LOM3227 - 2º semestre de 2016
Métodos Computacionais da Física

Parte 6. Modelagem em sistemas estocásticos

Distribuições de probabilidades;

Números pseudo-aleatórios;

Métodos de Monte Carlo (simulação discreta);

Caminhadas aleatórias;

Percolação;

Fractais;

Autômatos celulares;

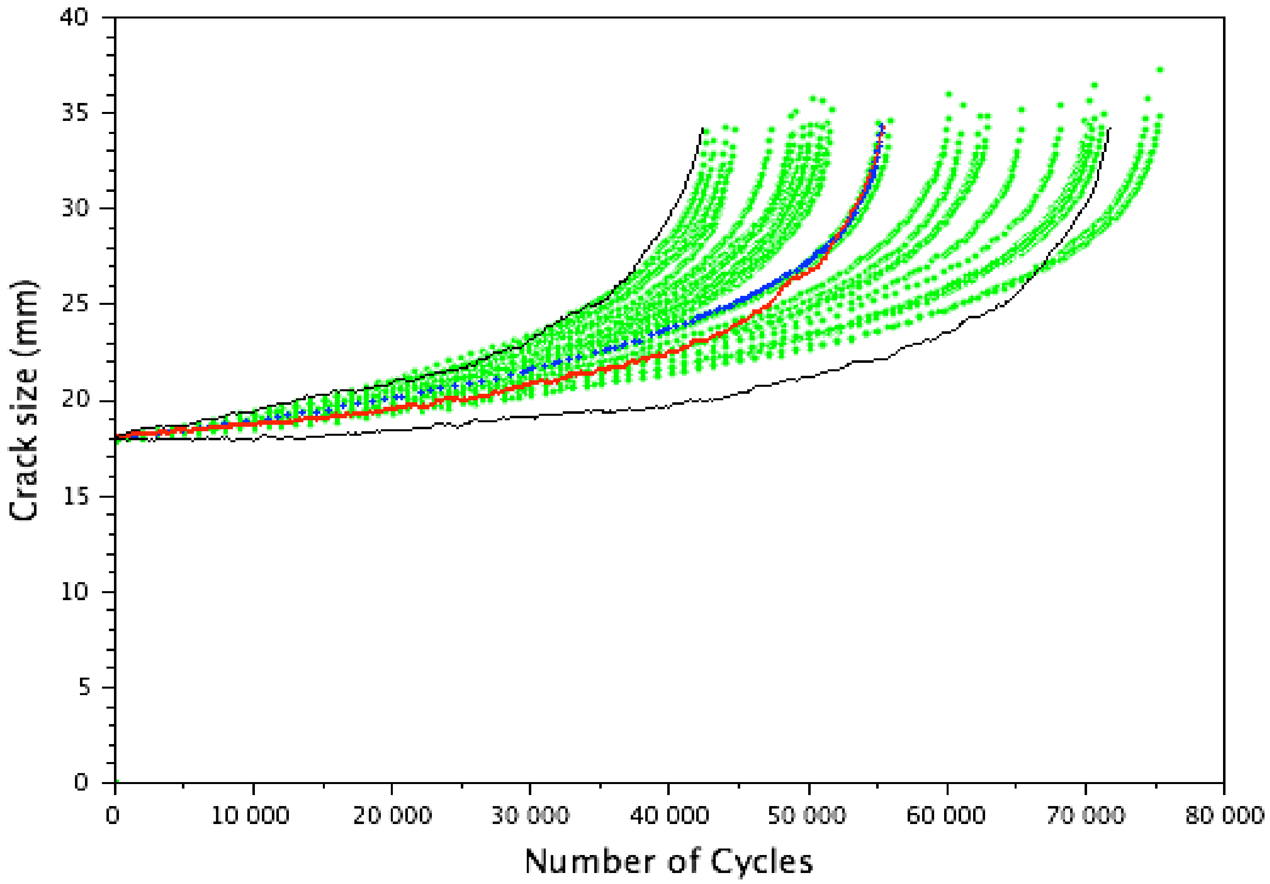
Modelagem estocástica: considerações gerais

- http://www.cin.ufpe.br/~rmcr/ESAP/arquivos/SimulacaoEstocastica.pdf (ref.6-01)

O que é Simulação Estocástica?

- Simulação:
  - ato ou efeito de simular
  - Disfarce, fingimento, ...
  - Experiência ou ensaio realizado com o auxílio de modelos.
- Aleatório: dependente de circunstâncias casuais ou fortuitas.
- Simulação estocástica é a arte de gerar amostras de variáveis aleatórias num ambiente computacional e usar as ditas amostras para a obtenção de um certo resultado.

Modelagem estocástica: exemplo de aplicação



Distribuições de probabilidades

- Contínuas:
- http://www.bertolo.pro.br/FinEst/Estatistica/DistribuicaoContinua.pdf (ref.6-12)
- http://pt.wikipedia.org/wiki/Distribui%C3%A7%C3%A3o_uniforme
- http://pt.wikipedia.org/wiki/Distribui%C3%A7%C3%A3o_normal
- Discretas:
- http://www.bertolo.pro.br/FinEst/Estatistica/DistribuicaoDiscreta.pdf (ref.6-13)
- http://cmq.esalq.usp.br/BIE5781/doku.php?id=01-discretas:01-discretas
- http://www.inf.ufsc.br/~anaclaudia/ine5108/notas_aula/texto_Bernoulli_Bin.pdf

Distribuições contínuas de probabilidades (ref.6-12)

Distribuições Contínuas

Variável aleatória contínua é aquela que pode assumir inúmeros valores num intervalo de números reais e é medida numa escala contínua. Por exemplo, uma variável aleatória contínua deve ser definida entre os números reais 0 e 1, ou números reais não negativos ou, para algumas distribuições, qualquer número real. A temperatura, a pressão, a precipitação ou qualquer elemento medido numa escala contínua é uma variável aleatória contínua.

