Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 0946 – 8633

On multivariate chi-square distributions and their applications in testing multiple hypotheses

Thorsten Dickhaus[1], Thomas Royen[2]

submitted: January 06, 2014

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
E-Mail: Thorsten.Dickhaus@wias-berlin.de

[2] University of Applied Sciences Bingen
Berlinstr. 109
55411 Bingen am Rhein
E-Mail: Thomas.Royen@t-online.de

No. 1913
Berlin 2014

2010 *Mathematics Subject Classification.* 62J15, 62E15.

Key words and phrases. Contingency tables, Kruskal-Wallis test, multiple Wald tests, multivariate analysis, multivariate gamma distributions, statistical genetics.

We thank Yuriy Kopanskyy and Jens Stange for help with LaTeX and computer simulations.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We are considered with three different types of multivariate chi-square distributions. Their members play important roles as limiting distributions of vectors of test statistics in several applications of multiple hypotheses testing. We explain these applications and provide formulas for computing multiplicity-adjusted p-values under the respective global hypothesis.

1 Introduction

Chi-square distributions play an important role in many areas of inferential statistics, at least for two reasons. First, chi-square distributions on $[0, \infty)$ are (limiting) distributions of quadratic forms and thus they occur naturally in the context of distance-based statistical methods in Euclidean geometry. One example of this type of application is the chi-square test of goodness-of-fit, where squared distances of observed and expected counts are evaluated in order to test the empirical distribution of a data sample against a given one. Second, likelihood ratio statistics in parametric models are under regularity conditions asymptotically chi-square distributed according to Wilks (1938), at least if nested models are considered.

In many modern application fields, however, several statistical hypotheses have to be tested simultaneously based on the same sample. This is typically referred to as a multiple test problem. For instance, one may want to test which genes from a potentially large list of candidates are associated with a clinically relevant outcome, or one may want to test which vareties of a certain agricultural product have (on average) the largest gross yield per unit. In such situations, typically a vector of test statistics is constructed, where every component corresponds to one (marginal) test problem; we will provide more details in Section 2.1. Hence, under the multiple testing framework, often vectors of chi-square distributed statistics are objects of interest. Particularly relevant cases occur if these marginal chi-square statistics exhibit certain dependencies, leading to the consideration of multivariate chi-square distributions.

In this work, we will be considered with different types of such multivariate chi-square distributions. We will explain typical multiple test problems in which these distributions play a role as (limiting) distributions of vectors of test statistics, and we will discuss methods for computing multivariate chi-square probabilities and, consequently, multiplicity-adjusted p-values corresponding to such multivariate chi-square distributed vectors of test statistics. The paper is organized as follows. After the preliminary Section 2, we study nonparametric rank-based multiple comparisons in Section 3, by exploiting permutational multivariate central limit theorems. Section 4 deals with the simultaneous analysis of several contingency tables, where data from different tables are dependent. This has important applications in statistical genetics, in particular in genetic association studies. In Section 5, multiple Wald tests (or asymptotically equivalently, multiple likelihood ratio tests) for dependent endpoints are considered. Section 6 summarizes computational methods for different types of multivariate chi-square distributions, and we conclude with a discussion in Section 7.

2 Preliminaries

2.1 Multiple hypotheses testing

The general setup of multiple testing theory assumes a statistical model $(\Omega, \mathcal{F}, (\mathbb{P}_\vartheta)_{\vartheta \in \Theta})$ parametrized by $\vartheta \in \Theta$ and is concerned with testing a family $\mathcal{H} = (H_i : i \in I)$ of hypotheses regarding the parameter ϑ with corresponding alternatives $K_i = \Theta \setminus H_i$, where I denotes an arbitrary index set. In the applications treated in this paper, I will be of finite cardinality m (say), such that $I = \{1, \ldots, m\}$ may be assumed without loss of generality. We identify hypotheses with subsets of the parameter space throughout the paper. Let $\varphi = (\varphi_i : i \in I)$ be a multiple test procedure for $\mathcal{H}$, meaning that each component φ_i, $i \in I$, is a (non-randomized) test for the test problem H_i versus K_i in the classical sense.

We restrict our attention to multiple tests which are defined via a family $(T_i, i \in I)$ of test statistics, where each $T_i : \Omega \to \mathbb{R}$ is a measurable mapping. We assume that each T_i tends to larger values under the respective alternative K_i. Thus, the marginal test φ_i is of the form $\varphi_i(x) = 1 \iff T_i(x) > c_i$, where the critical values c_i, $i \in I$, have to be chosen to ensure (type I) error control of given form at given level. More specifically, let $I_0 \equiv I_0(\vartheta) \subseteq I$ denote the index set of true hypotheses in $\mathcal{H}$ and $V(\varphi)$ the number of false rejections (type I errors) of φ, i. e., $V(\varphi) = \sum_{i \in I_0} \varphi_i$. The classical multiple type I error measure in multiple hypothesis testing is the family-wise error rate, FWER for short, and can (for a given $\vartheta \in \Theta$) be expressed as $\text{FWER}_\vartheta(\varphi) = \mathbb{P}_\vartheta(V(\varphi) > 0)$. The multiple test φ is said to control the FWER in the strong sense at a pre-defined significance level α, if $\sup_{\vartheta \in \Theta} \text{FWER}_\vartheta(\varphi) \leq \alpha$. In terms of the joint distribution of test statistics and the critical values, we can write

$$\text{FWER}_\vartheta(\varphi) = \mathbb{P}_\vartheta\left(\bigcup_{i \in I_0} \{T_i > c_i\}\right) = 1 - \mathbb{P}_\vartheta\left(\bigcap_{i \in I_0} \{T_i \leq c_i\}\right),$$

showing that suitable critical values c_i, $i \in I$, are given by quantiles of the joint distribution of test statistics. Let $H_0 = \bigcap_{i \in I} H_i$ denote the global (intersection) hypothesis of $\mathcal{H}$, which is assumed non-empty throughout the remainder. Often, FWER control under H_0 (in the weak sense) entails FWER control in the strong sense, namely, if parameter values in H_0 are least favorable configurations (LFCs) for the FWER of φ. Sufficient conditions for the latter have been provided by Gabriel (1969) and Dickhaus and Stange (2013), among others. In Sections 3 - 5, we will derive the joint distributions of test statistics only under H_0. Even if parameters in H_0 are not LFCs, this joint distribution under H_0 can straightforwardly be employed for strong FWER control by utilizing the closed test principle (cf. Marcus et al. (1976)), provided that the family $(T_i, i \in I)$ is joint in the sense of Gabriel (1969). Letting $t_i = T_i(x)$, $1 \leq i \leq m$, $x \in \Omega$, denote the actually observed values of the test statistics, multiplicity-adjusted p-values are given by

$$p_i(x) = \mathbb{P}_\vartheta\left(\max_{1\le j\le m} T_j > t_i\right), \quad 1 \le i \le m, \tag{1}$$

where we will again restrict our attention to parameter values $\vartheta \in H_0$ in (1).

2.2 Types of multivariate chi-square distributions

There is no general definition of a p-variate chi-square distribution, because there exist a variety of different ways in which univariate chi-square distributions can be combined by a copula. Although the stochastic representation of the resulting random vector $\mathbf{Q} = (Q_1, \ldots, Q_p)^\top$ (of any of the types considered in this work) only involves standard normal random variables (and can therefore be simulated straightforwardly), its (joint) distribution is not determined by the marginal degrees of freedom and the correlation matrix of the components $Q_1, \ldots, Q_p$, as analytically proved by Krishnaiah and Rao (1961). To illustrate this fact, let us consider the following elementary counterexample.

Example 2.1.

(i) Consider four real-valued, centered random variables $Z_{1,1}, Z_{1,2}, Z_{2,1}, Z_{2,2}$ having a joint normal distribution on $\mathbb{R}^4$, where each $Z_{i,j}$ has unit variance. Let $\rho(Z_{i,j}, Z_{k,\ell})$ denote Pearson's correlation coefficient of $Z_{i,j}$ and $Z_{k,\ell}$ for $1 \le i, j, k, \ell \le 2$ and assume that only the correlations $\rho(Z_{1,1}, Z_{2,1}) = \rho(Z_{1,2}, Z_{2,2}) = \rho$ are non-zero. Then, the random vector $\mathbf{Q} = (Q_1, Q_2)^\top$, given by $Q_1 = Z_{1,1}^2 + Z_{1,2}^2$ and $Q_2 = Z_{2,1}^2 + Z_{2,2}^2$, follows a bivariate chi-square distribution with two degrees of freedom in both marginals and with Cov$(Q_1, Q_2) = 4\rho^2$.

(ii) Consider three real-valued, independent and identically distributed (iid.) random variables Z_0, Z_1, Z_2, where Z_0 has the standard normal distribution on $\mathbb{R}$. Let the random vector $\tilde{\mathbf{Q}} = (\tilde{Q}_1, \tilde{Q}_2)^\top$ be given by $\tilde{Q}_1 = Z_1^2 + Z_0^2$ and $\tilde{Q}_2 = Z_2^2 + Z_0^2$. Then, $\tilde{\mathbf{Q}}$ also follows a bivariate chi-square distribution with two degrees of freedom in both marginals, and it holds Cov$(\tilde{Q}_1, \tilde{Q}_2) =$ Var$(Z_0^2) = 2$.

Letting $\rho^2 = 1/2$ in part (i), we obtain that Cov$(Q_1, Q_2) =$ Cov$(\tilde{Q}_1, \tilde{Q}_2) = 2$. However, even for this choice of ρ^2 the joint distributions of $\mathbf{Q}$ and $\tilde{\mathbf{Q}}$ do not coincide, as can be seen by comparing their Laplace transforms (Lt), which are given by $Lt_{(Q_1,Q_2)}(t_1, t_2) = [(1 + 2t_1)(1 + 2t_2) - 4\rho^2 t_1 t_2)]^{-1} = [1 + 2(t_1 + t_2 + t_1 t_2)]^{-1}$ in case of $\rho^2 = 1/2$, as well as $Lt_{(\tilde{Q}_1,\tilde{Q}_2)}(t_1, t_2) = [(1+2(t_1+t_2))(1+2t_1)(1+2t_2)]^{-1/2}$. In view of testing two hypotheses, notice that, assuming $\rho^2 = 1/2$, the equi-coordinate 95%-quantile of $\mathbf{Q}$ approximately equals 7.0802, while that of $\tilde{\mathbf{Q}}$ approximately equals 6.9776.

Example 2.1 demonstrates that the full joint stochastic representation of a multivariate chi-square distributed random vector is needed in order to compute its quantiles (and, hence, to calibrate associated multivariate multiple tests). For different statistical models and associated families of hypotheses that we are going to discuss in the following sections, these stochastic representations differ, giving rise to different types of multivariate chi-square distributions.

The “classical“ definition of a p-variate chi-square distribution is provided for instance by Timm (2002), see his Definition 3.5.7. It is the joint distribution of the diagonal elements of a Wishart-distributed random matrix $S \sim W_p(\nu, \Sigma)$. Therefore, we term this distribution a p-variate chi-square distribution of Wishart-type or a “Wishart-chi-square distribution“ for short. Its standardized form results whenever Σ is a correlation matrix, meaning that all its diagonal elements are equal to one. The distribution of $\mathbf{Q}$ in part (i) of Example 2.1 is a standardized bivariate Wishart-chi-square distribution. Wishart-chi-square distributions straightforwardly arise if a family of point hypotheses regarding the marginal variances of a multivariate normal distribution shall be tested based on an iid. sample; cf. Example 2.2. Multivariate chi-square distributions of Wishart-type will also play an important role in Examples 2.3 and 2.4, as well as in Section 3. Two different generalizations of Wishart-chi-square distributions have been introduced by Jensen (1970) (cf. our Section 4 for an application) and Dickhaus (2012) (see our Definition 5.1). Due to their relevance in many multiple test problems, these three types of multivariate chi-square distributions are the subject of this work.

On the other hand, the distribution of $\tilde{\mathbf{Q}}$ in part (ii) of Example 2.1 was considered by Simes (1986) and Sarkar and Chang (1997) in connection with the validity of Simes' global test under positive dependency, but is of none of the aforementioned three types. Further types of multivariate chi-square distributions, which are also not considered in this work, are compiled in Chapter 48 of Kotz et al. (2000).

2.3 Some first examples

Example 2.2 (Multiple tests of Gaussian variances)**.** *Assume that one can observe iid. random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_N$, where $\mathbf{X}_1 \sim \mathcal{N}_m(\mu, \Sigma)$. Consider the case that the mean vector μ and the diagonal elements of the covariance matrix $\Sigma = (\sigma_{ij})_{1 \leq i,j \leq m}$ of each of the observables are unknown. Assuming that $m \geq 2$ and $N > m$, a suitable estimator of Σ is the empirical covariance matrix S, given by $S = (N-1)^{-1} \sum_{1 \leq i \leq N} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\top$, where $\bar{\mathbf{X}}$ is the vector of component-wise arithmetic means. It is well-known (see, for instance, Corollary 7.2.3. of Anderson (1984)) that the distribution of S is a Wishart distribution with mean Σ and $N-1$ degrees of freedom. Consequently, the joint distribution of the diagonal elements $(S_{ii} : 1 \leq i \leq m)$ is an m-variate chi-square distribution of Wishart-type which is scaled by the (unknown) diagonal elements σ_{ii} for $1 \leq i \leq m$. Now, consider the system $\mathcal{H} = (H_i : 1 \leq i \leq m)$ of hypotheses, where $H_i : \{\sigma_{ii} = \sigma_{ii}^*\}$ for fixed, given constants σ_{ii}^*, $1 \leq i \leq m$, with one-sided alternatives $K_i : \{\sigma_{ii} > \sigma_{ii}^*\}$ (the one-sided alternative $\{\sigma_{ii} < \sigma_{ii}^*\}$ can be treated analogously and the two-sided alternative $\{\sigma_{ii} \neq \sigma_{ii}^*\}$ can be represented by two one-sided ones). A suitable vector of test statistics is then given by $(S_{11}/\sigma_{ii}^*, \ldots, S_{mm}/\sigma_{mm}^*)^\top$, and the respective m-variate chi-square distribution of Wishart-type under the global hypothesis $H_0 = \bigcap_{1 \leq i \leq m} H_i$ can be used for calibrating a multivariate multiple test procedure for testing $\mathcal{H}$.*

