Física Experimental III e IV

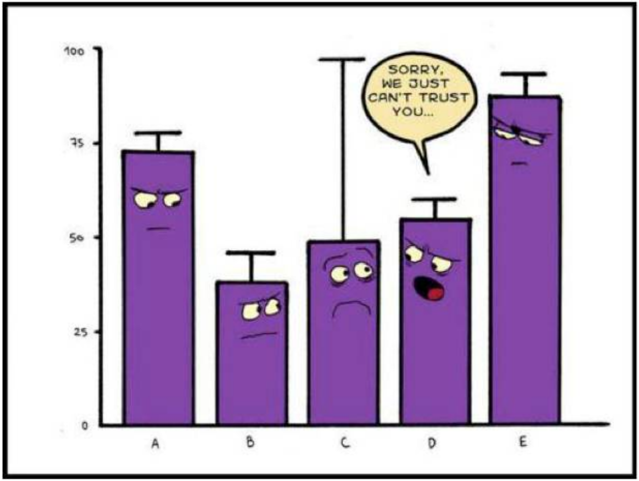
Estatística

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

2019

Sumário

Sumário



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

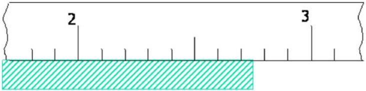
http://www.unicamp.br/~chibeni/textosdidaticos/metodocientifico.pdf

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)

O método científico

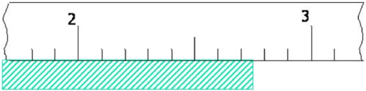
- Uma medida é sempre uma comparação com um padrão
- Sujeita a imperfeições e limitações
- Algarismos significativos
  - Todos que tenho certeza + primeiro duvidoso (estimado)



$$L = 2,74$$

O método científico

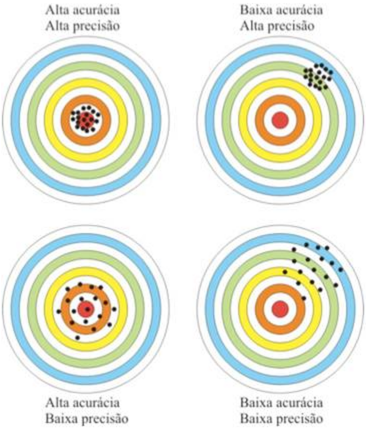
- Uma medida é sempre uma comparação com um padrão
- Sujeita a imperfeições e limitações
- Algarismos significativos
  - Todos que tenho certeza + primeiro duvidoso (estimado)



$$L = 2,74$$

- 2 e 7 tenho "certeza"
- 4 é uma estimativa $\rightarrow$ duvidoso

- **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

- 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

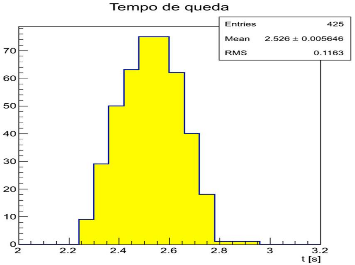
$$\sigma_m = \sqrt{\frac{1}{k}\sum_{i=1}^{k}(\bar{y}_i - \tilde{y})^2} \qquad \sigma_m = \frac{\sigma}{\sqrt{n}}$$

Sumário

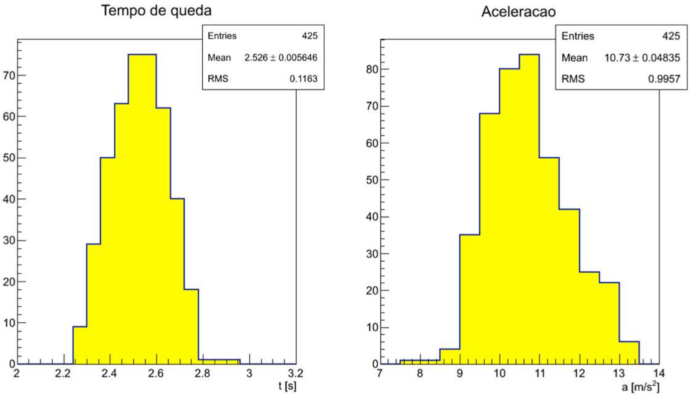
- Medida de tempo de queda de balões de água
- Quase quinhentas medidas
- Análise estatística
- A aceleração obtida é compatível com a gravidade?

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

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

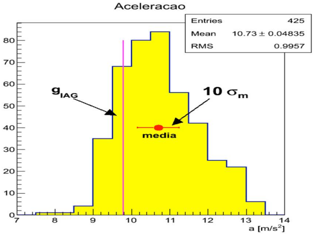
$$\sigma_{tempo} = 0,12 \text{ s}$$

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

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

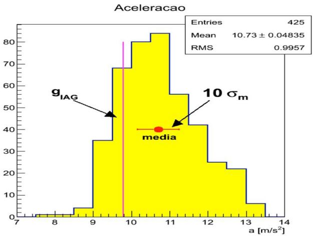
- Os valores de aceleração são precisos e acurados?
- **Precisão**:
  - Desvio padrão: $\sigma_{acel} = 1,0\ \text{m/s}^2$
  - $\sim 10\ \%$ do valor de g
  - $g_{IAG} = 9,7864\ \text{m/s}^2$
  - $g_{medio} = 10,73 \pm 0,05\ \text{m/s}^2$



Acurácia e precisão da medida

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

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



- 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

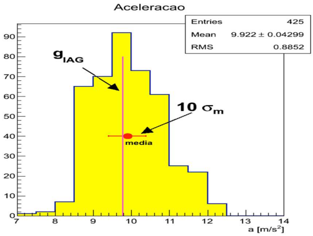
- Como investigar a diferença entre o valor médio e o valor do IAG?
- **Rever o procedimento experimental**
  - O disparo do cronômetro foi auditivo
    - Som tem velocidade de $\sim$340 m/s
    - Torre tem altura de 34 metros
    - Ouvimos o som 0,1 segundo depois que a bola começa a cair

- 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

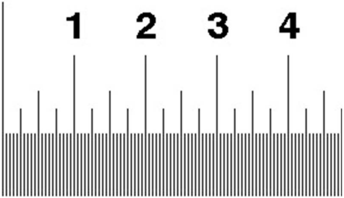
- A repetição exaustiva do experimento permitiu realizar uma análise estatística que evidenciava um problema no procedimento adotado para analisar os dados
- Isso só foi possível porque essa repetição permitiu avaliar as incertezas envolvidas, principalmente a incerteza na aceleração medida
- Em muitas 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

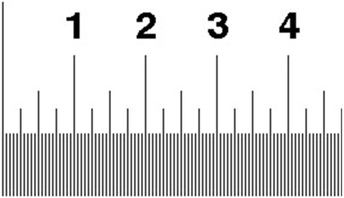
E se não for possível repetir?

- 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



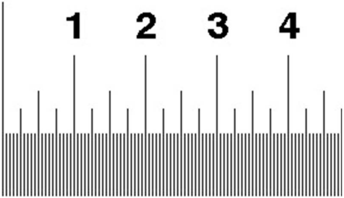
E se não for possível repetir?

- 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



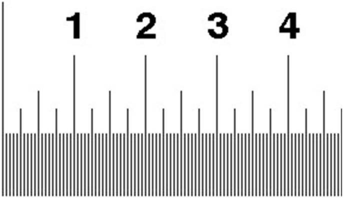
E se não for possível repetir?

- 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



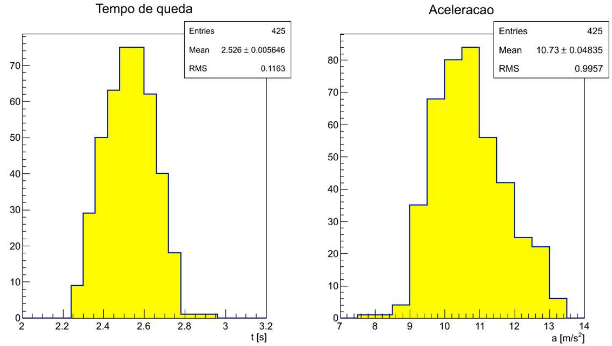
E se não for possível repetir?

- 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 ...

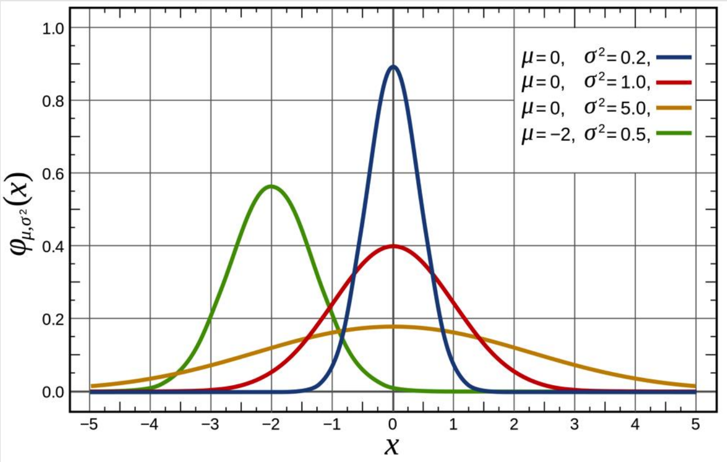
- ... não pudermos repetir este experimento de modo a calcular desvios padrão?

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

$$z = x + y$$

- Considere um exemplo simples onde se quer calcular

$$z = x + y$$

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

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:

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

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

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

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

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

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

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

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

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:

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

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

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

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

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

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

- 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?

- 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

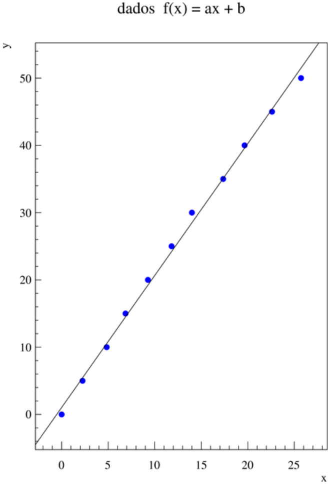
$$\mathrm{cov}_{xy} = \langle (x - \mu_x)(y - \mu_y) \rangle$$

- 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

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

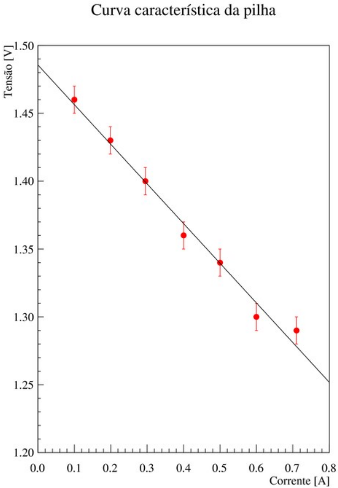
Resultados do ajuste

Número de parâmetros	2
Chi^2	3.42626
Número de graus de liberdade	5

parâmetro	**Valor**	**Incerteza**
0	1.4856	0.00837212
1	-0.292165	0.0186493

Matriz de covariância

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

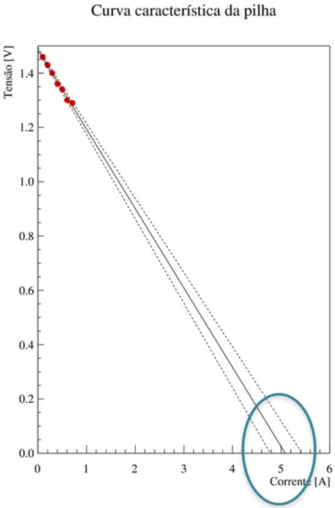


Um exemplo: a curva característica da pilha

- 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?

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

$$\sigma_{i_{max}}^2 = \left(\frac{\partial i_{max}}{\partial a}\right)^2 \sigma_a^2 + \left(\frac{\partial i_{max}}{\partial b}\right)^2 \sigma_b^2 + 2\frac{\partial i_{max}}{\partial a}\frac{\partial i_{max}}{\partial b}\mathrm{cov}_{ab}$$

- Calculando as derivadas, temos:

$$\sigma_{i_{max}}^2 = \frac{1}{b^2}\sigma_a^2 + \frac{a^2}{b^4}\sigma_b^2 - 2\frac{a}{b^3}\mathrm{cov}_{ab}$$

- 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.

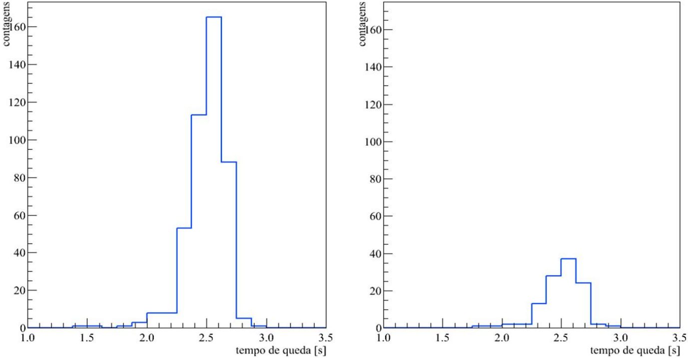
Cuidados importantes

- 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



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

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

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

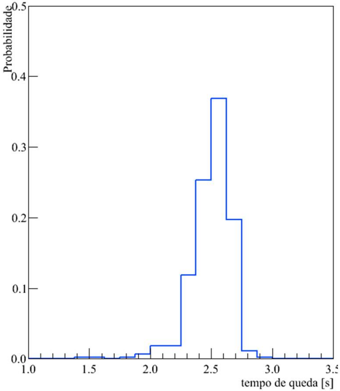
$$P(R) = \lim_{N \to \infty} \frac{N(R)}{N}$$

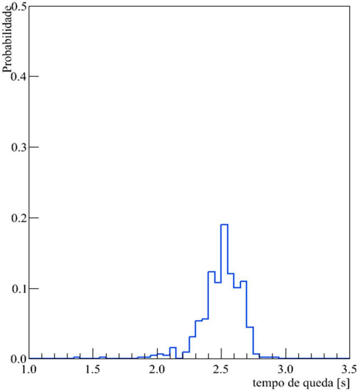
- Em um histograma definem-se canais e contam-se quantas ocorrências tem-se naquele canal

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

  - 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

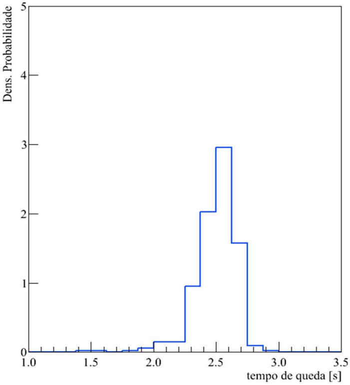
$$P(x, x + \Delta x) = \int_{x}^{x+\Delta x} H(x')dx'$$

- 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

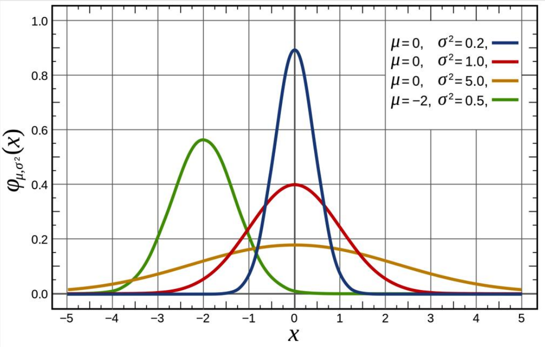
- Por ter significado de uma densidade é sempre positiva
- Como a probabilidade é sempre um número entre 0 e 1, a integral da densidade de probabilidade em todo o espaço deve ser a probabilidade de ter um evento, quaisquer que sejam suas características, ou seja, 100%. Deste modo

$$\int_{-\infty}^{\infty} H(x')dx' = 1$$

Sumário

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

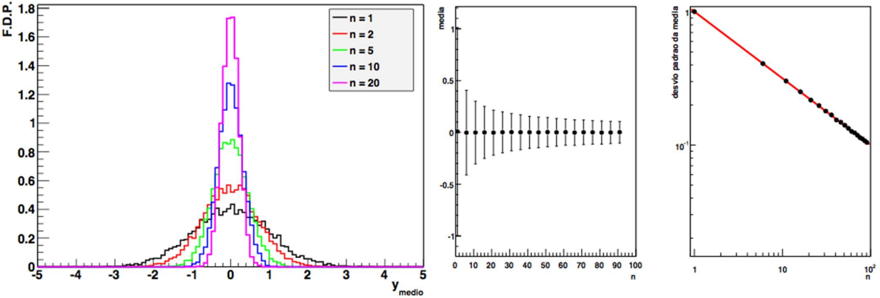
- Considere que um experimentador repita, diariamente, o experimento descrito anteriormente e calcule, para cada dia, o valor da média e o valor de χ^2 e coloque os resultados em um histograma. No final da vida dele ele vai ter repetido isso tantas vezes que é possível conhecer as FDPs dessas duas grandezas.
- 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:

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

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

Distribuição do valor médio

- 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



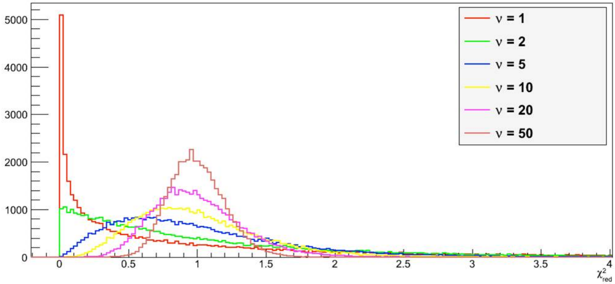
Distribuição da variância

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

- 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"

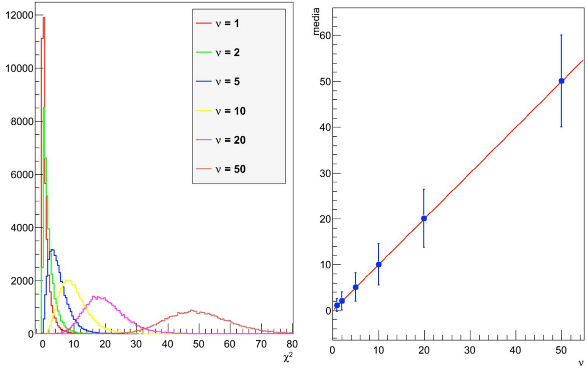
- Para um número pequeno de pontos a distribuição claramente não é gaussiana
- Para poucos pontos tem-se a tendência de SUBESTIMAR a variância dos dados



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.

