ROGER ALMEIDA PEREIRA MELO

TESTE DA RAZÃO DE VEROSSIMILHANÇAS PARA A VARIÂNCIA GENERALIZADA NORMAL MULTIVARIADA

LAVRAS – MG

2020

ROGER ALMEIDA PEREIRA MELO

TESTE DA RAZÃO DE VEROSSIMILHANÇAS PARA A VARIÂNCIA GENERALIZADA NORMAL MULTIVARIADA

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Prof. Dr. Daniel Furtado Ferreira
Orientador

LAVRAS – MG
2020

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Melo, Roger Almeida Pereira
Teste da razão de verossimilhanças para a variância generalizada normal multivariada / Roger Almeida Pereira Melo. – 2020.
65 p.

Orientador(a): Daniel Furtado Ferreira.

Dissertação (mestrado acadêmico)–Universidade Federal de Lavras, 2020.
Bibliografia.

1. Distribuição Wishart. 2. Método Delta. 3. Teste de Hipóteses. I. Ferreira, Daniel Furtado. II. Título.

ROGER ALMEIDA PEREIRA MELO

TESTE DA RAZÃO DE VEROSSIMILHANÇAS PARA A VARIÂNCIA GENERALIZADA NORMAL MULTIVARIADA

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 7 de Fevereiro de 2020.

Prof. Dr. Renato Ribeiro de Lima UFLA
Prof. Dr. Deive Ciro de Oliveira UNIFAL-MG

Prof. Dr. Daniel Furtado Ferreira
Orientador

LAVRAS – MG
2020

AGRADECIMENTOS

Ao meu orientador, Daniel Ferreira Furtado, pela orientação e toda ajuda prestada ao longo do mestrado.

Aos professores membros da banca, de qualificação e defesa, por contribuírem para o desenvolvimento desse trabalho.

Aos meus pais e irmãos, por me apoiarem sempre.

Ao Departamento de Estatística e à Universidade Federal de Lavras.

Aos meus amigos e demais colegas do Departamento.

A todas as pessoas que de algum modo contribuíram para que eu pudesse concluir o mestrado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

RESUMO

Uma medida interessante da variabilidade na população multivariada é o determinante da matriz de covariâncias $\mathbf{\Sigma}_{p\times p}$, $|\mathbf{\Sigma}|$, de uma população, conhecida como variância generalizada. Esta, é uma medida que resume a dispersão de uma população multivariada em um único valor, considerando a dependência entre as variáveis envolvidas. Devido a isso, possui aplicações em diversas áreas que almejam avaliar a dispersão existente em alguma população multivariada de interesse. Nas indústrias, por exemplo, existem diversas situações nas quais o monitoramento ou controle simultâneo de duas ou mais características relacionadas ao processo de qualidade são necessárias. Por assim dizer, avaliar se o processo está, estatisticamente, sob controle, consiste em analisar conjuntamente todas as variáveis relacionadas ao processo de qualidade, considerando a dependência entre as mesmas, como também a sua variabilidade. Além da indústria, o estudo da variabilidade multivariada através da variância generalizada está presente em processamento de sinal, análise de agrupamento, delineamentos ótimos e muitos outros campos. Deste modo, a construção de teste de hipóteses que avaliem a dispersão em populações multivariadas se fazem necessárias dado o seu amplo campo de atuação. Este trabalho divide-se em duas partes. A primeira, trata-se de uma revisão bibliográfica que engloba toda a teoria necessária para compreender a construção de um teste de hipótese para a variância generalizada da distribuição normal multivariada, que consiste na segunda parte e foi apresentada no formato de artigo. O artigo trata da proposição de dois novos testes de hipóteses, um construído via razão de verossimilhanças- o teste LRT- e o outro, também é construído via razão de verossimilhanças, porém é acrescido da Correção de Bartlett para testes da razão de verossimilhanças, denominado BCLRT. Tais testes de hipóteses são contruídos para testar a variância generalizada de uma população normal multivariada. Para a avaliação da taxa de erro tipo I e do poder dos testes, realizam-se simulações de Monte Carlo para diversos cenários em que são variados o tamanho da amostra n, o número de variáveis p e o nível de significância α para os testes propostos e para outros testes já existentes na literatura. Os desempenhos dos testes propostos nas avaliações da taxa de erro do tipo I e poder nos levaram a recomendação da utilização do teste BCLRT somente em cenários em que temos $p = 2$, sobretudo quando $n > 30$. Enquanto que para o teste LRT, recomendamos sua utilização em situações em que $p = 2$ e $p = 3$ para $n > 30$ e para $p = 5$ quando $n > 50$.

Palavras-chave: Correção de Bartlett, Distribuição Wishart, Método Delta, Teste de Hipóteses.

ABSTRACT

An interesting measure of variability in the multivariate population is the determinant of the covariance matrix $\mathbf{\Sigma}_{p\times p}$, $|\mathbf{\Sigma}|$, of a population, known as generalized variance. This is a measure that summarizes the dispersion of a multivariate population in a single value, considering the dependence between the variables involved. Because of this, it has applications in several areas that aim to evaluate the dispersion existing in some multivariate population of interest. In industries, for example, there are several situations in which the simultaneous monitoring or control of two or more characteristics related to the quality process is necessary. So to say, assessing whether the process is, statistically, under control, consists of jointly analyzing all the variables related to the quality process, considering the dependence between them, as well as their variability. In addition to the industry, the study of multivariate variability through generalized variance is present in signal processing, cluster analysis, optimal designs and many other fields. In this way, the construction of hypothesis tests that evaluate the dispersion in multivariate populations is necessary given its wide field of action. This work is divided into two parts. The first is a bibliographic review that encompasses all the theory necessary to understand the construction of a hypothesis test for the generalized variance of the multivariate normal distribution, which consists of the second part and was presented in article format. The article deals with the proposition of two new hypothesis tests, one built via the likelihood ratio - the LRT test - and the other, it is also built via the likelihood ratio, however it is added Bartlett's Correction for likelihood ratio tests, called BCLRT. Such hypothesis tests are designed to test the generalized variance of a normal multivariate population. For the evaluation of the type I error rate and the power of the tests, Monte Carlo simulations are performed for different scenarios in which are varied the sample size n, the number of variables p and the level of significance α for the proposed tests and for other tests already in the literature. The performance of the tests proposed in the evaluations of the type I error rate and power led us to recommend the use of the BCLRT test only in scenarios where we have $p = 2$, especially when $n > 30$. While for the LRT test, we recommend its use in situations where $p = 2$ and $p = 3$ for $n > 30$ and for $p = 5$ when $n > 50$.

Keywords: Bartlett's Correction, Wishart Distribution, Delta Method, Hypothesis Test.

LISTA DE FIGURAS

SUMÁRIO

1 INTRODUÇÃO

Uma medida interessante da variabilidade na população multivariada é o determinante da matriz de covariâncias $\mathbf{\Sigma}_{p\times p}$, $|\mathbf{\Sigma}|$, de uma população, conhecida como variância generalizada. Essa medida, como apontado por Najarzadeh (2019), também é amplamente utilizada, por exemplo, no gráfico de controle multivariado (YEH et al., 2003; YEH; J.; McGRATH, 2006; BERSIMIS; PSARAKIS; PANARETOS, 2007; DJAUHARI, 2005; DJAUHARI; MASHURI; HERWINDIATI, 2008; NOOR; DJAUHARI, 2014; LEE; KHOO, 2017), modelagem em análise de confiabilidade (TALLIS; LIGHT, 1968), processamento de sinal (BHANDARY, 1996), análise de agrupamentos (GUPTA, 1982), delineamentos ótimos (PUKELSHEIM, 2006) e alocação ótima na amostragem estratificada (ARVANITIS; AFONJA, 1971). A variância generalizada é uma medida do hipervolume que a distribuição das variáveis aleatórias ocupa no espaço p-dimensional. Uma medida correlacionada é a variância generalizada padronizada, que permite comparações de conjuntos de diferentes dimensões, definida pelas médias geométricas dos autovalores de $\mathbf{\Sigma}$, ou seja $|\mathbf{\Sigma}|^{1/p}$ (SENGUPTA, 1987a; SENGUPTA, 1987b).

Vários procedimentos de intervalos de confiança e testes de hipóteses foram considerados Jafari e Kazemi (2014). Eles avaliaram o desempenho dos procedimentos estudados utilizando simulações de Monte Carlo. Não foi encontrada na literatura científica menção a respeito do teste da razão de verossimilhanças para a variância generalizada de uma população normal. Por outro lado, Najarzadeh (2017), Najarzadeh (2019) propõe testes para comparações e intervalos de confiança para produtos de variâncias generalizadas (padronizadas).

O foco deste trabalho, é construir o teste da razão de verossimilhança para a hipótese nula $H_0 : |\mathbf{\Sigma}| = \eta$ sob normalidade multivariada. Além disso, pretende-se construir o teste da razão de verossimilhanças passo a passo e mostrar a teoria da distribuição da variância generalizada multivariada da amostra normal.

2 OBJETIVOS

2.1 Objetivo Geral

Construir um teste da razão de verossimilhança (LRT) e outro teste da razão de verossimilhança com correção de Bartlett (BCLRT) para a hipótese nula de $H_0 : |\mathbf{\Sigma}| = \eta$ sob normalidade multivariada.

2.2 Objetivos Específicos

- Construir os testes LRT e BCLRT passo a passo;
- Apresentar a teoria da distribuição da variância generalizada da amostra normal multivariada;
- Avaliar a taxa de erro do tipo I e o poder dos testes propostos através de simualação de Monte Carlo, bem como, compará-los aos testes já existentes na literatura.

3 REFERENCIAL TEÓRICO

Neste capítulo, serão abordadas algumas definições e resultados importantes.

3.1 Formas quadráticas

Seja $\mathbf{A}_{n\times n}$ uma matriz simétrica, sua forma quadrática é definida por (FERREIRA, 2018):

$$Q(x) = \mathbf{x}^\top \mathbf{A}\mathbf{x} = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i=1}^{n-1}\sum_{k=i+1}^{n} a_{ik}x_i x_k = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}x_i x_k,$$

em que $\mathbf{x} \neq \mathbf{0}$ é um vetor definido em $\mathbb{R}^n$.

Definição: Seja $\boldsymbol{A}_{p\times q}$ uma matriz, temos que $\text{vec}(\boldsymbol{T})$ é um vetor de dimensão $pq \times 1$, obtido através do empilhamento das colunas da matriz $\boldsymbol{T}$ uma sob a outra, isto é

$$\boldsymbol{T} = \begin{bmatrix} t_1 & t_2 & \dots & t_q \end{bmatrix},$$

em que t_i possui dimensão $p \times 1$, para $i = 1, 2, \dots, q$, então

$$\text{vec}(\boldsymbol{T}) = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_q \end{bmatrix}$$

3.2 Variáveis aleatórias multidimensionais

De acordo com Ferreira (2018), na realização de um experimento ou uma amostragem em que avaliam-se n indivíduos em p variáveis, os resultados são respostas mensuradas nas unidades amostrais ou experimentais. Quando o número de variáveis $p = 1$, tem-se o caso univariado. Para o caso multivariado $(p > 1)$, deve-se utilizar a notação de vetor aleatório para cada unidade amostral e cada vetor possuirá p componentes. Dessa forma, podemos representar uma amostra aleatória de tamanho n por um conjunto de vetores p-dimensionais

$$\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_j, \dots, \boldsymbol{X}_n.$$

O vetor n-dimensional da k-ésima variável aleatória é representada por

$$\boldsymbol{X}_k = \begin{bmatrix} X_{1k} \\ X_{2k} \\ \vdots \\ X_{jk} \\ \vdots \\ X_{nk} \end{bmatrix}$$

e o vetor p-dimensional da j-ésima unidade amostral por

$$\boldsymbol{X}_j^\top = \begin{bmatrix} X_{j1}, \, X_{j2}, \, \ldots, \, X_{jk}, \, \ldots, \, X_{jp} \end{bmatrix}.$$

Pode-se representar todo o conjunto multivariado avaliando n unidades experimentais em p variáveis , por uma matriz de dados $\boldsymbol{X}$, de dimensão $(n \times p)$, ou seja,

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \ldots & X_{1k} & \ldots & X_{1p} \\ X_{21} & X_{22} & \ldots & X_{2k} & \ldots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{j1} & X_{j2} & \ldots & X_{jk} & \ldots & X_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \ldots & X_{nk} & \ldots & X_{np} \end{bmatrix}, \tag{3.1}$$

em que X_{jk} é o valor da k-ésima variável aleatória na j-ésima unidade amostral.

A matriz de dados $\mathbf{X}$ representada em (3.1), torna cada linha da matriz um vetor p-dimensional, de observações multivariadas e a coluna da matriz um vetor n-dimensional, das n cópias de uma determinada variável.

Pode-se definir a esperança matemática para um vetor aleatório $\boldsymbol{X}$, para o caso contínuo, como

$$E(\boldsymbol{X}) = \begin{bmatrix} E(X_1) \\ \vdots \\ E(X_p) \end{bmatrix} = \begin{bmatrix} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_1 f(x_1, \ldots, x_p) dx_1 \ldots dx_p \\ \vdots \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_p f(x_1, \ldots, x_p) dx_1 \ldots dx_p \end{bmatrix} = \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}.$$

A covariância de um vetor aleatório é uma matriz definida por

$$\begin{aligned} Cov(\boldsymbol{X}) &= E(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^T \\ &= \begin{bmatrix} E\left[(X_1-\mu_1)^2\right] & \dots & E\left[(X_1-\mu_1)(X_p-\mu_p)\right] \\ E\left[(X_2-\mu_2)(X_1-\mu_1)\right] & \dots & E\left[(X_2-\mu_2)(X_p-\mu_p)\right] \\ \vdots & \ddots & \vdots \\ E\left[(X_p-\mu_p)(X_1-\mu_1)\right] & \dots & E\left[(X_p-\mu_p)^2\right] \end{bmatrix}, \end{aligned} \tag{3.2}$$

em que,

$$E(X_k-\mu_k)(X_l-\mu_l) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(x_k-\mu_k)(x_l-\mu_l)f(x_1,\dots,x_p)dx_1\dots dx_p,$$

para $k,l = 1,2,\dots,p$, é a covariância entre os vetores de variáveis aleatórias $\boldsymbol{X}_k$ e $\boldsymbol{X}_l$, também representada por σ_{kl}. A matriz definida em (3.2) é conhecida como *matriz de covariâncias populacional* e é dada por

$$Cov(\boldsymbol{X}) = \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}.$$

3.3 Função de verossimilhança

Seja a sequência de variáveis aleatórias independentes e identicamente distribuídas $\boldsymbol{X}_1$, $\boldsymbol{X}_2,\dots,$ $\boldsymbol{X}_j$, $\dots$, $\boldsymbol{X}_n$ uma amostra aleatória de tamanho n, sendo $\boldsymbol{X}_j$ uma variável aleatória p-dimensional com função densidade $f(\boldsymbol{x})$ (FERREIRA, 2018).

Dado uma amostra aleatória de uma população com uma distribuição $f(\boldsymbol{x};\boldsymbol{\theta})$, no qual $\boldsymbol{\theta}$ é o vetor de parâmetros, pode-se definir a função de verossimilhança, em geral denotada por $L(\boldsymbol{\theta};\boldsymbol{x}_1,\boldsymbol{x}_2,\dots,\boldsymbol{x}_n)$, sendo uma função de $\boldsymbol{\theta}$, e representada por

$$\begin{aligned} L(\boldsymbol{\theta};\boldsymbol{x}_1,\boldsymbol{x}_2,\dots,\boldsymbol{x}_n) &= f(\boldsymbol{x}_1,\boldsymbol{x}_2,\dots,\boldsymbol{x}_n;\boldsymbol{\theta}) = f(\boldsymbol{x}_1;\boldsymbol{\theta})f(\boldsymbol{x}_2;\boldsymbol{\theta})\times\cdots\times f(\boldsymbol{x}_n;\boldsymbol{\theta}) \\ &= \prod_{j=1}^{n} f(\boldsymbol{x}_j;\boldsymbol{\theta}). \end{aligned}$$

3.4 Estatísticas descritivas

Considere uma amostra aleatória $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_j, \ldots, \mathbf{X}_n$ em $\mathbb{R}^p$, de uma distribuição qualquer com média $\boldsymbol{\mu}$ e covariância comum $\boldsymbol{\Sigma}$. De acordo com Ferreira (2018), a média amostral, estimador de $\boldsymbol{\mu}$, é definida por

$$\bar{\boldsymbol{X}}_{.} = \frac{\sum_{j=1}^{n} \boldsymbol{X}_j}{n} = \frac{1}{n}\boldsymbol{X}^{\top}\mathbf{1} = \begin{bmatrix} \bar{X}_{.1} \\ \bar{X}_{.2} \\ \vdots \\ \bar{X}_{.k} \\ \vdots \\ \bar{X}_{.p} \end{bmatrix},$$

em que $\boldsymbol{X}$ é a matriz de dados de dimensões $(n \times p)$, $\mathbf{1}$ é um vetor de dimensões $(n \times 1)$ composto por elementos iguais a 1 e $\bar{X}_{.k} = \frac{\sum_{j=1}^{n} X_{jk}}{n}$ é a média amostral da k-ésima variável.

A matriz de covariâncias amostral ou estimador de $\boldsymbol{\Sigma}$ é dado por

$$\boldsymbol{S} = \begin{bmatrix} S_{11} & S_{12} & \ldots & S_{1p} \\ S_{21} & S_{22} & \ldots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \ldots & S_{pp} \end{bmatrix}$$

em que a covariância amostral entre a k-ésima e l-ésima variável é dado por

$$S_{kl} = \frac{1}{n-1}\sum_{j=1}^{n}(X_{jk} - \bar{X}_{.k})(X_{jl} - \bar{X}_{.l}).$$

A matriz de covariâncias amostral é obtida a partir dos vetores amostrais $\boldsymbol{X}_j$ da seguinte forma:

$$\boldsymbol{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_{.})(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_{.})^{\top}.$$

A matriz de somas de quadrados e produtos amostral é representada por $\boldsymbol{W}$ é definida por

$$\begin{aligned} \boldsymbol{W} &= \sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_.)(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_.)^\top \\ &= (n-1)\boldsymbol{S}. \end{aligned}$$

3.5 Aproximação de série de Taylor em torno da média e método Delta

Em muitas situações, nem sempre se tem o interesse apenas em conhecer uma variável aleatória específica, mas, também é desejado conhecer alguma função dessa variável aleatória propriamente dita. No entanto, em muita das vezes, a distribuição exata da variável aleatória é desconhecida ou ainda obter tais funções de variáveis aleatórias é uma tarefa árdua, pois pode ser uma função de difícil integração. O método Delta é uma maneira razoável de resolver tais questões apresentadas. Um procedimento baseado na aproximação da série de Taylor nos permite aproximar a média e a variância de uma função de uma variável aleatória. De maneira geral, uma série de Taylor de uma dada função $f(x)$ em torno do ponto $x = a$ é uma série de funções da forma:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

em que $f^{(n)}(a)$ é a n-ésima derivada da função f em relação à variável x avaliada no ponto a.

De maneira associada, podemos definir como $p(x)$ o Polinômio de Taylor de ordem n de uma dada função $f(x)$ n-vezes diferenciável em torno do ponto $x = a$ como:

$$p(x) = \sum_{n=0}^{n} \frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{3.3}$$

O polinômio de Taylor apresentado em (3.3) é útil em situações onde tem-se o interesse em obter média e/ou variâncias de funções de variáveis aleatórias cuja obtenção exata é intratável. Vejamos um exemplo:

Exemplo: *Seja X uma variável aleatória com $\mathbb{E}[X] = \mu$ e $g(\cdot)$ uma função diferenciável qualquer Uma maneira de aproximar $E[g(X)]$ é utilizando o polinômio de Taylor de ordem n em torno do ponto μ. Para $n = 4$, temos:*

$$g(X) \approx g(\mu) + \frac{g^{(1)}(\mu)}{1!}(X-\mu)^1 + \frac{g^{(2)}(\mu)}{2!}(X-\mu)^2 + \frac{g^{(3)}(\mu)}{3!}(X-\mu)^3 + \frac{g^{(4)}(\mu)}{4!}(X-\mu)^4. \tag{3.4}$$

Aplicando a função $\mathbb{E}[\cdot]$ em ambos lados de (3.4), obtém-se:

$$\mathbb{E}[g(X)] \approx g(\mu) + \frac{g^{(2)}(\mu)}{2}\mathbb{V}(X) + \frac{g^{(3)}(\mu)}{6}\mathbb{E}\left[(X-\mu)^3\right] + \frac{g^{(4)}(\mu)}{24}\mathbb{E}\left[(X-\mu)^4\right].$$

Supondo ainda, por exemplo, que $g(X) = \ln X$, temos que:

$$g^{(1)}(x) = \frac{1}{X}; \quad g^{(2)}(x) = -\frac{1}{X^2}$$
$$g^{(3)}(x) = \frac{2}{X^3}; \quad g^{(4)}(x) = -\frac{6}{X^4}.$$

Desse modo, temos que:

$$\mathbb{E}[\ln X] \approx \ln\mu - \frac{\mathbb{V}(X)}{2\mu^2} + \frac{\mathbb{E}\left[(X-\mu)^3\right]}{3\mu^3} - \frac{\mathbb{E}\left[(X-\mu)^4\right]}{4\mu^4}. \qquad \square$$

Utilizando o polinômio de Taylor para aproximar a média e a variância, podemos utilizar a seguinte generalização do Teorema Central do Limite, conhecido como Método Delta.

