Propagação de incertezas:

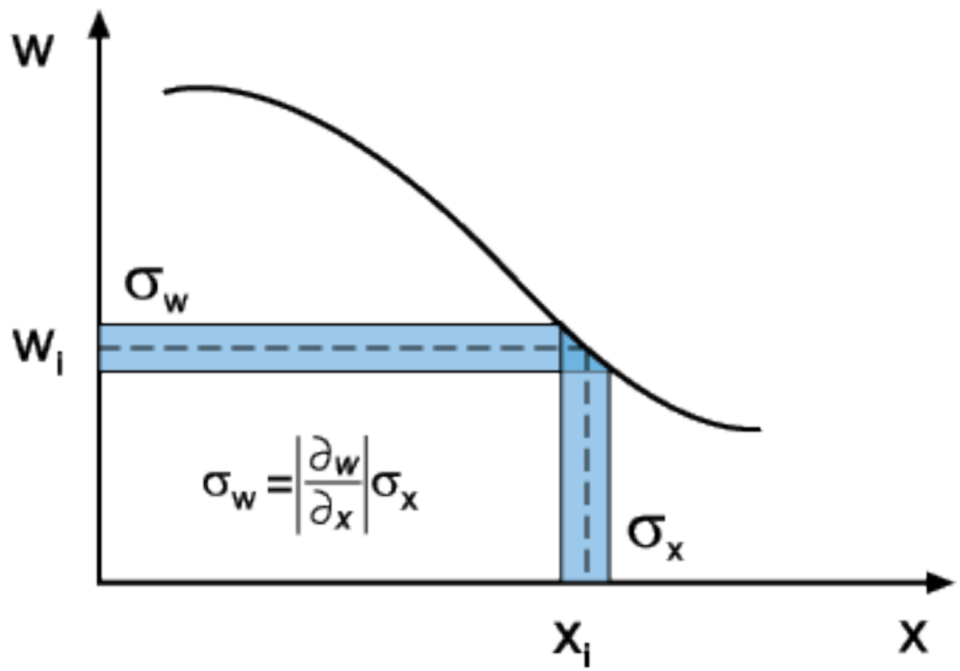
Método de Monte Carlo

Propagação de incertezas

- Fórmula geral de propagação de incertezas para incertezas pequenas e funções "bem comportadas".

$$\sigma^2 = C^T \Sigma C = \sum_i \sum_j \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \Sigma_{ij}, \quad \begin{cases} \sigma_i^2 = \Sigma_{ii} \\ \mathrm{cov}_{ij} = \Sigma_{ij} \end{cases}$$

- Como a variação de uma grandeza provoca variação em outra?



Propagar incertezas

- Se quero saber o quanto as incertezas de medidas afetam outras grandezas precisamos propagar as incertezas
- Em situações simples a avaliação é fácil

$$P = Vi$$

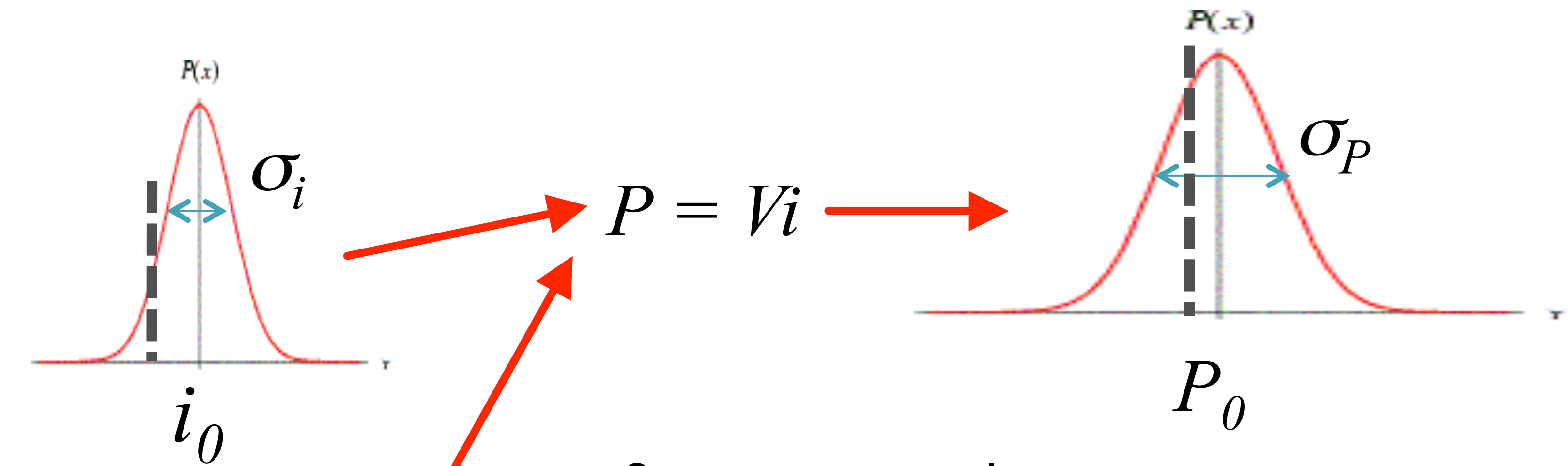
- Mas como fazer em situações mais complexas?

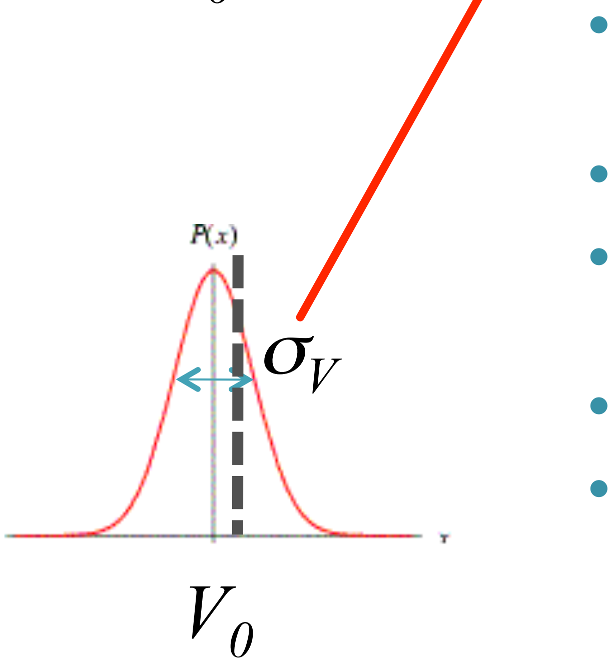
$$I_{visível} = \int_{400}^{800} I d\lambda \quad \begin{cases} \mathrm{I} = \mathrm{I} \pm \sigma_{\mathrm{I}} \\ \lambda = \lambda \pm \sigma_{\lambda} \end{cases} \Rightarrow \sigma_{I_{visível}} = ?$$

- Simulações de Monte Carlo

Caso simples, $P = Vi$

- $i = i_0 \pm \sigma_i; \quad V = V_0 \pm \sigma_V$
- $P = P_0 \pm \sigma_P$, quem é σ_P?





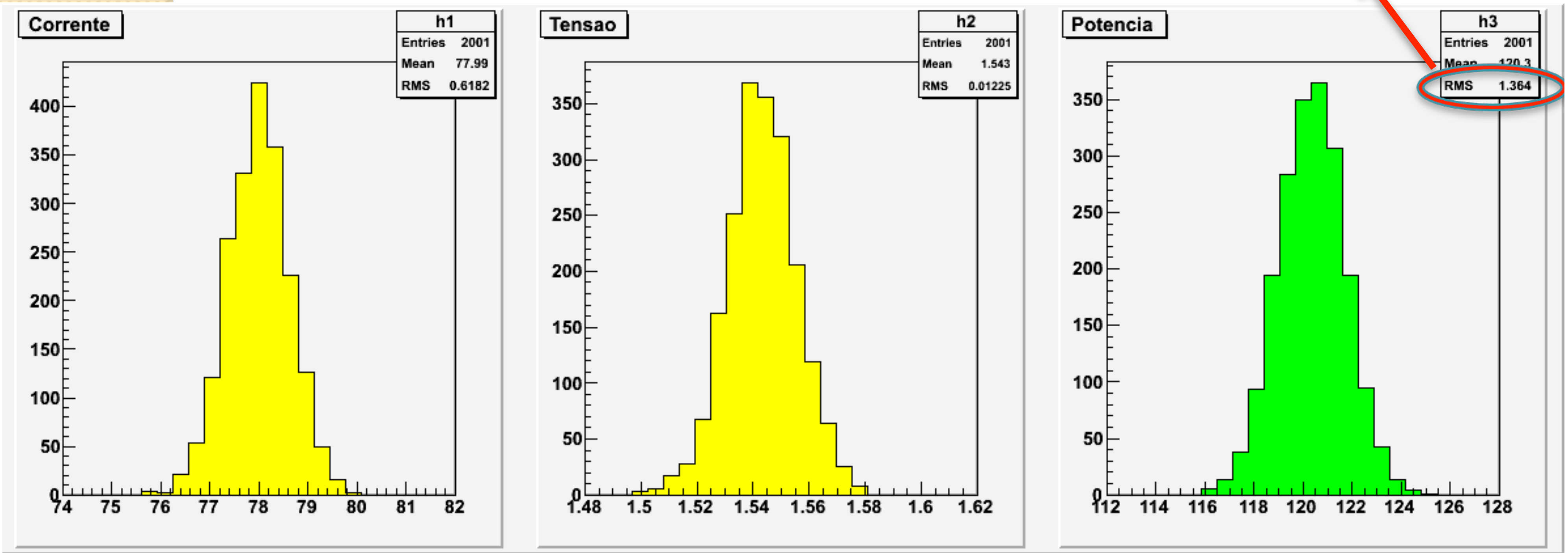
- Sorteia-se um valor para a primeira grandeza
- Sorteia-se um valor para a segunda grandeza
- Calcula-se a grandeza derivada com estes valores sorteados
- Repete-se o cálculo várias vezes
- A incerteza da grandeza derivada será o desvio padrão de todos os valores calculados

