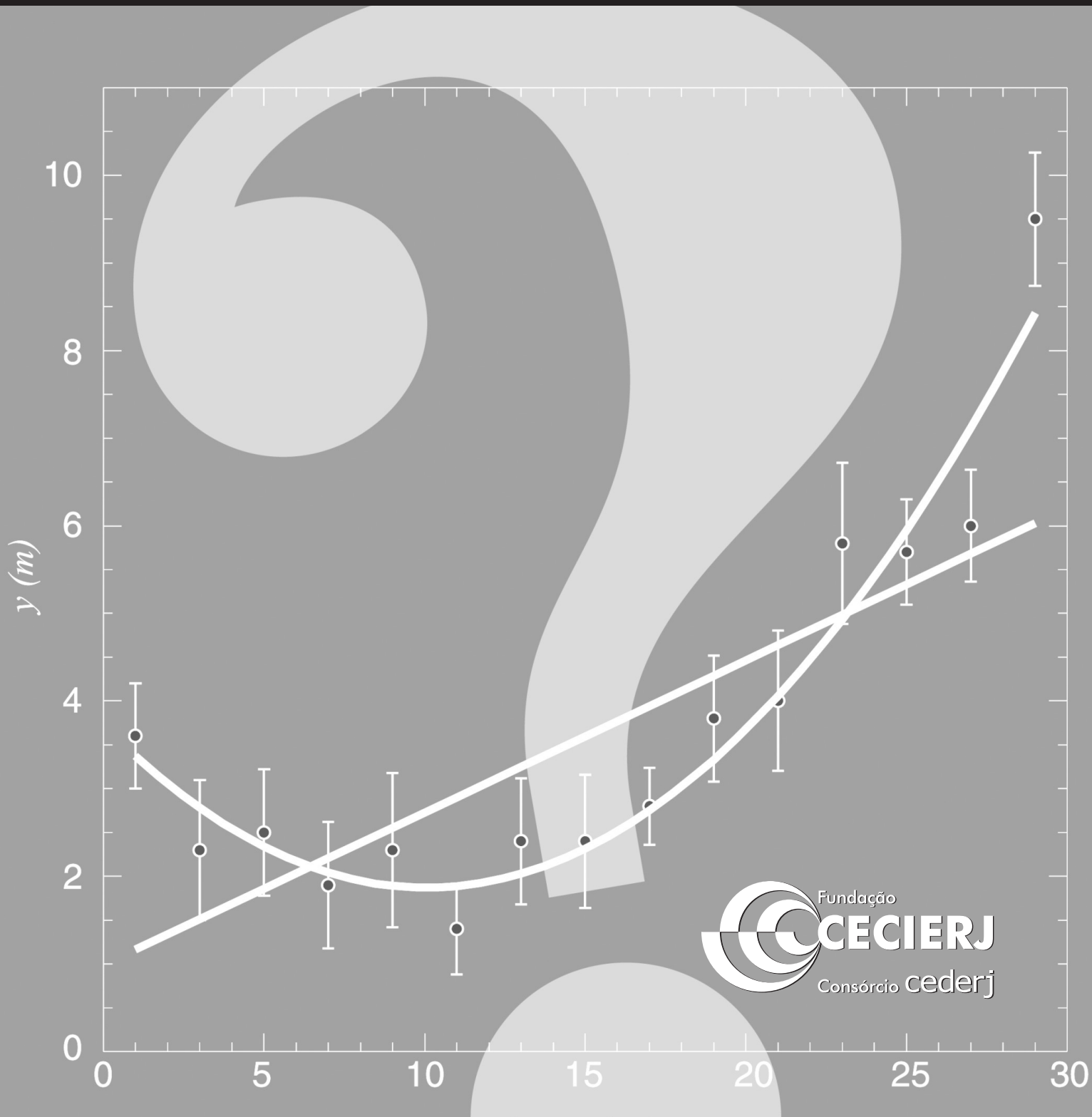


Tópicos de tratamento de dados experimentais





Fundação

CECIERJ

Consórcio **cederj**

Centro de Educação Superior a Distância do Estado do Rio de Janeiro

Tópicos de tratamento de dados experimentais

Volume 1

Bernard M. Marechal



**GOVERNO DO
Rio de Janeiro**

**SECRETARIA DE
CIÊNCIA E TECNOLOGIA**



**Ministério
da Educação**



Apoio:



FAPERJ

Fundação Carlos Chagas Filho de Amparo
à Pesquisa do Estado do Rio de Janeiro

Fundação Cecierj / Consórcio Cederj

Rua Visconde de Niterói, 1364 - Mangueira - Rio de Janeiro, RJ - CEP 20943-001

Tel.: (21) 2299-4565 Fax: (21) 2568-0725

Presidente

Carlos Eduardo Bielschowsky

Vice-Presidente de Educação Superior a Distância

Celso José da Costa

Diretor de Material Didático

Carlos Eduardo Bielschowsky

Coordenação do Curso de Física

Luiz Felipe Canto

Material Didático

ELABORAÇÃO DE CONTEÚDO

Bernard M. Marechal

EDITORA

Tereza Queiroz

COORDENAÇÃO EDITORIAL

Jane Castellani

COORDENAÇÃO DE DESENVOLVIMENTO INSTRUCIONAL

Cristine Costa Barreto

DESENVOLVIMENTO INSTRUCIONAL E REVISÃO

Ana Tereza de Andrade

Carmen Irene Correia de Oliveira

Leonardo Villela

José Meyohas

COORDENAÇÃO DE LINGUAGEM

Maria Angélica Alves

REVISÃO TIPOGRÁFICA

Patrícia Paula

COORDENAÇÃO GRÁFICA

Jorge Moura

PROGRAMAÇÃO VISUAL

Mirelle Mota

COORDENAÇÃO DE ILUSTRAÇÃO

Eduardo Bordoni

ILUSTRAÇÃO

Equipe CEDERJ

CAPA

Fabio Muniz

PRODUÇÃO GRÁFICA

Ana Paula Trece Pires

Andréa Dias Fiães

Copyright © 2005, Fundação Cecierj / Consórcio Cederj

Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada, por qualquer meio eletrônico, mecânico, por fotocópia e outros, sem a prévia autorização, por escrito, da Fundação.

M323t

Marechal, Bernard M.

Tópicos de tratamento de dados experimentais. v. 1 / Bernard M. Marechal. – Rio de Janeiro : Fundação CECIERJ, 2006.
53p.; 21 x 29,7 cm.

ISBN: 85-7648-026-3

1. Física. I. Título.

CDD: 530.1

2006/1

Referências Bibliográficas e catalogação na fonte, de acordo com as normas da ABNT.

Governo do Estado do Rio de Janeiro

Governadora
Rosinha Garotinho

Secretário de Estado de Ciência, Tecnologia e Inovação
Wanderley de Souza

Universidades Consorciadas

UENF - UNIVERSIDADE ESTADUAL DO
NORTE FLUMINENSE DARCY RIBEIRO
Reitor: Raimundo Braz Filho

UERJ - UNIVERSIDADE DO ESTADO DO
RIO DE JANEIRO
Reitor: Nival Nunes de Almeida

UFF - UNIVERSIDADE FEDERAL FLUMINENSE
Reitor: Cícero Mauro Fialho Rodrigues

UFRJ - UNIVERSIDADE FEDERAL DO
RIO DE JANEIRO
Reitor: Aloísio Teixeira

UFRRJ - UNIVERSIDADE FEDERAL RURAL
DO RIO DE JANEIRO
Reitor: Ricardo Motta Miranda

UNIRIO - UNIVERSIDADE FEDERAL DO ESTADO
DO RIO DE JANEIRO
Reitora: Malvina Tania Tuttman

Tópicos de tratamento de dados experimentais

Volume 1

Tópicos de tratamento de dados experimentais

Objetivos

As páginas que seguem, além de contribuírem para refrescar sua memória, fornecem as informações necessárias e, esperamos, suficientes para tratar de maneira correta e rigorosa os dados fornecidos, direta ou indiretamente, pelas experiências que você realizará durante sua **vida de experimentador**. O livro do José Henrique Vuolo, “*Fundamentos da Teoria de Erros*”, inspirou fortemente este texto complementar. Os exercícios do livro são importantes para melhor compreensão do conteúdo: não deixe de tentar resolvê-los. Se conseguir, ótimo. Se não, pode pedir socorro: faremos todo o possível para ajudá-lo!

Introdução

Uma experiência só está concluída após a apresentação **correta** e a análise (ou **tratamento**) dos dados. Nenhuma conclusão pode ser tirada, nenhum modelo pode ter suas predições verificadas sem uma discussão **rigorosa** dos **números** que a experiência forneceu.

A palavra **Tópicos** significa que este texto é somente um resumo de uma literatura extensa sobre a matéria e não um tratado exaustivo e auto-suficiente. Recomendamos, em particular, que o livro de José Henrique Vuolo, “*Fundamentos da Teoria de Erros*” [Vuolo(96)], esteja sempre ao seu alcance, tornando-se (e por que não?) seu “livro de cabeceira”. Espera-se que você esteja familiarizado com os princípios básicos e o vocabulário dessa área de conhecimento. Aconselhamos, fortemente, que você estude os capítulos 1 a 7 e o capítulo 9 deste livro, pois grande parte do conteúdo desses capítulos deve ser abordada, ou pelo menos mencionada, nos cursos de física.

Se você não conseguir acompanhar o que será visto agora, tem de dar um tempo e “começar pelo começo”, isto é, entender os conceitos e princípios básicos de estatística e probabilidade.

Vamos agora comentar uma experiência virtual para entender porque um resultado experimental não pode ser, simplesmente, por exemplo:

$$\alpha = 3,936$$

Exercício 1

Neste resultado, informações estão faltando! Você poderia dizer quais?

Experiência virtual “luso-britânica”

Um português, Joaquim, e um inglês, Charles, recebem duas amostras idênticas de uma cachaça com a tarefa delicada de determinar sua temperatura de ebulição. No exato momento do aparecimento das primeiras bolhas, nossos dois experimentadores anotam a temperatura fornecida por seu termômetro. Satisfeitos, eles confrontam seus resultados, mas começam a brigar porque cada um acha que o outro errou! O que aconteceu? Amostras idênticas, condições ambientais e procedimento experimental idênticos e ... resultados **diferentes**! Quem está certo?

Aí aparece o Zé-do-morro, brasileiro “malandro de praia há 40 anos”. Joaquim e Charles mostram seus respectivos resultados, $t_{joaquim} = 87.5$ e $t_{charles} = 190$, ao Zé-do-morro que, olhando para os dois “brigões”, começa a rir e convida todo mundo para beber essa cachaça (a original, e não as amostras fervidas!) pois, segundo ele, os dois **parecem** ter razão. Mas para poder afirmar que os dois **estão certos**, o nosso amigo esperto faz somente duas perguntas:

- Qual a unidade de temperatura utilizada?
- Qual a incerteza sobre sua medida?

Exercício 2

Você pode responder estas perguntas?

Ao ouvir as respostas, o malandro faz algumas contas, as que você deverá estar capaz de fazer ao final da leitura deste texto, e declara que tudo está em paz, pois, segundo ele, os resultados são perfeitamente compatíveis:

$$t_{joaquim} = (87.5 \pm 0.5)^{\circ}C$$

$$t'_{charles} = (87.78 \pm 0.28)^{\circ}C$$

Para resolver o problema da discrepância aparente, faltava indicar a unidade e a incerteza (ou barra de erro) nos resultados! Enquanto Joaquim utilizou um termômetro graduado em **graus centígrados** ($^{\circ}C$), o Charles, naturalmente, forneceu a temperatura em **graus fahrenheit** ($^{\circ}F$).

Sabendo disso, Zé-do-morro transformou os $190^{\circ}F$ em $87.78^{\circ}C$, usando a relação conhecida (você conhece?)

$$^{\circ}C = (^{\circ}F - 32) \frac{5}{9}$$

Em seguida, ele *propagou* a incerteza experimental de leitura

$$\sigma_{t_{charles}} = 0.5^{\circ}F$$

do termômetro do Charles, convertendo-a na incerteza

$$\sigma_{t'_{charles}} = \frac{5}{9} \sigma_{t_{charles}} = 0.28^{\circ}C$$

expressa em graus centígrados. Era só isso !!!

Paciência, daqui a pouco você saberá propagar incertezas e entenderá também porque Zé-do-morro chegou à conclusão de que os resultados dessa **brincadeira experimental** eram perfeitamente compatíveis. O importante, no momento, é você estar **definitivamente** convencido de que qualquer resultado só tem sentido se acompanhado de suas **incerteza e unidade**.

Lembretes Fundamentais

No intuito de poupar muitas idas e voltas entre este texto e a referência bibliográfica sugerida, vamos lembrar conceitos e definições de maior importância, para poder, em seguida, aprender como propagar incertezas, ajustar curvas teóricas a dados experimentais e testar hipóteses. Consideraremos

somente processos aleatórios **discretos** cujos resultados possíveis são os valores x_i de uma variável aleatória **discreta** x . Evidentemente, os resultados aplicam-se também a variáveis **contínuas**, transformando as somas discretas \sum em integrais \int .

