Física Experimental III e IV

Estatística

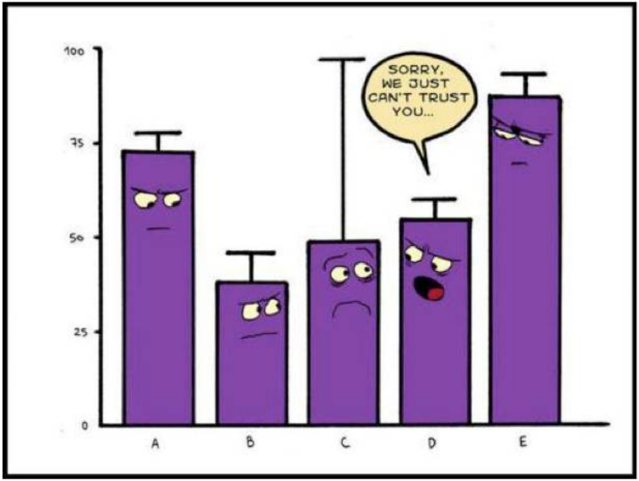
Página da disciplina:
https://edisciplinas.usp.br/course/view.php?id=66863

2019

Sumário

Sumário

Incertezas



Revendo alguns conceitos sobre incertezas

Sumário

O método científico

- A verificação e falsificação - Einstein: "No amount of experimentation can ever prove me right; a single experiment can prove me wrong."

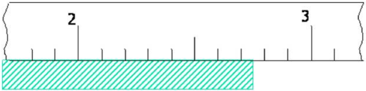
http://en.wikipedia.org/wiki/Scientific_method

http://www.unicamp.br/∼chibeni/textosdidaticos/metodocientifico.pdf

O método científico

- Uma medida é sempre uma comparação com um padrão
- Sujeita a imperfeições e limitações
- Algarismos significativos
  - Todos que tenho certeza + primeiro duvidoso (estimado)

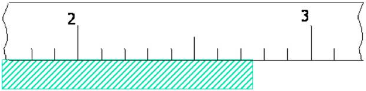
- Uma medida é sempre uma comparação com um padrão
- Sujeita a imperfeições e limitações
- Algarismos significativos
  - Todos que tenho certeza + primeiro duvidoso (estimado)



$$L = 2,74$$

O método científico

- Uma medida é sempre uma comparação com um padrão
- Sujeita a imperfeições e limitações
- Algarismos significativos
  - Todos que tenho certeza + primeiro duvidoso (estimado)

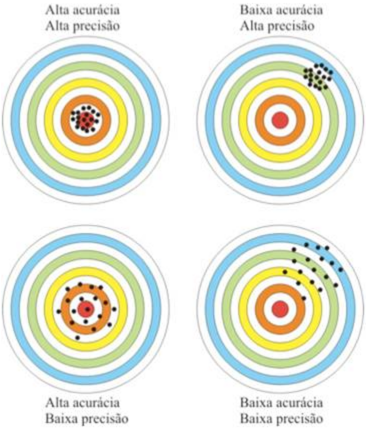


$$L = 2,74$$

- 2 e 7 tenho "certeza"
- 4 é uma estimativa → duvidoso

Acurácia e precisão

- **Precisão**: Relacionada à flutuação entre uma medida e outra
- **Acurácia**: Quão próximo você está do valor verdadeiro



Erro e incerteza

- Erro = valor verdadeiro - valor medido
  - Toda medida experimental apresenta um erro
  - O valor do erro não pode ser conhecido
  - Existem dois tipos de erro, um relacionado à precisão e outro, à acurácia
- Incerteza = melhor estimativa do valor do erro

Representando uma medida

- Faz-se a medida e avalia-se a incerteza
- Escreve-se a incerteza com, no máximo, 2 algarismos significativos
- A grandeza acompanha a precisão da incerteza
- Exemplo:
  - Obtive estes valores na calculadora

$$\text{Tempo médio} = 2{,}8764536952 \text{ s}$$

$$\text{Incerteza} = 0{,}0456485323 \text{ s}$$

  - Escrevo o resultado como:

Representando uma medida

- Faz-se a medida e avalia-se a incerteza
- Escreve-se a incerteza com, no máximo, 2 algarismos significativos
- A grandeza acompanha a precisão da incerteza
- Exemplo:
  - Obtive estes valores na calculadora

$$\text{Tempo médio} = 2{,}8764536952 \text{ s}$$

$$\text{Incerteza} = 0{,}0456485323 \text{ s}$$

  - Escrevo o resultado como:

$$\textbf{Tempo médio} = \mathbf{(2{,}876 \pm 0{,}046)\ s}$$

ou

$$\textbf{Tempo médio} = \mathbf{(2{,}88 \pm 0{,}05)\ s}$$

Sumário

Estimando incertezas: alguns conceitos de estatística

- Repetição de um experimento como ferramenta de avaliação da sua precisão
- Quanto mais eu repito, mais preciso se torna o valor médio
- Lei dos grandes números: se n tende ao infinito, o valor médio ($\bar{y}$) tende ao valor verdadeiro ($\tilde{y}$)
  - Não havendo problemas de acurácia

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

se $n \to \infty$, $\bar{y} \to \tilde{y}$

Desvio padrão

- Avaliação da flutuação dos dados em torno da média da amostra (por não conhecermos o valor verdadeiro)
- Não reflete problemas de acurácia
- O desvio padrão é o correspondente à incerteza estatística de uma única medida realizada
- Cada medida, além da incerteza instrumental, possui uma incerteza estatística dada pelo desvio padrão

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y})^2} \sim \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Desvio padrão da média

- De um conjunto de medidas, obtemos o seu valor médio
- Agora suponha que possamos repetir esse conjunto de medidas k vezes e, em cada caso, obtém-se um valor médio
- Incerteza estatística (precisão) do **valor médio** de uma amostra

$$\sigma_m = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\bar{y}_i - \tilde{y})^2} \qquad\qquad \sigma_m = \frac{\sigma}{\sqrt{n}}$$

Sumário

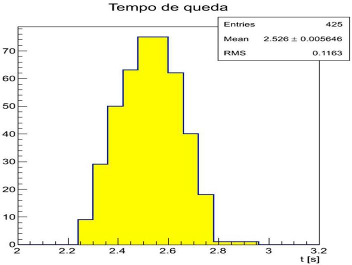
Revendo a análise de queda livre do Pelletron

- Medida de tempo de queda de balões de água
- Quase quinhentas medidas
- Análise estatística
- A aceleração obtida é compatível com a gravidade?

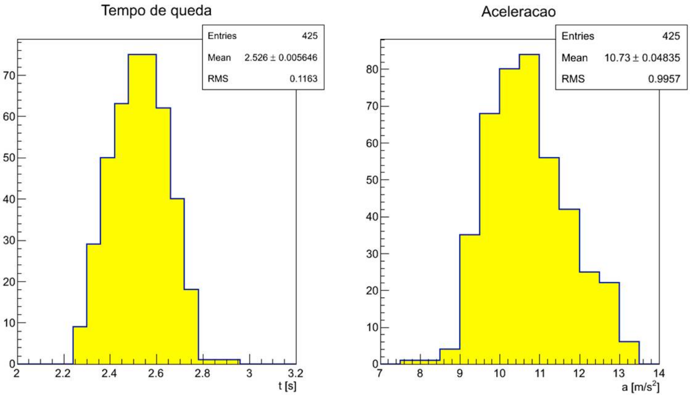
$$g_{IAG} = 9.7864\ \mathrm{m/s^2}$$

Medidas realizadas

Medida	Tempo (s)	Aceleração (m/s^2)
1	2,46	11,2
2	2,61	9,98
...	...	...
$\sim$500	2,73	9,12



Histogramas



Desvio padrão das medidas

- O desvio padrão é uma estimativa de quanto cada medida flutua em torno do valor médio da amostra
- Estimativa da incerteza de cada medida

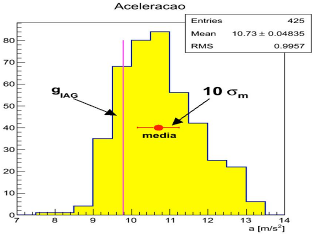
$$\sigma_{tempo} = 0,12 \text{ s}$$

$$\sigma_{acel} = 1,0 \text{ m/s}^2$$

Medida	Tempo (s)	Aceleração (m/s^2)
1	2,46	11,2
2	2,61	9,98
...	...	...
∼500	2,73	9,12

Acurácia e precisão da medida

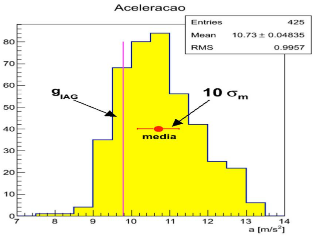
- Os valores de aceleração são precisos e acurados?
- **Precisão**:
  - Desvio padrão: $\sigma_{acel} = 1,0 \text{ m/s}^2$
  - ~ 10 % do valor de g
  - $g_{IAG} = 9,7864 \text{ m/s}^2$
  - $g_{medio} = 10,73 \pm 0,05 \text{ m/s}^2$



Acurácia e precisão da medida

- Os valores de aceleração são precisos e acurados?
- **Precisão**:
  - Desvio padrão: $\sigma_{acel} = 1,0\ \mathrm{m/s^2}$
  - $\sim 10\ \%$ do valor de g
  - $g_{IAG} = 9,7864\ \mathrm{m/s^2}$
  - $g_{medio} = 10,73 \pm 0,05\ \mathrm{m/s^2}$
- **Acurácia**:

$$\frac{g_{medio} - g_{IAG}}{\sigma_m} = 19$$



Revendo a acurácia do experimento

- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Hipóteses teóricas**
  - Desprezamos a resistência do ar
    - Se fosse importante iria diminuir a aceleração e não aumentar

Revendo a acurácia do experimento

- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Hipóteses teóricas**
  - Desprezamos a resistência do ar
    - Se fosse importante iria diminuir a aceleração e não aumentar
  - Velocidade inicial diferente de zero
    - Equação horária

$$y = v_0 t + \frac{1}{2} g t^2$$

    - Para a aceleração ser igual ao valor do IAG, teríamos que ter

$$v_0 = 1,1\,\mathrm{m/s}$$

    - Valor muito elevado se comparado ao método utilizado para lançar as bolas

Revendo a acurácia do experimento

- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Hipóteses teóricas**
  - Desprezamos a resistência do ar
    - Se fosse importante iria diminuir a aceleração e não aumentar
  - Velocidade inicial diferente de zero
    - Equação horária
$$y = v_0 t + \frac{1}{2} g t^2$$
    - Para a aceleração ser igual ao valor do IAG, teríamos que ter
$$v_0 = 1,1\,\mathrm{m/s}$$
    - Valor muito elevado se comparado ao método utilizado para lançar as bolas
  - A revisão das hipóteses teóricas não parece resolver a discrepância

Revendo a acurácia do experimento

- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Rever o procedimento experimental**
  - O disparo do cronômetro foi auditivo
    - Som tem velocidade de $\sim$340 m/s
    - Torre tem altura de 34 metros
    - Ouvimos o som 0,1 segundo depois que a bola começa a cair

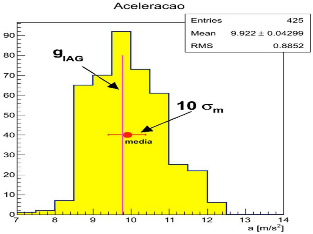
Revendo a acurácia do experimento

- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Rever o procedimento experimental**
  - O disparo do cronômetro foi auditivo
    - Som tem velocidade de $\sim$340 m/s
    - Torre tem altura de 34 metros
    - Ouvimos o som 0,1 segundo depois que a bola começa a cair
    - Ou seja, o tempo medido é sistematicamente menor que o tempo de queda por aproximadamente 0,1 segundo
    - O que acontece se somarmos 0,1 segundo em todos os tempos de queda?

