

# Interpretação gráfica de dados

---

Este texto foi baseado nas apostilas “Introdução à interpretação gráfica de dados, gráficos e equações”, 1990, dos Profs. Fuad Saad, Paulo Yamamura e Kazuo Watanabe; “Física Geral e Experimental para Engenharia I”, 2003, dos Profs. Ewout ter Haar e Valdir Bindilati.

## 1. Introdução

Nas atividades experimentais, muitas vezes, objetiva-se estudar a maneira como uma propriedade, ou quantidade, varia com relação a uma outra quantidade, por exemplo:

*“De que modo o comprimento de um pêndulo afeta o seu período?”*

ou ainda:

*“Como se comporta a força de atrito entre duas superfícies relativamente à força normal exercida por uma superfície sobre a outra?”*

Tais questões podem ser estudadas e mais bem respondidas, muitas vezes, através de métodos gráficos evidenciando, dessa forma, a dependência de uma grandeza em relação à outra. Neste capítulo apresentaremos os principais tipos de gráficos disponíveis bem como técnicas para a sua confecção. Apresentaremos também alguns métodos de análise gráfica de dados de forma a poder extrair informações e interpretar resultados experimentais.

## 2. Tipos de gráficos

Os gráficos, de modo geral, podem ser classificados em cinco tipos básicos, conforme o esquema apresentado na figura 2.1. Dependendo do tipo de análise a ser realizada um tipo de gráfico torna-se mais adequado que outro. Nos trabalhos experimentais em Ciências são frequentemente utilizados gráficos do tipo diagrama, ou linha, conforme o apresentado na figura 2.2. Nesse gráfico é mostrado o comportamento de uma grandeza física, nesse caso a velocidade de um corpo, em função do tempo. Pode-se perceber facilmente que a velocidade aumenta com o passar do tempo. A grande vantagem de análises gráficas é a interpretação direta e fácil de dados

experimentais. A linha tracejada, nesse caso, representa o comportamento médio dos dados obtidos e representa a tendência dos dados.

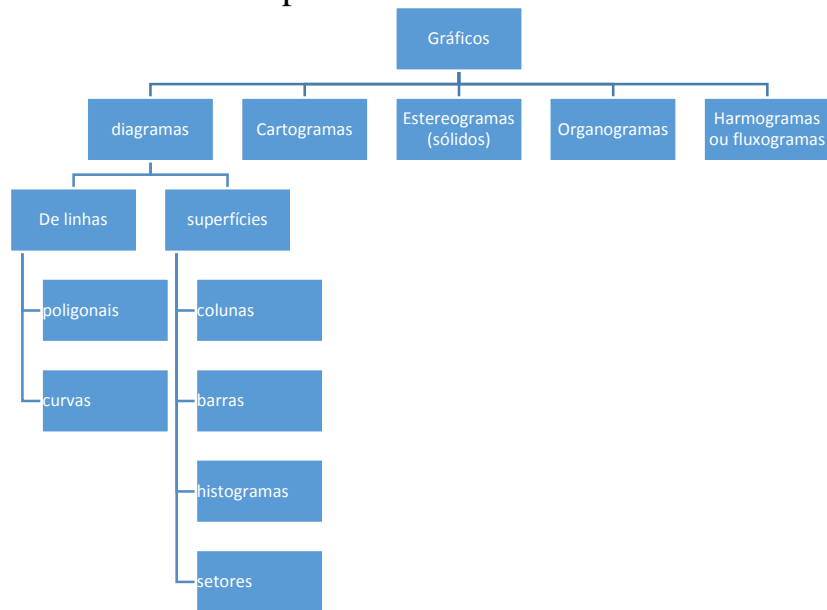


Figura 2.1: Principais tipos de gráficos

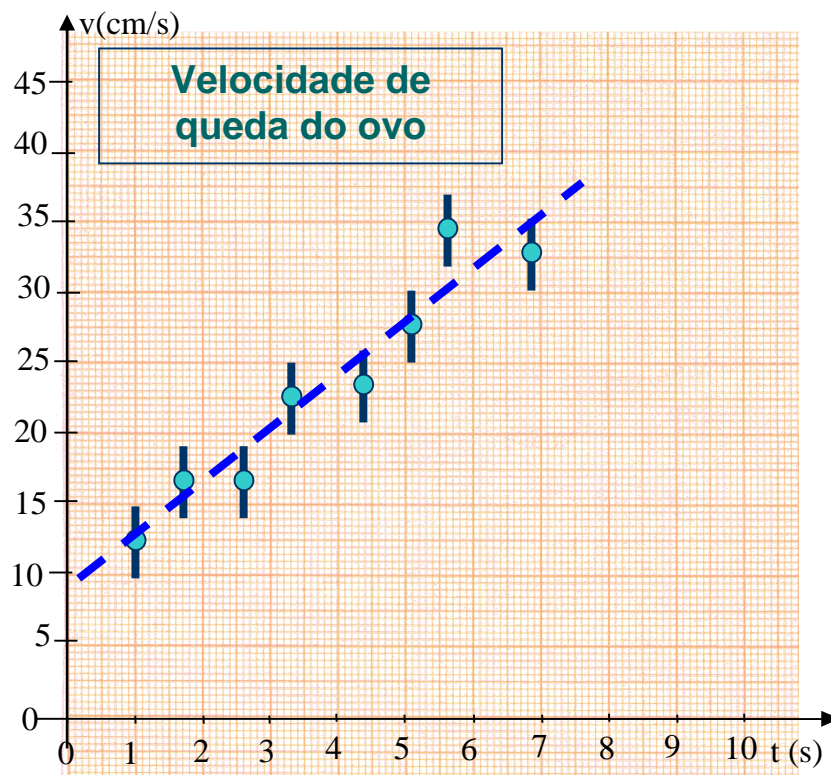


Figura 2.2: Exemplo de gráfico linear. Nesse gráfico, os pontos correspondem às medidas experimentais e a linha representa o comportamento médio.

### 3. Confeção de gráficos

Quando são realizados experimentos, os dados são adquiridos, geralmente, de dois modos:

No primeiro modo, quer-se examinar a dependência de uma grandeza em relação à outra, como, por exemplo, os dados apresentados na figura 2.2. Nesse caso, mede-se a velocidade do corpo em instantes consecutivos de tempo e analisa-se como a velocidade depende do tempo. Em medidas desse tipo, costuma-se denominar de variável independente aquela que se varia, nesse caso, o tempo. A grandeza na qual se quer estudar a dependência, nesse caso a velocidade, é denominada de variável dependente.

No segundo caso, o mesmo experimento é repetido muitas vezes nas mesmas condições e, em cada um desses experimentos, repete-se a medida de uma determinada grandeza. Nesse caso, querem-se estudar as variações de medidas devido às incertezas experimentais. Um caso típico é a medida do período de oscilação de um pêndulo simples. Dependendo dos instrumentos utilizados, a medida simples de um único período resulta, geralmente, em incertezas experimentais elevadas que podem ser minimizadas através da repetição do experimento muitas vezes. Assim, a medida final seria a média aritmética de todas as medidas efetuadas.

Em ambas as situações costuma-se organizar os dados em tabelas. Essas tabelas podem-se tornar demasiadamente longas e de difícil leitura. A representação desses dados em forma gráfica mostra, de forma mais clara, as propriedades das grandezas medidas. O gráfico mostra, igualmente, prováveis erros experimentais e permite realizar interpolações e extrapolações de modo visível e fácil.

No primeiro exemplo pode-se visualizar graficamente o comportamento da velocidade em função do tempo através de um gráfico de linhas. No segundo caso, contudo, a melhor visualização gráfica é feita através de um histograma. Nesse tipo de gráfico é muito simples obter grandezas como média e desvio padrão das medidas.

Antes de abordar os tipos de gráfico acima, devemos estabelecer algumas regras gerais de confeção de gráficos. Essas regras se aplicam a quase todos os tipos disponíveis.

### **3.1. Regras gerais para confecção de gráficos**

A construção de gráficos, quando feita sob regras universais, facilita significativamente a sua interpretação. Nesse sentido, regras rígidas (como regras de sintaxe de uma linguagem qualquer) são adotadas no mundo científico e tecnológico<sup>1</sup>.

Todo gráfico é composto dos seguintes itens:

1. Título e legenda do gráfico;
2. Eixos das variáveis com os nomes das variáveis, escalas e unidades;
3. Dados experimentais e incertezas;
4. Funções teóricas ou curvas médias (esse último item é opcional e, dependendo das circunstâncias, pode ser omitido);

A figura 3.1 mostra os principais componentes de um gráfico.

#### ***Título e legenda do gráfico***

Todo gráfico deve ter um título. Geralmente, o título do gráfico é colocado na parte superior do gráfico, em destaque. Títulos do tipo “gráfico de velocidade vs. tempo” são redundantes e não fornecem informação necessária para o entendimento do mesmo.

Caso o gráfico seja inserido dentro de um texto, o mesmo deve ser acompanhado de uma legenda, logo abaixo do gráfico, numerada, que explique de forma sucinta o seu conteúdo. No caso da presença de uma legenda, o título do gráfico torna-se opcional, já que a legenda acaba suprimindo o leitor de informação suficiente para o entendimento do gráfico.

#### ***Eixos, escalas e unidades***

Os eixos de um gráfico devem ser explicitamente desenhados. Cada um dos eixos deve conter o nome (ou símbolo) da variável representada, a escala de leitura e a unidade correspondente.

A escolha da escala utilizada deve ser tal que represente bem o intervalo medido para a variável correspondente. A regra prática para definir a escala a ser utilizada consiste em dividir a faixa de variação da variável a

---

<sup>1</sup> Programas computacionais de geração de gráficos não destinados à área científica, como o Excel, são muito limitados e possuem várias falhas no que diz respeito à confecção correta de gráficos e o seu uso é fortemente desaconselhado no mundo científico e tecnológico.

ser graficada pelo número de divisões principais disponíveis. Toma-se, então, um arredondamento para um valor superior e de fácil leitura. Esses valores são, em geral, 1, 2, 5 ou múltiplos/sub-múltiplos de 10 desses valores (10; 20; 500; 0,5; etc.). A figura 3.2 mostra alguns exemplos de escalas do eixo de um gráfico. Múltiplos de 3 são de difícil leitura e devem ser evitados.

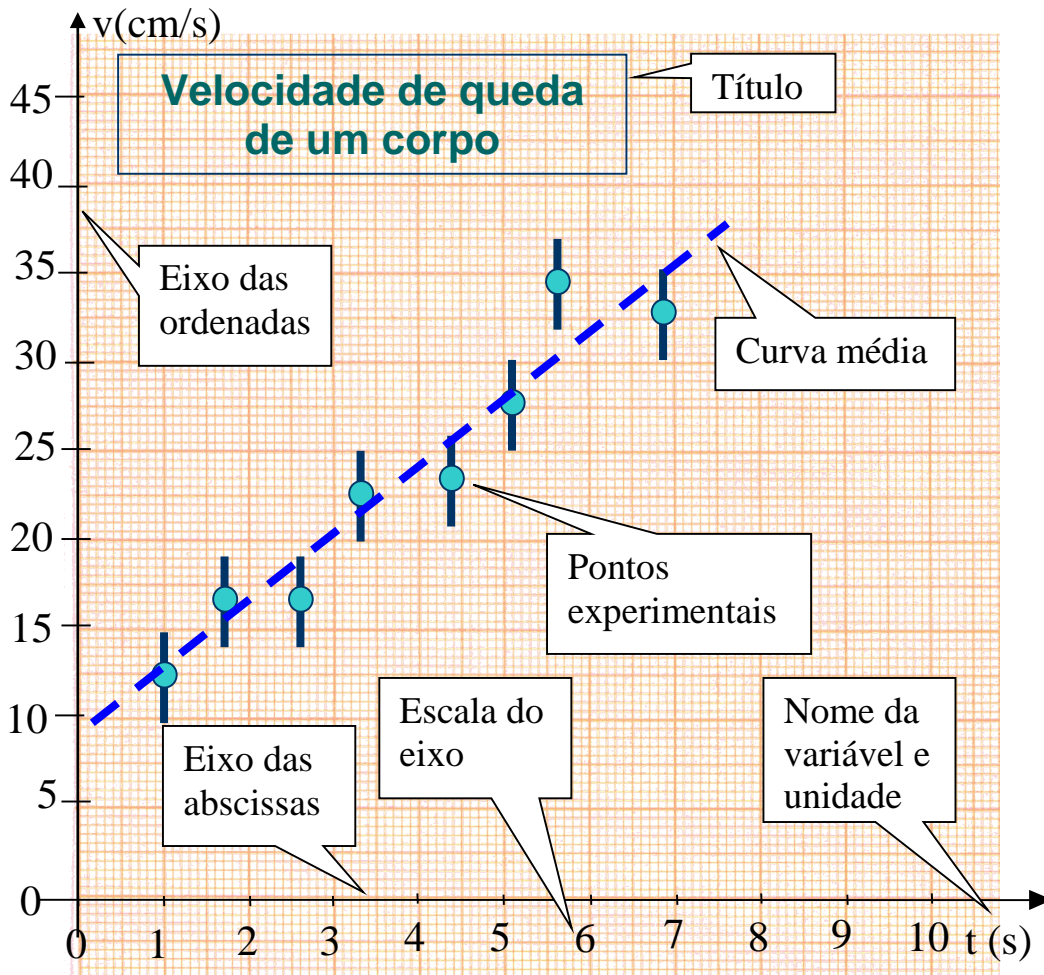


Figura 3.1. Componentes típicos de um gráfico científico padrão.