Observação: somente incertezas **estatísticas** serão consideradas neste texto, e você deverá consultar as referências bibliográficas para lidar com outros tipos de incertezas, como por exemplo as incertezas **sistemáticas**.

- **Valor Médio para Medidas *Idênticas***

Medidas idênticas de uma grandeza são as realizadas n vezes, onde n é um número finito, pelo **mesmo experimentador**, com os **mesmos instrumentos** e nas **mesmas condições experimentais**.

O valor médio \bar{x} do **conjunto** de valores discretos $x_i (i = 1, n)$ da variável x é definido como:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

O valor médio **verdadeiro** μ_x , obviamente impossível de conhecer exatamente, é dado por:

$$\mu_x \equiv \bar{x}_v = \lim_{n \rightarrow \infty} \bar{x} \quad (2)$$

Quando existe um **erro sistemático** nas medidas, em geral devido ao uso de instrumentos descalibrados, o valor médio verdadeiro μ_x não é igual ao **valor verdadeiro** X de x . Define-se, então, o **erro sistemático** δ_x como

$$\delta_x = |\mu_x - X| \quad (3)$$

o que não deixa de ser engraçado pois nem μ_x , nem X são conhecidos de maneira exata, isto é, com precisão infinita. O máximo que podemos fazer é obter a **melhor estimativa possível** para esses valores.

Assume-se que a melhor estimativa possível para μ_x , quando se realizam n medidas idênticas, é \bar{x} .

- **Variância para Medidas *Idênticas***

A **variância** σ^2 do **conjunto** de n medidas idênticas é definida por:

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2 \quad (4)$$

expressão que mostra que essa variância é a média dos quadrados dos **desvios** $d_i \equiv x_i - \mu_x$ das medidas, em torno de seu valor médio verdadeiro μ_x .

O **desvio padrão** $\sigma \equiv +\sqrt{\sigma^2}$ caracteriza a dispersão das medidas em relação a μ_x . Mostra-se que o desvio padrão no **valor médio** é dado por:

ver o § 7.3 do Vuolo (96)

$$\sigma_m = \frac{\sigma}{\sqrt{n}} \quad (5)$$

A melhor estimativa da variância σ^2 , de fato desconhecida, (**a propósito, você pode explicar porque?**) é:

ver o § 7.4 do Vuolo (96)

$$\sigma^2 \simeq \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \sum_{i=1}^n x_i^2 - \frac{n}{n-1} \bar{x}^2 \quad (6)$$

Exercício 3

Demonstre a Equação (6).

A **variância do processo de medida**, ou **variância verdadeira**, é o limite, quando o número de medidas tende ao ∞ , da variância do conjunto:

$$\sigma_v^2 = \lim_{n \rightarrow \infty} \sigma^2 \quad (7)$$

- **Covariância para Medidas *Idênticas***

Considere um conjunto de n medidas de várias variáveis x, y, z, \dots . Defina-se a **covariância desse conjunto** para um par qualquer de variáveis, x e y por exemplo, como:

$$\sigma_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (8)$$

A **covariância do processo de medida**, ou **covariância verdadeira**, é o limite da covariância do conjunto quando o número n de medidas tende ao ∞ :

$$\{\sigma_{xy}\}_v = \lim_{n \rightarrow \infty} \sigma_{xy} \quad (9)$$

Observação: enquanto a variância é sempre positiva, a covariância pode ser tanto positiva como negativa ou nula.

A melhor estimativa experimental da **covariância do conjunto** é:

$$\sigma_{xy} \simeq \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (10)$$

- **Correlação e Independência Estatísticas**

Se duas variáveis aleatórias x e y , com distribuições **simétricas** em torno de seus valores médios verdadeiros μ_x e μ_y , forem **estatisticamente independentes**, isto é se não existir **correlação** entre os desvios

$$d_{x_i} = x_i - \mu_x \quad e \quad d_{y_i} = y_i - \mu_y \quad (11)$$

então os produtos $d_{x_i} d_{y_i}$ estão distribuídos **aleatoriamente e simetricamente** em torno de 0. Portanto, nesse caso, a covariância de conjunto σ_{xy} deve ser pequena e tender a 0 conforme o número n de medidas aumenta e, pela definição da covariância de processo, tem-se:

$$\{\sigma_{xy}\}_v = 0 \quad (12)$$

Mas, CUIDADO, a covariância σ_{xy} pode ser nula, mesmo quando há correlação entre os desvios!!!

Exercício 4

Mostre que numa correlação **total** do tipo

$$d_{y_i} = d_{x_i}^2 + k \quad (13)$$

onde k é uma constante qualquer, a covariância verdadeira $\{\sigma_{xy}\}_v$ é **rigorosamente** nula. Você deverá usar um resultado que será demonstrado mais adiante (Equação 20): o valor médio verdadeiro μ_α de uma função $\alpha(x, y, z, \dots)$ de várias variáveis aleatórias é igual à função $\alpha(\mu_x, \mu_y, \mu_z, \dots)$ dos valores médios verdadeiros das variáveis.

- Coeficiente de Correlação

Define-se o **coeficiente de correlação** entre duas variáveis aleatórias x e y da seguinte maneira:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (14)$$

Vamos entender o significado mais profundo dessa simples definição algébrica.

Considere por exemplo a Figura 1 a seguir, na qual uma distribuição de desvios d_{x_i} e d_{y_i} é representada no plano $d_x - d_y$.

Apesar da dispersão observada nessa figura, você pode notar uma **tendência** dos pontos experimentais a se distribuírem em torno de uma certa direção. Seja Δ_i a distância vertical de um ponto qualquer à reta arbitrária \Re representada pela equação $d_y = a d_x$:

$$\Delta_i = d_{y_i} - a d_{x_i} \quad (15)$$

Vamos tentar descobrir qual o coeficiente angular a_{opt} que define de maneira mais correta a **direção** \Re_{opt} que os pontos experimentais parecem indicar.

A soma dos quadrados das distâncias Δ_i é

$$S = \sum_{i=1}^n \Delta_i^2 \quad (16)$$

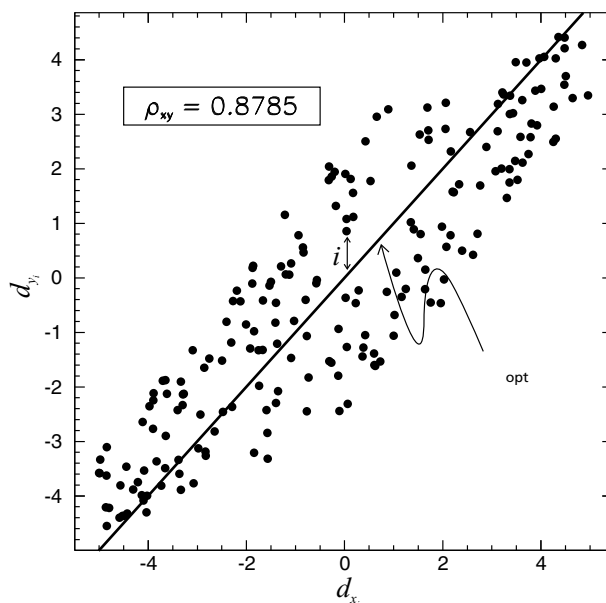


Figura 1: Exemplo de variáveis aleatórias estatisticamente correlacionadas.

Daqui a algumas páginas você vai entender o que afirmamos a seguir, ou seja:

o coeficiente angular a_{opt} é aquele que minimiza a soma S , isto é, aquele que anula a derivada $\frac{dS}{da}$.

veja o item Método dos
Mínimos Quadrados

Exercício 5

Calcule a derivada $\frac{dS}{da}$ e utilize as definições da covariância e da variância de conjunto para mostrar que:

$$a_{opt} = \frac{\sum_{i=1}^n d_{x_i} d_{y_i}}{\sum_{i=1}^n d_{x_i}^2} = \frac{\sigma_{xy}}{\sigma_x^2} \quad (17)$$

Agora é fácil mostrar, usando as definições das variâncias, covariâncias e coeficiente de correlação, que

$$S = n\sigma_y^2(1 - \rho_{xy}^2) \quad (18)$$

A soma S , sendo positiva ou nula, resulta em $|\rho_{xy}| \leq 1$.

Se o número n de medidas em qualquer experiência for finito, a covariância de conjunto σ_{xy} e, portanto, o coeficiente de correlação ρ_{xy} não são nulos, mesmo se as variáveis x e y forem estatisticamente independentes. Você teria de ter muita sorte para obter um coeficiente de correlação nulo com meia dúzia de medidas! Entretanto, quando $n \rightarrow \infty$, $\rho_{xy} \rightarrow 0$ para um par xy de variáveis **independentes**.

Mas cuidado, a recíproca não é necessariamente verdadeira!

As duas figuras a seguir são exemplos de coeficientes de correlação $|\rho_{xy}| \ll 1$, tanto para desvios estatisticamente independentes, ou não correlatos, Figura 2, como dependentes, ou correlatos, Figura 3.

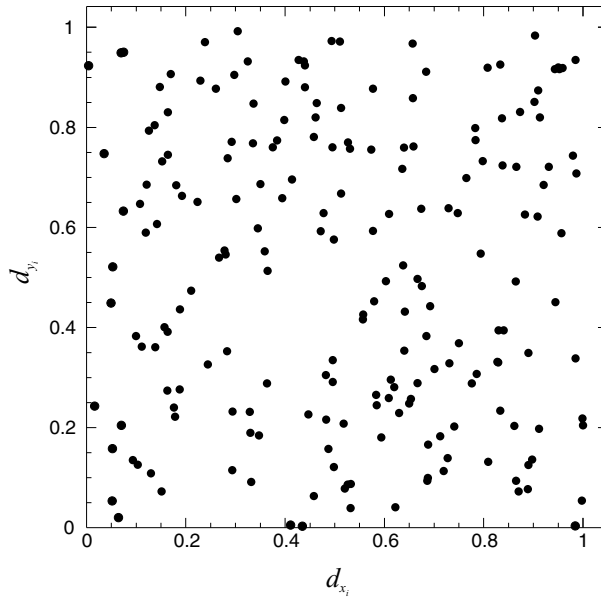


Figura 2: Exemplo de variáveis estatisticamente independentes, com coeficiente de correlação $\rho_{xy} = -0,0293$.

Contudo, se a correlação estatística for total (ou máxima), $\rho_{xy} = \pm 1$, $S = 0$ e os pontos experimentais no plano $d_x - d_y$ estão todos situados na reta \mathcal{R}_{opt} da Figura 1.

Neste ponto temos (quase!) todos ingredientes para atacar o problema da propagação de incertezas, desde que o conceito de derivada parcial esteja claro na sua cabeça!

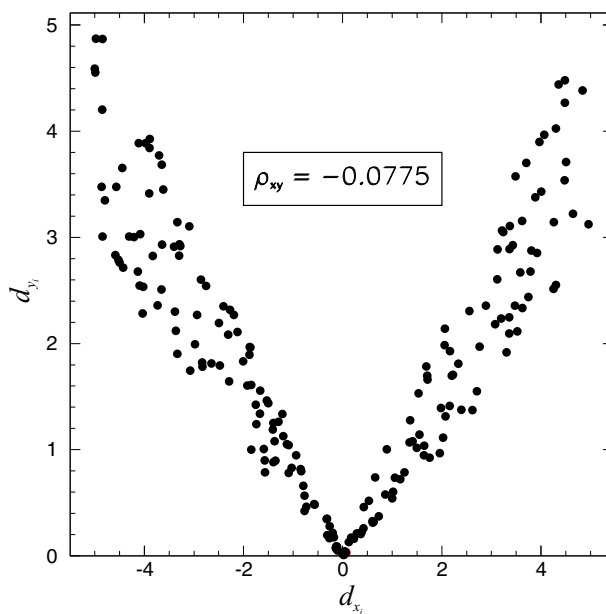


Figura 3: Exemplo de variáveis estatisticamente dependentes.

Propagação de Incertezas

Quando grandezas físicas são medidas **diretamente** com instrumentos, as incertezas experimentais são **estimadas**. A partir desses resultados, fórmulas matemáticas permitem medir **indiretamente** outras grandezas cujas incertezas são obtidas **propagando** as incertezas experimentais. Por exemplo, medindo o comprimento $l \pm \sigma_l$ e o período $T \pm \sigma_T$ de um pêndulo simples com uma trena e um cronômetro, respectivamente, podemos determinar o valor da aceleração da gravidade g , pois sabemos que $g = 4\pi^2 \frac{l}{T^2}$ (isto é demonstrado no curso de Física 2B). O novo desafio consiste em determinar a incerteza σ_g sobre o valor da medida indireta de g . Como? É isso que vamos descobrir agora!

