



Universidade Federal de Minas Gerais  
Instituto de Ciências Exatas  
Departamento de Estatística

---

## **Comparação de Dois Métodos de Análise Química Considerando a Modelagem Matemática da Precisão**

Márcio Veloso de Castilho



## AGRADECIMENTOS

O projeto para este trabalho foi uma idéia surgida lentamente, e jamais teria sido possível levá-lo a cabo sem a ajuda, o apoio, a compreensão e o estímulo de uma grande quantidade de pessoas. Foi um trabalho realizado simultaneamente a muitas responsabilidades inadiáveis, que competiram pelo tempo e pela dedicação necessária à sua conclusão com êxito. Entre estas estão as viagens a trabalho, projetos e reuniões de serviço, as tarefas (e alegrias) que envolvem ser Pai enquanto se estuda, as questões domésticas do dia-a-dia. Seria muito difícil citar aqui todos que, de alguma forma, contribuíram ou apoiaram neste trabalho. Portanto, agradeço de antemão a todos que porventura não sejam citados e tenham sido importantes nesta realização.

Agradeço em primeiro lugar a Deus, pelo dom da Vida, pela condução invisível, pela inspiração, por todas as forças que me ajudam, mesmo aquelas de que não me dei conta ainda.

A Elizabeth, pela compreensão, pelo apoio sempre, pelo companheirismo. A Ariel, pelo eterno carinho, e pela inspiração para as simulações do capítulo 4, que ocorreu nas madrugadas do ano 2000, enquanto eu andava pela sala com ela no colo, tentando acalmá-la nos momentos de cólica intestinal, quando eu não podia nem mesmo me sentar no sofá sem atrair uma crise de choro.

A meus pais, que sempre me mostraram o prazer e as vantagens do estudo. Pelo exemplo no estudo e interesse nas ciências matemáticas. Pelo despertar do interesse em aprender o Reto Pensar.

Aos professores Cibele Comini Cezar, Enrico Antônio Colósimo, Heleno Bolfarine, Mário de Castro. Graças a eles, este trabalho pôde ser levado a cabo com sucesso, o texto final melhorou em qualidade e objetividade de maneira a se tornar mais fácil de ler. Aos ensinamentos que esclareceram vários dos pontos importantes, e me levaram a uma compreensão mais profunda da estatística como ciência de investigação.

Aos amigos e mestres na normalização internacional. Em especial ao Sr. Jacques Marcel François Gillis, nosso guru; à Thásia Medeiros de Freitas, que primeiro me mostrou o interessante trabalho da química conjugada com a estatística nos testes interlaboratoriais; Ao Carlos Augusto Coutinho, pelo estímulo ao estudo da química analítica, e ao uso de



Universidade Federal de Minas Gerais  
Instituto de Ciências Exatas  
Departamento de Estatística

---

técnicas estatísticas associadas a ela. E pelos momentos divertidos em sua companhia, em diversos continentes, em tempos variados.

Este trabalho jamais poderia ter sido realizado com sucesso, não fosse o apoio contínuo de diversos profissionais da Companhia Vale do Rio Doce. Em particular, agradeço ao Luiz Paulo de Carvalho Serrano, meu primeiro amigo na empresa, e o que me apresentou os primeiros desafios profissionais. Ao Engenheiro Francisco Valadares Povia, que sempre me apoiou e estimulou a desenvolver trabalhos relevantes, e cujo apoio foi fundamental quando iniciei a trilhar o caminho da normalização internacional, que trouxe grande crescimento pessoal e profissional. Ao Arthur Napoleão de Souza Neto, que sempre me apresentou problemas interessantes para uso de técnicas estatísticas na indústria da mineração, além de ter sido um importante professor na técnica de redigir relatórios e textos especializados. Pelos ensinamentos sobre a Vida, da qual a profissão é apenas uma parte. Ao Tarcísio José da Silva, que apoiou o mestrado desde o início, mesmo nos momentos em que tive de me ausentar do trabalho para cumprir as exigências da escola. Ao Arnaldo Moreira Borges, um amigo inesperado, e que me fez voltar à universidade para terminar o trabalho, já quase completo.

Ao Luciano Manetta Martins Belém, que me auxiliou a colocar a apresentação na forma final, à altura dos mestres de todos os tempos, trazendo o equilíbrio final da ética e da estética.

Aos colegas e amigos que apoiaram e ajudaram ao longo desta caminhada. Em especial, agradeço aos colegas que compareceram à apresentação final do trabalho perante a Universidade, para aprovação final. Este gesto de carinho não será esquecido.

Muito obrigado.

## LISTA DE FÓRMULAS E SÍMBOLOS

$x_i$	resultados de análises químicas obtidos pelo método de referência
$y_i$	resultados de análises químicas obtidos pelo método de teste
$\mu_{x_i}$	verdadeiros valores de teor relacionados ao método de referência
$\mu_{y_i}$	verdadeiros valores de teor relacionados ao método de teste
$\varepsilon_i$	erros aleatórios associados às medidas $x_i$ , com variância $\sigma_{\varepsilon_i}^2 = \sigma_{x_i}^2$
$\delta_i$	erros aleatórios associados às medidas $y_i$ , com variância $\sigma_{\delta_i}^2 = \sigma_{y_i}^2$
$\alpha, \beta$	parâmetros da reta de regressão de $Y_i$ contra $X_i$ , estimados por $a$ e $b$
$\hat{\mu}_{x_i}, \hat{\mu}_{y_i}$	estimativas das verdadeiras concentrações, obtidas pela reta de regressão
$s_a, s_b$	estimativas dos erros padrão de regressão para as estimativas $a$ e $b$
$\lambda$	razão entre as variâncias $\sigma_{y_i}^2$ e $\sigma_{x_i}^2$
$S$	soma de quadrados que é minimizada para se obter estimativas dos parâmetros de regressão. É definido especificamente para cada modelo e técnica de regressão
$s$	erro padrão de regressão
$w_i$	pesos de regressão utilizados nos modelos com ponderação de erros
$\bar{x}_w, \bar{y}_w$	médias ponderadas obtidas para os dados $x_i$ e $y_i$ , aplicando-se os pesos $w_i$
$CV_R$	coeficiente de variação de reprodutibilidade
$C$	concentração expressa na forma de fração de massa (um número entre 0 e 1)
$\eta$	erro aleatório normal com variância $\sigma_{\eta}^2$
$F$	estatística de Fisher
$r_i$	resíduo de regressão ( $r_i = y_i - a - bx_i$ )
$r_i^*$	resíduo padronizado para regressão ponderada, $r_i^* = (y_i - a - bx_i)\sqrt{w_i}$
$\sigma_0, \kappa$	parâmetros das funções de precisão (desvio padrão de análise vs. teor)
$s_{0x}, k_x$	estimativas para $\sigma_0$ e $\kappa$ referentes ao método de referência ( $x_i$ )
$X_i, Y_i$	variáveis aleatórias, cujos valores observados são os valores $x_i$ e $y_i$ .
$S_e$	área da região de confiança elipsóide para os parâmetros da reta.



## LISTA DE TERMOS TÉCNICOS E ABREVIACÕES

IUPAC: “International Union of Pure and Applied Chemistry”. É uma organização internacional que define métodos de análise química e protocolos para comparação de resultados com especificações. Isto é feito através de normas técnicas.

ISO: “International Organization for Standardization”. Entidade internacional responsável pela elaboração de normas técnicas utilizadas em diversos segmentos da atividade econômica, com importância no mercado internacional. Produz normas de análise química, assim como normas de aplicações estatísticas e outras. Assumiu uma importância muito grande com a assinatura do Código de Normas do GATT, ao final da rodada de Tóquio, em 1980.

GATT: “General Agreement on Tariffs and Trade”. Grupo de negociações diplomáticas composto de vários países, e que trabalha em “rodadas” de negociação, onde se decidem acordos diversos de interesse no relacionamento entre os países participantes.

Código de Normas do GATT: Documento assinado pelos participantes do GATT em fins de 1979, que estabelece, dentre outras coisas, que o comércio internacional deverá utilizar de normas técnicas internacionais para dirimir possíveis conflitos entre parceiros comerciais. Ele estabelece formas de se evitar as barreiras técnicas ao comércio.

AOAC International: “Association of Official Analytical Chemists International”. Organização internacional que promove a elaboração de normas técnicas de análise química aplicadas a diversos setores, como a indústria de alimentos, análises usadas em diagnósticos médicos, fabricação de materiais de referência certificados para várias aplicações, promove testes de proficiência de laboratórios em níveis nacionais e internacionais.

CRM : “Certified Reference Materials”. Os materiais de referência certificados são substâncias produzidas por entidades de reconhecimento internacional, que possuem valores de certas características medidas informadas em certificado, onde também constam as incertezas de medição associadas a cada característica. São utilizados para se detectar se um método de análise quantitativa é exato, ou para estimar seu vício. Também são utilizados para calibração. O comitê da ISO responsável pela normalização dos assuntos referentes a CRMs é o ISO/REMCO (“Reference Materials Committee”).

Padrão Primário : “padrão (material) designado ou amplamente aceito como possuindo as mais altas qualidades metrológicas e cujo valor (de uma determinada característica) é aceito sem referência a outros padrões da mesma quantidade, dentro de um contexto específico”. Tradução do Guia ISO 30, equivalente à definição do VIM (Vocabulário Internacional de Metrologia).

**Incerteza de medição:** estimativa associada a um valor de medição, que caracteriza a faixa de valores dentro da qual o “verdadeiro valor” é assegurado estar, com um nível de confiança especificado.

**Precisão:** proximidade de concordância entre resultados de teste mutuamente independentes obtidos sob condições prescritas. É medida pela imprecisão. Por exemplo, pelo desvio padrão de repetibilidade ou pelo desvio padrão de reprodutibilidade, conforme o caso.

**Repetibilidade:** proximidade de concordância entre resultados de sucessivas medições da mesma característica, sujeitas às condições: mesmo procedimento de medição, mesmo observador (ou analista), mesmo instrumento de medição, usado nas mesmas condições, mesmo lugar, e em curto intervalo de tempo.

**Reprodutibilidade:** proximidade de concordância entre resultados de medição da mesma característica, onde as medições foram realizadas em condições diferentes, como: princípio ou método de medição, observador (ou analista), instrumento de medição, lugar, condições de uso, e tempo.

**Rastreabilidade:** propriedade do resultado de uma medida ou valor de um padrão que estabelece, dentro de uma incerteza especificada, que o valor pode ser relacionado a referências declaradas, usualmente padrões nacionais ou internacionais, através de uma cadeia ininterrupta de comparações (chamada cadeia de rastreabilidade).

**Acurácia:** proximidade de concordância entre o resultado de teste e o valor de referência aceito. Este termo, quando aplicado a um grupo de observações, descreve uma combinação de componentes aleatórios e um erro sistemático comum ou componente de vício.

**Exatidão:** proximidade de concordância entre o valor médio obtido de uma série de muitos resultados de teste e um valor de referência aceito. A medida de exatidão usualmente é expressa em termos de vício.

**Valor de referência aceito:** um valor que serve como uma referência de comum acordo para comparações, e que é obtido como:

- um valor teórico ou estabelecido, baseado em princípios científicos;
- um valor atribuído, baseado em trabalho experimental de alguma organização nacional ou internacional;
- um valor de consenso, baseado em trabalho experimental colaborativo sob os auspícios de um grupo de cientistas ou engenheiros.

**Diferença sistemática fixa:** diferença entre dois métodos de análise, que é independente da concentração da substância analisada. Quando a diferença é tomada em relação a um



padrão primário ou valor de referência aceito, a diferença sistemática fixa é chamada de vício fixo.

Diferença sistemática proporcional: diferença entre dois métodos de análise, que muda com a mudança da concentração da substância analisada. Quando a comparação é realizada com um padrão primário ou valor de referência aceito, esta medida é chamada de vício proporcional.

Limite de Detecção: é a menor quantidade de uma substância, que um determinado método de medição é capaz de distinguir do ruído experimental, com um nível especificado de confiança. Esta característica não está relacionada à precisão de medida, portanto é comum métodos que possuem baixos limites de detecção apresentarem alta variabilidade próximo a este valor. A quantidade que especifica o menor valor que se detecta com uma determinada precisão é o Limite de Quantificação.

Teste de Proficiência: o Guia ISO 43-1 define teste de proficiência como sendo a determinação do desempenho de um laboratório de teste por meio de comparações interlaboratoriais. Também pode prover informações a respeito de identificação de problemas e tomada de ações corretivas, pode ser uma fonte de garantia de qualidade analítica do laboratório perante seus clientes, pode ser usado para se determinar valores de referência aceitos em materiais de referência, além de outras.

As definições e termos aqui apresentados são baseadas em consenso internacional dentro das entidades citadas (IUPAC, ISO, AOAC). Para um maior aprofundamento nestas definições e conceitos, é recomendada a leitura dos Guias ISO 25, 30, 31, 32, 33, 34, 35, 36, 43-1 e 43-2, assim como as normas ISO 5725-1, ISO 3534-1 e o VIM (“International vocabulary of basic and general terms in metrology”). Estas publicações são apresentadas nas referências, ao final deste texto.

## SUMÁRIO

1. INTRODUÇÃO E OBJETIVOS	8
1.1 Introdução	8
1.2 Objetivos	12
2. REVISÃO BIBLIOGRÁFICA	14
2.1 Comparação entre resultados de dois métodos	14
2.2 Modelagem matemática da precisão	26
3. MÉTODOS ESTATÍSTICOS	37
3.1 Comparação de Resultados	40
3.1.1 Métodos de mínimos quadrados	40
3.1.1.1 Mínimos quadrados ordinários	40
3.1.1.2 Mínimos quadrados ponderados	41
3.1.1.2.1 Mínimos quadrados ponderados simples	42
3.1.1.2.2 Modificação de Williamson (1968) para o método de York (1966)	43
3.1.1.2.3 Mínimos quadrados bivariados de Riu e Rius (1996)	45
3.1.2 Métodos baseados em máxima verossimilhança	47
3.1.2.1 Método da razão de variâncias constante e conhecida	48
3.1.2.2 Método ponderado para coeficiente de variação constante	50
3.1.2.3 Relação funcional de máxima verossimilhança para variâncias não homogêneas	52
3.1.3 Métodos robustos	55
3.1.3.1 Mínimos quadrados aparados de Rousseeuw e Leroy (1987)	55
3.1.3.2 Regressão estrutural absoluta, de Feldmann (1992)	56
3.1.4 Método de Wald (1940)	59
4. APLICAÇÕES	62
4.1 Simulações	62
4.1.1 Avaliação de diversos fatores sobre o estimador de regressão. Estudo baseado na análise de cobre em minérios	63
4.1.1.1 Estudo das distâncias euclidianas entre as estimativas e os valores reais dos parâmetros	68
4.1.1.2 Estudo das áreas das elipses de confiança para as estimativas dos parâmetros	70
4.1.1.3 Alguns exemplos de efeito das variáveis de tratamento na regressão	71
4.1.2 Considerações sobre o tamanho de amostra e magnitude dos erros no nível de significância do teste, para vários estimadores	82
4.2 Exemplos da literatura	86
4.2.1 Dados de Pearson (1901) utilizando os pesos de York (1966)	86
4.2.2 Exemplos de Ripley e Thompson (1987)	88
4.3 Exemplos do autor	93



4.3.1	Análise de ouro utilizando abertura da amostra com água régia, extração por solvente e leitura por duas técnicas, absorção atômica em forno de grafite e emissão ótica em plasma de acoplamento indutivo _____	93
4.3.2	Análise de cloro em minérios de cobre utilizando espectrometria de fluorescência de raios X dispersivos em comprimento de onda. Dados de dois diferentes laboratórios _____	101
4.3.3	Comparação de dois métodos de análise de manganês em minérios de ferro por absorção atômica em chama _____	108
5.	CONCLUSÕES _____	113
6.	ANEXO Códigos-fonte dos programas de computador _____	116
7.	REFERÊNCIAS BIBLIOGRÁFICAS _____	126



## Capítulo 1

### Introdução e Objetivos

#### 1.1 INTRODUÇÃO

Os estudos de comparação de métodos em química analítica representam um assunto abrangente e amplamente discutido na literatura. Isto se deve ao fato de que decisões de grande responsabilidade são comumente tomadas com base em resultados de análise química.

Alguns exemplos podem esclarecer bem a importância desta questão.

Diagnósticos de doenças e conseqüente prescrição de remédios e tratamentos médicos são muitas vezes baseados em análises laboratoriais. Evidências de crimes podem advir de provas de laboratório, em que alguma tecnologia sofisticada é utilizada para desvendar a identidade do criminoso. Ensaios de laboratório, tanto químicos quanto físicos, são utilizados para determinação de constituintes dos materiais de construção em edificações e estradas. A avaliação econômica de negócios, na indústria, freqüentemente depende de ensaios nos produtos, em escala de laboratório ou em ensaios piloto. No caso especial da indústria mineral, a análise química é utilizada para o cálculo do potencial passível de exploração de um bem mineral, o que influi nos cálculos de investimentos necessários em equipamentos e infra-estrutura, assim como na previsão de lucro futuro do negócio.

O intercâmbio de bens de consumo e “commodities” no mercado internacional envolve uma parcela considerável de capital, em transações multilaterais regulamentadas por acordos, e que por sua vez muitas vezes envolvem o uso de métodos analíticos para garantia da qualidade. Um exemplo de acordo deste tipo é o Código de Normas do GATT (“General Agreement on Tariffs and Trade”), que favorece o uso de normas internacionais na avaliação de produtos de exportação. Alguns países, como os Estados Unidos da América, possuem entidades governamentais para controlar a qualidade de alimentos e medicamentos (no exemplo, o FDA – “Food and Drugs Administration”) produzidos internamente ou importados. Nestas situações é comum o uso de métodos de análise química *regulamentados* ou *normalizados*. Estes métodos são chamados em várias situações de *métodos de referência*, *métodos oficiais*, *métodos de arbitragem*, além de outras denominações.

Em qualquer das situações citadas como exemplo pode surgir a necessidade de comparar métodos. De um lado, quem toma a decisão a partir dos resultados de ensaios pode ter dúvidas quanto aos valores apresentados nos certificados ou laudos de análise. De outro, pode-se desejar publicar o método como um método oficial ou de referência, e para tanto é



exigido que este passe em uma série de avaliações, incluindo a comparação com os métodos considerados clássicos ou de referência existentes, nos quais a comunidade científica confia. Surgem então dois conceitos muito importantes dentro deste tipo de estudos: o da confiabilidade do método, que pressupõe a existência de meios para a sua validação.

Entende-se por confiabilidade do método analítico a sua capacidade em atender a uma série de exigências que garantam sua adequação ao uso em circunstâncias específicas. Dentre estas exigências estão aquelas usualmente demandadas em sistemas de garantia de qualidade, como a rastreabilidade a padrões primários, limite de detecção, limite de quantificação, registros adequados, precisão, exatidão, seletividade, robustez (no sentido da palavra inglesa “ruggedness”), linearidade, e outras. Em geral, para se estabelecer a confiabilidade de um método, é necessário se realizar sua validação.

Validação é um termo que se refere ao estabelecimento das características de qualidade do método, especificando bem o domínio em que ele é válido. Nestes estudos usualmente se estabelece os limites de aplicação (em termos de teores da substância a analisar), as equações de precisão necessárias ao seu uso, como por exemplo a repetibilidade e a reprodutibilidade. Também se estabelece em que faixa de concentrações o método pode ser considerado exato, além de se buscar também num estudo de validação especificar em que matrizes ele pode ser aplicado. Por exemplo, um método desenvolvido para determinação de íons de sódio em plasma sanguíneo não obrigatoriamente será útil para determinação dos mesmos íons em água.

Danzer et al.. (1995) apresentam um modelo matemático tridimensional para o estudo de métodos de análise, envolvendo, em suas projeções bidimensionais, três variações comuns em química analítica: calibração, medição (que os autores chamam de avaliação analítica), e recuperação. Como casos especiais deste último, citam os estudos de análises de amostras de controle (materiais de referência certificados) e de comparação com outro método de análise, considerado confiável e não viciado. O presente trabalho se concentra na comparação de métodos, e a comparação com resultados de CRMs, que são os dois últimos casos citados. Muitos dos resultados encontrados podem também ser aplicados nas outras variações apresentadas por estes autores. Onde pertinente, são feitas algumas considerações sobre estas outras aplicações.

Existem muitos métodos para o tratamento do problema de comparação de métodos. Altman e Bland (1997) propõem o uso de gráficos de diferença, com conseqüente tratamento dos resultados por um teste “t” de Student. Bolfarine, Rodrigues e Cordani (1992) apresentam abordagens bayesianas para utilização em problemas de regressão com erros nas variáveis. Técnicas desenvolvidas para o tratamento de modelos de regressão com erros nas variáveis poderiam ser adaptados a estes estudos. Existem muitos trabalhos que poderiam ser assim considerados. Schafer (1987) discute uma aplicação do algoritmo E-M



para erros nas variáveis em modelos lineares generalizados. Zeger e Karim (1991) apresentam uma solução de análise de modelos lineares generalizados por meio do amostrador de Gibbs. Whittemore e Keller (1988) utilizam estimativas de máxima verossimilhança e quase-verossimilhança, comparando com outros métodos, também no caso de regressão. Embora as três últimas citações feitas não sejam de propostas para estudos de comparação de métodos em química analítica, elas poderiam inspirar métodos alternativos de inferência para este caso.

O estudo completo do problema de comparação de métodos envolve a comparação dos *resultados* obtidos por dois métodos (e nesse caso se busca saber se há concordância, em média, entre os resultados obtidos por ambos), e a comparação entre as *precisões* de ambos. Embora não se encontre na literatura específica de química analítica propostas para o tratamento dos dois problemas em conjunto, é possível modelá-los desta maneira, utilizando o método de máxima verossimilhança, ou alternativamente, utilizando métodos bayesianos. A primeira destas fornece um sistema de equações não lineares em que os parâmetros a estimar não são explicitáveis (o que dificulta, portanto, a solução do sistema utilizando métodos iterativos), e não se consegue expressões “fechadas” para os estimadores. Seria necessário utilizar métodos numéricos de maximização para se encontrar as estimativas. A segunda opção, de se usar um método bayesiano, ainda é pouco conhecida dentro do meio analítico, e ambas as alternativas vão além do que se propõe no presente trabalho.

A comparação de métodos, na forma como é encontrada hoje, mesmo em comitês de normalização analítica internacionais, é dividida em duas etapas. Primeiramente se encontra as precisões dos dois métodos e se procede à sua avaliação. Só então se realizam comparações dos resultados encontrados pelos dois. Muitas vezes, mesmo a avaliação das precisões é colocada em segundo plano ou mesmo desconsiderada. Por isto, no presente trabalho se inclui esta etapa, já que tem importância fundamental na segunda instância, em que os resultados são testados em sua igualdade. Obviamente, esta igualdade se dá dentro de certa precisão, apesar de alguns autores não se preocuparem com isto.

Para o caso específico da literatura de química analítica e aplicações relacionadas, várias abordagens são discutidas no presente trabalho. Linnet (1990) cita uma revisão feita nas edições dos três primeiros meses de 1988 do periódico “Clinical Chemistry”, em que ele conta 50 artigos de comparação de métodos. Destes, 46 utilizam regressão linear simples por mínimos quadrados ordinários. Nos demais, é aplicado o método de Deming (1943).

Técnicas de regressão são mais frequentemente utilizadas neste tipo de estudos porque permitem a estimativa dos dois tipos de erros sistemáticos: fixo e proporcional. Entretanto, o uso de mínimos quadrados ordinários não é recomendado por causa de três aspectos que são características comuns dos métodos de análise química:



- Presença de erros nos dois métodos comparados;
- Estes erros não são constantes ao longo da faixa de concentrações analisadas (heteroscedasticidade);
- Presença de pontos discrepantes ou dados contaminados, comuns em situações práticas.

Desconsiderar estes aspectos, que são da natureza fundamental dos métodos de química analítica, leva a se obter estimativas viciadas dos parâmetros da reta, e as conclusões obtidas muitas vezes chegam a ser tiradas com erros superiores aos supostos.

Para tratar o primeiro aspecto enumerado (presença de erros nos dois métodos), grande quantidade de publicações é disponível, propondo métodos de regressão com erros nas variáveis. Riu e Rius (1995) apresentam uma extensa revisão destas propostas. Os métodos encontrados podem se subdividir em duas classes: os que consideram erros nas variáveis, porém constantes ou homogêneos em ambas (Altman e Bland, 1997; Hollis, 1996; Cornbleet e Gochman, 1979; Danzer et al., 1995; Hartmann et al., 1997; Petersen et al., 1997; Zwanziger e Sârbu, 1998 e outros); e os que consideram algum tipo de ponderação na regressão. Neste último caso se classificam os métodos propostos por Lisy et al. (1990), Riu e Rius (1996), Kalantar et al. (1995), Linnet (1990), Cecchi (1991), Neri et al. (1989), Ripley e Thompson (1987). Apenas os autores citados nesta segunda classe de propostas consideram a heteroscedasticidade, citado como segundo aspecto importante em comparação de métodos.

Em nenhuma destas publicações, entretanto, é utilizada alguma forma de modelagem matemática da precisão com a concentração. Ripley e Thompson (1987) e Linnet (1990) citam a modelagem funcional para a variância da análise química, mas no caso dos primeiros ela não é utilizada na abordagem sugerida. Eles partem do pressuposto de que as variâncias em cada ponto são conhecidas, e portanto não precisam possuir estrutura definida. Em Linnet (1990) esta modelagem é restrição do método proposto. Ambos comentam o modelo de relação linear entre o desvio padrão da análise e a concentração, que entretanto não explica esta característica adequadamente em todas as situações práticas.

Modelagens mais elaboradas do erro aleatório são vistas em outros autores, porém não para comparação de métodos, mas para calibração e para estudos de normalização e publicação de métodos. Contribuições importantes neste sentido são as de Horwitz (1982), Roche e Lorenzato (1995) e Thompson e Lowthian (1997).

Sendo assim, a precisão é considerada apenas para se definir fatores ponderadores de regressão. Trabalhos de comparação de métodos usualmente abordam a comparação dos *resultados* dos dois métodos, e não da precisão com que estes resultados são produzidos. Para isto seria necessário se dispor de modelos matemáticos que representem de forma aproximada a variação da precisão com a concentração, como encontrado na prática, e de



técnicas estatísticas para comparar estes modelos, através de inferência sobre seus parâmetros.

Para contornar o problema de dados contaminados, são propostos modelos de regressão robustos. Em aplicações de química analítica, os métodos mais citados são o da mínima mediana dos quadrados, de Rousseeuw e Leroy (1987), o da mínima mediana dos quadrados ortogonal, apresentada em Sarabia et al. (1997) e Hartmann et al. (1997), o método de Passing-Bablok (1983) e a regressão estrutural da mínima mediana e a regressão estrutural absoluta, as duas últimas propostas por Feldmann (1992).

Na revisão bibliográfica (capítulo 2) é apresentada uma breve discussão dos três tipos de métodos citados. Também são mostrados os modelos matemáticos usualmente utilizados para explicar a variação do desvio padrão do erro aleatório com a concentração.

No capítulo de métodos (capítulo 3), alguns dos procedimentos estatísticos citados na revisão são apresentados com maior detalhamento. Os procedimentos considerados para este trabalho são os de regressão. Estes métodos são posteriormente utilizados nos estudos de simulação, no capítulo de aplicações (capítulo 4). O propósito das simulações é o de comparar os diferentes estimadores de regressão propostos na literatura entre si, mesmo sabendo-se que alguns não são adequados para este tipo de comparações.

No capítulo de aplicações (capítulo 4) também são apresentados exemplos reais, alguns tirados da literatura, e outros referentes ao trabalho do autor como responsável pelo desenvolvimento e avaliação de métodos de análise química aplicados à indústria da mineração.

No capítulo 5 são apresentadas as conclusões do trabalho. No Anexo (capítulo 6), são apresentados os códigos-fonte dos programas de computador desenvolvidos para aplicação dos métodos descritos no capítulo 3 e os utilizados para as simulações e aplicações práticas do capítulo 4.

## 1.2. OBJETIVOS

O objetivo geral deste trabalho é o de fornecer uma sistemática que permita a comparação de dois métodos de análise química, em termos de concordância dos valores médios obtidos por ambos, assim como em termos de igualdade da precisão analítica nos dois.

Os passos tomados para se atingir este objetivo são:

- a) apresentar uma revisão dos métodos de regressão com erros nas variáveis, aplicados à comparação de métodos de análise química;



Universidade Federal de Minas Gerais  
Instituto de Ciências Exatas  
Departamento de Estatística

---

- b) apresentar uma revisão dos modelos que representam a variação do desvio padrão do erro aleatório (que mede a precisão analítica) com a concentração da substância analisada;
- c) comparar os métodos de regressão utilizados e propostos na literatura, através de simulações computacionais;
- d) finalizando, propor um método de regressão, que considere a modelagem do erro aleatório, baseado nos resultados da simulação e fundamentados teoricamente, apresentando exemplos práticos de aplicação.

## Capítulo 2

### Revisão Bibliográfica

Estudos de comparação de métodos em análise química têm por base duas seqüências de dados. Dispõe-se de um número  $n$  de amostras (por exemplo, provenientes de coleta de solos em regiões diferentes, de minérios provenientes de uma mesma ocorrência geológica, água, sangue, ou de alimentos de um mesmo tipo). Cada uma destas amostras é dividida em duas partes, que são submetidas a análise por dois diferentes procedimentos químicos, ou diferentes instrumentos, ou diferentes laboratórios. Os resultados assim obtidos são representados por dois vetores, por exemplo,  $X$  e  $Y$ .

#### 2.1 COMPARAÇÃO ENTRE RESULTADOS DE DOIS MÉTODOS

Na área de química clínica os estudos de comparação de métodos ganharam ao longo do tempo grande importância, que pode ser evidenciada pela profusão de publicações a respeito. Roy Barnett (1965) publicou um dos primeiros guias para orientar neste tipo de estudos. Este trabalho contou com a assistência do Dr. W. J. Youden, estatístico do antigo “National Bureau of Standards”. Nesta primeira proposta, e na que a sucede (Barnett e Youden, 1970), são considerados apenas três níveis de concentração em cada método, e a técnica de teste utilizada é um teste “ $t$ ” de Student para dados emparelhados. Apesar de sugerirem a apresentação dos dados em gráficos onde se poderia evidenciar uma relação linear entre os dois vetores comparados, não fazem referência a nenhum teste de hipótese para se confirmar os valores dos parâmetros. Na verdade, não propõem que se estime estes parâmetros. O gráfico seria apenas para visualização dos dados. No artigo de 1970 os autores sugerem que uma regressão utilizando o método de mínimos quadrados (ordinários) pode ser utilizada, mas apenas para se visualizar no gráfico se a reta obtida parece diferir da reta  $Y=X$ .

Altman e Bland (1983) propõem uma metodologia parecida, porém com destaque para a representação gráfica, utilizando uma figura onde se cruzam valores das diferenças entre métodos (respeitando os sinais) com as médias. Eles chamam a esta técnica “Gráficos de Diferença”. Hollis (1996) apresenta detalhadamente os cálculos necessários para se utilizar a proposta de Altman e Bland (1983), que na verdade é um teste “ $t$ ” de Student para dados emparelhados. Esta última autora não recomenda o uso de análise de regressão, supostamente devido às limitações da técnica, porém tais limitações se referem ao método de mínimos quadrados ordinários. Parece que ela, assim como outros autores, confunde regressão com método de mínimos quadrados ordinários.

Stöckl (1996) salienta que o importante do método “Gráfico de Diferenças” não é a representação gráfica, mas os métodos de inferência utilizados, o mesmo valendo para



análise de regressão. O que este e outros autores não citam, é que este tipo de métodos somente é válido se a variância é aproximadamente constante ao longo da faixa de concentrações estudada.

Infelizmente, o teste “t” de Student, técnica advogada por Altman e Bland (1983), Hollis (1996), Stöckl (1966), e reforçada por Petersen et al. (1997) não permite testar a existência de erros sistemáticos proporcionais, além de não tratar adequadamente a heteroscedasticidade presente em resultados de análise química.

Westgard e Hunt (1973) fazem uma avaliação por simulação para apontar as vantagens e desvantagens dos métodos até então comumente utilizados na área clínica. Em resumo, mostram a inadequação dos métodos de teste utilizando as estatísticas “t” de Student, F de Fisher para a razão de variâncias, o coeficiente de correlação de Pearson, e o método de mínimos quadrados ordinários. As simulações por eles realizadas não utilizam números aleatórios, mas partem de uma série de dados onde são sistematicamente introduzidos erros, conforme as situações que se deseja avaliar. Estes autores, após este estudo, recomendam o uso da análise de regressão. Neste trabalho, eles utilizaram o estimador de mínimos quadrados ordinários.

Entretanto, para a utilização da análise de regressão, é necessário se conhecer que tipo de dados estão disponíveis, pois existem inúmeras alternativas para se encontrar uma reta que passa pelos pontos experimentais.

Muitas das abordagens encontradas na literatura técnica do assunto são baseadas em regressão. Neste caso, a equivalência entre métodos é usualmente aceita se, após traçar-se uma reta de regressão que passe pelos pontos (X,Y), obtém-se um coeficiente angular próximo de um e uma intercessão próxima de zero. Várias abordagens são sugeridas para se obter as estimativas dos parâmetros desta reta, e o assunto é fonte de pesquisa há muito tempo. Referências de abordagem deste problema são Deming (1943), Youden (1955), Mandel (1964), Caulcutt e Boddy (1983), e Massart et al. (1988). Ripley e Thompson (1987) citam uma proposta de Adcock, encontrada em uma edição de 1878 da revista “The Analyst”, publicada pela “Royal Society of Chemistry”. Nesta proposta, é sugerida uma técnica de regressão ortogonal, onde se minimiza a soma dos quadrados das distâncias ortogonais dos pontos observados à reta de regressão, em contraposição ao método de mínimos quadrados ordinários, que minimiza a soma de quadrados das distâncias verticais.

A abordagem por regressão é geralmente favorecida entre diversos autores, como por exemplo Penninckx (1994), que sugere a utilização do estimador de mínimos quadrados ordinários, devido à possibilidade de se conhecer os dois componentes de erro sistemático, ou de vício, se for o caso: fixo e proporcional. Na verdade estes componentes podem ser estimados utilizando-se qualquer procedimento de regressão. O erro fixo é estimado pela

intercessão da reta de regressão e o proporcional pela inclinação. Os casos que se encontram podem ser resumidos na Figura 2.1.

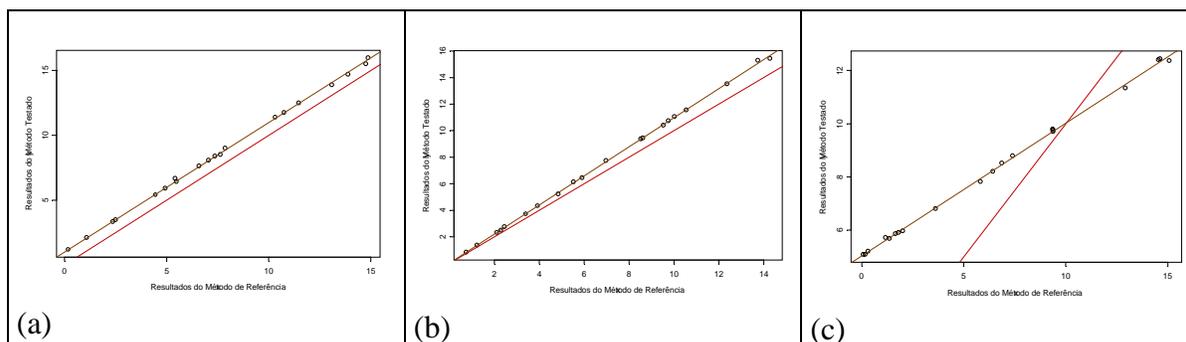


Figura 2.1: Exemplos de situações com presença de diferentes tipos de erros sistemáticos.

(a) erro sistemático fixo entre os dois métodos; (b) erro sistemático proporcional entre os dois métodos; (c) presença de erros sistemáticos, tanto fixos quanto proporcionais entre métodos. A linha contínua de referência representa a igualdade entre os dois métodos. Representa a reta  $Y=X$ .

Este autor alerta para o fato de que quando os erros de medida dos dois métodos forem da mesma ordem de grandeza, algum método de regressão que considere erros nos dois eixos deve ser utilizado.

Quando se sabe que a variância não muda ao longo da faixa de concentrações do método de análises, são sugeridas várias alternativas para comparar dois métodos. Esta suposição, embora não seja em geral verdadeira, pode ser aceita em algumas situações. Normalmente, quando a faixa analítica (intervalo de concentrações onde os métodos são válidos) é pequena, esta suposição é aproximadamente verdadeira.

Linnet (1993) cita um exemplo disto. Segundo este autor, a concentração de eletrólitos em soro sanguíneo é regulada rigidamente em organismos saudáveis, e desvios associados com doenças são pequenos a moderados. Para o teor de sódio, por exemplo, foi encontrado por Cornbleet e Gochman (1979) uma razão de 1,15 entre o maior valor comparado e o menor, usando o método de referência. Para faixas assim estreitas, a variância é aproximadamente constante em ambos os métodos comparados, e pode-se utilizar desta suposição no método estatístico.

Um dos estimadores mais comuns, utilizado nestas situações, é o de mínimos quadrados ordinários. Esta técnica, embora não considere o erro nas medidas do método de referência, é considerada no presente estudo, devido à sua popularidade.

O método de mínimos quadrados ordinários é apresentado com detalhes em Seber (1977), Draper e Smith (1981), e Montgomery e Peck (1992). No Capítulo 3 são apresentados os resultados mais importantes desta técnica para o estudo aqui proposto.



Zwanziger e Sârbu (1998) propõem um método de regressão que considera que a variância é constante ao longo da faixa de concentrações estudada, sem ignorar o fato de haver erro nas duas variáveis, e testam a razão entre os valores absolutos dos coeficientes A e B da reta  $Ax + By + C = 0$  através de uma técnica que estes autores chamam de análise de variância informacional. A hipótese básica neste problema é a de que a razão  $|A/B| = 1$ . No exemplo ilustrativo não fica clara qual a vantagem deste método em relação a outros, citados pelos autores, e usados como comparação. Os autores apenas citam que a análise de variância informacional utiliza de suposições mínimas a respeito da distribuição dos dados, e independe da faixa de variação dos mesmos.

Danzer et al. (1995) apresentam uma série de alternativas para se encontrar os parâmetros da reta de regressão, todas incluídas na descrição genérica de regressão ortogonal. O princípio básico dos métodos ortogonais é que a distância a ser minimizada (na verdade a soma dos quadrados desta distância é minimizada) é a distância ortogonal dos pontos à reta. Os autores contrapõem o método de mínimos quadrados ordinários (por eles denominado mínimos quadrados gaussianos) com propostas de Wald (1940), Mandel (1984), um método descrito pelos mesmos autores designado método da Média Geométrica, e também usam a técnica de Componentes Principais.

O método de Wald se assemelha ao método de três grupos de Bartlett (1949). Ambos particionam os dados a partir dos resultados ordenados da variável  $x$ . Bartlett propõe a partição em três grupos onde pelo menos o primeiro e terceiro devem possuir o mesmo tamanho. Wald divide os dados também ordenados, mas em dois grupos. O coeficiente angular em ambos os casos é calculado pela razão das diferenças entre as médias de  $y$  (maior média menos a menor média) e de  $x$ . Riu e Rius (1995) afirmam que o método de Bartlett tem importância apenas histórica, e que leva a grandes erros de estimativa em certos grupos de dados.

O método descrito como sendo o método de Mandel é uma variação do que Mandel descreve em seu livro (Mandel, 1964) como sendo o método de Deming. No caso, a variação citada é o fato de se considerar a razão de variâncias igual a um. Este método é apresentado em detalhe no capítulo de métodos, no presente estudo.

O método da média geométrica se resume em obter as inclinações por mínimos quadrados ordinários considerando  $y$  como função de  $x$ , e depois de  $x$  como função de  $y$ . Estas inclinações, representadas como  $b_x$  e  $b_y$ , são utilizadas para cálculo da média geométrica destes coeficientes.

O método de componentes principais produz dois componentes, o primeiro fornece uma aproximação para a linha de regressão buscada, e o segundo representa o espalhamento ortogonal dos dados ao redor da linha reta.



Em todos os casos citados acima a intercessão da reta é encontrada ao se utilizar do fato de que a reta passa pelo ponto representado pela média em  $y$  e em  $x$ , o ponto  $(\bar{x}, \bar{y})$ .

Danzer et al. (1995) apresentam um estudo de simulação utilizando o estimador de Wald, o de Mandel, o da Média Geométrica, o que utiliza Componentes Principais e o método de Mínimos Quadrados Ordinários. As simulações se baseiam em experimentos de calibração analítica, utilizando cinco pontos, com três réplicas de cada. Os erros aleatórios são primeiramente assumidos como apresentando variância constante e igual para  $X$  e  $Y$ . Numa segunda instância, faz-se variar a relação entre as variâncias, para se testar qual estimador apresentaria melhor desempenho. Para todas as simulações foi utilizada inclinação unitária para a reta teórica, o que significa considerar que não há desvios sistemáticos entre os métodos de análise química. Os melhores resultados foram obtidos com o estimador proposto por Wald.

O método mais citado na literatura de química analítica para a comparação de métodos, quando se deseja utilizar regressão com erros nas variáveis, portanto evitando o uso de mínimos quadrados ordinários, é o chamado método de Deming.

Em todas as publicações onde é citado, a referência dada é Deming (1943). Paiva Lima (1996) trabalha com este método sem fazer referência a Deming, e dá a ele a denominação de método clássico de modelo funcional de máxima verossimilhança com razão de variâncias conhecida. Paiva Lima (1996) deduz este estimador através do método de máxima verossimilhança, quando se considera as variâncias nos dois eixos constante. Reescrevendo a expressão da verossimilhança, esta autora mostra que não é necessário conhecer as duas variâncias separadamente, basta conhecer a razão entre elas, e torna-se possível a estimação dos parâmetros. No capítulo de métodos são apresentadas as expressões para se obter este estimador.

As referências mais comuns na literatura de química analítica para este método são Mandel (1964) e Cornbleet e Gochman (1979). Danzer et al. (1995), como citado anteriormente, e Hartmann et al. (1997) o chamam de método de Mandel, devido a uma publicação deste (Mandel, 1984), onde é apresentada uma maneira diferente de se obter as estimativas dos parâmetros e respectivos erros padrão, por meio de uma transformação de variáveis. Mandel faz uma rotação e translação de eixos nas variáveis originais, para obter um novo sistema coordenado onde os pressupostos do método de mínimos quadrados ordinários são aproximadamente válidos. Muitos autores citam o método de Mandel como sendo um caso especial do método de Deming, quando se considera a razão de variâncias constante e igual a um. Entretanto, Mandel (1984) deriva equações genéricas que podem ser utilizadas quando as variâncias forem constantes e diferentes. As fórmulas apresentadas por Paiva Lima (1996), Mandel (1964) e Cornbleet e Gochman (1979), para estimar os parâmetros da reta, são equivalentes.



Embora se saiba que em muitas situações os métodos de regressão que consideram variância constante são teoricamente injustificados, Cornbleet e Gochman (1979) mostram que o método de Deming pode ser aplicado, se o coeficiente de variação nos dois métodos for inferior a 20%, o que é freqüente em química analítica. Em suas simulações, estes autores exploram situações com dados seguindo distribuições normais e log-normais, variâncias constantes e coeficientes de variação constantes. Em todos os casos, o erro sistemático entre os métodos comparados é expresso pela equação  $y = 10 + 0,90x$ , ou seja, método  $y$  com diferença fixa de 10 unidades acima do método  $x$ , e erro proporcional de 10% para baixo do mesmo. São comparados os métodos de mínimos quadrados ordinários, o de Deming, o de Mandel (considerando razão de variâncias unitária) e o de Bartlett. Ao final, os autores recomendam o uso do método de Deming.

O caso mais genérico, quando se considera a comparação de métodos em química analítica, é aquele em que são considerados erros nos dois métodos, e a variância destes erros depende do valor de concentração da substância analisada. Muitos autores consideram, apesar disto, como satisfatória a utilização do método de mínimos quadrados ponderados. Christensen (1996), por exemplo, advoga o uso desta técnica para comparar resultados de um novo método com resultados de CRMs. O problema real nesta alternativa é que despreza os erros em  $x$ , introduzindo, portanto, vício nas estimativas.

Existem situações, entretanto, em que o estimador de mínimos quadrados ponderados é teoricamente correto, quando se sabe que os erros de medida no método  $x$  são bem menores que os do método  $y$ . Isto acontece em algumas situações de calibração, embora não se possa dizer que é uma suposição sempre válida mesmo então. Trabalhos propondo esta técnica para calibração podem ser vistos em Agterdenbos (1979), Schwartz (1979), Garden et al. (1980), e Watters, Carroll e Spiegelman (1987).