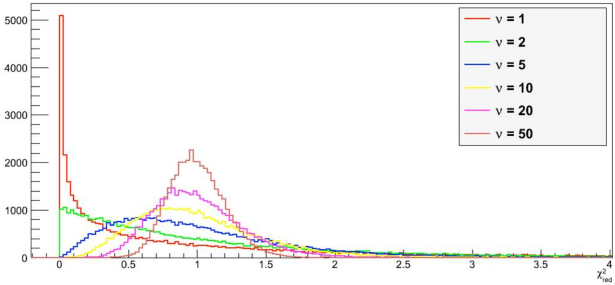


O χ^2 reduzido

- 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.

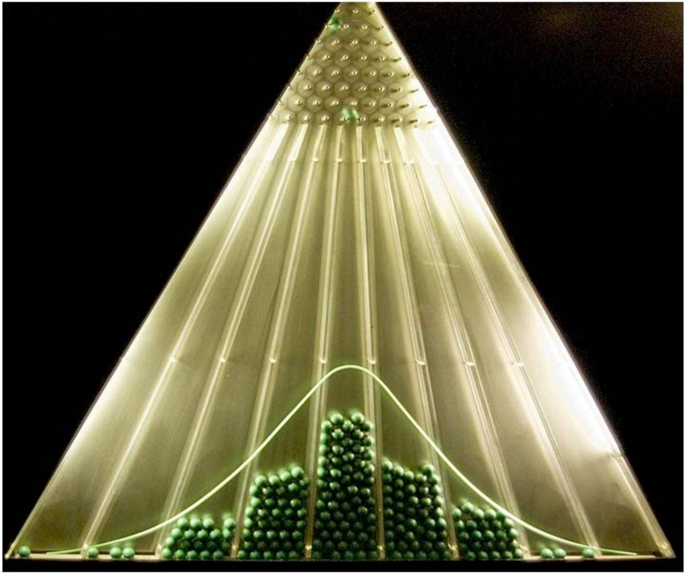
$$\sigma^2 = \frac{1}{N - 1} \sum_{i=1}^{N} (x_i - \bar{x})^2$$

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?

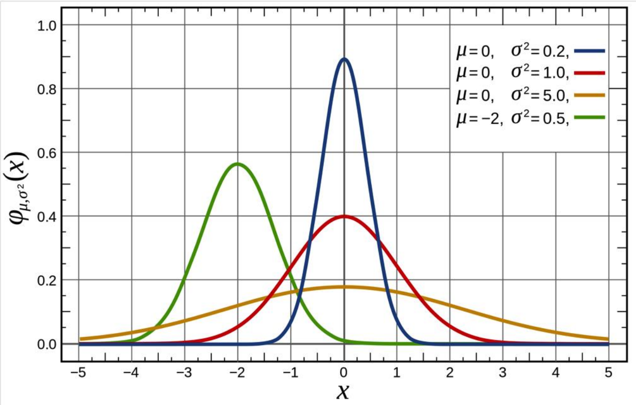
Uma pergunta recorrente

- 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?

Funções de densidade de probabilidade (FDPs)

- 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

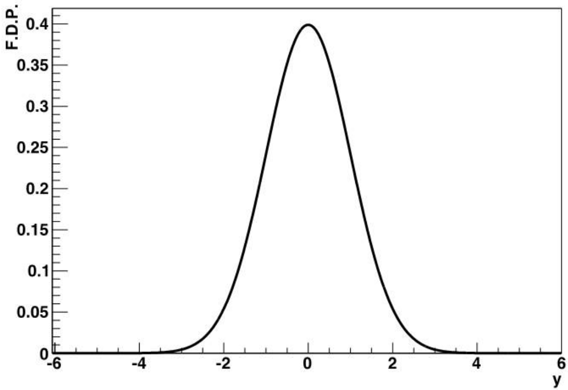
$$Y = \{y_i\} \quad \longrightarrow \quad y_i = \frac{x_i - \mu}{\sigma_\mu}$$

- 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

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



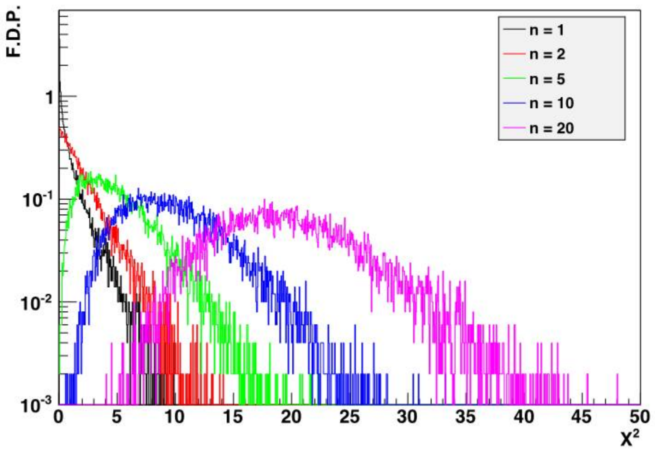
FDP de χ^2

- 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



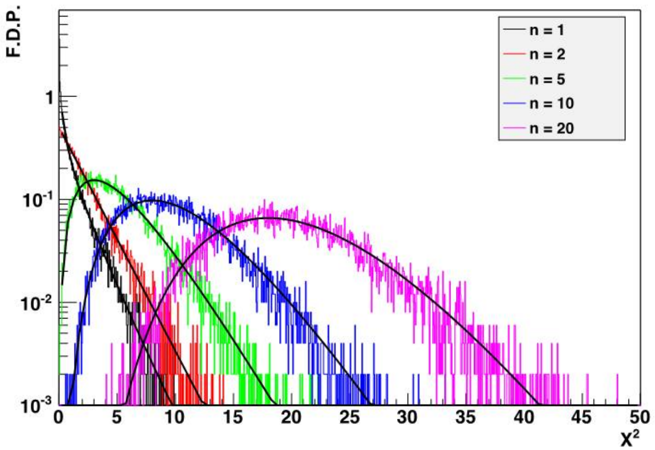
FDP de χ^2

- 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

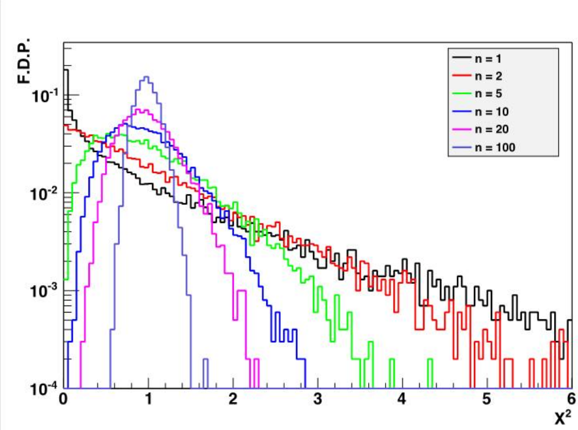
- A função χ^2_{red} é definida como:

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

- 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



- 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?

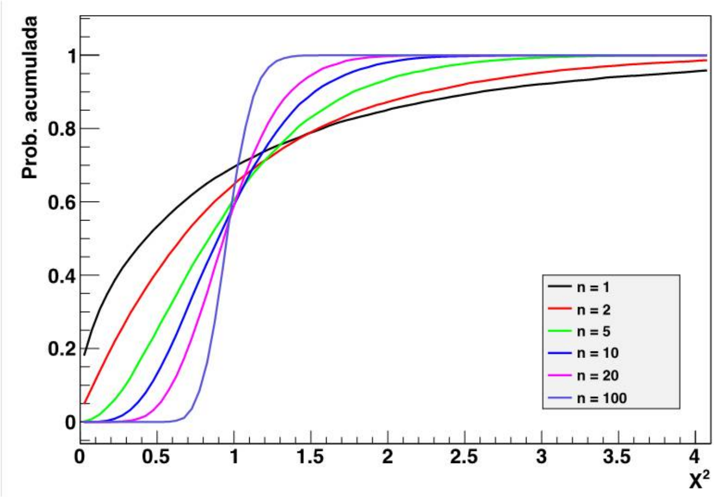
Probabilidade acumulada

- Podemos definir como probabilidade acumulada a grandeza:

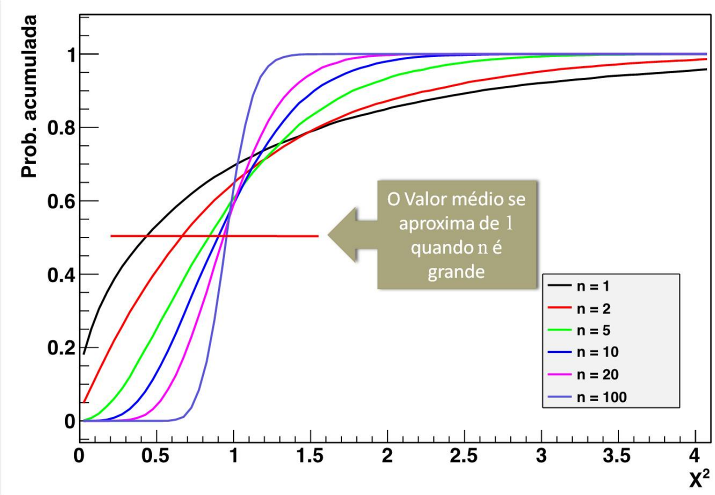
$$P(x) = \int_{-\infty}^{x} p(t)dt \longrightarrow \left\{ \begin{array}{l} \text{Probabilidade da variável } t \\ \text{assumir um valor menor ou} \\ \text{igual a } x \end{array} \right.$$

- 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)

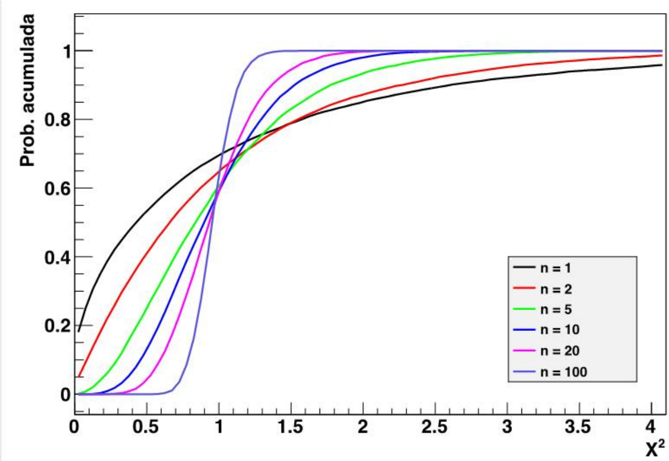


Probabilidade acumulada de χ^2_{red} (ou σ)



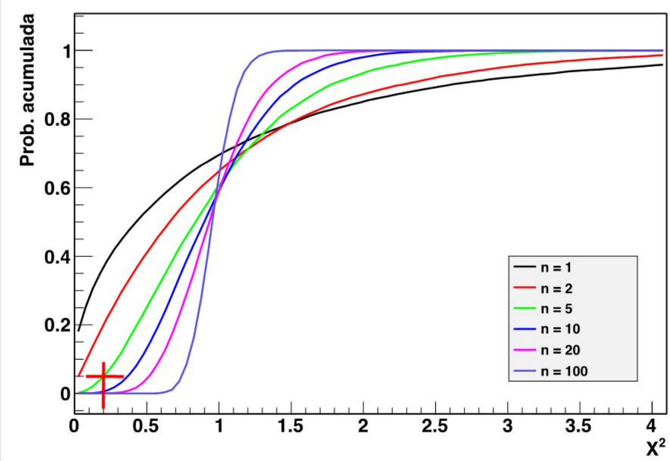
Probabilidade acumulada de χ^2_{red} (ou σ)

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



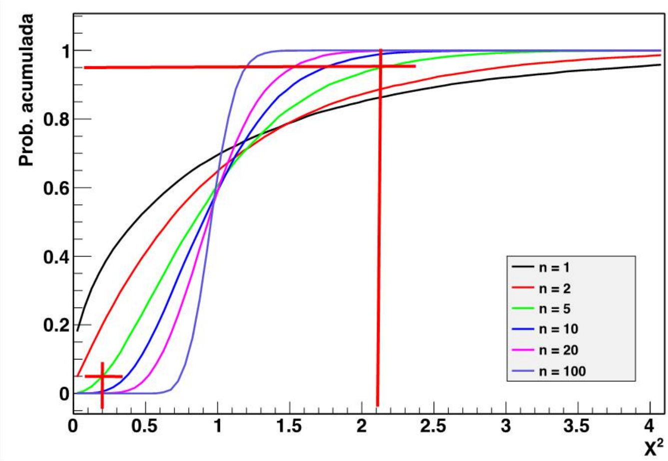
Probabilidade acumulada de χ^2_{red} (ou σ)

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



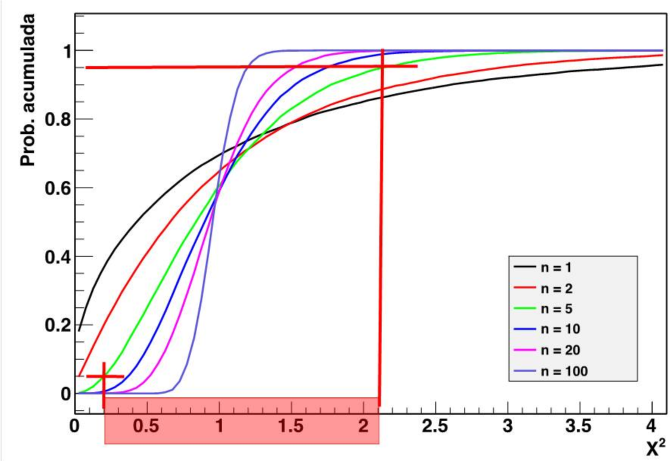
Probabilidade acumulada de χ^2_{red} (ou σ)

- 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)

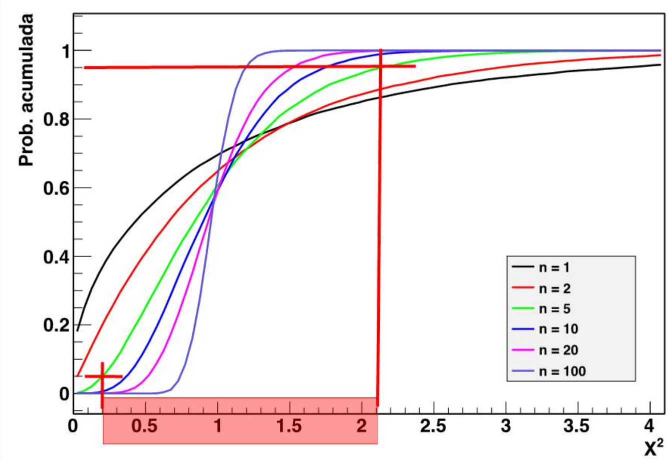
- Se fizermos um ajuste de 5 graus de liberdade, qual o intervalo esperado de χ^2_{red} com 90% de confiança?



