distributions, the two dependence measures Kendall's tau τ and the linear correlation coefficient ϱ are connected by the formula

$$\tau = \frac{2\, a_X}{\pi} \arcsin \varrho \,,$$

simply assuming that both components are non-degenerate (see Theorem 3.18). So we are given two ways of estimating the linear correlation, either using the standard estimator directly or estimating Kendall's tau and transforming it. In Chapter 3 we already showed that both methods provide estimators that are asymptotically normal under certain conditions.

A first remark on the comparison of the two methods concerns the assumptions that are needed to guarantee asymptotic normality. The transformation-estimator only needs non-degeneracy of the components, whereas the standard estimator additionally demands existing fourth moments (see Theorem 3.17), so in the case of elliptical distributions we must have $\mathbb{E}[R^4] < \infty$. This is not always given, e.g. not for the t-distribution with $\nu \leq 4$ degrees of freedom, see Section 6.4.

One further advantage of the transformation-estimator is, that the boundedness of the asymptotic variance of the tau-estimator (see Remark 2.17) implies that also the asymptotic variance of the transformation-estimator is bounded,

$$\sigma^2_{\varrho(\tau)} = \frac{\pi^2}{4}\, \sigma^2_\tau \,(1 - \varrho^2) \leq \pi^2\,(1 - \tau^2)\,(1 - \varrho^2)\,,$$

whereas the asymptotic variance of the standard estimator can have any non-negative value.

In the following we compare the two methods of estimation for several elliptical distributions where we can calculate both asymptotic variances.

6.1 Spherical distributions bounded on a disc

In a first consideration we look at spherical distributions where the densities are zero outside a circular disc with radius $r > 0$. Writing the density as

$$f(x,y) = \tilde{f}(x,y)\, 1_{\{x^2+y^2 \le r^2\}}\,, \quad x,y \in \mathbb{R}\,,$$

formula (4.9) for the asymptotic variance of the tau-estimator becomes

$$\sigma_\tau^2 = 4 \int_{-r}^{r} \int_{-\sqrt{r^2-y^2}}^{\sqrt{r^2-y^2}} \left(4 \int_0^y \int_0^x \tilde{f}(u,v)\, du\, dv \right)^2 \tilde{f}(x,y)\, dx\, dy\,.$$

Here are two examples:

Example 6.1. The simplest case is to have a uniform distribution on the circle, e.g. to have a density

$$f(x,y) = \frac{1}{\pi\, r^2}\, 1_{\{x^2+y^2 \le r^2\}}\,, \quad x,y \in \mathbb{R}\,.$$

In this case both estimators have the same asymptotic variance,

$$\sigma_\varrho^2 = \sigma_{\varrho(\tau)}^2 = \frac{2}{3}\,.$$

Example 6.2. Another possible density is the following:

$$f(x,y) = \frac{2}{\pi\, r^4} \Big(r^2 - (x^2+y^2) \Big)\, 1_{\{x^2+y^2 \le r^2\}}\,, \quad x,y \in \mathbb{R}\,.$$

In this case it is slightly better to use the standard estimator, as

$$\sigma_\varrho^2 = \frac{3}{4}$$

and

$$\sigma_{\varrho(\tau)}^2 = \frac{116}{135} \approx 0.859\,.$$

6.2 Example for a spherical decomposable density revisited

In Section 5.2 we looked at a spherical decomposable density where we could calculate the asymptotic variance of the tau-estimator. To determine also the asymptotic variance of the standard estimator we need a density of its radial vector:

$$f_R(r) = \frac{2\pi}{C_f}\, r \Big(a_n\, r^{2n} + a_{n-1}\, r^{2(n-1)} + \ldots + a_1\, r^2 + a_0 \Big)\, e^{-c\, r^2}, \quad r \ge 0\,,$$

with, as in (5.14),

$$C_f = \frac{\pi}{c} \sum_{i=0}^{n} \frac{i!\, a_i}{c^i}$$

and $c > 0$, $a_0 \geq 0$, $a_1, \dots, a_{n-1} \in \mathbb{R}$ and $a_n > 0$ such that $f_R(r) \geq 0$ for all $r \geq 0$. Using formula (3.13) we get

$$\begin{aligned}
\sigma_\varrho^2 &= \frac{\mathbb{E}[R^4]}{2\,\mathbb{E}[R^2]^2} \left(\varrho^2 - 1\right)^2 \\
&= \frac{1}{4c} \left(\sum_{i=0}^{n} \frac{i!\, a_i}{c^i}\right) \left(\sum_{i=0}^{n} a_i \int_0^\infty r^{5+2i}\, e^{-c\,r^2}\, dr\right) \left(\sum_{i=0}^{n} a_i \int_0^\infty r^{3+2i}\, e^{-c\,r^2}\, dr\right)^{-2} \\
&= \frac{1}{4c} \left(\sum_{i=0}^{n} \frac{i!\, a_i}{c^i}\right) \left(\sum_{i=0}^{n} \frac{a_i}{2\,c^{i+3}} \prod_{j=0}^{i+1} (i-j+2)\right) \left(\sum_{i=0}^{n} \frac{a_i}{2\,c^{i+2}} \prod_{j=0}^{i} (i-j+1)\right)^{-2} \\
&= \frac{1}{2} \left(\sum_{i=0}^{n} \frac{i!\, a_i}{c^i}\right) \left(\sum_{i=0}^{n} \frac{(i+2)!\, a_i}{c^i}\right) \left(\sum_{i=0}^{n} \frac{(i+1)!\, a_i}{c^i}\right)^{-2}.
\end{aligned}$$

We want to compare this variance to the variance of the transformation-estimator, which is (see (5.15))

$$\begin{aligned}
\sigma_{\varrho(\tau)}^2 = \frac{1}{3} \Big(\sum_{i=0}^{n} \frac{i!\, a_i}{c^i}\Big)^{-3} \Bigg(&\sum_{i,k,p=0}^{n} \frac{a_i\, a_k\, a_p}{(2c)^{i+k+p}} \\
&\times \sum_{j=0}^{i} \sum_{l=0}^{k} \sum_{q=0}^{p} \binom{i}{j} \binom{k}{l} \binom{p}{q}\, b\big(j,l,q\big)\, b\big(i-j, k-l, p-q\big) \Bigg).
\end{aligned}$$

The values of the asymptotic variances depend in a very complex way on the choice of the a_i and of c and there is no estimator which is preferable in any case. We want to look at a special case where we choose all a_i to be zero except for a_n, such that the two-dimensional density becomes

$$f(x,y) = \frac{c^{n+1}}{n!\,\pi} \,(x^2+y^2)^n\, e^{-c(x^2+y^2)}, \quad x, y \in \mathbb{R}. \tag{6.1}$$

The asymptotic variances simplify to

$$\begin{aligned}
\sigma_\varrho^2 &= \frac{1}{2} \left(\frac{n!\, a_n}{c^n}\right) \left(\frac{(n+2)!\, a_n}{c^n}\right) \left(\frac{(n+1)!\, a_n}{c^n}\right)^{-2} \\
&= \frac{1}{2}\, \frac{n!\,(n+2)!}{\big((n+1)!\big)^2} \\
&= \frac{(n+2)}{2\,(n+1)}
\end{aligned} \tag{6.2}$$

and

$$\sigma^2_{\varrho(\tau)} = \frac{1}{3}\left(\frac{n!\,a_n}{c^n}\right)^{-3}\left(\frac{a_n^3}{(2c)^{3n}}\sum_{j=0}^{n}\sum_{l=0}^{n}\sum_{q=0}^{n}\binom{n}{j}\binom{n}{l}\binom{n}{q}\,b(j,l,q)\,b\big(n-j,n-l,n-q\big)\right)$$
$$= \frac{1}{3\cdot 2^{3n}\,(n!)^3}\left(\sum_{j=0}^{n}\sum_{l=0}^{n}\sum_{q=0}^{n}\binom{n}{j}\binom{n}{l}\binom{n}{q}\,b(j,l,q)\,b\big(n-j,n-l,n-q\big)\right).$$

So both asymptotic variances only depend on n and no longer on the choice of c. We can calculate the values of the asymptotic variances for every $n \in \mathbb{N}$ and also for the limit $n \to \infty$. From equation (6.2) it can easily be seen that

$$\lim_{n\to\infty} \sigma^2_{\varrho}(n) = \frac{1}{2}\,.$$

To determine the limit for the asymptotic variance of the transformation-estimator we want to use Lemma 2.18. Instead of building the limit of the two-dimensional distribution (6.1) we can equivalently look at the distribution of the radial variable. As we want to show that the limit this distribution is degenerate, it is enough to show this limit property for the distribution of R^2/n. It has a density

$$f_{R^2/n}(r) = \frac{(cn)^{n+1}}{n!}\,r^n\,e^{-cnr}\,,$$

which is a density of the gamma distribution $\mathrm{Ga}(n+1, cn)$. The moments of this gamma distribution are

$$\mathbb{E} = \frac{n+1}{cn} \overset{n\to\infty}{\longrightarrow} \frac{1}{c} \quad \text{and} \quad \mathbb{V}\mathrm{ar} = \frac{n+1}{c^2\,n^2} \overset{n\to\infty}{\longrightarrow} 0\,.$$

So the random variable R^2/n converges in probability to a degenerate random variable and therefore the distribution with density (6.1) converges to a uniform distribution on a circle. We know from Theorem 3.32 that in this case the tau-kernel is degenerate and the asymptotic variance of the tau-estimator and also of the transformation-estimator equals zero.

The values of the variances for different n are:

n	0	1	2	3	4	5	6	7	8	...	∞
σ^2_{ϱ}	1	0.75	0.667	0.625	0.6	0.583	0.571	0.563	0.556	...	0.5
$\sigma^2_{\varrho(\tau)}$	1.097	0.763	0.602	0.502	0.434	0.383	0.344	0.312	0.287	...	0

So only in the case $n = 0$, which equals the normal distribution for $c = \frac{1}{2}$, the direct estimation is clearly better than the estimation using Kendall's tau. For $n = 1$ both estimators have nearly the same behaviour. If n becomes bigger than 1, then the transformation-estimator is clearly better, where the difference between the estimators grows when n becomes bigger. In the limit, the transformation-estimator has asymptotic variance zero, whereas the asymptotic variance of the standard estimator converges to $\frac{1}{2}$.

6.3 Standard normal distribution

When a multi-dimensional random vector is standard normally distributed, its components are independent. So we know from Section 5.1.1 that

$$\sigma_\tau^2 = \frac{4}{9}$$

and therefore

$$\sigma_{\varrho(\tau)}^2 = \frac{\pi^2}{9} \approx 1.0966\,.$$

The standard estimator performs slightly better as we have

$$\sigma_\varrho^2 = 1\,.$$

6.4 Student's t-distribution

6.4.1 Representation of the t-distribution as normal variance mixture distribution

Definition 6.3. A d-dimensional t-distribution with location vector $\boldsymbol{\mu}$, dispersion matrix $\boldsymbol{\Sigma}$, and $\nu > 0$ degrees of freedom is a normal variance mixture distribution, as defined in Definition 3.28, with mixing random variable $W \sim \mathrm{Ig}(\frac{\nu}{2}, \frac{\nu}{2})$.

Notation. We want to denote such a t-distribution by $t_d(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})$.

If the dispersion matrix $\boldsymbol{\Sigma}$ has full rank, then we can determine a density of the t-distribution. Using formula (3.24) we get

$$\begin{aligned} f_\nu(\boldsymbol{x}) &= \int_0^\infty \frac{1}{(2\pi w)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \exp\Big(-\frac{1}{2w}\,(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\Big)\,G(dw) \\ &= \frac{(\frac{\nu}{2})^{\nu/2}}{\Gamma(\frac{\nu}{2})\,(2\pi)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \int_0^\infty w^{-\frac{\nu+d}{2}-1} \exp\Big(-\frac{1}{2w}\,\big(\nu + (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\big)\Big)\,dw \\ &= \frac{\Gamma(\frac{\nu+d}{2})(\frac{\nu}{2})^{\nu/2}\,2^{(\nu+d)/2}}{\Gamma(\frac{\nu}{2})\,(2\pi)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \big(\nu + (\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\big)^{-\frac{\nu+d}{2}} \\ &= \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2})\,(\pi\nu)^{d/2}\,|\boldsymbol{\Sigma}|^{1/2}} \Big(1 + \frac{(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{\nu}\Big)^{-\frac{\nu+d}{2}}, \quad \boldsymbol{x} \in \mathbb{R}^d\,, \end{aligned} \tag{6.3}$$

where the integral is solved by realizing that, up to a constant, the integrand equals a density of an inverse gamma distribution.

