

PROCEDIMENTO ESTATÍSTICO PARA VALIDAÇÃO DE ANALISTAS EM ENSAIOS MICROBIOLÓGICOS QUANTITATIVOS

Veruska Norie TAKADA¹
Anderson Castro Soares de OLIVEIRA¹
Marcelino Alves Rosa de PASCOA¹
Marcelo Aparecido MARTINS²

- RESUMO: Laboratórios de microbiologia de alimentos realizam vários tipos de ensaios para detectar a presença e quantificar os microorganismos. Entretanto estes ensaios estão sujeitos a interferências e imperfeições. Assim, os laboratórios devem estabelecer um programa de controle de qualidade para garantir a qualidade das práticas realizadas no laboratório. Neste contexto, este trabalho tem como objetivo estabelecer um protocolo estatístico de validação de analistas, como parte de um sistema de controle de qualidade para análises microbiológicas quantitativas de alimentos. Para tanto foram realizadas simulações de validação de analistas, considerando um delineamento inteiramente casualizado e posterior análise dos dados por meio dos testes análise de variância (ANOVA) e análise de *deviance* (ANDEV). Os resultados mostraram que a ANDEV é melhor na detecção da variabilidade entre os analistas, tendo apresentado uma taxa menor de erro do tipo I, maior poder com um número menor de repetições, adequando-se melhor à rotina e à realidade dos laboratórios.
- PALAVRAS-CHAVE: Controle de qualidade; ANDEV; ANOVA; transformação de dados.

1 Introdução

A norma *ABNT NBR ISO/IEC 17025* (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2002) pode ser utilizada como um padrão para desenvolver

¹Universidade Federal de Mato Grosso - UFMT, Instituto de Ciências Exatas e da Terra, Departamento de Estatística, CEP: 78060-900, Cuiabá, MT, Brasil. E-mail: v3rusk4@gmail.com; andersoncso@gmail.com; marcelino.pascoa@gmail.com

²Centro Universitário Univag - UNIVAG, CEP: 78118-900, Varzea Grande, MT, Brasil. E-mail: marceloap.martins@hotmail.com

e estabelecer o sistema de qualidade em laboratórios de ensaios e calibração, de modo que todos os requisitos devem ser atendidos, para demonstrar que este é tecnicamente competente e capaz de gerar resultados válidos. Esta norma pode ser aplicada a diversos tipos de laboratórios inclusive, de microbiologia de alimentos, em que deve-se estabelecer uma estrutura organizacional criada para conduzir e garantir a qualidade, os procedimentos operacionais, os recursos necessários e as responsabilidades estabelecidas.

Todos os procedimentos analíticos dentro de um laboratório estão sujeitos a diversas fontes de erros sistemáticos e aleatórios que podem ser inerentes ao método ou relacionados às práticas de trabalho. Assim, um sistema de garantia da qualidade compreende todas as atividades de análise de laboratório de microbiologia, sendo as principais fontes de variabilidade do processo: pessoas, equipamentos, materiais, métodos e ambiente (HAYES, 1996).

Os ensaios microbiológicos devem ser realizados por ou sob a supervisão de uma pessoa qualificada em microbiologia, com formação adequada para as técnicas e manuseio dos instrumentos e necessário monitorar sua competência deve continuamente (HAYES, 1996). A realização de experimentos para monitoramento/validação dos analistas, pode ajudar a detectar a necessidade de treinamento para que se consigam resultados de alta qualidade nos ensaios. Estes experimentos devem ser realizados periodicamente ou quando da incorporação de um novo membro no quadro de colaboradores.

Os resultados de ensaios microbiológicos quantitativos indicam o nível microorganismos em uma matriz e podem refletir a dispersão dos organismos dentro, ou sobre, a matriz. Assim, a superdispersão é uma característica presente neste tipo de quantificação e os modelos probabilísticos mais utilizados para a análise deste tipo de dados microbiológicos são a Binomial negativa e Lognormal (JARVIS, 2008).

Na validação de analistas em ensaios laboratoriais são realizados experimentos em que os analistas são considerados os tratamentos, e é testado a hipótese nula de igualdade entre as médias dos resultados obtidos pelos analistas, ou seja se treinamentos são necessários ou não (COX, 2006; CASELLA e BERGER, 2010). Caso tenhamos que rejeitar a hipótese nula, isto significa que os analistas têm médias de resultados diferentes nas análises e treinamentos devem ser realizados (HAYES, 1996; JARVIS, 2008).

Para testar a igualdade entre os analistas em geral, utiliza-se o teste F por meio da análise de variância sob abordagem de modelos lineares. Entretanto para que o teste F obtenha resultados válidos os erros do modelo devem ser independentes, normalmente distribuído e com variância homogênea $N(0, \sigma^2)$ (DEAN e VOSS, 1999; QUINN e KEOUG, 2002; OLIVEIRA e FERREIRA, 2017). Quando esses pressupostos são violados, a taxa de erro tipo I pode aumentar e o poder do teste, diminuir (GLASS, 1972; HARWELL *et al.*, 1992; LIX *et al.*, 1996).

No intuito de corrigir uma falha dos pressupostos da ANOVA, muitas vezes, é utilizada a transformação de dados, entre as possíveis, a cerca de dados de contagem, faz-se o uso da transformação logarítmica de forma a estabilizar a variância dos dados e em seguida aplica-se novamente a ANOVA (RASMUSSEN, 1989;

YAMAMURA, 1999; O'HARA e KOTZE, 2010; OLIVEIRA e FERREIRA, 2017). Entretanto, para estes artifícios, há inúmeras limitações, mais notavelmente, questões complexas de interpretação, que devem ser feitas baseando-se nos valores transformados e não nas observações originais (LIX *et al.*, 1996; O'HARA e KOTZE, 2010).