Existem duas funções associadas a cada variável contínua X: a **função densidade de probabilidade**, simbolizada por **f(X)**, e a **função cumulativa de probabilidade**, ou **função de distribuição de probabilidade** representada por **F(X)**. A função **f(X)** é aquela cuja integral de X = a até X = b (b ≥ a) dá a probabilidade de que X assuma valores compreendidos no intervalo (a, b), ou seja,

$$P(a \le X \le b) = \int_a^b f(X)\, dX \qquad (1)$$

A função cumulativa de probabilidade F(b) é tal que:

$$F(b) = \mathrm{Prob}(X \le b) = \int_{-\infty}^{b} f(X)\, dX \qquad (2)$$

Distribuições contínuas de probabilidades (ref.6-12)

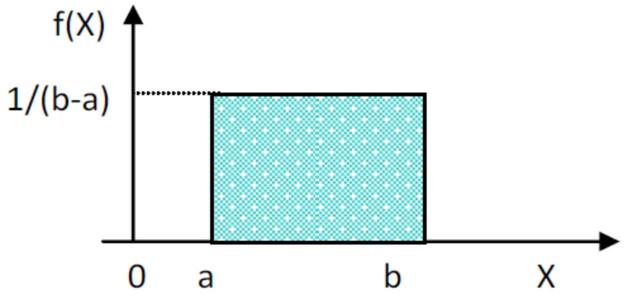
Distribuição Uniforme

Uma distribuição de variável aleatória contínua é a *distribuição uniforme* cuja função densidade de probabilidade é **constante** dentro de um intervalo de valores da variável aleatória **X.**

A variável aleatória X tem distribuição uniforme de probabilidades no intervalo (a, b) se a função densidade f(x) for:

$f(x) = \frac{1}{b-a}$, com as seguintes condições: $b \geq a$ e $a \leq x \leq b$.

A representação gráfica da distribuição uniforme é um retângulo com base definida pelos valores a e b que estabelecem os limites de valores possíveis da variável aleatória X, Figura XXXXX.



Da definição da *distribuição uniforme* deduzimos:

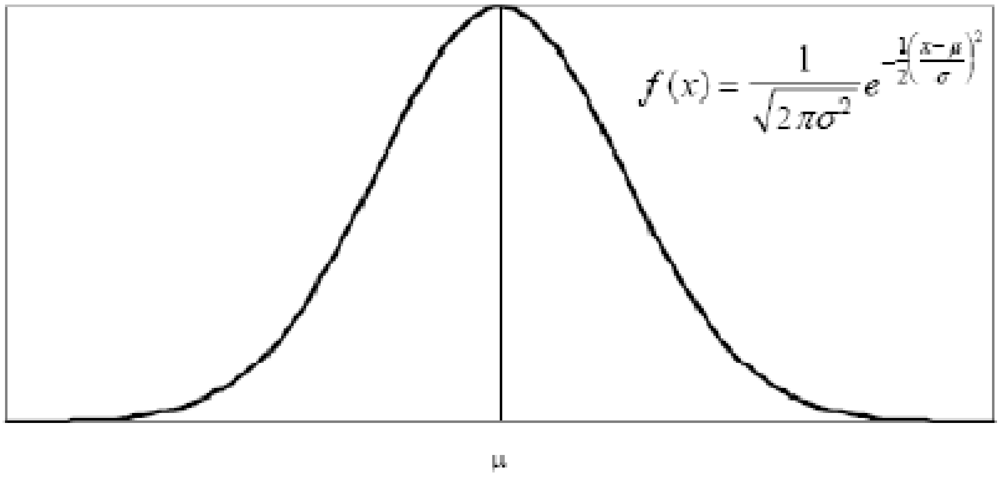
- A área do retângulo é igual a 1, pois a base é (b – a) e a altura 1/(b – a).
- A probabilidade da variável aleatória X ser igual ou maior que a e, ao mesmo tempo, menor ou igual a b é igual a 1 ou 100%

Distribuições contínuas de probabilidades (ref.6-12)

1. DISTRIBUIÇÃO NORMAL – CURVA NORMAL

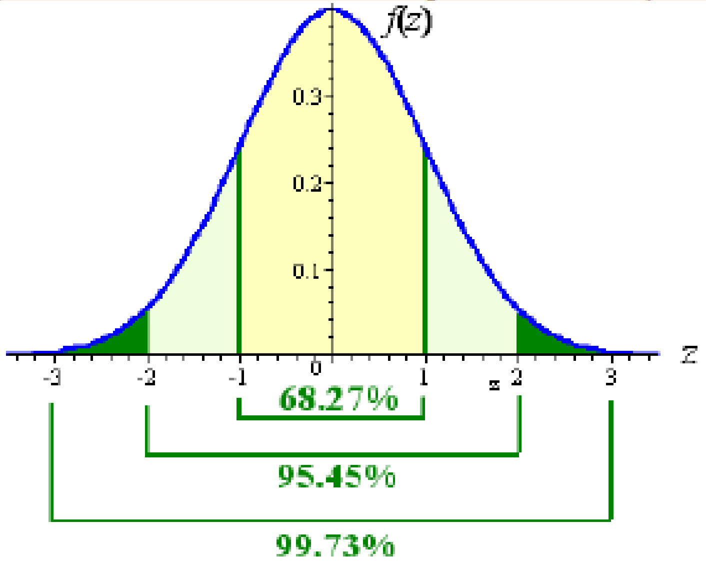
Entre as distribuições teóricas de variável aleatória contínua, uma das mais empregadas é a **distribuição normal.** Sua importância em análise matemática resulta do fato de que muitas técnicas estatísticas, como análise de variância, de regressão e alguns testes de hipótese, assumem e exigem a normalidade dos dados. Além disso, a ampla aplicação dessa distribuição vem em parte devido ao teorema do limite central. Este teorema declara que na medida em que o tamanho da amostra aumenta, a distribuição amostral das médias amostrais tende para uma distribuição normal (Triola, 1998).

