Fundamentos Estatísticos para Análise Estática de Instrumentos – Parte 1

Prof. Jean Tavares

19/10/2020

Universidade Federal de Uberlândia

SUMÁRIO

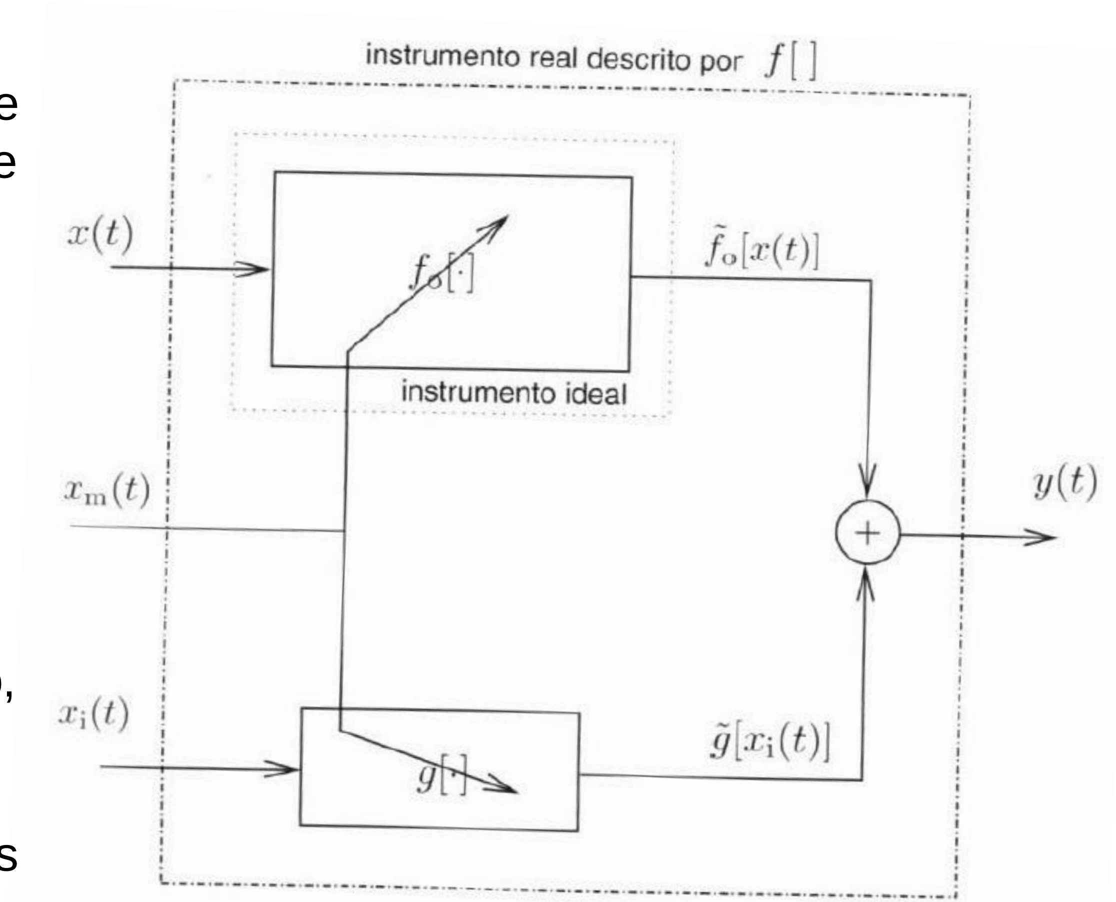
- Introdução
- Objetivo de um Sistema de Medição
- Requisitos para o Processo Estatístico
  - Fundamentos de Estatística
  - Matlab/ Octave
  - Histograma
  - Intervalo de Confiança
  - Critério de Chauvenet
  - Teste Qui-quadrado
  - Análise de Regressão
  - Propagação de Erro

Introdução: Fontes de Erros no Sistema de Medição

A saída do instrumento, *y(t)* pode ser vista como uma série de entradas.

3 Princiapais Tipos de Entrada:
Entrada Desejada (**ED**): *x(t)*
Entrada Modificante (**EM**): $x_m(t)$
Entrada de Interferência (**EI**): $x_i(t)$

Exemplo de EM: Temperatura, Umidade, Pressão, Vibração, Campos eletromagnéticos.
Exemplo de EI: Atrito, campos eletromagnéticos, tensões termoelétricas e galvanométricas



Introdução: Tipos de Incertezas Estáticas

Caracterização Estática de Instrumentos:

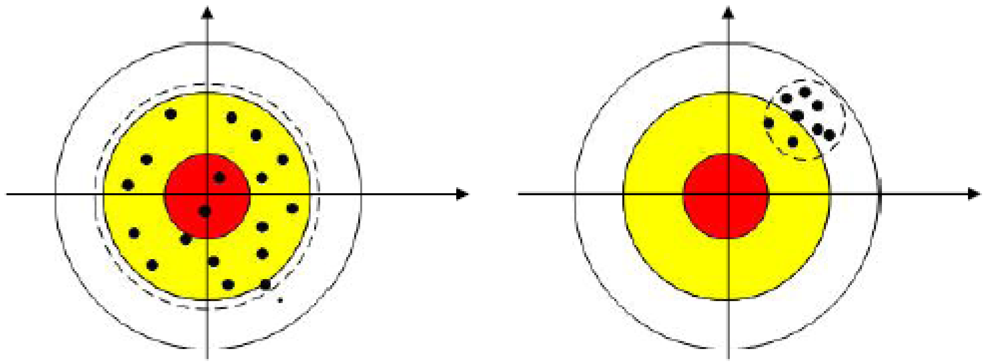
- Desempenho em estado estacionário
- Relação entre a Grandeza a ser Medida e o Sinal de Saída *y(t)* do Sensor
- Incertezas inerentes ao resultado:
  - Efeitos sistemáticos
  - Efeitos aleatórios

Introdução: Tipos de Incertezas Estáticas

CARACTERÍSTICAS DOS ERROS:

SISTEMÁTICO (ou BIAS): a **média** do estimador é diferente do **valor VERDADEIRO da variável → b**

ALEATÓRIO (ou RANDOM): O **valor do estimador** apresenta **DISPERSÃO** em torno do **valor médio. → Var**



BIAS PEQUENO
ALEATÓRIO GRANDE

BIAS GRANDE
ALEATÓRIO PEQUENO

Objetivo do Sistema de Medição

O objetivo de uma medição é determinar os valores do mesurando, ou seja, da grandeza a ser medida. Isso inclui um procedimento de medição e o resultado é apenas uma estimativa do valor do mesurando e, portanto, o resultado da medição deve incluir uma declaração da incerteza associada.

