

**TESTES ESTATÍSTICOS NA VALIDAÇÃO DE
MÉTODOS E ENSAIOS DE PROFICIÊNCIA
INTERLABORATORIAIS EM SEMENTES DE
NABO FORRAGEIRO**

VERÔNICA YUMI KATAOKA



2009

VERÔNICA YUMI KATAOKA

**TESTES ESTATÍSTICOS NA VALIDAÇÃO DE MÉTODOS E
ENSAIOS DE PROFICIÊNCIA INTERLABORATORIAIS EM
SEMENTES DE NABO FORRAGEIRO**

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

Orientador

Prof. Dr. Marcelo Silva de Oliveira

LAVRAS
MINAS GERAIS - BRASIL
2009

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Kataoka, Verônica Yumi.

Testes estatísticos na validação de métodos e ensaios de
proficiência interlaboratoriais em sementes de nabo forrageiro
/ Verônica Yumi Kataoka. – Lavras : UFLA, 2009.

244 p. : il.

Tese (Doutorado) – Universidade Federal de Lavras, 2009.

Orientador: Marcelo Silva de Oliveira.

Bibliografia.

1. Teste estatísticos. 2. Sementes. 3. Validação de métodos. 4
Ensaios de proficiência. I. Universidade Federal de Lavras. II. Título.

CDD – 519.56

VERÔNICA YUMI KATAOKA

**TESTES ESTATÍSTICOS NA VALIDAÇÃO DE MÉTODOS E ENSAIOS
DE PROFICIÊNCIA INTERLABORATORIAIS EM SEMENTES DE
NABO FORRAGEIRO**

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

APROVADA em 26 de janeiro de 2009


Prof. Dra. Maria Laene Moreira de Carvalho UFLA

Prof. Dr. Eric Batista Ferreira UNIFAL

Prof. Dr. Joel Augusto Muniz UFLA

Dr. Luiz Artur Costa do Vale MAPA

Prof. Dra. Thelma Sáfadi UFLA


Prof. Dr. Marcelo Silva de Oliveira
UFLA
(Orientador)

LAVRAS
MINAS GERAIS - BRASIL

Ao meu esposo, Rubens, e aos meus filhos, Alexander e Felipe,
pelo amor e compreensão,

OFEREÇO

Aos meus pais e irmãos que, mesmo a distância, sempre me deram
forças para que eu continuasse caminhando e superando todas as
dificuldades,

DEDICO.

AGRADECIMENTOS

Ao meu esposo, Rubens e aos meus filhos, Alexander e Felipe, que souberam ter paciência para compreender as minhas frequentes ausências e por terem sonhado comigo a conquista deste título.

Aos meus pais, Zilda e Osamu, que souberam conduzir com muita sabedoria a minha formação como cidadã e sempre deram prioridade à educação.

Aos meus irmãos, Patrícia, Flávio, Elisabete, Sheila e Frederico, fiéis companheiros em todos os momentos da minha vida.

Ao Prof. Marcelo, pela paciência, amizade e atenção durante todos esses anos, tendo um papel decisivo na condução de todos os meus trabalhos acadêmicos, mas principalmente pela contribuição, em forma de conselhos, para o meu crescimento intelectual, profissional e pessoal.

À Profa. Maria Laene, que sempre soube, com muita maestria, auxiliar na execução deste trabalho, bem como em outros momentos da minha vida.

Aos meus grandes amigos do Doutorado, Adriano, Claudiney, Eustáquio, Graziela, Josiane, Nádia e Rejane, que foram decisivos para que eu tivesse forças para concluir mais esta etapa. Agradeço sempre a Deus, por ter colocado no meu caminho essas pessoas. Espero que a nossa amizade, mesmo com toda distância, seja eterna.

Aos demais amigos do Mestrado e Doutorado do DEX, bem como ex-alunos da UFLA, atualmente morando em Lavras, Piracicaba, Belo Horizonte, Uberlândia e São Carlos, dos quais não citarei nomes, para não correr o risco de esquecer alguém, mas que sabem particularmente a quem estou me referindo. Afinal, foram inúmeras horas de convivência, tanto de trabalho como de lazer e

que me proporcionaram um conforto emocional tão necessário nos momentos decisivos.

A Luciana Souza, Marcela, Carla e Luciana Magda, que me conduziram pelo “Universo das Sementes” e às quais serei eternamente grata, pelo companheirismo, dedicação e ajuda.

Aos demais colegas de Mestrado e Doutorado do DEX e do DAG-Sementes, pelas palavras de incentivo e apoio nas dificuldades.

Aos amigos que fiz nesta cidade, Francisca, Dalva e Dona Itinha, que me deram apoio em todos os momentos e já são importantes na minha vida.

Aos professores e técnico-administrativos do DEX, pela competência, amizade e companheirismo.

A equipe do Laboratório de Análise de Sementes da UFLA, que me acolheu com carinho e sempre me auxiliou com presteza e disposição.

A todos os meus amigos de Petrolina (PE) e Juazeiro (BA), Marco e Ana, Rose, Betânia, Irene, Cláudia, Eurivalda e Hugo, que, mesmo de longe, com certeza, torceram pelo meu sucesso e os quais sempre trarei perto no meu coração.

A FAPEMIG e CNPq pelo suporte financeiro a esse trabalho, ao MAPA pela distribuição das sementes, e aos laboratórios participantes do processo de validação de métodos.

À Capes, pela concessão da bolsa de estudos.

Por fim, a Deus, que, na verdade, é o princípio, pois sem ele nada do que foi dito anteriormente teria sido possível.

SUMÁRIO

RESUMO.....	i
ABSTRACT	iii
CAPÍTULO 1.....	01
1 Introdução Geral	02
4 Referências Bibliográficas	05
CAPÍTULO 2 : Validação de Métodos.....	06
1 Resumo	07
2 Abstract.....	08
3 Introdução	09
4 Referencial Teórico	11
4.1 Programas de Validação de Métodos: aspectos gerais	13
4.2 Importância da cultura do nabo forrageiro.....	15
4.3 Testes de germinação.....	17
4.4 Seleção de lotes de acordo com a homogeneidade – Teste H.....	19
4.5 Caracterização dos lotes.....	22
4.6 Identificação de valores discrepantes (<i>outliers</i>).....	24
4.7 Identificação de <i>outliers</i> nas variâncias	30
4.8 Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel.....	34
4.9 Análise discriminante	41
4.10 Comparação de métodos.....	43
5 Material e Métodos	47
5.1 Seleção de lotes de acordo a homogeneidade – Teste H.....	47
5.2 Caracterização dos lotes - Teste de germinação inicial	50
5.3 Distribuição do material para os laboratórios	53
5.4 Das técnicas estatísticas	54
5.4.1 Identificação de valores discrepantes (<i>outliers</i>).....	56
5.4.2 Identificação de <i>outliers</i> nas variâncias	59

5.4.3	Análise de variância.....	62
5.4.4	Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel.....	63
5.4.5	Comparação dos métodos.....	69
5.4.6	Análise discriminante.....	71
5.5	Rotinas no R.....	79
6	Resultados e Discussão.....	81
6.1	Seleção de lotes de acordo a homogeneidade – Teste H.....	81
6.2	Caracterização dos lotes - Teste de germinação inicial.....	82
6.3	Teste de germinação nos laboratórios selecionados.....	84
6.4	Identificação de valores discrepantes (<i>outliers</i>).....	89
6.5	Identificação de <i>outliers</i> nas variâncias.....	91
6.6	Análise de variância.....	93
6.7	Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel.....	95
6.8	Comparação dos métodos.....	101
6.9	Análise discriminante.....	103
6.10	Discussão geral.....	106
7	Conclusões.....	109
8	Referências Bibliográficas.....	110
CAPÍTULO 3: Ensaio de Proficiência.....		118
1	Resumo.....	119
2	Abstract.....	120
3	Introdução.....	121
4	Referencial Teórico.....	123
4.1	Programas interlaboratoriais: aspectos gerais.....	123
4.2	Ensaio de Proficiência.....	128
4.3	Método do Z-score.....	131
4.4	Método Campos (1988).....	133
4.5	Análise de agrupamento.....	134
4.6	Análise generalizada de Procrustes.....	138

5.4.3	Análise de variância.....	62
5.4.4	Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel.....	63
5.4.5	Comparação dos métodos.....	69
5.4.6	Análise discriminante.....	71
5.5	Rotinas no R.....	79
6	Resultados e Discussão.....	81
6.1	Seleção de lotes de acordo a homogeneidade – Teste H.....	81
6.2	Caracterização dos lotes - Teste de germinação inicial.....	82
6.3	Teste de germinação nos laboratórios selecionados.....	84
6.4	Identificação de valores discrepantes (<i>outliers</i>).....	89
6.5	Identificação de <i>outliers</i> nas variâncias.....	91
6.6	Análise de variância.....	93
6.7	Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel.....	95
6.8	Comparação dos métodos.....	101
6.9	Análise discriminante.....	103
6.10	Discussão geral.....	106
7	Conclusões.....	109
8	Referências Bibliográficas.....	110
CAPÍTULO 3: Ensaio de Proficiência.....		118
1	Resumo.....	119
2	Abstract.....	120
3	Introdução.....	121
4	Referencial Teórico.....	123
4.1	Programas interlaboratoriais: aspectos gerais.....	123
4.2	Ensaio de Proficiência.....	128
4.3	Método do Z-score.....	131
4.4	Método Campos (1988).....	133
4.5	Análise de agrupamento.....	134
4.6	Análise generalizada de Procrustes.....	138

5 Material e Métodos	141
5.1 Dos dados.....	141
5.2 Das técnicas estatísticas	144
5.2.1 Método do Z-score.....	145
5.2.2 Método Campos (1988)	150
5.2.3 Análise de agrupamento.....	158
5.2.4 Análise generalizada de Procrustes.....	162
5.3 Rotinas no R	165
6 Resultados e Discussão	166
6.1 Simulação 1	166
6.2 Simulação 2	173
6.3 Simulação 3	177
6.4 Discussão geral	181
7 Conclusões	185
8 Referências Bibliográficas	186
Considerações Finais	190
ANEXOS	192

RESUMO

Kataoka, Verônica Yumi. **Testes estatísticos na validação de métodos e ensaios de proficiência interlaboratoriais em sementes de nabo forrageiro**. 2009. 244 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, Lavras, MG.*

Dentre as várias demandas de pesquisas na área de sementes, no que diz respeito à proposição e à investigação de técnicas estatísticas apropriadas para a avaliação de diversos processos nesta área, estão sendo consideradas neste trabalho duas delas. A primeira demanda refere-se ao processo de validação de métodos para teste de qualidade de sementes, em que uma avaliação estatística criteriosa dos resultados é essencial, uma vez que, após a validação, o método é incorporado na Regra para Análise de Sementes (RAS), passando a ser utilizado como base de comercialização e certificação de sementes. A segunda demanda está associada aos programas de ensaios de proficiência, já que as sucessivas checagens interlaboratoriais, além da implantação de sistemas de qualidade, fazem parte das exigências para o credenciamento de laboratórios de sementes junto a órgãos como a *International Seed Testing Association* (ISTA) e o Ministério de Agricultura Pecuária e Abastecimento (MAPA). Neste contexto, os problemas de pesquisa abordados neste trabalho foram: avaliar técnicas estatísticas complementares a um padrão de análise já determinado pela ISTA, num processo de validação de métodos para o teste de germinação em sementes de nabo forrageiro e investigar a adequação de três técnicas estatísticas, além do método do Z-score, para avaliar os resultados de um ensaio de proficiência, utilizando dados simulados, similares às condições reais. Foram desenvolvidas também rotinas específicas no software R, de código aberto e gratuito, que serviram para a otimização das análises. Como conclusão geral, verificou-se que as técnicas estatísticas propostas como complementares ao padrão da ISTA podem ser utilizadas em um programa de validação de métodos para sementes, uma vez que os resultados possibilitaram a verificação da eficiência desses procedimentos nas diversas etapas do processo. Foi possível comprovar que os métodos propostos para o teste de germinação em sementes de *Raphanus sativus* var. *Oleiferus*, com utilização de temperatura alternada 20°-30°C, em substrato papel ou areia, apresentam exatidão, robustez e precisão (repetibilidade e reprodutibilidade), podendo ser, assim, considerados validados. As técnicas de

* Comitê Orientador: Marcelo Silva de Oliveira – DEX/UFLA (orientador), Maria Laene Moreira de Carvalho – DAG/UFLA, Eric Batista Ferreira – DCE/UNIFAL (coorientadores).

análise generalizada de Procrustes, método campos e análise de agrupamento, quando utilizadas conjuntamente com o método do Z-score, são promissoras para a avaliação dos resultados de ensaios de proficiência na área de sementes.

ABSTRACT

KATAOKA, Verônica Yumi. **Statistical tests for methods validation and interlaboratory proficiency tests in oil radish seeds.** 2009. 244 p. Thesis (PhD in Statistics and Agricultural Experiments) – Universidade Federal de Lavras, Lavras, MG, Brazil.*

Among the various research demands in the area of seeds, regarding the proposition and investigation of proper statistical techniques for the assessment of several processes in the area, two of them are considered in this work. The first demand refers to the process of method validation for seed quality test, in which a solid statistical evaluation is essential, as, after validation, the method is incorporated into the Rules for Seed Testing, and is therefore used as basis for seed trade and certification. The second demand is associated with proficiency essay programs, given that successive inter-laboratory verifications, in addition to implementation of quality systems, are part of the requirements for seed laboratory accreditation at agencies such as the International Seed Testing Association (ISTA) and the Brazilian Ministry of Agriculture, Livestock and Supply (MAPA). In this context, the research problems approached in this work were to: assess complementary statistical techniques for an analysis standard already determined by ISTA, in a process for method validation for oil radish seed germination test, and investigate the appropriateness of three statistical techniques, in addition to the Z-score method, to evaluate the results of a proficiency essay, using simulated data, similar to real conditions. Specific routines have also been developed in the R software, open coded and free of charge, which served for analysis optimization. As a general conclusion, we verified that the statistical techniques proposed as complementary to the ISTA standard can be used in a program for validation of seed methods, once the results allowed for the verification of the efficiency of such procedures in the several stages of the process. It was possible to evidence that the proposed methods for germination test in seeds of *Raphanus sativus* var. *Oleiferus*, with the use of alternate temperatures of 20°-30°C, in substrates paper or sand, present accuracy, robustness and precision (repeatability and reproducibility), and therefore can be considered as validated. The techniques of generalized Procrustes analysis, Campos method and grouping analysis, whenever jointly

* Guidance Committee: Marcelo Silva de Oliveira – DEX/UFLA (Adviser), Maria Laene Moreira de Carvalho – DAG/UFLA (co-advisor), Eric Batista Ferreira – DSC/UNIFAL (Co-advisor).

used with the Z-score method, are promising in the evaluation of proficiency essay results in the area of seeds.

INTRODUÇÃO

Este livro é dedicado a todos os que desejam aprender a programar em Python.

Ele é escrito para quem está começando a aprender a programar.

Se você já sabe programar em outra linguagem, este livro também é para você.

Este livro é escrito em um estilo simples e direto, com exemplos práticos.

Ele é escrito em um estilo simples e direto, com exemplos práticos.

Este livro é escrito em um estilo simples e direto, com exemplos práticos.

Este livro é escrito em um estilo simples e direto, com exemplos práticos.

CAPÍTULO 1

Este capítulo apresenta o ambiente de desenvolvimento.

Ele é escrito em um estilo simples e direto, com exemplos práticos.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

Este capítulo apresenta o ambiente de desenvolvimento.

1 INTRODUÇÃO GERAL

Na pesquisa científica, a Estatística atua com as suas técnicas em todas as etapas, sendo que geralmente as evidências empíricas são analisadas por amostragem e estatística descritiva, testadas por meio da inferência estatística, e que ao final permitem algum tipo de tomada de decisão.

Machado et al. (2005) salientam que o uso da metodologia estatística não salva uma pesquisa mal planejada ou mal conduzida, e que a participação de um especialista em Estatística deve iniciar-se, pelo menos, na fase de planejamento.

Corroborando com essas idéias, especificamente no caso do programa do método de validação de testes na área de sementes da International Seed Testing Association (ISTA), esta organização preconiza que todo o processo deve ser acompanhado por dois estatísticos: um que é responsável por todas as orientações durante a execução e outro para revisar os resultados obtidos.

Esta recomendação se justifica pelo fato de que uma avaliação estatística criteriosa dos resultados é uma fase essencial nesse processo, uma vez que, após a validação, o método é incorporado nas Regras para Análise de Sementes da ISTA (RAS - ISTA), passando a ser utilizada como base de comercialização e certificação de sementes em todo o mundo.

Se o método não atender aos requisitos de exatidão, robustez, reprodutibilidade e repetibilidade, a consequência direta será a sua aplicação de forma inadequada em futuros testes.

Então, a discussão aqui não perpassa apenas pelo fato de que, no momento da tomada de decisões, podem-se cometer os erros tipo I ou II, já que esses estão sempre implícitos nos processos e podem ser controlados, mas,

também, pela utilização de métodos estatísticos inadequados e ou insatisfatórios para análise dos resultados.

No Brasil, para atender a uma demanda do mercado interno, o Ministério da Agricultura, Pecuária e Abastecimento (MAPA), em consonância com as regras internacionais, já está desenvolvendo um programa de validação de métodos que, associado à implantação de sistemas de qualidade como uma exigência para credenciamento de laboratórios de sementes, e as sucessivas checagens interlaboratoriais (denominadas de ensaios de proficiência) que serão realizadas, aumentarão a possibilidade de uniformização de testes e procedimentos.

Neste contexto, verifica-se a necessidade premente de investigar, propor e avaliar métodos estatísticos apropriados, tanto para o processo de validação de método, como para os ensaios de proficiência na área de sementes, bem como o desenvolvimento de rotinas específicas num programa estatístico para estes tipos de análises.

Este foi o problema de pesquisa abordado neste trabalho, o qual foi estruturado em dois capítulos. No capítulo 2, é apresentado o processo de validação de método para um teste de germinação com sementes de nabo forrageiro. No capítulo 3 são abordados os procedimentos de um ensaio de proficiência utilizando dados simulados, exemplificando uma situação de testes de germinação similares às condições reais. Um resumo esquemático das etapas que foram desenvolvidas em cada um dos capítulos pode ser observado na Figura 1.

Ressalta-se que, apesar de terem sido utilizadas sementes de nabo forrageiro, o propósito final deste trabalho é que as técnicas avaliadas e selecionadas sirvam como referência metodológica para programas de validação de método e ensaios de proficiência com sementes de qualquer espécie vegetal.

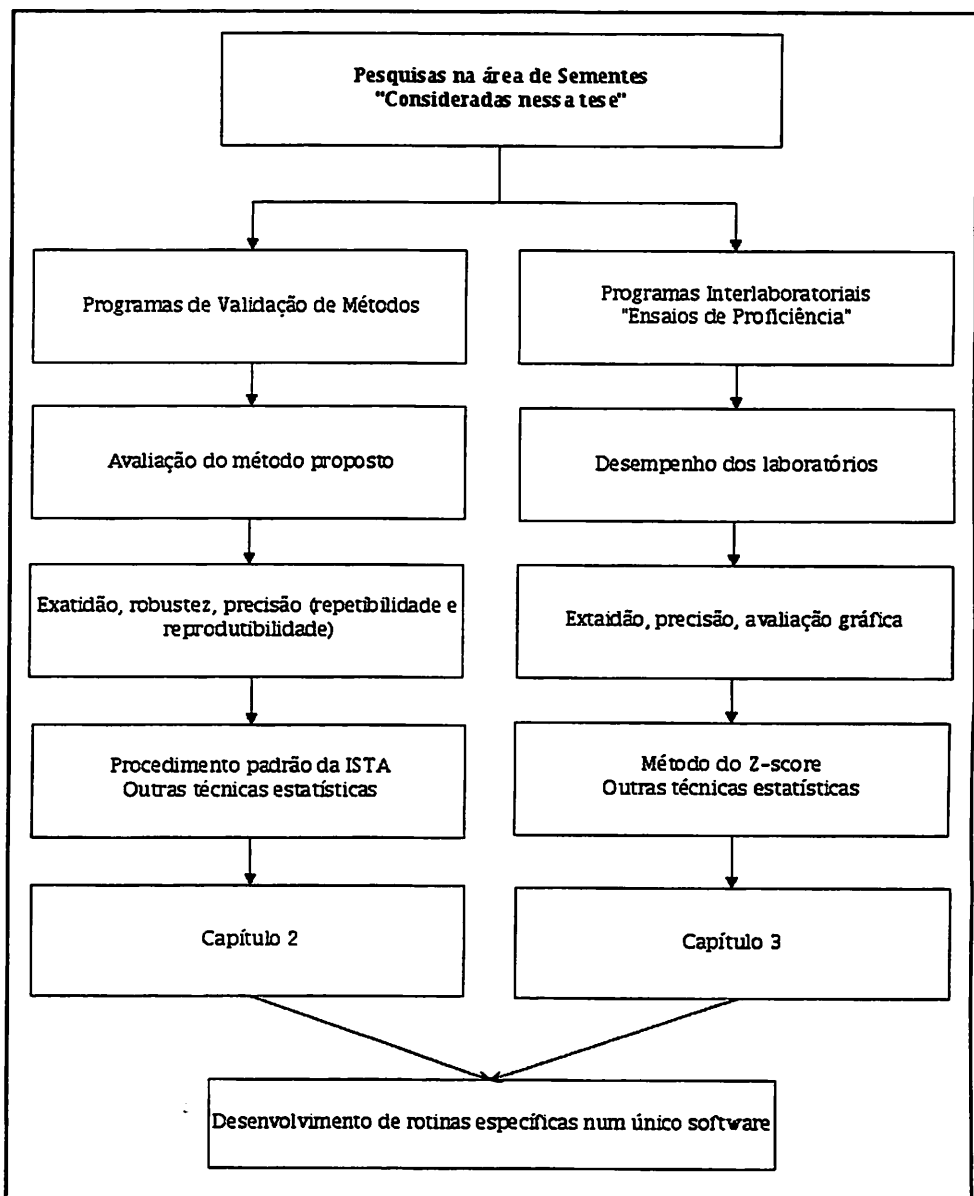


FIGURA 1 Resumo esquemático das etapas desenvolvidas nos capítulos 2 e 3

4 REFERÊNCIAS BIBLIOGRÁFICAS

MACHADO, A.A.; DEMÉTRIO, C.G.B.; FEREEIRA, D.F.; SILVA, J.G. C. da. Estatística experimental: uma abordagem fundamentada no planejamento e no uso de recursos computacionais. In: REUNIÃO ANUAL DA RBRAS, 50; SEAGRO,11., 2005, Londrina. *Anais...* Londrina: UEL, 2005. p. 290.

CAPÍTULO 2
VALIDAÇÃO DE MÉTODOS

1 RESUMO

Para várias espécies vegetais, o teste de germinação, bem como vários outros testes, não tem seus procedimentos descritos nas Regras para Análise de Sementes (RAS) utilizadas no Brasil pela falta de padronização e validação metodológica. A *International Seed Testing Association* (ISTA), para assegurar a padronização de métodos para o comércio internacional de sementes, estabeleceu critérios de validação de método e, neste processo, a descrição do método deve ser clara e completa e o procedimento deve fornecer exatidão, robustez, reprodutibilidade e repetibilidade, conseqüentemente, vários procedimentos estatísticos devem ser utilizados para analisar os resultados. O objetivo deste estudo foi propor e avaliar o uso de outras técnicas estatísticas, além do padrão ISTA, num processo de validação de métodos específicos para o teste de germinação em sementes de nabo forrageiro, utilizando os substratos areia e papel e temperatura alternada 20°-30°C. Os testes de germinação foram realizados em oito laboratórios. As técnicas estatísticas foram utilizadas para: verificar a homogeneidade dos lotes (teste H), identificar a presença de valores discrepantes (método de Hampel) e *outliers* nas variâncias (teste de Levene para média); avaliar os efeitos de laboratórios e lotes (análise de variância); verificar a repetibilidade, a reprodutibilidade, a exatidão e a robustez (limites críticos de repetibilidade, reprodutibilidade, estatísticas h e k de Mandel); comparar os diferentes métodos de germinação (testes F) e reavaliar o nível de qualidade dos lotes (análise discriminante). Tanto no aspecto estatístico como no fitotécnico, os métodos para o teste de germinação em sementes de nabo forrageiro com a utilização dos substratos areia e papel e temperatura alternada 20°-30°C podem ser consideradas validadas, pois apresentaram exatidão, robustez, repetibilidade e reprodutibilidade adequados. As técnicas estatísticas propostas, como teste de Levene, limites críticos de repetibilidade e reprodutibilidade, testes F e análise discriminante, contribuem para a melhoria e o aperfeiçoamento das análises no processo de validação de métodos.

2 ABSTRACT

For several vegetal species, the germination test, as well as many other tests, has no procedures described in the Rules for Seed Testing used in Brazil due to lack of standardization and methodological validation. The International Seed Testing Association (ISTA), to ensure the standardization of methods for the international seed trade, established criteria for the validation method and, in the process, the method description should be clear and comprehensive, and the procedure should provide for accuracy, robustness, reproducibility and repeatability, consequently, several statistical procedures shall be used to analyze the results. This study aims at proposing and evaluating the use of other statistical techniques, in addition to the ISTA standard, in a process of validation of specific methods for germination tests in oil radish seeds, using sand and paper as substrates at alternate temperatures of 20°-30°C. Germination tests were carried out in eight laboratories. The statistical techniques have been used to: verify the homogeneity of lots (H test), identify the presence of discrepant values (Hampel method) and outliers in variances (Levene's test for average); assess the effects of laboratories and lots (analysis of variance); verify repeatability, reproducibility, accuracy and robustness (critical limits of repeatability, reproducibility, Mandel's h and k statistics); compare different germination methods (F tests); and reevaluate the quality level of lots (discriminating analysis). Both in statistical and phytotechnical aspects, the methods for germination test in oil radish seeds with the use of substrates sand and paper and alternate temperatures of 20°-30°C can be considered validated, as they presented proper accuracy, robustness, repeatability and reproducibility. The proposed statistical techniques, such as the Levene's test, critical limits of repeatability and reproducibility, F tests, and discriminating analysis, contribute to improving and perfecting analyses in the process of method validation.

3 INTRODUÇÃO

Uma exigência do comércio de sementes é a produção das mesmas em quantidade e qualidade suficientes para viabilizar o cultivo de uma determinada cultura pelo produtor e, por conseguinte, possibilitar, por exemplo, a produção de grãos em larga escala.

Para averiguar os componentes de qualidade das sementes, são realizados testes, como o de germinação, sob condições ideais e artificiais que permitem a manifestação do máximo potencial de germinação (Association of official seed analysts, AOSA, 1983). Mas, para várias espécies vegetais, o teste de germinação, bem como vários outros testes, não tem seus procedimentos descritos nas Regras para Análise de Sementes (RAS) utilizadas no Brasil.

Segundo a literatura internacional (ISTA, 2007a; Sheppard & Cockerell, 2000; Eurachem, 1998), a designação de um teste ou método deve basear-se em princípios de metrologia. Dessa forma, a indicação de um determinado teste deve ser precedida de procedimentos de validação metodológica.

A International Seed Testing Association (ISTA), para assegurar a padronização de métodos para o comércio internacional de sementes, estabeleceu critérios de validação de métodos em um protocolo elaborado em 2002. Este programa consiste de uma avaliação crítica para assegurar que a descrição do método seja clara e completa, e que o procedimento forneça exatidão, robustez, reprodutibilidade e repetibilidade dos resultados em concordância com as especificações estabelecidas pelo próprio método. Quando apropriado, confirma a relação entre os resultados de um teste de qualidade e a expressão prática da qualidade fisiológica da semente.

Para a avaliação do processo de validação de métodos, já existe uma recomendação da ISTA para o desenvolvimento de uma análise padrão que

considera a utilização de algumas técnicas estatísticas, mas que não restringe a proposição de novos procedimentos (ISTA, 2007a).

O problema de pesquisa para esta etapa do estudo foi definir, primeiramente, uma cultura importante no cenário agrícola nacional de produção e que ainda não tivesse procedimentos específicos de testes descritos na RAS para avaliar a qualidade das suas sementes. Segundo, para a espécie em questão, definir qual ou quais métodos deveriam ser avaliados. Por fim, verificar se outras técnicas estatísticas, além das recomendadas pela ISTA, poderiam contribuir para a melhoria e/ou o aperfeiçoamento das análises do programa de validação de métodos.

4 REFERENCIAL TEÓRICO

Nesta seção são apresentadas, em linhas gerais, as duas etapas propostas neste estudo (Figuras 2 e 3) para o desenvolvimento de um programa de validação de métodos na área de sementes, bem como os conceitos envolvidos neste processo.

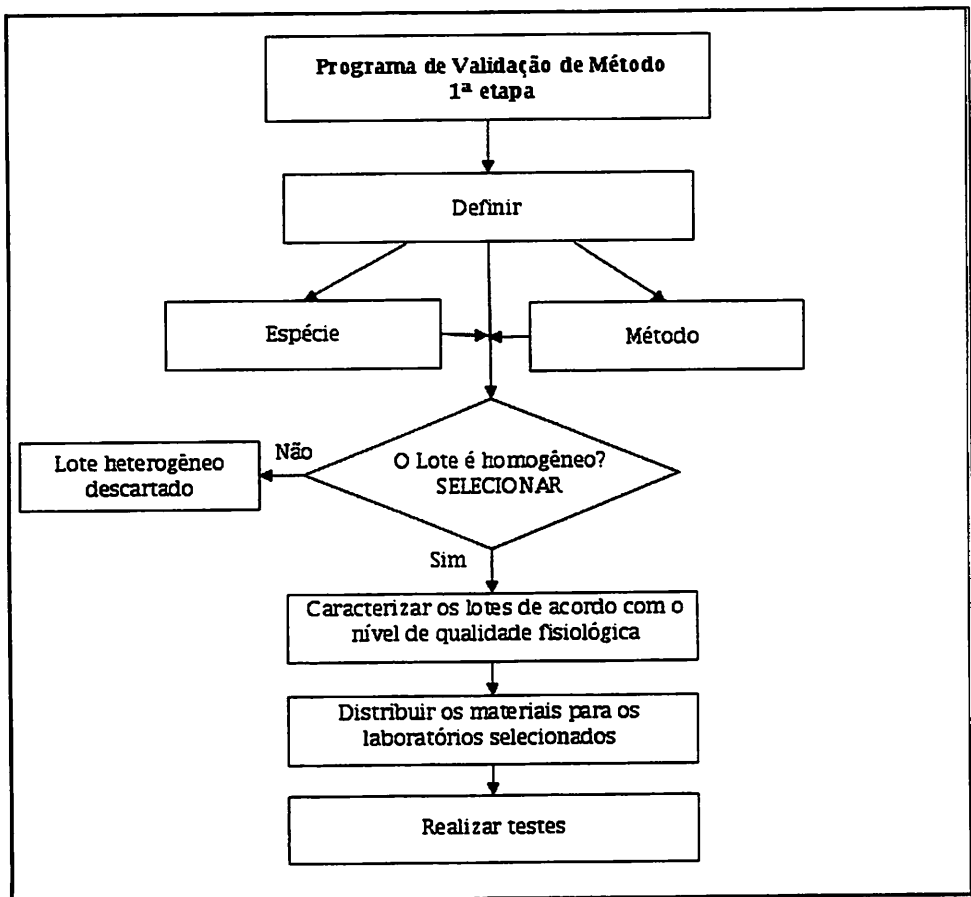


FIGURA 2 Esquema da 1ª etapa do programa de validação de métodos na área de sementes: procedimentos iniciais

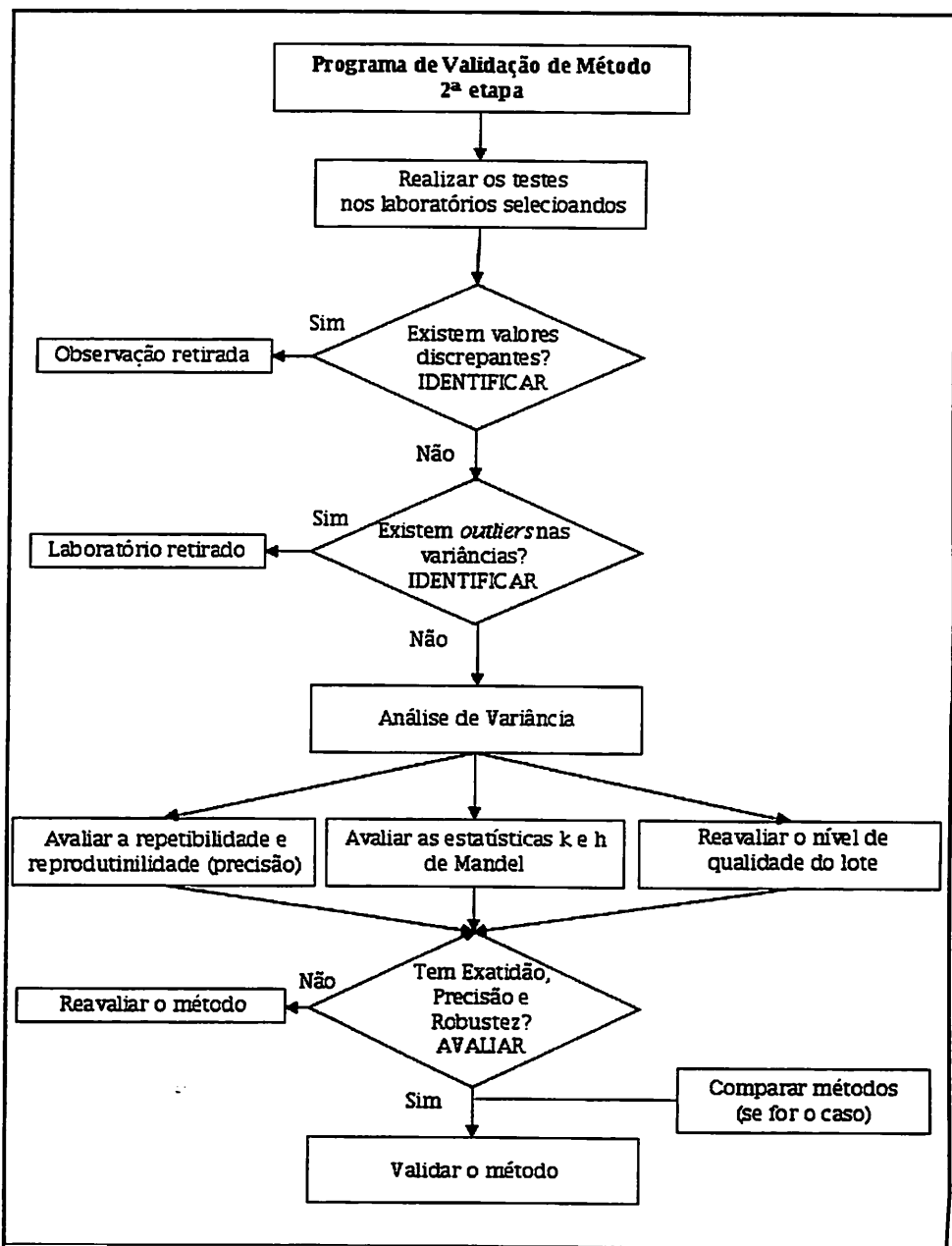


FIGURA 3 Esquema da 2ª etapa do programa de validação de métodos na área de sementes: avaliação dos resultados

As explicações sobre os elementos dos dois esquemas são dadas nas seções que se seguem.

4.1 Programas de validação de métodos: aspectos gerais

De acordo com a Associação Brasileira de Normas Técnicas, ABNT (2000), validação é a comprovação, por meio da apresentação de evidência objetiva, de que os requisitos para uma aplicação ou uso específicos pretendidos foram atendidos.

Esta definição pode ser interpretada para o método de validação em sementes como sendo uma investigação crítica de um teste de qualidade para assegurar que a descrição do método é clara e completa e que o procedimento fornece exatidão, precisão (reprodutibilidade e repetibilidade) e robustez (ISTA, 2007a).

O método de validação de um teste de qualidade de sementes é importante para a determinação de um resultado correto e que seja capaz de mostrar que é “correto”, dentro de limites de variação executáveis e aceitáveis. Importantes decisões poderão ser tomadas com base nesses resultados. Além disso, o custo de execução de um teste deve ser considerado quando da sua indicação.

No caso específico da área de sementes, a ISTA é a organização internacional responsável pela oficialização de métodos de testes de qualidade e pela promoção de seu uso seguro, conferindo, assim, confiança aos resultados e certificações.

Até 2002, não existia um sistema consistente para a incorporação dos métodos de teste em sementes nas regras da ISTA e cada comitê técnico tinha seu próprio procedimento.

Em 2002, o comitê executivo deste órgão elaborou um programa de validação de método único, tendo como objetivo fornecer um sistema que

permite validar novos métodos para teste em sementes, comparar métodos equivalentes e verificar a manutenção (revisão) de métodos já existentes para avaliação da qualidade da semente (ISTA, 2007a).

Salienta-se que, no caso deste trabalho, o programa de validação foi utilizado no intuito de validar novos métodos para teste de qualidade de sementes.

O programa de validação da ISTA vem sendo desenvolvido por meio de estudos colaborativos entre laboratórios (validação multilaboratorial colaborativa) para assegurar que o procedimento do teste em sementes forneça resultados confiáveis e reprodutíveis, de acordo com as especificações estabelecidas pelo método de teste. Existem alguns critérios para selecionar os laboratórios participantes (ISTA, 2007a). São eles:

- somente laboratórios com experiência em análise de teste de qualidade de semente ou aplicação das técnicas específicas a serem avaliadas no teste comparativo deverão ser convidados a participar do programa;
- estes laboratórios devem ter profissionais qualificados, instalações apropriadas e equipamentos recomendados para executar o teste em sementes; devem entender e aceitar as suas obrigações ao participar de um programa de validação de métodos;
- em casos nos quais exista alguma dúvida sobre a habilidade de um laboratório em executar o teste de qualidade de sementes, um pré-teste pode ser utilizado. Com base nos resultados, o organizador de teste decidirá quais laboratórios podem ser envolvidos no ensaio comparativo.

O processo completo de validação da ISTA é desenvolvido em cinco fases: seleção e desenvolvimento do método; validação por meio de testes comparativos; revisão dos resultados dos testes comparativos e preparação de

um relatório final; aprovação e aceitação final dos métodos pelos membros da ISTA e posterior publicação nas Regras Internacionais para Análise de Sementes.

No processo de validação de métodos proposto neste trabalho, somente as três primeiras fases citadas foram contempladas, já que nenhuma documentação foi enviada para a ISTA, pois não era este o foco principal do estudo.

Como apresentado no esquema da Figura 1, a 1ª etapa para o programa de validação proposta neste trabalho foi composta de alguns procedimentos iniciais, tendo como 1ª subetapa a definição da espécie a ser estudada.

Neste estudo, foi escolhido o nabo forrageiro (*Raphanus sativus* L. var. *oleiferus* Metzg.), já que não existem procedimentos específicos para teste de germinação para esta cultura nas RAS. O que vem sendo utilizado atualmente são métodos descritos para o teste de germinação para *Raphanus sativus*.

4.2 Importância da cultura do nabo forrageiro

A justificativa para o desenvolvimento de estudos mais aprofundados com o nabo forrageiro (*Raphanus sativus* L. var. *oleiferus* Metzg.) não decorre apenas do fato de não haver procedimentos específicos de teste descritos na RAS, mas porque as suas sementes têm se revelado uma opção promissora como matéria-prima para a produção do biodiesel¹, por possuir alto rendimento em óleo de baixa viscosidade, com teor médio estimado de 35%.

Vieira (2007) observa, ainda, que o óleo do nabo forrageiro é de fácil extração, mas o principal fator que atesta a qualidade dessa matéria-prima é a

¹ Segundo a Lei nº 11.097, de 13 de janeiro de 2005, biodiesel é um biocombustível para uso em motores a combustão interna com ignição por compressão ou, conforme o regulamento, para geração de outro tipo de energia, que possa substituir parcial ou totalmente combustíveis de origem fóssil (Brasil, 2005).

elevada presença de triglicerídeos no óleo, os quais vão dar origem a ésteres alquílicos de ácidos graxos e glicerina. São os ésteres que constituem o biodiesel.

Desde 2002, pesquisas vêm sendo realizadas no intuito de otimizar o aproveitamento das sementes de nabo forrageiro para a produção de biodiesel. Atualmente, essas pesquisas buscam, principalmente, o melhoramento genético dessa planta para a obtenção de uma maior produtividade, já que o rendimento médio é de apenas 1.000 kg de grão por hectare, sendo este ainda um fator limitante para seu uso em larga escala.

Esta cultura está incluída, principalmente, nos programas estaduais de biodiesel dos estados do Mato Grosso do Sul e do Paraná. Em Minas Gerais, apesar de não estar incluída como prioritária para pesquisas, foi inaugurada em Cássia, no ano de 2005, a primeira indústria de biodiesel no Brasil (Sol e Minas), produzido com sementes de nabo forrageiro.

A utilização de biodiesel no Brasil vem sendo incentivada por projetos definidos pelo governo federal, por ser uma alternativa viável para a redução de importação do óleo diesel que gera despesas anuais na ordem de US\$1,2 bilhão para o país. Existem, ainda, outros aspectos relevantes a favor da introdução do biodiesel, como a diminuição da poluição do ar, a sustentação de um grande programa de geração de emprego e renda com a participação da agricultura familiar, além da possibilidade de geração de excedentes de biocombustível exportáveis para outros países (Holanda, 2004).

Vale ressaltar que, além dos benefícios ambientais, a produção de biodiesel pode, ainda, propiciar vantagens econômicas e políticas ao Brasil pelo cumprimento dos acordos estabelecidos no Protocolo de Quioto e nas diretrizes do mecanismo de desenvolvimento limpo (MDL), pela geração de cotas de carbono e pelo fortalecimento de sua situação como um dos países de matriz energética limpa (Holanda, 2004).

Segundo Foster (2004), a diretriz do programa do governo federal é produzir biodiesel no Brasil de forma economicamente sustentável, fazendo a inclusão do pequeno produtor do campo, promovendo a sustentação social. Nesse sentido, o nabo forrageiro parece ser bem apropriado, já que seu custo de produção é baixo em relação a outras culturas, como a da soja ou a do girassol, e o seu teor de óleo é tão elevado quanto ao dessas culturas.

De acordo com Vieira (2007), a planta de nabo forrageiro se adapta bem a diferentes condições edafoclimáticas, fornece massa foliar para adubação verde, é indicada para a rotação de culturas, como cobertura do solo durante o inverno e, eventualmente, para compor a alimentação animal, além de não competir como matéria-prima na alimentação humana.

O desafio atual é a produção de sementes de nabo forrageiro em quantidade e qualidade suficientes para possibilitar um aumento na produtividade dessa cultura e, por conseguinte, viabilizar sua utilização como matéria-prima competitiva e rentável para produção de biodiesel.

Para o comércio de sementes, o resultado de um teste de germinação é o principal indicativo da qualidade fisiológica da semente, sendo considerado um dos testes obrigatórios a serem realizados.

No caso do nabo forrageiro, o método para o teste de germinação que vem sendo utilizado é o mesmo empregado para a espécie *Raphanus sativus*, não existindo um procedimento específico para variedade *Oleiferus*. Então, a proposta deste trabalho foi avaliar o processo de validação de métodos específicos para o teste de germinação em sementes desta cultura.

4.3 Testes de germinação

A germinação é um fenômeno biológico que pode ser considerado, em análise de sementes, como a emergência e o desenvolvimento das estruturas essenciais do embrião, manifestando a sua capacidade para dar origem a uma

planta normal, sob condições ambientais favoráveis (Instituto de Pesquisas e Estudos Florestais, IPEF, 1998).

A avaliação da germinação das sementes é efetuada pelo teste de germinação, conduzido em laboratório sob condições controladas e por meio de métodos padronizados que visam, principalmente, avaliar o valor dessas sementes para a semeadura e comparar a qualidade de diferentes lotes, servindo como base para a comercialização (Novembre, 1994; Marcos Filho et al., 2005).

O conhecimento das condições ideais para a germinação da semente de uma determinada espécie é de fundamental importância, principalmente pelas respostas diferenciadas que ela pode apresentar em função de diversos fatores, como viabilidade, dormência, condições de ambiente, envolvendo água, luz, temperatura, oxigênio e ausência de agentes patogênicos, associados ao tipo de substrato para a sua germinação (Brasil, 1992; Bewley & Black, 1994; Carvalho & Nakagawa, 2000).

Os substratos utilizados no teste de germinação, em geral, têm como principal função dar sustentação às sementes, tanto do ponto de vista físico como químico (Aguiar et al., 1993). Em função do tamanho e das exigências ecofisiológicas das sementes quanto à umidade e à luz, cada substrato é utilizado de maneira a oferecer maior praticidade nas contagens e na avaliação das plântulas (Popinigs, 1985).

Nery et al. (2007) e Nery (2008) realizaram testes de germinação para sementes de nabo forrageiro, visando definir a temperatura ótima e o substrato ideal para avaliação da qualidade de sementes das cultivares CATI AL-1000 e IPR 116. Estes autores verificaram que a temperatura alternada de 20°-30°C e uso do substrato areia seriam as condições ideais.

Apesar de essas duas pesquisas já apresentarem indicativos de um método adequado para o teste de germinação em sementes de nabo forrageiro,

como dito, não existe ainda uma recomendação específica, descrita nas RAS, para a variedade *Oleiferus*.

Foram avaliados, no programa de validação deste trabalho, dois métodos para o teste de germinação, um utilizando como substrato a areia, e o outro o papel, com temperatura alternada de 20°-30°C (Figura 4).

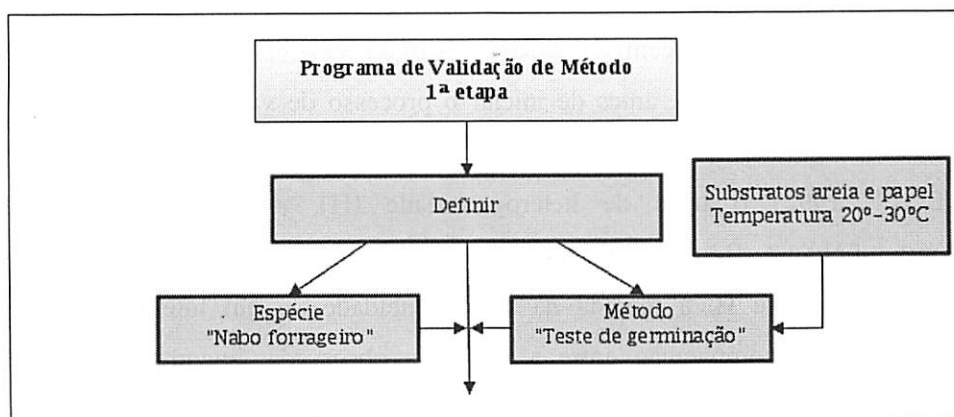


FIGURA 4 Esquema da 1ª etapa do programa de validação de métodos, com ênfase na subetapa definição da espécie e do método

Definidos a espécie a ser estudada e os métodos a serem validadas, a subetapa seguinte de um programa de validação é verificar se os lotes coletados para o processo podem ser considerados homogêneos. Para isso, são realizados testes correspondentes ao escopo em estudo (teste de germinação, no caso deste trabalho) e, com os resultados, aplica-se o teste de heterogeneidade (teste H).

4.4 Seleção de lotes de acordo com a homogeneidade – Teste H

A homogeneidade dos lotes e a representatividade da amostra constituem princípios básicos da amostragem para análise de sementes (Copeland & Liu, 2005).

A exigência para tal homogeneidade é a de que um lote de sementes é caracterizado pelo conjunto de todos os seus componentes. Assim sendo, se um lote for perfeitamente homogêneo, todos os seus componentes estarão uniformemente distribuídos em todas as suas partes. Dessa forma, uma amostra tomada em um ponto seria idêntica à outra tomada em qualquer outro ponto, no que se refere às diferentes características submetidas à análise (Brasil, 1992). Isto é, sendo um lote considerado como homogêneo, garante-se a aleatoriedade no processo de amostragem.

Neste contexto, antes de iniciar o processo de validação, é necessário selecionar os lotes que sejam mais homogêneos, sendo recomendado pela RAS (Brasil, 1992) o teste de heterogeneidade (H), desenvolvido por J.G. Tarttersfield e A. Bould, na década de 1970.

No teste H, a medida da heterogeneidade de um lote de sementes envolve uma comparação entre a variância observada, encontrada entre as amostras, e a variância teórica, se a distribuição for ao acaso. De fato, no cálculo, é utilizada a variância aceitável, que é obtida pela multiplicação da variância teórica pelo um fator “f” de variação adicional. No caso do teste de germinação, a variância teórica é calculada utilizando-se a distribuição binomial (Figura 5).

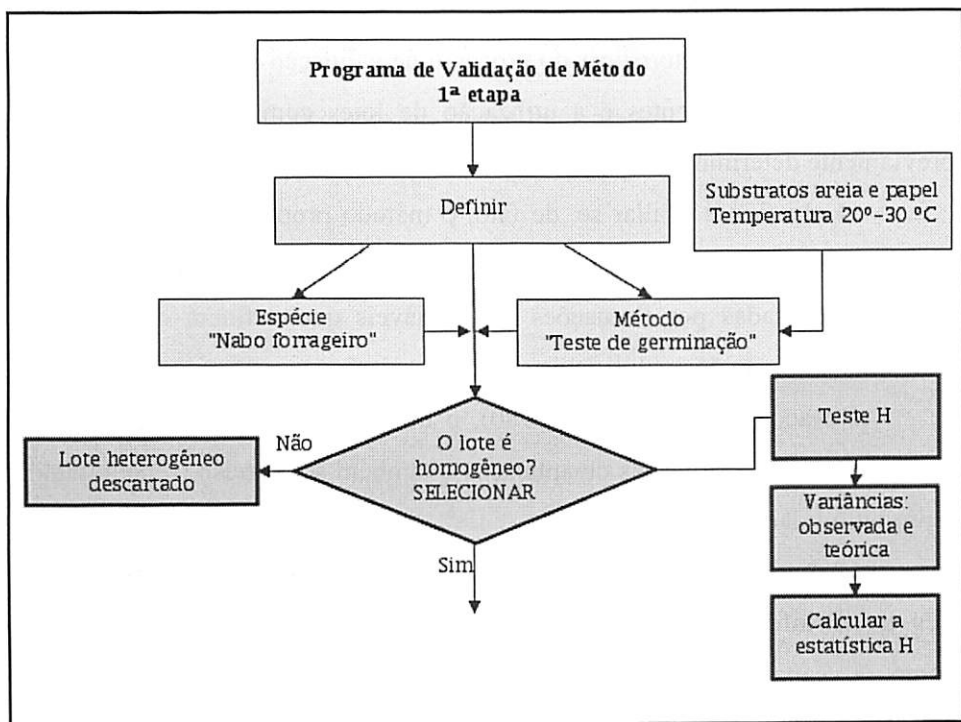


FIGURA 5 Esquema da 1ª etapa do programa de validação de métodos, com ênfase na subetapa seleção de lotes

Bould (1975) computou os valores de H críticos com base na distribuição χ^2 . As regras da ISTA incluem estes valores críticos de H na tomada de decisões no controle de qualidade de um lote de sementes (ISTA, 2007b).

No programa de validação, tão importante quanto trabalhar com lotes homogêneos é assegurar-se de que eles tenham níveis de qualidade fisiológica diferentes, determinando, assim, a 4ª subetapa, que é a caracterização dos lotes.

4.5 Caracterização dos lotes

Uma das características do processo de validação de métodos para testes de qualidade de sementes é a utilização de lotes com qualidade diferentes, previamente determinados antes do início do estudo (ISTA, 2007a).

O objetivo é avaliar se, de fato, o método proposto será robusto sob a condição de lotes com qualidade diferentes, isto é, se será pouco sensível a variações causadas por flutuações nas variáveis que definem e controlam o desenrolar do teste de qualidade de sementes.

De acordo com Waeny (1980), o método ideal deve ser robusto para suportar condições adversas de aplicação e também ser preciso (repetibilidade e reprodutibilidade) e exato.

Os lotes podem ser classificados, quanto à sua qualidade fisiológica, em três níveis: inferior, intermediário e superior. A qualidade fisiológica é definida como a capacidade das sementes para desempenharem funções vitais, caracterizadas pela sua germinação, vigor e longevidade (Popinigis, 1985) e tem sido avaliada rotineiramente por meio do teste de germinação (Figura 6).

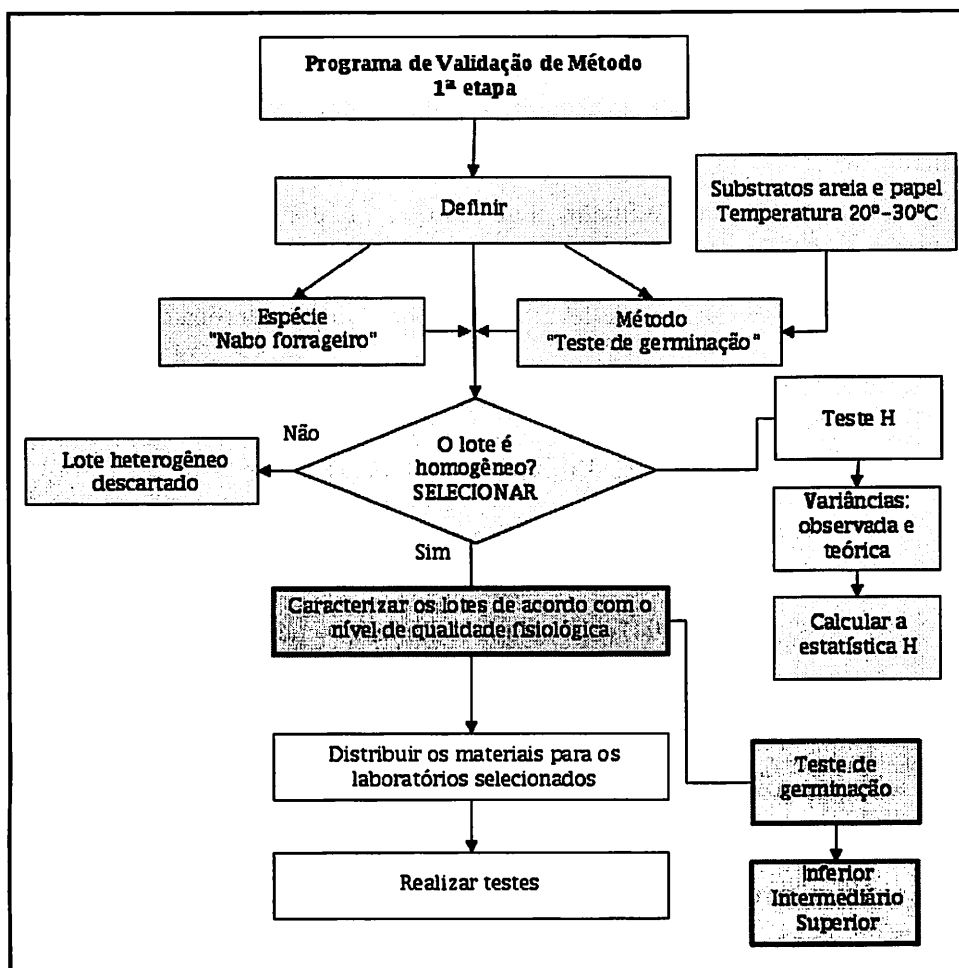


FIGURA 6 Esquema da 1ª etapa do programa de validação de métodos, com ênfase na subetapa caracterização dos lotes

Maiores detalhes sobre o procedimento de distribuição do material para os laboratórios selecionados (5ª subetapa), ainda desta 1ª etapa, podem ser observados na subseção 5.3.