O aspecto gráfico de uma **distribuição normal** é o da Figura 01:



Porcentagens da Area Sob a Curva Normal Padrão

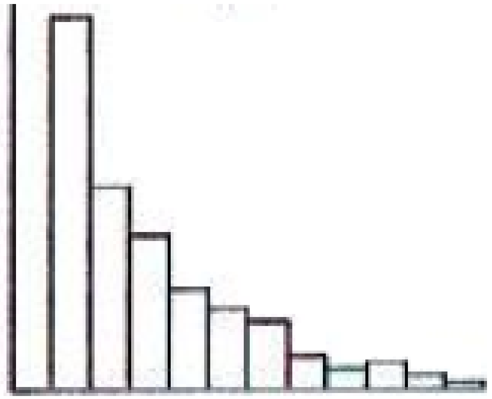
Um gráfico desta curva normal padronizada (média 0 e Variância 1) é:



Distribuições contínuas de probabilidades (ref.6-12)

Distribuição Exponencial

A distribuição exponencial é geralmente aplicada à dados com forte assimetria[3] como aqueles cujo histograma tem a forma da figura abaixo, ou seja, de J invertido. Quando os serviços prestados por uma empresa para clientes externos ou internos são de duração variável é esta distribuição a indicada para analisar esses experimentos, por exemplo, a duração do atendimento do caixa de um banco ou de postos de saúde, o tempo de operação sem interrupção de um equipamento, etc. Sua **densidade de probabilidade** tem a forma:



$$f(x) = \lambda e^{-\lambda X} \quad \text{com } \lambda > 0, x \geq 0 \tag{1}$$

e sua **função de distribuição de probabilidade** é do tipo:

$$F(x) = \int_0^{\infty} \lambda e^{-\lambda X} = 1 - e^{-\lambda X} \tag{2}$$

Distribuições contínuas de probabilidades (ref.6-12)

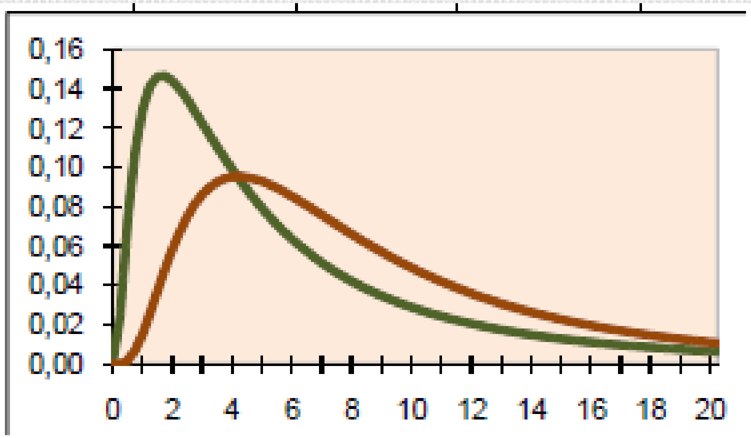
Distribuição Log-Normal

Nem todas as variáveis aleatórias têm distribuição normal. Há experiências com resultados **não simétricos**, por exemplo, o retorno das operações financeiras.

A variável aleatória X com valores positivos tem *distribuição log-normal* com função densidade de probabilidade:

$$\begin{cases} f(x) = \dfrac{1}{x\sigma_Y\sqrt{2\pi}} e^{-\frac{1}{2}x\frac{(\ln x - \mu_Y)^2}{\sigma_Y^2}} & para\ x \geq 0 \\ f(x) = 0 & para\ x < 0 \end{cases}$$

se a variável aleatória Y definida como $Y = ln(X)$ tiver distribuição normal com *média* $-\infty < \mu_Y < +\infty$ e *desvio padrão* $0 \leq \sigma_Y < \infty$.



Distribuições contínuas de probabilidades (ref.6-12)

Distribuição de Weibull

A distribuição de probabilidade Weibull é uma distribuição de probabilidade contínua amplamente utilizada na análise de dados de vida de equipamentos devido a sua flexibilidade – ela pode imitar outras distribuições de probabilidade, como a distribuição exponencial e a distribuição normal, dependendo do valor de seus parâmetros.

O seu nome se deve ao seu inventor, Waloddi Weibull, e é usada extensivamente em engenharia de confiabilidade e no cálculo do tempo médio de falha para determinado dispositivo.

As principais vantagens da utilização da distribuição de Weibull para análise da sobrevivência é que através da estimativa de apenas dois parâmetros (alfa e beta) são obtidas informações tanto de longevidade média quanto do tipo de curva de sobrevivência. Outra vantagem é que as observações não necessitam ser realizadas a intervalos constantes, como, por exemplo, com as tabelas de esperança de vida.

A fdp da distribuição Weibull é descrita pela Equação:

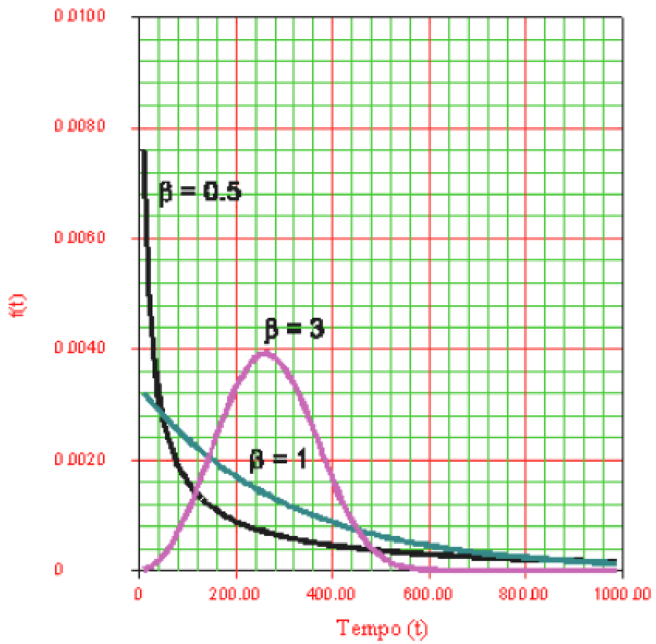
$$f(x;\beta;\alpha) = \frac{\beta}{\alpha^{\beta}} x^{\beta-1}\, e^{-\left(\frac{x}{\alpha}\right)^{\beta}} \qquad \beta>0 \text{ e } \alpha>0$$