Sistema de Medição (SM) → **Medidas** → **Processo Estatístico** → **Resultado Estatístico**

Sistema de Medição (SM)

Instrumentação Física
- Analógico
- Digital

Interface com Sistema Real

Medidas

Sistema de coleta de dados
Amostragem:
- Finita
- Discreta

Processo Estatístico

Sistema Computacional
- Função Densidade de Probabilidade
- Valores espúrios
- Testes de Aderência

Resultado Estatístico

- Média
- Desvio
- Variância
- Amplitude
- **Intervalo de Confiança ou Erro**
- Representa o sistema real*

Requisitos Para Processo Estatístico dos Dados

- Fundamentos de Estatística
- Matlab/ Octave
- Histograma
- Intervalo de Confiança
- Critério de Chauvenet
- Teste Qui-quadrado
- Análise de Regressão
- Propagação de Erro

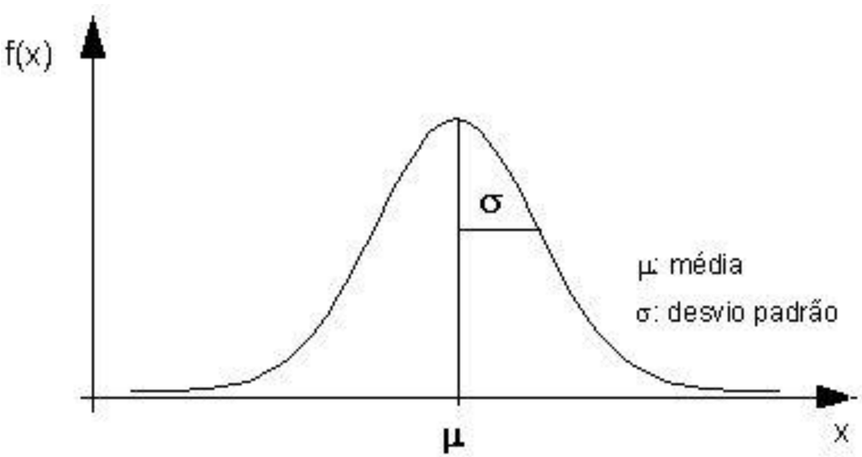
Requisitos: Fundamentos de Estatística

- Uma característica desejável do instrumento é não ser tendencioso, ou seja, não polarizado e sem viés.
- Um instrumento não tendencioso ou não polarizado tem a mesma probabilidade de indicar valores inferiores ou superiores.
- Matematicamente, a Função Densidade de Probabilidade (FDP) de um instrumento não tendencioso pode ser escrita como:

$$P(y < y_0) = \int_{-\infty}^{y_0} f_{\mathrm{p}}(y)dy = \int_{y_0}^{\infty} f_{\mathrm{p}}(y)dy = P(y_0 < y).$$

onde y_0 é o valor médio das medidas.

Requisitos: Fundamentos de Estatística



FUNÇÃO de DENSIDADE de PROBABILIDADE:

$$p(x) = \frac{dP(x)}{dx} \quad \leftrightarrow \quad P(x) = \int_{-\infty}^{x} p(x)dx$$

$$p(x) \in [0 \quad \infty] \quad e \quad \int_{-\infty}^{\infty} p(x)dx = 1$$

A Função Densidade de Probabilidade informa a probabilidade da variável *x* assumir um valor dentro de um determinado intervalo.

Variáveis Aleatórias:

1) Discretas. Exemplo: resultado de um dado;
2) Contínuas. Exemplo: temperatura de um motor.

FUNÇÃO de DISTRIBUIÇÃO de PROBABILIDADE:

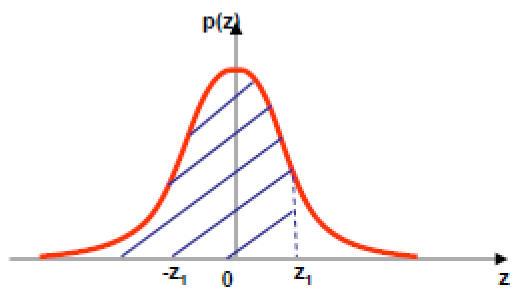
$$P(x) = PROB(x \leq X)$$

$$se \quad a \leq b \quad \Rightarrow \quad P(a) \leq P(b)$$

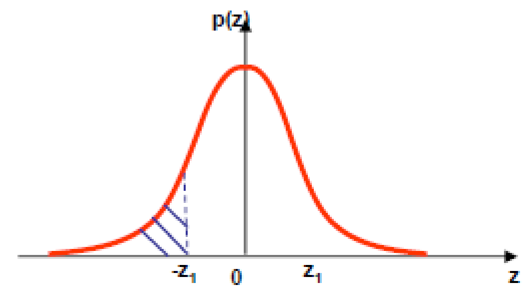
$$P(-\infty) = 0 \quad e \quad P(\infty) = 1$$

Requisitos:
Fundamentos de Estatística

Considerando uma fdp GAUSSIANA de MÉDIA NULA ➔ fdp NORMAL

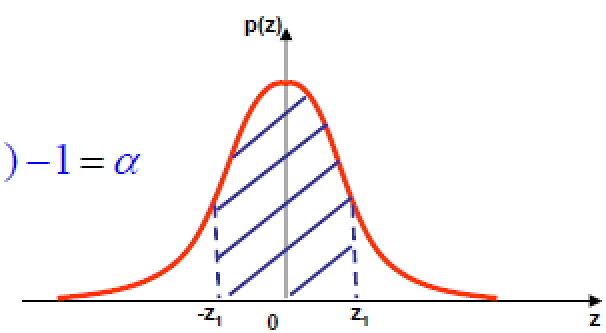


$$P[z \le z_1] = \int_{-\infty}^{z_1} p(z)\,dz = \phi(z_1)$$



$$P[z \le -z_1] = \int_{-\infty}^{-z_1} p(z)\,dz = \int_{z_1}^{\infty} p(z)\,dz = 1-\phi(z_1)$$

$$P[-z_1 \ge z \ge z_1] = \phi(z_1) - 1 + \phi(z_1) = 2\phi(z_1) - 1 = \alpha$$



Req

Fund

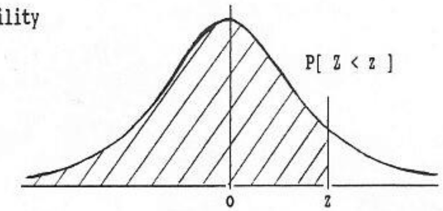
stica

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z i.e.