Quando a suposição de que os erros em  $x$  são muito menores que os erros em  $y$  não puder ser aceita como verdadeira (que é o mais comum no caso de comparação de métodos), torna-se necessário utilizar métodos de regressão ponderada com erros nas variáveis.

Uma introdução a estes métodos, aplicados a análises químicas, com uma extensa pesquisa bibliográfica, é apresentada em Riu e Rius (1995). Os métodos melhor avaliados por estes autores partem de uma extensão do critério de mínimos quadrados ponderados, ao se considerar também os erros em  $x$ , sendo os ponderadores funções das duas variáveis. As estimativas são produzidas então pela minimização deste critério em relação aos parâmetros.

Linnet (1990) propõe uma variação do método de Deming para situações onde o erro pode ser expresso como função linear da concentração (com intercessão nula). Isto equivale a dizer que o coeficiente de variação é constante, e que o limite de detecção é zero. O método parte da minimização de uma expressão similar à do método de Deming, substituindo a



razão de variâncias por uma constante que equivale a uma razão entre coeficientes de variação elevados ao quadrado. Desta forma, as fórmulas de cálculos obtidas são equivalentes às do método de razão de variâncias constante.

O método, como apresentado por Linnet, é especificado para o caso em que as comparações são feitas com dados em duplicata de cada método para cada amostra. Torna-se necessário adaptá-lo para um caso mais geral, por exemplo quando se deseja utilizar estimativas funcionais para os desvios padrão de medida em cada método, ou para estimativas obtidas com mais que duas réplicas por amostra.

Para estimar a razão entre os coeficientes de variação elevados ao quadrado, o autor propõe um método iterativo, que é testado por meio de simulações. Estes testes mostram que a estimação dos pesos de regressão (que dependem dos valores “verdadeiros” de teor encontrados pelos dois métodos) afeta a razão de coeficientes de variação, que por sua vez afeta a estimativa da inclinação da reta, introduzindo vício nesta última. Ao realizar estas estimativas, e utilizar a reta encontrada para atualizar os valores de concentração e assim estimar novamente os pesos, após algumas iterações são produzidas estimativas não viciadas dos parâmetros.

Em outro artigo (Linnet, 1993), ele utiliza simulações para comparar os dois estimadores de Deming com os métodos de mínimos quadrados ordinários, mínimos quadrados ponderados, e um método não paramétrico, proposto por Passing e Bablok (1983). São feitas três simulações, baseadas em problemas reais de química clínica. Na primeira os erros possuem variância constante, e a faixa de variação das concentrações é pequena. Os métodos comparados são o de mínimos quadrados ordinários, o método de Deming e o método não paramétrico de Passing e Bablok (1983) neste exemplo. Os resultados mostram que apenas o método de Deming fornece estimativas não viciadas para os parâmetros da reta e para o erro quadrático médio. O número de rejeições neste método está próximo de 5%, conforme especificado para o teste de hipóteses. Os outros dois métodos apresentam um grau de rejeição de aproximadamente 20%, mesmo usando um nível de significância de 5% para os testes. Na segunda simulação os erros possuem desvio padrão proporcional à concentração, e os resultados mostram que as cinco técnicas de regressão apresentam estimativas dos parâmetros com vício muito pequeno. Entretanto, os testes de hipótese para a igualdade entre os métodos de análise química apresentam rejeição de 5% apenas para o método ponderado de Deming. Para os demais, a rejeição é de no mínimo 7%, podendo chegar a 22% no caso de se utilizar mínimos quadrados ordinários. Na terceira simulação os dados são análogos aos do segundo caso, porém com inclusão de dados contaminados. Para esta situação, o autor propõe um método de detecção e eliminação de dados discrepantes para se proceder aos cálculos de estimação de parâmetros. Realizando esta eliminação de dados “outliers”, o método ponderado de Deming produz os menores erros quadráticos médios, sendo que os vícios de estimativa não são problema nos métodos usados (não paramétrico e ponderado de Deming).



Em outro artigo (Linnet, 1998) ele volta a usar de simulações, desta vez para comparar os dois métodos de Deming por ele advogados e o método de mínimos quadrados ordinários. Nestas simulações o autor se preocupa com a especificação da razão de variâncias, que é supostamente conhecida para aplicação do método de regressão. O autor comenta que embora o método seja derivado considerando esta suposição, muitas vezes os pesquisadores utilizam razão unitária, por não disporem de tal informação. Nestas situações, ele mostra que podem surgir vícios grandes nas estimativas, principalmente quando a faixa de concentrações avaliadas for pequena.

Thompson (1982), propõe um método por ele intitulado de regressão de máxima verossimilhança. Entretanto, o estimador de Thompson não é realmente o de máxima verossimilhança. Este autor propõe, posteriormente, juntamente com o Professor Brian Ripley, da Universidade de Strathclyde, Glasgow, um método de máxima verossimilhança. Este é encontrado em Ripley e Thompson (1987).

No artigo de 1982, Thompson realiza várias simulações de Monte Carlo para testar o estimador por ele apresentado, comparando-o com os de mínimos quadrados ordinários e mínimos quadrados ponderados. Os dados simulam diferentes faixas de concentração, com os valores distribuídos uniformemente, incluindo diferentes graus de vício fixo e proporcional no segundo método. Os erros aleatórios são modelados de forma linear com a concentração. Para atribuir estes erros aleatórios nas duas variáveis comparadas, são feitas combinações para se obter desde erros compatíveis nos dois, até erros bem diferentes. O autor também testa desvios da linearidade, introduzindo curvaturas variáveis nos dados, além de avaliar a influência do número de pontos da regressão nas estimativas de erros dos coeficientes da reta estimada. As comparações são favoráveis ao estimador proposto por ele, já que os outros dois estimadores usados apresentam restrições teóricas conhecidas.

Ripley e Thompson (1987) escrevem a função de verossimilhança para os dados de um estudo de comparação de métodos, considerando que as variâncias em cada ponto são conhecidas. Apresentam o algoritmo de cálculo e deduções no artigo. Também são apresentados exemplos de aplicação, onde se constata que as estimativas obtidas pelo método proposto são diferentes das obtidas por mínimos quadrados ordinários, mínimos quadrados ponderados e mínimos quadrados iterativamente reponderados. Este último método é uma adaptação do método de mínimos quadrados ponderados, utilizando como ponderador uma expressão que é função dos erros nos dois eixos. Eles não testam o estimador em experimentos de simulação. No capítulo 3 do presente trabalho é mostrado como se obter o estimador proposto por estes autores.

Outra proposta, citada por Riu e Rius (1995) e Cecchi (1991) é a de Neri et al. (1989), que propõem a minimização da soma dos quadrados das menores distâncias de cada ponto experimental à reta teórica, i.e., das distâncias perpendiculares à reta.

Através da aplicação da lei de propagação dos erros sem aproximações, Riu e Rius (1995) afirmam que a expressão tratada por Neri et al. (1989), se torna igual à utilizada por outros métodos apresentados por eles. Ainda relatam que, segundo Neri et al. (1989), a direção de minimização não é a característica mais importante do método de estimação, porém sim a escolha do fator ponderador na regressão.

Kalantar et al. (1995) apontam como um método particularmente conveniente de estimação, a modificação de Williamson sobre o método de York. Este método se resume a um algoritmo iterativo de solução do problema de minimização da soma dos quadrados dos resíduos ponderados, usando os fatores ponderadores propostos por Néri et al. (1989). A solução proposta por Williamson corrige um problema inerente ao algoritmo de York em seu trabalho original.

York, segundo MacTaggart e Farwell (1992), parte da minimização da mesma expressão utilizada por Kalantar et al. (1995), e obtém uma equação polinomial do terceiro grau em  $b$  (coeficiente angular da reta). Os coeficientes do polinômio obtido dependem dos pesos de regressão, que por sua vez dependem de  $b$ , o que faz com que na verdade a equação não seja cúbica em  $b$ , como ele afirma.

Tratando a expressão por ele derivada como um polinômio de terceiro grau, York apresenta uma solução genérica iterativa que fornece as três raízes para o mesmo. Ele afirma que a terceira solução obtida pelo método proposto por ele é sempre a buscada. O método de resolução das raízes do polinômio proposto por York converge muito lentamente, mas para certos dados ou certas estimativas iniciais para  $b$ , ele não converge de forma alguma. O uso de outras alternativas para a resolução do polinômio, como o método de Newton-Raphson, dependem da região inicial de busca, e não há como se garantir que a raiz encontrada é a desejada. Assim, o método de York apresenta sérios problemas para aplicação prática.

MacTaggart e Farwell (1992) comparam o método proposto por York com a modificação proposta por Williamson. Segundo estes, Williamson parte das mesmas expressões que York, porém reescreve a expressão em  $b$  como um polinômio do primeiro grau, ao invés de um do terceiro. O raciocínio dele seria que como o peso  $w_i$  inclui  $b$  no denominador, a expressão resultante poderia ser construída como um polinômio de qualquer grau. Sendo assim, escolheu o primeiro grau, o que faz com que se necessite de menos iterações para se estabilizar o resultado, e a convergência seja garantida. O método de Williamson apresenta uma característica adicional que ele apresenta como vantagem. Além das estimativas dos parâmetros não dependerem da escolha de métodos para  $x$  e  $y$ , também suas expressões para as variâncias das estimativas não dependem da atribuição de métodos aos eixos. O método de York, apesar de satisfazer a primeira condição, não o faz para a segunda.

Lisy et al. (1990) propõem outra solução para a o problema de minimização da soma dos quadrados dos resíduos ponderados, que é melhor descrita em Riu e Rius (1996). Utilizando um fator ponderador diferente do usado por todos os outros autores mencionados, estes autores chegam a uma equação matricial não linear, a partir da derivação das equações normais de regressão . Este procedimento é também detalhado na seção de métodos deste trabalho (capítulo 3). O fator ponderador proposto por estes autores difere pela inclusão de um termo que considera a covariância entre métodos, ponto a ponto. Na definição destes autores o fator ponderador é apresentado como denominador na expressão da soma de quadrados a ser minimizada, representando a variância no ponto e não o seu inverso como é comum.

Uma observação importante a respeito de alguns métodos aqui citados, é que, como partem das mesmas equações originais, representam não diferentes estimadores, mas diferentes procedimentos numéricos para se encontrar os estimadores desejados. Entende-se como enquadrados nesta situação os procedimentos propostos por Neri et al. (1989), York (1966), Williamson, descrito em MacTaggart e Farwell (1992) e em Kalantar et al. (1995), Lisy et al., descrito por Riu e Rius (1996). Todas estas abordagens não passam de diferentes algoritmos de cálculo desenvolvidos para se encontrar as estimativas a partir da mesma equação original. A racionalidade da expressão é adaptar o método de mínimos quadrados ponderados a uma situação onde a variável independente possui erro (e a rigor ela não tem por que ser chamada independente), e o peso dado ao  $i$ -ésimo resíduo é redefinido de acordo. Na abordagem de Lisy et al. e Riu e Rius (1996) o peso também é diferente, mas o algoritmo proposto poderia ser também utilizado com outras definições de ponderadores.

Outras propostas aparecem na literatura, porém os resultados de Riu e Rius (1995) desencorajam a sua utilização.

Alguns métodos que aparecem na literatura não levam em consideração a variação do desvio padrão do erro aleatório com a concentração, porém apresentam grande resistência à presença de dados contaminados ou discrepantes. Estes pertencem à classe de métodos chamados robustos.

Um dos mais citados trabalhos a respeito é o de Rousseeuw e Leroy (1987). Estes autores propõem a minimização da mediana dos quadrados dos resíduos de regressão.

Massart et al. (1986) mostram que este estimador é uma alternativa adequada para se detectar a presença de dados discrepantes, erros de ajuste de modelo em calibração, assim como para estudos de comparação de métodos.

Rutan e Carr (1988) apresentam um estudo de simulação onde se comparam o estimador de Rousseeuw (LMS = “least median of squares”) com o de mínimos quadrados ordinários, o de mínimos valores absolutos dos desvios, e dois algoritmos utilizando filtros de Kalman.



O estimador da mínima mediana dos quadrados dos resíduos de Rousseeuw (LMS) se mostra como sendo o mais eficiente para a detecção de pontos discrepantes em experimentos de calibração com poucas observações (4 a 10 pontos).

Passing e Bablok (1983) propõem um método que considera todas as retas possíveis obtidas através de dois pontos experimentais para obter seu estimador. Esta técnica é citada com frequência na literatura de química analítica quando se trata de métodos robustos. Alguns exemplos disto são Payne (1997), Zwanziger e Sârbu (1998), Feldmann (1992), dentre outros.

Feldmann (1992) apresenta duas alternativas para o caso bivariado, que ele denomina regressão estrutural da mínima mediana, e regressão estrutural da mediana absoluta. O primeiro destes estimadores é, segundo o autor, uma versão bivariada do estimador da mínima mediana dos quadrados, de Rousseeuw (1984). Feldmann afirma que o estimador proposto por ele é mais difícil de implementar computacionalmente do que o LMS, pois além de produzir a estimativa de um coeficiente angular robusto  $b$ , produz também a estimativa de um ponto focal robusto  $(\hat{\mu}_x, \hat{\mu}_y)$ . Para obter estas estimativas, ele utiliza um algoritmo de otimização pelo método Simplex.

O método da regressão estrutural da mediana absoluta é apresentado por Feldmann para contornar as dificuldades em se otimizar a expressão proposta para seu outro estimador. A adaptação por ele proposta para o estimador não necessita de resolução por nenhum método numérico iterativo e nem de algoritmos de otimização.

Feldmann apresenta em seu artigo a forma de se realizar a análise dos resíduos de regressão para detecção de pontos discrepantes, assim como exemplos de aplicação onde conclui que a facilidade do segundo estimador proposto justifica sua utilização em relação ao primeiro.

Sarabia et al. (1997) apresentam também sua versão bivariada do estimador por LMS de Rousseeuw. Estes autores apresentam este estimador como sendo o da mínima mediana dos quadrados ortogonais (forma reduzida para 'mínima mediana dos quadrados dos resíduos ortogonais'). Estes autores não apresentam o algoritmo de cálculo para se encontrar a solução da expressão que eles propõem, porém ela pode ser obtida a partir de uma sugestão de Rousseeuw (1984), também discutida em Rousseeuw e Leroy (1987). Para avaliar o desempenho do estimador, Sarabia et al. (1997) realizam simulações de Monte Carlo. Comparam seu estimador com o de mínimos quadrados ordinários, com o método de Deming, o LMS de Rousseeuw e Leroy (1987). A ênfase das simulações é a presença de contaminação nos dados. Os casos estudados para a simulação representam um experimento fatorial em que se variam: erros aleatórios em X e em Y, com desvio padrão constante com a razão de variâncias assumindo diferentes valores, sendo os dados distribuídos normalmente e uniformemente; número de pontos discrepantes e tamanho das discrepâncias. A simulação, feita com o programa MATLAB 4.0, leva os autores a concluir



que o método de LMS ortogonal, proposto por eles é o melhor a se utilizar quando se suspeita de contaminação nos dados.

Hartmann, Vankeerberghen, Smeyers-Verbeke e Massart (1997) também utilizam este estimador, mas apresentam em detalhes o algoritmo de cálculo, inclusive mostrando propostas de melhoria no algoritmo original proposto por Rousseeuw e Leroy (1987). Estes autores também apresentam uma variação ponderada para este estimador, que considera assim, além dos erros nos dois eixos, a variação destes erros com a concentração.

Embora sejam muito conhecidos outros estimadores robustos para regressão, como a regressão  $L_1$ , proposta por Edgeworth (1887), os M-estimadores de Huber (1973), os M-estimadores generalizados de Mallows (1975) e Schweppe (1977), a mediana repetida de Siegel (1982), estes não são utilizados na literatura de química analítica. Além disto, com o trabalho de Rousseeuw e Leroy (1987), estes estimadores ficam em aparente desvantagem, por exemplo quando se considera o “ponto de quebra”, definido por Hampel (1971). O “ponto de quebra” é a percentagem de pontos discrepantes necessária para mudar significativamente a estimativa. Para o método de mínimos quadrados, é mostrado que o ponto de quebra é zero, ou seja, se um dado estiver discrepante dos demais, isto já é o suficiente para mudar as estimativas da reta. Para o LMS ele é de ~50%, que é o máximo que um estimador pode apresentar em termos de ponto de quebra.

Um estimador importante a ser considerado é o de mínimos quadrados aparados (LTS = “least trimmed of squares”), também proposto por Rousseeuw e Leroy (1987), que segundo estes autores supera o de LMS em eficiência.



## 2.2 MODELAGEM MATEMÁTICA DA PRECISÃO

Quando se citou anteriormente que abordagens de regressão ponderada se aproximam melhor do problema de comparação de métodos de análise química, também foi dito que esta ponderação do erro nos eixos pode ser obtida pelo conhecimento de como este erro varia com a concentração.

As técnicas recomendadas na literatura para a utilização de estimativas de precisão nos fatores ponderadores de regressão não impõem uma forma funcional para as mesmas. Ripley e Thompson (1987) admitem que existe tal função matemática, citando o modelo linear, mas não o utilizam para derivar seu método. Linnet (1990) utiliza o modelo de coeficiente de variação constante quando as variâncias não são constantes ao longo da faixa de concentrações avaliada. Este modelo é equivalente ao modelo linear para o desvio padrão (com intercessão nula), que apesar de muito utilizado não se ajusta a todas as situações.

Feinberg (1995) apresenta a abordagem da norma ISO 5725-2 (1994) para se estimar esta relação. Pode-se ver nesta norma, no item 16.2, que a relação entre precisão e concentração pode ocorrer de três formas:

$$a) \sigma_x = \kappa \mu_x \quad (\text{uma reta passando pela origem}), \quad (1)$$

$$b) \sigma_x = \sigma_0 + \kappa \mu_x \quad (\text{uma reta com intercepto positivo}), \quad (2)$$

$$c) \sigma_x = \sigma_0 \mu_x^\kappa \quad \text{com } \kappa \leq 1 \quad (\text{uma relação de potência}). \quad (3)$$

O valor  $\sigma_x$  acima representa o desvio padrão de análise, seja para repetibilidade ou reprodutibilidade, ou desvio padrão entre laboratórios.  $\mu_x$  representa a concentração da substância analisada na amostra.

A interpretação dos parâmetros nestes modelos é direta. O valor  $\sigma_0$  representa o desvio padrão para a concentração zero, no segundo modelo, e pode ser utilizado para estimar o limite de detecção do método de análise. O primeiro modelo é um caso particular do segundo, e pode ser reescrito como um modelo para coeficiente de variação constante. O coeficiente de variação então é representado pelo parâmetro  $\kappa$ . No terceiro modelo, o pressuposto básico é o de que os erros de análise são log-normais, e o parâmetro  $\sigma_0$  representa o desvio padrão para uma concentração unitária. Como usualmente os valores de desvio padrão são estimados através de dados experimentais, os modelos são reescritos como:

$$i) \quad s_{x_i} = \sigma_0 + \kappa x_i + \xi_i \quad (4)$$

e

$$\text{ii) } \log(s_{x_i}) = \log(\sigma_0) + \kappa \log(x_i) + \zeta_i. \quad (5)$$

Os termos  $\xi_i$  e  $\zeta_i$  representam erros independentes normais, com médias nulas e variâncias  $\sigma_\xi^2$  e  $\sigma_\zeta^2$ , respectivamente. As estimativas para  $\sigma_0$  e  $\kappa$  são obtidas por mínimos quadrados. No primeiro modelo (4) é necessário utilizar-se de ponderação na regressão, já que a estimativa do desvio padrão possui variância proporcional ao valor do desvio padrão. A norma ISO 5725-2 recomenda um procedimento iterativo onde os pesos de regressão são obtidos a partir das estimativas  $\hat{s}_{x_i}$  produzidas pela reta de regressão. No modelo exponencial em (5), toma-se os logaritmos dos dados antes de se proceder à regressão. Com esta transformação, a variância é estabilizada, e é possível se utilizar o método de mínimos quadrados ordinários.

As estimativas de desvio padrão necessárias para se proceder ao estudo de regressão que modela o desvio padrão do erro em função da concentração através das opções acima, são obtidas por análise de variância de efeito aleatório, com um fator de classificação. O experimento proposto pela norma ISO 5725-2 para se produzir estas estimativas consiste em se escolher um número  $q$  de amostras a serem testadas pelo método de análise química em consideração. De cada uma destas amostras são retiradas  $p$  alíquotas, cada uma sendo enviada a um laboratório diferente. Cada laboratório deve realizar  $n$  análises repetidas na mesma alíquota. A análise de variância utilizada para tratamento dos dados oriundos deste experimento consideram o fator Laboratório como sendo um fator de efeito aleatório, e a variabilidade obtida pela combinação dos desvios padrão das réplicas de cada participante é usada como estimativa da repetibilidade do método de análise. O desvio padrão entre as médias de cada laboratório é utilizado para se produzir uma estimativa da reprodutibilidade. Após se dispor de estimativas destas duas medidas de precisão em cada amostra, estes valores são utilizados em uma análise de regressão, onde são estudadas contra a concentração média obtida por todos os laboratórios em cada amostra.

Existem críticas a esta abordagem proposta pela ISO, como Feinberg (1995) cita. Também é discutido o planejamento de experimentos proposto nesta norma por Zaalberg (1989), que propõe a utilização do esquema descrito na norma holandesa NEN 6303. Este esquema é um planejamento do tipo “split level”, em que duas amostras são selecionadas em cada nível de teste, com concentrações muito próximas. Como as identificações destas duas amostras são diferentes, os laboratórios participantes não sabem *a priori* que estão testando amostras para o mesmo nível do experimento. Também são incluídas repetições das mesmas amostras sem que os laboratórios participantes tomem conhecimento disto. Assim, obtém-se geralmente estimativas de variâncias superiores às obtidas com o experimento proposto na ISO 5725, o que, para Zaalberg significa ser mais realista. Ele afirma que o experimento do tipo “split-level” é melhor porque as realizações repetidas de uma mesma amostra ou de um mesmo nível são ignorados pelos analistas, evitando assim possíveis



vícios laboratoriais. Apesar das críticas, o modelo ISO continua sendo o mais amplamente utilizado para estimação dos erros aleatórios de métodos sendo publicados, e consequentemente, das curvas de precisão. Um exemplo de influência do modelo ISO pode ser visto no protocolo IUPAC (1988) para a condução de testes interlaboratoriais. Mesmo sendo um modelo ainda muito utilizado, a ISO incorporou a sugestão da norma holandesa em sua última revisão da ISO 5725, e esta consta na Parte 5 da atual norma.

Uma das vantagens da abordagem ISO é que ela apresenta a alternativa de se ajustar um modelo, para a precisão, do tipo proposto por Horwitz (1982), que é amplamente aceito. Horwitz et al. (1980), através da avaliação de mais de 50 testes interlaboratoriais conduzidos pela AOAC, verificaram que a precisão de análises químicas varia com a concentração segundo uma relação de potência como a descrita em (3). Estes autores preferiram trabalhar com o coeficiente de variação de reprodutibilidade, que segue uma relação como em (10), porém com  $\kappa < 0$ . A expressão por eles encontrada é a seguinte:

$$CV_R (\%) = 2^{(1-0,5 \log_{10} C)}, \quad (6)$$

onde o índice  $R$  simboliza reprodutibilidade, como definido na norma ISO 5725-1 (1994). A concentração  $C$  aqui é expressa em fração de massa, ou seja,  $C=1$  para uma substância pura, e  $C=10^{-6}$  representa uma parte por milhão ( $1 \text{ ppm} = 1 \mu\text{g} / \text{g}$ ). Esta expressão mostra de forma simples, que a imprecisão (em termos do erro relativo  $CV_R$ ) duplica a cada redução de duas ordens de grandeza na concentração. Hall e Selinger (1989) apelidaram esta relação de “Trompa de Horwitz”, devido à forma como é apresentada, o que pode ser visto na Figura 2.2.

Estes autores descrevem esta descoberta de Horwitz como sendo “uma das mais intrigantes relações empíricas da química analítica moderna”, e tentam explicá-la através de um modelo que aproxima a concentração ao resultado de um experimento binomial. Usando a lei de Zipf (1965), chamada “Princípio do menor esforço”, Hall e Selinger encontram uma expressão que se ajusta bem à curva de Horwitz, e sendo assim, na opinião deles, apoiam a idéia de que “o esforço colocado em detectar menores concentrações aumenta de tal forma que a precisão muda como se o tamanho da amostra aumentasse inversamente com a concentração<sup>1</sup>”. Horwitz e Albert (1997) retornam ao assunto, como que respondendo a Hall e Selinger (1989), e propõem uma derivação heurística para a relação.

<sup>1</sup> Tradução do original

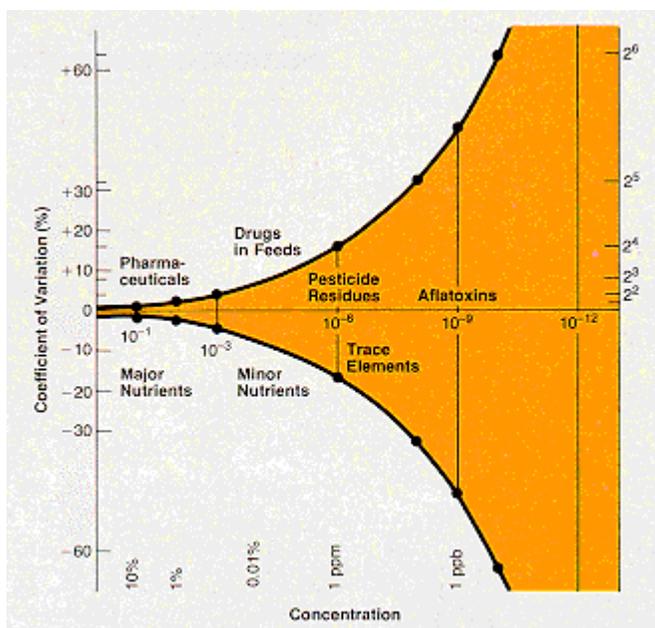


Figura 2.2: “Trompa” de Horwitz.

As concentrações no eixo das abcissas são apresentadas em escala logarítmica (base 10) e crescem da direita para a esquerda. Nas ordenadas, o coeficiente de variação de reprodutibilidade expresso em porcentagem é apresentado em escala logarítmica (base 2).

FONTE: Horwitz, W. (1982) - Evaluation of analytical methods used for regulation of foods and drugs. *Analytical Chemistry*, Vol. 54, No.1, 67A-76A.

De forma simples, propõem que uma variação infinitesimal na concentração, em relação ao valor da concentração, é acompanhada proporcionalmente por uma variação infinitesimal da precisão analítica em relação à própria precisão. Expressando isto de forma matemática,

$$\frac{d\sigma}{\sigma} = \alpha \frac{dC}{C}, \quad (7)$$

que por integração fornece a relação

$$\log \sigma = \alpha \log C + \beta. \quad (8)$$

Utilizando dados experimentais, as constantes  $\alpha$  e  $\beta$  são estimadas, e após transformações apropriadas, os autores chegam à expressão em (6). Pode-se perceber que esta expressão em (8) é equivalente à terceira alternativa da norma ISO 5725-2 (1994). Isto significa que é admitido que, embora o trabalho de Horwitz tenha produzido uma relação que representa o erro aleatório de métodos de análise química quando avaliados conjuntamente, misturando técnicas analíticas, diferentes matrizes e diferentes componentes analisados (analitos), esta



relação poderia representar a função de erro de um único método, avaliado em um único analito em um único tipo de matriz.

Esta suposição pode ser confirmada experimentalmente, e no capítulo de aplicações do presente trabalho isto é exemplificado.

Thompson e Lowthian (1997) retornam aos dados tratados por Horwitz, porém com acréscimos em relação aos artigos originais que Horwitz publica em 1980 e 1982. De posse agora de 7502 observações, coletadas entre 1915 e 1995 em testes interlaboratoriais de avaliação de desempenho de métodos de análise, estes autores testam a hipótese de que não há mudanças na precisão analítica ao longo do tempo. Seria de se esperar que com o surgimento de novas tecnologias os erros apresentassem alguma forma de decrescimento com o tempo, porém estes autores não conseguem rejeitar a hipótese de que isto não ocorre.

O acréscimo de dados ao longo do período citado aumenta a abrangência da relação empírica de Horwitz, devido à inclusão de diferentes matrizes além das originais da área de alimentos, produtos farmacêuticos e agrícolas, passando a incluir, por exemplo, amostras de estudos geológicos (Horwitz e Albert, 1995).

Horwitz dedicou uma série de estudos a verificar a validade de sua função matemática em diversos campos, como se pode ver em Boyer, Horwitz e Albert (1985), Grappin e Horwitz (1988), e levou à sua proposta de aceitação de resultados de testes interlaboratoriais. Esta proposta, expressa de forma simples, determina que resultados aceitáveis para o desvio padrão de reprodutibilidade, ou o coeficiente de variação de reprodutibilidade, se encontram entre metade e o dobro do encontrado a partir da curva de Horwitz. As considerações tomadas para se chegar a esta proposta podem ser vistas em Horwitz e Albert (1987) e Horwitz, Britton e Chirtel (1998).

A proposta de Horwitz para se aceitar um novo método de análise em testes de colaboração internacional dentro de comitês técnicos (como os da ISO, IUPAC e AOAC) em função do desempenho do método em relação à curva descrita por ele, traz mais objetividade ao julgamento de um novo método. Com isto, Horwitz tenta expressar objetivamente os critérios utilizados pelos químicos ao longo do tempo para este propósito. Estes critérios, antes da proposição de Horwitz, não eram produto de uma avaliação matemática, mas eram função da experiência prática dos profissionais de química, e de certa maneira, subjetiva, ao avaliar novos métodos. Segundo Thompson (1999), a curva de Horwitz representa um critério de seleção natural não calculado (análogo à proposta de evolução das espécies de Darwin), imposto pelos químicos devido às necessidades da sociedade em termos de erros analíticos nas diversas aplicações. Como tal, não existe garantia de que um dia esta relação não possa evoluir para outra diferente. De fato, ao se analisar, por exemplo, os dados produzidos pelo comitê de análises químicas de minério de ferro da ISO, o ISO/TC 102/SC

2, percebe-se uma outra realidade, de exigências diferentes daquelas encontradas pelos químicos que geraram os dados utilizados por Horwitz.

Embora Horwitz e seus colaboradores utilizem preferencialmente o coeficiente de variação, não há nada que impeça o uso do desvio padrão, como é de praxe, por exemplo, nas normas ISO de análise química. Além disto, Horwitz afirma que o desvio padrão dentro do laboratório (ou desvio padrão de repetibilidade) não apresenta comportamento tão definido como o desvio padrão de reprodutibilidade. Entretanto, a Figura 2.3 mostra que o primeiro pode ser relacionado com a concentração, de forma análoga. Os dados para a produção destes gráficos correspondem aos valores publicados em Margosis, Horwitz e Albert (1988), Peeler, Horwitz e Albert (1989), Horwitz, Albert, Deutsch e Thompson (1990), Horwitz e Albert (1991) e Horwitz, Albert, Deutsch e Thompson (1992). Estes se constituem de 2221 pontos para a repetibilidade ( $s_r$ ) e 2846 pontos para a reprodutibilidade ( $s_R$ ).

Thompson e Lowthian (1997) mostram que para concentrações muito baixas ( $C < 10^{-8}$ ), a relação de Horwitz já não é mais válida. De fato, dados de testes interlaboratoriais, para baixas concentrações, sofrem dois efeitos importantes. Em primeiro lugar, substâncias presentes em baixas concentrações normalmente são sujeitas a pressões de órgãos governamentais ou legislação, devido às suas implicações nas atividades humanas, como é o caso de tóxicos ou poluentes, que são muitas vezes prejudiciais em concentrações muito pequenas. Isto faz com que o esforço em se obter bons resultados nestes níveis seja muito diferente do relativo a outras áreas da análise química. Também a instrumentação e cuidados metodológicos são especiais para estes fins. Com isto, tem-se na verdade outra população, diferente da que produziu os resultados avaliados inicialmente por Horwitz.

Em segundo lugar, resultados muito baixos passam a sofrer censura, ora por se chegar próximo ao limite de detecção das técnicas utilizadas, ora devido ao grau de arredondamento utilizado na expressão dos resultados, para se informar resultados com precisão realista e de acordo com as exigências específicas ao campo de aplicação das análises.

De fato, parece que nesta faixa de concentrações, quando se aproxima dos limites de detecção, o modelo de Horwitz ou as equações apresentados pela ISO 5725, parecem não representar bem a realidade.

Uma crítica a estas aproximações é apresentada por Rocke e Lorenzato (1995). Estes autores propõem representar a precisão, nestas situações, por um modelo com dois componentes de erro, um válido em baixas concentrações, onde a experiência mostra que o desvio padrão da análise é aproximadamente constante, e outro válido em altas concentrações, onde eles propõem um coeficiente de variação constante.

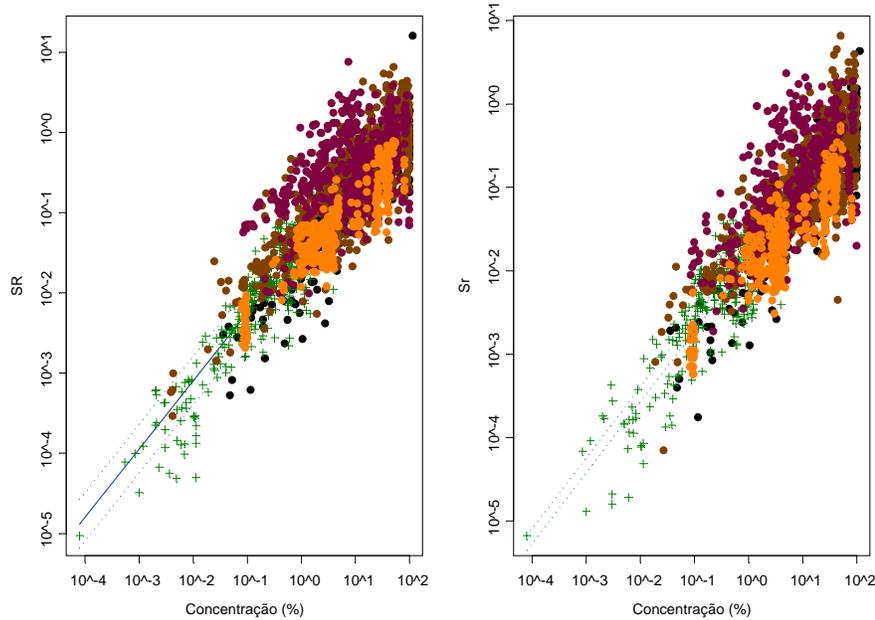


Figura 2.3: Variação dos desvios padrão de repetibilidade ( $S_r$ ) e reprodutibilidade ( $S_R$ ) estimados em experimentos apresentados em 5 publicações da autoria de Horwitz e colaboradores. As cores diferentes dos pontos se referem a diferentes publicações. As escalas em ambos eixos são logarítmicas.

Estas suposições afetam a definição do problema proposto inicialmente, já que implicam em que para baixas concentrações o valor medido seria:

$$x = \mu_x + \varepsilon, \quad (9)$$

sendo  $\varepsilon$  um erro aleatório normal com esperança igual a zero e variância  $\sigma_\varepsilon^2$ ,  $x$  é o valor observado experimentalmente e  $\mu_x$  a concentração real da amostra. Para altas concentrações,

$$x = \mu_x e^\eta. \quad (10)$$

Neste caso,  $\eta$  é o erro aleatório normal, com esperança igual a zero e variância  $\sigma_\eta$ , o que faz com que o erro em  $x$  neste caso seja log-normal.

Definindo em um só modelo,

$$x = \mu_x e^\eta + \varepsilon, \quad (11)$$

e o erro normal  $\varepsilon$  pode ser reescrito como

$$\varepsilon = x - \mu_x e^\eta \quad (12)$$

e a verossimilhança associada aos parâmetros  $(\sigma_\eta, \sigma_\varepsilon)$ , dadas  $n$  medidas  $x_i$  realizadas em amostras contendo as verdadeiras concentrações  $\mu_{x_i}$  é:

$$\prod_{i=1}^n \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_\varepsilon\sigma_\eta} e^{-\eta^2/(2\sigma_\eta^2)} e^{-(x_i - \mu_{x_i} e^\eta)^2/(2\sigma_\varepsilon^2)} d\eta. \quad (13)$$

Os autores calculam a distribuição marginal em  $\eta$  devido à impossibilidade de tratar a distribuição conjunta dos dois erros para estimação. Através da maximização da expressão acima, pode-se obter as estimativas dos componentes  $\sigma_\eta$  e  $\sigma_\varepsilon$ . Os autores solucionam o problema de estimação por um método de integração numérica e posterior otimização, usando um método Gauss-Hermite, que aproxima a integral a uma soma. Para isto, o logaritmo do integrando é aproximado por uma série de Taylor de segunda ordem.

Depois de obtidas as estimativas para os dois componentes de precisão, Rocke e Lorenzato propõem calcular a precisão total em função da concentração a partir das mesmas. Considerando que o primeiro componente é normal com variância  $\sigma_\varepsilon^2$  e o segundo componente é log-normal, se ambos são independentes a variância pode ser considerada aditiva em relação às variâncias dos dois componentes, o que produz a expressão

$$\sigma_x^2 = \sigma_\varepsilon^2 + \mu_x^2 e^{\sigma_\eta^2} (e^{\sigma_\eta^2} - 1), \quad (14)$$

com os índices  $i$  omitidos.

Na realidade, a proposta de Rocke e Lorenzato se aproxima bem dos resultados experimentais quando as concentrações partem do limite de detecção até valores que se afastam um pouco desta região. Entretanto, para concentrações que se espalham por ordens de grandeza maiores, percebe-se um afastamento do comportamento de coeficiente de variação constante, que não é novidade, como se pode ver pela expressão da curva de Horwitz. Esta é a principal limitação desta modelagem, já que o segundo componente dela representa o coeficiente de variação, que é assumido constante para altas concentrações.

Sendo assim, não se dispõe ainda de um modelo completo, que expresse o erro aleatório em função da concentração ao longo de uma faixa analítica compreendendo várias ordens de



grandeza, e.g., do limite de detecção até o nível de substância pura ou próximo disto. O que se faz é utilizar aproximações conforme os dados se apresentem.

Existem propostas diferentes de modelagem do erro, porém com pouca aceitação na comunidade de química analítica em geral. Watters, Carroll e Spiegelman (1987), por exemplo, propõem em algumas situações a modelagem do desvio padrão como uma função polinomial de segundo grau da concentração, ou ainda modelam a variância como um polinômio do segundo grau.

Independentemente do modelo matemático assumido, uma questão experimental deve ser levada em consideração para se calcular as estimativas de desvio padrão em cada amostra analisada. Para comparar dois métodos, vários autores dizem que a precisão (ou as variâncias) precisa ser conhecida, para cada um dos métodos comparados. Quando isto não acontece, recomenda-se que cada amostra seja repetida em cada método, gerando réplicas em cada caso. A utilização de dados em duplicata ou triplicata, como sugerido por Riu e Rius (1996), fornece estimativas de desvio padrão do erro que estes autores propõem utilizar na regressão ponderada para comparar os dois métodos em questão. Entretanto, sabe-se que com tal número de réplicas a estimativa do desvio padrão é muito pouco precisa. Schwartz (1979) afirma que, baseado na experiência, esta relação matemática não pode ser conhecida a partir de um número limitado de repetições. Garden, Mitchell e Mills (1980) sugerem realizar, no mínimo, dez repetições de cada amostra para se determinar a função do desvio padrão com a concentração. Richardson e Morrison (1995) utilizam de 7 a 10 réplicas para este mesmo fim. Kane (1997) sugere que no mínimo 15 repetições da mesma amostra sejam realizadas para se estimar o desvio padrão com muita precisão.

Devido a estas observações, corroboradas pela experiência prática, pode-se dizer que para um adequado uso da técnica de regressão sugerida por Riu e Rius (1996), ou qualquer outra técnica que utilize as variâncias de cada método em cada ponto, o melhor seria avaliar primeiramente como o desvio padrão do erro em cada método varia com a concentração. Em outras palavras, é melhor estudar e estimar as funções matemáticas do desvio padrão do erro aleatório contra concentração para então utilizá-las como fatores ponderadores de regressão, considerados conhecidos a partir destas expressões, na comparação dos métodos que produzem os resultados  $x_i$  e  $y_i$  observados.

O conhecimento destas relações é importante, também, para outra finalidade que geralmente a literatura de estudos de comparação de métodos não aborda. Quando se compara métodos, além de se testar a hipótese de igualdade entre eles, em termos de resposta, também é desejável testar a igualdade entre as precisões estimadas para os mesmos.

Em termos simples, não só é desejável saber se os métodos de análise se comparam favoravelmente em termos das concentrações estimadas, mas também em termos dos erros



cometidos ao se estimar estes valores. Por exemplo, se uma amostra possui teor de zinco de 0,1% em peso, supondo dois métodos exatos, apresentando as seguintes estimativas para esta amostra,

$$\bar{x} = 0,100 \pm 0,001\% \quad \text{e} \quad \bar{y} = 0,1 \pm 0,1\% ,$$

podemos concluir que o primeiro método apresenta desempenho melhor, mesmo que sendo ambos equivalentes em termos de esperança matemática.

Obviamente, na avaliação de métodos deseja-se avaliar comparativamente os dois métodos em questão não apenas para uma amostra, mas para toda uma faixa de concentrações onde os dois métodos são válidos.

Para a resolução deste problema, pouco desenvolvimento existe na literatura. Sutarno e Steger (1985) propõem a comparação ponto a ponto, ou seja, através da análise de uma mesma amostra  $n$  vezes por cada um dos métodos e a comparação das variâncias é feita por um teste F para a razão das mesmas. Esta abordagem corresponde à da norma ISO 5725 e é descrita por Feinberg (1995).

Spurrier (1992) apresenta várias alternativas para comparação de variâncias, porém contra uma única referência fixa. Fung e Tsang (1998) fazem o mesmo com relação ao coeficiente de variação. Kulasekera (1995) mostra uma alternativa para comparação de duas curvas de regressão genéricas, porém desenvolve metodologia para abordagens não-paramétricas. Para o caso da precisão em análises químicas, o usual é a modelagem paramétrica da relação. Nalimov (1963) apresenta uma proposta para comparação de duas retas obtidas por regressão, porém esta se constitui na comparação em passos separados, cada um com seu nível de significância, dos erros padrão de regressão das duas retas, prosseguindo com a comparação das inclinações, para ao fim comparar as intercessões. Com isto, o teste global deixa de ser um teste ótimo, não sendo um teste simultâneo para as três estimativas citadas.

No capítulo de métodos, no presente trabalho, é apresentada uma proposta para se comparar duas funções de precisão, utilizando regressão com variáveis indicadoras.

Todo o trabalho de revisão até aqui realizado mostra que existem inúmeras propostas para a comparação dos resultados de análise química obtidos por dois diferentes métodos, e que algumas apresentam formulação mais genérica, e são portanto teoricamente mais completas, respondendo a características específicas deste tipo de aplicações, quais sejam, a presença de erros nos dois vetores de dados comparados, com variância não homogênea ao longo da faixa de concentrações utilizada para os testes. A opinião da maioria dos autores que utilizam alguma forma de regressão robusta para avaliação de dados contaminados é unânime, no sentido de que os métodos robustos são úteis para se destacar pontos influentes suspeitos, que após análise adequada devem ser retirados ou substituídos, para se prosseguir



depois com uma análise paramétrica com propriedades conhecidas, como o método de Deming, ou o proposto por Riu e Rius (1996).

Vários autores procuram apresentar as vantagens e mostrar o desempenho dos estimadores que eles propõem através de estudos de simulação computacional. Entretanto, os métodos que teoricamente são os mais indicados não são comparados entre si, mesmo quando os autores os referenciam nos textos de seus trabalhos. Por exemplo, Riu e Rius (1995) citam o estimador de máxima verossimilhança de Ripley e Thompson (1987) e o proposto por Linnet (1990), mas em sua simulação (Riu e Rius, 1996) não os incluem para comparação. Usualmente os métodos comparados nestes estudos são o método proposto pelo autor, o método de mínimos quadrados ordinários e mínimos quadrados ponderados.

Poucos autores estudam o efeito do número de pontos na regressão sobre as estimativas, que seria útil para se aprender sobre o comportamento assintótico dos vários estimadores. A precisão, medida pelo desvio padrão dos dois métodos, é sempre considerada constante ou varia conforme um modelo linear com o teor.

A comparação de métodos se restringe normalmente à comparação entre resultados de análise de dois métodos em amostras gêmeas utilizadas em ambos. Não se avalia a equivalência entre as precisões da análise nos dois métodos. Quando existe esta preocupação, notadamente em comitês técnicos que publicam métodos oficiais, a abordagem usada é de comparação ponto a ponto, o que apresenta uma série de limitações.

No capítulo 3, a seguir, procura-se especificar bem os algoritmos para implementação dos métodos estatísticos mais relevantes encontrados na literatura, para realizar as comparações sugeridas acima, posteriormente, no capítulo de simulações.