• Propagação das Médias e das Variâncias

Considere um conjunto de grandezas experimentais representadas pelas variáveis aleatórias x, y, z, \dots , médias verdadeiras $\mu_x, \mu_y, \mu_z, \dots$, variâncias $\sigma_x^2, \sigma_y^2, \sigma_z^2, \dots$ e covariâncias $\sigma_{xy}, \sigma_{xz}, \sigma_{yz}, \dots$

Seja, agora, uma nova grandeza $\alpha = \alpha(x, y, z, \dots)$ função das primeiras. O desafio consiste em determinar seu valor médio $\bar{\alpha}$ e estimar

seu valor médio verdadeiro μ_α e sua variância σ_α^2 , ou, em outras palavras, **propagar as incertezas experimentais**.

Ao medir n vezes o conjunto x, y, z, \dots , obtém-se os resultados

$$[x_i, y_i, z_i, \dots; \alpha_i(x_i, y_i, z_i, \dots)] \quad \text{para } i = 1, \dots, n$$

Podemos expandir os α_i em séries de potências dos desvios:

$$\begin{aligned} \alpha_i &\simeq \alpha(\mu_x, \mu_y, \mu_z, \dots) \\ &+ \left. \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x} (x_i - \mu_x) + \left. \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y} (y_i - \mu_y) + \dots \\ &+ \frac{1}{2} \left. \frac{\partial^2 \alpha}{\partial x^2} \right|_{x=\mu_x} (x_i - \mu_x)^2 + \frac{1}{2} \left. \frac{\partial^2 \alpha}{\partial y^2} \right|_{y=\mu_y} (y_i - \mu_y)^2 + \dots \end{aligned} \quad (19)$$

É razoável desprezar os termos de ordem ≥ 2 se condições do tipo

$$\left. \frac{\partial^2 \alpha}{\partial x^2} \right|_{x=\mu_x} (x_i - \mu_x)^2 \simeq 0 \quad \text{para } (x_i - \mu_x) \simeq \sigma_x$$

são satisfeitas.

Nesse caso, as derivadas primeiras são praticamente constantes para desvios da ordem de $\pm\sigma_x, \pm\sigma_y, \pm\sigma_z, \dots$, em torno de $\mu_x, \mu_y, \mu_z, \dots$, respectivamente, isto é, a função α varia **lentamente em torno do ponto definido pelas coordenadas** $\mu_x, \mu_y, \mu_z, \dots$

Você pode se perguntar: o que acontece com as expansões dos α_i para desvios maiores que os considerados até agora? A resposta é simples e direta: não tem o menor interesse, pois a ocorrência de tais desvios é muito pouco provável!

ver os § 4.3 e § 8.3 do Vuolo(96).

Portanto, podemos escrever

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \alpha_i &\simeq \alpha(\mu_x, \mu_y, \mu_z, \dots) + \left. \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x} \frac{1}{n} \sum_{i=1}^n (x_i - \mu_x) \\ &+ \left. \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y} \frac{1}{n} \sum_{i=1}^n (y_i - \mu_y) + \dots \end{aligned}$$

Quando $n \rightarrow \infty$, é fácil mostrar que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \alpha_i \simeq \alpha(\mu_x, \mu_y, \mu_z, \dots)$$

pois os demais termos se anulam em consequência direta da definição do valor médio verdadeiro de uma variável.

O valor médio verdadeiro μ_α de α é, também por definição,

$$\mu_\alpha = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \alpha_i$$

Assim, obtemos

$$\mu_\alpha \simeq \alpha(\mu_x, \mu_y, \mu_z, \dots) \quad (20)$$

Da mesma forma, expandindo os α_i em séries de potências dos desvios, agora em torno dos valores médios de **conjunto** $\bar{x}, \bar{y}, \bar{z}, \dots$, chegamos à conclusão que

$$\bar{\alpha} \simeq \alpha(\bar{x}, \bar{y}, \bar{z}, \dots) \quad (21)$$

Exercício 6

Demonstre o resultado anterior e discuta-o com seu tutor.

Observação: se os valores médios verdadeiros forem desconhecidos, o valor médio de conjunto $\bar{\alpha}$ será a aproximação mais provável do valor médio de processo (ou verdadeiro) μ_α .

Temos agora de propagar as variâncias, sabendo que, por definição:

$$\sigma_\alpha^2 = \frac{1}{n} \sum_{i=1}^n (\alpha_i - \mu_\alpha)^2 \quad (22)$$

Usando as expansões dos α_i somente até a primeira ordem (ver Equação 19) e a expressão de μ_α (ver a Equação 20), podemos escrever

$$\begin{aligned} (\alpha_i - \mu_\alpha)^2 &= \left| \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x}^2 (x_i - \mu_x)^2 + \left| \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y}^2 (y_i - \mu_y)^2 + \dots \\ &+ 2 \left. \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x} \left. \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y} (x_i - \mu_x)(y_i - \mu_y) + \dots \\ &+ 2 \left. \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y} \left. \frac{\partial \alpha}{\partial z} \right|_{z=\mu_z} (y_i - \mu_y)(z_i - \mu_z) + \dots \end{aligned}$$

Somando os dois membros da equação anterior sobre $i = 1, \dots, n$ e dividindo-os por n , obtemos, lembrando as definições das variâncias e co-variâncias de conjunto:

$$\begin{aligned} \sigma_\alpha^2 &= \left| \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x}^2 \sigma_x^2 + \left| \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y}^2 \sigma_y^2 + \left| \frac{\partial \alpha}{\partial z} \right|_{z=\mu_z}^2 \sigma_z^2 + \dots \\ &+ 2 \frac{\partial \alpha}{\partial x} \bigg|_{x=\mu_x} \frac{\partial \alpha}{\partial y} \bigg|_{y=\mu_y} \sigma_{xy} + 2 \frac{\partial \alpha}{\partial x} \bigg|_{x=\mu_x} \frac{\partial \alpha}{\partial z} \bigg|_{z=\mu_z} \sigma_{xz} + \dots \end{aligned} \quad (23)$$

Podemos utilizar essa fórmula geral (23) para encontrar as variâncias de quatro funções que você já encontrou e *cansará* de encontrar. Os resultados estão na Tabela 1 a seguir.

Função	Variância de conjunto
$\alpha = x \pm y$	$\sigma_\alpha^2 = \sigma_x^2 + \sigma_y^2 \pm 2 \sigma_{xy}$
$\alpha = a x^m y^n$	$\left\{ \frac{\sigma_\alpha}{\alpha} \right\}^2 = m^2 \left\{ \frac{\sigma_x}{x} \right\}^2 + n^2 \left\{ \frac{\sigma_y}{y} \right\}^2 + 2 m n \frac{\sigma_{xy}}{xy}$
$\alpha = m \operatorname{sen}(nx)$	$\sigma_\alpha^2 = \{m n \cos(nx)\}^2 \sigma_x^2$
$\alpha = m \log_a x \equiv m \frac{\ln(x)}{\ln(a)}$	$\sigma_\alpha^2 = \left\{ \frac{m}{\ln(a)} \right\}^2 \left\{ \frac{\sigma_x}{x} \right\}^2$

Tabela 1: Exemplos de propagação de variâncias.

Exercício 7

Deduz os resultados da Tabela 1.

Considerando o limite dessa Equação (23) quando $n \rightarrow \infty$, obtemos a variância verdadeira de α :

$$\begin{aligned}
\{\sigma_\alpha^2\}_v &= \left| \frac{\partial \alpha}{\partial x} \right|_{x=\mu_x}^2 \{\sigma_x^2\}_v + \left| \frac{\partial \alpha}{\partial y} \right|_{y=\mu_y}^2 \{\sigma_y^2\}_v + \left| \frac{\partial \alpha}{\partial z} \right|_{z=\mu_z}^2 \{\sigma_z^2\}_v + \dots \\
&+ 2 \frac{\partial \alpha}{\partial x} \bigg|_{x=\mu_x} \frac{\partial \alpha}{\partial y} \bigg|_{y=\mu_y} \{\sigma_{xy}\}_v + 2 \frac{\partial \alpha}{\partial x} \bigg|_{x=\mu_x} \frac{\partial \alpha}{\partial z} \bigg|_{z=\mu_z} \{\sigma_{xz}\}_v + \dots
\end{aligned} \tag{24}$$

Obviamente, as fórmulas da Tabela 1 se aplicam ao caso de **variâncias verdadeiras**.

• Propagação das Covariâncias

Seja agora uma outra função $\beta = \beta(x, y, z, \dots)$ das mesmas variáveis x, y, z, \dots . Obviamente, existe uma correlação funcional entre as variáveis α e β e, portanto, uma covariância $\sigma_{\alpha\beta}$ propagada a partir das variâncias $\sigma_x^2, \sigma_y^2, \sigma_z^2, \dots$ e das covariâncias $\sigma_{xy}, \sigma_{xz}, \sigma_{yz}, \dots$. Por definição, sabemos que

$$\sigma_{\alpha\beta} = \frac{1}{n} \sum_{i=1}^n (\alpha_i - \mu_\alpha)(\beta_i - \mu_\beta) \tag{25}$$

onde μ_α e μ_β são os valores médios verdadeiros de α e β , respectivamente, $\alpha_i = \alpha(x_i, y_i, z_i, \dots)$ e $\beta_i = \beta(x_i, y_i, z_i, \dots)$. Outrossim, lembrando a propagação das médias (Equação 20), os valores médios verdadeiros de α e β são funções dos valores médios verdadeiros $\mu_x, \mu_y, \mu_z, \dots$:

$$\mu_\alpha = \alpha(\mu_x, \mu_y, \mu_z, \dots) \quad e \quad \mu_\beta = \beta(\mu_x, \mu_y, \mu_z, \dots) \tag{26}$$

Expandido, somente até a primeira ordem, os α_i e β_i em séries de potências de desvios das variáveis x_i, y_i, z_i, \dots em torno dos seus valores médios verdadeiros correspondentes $\mu_x, \mu_y, \mu_z, \dots$, temos:

$$\begin{aligned}
\alpha_i &\simeq \alpha(\mu_x, \mu_y, \mu_z, \dots) + \frac{\partial \alpha}{\partial x}(x_i - \mu_x) + \frac{\partial \alpha}{\partial y}(y_i - \mu_y) + \dots \\
\beta_i &\simeq \beta(\mu_x, \mu_y, \mu_z, \dots) + \frac{\partial \beta}{\partial x}(x_i - \mu_x) + \frac{\partial \beta}{\partial y}(y_i - \mu_y) + \dots
\end{aligned}$$

onde as derivadas parciais são calculadas nos valores médios verdadeiros de cada variável.

Usando as expressões (26), podemos reescrever as duas equações anteriores da seguinte forma:

$$\begin{aligned}\alpha_i - \mu_\alpha &\simeq \frac{\partial \alpha}{\partial x}(x_i - \mu_x) + \frac{\partial \alpha}{\partial y}(y_i - \mu_y) + \dots \\ \beta_i - \mu_\beta &\simeq \frac{\partial \beta}{\partial x}(x_i - \mu_x) + \frac{\partial \beta}{\partial y}(y_i - \mu_y) + \dots\end{aligned}$$

Transportando, agora, estas duas expressões na Equação (25), obtemos:

$$\begin{aligned}\sigma_{\alpha\beta} \simeq \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\partial \alpha}{\partial x} \frac{\partial \beta}{\partial x} (x_i - \mu_x)^2 + \frac{\partial \alpha}{\partial y} \frac{\partial \beta}{\partial y} (y_i - \mu_y)^2 + \dots \right. \\ \left. + \left(\frac{\partial \alpha}{\partial x} \frac{\partial \beta}{\partial y} + \frac{\partial \alpha}{\partial y} \frac{\partial \beta}{\partial x} \right) (x_i - \mu_x)(y_i - \mu_y) + \dots \right\}\end{aligned}$$

Lembrando, uma vez mais, as definições das variâncias e das covariâncias, podemos concluir a propagação das covariâncias:

$$\begin{aligned}\sigma_{\alpha\beta} &\simeq \frac{\partial \alpha}{\partial x} \frac{\partial \beta}{\partial x} \sigma_x^2 + \frac{\partial \alpha}{\partial y} \frac{\partial \beta}{\partial y} \sigma_y^2 + \dots \\ &+ \left\{ \frac{\partial \alpha}{\partial x} \frac{\partial \beta}{\partial y} + \frac{\partial \alpha}{\partial y} \frac{\partial \beta}{\partial x} \right\} \sigma_{xy} + \dots\end{aligned}\tag{27}$$

Se as variáveis x , y , z, \dots forem independentes, sabemos que suas covariâncias serão nulas. Portanto, o resultado anterior se torna mais simples:

$$\sigma_{\alpha\beta} \simeq \frac{\partial \alpha}{\partial x} \frac{\partial \beta}{\partial x} \sigma_x^2 + \frac{\partial \alpha}{\partial y} \frac{\partial \beta}{\partial y} \sigma_y^2 + \dots\tag{28}$$

Um pequeno lembrete: na prática, usaremos sempre as médias, variâncias e covariâncias de conjunto, uma vez que as de processo (ou verdadeiras) são, por essência, desconhecidas.

