

# Métodos estatísticos em Física de Partículas

Vitor Oguri

Departamento de Física Nuclear e Altas Energias (DFNAE)  
Instituto de Física Armando Dias Tavares (IFADT)  
Universidade do Estado do Rio de Janeiro (UERJ)

Genève - Suisse  
06 de setembro de 2010



# Sumário

- 1 A Física de Partículas
- 2 Métodos e testes estatísticos
- 3 Simulação de eventos



## O átomo filosófico e a ciência moderna

### ruptura com os mitos - reducionismo grego (600 a.C. – 350 a.C.)

$\left\{ \begin{array}{l} \text{Tales (água)} \\ \text{Anaxímenes (ar)} \\ \text{Xenófanés (terra)} \\ \text{Heráclito (fogo)} \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \text{Demócrito (átomos e} \\ \text{Leucipo (vazio)} \end{array} \right.$

### determinismo mecanicista

$\left\{ \begin{array}{l} \text{Galileu (1564 - 1642) – matemática e experimentos} \\ \text{Descartes (1569 - 1650) – mecanicismo não empírico} \\ \text{Newton (1643 - 1727) – síntese da mecânica} \end{array} \right.$

## O século XIX

### átomo químico

{	Dalton (1808)	argumentos de simetria
	Mendeleiev (1869)	

### apogeu e declínio dos modelos clássicos mecânicos

{	Teoria Cinética dos Gases (Maxwell-Boltzmann – 1859 - 1879)	argumentos estatísticos
	Eletromagnetismo (Maxwell – 1864)	campos eletromagnéticos



## O século XX

### novos pilares e conceitos da Física

{	teorias relativísticas (Einstein – 1905 e 1915)	espaço tempo
	teorias quânticas (Heisenberg, Dirac – 1925 e 1932)	incertezas antipartículas

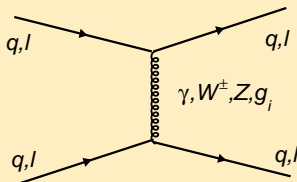


# O atomismo contemporâneo



## Modelo Padrão

- modelo dinâmico **probabilístico** das interações fundamentais (**eletro-frac**a e **forte**) entre **férmions** elementares (**quarks** e **léptons**) via os **bósons de calibre**  $\gamma$ ,  $W^\pm$ ,  $Z$  e glúons ( $g_i$ )



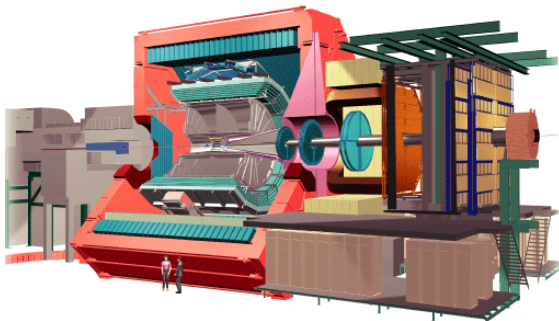
- características das interações fundamentais são ditadas por propriedades de **simetria**

# ALICE (A Large Ion Collider Experiment)

## plasma de quark-gluon

Universidade do Estado de São Paulo (USP)

Universidade Estadual de Campinas (UNICAMP)



26x16x16 (m)  
10.000 toneladas  
1000 físicos  
94 instituições  
28 países

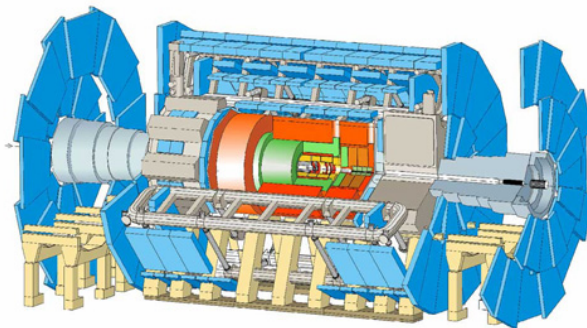




## ATLAS (A Toroidal LHC Apparatus)

**bóson de Higgs, dimensões extras e matéria escura**

Universidade Federal do Rio de Janeiro (UFRJ)



46x25x25 (m)  
7000 toneladas  
1700 físicos  
159 instituições  
37 países

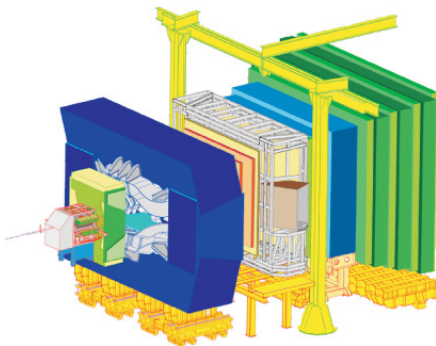


## LHCb (Large Hadron Collider beauty)

Por que existe mais matéria do que antimatéria?