Exemplo: $P = V\,i$

$$i = 78.0 \pm 0.6\ mA \qquad V = 1.543 \pm 0.012\ V$$

$$P = V\,i = 120.3 \pm ?\ mW$$

$$\sigma_P = 1.4\ mW$$



Como gerar números aleatórios

- Em geral, toda linguágem de programação possui um gerador de números aleatórios
  - Planilhas eletrônicas também
  - RND, RAND, etc.

- O problema é que, em geral, estas funções geram apenas distribuições uniformes
  - Como gerar uma distribuição gaussiana ou com outra F.D.P.?

Transformação inversa

- Probabilidade acumulada

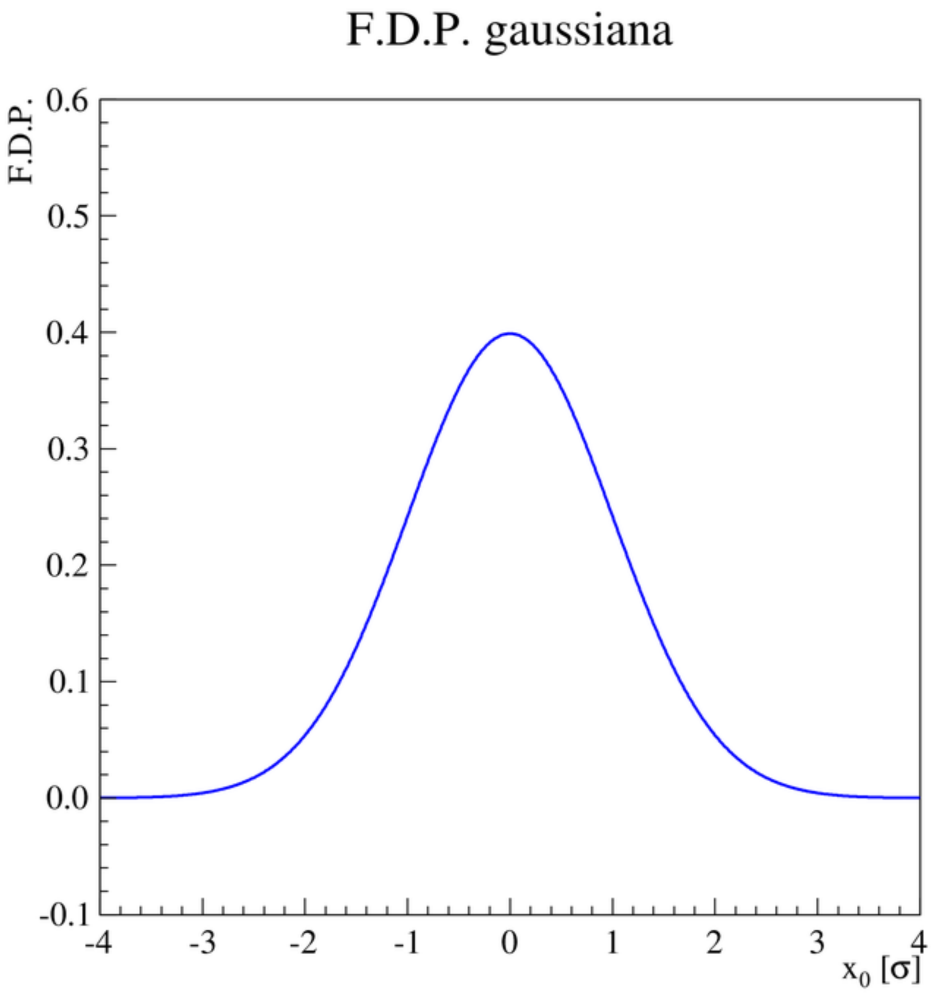
$$p = P(x < A) \rightarrow A = P^{-1}(p)$$

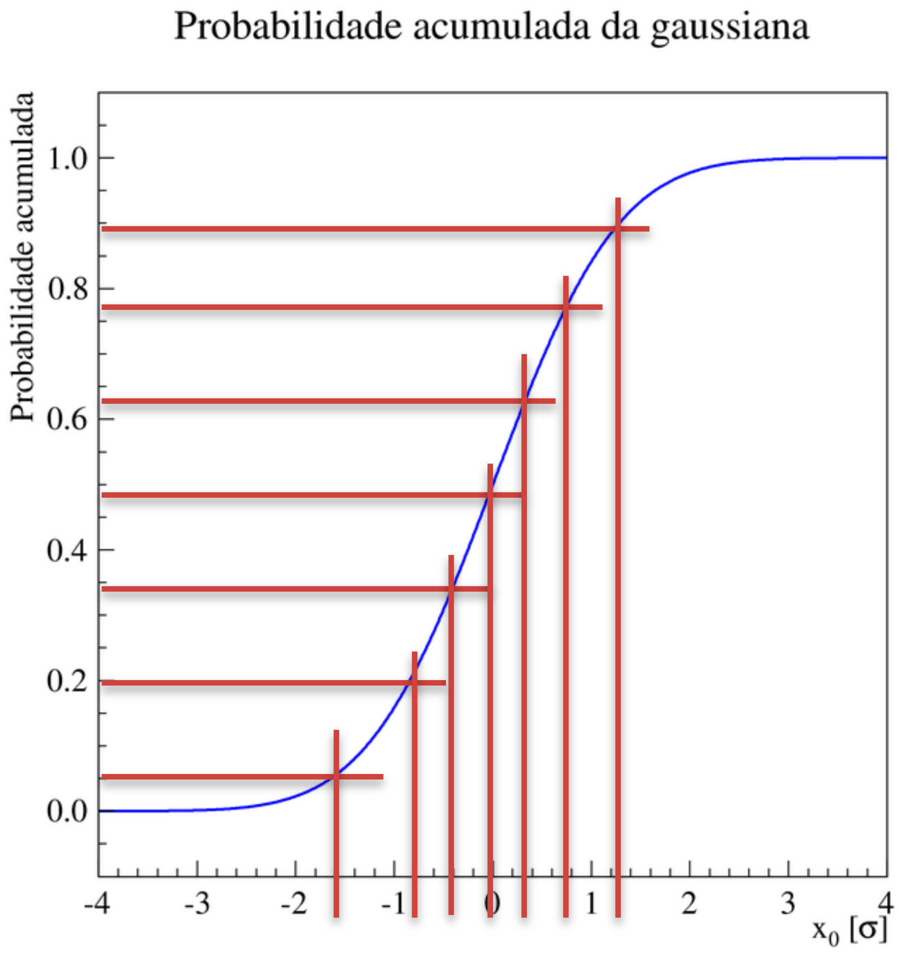
- Só que:

$$p = [0,1]$$

- Sorteia-se p com distribuição uniforme entre 0 e 1 e, conhecendo-se a inversa da probabilidade acumulada, determina-se A.
  - Ou simplesmente rebate do eixo y para o x

Ex: gaussiana





Cálculo no excel

- Para sortear um número aleatório, com distribuição Gaussiana no excel, dado

$$X = X_0 \pm \sigma_X$$

No excel usa-se a expressão

NORMINV(RAND(), média, sigma)

Ver planilha junto com as notas de aula

Vantagens deste método

- O conceito é bastante intuitivo.
- Fácil de implementar em planilhas eletrônicas (Excel, OO, etc.).
- Não é necessário fazer as derivadas parciais para propagar as incertezas.
- Independente da complexidade das contas, que podem tornar o cálculo de derivadas parciais muito complicados.
  - OBS: Estamos assumindo que não há covariância entre as grandezas a serem propagadas → Em geral é verdade para medidas diretas.
- Planilhas exemplos junto com notas de aula.

