

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
ESPECIALIZAÇÃO EM ANÁLISE INSTRUMENTAL

DANIEL EIJI KASHIWAGI

**REVISÃO BIBLIOGRÁFICA E COMPARAÇÃO DE TESTES  
ESTATÍSTICOS APLICADOS A VALIDAÇÃO DE METODOLOGIAS  
ANALÍTICAS NA INDÚSTRIA FARMACÊUTICA**

TRABALHO DE CONCLUSÃO DE CURSO

TOLEDO

2019

DANIEL EIJI KASHIWAGI

**REVISÃO BIBLIOGRÁFICA E COMPARAÇÃO DE TESTES  
ESTATÍSTICOS APLICADOS A VALIDAÇÃO DE METODOLOGIAS  
ANALÍTICAS NA INDÚSTRIA FARMACÊUTICA**

Trabalho de conclusão de curso ao Programa de Pós-Graduação em Especialização em Análise Instrumental da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título “Especialista em Análise Instrumental”

Orientador: Prof. Dr. Lincoln Figueira Marins Coutinho

TOLEDO

2019

## TERMO DE APROVAÇÃO

### REVISÃO BIBLIOGRÁFICA E COMPARAÇÃO DE TESTES ESTATÍSTICOS APLICADOS A VALIDAÇÃO DE METODOLOGIAS ANALÍTICAS NA INDÚSTRIA FARMACÊUTICA

Por

**DANIEL EIJI KASHIWAGI**

Esse trabalho de conclusão de curso foi apresentado no dia dois de agosto de dois mil e dezanove, como requisito parcial para a obtenção do título de Especialista em Análise Orgânica, no Curso de Especialização em Análise Orgânica, da Universidade Tecnológica Federal do Paraná, Câmpus Toledo. *O candidato foi arguido* pela Banca Examinadora composta pelos professores abaixo relacionados.

Após deliberação, a banca Examinadora considerou o ***trabalho Aprovado***.

---

Prof. Dr. Lincoln Figueira Marins Coutinho (Orientador)

---

Prof. Dra. Magali Silva Amorim (Membro – UTFPR)

---

Prof. Dra. Rosangela A. B. Assumpção (Membro – UTFPR)

***\*A versão assinada encontra-se arquivada na Coordenação da Especialização***

## SUMÁRIO

RESUMO .....	5
ABSTRACT .....	6
1. INTRODUÇÃO .....	7
2. OBJETIVOS .....	8
2.1. OBJETIVO GERAL .....	8
2.2. OBJETIVOS ESPECÍFICOS .....	8
2.3. JUSTIFICATIVA .....	8
3. CONTEXTUALIZAÇÃO .....	9
4. METODOLOGIA .....	12
5. REVISÃO BIBLIOGRÁFICA .....	13
5.1. HOMOCEDASTICIDADE/HETEROCEDASTICIDADE .....	13
5.2. TESTES DE COMPARAÇÃO DE VARIÂNCIAS .....	14
6. CONSIDERAÇÕES FINAIS .....	18
REFERÊNCIAS .....	19

## RESUMO

O objetivo deste trabalho foi avaliar teoricamente os diferentes testes estatísticos que podem ser aplicados para a avaliação da homocedasticidade de dados que são utilizados na validação de métodos analíticos empregados na indústria farmacêutica. Para tal, foram levantados alguns dos testes mais relatados em literatura científica seguido por uma verificação teórica da adequabilidade de cada teste em diferentes situações. Analisando-se tal revisão bibliográfica dos testes, é possível concluir que cada teste possui limitações e hipóteses a serem cumpridas e deve-se, então, realizar uma avaliação crítica para aplicação do teste mais adequado ao se validar uma metodologia analítica.

Palavras-chave: estatística, homocedasticidade, heterocedasticidade, validação analítica, linearidade.

## **ABSTRACT**

This objective of this work was to theoretically evaluate different statistical tests which could be applied for homoscedasticity evaluation of data used in analytical methods validation in the pharmaceutical industry. For this, some of the most reported tests related in scientific literature were reviewed through a theoretical verification of the suitability of each one in different situations. By analyzing this bibliographic review, it was possible to conclude that each test has limitations and hypotheses to be fulfilled and, therefore, a critical evaluation should be performed before choosing and applying the most adequate test when validating an analytical methodology.