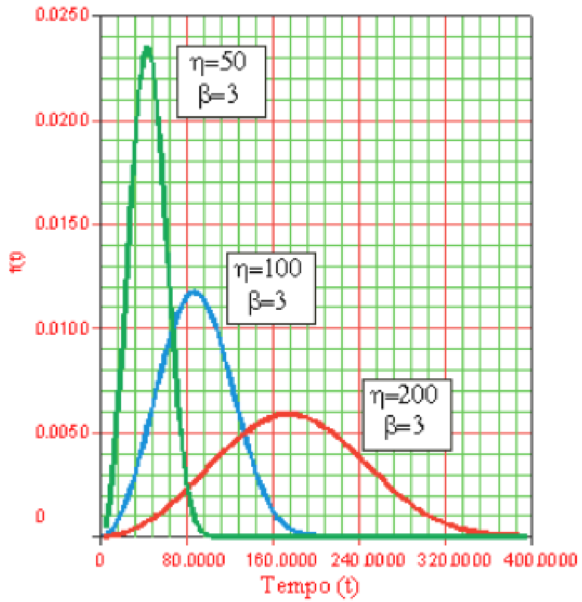
Onde:

- β é o parâmetro de forma (shape);
- α é o parâmetro de escala (scale);

A fda[9] da distribuição Weibull é descrita pela Equação:

$$F(x;\beta;\alpha) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}}$$





Distribuições discretas de probabilidades (ref. 6-13)

Distribuição de Bernoulli

Característica do modelo

Se uma variável aleatória X **só pode** assumir os valores **0** (fracasso) e **1** (sucesso) com $P(X = 0) = q$ e $P(X = 1) = p$ com $p + q = 1$, então diremos que a variável aleatória X admite distribuição de Bernoulli.

Discrição do modelo

1. $X = \{0,1\}$
2. $P(X = 0) = q$ e $P(X = 1) = p$;
3. $E(X) = p$;
4. $\sigma^2 = Var(X) = p \times q$ e $\sigma = Dp(X) = \sqrt{p \, x \, q}$

Podemos escrever o modelo do seguinte modo:

$$P(X = x) = p^x \, . \, q^{1-x}$$

onde $q = 1 - p$.

- Esperança (média) e Variância:

Distribuições discretas de probabilidades (ref. 2-13)

Distribuição Binomial

1. CONCEITUAÇÃO

Vamos, neste item, considerar experimentos que satisfaçam as seguintes condições:

a. O experimento deve ser repetido, nas mesmas condições, um número finito de vezes (**n**).

b. As provas repetidas devem ser independentes, isto é, o resultado de uma não deve afetar os resultados das sucessivas.

c. Em cada prova deve aparecer um dos dois possíveis resultados: **sucesso** e **insucesso**.

d. No decorrer do experimento, a probabilidade **p** do sucesso e a probabilidade **q** (q = 1 – p) do insucesso manter-se-ão constantes.

Resolveremos problemas do tipo: determinar a probabilidade de se obterem **k** *sucessos* em **n** tentativas.

O experimento "*obtenção de caras em cinco lançamentos sucessivos e independentes de uma moeda*" satisfaz essas condições.

Sabemos que, quando da realização de um experimento qualquer em uma única tentativa, se a probabilidade de realização de um evento (sucesso) é p, a probabilidade de não-realização desse mesmo evento (insucesso) é **1 – p = q**.

Suponhamos, agora, que realizemos a mesma prova n vezes sucessivas e independentes. A probabilidade de que um evento se realize **k** vezes nas provas é dada pela função:

$$f(x) = P(X = k) = \binom{n}{k} p^k q^{n-k}$$

Números pseudo-aleatórios - exemplos EXCEL

- Funções básicas: "=ALEATÓRIO()" ; "=ALEATÓRIOENTRE(G$5;I$5)"
- Modelagem de algumas distribuições: "=ALEATÓRIO()+ALEATÓRIO()“ etc.
- Funções avançadas: "=DIST.NORM.N(B51;B$49;C$49;FALSO)“ etc.

Métodos de Monte Carlo (simulação discreta)

- \- http://www.palisade-br.com/risk/monte_carlo_simulation.asp

A simulação de Monte Carlo efetua análise de risco por meio da construção de modelos de possíveis resultados, substituindo com um intervalo de valores – uma distribuição de probabilidade – todo fator com incerteza inerente. Em seguida, ela calcula os resultados repetidamente, cada vez com outro conjunto de valores aleatórios gerados por funções de probabilidades. Dependendo do número de incertezas e dos intervalos especificados para elas, uma simulação de Monte Carlo pode ter milhares ou dezenas de milhares de recálculos antes de terminar. A simulação de Monte Carlo produz distribuições de valores dos resultados possíveis.

Métodos de Monte Carlo (simulação discreta)

http://rassis.com/artigos/Resumo%20Simulacao.pdf (ref.6-20)

- Fase 1 – Definimos a função de probabilidade acumulada P(x), da variável aleatória x, a qual pode ser uma distribuição teórica (Uniforme, Triangular, Normal, Beta, Weibull, etc.) ou uma distribuição empírica qualquer.
- Fase 2 – Escolhemos um número aleatório equiprovável entre 0 e 1 numa tabela de números aleatórios (ou usando a função *RAND*() no EXCEL). Representamos este número y_p no eixo das ordenadas da função P(x);
- Fase 3 – Projetamos y_p horizontalmente até à curva P(x), definindo-se o ponto P. Projetamos este ponto, por sua vez, sobre o eixo das abcissas, definindo-se o valor x_p de uma amostra;
- Fase 4 – Repetimos o procedimento e obtemos uma amostra.

