

ISSN 1677-5473

*Empresa Brasileira de Pesquisa Agropecuária
Departamento de Pesquisa e Desenvolvimento
Ministério da Agricultura, Pecuária e Abastecimento*

Texto para Discussão 40

Questões críticas em validação de métodos analíticos

***Elisabeth Borges Gonçalves
Ana Paula Guedes Alves
Paula Alves Martins***

***Embrapa Informação Tecnológica
Brasília, DF
2011***

Exemplares desta publicação
podem ser solicitados na:

Empresa Brasileira de Pesquisa Agropecuária (Embrapa)

Departamento de Pesquisa e Desenvolvimento (DPD)

Parque Estação Biológica (PqEB)

Av. W3 Norte (final)

CEP 70770-901 Brasília, DF

Fone: (61) 3448-4451

Fax: (61) 3448-4887

textoparadiscussao@embrapa.br

Editor da série

Ivan Sergio Freire de Sousa

Coeditores

Eliane Gonçalves Gomes

Vicente Galileu Ferreira Guedes

Conselho editorial

Alberto Roseiro Cavalcanti

Assunta Helena Sicoli

Carlos Augusto Mattos Santana

Chang das Estrelas Wilches

Eliane Gonçalves Gomes

Geraldo B. Martha Jr.

Ivan Sergio Freire de Sousa

Lucilene Maria de Andrade

Marita Feres Cardillo

Otávio Valentim Balsadi

Paule Jeanne Mendes

Renato Cruz Silva

Roberto de Camargo Penteado Filho

Colégio de editores associados

Ademar Ribeiro Romeiro

Altair Toledo Machado

Antonio César Ortega

Antonio Duarte Guedes Neto

Arlson Favareto

Carlos Eduardo de Freitas Vian

Charles C. Mueller

Dalva Maria da Mota

Egídio Lessinger

Geraldo da Silva e Souza

Geraldo Stachetti Rodrigues

João Carlos Costa Gomes

John Wilkinson

José de Souza Silva

José Graziano da Silva

José Manuel Cabral de Sousa Dias

José Norberto Muniz

Josefa Salete Barbosa Cavalcanti

Léa Velho

Levon Yeghanyan

Manoel Moacir Costa Macêdo

Marcel Bursztyn

Maria Amália Gusmão Martins

Maria Lucia Maciel

Mauro Del Grossi

Oriowaldo Queda

Pedro Carlos Gama da Silva

Rui Albuquerque

Sergio Salles-Filho

Sergio Schneider

Suzana P. M. Mueller

Tarcizio Rego Quirino

Vera L. Divan Baldani

Zander Navarro

Supervisão editorial

Wesley José da Rocha

Normalização bibliográfica

Márcia Maria Pereira de Souza

Projeto gráfico

Tenisson Waldow de Souza

Revisão de texto

Maria Cristina Ramos Jubé

Editoração eletrônica

Júlio César da Silva Delfino

1ª edição

1ª impressão (2011): 600 exemplares

As opiniões emitidas nesta publicação são de exclusiva e de inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), vinculada ao Ministério da Agricultura, Pecuária e Abastecimento.

Todos os direitos reservados

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação (CIP)

Embrapa Informação Tecnológica

Questões críticas em validação de métodos analíticos / Elisabeth Borges Gonçalves, Ana Paula Guedes Alves, Paula Alves Martins. – Brasília, DF : Embrapa Informação Tecnológica, 2011.

69 p. : il. ; 21 cm x 15 cm - (Texto para Discussão / Embrapa. Departamento de Pesquisa e Desenvolvimento, ISSN 1677-5473 ; 40).

1. Metrologia. 2. Análise química. 3. Controle de qualidade. 4. Normalização. I. Gonçalves, Elisabeth Borges. II. Alves, Ana Paula Guedes. III. Martins, Paula Alves. IV. Embrapa. Departamento de Pesquisa e Desenvolvimento.

CDD 389.1

© Embrapa 2011

Apresentação

***Texto para Discussão** é uma publicação técnico-científica da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), de divulgação de resultados de pesquisas, cuja relevância os fazem merecedores de um espaço de reflexão e debate.*

Editada pelo Departamento de Pesquisa e Desenvolvimento (DPD), a série é dirigida a técnicos, pesquisadores, dirigentes, formuladores de políticas públicas, acadêmicos e público em geral que tenham, como área de atuação ou de interesse, as temáticas da ciência e tecnologia, da inovação, do agronegócio e do desenvolvimento rural sustentável.

*De caráter monográfico, **Texto para Discussão** publica e circula ideias e reflexões sobre assuntos contemporâneos de relevo para a sociedade brasileira, os quais são abordados tanto por analistas e pesquisadores dos quadros da Embrapa quanto por especialistas de instituições públicas e privadas que atuam com pesquisa, desenvolvimento e transferência de tecnologias.*

*Além da forma impressa, os leitores podem acessar todos os números da série **Texto para Discussão** no seguinte endereço: www.embrapa.br/embrapa/publicacoes/tecnico/folderTextoDiscussao.*

O Editor



Sumário

Resumo	9
Abstract.....	10
Introdução	11
Experimentação e testes de hipóteses	13
Dificuldades em treinamentos.....	19
Nível de validação necessário e responsabilidades cabíveis	20
Dificuldades comuns.....	23
Incerteza.....	24
Valor verdadeiro.....	25
Realização dos experimentos de validação.....	28
Número de repetições	33
Faixa de trabalho.....	35
Valores aberrantes	37
Experimentos de precisão (estudos colaborativos)	42
Amostragem.....	48

Incerteza de amostragem	50
Verificação de conformidade	58
Crítica da validação	61
Conclusões	62
Referências	64



Questões críticas em validação de métodos analíticos^{1,2,3}

Elisabeth Borges Gonçalves⁴

Ana Paula Guedes Alves⁵

Paula Alves Martins⁶

¹ Original recebido em 11/8/2010 e aprovado em 4/1/2011.

² Projeto de pesquisa em andamento.

³ Os autores agradecem o apoio financeiro recebido do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

⁴ Bacharel em Estatística, Mestre em Agronomia, Doutora em Tecnologia de Alimentos, Pesquisadora A, Embrapa Agroindústria de Alimentos, Av. Américas, 29.501, Guaratiba, CEP 23020-470, Rio de Janeiro, RJ, goncaleb@ctaa.embrapa.br

⁵ Bacharel em Estatística, Bolsista DTI - CNPq, Embrapa Agroindústria de Alimentos, Av. Américas, 29.501, Guaratiba, CEP 23020-470, Rio de Janeiro, RJ, ana.uerj@click21.com.br

⁶ Bacharel em Estatística, Especialista em Engenharia da Produção, Bolsista DTI - CNPq, Embrapa Agroindústria de Alimentos, Av. Américas, 29.501, Guaratiba, CEP 23020-470, Rio de Janeiro, RJ, pam_martins7@yahoo.com.br



Questões críticas em validação de métodos analíticos

Resumo

O mercado internacional vem adotando uma série de exigências em normas de modo a regular a qualidade de produtos. Normas internacionais estão sendo elaboradas ou revisadas e, muitas vezes, adotados limites de parâmetros de qualidade mais e mais restritivos. Posto que o exercício da tomada de decisão sobre esses parâmetros envolve qualidade de resultados laboratoriais, a acreditação de laboratórios de métodos analíticos vem sendo alvo de atenção cada vez mais intensa. Como resultado, a validação de métodos passou a ser alvo de exigências mais estreitas e documentadas, porém o processo de validação de métodos analíticos ainda carece de discussão em diversos de seus aspectos. Assim, apresenta-se neste texto uma abordagem de questões consideradas críticas a respeito deste tema, quer seja por resultarem em riscos para o mercado nacional e internacional, em interpretações equivocadas ou em outros tipos de problemas. São discutidos dificuldades no exercício e na leitura do tema, níveis necessários de validação, necessidade de testes de parâmetros de validação em métodos validados ou oficiais, tipos de experimentos de validação intralaboratorial e seu planejamento, experimentos de precisão (estudos colaborativos), limite de decisão ($CC\alpha$), capacidade de detecção ($CC\beta$), parte objeto de conclusão, realização de experimentos fracionados e classificação e extração de valores aberrantes. Admite-se que o conceito de amostra aleatória ainda não foi adequadamente absorvido na área; como consequência, a assim chamada “incerteza de amostragem” mostrou-se equivocada. A equivalência entre testes de hipóteses e intervalos de confiança (incerteza expandida), raramente encontrada na área, foi apresentada, bem como a interpretação correta, desses intervalos, terminando por mencionar que há uma posição brasileira sobre sua interpretação e uso na norma da NBR 11155: 1988. Finalmente, demonstrou-se que a validação de métodos analíticos é uma tarefa de conteúdo teórico avançado, e que exige uma análise crítica adequada dos resultados estatísticos.

Termos para indexação: gestão da qualidade, acreditação de laboratórios, metrologia química, análise crítica de validação, planejamento de validação.

Critical points in analytical methods validation

Abstract

The international market has been adopting a series of demands and procedures to regulate the quality of products. International standards have been developed or reviewed and, often, limits have been used for quality parameters with increasing restrictions. Since the decision making is based on these parameters involving quality of laboratory results, laboratory accreditation of analytical methods has been matter of enormous attention. As a result, the attention paid to validation of methods became more and more demanded, but the validation of analytical methods still needs discussion in several aspects. Therefore, this text shows an approach to issues considered as critical in this subject, either resulting in risks to the domestic and international markets, in erroneous interpretations or in other kind of problems. Difficulties in exercising and reading the topic are discussed, as well as the appropriate levels of validation of methods, the validation parameters needed to verify official or validated methods, the types of in house validation experiments and their planning, the precision experiments (collaborative studies), the part of the population which is the target of the conclusion, the decision limit ($CC\alpha$), the detection capability ($CC\beta$), the problems which occur when conducting fractionated experiments and the classification and extraction of outliers. It is recognized that the concept of random sampling has not been adequately understood in the area and, as a consequence, the so-called uncertainty sampling proved to be misinterpreted. The equivalence between hypothesis testing and confidence intervals (expanded uncertainty), not frequently known in the area, is presented, as well as their correct interpretation, and was pointed that there is a Brazilian position on their interpretation and use in NBR 11155: 1988. Finally, was demonstrated that the validation of analytical methods is an advanced theoretical content task, and that it demands a proper statistical critical analysis.

Index terms: quality management, laboratory accreditation, metrology in chemistry, planning and critical analysis of validation.

Introdução



rente à formação de blocos econômicos, à defesa de interesses financeiros e à exigência de manutenção de hábitos saudáveis, o mercado nacional e internacional vem impondo padrões estabelecidos em produtos de diversas naturezas, incluindo-se produtos agrícolas, alimentos ou não, mais especificamente quanto à qualidade, abrangendo segurança (toxicidade, defeitos, alergias, etc.), funcionalidade (ação bactericida, atributos que propiciem incremento da saúde corporal, etc.), e outras características.

Limites de contaminantes como agrotóxicos, resíduos de pesticidas, resíduos de diversos produtos químicos e outros estão sendo estabelecidos especialmente no comércio internacional, e normas internacionais estão sendo elaboradas ou revisadas para realização de seu controle. Ou seja, os importadores e o mercado em geral estão banindo aquilo que consideram ou que é comprovadamente prejudicial à saúde.

Como resultado disso, aumenta a exigência na medição de parâmetros que caracterizam os padrões de qualidade e segurança de produtos em geral. Laboratórios passam a ter sua atuação questionada, e, em função disto e das imposições de mercados internacionais, vêm sendo impelidos à acreditação, como aquela proposta na Norma Mercosur et al. (2006), ou em similares. Como consequência, esses laboratórios têm seus resultados analíticos e sua competência internacionalmente reconhecidos.

Claramente, um laboratório deve ter sua rotina de verificação de métodos sempre que os implanta, o que se chama de validação de métodos analíticos. Naturalmente que essa prática sempre foi seguida pelos analistas de qualquer laboratório; porém, a situação hoje é de internacionalização. Assim, a validação de métodos passou a ser requerida além daquilo exigido pelo exercício das funções analíticas. Hoje deve haver documentação rigorosa, e é necessário seguir os requisitos publicados nos textos internacionais de validação de métodos analíticos.

Os conceitos nesta área são métricos, a área é de estatística.

Normas técnicas, documentos orientativos, vários tipos de guias e diversas outras publicações sobre validação de métodos analíticos têm sido revisadas e divulgadas, mas desencontros ou equívocos entre definições, usos indicados e outros detalhes que trazem dificuldades de interpretação e emprego têm sido notados, sem que se tenha discussão suficiente e clareza de diversos conceitos, como será demonstrado no decorrer deste texto.

Pretendeu-se, portanto, trazer ao leitor um pouco mais de detalhe e de discussão de questões relevantes em validação, produtos de uma abordagem crítica da validação de métodos analíticos, quer seja na realização da validação intralaboratorial de métodos analíticos como na interlaboratorial.

Deve-se mencionar que o termo “amostra” e seus derivados serão grafados entre aspas. De maneira similar, sempre que se usar qualquer terminologia

cujo conceito se coloque em dúvida serão usadas aspas.

“Amostra” denotaria casos em que a amostra não foi obtida conforme técnicas de amostragem, não seguiu as regras estabelecidas pela teoria estatística. Por exemplo, um item em observação ou medição, como uma fatia de queijo ou uma peça de um equipamento, não é uma amostra, pois amostra estocástica é um conjunto de itens com $n > 2$ unidades amostrais obtidas de acordo com certas leis. “Amostra” de uma unidade não existe; então, casos desse tipo serão referidos entre aspas.