Teorema (Método Delta). *Seja Y_n uma sequência de variáveis aleatórias que satisfaz $\sqrt{n}(Y_n - \theta) \to N(0,\sigma^2)$. Para uma dada função g e um específico valor de θ, suponha que $g'(\theta)$ exista e seja diferente de 0. Então*

$$\sqrt{n}[g(Y_n) - g(\theta)] \to N(0,\sigma^2[g'(\theta)]^2).$$

3.6 Distribuição Normal Multivariada

Sejam $X_1, X_2, \ldots, X_p$, p variáveis normais independentes e com funções densidades definidas por:

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma_{ii}}} exp\left\{-\frac{(x_i-\mu_i)^2}{2\sigma_{ii}}\right\},$$

em que μ_i é a média e σ_{ii} é a variância para a i-ésima variável. Agrupando-se, as p variáveis independentes em um vetor aleatório $\mathbf{X}$, a função densidade de probabilidade conjunta é dada por

$$f_X(x) = \prod_{i=1}^{p} f_X(x_i) = (2\pi)^{-\frac{p}{2}} \left(\prod_{i=1}^{p} \sigma_{ii}\right)^{-\frac{1}{2}} exp\left\{-\sum_{i=1}^{p} \frac{(x_i-\mu_i)^2}{2\sigma_{ii}}\right\}. \tag{3.5}$$

Por meio de uma manipulação algébrica, é possível reescrever (3.5) dispondo de vetores e matrizes e considerando uma matriz de covariância positiva definida mais geral, $\mathbf{\Sigma}$, e desse modo pode-se obter a função densidade da distribuição normal multivariada, que tem a seguinte forma (FERREIRA , 2018):

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = (2\pi)^{-\frac{p}{2}} |\mathbf{\Sigma}|^{-\frac{1}{2}} exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \mathbf{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\}, \tag{3.6}$$

em que $\boldsymbol{\mu}$ é o vetor de médias e $\mathbf{\Sigma}$ é a matriz de covariâncias, expressos, matricialmente, da seguinte forma:

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} \quad \text{e} \quad \mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \dots & \sigma_{pp} \end{bmatrix}.$$

3.7 Distribuição Wishart

Seja $\mathbf{X}_j = [X_1, X_2, \dots, X_p]^\top$ o j-ésimo vetor $(j = 1, 2, \dots, \nu)$ de uma amostra aleatória de tamanho ν de uma normal p-variada com média $\mathbf{0}$ e covariância $\mathbf{\Sigma}$, então a matriz aleatória

$$\mathbf{W} = \sum_{j=1}^{\nu} \mathbf{X}_j \mathbf{X}_j^\top$$

possui distribuição Wishart com ν graus de liberdade e parâmetro $\mathbf{\Sigma}$ (matriz positiva definida).

Caso, seja uma amostra aleatória de tamanho n de uma distribuição normal multivariada com média $\boldsymbol{\mu}$ e a covariância $\mathbf{\Sigma}$, a distribuição da matriz aleatória

$$\mathbf{W} = \sum_{j=1}^{n} (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})^\top$$

é Wishart com $\nu = n-1$ graus de liberdade e parâmetro $\mathbf{\Sigma}$. A função densidade Wishart de uma matriz aleatória $\mathbf{W}$ de somas de quadrados e produtos é representada por $W_p(\nu, \mathbf{\Sigma})$ é definida por (FERREIRA, 2018)

$$f_{\mathbf{W}}(\boldsymbol{w}; \nu, \mathbf{\Sigma}) = \frac{|\mathbf{\Sigma}|^{-\nu/2} |\boldsymbol{w}|^{(\nu-p-1)/2}}{2^{\nu p/2} \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\frac{\nu-i+1}{2}\right)} exp\left\{-\frac{1}{2} tr(\mathbf{\Sigma}^{-1} \boldsymbol{w})\right\} \tag{3.7}$$

em que $\Gamma(x) = \int_0^\infty t^{x-1} e^{-x} dt$ é função gama.

3.8 Variâncias generalizadas

Uma importante quantidade que aparece na inferência multivariada é a variância generalizada. Esta é um sumário da matriz de covariâncias em um único valor, a qual contém p variâncias e $\frac{p(p-1)}{2}$ covariâncias e é uma medida de variabilidade aplicada para a determinação da dispersão em uma população multivariada.

A variância generalizada é definida por $|\mathbf{\Sigma}|$ para a matriz de covariâncias populacional e $|\boldsymbol{S}|$, para a matriz de covariâncias amostral. O determinante da matriz de covariâncias tem uma interpretação geométrica, em que os vetores de desvios para média formam um hiperparalelograma no espaço p-dimensional. O tamanho de cada aresta é dado por uma quantidade que é proporcional à variância e o ângulo entre cada par é função de uma quantidade proporcional à covariância (ou correlação) entre as duas variáveis contidas no par.

O máximo volume é obtido quando os ângulos são de 90° e as variâncias das p variáveis são iguais. Assim, no caso amostral, em uma amostra de tamanho n, temos que $|\boldsymbol{S}| = V^2(n-1)^{-p}$, em que V representa o volume desse hiperparalelograma.

3.9 Aplicações da variância generalizada

3.9.1 Gráfico de controle multivariado

Nas indústrias, existem diversas situações nas quais o monitoramento ou controle simultâneo de duas ou mais características relacionadas ao processo de qualidade é necessário. O monitoramento de processos em que várias variáveis relacionadas são de interesse, é conhecido como controle estatístico multivariado de processos. Uma importante ferramenta para o controle estatístico de processos multivariado é o gráfico de controle de qualidade.

Dois métodos de desenvolvimento do gráfico de controle baseado na variância generalizada, foram apontados por Bersimis, Psarakis e Panaretos (2007). Em uma abordagem utiliza-se a média e a variância de $|\boldsymbol{S}|$

$$E(|\boldsymbol{S}|) \pm 3Var(|\boldsymbol{S}|).$$

Outra abordagem mencionada por Bersimis, Psarakis e Panaretos (2007), indica que se houver duas características de qualidade, então

$$[2(n-1)|\boldsymbol{S}|^{1/2}]|\boldsymbol{\Sigma}_0|^{-1/2} \sim \boldsymbol{\chi}^2_{2n-4}.$$

Assim, os limites superior e inferior, respectivamente, de um gráfico de controle baseado na variância generalizada, é dado por

$$\begin{aligned} L_u &= [|\boldsymbol{\Sigma}_0|(\chi^2_{2n-4,1-\alpha})^2][2(n-1)]^{-2} \\ L_l &= [|\boldsymbol{\Sigma}_0|(\chi^2_{2n-4,\alpha/2})^2][2(n-1)]^{-2} \end{aligned}$$

3.9.2 Processamento de sinais

Na área de processamento de sinais, os sinais são observados em diferentes sensores de diferentes fontes em diferentes momentos. Em geral, o modelo no processamento de sinais, é dado por

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t),$$

em que,

- $\boldsymbol{x}(t) = (x_1(t), x_2(t), \ldots, x_p(t))^\top$ é o vetor de de observações $p \times 1$ no tempo t;
- $\boldsymbol{A} = [\boldsymbol{A}(\boldsymbol{\phi}_1), \boldsymbol{A}(\boldsymbol{\phi}_2), \ldots, \boldsymbol{A}(\boldsymbol{\phi}_q)]$ é uma matriz $p \times q$ de coeficientes desconhecidos; $\boldsymbol{A}(\boldsymbol{\phi}_i)$ é o vetor de funções dos elementos do vetor desconhecido $\boldsymbol{\phi}_i$ associado ao i-ésimo sinal e $q < p$;
- $\boldsymbol{s}(t) = (s_1(t), s_2(t), \ldots, s_q(t))^\top$ é o vetor $q \times 1$ de sinais aleatórios desconhecidos no tempo t;

- $\boldsymbol{n}(t) = (n_1(t), n_2(t), \ldots, n_p(t))^\top$ é o vetor $p \times 1$ de ruído aleatório no tempo t;

Assume-se $\boldsymbol{x}(t) \sim N_p(\boldsymbol{0}, \boldsymbol{A\Psi A}^\top + \sigma^2 \boldsymbol{I}_p)$, em que $\boldsymbol{\Psi}$ é a matriz de covariâncias de $\boldsymbol{s}(t)$ e $\sigma^2 \boldsymbol{I}_p$ é a matriz de covariâncias de $\boldsymbol{n}(t)$. De acordo com Bhandary (1996), se a matriz de covariâncias do vetor de ruídos $\boldsymbol{n}(t)$ for $\sigma^2 \boldsymbol{I}_p$, o modelo é chamado de modelo de ruído branco. Se a matriz de covariâncias de $\boldsymbol{n}(t)$ for $\sigma^2 \boldsymbol{\Sigma}_1$, em que $\boldsymbol{\Sigma}_1$ é uma matriz positva definida arbitrária, diz-se que o modelo é um modelo de ruído colorido.

Bhandary (1996) propõe testar a variância generalizada em processamento de sinais para os casos de modelo de ruído branco e modelo de ruído colorido, pelo fato de ser importante conhecer sobre a estrutura de variação dos dados multivariados e a variância generalizada fornecer uma ideia intuitiva geral sobre a dispersão dos pontos multidimensionais.

3.9.3 Análise de agrupamentos

Na análise de agrupamentos, objetiva-se classificar itens, objetos ou indivíduos de acordo com suas semelhanças. Desse modo, itens (objetos ou indivíduos) similares são alocados em um mesmo grupo e os que se encontram em diferentes grupos, são considerados dissimilares.

Dado um conjunto de pontos, pode-se utilizar de diversas técnicas de agrupamento baseadas em diferentes critérios, por exemplo, medidas de similaridade, funções de distância, correlações, etc. Gupta (1982) aponta algumas situações, em que dentro alguns agrupamentos existem subgrupos e que as variações dentro de cada subgrupo precisam ser consideradas para fornecer uma diferenciação precisa entre o agrupamento.

Nessas situações, dados os vetores de observações multidimensionais em cada subgrupo, uma medida de dispersão deve ser usada como critério para determinar agrupamentos, e portanto, a variância generalizada $|\boldsymbol{\Sigma}|$, pode ser considerada como um critério.

3.10 Testes de hipóteses

Uma hipótese é uma afirmação sobre um parâmetro populacional e o objetivo de um teste de hipóteses é decidir, com base em uma amostra da população, qual das duas hipóteses complementares é verdadeira. As duas hipóteses complementares em um problema de teste de hipótese são chamadas de hipótese nula e a hipótese alternativa, e são denotadas por H_0 e H_1 respectivamente (CASELLA; BERGER, 2001).

Para testar qualquer hipótese com base em uma amostra (aleatória) de observações, devemos dividir o espaço amostral em duas regiões. Seja R a região de rejeição para o teste, se o ponto de amostra observado $\boldsymbol{x}$ cair na região R, então rejeita-se a hipótese; se $\boldsymbol{x}$ cair na região complementar, R^c, não deve-se rejeitar a hipótese. R é conhecida como a região crítica do teste e R^c é chamada de região de aceitação (KENDALL; STUART, 1961).

Ao se realizar um teste de hipótese $H_0 : \theta \in \Theta_0$ contra $H_1 : \theta \in \Theta_0^c$, sendo θ um parâmetro populacional e Θ_0 um subconjuto do espaço do parâmetro, podem ocorrer dois tipos de erro, denominados de erros do tipo I e do tipo II.

3.10.1 Erros tipo I e II

A rejeição de H_0 quando a mesma é, de fato, verdadeira é chamada de erro tipo I, e o erro cometido quando não se rejeita a hipótese H_0 quando, na verdade, ela é uma hipótese falsa é chamada de erro tipo II.

Para $\theta \in \Theta_0$, o teste cometerá um erro se $\boldsymbol{x} \in R$, então a probabilidade de se cometer erro do tipo I é $P_\theta(\boldsymbol{X} \in R)$. Para $\theta \in \Theta_0^c$, a probabilidade de se cometer erro do tipo II é $P_\theta(\boldsymbol{X} \in R^c) = 1 - P_\theta(\boldsymbol{X} \in R)$ (CASELLA; BERGER, 2001).

3.11 Teste de razão de verossimilhança generalizada

Seja $L(\boldsymbol{\theta};\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n)$ a função de verossimilhança de uma amostra $\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n$ com função densidade conjunta $f(\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_n;\boldsymbol{\theta})$. A estatística do teste da razão de verossimilhanças para testar $H_0 : \boldsymbol{\theta} \in \Theta_0$ contra $H_1 : \boldsymbol{\theta} \in \Theta_0^c$ é definida por

$$\Lambda = \frac{L_{\Theta_0}(\boldsymbol{\theta};\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n)}{L_{\Theta}(\boldsymbol{\theta};\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n)},$$

em que, $L_{\Theta_0}(\boldsymbol{\theta};\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n)$ é o máximo da função de verossimilhança para o espaço restrito e $L_{\Theta}(\boldsymbol{\theta};\boldsymbol{X}_1,\boldsymbol{X}_2,\ldots,\boldsymbol{X}_n)$ é o máximo da função de verossimilhança para o espaço irrestrito. Se a razão de verossimilhanças em seu máximo é grande, H_0 é mais provável e caso a razão de verossimilhanças for pequena, H_1 deverá ser escolhida.

Muitas das vezes, a distribuição nula de Λ é muito complicada. Quando isso ocorre, é comum utilizar a aproximação qui-quadrado assintótica de $-2\ln\Lambda$. Para esta situação, a região de rejeição R da hipótese nula para o teste de razão de verossimilhanças é definida por:

$$R = \{\boldsymbol{X} | -2\ln[\Lambda(\boldsymbol{X})] > \chi^2_{\alpha,r-s}\}$$

em que $\chi^2_{\alpha,r-s}$ é quantil superior da distribuição qui-quadrado com $r-s$ graus de liberdade, sendo α o nível de significância estabelecido pelo pesquisador ao montar o teste de hipóteses, r o número de parâmetros da função de verossimilhança para o espaço irrestrito, enquanto que s é o número de parâmetros da função de verossimilhança restrita sob a hipótese nula.

4 REFERÊNCIAS BIBLIOGRÁFICAS

ARVANITIS, L. G.; AFONJA, B. Use of the generalized variance and the gradient projection method in multivariate stratified sampling. **Biometrics**, v. 27, n. 1, p. 119–127, 1971.

BERSIMIS, S.; PSARAKIS, S.; PANARETOS, J. Multivariate statistical process control charts: an overview. **Qual. Reliab. Eng. Int.**, v. 5, n. 23, p. 517–543, 2007.

BHANDARY, M. Test for generalized variance in signal processing. **Statatistical and Probability Letters**, v. 27, n. 2, p. 155–162, 1996.

CASELLA, G.; BERGER, R. L. **Statistical inference**. 2. ed. Austrália, Pacific Grove, CA: Thomson Learning, 2002. 660 p.

DJAUHARI, M. A. Improved monitoring of multivariate process variability. **Journal of Quality Technology**, v. 37, n. 1, p. 32–39, 2005.

DJAUHARI, M. A.; MASHURI, M.; HERWINDIATI, D. E. Multivariate process variability monitoring. **Communication in Statistics - Theory and Methods**, v. 37, n. 11, p. 1742–1754, 2008.

FERREIRA, D. F. **Estatística Multivariada**. 3. ed. Lavras, MG: Editora UFLA, 2018. 624 p.

GUPTA, A. S. Tests for simultaneously determining numbers of clusters and their shapewith multivariate data. **Statatistical and Probability Letters**, v. 1, n. 1, p. 46–50, 1982.

KENDALL, M. G.; STUART, A. **The advanced theory of statistics - Inference and Relationship**. London: Charles Griffin and Co. Ltd., 1961. 676 p.

LEE, M. H.; KHOO, M. B. C. Combined synthetic and |SS | chart for monitoring process dispersion. **Communications in Statistics - Simulation and Computation**, v. 46, n. 7, p. 5698–5711, 2017.

NAJARZADEH, D. Testing equality of generalized variances of k multivariate normal populations. **Communications in Statistics - Simulation and Computation**, Taylor and Francis, v. 46, n. 8, p. 6414–6423, 2017.

NAJARZADEH, D. Testing equality of standardized generalized variances of k multivariate normal populations with arbitrary dimensions. **Statistical Methods and Applications**, n. 1, p. 1–31, 2019.

NOOR, A. M.; DJAUHARI, M. A. Monitoring the variability of beltline moulding process using Wilks's statistic. **Malays J. Fundam. Appl. Sci**., v. 6, n. 2, p. 116–120, 2014.

PUKELSHEIM, F. **Optimal design of experiments**. New York: SIAM, 2006. 454 p.

SENGUPTA, A. Generalizations of Barlett's and Hartley's tests of homogeneity using overall variability. **Communication in Statistics - Theory and Methods**, v. 16, n. 4, p. 987–996, 1987.

SENGUPTA, A. Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions. **Journal Multivariate Analysis**, v. 23, n. 2, p. 209–219, 1987.

TALLIS, G. M.; LIGHT, R. The use of fractional moments for estimating the parameters of a mixed exponential distribution **Technometrics**, v. 10, n. 1, p. 161–175, 1968.

YEH, A. B. et al. A multivariate exponentially weighted moving average control chart for monitoring process variability. **Journal of Applied Statistics**, v. 30, n. 5, p. 507–536, 2003.

5 ARTIGO

Likelihood Ratio Test For The Multivariate Normal Generalized Variance

Abstract

An interesting measure of variability in the multivariate population is the determinant of the covariance matrix $\mathbf{\Sigma}_{p\times p}$, $|\mathbf{\Sigma}|$, of a population, known as generalized variance. This is a measure that summarizes the dispersion of a multivariate population in a single value, considering the dependence between the variables involved. Because of this, it has applications in several areas that aim to evaluate the dispersion existing in some multivariate population of interest. We present to construct a likelihood ratio test for multivariate normal generalized variance and show the theory of the distribution of the multivariate normal sample generalized variance developed. We proposed the LRT and BCLRT tests for testing the hypothesis that generalized variance is equal to a η, such that $\eta \in \mathbb{R}$. All development this test was build theorically. Monte Carlo simulations were performed to compare the performance of our test and confront it with the other tests considered in this study. Finally, we illustrate our method using a real example. We recommend using the BCLRT test only in scenarios in that we have $p = 2$, especially if $n > 30$. As for the LRT test, we recommend its use in situations where we have $p = 2$ and $p = 3$ for $n > 30$ and for $p = 5$ when $n > 50$.

Keywords: Monte Carlo; Standard generalized variance; Variability measure.

5.1 Introduction

A interesting measure of variability in a multivariate population is the determinant of the $p \times p$ covariance matrix $\mathbf{\Sigma}$, $|\mathbf{\Sigma}|$, of a population, known as a generalized variance. This measure, as pointed out by Najarzadeh (2019), is also extensively used, for example, in multivariate control chart (YEH et al., 2003; YEH; J.; McGRATH, 2006; BERSIMIS; PSARAKIS; PANARETOS, 2007; DJAUHARI, 2005; DJAUHARI; MASHURI; HERWINDIATI, 2008; NOOR; DJAUHARI, 2014; LEE; KHOO, 2017), modeling in reliability analysis (TALLIS; LIGHT, 1968), signal processing (BHANDARY, 1996), clustering (GUPTA, 1982), optimal design (PUKELSHEIM, 2006), and optimal allocation in stratified sampling (ARVANITIS; AFONJA, 1971). The generalized variance is a measure of the hypervolume that the distribution of the random variables occupied in the p-dimensional space. Another measure is the standardized generalized variances, that allow comparisons among sets of different dimensions, which is defined by the geometric means of the eigenvalues of $\mathbf{\Sigma}$, that is $|\mathbf{\Sigma}|^{1/p}$ (SENGUPTA, 1987a; SENGUPTA, 1987b).

Several procedures of confidence intervals and hypothesis tests were described by Jafari e Kazemi (2014). They evaluated the performance of the studied procedures by using Monte Carlo simulations. There was not found any mention in the scientific literature regarding the likelihood ratio test on generalized variance of a normal population. On the other hand, tests for comparisons and confidence intervals for product of several (standardized) generalized variances are proposed by Najarzadeh (2017), Najarzadeh (2019). Our focus is to construct the likelihood ratio test (LRT) for the null hypothesis of $H_0 : |\mathbf{\Sigma}| = \eta$, under multivariate normality.

Also, we show the LRT step by step and show the theory of the distribution of the multivariate normal sample generalized variance developed until now. Monte Carlo simulations were performed to compare the performance of our test and confront it with the other tests considered in this study. Finally, we illustrate our method using real data.

5.2 Normal sample generalized variance distribution

The theorem of Bartlett (1934) refers to the transformation of Wishart matrices using the Cholesky decomposition. This result is of great importance since this is the most commonly used procedure for generating Wishart, $\boldsymbol{W}$ $(p \times p)$, random variable realizations. Let $\boldsymbol{T}$ be a upper triangular matrix, the Cholesky factor of $\boldsymbol{W}$ and consider the transformation $\boldsymbol{W} = \boldsymbol{T}^\top \boldsymbol{T}$. The next result will be demonstrated using the Wishart density directly. An alternative proof can be seen at Kollo e Rosen (2005).