Example 2.3 (Multiple comparisons of multinomial probabilities)**.** *Consider k subpopulations Ω_i, $1 \leq i \leq k$, and assume that one can observe stochastically independent vectors $\mathbf{N}_i \sim \mathcal{M}_c(n_i, \vec{p}_i)$, $1 \leq i \leq k$, where $\mathcal{M}_c(n, \vec{p})$ denotes the multinomial distribution with c categories, sample size n and vector of probabilities $\vec{p}$. We assume that the sample sizes n_i are given constants and that the vectors $\vec{p}_i = (p_{i\ell} : 1 \leq \ell \leq c)^\top$ are unknown. Royen (1984) was* *considered with the problem of multiple comparisons with a control group, i. e., the system $\mathcal{H}^{\text{Dunnett}} = (H_i : 1 \leq i \leq k-1)$, where $H_i : \{\vec{p}_i = \vec{p}_k\}$ (group k is the control group), with two-sided alternatives $K_i : \{\vec{p}_i \neq \vec{p}_k\}$. A suitable test statistic for testing H_i is given by*

$$T_i = \sum_{\ell=1}^{c} \left\{ \frac{(N_{i\ell} - E_{i\ell})^2}{E_{i\ell}} + \frac{(N_{k\ell} - E_{k\ell})^2}{E_{k\ell}} \right\}, \quad E_{\gamma\ell} = \frac{n_\gamma (N_{i\ell} + N_{k\ell})}{n_i + n_k}, \quad \gamma \in \{i, k\}. \quad (2)$$

Under H_i, T_i is asymptotically (as $\min\{n_i : 1 \leq i \leq k\} \to \infty$) chi-square distributed on $[0, \infty)$ with $c-1$ degrees of freedom. However, the $(T_i : 1 \leq i \leq k-1)$ are correlated, because data from group k are used in all T_i. Let $\mathbb{P}_0$ denote any probability measure on $\Omega = \bigcup_{i=1}^{k} \Omega_i$ such that the global hypothesis $H_0 = \bigcap_{1 \leq i \leq k-1} H_i$ is true. Noticing that $\mathbb{P}_0(\forall 1 \leq i \leq k-1 : T_i \leq c_\alpha) = \mathbb{P}_0(\max_{1 \leq i \leq k-1} T_i \leq c_\alpha)$, the critical value c_α for calibrating a simultaneous test procedure in the sense of Gabriel (1969) for FWER control at level α can be chosen as a quantile of the distribution of $\max_{1 \leq i \leq k-1} T_i$ under H_0 (this distribution is invariant with respect to the parameter values in H_0). The latter distribution is a $(k-1)$-variate chi-square distribution of Wishart-type with $c-1$ marginal degrees of freedom and an associated one-factorial correlation matrix (see Section 6 for details) which only depends on the given sample sizes n_i for $1 \leq i \leq k$. The cdf. of this distribution has been computed by Royen (1984). We may also mention here that the cdf. of non-central multivariate chi-square distributions of Wishart-type with associated one-factorial correlation matrices has been computed by Royen (1995); see also Royen (1997). These non-central distributions play an important role in connection with power considerations for the simultaneous test procedure.

Furthermore, Royen (1984) also considered the problem of all pairwise group comparisons, i. e., the system $\mathcal{H}^{\text{Tukey}} = (H_{ij} : 1 \leq i < j \leq k)$, where $H_{ij} : \{\vec{p}_i = \vec{p}_j\}$ with two-sided alternatives $K_{ij} : \{\vec{p}_i \neq \vec{p}_j\}$. A suitable test statistic for testing H_{ij} is given by

$$T_{ij} = \sum_{\ell=1}^{c} \left\{ \frac{(N_{i\ell} - E_{i\ell})^2}{E_{i\ell}} + \frac{(N_{j\ell} - E_{j\ell})^2}{E_{j\ell}} \right\}, \quad E_{\gamma\ell} = \frac{n_\gamma (N_{i\ell} + N_{j\ell})}{n_i + n_j}, \quad \gamma \in \{i, j\}. \quad (3)$$

The distribution of $\max_{1 \leq i < j \leq k} T_{ij}$ under $\bigcap_{1 \leq i < j \leq k} H_{ij}$ is a generalization of the Gaussian range distribution to Gaussian random vectors and quantiles of it have been tabulated by Royen (1989, 1990).

Example 2.4 (Multiple comparisons of vectors of regression coefficients)**.** *Consider $k \geq 3$ stochastically independent response vectors $\mathbf{Y}_i$, $1 \leq i \leq k$, with values in $\mathbb{R}^n$ each, where $n \in \mathbb{N}$ denotes the common sample size in every group i. For each $1 \leq i \leq k$, assume a linear model for $\mathbf{Y}_i$ of the form*

$$\mathbf{Y}_i = X\beta_i + \varepsilon_i, \quad (4)$$

where the given $(n \times p)$ design matrix X is the same for all k groups and is assumed to have rank $p < n$. Royen (1995), Section 6, was considered with multiple comparisons regarding the vectors $\beta_i : 1 \leq i \leq k$. Differences between these vectors indicate that the influence of the covariates encoded by X on the (mean) response is different across groups. Assuming that all vectors ε_i, $1 \leq i \leq k$, of error terms are identically distributed as $\mathcal{N}_n(0, \sigma^2 I_n)$, the (in general unknown) error variance σ^2 can be estimated by the pooled estimator S^2 with $\nu = k(n-p)$

degrees of freedom, and it holds $\nu S^2/\sigma^2 \sim \chi^2_\nu$. Now, consider for illustration the problem of multiple comparisons with control group k, i. e., the system of hypotheses $\mathcal{H}^{\text{Dunnett}} = (H_i : 1 \leq i \leq k-1)$, where $H_i : \{\beta_i = \beta_k\}$. The usual least squares estimator of β_i is given by $\hat{\beta}_i = (X^\top X)^{-1} X^\top \mathbf{Y}_i$, leading to $\hat{\beta}_i - \hat{\beta}_k = (X^\top X)^{-1} X^\top (\mathbf{Y}_i - \mathbf{Y}_k)$ with covariance matrix Cov$(\hat{\beta}_i - \hat{\beta}_k) = 2\sigma^2 (X^\top X)^{-1}$. Hence, if σ^2 would be known, the normalized squared difference

$$T_i = \frac{1}{2\sigma^2} (\mathbf{Y}_i - \mathbf{Y}_k)^\top X (X^\top X)^{-1} X^\top (\mathbf{Y}_i - \mathbf{Y}_k)$$

would be a suitable test statistic for testing H_i, for $1 \leq i \leq k-1$. The joint distribution of $(T_1, \ldots, T_{k-1})^\top$ under the global hypothesis H_0 is a $(k-1)$-variate chi-square distribution of Wishart-type, cf. Example 2.3. In the practically relevant case of unknown σ^2, Studentization with S leads to the modified test statistics

$$\tilde{T}_i = \frac{1}{2S^2} (\mathbf{Y}_i - \mathbf{Y}_k)^\top X (X^\top X)^{-1} X^\top (\mathbf{Y}_i - \mathbf{Y}_k), \quad 1 \leq i \leq k-1.$$

Up to scaling with the degrees of freedom, the joint distribution of $(\tilde{T}_1, \ldots, \tilde{T}_{k-1})^\top$ under H_0 is a multivariate extension of Fisher's F-distribution the cdf. of which can be obtained by integrating the cdf. of $(T_1, \ldots, T_{k-1})^\top$ with respect to the distribution of S^2/σ^2. This results in a null distribution for the calibration of the multiple test based on $(\tilde{T}_1, \ldots, \tilde{T}_{k-1})^\top$ for FWER control at level α, see the derivations in Section 6 of Royen (1995). There, one can also find asymptotic expansions for large degrees of freedom of S^2 to avoid the additional integration over the density of S^2/σ^2, which can be used in the same way for the central multivariate F-distribution.

3 Multivariate nonparametric multiple comparisons

Puri and Sen (1971), Section 5.4, worked out a multivariate extension of the Kruskal-Wallis test. Assume that one can observe N stochastically independent vectors $\mathbf{X}_1, \ldots, \mathbf{X}_N$, each of which takes values in $\mathbb{R}^p$, and that the corresponding N observational units belong to c distinct groups, where n_k denotes the sample size for group $1 \leq k \leq c$ such that $N = \sum_{k=1}^c n_k$. The (global) null hypothesis H_0 (say) that Puri and Sen (1971) were concerned with states that there are no group differences, i. e., that $\mathbf{X}_1, \ldots, \mathbf{X}_N$ are identically distributed. For testing H_0, they proposed a rank-based method which works as follows. For every coordinate $1 \leq i \leq p$, all N observational units are ranked. Let $R_{i\ell}^{(k)}$ denote the resulting rank of the ℓ-th observational unit within group k in coordinate i, where $1 \leq \ell \leq n_k$. The scaled rank average of group k in coordinate i is given by $T_{N,i}^{(k)} = (N+1)^{-1} n_k^{-1} \sum_{\ell=1}^{n_k} R_{i\ell}^{(k)}$. Notice that one can equivalently write $T_{N,i}^{(k)} = (N+1)^{-1} n_k^{-1} \sum_{\ell=1}^{N} R_{i\ell} \mathbf{1}_{\{\text{group is } k\}}(\ell)$, where the index ℓ now runs over all N observational units and $(R_{i\ell} : 1 \leq \ell \leq N)$ denotes the pooled vector of ranks among all observational units in coordinate i. Both representations immediately entail that $\mathbb{E}_{H_0}[T_{N,i}^{(k)}] = 1/2$. Letting $T_N^{(k)} = (T_{N,1}^{(k)}, \ldots, T_{N,p}^{(k)})$, a reasonable test statistic for testing H_0 is given by

$$\mathcal{L}_N = \sum_{k=1}^{c} n_k \left[\left(T_N^{(k)} - \frac{1}{2}\mathbf{1} \right) V^{-1}(R_N^*) \left(T_N^{(k)} - \frac{1}{2}\mathbf{1} \right)^\top \right],$$

where the $(p \times p)$-matrix $V(R_N^*)$ has elements

$$v_{ii'}(R_N^*) = [N(N+1)^2]^{-1} \sum_{k=1}^{c} \sum_{\ell=1}^{n_k} R_{i\ell}^{(k)} R_{i'\ell}^{(k)} - 1/4$$

and is assumed to be invertible for ease of presentation. It is easy to check that, in the case of $p = 1$, the statistic $\mathcal{L}_N$ equals $NH/(N-1)$, where H is the test statistic for the Kruskal-Wallis test (see Kruskal and Wallis (1952)). Puri and Sen (1971) proved that, under H_0, the permutational distribution of $\mathcal{L}_N$ converges weakly to a chi-squared distribution with $p(c-1)$ degrees of freedom. For their proof, joint asymptotic normality of the statistics $T_{N,i}^{(k)}$ for $1 \le i \le p$ and $1 \le k \le c-1$ under the permutation distribution plays a crucial role. The latter can elegantly be deduced from the considerations by Sen (1983). We let $L_{Nik} = T_{N,i}^{(k)} - 1/2$ and represent this centered (under H_0) statistic as $L_{Nik} = \sum_{\ell=1}^{N} (c_{\ell,k} - \bar{c}_k) R_{i\ell}/(N+1)$ with regression coefficients

$$c_{\ell,k} = n_k^{-1} \mathbf{1}_{\{\text{group is } k\}}(\ell), \ \bar{c}_k = N^{-1} \sum_{\ell=1}^{N} c_{\ell,k} = 1/N,$$

which do not depend on the coordinate i. Following Sen (1983), we obtain that under H_0, it asymptotically (as $\min\{n_k : 1 \le k \le c\} \to \infty$) holds

$$(L_{Nik} : 1 \le i \le p, 1 \le k \le c-1) \sim \mathcal{N}_{p(c-1)}(0, V_N \otimes C_N) \tag{5}$$

for the joint permutational distribution of $(L_{Nik} : 1 \le i \le p, 1 \le k \le c-1)$, where the entries in the matrix $V_N \in \mathbb{R}^{p \times p}$ are given by

$$\begin{aligned} v_{N,i,i'} &= (N-1)^{-1} \sum_{\ell=1}^{N} \left(\frac{R_{i\ell}}{N+1} - \frac{1}{2} \right) \left(\frac{R_{i'\ell}}{N+1} - \frac{1}{2} \right) \\ &= \frac{N}{N-1} \left(\frac{1}{N(N+1)^2} \sum_{\ell=1}^{N} R_{i\ell} R_{i'\ell} - \frac{1}{4} \right), \ i, i' = 1, \ldots, p, \end{aligned}$$

and

$$\begin{aligned} C_N &= \sum_{\ell=1}^{N} \left[(c_{\ell,1}, \ldots, c_{\ell,c-1}) - \left(\frac{1}{N}, \ldots, \frac{1}{N} \right) \right] \left[(c_{\ell,1}, \ldots, c_{\ell,c-1}) - \left(\frac{1}{N}, \ldots, \frac{1}{N} \right) \right]^\top \\ &= \left(\frac{\delta_{k,k'}}{n_k} - \frac{1}{N} \right)_{k,k'=1,\ldots,c-1} \end{aligned} \tag{6}$$

with values in $\mathbb{R}^{(c-1)\times(c-1)}$.

Now, assume that group differences (if any) are to be localized in the sense that one is interested in inferring which of the p marginal distributions are heterogeneous between the c groups. For addressing this multiple test problem, a suitable vector of test statistics is given by

$\mathbf{H} = (H_1, \ldots, H_p)^\top$, where each H_i is a coordinate-specific statistic of Kruskal-Wallis type, i. e.,

$$
\begin{aligned}
H_i &= \frac{12}{N(N+1)} \sum_{k=1}^{c} n_k \left(\bar{R}_i^{(k)} - \frac{N+1}{2} \right)^2 = \frac{12(N+1)}{N} \sum_{k=1}^{c} n_k L_{Nik}^2 \qquad (7) \\
&= v_{N,i,i}^{-1} \sum_{k=1}^{c} n_k L_{Nik}^2, \; 1 \leq i \leq p.
\end{aligned}
$$

In (7), $\bar{R}_i^{(k)} = n_k^{-1} \sum_{\ell=1}^{n_k} R_{i\ell}^{(k)}$ denotes the (unscaled) rank average of group k in coordinate i.