Experimentação e testes de hipóteses

Generalidades



Em laboratório, ao empregar qualquer método analítico, necessita demonstrar que o opera de maneira adequada; assim uma validação de métodos busca meios e critérios objetivos para demonstrar que os métodos de ensaio usados conduzem a resultados confiáveis e adequados à qualidade pretendida, o que pode ser encontrado no Eurachem (1998), National Association of Testing Authorities (1998), Inmetro (2010), National Food Administration (2009) e outros.

A via utilizada para esse objetivo é a experimentação, ou seja, a demonstração, na prática, de que

parâmetros de validade são cumpridos pelo método analítico. Os experimentos para validação devem adotar procedimentos ou delineamentos estocásticos adequados.

Abre-se, aí, a grande área de Experimentação, abrangendo desde o teste de enunciados teóricos, quer seja por meio de modelos de delineamento de experimentos, de regressão, de análise de dados categóricos, etc., até a obtenção de conclusões para tomadas de decisão. Estamos, assim, frente à necessidade de comprovar enunciados teóricos na prática; portanto, defrontamo-nos com o Método Científico, também conhecido como método indutivo-dedutivo ou hipotético-dedutivo; a via estatística.

Enunciados teóricos são, então, traduzidos em hipóteses, que são afirmações que se pretende verificar, como: uma linha de produção está fornecendo itens dentro dos padrões de qualidade estipulados; está ocorrendo um surto de aftosa nesta cidade; a Aids chegou a seu máximo potencial de virulência em tal localidade; e outras.

Assim, o método é submetido a determinadas condições, hoje consensuais, que poderiam invalidá-lo e se verifica se seu comportamento se altera ou não. As hipóteses testadas, portanto, são sempre do tipo H_0 : o método permanece inalterado (em termos de média ou precisão), em determinada condição x H_1 : o método se altera.

A Tabela 1 traz um panorama geral envolvendo a realidade frente aos resultados dos testes de hipóteses baseados em amostras.

Tabela 1. Esquema de decisão nos testes de hipóteses.

Decisão a tomar	Realidade	
	H_0 verdadeira	H_0 falsa
Rejeitar H_0	Erro tipo I (nível de significância, α)	Decisão correta (poder de teste) $1 - \beta$
Não rejeitar H_0	Decisão correta	Erro tipo II β

As hipóteses H_0 e H_1 são ditas hipótese nula (ou de nulidade) e hipótese alternativa, respectivamente. A questão básica em validação de métodos analíticos é ter, ou não, evidências suficientes para se rejeitar H_0 , ou seja, para se rejeitar a hipótese de que o método não se altera nas condições que lhe foram impostas.

Desse modo, uma validação de método não consegue provar que um método é válido. Somente consegue não ter evidências suficientes para rejeitar hipóteses de invalidade e, portanto, concluir que o método atende a certos requisitos em determinadas circunstâncias e com determinada (im)precisão.

Pelo exposto, o que se adentra, em validação, é o mundo da estatística experimental, um mundo vasto e complexo, muito além das meras estatísticas básicas.

Cenários encontrados em testes de hipóteses

Como forma de tornar mais claras as situações encontradas em testes de hipóteses, considere-se que a probabilidade de se admitir que um surto de febre aftosa em gado ocorre em qualquer região seja $p = 2\%$.

Se uma amostra de gado bovino obtida no Mato Grosso do Sul com 500 (quinhentas) cabeças apresentou 8 (oito) ocorrências de aftosa, pode-se considerar que está ocorrendo um surto de aftosa neste estado?

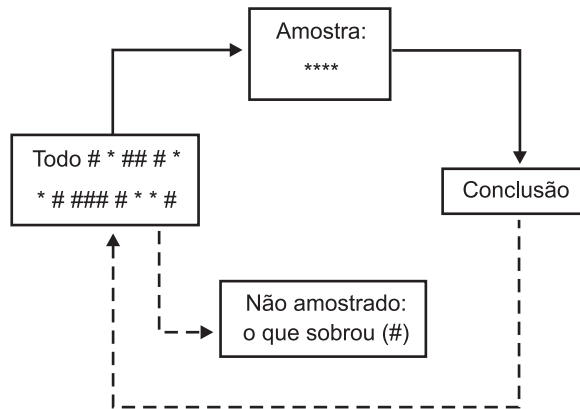
Seria considerada como a hipótese nula aquela em que a probabilidade de surto na região está próxima de 2%, ou seja, $H_0: p = 0,02$, ou que, ao contrário, este valor já foi atingido e pode ter sido superado, logo $H_1: p > 0,02$. Rejeitando-se H_0 em favor de H_1 será concluído que está ocorrendo um surto no estado. Mas, se a realidade for que $p < 0,02$, será admitido que há um surto quando, na realidade, não há e se comete o erro tipo I, também chamado nível de significância, com probabilidade expressa por α . O resultado disso seria criar clima de sobressalto e provocar gastos desnecessários.

Se a hipótese H_0 , porém, não for rejeitada em favor de H_1 , a conclusão será que não há surto; quando na realidade houver, será cometido o erro tipo II, cuja probabilidade se expressa por β . Providências preventivas e de contenção não serão tomadas, propiciando contaminação descontrolada e graves desdobramentos facilmente imagináveis. Nesse caso a probabilidade de não se cometer tal tipo de erro é notada por meio de $(1 - \beta)$ e chamada de poder de teste.

Teste de hipóteses e ciência

O campo que se descortina ao se verificarem hipóteses ou visar à obtenção de estimativas e à generalização/validade de conclusões é o da Inferência.

De acordo com a Figura 1, se a amostra empregada for tomada de acordo com as devidas bases esto-cásticas, então as conclusões nela baseadas têm valor generalizável também ao que não foi observado no conjunto todo. Isso é inferir, concluir inclusive sobre o que não se examinou, e nisso repousa o famoso risco amostral.



* ou #: unidades que formam todo o conjunto a ser amostrado.

Figura 1. Amostragem e validade de conclusão (a última abrange o todo * ou #, até mesmo aquilo que não foi amostrado #).

É o conhecimento sólido dos testes de hipóteses, na forma aqui apresentada, que traz ao pesquisador, cliente, vendedor, consumidor, etc. os riscos que envolvem o chamado poder de teste, riscos que geram curvas como as curvas características de operação em amostragem para qualidade, análises do tipo sequencial, e outras.

Além disso, esse conhecimento fornece visão apropriada de intervalos de confiança. Na ausência dessa visão, muitas vezes esses intervalos são

equivocadamente usados deixando de se fazer valer sua equivalência com testes de hipótese bilaterais.

Ainda, atualmente as publicações têm simplificado o verdadeiro teor da validação de métodos, tornando validação um campo perigoso para o cliente dos laboratórios e deixando de atribuir a devida responsabilidade a quem compete assinar os seus resultados estocásticos e as correspondentes conclusões, o estatístico.

A nomenclatura estocástica em qualidade é bastante confusa e pouco rigorosa. O leitor, porém, poderá encontrar diversos conceitos da área na própria internet usando órgãos normalizadores, como os glossários de Fuentes-Arderiu (2010), da Organization for Economic Cooperation and Development (2010), da European Statistics (2010), de National Institute of Standards and Technology (2010) ou outras fontes.

Ademais, é necessário acrescentar que cada técnica, cada teste de hipótese estatístico tem suas restrições quanto a sua propriedade, e estas devem ser seguidas. Assim, aplicar um teste de hipóteses requer também verificar se suas exigências são válidas ou não. Para não correr esse risco, deve-se sempre dar preferência ao uso de testes estatísticos menos exigentes, como os não paramétricos, e às investigações informais, como gráficos estocásticos e avaliações descritivas ou exploratórias; assim, deve-se sempre questionar a real necessidade de se realizar cada teste de hipóteses.

É possível que os problemas apresentados neste item e no decorrer de todo este texto tenham sido detectados porque a formação do estatístico no Brasil

é diferente do que acontece na esmagadora maioria dos outros países, que não têm a mesma formação de base tão ampla e de espectro tão abrangente, e isso é público. No exterior, ser estatístico é ter esse título por meio de uma pós-graduação, não é ter a formação superior na área.

Sem dúvida, isso constitui uma grande vantagem para nosso país, que detém maior poder de argumentação nos seus estatísticos para refutar exigências equivocadas de outros mercados, além do que ocorre no mundo da informação. Não é de se espantar, portanto, que muito venha sendo publicado sobre essa profissão.

Nas atividades de qualidade a estatística desempenha papel nevrálgico. O profissional que atua na atividade de validação, encaixada na qualidade para laboratórios, encontra a obrigatoriedade de detalhar todos os delineamentos experimentais, todos os métodos estatísticos em estudos acreditados em Boas Práticas de Laboratórios (BPL), Organization for Economic Cooperation and Development (1998), bem como de documentar os procedimentos de validação, o mesmo ocorrendo em acreditação de laboratórios do tipo Norma Mercosur et al. (2006).

Dificuldades em treinamentos



A principal dificuldade nos treinamentos em validação de métodos analíticos tem sido a extensão do conteúdo

técnico necessário para prover uma noção sólida da matéria, pois exige desde os cálculos diferenciais e integrais, cálculos de probabilidades, geometria linear, álgebra linear, e inferência, que são básicos em estatística, até métodos de análise complexos ministrados na pós-graduação.

Assim, validação não pode nunca ser suficientemente aprendida por um leigo em estatística em um curso de curta duração, nem em um semestre ou dois de matérias como quimiometria ou outras. Ainda, o que muitas escolas e profissionais chamam de “estatística básica” é, na realidade, um apanhado de introduções comumente destituídas da devida base estocástica para seu correto exercício.

Validação abrange a tomada de decisão e a imagem de laboratórios, além da formação adequada e a competência (poder legal para prática do ato) que essa confere, experiência técnica, rigor necessário para bem exercê-la, capacidade de análise crítica dos resultados, poder de argumentação teórica e facilidade de transmiti-la por escrito até para enfrentar disputas judiciais.

Nível de validação necessário e responsabilidades cabíveis



nível de validação necessário no caso de métodos normalizados ou oficiais, ou seja, obtidos em normas nacionais ou internacionais, ou publicados pelos órgãos

competentes, ainda gera dúvidas. Do ponto de vista metodológico, não se valida método normalizado; ele já foi validado ou é considerado válido há tempos. O Inmetro (2010), portanto, esclarece que, nesse caso, é necessário somente demonstrar que se pode operá-lo de maneira adequada, nas condições dos laboratórios.

O que esse documento deixa claro é que é preciso demonstrar que os métodos de ensaio usados conduzem a resultados confiáveis e adequados à qualidade pretendida no laboratório em que são utilizados. Assim, não é obrigatório o uso de todos os parâmetros de validação constantes em nenhum caso, mas somente daqueles que sejam necessários para comprovar a qualidade dos resultados. Por exemplo, não há necessidade de se validar um método para todas as matrizes (alimentos ou outro objeto de medição por método analítico) abrangidas por seu escopo, ou as regras que envolvem a amostragem estocástica e o próprio estabelecimento do escopo do método não fariam sentido.

Evidentemente, há métodos em que se usam técnicas diferentes dependendo de certos fatores, ou que mudam em algum detalhe, dependendo da matriz. Isso deve estar descrito no escopo do método – ou seja, naquilo em que ele se aplica –, o que é necessário é delimitá-lo adequadamente, e tomar o devido cuidado na sua redação.

Validação de métodos analíticos compreende pelo menos duas áreas de formação, assim é atividade de atuação e responsabilidade compartilhada. Ambos, estatístico e profissional responsável pelo método

analítico, devem discutir juntos todos os parâmetros de validação a serem pesquisados, e cada um deve emitir o seu parecer em documentos separados apresentando, em cada caso, as justificativas técnicas quanto a não pesquisar parâmetros julgados desnecessários.

Compete ao profissional responsável pelo método analítico a documentação que se refere a ele, o estabelecimento de locais apropriados para armazenamento e uso de reagentes, padrões, equipamentos, processos laboratoriais de validação, elaboração de análises críticas de ensaios, verificação de uso de rotinas ou documentos anteriores, conferição dos resultados de medições, inclusão de referências, escolha do método, execução e descrição de modificações, escolha de analistas, etc.

Complementarmente, cabe ao estatístico o planejamento dos experimentos necessários; o acompanhamento e a verificação do cumprimento dos delineamentos propostos; a análise crítica inicial das medições realizadas, do fornecimento e da organização da codificação do material; a garantia de que as medições serão cegas e de que a mesma porção teste não seja retestada; a análise das distribuições; a escolha, a aplicação e a análise dos testes estatísticos, dos gráficos estocásticos; a elaboração/escolha dos programas a serem usados; a validação dos cálculos estatísticos; a crítica dos resultados de validação; a emissão de documento abrangendo todos os métodos estatísticos usados; a inclusão de referências pertinentes; a conclusão sobre a validade, restrita ou não, do método analítico; etc.

Dificuldades comuns



dificuldade de manutenção de matrizes (material em que se realiza o método analítico) na área de alimentos é algo muito conhecido pelos profissionais da área, porque alimentos se deterioram. No entanto, nem sempre compreendido por clientes, auditores, etc.

A incapacidade de reproduzir a realidade da medição do analito em uma matriz, sua reatividade, suas características, etc. por meio do emprego de fortificação (adição de quantidade conhecida da substância teste ou microrganismo, adição do analito ou microrganismo, contaminação com a substância teste ou microrganismo) é conhecida em alimentos, porém subestimada por profissionais de outras áreas da qualidade.