Uma alternativa à transformação de dados é a utilização dos modelos lineares generalizados (MLG) introduzidos por (NELDER e WENDDERBURN, 1972). Os MLG é classe unificada de modelos para variáveis discretas e contínuas, em que a distribuição de probabilidade da variável resposta deve pertencer a família exponencial (NELDER e WENDDERBURN, 1972; DEMETRIO e CORDEIRO, 2007; MYERS *et al.*, 2012).

A Análise de *Deviance* (ANDEV) pode ser entendida como uma generalização da análise da variância para os modelos lineares generalizados, que permite comparar uma sequência de modelos encaixados visando verificar os efeitos dos tratamentos (DEMETRIO e CORDEIRO, 2007).

Neste contexto, este trabalho tem como objetivo estabelecer um protocolo estatístico de validação de analistas, como parte de um sistema de controle de qualidade para análises microbiológicas quantitativas de alimentos. Para tanto, foram realizadas simulações de validação de analistas a partir do modelo binomial negativo, considerando um delineamento inteiramente casualizado e posterior análise dos dados por meio dos testes análise de variância (ANOVA) e análise de *deviance* (ANDEV). A performance, dos teste de hipótese, foi avaliada por meio das taxas de erro tipo I e poder do teste.

2 Materiais e métodos

O estudo do protocolo de validação de analistas foi dividido em duas partes:

- Estudo de simulação;
- Estudo de dados reais.

2.1 Estudo de simulação

O estudo de simulação foi realizado em duas etapas: a primeira consiste na geração dos dados e a segunda a avaliação da performance dos teste de hipótese, por meio da obtenção das taxas de erro tipo I e poder do teste.

2.1.1 Geração dos conjuntos de dados

Os dados da simulação foram gerados considerando um delineamento inteiramente casualizado a partir do modelo binomial negativo utilizando a parametrização $E(Y) = \mu$ e $Var(Y) = \mu + \frac{\mu^2}{k}$ e considerando os analistas como tratamentos. Para verificar o efeito da variabilidade entre os tratamentos, foi

considerado razão (θ) entre a variância e a média entre dos tratamentos, dado por:

$$\theta = \frac{Var(Y_{.j})}{E(Y_{.j})}$$

Foram consideradas quatro diferentes situações:

- Médias e variâncias iguais entre os tratamentos;
- Médias iguais entre os tratamentos e pelo menos um tratamento com variância diferente;
- Pelo menos um tratamento com média diferente e variâncias iguais entre os tratamentos;
- Pelos menos um tratamento com média e variância diferentes.

Nas simulações foram considerados: número de tratamentos variando entre 3 e 10; número de repetições variando entre 3 e 20; θ variando entre 1,5 e 10, a cada 0,5. Desta forma em cada situação havia 2592 diferentes configurações, referentes as combinações de 8 números de tratamentos, 18 números de repetições e 18 valores de θ . Para cada uma dessas combinações foram realizadas 10000 simulações.

Todas as simulações foram realizadas utilizando o software ?.

2.1.2 Análise das simulações

Em todos os conjuntos de dados simulados foram utilizados a análise de variância (ANOVA), análise de variância considerando a transformação logarítmica $\log(x + 1)$ (ANOVAT) e a análise de deviance (ANDEV) considerando um modelo binomial negativo, para verificar a hipótese de não existência dos efeitos de tratamentos sobre a variável resposta.

Para verificar a pressuposições da análise de variância, foi realizado o teste de Komogorov-Smirnov para normalidade dos erros e do teste de Levene para homogeneidade de variância (QUINN e KEOUG, 2002). Quando as pressuposições não foram atendidas, aplicou-se transformação logarítmica e repetida a ANOVA.

Para obter a taxas de erro tipo I foram avaliadas no processo de simulação os casos em que a hipótese nula foi rejeitada erroneamente, ou seja quando a médias eram iguais entre si e o teste de hipótese rejeitou a hipótese nula, assim a taxa de erro tipo I foi dada por:

$$\hat{\alpha} = \frac{n_0}{10000}$$

em que n_0 representa o número de resultados simulados que foram considerados significativos, considerando nível de significância de 5%.

O poder do teste foi obtido de maneira análoga, foram verificados no processo de simulação os casos em que a hipótese nula foi rejeitada corretamente, ou seja quando havia pelo menos um tratamento com média diferente e o teste de hipótese

rejeitou a hipótese nula, assim o poder do teste foi dado por:

$$1 - \hat{\beta} = \frac{n_0}{10000}$$

2.2 Estudos dos dados reais

Foram realizados 30 experimentos de ensaios de contagens de bactérias dentro de um laboratório de análises de alimentos. Para cada ensaio, foram utilizadas amostras artificialmente contaminadas.

Para realização dos experimentos, todos os equipamentos utilizados foram calibrados, para não introduzir viés nos resultados experimentais.

Os experimentos realizados para a validação de analistas foram instalados segundo um delineamento inteiramente casualizado, sendo considerado como unidade experimental amostras homogeneamente contaminadas, e os tratamentos foram os 5 (cinco) analistas que realizaram os ensaios no laboratório. Os equipamentos, reagentes e meios de culturas foram os mesmos para todos os analistas. Para a aleatorização foi feito um sorteio na sequência em que os analistas fariam os ensaios.