Theorem 3.1. *As $\min\{n_k : 1 \leq k \leq c\} \to \infty$, the joint permutational distribution of $\mathbf{H} = (H_1, \ldots, H_p)^\top$ is under H_0 asymptotically a multivariate chi-square distribution of Wishart-type with associated correlation matrix*

$$
W_N = \left(\frac{v_{N,i,i'}}{\sqrt{v_{N,i,i} v_{N,i',i'}}} \right)_{1 \leq i,i' \leq p}.
$$

Proof. Throughout the proof, we assume that H_0 holds true. First, we notice that we can express $L_{Nic} = -n_c^{-1} \sum_{k=1}^{c-1} n_k L_{Nik}$, leading to

$$
\begin{aligned}
H_i &= v_{N,i,i}^{-1} \left(\sum_{k=1}^{c-1} n_k L_{Nik}^2 + \frac{1}{n_c} \sum_{k=1}^{c-1} \sum_{k'=1}^{c-1} n_k n_{k'} L_{Nik} L_{Nik'} \right) \\
&= v_{N,i,i}^{-1} \mathbf{L}_{Ni}^\top \left(n_k \delta_{k,k'} + \frac{n_k n_{k'}}{n_c} \right)_{k,k'=1,\ldots,c-1} \mathbf{L}_{Ni} \\
&= \mathbf{L}_{Ni}^\top \left(v_{N,i,i}^{-1} C_N^{-1} \right) \mathbf{L}_{Ni},
\end{aligned}
$$

where $\mathbf{L}_{Ni} = (L_{Ni1}, \ldots, L_{Ni(c-1)})^\top$ for $1 \leq i \leq p$ and $C_N^{-1} = \left(n_k \delta_{k,k'} + \frac{n_k n_{k'}}{n_c} \right)_{k,k'=1,\ldots,c-1}$ is the inverse of C_N from (6) (assuming invertibility of the latter). Hence, $H_i = \|\mathbf{Z}_i\|_2^2$, where $\mathbf{Z}_i = v_{N,i,i}^{-1/2} C_N^{-1/2} \mathbf{L}_{Ni}$ is asymptotically standard normal on $\mathbb{R}^{c-1}$ due to the limiting distributional result in (5) for $(L_{Nik} : 1 \leq i \leq p, 1 \leq k \leq c-1)$. Finally, it is easy to check that the full (conditional) covariance matrix of $(\mathbf{Z}_1^\top, \ldots, \mathbf{Z}_p^\top)^\top$ is asymptotically given by $W_N \otimes I_{c-1}$, implying the assertion. ∎

Remark 3.1. *The limiting result in* (5) *is not restricted to the particular form of regression coefficients leading to $\mathbf{H}$. Thus, many other, related problems in nonparametric multivariate analysis lead (under the global hypothesis) to analogous limiting multivariate chi-square distributions for the respective vectors of test statistics.*

4 Genetic association studies

From the statistical point of view, genetic association studies with case-control setup lead to the problem of simultaneous categorical data analysis, meaning that many contingency tables have

Table 1: Schematic representation of data for an association test problem at genetic locus j, where the two possible alleles are denoted by $A_{j,1}$ and $A_{j,2}$.

Genotype	$A_{j,1}A_{j,1}$	$A_{j,1}A_{j,2}$	$A_{j,2}A_{j,2}$	$\sum$
Phenotype 1	$x_{1,1}^{(j)}$	$x_{1,2}^{(j)}$	$x_{1,3}^{(j)}$	$n_{1.}$
Phenotype 0	$x_{2,1}^{(j)}$	$x_{2,2}^{(j)}$	$x_{2,3}^{(j)}$	$n_{2.}$
Absolute count	$n_{.1}{}^{(j)}$	$n_{.2}{}^{(j)}$	$n_{.3}{}^{(j)}$	N

to be analyzed simultaneously. Assuming a set of $m > 1$ bi-allelic genetic markers with exactly two possible values $A_{j,1}$ and $A_{j,2}$ (say) for $1 \leq j \leq m$, the data for genetic locus j can in such type of study be summarized as in Table 1.

The numbers $n_{1.}$ of cases (phenotype 1) and $n_{2.}$ of controls (phenotype 0) do not depend on j and are fixed by experimental design. The aim of the statistical analysis is to test the family of hypotheses $\mathcal{H} = (H_j : 1 \leq j \leq m)$, where the j-th null hypothesis H_j states that the genotype at locus j is stochastically independent of the (binary) phenotype of interest. A suitable marginal test statistic for testing H_j against its two-sided alternative K_j that the genotype at locus j is associated with the phenotype is given by

$$Q_{\text{assoc.}}(\mathbf{x}^{(j)}) = \sum_{r=1}^{2} \sum_{c=1}^{3} \frac{(x_{rc}^{(j)} - e_{rc}^{(j)})^2}{e_{rc}^{(j)}}, \tag{8}$$

where the numbers $e_{rc}^{(j)} = n_{r.}n_{.c}^{(j)}/N$ denote the expected cell counts under H_j, conditional to the marginal counts $n_{1.}$, $n_{2.}$, $n_{.1}{}^{(j)}$, $n_{.2}{}^{(j)}$, and $n_{.3}{}^{(j)}$.

It is well known that the distribution of $Q_{\text{assoc.}}(\mathbf{X}^{(j)})$ converges under H_j weakly to the (central) chi-square distribution on $[0, \infty)$ with two degrees of freedom, for all $1 \leq j \leq m$. However, the marginal test statistics typically exhibit strong correlations, at least in blocks, because of the biological mechanism of inheritance. These correlations can be described in terms of linkage disequilibrium (LD) matrices. The resulting stochastic representation of the asymptotic distribution of the vector $\mathbf{Q}_{\text{assoc.}}(\mathbf{X})$ of all m test statistics $Q_{\text{assoc.}}(\mathbf{X}^{(j)})$, $1 \leq j \leq m$, has been derived by Dickhaus and Stange (2013); see also Moskvina and Schmidt (2008) for a simpler disease risk model.

Lemma 4.1 (Dickhaus and Stange (2013))**.** *Let, for $1 \leq j \leq m$, $P_j = (P_{1j}, P_{2j}, P_{3j})^\top$ denote the vector of expected genotype frequencies at position j for cases in the target population, and define*

$$Z_{1,j} = \frac{X_{11}^{(j)} - n_{1.}P_{1j}}{\sqrt{n_{1.}P_{1j}(1 - P_{1j})}}, \tag{9}$$

$$Z_{2,j} = \frac{P_{2j}(X_{11}^{(j)} - n_{1.}P_{1j}) + (1 - P_{1j})(X_{12}^{(j)} - n_{1.}P_{2j})}{\sqrt{n_{1.}P_{2j}(1 - P_{1j})(1 - P_{1j} - P_{2j})}}. \tag{10}$$

Then, for $N \to \infty$, $(Z_{1,j}, Z_{2,j})^\top$ converges in distribution to $(Z_1, Z_2)^\top$ with $(Z_1, Z_2)^\top \sim \mathcal{N}_2(0, I_2)$, the standard normal distribution on $\mathbb{R}^2$. Furthermore, $Q_{\text{assoc.}}(\mathbf{X}^{(j)})$ converges in distribution to $Z_1^2 + Z_2^2$ under H_j. Finally, under the global hypothesis $H_0 = \bigcap_{j=1}^{m} H_j$, it holds *for all $1 \leq j,k \leq m$: For any tuple $(\ell_1, \ell_2) \in \{1,2\}^2$, the joint distribution of $(Z_{\ell_1,j}, Z_{\ell_2,k})^\top$ converges weakly to a bivariate normal distribution with correlation coefficient given by*

$$\lim_{N\to\infty} \textit{Cov}(Z_{\ell_1,j}, Z_{\ell_2,k}) = r_{j,k}(\ell_1, \ell_2) \text{ (say).} \tag{11}$$

Consequently, the vector $\mathbf{Q}_{\text{assoc.}}(\mathbf{X})$ asymptotically follows a multivariate central chi-squared distribution of generalized Wishart-type under H_0, with correlation structure given by

$$\lim_{N\to\infty} \textit{Cov}(Q_{\text{assoc.}}(\mathbf{X}^{(j)}), Q_{\text{assoc.}}(\mathbf{X}^{(k)})) = 2 \sum_{\ell_1=1}^{2} \sum_{\ell_2=1}^{2} r_{j,k}^2(\ell_1, \ell_2). \tag{12}$$

The correlation coefficients $r_{j,k}(\ell_1, \ell_2)$ in (11) only depend on the expected genotype frequencies P_{ij}, P_{ik}, $i = 1,2,3$, and on the second-order joint probabilities of genotype pairs. Thus, they can be deduced from appropriate LD matrices which are publicly available. Second-order product-type probability bounds (see Block et al. (1992) and Section 4.3 of Dickhaus (2014) for details) based on the bivariate marginal chi-square distributions of pairs of components of $\mathbf{Q}_{\text{assoc.}}(\mathbf{X})$ under H_0 can be calculated by making use of the derivations by Jensen (1970). These bounds can be used for approximating the joint m-variate distribution of $\mathbf{Q}_{\text{assoc.}}(\mathbf{X})$ under the global hypothesis in a computationally inexpensive manner. This strategy has originally been advocated by Moskvina and Schmidt (2008). In Section 6.4 below, a series expansion for the three-variate marginal distributions is provided. It allows for utilizing probability bounds (of sum- or product-type) of order 3.

5 Multiple Wald tests

In this section, we further generalize our definition of multivariate chi-square distributions of Wishart-type. In particular, we consider the following type of multivariate chi-square distributions.

Definition 5.1. *Let $m \geq 2$ and $\vec{\nu} = (\nu_1, \ldots, \nu_m)^\top$ be a vector of positive integers. Let $(Z_{1,1}, \ldots, Z_{1,\nu_1}, Z_{2,1}, \ldots, Z_{2,\nu_2}, \ldots, Z_{m,1}, \ldots, Z_{m,\nu_m})$ denote $\sum_{k=1}^m \nu_k$ jointly normally distributed random variables with joint correlation matrix $R = (\rho(Z_{k_1,\ell_1}, Z_{k_2,\ell_2}) : 1 \leq k_1, k_2 \leq m, 1 \leq \ell_1 \leq \nu_{k_1}, 1 \leq \ell_2 \leq \nu_{k_2})$ such that for any $1 \leq k \leq m$ the random vector $\mathbf{Z}_k = (Z_{k,1}, \ldots, Z_{k,\nu_k})^\top$ has a standard normal distribution on $\mathbb{R}^{\nu_k}$. Let $\mathbf{Q} = (Q_1, \ldots, Q_m)^\top$, where*

$$Q_k = \sum_{\ell=1}^{\nu_k} Z_{k,\ell}^2 \quad \text{for all} \quad 1 \leq k \leq m. \tag{13}$$

Then we call the distribution of $\mathbf{Q}$ a multivariate (central) chi-square distribution of generalized Wishart-type with parameters m, $\vec{\nu}$ and R and write $\mathbf{Q} \sim \chi^2(m, \vec{\nu}, R)$.

As demonstrated by Dickhaus (2012) in the context of likelihood-based simultaneous inference in dynamic factor models, multivariate chi-square distributions of generalized Wishart-type in the sense of Definition 5.1 occur as limit distributions of vectors of Wald statistics (or, asymptotically equivalently, likelihood ratio statistics) under H_0 if a statistical model with multiple correlated endpoints is considered. For concreteness, let us consider here a multivariate multiple linear regression model. We assume that one can observe stochastically independent vectors $\mathbf{Y}_1, \ldots, \mathbf{Y}_N$ for sample size N, where each $\mathbf{Y}_i$ takes values in $\mathbb{R}^m$ for $m > 1$. We make the model assumption that for all $1 \leq i \leq N$, $\mathbf{Y}_i \sim \mathcal{N}_m(\beta \mathbf{x}_i, \Sigma)$. In this, $\mathbf{x}_i \in \mathbb{R}^p$ denotes a (given) profile of covariates for observational unit i, β denotes an $(m \times p)$ matrix of unknown regression coefficients (the parameters of the model), and the covariance matrix $\Sigma \in \mathbb{R}^{m \times m}$ of each observable vector is regarded as a nuisance parameter. We denote by $\beta_k \in \mathbb{R}^p$ the vector of regression coefficients corresponding to coordinate k, where $1 \leq k \leq m$, which is given by the transpose of the k-th row of β. The following well-known result can for instance be found as Theorem 8.2.2. in the textbook by Anderson (1984).

Theorem 5.1. *Let $\hat{\beta}$ and $\hat{\Sigma}$ denote the maximum likelihood estimators of β and Σ. Then it holds:*

(a) vec$(\hat{\beta}^\top) = (\hat{\beta}_1^\top, \ldots, \hat{\beta}_m^\top)^\top \sim \mathcal{N}_{mp}(vec(\beta^\top), \Sigma \otimes A^{-1})$, where $A = \sum_{i=1}^N \mathbf{x}_i \mathbf{x}_i^\top$ is assumed to be invertible.

(b) $N\hat{\Sigma} \sim$ Wishart$(\Sigma, N - p)$ is stochastically independent of $\hat{\beta}$.

Now, we interpret each component of the observables as an endpoint and consider the system of hypotheses $\mathcal{H} = (H_k : 1 \leq k \leq m)$, where each endpoint-specific hypothesis H_k is a linear hypothesis, i. e., $H_k : C_k \beta_k = \xi_k$. The contrast matrices $C_k \in \mathbb{R}^{r_k \times p}$ are assumed to have rank r_k, for all $1 \leq k \leq m$, and the vectors $\xi_k \in \mathbb{R}^{r_k}$ are given. For instance, one may want to test if different subsets of the covariates have significant effects on different endpoints, while adjusting for the respectively remaining covariates. A suitable test statistic for testing H_k for each $1 \leq k \leq m$ is given by the Wald statistic $W_k = (C_k \hat{\beta}_k - \xi_k)^\top (C_k \hat{V}_k C_k^\top)^{-1} (C_k \hat{\beta}_k - \xi_k)$, where we denote by $\hat{V}_k$ the submatrix of $\hat{\Sigma} \otimes A^{-1}$ which corresponds to $\hat{\beta}_k$. Marginally, each W_k is asymptotically ($N \to \infty$) chi-square distributed with r_k degrees of freedom under H_k; cf., e. g., Section 12.4.2 in the textbook by Lehmann and Romano (2005). Hence, due to the joint normality of $\hat{\beta}$ according to Theorem 5.1, the vector $\mathbf{W} = (W_1, \ldots, W_m)^\top$ asymptotically follows a multivariate chi-square distribution in the sense of Definition 5.1 under the global hypothesis $H_0 = \bigcap_{k=1}^m H_k$.

Remark 5.1. *Limiting joint distributions of vectors of likelihood ratio statistics in more general models have been derived by Katayama (2008).*

6 Computational methods: Some representations and approximations for multivariate chi-square or gamma distributions

6.1 Notation and special functions

Since many formulas for multivariate chi-square distributions are scattered in the literature, some of them are compiled in this section. As mentioned in Section 2.2, there is no general definition of a “multivariate chi-square or gamma distribution“ but there are many well known families of distributions with one-dimensional marginal gamma distributions. In this paper we consider only the multivariate gamma distribution in the sense of Krishnamoorthy and Parthasarathy (1951), “Jensen's multivariate gamma“, derived from Jensen (1970), and the distribution from Definition 5.1. For dimension $p \in \mathbb{N}$, the Lt of the first one is given by

$$|I_p + RT|^{-\alpha} \tag{14}$$

with the $(p \times p)$-identity matrix I_p, the “associated“ correlation matrix R, $T = \text{diag}(t_1, ...t_p)$, $t_1, \ldots, t_p \geq 0$, $\alpha > p - 2$, $p \geq 2$ or $2\alpha \in \mathbb{N}$ (for $p - 1 \geq \alpha > p - 2$ see Section 2 in Royen (2007)). A characterization of R allowing all $\alpha > 0$ (i. e., infinite divisibility in (14)) is given in Griffiths (1984) and Bapat (1989). Throughout this section all correlation matrices are assumed to be regular. The distribution with Lt as in (14) is called a $\Gamma_p(\alpha, R)$-distribution. A former overview for this distribution, including some non-central extensions, is found in Royen (1997). The joint distribution of the diagonal elements of a $W_p(\nu, R)$ - Wishart matrix is a multivariate chi-square distribution with the Lt

$$|I_p + 2RT|^{-\nu/2}, \tag{15}$$

which we refer to as a “Wishart-chi-square“ with ν degrees of freedom, $\chi_p^2(\nu, R)$ for short. Here we are mainly interested in the cdf of this distribution, but formulas are given for the more general cdf derived from (14). Thus, α can be read as $\nu/2$ and a scale factor 2 can be inserted in the following formulas for the $\Gamma_p(\alpha, R)$ - cdf to obtain the corresponding $\chi_p^2(\nu, R)$-cdf.

