Revista da Estatística UFOP, Vol III(2), 2014, ISSN 2237-8111
Edição especial: 59ª Reunião Anual da Regional Brasileira da Sociedade Internacional de Biometria - RBRAS

IMPLEMENTAÇÃO DE TESTES PARA HOMOCEDASTICIDADE NO PACOTE *ExpDes*

Mateus Pimenta Siqueira Lima[1], **Marcos Costa de Paula**[1], **Eric Batista Ferreira**[2], **Denismar Alves Nogueira**[2]

Resumo: *O teste F associado à análise de variância possui pressuposições bem conhecidas na literatura. Uma das mais importantes é a homogeneidade de variância dos resíduos. Para experimentos simples conduzidos em Delineamento Inteiramente Casualizado (DIC) e Delineamento em Blocos Casualizados (DBC), podem ser encontrados na literatura testes para a homocedasticidade. Os principais exemplares desses testes foram implementados no pacote* ExpDes *que, a partir de agora, conta com mais essa funcionalidade.*

Palavras-chave: *ANAVA, homocedasticidade, software R.*

Abstract: *The F test, associated with analysis of variance, has well known assumptions in the literature. One of the most important is the residuals homogeneity of variances. For simple experiments conducted in a Completely Randomized Designs (CRD) and Randomized Block Designs (DBC), tests for homoscedasticity can be found in the literature. The main examples of these tests were implemented in* ExpDes *package which, from now on, has one more functionality.*

Keywords: *ANOVA, homocedasticity, R software.*

1 Introdução

A estatística experimental no mundo moderno é uma necessidade real, estando em todas as áreas do conhecimento humano como ferramenta para auxiliar as decisões a serem tomadas, permitindo maior segurança à pesquisa a ser transformada em tecnologia e utilizada pela humanidade ou por parte dela. Desde que se estabeleceram à base científica da estatística, novos conhecimentos envolvendo testes de médias, análises de regressão, análise multivariada, novos métodos para amostragens, estimativas de parâmetros, delineamento, entre outras determinações, estão sendo incorporadas ao estudo da estatística, facilitando a vida profissional de pesquisadores, estudantes, técnicos e outros grupos de profissionais que necessitam dos métodos estatísticos para uma elaboração dos seus trabalhos.

Em experimentação, a comparação entre médias, em geral, é feita por meio da comparação dos efeitos de tratamentos, por meio de testes de comparações múltiplas. Antes disso, porém, geralmente é feito um teste para detectar a existência de diferenças entre os tratamentos, no qual a hipótese nula de igualdade das médias é testada contra a hipótese alternativa de que haja, pelo menos, uma média diferente das demais.

[1]Graduando em Matemática Licenciatura, Universidade Federal de Alfenas.
[2]Instituto de Ciências Exatas, Universidade Federal de Alfenas. E-mail: *eric.ferreira@unifal-mg.edu.* *Agradecimento à FAPEMIG e à Unifal-MG apoio financeiro.*

Em um experimento, a forma com que as parcelas são casualizadas na área experimental, chamada de delineamento experimental, e o esquema de análise empregado determinam o modelo de análise de variância. Neste trabalho, os modelos estatísticos adotados são: Delineamento Inteiramente Casualizado (DIC) e Delineamento de Blocos Casualizados (DBC).

O (DIC) é tido como o mais simples delineamento usados em pesquisa, no qual, de forma geral, a atribuição das diversas repetições dos tratamentos é feita completamente ao acaso, isto é, sem qualquer restrição. Este delineamento é adequado quando a variação do material experimental é relativamente pequena e, em geral não aplica na experimentação de campo onde o material experimental pode variar acentuadamente. Seu uso é mais frequente em experimentos laboratoriais.

O (DBC) é talvez o tipo de delineamento experimental mais importante. O controle local é representado pelo agrupamento dos tratamentos em blocos. Esse delineamento é mais usado principalmente na experimentação de campo, e consiste na subdivisão do material experimental em grupos supostamente uniformes, em que o número de repetições é igual ao número de blocos. Esse tipo de delineamento é também empregado em experimentos zootécnicos e de laboratórios.

A análise de variância (ANAVA) foi o primeiro método para a análise de dados experimentais, desenvolvido por Ronald Fisher a partir da década de 1920 (FISHER, 1973).

A comparação de médias na análise de variância é feita por meio do teste F, que é considerado o teste mais poderoso para este fim no campo paramétrico. Um teste paramétrico é aquele que possui suposições fortes e, entre elas, geralmente sobre a distribuição que seguem os dados (SIEGEL; CASTELLAN, 2006).

A utilização do teste F depende da verificação de quatro pressuposições para que seja válido. Tais pressuposições, denominadas **hipóteses fundamentais da análise de variância**, são: aditividade dos efeitos admitidos no modelo, independência, homocedasticidade e normalidade dos erros. Se, pelo menos uma dessas hipóteses não for satisfeita, a análise de variância não tem validade como técnica de análise estatística e torna-se um simples tratamento matemático dos dados coletados (LIMA; ABREU, 2000).

1.1 Homogeneidade de variâncias

De acordo com Neyman e Pearson (1931) , testes para a igualdade de variâncias são de interesse em muitas situações, tais como análise de variância, controle de qualidade, etc. A abordagem clássica para testar a hipótese de homogeneidade de variâncias começa com o teste da razão de verossimilhança sob a suposição de distribuição normal. Como este teste é muito sensível à ausência desta, muitos testes alternativos foram propostos. Alguns destes testes são variações ou modificações do teste da razão de verossimilhança. Outros são alterações do teste F, que testa a igualdade de médias, para testar a de variâncias.

Estudos recentes mostram que alguns destes testes apresentam falta de robustez e baixo poder (Layard, 1973) . Este problema também foi estudado por Conover et al. (1981), que realizaram um estudo comparando testes paramétricos e não paramétricos, e chegaram a conclusão, utilizando simulação Monte Carlo, que poucos testes são robustos e poderosos.