Para cada tipo de ensaio, o protocolo experimental seguiu as seguintes etapas:

1. Preparou-se uma amostra inicial e cada analista realizou um ensaio nesta amostra;
2. Calculou-se a média e a variância entre os analistas;
3. Calculou-se a razão entre variância e média;
4. Determinou-se o número de repetições pela razão entre a variância e média;
5. Comparou os resultados entre os analistas utilizando Análise de Variância (ANOVA) e a Análise de Deviance (ANDEV) considerando um modelo binomial negativo. Em ambos os testes utilizou-se significância de 5%.

3 Resultados e discussão

3.1 Estudo de simulação

3.1.1 Taxa de erro tipo I

Na figura 1, é apresentada a taxa de erro tipo I em função do número de repetições e razão θ , considerando variâncias iguais com 3, 5, 7 e 10 tratamentos. Verifica-se que a ANDEV e ANOVAT apresentaram taxas de erro tipo I similares e se mostraram mais eficientes para todas as quantidades de repetições utilizadas, pois comete um menor erro em todas as situações simuladas. Em relação a ANOVA, observa-se que independente do número de tratamento, com poucas repetições, a taxa de erro tipo I é superior a ANDEV e ANOVAT, e conforme

se aumenta o número de repetições, tende a aproximar seus valores com as outras técnicas. Também se observou que, com maiores valores da razão θ , a três técnicas apresentaram maiores valores de taxa de erro tipo I em todas os números de repetições, indicando que o aumento da variabilidade entre os tratamentos afetam a probabilidade de aceitar H_0 .

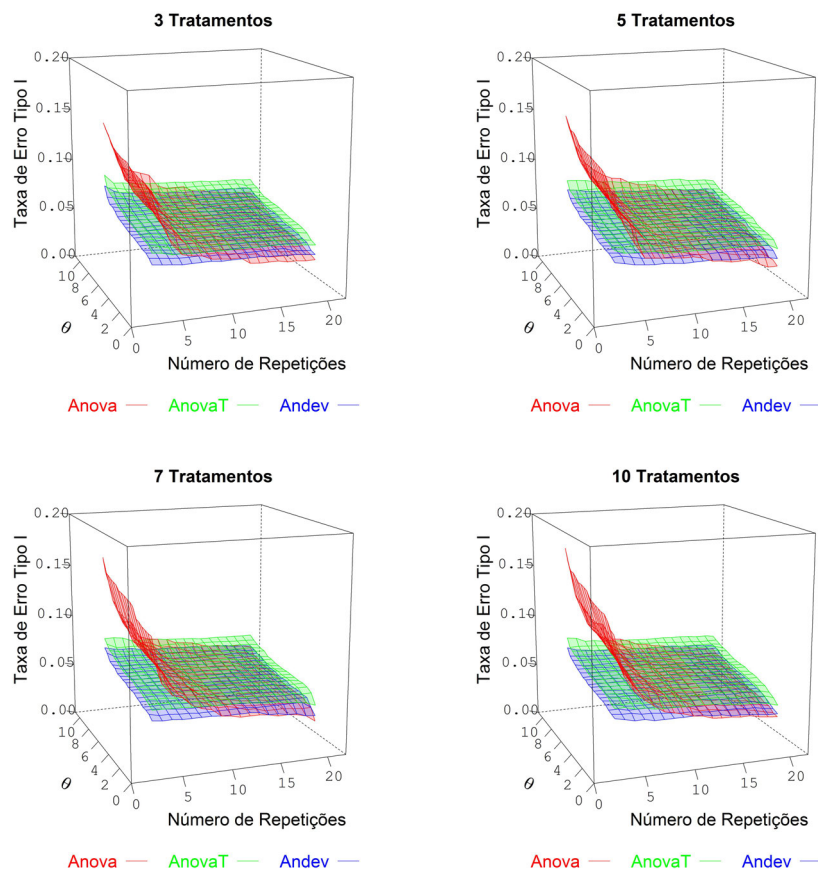


Figura 1 - Taxa de erro tipo I em função do número do número de repetições e razão θ com médias e variâncias iguais. Análise de variância - ANOVA, análise de variância com dados transformados - ANOVAT e análise de deviance modelo binomial negativo - ANDEV.

Quando considerado pelo menos um grupo com variância diferente (Figura 2), verifica-se que ocorre um leve aumento da taxa de erro tipo I na ANOVA, enquanto na ANDEV e ANOVAT tem seus valores de erro praticamente inalterados. Desta forma é possível observa uma maior influencia da heterogeneidade da variâncias na

ANOVA.