Centro Brasileiro de Pesquisas Físicas (CBPF)

Universidade Federal do Rio de Janeiro (UFRJ)

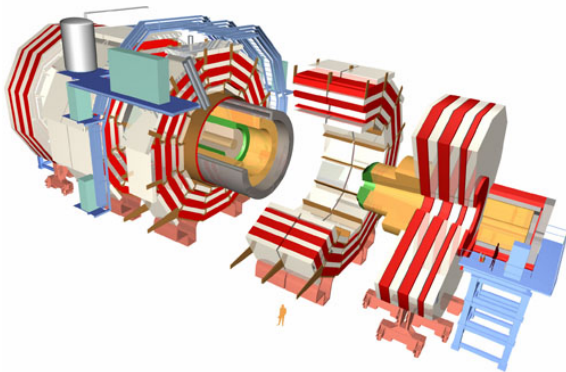


21x10x13 (m)  
5600 toneladas  
650 físicos  
48 instituições  
13 países



## CMS (Compact Muon Solenoid)

bóson de Higgs, dimensões extras e matéria escura



21x15x15 (m)  
12.500 toneladas  
2000 físicos  
155 instituições  
37 países



## Participação brasileira no CMS

- **Centro Brasileiro de Pesquisas Físicas (CBPF)**  
Moacyr Gomes e Souza, Gilvan Alves, Maria Elena Pol, Dilson de Jesus Damião, Marília Carneiro, Lucas Brito
- **Universidade do Estado do Rio de Janeiro (UERJ)**  
Alberto Santoro, Vitor Oguri, Andre Sznajder, Luiz Mundim, Wanda Prado da Silva, Carley de Oliveira Martins, Wagner de Carvalho, Helio Nogima, Eduardo Revoredo, Jose Afonso Sanches, Sandro de Souza, Jordan Martins, Ana Thereza, Diego Figueiredo, Sheila do Amaral, Eliza Melo da Costa, Luana Soares, Walter Alda
- **Universidade Estadual Paulista (UNESP)**  
Sergio Novaes, Sandra Padula, Flavia de Almeida Dias, Eduardo de Moraes Gregores, Thiago Tomei, Marco Andre Dias, Franciole Marinho



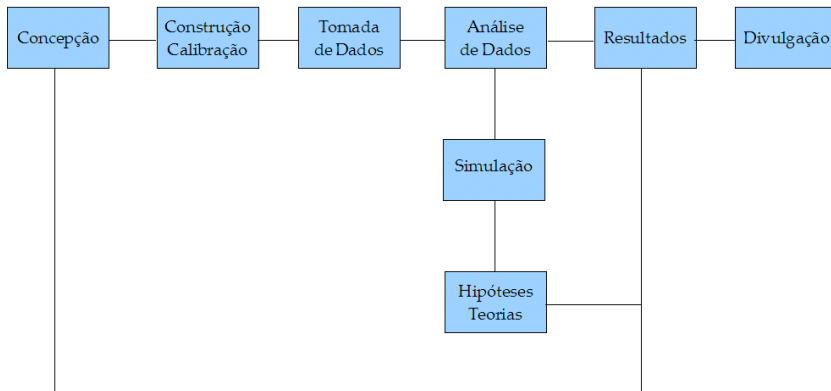
## Participação da UERJ no CMS

<http://www.hepgrid.uerj.br/>



- 150 milhões de sensores  $\rightarrow$  (filtro) 100 colisões/s
- 700 Mbytes/s  $\rightarrow$  15 milhões de Gbytes (15 Pbytes/ano)

# Experimento em Física



## The Grammar of Science – 1892

### Karl Pearson (1857-1936)

- modelos estatísticos como alternativa à visão determinística do séc. XIX
- resultados  $\{x_i\}$  de um experimento obedecem a certas distribuições,  $\rho(x)$ , que são caracterizadas por alguns parâmetros: média, desvio-padrão, simetria e curtose
- observáveis na Ciência são as distribuições de probabilidades (**pdf**) associadas aos valores (medidas) dos dados (grandezas)