• Transferência de Incertezas

Uma situação experimental bastante freqüente é a da medida de duas grandezas **funcionalmente dependentes mas estatisticamente independentes**. Por exemplo, o período T de um pêndulo simples é

uma função $T(l)$ do seu comprimento l . Existe, portanto, uma dependência funcional entre l e T . Porém, essas grandezas são medidas de maneira diferente e com instrumentos completamente diferentes (régua e cronômetro), sendo preciso atribuir-lhes desvios padrões σ_l e σ_T .

Vamos supor que você não saiba qual a dependência funcional $T(l)$, mas que você deseja descobri-la experimentalmente, realizando um conjunto de medidas $\{l_i, T_i\}$. A Figura 4 mostra os resultados dessas medidas com barras de erro $2\sigma_{l_i}$ e $2\sigma_{T_i}$.

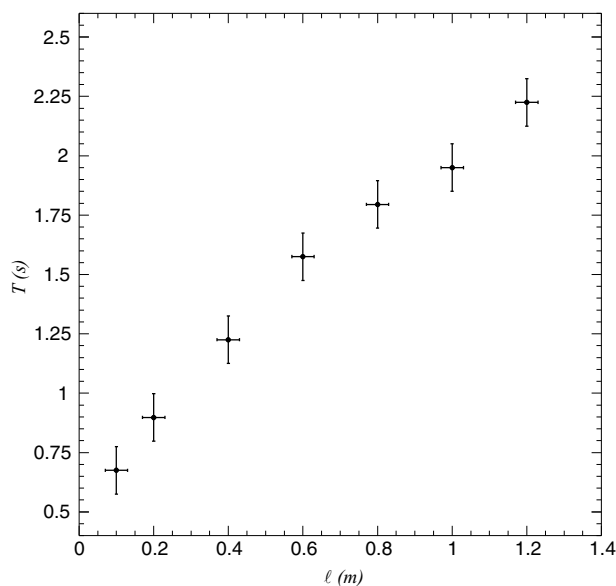


Figura 4: Período de um pêndulo simples em função de seu comprimento l . As barras de erro representam as incertezas experimentais.

Obviamente, você não está interessado em descobrir os **valores verdadeiros** de l e T !

O que vamos fazer agora é **transformar** essa situação experimental **real** em uma outra situação experimental que **poderia ser real** e na qual todos os desvios padrões σ_{l_i} sejam nulos. A representação gráfica dessa nova situação experimental pode ser vista na Figura 5, onde os pares $\{l_i, T_i\}$ são naturalmente os mesmos que os das Figura 4, mas onde as barras de erro $2\sigma'_{T_i}$, maiores que as $2\sigma_{T_i}$, **traduzem o desaparecimento** das incertezas sobre os l_i .

O desafio consiste em determinar qual o valor Υ^2 que satisfaz à equação

$$\sigma_{T_i}^{\prime 2} = \sigma_{T_i}^2 + \Upsilon^2$$

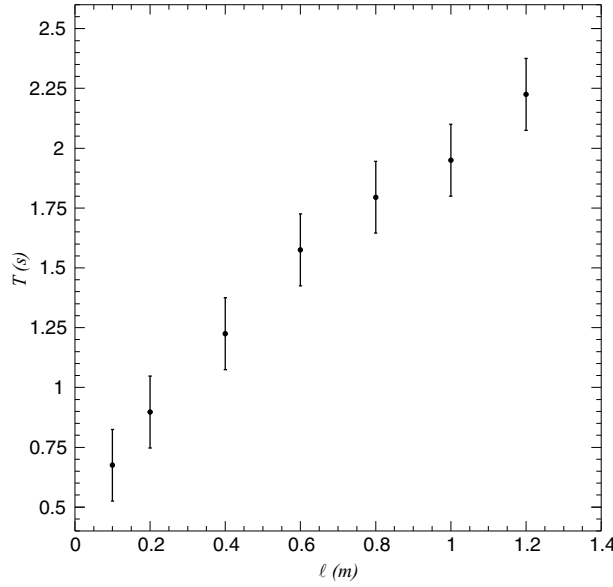


Figura 5: Período de um pêndulo simples em função de seu comprimento l , após a transferência das incertezas sobre o comprimento.

Para isso, introduzimos uma **variável aleatória auxiliar**

$$w = T - T(l)$$

onde T é a variável aleatória **medida** e $T(l)$ a dependência funcional entre a variável aleatória l medida e o período do pêndulo simples. Supondo l e T **estatisticamente independentes**, temos, usando a Equação (23) com covariância nula,

$$\sigma_w^2 = \left\{ \frac{\partial w}{\partial l} \right\}^2 \sigma_l^2 + \left\{ \frac{\partial w}{\partial T} \right\}^2 \sigma_T^2$$

ou seja,

$$\sigma_w^2 = \left\{ \frac{dT}{dl} \right\}^2 \sigma_l^2 + \sigma_T^2$$

Seja agora l^* o comprimento considerado **sem incerteza**. Podemos escrever

$$T^* = T(l^*) + (T - T(l)) = T(l^*) + w$$

onde T^* representa o período numa aquisição de dados **sem incerteza no comprimento l** .

Assim, $\sigma_{T(l^*)} = 0$ e $\sigma_{T^*} = \sigma_w$, o que demonstra que

$$\Upsilon^2 = \left\{ \frac{dT}{dl} \right\}^2 \sigma_l^2$$

Obtemos, então, a resposta final, isto é, sabemos como transferir as incertezas no comprimento l para as incertezas sobre o período T :

$$\sigma_{T*}^2 = \sigma_T^2 + \left\{ \frac{dT}{dl} \right\}^2 \sigma_l^2 \quad (29)$$

Após a transferência, a nova variância σ_{T*}^2 é chamada **variância efetiva**.

Entretanto, temos um **pequeno problema**, pois, não conhecendo (ainda!) qual a **dependência funcional** entre l e T , não podemos calcular a derivada $\frac{dT}{dl}$. O melhor que se pode esperar é **estimar** essa derivada, seja **chutando** uma função que **poderia** representar a **dependência funcional experimental** $\{l_i, T_i\}$, seja calculando os coeficientes angulares das tangentes a uma curva **suave** traçada **a olho** e passando o mais perto possível dos pontos experimentais, como indicado na Figura 6.

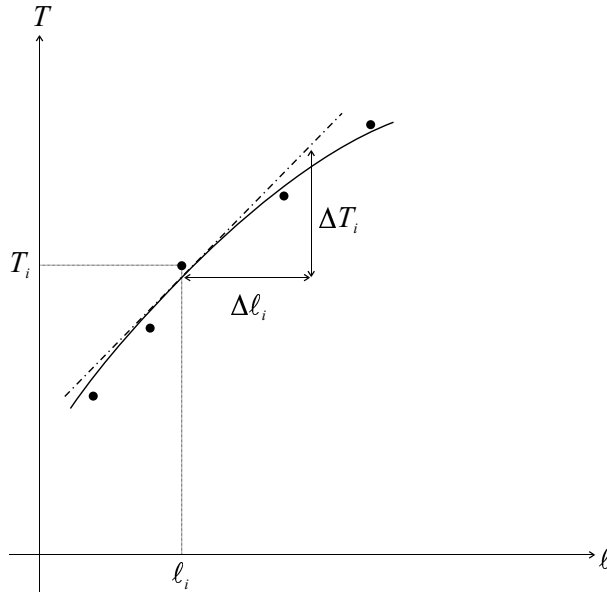


Figura 6: Estimativa “gráfica” das derivadas $\frac{dT}{dl}$.

Uma última pergunta: qual incerteza transferir? No exemplo anterior, transferimos a incerteza no comprimento l para a incerteza no período T , mas poderíamos ter feito o contrário. Como regra de conduta, é razoável transferir a incerteza experimental **menor para a maior**.

Nessa altura do campeonato, de posse de um conjunto de dados, sabemos estimar e propagar as incertezas experimentais, faltando agora descobrir que lei explica o que foi observado, isto é, falta **ajustar** uma função matemática aos dados.

Ajustes de Funções

Apresentaremos, a seguir, os dois métodos mais conhecidos, e consequentemente mais utilizados, para ajustar funções a um conjunto de dados experimentais, o de **Máxima Verossimilhança** e o dos **Mínimos Quadrados**, sendo, o segundo, um caso particular do primeiro.

• Método de Máxima Verossimilhança

A aquisição cuidadosa e o tratamento rigoroso de dados experimentais são necessários, mas não suficientes, para desvendar os mistérios da natureza. Falta **entender e/ou explicar** os resultados obtidos, isto é, falta **modelizar** ou, em outras palavras, obter a função que melhor descreve os dados. Essa tarefa, muitas vezes bastante delicada, chama-se **ajuste de uma função a um conjunto de dados** ou **regressão**. A priori, muitas funções podem ser **ajustadas** a um conjunto particular de dados, inclusive funções escolhidas arbitrariamente sem nenhuma justificativa. Porém, a modelização implica a escolha de funções com formas **predeterminadas**. O ajuste, por sua vez, consiste em determinar os parâmetros **livres** dessas funções. Vamos entender melhor com um exemplo simples.

Suponhamos que você realizou uma experiência de queda livre, registrando pares de dados **espaço-tempo**, supondo a variável independente “**tempo**” sem incerteza e a variável dependente “**espaço**” com incertezas. Examinando seu gráfico, você pode pensar em ajustar um polinômio. De que grau? Primeiro, segundo, terceiro, ...? Claro que, com seus conhecimentos em Física (Segunda Lei de Newton !), você vai tentar ajustar um polinômio de segundo grau, ou seja, uma forma parabólica. Pela “cara” do seu gráfico, você já descartou o de primeiro grau, pois seus pontos, com certeza, não estão alinhados numa reta. Mas porque não tentar uma regressão com polinômios de ordem maior que 2 ou com funções do tipo hiperbólico? A resposta vai ser dada logo a seguir.

Seja um conjunto de n pontos experimentais $\{x_i, y_i, \sigma_i\}$, a variável **independente** x sendo considerada isenta de incerteza (caso contrário, é só fazer uma transferência para a **variável dependente** y , como já foi explicado). O desvio padrão σ_i é a incerteza **estatística** na variável y . Para ajustar a **melhor** função $f(x)$ aos dados, é preciso ter um critério objetivo. O critério mais utilizado é o de **Máxima Verossimilhança**:

a função $f(x)$ que melhor descreve o conjunto $\{x_i, y_i, \sigma_i\}$ de dados experimentais é aquela para a qual esse conjunto é o mais verossímil possível.

Admite-se implicitamente que, do ponto de vista estatístico, a experiência forneceu os dados que tinham a maior probabilidade de ser obtidos, descartando qualquer hipótese de erros grosseiros ou sistemáticos. Se a forma funcional $f(x)$ depende de parâmetros a_j , com $j = 1, \dots, p$, o desafio é encontrar os valores dos a_j que maximizam a probabilidade de **acordo** entre os dados e a função. Para isso, temos de formular algebricamente essa maximização, o que faremos a seguir introduzindo o método dos mínimos quadrados derivado do método de máxima verossimilhança.

• Método dos Mínimos Quadrados

Considere o conjunto de n medidas $\{x_i, y_i, \sigma_i\}$, com distribuições **gaussianas** de erros na variável dependente y , e a **melhor função** $f(x; a_1, a_2, a_3, \dots, a_p)$ cuja forma e número de parâmetros já foram determinados por **boas razões físicas**.

A probabilidade P_i de obter um resultado (x_i, y_i, σ_i) é proporcional à probabilidade gaussiana

$$P_i = \frac{K}{\sigma_i} e^{-\frac{1}{2} \left\{ \frac{y_i - \mu_i}{\sigma_i} \right\}^2}$$

onde μ_i é o valor médio verdadeiro de y_i e K é uma constante de proporcionalidade.

A probabilidade de uma experiência fornecer o conjunto $\{x_i, y_i, \sigma_i\}$ é igual ao produto das probabilidades P_i :

$$P = \frac{K^n}{\sigma_1 \sigma_2 \dots \sigma_n} e^{-\frac{1}{2} \sum_{i=1}^n \left\{ \frac{y_i - \mu_i}{\sigma_i} \right\}^2}$$

De acordo com o Método de Máxima Verossimilhança, sendo $f(x; a_1, a_2, a_3, \dots, a_p)$ a função **verdadeira**, os parâmetros ótimos são os que maximizam a probabilidade P que pode ser reescrita como

$$P = \frac{K^n}{\sigma_1 \sigma_2 \dots \sigma_n} e^{-\frac{1}{2} \chi^2}$$

onde

$$\chi^2 = \sum_{i=1}^n \left\{ \frac{y_i - f(x_i; a_1, a_2, a_3, \dots, a_p)}{\sigma_i} \right\}^2 \quad (30)$$

O símbolo χ^2 é chamado “qui quadrado”.

Como a probabilidade P é uma função decrescente de χ^2 , temos que **minimizar o χ^2** , isto é, temos de impor as p condições a seguir para maximizar esta probabilidade.

$$\frac{\partial \chi^2}{\partial a_j} = 0 \quad \text{para} \quad j = 1, \dots, p \quad (31)$$

Vamos agora aplicar esse método à regressão de funções **lineares nos parâmetros** que são, em muitos casos, excelentes aproximações de funções quaisquer.