Lemma 6.4. *The sequence of t-distributions $\big(t_d(\nu, \boldsymbol{\mu}, \boldsymbol{\Sigma})\big)_{\nu>0}$ converges weakly to a normal distribution $\mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ as $\nu \to \infty$.*

Proof. From the definition we know that a random vector $\boldsymbol{X}$, that follows a t-distribution with ν degrees of freedom, can be written as $\boldsymbol{X} \overset{\mathrm{d}}{=} \sqrt{W_\nu}\,\boldsymbol{Z}$, with $W_\nu \sim \mathrm{Ig}(\frac{\nu}{2}, \frac{\nu}{2})$, $\boldsymbol{Z}$ is normally distributed and W_ν and $\boldsymbol{Z}$ are independent. As for $\nu \to \infty$ the radial variable W_ν converges in probability to a degenerate random variable with value 1 (see Lemma 4.14) we get the result by Slutsky's theorem. □

6.4.2 Asymptotic variance of the tau-estimator

Lemma 6.5. *For a two-dimensional uncorrelated t-distribution with $\nu > 0$ degrees of freedom, the asymptotic variance of the tau-estimator simplifies to*

$$\sigma_\tau^2 = \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty u^{\nu-1} \arctan^2 u \int_0^1 t^{\nu-1}\,(1-t)^{\nu-1}\,(u^2+t)^{-\nu}\,dt\,du\,. \tag{6.4}$$

Proof. Since the uncorrelated t-distribution is a standard normal variance mixture distribution where the mixing distribution function is the inverse gamma distribution with parameters $\alpha = \beta = \frac{\nu}{2}$, we can use formula (4.18) and get for the asymptotic variance

$$\sigma_\tau^2 = \frac{16\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty \int_0^\infty \frac{(xy)^{\nu-1}}{\big(x+y+xy\big)^{\frac{3\nu}{2}}} \arctan^2\Big(\frac{1}{\sqrt{x+y+xy}}\Big)\,dx\,dy\,.$$

To simplify the argument of the arc tangent we use the substitution $u := \big(x+y+xy\big)^{-\frac{1}{2}}$ and get

$$\begin{aligned}\sigma_\tau^2 &= \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty x^{\nu-1}\,(1+x)^{-\nu} \int_0^{\frac{1}{\sqrt{x}}} u^{3\nu-3}\,\Big(\frac{1}{u^2}-x\Big)^{\nu-1} \arctan^2 u\,du\,dx \\ &= \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty u^{3\nu-3} \arctan^2 u \int_0^{\frac{1}{u^2}} x^{\nu-1}\,(1+x^2)^{-\nu}\,\Big(\frac{1}{u^2}-x\Big)^{\nu-1}\,dx\,du\,.\end{aligned}$$

The substitution $t := u^2\,x$ leads to a nicer form of the inner integral:

$$\sigma_\tau^2 = \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty u^{-\nu-1} \arctan^2 u \int_0^1 t^{\nu-1}\,(1-t)^{\nu-1}\,\Big(1+\frac{t}{u^2}\Big)^{-\nu}\,dt\,du\,,$$

which is equivalent to formula (6.4). □

Remark 6.6. The integral over t is, up to a constant, a hypergeometric function. Hypergeometric functions are defined as the series

$${}_pF_q\left[\begin{matrix}a_1,\dots,a_p\\ b_1,\dots,b_q\end{matrix};\,z\right] = \sum_{n=0}^\infty \frac{(a_1)_n\dots(a_p)_n}{(b_1)_n\dots(b_q)_n}\,\frac{z^n}{n!}\,,$$

where $(a)_n = a\,(a+1)\dots(a+n-1)$, $n \in \mathbb{N}$, and $(a)_0 = 1$. We need the case $p = 2$ and $q = 1$ where the function has the following integral representation:

$$ {}_2\mathrm{F}_1\left[\begin{matrix} a, b \\ c \end{matrix}\,;\, z\right] = \frac{\Gamma(c)}{\Gamma(b)\,\Gamma(c-b)} \int_0^1 t^{b-1}\,(1-t)^{c-b-1}\,(1-t\,z)^{-a}\,dt $$

if $|z| < 1$ and $\mathrm{Re}(c) > \mathrm{Re}(b) > 0$. For details about hypergeometric functions see e.g. Bailey (1964) or Gasper and Rahman (1990). So in our case we have

$$ \int_0^1 t^{\nu-1}\,\big(1-t\big)^{\nu-1}\Big(1+\frac{1}{u^2}\,t\Big)^{-\nu}\,dt = \frac{\Gamma^2(\nu)}{\Gamma(2\nu)}\;{}_2\mathrm{F}_1\left[\begin{matrix} \nu, \nu \\ 2\nu \end{matrix}\,;\, -\frac{1}{u^2}\right]. $$

Anyway, this observation did not help for the solution, so we will not use it.

The formula in Lemma 6.5 holds for every $\nu > 0$. For $\nu \in \mathbb{N}$, even the last two integrals can be solved, where the solution depends on whether ν is even or odd.

Theorem 6.7. *For a two-dimensional uncorrelated t-distribution with $\nu \in \mathbb{N}$ degrees of freedom, the asymptotic variance of the tau-estimator has the following representation:*

(i) *If $\nu \in 2\mathbb{N}-1$, then*

$$ \begin{aligned} \sigma_\tau^2 = \frac{16}{\pi^2}\,\log^2(2) + (-1)^{\frac{\nu-1}{2}}\,\frac{32\,\Gamma(\frac{3\nu}{2})}{\pi\,\Gamma^3(\frac{\nu}{2})} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k}\binom{\nu-1}{k}\binom{\nu+k-1}{k} \\ \times \left(\sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} + \log(2) \sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h}\right); \end{aligned} \tag{6.5} $$

(ii) *If $\nu \in 2\mathbb{N}$, then*

$$ \begin{aligned} \sigma_\tau^2 = (-1)^{\frac{\nu}{2}+1}\,\frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k}\binom{\nu-1}{k}\binom{\nu+k-1}{k} \\ \times \left(\frac{\pi^2}{4}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{l} - \frac{\pi^2}{3}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1} - \sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1}\sum_{n=1}^{l-1}\frac{1}{n^2}\right). \end{aligned} \tag{6.6} $$

The proof is quite long and can be found in Chapter 7.

Although the solution looks lengthy, the values for fixed ν are quite tight, like shown in Table 6.1 for $\nu \in \{1, 3, \dots, 11\}$ and in Table 6.2 for $\nu \in \{2, 4, \dots, 12\}$. Note that we listed the asymptotic variance of the transformation-estimator, i.e. $\sigma^2_{\varrho(\tau)} = \frac{\pi^2}{4}\,\sigma_\tau^2$, to simplify the representation (to avoid to have π^2 in the denominator).

ν	$\sigma^2_{\varrho(\tau)}$
1	$4\log^2(2)$
3	$30 - 44\log(2) + 4\log^2(2)$
5	$-\frac{20\,221}{54} + \frac{1\,618}{3}\log(2) + 4\log^2(2)$
7	$\frac{342\,071}{50} - \frac{148\,066}{15}\log(2) + 4\log^2(2)$
9	$-\frac{1\,358\,296\,703}{9\,800} + \frac{20\,995\,691}{105}\log(2) + 4\log^2(2)$
11	$\frac{285\,183\,353\,759}{95\,256} - \frac{1\,360\,557\,907}{315}\log(2) + 4\log^2(2)$

Table 6.1: Asymptotic variance of the transformation-estimator for the uncorrelated t-distribution with ν degrees of freedom where ν is odd.

ν	$\sigma^2_{\varrho(\tau)}$
2	$\frac{8}{3} - \frac{1}{9}\pi^2$
4	$-\frac{1\,000}{27} + \frac{35}{9}\pi^2$
6	$\frac{401\,312}{675} - \frac{541}{9}\pi^2$
8	$-\frac{42\,307\,408}{3675} + \frac{10\,499}{9}\pi^2$
10	$\frac{71\,980\,077\,752}{297\,675} - \frac{220\,501}{9}\pi^2$
12	$-\frac{192\,375\,504\,097\,528}{36\,018\,675} + \frac{4\,870\,403}{9}\pi^2$

Table 6.2: Asymptotic variance of the transformation-estimator for the uncorrelated t-distribution with ν degrees of freedom where ν is even.

6.4.3 Asymptotic variance of the standard estimator

Lemma 6.8. *For a t-distribution with $\nu > 4$ degrees of freedom and with linear correlation coefficient ϱ, the asymptotic variance of the standard estimator equals*

$$\sigma_\varrho^2 = \left(1 + \frac{2}{\nu - 4}\right)\left(1 - \varrho^2\right)^2 . \tag{6.7}$$

Proof. Since the t-distribution is a normal variance mixture distribution, we can use formula (3.25) for the asymptotic variance of the standard estimator. Knowing that the mixing random variable W is inverse gamma distributed with parameters $\alpha = \beta = \frac{\nu}{2}$, we get from Lemma 4.13 that

$$\mathbb{E}[W] = \frac{\nu}{\nu - 2}$$

and

$$\mathbb{E}[W^2] = \frac{\nu^2}{(\nu - 2)(\nu - 4)} ,$$

and therefore

$$\sigma_\varrho^2 = \frac{\mathbb{E}[W^2]}{\mathbb{E}[W]^2}\left(1 - \varrho^2\right)^2 = \frac{(\nu - 2)}{(\nu - 4)}\left(1 - \varrho^2\right)^2 .$$

□

6.4.4 Comparison of both estimators

With the results from the previous two subsections we can calculate the asymptotic variances of the two dependence estimators, the standard estimator and the transformation-estimator, in the case of an uncorrelated t-distribution with $\nu \in \mathbb{N}$ degrees of freedom. The values of the asymptotic variances are listed in Table 6.3 and illustrated in Figure 6.1. Using the asymptotic variance to measure the performance of the estimators leads to a choice that depends heavily on the value of the parameter ν. The results can be summarized as follows:

- For heavy-tailed t-distributions ($\nu \leq 4$) the transformation-estimator is asymptotically normal with finite asymptotic variance, whereas Theorem 3.17 for the standard estimator is not applicable and formula (3.13) does not give a finite variance.

- For $\nu \in \{5, 6, \ldots, 16\}$ the transformation-estimator has a smaller asymptotic variance than the standard estimator and is in this sense better. Especially for small ν the difference is remarkable, like for $\nu = 5$ the asymptotic variance of the standard estimator is more than twice as big as the one of the transformation-estimator.

- The two estimation methods are approximately equivalent for $\nu \approx 17$, where the corresponding distribution is already quite similar to the normal distribution.

- For $\nu \geq 17$ the standard estimator performs better than the transformation-estimator, although the difference between the asymptotic variances is small.
- In the limit $\nu \to \infty$ the asymptotic variance of the standard estimator converges to 1 due to (6.7). Since the uncorrelated t-distributions converge weakly to the standard normal distribution (see Lemma 6.4) we can use Lemma 2.18 to determine the asymptotic variance of the transformation-estimator. So in the limit $\nu \to \infty$, the estimators have the same asymptotic variances as for the standard normal distribution.

ν	1	2	3	4	5	6	7	8	9	10
σ^2_ϱ	n.a.	n.a.	n.a.	n.a.	3	2	1.667	1.500	1.400	1.333
$\sigma^2_{\varrho(\tau)}$	1.922	1.570	1.423	1.345	1.296	1.263	1.240	1.222	1.208	1.197

ν	11	12	13	14	15	16	17	18	...	∞
σ^2_ϱ	1.286	1.250	1.222	1.200	1.182	1.167	1.154	1.143	...	1
$\sigma^2_{\varrho(\tau)}$	1.188	1.180	1.174	1.168	1.164	1.159	1.156	1.152	...	1.097

Table 6.3: Values of the asymptotic variances of the standard estimator and of the transformation-estimator for the uncorrelated t-distribution with ν degrees of freedom.

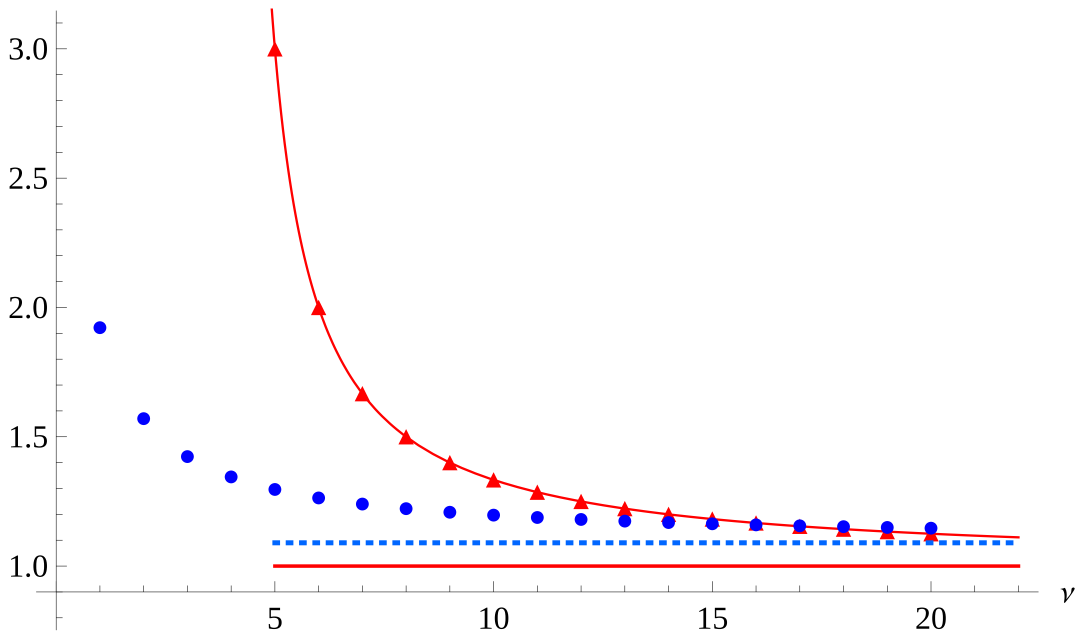

Figure 6.1: Asymptotic variances for the uncorrelated t-distribution, red triangles for the standard estimator, blue dots for the transformation-estimator. As the formula for the standard estimator is valid for every $\nu > 4$ there is further the red curve for the asymptotic variances of the standard estimator. The horizontal lines indicate the asymptotic variances for the limit $\nu \to \infty$, i.e. for the standard normal distribution, red for the standard estimator (with value 1), blue and dotted for the transformation-estimator (with value 1.097).