# Statistical Methods for Research Workers – 1925

## Ronald Fisher (1890-1962)

- todo experimento deve começar com um modelo matemático que estime os resultados esperados (simulação de eventos)



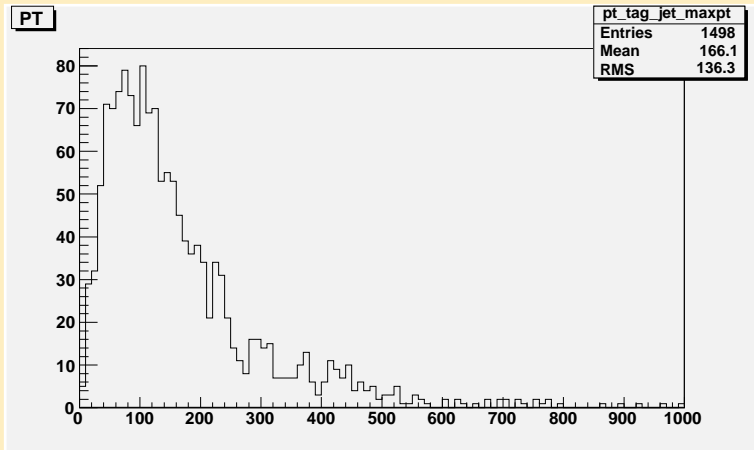


## Interpretações da Estatística

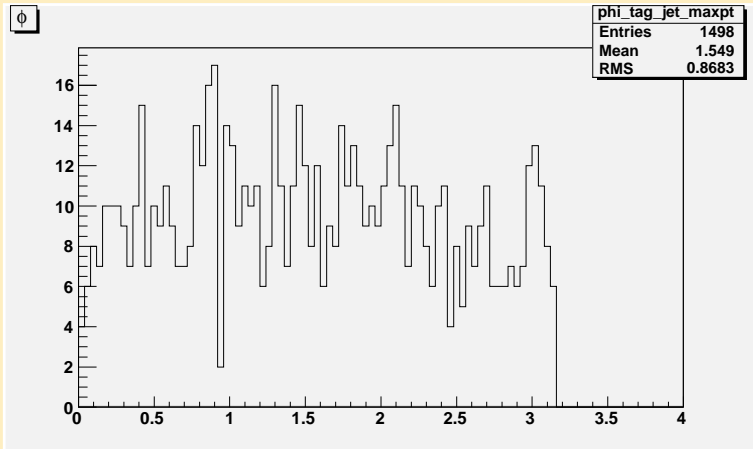
- a partir de um grande número de medições pode-se determinar os parâmetros da “verdadeira” distribuição das medidas ([Pearson](#))
- a partir de um experimento obtém-se apenas os estimadores dos parâmetros de distribuições hipotéticas dos dados ([Fisher](#))



## distribuição de momentum transverso

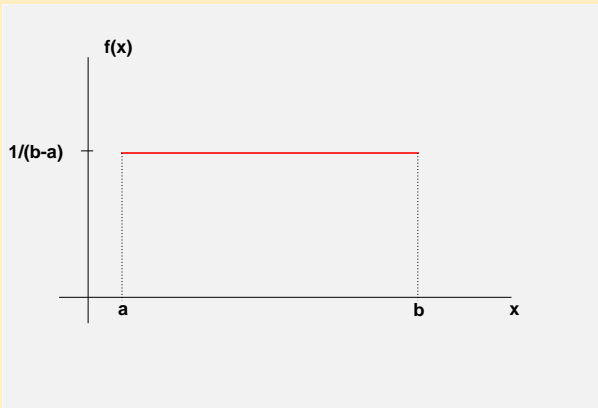


## distribuição de ângulo azimutal



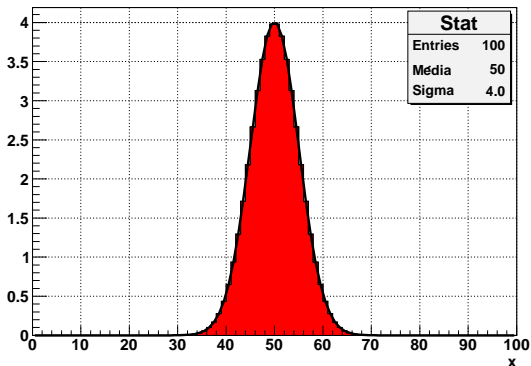
# probability density function (pdf)

