

# Quimiometria

**Iara de Fátima Gimenez  
Nivan Bezerra da Costa Jr.**



**São Cristóvão/SE  
2007**

# Quimiometria

## Elaboração de Conteúdo

Iara de Fátima Gimenez  
Nivan Bezerra da Costa Jr.

---

## Projeto Gráfico

Hermeson Alves de Menezes  
Leo Antonio Perrucho Mittaraquis  
Tatiane Heinemann Böhmer

## Capa

Hermeson Alves de Menezes

## Revisão

Edvar Freire Caetano

## Ilustração

Alysson Prado dos Santos, Edgar Pereira Santos Neto  
Gerri Sherlock Araújo, Manuel Messias de Albuquerque Neto

## Diagramação

Hermeson Alves de Menezes

Reimpressão

---

Copyright © 2007, Universidade Federal de Sergipe / CESAD.  
Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada por qualquer meio eletrônico, mecânico, por fotocópia e outros, sem a prévia autorização por escrito da UFS.

### FICHA CATALOGRÁFICA PRODUZIDA PELA BIBLIOTECA CENTRAL UNIVERSIDADE FEDERAL DE SERGIPE

G491q Gimenez, Iara de Fátima  
Quimiometria / Iara de Fátima Gimenez; Nivan  
Bezerra Costa Jr. - São Cristóvão : Universidade  
Federal de Sergipe, CESAD, 2007.

1. Quimiometria. 2. Química - Dados. 3. Métodos  
Matemáticos. I. Costa Jr, Nivan Bezerra da. II. Título.

CDU 54.08::51-7

**Presidente da República**  
Luiz Inácio Lula da Silva

**Chefe de Gabinete**  
Ednalva Freire Caetano

**Ministro da Educação**  
Fernando Haddad

**Coordenador Geral da UAB/UFS**  
**Diretor do CESAD**  
Antônio Ponciano Bezerra

**Secretário de Educação a Distância**  
Carlos Eduardo Bielschowsky

**Vice-coordenador da UAB/UFS**  
**Vice-diretor do CESAD**  
Fábio Alves dos Santos

**Reitor**  
Josué Modesto dos Passos Subrinho

**Vice-Reitor**  
Angelo Roberto Antonioli

---

**Diretoria Pedagógica**

Clotildes Farias (Diretora)  
Hérica dos Santos Mota  
Iara Macedo Reis  
Daniela Souza Santos  
Janaina de Oliveira Freitas

**Núcleo de Avaliação**

Guilhermina Ramos (Coordenadora)  
Carlos Alberto Vasconcelos  
Elizabete Santos  
Marialves Silva de Souza

**Diretoria Administrativa e Financeira**

Edélio Alves Costa Júnior (Diretor)  
Sylvia Helena de Almeida Soares  
Valter Siqueira Alves

**Núcleo de Serviços Gráficos e Audiovisuais**

Giselda Barros

**Coordenação de Cursos**

Djalma Andrade (Coordenadora)

**Núcleo de Tecnologia da Informação**

João Eduardo Batista de Deus Anselmo  
Marcel da Conceição Souza

**Núcleo de Formação Continuada**

Rosemeire Marcedo Costa (Coordenadora)

**Assessoria de Comunicação**

Guilherme Borba Gouy

---

**Coordenadores de Curso**

Denis Menezes (Letras Português)  
Eduardo Farias (Administração)  
Haroldo Dorea (Química)  
Hassan Sherafat (Matemática)  
Hélio Mario Araújo (Geografia)  
Lourival Santana (História)  
Marcelo Macedo (Física)  
Silmara Pantaleão (Ciências Biológicas)

**Coordenadores de Tutoria**

Edvan dos Santos Sousa (Física)  
Geraldo Ferreira Souza Júnior (Matemática)  
Janaína Couvo T. M. de Aguiar (Administração)  
Priscilla da Silva Góes (História)  
Rafael de Jesus Santana (Química)  
Ronilse Pereira de Aquino Torres (Geografia)  
Trícia C. P. de Sant'ana (Ciências Biológicas)  
Vanessa Santos Góes (Letras Português)