Testes de compatibilidade:

O teste-z e o teste-t

Testes de compatibilidade

- Teste z
  - Imagine que foi realizada uma medida e queremos comparar 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:

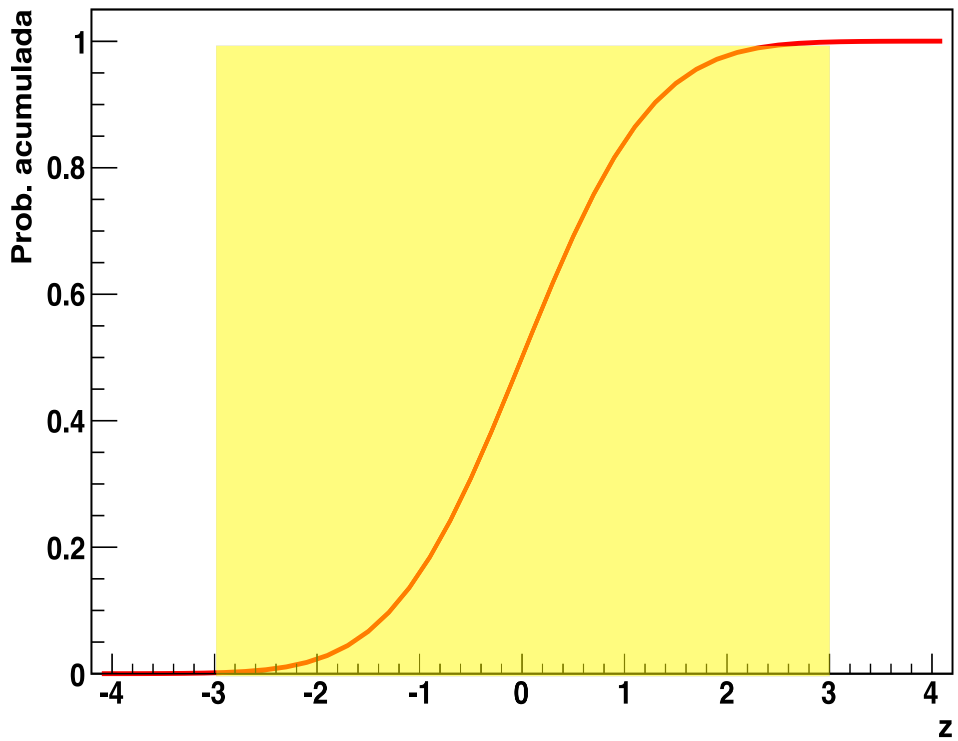
$$P(x) = \int_{-\infty}^{x} p(t)dt \rightarrow \begin{cases} \text{Probabilidade da variável } t \\ \text{assumir um valor menor ou} \\ \text{igual a } x. \end{cases}$$

- 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 F.D.P. normal
- Se a hipótese do teste for verdadeira z deve possuir valor verdadeiro zero e variância 1.



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

No teste-z, 3 significa que ~99.9% das medidas devem estar nesse intervalo, ou seja, a chance de obter um valor com $|z| > 3$ é desprezível

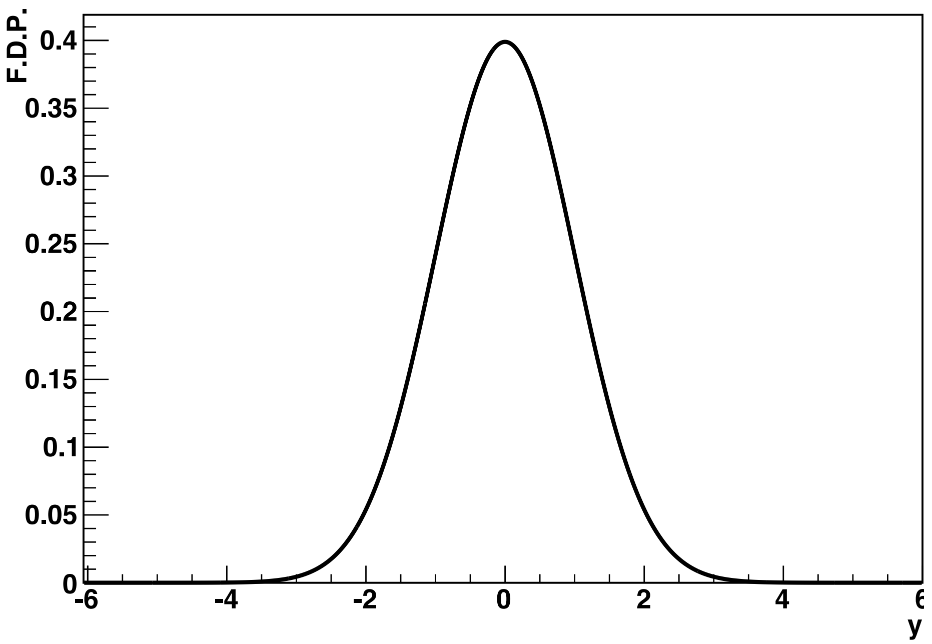
CUIDADO! Não usem o valor 3 a esmo. Entenda o seu significado

Teste-z

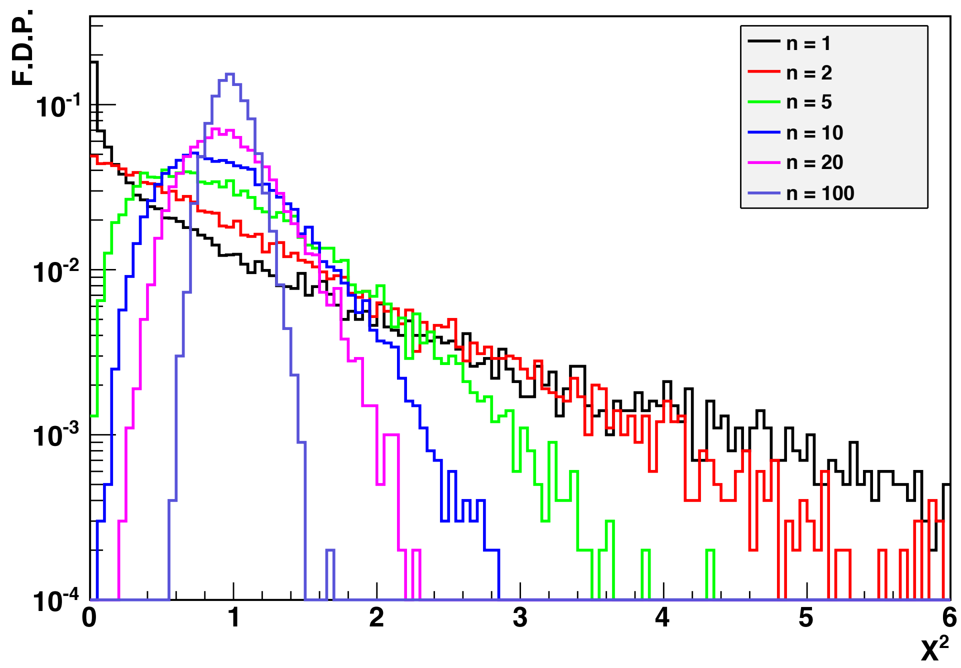
- Mas será que a variável z assume uma F.D.P. normal?