Key-words: statistic, homoscedasticity, heteroscedasticity, analytical validation, linearity.

## 1. INTRODUÇÃO

Com a recente publicação da RDC 166/2017, que dispõe sobre a validação de métodos analíticos e dá outras providências, e posterior publicação do Guia para tratamento estatístico da validação analítica (Guia nº 10/2017), cujo objetivo é recomendar modelos e procedimentos de análise estatística de alguns parâmetros da validação para cumprimento da resolução citada, as indústrias farmacêuticas foram forçadas a buscar conhecimentos para a aplicação de tais testes. Contudo, apesar da Agência Nacional de Vigilância Sanitária (ANVISA) ter exposto em seu Guia os principais testes a serem utilizados, vale ressaltar que o Guia não deixa de ser uma recomendação e é, exclusivamente, o entendimento da própria agência para atendimento a RDC 166/2017.

É sabido que há uma enorme gama de testes estatísticos que podem ser aplicados a diferentes situações, sendo que cada teste trabalha com hipóteses que devem ser cumpridas para garantir a validade do resultado do teste estatístico.

Diante deste cenário, as principais dificuldades encontradas pelas indústrias são:

- i) otimização da amostragem experimental levando em consideração a viabilidade para execução dos testes;
- ii) diferentes métodos analíticos apresentam diferentes formas de quantificação, o que leva a necessidade de aplicação de diferentes testes estatísticos para comprovação da adequabilidade dos mesmos;
- iii) falta de conhecimento técnico em estatística pelos analistas/pesquisadores que atuam no desenvolvimento e validação de métodos analíticos, o que leva a interpretações ambíguas e errôneas a respeito dos testes a serem aplicados;
- iv) a publicação de um Guia leva as indústrias a adotar uma postura conservadora de seguir exatamente o que consta neste (por falta de conhecimento técnico e receio de contradizer o entendimento da ANVISA).

## **2. OBJETIVOS**

### **2.1. OBJETIVO GERAL**

Realizar um levantamento bibliográfico acerca dos principais testes estatísticos que podem ser utilizados na validação de métodos analíticos, trazendo vantagens, desvantagens e conceitos envolvidos para utilização destes, considerando o cenário da indústria farmacêutica.

### **2.2. OBJETIVOS ESPECÍFICOS**

- Revisão dos testes aplicados no Guia 10/2017;
- Revisão de testes alternativos que constam em literatura, farmacopeias, etc.;
- Avaliação teórica comparativa dos testes;
- Verificação da adequação de cada teste para diferentes situações.

### **2.3. JUSTIFICATIVA**

Atender uma deficiência da indústria farmacêutica (profissionais com formação ou capacitados em estatística) visando o cumprimento de uma demanda da agência reguladora (ANVISA).



### 3. CONTEXTUALIZAÇÃO

No campo da pesquisa farmacêutica, a investigação analítica de insumos farmacêuticos ativos, intermediários, medicamentos acabados, formulações, impurezas, produtos de degradação e amostras biológicas contendo drogas e seus metabólitos é de extrema importância. Do início da análise farmacêutica oficial, métodos analíticos de doseamento foram incluídos em monografias compendiais com intuito de caracterizar a qualidade de insumos farmacêuticos ativos através da limitação do conteúdo de ingrediente ativo (SIDDIQUI, ALOTHMAN, RAHMAN, 2017).

Dos estágios de desenvolvimento de drogas até a comercialização e pós-comercialização, técnicas analíticas têm papel fundamental, seja no entendimento da estabilidade física e química da droga, impacto na seleção da forma farmacêutica, avaliação das moléculas, quantificação de impurezas, identificação destas que estão acima do limite estabelecido para determinação do perfil toxicológico e avaliação do conteúdo do composto ativo em produtos comercializados. A análise do composto principal ou metabólico, que pode ser quantitativa ou qualitativa, é extensivamente aplicada em estudos farmacocinéticos (SIDDIQUI, ALOTHMAN, RAHMAN, 2017).