Chapter 7

Proof of the asymptotic variance for the uncorrelated t-distribution

In this chapter we give the proof of Theorem 6.7, which claims the asymptotic variance of the tau-estimator for the uncorrelated t-distribution. By representing the t-distribution as a standard normal variance mixture distribution we could already simplify the formula to

$$\sigma_\tau^2 = \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})} \int_0^\infty u^{\nu-1} \arctan^2 u \int_0^1 t^{\nu-1}\,(1-t)^{\nu-1}\,(u^2+t)^{-\nu}\,dt\,du$$

(see Lemma 6.5). Some further simplifications are possible for general $\nu \in \mathbb{N}$, but soon we have to distinguish between even and odd ν.

7.1 Reduction of the exponents for general ν

To simplify notation we want to introduce the abbreviation

$$\gamma_\nu := \frac{\Gamma(\frac{3\nu}{2})}{\Gamma^3(\frac{\nu}{2})}$$

as this factor stays unchanged throughout the whole calculations. The first steps are similar for even and odd ν, but slightly different, so we want to introduce δ as

$$\delta = \begin{cases} 1\,, & \text{if } \nu \text{ is even,} \\ 0\,, & \text{if } \nu \text{ is odd.} \end{cases}$$

To simplify notation we define μ as

$$\mu := \frac{\nu - 1 + \delta}{2} = \left\lfloor \frac{\nu}{2} \right\rfloor \in \mathbb{N}\,,$$

such that the asymptotic variance has the form

$$\sigma_\tau^2 = \frac{32\,\gamma_\nu}{\pi^2} \int_0^\infty u^{2\mu-\delta} \arctan^2 u \int_0^1 t^{2\mu-\delta} (1-t)^{2\mu-\delta} (u^2+t)^{-2\mu-1+\delta}\, dt\, du\,.$$

Our goal is now to reduce $u^{2\mu-\delta}$ to $u^{2-\delta}$ and to simplify $(u^2+t)^{-2\mu-1+\delta}$ to $(u^2+t)^{-2}$. By writing

$$u^{2\mu-\delta} = u^{2-\delta}\,(u^2+t-t)^{\mu-1}$$

and dividing by $(u^2+t)^{2\mu+1-\delta}$ (see Section 7.5.1) we come to integrals of the kind

$$\sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \int_0^\infty u^{2-\delta} \arctan^2 u \int_0^1 t^{2\mu+j-\delta} (1-t)^{2\mu-\delta} (u^2+t)^{-\mu-j-2+\delta}\, dt\, du\,.$$

The integral over t can be further simplified by using the rule of Leibniz (see Section 7.5.2), which changes the inner integral for every $j = 0, \ldots, \mu-1$ to

$$\begin{aligned} &\int_0^1 t^{2\mu+j-\delta} (1-t)^{2\mu-\delta} (u^2+t)^{-\mu-j-2+\delta}\, dt \\ &\quad = \sum_{k=0}^{2\mu-\delta} \frac{(-1)^k}{\mu+k} \binom{2\mu-\delta}{k} \binom{2\mu+j+k-\delta}{\mu+j+1-\delta} \int_0^1 t^{\mu+k} (t+u^2)^{-2}\, dt\,. \end{aligned}$$

So the expression for the asymptotic variance becomes

$$\begin{aligned} \sigma_\tau^2 = \frac{32\,\gamma_\nu}{\pi^2} &\sum_{k=0}^{2\mu-\delta} \frac{(-1)^k}{\mu+k} \binom{2\mu-\delta}{k} \sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \binom{2\mu+j+k-\delta}{\mu+j+1-\delta} \\ &\times \int_0^\infty u^{2-\delta} \arctan^2 u \int_0^1 t^{\mu+k} (t+u^2)^{-2}\, dt\, du\,. \end{aligned}$$

We can further simplify the summation over j using Lemma 7.2, which states that

$$\sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \binom{2\mu+j+k-\delta}{\mu+j+1-\delta} = (-1)^{\mu-1} \binom{2\mu+k-\delta}{k}.$$

This leads to the following representation of the asymptotic variance:

$$\begin{aligned} \sigma_\tau^2 = (-1)^{\mu+1} \frac{32\,\gamma_\nu}{\pi^2} &\sum_{k=0}^{2\mu-\delta} \frac{(-1)^k}{\mu+k} \binom{2\mu-\delta}{k} \binom{2\mu+k-\delta}{k} \\ &\times \int_0^\infty u^{2-\delta} \arctan^2 u \int_0^1 t^{\mu+k} (t+u^2)^{-2}\, dt\, du\,. \end{aligned}$$

With this form we can now reduce $\arctan^2 u$ to $\arctan u$. This works differently for even and odd ν so we separate these two cases from now on.

7.2 Solution for even ν

If ν is even, then $\delta = 1$ and $\mu = \frac{\nu}{2}$, and with the simplifications developed so far the formula for the asymptotic variance equals

$$\sigma_\tau^2 = (-1)^{\mu+1} \frac{32\,\gamma_\nu}{\pi^2} \sum_{k=0}^{2\mu-1} \frac{(-1)^k}{\mu+k} \binom{2\mu-1}{k} \binom{2\mu+k-1}{k} \\ \times \int_0^\infty u \arctan^2 u \int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt\,du\,.$$

7.2.1 Reduction of the $\arctan^2$ for even ν

The next aim is to reduce $\arctan^2 u$ to $\arctan u$, which is done by the following formula (see Section 7.5.3):

$$\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)} \arctan u\,du = \int_0^\infty \frac{u}{(t+u^2)^2} \arctan^2 u\,du\,.$$

This brings us to

$$\sigma_\tau^2 = (-1)^{\mu+1} \frac{32\,\gamma_\nu}{\pi^2} \sum_{k=0}^{2\mu-1} \frac{(-1)^k}{\mu+k} \binom{2\mu-1}{k} \binom{2\mu+k-1}{k} \\ \times \int_0^1 t^{\mu+k} \int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)} \arctan u\,du\,dt\,.$$

7.2.2 Solution of $\int_0^1 t^{\mu+k} \int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)} \arctan u\,du\,dt$

Split of the fraction

To solve the remaining integral we use the following formula (see Section 7.5.4)

$$\frac{t^{\mu+k}}{(1+u^2)\,(t+u^2)} = \frac{1}{(1+u^2)\,(t+u^2)} + \left(\frac{1}{1+u^2} - \frac{1}{t+u^2}\right) \sum_{l=0}^{\mu+k-1} t^l\,.$$

In the complete formula for the asymptotic variance the first summand cancels out as it is independent of k. The integrals over the summands where t is just in the numerator can be calculated and we get

$$\sigma_\tau^2 = (-1)^{\mu+1} \frac{32\,\gamma_\nu}{\pi^2} \sum_{k=0}^{2\mu-1} \frac{(-1)^k}{\mu+k} \binom{2\mu-1}{k} \binom{2\mu+k-1}{k} \\ \times \left(\frac{\pi^2}{8} \sum_{l=1}^{\mu+k} \frac{1}{l} - \sum_{l=0}^{\mu+k-1} \int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du \right).$$

Solution of $\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du$

The solution of this integral is developed in Section 7.5.5. It involves several changes of the order of integration and again a split of the fraction. We finally get

$$\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du = \frac{1}{2l+1}\left(\frac{1}{2}\sum_{n=1}^{l}\frac{1}{n^2}+\frac{\pi^2}{6}\right).$$

7.2.3 Solution for the asymptotic variance for even ν

Using all the developed steps we get our final result for $\nu \in 2\mathbb{N}$:

$$\sigma_\tau^2 = (-1)^{\frac{\nu}{2}+1}\,\frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})}\sum_{k=0}^{\nu-1}\frac{(-1)^k}{\nu+2k}\binom{\nu-1}{k}\binom{\nu+k-1}{k}$$
$$\times\left(\frac{\pi^2}{4}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{l}-\frac{\pi^2}{3}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1}-\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1}\sum_{n=1}^{l-1}\frac{1}{n^2}\right).$$

7.3 Solution for odd ν

For odd ν we have $\delta = 0$ and $\mu = \frac{\nu-1}{2}$ and the asymptotic variance can be computed by solving

$$\sigma_\tau^2 = (-1)^{\mu-1}\,\frac{32\,\gamma_\nu}{\pi^2}\sum_{k=0}^{2\mu}\frac{(-1)^k}{\mu+k}\binom{2\mu}{k}\binom{2\mu+k}{k}$$
$$\times\int_0^\infty u^2\,\arctan^2 u\int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt\,du\,.$$

7.3.1 Reduction of the $\arctan^2$ for odd ν

Again the next step is reducing $\arctan^2 u$ to $\arctan u$. The formula, which we develop in Section 7.5.7, is slightly more complicated for odd ν than it was for even ν and looks like $(k = 0, \ldots, 2\mu)$

$$\int_0^1 t^{\mu+k}\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,dt$$
$$= \frac{\pi^3}{24\,(2\mu+2k+1)}+\frac{2\mu+2k}{2\mu+2k+1}\int_0^1 t^{\mu+k}\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\,.$$

The asymptotic variance becomes

$$\begin{aligned}\sigma_\tau^2 = (-1)^{\mu+1}\,\frac{4\,\pi\,\gamma_\nu}{3}\,\sum_{k=0}^{2\mu}\frac{(-1)^k}{(\mu+k)\,(2\mu+2k+1)}\binom{2\mu}{k}\binom{2\mu+k}{k}\\ +(-1)^{\mu+1}\,\frac{64\,\gamma_\nu}{\pi^2}\,\sum_{k=0}^{2\mu}\frac{(-1)^k}{2\mu+2k+1}\binom{2\mu}{k}\binom{2\mu+k}{k}\\ \times\int_0^1 t^{\mu+k}\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\,.\end{aligned}$$

7.3.2 Solution of $\int_0^1 t^{\mu+k}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt$

The main idea is to simplify the integral over u using the following result from an integration by parts (see Section 7.5.8):

$$\begin{aligned}&\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\\ &=\frac{\pi}{2\,(1-t)}\,\log(2)+\frac{1}{2\,(t-1)}\int_0^\infty\frac{1}{1+u^2}\,\log(t+u^2)\,du\,.\end{aligned}$$

As we get problems with the denominator $1-t$ if we integrate over t from 0 to 1, we first have to use the equation

$$t^{\mu+k}=(t-1)\sum_{h=0}^{\mu+k-1}t^h+1\,.$$

The summands can then be simplified and we get

$$\begin{aligned}&\int_0^1 t^{\mu+k}\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\\ &=-\frac{\pi}{2}\,\log(2)\sum_{h=1}^{\mu+k}\frac{1}{h}+\frac{1}{2}\sum_{h=1}^{\mu+k}\int_0^1 t^{h-1}\int_0^\infty\frac{1}{1+u^2}\,\log(t+u^2)\,du\,dt\\ &\quad+\int_0^\infty\frac{u}{(1+u^2)}\,\log(1+\frac{1}{u^2})\,\arctan u\,du\,.\end{aligned}$$

The solutions of the two remaining integrals can be found in the Sections 7.5.8 and 7.5.10, which finally leads to the solution

$$\begin{aligned}&\int_0^1 t^{\mu+k}\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\\ &=\frac{\pi}{2}\left(\frac{\pi^2}{12}-\log^2(2)-\log(2)\sum_{h=1}^{\mu+k}\frac{1}{h}-\sum_{h=1}^{\mu+k}\frac{1}{h}\sum_{l=1}^{2h}\frac{(-1)^l}{l}\right).\end{aligned}$$

Using this result the formula for the asymptotic variance becomes

$$\begin{aligned}\sigma_\tau^2 &= (-1)^{\mu+1}\,\frac{4\,\pi\,\gamma_\nu}{3}\sum_{k=0}^{2\mu}\frac{(-1)^k}{(\mu+k)\,(2\mu+2k+1)}\binom{2\mu}{k}\binom{2\mu+k}{k}\\ &\quad+(-1)^{\mu+1}\,\frac{32\,\gamma_\nu}{\pi}\sum_{k=0}^{2\mu}\frac{(-1)^k}{2\mu+2k+1}\binom{2\mu}{k}\binom{2\mu+k}{k}\\ &\qquad\times\left(\frac{\pi^2}{12}-\log^2(2)-\log(2)\sum_{h=1}^{\mu+k}\frac{1}{h}-\sum_{h=1}^{\mu+k}\frac{1}{h}\sum_{l=1}^{2h}\frac{(-1)^l}{l}\right).\end{aligned}$$

7.3.3 Solution for the asymptotic variance for odd ν

After recombining the two summands again (see Section 7.5.9) and calculating the coefficient of $\log^2(2)$, we get the final result for the asymptotic variance for odd ν as

$$\begin{aligned}\sigma_\tau^2 = \frac{16}{\pi^2}\log^2(2) &+ (-1)^{\frac{\nu-1}{2}}\,\frac{32\,\Gamma(\frac{3\nu}{2})}{\pi\,\Gamma^3(\frac{\nu}{2})}\sum_{k=0}^{\nu-1}\frac{(-1)^k}{\nu+2k}\binom{\nu-1}{k}\binom{\nu+k-1}{k}\\ &\times\left(\sum_{h=1}^{\frac{\nu-1}{2}+k}\frac{1}{h}\sum_{l=1}^{2h}\frac{(-1)^l}{l}+\log(2)\sum_{h=1}^{\frac{\nu-1}{2}+k}\frac{1}{h}\right).\end{aligned}$$

7.4 Helpful results

Before presenting the details of the proof we provide some results we will need for it. We first introduce the polylogarithm and some of its properties, then we show some simplifications for sums over binomial coefficients and finally develop the solutions for two definite integrals.