## Capítulo 3

### Métodos Estatísticos

A literatura da química analítica apresenta uma grande quantidade de publicações que se dedicam ao estudo de comparação de métodos por meio de análise de regressão. A grande quantidade de referências neste assunto pode fazer parecer que são utilizadas muitas opções diferentes de estimadores, mas em realidade são poucos os princípios básicos nos quais estes se fundamentam. Existe, sim, uma grande quantidade de algoritmos numéricos para se obter as estimativas, e infelizmente a fundamentação teórica nem sempre é bem apresentada, gerando confusão no entendimento. Métodos atribuídos a alguns autores na verdade não podem ser encontrados nas referências dadas, ou são encontrados como sendo apresentados por outros autores. Os princípios de derivação dos estimadores nem sempre são claros, ou às vezes são trocados. Como exemplo temos o método ponderado de Deming, como apresentado por Linnet (1990), em que este segundo autor afirma derivar as equações aplicando o método dos mínimos quadrados. Este método é apresentado no presente capítulo, e percebe-se que ele pode ser encontrado utilizando o método de máxima verossimilhança, e não da forma como o autor afirma.

No presente capítulo, alguns dos métodos citados na revisão bibliográfica foram escolhidos, e pretende-se mostrar como obter os estimadores referentes a eles. Algumas considerações sobre sua utilização em inferência são também mostradas, para uso no capítulo de aplicações, onde estes são comparados em casos reais e simulados. Nos casos simulados, é possível verificar o desempenho comparativo dos mesmos, e o objetivo final é de se escolher qual deles se aplica melhor ao problema de comparação de métodos em química analítica.

O modelo utilizado para todos os métodos a serem aqui descritos pode ser apresentado de uma forma comum, e em cada caso serão apontados os pressupostos básicos que divergem do modelo comum. A comparação de métodos supõe, de início, a igualdade entre dois candidatos, sendo que um deles pode ser considerado como sendo uma referência contra a qual o segundo é avaliado. Eles são comumente designados como método de referência e método alternativo (ou método de teste), respectivamente.

O estudo prevê então a coleta de  $n$  pares  $(x_i, y_i)$ , onde os valores  $x_i$  se referem a dados observados do primeiro método e  $y_i$  se referem ao segundo método, observações obtidas nas mesmas amostras. Como os valores observados contêm erro, estes se relacionam com os “verdadeiros” valores de concentração segundo as seguintes expressões:



$$x_i = \mu_{x_i} + \varepsilon_i \quad (15)$$

$$\text{e } y_i = \mu_{y_i} + \delta_i. \quad (16)$$

Os erros  $\varepsilon_i$  e  $\delta_i$  são variáveis aleatórias independentes entre si, e de tal forma que

$$\varepsilon_i \perp \varepsilon_j \forall i \neq j \quad \text{e} \quad \delta_i \perp \delta_j \forall i \neq j. \quad (17)$$

Se eles apresentam resultados iguais, pode-se afirmar que estão relacionados linearmente segundo a equação

$$\mu_{y_i} = \alpha + \beta \mu_{x_i}. \quad (18)$$

Neste caso, tem-se que a hipótese básica a ser testada é

$$H_0 : \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (19)$$

Para se testar esta hipótese, torna-se necessário estimar os parâmetros  $\alpha$  e  $\beta$ . Estas estimativas serão aqui denominadas de  $a$  e  $b$ , respectivamente. O conhecimento dos erros de estimação destes é importante na obtenção de regiões de confiança para se realizar inferência sobre os parâmetros.

Nem todos os métodos apresentados exigem alguma suposição acerca da distribuição dos erros  $\varepsilon_i$  e  $\delta_i$ , porém em geral são considerados como erros normais com média zero e variâncias  $\sigma_{\varepsilon_i}^2$  e  $\sigma_{\delta_i}^2$ , respectivamente.

Das definições acima, tem-se como consequência que

$$\sigma_{x_i}^2 = \sigma_{\varepsilon_i}^2 = \text{variância dos valores observados } x_i \text{ numa mesma amostra.}$$

Analogamente para o método simbolizado pelos dados  $y_i$ :

$$\sigma_{y_i}^2 = \sigma_{\delta_i}^2.$$

Estas variâncias são geralmente consideradas como sendo constantes conhecidas nos problemas de comparação de métodos. Em alguns casos considera-se que apenas a razão entre elas é conhecida.

A forma usual como se obtém o conhecimento sobre estas variâncias é utilizando-se experimentos isolados, onde cada método é avaliado apenas para se estimar estes valores e para se conhecer como eles variam com os teores analisados nas amostras.

Nestes experimentos isolados, dispõe-se de réplicas (aqui indicadas pelo índice  $j$ ) em cada um dos  $i$  níveis testados. Estas réplicas são utilizadas para se estimar os valores  $\sigma_{x_i}^2$  e  $\sigma_{y_i}^2$  para cada amostra, através das expressões

$$s_{x_i}^2 = \sum_{j=1}^{n_{x_i}} \frac{(x_{ij} - \bar{x}_i)^2}{n_{x_i} - 1} \quad \text{e} \quad s_{y_i}^2 = \sum_{j=1}^{n_{y_i}} \frac{(y_{ij} - \bar{y}_i)^2}{n_{y_i} - 1}, \quad (20)$$

$$\text{com} \quad \bar{x}_i = \frac{1}{n_{x_i}} \sum_{j=1}^{n_{x_i}} x_{ij} \quad \text{e} \quad \bar{y}_i = \frac{1}{n_{y_i}} \sum_{j=1}^{n_{y_i}} y_{ij}. \quad (21)$$

Os modelos aqui considerados para a expressão de como os desvios padrão estimados acima variam com a concentração são os chamados “intrinsecamente lineares” por Draper e Smith (1981), apresentados na revisão bibliográfica. São eles,

$$s_{x_i} = \sigma_0 + \kappa x_i + \xi_i \quad (22)$$

$$\log(s_{x_i}) = \log(\sigma_0) + \kappa \log(x_i) + \zeta_i, \quad (23)$$

sendo que

$$\xi_i \perp \xi_j \forall i \neq j; \quad \zeta_i \perp \zeta_j \forall i \neq j \quad \text{e} \quad \xi_i \perp \zeta_i \forall i. \quad (24)$$

A estimativa dos parâmetros  $\sigma_0$  e  $\kappa$  para cada método se faz por métodos de regressão, fornecendo os valores  $s_0$  e  $k$ , respectivamente com os índices  $x$  e  $y$ , conforme se trate do método de referência ou do método alternativo. Para o caso linear (modelo 22 acima) a regressão é iterativa e ponderada, usando as estimativas de  $s_{x_i}$  (valores preditos pela regressão) como sendo o inverso dos pesos de regressão. Para a equação (23) os logaritmos

dos valores são tomados e a regressão pode ser simples, pelo método de mínimos quadrados ordinários.

### 3.1. COMPARAÇÃO DE RESULTADOS

Os métodos aqui descritos podem se resumir em três idéias básicas: minimização da soma dos quadrados dos resíduos, maximização da função de verossimilhança e aplicação de estatísticas de ordem para obter estimadores robustos.

O método de Wald não se enquadra em nenhuma destas, porém é aqui apresentado devido a ser uma referência muito encontrada para o problema de regressão com erros nas variáveis, e também por ter sido apontado como sendo o melhor estimador nas simulações descritas em Danzer et al. (1995).

#### 3.1.1 Métodos de Mínimos Quadrados

Os métodos assim chamados partem da minimização da soma  $S$ , dos quadrados das distâncias verticais entre os pontos experimentais e a reta de regressão:

$$S = \sum_i w_i (y_i - \alpha - \beta x_i)^2. \quad (25)$$

Aqui os ponderadores  $w_i$  são utilizados para que seja possível diferenciar a influência de um determinado ponto sobre as estimativas dos parâmetros. Para as variações encontradas na aplicação deste método, estes pesos recebem atribuições diferentes. Normalmente, são utilizados pesos que se relacionam com a precisão das medidas  $(x_i, y_i)$ , através de suas variâncias  $\sigma_{x_i}^2$  e  $\sigma_{y_i}^2$ .

Para se obter as estimativas para  $\alpha$  e  $\beta$  que minimizam (25), escrevem-se as derivadas parciais de  $S$  em relação aos dois parâmetros e estas são igualadas a zero. Estas são as chamadas equações normais de regressão, e o sistema formado por elas deve ser resolvido para se obter os estimadores buscados. Dependendo da definição dos pesos  $w_i$  acima, este sistema pode ser mais imediatamente resolvido ou não.

##### 3.1.1.1 Mínimos Quadrados Ordinários

Este é o método mais conhecido dentro dos que utilizam mínimos quadrados, e é caracterizado pelo fato de que cada ponto da regressão recebe o mesmo peso. Isto se consegue fazendo  $w_i = 1$ ,  $\forall i$  em (25). Com isto, a resolução do sistema representado pelas equações normais é imediata. As estimativas são:

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \quad \text{e} \quad a = \bar{y} - b\bar{x}. \quad (26)$$

Para as aplicações propostas no capítulo a seguir, é importante conhecer-se os erros padrão com que estas estimativas são produzidas, para posterior inferência sobre os parâmetros. Para este método são:

$$s_a = s \sqrt{\frac{\sum_i x_i^2}{n \sum_i (x_i - \bar{x})^2}} \quad (27)$$

e

$$s_b = \frac{s}{\sqrt{\sum_i (x_i - \bar{x})^2}}. \quad (28)$$

Para se realizar inferência sobre os dois parâmetros simultaneamente (já que as estimativas produzidas por este método são correlacionadas), é usual considerar-se que os erros de regressão seguem distribuição normal, e com isto se usa a região de confiança definida pela elipse

$$\sum_{i=1}^n (a - \alpha)^2 + 2 \sum_{i=1}^n x_i (a - \alpha)(b - \beta) + \sum_{i=1}^n x_i^2 (b - \beta)^2 \leq 2s^2 F_{1-\alpha}(2, n-2) \quad (29)$$

Nesta expressão, assim como em (27) e (28), o erro padrão dos resíduos é representado por  $s$  e é obtido por

$$s = \frac{\sum_i (y_i - a - bx_i)^2}{n - 2} \quad (30)$$

$n$  é o número de pontos na regressão.

### 3.1.1.2 Mínimos Quadrados Ponderados

O estimador de mínimos quadrados ordinários desconsidera os erros  $\varepsilon_i$  do método de referência, e ainda considera que os erros  $\delta_i$ , do método testado são identicamente distribuídos e independentes entre si. Esta segunda suposição implica em que a variância destes resíduos é a mesma para todos os níveis avaliados. Como em geral isto não ocorre em métodos de análise química, a utilização daquele primeiro estimador resulta em estimativas viciadas para os parâmetros da reta.

A solução que é comum se encontrar na literatura é a de se ponderar as distâncias entre os pontos e a reta, o que é chamado de regressão ponderada.

### 3.1.1.2.1 Mínimos Quadrados Ponderados Simples

Este estimador corrige uma das limitações do método de mínimos quadrados ordinários, que é a que considera que a variância residual é constante para todos os níveis testados. Para isto, define-se o ponderador

$$w_i = \frac{1}{\sigma_{y_i}^2} \quad (31)$$

para a expressão  $S$  em (25), o que faz com que as equações normais produzam as seguintes expressões para o estimador:

$$b = \frac{\sum w_i (x_i - \bar{x}_w)(y_i - \bar{y}_w)}{\sum w_i (x_i - \bar{x}_w)^2} \quad \text{e} \quad a = \bar{y}_w - b\bar{x}_w, \quad (32)$$

$$s_a = \sqrt{\frac{\sum w_i x_i^2}{\sum w_i \sum w_i (x_i - \bar{x}_w)^2}} \quad (33)$$

e

$$s_b = \frac{1}{\sqrt{\sum w_i (x_i - \bar{x}_w)^2}}, \quad (34)$$

sendo

$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} \quad \text{e} \quad \bar{y}_w = \frac{\sum w_i y_i}{\sum w_i} \quad (35)$$

as médias ponderadas para os valores  $x_i$  e  $y_i$ , respectivamente.

A região de confiança conjunta para os dois parâmetros é

$$\sum_{i=1}^n w_i (a - \alpha)^2 + 2 \sum_{i=1}^n w_i x_i (a - \alpha)(b - \beta) + \sum_{i=1}^n w_i x_i^2 (b - \beta)^2 \leq 2s^2 F_{1-\alpha, 2, n-2} \quad (36)$$

### 3.1.1.2.2 Modificação de Williamson (1968) para o método de York (1966)

York (1966) propõe a minimização da soma dos quadrados  $S$  em (25), definindo o peso de regressão como

$$w_i = \frac{1}{\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2} \quad (37)$$

Desta forma, a ponderação leva em consideração os erros nos dois eixos. Para se obter este peso, que expressa o inverso da variância no ponto, parte-se da expressão da mesma, considerando independência entre os erros nos dois eixos. Da definição do problema, dada no início do presente capítulo, sabe-se que

$$x_i = \mu_{x_i} + \varepsilon_i, \text{ e } y_i = \alpha + \beta \mu_{x_i} + \delta_i = \alpha + \beta(x_i - \varepsilon_i) + \delta_i = \alpha + \beta x_i + (\delta_i - \beta \varepsilon_i).$$

No último termo à direita, a reta de regressão é expressa nas variáveis observadas diretamente,  $(x_i, y_i)$  onde o erro  $v_i = \delta_i - \beta \varepsilon_i$  possui variância

$$Var(v_i) = Var(\delta_i) + \beta^2 Var(\varepsilon_i) = \sigma_{\delta_i}^2 + \beta^2 \sigma_{\varepsilon_i}^2 = \sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2, \quad (38)$$

que é a variância no ponto.

York acha as equações normais de regressão para este caso, e após manipulações algébricas encontra a solução abaixo para  $b$ :

$$b^3 \sum_i w_i^2 \sigma_{x_i}^2 U_i^2 - 2b^2 \sum_i w_i^2 \sigma_{x_i}^2 U_i V_i - b \left[ \sum_i w_i U_i^2 - \sum_i w_i^2 \sigma_{x_i}^2 V_i^2 \right] + \sum_i w_i U_i V_i = 0 \quad (39)$$

onde

$$U_i = x_i - \bar{x}_w \quad (40)$$

e

$$V_i = y_i - \bar{y}_w \quad (41)$$

York chamou a esta expressão de “mínimos quadrados cúbica”, por ser um polinômio de ordem três em  $b$ . York cita que na verdade a expressão não é um polinômio de terceiro grau, devido à dependência de  $w_i$  em relação a  $b$ , porém ele propõe tratá-la como tal, utilizando uma estimativa inicial para  $b$ , para cálculo de  $w_i$ , que iterativamente é utilizada para se encontrar  $b$ . York apresenta uma maneira sistemática de se encontrar as raízes do polinômio, afirmando que a raiz desejada é sempre a última obtida por seu algoritmo.

O problema encontrado para se implementar a solução de York, é que o algoritmo por ele proposto é de convergência lenta, sendo que em algumas situações não há garantia de convergência. Esta, quando acontece, também depende muito do valor inicial de  $b$  utilizado para cálculo dos pesos necessários aos cálculos subseqüentes. MacTaggart e Farwell (1992) mostram, em um exemplo, os problemas de se aplicar a solução de York.

Williamson (1968) estudou a abordagem de York, e propôs outra maneira para se continuar a partir da equação cúbica obtida por este. Na verdade, ele a substituiu por

$$b \sum_i w_i z_i U_i = \sum_i w_i z_i V_i, \quad (42)$$

com 
$$z_i = w_i (\sigma_{y_i}^2 U_i + b \sigma_{x_i}^2 V_i). \quad (43)$$

Seu raciocínio se baseia no fato de que a equação “cúbica” obtida por York, não é uma equação cúbica. Escrevendo-se  $w_i$  na expressão (39) da forma como é definido em função de  $b$  e das variâncias em  $x$  e  $y$ , percebe-se que a expressão resultante apresenta  $b$  no numerador e no denominador de cada termo. Ao escrever a expressão em  $b$ , ele preferiu aproximá-la por uma função linear, que tornaria o problema de iteração e solução mais fácil. Com isto ele esperava também tornar a convergência do método mais eficiente e garantida.

Assim, Williamson apresenta as seguintes estimativas para os parâmetros e para as variâncias de  $a$  e  $b$ :

$$b = \frac{\sum w_i^2 \sigma_{y_i}^2 x_i y_i + b \sum w_i^2 \sigma_{x_i}^2 y_i^2}{\sum w_i^2 \sigma_{y_i}^2 x_i^2 + b \sum w_i^2 \sigma_{x_i}^2 x_i y_i}, \quad (44)$$

$$a = \bar{y}_w - b\bar{x}_w = \frac{\sum_i w_i y_i - b \sum_i w_i x_i}{\sum_i w_i} , \quad (45)$$

$$s_b^2 = Q^2 \sum_i w_i^2 (\sigma_{y_i}^2 U_i^2 + \sigma_{x_i}^2 V_i^2), \quad (46)$$

$$s_a^2 = \frac{1}{\sum_i w_i} + 2(\bar{x}_w + 2\bar{z})\bar{z}Q + (\bar{x}_w + 2\bar{z})^2 s_b^2 , \quad (47)$$

$$Q^{-1} = \sum_i w_i \left[ \frac{U_i V_i}{b} + 4z_i'(z_i - U_i) \right] , \quad (48)$$

$$\bar{z} = \frac{\sum_i w_i z_i}{\sum_i w_i} \quad e \quad z_i' = z_i - \bar{z} . \quad (49)$$

### 3.1.1.2.3 Mínimos Quadrados Bivariados de Riu e Rius (1996)

Este método, proposto por Lisy et al. (1990), não foi detalhado pelos mesmos autores em sua publicação. Tanto o algoritmo para sua implementação quanto a derivação da expressão para a região de confiança para os dois parâmetros da reta são apresentados por Riu e Rius (1996). Na verdade, quando esta técnica é apresentada dentro do contexto dos métodos ponderados de mínimos quadrados, a implementação da mesma fica bastante simplificada.

A alteração importante que estes autores propõem é a consideração da covariância entre os erros em  $x$  e em  $y$ , em cada ponto experimental. Eles propõem a seguinte definição para os pesos de regressão:

$$w_i = \frac{1}{\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2 - 2\beta \text{cov}(x_i, y_i)} . \quad (50)$$

Na verdade estes autores propõem um peso que é o inverso do mostrado acima, mas para efeito de uniformização de notação, no presente trabalho ele será usado da forma aqui definida. Isto faz com que se possa utilizar as fórmulas já deduzidas para o caso geral de mínimos quadrados ponderados.

Esta definição de ponderadores é a única encontrada na literatura que considera a possibilidade de haver correlação entre os erros das duas variáveis em cada ponto. Todos os

demais métodos de regressão ponderada considerando erro nos dois eixos assumem independência entre eles.

A solução encontrada por estes autores é análoga à resolução tradicional do problema de mínimos quadrados ponderados simples, expandida para incorporar o erro nos dois eixos, dada na forma matricial abaixo, que deve ser resolvida para o vetor  $\mathbf{b}^T = (a, b)$ .

$$\begin{bmatrix} \sum w_i & \sum w_i x_i \\ \sum w_i x_i & \sum w_i x_i^2 \end{bmatrix} \times \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum \left( w_i y_i - \frac{1}{2} r_i^2 \frac{\partial w_i}{\partial a} \right) \\ \sum \left( w_i x_i y_i - \frac{1}{2} r_i^2 \frac{\partial w_i}{\partial b} \right) \end{bmatrix} \quad (51)$$

Nesta expressão, o termo  $r_i = y_i - a - bx_i$ , é portanto o resíduo de regressão.

A diferença entre a resolução desta equação com a resolução tradicional encontrada em Draper e Smith (1981), por exemplo, é que o peso de regressão  $w_i$  aqui depende de  $b$ , assim como sua derivada parcial,  $\partial w_i / \partial b$ , portanto é necessária uma primeira estimativa deste parâmetro para se proceder aos cálculos, que assim se tornam um processo iterativo.

A expressão para a região de confiança para  $(\alpha, \beta)$  é então a mesma apresentada para o método de mínimos quadrados ponderados simples, usando o peso de regressão adequado.

Segundo Riu e Rius (1996), a elipse de confiança para os dois parâmetros simultaneamente possui os eixos orientados de acordo com os autovetores da matriz  $\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}$ , como definido em Draper e Smith (1981) e os comprimentos dos semi-eixos dados por

$$\sqrt{2s^2 F_{1-\alpha, 2, n-2}} \times \sqrt{\lambda_i}, \quad (52)$$

onde  $\lambda_i$  é um autovalor da matriz citada,  $i=1,2$ .

Pode-se perceber que a solução apresentada por estes autores, em (51), é de aplicação mais geral, já que nela não se especifica os pesos de regressão. Em realidade, é uma formulação mais geral, que fornece as estimativas de mínimos quadrados para todos os casos apresentados, bastando que os pesos sejam definidos de acordo (assim como suas derivadas). De forma geral, pode-se rescrever a expressão em (51) da forma geral apresentada por Draper e Smith (1981), página 79 (equação 2.11.10):

$$\mathbf{b} = (\mathbf{X}^T \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W} \mathbf{Y} \quad (53)$$

No caso de mínimos quadrados ordinários, a matriz  $\mathbf{W}$  de pesos é uma matriz identidade, para o caso de mínimos quadrados ponderados simples é uma matriz diagonal onde os elementos são obtidos pelo inverso da variância em  $y$  para cada ponto. O método de York poderia também ser avaliado por esta expressão, utilizando-se para esta matriz os pesos definidos por ele, e neste caso também é uma matriz diagonal. O método de Riu e Rius seria expresso fazendo a matriz de pesos ser uma matriz em que os elementos fora da diagonal principal possuam valor não-nulo, obtidos das covariâncias entre os valores  $x_i$  e  $y_i$ . Nos dois últimos casos, a matriz  $\mathbf{X}^T \mathbf{W} \mathbf{Y}$  dada por Draper e Smith precisa ser modificada para considerar os erros em  $x$ , pela introdução das derivadas parciais como apresentado em (51). Também para estes casos a solução da equação é iterativa, já que os pesos  $w_i$  (assim como suas derivadas) dependem dos parâmetros.

### 3.1.2. Métodos baseados em Máxima Verossimilhança

Usualmente os autores que utilizam o método de máxima verossimilhança para a regressão linear fazem uma distinção entre dois casos, o caso *estrutural* e o caso *funcional*. A diferença entre estas duas situações é que na primeira os “verdadeiros valores”  $X$  e  $Y$  são considerados variáveis aleatórias, e na segunda são considerados valores fixos. Linnet (1990) cita que esta distinção é de menor importância na prática, porque os cálculos são os mesmos nos dois casos.

A expressão da função de verossimilhança para um conjunto de  $n$  pares de valores  $(x_i, y_i)$ , dado o modelo funcional apresentado no início deste capítulo é apresentada a seguir:

$$L(\alpha, \beta, \mu_{x_i} | x_i, y_i) = \prod_{i=1}^n \left\{ \frac{1}{(2\pi\sigma_{y_i}^2)^{1/2}} \times \frac{1}{(2\pi\sigma_{x_i}^2)^{1/2}} \right\} \times \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left[ \frac{(x_i - \mu_{x_i})^2}{\sigma_{x_i}^2} + \frac{(y_i - \alpha - \beta\mu_{x_i})^2}{\sigma_{y_i}^2} \right] \right\} \quad (54)$$

Na montagem da expressão acima, considera-se que os erros  $\varepsilon_i$  do método de referência, e  $\delta_i$ , do método testado, seguem distribuição normal com média zero e variâncias  $\sigma_{x_i}^2 = \sigma_{\varepsilon_i}^2$  e  $\sigma_{y_i}^2 = \sigma_{\delta_i}^2$ , respectivamente.

Bofarine, Rodrigues e Cordani (1992) mostram que a formulação deste problema como acima possui número de parâmetros que cresce com o número de observações, pois os “verdadeiros” valores  $\mu_{x_i}$  precisam ser também estimados. Estes parâmetros são por eles chamados de *incidentais*, associados à  $i$ -ésima observação. Da forma como apresentada, também citando outros autores, mostram que não possui solução no espaço paramétrico, a não ser que se faça mais alguma suposição adicional.

Dependendo desta suposição ou suposições adicionais encontradas, tem-se diferentes estimadores.

### 3.1.2.1 Razão de variâncias constante e conhecida

A primeira variante obtida pelo método de máxima verossimilhança é chamada por Paiva Lima (1996) de modelo funcional clássico de máxima verossimilhança com razão de variâncias conhecida.

Para este caso, supõe-se, em adição ao que já foi considerado, que as variâncias de ambas as variáveis sendo comparadas são constantes, e conhece-se sua razão,  $\lambda$ . Esta razão é definida como

$$\lambda = \frac{\sigma_y^2}{\sigma_x^2} = \frac{\sigma_\delta^2}{\sigma_\varepsilon^2} = \frac{\lambda\theta}{\theta}, \quad (55)$$

se chamarmos  $\sigma_{x_i}^2$  de  $\theta$ .

Fazendo estas considerações, a função de verossimilhança em (54) pode ser reescrita como

$$L(\alpha, \beta, \theta, \mu_{x_i} | x_i, y_i) = \frac{1}{(2\pi\theta)^n \lambda^{n/2}} \exp\left\{-\frac{1}{2\lambda\theta} \sum_{i=1}^n [\lambda(x_i - \mu_{x_i})^2 + (y_i - \alpha - \beta\mu_{x_i})^2]\right\}. \quad (56)$$

O seu logaritmo é dado por

$$l(\alpha, \beta, \theta, \mu_{x_i} | x, y) = -n \log(2\pi\theta) - \frac{n}{2} \log(\lambda) - \frac{1}{2\lambda\theta} \left[ \sum_{i=1}^n \lambda(x_i - \mu_{x_i})^2 + \sum_{i=1}^n (y_i - \alpha - \beta\mu_{x_i})^2 \right]. \quad (57)$$

Como se pode perceber, não se dispõe de graus de liberdade suficientes para estimar consistentemente os parâmetros  $\alpha$ ,  $\beta$ ,  $\theta$ , e os  $n$  valores  $\mu_{x_i}$ . Uma possível saída para este problema pode ser vista em Garthwaite, Jolliffe e Jones (1995), pp.56. Estes autores chamam este problema de “inferência na presença de parâmetros perturbadores”. Quando se deseja estimar parte dos parâmetros de um modelo, não se interessando pelos demais, pode-se encontrar a função de verossimilhança perfilada. Para obtê-la no presente caso, deve-se primeiramente maximizar a função log-verossimilhança em (57) com relação aos parâmetros perturbadores (ou também chamados incidentais).

Isto fornece a seguinte solução:

$$\hat{\mu}_{x_i} = \frac{\lambda x_i + \beta(y_i - \alpha)}{\lambda + \beta^2}. \quad (58)$$

Substituindo-se esta expressão em (57) no lugar de  $\mu_{x_i}$  obtém-se a função log-verossimilhança perfilada, que é função apenas de  $\alpha$ ,  $\beta$ , e  $\theta$ .

Derivando-se a nova expressão em relação a estes três parâmetros e igualando as derivadas a zero, encontramos:

$$b = \frac{(S_{yy} - \lambda S_{xx}) + \sqrt{(\lambda S_{xx} - S_{yy})^2 + 4\lambda S_{xy}^2}}{2S_{xy}}, \quad (59)$$

$$a = \bar{y} - b\bar{x} \quad (60)$$

e

$$\hat{\theta} = \frac{S_{yy} + b^2 S_{xx} - 2b S_{xy}}{2n(\lambda + b^2)}, \quad (61)$$

com

$$S_{xx} = \sum_i (x_i - \bar{x})^2, \quad (62)$$

$$S_{yy} = \sum_i (y_i - \bar{y})^2 \quad (63)$$

$$e \quad S_{xy} = \sum_i (x_i - \bar{x})(y_i - \bar{y}). \quad (64)$$

Para se realizar inferências criando uma região de confiança, deve-se utilizar a matriz de covariâncias dos estimadores, e através dela se pode construir um teste de hipótese do tipo Wald, da seguinte forma.

$$W = n(\hat{\Phi} - \Phi_0)' \Omega_n^{-1} (\hat{\Phi} - \Phi_0), \quad (65)$$

em que  $\Phi$  é o vetor de parâmetros de interesse, sendo  $\hat{\Phi}$  seu estimador,  $\Phi_0$  o valor assumido na hipótese básica (ou nula) sendo testada e  $\Omega_n^{-1}(\hat{\Phi})$  é o inverso da matriz de covariâncias de  $\hat{\Phi}$  avaliada em  $\hat{\Phi}$ .

Patefield (1977) deriva a expressão correta para esta matriz, que é a seguinte:

$$\Omega_n(\hat{\Phi}) = \left\{ \left[ (\lambda + b^2) \theta b \right] / (n s_{xy}) \begin{bmatrix} \bar{x}^2 (1 + \hat{\tau}) + (s_{xy} / b) & -\bar{x}(1 + \hat{\tau}) \\ -\bar{x}(1 + \hat{\tau}) & (1 + \hat{\tau}) \end{bmatrix} \right\} \quad (66)$$

sendo  $s_{xy} = S_{xy} / n$  e  $\hat{\tau} = \lambda \theta b / \left\{ (\lambda + b^2) s_{xy} \right\}$ .

Este autor apresenta a matriz apenas para a parte que se refere a  $a$  e  $b$ . Para utilizar esta expressão, deve-se substituir  $\hat{\theta}$  pelo estimador consistente  $\frac{2n}{n-2} \hat{\theta}$ .

Gálea-Rojas *et al.* (2003) trabalham com uma generalização de (65) e (66).

### 3.1.2.2 Razão entre Coeficientes de Variação constante

Escrevendo as expressões para os coeficientes de variação em  $X$  e em  $Y$ , temos:

$$CV_x = f_x = \frac{\sigma_{x_i}}{\left( \frac{\mu_{x_i} + \mu_{y_i}}{2} \right)} \quad \forall i \quad \text{ou} \quad f_x = \sigma_{x_i} \cdot w^{1/2} \quad (67)$$

fazendo  $w = \frac{1}{\left( \frac{\mu_{x_i} + \mu_{y_i}}{2} \right)^2} \quad (68)$

Analogamente,  $f_y = \sigma_{y_i} \cdot w^{1/2} \quad (69)$

Definido desta forma, percebe-se que, nesta abordagem, proposta por Linnet (1990), o coeficiente de variação (tanto em  $X$  quanto em  $Y$ ) é a razão entre o desvio padrão e a média entre  $\mu_{x_i}$  e  $\mu_{y_i}$ . Sob  $H_0$  (19), este CV realmente será o CV de cada método, i.e.,  $CV_x = \sigma_{x_i} / \mu_{x_i}$  e  $CV_y = \sigma_{y_i} / \mu_{y_i}$ .

Fazendo  $\lambda = \frac{f_y^2}{f_x^2}$  (pode-se verificar que  $\lambda = \frac{\sigma_{y_i}^2}{\sigma_{x_i}^2}$ ), que é a razão entre os quadrados dos coeficientes de variação, suposta constante neste caso particular, e substituindo na função verossimilhança em (54), vem

$$L = \prod_{i=1}^n \left\{ \frac{1}{(2\pi f_y^2 / w_i)^{1/2}} \times \frac{1}{(2\pi f_y^2 / \lambda w_i)^{1/2}} \right\} \exp \left\{ -\frac{1}{2f_y^2} \sum_{i=1}^n [\lambda w_i (x_i - \mu_{x_i})^2 + w_i (y_i - \alpha - \beta \mu_{x_i})^2] \right\} \quad (70)$$

Para obter a verossimilhança perfilada, então, encontra-se  $\frac{\partial L}{\partial \mu_{x_i}}$ . Igualando esta derivada a zero e efetuando simplificações algébricas, chega-se a

$$\hat{\mu}_{x_i} = \frac{\lambda x_i + \beta(y_i - \alpha_0)}{\lambda + \beta^2} \quad (71)$$

Linnet (1990) centra os valores da variável  $X$  na média:

$$\mu_{y_i} = \alpha_0 + \beta \mu_{x_i} = \alpha + \beta(\mu_{x_i} - \bar{\mu}_x), \quad (72)$$

onde  $\bar{\mu}_x$  é a média de todos os valores  $\mu_{x_i}$ .

Após substituir a expressão (71) em (70), derivando em relação aos demais parâmetros e executando-se as simplificações necessárias, chega-se à seqüência de equações proposta por Linnet para se encontrar a solução

$$b = \frac{S_{wyy} - \hat{\lambda} S_{wxx} + \sqrt{(\hat{\lambda} S_{wxx} - S_{wyy})^2 + 4\hat{\lambda} S_{wxy}}}{2S_{wxy}}, \quad (73)$$

onde

$$S_{wxx} = \sum_i w_i (x_i - \bar{x}_w)^2, \quad (74)$$

$$S_{wyy} = \sum_i w_i (y_i - \bar{y}_w)^2 \quad (75)$$

e

$$S_{wxy} = \sum_i w_i (x_i - \bar{x}_w)(y_i - \bar{y}_w), \quad (76)$$

sendo 
$$\bar{x}_w = \frac{\sum w_i x_i}{\sum w_i} . \quad (77)$$

Os pesos  $w_i$  são estimados por

$$\hat{w}_i = \frac{1}{\left( \frac{\hat{X}_i + \hat{Y}_i}{2} \right)^2}, \quad (78)$$

com

$$\hat{X}_i = \frac{\hat{\lambda}x_i + b(y_i - a_0)}{\hat{\lambda} + b^2} \quad (79)$$

$$\hat{Y}_i = y_i - \frac{\hat{\lambda}d_i}{\hat{\lambda} + b^2} \quad (80)$$

$$d_i = y_i - (a_0 + bx_i) \quad (81)$$

e, finalmente,

$$a = \bar{y}_w . \quad (82)$$

Após a primeira estimativa dos parâmetros da reta como acima, estes valores são levados a (72), para se encontrar  $a_0$ . Estes valores são então levados às equações (73)-(82) onde novas estimativas são realizadas, de forma iterativa, até convergência. De início, deve-se utilizar uma primeira estimativa para os dois parâmetros da reta, por exemplo utilizando as fórmulas para mínimos quadrados ordinários.

### 3.1.2.3 Relação Funcional de Máxima Verossimilhança para variâncias não homogêneas

Este é o caso mais genérico. A proposta em Ripley e Thompson (1987) considera as variâncias em cada ponto ( $\sigma_{x_i}^2$  e  $\sigma_{y_i}^2$ ) conhecidas.

A função log-verossimilhança obtida a partir de (54) é

$$l(\alpha, \beta, \mu_{x_i} | x, y) = \text{constante} - \frac{1}{2} \sum \frac{(x_i - \mu_{x_i})^2}{\sigma_{x_i}^2} - \frac{1}{2} \sum \frac{(y_i - \alpha - \beta \mu_{x_i})^2}{\sigma_{y_i}^2} . \quad (83)$$

Obtém-se inicialmente a estimativa para os parâmetros incidentais

$$\hat{\mu}_{x_i} = w_i[\sigma_{y_i}^2 x_i + \sigma_{x_i}^2 \beta(y_i - \alpha)], \quad (84)$$

sendo  $w_i = (\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2)^{-1}$ . (85)

Substituindo na função log-verossimilhança e simplificando, a função log-verossimilhança perfilada fica da forma

$$l(\alpha, \beta) = \text{constante} - \sum_{i=1}^n \frac{(y_i - \alpha - \beta x_i)^2}{\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2}. \quad (86)$$

O problema passa a ser então a minimização de

$$Q_m(\alpha, \beta) = \min_{x_i} Q = \sum w_i (y_i - \alpha - \beta x_i)^2, \quad (87)$$

considerando que  $\hat{\alpha} = a$  e  $w_i$  dependem de  $\hat{\beta} = b$ .

Pode-se perceber que esta expressão é igual à proposta por Riu e Rius (1996), já apresentada no ítem de métodos de mínimos quadrados, se assumirmos que não há covariância entre  $x_i$  e  $y_i$ . Entretanto, Ripley e Thompson (1987) não solucionam esta expressão algebricamente da mesma forma que aqueles autores. Ao invés disto, fazem todas as substituições na expressão em (87) de forma a torná-la dependente apenas de  $\beta$ .

A estimativa para  $\alpha$  é:

$$a = \frac{\sum w_i (y_i - \beta x_i)}{\sum w_i} \quad (88)$$

Esta expressão é deduzida do fato de que a reta de regressão passa pelo ponto de coordenadas  $(\bar{x}_w, \bar{y}_w)$ , médias ponderadas em  $x$  e em  $y$ , respectivamente.

Assim, algum método numérico de otimização (minimização) deve ser utilizado para tratar a expressão

$$Q_m(\hat{\alpha}, \beta) = \sum w_i(\beta)[y_i - \hat{\alpha}(\beta) - \beta x_i]^2, \quad (89)$$

que deve ser reescrita incorporando as relações (85) e (88):

$$Q_m(\beta) = \sum \frac{\left[ y_i - \frac{\sum (\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2)^{-1} (y_i - \beta x_i)}{\sum (\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2)^{-1}} - \beta x_i \right]^2}{\sigma_{y_i}^2 + \beta^2 \sigma_{x_i}^2}, \quad (90)$$

que é uma expressão que considera apenas os dados e constantes conhecidas, em função de  $\beta$ .

Para o teste de  $\beta = 1$ , Ripley e Thompson sugerem que se calcule

$$\frac{(b-1)}{\sqrt{\sum w_i (x_i - \bar{x}_w)^2}}. \quad (91)$$

Para testar se  $\alpha = 0$ , utiliza-se a estatística

$$\frac{a}{\sqrt{\frac{\sum w_i x_i^2}{\sum w_i \sum w_i (x_i - \bar{x}_w)^2}}}. \quad (92)$$

Estas duas expressões, como se viu na primeira seção deste capítulo, representam os testes usuais para o método de mínimos quadrados ponderados. Estas duas estatísticas são comparadas com uma distribuição normal padronizada com nível de significância especificado. Os autores recomendam, para uma análise apropriada da regressão, testar os resíduos

$$r_i^* = (y_i - a - bx_i) \sqrt{w_i(b)}. \quad (93)$$

A soma dos quadrados dos resíduos padronizados, assim produzidos, deveria fornecer a relação aproximada

$$\sum (r_i^*)^2 \approx (n-2). \quad (94)$$

Isto se dá porque os resíduos, uma vez padronizados, seguem uma distribuição  $N(0,1)$ . A soma dos quadrados deles é obtida com  $(n-2)$  graus de liberdade, e é uma variável com

distribuição aproximada Qui-quadrado. A soma de variáveis qui-quadrado é uma variável com distribuição qui-quadrado com média igual ao número de graus de liberdade.

Quando se encontra um valor para esta soma ultrapassando  $(n-2)$ , é uma indicação de que as variâncias utilizadas para cada método, em cada ponto, são subestimadas. Além disso, o gráfico  $r_i^*$  vs.  $\hat{X}_i$  pode fornecer informações adicionais sobre a adequação do ajuste.

### 3.1.3 Métodos Robustos

Vários métodos são sugeridos na literatura de química analítica, quando se aborda o tema ‘comparação de métodos’, que procuram prover soluções para quando não se deseja fazer grandes suposições sobre a distribuição de probabilidade dos dados, ou se suspeita existirem resultados discrepantes ou contaminados.

Foram escolhidos para representar este grupo de métodos no atual trabalho, os estimadores LTS (“Least trimmed of squares”) de Rousseeuw e Leroy (1987) e a Regressão Estrutural Absoluta de Feldmann (1992).

O estudo mais aprofundado do uso de estimadores robustos em comparação de métodos analíticos fica como sugestão para trabalhos futuros e para outros pesquisadores.

#### 3.1.3.1 Mínimos Quadrados Aparados de Rousseeuw e Leroy (1987)

Este método tem como princípio, na verdade, o mesmo do método de mínimos quadrados ordinários. A mudança proposta é que, ao invés de se somar todos os erros quadrados de regressão, toma-se apenas  $h$  destes termos, eliminando os maiores termos da soma quadrática. Para se conseguir isto, deve-se primeiramente calcular todas as retas possíveis, tomando as  $n$  observações duas a duas.

Rousseeuw e Leroy (1987) então propõem que sejam ordenados os quadrados dos erros, em cada reta assim obtida, do menor para o maior, como abaixo:

$$(r^2)_{1:n} \leq (r^2)_{2:n} \leq \dots \leq (r^2)_{n:n} . \quad (95)$$

A expressão acima simboliza os quadrados dos resíduos de regressão, tomados de uma das possíveis retas calculadas, ordenados do menor para o maior. Os subscritos 1:n, 2:n, ..., n:n significam o menor quadrado de resíduo de  $n$  resíduos, o segundo menor de  $n$ , e assim por diante.

Para se obter a soma destes resíduos quadrados, escolhe-se um valor  $h$ , a partir do qual os termos da soma são desprezados ou aparados. Soma-se apenas até o  $h$ -ésimo resíduo quadrado ordenado.

O mesmo então é repetido para uma nova reta, produzida por outros dois pontos quaisquer dentre as observações disponíveis. A cada reta estudada, tem-se então uma soma de quadrados aparada. Aquela que fornecer o menor valor para esta soma, é a reta escolhida.

O critério de Rousseeuw, pode ser então resumido como tendo a seguinte função objetivo:

$$\min_{\theta} \sum_{i=1}^h (r^2)_{i:n} \quad (96)$$

Fazendo  $h=[n/2]+1$ , Rousseeuw demonstra que o estimador LTS atinge o máximo ponto de quebra, de 50%.

A expressão  $h=[n/2]+1$  é também conhecida como sendo a posição da “mediana alta”, e significa a posição do maior dos dois valores centrais, quando  $n$  é par. Quando  $n$  é ímpar,  $h$  representa a posição da própria mediana.

### 3.1.3.2 Regressão estrutural absoluta de Feldmann (1992)

Para se entender este estimador, proposto por Feldmann (1992), deve-se entender antes o estimador não robusto apresentado em seu artigo.

Feldmann parte de um método de máxima verossimilhança, em que se considera os pares  $(x_i, y_i)$  como sendo realizações independentes de uma distribuição normal bivariada  $(X, Y)$  tal que

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left( \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{bmatrix} \right) \quad (97)$$

A densidade para esta situação é então dada por

$$f(x, y) = \frac{1}{\gamma} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[ \frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right] \right\}, \quad (98)$$

$$\text{com } \gamma = 2\pi\sigma_x\sigma_y\sqrt{1-\rho^2} \quad (99)$$

e  $\rho$  é o coeficiente de correlação linear de Pearson entre  $x$  e  $y$ .

Colocado desta forma o problema, fica claro que Feldmann se dedica apenas ao caso em que os erros nos dois eixos possuem variâncias constantes, independentes do teor.

Feldmann assume que os valores  $(\hat{X}_i, \hat{Y}_i)$  são os valores preditos pela reta de regressão, de forma que

$$\hat{Y}_i = a + bx_i \quad \text{e} \quad y_i = a + b\hat{X}_i \quad . \quad (100)$$

Feldmann também assume que os valores preditos são equivalentes aos valores observados até seus momentos de segunda ordem. Com isto,  $(\hat{X}_i, \hat{Y}_i)$  possuem a mesma distribuição de probabilidades dada em (98).

Com base em (100) e nas suposições acima sobre os momentos, pode-se encontrar que

$$\alpha = \mu_x - \beta\mu_y \quad , \quad (101)$$

$$\beta^2 = \frac{\sigma_y^2}{\sigma_x^2} \quad (102)$$

e conseqüentemente,

$$\sigma_y = |\beta|\sigma_x \Rightarrow \sigma_x^2 = |\beta|^{-1}\sigma_y\sigma_x \quad \text{e} \quad \sigma_y^2 = |\beta|\sigma_x\sigma_y \quad . \quad (103)$$

Escrevendo  $f(\hat{X}_i, \hat{Y}_i)$  a partir de (98), substituindo por  $(x_i, y_i)$  usando (100) e usando as relações em (101), (102) e (103), vem

$$f(\hat{X}_i, \hat{Y}_i) = \frac{1}{\gamma} \exp \left\{ \frac{-\left[ |\beta|^{-1}(y_i - \mu_y)^2 - 2\rho(y_i - \mu_y)(x_i - \mu_x) + |\beta|(x_i - \mu_x)^2 \right]}{2\sigma_y\sigma_x(1 - \rho^2)} \right\} \quad , \quad (104)$$

que é uma expressão que representa a dependência da distribuição de probabilidades com relação ao parâmetro  $\beta$ , da reta a estimar.

A partir desta expressão pode-se chegar então à função log-verossimilhança

$$l(\mu_x, \mu_y, \beta) = \frac{-\left\{ \psi^2(\mu_x, \mu_y, \beta) - 2\rho\psi_0(\mu_x, \mu_y) \right\}}{2\sigma_x\sigma_y(1 - \rho^2)} - n \log(\gamma) \quad , \quad (105)$$

com 
$$\psi^2(\mu_x, \mu_y, \beta) = \sum_{i=1}^n |\beta| (x_i - \mu_x)^2 + |\beta|^{-1} (y_i - \mu_y)^2 \quad (106)$$

e 
$$\psi_0(\mu_x, \mu_y) = \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (107)$$

Maximizando a expressão em (105) com relação aos parâmetros, tem-se as estimativas

$$b = \text{sign}(r_{xy}) \frac{s_y}{s_x}, \quad (108)$$

$$\hat{\mu}_x = \bar{x} \quad (109)$$

e 
$$\hat{\mu}_y = \bar{y}, \quad (110)$$

em que  $r_{xy} = \hat{\rho}$  é a estimativa do coeficiente de correlação linear.