$$P[\,Z < z\,] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\, dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

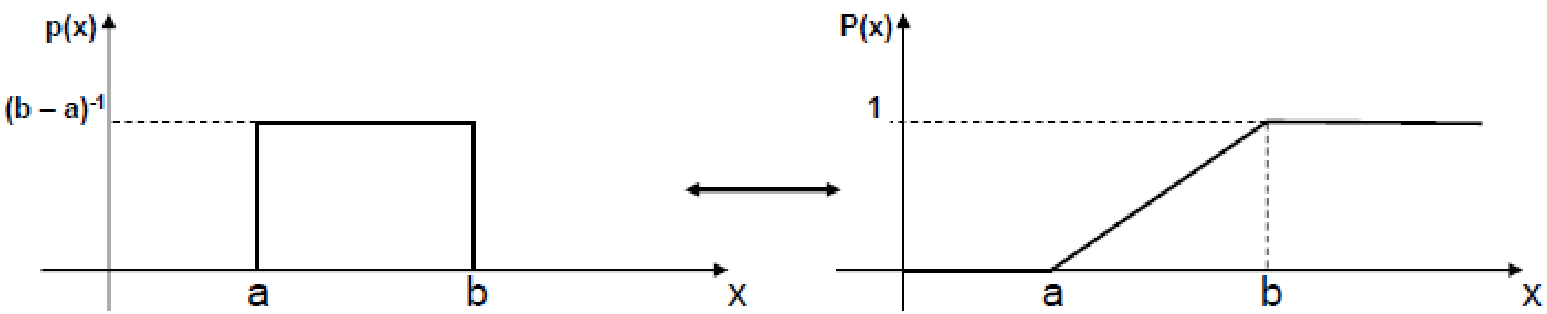
z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Requisitos:
Fundamentos de Estatística

Há outros tipos de Função de Distribuição de Probabilidade. Por exemplo a FDP Uniforme

$$p(x) = \begin{cases} (b - a)^{-1} & \text{para } a \leq x \leq b \\ 0 & \text{para } x \text{ fora} \end{cases}$$

$$P(x) = \begin{cases} 0 & \text{para } x < a; \\ (x - a)/(b - a) & \text{para } a \leq x \leq b; \\ 1 & \text{para } x > b; \end{cases}$$



Requisitos: Fundamentos de Estatística

- Em instrumentação, as duas características mencionadas no contexto de funções de densidade de probabilidade *p(x)* e de distribuição de probabilidade *P(x)* estão relacionadas às características de exatidão e precisão de instrumentos.
  - **Precisão** significa a aptidão de um instrumento de medição fornecer indicações muito próximas, quando se mede o mesmo mensurando, sob as mesmas condições (Aleatório). Por exemplo, quanto menos dispersa for a função (menor σ), mais preciso será o instrumento em questão.
  - **Exatidão** é a aptidão de um instrumento para dar respostas próximas ao valor verdadeiro do mensurando (Bias). Por exemplo, quanto mais próximo a média dos valores coletados de um valor desejado, mais exato é o instrumento.

Requisitos:
Matlab/ Octave

PROPRIEDADE	FUNÇÕES do MATLAB	FÓRMULAS
- MÍNIMO DA AMOSTRA:	min()	$L = \min(u_1\ u_2\ u_3\ \ldots.u_n)$
- MÁXIMO DA AMOSTRA:	max()	$U = \max(u_1\ u_2\ u_3\ \ldots.u_n)$
- AMPLITUDE DA AMOSTRA:	range()	$R = U - L$
- MEDIANA DA AMOSTRA:	median()	$MD=(\ u_{n/2-1}+u_{n/2+1}\ )/2$ n-par
(com valores ordenados)		$MD=\ u_{n/2}$ n-impar
- MÉDIA DA AMOSTRA:	mean()	$\bar{u}$ = SOMA(ui)/n
- VARIÂNCIA DA AMOSTRA:	var()	σ^2 = SOMA(ui – $\bar{u}$)^2 /(n-1)
- DESVIO PADRÃO:	std()	σ = SQRT($\sigma^{2)}$
- SIMETRIA DA AMOSTRA:	skewness()	S_K = SOMA(ui – $\bar{u}$)^3 /(n σ^3)
- AGUDEZA DA AMOSTRA:	kurtosis()	K_T = SOMA(ui – $\bar{u}$)^4 /(n σ^4)

Universidade Federal de Uberlândia

Requisitos: Matlab

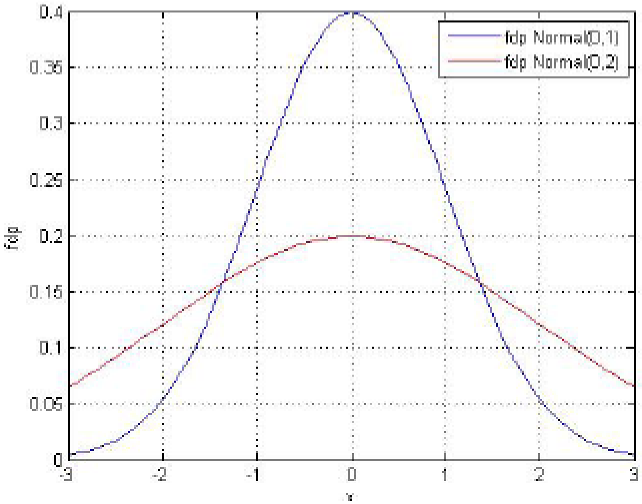
disttool: visualizar as fdp disponíveis e a influência dos parâmetros característicos.

xxxpdf e xxxcdf: funções para gerar as fdp e as acumuladas (xxx indica o tipo de distribuição - help pdf lista os nomes das fdp disponíveis)

xxxinv: inversa das fdp, sendo xxx o tipo de distribuição. Estas funções são úteis para o cálculo de Intervalos de Confiança e Testes de Hipóteses.