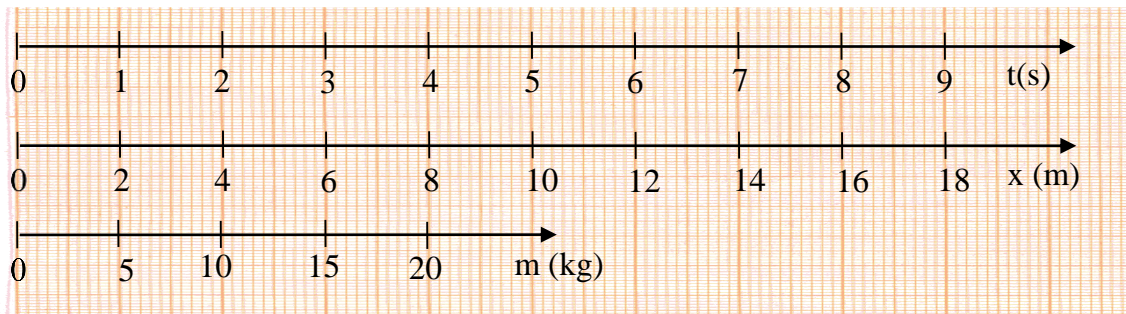


Figura 3.2. Alguns exemplos de formas CORRETAS de desenhar eixos em um gráfico.

As escalas de um gráfico não precisam começar na origem (0, 0). Elas devem abranger a faixa de variação que você quer representar. É conveniente que os limites da escala correspondam a um número inteiro de divisões principais. Indique os valores correspondentes às divisões principais abaixo (eixo-x) ou ao lado (eixo-y) da escala utilizando números legíveis. As unidades devem ser escolhidas de maneira a minimizar o número de dígitos utilizados na divisão principal (ver a terceira escala, de cima para baixo, na figura 3.2. Nesse caso, utilizou-se a escala de quilo-grama). Uma regra prática é utilizar no máximo 3 dígitos para representar esses valores. Pode-se também fazer o uso de potências de 10 na expressão das unidades para simplificar a escala.

Ao traçar os eixos em papel gráfico comum, não use a escala marcada no papel pelo fabricante. Você é quem define a escala. Também evite usar os eixos nas margens do papel. Desenhe os seus próprios eixos. Na figura 3.3 são mostradas algumas formas INCORRETAS de desenhar eixos de gráfico. Um erro muito comum é colocar nos eixos os valores medidos para cada variável. Esse é um erro MUITO grosseiro que torna o gráfico ilegível.

Por fim, escreva o nome (ou símbolo) da variável correspondente ao eixo e a unidade para leitura dos valores entre parêntesis (s, kg,  $10^5$  N/m<sup>2</sup>, etc.). No final das contas, o melhor critério para desenhar um eixo de um gráfico é o bom-senso. O teste final para saber se o eixo utilizado é adequado é a escolha aleatória de um ponto qualquer. O leitor deve ser capaz de identificar rapidamente o valor correspondente desse ponto através da leitura do eixo no gráfico.

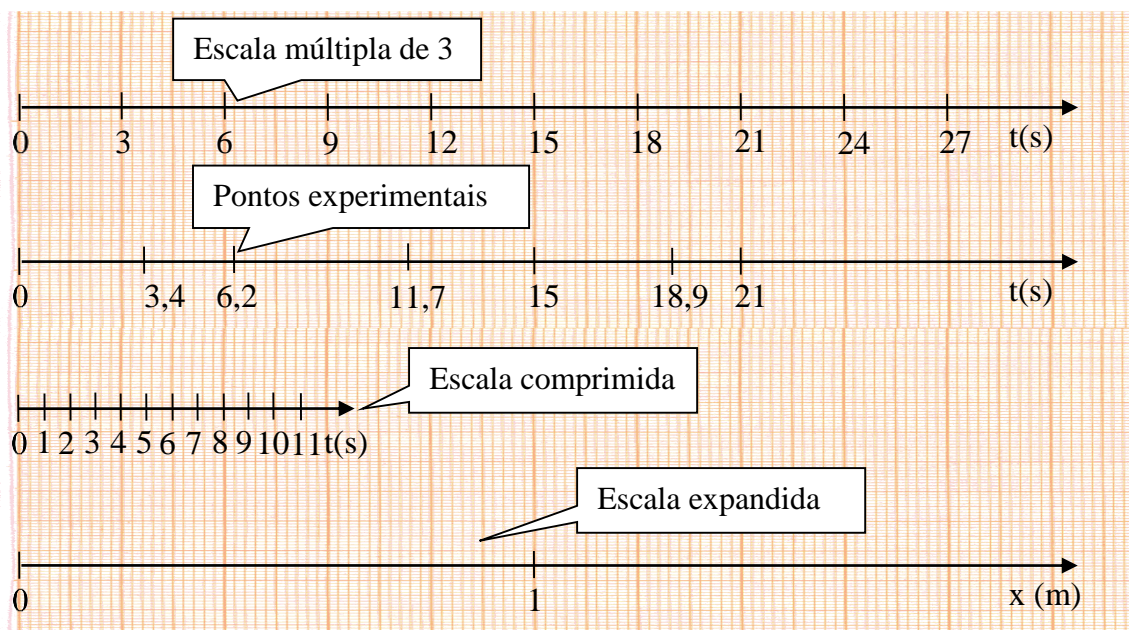


Figura 3.3. Algumas formas INCORRETAS de desenhar eixo em um gráfico.

### *Dados, funções teóricas e curvas médias*

Assinale no gráfico a posição dos pontos experimentais: use marcas bem visíveis (em geral círculos cheios). NUNCA indique as coordenadas dos pontos graficados no eixo. Coloque as barras de incerteza nos pontos, se for o caso. Se as incertezas são menores que o tamanho dos pontos, indique isso na legenda.

NUNCA LIGUE OS PONTOS. Esse é um erro grosseiro de confecção de gráficos, muito utilizado em programas de computadores. A figura 3.4 mostra como desenhar os pontos experimentais em um gráfico.

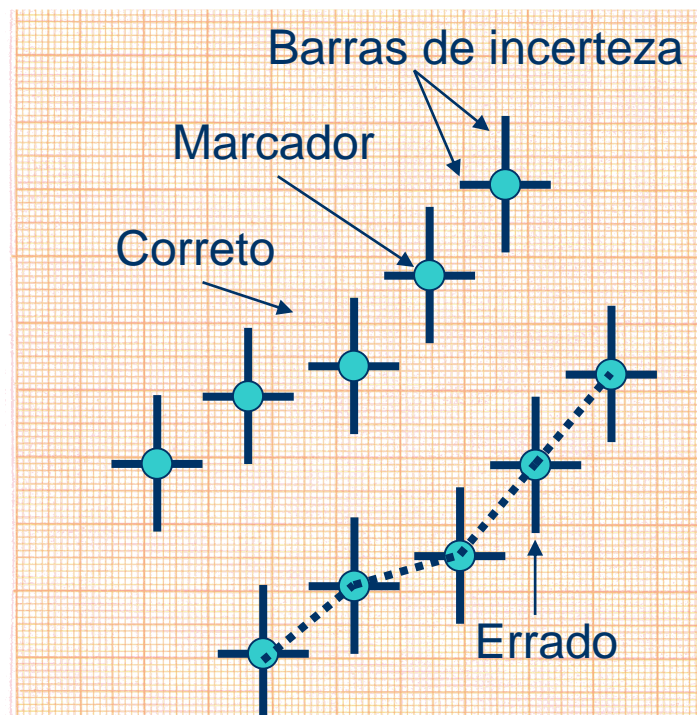


Figura 3.4. Representação de pontos experimentais em um gráfico. NUNCA LIGUE OS PONTOS. Indique as barras de incerteza (se for o caso) em cada ponto nos eixos x e y.

Às vezes, dependendo da análise a ser realizada com os dados, é necessário o desenho de curvas médias ou funções teóricas. Essas curvas têm como utilidade permitir a extrapolação e/ou interpolação de pontos, bem como a comparação entre os dados experimentais e uma previsão teórica. Esse ponto será discutido em detalhes adiante.

## **4. Gráficos de linhas**

Gráficos de linhas são normalmente utilizados para representar a dependência de uma grandeza em relação à outra, como o gráfico apresentado na figura 2.2 que mostra a dependência com o tempo da

velocidade de queda de um ovo. São muitos os tipos de gráficos de linhas que podem ser construídos. Dentre os vários se destacam três tipos comumente utilizados, conforme representado na figura 4.1.

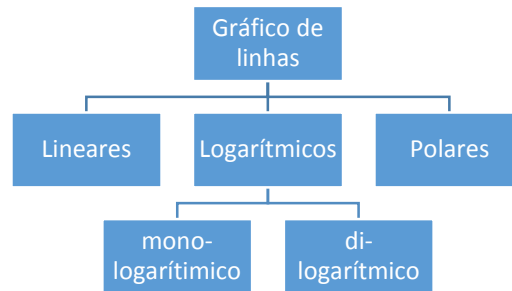


Figura 4.1. Principais tipos de gráficos de linhas utilizados no meio científico.

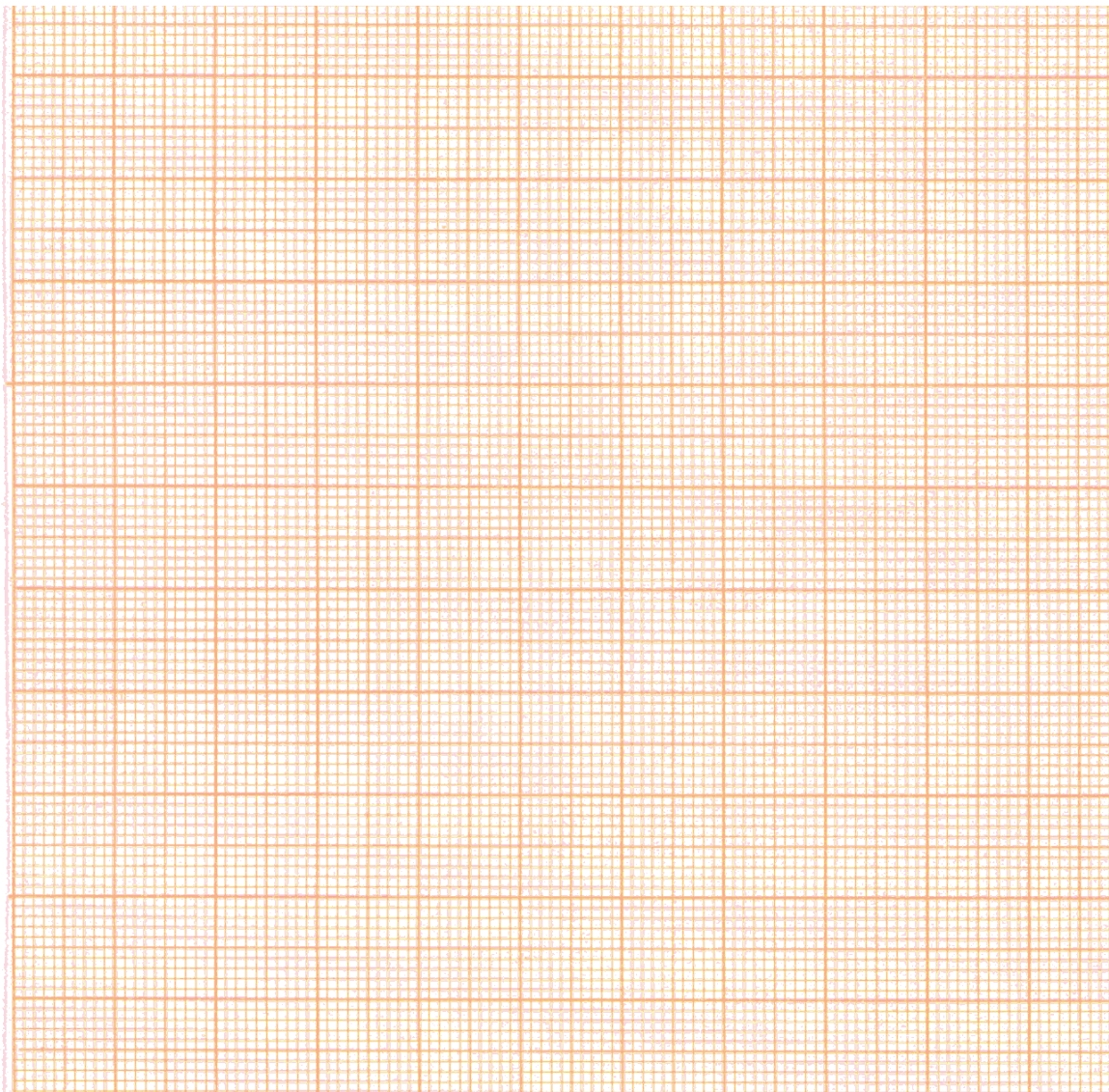


Figura 4.2. Papel em escala milimetrada. Nesse caso, ambas coordenadas são igualmente espaçadas em centímetros.



