Revisão de estatística

- O objetivo desta aula é apresentar técnicas e conceitos básicos de estatística que são comumente empregados em processamento de sinais.

- A ideia é revisar alguns princípios que são necessários para a compreensão das ferramentas serão apresentadas ao longo do curso

- Para complementar o que foi visto em sala sugere-se a leitura do livro: *Random data: Analysis and Measurement Procedures*. Dos autores: J. S. Bendat; A. G. Piersol.

Função densidade de probabilidade (PDF)

- O comportamento estatístico de diversos processos físicos e eventos pode ser associado à distribuições de probabilidade
- A compreensão e modelagem do problema pode ser bastante facilitada com a análise estatística
- Logo, o uso dessas ferramentas é muito útil para o processamento e análise de sinais
- Existem diversas particularidades, com relação a convergência e detalhes da amostragem dos dados para se representar de maneira adequada uma população.
  Esse conteúdo não é abordado nessa aula de conceitos básicos. Por isso, é importante utilizar referências do curso.

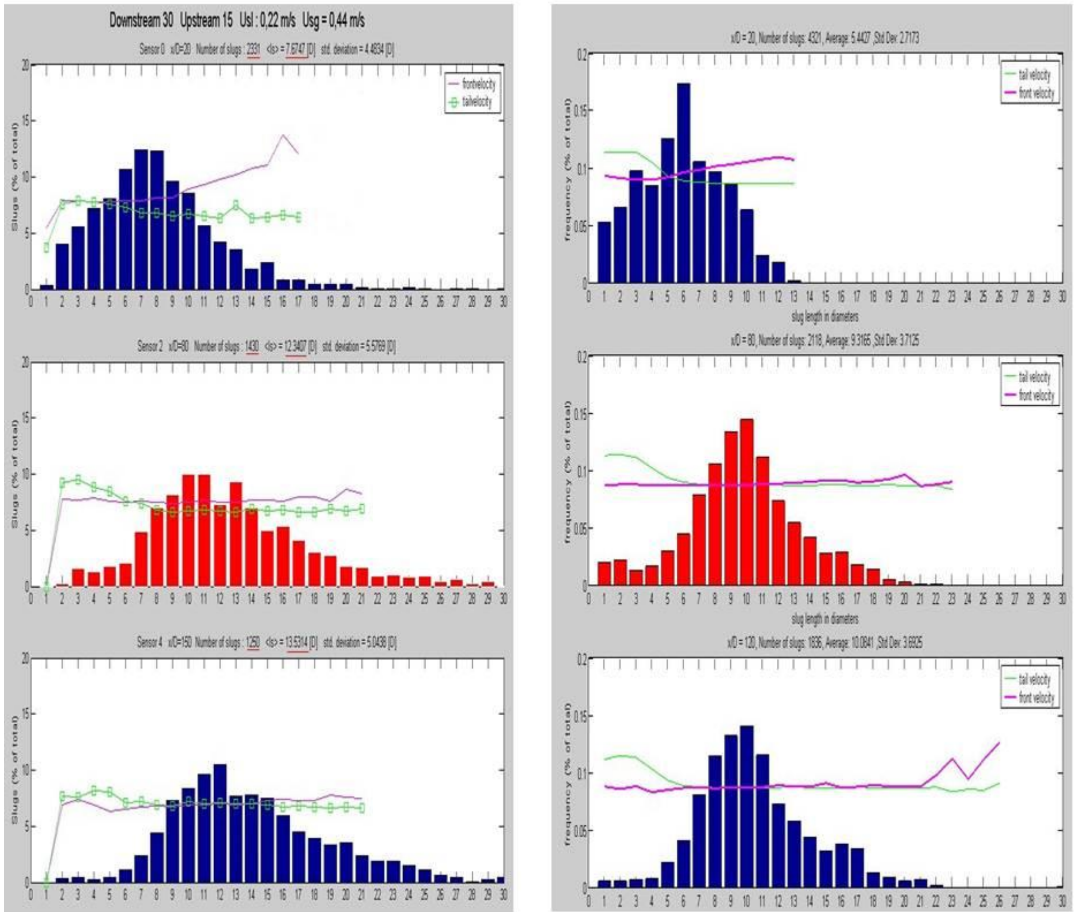
Exemplo (exploração de óleo e gás)

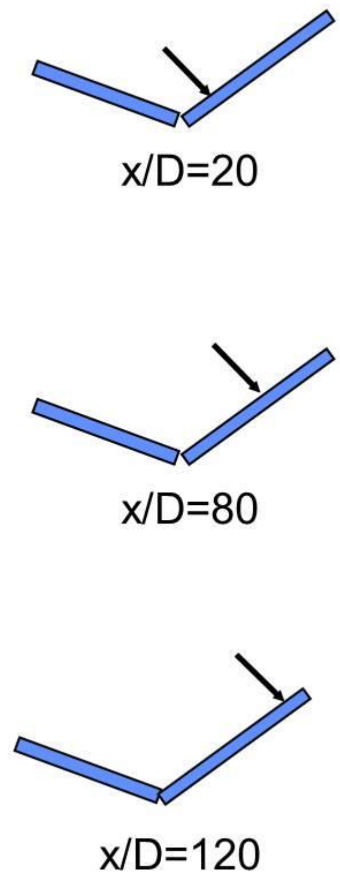
Slug length distributions. Normal slug flow

Experiments

Computations

Stratified flow upstream



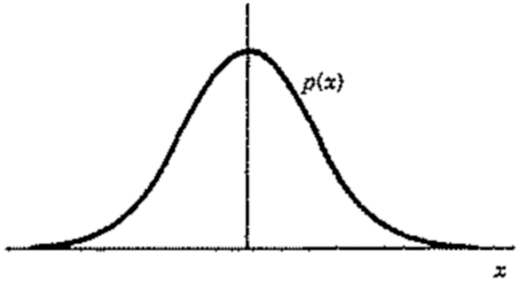


Slug length in diameters

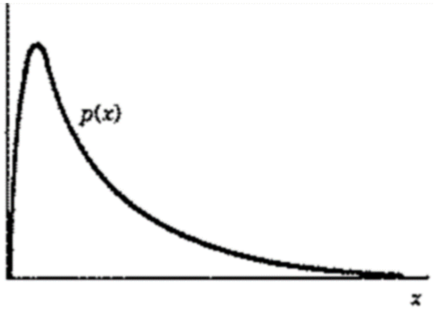
Exemplo (cálculo de incerteza)

- No cálculo de incerteza de medições é necessário associar a incerteza de cada variável medida a um tipo de distribuição estatística. Exemplos:

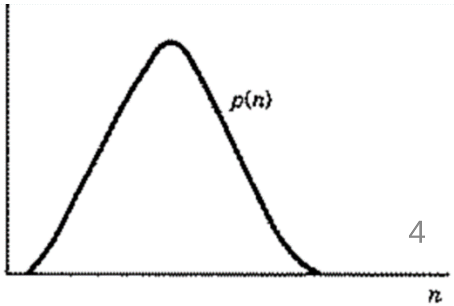
- Normal: Maioria dos casos de grandezas contínuas



- LogNormal: Falha ou projeções de durabilidade. Eventos tendem a ser distorcidos na extremidade da distribuição.



- Binomial: comum em situações que envolvem somente duas possibilidades. Ex.: Cara/Coroa; Transmissão de informação digital



Função densidade de probabilidade (PDF)

- Assumindo x(k) como uma variável de interesse qualquer. A função de densidade de probabilidade, contínua, para essa variável pode ser expressa como:

$$p(x) = \lim_{\Delta x \to 0} \left[\frac{Prob[x < x(k) < x + \Delta x]}{\Delta x} \right]$$

- A função p(x) mostra a probabilidade de se encontrar diferentes valores de x(k) dentro de uma faixa de valores possíveis de x.
- Logo, por definição:

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

Função densidade de probabilidade (PDF)

- Para dados discretos a função densidade de probabilidade pode ser ilustrada graficamente por um histograma
- Ex.: Medição de uma variável qualquer

Nº Medição	Valor Medido
1	5.3054
2	5.0934
3	4.9581
4	5.125
5	5.0366
6	4.794
7	5.1898
8	5.0614
9	5.027
10	5.103
11	5.0523
12	4.8117
13	4.9675
14	4.9708
15	4.8936

Função densidade de probabilidade (PDF)

- Para dados discretos a função densidade de probabilidade pode ser ilustrada graficamente por um histograma

$$n_{\text{intervalos}} = 1.87 \cdot (N-1)^{0.4} + 1$$

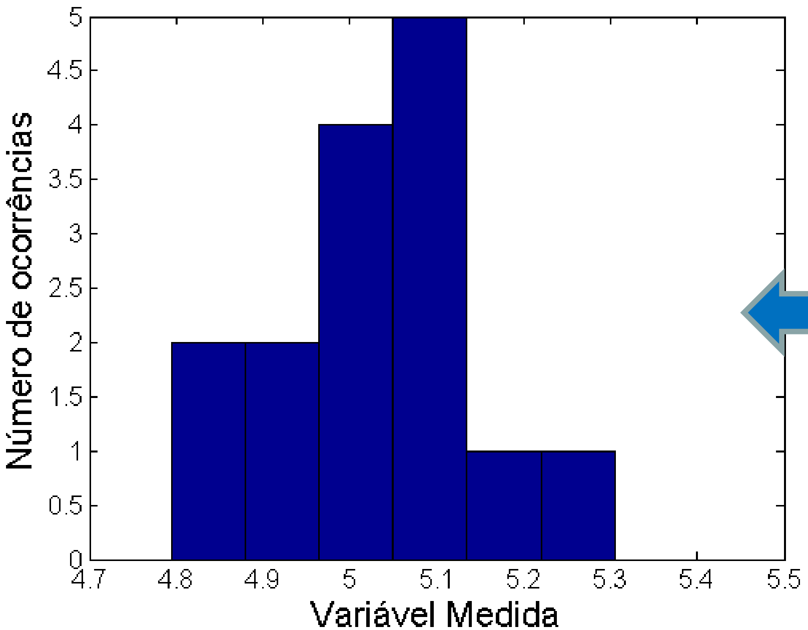
Nº Medição	Valor Medido
1	5.3054
2	5.0934
3	4.9581
4	5.125
5	5.0366
6	4.794
7	5.1898
8	5.0614
9	5.027
10	5.103
11	5.0523
12	4.8117
13	4.9675
14	4.9708
15	4.8936

Nº Intervalo	Intervalo	nº ocorrências	nº ocorrências/N $f(n_i)$
1	$4.8 \le x_i < 4.885$	2	0.13
2	$4.885 \le x_i < 4.97$	2	0.13
3	$4.97 \le x_i < 5.055$	4	0.27
4	$5.055 \le x_i < 5.14$	5	0.33
5	$5.14 \le x_i < 5.225$	1	0.07
6	$5.225 \le x_i < 5.31$	1	0.07

Função densidade de probabilidade (PDF)