As diferenças encontradas nesses casos não são comumente compreendidas, e muitos chegam a imaginar que é possível que alimentos colhidos em épocas diferentes reajam da mesma forma à fortificação e medição, tanto em termos de exatidão quanto em precisão. Isso, porém, muitas vezes não ocorre. Pequenas diferenças, por exemplo, de umidade em alimentos, podem alterar em muito a reatividade no emprego de métodos analíticos e resultar em formas completamente inesperadas de precisão (incerteza, variância) ou desvios nos resultados esperados.

Há que se adequar o método às circunstâncias em que é utilizado. Assim, é preciso incluir também na validação, além daqueles níveis do analito destinados a cobrir as aplicações do método no laboratório, as vizinhanças de limites impostos pela legislação, quando existentes.

A relação entre um limite legal e um limite de quantificação ou detecção é imediata quando da definição de um limite de detecção. Naturalmente, nada pode ser corretamente quantificado se estiver abaixo do limite de detecção.

Raramente se encontra, no entanto, uma validação envolvendo os limites de detecção, quantificação e limites legais em conjunto. Além disso, há declarações de limites de detecção ou de quantificação que não são propriamente compreendidas, ou adequadamente comunicadas. É reciso lembrar que os dois últimos, muitas vezes, são resultados de cálculos e não resultados de limites experimentalmente testados e validados.

Esses limites são pontos no eixo do analito. A probabilidade de um ponto, porém, é nula, ou seja, a probabilidade daquele valor de limite de detecção ou de quantificação ocorrer e corresponder à realidade é zero. Logo, obter um determinado limite teoricamente e comprovar, ao mesmo tempo, que ele é verdadeiro, raramente ocorre. Então, esses limites não podem meramente ser admitidos como válidos sem os testes laboratoriais e estatísticos necessários. Os resultados, certamente, não serão pontos, mas intervalo(s) que o(s) contenha(m).

Incerteza



Segundo a definição do Inmetro et al. (2000), publicação conhecida como Vocabulário Internacional de

Metrologia (VIM), incerteza é um parâmetro $u^2(x)$ associado ao resultado de uma medição, que caracteriza a dispersão de valores que poderiam ser razoavelmente atribuídos ao mensurando.

Conforme a literatura, esse parâmetro pode ser um desvio padrão, ou um múltiplo dele, ou a metade de um intervalo correspondente a um nível de confiança estabelecido e compreende, em geral, muitos componentes. Incerteza padrão $u(x)$ é um desvio padrão. Note-se que, então, incerteza $u^2(x)$ é variância (σ^2) de uma medição x .

Quando o laboratório calcular a incerteza expandida do seu resultado (U_{lab}), o valor verdadeiro (X_v) deve estar dentro do intervalo ($X_{lab} \pm U_{lab}$), Inmetro (2010).

Pergunta-se, então, quem é X_v ? Valor verdadeiro de quê? Do mensurando na porção amostral medida? Do mensurando no pacote disponível para medição? Do mensurando no lote? Do mensurando no processo produtivo?

Valor verdadeiro



Pelo exposto, outro problema pouco comentado é a definição do que seria o valor verdadeiro a que se referem as publicações. O Inmetro (2010) estabelece o valor verdadeiro como X_v , valor aceito como verdadeiro em relação a um valor certificado. Segundo

o National Institute of Standards and Technology (Nist), em Taylor e Kuyatt (1994), haveria um valor verdadeiro do mensurando. Essa obra, entretanto, não diz com relação a quê, e acrescenta que ainda existe dúvida sobre ele. No guia Eurachem e Cooperation on International Traceability in Analytical Chemistry, (2000), o valor verdadeiro é mencionado como um valor consistente com a definição de uma quantidade particular, porém essa mesma publicação admite que há muitos valores consistentes com essa definição.

Ainda, o limite de decisão $CC\alpha$ é definido como aquele ou acima do qual pode-se concluir que uma "amostra" é não conforme com probabilidade de erro α de acordo com a Commission Decision (2002). Nesse mesmo documento define-se capacidade de detecção $CC\alpha$ como o menor teor do analito que pode ser detectado, identificado ou quantificado em uma "amostra" com probabilidade de erro α . No caso de analitos sem mínimo estipulado, esse limite é a menor concentração na qual se pode detectar "amostras" verdadeiramente contaminadas com uma margem de certeza $1 - \alpha$.

O que se deve notar é que as "amostras" a que essa publicação se refere não são bem determinadas. Não é mencionado se a "amostra" é uma porção amostral procedente de um lote industrializado ou uma porção de uma determinada produção agrícola.

Esse valor, verdadeiro, muitas vezes tem sido associado a um valor que pode ser usado em uma expressão de incerteza expandida a ser aplicada para um lote, por exemplo, mas na maior parte das vezes não há base suficiente para isso.

O valor verdadeiro em muitas publicações é válido somente para a porção sob medição ou nem mesmo isso. Não pode ser usado em verificações de conformidade de produtos como agrícolas ou agroindustriais, a menos que seja uma evidência científica.

Há outros fatores, porém, operando sobre a medição, como a diferença de incertezas quando se alteram as matrizes, os processos produtivos agrícolas ou os industriais. Logo, falta poder inferencial nos dados laboratoriais para que seus resultados possam extrapolar os limites do laboratório para um processo, a menos que haja projeto especialmente desenvolvido com esse fim.

Infelizmente, não tem ficado claro nas publicações que muitas vezes a incerteza de medição, ou a incerteza em um *pool* (agregado, conjunto indivisível) é só o que os laboratórios podem conseguir, e que com ela raramente será obtido um intervalo de confiança válido para estimar o nível de um analito em determinado lote ou processo produtivo, especialmente se o material for regularmente heterogêneo quanto ao analito.

Aquilo que se refere a um processo de medição que é exclusivamente laboratorial não pode ser estendido além dele. Trata-se de um caso em que as medições são provenientes de um conjunto com unidades de medição que não estão na Figura 1. Em outras palavras, com base em itens do tipo ($\square_1, \square_2, \square_3, \square_4, \dots$), ou seja, itens que não são pertencentes à população estudada, pretende-se obter conclusões (inferir) para a população que é do tipo ($\#_1, \#_2, *_1, *_2, *_3, \#_3, *_4, \dots$). O correto, porém, seria usar ($*_1, *_2, *_3, *_4, \dots$), que é a amostra proveniente da população acima.

Na realidade, busca-se uma incerteza que possa ser utilizada em verificação de conformidade de lotes e processos. Isso é tema de item separado no decorrer deste texto.

De modo similar, a mesma situação é verificada com $CC\alpha$ e $CC\beta$. Ambos são provenientes de processos de medição intralaboratoriais, e nada além disso. Não têm poder inferencial em um sistema de produção industrial ou agrícola que não foi investigado.

Realização dos experimentos de validação

Bases da experimentação



Infelizmente a literatura em química tem dado pouca ênfase ao conceito de independência. Ela tem sido reportada por meio de uma medição não estar relacionada com a anterior.

É o esquema de dependência/independência em medições que determina o tipo de modelo que pode ser usado e o tipo de análise estatística apropriada. Ele é uma das estruturas que coordenam todo o arcabouço experimento – análise estatística – estimação (de médias, variâncias, desvios-padrão) – conclusão – tomada de decisão, etc.

Em condições de dependência, medidas básicas de qualquer sistema estocástico como média e variância não deverão ser estimadas pela média amostral e

variância amostral comuns. Nesses casos, por exemplo, a média amostral é tendenciosa (viciada, desviada) para estimar a média real de qualquer processo, e a variância da média difere daquela conhecida.

Como exemplos da importância do conceito em medições em ensaios analíticos, verificando-se as normas International Organization for Standardization (1994a, 1994b, 1994c, 1994d; 1998a; 1998b) pode-se notar que as definições de condições de repetitividade e reprodutibilidade trazem esse conceito implícito, e que o mesmo ocorre na definição de precisão. Sem independência, repetitividade e reprodutibilidade não podem ser definidas e, se calculadas nas formas tradicionais, disponíveis nas normas, resultarão em valores irrealistas que não refletirão a precisão de métodos, sem que disso se tenha qualquer indicação prática ou visível.

Se uma “subamostra” provém de determinado recipiente, toda e qualquer outra “subamostra” deverá resultar em medições próximas à primeira. Sem dúvida, ao se realizar “subamostragem”, não se garante que serão realizadas medições independentes, e não se poderá afirmar, de acordo com teoria estocástica, que repetições propriamente ditas tenham sido obtidas.

Da mesma forma, Inmetro (2010) menciona replicatas independentes e usa ensaios independentes na definição de precisão, o que implica também em exigência de ensaios independentes na repetitividade e na reprodutibilidade.

De modo similar, Taylor e Kuyatt (1994), Eurachem, Cooperation on International Traceability in Analytical Chemistry (2000) e Commission Decision (2002) incluem independência na precisão, repetitividade e reprodutibilidade. A National Food Administration (2009), no entanto, não a menciona. Definições e forma de obtenção de ensaios independentes raramente são encontradas em qualidade para laboratórios.

Uma definição de independência de ensaios consta do Inmetro (2010), conforme resultados obtidos de maneira não influenciada por qualquer resultado prévio na mesma ou em matriz similar testada. Esta, porém, não é uma definição suficiente.

Independência em resultados/ensaios implica em experimentos conhecidos como cegos, como mencionado em Pocklington (1985), Association Office of Analytical Chemists (1989) e Pocklington (1990). Esses experimentos são da maior importância em ciência e em validação. Podem ser obtidos resultados experimentais inválidos caso a independência não seja mantida, sem que se perceba, por muito tempo que os resultados obtidos tinham problemas.

Experimentos cegos são obtidos por meio de sistemas de geração que seguem distribuições probabilísticas e trazem observações organizadas e codificadas. São muito usados em medicina e evitam, nos laboratórios, interferências de analistas, equipamentos, ambiente, etc. São a base de qualquer validação, e a base de qualquer pesquisa científica. É necessário que os laboratórios sigam as ordens de ensaios estabelecidas no planejamento e que cada

porção a ser analisada não exceda em muito a quantidade necessária para a análise de modo que ela não possa ser repetida na mesma porção. Margens de até 20% a mais de material podem compor as porções a serem medidas e reposição deve ser fornecida quando da perda da porção, ao invés de se fornecer porções maiores e propiciar subamostragem.

Em muitas vezes será reportado exatamente o resultado obtido, especialmente em sistemas do tipo cromatográficos, em que o resultado é calculado pelo sistema. Em muitas outras o analista não vai alterar o resultado que obteve, mas poderá ser descartado algum resultado que venha a ser julgado incompatível com algum outro anterior. Felizmente, isso não é muito comum da parte do analista, mas é suficientemente realizado por parte do laboratório, que crê estar frente a um valor aberrante, quando isso não foi estocasticamente confirmado.

Sem dúvida, a capacidade de análise e a disponibilidade dos laboratórios devem ser consideradas no planejamento. Há, todavia, problemas com preços e padrões, por exemplo, em curvas de calibração em HPLC e similares, em que independência das concentrações não é sempre possível e diluições de uma mesma solução são preparadas, carregando-se os erros da solução-mãe para as diluições. Nesses casos é necessário fazê-las tão independentes quanto possível.

Sem independência nenhuma validação é válida, porém, muitas vezes serão aceitos certos “graus de dependência”.

Os experimentos devem ser realizados de acordo com o teste estatístico a ser usado em cada fase da validação e para cada parâmetro e deve ser tomado em consideração o tipo e o número de respostas a serem obtidas na medição. Por exemplo, há ensaios cujos resultados são dependentes porque foram obtidos na mesma porção teste e outros que são independentes por causa do sistema reativo empregado pelo método.

Tentar obter o máximo de informação possível o tempo todo é um fator comum de complicação. Sem dúvida, é mais econômico medir muitas coisas de uma vez que poucas coisas de cada vez, porém isso conhecidamente costuma gerar conjuntos de dados que terminam por confundir o analista, a validação, as conclusões e sua compreensão. Não é simples entender relações entre muitos analitos ao mesmo tempo, mas, sim, entre alguns que sejam os principais analitos de interesse. Dessa forma, o que seria economia pode confundir e até impossibilitar conclusão.

Fracionamento de delineamentos experimentais

Fracionamento de delineamentos experimentais é usado por muitos autores que visam economia de recursos e de trabalho experimental. O problema destes delineamentos, no entanto, é que, com o fracionamento, as interações são excluídas dos experimentos e não podem ser testadas. Cornell (1984) refere-se a esse problema em alimentos; Comissão Decision (2002) recomenda certo cuidado

com esses experimentos, e outros autores os apresentam e discutem. A maioria, contudo, dos documentos nacionais e internacionais em validação de métodos analíticos não o mencionam. Os resultados dos testes de hipóteses obtidos com essa exclusão podem ser enganosos. Efeitos significativos, então, deixam de ser detectados.

Na realidade, muito do que se testa em validação com emprego de experimentos fracionados são condições ambientais de realização de ensaios, por exemplo, as típicas variações de tempo de reação e de temperatura, tão comuns em química, notadamente em alimentos. Nesses casos interações existem e são significativas. São exatamente os casos de efeitos sinérgicos reais.

Caso essas interações sejam ignoradas, quando, na realidade elas existem, isso não será rapidamente detectado, e o laboratório correrá riscos na qualidade de seus resultados analíticos. O que se recomenda, portanto, é não fracionar, executando experimentos em partes, até a obtenção de experimentos completos.

Número de repetições



Quantas repetições eu uso? é a pergunta considerada folclórica e mais conhecida no mundo científico. Não há uma resposta honesta a essa pergunta a não ser: não sei.

Não há um número cabalístico ou consenso envolvendo essa questão. Os sistemas de amostragem dependem de dados como insumos para se poder fazer um cálculo adequado, ou que se julgue apropriado na fase inicial de uma validação. Ou seja, validação é um problema científico como qualquer outro, e essa informação não surge do vazio, mas é construída baseada em informações também experimentais.