7.4.1 The polylogarithm

Within our work we use several times a special function called polylogarithm. It is defined for $s \in \mathbb{C}$ and $|z| < 1$ as

$$\mathrm{Li}_s(z) = \sum_{k=1}^{\infty}\frac{z^k}{k^s}\,. \tag{7.1}$$

Observe that for all $s \in \mathbb{C}$ we have

$$\mathrm{Li}_s(0) = 0\,.$$

For $s \in \mathbb{N}$ the polylogarithm can be defined recursively:

$$\begin{aligned}\mathrm{Li}_0(z) &= \frac{z}{1-z}\,,\\ \mathrm{Li}_{n+1}(z) &= \int_0^z \frac{1}{t}\,\mathrm{Li}_n(t)\,dt\,,\quad n\in\mathbb{N}_0\,.\end{aligned}$$

So we know that

$$\mathrm{Li}_1(z) = -\log(1-z)$$

and

$$\mathrm{Li}_2(z) = -\int_0^z \frac{\log(1-t)}{t}\, dt\,.$$

This integral representation also allows an extension of the range outside the unit circle. We need the following well-known values of the dilogarithm (see e.g. Prudnikov et al., 1992, p. 498):

$$\mathrm{Li}_2(1) = \frac{\pi^2}{6}\,,$$

$$\mathrm{Li}_2\left(\frac{1}{2}\right) = \frac{\pi^2}{12} - \frac{1}{2}\log^2(2)\,,$$

and

$$\mathrm{Li}_2(-1) = -\frac{\pi^2}{12}\,.$$

7.4.2 Sums over binomial coefficients

At some points of this work we apply simplifications of sums over binomial coefficients. The proofs of these equations can be found in this subsection. We use the following definition of a binomial coefficient:

$$\binom{n}{k} = \begin{cases} \frac{n\,(n-1)\ldots(n-k+1)}{k!}\,, & \text{if } k \geq 0\,, \\ 0\,, & \text{if } k < 0\,, \end{cases} \qquad k \in \mathbb{Z},\ n \in \mathbb{R}\,.$$

We will need some well-known identities, like the upper negation

$$\binom{n}{k} = (-1)^k \binom{k-n-1}{k}\,, \qquad k \in \mathbb{Z},\ n \in \mathbb{R}\,, \tag{7.2}$$

the trinomial revision

$$\binom{r}{m}\binom{m}{k} = \binom{r}{k}\binom{r-k}{m-k}\,, \qquad m, k \in \mathbb{Z},\ r \in \mathbb{R}\,, \tag{7.3}$$

and also Vandermonde's identity

$$\sum_{k=0}^{r} \binom{m}{k}\binom{n}{r-k} = \binom{n+m}{r}\,, \qquad r \in \mathbb{Z},\ m, n \in \mathbb{R}\,. \tag{7.4}$$

For proofs see e.g. Graham et al. (1994, p. 174).

The result of our first lemma is used in the example for a spherical decomposable density (Remark 5.21).

Lemma 7.1. *Let $i \in \mathbb{N}_0$. Then*

$$\sum_{j=0}^{i} \binom{2j}{j} \binom{2\,(i-j)}{i-j} = 4^i \,.$$

Proof. Starting with the upper negation (7.2) and using the definition of a binomial coefficient, we get for every $k \in \mathbb{N}_0$

$$\begin{aligned}(-1)^k \, 2^{2k} \binom{-\frac{1}{2}}{k} &= 2^{2k} \binom{k-\frac{1}{2}}{k} = 2^{2k} \, \frac{(k-\frac{1}{2}) \cdot (k-\frac{3}{2}) \cdot \ldots \cdot \frac{3}{2} \cdot \frac{1}{2}}{k!} \\ &= \frac{(2k-1) \cdot (2k-3) \cdot \ldots \cdot 3 \cdot 1}{k!} = \frac{(2k)!}{(k!)^2} \\ &= \binom{2k}{k} \,. \end{aligned} \tag{7.5}$$

Using this equation and Vandermonde's identity (7.4) we get

$$\begin{aligned}\sum_{j=0}^{i} \binom{2j}{j} \binom{2\,(i-j)}{i-j} &= (-1)^i \, 2^{2i} \sum_{j=0}^{i} \binom{-\frac{1}{2}}{j} \binom{-\frac{1}{2}}{i-j} \\ &\overset{(7.4)}{=} (-1)^i \, 2^{2i} \binom{-1}{i} = 4^i \,, \end{aligned}$$

which we wanted to prove. □

The next four lemmata prove equations that are applied in the proofs for the asymptotic variance of the uncorrelated t-distribution. The first one is used in Section 7.1.

Lemma 7.2. *Let $\mu \in \mathbb{N}$, $\gamma \in \{0, \ldots, 2\mu\}$ and $k \in \mathbb{N}_0$. Then*

$$\sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \binom{2\mu+j+k-\gamma}{\mu+j+1-\gamma} = (-1)^{\mu-1} \binom{2\mu+k-\gamma}{k} \,. \tag{7.6}$$

Proof. By replacing j by $\mu-1-j$ on the left-hand side of (7.6), i.e. changing the order of summation, we get

$$\begin{aligned}&\sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \binom{2\mu+j+k-\gamma}{\mu+j+1-\gamma} \\ &\quad = (-1)^{\mu-1} \sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \binom{3\mu-j+k-1-\gamma}{2\mu-j-\gamma} \,.\end{aligned}$$

Now we us the upper negation (7.2) to rewrite

$$(-1)^j \binom{3\mu-j+k-1-\gamma}{2\mu-j-\gamma} = (-1)^{2\mu-\gamma} \binom{-\mu-k}{2\mu-j-\gamma}$$

and

$$\binom{2\mu+k-\gamma}{k} = \binom{2\mu+k-\gamma}{2\mu-\gamma} = (-1)^{2\mu-\gamma}\binom{-k-1}{2\mu-\gamma}.$$

Newton's binomial formula gives for $|x|<1$

$$\begin{aligned}
\sum_{h=0}^{\infty}\binom{-k-1}{h}x^h &= (1+x)^{-k-1} \\
&= (1+x)^{\mu-1}(1+x)^{-\mu-k} \\
&= \left(\sum_{i=0}^{\mu-1}\binom{\mu-1}{i}x^i\right)\left(\sum_{h=0}^{\infty}\binom{-\mu-k}{h}x^h\right) \\
&= \sum_{r=0}^{\infty}\left(\sum_{j=0}^{r}\binom{\mu-1}{j}\binom{-\mu-k}{r-j}\right)x^r .
\end{aligned}$$

All the coefficients have to coincide, especially the one of $x^{2\mu-\gamma}$. So we have

$$\begin{aligned}
\binom{-k-1}{2\mu-\gamma} &= \sum_{j=0}^{2\mu-\gamma}\binom{\mu-1}{j}\binom{-\mu-k}{2\mu-j-\gamma} \\
&= \sum_{j=0}^{\mu-1}\binom{\mu-1}{j}\binom{-\mu-k}{2\mu-j-\gamma},
\end{aligned}$$

where the last equation holds as the summands equal zero for $i > 2\mu-\gamma$ or $i > \mu-1$. □

The next two lemmata belong together, as the proof of the second one is based on the first one. The result of Lemma 7.3 is used in the proof for even ν in Section 7.5.4, the result of Lemma 7.4 is applied in the end of the proof for odd ν in Section 7.5.9.

Lemma 7.3. *Let $\mu \in \mathbb{N}$, $j \in \{0,\dots,\mu-1\}$ and $\gamma \in \{0,\dots,\mu\}$. Then*

$$\sum_{k=0}^{2\mu-\gamma}(-1)^k\binom{2\mu-\gamma}{k}\binom{2\mu+j+k-\gamma}{\mu+k} = 0 . \tag{7.7}$$

Proof. Using the upper negation (7.2), the second binomial coefficient can be written in the following way:

$$\binom{2\mu+j+k-\gamma}{\mu+k} = (-1)^{\mu+k}\binom{-\mu-j+\gamma-1}{\mu+k}.$$

We now change the order of summation by replacing k by $2\mu-\gamma-k$ and get

$$\sum_{k=0}^{2\mu-\gamma}(-1)^{\mu}\binom{2\mu-\gamma}{k}\binom{-\mu-j+\gamma-1}{\mu+k} = \sum_{k=0}^{2\mu-\gamma}(-1)^{\mu}\binom{2\mu-\gamma}{k}\binom{-\mu-j+\gamma-1}{3\mu-\gamma-k}.$$

Similar to the previous proof we apply Newton's binomial formula for $|x| < 1$:

$$\begin{aligned}\sum_{h=0}^{\mu-j-1} \binom{\mu-j-1}{h} x^h &= (1+x)^{\mu-j-1} \\ &= (1+x)^{2\mu-\gamma}(1+x)^{-\mu-j+\gamma-1} \\ &= \left(\sum_{i=0}^{2\mu-\gamma} \binom{2\mu-\gamma}{i} x^i\right)\left(\sum_{h=0}^{\infty} \binom{-\mu-j+\gamma-1}{h} x^h\right) \\ &= \sum_{r=0}^{\infty}\left(\sum_{k=0}^{r} \binom{2\mu-\gamma}{k}\binom{-\mu-j+\gamma-1}{r-k}\right) x^r .\end{aligned}$$

So for all $r \geq \mu - j$ we need to have a coefficient equal to zero,

$$\sum_{k=0}^{r} \binom{2\mu-\gamma}{k}\binom{-\mu-j+\gamma-1}{r-k} = 0 \quad \text{for all } r \geq \mu - j\,,$$

especially for $r = 3\mu - \gamma \geq \mu - j$, such that we get

$$\sum_{k=0}^{3\mu-\gamma} \binom{2\mu-\gamma}{k}\binom{-\mu-j+\gamma-1}{3\mu-\gamma-k} = 0\,.$$

Since the summands with $k > 2\mu - \gamma$ equal zero, we get the result. □

Lemma 7.4. *Let $\mu \in \mathbb{N}$ and $\gamma \in \{0, \dots, \mu\}$. Then*

$$\sum_{k=0}^{2\mu-\gamma} \frac{(-1)^k}{\mu+k}\binom{2\mu-\gamma}{k}\binom{2\mu+k-\gamma}{k} = 0\,. \tag{7.8}$$

Proof. We first rewrite the second binomial coefficient by using Lemma 7.2 and then we interchange the summations:

$$\begin{aligned}&\sum_{k=0}^{2\mu-\gamma} \frac{(-1)^k}{\mu+k}\binom{2\mu-\gamma}{k}\binom{2\mu+k-\gamma}{k} \\ &= (-1)^{\mu-1}\sum_{k=0}^{2\mu-\gamma} \frac{(-1)^k}{\mu+k}\binom{2\mu-\gamma}{k}\sum_{j=0}^{\mu-1}(-1)^j\binom{\mu-1}{j}\binom{2\mu+j+k-\gamma}{\mu+j+1-\gamma} \\ &= (-1)^{\mu-1}\sum_{j=0}^{\mu-1} \frac{(-1)^j}{\mu+j+1-\gamma}\binom{\mu-1}{j}\sum_{k=0}^{2\mu-\gamma}(-1)^k\binom{2\mu-\gamma}{k}\binom{2\mu+j+k-\gamma}{\mu+k} \\ &= 0\,,\end{aligned}$$

where we used Lemma 7.3 in the last step. □

The last lemma we show simplifies the coefficient of $\log^2(2)$ in the final solution for odd ν (see Section 7.5.10).

Lemma 7.5. *Let $\nu \in 2\mathbb{N}-1$. Then*

$$(-1)^{\frac{\nu-1}{2}} \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi\,\Gamma^3(\frac{\nu}{2})} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k} \binom{\nu-1}{k} \binom{\nu+k-1}{k} = \frac{16}{\pi^2}\,.$$

Proof. We show that

$$\sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k} \binom{\nu-1}{k} \binom{\nu+k-1}{k} = (-1)^{\frac{\nu-1}{2}} \frac{\Gamma^3(\frac{\nu}{2})}{2\pi\,\Gamma(\frac{3\nu}{2})}\,. \tag{7.9}$$

For $\nu = 1$ the assertion is true, so assume $\nu \geq 3$ in the following. We want to use a solution in Graham et al. (1994, pp. 184–185). We first use the trinomial revision (7.3) with $m = 2k$ and $r = \nu - 1 + k$ to rewrite our product of binomial coefficients:

$$\binom{\nu-1}{k} \binom{\nu-1+k}{k} = \binom{\nu-1+k}{2k} \binom{2k}{k}.$$

So the left-hand side of (7.9) can be rewritten, using S_m as defined in Graham et al. (1994, p. 184) with $n = \nu - 1$ and $m = \frac{\nu}{2} - 1$:

$$\begin{aligned}\sum_{k=0}^{\nu-1} \frac{(-1)^k}{2k+\nu} \binom{\nu-1}{k} \binom{\nu+k-1}{k} &= \frac{1}{2} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{k+\frac{\nu}{2}} \binom{\nu-1+k}{2k} \binom{2k}{k} \\ &= \frac{1}{2}\, S_{\frac{\nu}{2}-1}(\nu-1)\,.\end{aligned}$$

This choice of parameters is possible due to the last comment on page 185, although $m \notin \mathbb{N}$. Using the result

$$S_m(n) = (-1)^n \frac{m\,(m-1)\,\ldots\,(m-n+1)}{(m+n+1)\,(m+n)\,\ldots\,(m+1)}$$

from Graham et al. (1994, p. 185) and the property of the gamma distribution

$$\Gamma\Big(\frac{x}{2}\Big) = \Big(\frac{x}{2}-1\Big) \cdot \Big(\frac{x}{2}-2\Big) \cdot \ldots \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi}, \quad \text{if } x \in 2\mathbb{N}+1\,,$$

we get

$$\begin{aligned}\frac{1}{2}\, S_{\frac{\nu}{2}-1}(\nu-1) &= \frac{(-1)^{\nu-1}}{2} \frac{(\frac{\nu}{2}-1)\cdot(\frac{\nu}{2}-2)\cdot\ldots\cdot\frac{3}{2}\cdot\frac{1}{2}\cdot(-\frac{1}{2})\cdot(-\frac{3}{2})\cdot\ldots\cdot(-\frac{\nu}{2}+1)}{(\frac{3\nu}{2}-1)\cdot(\frac{3\nu}{2}-2)\cdot\ldots\cdot(\frac{\nu}{2}+1)\cdot\frac{\nu}{2}} \\ &= \frac{(-1)^{\frac{\nu-1}{2}}}{2\pi} \frac{\Gamma^3(\frac{\nu}{2})}{\Gamma(\frac{3\nu}{2})}\,,\end{aligned}$$

which is the right-hand side of (7.9). □

7.4.3 A special series

In Section 7.4.4 we will need the values of two special series which are stated in the following two lemmata.