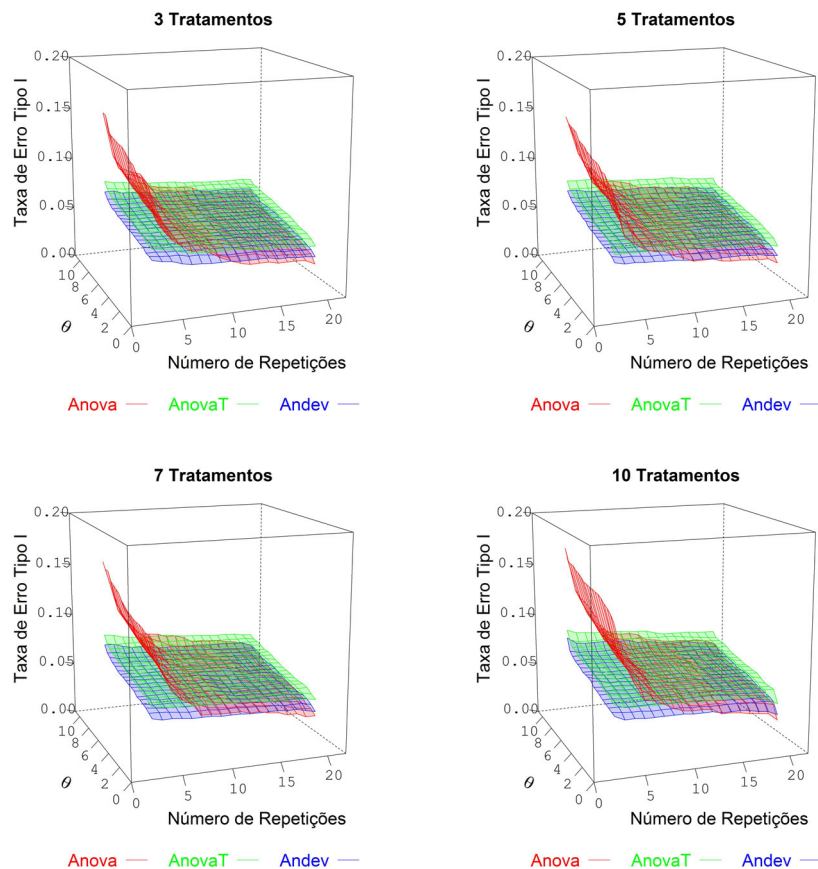


Figura 2 - Taxa de erro tipo I em função do número do número de repetições e razão θ com médias com médias e variâncias diferentes. Análise de variância - ANOVA, análise de variância com dados transformados - ANOVAT e análise de deviance modelo binomial negativo - ANDEV.

Lantz (2013) verificou que, para distribuições não normais, a taxa de erro tipo I na ANOVA é dependente de um grau de normalidade, ou seja, quanto a amostra se aproxima da normal. Assim, o aumento da taxa de erro tipo I na ANOVA com número de repetições maiores, é devido a média da amostras se tornarem aproximadamente normal (CASELLA e BERGER, 2010).

Box (1954) verificou que grupos com variâncias desiguais têm taxa de erro de Tipo I inflado, mesmo quando as dimensões das amostras eram iguais. Moder (2007) também verificou que a taxa de erro tipo I na ANOVA aumenta se os desvios não

são homogêneos, independente do número de repetições e tratamento. Liu (2015) afirma que a heterogeneidade das variâncias atua na inflação da taxa de erro tipo I e é sempre um problema na ANOVA, mesmo nos casos de heterogeneidade moderada. Yusof *et al.* (2013) observou que quando os problemas de não-normalidade e heterocedasticidade ocorrem simultaneamente, a taxa de erro Tipo I se infla, causando rejeições espúrias das hipóteses nulas.

Hubnerova (2007) comparou a ANOVA e ANDEV utilizando simulações de dados de contagem em delineamento inteiramente casualizado, considerando a distribuição Binomial Negativa, e seus resultados foram que a ANDEV têm melhor performance que a ANOVA, pois controla melhor a taxa de erro tipo I.

Rasmussen (1989) mostrou que quando é aplicada uma transformação de dados pode acarretar em uma taxa de erro de Tipo I apropriada. Ohara e Kotze (2010) observou que a transformação $\log(x + 1)$ pode ser uma boa alternativa para para situações em que dispersão dos dados é pequena e as contagens médias são grandes, nas demais situações é preferível utilizar modelos poisson e binômio negativo.

3.1.2 Poder do teste

O poder do teste em função do número de repetições e θ considerando variâncias iguais e ao menos uma com variância diferente são apresentados, respectivamente, nas figuras 3 e 4. Em todas as situações simuladas, verifica-se a ANDEV apresentou melhor desempenho, pois obteve o maior poder, independentemente do número de repetições. A ANOVA e ANOVAT tiveram seu desempenho relacionado as repetições, conforme aumenta o número de repetições ambas melhoram seu poder de detecção, porém somente a ANOVA tende a aproximar seus valores com os da ANDEV. Também é possível observar que quando considerado variâncias heterogêneas a ANOVA e ANOVAT tiveram perda do seu poder quando considerados poucas repetições.

Moder (2007) observou que se as variâncias são homogêneas, quanto maior o tamanho amostral ou número de repetições, maior é o poder da ANOVA. Liu (2015) verificou que sob condições de normalidade e com variâncias heterogêneas a ANOVA apresenta elevado poder. Hubnerova (2007) verificou que a ANDEV com modelo Binomial Negativo tem alto poder, independente do número de repetições.

Rasmussen (1989) verificou que o poder em geral é superior em dados transformados, quando a transformação adequada é utilizada. Nos resultados obtidos (Figura 3 e 4), a transformação logarítmica apresentou o menor poder em todas as situações simuladas, este um indicativo que está transformação utilizada não é apropriada para dados de uma distribuição binomial negativa. Yamamura (1999) sugere a transformação $\ln(x + \frac{k}{2})$ para ter poder alto do ANOVA.