## pdf uniforme

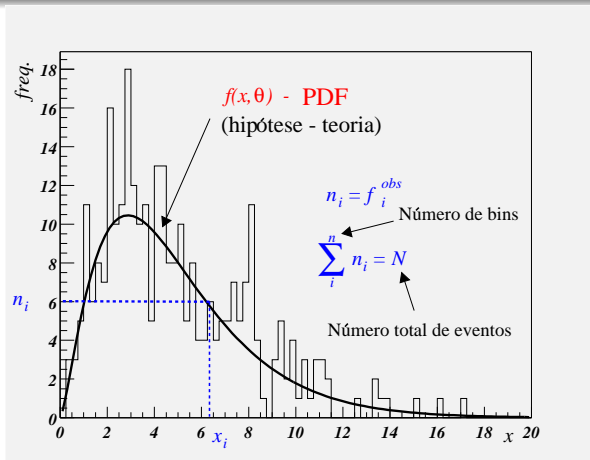


## distribuição de Gauss

$$p(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$



## teoria $\times$ experimento



- Quão bem a pdf  $f(x, \theta)$  se ajusta aos dados?



## Lançamento de dados

### processo aleatório



- evento ( $i$ )  $\Leftrightarrow$  ocorrência de uma face ou valor  $i$
- frequência ( $n_i$ )  $\Leftrightarrow$  número de ocorrências da face  $i$
- probabilidade a posteriori  $\rightarrow p_i = \frac{n_i}{N}$  (experimental)
- probabilidade a priori  $\rightarrow p_i = \frac{1}{6}$  (teórica)

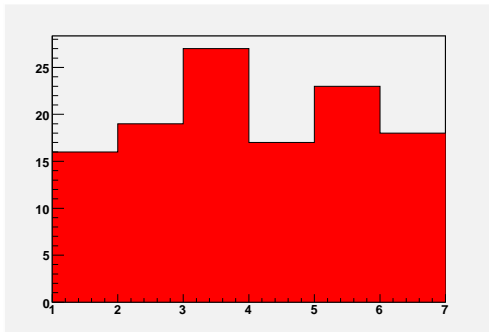


## Amostra de $N$ (120) lançamentos de um dado

$n_i$  (freq. observadas)

$\epsilon_j = Np_j$  (freq. esperadas)

$i$	$n_i$	$\epsilon_j$	$(n_i - \epsilon_j)^2$
1	16	20	16
2	19	20	1
3	27	20	49
4	17	20	9
5	23	20	9
6	18	20	4



- As diferenças são significativas?





## $\chi^2$ – medida de discrepância

- $\chi^2 = \sum_{i=1}^6 \frac{(n_i - \epsilon_i)^2}{\epsilon_i}$  (5 termos independentes)

- $\chi^2 = \frac{16}{20} + \frac{1}{20} + \frac{49}{20} + \frac{9}{20} + \frac{9}{20} + \frac{4}{20} = \frac{88}{20} = \frac{44}{10} = 4.4$

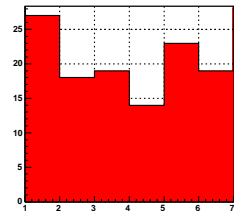
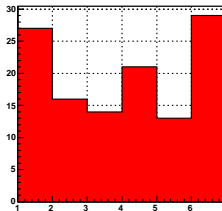
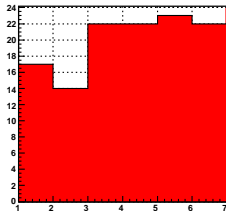
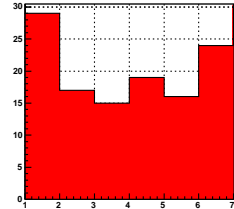
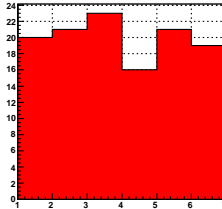
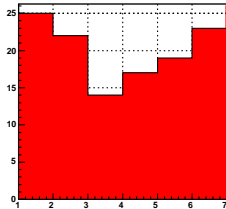
$$\sum_{i=1}^6 n_i = 120 \quad (\text{relação de vínculo})$$