Lemma 7.6.

$$\sum_{n=1}^{\infty} \frac{1}{n\, 2^{2n+1}} \binom{2n}{n} = \log(2)\,.$$

Proof. Starting with Newton's binomial formula we get for every $t \in (-1, 1)$

$$(1-t)^{-\frac{1}{2}} = \sum_{n=0}^{\infty} (-t)^n \binom{-\frac{1}{2}}{n} = \sum_{n=0}^{\infty} \frac{t^n}{2^{2n}} \binom{2n}{n}\,,$$

where we used (7.5). Subtracting 1 and dividing by t gives

$$\frac{1-\sqrt{1-t}}{t\sqrt{1-t}} = \sum_{n=1}^{\infty} \frac{t^{n-1}}{2^{2n}} \binom{2n}{n}\,.$$

If we integrate both sides with respect to t, we get

$$-2\log\bigl(2\,(1+\sqrt{1-x})\bigr) + 2\log(4) = \sum_{n=1}^{\infty} \frac{x^n}{n\, 2^{2n}} \binom{2n}{n} \quad \text{for all } x \in [0,1)\,. \tag{7.10}$$

From Abel's theorem we know that the series on the right-hand side converges for $x \nearrow 1$. As it is further bounded for all $x \in [0, 1)$, the limit must be finite and equals the value for $x = 1$ on the left-hand side, which gives the result. □

Lemma 7.7.

$$\sum_{n=1}^{\infty} \frac{1}{n^2\, 2^{2n+2}} \binom{2n}{n} = \frac{\pi^2}{24} - \frac{1}{2}\,\log^2(2)\,.$$

Proof. We start with equation (7.10) and divide again by x, such that we get for all $x \in [0, 1]$

$$-\frac{1}{x}\,\log\Bigl(\frac{1}{2}\,\bigl(1+\sqrt{1-x}\bigr)\Bigr) = \sum_{n=1}^{\infty} \frac{x^{n-1}}{n\, 2^{2n+1}} \binom{2n}{n}\,.$$

If we integrate both sides with respect to x, we get

$$-\frac{1}{2}\,\log^2\Bigl(\frac{1}{2}\,\bigl(1+\sqrt{1-u}\bigr)\Bigr) + \mathrm{Li}_2\Bigl(\frac{1}{2} - \frac{\sqrt{1-u}}{2}\Bigr) = \sum_{n=1}^{\infty} \frac{u^n}{n^2\, 2^{2n+1}} \binom{2n}{n} \quad \text{for all } u \in [0,1]\,.$$

Setting $u = 1$ and knowing that $\mathrm{Li}_2(\frac{1}{2}) = \frac{\pi^2}{12} - \frac{1}{2}\,\log^2(2)$ (see Section 7.4.1), we get the solution. □

7.4.4 Some integrals

Within our proof of the asymptotic variance there are two definite integrals that are a little more complicated. Both solutions are based on an integral that can be solved by complex integration, which is done in the first part. The two definite integrals can be found subsequently.

Solution of $\int_0^\infty (1+u^2)^{-n-1}\,du$

Lemma 7.8. *Let $n \in \mathbb{N}$. Then*

$$\int_0^\infty \frac{1}{(1+u^2)^{n+1}}\,du = \frac{\pi}{2^{2n+1}} \binom{2n}{n}. \tag{7.11}$$

Proof. We want to solve the integral using complex integration (for details see e.g. Freitag and Busam, 1993). The integrand can be written as

$$f(u) = \frac{1}{(u+i)^{n+1}} \, \frac{1}{(u-i)^{n+1}}\,, \quad u \in \mathbb{C} \setminus \{-i, i\}\,.$$

So we work on the complex plane but have to exclude the points $-i$ and i. We define the curves

$$\begin{aligned} \alpha_K :\quad & [\,-K, K\,] \to \mathbb{C} \\ & u \mapsto u \\ \beta_K :\quad & [0, \pi\,] \to \mathbb{C} \\ & u \mapsto K\,e^{iu} \end{aligned}$$

with $K \in (1, \infty)$ a large positive real number. The winding number of this curve around the pole at i is 1. The order of the pole is $n+1$, such that the residue can be calculated in the following way (see Freitag and Busam, 1993, p. 165):

$$\begin{aligned} \operatorname{Res}(f; i) &= \frac{1}{n!} \frac{\partial^n}{\partial u^n} \Big((u-i)^{n+1} f(u) \Big)\Big|_{u=i} \\ &= \frac{1}{n!} \frac{\partial^n}{\partial u^n} \Big(\frac{1}{(u+i)^{n+1}} \Big)\Big|_{u=i} \\ &= \frac{(-1)^n}{n!} (n+1)\,(n+2)\, \ldots\, (2n)\, \frac{1}{(u+i)^{2n+1}}\Big|_{u=i} \\ &= (-1)^n \binom{2n}{n} \frac{1}{(2i)^{2n+1}} \\ &= -\frac{i}{2^{2n+1}} \binom{2n}{n}. \end{aligned}$$

The residue theorem tells then that

$$\int_{\alpha_K} f(\zeta)\,d\zeta + \int_{\beta_K} f(\zeta)\,d\zeta = 2\pi i \operatorname{Res}(f; i) = \frac{\pi}{2^{2n}} \binom{2n}{n}.$$

For the first integral we have

$$\int_{\alpha_K} f(\zeta)\, d\zeta = \int_{-K}^{K} f(u)\, du = 2\int_0^K \frac{1}{(1+u^2)^{n+1}}\, du\,,$$

so in the limit $K \to \infty$ it is twice the one we are looking for. So it is just left to show that the limit of the second integral equals zero. Since $|1+z| \geq |z| - 1$ for all $z \in \mathbb{C}$ and since $K > 1$, we have

$$\sup_{t\in[0,\pi]} \left|\frac{1}{(1+K^2\, e^{2it})^{n+1}}\right| \leq \frac{1}{(K^2-1)^{n+1}}\,.$$

Knowing further that the length of β_K equals $l(\beta_K) = \pi K$, the estimation lemma (see Freitag and Busam, 1993, p. 65) tells

$$\left|\int_{\beta_K} f(\zeta)\, d\zeta\right| \leq \frac{\pi K}{(K^2-1)^{n+1}}$$

and since $n \geq 1$, the limit is

$$\lim_{K\to\infty} \left|\int_{\beta_K} f(\zeta)\, d\zeta\right| = 0\,.$$

□

Solution of $\int_0^\infty \frac{1}{(1+u^2)} \log(1+u^2)\, du$

In Section 7.5.8 we need the following solution:

Lemma 7.9.

$$\int_0^\infty \frac{1}{(1+u^2)} \log(1+u^2)\, du = \pi \log(2)\,.$$

Proof. We use the Taylor expansion of the logarithm, as for every $u \in \mathbb{R}$

$$\begin{aligned}\log(1+u^2) &= -\log\Big(\frac{1}{1+u^2}\Big) = -\log\Big(1 - \frac{u^2}{1+u^2}\Big)\\ &= \sum_{n=1}^{\infty} \frac{1}{n}\Big(\frac{u^2}{1+u^2}\Big)^n.\end{aligned}$$

Rewriting u^{2n} as

$$u^{2n} = (1+u^2-1)^n = \sum_{k=0}^{n} \binom{n}{k} (1+u^2)^{n-k} (-1)^k$$

we get

$$\log(1+u^2) = \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} (-1)^k \frac{1}{(1+u^2)^k}\,.$$

Monotone convergence allows to interchange the integral and the sum such that we get

$$\begin{aligned}\int_0^\infty \frac{1}{(1+u^2)} \log(1+u^2)\,du &= \sum_{n=1}^\infty \frac{1}{n} \sum_{k=0}^n \binom{n}{k} (-1)^k \int_0^\infty \frac{1}{(1+u^2)^{k+1}}\,du \\ &= \frac{\pi}{2} \sum_{n=1}^\infty \frac{1}{n} \sum_{k=0}^n \frac{(-1)^k}{2^{2k}} \binom{n}{k}\binom{2k}{k},\end{aligned}$$

where we used solution (7.11) with $n = k$. We can simplify the inner sum by applying a result from Graham et al. (1994, (5.23), p. 169):

$$\sum_{k=0}^n \frac{(-1)^k}{2^{2k}} \binom{n}{k}\binom{2k}{k} \overset{(7.5)}{=} \sum_{k=0}^n \binom{n}{k}\binom{-\frac{1}{2}}{k} = \binom{n-\frac{1}{2}}{n} \overset{(7.2)}{=} (-1)^n \binom{-\frac{1}{2}}{n} \overset{(7.5)}{=} \frac{1}{2^{2n}}\binom{2n}{n}.$$

Using Lemma 7.6 we get the result. □

Solution of $\int_0^\infty \frac{u}{(1+u^2)} \log(1+\frac{1}{u^2}) \arctan u\,du$

The solution of the next integral is used in the end of Section 7.5.8.

Lemma 7.10.

$$\int_0^\infty \frac{u}{(1+u^2)} \log\Big(1+\frac{1}{u^2}\Big) \arctan u\,du = \frac{\pi}{2}\left(\frac{\pi^2}{12} - \log^2(2)\right).$$

Proof. We apply again the Taylor expansion of the logarithm, where for every $u \in \mathbb{R}$

$$\begin{aligned}\log\Big(1+\frac{1}{u^2}\Big) &= -\log\Big(\frac{u^2}{1+u^2}\Big) = -\log\Big(1-\frac{1}{1+u^2}\Big) \\ &= \sum_{n=1}^\infty \frac{1}{n}\Big(\frac{1}{1+u^2}\Big)^n.\end{aligned}$$

Interchanging the integral and the sum is again possible due to monotone convergence, such that we get

$$\begin{aligned}&\int_0^\infty \frac{u}{(1+u^2)} \log\Big(1+\frac{1}{u^2}\Big) \arctan u\,du \\ &= \sum_{n=1}^\infty \frac{1}{n} \int_0^\infty \frac{u}{(1+u^2)^{n+1}} \arctan u\,du \\ &= \sum_{n=1}^\infty \frac{1}{n}\left(-\frac{1}{2n}\frac{1}{(1+u^2)^n} \arctan u\,\Big|_{u=0}^\infty + \frac{1}{2n}\int_0^\infty \frac{1}{(1+u^2)^{n+1}}\,du\right) \\ &= \frac{1}{2}\sum_{n=1}^\infty \frac{1}{n^2}\int_0^\infty \frac{1}{(1+u^2)^{n+1}}\,du.\end{aligned}$$

Using solution (7.11) this simplifies to

$$\int_0^\infty \frac{u}{(1+u^2)} \log\Big(1+\frac{1}{u^2}\Big) \arctan u \, du = \pi \sum_{n=1}^\infty \frac{1}{n^2\, 2^{2n+2}} \binom{2n}{n}.$$

Lemma 7.7 gives the result. □

7.5 Details of the proof of the asymptotic variance

7.5.1 Reduction of the power of $u^{2\mu-\delta}$

To divide $u^{2\mu-\delta}$ by $(u^2+t)^{2\mu+1-\delta}$ as far as possible we use the representation

$$\begin{aligned} u^{2\mu-\delta} &= u^{2-\delta}\,(u^2+t-t)^{\mu-1} \\ &= u^{2-\delta} \sum_{j=0}^{\mu-1} \binom{\mu-1}{j} (u^2+t)^{\mu-j-1}\,(-t)^j\,. \end{aligned} \tag{7.12}$$

We get

$$\begin{aligned} &u^{2\mu-\delta}\, t^{2\mu-\delta}\,(1-t)^{2\mu-\delta}\,\big(u^2+t\big)^{-2\mu-1+\delta} \\ &\quad = u^{2-\delta} \sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\,. \end{aligned}$$

Like this we come to integrals of the kind

$$\sum_{j=0}^{\mu-1} (-1)^j \binom{\mu-1}{j} \int_0^\infty u^{2-\delta} \arctan^2 u \int_0^1 t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\, dt\, du\,.$$

7.5.2 Rule of Leibniz for $\int_0^1 t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\,dt$

We use the rule of Leibniz to simplify the integral ($j=0,\dots,\mu-1$)

$$\int_0^1 t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\,dt\,.$$

For every j take $f(t) = t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}$ where we know for every $n=1,\dots,\mu+j-\delta-1$ that $f^{(n)}(0) = f^{(n)}(1) = 0$. Further we know

$$f(t) = \sum_{k=0}^{2\mu-\delta} (-1)^k \binom{2\mu-\delta}{k} t^{2\mu+j+k-\delta}$$

and therefore

$$f^{(\mu+j-\delta)}(t) = \sum_{k=0}^{2\mu-\delta} (-1)^k \binom{2\mu-\delta}{k} \frac{(2\mu+j+k-\delta)!}{(\mu+k)!}\, t^{\mu+k}\,.$$