Tentando corrigir problemas de acurácia

- Acrescentando 0,1 segundo em todos os tempos
- **Precisão**:
  - Desvio padrão: $\sigma_{acel} = 0,9$ m/s^2
  - ~ 10 % do valor de g
  - $g_{IAG} = 9,7864$ m/s^2
  - $g_{medio} = 9,92 \pm 0,04$ m/s^2
- **Acurácia**:

$$\frac{g_{medio} - g_{IAG}}{\sigma_m} = 3$$



Conclusões

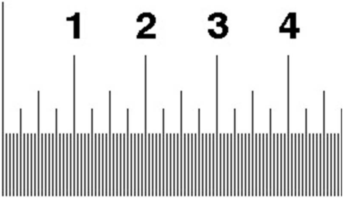
- A repetição exaustiva do experimento permitiu realizar uma análise estatística que evidenciava um problema no procedimento adotado para analisar os dados
- Isso só foi possível porque essa repetição permitiu avaliar as incertezas envolvidas, principalmente a incerteza na aceleração medida
- Em muitas situações não podemos repetir o experimento à exaustão
  - custa caro, leva muito tempo, etc.
- Como proceder se fizermos apenas uma medida de tempo?
- Qual a incerteza no tempo e aceleração?

E se não for possível repetir?

- Nas medidas diretas, tente estimar qual seria a flutuação obtida caso você repetisse o experimento

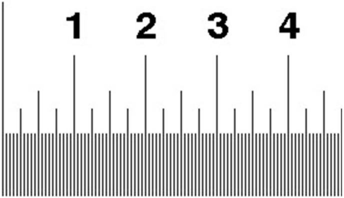
E se não for possível repetir?

- Nas medidas diretas, tente estimar qual seria a flutuação obtida caso você repetisse o experimento
- Em instrumentos com escalas simples desenhadas, como réguas, em geral utiliza-se metade da menor divisão



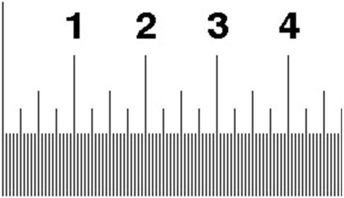
E se não for possível repetir?

- Nas medidas diretas, tente estimar qual seria a flutuação obtida caso você repetisse o experimento
- Em instrumentos com escalas simples desenhadas, como réguas, em geral utiliza-se metade da menor divisão
- Para instrumentos digitais a acurácia é fornecida no manual do instrumento



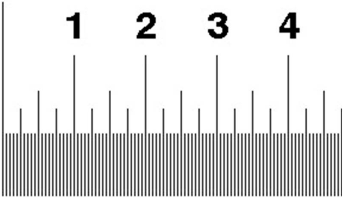
E se não for possível repetir?

- Nas medidas diretas, tente estimar qual seria a flutuação obtida caso você repetisse o experimento
- Em instrumentos com escalas simples desenhadas, como réguas, em geral utiliza-se metade da menor divisão
- Para instrumentos digitais a acurácia é fornecida no manual do instrumento
- Avalie a precisão humana
  - Por exemplo, o tempo de reação para disparar e parar um cronômetro



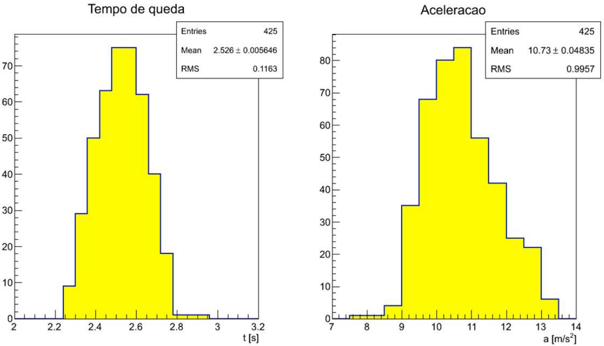
E se não for possível repetir?

- Nas medidas diretas, tente estimar qual seria a flutuação obtida caso você repetisse o experimento
- Em instrumentos com escalas simples desenhadas, como réguas, em geral utiliza-se metade da menor divisão
- Para instrumentos digitais a acurácia é fornecida no manual do instrumento
- Avalie a precisão humana
  - Por exemplo, o tempo de reação para disparar e parar um cronômetro
- E grandezas derivadas?
  - Propagação de incertezas



Sumário

Repetição de um experimento



- Mede-se a grandeza várias vezes. Neste caso, o tempo de queda.
- Calcula-se a grandeza derivada para cada medida e estuda-se sua distribuição

$$g = 2\frac{h}{t^2}$$

E se ...

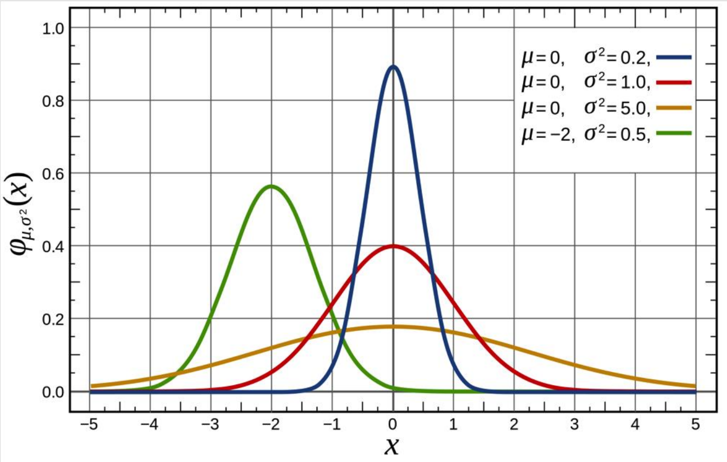
- ... não pudermos repetir este experimento de modo a calcular desvios padrão?

E se ...

- ... não pudermos repetir este experimento de modo a calcular desvios padrão?
- ... a grandeza derivada depender de muitas outras grandezas diferentes?

A gaussiana

$$G_{\mu,\sigma}(x) = G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$



Uma propriedade importante

- Cálculo de valores médios

$$\langle (x-\mu)^n \rangle = \int_{-\infty}^{\infty} (x-\mu)^n G(x) dx$$

Uma propriedade importante

- Cálculo de valores médios

$$\langle (x-\mu)^n \rangle = \int_{-\infty}^{\infty} (x-\mu)^n G(x) dx$$

- Fazendo uma mudança de variável

$$y = \frac{x-\mu}{\sigma}$$

Uma propriedade importante

- Cálculo de valores médios

$$\langle (x-\mu)^n \rangle = \int_{-\infty}^{\infty} (x-\mu)^n G(x) dx$$

- Fazendo uma mudança de variável

$$y = \frac{x-\mu}{\sigma} \quad \Rightarrow \quad \langle (x-\mu)^n \rangle = \frac{1}{\sqrt{2\pi}} \sigma^n \int_{-\infty}^{\infty} y^n \exp\left[-\frac{1}{2} y^2\right] dy$$

Uma propriedade importante

- Definindo

$$I_n = \int_{-\infty}^{\infty} y^n \exp\left[-\frac{1}{2}y^2\right] dy$$

Uma propriedade importante

- Definindo

$$I_n = \int_{-\infty}^{\infty} y^n \exp\left[-\frac{1}{2}y^2\right] dy$$

- Consultando uma tabela de integrais

$$I_0 = \sqrt{2\pi}$$

$$I_n = 0 \text{ se } n \text{ for ímpar}$$

$$\frac{I_n}{I_{n-2}} = (n-1) \text{ para } n \text{ par e } n > 1$$

Uma propriedade importante

- Substituindo

$$\langle (x-\mu)^n \rangle = \frac{I_n}{I_0}\sigma^n$$

Uma propriedade importante

- Substituindo
$$\langle (x-\mu)^n \rangle = \frac{I_n}{I_0}\sigma^n$$
- Exemplo:
$$\left\langle (x-\mu)^2 \right\rangle = \frac{I_2}{I_0}\sigma^2 = (2-1)\sigma^2 = \sigma^2$$

Uma propriedade importante

- Substituindo

$$\langle (x-\mu)^n \rangle = \frac{I_n}{I_0}\sigma^n$$

- Exemplo:

$$\left\langle (x-\mu)^2 \right\rangle = \frac{I_2}{I_0}\sigma^2 = (2-1)\sigma^2 = \sigma^2$$

$$\left\langle (x-\mu)^4 \right\rangle = \frac{I_4}{I_0}\sigma^4 = \frac{I_4}{I_2}\frac{I_2}{I_0}\sigma^4 = (4-1)(2-1)\sigma^4 = 3\sigma^4$$

Propagação de incertezas

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

Propagação de incertezas

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

$$\sigma_z^2 = \left\langle (z - \mu_z)^2 \right\rangle$$

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

$$\sigma_z^2 = \left\langle (z - \mu_z)^2 \right\rangle$$

- Substituindo a fórmula no cálculo da variância:

$$\sigma_z^2 = \left\langle [x + y - (\mu_x + \mu_y)]^2 \right\rangle$$

Propagação de incertezas

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

$$\sigma_z^2 = \left\langle (z - \mu_z)^2 \right\rangle$$

- Substituindo a fórmula no cálculo da variância:

$$\sigma_z^2 = \left\langle [x + y - (\mu_x + \mu_y)]^2 \right\rangle$$

$$\sigma_z^2 = \left\langle [(x - \mu_x) + (y - \mu_y)]^2 \right\rangle$$

Propagação de incertezas

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

$$\sigma_z^2 = \left\langle (z - \mu_z)^2 \right\rangle$$

- Substituindo a fórmula no cálculo da variância:

$$\sigma_z^2 = \left\langle [x + y - (\mu_x + \mu_y)]^2 \right\rangle$$

$$\sigma_z^2 = \left\langle [(x - \mu_x) + (y - \mu_y)]^2 \right\rangle$$

$$\sigma_z^2 = \left\langle (x - \mu_x)^2 \right\rangle + \left\langle (y - \mu_y)^2 \right\rangle + 2 \left\langle (x - \mu_x)(y - \mu_y) \right\rangle$$