O objetivo da validação de um procedimento analítico é demonstrar que o mesmo é adequado a seu propósito pretendido (ICH, 1994). Ainda, conforme a RCC 166/2017, a validação analítica “é a avaliação sistemática de um método por meio de ensaios experimentais de modo a confirmar e fornecer evidências objetivas de que os requisitos específicos para seu uso pretendido são atendidos”.

A Figura 1 traz o quadro retirado da RDC 166/2017 que traz os parâmetros típicos a serem avaliados em uma validação de método analítico a depender da natureza do método avaliado.

Parâmetro Avaliado	Identificação	Teste de Impurezas		Doseamento
		Quantitativo	Ensaio Limite	-dissolução (quantificação) -uniformidade de conteúdo -potência
Exatidão	não	sim	não	Sim
Precisão Repetibilidade	não	sim	não	Sim
Precisão Intermediária	não	sim <sup>(1)</sup>	não	sim <sup>(1)</sup>
Seletividade <sup>(2)</sup>	sim	sim	sim	sim
Limite de Detecção	não	não <sup>(3)</sup>	sim	não
Limite de quantificação	não	sim	não	não <sup>(3)</sup>
Linearidade	não	sim	não	sim
Intervalo	não	sim	não	sim

<sup>(1)</sup> Nos casos em que foi conduzida a reprodutibilidade, não é necessário conduzir a precisão intermediária.

<sup>(2)</sup> Nos casos de ensaios de identificação, pode ser necessária a combinação de dois ou mais procedimentos analíticos para atingir o nível necessário de discriminação.

<sup>(3)</sup> Pode ser necessário em alguns casos.

**Figura 1 - Parâmetros a serem considerados na validação analítica. Fonte: RDC 166/2017.**

Uma análise estatística dos dados obtidos durante a validação de um método deve ser feita para demonstrar a validade do método analítico. Os parâmetros primários usados para a interpretação de resultados de validação de métodos analíticos são o cálculo da média, desvio padrão, desvio padrão relativo, intervalos de confiança, análises de regressão e testes como teste-F e teste-t. O objetivo das análises estatísticas é de sumarizar um conjunto de dados que possa prover um entendimento da característica do método examinado (BELOUFA et al., 2017).

Os métodos estatísticos usados durante a validação de métodos analíticos envolvem conhecimentos básicos de estatística. Apesar de existirem pacotes estatísticos disponíveis para a realização dos cálculos, é importante o entendimento da base matemática por trás destes cálculos. É essencial que o analista esteja familiarizado com elementos básicos de estatística. A estatística usada para interpretações de dados de validação deveria ser incorporada em procedimentos operacionais das empresas e especificados no protocolo e relatório de validação (BELOUFA et al., 2017).

Linearidade de um método analítico é a suposição de que existe uma relação de linha reta entre uma variável x de entrada e uma variável y de saída e pode ser escrita

matematicamente pela expressão  $y = f(x)$  se a linha passa pela origem ou pela expressão  $y = f(x) + \delta$  se a linha não cruza pela origem (ARAUJO, 2009).

O método recomendado para a estimativa da equação é o método dos mínimos quadrados (MMQ). Este método pode ter um fator de ponderação, sendo então denominado método dos mínimos quadrados ponderados (MMQP) ou pode ser aplicado diretamente, sem fator de ponderação, sendo então denominado método dos mínimos quadrados ordinários (MMQO). O MMQO também pode ser entendido como um caso particular do MMQP no qual o fator de ponderação é igual a 1 (ANVISA, 2017).

#### **4. METODOLOGIA**

Será realizada uma revisão bibliográfica dos principais testes estatísticos possíveis de serem aplicados à validação de diferentes tipos de métodos analíticos. A revisão será realizada pela análise de livros, artigos, etc. e será direcionada para os testes estatísticos aplicados ao parâmetro linearidade de uma validação de método analítico, especificamente os testes relacionados à verificação da homocedasticidade de dados experimentais pois são relatados diversos testes para tal avaliação.

## 5. REVISÃO BIBLIOGRÁFICA

### 5.1. HOMOCEDASTICIDADE/HETEROCEDASTICIDADE

A classificação para a aplicação entre método dos mínimos quadrados ordinários (MMQO) ou método dos mínimos quadrados ponderados (MMQP) é baseado no comportamento dos desvios padrões ou variâncias das respostas sobre o intervalo de trabalho da curva de calibração. Existem diversos testes estatísticos que podem ser usados para se checar a homogeneidade/heterogeneidade dos desvios padrão ou variâncias (RAPOSO, 2016).