We further take

$$g^{(\mu+j-\delta)}(t) = (t+u^2)^{-\mu-j-2+\delta}$$

where we get

$$g(t) = \frac{(-1)^{\mu+j-\delta}}{(\mu+j+1-\delta)!}\,(t+u^2)^{-2}\,.$$

So in total we have

$$\begin{aligned}
&\int_0^1 t^{2\mu+j-\delta}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\,dt \\
&\quad = (-1)^{\mu+j-\delta}\int_0^1 \left(\sum_{k=0}^{2\mu-\delta}(-1)^k\binom{2\mu-\delta}{k}\frac{(2\mu+j+k-\delta)!}{(\mu+k)!}\,t^{\mu+k}\right) \\
&\qquad\qquad\qquad \times \frac{(-1)^{\mu+j-\delta}}{(\mu+j+1-\delta)!}\,(t+u^2)^{-2}\,dt \\
&\quad = \sum_{k=0}^{2\mu-\delta}(-1)^k\binom{2\mu-\delta}{k}\frac{1}{\mu+k}\,\frac{(2\mu+j+k-\delta)!}{(\mu+k-1)!\,(\mu+j+1-\delta)!}\int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt \\
&\quad = \sum_{k=0}^{2\mu-\delta}\frac{(-1)^k}{\mu+k}\binom{2\mu-\delta}{k}\binom{2\mu+j+k-\delta}{\mu+j+1-\delta}\int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt\,.
\end{aligned}$$

Our formula of the asymptotic variance becomes

$$\begin{aligned}
\sigma_\tau^2 &= \frac{32\,\gamma_\nu}{\pi^2}\sum_{j=0}^{\mu-1}(-1)^j\binom{\mu-1}{j}\int_0^\infty u^{2-\delta}\,\arctan^2 u \\
&\qquad\qquad \times \int_0^1 t^{2\mu-\delta+j}\,(1-t)^{2\mu-\delta}\,(u^2+t)^{-\mu-j-2+\delta}\,dt\,du \\
&= \frac{32\,\gamma_\nu}{\pi^2}\sum_{j=0}^{\mu-1}(-1)^j\binom{\mu-1}{j}\sum_{k=0}^{2\mu-\delta}\frac{(-1)^k}{\mu+k}\binom{2\mu-\delta}{k}\binom{2\mu+j+k-\delta}{\mu+j+1-\delta} \\
&\qquad\qquad \times \int_0^\infty u^{2-\delta}\,\arctan^2 u\int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt\,du \\
&= \frac{32\,\gamma_\nu}{\pi^2}\sum_{k=0}^{2\mu-\delta}\frac{(-1)^k}{\mu+k}\binom{2\mu-\delta}{k}\sum_{j=0}^{\mu-1}(-1)^j\binom{\mu-1}{j}\binom{2\mu+j+k-\delta}{\mu+j+1-\delta} \\
&\qquad\qquad \times \int_0^\infty u^{2-\delta}\,\arctan^2 u\int_0^1 t^{\mu+k}\,(t+u^2)^{-2}\,dt\,du\,.
\end{aligned}$$

7.5.3 Reduction of the $\arctan^2$ for even ν

To derive the formula to reduce $\arctan^2 u$ to $\arctan u$ for even ν we integrate

$$\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)} \arctan u \, du\,,$$

using integration by parts with $f(u) = \frac{1}{t+u^2}$, $g'(u) = \frac{1}{1+u^2} \arctan u$, $f'(u) = -\frac{2u}{(t+u^2)^2}$ and $g(u) = \frac{1}{2} \arctan^2 u$. We get

$$\begin{aligned}
&\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)} \arctan u \, du \\
&\quad = \frac{1}{2\,(t+u^2)} \arctan^2 u \,\Big|_{u=0}^{\infty} + \int_0^\infty \frac{u}{(t+u^2)^2} \arctan^2 u \, du \\
&\quad = \int_0^\infty \frac{u}{(t+u^2)^2} \arctan^2 u \, du\,.
\end{aligned}$$

7.5.4 Split of the fraction

For further calculations we need to split the fraction. We know that

$$t^{\mu+k} = 1 + (t-1) \sum_{l=0}^{\mu+k-1} t^l$$

and that

$$\frac{(t-1)}{(1+u^2)\,(t+u^2)} = \frac{1}{1+u^2} - \frac{1}{t+u^2}$$

and therefore get

$$\frac{t^{\mu+k}}{(1+u^2)\,(t+u^2)} = \frac{1}{(1+u^2)\,(t+u^2)} + \Big(\frac{1}{1+u^2} - \frac{1}{t+u^2}\Big) \sum_{l=0}^{\mu+k-1} t^l\,. \tag{7.13}$$

The first summand is independent of k so we can take it out of the sum. Using Lemma 7.3 with $\gamma = 1$ we know that

$$\sum_{k=0}^{2\mu-1} (-1)^k \binom{2\mu-1}{k} \binom{2\mu+j+k-1}{\mu+k} = 0$$

and so within the complete formula of the asymptotic variance the first summand in (7.13) cancels out. The integrals over the summands where t is only in the numerator can be solved as they can be split into a product where each factor just contains one of the variables:

$$\sum_{l=0}^{\mu+k-1} \Big(\int_0^\infty \frac{1}{1+u^2} \arctan u \, du \int_0^1 t^l \, dt \Big) = \frac{\pi^2}{8} \sum_{l=1}^{\mu+k} \frac{1}{l}\,.$$

So our formula for the asymptotic variance becomes

$$\begin{aligned}\sigma_\tau^2 = (-1)^{\mu+1}\,\frac{32\,\gamma_\nu}{\pi^2} &\sum_{k=0}^{2\mu-1} \frac{(-1)^k}{\mu+k}\binom{2\mu-1}{k}\binom{2\mu+k-1}{k}\\ &\times\left(\frac{\pi^2}{8}\sum_{l=1}^{\mu+k}\frac{1}{l} - \sum_{l=0}^{\mu+k-1}\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du\right).\end{aligned} \tag{7.14}$$

7.5.5 Solution of $\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du$

Integration by parts for $\int_0^1 t^l\,(t+u^2)^{-1}\,dt$

As a first simplification we use integration by parts for the integral over t. With $f(t) = (t+u^2)^{-1}$, $g'(t) = t^l$, $f'(t) = -(t+u^2)^{-2}$ and $g(t) = \frac{1}{l+1}\,t^{l+1}$ we get

$$\begin{aligned}\int_0^1 \frac{t^l}{t+u^2}\,dt &= \frac{1}{l+1}\,\frac{t^{l+1}}{t+u^2}\Big|_{t=0}^1 + \frac{1}{l+1}\int_0^1 \frac{t^{l+1}}{(t+u^2)^2}\,dt\\ &= \frac{1}{l+1}\,\frac{1}{1+u^2} + \frac{1}{l+1}\int_0^1 \frac{t^l}{t+u^2}\,dt - \frac{1}{l+1}\int_0^1 \frac{t^l\,u^2}{(t+u^2)^2}\,dt\,.\end{aligned}$$

So we can use the representation

$$\int_0^1 \frac{t^l}{t+u^2}\,dt = \frac{1}{l}\,\frac{1}{1+u^2} - \frac{1}{l}\int_0^1 \frac{t^l\,u^2}{(t+u^2)^2}\,dt$$

and our integral becomes

$$\begin{aligned}&\int_0^\infty \arctan u \int_0^1 \frac{t^l}{t+u^2}\,dt\,du\\ &\quad= \frac{1}{l}\int_0^\infty \frac{1}{1+u^2}\,\arctan u\,du - \frac{1}{l}\int_0^\infty \arctan u \int_0^1 \frac{t^l\,u^2}{(t+u^2)^2}\,dt\,du\\ &\quad= \frac{\pi^2}{8l} - \frac{1}{l}\int_0^1 t^l \int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan u\,du\,dt\,.\end{aligned}$$

Integration by parts for $\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan u\,du$

To simplify $\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan u\,du$ we use again integration by parts, here with $f(u) = u\,\arctan u$, $g'(u) = \frac{u}{(t+u^2)^2}$, $f'(u) = \arctan u + \frac{u}{1+u^2}$ and $g(u) = -\frac{1}{2}\,\frac{1}{t+u^2}$:

$$\begin{aligned}&\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan u\,du\\ &= -\frac{1}{2}\,\frac{u}{t+u^2}\,\arctan u\,\Big|_{u=0}^\infty + \frac{1}{2}\int_0^\infty \frac{1}{t+u^2}\,\arctan u\,du + \frac{1}{2}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\\ &= \frac{1}{2}\int_0^\infty \frac{1}{t+u^2}\,\arctan u\,du + \frac{1}{2}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,.\end{aligned}$$

Back to the whole integral

If we insert this solution we get

$$\begin{aligned}
&\int_0^\infty \arctan u \int_0^1 \frac{t^l}{t+u^2}\,dt\,du \\
&\quad = \frac{\pi^2}{8l} - \frac{1}{l}\int_0^1 t^l \left(\frac{1}{2}\int_0^\infty \frac{1}{t+u^2}\,\arctan u\,du + \frac{1}{2}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\right)dt \\
&\quad = \frac{\pi^2}{8l} - \frac{1}{2l}\int_0^\infty \arctan u \int_0^1 \frac{t^l}{t+u^2}\,dt\,du - \frac{1}{2l}\int_0^1 t^l \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,dt
\end{aligned}$$

and therefore

$$\int_0^\infty \arctan u \int_0^1 \frac{t^l}{t+u^2}\,dt\,du = \frac{1}{2l+1}\left(\frac{\pi^2}{4} - \int_0^1 t^l \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,dt\right).$$

Solution of the integral $\int_0^1 t^l \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,dt$

To solve this last integral we first make the substitution $z := u^2$. Then we use again the following representation of the fraction, similar to (7.13):

$$\frac{t^l}{(t+z)\,(1+z)} = \frac{1}{(t+z)\,(1+z)} + \left(\frac{1}{1+z} - \frac{1}{t+z}\right)\sum_{k=0}^{l-1} t^k\,.$$

We get the following:

$$\begin{aligned}
&\int_0^1 t^l \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,dt \\
&\quad = \frac{1}{2}\int_0^1 t^l \int_0^\infty \frac{1}{(1+z)\,(t+z)}\,dz\,dt \\
&\quad = \frac{1}{2}\int_0^\infty \frac{1}{1+z}\int_0^1 \frac{1}{t+z}\,dt\,dz + \frac{1}{2}\sum_{k=0}^{l-1}\int_0^1 t^k \int_0^\infty \left(\frac{1}{1+z} - \frac{1}{t+z}\right)dz\,dt \\
&\quad = \frac{1}{2}\int_0^\infty \frac{1}{1+z}\,\log\Big(\frac{1+z}{z}\Big)\,dz + \frac{1}{2}\sum_{k=0}^{l-1}\int_0^1 t^k\,\log(t)\,dt\,,
\end{aligned}$$

where we used

$$\int_0^1 \frac{1}{t+z}\,dt = \log(t+z)\,\Big|_{t=0}^1 = \log\Big(\frac{1+z}{z}\Big)$$

and

$$\int_0^\infty \left(\frac{1}{1+z} - \frac{1}{t+z}\right)dz = \log\Big(\frac{1+z}{t+z}\Big)\,\Big|_{z=0}^\infty = \log(t)\,.$$

The first remaining integral can be solved using the substitution $t := \frac{z}{1+z}$

$$\begin{aligned}\int_0^\infty \frac{1}{1+z} \log\Big(\frac{1+z}{z}\Big)\,dz &= -\int_0^1 \frac{1}{1-t} \log(t)\,dt \\ &= -\int_0^1 \frac{1}{t} \log(1-t)\,dt \\ &= \mathrm{Li}_2(1) = \frac{\pi^2}{6}\,,\end{aligned}$$

the second has the solution

$$\begin{aligned}\int_0^1 t^k \log(t)\,dt &= \frac{1}{k+1}\, t^{k+1} \log(t)\,\big|_{t=0}^1 - \frac{1}{k+1}\int_0^1 t^k\,dt \\ &= -\frac{1}{(k+1)^2}\,.\end{aligned}$$

This gives the value

$$\int_0^1 t^l \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,du\,dt = \frac{\pi^2}{12} - \frac{1}{2}\sum_{k=0}^{l-1} \frac{1}{(k+1)^2}\,.$$

Using all the developed steps we get the solution

$$\begin{aligned}\int_0^\infty \arctan u \int_0^1 \frac{t^l}{t+u^2}\,dt\,du &= \frac{1}{2l+1}\left(\frac{\pi^2}{4} - \frac{\pi^2}{12} + \frac{1}{2}\sum_{k=0}^{l-1}\frac{1}{(k+1)^2}\right) \\ &= \frac{1}{2l+1}\left(\frac{\pi^2}{6} + \frac{1}{2}\sum_{k=1}^{l}\frac{1}{k^2}\right).\end{aligned}$$

7.5.6 Solution for the asymptotic variance for even ν

From (7.14) we know that

$$\begin{aligned}\sigma_\tau^2 = (-1)^{\mu+1}\,\frac{32\,\gamma_\nu}{\pi^2} &\sum_{k=0}^{2\mu-1} \frac{(-1)^k}{\mu+k}\binom{2\mu-1}{k}\binom{2\mu+k-1}{k} \\ &\times\left(\frac{\pi^2}{8}\sum_{l=1}^{\mu+k}\frac{1}{l} - \sum_{l=0}^{\mu+k-1}\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du\right).\end{aligned}$$