Na 2ª etapa do programa de validação (Figura 3), os resultados dos laboratórios são avaliados de acordo com várias técnicas estatísticas, sendo a 1ª

subetapa do processo a investigação de *valores discrepantes*. A orientação do programa de validação, nos casos em que for detectado algum outlier, é promover a sua retirada do banco de dados, antes de proceder às próximas análises (ISTA, 2007 a).

De acordo com Williams et al. (2002) e Liu et al. (2004), é importante a identificação dos valores discrepantes antes da modelagem e análise, pois a presença desses valores pode levar a um modelo mal especificado, à estimação viesada de parâmetros e, por conseguinte, a resultados incorretos.

4.6 Identificação de valores discrepantes (*outliers*)

Podem ser considerados *outliers* aqueles elementos que não obedecem a um padrão do conjunto de dados ao qual eles pertencem.

Barnett & Lewis (1994) relatam que as fontes potenciais de *outliers* são:

- variabilidade inerente: variação natural das observações sobre o grupo, eventos inesperados durante o processo de gerações dos dados;
- erros de medição: instrumentos de medição inadequados, erros na transcrição dos dados, arredondamentos nos valores obtidos;
- erro praticado: variabilidade devido à coleta imperfeita dos dados, como, por exemplo, a escolha de uma amostra viesada.

Para identificar a presença valores discrepantes, podem ser utilizados diversos testes estatísticos e técnicas gráficas. A ISTA (2007a) utiliza, nos seus programas de validação de métodos para teste de qualidade de sementes, o teste de Hampel e o Boxplot.

O método de Hampel (Hampel, 1985) é uma técnica robusta para a detecção de *outlier* e que troca as estimativas da média e do desvio padrão, sensíveis à presença de valores discrepantes, respectivamente, pela mediana e o desvio absoluto mediano, robusto aos *outliers*.

O Boxplot (Tukey, 1977) é um tipo de gráfico que sumariza medidas estatísticas, como: mediana, quartis inferior e superior, valores máximo e mínimo, além de dar uma ideia da posição, da dispersão, da assimetria e da presença de valores discrepantes, permitindo também confrontar variáveis da mesma natureza em uma mesma escala (Bussab & Morettin, 2003).

Neste estudo, optou-se pelo uso do método de Hampel, para a retirada dos *outliers* em detrimento do Boxplot.

A decisão pela utilização do método de Hampel foi baseada no conceito de *breakdown point* (Hampel, 1974; Rousseeuw & Croux, 1993; Gnanadesikan, 1997; Carling, 2000; Cunha, 2002; Reiman et al., 2005; Davies & Gather, 2007; Hellerstein, 2008).

Dada uma amostra aleatória X_1, \dots, X_n e um estimador $\hat{\theta}(X_1, \dots, X_n)$ de um parâmetro θ , o *breakdown point* é a maior fração da amostra X_1, \dots, X_n em que os elementos X_i podem assumir valores extremos sem afetar a robustez do valor de $\hat{\theta}(X_1, \dots, X_n)$. Por exemplo, para a média amostral, o *breakdown point*, considerando uma amostra finita, é igual a $1/n$, pois, se os valores de x_1, \dots, x_{n-1} forem fixados e x_n assumir um valor tendendo ao infinito, a média amostral também tenderá para o infinito. Ou seja, para a média amostral, basta que um valor seja extremo para afetar a robustez da estimativa. Logo, assintoticamente, *breakdown point* para a média amostral é igual a zero $\left(\lim_{n \rightarrow \infty} \frac{1}{n} = 0 \right)$.

De acordo com Hampel (1974), Rousseeuw & Croux (1993), o desvio absoluto mediano $\left(MAD = \text{mediana} \left\{ \left| Y_i - \tilde{Y} \right| \right\} \right)$ é um estimador de escala mais

robusto² do que a amplitude interquartilica $(\hat{\delta} = (\hat{Y}_{0,25} - \hat{Y}_{0,75}))$, já que os *breakdown points* assintoticamente são iguais a 50% e 25%; respectivamente.

O *breakdown point* do desvio absoluto mediano é assintoticamente igual a 50% porque, considerando, uma amostra finita de tamanho n , se a minoria das observações $h((n-1)/2)$, em que em que $h(\cdot)$ representa o maior inteiro menor que o argumento, tenderem para o infinito e o resto das observações tiver um valor fixado, então, a mediana permanece com a maioria. Assim, o *breakdown point* para uma amostra finita é $h[(n-1)/2n]$ e assintoticamente igual a $\frac{1}{2}$, pois

$$\left(\lim_{n \rightarrow \infty} \frac{(n-1)}{2n} = \frac{1}{2} \right).$$

Os métodos diferem também na suas regras de decisão. No caso do método de Hampel, uma observação (Y_i) é considerada *outlier* se o desvio absoluto em relação à mediana (\tilde{Y}) for maior que $5,2MAD$, em que $MAD = \text{mediana} \{ |Y_i - \tilde{Y}| \}$ (para o esquema dessa subetapa ver Figura 7).

² A grosso modo, um estimador é considerado robusto, se for “insensível” a desvios de pressuposições básicas, por exemplo, mesmo na presença de *outliers* o seu resultado não é afetado de maneira significativa.

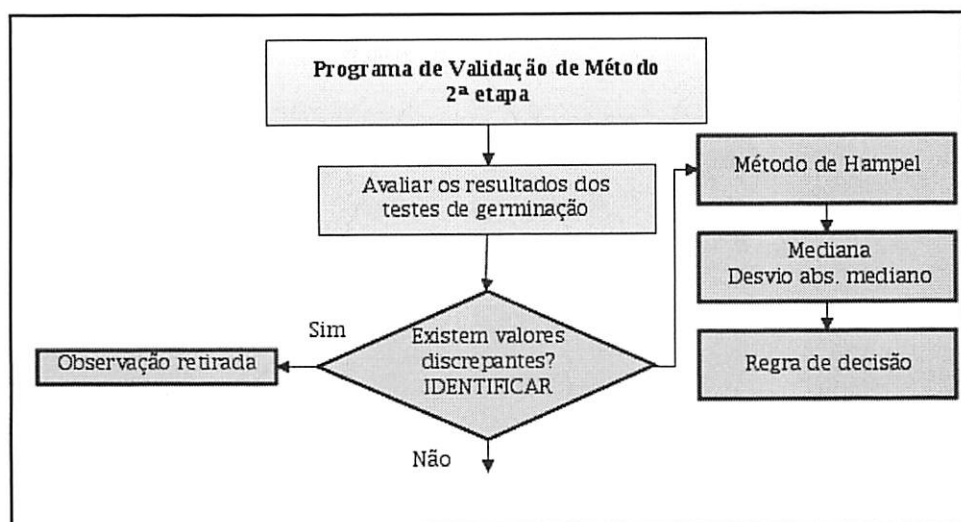


FIGURA 7 Esquema da 2ª etapa do programa de validação de métodos, com ênfase na subetapa identificação de valores discrepantes

No caso do Boxplot, o valor absoluto da observação deve ser maior que 1,5 vez a amplitude interquartilica $(\hat{Y}_{0,25} - \hat{Y}_{0,75})$, isto é, $|Y_i| > 1,5(\hat{Y}_{0,25} - \hat{Y}_{0,75})$.

Fazendo um paralelo entre essas duas técnicas, se os dados seguirem uma distribuição normal padrão, tem-se que $5,2MAD$ correspondem a aproximadamente $3,5\sigma$ (prova i), e que $1,5(Y_{0,25} - Y_{0,75})$ é equivalente a $2,2\sigma$ (prova ii) (Figura 8).

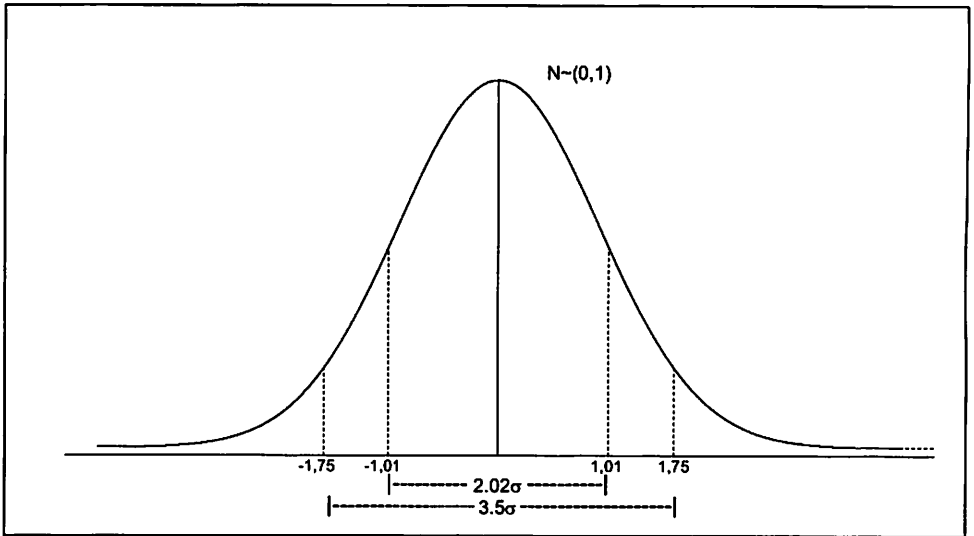


FIGURA 8 Curva da Normal Padrão, com delimitação das regiões correspondentes a $2,2\sigma$ e $3,5\sigma$

Prova (i)

$$P(|Y - \mu| \leq MAD) = \frac{1}{2} = P\left(\left|\frac{Y - \mu}{\sigma}\right| \leq \frac{MAD}{\sigma}\right) \Rightarrow \frac{1}{2} = P\left(|Z| \leq \frac{MAD}{\sigma}\right)$$

Para uma distribuição simétrica, a MAD é a distância entre o 1º e 2º quartis (equivalentemente, 2º e 3º quartis). Então, o fator escala para o uso da MAD para uma distribuição normal padrão é o 75º percentil, logo:

$$\frac{MAD}{\sigma} = \Phi^{-1}\left(\frac{3}{4}\right) = 0,6745 \Rightarrow \sigma \cong 1,4826MAD$$

Num paralelo com a regra de decisão do método de Hampel, tem-se:

$$|Y_i - \tilde{Y}| > 5,2MAD \Rightarrow |Y_i - \tilde{Y}| > 5,2 \frac{\hat{\sigma}}{1,4826} \Rightarrow |Y_i - \tilde{Y}| > 3,5\hat{\sigma}$$

Prova (ii)

Para demonstrar que $1,5(Y_{0,25} - Y_{0,75})$ é equivalente a, aproximadamente, $2,02\sigma$, deve-se levar em consideração que

$$\begin{aligned} P(|Y - \mu| \leq (Y_{0,25} - Y_{0,75})) &= \frac{1}{2} = P\left(\left|\frac{Y - \mu}{\sigma}\right| \leq \frac{(Y_{0,25} - Y_{0,75})}{\sigma}\right) \Rightarrow \\ \frac{1}{2} &= P\left(|Z| \leq \frac{(Y_{0,25} - Y_{0,75})}{\sigma}\right) \\ \frac{(Y_{0,25} - Y_{0,75})}{\sigma} &= 2\Phi^{-1}\left(\frac{3}{4}\right) = 2,0,6745 \Rightarrow \sigma \cong 0,7413(Y_{0,25} - Y_{0,75}) \end{aligned}$$

Num paralelo com a regra de decisão do Boxplot, tem-se:

$$|Y_i| > 1,5(Y_{0,25} - Y_{0,75}) \Rightarrow |Y_i| > 1,5 \frac{\sigma}{0,7413} \Rightarrow |Y_i| > 2,02\sigma$$

Reimann et al. (2005) investigaram três procedimentos para a detecção de outliers, o Boxplot, mediana $\pm 2MAD$ e a média $\pm 2\sigma$ e concluíram, por meio de simulações, que a função Boxplot era mais informativa se o número de outliers estivesse abaixo de 10% do total de observações. Quando esta proporção foi superior a 15%, somente a mediana $\pm 2MAD$ teve desempenho satisfatório.

Após a investigação de valores discrepantes, a próxima subetapa do processo de validação consiste em avaliar se a variância de algum laboratório é discrepante em relação aos demais e, caso seja identificado como sendo um outlier, recomenda-se a retirada do laboratório para as próximas etapas (ISTA, 2007a).

4.7 Identificação de *outliers* nas variâncias

Quanto às variâncias, um dos pressupostos fundamentais assumidos para o bom desempenho dos testes paramétricos na avaliação de resultados provenientes de diferentes fontes é considerar que as mesmas são homogêneas, a ponto de não ultrapassarem determinados limites de tolerância. Tais limites são, via de regra, com tolerância zero, isto é, variâncias homogêneas são variâncias iguais.

De acordo com as recomendações da ISTA (2007a), o teste de Cochran (Cochran, 1941) deve ser utilizado para verificar a hipótese nula de homogeneidade (igualdade) entre variâncias dos laboratórios, sendo a sua estatística expressa por:

$$g = \frac{\max(S_l^2)}{\sum_{l=1}^p S_l^2}, \quad (1)$$

em que p é o número de laboratórios e S_l^2 é o estimador da variância do l -ésimo laboratório.

Mas, observando-se os resultados do Capítulo 2, especificamente o lote 3, para os dois métodos, areia e papel, verificou-se haver certa restrição no uso deste teste³, que será explicitado posteriormente (Tabela 1). Foi então realizado um estudo investigativo para avaliar a adequabilidade de outros três testes: Levene para a média, designado de teste 2 (Levene, 1960); Bartlett - teste 3 (Bartlett, 1937) e Levene para mediana - teste 4 (Brown & Forsythe, 1974).

³ Este resultado está sendo apresentado antecipadamente nesta seção, com o objetivo de auxiliar a discussão sobre a escolha do teste para a verificação de homogeneidade das variâncias.

TABELA 1 Variâncias dos laboratórios e valor p dos testes 1 (Cochran), 2 (Levene para média), 3 (Bartlett) e 4 (Levene para mediana) para os resultados do teste de germinação em sementes de nabo forrageiro do lote 3, para os substratos areia e papel

Substrato	Lab.	s^2	Teste (valor p)							
			1	2	3	4				
Areia	1	4,50								
	2	2,33								
	3	0,00								
	4	8,33	p = 0,8897	p = 0,0099	-----	p = 0,8953				
	5	1,00								
	6	4,50								
	7	8,33								
	8	0,50								
<hr/>										
Papel	1	6,33								
	2	1,00								
	3	0,33								
	4	1,00	p = 0,0180	p = 0,0073	p = 0,2390	p = 0,0423				
	5	2,33								
	6	24,50								
	8	2,33								

De acordo com os resultados, observa-se que, pelo teste de Cochran, a 1% de significância, não há evidências para rejeitar a hipótese de homogeneidade das variâncias para os dois métodos. O mesmo resultado foi encontrado para os testes de Bartlett e Levene modificado para a mediana.

Pelos resultados, parece haver alguma condição de restrição no uso destes testes, uma vez que a maior variância é igual a 8,33 vezes a menor variância, no caso do substrato areia e 74,24 vezes, no caso do substrato papel.

Vale salientar que, para o teste de Bartlett, no caso do substrato areia, a estatística χ^2 do teste (expressão 2) nem foi calculada, já que o laboratório 3 tem variância igual a zero e, conseqüentemente, $\lim_{s_i^2 \rightarrow 0} \ln(S_3^2) = +\infty$:

$$\chi^2 = \frac{(N-p)\ln(S_p^2) - \sum_{l=1}^p (n_l - 1)\ln(S_l^2)}{1 + \frac{1}{3(p-1)} \left(\sum_{l=1}^p \left(\frac{1}{n_l - 1} \right) - \frac{1}{N-p} \right)}, \quad (2)$$

em que: $N = \sum_{l=1}^p n_l$, sendo n_l o número de repetições do l -ésimo laboratório;

$S_p^2 = \frac{1}{N-p} \sum_{l=1}^p (n_l - 1)S_l^2$ é o estimador da variância combinada, sendo S_l^2 o

estimador da variância do l -ésimo laboratório.

Vorapongsathorn et al. (2004) e Almeida et al. (2008) afirmam que os testes de Cochran e Bartlett são sensíveis a desvios de normalidade. De acordo com Hines & O'Hara Hines (2000), o teste de Levene modificado para a mediana (veja sua estatística W de teste na fórmula 3) é uma alternativa robusta com relação a desvios de normalidade, mas pode tornar-se ineficiente para detectar desigualdade de variâncias na presença de zeros estruturais dado por:

$$W = \frac{(N-p) \sum_{l=1}^p n_l (\bar{Z}_l - \bar{Z}_{..})^2}{(p-1) \sum_{l=1}^p \sum_{k=1}^{n_l} (Z_{lk} - \bar{Z}_l)^2}, \quad (3)$$

em que $Z_{lk} = |Y_{lk} - \bar{Y}_l|$, sendo Y_{lk} a porcentagem de germinação da k -ésima repetição no l -ésimo laboratório e \bar{Y}_l o estimador da mediana do l -ésimo laboratório.

Os zeros estruturais correspondem a valores de Z_{lk} nulos, que sempre existem para n_l ímpar, pois, nesse caso, a mediana coincide com um dos valores de Y_{lk} (Almeida et al., 2008).

O teste de Levene para média foi o único que rejeitou a hipótese nula para os dois substratos. De acordo com Vorapongsathorn et al. (2004), para uma grande variedade de distribuições, este teste é robusto.

A única diferença na estatística do teste de Levene para média em relação à mediana é que, na determinação de Z_{jk} , utiliza-se justamente o estimador da média (\bar{Y}_i) em vez da mediana (\tilde{Y}_i). Segundo Almeida et al. (2008), a estatística do teste de Levene é obtida a partir de uma análise de variância com um único fator, já que os níveis são os grupos e cada observação é k substituída pelo desvio absoluto da variável em relação à média do seu respectivo grupo (Figura 9).

Por essa análise preliminar dos quatro métodos, no contexto deste estudo, parece que o teste mais adequado foi o de Levene para média, já que foram detectados desvios de normalidade nos resultados de alguns laboratórios (teste de Shapiro-Wilk) e a presença de zeros estruturais (na maioria dos laboratórios, o número de repetições era ímpar).

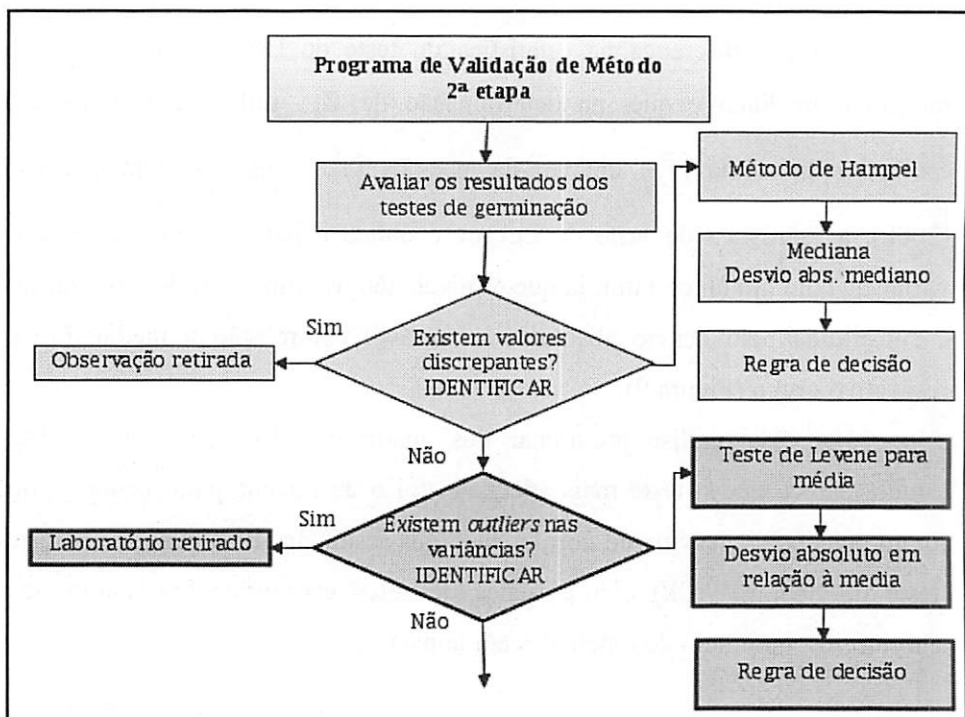


FIGURA 9 Esquema da 2ª etapa do programa de validação de métodos, com ênfase na subetapa identificação de *outliers* nas variâncias

4.8 Avaliação da repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel

Numa mesma medição podem ocorrer dois tipos de erros. Pode-se ter o erro aleatório, que é a componente do erro da medição que varia de uma forma imprevisível (variações inerentes ao objeto, ao instrumento ou observador) quando se efetuam várias medições da mesma grandeza (Vocabulário Internacional de Metrologia, VIM, 2000), ou, ainda, é um erro tal que os n resultados y_i se distribuem de maneira aleatória em torno valor verdadeiro y_v (na ausência de erro sistemático) e, quando $n \rightarrow \infty$, o valor médio $\bar{y} \rightarrow y_v$ (lei dos grande números).

O outro tipo de erro é o sistemático, que é o componente do erro da medição que se mantém constante ou muda de forma previsível quando se efetuam várias medições da mesma grandeza (VIM, 2000). Isto é, quando existe somente o erro sistemático, os n resultados y_i são iguais e a diferença para o valor verdadeiro y_v é a mesma (Vuolo, 2000).

A possibilidade de esses erros estarem presentes num estudo exige do pesquisador a realização de aferições de erro nas medidas de suas variáveis. Segundo Pereira (2004), essa aferição diz respeito à confiabilidade da medida e inclui avaliações de exatidão e precisão dessas medidas.

Exatidão é o grau de concordância entre o valor médio obtido de uma grande série de resultados de ensaio e um valor de referência aceito e a sua medida é usualmente expressa em termos de viés. Assim, quanto menor o viés, maior é a exatidão. Viés é um erro sistemático, em que um ou mais componentes do erro podem contribuir para o mesmo, por exemplo, viés do laboratório, de um método de medição (Association of Analytical Communities, AOAC, 1993).

De acordo com Mood et al. (1974), um estimador T é definido ser um estimador não viesado ou exato do parâmetro $\tau(\theta)$ se, e somente se, a esperança de T é igual a $\tau(\theta)$, ou seja, $E[T]_{\theta} = \tau(\theta)$. O termo $\tau(\theta) - E[T]_{\theta}$ é denominado de viés do estimador T .

Lupin & Schouenborg (2000) definem precisão como sendo o grau de concordância entre resultados de testes independentes obtidos sobre condições pré-estabelecidas e que dependem somente da distribuição dos erros aleatórios. Vuolo (2000) afirma que, em medições com boa precisão, obtêm-se resultados com muitos dígitos e bem repetitivos. Entretanto, pode existir um erro sistemático grande e, por conseguinte, a exatidão pode não ser boa (Figura 10).

Estatisticamente, a precisão é uma medida da proximidade dos valores oriundos de repetições de um processo em relação a uma medida de tendência

central desse mesmo processo. Por exemplo, uma medida possível para a precisão seria soma de quadrados ou um quadrado médio do tipo variância: $Var[Y_i]$.

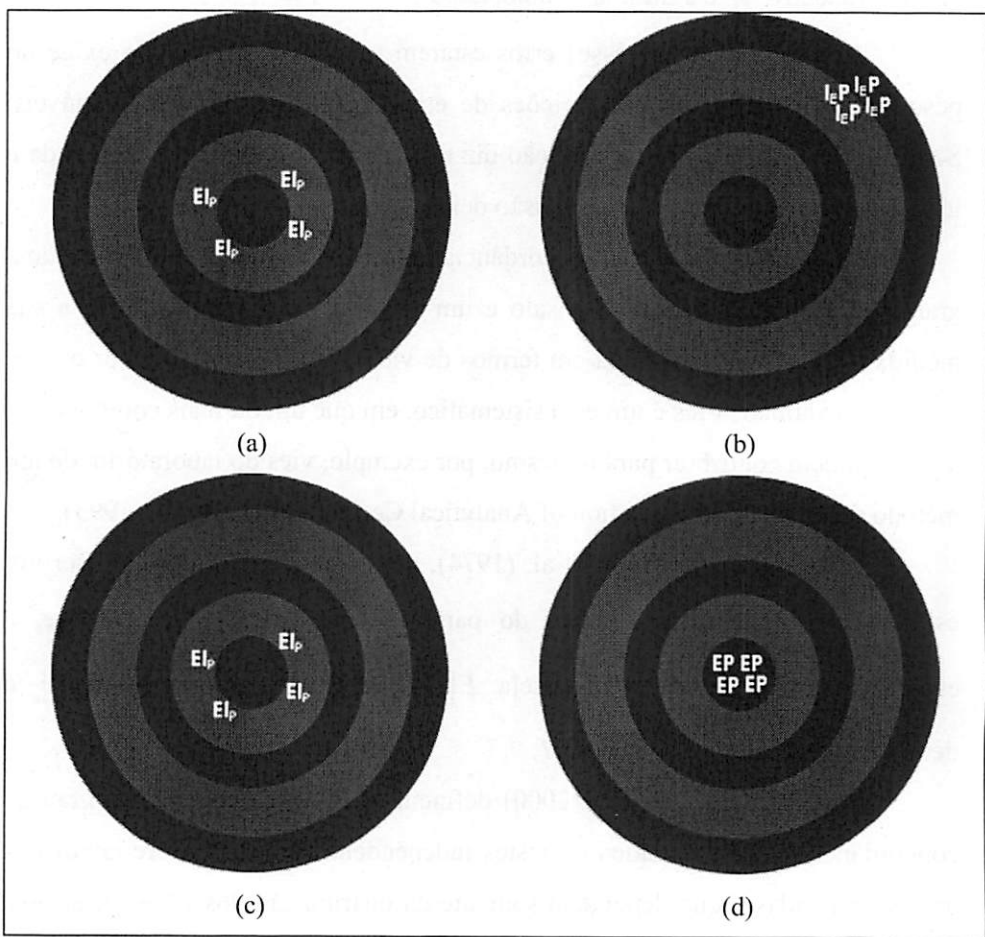


FIGURA 10 Representação da exatidão e precisão
Fonte: adaptado de Pereira (2004: p.49)

Na Figura 10 as simbologias significam: (a) I_{EI_P} - inexatidão (não se aproxima do alvo) e - imprecisão (as marcas são dispersas); (b) I_{EP} - inexatidão (está longe do alvo), mas precisão (pequena dispersão dos tiros); (c) EI_P - exatidão (aproxima-se do alvo), mas imprecisão (marcas dispersas) e (d) EP - exatidão (acerta o alvo) e precisão (pequena dispersão).

No caso do programa de validação de método para testes de qualidade de sementes, a precisão dos resultados é avaliada considerando-se tanto a precisão dentro do laboratório, denominada de repetibilidade, como a precisão geral (entre e dentro dos laboratórios), designada de reprodutibilidade.

Segundo Campos (1998), repetibilidade é a concordância entre os resultados de uma série de medidas de uma característica, quando há constância de um método, analista, instrumento, local e demais condições operacionais e ambientais. No caso de variar-se premeditadamente uma ou mais das condições que garantem a avaliação da repetibilidade, a concordância entre os resultados de uma série de medidas de uma característica representaria a reprodutibilidade.

As condições de repetibilidade e reprodutibilidade estão sumarizadas no Quadro 1.

QUADRO 1 Condições de repetibilidade e reprodutibilidade

Condições de repetibilidade	Condições de reprodutibilidade
o mesmo laboratório	diferentes laboratórios
o mesmo operador	diferentes operadores
o mesmo equipamento	diferentes equipamentos
usando o mesmo método de teste	
medindo materiais idênticos	
ocorrendo num pequeno intervalo de tempo	

Fonte: Lupin & Schouenborg (2000).

No protocolo da ISTA (2007a), para a validação de métodos, esses conceitos podem ser interpretados da seguinte forma: a repetibilidade irá quantificar a variabilidade média dos resultados dentro de cada laboratório, quando as repetições são feitas na amostra de um dado lote e a reprodutibilidade será a repetibilidade aumentada pela variabilidade dos resultados de um laboratório para outro laboratório de um dado lote (variância entre laboratórios), sendo calculadas de acordo com as normas da ISO 5725-2 (1994).

Com os valores da repetibilidade e reprodutibilidade, podem ser calculados os seus limites críticos, por meio da multiplicação por um fator numérico D , que se refere a uma amplitude padronizada baseada na tabela de Tukey (1977), com graus de liberdade tendendo para o infinito.

Esses limites críticos podem ser utilizados na comparação com as amplitudes determinadas entre as repetições dentro dos laboratórios e entre laboratórios, determinando, assim, se os índices obtidos de repetibilidade e reprodutibilidade são aceitáveis (Olivieri, 1988; Campos, 1998; Campos et al. 1999, Lupin & Schouenborg, 2000).

Outra medida que pode auxiliar na avaliação da precisão, além da repetibilidade e da reprodutibilidade, é a estatística k de Mandel, que indica o desvio de um laboratório, quando comparado com o desvio padrão da repetibilidade (ISTA, 2007a).

Já a estatística h está relacionada com a avaliação da exatidão, já que o seu valor indica a estimativa do viés, ou seja, o valor h é a medida do desvio da média de um laboratório, quando comparado com a média geral obtida de todos os laboratórios (Lupin & Schouenborg, 2000).

Assim, os valores de h podem ser calculados para verificar quais resultados dos laboratórios estão super ou subestimados, em comparação com a média geral (ISTA, 2007a).

Geralmente, as estatísticas k e h de Mandel são apresentadas graficamente. Um exemplo do gráfico k considerando oito laboratórios e quatro lotes pode ser observado na Figura 11.

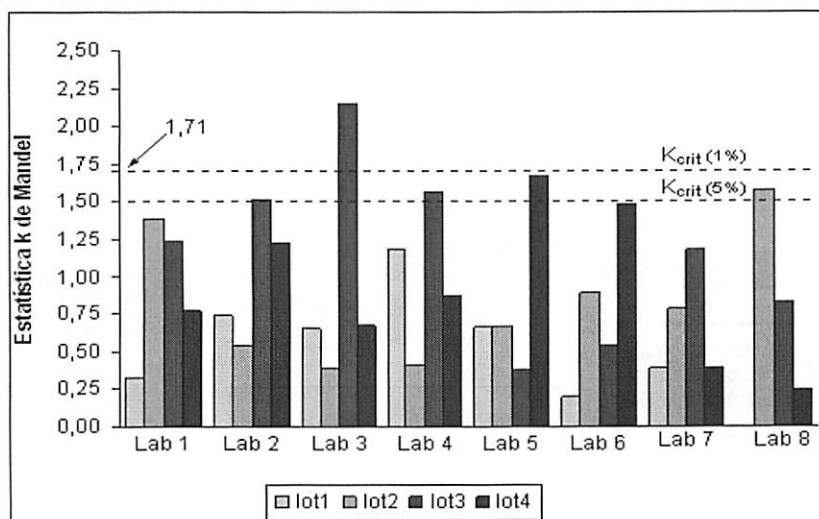


FIGURA 11 Exemplo do gráfico da estatística k de Mandel, comparado com os valores críticos (K_{crit}), a 1% e 5%, para um ensaio com oito laboratórios e quatro lotes

Uma síntese dos passos desta subetapa pode ser observada na Figura 12.

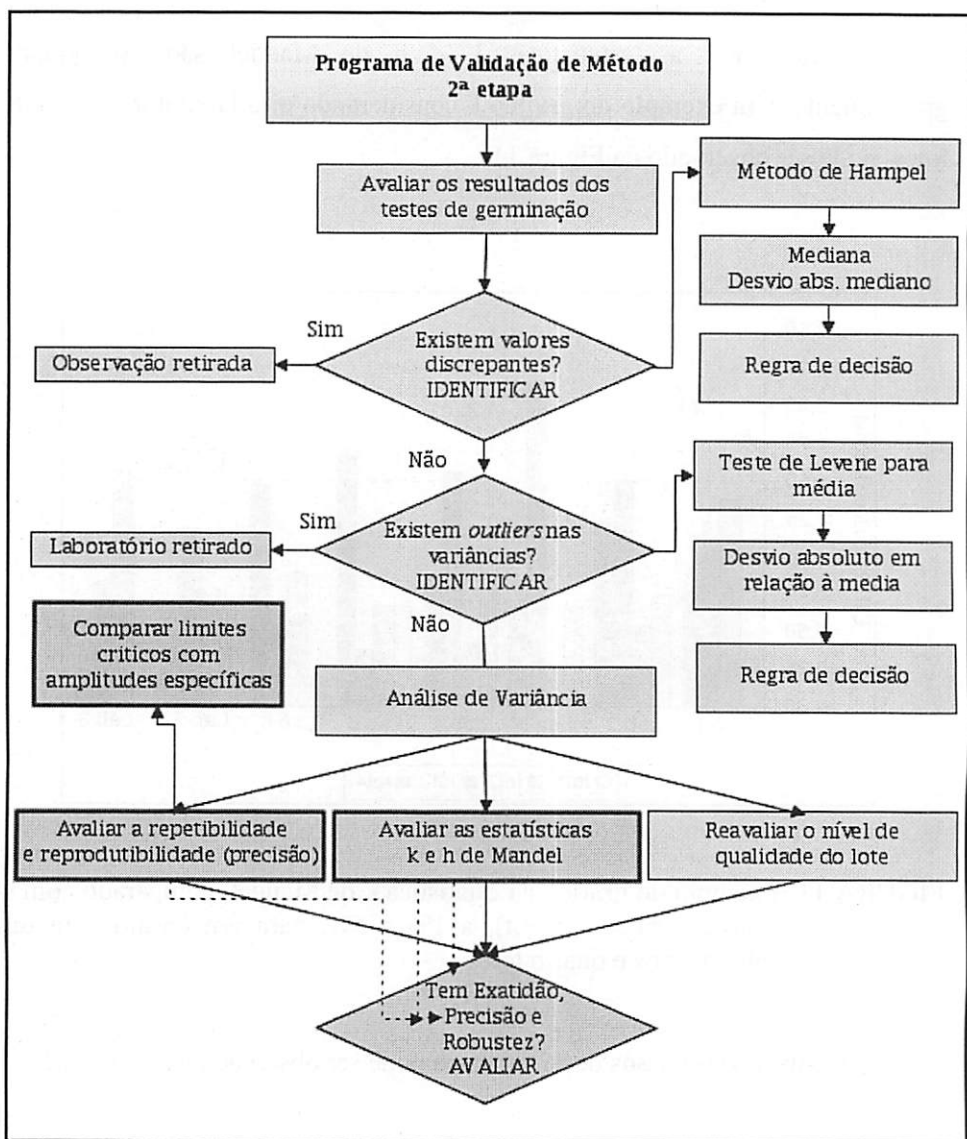


FIGURA 12 Esquema da 2ª etapa do programa de validação de métodos, com ênfase nas subetapas avaliação de repetibilidade, reprodutibilidade e das estatísticas h e k de Mandel

No que tange a robustez dos métodos⁴, neste contexto do processo de validação, ela passa a ser uma consequência direta da avaliação da precisão e exatidão dos resultados para cada um dos lotes. Isto é, se um lote de qualidade inferior não apresentar, por exemplo, reprodutibilidade, e se um lote intermediário for considerado reprodutível, pode ser um indício de que o método não seja suficientemente robusto.

Outro detalhe importante é que a classificação dos lotes é feita por meio de um processo de amostragem, sendo os resultados avaliados pelo laboratório que coordena o processo de validação. Sendo assim, é interessante averiguar se, de fato, esses resultados vão estar em consonância com os que serão obtidos nos demais laboratórios, ou seja, se um lote considerado previamente como sendo de baixa qualidade continuará mantendo a mesma classificação após o processo de validação.

Para tal, a técnica estatística que pareceu ser a mais apropriada para fazer essa reavaliação do nível de qualidade dos lotes foi a análise discriminante. Ressalta-se que esse procedimento não faz parte do padrão da ISTA para um programa de validação de método, sendo uma proposta desta tese.

4.9 Análise discriminante

A técnica multivariada de análise discriminante, segundo Barroso & Artes (2003) e Ferreira (2008), tem como principais objetivos:

- discriminação: consiste em encontrar funções das variáveis observadas (funções discriminantes) que são responsáveis ou que possam mensurar as diferenças entre as g populações;

⁴ Método robusto é aquele que suporta condições diferenciadas sem apresentar resultados descontroladamente errôneos, isto é, não há comprometimento da confiabilidade dos resultados.

- classificação ou alocação: consiste em determinar funções das variáveis que permitam classificar novos objetos em uma das g populações.

Esta técnica consiste em determinar uma combinação linear entre duas ou mais variáveis independentes que discriminarão melhor os grupos definidos a priori. A discriminação é conseguida estabelecendo-se os pesos da combinação linear que maximizam a variância entre grupos relativos à variância dentro dos grupos. A combinação linear para a análise discriminante é também conhecida como função discriminante (Hair et al., 2006).

Uma fase importante no procedimento da análise discriminante é a verificação de violação de hipóteses, como normalidade e multicolinearidade, dentre outras. A suposição de normalidade não é condição necessária para as análises quando aplicado o método de Fisher para a determinação da função discriminante. Entretanto, a suposição de igualdade das matrizes de variâncias e covariâncias é requerida, ou seja, $\Sigma_1 = \dots = \Sigma_g = \Sigma$ (Mingoti, 2005).

Uma vez que a função discriminante linear for calculada, uma observação pode ser alocada para uma das g populações com base em seu “*score* discriminante” e, como na amostra, a população de origem de cada observação é conhecida, será possível identificar quais deles foram classificados corretamente ou não. Assim, pode ser calculada a taxa estimada de erro (TEE), que é a fração das observações incorretamente classificadas, sendo uma estimativa da taxa de erro verdadeira.

De acordo com Ferreira (2008), uma boa regra de alocação deve levar a uma baixa TEE, mas este valor pode estar sendo subestimado, pois a mesma amostra é usada para determinar a função de classificação. Para contornar esse problema, pode-se usar um procedimento de validação cruzada (*cross-validation*), sugerido por Lachenbruch & Mickey (1968).

Portanto, no contexto do processo de validação, indiretamente, o uso da análise discriminante auxilia na avaliação da robustez do método estudado (Figura 13).

Após a aplicação de todas essas técnicas estatísticas, se o procedimento fornece exatidão, robustez e precisão (reprodutibilidade, repetibilidade), o método testado pode ser considerado como validado.

4.10 Comparação de métodos

No caso deste estudo, para finalizar o processo de validação proposto, foi necessário definir também uma técnica estatística para comparar os dois métodos testados (substratos areia e papel). Verificou-se que a utilização de testes F poderia ser um procedimento adequado para tal fim.

Lupin & Schouenborg (2000) afirmam que podem ser realizados testes F para comparar tanto a precisão dentro dos laboratórios como a precisão geral de dois métodos, utilizando-se como base os índices de repetibilidade de cada método e os de reprodutibilidade, respectivamente (Figura 14).

Salienta-se que esta subetapa da comparação não interfere diretamente no processo de validação de cada método, já que os testes F vão avaliar apenas se os índices de precisão de um método são diferentes do outro.

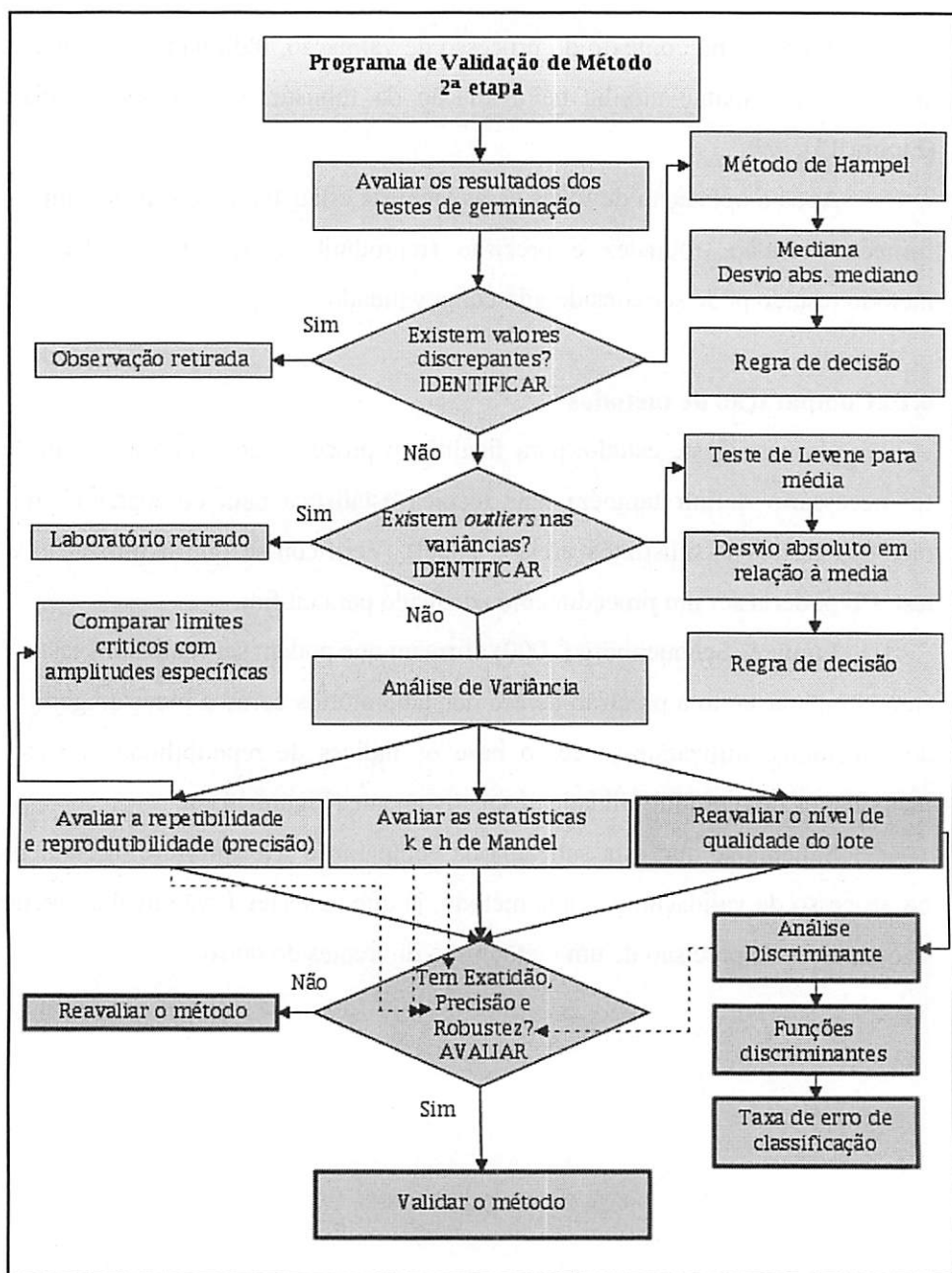


FIGURA 13 Esquema da 2ª etapa do programa de validação de métodos, com ênfase na subetapa reavaliação do nível de qualidade dos lotes.

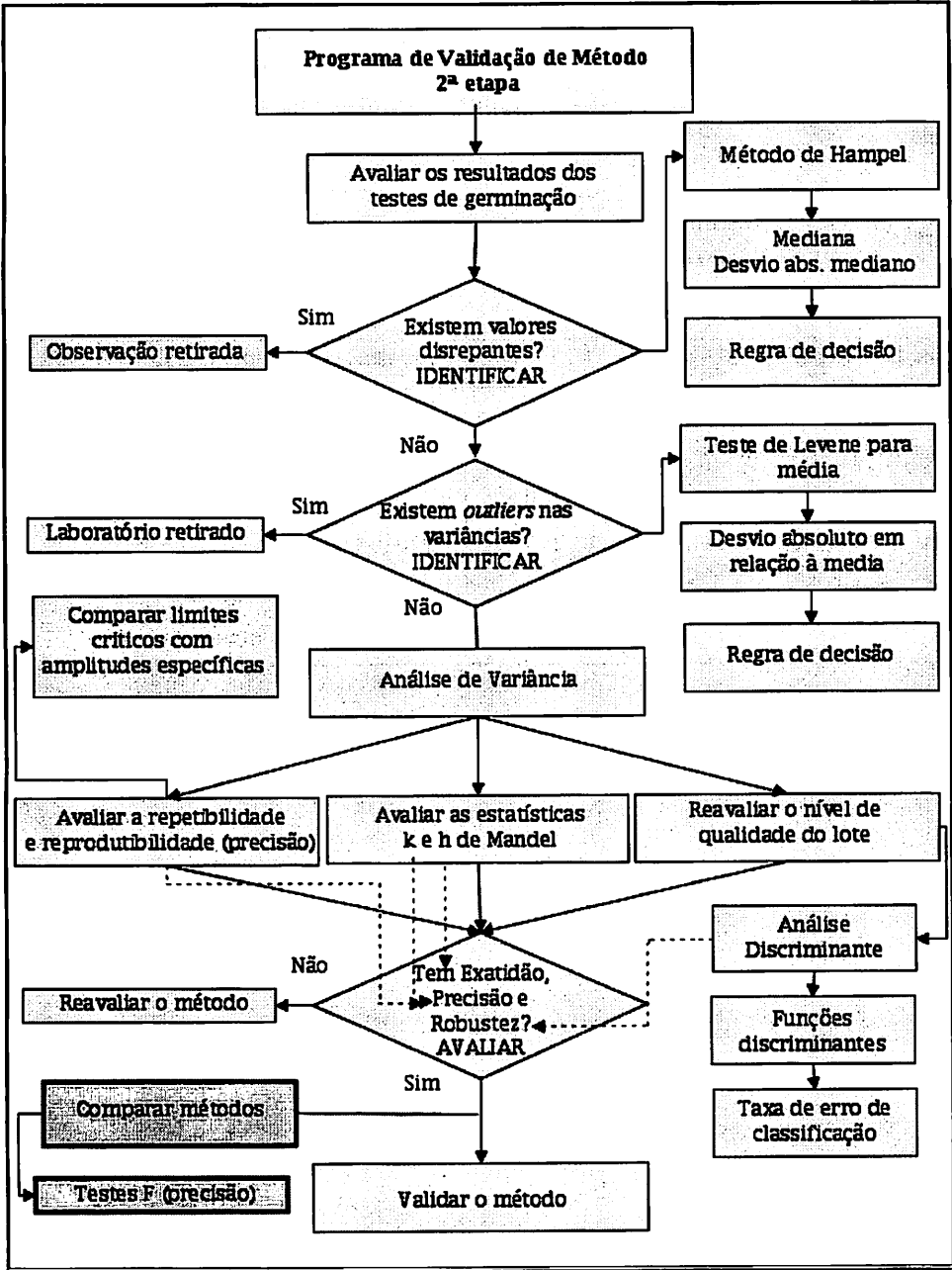


FIGURA 14 Esquema da 2ª etapa do programa de validação de métodos, com ênfase na subetapa comparação de métodos.

Como ressaltado na seção 3, um dos problemas de pesquisa foi propor o uso de técnicas estatísticas complementares ao procedimento padrão da ISTA, o que fica mais bem descrito observando-se o resumo apresentado no Quadro 2.

QUADRO 2 Resumo do procedimento padrão da ISTA e as técnicas estatísticas propostas complementares para o processo de validação de métodos deste trabalho.

Padrão ISTA	Propostas complementares
<ul style="list-style-type: none"> - Teste de Hampel, Boxplot - Análise de variância - Valores de repetibilidade, reprodutibilidade, estatísticas h e k de Mandel 	<ul style="list-style-type: none"> - Teste de Levene para média (em substituição ao teste de Cochran) - Limites críticos de repetibilidade e reprodutibilidade - Testes F para comparação dos métodos - Análise discriminante

5 MATERIAL E MÉTODOS

Foram adquiridas sementes de sete lotes de nabo forrageiro, cultivar CATI AL 1000, de diferentes safras e procedências, sendo coletados 40 kg de cada lote.

As etapas de seleção e caracterização dos lotes e preparação do material para distribuição aos laboratórios selecionados para processo de validação foram realizadas no Laboratório de Análise de Sementes, no Departamento de Agricultura da Universidade Federal de Lavras, em Lavras, MG.

5.1 Seleção de lotes de acordo com a homogeneidade – Teste H

Após a aquisição dos materiais, foram realizados testes de germinação para a avaliação da homogeneidade do lote por meio do teste H.

Cada lote de 40 kg foi dividido em 5 recipientes e de cada um deles foi retirada 1 subamostra de 30 g, com o auxílio de um amostrador tipo duplo para sementes. A coleta foi feita de acordo com as prescrições da RAS (Brasil, 1992).

Para a realização do teste de germinação, as subamostras foram classificadas de acordo com o tamanho das sementes, utilizando-se uma peneira de 8 mm.

A semeadura foi realizada em caixas acrílicas do tipo *gerbox* em substrato papel mata-borrão umedecido com água destilada 2,5 vezes a sua massa. Foram utilizadas 100 sementes de cada subamostra, divididas em duas caixas de 50.

Em seguida, as caixas foram transferidas para uma câmara de germinação tipo BOD (*biological oxygen demand*) regulada à temperatura alternada 20°-30°C, e procedida à contagem de plântulas normais aos três e sete dias. Com os resultados foi realizado, para cada lote, o teste H.

No teste H, a variável aleatória estudada foi a porcentagem de germinação (Q), considerando como hipótese nula (H_0) se o lote é homogêneo e, como hipótese alternativa (H_1) se o lote é heterogêneo.

Foi computada a estatística H (valor da heterogeneidade) para o j -ésimo lote, utilizando-se a expressão (ISTA, 2007b):

$$H_j = \frac{V_j}{W_j} - f, \quad (4)$$

em que $V_j = \frac{N \sum_{i=1}^N Q_{ji}^2 - \left(\sum_{i=1}^N Q_{ji} \right)^2}{N(N-1)}$ é o estimador da variância observada:, sendo

Q_{ji} a porcentagem de germinação na i -ésima subamostra e N o número de

subamostras; $W_j = \frac{\bar{Q}_j (100 - \bar{Q}_j)}{n} f$ é o estimador da variância aceitável:, sendo

n o número de sementes de cada subamostra, $\bar{Q}_j = \sum_{i=1}^N \frac{Q_i}{N}$ o estimador da

média e f o fator para variação adicional. Para o teste de germinação, f assume o valor constante 1,1.

Ressalta-se que $\frac{(H + f)(N - 1)}{f}$, sob H_0 , segue uma distribuição qui-

quadrada, com $N - 1$ graus de liberdade. Com os resultados do teste H, foram selecionados, para as próximas etapas, apenas os lotes homogêneos.

A sumarização das fases e um exemplo teórico da organização dos dados e dos cálculos auxiliares para o teste H podem ser observados na Figura 15 e Tabela 2, respectivamente.

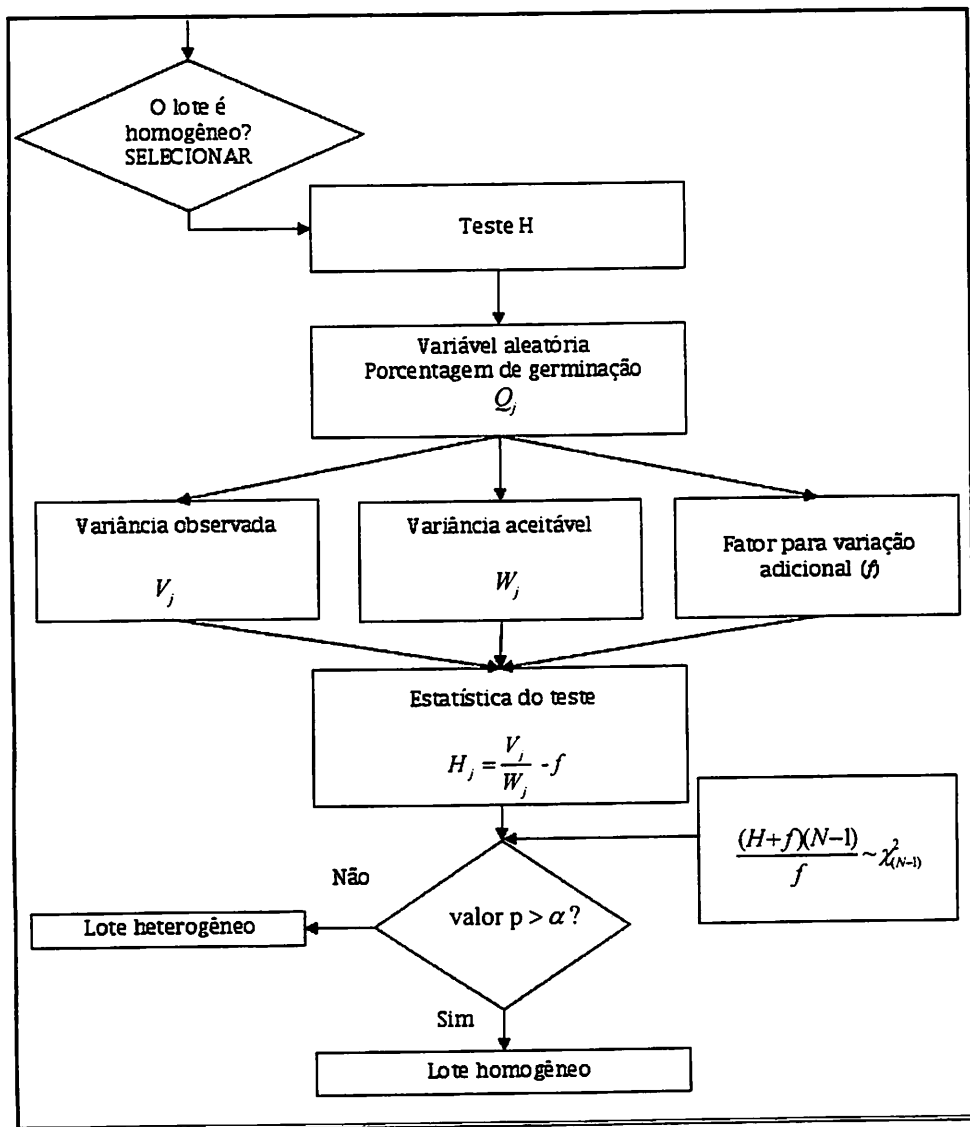


FIGURA 15 Resumo esquemático do teste H para a variável aleatória porcentagem de germinação (Q) no j -ésimo lote

TABELA 2 Exemplo teórico da organização dos dados e cálculos auxiliares para a realização do teste H utilizando os resultados de um teste de germinação no j -ésimo lote com 5 subamostras (N) de 100 sementes cada (n)

Lote	Subamostra	Q_{ji}	Q_{ji}^2	\bar{Q}_j	V_j	W_j
J	1	Q_{j1}				
				
	5	Q_{j5}				
		Σ	Σ			

* Um exemplo numérico pode ser observado no Anexo B (p. 208)

5.2 Caracterização dos lotes – Teste de germinação inicial

Após a realização do teste H, os lotes selecionados foram homogenizados com o auxílio de um divisor mecânico do tipo solo com oito canaletas (Brasil, 1992). Cada lote foi dividido em oito repetições de, aproximadamente, 5 kg cada.

Para cada repetição do lote foi retirada uma amostra de, aproximadamente, 100g para a realização do teste de germinação. As sementes foram classificadas de acordo com o tamanho, utilizando-se uma peneira de 8 mm.

Nessa etapa foram testadas dois métodos, um utilizando, como substrato, o papel e o outro, areia.

A areia passou por um processo de lavagem em água corrente, esterilizada em autoclave por 2 horas e foi umedecida em 50% da sua capacidade de retenção, conforme Brasil (1992).