Propagação de incertezas

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

$$\sigma_z^2 = \left\langle (z - \mu_z)^2 \right\rangle$$

- Substituindo a fórmula no cálculo da variância:

$$\sigma_z^2 = \left\langle [x + y - (\mu_x + \mu_y)]^2 \right\rangle$$

$$\sigma_z^2 = \left\langle [(x - \mu_x) + (y - \mu_y)]^2 \right\rangle$$

$$\sigma_z^2 = \left\langle (x - \mu_x)^2 \right\rangle + \left\langle (y - \mu_y)^2 \right\rangle + 2 \left\langle (x - \mu_x)(y - \mu_y) \right\rangle$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 + 2\mathrm{cov}_{xy} \qquad \mathrm{cov}_{xy} = \left\langle (x - \mu_x)(y - \mu_y) \right\rangle$$

Fórmula geral

- Seja uma grandeza calculada através de uma função qualquer

$$y = f(x_1, x_2, \ldots x_n) = f(\tilde{x})$$

Fórmula geral

- Seja uma grandeza calculada através de uma função qualquer

$$y = f(x_1, x_2, \dots x_n) = f(\tilde{x})$$

$$\sigma_y^2 = \left\langle [f(\tilde{x}) - f(\mu_x)]^2 \right\rangle$$

Fórmula geral

- Seja uma grandeza calculada através de uma função qualquer

$$y = f(x_1, x_2, \ldots x_n) = f(\tilde{x})$$

$$\sigma_y^2 = \left\langle \left[f(\tilde{x}) - f(\mu_x) \right]^2 \right\rangle$$

- Fazendo uma expansão em série de Taylor da função em primeira ordem:

$$f(\tilde{x}) = f(\mu_x) + \sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i)$$

Fórmula geral

- Seja uma grandeza calculada através de uma função qualquer

$$y = f(x_1, x_2, \dots x_n) = f(\tilde{x})$$

$$\sigma_y^2 = \left\langle [f(\tilde{x}) - f(\mu_x)]^2 \right\rangle$$

- Fazendo uma expansão em série de Taylor da função em primeira ordem:

$$f(\tilde{x}) = f(\mu_x) + \sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i)$$

$$\sigma_y^2 = \left\langle \left[f(\mu_x) + \sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) - f(\mu_x) \right]^2 \right\rangle$$

Fórmula geral

- Seja uma grandeza calculada através de uma função qualquer

$$y = f(x_1, x_2, \ldots x_n) = f(\tilde{x})$$

$$\sigma_y^2 = \left\langle [f(\tilde{x}) - f(\mu_x)]^2 \right\rangle$$

- Fazendo uma expansão em série de Taylor da função em primeira ordem:

$$f(\tilde{x}) = f(\mu_x) + \sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i)$$

$$\sigma_y^2 = \left\langle \left[f(\mu_x) + \sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) - f(\mu_x) \right]^2 \right\rangle$$

$$\sigma_y^2 = \left\langle \left[\sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) \right]^2 \right\rangle$$

Fórmula geral

$$\sigma_y^2 = \left\langle \left[\sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) \right]^2 \right\rangle$$

- Pegando esta equação e abrindo explicitamente o quadrado, podemos reagrupar os termos de tal forma que:

Fórmula geral

$$\sigma_y^2 = \left\langle \left[\sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) \right]^2 \right\rangle$$

- Pegando esta equação e abrindo explicitamente o quadrado, podemos reagrupar os termos de tal forma que:

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i} \right)^2 \left\langle (x_i - \mu_i)^2 \right\rangle + \sum_i \sum_j \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \langle (x_i - \mu_i)(x_j - \mu_j) \rangle$$

Fórmula geral

$$\sigma_y^2 = \left\langle \left[\sum_i \frac{\partial f}{\partial x_i}(x_i - \mu_i) \right]^2 \right\rangle$$

- Pegando esta equação e abrindo explicitamente o quadrado, podemos reagrupar os termos de tal forma que:

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i} \right)^2 \left\langle (x_i - \mu_i)^2 \right\rangle + \sum_i \sum_j \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \langle (x_i - \mu_i)(x_j - \mu_j) \rangle$$

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i} \right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \mathrm{cov}_{ij}$$

Fórmula geral

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}_{ij}$$

- Esta fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

Fórmula geral

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}_{ij}$$

- Esta fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

$$\sigma_y^2 = \Gamma^T \Sigma \Gamma$$

Fórmula geral

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}_{ij}$$

- Esta fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

$$\sigma_y^2 = \Gamma^T \Sigma \Gamma$$

$$\Gamma = \begin{pmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \dots \\ \partial_{x_n} \end{pmatrix} \ \text{com}\ \partial_{x_i} = \frac{\partial f}{\partial x_i}$$

Fórmula geral

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}_{ij}$$

- Esta fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

$$\sigma_y^2 = \Gamma^T \Sigma \Gamma$$

$$\Gamma = \begin{pmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \dots \\ \partial_{x_n} \end{pmatrix} \ \text{com } \partial_{x_i} = \frac{\partial f}{\partial x_i} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}_{12} & \cdots & \mathrm{cov}_{1n} \\ \mathrm{cov}_{21} & \sigma_2^2 & \cdots & \mathrm{cov}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}_{n1} & \mathrm{cov}_{n2} & \cdots & \sigma_n^2 \end{pmatrix}$$

Fórmula geral

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j}\mathrm{cov}_{ij}$$

- Esta fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

$$\sigma_y^2 = \Gamma^T \Sigma \Gamma$$

$$\Gamma = \begin{pmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \dots \\ \partial_{x_n} \end{pmatrix} \ \text{com } \partial_{x_i} = \frac{\partial f}{\partial x_i} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}_{12} & \cdots & \mathrm{cov}_{1n} \\ \mathrm{cov}_{21} & \sigma_2^2 & \cdots & \mathrm{cov}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}_{n1} & \mathrm{cov}_{n2} & \cdots & \sigma_n^2 \end{pmatrix}$$

Σ é chamada de matriz de covariância

Exemplo

- Mediu-se o tempo de queda de um balão de uma altura $h = 34,0 \pm 0,5$ m e obteve-se $t = 2,65 \pm 0,20$ s. Qual a aceleração?

Exemplo

- Mediu-se o tempo de queda de um balão de uma altura $h = 34,0 \pm 0,5$ m e obteve-se $t = 2,65 \pm 0,20$ s. Qual a aceleração?

$$a = 2\frac{h}{t^2}$$

Exemplo

- Mediu-se o tempo de queda de um balão de uma altura $h = 34,0 \pm 0,5$ m e obteve-se $t = 2,65 \pm 0,20$ s. Qual a aceleração?

$$a = 2\frac{h}{t^2}$$

$$\sigma_a^2 = \left(\frac{\partial f}{\partial h}\right)^2 \sigma_h^2 + \left(\frac{\partial f}{\partial t}\right)^2 \sigma_t^2$$

Exemplo

- Mediu-se o tempo de queda de um balão de uma altura $h = 34,0 \pm 0,5$ m e obteve-se $t = 2,65 \pm 0,20$ s. Qual a aceleração?

$$a = 2\frac{h}{t^2}$$

$$\sigma_a^2 = \left(\frac{\partial f}{\partial h}\right)^2 \sigma_h^2 + \left(\frac{\partial f}{\partial t}\right)^2 \sigma_t^2 = \left(\frac{2}{t^2}\right)^2 \sigma_h^2 + \left(-\frac{4h}{t^3}\right)^2 \sigma_t^2$$

Exemplo

- Mediu-se o tempo de queda de um balão de uma altura $h = 34,0 \pm 0,5$ m e obteve-se $t = 2,65 \pm 0,20$ s. Qual a aceleração?

$$a = 2\frac{h}{t^2}$$

$$\sigma_a^2 = \left(\frac{\partial f}{\partial h}\right)^2 \sigma_h^2 + \left(\frac{\partial f}{\partial t}\right)^2 \sigma_t^2 = \left(\frac{2}{t^2}\right)^2 \sigma_h^2 + \left(-\frac{4h}{t^3}\right)^2 \sigma_t^2$$

- Substituindo os valores

$$a = 9,7 \pm 1,5 \; \mathrm{m/s}^2$$

E a covariância?

- A covariância é dada por

$$\mathrm{cov}_{xy} = \langle (x - \mu_x)(y - \mu_y) \rangle$$

E a covariância?

- A covariância é dada por

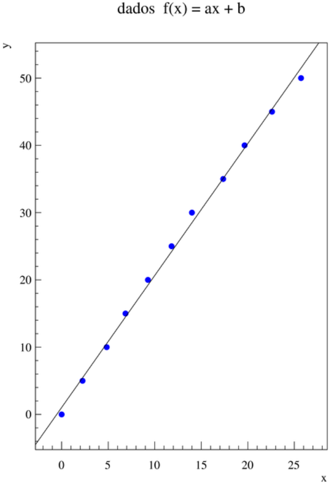
$$\mathrm{cov}_{xy} = \langle (x - \mu_x)(y - \mu_y) \rangle$$

- Se as variáveis x e y são totalmente independentes, elas podem flutuar de maneira independente e a somatória tende a se anular, de modo que:

$$\mathrm{cov}_{xy} = 0$$

Quando grandezas são dependentes entre si?

- A situação mais comum é no ajuste de uma curva.
  - Os parâmetros ajustados possuem, em geral, covariância, pois estão vinculados entre si através dos pontos experimentais.
  - Se eu forçar o coeficiente linear para um valor maior eu devo diminuir o coeficiente angular para continuar passando pelos pontos experimentais.