- Para dados discretos a função densidade de probabilidade pode ser ilustrada graficamente por um histograma

$$n_{\text{intervalos}} = 1.87 \cdot (N-1)^{0.4} + 1$$

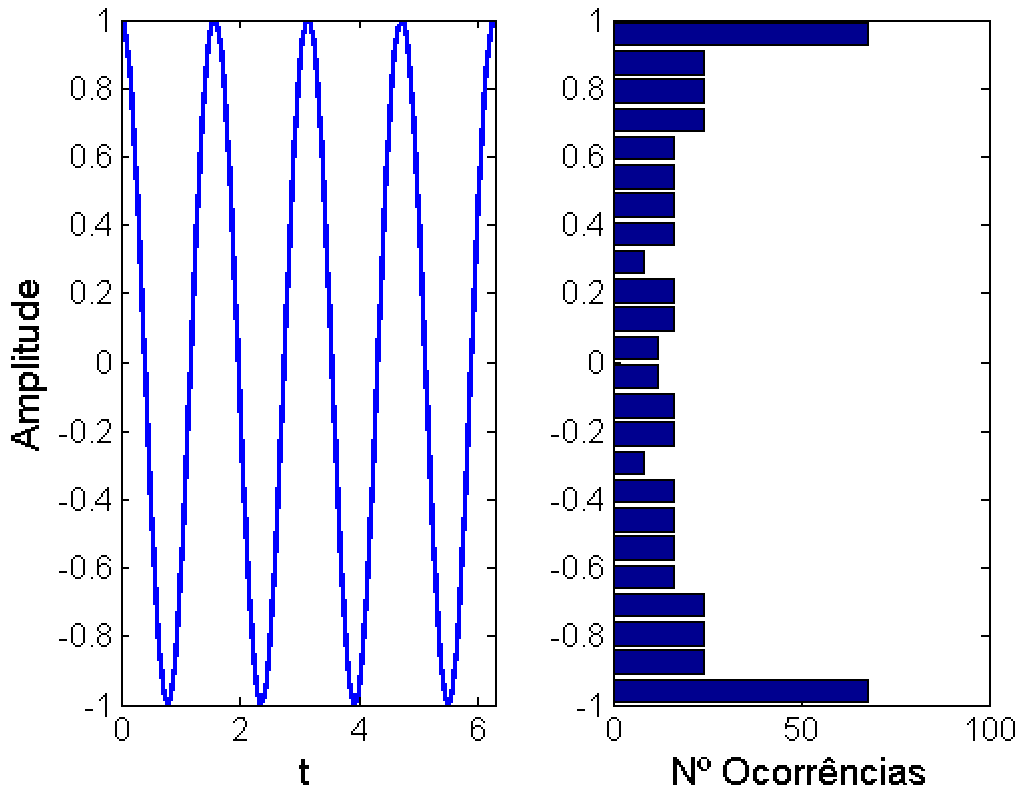


ou $x(k)$

Nº Intervalo	Intervalo	nº ocorrências	nº ocorrências/N $f(n_i)$
1	$4.8 \leq x_i < 4.885$	2	0.13
2	$4.885 \leq x_i < 4.97$	2	0.13
3	$4.97 \leq x_i < 5.055$	4	0.27
4	$5.055 \leq x_i < 5.14$	5	0.33
5	$5.14 \leq x_i < 5.225$	1	0.07
6	$5.225 \leq x_i < 5.31$	1	0.07

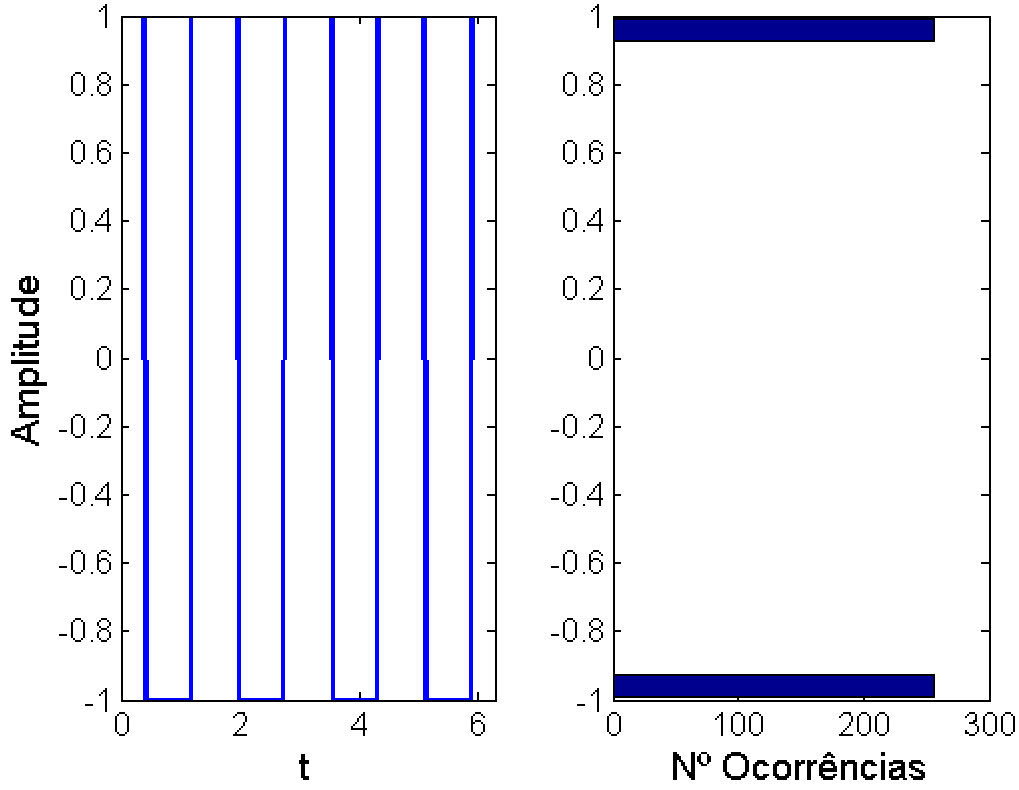
Função densidade de probabilidade (PDF)

- Exemplo de histogramas para diferentes sinais.
  - Sinal senoidal



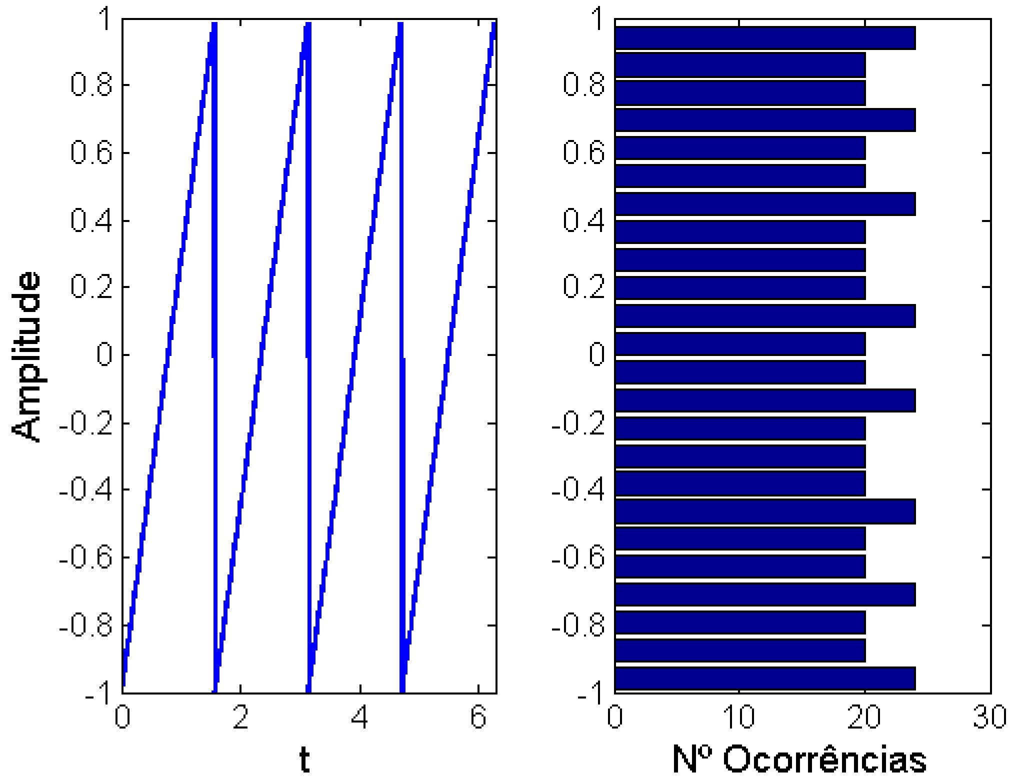
Função densidade de probabilidade (PDF)

- Exemplo de histogramas para diferentes sinais.
  - Onda quadrada



Função densidade de probabilidade (PDF)

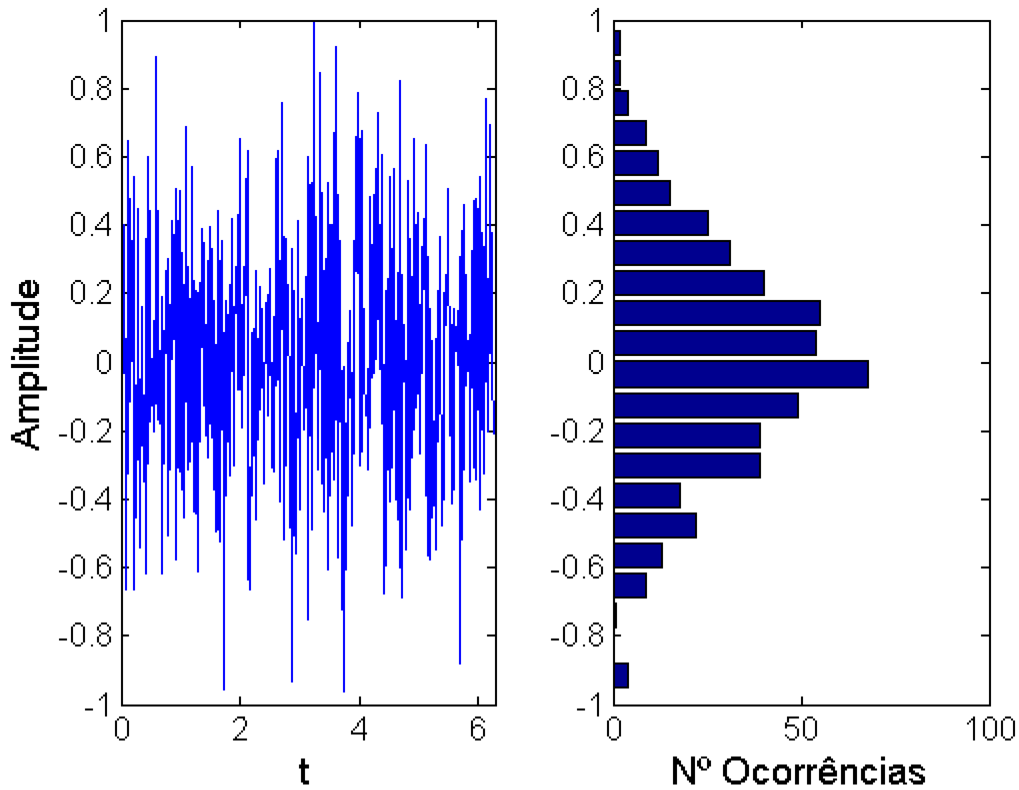
- Exemplo de histogramas para diferentes sinais.
  - Onda dente de serra



Função densidade de probabilidade (PDF)