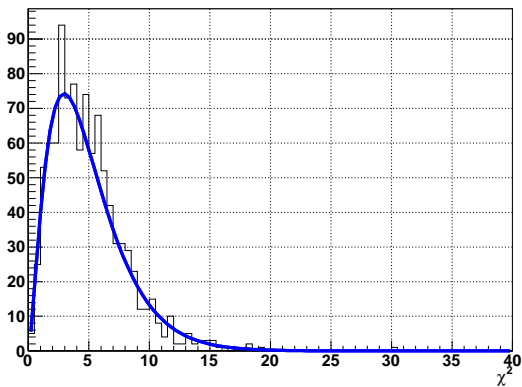
$$\nu = 6 - 1 = 5 \quad (\text{graus de liberdade})$$



## Amostras de 120 lançamentos de um dado

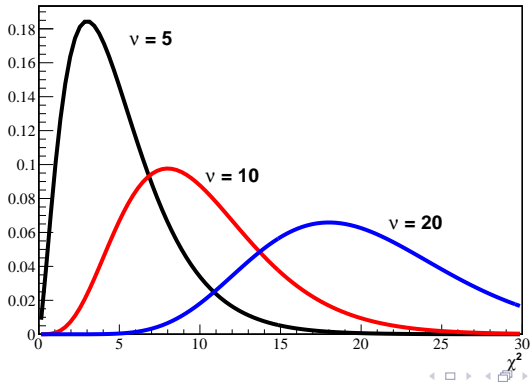


## Distribuição de $\chi^2$ para 1000 amostras de 120 lançamentos de um dado



## Teste de $\chi^2$

$$\frac{\chi^2}{\nu} = \frac{\sum \frac{(n_{obs} - n_{esp})^2}{n_{esp}}}{\nu} \simeq 1$$



# O experimento de Rutherford-Geiger

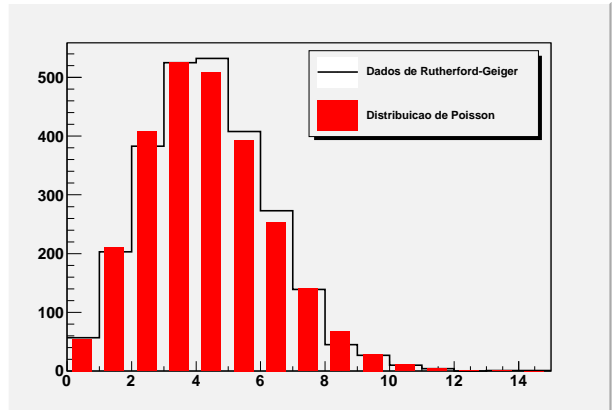
## Ernest Rutherford (1871-1937)

O exemplo clássico, que envolve uma distribuição de probabilidades (de Poisson) é o experimento de Rutherford-Geiger (1910), de contagem do número de partículas  $\alpha$  emitidas por uma amostra de polônio, em intervalos de 7.5 s (num total de 2608 intervalos).



## Distribuição dos dados de Rutherford-Geiger

$m$	$f_m$
0	57
1	203
2	383
3	525
4	532
5	408
6	273
7	139
8	45
9	27
10	10
11	4
12	0
13	1
14	1
	2608



## Resultados e questões

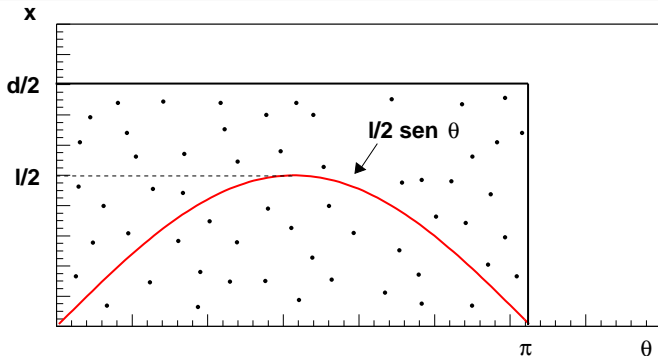
- o número médio ( $\bar{m}$ ) de contagens é da ordem de 3.87
- a distribuição das contagens pode ser comparada com uma distribuição de Poisson

$$P_m(\mu) = \frac{\mu^m}{m!} e^{-\mu}$$

de mesma média ( $\mu = \bar{m}$ )







$$\left\{ \begin{array}{l} \theta \rightarrow (0, \pi) \quad x \rightarrow (0, d/2) \\ \text{(domínio)} \\ x \leq (l/2)\text{sen}\theta \\ \text{(condição para interceptação)} \end{array} \right.$$

$$\Rightarrow \rho_{\text{teor}} = \frac{\text{área sob a curva}}{\text{área da região retangular}}$$