Um exemplo: a curva característica da pilha

- Matriz de covariância do ajuste

$$y = [0] + [1] * x = a + bx$$

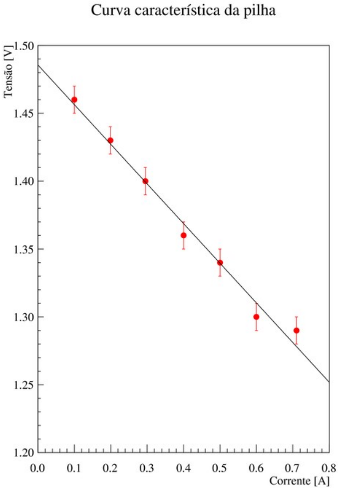
Resultados do ajuste

Número de parâmetros	2
Chi²	3.42626
Número de graus de liberdade	5

parâmetro	Valor	Incerteza
0	1.4856	0.00837212
1	-0.292165	0.0186493

Matriz de covariância

$$\begin{bmatrix} 7.00924E\text{-}05 & -0.000139318 \\ -0.000139318 & 0.000347797 \end{bmatrix}$$

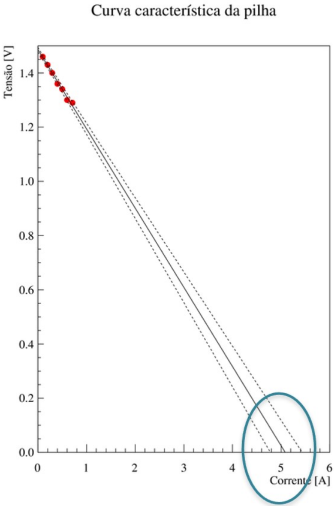


Um exemplo: a curva característica da pilha

- Qual a corrente máxima?
  - Extrapola para tensão = 0

$$i_{max} = -\frac{a}{b} = -\frac{[0]}{[1]}$$

- Qual a incerteza na corrente máxima?



Um exemplo: a curva característica da pilha

- Qual a incerteza na corrente máxima?

$$i_{max} = -\frac{a}{b}$$

Um exemplo: a curva característica da pilha

- Qual a incerteza na corrente máxima?

$$i_{max} = -\frac{a}{b}$$

$$\sigma^2_{i_{max}} = \left(\frac{\partial i_{max}}{\partial a}\right)^2 \sigma^2_a + \left(\frac{\partial i_{max}}{\partial b}\right)^2 \sigma^2_b + 2\frac{\partial i_{max}}{\partial a}\frac{\partial i_{max}}{\partial b}\mathrm{cov}_{ab}$$

- Calculando as derivadas, temos:

$$\sigma^2_{i_{max}} = \frac{1}{b^2}\sigma^2_a + \frac{a^2}{b^4}\sigma^2_b - 2\frac{a}{b^3}\mathrm{cov}_{ab}$$

- E é só substituir os valores

Faz diferença utilizar a covariância?

- Depende da situação
- Vamos estudar a covariância em detalhes mais tarde
  - Por enquanto vamos nos acostumar com a existência dela e utilizar quando necessário.
  - Sempre que formos utilizar parâmetros de um ajuste para fazer contas, fiquem atentos à covariância.

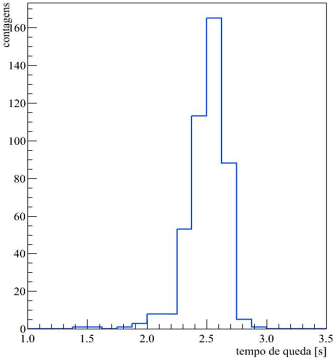
Cuidados importantes

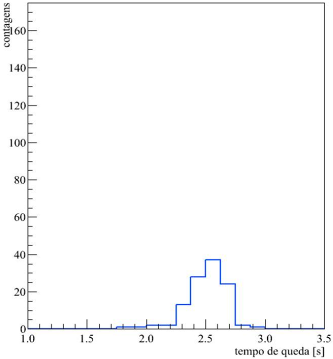
- Consideramos que todas as grandezas envolvidas possuem distribuições gaussianas
  - E se não possuírem?
    - Veremos como lidar com isto em breve
- Fizemos uma expansão em Taylor de primeira ordem.
  - Esta expansão é sempre razoável?
  - E se não for? O que devemos fazer?
    - Respostas em um futuro próximo

Sumário

Repetição de um experimento

- Histogramas simples de contagens não são interessantes pois a comparação entre dois conjuntos de dados diferentes nem sempre é possível de forma direta
  - Depende do número de entradas no histograma





- Define-se a probabilidade de se obter um determinado resultado como sendo a relação entre o número de vezes que esse resultado é obtido dividido pelo número total de dados, quando este é suficientemente grande

$$P(R) = \lim_{N\to\infty} \frac{N(R)}{N}$$

Probabilidade

- Define-se a probabilidade de se obter um determinado resultado como sendo a relação entre o número de vezes que esse resultado é obtido dividido pelo número total de dados, quando este é suficientemente grande

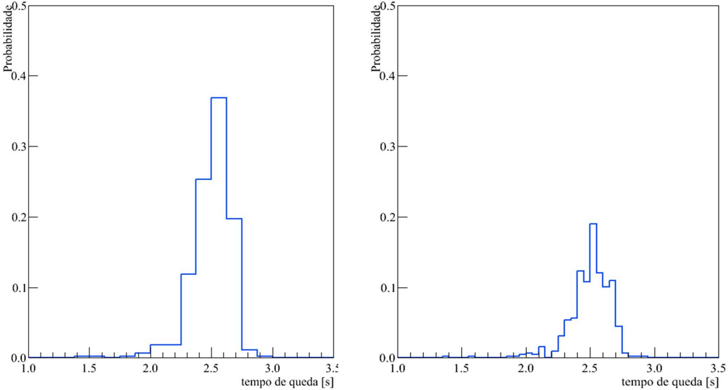
$$P(R) = \lim_{N \to \infty} \frac{N(R)}{N}$$

- Em um histograma definem-se canais e contam-se quantas ocorrências tem-se naquele canal

$$N(R) = N(x, x + \Delta x)$$

  - Em um histograma a probabilidade depende da escolha do tamanho do canal do histograma

Histogramas de probabilidade



Sumário

Densidade de probabilidade

- A Função Densidade de Probabilidade (FDP) é definida de tal forma que a probabilidade de encontrar um resultado em um intervalo é tal que

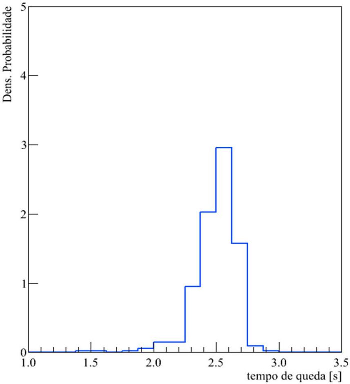
$$P(x, x + \Delta x) = \int_x^{x+\Delta x} H(x')dx'$$

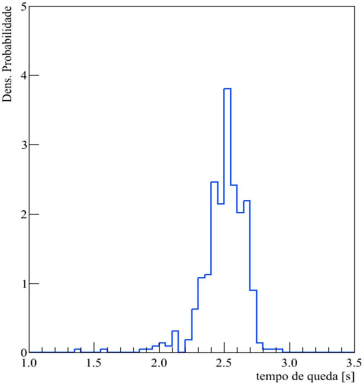
- Ou seja

$$H(x) = \lim_{\Delta x \to 0} \frac{P(x, x + \Delta x)}{\Delta x}$$

  - A densidade de probabilidade não depende da escolha do tamanho do canal em um histograma
    - A menos de flutuações por conta da amostra ser limitada

Histogramas de densidade de probabilidade





Algumas características das densidades de probabilidade

- Por ter significado de uma densidade é sempre positiva

Algumas características das densidades de probabilidade

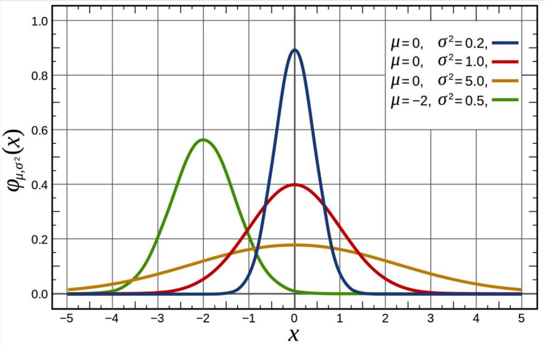
- Por ter significado de uma densidade é sempre positiva
- Como a probabilidade é sempre um número entre 0 e 1, a integral da densidade de probabilidade em todo o espaço deve ser a probabilidade de ter um evento, quaisquer que sejam suas características, ou seja, 100%. Deste modo
$$\int_{-\infty}^{\infty} H(x')dx' = 1$$

Sumário

1. Incertezas
   - Representação de uma medida
   - Estatística
   - Exemplo
2. Propagação de incertezas
3. Tratamento estatístico
   - Função densidade de probabilidade
   - Densidade de probabilidade gaussiana
4. Mais um pouco de estatística
   - O χ^2
   - Exemplos
5. Propagação de incertezas

A gaussiana

$$H(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



- Será que tudo que é derivado de uma grandeza gaussiana também possui FDP gaussiana?

Experimento virtual

- Seja uma grandeza qualquer (x) que possua FDP gaussiana de valor verdadeiro μ e variância σ_μ conhecidos.

Experimento virtual

- Seja uma grandeza qualquer (x) que possua FDP gaussiana de valor verdadeiro μ e variância σ_μ conhecidos.
- Realiza-se um experimento no qual são tomadas ν medidas independentes desta grandeza e o resultado do experimento é o valor médio dessas medidas

$$\bar{x} = \frac{1}{\nu}\sum_{i=1}^{\nu} x_i$$

Experimento virtual

- Seja uma grandeza qualquer (x) que possua FDP gaussiana de valor verdadeiro μ e variância σ_μ conhecidos.
- Realiza-se um experimento no qual são tomadas ν medidas independentes desta grandeza e o resultado do experimento é o valor médio dessas medidas

$$\bar{x} = \frac{1}{\nu}\sum_{i=1}^{\nu} x_i$$

- Calcula-se também a variância "experimental" destas medidas em relação ao valor verdadeiro

$$\sigma^2 = \frac{1}{\nu}\sum_{i=1}^{\nu}(x_i - \mu)^2$$

Experimento virtual

- Seja uma grandeza qualquer (x) que possua FDP gaussiana de valor verdadeiro μ e variância σ_μ conhecidos.
- Realiza-se um experimento no qual são tomadas ν medidas independentes desta grandeza e o resultado do experimento é o valor médio dessas medidas

$$\bar{x} = \frac{1}{\nu}\sum_{i=1}^{\nu} x_i$$

- Calcula-se também a variância "experimental" destas medidas em relação ao valor verdadeiro

$$\sigma^2 = \frac{1}{\nu}\sum_{i=1}^{\nu}(x_i - \mu)^2$$

- Como são as FDPs do valor médio e da variância? São gaussianas? O valor médio da variância experimental coincide com a variância real? Qual a dependência dessas FDPs com o número de medidas realizadas, ν?