Theorem 1 (Bartlett's theorem)**.** *Let $\boldsymbol{W} \sim W_p(\nu, \boldsymbol{I})$ $(\nu \geq p)$ and $\boldsymbol{W} = \boldsymbol{T}^\top \boldsymbol{T}$, where $\boldsymbol{T}$ is an upper triangular matrix $p \times p$ with positive diagonal entries, then the elements t_{ij} $(1 \leq i \leq j \leq p)$ of $\boldsymbol{T}$ are independents and $t_{ij} \sim N(0,1)$ $(1 \leq i < j \leq p)$ and $t_{ii}^2 \sim \chi^2_{\nu-i+1}$ $(i = 1, 2, \cdots, p)$.*

Proof. The density of $\boldsymbol{W}$, when $\mathbf{\Sigma} = \boldsymbol{I}$ is given by

$$f_{\boldsymbol{W}}(\boldsymbol{w}; n, \boldsymbol{I}) = \frac{|\boldsymbol{w}|^{(\nu-p-1)/2}}{2^{\nu p/2}\Gamma_p\left(\frac{\nu}{2}\right)} \exp\left\{-\frac{1}{2}\operatorname{tr}(\boldsymbol{w})\right\}, \tag{5.1}$$

where

$$\Gamma_p\left(\frac{\nu}{2}\right) = \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(\frac{\nu - i + 1}{2}\right).$$

We should notice to use the Jacobian transformation method that we have essentially $p(p+1)/2$ variables in $\boldsymbol{W}$ (symmetric). Therefore

$$
\begin{aligned}
\boldsymbol{W} =& \boldsymbol{T}^\top \boldsymbol{T} \\
=& \begin{bmatrix} T_{11} & 0 & 0 & \cdots & 0 \\ T_{12} & T_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_{1p} & T_{2p} & T_{3p} & \cdots & T_{pp} \end{bmatrix} \begin{bmatrix} T_{11} & T_{12} & T_{13} & \cdots & T_{1p} \\ 0 & T_{22} & T_{23} & \cdots & T_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & T_{pp} \end{bmatrix}.
\end{aligned}
$$

Considering the vec operator ignoring the lower triangular matrix elements of this product we get

$$
\operatorname{vec}\left(\boldsymbol{T}^\top \boldsymbol{T}\right) = \begin{bmatrix} T_{11}^2 \\ T_{11}T_{12} \\ \vdots \\ T_{11}T_{1p} \\ T_{12}^2+T_{22}^2 \\ T_{12}T_{13}+T_{22}T_{23} \\ \vdots \\ T_{12}T_{1p}+T_{22}T_{2p} \\ \vdots \\ T_{1p}^2+T_{2p}^2+\cdots+T_{pp}^2 \end{bmatrix}.
$$

Taking the first derivative in respect to $\operatorname{vec}(\boldsymbol{T}^\top)$, the following $p(p+1)/2 \times p(p+1)/2$ lower triangular Jacobian matrix is obtained. This,

$$
\frac{\partial \operatorname{vec}\left(\boldsymbol{T}^\top \boldsymbol{T}\right)}{\partial \operatorname{vec}\left(\boldsymbol{T}^\top\right)} = \begin{bmatrix} 2t_{11} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ t_{12} & t_{11} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ t_{1p} & 0 & \cdots & t_{11} & 0 & \cdots & 0 \\ 0 & 2t_{12} & \cdots & 0 & 2t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 2t_{pp} \end{bmatrix}.
$$

This matrix has on its diagonal an element equal to $2t_{11}$ and other $p-1$ equal to t_{11}, an element equal to $2t_{22}$ and other $p-2$ equals t_{22} and so on. So the Jacobian of transformation is

$$
J = 2t_{11}\prod_{i=1}^{p-1} t_{11} \times 2t_{22}\prod_{i=1}^{p-2} t_{22} \times 2t_{33}\prod_{i=1}^{p-3} t_{33} \times \cdots \times 2t_{pp} = 2^p \prod_{i=1}^{p} t_{ii}^{p-i+1}.
$$

Also,

$$\mathrm{tr}(\boldsymbol{w}) = \mathrm{tr}(\boldsymbol{t}^\top \boldsymbol{t}) = \sum_{i \le j}^{p} t_{ij}^2$$

and

$$|\boldsymbol{w}| = |\boldsymbol{t}^\top \boldsymbol{t}| = |\boldsymbol{t}|^2 = \prod_{i=1}^{p} t_{ii}^2.$$

Using the Jacobian transformation method we get

$$\begin{aligned}
f_{\boldsymbol{T}}(\boldsymbol{t};\nu) =& f_{\boldsymbol{W}}(\boldsymbol{w};n)|J| \\
=& \frac{\left(\prod_{i=1}^{p} t_{ii}^2\right)^{(\nu-p-1)/2}}{2^{\nu p/2}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left(\frac{\nu-i+1}{2}\right)} \times \\
& \times \exp\left\{-\frac{1}{2}\sum_{i\le j}^{p} t_{ij}^2\right\} 2^p \prod_{i=1}^{p} t_{ii}^{p-i+1} \\
=& \frac{\prod_{i=1}^{p}\left(t_{ii}^{\nu-p-1} t_{ii}^{p-i+1}\right)}{2^{\nu p/2}2^{-p}\pi^{p(p-1)/4}\prod_{i=1}^{p}\Gamma\left(\frac{\nu-i+1}{2}\right)} \prod_{i\le j}^{p} \exp\left\{-\frac{t_{ij}^2}{2}\right\} \\
=& \frac{1}{2^{\nu p/2}2^{-p}2^{p/2-p/2}2^{-p(p-1)/4}} \prod_{i<j}^{p}\left[(2\pi)^{-1/2} e^{-t_{ij}^2/2}\right] \times \\
& \times \prod_{i=1}^{p}\left[\frac{1}{\Gamma\left(\frac{\nu-i+1}{2}\right)}\left(t_{ii}^2\right)^{(\nu-i+1-1)/2} e^{-t_{ii}^2/2}\right] \\
=& \prod_{i<j}^{p}\left[(2\pi)^{-1/2} e^{-t_{ij}^2/2}\right] \times \\
& \times \prod_{i=1}^{p}\left[\frac{1}{2^{(\nu-i+1)/2}\Gamma\left(\frac{\nu-i+1}{2}\right)}\left(t_{ii}^2\right)^{(\nu-i+1-1)/2} e^{-t_{ii}^2/2}\right],
\end{aligned}$$

that correspond to the product of independent chi-square variables T_{ii}^2's with $\nu - i + 1$ degrees of freedom and of standard normal variables T_{ij}'s, i.e., $N(0,\ 1)$. As the joint probability density of the T_{ii}^2's and T_{ij}'s $(i < j)$ is product of their marginal densities, they are independent distributed.

$\square$

The chi-square moment generating function of a chi-square X variable with ν degrees of freedom is $M_X(t) = (1-2t)^{-\nu/2}$ (MOOD; GRAYBILL; BOES, 1974; MITTELHAMMER, 2013). Thus, the rth moment about the origin can be deduced, as showed in the next theorem.

Theorem 2 (Chi-square moments about the origin)**.** *Let $X \sim \chi^2_\nu$ with $\nu > 0$ degrees of freedom and moment generating function of $M_X(t) = (1-2t)^{-\nu/2}$, then the rth moment about the origin from the distribution of X is*

$$\mathbb{E}[X^r] = \nu(\nu+2)(\nu+4)(\nu+6)\times\cdots\times(\nu+2r-2) = 2^r\frac{\Gamma\left(\frac{\nu}{2}+r\right)}{\Gamma\left(\frac{\nu}{2}\right)}. \tag{5.2}$$

Proof. We can see that

$$M_X^{(1)}(t) = \frac{dM_X(t)}{dt} = \frac{d(1-2t)^{-\nu/2}}{dt} = \nu(1-2t)^{-\nu/2-1},$$

that if it was evaluated in $t=0$ results in

$$M_X^{(1)}(0) = \nu.$$

Repeating this procedure, to obtain the second derivative, we get

$$M_X^{(2)}(t) = \frac{d^2M_X(t)}{dt^2} = \frac{d\nu(1-2t)^{-\nu/2-1}}{dt} = \nu(\nu+2)(1-2t)^{-\nu/2-2},$$

that simplifies in

$$M_X^{(2)}(0) = \nu(\nu+2).$$

Similarly, for the third derivative, we have

$$M_X^{(3)}(t) = \frac{d^3M_X(t)}{dt^3} = \frac{d\nu(\nu+2)(1-2t)^{-\nu/2-2}}{dt} = \nu(\nu+2)(\nu+4)(1-2t)^{-\nu/2-3},$$

that gives

$$M_X^{(3)}(0) = \nu(\nu+2)(\nu+4).$$

Thus, repeating this procedure several times until the rth derivative, we get the final result given by

$$\begin{aligned} M_X^{(r)}(0) =& \nu(\nu+2)(\nu+4)\times\cdots\times(\nu+2r-2) \\ =& 2^r(\nu/2)(\nu/2+1)(\nu/2+2)\times\cdots\times(\nu/2+r-1) \\ =& = 2^r\frac{\Gamma\left(\frac{\nu}{2}+r\right)}{\Gamma\left(\frac{\nu}{2}\right)} = \mathbb{E}\left[X^r\right], \end{aligned}$$

by the gamma function properties. □

The generalized variance is a single-value summary of the covariance matrix, which contains p variances and $p(p-1)/2$ covariances. The generalized variance is defined by $|\boldsymbol{\Sigma}|$ for the population covariance matrix and $|\boldsymbol{S}|$ for the sample covariance matrix. The determinant of the covariance matrix has a geometric interpretation, where the vectors of deviations of each observation to the mean form a hyperparallelogram in a p-dimensional space. The size of each side is given by an amount proportional to the variance and with the angle between each pair a function of a quantity proportional to the covariance (or correlation) between the two variables contained in the pair. The maximum volume is obtained when the angles are 90^0 and the variances of the p variables are equal. Thus, in the sample case, with sample size n, we have $|\boldsymbol{S}| = V^2(n-1)^{-p}$, where V represents the volume of this hyperparallelogram. Generalized variance appears in the statistics of many multivariate hypothesis tests. The probability density function of its exact distribution, in the normal multivariate case, is very difficult to obtain in practice, although its distribution is easily obtained with the help of Bartlett's theorem.

Theorem 3 (Distribution of $|\boldsymbol{W}|$ from the Wishart distribution)**.** *Let $\boldsymbol{W} \sim W_p(\nu,\, \boldsymbol{\Sigma})$ $(\nu \geq p)$, then the distribution of the random variable $|\boldsymbol{W}|/|\boldsymbol{\Sigma}|$ is the same of the $\prod_{i=1}^{p}\chi^2_{\nu-i+1}$, where $\chi^2_{\nu-i+1}$, $i = 1,\, 2,\, \cdots,\, p$ are independent random chi-square variables with $\nu-i+1$ degrees of freedom, for $i = 1,\, 2,\, \cdots,\, p$.*

Proof. Consider that $|\boldsymbol{W}|/|\boldsymbol{\Sigma}| = |\boldsymbol{W}||\boldsymbol{\Sigma}|^{-1} = |\boldsymbol{\Sigma}|^{-1/2}|\boldsymbol{W}||\boldsymbol{\Sigma}|^{-1/2} = |\boldsymbol{\Sigma}^{-1/2}\boldsymbol{W}\boldsymbol{\Sigma}^{-1/2}|$. As $\boldsymbol{W}$ is $W_p(\nu, \boldsymbol{\Sigma})$, then by the linear transformation of Wishart random variables, $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{W}\boldsymbol{\Sigma}^{-1/2}$ is $W_p(\nu, \boldsymbol{I}_p)$ (JOHNSON; WICHERN, 1998). Thus

$$\boldsymbol{\Sigma}^{-1/2}\boldsymbol{W}\boldsymbol{\Sigma}^{-1/2} = \boldsymbol{T}^\top\boldsymbol{T},$$

where $\boldsymbol{T}$ is an upper triangular matrix. By the Bartlett's theorem 1, we have $|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{W}\boldsymbol{\Sigma}^{-1/2}| = \prod_{i=1}^{p} T_{ii}^2$, where T_{ii}^2 has a chi-square distribution with $\nu - i + 1$ degrees of freedom. By the same theorem, the T_{ii}^2, for $i = 1, 2, \cdots, p$, are independent random variables. Hence,

$$|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{W}\boldsymbol{\Sigma}^{-1/2}| = \prod_{i=1}^{p} T_{ii}^2 = \prod_{i=1}^{p} \chi^2_{\nu-i+1}.$$

as expected. □

We know that the distribution of $|\boldsymbol{W}|/|\boldsymbol{\Sigma}|$ is the product of independent chi-square variables, which does not mean that we know its probability density function. How Muirhead (1982) points out, find the probability density function, in this case, is not an easy task, even we know it is the product of independent chi-square variables. This result is very important if we use Monte Carlo simulation in some inference process. It is also important to determine some distributional properties, such as moments and asymptotic approximations.

Theorem 4 (Moments of $|\boldsymbol{W}|$). *Let $\boldsymbol{W} \sim W_p(\nu, \boldsymbol{\Sigma})$ $(\nu \geq p)$, then the rth $(r \geq 1)$ moment about the origin from the distribution of $|\boldsymbol{W}|$ is*

$$\mathbb{E}\left[|\boldsymbol{W}|^r\right] = |\boldsymbol{\Sigma}|^r \prod_{i=1}^{p} \left(\frac{2^r \Gamma[((\nu - i + 1)/2) + r]}{\Gamma[(\nu - i + 1)/2]} \right). \tag{5.3}$$

Proof. Considering that $|\boldsymbol{W}|/|\boldsymbol{\Sigma}|$ has the distribution of the product of independent chi-square variables, then

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{|\boldsymbol{W}|}{|\boldsymbol{\Sigma}|}\right)^r\right] &= \mathbb{E}\left[\left(\prod_{i=1}^{p} \chi^2_{\nu-i+1}\right)^r\right] \\
&= \mathbb{E}\left[\prod_{i=1}^{p} \left(\chi^2_{\nu-i+1}\right)^r\right] \\
&= \prod_{i=1}^{p} \mathbb{E}\left[\left(\chi^2_{\nu-i+1}\right)^r\right] \qquad \text{(by the independence)} \\
&= \prod_{i=1}^{p} \frac{2^r \Gamma((\nu - i + 1)/2 + r)}{\Gamma((\nu - i + 1)/2)}, \qquad \text{(by theorem 2, expression (5.2))}.
\end{aligned}$$

But $\mathbb{E}\left[(|\boldsymbol{W}|/|\boldsymbol{\Sigma}|)^r\right]$ by the expectation linearity is $|\boldsymbol{\Sigma}|^{-r}\mathbb{E}\left[|\boldsymbol{W}|^r\right]$. Hence

$$\mathbb{E}\left[|\boldsymbol{W}|^r\right] = |\boldsymbol{\Sigma}|^r \prod_{i=1}^{p} \frac{2^r \Gamma((\nu - i + 1)/2 + r)}{\Gamma((\nu - i + 1)/2)},$$

as pointed out. □

The mean and variance of the determinant of a Wishart matrix can be provided considering the results of theorem 4.

Corollary 4.1 (Mean and variance of $|\boldsymbol{W}|$). *Let $\boldsymbol{W} \sim W_p(\nu,\, \boldsymbol{\Sigma})$, then the mean and variance of $|\boldsymbol{W}|$ are, respectively, given by*

$$\mathbb{E}\left[|\boldsymbol{W}|\right] = |\boldsymbol{\Sigma}| \prod_{i=1}^{p} (\nu - i + 1) \tag{5.4}$$

and

$$\mathbb{V}\left(|\boldsymbol{W}|\right) = |\boldsymbol{\Sigma}|^2 \prod_{i=1}^{p} (\nu - i + 1) \left[\prod_{k=1}^{p} (\nu - k + 3) - \prod_{k=1}^{p} (\nu - k + 1) \right]. \tag{5.5}$$

Proof. For $r = 1$ and using (5.3), we have

$$\begin{aligned}
\mathbb{E}\left[|\boldsymbol{W}|\right] &= |\boldsymbol{\Sigma}|^1 \prod_{i=1}^{p} \frac{2^1 \Gamma((\nu - i + 1)/2 + 1)}{\Gamma((\nu - i + 1)/2)} \\
&= |\boldsymbol{\Sigma}| \prod_{i=1}^{p} \frac{2(\nu - i + 1)/2 \Gamma((\nu - i + 1)/2)}{\Gamma((\nu - i + 1)/2)} \\
&= |\boldsymbol{\Sigma}| \prod_{i=1}^{p} (\nu - i + 1).
\end{aligned}$$

In the same way, for $r = 2$, we have

$$\begin{aligned}
\mathbb{E}\left[|\boldsymbol{W}|^2\right] &= |\boldsymbol{\Sigma}|^2 \prod_{i=1}^{p} \frac{2^2 \Gamma((\nu - i + 1)/2 + 2)}{\Gamma((\nu - i + 1)/2)} \\
&= |\boldsymbol{\Sigma}|^2 \prod_{i=1}^{p} \frac{2^2 [(\nu - i + 1)/2 + 1](\nu - i + 1)/2 \Gamma((\nu - i + 1)/2)}{\Gamma((\nu - i + 1)/2)} \\
&= |\boldsymbol{\Sigma}|^2 \prod_{i=1}^{p} (\nu - i + 3)(\nu - i + 1).
\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{V}\left(|\boldsymbol{W}|^2\right) =& \mathbb{E}\left[|\boldsymbol{W}|^2\right] - \mathbb{E}^2\left[|\boldsymbol{W}|\right]|\\ =& |\boldsymbol{\Sigma}|^2\prod_{i=1}^{p}(\nu-i+3)(\nu-i+1) - \left[|\boldsymbol{\Sigma}|\prod_{i=1}^{p}(\nu-i+1)\right]^2\\ =& |\boldsymbol{\Sigma}|^2\prod_{i=1}^{p}(\nu-i+3)(\nu-i+1) - |\boldsymbol{\Sigma}|^2\prod_{i=1}^{p}(\nu-i+1)^2\\ =& |\boldsymbol{\Sigma}|^2\prod_{i=1}^{p}(\nu-i+1)\left[\prod_{k=1}^{p}(\nu-k+3) - \prod_{k=1}^{p}(\nu-k+1)\right],\end{aligned}$$

as stated. □

Considering the sample covariance matrix $\boldsymbol{S}$ from a multivariate normal random sample of size n, then $\boldsymbol{S} = \boldsymbol{W}/\nu$ and the distribution of $|\boldsymbol{S}|$ is the same of the $|\boldsymbol{W}|\nu^{-p}$, making use of the determinant properties, , where $\nu = n-1$ and $\boldsymbol{W} \sim W_p(\nu, \boldsymbol{\Sigma})$. The distribution of $\boldsymbol{S}$ from a sample of the multivariate normal distribution is $W_p(\nu, \boldsymbol{\Sigma}\nu^{-1})$, as shown in the next result.

Theorem 5 (Distribution of $|\boldsymbol{S}|$)**.** *Let $\boldsymbol{S} \sim W_p(\nu, \boldsymbol{\Sigma}\nu^{-1})$ $(\nu \geq p)$, then $|\boldsymbol{S}| \sim |\boldsymbol{\Sigma}|\nu^{-p}\prod_{i=1}^{p}\chi^2_{\nu-i+1}$, where $\chi^2_{\nu-i+1}$, $i = 1, 2, \cdots, p$ are independent random chi-square variables with degrees of freedom of $\nu-i+1$ to the ith factor of the product.*

Proof. The proof is immediate, since $\boldsymbol{S} \sim W_p(\nu, \boldsymbol{\Sigma}\nu^{-1})$. Therefore, replacing $|\boldsymbol{\Sigma}|$ by $|\boldsymbol{\Sigma}\nu^{-1}|$ in theorem 3, and using the fact that $|\boldsymbol{\Sigma}\nu^{-1}| = |\boldsymbol{\Sigma}|\nu^{-p}$, the result is obtained immediately. □

Anderson (2003) shows the exact distribution for two particular cases. The first one, for $p = 1$, is trivial and the second, for $p = 2$, results in the distribution of the product of two independent chi-square variables, which, according to the author, is also a chi-square. The next corollary presents these two cases.

Corollary 5.1 (Distribution of $|\boldsymbol{S}|$ for $p = 1$ and $p = 2$)**.** *Let $\boldsymbol{S} \sim W_p(\nu, \boldsymbol{\Sigma}\nu^{-1})$, then the distribution of $|\boldsymbol{S}|$ for $p = 1$ is $\sigma^2\chi^2_\nu\nu^{-1}$ and for $p = 2$ is $|\boldsymbol{\Sigma}|\ (\chi^2_{2(\nu-1)})^2(4\nu^2)^{-1}$.*

Proof. For $p = 1$, by the theorem 5 we have $\boldsymbol{\Sigma} = \sigma^2$, $\nu^{-p} = \nu^{-1}$, and $\prod_{i=1}^{1}\chi^2_{\nu-i+1} = \chi^2_\nu$. Therefore, the result follows immediately, i.e., the distribution is scaled chi-square with ν degrees of freedom and scale factor of σ^2/ν. The proof for $p = 2$ can be seen in Anderson (2003). □

The exact distribution of $|\boldsymbol{S}|$ is very complicated to process in real data, because it involves the distribution of products of independent chi-square variables. Hence, it is advisable to use some approximations to this distribution. An asymptotic normal approximation of this distribution is presented in Muirhead (1982) and Anderson (2003). The derivation of this approximation makes use of the delta method. On the other hand, Muirhead (1982) presents a method based on the characteristic function.

Theorem 6 (Asymptotic distribution of $|\boldsymbol{S}|$)**.** *Let* $\boldsymbol{S} \sim W_p(\nu, \boldsymbol{\Sigma}\nu^{-1})$ $(\nu \geq p)$*, then* $\sqrt{\frac{\nu}{2p}}\left(\frac{|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|} - 1\right)$ *has an asymptotically standard normal distribution,* $N(0{,}1)$.