A escolha do tipo de gráfico está relacionada com os objetivos que se pretende alcançar. Um dos fatores que pode fornecer a ajuda na escolha é analisar a variação dos dados adquiridos. Por exemplo, uma grandeza que varia entre 10 Hz e 100 kHz (100000 Hz) torna-se impossível de ser graficada de forma eficiente em um gráfico linear, devido à grande variação entre um extremo e outro. Nesse caso, gráficos logarítmicos são mais adequados para representar dados desse tipo.

#### 4.1. Escalas lineares

Gráficos em escalas lineares são os mais simples de serem realizados. Como o próprio nome diz, gráficos em escalas lineares são aqueles nos quais ambos os eixos (x e y) são lineares, ou seja, a escala representada no eixo é diretamente proporcional à distância do ponto em relação à origem do eixo.

Gráficos em escalas lineares são desenhados normalmente em papéis milimetrados, conforme mostra a figura 4.2. Você pode usar a figura 4.2 como modelo para gráficos lineares. Basta fazer cópias xérox da figura e utilizar para os seus gráficos.

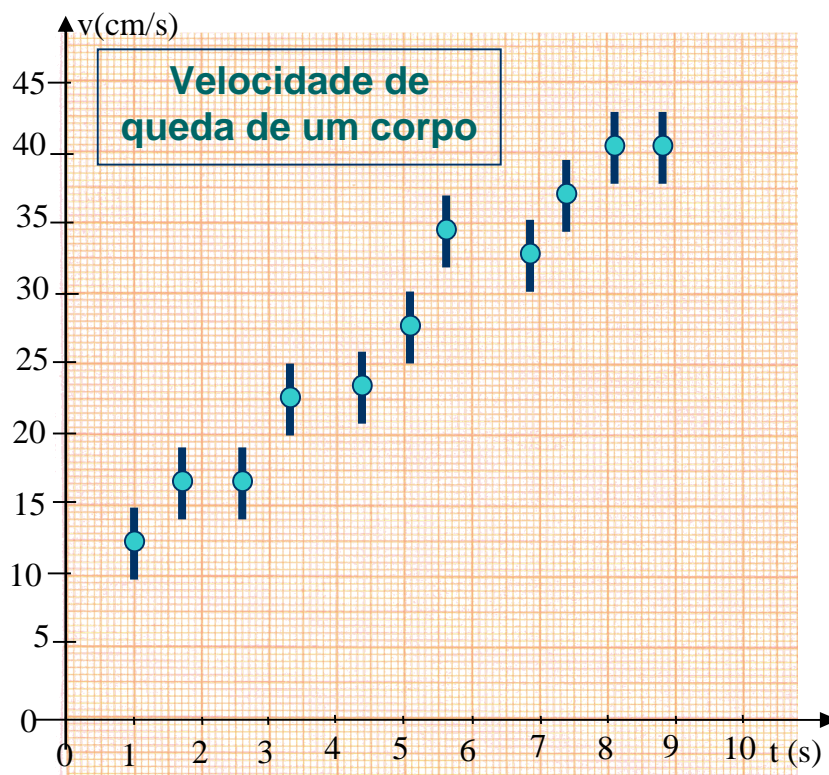


Figura 4.3. Velocidade de queda de um ovo.

Um exemplo de gráfico em escala linear é mostrado na figura 4.3. Nesse caso, grafica-se a velocidade instantânea de queda de um ovo como função do tempo de queda.

### *Traçando curvas médias*

Muitas vezes quer-se extrair informações mais complexas de um gráfico. Poderíamos perguntar, por exemplo, utilizando o gráfico da figura 4.3, qual seria a velocidade do ovo no instante 15 segundos, caso o tipo de movimento não se altere? Qual é a velocidade inicial de queda desse ovo e qual a sua aceleração média? Perguntas como essas podem ser respondidas combinando-se o conhecimento adquirido de Física com algumas técnicas de análise gráfica.

Existem técnicas matemáticas e testes sofisticados<sup>2</sup> para determinar o comportamento de dados e permitir extrapolações e interpolações. O aprendizado dessas técnicas foge ao escopo desta disciplina introdutória. Contudo, o método descrito a seguir pode, se executado de forma criteriosa, fornecer resultados muito próximos daqueles obtidos a partir de métodos matemáticos rigorosos.

De modo geral, pode-se desenhar curvas médias sobre conjunto de dados utilizando-se a curva francesa (ver figuras 4.4 e 4.5). O uso de curva francesa exige prática, porém pode-se conseguir resultados bastante satisfatórios.



Figura 4.4. Alguns exemplos de curva francesa. A curva francesa é comumente utilizada para traçar curvas médias de gráficos científicos.

---

<sup>2</sup> Para mais detalhes ver o livro “Fundamentos da Teoria de Erros”, José Henrique Vuolo, Editora Edgard Blücher Ltda.

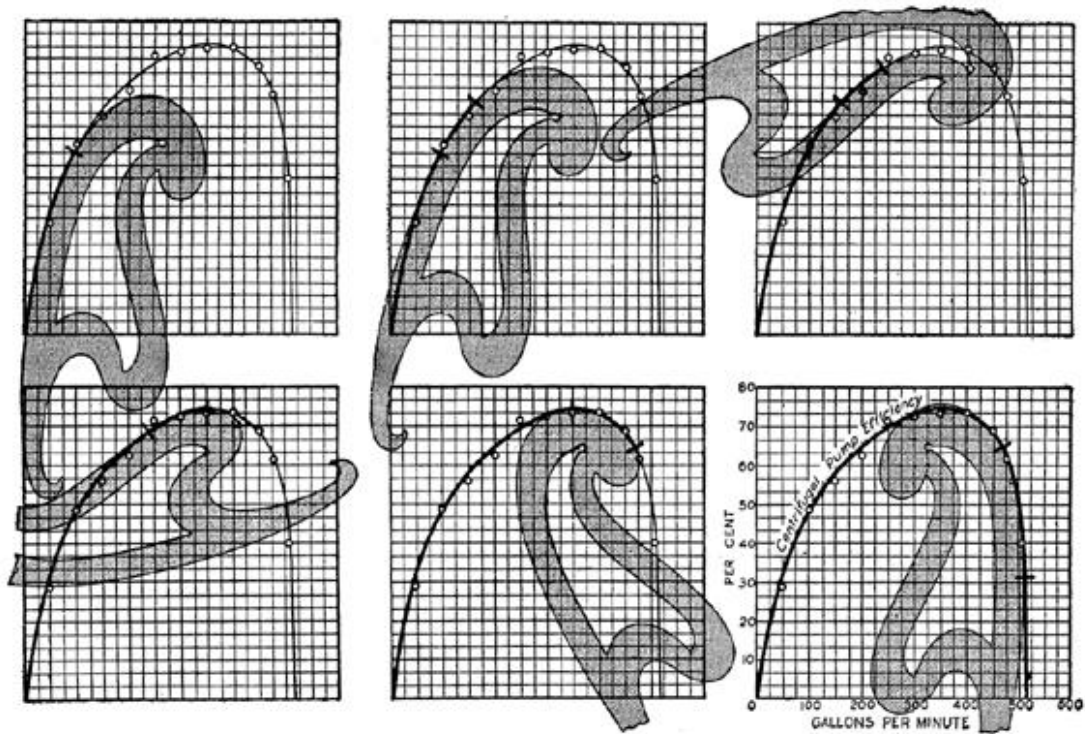


Figura 4.5. Exemplo da utilização da curva francesa para traçar uma curva média em um gráfico científico.

Um inconveniente do uso geral de curvas francesas é o fato de, apesar das curvas médias serem bastante satisfatórias, é difícil obter informações numéricas de forma direta. Além disso, pelo fato da curva obtida ser um guia visual, extrapolações para valores fora do intervalo onde os dados foram medidos são muito imprecisas e não devem ser feitas.

Contudo, existe um caso particular onde o traçado de curvas médias fornece várias informações sobre os dados graficados. Isso ocorre quando o gráfico entre duas grandezas pode ser representado por uma reta. Assim, a curva média obtida é uma reta, que pode ser desenhada utilizando-se uma régua simples.

Vamos re-examinar os dados na figura 4.3. Percebe-se que a dependência entre velocidade e tempo ocorre de forma mais ou menos linear (lembre-se de considerar as incertezas dos pontos experimentais). Para traçar uma reta média, nesse caso, deve-se utilizar uma régua e a reta desenhada deve ser tal que os pontos fiquem aleatoriamente distribuídos em torno dessa reta. Esse desenho é feito de forma manual e exige senso crítico por parte da pessoa que está realizando a análise. A figura 4.6 mostra o mesmo conjunto de dados com a reta média correspondente.

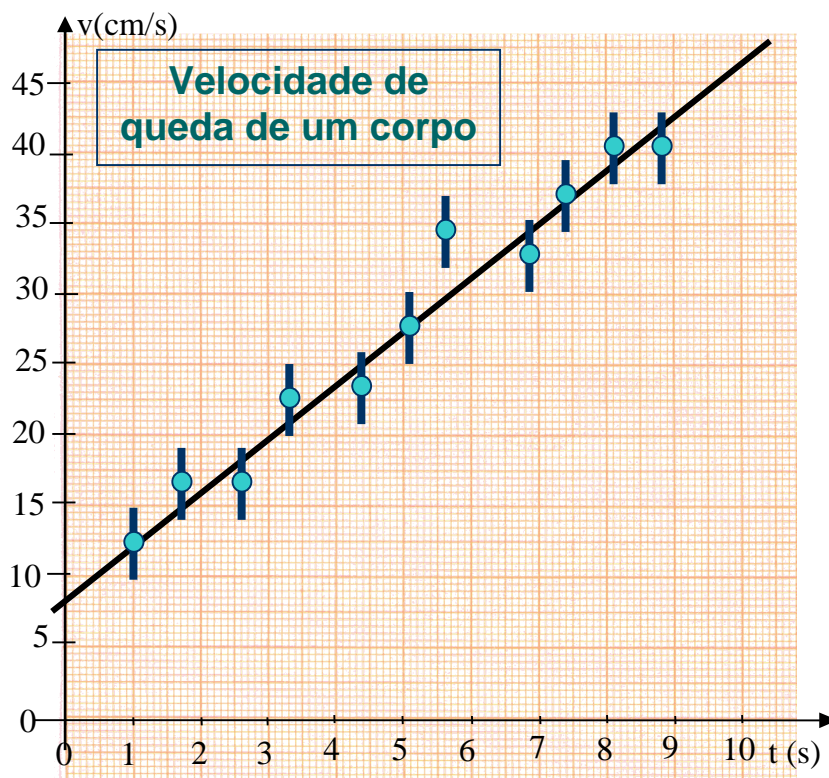


Figura 4.6. Velocidade de queda de um ovo com a sua respectiva reta média que é utilizada para extrair informações numéricas a respeito do movimento de queda.