• Ajuste de Funções Lineares nos Parâmetros

Essas funções são, por definição, do tipo

$$f(x; a_1, a_2, a_3, \dots, a_p) = \sum_{j=1}^p a_j f_j(x) \quad (32)$$

sendo as funções $f_j(x)$ funções e **linearmente independentes**.

As p condições simultâneas (31) que minimizam o χ^2 escrevem-se:

$$-2 \sum_{i=1}^n \frac{1}{\sigma_i^2} \left\{ y_i - \sum_{k=1}^p a_k f_k(x_i) \right\} f_j(x_i) = 0 \quad \text{para} \quad j = 1, \dots, p$$

Colocando em evidência os parâmetros a_j , podemos ver que temos de resolver um sistema de p equações lineares

$$\sum_{k=1}^p \left\{ \sum_{i=1}^n \frac{f_j(x_i) f_k(x_i)}{\sigma_i^2} \right\} a_k = \sum_{i=1}^n \frac{y_i f_j(x_i)}{\sigma_i^2} \quad \text{para } j = 1, \dots, p$$

No intuito de tornar as expressões menos **pesadas**, é conveniente introduzir a notação seguinte nos somatórios sobre o index i :

$$f_j \equiv f_j(x_i)$$

Introduzindo as matrizes colunas

$$\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \dots \\ \dots \\ \dots \\ a_p \end{pmatrix} \quad \text{e} \quad \mathbf{B} = \begin{pmatrix} \sum_{i=1}^n \frac{y_i f_1}{\sigma_i^2} \\ \sum_{i=1}^n \frac{y_i f_2}{\sigma_i^2} \\ \sum_{i=1}^n \frac{y_i f_3}{\sigma_i^2} \\ \sum_{i=1}^n \frac{y_i f_4}{\sigma_i^2} \\ \dots \\ \dots \\ \dots \\ \sum_{i=1}^n \frac{y_i f_p}{\sigma_i^2} \end{pmatrix} \quad (33)$$

e a matriz quadrada **simétrica** de ordem p

$$\mathbf{M} = \begin{pmatrix} \sum_{i=1}^n \frac{f_1^2}{\sigma_i^2} & \sum_{i=1}^n \frac{f_1 f_2}{\sigma_i^2} & \dots & \sum_{i=1}^n \frac{f_1 f_p}{\sigma_i^2} \\ \sum_{i=1}^n \frac{f_2 f_1}{\sigma_i^2} & \sum_{i=1}^n \frac{f_2^2}{\sigma_i^2} & \dots & \sum_{i=1}^n \frac{f_2 f_p}{\sigma_i^2} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n \frac{f_p f_1}{\sigma_i^2} & \sum_{i=1}^n \frac{f_p f_2}{\sigma_i^2} & \dots & \sum_{i=1}^n \frac{f_p^2}{\sigma_i^2} \end{pmatrix} \quad (34)$$

o sistema de equações lineares pode ser **condensado** sob a forma matricial

$$\mathbf{M} \mathbf{A} = \mathbf{B} \quad (35)$$

A solução deste sistema, única se as funções f_j forem linearmente independentes, é dada por

$$\mathbf{A} = \mathbf{M}^{-1} \mathbf{B} \quad (36)$$

o que significa que os parâmetros que minimizam o χ^2 são dados por

$$a_j = \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} b_k \quad \text{para } j = 1, \dots, p \quad (37)$$

$(\mathbf{M}^{-1})_{jk}$ sendo o elemento da linha j e da coluna k da matriz **inversa** da matriz \mathbf{M} e b_k o elemento da linha k da matriz coluna \mathbf{B} .

Neste ponto, supomos que você saiba inverter uma matriz. Caso contrário, seria muito oportuno consultar seu curso de álgebra linear!

• Variâncias e Covariâncias dos Parâmetros

Usando a definição (33) da matriz coluna \mathbf{B} , podemos reescrever os valores ótimos dos parâmetros sob a forma:

$$a_j = \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} \sum_{i=1}^n \frac{y_i f_k}{\sigma_i^2} \quad \text{para } j = 1, \dots, p \quad (38)$$

Se as medidas y_i forem consideradas **independentes**, isto é, se as covariâncias σ_{ij} forem nulas, as variâncias dos parâmetros a_j são obtidas aplicando a fórmula geral de propagação das variâncias (23):

$$\sigma_{a_j}^2 = \sum_{i=1}^n \left(\frac{\partial a_j}{\partial y_i} \right)^2 \sigma_i^2$$

Mas, derivando as equações (38), temos:

$$\frac{\partial a_j}{\partial y_i} = \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} \frac{f_k}{\sigma_i^2} \quad (39)$$

Assim,

$$\sigma_{a_j}^2 = \sum_{i=1}^n \frac{1}{\sigma_i^2} \left\{ \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} f_k \right\}^2$$

que pode também ser escrito como

$$\sigma_{a_j}^2 = \sum_{i=1}^n \frac{1}{\sigma_i^2} \left\{ \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} f_k \right\} \left\{ \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} f_l \right\}$$

ou, permutando os somatórios,

$$\sigma_{a_j}^2 = \sum_{k=1}^p \sum_{l=1}^p (\mathbf{M}^{-1})_{jk} (\mathbf{M}^{-1})_{jl} \sum_{i=1}^n \frac{f_k f_l}{\sigma_i^2}$$

Na expressão anterior, o último somatório é simplesmente o elemento $(\mathbf{M})_{kl}$ da matriz \mathbf{M} . Portanto, podemos escrever

$$\sigma_{a_j}^2 = \sum_{k=1}^p \sum_{l=1}^p (\mathbf{M}^{-1})_{jk} (\mathbf{M}^{-1})_{jl} (\mathbf{M})_{kl}$$

ou, ainda, permutando os dois somatórios,

$$\sigma_{a_j}^2 = \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} \left\{ \sum_{k=1}^p (\mathbf{M}^{-1})_{jk} (\mathbf{M})_{kl} \right\}$$

Você se lembra da regra de multiplicação de matrizes? Pois bem, então você deve ter notado que o último somatório é o elemento da linha j e da coluna l da matriz **unidade** \mathbf{I} , pois o produto de uma matriz pela sua inversa é, por definição, igual à matriz **unidade**. Você sabe também que este elemento, pode ser escrito como

$$(\mathbf{I})_{jl} = \delta_{jl}$$

onde o símbolo de Kroenecker δ_{jl} é definido por:

$$\begin{aligned} \delta_{jj} &= 1 && \text{e} \\ \delta_{jl} &= 0 && \text{para } j \neq l \end{aligned}$$

Portanto,

$$\sigma_{a_j}^2 = \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} \delta_{jl}$$

e, assim, chegamos à expressão final da variância dos parâmetros ajustados a_j :

$$\sigma_{a_j}^2 = (\mathbf{M}^{-1})_{jj} \quad (40)$$

Seguindo exatamente o mesmo procedimento, podemos calcular as covariâncias $\sigma_{a_j a_m}$, supondo as medidas y_i **independentes**. Para isso, utilizamos a Equação (28) que escrevemos agora:

$$\sigma_{a_j a_m} = \sum_{i=1}^n \frac{\partial a_m}{\partial y_i} \frac{\partial a_j}{\partial y_i} \sigma_i^2$$

Substituindo nessa expressão as derivadas parciais dadas pelas Equações (39), temos:

$$\sigma_{a_j a_m} = \sum_{i=1}^n \left\{ \sum_{k=1}^p (\mathbf{M}^{-1})_{mk} \frac{f_k}{\sigma_i^2} \right\} \left\{ \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} \frac{f_l}{\sigma_i^2} \right\} \sigma_i^2$$

Reordenando os somatórios

$$\sigma_{a_j a_m} = \sum_{k=1}^p \sum_{l=1}^p (\mathbf{M}^{-1})_{mk} (\mathbf{M}^{-1})_{jl} \sum_{i=1}^n \frac{f_k f_l}{\sigma_i^2}$$

e lembrando a definição da matriz \mathbf{M} (ver a Equação 34),

$$\begin{aligned} \sigma_{a_j a_m} &= \sum_{k=1}^p \sum_{l=1}^p (\mathbf{M}^{-1})_{mk} (\mathbf{M}^{-1})_{jl} (\mathbf{M})_{kl} \\ &= \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} \left\{ \sum_{k=1}^p (\mathbf{M}^{-1})_{mk} (\mathbf{M})_{kl} \right\} \end{aligned}$$

isto é,

$$\sigma_{a_j a_m} = \sum_{l=1}^p (\mathbf{M}^{-1})_{jl} \delta_{ml}$$

Assim, chegamos finalmente à expressão da covariância dos parâmetros ajustados:

$$\sigma_{a_j a_m} = (\mathbf{M}^{-1})_{jm} \quad (41)$$

As duas Equações (40) e (41), justificam o nome de **matriz das covariâncias** dado à matriz \mathbf{M}^{-1} . Seus elementos **diagonais** são as **variâncias** dos parâmetros e os **não diagonais** são as **covariâncias**.

A independência dos parâmetros implica covariâncias nulas, ou seja, uma matriz das covariâncias diagonal. Nesse caso, sua inversa, a matriz \mathbf{M} , é também diagonal. Voltando à definição dessa matriz (Equação

34), podemos concluir que a independência dos parâmetros é a consequência de uma escolha **judiciosa** das funções $f_j(x)$, pois essas funções devem satisfazer simultaneamente às seguintes condições:

$$\sum_{i=1}^n \frac{f_j(x_i)f_k(x_i)}{\sigma_i^2} = 0 \quad \text{para } (j \neq k)$$

- **Caso particular de incertezas iguais**

Quando as incertezas σ_i são todas iguais a σ , as expressões das matrizes **B** e **M** se simplificam:

$$\mathbf{B} = \frac{1}{\sigma^2} \begin{pmatrix} \sum_{i=1}^n y_i f_1 \\ \sum_{i=1}^n y_i f_2 \\ \sum_{i=1}^n y_i f_3 \\ \sum_{i=1}^n y_i f_4 \\ \dots \\ \dots \\ \dots \\ \sum_{i=1}^n y_i f_p \end{pmatrix} \quad (42)$$

$$\mathbf{M} = \frac{1}{\sigma^2} \begin{pmatrix} \sum_{i=1}^n f_1^2 & \sum_{i=1}^n f_1 f_2 & \dots & \sum_{i=1}^n f_1 f_p \\ \sum_{i=1}^n f_2 f_1 & \sum_{i=1}^n f_2^2 & \dots & \sum_{i=1}^n f_2 f_p \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^n f_p f_1 & \sum_{i=1}^n f_p f_2 & \dots & \sum_{i=1}^n f_p^2 \end{pmatrix} \quad (43)$$

A matriz das covariâncias \mathbf{M}^{-1} é, agora, proporcional a σ^2 e, portanto, a matriz \mathbf{A} dos parâmetros ajustados, solução da Equação matricial (36), não depende da incerteza experimental **única** σ . Entretanto, devemos notar que variâncias e covariâncias dos parâmetros, por serem os elementos da matriz das covariâncias, são proporcionais a σ^2 .

• Valor Médio e Variância do χ^2

Na Equação (30) que define o χ^2 , notamos a presença de n medidas e p parâmetros. Temos também um conjunto de p condições (31) impostas para minimizar o χ^2 . Define-se o **número de graus de liberdade** ν como

$$\nu = n - p \quad (44)$$

É possível mostrar que o χ^2 , com ν graus de liberdade e considerado como uma variável aleatória, é distribuído segundo uma função densidade de probabilidade cuja forma depende de ν . Também é possível demonstrar que o valor médio e a variância do χ^2 são dados, respectivamente, por

$$\overline{\chi^2} = \nu \quad e \quad \sigma_{\chi^2}^2 = 2\nu$$

Como será mostrado mais adiante, essas propriedades podem ser utilizadas para avaliar a qualidade de um ajuste pelo método dos mínimos quadrados.

No caso de incertezas iguais ($\sigma_i = \sigma$) e de um grande número n de medidas, o χ^2 pode ser aproximado por seu valor médio:

$$\frac{1}{\sigma^2} \sum_{i=1}^n \{y_i - f(x_i)\}^2 \simeq \nu$$

Assim, podemos estimar a variância σ^2 nos pontos experimentais, uma vez ajustada a função $f(x)$:

$$\sigma^2 \simeq \frac{1}{\nu} \sum_{i=1}^n \{y_i - f(x_i)\}^2 \quad (45)$$

Pois bem, vamos considerar agora uma categoria muito importante de funções lineares nos parâmetros, ou seja, polinômios de grau arbitrário.

• Regressão Polinomial

Uma regressão polinomial consiste em ajustar, pelo método do χ^2 , um polinômio de grau m a um conjunto de n pontos experimentais. Trata-se de um caso particular de ajuste de funções lineares nos parâmetros, pois podemos escrever:

$$f(x) = a_1 + a_2x + a_3x^2 + \dots + a_{m+1}x^m$$

Comparando este polinômio à função linear definida pela Equação (32), deduzimos que:

$$f_{l+1}(x) = x^l \quad \text{para } l = 0, \dots, m$$

Observamos que o número de parâmetros a serem ajustados é $m + 1$.