3.1.3 Número de repetições

Foi estabelecido um número de repetições necessárias em função do número de analistas e razão θ , para isto considerou-se o número de repetições em que a taxa de erro tipo I seja inferior a 5% e o poder do teste superior a 90%, estes resultados

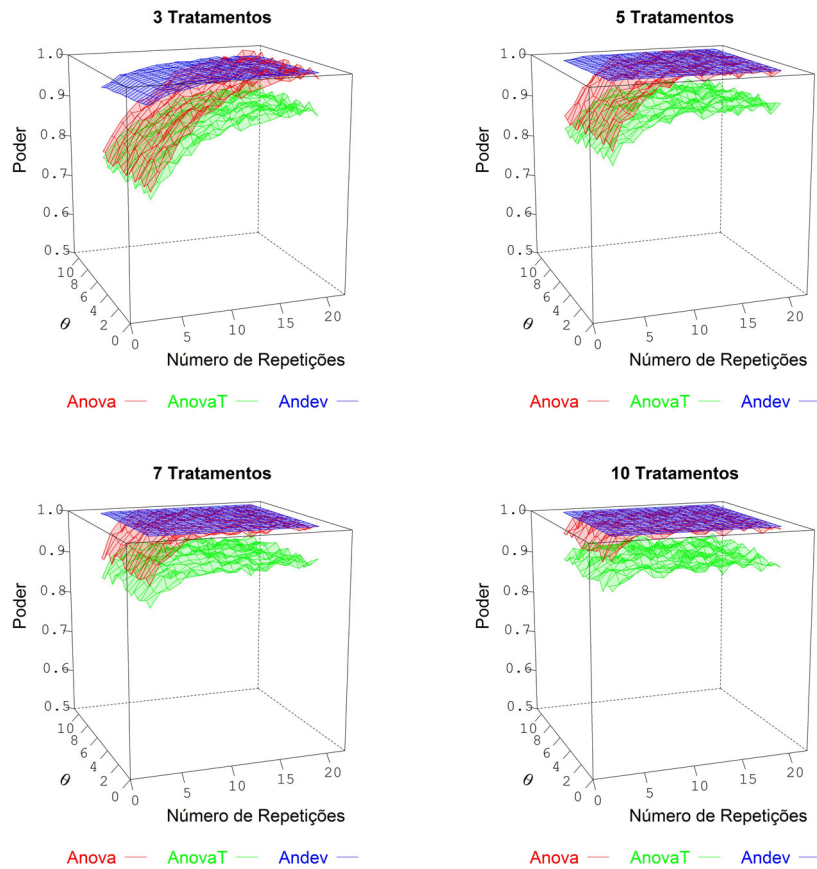


Figura 3 - Poder do teste em função do número em função do número do número de repetições e razão θ com médias diferentes e variâncias iguais. Análise de variância - ANOVA, análise de variância com dados transformados - ANOVAT e análise de deviance modelo binomial negativo - ANDEV.

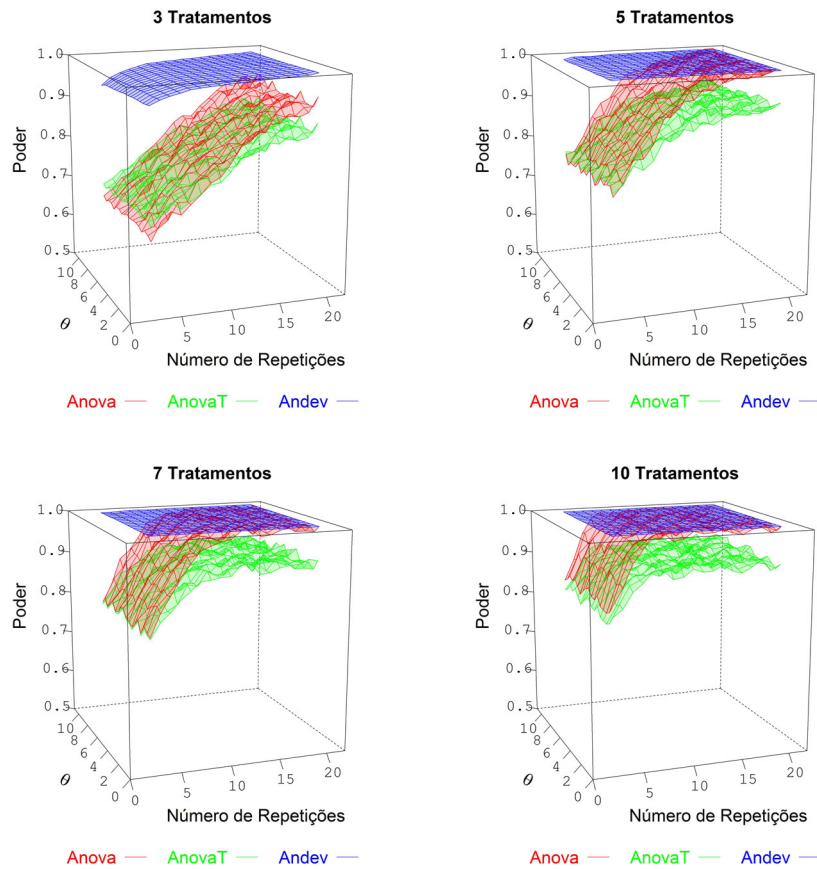


Figura 4 - Poder do teste em função do número em função do número do número de repetições e razão θ com médias e variâncias diferentes. Análise de variância - ANOVA, análise de variância com dados transformados - ANOVAT e análise de deviance modelo binomial negativo - ANDEV.