EXEMPLO: gerar fdp GAUSSIANA de média NULA e DESVIOS PADRÃO 1 e 2

```
x = [-3 : 0.1 :3];           %um vetor de dados
f1 = normpdf (x, 0, 1);      %gera fdp Normal
f2 = normpdf (x, 0, 2);      %gera fdp Normal
plot (x, f1, x, f2)          % gráficos das fdp
```



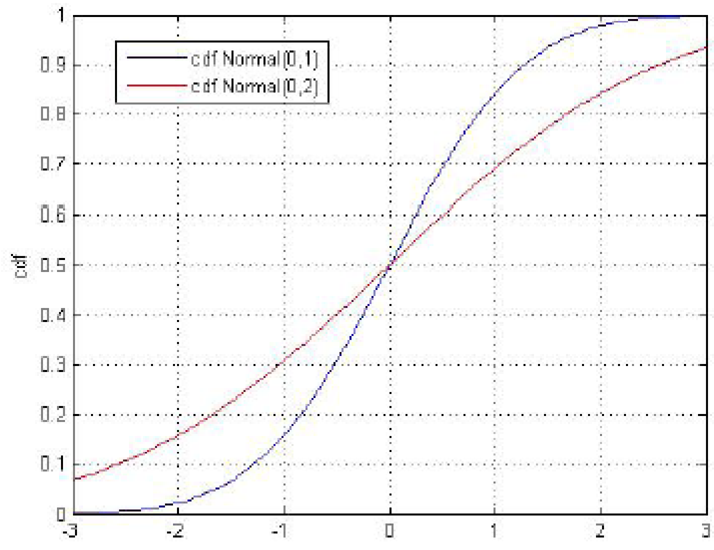
Requisitos:
Matlab

A PROBABILIDADE ACUMULADA:

cdf1 = normcdf (x,0,1);
$\mu = 0$ e $\sigma = 1$

cdf2 = normcdf (x,0,2);
$\mu = 0$ e $\sigma = 2$

plot (x, cdf1, x, cdf2)   % gráficos das ACUMULADAS



Requisitos:
Matlab

FUNÇÃO INVERSA da PROBABILIDADE

Determinar o INTERVALO [x] com probabilidade α:

x = xxxinv (α, média, desvio padrão)

Seja P = [0.005 0.995]; um vetor de probabilidades entre 0.005 e 0.995, ou seja um FAIXA de PROBABILIDADE α = 0.995 – 0.005 ➔ 99 %

x = norminv (P, 0, 1); gera o INTERVALO x com fdp NORMAL de média 0 e desvio padrão 1, com 99% de probabilidade (CONFIANÇA)

RESULTA o INTERVALO: x = [-2.5758 2.5758]

Universidade Federal de Uberlândia

Requisitos: Histograma

HISTOGRAMA (ESTIMADOR da fdp)

Seja u = { u_1 u_2 ….u_N } uma amostra **FINITA e DISCRETA** das medidas obtidas num experimento.

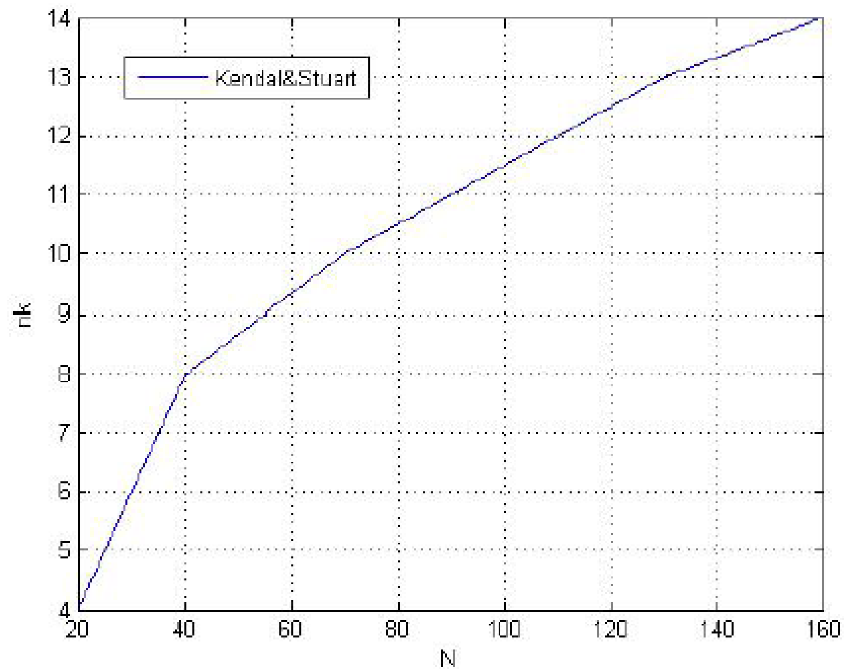
O HISTOGRAMA da amostra pode ser construído para **ESTIMAR** a **fdp das medidas u, de acordo com as seguintes etapas:**

1. **Ordenar u e determinar a amplitude da amostra (R).**
2. **Agrupar as medidas em k CLASSES (bins) de amplitudes calculadas por um critério estatístico. (ver slide seguinte)**
3. **Determinar o número (n_{ok}) de ocorrências das medidas na CLASSE k.**
4. **Calcular a frequência relativa em cada CLASSE: $f_k = n_{ok}/N$.**
5. **Calcular a frequência acumulada f_a até cada CLASSE k.**
6. **Fazer os gráficos: $n_{ok} \times k$, $f_k \times k$, $f_a \times k$.**
7. **Testar a fdp teórica que melhor se ajusta ao histograma.**

Requisitos: Histograma

HISTOGRAMA: CONSIDERAÇÕES SOBRE O NÚMERO DE CLASSES (n_k):

1. **Se N ≤ 20** ➔ $n_k = 5$
2. **Se 20 < N < 40** ➔ número de OCORRÊNCIAS em cada classe ≥ 5
3. **Se N > 40** ➔ critério de KENDAL & STUART: $n_k = 1.87*(N-1)^{0.4}$



Requisitos: Histograma

HISTOGRAMA – EXEMPLO: Foram realizadas 20 testes experimentais para medir a tensão de escoamento de um aço, resultando: S_e [MPa] (Na tabela abaixo os dados já estão em ordem crescente):

teste	S_e	teste	S_e	teste	S_e	teste	S_e	teste	S_e
1	**448**	5	**510**	9	**539**	13	**551**	17	**579**
2	**471**	6	**519**	10	**543**	14	**554**	18	**590**
3	**498**	7	**530**	11	**545**	15	**559**	19	**600**
4	**507**	8	**536**	12	**546**	16	**570**	20	**619**

N = 20 **LIMITES: L = 448 e U = 619**

RANGE: R = U – L = 171 **MEDIANA: MD = 533.5**

MÉDIA: mean = 540.7 **DESVIO PADRÃO: Std = 41.7**

AMPLITUDE das CLASSES: para n_k = 5 → A_k = R/5 = 34.2;