O papel foi umedecido conforme já descrito no procedimento de germinação do teste H, bem como o processo de semeadura em caixas acrílicas do tipo *gerbox*; as condições da câmara de germinação e a contagem aos três e

sete dias de plântulas normais. As sementes dormentes, remanescentes ao final do teste, foram submetidas ao teste de tetrazólio (Brasil, 1992).

Durante a contagem, as plântulas normais, as plântulas anormais, e as sementes duras, mortas e dormentes foram descritas e fotografadas, para a padronização dos resultados do teste de germinação.

O experimento, para cada método, foi instalado segundo um delineamento inteiramente casualizado (DIC), apenas para a variável resposta porcentagem de germinação, com 8 repetições de 50 sementes. Os dados foram submetidos à análise de variância e o efeito dos lotes, quando significativo, teve suas médias comparadas pelo teste de Scott-Knott, a 5% de significância (Figura 16).

Silva et al. (1999) recomendam a utilização do teste de Scott-Knott para os procedimentos de comparações múltiplas, devido ao fato de ele possuir poder⁵ elevado, taxas de erro tipo I quase sempre de acordo com os níveis nominais e por não apresentar resultados com ambiguidade

De acordo com os resultados deste teste de germinação inicial, os lotes foram classificados em três níveis de qualidade: inferior, intermediária e superior.

Neste trabalho, o lote foi considerado de qualidade inferior se a taxa de germinação fosse menor que 60%, intermediária entre 60% e 70%, e superior acima de 70%. Essas faixas foram escolhidas com base na porcentagem de germinação mínima para comercialização, que é de 60% para *Raphanus sativus* (Brasil, 2008).

⁵ O poder de um teste é a capacidade deste de identificar todas as reais diferenças entre os tratamentos (Ramalho et al., 2000), ou seja, a probabilidade de rejeitar a hipótese nula H_0 , quando ela é falsa ($1-\beta$). (Mood et al., 1974)

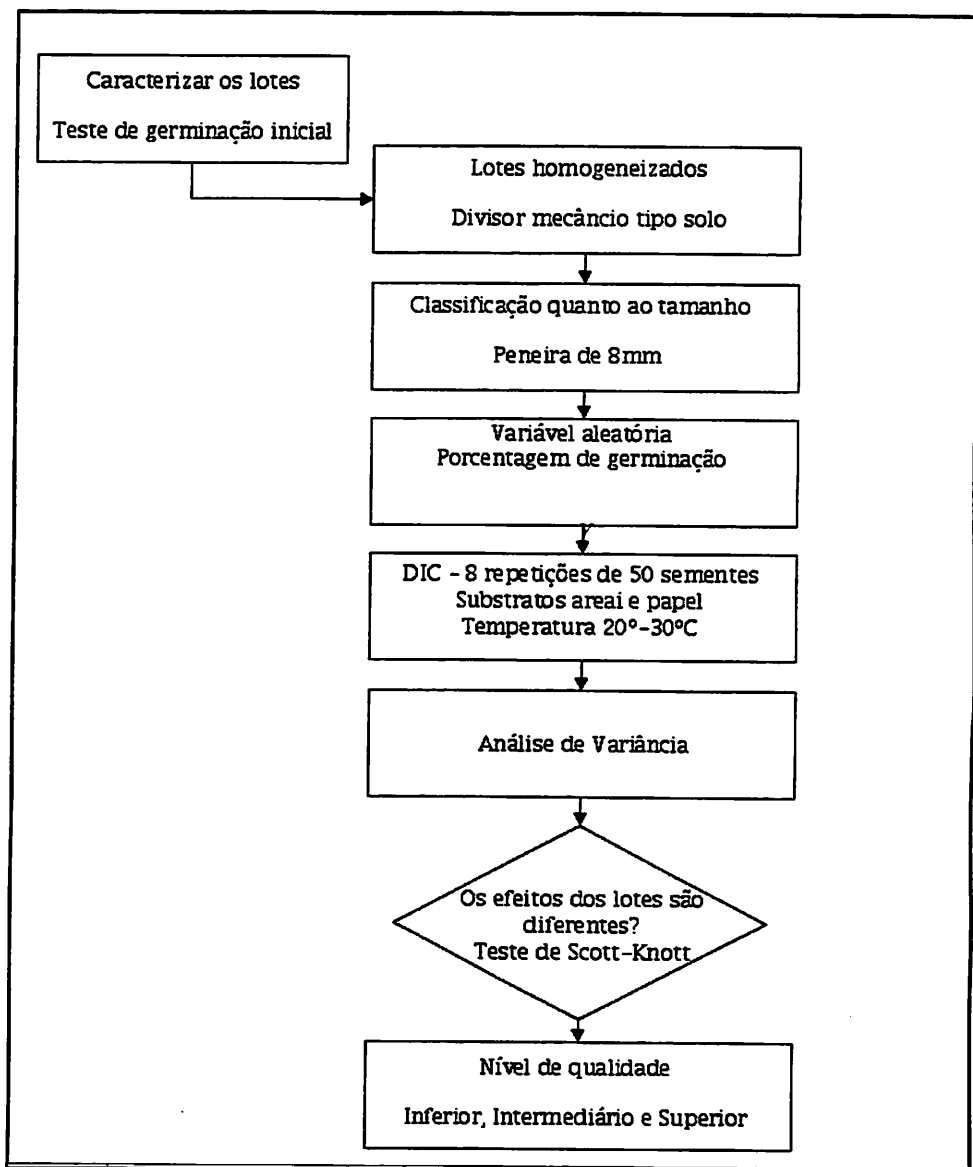


FIGURA 16 Resumo esquemático do teste de germinação inicial da variável aleatória porcentagem de germinação (Y)

5.3 Distribuição do material para os laboratórios

Nessa etapa foram coletadas, de cada lote, amostras com cerca de 4 kg, utilizando-se um amostrador tipo duplo, classificadas de acordo com o tamanho por meio de uma peneira de 8 mm, ficando reduzidas a 1,2 kg. As amostras enviadas aos oito laboratórios foram resultantes da subdivisão desse total, ou 150g por laboratório.

Deve-se ressaltar que os laboratórios selecionados eram de prestação de serviço ou de pesquisa, credenciados ou não pelo MAPA e com diferentes níveis de experiência com o teste de germinação para nabo forrageiro.

As sementes foram acondicionadas em sacos de papel multifoliado e caixas de papelão e enviadas aos laboratórios, juntamente com as instruções para a realização e a avaliação do teste de germinação, além de um questionário para avaliar a execução prática de cada um dos métodos (Anexo A). Os laboratórios foram informados da data de envio das amostras, sendo solicitada a rápida execução do teste de germinação para evitar a deterioração das sementes (Figura 17).

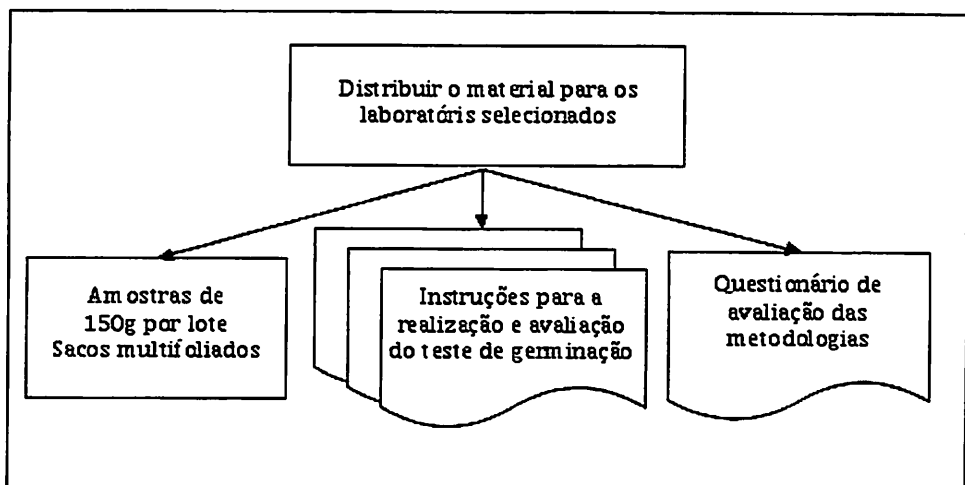


FIGURA 17 Resumo esquemático da distribuição de material para os laboratórios selecionados.

Cada método proposto foi realizado três vezes pelos laboratórios, que foram denominado de ensaios 1, 2 e 3.

Todos os procedimentos descritos para os dois métodos propostos neste trabalho foram estabelecidos com base nos estudos de Nery (2008).

5.4 Das técnicas estatísticas

Nesta seção serão descritas as técnicas estatísticas que foram utilizadas para: identificar a presença de valores discrepantes e *outliers* nas variâncias; avaliar os efeitos de laboratórios e lotes; verificar a repetibilidade, a reprodutibilidade, a exatidão e a robustez; comparar os diferentes métodos de germinação e reavaliar o nível de qualidade dos lotes (Figura 14, p. 45).

Exemplos numéricos para apresentar a aplicabilidade de algumas dessas técnicas podem ser observados no Anexo B.

Ressalta-se que, para proceder as análises estatísticas, foram calculadas as médias de cada ensaio (médias de 8 repetições) e estas passaram a ser

denominadas de repetições, ou seja, de cada lote nas análises foram consideradas apenas as três repetições. A variável aleatória em estudo foi a porcentagem de germinação (Y).

Cada etapa foi realizada separadamente para os substratos papel e areia, com exceção do teste F (subseção 5.4.5), que foi utilizado justamente para comparar os resultados dos dois métodos.

A estrutura padrão dos dados para esse processo de validação dos métodos pode ser observada na Tabela 3. Exemplos de como esta estrutura foi sendo modificada de acordo com as exigências de cálculo de cada técnica estatística podem ser observados tanto nos exemplos teóricos que vão ser apresentados a seguir, bem como por meio dos exemplos numéricos no Anexo B.

TABELA 3 Estrutura padrão dos dados para o processo de validação de métodos

Lote (j)	Laboratório (l)	Repetição (k)	Germinação (Y_{jlk})
1	1	1	Y_{111}
		2	Y_{112}
		3	Y_{113}

	8	1	Y_{181}
		2	Y_{182}
3		Y_{183}	
...
5	1	1	Y_{511}
		2	Y_{512}
		3	Y_{513}

	8	1	Y_{581}
		2	Y_{582}
3		Y_{583}	

5.4.1 Identificação de valores discrepantes (*outlier*)

Para a detecção de valores discrepantes em todos os laboratórios do j -ésimo lote, foram utilizadas dois tipos de técnicas, uma gráfica pelo esboço do Boxplot (Tukey, 1977) e outra quantitativa, por meio de método de Hampel (Hampel, 1985).

O método de Hampel para o j -ésimo lote consistiu em:

- computar a mediana ($\tilde{Y}_{j..}$) e o desvio absoluto mediano $MAD_j = \text{mediana} \{ |Y_{jlk} - \tilde{Y}_{j..}| \}$, em que Y_{jlk} é a j -ésima porcentagem de germinação observada na k -ésima repetição (ensaio) do l -ésimo laboratório;

- utilizar a regra $|Y_{jlk} - \tilde{Y}_{j..}| > 5,2MAD_j$, para identificar se uma determinada observação era um *outlier*.

Deve-se ressaltar que, apenas nos casos em que foi detectada a presença de *outliers* por meio do método de Hampel, esses valores foram retirados das amostras para as próximas fases do processo de validação.

O Boxplot foi utilizado tanto como uma técnica auxiliar ou confirmatória do método de Hampel, como para permitir uma avaliação visual do comportamento dos resultados (a descrição das medidas necessárias para sua construção pode ser observado no Anexo C).

A sumarização das fases e um exemplo teórico da organização dos dados e dos cálculos auxiliares para o método de Hampel podem ser observados na Figura 18 e na Tabela 4, respectivamente.

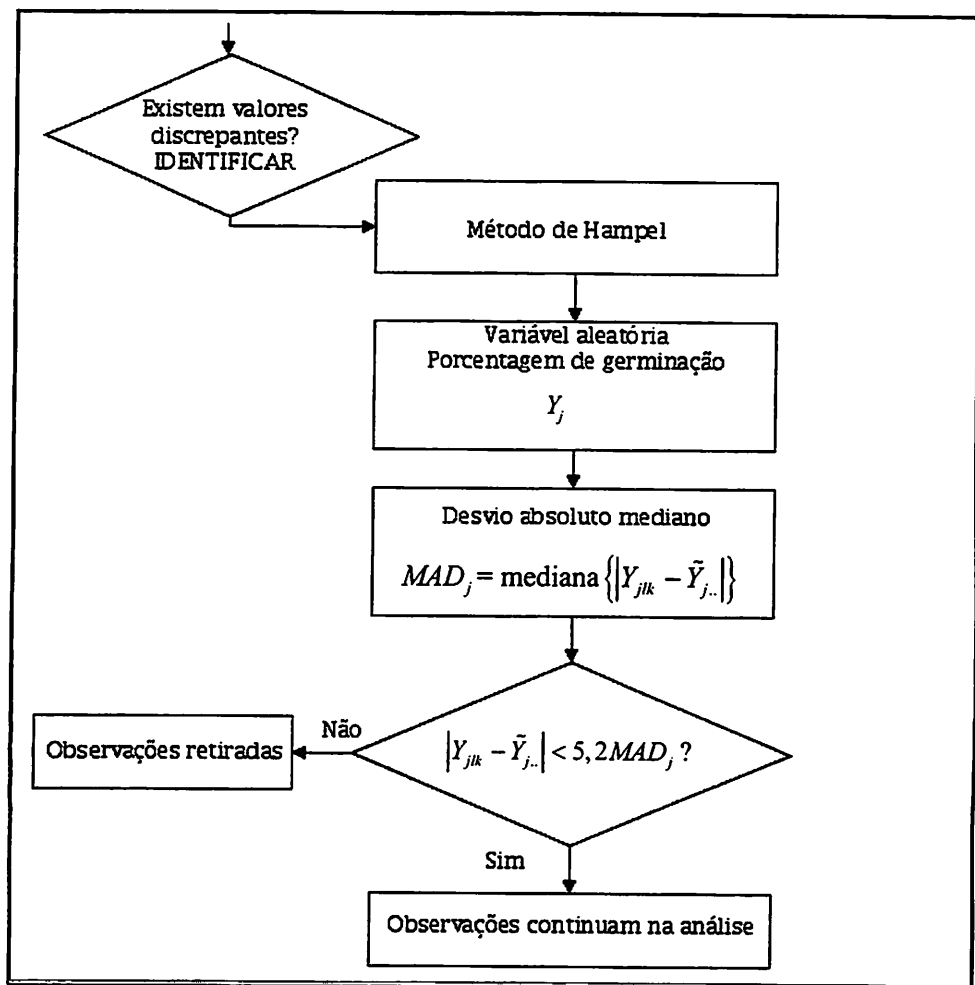


FIGURA 18 Resumo esquemático do método de Hampel para a variável aleatória porcentagem de germinação (Y_j) no j -ésimo lote

TABELA 4 Exemplo teórico da organização dos dados e cálculos auxiliares para a realização do método de Hampel utilizando os resultados de um teste de germinação no j -ésimo lote com p laboratórios e três repetições

Lote	Lab.	Rep.	Y_{jlk}	$ Y_{jlk} - \bar{Y}_{j..} $
J	1	1	Y_{j11}	
		2	Y_{j12}	
		3	Y_{j13}	

	p	1	Y_{jp1}	
		2	Y_{jp2}	
3		Y_{jp3}		
\tilde{Y}_j Mediana →			Mediana ↓	
MAD_j				
$5,2MAD_j$				

* Um exemplo numérico pode ser observado no Anexo B (p. 210)

5.4.2 Identificação de *outlier* nas variâncias

Para avaliar a homogeneidade nas variâncias dos laboratórios, em cada lote foi realizado o teste de Levene (utilizando à média), com as seguintes hipóteses estabelecidas para o j -ésimo lote:

$$\begin{cases} H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_{p_j}^2 \\ H_1: \sigma_b^2 \neq \sigma_m^2, \text{ para pelo menos um par } (b, m), \end{cases}$$

em que p_j é o número de laboratórios no j -ésimo lote.

Foi computada a estatística W_j para o j -ésimo lote, utilizando-se a seguinte expressão (Levene, 1960):

$$W_j = \frac{(N_j - p_j) \sum_{l=1}^{p_j} n_{jl} (\bar{Z}_{jl} - \bar{Z}_{j..})^2}{(p_j - 1) \sum_{l=1}^{p_j} \sum_{k=1}^{n_{jl}} (Z_{jlk} - \bar{Z}_{jl})^2}, \quad (5)$$

em que $Z_{jlk} = |Y_{jlk} - \bar{Y}_{jl}|$, sendo \bar{Y}_{jl} o estimador da média do l -ésimo laboratório e Y_{jlk} a porcentagem de germinação da k -ésima repetição no l -ésimo laboratório; n_{jl} é o número de repetições do l -ésimo laboratório e N_j é o número total de observações $(N_j = \sum_{l=1}^{p_j} n_{jl})$.

A estatística W_j , sob H_0 , segue uma distribuição F com $(p_j - 1)$ e $(N_j - p_j)$ graus de liberdade.

No caso em que era detectada a presença de *outlier*, o laboratório com maior variância era retirado e o processo repetido, até que ficassem apenas os laboratórios com variâncias homogêneas, de acordo com o teste Levene.

A sumarização das fases e um exemplo teórico da organização dos dados e dos cálculos auxiliares para o teste de Levene podem ser observados na Figura 19 e na Tabela 5, respectivamente.

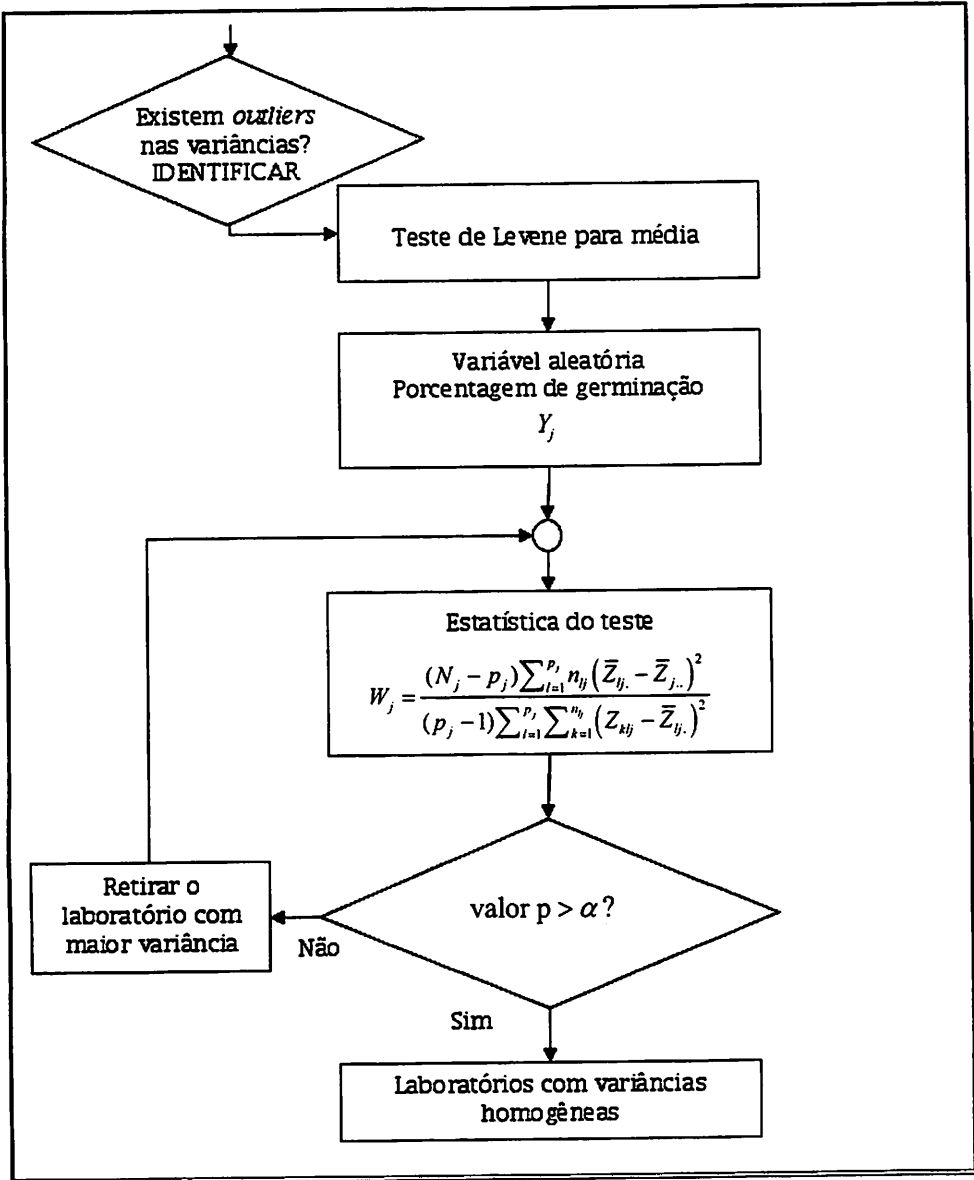


FIGURA 19 Resumo esquemático do teste de Levene para a variável aleatória porcentagem de germinação (Y_j) no j -ésimo lote.

TABELA 5 Exemplo teórico da organização dos dados e cálculos auxiliares para a identificação de *outliers* pelo teste de Levene utilizando os resultados de um teste de germinação no j -ésimo lote com p laboratórios e três repetições

Lote	Lab	Rep	Y_{jlk}	$Z_{jlk} = Y_{jlk} - \bar{Y}_{jl} $	$(Z_{jlk} - \bar{Z}_{jl})^2$	$(\bar{Z}_{jl} - \bar{Z}_{j..})^2$	
j	1	1	Y_{j11}				
		2	Y_{j12}				
		3	Y_{j13}				
			\bar{Y}_{j1}	$\bar{Z}_{j1.}$	Σ	Σ	
	...	1	Y_{jp1}				
		2	Y_{jp2}				
		3	Y_{jp3}				
				\bar{Y}_{jp}	$\bar{Z}_{jp.}$	Σ	Σ
				$\bar{Z}_{j..}$	Σ	Σ	

* Um exemplo numérico pode ser observado no Anexo B (p. 211)

5.4.3 Análise de variância

Para avaliar os efeitos dos laboratórios e lotes, foi aplicada a teoria da análise de variância num delineamento inteiramente casualizado em esquema fatorial com oito repetições, entretanto considerou-se a interação desprezível, ficando seu efeito confundido nos resíduos. Foi proposto então o seguinte modelo estatístico:

$$Y_{jlk} = \mu + \beta_j + \gamma_l + \varepsilon_{jlk}, \quad (6)$$

em que: Y_{jlk} é a porcentagem de germinação da k -ésima repetição, no l -ésimo laboratório e j -ésimo lote ($\forall k = 1, \dots, n_r, l = 1, 2, \dots, p$ e $j = 1, 2, \dots, n_j$); μ é a media geral; β_j é o efeito do j -ésimo lote; γ_l é o efeito do l -ésimo laboratório ; ε_{jlk} é o erro experimental da k -ésima repetição, no l -ésimo laboratório e j -ésimo lote

$(\varepsilon_{kjl} \sim NID(0, \sigma^2))^6$; n_j = número de lotes; n_r = número de repetições; p = número de laboratórios.

5.4.4 Avaliação da repetibilidade, reprodutibilidade e estatísticas h e k de Mandel

a) Repetibilidade

A variância de repetibilidade representa uma medida da variabilidade dentro dos laboratórios. Para o j -ésimo lote, o seu estimador foi obtido por meio da expressão (ISO 5725-2, 1994):

$$S_{rj}^2 = \frac{\sum_{l=1}^{p_j} (n_{jl} - 1) S_{jl}^2}{\sum_{l=1}^{p_j} (n_{jl} - 1)}, \quad (7)$$

em que n_{jl} é o número de repetições do l -ésimo laboratório; p_j é o número de laboratórios e S_{jl}^2 é o estimador da variância do l -ésimo laboratório.

O limite crítico r_j de repetibilidade para cada lote foi estimado por $r_j = S_{rj} D_j$, em que D_j foi obtido da tabela de Tukey, com graus de liberdade tendendo para o infinito (Banzatto & Konkra, 2006), podendo também ser calculado pela expressão $D_j = z_\gamma \sqrt{\bar{n}_j}$, em que z_γ é o quantil 100 γ % da distribuição normal padronizada e \bar{n}_j é o número médio das repetições dos laboratórios no j -ésimo lote.

⁶ A sigla NID significa normal e independentemente distribuído.

Esse valor de r_j foi comparado com a amplitude entre as repetições de cada laboratório, em cada lote (Lr_{jl}), avaliando-se, então, quais laboratórios tiveram repetibilidade aceitável.

Um exemplo teórico da organização dos dados e dos cálculos auxiliares para a avaliação da repetibilidade pode ser observada na Tabela 6.

TABELA 6 Exemplo teórico da organização dos dados e cálculos auxiliares para a avaliação da repetibilidade nos resultados de um teste de germinação no j -ésimo lote com p laboratórios e três repetições

Lote	Lab	Rep	Y_{jlk}	n_{jl}	s_{ij}^2	Lr_{jl}	s_{rj}^2	D_j	r_j
1		1	Y_{j11}	n_{j1}	s_{j1}^2	Lr_{j1}			
		2	Y_{j12}						
		3	Y_{j13}						
		\bar{Y}_{j1}							
j	
		1	Y_{jp1}	n_{jp}	s_{jp}^2	Lr_{jp}			
		2	Y_{jp2}						
		3	Y_{jp3}						
\bar{Y}_{jp}									
				\bar{n}_j					

* Um exemplo numérico pode ser observado no Anexo B (p. 212)

b) Reprodutibilidade

A variância de reprodutibilidade representa uma medida da variabilidade entre e dentro dos laboratórios. Para o j -ésimo lote, o seu estimador pôde ser obtido por meio da expressão 8 (ISO 5725-2, 1994).

$$S_{Rj}^2 = S_{rj}^2 + S_{Lj}^2, \quad (8)$$

em que $S_{Lj}^2 = \frac{S_{dj}^2 - S_{rj}^2}{\bar{n}_j}$ é o estimador da variância entre laboratórios;

$S_{dj}^2 = \frac{1}{(p_j - 1)} \sum_{l=1}^{p_j} n_{jl} (\bar{Y}_{jl} - \bar{Y}_{j..})^2$; sendo $\bar{Y}_{j..}$ o estimador da média geral de todas

as observações e \bar{Y}_{jl} o estimador da média do l -ésimo laboratório;

$$\bar{n}_j = \frac{1}{(p_j - 1)} \left(\sum_{l=1}^{p_j} n_{jl} - \frac{\sum_{l=1}^{p_j} n_{jl}^2}{\sum_{l=1}^{p_j} n_{jl}} \right).$$

O limite crítico R_j de reprodutibilidade para cada lote foi estimado por $R_j = S_{Rj} D_j$, sendo adotado o mesmo procedimento descrito anteriormente para a repetibilidade, com a diferença apenas de que a amplitude foi calculada entre os laboratórios para cada repetição LR_{jk} , determinando, assim, quais lotes tiveram reprodutibilidade aceitável.

A sumarização das fases para a avaliação da repetibilidade e reprodutibilidade pode ser observada na Figura 20, e um exemplo teórico da organização dos dados e dos cálculos auxiliares para a avaliação da reprodutibilidade pode ser observada na Tabela 7.

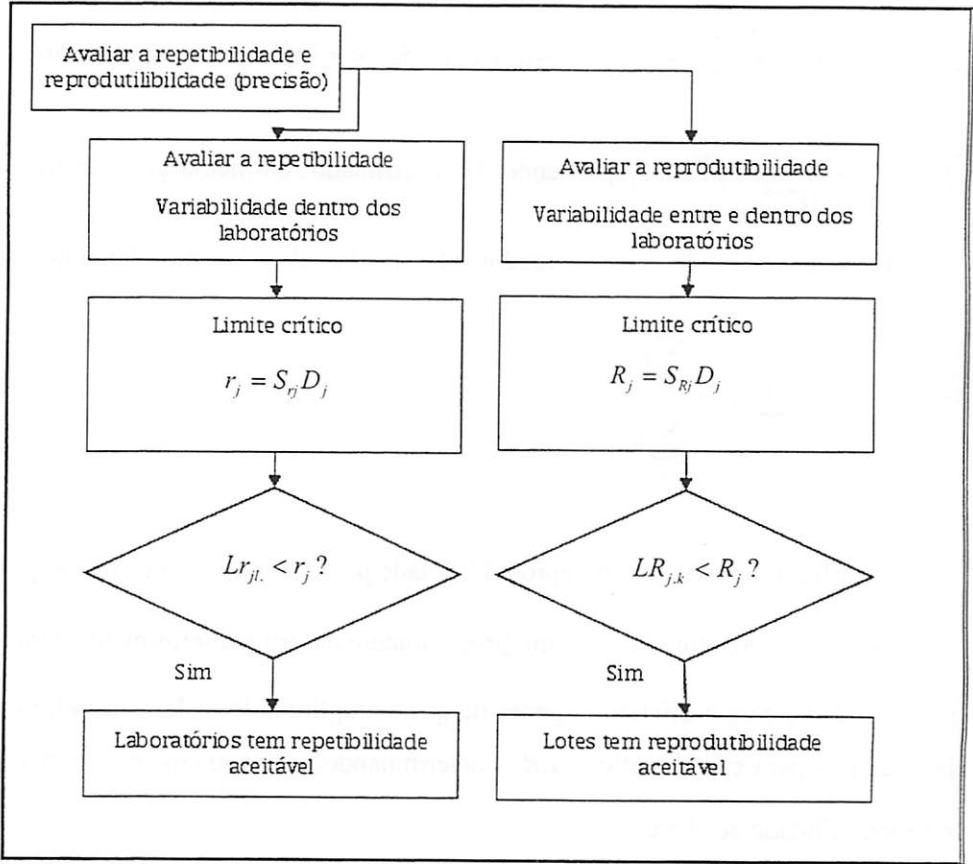


FIGURA 20 Resumo esquemático de avaliação da repetibilidade e da reprodutibilidade para o j -ésimo lote.

TABELA 7 Apresentação teórica dos cálculos auxiliares para a avaliação da reprodutibilidade nos resultados de um teste de germinação no j -ésimo lote com p laboratórios e três repetições

Lote	Rep.	Lab.	Y_{jlk}	$LR_{j,k}$	S_{dj}^2	$S_{L_i}^2$	$S_{R_i}^2$	D_j	R_j	
j	1	1	Y_{j11}	$LR_{j,1}$						
		...								
		p	Y_{jp1}							
	2	1	Y_{j12}	$LR_{j,2}$						
		...								
		p	Y_{jp2}							
	3	1	Y_{j13}	$LR_{j,3}$						
		...								
		p	Y_{jp3}							

* Um exemplo numérico pode ser observado no Anexo B (p. 213)

c) Estatística k de Mandel (k_j) e o valor crítico (k_c)

Uma das medidas utilizadas para auxiliar na avaliação da precisão dos resultados foi a estatística k de Mandel que, após determinada, foi comparada ao valor crítico (k_c), determinados para o j -ésimo lote por (ISO 5725-2, 1994):

$$k_{jl} = \frac{S_{jl}}{S_r}, \tag{9}$$

$$k_c = \sqrt{\frac{n_l F_{(\alpha, \eta_1, \eta_2)}}{F_{(\alpha, \eta_1, \eta_2)} + (n_l - 1)}}, \tag{10}$$

em que $F_{(\alpha, \eta_1, \eta_2)}$ é o quantil $100(1 - \alpha)\%$ da distribuição F com $\eta_1 = (n_r - 1)$ e $\eta_2 = (p - 1)(n_r - 1)$ graus de liberdade, sendo n_r o número de repetições e p o número de laboratórios.

d) Estatística h de Mandel (h_{jl}) e do valor crítico (h_c)

A medida utilizada para avaliar graficamente a estimativa do viés dos resultados foi a estatística h de Mandel para o j -ésimo lote que foi obtida, bem como o valor crítico (h_c), por (ISO 5725-2, 1994):

$$h_{jl} = \frac{(\bar{Y}_{jl} - \bar{Y}_{j..})}{\sqrt{\frac{\sum_{l=1}^{p_j} (\bar{Y}_{jl} - \bar{Y}_{j..})^2}{(p_j - 1)\bar{n}_j}}}, \quad (11)$$

$$h_c = \frac{(p-1)t_{(\alpha,\eta)}}{\sqrt{p(t_{(\alpha,\eta)}^2 + p - 2)}}, \quad (12)$$

em que $t_{(\alpha,\eta)}$ é o quantil $100(1-\alpha)\%$ da distribuição t de Student, com graus de liberdade $\eta = (p - 2)$.

Após os cálculos, foram esboçados os gráficos dos valores de h e k para cada lote e laboratório, como no exemplo da Figura 10 (subseção 4.8), e comparados com os valores críticos para os níveis de significância de 1% e 5%.

Exemplo teórico da organização dos dados e cálculos auxiliares para a determinação dos valores de h e k podem ser observados na Tabela 8.

TABELA 8 Exemplo teórico da organização dos dados e cálculos auxiliares para a determinação das estatísticas h e k de Mandel nos resultados de um teste de germinação no j -ésimo lote com p laboratórios e três repetições

Lote	Lab	Rep	y_{kj}	s_{jl}	s_{rj}	k_{jl}	$\bar{y}_{j..}$	h_{jl}	
j	1	1	Y_{j11}						
		2	Y_{j12}	s_{j1}		k_{j1}		h_{j1}	
		3	Y_{j13}						
			$\bar{y}_{j1.}$						
		
	p	1							
		2			s_{jp}		k_{jp}		h_{jp}
		3							
			$\bar{y}_{jp.}$						

* Um exemplo numérico pode ser observado no Anexo B (p. 214)

5.4.5 Comparação dos métodos

a) Comparação da precisão dos métodos dentro dos laboratórios

Foi calculada uma estatística F_{rj} para cada lote $j = (1, 2, \dots, n_j)$ e feita a comparação com dois limites tabelados $F_{rj\text{inf}}$ e $F_{rj\text{sup}}$, utilizando-se as estimativas das variâncias da repetibilidade, considerando os dois substratos, areia e papel (Lupin & Schouenborg, 2000):

$$F_{rj} = \frac{S_{rj1}^2}{S_{rj2}^2}, \quad (13)$$

$$F_{rj\text{inf}} = F_{\{(1-\alpha/2), \eta_{j1}, \eta_{j2}\}} \text{ e } F_{rj\text{sup}} = F_{\{\alpha/2, \eta_{j1}, \eta_{j2}\}},$$

em que $F_{(\alpha, \eta_1, \eta_2)}$ é o quantil $100(1-\alpha)\%$ da distribuição F com $\eta_{j1} = p_{j1}(n_{rj1} - 1)$ e $\eta_{j2} = p_{j2}(n_{rj2} - 1)$ graus de liberdade, sendo n_{rj1} o número

de repetições do método 1 no j -ésimo lote, p_{j1} o número de laboratórios do método 1 no j -ésimo lote, a designação para o método 2 é equivalente.

Sendo as linhas de decisão como seguem:

- se $F_{Rj\text{inf}} \leq F_{Rj} \leq F_{Rj\text{sup}}$, não há evidências de que os métodos tenham precisões diferentes dentro de cada laboratório;
- Se $F_{Rj} < F_{Rj\text{inf}}$, há evidências de que o método 1 tenha precisão superior ao método 2 dentro de cada laboratório;
- Se $F_{Rj} > F_{Rj\text{sup}}$, há evidências de que o método 1 tenha precisão inferior ao método 2, dentro de cada laboratório;

b) Comparação da precisão geral dos métodos

Foi calculada uma estatística F_{Rj} para cada lote ($j = 1, \dots, n_j$) e feita a comparação com os valores críticos, inferior e superior, similares aos apresentados no item a, bem como as linhas de decisão. A única diferença é que a interpretação foi feita para a precisão geral, sendo, então, utilizadas as estimativas das variâncias de reprodutibilidade e repetibilidade dos substratos areia e papel (Lupin & Schouenborg, 2000):

$$F_{Rj} = \left(S_{Rj1}^2 - \left(1 - \frac{1}{n_{kj1}} \right) S_{rj1}^2 \right) / \left(S_{Rj2}^2 - \left(1 - \frac{1}{n_{kj2}} \right) S_{rj2}^2 \right), \quad (14)$$

$$F_{Rj\text{inf}} = F_{\{(1-\alpha/2), \eta_{j1}, \eta_{j2}\}} \text{ e } F_{Rj\text{sup}} = F_{\{\alpha/2, \eta_{j1}, \eta_{j2}\}},$$

em que $F_{(\alpha, \eta_1, \eta_2)}$ é o quantil $100(1-\alpha)\%$ da distribuição F , com $\eta_{j1} = (p_{j1} - 1)$ e $\eta_{j2} = (p_{j2} - 1)$ graus de liberdade.

A sumarização das fases de execução para a comparação de métodos pode ser observada na Figura 21.

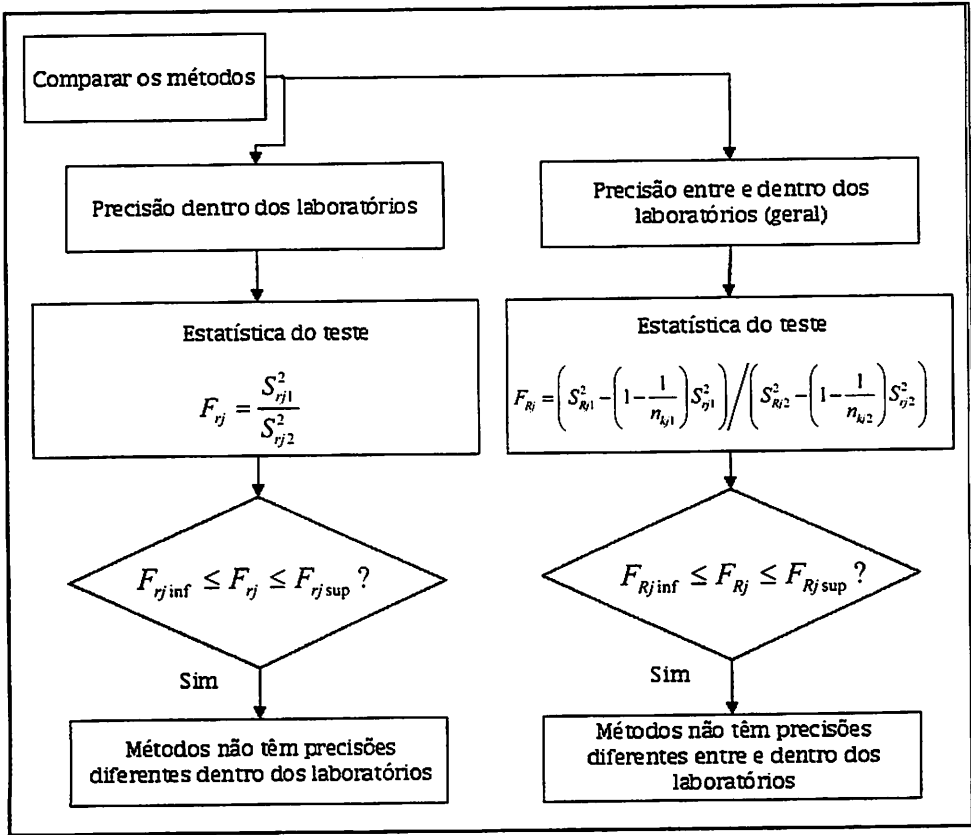


FIGURA 21 Resumo esquemático da comparação entre métodos para o j -ésimo lote

5.4.6 Análise discriminante

Como foram selecionados lotes com níveis de qualidade diferentes, utilizou-se a técnica de análise discriminante para averiguar, com base nos resultados obtidos pelos laboratórios, a adequação da classificação feita *a priori*. Assim, foram formados três grupos, correspondentes aos lotes de qualidade inferior, intermediária e superior.

A aplicação do método de Fisher⁷ foi dividida em quatro etapas, considerando:

- três grupos, denotados por τ_1 (lote inferior), τ_2 (lotes intermediários) e τ_3 (lote superior),
- o vetor da m -ésima observação (porcentagem de germinação) no i -ésimo grupo⁸ representado por X_{im} e composto das medidas das r -variáveis (repetições), logo $X_{im}^T = [X_{im1} \dots X_{imr}]$ (Figura 1.22);

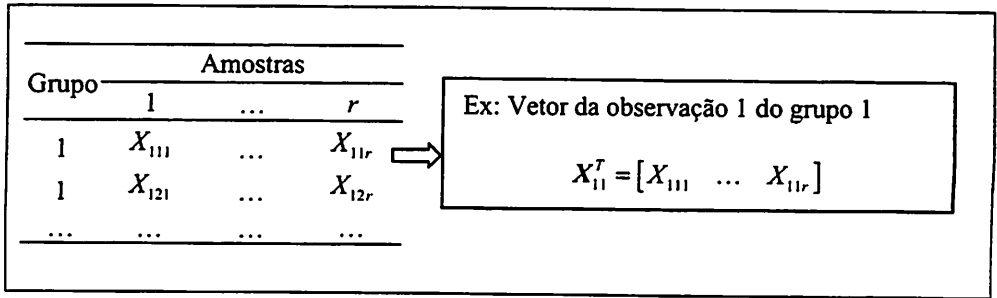


FIGURA 22 Exemplo teórico da representação de um vetor de observação gerado a partir de um banco de dados (um exemplo numérico pode ser observado no Anexo B, p. 214)

⁷ Os passos e a simbologia (em negrito, para representar uma matriz ou um vetor) utilizados estão de acordo com Barroso & Artes (2003), com algumas adaptações para o contexto deste estudo.

⁸ Não há correspondência direta entre a variável Y (utilizada nas etapas anteriores) e a variável X , no que tange à indicação do lote e/ou do laboratório. Ou seja, o valor, por exemplo, da observação y_{111} , não é igual ao da observação x_{111} , já que a primeira representa a porcentagem de germinação do lote 1, laboratório 1 e repetição 1 e a segunda representa a porcentagem de germinação do grupo 1, laboratório 1 e repetição 1. Por este motivo, elas foram indexadas com letras diferentes.

- cada grupo foi composto de n_i observações r -variadas: $X_{11}, X_{12}, \dots, X_{1n_1}$ amostradas da população τ_1 , $X_{21}, X_{22}, \dots, X_{2n_2}$ da população τ_2 e $X_{31}, X_{32}, \dots, X_{3n_3}$ da população τ_3 ;
- estimadores dos vetores de médias (μ_i) denotadas por $\bar{X}_1, \bar{X}_2, \bar{X}_3$,
- estimadores das matrizes de variâncias e covariâncias (Σ_i) do vetor X para os três grupos, representados por S_1, S_2, S_3 .

1ª etapa: avaliação da suposição de igualdade das matrizes de covariância

Para avaliar a suposição que $S_1 = S_2 = S_3$, foi utilizado o teste de Box M^9 , sendo S_i calculado como segue:

$$S_i = \frac{1}{n_i - 1} \sum_{m=1}^{n_i} (X_{im} - \bar{X}_i)(X_{im} - \bar{X}_i)^T, \quad (15)$$

em que $\bar{X}_i = \frac{1}{n_i} \sum_{m=1}^{n_i} X_{im}$ é o estimador do vetor de médias do i -ésimo grupo.

2ª etapa: determinação das funções discriminantes lineares de Fisher

As s funções discriminantes lineares de Fisher são determinadas por meio de combinações lineares das variáveis em estudo e representadas por $T_u = l_u^T X$, em que $u = 1, 2, \dots, s', \dots, s$, sendo $s = \min(r, g - 1)$, isto é, existem s soluções não-nulas de combinações lineares igual ao mínimo entre o número de r variáveis e o número de grupos (g) menos 1 (Ferreira, 2008).

A ideia inicial de Fisher (1938), para obter uma função discriminante T , foi estabelecer combinações lineares que maximizassem uma razão em que o numerador era a soma das distâncias ao quadrado entre as médias de T de cada

⁹ Para maiores detalhes veja Anexo D.

grupo e a média global de T e o denominador, a variância de T . Ou seja, a razão entre a variabilidade entre grupos e a variabilidade comum dentro de grupos, representada por:

$$\frac{\sum_{i=1}^{n_i} (\bar{X}_{iT} - \bar{X}_T)^2}{S_T^2} = \frac{l^T \hat{B}_0 l}{l^T S_p l}, \quad (16)$$

em que: $\bar{X}_{iT} = l^T \bar{X}_i$ é o estimador da média de T para o i -ésimo grupo;

$\bar{X}_T = l^T \bar{X}$ é o estimador da média geral de T , sendo $\bar{X} = \frac{1}{g} \sum_{i=1}^g \bar{X}_i$;

$S_T^2 = l^T S_p l$ é o estimador da variância de T ; $S_p = \left(\frac{1}{\sum_{i=1}^g n_i - g} \right) \sum_{i=1}^g (n_i - 1) S_i$ é

o estimador de Σ , ou seja, a matriz de variâncias e covariâncias combinada

e \hat{B}_0 é o estimador da soma de produtos cruzados: $\hat{B}_0 = \sum_{i=1}^{n_i} (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^T$.

Logo, o problema em questão consistia em obter o vetor l que maximizasse a expressão (1.16), o que equivale à maximização de uma razão entre formas quadráticas (Ferreira, 2008 e Barros & Artes, 2003). Dessa forma, os vetores l_u foram obtidos dos autovetores do sistema de equações homogêneas:

$$\left(\hat{B}_0 - \lambda_u S_p \right) l_u = 0 \quad (17)$$

Para solucionar este sistema, adotou-se a transformação linear não-singular do vetor l_u , dada por $l_u = S_p^{-1/2} e_u$ e pré-multiplicando-se o sistema por $S_p^{-1/2}$, obtendo-se:

$$\left(S_p^{-1/2} \hat{B}_0 S_p^{-1/2} - \lambda_u I \right) e_u = 0 \quad (18)$$

em que λ_u e e_u são os autovalores e autovetores da matriz¹⁰ $S_p^{-1/2} \hat{B}_0 S_p^{-1/2}$

Calculados os autovetores, determinou-se o vetor de combinação linear, retomando-se a expressão $l_u = S_p^{-1/2} e_u$.

Assim, os coeficientes da primeira função discriminante (l_1) são os elementos do autovetor padronizado de $S_p^{-1} \hat{B}_0$, associado ao maior autovalor dessa matriz.

Para determinar a segunda função discriminante, deve se buscar o vetor l que maximiza a referida razão sujeita à restrição $Cov(l_1^T X, l_2^T X) = 0$, que é o autovetor padronizado associado ao segundo maior autovalor de $S_p^{-1} \hat{B}_0$. Generalizando, a s' -ésima função discriminante é $l_s^T X$, em que l_s é o correspondente s' -ésimo autovetor padronizado, sujeito à condição $Cov(l_b^T X, l_s^T X) = 0$, $b < s'$.

3ª etapa: regra linear de classificação

Considerou-se T o vetor de dimensões $s \times 1$, cujos elementos foram as s funções discriminantes e \bar{X}_{it} o correspondente vetor de estimadores das médias

¹⁰ As matrizes $S_p^{-1} \hat{B}_0$ e $S_p^{-1/2} \hat{B}_0 S_p^{-1/2}$ têm os mesmos autovalores e os autovetores de $S_p^{-1} \hat{B}_0$ são iguais a $S_p^{-1/2}$ vezes os autovalores de $S_p^{-1/2} \hat{B}_0 S_p^{-1/2}$ (Barroso & Artes, 2003).

no i -ésimo grupo. A regra de classificação consistiu em alocar X no grupo τ_k , se a distância ao quadrado entre T_k e \bar{X}_{kT_k} fosse menor que a mesma distância entre T_i e \bar{X}_{iT_i} , para todo $i \neq k$, ou seja, alocar X em τ_k se:

$$\sum_{u=1}^s (T_u - \bar{X}_{kT_u})^2 = \sum_{u=1}^s [l_u^T (X - \bar{X}_k)]^2 \leq \sum_{u=1}^s [l_u^T (X - \bar{X}_i)]^2, \quad (19)$$

Para todo $i \neq k$, em que l_u é o u -ésimo autovetor padronizado de $S_p^{-1} \hat{B}_0$.

4ª etapa: avaliação da função de classificação

Para avaliar a função de classificação, foi utilizado o procedimento de validação cruzada ou método Pseudo-Jackknife (Lachenbruch & Mickey, 1968), para o cálculo da taxa de erro estimada (TEE).

De acordo com Ferreira (2008), a idéia deste método é omitir uma observação por vez das $\sum_{i=1}^g n_i$ originais na amostra de treinamento e estimar as regras de classificação utilizando-se as $\sum_{i=1}^g n_i - 1$ observações remanescentes. A observação omitida é classificada em um dos grupos. Como se sabe qual o grupo foi originalmente amostrado, pode-se determinar, para cada grupo, quantas observações foram classificadas corretamente. Monta-se, então, a denominada “matriz de confusão” (Tabela 9) e calcula-se a TEE a partir da mesma:

TABELA 9 Exemplo de uma matriz de confusão para três grupos

grupo verdadeiro	classificado			Total
	τ_1	τ_2	τ_3	
τ_1	n_{11}	n_{12}	n_{13}	n_1
τ_2	n_{21}	n_{22}	n_{23}	n_2
τ_3	n_{31}	n_{32}	n_{33}	n_3
Total	\hat{n}_1	\hat{n}_2	\hat{n}_3	N

A TEE geral foi calculada, então, da seguinte forma

$$TEE = \frac{n - \sum_{i=1}^g n_{ii}}{n}, \quad (20)$$

em que n_{ij} = número de observações de g_i classificados em τ_j ; \hat{n}_i = número de observações classificadas em τ_i ; n_i = número de observações em τ_i e n = número total de observações na amostra.

Essa taxa de erro geral abrange todos os tipos de erros de classificação que podem ocorrer, mas foi possível estimar também a taxa de erro de qualquer um dos grupos. Por exemplo, para calcular a TEE de se classificar uma observação do grupo τ_i no grupo τ_j , utilizou-se a seguinte expressão:

$$TEE(j | i) = \frac{n_{ij}}{n_i}, \quad (21)$$

Além disso, foi calculada a TEE de se classificar uma observação da grupo τ_i incorretamente, dada por:

$$TEE(i) = \frac{n_i - n_{ii}}{n_i}, \quad (22)$$

A sumarização das fases de execução para a utilização da técnica de análise discriminante pode ser observada na Figura 23.

Um exemplo numérico de todos os passos apresentados para esta técnica pode ser observado no Anexo B.

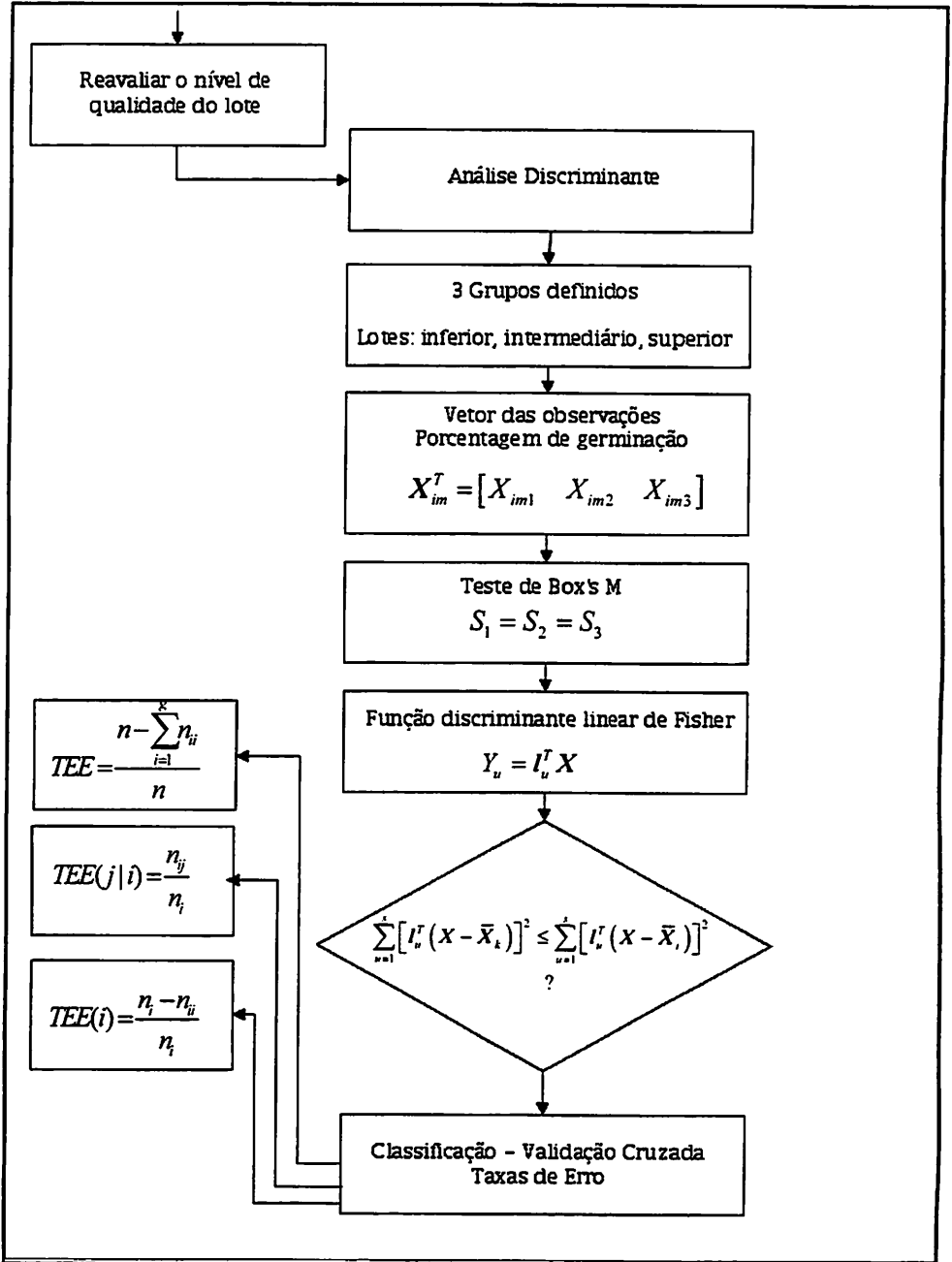


FIGURA 23 Resumo esquemático da técnica de análise discriminante no contexto de um processo de validação de método

5.5 Rotinas no R

Para todas as análises estatísticas foram desenvolvidas funções específicas no software R (R Development Core Team, 2008), no intuito de facilitar a manipulação e a interpretação dos resultados não apenas para estatísticos, mas para pesquisadores da área de sementes.

As funções foram estruturadas de tal forma que o usuário deve entrar apenas com um banco de dados inicial e, por exemplo, se for detectado um valor discrepante pelo método de Hampel, o mesmo era retirado e um segundo banco gerado automaticamente para a execução do teste de Levene. Novamente, se algum laboratório for considerado como um *outlier*, um terceiro banco era gerado para as etapas seguintes.

Para a comparação dos métodos, são necessários os valores dos desvios padrões de reprodutibilidade e repetibilidade, bem como o número de laboratórios e um número de repetições por lote. É gerado também, automaticamente, um banco de dados após a execução dessa etapa correspondente.

Os nomes e os argumentos necessários para cada função podem ser observados na Tabela 10.