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?

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



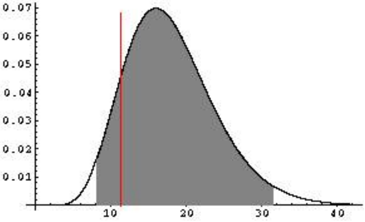
Moral da história

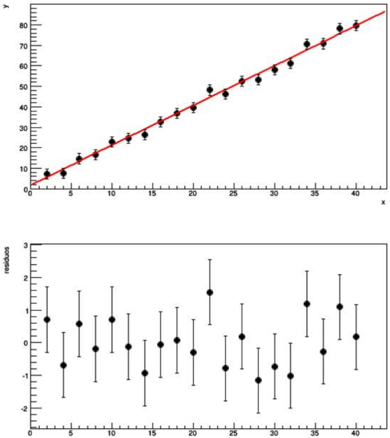
- Testes de χ^2 e análise de resíduos (análise mesmo, não apenas fazer um gráfico) constituem ferramentas poderosas na validação de resultados experimentais
- Cálculo de intervalo de confiança para χ^2, teste-z, teste-t, etc.
  - No WebROOT $\rightarrow$ Calculadoras
- E se o valor de χ^2 estiver fora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

Sumário

Um bom ajuste

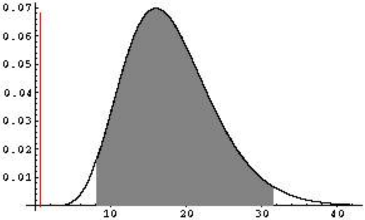
- $\chi^2 = 11,45$
- $ngl = 18$
- 95% CL $= 8,2 < \chi^2 < 32$

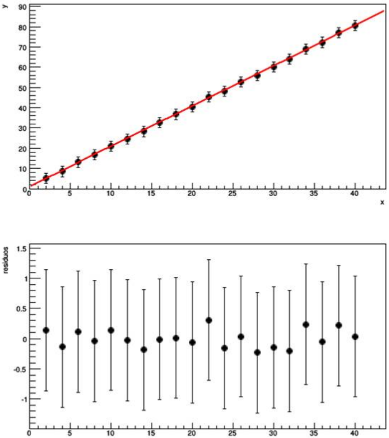




Incertezas superestimadas

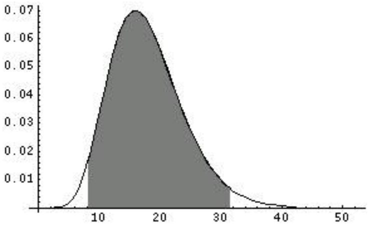
- $\chi^2 = 0,45$
- $ngl = 18$
- $95\%\ \text{CL} = 8,2 < \chi^2 < 32$

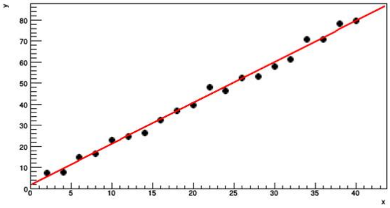


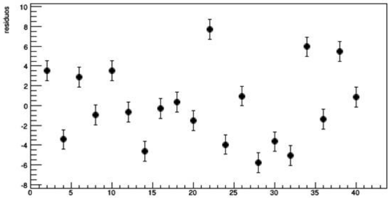


Incertezas subestimadas

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

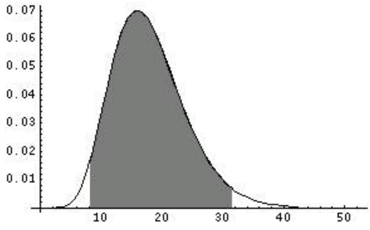


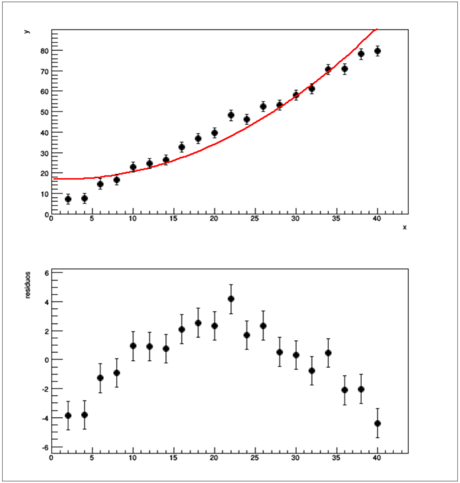




Função incorreta

- $\chi^2 = 105$
- $ngl = 18$
- 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

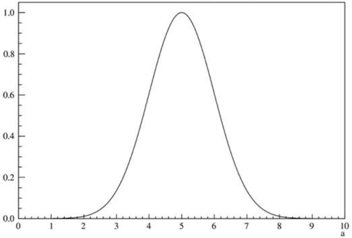
- Supomos que a e b seguem uma FDP gaussiana
- Supomos que y também será gaussiana
- De fato, tomar uma expansão de Taylor de primeira ordem apenas faz transformações lineares de $a, b \rightarrow y$
  - Valor médio de y calculado a partir dos valores médios de a e b
  - Variância de y calculada a partir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

Incerteza

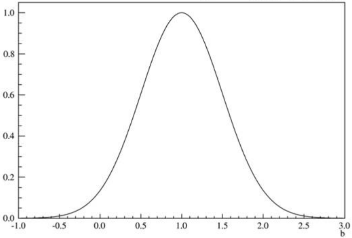
- Somente estatística por enquanto
- Aprendemos que incerteza é uma estimativa do erro da medida
- Uma forma de obter a incerteza é repetir o experimento muitas vezes, tomar os dados e estudar a flutuação destes dados em torno da média
  - Desvio padrão dos dados!

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

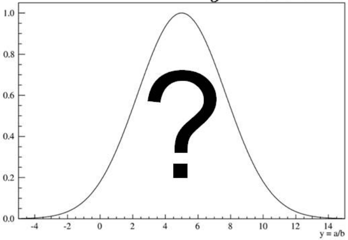
$a = 5.0 \pm 1.0$



$b = 1.0 \pm 0.5$



$y = 5.0 \pm 2.7$



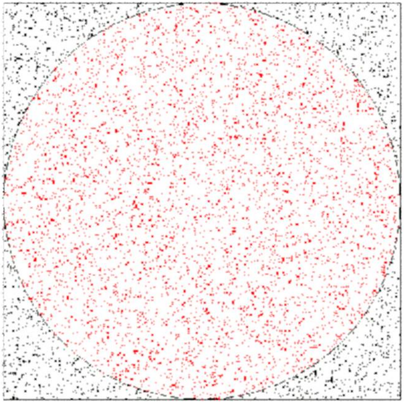
Como achar a FDP de y?

- 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).

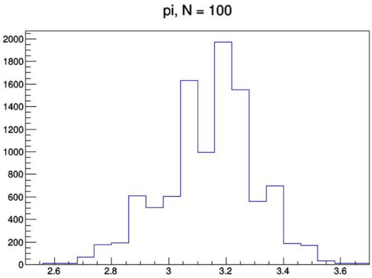
$$\frac{n}{N} = \frac{A_{circ}}{A_{quad}} = \frac{\pi R^2}{L^2} = \frac{\pi}{4}$$

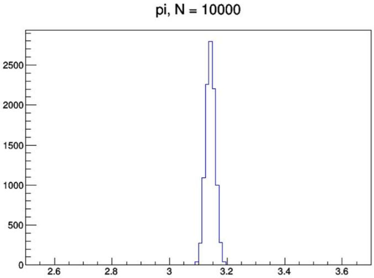


Monte Carlo possui incertezas

- 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



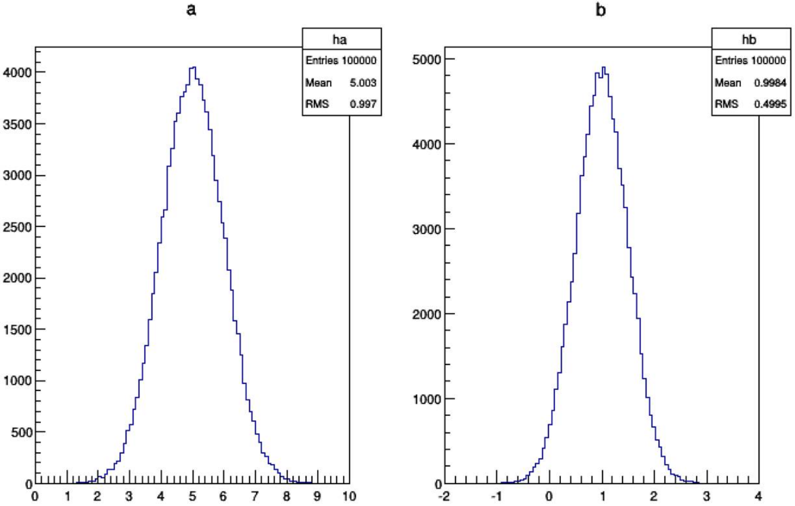


Note a diferença dos resultados!

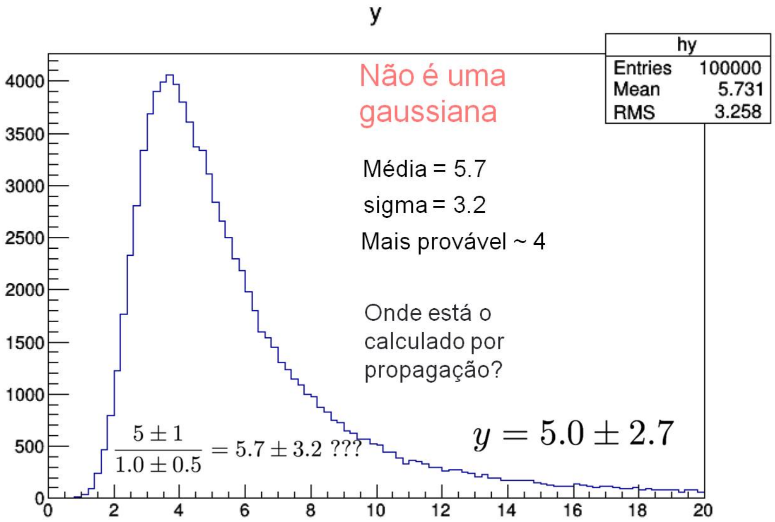
Como achar a FDP de y?

- 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



Após muitas iterações



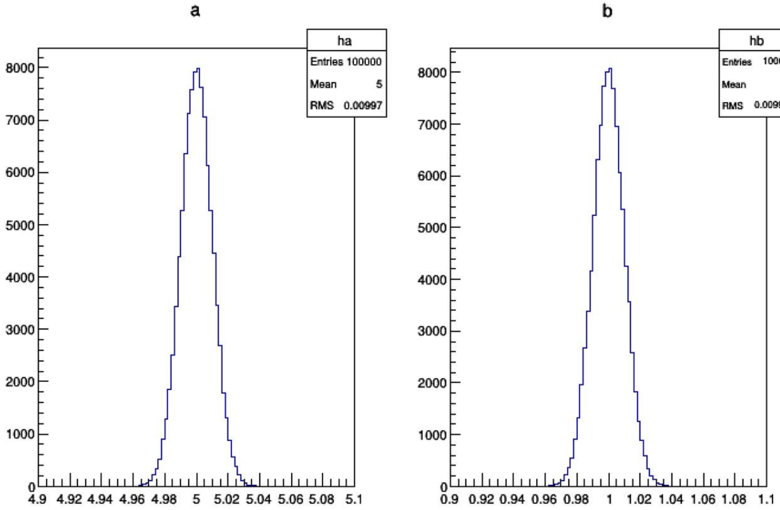
Conclusão

- 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)

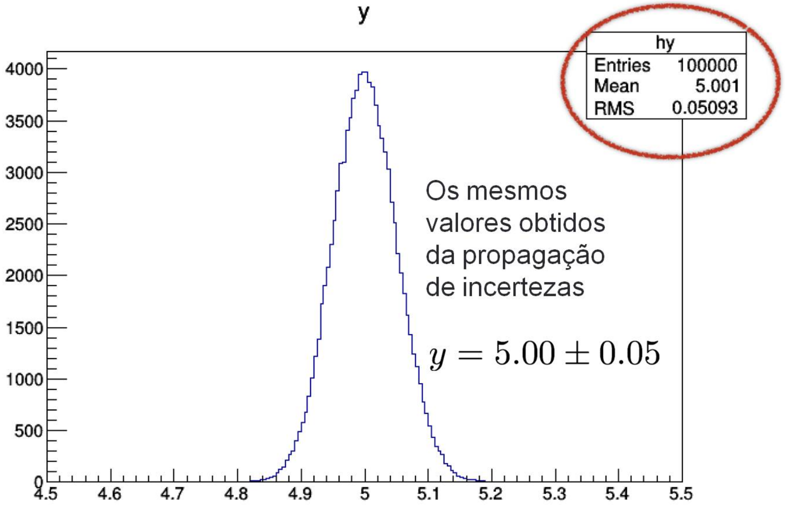
Refazendo o exercício

- 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



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

Sumário

Problema básico

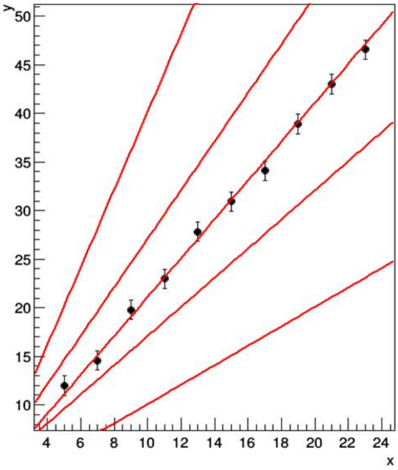
- Imagine um conjunto de dados

$$\{x_i, y_i, \sigma_i\}$$

- E uma função

$$y = f(x, \vec{a})$$

- Onde o vetor $\vec{a}$ corresponde ao conjunto de parâmetros desta função
- Qual o vetor $\vec{a}$ que descreve melhor o conjunto de dados?



Vamos dar um passo para trás

- Imagine um conjunto de medidas realizadas y_i com valores verdadeiros μ_i, cujas funções densidade de probabilidade são:

$$H(y_i, \mu_i)$$

- Os valores verdadeiros podem ser função de outra grandeza, x_i, ou seja:

$$\mu_i = f(x_i, \vec{a})$$

- Onde os elementos do vetor $\vec{a}$ são os parâmetros desta função

Método da máxima verossimilhança

- Podemos definir uma função batizada de verossimilhança como sendo:

$$L = \prod_i H(y_i, \mu_i) = \prod_i H(y_i, f(x_i, \vec{a}))$$

- Esta função não pode ser obtida pois não conhecemos o vetor $\vec{a}$
- O método da máxima verossimilhança consiste em encontrar o vetor $\vec{a}$ de tal forma que L seja máxima

Método da máxima verossimilhança

- Em geral (vai ficar claro porque) trabalha-se com o logaritmo de L

$$L \rightarrow \ln L$$

- Que deve ser maximizado de modo a encontrar o vetor $\vec{a}$.