Tabela 1 - Número mínimo de repetições de acordo com o teste, número de analistas e razão θ , considerando que as variâncias são homogêneas

θ	ANOVA								ANDEV							
	3	4	5	6	7	8	9	10	3	4	5	6	7	8	9	10
1,5	8	6	6	6	5	5	5	4	5	4	3	3	3	3	3	3
2,0	8	6	6	6	6	6	6	4	5	4	3	3	3	3	3	3
2,5	9	6	6	6	6	6	6	4	5	4	3	3	3	3	3	3
3,0	9	7	6	6	6	6	6	5	6	5	4	3	3	3	3	3
3,5	9	7	7	6	7	6	6	6	6	5	4	3	3	3	3	3
4,0	9	7	7	7	7	7	7	6	6	5	4	4	3	3	3	3
4,5	9	7	7	8	7	7	7	6	6	5	5	4	3	3	3	3
5,0	9	7	7	8	7	8	8	6	7	5	5	4	3	3	3	3
5,5	9	8	8	8	7	8	8	7	7	6	5	5	3	3	3	3
6,0	9	8	8	8	7	8	8	7	7	6	5	5	3	3	3	3
6,5	10	8	8	8	8	8	8	7	7	6	5	5	3	3	3	3
7,0	10	8	8	8	8	8	8	8	8	6	6	5	4	3	3	3
7,5	10	9	8	9	8	8	8	8	8	7	6	5	4	4	3	3
8,0	11	9	9	10	8	8	8	8	8	7	6	6	4	4	4	4
8,5	12	10	10	10	9	8	8	8	8	7	6	6	5	4	4	4
9,0	13	11	10	10	10	9	9	8	9	8	7	6	5	5	5	4
9,5	13	11	11	11	11	10	9	8	9	8	7	6	5	5	5	5
10,0	14	13	12	11	11	11	10	9	9	9	8	7	6	5	5	5

foram divididos em duas situações com variâncias homogêneas e heterogêneas. Em nenhuma situação simulada a ANOVAT apresentou poder superior a 90%, desta forma foram determinados número de repetições apenas para ANOVA e ANDEV.

Na Tabela 1, são apresentados os número de repetições quando as variâncias são homogêneas. Verificou-se que os menores números de repetições são para $\theta = 1,5$, em que para a ANDEV é necessário de 3 repetições para 10 analistas a 5 repetições para 3 analistas, já ANOVA necessita de 4 a 8 repetições, com 10 e 3 analistas respectivamente. Quando $\theta = 10$ é necessário um número maior de repetições, em que a ANDEV variou de 5 a 9 e a ANOVA 9 a 14. Assim, verifica-se que o número de repetições depende da variabilidade entre analistas e do número de tratamentos.

Quando considerados grupos com variância heterogênea (Tabela 2), verificou-se que, para uma variabilidade pequena entre as unidades experimentais ($\theta = 1,5$), a quantidade mínima de repetições na ANDEV varia de 3 a 6, e para a ANOVA de 7 a 15. Esta diferença se mantém nas razões entre todos os valores de θ , sendo que a ANDEV sempre necessita de um número menor de repetições.

Comparando as tabelas 1 e 2, verifica-se que na ANOVA o número mínimo de repetições é superior a quando considera variâncias heterogêneas. Já para a ANDEV verifica-se que o número de repetições não é tão afetado pela heterogeneidade das variâncias.

Tabela 2 - Número mínimo de repetições de acordo com o teste, número de analistas e razão θ , considerando que as variâncias são heterogêneas

θ	ANOVA									ANDEV						
	3	4	5	6	7	8	9	10	3	4	5	6	7	8	9	10
1,5	15	14	13	12	11	9	7	6	6	5	4	3	3	3	3	3
2,0	15	14	13	12	11	9	8	7	6	5	4	3	3	3	3	3
2,5	15	14	13	12	11	9	8	7	6	5	4	3	3	3	3	3
3,0	16	14	13	12	11	10	8	7	7	6	5	3	3	3	3	3
3,5	16	15	13	12	12	10	8	7	7	6	5	4	3	3	3	3
4,0	16	15	13	13	12	10	8	7	7	6	5	4	4	3	3	3
4,5	16	16	14	13	12	10	8	7	7	6	6	4	4	3	3	3
5,0	17	16	14	13	12	10	8	7	8	6	6	5	4	3	3	3
5,5	17	16	14	13	12	10	8	8	8	7	6	5	4	4	3	3
6,0	17	16	15	13	13	11	8	8	8	7	6	5	4	4	4	3
6,5	19	17	15	14	13	11	8	8	8	7	6	5	4	4	4	4
7,0	20	17	15	14	13	11	9	9	9	7	7	5	5	4	4	4
7,5	20	17	16	14	13	12	9	9	9	8	7	6	5	5	4	4
8,0	20	19	16	15	13	12	9	9	9	8	7	6	5	5	4	4
8,5	20	19	17	15	14	12	10	9	9	8	7	7	6	5	5	4
9,0	20	19	17	16	14	12	11	10	10	9	8	7	6	6	5	4
9,5	20	19	17	17	14	13	11	11	10	9	8	8	7	6	6	5
10,0	20	19	18	17	15	14	12	11	10	10	9	8	7	6	6	6

Tabela 3 - Valores de média, variância e razão θ para os 30 ensaios de contagem de bactérias