TABELA 10 Nome da função e argumentos necessários, de acordo com a técnica estatística

Técnica	Nome da função	Argumentos
Teste H	FunçãoTesteH	- nome do banco de dados. - n° de amostras simples; - n° de sementes por amostra; - fator de correção (f) - nível de significância.
Método de Hampel e Boxplot	FunçãoHampelBXP	- nome do banco de dados.
Teste de Levene	FunçãoLevene	- nome do banco de dados (pré-estabelecido ¹); - nível de significância; - número de vezes que o procedimento deve ser repetido.
Análise de variância	FunçãoANAVA	- nome do banco de dados (pré-estabelecido).
Repetibilidade, Reprodutibilidade Valores de h e k	FunçãoRRhk	- nome do banco de dados (pré-estabelecido); - grau de confiança.
Comparação dos métodos	FunçãoCompmetod	- nome do banco de dados (pré-estabelecido); - nível de significância.
Análise discriminante	Funçãodiscrim	- nome do banco de dados; - probabilidades a priori.

¹ O banco de dados era gerado na etapa anterior, com um nome já pré-estabelecido.

Exemplos de alguns relatórios gerados, tipo de banco de dados e os comandos para execução da FunçãoTesteH podem ser observados no Anexo E.

6 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados tanto os resultados obtidos nos procedimentos iniciais no Laboratório de Análise de Sementes (LAS) da UFLA (1ª etapa), como na aplicação do processo de validação nos oitos laboratórios selecionados (2ª etapa).

6.1 Seleção de lotes de acordo com a homogeneidade - Teste H

De acordo com os resultados do teste H, cinco dos sete lotes avaliados não apresentaram evidências estatísticas para rejeitar a hipótese de homogeneidade do lote, a 1% de significância (Tabela 11 e Figura 24).

Estes cinco lotes foram selecionados atendendo ao critério de homogeneidade dos lotes, que é um dos princípios básicos da amostragem para análise de sementes (Copeland & Liu, 2005).

TABELA 11 Resultados do teste H (H calculado e valor p) para os sete lotes de sementes de nabo forrageiro

Lote	H calculado	valor p
1	-0,30	0,5731
2	5,58	< 0,0001
3	-0,83	0,9125
4	0,52	0,2074
5	2,29	0,0151
6	3,21	0,0035
7	1,63	0,0417

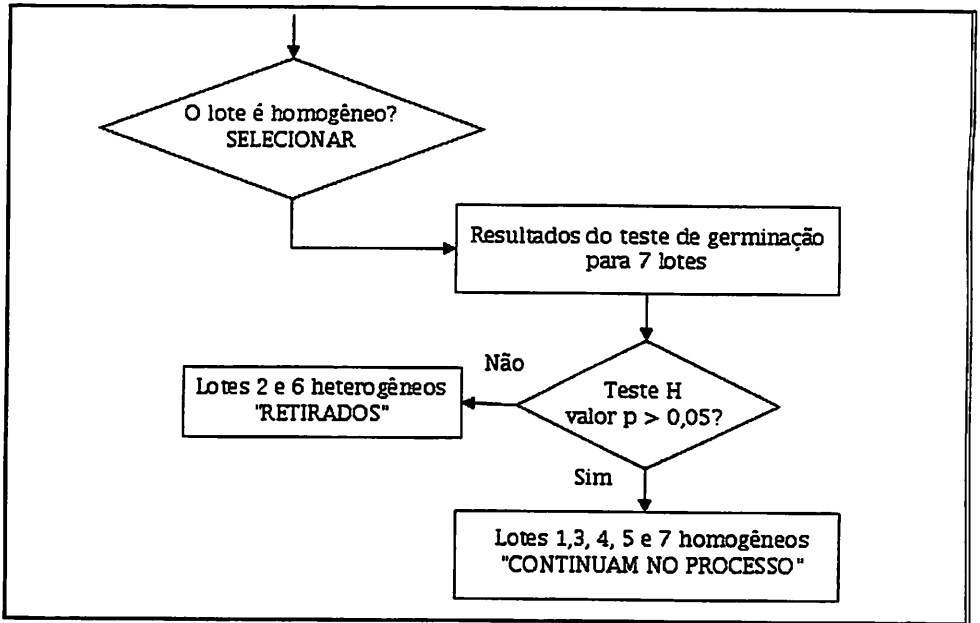


FIGURA 24 Resumo esquemático dos resultados do teste H, a 1% de significância

Para etapas posteriores, somente os lotes considerados homogêneos 1, 3, 4, 5 e 7 foram utilizados e renumerados de 1 a 5, mantendo-se a mesma ordem.

6.2 Caracterização dos lotes - Teste de germinação inicial

Os resultados obtidos no teste de germinação indicam haver diferenças estatisticamente significativas entre os efeitos dos lotes, tanto na utilização do substrato areia como papel, sendo aplicado o teste de Scott-Knott, a 5% significância. Esses lotes foram classificados, quanto ao nível de porcentagem de germinação, de acordo a qualidade fisiológica, em grupos: inferior, intermediária e superior (Tabela 12 e Figura 25).

TABELA 12 Resultados médios¹¹ (%) obtidos no teste de germinação inicial com sementes de cinco lotes de nabo forrageiro para os métodos areia e papel

Lote	'Médias	
	Areia	Papel
1	76 a	75 a
2	55 c	44 c
3	76 a	70 a
4	66 b	57 b
5	67 b	67 a

¹ Médias seguidas de mesma letra minúscula na coluna não diferem entre si pelo teste de Scott-Knott, a 5% de significância.

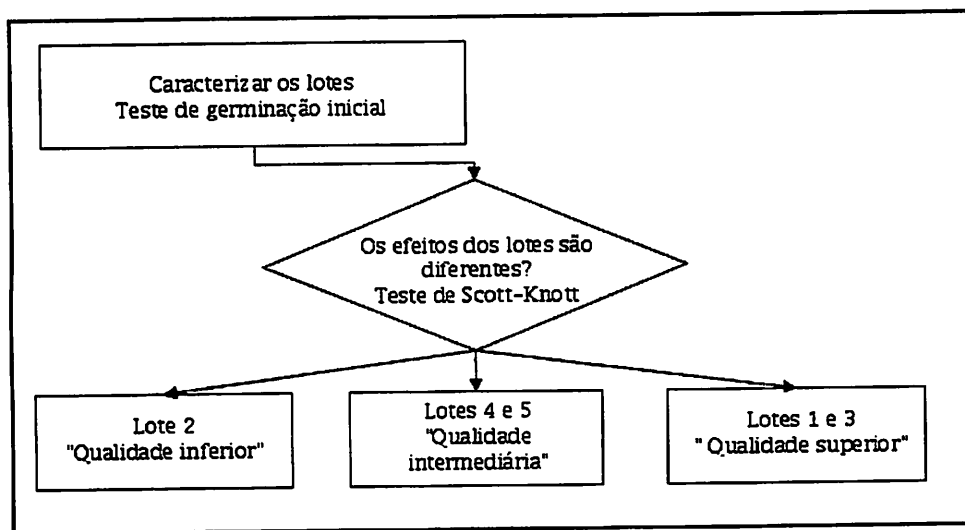


FIGURA 25 Resumo esquemático do resultado do teste de Scott-Knott, a 5% de significância

¹¹ No cálculo das médias apresentadas neste estudo foi utilizado o critério de arredondamento estabelecido nas Regras Internacionais para Análise de Sementes (ISTA, 2007b). Deve-se ressaltar que essa aproximação, para a Estatística, é desnecessária e desaconselhável, por provocar mudanças, algumas vezes significativas num resultado de um teste.

Apesar de a média do lote 5, para o métodos papel, ter sido considerada estatisticamente igual à dos lotes 1 e 3, o mesmo foi classificado como nível intermediário, em razão do resultado obtido na areia e por apresentar valor médio de germinação entre 60% a 70%, próximo ao valor mínimo para comercialização de sementes da espécie *Raphanus sativus*, que é de 60% (Brasil, 2008).

A utilização de lotes com níveis de qualidades diferentes no processo de validação de métodos é necessária porque o método ideal deve ser robusto para suportar condições adversas de aplicação e também, ser preciso e exato (Waeny, 1980).

Ressalta-se que foram realizados mais dois testes de germinação com objetivo de confirmar as classificação dos lotes encontrados nesse teste, mas seus resultados não estão sendo apresentados nesse trabalho.

6.3 Teste de germinação nos laboratórios selecionados

Observa-se, pelo resultado geral dos oito laboratórios, que, para todos os lotes no substrato areia, a média de germinação foi maior e o coeficiente de variação foi menor em relação ao papel (Tabela 13). Resultados similares, sob as mesmas condições de teste, foram encontrados por Nery et al. (2007) e Nery (2008).

TABELA 13 Resultados médios (%) e coeficientes de variação (%) obtidos no teste de germinação com sementes de cinco lotes de nabo forrageiro para os métodos areia e papel

Substrato	Lab.	Lote 1		Lote 2		Lote 3		Lote 4		Lote 5	
		\bar{y}	CV	\bar{y}	CV	\bar{y}	CV	\bar{y}	CV	\bar{y}	CV
Areia	1	89	4,90	55	8,33	87	6,54	64	0,91	80	3,31
	2	94	4,87	62	6,52	95	1,61	78	5,18	87	2,90
	3	98	0,59	64	6,31	96	0,00	85	1,18	92	0,00
	4	94	2,81	62	7,95	94	3,07	71	16,32	85	4,25
	5	92	2,51	68	5,94	95	1,05	72	9,11	86	4,08
	6	92	4,52	63	11,14	91	6,25	66	18,08	79	5,80
	7	91	6,34	61	8,67	93	3,11	69	8,25	85	0,68
	8	93	5,69	63	16,06	91	12,38	79	14,38	85	13,94
	Geral	93	4,51	62	9,60	93	5,45	73	12,90	86	6,31
Papel	1	88	4,55	53	17,02	89	2,83	59	11,29	83	5,52
	2	91	4,96	56	8,80	94	1,06	61	14,13	85	3,78
	3	98	0,59	68	3,06	93	0,62	83	3,69	94	2,13
	4	94	4,87	61	11,16	94	1,06	65	19,66	85	9,51
	5	95	2,19	66	10,61	96	1,59	81	7,44	93	4,06
	6	92	3,49	61	4,34	84	9,55	68	13,26	77	8,45
	7	76	8,36	40	29,63	74	14,04	45	37,47	60	20,88
	8	95	1,05	64	8,27	95	1,61	77	8,52	88	0,66
	Geral	91	7,78	59	17,54	90	9,19	68	21,49	83	13,95

Nos casos de espécies com nível de domesticação inferior, ou com altos índices de infestação com microorganismos, segundo Ashton (2005), recomenda-se a utilização do substrato areia, uma vez que este material possibilita maior taxa e velocidade de germinação. A maior velocidade de germinação pode ser explicada por autores, tais como Lopes et al. (2005), Pacheco et al. (2006) e Nery (2008), como sendo consequência da maior área de contato da semente com a areia, o que propicia maior velocidade na absorção de água. Da mesma forma, as menores médias de germinação obtidas no substrato papel parecem estar associadas, segundo Andrade et al. (2006), a uma menor capacidade de retenção de água deste material ou a proximidade das plântulas no substrato, o que pode facilitar a contaminação por microorganismos.

Para os dois métodos avaliados, o lote 2, de modo geral, apresentou menores médias de germinação em relação aos demais lotes, o que corrobora com a classificação estabelecida no teste de germinação inicial, como sendo este de qualidade inferior.

Em relação ao coeficiente de variação, de maneira geral, observam-se maiores valores para os lotes 2, 4 e 5, que foram considerados em estados mais avançados de deterioração. Resultados semelhantes foram encontrados por Illipront Junior (1997), trabalhando com sementes de soja, que concluiu existir maior variação individual de sementes em lotes com estados progressivos de deterioração em relação aos deteriorados ou aqueles de qualidade superior.

Este resultado era esperado também do ponto de vista estatístico, pois a variância da proporção é função quadrática da própria proporção (p), como pode ser observado pela curva apresentada na Figura 26.

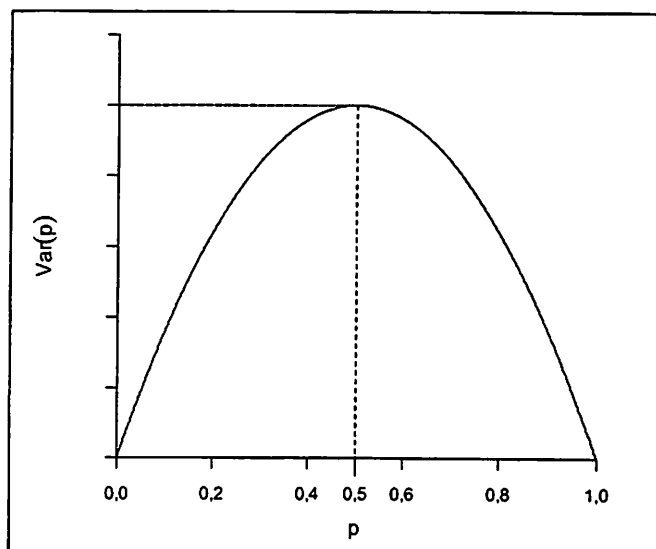


FIGURA 26 Curva da variância, em função da proporção

Pode-se observar que, quanto mais próximo de 0,5 (50%), maior a variância e, conseqüentemente, maior o coeficiente de variação.

Essas considerações acerca dos resultados podem ser verificadas também por meio de uma inspeção visual ao Boxplot (Figura 27). Observa-se a menor variação de resultados para os lotes 1 e 3, de qualidade superior.

Vale salientar que as taxas de germinação encontrados nessa etapa pelos oito laboratórios foram superiores às determinadas na etapa de caracterização dos lotes (Tabela 12). A baixa germinação na etapa anterior foi atribuída a uma alta incidência de microrganismos, comprovada por uma elevada porcentagem de plantas anormais infectadas nos lotes recém-colhidos. Após um período de armazenamento, observou-se redução dessa infestação, o que propiciou, por conseguinte, elevação nas taxas de germinação nesta fase do processo.

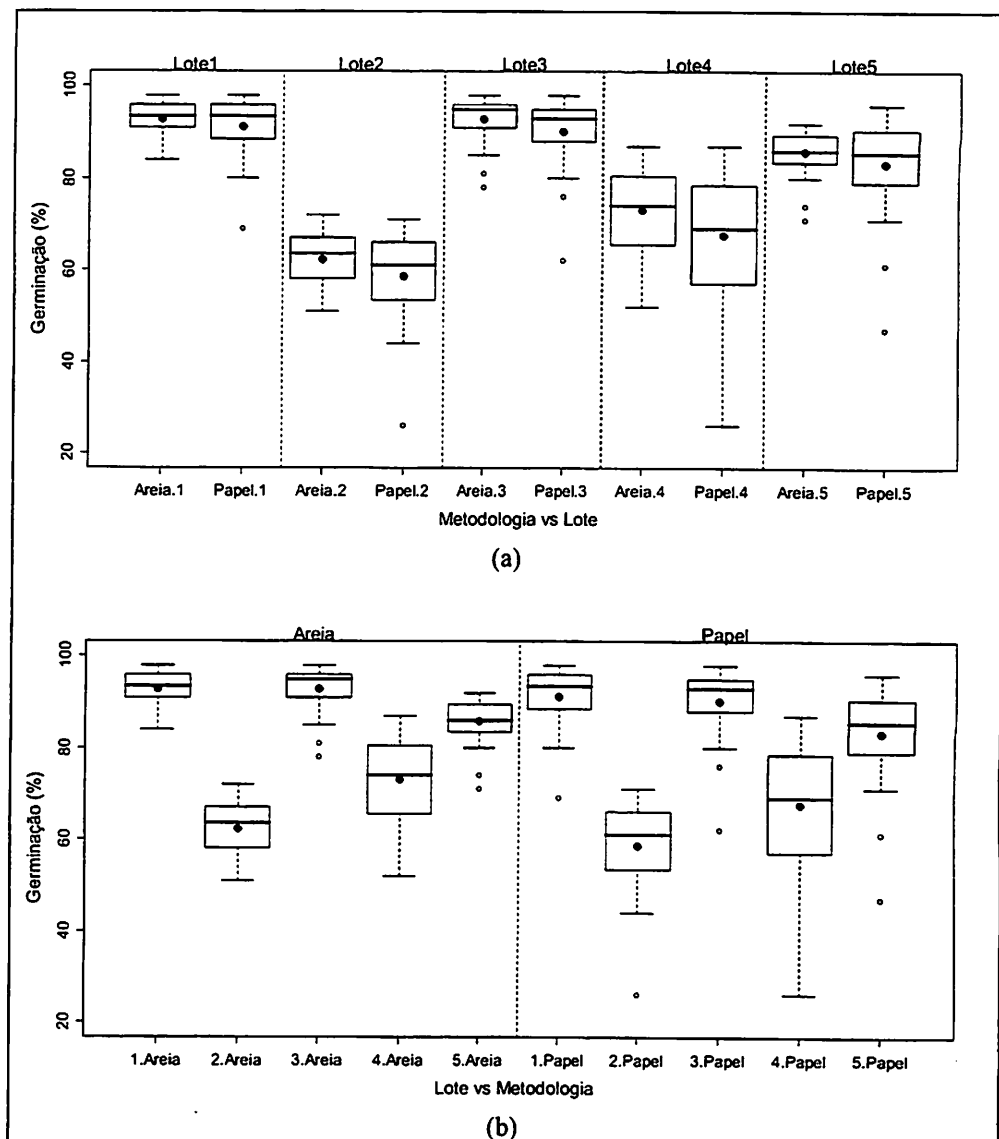


FIGURA 27 Boxplot de acordo com o método para cada lote (a) e lote para cada método (b), com as médias representadas pelo símbolo (•)

6.4 Identificação de valores discrepantes (*outliers*)

Para o método areia, foi identificado, pelo método de Hampel, quantidade menor de *outliers* em relação ao papel, três e sete, respectivamente (Tabela 14).

TABELA 14 Identificação de valores discrepantes nos resultados do teste de germinação em sementes de nabo forrageiro, pelo método de Hampel e do Boxplot para os métodos areia e papel

Substrato	Método de Hampel				Boxplot			
	Lote	Lab	Rep.	Germ.	Lote	Lab.	Rep.	Germ.
Areia	3	1	2	81	3	1	2	81
	3	6	3	85	3	8	3	78
	3	8	3	78	5	6	3	74
					5	8	3	71
Papel	1	7	1	69	1	7	1	69
	2	7	1	26	2	7	1	26
	3	6	1	76	3	6	1	76
	3	7	1	62	3	7	1	62
	3	7	2	80	5	7	1	47
	3	7	3	80	5	7	3	61
	5	7	1	47				

Para o substrato areia dos três *outliers* identificados pelo método de Hampel, apenas dois foram detectados também pelo Boxplot. Já para o papel dos sete lotes detectados pelo Hampel, cinco são concordantes com o Boxplot. Esses resultados reforçam as diferenças entre as duas técnicas apresentadas na subseção 4.6.

Observa-se, ainda com base nos resultados, a eliminação de todas as repetições do laboratório 7 no lote 3 para o substrato papel. De forma geral, este laboratório, principalmente para o substrato papel, obteve resultados bem diferenciados dos demais (Figura 28), tendo sido, inclusive, identificadas

algumas plântulas com danos de umidade, provocados, possivelmente, por condições inadequadas na realização do teste de germinação.

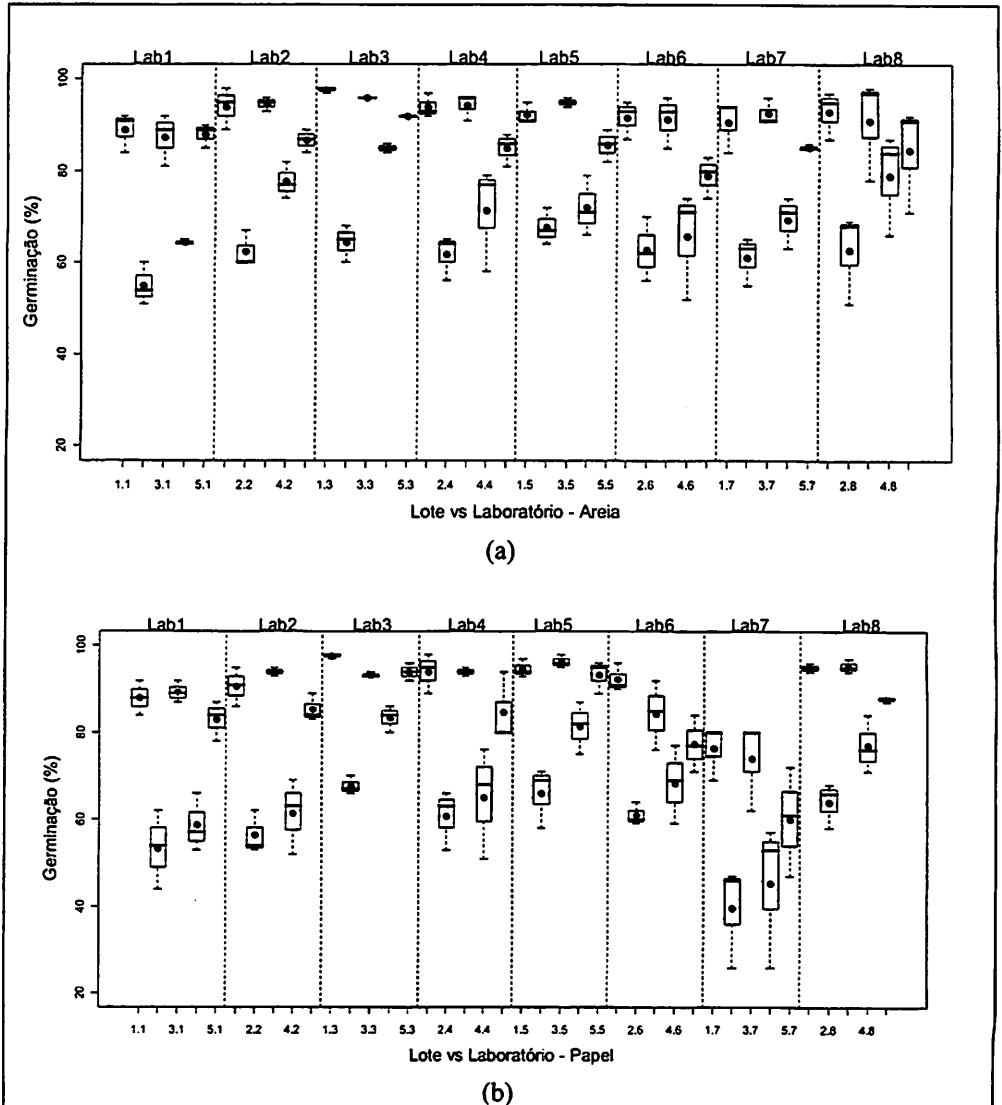


FIGURA 28 Boxplot de acordo com o laboratório para os substratos areia (a) e papel (b), com as médias representadas pelo símbolo (●)

6.5 Identificação de *outlier* nas variâncias

Para o método areia, foram identificados, pelo teste de Levene, *outliers* nos lotes 3, 4 e 5 e, para o papel, apenas no lote 3 (Tabela 15).

TABELA 15 Resultados do teste de Levene, a 1% de significância, para cada combinação de lote e substrato

Substrato	Lote	valor p	Condição	Indicação
Areia	1	0,0915	Variâncias homogêneas	Continuar a análise
	2	0,2698	Variâncias homogêneas	Continuar a análise
	3	0,0099	Variâncias heterogêneas	Repetir o teste
	4	0,0067	Variâncias heterogêneas	Repetir o teste
	5	0,0009	Variâncias heterogêneas	Repetir o teste
Papel	1	0,1419	Variâncias homogêneas	Continuar a análise
	2	0,1602	Variâncias homogêneas	Continuar a análise
	3	0,0073	Variâncias heterogêneas	Repetir o teste
	4	0,1268	Variâncias homogêneas	Continuar a análise
	5	0,0399	Variâncias homogêneas	Continuar a análise

A princípio, esses resultados parecem ser contraditórios ou incoerentes com toda a discussão feita anteriormente, de que houve uma menor variação nos resultados para o substrato areia, já que foram identificados três lotes com variâncias heterogêneas para este método e apenas um lote para o papel. De fato, o que acontece é que comparações dessa natureza (entre medidas descritivas e um teste estatístico) não procedem, uma vez que medidas como coeficiente de variação e desvio padrão são meramente descritivas e mostram apenas tendências de comportamento. Além disso, seus valores são calculados assumindo a hipótese de não haver diferenças entre laboratórios.

Já o teste de Levene é um processo inferencial, em que estão associados grau de confiança e erros tipo I e II, e o que está sendo testado é justamente se os

laboratórios podem ser considerados, quanto à sua variação, homogêneos ou não.

No caso dos lotes 4 e 5, por serem considerados intermediários, espera-se que, na maioria das vezes, este tipo de lote apresente sementes em diferentes estágios de deterioração, propiciando assim uma maior variação no resultados entre repetições e laboratórios (Illipronti Junior, 1997).

Para cada lote em que foram identificados *outliers*, retirou-se o laboratório com maior variância e repetiu-se o teste de Levene e, com exceção do lote 4 – areia, para os demais lotes não foi rejeitada a hipótese nula de variâncias heterogêneas nos laboratórios restantes (Tabela 16).

TABELA 16 Resultados da reaplicação do teste de Levene, 1% de significância, para as combinações de lote e substrato

Substrato	Lote	valor p	Condição	Laboratório retirado
	3	0,0198	Variâncias homogêneas	4 e 7
Areia	4	0,0080	Variâncias heterogêneas	6
	5	0,0845	Variâncias homogêneas	8
Papel	3	0,3372	Variâncias homogêneas	6

No caso do lote 3 - areia, foram retirados, simultaneamente, os laboratórios 4 e 7, por apresentarem variâncias iguais. Para o lote 4 – areia, mesmo com a retirada do laboratório 6, foi necessário reapplicar o teste de Levene pela segunda vez, sendo excluído também o laboratório 4 (Figura 29). Finalmente, as variâncias dos laboratórios restantes para este lote foram consideradas homogêneas (valor p = 0,0145), salientando que são retirados apenas os resultados dessas combinações lote-laboratório e não laboratório como um todo.

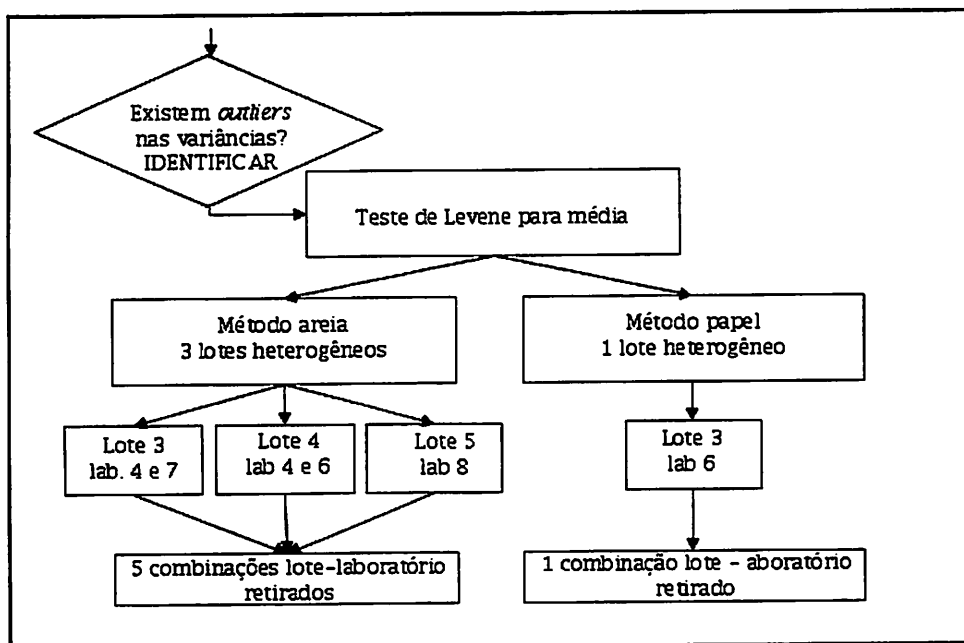


FIGURA 29 Resumo esquemático dos resultados do teste de Levene para os substratos areia e papel, a 1% de significância

6.6 Análise de variância

Com as retiradas de *outliers*, o delineamento experimental passou a ser não balanceado. Logo, para a obtenção dos quadrados médios, foram utilizadas as somas de quadrados tipo III (Tabela 17).

TABELA 17 Resumo da análise de variância e teste de Shapiro-Wilk para a variável porcentagem de germinação em substrato areia e papel, de acordo com oito laboratórios, em ensaios com sementes de cinco lotes de nabo forrageiro

Sub.	Fonte de variação	G.L.	Quadrado médio	F	Valor p
Areia	Laboratório	7	122,8865	5,27	<0,0001
	Lote	4	3.912,8757	167,94	<0,0001
	Resíduo	90	23,2987		
	C.V (%)		5,94		
	Shapiro – Wilk (valor p)		0,4583		
Papel	Laboratório	7	717,1304	19,05	<0,0001
	Lote	4	4.588,0339	121,87	<0,0001
	Resíduo	99			
	C.V (%)		7,78		
	Shapiro – Wilk (valor p)		0,0001		

Pelos resultados da análise de variância, para os dois substratos, foram identificadas diferenças estatisticamente significativas entre os efeitos de laboratórios.

Algumas hipóteses podem ser levantadas para explicar tais diferenças nos resultados, como, por exemplo, a alta variação no estágio de maturação das sementes no momento da colheita, grande número de sementes imaturas, mal formadas e chochas (Bittencourt et al., 1991). Uma possível dificuldade na análise e interpretação do teste pelos laboratórios é o fato de ser uma planta pouco avaliada e conhecida pelos mesmos ou, ainda, o excesso de microrganismos, pois já foi constatada alta associação entre patógenos e nabo forrageiro (Nery, 2008). Ou, ainda, o fato de que nem todos os laboratórios selecionados são credenciados pelo MAPA ou tem um programa ativo de controle de qualidade.

Ao contrário dos laboratórios, era previsível a presença de diferenças estatisticamente significativas para o efeito dos lotes, já que, por princípio,

inicia-se o processo de validação com lotes de qualidades diferentes (ISTA, 2007a).

Pelo teste de Shapiro-Wilk, a 1% de significância, foi rejeitada a hipótese de normalidade nos resíduos para o substrato papel, o que pode ser confirmado pela inspeção visual do gráfico dos resíduos padronizados pelos quantis teóricos (Figura 30). Observa-se que os maiores desvios de normalidade encontram-se nas caudas inferiores.

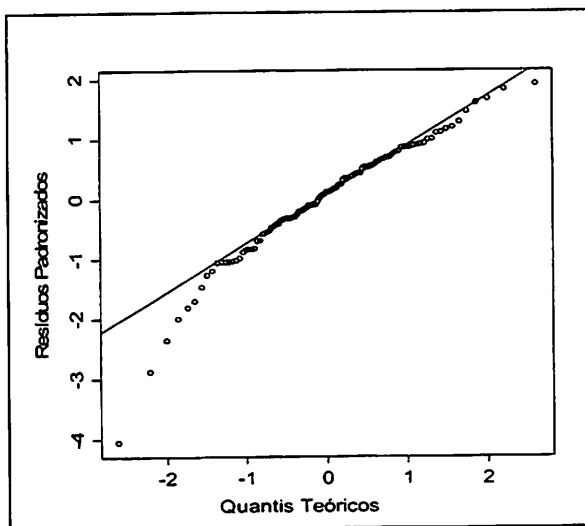


FIGURA 30 Gráficos para a verificação das suposições de normalidade dos resíduos para o substrato papel

6.7 Avaliação da repetibilidade, reprodutibilidade e dos valores de h e k

1) Repetibilidade

No caso do substrato areia, comparando-se a amplitude de cada laboratório (Lr_j) com os limites críticos de repetibilidade por lote, grau de confiança de 99%, pôde-se verificar que, com exceção dos laboratórios 4 e 7, em

relação ao lote 3 e laboratório 8 para o lote 5, todos os demais apresentaram repetibilidade aceitável (Tabela 18).

De fato, esses laboratórios que não apresentaram repetibilidade não tinham sido incluídos no cálculo do limite crítico, já que foram retirados da análise em etapas anteriores, sendo reapresentado apenas para confirmar a situação de não adequabilidade dos seus resultados, para os lotes mencionados, no processo de validação.

TABELA 18 Amplitude (Lr_{ji}), limite crítico, desvio padrão de repetibilidade e resultados da determinação da condição de repetibilidade (Re) ou não (Nr) para os laboratórios em cada lote no substrato areia

Laboratório		Lote				
		1	2	3	4	5
1	Lr_{j1}	8	9	3	1	5
	Condição	Re	Re	Re	Re	Re
2	Lr_{j2}	9	7	3	8	5
	Condição	Re	Re	Re	Re	Re
3	Lr_{j3}	1	8	0	2	0
	Condição	Re	Re	Re	Re	Re
4	Lr_{j4}	5	9	5	21	7
	Condição	Re	Re	Nr	Re	Re
5	Lr_{j5}	5	8	2	13	7
	Condição	Re	Re	Re	Re	Re
6	Lr_{j6}	8	14	3	22	9
	Condição	Re	Re	Re	Re	Re
7	Lr_{j7}	10	10	5	11	1
	Condição	Re	Re	Nr	Re	Re
8	Lr_{j8}	10	18	1	21	21
	Condição	Re	Re	Re	Re	Nr
Limite crítico		16,32	23,58	4,93	24,51	11,79
s_r		4,05	5,85	1,34	6,08	2,93

Apesar de os laboratórios 4 e 6, com relação ao lote 4, estarem na mesma situação dos outros três citados anteriormente, ainda conseguiram

apresentar condições de repetibilidade. A explicação parece estar associada à alta variabilidade apresentada por este lote (Figura 28(a)).

Já para o substrato papel, as exceções foram os laboratórios 6 e 7, para o lote 3, que também tinham sido retirados da análise nas outras etapas (Tabela 19).

TABELA 19 Amplitude (Lr_{ji}), limite crítico, desvio padrão de repetibilidade e resultados da determinação da condição de repetibilidade (Re) ou não (Nr) para os laboratórios em cada lote no substrato papel

Laboratório		Lote				
		1	2	3	4	5
1	Lr_{11}	8	18	5	13	9
	Condição	Re	Re	Re	Re	Re
2	Lr_{12}	9	9	2	17	6
	Condição	Re	Re	Re	Re	Re
3	Lr_{13}	1	4	1	6	4
	Condição	Re	Re	Re	Re	Re
4	Lr_{41}	9	13	2	25	14
	Condição	Re	Re	Re	Re	Re
5	Lr_{15}	4	13	3	12	7
	Condição	Re	Re	Re	Re	Re
6	Lr_{16}	6	5	16	18	13
	Condição	Re	Re	Nr	Re	Re
7	Lr_{17}	0	1	18	31	11
	Condição	Re	Re	Nr	Re	Re
8	Lr_{18}	2	10	3	13	1
	Condição	Re	Re	Re	Re	Re
Limite crítico		12,33	22,34	6,00	38,64	19,74
s_r		3,13	5,66	1,49	9,59	5,00

2) Reprodutibilidade

Quanto às condições de reprodutibilidade, no caso do substrato areia, comparando-se a amplitude de cada repetição (LR_j) com os limites críticos por

lote, grau de confiança de 99%, pôde-se verificar que todos os lotes apresentaram reprodutibilidade aceitável (Tabela 20). No caso do substrato papel, a exceção foi novamente o lote 4.

TABELA 20 Amplitude (Lr_{jk}), limite crítico, desvio padrão de reprodutibilidade e resultados da determinação da condição de reprodutibilidade (Rr) ou não (NR) para os lotes em cada repetição nos substratos areia e papel

Substrato	Lote		Repetição			Limite crítico	s_R
			1	2	3		
Areia	1	LR_{1k}	14	7	13	16,94	4,20
		Condição	Rr	Rr	Rr		
	2	LR_{2k}	14	16	21	23,99	5,96
		Condição	Rr	Rr	Rr		
	3	LR_{3k}	6	4	7	8,49	2,31
		Condição	Rr	Rr	Rr		
	4	LR_{4k}	23	23	20	34,92	8,66
		Condição	Rr	Rr	Rr		
	5	LR_{5k}	12	9	18	18,20	4,52
		Condição	Rr	Rr	Rr		
Papel	1	LR_{1k}	8	18	18	22,12	5,61
		Condição	Rr	Rr	Rr		
	2	LR_{2k}	25	24	24	31,76	8,05
		Condição	Rr	Rr	Rr		
	3	LR_{3k}	6	6	9	10,64	2,64
		Condição	Rr	Rr	Rr		
	4	LR_{4k}	61	34	33	60,46	15,00
		Condição	NR	Rr	Rr		
	5	LR_{5k}	21	23	35	36,13	9,16
		Condição	Rr	Rr	Rr		

3) Valores de h e k

Pela observação dos gráficos dos valores de h e k para o substrato areia (Figura 31), verifica-se que os h e k críticos a 1% não foram ultrapassados, evidenciando a consistência dos resultados de cada lote por laboratório.

Ocorreram apenas uma tendência à subestimação pelo laboratório 1 e superestimação pelo laboratório 3, comparados com a média geral, de acordo com os valores de h e maior variabilidade entre os lotes para o laboratório 8, de acordo com os valores de k.

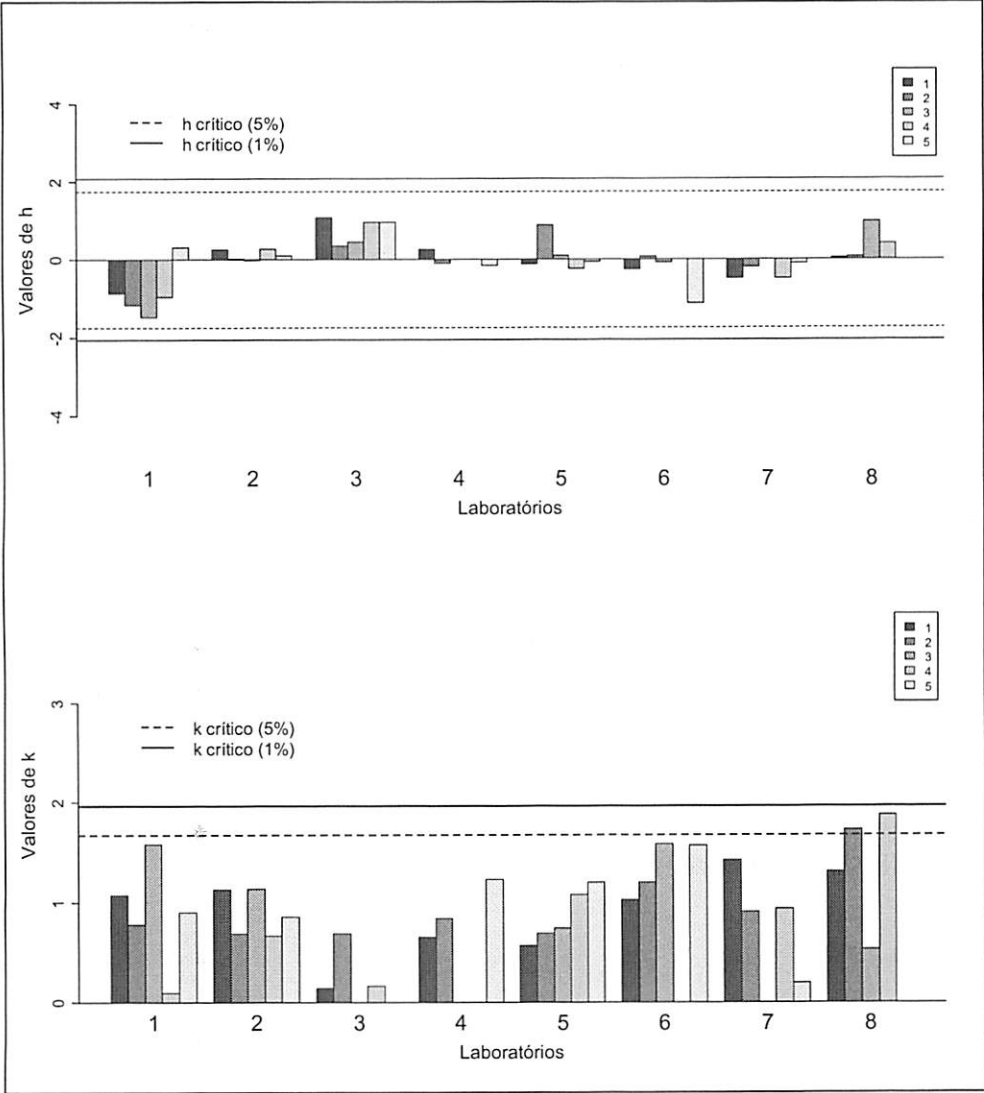


FIGURA 31 Gráficos dos valores de h e k, para o substrato areia

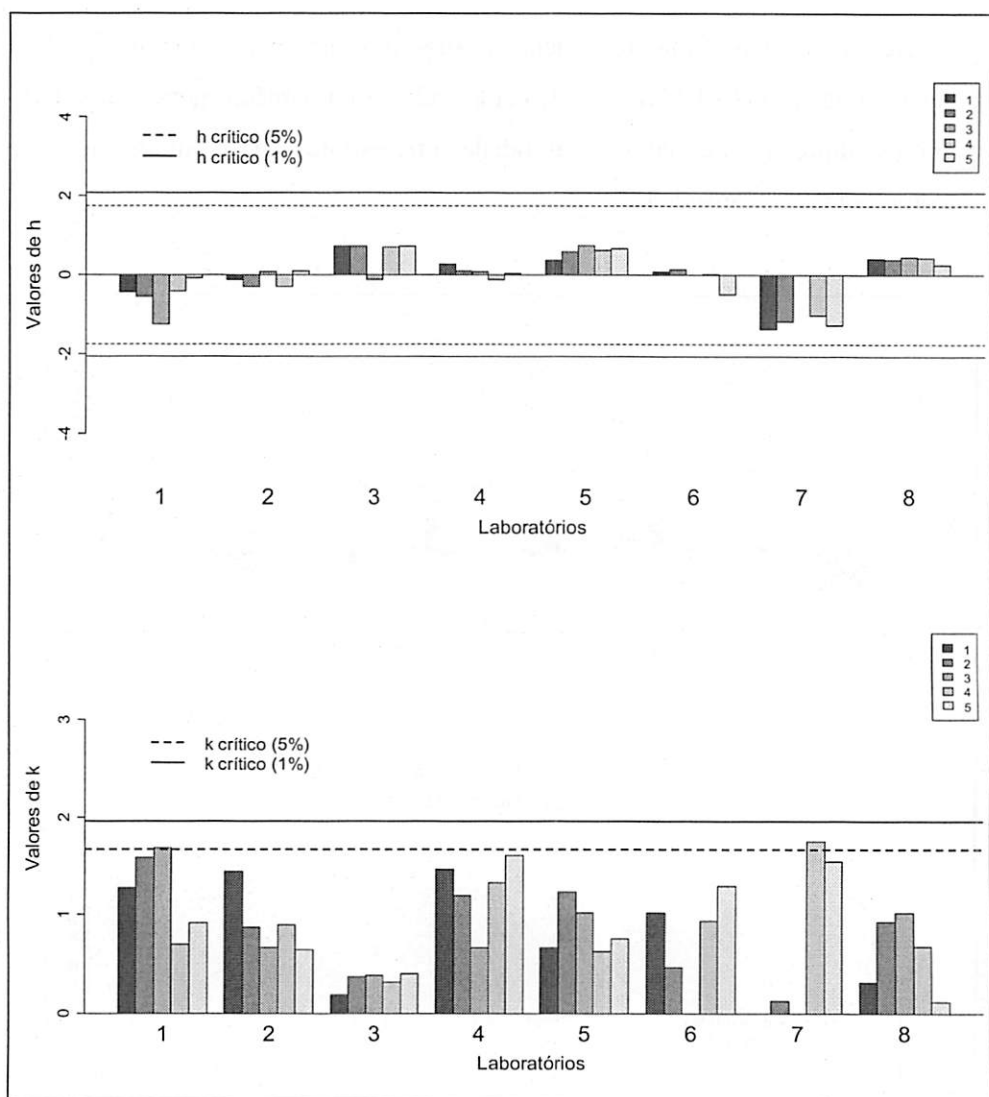


FIGURA 32 Gráficos dos valores de h e k, para o substrato papel

Para o substrato papel verifica-se também que os h e k críticos a 1% não foram ultrapassados (Figura 32), com tendência à subestimação dos resultados pelos laboratórios 1 e 7, superestimação pelos laboratórios 3 e 5, comparados

com a média geral, baseado nos valores de h e maior variabilidade entre os lotes dos laboratórios 1 e 7, de acordo com os valores de k.

6.8 Comparação dos métodos

De acordo com o teste F, não existiram evidências estatísticas para considerar diferenças na precisão nem dentro dos laboratórios e nem no geral, para os dois métodos em cada lote (Tabela 21).

TABELA 21 Resultados da comparação dos dois métodos pelo teste F, a 1% de significância

Comparação	Lote	Substrato	F calculado	F inferior	F superior
Dentro dos laboratórios	1	Areia Papel	0,59	0,32	3,22
	2	Areia Papel	0,94	0,32	3,22
	3	Areia Papel	1,23	0,25	4,07
	4	Areia Papel	2,49	0,29	3,86
	5	Areia Papel	2,92	0,30	3,47
Geral	1	Areia Papel	6,93	0,11	8,88
	2	Areia Papel	6,68	0,11	8,88
	3	Areia Papel	1,34	0,07	14,94
	4	Areia Papel	3,39	0,10	14,20
	5	Areia Papel	4,81	0,11	10,79

Além dos resultados estatísticos, cada laboratório preencheu uma ficha de avaliação (Anexo A) na qual foram comparados os dois métodos, sempre

indicando qual dos dois seria mais fácil na execução de alguns procedimentos. Para todas as respostas, o método papel foi eleito, pela maioria dos laboratórios, como o mais fácil (Tabela 22).

TABELA 22 Respostas para a ficha de avaliação sobre facilidade dos dois métodos, de acordo com a etapa de execução

Etapa	n° de laboratórios		
	Areia	Papel	Ambos
Pré-montagem	0	7	0
Montagem	1	5	1
Manipulação	1	6	0
Avaliação	2	4	1

* Apenas um laboratório não enviou a ficha de avaliação.

As principais dificuldades apontadas por alguns laboratórios foram: na pré-montagem, o método areia foi de difícil manipulação e esterilização; já na avaliação, o papel possibilitou maior incidência de fungos nas sementes.

Os comentários gerais enviados pelos os laboratórios foram:

- “Embora o método papel tenha sido de mais fácil montagem à manipulação, observou-se que as plântulas estavam mais definidas no método da areia”;
- “Como as sementes estavam sem tratamento, durante a avaliação do teste de germinação em papel, verificou-se uma maior contaminação de fungos”;
- “Para a nossa rotina, preferimos o papel, por questão de facilidade”;
- “Para manipulação, a montagem em papel é mais fácil. Já para avaliação, é mais fácil em areia.”

6.9 Análise discriminante

Os cinco lotes foram divididos em três grupos, cada grupo correspondendo aos níveis de qualidade: grupo τ_1 – inferior (lote 2); grupo τ_2 – intermediário (lotes 4 e 5) e grupo τ_3 – superior (lotes 1 e 3).

Para a utilização da técnica de análise discriminante, foram retirados, de cada lote, os laboratórios que tinham sido considerados *outliers* pelo teste de Levene.

Pelos resultado do teste de Box M, a 1% de significância, para o método areia (valor $p=0,1258$) pareceu não haver evidências para rejeitar a hipótese de igualdade das matrizes de variâncias e covariâncias, sugerindo a utilização da regra linear para a classificação. Já no caso do papel (valor $p=0,0004$), esta hipótese foi rejeitada e a primeira indicação seria utilizar a função quadrática (Barroso & Artes, 2003).

Apesar dos resultados do teste de Box M, foram determinadas, para os dois métodos, as funções discriminantes lineares de Fisher. A justificativa da utilização da regra linear, no caso do papel, será apresentada posteriormente.

Para avaliar a função de classificação, foi utilizado o processo de validação cruzada, obtendo-se as matrizes de confusão para os dois métodos (Tabelas 23 e 24).

TABELA 23 Matriz de confusão com validação cruzada – regra linear de Fisher, para os resultados do teste de germinação em sementes de nabo forrageiro, para o método areia

Grupo verdadeiro	Grupo classificado			total
	τ_1	τ_2	τ_3	
τ_1	8	0	0	8
τ_2	2	8	3	13
τ_3	0	0	14	14
Total	10	8	17	35

Taxas de erro	Areia	Papel
TEF (12)	15,38	18,75
TEF (13)	0,00	0,00
TEF (21)	0,00	25,00
TEF (23)	0,00	13,33
TEF (31)	0,00	0,00
TEF (32)	23,08	43,75
TEF (1)	0,00	18,75
TEF (2)	38,46	62,50
TEF (3)	0,00	15,38

TABELA 25 Taxas de erro (%) para classificações pelo processo de validação cruzada – regra linear, para os métodos areia e papel

Com os resultados das matrizes de confusão obtiveram-se, para os dois métodos, as três taxas de erro de se classificar: no geral uma observação incorretamente - TEF; uma observação do grupo t_i no grupo t_j - TEF (f_{ij}) e uma observação do grupo t_j incorretamente - TEF (f_j) (Tabela 25).

Grupo verdadeiro	t_1	t_2	t_3	total
Grupo classificado	t_1	t_2	t_3	
t_1	6	2	0	8
t_2	3	6	7	16
t_3	0	2	13	15
Total	9	10	20	39

TABELA 24 Matriz de confusão com validação cruzada – regra linear de Fisher, para os resultados do teste de germinação em sementes de nabo forrageiro, para o método papel

Para o método areia, somente o grupo 2 apresentou taxa de erro de classificação, o que era em parte esperado por serem lotes de qualidade intermediária (Illipronti, 1997).

Pelos resultados apresentados na matriz de confusão para o substrato areia (Tabela 23), verifica-se uma tendência do grupo 2, tanto à sub como à superestimação, uma vez que foram duas observações classificadas no grupo 1 e três no grupo 3. E, desses cinco valores confundidos, quatro deles são observações do lote 4, reforçando também o que já foi dito sobre este lote, no que se refere à sua alta variabilidade em relação aos demais lotes.

Para o método papel, a primeira consideração a ser feita é que a escolha pela utilização da regra linear em detrimento da regra quadrática vem do fato de que a taxa de erro geral obtida foi menor, já que, no caso quadrático, a TEE seria igual a 46,15%.

Esses melhores resultados da regra linear, mesmo que as matrizes de variâncias e covariâncias sejam diferentes, podem estar associados a desvios de normalidade. De acordo com Johnson & Wichern (1998), a regra quadrática é mais afetada pela falta de normalidade multivariada do que a regra linear. Já o método de validação cruzada, segundo Giri (2004), não é muito sensível à violação de normalidade.

Essa suposição de rejeição à hipótese de normalidade multivariada foi verificada no caso do grupo 3 do método papel (valor $p = 0,00001$), pelo teste de Shapiro Wilk multivariado de Royston (Royston, 1983), a 1% significância.

Comparando-se os dois métodos, verifica-se que os três tipos de taxa de erro para o substrato papel foram bem superiores ao da areia, evidenciando um fato já constatado de uma variação maior dos resultados obtidos no papel. Segundo Ferreira (2008), o desempenho de uma regra de classificação depende das distâncias dos grupos e essa distância depende da variabilidade. Quanto

menor for a dispersão do grupo, menores serão os erros de classificação incorreta.

Pode-se observar que, independente do método, o grupo 2 apresentou as maiores taxas de erro de classificação, o que parece estar associado a uma maior dispersão nos resultados obtidos num lote de qualidade intermediária. Como afirma Ferreira (2008), “por melhores que sejam os métodos matemáticos uma boa regra de decisão dependerá de outros fatores exógenos ao controle.”

No geral, foi possível sustentar a confirmação da classificação feita a priori, dessa forma os métodos puderam ser considerados de fato robusto, por discriminar os lotes com diferentes níveis de qualidade.

6.10 Discussão geral

Para os dois métodos, os resultados indicaram que os procedimentos forneceram exatidão, robustez, reprodutibilidade e repetibilidade. Assim, numa avaliação global, em termos estatísticos e considerando o nível de significância de 1% para todos os testes e análises, verifica-se que métodos com a utilização dos substratos areia e papel, temperatura alternada 20°-30°C, podem ser validadas para o teste de germinação em sementes de nabo forrageiro (Figura 33).

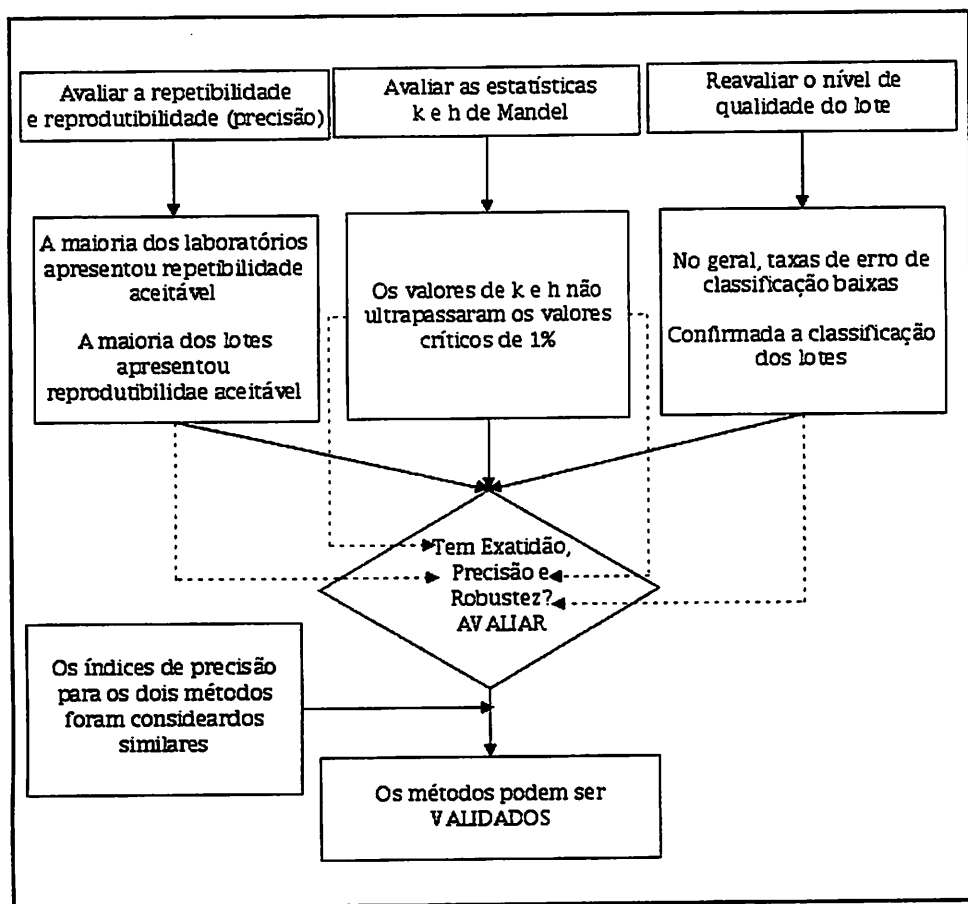


FIGURA 33 Resumo esquemático dos resultados para avaliação da exatidão, precisão e robustez no processo de validação dos métodos que utilizam substratos areia e papel, temperatura alternada 20°-30°C, para o teste de germinação em sementes de nabo forrageiro

Em termos fitotécnicos, observou-se que a eficiência do método depende do nível de deterioração do lote. Aqueles de qualidade intermediária apresentam maior variação de resultados entre e dentro dos laboratórios. Mas, apesar de não ter apresentando, como dito, diferenças estatísticas entre os dois métodos, parece haver uma preferência dos laboratórios pelo uso do substrato papel.