Note que a reta média não necessariamente deve passar por todos os pontos experimentais (veja ponto com  $t = 5,6$  s) e, não necessariamente, deve passar pelo primeiro e último pontos do gráfico. O critério é que os pontos fiquem distribuídos em torno da reta da forma mais aleatória possível.

Deve-se ter cuidado com o uso dessa técnica para traçar retas médias. Em muitos casos, apesar das incertezas experimentais serem suficientemente grandes, os pontos não ficam aleatoriamente distribuídos em torno da reta. Nesse caso, é evidente que a função que descreve a curva média não deve ser uma reta. Um exemplo é mostrado na figura 4.7. Note que os pontos não estão igualmente distribuídos em torno da reta média. Nota-se que, apesar do número de pontos sobre a reta ser equivalente ao número de pontos sob a reta, há a tendência de haver pontos na parte inferior somente nos extremos do gráfico enquanto os pontos superiores encontram-se na região central do gráfico. Esse é um exemplo claro de que a curva média selecionada (reta) não é adequada para descrever os dados experimentais. Mais uma vez, existem métodos matemáticos para avaliar se a função utilizada é a que melhor descreve os dados experimentais, porém o aprendizado desse método foge ao escopo da disciplina. O desenvolvimento da intuição, nesse caso, é importante no julgamento dos resultados obtidos.

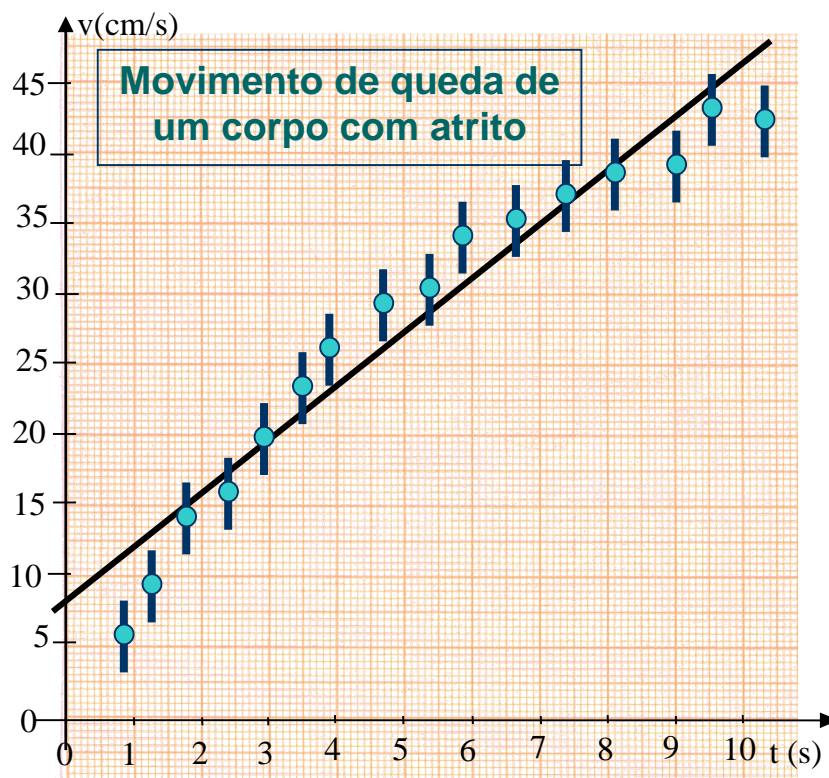


Figura 4.7. Conjunto de dados no qual o uso de uma reta média não é adequado para descrever o comportamento dos dados.

Em um gráfico de escalas lineares (papel milimetrado) retas são objetos geométricos simples de serem representados matematicamente. Nesse caso, a equação de uma reta pode ser escrita como:

$$y = ax + b$$

Onde  $y$  é a variável dependente e  $x$  é a variável independente.  $a$  e  $b$  são constantes, respectivamente denominadas coeficientes angular e linear.

Para obter os coeficientes  $a$  e  $b$  é necessário escolher dois pontos da reta média desenhada no gráfico. **ESCOLHA PONTOS BASTANTE DISTANTES!!!!** Pontos muito próximos acarretam em incertezas bastante elevadas e, muitas vezes, fora de controle. De preferência, escolha um ponto anterior ao intervalo dos dados e um ponto após o intervalo das medidas efetuadas. Vamos denominar esses pontos como sendo  $(x_1, y_1)$  e  $(x_2, y_2)$ . Utilizando a equação de reta acima, podemos escrever que:

$$y_1 = ax_1 + b \quad \text{e} \quad y_2 = ax_2 + b$$

Temos, nesse caso, duas equações e duas incógnitas ( $a$  e  $b$ ). Podemos resolver o sistema acima de tal modo que:

$$a = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \quad \text{e} \quad b = y_1 - ax_1$$

Note que os parâmetros  $a$  e  $b$  possuem unidades. A unidade de  $a$  é [unidade de  $y$ ]/[unidade de  $x$ ] enquanto a unidade de  $b$  é [unidade de  $y$ ].

Note que, apesar do nome, o coeficiente angular não é igual à tangente do ângulo entre a reta e o eixo- $x$ , porque as escalas de um gráfico são, em geral, diferentes nos eixos  $x$  e  $y$ , ao contrário do caso geométrico. Lembre-se que o coeficiente angular possui unidade enquanto tangente de um ângulo é um número adimensional. Em geral:

$$\frac{\Delta y}{\Delta x} \neq \tan \theta$$

### *Avaliação de incertezas nos coeficientes angular e linear*

A representação gráfica, como vimos, é importante no sentido de ilustrar e sintetizar as relações entre grandezas representativas de um fenômeno. Contudo, medidas experimentais são sempre acompanhadas de suas respectivas incertezas, avaliadas pelos experimentadores. Essas incertezas são representadas graficamente através de barras de erro em cada ponto experimental, conforme mostrado nas figuras anteriores.

Uma pergunta natural que surge do ajuste da reta média, como o realizado na figura 4.6 reflete o fato das incertezas, bem como as flutuações nos pontos experimentais, permitirem que mais do que uma reta média possa ajustar razoavelmente os dados experimentais. É razoável pensar que os coeficientes angular e linear obtidos para a reta média possuem incertezas associadas. Como avaliar a incertezas desses coeficientes?

Tanto a escolha da melhor curva, como mencionado, como o cálculo das incertezas nos coeficientes, pode ser feito de forma rigorosa. Contudo, assim como há um método gráfico razoável para traçar a reta média, há também um método gráfico que pode ser utilizado para estimar as incertezas nos coeficientes obtidos. Esse método consiste em estimar duas retas, uma de máxima inclinação e outra de mínima inclinação, que ainda se adaptem de forma razoável aos dados experimentais. O procedimento a seguir tenta sistematizar esse método de tal forma que as incertezas obtidas sejam razoáveis.

Vamos voltar aos dados apresentados na figura 4.6. Imagine agora dois conjuntos de pontos. Um desses conjuntos tem coordenadas  $(x, y+\sigma)$  enquanto o outro conjunto de pontos tem coordenadas  $(x, y-\sigma)$ , sendo  $\sigma$  a incerteza de cada um dos pontos do conjunto original, conforme mostrado na figura 4.8-a. Nessa figura esses conjuntos estão representados por quadrados e triângulos, respectivamente. **VOCÊ NÃO PRECISA DESENHAR ESSES PONTOS NOS SEUS GRÁFICOS!** Eles são apenas

guias visuais para fins didáticos. A seguir, traça-se duas retas, uma que melhor se adapte ao conjunto  $(x, y + \sigma)$  e outra que melhor se adapte ao conjunto  $(x, y - \sigma)$ , conforme mostrado na figura 4.8-b. Note que essas retas não precisam ser paralelas entre si e nem mesmo paralelas à reta média ajustada.

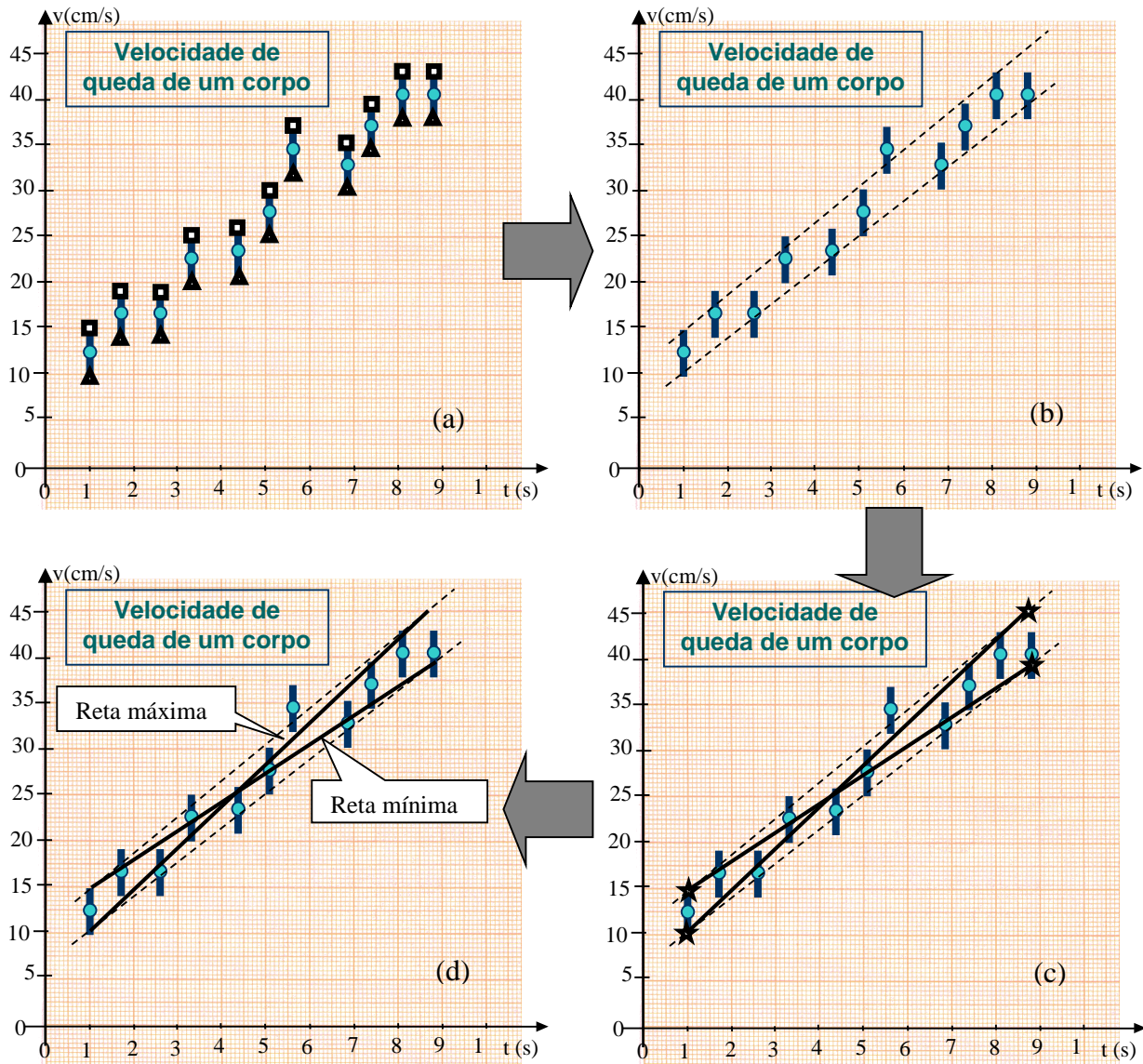


Figura 4.8. Procedimento para estimar as incertezas nos coeficientes da reta média.

A seguir, tomam-se os pontos nessas retas correspondentes ao menor e maior valor da variável  $x$  no conjunto de dados experimentais (ver estrelas na figura 4.8-c). Esses pontos servem de referência para traçar as retas máxima e mínima. Para traçar as retas máxima e mínima, ligam-se os pontos marcados por estrelas, conforme mostrado na figura 4.8-c por retas contínuas.