Jensen's multivariate gamma distribution is derived from the Lt

$$\prod_{\mu=1}^{\nu} |I_p + R_\mu T|^{-1/2}. \tag{16}$$

Series representations for the corresponding $\Gamma_p(R_1, ..., R_\nu)$-pdf were derived by Jensen (1970) for $p = 2$ and $p = 3$, but the trivariate series are not always convergent. Jensen (1970) has also given a formula for $(p \times p)$ tridiagonal matrices R_μ, but unfortunately, it is based on a formula for determinants of tridiagonal matrices, which is incorrect for $p > 3$. Always absolutely convergent series for convolutions of not identically scaled multivariate gamma distributions (i. e., $\Gamma_p(\alpha_1, ..., \alpha_\nu, \Sigma_1, ..., \Sigma_\nu)$ - distributions) with Lt

$$\prod_{\mu=1}^{\nu} |I_p + \Sigma_\mu T|^{-\alpha_\mu}$$

and regular covariance matrices Σ_μ, $1 \leq \mu \leq \nu$, are given in Royen (2013b). Actual computations are feasible at least for $p \leq 3$.

The cdfs corresponding to (14) and (16) are denoted by $F(x_1, ..., x_p, \alpha, R)$ and $F(x_1, ..., x_p, R_1, ..., R_\nu)$ respectively. For their representations we use the following notations: The spectral norm of any $(p \times p)$-matrix $A = (a_{ik})$ ist denoted by $||A||$ and its determinant by $|A|$, $\dot{A}$ is defined by $A - \text{diag}(a_{11}, ..., a_{pp})$, $a^{ik} : 1 \leq i, k \leq p$ are the elements of A^{-1} and $A > 0$ means positive definiteness of A. An identity matrix is always denoted by I. The notation $\sum_{(n)}$ stands for a summation over all decompositions of a non-negative integer $n = \sum n_i$ with non-negative integers $n_i, i = 1, \ldots, p$. Formulas from the NIST Handbook of Mathematical Functions (Olver et al. (2010)) are cited by HMF and their number. The pdf of a gamma distribution with shape parameter α is given by

$$g_\alpha(x) = e^{-x}x^{\alpha-1}/\Gamma(\alpha), \quad x > 0, \alpha > 0, \text{ with cdf } G_\alpha(x) = \int_0^x g(\xi)d\xi. \tag{17}$$

Furthermore, we need the derivatives

$$G_{\alpha+n}^{(n)}(x) = \left(\frac{d}{dx}\right)^n G_{\alpha+n}(x) = \binom{\alpha+n-1}{n-1}^{-1} L_{n-1}^{(\alpha)}(x)g_{\alpha+1}(x), n \geq 1, \tag{18}$$

with the Laguerre polynomials $L_n^{(\alpha)}$ (HMF 18.3,18.5) and the functions

$$H_{\alpha,n}(x) = \sum_{m=0}^{n}(-1)^{n-m}\binom{n}{m}2^m G_{\alpha+m}(x) \tag{19}$$

with

$$h_{\alpha,n}(x) = \frac{d}{dx}H_{\alpha,n}(x) = (-1)^n\binom{\alpha+n-1}{n}^{-1} L_n^{(\alpha-1)}(2x)g_\alpha(x).$$

Alternatively, we can also use the relation $H_{\alpha,n+1} = H_{\alpha,n} - 2h_{\alpha+1,n}$ which can be verified by Lt. Besides, $\lim\limits_{n\to\infty} H_{\alpha,n}(x) = 0$ for every $x \in (0,\infty)$, see Section 2 of Royen (1991). Moreover, we need the extension of the non-central gamma cdf

$$G_\alpha(x,y) = e^{-y}\sum_{n=0}^{\infty} G_{\alpha+n}(x)\frac{y^n}{n!} = \sum_{n=0}^{\infty} G_{\alpha+n}^{(n)}(x)\frac{(-y)^n}{n!} \tag{20}$$

to $x \in \mathbb{C}, y \in \mathbb{C}$ (actually, we only need $2\alpha \in \mathbb{N}, x \in \mathbb{R}, y \in \mathbb{C}$). For $\alpha = 1/2 + n$ we have

$$\begin{aligned} &G_{1/2+n}(x,y) - \frac{1}{2}\left(\mathrm{erf}(\sqrt{x}+\sqrt{y}) + \mathrm{erf}(\sqrt{x}-\sqrt{y})\right) \\ &= -e^{-x-y}\sum_{k=1}^{n}\left(\frac{x}{y}\right)^{(k-1/2)/2} I_{k-1/2}(2\sqrt{xy}) = -e^{-y}\sum_{k=1}^{n} g_{1/2+k}(x)\,{}_0F_1(1/2+k;xy) \end{aligned} \tag{21}$$

with the principal value $\sqrt{x}$ and the modified Bessel functions $I_{k-1/2}$, which are elementary functions (HMF 10.49 (ii)). In a similar way

$$G_{1+n}(x,y) - G_1(x,y) = -e^{-y}\sum_{k=1}^{n} g_{1+k}(x)\,{}_0F_1(1+k;xy)$$

with

$$G_1(x,y) = e^{-y}\int_0^x \exp(-\xi)I_0(2\sqrt{\xi y})d\xi \tag{22}$$

$$= 1 - e^{-x-y}\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\frac{x^k}{k!}\right)\frac{y^n}{n!} = e^{-x-y}\sum_{k=1}^{\infty}\left(\sum_{n=0}^{k-1}\frac{y^n}{n!}\right)\frac{x^k}{k!}. \tag{23}$$

6.2 Some representations for the $\Gamma_p(\alpha, R)$-cdf

Explicit and actually computable formulas for $F(x_1, ..., x_p; \alpha, R)$ are only available for special structures of R and in the general case only for low dimensions p.

Definition 6.1. *A regular $(p \times p)$-covariance matrix Σ is called "real m-factorial" if m is the lowest integer allowing a representation*

$$\Sigma = D + AA^\top$$

with a real invertible diagonal matrix D and a $(p \times m)$-matrix A of rank m with real or purely imaginary columns. Σ is called "m-factorial" if the real D in the above definition is replaced by a positive definite D.

A regular $(p\times p)$-correlation matrix R with eigenvalues $\lambda_1, \ldots, \lambda_p > 0$ has an at most $(p-1)$-factorial representation with $D = \lambda_{\min} I_p$. For an m-factorial R with $D = W^{-2} > 0$ we have $WRW = I + BB^\top, B = (b_{j\mu}) = WA$ with rows b^j, columns b_μ, and $b_\mu^\top b_\kappa = 0, \mu \neq \kappa$, w.l.o.g. Then the $\Gamma_p(\alpha, R)$ - cdf is given by

$$F(x_1, ..., x_p; \alpha, R) = \mathbb{E}\left[\prod_{j=1}^{p} G_\alpha\left(w_j^2 x_j, \frac{1}{2} b^j S\left(b^j\right)^\top\right)\right], \tag{24}$$

where the expectation refers to the $W_m(2\alpha, I_m)$-distributed Wishart matrix S, see Royen (1995). With m = 1 and $R = \text{diag}(..., 1 - a_j^2, ...) + aa^\top, \max(a_j^2) < 1$, we obtain

$$\begin{aligned} F(x_1, ..., x_p; \alpha, R) &= \int_0^\infty \left[\prod_{j=1}^{p} G_\alpha\left(\frac{x_j}{1-a_j^2}, \frac{a_j^2 y}{1-a_j^2}\right)\right] g_\alpha(y) dy \\ &= \sum_{n=0}^{\infty} (\alpha)_n \lambda^{-\alpha-n} \sum_{(n)} \prod_{j=1}^{p} G_{\alpha+n_j}\left(\frac{x_j}{1-a_j^2}\right) \frac{\left(\frac{a_j^2}{1-a_j^2}\right)^{n_j}}{n_j!}, \end{aligned} \tag{25}$$

where $(\alpha)_n = \alpha(\alpha+1)\cdots(\alpha+n-1)$ and $\lambda = 1 + \sum_{j=1}^{p} a_j^2/(1-a_j^2) > 0$, since $|R| = \lambda \prod_{j=1}^{p}(1-a_j^2) > 0$. For a real vector a all $\alpha > 0$ are admissible in (25). In the limit case with a real a, $a_k^2 = 1$, (and consequently $|R| = \prod_{j\neq k}(1-a_j^2) > 0$) we find

$$\begin{aligned} F(x_1, ..., x_p; \alpha, R) &= \int_0^{x_k} \prod_{j\neq k} G_\alpha\left(\frac{x_j}{1-a_j^2}, \frac{a_j^2 y}{1-a_j^2}\right) g_\alpha(y) dy \\ &= \sum_{n=0}^{\infty} (\alpha)_n \lambda_k^{-\alpha-n} G_{\alpha+n}(\lambda_k x_k) \sum_{(n)} \prod_{j\neq k} G_{\alpha+n_j}\left(\frac{x_j}{1-a_j^2}\right) \frac{\left(\frac{a_j^2}{1-a_j^2}\right)^{n_j}}{n_j!}, \end{aligned} \tag{26}$$

where $\lambda_k = 1 + \sum_{j\neq k} a_j^2/(1-a_j^2)$. For non-central $\chi_p^2(\nu, R, \Delta)$-distributions with one-factorial R see Royen (1995, 1997).

Now let R be real one-factorial. If ${a_k}^2 = \max a_j^2 > 1$, then λ from (25) becomes negative and we get with

$$G_\alpha^*(x,y) = \exp(y) G_\alpha(x,y) \tag{27}$$

that

$$F(x_1,\ldots,x_p;\alpha,R) = \lambda^{-\alpha}\int_0^\infty \left[\prod_{j=1}^p G_\alpha^*\left(\frac{x_j}{1-a_j^2}, \frac{a_j^2}{1-a_j^2}\frac{y}{\lambda}\right)\right] g_\alpha(y)dy \tag{28}$$

$$= (2(1-a_k^{-2}))^\alpha \sum_{n=0}^\infty (\alpha)_n H_{\alpha,n}\left(\frac{x_k}{a_k^2-1}\right) \times$$

$$\sum_{(n)} \frac{c_{kk}^{n_k}}{n_k!} \prod_{j\neq k} \frac{c_{jk}^{2n_j}}{n_j!} H_{\alpha,n_j}\left(\frac{x_j}{1-a_j^2}\right) \tag{29}$$

with the functions $H_{\alpha,n}$ from (19) and the elements c_{jk} from the matrix

$$C = I_p - 2(I_p + WRW)^{-1},\ W = \text{diag}(w_1,\ldots,w_p),\ w_j^2 = |1-a_j^2|^{-1},$$

$(c_{ij} = 0,\ i,j \neq k)$. For the integral representation see formula (21) in Royen (2007) and for the series in (29) see (15) in Royen (1997) or (3.12) in Royen (1991) if $p = 3$. If $r_{ij} = 0$, for $i \neq j$ and $i,j \neq k$, then

$$F(x_1,\ldots,x_p;\alpha,R) = \int_0^\infty \left[\prod_{j\neq k} G_\alpha(x_j, -r_{jk}^2 y)\left(\frac{x_k}{y}\right)^{\alpha/2} J_\alpha(2\sqrt{x_k y})\right] g_\alpha(y)dy$$

$$= \sum_{n=0}^\infty (\alpha)_n G_{\alpha+n}^{(n)}(x_k) \sum_{(n)} \prod_{j\neq k} \frac{r_{jk}^{2n_j}}{n_j!} G_{\alpha+n_j}^{(n_j)}(x_j). \tag{30}$$

All the series in (25), (26), (29) and (30) are absolutely convergent. The formulas (25), (26) and (29) include all $\Gamma_3(\alpha,R)$-cdfs with $R = (r_{ik}), r_{12}r_{13}r_{23} \neq 0$, since $r_{ij} = s_{ij}|r_{ij}| = a_i a_j, i \neq j$, with $a_i = s^{1/2} s_{jk} |r_{ij} r_{ik}/r_{jk}|^{1/2}$ for each permutation (i,j,k) of $(1,2,3)$, where $s = s_{12}s_{13}s_{23}$. Formula (30) can be applied if there is exactly one correlation $r_{ij} = 0$ in the (3×3)-matrix R.

The $\Gamma_2(\alpha,R)$-cdf is given by

$$\begin{aligned}
F(x_1,x_2;\alpha,R) &= F(x_1,x_2;\alpha,r)\\
&= \sum_{n=0}^\infty \binom{\alpha+n-1}{n} r^{2n} G_{\alpha+n}^{(n)}(x_1) G_{\alpha+n}^{(n)}(x_2)\\
&= (1-r^2)^\alpha \sum_{n=0}^\infty \binom{\alpha+n-1}{n} r^{2n} G_{\alpha+n}\left(\frac{x_1}{1-r^2}\right) G_{\alpha+n}\left(\frac{x_2}{1-r^2}\right)\\
&= (1-c^2)^\alpha \sum_{n=0}^\infty \binom{\alpha+n-1}{n} c^{2n} H_{\alpha,n}\left(\frac{x_1}{\sqrt{1-r^2}}\right) H_{\alpha,n}\left(\frac{x_2}{\sqrt{1-r^2}}\right),
\end{aligned}$$

where $c = r/(1+\sqrt{1-r^2})$ and $H_{\alpha,n}$ is as in (19). The former two series are well known and the latter one follows from the general p-variate series by Royen (1991).