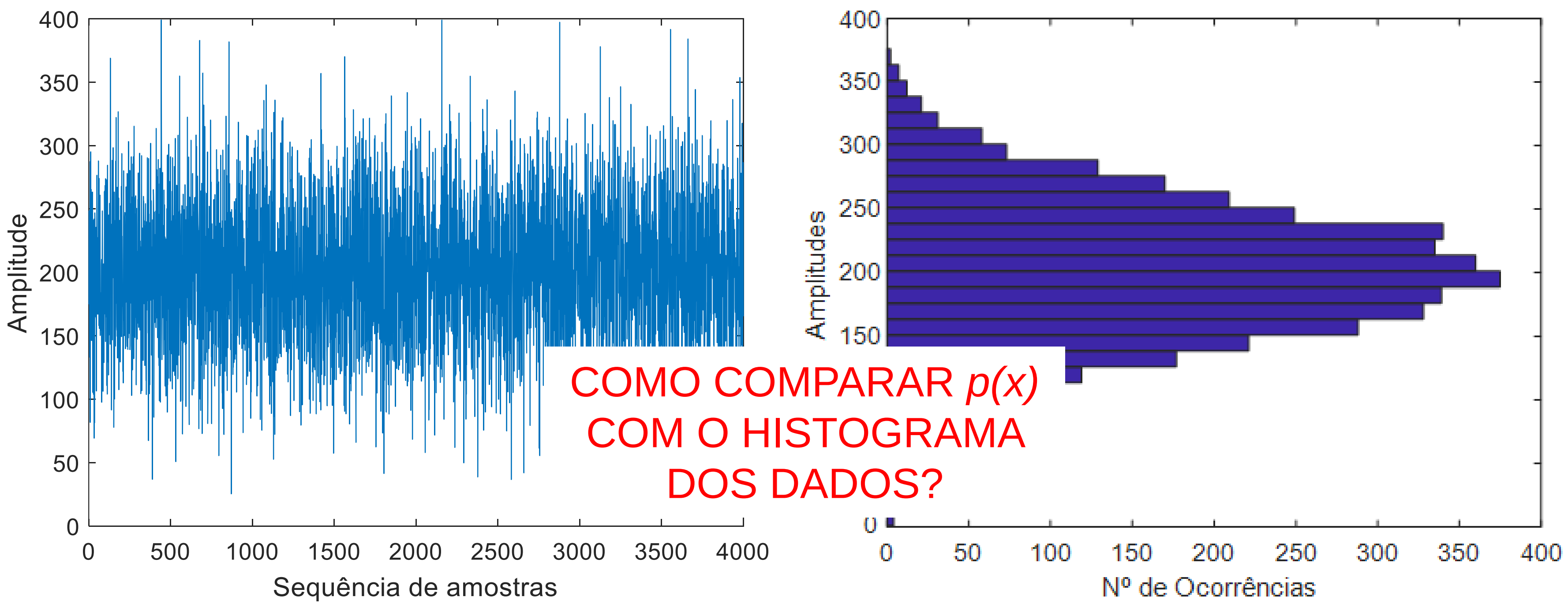
➢Exemplo de histogramas para diferentes sinais.

- Ruído branco (normal ou gaussiano)



Função densidade de probabilidade (PDF)

- Ex.: Distribuição Normal



- Estimando-se a média (μ) e o desvio padrão (σ) das amostras e sabendo-se que p(x) para uma distribuição normal é:

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

Função densidade de probabilidade (PDF)

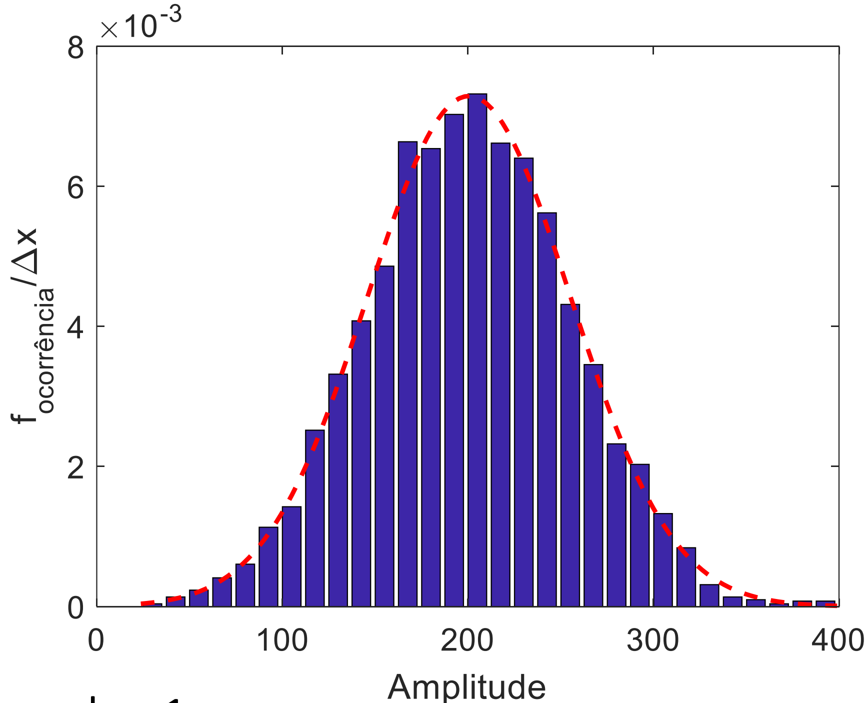
- Pode-se notar que os histogramas podem ilustrar qualitativamente o conceito de uma PDF
- Por que, então não são considerados como PDF’s?
- A amplitude do histograma é normalmente expressa em número de ocorrências e a dispersão da variável x(k) em valores dimensionais, logo a integral do histograma é diferente de 1.
- Sendo assim, é necessário normalizar o histograma para que ele possa vir a representar uma PDF.

Função densidade de probabilidade (PDF)

- Para normalizar um histograma é necessário simplesmente considerar que a frequência de ocorrências (ou nº ocorrências/nº de amostras) é obtido para um dado intervalo.
- Logo, deve-se normalizar a frequência de ocorrências pelo intervalo.
- Calculando-se a média e desvio padrão dos dados é possível comparar, qualitativamente, os dados obtidos com a função densidade de probabilidade correspondente

Função densidade de probabilidade (PDF)

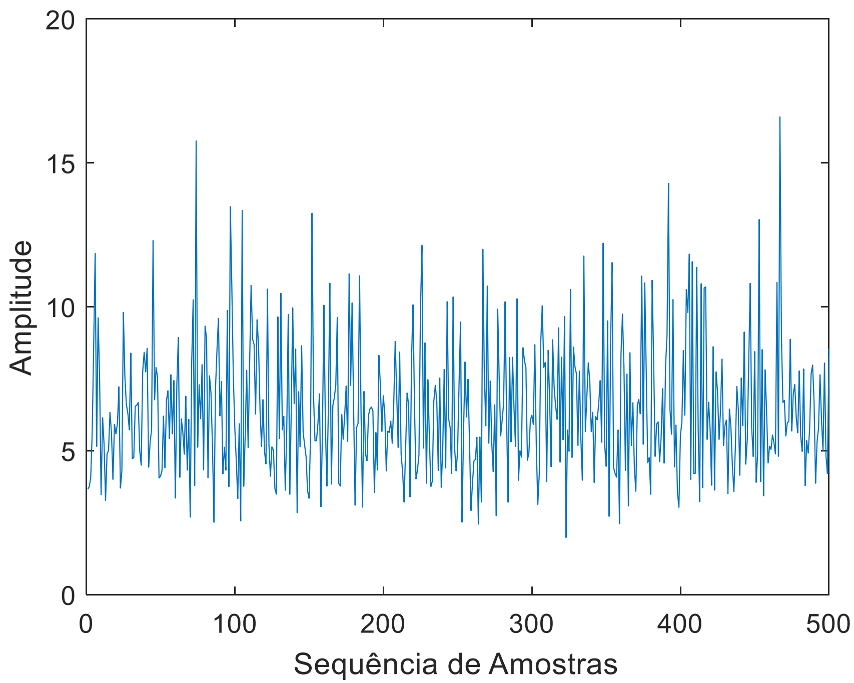
- A figura ilustra a comparação entre uma PDF obtida a partir de uma série de dados e a função p(x)
- Onde Δx é o intervalo escolhido para o cálculo do nº de ocorrências
- Para verificar se o procedimento está correto, pode-se realizar a integral da área do histograma.

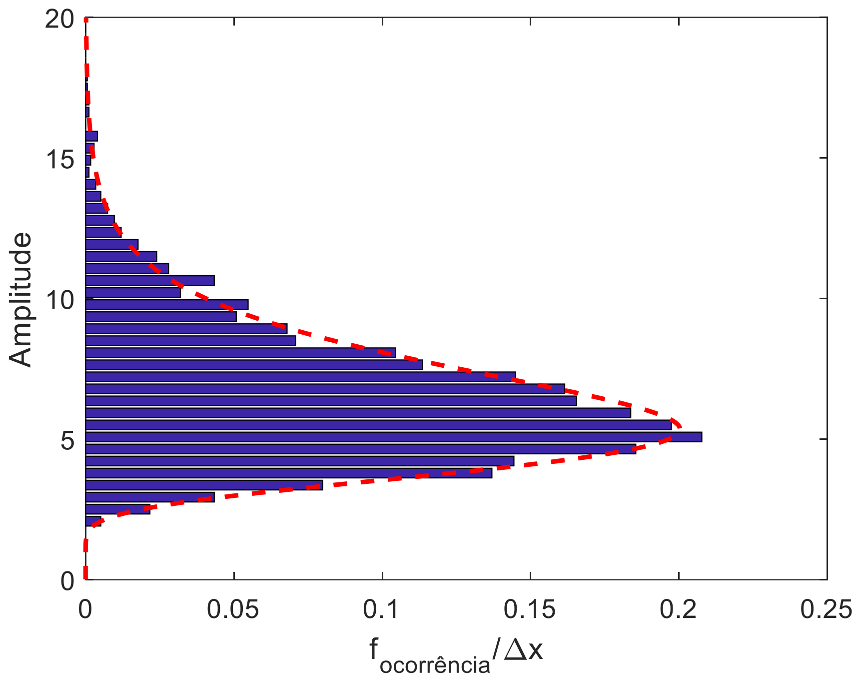


- O resultado tem que ser igual a 1

Função densidade de probabilidade (PDF)

- Ex.: Distribuição LogNormal





$$p(x) = \frac{1}{x\sqrt{2\pi\sigma^{*2}}} e^{\left[-\frac{1}{2}\left(\frac{\ln(x)-\mu*}{\sigma*}\right)^2\right]}$$