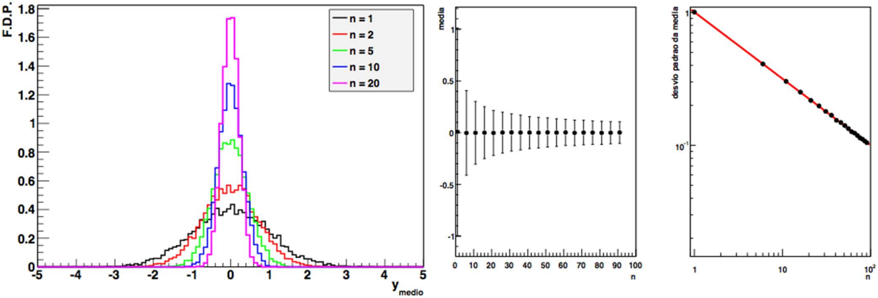
Experimento virtual

- Considere que um experimentador repita, diariamente, o experimento descrito anteriormente e calcule, para cada dia, o valor da média e o valor de χ^2 e coloque os resultados em um histograma. No final da vida dele ele vai ter repetido isso tantas vezes que é possível conhecer as FDPs dessas duas grandezas.
- Para facilitar, vamos considerar $\mu = 0$ e $\sigma_\mu = 1$. Não há perda de generalidade pois sempre podemos fazer uma mudança de variáveis do tipo:

$$\frac{x - \mu}{\sigma_\mu} \rightarrow y$$

- Por sorte existem computadores e pode-se simular este experimento computacionalmente, bem como a repetição do mesmo indefinidamente.

- Continua sendo uma gaussiana
  - A média não se altera com ν (*n* na figura)
  - Na medida em que ν (*n* na figura) aumenta, a gaussiana fica mais estreita



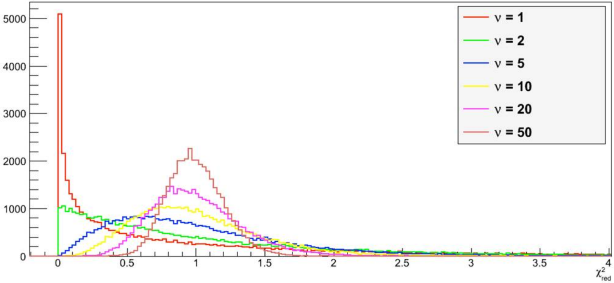
Distribuição da variância

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

- Como os termos da somatória são sempre positivos não é esperado que ela tenha média zero.
- O valor médio da variância é 1 (um), como esperado para o seu valor verdadeiro?

A distribuição da variância "experimental"

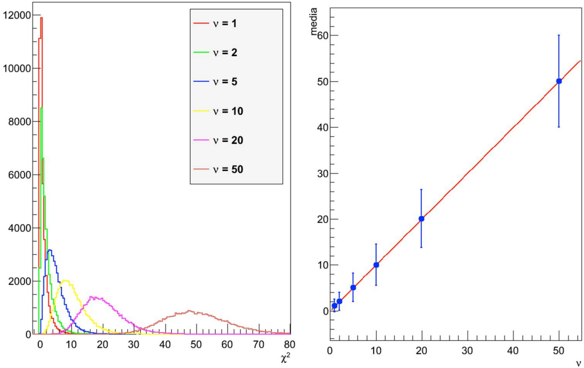
- Para um número pequeno de pontos a distribuição claramente não é gaussiana
- Para poucos pontos tem-se a tendência de SUBESTIMAR a variância dos dados



Distribuição de χ^2

$$\chi^2 = \sum_{i=1}^{\nu} \left(\frac{x_i - \mu}{\sigma_\mu} \right)^2 = \sum_{i=1}^{\nu} y_i^2$$

- A FDP do χ^2 não é gaussiana
- O valor médio da distribuição de χ^2 é ν.

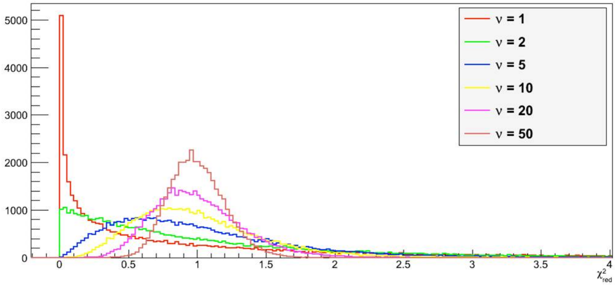


O χ^2 reduzido

- χ^2 reduzido tem a mesma distribuição da variância!!!!

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

$$\chi^2_{red} = \frac{1}{\nu}\sum_{i=1}^{\nu}\left(\frac{x_i-\mu}{\sigma_\mu}\right)^2 = \frac{1}{\nu}\sum_{i=1}^{\nu} y_i^2$$



Quem é ν?

- Relembrando, "... Imagine um experimento no qual se realizam ν medidas independentes desta grandeza e o resultado do experimento é o valor médio dessas medidas.."
  - ν é o número de medidas estatisticamente independentes no experimento. É chamado de **número de graus de liberdade** da amostra.
    - Número de graus de liberdade corresponde à quantidade de valores independentes que podem variar livremente no cálculo de uma grandeza estatística.
    - O que é independente em uma amostra?

Número de graus de liberdade

- Toma-se um conjunto de dados e calcula-se a sua variância

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

- Considere a situação na qual o valor verdadeiro da amostra não é conhecido
  - A melhor estimativa do valor verdadeiro da amostra corresponde ao valor médio da mesma
  - A variância seria calculada como:

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

  - Note que $N \rightarrow N - 1$ na expressão acima. Porque disso?

Número de graus de liberdade

- No momento em que a média é calculada, somente $N-1$ pontos são totalmente livres para variar. O último ponto da amostra necessariamente deve ter o valor:

$$x_n = N\bar{x} - \sum_{i=1}^{N-1} x_i$$

- Um dos dados da amostra não é independente dos outros. Ele está vinculado aos demais por conta do valor médio calculado.
- O número de graus de liberdade disponíveis é $\mathbf{N-1}$. Esta é uma interpretação pela qual aparece o $N-1$ no cálculo da variância quando não se conhece o valor verdadeiro da amostra.

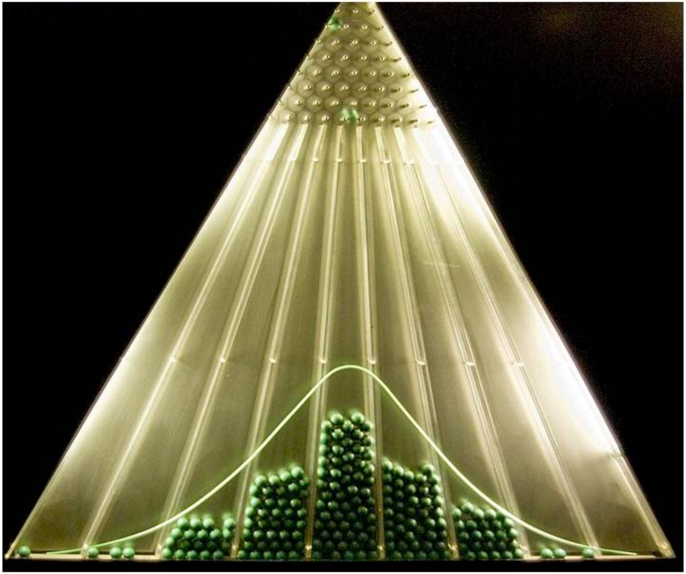
$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$$

Porque distribuições de probabilidade são importantes?

- Quando é feita uma medida, ou um conjunto de medidas, queremos saber quão provável é o resultado obtido ou quão confiável é a análise realizada.
    - A medida é compatível com o valor teórico esperado?
    - Duas medidas são compatíveis entre si?
    - O número de pontos neste gráfico é suficiente?
    - O ajuste de reta que foi feito é bom?
- Para responder estas perguntas precisamos conhecer as distribuições de probabilidade das muitas grandezas envolvidas.

Sumário

Caixa de Galton



Uma pergunta recorrente

- Quando eu sei que o χ^2 de um ajuste é bom ou ruim?

Uma pergunta recorrente

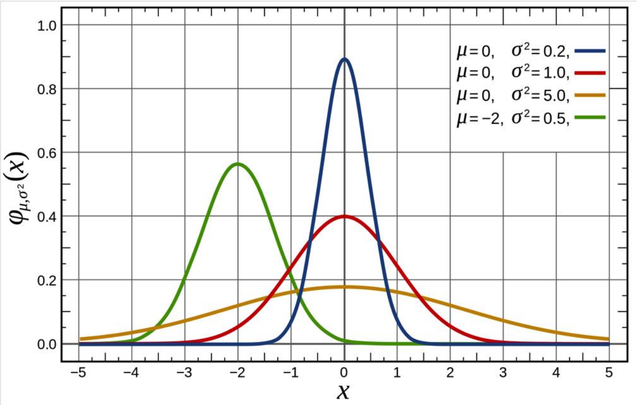
- Quando eu sei que o χ^2 de um ajuste é bom ou ruim?
  - $\chi^2_{red} \sim 1$
    - Ordem de grandeza apenas ou tem algum intervalo numérico definido?

Uma pergunta recorrente

- Quando eu sei que o χ^2 de um ajuste é bom ou ruim?
  - $\chi^2_{red} \sim 1$
    - Ordem de grandeza apenas ou tem algum intervalo numérico definido?
  - E se o χ^2 for muito grande ou muito pequeno?

- Normal ou gaussiana

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$$



Sumário

FDP de χ^2

- Vamos realizar um experimento virtual, no qual simularemos a obtenção da amostra $X_{ngl} = \{x_i\}$

FDP de χ^2

- Vamos realizar um experimento virtual, no qual simularemos a obtenção da amostra $X_{ngl} = \{x_i\}$
- A partir dessa amostra vamos calcular a variável de interesse (no nosso caso, o χ^2)

FDP de χ^2

- Vamos realizar um experimento virtual, no qual simularemos a obtenção da amostra $X_{ngl} = \{x_i\}$
- A partir dessa amostra vamos calcular a variável de interesse (no nosso caso, o χ^2)
- Vamos repetir esse experimento virtual um número muito grande de vezes de modo a obter as FDPs das variáveis estudadas.