---

**NÚCLEO DE MATERIAL DIDÁTICO**

Hermeson Menezes (Coordenador)  
Edvar Freire Caetano  
Isabela Pinheiro Ewerton  
Lucas Barros Oliveira

Neverton Correia da Silva  
Nicolos Menezes Melo  
Tadeu Santana Tartum

---

**UNIVERSIDADE FEDERAL DE SERGIPE**  
Cidade Universitária Prof. "José Aloísio de Campos"  
Av. Marechal Rondon, s/n - Jardim Rosa Elze  
CEP 49100-000 - São Cristóvão - SE  
Fone(79) 2105 - 6600 - Fax(79) 2105- 6474



# Sumário

## **AULA 1**

Introdução à Quimiometria ..... 07

## **AULA 2**

Parâmetros estatísticos ..... 17

## **AULA 3**

Amostras, populações e histogramas ..... 29

## **AULA 4**

Distribuição normal - Parte 1 ..... 39

## **AULA 5**

Distribuição normal - Parte 2 ..... 49

## **AULA 6**

Teorema do Limite Central e Intervalo de Confiança ..... 59

## **AULA 7**

Covariância e correlação ..... 69

## **AULA 8**

Combinação linear de variáveis aleatórias e distribuição das médias.....79

## **AULA 9**

Distribuição de Student ..... 89

## **AULA 10**

Planejamento fatorial  $2^2$ .....99

## **AULA 11**

Interpretação dos dados do planejamento fatorial  $2^2$  ..... 109

## **AULA 12**

Planejamento fatorial  $2^3$  ..... 123

**AULA 13**

Interpretação dos dados do planejamento fatorial  $2^4$  ..... 135

**AULA 14**

Planejamento fatorial  $2^4$  ..... 145

**AULA 15**

Planejamentos fatoriais fracionários ..... 155

**AULA 16**

Correlação e regressão: como construir modelos empíricos ..... 165

**AULA 17**

Ajustes por mínimos quadrados ..... 177

**AULA 18**

Análise do desempenho do modelo em regressões ..... 187

**AULA 19**

Exemplo de aplicação de regressão ..... 197

**AULA 20**

Exemplo de proposta final de atividade prática ..... 209

# INTRODUÇÃO À QUIMIOMETRIA

**1**  
aula

## **META**

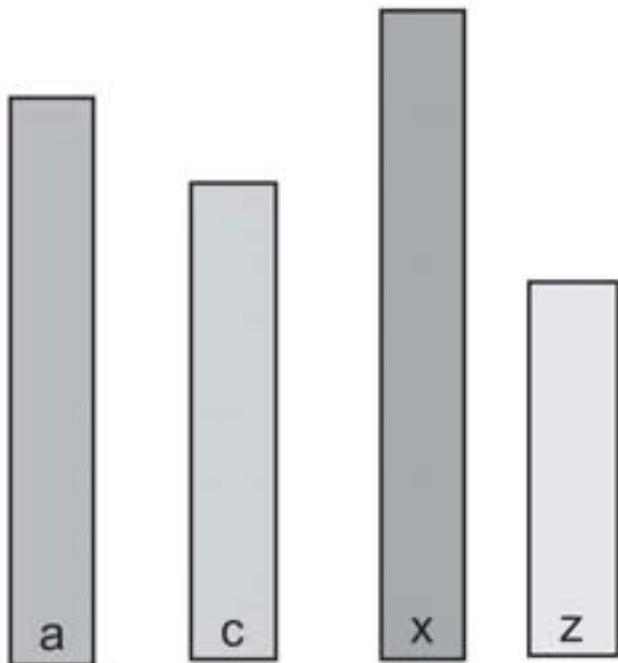
Demonstrar situações em que a estatística é fundamental e comparar casos em que ela foi usada com casos em que não foi.

## **OBJETIVOS**

Ao final desta aula, o aluno deverá:

identifique claramente o contexto e ter uma visão geral da abrangência da disciplina;

vislumbre as aplicações dos métodos a serem abordados e sua importância.