Apesar da maioria dos relatos na literatura sejam referentes a dados em MMQO, é importante se notar que o uso do MMQO deve ser usado apenas quando os dados experimentais têm variância constante (homocedasticidade) enquanto o MMQP é apropriado quando a variância varia (heterocedasticidade). Normalmente, a escolha da técnica de adaptação dependerá da faixa de concentrações de interesse. Quando uma faixa estreita é considerada, um modelo linear comum ou não ponderado é usualmente adaptado, enquanto uma faixa maior pode exigir um modelo mais complexo ou ponderado (RAPOSO, 2016).

A heterocedasticidade pode ser considerada como uma propriedade de um sistema de medição (usado com um procedimento de medição específico) e definido como uma propriedade pela qual a variância metrológica depende do valor do mensurando. Por outro lado, homocedasticidade pode ser definido como uma propriedade pela qual a variância metrológica é constante para qualquer valor do mensurando (ÁLVAREZ; CORRES; COMESAÑA, 2011).

Por muitas razões, a maioria das quais são de natureza estatística, é desejável lidar com dados homocedásticos. As variáveis de resposta  $y_1...y_n$  são amostradas de forma independente ou, para uma restrição menor, possuem covariâncias nulas e são constituídas por uma população com uma variância comum  $\sigma^2$  (McCORMICK; ROACH, 1987; SAYAGO; ASUERO, 2004).

Para poder decidir se a variância das respostas é constante ou não é constante para os diferentes valores da variável independente, é necessário possuir várias respostas medidas  $y_i$  para o mesmo valor da variável preditora  $x_i$ , pelo menos em dois níveis de concentração (ORTIZ, SÁNCHEZ, SARABIA, 2009).

Diversos métodos de comparação de variâncias em diferentes níveis de fator são reportados em literatura, tais como: teste F, teste de Bartlett, teste de Cochran, teste de Breusch Pagan, teste de Burr e Foster, teste de Hartley, teste de Levene, teste de Brown-Forsythe, etc.

## 5.2. TESTES DE COMPARAÇÃO DE VARIÂNCIAS

Os testes de comparação de variâncias são testes de hipóteses na qual se testam:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 \\ H_1: \text{pelo menos um dos } \sigma_i^2 \text{ é diferente, } i = 1, \dots, n \end{cases}$$

O teste F é usado para a comparação entre duas variâncias de populações normais. A hipótese nula  $H_0: \sigma_1^2 = \sigma_2^2$  é testada através de uma estatística em forma de razão de estimativas  $s_1^2/s_2^2$  (onde  $s_1^2$  é a maior variância estimada e  $s_2^2$  é a menor variância estimada), que segue uma distribuição conhecida como distribuição F (teste unicaudal quando a alternativa para a hipótese nula é  $s_1^2 > s_2^2$  ou teste bicaudal quando a alternativa é simplesmente  $s_1^2 \neq s_2^2$ ). Se o valor computado de F exceder o valor tabelado de  $F_{1-\alpha}(f_1, f_2)$ , a probabilidade de que a diferença entre as duas variâncias seja causada apenas pelo acaso é menor do que a probabilidade especificada, sendo possível a rejeição da hipótese nula, com a probabilidade de se cometer um erro do tipo I, ou seja, o nível crítico tem um falso positivo especificamente definido, ao contrário de um falso negativo indefinido ou erro do tipo II (SAYAGO; ASUERO, 2004).

Como a variância muda suavemente com a resposta em aplicações de química analítica, a ISO propõe que seja realizado um teste F unicaudal para checar se a variância no nível de concentração mais alto é significativamente maior do que no menor nível de concentração (SAYAGO; ASUERO, 2004).

Porém, a abordagem de aplicação do teste F pela comparação entre replicatas no menor e maior valor de uma variável  $x$  não detectará oscilações de variância em níveis intermediários da variável  $x$ , implicando em aumento de custo e tempo de análise dos dados (DESIMONI, 1999). Visando evitar tal situação, existem diversos testes para a comparação de mais de duas variâncias se o experimentador temer um grande

afastamento da homogeneidade (teste de homogeneidade de variâncias), embora, a maioria destes depende da suposição da normalidade (SAYAGO; ASUERO, 2004).