Royen (1991) provided three further types (a), (b), (c) of absolutely convergent series for the general $\chi_p^2(\nu, R)$-cdf. Here only the resulting series for the $\Gamma_3(\alpha, R)$-cdf are given. The off-diagonal elements c_{ik} of a symmetrical (3×3)-matrix (c_{ik}) are also denoted here by c_j for each permutation (i, j, k) of $(1, 2, 3)$. We get that

$$F(x_1, x_2, x_3; \alpha, R) = c \left(\prod_{j=1}^{3} G_\alpha(d_j x_j) + \sum_{N=2}^{\infty} P_N(x_1, x_2, x_3) \right) \text{ with} \tag{31}$$

$$\Gamma(\alpha) P_N(x_1, x_2, x_3) = \begin{cases} \displaystyle\sum_{m_1+m_2+m_3=n} \left(\sum_{m=0}^{\min(m_j)} \frac{2^{2m}\Gamma(\alpha+n-m)}{(2m)! \prod_{j=1}^{3} (m_j-m)!} \right) \prod_{j=1}^{3} c_j^{2m_j} F_{\alpha,n-m_j}(d_j x_j), \\ \displaystyle\sum_{m_1+m_2+m_3=n-1} \left(\sum_{m=0}^{\min(m_j)} \frac{2^{2m+1}\Gamma(\alpha+n-m)}{(2m+1)! \prod_{j=1}^{3} (m_j-m)!} \right) \times \\ \displaystyle\qquad \prod_{j=1}^{3} c_j^{2m_j+1} F_{\alpha,n-m_j}(d_j x_j), \end{cases} \tag{32}$$

for $N = 2n \geq 2$ (upper branch in (32)) and $N = 2n + 1 \geq 3$ (lower branch in (32)), respectively. Three alternative choices for the quantities in (31) and (32) are

$$c = 1,\ d_j = 1,\ c_j = -r_j,\ F_{\alpha,n} = G_{\alpha+n}^{(n)},\ \text{provided that } ||\dot{R}|| < 1, \tag{33a}$$

$$c = |Q|^\alpha,\ Q = (q_{ik}),\ q_{ik} = r^{ik}/(r^{ii}r^{kk})^{1/2},\ (r^{ik}) = R^{-1},\ d_j = r^{jj},$$
$$c_j = -q_j = -q_{ik},\ F_{\alpha,n} = G_{\alpha+n}, \tag{33b}$$

$$c = |2(I_3 + WRW)^{-1}|^\alpha,\ W = \text{diag}(w_1, w_2, w_3),\ d_j = w_j^2,\ c_j = c_{ik},\ F_{\alpha,n} = H_{\alpha,n}, \tag{33c}$$

with unique positive scale factors w_j implying $c_{11} = c_{22} = c_{33} = 0$ in $C = (c_{ik}) = I_3 - 2(I_3 + WRW)^{-1}$. For existence, uniqueness and computation of the w_j see formulas (3.1) and (3.2) by Royen (1991).

Now we consider real two-factorial correlation matrices and four-variate gamma distributions. A real two-factorial representation $R = D^{-1} + AA^\top$ of a $(p \times p)$-correlation matrix R with $p \geq 4$ is equivalent to a real two-factorial representation

$$R^{-1} = D + BB^\top, \quad B = (b_{j\mu}),\ j = 1, \ldots, p,\ \mu = 1, 2. \tag{34}$$

According to Lemma 2 of Royen (2007) a regular irreducible (4×4)-covariance matrix Σ with at least four off-diagonal elements $\sigma_{ik} \neq 0$ is real m-factorial with $m \leq 2$. Moreover, it is always possible to obtain at most one negative element d_j in D. Let Σ be two-factorial and (i, j, k, ℓ) any permutation of $(1, 2, 3, 4)$ with $\sigma_{ij}\sigma_{ik}\sigma_{jk} \neq 0$. Then

$$d_k = \sigma_{kk} - \sigma_{ik}\sigma_{jk}\sigma_{ij}^{-1} + (\sigma_{k\ell} - \sigma_{ik}\sigma_{j\ell}\sigma_{ij}^{-1})(\sigma_{k\ell} - \sigma_{i\ell}\sigma_{jk}\sigma_{ij}^{-1})/[d_\ell - (\sigma_{\ell\ell} - \sigma_{i\ell}\sigma_{j\ell}\sigma_{ij}^{-1})].$$

For a real two-factorial $R^{-1} = (r^{ik})$ we obtain with G_α^* from (27) and the density f_α, given by

$$f_\alpha(\varphi) = (\sin^2 \varphi)^{\alpha-1}/B(1/2, \alpha - 1/2),\ 0 < \varphi < \pi,\ \alpha > 1/2, \tag{35}$$

that

$$
\begin{aligned}
F(x_1,\ldots,x_p;\alpha,R) &= |DR|^{-\alpha}\int_0^{\pi}\int_0^{\infty}\int_0^{\infty}\left[\prod_{j=1}^{p} G_\alpha^*\left(d_jx_j,-d_j^{-1}h_j(y_1,y_2;\varphi)\right)\right]\times \\
&\quad g_\alpha(y_1)g_\alpha(y_2)f_\alpha(\varphi)dy_1dy_2d\varphi, \qquad (36)\\
h_j(y_1,y_2;\varphi) &= b_{j1}^2y_1+b_{j2}^2y_2+2b_{j1}b_{j2}\sqrt{y_1y_2}\cos(\varphi). \qquad (37)
\end{aligned}
$$

If $p=4$ in (36), one element d_ℓ in D - determining the remaining ones - can be chosen within a certain set of possible values. Then B in $BB^\top = R^{-1}-D$ has rank 2 and is e. g. available by the eigenvectors and the two eigenvalues $\lambda_1,\lambda_2\neq 0$ of $R^{-1}-D$. If there exists an index ℓ, w.l.o.g. $\ell=p=4$, with a real one-factorial (3×3)-covariance matrix $(r^{jk|4})=(r^{jk}-r^{j4}r^{k4}/r^{4,4})$, i. e., $\prod_{1\le j<k\le 3} r^{jk|4}\neq 0$, then we can choose the limit value $d_4=0$. It is

$$
\begin{aligned}
\lim_{d\to 0} d^{-\alpha}G_\alpha^*(dx,d^{-1}y) &= \lim_{d\to 0}\frac{1}{\Gamma(\alpha)}\int_0^x {}_0F_1(\alpha;\xi y)\exp(-d\xi)\xi^{\alpha-1}d\xi \\
&= \frac{x^\alpha}{\Gamma(\alpha+1)}{}_0F_1(\alpha+1;xy)=\left(\frac{x}{y}\right)^{\alpha/2} I_\alpha(2\sqrt{xy}).
\end{aligned}
$$

Together with

$$
\begin{aligned}
b_{j1} &= b_j = s_{ik}\sqrt{s}|r^{ij|4}r^{jk|4}/r^{ik|4}|^{1/2}, s_{ik}=\mathrm{sgn}(r^{ik|4}), s=s_{12}s_{13}s_{23}, j\le 3,\\
b_{j2} &= (r^{4,4})^{-1/2}r^{j4}, j\le 3, b_{4,1}=0, b_{4,2}=\sqrt{r^{4,4}},\\
d_j &= r^{jj}-b_{j1}^2-b_{j2}^2, j\le 3, \text{ and } d_4=0
\end{aligned}
$$

in (36) and (37), this leads to the representation

$$
\begin{aligned}
F(x_1,...,x_4;\alpha,R) &= \left(|R|\prod_{j\le 3}d_j\right)^{-\alpha}\int_0^{\pi}\int_0^{\infty}\int_0^{\infty}\left[\prod_{j=1}^{3} G_\alpha^*\left(d_jx_j,-d_j^{-1}h_j(y_1,y_2;\varphi)\right)\times\right. \\
&\quad \left.(\Gamma(\alpha+1))^{-1}x_4^\alpha{}_0F_1(\alpha+1;-r^{4,4}x_4y_2)\right]\times \\
&\quad g_\alpha(y_1)g_\alpha(y_2)f_\alpha(\varphi)dy_1dy_2d\varphi. \qquad (38)
\end{aligned}
$$

In the case of its existence we can also get formulas from a two-factorial representation $R=D^{-1}+AA^\top$, which is equivalent to $WRW=I_4+BB^\top$ with $D=W^2=\mathrm{diag}(w_1^2,\ldots,w_4^2)>0$, $B=WA=(b_1,b_2)=(b_{j\mu})$ and, w.l.o.g., $b_1^\top b_2=0$ since $BB^\top=U\Lambda^{1/2}\Lambda^{1/2}U^\top$ with an orthogonal matrix U and $\Lambda=\mathrm{diag}(\lambda_1,\lambda_2,0,0)$, where $\lambda_1,\lambda_2\neq 0$. Then

$$
\begin{aligned}
F(x_1,...,x_4;\alpha,R) &= \int_0^{\pi}\left(\int_0^{\infty}\int_0^{\infty}\left[\prod_{j=1}^{4} G_\alpha\left(w_j^2x_j,h_j(y_1,y_2;\varphi)\right)\right]\times\right. \\
&\quad \left. g_\alpha(y_1)g_\alpha(y_2)dy_1dy_2\right)f_\alpha(\varphi)d\varphi. \qquad (39)
\end{aligned}
$$

If there exists an index ℓ - w.l.o.g. $\ell = 4$ - with a one-factorial conditional covariance matrix $(r_{jk|4}) = (r_{jk} - r_{j4}r_{k4})_{1\le j,k\le 3} = W^{-2} + aa^\top$, then, with $b_{j1} = w_j a_j$ and $b_{j2} = w_j r_{j4}$ in (37) for $j \le 3$, the cdf is given by

$$\begin{aligned} F(x_1, ..., x_4; \alpha, R) \;&=\; \int_0^{\pi} \left(\int_0^{x_4} \int_0^{\infty} \left[\prod_{j=1}^{3} G_\alpha \left(w_j^2 x_j, h_j(y_1, y_2; \varphi) \right) \right] \times \right. \\ &\quad \left. g_\alpha(y_1) g_\alpha(y_2) dy_1 dy_2 \right) f_\alpha(\varphi) d\varphi. \qquad (40) \end{aligned}$$

For a numerical evaluation the formulas (39) and (40) are more favourable, but formulas (36) and (38) are more general. They are supplemented by the following formulas for some special cases. If $r^{ij} = 0$ then

$$\begin{aligned} F(x_1, ..., x_4; \alpha, R) \;&=\; \left[(1 - r_{k\ell}^2)^\alpha \pi^{1/2} \Gamma(\alpha) \Gamma(\alpha - 1/2) \right]^{-1} \times \\ &\quad \int_0^{\pi} \int_0^{x_k} \int_0^{x_\ell} \left[\prod_{m=i,j} G_\alpha(\sigma_m^{-2} x_m, Q_m) \right] \times \\ &\quad \exp\left(- \frac{y_k + y_\ell - 2 r_{k\ell} \sqrt{y_k y_\ell} \cos(\varphi)}{1 - r_{k\ell}^2} \right) (y_k y_\ell \sin^2 \varphi)^{\alpha-1} dy_\ell dy_k d\varphi \\ &= \; (1 - r_{k\ell}^2)^\alpha \times \\ &\quad \int_0^{\pi} \int_0^{x_k/(1-r_{k\ell}^2)} \int_0^{x_\ell/(1-r_{k\ell}^2)} \prod_{m=i,j} G_\alpha(\sigma_m^{-2} x_m, (1 - r_{k\ell}^2) Q_m) \times \\ &\quad \exp(2 r_{k\ell} \sqrt{y_k y_\ell} \cos(\varphi)) g_\alpha(y_\ell) g_\alpha(y_k) f_\alpha(\varphi) dy_\ell dy_k d\varphi, \qquad (41) \end{aligned}$$

with f_α from (35), $\sigma_m^2 = \sigma_{mm|k\ell} = 1 - (r_{mk}^2 + r_{m\ell}^2 - 2 r_{k\ell} r_{mk} r_{m\ell})/(1 - r_{k\ell}^2)$, and

$$Q_m = \frac{(r_{mk} - r_{m\ell} r_{k\ell})^2 y_k + (r_{m\ell} - r_{mk} r_{k\ell})^2 y_\ell + 2(r_{mk} - r_{m\ell} r_{k\ell})(r_{m\ell} - r_{mk} r_{k\ell}) \sqrt{y_k y_\ell} \cos(\varphi)}{\sigma_m^2 (1 - r_{k\ell}^2)^2},$$

$m = i, j$; see also Section 5 by Royen (1995).

Now let $R^{-1} = (r^{ik})$, or equivalently the “standardized“ inverse

$$Q = (q_{ik}) = (r^{ik}/(r^{ii} r^{kk})^{1/2}) \qquad (42)$$

of R, be a tridiagonal matrix, possibly after a permutation of the variables. Then the $\Gamma_4(\alpha, R)$-

cdf is given by

$$\begin{aligned}
F(x_1,...,x_4;\alpha,R) &= |Q|^\alpha \int_0^{r^{22}x_2}\int_0^{r^{33}x_3} G_\alpha^*\left(r^{11}x_1, q_{12}^2 y_2\right) G_\alpha^*\left(r^{44}x_4, q_{34}^2 y_3\right) \times \\
&\qquad {}_0F_1(\alpha; q_{23}^2 y_2 y_3) \prod_{i=2}^{3} g_\alpha(y_i) dy_i \\
&= |Q|^\alpha \sum_{n=0}^{\infty} \binom{\alpha+n-1}{n} q_{23}^{2n} \times \\
&\int_0^{r^{22}x_2} G_\alpha^*(r^{11}x_1, q_{12}^2 y) g_{\alpha+n}(y) dy \int_0^{r^{33}x_3} G_\alpha^*(r^{44}x_4, q_{34}^2 y) g_{\alpha+n}(y) dy \\
&= \frac{|Q|^\alpha}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \sum_{n_{12}+n_{23}+n_{34}=n} \prod_{i<j} \frac{q_{ij}^{2n_{ij}}}{\Gamma(\alpha+n_{ij}) n_{ij}!} \times \\
&\prod_{i=1}^{4} \Gamma(\alpha+N_i) G_{\alpha+N_i}(r^{ii}x_i), \qquad (43)
\end{aligned}$$

with

$$N_1 = n_{12}, N_2 = n_{12}+n_{23}, N_3 = n_{23}+n_{34}, N_4 = n_{34}. \qquad (44)$$

If, after a suitable permutation, R itself is tridiagonal, then the $\Gamma_4(\alpha, R)$-cdf is given by

$$\begin{aligned}
F(x_1,...,x_4;\alpha,R) &= \int_0^\infty \int_0^\infty G_\alpha(x_1, -r_{12}^2 y_2) G_\alpha(x_4, -r_{34}^2 y_3) {}_0F_1(\alpha; r_{23}^2 y_2 y_3) \times \\
&\prod_{i=2}^{3} (\Gamma(\alpha+1))^{-1} x_i^\alpha {}_0F_1(\alpha+1; -x_i y_i) g_\alpha(y_i) dy_i \\
&= \sum_{n=0}^{\infty} \binom{\alpha+n-1}{n} r_{23}^{2n} L_n(x_1, x_2; \alpha, r_{12}^2) L_n(x_4, x_3; \alpha, r_{34}^2) \\
&= \frac{1}{\Gamma(\alpha)} \sum_{n=0}^{\infty} \sum_{n_{12}+n_{23}+n_{34}=n} \prod_{i<j} \frac{r_{ij}^{2n_{ij}}}{\Gamma(\alpha+n_{ij}) n_{ij}!} \times \\
&\prod_{i=1}^{4} \Gamma(\alpha+N_i) G_{\alpha+N_i}^{(N_i)}(x_i), \qquad (45)
\end{aligned}$$

with N_i as in (44) and

$$L_n(x_1, x_2; \alpha, r^2) = \int_0^\infty G_\alpha(x_1, -r^2 y)(\Gamma(\alpha+1))^{-1} x_2^\alpha {}_0F_1(\alpha+1; -x_2 y) g_{\alpha+n}(y) dy. \qquad (46)$$

For the case of $r_{ij} = 0$,$i \neq j$, $i, j \neq k$, $\prod_{i\neq k} r_{ik} \neq 0$, see (30). If $q_{ij} = 0$, $i \neq j$, $i, j \neq k$, $\prod_{i\neq k} q_{ik} \neq 0$, then R is the limit case of a one-factorial correlation matrix with a real vector a, $a_k^2 \to 1$, see (26). Thus, all regular irreducible not real one-factorial (4×4)-correlation matrices are included by formulas (36), (38) - (41), (43), (45), (26) and (30).