## Cálculo de áreas

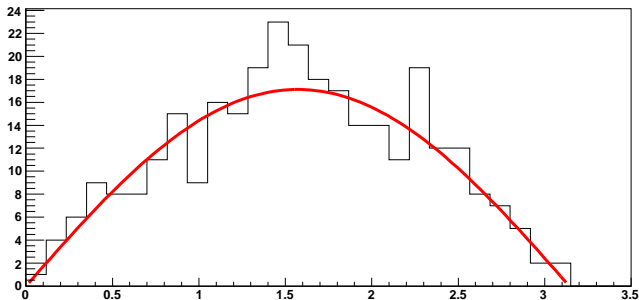
$$\rho_{\text{teor}} = \rho_{\text{exp}} \Leftrightarrow \frac{\text{área sob a curva}}{\text{área da região retangular}} = \frac{m}{N}$$



$$\text{área sob a curva} = \text{área da região retangular} \times \frac{m}{N}$$



## Geração de eventos



- distribuição de valores de  $\theta$  aceitos: senoidal



## Métodos de Monte Carlo

- métodos numéricos probabilísticos  
difusão de nêutrons (Fermi -1934)  
von Neumann, Metropolis e Ulam (1949)
- distribuições uniformes de números aleatórios



- geração de eventos
- cálculo de integrais



## Método da Rejeição

- gera-se uma seqüência  $\{r_1, r_2\}$  de pares de números aleatórios distribuídos uniformemente no intervalo  $(0,1)$

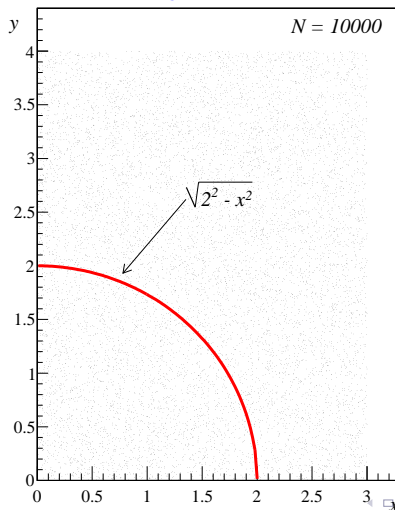
- a partir das expressões 
$$\begin{cases} x = r_1 \times (b - a) + a \\ y = r_2 \times f_{\max} \end{cases}$$

↓

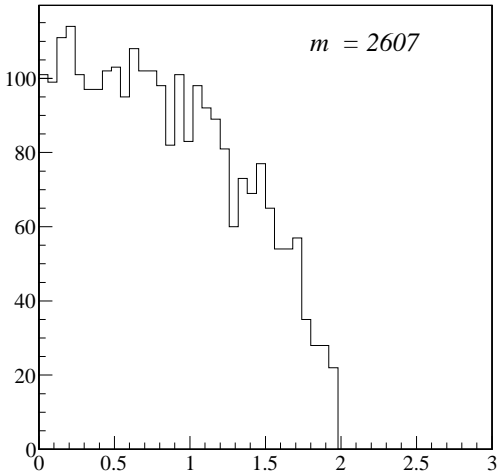
obtém-se uma outra seqüência  $\{x_i, y_i\}$  de pares uniformemente distribuídos nos intervalos  $(a, b)$  e  $(0, f_{\max})$

- os números  $x_i$  dos pares de  $\{x_i, y_i\}$  que passarem pelo critério  $y_i \leq f(x_i)$ , estarão distribuídos segundo a **pdf**  $f(x)$

## Simulação de eventos



## Simulação de eventos



## Sugestões de leitura



João Varela

*O século dos quanta*

Gradiva, Lisboa, 1996.



Deborah J. Bennett

*Aleatoriedade*

Martins Fontes, São Paulo, 2003.



Joaquim Marques de Sá

*O acaso: a vida do jogo e o jogo da vida*

Gradiva, Lisboa, 2006.