A figura 4.8-d mostra a figura final obtida. As duas retas contínuas obtidas são denominadas retas máxima e mínima por possuírem, respectivamente, máxima e mínima inclinações. Para cada uma dessas retas calcula-se os coeficientes angulares e lineares, denominados, respectivamente  $a_{max}$ ,  $b_{max}$ ,  $a_{min}$ ,  $b_{min}$ . As incertezas nos coeficientes da reta média podem ser obtidas através das expressões:

$$\sigma_a = \frac{|a_{max} - a_{min}|}{2} \quad \text{e} \quad \sigma_b = \frac{|b_{max} - b_{min}|}{2}$$

### ***Linearização de dados***

Provavelmente por razões biológicas, o ser humano sabe distinguir bem entre uma curva e uma reta. Porém, é muito difícil para o ser humano perceber, graficamente, a diferença entre uma curva dada por  $y = x^2$  e outra dada por  $y = x^4$ . Em trabalhos técnico-científicos, os dados experimentais, nem sempre, produzem uma curva linear do tipo  $y = ax + b$ , fácil de extrair informações quantitativas, como descritas anteriormente. Nesse caso faz-se uso de técnicas de linearização de dados, de tal forma que os dados finais obtidos, quando graficados, forneçam uma linha reta, fácil de ser analisada. Experiência e bom senso são elementos importantes para essa operação, bem como o conhecimento da equação esperada para os dados originais.

O ingrediente básico para linearização de dados é o conhecimento da equação esperada para descrever os dados originais. A técnica consiste no uso dessa equação para realizar mudanças de variáveis de tal forma que o gráfico dessas novas variáveis seja uma reta.

Vamos tomar como exemplo um corpo em queda livre. Em um experimento, realizou-se a medida da altura desse corpo ( $h$ ) para diversos instantes de tempo ( $t$ ), conforme mostrado na tabela 4.1. Fazendo o gráfico de altura como função do tempo de queda, obtém-se a figura 4.9. Observando esse gráfico, percebe-se que ele tem uma forma de parábola com a concavidade para baixo. De fato, esse é o comportamento esperado para um corpo em queda livre. Assim, podemos supor que a equação que melhor descreveria o comportamento da altura em função do tempo pode ser escrita como:

$$h(t) = C + At^2$$

Onde  $C$  e  $A$  são constantes que devem ser obtidas a partir da análise dos dados. Como obtê-las?

| $t$ (s) | $h$ (cm) | $z = t^2$ (s <sup>2</sup> ) |
|---------|----------|-----------------------------|
| 0,010   | 200      | 0,00010                     |
| 0,225   | 173      | 0,0506                      |



|       |     |        |
|-------|-----|--------|
| 0,319 | 151 | 0,1018 |
| 0,390 | 124 | 0,1521 |
| 0,450 | 99  | 0,2025 |
| 0,504 | 76  | 0,2540 |
| 0,552 | 48  | 0,3047 |
| 0,596 | 26  | 0,3552 |
| 0,637 | 1   | 0,4058 |

Tabela 4.1. Altura ( $h$ ) em função do tempo ( $t$ ) para um corpo em queda livre.

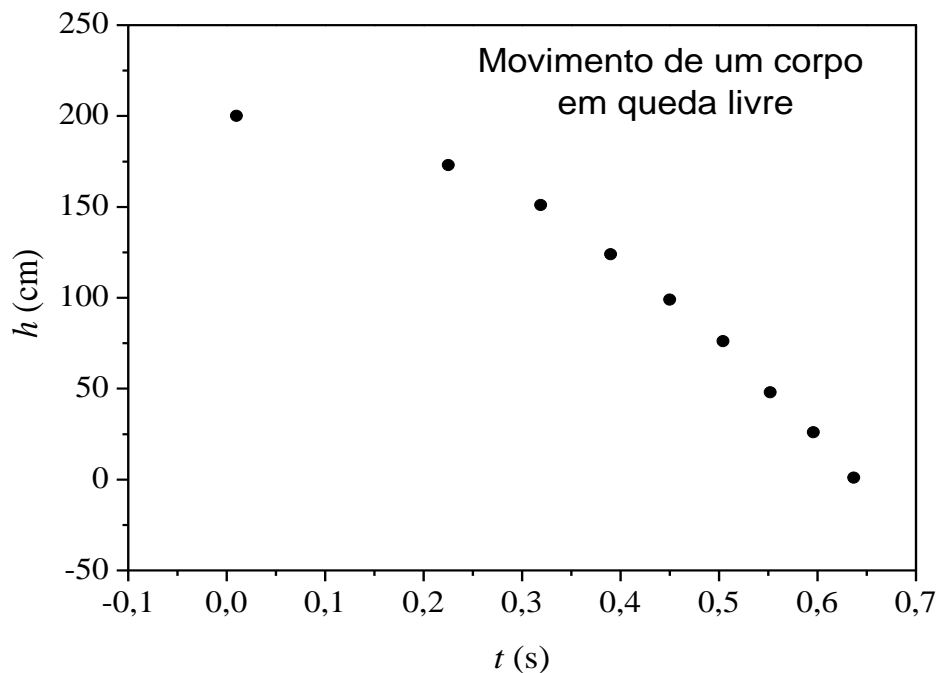


Figura 4.9 – Altura de um corpo em queda livre como função do tempo de queda.

Podemos testar se, de fato, a expressão  $h(t) = C + At^2$  representa bem os dados obtidos utilizando técnicas de linearização. Para transformar essa expressão em uma reta, devemos fazer a mudança de variável  $z = t^2$ . Realizando essa mudança de variáveis obtemos a expressão:

$$h(t) = C + Az,$$

que é a equação para uma reta. A terceira coluna na tabela 4.1 mostra o valor da variável  $z$ , calculada a partir dos dados obtidos para o tempo de queda. A figura 4.10 mostra o gráfico da altura de queda em função da variável  $z$ . Pode-se descrever o gráfico obtido através de uma reta, mostrando que a suposição utilizada para a linearização funciona adequadamente.

A partir de um ajuste de reta média, como descrita anteriormente, pode-se obter, sem complicações, os valores para os coeficientes  $C$  e  $A$ .

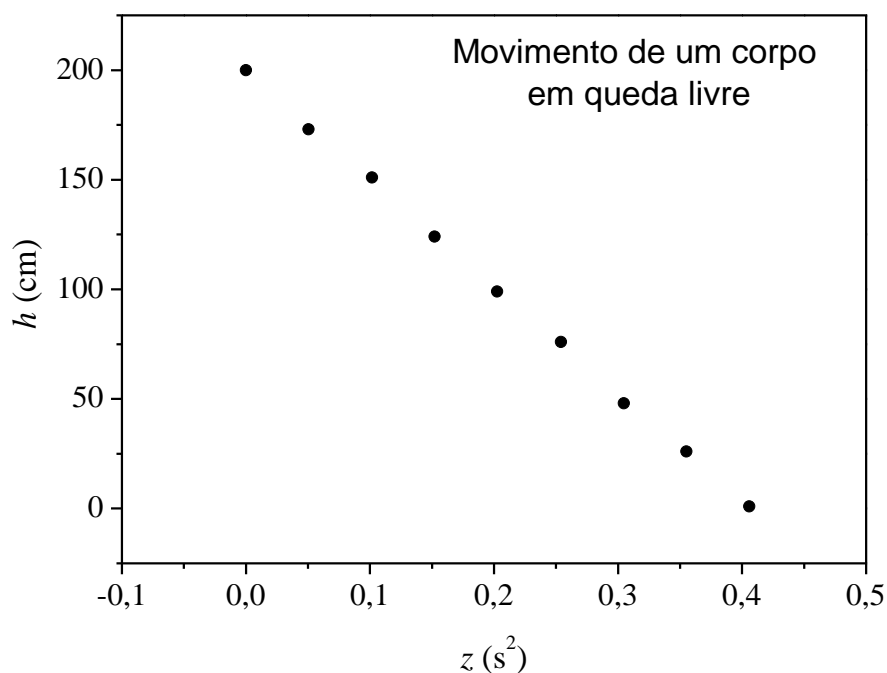


Figura 4.10 – Altura de um corpo em queda livre como função do tempo de queda ao quadrado.

Técnicas de linearização são muito utilizadas na análise gráfica de dados e simplificam consideravelmente o tratamento desses dados. Deve-se lembrar que, caso a mudança de variáveis ocorra sobre uma grandeza que possua incertezas, as incertezas associadas à nova variável devem ser obtidas através de técnicas de propagação de erros, como descritas nessa apostila.

## 4.2. Escalas logarítmicas

Em muitas situações é comum fazer gráficos de grandezas onde a dependência com uma outra variável é dada por expressões do tipo:

$$y(x) = A^{Bx} \quad \text{ou} \quad y(x) = Ax^B$$

Nesse caso, dependendo das constantes  $A$  e  $B$ , a grandeza  $y(x)$  pode variar muitas ordens de grandeza a partir de pequenas variações de  $x$ . É claro que, nesse caso, mudanças de variáveis podem ser realizadas para tornar as equações acima retas. Em geral, as mudanças de variáveis mais comuns envolvem funções logarítmicas. No passado, o cálculo de logaritmos era bastante trabalhoso e envolvia consulta a tabelas (ou tábuas) de logaritmos, nem sempre disponíveis. Nesse sentido, foram criados papéis gráficos especiais nos quais uma (ou ambas) das escalas é graduada logaritmicamente. A escala logarítmica é construída de tal forma que quando uma quantidade  $x$  é marcada nessa escala o comprimento (distância em relação à origem do eixo) é proporcional à  $\log(x)$ . Um trecho de uma escala logarítmica é mostrado na figura 4.11. Assim, a escala logarítmica é útil

quando a mudança de variável necessária para linearizar o gráfico envolver o logaritmo de um número.

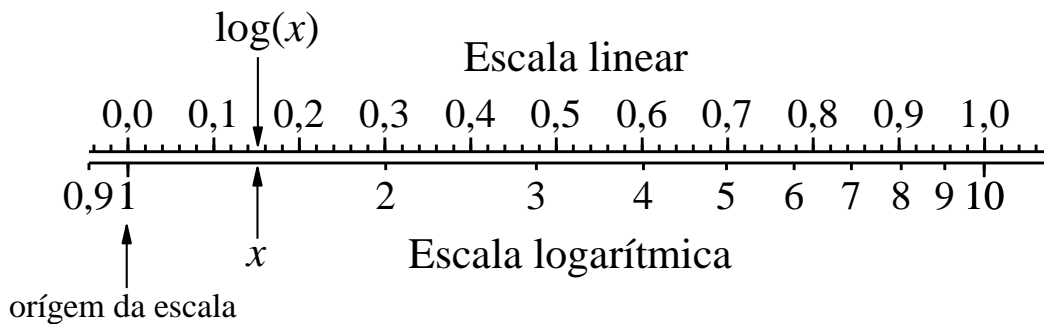


Figura 4.11. Escala logarítmica (abaixo) em comparação com a escala linear (acima). A escala logarítmica é construída de tal forma que quando uma quantidade  $x$  é marcada nessa escala o comprimento (distância em relação à origem do eixo) é proporcional a  $\log(x)$ .

Devido à forma na qual a escala logarítmica é construída, deve-se ficar atento para algumas regras de uso:

1. Não existe zero em escala logarítmica. Devido ao fato de  $\lim_{x \rightarrow 0} (\log(x)) = -\infty$  é impossível definir o valor zero na escala.
2. A escala logarítmica é dividida em décadas. Cada década corresponde a uma ordem de grandeza decimal. A divisão da escala, em cada década, é idêntica de uma década para outra.
3. Pelo fato da posição da escala ser proporcional a  $\log(x)$  não podemos escolher qualquer escala para fazer o gráfico. A posição equivalente ao 1 na escala logarítmica da figura 4.11 pode ser atribuída somente a números do tipo 1; 0,1; 10; 1000; etc. Do mesmo modo, a posição 3 só pode ser atribuída a números do tipo 3; 0,3; 30; 3000; etc.
4. Uma década subsequente tem que, necessariamente, possuir escala de tal forma que os números são marcados uma ordem de grandeza acima da década anterior. Por exemplo, caso a década anterior varie de 0,01 à 0,1; a década subsequente deve variar de 0,1 à 1 e assim sucessivamente.