Os engenheiros, os químicos e outros profissionais que atuam na indústria e em pesquisas científicas depararam-se constantemente com a necessidade de resolver problemas práticos através da realização de experimentos novos, ou seja, que não serão encontrados em nenhum manual

## INTRODUÇÃO

de laboratório. Com isso, precisam obter dados que serão, em seguida, analisados para propiciar a solução do problema de modo confiável. Se você estiver pensando: “é só usar a lógica e o bom senso, afinal de contas, qualquer pessoa sensata pode parar e pensar um pouco”, está matriculado na disciplina certa, porque irá verificar como isso quase sempre esconde armadilhas. Pode ser que você gaste um tempo valioso, além de materiais valiosos, realizando inúmeros experimentos que não trarão as respostas que você precisa, mesmo agindo com a melhor das intenções. Usando os métodos adequados de planejamento experimental/tratamento de dados, você pode planejar um número mínimo de experimentos que sejam suficientes para fornecer as respostas necessárias, de forma confiável e segura.



Vamos começar esta aula propondo uma atividade. Imagine a seguinte situação: uma indústria fabrica um produto na forma de um pó, pela adição de uma base a uma solução da matéria-prima, seguida de aquecimento por um período de tempo. Sabe-se que o produto adquire um valor elevado se todos os grãos se formarem com o mesmo tamanho. Você recebeu a tarefa de descobrir qual a melhor condição experimental (concentração de matéria-prima, tipo da base, temperatura e tempo) que leve à formação do pó com partículas com a mínima variação de tamanho.

**EXPERIMENTOS****ATIVIDADES**

Pense sobre este problema e planeje uma série de experimentos.

**COMENTÁRIO SOBRE AS ATIVIDADES**

A tendência dos experimentalistas que desconhecem os métodos quimiométricos, em uma situação destas, seria estudar separadamente qual a melhor condição para cada variável. Para isto, iriam fixar todas as outras variáveis no valor que o senso comum diz ser o melhor e variar aquela que se tem mais dúvida. Por exemplo, estudariam qual a melhor base ( $\text{NaOH}$  ou  $\text{NH}_4\text{OH}$ ) para uma certa concentração da matéria-prima à temperatura ambiente (menos custo para a indústria) e por um tempo longo (aparentemente mais garantido para o término da reação). A base que fosse melhor nestas condições seria escolhida como sendo a melhor em todas as condições possíveis,

descartando-se a outra. Em seguida, tentariam descobrir qual a melhor concentração de matéria-prima, usando a base já escolhida, temperatura ambiente e tempo longo. A investigação prosseguiria com base neste mesmo método até que se chegasse a alguma conclusão a respeito da melhor condição possível, pois se variássemos mais de uma variável ao mesmo tempo, seria impossível chegar a alguma conclusão, pois como iríamos saber qual delas foi responsável pelo resultado obtido? Além disso, para variarmos tudo ao mesmo tempo, fazendo todas as combinações possíveis, seríamos obrigados a fazer um número infinito de experimentos, uma vez que algumas das variáveis podem assumir valores contínuos. Por fim, ao ter em mãos uma quantidade tão grande de dados, não seria nada prático espalhar os dados sobre a mesa e realizar um procedimento semelhante ao “jogo dos sete erros”. Para quem não se lembra, é aquela brincadeira que vem em revistas infantis, em que duas figuras aparentemente idênticas apresentam diferenças sutis (ou nem tanto) e o leitor é convidado a procurar as sete diferenças entre ambas.

---

Este é um exemplo de como o senso comum pode ser capaz de nos induzir a erro. Por incrível que pareça, o procedimento correto é variar todas as variáveis ao mesmo tempo, mas sem a necessidade de usar muitos valores diferentes para cada uma: é isso o que reduz o número de experimentos. Outro erro de se variar uma variável de cada vez é que, com isto, parte-se do pressuposto de que as variáveis não se influenciam mutuamente, ou seja, que o valor ótimo de uma não depende do valor da(s) outra(s). Pode ser que para um tempo curto, a temperatura alta dê bons resultados, mas, para tempos longos, temperaturas baixas dêem resultados ainda melhores. Ou seja, você havia chega-

do a um resultado aparentemente bom, mas deixou de descobrir qual o melhor.