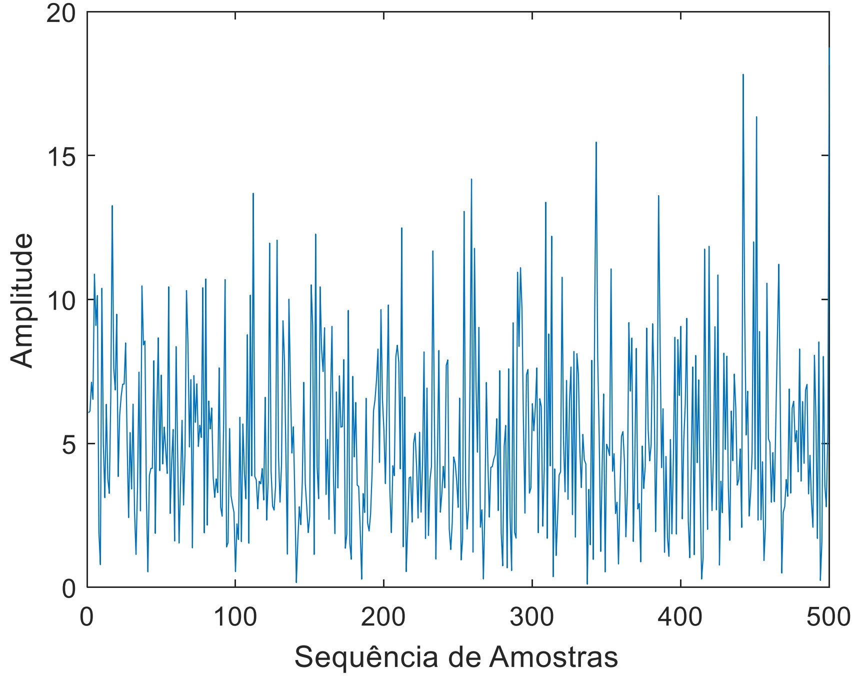
onde

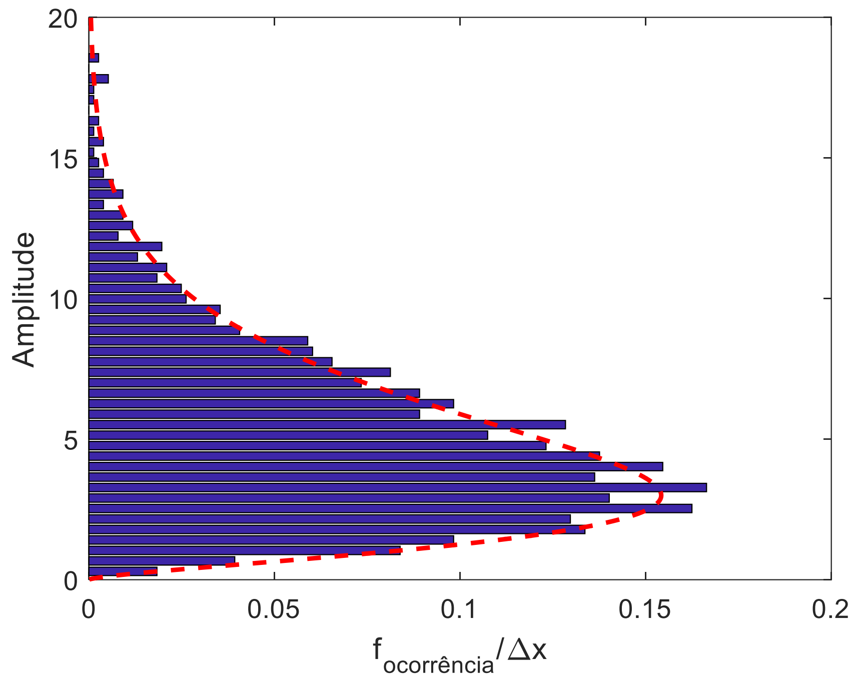
$$\mu^* = ln\left(\frac{\mu^2}{\sqrt{\sigma+\mu^2}}\right)$$

$$\sigma^{*2} = ln\left(\frac{\sigma}{\mu^2}+1\right)$$

Função densidade de probabilidade (PDF)

- Ex.: Distribuição χ2





$$p(x) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$

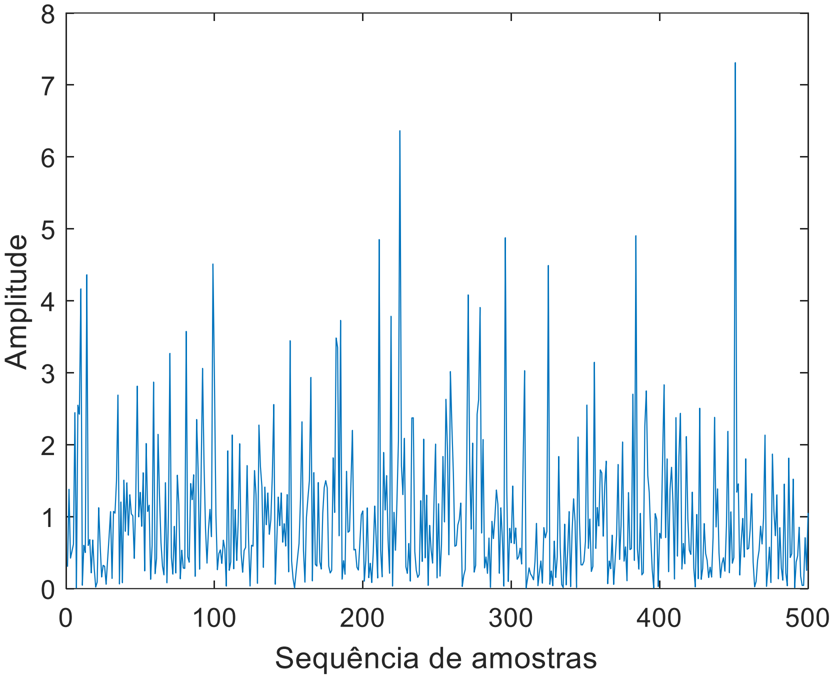
onde

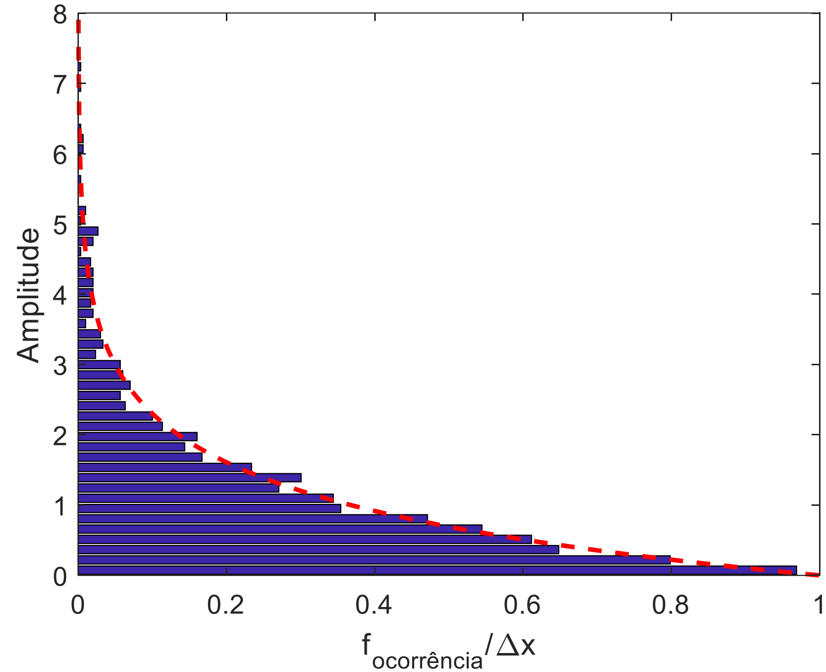
$$\Gamma(n) = (n-1)!$$

$$k = \mu$$

Função densidade de probabilidade (PDF)

➢ Ex.: Distribuição exponencial





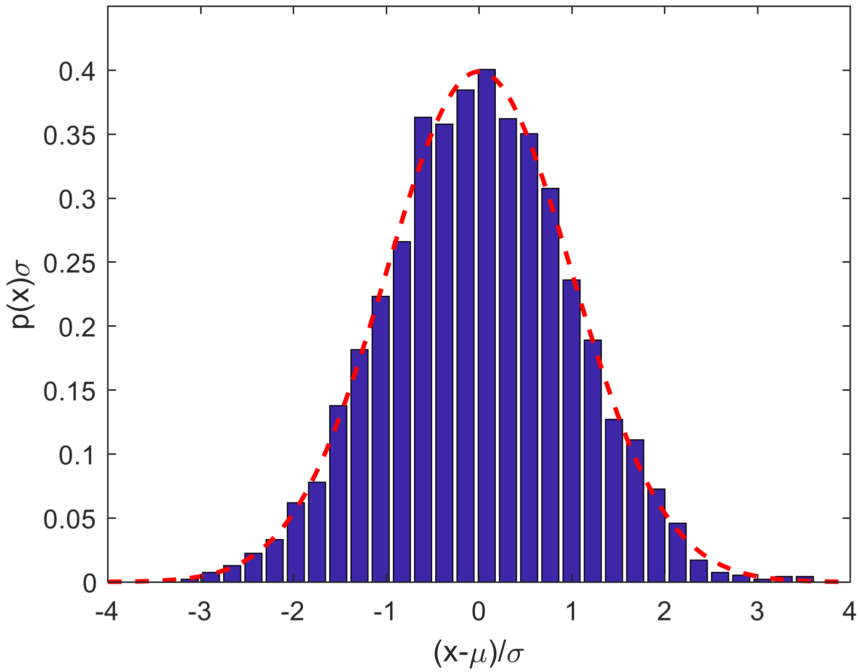
$$p(x) = \lambda e^{-\lambda x}$$

onde $\lambda = \frac{1}{\mu}$

Função densidade de probabilidade (PDF)

- Em alguns casos é possível normalizar os resultados para que sejam comparados com uma distribuição normal de média 0 e desvio padrão unitário

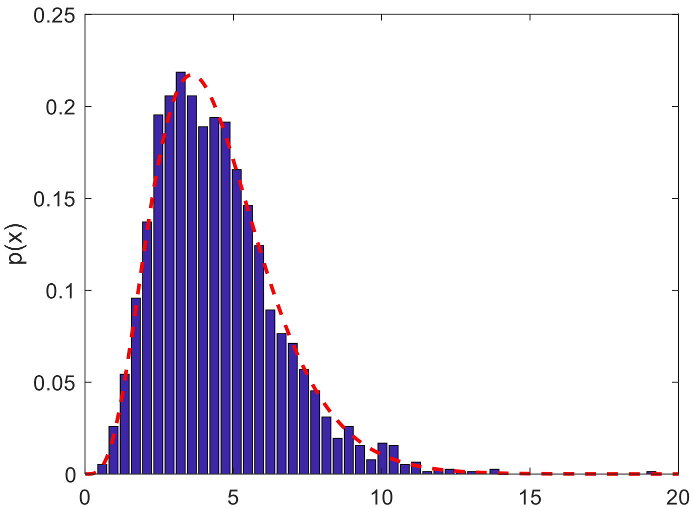
$$p(x) = \left(\frac{n_{ocorrências}}{N_{amostras}} \right) \Big/ \Delta x \quad ; \quad p(x)_{Normalizada} = p(x) \cdot \sigma \qquad x_{normalizado} = (x - \mu) \big/ \sigma$$