A importância do teste de homogeneidade de variâncias em muitas áreas da Experimentação é baseada na premissa de que muitos testes de hipóteses sobre médias ou efeitos de tratamentos são realizados pressupondo que as variâncias das populações amostradas sejam iguais.

A violação dessa hipótese pode afetar o desempenho do método e comprometer os resultados de diferentes formas, segundo Johnson e Wichern (1998). Vários modelos estão disponíveis na literatura para este tipo de estudo e é sabido que a heterocedasticidade dos resíduos e um fator que pode afetar a inferência, podendo impactar diretamente nas conclusões (FERREIRA et al., 2006). A presença de heterogeneidade de variâncias pode também ter um sério efeito na validade do teste F, especialmente quando os tamanhos dessas amostras são desbalanceados (O'BRIEN, 1978;KEYES; LEVY, 1997).

Uma alternativa à violação de pressuposições exigidas pela análise de variância é o uso da

Estatística não-paramétrica para a análise dos dados. A Estatística não-paramétrica é relativamente recente. Seus primeiros testes datam do início do século XX, embora tenha sido mais difundida a partir da década de 1940 .

Segundo Gomez e Gomez (1984), a heterogeneidade de variâncias pode ocorrer de duas maneiras. Em uma delas, a variâncias ocorre sem nenhuma relação com a média e na outra existe uma relação entre estas. Em ciências biológicas é comum a presença de correlação positiva entre média e variâncias. Grupos com grandes médias tendem a apresentar grandes variâncias e grupos de pequenas médias apresentam pequenas variâncias.

Nas simulações foram separados em duas etapas: DIC e DBC. No DIC, foram encontrados na literatura vários testes para analisar a homogeneidade de variâncias, quais sejam: Bartlett (1937), Boes e Brownie(1989), Boes e Brownie(1989), Dixon e Massey(1969), Levene(1960), Levene(1960), Brown e Forsythe(1974), Rubin(2005) e Mehrotra(1997), Samiuddin(1976), O'Neill e Mathews(2000), Layard(1973), O'Brien(1978), James(1951), Welch(1951), Hartung et al.(2002).

No trabalho de Nogueira e Pereira (2013), no qual foram analisados esses quinze testes encontrados na literatura, cinco deles apresentaram melhor desempenho, que são: Bartlett(1937), Levene(1960), Samiuddin(1976), O'Neill e Mathews(2000), Layard(1973). Que neste trabalho serão nossos objetos de estudo.

A performance do teste de Bartlett e da modificação do teste de Bartlett proposta por Dixon e Massey (1969) foi avaliada e, no caso de duas amostras, foi comparada com o tete F. O teste de Bartlett foi usado quando mais de duas amostras eram independentes e não foi considerado bom quando amostras eram de tamanhos diferentes. A modificação proposta por Dixon e Massey (1969) foi considerada boa para amostras de tamanhos iguais e diferentes.

Em Silva (1999) foram estudados testes multivariados para comparações de homogeneidade de variância. O teste de Bartlett multivariado foi comparado com uma generalização, proposta pelos autores, do teste de Samiuddin (1976) para o caso multivariado e também com as versões bootstrap destes. Assim, foi avaliado o poder e a taxa de erro tipo I em simulações utilizando a distribuição normal e não-normal. Os resultados foram interessantes, pois mostraram que a versão do teste de Samiuddin, sob normalidade, apresentaram resultados melhores que os de Bartlett e sua versão bootstrap. Melhores resultados foram apresentados pelas versões bootstrap para o controle da taxa de erro tipo I e robustez, que as versões assintóticas.

O teste da razão de verossimilhança modificado por Bartlett (Bartlett, 1937 é o teste mais utilizado para testar a hipótese de homogeneidade. O problema é que este se limita a modelos de um único fator sem blocos ("one-way"). Junto deste estão o teste de Levene (média) (Levene, 1960) e sua modificação (Brown e Forsythe (1974) (mediana), pois suas versões também são utilizadas para modelos "one-way", entretanto conceitualmente pode ser utilizado para qualquer delineamento.

O teste de Bartlett apresenta sua performance comprometida em situações onde não existe normalidade. O teste de Levene por utilizar a estatística do teste F, apresenta um comportamento mais robusto quanto a ausência da normalidade. Muitos autores têm afirmado que o procedimento do teste de Levene é apropriado somente em situações em que o delineamento é balanceado (O'Brien, 1978; Keyes e Levy, 1997). Estes mesmos autores propuseram algumas modificações para contornar o problema, pois a hipótese para a situação de desbalanceamento é diferente da situação de balanceamento. Segundo O'Neill e Mathews (2000) os softwares estatísticos que oferecem as variadas formas do teste de Levene ignoram o fato dos delineamentos serem desbalanceados, ignoram também o fato de que as variáveis analisadas são não-normais e por isso a estatística do teste F, utilizada pelo teste de Levene não segue uma distribuição F. O'Neill e Mathews (2000) propuseram uma alteração no modo de se analizar o teste de Levene para contornar estes problemas. A proposta é a estimação utilizando o método de mínimos quadrados ponderados na análise de variância ao invés de quadrados mínimos ordinários. É mostrado também que o teste de O'Brien (O'Brien, 1979) apresenta resultados muito semelhantes da proposta deles.

O'Neill e Mathews (2002) estendem o trabalho deles de 2000 para qualquer delineamento e para os casos uni e multivariados.