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

F.D.P. para y



F.D.P. para σ



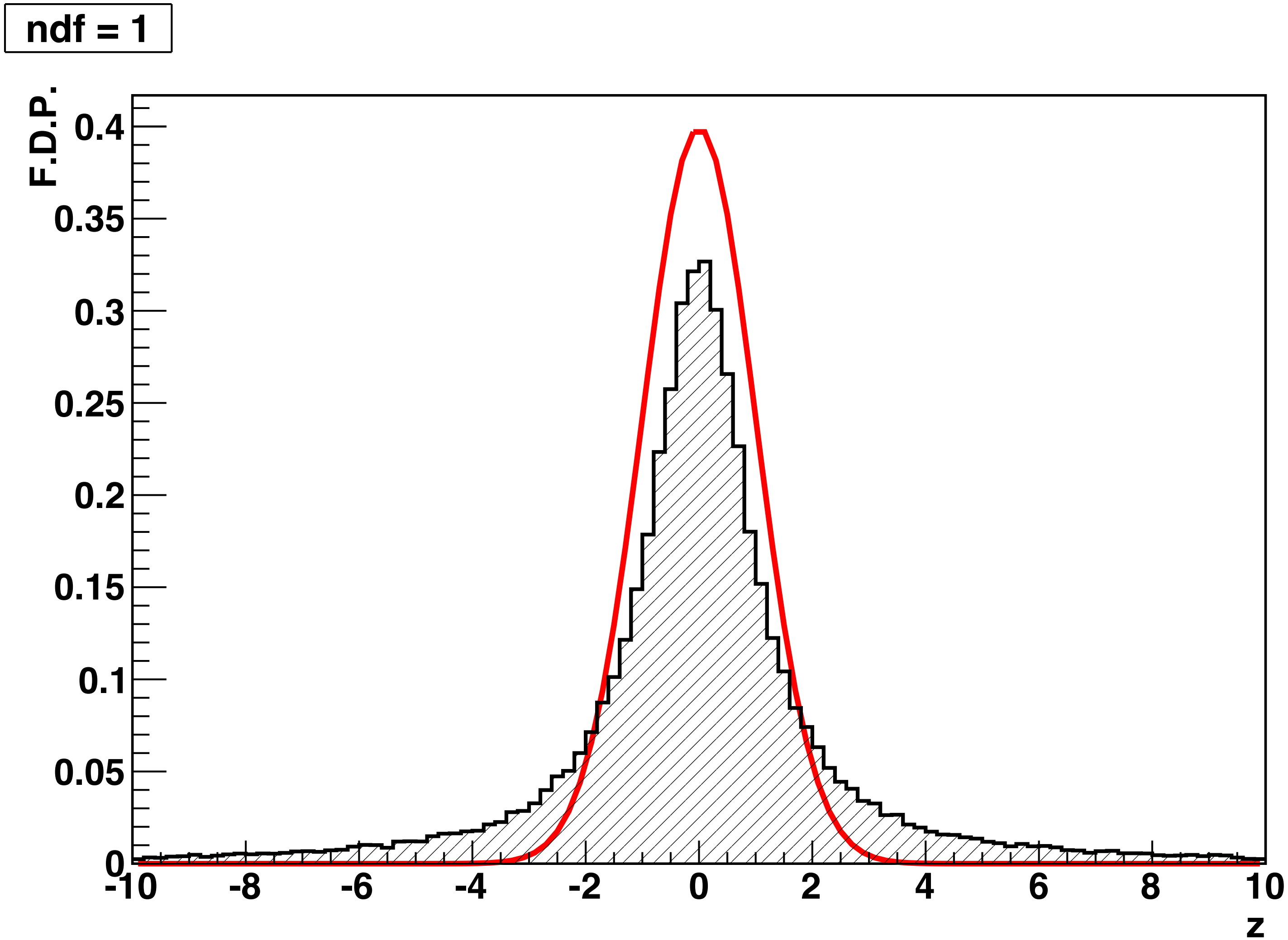
Teste-z

- Vamos fazer o mesmo procedimento (exp. virtuais) que fizemos na aula passada e obter a F.D.P. de z para diferentes números de graus de liberdade (ndf)

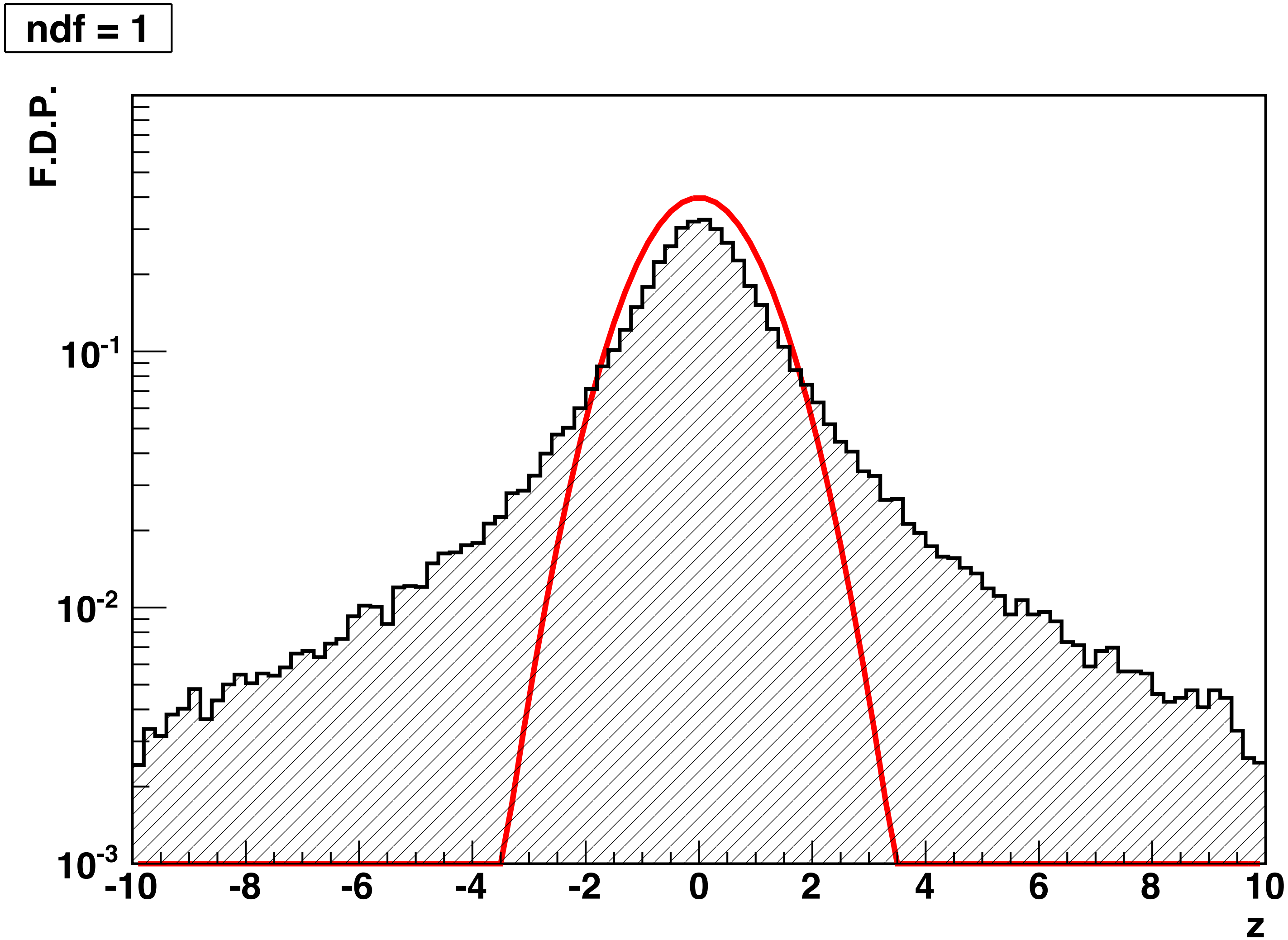
$$z = \frac{\bar{y} - \mu}{\sigma_{\bar{y}}}$$

- Ou seja, vamos simular, para um dado ndf 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.