For generalizations of (43) and (45) see Royen (1994). Namely, each symmetrical $(p\times p)$-matrix $C=(c_{ij})$ can be mapped to a graph $\mathcal{G}(C)$ with vertices $1,...,p$ and edges $[i,j]$ corresponding to the non-zero c_{ij}, $i\neq j$. The matrix C is of a “tree type“ if $\mathcal{G}(C)$ is a spanning tree, meaning that $\mathcal{G}(C)$ is connected with exactly $p-1$ edges and therefore without cycles. If the standardized inverse Q from (42) has a tree type then we obtain the infinitely divisible $\Gamma_p(\alpha,R)$-cdf

$$\begin{aligned} F(x_1,...,x_p;\alpha,R) &= \frac{|Q|^\alpha}{\Gamma(\alpha)}\sum_{n=0}^{\infty}\sum_{(n)}\prod_{i<j,q_{ij}\neq 0}\frac{q_{ij}^{2n_{ij}}}{\Gamma(\alpha+n_{ij})n_{ij}!}\times \\ & \prod_{i=1}^{p}\Gamma(\alpha+N_i)G_{\alpha+N_i}(r^{ii}x_i) \end{aligned} \tag{47}$$

with $N_i=\sum_{j=1}^{p}n_{ij}$, $n_{ij}=n_{ji}$, $n_{ii}=0$, $n_{ij}=0$ if $q_{ij}=0$, and the inner sums taken over all decompositions $n=\sum_{i<j}n_{ij}$. This follows from the density

$$f(x_1,...,x_p;\alpha,R)=|Q|^\alpha\prod_{i=1}^{p}r^{ii}g_\alpha(r^{ii}x_i)\prod_{i<j}{}_0F_1(\alpha;r^{ij2}x_ix_j),$$

which was already derived by Blumenson and Miller (1963) for a tridiagonal R^{-1}. For R of a tree type we find the not infinitely divisible cdf

$$F(x_1,...,x_p;\alpha,R)=\frac{1}{\Gamma(\alpha)}\sum_{n=0}^{\infty}\sum_{(n)}\prod_{i<j,r_{ij}\neq 0}\frac{r_{ij}^{2n_{ij}}}{\Gamma(\alpha+n_{ij})n_{ij}!}\prod_{i=1}^{p}\Gamma(\alpha+N_i)G_{\alpha+N_i}^{(N_i)}(x_i). \tag{48}$$

6.3 Some formulas for Jensen’s multivariate gamma distribution

The bivariate cdf, derived from the Lt in (16), is given by

$$\begin{aligned} F(x_1,x_2;R_1,\dots,R_\nu) &= F(x_1,x_2;r_1,...,r_\nu) \\ &= \sum_{n=0}^{\infty}a_n(r_1^2,\dots,r_\nu^2)G_{\nu/2+n}^{(n)}(x_1)G_{\nu/2+n}^{(n)}(x_2), \end{aligned} \tag{49}$$

where $a_n\equiv a_n(r_1^2,\dots,r_\nu^2)$ is given by

$$\begin{aligned} a_n &= \sum_{(n)}\prod_{\mu=1}^{\nu}\binom{n_\mu-1/2}{n_\mu}r_\mu^{2n_\mu}, \text{ or recursively by} \\ a_0 &= 1,\ a_{n+1}=\frac{1}{2(n+1)}\sum_{m=0}^{n}a_{n-m}\sum_{\mu=1}^{\nu}r_\mu^{2(m+1)}. \end{aligned}$$

If $(X_1,X_2)\sim\Gamma_2(r_1,\dots,r_\nu)$ and $(Y_1,Y_2)\sim\Gamma_2(\nu/2,r)$ with $r=\left(\nu^{-1}\sum_{\mu=1}^{\nu}r_\mu^2\right)^{1/2}$ then

$$\text{Pr}(a\leq X_1\leq b,a\leq X_2\leq b)\geq\text{Pr}(a\leq Y_1\leq b,a\leq Y_2\leq b),0\leq a<b\leq\infty, \tag{50}$$

with the equal sign only for identical $r_\mu = r$, see Theorem 5 by Royen (2013b).

For the trivariate case we start from the Lt

$$\prod_{\mu=1}^{\nu} |I_3 + R_\mu T|^{-1/2} = \prod_{j=1}^{3} z_j^{\nu/2} \prod_{\mu=1}^{\nu} |I_3 + \dot{R}_\mu U|^{-1/2}, \tag{51}$$

where $z_j = (1+t_j)^{-1}$, $U = \text{diag}(u_1, u_2, u_3)$, and $u_j = 1 - z_j = t_j(1+t_j)^{-1}$. For each permutation (i,j,k) of $(1,2,3)$ we set $r_{\mu,j} = r_{\mu,ik}$. Then we obtain

$$|I_3 + \dot{R}_\mu U|^{-1/2} = \sum_{N=0}^{\infty} T_N(U; \frac{1}{2}, \dot{R}_\mu), \quad T_0 = 1, T_1 = 0, \tag{52}$$

with Taylor polynomials

$$\begin{aligned} T_N &= (-1)^N \sum_{n_1+n_2+n_3=N,\, n_j \le [N/2]} \prod_{j=1}^{3} \frac{(2n_j)! r_{\mu,j}^{N-2n_j}}{2^{n_j} n_j! (N-2n_j)!} u_j^{n_j} \\ &= \begin{cases} \sum\limits_{m_1+m_2+m_3=n} \prod\limits_{j=1}^{3} \frac{(2(n-m_j))! r_{\mu,j}^{2m_j}}{2^{n-m_j}(n-m_j)!(2m_j)!} u_j^{n-m_j}, & N = 2n \ge 2, \\ -\sum\limits_{m_1+m_2+m_3=n-1} \prod\limits_{j=1}^{3} \frac{(2(n-m_j))! r_{\mu,j}^{2m_j+1}}{2^{n-m_j}(n-m_j)!(2m_j+1)!} u_j^{n-m_j}, & N = 2n+1 \ge 3. \end{cases} \end{aligned} \tag{53}$$

For dimension $p = 3$ these formulas follow from formula (11) by Royen (1997). At least for a low degree of freedom we can compute the Taylor polynomials T_n in the expansion

$$\begin{aligned} \prod_{\mu=1}^{\nu} |I_3 + \dot{R}_\mu U|^{-1/2} &= 1 + \sum_{n=2}^{\infty} T_n(U; \dot{R}_1, \ldots, \dot{R}_\nu) \\ &= \sum_{n=0}^{\infty} \sum_{(n)} t(n_1, n_2, n_3, \dot{R}_1, \ldots, \dot{R}_\nu) u_1^{n_1} u_2^{n_2} u_3^{n_3} \end{aligned} \tag{54}$$

by direct multiplication of the Taylor series in (52), and find by inversion of the Lt, followed by termwise integration, that

$$\begin{aligned} F(x_1, x_2, x_3; R_1, \ldots, , R_\nu) &= \prod_{j=1}^{3} G_{\nu/2}(x_j) + \sum_{n=2}^{\infty} \sum_{(n)} t(n_1, n_2, n_3, \dot{R}_1, ..., \dot{R}_\nu) \times \\ & \quad \prod_{j=1}^{3} G_{\nu/2+n_j}^{(n_j)}(x_j). \end{aligned} \tag{55}$$

These series are absolutely convergent if

$$\max\left\{ ||\dot{R}_\mu||, 1 \le \mu \le \nu \right\} < 1 \Longleftrightarrow \max\left\{ \sum_{j=1}^{3} r_{\mu,j}^2 + 2\prod_{j=1}^{3} |r_{\mu,j}|, 1 \le \mu \le \nu \right\} < 1. \tag{56}$$

A modified computation uses

$$|I_3 + \dot{R}_\mu U| = 1 - \zeta(U; \dot{R}_\mu) = 1 - \zeta_\mu,\ \zeta_\mu = \sum_{i<k} r_{\mu,ik}^2 u_i u_k - 2\prod_{j=1}^{3} r_{\mu,j} u_j, \quad (57)$$

$$|I_3 + \dot{R}_\mu U|^{-1/2} = \exp\left(\frac{1}{2}\sum_{N=1}^{\infty}(-1)^N N^{-1}\mathrm{tr}(\dot{R}_\mu U)^N\right) = \exp\left(\frac{1}{2}\sum_{n=1}^{\infty} n^{-1}\zeta_\mu^n\right)$$

and therefore

$$N^{-1}\mathrm{tr}(\dot{R}_\mu U)^N = \sum_{2n_2+3n_3=N} 2^{n_3}\frac{(n_2+n_3-1)!}{n_2!n_3!}\left(\sum_{i<k} r_{\mu,ik}^2 u_i u_k\right)^{n_2}\left(\prod_j r_{\mu,j}u_j\right)^{n_3}$$

$$= \begin{cases} \sum\limits_{m_1+m_2+m_3=n}\left(\sum\limits_{m=0}^{\min(m_j)} \frac{2^{2m}(n-m-1)!}{(2m)!\prod_{j=1}^{3}(m_j-m)!}\right)\prod\limits_{j=1}^{3} r_{\mu,j}^{2m_j} u_j^{n-m_j}, \\ \sum\limits_{m_1+m_2+m_3=n-1}\left(\sum\limits_{m=0}^{\min(m_j)} \frac{2^{2m+1}(n-m-1)!}{(2m+1)!\prod_{j=1}^{3}(m_j-m)!}\right)\prod\limits_{j=1}^{3} r_{\mu,j}^{2m_j+1} u_j^{n-m_j} \end{cases} \quad (58)$$

for $N = 2n \geq 2$ (upper branch in (58)) and $N = 2n+1 \geq 3$ (lower branch in (58)), respectively. Then, the Taylor polynomials in (54) are computed recursively by

$$T_0 = 1,\ T_1 = 0,\ T_{n+1} = \frac{1}{2(n+1)}\sum_{m=1}^{n} T_{n-m}\sum_{\mu=1}^{\nu}\mathrm{tr}(-\dot{R}_\mu U)^{m+1},\ n \geq 1. \quad (59)$$

Utilizing Jensen's original trivariate series for the $\Gamma_3(R_1, ..., R_\nu)$-pdf (Section 3 by Jensen (1970)), we obtain from

$$\prod_{\mu=1}^{\nu}|I_p + \dot{R}_\mu U|^{-1/2} = \prod_{\mu=1}^{\nu}\sum_{n=0}^{\infty}\binom{n-1/2}{n}\zeta_\mu^n = \sum_{n=0}^{\infty} Q_n(U; \dot{R}_1, ..., \dot{R}_\nu)$$

with inhomogeneous polynomials

$$Q_n = \sum_{(n)}\prod_{\mu=1}^{\nu}\binom{n_\mu - 1/2}{n_\mu}\zeta_\mu^{n_\mu} = \sum_{2n\leq\sum_k n_k\leq 3n} q_n(n_1, n_2, n_3, \dot{R}_1, \ldots, \dot{R}_\nu)\prod_{k=1}^{3} u_k^{n_k}$$

computed recursively by

$$Q_0 = 1,\ Q_{n+1} = \frac{1}{2(n+1)}\sum_{m=0}^{n} Q_{n-m}\sum_{\mu=1}^{\nu}\zeta_\mu^{m+1},\ n \geq 0,$$

the series

$$\begin{aligned} F(x_1, x_2, x_3; R_1, \ldots, R_\nu) = &\prod_{k=1}^{3} G_{\nu/2}(x_k) + \\ &\sum_{n=1}^{\infty}\sum_{2n\leq\sum_k n_k\leq 3n} q_n(n_1, ..., n_p, \dot{R}_1, \ldots, \dot{R}_\nu) \times \\ &\prod_{k=1}^{3} G_{\nu/2+n_k}^{(n_k)}(x_k). \end{aligned} \quad (60)$$

The outer series $\sum_{n=1}^{\infty} S_n$ in (60) with the inner sums S_n viewed as single terms is convergent under the condition

$$\max_{1\leq\mu\leq\nu}\Big(\sup\{|\zeta(U;\dot{R}_\mu)| : U = \text{diag}(u_1,u_2,u_3),\ u_k = -it_k(1-it_k)^{-1},\ t_k\in\mathbb{R}\}\Big) < 1, \tag{61}$$

with $\zeta(U;\dot{R}_\mu)$ as in (57). For dimensions $p\geq 3$ see Theorem 4 by Royen (2013b). This condition is weaker than (56), but for numerical evaluations (55) is more suitable.