Outra consideração a ser feita é quanto à utilização do nível de significância de 1% em detrimento, por exemplo, de um valor mais tradicional, como 5%. A justificativa está associada ao fato que, no processo de validação, o controle do erro tipo I é importante e necessário, mesmo ocasionando um aumento no erro tipo II. Além disso, para a maioria dos testes estatísticos, existe uma dificuldade metodológica de definir o erro tipo II.

Nesse contexto de validação, deve-se diminuir o risco de considerar que um laboratório ou método estejam fora do padrão quando de fato não estejam, isto é, rejeitar hipóteses nulas, como os laboratórios têm repetibilidade, os métodos têm mesma precisão, quando elas são verdadeiras. Por isso, a opção de tornar os testes mais conservativos¹².

Para o teste Scott-Knott, apresentado na seção 6.2, optou-se por considerar o nível de significância de 5%, tornando o teste mais liberal¹³. Naquele contexto, um maior controle do erro tipo II era importante, já que caracterizar o lote com um determinado nível de qualidade, sem que de fato seja a classificação verdadeira, pode até mesmo inviabilizar o processo de validação. Como dito, ter lotes com características diferentes é um pré-requisito para este tipo de processo.

¹² Um teste é conservativo quando o erro tipo I é controlado de forma excessiva, aumentando-se a taxa de erro tipo II e diminuindo-se o poder do teste (Carmer & Swanson, 1973). Em um teste conservativo, a probabilidade de se encontrar um resultado significativo falso (erro tipo I) é inferior ao valor α estabelecido (Snedecor & Cochran, 1980).

¹³ Um teste é liberal quando se aumenta a taxa de erro tipo I, diminuindo a taxa de erro tipo II e aumentando o poder do teste (Machado et al., 2005).

7 CONCLUSÕES

As técnicas estatísticas propostas contribuem para a melhoria e o aperfeiçoamento das análises no processo de validação, no sentido de que:

- o teste de Levene para média permite identificar alguns laboratórios discrepantes que não são considerados no uso do teste de Cochran;
- a determinação dos limites críticos de repetibilidade e reprodutibilidade permite afirmar mais pontualmente que laboratórios e lotes têm essas condições satisfeitas;
- quanto aos seus índices de precisão, os métodos de germinação podem ser comparadas, com base em resultados de um teste estatístico;
- a técnica da Análise Discriminante auxilia o processo de validação de método, sendo instrumentação de confirmação dos resultados obtidos nos laboratórios relativos a classificação de um lote quanto ao seu potencial de germinação. Pode-se então considerar este resultado como um indicativo que o método a ser validado seja de fato robusto, por discriminar lotes com diferentes níveis de qualidade.

O controle da taxa de erro tipo I, em todos os testes estatísticos, é determinante para a viabilidade do processo de validação de método.

Por fim, tanto no aspecto fitotécnico como no estatístico, os métodos para o teste de germinação em sementes de nabo forrageiro com a utilização dos substratos areia e papel temperatura alternada 20°-30°C podem ser consideradas validadas, pois a análise dos resultados obtidos em oito diferentes laboratórios confirmaram que os mesmos apresentarem exatidão, precisão (repetibilidade e reprodutibilidade) e robustez.

8 REFERÊNCIAS BIBLIOGRÁFICAS

AGUIAR, I.B.; PIÑA-RODRIGUES, F.C.M.; FIGLIOLIA, M.B. **Sementes florestais tropicais**. Brasília: ABRATES, 1993. 350 p.

ALMEIDA, A. de; ELIAM, S; NOBRE, J. Modificações alternativas aos testes de Levene e de Brown e Forysthe para igualdade de variâncias e médias. **Revista Colombiana de Estadística**, Bogotá, v. 31, n. 2, p. 241-260, 2008.

ANDRADE, A.C.S.; PEREIRA, T.S.; FERNANDES, M.J.; CRUZ, A.P.M.; CARVALHO, A.S.R. Substrato, temperatura de germinação e desenvolvimento pós-seminal de sementes de *Dalbergia nigra*. **Pesquisa Agropecuária Brasileira**, Brasília, v. 41, n. 3, p. 517-523, mar. 2006.

ASHTON, P.M. Seed germinatio testing. In: SEED technologists training manual. Ithaca: The Society of Commercial Seed Technologists, 2005. 450 p.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Sistema de gestão da qualidade – Fundamentos e Vocabulário**, 2000 (NBR ISO 9000).

ASSOCIATION OF ANALYTICAL COMMUNITIES. International Harmonized Protocol for Proficiency Testing of (Chemical) Analytical Laboratories. **Journal of AOAC International**, v. 76, n. 4, p. 1-27, 1993. Disponível em: <http://www.inmetro.gov.br/kits/ctle05_protocolo.pdf>. Acesso em: 10 dez. 2007.

ASSOCIATION OF OFFICIAL SEED ANALYSTS. **Seed vigour testing handbook**. Washington, 1983. 88 p. (AOSA. Handbook on Seed Testing Contribution, 32)

BANZATTO, D.A.; KONKRA, S.N. **Experimentação agrícola**. 4.ed. Jaboticabal: FUNEP. 2006. 237 p.

BARNETT, V.; LEWIS, T. **Outliers in statistical data**. 3.ed. Chichester: J. Wiley, 1994. 604 p.

BARROSO, L. P.; ARTES, R. Análise multivariada. In: REUNIÃO ANUAL DA RBRAS, 48.; SEAGRO,10., 2003, Londrina. **Anais...** Lavras: UFLA, 2003. p. 151.

BARTLETT, M. S. Properties of sufficiency and statistical tests. **Proceedings of the Royal Statistical Society Series A**, London, n. 160, p. 268–282, 1937.

BEWLEY, J.D.; BLACK, M. **Seeds: physiology of development and germination**. 2.ed. New York/London: Plenum, 1994. 445 p.

BITTENCOURT, J.F.N.; SADER, R.; UNAGRO, M.R.G.; TOLEDO, N.M. P. Maturação fisiológica de sementes de girassol cv. Contisol. **Revista Brasileira de Sementes**, Brasília, v. 13, n. 2, p. 81-85, 1991.

BOULD, A. The distribution of the heterogeneity value H. **Seed Science. & Technology**, Switzerland, v. 3, p. 439-449, 1975.

BRASIL. Ministério da Agricultura e Reforma Agrária. **Regras para análise de sementes**. Brasília: SNDA/DNDV/CLAV, 1992. 365 p.

BRASIL. Ministério da Ciência e Tecnologia. **Programa Nacional de Produção e Uso de Biodiesel**. 2005. Disponível em: <<http://www.biodiesel.gov.br/>>. Acesso em: 21 de jan. 2008.

BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Instrução Normativa Nº 30**. 2008. Disponível em: <http://www.puntofocal.gov.ar/notific_otros_miembros/bra288_t.pdf>. Acesso em: 2 jan. 2009.

BROWN, M. B.; FORSYTHE, A. B. Robust tests for equality of variances. **Journal of the American Statistical Association**, Alexandria, n. 69, p. 364-367, 1974.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2003. 526 p.

CAMPOS, P. M. **Comparação de métodos para a avaliação do controle de qualidade de laboratório de análises de solo**. 1998. 99 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras.

CAMPOS, P. M.; MUNIZ, J.A.; OLIVEIRA, M.S. de; FERREIRA, D.F. Estimativa da repetibilidade e reprodutividade de laboratórios e via de aplicação no controle de qualidade de solo. **Ciência e Agrotecnologia**, Lavras, v. 23, n. 2, p. 404-409, 1999.

- CARLING, K. Resistant outlier rules and the non-gaussian case. **Computational Statistics & Data Analysis**, Netherlands, n. 33, p. 249-258, 2000.
- CARVALHO, N.M.; NAKAGAWA, J. **Sementes: ciência, tecnologia e produção**. Jaboticabal: FUNEP, 2000. 588 p.
- CARMER, S.G.; SWANSON, M.R. An evaluation of ten pairwise multiple comparison procedure by Monte Carlo methods. **Journal American Statistical Association**, Washington, v. 68, n. 341, p. 66-74, Mar. 1973.
- COCHRAN, W.G. The distribution of the largest of a set of estimated variance as a fraction of their total. **Annals of Eugenics**, London, n. 11, p. 47-52, 1941.
- COPELAND, L.; LIU, H. Seed sampling and sub sampling. In: THE SOCIETY OF COMMERCIAL SEED TECHNOLOGISTS. **Seed technologists training manual**, Switzerland, 2005. 450p.
- CUNHA, U. S. da; MACHADO, S. do A.; FIGUEIREDO FILHO, A. Uso de análise exploratória de dados e de regressão robusta na avaliação do crescimento de espécies comerciais de terra firme da Amazônia. **Revista Árvore**, Viçosa, v. 26, n. 4, p. 391-402, ago. 2002.
- DAVIES, P. L.; GATHER, U. The breakdown point- examples and counterexamples. **REVSTAT – Statistical Journal**, v. 5, n. 1, p. 1-17, mar. 2007.
- EURACHEM Guide. **The fitness for purpose of analytical methods. A laboratory guide to method validation and related topics**. Middlesex, 1998. Disponível em: <http://www.eurachem.org/guides/valid.pdf>. Acesso em : dez. 2008.
- FERREIRA, D. F. **Estatística multivariada**. Lavras: UFLA/DEX, 2008. 661 p.
- FISHER, R.A. The statistical utilization of multiple measurements. **Annals of Eugenics**, London, v. 8, p. 376-386, 1938.
- FOSTER, M. das G. S. **Estados brasileiros se articulam para produzir biodiesel**. 2004. Disponível em: <http://www.secti.ba.gov.br/secti.page?module=noticias&action=verNoticia&id=221>>. Acesso em: 13 de jan. 2008.

GNANADESIKAN, R. **Methods for statistical data analysis of multivariate observations**. New Jersey: J. Wiley, 1997. 353 p.

GIRI, N.C. **Multivariate statistical analysis**. 2.ed. New York: M. Dekker, 2004. 558 p.

HAIR, J. F. Jr.; ANDERSON, R.E.; TATHAN, R.L.; BLACK, W.C. **Multivariate data analysis**. 5.ed. Upper Saddle River: Prentice Hall, 2006. 593p.

HAMPEL, F.R. The influence curve and its role in robust estimation. **Journal of the American Statistical Association**, Washington, n. 69, p. 383-393, 1974.

HAMPEL, F.R. The breakdown points of the mean combined with some rejection rules. **Technometrics**, Alexandria, n. 27, p. 95 – 107, 1985.

HELLERSTEIN, J.M. **Quantitative data cleaning for large databases**. 2008. p. 1- 42. Disponível em: < <http://db.cs.berkeley.edu/jmh>>. Acesso em: 20 dez. 2008.

HINES, W. G. S.; O'HARA HINES, R. J. 'Increased power with modified forms of the levene (Med) test for heterogeneity of variance'. **Biometrics**, Arlington , n. 56, p. 451-454, 2000.

HOLANDA, A. **Biodiesel e inclusão social**. Brasília: Câmara dos Deputados, Coordenação de Publicações, 2004. 200 p.

ILLIPRONTI JUNIOR, R.A. **Variation in quality of individual seeds within a seed lot of soybean (Glycine Max (L.) Merrill)**. 1997. 157 p. Thesis (Doctoral in Agronomy) Wageningen Agricultural University, Wageningen.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725/2. Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method**. Genève, 1994.

INTERNATIONAL SEED TESTING ASSOCIATION. **Method validation for seed testing**. Switzerland, 2007a.

INTERNATIONAL SEED TESTING ASSOCIATION. **International rules for seed testing**. Switzerland, 2007b.

INSTITUTO DE PESQUISAS E ESTUDOS FLORESTAIS. **Informativo Sementes IPEF**, abr. 1998. 2 p. Disponível em: <<http://www.ipef.br/especies/germinacaoambiental.html>>. Acesso em: 10 nov. 2008.

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

LACHENBRUCH, P.A.; MICKEY, M.R. Estimation of error rates in discriminant analysis. **Technometrics**, Washington, v. 10, n. 1, p. 1-11, 1968.

LEVENE, H. **Contributions to probability and statistics: Essays in Honor of Harold Hotelling**. In: OLKIN, I.; GHURYE, S. G.; HOEFFDING, W.; MADOW, W.G.; MANN, H.B. (Ed.). Stanford: Stanford University, 1960. p. 278-292.

LIU, H.; SHAH, S.; JIANG, W. On-line outlier detection and data cleaning. **Computers and Chemical Engineering**, United Kingdom, n. 28, p. 1635-1647, 2004.

LOPES, J.C.; CAPUCHO, M.T.; MARTINS FILHO, S.; REPOSSI, P.A. Influência da temperatura, substrato e luz na germinação de sementes de bortalha. **Revista Brasileira de Sementes**, Pelotas, v. 27, n. 2, p. 18-24, dez. 2005.

LUPING, T.; SCHOUENBORG, B. Methodology of Inter-comparison tests and statistical analysis of the results. **SP Swedish National Testing and Research Institute**. SP REPORT. 2000. p. 35. Disponível em: <<http://www.nordicinnovation.net/nordtestfiler/tec482.pdf>>. Acesso em 10 jan. 2008.

MACHADO, A.A.; DEMÉTRIO, C.G.B.; FEREEIRA, D.F.; SILVA, J.G. C. da. Estatística experimental: uma abordagem fundamentada no planejamento e no uso de recursos computacionais. In: REUNIÃO ANUAL DA RBRAS, 50; SEAGRO, 11., 2005, Londrina. **Anais...** Londrina: UEL, 2005. p. 290.

MARCOS FILHO, J. **Fisiologia de sementes de plantas cultivadas**. Piracicaba: FEALQ, 2005. 495 p.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 295 p.

MOOD, A. M.; GRAYBILL, F.A.; BOES, D.C. **Introduction to the theory of statistics**. 3 ed. New York: J. Wiley, 1974. 564 p.

NERY, M.C. **Colheita, beneficiamento e controle de qualidade de sementes de nabo forrageiro** 2008. 180 p. Tese (Doutorado em Fitotecnia) - Universidade Federal de Lavras, Lavras.

NERY, M.C.; CARVALHO, M.L.M. de; GRIS, C.F.; NERY, F.C.; LEAL FILHO, W.R.; FRAGA, A.C. Germinação de sementes de nabo forrageiro cultivar IPR 116. In: CONGRESSO BRASILEIRO DE PLANTAS OLEAGINOSAS, ÓLEOS, GORDURAS E BIODIESEL. 4., 2007, Varginha. **Resumos...** Varginha: UFLA e Prefeitura Municipal de Varginha, 2007. 1CD-ROM.

NOVEMBRE, A.D.L. C. **Estudo da metodologia para a condução do teste de germinação em sementes de algodão (*Gossypium hirsutum* L.) deslindadas mecanicamente**. 1994. 133 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba, SP.

OLIVIERI, J.C. **Método gráfico para interpretação de resultados em programa interlaboratorial: elipse de confiança, técnicas e métodos**. São Paulo: IPT. 1988. 20 p.

PACHECO, M.V.; MATOS, V.P.; FERREIRA, R.L.; FELICIANO, A.L.P.; PINTO, K.M.S. Efeito de temperatura e substratos na germinação de sementes de *Myracrodruon urundeuva* Fr. All, (Anarcadiace). **Revista Árvore**, Viçosa, MG, v. 30, n. 3, p. 359-367, maio/jun. 2006.

PEREIRA, J.C.R. **Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais**. São Paulo: EDUSP, 2004. 156 p.

POPINIGIS, F. **Fisiologia da semente**. Brasília: AGIPLAN, 1985. 289 p.

RAMALHO, M.A. P.; FERREIRA, D.F. ; OLIVEIRA, A.C. **Experimentação em genética e melhoramento de plantas**. Lavras, UFLA, 2000. 303 p.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. R Vienna, Áustria: Foundation for Statistical Computing. 2008. Disponível em: <<http://www.R-project.org>>. Acesso em: 10 jan. 2008.

REIMANN, C.; FILZMOSE, P.; GARRETT, R.G. Background and threshold: critical comparison of methods of determination. **Science of the Total Environment**, Netherlands, n. 346, p. 1-16, 2005.

ROUSSEEUW, P.J.; CROUX, C. Alternatives to the median absolute deviation. **Journal of the American Statistical Association**, Washington, v. 88, n. 424, p. 1273-1283. 1993

ROYSTON, J.P. Some techniques for assessing multivariate normality based on the Shapiro –Wilk W. **Applied Statistics Journal of the Royal Statistical Society – Series C**, London, v. 32, n. 2, p. 121-1333, 1983.

SHEPPARD, J. W.; COCKERELL, V. **International Seed Testing Association handbook of method validation for the detection of seedborne pathogens**. Switzerland, International Seed Testing Association 2000.

SILVA, E.C.; FERREIRA, D.F.; BEAZORTI, E. Avaliação do poder e taxas de erro tipo I do teste de Scott-Knott por meio do método de Monte Carlo. **Ciência e Agrotecnologia**, Lavras, v. 23, n. 2, p. 687-696, 1999.

SNEDECOR, G.W.; COCHRAN, W.G. **Statistical methods**. Iowa: University of Ames, 1980. 507 p.

TUKEY, J.W. **Exploratory data analysis**. Reading, Mass: Addison-Wesley, 1977.

VIEIRA, A. M. Química pesquisa uso de nabo forrageiro na produção de biodiesel. **Boletim Informativo da UFMG**, v. 33, n. 1574, 2007. Disponível em: <<http://www.ufmg.br/boletim/bol1574/sexta.shtml>>. Acesso em 10 jan. de 2007.

VOCABULÁRIO Internacional de Metrologia. **Erro aleatório e erro sistemático**. 2.ed. Brasília: INMETRO/SENAI, 2000.

VORAPONGSATHORN, T.; TAEJAROENKUL, S.; VIWATWONGKASEM, C.

A comparison of type I error and power of Bartlett's test, Levene's test and Cochran's test under violations of assumptions. **Journal Science Technology**, Lusaka, v. 26 n. 4, p. 537-547, 2004.

VUOLO, J. H. **Fundamentos da teoria de erros**. 2.ed. São Paulo: E. Blucher, 2000. 249 p.

WAENY, J.C.C. **Repetitividade e reprodutividade II**. São Paulo: IPT. 1980. 14 p.

WILLIAMS, G. J.; BAXTER, R. A.; He H. X.; HAWKINS S.; GU, L. A comparative study of RNN for outlier detection in data mining. In: IEEE INTERNATIONAL CONFERENCE ON DATA-MINING, 2002, Maebashi City. **Proceedings...** Maebashi City, Japan: CSIRO, 2002. (Technical Report CMIS-02/102).

CAPÍTULO 3
ENSAIOS DE PROFICIÊNCIA

1 RESUMO

O ensaio de proficiência é um dos tipos de programas interlaboratoriais que permitem determinar o desempenho de um laboratório por meio de uma comparação que utiliza medidas em materiais homogêneos, ou similares, em pelo menos dois laboratórios, sob condições predeterminadas. Na área de sementes, para o processo de manutenção do credenciamento ou acreditação, a *International Seed Testing Association* (ISTA) exige a participação dos laboratórios de sementes credenciados em ensaios de proficiência. A técnica estatística utilizada pela ISTA na avaliação dos resultados de um ensaio de proficiência é o método do Z-score. O problema de pesquisa abordado neste capítulo foi investigar e avaliar técnicas estatísticas, além do método do Z-score, que pudessem ser adequadas para analisar os resultados de ensaios de proficiência na área de sementes. Foram avaliados então três procedimentos: método Campos, análise de agrupamento e análise generalizada de Procrustes. Para gerar o banco de dados foram realizadas três simulações, cada uma representando os resultados de um teste de germinação com doze laboratórios participantes. Esses doze laboratórios formaram uma composição de dois grupos discrepantes em relação a algum(ns) parâmetro(s) estatístico(s), como a média e a matriz de variâncias e covariâncias. No caso do método do Z-score, foram calculados, para cada laboratório e simulação, a precisão e a exatidão, e atribuído uma classificação e uma nota. No método Campos também foram calculados para cada laboratório e simulação a precisão e a exatidão e atribuído um selo. Na análise de agrupamento foram construídos dendrogramas e calculadas as correlações cofenéticas para cada simulação. Na análise generalizada de Procrustes, para cada simulação foram construídos elipses de confiança de 95%. As três técnicas propostas, no geral, conseguiram distinguir os dois grupos, mostrando-se como potenciais procedimentos para a utilização em ensaios de proficiência na área de sementes. O método do Z-score não conseguiu separar adequadamente os grupos, sendo necessários maiores estudos para avaliar o desempenho deste procedimento nas situações simuladas.

2 ABSTRACT

Proficiency test is a type of inter-laboratory program that allows determining the performance of a laboratory by comparing to at least other two laboratories through measurements in homogeneous or similar materials under predetermined conditions. In the area of seeds, for the process of maintaining accreditation, the International Seed Testing Association (ISTA) requires the participation of accredited seed laboratories in proficiency tests. The statistical technique used by ISTA for the assessment of proficiency test results is the Z-score method. The research problem approached in this chapter was to investigate and assess statistical techniques, besides the Z-score method, that would prove adequate to analyze the results of proficiency tests in the area of seeds. Therefore, three procedures have been evaluated: Campos method, Cluster Analysis, and Generalized Procrustes Analysis. In order to generate the database three simulations have been carried out, each representing the results of germination test with twelve participating laboratories. The twelve laboratories were arranged in two discrepant groups regarding some statistical parameter(s) such as variance and covariance average and matrix. In the case of Z-score method, precision and accuracy were calculated for each laboratory and simulation, and rating and grade were attributed. In the Campos method, precision and accuracy were also calculated for each laboratory and simulation, and a label was attributed. In the grouping analysis, dendrograms were built and cophenetic correlations were calculated for each simulation. In the Generalized Procrustes Analysis, confidence ellipses of 95% were built for each simulation. The three proposed techniques, in general, managed to distinguish the two groups, presenting themselves as potential procedures to be used for proficiency tests in the area of seeds. The Z-score method did not manage to adequately separate the groups, and further studies become necessary to assess the performance of that procedure in the simulated situations.

3 INTRODUÇÃO

A participação dos laboratórios em programas interlaboratoriais é um dos mecanismos de controle de qualidade dos seus resultados previstos na NBR ISO/IEC 17025 (Association of Analytical Communities, AOAC, 1993). Estes tipos de ensaios comparativos podem auxiliar os laboratórios na detecção e, por conseguinte, na correção de possíveis fontes de erros. Esses erros podem ser aleatórios ou sistemáticos, provenientes de deficiências nos procedimentos de análise, nos equipamentos e nos analistas, dentre outras fontes.

De acordo com a indicação dessa mesma norma, para garantir a qualidade dos resultados de um laboratório, os mesmos devem ser provenientes de um método validado, seguido da estimativa da incerteza de medição, ou seja, de uma avaliação da exatidão e precisão.

O ensaio de proficiência é um dos tipos de programas interlaboratoriais que permitem determinar o desempenho de um laboratório por meio de uma comparação que utiliza medidas em materiais homogêneos, ou similares, em pelo menos dois laboratórios, sob condições predeterminadas.

Para o processo de manutenção do credenciamento ou acreditação, a International Seed Testing Association (ISTA) exige a participação dos laboratórios de sementes credenciados em ensaios de proficiência.

O objetivo dos ensaios de proficiência da ISTA não é identificar os melhores laboratórios de análises de sementes no mundo, mas sim destacar aqueles que não satisfazem ao desempenho mínimo, que é razoável esperar de um laboratório credenciado e determinar se estão, de fato, tomando ações corretivas para que seu desempenho atinja, pelo menos, o nível mínimo (ISTA, 2007).

A técnica estatística utilizada pela ISTA na avaliação dos resultados de um ensaio de proficiência é o método do Z-score. Em outras áreas (química, meio ambiente, solos, etc.) são utilizados, além desse método, procedimentos estatísticos, como, por exemplo, os métodos da elipse de confiança de Youden (Olivieri, 1988) e do Profert (Oliveira et al., 1995).

Atualmente, no Brasil, a Rede Metrológica RS e a Lanagro/RS vêm desenvolvendo um programa de ensaio de proficiência em sementes de soja, com os seguintes itens de ensaios: análise de pureza e determinação de outras sementes por número (determinação do número total por categoria e determinação do número por espécie). O desempenho de cada laboratório será avaliado também pelo método do Z-score (Rede Metrológica, 2008).

O problema de pesquisa abordado neste capítulo foi investigar e avaliar técnicas estatísticas, além da método do Z-score, que pudessem ser adequadas para analisar os resultados de ensaios de proficiência na área de sementes.

4 REFERENCIAL TEÓRICO

4.1 Programas interlaboratoriais: aspectos gerais

De acordo com a norma Associação Brasileira de Normas Técnicas, ABNT NBR ISO 9000 (ABNT, 2000), **qualidade** pode ser definida como sendo a totalidade das propriedades e das características de um produto ou serviço que lhe conferem aptidão para satisfazer às necessidades implícitas ou explícitas dos clientes.

No caso do laboratório de sementes, o serviço prestado é a avaliação da qualidade de sementes, por meio de procedimentos de análise, até a emissão do respectivo boletim de análise.

De acordo com a legislação pertinente, para proceder às análises, os laboratórios baseiam-se nas Regras para Análise de Sementes (Brasil, 1992), que se fundamentam na uniformidade dos procedimentos e especificam os diferentes métodos de análises que devem ser empregados, os tamanhos máximos para os lotes de sementes, o peso mínimo da amostra média ou submetida e da amostra de trabalho e a utilização de tabelas de tolerância para os diferentes tipos de testes em sementes (Zorato, 2005).

No que tange às tabelas de tolerância¹⁴, calculadas por Miles, em 1963, os laboratórios fundamentam-se nas mesmas como um método estatístico para controle interno de qualidade em testes de rotina, bem como para comparar resultados de diferentes laboratórios, uma vez que é possível garantir que os resultados estejam dentro dos padrões exigidos pela RAS.

¹⁴ As tabelas de tolerância determinam a amplitude tolerável para as diferenças entre repetições de um teste com $(1-\alpha)$ de confiança, ou seja, no caso do teste de germinação, são equivalentes as amplitudes de intervalos de confiança unilateral ou bilateral para a proporção.

Contudo, no cálculo dessas tabelas, levou-se em consideração apenas a variação aleatória das amostras obtidas em diferentes laboratórios e, de acordo com os estudos de Jorgensen et al. (1992), a variação entre testes de germinação conduzidos no mesmo laboratório não pode ser explicada exclusivamente pela variação ao acaso da amostragem.

Miles (1963), citado por Oliveira & Cicero (1996), afirma que as diferenças significativas dos resultados no teste ou entre testes de germinação podem ter como causas: o acaso, deficiência de equipamentos, incluindo variação de ambiente dentro do germinador, deficiência de método, deficiência técnica, erros ou inconsistência na distinção entre plântulas normais e anormais, fungos e bactérias, produtos químicos nas sementes, contagens ou registros imprecisos, amostragem não casualizada das sementes para o teste e mudança na porcentagem de germinação entre os testes.

Em consonância com essas afirmações, Rennie & Tomlin (1984) observaram que variações ocorridas em testes de germinação podem ser devido às diferenças normais de amostragem, ou dos procedimentos e condições para o teste, ou, ainda, às variações nas avaliações entre analistas. Esses mesmos autores verificaram também, em seus estudos, que a variação de resultados entre laboratórios foi maior que a variação entre repetições ou entre testes, em um mesmo laboratório.

Nesse sentido, a participação dos laboratórios em trabalhos de comparação pode auxiliar na busca de fontes de erros, para eliminá-los ou minimizar seus efeitos, podendo, assim, auxiliar o controle interno de qualidade laboratorial (Chuí et. al., 2004).

Segundo Olivieri & D'elia (1996), citados por Campos (1998), os programas interlaboratoriais devem permitir aos laboratórios participantes a comparação dos seus resultados, para verificar se a exatidão e a precisão obtidas

em seus processos metrológicos são coerentes com os resultados de outros laboratórios que atuam em um mesmo setor tecnológico.

As comparações interlaboratoriais são utilizadas para determinar características de métodos, materiais ou laboratórios. Se o interesse for conhecer características de métodos, é necessário utilizar materiais adequados e laboratórios competentes para realizar as medições propostas. Se o problema é atribuir valores para certificar materiais de referência, o estudo deverá ser apoiado por laboratórios confiáveis de mais alta qualidade metrológica. Por outro lado, não é certo avaliar o desempenho de laboratórios, a menos que estejam disponíveis amostras homogêneas a serem distribuídas para laboratórios, utilizando-se de métodos validados (Chuí et al., 2002).

Pelo relatado, o termo “Programa Interlaboratorial” é muito abrangente. O planejamento, a organização, a operacionalização e a interpretação dos resultados dependem da finalidade a que se destinam, resultando, por conseguinte, na utilização de diversas técnicas estatísticas para avaliação dos resultados (Quadro 3).

QUADRO 3 Tipos de programas interlaboratoriais (PI), descrição e técnicas estatísticas associadas

Tipos de PI	Objeto de estudo	Descrição	Técnicas estatísticas
Ensaio de proficiência	Laboratório	Avaliação do desempenho de laboratórios, reconhecimento formal da sua competência técnica para um determinado ensaio	Método do <i>Z-score</i>
Compatibilização de resultados	Laboratório	Comparação dos resultados entre laboratórios	Método da Elipse de Confiança de Youden
Estudo de precisão de métodos – estudo colaborativo	Método	Avaliação do desempenho de um método	ANOVA/ISO 5725
Certificação de materiais de referência	Material	Estimativa do “valor real” de uma ou mais amostras em um material	ANOVA/ABNT/ISO 35
Calibração	Equipamento	Calibração de equipamentos, utilizando padrões ou materiais de referência	Método do erro normalizado
Aferição interlaboratorial	Laboratório	Visa o aprimoramento das técnicas operacionais, orientando os laboratórios participantes a atentarem ao método, à calibração ou aferição do instrumental, bem como treinamento do operador	Método da Elipse de Confiança de Youden Método Profert

Fonte: Quadro adaptado de Chui (2006).

As técnicas estatísticas identificadas no Quadro 3, para cada tipo de programa interlaboratorial, são os padrões utilizados na maioria das análises, mas não são únicas, podendo acontecer algumas alterações e ou adaptações, de acordo com a conveniência do estudo. Outro ponto importante a ser ressaltado é

que, em quaisquer das situações mencionadas, o mais complexo numa análise é determinar quais valores devem ser designados para as avaliações estatísticas dos resultados oferecidos pelos laboratórios participantes de um programa de comparação interlaboratorial. Como valor designado entende-se aquele que melhor estima o valor verdadeiro (AOAC, 1993).

Geralmente, em ensaios de proficiência, estes valores são estabelecidos como os "verdadeiros", pelo órgão competente que o patrocina. Os laboratórios que se submetem não questionam os valores designados utilizados na avaliação estatística, por entenderem que a entidade que oferece o programa esteja imbuída de sua autoridade. Já no caso da compatibilização de resultados, estes valores podem ser estabelecidos por consenso.

Nos programas para certificação de materiais de referência, geralmente, são escolhidos os laboratórios de maior hierarquia na escala metrológica, ou seja, laboratórios de metrologia, de instituições públicas reconhecidas, ou especialistas que podem demonstrar sua capacidade metrológica.

A princípio, não existem os valores verdadeiros e, não havendo critério suficiente para defini-los, utilizam-se, por consenso, a média e o desvio padrão do conjunto de resultados apresentados pelos laboratórios, correndo-se o risco de aceitação de quase todos eles (tanto faz usar, por exemplo, o método do Z-score como a Elipse de Confiança de Youden). De acordo com Wong (2005), muitas variáveis influenciam estes valores, como, por exemplo, número de participantes, homogeneidade dos itens de ensaios distribuídos, precisão dos métodos e repetibilidade dos resultados dos laboratórios.

Enfim, pode-se dizer que todos os programas citados têm o propósito de avaliar o desempenho de laboratórios. O que distingue uma situação da outra são os objetivos para os quais a avaliação se propõe. Nesse trabalho, o programa interlaboratorial proposto seguiu as características de um ensaio de proficiência para área de sementes.

4.2 Ensaio de proficiência

Apesar de existirem vários tipos de programas de ensaios de proficiência, como descrito no ABNT ISO/IEC guia 43-1 (ABNT, 1999), todos possuem uma característica em comum: os resultados de um laboratório de ensaio são comparados com aqueles obtidos por um ou mais participantes do programa (AOAC, 1993).

Os benefícios advindos da inserção neste tipo de programa interlaboratorial incluem que o laboratório participante dispõe de uma avaliação externa regular e independente da qualidade de seus resultados, podendo comparar o seu desempenho com o de outros laboratórios semelhantes e os dados obtidos podem servir de subsídio para a implementação de ações preventivas para a melhoria dos procedimentos do laboratório (Instituto Nacional de Metrologia, INMETRO, 2005)

No ensaio de proficiência, os materiais enviados aos laboratórios devem ser estáveis e homogêneos, para que qualquer variabilidade significativa observada não seja atribuída aos mesmos.

No caso da área de sementes, a ISTA promove ensaios de proficiência que são realizados por escopo¹⁵. Os laboratórios credenciados da área são obrigados a participar, sendo permitida também a adesão voluntária de laboratórios credenciados em outros escopos que queiram aferir-se, bem como os não credenciados (ISTA, 2007).

Outra característica dos ensaios de proficiência da ISTA é que são realizadas pelo menos duas avaliações anuais, sendo cada uma delas chamada de

¹⁵ Escopo - Define as atividades de um laboratório para o qual demonstrou a sua competência e que são abrangidas por um sistema que proporciona confiança na capacidade do laboratório para cumprir os critérios estabelecidos. O escopo do credenciamento é um resumo do trabalho do laboratório, mas detalhado o suficiente para tornar a distinção entre os seus diferentes elementos. É baseada em métodos e espécies individuais ou grupos de espécies, conforme o caso (ISTA, 2007).

rodada. Para o teste de germinação, geralmente, realizam-se três rodadas por ano e a avaliação final é feita no período de três anos, ou seja, a cada rodada, o laboratório vai acompanhando o seu desempenho, recebendo classificações parciais e, ao final, ainda recebe uma classificação geral.

Nos ensaios de proficiência da ISTA, o procedimento estatístico utilizado é o método do Z-score. Assim, a proposta deste trabalho foi verificar se outras técnicas estatísticas poderiam auxiliar e/ou aperfeiçoar a análise dos resultados de um ensaio de proficiência na área de sementes. Para tal, foram avaliados mais três procedimentos: método Campos, análise de agrupamento e análise generalizada de Procrustes. O primeiro foi proposto, inicialmente, para a avaliação de um programa interlaboratorial na área de solos e os outros dois são técnicas multivariadas usados em diversos contextos (Figura 34).

Vale ressaltar que é possível utilizar o contexto multivariado no ensaio de proficiência, no sentido de que, para cada laboratório, sempre são distribuídas duas ou mais amostras e essas podem ser tratadas como as variáveis explicativas.

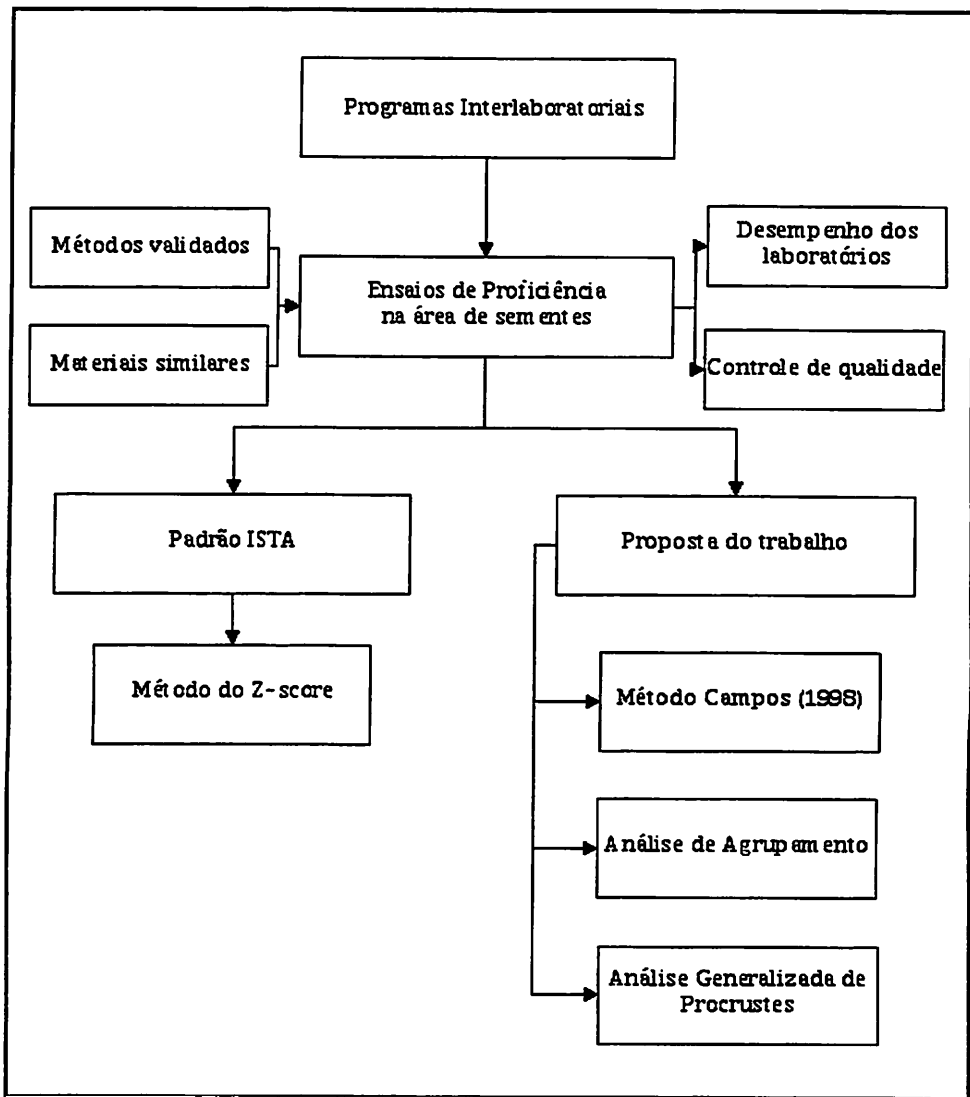


FIGURA 34 Resumo esquemático das técnicas estatísticas utilizadas na avaliação de um ensaio de proficiência na área de sementes

4.3 Método do Z-score

A técnica de Z-score é adotada quando se deseja avaliar o desempenho técnico de laboratórios, principalmente em ensaios de proficiência, fornecendo boa indicação da competência analítica do laboratório. O princípio deste método é calcular Z-scores baseados na determinação do valor de referência aceito, representado pela estimativa da média, após a retirada dos valores discrepantes (ISTA, 2007).

Com os valores do Z-scores são calculados os índices de viés, precisão e exatidão para cada laboratório (Figura 35).

A avaliação de desempenho do laboratório, por exemplo, para o teste de germinação, segue o seguinte critério: $|Z| \leq 2$ indica desempenho satisfatório; se $2 < |Z| < 3$, o desempenho é questionável e, quando $|Z| \geq 3$, o desempenho é considerado não satisfatório.

No programa do ensaio de proficiência da ISTA, estes critérios são adaptados, pois, a cada rodada, são enviadas três amostras para cada laboratório, senso assim somados os Z-scores calculados para cada uma delas e atribuída a seguinte classificação:

- A - soma absoluta do Z-scores $\leq 3,5$;
- B- $3,5 < \text{soma absoluta do Z-scores} \leq 5,3$;
- C - $5,3 < \text{soma absoluta do Z-scores} \leq 7,0$ e
- BMP (*below minimum performance*) - soma absoluta do Z-scores $> 7,0$.

Estas classificações são convertidas em “notas”, em cada rodada e a classificação, ao final de seis rodadas, fica definida como descrito na Tabela 26.

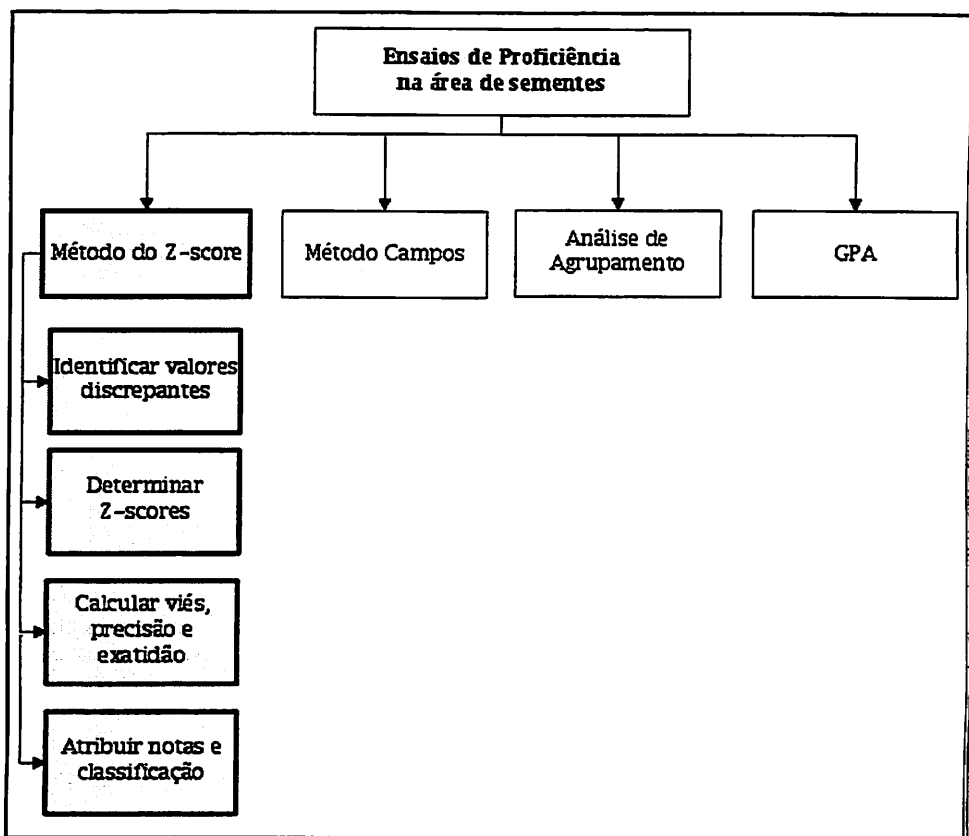


FIGURA 35 Esquema de um ensaio de proficiência na área de sementes, com ênfase no Método do Z-score

TABELA 26 Classificação geral de todos os testes, baseados nas notas após 6 rodadas

Classificação por rodada	Nota atribuída	Intervalo	Classificação geral
A	5	28-30	A
B	4	21-27	B
C	3	16-20	C
BMP	0	Abaixo de 16	BMP

Fonte: The ISTA Proficiency Test Program (ISTA, 2007).

4.4 Método Campos

O método denominado Campos, neste estudo, se refere à proposta metodológica estatística apresentada por Campos (1998). De fato este autor, já tinha realizado algumas alterações no método originalmente denominado de PROFERT, proposto por Oliveira et. al. (1995).

Segundo Campos et. al. (1999), o método do PROFERT pode ser utilizado para avaliar resultados de um programa de controle de qualidade das análises de laboratórios de solos, por meio de técnicas da estatística básica.

Nos estudos de Campos (1998), foram comparados os métodos de Waeny e PROFERT, associados ao método gráfico da elipse de confiança de Youden em dois conjuntos de dados: um ensaio real e outro simulado no contexto de um programa interlaboratorial na área de solos, sendo verificada uma maior adequação nas situações propostas do método PROFERT em detrimento das outras duas técnicas.

O método PROFERT baseia-se em propriedades da distribuição normal de probabilidade, para a construção de faixas de confiança para mediana (considerada como valor de referência), obtendo-se a estimativa da exatidão de um laboratório. A precisão pode ser determinada por meio de um índice de imprecisão, de acordo com um intervalo de confiança para a razão de variâncias.

No PROFERT, a exatidão e a precisão são quantificadas, sendo considerados exatos todos os laboratórios que possuam, pelo menos, 67% de exatidão e precisos se apresentam também, no mínimo, 67% de índice de precisão. Ao final, recebem uma atribuição de selos com os seguintes critérios: nenhum selo, se forem inexatos; selo prata, se forem exatos, porém, imprecisos e selo ouro, classificados como exatos e precisos (Figura 36).

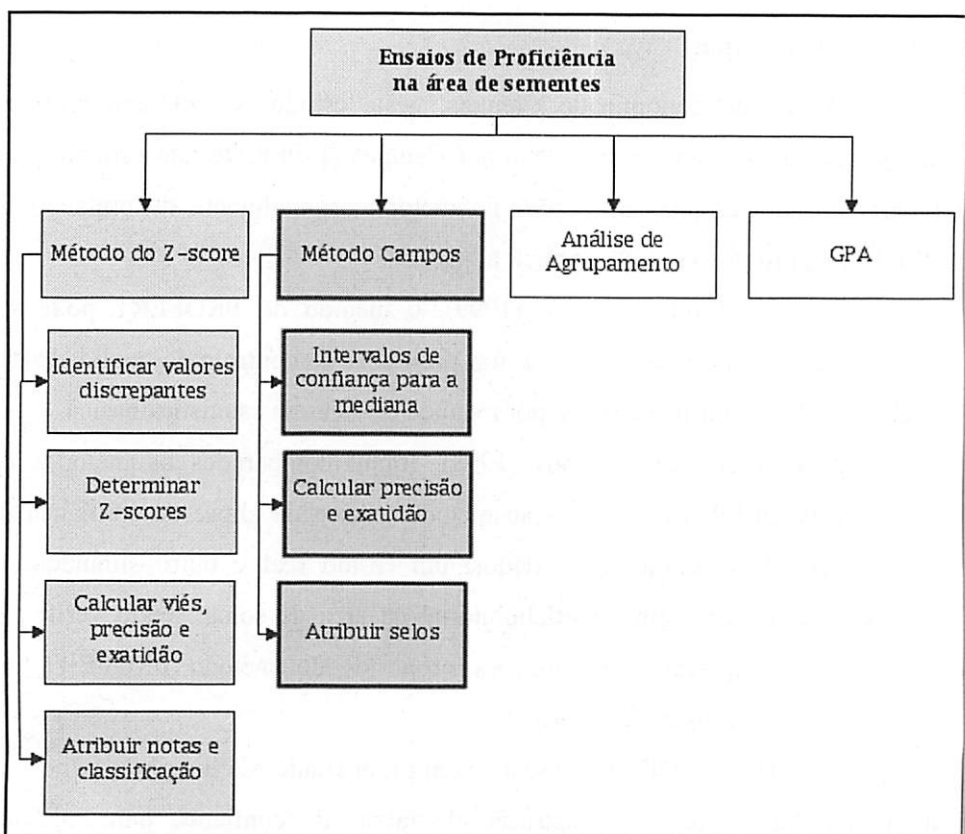


FIGURA 36 Esquema de um ensaio de proficiência na área de sementes, com ênfase no método Campos

4.5 Análise de agrupamento

Segundo Mingoti (2005), o objetivo da análise de agrupamento é dividir os elementos de uma amostra ou população em grupos, de forma que os elementos pertencentes a um mesmo grupo sejam similares entre si com relação às variáveis avaliadas e as observações em grupos diferentes sejam heterogêneas para estas mesmas características.

Nenhuma suposição é feita a respeito do número de grupos ou de sua estrutura. Assim, o princípio dessa técnica é agrupar elementos de acordo com suas similaridades.

Para esse procedimento, é necessário definir, primeiro, as medidas de similaridade ou de dissimilaridade que serão utilizadas. Se as medidas forem de dissimilaridade, quanto menores os seus valores, mais similares são os elementos que estão sendo comparados. A segunda etapa se refere à determinação do método de agrupamento.

Existem diversos tipos de medidas de dissimilaridade, e a escolha por uma delas depende de alguns fatores, dentre eles, o tipo de variável em estudo. No caso das variáveis quantitativas, podem ser utilizadas, por exemplo, distância de Minkowsky, distância generalizada ou ponderada, distância euclidiana e distância Manhattan ou quarteirão (*city block*), sendo essas duas últimas casos particulares da distância de Minkowsky.

De acordo com Ferreira (2008), a distância euclidiana é apropriada para os casos em que os grupos de variáveis possuem escalas similares, pois, caso contrário, variáveis com maior variabilidade irão dominar a classificação das distâncias. No contexto de um ensaio de proficiência, as medidas das variáveis (amostras) sempre têm a mesma escala.

Quanto aos métodos de agrupamento, existem também diversos tipos e o pesquisador deve decidir qual o mais adequado ao seu propósito, uma vez que diferentes técnicas podem levar a diferentes soluções (Souza et al., 1997). Segundo Bussab et al. (1990), todos os métodos baseiam-se em duas ideias básicas: coesão interna dos elementos e isolamento externo entre os grupos.

Os métodos de agrupamentos são divididos em hierárquicos e não hierárquicos. Segundo Rencher (2002), nos métodos não hierárquicos, o número de grupos deve ser definido previamente e os elementos são alocados de forma otimizada; já os hierárquicos são aqueles em que os elementos são classificados em grupos em etapas diferentes, ou seja, de modo hierárquico, produzindo, como resultado final, um gráfico denominado dendrograma (Figura 37). Os

dendrogramas apresentam os elementos e os respectivos pontos de junção ou de divisão dos grupos formados em estágio (Ferreira, 2008).

Um método hierárquico é dito aglomerativo quando o processo de agrupamento inicia-se com g grupos, cada um contendo um dos elementos e termina com um único grupo formado por todos os elementos.

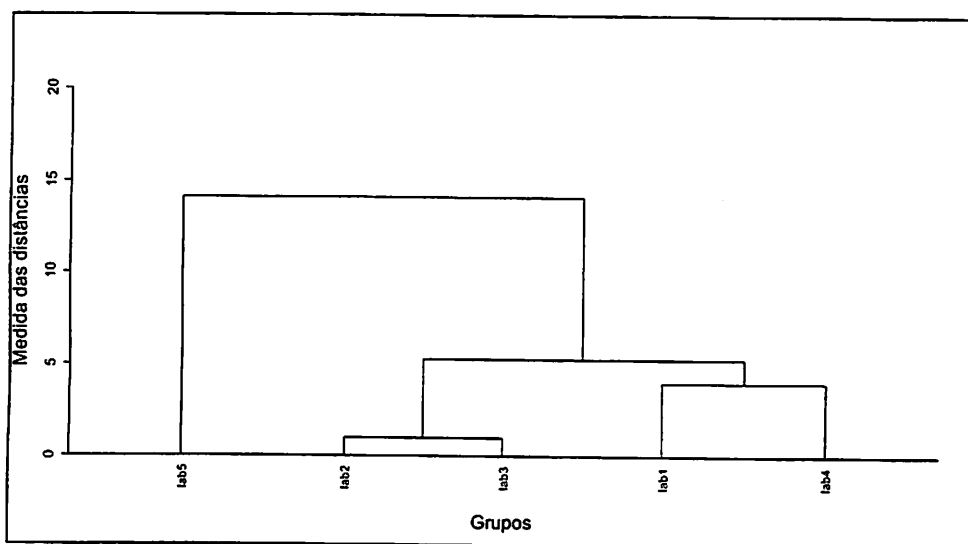


FIGURA 37 Exemplo de um dendrograma para o método de agrupamento hierárquico aglomerativo com cinco laboratórios, a partir dos dados de um teste de germinação

As técnicas aglomerativas mais utilizadas são: métodos do vizinho mais próximo, do vizinho mais distante, da ligação média, do centroide, da mediana e de Ward. Não existe uma recomendação específica sobre qual método é o mais adequado para um conjunto de dados, mas, de acordo com Totti et al. (2001), o método da ligação média é bastante utilizado.

Uma das formas de escolher o método de agrupamento pode ser por meio de comparação dos valores da correlação cofenética, isto é, o método selecionado será aquele que apresentar a maior correlação cofenética. Essa correlação é uma medida da qualidade do ajuste, pois mede o grau de preservação das distâncias emparelhadas pelo dendrograma resultante do agrupamento em relação às distâncias originais (Ferreira, 2008).

A determinação do número final de grupos também é uma etapa importante no procedimento de agrupamento, ou seja, apesar de o método aglomerativo terminar sempre com a formação de um único grupo, essa pode não ser a junção ideal ou natural. Mingoti (2005) afirma que, em termos de variabilidade, no estágio inicial tem-se a partição com menor dispersão possível, já que todos os grupos têm um único elemento. Logo, a variância de cada um deles é igual a zero e, no estágio final, tem-se a maior dispersão interna possível.

Desse fato decorre a necessidade de buscar uma partição natural dos elementos, estabelecendo um ponto de corte. Um critério que vem sendo utilizado é o proposto por Mojena (1977), que leva em consideração a média e o desvio padrão das distâncias para a determinação de uma regra de decisão (Figura 38).

Em síntese, a utilização da técnica de análise de agrupamento para análise dos resultados no contexto de ensaio de proficiência parece ser apropriada, uma vez que os laboratórios que apresentem desempenhos similares devem se agrupar naturalmente e, por conseguinte, devem formar grupo(s) isolado(s) os de piores desempenhos. Quanto ao tipo de agrupamento, pelo exposto, o hierárquico aglomerativo pode ser indicado, porque, de fato, os laboratórios não estariam reunidos em grupos pré-estabelecidos e tê-los inicialmente em grupos separados, verificando passo a passo as junções, pode ser mais adequado.

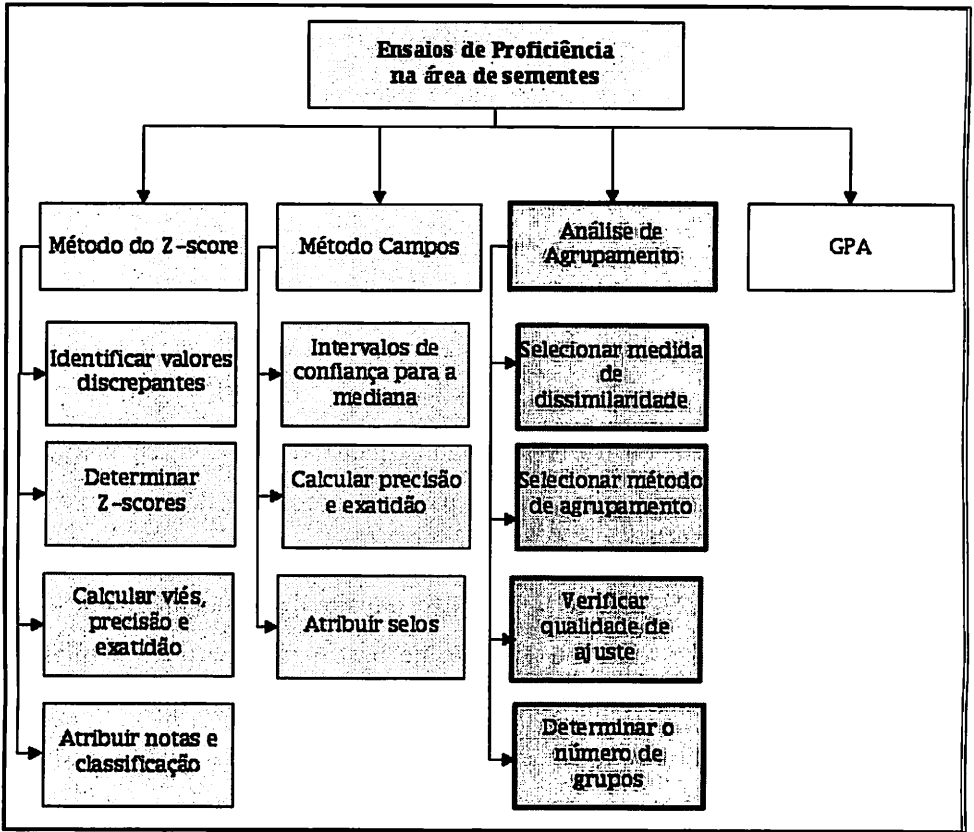


FIGURA 38 Esquema de um ensaio de proficiência na área de sementes com ênfase na análise de agrupamento

4.6 Análise generalizada de Procrustes

Desenvolvida em meados de 1970, a análise generalizada de Procrustes (cuja sigla em inglês é GPA) surgiu para resolver um problema puramente matemático e acabou tendo enorme impulso pela sua aplicação prática, particularmente no campo da análise sensorial (Ferreira & Oliveira, 2007).