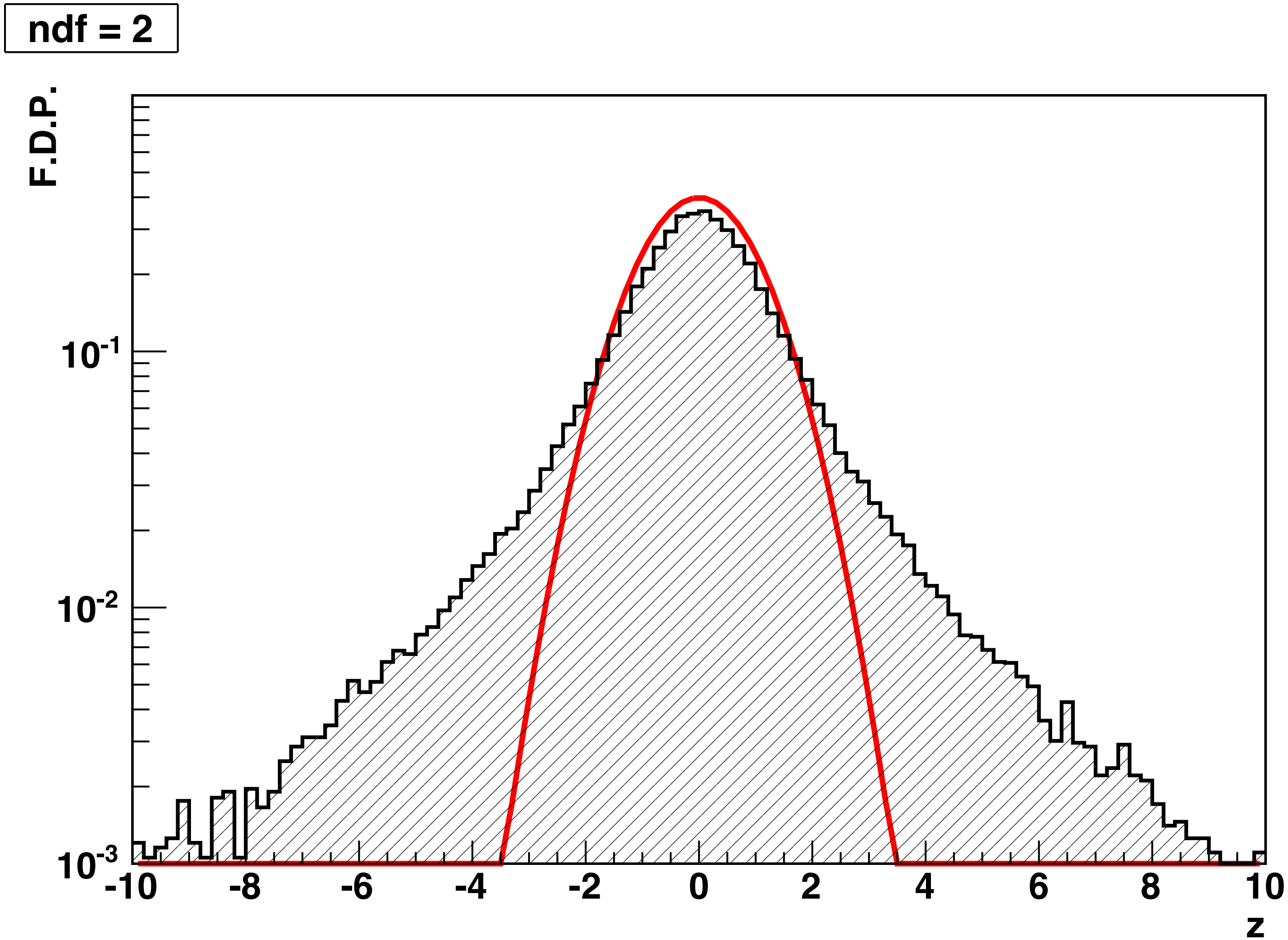
F.D.P. de z para ndf = 1



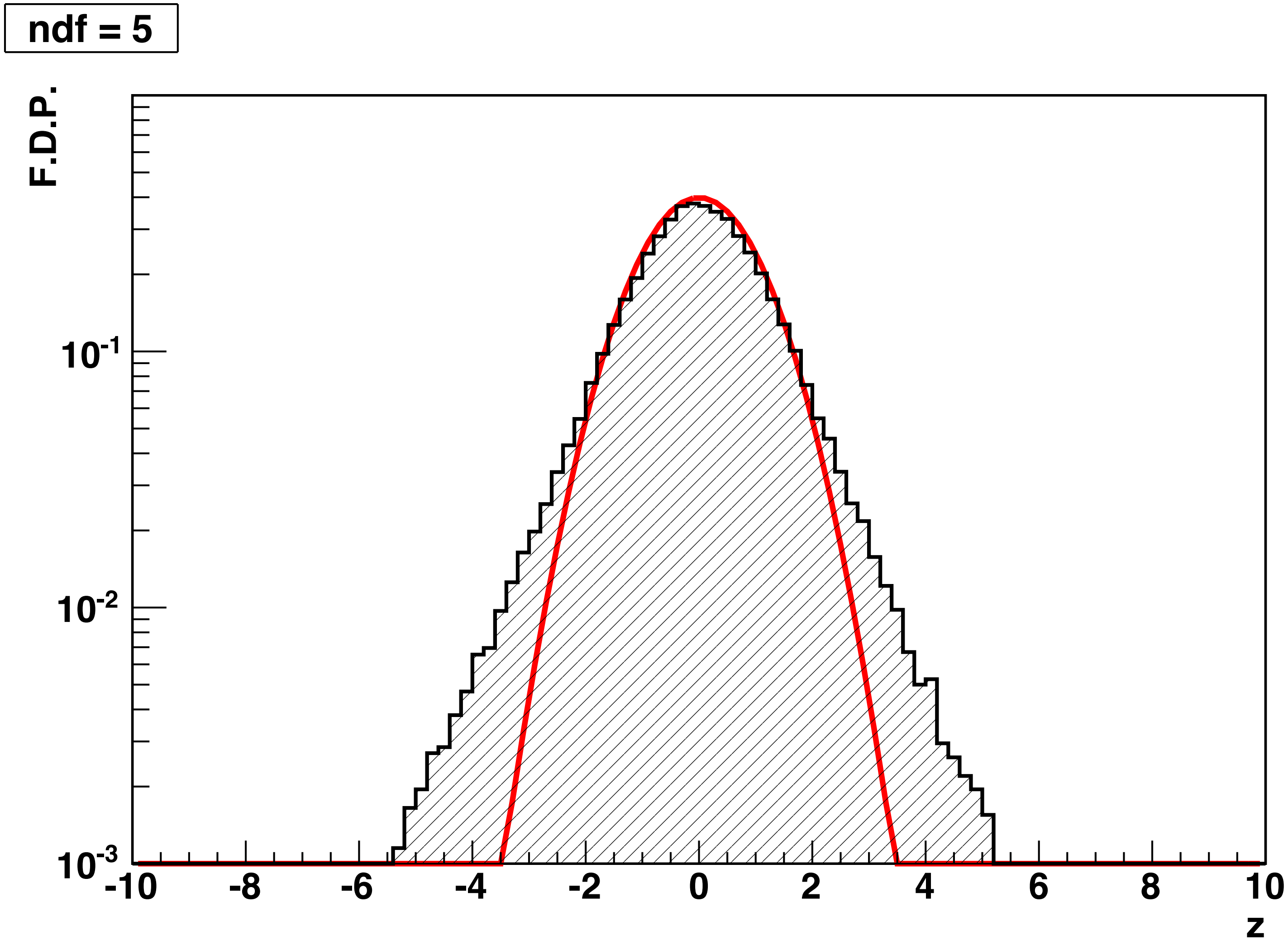
F.D.P. de z para ndf = 1



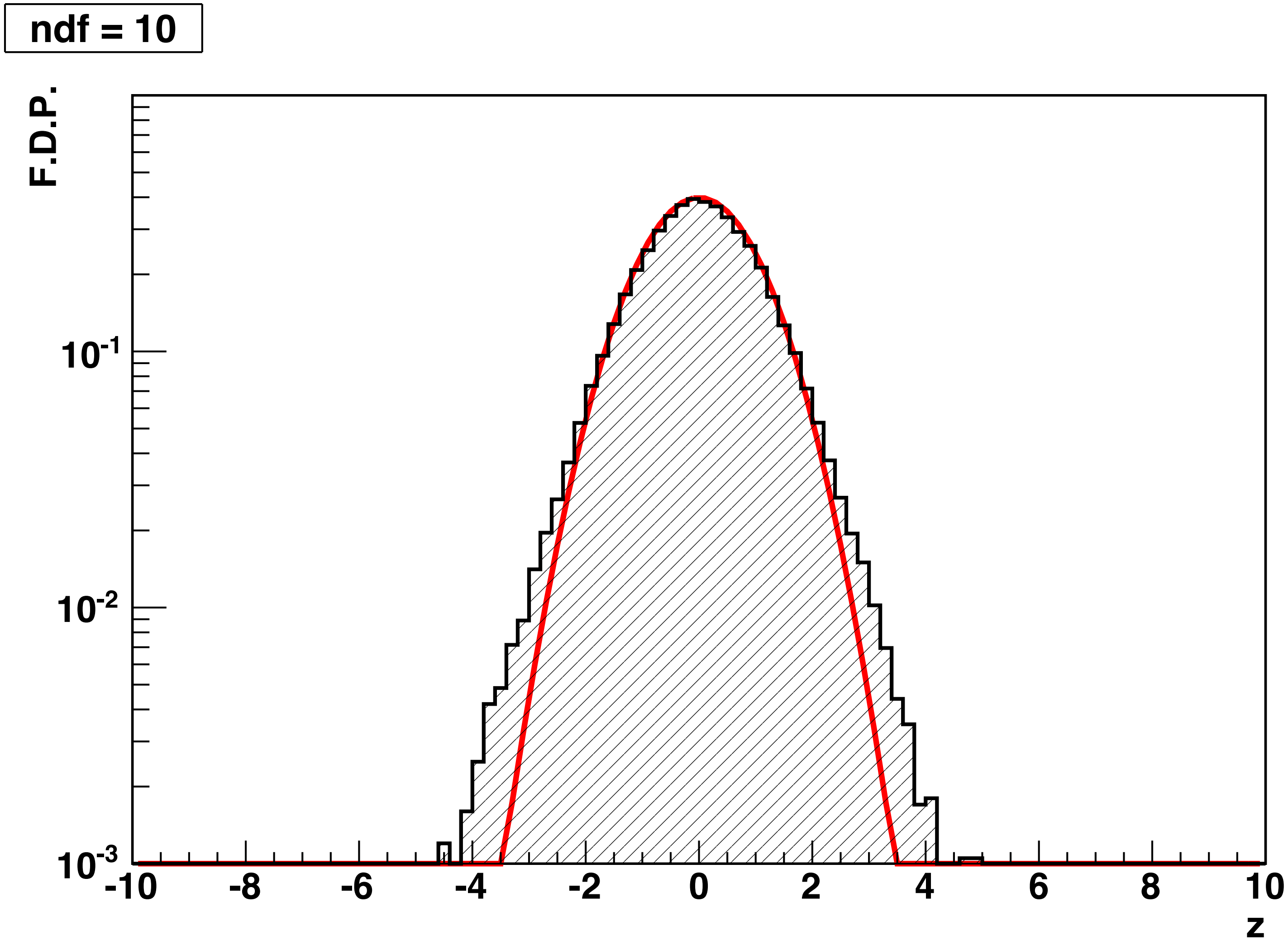
F.D.P. de z para ndf = 2



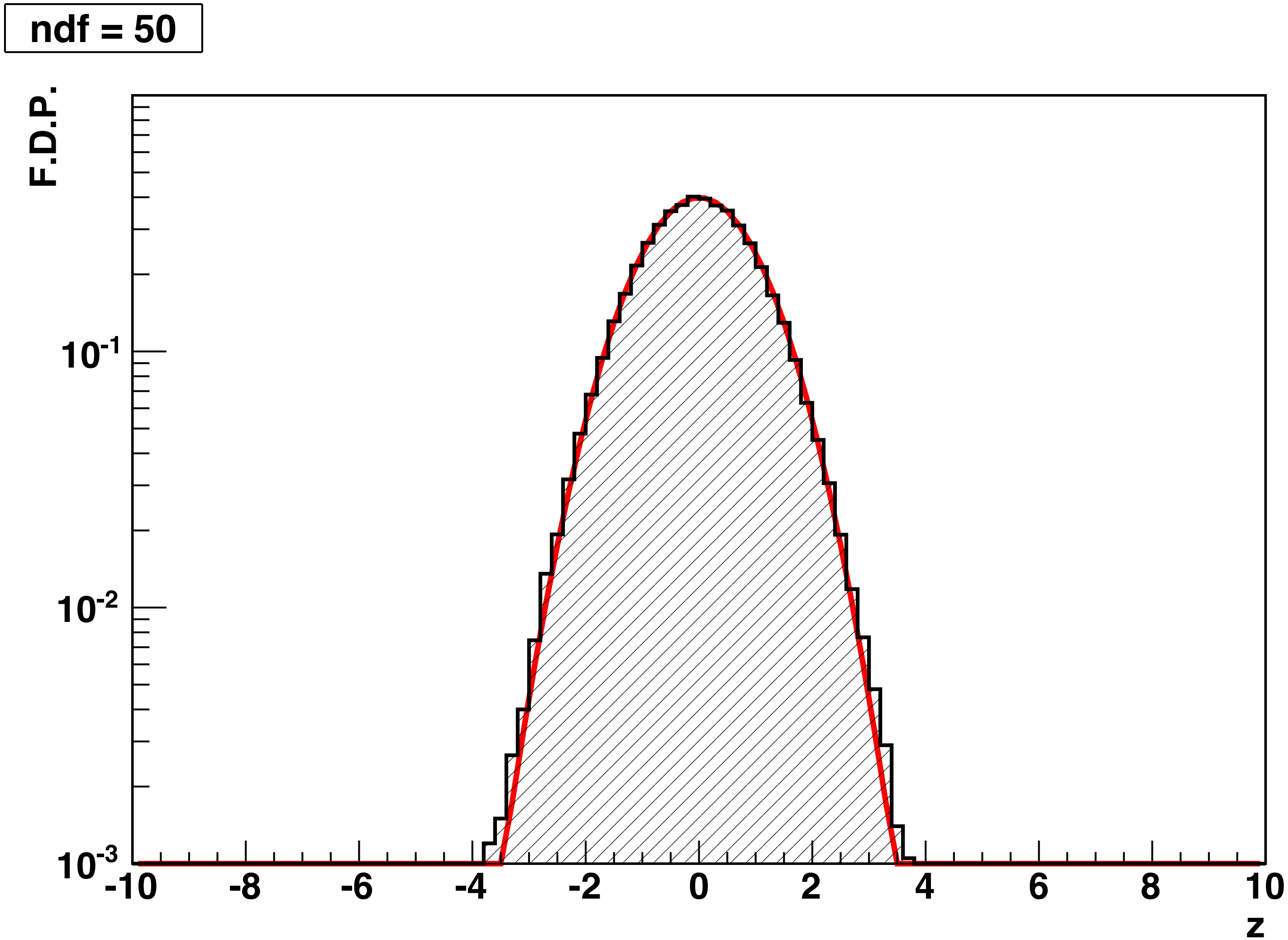
F.D.P. de z para ndf = 5



F.D.P. de z para ndf = 10



F.D.P. de z para ndf = 50



Conclusões

- A variável z somente possui F.D.P. normal para medidas com elevado número de graus de liberdade
  - Tipicamente ndf > 20 já dá para aproximar com um pouco de ressalva.
- Ou seja, o teste-z só pode ser utilizado em situações com grande ndf.
  - A situação é crítica para ndf ~ 1-2
- Que tipo de teste utilizar para baixos ndf?

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 F.D.P. da variância não seguir uma distribuição normal, nesses casos, faz com que seja necessário utilizar distribuições t de probabilidade.

Teste-t de Student

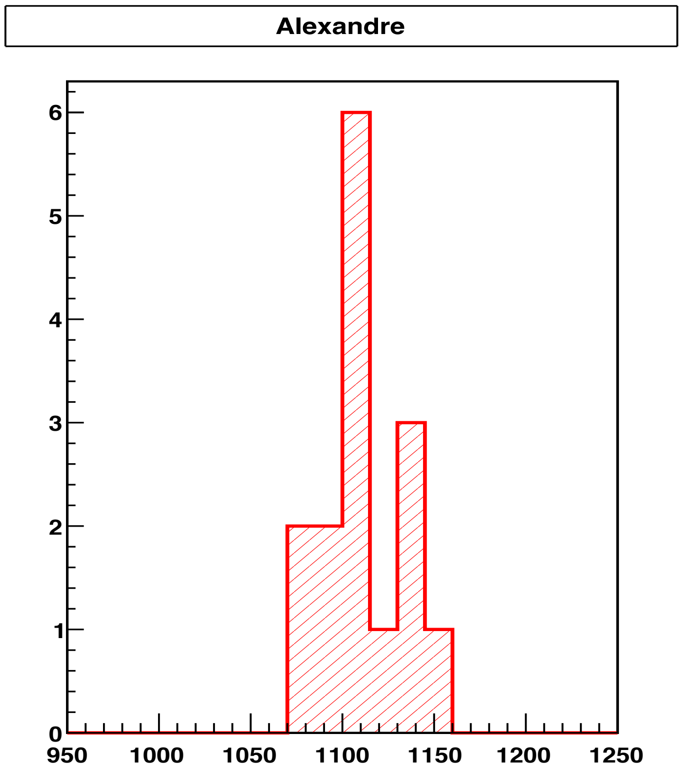