Three always absolutely convergent series for Jensen's $\Gamma_p(R_1,...,R_\nu)$-cdf follow from the general formulas in Theorem 1 of Royen (2013b) for convolutions of differently scaled multivariate gamma distributions ($\Gamma_p(\alpha_1,...,\alpha_\nu,\Sigma_1,...,\Sigma_\nu)$-distributions). The Lt in (16) can be written in three ways:

$$\prod_{\mu=1}^{\nu}|I_p+R_\mu T|^{-1/2} = \prod_{k=1}^{p} z_k^{\nu/2}\prod_{\mu=1}^{\nu} c_\mu^{1/2}|I_p - C_\mu Y|^{-1/2},\ Y=\text{diag}(y_1,...,y_p), \tag{62}$$

where, alternatively,

$$Y=U,\ C_\mu = I_p - vR_\mu,\ c_\mu = 1, \tag{63a}$$
$$Y=Z,\ C_\mu = I_p - (vR_\mu)^{-1},\ c_\mu = |I_p - C_\mu|, \tag{63b}$$
$$Y=\Omega,\ C_\mu = I_p - 2(I_p+vR_\mu)^{-1},\ c_\mu = |I_p - C_\mu|, \tag{63c}$$

and $z_k = (1+v^{-1}t_k)^{-1}$, $u_k = 1-z_k$, $\omega_k = z_k - u_k$, $k=1,\ldots,p$. If R_μ has the eigenvalues $\lambda_{\mu,k}$ we obtain with

$$\lambda_{\min} = \min\{\lambda_{\mu,k};\ 1\leq\mu\leq\nu,\ 1\leq k\leq p\} > 0,\ \lambda_{\max} = \max\lambda_{\mu,k},\ \vartheta = \lambda_{\min}/\lambda_{\max} \tag{64}$$

that, respectively,

$$\max||C_\mu|| = (1-\vartheta)/(1+\vartheta) \text{ with } v = 2(\lambda_{\min}+\lambda_{\max})^{-1} \tag{65a}$$
$$\max||C_\mu|| = (1-\vartheta)/(1+\vartheta) \text{ with } v = (\lambda_{\min}^{-1}+\lambda_{\max}^{-1})/2 \tag{65b}$$
$$\max||C_\mu|| = (1-\vartheta^{1/2})/(1+\vartheta^{1/2}) \text{ with } v = (\lambda_{\min}\lambda_{\max})^{-1/2}. \tag{65c}$$

Remark 6.1. *With suitable values of the scaling factors v also $C_\mu\geq 0$ and $||C_\mu||<1$ can be accomplished.*

For $p=3$, we now obtain from $-\frac{1}{2}\sum_{\mu=1}^{\nu}\log|I_3 - C_\mu Y| = \sum_{m=1}^{\infty}\frac{1}{2m}\sum_{\mu=1}^{\nu}\text{tr}(C_\mu Y)^m$ and by setting Y, C_μ and c_μ as in (63a), (63b) and (63c), respectively, the three series

$$\prod_{\mu=1}^{\nu}|I_3+R_\mu T|^{-1/2} = \left(\prod_{\mu=1}^{\nu}c_\mu^{1/2}\right)\left(\prod_{k=1}^{3}z_k^{\nu/2}\right)\sum_{n=0}^{\infty}T_n(Y;C_1,...,C_\nu),$$
$$T_0 = 1, T_{n+1} = \frac{1}{2(n+1)}\sum_{m=0}^{n}T_{n-m}\sum_{\mu=1}^{\nu}tr(C_\mu Y)^{m+1},\ n\geq 0,$$

with Taylor polynomials

$$T_n = \sum_{(n)} t(n_1, n_2, n_3; C_1, ..., C_\nu) \prod_{k=1}^{3} y_k^{n_k},$$

and consequently the three corresponding always absolutely convergent series for the $\Gamma_3(R_1, ..., R_\nu)$-cdf, given by

$$F(x_1, x_2, x_3; R_1, ..., R_\nu) = \left(\prod_{\mu=1}^{\nu} c_\mu^{1/2}\right) \sum_{n=0}^{\infty} \sum_{(n)} t(n_1, n_2, n_3; C_1, ..., C_\nu) \prod_{k=1}^{3} F_{\nu/2, n_k}(vx_k), \tag{66}$$

where we respectively take

$$\begin{aligned} F_{\alpha,n} &= G_{\alpha+n}^{(n)} \text{ from (18)}, && (67a)\\ F_{\alpha,n} &= G_{\alpha+n} \text{ from (17)}, && (67b)\\ F_{\alpha,n} &= H_{\alpha,n} \text{ from (19)}. && (67c) \end{aligned}$$

For numerical evaluations a high number of the polynomials $\text{tr}(C_\mu Y)^m$ is available by computer algebra systems. Alternatively, we can also obtain the polynomials

$$\text{tr}(CY)^{n+1} = (n+1)T_{n+1} - \sum_{m=0}^{n-1} T_{n-m}\text{tr}(CY)^{m+1}$$

recursively from

$$|I_3 - CY|^{-1} = \left[1 - \left(tr(CY) - \sum_{i<k} |C_{ik}| y_i y_k + |C| y_1 y_2 y_3\right)\right]^{-1} = \sum_{n=0}^{\infty} T_n(Y; 1, C),$$

where $|C_{ik}| = c_{ii}c_{kk} - c_{ik}^2$, $T_0 = 1$, $T_1 = \text{tr}(CY)$,

$$\begin{aligned} T_2 &= (\text{tr}(CY))^2 - \sum_{i<k} |C_{ik}| y_i y_k = \sum_{k=1}^{3} c_{kk}^2 y_k^2 + \sum_{i<k} (c_{ii}c_{kk} + c_{ik}^2) y_i y_k,\\ T_{n+1} &= T_n \text{tr}(CY) - T_{n-1} \sum_{i<k} |C_{ik}| y_i y_k + T_{n-2} |C| y_1 y_2 y_3, \ n \geq 2. \end{aligned}$$

More explicitly, but less useful, we have for $p = 3$ that

$$\begin{aligned} T_n(Y; 1, C) &= \\ &\sum_{n_1+2n_2+3n_3=n} (-1)^{n_2} \frac{(n_1+n_2+n_3)!}{n_1! n_2! n_3!} (\text{tr}(CY))^{n_1} \left(\sum_{i<k} |C_{ik}| y_i y_k\right)^{n_2} (|C| y_1 y_2 y_3)^{n_3} \end{aligned}$$

and $n^{-1}\text{tr}(CY)^n$ is obtained if $(n_1 + n_2 + n_3)!$ is replaced by $(n_1 + n_2 + n_3 - 1)!$.

In particular, for $\nu = 2$ degrees of freedom a bivariate integral representation for Jensen's p-variate gamma cdf is obtained as a special case from Theorem 3 by Royen (2013b) with one-factorial $(p\times p)$-correlation matrices $R_\mu = \text{diag}(\ldots,1-a_{\mu,j}^2,\ldots)+a_\mu a_\mu^\top$, $\mu=1,2$. Then

$$\begin{aligned}F(x_1,...,x_p;R_1,R_2) &= \int\limits_0^\infty\int\limits_0^\infty\left[\prod_{j=1}^p \exp\left(-\frac{a_{1j}^2 y_1}{1-a_{1j}^2}-\frac{a_{2j}^2 y_2}{1-a_{2j}^2}\right)\times\right.\\ &\qquad \left.\sum_{n=0}^\infty P_n(y_1,y_2;a_{1j}^2,a_{2j}^2)G_{1+n}\left(\frac{x_j}{\min(1-a_{1j}^2,1-a_{2j}^2)}\right)\right]\prod_{\mu=1}^2 g_{1/2}(y_\mu)dy_\mu\\ &= \frac{4}{\pi}\int\limits_0^\infty\int\limits_0^\infty\left[\prod_{j=1}^p\sum_{n=0}^\infty P_n(y_1^2,y_2^2;a_{1j}^2,a_{2j}^2)\times\right.\\ &\qquad \left. G_{1+n}\left(\frac{x_j}{\min(1-a_{1j}^2,1-a_{2j}^2)}\right)\right]\prod_{\mu=1}^2\exp(-\lambda_\mu y_\mu^2)dy_\mu, \end{aligned}\tag{68}$$

where $\lambda_\mu = 1+\sum\limits_{j=1}^p a_{\mu,j}^2/(1-a_{\mu,j}^2)>0$ and

$$\begin{aligned}P_n(y_1,y_2;a_{1j}^2,a_{2j}^2) &= q_j^{1/2}\sum_{0\le k_1+k_2\le n}\binom{n-k_1-1/2}{n-k_1-k_2}\frac{q_j^{k_2}(1-q_j)^{n-k_1-k_2}}{k_1!k_2!}\times\\ &\qquad \left(\frac{a_{1j}^2 y_1}{1-a_{1j}^2}\right)^{k_1}\left(\frac{a_{2j}^2 y_2}{1-a_{2j}^2}\right)^{k_2}\end{aligned}$$

if $q_j=\min(1-a_{1j}^2,1-a_{2j}^2)/\max(1-a_{1j}^2,1-a_{2j}^2)=(1-a_{1j}^2)/(1-a_{2j}^2)$.
Otherwise, $\binom{n-k_1-1/2}{n-k_1-k_2}q_j^{k_2}$ has to be replaced by $\binom{n-k_2-1/2}{n-k_1-k_2}q_j^{k_1}$.

6.4 A series for the generalized multivariate chi-square distribution from Definition 5.1

The p-variate generalized chi-square distribution from Definition 5.1 has the Lt

$$|I_\nu+2RT|^{-1/2}\tag{69}$$

with $\nu=\nu_1+...+\nu_p$, $T=t_1I_{\nu_1}\oplus\ldots\oplus t_pI_{\nu_p}$ and a regular $(\nu\times\nu)$-correlation matrix R with diagonal blocks $R_{ii}=I_{\nu_i}$ and off-diagonal blocks R_{ik}.

For $p=2$ and $\nu_1\le\nu_2$ the corresponding cdf is given by

$$\begin{aligned}\Pr\{\chi_1^2\le x_1,\chi_2^2\le x_2;R_{12}\} &= F\left(\frac{x_1}{2},\frac{x_2}{2};r_{12,1},...,r_{12,\nu_1}\right) \\ &= \sum_{n=0}^\infty a_n(r_{12,1}^2,...,r_{12,\nu_1}^2)G_{\nu_1/2+n}^{(n)}\left(\frac{x_1}{2}\right)G_{\nu_2/2+n}^{(n)}\left(\frac{x_2}{2}\right),\end{aligned}\tag{70}$$

with the canonical correlations $r_{12,1},...,r_{12,\nu_1}$ of R_{12} which are the roots of the eigenvalues of $R_{12}R_{21}$, and the coefficients a_n as in (49); see, e. g., Section 5 of Royen (2013b). For $\nu_1=\nu_2$ this distribution coincides with a bivariate chi-square distribution of Jensen's type.

For $p = 3$ a series for the cdf, suitable for actual computations, seems to be feasible at most for very small degrees of freedom ν_j, for instance for the distribution discussed at the end of Section 4 where $\nu_1 = \nu_2 = \nu_3 = 2$. For an approximation see the end of Section 6.5.

With $C = I_\nu - vR$, $||C|| < 1$, (e. g., $v = 2(\lambda_{\min} + \lambda_{\max})^{-1}$ with the maximal and the minimal eigenvalue of R), the principal minor arrays C_J of C with indices i, j from J, where $\emptyset \neq J \subseteq \{1, ..., \nu\}$ and $U = u_1 I_{\nu_1} \oplus u_2 I_{\nu_2} \oplus u_3 I_{\nu_3}$, $u_j = 2v^{-1}t_j(1 + 2v^{-1}t_j)^{-1}$, we obtain from the polynomial

$$|I_\nu - CU| = 1 - \sum_{k=1}^{\nu} D_k(u_1, u_2, u_3; C),\ D_k = (-1)^{k-1} \sum_{\text{size}(J)=k} |C_J| \cdot |U_J|$$

the Taylor polynomials $P_n = \sum_{k=1}^{\min(\nu,n)} D_k P_{n-k}$ in the expansion

$$|I_\nu - CU|^{-1} = 1 + \sum_{n=1}^{\infty} P_n(u_1, u_2, u_3; C),$$

and then again recursively the Taylor polynomials

$$T_n(u_1, u_2, u_3; C) = \sum_{(n)} t(n_1, n_2, n_3; C) \prod_{j=1}^{3} u_j^{n_j}$$

in

$$|I_\nu - CU|^{-1/2} = 1 + \sum_{n=1}^{\infty} T_n(u_1, u_2, u_3; C),$$

namely by

$$T_1 = \frac{1}{2}P_1,\ T_{2m} = \frac{1}{2}(P_{2m} - T_m^2) - \sum_{k=1}^{m-1} T_k T_{2m-k},\ T_{2m+1} = \frac{1}{2}P_{2m+1} - \sum_{k=1}^{m} T_k T_{2m+1-k}.$$

Transforming the Lt given in (69) into $\left(\prod_{j=1}^{3} z_j^{\nu_j/2}\right) |I_\nu - CU|^{-1/2}$ with $z_j = (1 + 2v^{-1}t_j)^{-1}$, the absolutely convergent series

$$\Pr\{\chi_1^2 \leq x_1, \chi_2^2 \leq x_2, \chi_3^2 \leq x_3; R\} = \sum_{n=0}^{\infty} \sum_{(n)} t(n_1, n_2, n_3; C) \prod_{j=1}^{3} G_{\nu_j/2+n_j}^{(n_j)} \left(\frac{v}{2} x_j\right) \quad (71)$$

follows by inversion and termwise integration.

6.5 Some approximations

In this section we are mainly interested in approximations for small exceedance probabilities $p = 1 - F_p(x, \dots, x; \alpha, R)$. Notice that setting $x = t_i$ leads to the multiplicity-adjusted p-values defined in (1). For the more general distributions from Sections 6.3 and 6.4 only the three-variate case is considered here. Classical conservative approximations make use of refined Bonferroni inequalities with marginal distributions of low orders or sometimes - under the required assumptions - the “product-type probability bounds“ of Glaz and Johnson (1984) and Block et al. (1992), cf. our respective remarks at the end of Section 4. However, frequently comparatively large differences between the true p-values and the approximated ones resulting from the aforementioned bounds are the price for guaranteed conservatism (i. e., strict control of the multiple type I error level). Here, a different method from Royen (2013a) is presented, providing very accurate approximations of the probability p for many correlation matrices R. It is based on the following two formulas (73) and (75) for the $\Gamma_p(\alpha, R)$-cdf. Let R be the special regular $(p \times p)$-correlation matrix

$$\begin{pmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{pmatrix} \tag{72}$$

with a $(q \times q)$-submatrix R_{11} for $q \geq 2$, $p - q \geq 2$, and with identical correlations $r_1 > 0$, $r_2 > 0$ and r within the corresponding submatrices R_{11}, R_{22} and R_{12}. If $r^2 \leq r_1 r_2$ then

$$\begin{aligned} F(x_1, ..., x_p; \alpha, R) &= \sum_{k=0}^{\infty} \binom{\alpha + k - 1}{k}^{-1} \frac{r^{2k}}{(r_1 r_2)^k} \times \\ & \int_0^{\infty} \left[\prod_{j=1}^{q} G_\alpha \left(\frac{x_j}{1 - r_1}, \frac{r_1 y}{1 - r_1} \right) \right] L_k^{(\alpha-1)}(y) g_\alpha(y) dy \times \\ & \int_0^{\infty} \left[\prod_{j=q+1}^{p} G_\alpha \left(\frac{x_j}{1 - r_2}, \frac{r_2 y}{1 - r_2} \right) \right] L_k^{(\alpha-1)}(y) g_\alpha(y) dy. \end{aligned} \tag{73}$$