Para derivar uma alternativa robusta para o estimador acima apresentado, Feldmann propõe substituir, em (108), a razão de desvios padrão pela mediana da razão absoluta entre os erros robustos em  $y$  e em  $x$ , definidos da seguinte maneira:

$$b = \pm \text{med}_i \left\{ \frac{|y_i - \text{med}_k(y_k)|}{|x_i - \text{med}_k(x_k)|} \right\} \quad (111)$$

A interseção da reta de regressão é obtida, por sua vez, por

$$a = \text{med}_i (y_i - bx_i), \quad (112)$$

e o sinal de  $b$  pode ser encontrado pelo sinal de  $\tilde{b}$ :

$$\tilde{b} = \text{med}_i \left\{ \frac{y_i - \text{med}_k(y_k)}{x_i - \text{med}_k(x_k)} \right\}. \quad (113)$$

Feldmann comenta que o estimador  $\tilde{b}$  apresentado acima é apresentado inicialmente por Hampel, que o desconsidera por apresentar resultados pouco satisfatórios em muitas situações. Feldmann só o utiliza para determinação do sinal.

Feldmann apresenta em seu artigo a forma de se realizar a análise dos resíduos de regressão para detecção de pontos discrepantes, assim como exemplos de aplicação onde conclui que este estimador apresenta resultados satisfatórios quando os dados forem contaminados.

### 3.1.4 Método de Wald (1940)

Este método, apesar de ser mais antigo que várias propostas apresentadas anteriormente, é apresentado em separado, já que não se enquadra em nenhum dos princípios já apresentados, utilizados para se encontrar estimadores, quais sejam: mínimos quadrados, máxima verossimilhança e estatísticas de ordem (métodos robustos).

Wald (1940) apresenta uma alternativa de estimador para o problema de regressão com erros nas variáveis, que não é justificado por nenhum princípio. Wald apresenta o estimador e então se preocupa em demonstrar sua consistência.

Wald faz as seguintes suposições básicas:

- i) As variáveis aleatórias  $\varepsilon_i$  em (15) têm cada uma a mesma distribuição e são não correlacionadas. A variância  $\sigma_\varepsilon^2$  é finita;
- ii) As variáveis aleatórias  $\delta_i$  em (16) têm cada uma a mesma distribuição e são não correlacionadas. A variância  $\sigma_\delta^2$  é finita;
- iii) As variáveis aleatórias  $\varepsilon_i$  e  $\delta_i$  são não correlacionadas;
- iv) Existe uma relação linear simples entre os valores verdadeiros  $\mu_{x_i}$  e  $\mu_{y_i}$ ;
- v) O limite inferior de

$$\left| \frac{(X_1 + \dots + X_m) - (X_{m+1} + \dots + X_n)}{n} \right| \quad (n = 2, 3, \dots)$$

é positivo. Os valores de  $X_i$  aqui estão ordenados.

Wald define as seguintes expressões:

$$b_1 = \frac{(x_1 + \dots + x_m) - (x_{m+1} + \dots + x_n)}{n} \quad (114)$$

e

$$b_2 = \frac{(y_1 + \dots + y_m) - (y_{m+1} + \dots + y_n)}{n}. \quad (115)$$

com  $m = n/2$  e  $n$  é o número de pontos da regressão. Por simplicidade, Wald considera, para apresentar seu estimador, que  $n$  é um número par.

O estimador proposto para o coeficiente angular da reta é então

$$b = \frac{b_1}{b_2} = \frac{(y_1 + \dots + y_m) - (y_{m+1} + \dots + y_n)}{(x_1 + \dots + x_m) - (x_{m+1} + \dots + x_n)}. \quad (116)$$

A interseção é obtida a partir do ponto  $(\bar{x}, \bar{y})$ , que é um ponto da reta.

É fácil ver que

$$E(b_1) = \frac{(X_1 + \dots + X_m) - (X_{m+1} + \dots + X_n)}{n}$$

e

$$E(b_2) = \frac{(Y_1 + \dots + Y_m) - (Y_{m+1} + \dots + Y_n)}{n}.$$

Usando a condição iv dada acima, pode-se ver que

$$E(b_2) = \beta E(b_1) \quad \text{ou} \quad \beta = \frac{E(b_2)}{E(b_1)}. \quad (117)$$

A variância de  $b_1 - E(b_1)$  é igual a  $\sigma_\epsilon^2/n$  e a variância de  $b_2 - E(b_2)$  é igual a  $\sigma_\delta^2/n$ . Assim, Wald mostra que  $b_1$  e  $b_2$  convergem para  $E(b_1)$  e  $E(b_2)$ , respectivamente. Do pressuposto v, acima, Wald conclui que a razão  $b_2/b_1$  converge para  $E(b_2)/E(b_1) = \beta$ .

Wald apresenta em seu trabalho também expressões que ele deduz para estimar as variâncias das estimativas, e logo os intervalos de confiança para os parâmetros  $\alpha$  e  $\beta$ .



É importante verificar-se que a proposta de Wald assume um modelo linear entre os resultados dos dois métodos sendo comparados (se aplicado a este problema específico), cujas variâncias são constantes ao longo da faixa de concentrações.

Obviamente, isto se torna uma limitação em sua aplicação. Como no caso de comparação de métodos de química analítica isto não é em geral verdadeiro, espera-se produzir com a aplicação deste estimador a casos de aplicação, estimativas viciadas.

Entretanto, o método de Wald é uma referência importante na literatura de regressão com erros nas variáveis, além do que, Danzer et al (1995) o utilizam comparando em simulações com quatro outros métodos, e ele apresenta os melhores resultados.

## Capítulo 4

### Aplicações

Este capítulo se desenvolve em duas linhas importantes. Em primeiro lugar, são realizados estudos de Monte Carlo, baseados em casos reais da indústria mineral, em particular da Companhia Vale do Rio Doce (CVRD). Nestes estudos se pretende verificar o desempenho comparativo dos vários estimadores descritos no capítulo 3. Num segundo momento, os estimadores são aplicados a problemas reais, alguns tirados da literatura específica da química analítica e relativa a modelos de regressão com erros nas variáveis. Outros exemplos reais são fornecidos pela experiência do autor em projetos de trabalho da CVRD. Os estimadores comparados são: mínimos quadrados ordinários, mínimos quadrados ponderados, o estimador de York (1966) modificado por Williamson (1968), mínimos quadrados bivariados de Riu e Rius (1996), os modelos de máxima verossimilhança propostos por Deming (1943), Linnet (1990) e Ripley e Thompson (1987), além dos métodos robustos propostos por Rousseeuw e Leroy (1987) e Feldmann (1992) e o método de Wald (1940).

#### 4.1 SIMULAÇÕES

Para se realizar as simulações a seguir, baseou-se em uma rotina criada usando a linguagem de programação do S-Plus 2000 ©, que produz números pseudo-aleatórios de acordo com especificações definidas no início do estudo.

O simulador, basicamente, funciona gerando um par de vetores, X e Y, que serão usados para cálculos de regressão, utilizando os estimadores escolhidos. Os parâmetros que precisam ser especificados para que ele gere os dados são:

- a) tamanho (n) dos vetores X e Y a serem produzidos;
- b) Limites inferior e superior para X;
- c) Distribuição de frequência para os dados X (uniforme ou log-normal);
- d) Parâmetros  $\alpha$  e  $\beta$ , da reta que relaciona Y com X;
- e) Tipo de modelo matemático que relaciona o desvio padrão do erro com a concentração (linear ou potência); este é especificado para X e para Y;
- f) Parâmetros para as funções matemáticas do item anterior ( $\sigma_{0_x}, \sigma_{0_y}, \kappa_x, \kappa_y$ ).

A nomenclatura usada aqui está de acordo com as definições utilizadas ao longo desta dissertação, e resumidas na Lista de Fórmulas e Símbolos.

Além dos valores de X e Y, gerados, o simulador também devolve o desvio padrão do erro em X e em Y, para cada ponto gerado. Todas as simulações feitas consideram que a variância do erro em X e em Y são conhecidas, quando estas informações são necessárias ao estimador. Por isto, os valores usados são os especificados para gerar os vetores aleatórios.

O programa desenvolvido para gerar os vetores aleatórios pode ser visto no ANEXO.

#### 4.1.1. Avaliação de diversos fatores sobre os estimadores de regressão. Estudo baseado na análise de cobre em minérios

Neste primeiro estudo, foi escolhida uma faixa de concentrações variando de 0,01 a 10% de cobre contido, e os erros variaram de acordo com a prática do mercado, para análises químicas deste bem mineral. Foi assumido o modelo de coeficiente de variação constante, e o que se vê na prática é que este varia entre 0,5 e 10% em análises químicas para a faixa de concentrações citada.

A avaliação foi concebida considerando quatro características dos dados:

- i) vício na inclinação  $\beta$ ;
- ii) distribuição dos dados X;
- iii) CV para a variável X;
- iv) CV para Y.

O vício na variável Y é especificado através da atribuição de valores para os parâmetros  $\alpha$  e  $\beta$ , da reta que a relaciona com a variável X. Para efeito deste estudo,  $\alpha=0$  em todas as simulações (é o que mais se encontra na prática da mineração e da pesquisa geológica). Já o parâmetro  $\beta$  assume nas simulações dois valores, 1 e 1,05. Isto representa uma situação sem vício ( $\beta=1$ ) e uma situação em que Y possui vício proporcional 5% acima de X ( $\beta=1,05$ ).

Para a distribuição de probabilidade de X, foram tomados dois casos base. O caso em que os dados a serem comparados se referem a dois laboratórios químicos analisando as mesmas amostras em um programa de sondagem geológica para cálculo de recursos em um projeto avançado (estudo conceitual, de pré-viabilidade ou viabilidade), em que os teores se distribuem de forma aproximadamente log-normal. E o caso de validação de métodos de análise química em laboratório, onde se escolhe amostras com teor aproximado conhecido, procurando cobrir a faixa de teores de interesse de forma equiprovável. Neste segundo caso, os dados então são gerados de acordo com uma distribuição uniforme.

Os CVs, tanto para X quanto para Y, foram assumidos ora valendo 0,8% ( $\kappa = 0,008$ ), ora 8% ( $\kappa = 0,08$ ), que representam situações quase extremas de precisão analítica, uma considerada muito boa, e outra considerada fora do desejado, embora ainda presente na realidade de mercado.

Como cada uma das quatro variáveis de especificação citadas para se gerar os dados simulados pode assumir dois níveis, estas simulações representam um estudo fatorial  $2^k$ , com  $k=4$ , portanto são 16 combinações possíveis para estas especificações. Cada uma destas 16 combinações é chamada de um teste. Cada um dos 16 testes foi repetido mil

vezes. Para cada uma das 16 mil simulações, foram aplicados os dez estimadores descritos no capítulo 3.

Alguns estimadores propostos utilizam testes de hipótese separados para os parâmetros  $\alpha$  e  $\beta$  da reta de regressão. Alguns utilizam um teste conjunto. No presente estudo, foram utilizados os testes de hipótese sugeridos pelos próprios autores, com exceção do método de Deming, que é muito citado na literatura, e cada autor realiza de uma forma diferente. Para este, foi escolhido o teste conjunto, utilizando uma estatística do tipo Wald, usando a matriz de covariâncias proposta por Patefield, citada no capítulo 3, equação (66).

Cada teste foi gerado com  $n = 500$  observações, mas foram realizados cálculos de regressão apenas para  $n = 10$  e  $n = 100$ . Um resumo das condições dos 16 testes pode ser visto na tabela abaixo.

Tabela 4.1 – Condições estabelecidas para as simulações

Teste	1	2	3	4	5	6	7	8
$\beta$	1	1	1	1	1,05	1,05	1,05	1,05
$CV_x$	0,008	0,08	0,08	0,008	0,008	0,08	0,08	0,008
$CV_y$	0,008	0,08	0,008	0,08	0,008	0,08	0,008	0,08
Distribuição	“u”	“u”	“u”	“u”	“u”	“u”	“u”	“u”

Para os testes 9 a 16, a especificação é igual aos testes 1 a 8, respectivamente, porém substituindo a distribuição por “ln” (dados log-normais ao invés de uniformes “u”).

Para se realizar a comparação entre os estimadores aplicados a estes dados simulados, foram calculadas algumas estatísticas auxiliares. Para cada teste, foram calculados o número de vezes que a hipótese básica (de igualdade entre Y e X) foi aceita. Com base nisto, pode-se calcular o nível de significância empírico (para as simulações em que  $\beta=1$ ) e o poder do teste (quando  $\beta = 1,05$ ). Também foram calculados os erros quadráticos médios, em cada um dos 16 testes.

Para cada simulação, foi calculada a distância euclidiana entre o ponto (a, b) e o ponto ( $\alpha$ ,  $\beta$ ), representados no espaço paramétrico. Esta distância representa uma espécie de soma vetorial entre os erros quadráticos médios para os dois parâmetros da reta estimados. Portanto, estas distâncias permitem uma comparação entre estimadores, de quanto suas estimativas se aproximam do valor verdadeiro dos parâmetros.

Para alguns estimadores, dispõe-se de uma região de confiança para o teste de hipótese, que consiste de um elipsóide nos dois parâmetros. Para comparar estes estimadores, foram calculadas as áreas destes elipsóides.

Entende-se que o estimador ideal seria aquele que produz estimativas mais próximas dos valores reais dos parâmetros (as menores distâncias euclidianas), com as menores

áreas elípticas para a região de confiança. Todos os testes de hipótese foram realizados considerando-se um coeficiente de confiança de 95%.

As Tabelas 4.2, 4.3 e 4.4 mostram um resumo dos testes. Na primeira delas se mostra a quantidade de vezes que  $H_0$  foi aceita. Na segunda, os valores de EQM(a) – erros quadráticos médios para a interseção da reta. Na terceira, os valores de EQM(b) – erros quadráticos médios para a inclinação da reta calculada. A codificação dos estimadores utilizada ao longo deste capítulo é a que segue:

**MQ** : mínimos quadrados ordinários;  
**MQP** : mínimos quadrados ponderados;  
**YW** : método de York-Williamson;  
**RR** : método proposto por Riu e Rius;  
**Deming** : método de Deming;  
**Linnet** : método de CV constante proposto por Linnet  
**RT** : método proposto por Ripley e Thompson;  
**MQA** : mínimos quadrados aparados, de Rousseeuw e Leroy;  
**Feldmann** : regressão estrutural absoluta, de Feldmann;  
**Wald** : método de Wald.

Os métodos robustos (MQA e Feldmann) não produzem estatísticas de testes, portanto só foram comparados aos demais através das distâncias euclidianas entre as estimativas e os parâmetros.

Os métodos que produzem uma região de confiança para os dois parâmetros são: MQ, MQP, RR, Deming, e Wald. Para efeito de cálculo das áreas dos elipsóides que representam estes intervalos de confiança, não foi considerado o método de Wald, já que a expressão derivada por este autor não é da forma

$$A(a - \alpha)^2 + B(a - \alpha)(b - \beta) + C(b - \beta)^2 + D = 0.$$

Para os demais estimadores, é possível representar o elipsóide desta forma, e então fica simples o cálculo de sua área. Pode-se mostrar, através da geometria analítica das cônicas, que a área de uma elipse, quando expressa da forma acima pode ser encontrada através de

$$S_e = \frac{-2D\pi}{\sqrt{(A+C-B)\sqrt{A^2+B^2+C^2-2AC}}}.$$

Tabela 4.2 – Resumo dos Resultados Obtidos nas Simulações – Aceitação de H0 (n=100)

<i>Teste</i>	$\beta$	$CV_x$	$CV_y$	<i>Distr</i>	<i>MQ</i>	<i>MQP</i>	<i>YW</i>	<i>RR</i>	<i>Dem</i>	<i>Lin</i>	<i>RT</i>	<i>Wald</i>
1	1	0,008	0,008	Unif	988	962	922	964	916	859	888	729
2	1	0,08	0,08	Unif	979	941	909	955	905	892	934	660
3	1	0,08	0,008	Unif	989	971	918	968	907	899	962	619
4	1	0,008	0,08	Unif	984	950	913	954	923	884	854	713
5	1,05	0,008	0,008	Unif	0	0	0	0	0	0	0	0
6	1,05	0,08	0,08	Unif	169	118	31	61	80	78	48	418
7	1,05	0,08	0,008	Unif	35	2	0	2	1	0	0	220
8	1,05	0,008	0,08	Unif	28	3	0	3	2	5	0	78
9	1	0,008	0,008	Log-n	641	973	921	973	431	910	807	834
10	1	0,08	0,08	Log-n	636	947	910	965	387	898	762	855
11	1	0,08	0,008	Log-n	640	987	921	969	435	912	804	848
12	1	0,008	0,08	Log-n	658	974	921	978	424	909	728	850
13	1,05	0,008	0,008	Log-n	0	0	0	0	0	30	0	0
14	1,05	0,08	0,08	Log-n	233	216	140	178	142	324	134	402
15	1,05	0,08	0,008	Log-n	109	4	9	40	42	100	20	176
16	1,05	0,008	0,08	Log-n	113	47	10	41	59	143	17	166

As primeiras 5 colunas da tabela mostram qual o teste e as condições de simulação empregadas. A partir da coluna MQ se vê os resultados da aplicação dos diversos estimadores. Os testes de 1 a 4 e de 9 a 12 são aqueles em que não há vício no método Y, e portanto o número ideal de vezes em que se aceita H0 é 950.

Para estes testes, pode-se ver que a melhor aproximação do ideal se dá no estimador RR, entretanto MQP teve também um bom desempenho. Em geral, nota-se que todos os estimadores apresentam desempenho pior no caso log-normal que no caso uniforme, à exceção de Wald. Os piores estimadores foram os de Wald, Linnet e RT para o caso uniforme, e MQ e Deming para o caso log-normal.

Para os demais testes H0 é falsa, portanto o ideal é não aceitá-la nenhuma vez. Se aceitamos um poder de teste de pelo menos 90%, valores entre 0 e 100 seriam aceitáveis para estes resultados. Os estimadores bem sucedidos nestas condições foram YW, RR, Deming, Linnet e RT, sendo que Linnet falha bem mais no caso log-normal. Estes estimadores são exatamente aqueles que consideram erros nos dois eixos, e variâncias não homogêneas, à exceção de Deming. Este último desempenha melhor que o esperado teoricamente, mas como foi visto na revisão bibliográfica, ainda é recomendado nos casos em que o CV é inferior a 20%.

No Teste 14, todos os estimadores apresentam indicação de desempenho pior. É justamente neste teste que os CVs são maiores em ambos os eixos, e os dados têm distribuições log-normais.

Tabela 4.3 – Resumo dos Resultados Obtidos nas Simulações – EQM(a) - (n=100)

<i>Teste</i>	$\beta$	$CV_x$	$CV_y$	<i>Distr</i>	<i>MQ</i>	<i>MQP</i>	<i>YW</i>	<i>RR</i>	<i>Dem</i>	<i>Lin</i>	<i>RT</i>	<i>MQA</i>	<i>Feld</i>	<i>Wald</i>
1	1	0,008	0,008	Unif	0,0088	0,0024	0,0024	0,0024	0,0087	0,0029	0,0027	0,0093	0,0244	0,0102
2	1	0,08	0,08	Unif	0,1525	0,0137	0,0134	0,0134	0,0863	0,0136	0,0135	0,1193	0,1157	0,1205
3	1	0,08	0,008	Unif	0,1421	0,0131	0,0092	0,0092	0,0601	0,0092	0,0091	0,0980	0,0958	0,0996
4	1	0,008	0,08	Unif	0,0606	0,0096	0,0096	0,0096	0,0606	0,0097	0,0108	0,0644	0,0991	0,0683
5	1,05	0,008	0,008	Unif	0,0092	0,0025	0,0025	0,0025	0,0090	0,0031	0,0028	0,0097	0,0259	0,0106
6	1,05	0,08	0,08	Unif	0,1683	0,0131	0,0121	0,0121	0,0912	0,0122	0,0123	0,1299	0,1183	0,1298
7	1,05	0,08	0,008	Unif	0,1462	0,0137	0,0096	0,0096	0,0619	0,0097	0,0096	0,1031	0,1029	0,1037
8	1,05	0,008	0,08	Unif	0,0641	0,0097	0,0097	0,0097	0,0642	0,0098	0,0115	0,0653	0,1039	0,0731
9	1	0,008	0,008	Log-n	0,0030	0,0007	0,0007	0,0007	0,0030	0,0016	0,0011	0,0015	0,0023	0,0009
10	1	0,08	0,08	Log-n	0,0278	0,0027	0,0025	0,0025	0,0279	0,0030	0,0041	0,0121	0,0101	0,0063
11	1	0,08	0,008	Log-n	0,0189	0,0034	0,0018	0,0018	0,0188	0,0023	0,0030	0,0091	0,0082	0,0045
12	1	0,008	0,08	Log-n	0,0190	0,0018	0,0018	0,0018	0,0190	0,0022	0,0033	0,0085	0,0082	0,0042
13	1,05	0,008	0,008	Log-n	0,0030	0,0007	0,0007	0,0007	0,0030	0,0016	0,0010	0,0015	0,0026	0,0010
14	1,05	0,08	0,08	Log-n	0,0294	0,0027	0,0024	0,0024	0,0293	0,0029	0,0042	0,0129	0,0103	0,0064
15	1,05	0,08	0,008	Log-n	0,0209	0,0036	0,0019	0,0019	0,0209	0,0023	0,0032	0,0091	0,0090	0,0050
16	1,05	0,008	0,08	Log-n	0,0225	0,0019	0,0019	0,0019	0,0225	0,0023	0,0034	0,0086	0,0085	0,0042

Tabela 4.4 – Resumo dos Resultados Obtidos nas Simulações – EQM(b) - (n=100)

<i>Teste</i>	$\beta$	$CV_x$	$CV_y$	<i>Distr</i>	<i>MQ</i>	<i>MQP</i>	<i>YW</i>	<i>RR</i>	<i>Dem</i>	<i>Lin</i>	<i>RT</i>	<i>MQA</i>	<i>Feld</i>	<i>Wald</i>
1	1	0,008	0,008	Unif	0,0026	0,0014	0,0014	0,0014	0,0026	0,0016	0,0020	0,0034	0,0070	0,0028
2	1	0,08	0,08	Unif	0,0352	0,0145	0,0124	0,0124	0,0255	0,0125	0,0133	0,0393	0,0401	0,0302
3	1	0,08	0,008	Unif	0,0309	0,0128	0,0088	0,0088	0,0180	0,0089	0,0090	0,0301	0,0323	0,0237
4	1	0,008	0,08	Unif	0,0174	0,0088	0,0088	0,0088	0,0174	0,0088	0,0135	0,0241	0,0325	0,0184
5	1,05	0,008	0,008	Unif	0,0026	0,0015	0,0015	0,0015	0,0026	0,0018	0,0020	0,0035	0,0074	0,0029
6	1,05	0,08	0,08	Unif	0,0385	0,0153	0,0131	0,0131	0,0267	0,0131	0,0139	0,0417	0,0423	0,0323
7	1,05	0,08	0,008	Unif	0,0315	0,0133	0,0094	0,0094	0,0183	0,0094	0,0096	0,0318	0,0349	0,0240
8	1,05	0,008	0,08	Unif	0,0188	0,0091	0,0091	0,0091	0,0188	0,0092	0,0149	0,0248	0,0342	0,0199
9	1	0,008	0,008	Log-n	0,0062	0,0025	0,0025	0,0025	0,0062	0,0075	0,0047	0,0057	0,0078	0,0032
10	1	0,08	0,08	Log-n	0,0591	0,0208	0,0166	0,0166	0,0594	0,0186	0,0403	0,0475	0,0464	0,0280
11	1	0,08	0,008	Log-n	0,0403	0,0179	0,0117	0,0117	0,0405	0,0138	0,0270	0,0347	0,0366	0,0195
12	1	0,008	0,08	Log-n	0,0418	0,0115	0,0114	0,0114	0,0418	0,0136	0,0305	0,0339	0,0360	0,0205
13	1,05	0,008	0,008	Log-n	0,0064	0,0027	0,0027	0,0027	0,0064	0,0077	0,0046	0,0059	0,0086	0,0034
14	1,05	0,08	0,08	Log-n	0,0608	0,0206	0,0167	0,0167	0,0612	0,0188	0,0410	0,0493	0,0467	0,0294
15	1,05	0,08	0,008	Log-n	0,0449	0,0190	0,0126	0,0126	0,0455	0,0150	0,0302	0,0352	0,0400	0,0218
16	1,05	0,008	0,08	Log-n	0,0462	0,0123	0,0123	0,0123	0,0462	0,0144	0,0319	0,0341	0,0383	0,0209

Quando avaliamos os erros quadráticos médios simulados para cada estimador, verificamos que estes são em geral maiores para os testes em que os coeficientes de variação analíticos são maiores (valores iguais a 0,08 ou 8%). Os valores obtidos não variam muito se são comparados os casos com vício e sem vício em  $\beta$ . Também se percebe que os resultados são piores nos casos em que a regressão é feita sem ponderadores. Todos os tipos de regressão ponderada apresentam valores de EQM mais baixos. Os estimadores robustos se equiparam em termos de EQM aos métodos sem ponderação, entretanto o método de Feldmann parece apresentar resultados sempre piores que os demais.

#### 4.1.1.1 Estudo das distâncias euclidianas entre as estimativas e os valores reais dos parâmetros

Para se avaliar as distâncias euclidianas, comparando os estimadores dois a dois, foram feitos alguns gráficos de pares, como se pode ver na Figura 4.1. Nestes gráficos, a linha vermelha que aparece representa os pontos onde ambos estimadores encontram a mesma reta. Pontos acima da reta significam que o estimador locado no eixo das ordenadas encontra resultado mais afastado do real, em relação ao estimador locado nas abscissas.

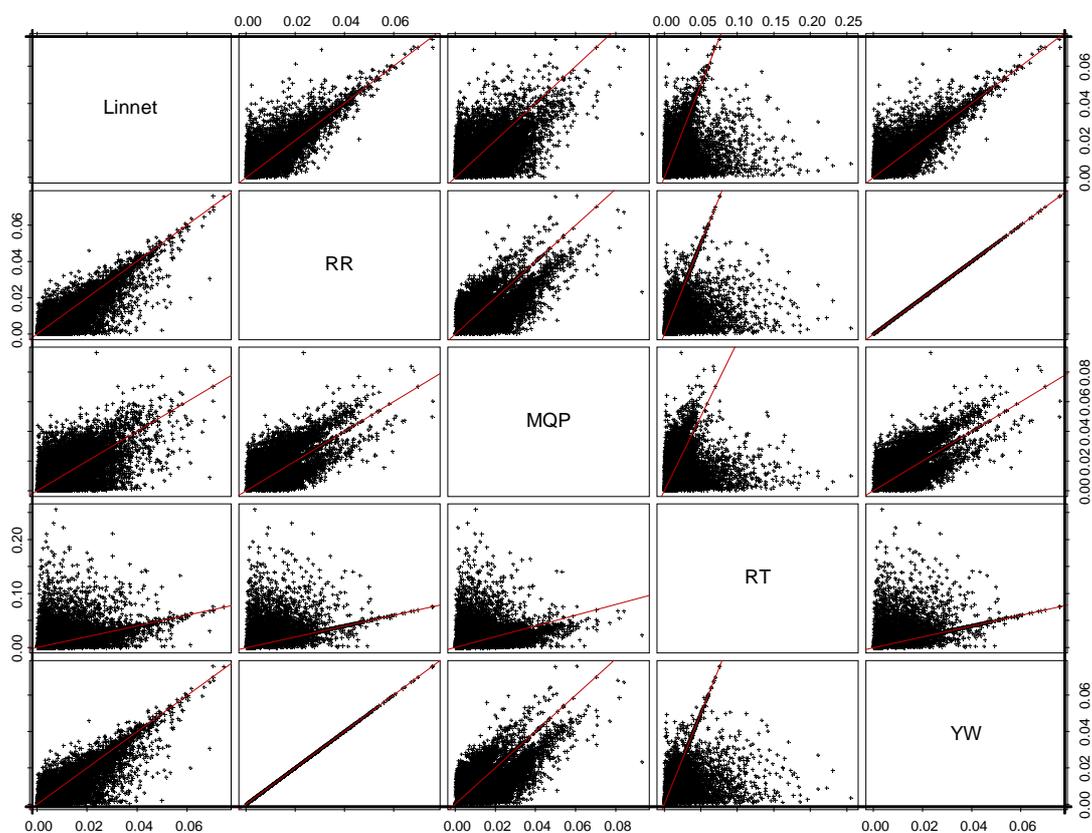


Figura 4.1: Comparação entre estimadores de regressão ponderada, dois a dois. Estatística usada para as comparações é a distância euclidiana entre as estimativas e os parâmetros da reta de regressão. A linha vermelha em cada gráfico mostra a região onde as distâncias são iguais para os dois estimadores, i.e., a reta encontrada é a mesma. Acima dela, as distâncias são maiores para o estimador das ordenadas.

Partindo-se de comparações dois a dois entre os estimadores, pode-se estabelecer uma gradação dos mesmos. A Figura 4.1 mostra que os melhores estimadores dentro deste critério de comparação, são RR e Linnet, nesta ordem. YW fornece exatamente os mesmos resultados que RR, o que é de se esperar pela discussão teórica no capítulo 3. RT deveria fornecer teoricamente os mesmos resultados que RR e YW, como se pode ver na dedução do estimador no capítulo 3. O motivo pelo qual aqui ele apresenta comportamento pior só pode ser explicado pelo método computacional utilizado. Os autores não apresentam um método computacional para se encontrar o mínimo da

expressão que resume a função objetivo, e deixam esta questão aberta ao leitor. Aqui foi utilizado um método de otimização pouco eficiente computacionalmente, mas facilmente implementável em termos de algoritmo. Foi inspirado em Kimanani e Lavigne (1997), e utiliza uma variação do método de bisseção, só que aplicado à procura de um mínimo, ao invés de um zero de função. Ao que tudo indica, algum problema na rotina computacional impede o estimador de encontrar os valores de mínimo de função conforme desejado. Por este motivo, os resultados deste estimador não serão referenciados na seqüência destes estudos. Uma vez que se sabe que corretamente implementado, ele vai fornecer a mesma reta dos métodos RR e YW, a única questão que fica é o teste de hipótese separado para os dois parâmetros, que falha mais vezes na aceitação de  $H_0$  que o método RR, que considera um teste conjunto para os dois, simultaneamente.

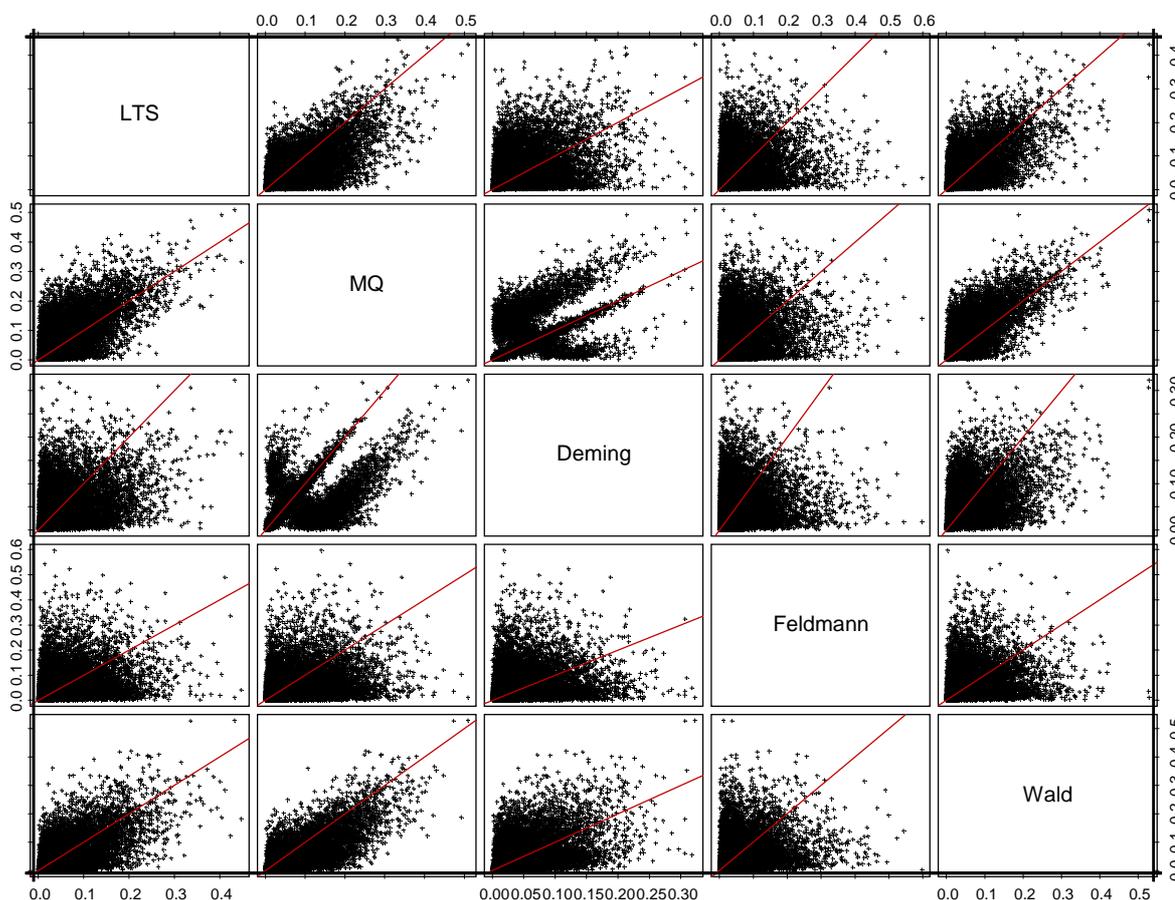


Figura 4.2: Comparação dos estimadores que não consideram regressão ponderada, em termos das distâncias euclidianas entre as estimativas e os parâmetros.

A Figura 4.2 mostra os demais estimadores testados, que não usam de ponderação na regressão. Pode-se perceber que o melhor dentre os cinco mostrados aqui, é o estimador de Deming. O pior de todos é sem dúvida o de mínimos quadrados ordinários (MQ), o que não representa nenhuma surpresa. Fica difícil discernir por estes gráficos qual o melhor desempenho entre LTS, Feldmann e Wald.

Fazendo-se os cálculos a partir dos dados individuais em cada uma das 16000 simulações, constata-se que o estimador de Feldmann apresenta distâncias maiores que o LTS em 53,0% dos casos. Já o estimador de Wald apresenta distâncias maiores em 39,6% dos casos. Isto mostra que o melhor estimador entre os três é o de Wald, e que LTS e Feldmann são equivalentes. Considerando então apenas os melhores estimadores de cada comparação acima, a Figura 4.3 é produzida para mostrar como fica no conjunto.

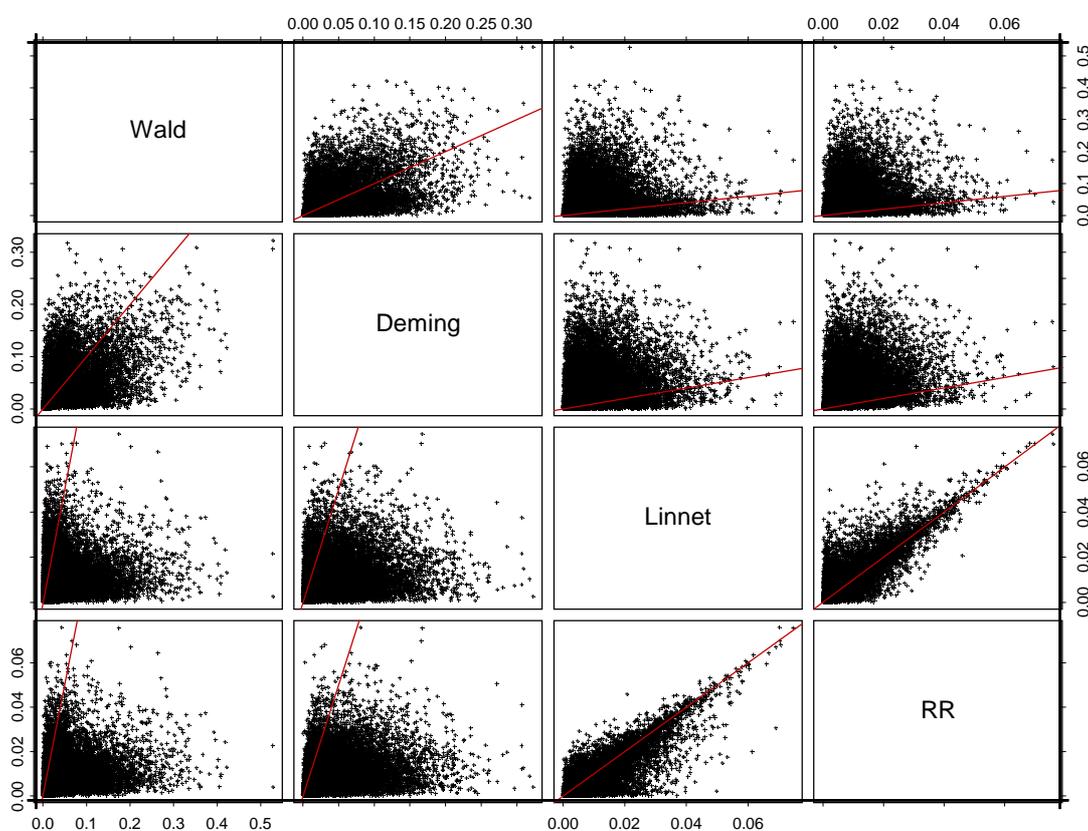


Figura 4.3: Resumo das comparações, com os quatro melhores estimadores encontrados, considerando-se as distâncias euclidianas. Estimadores apresentados, da esquerda para a direita, do pior para o melhor, respectivamente.

#### 4.1.1.2 Estudo das áreas das elipses de confiança para as estimativas dos parâmetros

O segundo estudo, realizado para os estimadores que apresentam testes de hipóteses simultâneos para os dois parâmetros, visa concluir sobre o seu desempenho, de forma análoga ao realizado para as distâncias euclidianas no item anterior. Pelos resultados da Tabela 4.2, percebe-se, como aliás já seria de se esperar pela teoria, que quando se faz os testes em separado para cada parâmetro, os resultados são piores, e o nível de significância dos testes como um todo ficam maiores que os especificados para a construção dos intervalos de confiança.

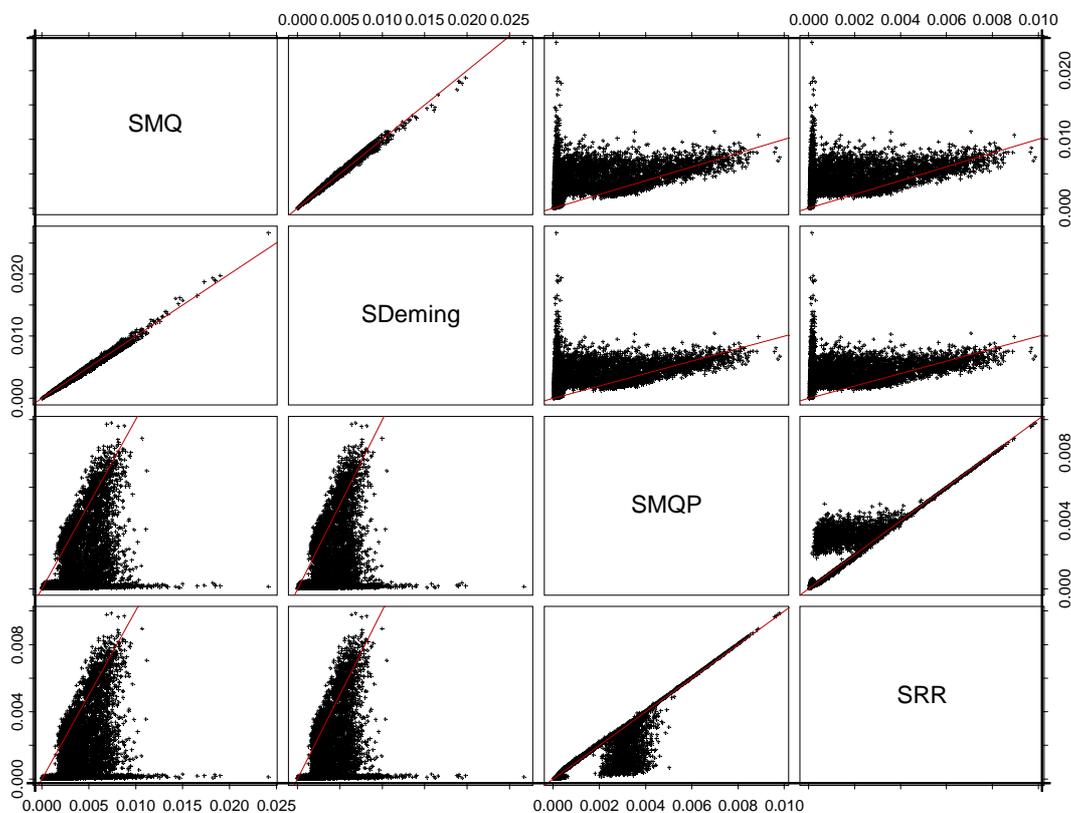


Figura 4.4: Comparação de estimadores dois a dois, considerando as áreas das elipses de confiança de 95% construídas para cada simulação. A linha reta vermelha representa a equivalência entre áreas.

A Figura 4.4 mostra que os estimadores de MQ e Deming produzem áreas praticamente equivalentes, sendo que o de MQ é melhor para as maiores áreas (i.e., testes com CV maior). Isto se explica pelo fato de que MQ só considera erros em um eixo, portanto o intervalo de confiança gerado por ele é mais otimista, já que despreza parte dos erros da comparação. Já no caso dos métodos ponderados, vê-se que RR sempre apresenta áreas menores, ou na pior das situações, iguais, às produzidas por MQP. A superioridade dos métodos ponderados em relação aos que não o são fica evidente também nesta figura.

#### 4.1.1.3 Alguns exemplos de efeito das variáveis de tratamento na regressão.

Quando se avalia conjuntamente as figuras comparando os estimadores através das distâncias euclidianas e áreas das elipses de confiança, conclui-se que em geral RR é o melhor estimador. Entretanto, vê-se que o método de Linnet apresenta distâncias melhores em algumas situações, e MQP apresenta áreas equivalentes às de RR em certas simulações.

Para se entender melhor onde estes estimadores podem representar boas opções de uso, é interessante avaliar estes resultados em termos das condições de simulação (níveis dos fatores usados para se gerar os números aleatórios).

A título de ilustração, são também apresentados alguns resultados que ilustram fatos conhecidos da pesquisa bibliográfica e do aprofundamento teórico (e.g., comparação das inclinações da reta – MQ vs. Deming , ou MQP vs. RR).

Inicialmente, comparamos RR e Linnet. Avaliando separadamente cada um dos 16 testes, verifica-se que apenas nos testes 9 e 13 o estimador de Linnet apresenta resultados substancialmente piores, em termos de distância euclidiana das estimativas aos parâmetros. Estes testes são aqueles em que os dados de X foram gerados segundo uma distribuição log-normal, com  $CV_x = CV_y = 0,8\%$ . Os casos em que o estimador de Linnet apresenta distâncias menores são sempre em testes onde ele praticamente empata com RR, ou seja, os valores maiores ou menores são aleatórios.

A conclusão desta observação é que o estimador de Linnet, nestas simulações, apresenta desempenho equivalente ao de RR em todas as situações simuladas, à exceção de quando os dados são log-normais e os coeficientes de variação em ambos os eixos são bem pequenos. Quando estes CVs aumentam, em um ou outro eixo, ou em ambos, os dois estimadores passam a ser equivalentes, mesmo no caso log-normal.

As Figuras 4.5 e 4.6 mostram isto graficamente.