No contexto da GPA (Gower, 1975) há sempre três elementos presentes: os indivíduos, os objetos e as variáveis. Por exemplo, para a análise sensorial, geralmente, os indivíduos são os provadores, os objetos são os alimentos e as

variáveis são os atributos desses alimentos (sabor, aroma, etc.). Fazendo um paralelo para um ensaio de proficiência na área de sementes, os indivíduos podem ser considerados como sendo as repetições de um determinado teste, os objetos, os laboratórios e as variáveis, as amostras. Portanto, assim como a análise sensorial busca, por meio da sensimetria, caracterizar um grupo de alimentos ao longo dos provadores, a GPA, neste contexto, busca caracterizar os laboratórios ao longo das repetições.

A GPA, em linhas gerais, procura ajustar os dados vindos de vários indivíduos para obter uma configuração de consenso, minimizando as diferenças geométricas. Para tal, todos os m conjuntos de dados são simultaneamente transladados, rotacionados e multiplicados por um fator de escala, enquanto um critério de qualidade de ajuste é otimizado (Ferreira et al., 2008).

Um dos resultados da GPA é um plano-resposta que é formado por dois eixos principais representando combinações lineares das variáveis em que os objetos foram originalmente avaliados. Esse plano mostra cada objeto como um ponto (Ferreira & Oliveira, 2007).

Ferreira et al. (2007) verificaram, em seus estudos, a importância de determinar uma distância mínima significativa entre objetos no plano-resposta, sendo, então, proposta por esses autores a estimação de uma região de confiança ao redor de cada ponto (objeto) via técnica de reamostragem Bootstrap. Nos casos em que ocorria a interseção entre elipses, um teste de hipóteses Bootstrap era então estabelecido para decidir a significância sobre a distância euclidiana entre tais pontos (Ferreira & Oliveira, 2007).

Logo, parece promissora a proposição do uso dessa técnica para análise dos resultados de um ensaio de proficiência, uma vez que é razoável supor que laboratórios (objetos) que tenham desempenhos similares serão identificados pelas interseções entre as suas elipses e, por conseguinte, os mais discordantes

ficarão geometricamente distantes dos demais, com um determinado grau de confiança pré-estabelecido (Figura 39).

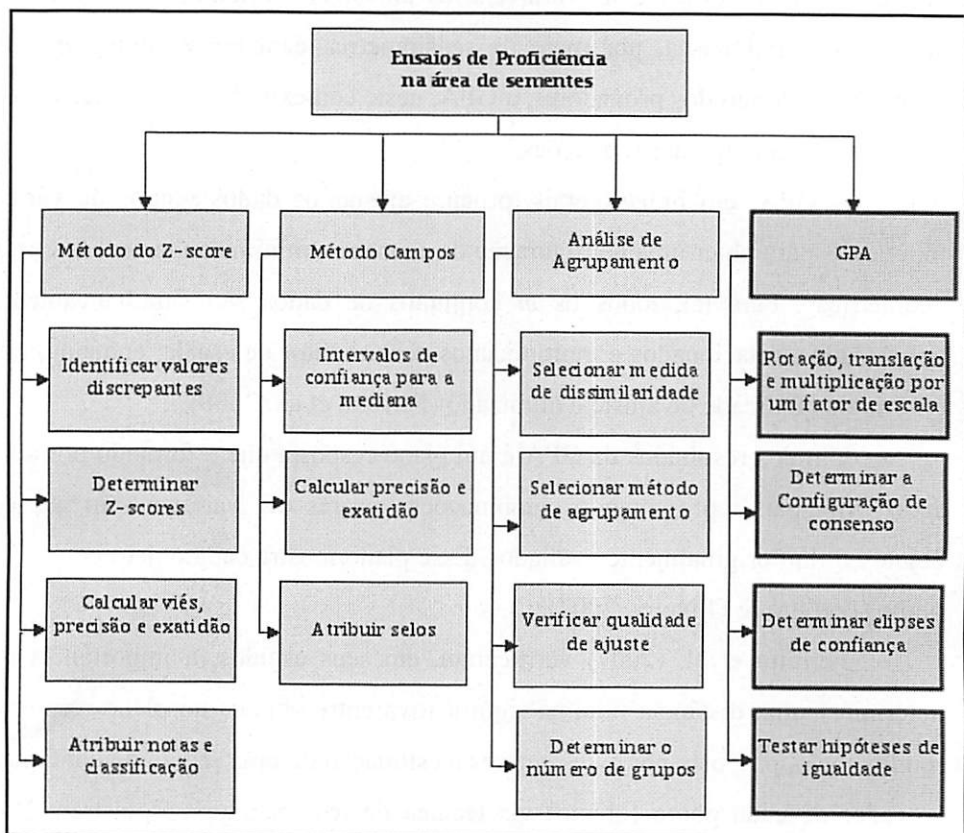


FIGURA 39 Esquema de um ensaio de proficiência na área de sementes com ênfase na GPA

5 MATERIAL E MÉTODOS

Nesta seção são apresentados os procedimentos utilizados para a geração do banco de dados, bem como a descrição metodológica das técnicas estatísticas adotadas nesse estudo.

5.1 Dos dados

Como não houve tempo hábil para a coleta de dados reais, foi feita a opção pela utilização de um exemplo ilustrativo com dados simulados. Deve-se ressaltar que o procedimento de simulação utilizado aqui não se presta a comparação entre métodos, quando então simulação implica em reamostragem. Nesta tese, simulação é um procedimento que gera uma massa de dados apenas para efeito ilustrativo. Ainda deve se salientar que esta geração de dados ilustrativos utilizou valores de parâmetros condizentes com a realidade, como será explicitado a seguir. Assim, para gerar o banco de dados, foram realizadas três simulações, cada uma representando os resultados de um teste de germinação com doze laboratórios participantes.

Esses doze laboratórios formaram uma composição de dois grupos discrepantes em relação a algum(ns) parâmetro(s) estatístico(s), como a média e a matriz de variâncias e covariâncias. Para a definição dos parâmetros da simulação, foram utilizados os resultados de germinação de alguns lotes e laboratórios selecionados do capítulo 2 (Figura 40).

O grupo 1 foi composto por dez laboratórios e o grupo 2, por apenas dois laboratórios. A determinação do número de laboratórios por grupo foi estabelecida por uma decisão empírica, já que, num ensaio de proficiência, espera-se que os resultados de laboratórios credenciados não sejam muito discordantes. Assim, foram inseridos apenas dois laboratórios com alguma

medida de discrepância para avaliar se as técnicas seriam adequadas para detectar tais diferenças.

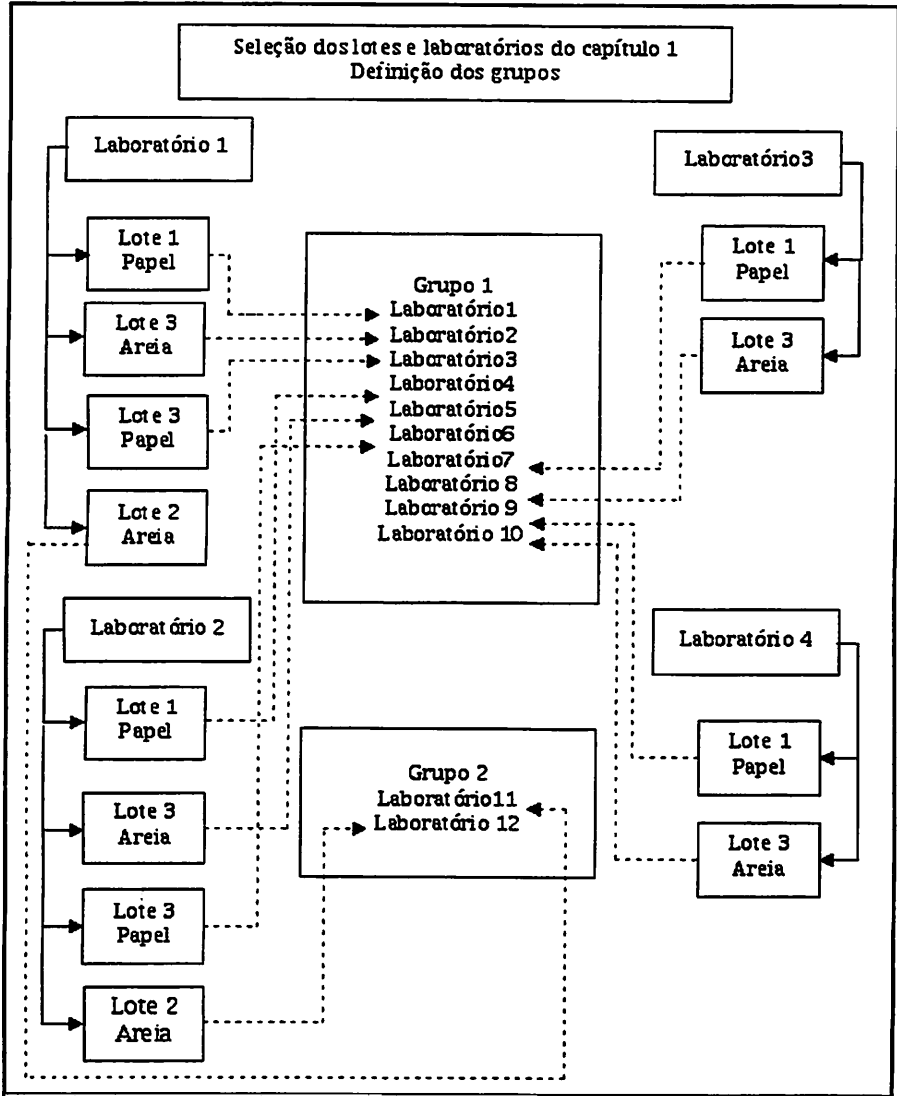


FIGURA 40 Resumo esquemático da seleção do laboratórios e lotes do capítulo 2, para a determinação dos parâmetros das simulações

O critério para a escolha dos lotes foi baseado no princípio, já comentado anteriormente, de que o grupo 1 ficasse discrepante do grupo 2. Logo, para o primeiro grupo, optou-se pelos lotes que foram caracterizados, no capítulo 2, como de qualidade superior (lotes 1 e 3) e, para o outro grupo, foi selecionado o lote considerado inferior (lote 2). Quanto aos laboratórios, foram selecionados os quatro primeiros que finalizaram as suas análises.

Escolhidos os lotes e os laboratórios do capítulo 2, foram calculadas as médias e as matrizes de variâncias e covariâncias para a definição dos parâmetros das simulações, considerando as três repetições, denominadas, neste contexto, de amostras (Tabelas 27 e 28).

TABELA 27 Proporções médias de germinação das três amostras, para os grupos 1 e 2, de acordo com os dados do capítulo 2

Grupo	Amostra		
	1	2	3
1	0,94375	0,93712	0,94137
2	0,63875	0,63750	0,62250

¹ Os dados não foram simulados na forma percentual.

TABELA 28 Variâncias e covariâncias das três amostras, para os grupos 1 e 2, de acordo com os dados do Capítulo 2

Grupos	Amostra	Amostra		
		1	2	3
1	1	0,00136	0,00064	0,00077
	2		0,00162	0,00108
	3			0,00136
2	1	0,00792	-0,00001	0,00149
	2		0,00890	-0,00010
	3			0,00767

*Para cada grupo, os valores da diagonal correspondem às variâncias e os demais valores as covariâncias, por exemplo, $Cov(a_1, a_2) = -0,00001$.

Com os parâmetros iniciais definidos, foram geradas as três simulações, sob as seguintes condições:

- simulação 1: médias iguais entre os dois grupos e matriz de variâncias e covariâncias diferentes;
- simulação 2: médias diferentes e matrizes de variâncias e covariâncias iguais entre os grupos;
- simulação 3: médias e matrizes de variâncias e covariâncias diferentes.

As simulações foram realizadas considerando, para o i -ésimo grupo ($i=1,2$), a variável aleatória porcentagem de germinação como normais multivariadas $Y_i \sim N(\mu_i, \Sigma_i)$, em que Y_i é a matriz de proporção de germinação com dimensões $v \times q$, sendo $v = n_r p$ (n_r é o número de repetições e p é o número de laboratórios) e q o número de amostras; μ_i é o vetor de médias das q amostras e Σ_i é a matriz de variâncias e covariâncias das q amostras.

Apesar da recomendação da ISTA (2007) para a realização de seis rodadas nos ensaios de proficiência, neste estudo foi considerada apenas uma rodada, pois não teria sentido prático utilizar, nas simulações, os mesmos parâmetros estatísticos para a geração dos resultados das demais rodadas.

5.2 Das técnicas estatísticas

Foram estabelecidos quatro procedimentos estatísticos para avaliar o desempenho dos laboratórios nesse ensaio de proficiência simulado: o método do Z-score, o método Campos, a análise de agrupamento e a análise generalizada de Procrustes. A porcentagem de germinação (Y) foi a variável aleatória em estudo.

Os procedimentos, com exceção da análise generalizada de Procrustes, foram descritos em forma de etapas apenas para fins didáticos.

5.2.1 Método do Z-score

Para calcular os índices de viés, exatidão, precisão de cada laboratório utilizando o método do *Z-score*, foi seguida a metodologia adotada nos ensaios de proficiência da ISTA (2002):

1ª etapa: identificação de valores discrepantes (*outliers*)

Nesta etapa, o laboratório foi considerado como *outlier*, se a estimativa da sua média, numa determinada amostra (*a*-ésima amostra), não estivesse contida num intervalo de confiança para a mediana dado por:

$$IC_{a(\gamma)} = \left[\tilde{Y}_a - S_a^* t_{a(\alpha/2; \eta)}; \tilde{Y}_a + S_a^* t_{a(\alpha/2; \eta)} \right], \quad (23)$$

em que \tilde{Y}_a é o estimador da mediana para a variável aleatória porcentagem de germinação (*Y*); S_a^* é o estimador do desvio padrão dos valores acima da mediana; $t_{a(\alpha/2; \eta)}$ é o quantil 100(1 - $\alpha/2$)% da distribuição t de Student com graus de liberdade $\eta = (n_M - 1)$, sendo n_M o número de laboratórios que têm valores acima da mediana, γ é o grau de confiança e $\alpha = (1 - \gamma)$.

Um exemplo teórico da organização dos dados e dos cálculos auxiliares para esta etapa do método do Z-score pode ser observado na Tabela 29

TABELA 29 Exemplo teórico da organização dos dados e cálculos auxiliares para a identificação de valores discrepantes pelo método do Z-score nos resultados de um teste de germinação em q amostras com p laboratórios

	Amostra 1		...	Amostra q
	Lab	Média		Média
	1	\bar{Y}_{11}		\bar{Y}_{q1}

	p	\bar{Y}_{1p}		\bar{Y}_{qp}
Mediana - \bar{Y}_a		\bar{Y}_1		\bar{Y}_q
Desvio padrão (acima da mediana) - S_a^*		S_1^*		S_q^*

* Um exemplo numérico pode ser observado no ANEXO B (p. 220)

2ª etapa: determinação do Z-score

Considerando apenas os laboratórios aceitáveis, ou seja, desconsiderando os *outliers*, em cada amostra foram calculados os valores de \bar{Y}_a (estimador da porcentagem média de germinação da a -ésima amostra) e S_a (estimador do desvio padrão da a -ésima amostra).

Em seguida, foi calculado o Z-score para cada laboratório numa determinada amostra (Z_{al}), inclusive para os valores discrepantes, de acordo com a expressão (24).

$$Z_{al} = \frac{(\bar{Y}_{al} - \bar{Y}_a)}{S_a}, \quad (24)$$

em que \bar{Y}_{al} é o estimador da porcentagem média de germinação do l -ésimo laboratório na a -ésima amostra.

3ª etapa: classificação numa rodada

A classificação dos laboratórios com as designações A, B, C ou BMP foi feita de acordo com as regras da ISTA (Tabela 26, seção 4.3), somando-se o *Z score* absoluto de todas as amostras:

$$|Z_l| = \sum_{a=1}^q Z_{al}, \quad (25)$$

em que q é o número de amostras.

Com o valor de Z_l , foi atribuída à nota n_{rl} : 5, 4, 3 ou 0, para cada laboratório, também de acordo com as regras da ISTA para apenas uma rodada.

4ª etapa: cálculo do viés

Para o cálculo do viés de cada laboratório foi determinada a média do *Z-score* utilizando a expressão (2.4).

$$V_l = \frac{\sum_{a=1}^q Z_{al}}{q} \quad (26)$$

Para valores de V superiores a 1,5 recomenda-se que o laboratório revise todos os seus procedimentos (ISTA, 2002).

5ª etapa: cálculo da precisão

A medida da precisão para o l -ésimo laboratório foi determinada em função do viés e do *Z-score*, de acordo com a expressão (27).

$$P_l^{(l)} = \frac{\sqrt{\sum_{a=1}^q (Z_{al} - V_l)^2}}{q}, \quad (27)$$

6ª etapa: cálculo da exatidão

A medida da exatidão do l -ésimo laboratório foi obtida, em função do viés e da precisão, pela expressão (28).

$$E_l^{(l)} = \sqrt{(V_l)^2 + (P_l^{(l)})^2} \quad (28)$$

Para $0,000 < E_l^{(l)} < 1,499$, o laboratório teve exatidão considerada como aceitável; para $1,500 < E_l^{(l)} < 1,999$, crítica e para $E_l^{(l)} \geq 2,000$, inaceitável (ISTA, 2002).

Um exemplo teórico da organização dos dados e dos cálculos auxiliares, bem como a sumarização das fases para a aplicação do método do Z-score, pode ser observado na Tabela 30 e na Figura 41, respectivamente.

TABELA 30 Exemplo teórico da organização dos dados e cálculos auxiliares para determinação do Z-score ($|Z_l|$), viés (V_l), precisão ($P_l^{(l)}$) e exatidão ($E_l^{(l)}$), pelo método do Z-score, nos resultados de um teste de germinação em q amostras com p laboratórios

	Amostra 1 ... Amostra q		$\sum_{a=1}^q Z_{al}$	$ Z_l $	V_l	$P_l^{(l)}$	$E_l^{(l)}$
Lab	Z-score	Z-score					
1	Z_{1l}	Z_{ql}	Z_1	$ Z_1 $	V_1	$P_1^{(l)}$	$E_1^{(l)}$
...
p	Z_{1p}	Z_{qp}	Z_p	$ Z_p $	V_p	$P_p^{(l)}$	$E_p^{(l)}$
Média - \bar{Y}_a	\bar{Y}_1	\bar{Y}_q					
Desvio padrão - S_a	S_1	S_q					

* Um exemplo numérico pode ser observado no Anexo B (p. 221-222)

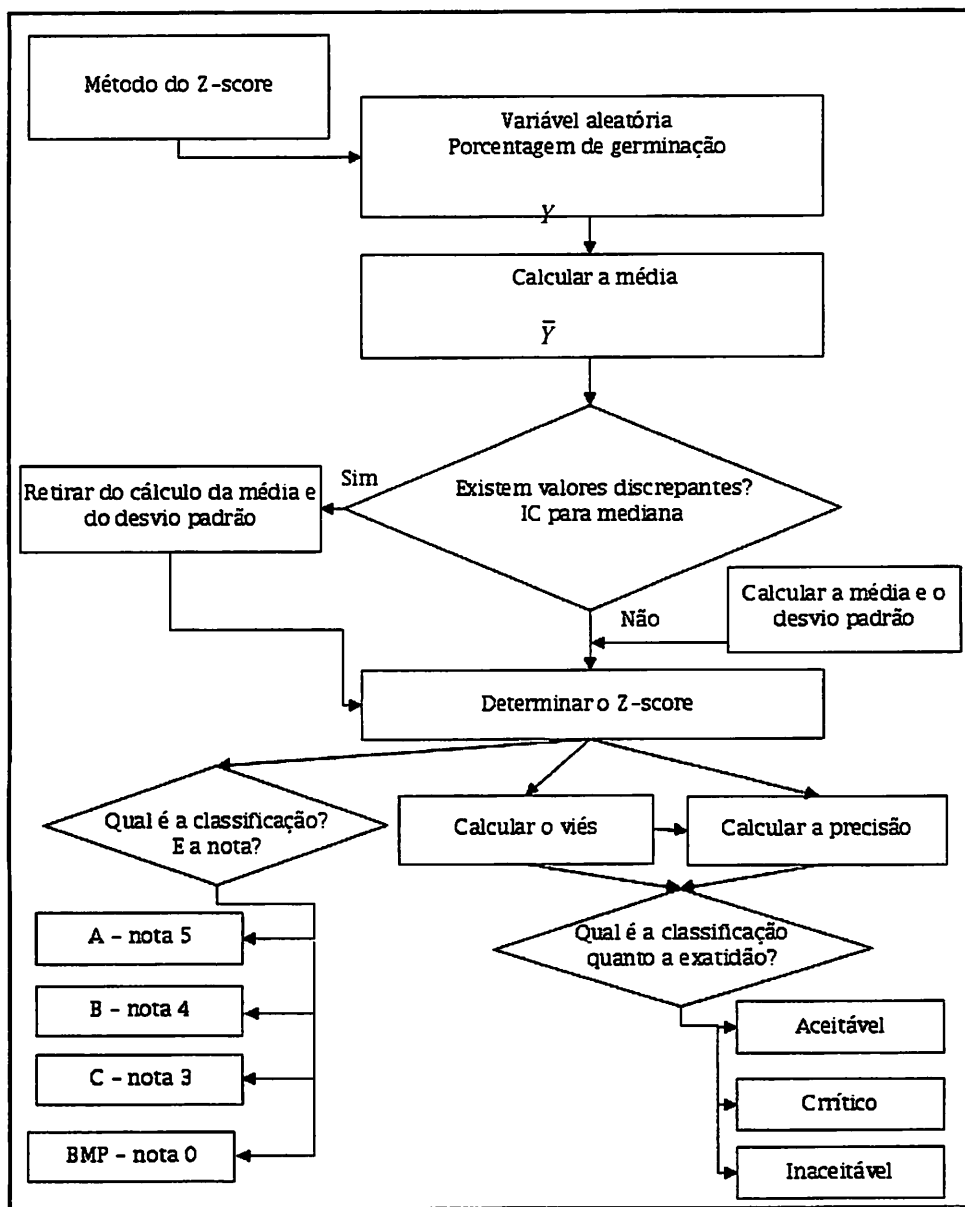


FIGURA 41 Resumo esquemático do método do Z-score no contexto de um ensaio de proficiência

5.2.2 Método Campos (1998)

Para calcular a precisão e a exatidão de cada laboratório, utilizando o método Campos, foi seguida a metodologia adotada por Campos (1998), com algumas adaptações para o contexto deste estudo.

1ª etapa: identificar valores discrepantes (*outliers*)

Para identificar e eliminar os *outliers*, foram utilizadas estatísticas robustas, especificamente os quartis e o intervalo interquartil para k -ésima repetição e a -ésima amostra

$$\left(\hat{Y}_{(ak)0,75} - \frac{B}{2} \hat{\delta}_{(ak)}; \hat{Y}_{(ak)0,25} + \frac{B}{2} \hat{\delta}_{(ak)} \right), \quad (29)$$

em que $\hat{Y}_{(ak)0,25}$ é o estimador do quantil para porcentagem de germinação com frequência de 25% acima (3º quartil), $\hat{Y}_{(ak)0,75}$ é o estimador do quantil com frequência de 75% acima (1º quartil), $\hat{\delta}_{(ak)} = \hat{Y}_{(ak)0,25} - \hat{Y}_{(ak)0,75}$ é o estimador do intervalo interquartil; B é definido como sendo $d_1 = B - 1$, para $B > 0$.

Fazendo d_1 variar de 0 a 9, e conseqüentemente B variando de 1 a 10, foi considerada como *outlier* toda a determinação que tivesse discrepância¹⁶ $d_1 > 2$,

ou seja, que estivesse fora do intervalo $\left(\hat{Y}_{(ak)0,75} - \frac{3}{2} \hat{\delta}_{(ak)}; \hat{Y}_{(ak)0,25} + \frac{3}{2} \hat{\delta}_{(ak)} \right)$,

devendo ser retirado para o procedimento da 2ª etapa.

¹⁶ A medida de discrepância para uma determinação Y indica que Y não é aceito como sendo um valor próximo do verdadeiro valor que, neste caso, é a mediana, isto é, Y não é um valor aceitável para a característica em questão (Campos, 1998).

Neste caso, entendeu-se como determinação de Y como sendo Y_{alk} , que é porcentagem de germinação da k -ésima repetição do l -ésimo laboratório na a -ésima amostra.

Um exemplo teórico da organização dos dados e dos cálculos auxiliares para esta etapa do método de Campos pode ser observado na Tabela 31.

TABELA 31 Exemplo teórico da organização dos dados e cálculos auxiliares para a determinação dos limites inferiores e superiores e valores de d_l , pelo método Campos, nos resultados de um teste de germinação em q amostras com p laboratórios e n_r repetições

Amostra 1						...	Amostra q					
B	Repetição				d_l		B	Repetição				d_l
	1	2	...	n_r				1	2	...	n_r	
1	Li				0	...	1	Li				0
	Ls							Ls				
...	Li					Li				1
	Ls							Ls				
10	Li				9	...	10	Li				9
	Ls							Ls				
	$\hat{Y}_{(1k)0.75}$...		$\hat{Y}_{(\rho k)0.75}$				
	$\hat{Y}_{(1k)0.25}$...		$\hat{Y}_{(\rho k)0.25}$				
	$\hat{\delta}_{(1r)}$...		$\hat{\delta}_{(\rho k)}$				

* Um exemplo numérico pode ser observado no Anexo B (p. 223-224)

2ª etapa: cálculo da medida de discrepância d_2

a) Foram estabelecidos intervalos de confiança para a mediana da k -ésima repetição e a a -ésima amostra:

$$(L_{i1}, L_{s1}) \cup (L_{i2}, L_{s2}) \\ \left(\bar{Y}_{ak} - (B+1)S_{md(ak)}; \bar{Y}_{ak} - BS_{md(ak)} \right) \cup \left(\bar{Y}_{ak} + BS_{md(ak)}; \bar{Y}_{ak} + (B+1)S_{md(ak)} \right), \quad (30)$$

em que \bar{Y}_{ak} é o estimador da mediana; $S_{md(ak)} = \sqrt{\frac{\pi}{2} \cdot \frac{S_{ak}^2}{n_{ak}'}}$ é o estimador do desvio padrão da mediana, sendo S_{ak}^2 o estimador da variância, n_{ak}' o novo tamanho da amostra após a retirada dos *outliers* na 1ª fase ($n_{ak}' \leq n_{ak}$), com n_{ak} sendo o número de observações iniciais.

b) Foi determinado o valor de outra medida de discrepância para a determinação de Y , denominada $d_2 = \begin{cases} 0, & \text{se } B = 0 \\ B - 1, & \text{se } B > 0 \end{cases}$. Assim, fazendo d_2 variar de 0 a 9, cada determinação de Y recebeu uma classificação, de acordo com o intervalo em que estivesse contido.

Um exemplo teórico da organização dos dados e dos cálculos auxiliares para esta etapa do método de Campos pode ser observado na Tabela 32.

TABELA 32 Exemplo teórico da organização dos dados e de cálculos auxiliares para a determinação dos limites inferiores e superiores e valores de d_2 , pelo método Campos, nos resultados de um teste de germinação em q amostras com p laboratórios e n_r repetições

Amostra 1					...	Amostra q					
B	Repetição				d_2	B	Repetição				d_2
	1	2	...	n_r			1	2	...	n_r	
	L_{i1}						L_{i1}				
0	$L_{s1} = L_{i2}$				0	0	$L_{s1} = L_{i2}$				0
	L_{s2}						L_{s2}				
...				
	L_{i1}						L_{i1}				
9	$L_{s1} = L_{i2}$				10	9	$L_{s1} = L_{i2}$				10
	L_{s2}						L_{s2}				
	\bar{Y}_{1k}				...		\bar{Y}_{qk}				
	S_{1k}^2				...		S_{qk}^2				
	n_{1k}				...		n_{qk}				
	$S_{md(1k)}$...		$S_{md(qk)}$				

* Um exemplo numérico pode ser observado no Anexo B (p. 225)

3ª etapa: cálculo da exatidão

Após o cálculo de d_1 e d_2 , foi proposto o cálculo da exatidão do l -ésimo laboratório como segue:

$$E_l^{(m)} = 100\% - \frac{T_l}{(18n_l)}, \quad (31)$$

em que T_l é a soma das discrepâncias d_1 e d_2 de todas as repetições (n_r) e amostras (q) numa única rodada; n_l é o número de resultados total ($n_r q$).

4ª etapa: cálculo da precisão

Para determinar a precisão do l -ésimo laboratório, foi necessária a utilização de três medidas auxiliares: variância mínima inerente, variância observada e imprecisão, como descrito a seguir.

a) Variância mínima inerente – variância das determinações em que a soma das discrepâncias d_1 com d_2 era igual a zero

Foi calculado primeiro o efeito para cada amostra ($S_{\min(a)}^2$) e, depois, o efeito geral (S_{\min}^2).

$$S_{\min(a)}^2 = \frac{\sum_{i=1}^{n_a} Y_{ai}^2 - \frac{\left(\sum_{i=1}^{n_a} Y_{ai}\right)^2}{n_a}}{n_a - 1}, \quad (32)$$

$$S_{\min}^2 = \frac{\sum_{a=1}^q n_a S_{\min(a)}^2}{\sum_{a=1}^q n_a}, \quad (33)$$

em que n_a ($1 < n_a < n_a$) será o numero de determinações em que a soma das discrepâncias d_1 e d_2 é zero para a -ésima amostra.

b) Variância observada – variância para cada laboratório com os dados de todas as amostras, independente das discrepâncias d_1 e d_2

$$S_{obs(l)}^2 = \frac{\sum_{a=1}^q \sum_{k=1}^{n_r} Y_{alk}^2 - \frac{\left(\sum_{a=1}^q \sum_{k=1}^{n_r} Y_{alk}\right)^2}{n_l}}{n_l - 1}, \quad (34)$$

c) Cálculo do índice de imprecisão

$$I_I = h \left(\frac{S_{obs(I)}^2}{S_{min}^2} \cdot \frac{1}{F_{\alpha(\eta_1, \eta_2)}} \right), \quad (35)$$

em que $h(\cdot)$ representa o maior inteiro menor que o argumento; $F_{\alpha(\eta_1, \eta_2)}$ é o quantil $100(1-\alpha)\%$ da distribuição F com $\eta_1 = (n_I - 1)$ e $\eta_2 = \left(\sum_{a=1}^q n_a - 1 \right)$ graus de liberdade.

d) Precisão:

$$P_I^{(II)} = 100 - \left(\frac{I_I \cdot 100}{9} \right) \quad (36)$$

5ª etapa: classificação dos laboratórios

O laboratório foi considerado exato e preciso se os valores de $E_I^{(II)}$ e $P_I^{(II)}$, respectivamente, tivessem sido superiores a 67%¹⁷, recebendo a seguinte classificação por selos:

- nenhum selo – inexatos e imprecisos;
- selo bronze¹⁸ – inexatos e precisos;
- selo prata – exatos e imprecisos;
- selo ouro – exatos e precisos.

¹⁷ O valor determinado como limite para a exatidão e precisão (67%) foi resultado de discussões e de um consenso entre profissionais das áreas de estatística e de fertilidade dos solos (Campos, 1998).

¹⁸ Esta classificação com atribuição do selo bronze não existia na proposta original de Campos (1998).

Um exemplo teórico da organização dos dados e dos cálculos auxiliares para as etapas de 3 a 5 do método de Campos, bem como a sumarização dos procedimentos, pode ser observado na Tabela 33 e na Figura 42, respectivamente.

TABELA 33 Exemplo teórico da organização dos dados e cálculos auxiliares para a determinação da exatidão ($E_i^{(II)}$), da precisão ($P_i^{(II)}$) e da classificação dos laboratórios pelo método Campos nos resultados de um teste de germinação em q amostras com p laboratórios e n_r repetições

Lab.	T_i	n_i	$E_i^{(II)}$	\bar{n}_1	\bar{n}_2	$S_{\min(1)}^2$	$S_{\min(2)}^2$	S_{\min}^2	$S_{obs(1)}^2$	I_i	$P_i^{(II)}$	Clãs.
1												
...												
p												

* Um exemplo numérico pode ser observado no Anexo B (p. 226)

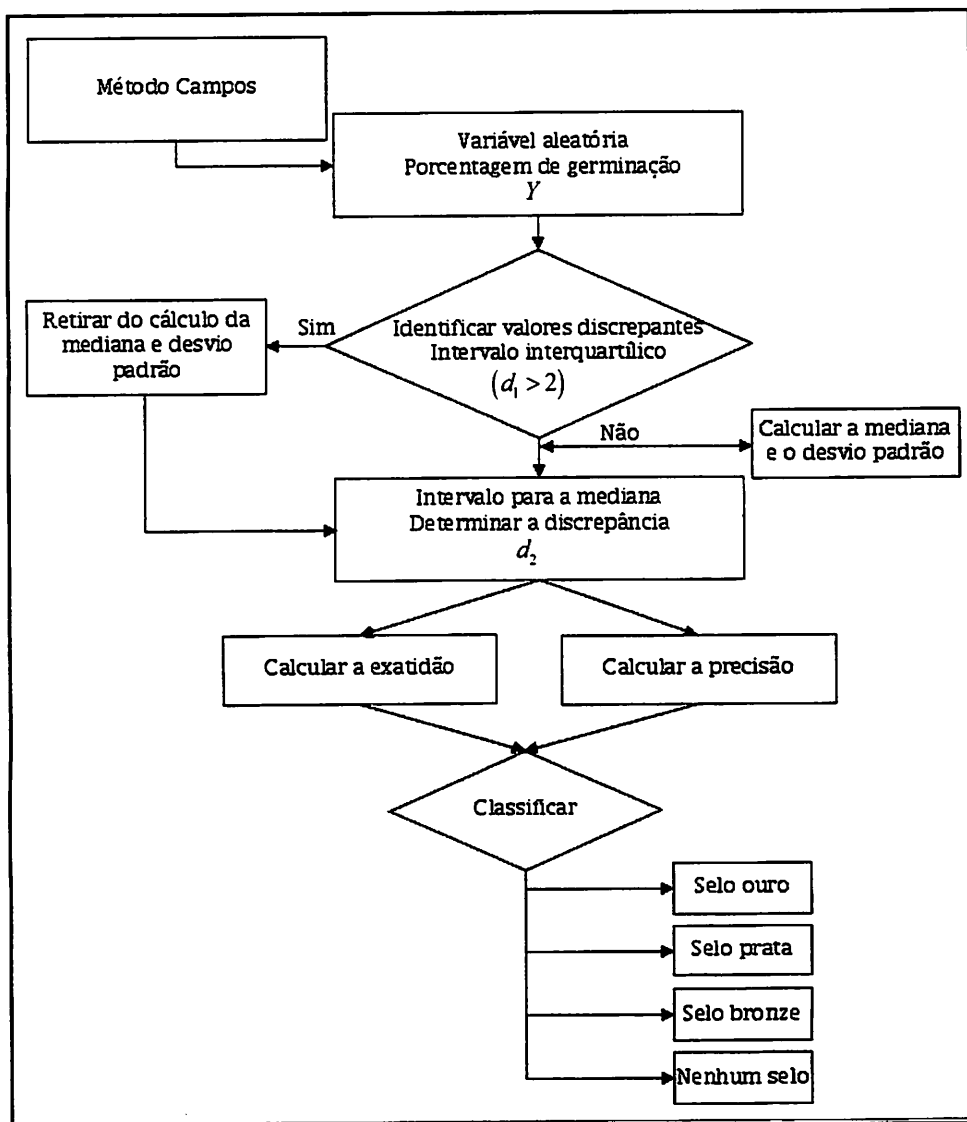


FIGURA 42 Resumo esquemático do método de Campos no contexto de um ensaio de proficiência

5.2.3 Análise de agrupamento

Para o uso desta técnica foram consideradas apenas as médias de cada laboratório por amostra. A metodologia utilizada está de acordo com Ferreira (2008) e Mingotti (2005), com algumas adaptações para o contexto deste estudo.

1ª etapa: escolha do critério de parença

Neste estudo adotou-se a distância euclidiana como medida de dissimilaridade entre dois objetos k e j , e que determina a distância física entre dois objetos, considerando o espaço euclidiano.

2ª etapa: escolha do algoritmo de agrupamento

Foi selecionado o método da ligação média¹⁹ ou média das distâncias, no qual a distância entre grupos é calculada pela média das distâncias entre todos os pares de elementos dos dois grupos que estão sendo comparados.

As fases para o desenvolvimento do algoritmo estão descritas a seguir.

1ª fase

Inicialmente, o número de grupos estabelecidos foi igual ao número de laboratórios, ou seja, 12 grupos com um laboratório cada, sendo calculada a distância euclidiana entre dois grupos quaisquer (i e j), de um laboratório cada (expressão 37).

$$d_{ij} = \sqrt{\sum_{a=1}^q (\bar{Y}_{ai.} - \bar{Y}_{aj.})^2}, \quad (37)$$

¹⁹ A escolha por este método foi baseada em estudo comparativo, no qual o valor da correlação cofenética para este método foi superior aos dos métodos do vizinho mais próximo, do vizinho mais distante, centroide, mediana e Ward. Esses resultados não serão apresentados neste trabalho.

em que q é igual ao número de amostras; \bar{Y}_{ia} é o estimador da porcentagem média de germinação do i -ésimo grupo (laboratório) na a -ésima amostra; \bar{Y}_{ja} é o estimador da porcentagem média de germinação do j -ésimo grupo na a -ésima amostra.

Em seguida, foi montada uma matriz de distâncias e observada a combinação com o menor valor de d_{ij} , estabelecendo-se, assim, os dois laboratórios que passaram a constituir um novo grupo.

2ª fase

Com o primeiro grupo formado, atualizou-se a matriz de distâncias calculando-se a distância entre o grupo recém-criado e os grupos remanescentes, utilizando-se o critério de tomar a média das distâncias entre todos os pares de elementos que podem ser formados com os elementos dos dois grupos que estão sendo comparados. Assim, se o grupo G_i tem n_i elementos e o grupo G_j tem n_j elementos, a distância entre eles pode ser definida por:

$$d(G_i, G_j) = \sum_{m \in G_i} \sum_{q \in G_j} \frac{d_{mq}}{n_i n_j}. \quad (38)$$

Este processo foi então repetido, até se formar apenas um único grupo ($g=1$) e, com os valores de distâncias obtidos em cada passo do processo iterativo, foi construído o dendrograma, indicando-se os grupos que foram formados e as distâncias correspondentes.

3ª etapa: determinação do número de grupos

Para a determinação do número de grupos g , foi utilizado o critério proposto por Mojena (1977), que aperfeiçoa a qualidade do ajuste do

agrupamento aos dados. Isto é, foi escolhido o número de grupos dado pelo primeiro estágio no dendrograma, no qual

$$\alpha_m > \bar{\alpha} + \phi S_\alpha \quad (39)$$

em que $m=1,2,\dots,p$, sendo p o número máximo de grupos (neste contexto o número de laboratórios); α_m é o valor da distância para o estágio da junção correspondente a $p-m+1$ grupos, $\bar{\alpha}$ e S_α são estimadores da média e o desvio padrão dos α 's distâncias; ϕ é uma constante.

Milligan & Cooper (1985) sugerem utilizar $\phi=1,25$, baseados em resultados de simulação.

4ª etapa: validação do agrupamento

Para validar o agrupamento foram comparadas as distâncias efetivamente observadas entre os objetos (distâncias euclidianas) e as distâncias previstas a partir do processo de agrupamento, utilizando-se a correlação de Pearson, mas que, neste contexto, é denominada de correlação cofenética. Quanto mais próximo de 1 melhor a qualidade de ajuste do agrupamento.

A sumarização das fases para a utilização da técnica de análise de agrupamento pode ser observada na Figura 43.

Um exemplo numérico de todos os passos apresentados para esta técnica pode ser observado no Anexo B.

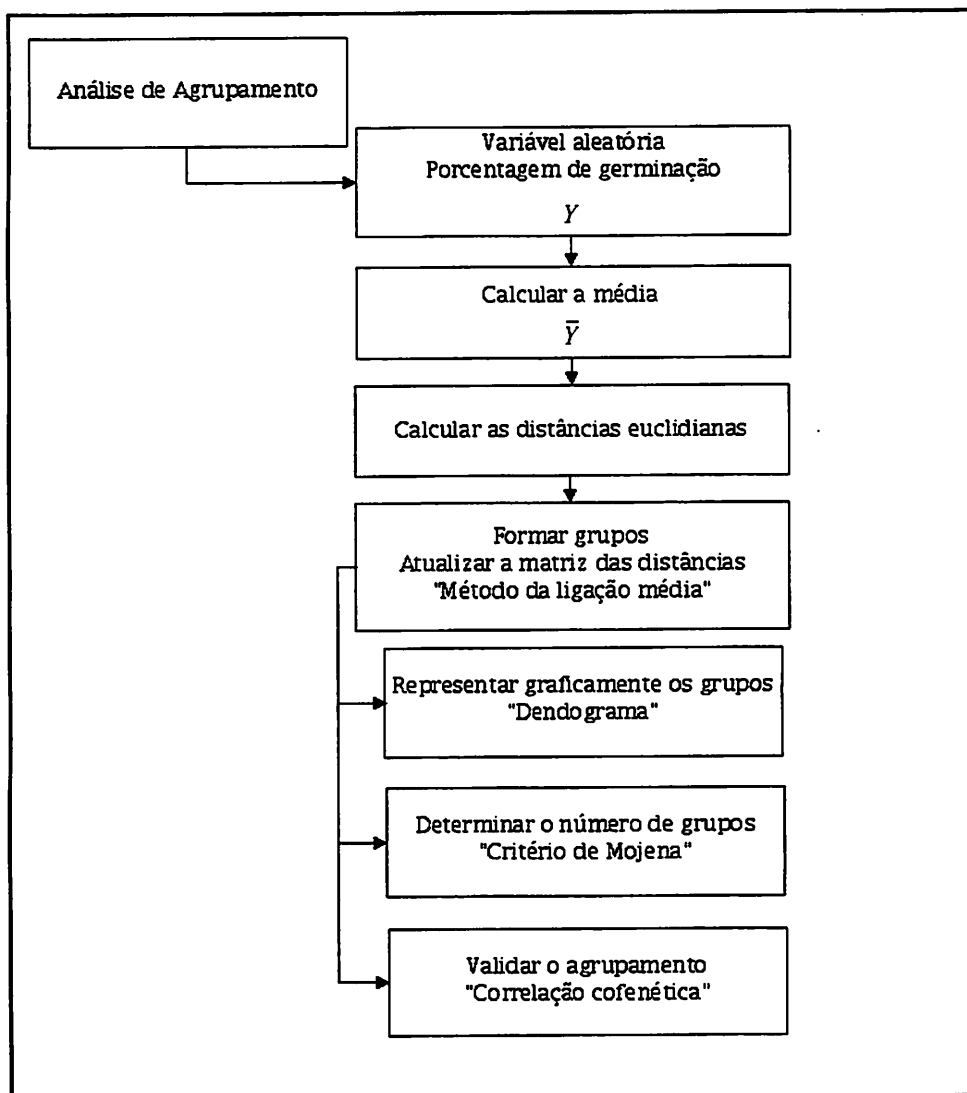


FIGURA 43 Resumo esquemático da técnica de análise de agrupamento, no contexto de um ensaio de proficiência

5.2.4 Análise generalizada de Procrustes (GPA)

A seguir são apresentadas, de forma sucinta, as fases de organização e análise de dados na GPA. A metodologia utilizada está de acordo com Ferreira et al. (2007), com algumas adaptações para o contexto deste estudo.

Organização dos dados

Os dados foram organizados em n_r matrizes de dimensões pxq , em que n_r é o número de repetições; p é o número de laboratórios e q é o número de amostras.

Cada matriz foi denotada por X_k ($k=1, \dots, r', \dots, n_r$), tendo na linha a porcentagem de germinação correspondente ao l -ésimo laboratório ($l=1, \dots, p$) e, na coluna, a a -ésima amostra ($a=1, \dots, q$).

$$X_r = \begin{bmatrix} x_{11} & \cdots & x_{1q} \\ \cdots & \cdots & \cdots \\ x_{p1} & \cdots & x_{pq} \end{bmatrix}$$

em que x_{ij} é a porcentagem de germinação do l -ésimo laboratório e da a -ésima amostra.

Análise dos dados

Como comentado na seção 4.6, geometricamente, a ideia básica da técnica de Procrustes é procurar uma configuração de consenso que se ajuste simultaneamente a todos os n_r conjuntos de dados.

Para tal, as matrizes X_k devem ser transformadas aplicando-se sequencialmente operações matemáticas de translação, rotação e escala, de forma que fiquem as mais parecidas possíveis.

Essa transformação pode ser expressa por:

$$X'_k \rightarrow s_k (X_k - I u'_k) Q_k \quad (40)$$

em que X'_k é a matriz X_k transformada; s_k é um fator de escala; u_k é um vetor coluna de translação (dimensão $1 \times p$) é que contém as médias das colunas de X_k ; I é um vetor linha de uns e Q_k é uma matriz ortogonal que promove a rotação.

A maneira de verificar a similaridade entre as matrizes X_k é adotar um critério de ajuste a ser otimizado, como a minimização da soma de quadrados dos resíduos, representada por:

$$S = \sum_{k=1}^{n_r} \|X'_k - G\| \quad (41)$$

em que $G = \frac{1}{n_r} \sum_{k=1}^{n_r} s_k (X_k - I u'_k) Q_k = \frac{1}{n_r} \sum_{k=1}^{n_r} X'_k$, representando a média das matrizes transformadas (com a mesma dimensão de X_k).

Assim, para resolver o problema da minimização de S , é necessário estimar s_k , u_k , Q_k , e determinar a matriz G , sendo utilizado, para tal, um algoritmo iterativo.

A reamostragem Bootstrap é utilizada para permitir a inferência estatística (intervalos de confiança e teste de hipóteses), como descrito por Ferreira et al. (2007). Cada unidade reamostral, neste contexto, é uma das matrizes X_k que geram as estimativas de Bootstrap de localização dos objetos (laboratórios, neste contexto).

Neste estudo, foram compostas 100 amostras de Bootstrap, resultando, conseqüentemente, em 100 matrizes G , e que, para fins de visualização gráfica, utilizaram-se apenas as duas primeiras colunas de sua projeção no espaço gerado pelos eixos principais.

Vale salientar que, além dos 100 conjuntos de pontos gerados por reamostragem Bootstrap, o conjunto de pontos proveniente da amostra original também foi utilizado para a tomada de decisões.

Determinados os 100 pontos, foram esboçadas p elipses, com 95% de confiança, cada uma representando geometricamente a configuração de um laboratório. A sumarização das fases para a utilização da técnica de GPA pode ser observada na Figura 44.

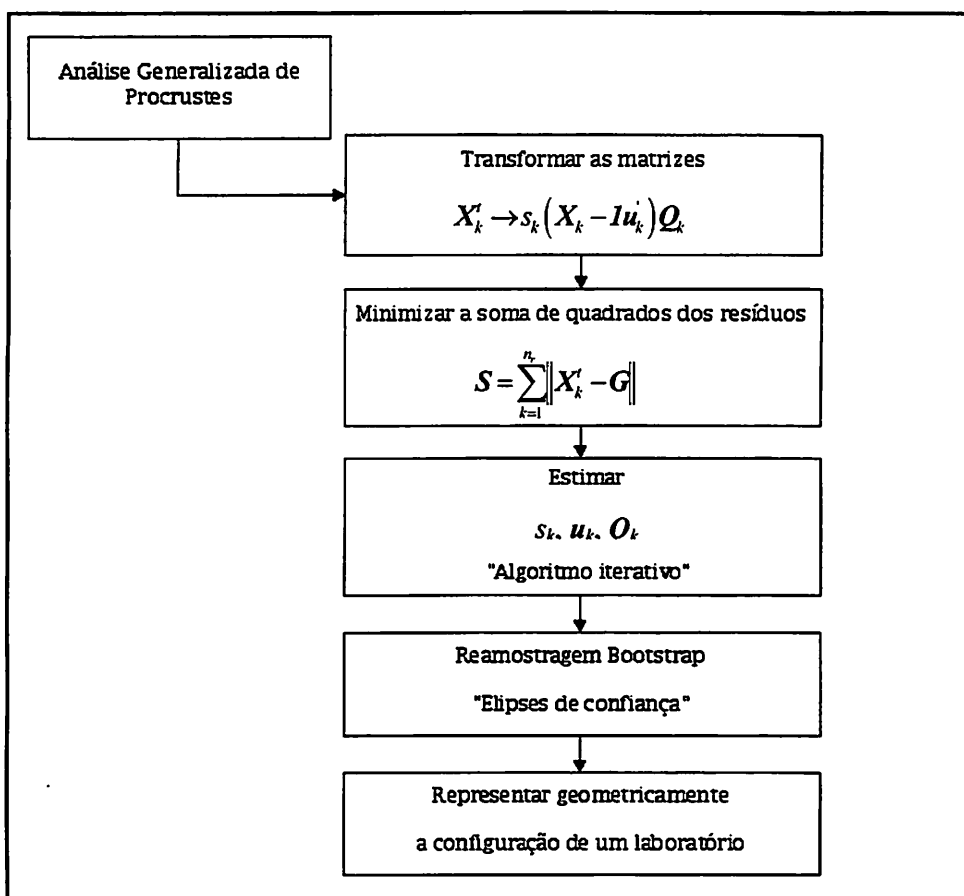


FIGURA 44 Resumo esquemático da análise generalizada de Procrustes, no contexto de um ensaio de proficiência

5.3 Rotinas no R

Para as análises estatísticas, foram desenvolvidas funções específicas no software R (R Development Core Team, 2008), com exceção da análise de agrupamento, em que foram usadas as funções já implementadas. Os nomes e os argumentos necessários para cada função podem ser observados na Tabela 34.

Vale salientar que Campos (1998) desenvolveu, numa planilha de cálculo, as rotinas para a execução do método PROFERT e verificou que o tempo gasto era de 32 minutos, em média, para realizar uma análise completa. Com as rotinas desenvolvidas no *software* R, não foram gastos nem 2 segundos para proceder à mesma análise, incluindo a geração automática de um relatório completo.

TABELA 34 Nome da função e argumentos necessários, de acordo com a técnica estatística

Técnica	Nome da função	Argumentos
Método do Z score	FunçãoZscore	- nome do banco de dados. - grau de confiança
Método Campos	FunçãoCampos	- nome do banco de dados - array de dados - nome das dimensões - lógico, diz se a transformação deve ser ortogonal (rotação)
Análise generalizada de Procrustes	bootGPA (desenvolvida por Ferreira, 2007)	- número de reamostragens bootstrap - grau de confiança das elipses

Exemplos de alguns relatórios gerados e tipo de banco de dados podem ser observados no Anexo E.

6 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados os resultados obtidos nas três simulações, bem como a aplicação, para cada caso, das quatro técnicas estatísticas propostas neste estudo.

Cada simulação correspondeu a um ensaio de proficiência, tendo como escopo o teste de germinação realizado em doze laboratórios com três amostras e oito repetições.

6.1 Simulação 1

Na simulação 1, os dados foram gerados utilizando-se as matrizes de variâncias e covariâncias diferentes e médias iguais para os dois grupos, um com 10 (grupo 1) e outro com 2 (grupo 2) laboratórios.

Dados simulados

Os resultados simulados das proporções médias de germinação e das variâncias e covariâncias para os dois grupos podem ser observados nas Tabelas 35 e 36, respectivamente.

TABELA 35 Proporções médias de germinação das três amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 1.

Simulação	Grupo	Amostra		
		1	2	3
1	1	0,94132	0,93708	0,93872
	2	0,95737	0,92561	0,94317

TABELA 36 Variâncias e covariâncias da proporção de germinação das três amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 1.

Simulação	Grupo	Amostra	Amostra		
			1	2	3
1	1	1	0,00131	0,00052	0,00061
		2		0,00154	0,00091
		3			0,00107
	2	1	0,00307	0,00177	0,00009
		2		0,00418	-0,00124
		3			0,00487

Método do Z-score

Pelos valores do Z-score dos doze laboratórios para as três amostras, observa-se que o laboratório 3 teve dois valores na faixa de desempenho considerada como insatisfatória ($|Z| \geq 3$), assim como os laboratórios 6, 7, 10, 11 e 12, com apenas um valor cada. Os melhores desempenhos foram os dos laboratórios 1, 4, 5 e 9 - $|Z| < 2$ (Figura 45).

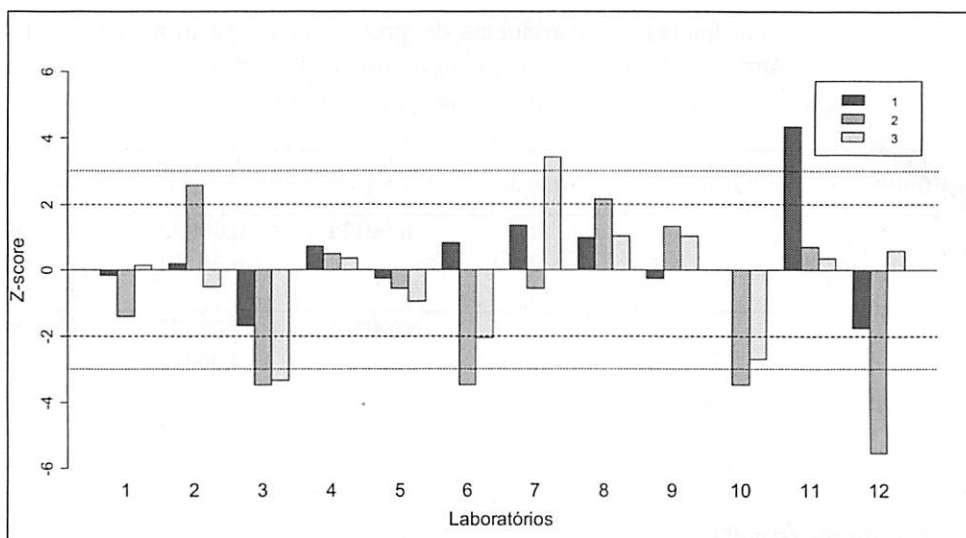


FIGURA 45 Valores de Z-score para um ensaio de proficiência em doze laboratórios, com três amostras, de acordo com os critérios estabelecidos para a simulação 1

Levando em consideração tanto as notas como a classificação da exatidão, verifica-se que não foi possível distinguir satisfatoriamente os dois grupos para a simulação 1 (Tabela 37).