Proof. Considering the delta method, let $|\boldsymbol{S}|/|\boldsymbol{\Sigma}|$ be $\prod_{i=1}^{p} \chi^2_{\nu-i+1}/\nu$, as shown in theorem 5. Notice, by the properties of a chi-square variable that $\mathbb{E}\left[\chi^2_{\nu-i+1}/\nu\right] = 1-(i-1)/\nu$ and $\mathbb{V}\left(\chi^2_{\nu-i+1}/\nu\right) = 2\nu^{-1} - 2(i-1)\nu^{-2}$. Let $\chi^2_{\nu-i+1}$ be the sum of square of $\nu-i+1$ standard normal variables $N(0{,}1)$, for $\nu \geq p$, then by the central limit theorem the asymptotic distribution of $\chi^2_{\nu-i+1}/\nu$ is a normal $N(1,\, 2\nu^{-1})$, since $(i-1)\nu^{-1}$ and $2(i-1)\nu^{-2}$ has limit equal to zero when $\nu \to \infty$. Considering the random vector, whose components are independent distributed, given by

$$\boldsymbol{U} = \begin{bmatrix} \chi^2_{\nu}/\nu \\ \chi^2_{\nu-1}/\nu \\ \cdots \\ \chi^2_{\nu-p+1}/\nu \end{bmatrix},$$

we noticed that $\boldsymbol{U}$ has a asymptotic multivariate normal distribution given by $N_p(\mathbf{1}_p,\, 2\nu^{-1}\boldsymbol{I})$.

Let a real value function be $h(\boldsymbol{U}) = \prod_{i=1}^{p} U_i = \prod_{i=1}^{p} \chi^2_{\nu-i+1}/\nu = |\boldsymbol{S}|/|\boldsymbol{\Sigma}|$, then we get $\boldsymbol{h}'(\boldsymbol{u}) = [\prod_{j\neq i=1}^{p} U_j]_i$ $(p \times 1)$, $i = 1, 2, \cdots, p$. By the delta method we have

$$\mathbb{E}\left[h(\boldsymbol{U})\right] \simeq h(\boldsymbol{\mu_U}) = 1$$

and

$$\mathbb{V}\left(h(\boldsymbol{U})\right) \simeq \boldsymbol{h}'^{\top}(\boldsymbol{\mu_U})\boldsymbol{\Sigma_U}\boldsymbol{h}'(\boldsymbol{\mu_U}) = 2\nu^{-1}\mathbf{1}_p^{\top}\mathbf{1}_p = 2p\nu^{-1}.$$

Considering that U_i is asymptotically normal, the first order approximation of $h(\boldsymbol{U})$ in the Taylor series will also be asymptotic normal. The consequence is that the asymptotic distribution of $|\boldsymbol{S}|/|\boldsymbol{\Sigma}|$ is $N_1(1,\, 2p\nu^{-1})$. We can also apprehend that $\sqrt{\nu}|\boldsymbol{S}|/|\boldsymbol{\Sigma}|$ has asymptotic normal distribution $N_1(\sqrt{\nu},\, 2p)$. Therefore, the result is achieved immediately using the $\sqrt{\nu/(2p)}(|\boldsymbol{S}|/|\boldsymbol{\Sigma}|-1)$ transformation, as we wanted to show. □

The normal approximation from theorem 5 is attributed to Anderson (2003). We noticed that the random variable $|\boldsymbol{W}|/|\boldsymbol{\Sigma}|$ has a distribution given by the product of chi-square random variables. Therefore,

$$U = \frac{\nu^p|\boldsymbol{S}|}{|\boldsymbol{\Sigma}|} = \frac{|\boldsymbol{W}|}{|\boldsymbol{\Sigma}|} \sim \prod_{i=1}^{p} \chi^2_{\nu-i+1}. \tag{5.6}$$

Two other normal approximations of the U distribution are reported in the literature. One of them needs the following result:

$$\ln(\chi^2_\nu) \sim N\left(\psi(\nu/2) + \ln(2),\, \psi'(\nu/2)\right), \tag{5.7}$$

where $\psi()$ and $\psi'()$ are the digamma (the logarithmic derivative of the gamma function) and trigamma (the first derivative of the digamma function) functions, respectively. Thus, let $Y = \ln(U) = p\ln(\nu) + \ln(|\boldsymbol{S}|) - \ln(|\boldsymbol{\Sigma}|)$, then we have

$$Y \sim N\left(\sum_{i=1}^{p} \psi((\nu-p+1)/2) + p\ln(2),\, \sum_{i=1}^{p} \psi'((\nu-p+1)/2)\right). \tag{5.8}$$

This normal approximation is due to Sarkar (1989) and recommended for $p > 3$.

The second normal approximation was developed by Djauhari (2009), and is given by

$$|\boldsymbol{S}| \sim N\left(b_1|\boldsymbol{\Sigma}|,\, b_2|\boldsymbol{\Sigma}|^2\right), \tag{5.9}$$

where

$$b_1 = \frac{1}{\nu^p}\prod_{i=1}^{p}(\nu-i+1) \qquad \text{and} \qquad b_2 = \frac{b_1}{\nu^p}\prod_{i=1}^{p}(\nu-i+3) - b_1^2.$$

5.3 Inferences on Normal sample generalized variance

Considering the exact and approximate distributions for some functions of the $|\boldsymbol{S}|$ presented previously, we can consider the hypothesis tests for $|\boldsymbol{\Sigma}|$ from normal populations. The null and alternative hypotheses are

$$\begin{cases} H_0^{(a)} : |\boldsymbol{\Sigma}| = \eta & \text{against} \quad H_1^{(a)} : |\boldsymbol{\Sigma}| \neq \eta \\ H_0^{(b)} : |\boldsymbol{\Sigma}| \leq \eta & \text{against} \quad H_1^{(b)} : |\boldsymbol{\Sigma}| > \eta \\ H_0^{(c)} : |\boldsymbol{\Sigma}| \geq \eta & \text{against} \quad H_1^{(c)} : |\boldsymbol{\Sigma}| < \eta, \end{cases} \tag{5.10}$$

where $\eta > 0$ is a previously specified real value derived from some real problem of interest. For an exact test, a similar Monte Carlo version to the one proposed by Jafari e Kazemi (2014) was used. Thus, in the three hypothesis cases (5.10), we initially computed the value of the test statistic by

$$U_c = \frac{\nu^p |\boldsymbol{S}|}{|\boldsymbol{\Sigma}_0|} = \frac{\nu^p |\boldsymbol{S}|}{\eta}, \tag{5.11}$$

that under H_0 has distribution of $\prod_{i=1}^{p} \chi^2_{\nu-i+1}$, where $\nu = n-1$.

We considered a computational alternative to avoid overflow issues. Hence, the following Monte Carlo algorithm was used to obtain p-values for testing the above null hypotheses on $\boldsymbol{\Sigma}$:

Algorithm 1**:** Given p, n, and $|\boldsymbol{s}|$:

1. Generate $\ln(U) = \sum_{i=1}^{p} \ln(\chi^2_{\nu-i+1})$, simulating $\chi^2_{\nu-i+1}$ for each $i = 1, 2, \cdots, p$.

2. Calculate $V_0 = p\ln(\nu) + \ln(|\boldsymbol{s}|) - \ln(U)$ and $V = \exp(V_0)$.

3. Repeat steps $1-2$ for a large number os times, i.e., $m = 5000$ and obtain m values of V, denoting them by V_j, $j = 1, 2, \cdots, m$.

4. Calculate the $p$value for each case of (5.10), respectively, by

$$\begin{cases} q = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} I_{[0,\eta]}(V_j), & p\text{-value} = 2\min(q, 1-q) \quad \text{for } H_0^{(a)}, \\ q = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} I_{[0,\eta]}(V_j), & p\text{-value} = q \quad \text{for } H_0^{(b)}, \\ q = \dfrac{1}{m}\displaystyle\sum_{j=1}^{m} I_{[0,\eta]}(V_j), & p\text{-value} = 1-q \quad \text{for } H_0^{(c)}, \end{cases} \tag{5.12}$$

   where $I_{[0,\eta]}(V_j)$ is the indicator function that é a função indicadora that returns 1 if $V \leq \eta$ and 0, otherwise.

We can also apply the test using any of the three normal approximations showed. We will start by presenting Anderson's approach (ANDERSON, 2003) in full details. Subsequently, we will present only essential results for the other two approaches. Thus, in the case of the normal approximation of Anderson (2003), for testing one of the three cases of the null hypothesis in (5.10), we initially compute the test statistic by

$$Z_c = \sqrt{\frac{\nu}{2p}}\left(\frac{|\boldsymbol{S}|}{\eta} - 1\right), \tag{5.13}$$

where $\nu = n-1$. The corresponding p-value depends on the hypothesis being tested. For $H_0^{(a)}$, we have

$$p\text{-value} = 2(1-\Phi(|Z_c|)), \tag{5.14}$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution evaluated at x. For $H_0^{(b)}$ and $H_0^{(c)}$, we have

$$p\text{-value} = 1-\Phi(Z_c) \qquad \text{and} \qquad p\text{-value} = \Phi(Z_c), \tag{5.15}$$

respectively. If the p-value was less or equal to the nominal significance level α, the null hypothesis H_0 should be rejected.

A similar approach was used for the Sarkar (1989)'s test, with $p > 3$. Initially, the test statistic, given by

$$Z_c = \frac{p\ln(\nu) + \ln(|\boldsymbol{S}|) - \ln(\eta) - \mu_Y}{\sigma_Y} \tag{5.16}$$

should be computed, where $\mu_Y = \sum_{i=1}^{p} \psi((\nu - p + 1)/2) + p\ln(2)$ and $\sigma_Y^2 = \sum_{i=1}^{p} \psi'((\nu - p + 1)/2)$. The corresponding p-value depends on the null hypothesis being tested. For $H_0^{(a)}$, we have

$$p\text{-value} = 2(1-\Phi(|Z_c|)) \tag{5.17}$$

where $\Phi(x)$ is the cumulative distribution function of the standard normal distribution evaluated at x. For $H_0^{(b)}$ and $H_0^{(c)}$, we have

$$p\text{-value} = 1-\Phi(Z_c) \qquad \text{and} \qquad p\text{-value} = \Phi(Z_c), \tag{5.18}$$

respectively. If the p-value is less or equal to the nominal significance level α, the null hypothesis H_0 should be rejected.

Again for the Djauhari (2009)'s approach, the process is similar to the previous cases. The test statistic is given by

$$Z_c = \frac{|\boldsymbol{S}| - b_1\eta}{\eta\sqrt{b_2}}, \tag{5.19}$$

where

$$b_1 = \frac{1}{\nu^p}\prod_{i=1}^{p}(\nu-i+1) \qquad \text{and} \qquad b_2 = \frac{b_1}{\nu^p}\prod_{i=1}^{p}(\nu-i+3) - b_1^2.$$

For $H_0^{(a)}$, the p-value is given by

$$p\text{-value} = 2(1-\Phi(|Z_c|)). \tag{5.20}$$

For $H_0^{(b)}$ and $H_0^{(c)}$, the p-values are

$$p\text{-value} = 1-\Phi(Z_c) \qquad \text{and} \qquad p\text{-value} = \Phi(Z_c), \tag{5.21}$$

respectively.

5.4 Likelihood ratio test on normal generalized variance

The likelihood ratio test for the null hypothesis H_0: $|\boldsymbol{\Sigma}| = \eta$, $\eta > 0$, is developed in this work under multivariate normality. Let $\boldsymbol{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n$ be a random sample from this distribution. We know that the unrestricted maximum likelihood estimators of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are well known and given $\bar{\boldsymbol{X}}_.$ and $\hat{\boldsymbol{\Sigma}} = n^{-1}\sum_{j=1}^{n}(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_.)(\boldsymbol{X}_j - \bar{\boldsymbol{X}}_.)^\top$, respectively

(MUIRHEAD, 1982). Moreover, the unrestricted likelihood function is

$$L_{\Omega}(\boldsymbol{X};\boldsymbol{\mu},\boldsymbol{\Sigma}) =(2\pi)^{-np/2}|\boldsymbol{\Sigma}|^{-n/2}\exp\left\{-\frac{1}{2}tr\left[\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{W}+ + n\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)^{\top}\right]\right]\right\}, \tag{5.22}$$

and its maximum is

$$L_{\Omega}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}}) =(2\pi)^{-np/2}|\hat{\boldsymbol{\Sigma}}|^{-n/2}\exp\left\{-\frac{np}{2}\right\}. \tag{5.23}$$

Under H_0, with $|\boldsymbol{\Sigma}| = \eta$ and denoting $\boldsymbol{\Sigma}$ by $\boldsymbol{\Sigma}_0$ to differentiate from the unrestricted case, the restricted likelihood function is given by

$$L_{\Omega_0}(\boldsymbol{X};\boldsymbol{\mu},\boldsymbol{\Sigma}_0,\eta) =(2\pi)^{-np/2}\eta^{-n/2}\exp\left\{-\frac{1}{2}tr\left[\boldsymbol{\Sigma}_0^{-1}\left[\boldsymbol{W}+ + n\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)^{\top}\right]\right]\right\}. \tag{5.24}$$

The log-likelihood function is given by

$$g_{\Omega_0}(\boldsymbol{X};\boldsymbol{\mu},\boldsymbol{\Sigma}_0,\eta) =-\frac{np}{2}\ln(2\pi)-\frac{n}{2}\ln(\eta)-\frac{1}{2}tr\left[\boldsymbol{\Sigma}_0^{-1}\left[\boldsymbol{W}+ + n\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right)^{\top}\right]\right]. \tag{5.25}$$

Taking the first derivative of the log-likelihood function (5.25) with respect to $\boldsymbol{\mu}$ and equating it to zero, we will have the $\boldsymbol{\mu}$ estimator that maximizes (5.24), for $\boldsymbol{\Sigma}_0$ fixed. Therefore, the derivative is

$$\frac{\partial g_{\Omega_0}(\boldsymbol{X};\boldsymbol{\mu},\boldsymbol{\Sigma}_0,\eta)}{\partial\boldsymbol{\mu}} =\boldsymbol{\Sigma}_0^{-1}\left(\bar{\boldsymbol{X}}_{\cdot}-\boldsymbol{\mu}\right),$$

where the solution when equated to $\mathbf{0}$ results in the maximum likelihood estimator given by

$$\hat{\boldsymbol{\mu}} =\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{X}_j=\bar{\boldsymbol{X}}_{\cdot}. \tag{5.26}$$

Therefore, the function

$$L_{\Omega_0}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\boldsymbol{\Sigma}_0,\eta) = (2\pi)^{-np/2}\eta^{-n/2}\exp\left\{-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}\right)\right\} \tag{5.27}$$

is such that $L_{\Omega_0}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\boldsymbol{\Sigma}_0,\eta) \geq L_{\Omega_0}(\boldsymbol{X};\boldsymbol{\mu},\boldsymbol{\Sigma}_0,\eta)$.

The corresponding log-likelihood function is

$$g_{\Omega_0}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\boldsymbol{\Sigma}_0,\eta) = -\frac{np}{2}\ln(2\pi) - \frac{n}{2}\ln(\eta) - \frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}\right). \tag{5.28}$$

We should maximize the function (5.27) or (5.28) in respect to the only remaining parameter, which is $\boldsymbol{\Sigma}_0$. This is equivalent to minimizing $\operatorname{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}\right)$ subject to the restriction imposed by H_0 given by $|\boldsymbol{\Sigma}_0| = \eta$, $\eta > 0$. Using Lagrange multipliers, we have the Lagrangian function, denoted by Φ and given by

$$\Phi(\boldsymbol{\Sigma}_0;\eta) = \operatorname{tr}\left(\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}\right) + \lambda\left(|\boldsymbol{\Sigma}_0| - \eta\right). \tag{5.29}$$

The partial derivatives by respect to $\boldsymbol{\Sigma}_0$ and λ are

$$\frac{\partial\Phi(\boldsymbol{\Sigma}_0;\eta)}{\partial\boldsymbol{\Sigma}_0} = -\boldsymbol{\Sigma}_0^{-1}\boldsymbol{W}\boldsymbol{\Sigma}_0^{-1} + \lambda|\boldsymbol{\Sigma}_0|\boldsymbol{\Sigma}_0^{-1} \qquad \text{and} \qquad \frac{\partial\Phi(\boldsymbol{\Sigma}_0;\eta)}{\partial\lambda} = |\boldsymbol{\Sigma}_0| - \eta,$$

that when equal to zero, we have from the second part that

$$|\hat{\boldsymbol{\Sigma}}_0| = \eta.$$

Replacing this result in the first part, we get

$$\begin{aligned} -\hat{\boldsymbol{\Sigma}}_0^{-1}\boldsymbol{W}\hat{\boldsymbol{\Sigma}}_0^{-1} + \lambda|\hat{\boldsymbol{\Sigma}}_0|\hat{\boldsymbol{\Sigma}}_0^{-1} &= \boldsymbol{0} \\ \lambda|\hat{\boldsymbol{\Sigma}}_0|\hat{\boldsymbol{\Sigma}}_0 &= \boldsymbol{W} \quad \text{(after some algebra)} \\ \lambda\eta\hat{\boldsymbol{\Sigma}}_0 &= \boldsymbol{W} \quad (\text{replacing } |\hat{\boldsymbol{\Sigma}}|_0 = \eta). \end{aligned} \tag{5.30}$$

Taking the determinant on both sides of the last equation, we have

$$\lambda^p\eta^p|\hat{\boldsymbol{\Sigma}}_0| = |\boldsymbol{W}| = n^p|\hat{\boldsymbol{\Sigma}}|,$$

that results in

$$\begin{aligned}\lambda =& \frac{\sqrt[p]{|\boldsymbol{W}|}}{\eta^{(p+1)/p}} \quad \text{(replacing } |\hat{\boldsymbol{\Sigma}}|_0 = \eta)\\ =& \frac{n\sqrt[p]{|\hat{\boldsymbol{\Sigma}}|}}{\eta^{(p+1)/p}}.\end{aligned}$$

Replacing this solution of λ in (5.30), we get

$$\frac{n\sqrt[p]{|\hat{\boldsymbol{\Sigma}}|}}{\eta^{(p+1)/p}}\eta\hat{\boldsymbol{\Sigma}}_0 = n\hat{\boldsymbol{\Sigma}} = \boldsymbol{W},$$

resulting in the maximum likelihood estimator of $\boldsymbol{\Sigma}_0$, given by

$$\hat{\boldsymbol{\Sigma}}_0 = \frac{\sqrt[p]{\eta}}{\sqrt[p]{|\hat{\boldsymbol{\Sigma}}|}}\hat{\boldsymbol{\Sigma}} = \frac{\sqrt[p]{\eta}}{\sqrt[p]{|W|}}\boldsymbol{W}. \tag{5.31}$$

Therefore, the maximum of the restricted likelihood function is

$$\begin{aligned}L_{\Omega_0}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}}_0,\eta) =& (2\pi)^{-np/2}\eta^{-n/2}\exp\left\{-\frac{p\sqrt[p]{|\boldsymbol{W}|}}{2\sqrt[p]{\eta}}\right\}\\ =& (2\pi)^{-np/2}\eta^{-n/2}\exp\left\{-\frac{np\sqrt[p]{|\hat{\boldsymbol{\Sigma}}|}}{2\sqrt[p]{\eta}}\right\}.\end{aligned} \tag{5.32}$$

The likelihood ratio test statistic is given by

$$\Lambda = \frac{L_{\Omega_0}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}}_0,\eta)}{L_{\Omega}(\boldsymbol{X};\hat{\boldsymbol{\mu}},\hat{\boldsymbol{\Sigma}})} = \left(\frac{\eta}{|\hat{\boldsymbol{\Sigma}}|}\right)^{-n/2}\exp\left\{-\frac{np}{2}\left[\sqrt[p]{\frac{|\hat{\boldsymbol{\Sigma}}|}{\eta}}-1\right]\right\}. \tag{5.33}$$

In the unrestricted model we have a dimension given by $p+p(p+1)/2$ and in the constrained model we have a dimension given by $p+p(p+1)/2-1$. Therefore, under H_0: $|\boldsymbol{\Sigma}| = \eta$, we have, using the general theory of likelihood ratio tests, that $-2\ln(\Lambda)$, given by

$$\chi^2_c = n\left[\ln(\eta)-\ln(|\hat{\boldsymbol{\Sigma}}|)\right]+np\left[\sqrt[p]{\frac{|\hat{\boldsymbol{\Sigma}}|}{\eta}}-1\right],$$

has asymptotic chi-square distribution with $\nu = 1$ degree of freedom.

Example 1. *Use the example data where six hematological variables were measured at $n =$ 103 individuals (ROYSTON, 1983; JAFARI; KAZEMI, 2014) and apply the test for the null hypothesis H_0: $|\boldsymbol{\Sigma}| = 6.0$ against H_1: $|\boldsymbol{\Sigma}| \neq 6.0$ considering a confidence coefficient of 95%. The sample estimate was $|\boldsymbol{s}| = 6.2453$.*

The unrestricted maximum likelihood estimate of $|\boldsymbol{\Sigma}|$ is obtained by

$$|\hat{\boldsymbol{\Sigma}}| = \frac{(n-1)^p}{n^p}|\boldsymbol{s}| = 5.890213.$$

Thus, the test statistic is

$$\begin{aligned}\chi_c^2 =& n\left[\ln(\eta) - \ln(|\hat{\boldsymbol{\Sigma}}|)\right] + np\left[\sqrt[p]{\frac{|\hat{\boldsymbol{\Sigma}}|}{\eta}} - 1\right]\\ =& 103\left[\ln(6) - \ln(5.890213)\right] + 103 \times 6 \times \left[\sqrt[6]{\frac{5.890213}{6}} - 1\right] = 0.00292,\end{aligned}$$

whose p-value is 0.9569, which leads to non-rejection of H_0 at the 5% significance nominal level.