O Método dos mínimos quadrados

- Vamos assumir que as medidas y_i tenham distribuições gaussianas de tal modo que:

$$H(y_i, \mu_i) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i - f(x_i,\vec{a})}{\sigma_i}\right)^2}$$

- A função de verossimilhança vale (para N medidas)

$$L = \prod_i \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i - f(x_i,\vec{a})}{\sigma_i}\right)^2}$$

O Método dos mínimos quadrados

$$L = \prod_i \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i - f(x_i,\vec{a})}{\sigma_i}\right)^2}$$

- Então;

$$\ln L = \sum_i \left(\frac{1}{\sigma_i\sqrt{2\pi}}\right) - \frac{1}{2}\sum_i \left(\frac{y_i - f(x_i,\vec{a})}{\sigma_i}\right)^2$$

- Lembrando da definição de χ^2, temos:

$$\ln L = \text{const} - \frac{1}{2}\chi^2$$

O Método dos mínimos quadrados

- Deste modo, para L ser máxima

$$\ln L = \text{const} - \frac{1}{2}\chi^2 \quad \text{com} \quad \chi^2 = \sum_i \left(\frac{y_i - f(x_i, \vec{a})}{\sigma_i} \right)^2$$

- O χ^2, deve ser mínimo. Encontrar o vetor $\vec{a}$ que torna o χ^2 mínimo é resolver um sistema de equações que parte de

$$\frac{\partial \chi^2}{\partial a_j} = 0$$

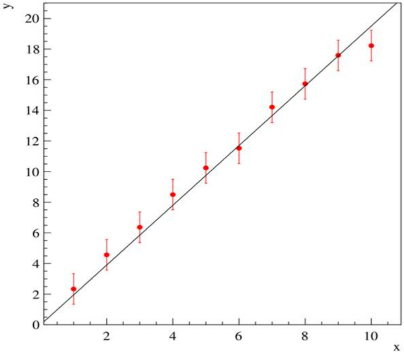
Exemplo: $f(x_i, \vec{a}) = ax$

- Resolvendo o problema:

$$\chi^2 = \sum_i \left(\frac{y_i - ax_i}{\sigma_i} \right)^2$$

$$\frac{d\chi^2}{da} = 0 \quad \Rightarrow \quad a = \frac{\sum_i \frac{x_i y_i}{\sigma_i^2}}{\sum_i \frac{x_i^2}{\sigma_i^2}}$$

- Que tem solução única



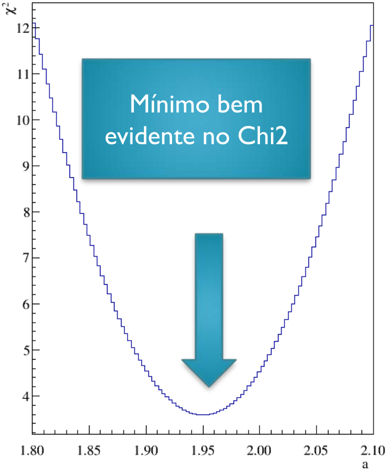
Exemplo: $f(x_i, \vec{a}) = ax$

- Resolvendo o problema:

$$\chi^2 = \sum_i \left(\frac{y_i - ax_i}{\sigma_i} \right)^2$$

$$\frac{d\chi^2}{da} = 0 \quad \Rightarrow \quad a = \frac{\sum_i \frac{x_i y_i}{\sigma_i^2}}{\sum_i \frac{x_i^2}{\sigma_i^2}}$$

- Que tem solução única



Funções lineares nos parâmetros

- No caso de funções lineares nos parâmetros

$$f(x, \vec{a}) = \sum_k a_k g_k(x)$$

- Onde $g_k(x)$ pode ser uma função qualquer, desde que não contenha nenhum parâmetro a ser ajustado
- Neste caso o MMQ tem resolução analítica fechada
  - Ver, por exemplo, Fundamentos da Teoria de Erros, J. H. Vuolo para fórmula geral

Para um bom ajuste, quantos pontos e onde medir?

- Vamos supor um experimento virtual onde os dados se comportam segundo o modelo

$$G(\omega) = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_0}\right)^2}}$$

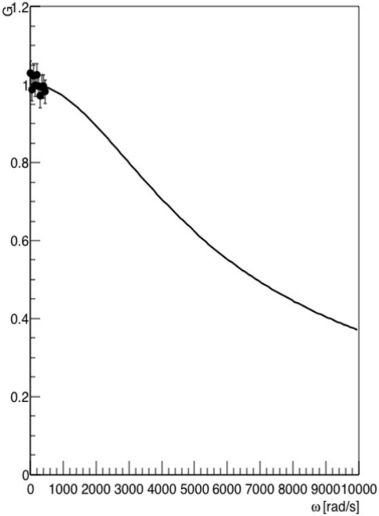
- E o nosso objetivo é determinar, experimentalmente, ω_0
  - Sabemos, por estimativas, que $\omega_0 \sim 4000$ rad/s

Quantos pontos e onde medir?

- Vamos fazer dois experimentos virtuais:
  - Exp 1
    - Medimos 10 pontos em diferentes regiões de ω
  - Exp 2
    - Medimos 5, 10, 15, 20, 25 e 30 pontos na mesma região de ω
- Como as curvas de χ^2 se comportam?
- Quão bem determinado é ω_0 em cada medida?

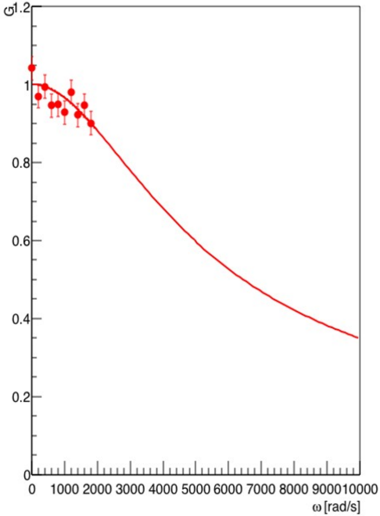
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



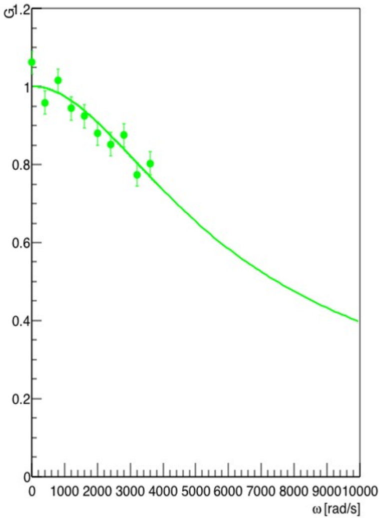
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



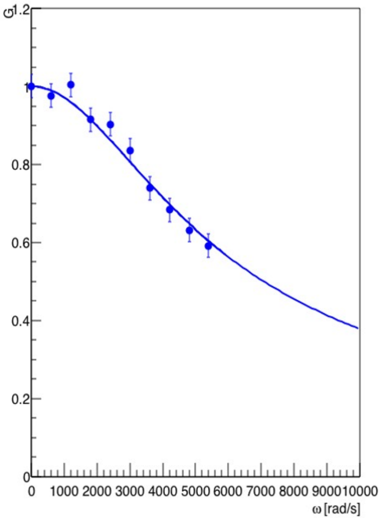
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



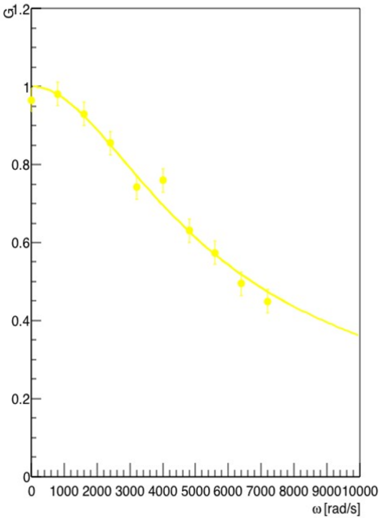
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



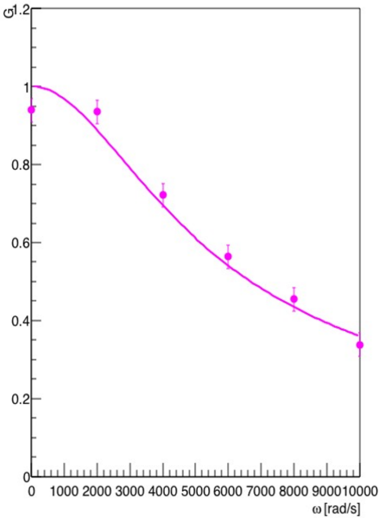
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



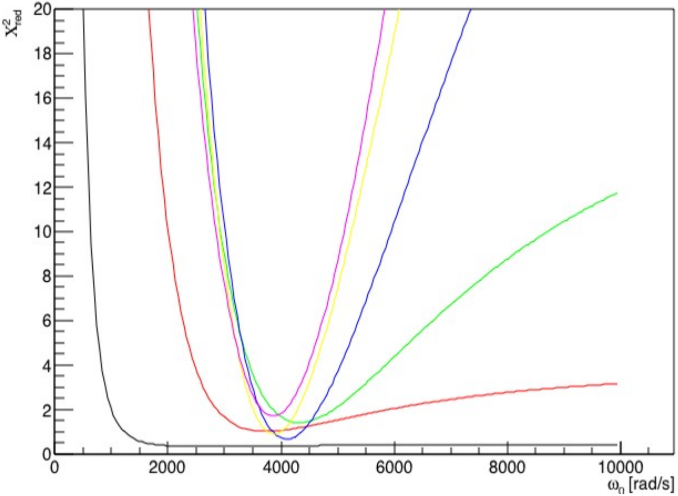
Qual região medir?

- Vamos medir sempre 10 pontos em vários intervalos
  - $\omega = 0 \rightarrow 500$ rad/s
  - $\omega = 0 \rightarrow 2000$ rad/s
  - $\omega = 0 \rightarrow 4000$ rad/s
  - $\omega = 0 \rightarrow 6000$ rad/s
  - $\omega = 0 \rightarrow 8000$ rad/s
  - $\omega = 0 \rightarrow 20000$ rad/s



Como é a curva de χ^2?

- Fizemos χ^2_{red} para caber tudo no mesmo gráfico



Como é a curva de χ^2?

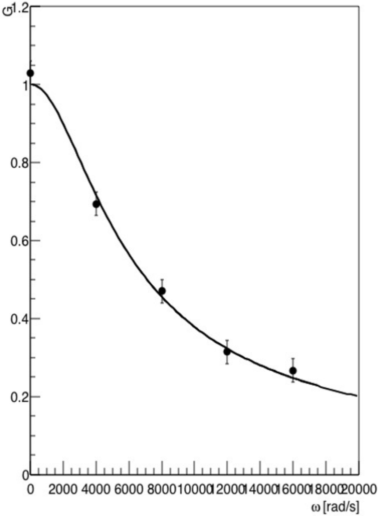
- Fizemos χ^2_{red} para caber tudo no mesmo gráfico

- A escolha do intervalo de medida define quão bem você define a região de mínimo do χ^2
  - No nosso caso, medir somente o começo da curva (dados "pretos") não consegue definir o valor do parâmetro
  - Reflexo na incerteza do parâmetro ajustado

- Parâmetros ajustados
  - $\omega = (2.5 \pm 1.7) \times 10^3$ rad/s
  - $\omega = (3.72 \pm 0.40) \times 10^3$ rad/s
  - $\omega = (4.32 \pm 0.24) \times 10^3$ rad/s
  - $\omega = (4.09 \pm 0.14) \times 10^3$ rad/s
  - $\omega = (3.87 \pm 0.12) \times 10^3$ rad/s
  - $\omega = (3.86 \pm 0.12) \times 10^3$ rad/s

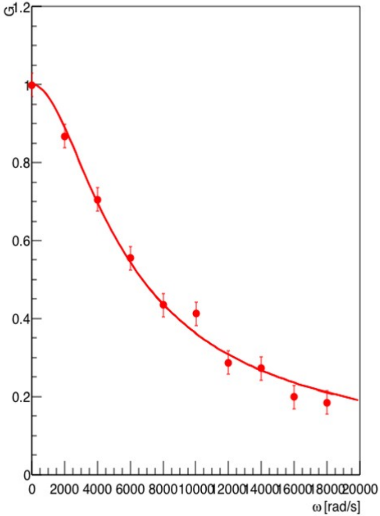
Quantos pontos medir?

- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos

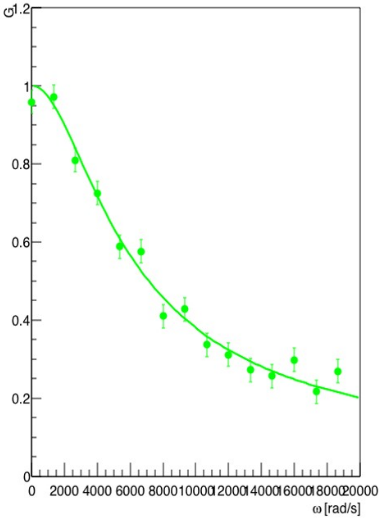


Quantos pontos medir?

- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos

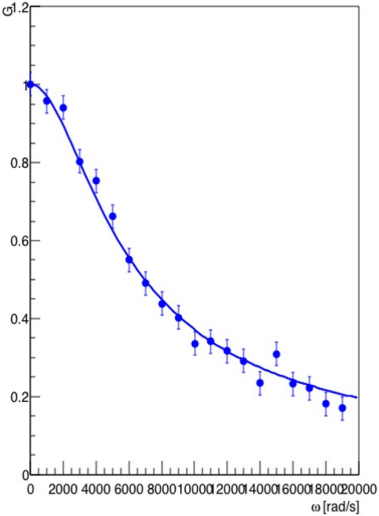


- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos

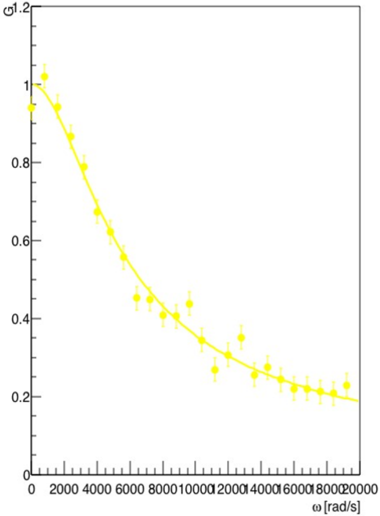


Quantos pontos medir?

- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos

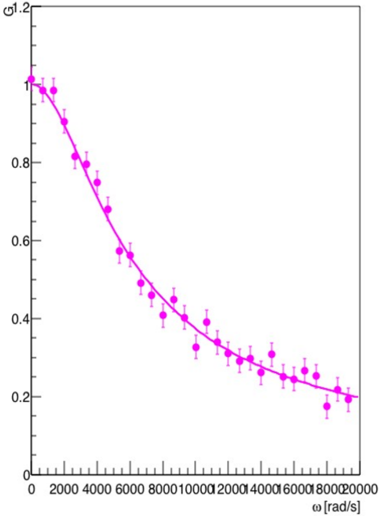


- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos



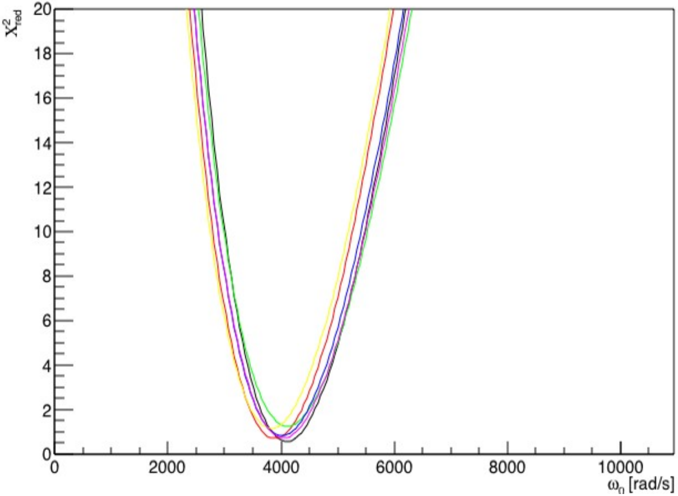
Quantos pontos medir?