O teste de Barlett é o único teste que pode ser aplicado para comparação de duas ou mais variâncias cujos graus de liberdade não são iguais, ou seja, pode ser utilizado em conjunto de dados não balanceados. Porém, o teste de Barlett possui a desvantagem de ser um teste relativamente complicado de ser aplicado e de ser muito sensível a desvios de normalidade da distribuição (SAYAGO; ASUERO, 2004). Dessa forma, quando aplicados a dados não normais, o teste de Barlett pode dar um veredito errôneo de não homogeneidade das variâncias (BOX, 1953).

O teste de Cochran (COCHRAN, 1941) é designado a ser aplicado em situações nas quais o número de replicatas em grupos é igual. É um teste apropriado se existe a suspeita de que uma única variância é apreciavelmente maior do que as  $k - 1$  restantes, e cuja hipótese nula é de que as outras variâncias são igualmente variáveis. O teste é particularmente sensível para o caso em que todas as variâncias  $\sigma_k^2$  são esperadas de serem iguais, exceto uma variância que pode ser maior (SAYAGO; ASUERO, 2004). Este teste é relativamente simples, sendo possível de ser executado em softwares de tabelas eletrônicas a partir de uma estatística  $G$  frente a um valor tabelado. Se  $G < G_{1-\alpha}(n - 1, k)$ , onde  $k$  é o número de variâncias, a diferença entre as variâncias no nível de significância selecionado é explicada por probabilidades.

Apesar do teste de Cochran ser indicado em casos onde um número igual de pontos de medições para cada nível  $x$ , a ISO indica que pequenas diferenças no número de pontos podem ser ignoradas neste teste e aplica o critério de Cochran para o número de medições que ocorre na maioria dos  $x$  níveis. Ainda, o erro introduzido com arranjos de dados não balanceados, como em um caso onde um dado resultado foi descartado, pode ser pequeno e não afetaria a estatística do teste (SAYAGO; ASUERO, 2004). Existe ainda a possibilidade de se utilizar um método para repor dados onde um experimento balanceado apresentou dados rejeitados (DESIMONI, DAGHETTA, VALSECCHI, 1998).

O teste de Hartley pode ser aplicado quando os grupos de tratamento possuem o mesmo tamanho amostral e envolve a comparação da maior variância com a menor variância, sendo um teste cuja simplicidade pode ser vantajosa (DOWY, WEARDEN, 1991). O teste consiste em obter um valor de  $F_{max}$  que consiste na razão da maior

variância obtida pela menor variância e este valor é comparado numa distribuição de probabilidade que depende de  $k$  e  $f = n - 1$  graus de liberdade. Se  $F_{max}$  satisfaz a condição  $F_{max} \geq F_{max,k,f,\alpha}$ , rejeita-se  $H_0$  em favor de  $H_1: \sigma_1^2 \neq \sigma_2^2 \neq \dots \neq \sigma_k^2$ . Porém, devido a sensibilidade do teste a desvios de normalidade, se  $F_{max}$  for significativo, o teste pode indicar a desigualdade de variâncias ou falta de normalidade dos dados. Ainda, o teste tem menor poder estatístico do que os testes de Barlett ou Levene (SAYAGO; ASUERO, 2004).

O teste de Levene é um teste F de análise de variância para 1 fator em  $|y_{iv} - \bar{y}_i|$  que corresponde ao valor absoluto dos desvios em  $y_{iv}$  da média do grupo  $\bar{y}_i$ . Isso essencialmente converte um teste de variâncias em um teste de médias, que é relativamente não afetado por dados não normais. O teste se demonstrou robusto para não normalidade quando o tamanho das amostras é igual (SCHULTZ, 1983). Considerando o desvio absoluto de cada observação da média do grupo, em um grupo com pequenas observações de variância o valor de  $|y_{iv} - \bar{y}_i|$  será pequeno. Em contraste, um grupo com grandes variâncias, o valor de  $|y_{iv} - \bar{y}_i|$  será grande. Assim, para se testar a igualdade das variâncias dos grupos, Levene sugeriu que fosse realizado um teste ANOVA em  $|y_{iv} - \bar{y}_i|$  (GREEN, MARGERISON, 1977).