Para exemplificar uma hipótese de homogeneidade de variâncias, sejam t amostras de tratamentos de tamanhos n cada, com i = 1,..., t. Seja y_{ij} a j-ésima observação da i-ésima população, considerada normal com média μ_i e variância σ^2, com mesma distribuição e independentes para i = 1, ..., t e j = 1, ..., n. Como definido deseja-se testar a seguinte hipótese:

$$\begin{cases} H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \ldots = \sigma_t^2 = \sigma; \\ H_1 : \sigma_i^2 \neq \sigma_l^2, \text{ para qualquer i} \neq l. \end{cases} \tag{1}$$

Considerou-se o seguinte modelo:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \tag{2}$$

para i = 1, ..., t e j = 1, ..., n_i. Sendo a média geral, i o efeito do i-ésimo tratamento e ε_{ij} o erro experimental associado a cada observação, assumindo distribuição normal com média 0 e variância σ^2. A partir destas definições serão apresentados os testes que serão utilizados neste trabalho para testar esta hipótese.

Já no DBC, foram utilizados três testes para verificação da homogeneidade de variâncias, que são: Teste de correlação múltipla de Han(1969), teste de Anscombe e Tukey(1963), O'Neill e Mathews(2002), os quais também serão objetos de estudo desse trabalho.

2 Material e métodos

Para a realização deste trabalho foi utilizado programação em código R (R CORE TEAM, 2013).

Foram implementados no pacote ExpeDes (FERREIRA; CAVALCANTI; NOGUEIRA, 2013) apenas os testes que tiveram melhor desempenho nas avaliações de erro tipo I e poder feitas por Nogueira e Pereira (2013).

2.1 Testes Implementados no Delineamento Inteiramente Casualizados

2.1.1 Teste de Bartlett (1937)

Para testar a hipótese de homogeneidade (H_0) inicia-se com a proposta apresentada por Bartlett, que é considerada por muitos como o melhor teste para comparação de variâncias que se baseia na razão de verossimilhanças dada por:

$$\Lambda = \frac{(2\pi)^{-\frac{2}{n}}(\sigma^2)^{-\frac{2}{n}}exp^{-\frac{n}{2}}}{(2\pi)^{-\frac{2}{n}}\prod_{i=1}^{t}(\sigma_i^2)^{-\frac{2}{n_i}}exp^{-\frac{n}{2}}} = \frac{(\sigma^2)^{-\frac{2}{n}}}{\prod_{i=1}^{t}(\sigma_i^2)^{-\frac{2}{n_i}}} = \frac{\prod_{i=1}^{t}(\sigma_i^2)^{\frac{2}{n_i}}}{(\sigma^2)^{\frac{2}{n}}} \tag{3}$$

Sob H_0, $-2Ln(\Lambda)$ tem distribuição assintótica de qui-quadrado com $\nu = t - 1$ graus de liberdade. Sob H_1 têm-se t médias e t variâncias e sob H_0, t médias e 1 variância comum a todas. Assim,

$$\chi_c^2 = nLn(\widehat{\sigma}^2) - \sum_{i=1}^{t}(n_i Ln(\widehat{\sigma}^2)) \tag{4}$$

tem distribuição assintótica de qui-quadrado com $\nu = t-1$ graus de liberdade, sob H_0. Bartlett (1937) propôs uma correção e mudanças para melhorar a aproximação, desta forma a estatística de Bartlett (1937) para o teste da hipótese é:

$$B_1 = \chi^2_{c1} = \sum_{i=1}^{t} \frac{n_i}{S_i^2}\left[\overline{y}_{i.} - \sum_{j=1}^{t} h_j \overline{y}_{j.}\right]^2,$$

sendo $S_p^2 = \sum_{i=1}^{n} \nu_i S_i^2 \Big/ (n-t)$ o estimador não viesado da variância comum. Sobre H_0 a estatística tem distribuição de qui-quadrado com $\nu = t-1$ graus de liberdade.

2.1.2 Teste de Levene (1960)

O teste proposto por Levene (1960) baseia-se em uma transformação nos dados originais e assim realiza uma análise de variância, com um fator, utilizando o teste F para testar a existência de efeitos entre os tratamentos. Na verdade estes efeitos são as variâncias, então se testa assim a presença de homogeneidade de variâncias. A transformação, nada mais é do que a obtenção dos resíduos. Para tanto seja $\overline{y}_i$ a média amostral da i-ésima população de tratamentos e seja $z_{ij} = |y_{ij} - \overline{y}_{i.}|$ uma transformação realizada nos valores originais. A estatística é:

$$L_5 = F_{c5} = \frac{\sum_{i=1}^{t} n_i \left(\overline{z}_{i.} - \overline{z}_{..}\right)^2 \Big/ (t-1)}{\sum_{i=1}^{t}\sum_{j=1}^{n_i} \left(\overline{z}_{i.} - \overline{z}_{..}\right)^2 \Big/ (n-t)}, \tag{5}$$

sendo $\overline{z}_{i.}$ a média de cada tratamento da variável transformada e $\overline{z}_{..}$ a média geral da variável transformada. Sob H_0 a estatística F_{c_5} segue uma distribuição F, com $\nu 1 = t-1$ e $\nu 2 = n-t$ graus de liberdade. A hipótese é rejeitada para $F_{c5} > F_{\nu 1, \nu 2, \alpha - 1}$.