- Vamos medir diferente número de pontos no mesmo intervalo
  - 5 pontos
  - 10 pontos
  - 15 pontos
  - 20 pontos
  - 25 pontos
  - 30 pontos



Como é a curva de χ^2?

- Fizemos χ^2_{red} para caber tudo no mesmo gráfico



Como é a curva de χ^2?

- Fizemos χ^2_{red} para caber tudo no mesmo gráfico

- Fixando-se o intervalo que se toma os dados a curva de χ^2_{red} quase não muda de forma
- Porque é melhor tomar mais dados?

$$\sigma_f \sim \frac{\langle\sigma\rangle}{\sqrt{ngl}}$$

- Parâmetros ajustados
  - $\omega = (4.09 \pm 0.19) \times 10^3$ rad/s
  - $\omega = (3.88 \pm 0.13) \times 10^3$ rad/s
  - $\omega = (4.11 \pm 0.11) \times 10^3$ rad/s
  - $\omega = (4.00 \pm 0.09) \times 10^3$ rad/s
  - $\omega = (3.83 \pm 0.08) \times 10^3$ rad/s
  - $\omega = (4.03 \pm 0.07) \times 10^3$ rad/s

Resumindo

- Intervalo de medida
  - Define a forma da curva de χ^2
  - Define a sensibilidade que temos em definir o valor mínimo de χ^2 e, consequentemente o valor do parâmetro ajustado
    - Isto afeta também a incerteza no parâmetro
- Número de pontos
  - Define a incerteza no parâmetro ajustado
- A combinação dos dois define a qualidade da sua análise!

E funções não lineares nos parâmetros?

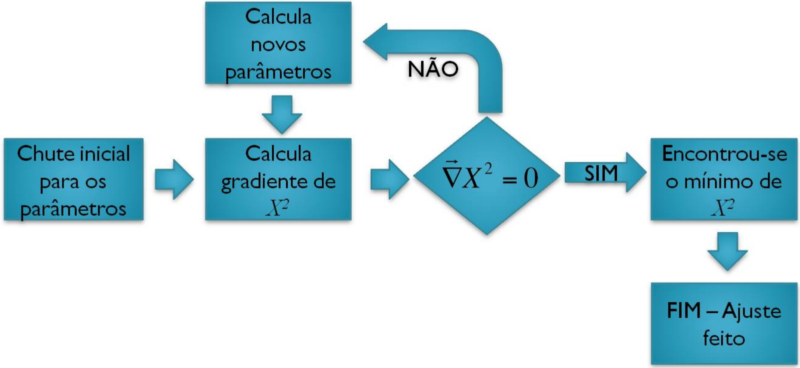
- São aquelas onde os parâmetros aparecem dentro de expressões não lineares, por exemplo

$$f(x, \vec{a}) = a_0 \mathrm{sen}(a_1 x + a_2)$$

- Em geral não dá para resolver analiticamente as expressões de minimização do χ^2.
  - Utilizam-se métodos numéricos para isto

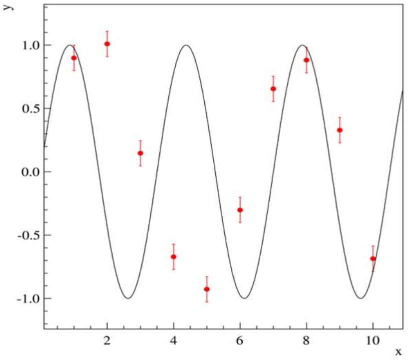
Método típico de ajuste de funções não lineares

- Em geral utiliza-se algorítmos baseados em gradiente



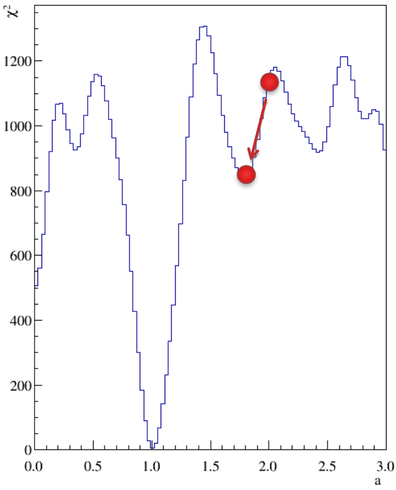
Exemplo: $f(x, \vec{a}) = \mathrm{sen}(ax)$

- Quem nunca se deparou com o problema ao lado?
  - O ajuste não é bom
- Chute inicial $a = 2$
- Mas o método não encontrou um mínimo de χ^2?
  - Encontrou sim, e dai?



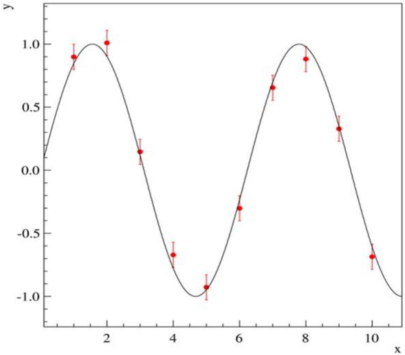
Exemplo: $f(x, \vec{a}) = \mathrm{sen}(ax)$

- Funções não lineares podem gerar função de χ^2 com mínimos locais
  - Dependendo do chute inicial acabamos caindo em um destes mínimos
- Estamos interessados no mínimo global



Exemplo: $f(x, \vec{a}) = \mathrm{sen}(ax)$

- O chute inicial torna-se muito importante no ajuste de funções não lineares
- Chute inicial $a = 0.8$
- Encontramos o mínimo geral



Algumas considerações

- Método de máxima verossimilhança
  - Maximiza-se a probabilidade com base nas densidades de probabilidade dos pontos
  - Se forem gaussianas e independentes $\rightarrow$ máxima verossimilhança resulta em achar o mínimo de χ^2 (MMQ)
- Ajustes de MMQ com funções não lineares podem ser delicados
  - Chute inicial dos parâmetros é importante
  - E as incertezas nos parâmetros ajustados $\rightarrow$ próxima semana

Leitura recomendada

- Tratamento Estatístico de Dados em Física Experimental, capítulo V, O. Helene e V. Vanin. Ed. Edgard Blücher.
- Fundamentos da Teoria de Erros, capítulo 10, J. H. Vuolo. Ed. Edgard Blücher.

Sumário

Método da máxima verossimilhança

- Podemos definir uma função batizada de verossimilhança como sendo:

$$L = \prod_i H(y_i, \mu_i) = \prod_i H(y_i, f(x_i, \vec{a}))$$

- Vamos definir

$$\xi = -\ln L$$

- Note o sinal foi mudado, vai ficar óbvio porque

Método da máxima verossimilhança

- Maximizar a verossimilhança significa minimizar a grandeza

$$\xi = -\ln L = -\sum_i \ln\left(H(y_i, f(x_i, \vec{a}))\right)$$

E isto pode ser feito resolvendo um sistema de equações tal que:

$$\frac{\partial \xi}{\partial a_j} = 0$$

Método dos mínimos quadrados

- No caso das medidas y_i terem distribuições gaussianas temos que (note a inversão de sinal)

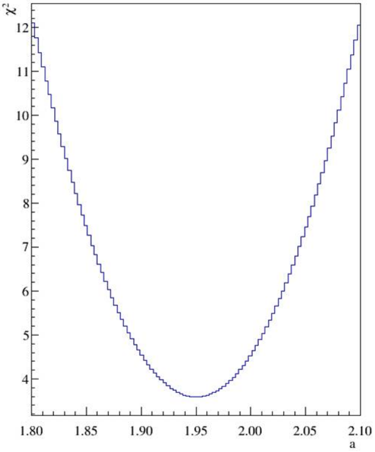
$$\xi = \text{const.} + \frac{1}{2}\chi^2$$

E minimizar esta grandeza é o mesmo que minimizar o χ^2, que é feito através da resolução de um sistema de equações tal que:

$$\frac{\partial \chi^2}{\partial a_j} = 0$$

Neste caso

- Ajustar uma função de pontos gaussianos significa encontrar o mínimo global do χ^2 para os parâmetros
- E as incertezas dos parâmetros? De onde elas vêem?



Incertezas nos parâmetros ajustados

- Vamos iniciar com o método da máxima verossimilhança. Vamos expandir ξ em uma série de Taylor em torno do mínimo ajustado
  - Vamos admitir que as incertezas nos parâmetros são pequenas e que podemos truncar esta expansão nos primeiros termos

$$\xi = \xi_0 + \underbrace{\sum_i \frac{\partial \xi}{\partial a_j}(a_j - \bar{a}_j)}_{0 \text{ (ponto de mínimo)}} + \frac{1}{2}\sum_j \frac{\partial^2 \xi}{\partial a_j^2}(a_j - \bar{a}_j)^2 + \frac{1}{2}\sum_j \sum_{k \neq j} \frac{\partial^2 \xi}{\partial a_j \partial a_k}(a_j - \bar{a}_j)(a_k - \bar{a}_k) + \ldots$$

  - O termo ξ_0 é o valor de ξ calculado no mínimo

- Podemos então calcular a função verossimilhança

$$L = e^{-\xi}$$

- Que resulta em

$$L = e^{-\xi_0} \times e^{-\frac{1}{2}\sum_j \frac{\partial^2 \xi}{\partial a_j^2}(a_j - \bar{a}_j)^2} \times e^{-\frac{1}{2}\sum_j \sum_{k \neq j} \frac{\partial^2 \xi}{\partial a_j \partial a_k}(a_j - \bar{a}_j)(a_k - \bar{a}_k)} \times \ldots$$

- Se ξ for suficientemente parabólica em torno do mínimo os termos de ordem superiores são praticamente nulos
  - Exponencial destes termos ~ 1

$$L \sim e^{-\xi_0} \times \underbrace{e^{-\frac{1}{2}\sum_j \frac{\partial^2 \xi}{\partial a_j^2}(a_j - \bar{a}_j)^2}}_{\text{produto de gaussianas}} \times \underbrace{e^{-\frac{1}{2}\sum_j \sum_{k \neq j} \frac{\partial^2 \xi}{\partial a_j \partial a_k}(a_j - \bar{a}_j)(a_k - \bar{a}_k)}}_{\text{termos de covariância entre parâmetros}}$$

- A função verossimilhança se assemelha a um produto de distribuições de probabilidade gaussianas com covariância

Incertezas nos parâmetros ajustados

- Por comparação

$$L \sim e^{-\xi_0} \times \underbrace{e^{-\frac{1}{2}\sum_j \frac{\partial^2 \xi}{\partial a_j^2}(a_j - \bar{a}_j)^2}}_{\text{produto de gaussianas}} \times \underbrace{e^{-\frac{1}{2}\sum_j \sum_{k \neq j} \frac{\partial^2 \xi}{\partial a_j \partial a_k}(a_j - \bar{a}_j)(a_k - \bar{a}_k)}}_{\text{termos de covariância entre parâmetros}}$$

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

- O que faz com que:

$$\sigma_{a_j}^2 = \left(\frac{\partial^2 \xi}{\partial a_j^2}\right)^{-1}$$

O método dos mínimos quadrados

- No método dos mínimos quadrados, sabemos que

$$\xi = \text{const.} + \frac{1}{2}\chi^2$$

- De modo que:

$$\sigma_{a_j}^2 = \left(\frac{\partial^2 \xi}{\partial a_j^2}\right)^{-1} = \left(\frac{1}{2}\frac{\partial^2 \chi^2}{\partial a_j^2}\right)^{-1}$$

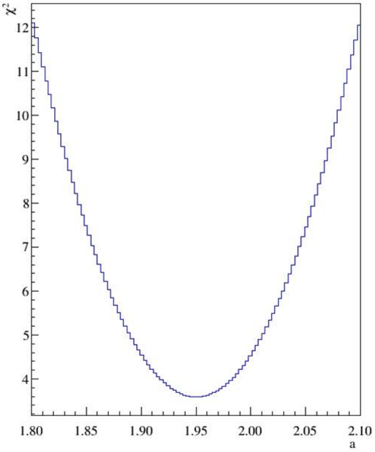
- Expandir em Taylor

$$\chi^2 = \chi^2_{min} + \frac{1}{2}\frac{\partial^2\chi^2}{\partial a^2}(a - \bar{a})^2 + \dots$$

$$= \chi^2_{min} + \frac{1}{\sigma_a^2}(a - \bar{a})^2 + \dots$$

- Ou seja

$$\Delta\chi^2 = \chi^2 - \chi^2_{min} \sim \frac{1}{\sigma_a^2}(a - \bar{a})^2$$



Vamos olhar a curva de χ^2

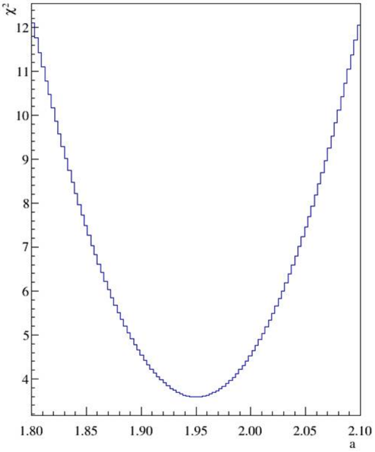
- De modo que

$$\Delta\chi^2 = \chi^2 - \chi^2_{min} \sim \frac{1}{\sigma_a^2}(a - \bar{a})^2$$

- Se

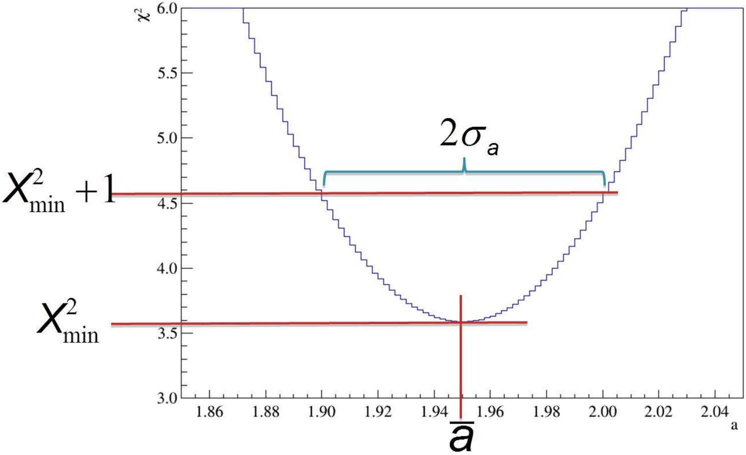
$$(a - \bar{a})^2 = \sigma_a^2 \quad \rightarrow \quad \Delta\chi^2 \sim 1$$

- Posso graficamente estimar a incerteza em a



Incerteza de parâmetros do MMQ graficamente

$$(a - \bar{a})^2 = \sigma_a^2 \quad \rightarrow \quad \Delta\chi^2 \sim 1$$



Algumas considerações

- Supomos que a grandeza ξ seja suficientemente parabólica em torno do seu mínimo
  - Consequentemente, também o χ^2 quando o MMQ se aplica
  - Isto nem sempre é verdade mas em geral dá uma boa “estimativa” das incertezas nos parâmetros ajustados
    - ⋆ Para fazer direito utiliza-se método de Monte Carlo
- Note que estamos olhando variação no χ^2 e não no χ^2_{red}

Sumário

Teste z

- Imagine que foi realizada uma medida e queremos compará-la ao seu "valor verdadeiro" - muitas vezes uma previsão teórica - calculamos z

$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$

- y será compatível com μ se $|z| < 3$
- Isso é razoável? Qual a origem do valor 3?