TABELA 37 Resultados do método do Z-score para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 1

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classificação exatidão
1	1	A	5	0,6	1,2	1,4	Aceitável
	2	A	5	1,1	1,4	1,7	Crítico
	3	BMP	0	2,8	5,7	6,4	Não aceitável
	4	A	5	0,5	0,1	0,5	Aceitável
	5	A	5	0,6	1,2	1,3	Aceitável
	6	C	3	2,1	4,1	4,6	Não aceitável
	7	C	3	1,8	1,7	2,4	Não aceitável
	8	B	4	1,4	0,5	1,5	Aceitável
	9	A	5	0,9	0,7	1,1	Aceitável
	10	C	3	2,1	4,4	4,8	Não aceitável
2	11	C	3	1,8	1,8	2,6	Não aceitável
	12	BMP	0	2,6	5,5	6,1	Não aceitável

Neste contexto, o que se pode questionar é o risco que se ocorre de classificar um laboratório como BMP, por exemplo, o laboratório 3, quando ele deveria ter desempenho similar aos demais laboratórios do seu grupo inicial. Esta estranheza existe pelo menos em tese, já que se está trabalhando com dados simulados e espera-se, conseqüentemente, que em condições relativamente controladas.

Assim, parece que esta técnica talvez não seja a mais adequada para a situação em que as matrizes de variâncias e covariâncias sejam diferentes.

Mas, sem levar em consideração a discussão sobre a distinção dos grupos, esses resultados indicam que, no contexto de um ensaio de proficiência, os laboratórios que receberam classificação C e BMP devem ser alertados para fazer uma revisão geral em todos os seus procedimentos e equipamentos, antes de uma próxima rodada (ISTA, 2007), bem como aqueles que apresentaram índice de viés acima de 1,5 (ISTA, 2002).

Método Campos (1998)

Pelos resultados do método de Campos foi possível classificar os laboratórios do grupo 1 com um selo diferente do grupo 2, evidenciando as diferenças entre os valores de precisão e exatidão para os dois grupos (Tabela 38).

TABELA 38 Resultados do método Campos para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 1

Grupo	Lab.	Precisão (%)	Exatidão (%)	Selo
1	1	93,29	88,89	Ouro
	2	96,30	100,00	Ouro
	3	88,89	88,89	Ouro
	4	91,90	77,78	Ouro
	5	97,92	100,00	Ouro
	6	91,90	88,89	Ouro
	7	93,29	88,89	Ouro
	8	95,60	88,89	Ouro
	9	93,06	100,00	Ouro
	10	94,68	88,89	Ouro
2	11	81,25	66,67	Prata
	12	86,57	55,56	Prata

Análise de Agrupamento

Inicialmente, utilizando-se o critério proposto por Mojena (1977), só seria formado um único grupo, pois o ponto de corte foi igual a 4,63 ($\bar{\alpha} = 2,51; s_{\alpha} = 1,70$) e o valor da distância do estágio da junção correspondente a um grupo foi 6,49, de acordo com o dendrograma (Figura 46)

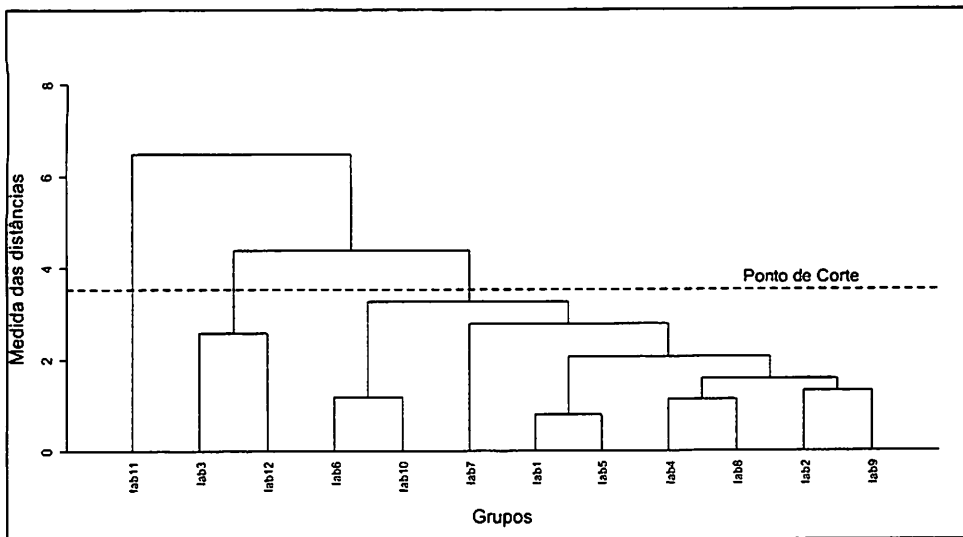


FIGURA 46 Dendrograma para o agrupamento de doze laboratórios utilizando o método da ligação média, de acordo com os critérios estabelecidos para a simulação 1

Contundo, este valor de distância foi considerado *outlier* pelo método de Hampel. Logo, neste estudo, está sendo proposta uma modificação no procedimento original do critério de Mojena, isto é, primeiro deve-se verificar uma possível presença de um valor discrepante na medida das distâncias, para depois calcular a média, o desvio padrão e, conseqüentemente, o ponto de corte.

Neste caso, o novo ponto de corte foi igual a 3,52 ($\bar{\alpha} = 2,10; s_{\alpha} = 1,13$) e a medida da distância correspondente à junção da formação de dois grupos foi de 4,39. Assim, o laboratório 11 ficou separado dos demais laboratórios, mas não foi possível, com o uso da técnica de análise de agrupamento, separar também o laboratório 12.

O agrupamento pôde ser considerado de boa qualidade, pois o valor da correlação cofenética foi igual a 86,52%.

Análise Generalizada de Procrustes (GPA)

De acordo com a inspeção visual do gráfico gerado pela utilização da técnica de GPA, verifica-se que os laboratórios do grupo 2 estão relativamente afastados do grupo 1. Este afastamento nota-se, principalmente, no laboratório 12, já que a elipse do laboratório 11 ainda intercepta numa pequena região da elipse do laboratório 7, porém, estatisticamente (valor $p = 0,2920$) o suficiente para não se poder afirmar que há discordância entre seus resultados. Contudo, parece ser ainda uma técnica apropriada para a distinção dos laboratórios na situação da simulação 1 (Figura 47).

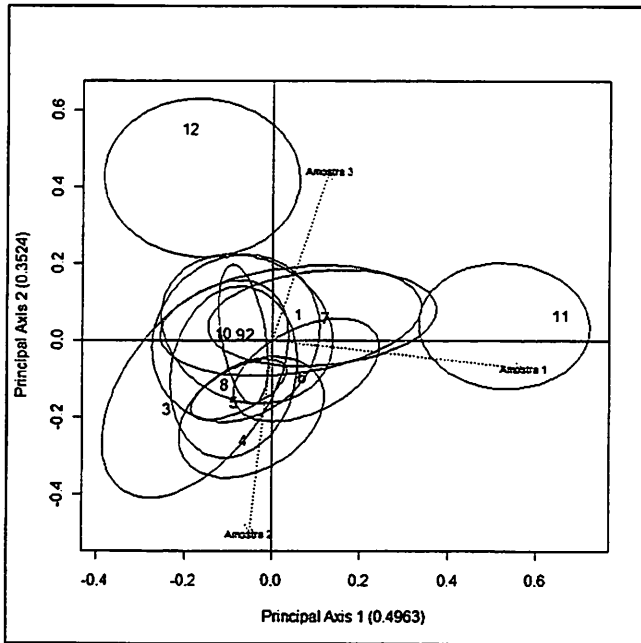


FIGURA 47 Elipses de confiança de 95%, geradas pela técnica de GPA para os doze laboratórios, de acordo com os critérios estabelecidos para a simulação 1

6.2 Simulação 2

Na simulação 2 os dados foram gerados utilizando-se médias diferentes e matrizes de variâncias e covariâncias iguais para os dois grupos.

Dados simulados

Os resultados simulados das médias e das matrizes de variâncias e covariâncias para os dois grupos podem ser observados nas Tabelas 39 e 40, respectivamente.

TABELA 39 Proporções médias de germinação das três amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 2

Simulação	Grupo	Amostra		
		1	2	3
2	1	0,94563	0,93079	0,93327
	2	0,63152	0,62662	0,61657

TABELA 40 Variâncias e covariâncias da proporção de germinação das três amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 2

Simulação	Grupo	Amostra	Amostra		
			1	2	3
2	1	1	0,00128	0,00047	0,00062
		2		0,00145	0,00113
		3			0,00143
	2	1	0,00183	0,00049	0,00110
		2		0,00138	0,00086
		3			0,00103

Técnicas estatísticas

No caso da simulação 2, apesar de o método Z-score não ter conseguido distinguir satisfatoriamente os dois grupos (Tabela 41), deve-se observar que os

scores para as três amostras dos laboratórios 11 e 12 são bastante elevados negativamente (Figura 48). Tal comportamento não foi observado os resultados desses dois laboratórios para as condições da simulação 1 (Figura 45). Assim, essas considerações parecem ser um indicativo que poderia ser proposta uma classificação intermediária entre C e BMP.

Todas as demais técnicas diferenciaram, de forma satisfatória, os dois grupos, como pode ser observado pelos resultados da Tabela 42, no caso do Método Campos, o dendrograma para Análise de Agrupamento (Figura 49) e o gráfico gerado pela GPA (Figura 50).

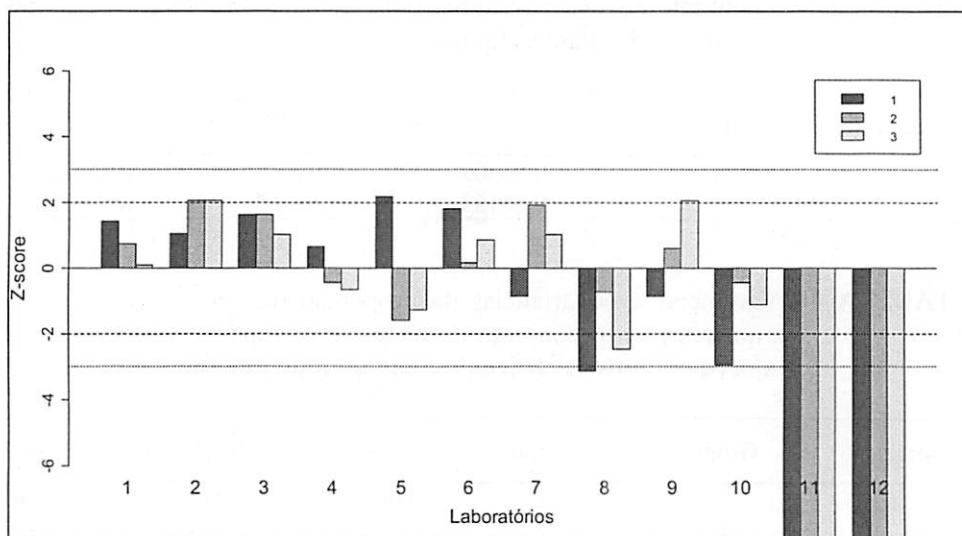


FIGURA 48 Valores de Z-score²⁰ para um ensaio de proficiência em doze laboratórios com três amostras, de acordo com os critérios estabelecidos para a simulação 2

²⁰ Para os laboratórios 11 e 12, os valores de Z-score eram muito elevados, não tendo sido adotada a escala do eixo y em função dos mesmos, para não provocar uma distorção visual dos demais valores, bem como torná-los comparáveis com as outras simulações.

TABELA 41 Resultados do método do Z-score para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 2

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classificação exatidão
1	1	A	5	0,8	0,5	0,9	Aceitável
	2	B	4	1,7	0,5	1,8	Crítico
	3	B	4	1,4	0,3	1,4	Aceitável
	4	A	5	0,6	0,9	1,1	Aceitável
	5	B	4	1,7	2,5	3,1	Não aceitável
	6	A	5	0,9	0,7	1,2	Aceitável
	7	B	4	1,3	1,3	1,8	Crítico
	8	C	3	2,1	4,3	4,8	Não aceitável
	9	B	4	1,2	1,3	1,8	Crítico
	10	B	4	1,5	3,2	3,5	Não aceitável
2	11	BMP	0	41,2	82,5	92,2	Não aceitável
	12	BMP	0	39,4	79,1	88,3	Não aceitável

TABELA 42 Resultados do método Campos para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 2

Grupo	Lab.	Precisão (%)	Exatidão (%)	Selo
1	1	93,98	88,89	Ouro
	2	93,29	88,89	Ouro
	3	94,91	88,89	Ouro
	4	96,06	88,89	Ouro
	5	96,30	100,00	Ouro
	6	97,69	100,00	Ouro
	7	90,74	88,89	Ouro
	8	92,82	88,89	Ouro
	9	93,89	88,89	Ouro
	10	93,89	88,89	Ouro
2	11	7,18	100,00	Bronze
	12	9,26	88,89	Bronze

Dados simulados

Os resultados simulados das médias e das variâncias e covariâncias para os dois grupos podem ser observados nas Tabelas 43 e 44, respectivamente.

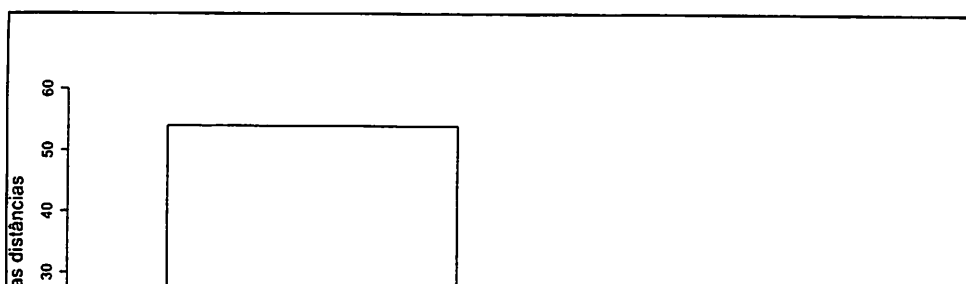


TABELA 43 Proporções médias de germinação das três amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 3

Simulação	Grupo	Amostra		
		1	2	3
2	1	0,93958	0,93590	0,93529
	2	0,65235	0,63809	0,61254

TABELA 44 Variâncias e covariâncias da proporção de germinação das TRÊS amostras para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 3

Simulação	Grupo	Amostra	Amostra		
			1	2	3
2	1	1	0,00726	0,00031	0,00481
		2		0,00533	-0,00168
		3			0,00954
	2	1	0,00144	0,00079	0,00086
		2		0,00200	0,00135
		3			0,00144

Técnicas estatísticas

No caso da simulação 3, também com exceção do método do Z-score (Figura 51 e Tabela 45), todas as demais técnicas diferenciaram de forma satisfatória os dois grupos, no caso do Método Campos com os resultados apresentados na Tabela 46, o dendrograma para Análise de Agrupamento (Figura 52) e o gráfico gerado pela GPA (Figura 53).

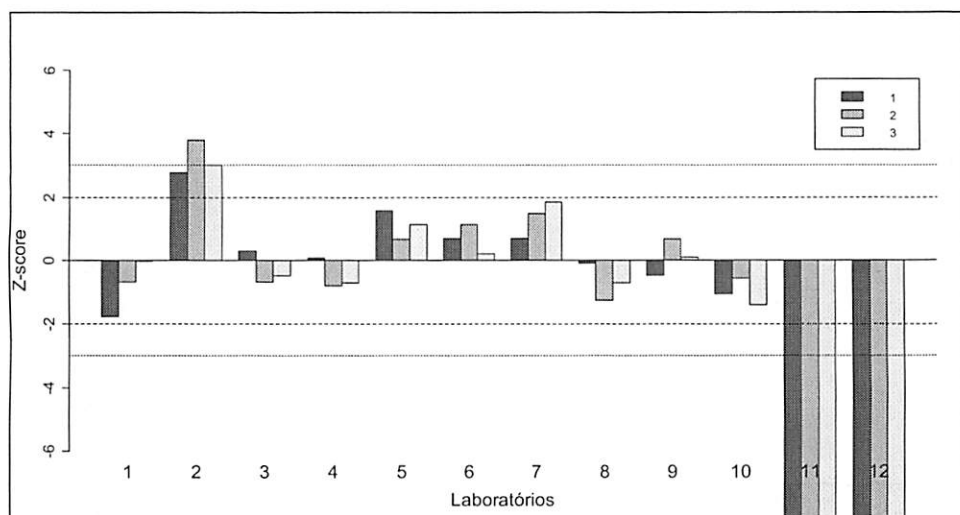


FIGURA 51 Valores de Z-score²¹ para um ensaios de proficiência em doze laboratórios, com três amostras, de acordo com os critérios estabelecidos para a simulação 3

TABELA 45 Resultados do método do Z-score para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 3

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classificação exatidão
1	1	A	5	0,8	1,8	2,0	Crítico
	2	BMP	0	3,2	0,4	3,2	Não aceitável
	3	A	5	0,5	0,9	1,0	Aceitável
	4	A	5	0,5	1,1	1,2	Aceitável
	5	A	5	1,1	0,4	1,2	Aceitável
	6	A	5	0,7	0,4	0,8	Aceitável
	7	B	4	,3	0,5	1,4	Aceitável
	8	A	5	0,7	1,5	1,6	Crítico
	9	A	5	0,4	0,6	0,7	Aceitável
	10	A	5	1,0	2,0	2,3	Não aceitável
2	11	BMP	0	23,4	47,1	52,6	Não aceitável
	12	BMP	0	26,4	52,9	59,1	Não aceitável

²¹ Para os laboratórios 11 e 12, os valores de Z-score eram muito elevados, não tendo sido adotada a escala do eixo y em função dos mesmos para não provocar uma distorção visual dos demais valores, bem como torná-los comparáveis com as outras simulações.

TABELA 46 Resultados do método Campos para o conjunto de dados dos grupos 1 e 2, de acordo com os critérios estabelecidos para a simulação 3

Grupo	Lab.	Precisão (%)	Exatidão (%)	Selo
1	1	97,45	100,00	Ouro
	2	91,20	100,00	Ouro
	3	95,60	88,89	Ouro
	4	93,52	88,89	Ouro
	5	97,45	88,89	Ouro
	6	97,45	88,89	Ouro
	7	95,14	88,89	Ouro
	8	90,74	88,89	Ouro
	9	97,45	100,00	Ouro
	10	96,76	88,89	Ouro
2	11	19,91	55,56	Nenhum selo
	12	15,97	22,22	Nenhum selo

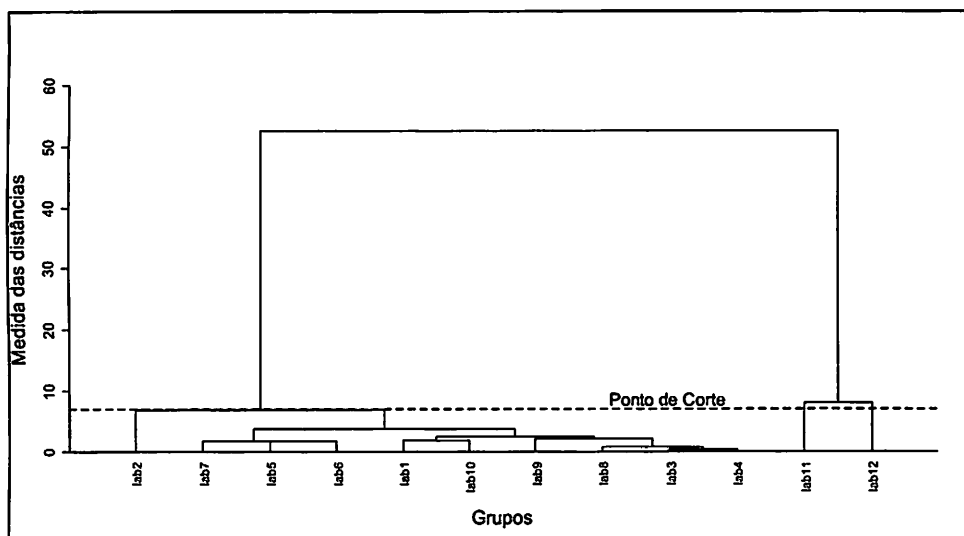


FIGURA 52 Dendrograma para o agrupamento de doze laboratórios, utilizando o método da ligação média, de acordo com os critérios estabelecidos para a simulação 3

Para a técnica de agrupamento, o ponto de corte estabelecido, após a retirada da distância 52,56 (*outlier* pelo método de Hampel), foi igual a 6,18 ($\bar{\alpha} = 2,99; s_{\alpha} = 2,54$) e a medida da distância correspondente à junção da formação de dois grupos de 8,09. Assim, os laboratórios 11 e 12 ficam num grupo diferente dos demais. O agrupamento pode ser considerado de boa qualidade, pois o valor da correlação cofenética foi igual a 99,39%.

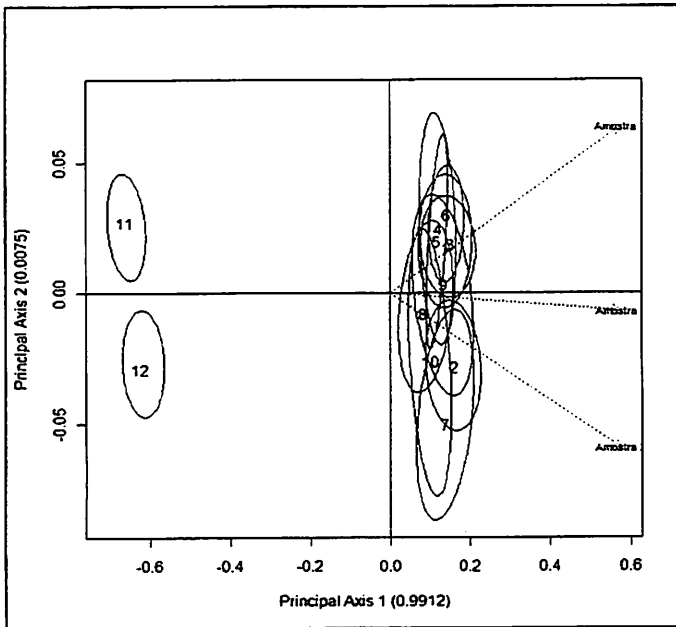


FIGURA 53 Elipses de confiança de 95%, geradas pela técnica de GPA para os doze laboratórios, de acordo com os critérios estabelecidos para a simulação 3

6.4 Discussão geral

Como síntese dos resultados, verificou-se que, no caso da simulação 1, para a maioria das técnicas, a distinção dos dois grupos não foi considerada totalmente satisfatória, apresentando, por exemplo, no caso da análise de

agrupamento, menor valor de correlação cofenética em relação às demais simulações. Já no caso da GPA, como comentado, ocorreu até mesmo a intercepção entre dois laboratórios de grupos distintos (Figura 54).

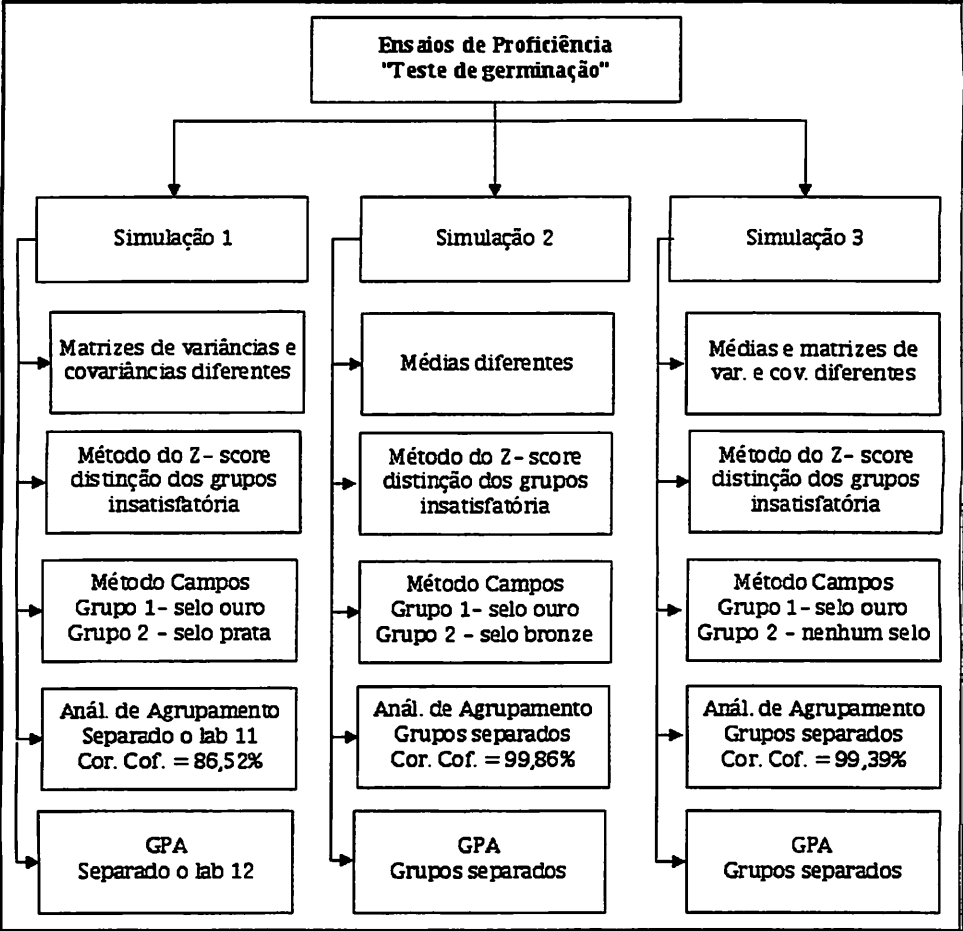


FIGURA 54 Resumo esquemático dos resultados das quatro técnicas estatísticas utilizadas na avaliação dos três ensaios de proficiência simulado para um teste de germinação, sendo os laboratórios de 1 a 10 considerados como grupo 1 e os laboratórios 11 e 12, como grupo 2

No caso do método do Z-score, para as três simulações, foi reavaliado o uso desta técnica considerando apenas o grupo 1, isto é, as notas e as classificações atribuídas foram recalculadas retirando-se os dados do grupo 2. Nesta nova avaliação ocorreram poucas mudanças tanto na classificação como nas medidas de viés, precisão e exatidão (resultados no Anexo F). Em alguns casos, aconteceu até mesmo redução da nota, parecendo indicar que esta técnica tem certa restrição de uso em algumas situações e que novos estudos devem ser feitos para avaliar o desempenho desse procedimento.

Numa discussão mais geral acerca do emprego do método do Z-score para análise dos resultados de um ensaio de proficiência, deve-se ressaltar que a utilização exclusiva desta técnica deve ser feita com cautela, pois pode colocar em risco algum laboratório quanto à sua real condição de qualidade. Por outro lado, o seu uso conjunto com outros procedimentos pode provocar dúvidas também quanto à validade dos resultados, como é o caso neste estudo.

A baixa distinção dos grupos, no caso do método do Z-score, pode estar associada ao tamanho da amostra, pois os intervalos de confiança para mediana são construídos utilizando-se um valor de t de Student alto (graus de liberdade igual a 4), provocando, conseqüentemente, um aumento da taxa de erro para um grau de confiança pré-fixado.

Vale salientar que, mesmo com os grupos já pré-estabelecidos na simulação, não foram utilizadas nem a técnica de análise discriminante nem a própria técnica de análise de agrupamento pelo método, por exemplo, das k-médias (agrupamento não hierárquico) para a confirmação dos grupos, pois o objetivo era justamente verificar quanto uma determinada técnica poderia detectar essas diferenças naturalmente, ou seja, tentando reproduzir uma situação com dados reais em que número de grupos não é pré-estabelecido, nem tampouco os seus elementos.

Se fosse feito um paralelo dos resultados obtidos com testes de germinação real para essas três situações de simulação, poderia se considerar que:

- caso da simulação 1 é, talvez, o menos esperado de acontecer, isto é, laboratórios apresentarem, para uma mesma espécie testada, médias iguais e variabilidade tão diferentes, a não ser que existam realmente diferenças quanto à qualidade dos laboratórios;
- para a simulação 2, em que as médias são diferentes, esta situação pode ocorrer na prática se os laboratórios também forem bem diferentes ou se houver algum problema no procedimento do teste ou no material enviado;
- a situação da simulação 3 não seria incomum em testes de germinação com espécies silvestres, pois, geralmente, ocorre maior variabilidade entre repetições e laboratórios, como verificado nos resultados do capítulo 2.

7 CONCLUSÕES

Para os três casos de simulação, existem limitações na utilização do método do Z-score, uma vez que não foi possível a distinção dos grupos de forma satisfatória. Já o método de Campos sempre atribuiu selos diferentes para os dois grupos.

A modificação proposta para o teste de Mojena mostrou ser apropriada para a definição do número de grupos, na presença de distâncias discrepantes.

As técnicas de análise de agrupamento e da análise generalizada de Procrustes, de forma geral, distinguiram bem os grupos. Deve-se ressaltar que, com a utilização destes dois procedimentos, foi possível fazer considerações acerca do desempenho dos laboratórios, com base apenas na representação gráfica dos resultados (dendrograma e biplot, respectivamente). Não foi necessária, então, nestes casos, a obtenção de índices de precisão e exatidão para avaliar os laboratórios no contexto de um ensaio de proficiência.

De forma geral, os laboratórios que apresentaram distorção nos resultados foram identificados por pelo menos uma das técnicas. Logo, esses resultados podem servir para nortear, por exemplo, as diretrizes do programa de Controle de Qualidade de um laboratório e o processo de credenciamento de laboratórios de sementes junto aos órgãos certificadores.

8 REFERÊNCIAS BIBLIOGRÁFICAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Ensaio de Proficiência por comparações interlaboratoriais**. Rio de Janeiro, 1999 (NBR ISO/IEC Guia 43-1).

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **Sistema de gestão da qualidade: fundamentos e vocabulário**. Rio de Janeiro, 2000 (NBR ISO 9000).

ASSOCIATION OF ANALYTICAL COMMUNITIES. International harmonized protocol for proficiency testing of (Chemical) analytical laboratories. **Journal of AOAC International**, v. 76, n. 4, 1993. Disponível em: <http://www.inmetro.gov.br/kits/ctle05_protocolo.pdf>. Acesso em: 10 dez. 2007.

BRASIL. Ministério da Agricultura e Reforma Agrária. **Regras para análise de sementes**. Brasília: SNDA/DNDV/CLAV, 1992. 365 p.

BUSSAB, W.O.; MIAZAKI, É.S.; ANDRADE, D.F. **Introdução à análise de agrupamentos**. São Paulo: ABE, 1990. 105 p.

CAMPOS, P.M. **Comparação de métodos para a avaliação do controle de qualidade de laboratório de análises de solo**. 1998. 99 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras.

CAMPOS, P.M.; MUNIZ, J.A.; OLIVEIRA, M.S. de; FERREIRA, D.F. Estimativa da repetibilidade e reprodutividade de laboratórios e via de aplicação no controle de qualidade de solo. **Ciência e Agrotecnologia**, Lavras, v. 23, n. 2, p.404-409, 1999.

CHUI, Q.S. **Programas interlaboratoriais**. Associação Brasileira de Polímeros (ABPol). 2006. Disponível em: <<http://www.abpol.com.br/pdf/PIABPo.pdf>>. Acesso em: 10 jan. 2008.

CHUI, Q.S.H.; ANTONOFF, H.B.; OLIVIERI, J.C. Utilização de índices r e R obtidos de Programas Interlaboratoriais para o controle de precisão de método

analítico: determinação de água por Karl Fische. **Química Nova**, v. 25, n. 4, p. 657-659, 2002.

CHUI, Q.S.; BISPO, J.M. de A.; IAMASHITA, C.O. O papel dos programas interlaboratoriais para a qualidade dos resultados analíticos. **Química Nova**, São Paulo, v. 27, n. 6, p. 9993-1003, 2004

FERREIRA, D.F. **Estatística multivariada**. Lavras: UFLA/DEX, 2008. 661p.

FERREIRA, E.B. **Some topics in Procrustes analysis applied to sensometrics**. 2007. 131 p. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras.

FERREIRA, E.B.; MAGALHÃES, F.A.R.; OLIVEIRA, M.S. de; FERREIRA, D.F. Sensory profile of the gorgonzola type cheese through bootstrap confidence regions in generalized procrustes analysis. **Revista do Instituto de Laticínios Cândido Tostes**, Juiz de Fora, v. 62, p. 535-540, 2007.

FERREIRA, E.B.; OLIVEIRA, M.S. de. **Sensiometria: uma abordagem com ênfase em Procrustes**. Santa Maria: UFSM, 2007. 71 p.

FERREIRA, E.B.; OLIVEIRA, M.S. de; FERREIRA, D.F.; MAGALHÃES, F.A.R. Sensory profile of gorgonzola via generalized procrustes analysis using R. **Boletim CEPPA**, Curitiba, v. 26, n. 1, p. 151-159, jan./jul. 2008.

GOWER, J.C. Generalized Procrustes analysis. **Psychometrika**, Berkeley, v. 40, p. 33-51, 1975.

INSTITUTO NACIONAL DE METROLOGIA. **Informações sobre ensaios de proficiência**. In: ENQUALB - Encontro para qualidade. São Paulo: 2005. Disponível em: <<http://www.inmetro.gov.br/laboratorios/ensaioProf.asp>>. Acesso em: 10 jan. 2008.

INTERNATIONAL SEED TESTING ASSOCIATION. **ISTA Referee Test Programme**: ISTA statistics seminar. Corvallis, Oregon, 2002.

INTERNATIONAL SEED TESTING ASSOCIATION. **The ISTA Proficiency Test Programme**. Switzerland, 2007.

JORGENSEN, J.; KRISTENSEN, K.; DEPRINS, H.; STEINER, A.M. Studies on variation between results of repeated germination tests made in the same laboratory. **Seed Science & Technology**, Zürich, v. 20, n. 2, p. 281-287, 1992.

MILES, S.R. Handbook of tolerances and of measures of precision for seed testing. **Proceedings of the International Seed Testing Association**, Zürich, v. 28, n. 3, p. 525-686, 1963.

MILLIGAN, G.W.; COOPER, M.C. An examination of procedures for determining the number of clusters in a data set. **Psychometrika**, Williamsburg, v. 50, p. 159-179, 1985.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2005. 295 p.

MOJENA, R. Hierarchical grouping methods and stopping rules: an evaluation. **Computer Journal**, London, v. 20, p. 359-363, 1977.

OLIVEIRA, M.S. de; CORRÊA, J.B.; CAMPOS, P.M. Programa interlaboratorial de controle de qualidade de análise de solo do estado de Minas Gerais (PROFERT-MG). **Boletim Informativo da SBSCS**, Campinas, v. 20, n. 1, p. 18-23, 1995.

OLIVEIRA, P.R.P.; CICERO, S.M. Causas de variação dos resultados das análises de sementes de capim colônia (*Panicum maximum Jacq.*). **Revista Brasileira de Sementes**, v. 18, n. 1, p. 122-128, 1996.

OLIVIERI, J.C. **Método gráfico para interpretação de resultados em programa interlaboratorial: elipse de confiança, técnicas e métodos**. São Paulo: IPT. 1988. 20p.

R Development Core Team (2008). **R: a language and environment for statistical computing**. Vienna, Áustria: R Foundation for Statistical Computing. Disponível em: <<http://www.R-project.org>>. Acesso em: 10 jan. 2008.

REDE METROLÓGICA. **Programa piloto de comparação interlaboratorial em sementes**. Porto Alegre, 2008. Disponível em: <http://www.redemetrologica.com.br/arquivo/laboratorial/ARQ_136.pdf>. Acesso em: 10 jan. 2009.

RENCHE, A.C. **Methods of multivariate analysis**. 2thed. New York: J. Wiley, 2002. 708 p.

RENNIE, W.J.; TOMLIN, M.N. Repeatability, reproducibility and interrelationship of results of tests on wheat seed samples infected with *Septoria nodorum*. **Seed Science & Technology**, Zürich, v. 12, n. 3, p. 863-880, 1984.

SOUZA, A.L.; FERREIRA, R.L.C.; XAVIER, A. **Análise de agrupamento aplicada à ciência florestal**. Viçosa, MG: SIF, 1997. 109 p.

TOTTI, R.; VENCOSKY, R.; BATISTA, L.A.R. Utilização do método de agrupamentos hierárquicos em acessos de *Paspalum* (*Graminea*(*Poaceae*)). **Semina: Ciências Exatas Tecnológicas**, Londrina, v.22, p. 25-35, dez. 2001.

WONG, S.K. Evaluation of the use of consensus values in proficiency testing programmes. **Accreditation and Quality Assurance**, Heidelberg, v.10, p.409-414, 2005.

ZORATO, F. Evolução do laboratório de análise de sementes. **Revista Seed News**, v. 9, n. 6, 2005. Disponível em: <
<http://www.seednews.inf.br/portugues/seed96/artigocapa96.shtml>>. Acesso em: 10 nov. 2007.

CONSIDERAÇÕES FINAIS

Neste estudo, de forma geral, deve-se considerar que:

- O desenvolvimento de rotinas específicas no software R serve para otimização das análises, sendo todos os relatórios gerados automaticamente e de forma rápida, além de possibilitar o uso de um software único, de código aberto, grátis e livre.
- As técnicas estatísticas propostas no capítulo 2, como, por exemplo, a análise discriminante para avaliar a classificação dos lotes e os testes F para comparação de dois métodos, podem ser utilizados em um programa de validação de métodos para sementes, uma vez que as análises possibilitam mostrar a eficiência desses procedimentos nas diversas etapas do processo.
- Para o teste de germinação em sementes de *Raphanus sativus* var *oleiferus*, os métodos utilizando temperatura alternada 20°-30°C e substratos areia ou papel estão validados por terem apresentado exatidão, precisão (repetibilidade e reprodutibilidade) e robustez.
- A análise generalizada de Procrustes, o método Campos e a análise de agrupamento, apresentadas no capítulo 3, são técnicas promissoras para a avaliação dos resultados de ensaios de proficiência na área de sementes.
- No caso dos ensaios de proficiência na área de sementes, a utilização simultânea das quatro técnicas estatísticas (as três citadas no item anterior, acrescidas do método do Z-score) é inédita, sendo recomendável para potencializar a avaliação do desempenho dos laboratórios.

- Apesar dos indicativos positivos da adequabilidade das técnicas estatísticas propostas neste trabalho, tanto para programas de validação de métodos como para ensaios de proficiência, verifica-se a necessidade do desenvolvimento de mais alguns estudos teóricos, como, por exemplo, a avaliação do tamanho da amostra, do efeito de lote, da espécie, bem como outras técnicas para a detecção de valores e variâncias discrepantes.

ANEXOS

ANEXO A

Página

- FIGURA 1A Categorias de sementes encontradas no teste de germinação em lotes de sementes de nabo forrageiro. A - Plântula Normal; B – Anormal deformada; C – Anormal infeccionada; D – dormente e E – morta..... 205
- FIGURA 2A Categorias de sementes encontradas no teste de tetrazólio em lotes de sementes de nabo forrageiro. **Viáveis:** A - embrião com coloração rosa ou menos de 50% do embrião descolorido sem atingir o eixo embrionário e tecido com aspecto normal e firme. **Inviáveis:** B – mais de 50% dos cotilédones descoloridos; C - região do eixo embrionário descolorida; D – embrião com coloração ~~vermelho~~ carmim e E - embrião completamente descolorido..... ~~206~~
- FIGURA 3A Métodos adotadas para o teste de germinação em lotes de sementes de nabo forrageiro. A – método utilizando como substrato o papel germitest; B – nivelamento da areia fina; C – método utilizando como substrato areia fina..... 206

ANEXO B**Página**

TABELA 1B	Determinação pelo teste H_j da condição de homogeneidade de um lote utilizando os resultados de um teste de germinação de 5 subamostras (N) com 100 sementes cada (n).....	208
TABELA 2B	Identificação de <i>outliers</i> pelo método de Hampel para os resultados de um teste de germinação em dois laboratórios com dois lotes e três repetições cada	210
TABELA 3B	Identificação de <i>outliers</i> nas variâncias pelo teste de Levene utilizando os resultados do teste de germinação apresentados na Tabela 2B.....	211
TABELA 4B	Determinação da amplitude dentro do laboratório, da variância e dos limites críticos de repetibilidade utilizando os resultados do teste de germinação apresentados na Tabela 2B.....	212
TABELA 5B	Determinação da amplitude entre laboratórios, da variância e dos limites críticos de reprodutibilidade utilizando os resultados do teste de germinação apresentados na Tabela 2B.....	213
TABELA 6B	Determinação dos valores de h e k.....	214
TABELA 7B	Resultados de um teste de germinação em oito laboratórios para dois lotes (grupo) e duas repetições cada.....	215
TABELA 8B	Valores da função discriminante para os dois grupos.....	218
TABELA 9B	Matriz de confusão simples – regra linear.....	219
TABELA 10B	Identificação de <i>outlier</i> pelo método do Z-score nos resultados médios de um teste de germinação com nove laboratórios em duas amostras.....	220

ANEXO B**Página**

TABELA 11B	Determinação do Z-score para os resultados médios de um teste de germinação com nove laboratórios em duas amostras.....	221
TABELA 12B	Determinação do viés, precisão, exatidão pelo método do Z-score dos nove laboratórios em duas amostras.....	222
TABELA 13B	Resultados de um teste de germinação em três laboratórios com duas amostras e quatro repetições cada.....	222
TABELA 14B	Determinação dos limites inferiores e superiores e valores de d_1 para os três laboratórios em duas amostras e quatro repetições cada.....	223
TABELA 15B	Determinação de d_1 (com base nos limites da Tabela 14B) para os três laboratórios em duas amostras e quatro repetições cada	224
TABELA 16B	Determinação dos limites inferiores e superiores e valores de d_2 para os três laboratórios em duas amostras e quatro repetições cada.....	225
TABELA 17B	Determinação de d_2 (com base nos limites da Tabela 16B) para os três laboratórios em duas amostras e quatro repetições cada.....	225
TABELA 18B	Determinação de medidas auxiliares, exatidão, precisão e classificação pelo método de Campos para os três laboratórios em duas amostras.....	226
TABELA 19B	Resultados médios de um teste de germinação com cinco laboratórios em duas amostras.....	226
TABELA 20B	Resultados das distâncias euclidianas para combinações duas a duas de quatro laboratórios em duas amostras.....	226
FIGURA 1B	Dendrograma para o agrupamento de 5 laboratórios a partir dos dados de germinação utilizando o método da ligação média.....	230

ANEXO B**Página**

TABELA 21B	Determinação do ponto de corte para determinação do número de grupos, procedimento de Mojena (1977).....	230
TABELA 22B	Medidas das distâncias euclidianas e previstas para o cálculo da correlação cofenética.....	231

ANEXO C

Página

FIGURA 1C	Descrição e interpretação das medidas utilizadas na construção de um boxplot.....	232
------------------	--	------------

ANEXO E**Página**

TABELA 1E	Exemplo da montagem de um banco de dados para a realização do teste H.....	236
TABELA 2E	Exemplo da montagem de um banco de dados para a realização do método de Hampel.....	237
TABELA 3E	Exemplo da montagem de um banco de dados para a Análise Discriminante.....	237
TABELA 4E	Exemplo da montagem de um banco de dados para o método do Z-score e método de Campos (1998).....	238
TABELA 5E	Exemplo da montagem de um banco de dados para a Análise de Agrupamento.....	239
FIGURA 1E	Exemplo de alguns relatórios gerados no <i>software</i> R para o Capítulo 2.....	240
FIGURA 2E	Exemplo de alguns relatórios gerados no <i>software</i> R para o Capítulo 3.....	241
FIGURA 3E	Programa para execução da FunçãoTesteH no <i>software</i> R.....	242

ANEXO F**Página**

TABELA 1F	Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 1.....	243
TABELA 2F	Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 2	243
TABELA 3F	Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 3.....	244

ANEXO A

Neste anexo são apresentadas as instruções para a realização e avaliação do teste de germinação, e um questionário para verificar a execução prática de cada um dos métodos.

Etapas para realização do teste de germinação em sementes de nabo forrageiro (*Raphanus sativus* L. var. *oleiferus* Metzg).

Organizadores do teste: UFLA/MAPA

e-mail:sementes@ufla.br

Endereço:Campus da UFLA/DAG/Setor de Sementes-CP3037-CEP37200-000
Lavras MG

Telefone: 35 3829 1055 ou 35 3829 1084

1) Orientações gerais

- a) O teste deverá ser iniciado imediatamente após o recebimento das amostras dos cinco lotes de sementes de nabo forrageiro. Caso não seja possível, as amostras devem ser armazenadas em local fresco e seco.
- b) O teste de germinação deverá ser realizado em dois tipos de substrato: areia fina e papel mata-borrão; e com 8 repetições de 50 sementes por lote;
- b) As caixas acrílicas tipo gerbox deverão ser acondicionadas em câmara de germinação regulada à temperatura alternada 20°C-30°C, com fotoperíodo de 8/16 horas;

- c) Após 3 dias deverão ser retiradas as plântulas normais, e os resultados dessa primeira contagem (expressos em porcentagem) registrados na ficha de avaliação (Anexo A1);
- d) Após 7 dias as sementes deverão ser avaliadas, em porcentagem, de acordo com as seguintes categorias: plântulas normais (N), anormais defeituosas (AD), anormais infeccionadas (AI), mortas (M) e dormentes (D); e os resultados preenchidos na ficha de avaliação (Anexo A1).
- e) Para padronização dos resultados do teste observe a Figura 1A (Anexo A3), correspondente as categorias N, AD, AI, M e D.
- f) Para uma categorização adequada das sementes remanescentes, em mortas ou dormentes, deverá ser utilizado o teste de Tetrazólio com solução de 0,5% por 3 horas a 40 °C, sendo as viáveis classificadas como dormentes e inviáveis como mortas. (Figura 2A – Anexo A3) e registrados na ficha de avaliação (Anexo A2)

2) Procedimento inicial

Determinação e uniformização do grau de umidade das sementes

- a) O grau de umidade deverá ser determinado pelo método de estufa a $105 \pm 3^\circ\text{C}$ por 24 horas (BRASIL, 1992), com 2 repetições de 10 gramas de sementes por lote.

3) Substrato papel mata-borrão

- a) O papel deverá ser umedecido com água destilada numa quantidade correspondente a 2,5 vezes da sua massa;
- b) As sementes em numero de 50 por repetição ou gerbox, deverão ser dispostas de forma alternada sobre o papel (Figura 3A – Anexo A3);



c) Após a montagem as caixas deverão ser transferidas para o germinador, seguindo as orientações gerais já descritas no item 1.

4) Substrato areia

a) 500 gramas de areia fina deverão ser peneiradas, lavadas e secas ao livre por um 24h.

b) Esterilizar a areia em estufa a 200°C/2h, devendo ser utilizada no máximo em até 15 dias.

c) Antes da montagem do teste a areia deverá ser umedecida em 50% da sua capacidade de campo;

d) Para cada caixa acrílica tipo “gerbox” deverá ser utilizado 215 gramas de areia;

e) A areia deverá ser nivelada, utilizando um pedaço de papel germitest dobrado. O papel deverá ser passado no sentido vertical, numa quantidade de vezes suficiente para o seu nivelamento completo (Figura 3A– Anexo A3).

f) Após o nivelamento, as sementes deverão ser dispostas alternadamente sobre a areia (Figura 3A– Anexo A3).

g) As caixas deverão ser transferidas para o germinador, seguindo as orientações gerais já descritas no item 1.

5) Comunicação dos resultados

a) Deverão ser enviadas via correio (endereço acima) uma cópia das fichas de avaliação, bem como o questionário (Anexo A4) devidamente preenchidos. Os originais deverão ser mantidos no laboratório até o final do processo.

b) Os resultados deverão também ser enviados via mail, pelo preenchimento dos anexos e encaminhamento ao organizador do teste. Endereço: sementes@ufla.br

c) Cada laboratório receberá um código identificador para poder verificar os seus resultados no boletim parcial. Qualquer inconsistência nos resultados o laboratório deverá ser comunicada imediatamente ao organizador do teste.

6) Orientações finais

a) Caso seja necessário um re-teste o organizador deverá ser informado. Esse procedimento deverá ser adotado nas seguintes circunstâncias: problemas no equipamento durante a condução do teste; resultados aparentemente inconsistentes, sendo que neste caso as possíveis causas devem ser investigadas imediatamente e caso haja dúvidas o organizador do teste deverá ser consultado.

b) O material sobressalente só deverá ser descartado após o término do processo de validação, data que será informada via e-mail.

c) Qualquer dúvida e/ou possíveis problemas no material enviado poderão ser encaminhadas via e-mail para o organizador do teste.

ANEXO A1

FICHA DE AVALIAÇÃO

Teste para validação de método

Espécie: *Raphanus sativus* L. var. *oleiferus* Metzg

Escopo: Germinação

Código do laboratório: _____

Método – Substrato: Areia Papel

Lote: 1 2 3 4 5

Rep	PN Contagem ao 3º dia	PN Contagem ao 7º dia	PN total	PAD	PAI	SM	SD
1							
2							
3							
4							
5							
6							
7							
8							
total							

*PN = Plântulas normais; PAD = Plântulas anormais deformadas; PAI = Plântulas anormais infeccionadas, SM = Sementes mortas; SD = Sementes duras.

Observações:

ANEXO A2
FICHA DE AVALIAÇÃO

Teste para validação de método

Espécie: *Raphanus sativus* L. var. *oleiferus* Metzg

Escopo: Tetrázólio

Código do laboratório:

Método – Substrato: Areia Papel

Lote: 1 2 3 4 5

Repetição	Viáveis	Inviáveis
1		
2		
3		
4		
5		
6		
7		
8		

Observações:

ANEXO A3

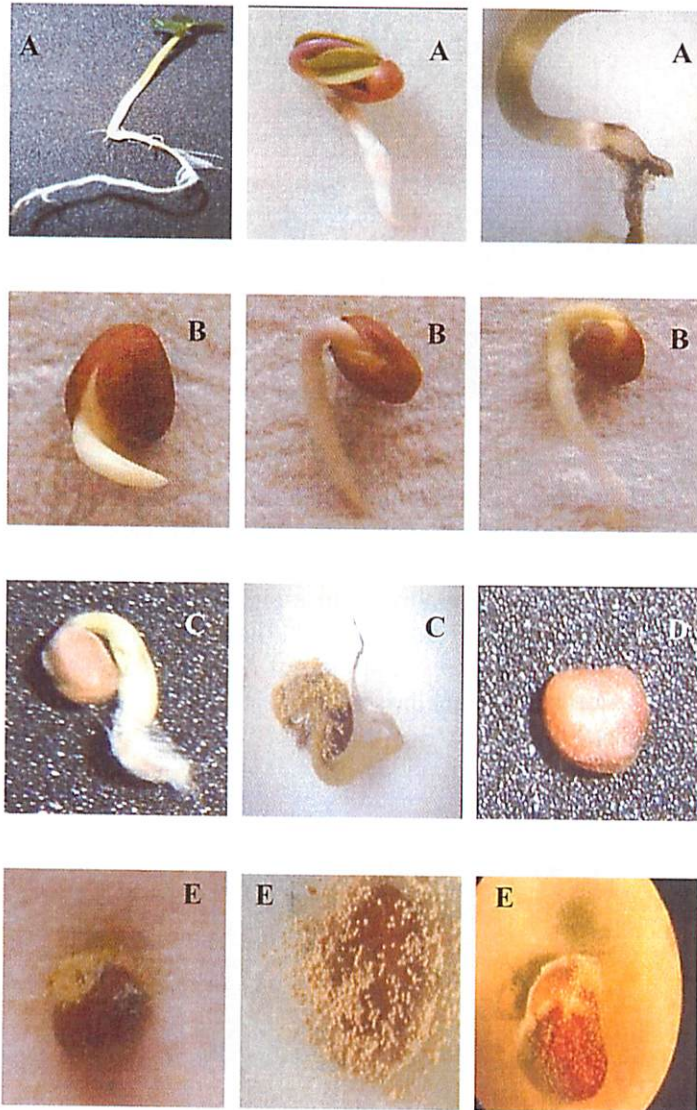


FIGURA 1A Categorias de sementes encontradas no teste de germinação em lotes de sementes de nabo forrageiro. A - Plântula Normal; B - Anormal deformada; C - Anormal infeccionada; D - dormente e E - morta ao final de 7 dias

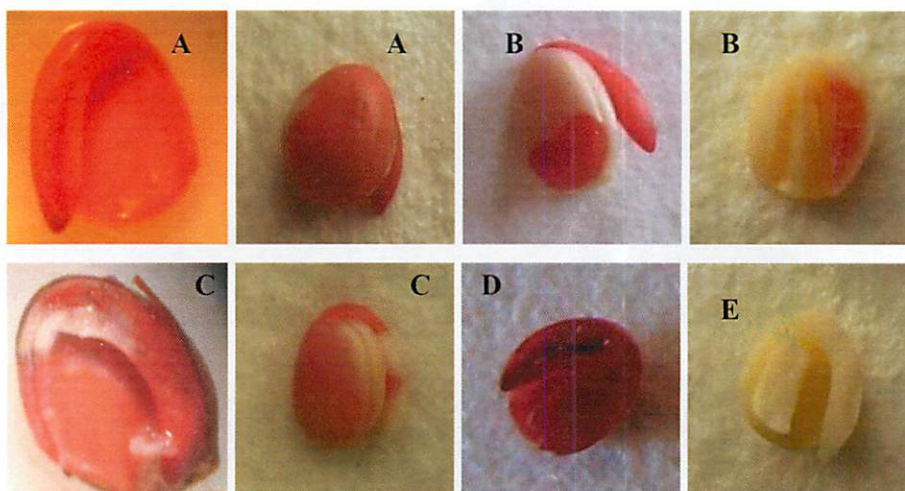


FIGURA 2A. Categorias de sementes encontradas no teste de tetrazólio em lotes de sementes de nabo forrageiro. **Viáveis:** A - embrião com coloração rosa ou menos de 50% do embrião descolorido sem atingir o eixo embrionário e tecido com aspecto normal e firme. **Inviáveis:** B - mais de 50% dos cotilédones descoloridos; C - região do eixo embrionário descolorida; D - embrião com coloração vermelho carmim e E - embrião completamente descolorido

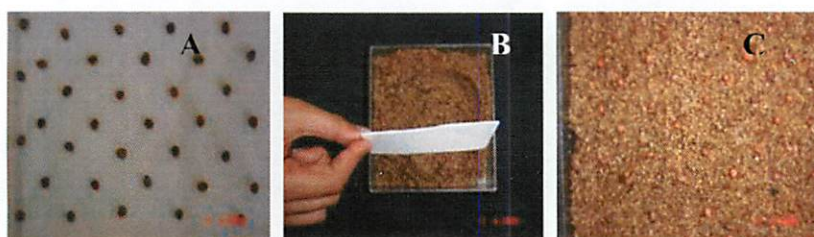


FIGURA 3A Métodos adotados para o teste de germinação em lotes de sementes de nabo forrageiro. A - método utilizando como substrato o papel germitest; B - nivelamento da areia fina; C - método utilizando como substrato areia fina

ANEXO A4
QUESTIONÁRIO DE AVALIAÇÃO

Teste para validação de método

Espécie: *Raphanus sativus* L. var. *oleiferus* Metzg

Escopo: Germinação

Código do laboratório:

1) Os métodos estão descritos de forma clara e sem ambigüidades?

Sim Não

2) Qual dos dois métodos você considerou mais “fácil” no momento da:

a) Pré-montagem (Preparação do material)

Areia Papel Ambas Nenhuma delas

Se for o caso, indique as principais dificuldades:

b) Montagem

Areia Papel Ambas Nenhuma delas

Se for o caso, indique as principais dificuldades:

c) Manipulação

Areia Papel Ambas Nenhuma delas

Se for o caso, indique as principais dificuldades:

d) Avaliação (Contagem – visualização)

Areia Papel Ambas Nenhuma delas

Se for o caso, indique as principais dificuldades:

3) Comentários gerais:

ANEXO B

Neste anexo são apresentadas passo a passo algumas técnicas estatísticas utilizadas neste estudo por meio de exemplos numéricos. Ressalta-se que os dados utilizados são apenas para fins didáticos, não havendo, portanto uma preocupação excessiva quanto ao tamanho da amostra, número de laboratórios selecionados, dentre outras medidas.