2.1.3 Teste de Samiuddin (1976)

Para o desenvolvimento do teste proposto por Samiuddin (1976) considerou uma distribuição a priori não informativa para μ_i e σ_i^2 sendo i = 1, 2, ..., t. A verossimilhança assumida é proporcional a:

$$\prod_{i=1}^{t}\left(\frac{1}{\sigma_i}\right) exp\left[-\frac{1}{2}\sum_{j=1}^{n_i}(y_{ij} - \mu_i)^2 \Big/ \sigma_i^2\right].$$

A posteriori conjunta pode ser definida como o produto das duas distribuições (Priori e Verossimilhança). A integração da posteriori em relação aos μ_i's nos permitirá obter a distribuição marginal de σ_i^2's. Samiuddin utiliza uma transformação de Wilson-Hilferty (Wilson e Hilferty, 1931) para aproximar uma qui-quadrado pela normal. O autor mostra que $\phi_i = (1/\sigma_i^2)^{\frac{1}{2}}$ segue aproximadamente uma distribuição normal com média m_i e variância a_i^2, sendo portanto:

$$mi = \left((n_i - 1) \Big/ \sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{i.})^2\right)^{\frac{1}{3}} [1 - (2/9(n_i - 1))]$$

$$a_i^2 = 2 \Big/ \left[9\left(\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{i.})^2\right)^{\frac{2}{3}} (n_i - 1)^{\frac{1}{3}}\right].$$

Quando a hipótese de homogeneidade de variância é verdadeira então a estatística do teste Bayesiano é dado por:

$$S_9 = \chi^2_{c12} = \sum_{i=1}^{t} (m_i - m)^2 / a_i^2$$

sendo $m = \left(\sum_{i=1}^{t} m_i/a_i^2\right) \Big/ \left(\sum_{i=1}^{t} 1/a_i^2\right)$. Sob H_0 a estatística do teste Bayesiano tem distribuição assintótica de qui-quadrado com $\nu = t-1$ graus de liberdade. Rejeita-se H_0 para $\chi^2_{c12} > \chi^2_{\nu,\alpha-1}$

2.1.4 Teste de O'Neill e Mathews (2000)

O'Neill e Mathews em 2000 propuseram, para delineamentos "one-way", a realização do teste de Levene (Levene, 1960), onde tem-se como variável $z_{ij} = |y_{ij} - \overline{y}_{i.}|$, utilizando quadrados mínimos ponderados. Os autores calcularam a matriz de covariâncias de z_{ij} e mostraram que o teste F da análise de variância por quadrados mínimos ponderados é dada por:

$$OM_{10} = F_{c13} = \frac{N-t}{t-1} \frac{\sum_{i=1}^{t} w_{0i}(\overline{z}_{i.} - \overline{\overline{z}})^2}{\sum_{i=1}^{t} w_{1i} \sum_{j=1}^{n} (z_{ij} - \overline{z}_{i.})^2},$$

sendo $N = \sum n_i$ e w_{0i} e w_{1i} os pesos dados por:

$$\begin{aligned} w_{0i} &= n_i \left[1 + \frac{2}{\pi}\left(\sqrt{n_i(n_i-2)} + sen^{-1}\frac{1}{n_i-1} - n_i\right)\right]^{-1}; \\ w_{1i} &= \left[1 + \frac{2}{\pi(n_i-1)}\left(\sqrt{n_i(n_i-2)} + sen^{-1}\frac{1}{n_i-1}\right)\right]^{-1}, \end{aligned}$$

sendo $\overline{z}_{i.}$ a i-ésima média de z_{ij}, e $\overline{\overline{z}}$ a média ponderada de $\overline{z}_{i.}$ usando w_{0i}. Para a situação de delineamentos balanceados ($n_i = n$ para todo i) a estatística é simplesmente um múltiplo da análise de variância (ANAVA) por quadrados mínimos ordinários (QMO) baseada na média.

$$F_{QMP} = m \times F_{QMO}$$

sendo m dado por:

$$m = \frac{b-c}{b+(n-1)c}$$

sendo $b = (1 - \frac{2}{\pi})$ e $c = \frac{2}{\pi}(\frac{1}{n-1})(\sqrt{n(n-2)} + sin^{-1}\frac{1}{n-1} - (n-1))$ e o m como sendo o multiplicador que converte a estatística por quadrados mínimos ordinários em uma estatística por quadrados mínimos ponderados.

Este ajuste realizado no teste F por QMO é importante, especialmente para amostras pequenas, pois o valor de $m \to 1$ quando $n \to \inf$.

2.1.5 Teste de Layard

Layard propôs mudanças no procedimento de Jackknife Miller (1968) para testar a hipótese de igualdade de variâncias. O teste é baseado no procedimento de Levene (1960), porém considerando pseudovalores, a estatística do teste é dada por:

$$L_{11} = F_{c14} = \frac{\sum_{i=1}^{t} n_i \left(\overline{U}_{i.} - \overline{U}_{..}\right)^2 \Big/ (t-1)}{\sum_{i=1}^{t}\sum_{j=1}^{n_i} \left(\overline{U}_{i.} - \overline{U}_{..}\right)^2 \Big/ (n-t)}, \tag{6}$$

2.2 Testes Implementados em Delineamento de Blocos Casualizados

2.2.1 Teste de correlação múltipla de Han (1969)

Seja $\overline{y}_{.j}$ a média do j-ésimo bloco e com isso pode-se definir uma nova variável $z_{ij} = y_{ij} - \overline{y}_{.j}$. Segundo Han (1969) só existe $\sigma_1^2 = \sigma_2^2 = ... = \sigma_t^2$ se e somente se $\overline{y}_{.j}$ for não correlacionado com z_{ij} para i = 1, ..., t. Se $\sum_{i=1}^{t} z_{ij} = 0$, z_{1j} é uma combinação linear de $z_{2j} = z_{3j} = ... = z_{tj}$. Se $\overline{y}_{.j}$ for não correlacionado com z_{kj} k = 2, 3, ..., t, então é não correlacionado com z_{1j}. Caso seja verdade, então basta somente considerar a correlação múltipla de $\overline{y}_{.j}$ com $z_{2j} = z_{3j} = ... = z_{tj}$. A hipótese de nulidade de homogeneidade de variância será equivalente a testar que o (R^2) (coeficiente de correlação múltipla entre a média do bloco e qualquer (t-1) observações ajustadas para a média) é zero. Então a estatística do teste é:

$$F_{c16} = \frac{R^{*2}}{1 - R^{*2}} \frac{n - t}{t - 1},$$

sendo (R^{*2}) o coeficiente de correlação múltipla amostral utilizando n blocos. A estatística segue uma distribuição F com $\nu 1 = t - 1$ e $\nu 2 = n - t$ graus de liberdade para a situação de quando os dados seguem uma distribuição normal. A hipótese é rejeitada para $F_{c16} > F_{\nu 1, \nu 2, \alpha - 1}$. Segundo Han (1969) quando n é grande, nR^{*2} tende a uma distribuição de qui-quadrado com $\nu 1 = t - 1$ graus de liberdade.