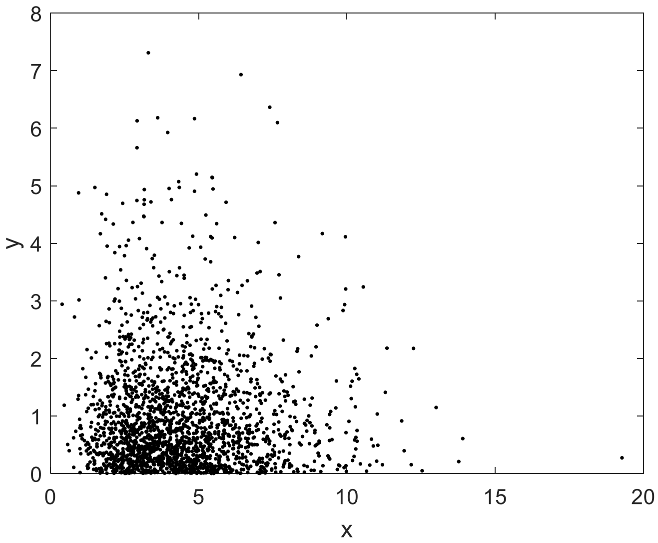
PDF – área sob a curva continua =1

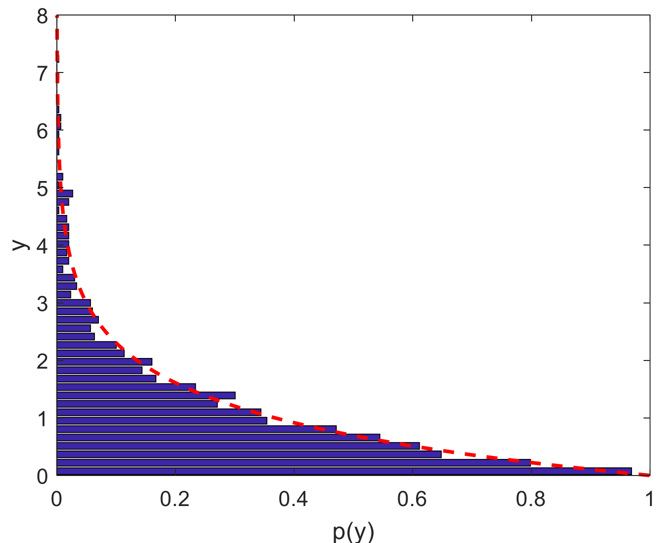
Função densidade de probabilidade (PDF)

➢Casos com mais de uma variável independente (ex.: *x,y*) podem ter distribuições combinadas: *p(x,y)*



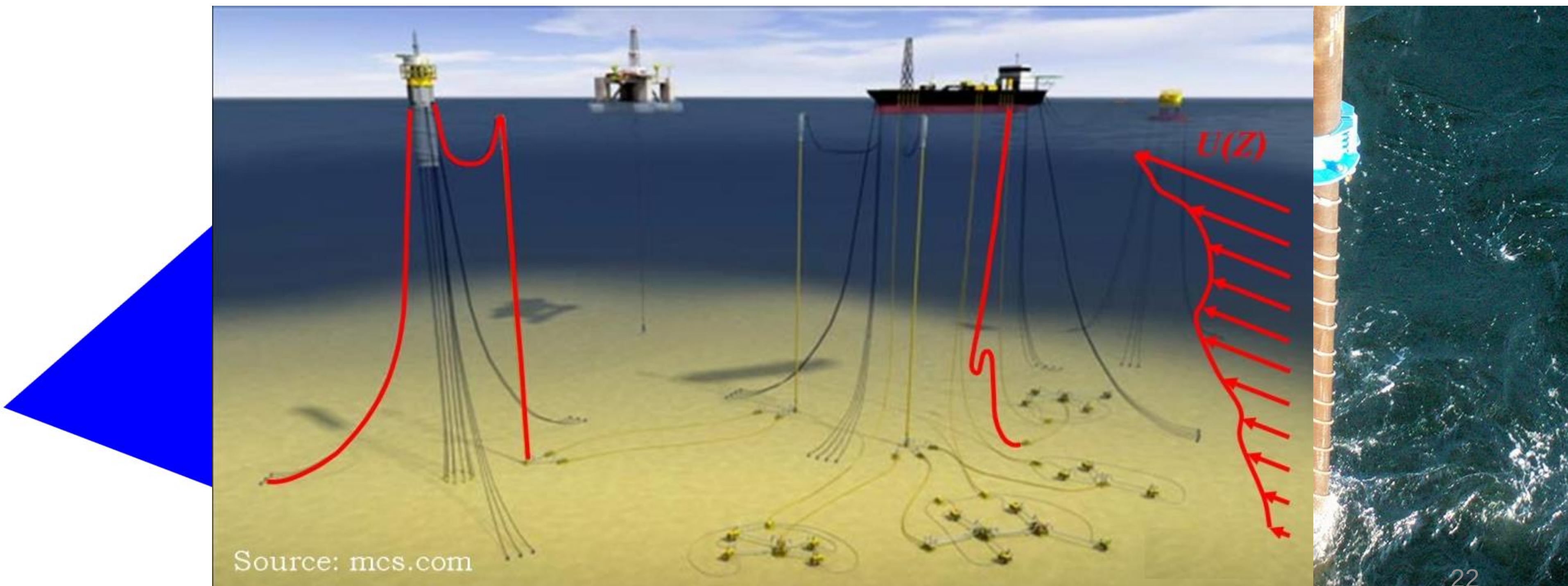
No exemplo têm-se uma distribuição chi2 em x e exponencial em y





Análise básica de sinais – Motivação

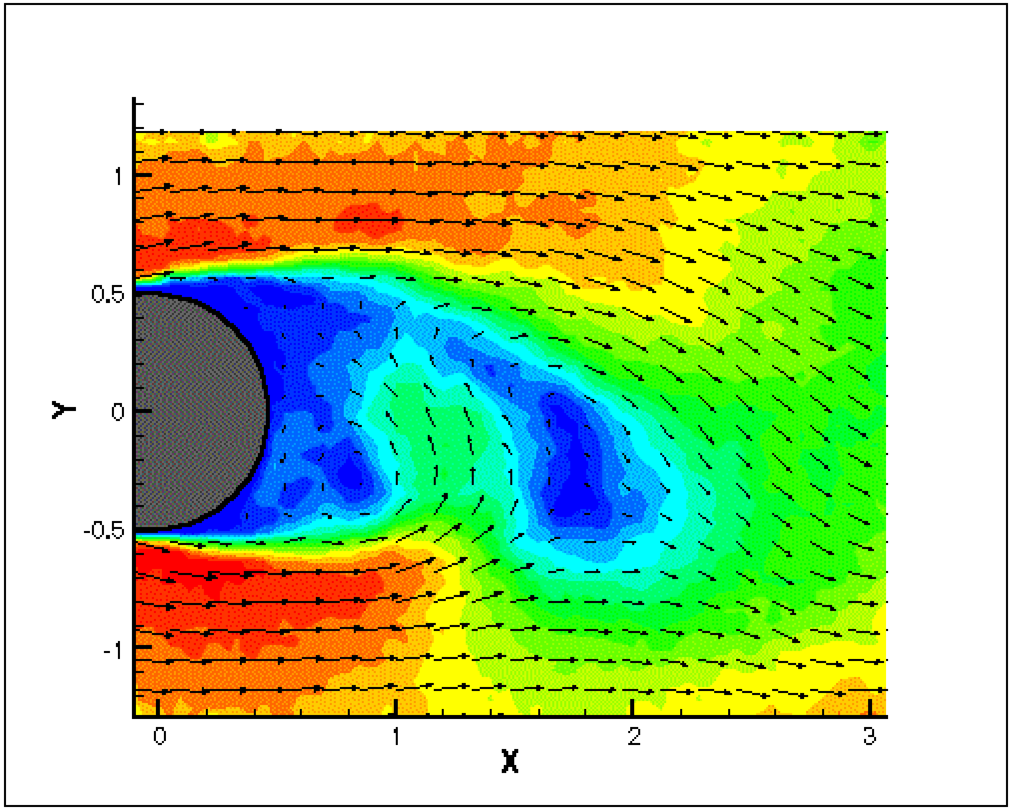
- Imagine que o cilindro abaixo, representa um corte transversal de uma tubulação de um riser, por onde escoa um fluido internamente e há uma corrente marinha. O dispositivo dispõe de medidores de deformação da linha e um sensor de velocidade.



Análise básica de sinais – Motivação

- A tubulação oscila e com os dados dos medidores de velocidade e deformação deseja-se saber :

  - A oscilação é devida ao escoamento no interior da tubulação ou ao escoamento externo?
  - O problema pode ser modelado através de pequenas perturbações (linear)?
  - Em algum momento houve falha ne leitura dos equipamentos?



COMO VOCÊ FARIA PARA TENTAR RESPONDER ALGUMA DESSAS PERGUNTAS?

Covariância e coeficiente de correlação

- Nos casos com mais de uma variável independente (ex.: *x*,*y*) a densidade de probabilidade combinada pode ser expressa como: p(x,y)=p(x)p(y)
- Onde

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)dxdy = 1$$

- O valor esperado para para uma função contínua g(x,y) de duas variáveis x(k) e y(k) é dado por

$$\mathrm{E[g(x,y)]} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)p(x,y)dxdy = 1$$

- No caso da função g(x,y) ser simplesmente o produto das variáveis independentes, sem a média

Covariância e coeficiente de correlação

- No caso onde a função g(x,y) é simplesmente o produto das variáveis independentes, subtraídas das respectivas médias, tem-se:

$$g(x,y) = [x(k) - \mu_x][y(k) - \mu_y]$$

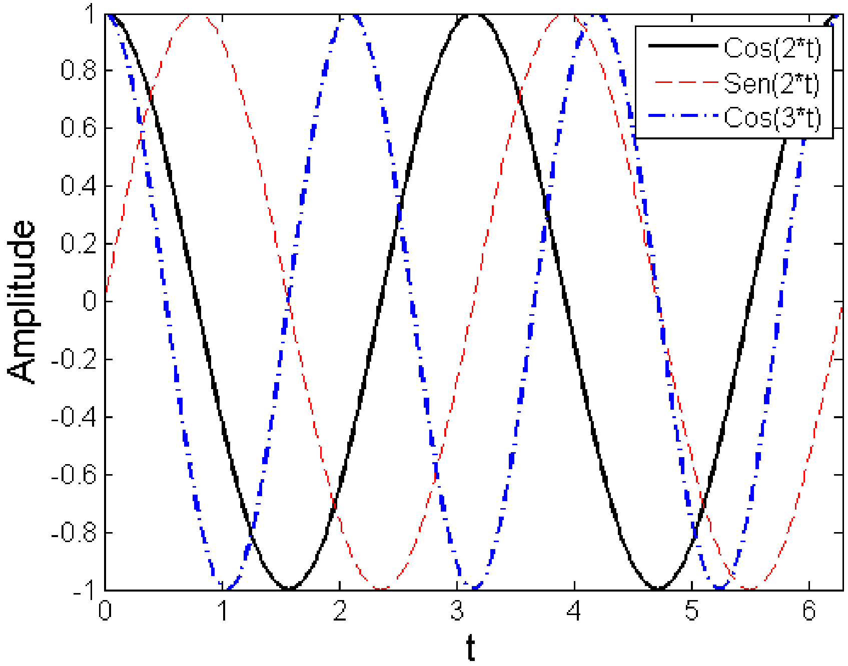
- Assim, tem-se a covariância.
- Pode-se notar que se x(k)=y(k), tem-se a variância de x(k)
- Para dados amostrados, a integral de g(x,y)p(x,y) pode ser escrita na forma de somatório:

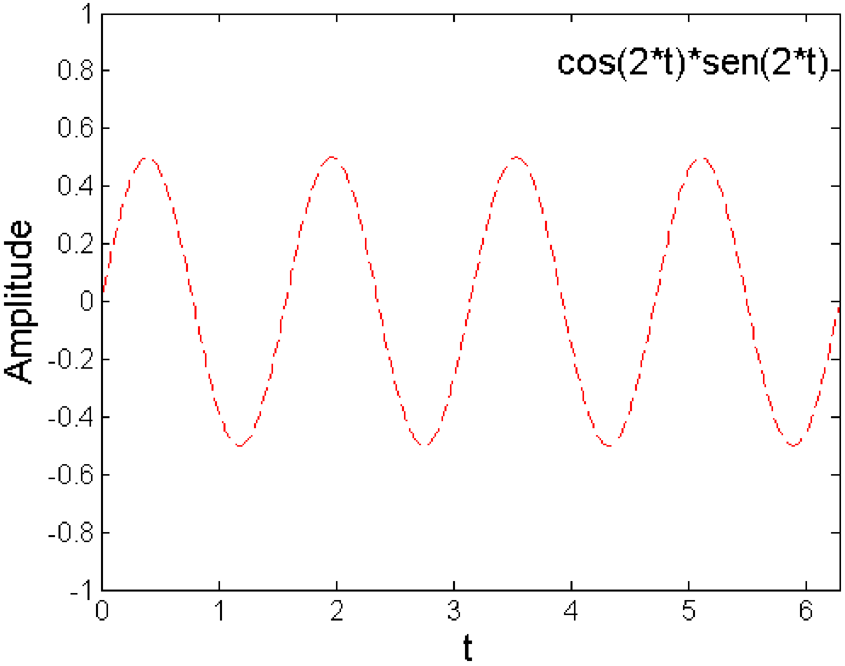
$$\operatorname{cov}(x,y) = \sum_{j=1}^{N} \frac{(x_j - \mu_X)(y_j - \mu_y)}{N}$$

Covariância e coeficiente de correlação

- Covariância (medida da correlação entre 2 sinais).

$$\mathrm{cov}(X,Y)=\sum_{j=1}^{N}\frac{(X_j-\mu_X)(Y_j-\mu_j)}{N}$$

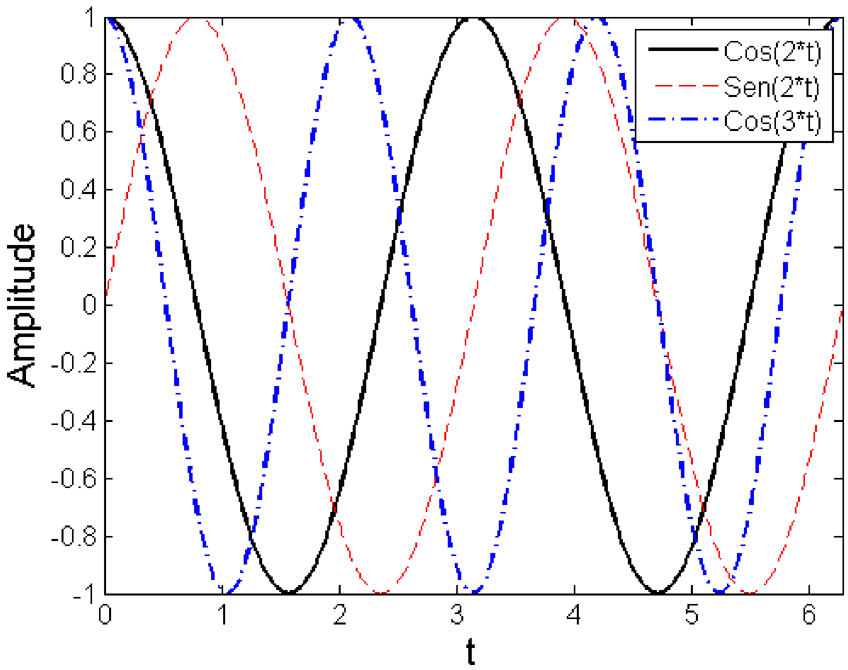


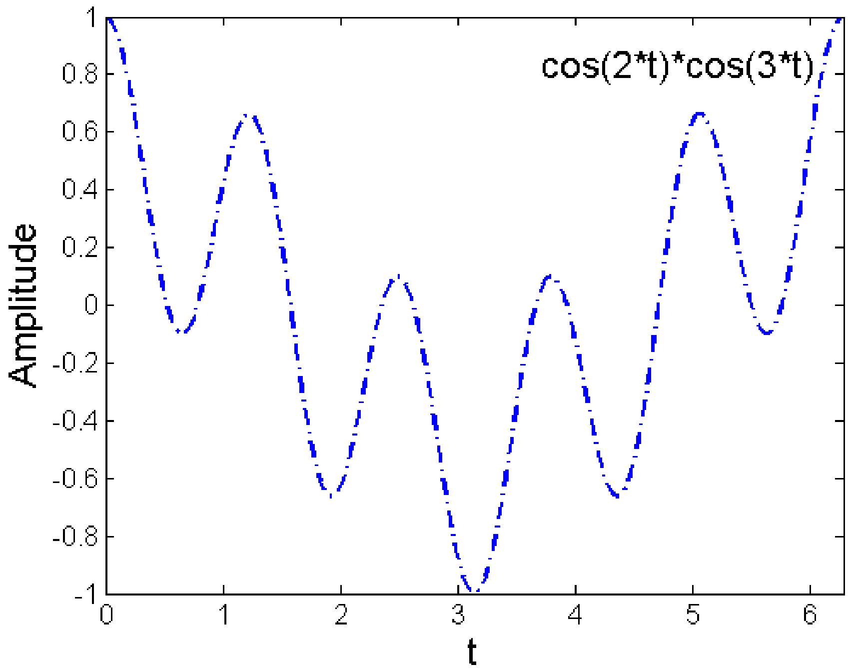


Covariância e coeficiente de correlação

➢Covariância (medida da correlação entre 2 sinais).

$$\mathrm{cov}(X,Y)=\sum_{j=1}^{N}\frac{(X_j-\mu_X)(Y_j-\mu_j)}{N}$$



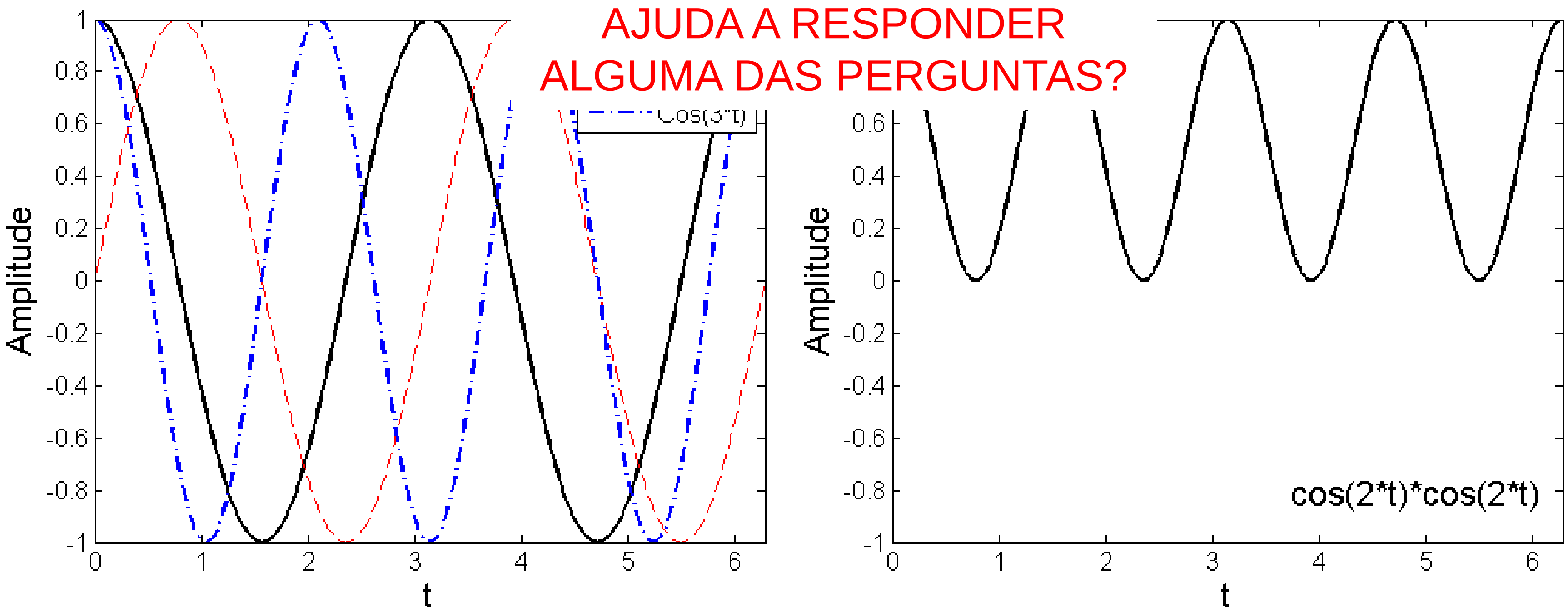


Covariância e coeficiente de correlação

➢ Covariância (medida da correlação entre 2 sinais).

$$\operatorname{cov}(X,Y) = \sum_{j=1}^{N} \frac{(X_j - \mu_X)(Y_j - \mu_j)}{N}$$

AJUDA A RESPONDER ALGUMA DAS PERGUNTAS?



A covariância só é diferente de zero quando os sinais x e y tiverem alguma correlação

Covariância e coeficiente de correlação

- Pode-se deduzir, facilmente, uma relação entre covariância e os desvios padrões de x e y.

$$cov(x,x) = E[(x(k) - \mu_x)^2] = {\sigma_x}^2$$

$$cov(y,y) = E[(y(k) - \mu_y)^2] = {\sigma_y}^2$$

$$E[\sqrt{(x(k) - \mu_x)^2 \left(y(k) - \mu_y\right)^2}] = \sqrt{{\sigma_x}^2 {\sigma_y}^2 =} \sigma_x \sigma_y$$

- Logo, o máximo valor que a covariância pode assumir, ocorre quando somente o produto entre as variâncias é considerado. Logo, a covariância pode ser normalizada por esse produto, assumindo valores entre -1 e 1. Essa razão é conhecida como coeficiente de correlação (c.c)

$$\mathrm{c.c} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

Covariância– notação matricial

- Para o cálculo da covariância entre múltiplos dados, conforme ilustrado na matriz x, é conveniente utilizar a notação matricial.

$$x = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix};$$

- Nesse caso, é necessário criar uma matriz dos desvios. O vetor das médias de cada coluna (j) é dado por x_m (assumindo dados estão dispostos em colunas, e 1_n é um vetor de nx1)

$$x_{mj} = \left(\bar{x}_1, \bar{x}_2, \bar{x}_3, \dots, \bar{x}_p\right) = \frac{1}{n} 1_n^{'} x_j$$

Covariância– notação matricial

- A partir da média pode-se calcular a matriz de diferenças através da relação:

$$x_d = x - 1_n x_m = \begin{bmatrix} x_{11} - \bar{x}_1 & x_{12} - \bar{x}_2 & x_{13} - \bar{x}_3 & \dots & x_{1p} - \bar{x}_p \\ x_{21} - \bar{x}_1 & x_{22} - \bar{x}_2 & x_{23} - \bar{x}_3 & \dots & x_{2p} - \bar{x}_p \\ \dots & \dots & \dots & \dots & \dots \\ x_{n1} - \bar{x}_1 & x_{n2} - \bar{x}_2 & x_{n3} - \bar{x}_3 & \dots & x_{np} - \bar{x}_p \end{bmatrix};$$