By the general theory of LRTs, this testing procedure is not appropriate for small sample sizes. Najarzadeh (2017) points out the standard approach to this problem, which consists of modifying the LRT statistic using Bartlett's correction. In this approach, to adjust the LRT, we use statistics $-2\phi\ln(\Lambda)$, where $\phi = \frac{n-1}{\mathbb{E}[-2\ln(\Lambda)]}$ is Bartlett correction factor. Using the 4th order Taylor polynomial approximation on the functions $|\boldsymbol{S}|^{1/p}$ and $\ln(|\boldsymbol{S}|)$ centered about the point $\mathbb{E}[|\boldsymbol{S}|]$, we find the following results

$$\begin{aligned}\mathbb{E}[-2\ln(\Lambda)] &= \mathbb{E}\left[n\ln(\eta) - n\ln(|\hat{\boldsymbol{\Sigma}}|) + np\left(\frac{|\hat{\boldsymbol{\Sigma}}|}{\eta}\right)^{1/p} - np\right]\\ &= n\ln(\eta) - n\mathbb{E}\left[\ln(|\hat{\boldsymbol{\Sigma}}|)\right] + \frac{np}{\eta^{1/p}}\mathbb{E}\left[|\hat{\boldsymbol{\Sigma}}|^{1/p}\right] - np\\ &= n\ln(\eta) - np + \frac{np}{\eta^{1/p}}\mathbb{E}\left[|\hat{\boldsymbol{\Sigma}}|^{1/p}\right] - n\mathbb{E}\left[\ln(|\hat{\boldsymbol{\Sigma}}|)\right]\\ &= n\ln(\eta) - np + \frac{np}{\eta^{1/p}}\mathbb{E}\left[\left(\frac{(n-1)^p}{n^p}|\boldsymbol{S}|\right)^{1/p}\right] - n\mathbb{E}\left[\ln\left(\frac{(n-1)^p}{n^p}|\boldsymbol{S}|\right)\right]\\ &= n\ln(\eta) - np + \frac{np}{\eta^{1/p}}\frac{(n-1)}{n}\mathbb{E}\left[|\boldsymbol{S}|^{1/p}\right] - np\ln\left(\frac{n-1}{n}\right) - n\mathbb{E}\left[\ln(|\boldsymbol{S}|)\right],\end{aligned}$$

where

$$\mathbb{E}\left[|\boldsymbol{S}|^{1/p}\right] \simeq (\mathbb{E}[|\boldsymbol{S}|])^{\frac{1}{p}} + \frac{1}{2p}\left(\frac{1}{p}-1\right)\mathbb{E}[|\boldsymbol{S}|]^{\frac{1}{p}-2}\mathbb{V}(|\boldsymbol{S}|) + \\ + \frac{1}{6p}\left(\frac{1}{p}-1\right)\left(\frac{1}{p}-2\right)\mathbb{E}[|\boldsymbol{S}|]^{\frac{1}{p}-3}\mathbb{E}\left[(|\boldsymbol{S}|-E|\boldsymbol{S}|)^3\right] \\ + \frac{1}{24p}\left(\frac{1}{p}-1\right)\left(\frac{1}{p}-2\right)\left(\frac{1}{p}-3\right)\mathbb{E}[|\boldsymbol{S}|]^{\frac{1}{p}-4}\mathbb{E}\left[(|\boldsymbol{S}|-E|\boldsymbol{S}|)^4\right]$$

and

$$\mathbb{E}[\ln(|\boldsymbol{S}|)] \simeq \ln(\mathbb{E}[|\boldsymbol{S}|]) - \frac{\mathbb{V}(|\boldsymbol{S}|)}{2(\mathbb{E}[|\boldsymbol{S}|])^2} + \frac{2}{6(\mathbb{E}[|\boldsymbol{S}|])^3}\mathbb{E}\left[(|\boldsymbol{S}|-\mathbb{E}[|\boldsymbol{S}|])^3\right] - \\ - \frac{6}{24(\mathbb{E}[|\boldsymbol{S}|])^4}\mathbb{E}\left[(|\boldsymbol{S}|-\mathbb{E}[|\boldsymbol{S}|])^4\right]$$

Here, we can use the fact that $\mathbb{E}[|\boldsymbol{S}|] = \dfrac{\mathbb{E}[|\boldsymbol{W}|]}{\nu^p}$ and the theorem 4 to calculate $\mathbb{E}[|\boldsymbol{W}|^r]$, $r \geq 1$. Furthermore, all the central moments needed to find Taylor's approximation was expanded to non-central moments in which it was possible to apply theorem 4 directly. Thus, we can find Bartlett's correction $\phi = \dfrac{n-1}{\mathbb{E}[-2\ln(\Lambda)]}$. Therefore, we reject the null hypothesis of H_0 at the nominal significance level α, if the value of the modified Bartlett correction statistic $-2\phi\ln(\Lambda)$, be greater than the upper-tail α critical value of the chi-square distribution with 1 degrees of freedom.

5.5 Monte Carlo performance evaluation

Monte Carlo simulations were used to compare the actual sizes and powers of the following tests: i) Monte Carlo exact test (MCET) using algorithm 1, with $m = 2000$ replications, ii) Anderson test (AT), iii) Sarkar test (ST), iv) Djauhari test, v) the proposed likelihood ratio test (LRT), and vi) Bartlett corrected likelihood ratio test (BCLRT). The two-sided hypothesis

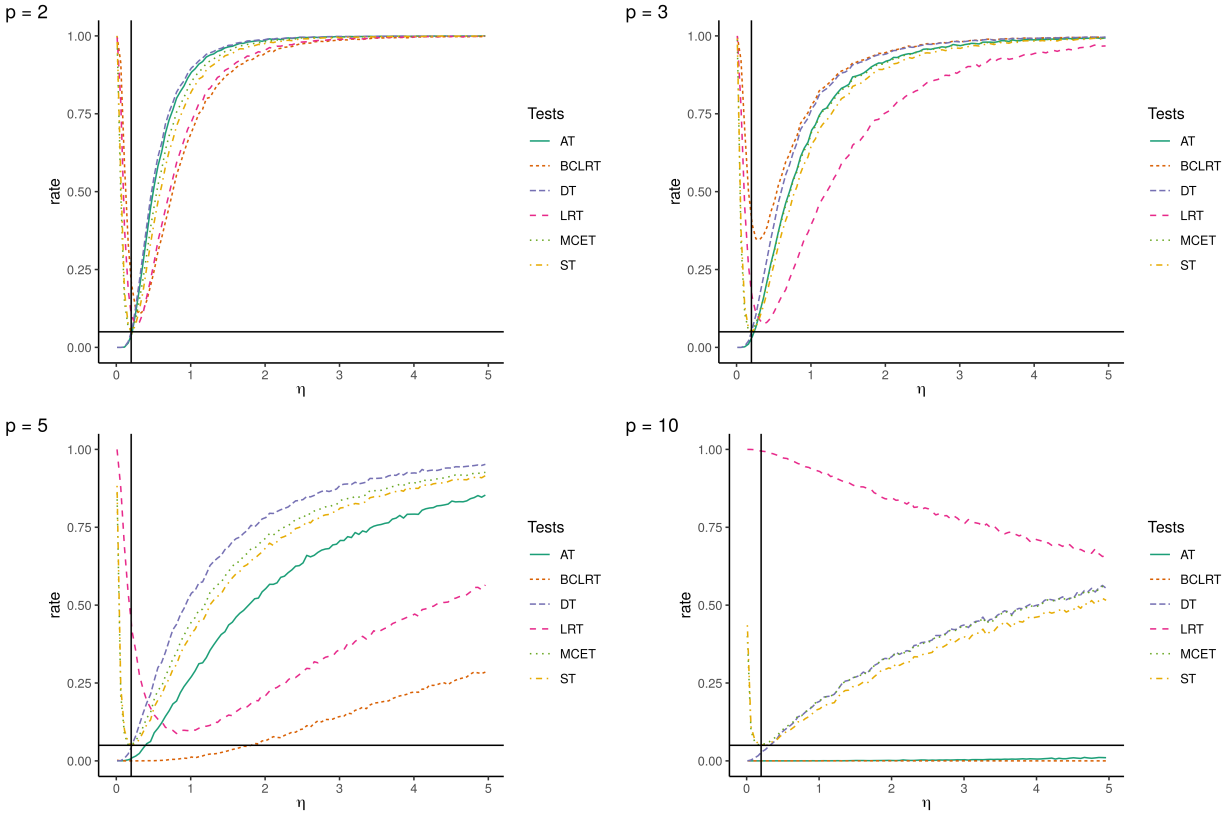
$$H_0^{(a)}:\ |\boldsymbol{\Sigma}| = \eta \qquad \text{against} \qquad H_1^{(a)}:\ |\boldsymbol{\Sigma}| \neq \eta$$

were considered values of η, as 0.2 and 1.0, sample sizes n $(15, 30, 50)$ and dimensions p $(2, 3, 5, 10)$. Also, were considered in 10000 Monte Carlo replications, in the same cases of Jafari e Kazemi (2014). Samples of size n and dimension p were generate from the multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$. The six tests were applied in each case at the nominal significance level of α $(0.01, 0.05\ 0.10)$. In the two cases with respect to η (0.2 and 1), the samples were generated from multivariate normal populations with actual $|\mathbf{\Sigma}|$ ranging from 0.01 to 5.00 with step size of 0.05. The empirical test powers and sizes were computed in each configuration for each set of 10000 simulations.

The evaluation of the tests will be performed graphically, more precisely through the empirical graph of the power function for each evalueted test. Three figures will be presented, each containing 4 graphs. In the first figure, the sample for the test is $n = 15$, with the number of p variables chosen being $2, 3, 5$ and 10. The level of significance was set at 5%. $N = 10{,}000$ normal p-varied samples were generated for each p described and for each covariance matrix $\mathbf{\Sigma}$ such that their $|\mathbf{\Sigma}|$ values form a sequence between the numbers 0.01 and 5.00 in increments of 0.05. All analysis was build using *R software* (R CORE TEAM, 2019).

Under H_0, $|\mathbf{\Sigma}| = 0.2$ was considered, so the type I error rate will be estimated when the sample of the normal p-varied is generated from a normal p-varied with covariance matrix, whose determinant is equal to 0.2. In Figure 5.1, we show the performance of the tests when the sample is of size 15 and $\eta = 0.2$.

Figura 5.1 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 0.2$, $n = 15$ and $p = 2$, 3, 5 and 10.



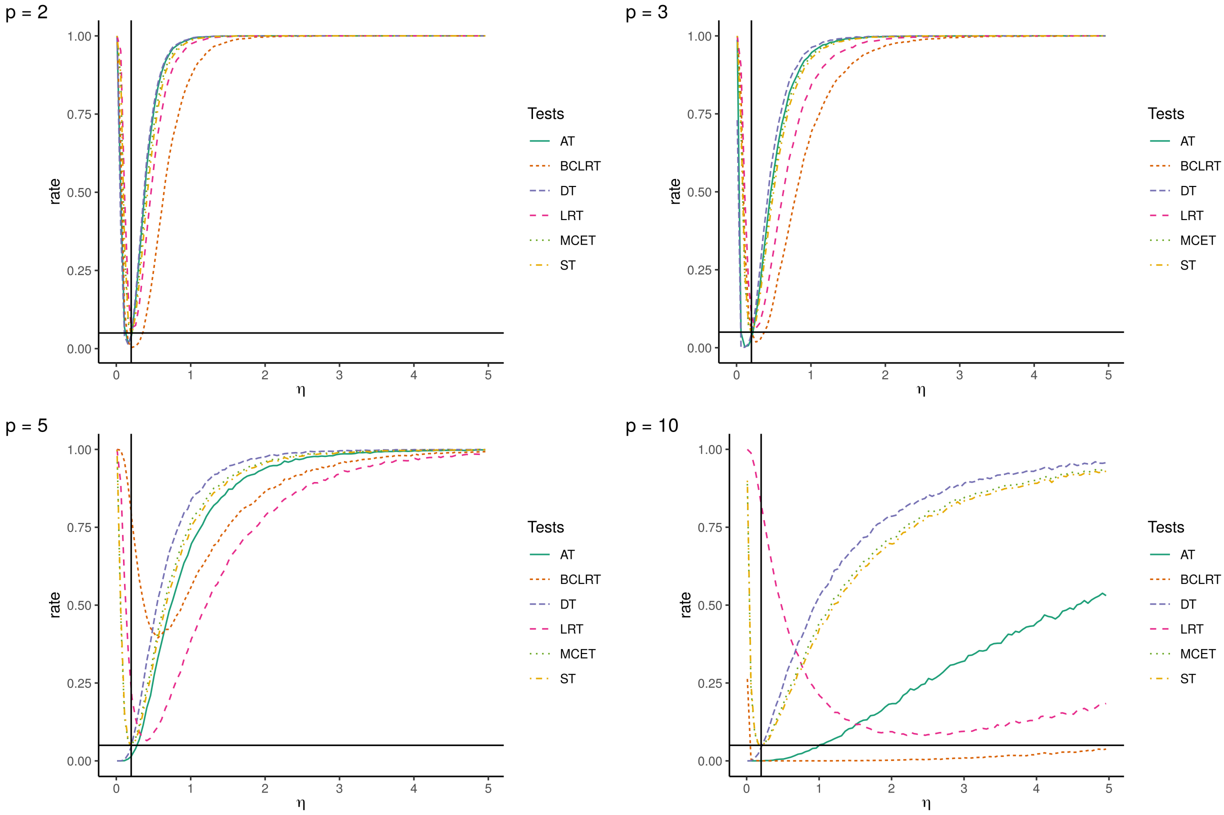
It can be seen in Figure 5.1 that when $p = 2$, all tests are able to control the type I error rate, while the form of the function of every tests has a similar behavior in relation to power. Regarding power, the AT and DT tests stood out in relation to the others, while the proposed LRT and BCLRT tests have less power than the other tests for most situations.

For the other cases of p, the BCLRT test was very far from controlling the type I error rate. Hence, the power of the test is not be must considered, while the LRT test for the number of p variables equal to 3 and 5 can be considered liberal, since its estimated type I error rate is greater than the 5% significance level. In addition, the power of the LRT test is worse than the other tests.

For $p = 10$, the LRT test did not control the type I error rate, whereas the AT and BCLRT tests are close to zero for all situations. Although not showing high power, the DT, MCET and ST tests controlled the type I error rate, with ST having lower power than the other tests.

The next figure show the performance of the tests when the sample is of size 30 and $\eta = 0.2$.

Figura 5.2 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 0.2$, $n = 30$ and $p = 2$, 3, 5 and 10.

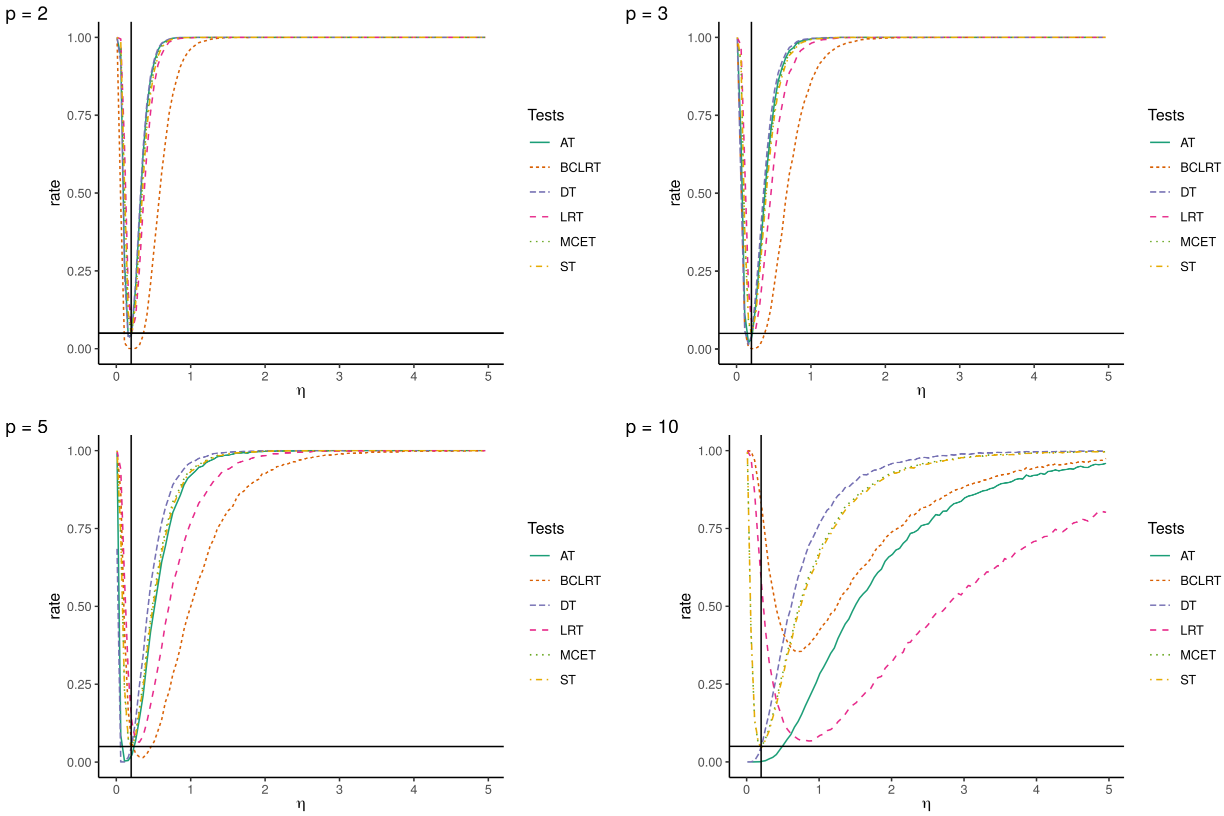


For the $p = 2$ and $p = 3$, all tests are able to control the type I error rate, while the form of the function of all tests has a similar behavior in relation to power. About the power, the DT and MCET tests stood out in relation to the others, while the proposed LRT and BCLRT tests have less power than the other tests for most situations. For the other cases of p, the BCLRT test was very far from controlling the type I error rate, while the LRT test for the number of p variables equal to 3 and 5 can be considered liberal, since its estimated type I error rate is greater than the 5% significance level.

For $p = 10$, the LRT test did not control the type I error rate, whereas the AT and BCLRT tests are close to zero for all situations. Although not showing high power, the DT, MCET and ST tests controlled the type I error rate, with ST having lower power than the other tests.

The next figure show the performance of the tests when the sample is of size 50 and $\eta = 0.2$.

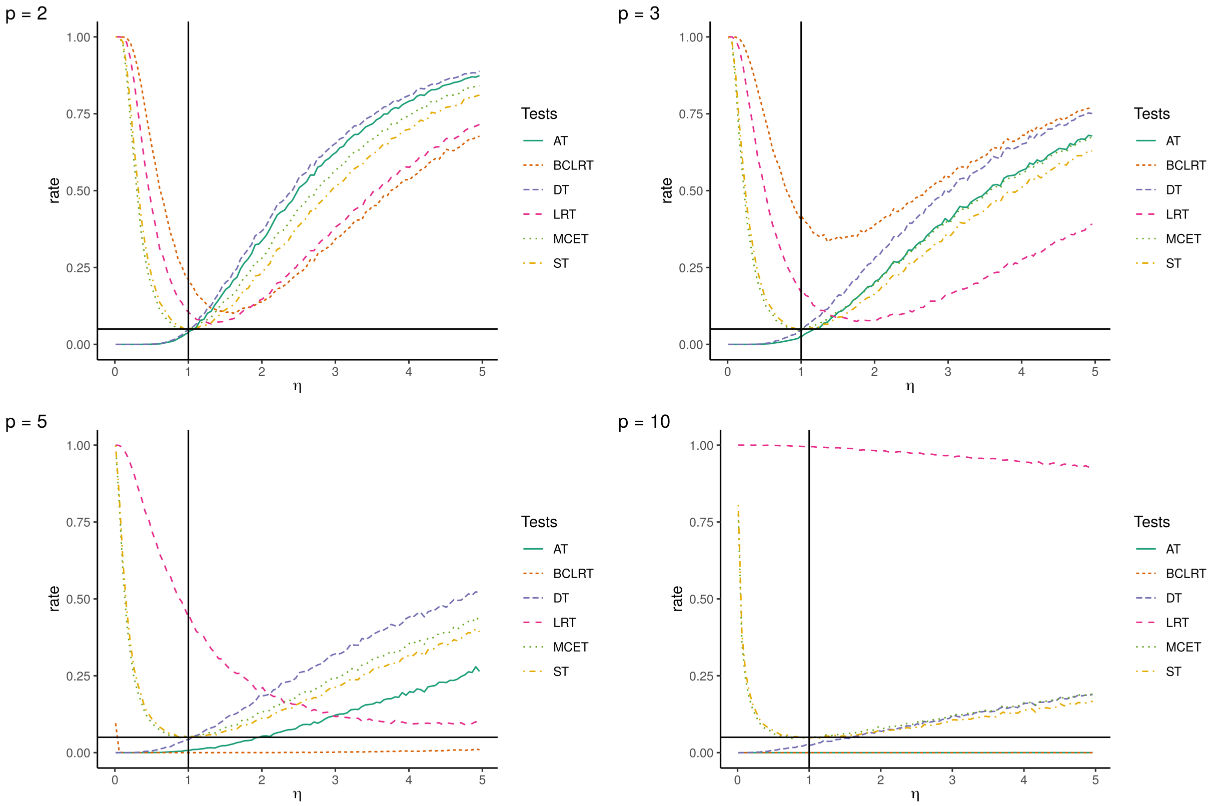
Figura 5.3 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 0.2$, $n = 50$ and $p = 2$, 3, 5 and 10.



For the $p = 2$, 3, and 5, all tests are able to control the type I error rate, while the form of the function of all tests has a similar behavior in relation to power. About the power, the DT and MCET tests stood out in relation to the others, while the proposed LRT and BCLRT tests have less power than the other tests for most situations. For the $p = 10$, the LRT and BCLRT test did not control the type I error rate. Regarding power, the DT, ST and MCET tests obtained the best results.