Sua aplicação tem uma desvantagem pois tem como pressuposição assumido que as variâncias entre blocos são homogêneas.

2.2.2 Teste de Anscombe e Tukey (1963)

Em 1993 Anscombe e Tukey (1963) propuseram um método que ajusta um modelo de regressão entre o erro experimental e os desvios das observações em relação à média, que significa estimar a relação entre as parcelas e o erro associado a elas. A estatística do teste é:

$$F_{c17} = \frac{n\left[\sum_{i=t}^{t} \sum_{j=1}^{n_i} \varepsilon_{ij}^2 (\widehat{y}_{ij} - \overline{y})\right]^2}{[2GLerro(QMerro)^2/GLerro + 2][(t-2)(n-1)SQtrat + (t-1)(n-2)SQbloco]}$$

sendo ε_{ij} o erro experimental associado a parcela i e bloco j; GLerro é o grau de liberdade do erro experimental; SQtrat a soma de quadrados dos tratamentos e SQbloco a soma de quadrado de blocos. A estatística obtida segue uma distribuição F com $\nu 1 = 1$ e $\nu 2 = (t-1)(n-1)$ graus de liberdade, sendo $n = \sum n_i/n$ ou a média harmônica.

2.2.3 Teste de O'Neill e Mathews (2002)

O'Neill e Mathews (2002) estenderam o trabalho deles de 2000 propondo a extensão da modificação do teste de Levene para vários delineamentos. Para a situação de delineamentos em blocos casualizados a matriz de correlação dos resíduos. Maiores detalhes ver O'Neill e Mathews (2002) . Nesta matriz existem três distintas correlações: $\rho_1 = -1/(t-1)$ para dois resíduos no mesmo bloco, $\rho_2 = -1/(n-1)$ para dois resíduos no mesmo tratamento e $\rho_3 = 1/[(t-1)(n-1)]$ para dois resíduos em diferentes blocos e diferentes tratamentos. De maneira resumida, para delineamentos balanceados, a forma mais simples da estatística do teste de Levene (QMP) é:

$$F_{QMP} = \frac{z'V^{-1}(Q_0 - Q_1)z/s}{z'V^{-1}Q_1 z} = \frac{k_0^{-1}}{k_1^{-1}} F_{QMO} = m \times F_{QMO},$$

sendo F_{QMO} a estatística F provida de uma ANAVA por quadrados mínimos ordinários, de valores absolutos dos resíduos padronizados que vieram da análise dos dados originais. O m é o multiplicador que converte a estatística por quadrados mínimos ordinários em uma estatística por quadrados mínimos ponderados.

Os elementos da diagonal de V são todos $w_0 = 1 - 2/\pi$ e existem três distintos termos fora da diagonal dados por $w_i = (2/\pi)[(1-\rho_i^2)^{1/2} + \rho_i sen^{-1}(\rho_i) - 1]$, para i=1, 2 e 3.

A estatística apropriada para F_{QMP} para testar H_0 de tratamentos é $F_{trat(QMP)} = m \times F_{trat(QMO)}$, sendo

$$m = \frac{w_0 - w_1 - w_2 + w_3}{w_0 - w_1 + (b-1)(w_2 - w_3)}$$

$F_{trat(QMO)}$ é o valor do teste F para tratamentos de uma ANAVA de uma delineamento em blocos casualizados utilizando quadrados mínimos ordinários dos valores absolutos do resíduo padronizado originalmente adquirido da análise original dos dados. A estatística $F_{trat(QMO)}$ segue uma distribuição F com $\nu 1 = t-1$ e $\nu 2 = (t-1)(n-1)$ graus de liberdade. A hipótese é rejeitada para $F_{trat(QMO)} > F_{\nu 1, \nu 2, \alpha - 1}$.

Para completar o teste é necessário verificar se também as variâncias entre blocos são homogêneas. Por simetria a hipótese de homogeneidade de variância entre blocos é simplesmente $F_{bloc(QMP)} = m \times F_{bloc(QMO)}$, sendo

$$m = \frac{w_0 - w_2 - w_1 + w_3}{w_0 - w_2 + (t-1)(w_1 - w_3)}.$$

2.3 Objetivo

Implementar no pacote ExpDes os principais testes de verificação de homogeneidade de variâncias para DIC e DBC.

3 Resultados e discussão

As funções que tiveram melhor desempenho nas avaliações de erro tipo I e poder feitas por Nogueira e Pereira (2013) foram implementadas no pacote ExpDes com os seguintes códigos:

Em DIC, foram implementadas as funções de: Bartlett, Leven, Samiuddin, O'Neill e Mathews, e Layard.