Transportando essas funções $f_l(x)$ nas Equações (33) e (34), os parâmetros ajustados serão dados pela matriz coluna **A**, solução da Equação matricial (36), sendo que as matrizes **B** e **M** escrevem-se agora:

$$\mathbf{B} = \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i^2 \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i^3 \\ \dots \\ \dots \\ \dots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i^m \end{pmatrix} \quad (46)$$

$$\mathbf{M} = \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^m \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^2 & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^m & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^{m+1} & \cdots & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^{2m} \end{pmatrix} \quad (47)$$

As variâncias e covariâncias dos parâmetros continuam sendo os elementos diagonais e não diagonais, respectivamente, da nova matriz das covariâncias \mathbf{M}^{-1} .

No caso particular de incertezas experimentais iguais, temos $\sigma_i = \sigma$ qualquer que seja $i = 1, \dots, n$. As matrizes \mathbf{B} e \mathbf{M} anteriores escrevem-se

$$\mathbf{B} = \frac{1}{\sigma^2} \begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n y_i x_i \\ \sum_{i=1}^n y_i x_i^2 \\ \sum_{i=1}^n y_i x_i^3 \\ \cdots \\ \cdots \\ \cdots \\ \sum_{i=1}^n y_i x_i^m \end{pmatrix} \quad (48)$$

$$\mathbf{M} = \frac{1}{\sigma^2} \begin{pmatrix} n & \sum_{i=1}^n x_i & \cdots & \sum_{i=1}^n x_i^m \\ \sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 & \cdots & \sum_{i=1}^n x_i^{m+1} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \sum_{i=1}^n x_i^m & \sum_{i=1}^n x_i^{m+1} & \cdots & \sum_{i=1}^n x_i^{2m} \end{pmatrix} \quad (49)$$

Invertendo essa matriz, obtemos a matriz das covariâncias e, portanto, as variâncias e covariâncias dos parâmetros ajustados.

Uma caso particular de regressão polinomial, entretanto muito importante, é o de regressão linear que consiste em ajustar um polinômio de primeiro grau a dados experimentais. Em outras palavras, procura-se a reta que melhor representa um conjunto de medidas $\{x_i, y_i\}$.

• Regressão Linear

Temos pela frente uma tarefa mais fácil, a de ajustar uma função linear $f(x; a_1, a_2) = a_1 f_1(x) + a_2 f_2(x)$, onde $f_1(x) = 1$ e $f_2(x) = x$.

As Equações (46) e (47) reduzem-se, agora, a:

$$\mathbf{B} = \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i \end{pmatrix} \quad (50)$$

$$\mathbf{M} = \begin{pmatrix} \sum_{i=1}^n \frac{1}{\sigma_i^2} & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i \\ \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i & \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^2 \end{pmatrix} \quad (51)$$

Para tornar as equações mais “leves”, introduzimos as notações seguintes:

$$S = \sum_{i=1}^n \frac{1}{\sigma_i^2}$$

$$S_x = \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i \quad S_{x^2} = \sum_{i=1}^n \frac{1}{\sigma_i^2} x_i^2 \quad (52)$$

$$S_y = \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i \quad S_{xy} = \sum_{i=1}^n \frac{1}{\sigma_i^2} y_i x_i$$

A matriz das covariâncias, inversa da matriz \mathbf{M} , escreve-se, então,

$$\mathbf{M}^{-1} = \frac{1}{\Delta} \begin{pmatrix} S_{x^2} & -S_x \\ -S_x & S \end{pmatrix} \quad (53)$$

onde

$$\Delta = S S_{x^2} - S_x^2 \quad (54)$$

é o determinante da matriz \mathbf{M} .

A regressão linear está concluída, pois encontramos:

os parâmetros a_1 e a_2 , soluções da Equação matricial (36)

$$a_1 = \frac{1}{\Delta} \{S_{x^2} S_y - S_x S_{xy}\}$$

$$a_2 = \frac{1}{\Delta} \{S S_{xy} - S_x S_y\} \quad (55)$$

as variâncias

$$\sigma_{a_1}^2 = \frac{S_{x^2}}{\Delta} \quad \sigma_{a_2}^2 = \frac{S}{\Delta} \quad (56)$$

e a covariância

$$\sigma_{a_1 a_2} = -\frac{S_x}{\Delta} \quad (57)$$

Vamos, agora, simplificar ainda mais nossa tarefa, estudando a regressão linear no caso em que as incertezas são iguais.

• Regressão Linear para Incertezas Iguais

Se $\sigma_1 = \sigma_2 = \sigma_3 = \dots = \sigma$, as somas S , S_x , S_{x^2} , S_y e S_{xy} , definidas na Equação (52), escrevem-se:

$$\begin{aligned} S &= \sigma^{-2} n \\ S_x &= \sigma^{-2} \sum_{i=1}^n x_i & S_{x^2} &= \sigma^{-2} \sum_{i=1}^n x_i^2 \\ S_y &= \sigma^{-2} \sum_{i=1}^n y_i & S_{xy} &= \sigma^{-2} \sum_{i=1}^n y_i x_i \end{aligned} \quad (58)$$

Portanto, introduzindo as notações

$$\begin{aligned} s &= \sigma^2 S & s_x &= \sigma^2 S_x & s_{x^2} &= \sigma^2 S_{x^2} \\ s_y &= \sigma^2 S_y & s_{xy} &= \sigma^2 S_{xy} \end{aligned} \quad (59)$$

temos $\Delta = \sigma^{-4} (n s_{x^2} - s_x^2) \equiv \sigma^{-4} \Delta'$ e obtemos as novas expressões dos parâmetros ajustados,

$$\begin{aligned} a_1 &= \frac{1}{\Delta'} \{s_{x^2} s_y - s_x s_{xy}\} \\ a_2 &= \frac{1}{\Delta'} \{n s_{xy} - s_x s_y\} \end{aligned} \quad (60)$$

das suas variâncias,

$$\sigma_{a_1}^2 = \frac{s_{x^2}}{\Delta'} \sigma^2 \quad \sigma_{a_2}^2 = \frac{n}{\Delta'} \sigma^2 \quad (61)$$

e da covariância,

$$\sigma_{a_1 a_2} = -\frac{s_x}{\Delta'} \sigma^2 \quad (62)$$

Podemos notar que os parâmetros ajustados **não dependem** da incerteza experimental σ , o que não ocorre para as variâncias e a covariância. Esta incerteza, mesmo sendo **a priori** desconhecida, pode ser determinada após a regressão, usando a relação (45):

$$\sigma^2 \simeq \frac{1}{n-2} \sum_{i=1}^n \{y_i - a_1 - a_2 x_i\}^2 \quad (63)$$

Nessa última expressão, os parâmetros ajustados são fornecidos pelas equações (60).

Só falta considerar o caso mais simples, isto é, o ajuste de uma reta passando pela origem.

• Regressão Linear com um só Parâmetro

Trata-se agora de ajustar uma reta $y = f(x) = a x$ a um conjunto de n pontos experimentais. Obviamente, as matrizes (33) e (34) se reduzem a números reais. O parâmetro ajustado e sua variância são, respectivamente:

$$a = \frac{S_{xy}}{S_{x^2}} \quad e \quad \sigma_a^2 = \frac{1}{S_{x^2}} \quad (64)$$

Se as incertezas forem iguais, obtemos, usando de novo as Equações (58) e (59):

$$a = \frac{s_{xy}}{s_{x^2}} \quad e \quad \sigma_a^2 = \frac{\sigma^2}{s_{x^2}} \quad (65)$$

E agora? Saber ajustar uma função a um conjunto de pontos experimentais é, sem dúvida, uma necessidade para qualquer experimentador. Entretanto, isso não é suficiente. É muito importante poder avaliar a qualidade do ajuste, tarefa nem sempre trivial, pois um excesso de subjetividade pode levar a conclusões errôneas. Por exemplo, muita gente fica feliz e satisfeita quando a regressão fornece um χ^2 muito pequeno. Infelizmente, como veremos em breve, essa felicidade só esconde uma grande ignorância! Vamos, então, apresentar critérios **objetivos** que permitem decidir se, após ajustar uma função, há ou não motivo para felicidade e satisfação do dever cumprido.

Como avaliar a qualidade de uma regressão?

De posse de dados experimentais e usando o método dos mínimos quadrados, **impõe-se** a forma $f(x; a_1, a_2, \dots, a_p)$ da função a ser ajustada, e a regressão fornece os valores dos parâmetros e suas variâncias e covariâncias. Se, por motivos diversos, não há dúvida alguma sobre a forma geral da função, o resultado do ajuste deve indicar se a experiência, isto é, a tomada de dados foi bem realizada e se as incertezas foram estimadas e/ou propagadas de maneira correta.

Uma consulta aos exemplos 3 e 4 do capítulo 13 e ao exemplo 1 do capítulo 14 do livro do Vuolo(96) é fortemente recomendada.

Entretanto, condições experimentais particulares e/ou dúvidas quanto à forma da função podem obrigar a **testar** a verossimilhança de vários tipos de função, sendo, portanto, importante dispor de critérios de avaliação da qualidade dos ajustes para poder escolher **a função verdadeira** e rejeitar as demais. Por exemplo, ao medir a velocidade de um corpo deslizando **sem atrito de escorregamento** sobre um plano inclinado, os dados experimentais podem sugerir uma dependência temporal linear ou quadrática, a dependência quadrática traduzindo a existência de um **atrito viscoso** devido à resistência do ar: o experimentador deve analisar suas regressões para avaliar se os dados revelam a influência desta resistência.

Antes de ir mais adiante, vamos descrever algumas situações experimentais, introduzindo, por comodidade, uma nova grandeza chamada “ χ^2 **reduzido**” igual, por definição, ao valor do χ^2 dividido pelo número de graus de liberdade ν da regressão (ver a Equação 44):

$$\chi_{red}^2 = \frac{\chi^2}{\nu} \quad (66)$$

Observe atentamente as Figuras 7, 8 e 9 a seguir que ilustram os resultados de experimentos realizados por três experimentadores, o João (JO), a Sandra (SA) e o Leandro (LE). Os pontos experimentais $\{x_i, y_i\}$ dessas figuras são idênticos, porém as incertezas do João são duas vezes maiores que as da Sandra, que são duas vezes maiores que as do Leandro. Cada experimentador ajustou uma reta aos seus dados, obtendo os χ_{red}^2 seguintes:

$$\chi_{red}^2(JO) = 0.36 \quad \chi_{red}^2(SA) = 1.44 \quad \chi_{red}^2(LE) = 5.76 \quad (67)$$

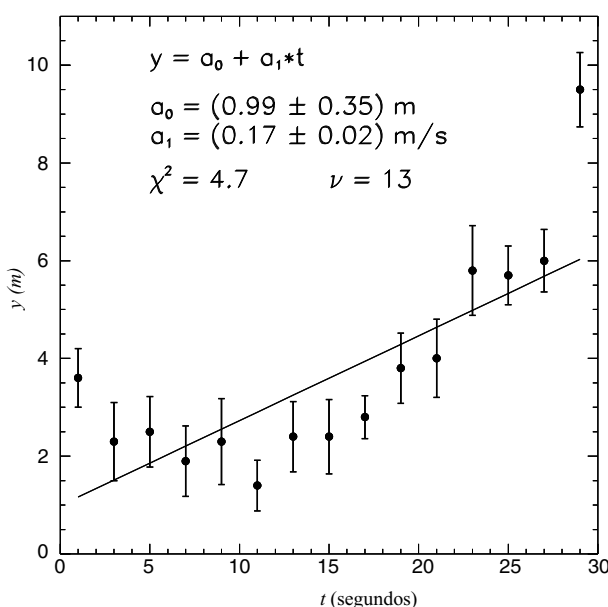


Figura 7: Ajuste linear realizado por João. A covariância entre os parâmetros é igual a $-0,006$, ou seja, não há correlação estatística entre os parâmetros.

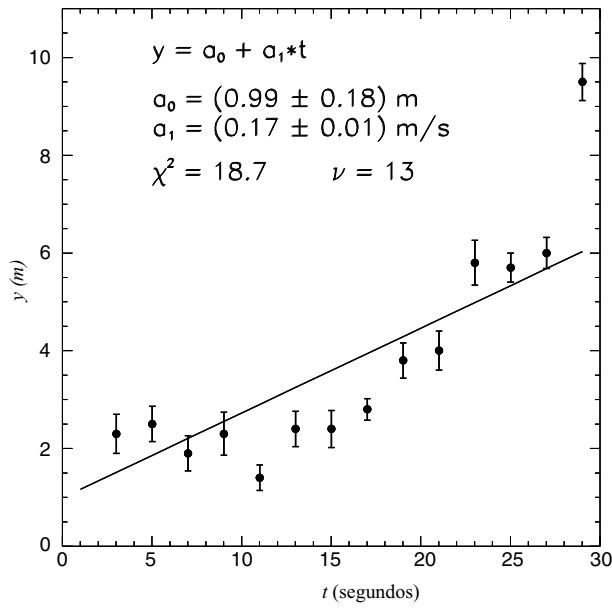


Figura 8: Ajuste linear realizado por Sandra. A covariância entre os parâmetros é igual a $-0,002$.