Under H_0, $|\mathbf{\Sigma}| = 1$ was considered, so the type I error rate will be estimated when the sample of the normal p-varied is generated from a normal p-varied with covariance matrix, whose determinant is equal to 1. In figure 5.4, we show the performance of the tests when the sample is of size 15 and $\eta = 1$.

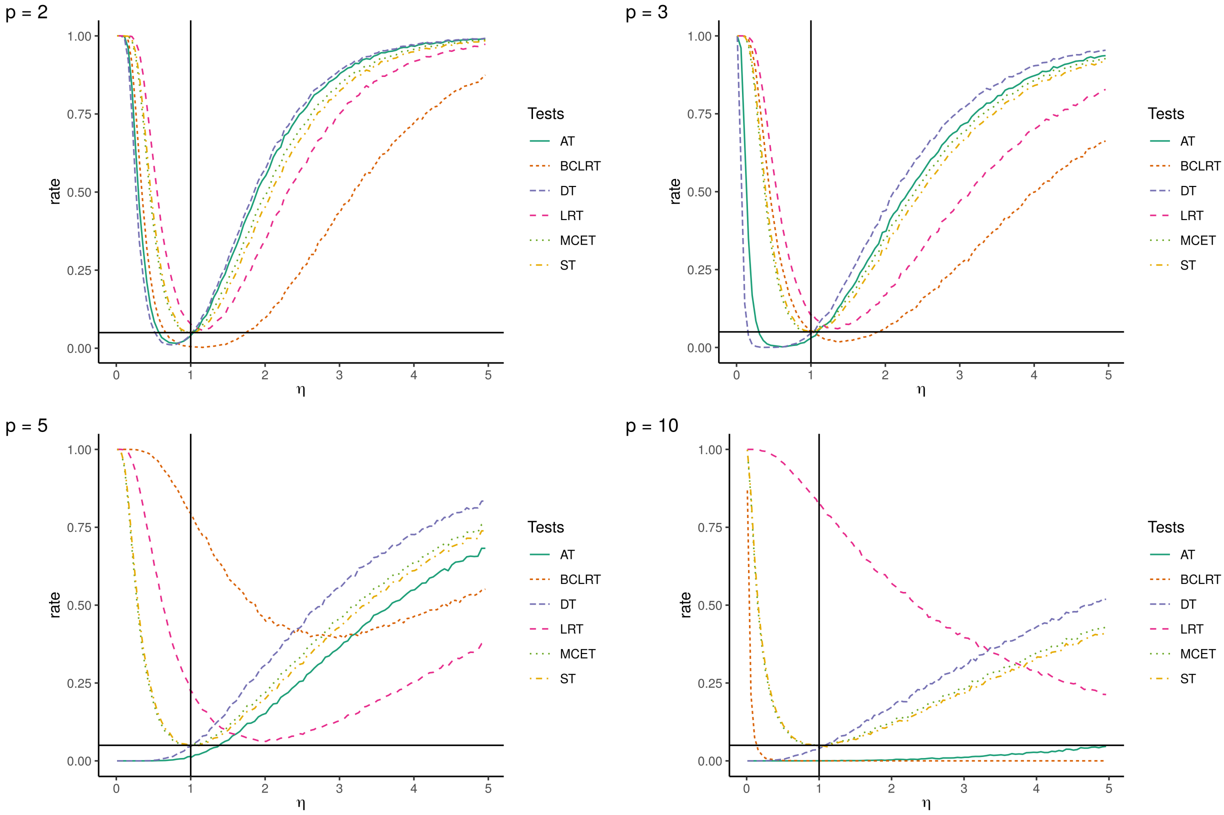
Figura 5.4 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 1$, $n = 15$ and $p = 2, 3, 5$ and 10.



For the $p = 2$, the LRT and BCLRT tests can be considered liberal, ST and MCET controlled the type I error rate. When $p = 3$, the BCLRT test did not control the type I error rate, the LRT test can be considered liberal, ST and MCET obtained the best performances. For $p = 5$, the performance of the AT, LRT and BCLRT tests were poor. The only tests that managed to control the type I error rate for $p = 10$, were the ST and MCET tests, however they did not obtain high power.

The next figure show the performance of the tests when the sample is of size 30 and $\eta = 1$.

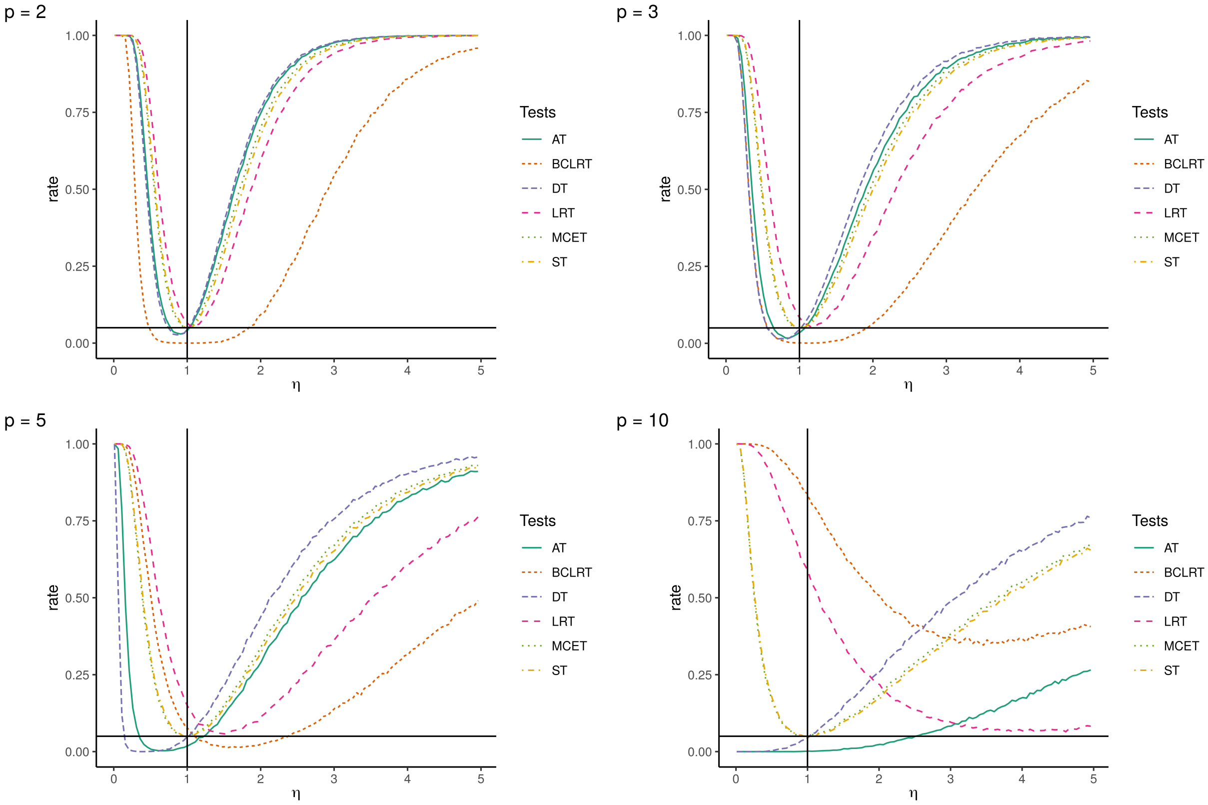
Figura 5.5 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 1$, $n = 30$ and $p = 2$, 3, 5 and 10.



For the p equal to 2 and 3, the ST and MCET tests controlled the type I error rate, while the AT, DT and BCLRT tests can be considered conservative and the LRT test considered liberal. AT, DT and MCET obtained the best results in relation to power. For $p = 5$, the DT test obtained greater power and the ST and MCET tests well controlled the type I error rate. For $p = 10$, the AT, LRT and BCLRT tests were worse than the others.

The next figure show the performance of the tests when the sample is of size 50 and $\eta = 1$.

Figura 5.6 – Performance evaluation of hypothesis tests at the 5% significance level for $\eta = 1$, $n = 50$ and $p = 2$, 3, 5 and 10.



For the p equal to 2, 3 and 5, the ST and MCET tests controlled the type I error rate, while the AT, DT and BCLRT tests can be considered conservative and the LRT test considered liberal. For $p = 10$, ST and MCET they obtained better performance.

The other results, as well as the implementation of the tests and evaluation are in the complementary materials.

5.6 Conclusion

We proposed the LRT and BCLRT tests for testing the hypothesis that generalized variance is equal to a η, such that $\eta \in \mathbb{R}$. All development this test was build theorically. The LRT and BCLRT tests do not control the type I error rate as the number of p variables increases, while still having low power. In addition, they were inferior in these questions to tests already existing in the literature, specifically the ST and MCET tests. However, the LRT test controls the type I error rate for the number of variables $p = 2$ and $p = 3$ and has good power performance even if n is small. The LRT test also has control of the type I error rate for p = 5, when $n \geq 50$. The BCLRT test has the same good performance only for $p = 2$. Both tests perform better as n increases. Therefore, we recommend using the BCLRT test only in scenarios in that we have $p = 2$, especially if $n > 30$. As for the LRT test, we recommend its use in situations where we have $p = 2$ and $p = 3$ for $n > 30$ and for $p = 5$ when $n > 50$.

5.7 Complementary Material

```

# Exact test for the normal generalized variance - MCET
# H_0: |Sig| = Delta_0 and confidence interval
exactGV.Test <- function(Delta_0, n, p, detS, alpha = 0.05,
alternative="two.sided", m = 5000)
{
j <- 1:p
nu <- n - j
nullDist <- function(n, p, nu)
{
lnU <- sum(log(rchisq(p, nu)))
V <- p * log(n - 1) - lnU + log(detS)
return(exp(V))
}
V <- matrix(n, m, 1)
V <- apply(V, 1, nullDist, p, nu)
#hist(V)
p.value <- length(V[V <= Delta_0]) / m
if (alternative == "two.sided") {
p.value <- 2 * min(p.value, 1 - p.value)
LL <- quantile(V, alpha / 2)
UL <- quantile(V, 1 - alpha / 2)
} else
if (alternative == "less") {
p.value <- 1 - p.value
LL <- 0
UL <- quantile(V, 1 - alpha)
} else
```

```
if (alternative == "greater") {
LL <- quantile(V, alpha)
UL <- Inf
}
return(list(p.value = p.value, LL = LL, UL = UL))
}



# AT
# Anderson  approximation test for the normal generalized variance
# H_0: |Sig| = Delta_0 and approximated confidence interval
andersonGV.Test <- function(Delta_0, n, p, detS, alpha = 0.05,
alternative="two.sided")
{
Zc <- sqrt((n - 1) /(2 * p)) * (detS / Delta_0 - 1)
if (alternative == "two.sided") p.value <- 2 * (1 - pnorm(abs(Zc)))
else
p.value <- pnorm(Zc, lower.tail = (alternative == "less"))
aux1 <- 2 * p * qnorm(1 - alpha / 2)^2 + 1
aux2 <- 2 * p * qnorm(1 - alpha)^2 + 1
if (alternative == "two.sided") {
z <- qnorm(1 - alpha / 2)
LL <- ((n - 1)^0.5 * detS) / ((n - 1)^0.5 + sqrt(2 * p) * z)
if (n >  aux1)
UL <- ((n - 1)^0.5 * detS) / ((n - 1)^0.5 - sqrt(2 * p) * z) else
UL <- Inf
} else
if (alternative == "less") {
z <- qnorm(1 - alpha)
LL <- 0
if (n > aux2)
UL <- (n - 1)^0.5 * detS / ((n - 1)^0.5 - sqrt(2 * p) * z) else
UL <- Inf
```

```
} else
if (alternative == "greater") {
z <- qnorm(1 - alpha)
LL <- (n - 1)^0.5 * detS / ((n - 1)^0.5 + sqrt(2 * p) * z)
UL <- Inf
}
return(list(Zc = Zc, p.value = p.value, LL = LL, UL = UL))
}



# Sarkar approximation test for the normal generalized variance - ST
# H_0: |Sig| = Delta_0 and approximated confidence interval
sarkarGV.Test <- function(Delta_0, n, p, detS, alpha = 0.05,
alternative="two.sided")
{
j <- 1:p
muy <- sum(digamma((n - j) / 2)) + p * log(2)
sigy <- sqrt(sum(trigamma((n - j)/2)))
Zc <- (p * log(n - 1) + log(detS) - log(Delta_0) - muy) / sigy
if (alternative == "two.sided") p.value <- 2 * (1 - pnorm(abs(Zc)))
else
p.value <- pnorm(Zc, lower.tail = (alternative == "less"))
if (alternative == "two.sided") {
z <- qnorm(1 - alpha / 2)
LL <- exp(p * log(n - 1) + log(detS) - muy - sigy * z)
UL <- exp(p * log(n - 1) + log(detS) - muy + sigy * z)
} else
if (alternative == "less") {
z <- qnorm(1 - alpha)
LL <- 0
UL <- exp(p * log(n - 1) + log(detS) - muy + sigy * z)
} else
if (alternative == "greater") {
```

```
z <- qnorm(1 - alpha)
LL <- exp(p * log(n - 1) + log(detS) - muy - sigy * z)
UL <- Inf
}
return(list(Zc = Zc, p.value = p.value, LL = LL, UL = UL))
}


# DT
# Djauhari approximation test for the normal generalized variance
# H_0: |Sig| = Delta_0 and approximated confidence interval
djauhariGV.Test <- function(Delta_0, n, p, detS, alpha = 0.05,
alternative="two.sided")
{
j <- 1:p
b1 <- exp(sum(log(n - j)) - p * log(n - 1))
rb2 <- sqrt(exp(sum(log(n - j + 2)) -
p * log(n - 1) + log(b1)) - b1^2)
Zc <- (detS / Delta_0 - b1) / rb2
aux1 <-  qnorm(1 - alpha / 2)^2 * rb2^2
aux2 <- qnorm(1 - alpha)^2 * rb2^2
if (alternative == "two.sided") p.value <- 2 * (1 - pnorm(abs(Zc)))
else
p.value <- pnorm(Zc, lower.tail = (alternative == "less"))
if (alternative == "two.sided") {
z <- qnorm(1 - alpha / 2)
LL <- detS / (b1 + rb2 * z)
if (b1^2 > aux1 ) UL <- detS / (b1 - rb2 * z) else
UL <- Inf
} else
if (alternative == "less") {
z <- qnorm(1 - alpha)
LL <- 0
if (b1^2 > aux2) UL <- detS / (b1 - rb2 * z) else
```

```
UL <- Inf
} else
if (alternative == "greater") {
z <- qnorm(1 - alpha)
LL <- detS / (b1 + rb2 * z)
UL <- Inf
}
return(list(Zc = Zc, p.value = p.value, LL = LL, UL = UL))
}



# LRT  for the normal generalized variance - LRT
# H_0: |Sig| = Delta_0 and approximated confidence interval
LRT.Test <- function(Delta_0, n, p, detS, alternative="two.sided")
{
detSigHat <- detS * ((n - 1) / n)^p
nu <- n  - 1
chi2c <- n * (log(Delta_0) - log(detSigHat)) +  n * p *
((detSigHat / Delta_0)^(1/p) - 1)
if (alternative == "two.sided") p.value <- 1 - pchisq(chi2c, 1) else
p.value <- pchisq(chi2c, 1, lower.tail = FALSE)
return(list(chi2c = chi2c, p.value = p.value))
}




BCLRT4.Test <- function(x, Delta_0, alternative="two.sided") - BCLRT
{
n <-    nrow(x)
p <-    ncol(x)
detS <- det(var(x))
detSigHat <- detS  * ((n - 1) / n)^p
```

```
nu <- n  - 1
EdetW  <- detSigHat   * prod(nu:(nu-p+1))
EdetW2 <- detSigHat^2 * prod(nu:(nu-p+1)) * prod((nu+2):(nu-p+3))
EdetW3 <- detSigHat^3 * prod(2^3 *
gamma((nu:(nu-p+1))/2 + 3) / gamma((nu:(nu-p+1))/2))
EdetW4 <- detSigHat^4 * prod(2^4 *
gamma((nu:(nu-p+1))/2 + 4) / gamma((nu:(nu-p+1))/2))
mc2    <- EdetW2 - EdetW^2
mc3    <- 1/(nu^(3*p)) * (EdetW3 - EdetW^3 - 3*EdetW * mc2)
mc4    <- 1/(nu^(4*p)) * (EdetW4 - 5*EdetW^4 - 4* EdetW * EdetW3 +
6 * EdetW^2 * EdetW2)
EdetS  <- EdetW/(nu^p)
ElogDetS <- -p*log(nu) + log(EdetW) -0.5*(EdetW2/EdetW^2 - 1) +
1/3 * mc3/EdetS^3 - 0.25*mc4/EdetS^4
EDetS1p <- 1/nu*(EdetW^(1/p) + 0.5*prod(1/p - 0:1) *
EdetW^(1/p - 2) * mc2) +
1/(6*nu^(1 - 3*p)) * (prod(1/p - 0:2) * EdetW^(1/p - 3) * mc3) +
1/(24*nu^(1 - 4*p)) * (prod(1/p - 0:3) * EdetW^(1/p - 4) * mc4)
Eminus2logLamb <- n*log(Delta_0) - n*p - n*p*log((n-1)/n) -
n*ElogDetS + n*p/Delta_0^(1/p) * (n - 1) / n * EDetS1p
phi     <- nu / Eminus2logLamb
chi2c <- phi * (n * (log(Delta_0) - log(detSigHat)) +  n * p *
((detSigHat / Delta_0)^(1/p) - 1))
if (alternative == "two.sided") p.value <- 1 - pchisq(chi2c, 1) else
p.value <- pchisq(chi2c, 1, lower.tail = FALSE)
return(list(chi2c = chi2c, p.value = p.value))
}




# Monte Carlo Simulation Function to evaluate the
# test performance. Dependence: MASS
library(MASS)
```

```
evalMC <- function(N = 10000, n = 15, p = 2, eta = 0.2, m = 10000)
{
Rej <- matrix(0, 101, 18)
colnames(Rej) <- c("MCET10","AT10","ST10","DT10","LRT10","BCLRT10",
"MCET5","AT5","ST5","DT5","LRT5","BCLRT5",
"MCET1","AT1","ST1","DT1","LRT1","BCLRT1")

rDetSig <- seq(0.01, 5.00, by = 0.05)
rDetSig <- c(rDetSig[rDetSig<eta],eta, rDetSig[rDetSig > eta])
Rej  <- cbind(rDetSig, Rej)
mu <- rep(c(0), times  = p)
alternative <- "two.sided"
st <- 1.0 / N
ct <- 1
alpha <- 0.05
for (D in rDetSig)
{
#print(D)
Sigma <-  D^(1/p) * diag(p)
#print(det(Sigma))
rej10 <- rep(0.0, times = 6)
rej05 <- rep(0.0, times = 6)
rej01 <- rep(0.0, times = 6)
for (i in 1:N)
{
X <- mvrnorm(n, mu, Sigma)
detS <- det(cov(X))
MCET <- exactGV.Test(eta, n, p, detS, alpha, alternative, m)
AT <- andersonGV.Test(eta, n, p, detS, alpha, alternative)
ST <- sarkarGV.Test(eta, n, p, detS, alpha, alternative)
DT <- djauhariGV.Test(eta, n, p, detS, alpha, alternative)
LRT <- LRT.Test(eta, n, p, detS, alternative)
BCLRT <- BCLRT4.Test(X, eta, alternative = "two.sided")
if (MCET$p.value <= 0.10)    rej10[1] <- rej10[1] + st
```

```
if (MCET$p.value <= 0.05)     rej05[1] <- rej05[1] + st
if (MCET$p.value <= 0.01)     rej01[1] <- rej01[1] + st
if (AT$p.value <= 0.10)       rej10[2] <- rej10[2] + st
if (AT$p.value <= 0.05)       rej05[2] <- rej05[2] + st
if (AT$p.value <= 0.01)       rej01[2] <- rej01[2] + st
if (ST$p.value <= 0.10)       rej10[3] <- rej10[3] + st
if (ST$p.value <= 0.05)       rej05[3] <- rej05[3] + st
if (ST$p.value <= 0.01)       rej01[3] <- rej01[3] + st
if (DT$p.value <= 0.10)       rej10[4] <- rej10[4] + st
if (DT$p.value <= 0.05)       rej05[4] <- rej05[4] + st
if (DT$p.value <= 0.01)       rej01[4] <- rej01[4] + st
if (LRT$p.value <= 0.10)      rej10[5] <- rej10[5] + st
if (LRT$p.value <= 0.05)      rej05[5] <- rej05[5] + st
if (LRT$p.value <= 0.01)      rej01[5] <- rej01[5] + st
if (BCLRT$p.value <= 0.10)  rej10[6] <- rej10[6] + st
if (BCLRT$p.value <= 0.05)  rej05[6] <- rej05[6] + st
if (BCLRT$p.value <= 0.01)  rej01[6] <- rej01[6] + st
}
rej <- c(rej10, rej05, rej01)
Rej[ct,2:19] <- rej
ct <- ct + 1
}
return(Rej)
}
```

Graphics

Figura 5.7 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 0.2$, $n = 15$ and $p = 2$, 3, 5 and 10.