Requisitos: Histograma

Organizando as CLASSES: N = 20 ➔ n_k = 5

CLASSE	A_k	n_{ok}	$f_k = n_{ok}/20$	f_a
1	448 - 482	2	2/20	2/20
2	482 - 516	3	3/20	5/20
3	516 - 550	7	7/20	12/20
4	550 - 584	5	5/20	17/20
5	584 - 619	3	3/20	20/20

USO do MATLAB:

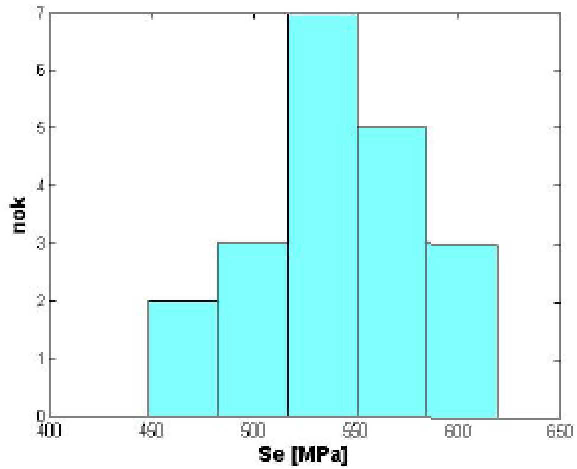
se = [se_1 se_2 ......se_{20}]; % vetor das medidas.

% A função **hist (se, nk)** gera o histograma com as medidas contidas no vetor **se** com **nk** classes.

hist (se,5)

% gera o histograma com 5 classes

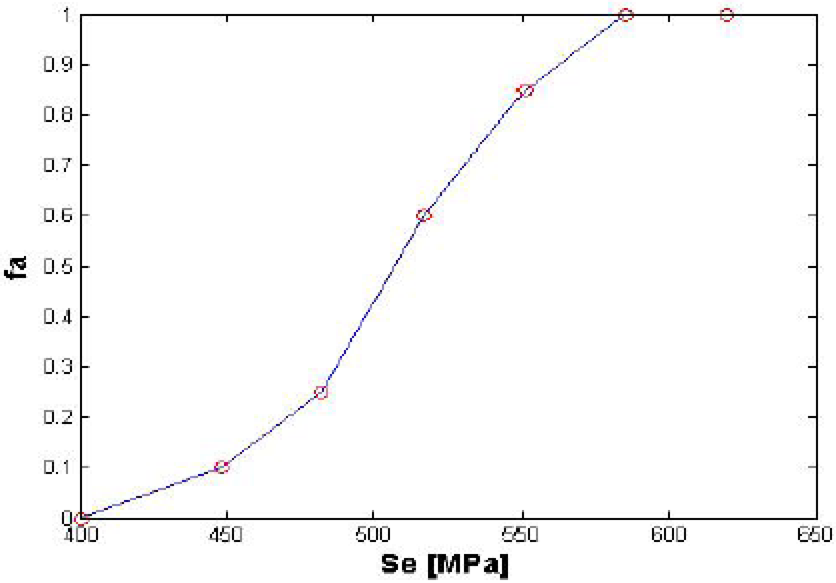
% e plota (se, nok)



Requisitos: Histograma

Uso do Matlab para criar o HISTOGRAMA: Frequência ACUMULADA

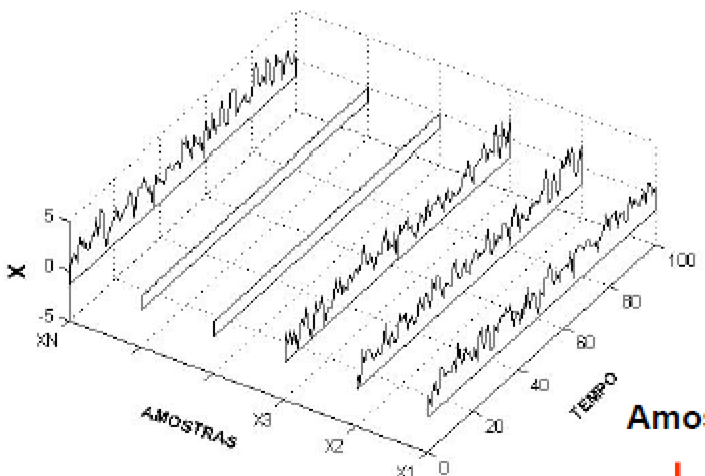
```
% com nk = 5 (cinco classes)
[nok, x] = hist (se,5) % retorna nok e x = ponto médio da classe
na(1)=0;
for j=1 : 5,          % cria nac
na(j+1)=na(j)+nok(j);
end
fa = na/20;           % frequência acumulada
plot (se, fa, x, fa, ' or ')   % plota
```



A frequência acumulada é usada para realizar o **TESTE de HIPÓTESE** sobre a fdp assumida para os dados

Requisitos:
Intervalo de Confiança

Um processo de medição gera VÁRIAS AMOSTRAS para a variável (x):



CADA AMOSTRA permite calcular estimadores estatísticos para (x)

POR EXEMPLO :

Amostras: $\{x_1\}, \{x_2\}, \{x_3\}$ ..... $\{x_N\}$

Médias amostrais: $\overline{x}_1, \overline{x}_2, \overline{x}_3$ ..... $\overline{x}_N$

Desvios Padrão amostrais: s_1, s_2, s_3 ..... s_N

Requisitos: Intervalo de Confiança

<u>QUESTÃO:</u> a MÉDIA das MÉDIAS AMOSTRAIS é um bom estimador da média de x?

<u>CASO 1</u> : Se x é uma variável COM fdp DESCONHECIDA tal que E[x] = μ e VAR[x] = σ^2 são CONHECIDAS. Colhidas amostras discretas e finitas {x} de tamanho N, e calculados os estimadores das médias amostrais $\bar{x}$ tem-se:

$$se\, N\, é\, grande \quad (N > 20):$$

$$E[\bar{x}] = \mu \quad e \quad VAR[\bar{x}] = \frac{\sigma^2}{N}$$

$$Como \quad \mu \quad e \quad \sigma \quad são\ ctes \quad \Rightarrow \quad \bar{x}\ tem\ fdp\ GAUSSIANA.$$

$$Neste\, caso: \quad z = \frac{(\bar{x} - \mu)}{\left(\frac{\sigma}{\sqrt{N}}\right)} \quad tem\ fdp\ Gaussiana\ NORMAL.$$

Requisitos: Intervalo de Confiança

Este fato permite calcular o INTERVALO (D) de $\bar{x}$ que contém μ com um determinado nível de probabilidade (CONFIANÇA = α)

$$P(-z_1 \le z \le z_1) = \alpha = 2\,\phi(z_1) - 1 \quad como: z = \frac{\sqrt{N}(\bar{x}-\mu)}{\sigma}$$

$$\alpha = P\left(\bar{x} - \frac{z\sigma}{\sqrt{N}} \le \quad \mu \quad \le \bar{x} + \frac{z\sigma}{\sqrt{N}}\right) \quad \rightarrow \quad \boxed{D = \frac{z\sigma}{\sqrt{N}}}$$

EXEMPLO:

Um componente elétrico tem duração (vida) cujo valor tem desvio padrão σ = 4 horas. Montou-se uma amostra com N = 100 componentes que foram ensaiados para determinar suas durações. A média amostral da vida do componente resultou $\bar{x}$ = 501.2 horas.