É necessário deixar claro, contudo, que nada substitui o conhecimento técnico a respeito do problema. Segundo BOX, HUNTER E HUNTER: ...”É possível para um pesquisador fazer uma investigação sem usar estatística, e é impossível para um estatístico fazer o mesmo sem algum conhecimento de pesquisa. No entanto, um bom pesquisador torna-se um pesquisador muito melhor se usar métodos estatísticos.”

Para cada tipo de problema há um método específico de planejamento experimental e de análise dos dados. Por exemplo, nosso problema pode ser a otimização de um rendimento, ou descobrir qual matéria-prima está prejudicando a qualidade de um produto, entre outros. Podem parecer problemas cuja solução poderá necessitar de um número imenso de testes altamente específicos a cada caso, mas o uso dos métodos de planejamento experimental e tratamento de dados permite solucionar com o uso de técnicas já bem estabelecidas e eficientes. Você só precisa conhecer os métodos existentes e saber escolher o mais adequado a cada caso.

Os métodos mais comuns são: i) planejamentos fatoriais completos ou fracionários (aplicam-se à triagem de variáveis e avaliação da influência de cada uma); ii) modelagem por mínimos quadrados (aplica-se à construção de modelos empíricos). Há outros métodos mais específicos, que não abordaremos neste curso, tais como: superfície de resposta e simplex, que podem ser encontrados na bibliografia sugerida.

Sempre que vamos tratar dados experimentais e tentamos descrever e explicar seu comportamento, este procedimento é



chamado de modelagem. Estaremos construindo um modelo empírico, com descrição baseada nos dados experimentais. Por exemplo, observamos que o rendimento de uma reação aumenta quando aumentamos a temperatura. Construir um modelo significa descobrir, por exemplo, que o rendimento aumenta de forma linear com a temperatura, e chegar equação da reta que se aplica ao caso.

É importante neste ponto fazer uma distinção entre modelos empíricos e modelos mecanísticos. Os modelos de que se ocupa a estatística de dados experimentais são os modelos empíricos, que se baseiam inteiramente em dados experimentais e se limitam a descrever o comportamento do sistema. Por outro lado, existem os modelos mecanísticos, que se baseiam em



explicações para o comportamento observado nas teorias existentes e na possibilidade de prever o comportamento de outros sistemas. Por exemplo, podemos descobrir empiricamente como o rendimento de uma reação aumenta com a variação da temperatura, propondo inclusive uma equação que relaciona estas duas variáveis: este é um modelo empírico. Por outro lado, se partir-

mos para elaborar uma explicação de como a reação ocorre, quais ligações são rompidas, quais são as substâncias intermediárias formadas durante a reação e que se transformam no produto final, explicando, inclusive, porque a temperatura é importante, estamos elaborando um modelo mecanístico.

Outra questão fundamental é que nossos modelos precisam ser confiáveis e, neste caso, poderemos com eles prever o que vai ocorrer quando fizermos experimentos análogos. Isto é muito vantajoso para um experimentalista, mas só haverá garantia

de vantagens se o modelo for confiável e, novamente, a estatística tiver ferramentas para avaliação dos modelos. A boa notícia é que nenhum experimentalista precisa se transformar em um especialista em estatística, pois o uso de algumas noções básicas de planejamento experimental e análise de dados é suficiente. Este curso se propõe, portanto, a fornecer estas noções básicas suficientes para que um experimentalista possa realizar seu trabalho de maneira prática e confiável, minimizando os riscos de se tirar conclusões incorretas.

---

Muitas vezes, planejar, realizar experimentos e tirar conclusões a respeito dos dados obtidos esconde armadilhas, tais como, ao final de um número grande de experimentos, concluir que foram inúteis ou, o que é pior: tirar muitas conclusões errôneas e assinar embaixo. Isto se torna particularmente perigoso quando consideramos a tendência natural do ser humano a criar expectativas acerca de um resultado, que vai ao encontro de nossas pré-concepções. Os métodos quimiométricos, neste contexto, são ferramentas valiosas para que os experimentos sejam planejados adequadamente e, após a sua realização dentro do rigor científico, os resultados sejam manipulados e interpretados de maneira correta.