Deve-se, então, questionar que tipo de informações está à disposição para que se possa calcular a quantidade de medições necessárias? Deve-se também verificar se tais informações são suficientes e apropriadas; logo, a validação pode necessitar de complementação de medições em qualquer um de seus parâmetros, a qualquer tempo.

As publicações que trazem os números de repetições a serem usados para, por exemplo, verificação de linearidade e calibração, como as publicações de Inmetro (2010), Taylor e Kuyatt (1994), Eurachem, Cooperation on International Traceability in Analytical Chemistry, Eurachem e Citac Working Group (2000), Comissão Decision (2002), National Food Administration (2009), Eurachem (1998) e National Association of Testing Authorities (1998), não são consensuais, e os textos que as emitem não fornecem as bases científicas segundo as quais foram obtidos. Na verdade não passam de números que alguns consensos julgam razoáveis, sem base teórica e não representam qualidades necessariamente suficientes.

Caso se opte pelo mínimo de seis repetições/matriz ou concentração requerido por essas publicações, o

laboratório corre risco de não ter testes de hipóteses suficientemente poderosos pelo baixo número de repetições, situação bastante arriscada nas regiões críticas das curvas de calibração, faixas de trabalho, limite de detecção, limite de quantificação, etc.

O real número de repetições necessário para cada parâmetro de validação é algo a ser estudado. Assim, quando se necessita de maior sensibilidade nesses parâmetros não há outra recomendação a não ser um dimensionamento específico para as diferenças que se pretende detectar, ou empregar melhoramento de experimentos.

Faixa de trabalho



Estabelecer a faixa de trabalho pode parecer simples, mas nem sempre é. Devem-se incluir limites de legislações nas faixas para serem testados; bem como suas vizinhanças. Ainda, a real detecção também deve ser considerada quando se estuda essa faixa.

As retas ajustadas nessa faixa não devem ser forçadas a passar no ponto zero, nem valores zero do analito não ensaiados devem ser incluídos nos cálculos. O intercepto deve ser obtido e é preciso que sua nulidade seja testada do ponto de vista estocástico. É necessário, ainda, realizar os testes de hipóteses de não existência de regressão e de falta de ajuste. A última é lembrada por Thompson et al. (2002) e deve ser incluída no modelo de regressão. Finalmente,

há que se apresentar a análise de regressão completa, incluindo-se um diagnóstico realizado por profissional experiente em reconhecimento de padrões de comportamento estocástico, bem como a correspondente análise crítica.

Vários outros aspectos, contudo, devem ser examinados. Na realidade, a maior parte das respostas dos métodos não descreve uma reta no decurso dos analitos; ou seja, não há reta, o que há é uma região em que o método tem comportamento que pode ser aproximado por uma reta. O problema é descobrir onde fica essa região, isto é, seus limites.

Muito comumente o que se tem é uma sigmoide com uma assíntota horizontal na parte superior do eixo x. A primeira parte é de crescimento rápido, aproximadamente exponencial, dita log fase, e se poderia dizer que uma segunda parte seria aproximadamente linear. A velocidade com que a curva passa da primeira para a segunda fase, todavia, muitas vezes é tão grande que o que se tem é quase que um ponto a partir do qual o método passa a detectar ou quantificar quase linearmente. Um ponto, no entanto, tem probabilidade nula; alcançá-lo, na prática, aproxima-se do impossível.

É necessário acrescentar que uma regressão somente é válida no seu intervalo de observação; logo, extrapolações não são permitidas, como pode ser encontrado na literatura clássica da área ou mesmo em Association Office of Analytical Chemists (1989).

Finalmente, é usual ampliar a região de trabalho ao longo do uso do método como resultado de

investigação adicional. Assim, o trabalho de validação não tem um início ou fim exatamente demarcados, como já exposto anteriormente.

Valores aberrantes



detecção de valores aberrantes (*outliers*) em validação de métodos analíticos permeia tanto a validação intralaboratorial quanto os experimentos de precisão (estudos colaborativos) e os ensaios de proficiência. O descarte desses valores era realizado há aproximadamente 200 anos, porém, com o uso de estimadores robustos e com novas abordagens do assunto, o descarte vem sendo menos frequentemente praticado.

Em todos os experimentos envolvidos em validação a preocupação com esses valores é justificada. O principal objetivo experimental é obter estimativas de médias e variâncias – incertezas –, e correspondentes desvios padrão – incertezas padrão – de modo a verificar, sem influência de erros grosseiros, o comportamento de métodos analíticos e o desempenho de laboratórios de análise.

Os testes estocásticos mais populares para valores aberrantes em química são os de Grubbs, Dixon e Cochran, como em Horwitz (1988, 1995) e outros.

O teste de Grubbs é aplicável a dados com distribuição normal ou log-normal, como medições

de fenômenos naturais, como parâmetros de fertilidade de solos, teores de diversos analitos em alimentos, diâmetros de árvores, etc., e tem sido recomendado por várias normas e instituições internacionais. É teste para um valor aberrante, podendo ser aplicado ao maior ou menor valor do mensurando. Foi ampliado por Rosner (1983), que estendeu o método de Grubbs para detectar k valores aberrantes em uma amostra.

$$\text{Grubbs} = \frac{|\text{valor estranho} - \text{média}|}{\text{desvio padrão}}$$

O teste de Dixon é usado para 3 a 25-30 medições com distribuição normal; baseia-se em distâncias entre extremos e detecta um valor aberrante.

$$Q_{\text{Dixon}} = \frac{|\text{valor aberrante} - \text{valor mais próximo após ordenação}|}{\text{maior valor} - \text{menor valor após ordenação}}$$

O teste de Cochran é utilizado para valores aberrantes em casos de medições em duplicatas e é expresso como se segue.

$$\text{Cochran} = \frac{(\text{maior diferença entre duplicatas})^2}{\Sigma(\text{diferenças entre duplicatas})^2}$$

Menos conhecidos são os testes não paramétricos de Rosner e de Walsh. As estatísticas de testes referentes a cada um deles são:

$$R_{i+1} = \frac{|x^{(i)} - \bar{x}^{(i)}|}{s^{(i)}}$$

em que $x^{(i)}$ é estatística de ordem obtida da distribuição disponível, $\bar{x}^{(i)}$ é a média, $s^{(i)}$ é o desvio padrão, ambos calculados com exclusão do valor aberrante

pesquisado $x^{(i)}$. Para número de medições ≥ 60 pode-se definir a estatística de teste de Walsh, em Lohninger (2010), conforme:

$$a = \frac{1 + b \sqrt{\frac{c - b^2}{c - 1}}}{c - b^2 - 1}$$

com

$$c = IFIX(\sqrt{2n})$$

em que $IFIX$ é o maior inteiro do argumento na expressão, $b^2 = 1/\alpha$, α é o nível de significância, r o número de valores aberrantes na parte superior ou inferior da distribuição, $k = r + c$ para:

$$X_r - (1 + a) X_{r+1} + aX_k < 0, X_{n+1-r} - (1 + a) X_{n-r} + aX_{n+1-k} > 0$$

Se ambas as igualdades são verdadeiras então ambos, menor e maior valor, são verdadeiramente aberrantes.

A Association Office of Analytical Chemists (1994) recomenda que se usem dados válidos de um mínimo de 8 laboratórios, mesmo número recomendado por Pocklington (1990), bem como que se estudem 5 matrizes, admitindo um máximo de 2/9 valores/matriz como aberrantes a serem descartados, sem levantar que motivos os originaram. Pocklington (1990) considera que não deve haver mais do que um conjunto (de matrizes diferentes, níveis de analito, etc.) em cada 5 que apresente mais que 20% de valores classificados como aberrantes.

Sem dúvida, esses valores podem ser generalizados a diversos parâmetros de validação.

A Association Office of Analytical Chemists (1994) instrui determinar a probabilidade de valores aberrantes no conjunto de resultados considerados normais aplicando os testes de Cochran ou Grubbs. Recomenda que se os valores aberrantes foram removidos após o teste de Grubbs, verifique-se o conjunto restante por meio do teste de Cochran.

Embora a área de química tenha dado muita importância a esses valores, tem deixado de lado os conceitos e probabilidades envolvidos nesse assunto. Valores aberrantes não são rigorosamente definidos e têm sido, historicamente, tema de uma enorme variedade de publicações. A noção mais comum de um valor desses seria um resultado que difere exageradamente dos demais. Questões comuns nesse caso são: foi um erro de anotação/digitação? Algo foi excluído ou incluído equivocadamente nos cálculos? Há diversidade química ou biológica no caso? Se não houver explicação, também é possível que o valor seja aleatório. O problema, portanto, passa a ser descobrir qual dessas, ou outras, hipóteses é verdadeira.

Nas áreas estocásticas, a detecção de valores aberrantes é comumente realizada com certa massa de dados. Do ponto de vista probabilístico, esse tipo de valores tem probabilidade muito baixa, muitas vezes menor que 1%. Desse modo, como classificar como aberrante um valor um pouco desviado em um conjunto de 10 medições? Ou de 20? Nesses casos as probabilidades associadas são de 10% e 5%, muito altas para se poder classificar tais valores como

aberrantes. Assim, extrair tais valores é um problema amostral maior do que imagina o analista comum.

A extração reiterada de valores aberrantes é algo que merece ser comentado. Uma vez que se detecte com uma determinada técnica ou teste estatístico um ou vários valores aberrantes em uma só vez, é possível realizar descarte. Não se deve, porém, reaplicar as verificações de valores aberrantes e reextraí-los.

Retornando ao aspecto probabilístico, se um valor aberrante tem probabilidade inferior a 1%, por exemplo, dois valores aberrantes em sequência teriam probabilidade muitíssimo menor que 1%, por isso mesmo ambos não poderiam ter ocorrido por acaso; logo, o segundo valor não seria aberrante. Reaplicar, portanto, verificações e realizar extrações sucessivas constituem prática sem fundamentação probabilística válida.

Além disso, testes estatísticos para valores aberrantes com um grupo limitado de dados em experimentos de precisão podem, comparativamente, demonstrar grandes discrepâncias, Hamaker (1986).

É prática comum aceitar a suposição de normalidade e independência mútua, logo verificações informais de normalidade começam a ser empregadas como ferramentas de exame de valores aberrantes. É preciso, então, que verificações desse tipo, ou seja, livres da carga de exigências dos testes de hipóteses formais, venham a ser mais intensamente estudados na área e relatados.

Experimentos de precisão (estudos colaborativos)



Muito se fala sobre experimentos de precisão (estudos colaborativos), alguns até julgam que um método não está suficientemente verificado se experimentos desse tipo não foram realizados. Realmente, deles é possível extrair muita informação, mas é preciso mais.

Eles se destinam a estudar métodos desde seu procedimento escrito, verificar seu comportamento em diversas matrizes, estudar seus resultados em diferentes ambientes, propiciar uma crítica de medições detalhada e obter precisões. Uma vez que se pretende estudar um método, são usados laboratórios em que se tem confiança. Então, o que se pretende é uma amostra, que possa ser aceita, de bons laboratórios para o experimento.

O principal nesses experimentos é, segundo as normas da International Organization for Standardization (1994a, 1994b, 1994c, 1994d; 1998a, 1998b), estimar a precisão de métodos. Isso é feito usando o modelo mais comum em delineamento de experimentos: o modelo a um critério de classificação. Nesse modelo, o critério de classificação conhecido como tratamentos em diversas áreas nada mais é que laboratórios. As medições realizadas pelos laboratórios têm suas variações (incertezas) totais decompostas em variação entre laboratórios, chamada de variação da reprodutibilidade, e variação dentro de laboratórios, dita variação da repetitividade.

Da estimação dos componentes de variância desse modelo resultam os componentes mencionados, repetitividade e reprodutibilidade, nomenclatura que comumente causa estranheza em profissionais de estatística, pois ambas se referem ao conteúdo teórico de variância, imprecisão.

O modelo que é usado, ou seja, com laboratórios como único critério de classificação, é do tipo:

$$y_{ij} = \mu + \ell_i + \varepsilon_{ij}$$

em que y_{ij} é o resultado da ij -ésima medição, μ é a média geral, ℓ_i o efeito (desvio, vício) do laboratório i , $i = 1, 2, \dots, n$, e ε_{ij} o erro aleatório, com distribuição $N(0, \sigma^2)$, $j = 1, 2, \dots, J$, ou seja, distribuição normal de média zero e variância (incerteza) σ^2 . Um esquema do experimento mais simples, com esse modelo, teria duas repetições em cada laboratório e poderia ser representado como na Figura 2.

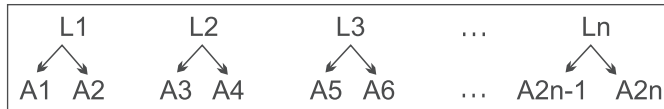


Figura 2. Esquema de realização do experimento mais simples para obtenção de estimativas de precisão em estudos colaborativos.