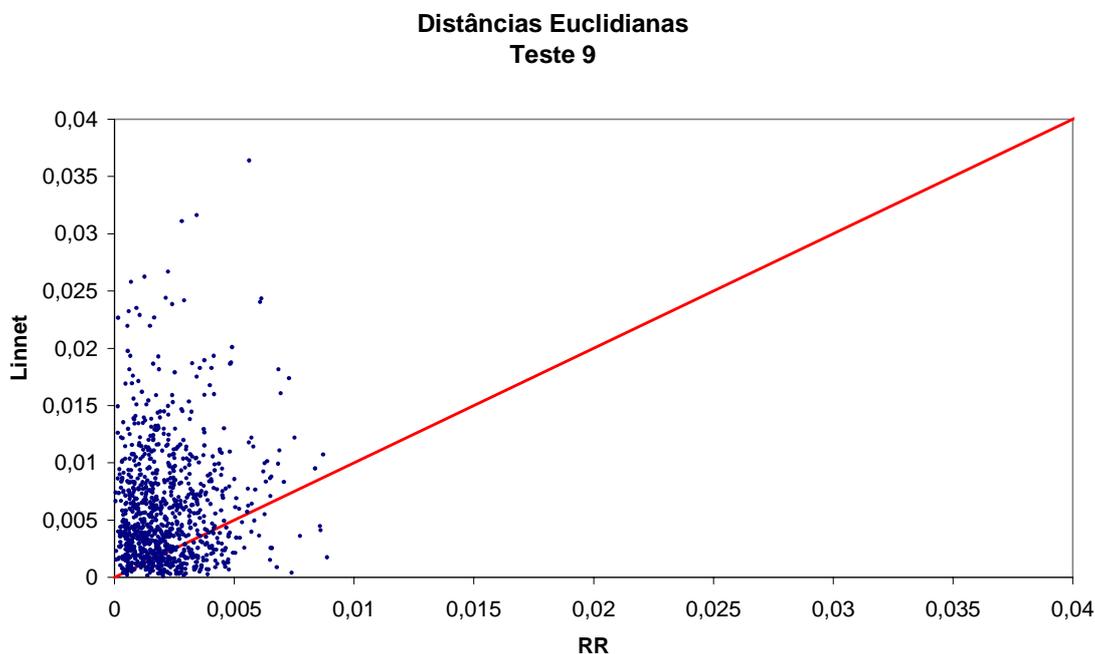


Figura 4.5: Comparação entre as distâncias euclidianas das estimativas produzidas pelos métodos RR e Linnet. Dados do Teste 9, onde  $\beta=1$ ,  $CV_x = CV_y = 0,8\%$  e os dados são log-normais.

**Distâncias Euclidianas  
Teste 12**

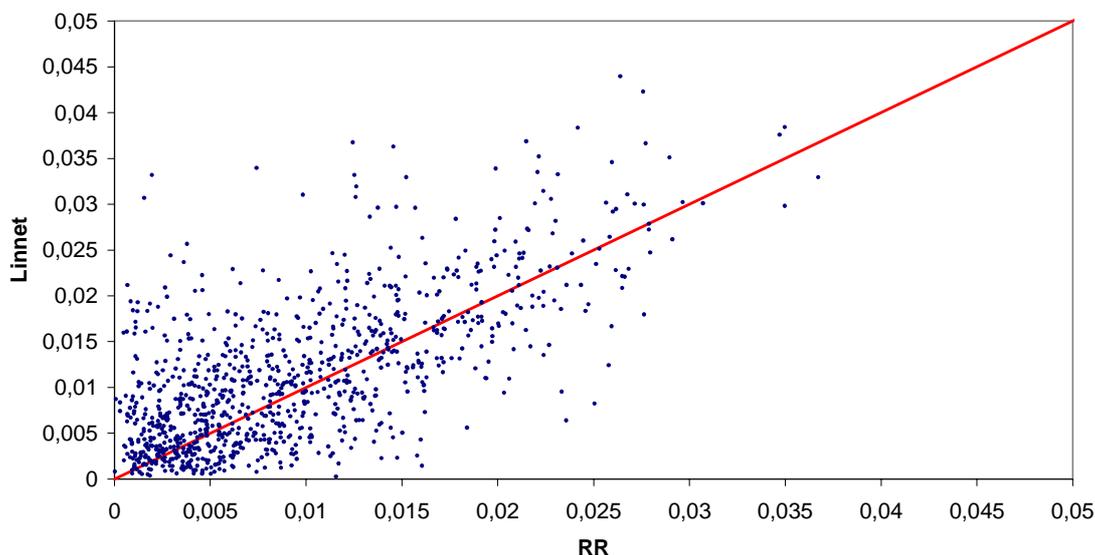


Figura 4.6: Comparação dos estimadores de RR e Linnet usando as distâncias euclidianas entre estimativas e parâmetros. Dados do Teste 12, onde  $\beta=1$ ,  $CV_x = 0,8\%$ ,  $CV_y = 8\%$  e dados são log-normais.

Para verificarmos em que situação os estimadores avaliados pelas regiões elípticas de confiança se distinguem, foram feitos gráficos cruzando as distâncias euclidianas e as áreas das elipses, e os pontos de cada estimador representados de forma diferente. A Figura 4.7 mostra um caso em que o estimador de RR se destaca bastante dos demais, inclusive MQP. Em geral RR e MQP se apresentam equivalentes nestes dois critérios, mas no caso apresentado, no Teste 3, o estimador RR se mostra superior a todos os outros.

Este teste se diferencia dos demais pelo fato de que os erros em X são maiores que os erros em Y. O mesmo comportamento se nota nos Testes 7, 11 e 15. Nestes quatro testes, os estimadores ficam bastante mais separados neste gráfico quando os dados são log-normais, como se pode ver na Figura 4.8, onde são mostrados os resultados para o Teste 15.

**Estudo de Consistência dos Estimadores**

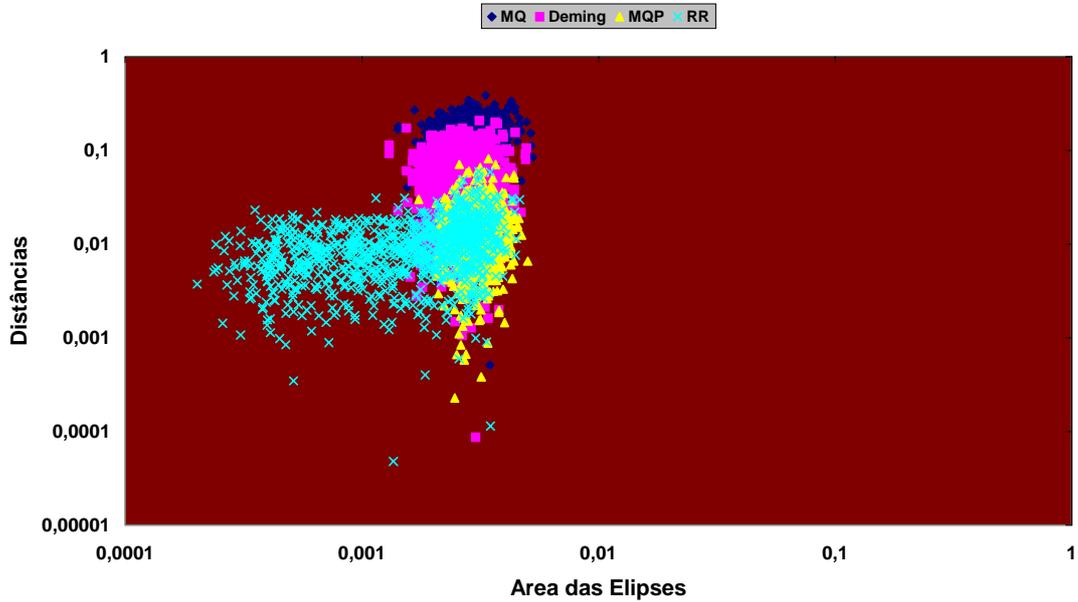


Figura 4.7: Representação das distâncias euclidianas *versus* áreas das elipses para os quatro estimadores onde se dispõe destes resultados. Dados do Teste 3, onde se pode ver a superioridade do estimador RR. Neste teste,  $\beta=1$ ,  $CV_x = 8\%$ ,  $CV_y = 0,8\%$  e dados são uniformes.

**Estudo de Consistência dos Estimadores**

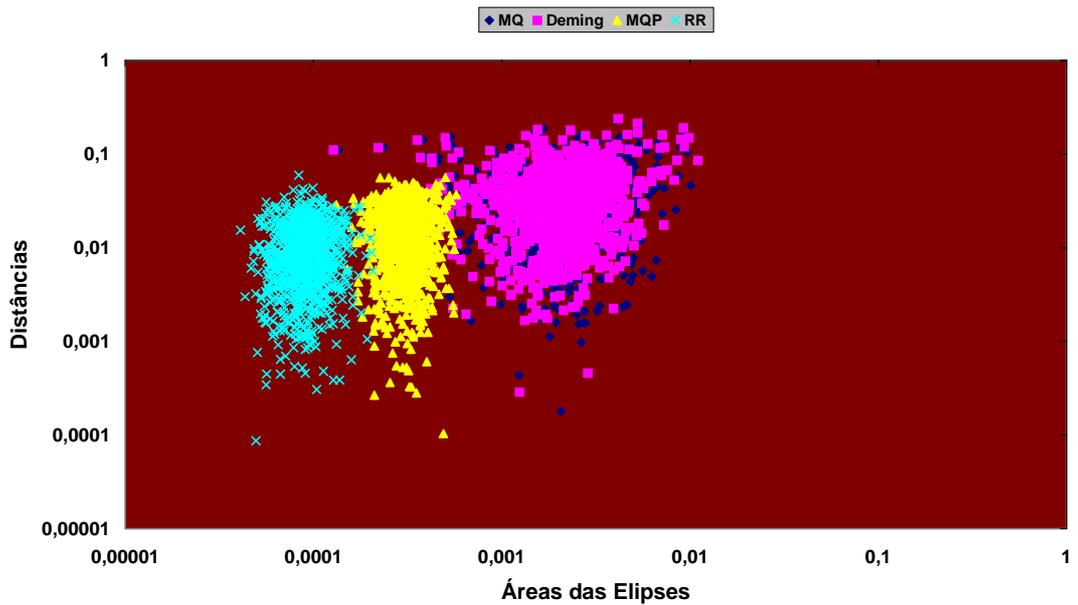


Figura 4.8: Comparação das áreas das elipses e distâncias euclidianas para os estimadores MQ, Deming, MQP e RR. Teste 15, onde  $\beta=1,05$ ,  $CV_x = 8\%$ ,  $CV_y = 0,8\%$  e dados são log-normais.

Nos demais testes simulados, as diferenças entre os estimadores RR e MQP não são tão sensíveis, sendo sempre RR ligeiramente melhor que MQP.

Uma outra forma de se comparar as estimativas por RR e MQP pode ser vista na Figura 4.9, onde se representa, no espaço paramétrico, as estimativas para o caso do Teste 11. As retas vermelhas se cruzam no valor verdadeiro dos parâmetros para estas simulações. Verificando a posição dos pontos em relação a estas retas, pode-se perceber que o estimador MQP é viciado, enquanto que RR não parece ser.

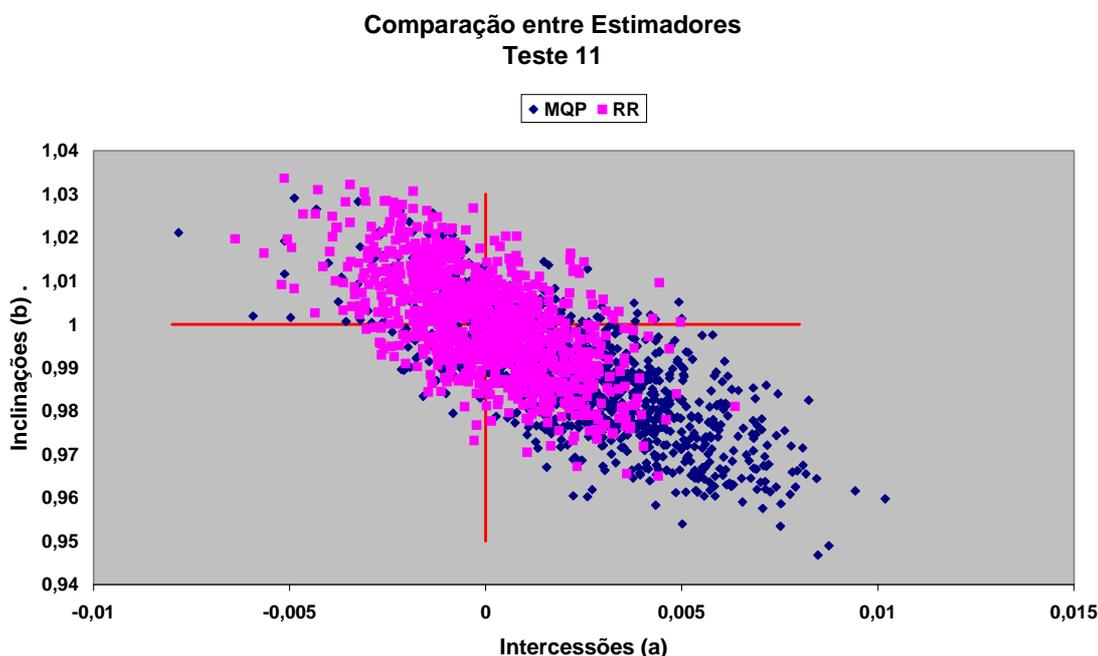


Figura 4.9: Visão das estimativas de MQP e RR no espaço paramétrico. Percebe-se a consistência do estimador RR, que fica centrado nas duas retas vermelhas, cuja interseção é a verdadeira solução em todas estas simulações do Teste 11.

A mesma figura, comparando RR e Linnet no Teste 9, que é onde o método de Linnet se mostrou pior que RR na avaliação anterior (maiores distâncias euclidianas, na Figura 4.5), mostra melhor os resultados superiores de RR em termos de desempenho. Pode-se ver que, apesar de serem ambos consistentes, o estimador RR é mais eficiente. Isto está ilustrado na Figura 4.10.

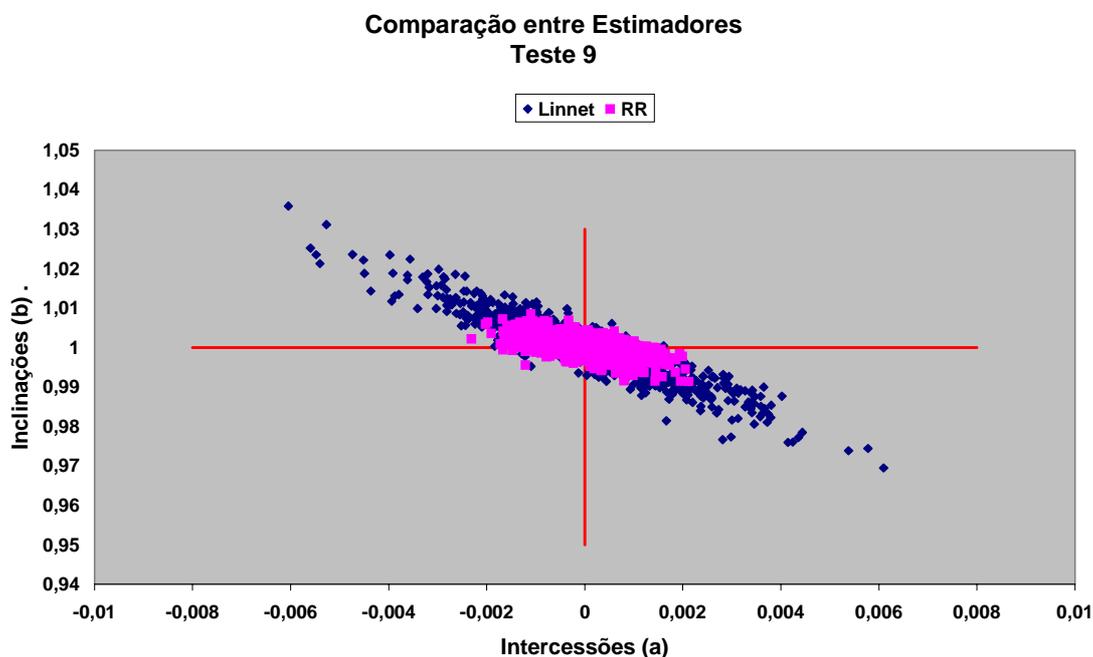


Figura 4.10: Comparação entre as estimativas obtidas pelos métodos de Linnet e RR. Ambos os conjuntos de pontos estão centrados no encontro das duas retas vermelhas, porém a variabilidade das estimativas RR é muito menor, o que mostra a superioridade deste estimador.

A Figura 4.11 mostra um resultado que é esperado, visto na literatura (e.g. Parvin, 1984). As estimativas do coeficiente angular da reta, quando se aplica o método de Deming, são sempre maiores que as estimativas por MQ. Aqui são apresentados os dados referentes a todos os testes.

Na Figura 4.12 se pode ver semelhante comparação das inclinações no caso dos estimadores de RR e MQP. Neste segundo caso, as estimativas MQP são viciadas para baixo em relação a RR, considerado o mais consistente. Como se pode ver pela Tabela 4.4, os erros quadráticos médios para as inclinações são sempre menores para RR, sendo em alguns casos iguais aos obtidos por MQP. Assim, conclui-se que o método MQP é viciado, o que está de acordo com o teoricamente esperado.

### Comparação das Inclinações

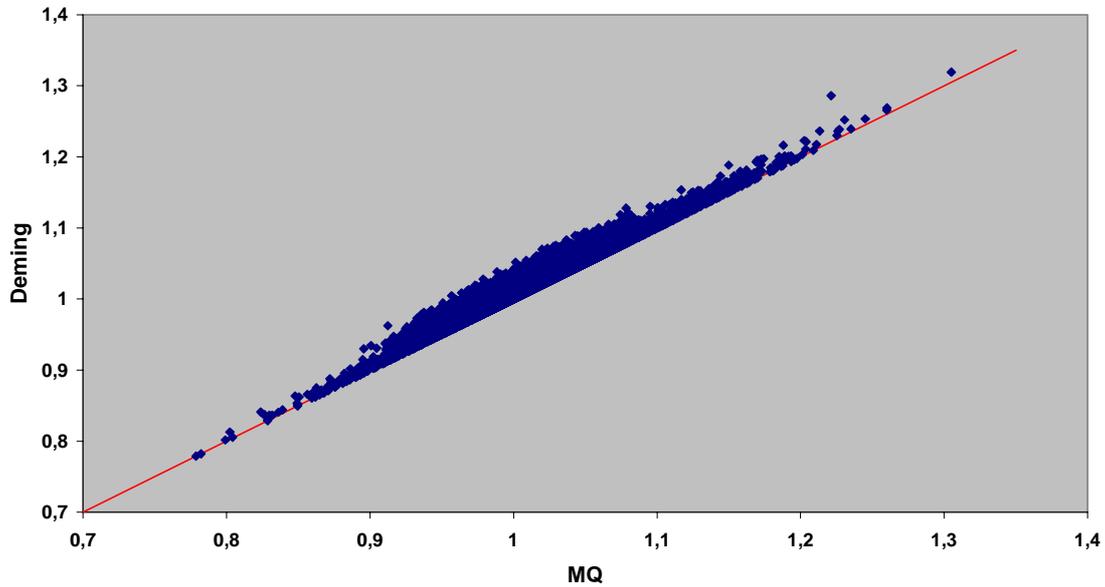


Figura 4.11: Comparação entre as inclinações obtidas por MQ e Deming.

### Comparação das Inclinações

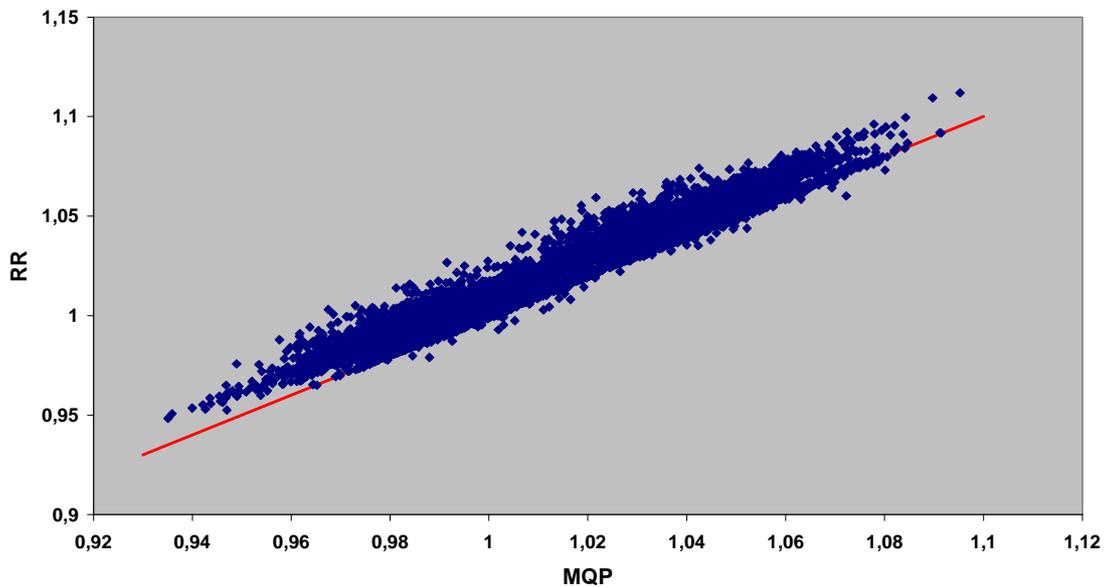


Figura 4.12: Comparação entre as inclinações obtidas por MQP e RR.

A Figura 4.13 mostra as estimativas de interseção da reta para todas as simulações, utilizando-se o método de Linnet. Os dados são mostrados como 'boxplots', e separados por teste. Pode-se perceber que as estimativas estão sempre ao redor de zero, que é o

valor correto, e que a variabilidade é menor para dados log-normais, especialmente quando os  $CV_x$  e  $CV_y$  são menores.

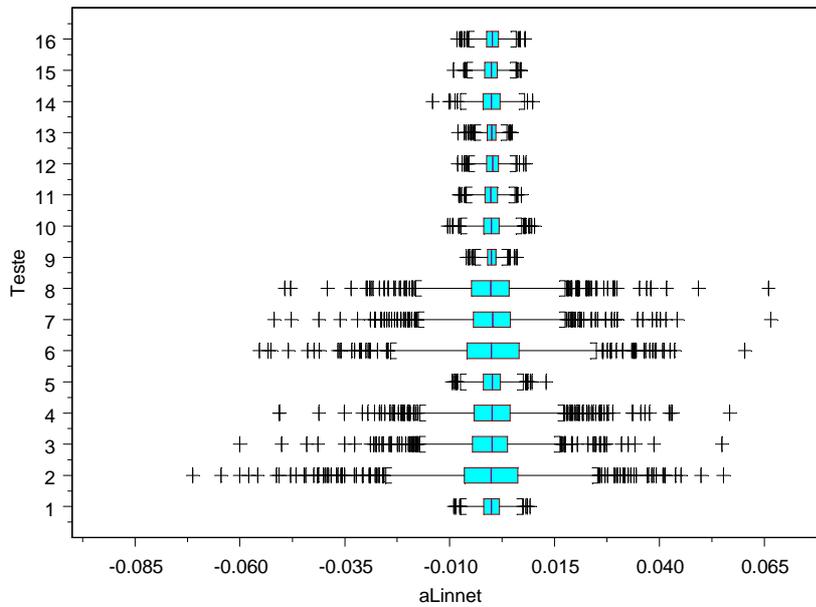


Figura 4.13: Estimativas pelo método Linnet para a interseção da reta, vistas separadamente por teste.

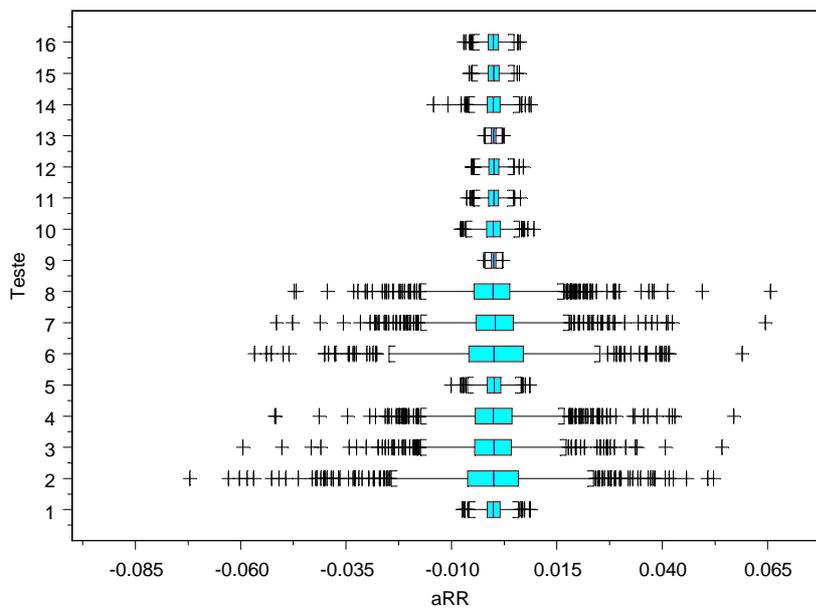


Figura 4.14: Estimativas pelo método RR para a interseção.

A Figura 4.14 mostra a mesma análise, feita aqui com resultados do estimador RR. A escala do gráfico foi feita igual à da Figura 4.13, para facilitar a comparação visual. Pode-se assim perceber que os métodos representados nestas duas figuras apresentam resultados equivalentes.

Se é construído o mesmo gráfico para MQ, pode-se ver que no caso uniforme as estimativas são viciadas, a não ser quando os erros em X são pequenos. No caso log-normal, a própria distribuição dos dados funciona como uma ponderação, o que auxilia o estimador a encontrar valores mais próximos ao correto. Isto está ilustrado na Figura 4.15.

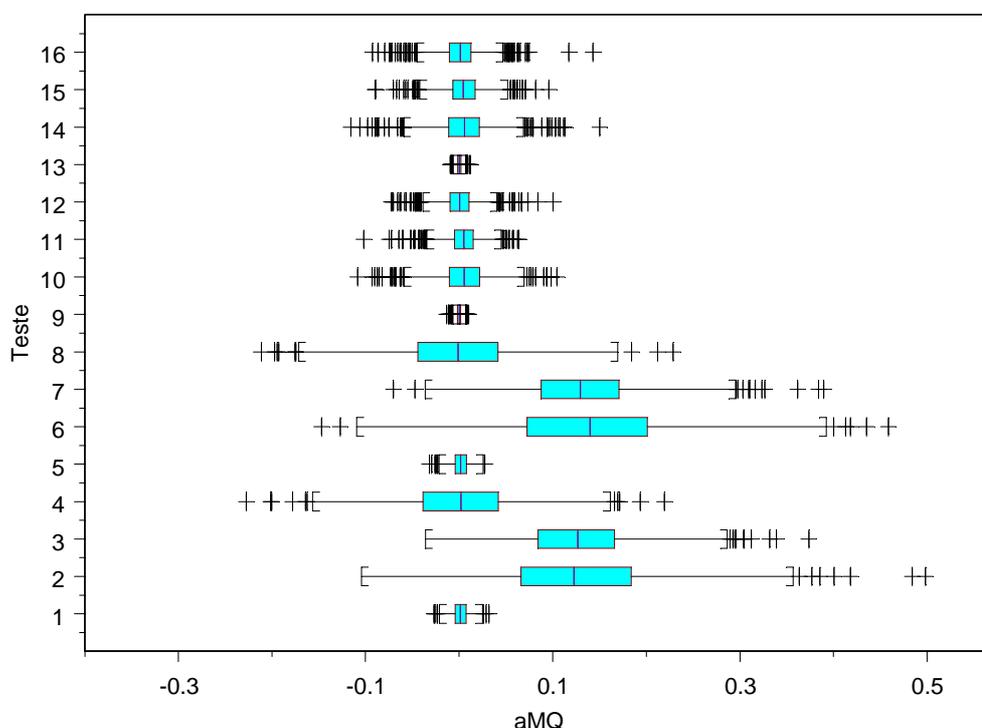


Figura 4.15: Boxplots para as interseções calculadas pelo método MQ, separadas por teste.

A seguir seguem os gráficos referentes aos demais métodos, de onde se pode concluir sobre a presença de vício em alguns estimadores.

No caso do MQP, ocorre algo semelhante ao método MQ, porém os valores de  $a$  são bem menores. Mas equivalentemente, no caso uniforme há tendência a valores viciados para cima. As condições onde isto não é muito notável são as condições usuais de mínimos quadrados (erros em X muito menores que os erros em Y) ou então quando todos os erros são pequenos.

O método de Deming produz resultados não viciados, porém com alta variabilidade.

Os métodos LTS e Wald se comportam também de forma parecida ao MQ, com resultados viciados para o caso de dados uniformes, a não ser quando os erros são pequenos.

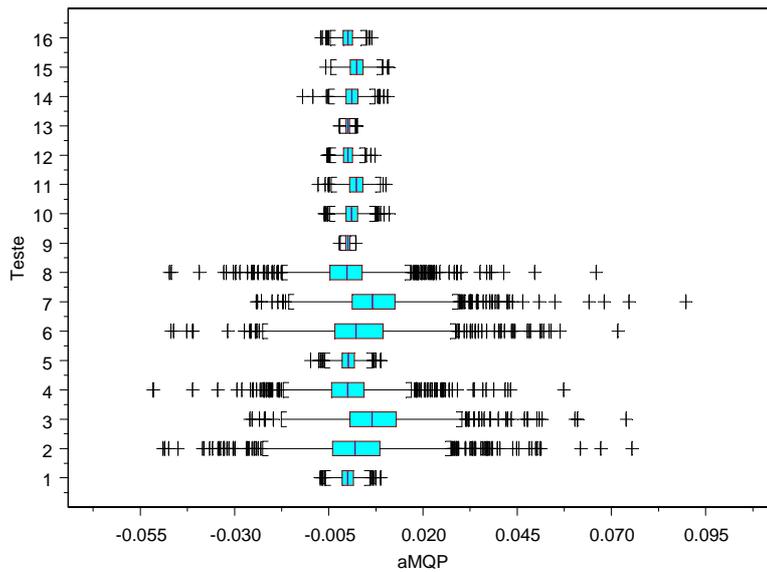


Figura 4.16: Interseções pelo método MQP.

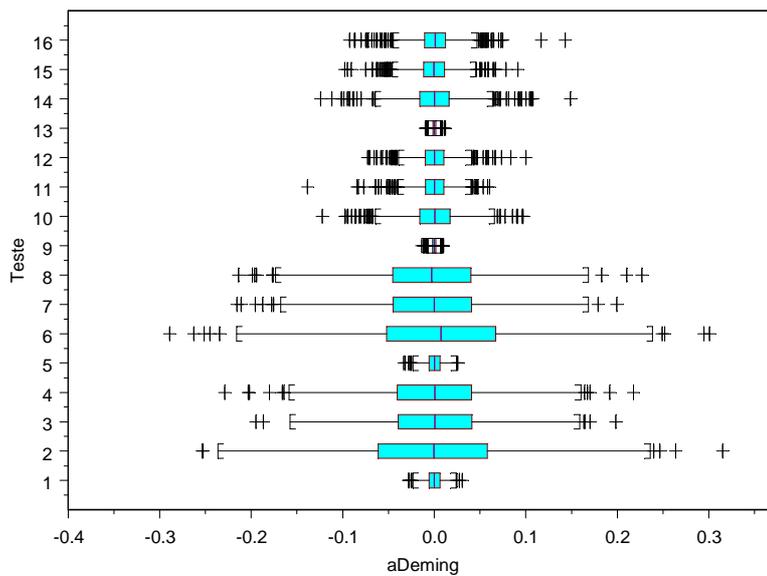


Figura 4.17 : Interseções pelo método de Deming.

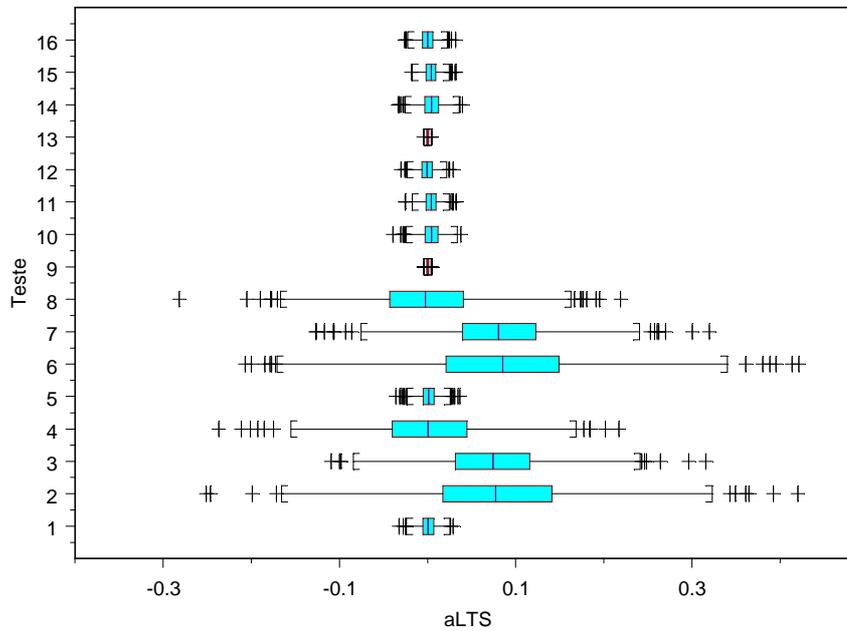


Figura 4.18: Interseções pelo método LTS, de Rousseeuw.

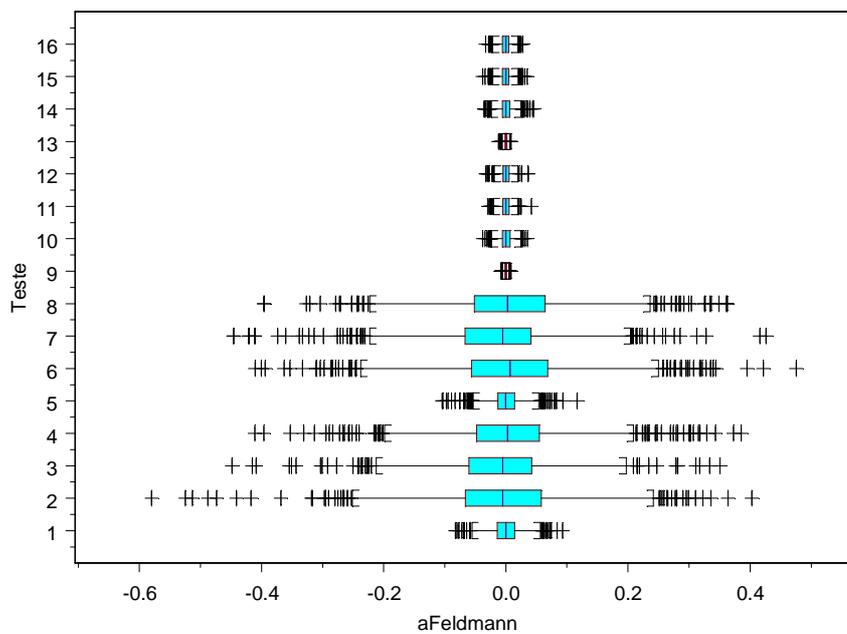


Figura 4.19: Interseções pelo método de Feldmann.

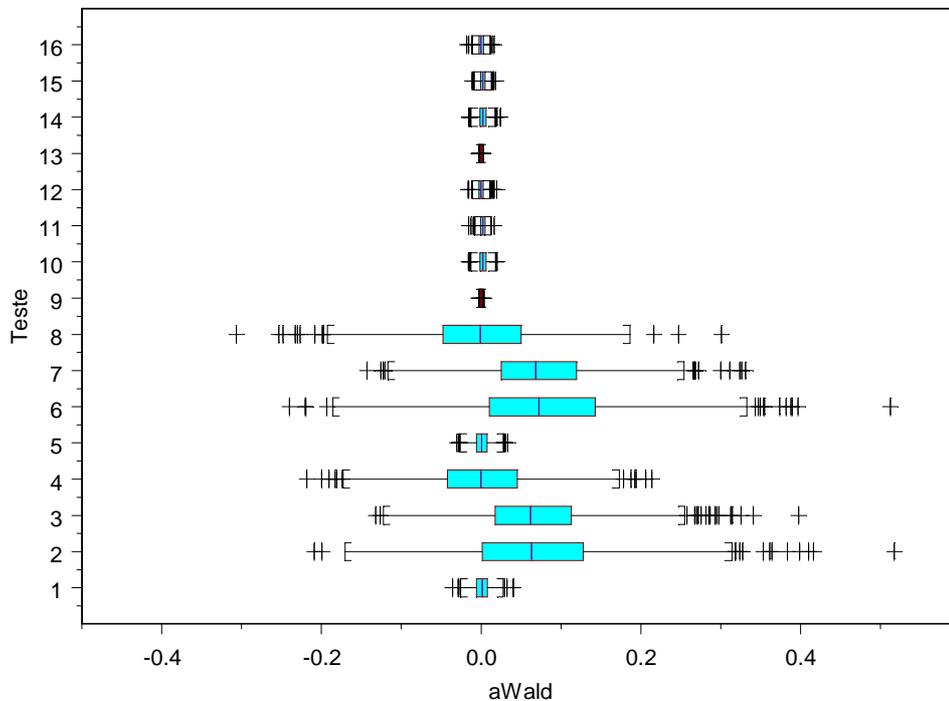


Figura 4.20: Interseções pelo método de Wald.

O método de Feldmann parece apresentar estimativas não viciadas, o que parece indicar que é um método melhor que o LTS, se consideramos apenas os estimadores robustos. É importante lembrar que as estimativas de interseção e inclinação são sempre correlacionadas. Portanto, resultados viciados para um deles implicam em vício no outro.

Considerando os métodos não ponderados e robustos, os melhores sob este ponto de vista seriam o de Deming e depois o de Feldmann. Para o caso dos métodos ponderados, somente são não viciados os estimadores RR e Linnet, que são quase equivalentes.

#### 4.1.2. Considerações sobre o tamanho de amostra e magnitude dos erros no nível de significância do teste.

Em função dos resultados encontrados na Seção 4.1.1, uma nova série de simulações foi realizada, procurando-se explorar a pior condição encontrada para os melhores estimadores avaliados, que foi a do Teste 14.

Nesta nova seqüência de simulações, foram então tomadas as seguintes condições:

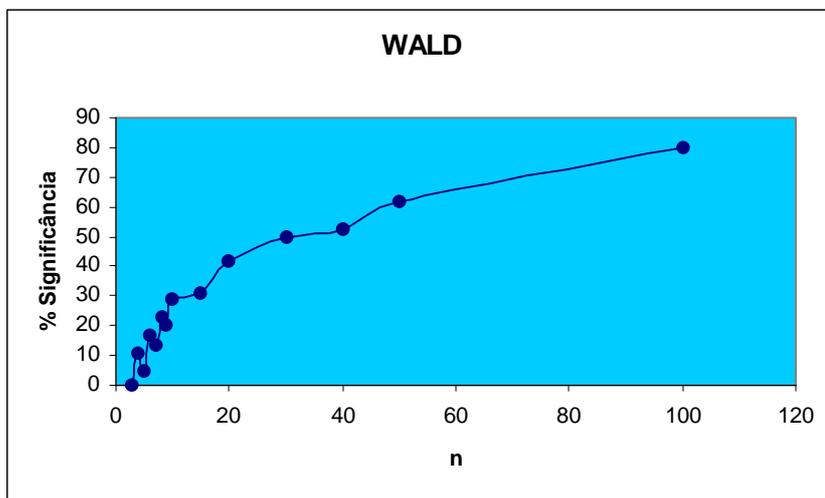
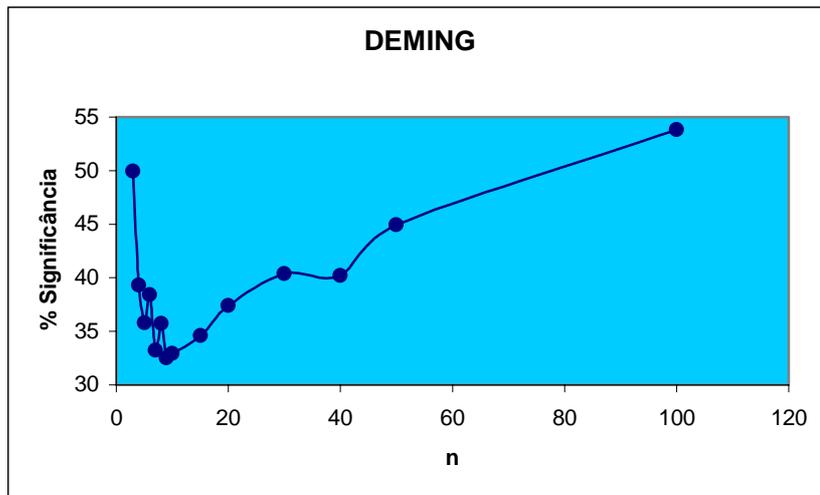
- i) em todos os casos, dados seguindo distribuição log-normal;
- ii)  $\beta = 1$ ;
- iii)  $CV_y = 20\%$  ou  $0,2$ ;



- iv)  $CV_x = 50\%$  ou  $0,5$ .  
 v)  $n$  variando ao longo dos valores:  
 3,4,5,6,7,8,9,10,15,20,30,40,50,100,200,300,400,500,1000

Também aqui foram realizadas 1000 simulações para cada um dos testes (para cada valor de  $n$ ). Estimadores adotados nesta fase foram: RR, MQP, Deming, e Wald.

Os resultados obtidos são mostrados na Figura 4.21.



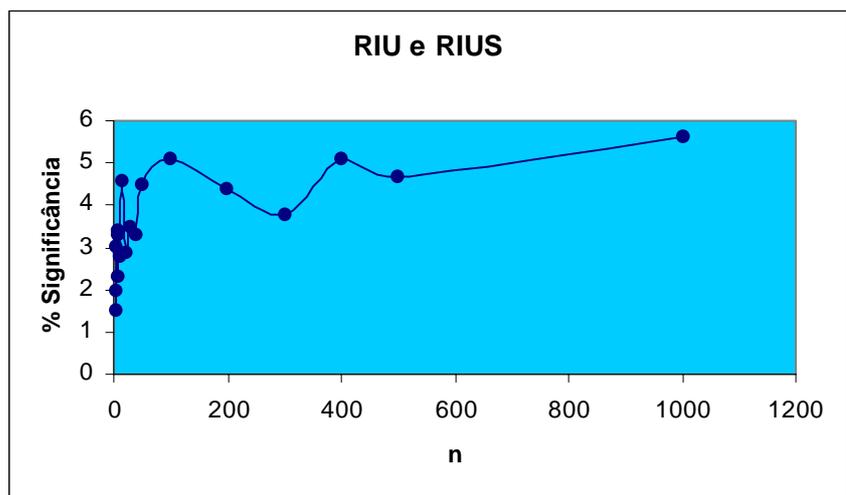
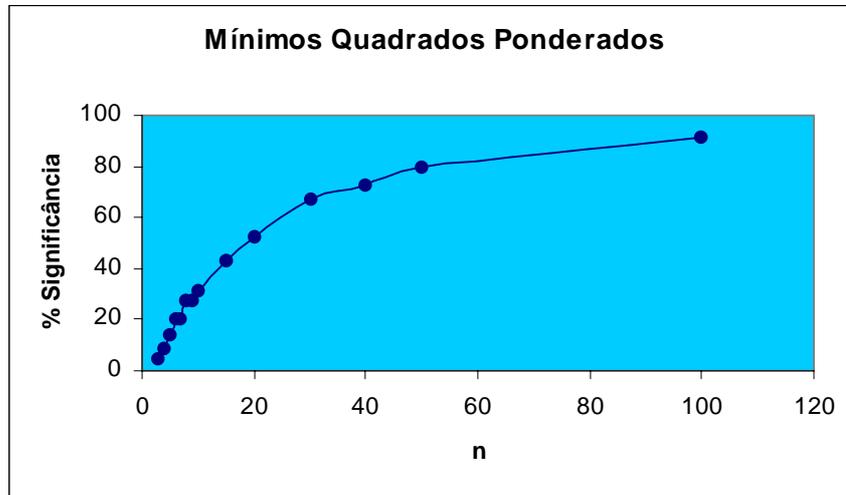


Figura 4.21: Variação do Nível de Significância com o tamanho da amostra para alguns estimadores

Todos os testes foram feitos para um nível de significância de 5%. Apenas o estimador de Riu e Rius chega a este nível, para valores de  $n$  a partir de 15. Entretanto, mesmo a partir deste valor de  $n$  o nível de significância oscila, a grosso modo, entre 4 e 5%. Os demais estimadores falham notavelmente nestas condições de simulação.

Uma outra investigação que foi feita buscou informação sobre o nível de significância em relação ao aumento da grandeza dos erros em  $X$ . Variou-se o coeficiente de variação em  $X$  de 0,5 até 1,0 (com passos de 0,1). Para estes testes, manteve-se  $n=500$ . Como o único estimador que se manteve estável em termos de nível de significância, em termos assintóticos, foi o RR, apenas este foi usado nesta última investigação. Os resultados podem ser vistos na Figura 4.22.