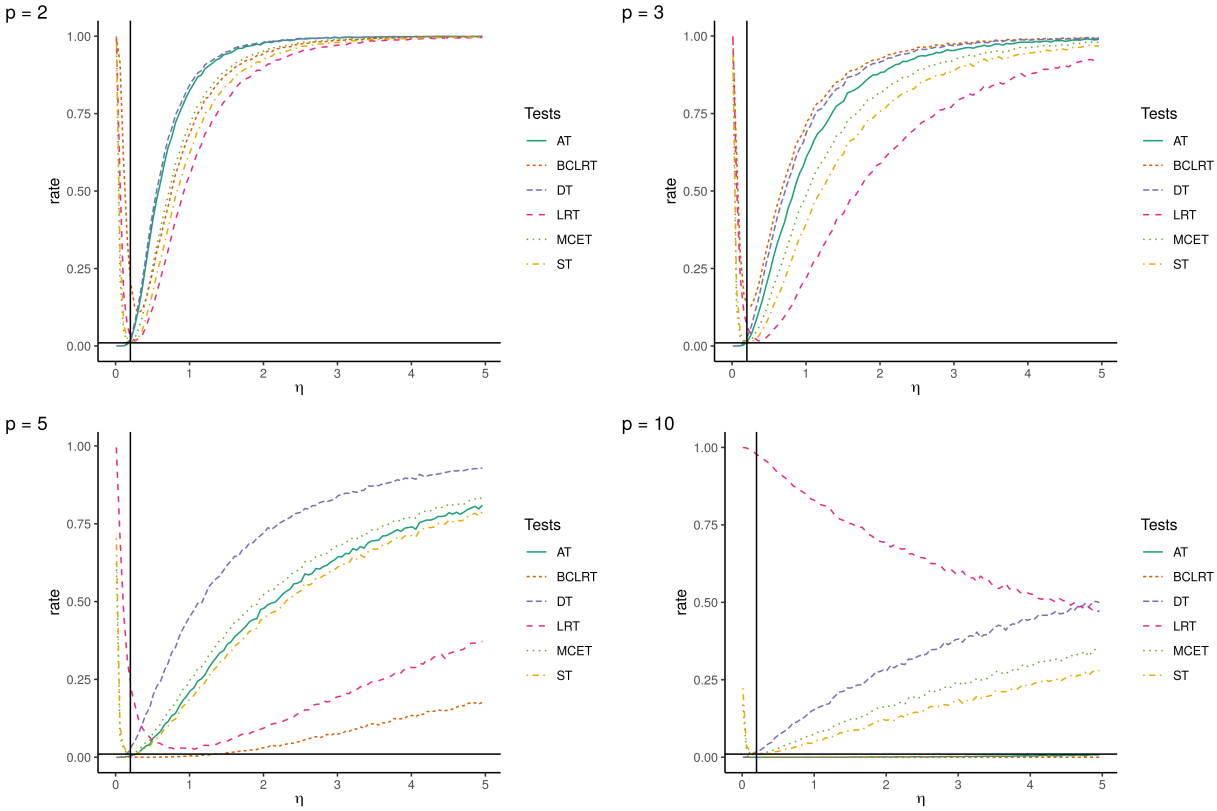


Figura 5.8 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 0.2$, $n = 30$ and $p = 2$, 3, 5 and 10.

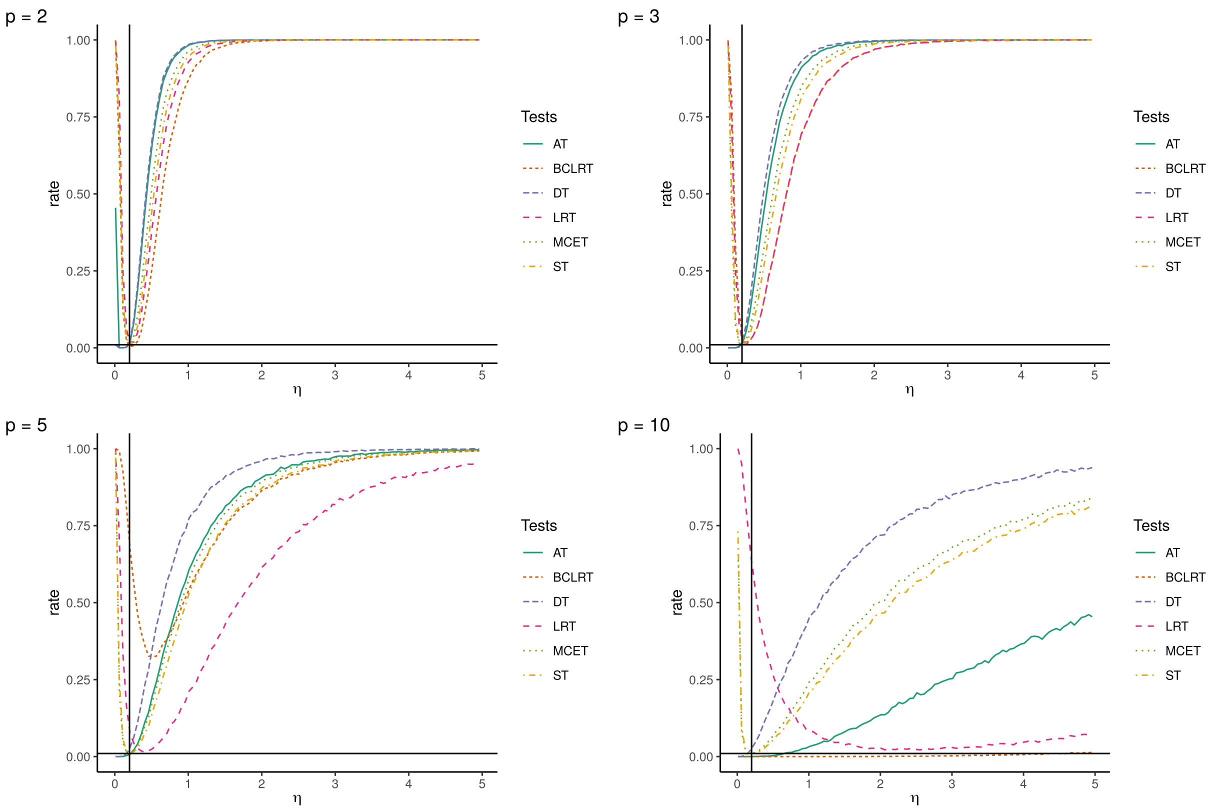


Figura 5.9 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 0.2$, $n = 50$ and $p = 2$, 3, 5 and 10.

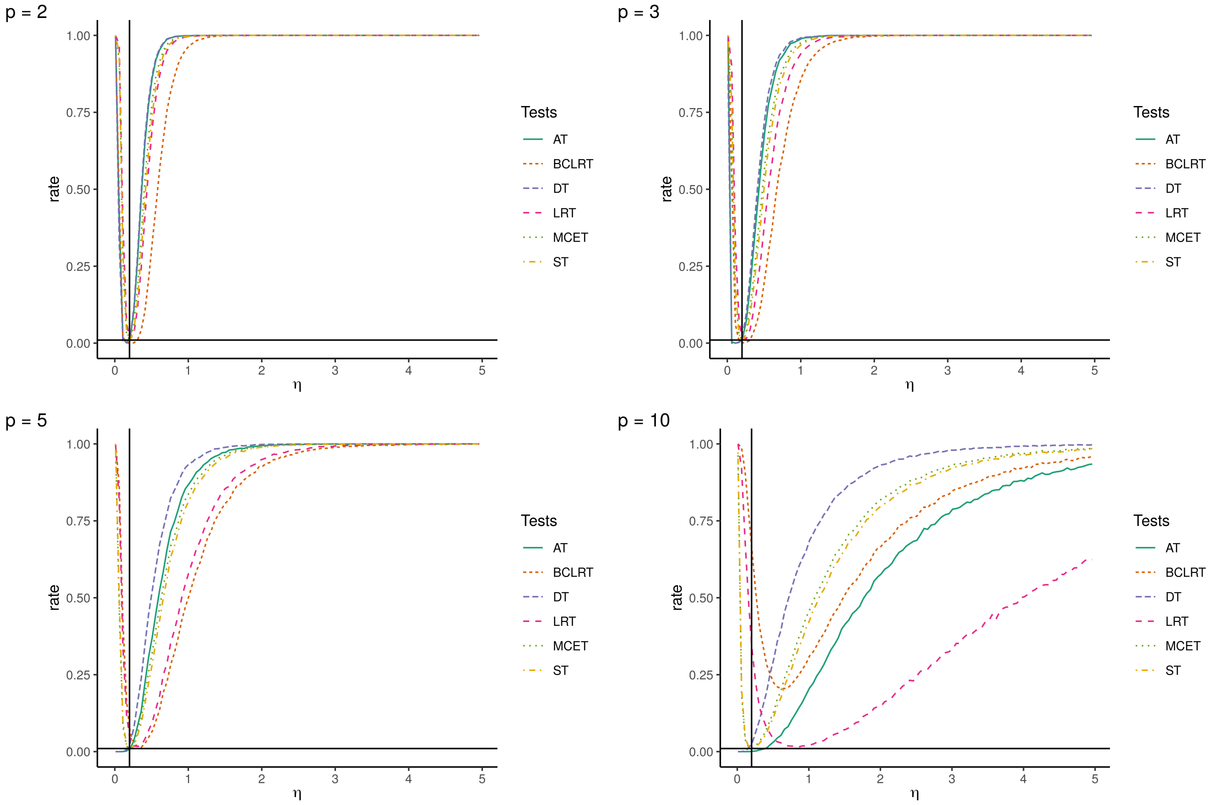


Figura 5.10 – Performance evaluation of hypothesis tests at the 10% significance level for $\eta = 0.2$, $n = 15$ and $p = 2$, 3, 5 and 10.

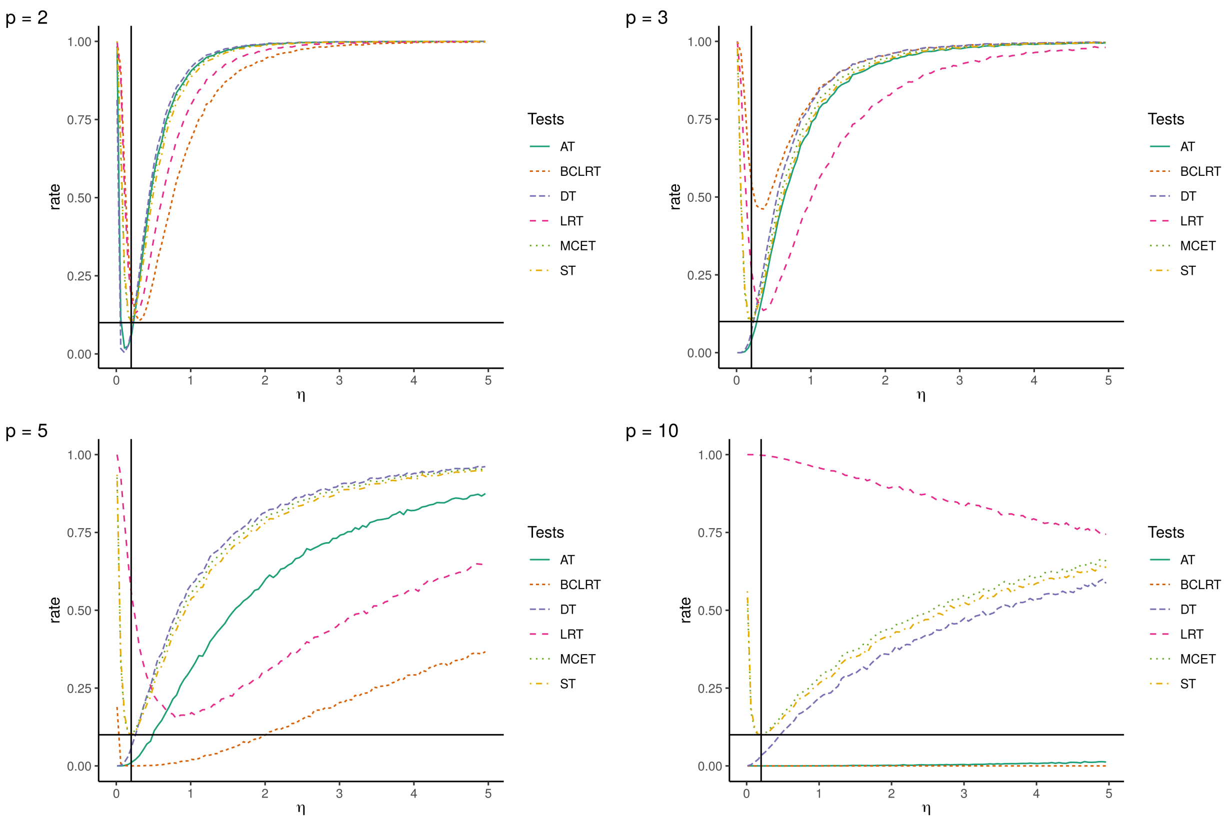


Figura 5.11 – Performance evaluation of hypothesis tests at the 10% significance level for $n = 30$ and $p = 2$, 3, 5 and 10.

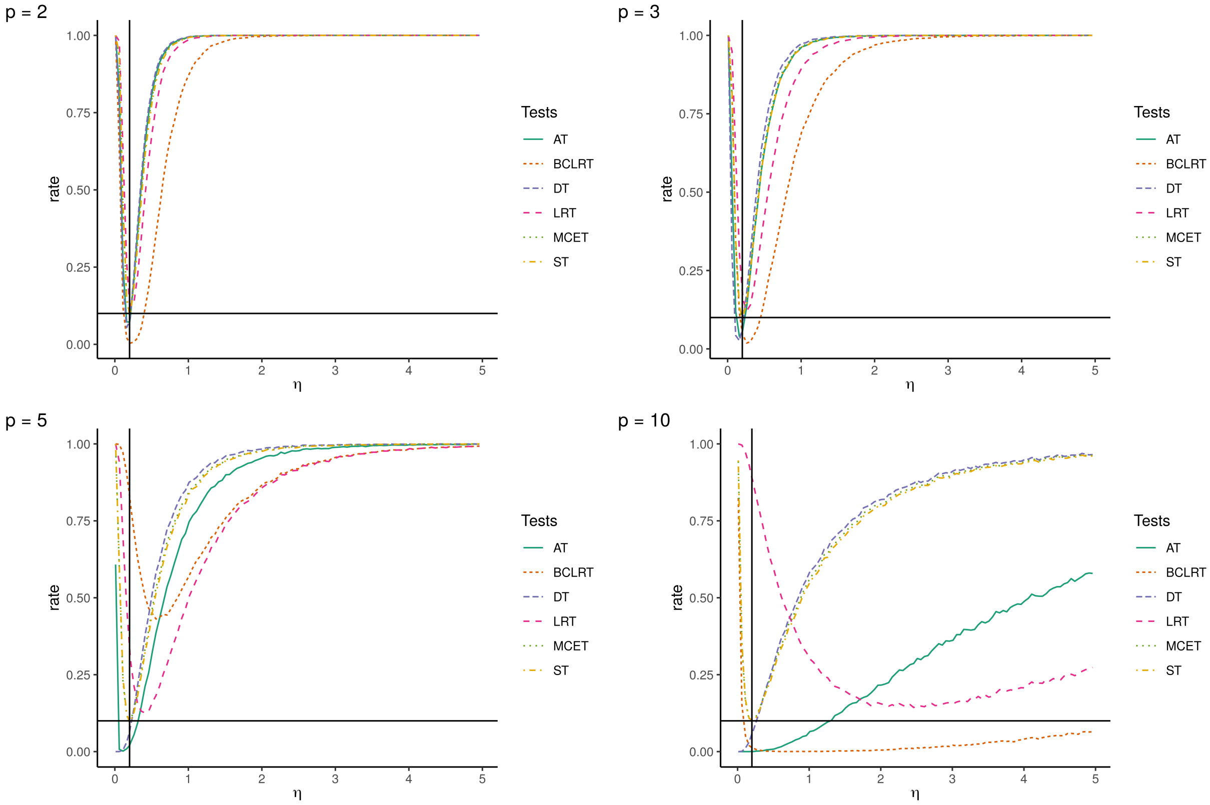


Figura 5.12 – Performance evaluation of hypothesis tests at the 10% significance level for $\eta = 0.2$, $n = 50$ and $p = 2$, 3, 5 and 10.

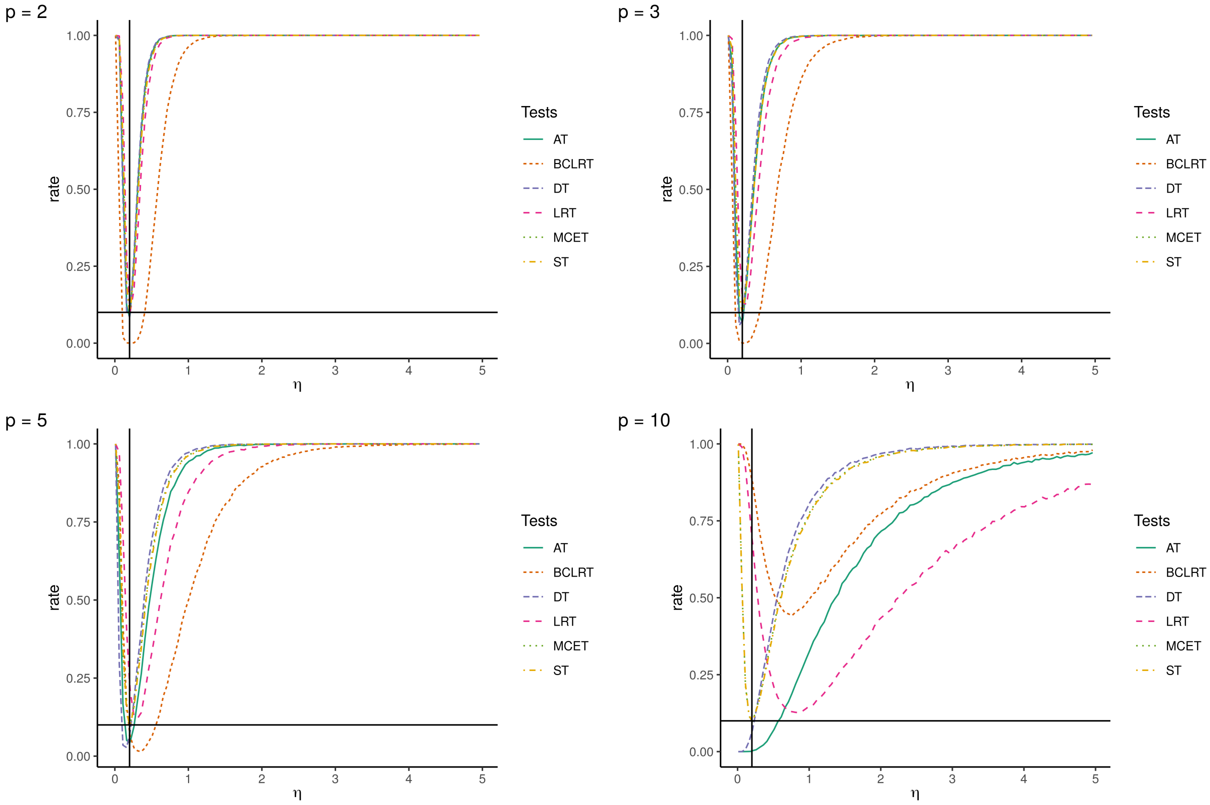


Figura 5.13 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 1$, $n = 15$ and $p = 2, 3, 5$ and 10.

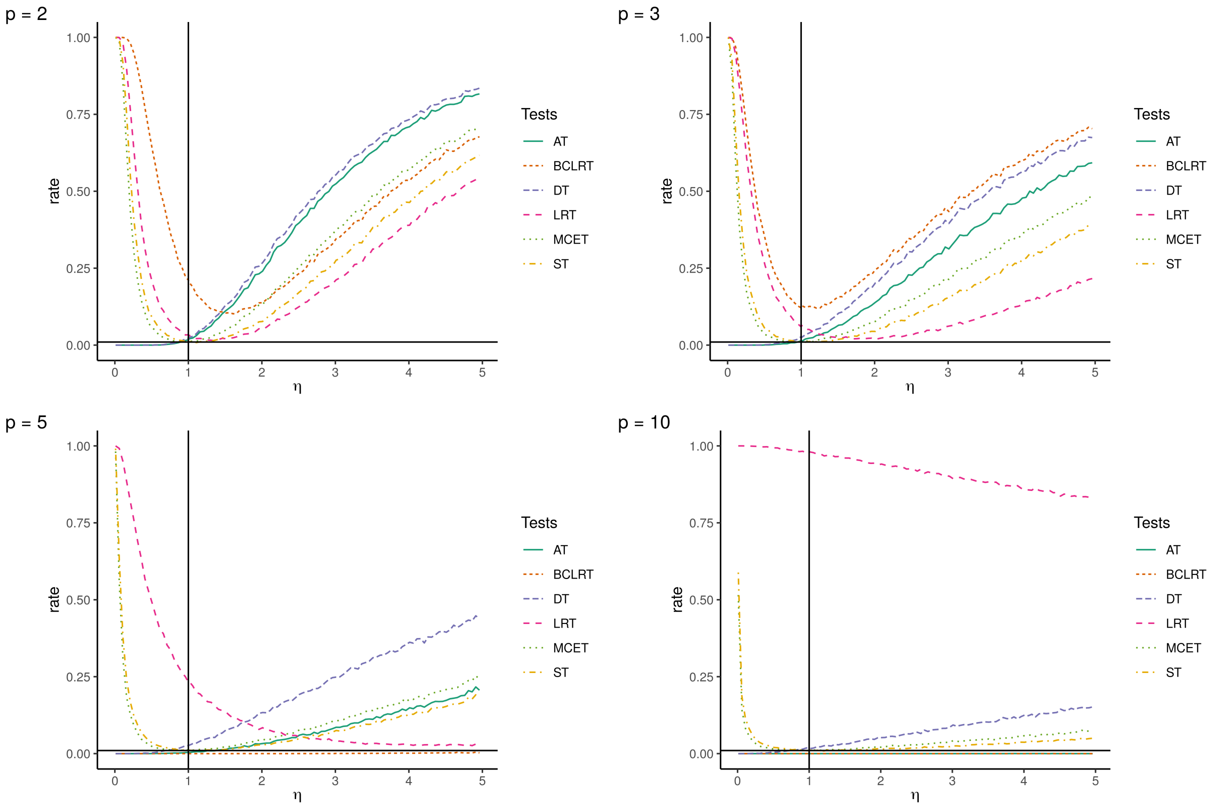


Figura 5.14 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 1$, $n = 30$ and $p = 2, 3, 5$ and 10.