Probabilidade acumulada

- Podemos definir como probabilidade acumulada a grandeza:

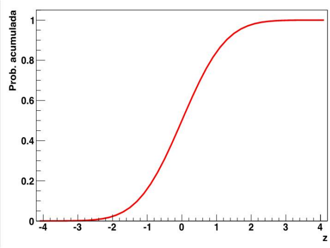
$$P(x) = \int_{-\infty}^{x} p(t)dt \longrightarrow \left\{ \begin{array}{l} \text{Probabilidade da variável } t \\ \text{assumir um valor menor ou} \\ \text{igual a } x \end{array} \right.$$

- Nesse caso, $p(t)$ é a função densidade de probabilidade da grandeza estudada

Teste z

$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$

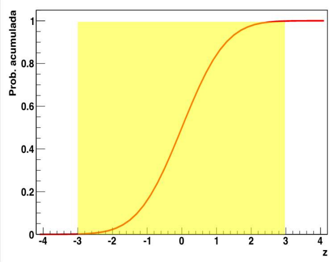
- No teste-z, assume-se que a variável z possui FDP normal
- Se a hipótese do teste for verdadeira z deve possuir valor verdadeiro zero e variância 1



Teste z

$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$

- No teste-z, assume-se que a variável z possui FDP normal
- Se a hipótese do teste for verdadeira z deve possuir valor verdadeiro zero e variância 1



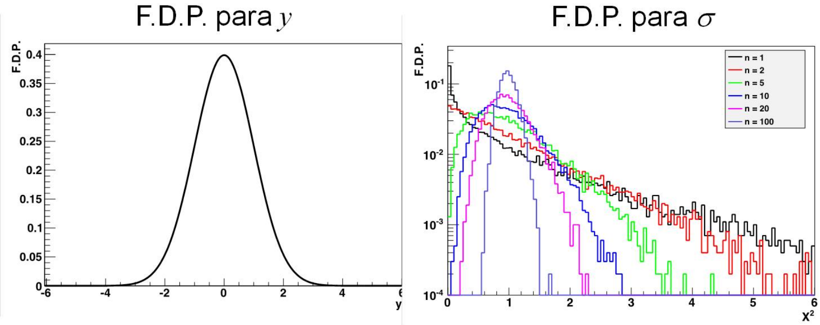
$$|z| < 3 \quad \Rightarrow \quad \sim 99.9\%$$

- No teste-z, 3 significa que $\sim 99.9\%$ das medidas devem estar nesse intervalo, ou seja, a chance de obter um valor com $|z| > 3$ é desprezível
- **CUIDADO!** Não use o valor 3 a esmo. Entenda o seu significado

Teste z

- Mas será que a variável z assume uma FDP normal?

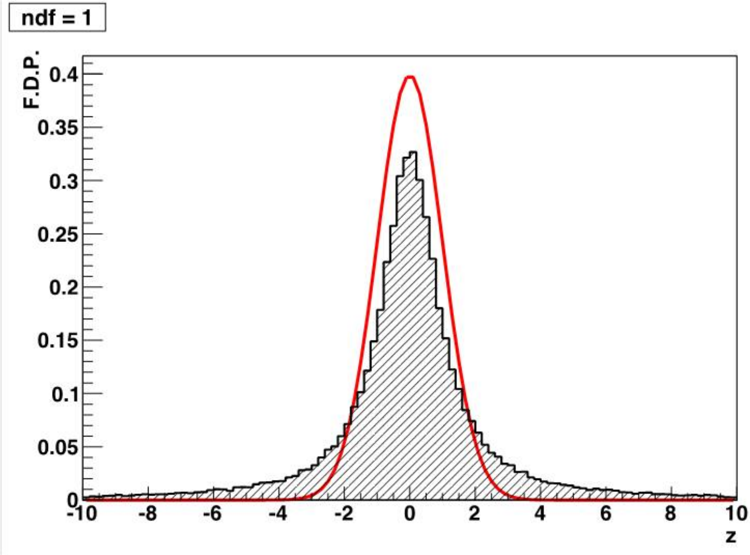
$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$



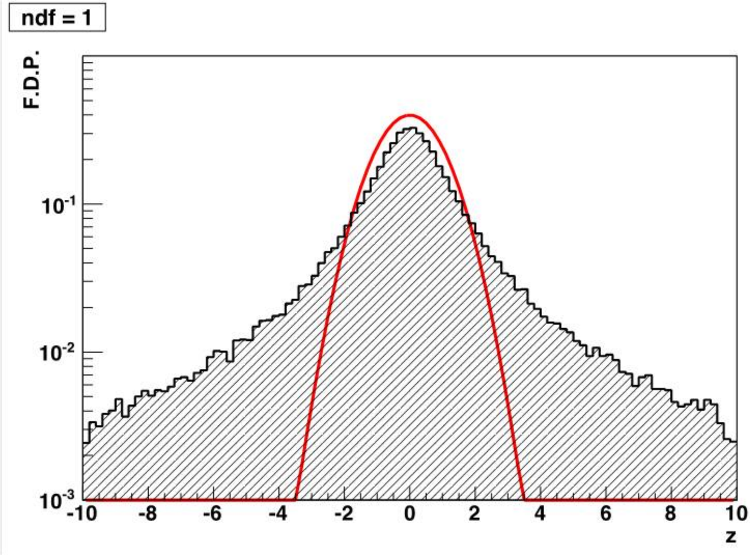
Teste z

- Vamos realizar o mesmo procedimento (experimentos virtuais) de aulas passadas e obter a FDP de z para diferentes números de graus de liberdade (ngl)
$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$
- Ou seja, vamos simular, para um dado ngl qual é o valor médio de y, a sua incerteza e calcular z.
- Vamos comparar a distribuição de z com uma distribuição normal de média 0 e variância 1.
  - O teste-z supõe que a variável z tem distribuição normal de média 0 e desvio padrão 1.

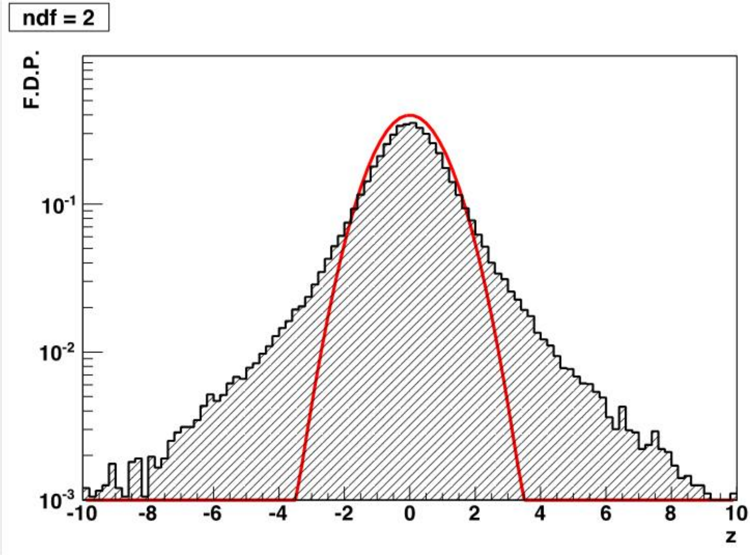
FDP de z para ngl $= 1$



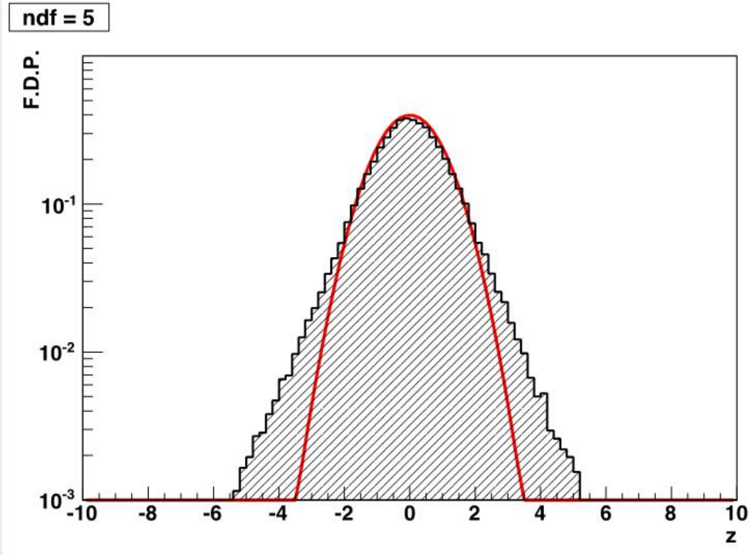
FDP de z para ngl = 1

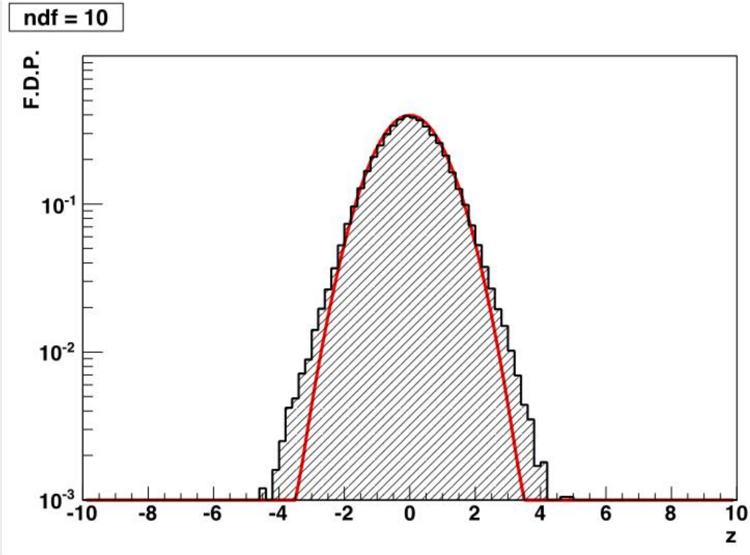


FDP de z para ngl $= 2$



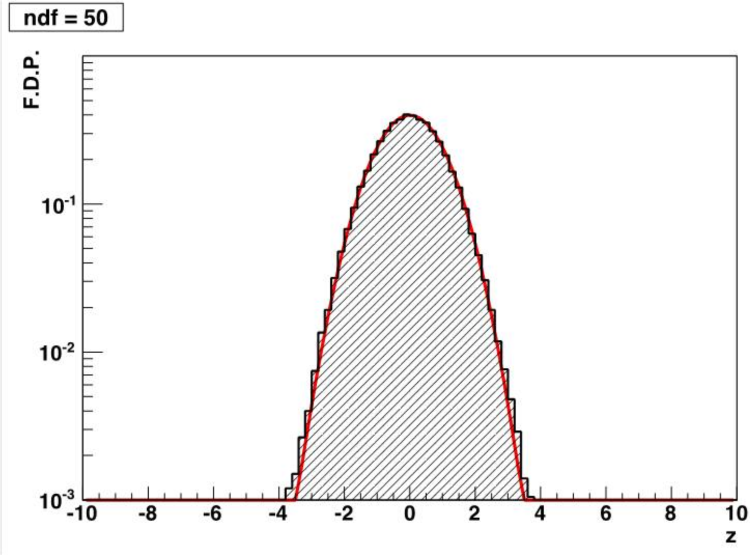
FDP de z para ngl $= 5$





ndf = 10
F.D.P.
10-1
10-2
10-3
-10
-8
-6
-4
-2
0
2
4
6
8
10
z

FDP de z para ngl $= 50$!!



Conclusões

- A variável z somente possui FDP normal para medidas com elevado número de graus de liberdade
  - Tipicamente ngl > 20 já dá para aproximar com um pouco de ressalva
- Ou seja, o teste-z só pode ser utilizado em situações com grande ngl
  - A situação é crítica para ngl $\sim$ 1-2
- Que tipo de teste utilizar para pequenos valores de ngl?

Teste-t de Student

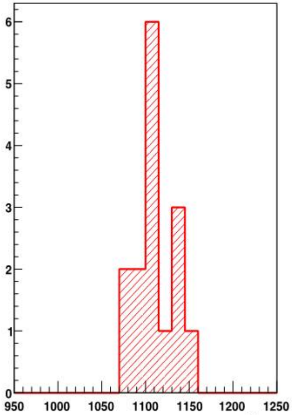
- Willian Gosset (Student) - 1908
  - Distribuição de probabilidades de Student (ou distribuição t)
- A distribuição t surge quando se quer comparar valores médios de distribuições com poucos graus de liberdade. O fato da FDP da variância não seguir uma distribuição normal, nesses casos, faz com que seja necessário utilizar distribuições t de probabilidade

Teste-t de Student

- Vamos supor que queremos comparar duas amostras diferentes e testar se elas são compatíveis. Cada amostra foi medida com um certo número de graus de liberdade, possui uma média e um desvio padrão estimado da amostra. Define-se a grandeza t, de modo geral, como sendo:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \longrightarrow \left\{ \begin{array}{l} \bar{y}_i = \text{média da amostra } i \\ \sigma_i = \text{variância da amostra } i \\ n_i = \text{ngl da amostra } i \end{array} \right.$$

- Duas turmas mediram uma constante C e obtiveram as seguintes distribuições. Podemos dizer que os valores médios das salas são compatíveis?



Um caso particular

- Queremos comparar o valor médio de uma amostra com uma expectativa para o seu valor verdadeiro (teórico, por exemplo). Nesse caso, t pode ser escrito como:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \longrightarrow t = \frac{\bar{y}_1 - \mu}{\frac{\sigma_1}{\sqrt{n_1}}}$$

- Que é a mesma expressão para z

- O teste-t de Student consiste em verificar se a hipótese de igualdade entre valores médios de duas amostras (ou entre o valor médio de uma amostra e uma expectativa "verdadeira") é válida, ou seja, verificar se:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \text{ ou } t = \frac{\bar{y}_1 - \mu}{\frac{\sigma_1}{\sqrt{n_1}}}$$

- É compatível com zero
- Contudo, para testar essa compatibilidade, devemos levar em conta que a distribuição de t não é mais normal
- Qual a FDP de t?