Ensaio	Media	Variância	θ	Ensaio	Media	Variância	θ
1	11611,1	29633,5	2,55	16	11459,3	16449,8	1,44
2	9248,7	26127,3	2,82	17	8744,5	20893,7	2,39
3	11434,5	28441,1	2,49	18	3472,5	11198,7	2,03
4	10020,0	17770,2	1,77	19	9845,0	16479,6	1,67
5	10486,7	15758,4	1,50	20	10372,7	17624,2	1,70
6	9621,1	16078,3	1,67	21	3586,1	8238,1	2,30
7	7264,7	9659,8	1,33	22	8704,5	29782,8	3,42
8	5940,3	10481,6	1,76	23	8043,1	9424,1	1,17
9	10788,0	19940,6	1,85	24	10387,2	14415,2	1,39
10	9762,6	24934,6	2,55	25	8869,1	19015,6	2,14
11	4932,0	10442,2	2,12	26	11533,2	16268,0	1,41
12	7389,3	21653,6	2,93	27	4090,9	8686,7	2,12
13	11185,0	34392,1	3,07	28	9348,7	24829,7	2,66
14	3472,5	11198,7	3,22	29	9652,3	19557,2	2,03
15	9862,7	21536,7	2,18	30	9206,2	13752,2	1,49

3.2 Estudo de dados reais

Seguindo o protocolo experimental, inicialmente cada analista realizou o ensaio em uma única amostra, a fim de comparar a média e a variância entre os analistas. Na Tabela 3, é apresentada a média, variância e razão θ para cada ensaio realizado, verificou-se que valores de θ entre 1,17 e 3,42. Como o laboratório dispunha de 5 analistas estabeleceu-se que em todos os ensaios seriam realizadas treze repetições, número mínimo necessário para aplicação da ANOVA (Tabela 2).

Na tabela 4, é apresentado o valor-p para ensaio de comparação entre analistas considerando ANOVA e ANDEV referente ao 30 ensaios. Considerando o resultado da ANOVA, observa-se que para todos os ensaios realizados, que não existe evidências de diferença entre os analistas (*valor - p* > 0,05). Já para a ANDEV nos ensaios 14 e 22, rejeitou-se a hipótese nula (*valor - p* < 0,05), indicando que pelo menos um analista difere dos demais. Comparando os resultados dos ANOVA e ANDEV, verifica-se que nos ensaios 14 e 22 os dois testes discordaram com relação à significância. Ao analisar a Tabela 3, a razão (θ) para ambos os ensaios (14 e 22) são os maiores valores encontrados, 3,22 e 3,42 respectivamente, dando indícios de que o aumento na variabilidade entre os analistas é um fator que contribui para a divergência de conclusão entre os testes de hipótese.

Tabela 4 - Valor-p para ensaio de comparação entre analistas considerando ANOVA e ANDEV

Ensaio	ANOVA	ANDEV	Ensaio	ANOVA	ANDEV	Ensaio	ANOVA	ANDEV
1	0,9949	0,9998	11	0,9869	0,9904	21	0,6536	0,6133
2	0,9982	0,9997	12	0,8551	0,8127	22	0,3356	0,0393
3	0,9983	0,9998	13	0,7445	0,6650	23	0,3171	0,3695
4	0,9968	0,9998	14	0,0865	0,0419	24	0,8164	0,7479
5	0,9946	0,9987	15	0,9099	0,7363	25	0,8595	0,8762
6	0,9952	0,9999	16	0,9978	0,9917	26	0,4405	0,5300
7	0,9980	0,9996	17	0,7189	0,6101	27	0,9274	0,9135
8	0,9985	0,9997	18	0,9392	0,8482	28	0,3423	0,4107
9	0,9963	0,9998	19	0,4979	0,4584	29	0,9909	0,9849
10	0,7638	0,7500	20	0,2863	0,0863	30	0,1614	0,0607

Conclusões

Este é um estudo aplicado à laboratórios de análises de microbiologia de alimentos, desta forma deve ser compatível tanto com a rotina dos analistas, quanto com a realidade entre custo-benefício dos laboratórios.

Verificou-se nas situações simuladas que para todas as quantidades de repetições, a ANDEV se mostra mais eficiente, e cometeu menor taxa do erro tipo I quando comparada à ANOVA. A ANDEV também se mostra mais poderosa, detectando melhor a diferença entre as médias. O poder da ANOVA apenas se assemelha aos da ANDEV com uma quantidade grande de repetições, o que a torna inviável tanto na execução, quanto nos custos. Também observou-se que a ANOVAT apresenta uma taxa de erro tipo I semelhante a ANDEV, porém apresenta um poder bem inferior.

Para os dados reais conclui-se que o protocolo estabelecido para validação dos analistas, é viável para verificar a semelhança entre eles. Vale salientar a importância da preparação de uma amostra inicial, para análise que será realizada por todos os analistas. Por meio da amostra inicial, deve-se determina a razão entre a média e a variância (θ), e isto auxiliará na determinação do número de repetições.

Para realização de experimentos de qualificação de analistas em microbiologia, deve-se preferir a utilização análise de deviance com modelo binomial negativo, pois este é um teste mais poderoso e com taxa de erro tipo I menor, e necessita de um menor número de repetições para manter essas propriedades. Do ponto de vista prático, um menor número de repetições, na realização de um experimento, além de diminuir o custo da sua realização, também sofre menor influência de fatores não controlados, visto que o tempo necessário para realiza-lo é menor.

O protocolo estabelecido neste trabalho, deve ser utilizado apenas quando as amostras da validação forem homogêneas. Para trabalhos futuros, pode-se investigar um protocolo que considere amostras heterogêneas, aplicando-se assim outros delineamentos experimentais. E também a utilização de outros testes para comparar os analistas.