A função de Bartlett:

```
bartlett<-function(trat, resp, t, r) {
vari<-matrix(0,t,1)
rp<-0
for(i in 1:t) {
vari[i]<-var(resp[(rp+1):(rp+r[i])])
rp<-sum(r[1:i]) }
S2p<-sum((r-1)*vari)/(length(resp)-t)
A<-(length(resp)-t)*log(S2p)-sum((r-1)*log(vari))
B<-(1/(3*(t-1)))*(sum(1/(r-1))-(1/(length(resp)-t)))
Xc1<-A/(1+B)
pvalor<-1-pchisq(Xc1, t-1)
output<- pvalor
return(output) }
```

A função de Levene:

```
levene<-function (trat, resp, t, r) {
Trat<-factor(trat)
 zdados1<-matrix(0,length(resp),1)
 rp<-0
 for(k in 1:length(resp)) {
 zdados1[k]<-abs(resp[k]-mean(resp[(rp+1):(rp+r[Trat[k]])]))
 if(k<length(resp)){if(trat[k]!=trat[k+1])
 {rp<-sum(r[1:Trat[k]])}} }
  pvalor<-summary(aov(zdados1 ~ Trat))[[1]][1,5]
  output <- pvalor
  return(output) }
```

A função de Samiuddin:

```
Samiuddin<-function(trat, resp, t, r) {
Trat<-factor(trat)
t=length(levels(Trat))
m<-matrix(0,t,1)
somma<-matrix(0,t,1)
a2<-matrix(0,t,1)
rp<-0
for(i in 1:t) {
  dife<-0
  soma<-0
  for(j in 1:r[i]) {
   dife<-(resp[rp+j]-mean(resp[(rp+1):(rp+r[i])]))^2
    soma<-soma+dife }
  somma[i]<-soma
  rp<-sum(r[1:i])
  m[i]<-(((r[i]-1)/somma[i])^(1/3))*(1-(2/(9*(r[i]-1))))
  a2[i]<-2/(9*(somma[i]^(2/3))*(r[i]-1)^(1/3)) }
mm<-sum(m/a2)/sum(1/a2)
pvalor<-pchisq(sum(((m-mm)^2)/a2), (t-1)) }
```

A função de O'Neill e Mathews:

```
ONeill<-function(trat, resp, t, r) {
Trat<-factor(trat)
zdados1.1<-matrix(0,length(resp),1)
rr<-t/sum(1/r)
rp<-0
for(k in 1:length(resp)) {
  zdados1.1[k]<-abs(resp[k]-mean(resp[(rp+1):(rp+r[Trat[k]])]))/ \sqrt(1-(1/rr))
  if(k< length(resp)){if(trat[k]< trat[k+1]){rp<-sum(r[1:Trat[k]])}} }
Fc5.1<-summary(aov(zdados1.1 ~ trat))[[1]][1,4]
b<-(1-2/pi)
c<-(2/pi)*(1/(rr-1))*(sqrt(rr*(rr-2))+asin(1/(rr-1))-(rr-1))
m<-(b-c)/(b+(rr-1)*c)
Fc13<-m*Fc5.1
pvalor<-(1-pf(Fc13, (t-1), summary(aov(zdados1.1 ~ trat))[[1]][2,1]))   }
```

Vale lembrar que a função de Layard é uma modificação da função de Jackknife, já que existe outras modificações na função de Jackknife. Então a função de Layard foi implementada com o seguinte código:

```
Jackknife<-function(trat, resp, t, r) {
vari<-matrix(0,t,1)
varij<-matrix(0,max(r),t)
U<-matrix(0,max(r),t)
rp<-0
for(i in 1:t) {
  vv<-resp[(rp+1):(rp+r[i])]
  vari[i]<-var(vv)
  for(j in 1:r[i]) {
    varij[j,i]<-var(vv[-j])
    U[j,i]<-(r[i]*log(vari[i]))-((r[i]-1)*log(varij[j,i])) }
  rp<-sum(r[1:i]) }
Uij<-as.vector(U)
dadosUij<-cbind(trat,Uij)
dadosUij<-as.data.frame(dadosUij)
pvalor<-summary(aov(dadosUij\$Uij ~ trat))[[1]][1,5] }
```

A função de Han:

```
han<-function(resp, Trat, Bloco) {
Trat<-length(levels(Trat))
  Block<-length(levels(Bloco))
```

```
if (Block\textgreater Trat){
  dife<-matrix(0,Block,Trat)
  ymedia<-matrix(0,Block,1)
  rp<-0
  for(j in 1:Block) {
    for(i in 1:Trat) {
      ymedia[j]<-mean(resp[(rp+1):(rp+Trat)])
      dife[j,i]<-(resp[rp+i]-ymedia[j])  }
    rp<-Trat*j }
  modelohan<-lm(ymedia ~ dife[,2:Trat])
  pvalor.hvar<-1-pf(summary(modelohan)[[10]][1],
  summary(modelohan)[[10]][2],summary(modelohan)[[10]][3])
  output <- pvalor.hvar
  return(output) } }
```

A função de Anscombe e Tukey:

```
anscombetukey<-function(resp, Trat, Bloco, glres, msres, sstrat, ssbloco,
residuals, fitted.values) {
Trat<-length(Trat)
Bloco<-length(Bloco)
div1<-(2*glres*(msres)^2)/glres+2
div2<-((((Trat-2)*(Bloco-1))/Trat*Bloco)*sstrat)+((((Trat-1)*(Bloco-2))/Trat*Bloco)*ssbloco)
Fc17<-((sum((residuals^2)*(fitted.values-mean(resp))))^2)/(div1*div2)
pvalor.hvar<-1-pf(Fc17,1,((Trat-1)*(Block-1)))
output <- pvalor.hvar
return(output) }
```