Simplificando o problema

- Ao invés de usar a amostra X, vamos fazer uma mudança de variável tal que

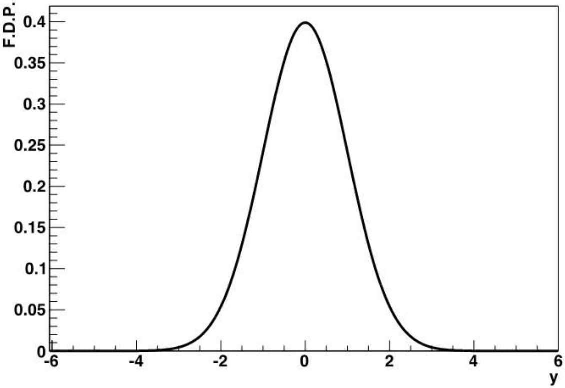
$$Y = \{y_i\} \quad \longrightarrow \quad y_i = \frac{x_i - \mu}{\sigma_\mu}$$

- Ou seja, vamos estudar uma amostra de valor verdadeiro 0 e variância 1.
    - Para FDP com médias e variâncias diferentes, basta uma mudança de escala

FDP de y

- Y segue uma distribuição normal de valor verdadeiro 0 e variância 1

$$p(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$



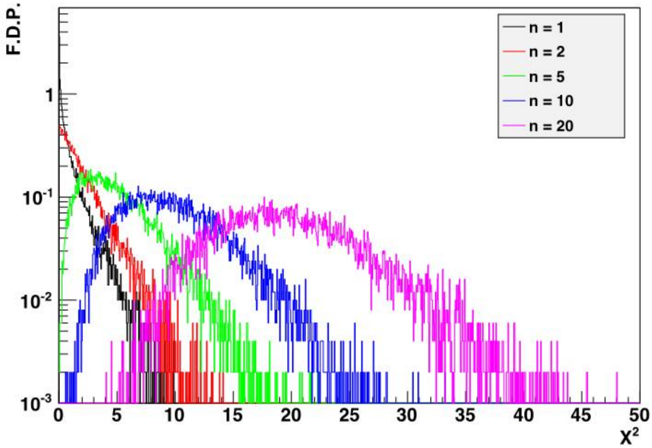
FDP de χ^2

- A função χ^2 é definida como:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma_\mu} \right)^2 = \sum_{i=1}^{n} y_i^2$$

- A FDP é obtida calculando o valor de χ^2 para cada conjunto de dados simulado

FDP de χ^2



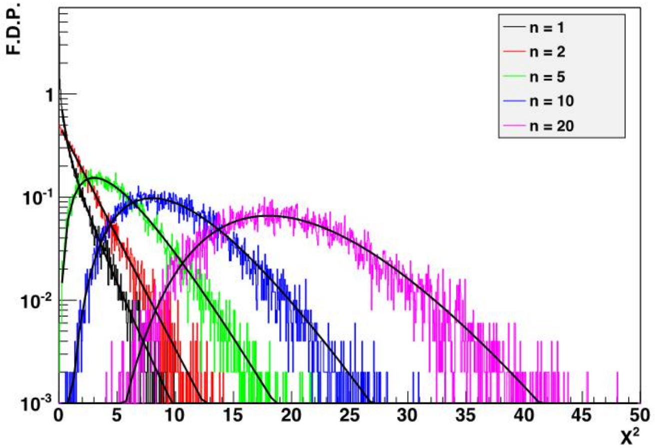
FDP de χ^2

- A FDP não segue mais uma distribuição normal, mas sim:

$$p(\xi) = \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}\xi^{\frac{n}{2}-1}e^{-\frac{\xi}{2}}$$

- Onde Γ é a função gama, n é o número de graus de liberdade e ξ, o valor de χ^2.

FDP de χ^2



FDP de χ^2_{red} e σ

- A função χ^2_{red} é definida como:

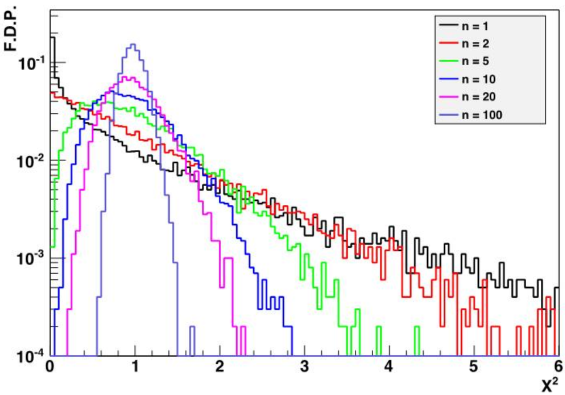
$$\chi^2_{red} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma_\mu}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2$$

- Por outro lado, a variância de um conjunto de medidas é:

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

- Ou seja, essas grandezas são muito similares e seguem a mesma FDP

FDP de χ^2_{red} e σ



- Note que quanto maior o número de graus de liberdade, mais estreita é a distribuição.

FDP de χ^2_{red} e σ

- χ^2_{red} e σ são importantes em testes de significância
  - A função χ^2_{red} é calculada quando se faz um ajuste de curvas. Como avaliar se o ajuste é bom?
  - Em uma medida estatística, como saber se a variância que estou obtendo é representativa?

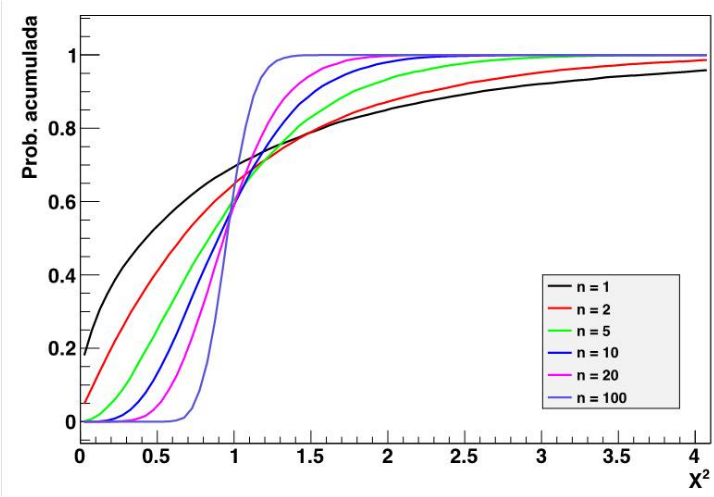
Probabilidade acumulada

- Podemos definir como probabilidade acumulada a grandeza:

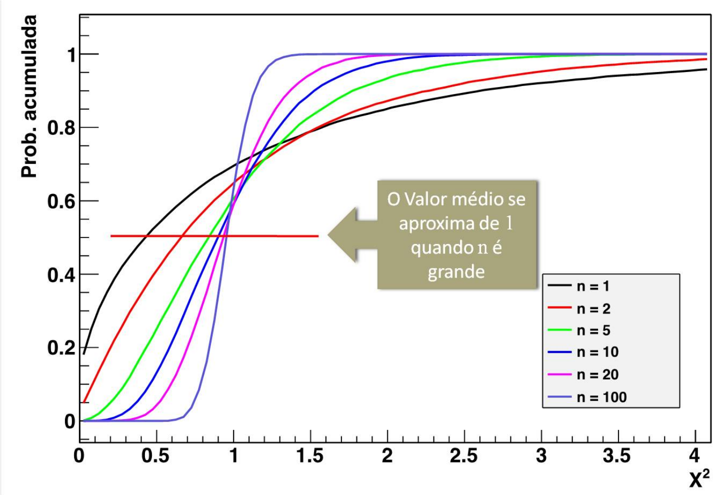
$$P(x) = \int_{-\infty}^{x} p(t)dt \longrightarrow \left\{ \begin{array}{l} \text{Probabilidade da variável } t \\ \text{assumir um valor menor ou} \\ \text{igual a } x \end{array} \right.$$

- Essa grandeza é particularmente útil para definir intervalos de confiança
  - Ex: qual o intervalo de 90% de confiança para a distribuição de χ^2_{red} de um ajuste com 5 graus de liberdade?

Probabilidade acumulada de χ^2_{red} (ou σ)

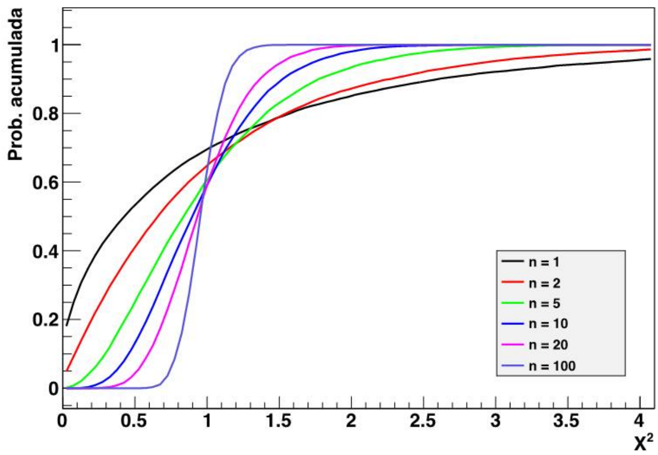


Probabilidade acumulada de χ^2_{red} (ou σ)



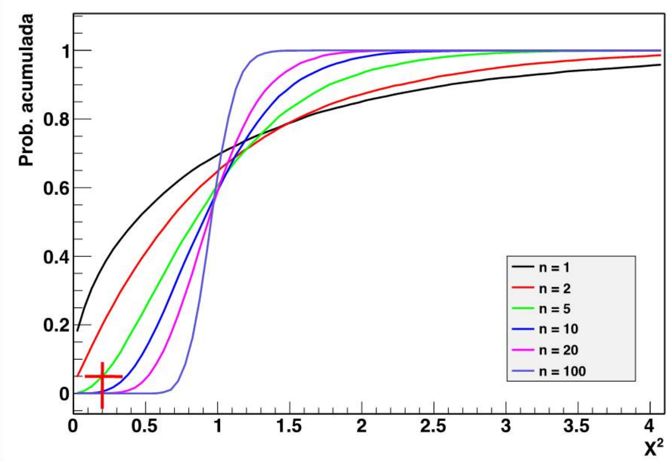
Probabilidade acumulada de χ^2_{red} (ou σ)

- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?



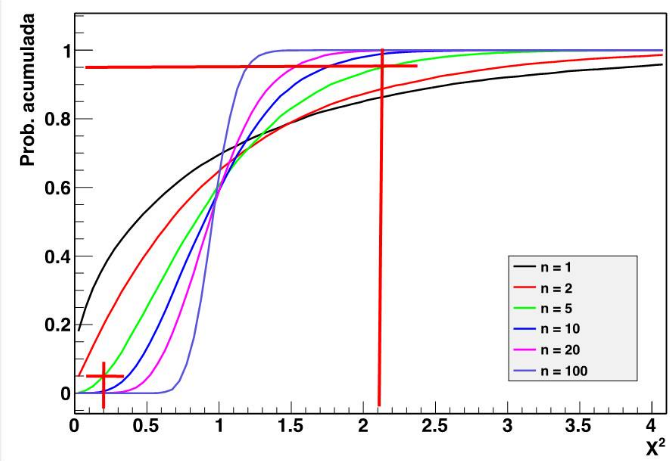
Probabilidade acumulada de χ^2_{red} (ou σ)

- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?