Um uso interessante para a escala logarítmica diferente de fazer gráficos é a forma simples de calcular logaritmos. Como a posição de um valor  $x$ , na escala, é proporcional a  $\log(x)$ , e como o tamanho de uma década corresponde a variação de 1 em logaritmos ( $\log(10x) - \log(x) = 1$ , qualquer

que seja  $x$ ) podemos usar essa informação para o cálculo de logaritmos. Para isso, basta medir a distância  $d$  (em centímetros) da posição de  $x$  na escala logarítmica e o tamanho da década  $D$ , conforme mostra a figura 4.12. Desse modo,  $\log(x)$  vale:

$$\log(x) = \frac{d \text{ (cm)}}{D \text{ (cm)}}$$

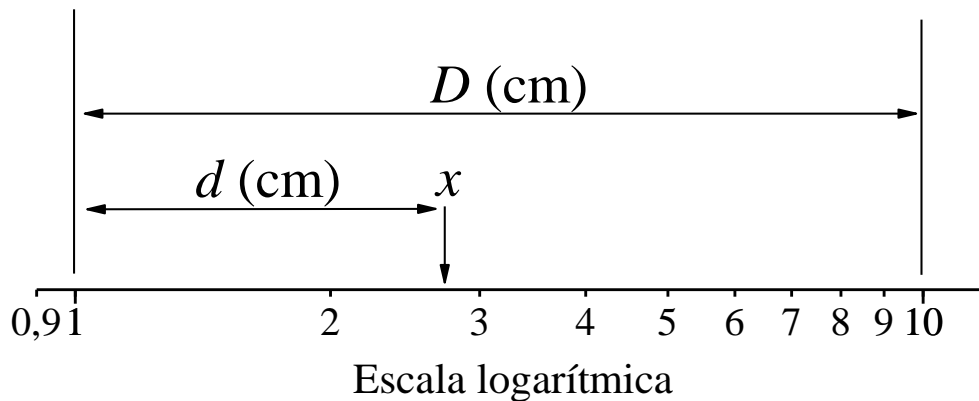


Figura 4.12. Cálculo de  $\log(x)$  utilizando a escala logarítmica

### **Gráfico mono-log**

O gráfico mono-log é um gráfico com escala linear no eixo-x e escala logarítmica no eixo-y, conforme mostra a figura 4.14. Esse tipo de escala é bastante útil para gráficos com comportamentos exponenciais, do tipo:

$$y(x) = CB^{Ax}$$

onde  $A$  e  $B$  são os coeficientes da expressão. Vamos agora calcular o logaritmo da expressão acima. Desse modo:

$$\log(y(x)) = \log(CB^{Ax}) = \log(C) + \log(B^{Ax})$$

ou

$$\log(y(x)) = Ax \log(B) + \log(C)$$

Fazendo uma mudança de variáveis  $z(x) = \log(y(x))$ , podemos reescrever a equação acima como sendo:

$$z(x) = ax + c,$$

onde  $a = A \log(B)$  e  $c = \log(C)$ .

Desse modo, situações nas quais os dados se comportam como funções exponenciais tornam-se retas quando graficados em papel mono-log. Pode-se, a partir desse gráfico, desenhar a reta média, bem como as retas mínima e máxima para cálculo das incertezas nos coeficientes. Depois de desenhada as retas ajustadas aos dados, o coeficiente angular ( $a$ ) pode ser calculado a partir de dois pontos quaisquer sobre a reta ajustada ( $x_1, y_1$ ) e ( $x_2, y_2$ ) utilizando a expressão (ver figura 4.13):

$$a = \frac{z_2 - z_1}{x_2 - x_1} = \frac{\log(y_2) - \log(y_1)}{x_2 - x_1}$$

Ou, simplesmente, medindo-se a distância, em centímetros, entre os pontos  $y_1$  e  $y_2$  ( $d$ ) bem como o tamanho da década no gráfico ( $D$ ) e utilizando a expressão:

$$a = \frac{d/D}{x_2 - x_1}$$

A constante  $C$  pode ser obtida diretamente pela leitura da escala no eixo-y para o qual  $x = 0$ .

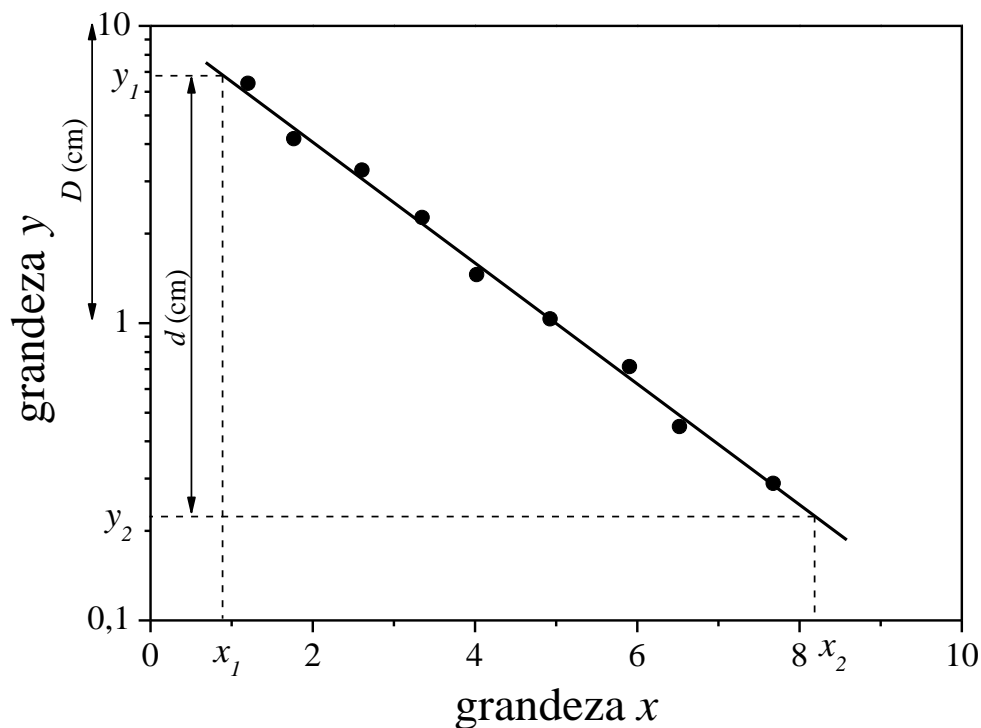


Figura 4.13. Cálculo do coeficiente angular em um papel mono-log.

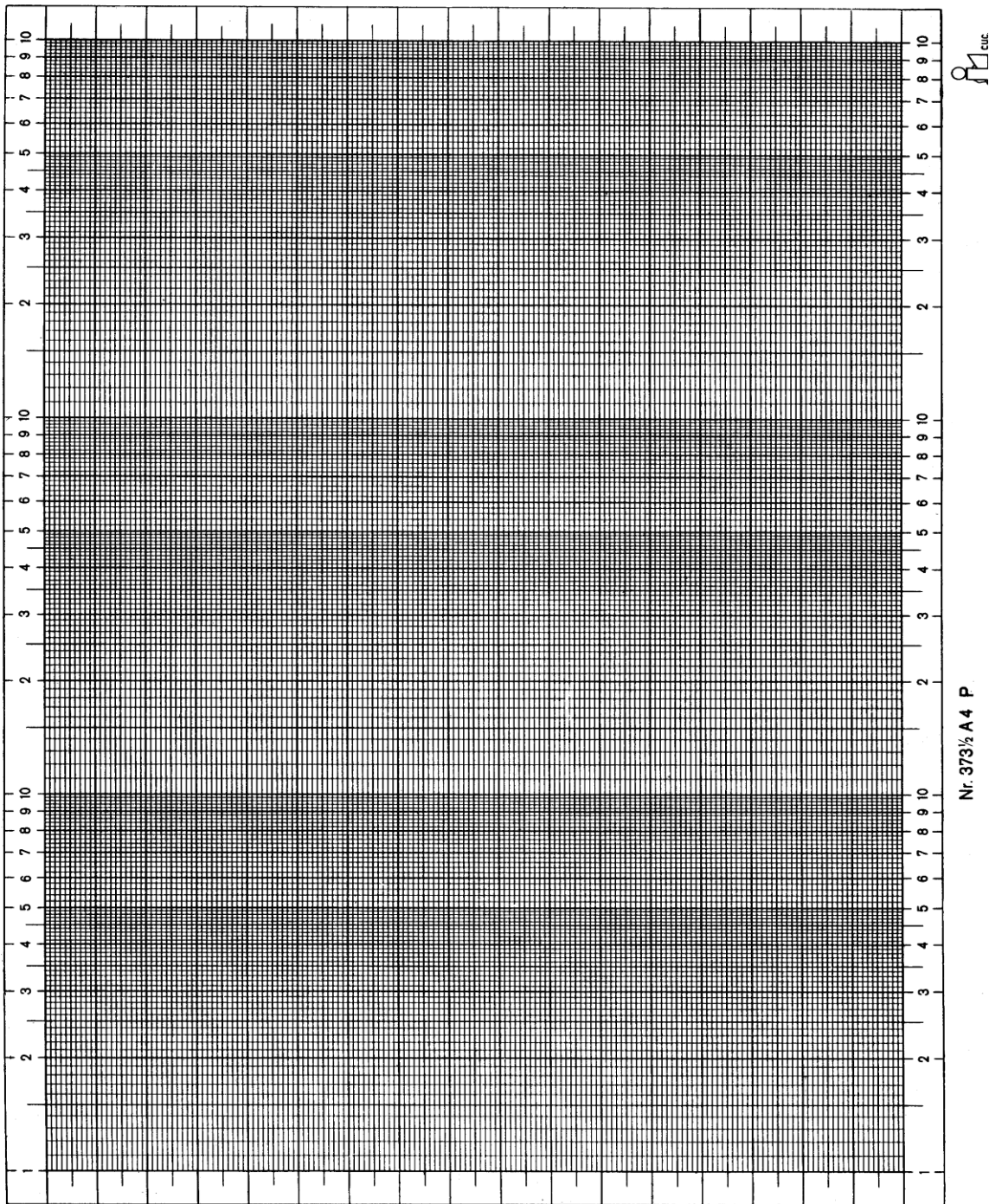


Figura 4.14. Papel mono-log. Você pode usar essa figura como modelo para gráficos mono-logs. Basta fazer cópias xérox.

### ***Gráfico di-log***

Como o próprio nome diz, o gráfico di-log é aquele onde ambos os eixos x e y estão em escala logarítmica (figura 4.16). Esse gráfico é útil para linearizar expressões do tipo:

$$y(x) = Bx^A .$$

Aplicando-se log na equação acima obtemos:

$$\log(y(x)) = \log(Bx^A) = \log(B) + A\log(x)$$

Fazendo as mudanças de variáveis

$$z(x) = \log(y(x))$$

e

$$k(x) = \log(x)$$

Podemos escrever a equação acima como sendo

$$z(x) = ak(x) + b$$

Ou seja, a equação de uma reta. Nesse caso, as constantes  $a$  e  $b$  valem, respectivamente,  $a = A$  e  $b = \log(B)$ .

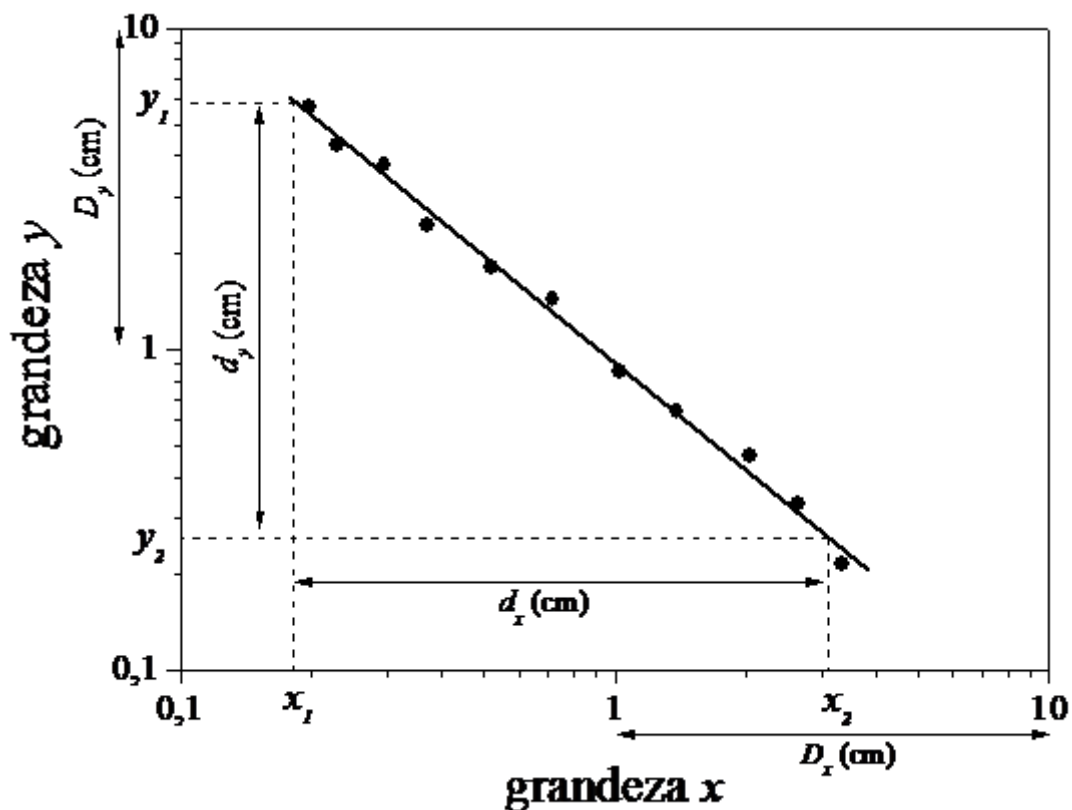


Figura 4.15. Cálculo do coeficiente angular em um papel di-log.