Qual o INTERVALO de CONFIANÇA para a MÉDIA com confiança 95%?

Requisitos:
Intervalo de Confiança

<u>CASO 2:</u> NÃO SE CONHEÇE PREVIAMENTE o valor do DESVIO PADRÃO ➔ ele pode ser estimado pela VARIÂNCIA AMOSTRAL:

$$\boxed{S^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2 \neq \sigma^2}$$

A MÉDIA AMOSTRAL NÃO tem fdp GAUSSIANA, mesmo se N é grande. MAS tem fdp t-Student:

$$t = \frac{\bar{x} - \mu}{S/\sqrt{N}} \quad \text{tem fdp} \quad t - \text{student}$$

$$\alpha = P\left(\bar{x} - t_{d,\frac{1+\alpha}{2}}\frac{S}{\sqrt{N}} \le \mu \le \bar{x} + t_{d,\frac{1+\alpha}{2}}\frac{S}{\sqrt{N}}\right)$$

$$\rightarrow \quad D = t_{d,\frac{1+\alpha}{2}}\frac{S}{\sqrt{N}} \quad \text{é o DESVIO de } \mu.$$

$$\text{sendo} \quad d = N - 1 \quad e \quad \alpha \text{ a Confiança}$$

$$\boxed{\text{dados}: \quad N \text{ e } \alpha \quad \rightarrow \quad D}$$

$$\boxed{\text{dados}: \quad D \text{ e } \alpha \quad \rightarrow \quad N}$$

Requisitos: Intervalo de Confiança

EXEMPLO: **Foram realizadas 10 medidas da resistência elétrica de um componente elétrico resultando: Média amostral = 10.48 Ω e Desvio Padrão Amostral = 1.36 Ω. (o desvio padrão é desconhecido)**

Qual é o intervalo de confiança para a MÉDIA com probabilidade de α = 90%?

$(1 + \alpha)/2 = 0.95$, $d = N - 1 = 9$ ➔ TABELA ou tinv((1+ α)/2 ,d)

Resulta: $t_{9,0.95} = 1.833$

Requisitos:
Intervalo de Confiança

EXEMPLO: Foram realizadas 10 medidas da resistência elétrica de um componente elétrico resultando: Média amostral = 10.48 Ω e Desvio Padrão Amostral = 1.36 Ω. (o desvio padrão é desconhecido)

Qual é o intervalo de confiança para a MÉDIA com probabilidade de $\alpha = 90\%$?

$(1 + \alpha)/2 = 0.95$, $d = N - 1 = 9$ ➔ TABELA ou tinv((1+ α)/2 ,d)

Resulta: $t_{9,0.95} = 1.833$

Logo: $D = 1.833 \times 1.36 / 10^{1/2} = 0.79\ \Omega$

O intervalo será: $9.69 \leq \mu \leq 11.27$ com 90% de confiança

Universidade Federal de Uberlândia

Requisitos:
Critério de Chauvenet

OBJETIVO: REMOVER da amostra os valores que tenham dispersão em relação à MÉDIA superior a um VALOR PADRÃO.

Seja a amostra $\{x\} = \{x_1\ x_2\ x_3\ ....x_N\}$ com média $\overline{x}$ e desvio padrão S_x

Define-se o desvio relativo à média: $DR(x_i) = |\ x_i - \overline{x}\ |\ /\ S_x$.

Se $DR(x_i) > DRo$ ➔ x_i é REMOVIDO da amostra, e

RECALCULAM-SE os novos valores da Média e do Desvio amostrais.

DRo é a Taxa PADRÃO de desvio em relação à Média Amostral, definida por CHAUVENET , em função do tamanho da amostra (N).

HIPÓTESE: DRo tem fdp NORMAL

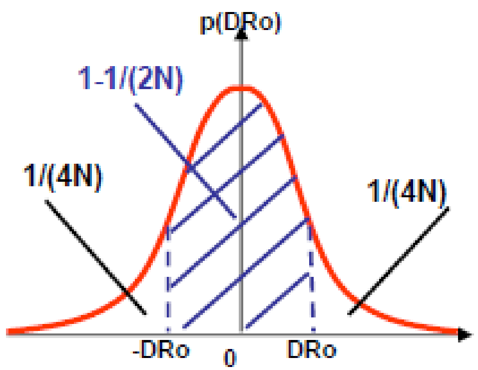
O critério de CHAUVENET pode ser aplicado apenas UMA vez.

Requisitos:
Critério de Chauvenet

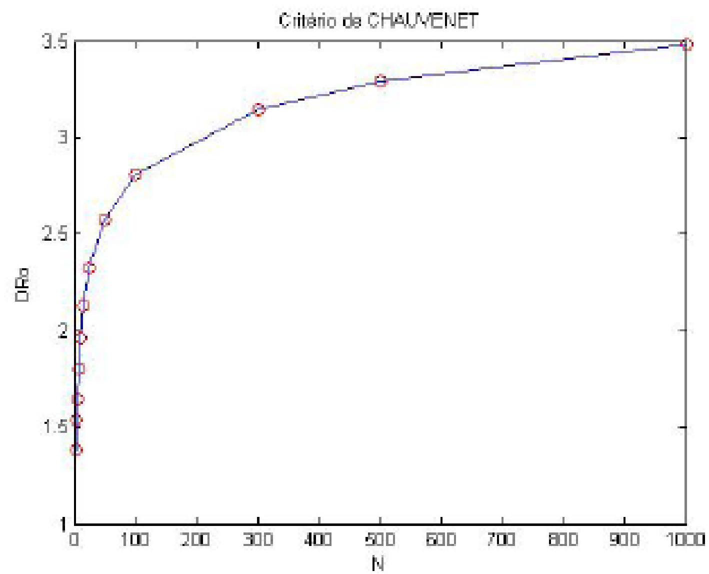
Usando o Matlab:

p(Dro) é Normal

Dro(N) = norminv (1-1/(4*N), 0, 1)



N	DRo
3	1.383
4	1.534
5	1.645
7	1.803
10	1.960
15	2.128
25	2.326
50	2.576
100	2.807
300	3.144
500	3.291
1000	3.481



Requisitos:
Teste Qui-quadrado

OBJETIVO: Verificar a HIPÓTESE feita sobre a fdp ADOTADA para uma variável, a partir de uma amostra finita e discreta (tamanho N).