Distribuição t de Student

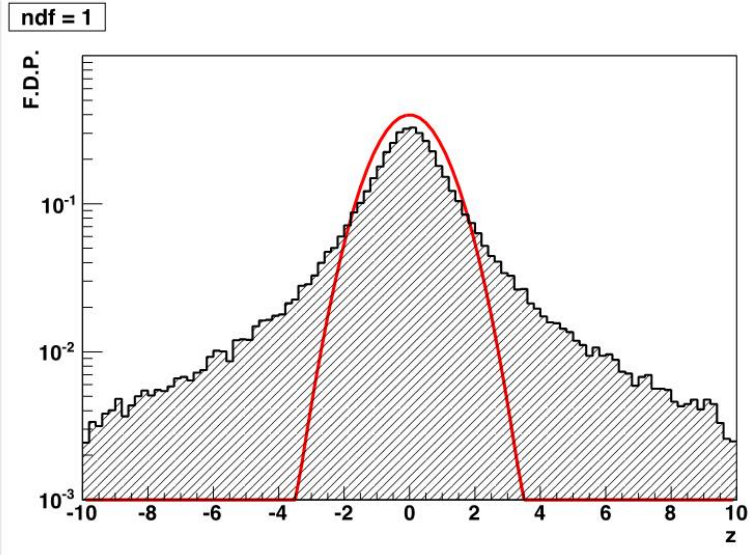
- Grandezas que são obtidas como razões entre uma distribuição normal (valor médio) e uma distribuição de χ^2 (variância, por exemplo), possuem FDP de Student.
  - Grandezas t e z se enquadram nessa relação

$$p(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$$

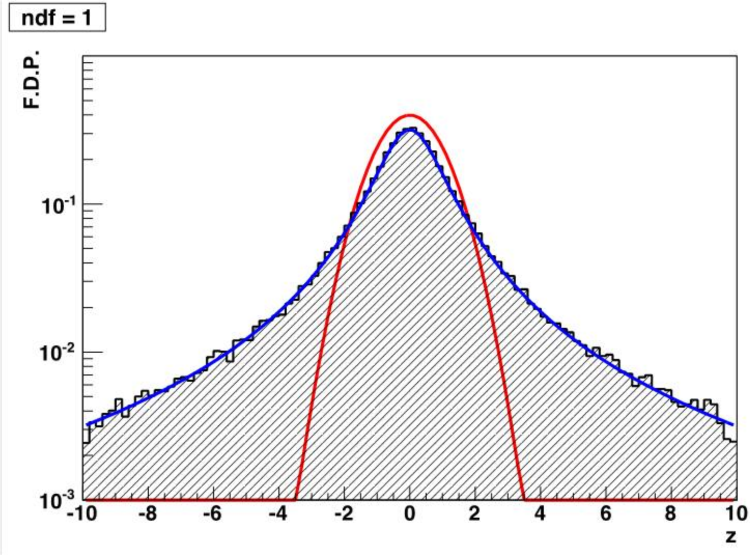
- Sendo n o número de graus de liberdade da distribuição e Γ a função Gama

http://en.wikipedia.org/wiki/Student%27s_t-distribution

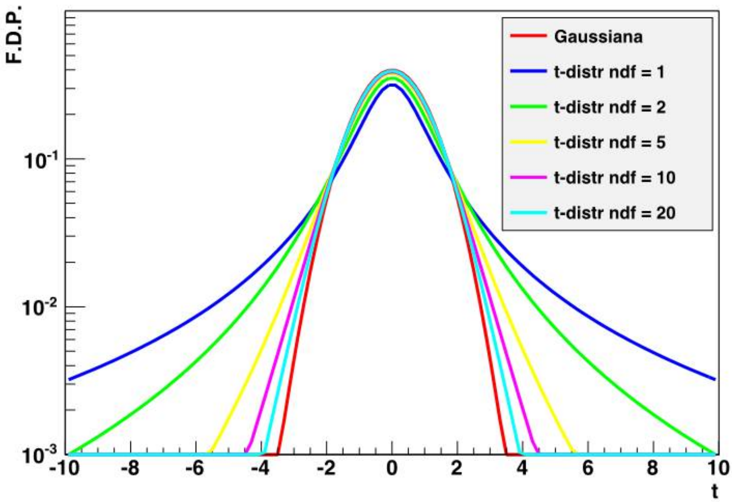
FDP de t para ngl $= 1$



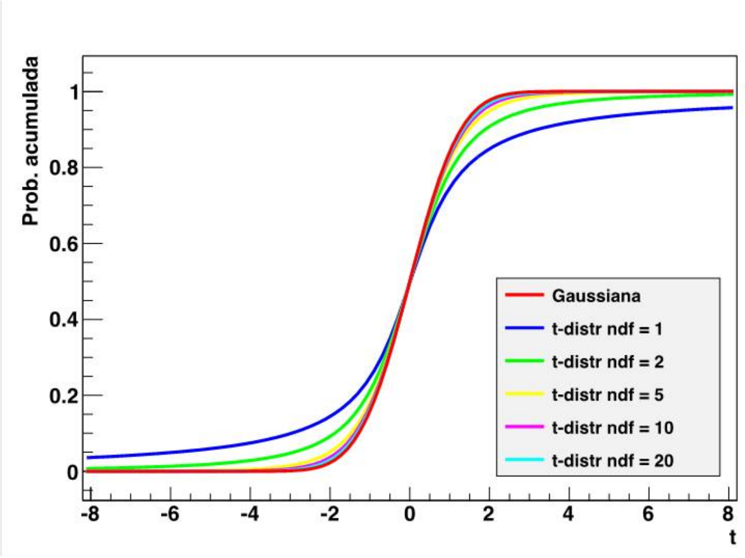
FDP de t para ngl $= 1$



Distribuições-t para diferentes ngl

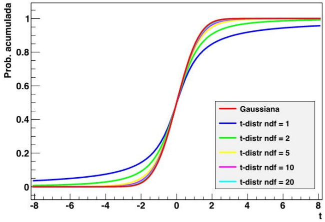


Probabilidade acumulada

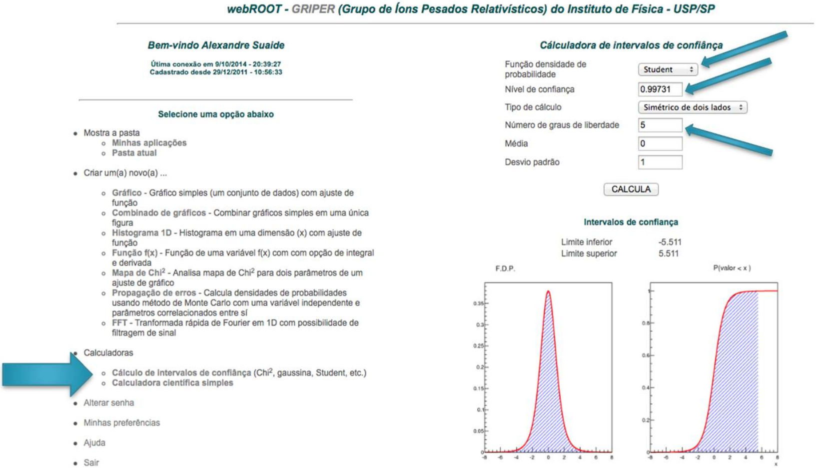


Intervalo de confiança

- Do gráfico de probabilidade acumulada fica claro que, o intervalo de compatibilidade de t para o mesmo nível de confiança (por exemplo, 99%) depende do número de graus de liberdade da amostra
  - Podemos ter $|t| < 8$, em alguns casos, para $\sim$95% de confiança!



Teste-z e teste-t no webroot



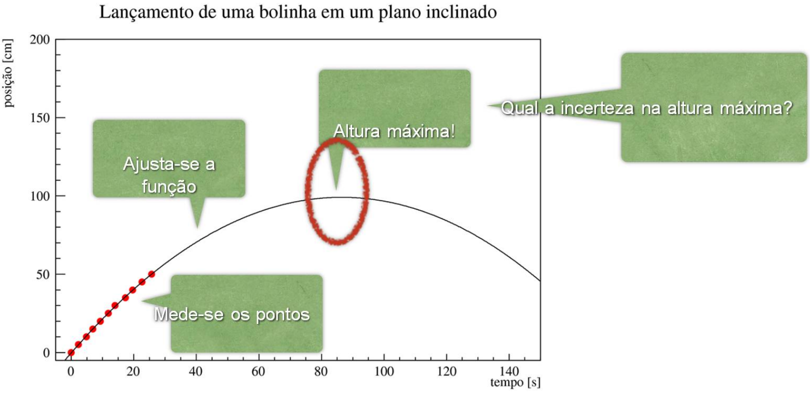
Teste-t de Student

- Leva em conta o fato da FDP para t (ou z) desviar de uma distribuição normal para poucos graus de liberdade
  - Importante para fazer uma comparação justa entre conjuntos de dados
- Para saber mais, olhe em qualquer livro de estatística. Esse é um teste padrão.
  - http://en.wikipedia.org/wiki/Student%27s_t-test
  - http://en.wikipedia.org/wiki/T_distribution

Sumário

O problema clássico

- Extrapolações de funções
  - Ex: lançamento de uma bolinha em um plano inclinado. Qual altura máxima ela atinge?



Fórmula geral de propagação de incertezas

(ver tópicos anteriores)

- Dada uma função

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

Fórmula geral de propagação de incertezas

(ver tópicos anteriores)

- Dada uma função

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

- A variância de y é dada por

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

Fórmula geral de propagação de incertezas

(ver tópicos anteriores)

- Dada uma função

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

- A variância de y é dada por

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

- Expandindo f em Taylor em 1^a ordem

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

Fórmula geral de propagação de incertezas

(ver tópicos anteriores)

- Dada uma função

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

- A variância de y é dada por

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

- Expandindo f em Taylor em 1^a ordem

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

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

Fórmula geral de propagação de incertezas

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

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

Fórmula geral de propagação de incertezas

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

- 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 de propagação de incertezas

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

- 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 de propagação de incertezas

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

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

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

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

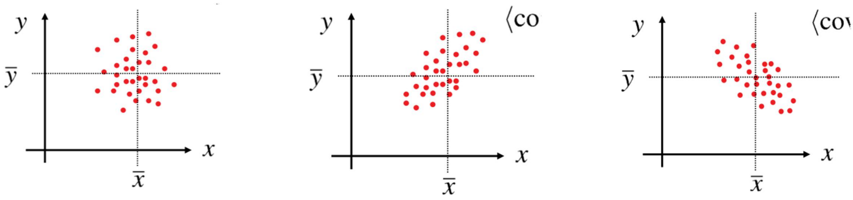
Σ é chamada de matriz de covariância

O significado da covariância

- Considere duas medidas x, y
- Considere que podemos repetir o experimento e medir várias vezes x e y
- Considere que a cada medida, colocamos um ponto no gráfico de y em função de x
  - Calculamos o valor médio de x e de y
  - Calculamos a covariância entre x e y

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

Há três possibilidades



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

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

Coeficiente de correlação

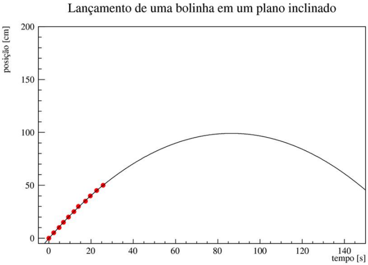
- Muitas vezes é melhor expressar a dependência entre duas grandezas através do coeficiente de correlação, definido como:

$$\rho_{xy} = \frac{\mathrm{cov}_{xy}}{\sigma_x \sigma_y}$$

$$-1 \leqslant \rho_{xy} \leqslant 1$$

Vamos voltar ao problema inicial

- Um bom programa de ajuste fornece a matriz de covariância dos parâmetros ajustados



$$f(x) = [0] + [1]x + [2]x^2$$

Resultados do ajuste

Número de parâmetros	3
Chi^2	5.3756
Número de graus de liberdade	8

parâmetro	Valor	Incerteza
0	-0.211885	0.39859
1	2.29672	0.0700687
2	-0.0132831	0.00256606

Matriz de covariância

$$\begin{bmatrix} 0.158874 & -0.0232467 & 0.000711744 \\ -0.0232467 & 0.00490963 & -0.000173525 \\ 0.000711744 & -0.000173525 & 6.58469E-06 \end{bmatrix}$$

- Quão correlacionados estão os parâmetros [1] e [2]?

Resultados do ajuste

Número de parâmetros	3
Chi^2	5.3756
Número de graus de liberdade	8

parâmetro	Valor	Incerteza
0	-0.211885	0.39859
1	2.29672	0.0700687
2	-0.0132831	0.00256606

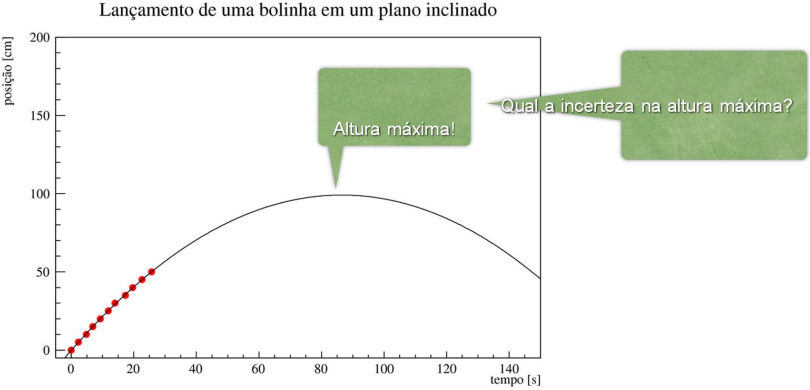
Matriz de covariância

$$\begin{bmatrix} 0.158874 & -0.0232467 & 0.000711744 \\ -0.0232467 & 0.00490963 & -0.000173525 \\ 0.000711744 & -0.000173525 & 6.58469E-06 \end{bmatrix}$$

$$\rho_{12} = \frac{\mathrm{cov}_{12}}{\sigma_1 \sigma_2} = -0.96$$

- Estes parâmetros estão altamente correlacionados

Voltando ao problema original



$$f(x) = [0] + [1]x + [2]x^2 \quad \longrightarrow \quad H_{max} = [0] - \frac{1}{4}\frac{[1]^2}{[2]}$$

Voltando ao problema original

$$H_{max} = [0] - \frac{1}{4}\frac{[1]^2}{[2]}$$

$$\sigma^2_{H_{max}} = \Gamma^T \Sigma \Gamma$$

$$\Gamma = \begin{pmatrix} \partial_0 \\ \partial_1 \\ \partial_2 \end{pmatrix} \text{ com } \partial_i = \frac{\partial H_{max}}{\partial_{[i]}} \quad \Sigma = \begin{pmatrix} 0.158874 & -0.0232467 & 0.000711744 \\ -0.0232467 & 0.00490963 & -0.000173525 \\ 0.000711744 & -0.000173525 & 6.58469E-06 \end{pmatrix}$$

Façam esta propagação como exercício!
Façam também considerando $\mathrm{cov} = 0$ e vejam a diferença.

Perguntas em aberto

- Como a matriz de covariância é calculada? O que ela tem a ver com o ajuste de uma função?
  - O que o χ^2 tem a ver com isto?
- Como eu sei se duas grandezas estão correlacionadas ou não?
  - No caso de um ajuste é fácil mas e no caso de duas grandezas medidas em um experimento?

Sumário

Matriz de covariância

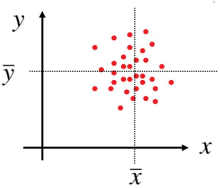
- Fórmula geral de propagação de incertezas pode ser escrita em uma forma matricial do tipo

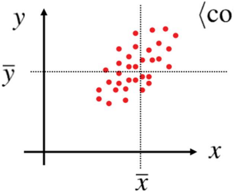
$$\sigma_y^2 = \Gamma^T \Sigma \Gamma$$

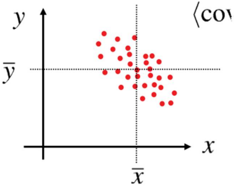
$$\Gamma = \begin{pmatrix} \partial_{x_1} \\ \partial_{x_2} \\ \dots \\ \partial_{x_n} \end{pmatrix} \text{ com } \partial_{x_i} = \frac{\partial f}{\partial x_i} \qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \mathrm{cov}_{12} & \cdots & \mathrm{cov}_{1n} \\ \mathrm{cov}_{21} & \sigma_2^2 & \cdots & \mathrm{cov}_{2n}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

Σ é chamada de matriz de covariância

Covariância e correlação





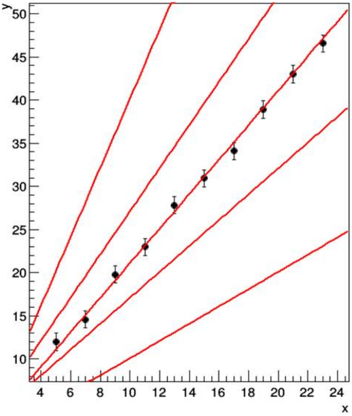


$\mathrm{cov}_{xy} = 0$ $\qquad$ $\mathrm{cov}_{xy} > 0$ $\qquad$ $\mathrm{cov}_{xy} < 0$

$$\rho_{xy} = \frac{\mathrm{cov}_{xy}}{\sigma_x \sigma_y} \qquad -1 \leqslant \rho_{xy} \leqslant 1$$

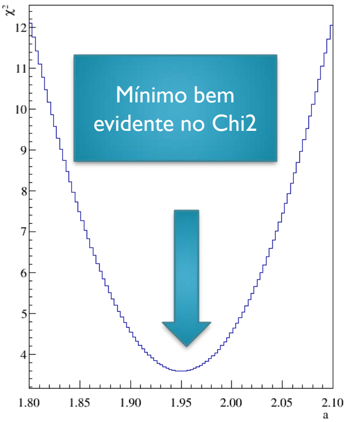
Mapa de χ^2

- Método dos mínimos quadrados
- Maximizar a probabilidade da função descrever os dados $=$ minimizar o χ^2