Da mesma forma que no gráfico mono-log, caso o gráfico resulte em uma reta, pode-se traçar a reta média para o cálculo dos coeficientes  $a$  e  $b$ , bem como as retas máxima e mínima para a estimativa das incertezas nos coeficientes. Escolhendo-se dois pontos sobre as retas ajustadas  $(x_1, y_1)$  e  $(x_2, y_2)$ , o coeficiente  $a$ , vale, nesse caso:

$$a = \frac{z_2 - z_1}{k_2 - k_1} = \frac{\log(y_2) - \log(y_1)}{\log(x_2) - \log(x_1)}$$

Ou, simplesmente, medindo-se a distância, em centímetros, entre os pontos  $y_1$  e  $y_2$  ( $d_y$ );  $x_1$  e  $x_2$  ( $d_x$ ) bem como o tamanho das décadas no gráfico ( $D_y$  e  $D_x$ ) e utilizando a expressão:

$$a = \frac{d_y/D_y}{d_x/D_x}$$

A constante B pode ser obtida diretamente pela leitura da escala no eixo-y para o qual  $x = 1$  (caso onde  $\log(x) = 0$ ).

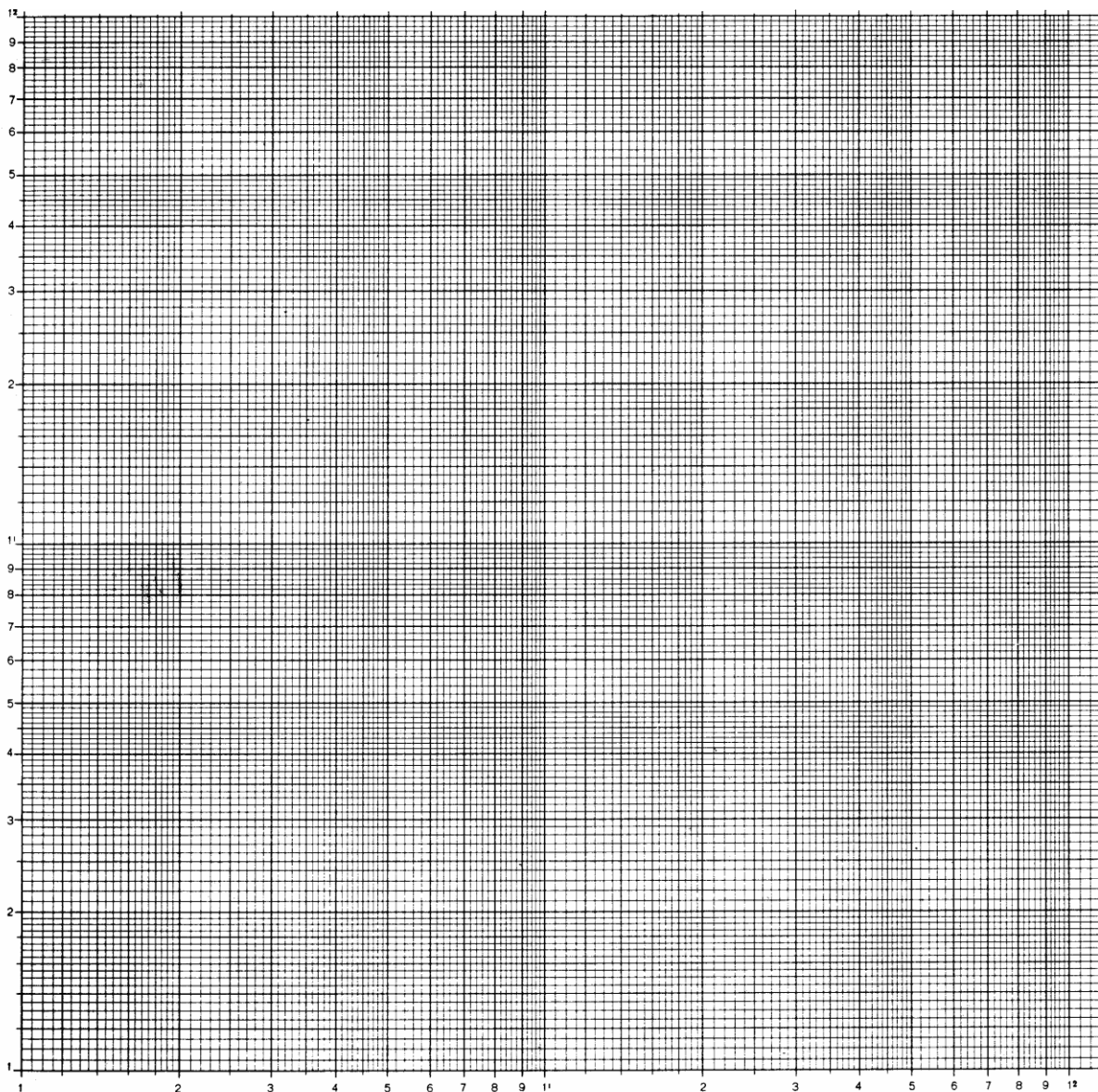


Figura 4.16. Papel di-log. Você pode usar essa figura como modelo para gráficos di-log. Basta fazer cópias xérox.



## 5. Histogramas

Vamos imaginar o seguinte experimento. Um cientista resolve medir o período de oscilação de um pêndulo. Após realizar o experimento uma única vez ele obtém um determinado valor  $T$  para o período de oscilação desse pêndulo. Contudo, após repetir o experimento várias vezes ele observa que cada experimento, mesmo que efetuado sob as mesmas condições experimentais (aquelas controladas pelo experimentador), fornece um valor diferente para o período de oscilação. Nesse caso, o experimentador conclui que o período de oscilação do pêndulo pode ser dado pela média de todas as medidas efetuadas. Contudo, outras questões podem ser igualmente importantes: como as medidas se distribuem em torno desse valor médio? O valor médio é também o valor mais provável de ser medido? Qual a probabilidade de realizar uma medida na qual o período de oscilação obtido é duas vezes maior que o valor médio?

Muitas dessas questões podem ser resolvidas através da análise estatística das medidas efetuadas. Contudo, uma ferramenta importante para análise estatística é o histograma das medidas. Um histograma é um gráfico no qual o conjunto de pontos  $(x, y)$  tem um significado específico. Um certo valor  $y$  está diretamente relacionado com a probabilidade de efetuar uma determinada medida e obter, como resultado, o valor  $x$ . Voltando ao nosso exemplo do pêndulo, a variável graficada no eixo- $x$  poderia ser o período de oscilação enquanto que a variável no eixo- $y$  pode ser o número de vezes que aquele determinado período foi medido pelo experimentador.

Por ter um significado específico, muitas vezes um histograma não é graficado colocando pontos nas coordenadas  $(x, y)$  de um papel milimetrado e sim através dos desenhos de barras verticais cuja altura corresponde ao valor  $y$  obtido para o ponto  $x$ .

A figura 5.1 mostra um histograma típico para o nosso experimento fictício. Nesse caso, o experimentador realizou a mesma medida 200 vezes. Cada barra vertical no histograma corresponde a um intervalo de períodos. Por exemplo, a barra mais alta corresponde a medidas cujo período de oscilação estava entre 0,40 e 0,43 segundos. Após repetir 200 vezes o experimento, o experimentador obteve 39 medidas cujo período de oscilação do pêndulo encontrava-se nesse intervalo de tempo. Para o intervalo de tempo entre 0,50 e 0,53 segundos, o experimentador obteve somente 6 medidas nesse intervalo. Cada um desses intervalos de medidas, que corresponde a uma barra no histograma é denominado de um canal do histograma. Em geral, histogramas possuem canais cujas larguras são fixas para todo o histograma. Casos especiais de histograma possuem canais de larguras variadas, porém são mais difíceis de serem analisados.

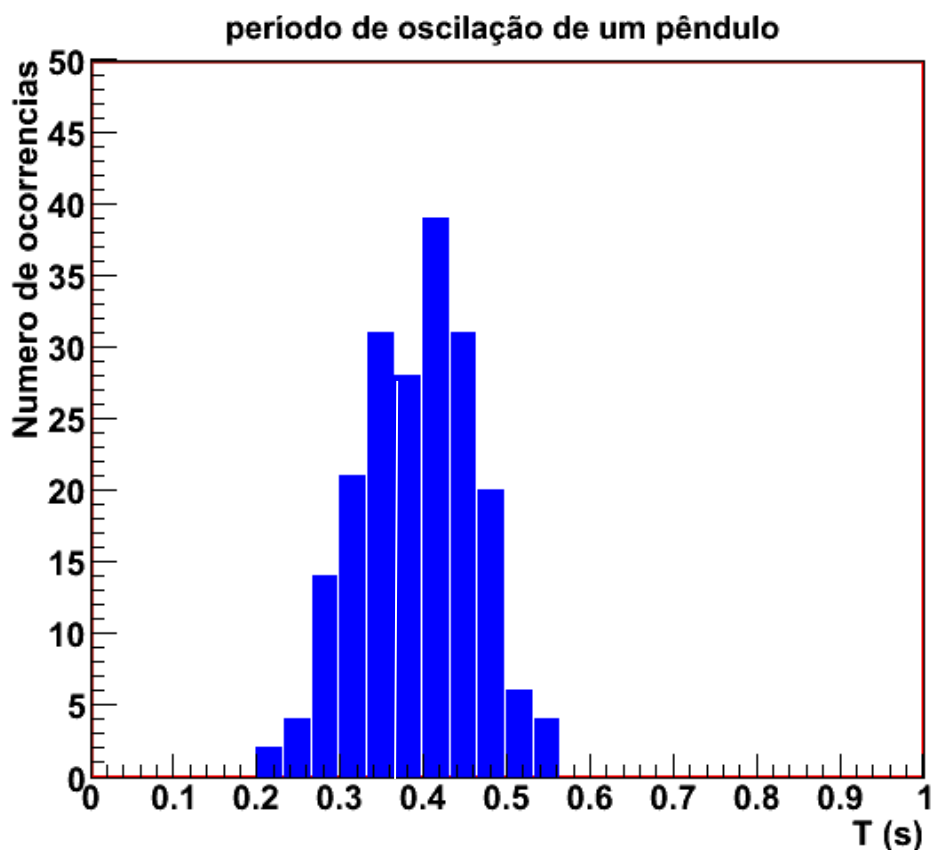


Figura 5.1. Histograma do período de oscilação de um pêndulo simples para um experimento realizado 200 vezes.

A amplitude a ser graficada em um histograma, para cada intervalo de variação da medida, depende de como esse histograma será utilizado posteriormente. É comum, contudo, utilizar uma das seguintes opções:

#### ***Histograma de número de ocorrências ( $N$ )***

A amplitude do histograma,  $N(x)$ , é simplesmente o número de ocorrências verificadas em cada canal do histograma cujo centro vale  $x$ . Apesar de ser o histograma mais simples de se construir, pois exige apenas a contagem do número de ocorrências, a análise do mesmo é mais trabalhosa. Por exemplo, para calcular a probabilidade de efetuar uma medida em um intervalo é necessário saber o número total de medidas utilizadas no histograma.

#### ***Histograma de frequência de ocorrência ( $F$ )***

A frequência na qual ocorre uma determinada medida é definida como sendo a razão entre o número de ocorrências em um canal do histograma cujo centro vale  $x$  e o número total de medidas efetuada, ou seja:

$$F(x) = \frac{N(x)}{N_{total}}$$

A vantagem de utilizar essa variável como amplitude do histograma é óbvia. A simples leitura da amplitude do histograma em um determinado canal, no limite de um grande número de medidas,  $N_{total}$ , tende à probabilidade de realizar uma medida no intervalo correspondente ao canal estudado. No caso mostrado na figura 5.1, como o experimento foi realizado 200 vezes, a frequência de ocorrência para um dado canal é o número de contagens daquele canal, dividido por 200.