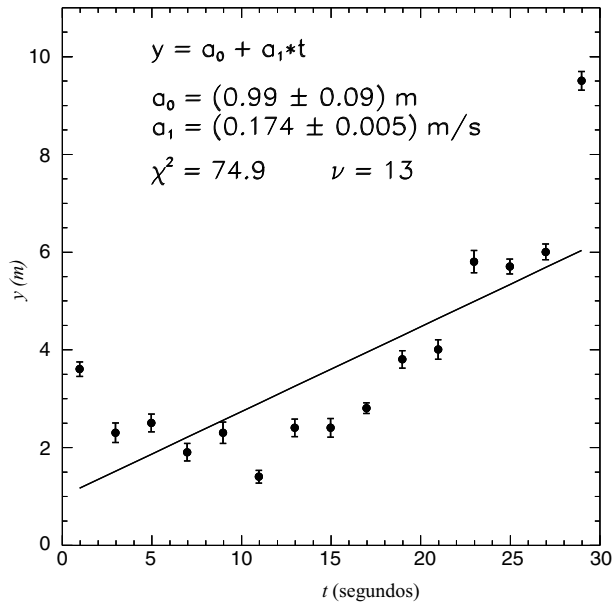


Figura 9: Ajuste linear do Leandro, com covariância igual a -0.0004 .

Estranhando a **discrepância visual** entre os pontos experimentais e as retas ajustadas, cada experimentador realizou uma regressão quadrática usando os mesmos dados. As curvas ajustadas aparecem nas Figuras 10, 11 e 12 e os novos valores de χ^2_{red} são:

$$\chi^2_{red}(JO) = 0.06 \quad \chi^2_{red}(SA) = 0.23 \quad \chi^2_{red}(LE) = 0.94 \quad (68)$$

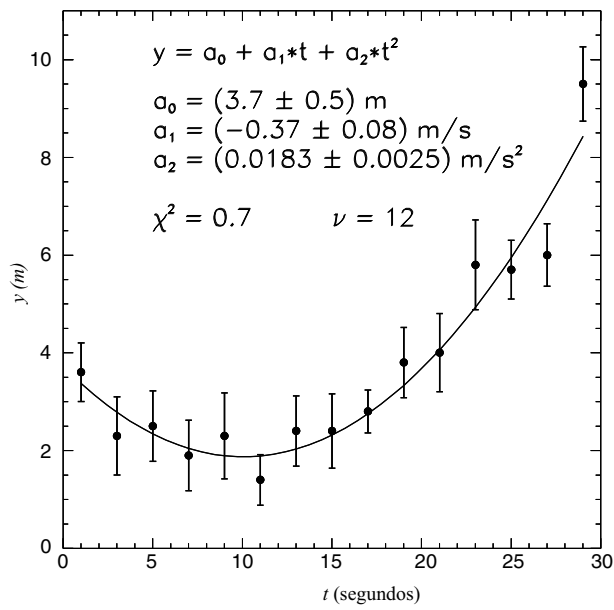


Figura 10: Ajuste quadrático realizado por João, com covariâncias muito pequenas, $\sigma_{a_0 a_1} = -0,035$, $\sigma_{a_0 a_2} = -0,001$ e $\sigma_{a_1 a_2} = -0,0002$.

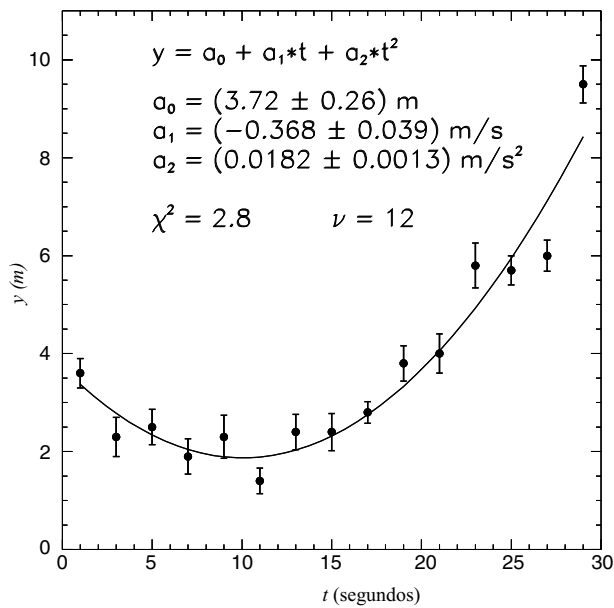


Figura 11: Ajuste quadrático realizado por Sandra. As covariâncias entre os parâmetros são $\sigma_{a_0 a_1} = -0,009$, $\sigma_{a_0 a_2} = -0,0002$ e $\sigma_{a_1 a_2} = -0,00005$.

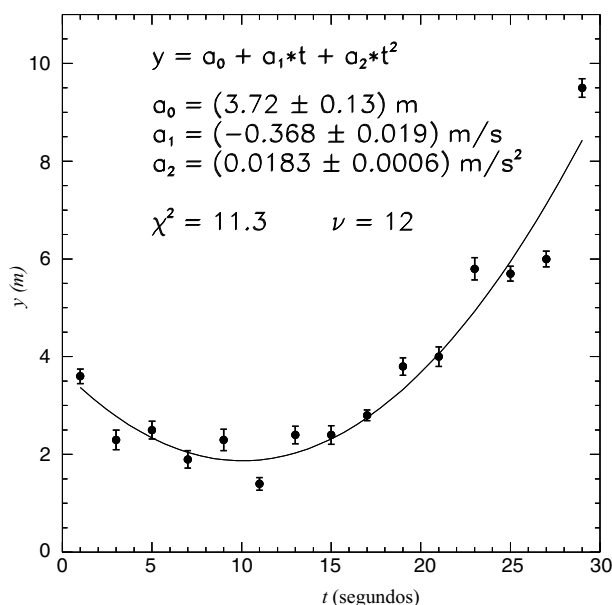


Figura 12: Ajuste quadrático realizado por Leandro. As covariâncias entre os parâmetros são $\sigma_{a_0 a_1} = -0,0005$, $\sigma_{a_0 a_2} = -0,00002$ e $\sigma_{a_1 a_2} = -0,00001$.

À luz desses resultados, duas perguntas surgem:

- Qual a função que melhor representa os dados, a linear ou a quadrática?
- Qual a melhor estimativa das incertezas experimentais: a do João, a da Sandra ou a do Leandro?

Infelizmente, você não tem como responder, de maneira segura e definitiva, a estas perguntas, e a única **conclusão tímida** é que, **a olho**, os pontos experimentais **sugerem** um comportamento quadrático. Precisamos, então, encontrar critérios objetivos para avaliar os resultados experimentais e sua adequação a modelos teóricos.

• Critério dos Cruzamento

Quando usamos o método dos mínimos quadrados, assumimos que as distribuições dos desvios $\Delta_i = |y_i - f(x_i)|$ são gaussianas ou, em outras palavras, normais.

Sendo assim, sabemos que cada um desses desvios tem uma probabilidade, de aproximadamente 0,68 (ou seja, 2/3), de ser menor que seu respectivo desvio padrão σ_i .

Veja o § 4.4 do Vuelo(96)

A princípio, isso significa que a curva ajustada deve cruzar, **em média**, $2/3$ das barras de erro. Entretanto, nem todos os n pontos experimentais devem ser considerados independentes, pois p entre eles - p sendo o número de parâmetros ajustados - são usados para definir a curva. Por exemplo, numa regressão linear com dois parâmetros, os dois pontos experimentais com **menor desvio** podem ser usados para traçar a reta **ajustada**.

Assim, é mais correto considerar que o número total **esperado** \overline{N}_{cr} de cruzamentos “**curva ajustada - barra de erro**” é, **em média**, igual ao número de cruzamentos **estatísticos**, $\frac{2}{3}\nu$, acrescentado do número de cruzamentos **certos**, p :

$$\overline{N}_{cr} = \frac{2}{3}\nu + p = \frac{2}{3}n + \frac{1}{3}p \quad (69)$$

onde $\nu = n - p$ é o número de graus de liberdade da regressão.

Podemos usar este critério, que chamaremos **critério dos cruzamentos**, para avaliar a verossimilhança das regressões lineares (Figuras 7, 8 e 9) e quadráticas (Figuras 10, 11 e 12).

Contando o número de cruzamentos em cada uma dessas figuras, obtemos os seguintes resultados:

$$N_{cr}^{(reta)}(JO) = 9 \quad N_{cr}^{(parab)}(JO) = 13$$

$$N_{cr}^{(reta)}(SA) = 3 \quad N_{cr}^{(parab)}(SA) = 9$$

$$N_{cr}^{(reta)}(LE) = 0 \quad N_{cr}^{(parab)}(LE) = 5$$

Se uma dependência linear $y(t) = a_0 + a_1 t$ for imposta por um modelo físico, usando a Equação (69), vemos que o número **esperado** de cruzamentos é $\overline{N}_{cr}^{(reta)} = 10.7$. Portanto, assumindo uma dependência linear, podemos dizer que o ajuste do João é bastante verossímil com 9 cruzamentos. Sandra e Leandro subestimaram suas incertezas e seus ajustes lineares são pouco verossímeis.

E se a dependência funcional for quadrática? Neste caso, o número **esperado** de cruzamentos $\overline{N}_{cr}^{(parab)}$ é igual a 11, o que torna os resultados da Sandra e do João os mais verossímeis.

Na ausência de um modelo, se cada experimentador confiar na estimativa das suas incertezas e usar o critério dos cruzamentos, as seguintes conclusões são **verossímeis**: para João, ambas as dependências linear e quadrática são verossímeis. A Sandra rejeita categoricamente a dependência linear mas acredita que subestimou um pouco suas barras de erro. Leandro concorda com Sandra, acreditando numa dependência quadrática, mas acha que suas incertezas foram bastante subestimadas.

Para resumir, podemos dizer que o critério dos cruzamentos, simples de ser utilizado, permite testar hipóteses (ou modelos) se as incertezas experimentais forem corretamente estimadas, ou estimar a verossimilhança da estimativa das incertezas se a dependência funcional for conhecida **a priori**.

Entretanto, existem outros critérios como o do Nível de Confiança do χ^2 reduzido, descrito a seguir.

- **Critério do Nível de Confiança do χ^2 reduzido**

Tendo realizado um ajuste pelo método do χ^2 , podemos reescrever a Equação (30), que define o χ^2 , introduzindo as distâncias normalizadas dos pontos experimentais à curva ajustada

$$D_i = \frac{y_i - f(x_i)}{\sigma_i}$$

O χ^2 é, então, igual à soma dos quadrados dessas distâncias:

$$\chi^2 = \sum_{i=1}^n D_i^2$$

Cada ponto experimental y_i flutua em torno do seu valor médio \bar{y}_i com um desvio padrão σ_i . As distâncias normalizadas D_i , o χ^2 e, obviamente, o χ^2 reduzido definido pela Equação (66), flutuam também.

É possível, embora não trivial, mostrar que a função de densidade de probabilidade da **variável aleatória** χ^2 é

$$h(\chi^2) = \frac{1}{\sqrt{2^\nu} \Gamma(\frac{\nu}{2})} \{\chi^2\}^{\frac{\nu-2}{2}} e^{-\frac{1}{2}\chi^2} \quad (70)$$

onde ν é o número de graus de liberdade do ajuste.

A função $\Gamma(q)$, definida pela equação

$$\Gamma(q) = \int_0^{\infty} e^{-z} z^{q-1} dz \quad (71)$$

pode ser também expressa por

$$\Gamma(q) = (q-1) \Gamma(q-1) = (q-1)!$$

$$\text{com} \quad \Gamma(1) = 1 \quad e \quad \Gamma(1/2) = \sqrt{\pi}$$

Assim, podemos mostrar que

$$\overline{\chi^2} = \nu \quad e \quad \overline{\chi_{red}^2} = 1 \quad (72)$$

Exercício 8

Demonstre os resultados 72, utilizando a definição do valor médio de uma variável aleatória **contínua**:

$$\overline{\chi^2} = \int_0^{\infty} \chi^2 h(\chi^2) d\chi^2$$

e as Equações 70 e 71.

Exercício 9

Resolvendo a Equação

$$\frac{dh(\chi^2)}{d\chi^2} = 0,$$

mostre que o valor mais provável do χ^2 para $\nu \geq 2$, é:

$$(\chi^2)_{mp} = \nu - 2$$

Exercício 10

Mostre que

$$(\chi_{red}^2)_{mp} = 1 - \frac{2}{\nu}$$

Exercício 11

Mostre que a função de densidade de probabilidade do χ^2 , definida pela Equação (70), diverge na origem para $\nu = 1$.