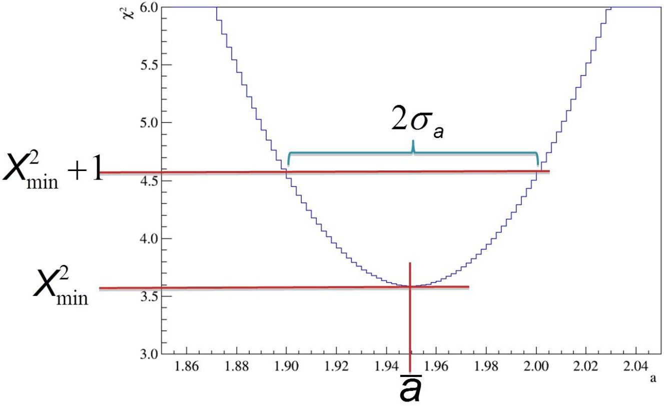


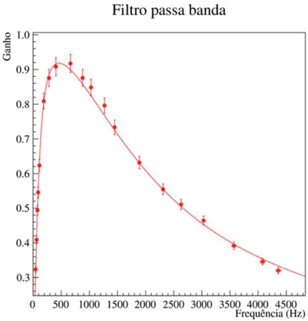
Mapa de χ^2

- Método dos mínimos quadrados
- Maximizar a probabilidade da função descrever os dados = minimizar o χ^2



Incerteza no parâmetro ajustado





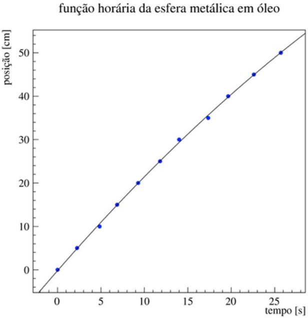
Resultados do ajuste

Número de parâmetros	2
Chi^2	8.48087
Número de graus de liberdade	18

parâmetro	Valor	Incerteza
0	137.519	2.29791
1	1541	20.2911

Matriz de covariância

$$\begin{bmatrix} 5.28038 & 0.814965 \\ 0.814965 & 411.73 \end{bmatrix}$$



Resultados do ajuste

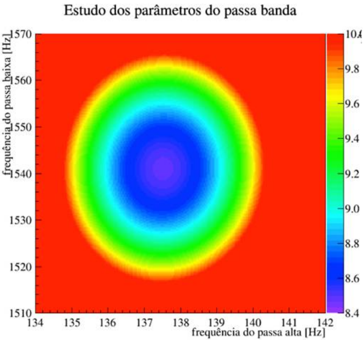
Número de parâmetros	3
Chi^2	6.77906
Número de graus de liberdade	8

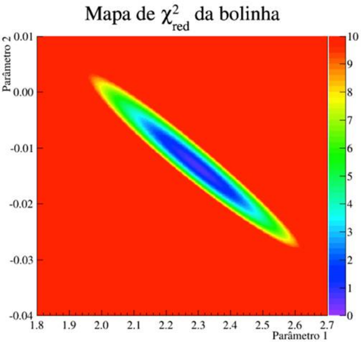
parâmetro	Valor	Incerteza
0	-0.216646	0.362052
1	2.29752	0.062854
2	-0.0133085	0.00228345

Matriz de covariância

$$\begin{bmatrix} 0.131082 & -0.0190145 & 0.000579051 \\ -0.0190145 & 0.00395063 & -0.000138616 \\ 0.000579051 & -0.000138616 & 5.21416E-06 \end{bmatrix}$$

Mapa de χ^2





$$\rho_{01} = \frac{\text{cov}_{01}}{\sigma_0 \sigma_1} = 0.02$$

$$\rho_{12} = \frac{\text{cov}_{12}}{\sigma_1 \sigma_2} = -0.96$$

- Como extrair a covariância ou correlação entre dois parâmetros do mapa de χ^2?

Estudando o mapa de χ^2

- Vamos começar com duas grandezas gaussianas, independentes entre si, cada uma com uma variância conhecida. A probabilidade de obtermos, simultaneamente, um determinado valor de a e b é:

$$P(a,b) = P(a)P(b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left[-\frac{1}{2}\left(\frac{a-\mu_a}{\sigma_a}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{b-\mu_b}{\sigma_b}\right)^2\right]$$

$$P(a,b) = P(a)P(b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left\{-\frac{1}{2}\left[\left(\frac{a-\mu_a}{\sigma_a}\right)^2 + \left(\frac{b-\mu_b}{\sigma_b}\right)^2\right]\right\}$$

Estudando o mapa de χ^2

$$P(a,b) = P(a)P(b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left\{-\frac{1}{2}\left[\left(\frac{a-\mu_a}{\sigma_a}\right)^2 + \left(\frac{b-\mu_b}{\sigma_b}\right)^2\right]\right\}$$

$$P(a,b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left(-\frac{1}{2}\chi^2\right)$$

- A probabilidade é máxima quando o χ^2 é mínimo

Os contornos do mapa de χ^2

- Limites no mapa de χ^2

$$1\sigma \rightarrow \chi^2 = \chi^2_{min} + 1$$

$$2\sigma \rightarrow \chi^2 = \chi^2_{min} + 4$$

- Podemos desenhar estas linhas

Mapa para duas grandezas independentes

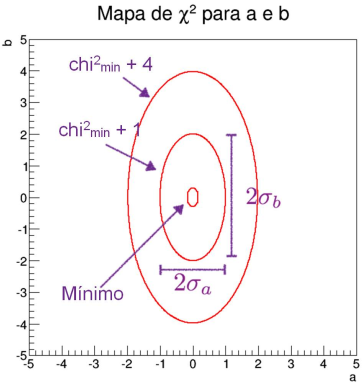
- Assumindo valor médio zero para ambas e

$$\sigma_a = 1$$

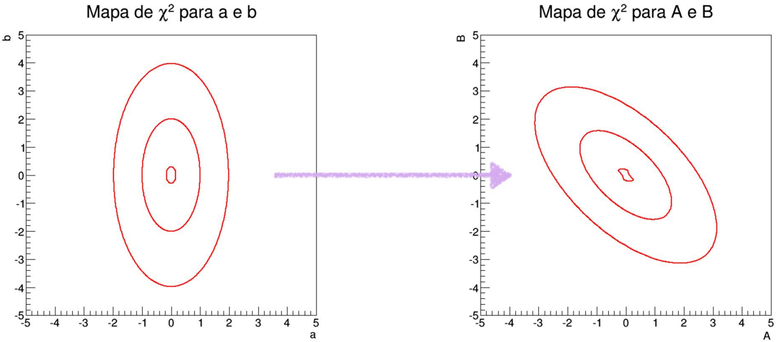
$$\sigma_b = 2$$

- Contornos em 1 e 2 sigmas

$$P(a,b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left(-\frac{1}{2}\chi^2\right)$$

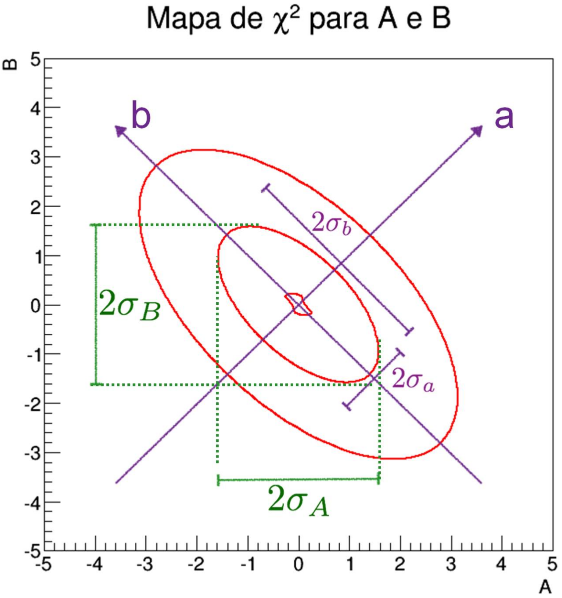


Introduzir covariância significa girar estas elipses



$$a, b \rightarrow A, B$$

Relação de A, B com a, b



Muito bem...

- Como eu matematizo esta rotação?
- Como eu extraio as relações entre as incertezas e as covariâncias?

Girando uma elipse

- Rotação de um ângulo θ

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \cos\theta & \mathrm{sen}\theta \\ -\mathrm{sen}\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

- Simplificando a notação

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

$$A = ca + sb$$

$$B = -sa + cb$$

Calculando a covariância entre A e B

- Calculando a covariância

$$\mathrm{cov}_{AB} = \langle (A - \mu_A)(B - \mu_B) \rangle = \langle AB \rangle$$

$$\mathrm{cov}_{AB} = \langle (ca + sb)(-sa + cb) \rangle$$

$$\mathrm{cov}_{AB} = \left\langle scb^2 - sca^2 + (c^2 - s^2)ab \right\rangle$$

- Como a e b são independentes

$$\mathrm{cov}_{AB} = sc\left(\sigma_b^2 - \sigma_a^2\right)$$

Calculando a covariância entre A e B

- Escrevendo a covariância em termos do coeficiente de correlação

$$\rho_{AB}\sigma_A\sigma_B = sc\left(\sigma_b^2 - \sigma_a^2\right)$$

- Como eliminar a dependência com o ângulo?
  - Podemos calcular as variâncias de A e B

Calculando variâncias

- Da definição

$$\sigma_A^2 = \left\langle (A - \mu_A)^2 \right\rangle = \left\langle A^2 \right\rangle$$

$$\sigma_A^2 = \left\langle (ca + sb)^2 \right\rangle$$

$$\sigma_A^2 = \left\langle c^2 a^2 + s^2 b^2 + 2csab \right\rangle$$

- Como a e b não possuem covariância

$$\sigma_A^2 = c^2 \sigma_a^2 + s^2 \sigma_b^2$$

- Similarmente para B

$$\sigma_B^2 = s^2 \sigma_a^2 + c^2 \sigma_b^2$$

Sendo assim

- Calculando o produto das variâncias

$$\sigma_A^2\sigma_B^2 = s^2c^2(\sigma_a^4 + \sigma_b^4) + (c^4 + s^4)\sigma_a^2\sigma_b^2$$

- Comparando ao quadrado de

$$\rho_{AB}\sigma_A\sigma_B = sc\left(\sigma_b^2 - \sigma_a^2\right)$$

- Com pouca álgebra, chega-se a

$$\rho_{AB}^2 = 1 - \left(\frac{\sigma_a\sigma_b}{\sigma_A\sigma_B}\right)^2$$

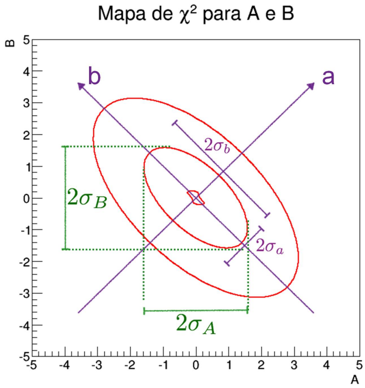
Em resumo...

$$A = ca + sb \quad \text{e} \quad B = -sa + cb$$

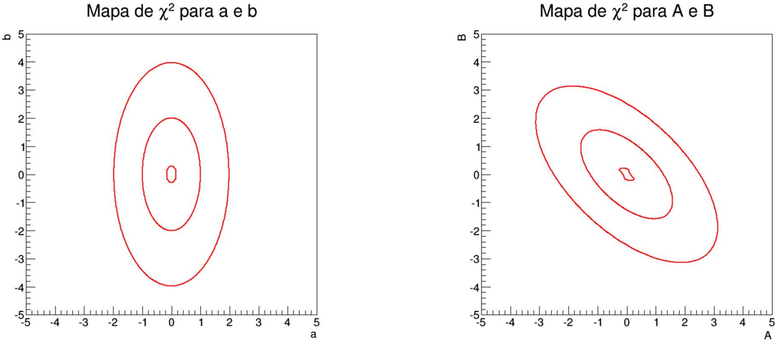
$$\sigma_A^2 = c^2\sigma_a^2 + s^2\sigma_b^2$$

$$\sigma_B^2 = s^2\sigma_a^2 + c^2\sigma_b^2$$

$$\rho_{AB}^2 = 1 - \left(\frac{\sigma_a\sigma_b}{\sigma_A\sigma_B}\right)^2$$



Como descrever as curvas de chi2 com covariância?



$$P(a,b) = \frac{1}{2\pi\sigma_a\sigma_b}\exp\left\{-\frac{1}{2}\left[\left(\frac{a}{\sigma_a}\right)^2 + \left(\frac{b}{\sigma_b}\right)^2\right]\right\} \qquad P(A,B) = ?$$

Descrevendo $P(A, B)$

- Começamos escrevendo a e b em função de A e B

$$a = cA - sB \quad \text{e} \quad b = sA + cB$$

- Substituímos em

$$P(a, b) = \frac{1}{2\pi\sigma_a\sigma_b} \exp\left\{-\frac{1}{2}\left[\left(\frac{a}{\sigma_a}\right)^2 + \left(\frac{b}{\sigma_b}\right)^2\right]\right\}$$

- É necessário que (MOSTRE ISTO)

$$P(a, b)\, da\, db = P(A, B)\, dA\, dB$$

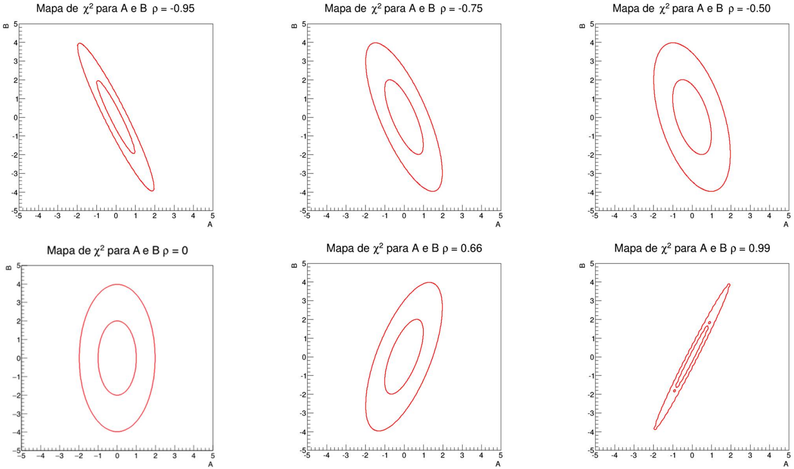
Descrevendo $P(A, B)$

- Com um pouco de álgebra e substituindo as relações necessárias entre as variâncias de a, b e A, B

$$P(A,B) = \frac{1}{2\pi\sqrt{1-\rho^2}\sigma_A\sigma_B} \exp\left\{-\frac{1}{2}\left(\frac{1}{1-\rho^2}\right)\left[\left(\frac{A}{\sigma_A}\right)^2 + \left(\frac{B}{\sigma_B}\right)^2 - \frac{2\rho AB}{\sigma_A\sigma_B}\right]\right\}$$

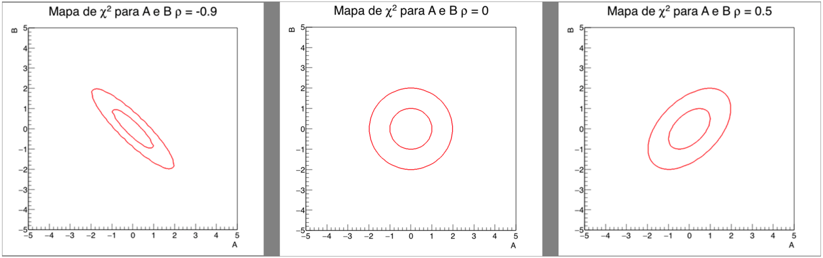
- Note que, se a correlação for nula ($\rho = 0$), voltamos à expressão para duas grandezas independentes

$$\sigma_A = 1 \text{ e } \sigma_B = 2$$



Algumas curvas

$$\sigma_A = \sigma_B = 1$$



- Incertezas iguais não significa falta de correlação $\Rightarrow$ preste atenção nisto

Alguns comentários

- Introduzir correlações entre duas grandezas significa "girar" uma curva de χ^2
  - Sempre podemos redefinir variáveis de modo a eliminar correlações entre elas
- Do mapa de χ^2 é possível extrair a correlação entre duas grandezas "geometricamente"