- Com isso a matriz de covariâncias pode ser rescrita na forma

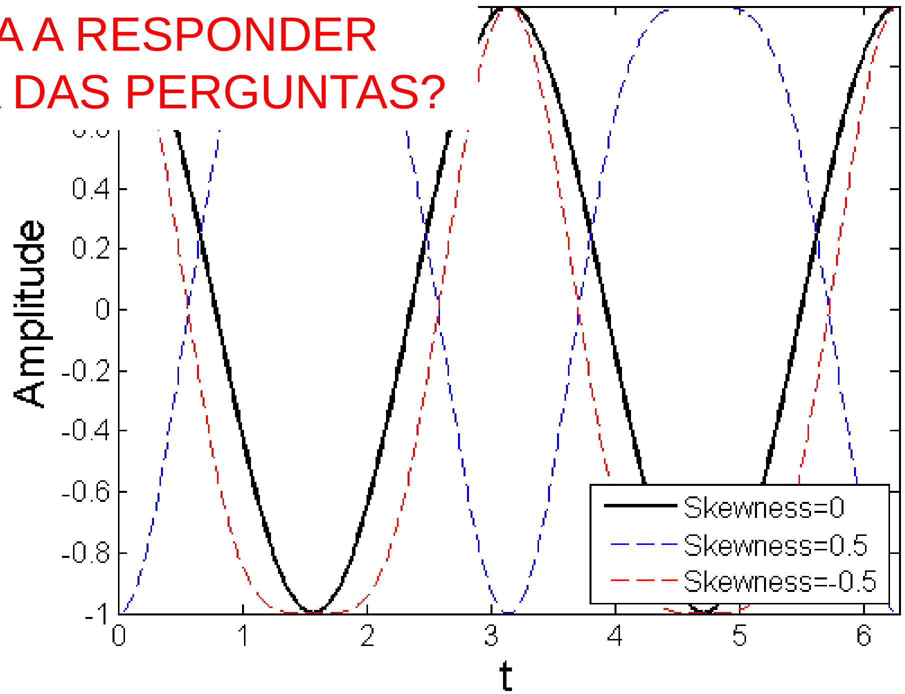
$$\mathrm{cov} = \frac{1}{n} x_d{}' x_d = \begin{bmatrix} \sigma_{11}^2 & \mathrm{cov}_{12} & \mathrm{cov}_{13} & \dots & \mathrm{cov}_{1p} \\ \mathrm{cov}_{21} & \sigma_{22}^2 & \mathrm{cov}_{23} & \dots & \mathrm{cov}_{2p} \\ \dots & \dots & \dots & \dots & \dots \\ \mathrm{cov}_{n1} & \mathrm{cov}_{n2} & \mathrm{cov}_{n3} & \dots & \sigma_{np}^2 \end{bmatrix};$$

Coeficiente de assimetria (skewness)

- É uma medida para quantificar a assimetria de um sinal. É comum a utilização para se identificar saturação de um sinal devido a efeitos não lineares. Ex.: ondas se aproximando da praia

AJUDA A RESPONDER
ALGUMA DAS PERGUNTAS?

$$skewness = \frac{\frac{1}{N}\sum_{j=1}^{N}\left(X_j - \mu_X\right)^3}{\left(\sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(X_j - \mu_X\right)^2}\right)^3}$$

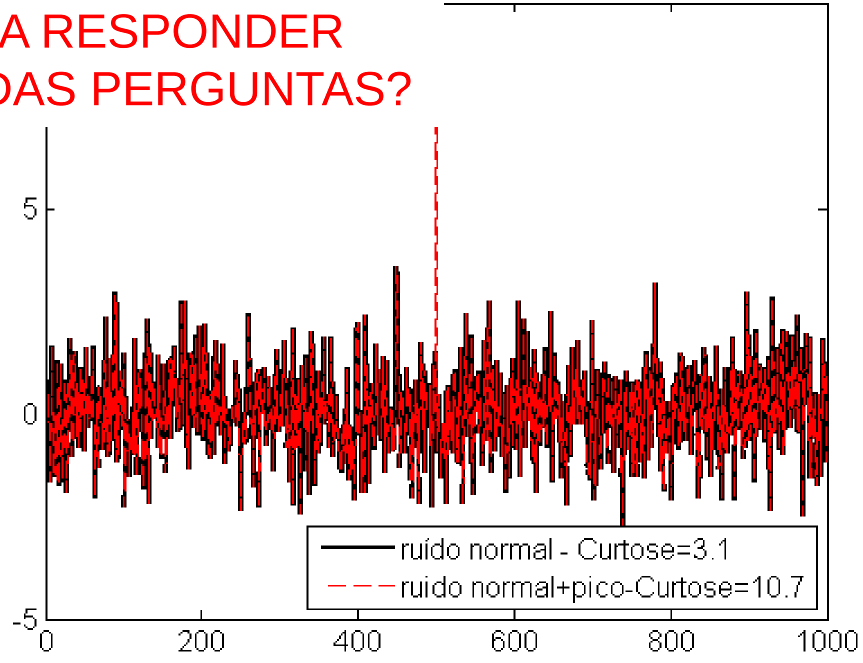


Curtose (Kurtosis)

➢ É uma medida do grau de espalhamento da distribuição de probabilidades. Isso faz com que seja útil para identificar dados espúrios (outliers) em uma sequência de dados.

AJUDA A RESPONDER ALGUMA DAS PERGUNTAS?

$$Curtose = \frac{\frac{1}{N}\sum_{j=1}^{N}\left(X_j - \mu_X\right)^4}{\left(\frac{1}{N}\sum_{j=1}^{N}\left(X_j - \mu_X\right)^2\right)^2}$$



Regressão linear

Regressão (Método dos Mínimos Quadrados)

Aproximação polinomial de ordem *m* dos dados:

$$f(x) = y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \ldots a_m x^m$$

A relação polinomial deve ser encontrada para um conjunto de N pontos de dados [da forma (x_i, y_i)].

Dedução para eq. do 1º grau:

$$f(x) = y = a_0 + a_1 x$$

onde y é a resultado da equação de ajuste e a_0 e a_1 são, respectivamente, os coeficientes linear e angular.

O desvio do ajuste pode ser dado por:

$$\Delta^2 = \sum (Y_i - y)^2$$

onde Y_i é um dos pontos medidos e y_i o ponto fornecido pelo ajuste

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

$$\Delta^2 = \sum (Y_i - y)^2$$

O objetivo é obter coeficientes que minimizem a soma dos erro quadrados, de modo que:

$$\frac{\partial \Delta^2}{\partial a_1} = \frac{\partial}{\partial a_1} \sum (Y_i - y)^2 = \frac{\partial}{\partial a_1} \sum (Y_i - a_1 x - a_0)^2 = 0$$

$$= \frac{\partial}{\partial a_1} \sum \left(Y_i^2 - 2Y_i a_1 x - 2Y_i a_0 + a_1^2 x^2 + 2a_1 x a_0 + a_0^2 \right) = 0$$

$$= 2\sum (Y_i - a_1 x - a_0)(-x) = 0$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

$$\Delta^2 = \sum (Y_i - y)^2$$

O objetivo é obter coeficientes que minimizem a soma dos erro quadrados, de modo que:

$$\frac{\partial \Delta^2}{\partial a_0} = \frac{\partial}{\partial a_0} \sum (Y_i - y)^2 = \frac{\partial}{\partial a_0} \sum (Y_i - a_1 x - a_0)^2 = 0$$

$$= \frac{\partial}{\partial a_0} \sum \left(Y_i^2 - 2Y_i a_1 x - 2Y_i a_0 + a_1^2 x^2 + 2a_1 x a_0 + a_0^2\right) = 0$$

$$= 2\sum (Y_i - a_1 x - a_0)(-1) = 0$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

Resolvendo para a_0:

$$\sum Y_i - a_1 \sum x - n a_0 = 0$$

$$a_0 = \frac{\sum Y_i - a_1 \sum x}{n}$$

E substituindo para encontrar a_1 :

$$\sum Y_i x - a_1 \sum x^2 - a_0 \sum x = 0$$

$$\sum Y_i x - a_1 \sum x^2 - \frac{\left(\sum Y_i - a_1 \sum x\right)}{n} \sum x = 0$$

$$a_1 = \frac{\sum Y_i \sum x - n \sum Y_i x}{\left(\sum x\right)^2 - n \sum x^2}$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

Erro padrão do ajuste (1ª ordem):

$$S = \sqrt{\frac{\Delta^2}{N-2}} \rightarrow \quad S = \sqrt{\frac{\sum (Y_i - y)^2}{N-2}}$$

$$S_{a0} = S\sqrt{\frac{\sum x^2}{n\sum x^2 - \left(\sum x\right)^2}} \qquad S_{a1} = S\sqrt{\frac{n}{n\sum x^2 - \left(\sum x\right)^2}}$$

De maneira simplificada, é possível estimar um intervalo de confiança para os dados em torno do ajuste a partir do erro padrão experimental

$$\pm t_{N-1,P}\frac{S}{\sqrt{N}} \ (\%P)$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

  Existe uma solução para o problema envolvendo produto de matrizes, na forma:

$$a = \left(x^T x\right)^{-1} x^T y$$

  Onde o expoente T se refere a transposta da matriz e -1 a inversa, com as matrizes *x*,*y* e *a* sendo :

$$y = \begin{bmatrix} y_0 \\ y_1 \\ \ldots \\ y_n \end{bmatrix}; \quad x = \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \ldots & \ldots \\ 1 & x_n \end{bmatrix}; \quad a = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

  Existe uma solução para o problema envolvendo produto de matrizes, na forma:

$$a = \left(x^T x\right)^{-1} x^T y$$

  Prova :

$$x^T x = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} 1 & x_0 \\ 1 & x_1 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} = \begin{bmatrix} n & \sum x \\ \sum x & \sum x^2 \end{bmatrix}$$

$$x^T y = \begin{bmatrix} 1 & 1 & \dots & 1 \\ x_0 & x_1 & \dots & x_n \end{bmatrix} \cdot \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum xy \end{bmatrix}$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

  Existe uma solução para o problema envolvendo produto de matrizes, na forma:

$$a = \left(x^T x\right)^{-1} x^T y$$

  Prova :

$$\left(x^T x\right)^{-1} = \frac{1}{n\sum x^2 - \left(\sum x\right)^2} \begin{bmatrix} \sum x^2 & -\sum x \\ -\sum x & n \end{bmatrix}$$

  Substituindo:

$$\left(x^T x\right)^{-1} x^T y = \frac{1}{n\sum x^2 - \left(\sum x\right)^2} \begin{bmatrix} \sum x^2 & -\sum x \\ -\sum x & n \end{bmatrix} \cdot \begin{bmatrix} \sum y \\ \sum xy \end{bmatrix}$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

  Existe uma solução para o problema envolvendo produto de matrizes, na forma:

$$a = \left(x^T x\right)^{-1} x^T y$$

  Prova :

$$\left(x^T x\right)^{-1} x^T y = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \frac{1}{n\sum x^2 - \left(\sum x\right)^2} \begin{bmatrix} \sum x^2 \sum y - \sum x \sum xy \\ -\sum x \sum y + n \sum xy \end{bmatrix}$$

Equação encontrada anteriormente:

$$a_1 = \frac{\sum Y_i \sum x - n \sum Y_i x}{\left(\sum x\right)^2 - n \sum x^2}$$

Regressão linear

- Regressão (Método dos Mínimos Quadrados)

Polinômios de mais alta ordem também podem ser ajustados usando essa operação:

$$a = \left(x^T x\right)^{-1} x^T y$$

Só que nesses casos as matrizes ficam:

$$y = \begin{bmatrix} y_0 \\ y_1 \\ \dots \\ y_n \end{bmatrix}; \quad x = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^m \\ 1 & x_1 & x_1^2 & \dots & x_1^m \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{bmatrix}; \quad a = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix}$$