Métodos de Monte Carlo (simulação discreta)

http://www.bertolo.pro.br/FinEst/Estatistica/Introducao%20a%20simulacao%20de%20Monte%20Carlo.pdf (ref. 6-21)

- Exemplos de modelagem com implementação no EXCEL (ref. 6-22)

http://www.comp.ita.br/~gian/teep37/cap_2-como_funciona_a_simulacao_2.signed.pdf

- Exemplos de modelagem

http://www.mpsantos.com.br/simul/arquivos/simul.pdf (ref. 6-23)

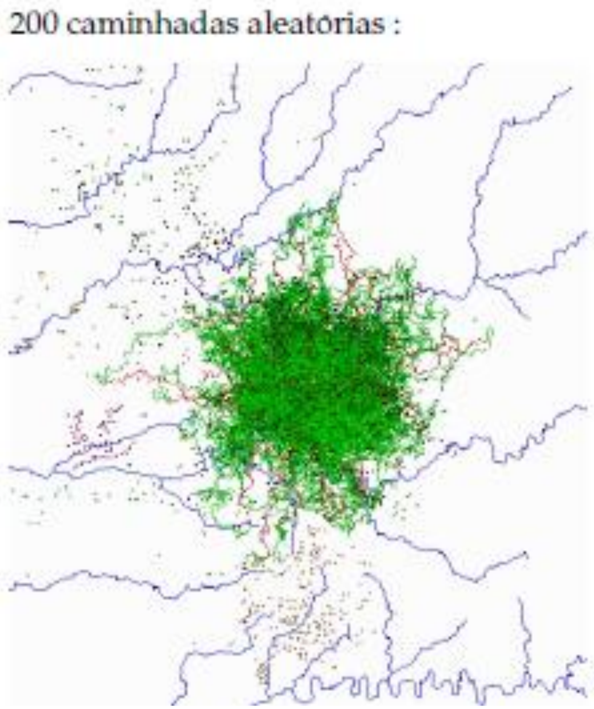
- Livro: números aleatórios – distribuições - simulação

http://www.dad.uem.br/especs/mba15/download/MBA15-PO%205.pdf (ref. 6-24)

- MBA - Pesquisa operacional - simulação

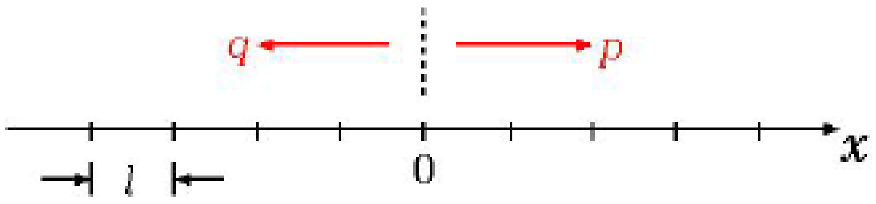
Caminhadas aleatórias

- http://pt.wikipedia.org/wiki/Passeio_aleat%C3%B3rio
- http://minerva.ufpel.edu.br/~diehl/class/fisica_estatistica/rand_lec1.pdf (ref.6-31) - Física Estatística, UFPel



Caminhadas aleatórias (ref.6-31)

Características dos deslocamentos :



- N passos sucessivos.
- Independência estatística.
- Mesmo comprimento l.
- Probabilidade p para a direita
- Probabilidade q para a esquerda.
- n_1 passos para a direita.
- n_2 passos para a esquerda.

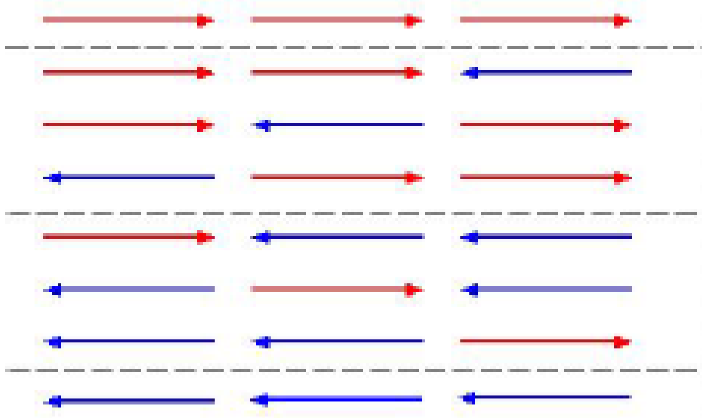
$$p + q = 1 \, ,$$

e

$$N = n_1 + n_2 \, .$$

Caminhadas aleatórias (ref.6-31)

Caminhante com $N = 3$ passos :



n_1	n_2	m
3	0	3
2	1	1
2	1	1
2	1	1
1	2	-1
1	2	-1
1	2	-1
0	3	-3

Deslocamento líquido,

$$m = n_1 - n_2 \,,$$

com

$$-N \leqslant m \leqslant N \,,$$

Neste caso o espaço amostral tem 8 elementos.

Caminhadas aleatórias (ref.6-31)

Probabilidade de uma determinada seqüência de N passos :

$$P_N(n_1, n_2) = \underbrace{p\,p\ldots p}_{n_1 \text{ fatores}} \;\; \underbrace{q\,q\ldots q}_{n_2 \text{ fatores}} = p^{n_1}\, q^{n_2} \; .$$