Para obtenção dos componentes de variância calcula-se:

$$E(SQT) = E\left(\sum_{ij} y_{ij}^2 - C\right) = E\left(\sum_{ij} y_{ij}^2\right) - E(C)$$

logo:

$$E\left(\sum_{ij} y_{ij}^2\right) = \sum_{ij} (\mu^2 + \ell_i^2 + \varepsilon_{ij}^2)$$

mas,

$$E(C) = IJ\mu^2 + J\sigma_\ell^2 + \sigma^2, \text{ portanto}$$

$$E(SQT) = (IJ\mu^2 + IJ\sigma_\ell^2 + IJ\sigma^2) - (IJ\mu^2 + J\sigma_\ell^2 + \sigma^2) = J(I-1)\sigma_\ell^2 + (IJ-1)\sigma^2$$

Ainda,

$$E(SQLab) = E\left(\frac{1}{J} \sum_i (\sum_j y_{ij})^2 - C\right) = \frac{1}{J} (IJ^2\mu^2 + IJ^2\sigma_\ell^2 + IJ\sigma^2) - (IJ\mu^2 + J\sigma_\ell^2 + \sigma^2)$$

mas,

$$E\left(\sum_i (\sum_j y_{ij})^2\right) = \sum_i E\left(\sum_j \mu + \ell_i + \varepsilon_{ij}\right)^2 = IJ^2\mu^2 + IJ^2\sigma_\ell^2 + IJ\sigma^2$$

logo,

$$\begin{aligned} E(SQLab) &= \frac{1}{J} (IJ^2\mu^2 + IJ^2\sigma_\ell^2 + IJ\sigma^2) - (IJ\mu^2 + J\sigma_\ell^2 + \sigma^2) \\ &= J(I-1)\sigma_\ell^2 + (I-1)\sigma^2 = (I-1)(J\sigma_\ell^2 + \sigma^2) \end{aligned}$$

então,

$$E(SQRes) = \sigma^2(IJ-1-I+1) = I(J-1)\sigma^2$$

assim,

$$E(QMRes) = E\left(\frac{SQRes}{I(J-1)}\right) = \sigma^2 \text{ e } E(QMLab) = \sigma^2 + J\sigma_\ell^2$$

Frente aos valores esperados obtidos montou-se a Tabela 2.

Tabela 2. Esquema de análise de variância e valores esperados.

Causas de variação	GL	SQ	QM	E(QM)
Entre laboratórios	I - 1	SQLab	QMLab	$\sigma^2 + J\sigma_\ell^2$
Dentro laboratórios	I(J-1)	SQRes	QMRes	σ^2
Total	IJ - 1	SQT		

GL: graus de liberdade; SQ: soma de quadrados; QM: quadrado médio; Lab: laboratórios; Res: resíduo; σ^2 : repetitividade; σ_ℓ^2 : variância entre laboratórios; E(QM): esperança do quadrado médio correspondente.

A partir dos componentes de variância define-se a reprodutibilidade conforme abaixo.

$$\sigma_R^2 = \sigma^2 + \sigma_\ell^2$$

Repetitividade e reprodutibilidade fisicamente, portanto, não existem, são componentes de variância, matéria de delineamento de experimentos, de um modelo comum, cuja nomenclatura originariamente causou problemas na American Society for Testing Materials (1987), e que raramente deixará de causar estranheza aos estatísticos.

Comumente se diz que métodos normalizados são robustos. Quem se refere à reprodutibilidade, porém, está se referindo exatamente a uma média de desvios de laboratórios com relação a um valor consensual, está se referindo a uma média de pequenas *faltas* de robustez geradas pelas condições experimentais, equipamentos e analistas de diversos laboratórios. De modo similar, repetitividade refere-se a uma média de pequenos erros intralaboratoriais.

Uma vez que repetitividade e reprodutibilidade são médias de valores de precisão obtidos em uma

amostra de laboratórios, não podem ser aplicadas a um único laboratório. A repetitividade de um laboratório não é definida nessas normas e, sim, a repetitividade de um método obtida em uma amostra de vários laboratórios. Para manter fidelidade às definições, portanto, repetitividade de um laboratório não existe, embora um laboratório possa operar em condições de repetitividade. Assim é que a International Organization for Standardization (2004) define σ_w^2 como variância (imprecisão) dentro (*within*) de um laboratório. Essa distinção, no entanto, não é mantida em outros documentos.

Algo que se deve dizer sobre repetitividade e reprodutibilidade é que elas de nada servem se não houver algo com que compará-las. Ou seja, dizer que a repetitividade de um método é, por exemplo, 0,0005, e a reprodutibilidade é 0,0009, não significa absolutamente nada. Não se sabe se esses valores indicam se o método é bom ou não, a menos que se tenha a repetitividade e a reprodutibilidade de um concorrente seu. Então, repetitividade e reprodutibilidade, sozinhas, não dizem absolutamente nada.

Por outro lado, os resultados do modelo que gera essas imprecisões têm sido pouco explorados. Por exemplo, estudos colaborativos raramente trazem se a reprodutibilidade foi ou não significativa. Aqui repousaria a maior importância desse tipo de experimento; se o teste para a causa de variação correspondente à reprodutibilidade, ou seja, à variação interlaboratórios, foi ou não significativo, o método é, ou não, robusto.

Frequentemente a variação da reprodutibilidade é significativa sem que isso tenha sido verificado, o que ocorre em métodos de qualquer tipo. Assim, muitos métodos normalizados não são exatamente robustos, mas têm certo grau de robustez.

Ademais, é possível aproveitar o experimento de precisão e incluir outros parâmetros a serem testados e ter resultados mais amplos, como faixa de trabalho; ou ainda, obter materiais de referência.

Finalmente, apresenta-se na Tabela 3 resultados de estudo colaborativo (experimento de precisão), realizado com o exemplo de Youden e Steiner (1975) citado por Hamaker (1986) no manual da AOAC.

Esses resultados mostram reprodutibilidade significativa no nível 2 e não nos demais níveis.

Tabela 3. Resultados obtidos na modelagem e variâncias (S^2) da repetitividade (r) e reprodutibilidade (R).

Nível	Causa	QM	F-Snedecor	Cálculo	Variância	Resultado
1	Entre lab.	0,0269444	1,54	$(0,0269444 - 0,0175)/2$	S_R^2	0,0222
	D. lab.	0,0175		0,0175	S_r^2	0,0175
2	Entre lab.	0,150222	5,56 *	$(0,150222 - 0,027)/2$	S_R^2	0,088611
	D. lab.	0,027		0,027	S_r^2	0,027
3	Entre lab.	0,1905	2,18	$(0,1905 - 0,0875)/2$	S_R^2	0,139
	D. lab.	0,0875		0,0875	S_r^2	0,0875

Causa: causa de variação; Lab: laboratórios; D.: dentro; QM: quadrado médio; S_R^2 : variação da reprodutibilidade; S_r^2 : variação da repetitividade.

Assim, é necessária uma análise crítica nos dados obtidos nesse nível no analito para que se verifique se houve algum problema referente à matriz, a algum analista, às condições dos laboratórios, à redação do método, etc., para que se possa concluir sobre a possível robustez do método analítico.

Desse modo, recomenda-se que o teste estatístico da reprodutibilidade seja realizado em qualquer experimento de precisão.

Amostragem



mostra, como cientificamente conhecida e exposta na Organization for Economic Cooperation and Development (2010) – organização que traz as normas de boas práticas de laboratórios –, é um subconjunto de elementos selecionados de uma população com uma probabilidade conhecida de seleção, ou o conjunto de unidades observadas cujas propriedades se estuda, National Institute of Standards and Technology (2010), ou a conhecida amostra aleatória, como em Duncan (1974), não são comumente abordadas em muitas normas.

A Commission of the European Communities (2006), como outras normas, não fornece método para se obter uma amostra de acordo com a conceituação estocástica. Apresenta, porém, o que se pode chamar de uma unidade de observação, ou

porção amostral, do tipo composta, formada por subporções agregadas e homogeneizadas. Normas desse tipo não trazem uma sistemática estocástica – científica –, são resultados de acordos internacionais a respeito de esquemas de coleta de porções a serem estudadas, não de amostras representativas de um lote, de uma cultivar, de um processo, etc.

O que rege a tomada de decisão nessas normas não é uma regra científica, mas um acordo envolvendo um valor consensual e nenhuma regra de amostragem explícita.

Nesses casos, os resultados de medição da unidade (ou porção) obtida podem representar uma evidência. Por exemplo, se for obtido 30 µg/kg de aflatoxina A1 em uma porção de milho de um lote, não se poderá dizer que o lote tem em média 30 µg/kg de aflatoxina A1, mas que pode apresentar até 30 µg/kg de aflatoxina A1 em algumas de suas porções. Trata-se de um caso de evidência, não de estimação. Assim, não se pode dizer qual é a média de contaminação do lote, e esse pode, até mesmo, ter média menor que 30 µg/kg de aflatoxina A1. Então se pode condenar injustamente um lote com base nesse resultado, e sem saber com que probabilidade esse resultado está ou não correto. Seria recomendável, portanto, uma contraprova, ou melhor, um processo científico de amostragem estocástica.

Tem sido comum o emprego de amostras sem sistemas de geração aleatórios adequados. As amostras coletadas conforme este aqui, aquele não, aquele outro sim, etc. são ditas amostras a esmo. Não são aleatórias e são um tipo de amostra muito susceptível

a desvios não intencionais. Há ainda as amostras intencionais, em que se busca a irregularidade pretendida em *locus* suspeitos, como procurar madeira de lei fora do centro do invólucro, onde estaria com maior probabilidade, para fugir à fiscalização que comumente investiga a região central.

Incerteza de amostragem



este item é adotada a grafia normal para amostra estocástica e "amostra" para o que se considera amostra na referida publicação.

Com o decorrer do tempo, frente a confrontos com estatísticos conhecedores de modelagem que discordavam do cálculo de incerteza por componentes e clamavam por uma abordagem por modelagem estocástica como Working Party on Chemical Contaminants in Foods (2003), com o surgimento de inevitáveis publicações mais detalhadas envolvendo incerteza e suas propriedades como a International Organization for Standardization (2004), e outras, foi-se tornando cada vez mais claro que a incerteza por componentes tem validade limitada para determinados propósitos.

Publicou-se, então, a obra de Ramsey e Ellison (2007), referente a uma pretensa incerteza, então chamada “incerteza de amostragem”. Essa obra considera a possibilidade de uma “amostra” obtida em um processo de produção agrícola de modo a se poder

estimar uma “incerteza” válida para tomada de decisão que não seja baseada somente em informações intralaboratoriais.

Esses autores comentam a necessidade de uma incerteza total, que inclua, por exemplo, a heterogeneidade das matrizes, objeto de conclusão. Prosseguem admitindo modelos de delineamento experimentais pretendendo obter os componentes de variância dos modelos. A principal causa de variação incluída nos modelos é a “amostra”, ou seja, a unidade coletada a ser estudada que deve representar o processo estudado. Essa unidade é subdividida, porções de teste são retiradas e medidas. Os resultados de medição, então, são considerados como respostas nos modelos.

Do ponto de vista desses autores, a heterogeneidade natural dominaria as causas de variação, seria chamada de “incerteza de amostragem”, e poderia ser isolada e estimada. A estimação da “incerteza de amostragem” seria obtida em modelo em que os critérios de classificação seriam “amostras” e “subamostras”.

De fato, o que esses autores trazem é o emprego de modelos de delineamento de classificação hierárquica (*nested*) e seus derivados em que o tronco do modelo são as “amostras”, na realidade, unidades de observação. Não sejam confundidas neste item as “amostras” conforme conceituadas por esses autores e as amostras obtidas em processos estatísticos de amostragem do item Amostragem.

Apresenta-se aqui o exemplo 1 desses autores de modo que o leitor possa compreendê-lo e refazer os cálculos para seu aprendizado ou conferência.

O esquema experimental usado é o que se pode ver na Figura 3.

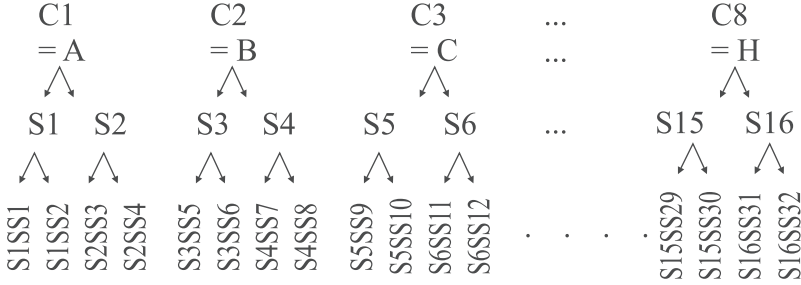


Figura 3. Esquema real do delineamento experimental no exemplo 1 de Ramsey, Ellison.

Fonte: Ramsey e Ellison (2007).

Em cada “amostra” (canteiros) A a H, nesse esquema, colhem-se duas “subamostras” (pés de alface) obtendo-se S1 a S16 (dois pés em cada um dos 8 canteiros), cada uma das quais é repartida em duas, gerando 32 “subsubamostras”. O modelo de delineamento experimental é de classificação hierárquica, conforme:

$$y_{i,j,k} = \mu + c_i + s_{ij} + \varepsilon_{ijk}$$

em que $y_{i,j,k}$ é o valor observado (medido), μ é a média, c_i o efeito do canteiro_{*i*} (ou “amostras”, ou etc.), $i = 1, 2, \dots, 8$, s_{ij} o erro da “subamostra_{*ij*}” (alface_{*ij*}), $j = 1, 2$, e ε_{ijk} o erro de análises, $k = 1, 2$ – ou efeito da “subsubamostra_{*k*}” –, dentro da “subamostra_{*ij*}”. A variável medida é o teor de nitrato.