A função de O'Neill e Mathews:

```
oneillmathews<-function(resp, Trat, Bloco){
ntrat<-length(levels(factor(Trat)))
nbloc<-length(levels(factor(Bloco)))
data<-data.frame(Trat,Bloco,resp)
data<-data[order(Trat),]
zdados<-y<-matrix(0,ntrat,nbloc)

for(i in 1:ntrat) {
 y[i,]<-data\$resp[((i-1)*nbloc + 1) : (i*nbloc)]  }

trat.mean<-apply(y,1,mean)
bloc.mean<-apply(y,2,mean)
g.mean<-mean(resp)
  for(i in 1:ntrat){
    for(j in 1:nbloc){
      zdados[i,j]=abs(y[i,j]- trat.mean[i] - bloc.mean[j] + g.mean) } }

zdados<-as.vector(zdados)
dadosz<-data.frame('z'=zdados,'blocagem'=rep(1:nbloc,
each=ntrat),'tratamento'=rep(seq(1:ntrat),nbloc))
attach(dadosz)
Fc6<-summary(aov(z ~ factor(tratamento) + factor(blocagem)))[[1]][1,4]
rho<-c(-1/(ntrat-1), -1/(nbloc-1), 1/((nbloc-1)*(ntrat-1)))
w0<-1-(2/pi)
w1<-(2/pi)*(sqrt(1-rho[1]^2)+rho[1]*asin(rho[1])-1)
w2<-(2/pi)*(sqrt(1-rho[2]^2)+rho[2]*asin(rho[2])-1)
w3<-(2/pi)*(sqrt(1-rho[3]^2)+rho[3]*asin(rho[3])-1)
m<-(w0-w1-w2+w3)/(w0-w1+(nbloc-1)*(w2-w3))
Fc18<-m*Fc6
pvalor.hvar<-(1-pf(Fc18, (ntrat-1), (nbloc-1)*(ntrat-1)))
output <- pvalor.hvar
return(output) }
```

3.1 Exemplo de DIC

De acordo com Nogueira e Pereira (2013), a função de Bartlett foi a que apresentou melhor desempenho quanto ao erro tipo I e poder, e por esse motivo ela foi implementada como *default* na função DIC. Mas fica a cargo do usuário poder escolher qual teste usar, basta trocar o parâmetro `hvar` e escolher um dos outros testes implementados no pacote `ExpDes`.

Para exemplificar o teste de homogeneidade de variância dentro da função DIC, foi utilizado um experimento que teve como objetivo avaliar o consumo da farinha da polpa de yacon sobre o índice glicêmico das dietas experimentais (PEREIRA et al., 2013).

```
data(ex1)
attach(ex1)
dic(trat, ig, quali = FALSE, hvar='Bartlett', sigF = 0.05)
```

```
------------------------------------------------------------------------
Quadro da analise de variancia
------------------------------------------------------------------------
           GL     SQ     QM     Fc     Pr>Fc
Tratamento  3 214.88 71.626 6.5212 0.0029622
Residuo    20 219.67 10.984
Total      23 434.55
------------------------------------------------------------------------
CV = 3.41 %
------------------------------------------------------------------------
...
------------------------------------------------------------------------
Teste de homogeneidade de variância ( Bartlett )
p-valor:  0.1863216
De acordo com o teste de Bartlett a 5% de significancia, as variâncias
podem ser consideradas homogêneas.
------------------------------------------------------------------------
...
```

No exemplo apresentado podemos verificar que de acordo com o teste de Bartlett a 5 % de significância, as variâncias podem ser consideradas homogêneas, atendendo ao exigido pelo teste F. Caso fosse rejeitado, a análise não seria abortada automaticamente pelo programa, ficaria a cargo do usuário levar os resultados em consideração ou não.

3.2 Exemplo de DBC

De acordo com Nogueira e Pereira (2013), a função de O'Neill e Mathews foi a que apresentou melhor desempenho de erro tipo I e poder, e por esse motivo ela foi implementada como *default* na função DBC. Mas fica a cargo do usuário poder escolher qual teste usar, basta trocar o parâmetro `hvar` e escolher um dos outros testes implementados no pacote `ExpDes`.

Para exemplificar o teste de homogeneidade de variância dentro da função DBC, foi utilizado um experimento de uma avaliacao sensorial, em que provadores (blocos) avaliaram a aparencia das barras alimenticias (PAIVA, 2008).

```
data(ex1)
attach(ex1)
dbc(trat, provador, aparencia, quali = TRUE, mcomp="lsd",
hvar='ONeill&Mathews', sigT = 0.05, sigF = 0.05)
```

```
------------------------------------------------------------------------
Quadro da analise de variancia
------------------------------------------------------------------------
            GL      SQ      QM     Fc      Pr>Fc
Tratamento   4  720.38 180.096 71.156 0.00000000
```

```
Bloco       69  324.00   4.696  1.855 0.00025852
Residuo    276  698.55   2.531
Total      349 1742.94
------------------------------------------------------------------------
CV = 29.14 %
------------------------------------------------------------------------
...
------------------------------------------------------------------------
Teste de homogeneidade de variância ( ONeill&Mathews )
p-valor:  0.2059067
De acordo com o teste de ONeill&Mathews a 5% de significancia,
as variâncias podem ser consideradas homogêneas.
------------------------------------------------------------------------
...
```

No exemplo apresentado podemos verificar que de acordo com o teste de O'Neill e Mathews com 5% de significância, as variâncias podem ser consideradas homogêneas, atendendo ao exigido pelo teste F. Caso fosse rejeitado, a analise não seria abortada automaticamente pelo programa, ficaria a cargo do usuário levar os resultados em consideração ou não.

4 Conclusões

Toda programação e implementação dos testes ocorreram com sucesso, e a partir de agora o pacote ExpDes tem uma função que verifica a homogeneidade de variâncias para DIC e DBC simples. E estudos futuros irão buscar implementar e proporcionar aos usuários os testes de homogeneidade de variâncias para experimentos fatoriais.

Referências

[1] ANSCOMBE, F. J.; TUKEY, J. W. **The examination and analysis of residuals.** Technometrics, 5:141-160, 1963.

[2] BARTLETT, M. S. **Properties of sufficiency and statistical tests. Proceedings of the Royal Statistical Society - Serie A**, 60:268-282, 1937.

[3] BOOS, D. D.; BROWNIE, C. **Bootstrap methods for testing homogeneity of variances. Technometrics**, 31(1):69-82, 1989.