- Vamos supor que queremos comparar duas amostras diferentes e testar se elas são compatíveis. Cada amostra foi medida com um certo número de graus de liberdade, possui uma média e um desvio padrão estimado da amostra. Define-se a grandeza, de modo geral, t como sendo:

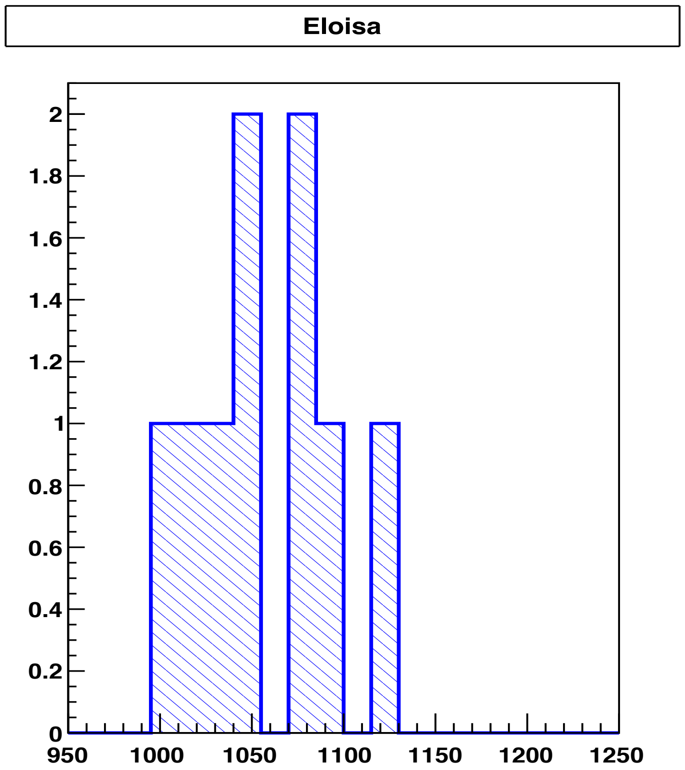
$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \qquad \begin{cases} \bar{y}_i = \text{média da amostra } i \\ \sigma_i = \text{variância da amostra } i \\ n_i = \text{ndf da amostra } i \end{cases}$$

Exemplo

Responder essa pergunta fica como exercício para a sala. Usem os histogramas abaixo como dados.

- As classes o Alexandre e Eloisa mediram a constante C e obtiveram as seguintes distribuições. Podemos dizer que os valores médios das salas do Alexandre e Eloisa 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:

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

- Que é a mesma expressão para z. Sigma não é a incerteza na média. É o desvio padrão da distribuição. Por isto dividimos por raiz de n.