Apesar de os histogramas de ocorrências (N) e frequências (F) serem simples de construir eles possuem algumas limitações. A maior delas é o fato das amplitudes nesses histogramas serem fortemente dependentes da largura escolhida para os canais. Caso a largura escolhida seja duas vezes maior, tanto os números de ocorrências como as frequências serão também duas vezes maiores. Esse aspecto torna histogramas de ocorrências e frequências difíceis de serem comparados com outros histogramas, bem como com curvas teóricas. Um terceiro tipo de histograma, definido como histograma de densidades de probabilidade, elimina essa limitação.

### ***Histograma de densidade de probabilidades (H)***

A densidade de probabilidade é definida como sendo a razão entre a probabilidade de realizar uma medida no intervalo  $x$  e  $x+dx$  e o tamanho do intervalo,  $dx$ , no limite no qual esse intervalo é muito pequeno, ou seja:

$$H(x) = \frac{dP}{dx}$$

Se a densidade de probabilidade é conhecida, a probabilidade de ocorrer um resultado em um intervalo  $(x, x+\Delta x)$ , com  $\Delta x$  pequeno, é, aproximadamente:

$$P(x, x+\Delta x) \approx H(x)\Delta x$$

A grande vantagem de utilizar a densidade de probabilidade para montar histogramas é o fato das amplitudes em cada canal ser independente do número de medidas efetuadas bem como da largura escolhida para os canais do histograma. Experimentalmente, a densidade de probabilidade pode ser obtida como sendo a frequência de ocorrência de eventos em um canal, dividida pela largura do canal no histograma, ou seja:

$$H(x) \approx \frac{F(x)}{\Delta x} = \frac{N(x)}{N_{total}\Delta x}$$

## 5.1. Construção de histogramas

Depois de realizadas as medidas, o experimentador tem em mãos uma tabela na qual estão listados os valores obtidos para a grandeza que se quer histogramar. Construir um histograma consiste nos seguintes passos:

1. Escolher a largura dos canais do histograma,  $\Delta x$ ;
2. Escolher os centros de cada canal, tomando o cuidado que não sobre espaços vazios entre os canais.
3. Contar o número de ocorrências para cada um dos canais,  $N(x)$ . Nesse ponto é possível construir o histograma de número de ocorrências. Caso uma ocorrência ocorra na borda entre dois canais, considere a ocorrência como pertencendo ao canal cujo centro possua maior valor.
4. Caso queira-se construir o histograma de frequências,  $F(x)$  dividir o número de ocorrências em cada canal pelo total de medidas efetuadas.
5. Caso queira-se construir o histograma de densidade de probabilidades,  $H(x)$ , dividir a frequência de cada canal pela largura de cada um dos canais.

Alguns problemas ocorrem na criação do histograma, principalmente quando o número total de medidas ( $N_{total}$ ) é estatisticamente pequeno.

O problema mais frequente é a escolha da largura do canal,  $\Delta x$ . Evidentemente, para que a densidade de probabilidade experimental seja o mais próxima possível da definição teórica, deve-se escolher  $\Delta x$  de tal forma a ser o menor valor possível. Entretanto, diminuindo  $\Delta x$  estamos também diminuindo o número de ocorrências em cada canal do histograma, correndo o risco de que, em casos extremos, ocorram canais onde não seja registrada nenhuma ocorrência.

A figura 5.2 mostra dois histogramas onde foram realizadas 20 medidas. No histograma da esquerda, a largura do canal utilizada é cinco vezes mais larga que no histograma da direita. Note que o histograma com largura de canal menor apresenta flutuações elevadas de um canal para outro, além de haver canais onde não há ocorrências. Isso resulta em alguns canais com elevada densidade de probabilidade enquanto outros canais apresentam densidade de probabilidade nula.

Esse fator deixa de ser um problema quando o número de medidas é bastante elevado, como mostrado na figura 5.3. Nesse caso, o experimento hipotético foi realizado 20 mil vezes. Note que, além do tamanho dos canais, não há diferença entre as densidades de probabilidade entre os histogramas.

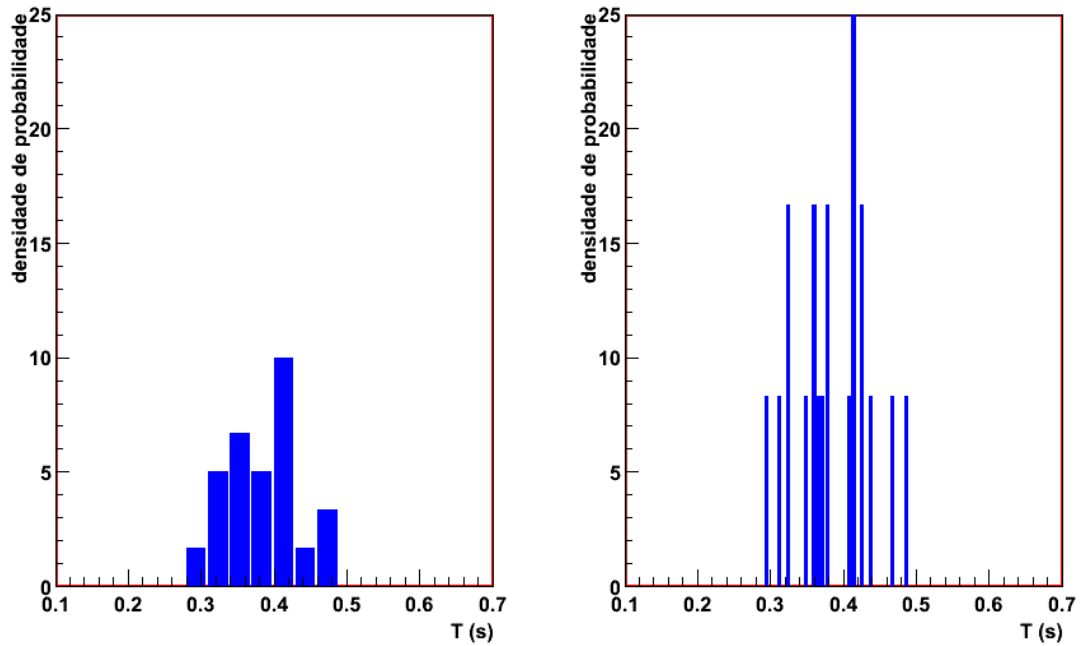


Figura 5.2. Histogramas de densidade de probabilidades para medidas do período de um pêndulo simples. O conjunto de dados utilizado é o mesmo em ambos os casos. O histograma da esquerda foi montado de tal forma que a largura do canal seja 5 vezes maior que no caso da direita. O total de medidas utilizadas para montar os histogramas ( $N_{total}$ ) foi 20.

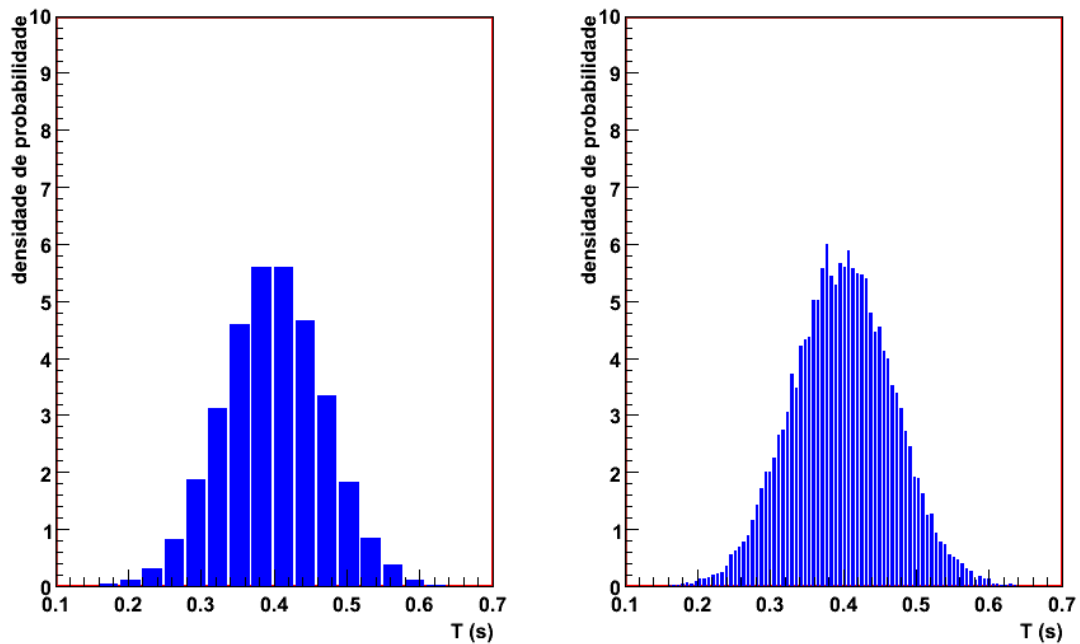


Figura 5.3. Histogramas de densidade de probabilidades para medidas do período de um pêndulo simples, conforme explicado na figura 5.2. Nesse caso, o total de medidas utilizadas para montar os histogramas ( $N_{total}$ ) foi 20000.

Em muitas situações experimentais é muito difícil realizar um número elevado de medidas de tal forma que a escolha da largura dos canais no

histograma possa ser arbitrariamente pequena. Como regra prática, a largura dos canais,  $\Delta x$ , deve ser escolhida de tal forma que o número de ocorrências,  $N(x)$ , seja pelo menos 10 para os canais próximos ao valor médio das medidas. Outro fator importante é a escolha das posições centrais dos canais do histograma. Deve-se, nesse caso, escolher as posições centrais de tal forma que uma delas seja aproximadamente igual ao valor médio das medidas.

## 5.2. Interpretação de um Histograma

Quando medimos  $N$  vezes uma grandeza, normalmente obtemos valores diferentes para cada medida devido à incerteza estatística ou aleatória associada ao procedimento de medida. Se a incerteza é aleatória, é razoável supor que ela pode fazer com que o resultado da medida seja igualmente maior, ou menor, que o valor verdadeiro da grandeza. Portanto, esperamos que um histograma tenha uma forma simétrica em torno do valor que representa a melhor estimativa para o valor verdadeiro da medida, como podemos observar no histograma da figura 5.4.

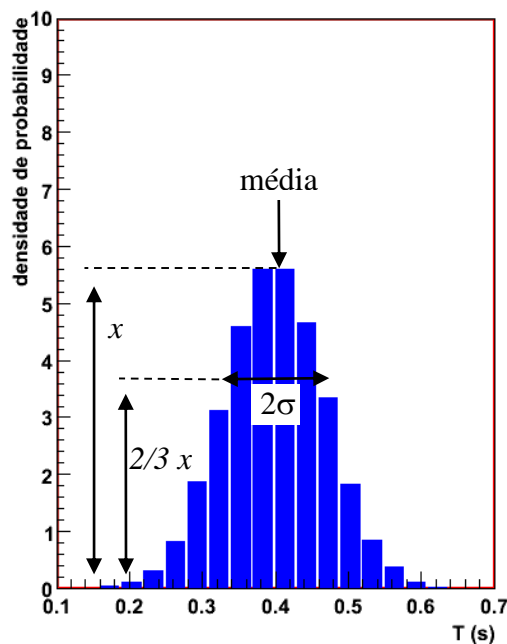


Figura 5.4 – Obtenção de média e desvio padrão a partir da análise gráfica do histograma.

A largura do histograma deve refletir a precisão da medida, pois ela mostra o quanto as medidas variaram em torno da estimativa do valor verdadeiro. Um histograma mais largo significa uma medida menos precisa e vice-versa. Como discutido no capítulo 5 da apostila “Introdução à Teoria

de Erros” de J. H. Vuolo, a melhor estimativa do valor verdadeiro de uma medida é dada pela média e a variação (ou variância) das medidas é dada pelo desvio padrão. Portanto, podemos estimar o valor da média e do desvio padrão de um conjunto de medidas a partir do seu histograma, somente observando o valor central do mesmo e a largura do histograma  $a$ , aproximadamente,  $2/3$  de sua altura máxima, conforme mostra a figura 5.4. Uma discussão mais formal sobre essa interpretação do significado do valor central e da largura de um histograma pode ser encontrada no capítulo 7 da apostila “Introdução à Teoria de Erros”.