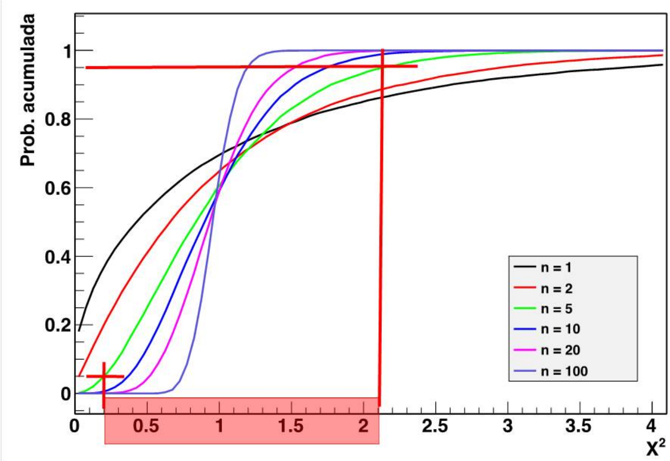
Probabilidade acumulada de χ^2_{red} (ou σ)

- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?



Probabilidade acumulada de χ^2_{red} (ou σ)

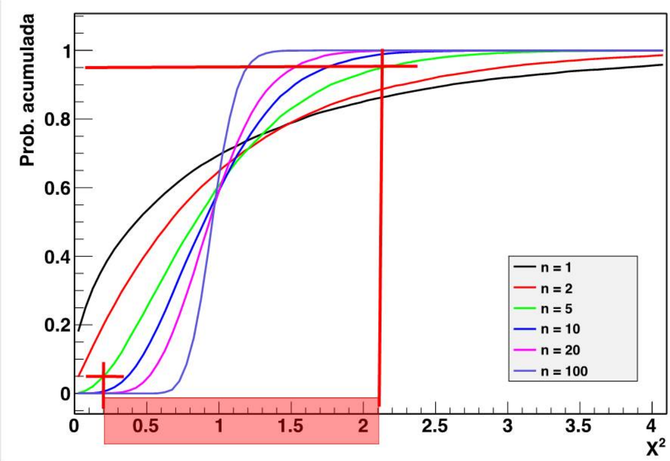
- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?



Probabilidade acumulada de χ^2_{red} (ou σ)

- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?

$$0,2 < \chi^2_{red} < 2,1$$



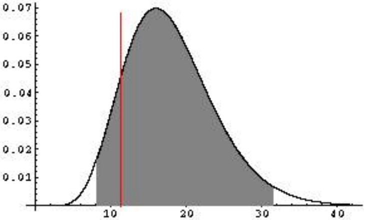
Moral da história

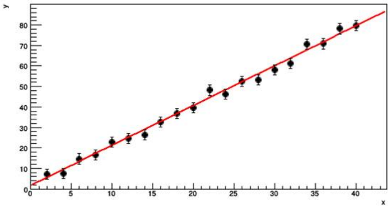
- Testes de χ^2 e análise de resíduos (análise mesmo, não apenas fazer um gráfico) constituem ferramentas poderosas na validação de resultados experimentais
- Cálculo de intervalo de confiança para χ^2, teste-z, teste-t, etc.
  - No WebROOT $\rightarrow$ Calculadoras
- E se o valor de χ^2 estiver fora do intervalo de significância?
  - Em geral:
    - Incertezas super/subestimadas
    - Modelo teórico não se aplica aos dados

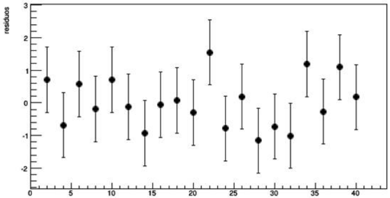
Sumário

Um bom ajuste

- $\chi^2 = 11,45$
- $ngl = 18$
- $95\%\ \text{CL} = 8,2 < \chi^2 < 32$

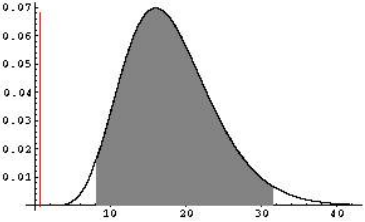


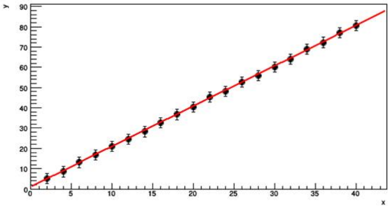


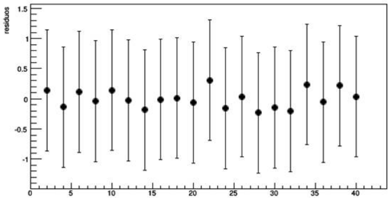


Incertezas superestimadas

- $\chi^2 = 0,45$
- $ngl = 18$
- $95\%\ \text{CL} = 8,2 < \chi^2 < 32$

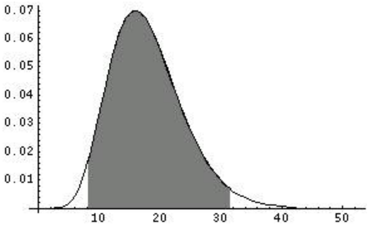


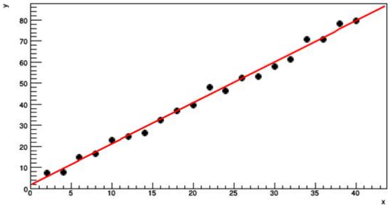


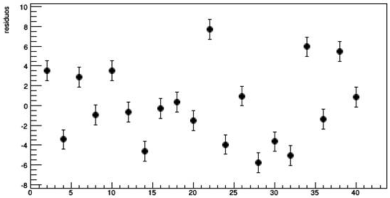


Incertezas subestimadas

- $\chi^2 = 286$
- $ngl = 18$
- 95% CL $= 8,2 < \chi^2 < 32$

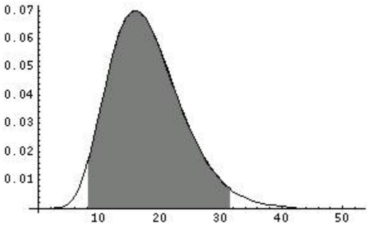


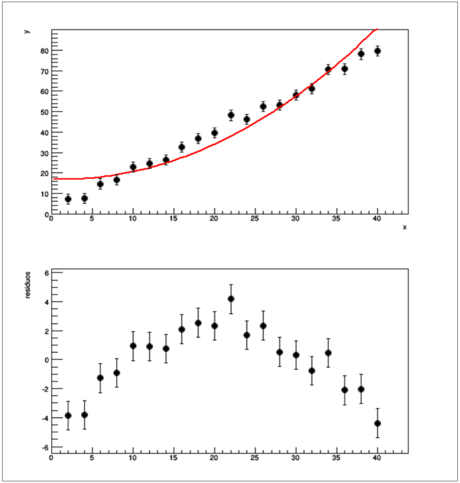




Função incorreta

- $\chi^2 = 105$
- $ngl = 18$
- 95% CL $= 8,2 < \chi^2 < 32$





Sumário

Propagação de incertezas

- Fórmula "geral"

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} cov_{ij}$$

Propagação de incertezas

- Fórmula "geral"

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} cov_{ij}$$

- Exemplo (sem covariância)
  - $a = 5.0 \pm 1.0$ e $b = 1.0 \pm 0.5$

$$y = \frac{a}{b}$$

Propagação de incertezas

- Fórmula "geral"

$$\sigma_y^2 = \sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 \sigma_i^2 + \sum_i \sum_j \frac{\partial f}{\partial x_i}\frac{\partial f}{\partial x_j} cov_{ij}$$

- Exemplo (sem covariância)
  - $a = 5.0 \pm 1.0$ e $b = 1.0 \pm 0.5$

$$y = \frac{a}{b} \quad \Rightarrow \quad y = 5.0 \pm 2.7$$

O que está por trás desta conta

- Supomos que a e b seguem uma FDP gaussiana

O que está por trás desta conta

- Supomos que a e b seguem uma FDP gaussiana
- Supomos que y também será gaussiana

O que está por trás desta conta

- Supomos que a e b seguem uma FDP gaussiana
- Supomos que y também será gaussiana
- De fato, tomar uma expansão de Taylor de primeira ordem apenas faz transformações lineares de a, $b \rightarrow y$
  - Valor médio de y calculado a partir dos valores médios de a e b
  - Variância de y calculada a partir da propagação de incertezas

O que está por trás desta conta

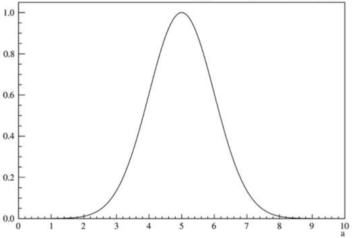
- Supomos que a e b seguem uma FDP gaussiana
- Supomos que y também será gaussiana
- De fato, tomar uma expansão de Taylor de primeira ordem apenas faz transformações lineares de $a, b \to y$
  - Valor médio de y calculado a partir dos valores médios de a e b
  - Variância de y calculada a partir da propagação de incertezas
- Isto é válido sempre? y será sempre gaussiano, se a e b forem?

Incerteza

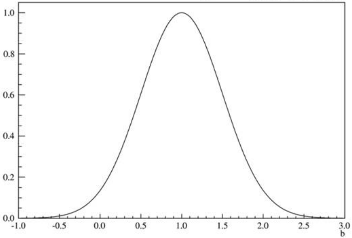
- Somente estatística por enquanto
- Aprendemos que incerteza é uma estimativa do erro da medida
- Uma forma de obter a incerteza é repetir o experimento muitas vezes, tomar os dados e estudar a flutuação destes dados em torno da média
    - Desvio padrão dos dados!

O que está por trás no nosso exemplo? Isto é válido?