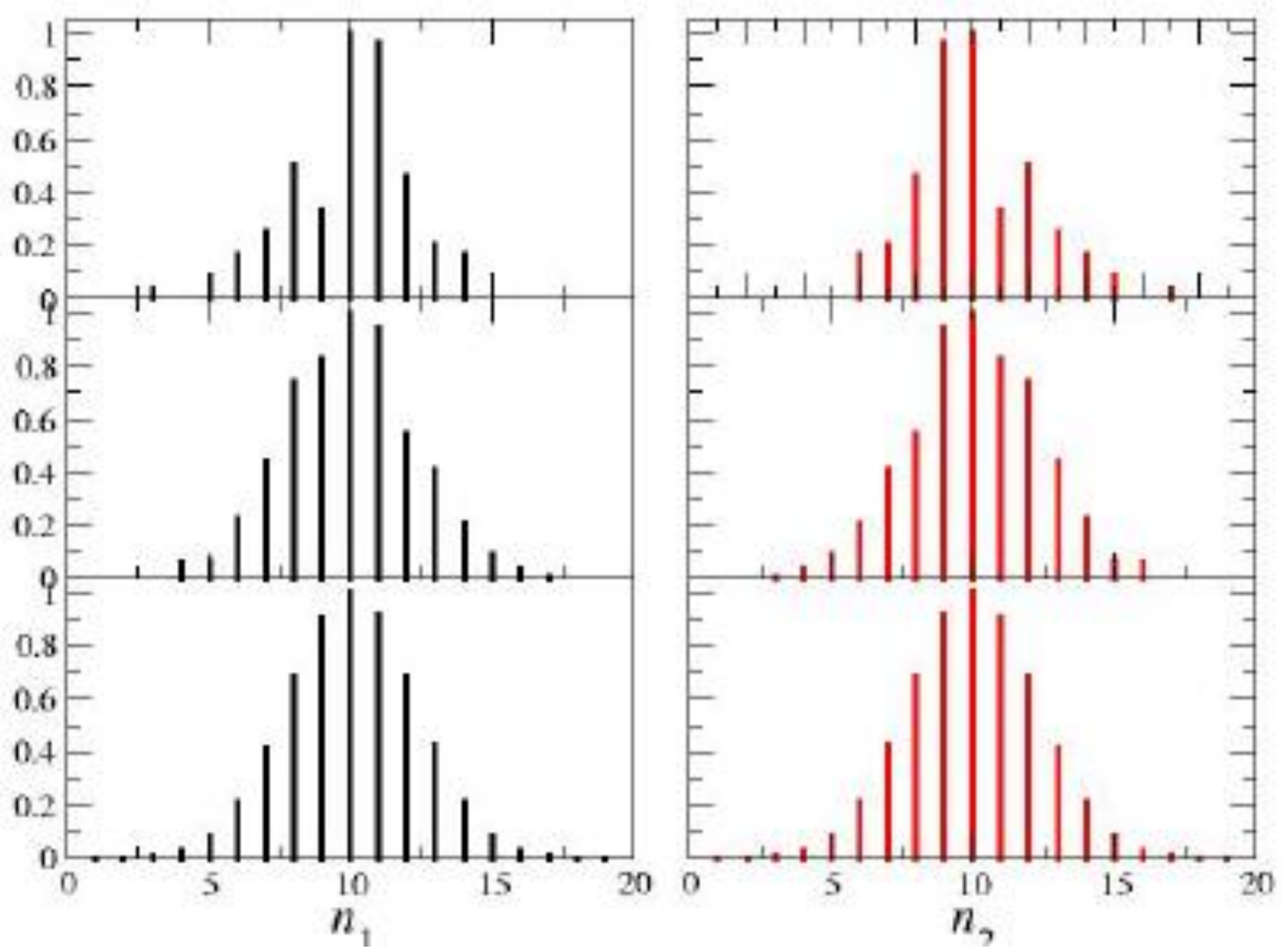
Distribuição binomial

$$W_N(n_1) = \frac{N!}{n_1!\, n_2!}\, p^{n_1}\, q^{n_2} \; ,$$

Expansão binomial (distribuição de Bernoulli),

$$(q+p)^n = \sum_{x=0}^{n} \binom{n}{x} p^x\, q^{n-x} = \sum_{x=0}^{n} \frac{n!}{x!\,(n-x)!}\, p^x\, q^{n-x} \; ,$$

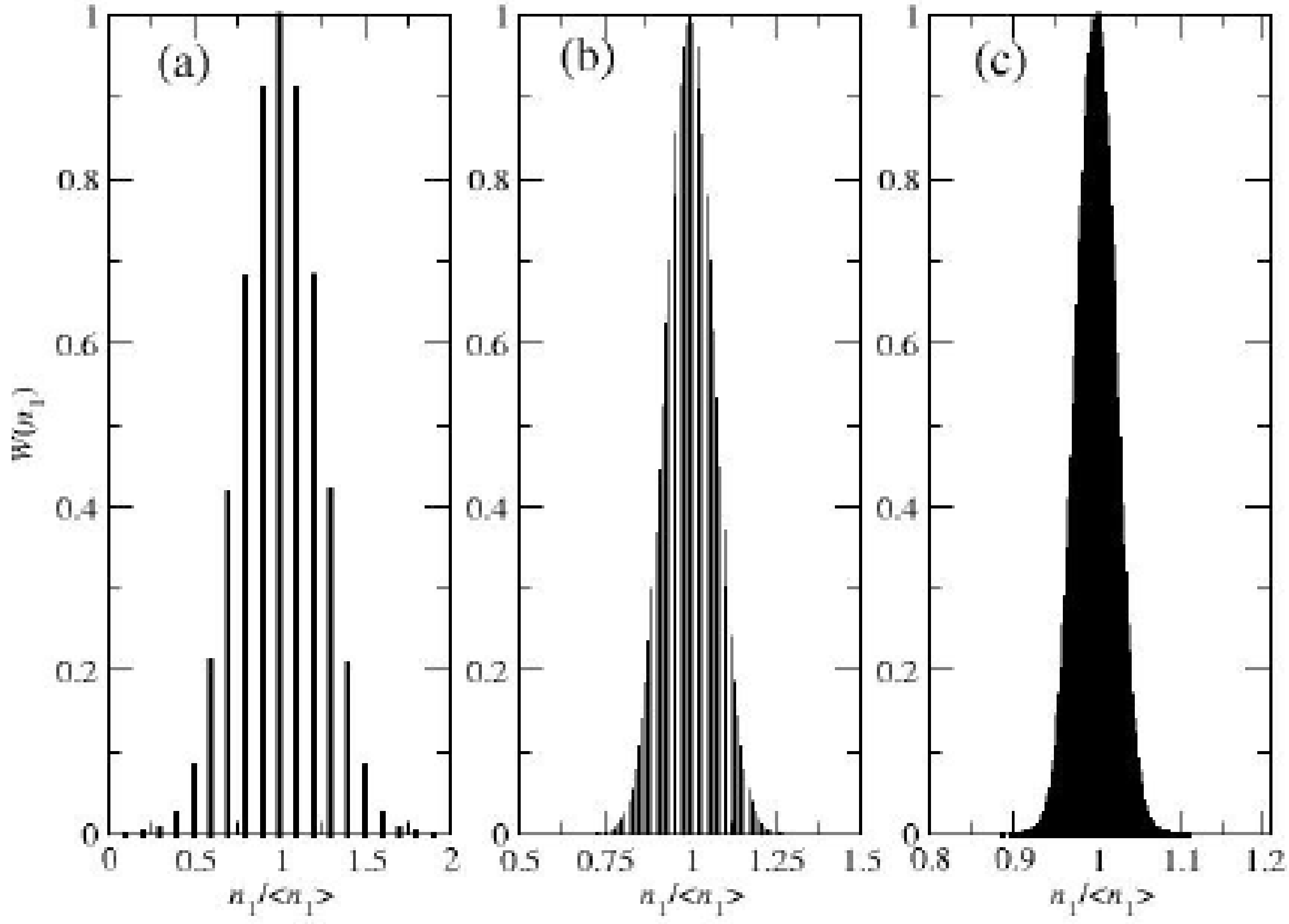
Caminhadas aleatórias (ref.6-31)