Procedendo-se a obtenção dos componentes de variância desse modelo, quanto à soma de quadrados total obteve-se:

$$E \left(\sum y_{ijk}^2 \right) = \sum_{i,j,k} (\mu^2 + c_i^2 + s_{ij}^2 + \varepsilon_{ij}^2 + 2 \text{ duplos}) = IJK(\mu^2 + \sigma_c^2 + \sigma_s^2 + \sigma^2)$$

$$E(C) = IJK\mu^2 + JK\sigma_c^2 + K\sigma_s^2 + \sigma^2, \text{ logo}$$

$$\begin{aligned} E(SQ_{Tot}) &= (IJK\mu^2 + IJK\sigma_c^2 + IK\sigma_s^2 + I\sigma^2) - (IJK\mu^2 + JK\sigma_c^2 + K\sigma_s^2 + \sigma^2) \\ &= JK(I-1)\sigma_c^2 + K(I-1)\sigma_s^2 + (I-1)\sigma^2 \end{aligned}$$

No que se refere à soma de quadrados de canteiros,

$$E(SQ_{Cant}) = E\left[\frac{1}{JK} \sum_i y_{i..}^2 - C\right] = \frac{1}{JK} E\left(\sum_i y_{i..}^2\right) - E(C), \text{ mas}$$

$$\begin{aligned} \frac{1}{JK} E\left(\sum_i y_{i..}^2\right) &= \frac{1}{JK} \sum_i E\left[JK\mu + JKc_i + K \sum_j S_{ij} + \sum_{j,k} \varepsilon_{ijk}\right]^2 \\ &= (IJK\mu^2 + IJK\sigma_c^2 + IK\sigma_s^2 + I\sigma^2), \text{ logo} \end{aligned}$$

$$\begin{aligned} E(SQ_{Cant}) &= (IJK\mu^2 + IJK\sigma_c^2 + IK\sigma_s^2 + I\sigma^2) - (IJK\mu^2 + JK\sigma_c^2 + K\sigma_s^2 + \sigma^2) \\ &= JK(I-1)\sigma_c^2 + K(I-1)\sigma_s^2 + (I-1)\sigma^2, \text{ então} \end{aligned}$$

$$E(QM_{Cant}) = JK\sigma_c^2 + K\sigma_s^2 + \sigma^2$$

Prosseguindo os cálculos obtém-se:

$$E(SQ_{Subamost d. Cant}) = E\left(\sum_{i,j} \frac{y_{ij.}^2}{K} - \sum_i \frac{y_{i..}^2}{JK}\right), \text{ onde}$$

$$E\left(\sum_{i,j} \frac{y_{ij.}^2}{K}\right) = IJK\mu^2 + IJK\sigma_c^2 + IJK\sigma_s^2 + IJ\sigma^2, \text{ portanto}$$

$$\begin{aligned} E(SQ_{Subamost d. Cant}) &= (IJK\mu^2 + IJK\sigma_c^2 + IJK\sigma_s^2 + IJ\sigma^2) \\ &\quad - (IJK\mu^2 + IJK\sigma_c^2 + IK\sigma_s^2 + I\sigma^2) = IK(J-1)\sigma_s^2 + I(J-1)\sigma^2 \end{aligned}$$

então,

$E(QM_{Subamost d. Cant}) = k\sigma_s^2 + \sigma^2$, assim a incerteza padrão de “amostragem” seria dada por:

$$\sigma_{\text{amost}} = \sqrt{\sigma_s^2 + \sigma^2}$$

e se pode colocar os resultados em um esquema conforme a Tabela 4 e estimar os componentes obtidos.

Tabela 4. Esquema de análise de variância do modelo e esperanças matemáticas.

Causas variação	GL	S. Quadr.	QMédios	E(SQ)
Canteiros	I-1	S1	QMCant	$\sigma^2 + K\sigma_s^2 + JK\sigma_c^2$
“Subamostras” d. Canteiros	I(J-1)	S2	QMSubd.Cant	$\sigma^2 + K\sigma_s^2$
Res. (med. d. sub d. cant)	IJ(K-1)	S3	QMRes	σ^2
Total	IJK-1	S4		

GL: graus de liberdade; QMédios: quadrados médios; S. Quadr.: soma de quadrados; d.: dentro; med.: medição; E(SQ): esperança da soma de quadrados; σ^2 : variância residual; σ_c^2 : variância de canteiros; σ_s^2 : variância de amostragem.

Empregando-se os componentes obtidos no exemplo 1 do mencionado guia foi elaborada a Tabela 5.

Tabela 5. Resultados obtidos para o exemplo 1 do guia de incerteza de “amostragem”.

Causas variação	GL	SQ	QM	F
Entre canteiros	7	12.577.112,88	1.796.730,41	81,83 *
“Subamostras” d. canteiros	8	4.471.511,00	558.938,87	25,46 *
Análises (erros aleatórios)	16	351.320,00	21.957,50	
Total	31	17.399.943,88		

GL: graus de liberdade; SQ: soma de quadrados; QM: quadrado médio;

F: F – Snedecor; *: significativo com $p < 0,05$; d.: dentro.

Nota: tabela não apresentada no Guia.

A partir dessa Tabela foram estimados:

$$S^2_{análises} = 21.957,50 \Rightarrow S_{análises} = \sqrt{21.957,50} = 148,1806$$

$$s^2_{canteiros} = \frac{QM_{Entre\ canteiros} - QM_{Subamostras}}{n_{observações\ canteiros}} = (1.796.730,41 - 558.938,87) / 4 = 309.447,885, \text{ logo}$$

$$S_{canteiros} = \sqrt{S^2_{canteiros}} = \sqrt{309.447,885} = 556,2804, \text{ e}$$

$$S^2_{subamostras} = S^2_{subamostras\ d.\ canteiros} - S^2_{análises} / 2$$

$$= \frac{558.938,87 - 21.957,50}{2} = 268.490,685, \text{ então,}$$

$$S_{subamostras} = \sqrt{S^2_{subamostras}} = \sqrt{268.490,685} = 518,1609$$

portanto,

$$S^2_{amost} = S^2_{análises} + S^2_{subamostras} = 21.957,50 + 268.490,685 = 48.448,18$$

$$\text{e } S_{amos} = 220,11$$

Os resultados dos cálculos presentemente realizados foram, então, arranjados; e foram obtidos novos resultados, todos apresentados na Tabela 6.

Assim, pelo confronto com os resultados publicados por esses autores, foram validadas as estimativas de componentes de variâncias presentemente obtidas e se pôde prosseguir a discussão. Retornando à Tabela 5, dela constam efeitos significativos entre canteiros e “subamostras” dentro de canteiros. Esses resultados são absolutamente inesperados do ponto de vista da amostragem estocástica, dos fundamentos do conceito de distribuição probabilística e da possibilidade de existência de alguma amostra representativa.

Tabela 6. Resultados dos cálculos dos componentes de variância do Exemplo 1 do guia de “incerteza de amostragem”.

Causa variação	Nome no guia	QM	Variância	Desvio padrão	% VAR
Entre canteiros	Entre <i>target</i> (canteiros)	1.796.730,41	309.447,885	556,2804	51,58
"Subamostras" d. canteiros	"Amostragem"	558.938,87	268.490,685	518,1609	44,76
Análises (erros aleatórios)	Análises	21.957,5	21.957,5	148,1806	3,66
Total			599.896,07		100,00

QM: quadrado médio; VAR: variação; d.: dentro.

Do ponto de vista probabilístico, se há efeito significativo de canteiros, ou seja, médias diferentes dos teores do analito entre canteiros, eles não poderiam compor uma população uniforme referentemente ao analito e não poderiam constar juntos em uma amostra. O mesmo se aplica às subamostras dentro dos canteiros.

Cabe lembrar a função de distribuição de medições. Tomando-se uma amostra aleatória de unidades em que se realizaram medições de determinado analito, amostra de n medições, $n > 2$, com observações independentes e provenientes de uma população específica, somente uma, com uma única média, pode-se obter uma curva de densidade representada conforme a Figura 4.

Nessa curva o intervalo representado concentra 95% dos resultados de medições no espaço delimitado pela seta. Ou seja, qualquer valor observado nesse intervalo faz parte dessa curva que tem uma, e uma única, média. Então, observações obtidas nesse intervalo não

serão significativamente diferentes umas das outras com nível de significância de 5%.

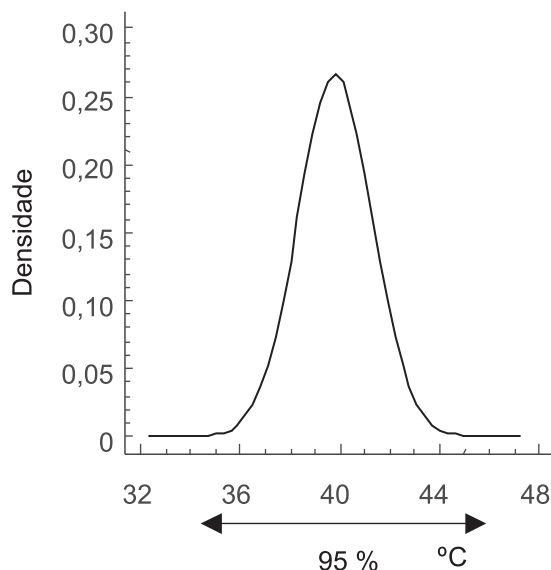


Figura 4. Função de densidade e intervalo de concentração de 95% das medições.

Uma amostra probabilística é constituída de unidades amostrais obtidas de uma única população, ou seja, uma população com uma média, e somente uma. Obter efeitos significativos de canteiros ou de “subamostras” nos canteiros seria ter unidades amostrais que não viriam de uma única população, portanto os canteiros ou “subamostras” jamais constituiriam uma amostra. Aceitar efeitos desse tipo, significativos, seria renegar a estatística toda como ciência; desde suas bases, todas as pesquisas amostrais seriam inválidas, o que constituiria um absurdo.

Por outro lado, se esses efeitos jamais podem ser significativos para que se tenha uma amostra

estocástica, eles não poderiam diferir de zero e, como consequência, o modelo de delineamento, quer seja, classificação hierárquica, usado, se torna

$$y_i = \mu + \varepsilon_i$$

em que y_i é o valor observado (medido), $i = 1, 2, \dots, n$, $n > 2$, μ é a média, e ε_i é um erro aleatório, $\varepsilon_i \in \mathbb{R}$. Esse é um dos modelos mais simples para se expressar qualquer tipo de medição, como uma média e um erro aleatório, real, positivo ou não.

É necessário, portanto, compreender que variação, incerteza, é intrínseca não somente ao processo de medição, mas a qualquer fenômeno, item, etc. que seja medido, porque, simplesmente, nada é absolutamente igual ou absolutamente controlado nos fenômenos naturais ou sistemas. Ainda, variações nos resultados não são somente provenientes das causas de variação internas aos laboratórios, mas de externas também, como de variedade, de safra, de clima, etc.

Variação é a regra, exatidão é o impossível, exceto quando motivada por imprecisão na medição, como obter o mesmo pH em duas soluções. Assim, a chamada “incerteza de amostragem” constitui um equívoco.

Verificação de conformidade



o caso de uma distribuição D , simétrica, com gl graus de liberdade, admitindo-se α como a margem de erro, um valor x seria pertencente a essa distribuição do

ponto de vista da realização de um teste de hipótese, se a hipótese $H_0 : x \in \Omega$ (Ω conjunto de realização, população) não fosse rejeitada em favor de $H_1 : x \notin \Omega$. Para tanto é necessário que:

$$\frac{|x - \bar{x}|}{s / \sqrt{n}} < D_{gl, \alpha/2}$$

em que \bar{x} é a média amostral, s é o desvio padrão, n é o número de observações na amostra (estocástica), D a distribuição (simétrica) seguida por x na validade da hipótese nula, gl é o número de graus de liberdade, e α o nível de significância. Note-se que no caso de distribuições assimétricas o argumento será similar.

Desse modo,

$$|x - \bar{x}| < D_{gl, \alpha/2} X S / \sqrt{n}$$

então uma vez que se chame,

$$k = D_{gl, \alpha/2}$$

será obtido

$$\bar{x} - k X S / \sqrt{n} < x < \bar{x} + k X S / \sqrt{n}$$

que será um intervalo de confiança a 95% para x . Esse intervalo está representado na Figura 5.

Assim, realizar um teste de hipóteses equivale a construir um intervalo de confiança. Intervalos desse tipo, também chamados incerteza expandida, têm sido a base de muitos documentos em qualidade que o empregam em verificação de conformidade.

Eurachem, Cooperation on International Traceability in Analytical Chemistry (2007) e outros trazem formas de interpretação desse intervalo para uso em verificações de conformidade. A relação entre testes de hipóteses e intervalos de confiança, porém, não

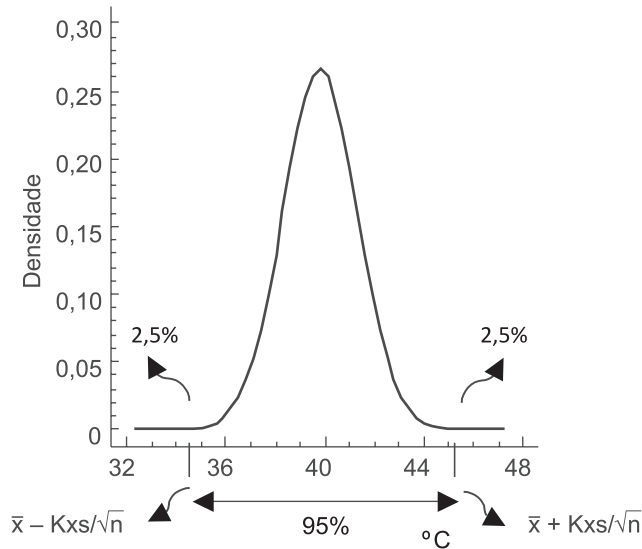


Figura 5. Função de densidade e limites para intervalo de concentração a 95% de confiança, ou teste de hipótese de $H_0: x \in \Omega$ para as medições.

tem sido notada em qualidade nas engenharias, e o uso que se faz desse intervalo, por vezes, é equivocado.