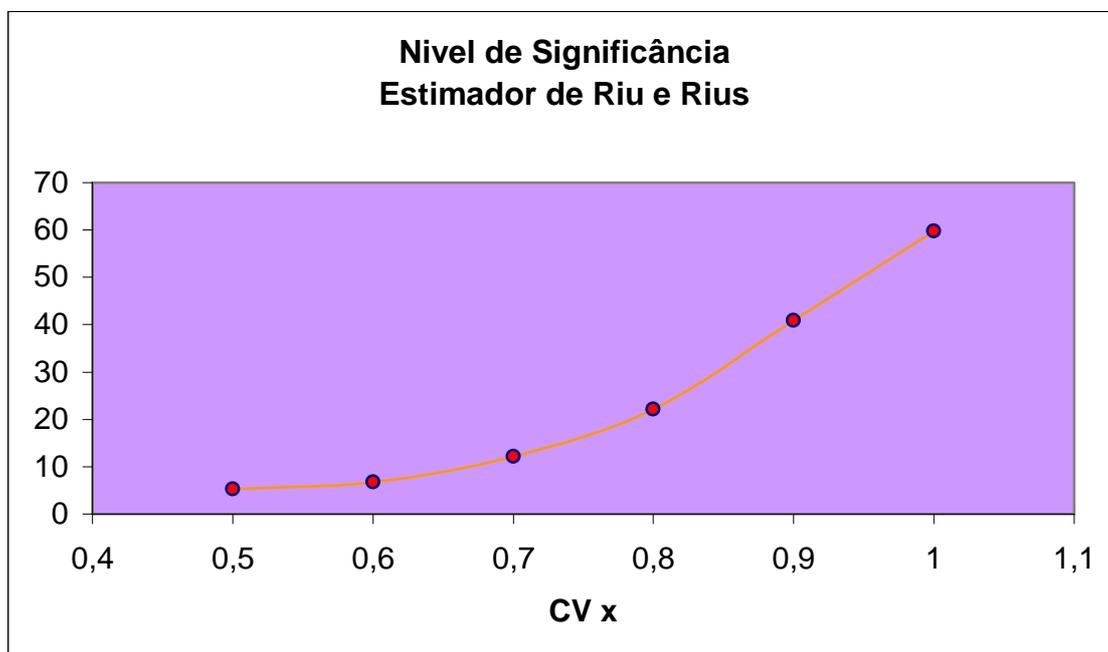


Figura 4.22: Comportamento do nível de significância de teste com o aumento do coeficiente de variação em X. Estimador de Riu e Rius (1996).

O que se percebe neste gráfico é que acima de  $CV=50\%$ , o estimador não consegue manter mais o nível de significância estabelecido para o teste. Isto mostra que para erros muito grandes (considerando como referência os métodos de análise química usualmente utilizados na indústria mineral), nem mesmo o teste de Riu e Rius consegue manter os níveis nominais. Para estes casos, tudo indica que se torna necessário um novo procedimento.

## 4.2. EXEMPLOS DA LITERATURA

### 4.2.1. Dados de Pearson (1901) utilizando os pesos de York (1966)

Constantemente na literatura são propostos métodos para cálculo das estimativas da reta em situações onde há erros nas duas variáveis, com exemplos. Os dados publicados por Pearson (1901) são uma das referências. York (1966) ao apresentar sua proposta, exemplifica com estes dados, atribuindo pesos para a regressão.

Posteriormente, outros autores aproveitam os pesos atribuídos por York e testam outros métodos. Cecchi (1991) usa estes dados para testar sua proposta de cálculo dos erros padrão de regressão quando a estimação se faz com o método de Neri et al. (1989). Riu e Rius (1995) usam os mesmos dados (com os mesmos pesos) para uma comparação entre 15 métodos de regressão por eles pesquisados na literatura.

O atual estudo acrescenta ao trabalho de Riu e Rius (1995) as estimativas de erro padrão, além dos coeficientes da reta, e mostra os resultados para métodos que naquele estudo não foram usados. Os métodos aqui apresentados são:

- Mínimos quadrados bivariados, de Riu e Rius (1996);
- Mínimos quadrados ponderados;
- Mínimos quadrados ordinários;
- Relação funcional de máxima verossimilhança, de Ripley e Thompson (1987);
- Variação do método de Deming para Coeficiente de Variação constante, de Linnet (1990).

Os dois últimos métodos, avaliados como apresentando bom desempenho no capítulo de simulações do presente trabalho, não foram incluídos no estudo de Riu e Rius (1995), embora ambos sejam citados nas referências bibliográficas destes autores.

A Tabela 4.5 mostra os dados em que foram aplicados os métodos de regressão.

Tabela 4.5: Dados de Pearson (1901) e pesos de York (1966)

$i$	$x$	$w_x = 1/s_x^2$	$y$	$w_y = 1/s_y^2$
1	0,0	1000	5,9	1
2	0,9	1000	5,4	1,8
3	1,8	500	4,4	4
4	2,6	800	4,6	8
5	3,3	200	3,5	20
6	4,4	80	3,7	20
7	5,2	60	2,8	70
8	6,1	20	2,8	70
9	6,5	1,8	2,4	100
10	7,4	1	1,5	500

Como se pode ver, estes dados não correspondem a um estudo de comparação de métodos, já que os valores  $x$  e  $y$  não são equivalentes, evidentemente. Eles são, inclusive, correlacionados negativamente, como uma inspeção visual da tabela pode mostrar. Entretanto, são um exemplo muito utilizado na literatura para verificar os resultados comparativos dos métodos de regressão propostos por vários pesquisadores, quando existem erros nos dois eixos.

A Tabela 4.6 mostra os resultados de aplicação dos métodos de regressão de interesse no presente trabalho, e a Figura 4.23 mostra o gráfico com os pontos e retas calculadas. Na tabela,  $a$  é a estimativa da intercessão da reta de regressão,  $b$  o coeficiente angular, e  $Sa$  e  $Sb$  são os erros padrão de estimativa para os mesmos parâmetros.

Tabela 4.6: Resultados de regressão por vários métodos da literatura – Dados de Pearson

Método	a	b	Sa	Sb
MQ	5,76118519	-0,53957727	0,18948520	0,04212655
MQP	6,10010932	-0,61081296	0,20466269	0,03008745
RR	5,47991022	-0,48053341	0,19410920	0,04295442
RT	5,76450700	-0,53957727	0,30186495	0,05995857
Linnet	5,78667166	-0,54648129	0,14079146	0,03224749
<b>Solução Exata</b>	5,47991022	-0,480533407	0,2919335	0,057617

A última linha da tabela se refere à solução apresentada por Riu e Rius (1995) como sendo a encontrada por outros autores e aceita como a solução exata de mínimos quadrados generalizados. Os erros padrão correspondentes a esta solução são os publicados por Williamson, considerados como referência por Riu e Rius (1995) e por Kalantar (1992).

Pode-se perceber que a solução que melhor se aproxima da chamada solução exata é a de RR. Também se nota que praticamente as estimativas MQ, RT e Linnet são equivalentes, porém Linnet apresenta os menores erros padrão de estimação, usando o método “jackknife” para sua obtenção. O método RT apresenta erros padrão bastante próximos da solução de Williamson.

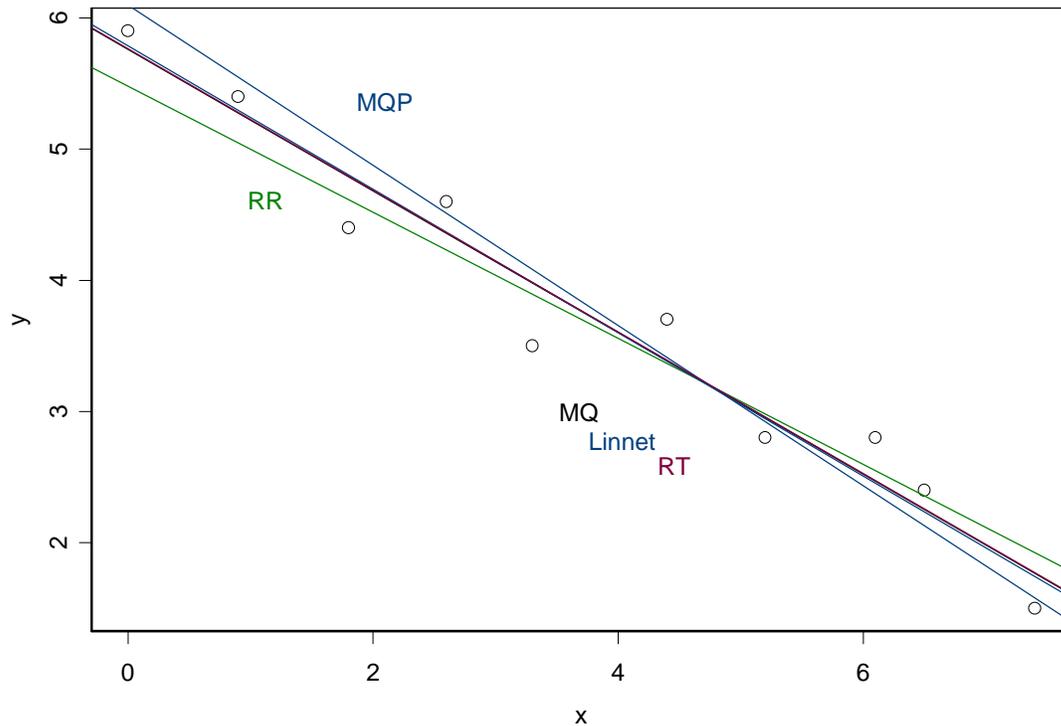


Figura 4.23: Regressão utilizando vários métodos encontrados na literatura. Dados de Pearson (1901) com pesos de York (1966) mostrados na Tabela 4.5.

#### 4.2.2. Exemplos de Ripley e Thompson (1987)

Neste caso, considera-se os dados fornecidos por Ripley e Thompson (1987) em estudos de comparação de métodos, onde eles apresentam o método deles (RT). No primeiro exemplo são apresentados resultados de determinação de Arsênio (V) em água por (i) redução seletiva contínua com absorção atômica e (ii) redução não-seletiva, extração a frio e emissão ótica.

Os dados são apresentados na Tabela 4.7, em  $\mu\text{g/L}$ . A Tabela 4.8 apresenta os resultados de aplicação dos métodos de regressão, que são ilustrados na Figura 4.24.

Tabela 4.7: Comparação de métodos – Exemplo no. 1 de Ripley e Thompson (1987)

<i>I</i>	<i>Redução seletiva / AA</i>				<i>Extração a frio / EA</i>			
	<i>x</i>	<i>s(x)</i>	<i>x</i>	<i>s(x)</i>	<i>y</i>	<i>s(y)</i>	<i>y</i>	<i>s(y)</i>
1	8,71	1,92	1,38	0,36	7,35	2,07	1,17	0,33
2	7,01	1,56	1,81	0,45	7,92	2,23	2,31	0,66
3	3,28	0,76	1,27	0,33	3,40	0,96	1,88	0,54
4	5,60	1,26	0,82	0,23	5,44	1,53	0,44	0,13
5	1,55	0,39	1,88	0,46	2,07	0,59	1,37	0,40
6	1,75	0,43	5,66	1,27	2,29	0,65	7,04	1,98
7	0,73	0,22	0,00	0,06	0,66	0,19	0,00	0,01
8	3,66	0,84	0,00	0,06	3,43	0,97	0,49	0,15
9	0,90	0,25	0,40	0,15	1,25	0,36	1,29	0,37
10	9,39	2,07	0,00	0,06	6,58	1,85	0,37	0,12
11	4,39	1,00	1,98	0,48	3,31	0,93	2,16	0,62
12	3,69	0,84	10,21	2,24	2,72	0,77	12,53	3,51
13	0,34	0,13	4,64	1,05	2,32	0,66	3,90	1,10
14	1,94	0,47	5,66	1,27	1,50	0,43	4,66	1,31
15	2,07	0,50	19,25	4,18	3,50	0,99	15,86	4,45

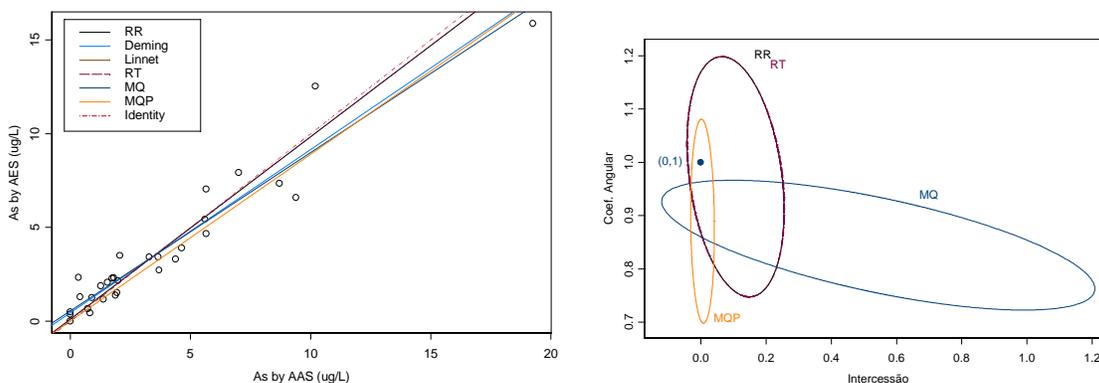


Figura 4.24: Utilização de vários métodos de regressão para análise dos dados do primeiro exemplo de Ripley e Thompson (1987). Estudo com 30 pontos, apresentando coeficiente de correlação de 0,959. No gráfico da direita se vê as elipses de 95% de confiança para os parâmetros.

Tabela 4.8: Resultados das Análises de Regressão – Dados de Ripley e Thompson

Método	a	b	Sa	Sb
<b>MQ</b>	0,5441525	0,8446433	0,25696640	0,04712432
<b>MQP</b>	0,0050555	0,8895175	0,00991208	0,05194422
<b>RR</b>	0,1064500	0,9729924	0,04941785	0,07481156
<b>RT</b>	0,1064482	0,9729888	0,04941721	0,07481096
<b>Deming</b>	0,4515020	0,8699185	0,25884980	0,04759547

O método de Deming (MQD) não é apresentado no gráfico da direita, já que os autores pesquisados não apresentam expressão para o intervalo conjunto para os dois parâmetros. Entretanto, pode-se ver pelo gráfico da direita e pela tabela acima que ele se aproxima mais do resultado de aplicação de mínimos quadrados. Ambos rejeitam, com 95% de confiança, a hipótese de igualdade entre os métodos comparados. Os demais testes, utilizando regressão ponderada, não a rejeitam. Percebe-se, por meio da elipse de confiança, que o método de mínimos quadrados apresenta uma incerteza muito grande na determinação do coeficiente linear da reta de regressão.

O segundo exemplo encontrado em Ripley e Thompson (1987) contém resultados de análises de berílio em rochas e solos, realizadas com alguns CRMs. No primeiro exercício, os autores usam apenas os CRMs do USGS (“United States Geological Survey”), cujos resultados se encontram na Tabela 4.9. O primeiro método usado consiste de fusão das amostras em metaborato de lítio e posterior dissolução em ácido nítrico diluído. A leitura é feita usando emissão ótica em plasma de acoplamento indutivo (ICP-AES). O segundo método consiste de abertura ácida e extração por solvente, seguida de leitura em espectrômetro de absorção atômica (AAS).

Tabela 4.9 – Análise de Berílio em amostras de Rocha e Solos (padrões USGS)

<i>Amostra</i>	<i>ICP-AES</i>		<i>AAS</i>	
	<i>x</i>	<i>s(x)</i>	<i>y</i>	<i>s(y)</i>
G2	2,34	0,08	2,48	0,05
GSP-1	1,20	0,09	1,22	0,04
AGV-1	1,88	0,08	2,14	0,045
DTS-1	0,08	0,07	0,0026	0,0006
PCC-1	0,12	0,07	0,0023	0,0006
GXR-1	1,12	0,10	1,05	0,035
GXR-2	1,60	0,10	1,42	0,04
GXR-4	2,16	0,08	1,99	0,045
GXR-5	1,34	0,09	1,06	0,04
GXR-6	1,35	0,09	1,04	0,04

Neste exemplo, pode-se verificar a importância da avaliação das precisões, para melhor compreender os resultados da comparação. A Figura 4.25, mostra os desvios padrão dos dois métodos em função da concentração de Be (dada em ppm ou  $\mu\text{g/g}$ ). As linhas acrescentadas ao gráfico são linhas de regressão, traçadas para visualizar melhor a tendência de variação do erro aleatório com o teor. Pode-se perceber neste caso bem nitidamente que o método por absorção atômica é mais preciso. Os erros do ICP-AES apresentam comportamento de desvio padrão aproximadamente constante, o que sugere que estas leituras se encontram próximas ao limite de detecção do método. O mesmo não

acontece com o AAS. Para este último, parece que a relação linear para a precisão não é a mais adequada.

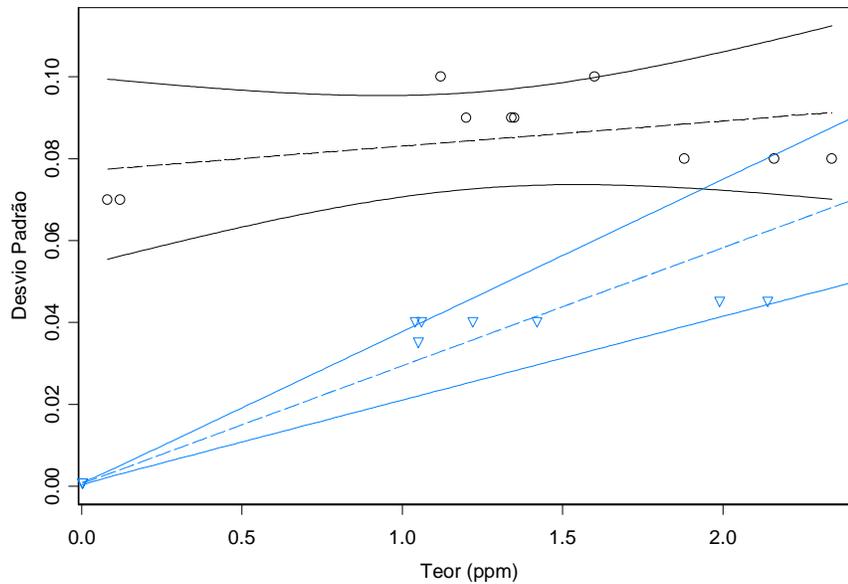


Figura 4.25: Comparação das precisões. Bolas pretas (ICP) e triângulos azuis (AAS). O gráfico mostra que o método de Absorção Atômica é notavelmente mais preciso que o ICP.

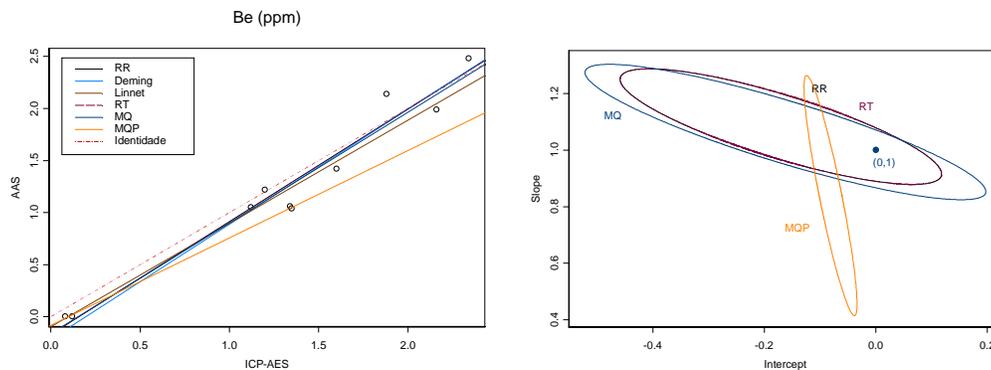


Figura 4.26: Comparação dos resultados do segundo exemplo de Ripley e Thompson (1987) de análise de Be por ICP-AES e AAS.

A comparação dos resultados dos dois métodos é feita na Figura 4.26. Todos os métodos de regressão, à exceção do de mínimos quadrados ponderados, aceitam a hipótese de igualdade (ou melhor, não a rejeitam) no nível de significância de 5%. Como o método ICP-AES é colocado no eixo  $x$ , e é o menos preciso, a conclusão por MQP é questionável.



A Tabela 4.10 mostra os resultados das regressões.

Tabela 4.10: Resultados das Análises de Regressão

Método	a	b	Sa	Sb
MQ	-0,1622548	1,063491	0,12038540	0,08013430
MQP	-0,0812414	0,838946	0,00098916	0,00882831
RR	-0,1704758	1,082743	0,09658638	0,06837337
RT	-0,1704763	1,082744	0,05298407	0,03750736
Deming	-0,2169653	1,104970	0,12374250	0,08263575
Linnet	-0,0969512	0,991467	0,02929366	0,05058957

Em todos os exemplos citados nesta Seção, pode-se evidenciar que é necessário muito cuidado ao se escolher um método de regressão para se comparar metodologias de análise química. Existem situações, algumas descritas nas simulações apresentadas anteriormente, em que mais de uma metodologia é aceitável, e as conclusões não são afetadas. Em outras situações podem ocorrer diferenças muito grandes. O mais seguro então, é utilizar uma metodologia cujas suposições básicas não sejam violadas facilmente, como RR e RT, que como se vê nestes exemplos, apresentam resultados bem semelhantes. O número de casas decimais reportadas nas tabelas é muito superior ao necessário, mas permitem verificar a concordância dos algoritmos de cálculo utilizados.

### 4.3. EXEMPLOS DO AUTOR

#### 4.3.1. Análise de ouro utilizando abertura da amostra com água régia, extração por solvente e leitura por duas técnicas, absorção atômica em forno de grafite e emissão ótica em plasma de acoplamento indutivo

Este estudo foi feito para se verificar a equivalência de duas técnicas de leitura comumente utilizadas em análises de ouro, absorção atômica em forno de grafite (AA-FG) e espectrometria de plasma (ICP-AES). As amostras foram selecionadas a partir de vários projetos de pesquisa de ouro da Companhia Vale do Rio Doce, na fase de levantamento geoquímico em solos e sedimentos de corrente.

Cada amostra foi dividida em dez alíquotas através de quarteamento do material pulverizado abaixo de 0,075 mm, de forma a se dispor de dez repetições de todo o procedimento em amostras “equivalentes”. As espas são apenas uma lembrança de que para se obter as diferentes alíquotas da mesma amostra, as amostras foram divididas. Este procedimento introduz um erro que pode ser significativo. A abertura utilizada foi a de solubilização em água régia sob refluxo em 40 g de amostra. Cada alíquota, após ataque ácido, gerou duas porções-teste. Uma porção para ser levada ao ICP e outra para AA. A etapa de extração por solvente foi realizada independentemente nas duas alíquotas. Para o ICP, após a extração é feita uma secagem e retomada com água régia em pequeno volume.

As amostras selecionadas são apresentadas na Tabela 4.11, juntamente com os resultados encontrados no estudo. A última coluna, onde se apresenta a covariância entre os resultados dos dois métodos, representa o detalhe de maior interesse no presente exemplo. Pode-se perceber que para várias amostras, existe uma covariância expressiva entre os resultados dos dois métodos. A Figura 4.27 mostra para cada uma das amostras, gráficos de correlação, onde os pontos representam os resultados em cada alíquota testada nos dois métodos. Cada gráfico representa uma amostra. Os dez pontos em cada um deles são as dez alíquotas quarteadas em cada amostra. A linha sólida representa a igualdade entre os resultados. As concentrações estão expressas em partes por bilhão (ppb ou ng/g). Os dados apresentados como PL se referem ao plasma (ICP-AES) e FG significa forno de grafite.

As amostras estão colocadas aproximadamente por ordem de teor. Pode-se perceber que os teores variam consideravelmente de uma alíquota para outra na mesma amostra. Para teores mais baixos, por exemplo, abaixo de 50 ppb (50 ng/g), não parece haver correlação expressiva entre as duas seqüências de dados. Isto parece indicar que o erro devido à preparação de amostras é da mesma ordem de grandeza que o erro da análise química nesta faixa de teores. À medida em que o teor de ouro aumenta, parece predominar o erro de preparação e quarteamento das alíquotas.

Nestas condições, comparar os métodos sem considerar a correlação entre as alíquotas da mesma amostra pode levar a conclusões viciadas.

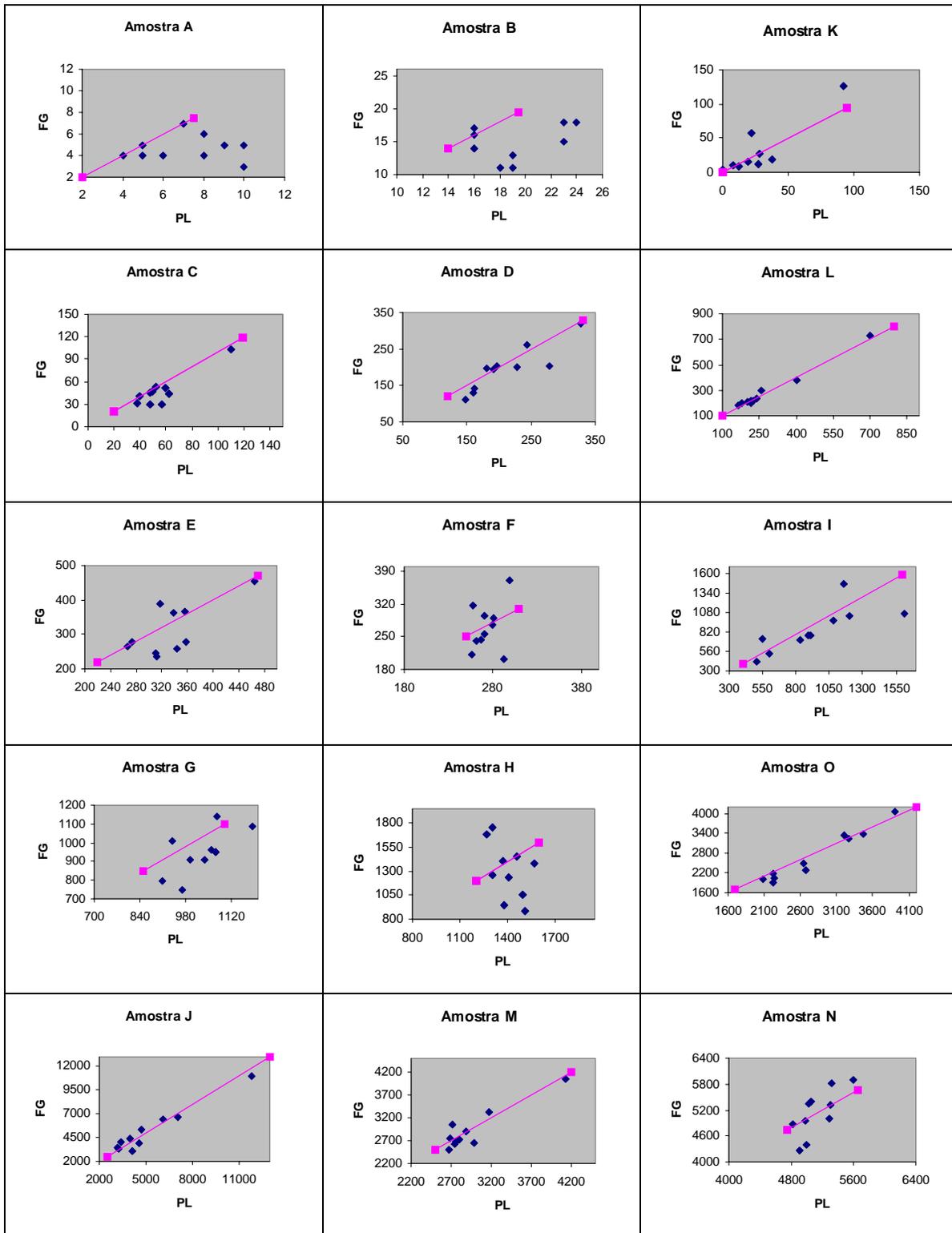


Figura 4.27: Gráficos para se visualizar as correlações entre alíquotas de uma mesma amostra.

No capítulo 2 é apresentado como único método de comparação que permite levar em conta este efeito de correlação expressiva dentro do mesmo nível experimental, o proposto por Riu e Rius (1996). Os resultados de aplicação dos métodos estatísticos se encontram na Tabela 4.12. Para melhorar a eficiência dos cálculos, nesta tabela e gráficos a seguir, os dados foram divididos por mil (o que equivale a transforma-los para a unidade em  $\mu\text{g/g}$ ).

Tabela 4.11 – Resultados de análise de Au (ng/g) por ICP-AES e AA-FG

<i>Amostra</i>	<i>ICP-AES</i>	<i>s(ICP)</i>	<i>AA-FG</i>	<i>s(AA)</i>	<i>Cov(ICP,AA)</i>
A	7,2	2,1499	4,7	1,1595	0,06
B	19,0	3,2318	14,7	2,5841	2,80
C	56,5	20,3484	47,4	21,4952	357,80
D	211,5	57,6161	196,5	61,4605	2838,15
E	334,2	55,7371	313,4	73,9462	2607,82
F	273,4	14,3077	270,5	50,3879	221,20
G	1030,8	80,0844	945,7	118,1290	5629,74
H	1407,5	99,0502	1306,8	288,7940	-13237,00
I	934,2	343,4381	852,7	299,7069	71013,56
J	5228,2	2634,1240	5163,7	2403,0290	5552560,76
K	24,0	27,3496	28,3	37,8449	816,22
L	281,0	160,9824	287,3	167,2151	24041,60
M	2955,8	440,0098	2928,6	464,2031	168560,82
N	5129,8	239,7961	5128,7	540,0442	85420,64
O	2791,1	629,5708	2698,6	749,7493	417494,04

Tabela 4.12 – Resultados dos testes de comparação usando vários estimadores de regressão

<b>Método</b>	<b>a</b>	<b>b</b>	<b>S<sub>a</sub></b>	<b>S<sub>b</sub></b>
<b>MQ</b>	-23,733	0,993	12,817	0,0057
<b>MQP</b>	-2,423	0,958	1,1801	0,0543
<b>RR</b>	-2,800	0,967	0,5615	0,0115
<b>RT</b>	-2,941	0,993	2,2160	0,0651
<b>Deming</b>	-24,044	0,993	12,817	0,0057
<b>Linnet</b>	-2,164	0,971	0,8416	0,0211

A aplicação dos métodos RR e MQP levam à conclusão de que tanto os erros fixos quanto proporcionais são significativos entre métodos. Os erros fixos são da ordem de 3 ppb (forno de grafite mais baixo que ICP), e os erro proporcionais são de 3 a 4% relativo (FG mais baixo). Neste exemplo, RR e MQP fornecem conclusões semelhantes. Também há uma sugestão de que métodos não-ponderados (MQ e Deming) estimam a interseção da reta com grande incerteza.

O método de MQ levaria à conclusão de que ou há diferença fixa ou proporcional entre as técnicas comparadas. Já o método RT aceita a hipótese de erro proporcional não significativo, porém com erro fixo significativo no intervalo  $[-4,3; -1,8]$  ppb (isto significa que o método por AA-FG apresenta resultados ligeiramente menores). Na prática uma estimativa de erro de até 5 ppb entre métodos de análise válidos na faixa de 1 a 5000 ppb não é uma forte indicação de vício entre métodos.

A Figura 4.28 mostra os resultados informados na Tabela 4.12 de forma gráfica, permitindo uma melhor visualização do desempenho dos diversos estimadores usados neste exemplo. As barras em cada ponto representam erros de mais ou menos um erro padrão para a estimativa de  $a$ .

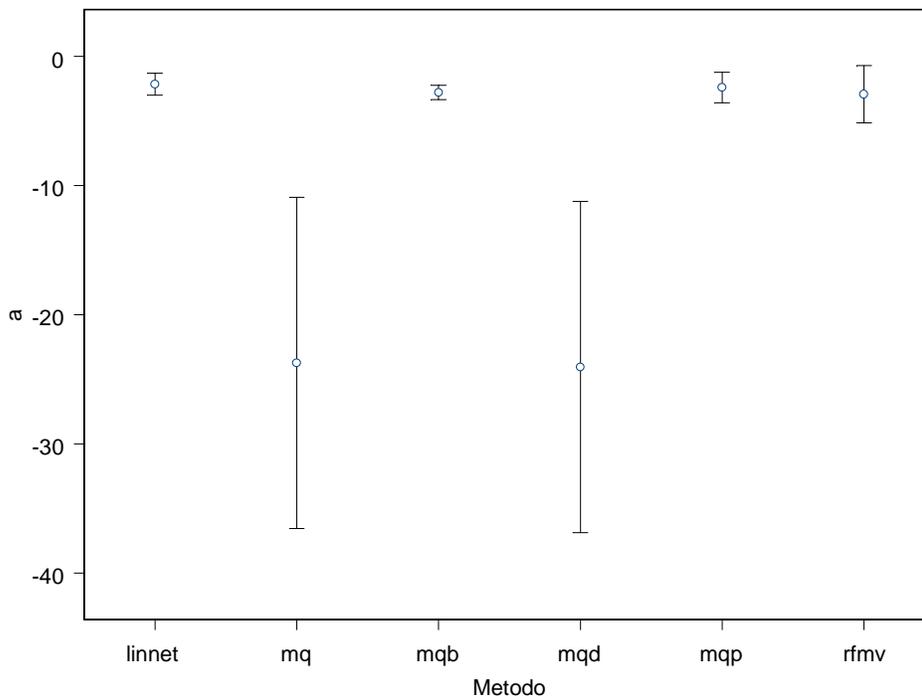


Figura 4.28: Resultados obtidos com a aplicação dos diversos estimadores de regressão no exemplo de análise de ouro. As estimativas apresentadas são da intercessão da reta ( $a$ ) colocando como barra de erros no gráfico o valor de uma vez o respectivo erro padrão. A Figura 4.29 mostra o mesmo para as estimativas do coeficiente angular ( $b$ ). Neste gráfico, mqd se refere ao método de Deming e rfmv ao de Ripley e Thompson. O método mqb é ode RR.

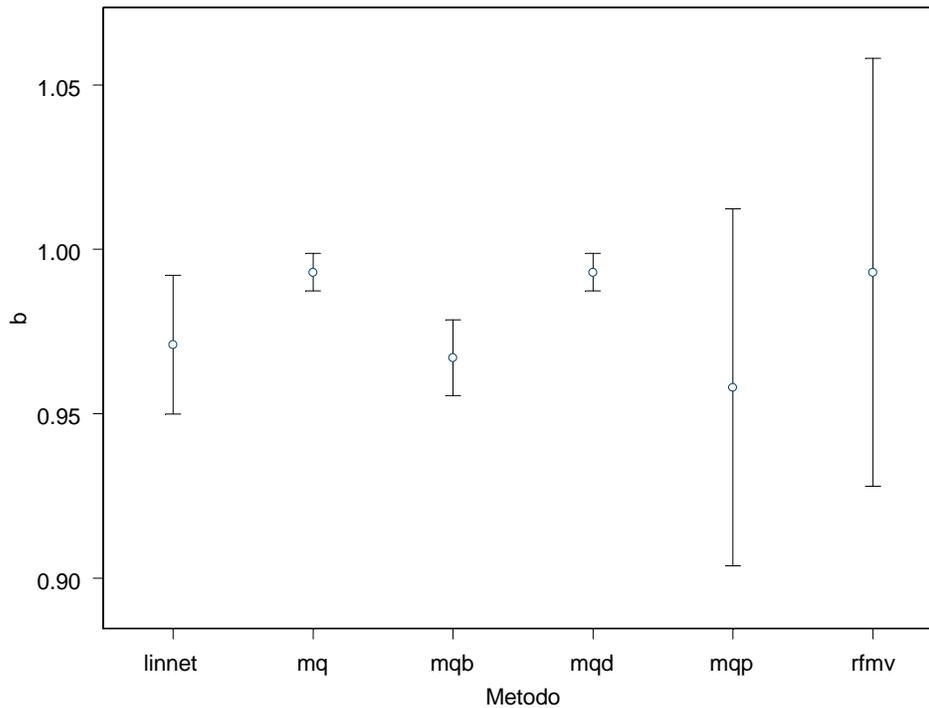


Figura 4.29: Representação gráfico dos resultados apresentados na Tabela 4.12

Por esta figura, percebe-se que os estimadores para os dois parâmetros, com menor erro, são os de mínimos quadrados bivariados (mqb), de Riu e Rius, e o estimador proposto por Linnet (1990).

As Figuras 4.30 e 4.31 ilustram a comparação dos resultados médios obtidos pelos dois métodos de análise química. Os resultados aqui são divididos por mil, portanto a unidade de medida passa de ng/g a  $\mu\text{g/g}$  (que é equivalente a partes por milhão – ppm), o que evita problemas computacionais na geração das elipses de confiança da Figura 4.31.

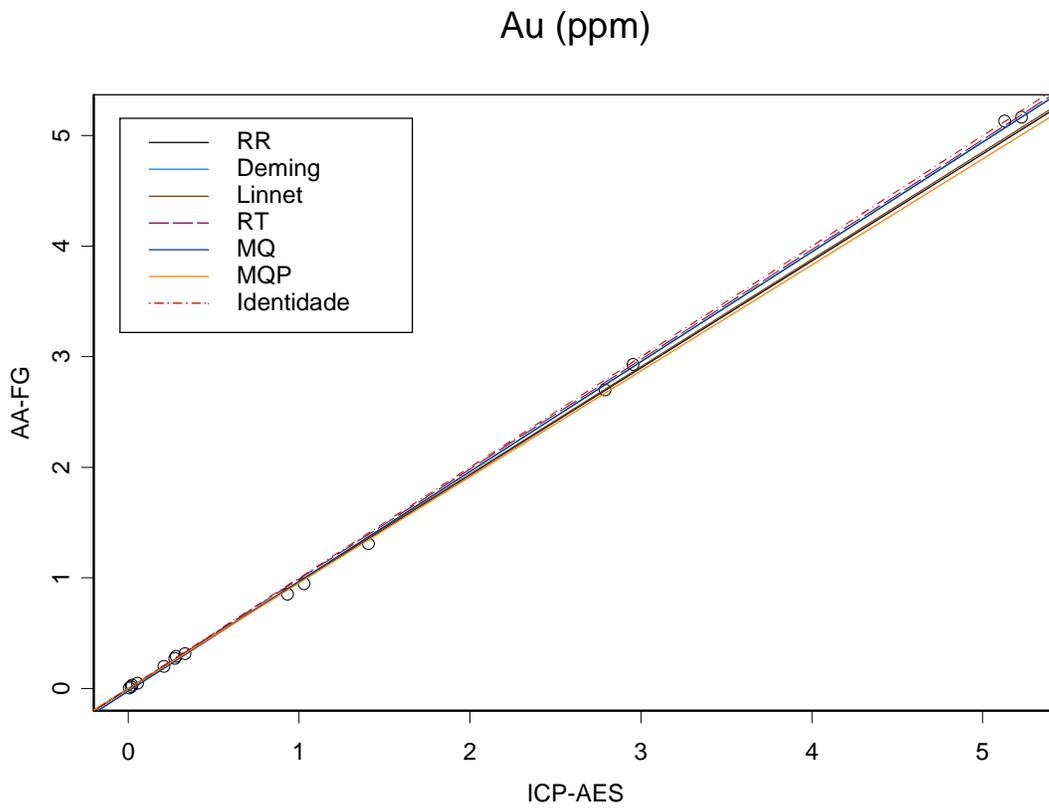


Figura 4.30: Retas calculadas por vários métodos para comparação de análises de Au ( $\mu\text{g/g}$ ) por plasma de acoplamento indutivo (ICP-AES) e absorção atômica em forno de grafite (AA-FG).

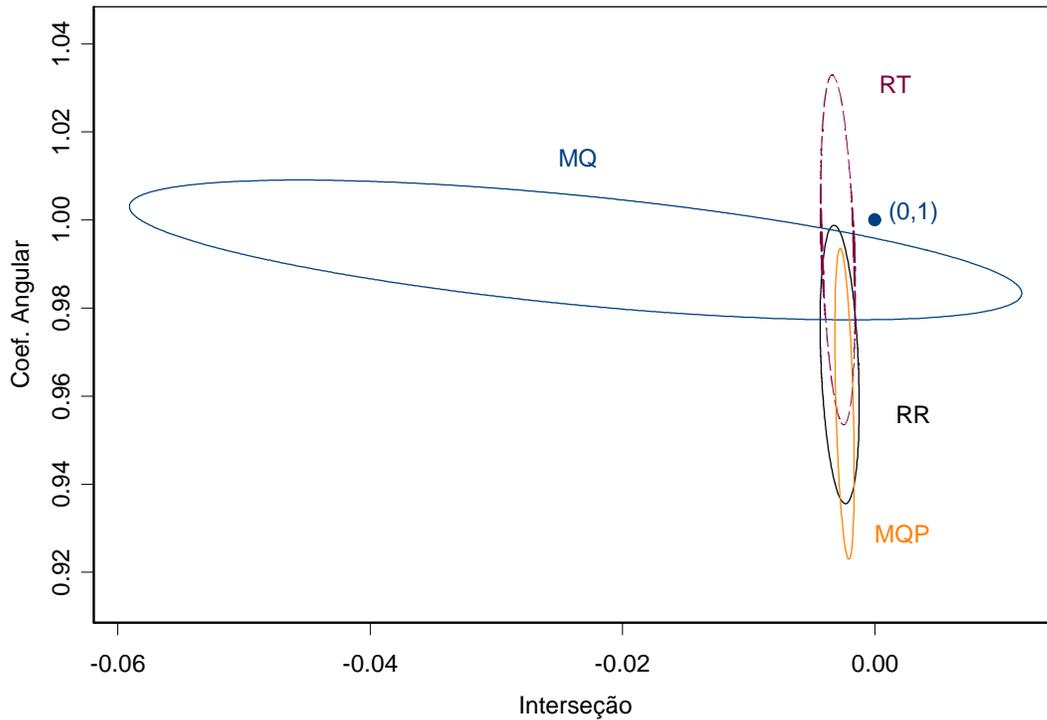


Figura 4.31: Regiões de 95% de confiança para alguns métodos usados na Figura 4.30.

Neste exemplo pode-se perceber a grande incerteza na estimação do coeficiente linear da reta quando se utiliza o método de mínimos quadrados ordinário.

Apesar de que as regiões de confiança para cada estimador diferirem bastante, pode-se ver na Tabela 4.12 que todos apontam para uma inclinação menor que a unidade e uma intercessão menor que zero. Isto indica que o método com forno de grafite encontra sistematicamente menos ouro que o método por emissão atômica em plasma.

As precisões dos dois métodos sendo comparados foram obtidas através das dez réplicas realizadas em cada amostra por cada método. A Figura 4.32 mostra graficamente as duas curvas de precisão obtidas.

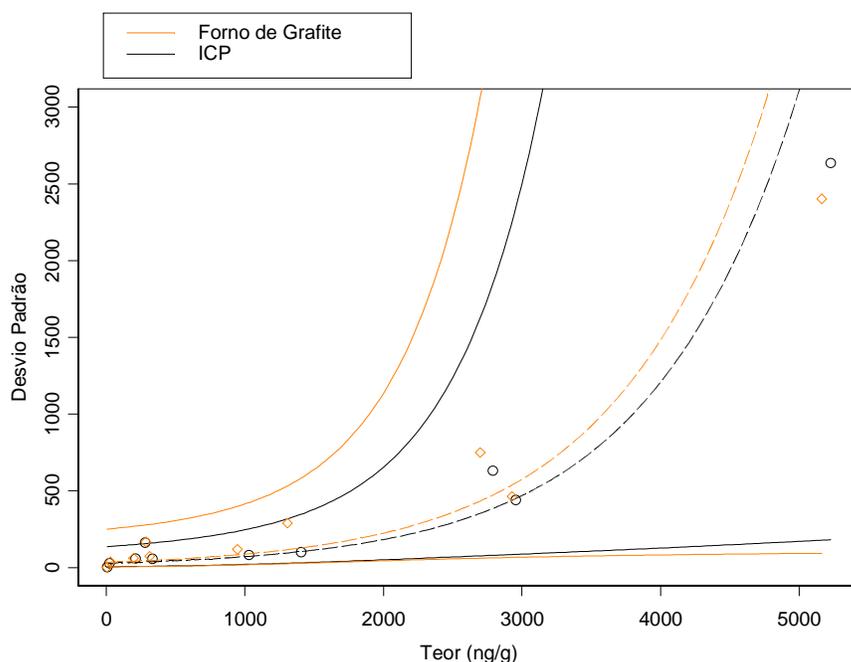


Figura 4.32: Modelagem do desvio padrão de análise em função da concentração para os dois métodos comparados.

Nesta modelagem foram excluídas algumas amostras, pois estas se agrupam em duas categorias distintas, cada uma proveniente de uma região geológica diferente. A tentativa de modelar todos os dados juntos indica que o comportamento das amostras quanto à precisão da análise química é diferente para diferentes origens do material. As amostras mantidas foram as identificadas como A, D, E, G, H, J, K, L, M e O, que são oriundas da mesma origem, e constituem a maioria das amostras.