- caminhante aleatório com $N = 20$ passos
- $p = q = 1/2$
- experimentos com 100, 1000 e 1000000 repetições

"Boca de sino" invertida!

Caminhadas aleatórias (ref.6-31)



- (a) $N = 20$
- (b) $N = 200$
- (c) $N = 2000$

Percolação

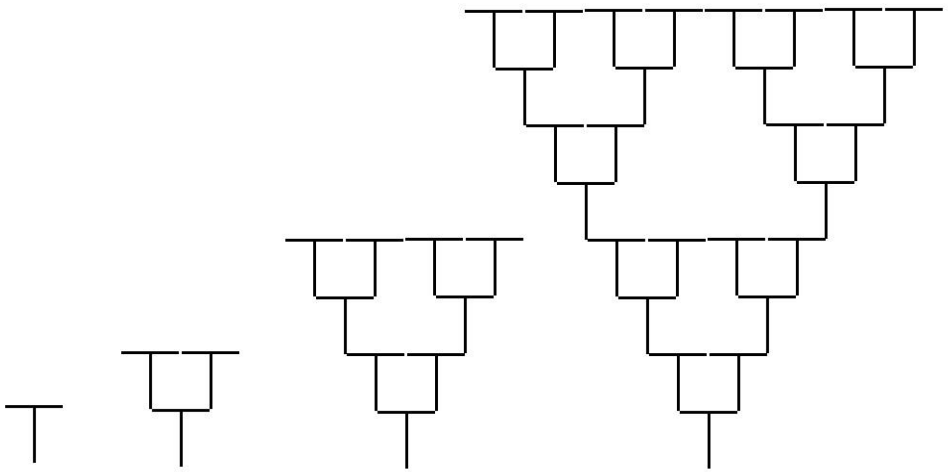
- http://pt.wikipedia.org/wiki/Percola%C3%A7%C3%A3o

- Percolação (do Lat. percōlāre, filtrar) em farmacologia e ciência dos materiais, se refere a extração de componentes solúveis passando solventes por materiais porosos. Na geologia se refere a passagem de água pelo solo e pedras permeáveis fluindo para reservatórios subterrâneos. Durante as últimas cinco décadas, o desenvolvimento de modelos matemáticos para percolação tem expandido sua aplicação também para geotecnia e redes complexas.

- http://rmu.sbm.org.br/Conteudo/n28/n28_Artigo01.pdf (ref.6-41)

- Modelos para meio poroso, neste caso o movimento em si não é aleatório, mas a distribuição das características locais do material é aleatória.

Fractais



http://hypescience.com/fractais-o-que-sao/

- Com o advento da computação gráfica, a geometria fractal escapou do campo da matemática pura e ganhou ares de concepção artística e vedete da tecnologia de ponta. Hoje é aplicada nas mais diversas áreas do conhecimento humano
- **Autossimilaridade (também denominada egossimilaridade):** existe um padrão que se repete tanto na parte quanto no todo. Nesse caso o padrão é a letra T.
- **Recursividade ou iteratividade:** é a própria repetição do padrão em si.
- **Holismo (ou sinergia):** o todo é superior à soma das partes. A partir de figuras de uma dimensão (duas retas) se constrói uma figura (quase) bidimensional. É evidente que quanto maior o número de repetições do padrão (iteração) mais próximo de 2 chegará o valor do número de dimensões topológicas dessa figura.
- **Amplificação:** uma figura fractal poderá sempre ser “ampliada” ou “amplificada” se aumentarmos o número de repetições (iterações) — daí a necessidade da utilização da computação para a construção de modelos mais aproximados dos fractais.

Fractais

- http://www.geocities.ws/projeto_caos_ufg/fractais/fractais4.html

Existem muitos métodos de geração de estruturas fractais através de modelos matemáticos simples. ...destaque a dois deles, pela sua importância e caráter bastante genérico: a geração de fractais por agregação e o método IFS *(Iterated Function System). Outro importante método de geração algorítmica de fractais provém dos* ***autômatos celulares***

Figura construída pelo método

de junção de agregados ->

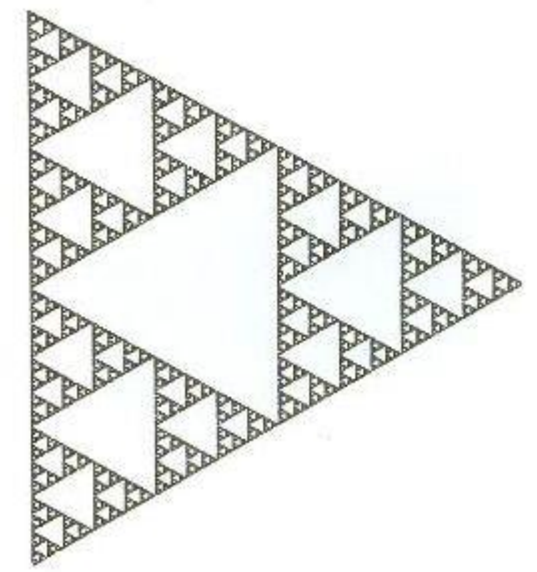
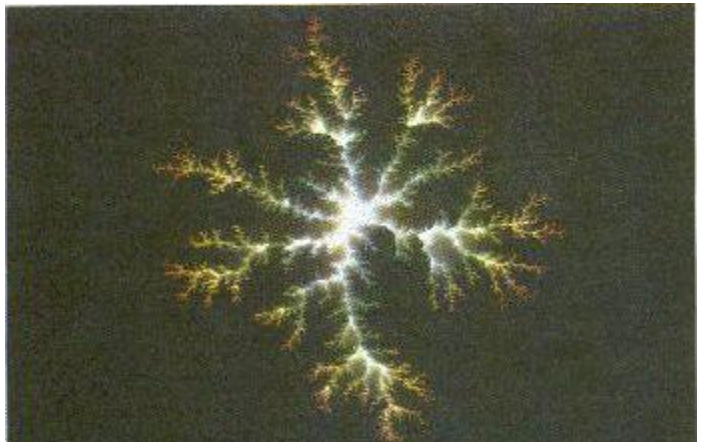