Em verificações desse tipo, pela estrutura dos testes de hipóteses e sua relação com esse intervalo, estão conformes os itens que apresentarem medições dentro do intervalo, e os demais não estarão.

Não se pode classificar um item como não conforme porque o intervalo de confiança equivocadamente traçado pela incerteza (variância) disponível tem parte fora da região de não rejeição. Por essa relação que se manifesta como equivalência, esse item não pode, de forma alguma, ser classificado como não conforme, como se encontra em certas publicações. Por essa razão, são geradas perdas

financeiras, industriais e ambientais imensas, perdas de empregos, credibilidade e muitas outras, e coloca-se em risco, sem base suficiente, a imagem de instituições, países, produtores e profissionais.

A verificação de conformidade poderá ser realizada em condições como as previstas pela Associação Brasileira de Normas Técnicas (1988), que já traz a posição brasileira da interpretação desses intervalos, concordante com a aqui apresentada, e menciona processo amostral apropriado.

Crítica da validação



Todo trabalho estocástico deve sofrer uma revisão crítica, método a método, resultado a resultado. Na prática estatística, um profissional experiente o faz no decorrer da elaboração de todo projeto, até chegar a conclusões sobre métodos e resultados, mesmo que anteriormente verificados. Apesar disso, corre-se o risco de não verificar todas as condições de realização do método, de faltar observação/estudo de uma região crítica ou, principalmente, de equívocos provocados por baixo poder de teste causado por não se dispor de observações suficientes para que uma condição de invalidade fosse considerada evidência estatística suficiente para sua rejeição.

Desse modo, apesar da crítica estocástica, o método deve ser monitorado por mais algum tempo a fim de se verificar se alguma condição permaneceu ou

não inalterada, e se ele permanece ou não válido (revalidação). Deve-se incluir nesse monitoramento desde a unidade amostral, extensão de conclusões, considerações sobre poder de teste, testes alternativos, casos qualitativos (como acompanhamento de EP's), etc.

Nos testes estocásticos não há, na verdade, aceitação, e, sim, falta de evidência suficiente para se rejeitar uma hipótese. Assim, concluir sobre validade e conformidade é rigorosamente impossível; pode-se, somente, concluir sobre invalidade e não conformidade. Validade não se prova, validade se crê, ou se refuta. E validação sem crítica não é aceitável.

Conclusões



Foram colocados nesse item alguns aspectos críticos considerados absolutamente relevantes em validação de métodos analíticos, quer sejam:

- 1) As publicações em validação simplificam este tema tornando-o um campo perigoso para o cliente e deixando de atribuir a devida responsabilidade àquele a quem compete assinar os seus resultados estocásticos e conclusões, o estatístico. Ademais, a existência desses profissionais no País é um diferencial e constitui um enorme fator de competitividade.
- 2) Há grandes dificuldades no exercício e na leitura do material de validação pela extensão, pela

profundidade do conhecimento exigido em validação de métodos analíticos e pela nomenclatura usada em qualidade para laboratórios, diversa da clássica.

- 3) O nível necessário de validação recomendado inclui testes de limites impostos por legislação, testes de alguma forma de precisão e outros de alguma forma de vício que sejam necessários para que o método seja considerado adequado ao uso.
- 4) Isso, claramente, resulta na não necessidade de testar todos os parâmetros de validação quando se realiza uma modificação em um método validado ou oficial.
- 5) Os experimentos devem ser cegos, e planejados para ter um número razoável de repetições. Dependendo do caso, melhoramento de experimentos deve ser empregado.
- 6) Experimentos de precisão (estudos colaborativos) devem ter a variação entre laboratórios do modelo do delineamento devidamente testada.
- 7) Experimentos fracionados somente devem ser realizados em casos de interações inexistentes (não podem ser, meramente, consideradas inexistentes).
- 8) É necessário esclarecer de forma inequívoca a que tipo de amostra uma medição se refere.
- 9) O conceito de amostra aleatória não está adequadamente difundido em qualidade para alimentos.

- 10) A assim chamada incerteza de amostragem revelou-se equivocada.
- 11) A realização de testes de hipóteses se equivale à construção de intervalos de confiança (incerteza expandida), e implica sua interpretação correta.
- 12) Há uma posição brasileira em torno da interpretação e uso de intervalos de confiança na forma de norma ABNT, quer seja Associação Brasileira de Normas Técnicas (1988).
- 13) Validação de métodos analíticos é uma tarefa rigorosa, de conteúdo teórico avançado, e exige uma análise crítica adequada.

Referências



AMERICAN SOCIETY FOR TESTING MATERIALS. D 4392: ASTM Standard terminology for statistically related terms. **Annual Book of ASTM Standards**, Philadelphia, v. 7, n. 1, 1987.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 11155**: interpretação estatística de dados: determinação de um intervalo de tolerância estatístico. Rio de Janeiro, 1988.

ASSOCIATION OFFICE OF ANALYTICAL CHEMISTS. Guidelines for collaborative procedure to validate characteristics of a method of analysis. **Journal of the Association Office of Analytical Chemists**, Arlington, v. 72, n. 4, p. 694-704, 1989.

ASSOCIATION OFFICE OF ANALYTICAL CHEMISTS. **Guidelines for collaborative study procedures to validate**

characteristics of a method of analysis. Gaithersburg: AOAC, 1994. 24 p.

COMMISSION DECISION. EC 657: implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results. **Official Journal of the European Communities**, Luxemburgo, 2002. p. 8-36. Disponível em: <http://eur-lex.europa.eu/RECH_naturel.do>. Acesso em: 20 jul. 2010.

COMMISSION OF THE EUROPEAN COMMUNITIES. **EC 401**: laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. 2006. Disponível em: <<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:070:0012:0034:EN:PDF>>. Acesso em: 20 jul. 2010.

CORNELL, J. A. **How to apply response surface methodology.** Milwaukee: ASQC, 1984. 49 p.

DUNCAN, A. J. **Quality control and industrial statistics.** Homewood: Richard D. Irwin, 1974. 1047 p.

EURACHEM, COOPERATION ON INTERNATIONAL TRACEABILITY IN ANALYTICAL CHEMISTRY; EURACHEM; CITAC WORKING GROUP. **Eurachem/Citac guide, quantifying uncertainty in analytical measurement.** 2nd ed. Teddington: Eurachem, 2000. 122 p.

EURACHEM, COOPERATION ON INTERNATIONAL TRACEABILITY IN ANALYTICAL CHEMISTRY. **Eurachem/Citac guide use of uncertainty information in compliance assessment.** Teddington: Eurachem, 2007. 15 p.

EURACHEM. **The fitness for purpose of analytical methods, a laboratory guide to method validation and related topics.** Teddington: Eurachem, 1998. 61 p.

EUROPEAN STATISTICS, EUROPEAN COMMISSION. **Ramon Eurostat's metadata server, CODED (Eurostat's concepts and definitions database).** Disponível em: <http://ec.europa.eu/eurostat/ramon/index.cfm?TargetUrl=DSP_PUB_WELC>. Acesso em: 15 jul. 2010.

FUENTES-ARDERIU, X. **Glossary of ISO terms**. Disponível em: <<http://www.westgard.com/isoglossary.htm>> Acesso em: 30 jun. 2010.

HAMAKER, H. C. A statistician's approach to repeatability and reproducibility. **Journal of the Association of Analytical Chemists**, Arlington, v. 69, n. 3, p. 417-428, 1986.

HORWITZ, W. Protocol for the design, conduct and interpretation of collaborative studies. **Pure and Applied Chemistry**, London, GB, v. 60, n. 6, p. 855-864, 1988.

HORWITZ, W. Protocol for the design, conduct and interpretation of method-performance studies. **Pure and Applied Chemistry**, London, GB, v. 67, n. 2, p. 331-343, 1995.

INMETRO. INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL. **DOQ-CGCRE-008**: orientação sobre validação de métodos de ensaios químicos. Rio de Janeiro: Inmetro, 2010. 20 p.

INMETRO. INSTITUTO NACIONAL DE METROLOGIA, NORMALIZAÇÃO E QUALIDADE INDUSTRIAL; CONFEDERAÇÃO NACIONAL DA INDÚSTRIA; SERVIÇO NACIONAL DE APRENDIZAGEM INDUSTRIAL. **Vocabulário internacional de termos fundamentais e gerais de metrologia**. Brasília, DF: Inmetro: CNI: Senai, 2000. 75 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 21748**: guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation. Genève, 2004.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725-1**: accuracy (trueness and precision) of measurement methods and results: part 1: general principles and definitions. Genève, 1994a.

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, 1994b.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725-3**: accuracy (trueness and precision) of measurement methods and results; part 3: intermediate measures of the precision of a standard measurement method. Genève, 1994c.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725-4**: accuracy (trueness and precision) of measurement methods and results; part 4: basic methods for the determination of the trueness of a standard measurement method. Genève, 1994d.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725-5**: accuracy (trueness and precision) of measurement methods and results; part 5: alternative methods for the determination of the precision of a standard measurement method. Genève, 1998a.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 5725-6**: accuracy (trueness and precision) of measurement methods and results; part 6: use in practice of accuracy values. Genève, 1998b.

LOHNINGER, H. **Fundamentals of statistics**. Disponível em: <http://www.statistics4u.info/fundstat_eng/copyright.html>. Acesso em: 20 jul. 2010.

NATIONAL ASSOCIATION OF TESTING AUTHORITIES. **NATA Technical note 17**: format and contents of test methods and procedures for validation and verification of chemical tests methods. Sidney, 1998.

NATIONAL FOOD ADMINISTRATION. **Method validation and quality control procedures for pesticide residues analysis in food and feed**: document no. Sanco/10684/2009. Upsalla: National Food Administration, 2009. 42 p.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY; SEMATECH. **NIST/SEMATECH**

e-handbook of statistical methods. Disponível em: <<http://www.itl.nist.gov/div898/handbook/glossary.htm#S>>. Acesso em: 20 jun. 2010.

NORMA MERCOSUR; INTERNATIONAL ORGANIZATION FOR STANDARDIZATION; INTERNATIONAL ELETROTECHNICAL COMMISSION. **NM/ ISO/IEC 17025:** requisitos gerais para a competência de laboratórios de ensaio e calibração. São Paulo, 2006.

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. **OECD glossary of statistical terms.** Disponível em: <<http://stats.oecd.org/glossary/>>. Acesso em: 20 jun. 2010.

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. **OECD series on principles of good laboratory practice and compliance monitoring number 1, OECD principles on good laboratory practice.** Paris, FR: OECD, 1998. 41 p.

POCKLINGTON, W. D. **IUPAC commission on oils, fats and derivatives, guidelines for interlaboratory collaborative studies.** London, KG: IUPAC, 1985. 16 p.

POCKLINGTON, W. D. Harmonized protocols for the adoption of standardized analytical methods and for the presentation of their performance characteristics. **Pure and Applied Chemistry**, London, GB, v. 62, n. 1, p. 149-162, 1990.

RAMSEY, M. H.; ELLISON, S. L. R. **Eurachem/Eurolab/Citac/Nordtest/AMC guide:** measurement uncertainty arising from sampling: a guide to methods and approaches. Teddington: Eurachem, 2007. 102 p.

ROSNER, B. Percentage points for a generalized ESD many-outlier procedure. **Technometrics**, Washington, DC, v. 25, n. 1, p. 165-172, 1983.

TAYLOR, B. N.; KUYATT, C. E. **NIST Technical Note 1297, guidelines for evaluating and expressing the uncertainty of NIST measurement results.** Gaithersburg, NIST, 1994. 25 p.

THOMPSON, N.; STEPHEN, L. R.; WOOD, R. Harmonized guidelines for the single laboratory validation of methods of analysis, IUPAC Technical Report. **Pure and Applied Chemistry**, London, GB, v. 74, n. 5, p. 835-855, 2002.

WORKING PARTY ON CHEMICAL CONTAMINANTS IN FOODS 2003. **Annex to WPCC/2003/9**: report of the Standing Committee on the Food Chain and Animal, health on the relationship between analytical results, the measurement of uncertainty recovery, factors and the provisions in EU food and feed legislation. Disponível em: <www.food.gov.uk/multimedia/pdfs/wpmin36final.pdf>. Acesso em: 20 jul. 2010.



Títulos lançados



1998

Nº 1 – A pesquisa e o problema de pesquisa:
quem os determina?

Ivan Sergio Freire de Sousa

Nº 2 – Projeção da demanda regional de grãos no Brasil: 1996 a 2005

*Yoshihiko Sugai, Antonio Raphael Teixeira Filho, Rita de Cássia
Milagres Teixeira Vieira e Antonio Jorge de Oliveira,*

1999

Nº 3 – Impacto das cultivares de soja da Embrapa e rentabilidade
dos investimentos em melhoramento

*Fábio Afonso de Almeida, Clóvis Terra Wetzel e
Antonio Flávio Dias Ávila*

2000

Nº 4 – Análise e gestão de sistemas de inovação em organizações
públicas de P&D no agronegócio

Maria Lúcia D'Apice Paez

Nº 5 – Política nacional de C&T e o programa de biotecnologia
do MCT

Ronaldo Mota Sardenberg

Nº 6 – Populações indígenas e resgate de tradições agrícolas

José Pereira da Silva

2001

Nº 7 – Seleção de áreas adaptativas ao desenvolvimento agrícola,
usando-se algoritmos genéticos

Jaime Hidehiko Tsuruta, Takashi Hoshi e Yoshihiko Sugai

Nº 8 – O papel da soja com referência à oferta de alimento
e demanda global

Hideki Ozeki, Yoshihiko Sugai e Antonio Raphael Teixeira Filho

Nº 9 – Agricultura familiar: prioridade da Embrapa
Eliseu Alves

Nº 10 – Classificação e padronização de produtos, com ênfase na agropecuária: uma análise histórico-conceitual
Ivan Sergio Freire de Sousa

2002

Nº 11 – A Embrapa e a aquicultura: demandas e prioridades de pesquisa
Júlio Ferraz de Queiroz, José Nestor de Paula Lourenço e Paulo Choji Kitamura (Eds.)