De acordo com a teoria das probabilidades, a probabilidade de que a variável aleatória χ^2 seja maior que o valor χ_0^2 obtido num ajuste com ν graus de liberdade é dada por

$$P(\chi^2 > \chi_0^2) = \int_{\chi_0^2}^{\infty} h(\chi^2) d\chi^2 \quad (73)$$

onde $h(\chi^2)$ é a função de densidade de probabilidade definida pela equação (70). Em outras palavras, $P(\chi^2 > \chi_0^2)$ mede a probabilidade de se ter um acordo entre dados experimentais e função ajustada pior que o obtido na regressão. Isso significa que, quanto maior for esta probabilidade, maior será o nível de confiança nos resultados.

Obviamente, essa probabilidade é igual à probabilidade de que o χ_{red}^2 seja maior que $\chi_{red(0)}^2 \equiv \chi_0^2/\nu$.

Tabelas e gráficos de $P(\chi^2 > \chi_0^2)$ podem ser encontrados na literatura da área.

O critério do nível de confiança consiste em definir um valor mínimo **arbitrário** da probabilidade $P(\chi^2 > \chi_0^2)$ para aceitar um ajuste pelo método dos mínimos quadrados. O valor mínimo geralmente adotado é de 5%.

As Tabelas 2 e 3 a seguir permitem analisar os resultados obtidos pelos nossos experimentadores Sandra, Leandro e João.

Veja por exemplo, Meyer(75) ou Helene(81).

Regressões Lineares	João	Sandra	Leandro
$\chi_{red(0)}^2$	0.36	1.44	5.76
$P(\chi_{red}^2 > \chi_{red(0)}^2)$	98%	18%	$\approx 10^{-7}\%$

Tabela 2: Níveis de confiança dos ajustes lineares.

Regressões Quadráticas	João	Sandra	Leandro
$\chi_{red(0)}^2$	0.06	0.23	0.94
$P(\chi_{red}^2 > \chi_{red(0)}^2)$	$\approx 100\%$	99%	61%

Tabela 3: Níveis de confiança dos ajustes quadráticos.

Pelo critério do nível de confiança, a estimativa das incertezas do Leandro torna seus dados experimentais definitivamente incompatíveis com a hipótese de uma dependência $y(t)$ linear. Em contrapartida, as incertezas do João parecem nitidamente superestimadas, pois os níveis de confiança obtidos são “bons demais”, tanto no caso da regressão linear quanto quadrática: não há como escolher uma hipótese rejeitando a outra! As barras de erro propostas pela Sandra favorecem a hipótese quadrática cujo nível de confiança é exageradamente alto, fato este que, de novo, sugere uma superestimativa das incertezas experimentais. Do ponto de vista estatístico, as incertezas do Leandro parecem mais realistas e sua regressão quadrática bastante verossímil.

Define-se também um outro critério, o do Intervalo de Confiança do χ^2 reduzido, preferido por muitos experimentadores e definido a seguir.

• Critério do Intervalo de Confiança do χ^2 reduzido

Usando duas vezes a probabilidade definida pela Equação (73), define-se a probabilidade de que o χ^2 esteja no intervalo $[\chi_0^2, \chi_1^2]$

$$P(\chi_0^2 < \chi^2 < \chi_1^2) = \int_{\chi_0^2}^{\chi_1^2} h(\chi^2) d\chi^2 \quad (74)$$

pois tem-se

$$\int_{\chi_0^2}^{\chi_1^2} h(\chi^2) d\chi^2 = \int_{\chi_0^2}^{\infty} h(\chi^2) d\chi^2 - \int_{\chi_1^2}^{\infty} h(\chi^2) d\chi^2 \quad (75)$$

Evidentemente, esta probabilidade é igual à probabilidade de que o χ_{red}^2 esteja no intervalo $[\chi_{red(0)}^2, \chi_{red(1)}^2]$. Por exemplo, se $\chi_{red(0)}^2$ e $\chi_{red(1)}^2$ correspondem aos **Níveis de Confiança** de 95 % e 5 %, respectivamente, o **Intervalo de Confiança** correspondente será de 90 %.

Uma maneira muito prática de avaliar a qualidade de uma regressão é a de utilizar figuras como a Figura (13) na qual o χ^2 reduzido aparece como função do número de graus de liberdade ν , em nível de confiança constante.

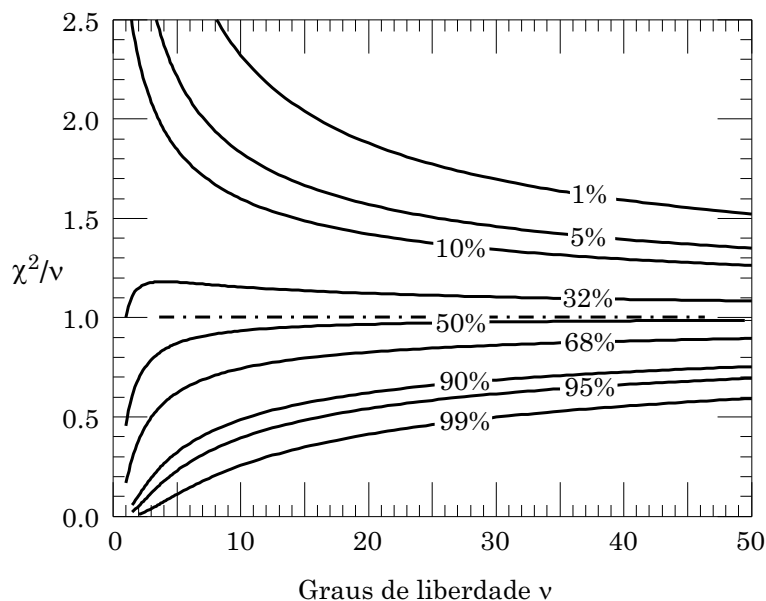


Figura 13: χ^2 reduzido como função do número de graus de liberdade para níveis de confiança entre 1% e 99%.

Assim, com 90% de confiança, verificamos que

$$0.45 < \chi_{red}^2 < 1.72 \quad \text{para } \nu = 13$$

$$0.43 < \chi_{red}^2 < 1.75 \quad \text{para } \nu = 12$$

Utilizando o **Critério do Intervalo de Confiança**, podemos analisar de novo os resultados dos ajustes do João, da Sandra e do Leandro: com 90 % de confiança, rejeitamos as duas regressões do João, a regressão linear do Leandro e a quadrática da Sandra. Sobram então a regressão linear da Sandra e a quadrática do Leandro. Enquanto o χ^2 reduzido da regressão linear da Sandra encontra-se perto do limite superior do intervalo de confiança de 90% para 13 graus de liberdade, o da regressão quadrática do Leandro está no meio desse mesmo intervalo para 12 graus de liberdade. Assim, chegamos de novo à conclusão de que a hipótese mais verossímil é a quadrática, e que a melhor estimativa de erros experimentais é a do Leandro.

Para concluir, podemos dizer que avaliar a qualidade de uma regressão significa:

- verificar a verossimilhança da função ajustada, **desde que as incertezas sejam corretamente estimadas e/ou calculadas**, permitindo, assim, escolher entre várias hipóteses físicas, ou
- verificar se as incertezas foram corretamente estimadas e/ou calculadas, **desde que não haja dúvida sobre a forma da função ajustada**.

Importância das Barras de Erro

Considere as duas regressões lineares realizadas pela Tatiana (Figura 14) e pela Érica (Figura 15).

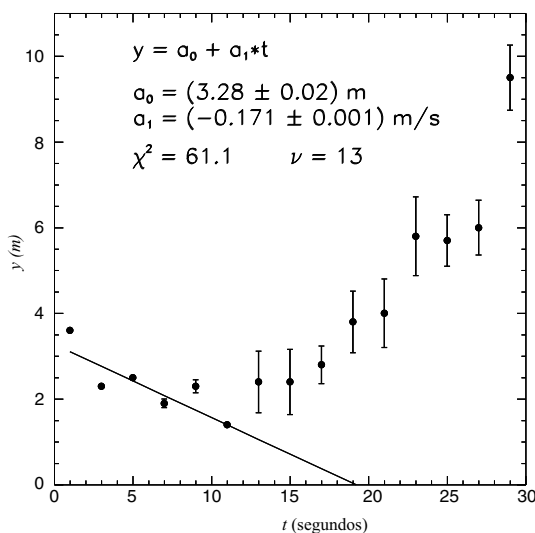


Figura 14: Ajuste linear feito por Tatiana.

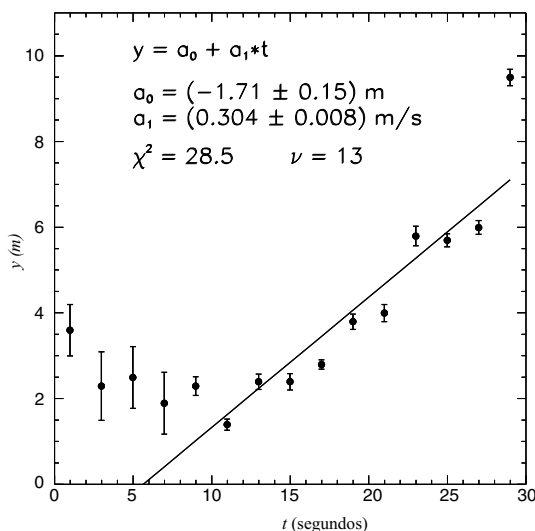


Figura 15: Ajuste linear feito por Érica.

Os pontos experimentais são aqueles já utilizados por João, Sandra e Leandro (ver Figuras 7 a 12) mas as incertezas foram estimadas de maneira **estranha, porém didaticamente interessante**. Em particular, na regressão da Tatiana, a barra de erro do sexto ponto experimental, invisível na figura, é igual a 0.00001 metro, o que confere a este ponto um peso enorme no cálculo do χ^2 (ver a Equação 30). Não há dúvida quanto à importância das incertezas, pois os resultados das regressões, definitivamente inverossímeis, são completamente diferentes, as retas tendendo a se ajustar aos pontos de menor incerteza.

Para terminar este texto complementar, vamos indicar, **sem demonstração**, em que circunstâncias, medidas podem ser descartadas de um conjunto de pontos experimentais, **antes** de estimar médias, variâncias etc ou de realizar uma regressão.

Critério de Rejeição de Chauvenet

Sendo um conjunto de n medidas y_i com média de conjunto \bar{y} e desvio padrão σ , o critério de rejeição de Chauvenet estipula que, se um desvio $d_i = |y_i - \bar{y}| > d_{rc}$, onde d_{rc} é o limite de rejeição de Chauvenet, então a medida y_i deve ser descartada do conjunto. A Figura 16 permite determinar d_{rc} em função da população n do conjunto.

Veja Pugh(66).

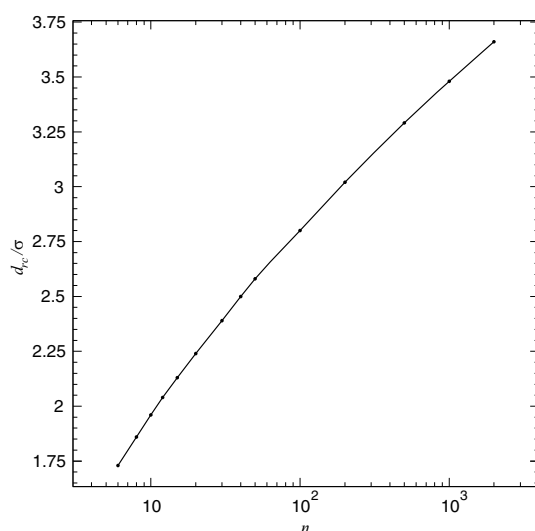


Figura 16: Limite de rejeição normalizado de Chauvenet como função do número de medidas.

Conclusão

Você leu atentamente este texto? Consultou as referências sugeridas? Tentou resolver os exercícios propostos? Muito bem! Com certeza, isso não é suficiente para fazer de você um *expert*, mas acreditamos que deve estar pronto para tratar corretamente dados experimentais e concluir suas experiências.

BOM TRABALHO ...E NÃO TENHA MEDO DE PEDIR
AJUDA!

Agradecimentos

– Aos estudantes do Instituto de Física da UFRJ, Tatiana da Silva, Cristina Schoch Vianna e Kazuyoshi Akiba que ajudaram a simular experiências e a realizar regressões.

– À minha colega, Professora Thereza Cristina de Lacerda Paiva, que teve a coragem de fazer a primeira revisão do texto.

– Ao Fabio Muniz, que produziu as figuras.

– À Anna Maria Osborne e Ana Tereza de Andrade, que revisaram o texto, ajudando a torná-lo mais atraente.

– Às diagramadoras Mirelle Mota e Aline Brondani, que deixaram de almoçar mas produziram, com gentileza e competência, a versão final do texto.

Que todos recebam aqui meus agradecimentos.

ISBN 85-7648-026-3



9 788576 480266



UENF
Universidade Estadual
do Norte Fluminense



Universidade Federal Fluminense



**GOVERNO DO
Rio de Janeiro**

SECRETARIA DE
CIÊNCIA E TECNOLOGIA



**Ministério
da Educação**