O teste de Brown e Forsythe (BROWN, FORSYTHE, 1974) é uma modificação do teste de Levene e sugere a utilização dos desvios absolutos das medianas dos grupos, que é uma estimativa mais robusta da tendência central, como uma alternativa à média para se computar os desvios absolutos. O teste de Levene utilizando-se as medianas dos grupos é mais robusto do que o teste original quando os grupos não possuem o mesmo tamanho amostral e quando os valores dos desvios absolutos possuem alta assimetria, causando violação da normalidade dos dados (SAYAGO; ASUERO, 2004).

Lim e Loh (1996) realizaram uma comparação de sete testes de igualdade de variâncias em termos de robustez e poder estatístico em um experimento simulando amostragens pequenas a moderadas. Os autores concluíram que o teste de Levene e o teste de Barlett com ajustes de curtose são robustos. Entre os testes, a versão *bootstrap* do teste de Levene demonstrou possuir maior poder estatístico.

Em modelos de regressão, o teste de Breusch-Pagan é amplamente utilizado para avaliação de heterocedasticidade. Breusch e Pagan (1979) derivaram um teste a partir



de um multiplicador de Lagrange para uma hipótese nula de homocedasticidade versus a hipótese alternativa que as variâncias dependem, de uma forma em alguns regressores que explicam a diferença das variâncias (ZAMAN, 2000). A estatística do teste se dá pela obtenção do quadrado dos resíduos obtidos de uma regressão linear e pela padronização de um vetor de resíduos, denotado por  $u_i$ , na qual cada vetor é a razão do resíduo ao quadrado pela razão da soma dos quadrados dos resíduos dividido pelo número de observações. Realiza-se, então, a regressão entre as variáveis respostas  $u$  pela variável explicativa  $\hat{y}$ , obtendo-se a estatística do teste calculando a soma dos quadrados da regressão de  $u$  sobre  $\hat{y}$  e dividindo o valor encontrado por 2. Sob a hipótese nula, esta estatística tem distribuição qui-quadrada com 1 grau de liberdade (PORTAL ACTION).

O teste de Breusch-Pagan tem sido difundido como o principal teste a ser utilizado na indústria farmacêutica para a avaliação da homocedasticidade no parâmetro linearidade de uma validação de método analítico. Porém, Zaman (2000), Li e Yao (2015) e Halunga, Orme e Yamagata (2017) reportam algumas limitações do teste.

## 6. CONSIDERAÇÕES FINAIS

Diversos testes para verificação da homogeneidade de variâncias (homocedasticidade) estão presentes em literatura. Contudo, cada teste possui limitações que devem ser consideradas a depender da aplicação.

Pode ser difícil na prática dizer se valores significativos do teste F são devidos à heterogeneidade da variância ou ao fato de que as populações não são normais. O teste de Cochran é recomendado pelo ISO, mesmo no caso de conjuntos de dados não balanceados. O teste de Bartlett é dependente do cumprimento da suposição de normalidade e, portanto, o teste de Levene agora o substituiu em grande parte. O teste de Levene é o teste F usual para igualdade de médias computadas em pseudo-observações, o qual se define como desvios absolutos dos dados de uma estimativa do grupo “centro”. Um teste mais robusto muito semelhante ao teste de Levene foi proposto por Brown e Forsythe, no qual se analisam os desvios das medianas dos grupos ao invés dos desvios em relação a média. O teste de Breusch-Pagan foi difundido na indústria farmacêutica, mas observa-se que a utilização do mesmo não é completamente fundamentada, corroborando com a ideia de que não há um conhecimento difundido acerca do tema entre analistas designados a realizar validações de métodos analíticos na indústria farmacêutica.

Assim, recomenda-se que cada teste seja utilizado após avaliação crítica de cada condição experimental e método analítico a ser validado.