Agradecimentos

Agradecemos aos revisores e editores pelas sugestões apresentadas que ajudaram a melhorar a apresentação do trabalho.

TAKADA, V. M.; OLIVEIRA, A. C. S.; PASCOA, M. A .R.; MARTINS, M. A. Statistical methodology for analyst validation in microbiological quantitative assays. *Rev. Bras. Biom.*, Lavras, v.36, n.2, p.241-257, 2018.

■ **ABSTRACT:** *Food microbiology laboratories perform various types of tests to detect and quantify the presence of microorganisms. However these tests are subject to interference and imperfections. Thus, laboratories should establish a quality control program to ensure quality of practices carried out in the laboratory. In this context, the objective of this paper is establish a statistical protocol for validation of analysts as part of a quality control system for quantitative microbiological analysis. Thus, simulations of analyst validation were performed considering a completely randomized design and subsequent analysis of the data by means of ANOVA and ANDEV tests. The results show that ANDEV is better at detecting the variability among analysts and presented a lower rate of type I error, more power with a smaller number of repetitions, adapting better to the routine and the reality of laboratories.*

■ **KEYWORDS:** *Quality control; ANDEV; ANOVA; data transformation.*

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *ABNT NBR ISO/IEC 17025: Requisitos gerais para competência de laboratórios de ensaio e calibração.* Rio de Janeiro, 2002. 31p.

BOX, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems, i. effect of inequality of variance in the one-way classification. *Annals of Mathematical Statistics*, v.25, n.2, p.290-302, 1954.

CASELLA, G.; BERGER, R.L. *Inferência estatística.* 1.ed. Porto Alegre: Cengage Learning, 2010.

COX, D. R. *Principles of statistical inference.* 1.ed. Cambridge: Cambridge University Press, 2006.

DEAN, A. M.; VOSS, D. T. *Design and analysis of experiments.* New York: Springer, 1999.

CORDEIRO, G. M.; DEMÉTRIO, C. G. B. Modelos lineares generalizados. In: *Simposio Nacional de Experimentação Agronomica*, 2007.

GLASS, G.V; PERCY D.; PECKHAM, J. R. S. G. V. Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance. *Review of Educational Research*, v.42, n.3, p.237-288, 1972.

HARWELL, M. R.; RUBINSTEIN, E. N.; HAYES, W. S.; OLDS, C. C. Summarizing monte carlo results in methodological research: The one- and two-factor fixed effects anova cases. *Journal of Educational and Behavioral Statistics*, v.17, n.4, p.315-339, 1992.

HAYES, D. Quality assurance in the microbiology laboratory. *Accreditation and Quality Assurance*, v.1, n.1, p.18-23, 1996.

HÜBNEROVÁ, D. One-way ANOVA type model with negative binomial distribution. In: *International Environmetrics Society North American Regional Meeting*, 2007.

JARVIS, B. *Statistical aspects of the microbiological examination of foods*. 2.ed. New York: Elsevier Science, 2008.

LANTZ, B. The impact of sample non-normality on ANOVA and alternative methods. *British Journal of Mathematical and Statistical Psychology*, v.66, n.2, p.224-244, 2013.

LIU, H. *Comparing Welch's ANOVA, a Kruskal-Wallis test and traditional ANOVA in case of heterogeneity of variance*. Dissertação (Mestrado em Bioestatística), Virginia Commonwealth University, Richmond, 2015.

LIX, L. M.; KESELMAN, J. C.; KESELMAN, H. J. Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance f test. *Review of Educational Research*, v.66, n.4, p. 579-619, 1996.

MODER, K. How to keep the type i error rate in anova if variances are heteroscedastic. *Austrian Journal of Statistics*, v.36, n.3, p.179-188, 2007.

MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G.; ROBINSON, T. J. *Generalized linear models: with applications in engineering and the sciences*. New York: Wiley, 2012.

NELDER, J. A.; WEDDERBURN, R. W.M. Generalized linear models. *Journal of the Royal Statistical Society*, v.135, n.3, p. 370-384, 1972.

O'HARA, R. B.; KOTZE, D. J. Do not log-transform count data. *Methods in Ecology and Evolution*, v.1, n.2, p.118-122, 2010.

OLIVEIRA, H. M. P.; FERREIRA, E. B. Comparação de testes para igualdades de médias sob heterocedasticidade: simulação e aplicação. *Revista Brasileira de Biometria*, v.35, n.1, p.132-159, 2017.

QUINN, G.P.; KEOUGH, M. J. *Experimental design and data analysis for biologists*. Cambridge: Cambridge University Press, 2002.

RASMUSSEN, J. L. Data transformation, Type I error rate and power. *British Journal of Mathematical and Statistical Psychology*, v.42, n.2, p.203-213, 1989.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2017. Disponível em: <https://www.R-project.org/>.

YAMAMURA, K. Transformation using $(x + 0.5)$ to stabilize the variance of populations. *Researches on Population Ecology*, v.41, n.3, p.229-234, 1999.

YUSOF, Z. M., ABDULLAH, S. YAHAYA, S. S. Comparing the performance of modified F-t statistic with ANOVA and Kruskal Wallis test. *Applied Mathematics & Information Sciences*, v.7, n.2, p.403-408, 2013.

Recebido em 08.09.2016.

Aprovado após revisão em 09.06.2017.