Teste-t de Student

- 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:

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

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

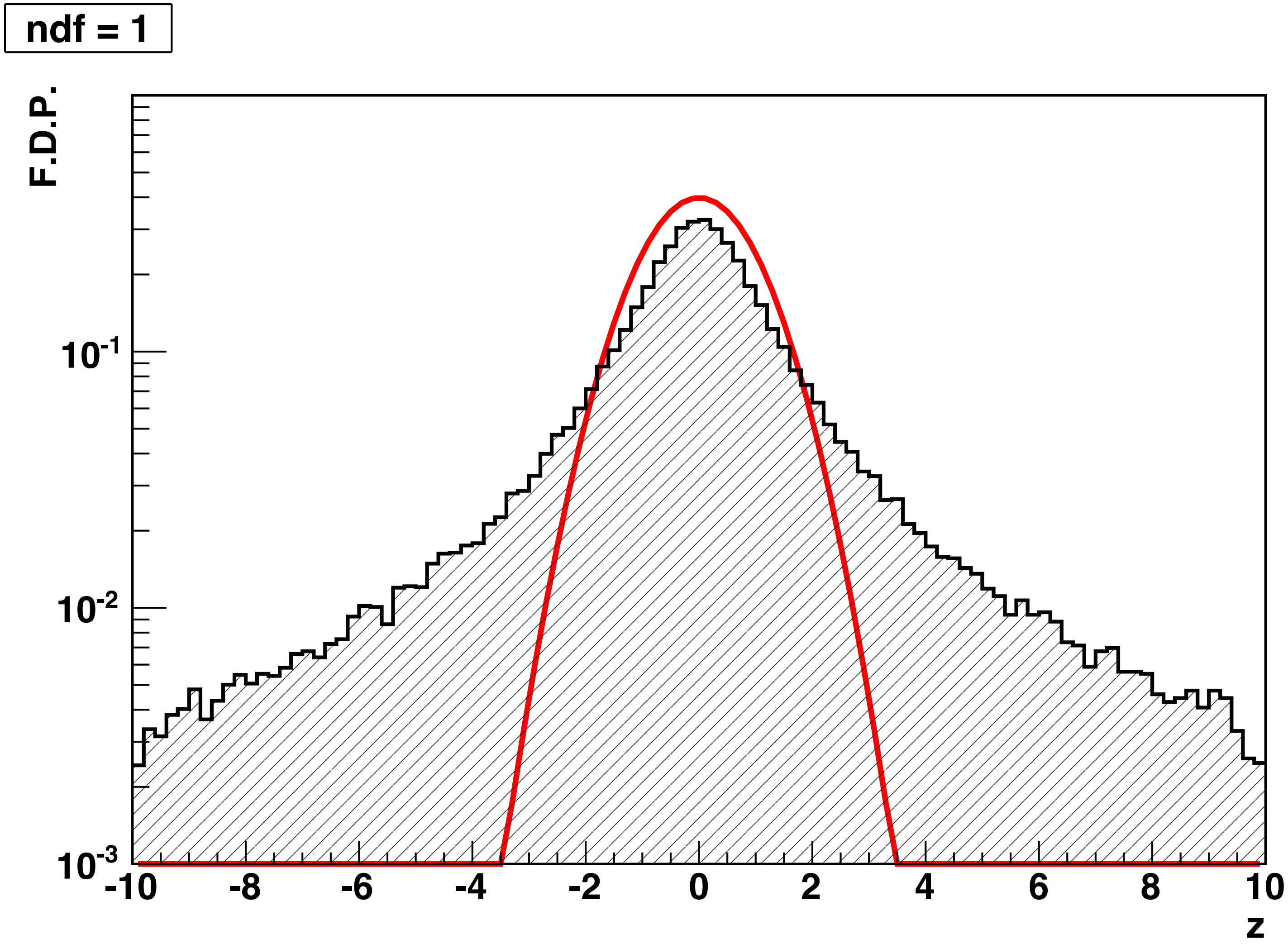
Distribuição t de Student

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

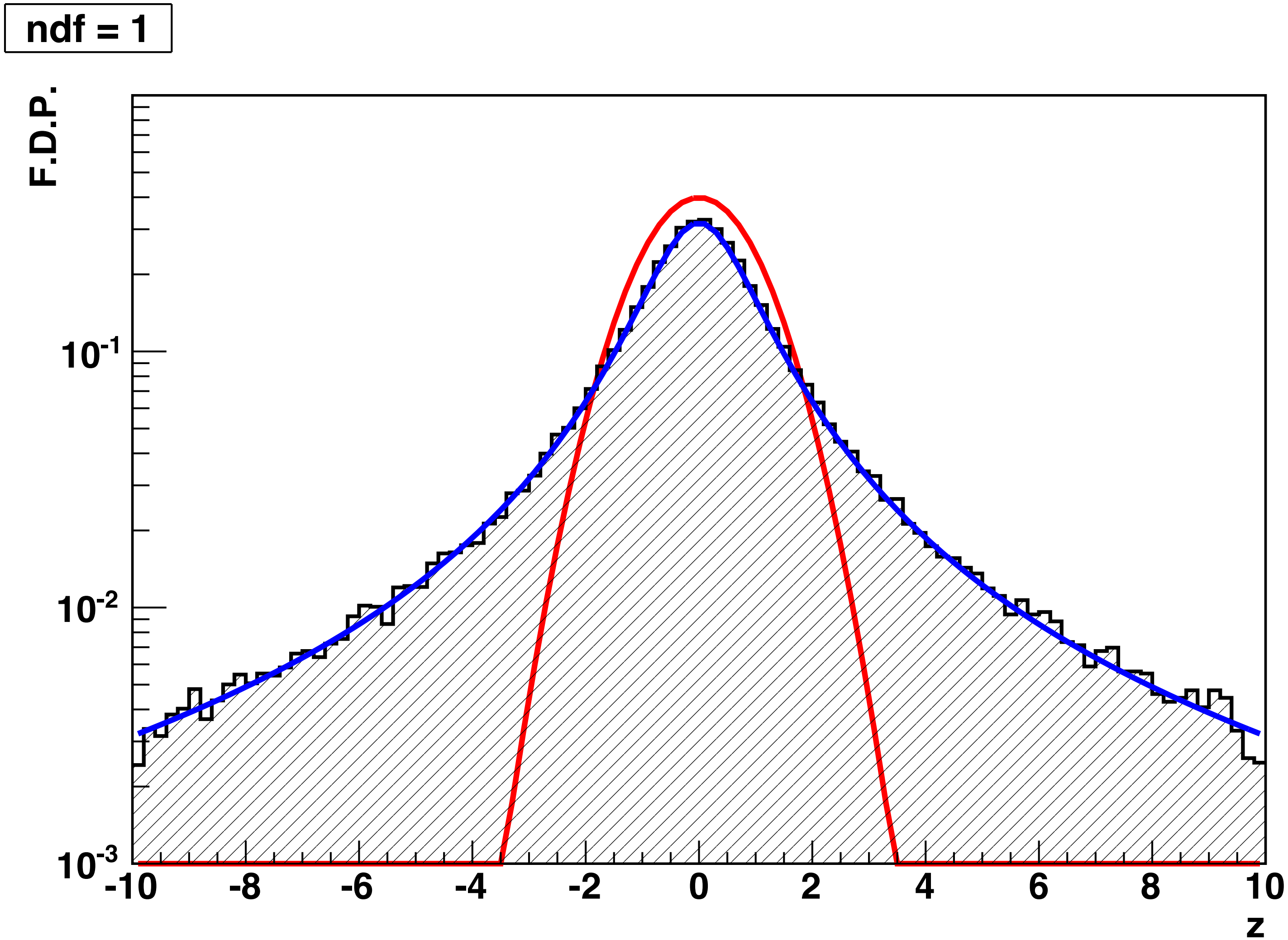
$$p(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\left(1+\frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)}$$

  - Sendo n o número de graus de liberdade da distribuição.