## REFERÊNCIAS

ÁLVAREZ, J. S.; CORRES, R. C.; COMESAÑA, S. C. Letter to the Editor. Heteroscedasticity and homoscedasticity, and precision profiles in clinical laboratory sciences. *Clinica Chimica Acta*, 412 (2011) 2351-2352

ARAUJO, P. Key aspects of analytical method validation and linearity evaluation. *Journal of Chromatography B*, 877 (2009) 2224–2234.

BELOUAFIA, S. et al. Statistical tools and approaches to validate analytical methods: methodology and practical examples. *International Journal of Metrology and Quality Engineering*, 2017.

BOX, G.E.P. Non-normality and tests on variances. *Biometrika* 40 (1953):318–335

BRASIL, ANVISA. Agência Nacional de Vigilância Sanitária. Guia para tratamento estatístico da validação analítica – Guia nº 10, versão 1, de 30 de agosto de 2017.

BRASIL, ANVISA. Agência Nacional de Vigilância Sanitária. Resolução da Diretoria Colegiada - RDC Nº 166, 24/07/2017. Guia para validação de métodos analíticos - Julho, 2017.

BREUSCH, T.; PAGAN, A. A Simple Test of Heteroskedasticity and Random Coefficient Variation. *Econometrica* 47 (1979):1287-1294.

BROWN, M. B.; FORSYTHE, A. B. The small sample behavior of some statistics which test the equality of several means. *Technometrics* 16 (1974):129–132

COCHRAN, W.G. The distribution of the largest of a set of estimated variances as a fraction of the total. *Ann. Eugenics* 11 (1941):47–53.

DESIMONI, E. A program for the weighted least squares regression of unbalanced response arrays. *Analyst*124 (1999):1191–1196.

DESIMONI, E; DAGHETTA, A.; VALSECCHI, S. Evaluation of uncertainty by ordinary linear least squares regression of replicated data. A revised formulation to deal with unbalanced data set. *Ann. Chim. (Roma)* 88 (1998):601–609.

DOWY, S.; WEARDEN, S. *Statistics for Research*, 2nd ed. (New York: Wiley, 1991), 348–356.

GREEN, J. R.; MARGERISON, D. *Statistical Treatment of Experimental Data* (Amsterdam: Elsevier, 1977).

HALUNGA, A.; ORME, C., YAMAGATA, T. A Heteroskedasticity Robust Breusch-Pagan Test for Contemporaneous Correlation in Dynamic Panel Data Models. *Journal of Econometrics*. 2017. pp. 209-230. ISSN 0304-4076.

ICH Q2 (R1). *Validation of Analytical Procedures: Text and Methodology*. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2005.

LIM, T.-S; LOH, W. Y. A comparison of tests of equality of variances. *Comp. Stat. Data Anal.*22 (1996):287–301.

LI, Z.; YAO, J. Testing for Heteroscedasticity in High-dimensional Regressions. *arXiv:1510.00097 [stat.ME]*, 2017.

McCORMICK, D.; ROACH, A. *Measurement Statistics and Computation* (New York: Wiley, 1987), 303–321.

ORTIZ, M.C.; SÁNCHEZ, M.S.; SARABIA, L.A. Quality of Analytical Measurements: Univariate Regression. Comprehensive Chemometrics, Volume 1, 2009, pág. 127-169.

PORTAL ACTION. Site disponível em: <http://www.portalaction.com.br/analise-de-regressao/32-diagnostico-de-homocedasticidade>. Acessado em: 1 de Julho de 2019.

RAPOSO, F. Evaluation of analytical calibration based on least-squares linear regression for instrumental techniques: A tutorial review. Trends in Analytical Chemistry, 77 (2016) 167-185.

SAYAGO, A.; ASUERO, A.G. Fitting Straight Lines with Replicated Observations by Linear Regression: Part II. Testing for Homogeneity of Variances. Critical Reviews in Analytical Chemistry (2004), 34:3-4, 133-146.

SCHULTZ, B. On Levene's test and other statistics of variation. Evolutionary Theory 6 (1983):197-203

SIDDIQUI, M.R.; ALOTHMAN, Z. A.; RAHMAN, N. Analytical Techniques in Pharmaceutical Analysis: A review. Arabian Journal of Chemistry, 2013.

ZAMAN, A. The Inconsistency of the Breusch-Pagan Test. Journal of Economic and Social Research 2 (1) 2000, 1-11.