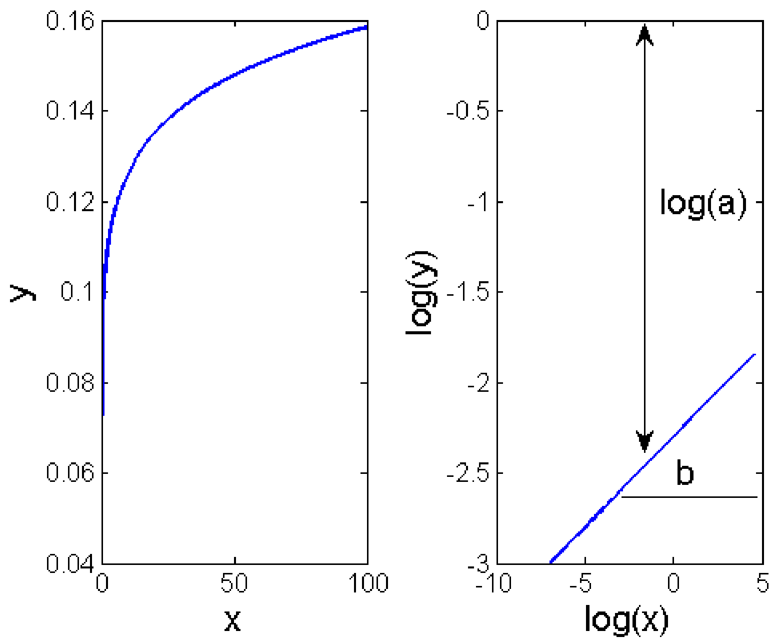
Regressões de múltiplas variáveis também podem ser obtidas

Regressão linear

- **Exemplos**
  - Regressão (Método dos Mínimos Quadrados)

    O método pode ser aplicado para a regressão de diversas funções, para isso é necessário somente encontrar uma escala adequada para os dados.

    Ex.: $y = ax^b$

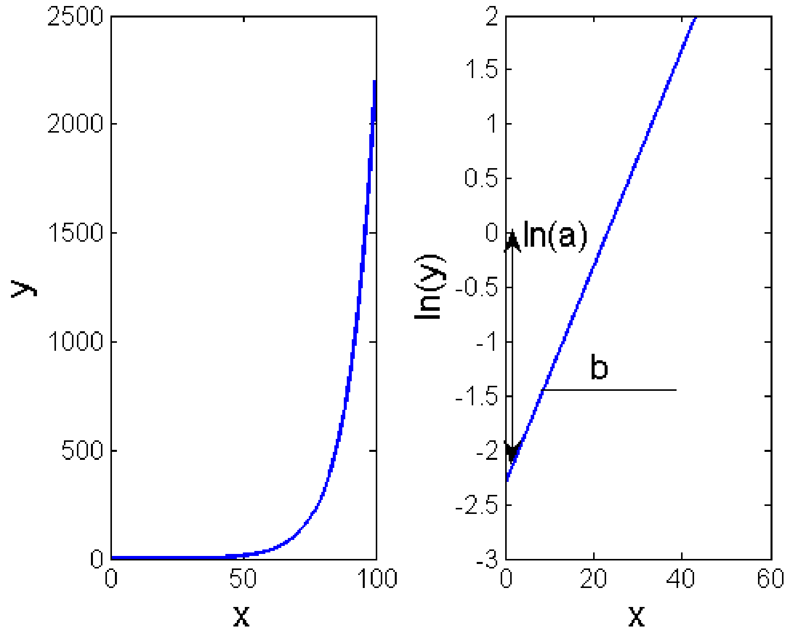


Regressão linear

- **Exemplos**
  - Regressão (Método dos Mínimos Quadrados)

    O método pode ser aplicado para a regressão de diversas funções, para isso é necessário somente encontrar uma escala adequada para os dados.

    Ex.: $y = ae^{bx}$

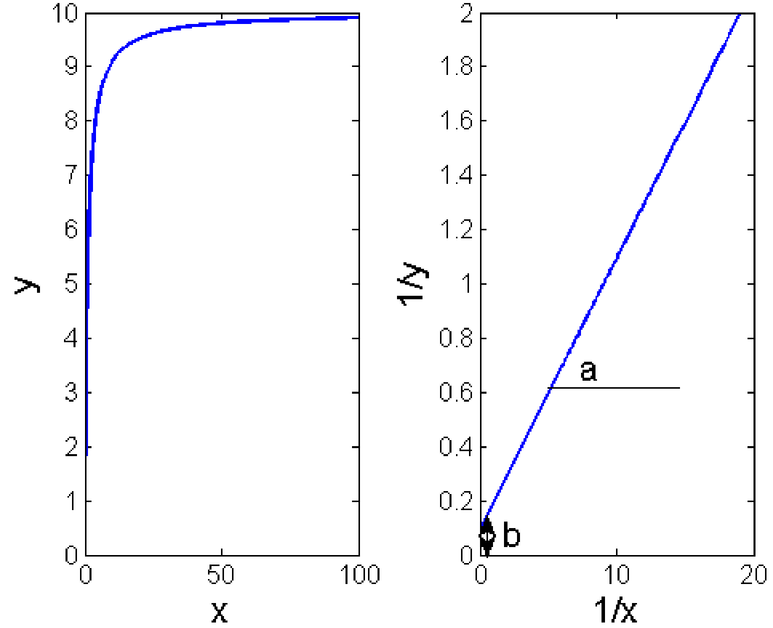


Regressão linear

- **Exemplos**
  - Regressão (Método dos Mínimos Quadrados)

    O método pode ser aplicado para a regressão de diversas funções, para isso é necessário somente encontrar uma escala adequada para os dados.

    Ex.: $y = \dfrac{x}{a + bx}$

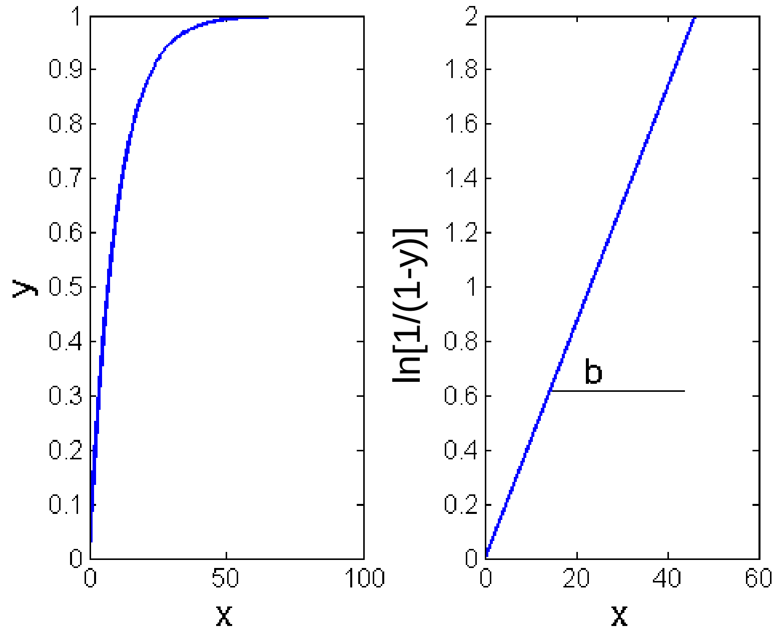


Regressão linear

- **Exemplos**
  - Regressão (Método dos Mínimos Quadrados)

    O método pode ser aplicado para a regressão de diversas funções, para isso é necessário somente encontrar uma escala adequada para os dados.

    Ex.: $y = 1 - e^{-bx}$

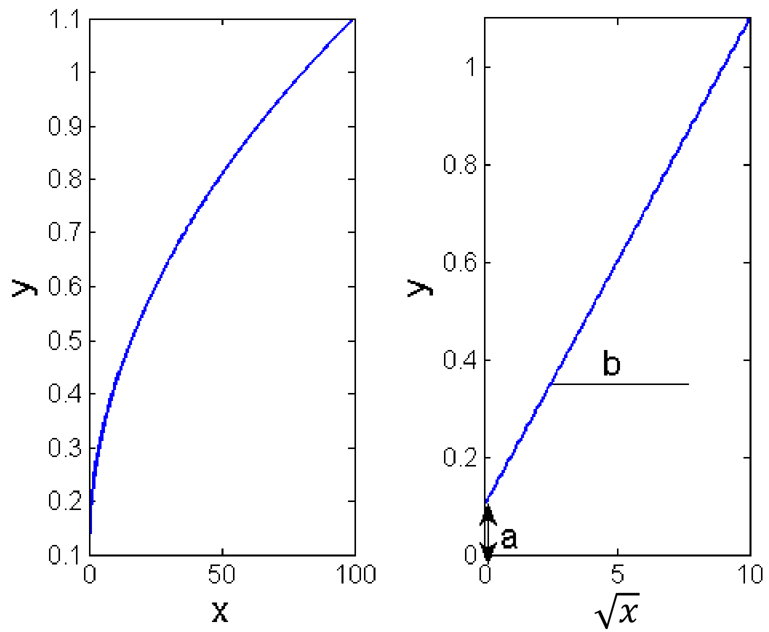


Regressão linear

- **Exemplos**
  - Regressão (Método dos Mínimos Quadrados)

    O método pode ser aplicado para a regressão de diversas funções, para isso é necessário somente encontrar uma escala adequada para os dados.

    Ex. $y = a + b\sqrt{x}$



Exercícios em sala que devem ser finalizados e entregues na aula seguinte

Os dados para os exercícios estão disponíveis no site do curso

- Questão 1) Construir PDF's a partir de séries de dados (1D e 2D);
- Guia para demais tarefas:

Acessar os dados do arquivo fornecido. Para isso, o arquivo 'exercício.mat' deve ser copiado para o diretório corrente do matlab. Para carregar os dados digite no workspace o comando: load('exercicio')

____

Mapa de variáveis

time = vetor de tempo dos sinais (cada posição do vetor indica o instante em que o dado foi amostrado)

- signal = 4 diferentes tipos de sinais para serem classificados como: determinísticos periódicos ou transientes, aleatórios estacionários ergódicos, aleatórios estacionários não-ergódicos, e aleatórios não estacionários).
- signal_seq = sequência de 100 formas de onda para serem analisadas
- ref_signal = sinal de referência para comparação com a sequência de sinais amostrados
- x_1D, y_1D = dados para ajuste de uma curva de uma calibração (1D e 2D)
  x1_2D,x1_2D e y_2D

Guia para o trabalho

- Questão 2) Mostrar gráfico dos sinais salvos na variável signal e classifica-los (ex. para visualizar o 1º. sinal: plot(time,signal(:,1))

- Questão 3) Encontrar dentre as 100 formas de onda contidas na variável signal_seq qual se correlaciona mais com o sinal de referência (ref_signal).

Obs: Os dados de cada uma das 100 formas de onda estão salvos em uma coluna da variável signal_seq. (ex. para acessar os dados da 1ª forma de onda: signal_seq(:,1)).

- Questão 4) Fazer ajuste dos dados usando mínimos quadrados

- a) Ajustar os dados x_1D e y_1D a uma função polinomial de 1ª ordem. Mostrar valor dos coeficientes linear a angular do ajuste.

- b) Ajustar os dados x2_1D e y2_1D a uma função do tipo: y=x/(a+b*x).

- c) Ajustar os dados x1_2D, x2_2D e y_2D a uma função do tipo: y=a+b*x1+c*x1*x2+d*x3