Teste H

Para verificar se um lote pode ser considerado homogêneo pelo teste H é necessário calcular as variâncias: observada (V_j) e teórica (W_j) de um teste de germinação (Tabela 1B). Determinar a estatística H_j e comparar com um nível de significância estabelecido.

No caso do teste de germinação, o valor do fator para variação adicional é sempre igual a 1,1. Sendo a variável aleatória em questão a porcentagem de germinação (Q)

TABELA 1B Determinação pelo teste H_j da condição de homogeneidade de um lote utilizando os resultados de um teste de germinação de 5 subamostras (N) com 100 sementes cada (n)

Subamostra	q_i	q_i^2	\bar{y}	v	w
1	98	9.604			
2	96	9.216			
3	89	7.921	93,40	14,80	6,78
4	94	8.836			
5	90	8.100			
	$\Sigma = 467$	$\Sigma = 43.667$			

Cálculo de v e w

$$v = \frac{N \sum_{i=1}^N q_i^2 - \left(\sum_{i=1}^N q_i \right)^2}{N(N-1)} = \frac{5(98^2 + \dots + 90^2) - (98 + \dots + 90)^2}{5(5-1)} = 14,80$$

$$w = \frac{\bar{q}(100 - \bar{q})}{n} \cdot f = \frac{93,40(100 - 93,40)}{100} \cdot 1,1 = 6,78$$

Determinação da estatística H

$$H = \frac{v}{w} \cdot f = \frac{14,80}{6,78} \cdot 1,1 = 1,08 \Rightarrow \text{valor } p = 0,0942$$

O lote avaliado pode ser considerado homogêneo pelo teste H, ao nível de significância de 1%.

Método de Hampel

Para a detecção de um *outlier* pelo método de Hampel é necessário calcular: mediana de um conjunto de dados para um determinado nível do estudo, neste exemplo, o nível é o lote; a diferença entre cada observação e a mediana (desvios); a mediana desses desvios (MAD) e a regra de decisão - *outlier* > 5,2MAD (Tabela 2B).

TABELA 2B Identificação de *outliers* pelo método de Hampel para os resultados de um teste de germinação em dois laboratórios com dois lotes e três repetições cada

Lote 1					Lote 2				
Lab.	Rep.	y_{1lk}	$ y_{1lk} - \bar{y}_{1..} $	Condição	Lab.	Rep.	y_{2lk}	$ y_{2lk} - \bar{y}_{2..} $	Condição
	1	88	2	não		1	92	1,5	não
1	2	91	1	não	1	2	89	1,5	não
	3	93	3	não		3	93	2,5	não
	1	93	3	não		1	65	25,5	outlier
2	2	89	1	não	2	2	85	5,5	não
	3	86	4	não		3	92	1,5	não
	$\bar{y}_{1..}$	90,0	↓	Mediana		$\bar{y}_{2..}$	90,5	↓	Mediana
	MAD_1		2,5			MAD_2		2,0	
	$5,2 MAD_1$		13,0			$5,2 MAD_2$		10,4	

Somente a observação y_{221} foi considerada um outlier, uma vez que $|y_{221} - \bar{y}_{2..}| = 25,5 > 5,2MAD_2 = 10,4$

Teste de Levene

Para o uso desta técnica é necessário calcular vários tipos de desvios em relação: a média de um subnível (laboratório), a média geral dos desvios de um subnível, a média geral dos desvios de um nível - lote (Tabela 3B). Determinar a estatística W e comparar com um determinado nível de significância.

$$\text{Cálculo da estatística: } W_j = \frac{(N_j - p_j) \sum_{l=1}^{p_j} n_{jl} (\bar{Z}_{jl.} - \bar{Z}_{j..})^2}{(p_j - 1) \sum_{l=1}^{p_j} \sum_{k=1}^{n_{jl}} (Z_{jlk} - \bar{Z}_{jl.})^2}$$

$$\text{a) } W_1 = \frac{(6 - 2) \cdot [3 \cdot (0,11) + 3 \cdot (0,11)]}{(2 - 1) \cdot (3,19 + 6,74)} = \frac{2,68}{9,93} = 0,2686$$

valor $p = 0,6316$ (distribuição F com 1 e 4 graus de liberdade)

$$b) W_2 = \frac{(5-2) \cdot [3 \cdot (0,60) + 2 \cdot (1,36)]}{(2-1) \cdot (1,41 + 0,00)} = \frac{13,61}{1,41} = 9,6710$$

valor $p = 0,0529$ (distribuição F com 1 e 3 graus de liberdade)

TABELA 3B Identificação de *outliers* nas variâncias pelo teste de Levene utilizando os resultados do teste de germinação apresentados na Tabela 2B

Lote	Lab	Rep	y_{jlk}	$z_{jlk} = y_{jlk} - \bar{y}_{jl} $	$(z_{jlk} - \bar{z}_{jl})^2$	$(\bar{z}_{jl} - \bar{z}_{j..})^2$	
1	1	1	88	$ 88 - 90,67 = 2,67$	$(2,67 - 1,78)^2 = 0,79$	0,11	
		2	91	0,33	2,09		
		3	93	2,33	0,31		
			$\bar{y}_{11.}$	90,67	$\bar{z}_{11.} = 1,78$	$\Sigma = 3,19$	
	2	1	93	3,67	1,49	0,11	
		2	89	0,33	4,46		
3		86	3,33	0,79			
		$\bar{y}_{21.}$	89,33	$\bar{z}_{21.} = 2,44$	$\Sigma = 6,74$		
				$\bar{z}_{1..} = 2,11$			
2	1	1	92	0,67	0,79	0,60	
		2	89	2,33	0,60		
		3	93	1,67	0,01		
			$\bar{y}_{12.}$	91,33	$\bar{z}_{12.} = 1,56$	$\Sigma = 1,41$	
	2	1				0,00	
		2	85	3,50	0,00		
3		92	3,50	0,00			
		$\bar{y}_{22.}$	88,50	$\bar{z}_{22.} = 3,50$	$\Sigma = 0,00$	1,36	
				$\bar{z}_{2..} = 1,87$			

Nenhum laboratório, ao nível de significância de 1%, é considerado um *outlier*.

Condições de repetibilidade

Para avaliar as condições de repetibilidade de um determinado laboratório e lote, é necessário calcular: a variância dos dados, a variância de repetibilidade, o limite crítico e a amplitude dentro dos laboratórios (Tabela 4B).

Exemplo para o cálculo da variância do lab1 e lote 1:

$$s_{11}^2 = \frac{\sum_{k=1}^{n_{j1}} (y_{j1k} - \bar{y}_{j1.})^2}{(n_{j1} - 1)} = \frac{(88 - 90,67)^2 + (91 - 90,67)^2 + (93 - 90,67)^2}{(3 - 1)} = 6,33$$

TABELA 4B Determinação da amplitude dentro do laboratório, da variância e dos limites críticos de repetibilidade utilizando os resultados do teste de germinação apresentados na Tabela 2B

Lote	Lab	Rep	y_{j1k}	n_{j1}	s_{j1}^2	Lr_{j1}	s_{ij}^2	D_j	r_j
1	1	1	88	3	6,33	5	9,33	4,03	12,31
		2	91						
		3	93						
			$\bar{y}_{11.}$	90,67					
	2	1	93	3	12,33	7			
		2	89						
3		86							
		$\bar{y}_{21.}$	89,33	$\bar{n}_{1.} = 3$					
2	1	1	92	3	4,33	4	11,06	3,68	12,24
		2	89						
		3	93						
			$\bar{y}_{12.}$	91,33					
	2	1	85	2	24,50	7			
		2	85						
3		92							
		$\bar{y}_{22.}$	88,50	$\bar{n}_{2.} = 2,5$					

Os laboratórios apresentam repetibilidade aceitável, com grau de confiança de 99%, pois as amplitudes ($LR_{j,i}$) são menores que os limites críticos (r_j).

Condições de reprodutibilidade

Para verificar as condições de reprodutibilidade de uma determinada “repetição” e lote, é necessário calcular: a variância entre laboratórios, a variância de reprodutibilidade, o limite crítico e a amplitude entre os laboratórios (Tabela 5B).

TABELA 5B Determinação da amplitude entre laboratórios, da variância e dos limites críticos de reprodutibilidade utilizando os resultados do teste de germinação apresentados na Tabela 2B

Lote	Rep.	Lab.	y_{ijk}	$LR_{j,k}$	S_{dj}^2	$S_{L_j}^2$	$S_{R_j}^2$	D_j	R_j
1	1	1	88	5	2,66	-2,22	7,11	4,03	10,75
		2	93						
	2	1	91	2					
		2	89						
	3	1	93	7					
		2	86						
2	1	1	92	0	10,03	-0,42	10,63	3,68	11,99
		2							
	2	1	89	4					
		2	85						
	3	1	93	1					
		2	92						

Exemplo do cálculo S_{dj}^2 e \bar{n}_j para o lote 1, considerando as informações da Tabela 5B.

$$s_{d1}^2 = \frac{\sum_{i=1}^2 n_{i1} (\bar{y}_{i1} - \bar{y}_{1..})^2}{(p_1 - 1)} = \frac{[3(90,67 - 90)^2 + 3(89,93 - 90)^2]}{(2 - 1)} = 2,66$$

$$\bar{n}_1 = \frac{1}{(p_1 - 1)} \left(\sum_{i=1}^2 n_{i1} - \frac{\sum_{i=1}^2 n_{i1}^2}{\sum_{i=1}^2 n_{i1}} \right) = \frac{1}{(2 - 1)} \left[(3 + 3) - \frac{3^2 + 3^2}{3 + 3} \right] = 3$$

As repetições apresentam reprodutibilidade aceitável, com garu de confiança de 99%, pois as amplitudes são menores que os limites críticos (R_j).

Valores de h e k

O valor de k para cada laboratório e lote é determinado pela divisão da variância entre laboratórios pela variância de reprodutibilidade. Para obter o valor de h são calculados dois tipos de desvios em relação: a média do laboratório, a média geral do lote (Tabela 6B)

TABELA 6B Determinação dos valores de h e k

Lote	Lab	Rep	y_{jik}	s_{ij}	s_{ij}	k_{jl}	$\bar{y}_{j..}$	h_{jl}
1	1	1	88	2,52	3,05	0,8239	90,00	1,2247
		2	91					
		3	93					
			$\bar{y}_{11.}$	90,67				
	2	1	93	3,51	3,33	1,1496	89,92	-1,2247
		2	89					
3		86						
		$\bar{y}_{12.}$	89,33					
2	1	1	92	2,08	3,33	0,6257	89,92	1,0954
		2	89					
		3	93					
			$\bar{y}_{21.}$	91,33				
	2	1	85	4,95	3,33	1,4883	88,50	-1,0954
		2	85					
3		92						
		$\bar{y}_{22.}$	88,50					

Exemplo para o valor de h do lab1 e lote 1:

$$h_{11} = \frac{(\bar{Y}_{11} - \bar{Y}_{1..})}{\sqrt{\sum_{i=1}^{p_1} (\bar{Y}_{1i} - \bar{Y}_{1..})^2 / (p_1 - 1) \bar{n}_1}} = \frac{(90,67 - 90,00)}{\sqrt{(90,67 - 90,00)^2 + (89,33 - 90,00)^2 / (2 - 1) \cdot 3}} = 1,2247$$

Análise Discriminante

Como descrito na seção 5.3.6 do capítulo 2, esta técnica pode ser desenvolvida em 4 etapas.

1ª etapa) Avaliação da suposição de igualdade das matrizes de covariância

São calculados os vetores de médias, as matrizes de variâncias e covariâncias (Tabela 7B) e aplicado o teste de Box's M (que não será apresentado).

TABELA 7B Resultados de um teste de germinação em oito laboratórios para dois lotes (grupo) e duas repetições cada

Lab	Grupo	Amostra 1 (a ₁)	Amostra 2 (a ₂)
1	1	94	92
2	1	90	94
3	1	88	92
4	1	90	98
5	1	92	96
6	1	90	98
7	1	94	96
8	1	96	96
1	2	72	74
2	2	98	88
3	2	86	84
4	2	64	78
5	2	78	88
6	2	76	84
7	2	64	84
8	2	66	74



Ex: Vetor da observação 1 do grupo 1

$$x_{11}^T = \begin{bmatrix} x_{111} & x_{112} \\ 94 & 92 \end{bmatrix}$$

Vetor das médias para cada grupo

$$\bar{x}_1 = \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} \quad \bar{x}_2 = \begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix}$$

Matrizes de covariâncias e variâncias para cada grupo

$$s_1 = \begin{bmatrix} 7,36 & 0,36 \\ 0,36 & 5,64 \end{bmatrix} \quad s_2 = \begin{bmatrix} 141,43 & 42,71 \\ 42,71 & 32,50 \end{bmatrix}$$

Exemplo do cálculo da determinação de \bar{x}_1

$$\bar{x}_1 = \frac{1}{n_1} \sum_{m=1}^{n_1} x_{1,m} = \frac{1}{8} \left(\begin{bmatrix} 94 \\ 92 \end{bmatrix} + \begin{bmatrix} 90 \\ 94 \end{bmatrix} + \begin{bmatrix} 88 \\ 92 \end{bmatrix} + \begin{bmatrix} 90 \\ 98 \end{bmatrix} + \begin{bmatrix} 92 \\ 96 \end{bmatrix} + \begin{bmatrix} 90 \\ 98 \end{bmatrix} + \begin{bmatrix} 94 \\ 96 \end{bmatrix} + \begin{bmatrix} 96 \\ 96 \end{bmatrix} \right)$$

$$\bar{x}_1 = \frac{1}{8} \begin{bmatrix} 734 \\ 762 \end{bmatrix} = \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix}$$

Exemplo do cálculo da determinação de s_1

$$s_1 = \frac{1}{7} \left\{ \left(\begin{bmatrix} 94 \\ 92 \end{bmatrix} - \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} \right) \left(\begin{bmatrix} 94 \\ 92 \end{bmatrix} - \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} \right)^T + \dots + \left(\begin{bmatrix} 96 \\ 96 \end{bmatrix} - \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} \right) \left(\begin{bmatrix} 96 \\ 96 \end{bmatrix} - \begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} \right)^T \right\}$$

$$s_1 = \frac{1}{7} \left(\begin{bmatrix} 2,25 \\ -3,25 \end{bmatrix} \begin{bmatrix} 2,25 & -3,25 \end{bmatrix} + \dots + \begin{bmatrix} 4,25 \\ 0,75 \end{bmatrix} \begin{bmatrix} 4,25 & 0,75 \end{bmatrix} \right)$$

$$s_1 = \frac{1}{7} \left(\begin{bmatrix} 5,06 & -7,31 \\ -7,31 & 10,56 \end{bmatrix} + \dots + \begin{bmatrix} 18,06 & 3,19 \\ 3,19 & 0,56 \end{bmatrix} \right)$$

$$s_1 = \frac{1}{7} \begin{bmatrix} 51,50 & 2,50 \\ 2,50 & 39,50 \end{bmatrix} = \begin{bmatrix} 7,36 & 0,36 \\ 0,36 & 5,64 \end{bmatrix}$$

2ª etapa) Determinação das funções discriminantes lineares de Fisher

Supondo a igualdade das matrizes das variâncias e covariâncias, determinam-se as funções discriminantes lineares de Fisher, por meio de algumas medidas auxiliares: \bar{X} , \hat{B}_0 e S_p .

Cálculo de \bar{X} , \hat{B}_0 e S_p com os dados apresentados na Tabela 7B

$$\bar{X} = \frac{1}{2} \sum_{i=1}^2 \bar{X}_i = \frac{1}{2} \left(\begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} + \begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} 167,25 \\ 177,00 \end{bmatrix} = \begin{bmatrix} 83,62 \\ 88,50 \end{bmatrix}$$

$$\hat{B}_0 = \sum_{i=1}^2 (\bar{X}_i - \bar{X})(\bar{X}_i - \bar{X})^T$$

$$\hat{B}_0 = \left(\begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} - \begin{bmatrix} 83,62 \\ 88,50 \end{bmatrix} \right) \left(\begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} - \begin{bmatrix} 83,62 \\ 88,50 \end{bmatrix} \right)^T + \left(\begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix} - \begin{bmatrix} 83,62 \\ 88,50 \end{bmatrix} \right) \left(\begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix} - \begin{bmatrix} 83,62 \\ 88,50 \end{bmatrix} \right)^T$$

$$\hat{B}_0 = \left(\begin{bmatrix} 66,015 & 54,845 \\ 54,845 & 45,560 \end{bmatrix} + \begin{bmatrix} 66,015 & 54,845 \\ 54,845 & 45,560 \end{bmatrix} \right) = \begin{bmatrix} 132,03 & 109,69 \\ 109,76 & 91,12 \end{bmatrix}$$

$$S_p = \left(\frac{1}{\sum_{i=1}^g n_i - g} \right) \sum_{i=1}^g (n_i - 1) S_i = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{(n_1 + n_2) - 2}$$

$$S_p = \frac{(8-1) \begin{bmatrix} 7,36 & 0,36 \\ 0,36 & 5,64 \end{bmatrix} + (8-1) \begin{bmatrix} 141,43 & 42,71 \\ 42,71 & 32,50 \end{bmatrix}}{(8+8-2)}$$

$$S_p = \begin{bmatrix} 74,39 & 21,53 \\ 21,53 & 19,07 \end{bmatrix}$$

No caso em que $g = 2$, o vetor l pode ser obtido diretamente por $l = (\bar{X}_1 - \bar{X}_2)^T S_p^{-1}$, logo para este exemplo:

$$l = [0,019 \quad 0,686] \Rightarrow y = 0,019a_1 + 0,686a_2$$

3ª etapa) Regra de classificação

Para o caso em que $g=2$, calcula-se o ponto médio m entre duas médias univariadas amostrais dado por $\hat{m} = \frac{1}{2}(\bar{y}_1 + \bar{y}_2) = \frac{1}{2}(\bar{x}_1 - \bar{x}_2)^T s_p^{-1} (\bar{x}_1 + \bar{x}_2)$ e estabelecesse a regra de classificação para x como sendo: alocar x em τ_1 se $l^T x \geq \hat{m}$ e alocar em τ_2 caso contrário (Tabela 8B).

TABELA 8B Valores da função discriminante para os dois grupos

Lab	Grupo	função
1	1	64,89
2	1	66,19
3	1	64,78
4	1	68,94
5	1	67,60
6	1	68,94
7	1	67,64
8	1	67,68
1	2	52,13
2	2	62,23
3	2	59,26
4	2	54,72
5	2	61,85
6	2	59,07
7	2	58,84
8	2	52,02

Assim, tomando com exemplo vetor da observação 1 do grupo 1

$x_{1,1} = \begin{bmatrix} 94 \\ 92 \end{bmatrix}$ a classificação pode ser obtida da seguinte forma:

$$\hat{m} = \frac{1}{2} \left(\begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} - \begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix} \right)^T \begin{bmatrix} 74,39 & 21,53 \\ 21,53 & 19,07 \end{bmatrix}^{-1} \left(\begin{bmatrix} 91,75 \\ 95,25 \end{bmatrix} + \begin{bmatrix} 75,50 \\ 81,75 \end{bmatrix} \right) = 62,33$$

$$I^T x = \begin{bmatrix} 0,019 \\ 0,686 \end{bmatrix} \begin{bmatrix} 94 & 92 \end{bmatrix} = 64,89 > 62,33$$

Pela regra de classificação adotada esta observação é considerada sendo do grupo 1.

4ª etapa) Avaliação da função de classificação

Com as observações classificadas é construída a matriz de confusão (Tabela 9B) e determinada a taxa de erro geral. Neste exemplo, não está sendo apresentado o método de validação cruzada utilizado no capítulo 2.

TABELA 9B Matriz de confusão simples – regra linear

grupo verdadeiro	grupo classificado		total
	τ_1	τ_2	
τ_1	8	0	8
τ_2	0	8	8
Total	8	8	16

Como todas as observações foram classificadas corretamente não há taxas de erro a serem determinadas.

Método do Z score

Como apresentado na seção 5.2.1 do capítulo 3, esta técnica pode ser desenvolvida em 6 etapas.

1ª etapa) Identificação de valores discrepantes (*outliers*)

São estabelecidos, para cada nível de estudo (neste caso a amostra), intervalos de confiança para mediana, e considera-se um *outlier* se o resultado médio de um teste de germinação não estiver contido neste intervalo (Tabela 10B).

TABELA 10B Identificação de *outlier* pelo método do Z-score nos resultados médios de um teste de germinação com nove laboratórios em duas amostras

	Amostra 1		Amostra 2	
	Lab.	Média	Lab.	Média
	6	98	6	97
	5	95	5	95
	1	92	1	91
	3	92	4	89
Mediana	4	90	9	89
	7	86	8	86
	9	86	3	81
	8	85	7	80
(Outlier)	2	60	2	54
Média	87,11		84,67	
Mediana - \tilde{y}_a	90,00		89	
Desvio padrão (acima da mediana) - s_a^*	2,87		3,65	
Graus de liberdade	3		3	
$t_{a(0,005;3)}$	5,84		5,84	
Limite Inferior - $\tilde{y}_a - s_a^* t_{a(0,005;3)}$	73,24		67,67	
Limite Superior - $\tilde{y}_a + s_a^* t_{a(0,005;3)}$	106,76		110,32	

¹ Nível de confiança de 99%.

2ª etapa) Determinação do Z- score

Calcula-se a média e o desvio padrão para cada amostra, desconsiderando os *outliers* identificados na etapa anterior, e determina-se o valor do Z-score para cada laboratório (Tabela 11B).

TABELA 11B Determinação do Z-score para os resultados médios de um teste de germinação com nove laboratórios em duas amostras

Amostra 1			Amostra 2		
Lab.	Média	Z_{1i}	Lab.	Média	Z_{2i}
6	98	1,61	6	97	1,41
5	95	0,97	5	95	1,07
1	92	0,32	1	91	0,41
3	92	0,32	4	89	0,08
4	90	-0,11	9	89	0,08
7	86	-0,97	8	86	-0,41
9	86	-0,97	3	81	-1,24
8	85	-1,18	7	80	-1,41
2	60	-6,55	2	54	-5,70
Média - \bar{y}_a	87,11		84,67		
Desvio padrão - s_a	2,87		3,65		

3ª etapa) Classificação numa rodada

Não estão sendo apresentadas nem a classificação, nem a nota que seria atribuída a cada laboratório, uma vez que, neste exemplo, não há três amostras.

4ª a 6ª etapas) Cálculo do viés, precisão e exatidão

Pelos resultados obtidos, o laboratório 6 tem exatidão considerada como crítica ($1,500 < E_6^{(t)} < 1,999$) e o laboratório 2 como inaceitável. ($E_2^{(t)} \geq 2,000$)

Todos os laboratórios apresentam boa precisão. (Tabela 12B)

TABELA 12B Determinação do viés, precisão, exatidão pelo método do Z-score dos nove laboratórios em duas amostras

Amostra 1			Amostra 2			4ª a 6ª etapas			
Lab.	Média	Z_{1i}	Lab.	Média	Z_{2i}	$\sum_{\alpha=1}^2 Z_{\alpha i}$	V_i	$P_i^{(t)}$	$E_i^{(t)}$
6	98	1,61	6	97	1,41	3,02	1,51	0,07	1,51
5	95	0,97	5	95	1,07	2,04	1,02	0,04	1,02
1	92	0,32	1	91	0,41	0,74	0,37	0,03	0,37
3	92	0,32	4	89	0,08	0,40	0,20	0,08	0,22
4	90	-0,11	9	89	0,08	-0,02	-0,01	0,07	0,07
7	86	-0,97	8	86	-0,41	-1,38	-0,69	0,20	0,72
9	86	-0,97	3	81	-1,24	-2,21	-1,10	0,10	1,11
8	85	-1,18	7	80	-1,41	-2,59	-1,29	0,08	1,30
2	60	-6,55	2	54	-5,70	-12,25	-6,13	0,30	6,13

Exemplo para o cálculo da precisão do lab 1:

$$P_1^{(t)} = \frac{\sqrt{(1,61 - 1,51)^2 + (1,41 - 1,51)^2}}{2} = 0,07$$

Método Campos (1998)

O banco de dados utilizado como exemplo para aplicação do método Campos encontra-se na Tabela 13B.

TABELA 13B Resultados de um teste de germinação em três laboratórios com duas amostras e quatro repetições cada

Amostra 1					Amostra 2				
Lab.	Repetição				Lab.	Repetição			
	1	2	3	4		1	2	3	4
1	88	91	93	97	1	92	89	93	96
2	97	96	99	99	2	100	99	98	100
3	73	59	66	68	3	65	65	52	65

Como descrito na seção 5.2.2 do capítulo 3, esta técnica pode ser desenvolvida em 5 etapas.

1ª etapa) Identificação de valores discrepantes (*outliers*)

São construídos intervalos interquartílicos (Tabela 14B) e verifica-se em qual deles a observação está contida, sendo atribuído um índice de discrepância d_i com valor correspondente ao número do intervalo menos uma unidade (Tabela 15B).

TABELA 14B Determinação dos limites inferiores e superiores e valores de d_i para os três laboratórios em duas amostras e quatro repetições cada

Amostra 1						Amostra 2							
B		Repetição				d_i	B		Repetição				d_i
		1	2	3	4				1	2	3	4	
1	Li	74,5	65,8	71,3	74,8	0	1	Li	69,8	68,5	61,0	71,8	0
	Ls	98,5	100	100	100			Ls	100	100	100	100	
2	Li	68,5	56,5	63,0	67,0	1	2	Li	61,0	60,0	49,5	63,0	1
	Ls	100	100	100	100			Ls	100	100	100	100	
3	Li	62,5	47,3	54,8	59,3	2	3	Li	52,3	51,5	38,0	54,3	2
	Ls	100	100	100	100			Ls	100	100	100	100	
...		
	$\hat{y}_{(1k)0,75}$	80,5	75,0	79,5	82,5			$\hat{y}_{(2k)0,75}$	78,5	77,0	72,5	80,5	
	$\hat{y}_{(1k)0,25}$	92,5	93,5	96,0	98,0			$\hat{y}_{(2k)0,25}$	96,0	94,0	95,5	98,0	
	$\hat{\delta}_{(1k)}$	12,0	18,5	16,5	15,5			$\hat{\delta}_{(2k)}$	17,5	17,0	23,0	17,5	

¹ Quando o limite superior ultrapassava a 100 foi feita a substituição por 100.

Exemplo para o cálculo dos limites da repetição 3, amostra 2 e $B = 3$:

$$\left(\hat{y}_{\{23\}0,75} - \frac{B}{2} \hat{\delta}_{(23)}; \hat{y}_{\{23\}0,25} + \frac{B}{2} \hat{\delta}_{(23)} \right) = \left(72,5 - \frac{3}{2} \cdot (15,5); 95,5 + \frac{3}{2} \cdot (15,5) \right) = (38,0; 132,5)$$

TABELA 15B Determinação de d_j (com base nos limites da Tabela 14C) para os três laboratórios em duas amostras e quatro repetições cada

Amostra 1					Amostra 2				
Lab.	Repetição				Lab.	Repetição			
	1	2	3	4		1	2	3	4
y_{11k}	88	91	93	97	y_{21k}	92	89	93	96
d_j	0	0	0	0	d_j	0	0	0	0
y_{12k}	97	96	99	99	y_{22k}	100	99	98	100
d_j	0	0	0	0	d_j	0	0	0	0
y_{13k}	73	59	66	68	y_{23k}	65	65	52	65
d_j	1	1	1	1	d_j	1	1	1	1

Neste caso não é identificado nenhum outlier, pois todos os valores de d_j são menores que 3, permanecendo todos os valores para o 2º passo.

2ª etapa) Cálculo da medida de discrepância d_2

São construídos intervalos de confiança para mediana (Tabela 16B) e verifica-se em qual deles a observação está contida, sendo atribuído um índice de discrepância d_2 com valor correspondente ao número do intervalo menos uma unidade, com exceção de $d_2 = 0$, que é determinado pelos dois primeiros intervalos (Tabela 17B).

TABELA 16B Determinação dos limites inferiores e superiores e valores de d_2 para os três laboratórios em duas amostras e quatro repetições cada

Amostra 1						Amostra 2							
B	Repetição				d_1	B	Repetição				d_1		
	1	2	3	4			1	2	3	4			
	L_{i1}	79,2	76,5	80,3	84,4		L_{i1}	78,7	76,3	74,7	82,1		
0	$L_{s1} = L_{i2}$	88,0	91,0	93,0	97,0	0	0	$L_{s1} = L_{i2}$	92,0	89,0	93,0	96,0	0
	L_{s2}	96,7	100	100	100			L_{s2}	100	100	100	100	
	L_{i1}	70,5	61,9	67,6	71,9			L_{i1}	65,5	63,7	56,5	68,3	
1	$L_{s1} = L_{i2}$	79,2	76,5	80,3	84,4	0	1	$L_{s1} = L_{i2}$	78,7	76,3	74,7	82,1	0
	L_{s2}	100	100	100	100			L_{s2}	100	100	100	100	
	L_{i1}	61,7	47,4	54,8	59,3			L_{i1}	52,2	51,0	38,2	54,4	
2	$L_{s1} = L_{i2}$	70,4	61,9	67,6	71,9	1	2	$L_{s1} = L_{i2}$	65,5	63,7	56,5	68,3	1
	L_{s2}	100	100	100	100			L_{s2}	100	100	100	100	
...	
	\tilde{y}_{1k}	88,0	91,0	93,0	97,0			\tilde{y}_{2k}	92,0	89,0	93,0	96,0	
	s_{1k}^2	147	403	309	301			s_{2k}^2	336	305	637	367	
	n_{1k}	3	3	3	3			n_{2k}	3	3	3	3	
	$s_{md(1k)}$	8,8	14,5	12,7	12,6			$s_{md(2k)}$	13,3	12,6	18,3	13,9	

TABELA 17B Determinação de d_2 (com base nos limites da Tabela 16B) para os três laboratórios em duas amostras e quatro repetições cada

Amostra 1					Amostra 2				
Lab.	Repetição				Lab.	Repetição			
	1	2	3	4		1	2	3	4
y_{11k}	88	91	93	97	y_{21k}	92	89	93	96
d_1	0	0	0	0	d_1	0	0	0	0
d_2	0	0	0	0	d_2	0	0	0	0
y_{12k}	97	96	99	99	y_{22k}	100	99	98	100
d_1	0	0	0	0	d_1	0	0	0	0
d_2	0	0	0	0	d_2	0	0	0	0
y_{13k}	73	59	66	68	y_{23k}	65	65	52	65
d_1	1	1	1	1	d_1	1	1	1	1
d_2	0	1	1	1	d_2	1	1	1	1

3ª a 5ª etapas) Classificação, cálculo da exatidão e precisão (Tabela 18B)

TABELA 18B Determinação de medidas auxiliares, exatidão, precisão e classificação pelo método de Campos para os três laboratórios em duas amostras

Lab.	T_l	n_l	$E_l^{(II)}$	\bar{n}_1	\bar{n}_2	$s_{\min(1)}^2$	$s_{\min(2)}^2$	s_{\min}^2	$s_{abs(l)}^2$	I_l	$P_l^{(II)}$	Clas.
1	0	8	100						9,7	0	100	Ouro
2	0	8	100	8	8	15,7	17,0	16,3	2,0	0	100	Ouro
3	13	8	90,3						39,0	0	100	Ouro

Para os cálculos de $s_{\min(1)}^2$ e $s_{\min(2)}^2$ são excluídas todas as determinações do laboratório 3, uma vez que a soma de d_1 com d_2 é diferente de zero.

Análise de agrupamento

Para a aplicação da técnica de análise de agrupamento, estão sendo utilizados resultados de um teste de germinação, com cinco laboratórios e duas amostras (Tabela 19B).

TABELA 19B Resultados médios de um teste de germinação com cinco laboratórios em duas amostras

Lab.	Amostra 1	Amostra 2
1	86	85
2	82	80
3	82	81
4	82	85
5	72	74

A primeira etapa consiste no cálculo da distância euclidiana entre cada par de laboratórios – grupos (Tabela 20B). Determina-se um novo grupo para a combinação que apresentar a menor distância.

TABELA 20B Resultados das distâncias euclidianas para combinações duas a duas de quatro laboratórios em duas amostras

Combinações	d_{ij}	Novo grupo
Lab1 – Lab2	6,43	
Lab1 – Lab3	5,66	
Lab1 – Lab4	4,00	
Lab1 – Lab5	17,80	
Lab2 – Lab3	1,00	Lab2 e Lab3, representado por Lab23
Lab2 – Lab4	5,00	
Lab2 – Lab5	11,66	
Lab3 – Lab4	4,00	
Lab3 – Lab5	12,21	
Lab4 – Lab5	14,87	

Exemplo do cálculo da distância entre os laboratórios 1 e 2

$$d_{12} = \sqrt{\sum_{k=1}^p (\bar{y}_{1k} - \bar{y}_{2k})^2} = \sqrt{(86 - 82)^2 + (85 - 80)^2} = 6,43$$

1ª matriz de distâncias:

	<i>Lab1</i>	<i>Lab2</i>	<i>Lab3</i>	<i>Lab4</i>	<i>Lab5</i>
<i>Lab1</i>	0,00	6,43	5,66	4,00	17,80
<i>Lab2</i>	6,43	0,00	1,00	5,00	11,66
<i>Lab3</i>	5,66	1,00	0,00	4,00	12,21
<i>Lab4</i>	4,00	5,00	4,00	0,00	14,87
<i>Lab5</i>	17,80	11,66	12,21	14,87	0,00

Processo 1: Recalcular a matriz de distâncias pelo método de ligação média, considerando o novo grupo Lab23 com os demais grupos.

Exemplo do cálculo da distância do Lab1 (grupo G_1) com o novo grupo Lab23 (grupo G_2) utilizando os resultados da 1ª matriz:

$$d(G_1, G_2) = d(\text{Lab1}, \text{Lab23}) = \frac{(d_{12} + d_{13})}{(1.2)} = \frac{6,43 + 5,66}{2} = 6,05$$

Calculam-se também as distâncias entre Lab4 e Lab23; e entre Lab5 com Lab23, atualizando-se assim a matriz de distâncias com esses novos valores e mantendo as demais distâncias apresentadas na 1ª matriz de distâncias.

2ª matriz de distâncias:

	<i>Lab1</i>	<i>Lab23</i>	<i>Lab4</i>	<i>Lab5</i>
<i>Lab1</i>	0,00	6,05	4,00	17,80
<i>Lab23</i>	6,05	0,00	4,50	11,94
<i>Lab4</i>	4,00	4,50	0,00	14,87
<i>Lab5</i>	17,80	11,94	14,87	0,00

Processo 2: Recalcular a matriz de distância, considerando o novo grupo Lab14 (menor distância apresentada na 2ª matriz) com os grupos remanescentes Lab23 e Lab5, utilizando ainda alguns resultados da 1ª matriz de distâncias:

$$d(\text{Lab14}, \text{Lab23}) = \frac{(d_{12} + d_{13} + d_{42} + d_{43})}{(2.2)} = \frac{6,43 + 5,66 + 5,00 + 4,00}{4} = 5,27$$

3ª matriz de distâncias:

$$\begin{array}{c} \text{Lab14} \quad \text{Lab23} \quad \text{Lab5} \\ \text{Lab14} \left[\begin{array}{ccc} 0,00 & 5,27 & 16,43 \\ \text{Lab23} & 5,27 & 0,00 & 11,94 \\ \text{Lab5} & 16,43 & 11,94 & 0,00 \end{array} \right] \end{array}$$

Processo 3: Recalcular a matriz de distância, considerando o novo grupo Lab2314 (menor distância apresentada na 3ª matriz) com o grupo remanescente Lab5.

$$d(\text{Lab2314}, \text{Lab5}) = \frac{(d_{25} + d_{35} + d_{15} + d_{45})}{(4.1)} = \frac{12,21 + 14,87 + 17,80 + 11,66}{4} = 14,14$$

4ª matriz de distâncias:

$$\begin{array}{c} \text{Lab2314} \quad \text{Lab5} \\ \text{Lab2314} \left[\begin{array}{cc} 0,00 & 14,14 \\ \text{Lab5} & 14,14 & 0,00 \end{array} \right] \end{array}$$

O processo é finalizado com a formação de um único grupo Lab23145, sendo construído o dendrograma considerando as menores distâncias em cada processo, neste exemplo, 1,00; 4,00; 5,27; 14,14 (Figura 1B).

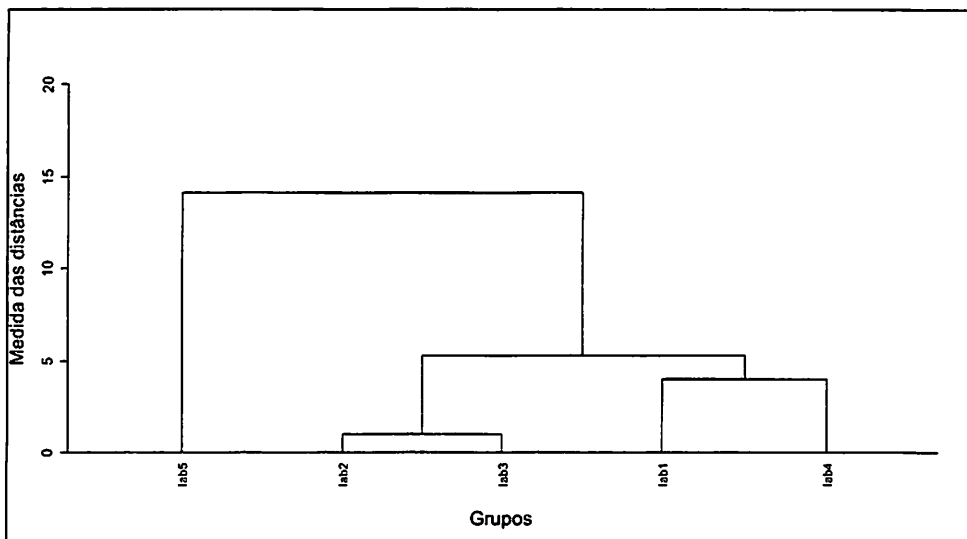


FIGURA 1B Dendrograma para o agrupamento de 5 laboratórios a partir dos dados de germinação utilizando o método da ligação média

Para verificar quanto grupos de fato devem ser considerados no final do processo, pode se utilizar o procedimento de Mojena (1977) (Tabela 21B).

TABELA 21B Determinação do ponto de corte para determinação do número de grupos, procedimento de Mojena (1977)

Estágio (m)	Grupos	α_m
1	Lab1, Lab2, Lab3, Lab4, Lab5	
2	Lab1, Lab23, Lab4, Lab5	1,00
3	Lab14, Lab23, Lab5	4,00
4	Lab2314, Lab5	5,27
5	Lab23145	14,14
Média ($\bar{\alpha}$)		6,10
Desvio padrão (s_{α})		5,65
$\bar{\alpha} + \phi s_{\alpha}$ - Ponto de corte		13,16

Neste exemplo, por este critério, seria considerado apenas um único grupo com todos os laboratórios.

Para validar o agrupamento são comparadas as distâncias efetivamente observadas entre os objetos (distâncias euclidianas) e as distâncias previstas a partir do processo de agrupamento, utilizando-se a correlação cofenética (Tabela 22B).

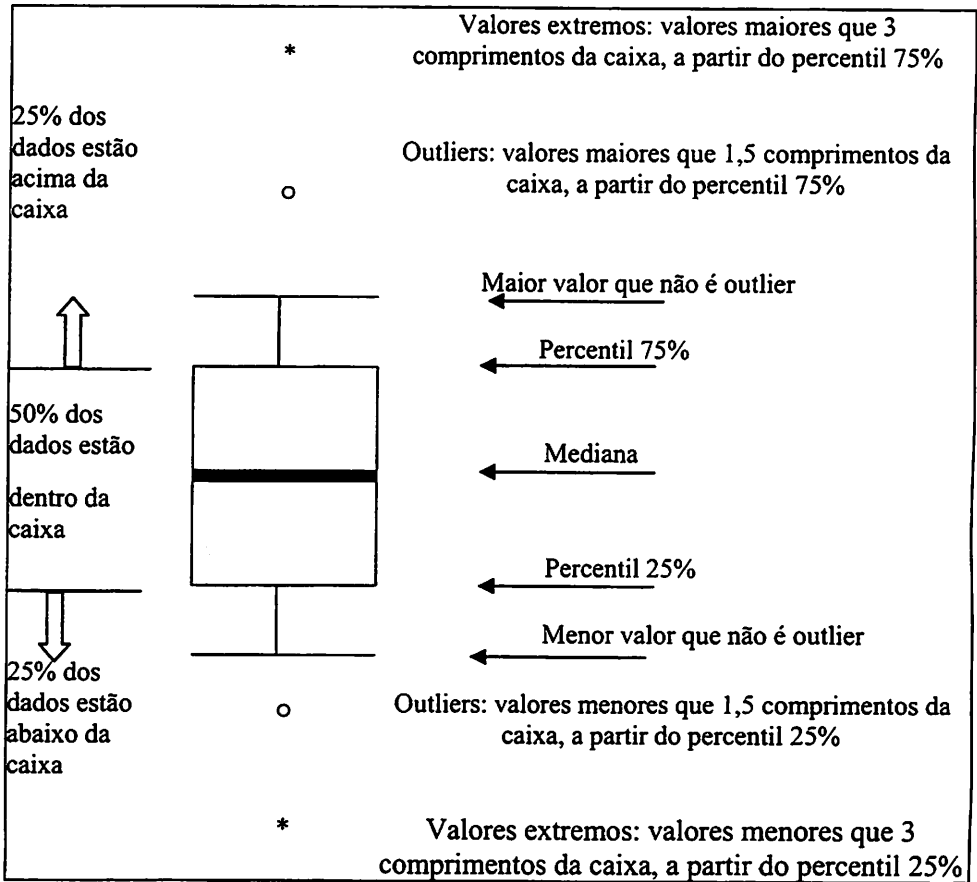
TABELA 22B Medidas das distâncias euclidianas e previstas para o cálculo da correlação cofenética

Combinações	Distância euclidiana	Distância prevista
Lab1 – Lab2	6,43	5,27
Lab1 – Lab3	5,66	5,27
Lab1 – Lab4	4,00	4,00
Lab1 – Lab5	17,80	14,14
Lab2 - Lab3	1,00	1,000
Lab2 – Lab4	5,00	5,27
Lab2 - Lab5	11,66	14,14
Lab3 - Lab4	4,00	5,24
Lab3 – Lab5	12,21	14,14
Lab4 – Lab5	14,87	14,14
Correlação cofenética	0,95	

O agrupamento pode ser considerado de boa qualidade, pois o valor da correlação cofenética é próximo de 1.

ANEXO C

Para esboçar um boxplot devem ser calculadas várias medidas descritivas (Figura 1C).



Fonte: Cazorla (2001)²²

FIGURA 1C Descrição das medidas utilizadas na construção de um boxplot

²² CAZORLA, I. C. *Educação estatística: análise exploratória*. Ilhéus, BA, 2001. Cap. III, 41 p.

Disponível em: <<http://www.socio-estatistica.com.br/Edestatistica/apostila.htm>>. Acesso em: 10 abr. 2009.

Calculadas as medidas descritivas como: mediana, o percentil 25%, o percentil 75% e a amplitude interquartílica é possível esboçar o boxplot, sendo que o comprimento da caixa é determinado pela amplitude interquartílica = $Q_3 - Q_1$

ANEXO D

Teste de Box's M (Ferreira, 2008²³)

Um dos testes para verificar a homogeneidade de matrizes de covariâncias de $I \geq 2$ populações normais multivariadas é conhecido como teste de Box's M.

Esse teste foi proposto por Box (1949) como uma modificação ao teste de Bartlett (1947).

O interesse é testar a hipótese de homogeneidade sobre as covariâncias ($p \times p$) das I populações dada por:

$$H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_I$$

Considerando amostras aleatórias de I populações normais p -variadas $N_p(\mu_i, \Sigma_i)$ de tamanho n_i , cada uma $i = (1, 2, \dots, I)$. A amostra da i -ésima população é dada por $Y_{i1}, Y_{i2}, \dots, Y_{in_i}$ em que $Y_{ij} \in \mathbb{R}^p$ corresponde a j -ésima da unidade da i -ésima população. Considerando ainda que $n = \sum_{i=1}^I n_i$, a estatística do teste de Box's M é dada por:

$$\chi^2 = - \left[\left(\sum_{i=1}^I \frac{1}{v_i} - \frac{1}{n-I} \right) \left(\frac{2p^2 + 3p - 1}{6(p+1)(I-1)} \right) \right] \times \left[\sum_{i=1}^I v_i \ln |S_i| - (n-I) \ln |S_c| \right],$$

possui distribuição assintótica qui-quadrado com $f = (I-1)p(p+2)/2$ graus de liberdade, sendo que os estimadores (S_i e S_c) não-viesados das matrizes de covariâncias populacionais são definidos por:

²³ FERREIRA, D.F. **Estatística multivariada**. Lavras: UFLA/DEX, 2008. 661 p.

$$S_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_i)(Y_{ij} - \hat{Y}_i)^T \text{ e } S_c = \frac{1}{n-1} (n_i - 1) S_i$$

ANEXO E

Neste anexo são apresentados os tipos de bancos de dados que foram montados para análise no R e variaram conforme a técnica estatística empregada; um exemplo com alguns relatórios gerados e os comandos para a execução da Função `testeH`.

A justificativa para apresentação de tais exemplos decorre de dois fatos. Primeiro, as rotinas desenvolvidas no R estão na forma de linguagem de programação e, conseqüentemente, apresentam uma lógica para a montagem do banco de dados.

Segundo, exemplificar para pesquisadores a montagem de alguns tipos de banco de dados para viabilizar a utilização dos mesmos no *software* R.

Teste H

TABELA 1E Exemplo da montagem de um banco de dados para a realização do teste H

Lote ¹	Subamostra	Germinação
LT1	S1	98
LT1	S2	92
LT1	S3	96
LT1	S4	95
LT1	S5	94
LT2	S1	85
...
LT5	S5	74

¹ Não utilizar, por exemplo, a simbologia L T 1, ou seja, não devem ser dados espaço entre letras, senão o *software* R fará a leitura considerando como sendo 3 colunas separadas

Método de Hampel

Foi montado um banco de dados inicial somente para a detecção de *outliers* (seção 5.4.1 do capítulo 2).

TABELA 2E Exemplo da montagem de um banco de dados para a realização do método de Hampel

Laboratório	Lote	Repetição	Germinação
LB1	LT 1	R1	98
LB1	LT 1	R2	92
LB1	LT 1	R3	96
LB2	LT 1	R1	95
LB2	LT 1	R2	94
LB2	LT 1	R3	85
...
LB8	LT 5	R3	75

Análise Discriminante

TABELA 3E Exemplo da montagem de um banco de dados para a Análise Discriminante

Grupo	R1	R2	R3
G1	90	95	98
G1	89	94	92
G1	92	91	96
G1	94	93	95
G1	91	96	94
G1	89	97	85
...
G3	75	70	72

Método do Z-score e Método Campos (1998)

TABELA 4E Exemplo da montagem de um banco de dados para o método do Z-score e método de Campos (1998)

Laboratório	Lote	Repetição	Germinação	Rodada ¹
LB1	LT1	R1	98	1
LB1	LT1	R2	92	1
LB1	LT1	R3	96	1
LB2	LT1	R1	95	1
LB2	LT1	R2	94	1
LB2	LT1	R3	85	1
...
LB12	LT5	R3	79	6

¹ No caso desse estudo foi realizada apenas uma rodada.

Análise de Agrupamento

TABELA 5E Exemplo da montagem de um banco de dados para a Análise de Agrupamento

¹	Amostra1	Amostra2	Amostra3
LB1	90	95	98
LB2	89	94	92
LB3	92	91	96
LB4	94	93	95
LB5	91	96	94
LB6	89	97	85
...
LB12	75	70	72

¹ Na utilização das funções já implementadas no R para esta técnica, não se nomeia a coluna do fator

² São utilizadas as médias de cada laboratório de acordo com a amostra.

Análise Generalizada de Procrustes (GPA)

Para a Análise Generalizada de Procrustes o banco de dados inicial é formado por um vetor único. No próprio R são estruturadas oito (repetições) matrizes (*arrays*) tendo nas linhas os laboratórios e nas colunas as amostras, de acordo com os seguintes comandos:

```
gvar=read.table("gvar.txt",h=T)
attach(gvar)
testvar<-array(gv,
  dim=c(12,3,8),
  dimnames=list(Lab=c("1","2","3","4","5","6","7","8","9","10","11","12"),
    Amostra=c("1","2","3"),
    Repetição=c("1","2","3","4","5","6","7","8")))
```

Outliers detectados pelo Método de Hampel - Areia

Lab.	Lote	Rep.	Germ.
1	3	2	81
6	3	3	85
8	3	3	78

Outliers detectados pelo Box plot - Areia

Lab.	Lote	Rep.	Germ.
1	3	2	81
8	3	3	78
6	5	3	74
8	5	3	71

Deteção de outliers nas variâncias para cada lote pelo Teste de Levene - Areia

Lote	valorp	Condição ¹
1	0.0915	Var. Homogêneas - continuar a análise
2	0.2698	Var. Homogêneas - continuar a análise
3	0.0099	Var. Heterogêneas - repetir o teste
4	0.0067	Var. Heterogêneas - repetir o teste
5	0.0009	Var. Heterogêneas - repetir o teste

¹Nível de significância = 0.01

Comparação com Limite crítico de Reprodutibilidade - Areia

Lote	Amostra.1	Amostra.2	Amostra.3	limite crítico	sR
1	14	7	13	16.93892	4.203881
2	14	16	21	23.99690	5.955523
3	6	4	7	8.49251	2.308827
4	23	23	20	34.91761	8.665812
5	12	9	18	18.20006	4.516870

FIGURA 1E Exemplo de alguns relatórios gerados no *software* R para o Capítulo 2

Resultado do Z score¹ - Exatidão

Lab.	Viés.rod	Precisão.rod	Exat.rod	Clas.Exat.rod ²
1	0.6	1.2	1.4	Aceitável
2	1.1	1.4	1.7	Crítico
3	2.8	5.7	6.4	Não aceitável
4	0.5	0.1	0.5	Aceitável
5	0.6	1.2	1.3	Aceitável
6	2.1	4.1	4.6	Não aceitável
7	1.8	1.7	2.4	Não aceitável
8	1.4	0.5	1.5	Aceitável
9	0.9	0.7	1.1	Aceitável
10	2.1	4.4	4.8	Não aceitável
11	1.8	1.8	2.6	Não aceitável
12	2.6	5.5	6.1	Não aceitável

¹Nível de confiança = 0.95

² Aceitável - Exatidão menor que 1,5
 Crítico - Exatidão maior ou igual a 1,5 e menor que 2,0
 Não aceitável - Exatidão maior do que 2,0

Resultado do Método do Profert

Lab.	Exatidão	Precisão	Selo ¹
1	93.29	88.89	Ouro
2	96.3	100	Ouro
3	88.89	88.89	Ouro
4	91.9	77.78	Ouro
5	97.92	100	Ouro
6	91.9	88.89	Ouro
7	93.29	88.89	Ouro
8	95.6	88.89	Ouro
9	93.06	100	Ouro
10	94.68	88.89	Ouro
11	81.25	66.67	Prata
12	86.57	55.56	Prata

¹ Selo Ouro - Exatidão e Precisão maiores que 67%
 Selo Prata - Exatidão maior que 67% e Precisão menor ou igual a 67%
 Selo Bronze - Exatidão menor que 67% e Precisão maior que 67%
 Nenhum selo - Exatidão e Precisão menores ou iguais a 67%

FIGURA 2E Exemplo de alguns relatórios gerados no *software* R para o Capítulo 3

```

#-----
# Nome      : FunçãoTesteH
# Autora    : Verônica Yumi Kataoka
# Data      : 18/08/2008
# Versão    : v1
# Objetivo   : Realizar o teste H de acordo com as normas ISTA., 2006
#-----
FunçãoTesteH =function (a,N,n,fc,prob)
{
mg=tapply(a[,3],a[,1],mean)
s2=a[,3]^2
stot=tapply(s2,a[,1],sum)
stotq=(tapply(a[,3],a[,1],sum))^2
varobs=(N*stot-stotq)/(N*(N-1))
varaceit=(mg*(100-mg)*fc)/n
H=round(((varobs/varaceit)-fc),2)
qui=qchisq((1-prob),N-1)
Htabela=(fc*qui)/(N-1)
Htabelaf=round((Htabela -fc),2)
lt=as.factor(a[,1])
nL=nlevels(lt)
j=matrix(0,nL,1)
for(i in 1:nL)
{
j[i]<- ifelse(H[i] < Htabelaf, "Homogêneo", "Heterogêneo")
}
d<-data.frame(Lote=levels(a[,1]),Hcalculado=H,Htabelado'=Htabelaf,Condição = j)
cat("_____", "\n")
cat("     Teste H     ", "\n")
cat("_____", "\n")
print(d,row.names=FALSE)
cat("_____", "\n")
cat("'Nível de significância =',(prob),'", "\n")
cat("_____", "\n")
Sys.time( )
cat("_____", "\n")
cat("#####", "\n")
}

```

FIGURA 3E Programa para execução da FunçãoTesteH no *software* R

ANEXO F

Neste anexo são apresentados resultados complementares do Capítulo 3.

TABELA 1F Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 1

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classif. Exatidão
1	1	A	5	0,8	1,5	1,8	Crítico
	2	A	5	0,9	1,2	1,5	Crítico
	3	BMP	0	2,9	5,9	6,6	Não aceitável
	4	A	5	0,7	0,1	0,7	Aceitável
	5	A	5	0,5	1,1	1,3	Aceitável
	6	B	4	1,7	3,3	3,7	Não aceitável
	7	C	3	1,8	1,4	2,3	Não aceitável
	8	B	4	1,6	0,4	1,6	Crítico
	9	B	4	1,2	1,3	1,7	Crítico
	10	B	4	1,4	3,0	3,3	Não aceitável

TABELA 2F Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 2

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classif. Exatidão
1	1	A	5	0,3	0,5	0,5	Aceitável
	2	B	4	1,4	1,2	1,8	Crítico
	3	A	5	0,8	0,1	0,8	Aceitável
	4	B	4	1,3	2,6	2,9	Não aceitável
	5	C	3	2,1	3,6	4,1	Não aceitável
	6	A	5	0,7	0,7	1,0	Aceitável
	7	C	3	2,2	4,2	4,7	Não aceitável
	8	BMP	0	5,0	10,6	11,7	Não aceitável
	9	C	3	2,3	4,3	4,9	Não aceitável
	10	BMP	0	4,1	9,1	10,0	Não aceitável

TABELA 3F Resultados do método do Z-score para o conjunto de dados do grupo 1, de acordo com a simulação 3

Grupo	Lab.	Classificação	Nota	Viés	Precisão	Exatidão	Classif. Exatidão
1	1	A	5	0,9	1,9	2,1	Não aceitável
	2	BMP	0	2,9	0,7	2,9	Não aceitável
	3	A	5	0,5	1,0	1,1	Aceitável
	4	A	5	0,5	1,1	1,2	Aceitável
	5	A	5	1,0	0,4	1,1	Aceitável
	6	A	5	0,6	0,5	0,8	Aceitável
	7	A	5	1,1	0,3	1,2	Aceitável
	8	A	5	0,7	1,5	1,6	Crítico
	9	A	5	0,4	0,6	0,8	Aceitável
	10	A	5	1,0	1,9	2,2	Não aceitável