TÉCNICA: Construir o histograma com n_k CLASSES e comparar o número de ocorrências OBSERVADO com aquele OBTIDO com a fdp ASSUMIDA. A variável X^2 definida abaixo tem fdp QUI-QUADRADO:

$$X^2 = \sum_{k=1}^{n_k} \frac{(n_o - n_e)_k^2}{n_{ek}}$$

n_o – número **OBSERVADO** de ocorrências na **CLASSE k**

n_e – número de ocorrências **ESPERADO** na **CLASSE k**, para a **fdp** sob teste.

Requisitos:
Teste Qui-quadrado

X_n^2 – variável com fdp QUI-QUADRADO

n - número de graus de liberdade estatísticos

$$n = n_k - 1 - n_p$$

n_k – número de Classes
n_p – número de parâmetros da fdp testada

Dados X^2 e n ➔ TABELA ➔ α = Confiança do Teste de Hipótese

Usando o Matlab: β = chi2cdf(X^2, n) ➔ $\alpha = 1 - \beta$ = CONFIANÇA

Requisitos:
Teste Qui-quadrado

Verificar se {x} tem fdp GAUSSIANA (μ, σ ➔ n_p = 2), e calcular a CONFIANÇA (α)

```
x=[10.02 10.2 10.26 10.2 10.22 10.13 9.97 10.12 10.09 9.9...
10.05 10.17 10.42 10.21 10.23 10.11 9.98 10.1 10.04 9.81];

N= length (x);        % tamanho da amostra = 20
nk= 6;                % número de classes
xs= sort (x);         % ordenando x
R=xs(20) - xs(1);     % amplitude da amostra
ak= R/nk;             % amplitude da classe
xb= mean (x);         % media amostral
sx= std (x);          % desvio padrão amostral
hist (x, nk)          % cálculo do Histograma com nk classes
no=hist(x, nk);       % número observado de ocorrências
lik= zeros(1,nk);     % limites das classes (aloca espaço)
for j=1:nk,
lik (j)=xs(1) + j*ak;
end
ne=zeros(1,nk);   % número esperado de ocorrências
ne(1)=normcdf (lik (1), xb, sx)*N;
for j=2:nk,
   ne(j)=(normcdf(lik(j), xb, sigx)-normcdf(lik(j-1), xb, sigx))*N;
end
```

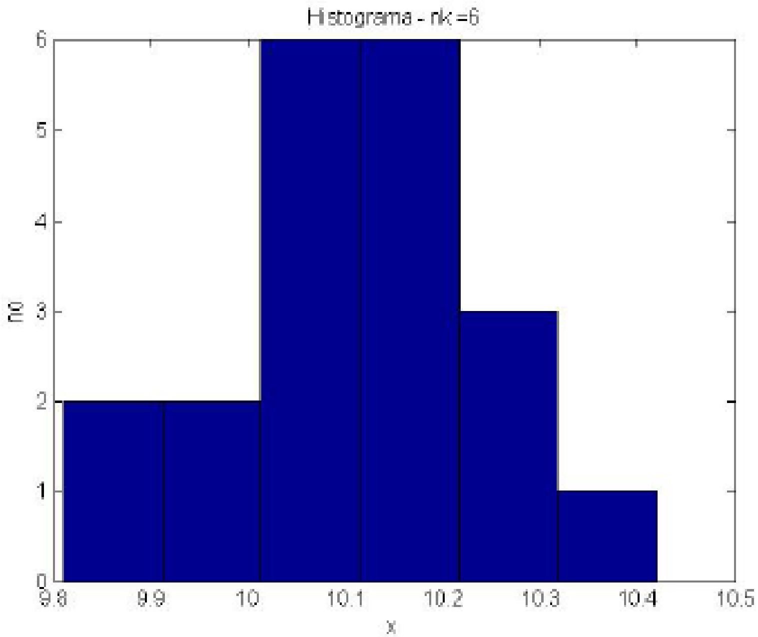
N = 20
n_k = 6

a_k = 0.1017
xb = 10.112
sigx = 0.138

Requisitos:
Teste Qui-quadrado

VERIFICAR se {x} tem fdp GAUSSIANA (μ, σ → n_p = 2), e calcular a CONFIANÇA (α)

CLASSE	n_o	n_e	$(n_o - n_e)^2/n_e$
1 (9.8100, 9.9117)	2	1.48	0.1804
2 (9.9117, 10.0133)	2	3.29	0. 5079
3 (10.0133, 10.1150)	6	5.43	0.0607
4 (10.1150, 10.2167)	6	5.33	0.0842
5 (10.2167, 10.3183)	3	3.12	0.0048
6 (10.3183, 10.4200)	1	1.09	0.0074
		$\sum$=	0.8453



Universidade Federal de Uberlândia

Requisitos:
Teste Qui-quadrado

Cálculo da Confiança

```
dr= (no - ne).^2./ne;       % diferenças normalizadas
xi2= sum(dr);               % variável Xi2
np = 2;                     % número de parâmetros da gaussiana
ngl= nk - np - 1;           % graus de liberdade estatístico

alfa=1-chi2cdf(xi2,ngl);    % confiança da hipótese = 0.8386
```

↓

ACEITA-SE a HIPÓTESE que {x} tem fdp GAUSSIANA com CONFIANÇA: alfa = 83.86%

Universidade Federal de Uberlândia

Requisitos:
Teste Qui-quadrado

OBSERVAÇÕES:

1. **O NÍVEL de CONFIANÇA da variável Qui - quadrado pode ser calculada pela função chi2cdf(xi2,ngl) que fornece a probabilidade acumulada P(chi2 < xi2) = 1 – α**
2. **A função [h, P, stats] = chi2gof(x) fornece:**

   - h = 0 ➔ hipótese aceita com 95 % de confiança (=1 rejeitada)

   - P: valor da probabilidade ou confiança do resultado

   - stats: informações adicionais....