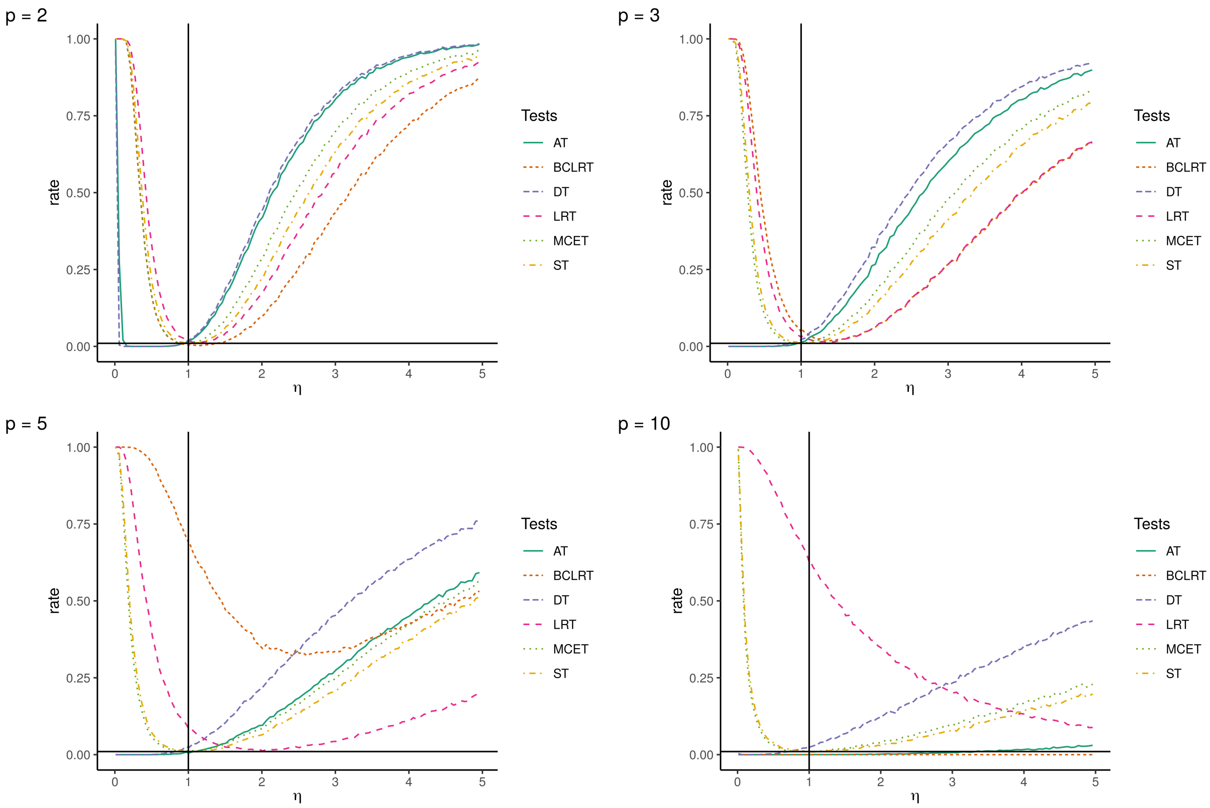


Figura 5.15 – Performance evaluation of hypothesis tests at the 1% significance level for $\eta = 1$, $n = 50$ and $p = 2$, 3, 5 and 10.

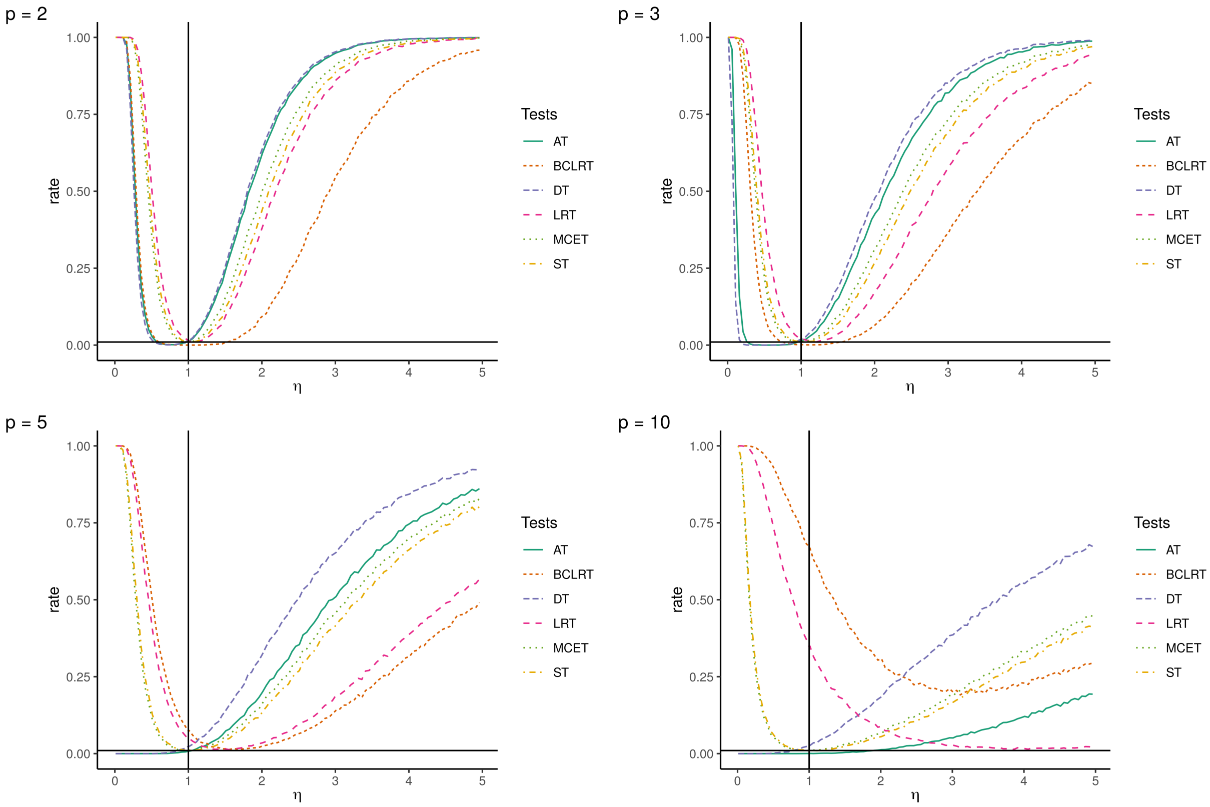


Figura 5.16 – Performance evaluation of hypothesis tests at the 10% significance level for $\eta = 1$, $n = 15$ and $p = 2$, 3, 5 and 10.

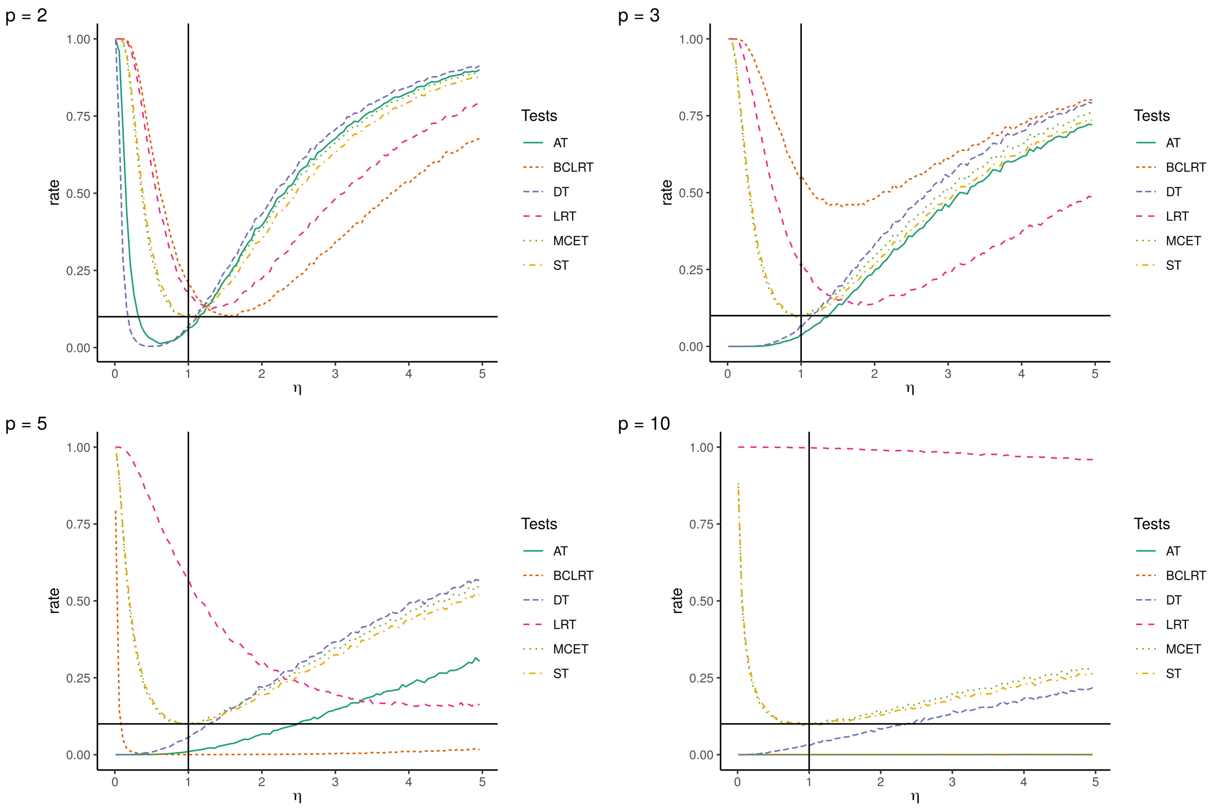


Figura 5.17 – Performance evaluation of hypothesis tests at the 10% significance level for $\eta = 1$, $n = 30$ and $p = 2$, 3, 5 and 10.

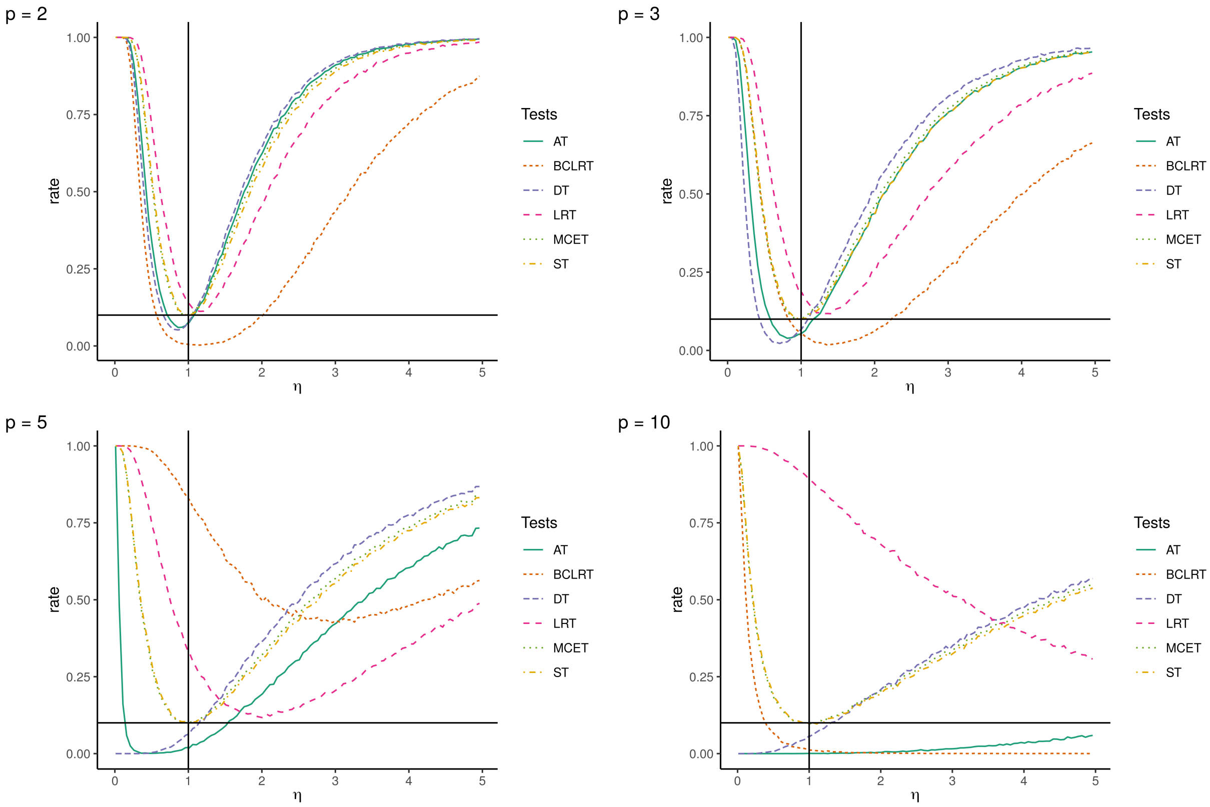
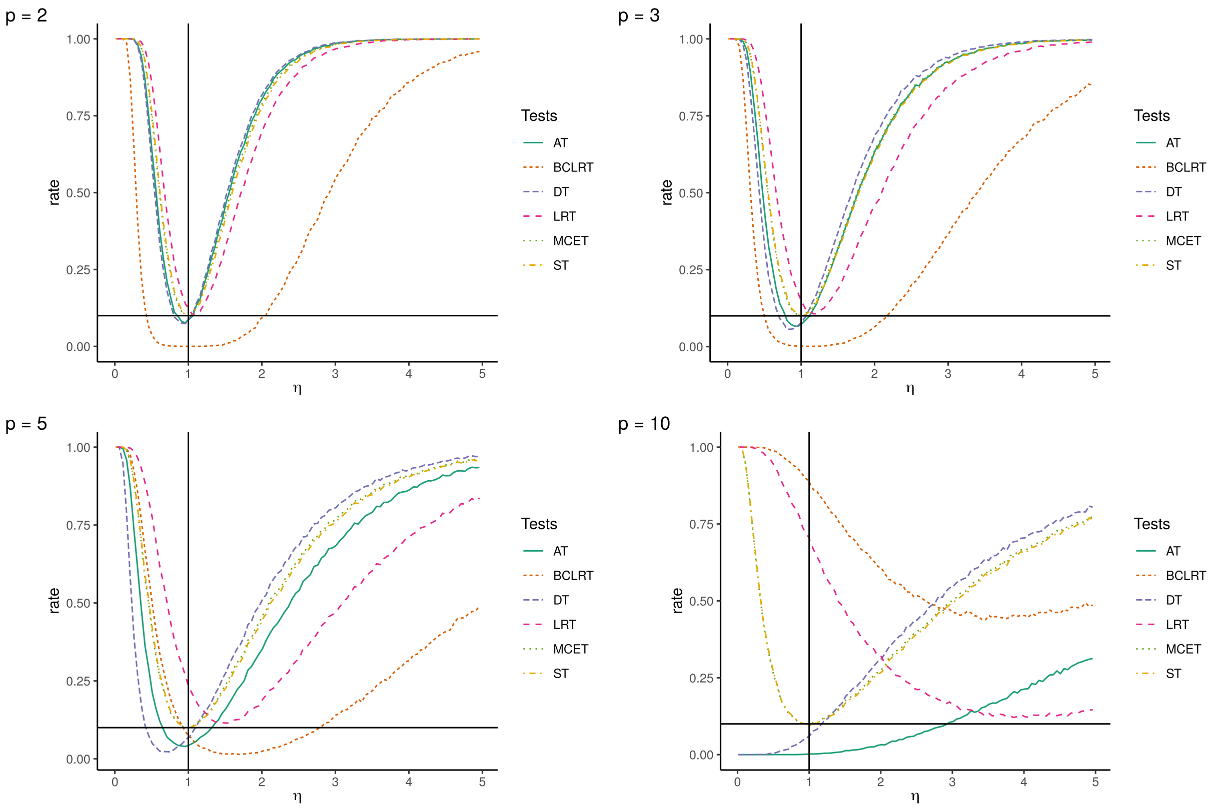


Figura 5.18 – Performance evaluation of hypothesis tests at the 10% significance level for $\eta = 1$, $n = 50$ and $p = 2$, 3, 5 and 10.



REFERÊNCIAS

ANDERSON, T. W. **An introduction to multivariate statistical analysis**. 3. ed. Hoboken: John Wiley and Sons, 2003. 722 p.

ARVANITIS, L. G.; AFONJA, B. Use of the generalized variance and the gradient projection method in multivariate stratified sampling. **Biometrics**, v. 27, n. 1, p. 119–127, 1971.

BARTLETT, M. S. On the theory of statistical regression. **Proceedings of the Royal Society of Edinburgh**, Royal Society of Edinburgh Scotland Foundation, v. 53, p. 260–283, 1934.

BERSIMIS, S.; PSARAKIS, S.; PANARETOS, J. Multivariate statistical process control charts: an overview. **Qual. Reliab. Eng. Int.**, v. 5, n. 23, p. 517–543, 2007.

BHANDARY, M. Test for generalized variance in signal processing. **Statatistical and Probability Letters**, v. 27, n. 2, p. 155–162, 1996.

DJAUHARI, M. A. Improved monitoring of multivariate process variability. **Journal of Quality Technology**, v. 37, n. 1, p. 32–39, 2005.

DJAUHARI, M. A. Asymptotic distribution of sample covariance determinant. **Mathematika**, Johor, v. 25, n. 1, p. 79–85, 2009.

DJAUHARI, M. A.; MASHURI, M.; HERWINDIATI, D. E. Multivariate process variability monitoring. **Communication in Statistics - Theory and Methods**, v. 37, n. 11, p. 1742–1754, 2008.

GUPTA, A. S. Tests for simultaneously determining numbers of clusters and their shapewith multivariate data. **Statatistical and Probability Letters**, v. 1, n. 1, p. 46–50, 1982.

JAFARI, A. A.; KAZEMI, M. R. Inference on the generalized variance under normality. **JIRSS**, Tehran, v. 13, n. 1, p. 57–67, 2014.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4. ed. New Jersey: Prentice Hall, 1998. 816 p.

KOLLO, T.; ROSEN, D. von. **Advanced multivariate statistics with matrices**. 1. ed. Dordrecht: Springer Netherlands, 2005. (Mathematics and Its Applications, 579).

LEE, M. H.; KHOO, M. B. C. Combined synthetic and $|\boldsymbol{S}|$ chart for monitoring process dispersion. **Communications in Statistics - Simulation and Computation**, v. 46, n. 7, p. 5698–5711, 2017.

MITTELHAMMER, R. C. **Mathematical statistics for economics and business**. 2. ed. New York: Springer, 2013. 742 p.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. Singapore: McGraw-Hill, 1974. 480 p.

MUIRHEAD, R. J. **Aspects of multivariate statistical theory**. New York: John Wiley, 1982. 673 p.

NAJARZADEH, D. Testing equality of generalized variances of k multivariate normal populations. **Communications in Statistics - Simulation and Computation**, Taylor and Francis, v. 46, n. 8, p. 6414–6423, 2017.

NAJARZADEH, D. Testing equality of standardized generalized variances of k multivariate normal populations with arbitrary dimensions. **Statistical Methods and Applications**, n. 1, p. 1–31, 2019.

NOOR, A. M.; DJAUHARI, M. A. Monitoring the variability of beltline moulding process using Wilks's statistic. **Malays J. Fundam. Appl. Sci.**, v. 6, n. 2, p. 116–120, 2014.

PUKELSHEIM, F. **Optimal design of experiments**. New York: SIAM, 2006. 454 p.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2019. Disponível em: <https://www.R-project.org/>.

ROYSTON, J. P. Some techniques for assessing multivariate normality based on the Behrens-Fisher W. **Applied Statistics - Journal of the Royal Statistical Society - Series C**, London, v. 32, n. 2, p. 121–133, 1983.

SARKAR, S. K. On improving the shortest length confidence interval for the generalized variance. **Journal of Multivariate Analysis**, v. 31, n. 1, p. 136–147, 1989.

SENGUPTA, A. Generalizations of Barlett's and Hartley's tests of homogeneity using overall variability. **Communication in Statistics - Theory and Methods**, v. 16, n. 4, p. 987–996, 1987.

SENGUPTA, A. Tests for standardized generalized variances of multivariate normal populations of possibly different dimensions. **Journal Multivariate Analysis**, v. 23, n. 2, p. 209–219, 1987.

TALLIS, G. M.; LIGHT, R. The use of fractional moments for estimating the parameters of a mixed exponential distribution. **Technometrics**, v. 10, n. 1, p. 161–175, 1968.

YEH, A. B.; J., L. D. K.; McGRATH, R. N. Multivariate control charts for monitoring covariance matrix: a review. **Qual. Technol. Quant. Manag.**, v. 4, n. 3, p. 415–436, 2006.

YEH, A. B. et al. A multivariate exponentially weighted moving average control chart for monitoring process variability. **Journal of Applied Statistics**, v. 30, n. 5, p. 507–536, 2003.