[4] BROWN, M. B.; FORSYTHE, A. B. Robust tests for equality of variances. **Journal of the American Statistical Association**, 69(346):364-367, Jun. 1974.

[5] CONOVER, W. J.; JOHNSON, M. E.; JOHNSON, M. M. A comparative study of tests for heterogeneity of variances, with applications to the outer continental shelf bidding data. **Technometrics**, 23(4):351-361, 1981.

[6] DIXON, W. J.; MASSEY, F. J. **Introduction to statistical analysis. McGraw-Hill Book, New York, 3th edition**, 1969. 308-309 p.

[7] FERREIRA, D. F.; DEMÉTRIO, C. G. B.; MANLY, B. F. J.; MACHADO, A. DE A.; VENCOVSKY, R. **Statistical models in agriculture: biometrical methods for evaluating phenotypic stability in plant breeding.** Cerne, 12(4): 373 - 388, 2006.

[8] FERREIRA, E. B., CAVALCANTI, P. P., NOGUEIRA, D. A. . **ExpDes.pt: Experimental Designs pacakge (Portuguese).** 2013.

[9] FISHER, R. A. Statistical methods for research workers. 14 ed. New York: Hafner. 1973. 354p.

[10] GOMEZ, K. A.; GOMEZ, A. A. **Statistical procedures for agricultural research**. John Wiley, New York, 2th edition, 1984. 680 p.

[11] HAN, C. P. **Testing the homogeneity of variances in a two-way classification.** Biometrics, 25:153-158, Mar. 1969.

[12] HARTUNG, J.; ARGAÇ, D.; MAKAMBI, K. H. **Tsmall sample properties of tests on homogeneity in one-way anova and meta-analysis. Statist**. Papers, 43:197-235, 2002.

[13] JAMES, G. S. **The comparison of several groups of observations when the ratios of population variances are unknown. Biometrika**, 38:324-329, 1951.

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

[15] KEYES, T. K.; LEVY, M. S. **Analysis of levene test under design imbalance.** Journal of Educational and Behavioral Statistics, 22:227-236, 1997.

[16] LAYARD, M. N. J. **Robust large-sample tests for homogeneity of variances.** Journal of the American Statistical Association, v.68, n.341, p.195-198, 1973.

[17] LEVENE, H. Robust tests for equality of variances. in: Olkin, i.; ghurye, s.g.; hoeffding, w.; madow, w.g.; mann, h.b. (eds.). **Contribution to Probability and Statistics. Stanford**, CA: Stanford University Press, pages 278-292, 1960.

[18] LIMA, P. C.; ABREU, A. R. de. **Estatástica Experimental: ensaios balanceados.** Lavras: UFLA, 2000. 99 p.

[19] MEHROTRA, D. V. **Improving the brown-forsythe solution to the generalized behrens-fisher problem. Comm. Statist. Simulation Comput**., 26:1139-1145, 1997.

[20] MILLER, R. G., Jr. **Jackknifing variances. Annals of Mathematical Statistics**, 39(2):567-582, Apr. 1968.

[21] NEYMAN, J.; PEARSON, E. S. On the problem of k samples. Bull. Plish Acad. Sci. A, 1931. 460-481 p.

[22] NOGUEIRA, D, P.; PEREIRA, G, M. **Desempenho de testes para homogeneidade de variâncias em delineamentos inteiramente casualizados**. Sigmae, Alfenas, v.2, n.1, p. 7-22. 2013.

[23] O´BRIEN, R. G. **A robust technique for testing heterogeneity of variance effects in factorial design**. Psychometrika, 43(3):327 - 342, 1978.

[24] O´NEILL, M. E.; MATHEWS, K. L. **A weighted least squares approach to levene test of homogeneity of variance**. Australian e New Zealand Journal Statistical, 42(1):81-100, 2000.

[25] O´NEILL, M. E.; MATHEWS, K. L. **Levene tests of homogeneity of variance for general block and treatment designs**. Biometrics, 58:216-224, Mar. 2002.

[26] PAIVA, A. P. de. Estudos tecnologicos, quimico, fisico-quimico e sensorial de barras alimenticias elaboradas com subprodutos e residuos agoindustriais. 2008. 131p. **Dissertacao (Mestrado em Ciencias dos Alimentos)** - Universidade Federal de Lavras, UFLA, Lavras, 2008.

[27] PEREIRA, J. A. R. ; BARCELOS, M. F. P. ; PEREIRA, M. C. A. ; Ferreira, E. B. Studies of chemical and enzimatic characteristics of Yacon (Smallanthus sonchifolius) and its flour. **Ciência e Tecnologia de Alimentos**, v. 33, p. 75-83, 2013.

[28] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria (2013).

[29] RIBEIRO, R. Proposta e comparação do desempenho de testes para homogeneidade de variância de modelos de classi cação one-way e two-way. Iniciação Científica. (Iniciação Científica) - Universidade Federal de Alfenas. 2012.

[30] RUBIN A. S. **Experimentção em genética**. Editora UFLA, Lavras - MG, n.2, p.303, 2005.

[31] SAMIUDDIN, M. Bayesian test of homogeneity of variance. **Journal of the American Statistical Association**, 71(354):515-517, Jun. 1976.

[32] SIEGEL, S.; CASTELLAN, N. J. Jr.**Estatística não-paramétrica para ciências do comportamento (2° Edição)**. Trad. S. I. C. Carmona. Porto Alegre: Artmed, 2006. 448p.

[33] WELCH, B. L. **On the comparison of several mean values: an alternative** approach. Biometrika, 38:330-336, 1951.

[34] WILSON, E. B.; HILFERTY, M. M. **The distribution of chi-square. Proceedingdof the National Academy of Science**, 17:684-688, 1931.