Nº 12 – Adição de derivados da mandioca à farinha de trigo: algumas reflexões
Carlos Estevão Leite Cardoso e Augusto Hauber Gameiro

Nº 13 – Avaliação de impacto social de pesquisa agropecuária: a busca de uma metodologia baseada em indicadores
Levon Yeganiantz e Manoel Moacir Costa Macêdo

Nº 14 – Qualidade e certificação de produtos agropecuários
Maria Conceição Peres Young Pessoa, Aderaldo de Souza Silva e Cilas Pacheco Camargo

Nº 15 – Considerações estatísticas sobre a lei dos julgamentos categóricos
Geraldo da Silva e Souza

Nº 16 – Comércio internacional, Brasil e agronegócio
Luiz Jésus d'Ávila Magalhães

2003

Nº 17 – Funções de produção – uma abordagem estatística com o uso de modelos de encapsulamento de dados
Geraldo da Silva e Souza

Nº 18 – Benefícios e estratégias de utilização sustentável da Amazônia
Afonso Celso Candeira Valois

Nº 19 – Possibilidades de uso de genótipos modificados e seus benefícios
Afonso Celso Candeira Valois

2004

Nº 20 – Impacto de exportação do café na economia do Brasil – análise da matriz de insumo-produto

Yoshihiko Sugai, Antônio R. Teixeira Filho e Elisio Contini

Nº 21 – Breve história da estatística

José Maria Pompeu Memória

Nº 22 – A liberalização econômica da China e sua importância para as exportações do agronegócio brasileiro

Antônio Luiz Machado de Moraes

2005

Nº 23 – Projetos de implantação do desenvolvimento sustentável no Plano Plurianual 2000 a 2003 – análise de gestão e política pública em C&T

Marlene de Araújo

2006

Nº 24 – Educação, tecnologia e desenvolvimento rural – relato de um caso em construção

Elisa Guedes Duarte e Vicente G. F. Guedes

2007

Nº 25 – Qualidade do emprego e condições de vida das famílias dos empregados na agricultura brasileira no período 1992–2004

Otávio Valentim Balsadi

Nº 26 – Sistemas de gestão da qualidade no campo

Vitor Hugo de Oliveira, Janice Ribeiro Lima, Renata Tieko Nassu, Maria do Socorro Rocha Bastos, Andréia Hansen Oster e Luzia Maria de Souza Oliveira

2008

Nº 27 – Extrativismo, biodiversidade e biopirataria na Amazônia

Alfredo Kingo Oyama Homma

Nº 28 – A construção das alegações de saúde para alimentos funcionais

André Luiz Bianco

Nº 29 – Algumas reflexões sobre a polêmica agronegócio versus agricultura familiar

Ana Lúcia E. F. Valente

Nº 30 – Agricultura familiar versus agronegócio: a dinâmica sociopolítica do campo brasileiro

Sérgio Sauer

Nº 31 – O conteúdo social da tecnologia

Michelangelo Giotto Santoro Trigueiro

Nº 32 – Dimensões, riscos e desafios da atual expansão canavieira

Tamás Szmrécsányi, Pedro Ramos, Luiz Octávio Ramos Filho e Alceu de Arruda Veiga Filho

Nº 33 – Procedimentos de sustentabilidade no sistema de produção de grãos

Carlos Magri Ferreira

Nº 34 – A agrobiodiversidade com enfoque agroecológico: implicações conceituais e jurídicas

Altair Toledo Machado, Juliana Santilli e Rogério Magalhães

2009

Nº 35 – As indicações geográficas como estratégia mercadológica para vinhos

Rogério Fabrício Glass e Antônio Maria Gomes de Castro

Nº 36 – Embrapa Brasil: análise bibliométrica dos artigos na Web of Science (1977–2006)

Roberto de Camargo Penteado Filho e Antonio Flavio Dias Avila

Nº 37 – Estudo das citações dos artigos da Embrapa na Web of Science de 1977 a 2006

Roberto de Camargo Penteado Filho e Antonio Flavio Dias Avila

2010

Nº 38 – Rumo a uma sociologia da agroenergia

Ivan Sergio Freire de Sousa

Nº 39 – Fatores de influência no preço do milho no Brasil

Carlos Eduardo Caldarelli e Mirian Rumenos Piedade Bacchi

Instruções aos autores

Processo editorial: submissão de originais e informações gerais

Processo editorial

Os originais submetidos à série *Texto para Discussão* passam, a princípio, por uma avaliação vestibular na editoria. Formula-se, nessa etapa, um juízo de admissibilidade depois de conferido se o artigo atende aos requisitos formais para sua admissão no processo editorial. Posteriormente, mediante correspondência a Editoria comunica o resultado ao(s) autor(es).

No caso de admissão nessa etapa vestibular, os originais são submetidos à avaliação de editores associados (pareceristas), que, após analisarem tanto a forma quanto o conteúdo do artigo, enviam à editoria o resultado da avaliação.

De posse da avaliação dos pareceristas, a editoria elabora um comunicado síntese endereçado ao autor, ou ao primeiro autor, cientificando-lhe a decisão, que consistirá em uma das três possibilidades a seguir: **aprovação plena**; **aprovação parcial**, com recomendação de ajustes e/ou de correções; ou **rejeição**.

No caso de aprovação parcial, depois de feitos as correções e os ajustes solicitados pela editoria, o autor, ou primeiro autor, devolve-lhe o trabalho para verificação da pertinência das modificações processadas. Se aprovada, a nova versão do artigo é repassada para a revisão de textos e referências.

O artigo revisado retorna então para a editoria, que define se ela própria valida a revisão, ou se é o caso de enviá-la ao autor, ou ao primeiro autor, para que ele(s) valide(m) as alterações, as sugestões e as recomendações feitas pelos revisores. Caso a validação seja encaminhada ao(s) autor(es), este(s) deve(m) retornar a versão validada para a editoria, com as respectivas posições.

Cabe à editoria fazer uma nova verificação da versão revisada validada; manifestar-se a respeito dela, se necessário; e, posteriormente, repassar todo o material para finalização e impressão gráfica.

Orientações relativas à preparação e à apresentação dos originais

Quanto à **forma** – Independentemente do número de autores, da complexidade ou da extensão do tema em enfoque, para ser editado na série o artigo original deve ser único e inédito.

O texto deve ser digitado em Word, em papel no formato A4, com margens superior e lateral direita de 3 cm, e inferior e lateral esquerda de 2,5 cm. O espaçamento entre linhas e o de recuo de parágrafo devem ser ambos de 1,5 cm. Além disso, o artigo deve ser redigido em fonte Times New Roman, e em corpo 12; com número de páginas (numeradas sequencialmente em algarismos arábicos) limitado entre 30 e 200 (já com a inclusão de tabelas, figuras e referências).

Autores que operam programas de edição de texto diferentes do padrão Microsoft (como o BrOffice.org) devem ter o cuidado de gravar o material a ser enviado para submissão no formato documento (*.doc).

Quanto ao **estilo** – O texto deve ser escrito em linguagem técnico-científica. Não deve ter a forma de um relatório e tampouco de um artigo de opinião destinado à mídia, por exemplo.

Devem ser enviadas, à editoria, quatro cópias impressas do arquivo original, assim como o seu arquivo eletrônico gravado em CD (preferencialmente).

No arquivo eletrônico, e em uma das cópias impressas, deve constar o nome completo e demais dados que possibilitem a identificação do(s) autor(es). Nas outras três cópias impressas, no entanto, esses elementos devem ser excluídos.

Quanto à **especificação de autoria** – No rol de autores, o nome completo de cada um deles deve ser separado por vírgulas, e limitar-se a um máximo de 160 (cento e sessenta) caracteres, incluídos os espaços entre palavras. Portanto, se necessário, os próprios autores devem abreviar seu nome e sobrenome de modo a respeitar esse limite.

As nota(s) de rodapé (uma para cada autor), que deve(m) constar da primeira página do artigo, deve(m) apresentar a **qualificação dos autores**. Tal(is) nota(s) deve(m) ser vinculada(s) ao nome do(s) autor(es) e conter: formação e grau acadêmico, tipo de vínculo institucional (se for o caso), endereço postal completo e endereço eletrônico.

Na primeira nota de rodapé, vinculada ao título geral, a editoria recomenda registrar informação sobre a procedência do artigo, caso ele tenha se originado de um trabalho anterior: monografia, dissertação, tese, livre docência, pós-doutoramento, projeto de pesquisa encerrado ou em andamento, entre outros.

Quanto à **estrutura** – O artigo deve conter, ordenados, os seguintes elementos: título geral, autoria, resumo e termos para indexação, título em inglês, *abstract* e *index terms*, introdução, desenvolvimento (em que o conteúdo deve ser hierarquizado em subtítulos), conclusões e referências (bibliográficas, eletrônicas, pictográficas, entre outras, que contenham, exclusivamente, as fontes citadas).

As partes “desenvolvimento” e “conclusões” devem estar claramente definidas; entretanto, não precisam, necessariamente, ser assim intituladas.

Especificações importantes

Título: Deve ser claro e objetivo, sintetizar o conteúdo e ser grafado com, no máximo, 83 (oitenta e três) caracteres, incluídos os espaços entre palavras.

Resumo: Deve vir na primeira página, logo abaixo do título e da indicação de autoria, e ser grafado com, no máximo, 300 (trezentas) palavras, incluindo-se artigos, preposições e conjunções.

Deve ser redigido com frases curtas, claras e objetivas, que enfoquem o objetivo central do trabalho, os métodos empregados na pesquisa (se for o caso), além de seus resultados e conclusões. É altamente recomendável evitar, no resumo, citações bibliográficas, agradecimentos e siglas.

Termos para indexação: Logo após o resumo devem vir citados de 3 (três) a 5 (cinco) termos para indexação, que possam ser empregados, se necessário, na composição da ficha de catalogação. Deve-se evitar a seleção de palavras que já constem do título do artigo e da série, bem como do nome dos autores.

Title, abstract e index terms: Logo após a apresentação, em português, do título, do resumo e dos termos para indexação, deve vir a tradução de todos esses elementos para o idioma inglês.

Notas de rodapé: Devem ser em número reduzido e constar da mesma página de sua chamada, cuja indicação deve ser feita por número em algarismo arábico e sobrescrito. Recomenda-se que seu texto – que deve vir grafado no pé da página, sob um fio – seja de natureza substantiva (e não bibliográfica).

Citações: Tanto as diretas quanto as indiretas devem ser feitas em conformidade com normas da ABNT.

Referências: São indicações de dados completos de obras citadas ao longo do artigo, as quais devem ser elaboradas em conformidade com normas da ABNT.

Figuras: São gráficos, desenhos, mapas, fotografias, lâminas ou outras formas pictográficas usadas no trabalho, as quais devem ser produzidas em escala de cinza. Devem ser numeradas em algarismos arábicos e em ordem sequencial, trazer legenda elucidativa em que, além das especificações próprias, contenham também título, fonte e/ou, se for o caso, crédito (nome de fotógrafo, ilustrador, etc.). Nos gráficos, as designações das variáveis dos eixos X e Y devem ter iniciais maiúsculas, e devem ser seguidas das unidades entre parênteses.

Tabelas: Devem ser produzidas em escala de cinza, e, se for o caso, com diferenciação com cores; e ser numeradas em algarismos arábicos. Além disso, devem ter tanto sua chamada quanto sua inserção em ordem sequencial no texto, e conter fonte e títulos (geral e de cada coluna).

Orientações para o envio dos artigos

O documento de encaminhamento dos originais para submissão, análise e seleção na série deve ser em forma de **carta**, assinada pelo autor, ou pelo primeiro autor, na qual devem constar:

- Título do trabalho.
- Nome completo do(s) autor(es), seguido da indicação dos seguintes dados: formação e grau acadêmico, tipo de vínculo institucional (se for o caso), endereço institucional completo e endereço eletrônico.
- Concordância expressa do(s) autor(es) em relação à submissão do trabalho.
- Declaração de que o trabalho é original e de que não foi submetido à edição em outra publicação, quer seja impressa, quer seja eletrônica.

- Autorização para que, na condição de detentora dos direitos patrimoniais de artigo editado da série *Texto para Discussão*, assim como de garantidora de direitos morais de seu(s) autor(es), a Embrapa possa:
 - a) Reproduzi-lo por qualquer meio, a qualquer tempo, em qualquer suporte físico, no todo ou em parte.
 - b) Divulgá-lo e publicá-lo.
 - c) Utilizá-lo de forma onerosa ou não, sem limite de quantidade de exemplares, de impressão ou de edição.
 - d) Disponibilizá-lo na internet.
 - e) Autorizar terceiro a praticar quaisquer dos atos relacionados nos itens anteriores.

Os artigos devem ser encaminhados para o seguinte endereço:

Série *Texto para Discussão*

Editoria

Empresa Brasileira de Pesquisa Agropecuária (Embrapa)

Departamento de Pesquisa e Desenvolvimento (DPD)

Parque Estação Biológica (PqEB), Av. W3 Norte (final)

CEP 70770-901 Brasília, DF

Fone: (61) 3448-4451

Fax: (61) 3448-4887

textoparadiscussao@embrapa.br



Impressão e acabamento
Embrapa Informação Tecnológica

*O papel utilizado nesta publicação foi produzido conforme
a certificação do Bureau Veritas Quality International (BVQI) de Manejo Florestal.*