We can now insert

$$\int_0^\infty \arctan u \int_0^1 t^l\,(t+u^2)^{-1}\,dt\,du = \frac{1}{2l+1}\left(\frac{1}{2}\sum_{n=1}^{l}\frac{1}{n^2} + \frac{\pi^2}{6}\right)$$

and get our final result for $\nu \in 2\mathbb{N}$:

$$\begin{aligned}\sigma_\tau^2 &= (-1)^{\mu+1}\,\frac{32\,\gamma_\nu}{\pi^2}\sum_{k=0}^{2\mu-1}\frac{(-1)^k}{\mu+k}\binom{2\mu-1}{k}\binom{2\mu+k-1}{k}\\ &\qquad\times\left(\frac{\pi^2}{8}\sum_{l=1}^{\mu+k}\frac{1}{l}-\sum_{l=0}^{\mu+k-1}\frac{1}{2l+1}\left(\frac{1}{2}\sum_{n=1}^{l}\frac{1}{n^2}+\frac{\pi^2}{6}\right)\right)\\ &= (-1)^{\frac{\nu}{2}+1}\,\frac{32\,\Gamma(\frac{3\nu}{2})}{\pi^2\,\Gamma^3(\frac{\nu}{2})}\sum_{k=0}^{\nu-1}\frac{(-1)^k}{\nu+2k}\binom{\nu-1}{k}\binom{\nu+k-1}{k}\\ &\qquad\times\left(\frac{\pi^2}{4}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{l}-\frac{\pi^2}{3}\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1}-\sum_{l=1}^{\frac{\nu}{2}+k}\frac{1}{2l-1}\sum_{n=1}^{l-1}\frac{1}{n^2}\right).\end{aligned}$$

7.5.7 Reduction of the $\arctan^2$ for odd ν

This step is a little more complicated for odd ν than it was for even ν. To derive the formula to reduce $\arctan^2 u$ to $\arctan u$ we first integrate

$$\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,,$$

using integration by parts with $f(u) = u\,\arctan^2 u$, $g'(u) = \frac{u}{(t+u^2)^2}$, $f'(u) = \arctan^2 u + \frac{2\,u}{1+u^2}\,\arctan u$ and $g(u) = -\frac{1}{2}\,\frac{1}{t+u^2}$. We get

$$\begin{aligned}&\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\\ &\quad= -\frac{u}{2\,(t+u^2)}\,\arctan^2 u\,\Big|_{u=0}^{\infty}\\ &\quad\quad+\frac{1}{2}\int_0^\infty\frac{1}{t+u^2}\,\arctan^2 u\,du+\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\\ &\quad=\frac{1}{2}\int_0^\infty\frac{1}{t+u^2}\,\arctan^2 u\,du+\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,.\end{aligned}$$

For the whole double-integral we come to

$$\begin{aligned}&\int_0^1 t^{\mu+k}\int_0^\infty\frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,dt\\ &\quad=\frac{1}{2}\int_0^1 t^{\mu+k}\int_0^\infty\frac{1}{t+u^2}\,\arctan^2 u\,du\,dt\\ &\quad\quad+\int_0^1 t^{\mu+k}\int_0^\infty\frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\,.\end{aligned}$$

To further simplify the first of the two remaining integrals we change the order of integration and modify the integral

$$\int_0^1 \frac{t^{\mu+k}}{t+u^2}\,dt$$

with integration by parts, using $f(t) = \frac{1}{t+u^2}$, $g'(t) = t^{\mu+k}$, $f'(t) = -\frac{1}{(t+u^2)^2}$ and $g(t) = \frac{1}{\mu+k+1}\,t^{\mu+k+1}$, and get

$$\begin{aligned}\int_0^1 \frac{t^{\mu+k}}{t+u^2}\,dt &= \frac{1}{\mu+k+1}\,\frac{t^{\mu+k+1}}{t+u^2}\Big|_{t=0}^{1} + \frac{1}{\mu+k+1}\int_0^1 \frac{t^{\mu+k+1}}{(t+u^2)^2}\,dt \\ &= \frac{1}{\mu+k+1}\left(\frac{1}{1+u^2} + \int_0^1 \frac{t^{\mu+k}}{t+u^2}\,dt - \int_0^1 \frac{t^{\mu+k}\,u^2}{(t+u^2)^2}\,dt\right)\end{aligned}$$

or, equivalently,

$$\int_0^1 \frac{t^{\mu+k}}{t+u^2}\,dt = \frac{1}{\mu+k}\,\frac{1}{1+u^2} - \frac{1}{\mu+k}\int_0^1 \frac{t^{\mu+k}\,u^2}{(t+u^2)^2}\,dt\,.$$

If we insert this above we have

$$\begin{aligned}&\int_0^1 t^{\mu+k}\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,dt \\ &\quad= \frac{1}{2\,(\mu+k)}\int_0^\infty \frac{1}{1+u^2}\,\arctan^2 u\,du - \frac{1}{2\,(\mu+k)}\int_0^\infty \arctan^2 u\int_0^1 \frac{t^{\mu+k}\,u^2}{(t+u^2)^2}\,dt\,du \\ &\qquad+ \int_0^1 t^{\mu+k}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt \\ &\quad= \frac{1}{2\,(\mu+k)}\,\frac{\pi^3}{24} - \frac{1}{2\,(\mu+k)}\int_0^1 t^{\mu+k}\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,dt \\ &\qquad+ \int_0^1 t^{\mu+k}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\end{aligned}$$

and finally get

$$\begin{aligned}&\int_0^1 t^{\mu+k}\int_0^\infty \frac{u^2}{(t+u^2)^2}\,\arctan^2 u\,du\,dt \\ &\quad= \frac{\pi^3}{24\,(2\mu+2k+1)} + \frac{2\mu+2k}{2\mu+2k+1}\int_0^1 t^{\mu+k}\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)}\,\arctan u\,du\,dt\,.\end{aligned}$$

The asymptotic variance becomes

$$\sigma_\tau^2 = (-1)^{\mu-1} \frac{32\,\gamma_\nu}{\pi^2} \sum_{k=0}^{2\mu} \frac{(-1)^k}{\mu+k} \binom{2\mu}{k} \binom{2\mu+k}{k}$$
$$\left(\frac{\pi^3}{24\,(2\mu+2k+1)} + \frac{2\mu+2k}{2\mu+2k+1} \int_0^1 t^{\mu+k} \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \arctan u \, du \, dt \right)$$
$$= (-1)^{\mu-1} \frac{4\,\pi\,\gamma_\nu}{3} \sum_{k=0}^{2\mu} \frac{(-1)^k}{(\mu+k)\,(2\mu+2k+1)} \binom{2\mu}{k} \binom{2\mu+k}{k}$$
$$+ (-1)^{\mu-1} \frac{64\,\gamma_\nu}{\pi^2} \sum_{k=0}^{2\mu} \frac{(-1)^k}{2\mu+2k+1} \binom{2\mu}{k} \binom{2\mu+k}{k}$$
$$\int_0^1 t^{\mu+k} \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \arctan u \, du \, dt \,. \tag{7.15}$$

7.5.8 Solution of $\int_0^1 t^{\mu+k} \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \arctan u \, du \, dt$

Simplification for the integral over u

To solve the remaining integral we first want to look at the integration over u and use integration by parts with $f(u) = \arctan u$, $g'(u) = \frac{u}{(1+u^2)\,(t+u^2)}$, $f'(u) = \frac{1}{1+u^2}$ and $g(u) = \frac{1}{2\,(t-1)} \left(\log(1+u^2) - \log(t+u^2)\right)$:

$$\int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \arctan u \, du$$
$$= \frac{1}{2\,(t-1)} \arctan u \left(\log(1+u^2) - \log(t+u^2)\right) \Big|_{u=0}^\infty$$
$$- \frac{1}{2\,(t-1)} \int_0^\infty \frac{1}{1+u^2} \log(1+u^2)\, du + \frac{1}{2\,(t-1)} \int_0^\infty \frac{1}{1+u^2} \log(t+u^2)\, du$$
$$= \frac{\pi}{2\,(1-t)} \log(2) + \frac{1}{2\,(t-1)} \int_0^\infty \frac{1}{1+u^2} \log(t+u^2)\, du \,, \tag{7.16}$$

where we know

$$\int_0^\infty \frac{1}{1+u^2} \log(1+u^2)\, du = \pi \log(2)$$

from Lemma 7.9.

Rewriting $t^{\mu+k}$

If we look at the simplification for the integral over u we see that we get problems with the denominator $1-t$ if we integrate over t. We therefore use the formula

$$t^{\mu+k} = (t-1)\sum_{h=0}^{\mu+k-1} t^h + 1$$

and can write by using the result (7.16)

$$\begin{aligned}
&\int_0^1 t^{\mu+k} \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \,\arctan u \, du\, dt \\
&\quad = \sum_{h=0}^{\mu+k-1} \int_0^1 t^h\,(t-1) \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \,\arctan u \, du\, dt \\
&\quad + \int_0^1 \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \,\arctan u \, du\, dt \\
&\quad = -\frac{\pi}{2}\,\log(2) \sum_{h=0}^{\mu+k-1} \int_0^1 t^h\, dt + \frac{1}{2}\sum_{h=0}^{\mu+k-1} \int_0^1 t^h \int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\, du\, dt \\
&\quad + \int_0^\infty \frac{u}{(1+u^2)} \,\arctan u \int_0^1 \frac{1}{(t+u^2)}\, dt\, du \\
&\quad = -\frac{\pi}{2}\,\log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} + \frac{1}{2}\sum_{h=1}^{\mu+k} \int_0^1 t^{h-1} \int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\, du\, dt \\
&\quad + \int_0^\infty \frac{u}{(1+u^2)}\,\log\bigl(1+\frac{1}{u^2}\bigr)\,\arctan u\, du\,.
\end{aligned}$$

Solution of $\int_0^1 t^{h-1}\int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\,du\,dt$

To solve the integral ($h = 1,\ldots,3\mu$)

$$\int_0^1 t^{h-1} \int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\, du\, dt$$

we do once integration by parts for the integral over t

$$\int_0^1 t^{h-1}\,\log(t+u^2)\, dt = \frac{1}{h}\,\log(1+u^2) - \frac{1}{h}\int_0^1 \frac{t^h}{t+u^2}\, dt$$

before we try again to solve the integral over u:

$$\begin{aligned}
&\int_0^1 t^{h-1} \int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\, du\, dt \\
&\quad = \frac{1}{h}\int_0^\infty \frac{1}{1+u^2}\,\log(1+u^2)\, du - \frac{1}{h}\int_0^1 t^h \int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)}\, du\, dt \\
&\quad = \frac{\pi}{h}\,\log(2) - \frac{1}{h}\int_0^1 t^h \int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)}\, du\, dt\,,
\end{aligned}$$

where we used again

$$\int_0^\infty \frac{1}{1+u^2}\,\log(1+u^2)\,du = \pi\,\log(2)$$

from Lemma 7.9. To solve

$$\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)}\,du$$

we write the fraction as

$$\begin{aligned}\frac{1}{(1+u^2)\,(t+u^2)} &= \frac{1}{t-1}\,\frac{(t+u^2)-(1+u^2)}{(1+u^2)\,(t+u^2)}\\ &= \frac{1}{t-1}\left(\frac{1}{1+u^2}-\frac{1}{t+u^2}\right)\end{aligned}$$

and get

$$\begin{aligned}\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)}\,du &= \frac{1}{t-1}\left(\int_0^\infty \frac{1}{1+u^2}\,du - \int_0^\infty \frac{1}{t+u^2}\,du\right)\\ &= \frac{1}{t-1}\left(\arctan u - \frac{1}{\sqrt{t}}\,\arctan\Big(\frac{u}{\sqrt{t}}\Big)\right)\Big|_{u=0}^{\infty}\\ &= \frac{1}{t-1}\,\frac{\pi}{2}\left(1-\frac{1}{\sqrt{t}}\right)\\ &= \frac{\pi}{2\,(\sqrt{t}+t)}\,.\end{aligned}$$

So it only remains to solve

$$\int_0^1 \frac{t^h}{\sqrt{t}+t}\,dt\,.$$

Using the substitution $z := \sqrt{t}$ we get

$$\int_0^1 \frac{t^h}{\sqrt{t}+t}\,dt = 2\int_0^1 \frac{z^{2h}}{z+1}\,dz = 2\sum_{l=1}^{2h}\frac{(-1)^l}{l} + 2\,\log(2)\,.$$

The whole integral becomes

$$\begin{aligned}&\int_0^1 t^{h-1}\int_0^\infty \frac{1}{1+u^2}\,\log(t+u^2)\,du\,dt\\ &\quad= \frac{\pi\,\log(2)}{h} - \frac{1}{h}\int_0^1 t^h\int_0^\infty \frac{1}{(1+u^2)\,(t+u^2)}\,du\,dt\\ &\quad= \frac{\pi\,\log(2)}{h} - \frac{\pi}{2h}\int_0^1 \frac{t^h}{\sqrt{t}+t}\,dt\\ &\quad= -\frac{\pi}{h}\sum_{l=1}^{2h}\frac{(-1)^l}{l}\,.\end{aligned}$$

Total solution

With the value

$$\int_0^\infty \frac{u}{(1+u^2)} \log\Big(1+\frac{1}{u^2}\Big) \arctan u \, du = \frac{\pi}{2}\left(\frac{\pi^2}{12} - \log^2(2)\right),$$

as shown in Lemma 7.10, the solution of the whole integral is

$$\begin{aligned}
&\int_0^1 t^{\mu+k} \int_0^\infty \frac{u}{(1+u^2)\,(t+u^2)} \arctan u \, du \, dt \\
&= -\frac{\pi}{2} \log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} - \frac{\pi}{2} \sum_{h=1}^{\mu+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} + \frac{\pi^3}{24} - \frac{\pi}{2} \log^2(2) \\
&= \frac{\pi}{2}\left(\frac{\pi^2}{12} - \log^2(2) - \log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} - \sum_{h=1}^{\mu+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l}\right).
\end{aligned}$$

7.5.9 Further simplifications

The two summands can be recombined by rewriting the sum over binomial coefficients and using Lemma 7.4. We have

$$\begin{aligned}
&\sum_{k=0}^{2\mu} \frac{(-1)^k}{(\mu+k)\,(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&= \sum_{k=0}^{2\mu} \frac{(-1)^k\,(2\mu+2k+1)}{(\mu+k)\,(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&\quad - \sum_{k=0}^{2\mu} \frac{(-1)^k\,(2\mu+2k)}{(\mu+k)\,(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&= \sum_{k=0}^{2\mu} \frac{(-1)^k}{(\mu+k)} \binom{2\mu}{k}\binom{2\mu+k}{k} - 2 \sum_{k=0}^{2\mu} \frac{(-1)^k}{(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&= -2 \sum_{k=0}^{2\mu} \frac{(-1)^k}{(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k},
\end{aligned}$$

where we used Lemma 7.4 with $\gamma = 0$ for the last step.