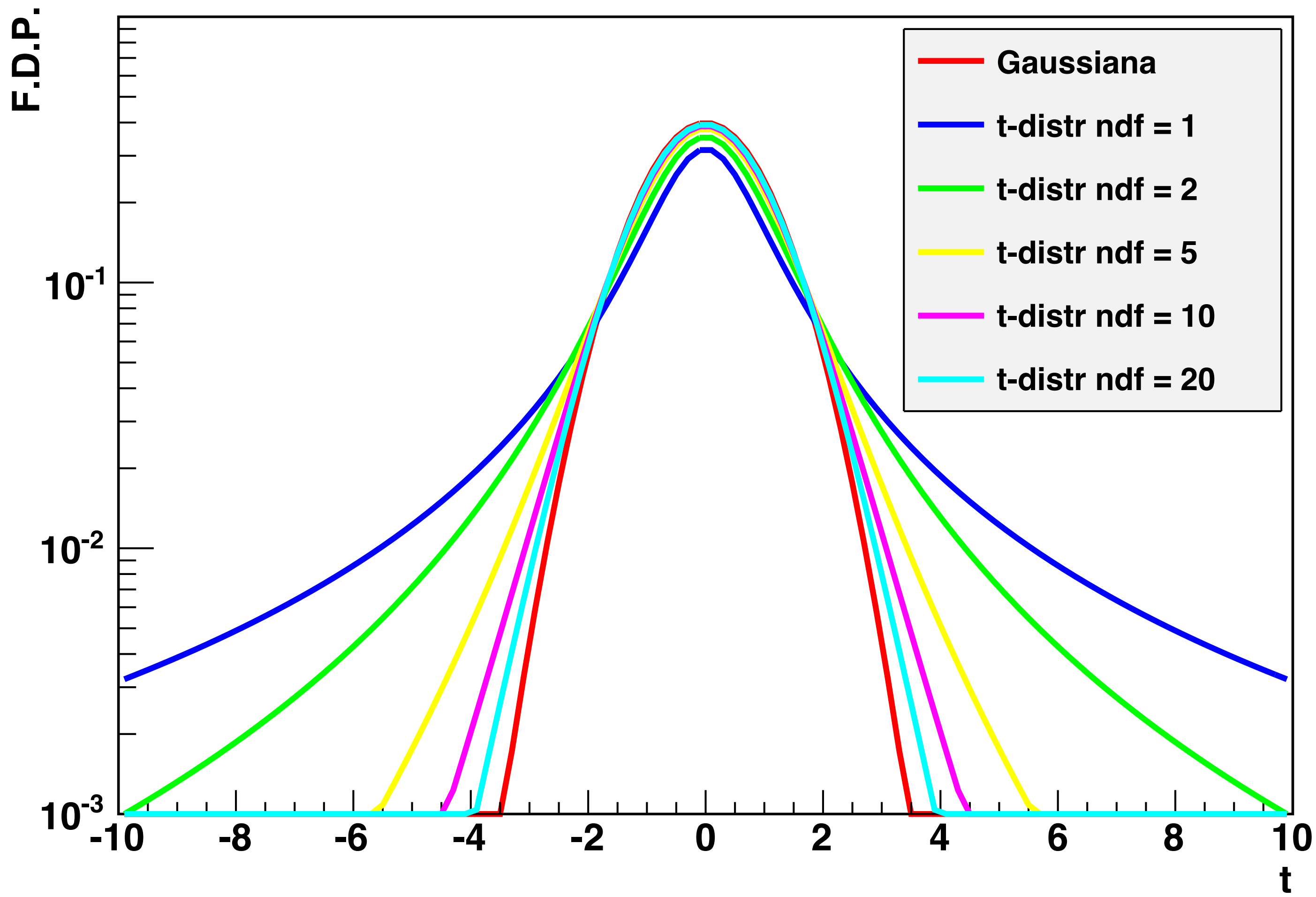
F.D.P. de t para ndf = 1



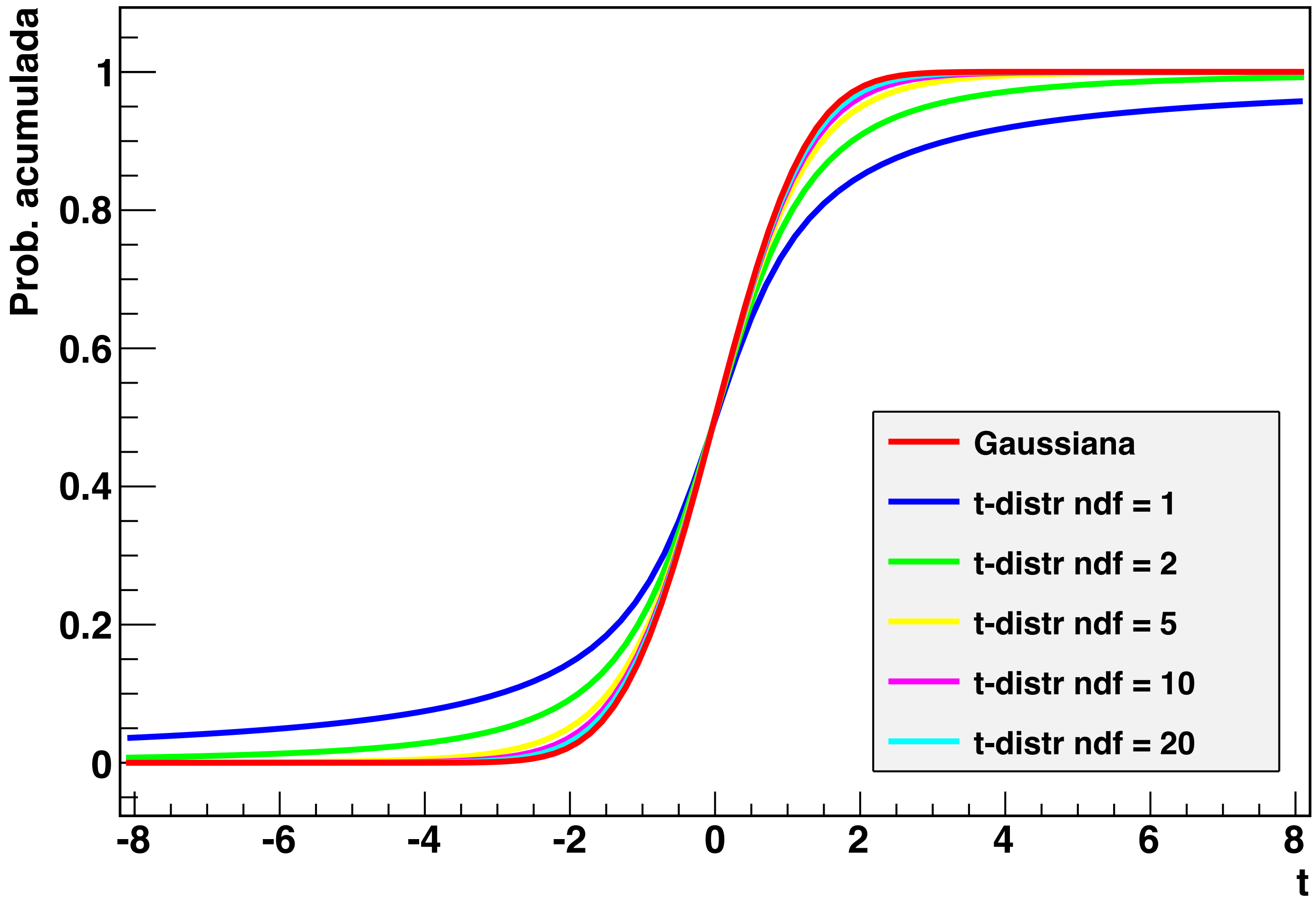
F.D.P. de *t* para ndf = 1



Distribuições-*t* para diferentes ndf

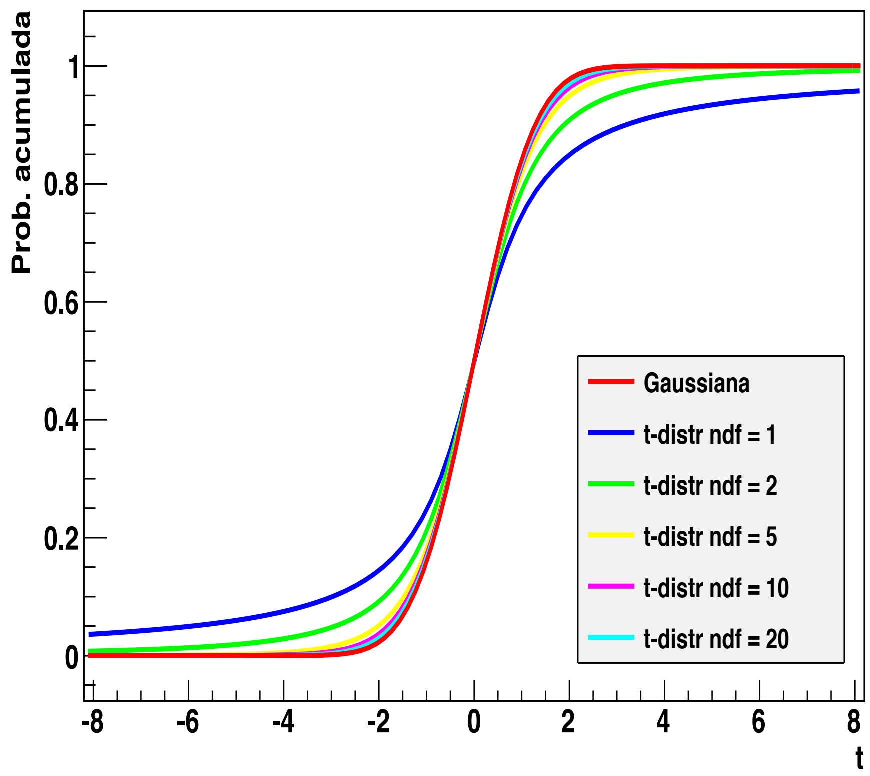


Probabilidade acumulada



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 $|t|<8$ em alguns casos para ~95% confiança!



Teste-t de Student

- Leva em conta o fato da F.D.P. 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