Um aspecto interessante nesta modelagem foi que o melhor modelo matemático se apresentou como sendo da forma  $\sigma_i = e^{\sigma_0 + \kappa \cdot x_i}$ , que não é encontrada na pesquisa bibliográfica do presente trabalho. Para a análise por plasma, a equação obtida foi

$$\sigma_i = e^{3,3114 + 0,000946 \cdot C_i}$$

Nesta expressão,  $C_i$  se refere à concentração na amostra, dada em ppb (ng/g). Para o forno de grafite, a equação de precisão obtida foi

$$\sigma_i = e^{3,5186 + 0,000946 \cdot C_i}$$

o que parece indicar que a precisão do método por forno de grafite é ligeiramente pior.

#### 4.3.2. Análise de cloro em minérios de cobre utilizando espectrometria de fluorescência de raios X dispersivos em comprimento de onda. Dados de dois diferentes laboratórios.

Para este estudo de comparação, dispõe-se de 1057 amostras analisadas em dois laboratórios diferentes. Os dois laboratórios analisam cloro total por espectrometria de fluorescência de raios X dispersivos em comprimento de onda (FRX). Ambos utilizam amostras preparadas na forma de pós prensados, em pastilhas de aproximadamente 40 mm de diâmetro. No primeiro laboratório a massa utilizada é de 10g, prensada em fôrmas de alumínio com aglomerante comercial (cera sintética, específica para FRX), e o material é pulverizado em condições controladas, em moinho de panelas de carbetto de tungstênio. No segundo laboratório, 1 g de material é pulverizado em graal de cerâmica e espalhado sobre ácido bórico no pistão interno da prensa hidráulica.

No primeiro laboratório, a calibração é feita contra resultados de análise de cloro por colorimetria. No segundo, utilizam-se alguns materiais de referência comerciais, com matrizes de rochas e solos variados.

A fonte de que se dispõe para conhecimento dos erros de análise em ambos os laboratórios é o banco de dados de controle de qualidade de rotina, onde 5% das amostras que passam por análise química são repetidas para efeito de avaliação. Do primeiro laboratório dispõe-se de 990 análises em duplicata, e do segundo, 498 duplicatas.

A idéia, para se aproveitar estes dados, é buscar conhecer como a diferença entre duplicatas varia com a média das duas. Tendo-se uma relação funcional que explique a variação do desvio padrão das diferenças em relação à concentração, pode-se deduzir facilmente a equação de precisão do método, por regressão. Para se conhecer como o desvio padrão das duplicatas varia com o teor, cria-se uma janela móvel, por exemplo de tamanho 100, e dentro desta janela se calcula o desvio padrão das diferenças entre as duas análises, e a média dos teores. Esta linha não paramétrica é posteriormente utilizada para uma análise de regressão onde se estima os parâmetros para a relação funcional. No presente caso a regressão realizada utilizou o método de mínimos quadrados ordinários.

As Figuras 4.33 e 4.34 mostram este estudo, feito para o primeiro e segundo laboratórios, respectivamente. O gráfico da esquerda mostra os pontos (diferenças absolutas *versus* médias), a linha não-paramétrica encontrada pelas janelas móveis, e linhas retas calculadas a partir desta última, por regressão simples. As linhas retas adicionais representam os intervalos de confiança de 95% e 99% para as diferenças entre dois resultados (as duas retas superiores nos gráficos). Estas últimas permitem verificar se os pontos ficam contidos nos intervalos calculados. O gráfico da direita mostra a diferença de teores nos extremos da janela móvel em função do teor do início da janela. A estimativa feita desta forma se baseia na suposição de que o teor varia muito devagar em uma base de dados como esta, devido ao grande número de pontos, logo o desvio padrão do erro é aproximadamente constante dentro da janela, podendo-se assim utilizar o desvio padrão das diferenças dentro da janela

como sendo uma estimativa do erro de análise para a concentração média dentro da janela. A importância do gráfico da direita é a de verificar se esta diferença de teores dentro da janela realmente é pequena.

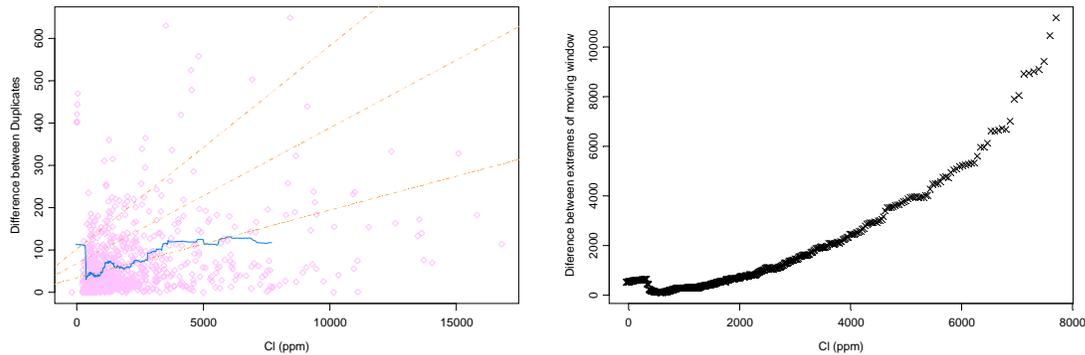


Figura 4.33: Avaliação das duplicatas do primeiro laboratório.

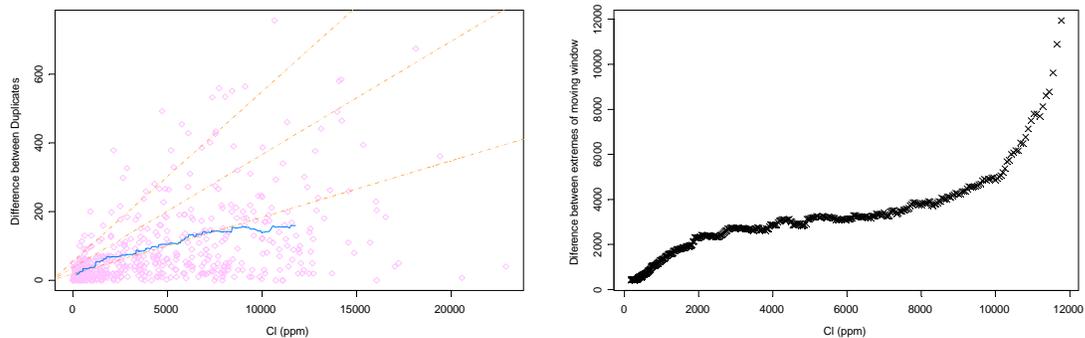


Figura 4.34: Avaliação das duplicatas do segundo laboratório.

Em ambas as figuras percebe-se que nas faixas mais altas de teores a diferença dentro da janela móvel não é pequena. Entretanto, a reta calculada para os pontos gerados pelas médias e desvios padrão móveis parece explicar bem o comportamento do erro mesmo nestes teores, embora ao final exista uma possível curvatura. Provavelmente isto se dá devido ao fato de se dispor de muitos pontos na faixa onde as suposições feitas são válidas, e estes pontos por sua vez dominam na definição dos parâmetros da reta de regressão. Partindo-se do pressuposto de que a relação linear é a verdadeira relação subjacente aos erros de análise, a extrapolação para teores um pouco além da faixa válida não acarreta grandes erros. Pode-se ver que as linhas de 95 e 99% de confiança para as diferenças contêm realmente a maior parte dos pontos.

Entretanto, a suposição de linearidade pode ser melhor investigada. Uma inspeção mais detalhada dos pontos gerados pelas janelas móveis permite verificar que o ajuste de potência é melhor. A Figura 4.35 mostra este ajuste aos pontos gerados para os dois laboratórios simultaneamente. Para o primeiro laboratório foram considerados apenas os

dados acima de 365 ppm de Cl, pois abaixo desta concentração a presença de pontos anômalos prejudica a estimação da relação funcional da precisão.

### Modelagem da Precisão

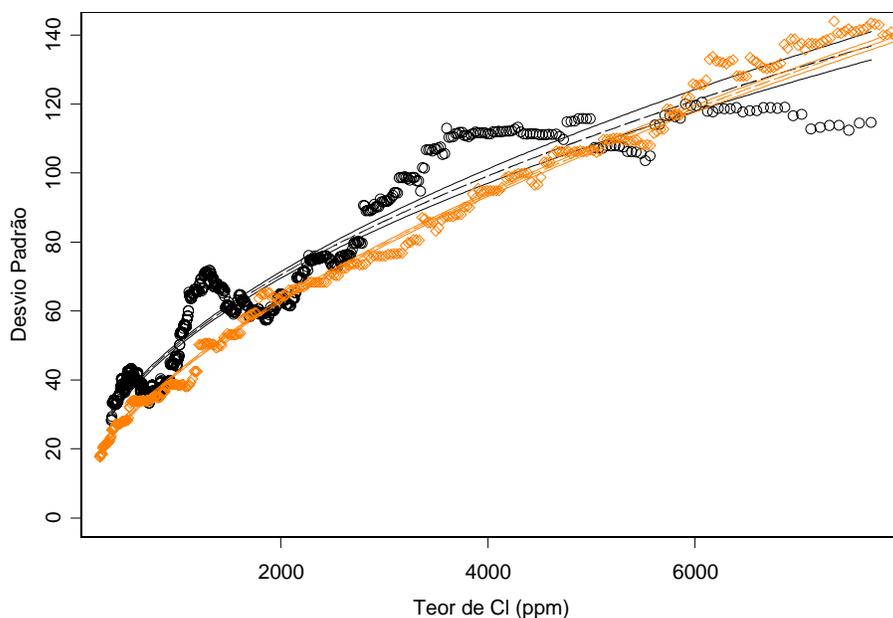


Figura 4.35: Comparação das precisões dos dois laboratórios

As equações de precisão obtidas na modelagem são:

$$s_{x_i} = 1,669555 \cdot x_i^{0,49238}, \quad \text{Primeiro laboratório (pontos pretos)}$$

$$\text{Correlação} = 0,952$$

e

$$s_{x_i} = 0,778367 \cdot x_i^{0,57783}, \quad \text{Segundo laboratório (pontos alaranjados)}$$

$$\text{Correlação} = 0,997 .$$

Uma vez de posse das equações de precisão, pode-se aplicá-las para o cálculo dos desvios padrão das 1057 amostras analisadas nos dois laboratórios, e posteriormente aplicar o método de comparação. A Figura 4.36 mostra a comparação destas análises e a Figura 4.37, o teste de hipótese simultâneo para os parâmetros.

### Cl (% em massa)

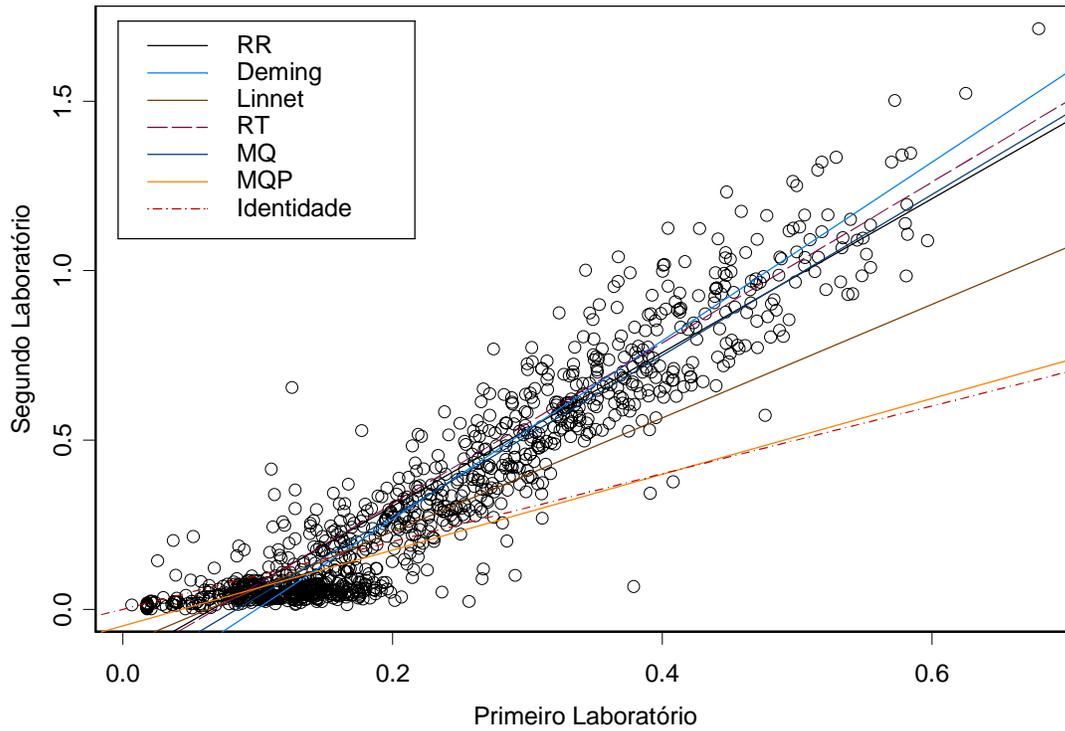


Figura 4.36: Comparação dos dois métodos de fluorescência de raios X para análises de cloro total em minérios de cobre. Aqui os dados foram transformados para porcentagem em massa (% m/m). Isto se deve ao fato de que com valores muito altos, a rotina de cálculo usada deforma os elipsóides de confiança no gráfico com o teste de hipótese.

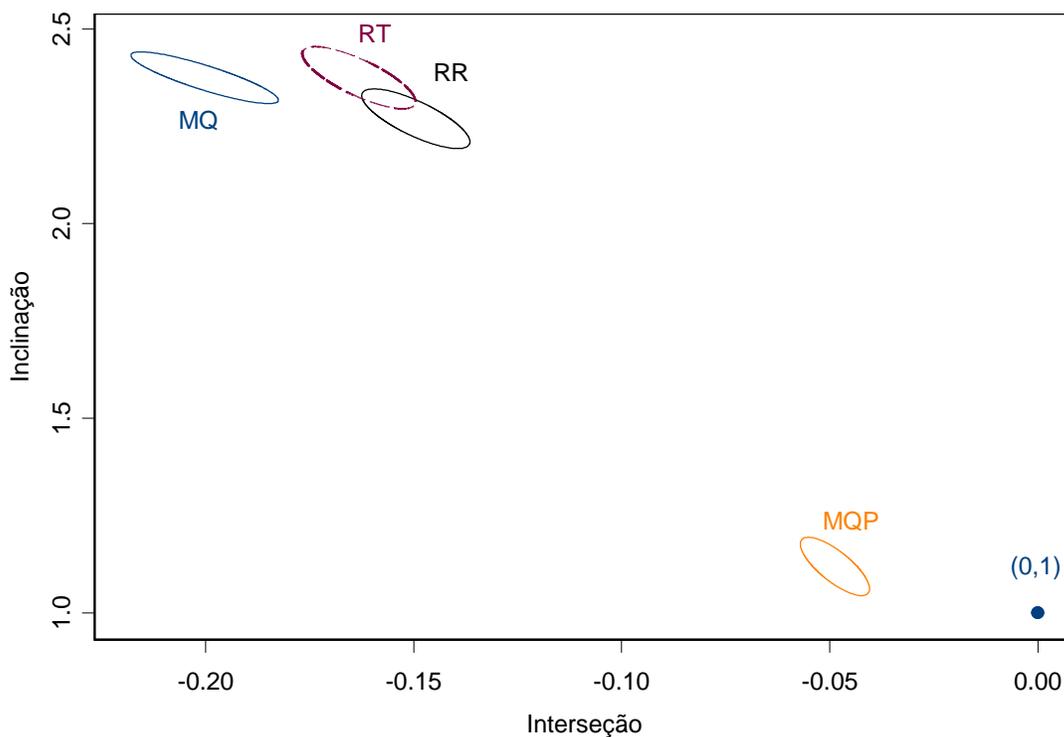


Figura 4.37 – Elipsóides de confiança para os parâmetros da reta, obtidos pelos diversos métodos de regressão.

A Tabela 4.13 apresenta o resumo das avaliações estatísticas. A correlação entre os dois métodos é de 0,939. Resultados em parte por milhão ( $\mu\text{g/g}$ ).

Tabela 4.13 – Resultados dos testes de comparação usando vários estimadores de regressão

Método	a	b	Sa	Sb
RR	-1314,53	2,2206	49,8323	0,02987
Deming	-2638,48	2,6499	78,6078	0,02983
Linnet	-1055,99	1,6800	186,9423	0,11963
RT	-1669,21	2,3746	5,4459	0,00304
MQ	-2002,20	2,3746	72,3049	0,02712
MQP	-778,222	1,4396	1,3064	0,00098

As Figuras 4.38 e 4.39 mostram estes mesmos resultados de forma gráfica, facilitando a interpretação.

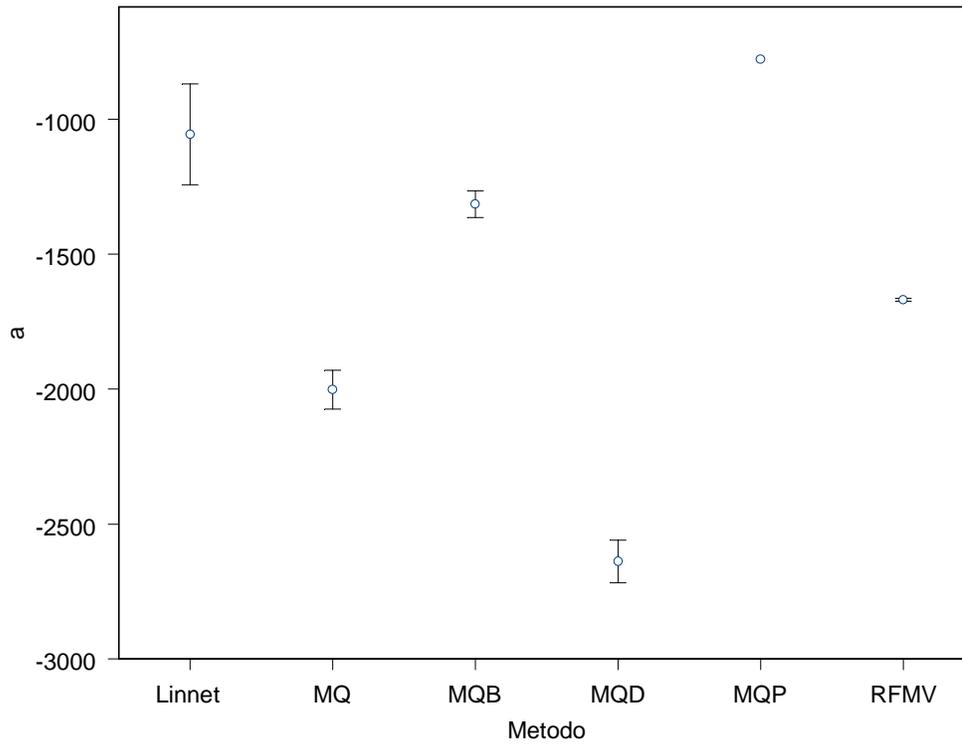


Figura 4.38: Apresentação dos resultados de regressão da Tabela 4.13 de forma gráfica. Pontos representam estimativas da interseção da reta, e as barras associadas mostram o erro padrão calculado para os interceptos.

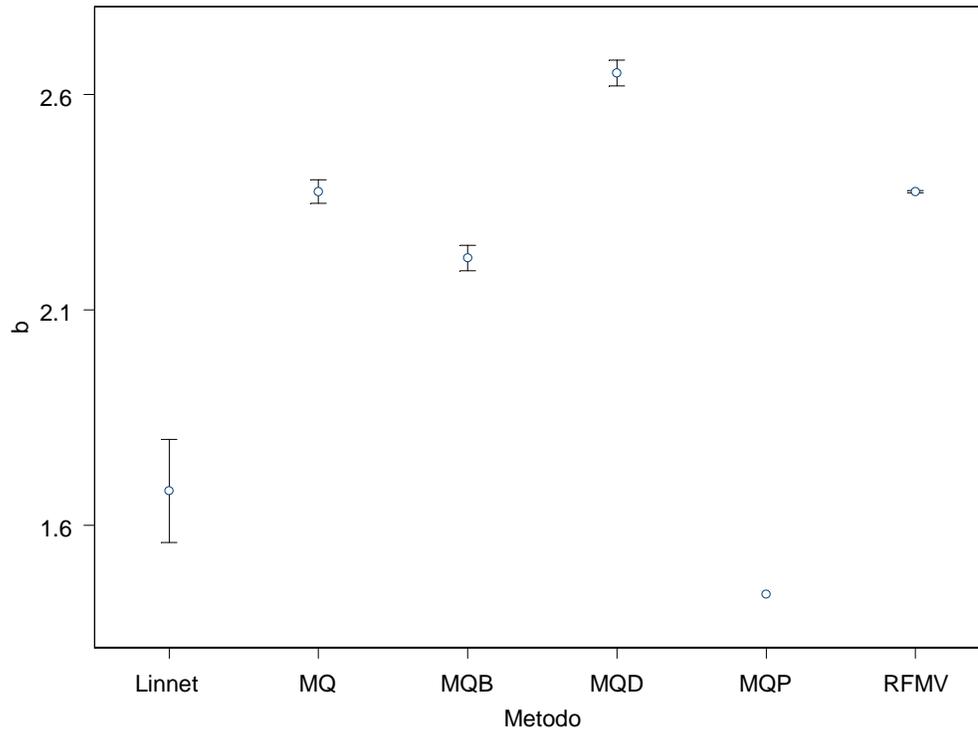


Figura 4.39 – Comparação entre as estimativas do coeficiente angular com respectivos erros padrão.

Os métodos que apresentam os menores erros de estimação são RT e MQP, entretanto os resultados por eles obtidos são muito diferentes. Neste exemplo, pode-se verificar que a escolha do método de regressão é crucial para a interpretação dos resultados. Não há muita semelhança entre estimativas. Para o coeficiente angular, pode-se perceber que se aproximam mais os métodos RR, RT e MQ. Como os demais (MQP, Deming e Linnet) não utilizam a modelagem adequada da precisão, sendo que no primeiro esta só é considerada para o método do segundo laboratório, não é de se esperar que produzam resultados confiáveis. O método MQ, sabidamente produz resultados viciados, portanto não é confiável.

#### 4.3.3. Comparação de dois métodos de análise de manganês em minério de ferro por absorção atômica em chama.

Este exemplo é um caso de aprovação de métodos de análise química num comitê de normalização internacional. Uma vez aprovados, estes dois métodos passam a ser considerados métodos de referência para comparação com outros que venham a ser utilizados ou desenvolvidos. Periodicamente o comitê que gerou a norma onde os métodos estão descritos conduz uma revisão dos mesmos. Nestas ocasiões podem surgir modificações nos métodos, o que muitas vezes suscita a necessidade de se conduzir outro teste interlaboratorial para rever as cláusulas de precisão, ou ainda mesmo alteração do 'status quo' do método (por exemplo, passando de método de referência para método de rotina). Dependendo da aplicação da norma, estas mudanças trazem conseqüências comerciais. Por isso estes métodos são avaliados com mais rigor que em outros casos encontrados na literatura. No presente estudo apenas se faz a comparação entre os dois procedimentos descritos na norma, o que representa apenas uma parte do que seria necessário para sua aprovação.

A análise de manganês em minérios de ferro por técnicas de absorção atômica é descrita na norma ISO 9682-1:1991. Os testes interlaboratoriais que geraram as estatísticas de precisão para inclusão nesta norma foram finalizados em 2005 e são apresentados no documento ISO/TC 102/SC 2 N 1417 E. Na Tabela 4.14 estão os dados resumidos após tratamento dos resultados do teste. O comitê técnico TC 102 da ISO é o responsável pela normalização dos métodos de avaliação da qualidade de minérios de ferro e enfoca produtos finais (minérios) comercializados no mercado transoceânico.

Este comitê cedeu os dados para o presente estudo. A técnica de absorção atômica em chama foi estudada pelo grupo de trabalho que gerou a norma ISO 9682-1, utilizando dois procedimentos diferentes de abertura: solubilização em meio ácido e fusão alcalina. Os resultados produzidos pelo grupo são obtidos e tratados de acordo a norma ISO 5725-2, que trata de experimentos de precisão. Basicamente, esta norma apresenta um planejamento de experimento para tratamento de dados por meio de uma análise de variância de efeito aleatório com um fator de classificação. Esta norma foi desenvolvida pelo comitê de aplicações estatísticas da ISO, o TC 69.

Após os cálculos efetuados, seguindo-se as recomendações desta norma, são produzidas estimativas do teor de manganês em cada amostra, assim como os desvios padrão de repetibilidade ( $S_r$ ), reprodutibilidade ( $S_R$ ) e desvio entre laboratórios ( $S_L$ ). Com isto é possível encontrar expressões para a tolerância entre análises feitas no mesmo laboratório e entre laboratórios. Para efeito do presente estudo, serão utilizados apenas os resultados médios obtidos utilizando cada um dos dois procedimentos, e os desvios padrão de repetibilidade. Estes resultados são apresentados na Tabela 4.14.

Tabela 4.14 – Resultados do Teste Interlaboratorial para Análise de Mn por Absorção Atômica. Dados obtidos usando a norma ISO 9682-1, expressos em % em massa.

<i>Amostra</i>	<i>Fusão</i>	$S_r(\text{Fusão})$	<i>Abertura Ácida</i>	$S_r(\text{Ab. Ácida})$
76-16	0,01127	0,0006744	0,01099	0,0004824
83-7	0,08738	0,0013818	0,08665	0,0011051
83-4	0,40128	0,0043719	0,40196	0,0029639
76-12	0,77221	0,0089413	0,76894	0,0097379
83-8	2,51709	0,0241675	2,52821	0,0153892

A modelagem e comparação das precisões analíticas dos dois métodos são ilustradas na Figura 4.40.

### Modelagem da Precisão

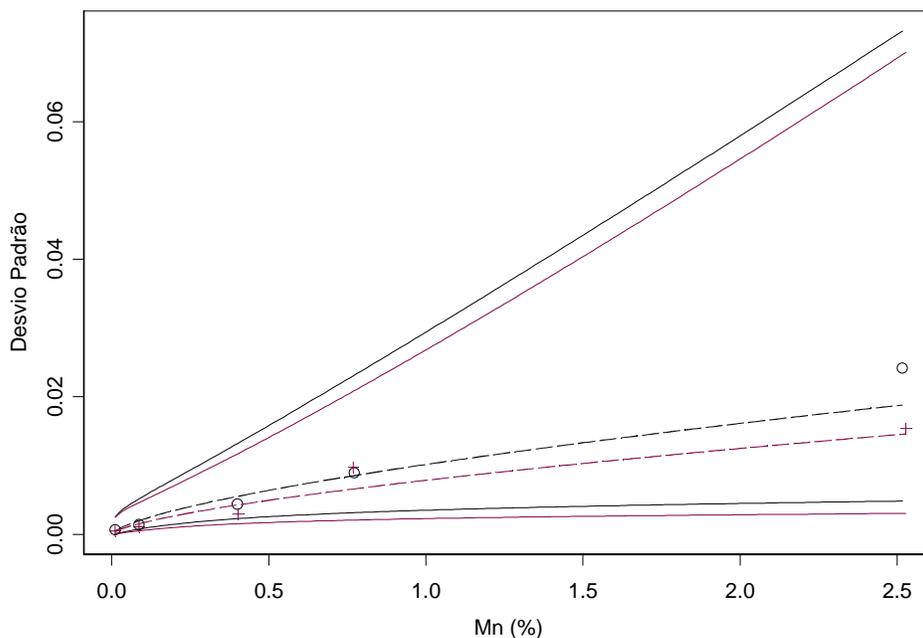


Figura 4.40: Comparação dos dois procedimentos de solubilização de amostra para análises via espectrometria de absorção atômica em chama, em termos de sua precisão de repetibilidade. Linhas e bolas pretas representam o método de fusão alcalina. Linhas e cruces vermelhas representam a abertura ácida.

A comparação das médias obtidas por cada um, utilizando alguns estimadores de regressão, fornece o gráfico da Figura 4.41, e seus resultados são resumidos na Tabela 4.15.

Neste exemplo, o uso de qualquer um dos métodos de análise de regressão leva à mesma conclusão, de que não há evidências de diferenças significativas entre os dois

procedimentos. Entretanto, pode-se perceber que a precisão das estimativas não é a mesma, logo também diferem poder do teste e o intervalo de confiança associado a cada estimador.

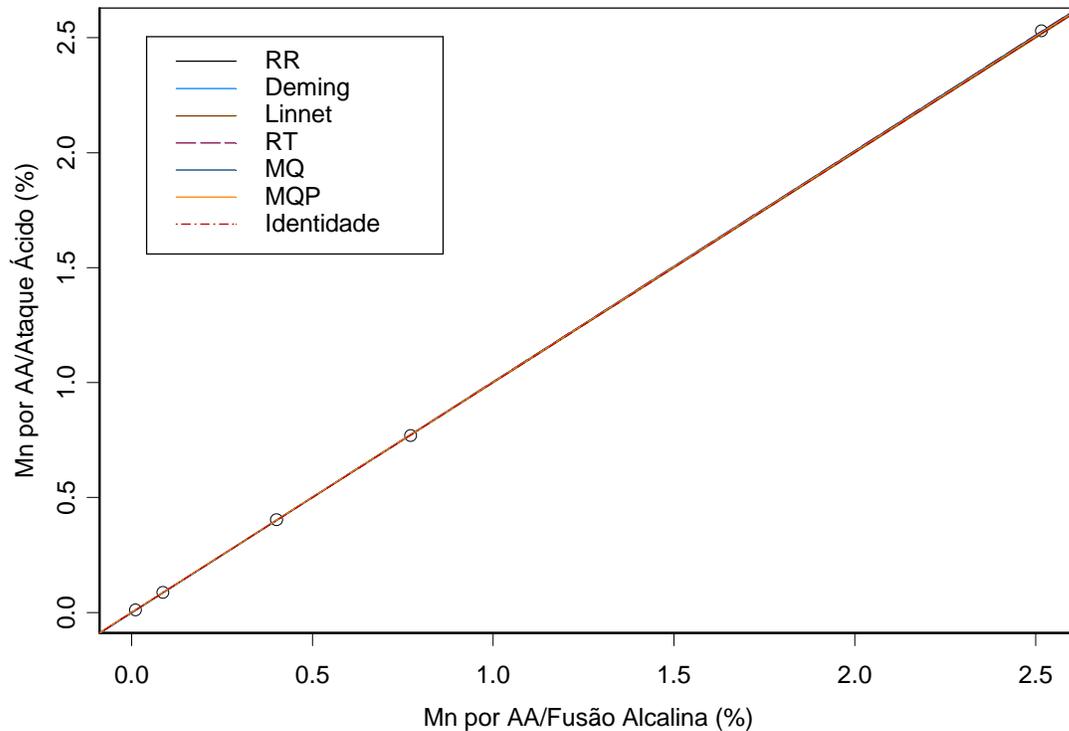


Figura 4.41: Comparação das médias dos dois procedimentos de abertura para determinação de Mn por AA em minérios de ferro. O coeficiente de correlação de Pearson para esta comparação é de 0,939.

A Figura 4.42 mostra os elipsóides de 95% de confiança para os parâmetros, e aí se vê que todos os testes de hipótese aceitam a hipótese básica  $H_0$ , de igualdade. Pode-se ver claramente a grande incerteza associada à estimativa do intercepto de regressão para o método de mínimos quadrados ordinários. Todos os demais estimadores apresentam regiões de confiança mais próximas entre si.

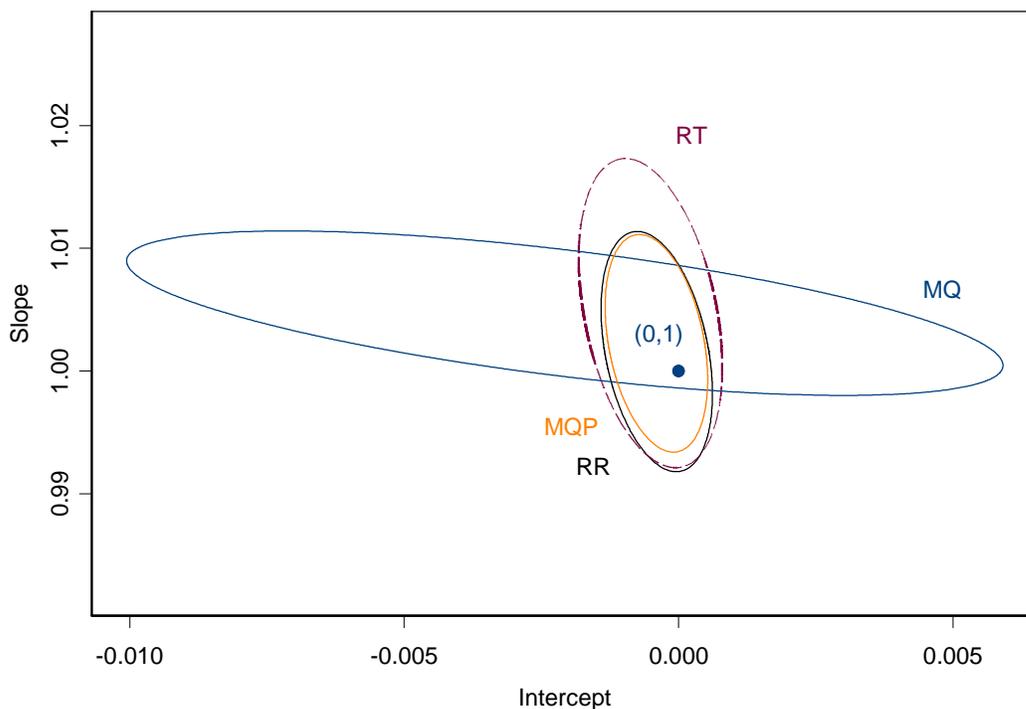


Figura 4.42: Elipsóides de 95% de confiança para os parâmetros.

Tabela 4.15 – Resultados dos testes de comparação usando vários estimadores de regressão.

Método	A	b	Sa	Sb
<b>MQ</b>	-0,0020748	1,004722	0,00182626	0,00153247
<b>MQP</b>	-0,0004059	1,002264	0,00046325	0,00439870
<b>RR</b>	-0,0003999	1,001591	0,00023211	0,00223891
<b>RT</b>	-0,0005141	1,004722	0,00079490	0,00766728
<b>Deming</b>	-0,0020784	1,004727	0,00182626	0,00153247
<b>Linnet</b>	-0,0002818	0,999610	0,00049328	0,00312866

A Figura 4.43 mostra estes mesmos resultados de forma gráfica, facilitando a interpretação.

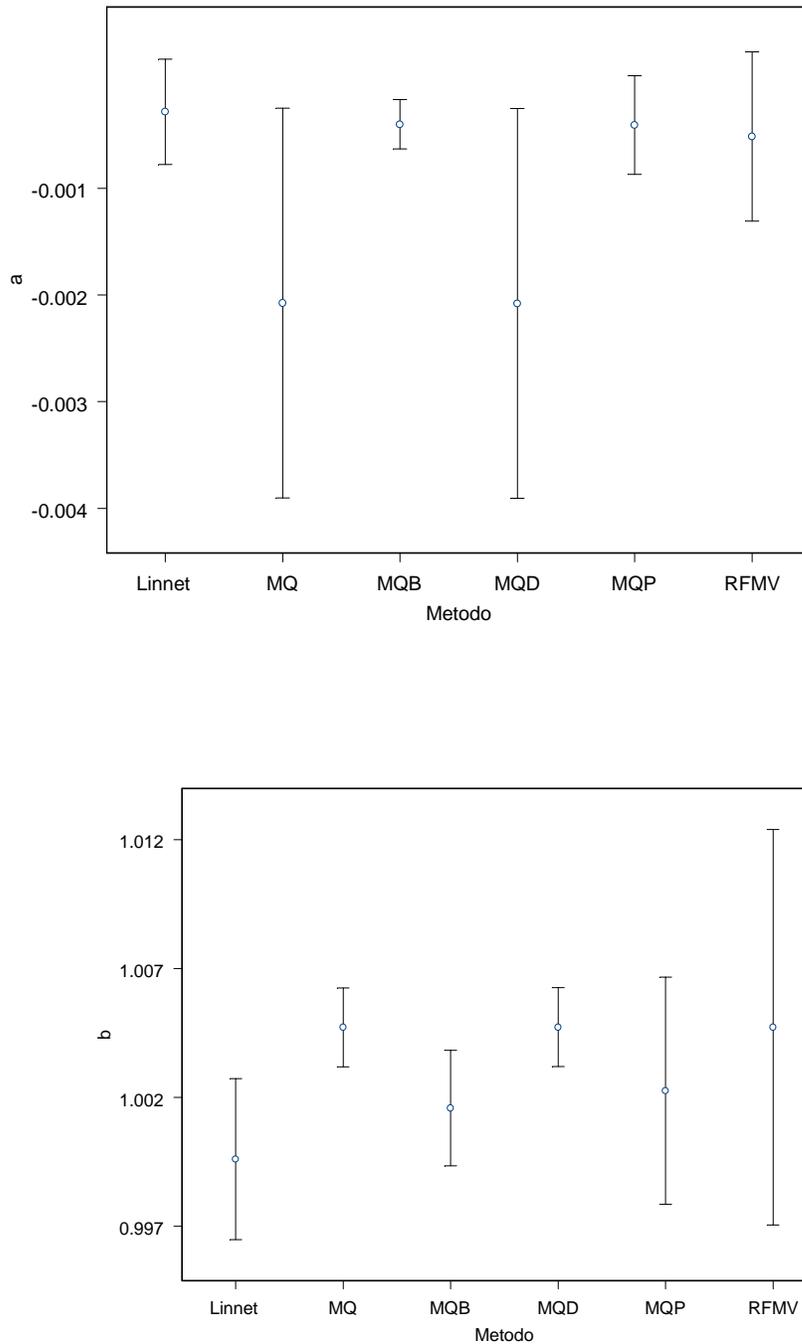


Figura 4.43: Apresentação dos resultados da Tabela 4.15 de forma gráfica. O gráfico de cima contém as estimativas do termo independente na regressão para os vários métodos usados. O gráfico de baixo contém as estimativas para o coeficiente angular. As barras de erro coincidem com o erro padrão estimado para cada parâmetro. Nestes gráficos MQB é o método de RR, MQD é o de Deming, e RFMV o de RT.

## Capítulo 5

### Conclusões

O estudo das publicações na área de química analítica que fazem referência à comparação de métodos mostra uma grande quantidade de alternativas, quando se refere à comparação dos resultados obtidos pelos dois métodos sendo estudados. Nem todas as abordagens encontradas são recomendadas para utilização neste tipo de problemas. Parte das limitações encontradas por vários autores, que faz com que utilizem técnicas estatísticas não adequadas, se refere aos recursos computacionais necessários para uma solução mais rigorosa.

Hoje em dia, isto não é mais uma restrição, já que os microcomputadores disponíveis aos profissionais são capazes de realizar cálculos complexos em grande velocidade.

Sendo assim, não há motivos para não utilizar métodos de aplicação mais genérica, e que apresentam melhor desempenho.

No presente estudo, fica evidente que as técnicas de regressão ponderada são as mais indicadas para a comparação de métodos em análise química. Mesmo no caso em que os coeficientes de variação são pequenos, onde Cornbleet e Gochman (1979) sugerem a utilização da regressão de Deming, os resultados são melhores ainda se utilizarmos o método proposto por Riu e Rius (1996). Os métodos que não consideram erros nas abscissas são limitados, pois introduzem vício nas estimativas. O método de mínimos quadrados ponderados apresenta vício compatível com os métodos ponderados que consideram erro nas duas variáveis, porém as simulações mostram que o erro do tipo I ocorre com mais frequência.

De todos, o método de regressão que apresentou as melhores características para uso nos problemas em questão é o de mínimos quadrados bivariados, proposto por Riu e Rius (1996). Além de seu desempenho ser melhor em termos de vício, precisão, consistência e eficiência, ele apresenta vantagens adicionais:

- i) rápida convergência, melhor que o método de máxima verossimilhança proposto por Ripley e Thompson (1987);
- ii) produz uma estimativa da matriz de covariâncias, que pode ser utilizada para inferências posteriores;
- iii) é apresentado pelos autores que o propõem já com o teste conjunto sobre os parâmetros;
- iv) é facilmente adaptado para problemas não-lineares. A extensão para polinômios de ordem maior que um é imediata. Para definição dos pesos, entretanto, a melhor alternativa é a de utilização da expressão proposta por Cecchi (1991).

O algoritmo proposto por Williamson fornece resultados idênticos aos obtidos pelo método de Riu e Rius (1996), porém ele não fornece as equações para o cálculo da região de confiança simultânea para as estimativas dos parâmetros além de não fornecer a matriz de covariâncias dos estimadores dos parâmetros.

A conclusão geral dos métodos ponderados é que o único que se saiu bem comparativamente em todas as situações foi o de mínimos quadrados bivariados. O método proposto por Linnet (1990) acompanhou de perto este desempenho, porém ele não apresenta as vantagens adicionais citadas para o primeiro, além de apresentar erros padrão de estimativas maiores que ele. O método de máxima verossimilhança parece ter vocação para dados com distribuição uniforme, e os problemas para dados log-normais parecem piorar quando as precisões dos métodos químicos são piores, o número de pontos na regressão é muito pequeno (menor que 10). Em geral pode-se dizer que o método RR apresenta melhor desempenho que os demais nos testes de hipótese por ser o único que considera erros heteroscedásticos nos dois eixos e ainda apresenta um teste de hipótese conjunto para os dois parâmetros. Entretanto, mesmo ele falha quando os erros em X aumentam muito (acima de um coeficiente de variação de 50%). A provável explicação desta falha no desempenho é que a região de confiança derivada para o teste de hipótese é deduzida de forma 'ad-hoc', a partir da teoria de mínimos quadrados, sem considerar totalmente a estrutura dos dados neste tipo de problemas. O autor do presente trabalho investigou esta questão e juntamente com outros pesquisadores propõe alternativas que corrigem esta tendência neste estimador, ou que não dependem dele. Estas pesquisas estão publicadas (De Castro et al.- 2005a e Galea-Rojas et al. – 2003). Os mesmos autores avançaram neste trabalho, propondo um método modificado para comparação de mais de dois métodos de análise química (De Castro et al., 2004), propondo um teste exato (De Castro et al., 2005b), e avaliando o modelo de máxima verossimilhança para conhecer os efeitos de perturbações locais nos pontos de regressão (De Castro et al., 2006).

Os métodos não ponderados estudados sempre apresentam pior desempenho quando comparados aos ponderados, sendo que no caso dos métodos robustos, não foi feita simulação para dados contaminados, o que teoricamente mostraria seus pontos fortes. O método LTS de Rousseeuw, entretanto, apresentou bons resultados em termos de vício, quando comparado com o método de Deming para variância constante. O estudo de dados contaminados fica como exercício futuro.

Os estudos de dados reais, da literatura ou fornecidos pelo presente autor, mostram as diferenças que se pode encontrar na prática quando se desconsidera as limitações de cada estimador. Como nem sempre é possível conhecer com segurança a estrutura dos dados (distribuição e erros associados), é mais recomendável a utilização de métodos com melhor desempenho e que fazem menos suposições sobre os dados.

O estudo do erro aleatório em química analítica se mostra como sendo o terreno mais fértil para novos trabalhos, em vista da escassez de material literário que o trate. Isto é tratado de forma resumida em De Castilho (2004). Existem numerosos exemplos



práticos conhecidos pelo autor, em que os modelos teóricos disponíveis ainda não satisfazem as necessidades práticas dos laboratórios. Roche e Lorenzato (1995) previram limitações do seu modelo, e fazem algumas sugestões de ampliação do trabalho, mas muito está para ser feito a este respeito ainda. Não foi apresentada proposta para comparação de métodos em termos de precisão.

Finalmente, a análise conjunta das expressões de precisão e da comparação de resultados entre métodos de análise química, como aqui proposta e demonstrada em casos práticos, permite um conhecimento mais profundo e confiável dos métodos analíticos estudados. Espera-se que contribua para se obter mais um avanço no campo da pesquisa da química analítica.