Now, we regard $F(x_1, \ldots, , x_p; \alpha, R)$, $p \geq 4$, as a function of the (general) correlation matrix $R = (r_{ik}) = R_0 + H$, $H = (h_{ik})$, and approximate it by a Taylor polynomial of second degree with a one-factorial correlation matrix R_0 if $(p(p-1))^{-1}\mathrm{tr}(H^2) \leq h^2_{\max}$ with a sufficiently small value $h^2_{\max}$, e. g., $h^2_{\max} \leq 0.01$. For the computation of R_0, the general Taylor formula and numerical examples up to dimension $p = 10$ see Royen (2013a). Here only the special case is given where

$$R_0 = (1 - r) I_p + r \vec{1}\vec{1}^\top, \; \vec{1} = (1, \ldots, 1)^\top, \; r = \frac{2}{p(p-1)} \sum_{1 \leq i < k \leq p} r_{ik}, \tag{74}$$

and consequently

$$\begin{aligned} \left(\sum_{1 \leq i < k \leq p} h_{ik} \right)^2 &= H_2 + H_3 + H_4 = 0, \; H_2 = \sum_{i<k} h_{ik}^2, \\ H_4 &= \sum_{\substack{i<k, \ell<m \\ \{i,k\} \cap \{\ell,m\} = \emptyset}} h_{ik} h_{\ell m}, \; H_3 = -H_2 - H_4. \end{aligned}$$

Then, with identical $x_j = x$,

$$F := G_\alpha \left(\frac{x}{1 - r}, \frac{ry}{1 - r} \right), \; f_n := \frac{\partial^n}{\partial x^n} G_{\alpha+n} \left(\frac{x}{1 - r}, \frac{ry}{1 - r} \right), \; n = 1, 2,$$

and the coefficients

$$
\begin{aligned}
a &= \int_0^\infty (\alpha f_1^2 - 2ryf_1f_2 + 2r^2y^2f_2^2)F^{p-2}g_\alpha(y)dy, \\
b &= -r\int_0^\infty f_1^2(f_1 - 2ryf_2)F^{p-3}yg_\alpha(y)dy, \\
c &= 2r^2\int_0^\infty f_1^4F^{p-4}y^2g_\alpha(y)dy,
\end{aligned}
$$

we obtain the Taylor polynomial

$$
T_2(x;\alpha,r,H) = \int_0^\infty F^pg_\alpha(y)dy + aH_2 + bH_3 + cH_4 \tag{75}
$$

and the approximation

$$
F_p(x,\ldots,x;\alpha,R = R_0 + H) \approx T_2(x;\alpha,r,H), \tag{76}
$$

which is in particularly useful for larger values of x. Besides, $F_p(x,...,x;\alpha,R)$, where $\alpha \geq 1/2$ and $p \geq 3$, has - as a function of R - a local minimum at $R = R_0$ from (74), at least for $r \geq 0$ and $a + (p-4)b - (p-3)c > 0$, see Theorem 3 of Royen (2013a). Frequently, the latter condition can already be verified by a plot of the corresponding integrand.

Now suppose that R - possibly after a suitable renumbering of the variables - has a partitioned form as in (72) with two blocks $B_1 = \{1,...,b_1\}$, $B_2 = \{b_1+1,...,b_1+b_2 = p\}$ of indices, mean correlations $r_\mu > 0$ of $R_{\mu\mu}$, $\mu = 1,2$, and a mean correlation r of R_{12} with $r^2 \leq r_1r_2$. Then, with $P_{B_\mu}(x) := \mathrm{Pr}\{\max X_j \leq x, j \in B_\mu;\alpha,R_{\mu\mu}\}$ and

$$
c_k(x;\alpha,b_\mu,r_\mu) := \binom{\alpha+k-1}{k}^{-1}\int_0^\infty \left[G_\alpha\left(\frac{x}{1-r_\mu},\frac{r_\mu y}{1-r_\mu}\right)\right]^{b_\mu} L_k^{(\alpha-1)}(y)g_\alpha(y)dy,
$$

the approximation

$$
\begin{aligned}
F_p(x,\ldots,x;\alpha,R) &= P_{B_1\cup B_2}(x) \approx P_{B_1}(x)P_{B_2}(x) + \\
&\sum_{k=1}^\infty \binom{\alpha+k-1}{k}\frac{r^{2k}}{(r_1r_2)^k}c_k(x;\alpha,b_1,r_1)c_k(x;\alpha,b_2,r_2) \quad (77)
\end{aligned}
$$

with an always positive series is proposed if the $P_{B_\mu}(x)$ are computable by an exact representation. The corresponding hypothetical inequality with “$\geq$“ instead of “$\approx$“ seems to be frequently satisfied, but presumably it holds true only under additional assumptions. It should be noted, that not even the inequality $P_{B_1\cup B_2}(x) \geq P_{B_1}(x)P_{B_2}(x)$, following from the famous Gaussian correlation conjecture, has been proved until now for all R.

If $P_{B_\mu}(x)$ in (77) is not computable by an exact representation then it should be replaced by the best available approximation $\tilde{P}_{B_\mu}(x)$. This can be a Taylor approximation as in (76) if the correlations in $R_{\mu\mu}$ have a sufficiently low variability or the more general Taylor approximation with a one-factorial $R_{0\mu\mu}$ in $R_{\mu\mu} = R_{0\mu\mu} + H_{\mu\mu}$. Royen (2007) has described such Taylor approximations with a "low-factorial" R_0 in a general form, but even the search for a two-factorial approximation R_0 of R and in particular the computation of the resulting Taylor polynomials by three-variate integrals is rather intricate. However, very frequently a (5×5)-correlation matrix R can be approximated rather accurately by a two-factorial R_0. Then, the additional correction terms of the Taylor polynomial may be dropped in many cases.

The approximation $\tilde{P}_{B_\mu}(x)$ can also be obtained by a further decomposition of B_μ into two smaller blocks $B_{\mu 1}$, $B_{\mu 2}$ and application of (77) to $\tilde{P}_{B_{\mu 1}}$, $\tilde{P}_{B_{\mu 2}}$ or $P_{B_{\mu 1}}$, $P_{B_{\mu 2}}$, but more than one such iterated application of (77) is not recommended in general.

The $\Gamma_P(\alpha, R)$-distribution can also be used for an approximation of Jensen's $\Gamma_p(R_1, ..., R_\nu)$-distribution. A generalization of inequality (50) is not proved here, but for $p = 3$ and at least for correlation matrices R_μ with $r_{\mu,12}r_{\mu,13}r_{\mu,23} > 0$, $1 \leq \mu \leq \nu$, the approximation

$$p := 1 - F(x, x, x; R_1, ..., R_\nu) \approx 1 - F(x, x, x; \nu/2, R) \tag{78}$$

with $R = (r_{ik})$, $r_{ik} = \left(\nu^{-1} \sum_{\mu=1}^{\nu} r_{\mu,ik}^2\right)^{1/2}$, $i \neq k$, is recommended for small exceedance probabilities. That R is always a correlation matrix follows from the convexity of the intersection of the unit cube with the body whose surface is determined by the singular correlation matrices C with correlations $\sqrt{x}$, $\sqrt{y}$, $\sqrt{z}$ and the equation $|C| = 1 + 2\sqrt{xyz} - x - y - z = 0$.

For the more general three-variate χ^2-distribution with Lt given in (69), now with $\nu_1 = \nu_2 = \nu_3 = \nu$ and the $(3\nu \times 3\nu)$-correlation matrix R with the diagonal blocks $R_{ii} = I_\nu$ and the off-diagonal blocks R_{ik}, $1 \leq i, k \leq 3$, $i \neq k$, small values of Pr$\{\max \chi_j^2 > x, j = 1, 2, 3; \nu, R\}$ can be approximated by $1 - F(x/2, x/2, x/2; R_1, ..., R_\nu)$, provided that all the symmetrical (3×3)-matrices $R_\mu = (r_{\mu,ik})$, $r_{\mu,ii} = 1$ containing the canonical correlations (i. e., singular values) $r_{\mu,ik}$ of the R_{ik}, are positive definite. Afterwards, approximation (78) can be applied.

7 Concluding Remarks

We have demonstrated the relevance of three types of multivariate chi-square distributions for a variety of multiple test problems. Our computational methods allow for addressing these problems by multivariate methods, meaning that the joint (limiting) distribution of test statistics is employed for the calibration of a multiple test with respect to type I error control, rather than just the marginal distributions like, for instance, in a Bonferroni or a Šidák correction. Since chi-square distributed random variables are necessarily non-negatively correlated, it is to be expected that the utilization of their joint distribution will typically lead to a better exhaustion of the FWER level and, due to the structure of the decision rule, to higher power in comparison with the latter margin-based approaches.

Up to present, the drawback of the computational methods described in Sections 6.2 - 6.4 is that they are only feasible for low dimensions $p \leq 4$. However, statistical methodology for multiple test problems with $m > 4$ hypotheses can also profit from these methods. Namely, the m-variate limiting joint distribution of test statistics under the global hypothesis can be approximated conservatively by probability bounds of sum- or product-type, as outlined in Section 4 and Section 6.5, which only require the computation of p-variate marginal chi-square probabilities for $p < m$. For instance, computer simulations (not shown here) under the model considered in Section 4 indicate that product-type probability bounds of order 4 in the sense of Block et al. (1992) often approximate the true p-value already markedly tighter than a simple Bonferroni or Šidák correction, at least if pronounced correlations are present among test statistics. Such pronounced correlations typically occur in modern applications from the life sciences like genetic association studies (strong linkage disequilibrium among genetic markers), gene expression studies (co-activation of several genes), or functional magnetic resonance imaging (highly correlated voxels within regions of interest); cf., e. g., Part II of Dickhaus (2014) and references therein. Furthermore, formula (77) - in combination with the Taylor approximation from (76) where appropriate - has turned out to be accurate for small exceedance probabilities at least for dimensions $p \leq 12$, albeit conservativity in terms of strict FWER control is not guaranteed.

Future work will aim at implementing the formulas given in Section 6 into easy-to-access software for practitioners, with special emphasis on user-friendlyness.

References

Anderson, T., 1984. An introduction to multivariate statistical analysis. 2nd ed. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons, Inc.

Bapat, R., 1989. Infinite divisibility of multivariate gamma distributions and M-matrices. Sankhyā, Ser. A 51 (1), 73–78.

Block, H. W., Costigan, T., Sampson, A. R., 1992. Product-type probability bounds of higher order. Probab. Eng. Inf. Sci. 6 (3), 349–370.

Blumenson, L., Miller, K., 1963. Properties of generalized Rayleigh distributions. Ann. Math. Stat. 34, 903–910.

Dickhaus, T., 2012. Simultaneous Statistical Inference in Dynamic Factor Models. SFB 649 Discussion Paper 2012-033, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany, available at http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2012-033.pdf.

Dickhaus, T., 2014. Simultaneous Statistical Inference with Applications in the Life Sciences. Springer, forthcoming.

Dickhaus, T., Stange, J., 2013. Multiple point hypothesis test problems and effective numbers of tests for control of the family-wise error rate. Calcutta Statistical Association Bulletin, to appear.

Gabriel, K. R., 1969. Simultaneous test procedures - some theory of multiple comparisons. Ann. Math. Stat. 40, 224–250.

Glaz, J., Johnson, B. M., 1984. Probability inequalities for multivariate distributions with dependence structures. J. Am. Stat. Assoc. 79, 436–440.

Griffiths, R., 1984. Characterization of infinitely divisible multivariate gamma distributions. J. Multivariate Anal. 15, 13–20.

Jensen, D. R., 1970. A generalization of the multivariate Rayleigh distribution. Sankhyā, Series A 32 (2), 193–208.

Katayama, N., 2008. Portmanteau Likelihood Ratio Tests for Model Selection. Discussion Paper Series 2008-1, Faculty of Economics, Kyushu University.

Kotz, S., Balakrishnan, N., Johnson, N. L., 2000. Continuous multivariate distributions. Vol. 1: Models and applications, 2nd Edition. New York, NY: Wiley.

Krishnaiah, P., Rao, M., 1961. Remarks on a multivariate gamma distribution. Am. Math. Mon. 68, 342–346.

Krishnamoorthy, A., Parthasarathy, M., 1951. A multivariate gamma-type distribution. Ann. Math. Stat. 22, 549–557.

Kruskal, W. H., Wallis, W., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47, 583–621.

Lehmann, E. L., Romano, J. P., 2005. Testing statistical hypotheses. 3rd ed. Springer Texts in Statistics. New York, NY: Springer.

Marcus, R., Peritz, E., Gabriel, K. R., 1976. On closed test procedures with special reference to ordered analysis of variance. Biometrika 63 (3), 655–660.

Moskvina, V., Schmidt, K. M., 2008. On multiple-testing correction in genome-wide association studies. Genetic Epidemiology 32, 567–573.

Olver, F. W. J., Lozier, D. W., Boisvert, R. F., Clark, C. W. (Eds.), 2010. NIST handbook of mathematical functions. Cambridge: Cambridge University Press.

Puri, M. L., Sen, P. K., 1971. Nonparametric methods in multivariate analysis. Wiley Series in Probability and Mathematical Statistics. New York etc.: John Wiley & Sons, Inc.

Royen, T., 1984. Multiple comparisons of polynomial distributions. Biom. J. 26, 319–332.

Royen, T., 1989. Generalized maximum range tests for pairwise comparisons of several populations. Biom. J. 31 (8), 905–929.

Royen, T., 1990. Tables for studentized multivariate maximum ranges and their application by maximum range tests. Biom. J. 32 (6), 643–680.

Royen, T., 1991. Expansions for the multivariate chi-square distribution. J. Multivariate Anal. 38 (2), 213–232.

Royen, T., 1994. On some multivariate gamma distributions connected with spanning trees. Ann. Inst. Stat. Math. 46 (2), 361–371.

Royen, T., 1995. On some central and non-central multivariate chi-square distributions. Stat. Sin. 5 (1), 373–397.

Royen, T., 1997. Multivariate gamma distributions. In: Samuel Kotz, Campbell B. Read, David L. Banks (eds.): Encyclopedia of Statistical Sciences, Update Volume 1. Wiley-Interscience, pp. 419 – 425.

Royen, T., 2007. Integral representations and approximations for multivariate gamma distributions. Ann. Inst. Stat. Math. 59 (3), 499–513.

Royen, T., 2013a. Some upper tail approximations for the distribution of the maximum of correlated chi-square or gamma random variables. Far East J. Theor. Stat. 43 (1), 27–56.

Royen, T., 2013b. Some representations for convolutions of multivariate gamma distributions. Far East J. Theor. Stat. 45, 21–49.

Sarkar, S. K., Chang, C.-K., 1997. The Simes method for multiple hypothesis testing with positively dependent test statistics. J. Am. Stat. Assoc. 92 (440), 1601–1608.

Sen, P. K., 1983. On permutational central limit theorems for general multivariate linear rank statistics. Sankhyā, Ser. A 45, 141–149.

Simes, R. J., 1986. An improved Bonferroni procedure for multiple tests of significance. Biometrika 73, 751–754.

Timm, N. H., 2002. Applied multivariate analysis. New York, NY: Springer.

Wilks, S., 1938. The large-sample distribution of the likelihood ratio for testing composite hypotheses. Ann. Math. Stat. 9, 60–62.