7.5.10 Solution for the asymptotic variance for odd ν

If we insert the solution developed in Section 7.5.8 into formula (7.15) we get

$$\begin{aligned}
\sigma_\tau^2 = (-1)^{\mu-1} \frac{4\,\pi\,\gamma_\nu}{3} \sum_{k=0}^{2\mu} \frac{(-1)^k}{(\mu+k)\,(2\mu+2k+1)} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
+ (-1)^{\mu-1} \frac{32\,\gamma_\nu}{\pi} \sum_{k=0}^{2\mu} \frac{(-1)^k}{2\mu+2k+1} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
\times \left(\frac{\pi^2}{12} - \log^2(2) - \log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} - \sum_{h=1}^{\mu+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} \right).
\end{aligned}$$

By Section 7.5.9 we can recombine the two summands and get

$$\begin{aligned}
\sigma_\tau^2 &= (-1)^{\mu-1} \frac{32\,\gamma_\nu}{\pi} \sum_{k=0}^{2\mu} \frac{(-1)^k}{2\mu+2k+1} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&\quad \times \left(-\frac{\pi^2}{12} + \frac{\pi^2}{12} - \log^2(2) - \log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} - \sum_{h=1}^{\mu+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} \right) \\
&= (-1)^{\mu-1} \frac{32\,\gamma_\nu}{\pi} \sum_{k=0}^{2\mu} \frac{(-1)^k}{2\mu+2k+1} \binom{2\mu}{k}\binom{2\mu+k}{k} \\
&\quad \times \left(-\sum_{h=1}^{\mu+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} - \log(2) \sum_{h=1}^{\mu+k} \frac{1}{h} - \log^2(2) \right).
\end{aligned}$$

We can now insert the definition $\mu = \frac{\nu-1}{2}$:

$$\begin{aligned}
\sigma_\tau^2 = (-1)^{\frac{\nu+1}{2}} \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi\,\Gamma^3(\frac{\nu}{2})} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k} \binom{\nu-1}{k}\binom{\nu+k-1}{k} \\
\times \left(-\sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} - \log(2) \sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h} - \log^2(2) \right).
\end{aligned}$$

Lemma 7.5 tells us the coefficient of $\log^2(2)$ which brings us to the final result

$$\begin{aligned}
\sigma_\tau^2 = \frac{16}{\pi^2} \log^2(2) + (-1)^{\frac{\nu-1}{2}} \frac{32\,\Gamma(\frac{3\nu}{2})}{\pi\,\Gamma^3(\frac{\nu}{2})} \sum_{k=0}^{\nu-1} \frac{(-1)^k}{\nu+2k} \binom{\nu-1}{k}\binom{\nu+k-1}{k} \\
\times \left(\sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h} \sum_{l=1}^{2h} \frac{(-1)^l}{l} + \log(2) \sum_{h=1}^{\frac{\nu-1}{2}+k} \frac{1}{h} \right).
\end{aligned}$$

Bibliography

Mir M. Ali, N. N. Mikhail, and M. Safiul Haq. A class of bivariate distributions including the bivariate logistic. *Journal of Multivariate Analysis*, 8:405–412, 1978.

Herbert Amann and Joachim Escher. *Analysis II*. Birkhäuser, Basel, 1999.

David F. Andrews and Colin L. Mallows. Scale mixtures of normal distributions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(1):99–102, 1974.

Wilfrid N. Bailey. *Generalized hypergeometric series*, volume 32 of *Cambridge Tracts in Mathematics and Mathematical Physics*. Stechert-Hafner Service Agency, New York, 1964.

Evelyn Martin L. Beale and Colin L. Mallows. Scale mixing of symmetric distributions with zero means. *The Annals of Mathematical Statistics*, 1959.

Vidmantas Bentkus, Friedrich Götze, and Ričardas Zitkis. Lower estimates of the convergence rate for U-statistics. *The Annals of Probability*, 22(4):1707–1714, 1994.

Vidmantas Bentkus, Bing-Yi Jing, and Wang Zhou. On normal approximations to U-statistics. *The Annals of Probability*, 37(6), 2009.

Andrew C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. *Transactions of the American Mathematical Society*, 1941.

Peter J. Bickel. Edgeworth expansions in nonparametric statistics. *The Annals of Statistics*, 2(1):1–20, 1974.

Patrick Billingsley. *Convergence of probability measures*. Wiley series in probability and mathematical statistics. John Wiley & Sons, New York, 1968.

Yuri V. Borovskikh. *U-Statistics in Banach spaces*. VSP, Utrecht, 1996.

Stamatis Cambanis, Steel Huang, and Gordon Simons. On the theory of elliptically contoured distributions. *Journal of Multivariate Analysis*, 11:368–385, 1981.

Carles M. Cuadras and Joan Augé. A continuous general multivariate distribution and its properties. *Communications in Statistics: Theory and Methods*, 10:339–353, 1981.

Paul Embrechts, Alexander McNeil, and Daniel Straumann. Correlation: Pitfalls and alternatives. *RISK Magazine*, 1999.

Paul Embrechts, Filip Lindskog, and Alexander McNeil. Modelling dependence with copulas and applications to risk management. In S.T. Rachev, editor, *Handbook of heavy tailed distributions in finance*, chapter 8, pages 329–384. Elsevier, Amsterdam, 2003.

Carl-Gustav Esseen. On the Liapounoff limit of error in the theory of probability. *Arkiv för Matematik, Astronomi och Fysik*, 28A(9):1–19, 1942.

Carl-Gustav Esseen. Fourier analysis of distribution functions. a mathematical study of the Laplace-Gaussian law. *Acta Mathematica*, 1945.

Carl-Gustav Esseen. A moment inequality with an application to the central limit theorem. *Skandinavisk Aktuarietidskrift*, 39:160–170, 1956.

Hong-Bin Fang, Kai-Tai Fang, and Samuel Kotz. The meta-elliptical distributions with given marginals. *Journal of Multivariate Analysis*, 82(1):1–16, 2002.

Kai-Tai Fang, Samuel Kotz, and Kai Wang Ng. *Symmetric multivariate and related distributions*. Chapman & Hall, London, 1990.

William Feller. *An introduction to probability theory and its applications, volume II*. Wiley series in probability and mathematical statistics. John Wiley & Sons, New York, second edition, 1971.

Eberhard Freitag and Rolf Busam. *Funktionentheorie*. Springer, Berlin, 1993.

Karl O. Friedrich. A Berry–Esseen bound for functions of independent random variables. *The Annals of Statistics*, 17(1):170–183, 1989.

Peter Gänssler and Winfried Stute. *Wahrscheinlichkeitstheorie*. Springer-Verlag, Berlin, 1977.

George Gasper and Mizan Rahman. *Basic hypergeometric series*, volume 35 of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press, Cambridge, 1990.

Ronald L. Graham, Donald E. Knuth, and Oren Patashnik. *Concrete mathematics: a foundation for computer science*. Addison-Wesley, Reading, second edition, 1994.

Arjun K. Gupta and Tamas Varga. *Elliptically contoured models in statistics*, volume 240 of *Mathematics and Its Applications*. Kluwer Academic Publisher, Dordrecht, 1993.

Paul R. Halmos. The theory of unbiased estimation. *The Annals of Mathematical Statistics*, 1946.

Joachim Hartung, Bärbel Elpelt, and Karl-Heinz Klösener. *Statistik: Lehr- und Handbuch der angewandten Statistik*. Oldenburg Verlag, München, 2005.

Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *The Annals of Mathematical Statistics*, 1948.

Wassily Hoeffding. The strong law of large numbers for U-statistics. *Institute of Statistics Mimeo Series*, 1961.

Olav Kallenberg. *Foundations of modern probability*. Probability and its applications. Springer, New York, second edition, 2002.

Claudia Klüppelberg and Gabriel Kuhn. Copula structure analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(3):737–753, 2009.

Volodymyr S. Koroljuk and Yuri V. Borovskich. *Theory of U-statistics*, volume 273 of *Mathematics and Its Applications*. Kluwer Academic Publisher, Dordrecht, 1994.

Volodymyr S. Korolyuk and Yuri V. Borovskikh. Approximation of nondegenerate U-statistics. *Theory of Probability and its Applications*, 30(3):439–450, 1985.

Gabriel Kuhn. *On dependence and extremes*. PhD thesis, Munich University of Technology, 2006.

A. J. Lee. *U-statistics: theory and practice*, volume 110 of *Statistics: Textbooks and Monographs*. Marcel Dekker, New York, 1990.

Erich L. Lehmann. *Elements of Large-Sample Theory*. Springer, New York, 1999.

Erich L. Lehmann and George Casella. *Theory of point estimation*. Springer, New York, second edition, 1998.

Filip Lindskog, Alexander McNeil, and Uwe Schmock. Kendall's tau for elliptical distributions. In Georg Bol et al., editor, *Credit risk measurement, evaluation and management*, page 149. Physica-Verlag, Heidelberg, 2003.

Albert W. Marshall and Ingram Olkin. A multivariate exponential distribution. *Journal of the American Statistical Association*, 1967.

Alexander J. McNeil, Rüdiger Frey, and Paul Embrechts. *Quantitative risk management: concepts, techniques and tools*. Princeton University Press, Princeton, 2005.

Roger B. Nelsen. *An introduction to copulas.* Springer series in statistics. Springer, New York, second edition, 2006.

Anatolii P. Prudnikov, Yuri A. Brychkov, and Oleg I. Marichev. *Integrals and series, volume 1: elementary functions.* Gordon and Breach Schience Publishers, New York, 1992.

Herbert Robbins. Mixture of distributions. *The Annals of Mathematical Statistics*, 1948.

Arthur Wayne Roberts and Dale Elthon Varberg. *Convex functions.* Academic Press, New York, 1973.

Walter Rudin. *Real and complex analysis.* McGraw-Hill, New York, third edition, 1987.

Irina G. Shevtsova. Sharpening of the upper bound of the absolute constant in the Berry–Esseen inequality. *Theory of Probability and its Applications*, 51(3):549–553, 2007.

I. S. Shiganov. Refinement of the upper bound of the constant in the central limit theorem. *Journal of Mathematical Sciences*, 1986.

Paul van Beek. An application of Fourier methods to the problem of sharpening the Berry–Esseen inequality. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1972.

David Williams. *Probability with martingales.* Cambridge University Press, Cambridge, 1991.

Hermann Witting and Ulrich Müller-Funk. *Mathematische Statistik II.* Teubner, Stuttgart, 1995.

Barbara Dengler

Curriculum Vitae

Krongasse 4/5
1050 Vienna, Austria
☎ *+43 1 58801 10532 (office)*
✉ *bdengler@fam.tuwien.ac.at*

Personal Data

Date and place of birth	August 13, 1981 in Munich
Nationality	German

Education

March 2006–present — **PhD in Financial Mathematics**, *Vienna University of Technology*.
Supervisor: Prof. Uwe Schmock

Oct. 2000–Dec. 2005 — **Studies of Financial Mathematics and Actuarial Sciences**, *Ludwig Maximilian University Munich*.

- Diploma in Dec. 2005, grade: very good.
- Diploma thesis: Darstellung von Copulae zur Beschreibung der Abhängigkeitsstruktur versicherungsbetrieblicher Risiken und systematischer Vergleich mit skalaren Abhängigkeitsmaßen
- Bachelor in Oct. 2002, grade: very good

June 2000 — **Abitur**, *Gymnasium Kirchheim*.
grade: 1,0

Professional Experience

Feb. 2010–present — **Graduate Teaching Assistant / Assistant Professor**, *Vienna University of Technology, Research Group Financial and Actuarial Mathematics*.
Teaching: Advanced life insurance mathematics

March 2006–present — **Research Assistant**, *Vienna University of Technology, Research Group Financial and Actuarial Mathematics*.

- Project: Christian Doppler Laboratory for Portfolio Risk Management;
  Module: Dependence modelling for pricing and risk management
- Teaching: - Mathematical finance: discrete-time models (Exercises)
  - Life insurance mathematics (Exercises)

2003–2005 — **Actuarial Analyst**, *Allianz, Health Insurance Company, Munich*.
Internship and following part time work