Figura construída pelo método DLA ->

(Diffusion-Limited Aggregation)



Triângulo de Sierpinsky (IFS)

Autômatos celulares

- http://pt.wikipedia.org/wiki/Aut%C3%B3mato_celular
- http://www.dcc.fc.up.pt/~sssousa/trabs/monografia/aut
- http://www.ime.usp.br/~slago/sia-ac.pdf (ref.6-51)

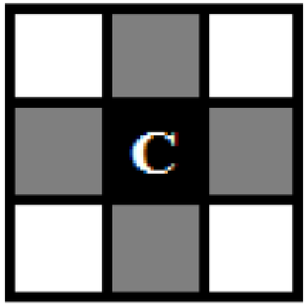
- Os autômatos celulares começaram a ser estudados em torno da década de 40, pelo matemático John Von Neumann. Através deste estudo Neumann tinha por objetivo representar matematicamente a evolução natural, tentando construir uma máquina com auto-replicação.
- Em 1968 o matemático John Conway desenvolve o Jogo da Vida, autômato celular que tem por objetivo representar através de regras matemáticas sistemas complexos da evolução da vida, que o popularizou, provocando o interesse de pesquisadores de diversas áreas.

Autômatos celulares (ref.6-51)

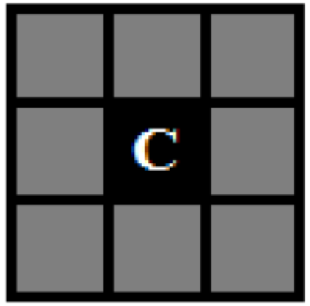
- ***Definição:*** Autômato celular é uma coleção de células com valores próprios, dispostas em uma grade de formato especificado, que evolui em um número discreto de passos de tempo, de acordo com um conjunto de regras baseados nos estados das células vizinhas.
- O estado de um autômato celular é especificado pelos valores das variáveis em cada sítio e esse autômato celular evoluirá com o passo do tempo obedecendo determinadas regras locais que fazem com que os valores dos sítios sejam atualizados simultaneamente dependendo das suas configurações iniciais, sendo que o valor da variável de um sítio em um passo de tempo posterior será afetado pelos valores das variáveis dos sítios vizinhos no passo de tempo anterior, ou seja, o objetivo de tais regras é definir o estado do sítio no passo de tempo seguinte.

Autômatos celulares (ref.6-51)

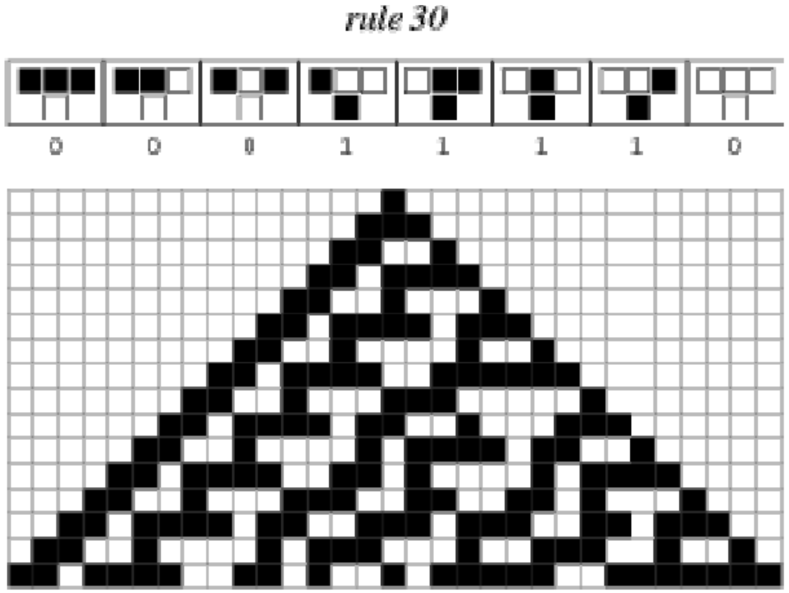
- Assim, os autômatos celulares possuem quatro elementos fundamentais que os caracterizam: o estado inicial, os estados possíveis de cada célula, a sua vizinhança e a regra de transição local.
- Os autômatos celulares podem ser n-dimensionais, porém os mais comuns são os unidimensionais e os bidimensionais. Os autômatos unidimensionais são mais simples e podem ser representados por um vetor de células, e cada geração posterior é representada em uma matriz. Já os autômatos bidimensionais podem ser representados por uma matriz e atualizado em função do tempo.



Vizinhança de Von Neumann



Vizinhança de Moore



Autômatos celulares (ref.6-51)

- ***Aplicações:*** Os autômatos celulares possuem como aplicação na formação de fractais, que podem representar a formação de cristais ou a penetração de fluídos em outro material. Através do Jogo da Vida, podem ser aplicados em diversas áreas e auxiliar no estudo de processos evolutivos de um determinado grupo. Entre os estudos que utilizam o jogo da vida podemos destacar modelagem processos de urbanização e o mapeamento de mutações que levam ao câncer.
- Além dessas aplicações, podem ser citadas implementações na área de criptografia de dados, incêndios florestais, ocupação territorial, modelagem de epidemias etc.