## Capítulo 6

### Anexo Códigos-fonte dos programas de computador

# Programa para gerar a função 'mq.fcn', que usa os mínimos quadrados simples

```
mq.fcn <- function(x, y){  
  
  n <- length(x)  
  mq <- lsfit(x, y)  
  a <- mq$coef[1]  
  b <- mq$coef[2]  
  R <- y - a - b*x  
  w <- 1  
  rms <- sqrt(sum(R^2)/(n-2))  
  sa <- rms*sqrt(sum(x^2)/(n*sum((x-mean(x))^2)))  
  sb <- rms/sqrt(sum((x-mean(x))^2))  
  Elipse <- CalcEli.fcn(x, y, a, b)  
  
  final <- list(a=a, b=b, A=Elipse$A, B=Elipse$B, C=Elipse$C, D=Elipse$D, SomaR2=Elipse$SomaRes2,  
               RMS=rms, S=Elipse$S, sa=sa, sb=sb, Resid.esc=Elipse$Resid.esc,  
               Area=Elipse$Area, Acertos=Elipse$Acerta)  
  
  final  
}
```

# Programa para gerar a função 'MQPond.fcn', que faz a regressão ponderada,  
# gerando equação da elipse de mínimos quadrados ponderados

```
mqpond.fcn <- function(x, y, var.y){  
  
  mqp <- lsfit(x, y, wt=1/var.y)  
  a <- mqp$coef[1]  
  b <- mqp$coef[2]  
  w <- 1/var.y  
  Elipse <- CalcEli.fcn(x, y, a, b, peso=1/w)  
  
  x.til <- sum(w*x)/sum(w)  
  y.til <- sum(w*y)/sum(w)  
  sa <- sqrt(sum(w*x^2)/(sum(w)*sum(w*(x-x.til)^2)))  
  sb <- 1/sqrt(sum(w*(x-x.til)^2))  
  
  final <- list(a=a, b=b, A=Elipse$A, B=Elipse$B, C=Elipse$C, D=Elipse$D,  
               RMS=Elipse$RMS.y, S=Elipse$S, Area=Elipse$Area, sa=sa, sb=sb,  
               Acertos=Elipse$Acerta, Resid.esc=Elipse$Resid.esc, SomaR2=Elipse$SomaRes2)  
  
  final  
}
```

```
# Programa para gerar a função Rius, que aplica o método de Riu e Rius (1996)

rius.fcn <- function(x, y, var.x, var.y, covar=rep(0, length(x)), itera=10){

mq <- lsfit(x, y)
a <- mq$coef[1]
b <- mq$coef[2]
old.a <- 0
old.b <- 0
w <- 0
dw.db <- 0
R <- 0
s1 <- s2 <- s3 <- s4 <- s5 <- 0
# while (abs(old.b - b)>0.00000001 & abs(old.a - a)>0.00000001)
  for (i in 1:itera) {
    w <- var.y + (b^2)*var.x - 2*b*covar
    dw.db <- 2*b*var.x - 2*covar
    R <- y - a - b*x
    s1 <- sum(1/w)
    s2 <- sum(x/w)
    s3 <- sum(x^2/w)
    s4 <- sum(y/w)
    s5 <- sum(x*y/w + 0.5*(R/w)^2*dw.db)
    old.a <- a
    old.b <- b
    b <- (s5*s1 - s2*s4)/(s1*s3 - s2*s2)
    a <- (s4 - s2*b)/s1
  }
desvio.x <- sqrt(var.x)
desvio.y <- sqrt(var.y)
tam <- cbind(desvio.x, desvio.y)
Ellipse <- CalcEli.fcn(x, y, a, b, peso=w)

dif.a <- abs(old.a - a)
dif.b <- abs(old.b - b)

final <- list(a=a, b=b,
             A=Ellipse$A, B=Ellipse$B, C=Ellipse$C, D=Ellipse$D, Dif.a=dif.a, Dif.b=dif.b,
             RMS=Ellipse$RMS.y, S=Ellipse$S, desvios=tam, Area=Ellipse$Area,
             s=Ellipse$s, Acertos=Ellipse$Acerta, Resid.esc=Ellipse$Resid.esc, sa=Ellipse$SE.a,
             sb=Ellipse$SE.b, SomaR2=Ellipse$SomaRes2)

final
}

# Função para cálculo da regressão de Deming, com Teoria de Mandel (Journal of Quality Techno-
# logy, 1984) para estimar erros padrão e teste simultâneo de Hartmann et al (Analytica
# Chimica Acta, 1997) para as estimativas dos parâmetros

deming.fcn <- function(x, y, sx, sy){

n <- length(x)
SeX <- sqrt(sum(sx^2)/(length(sx)))
SeY <- sqrt(sum(sy^2)/(length(sy)))
```

```

Fisher <- qf(.95, 2, n-2)
lambda <- SeY^2/SeX^2
Sxx <- sum((x-mean(x))^2)
Syy <- sum((y-mean(y))^2)
Sxy <- sum((x-mean(x))*(y-mean(y)))
b <- (Syy/lambda-Sxx+sqrt((Sxx-Syy/lambda)^2+(4/lambda)*Sxy^2))/(Sxy*2/lambda)
k <- b/lambda
a <- mean(y) - b*mean(x)
u <- x + k*y
v <- y - b*x
u.bar <- mean(u)
v.bar <- mean(v)
Suu <- Sxx+2*k*Sxy+k^2*Syy
Svv <- b^2*Sxx-2*b*Sxy+Syy
Se2 <- Svv/(n-2)
Se <- sqrt(Se2)
Sb <- Se*(abs(1+k*b)/sqrt(Suu))
Sa <- Se*sqrt(1/n+mean(x)^2*(1+k*b)^2/Suu)
S <- sum((y-a-b*x)^2)
d <- y-a-b*x
RMS <- S/(n-2)
Fcalc <- (-v.bar^2+2*u.bar*(-v.bar)*(1-b)/(1+b)+(1-b)^2/(1+b)^2*sum(u^2)/n)/(2*Se2/n)

saida <- list(a=a, b=b, n=n, RMS=RMS, S=S, sa=Sa, sb=Sb, Resid=d, Fcalc=Fcalc)
saida
}

# Função para implementar o método ponderado de Deming
# Ref.: Linnet, K. (1990), Statistics in Medicine, Vol 9, pp. 1463-1473

linnet.fcn <- function(x, y, var.x, var.y, itera=10){
  cvx2 <- sum(var.x/(((x+y)/2)^2))
  cvy2 <- sum(var.y/(((x+y)/2)^2))
  lambda <- cvx2/cvy2
  w <- 1
  estima <- mq.fcn(x, y)
  a0 <- estima$a
  b <- estima$b
  for (i in 1:itera){
    xw.bar <- sum(w*x)/sum(w)
    yw.bar <- sum(w*y)/sum(w)
    uw <- sum(w*(x-xw.bar)^2)
    qw <- sum(w*(y-yw.bar)^2)
    pw <- sum(w*(x-xw.bar)*(y-yw.bar))
    old.a <- a0
    old.b <- b
    b <- ((lambda*qw-uw)+sqrt((uw-lambda*qw)^2+4*lambda*pw^2))/(2*lambda*pw)
    a <- yw.bar
    a0 <- a-b*xw.bar
    d <- y-(a0+b*x)
    X <- x+lambda*b*d/(1+lambda*b^2)
    Y <- y-d/(1+lambda*b^2)
    cvx2 <- sum(var.x/(((X+Y)/2)^2))
  }
}

```

```

    cvy2 <- sum(var.y/(((X+Y)/2)^2))
    lambda <- cvx2/cvy2
    w <- 1/((X+Y)/2)^2
  }
  Elipse <- CalcEli.fcn(x, y, a, b, peso=1/w)

  dif.a <- abs(old.a - a0)
  dif.b <- abs(old.b - b)
  desvio.x <- sqrt(var.x)
  desvio.y <- sqrt(var.y)
  tam <- cbind(desvio.x, desvio.y)

  final <- list(a=a0, b=b, A=Elipse$A, B=Elipse$B, C=Elipse$C, D=Elipse$D,
    Dif.a=dif.a, Dif.b=dif.b,
    RMS=Elipse$RMS.y, S=Elipse$S, desvios=tam, Area=Elipse$Area,
    Acertos=Elipse$Acerta, Resid.esc=Elipse$Resid.esc, SomaR2=Elipse$SomaRes2)
  final
}

```

# Programa para cálculo dos erros padrão sa e sb usando o método de Linnet, usando  
 # o método "jackknife" proposto pelo autor em Statistics in Medicine (1990)

```

jack <- function(lista, x, y, var.x, var.y, it=10){
  a <- lista$a
  b <- lista$b
  k <- length(x)
  ai <- bi <- numeric(k)

  for (i in 1: k){
    dados <- linnet.fcn(x[-i], y[-i], var.x[-i], var.y[-i], it)
    bi[i] <- k*b-(k-1)*dados$b
    ai[i] <- k*a-(k-1)*dados$a
  }
  b.til <- sum(bi)/k
  a.til <- sum(ai)/k
  Vb <- sum((bi-b.til)^2)/(k-1)
  Va <- sum((ai-a.til)^2)/(k-1)
  sb <- sqrt(Vb/k)
  sa <- sqrt(Va/k)
  saida <- list(a=a, b=b, sa=sa, sb=sb, A=lista$A, B=lista$B, C=lista$C, D=lista$D, RMS=lista$RMS,
    Dif.a=lista$Dif.a, Dif.b=lista$Dif.b)
  saida
}

```

# Programa para gerar a função MLFR, que aplica o método de Ripley e Thompson (1987)

```

mlfr.fcn <- function(x, y, dpx, dpy, passo=10, itera=100){

mq <- lsfit(x, y)
a.mq <- mq$coef[1]
b.mq <- mq$coef[2]
b.ini <- b.mq
b0 <- Qm <- numeric(2)

```

```

Qm.min <- 0

for (i in 1:itera){
  b0[1] <- b.ini-passo
  b0[2] <- b.ini+passo
  Qm[1] <- sum((1/(dpy^2 + (b0[1]^2)*dpx^2))*(y-(sum((1/(dpy^2 + (b0[1]^2)*dpx^2))*(y-
  b0[1]*x))/sum((1/(dpy^2 + (b0[1]^2)*dpx^2))))-b0[1]*x)^2)
  Qm[2] <- sum((1/(dpy^2 + (b0[2]^2)*dpx^2))*(y-(sum((1/(dpy^2 + (b0[2]^2)*dpx^2))*(y-
  b0[2]*x))/sum((1/(dpy^2 + (b0[2]^2)*dpx^2))))-b0[2]*x)^2)
  Qm.min <- min(Qm)
  b.ini <- b0[Qm == Qm.min]
  passo <- passo/2
}
b <- min(b.ini)
w <- (1/(dpy^2 + (b^2)*dpx^2))
a <- (sum(w*(y-b*x))/sum(w))
u <- w*(x*dpy^2+b*(y-a)*dpx^2)
Elipse <- CalcEli.fcn(x, y, a, b, peso=1/w)
x.til <- sum(w*x)/sum(w)
y.til <- sum(w*y)/sum(w)
sa <- sqrt(sum(w*x^2)/(sum(w)*sum(w*(x-x.til)^2)))
sb <- 1/sqrt(sum(w*(x-x.til)^2))

final <- list(a=a, b=b, A=Elipse$A, B=Elipse$B, C=Elipse$C, D=Elipse$D, n=length(x),
  RMS=Elipse$RMS.y, S=Elipse$S, SomaR2=Elipse$SomaRes2, sa=sa, sb=sb, w=w,
  Resid.esc=Elipse$Resid.esc, u=u, Area=Elipse$Area, Acertos=Elipse$Acerta)

final
}

# Função para gerar a reta do método LTS de Rousseew e Leroy

lts.fcn <- function(x, y){
  reg <- ltsreg(x, y)
  a <- reg$coef[1]
  b <- reg$coef[2]
  rms <- sqrt(sum(((y-a-b*x)^2)/(length(x)-2)))
  S <- sum((y-a-b*x)^2)
  saida <- list(a=a, b=b, RMS=rms, S=S)

saida
}

# Função para cálculo do coeficiente angular em regressão linear ortogonal
# Aproximação de Wald
# Ref.: Danzer, K.; Wagner, M.; Fischbacher, C. (1995) - Calibration by orthogonal and common
# least squares - theoretical and practical aspects. Fresenius' Journal of Analytical
# Chemistry, Vol. 352, pp. 407-412.

wald.fcn <- function(x, y){
  n <- length(x)
  if (floor(n/2) == n/2) {
    g <- n/2
    h <- g+1
  }
  else g <- h <- (n+1)/2

```

```

ordem <- order(x)
y.ord <- y[ordem]
x.ord <- x[ordem]
yi <- y.ord[1:g]
xi <- x.ord[1:g]
yj <- y.ord[h:n]
xj <- x.ord[h:n]
b <- (sum(yi)-sum(yj))/(sum(xi)-sum(xj))
a <- mean(y) - b*mean(x)
S <- sum((y-a-b*x)^2)
rms <- sqrt(sum(((y-a-b*x)^2)/(length(x)-2)))
saida <- list(a=a, b=b, RMS=rms, S=S)
saida
}

# Função para uso do método de Williamson, correção sobre o de York
# Ref.: Kalantar, A.H.; Gelb, R.L.; Alper, J.S. (1995) - Biases in summary statistics of
#       slopes and intercepts in linear regression with errors in both variables.
#       Talanta, Vol. 42, No. 4, pp. 597-603.

yw.fcn <- function(x, y, var.x, var.y, itera=10){

n <- length(x)
mq <- lsfit(x, y)
b <- mq$coef[2]
a <- mq$coef[1]
old.a <- 0
old.b <- 0
w <- numeric(n)
for (i in 1:itera) {
  w <- 1/(var.y + (b^2)*var.x)
  xw.bar <- sum(w*x)/sum(w)
  yw.bar <- sum(w*y)/sum(w)
  X <- x-xw.bar
  Y <- y-yw.bar
  old.a <- a
  old.b <- b
  b <-
  (sum(w^2*var.y*X*Y)+b*sum(w^2*var.x*Y^2))/(sum(w^2*var.y*X^2)+b*sum(w^2*var.x*X*Y))
  a <- yw.bar - b*xw.bar
}
  Elipse <- CalcEli.fcn(x, y, a, b, peso=1/w)
dpx <- sqrt(var.x)
dpy <- sqrt(var.y)
dif.a <- abs(old.a - a)
dif.b <- abs(old.b - b)
tam <- cbind(dpx, dpy)
final <- list(a=a, b=b, Dif.a=dif.a, Dif.b=dif.b, RMS=Elipse$RMS.y, S=Elipse$S, desvios=tam,
  Resid.esc=Elipse$Resid.esc, SomaR2=Elipse$SomaRes2)
final
}

# Função para o método da Regressão Estrutural Absoluta, de Feldmann

```

```
# Ref.: Feldmann, U. (1992) - Robust bivariate errors-in-variables regression and outlier
# detection. European Journal of Clinical Chemistry and Clinical Biochemistry, Vol. 30
# pp. 405-414.

rea.fcn <- function(x,y){
  aux <- median((y-median(y))/(x-median(x)))
  sinal <- aux/abs(aux)
  b <- sinal*median(abs(y-median(y))/abs(x-median(x)))
  a <- median(y-b*x)
  S <- sum((y-a-b*x)^2)
  rms <- sqrt(sum(((y-a-b*x)^2)/(length(x)-2)))
  saida <- list(a=a, b=b, RMS=rms, S=S)
saida
}

#Modelagem da precisão analítica

# Função para encontrar equações de precisão, dados x e sx

modela.fcn <- function(x, sx, tipo="I", itera=4, lab="Teor (%)", tit=" ", Graf=T, Overlay=F,
  tipocor=1){

# Calcula a curva, com as bandas de confiança, e faz o gráfico

nx <- length(x)
x0 <- seq(from=min(x), to=max(x), length=200)

if (tipo == "I") {

  regx <- lm(sx ~ x)
  corr.x <- cor(sx, x)
  old.b0x <- regx$coefficients[1]
  old.b1x <- regx$coefficients[2]
  predx <- predict(regx)
  wx <- 1/predx^2
  for (i in 1:itera){
    medx <- sum(wx*x)/sum(wx)
    b0x<-(sum(wx*sx)*sum(wx*x^2)-sum(wx*x)*sum(wx*x*sx))/(sum(wx)*sum(wx*x^2)-
      (sum(wx*x))^2)
    b1x <- (sum(wx)*sum(wx*x*sx)-sum(wx*x)*sum(wx*sx))/(sum(wx)*sum(wx*x^2)-
      (sum(wx*x))^2)
    pred.novx <- b0x + b1x*x
    wx <- 1/pred.novx^2
    old.b0x <- b0x
    old.b1x <- b1x
  }
  dif.b0x <- abs(b0x - old.b0x)
  dif.b1x <- abs(b1x - old.b1x)
  linha.x <- b0x + b1x*x0
  rms.x <- sqrt((sum(wx*sx^2) - b0x*sum(wx*sx) - b1x*sum(wx*x*sx))/(nx - 2))
  erro.x <- (2*qt(0.975, 2, nx-2))^0.5*rms.x*(1/sum(wx)+(x0 - medx)^2/sum(wx*(x-medx)^2))^0.5
  sup.x <- linha.x + erro.x
  inf.x <- linha.x - erro.x
```

```
    } # término do 'if'

else if (tipo == "p") {
  zx <- log(x)
  z0 <- log(x0)
  zx.med <- sum(log(x))/nx
  regx <- lm(log(sx) ~ zx)
  corr.x <- cor(log(sx), zx)
  b0x <- exp(regx$coefficients[1])
  b1x <- regx$coefficients[2]
  linha.x <- b0x * x0^b1x
  rms.x <- sqrt((sum(log(sx)^2) - log(b0x)*sum(log(sx)) - b1x*sum(zx*log(sx)))/(nx - 2))
  erro.x <- (2*qt(0.975, 2, nx-2))^0.5*rms.x*(1/nx+(z0 - zx.med)^2/sum((zx-zx.med)^2))^0.5
  sup.x <- exp(log(linha.x) + erro.x)
  inf.x <- exp(log(linha.x) - erro.x)
  wx <- rep(1, nx)

  } # término do 'if' dentro do 'else'

else {

  regx <- lm(log(sx) ~ x)
  corr.x <- cor(log(sx), x)
  b0x <- regx$coefficients[1]
  b1x <- regx$coefficients[2]
  linha.x <- exp(b0x + b1x*x0)
  rms.x <- sqrt((sum(log(sx)^2) - b0x*sum(log(sx)) - b1x*sum(x*log(sx)))/(nx - 2))
  erro.x <- (2*qt(0.975, 2, nx-2))^0.5*rms.x*(1/nx+(x0 - mean(x))^2/sum((x-mean(x))^2))^0.5
  sup.x <- exp(log(linha.x) + erro.x)
  inf.x <- exp(log(linha.x) - erro.x)
  wx <- rep(1, nx)

  } # fim do 'else'

# Gráfico

if (Graf){
  if (!Overlay){
    win.graph()
    esc.x <- range(x)
    esc.y <- range(sx, sup.x, inf.x, 0)
    plot(x, sx, xlab=lab, ylab="Desvio Padrão", pch=1, xlim=esc.x, ylim=esc.y, main=tit)
    }
  else { points(x, sx, pch=tipocor, col=tipocor) }
  lines(x0, linha.x, lty=4, col=tipocor)
  lines(x0, sup.x, lty=1, col=tipocor)
  lines(x0, inf.x, lty=1, col=tipocor)
  }

# Saídas
```

```
#saida <- list(Dif.Metodo.1=c(dif.b0x, dif.b1x))
saida <- list(b0=b0x, b1=b1x, correlacao=corr.x, rms=rms.x)#, w=wx)#,
#           x0=x0, linha=linha.x, sup=sup.x, inf=inf.x)
saida
} # término do comando 'function'

# Função para gerar números aleatórios que simulem duas sequências de dados a serem comparados
# Ex.: Duas séries de resultados de análises provenientes de dois laboratórios diferentes nas
#       mesmas amostras.
#
# tipo é uma variável que pode assumir os valores "l" (=linear), "p" (=potência) ou "c" (=composta
# = função dos dois tipos de erro). É o tipo de relação matemática entre a precisão e teor.
#
# distrib é a distribuição dos dados do método de referência (u=uniforme, n=normal,
# ln=log-normal, g=gamma)
#
# out.perc é a porcentagem de pontos que se deseja serem outliers - inclusão de outliers
# não finalizada
#

simula.fcn <- function(n=1000, min=0, max=100, a=0, b=1, alfa.x, beta.x,
                      alfa.y, beta.y, tipo.x="l", tipo.y="l", distrib="u", out.perc=0,
                      out.min=3, out.max=10){

# Geração dos valores para o vetor X

if (min == 0) minln <- 1+abs(max/10000)
else minln <- min
if (distrib=="u") X <- runif(n, min, max)
else if (distrib=="n") X <- abs(rnorm(n, mean=mean(c(min, max)), sd=(((max-min)/6))))
else if (distrib=="ln") X <- exp(rnorm(n, mean=mean(c(log(minln), log(max))),
                                     sd=((log(max)-log(minln))/6))))

else stop("Distribuição Inválida")

# Geração dos valores para o vetor Y

Y <- a + b*X

# Geração dos erros para cada ponto de X e de Y

if (tipo.x == "l") erro.x <- alfa.x + beta.x*X
else if (tipo.x == "p") erro.x <- alfa.x*X^beta.x
else if (tipo.x == "c") erro.x <- sqrt(alfa.x^2 + X^2*exp(beta.x^2)*(exp(beta.x^2)-1))
                                     # alfa= sigma.epsilon, e beta=sigma.eta
else stop("Valor para o argumento 'tipo' inválido. Deve ser 'l', 'p' ou 'c' ")

if (tipo.y == "l") erro.y <- alfa.y + beta.y*Y
else if (tipo.y == "p") erro.y <- alfa.y*Y^beta.y
else if (tipo.y == "c") erro.y <- sqrt(alfa.y^2 + Y^2*exp(beta.y^2)*(exp(beta.y^2)-1))
else stop("Valor para o argumento 'tipo' inválido. Deve ser 'l', 'p' ou 'c' ")

# Incorporação do erro aleatório aos vetores X e Y, gerando x e y
```



```
x <- y <- rep(-1, n)
for (i in 1:n) {
  x[i] <- rnorm(1, mean=X[i], sd=erro.x[i])
  y[i] <- rnorm(1, mean=Y[i], sd=erro.y[i])
}

resultado <- list(X=X, Y=Y, x=x, y=y, ErroX=erro.x, ErroY=erro.y)
resultado
}
```

## Capítulo 7

### Referências bibliográficas

- Agterdenbos, J. (1979) – Calibration in quantitative analysis – Part 1. General considerations. *Analytica Chimica Acta*, 108, 315-323.
- Albert, R. & Horwitz, W. (1997) – A heuristic derivation of the Horwitz Curve. *Analytical Chemistry*, 69, 789-790.
- Altman, D. G. & Bland, J. M. (1983) – Measurement in medicine: the analysis of method comparison studies. *The Statistician*, 32, 307-317.
- Altman, D. G. & Bland, J. M. (1997) – Difference versus mean plots. *Annals of Clinical Biochemistry*, 34, 570-571.
- Bablok, W.; Passing, H.; Poppe, W. (1990) – Statistical procedures for method comparison studies. *GIT Labor-Medizin*, Vol. 12, 732-736.
- Barnett, R. N. (1965) – A scheme for the comparison of quantitative methods. *The American Journal of Clinical Pathology*, Vol. 43, No. 6, 562-569.
- Barnett, R. N.; Youden, W. J. (1970) – A revised scheme for the comparison of quantitative methods. *The American Journal of Clinical Pathology*, Vol. 54, 454-462.
- Bartlett, M. S. (1949) – Fitting a straight line when both variables are subject to error. *Biometrics*, Vol. 5, 207-212.
- Bolfarine, H.; Rodrigues, J. e Cordani, L. (1992) – *O modelo de regressão com erros nas variáveis*. 10º Simpósio Nacional de Probabilidade e Estatística.
- Boyer, K.W.; Horwitz, W. e Albert, R. (1985) – Interlaboratory variability in trace element analysis. *Analytical Chemistry*, Vol.57, N° 2, 454-459.
- Bruce, P.; Minkinen, P. and Riekkola, M.-L. (1998) – Practical method validation: validation sufficient for an analysis method. *Mikrochimica Acta*, 128, 93-106.
- Carey, R. N.; Wold, S. and Westgard, J. O. (1975) – Principal component analysis: an alternative to “referee” methods in method comparison studies. *Analytical Chemistry*, Vol. 47, No. 11, 1824-1829.

Carré, M.; Excoffier, S. and Mermet, J. M. (1997) – A study of the relation between the limit of detection and the limit of quantitation in inductively coupled plasma spectrochemistry. *Spectrochimica Acta Part B*, 52, 2043-2049.

Carroll, R. J. and Spiegelman, C. H. (1992) – Diagnostics for nonlinearity and heteroscedasticity in errors-in-variables regression. *Technometrics*, Vol. 34, No. 2, 186-196.

Caulcutt, R. e Boddy, R. (1983) – *Statistics for Analytical Chemists*. London – New York: Chapman and Hall.

Cecchi, G. C. (1991) – Error analysis of the parameters of a least-squares determined curve when both variables have uncertainties. *Measurement Science and Technology*, 2, 1127-1128.

Centner, V.; Massart, D.L. e De Jong, S. (1998) – Inverse calibration predicts better than classical calibration. *Fresenius Journal of Analytical Chemistry*, Vol. 361, 2-9.

Christensen, J. M. (1996) – Correcting measurement errors using reference materials in method validation. *Mikrochimica Acta*, 123, 231-240.

Cornbleet, P. J. & Gochman, N. (1979) – Incorrect least-squares regression coefficients in method comparison analysis. *Clinical Chemistry*, 25/3, 432-438.

Currie, L. A. and Horwitz, W. (1994) – IUPAC recommendations for defining and measuring detection and quantification limits. *Analisis Magazine*, v 22, no. 5, 24-26.

Danzer, K.; Wagner, M.; Fischbacher, C. (1995) – Calibration by orthogonal and common least squares – Theoretical and practical aspects. *Fresenius Journal of Analytical Chemistry*, 352: 407-412.

De Castilho, M. V. (2004) – A comparison of statistical techniques for detecting analytical bias in geoanalysis. *Geostandards and Geoanalytical Research*. Vol.28, No. 2, 277-290.

De Castro, M.; Galea-Rojas, M.; Bolfarine, H. and De Castilho, M. V. (2004)- Detection of analytical bias when comparing two or more measuring methods. *Journal of Chemometrics*, Vol. 18, Issue 10, pp 431-440.

De Castro, M., Bolfarine, H. and De Castilho, M. V. (2005) – Consistent estimation and testing in comparing analytical bias models. *Environmetrics*, vol. 16, 1-16.

De Castro, M.; Bolfarine, H.; Galea-Rojas, M. and De Castilho, M. V. (2005) - An exact test for analytical bias detection. *Analytica Chimica Acta*, Vol 538, 375-381.

De Castro, M.; Galea-Rojas, M.; Bolfarine, H. and De Castilho, M. V. (2006)- Local influence in regression models for the detection of analytical bias. *Chemometrics and Intelligent Laboratory Systems*, Vol. 83, pp 139-147.

Dellaportas, P.; Stephens, D. A. (1995) – Bayesian analysis of errors-in-variables regression models. *Biometrics*, Vol. 51, 1085-1095.

Deming, W. E. (1943) – *Statistical adjustment of data*. New York: John Wiley & Sons, Inc.

Dixon, S. L. and McKean, J. W. (1996) – Rank-based analysis of the heteroscedastic linear model. *Journal of the American Statistical Association*, Vol. 91, No. 434, 699-712.

Draper, N. R. & Smith, H. (1981) - *Applied regression analysis*, 2nd. ed. New York: John Wiley & Sons, Inc.

Edgeworth, F.Y. (1887) – On observations relating to several quantities. *Hermathena*, 6, 279-285.

Ellison, S. L. R. (1998) – ISO uncertainty and collaborative trial data. *Accreditation and Quality Assurance*, 3, 95-100.

Feinberg, M. (1995) - Basics of interlaboratory studies: the trends in the new ISO 5725 standard edition. *Trends in analytical chemistry*, vol. 14, No. 9, 450-457.

Feldmann, U. (1992) – Robust bivariate errors-in-variables regression and outlier detection. *European Journal of Clinical Chemistry and Clinical Biochemistry*, Vol. 30, 405-414.

Fung, W.K.; Tsang, T.S. (1998) – A simulation study comparing tests for the equality of coefficients of variation. *Statistics in Medicine*, 17, 2003-2014

Galea-Rojas, M.; De Castilho, M. V.; Bolfarine, H. and De Castro, M. (2003) – Detection of analytical bias. *The Analyst*, Vol. 128, 1073-1081.

Garden, J.S.; Mitchell, D.G. e Mills, W.N. (1980) – Nonconstant variance regression techniques for calibration-curve-based analysis. *Analytical Chemistry*, Vol. 52, N<sup>o</sup> 14, 2310-2315.

Garthwaite, P.H.; Jolliffe, I.T. e Jones, B. (1995) – *Statistical Inference*. Prentice Hall International (UK) Limited.

Gibbons, R.; Coleman, D. E. and Maddalone, R. F. (1997) – An alternative minimum level definition for analytical quantification. *Environmental Science & Technology*, Vol. 31, No. 7, 2071-2077.

Hall, P. & Selinger, B. (1989) - A statistical justification relating interlaboratory coefficients of variation with concentration levels. *Analytical Chemistry*, 61, 1465-1466.

Hampel, F.R. (1971) – A general qualitative definition of robustness. *Annals of Mathematical Statistics*, 42, 1887-1896.

Hartmann, C.; Smeyers-Verbeke, J.; Massart, D.L. (1993) – Problems in method-comparison studies. *Analisis*, 21, 125-132.

Hartmann, C.; Smeyers-Verbeke, J.; Penninck, W. & Massart, D.L. (1997) - Detection of bias in method comparison by regression analysis. *Analytica Chimica Acta*, 338, 19-40.

Hartmann, C.; Vankeerberghen, J.; Smeyers-Verbeke, J. & Massart, D. L. (1997) – Robust Orthogonal regression for the outlier detection when comparing two series of measurement results. *Analytica Chimica Acta*, 344, 17-28.

Hayashi, Y; Matsuda, R and Poe, R. B. (1996) – Probabilistic approach to confidence intervals of linear calibration. *Analyst*, Vol. 121, 591-599.

Hollis, S. (1996) – Analysis of method comparison studies. *Annals of Clinical Biochemistry*, 33, 1-4.

Horwitz, W.; Kamps, L. R. and Boyer, K. W. (1980) – Quality assurance in the analysis of food for trace constituents. *Journal of the Association of Official Analytical Chemists*, Vol. 63, No. 6, 1344-1354.

Horwitz, W. (1982) - Evaluation of analytical methods used for regulation of foods and drugs. *Analytical Chemistry*, Vol. 54, No.1, 67A-76A.

Horwitz, W. (1988) – Protocol for the design, conduct and interpretation of collaborative studies. *Pure & Applied Chemistry*, Vol. 60, No. 6, 855-864.

Horwitz, W. e Albert, R. (1987) – Standardisation of analytical methods. *Analytical Proceedings*, Vol. 24, 49-55.

Horwitz, W. e Albert, R. (1991) – Special reports – Performance characteristics of methods of analysis used for regulatory purposes. Part II. Pesticide formulations. *Journal of AOAC International*, Vol. 74, N° 5, 718-744.

Horwitz, W.; Albert, R.; Deutsch, M.J. e Thompson, J.N. (1992) – Precision parameters of methods of analysis required for nutrition labeling. Part II. Macro elements – calcium, magnesium, phosphorus, potassium, sodium and sulfur. *Journal of AOAC International*, Vol. 75, N° 2, 227-239.

Horwitz, W. e Albert, R. (1995) – Precision in analytical measurements: expected values and consequences in geochemical analyses. *Fresenius' Journal of Analytical Chemistry*. Vol. 351, 507-513.

Horwitz, W.; Britton, P. e Chirtel, S.J. (1998) – A simple method for evaluating data from an interlaboratory study. *Journal of AOAC International*, Vol. 81, N° 6, 1257-1265.

Huber, L. (1998) – Validation of analytical methods: review and strategy. *LC.GC International*. February edition, pp. 96-105.

Huber, P. J. (1973) – Robust regression: asymptotics, conjectures and Monte Carlo. *Annals of Statistics*, 1, 799-821.

International Organization for Standardization, Geneva. ISO 5725-1; Accuracy (trueness and precision) of measurement methods and results; Part 1: General Principles and Definitions. Geneva, 1994.

International Organization for Standardization, Geneva. ISO 5725-2; Accuracy (trueness and precision) of measurement methods and results; Part 2: A basic method for the determination of repeatability and reproducibility of a standard reference method. Geneva, 1994.

International Organization for Standardization, Geneva. 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. Geneva, 1994.

International Organization for Standardization, Geneva. 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. Geneva, 1998.

International Organization for Standardization, Geneva. ISO 11843-1; Capability of detection; Part 1: Terms and definitions. Geneva, 1997.

International Organization for Standardization, Geneva. ISO/DIS 11843-2; Capability of detection; Part 2: Methodology in the linear calibration case. Geneva, 1998.

International Organization for Standardization, Geneva. ISO Guide 30; Terms and definitions used in connection with reference materials. Geneva, 1992.

International Organization for Standardization, Geneva. ISO Guide 31; Contents of certificates of reference materials. Geneva, 1981.

International Organization for Standardization, Geneva. ISO Guide 32; Calibration in analytical chemistry and use of certified reference materials. Geneva, 1997.

International Organization for Standardization, Geneva. ISO Guide 33; Uses of certified reference materials. Geneva, 1989.

International Organization for Standardization, Geneva. ISO Guide 34; Quality system guidelines for the production of reference materials. Geneva, 1996.

International Organization for Standardization, Geneva. ISO Guide 35; Certification of reference materials – general and statistical principles. Geneva, 1989.

Johnson, W. (1998) – The QA/QC Corner #4. *Alchemist's Digest*, Edition 98-1, pp. 2-5.

Jülicher, B.; Gowik, P. and Uhlig, S. (1998) – Assessment of detection methods in trace analysis by means of a statistically based in-house validation concept. *Analyst*, Vol. 123, 173-179.

Kalantar, A. H.; Gelb, R. I.; Alper, J. S. (1995) – Biases in summary statistics of slopes and intercepts in linear regression with errors in both variables. *Talanta*, Vol. 42, No. 4, 597-603

Kane, J. S. (1997) – Analytical bias: the neglected component of measurement uncertainty. *Analyst*, Vol. 122, 1283-1288.

Kulasekera, K.B. (1995) – Comparison of regression curves using quase-residuals. *Journal of the American Statistical Association*, Vol. 90, No. 431, 1085-1093.

Linnet, K. (1990) – Estimation of the linear relationship between the measurements of two methods with proportional errors. *Statistics in Medicine*, Vol. 9, 1463-1473.

Linnet, K. (1993) – Evaluation of regression procedures for methods comparison studies. *Clinical Chemistry*, Vol. 39, N° 3, 424-432.

Linnet, K.; (1998) – Performance of Deming regression analysis in case of misspecified analytical error ratio in method comparison studies. *Clinical Chemistry*, 44:5, 1024-1031.

Linnet, K. (1999) – Necessary sample size for method comparison studies based on regression analysis. *Clinical Chemistry*, Vol. 45, N° 6, 882-894.

Lisy, J. M.; Cholvadova, A.; Kutej, J. (1990) – *Computer Chemistry*, 14, 189-192.

Long, G. L. & Winefordner, J. D. (1983) – Limit of detection. A closer look at the IUPAC definition. *Analytical Chemistry*, vol. 55, no. 7, 712 A – 724 A.

MacTaggart, D. L.; Farwell, S. O. (1992) – Analytical use of linear regression. Part II: statistical error in both variables. *Journal of AOAC International*, Vol. 75, No. 4, 608-614.

Mandel, J. e Linnig, F. (1957) – Study of accuracy in chemical analysis using linear calibration curves. *Analytical Chemistry*, Vol 29, N° 5, 743-749.

Mandel, J. (1964) - *The statistical analysis of experimental data*. New York: Dover Publications, Inc.

Mandel, J. (1984) – Fitting straight lines when both variables are subject to error. *Journal of Quality Technology*, 16, 1-13.

Mandel, J. (1996) – Interlaboratory testing and the rejection of observations. *ISO/REMCO N 184*.

Margosis, M.; Horwitz, W. e Albert, R. (1988) – Method performance – Performance characteristics of methods of analysis used for regulatory purposes. I. Drug dosage forms. F. Gravimetric and titrimetric methods. *Journal of AOAC International*, Vol. 71, N° 3, 619-635.

Marschal, A. (1997) – Problematic of the selection and use of certified reference materials for calibration of chemical analyses – general guidance. *Accreditation and Quality Assurance*, 2, 259-261.

Massart, D.L.; Kaufman, L.; Rousseeuw, P.J. e Leroy, A. (1986) – Least median of squares: a robust method for outlier and model error detection in regression and calibration. *Analytica Chimica Acta*, Vol. 187, 171-179.

Massart, D. L.; Vandeginste, B.G.M.; Deming, S.N.; Michotte, Y. e Kaufman, L.(1988) – *Chemometrics: a textbook*. Elsevier Science Publishing Company, Inc.

Minkinen, P. (1995) – Estimation of variance components from the results of interlaboratory comparisons. *Chemometrics and Intelligent Laboratory Systems*, 29, 263-270.

Montgomery, D. C. (1991) – *Design and analysis of experiments*. 3<sup>rd</sup>. Ed. New York: John Wiley & Sons.

Montgomery, D. C. & Peck, E. A. (1992) – *Introduction to linear regression analysis*, 2<sup>nd</sup>. Edition. New York: John Wiley & Sons, Inc.

Nalimov, V. V. (1963) – *The application of mathematical statistics to chemical analysis*. Addison-Wesley Publishing Company, Inc.

Neri, F.; Saitta, G.; Chiofalo, S. (1989) – An accurate and straightforward approach to linear regression analysis of error-affected experimental data. *Journal of Physics Education: Scientific Instrumentation*, Vol. 22, 215.

Neri, F.; Patanè, S.; Saitta, G. (1990) – Error-affected experimental data analysis: application to fitting procedures. *Measurement Science and Technology*, Vol.1, 1007.

Ortiz, M. C.; Arcos, J.; Juarros, J. V.; López-Palacios, J. and Sarabia, L. A. (1993) – Robust procedure for calibration and calculation of the detection limit of trimipramine by adsorptive stripping voltammetry at a carbon paste electrode. *Analytical Chemistry*, 65, 678-682.

Paiva Lima, C. R. O. (1996) – *Calibração absoluta com erros nas variáveis*. São Paulo: Instituto de Matemática e Estatística da Universidade de São Paulo. (Tese, Doutorado em Estatística).

Parvin, C.A. (1984) – A direct comparison of two slope-estimation techniques used in method comparison studies. *Clinical Chemistry*, Vol. 30, N<sup>o</sup> 5, 751-754.

Passing, H.; Bablok, W. (1983) – A new biometrical procedure for testing the equality of measurements from two different analytical methods – Application of linear regression procedures for method comparison studies in Clinical Chemistry, Part I. *Journal of Clinical Chemistry and Clinical Biochemistry*, Vol. 21, 709-720.

Patefield, W. M. (1977) – On the information matrix in the linear functional relationship problem. *Applied Statistics*, 26, No. 1, p. 67.

Pauwels, J.; Lambert, A. and Schimmel, H. (1998) – The determination of the uncertainty of reference materials certified by laboratory intercomparison. *Accreditation and Quality Assurance*, 3, 180-184.

Payne, R. B. (1997) – Method comparison: evaluation of least squares, Deming and Passing/Bablok regression procedures using computer simulation. *Annals of Clinical Biochemistry*, 34, 319-320.

Peeler, J. T.; Horwitz, W. e Albert, R. (1989) – Method performance – Precision parameters of standard methods of analysis for dairy products. *Journal of AOAC International*, Vol. 72, N° 5, 784-806.

Penninckx, W. (1994) – Detection of the bias in analytical methods. *Analisis Magazine*, vol. 22, n° 5.

Penninckx, W.; Hartmann, C.; Massart, D. L.; Smeyers-Verbeke, J. (1996) – Validation of the calibration procedure in atomic absorption spectrometric methods. *Journal of Analytical Atomic Spectrometry*, Vol. 11 , 237-246.

Petersen, P. H.; Stöckl, D.; Blaabjerg, O.; Pedersen, B.; Birkemose, E.; Thienpont, L.; Lassen, J. F. and Kjeldsen, J. (1997) – Graphical interpretation of analytical data from comparison of a field method with a reference method by use of difference plots. *Clinical Chemistry*, 43:11, 2039-2046.

Richardson, J. M. & Morrison, J. L. (1995) – Method validation in the Ontario Geoscience Laboratories. *Analyst*, Vol 120, 1519-1526.

Rifkin, R. D. (1995) – Effects of correlated and uncorrelated measurement error on linear regression and correlation in medical method comparison studies. *Statistics in Medicine*. Vol. 14, 789-798.

Ripley, B.D. e Thompson, M. (1987) – Regression techniques for the detection of analytical bias. *Analyst*. Vol. 112, 377-383.

Riu, J. & Rius, F. X. (1995) – Univariate regression models with errors in both axes. *Journal of Chemometrics*, Vol. 9, 343-362.

Riu, J. & Rius, F. X. (1996) - Assessing the accuracy of analytical methods using linear regression with errors in both axes. *Analytical Chemistry*, 68, 1851-1857.

Rocke, D. M. & Lorenzato, S. (1995) – A two-component model for measurement error in analytical chemistry. *Technometrics*, Vol. 37, No. 2, 176-184.

Rousseeuw, P. J. (1984) – Least median of squares regression. *Journal of the American Statistical Association*, Vol. 79, No. 388, 871-880.

Rousseeuw, P. J.; Leroy, A. M. (1987) – *Robust regression and outlier detection*. John Wiley & Sons, Inc.

Rutan, S.C. e Carr, P.W. (1988) – Comparison of robust regression methods based on least-median and adaptive kalman filtering approaches applied to linear calibration data. *Analytica Chimica Acta*, Vol. 215, 131-142.

Sarabia, L. A.; Ortiz, M. C.; Tomás, X. (1997) – Performance of the orthogonal least median squares regression. *Analytica Chimica Acta*, 348, 11-18.

Schafer, D. W. (1987) – Covariate measurement error in generalized linear models. *Biometrika*, 74, 2, 385-91.

Schwartz, L.M. (1979) – Calibration curves with nonuniform variance. *Analytical Chemistry*, Vol. 51, N<sup>o</sup> 6, 723-727.

Seber, G.A.F. (1977) – *Linear Regression Analysis*. John Wiley & Sons.

Siegel, A. F. (1982) – Robust regression using repeated medians. *Biometrika*, 69, 242-244.

Smagunova, A. N. (1997) – Methods for estimating the accuracy of analysis results. *Journal of Analytical Chemistry*, Vol. 52, No. 10, 926-933.

Snedecor, G.W. e Cochran, W.G. (1996) – *Statistical Methods*. Eighth edition, seventh printing. Iowa State University Press, Ames.

Spurrer, J. D. (1992) – Optimal designs for comparing the variances of several treatments with that of a standard treatment. *Technometrics*, Vol. 34, No.3, 332-339.

Stöckl, D. (1996) – Beyond the myths of difference plots. *Annals of Clinical Biochemistry*, 33, 575-577.

Sutarno, R.; Steger, H. F. (1985) – The use of certified reference materials in the verification of analytical data and methods, pp 439-445. Pergamon Press Ltd.

Thompson, M. (1982) – Regression methods in the comparison of accuracy. *Analyst*, Vol. 107, 1169-1180.

Thompson, M. (1989) – Robust statistics and functional relationship estimation for comparing the bias of analytical procedures over extended concentration ranges. *Analytical Chemistry*, Vol. 61, 1942-1945.

Thompson, M. (1998) – Do we really need detection limits? *Analyst*, Vol. 123, 405-407.

Thompson, M. (1999) – A natural history of analytical methods. *Analyst*, Vol. 124, 991.

Thompson, M. & Lowthian, P. J. (1997) - The Horwitz function revisited. *Journal of AOAC International*, vol. 80, No. 3, 676-679.

Venables & Ripley (1994) – *Modern Applied Statistics with S-Plus* . New York: Springer- Verlag, Inc.

Vogelgesang, J. & Hädrich, J. (1998) – Limits of detection, identification and determination: a statistical approach for practitioners. *Accreditation and Quality Assurance*, 3, 242-255.

Wald, A. (1940) – The fitting of straight lines if both variables are subject to error. *Annals of Mathematical Statistics*, Vol. 11: 284-300.

Watters, Jr., R. L.; Carroll, R. J. and Spiegelman, C. H. (1987) – Error modeling and confidence interval estimation for inductively coupled plasma calibration curves. *Analytical Chemistry*, 59, 1639-1643.

Westgard, J. O. and Hunt, M. R. (1973) – Use and interpretation of common statistical tests in method-comparison studies. *Clinical Chemistry*, Vol. 19, No. 1, 49-57.

Whittemore, A. S.; Keller, J. B. (1988) – Approximations for regression with covariate measurement error. *Journal of the American Statistical Association*, Vol. 83, No. 404, 1057-1066.

Williamson, J. H. (1968) – Least squares fitting of a straight line – *Canadian Journal of Physics*, Vol. 46, p. 1845-1847.

York, D. (1966) - Least squares fitting of a straight line – *Canadian Journal of Physics*, Vol. 44, p. 1079-1086.

Zaalberg, J. (1989) – Experimental designs for interlaboratory precision experiments: comparison of ISO 5725 with draft NEN 6303. *Journal of the Association of Official Analytical Chemists*, Vol. 72, No. 1, 34-37.

Zeger, S. L.; Karim, M. R. (1991) – Generalized linear models with random effects; a Gibbs sampling approach. *Journal of the American Statistical Association*, Vol. 86, No. 413, 79-86.

Zwanziger, H. W. and Sârbu, C. (1998) – Validation of analytical methods using a regression procedure. *Analytical Chemistry*, 70, 1277-1280.