## CONCLUSÃO

## RESUMO



Faz parte da rotina diária de engenheiros, de químicos e outros profissionais que atuam nas indústrias e em laboratórios de pesquisa ter que encontrar as melhores condições experimentais para um dado processo ou mesmo descobrir uma variável que está causando problema. O chamado senso comum poderia nos levar a considerar que basta “mexer” em uma variável de cada vez e ver o que acontece. Por exemplo, se fosse na cozinha: mantemos o tempo de aquecimento de um bolo fixo e vamos testando diferentes temperaturas até ver qual “dá mais certo”. Depois, já sabendo qual a melhor temperatura, vemos se dá para diminuir o tempo. Afinal, tempo é dinheiro, certo? ERRADO. Pode ser que você gaste um tempo valioso, além de materiais valiosos, realizando inúmeros experimentos que não trarão as respostas de que você precisa, mesmo agindo com a melhor das intenções. Usando os métodos adequados de planejamento experimental/tratamento de dados, você pode planejar um número mínimo de experimentos que sejam suficientes para fornecer as respostas necessárias, de forma confiável e segura.

Usando o exemplo do bolo, o uso de métodos estatísticos pode nos levar a concluir que o melhor resultado vem da combinação de um tempo médio com uma temperatura média... Logo, em se tratando de um problema profissional, na indústria, por exemplo, você poderia ter corrido o risco de assinar um laudo errado, arriscando sua carreira. Poderia até ser acusado de má-fé ou de fraude. Nada disso está sendo dito para assustar você mas, desde cedo, é importante que um aluno de curso experimental tenha consciência das responsabilidades envolvidas com esta atividade. Ao sair da universidade, você será um profissional que terá que responder por possíveis conseqüências de sua atividade. Neste sentido, o uso de métodos estatísticos pode assegurar a confiabilidade dos trabalhos.

## QUIMIOMETRIA NO BRASIL

A importância prática da Quimiometria nas atividades industriais e de pesquisa foi levantada no excelente artigo *25 Anos de Quimiometria no Brasil*, publicado na revista *Química Nova*, 29 (2006) 1401, de autoria dos pesquisadores B. B. Neto, I. E. Scarminio e R. E. Bruns. Veja um trecho retirado do artigo:

*...Com o advento da quimiometria, o uso de experimentos estatisticamente planejados cresceu rapidamente em vários campos de pesquisa no Brasil, especialmente Química, Engenharia Química, Engenharia de Alimentos e Biotecnologia. Além disto, a utilização dos métodos quimiométricos de planejamento e análise não ficou restrita às atividades acadêmicas. Indústrias como Oxiteno, Pirelli, Braskem, Clariant, Nitroquímica, 3M do Brasil, Unilever, Petrobrás, Petroflex e Masterfoods do Brasil têm empregado – e continuam empregando – técnicas de planejamento e otimização de experimentos.*

*Uma busca no ISI, o Instituto de Informação Científica, mostrou que, dos 4099 trabalhos científicos localizados com a palavra-chave “factorial design”, 225 (isto é, 5,5%) tiveram participação de autor(es) brasileiro(s). Dos 3617 trabalhos com a palavra-chave “response surface”, 137 (3,8%) foram feitos, pelo menos parcialmente, no Brasil. Estas proporções são superiores à participação geral do Brasil na produção científica mundial, que fica entre 1 e 2%...*

## PRÓXIMA AULA



Na próxima aula, daremos início ao estudo propriamente dito da Quimiometria, destacando os parâmetros estatísticos mais relevantes dentro deste contexto.

---

## REFERÊNCIAS

- BARROS NETO, B.; SCARMINIO, I. E.; BRUNS, R. E.; **Planejamento e otimização de experimentos**. Campinas: Editora da Unicamp, 1995.
- BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S.; **Statistics for experimenters. An introduction to design, data analysis and model building**. New York: Wiley, 1978.
- BUSSAB, W. O.; MORETIN, P.A. **Estatística básica**. São Paulo, Ed. Atual, 1985.