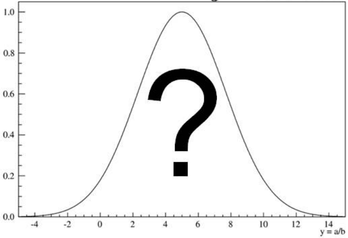
$a = 5.0 \pm 1.0$



$b = 1.0 \pm 0.5$



$y = 5.0 \pm 2.7$



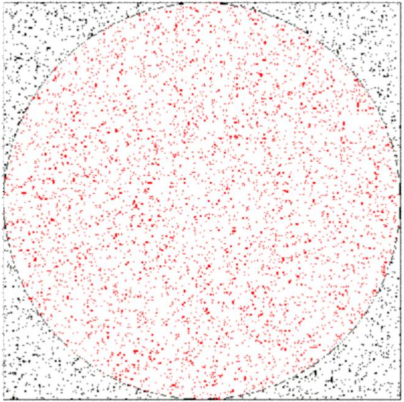
Como achar a FDP de y?

- Método de Monte Carlo
  - Conjunto de métodos computacionais, baseados na repetição, dependente de números aleatórios, para calcular um determinado resultado.

Exemplo: cálculo de pi

- Seja um quadrado de lado 2.
- Sorteia-se um par de números, aleatórios e uniformes, x e y entre -1 e 1
- Se $x^2 + y^2 < R$ com $R = 1$, o par está dentro de um círculo inscrito ao quadrado
- Conta-se quantos pares estão inscritos (n) e quantos são sorteados (N).

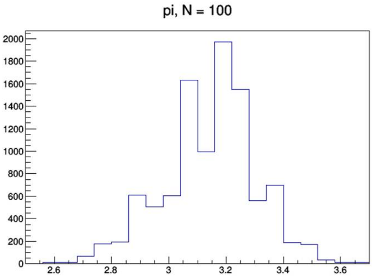
$$\frac{n}{N} = \frac{A_{circ}}{A_{quad}} = \frac{\pi R^2}{L^2} = \frac{\pi}{4}$$

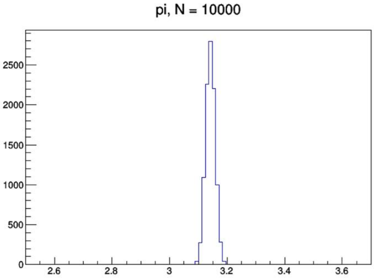


Monte Carlo possui incertezas

- O método "simula" a repetição de um experimento virtual
- Quanto maior a repetição, menor a incerteza no resultado
- Ex: cálculo de pi
    - $N = 100 \rightarrow pi = 3.4$
    - $N = 10000000 \rightarrow pi = 3.14143$
- Repetir o cálculo não deve dar o mesmo resultado
    - $N = 100 \rightarrow pi = 3.4, 3.0, 3.3, etc.$
    - $N = 10000000 \rightarrow pi = 3.14143, 3.14119, 3.14215, etc.$
    - Note que a incerteza diminui.

Incertezas de pi



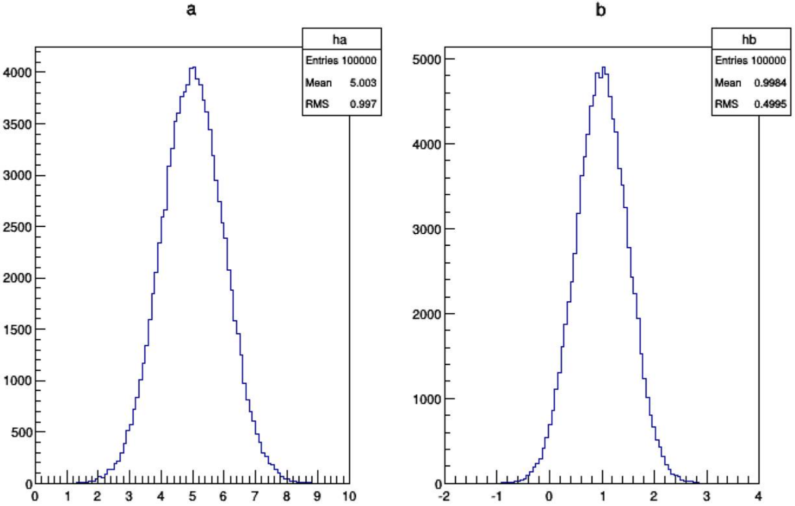


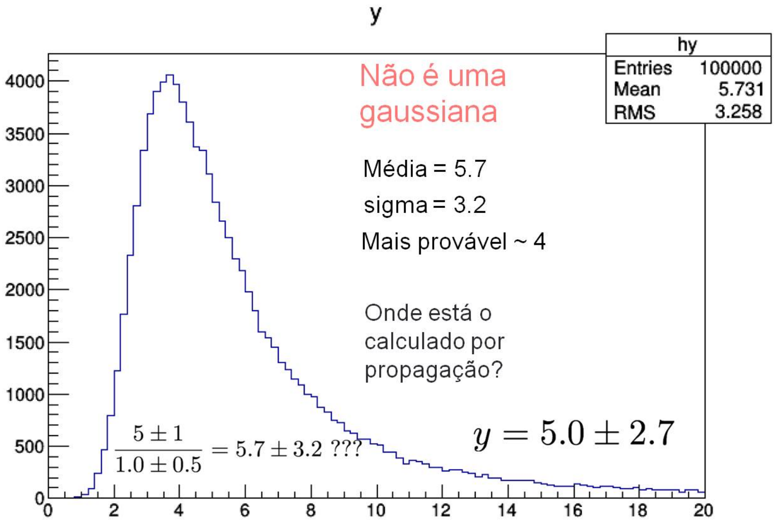
Note a diferença dos resultados!

Como achar a FDP de y?

- Vamos usar o método de Monte Carlo
  - Sorteamos a com uma distribuição gaussiana de média 5 e desvio padrão 1
  - Sorteamos b com uma distribuição gaussiana com média 1 e desvio padrão 0.5
  - Calculamos $y = a/b$ e colocamos em um histograma
  - Repetimos o cálculo N vezes, com N grande

Após muitas iterações




y
hy
Entries 100000
Mean 5.731
RMS 3.258
Não é uma gaussiana
Média = 5.7
sigma = 3.2
Mais provável ~ 4
Onde está o calculado por propagação?
$y = 5.0 \pm 2.7$
$\frac{5 \pm 1}{1.0 \pm 0.5} = 5.7 \pm 3.2$???
0 500 1000 1500 2000 2500 3000 3500 4000
0 2 4 6 8 10 12 14 16 18 20

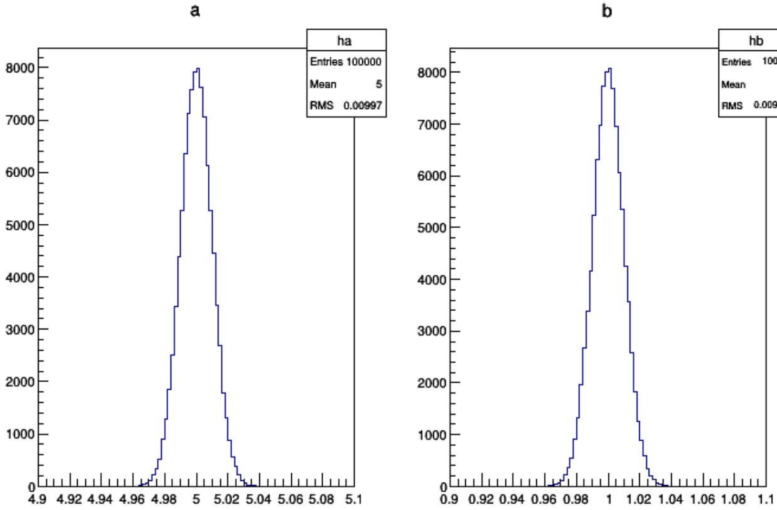
Conclusão

- A fórmula de propagação de incertezas pode dar a falsa impressão de que basta conhecer médias e variâncias e aplicar fórmulas
- A fórmula de propagação vale para incertezas pequenas em torno de uma média (expansão de Taylor de primeira ordem)

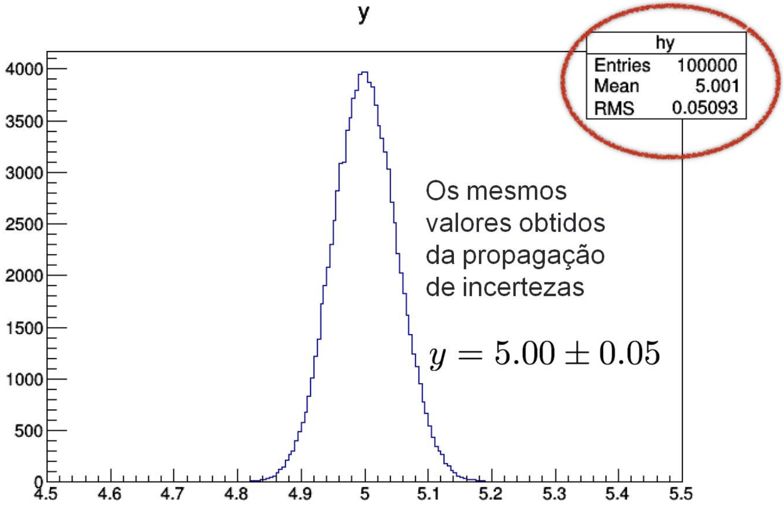
Refazendo o exercício

- Vamos retomar o mesmo exercício com:
  - $a = 5.00 \pm 0.01$
  - $b = 1.00 \pm 0.01$
  - $y = 5.00 \pm 0.05$ (por propagação)
- Como ficam as distribuições?

Após muitas iterações



Após muitas iterações



Os mesmos valores obtidos da propagação de incertezas

$$y = 5.00 \pm 0.05$$

Método de MC para propagação de incertezas

- O que fizemos é chamado de método de Monte Carlo para propagação de incertezas
- Conhecendo-se as FDPs das variáveis, realizamos sorteios das mesmas, calculamos a variável desconhecida e obtemos a FDP desta
- A análise da FDP fornece informações sobre valores mais prováveis, médias, variâncias, etc.

Alguns comentários

- Nem sempre obtemos FDPs gaussianas! CUIDADO!
  - Sendo assim, médias e desvios padrão podem não fornecer toda informação.
- É comum (em algumas áreas) apresentar as FDPs e não apenas números com incertezas.
- MUITOS físicos passam batido por isto. É extremamente comum negligenciar isto e assumir que tudo no universo é gaussiano. Não quer dizer que é correto!

Outros comentários

- No método que desenvolvemos, supomos que a e b são independentes
  - Sorteamos a e b de forma independente
  - Isto deixa implícito que $cov_{ab} = 0$
- Como fazer quando houver covariância entre as grandezas? Como sortear os valores?
  - Vamos abordar covariâncias em detalhe mais adiante