Requisitos
Análise de Regressão

OBJETIVO: Criar um MODELO MATEMÁTICO (f) que representa a relação entre pelo menos DOIS CONJUNTOS DE DADOS.

{x} é uma amostra da **ENTRADA** do **SM**

{y} uma amostra da **SAÍDA** do **SM**

DETERMINAR $y = f(x)$, tal que $| y_{medido} - y |$ seja **MÍNIMO.**

FUNÇÕES do MATLAB: ajuste polinomial

p = polyfit (x, y, n);	%retorna os coeficientes do polinômio (p) de grau n ajustado aos dados x e y.
ya = polyval (p, x);	%retorna os valores do polinômio de coeficientes p em cada x.
Dr = y – ya;	%calcula diferença entre os valores ajustados e os medidos.
R = sum (Dr./ya).^2;	%calcula o resíduo quadrático de (y) em relação à função ajustada (ya).

Universidade Federal de Uberlândia

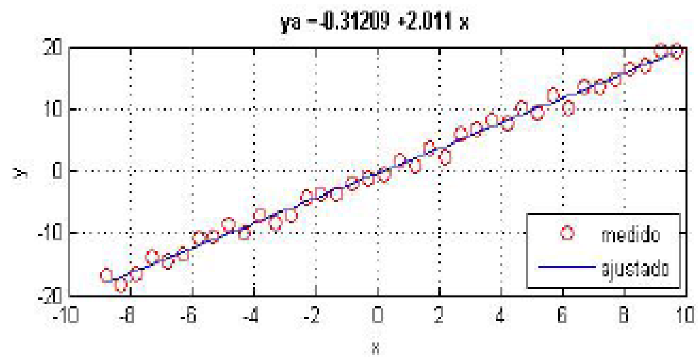
Requisitos
Análise de Regressão

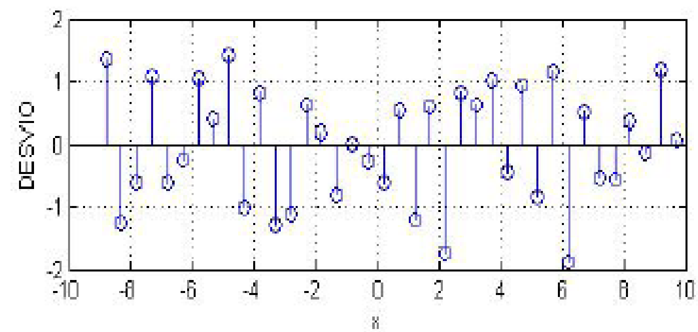
EXEMPLO: Num ensaio de CALIBRAÇÃO ESTÁTICA foi medida a saída {y} do SM para entradas PADRÃO {x}. Neste caso os valores y são contaminados por ruído, porém têm valores crescentes nos limites da faixa de operação do SM (tendência).

```
x = -8.8:0.5:10;            % 38 valores de x
y = 2*x + randn (1,38);     %saída medida
p = polyfit (x, y, 1);      %ajuste linear
ya = polyval (p, x);        %saída ajustada
dr = (y – ya)./ya;          %desvio relativo
```

GRÁFICOS

```
plot (x, ya, x, y, ' or ' )      %ajuste
stem (x, dr)                     %desvio
```





Requisitos
Análise de Regressão

EXEMPLO: Com os mesmos dados do exemplo anterior, Uma forma alternativa é o uso da função regress() do Matlab.

Monta-se uma matriz X = [ones (n,1) x] com a 1ª coluna com n valores 1 e a 2ª coluna com os n valores de x.

[pr ,inter95, resi] = regress (y, X);

RETORNAM:

pr – coeficientes do polinômio (α=95 %)

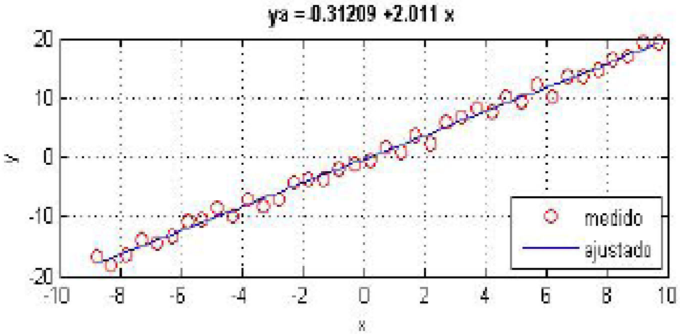
resi – os resíduos do ajuste em cada **x**

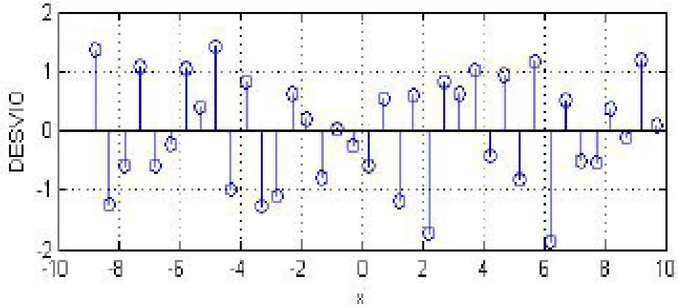
inter95 – os intervalos de confiança de **pr**

Se y = m x +b, com α = 95 %:

b = -0.312 ∈ [-0.6178 -0.0064]

m = 2.011 ∈ [1.9554 2.0665]

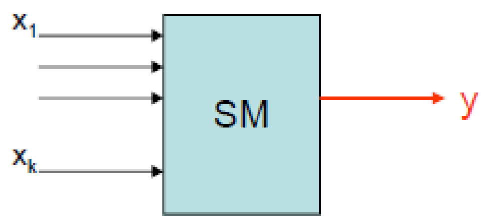




Requisitos
Análise de Regressão

REGRESSÃO MÚLTIPLA LINEAR

A saída {y} do SM é devida a várias entradas SIMULTÂNEAS
$\{x_1\}, \{x_2\}, \ldots, \{x_k\}$: ➔ $y = f(x_1, x_2, \ldots, x_k)$.



OBJETIVO: DETERMINAR A SUPERFÍCIE QUE MELHOR AJUSTA A RESPOSTA DO SM PARA AS ENTRADAS.

METODOLOGIA:

MODELO: $y = a + b_1 x_1 + b_2 x_2 + \ldots + b_k x_k$, sendo $a, b_1, \ldots, b_k$ constantes

FUNCIONAL: $F^2 = [\, y_i - (a + b_1 x_{1i} + b_2 x_{2i} + \ldots + b_k x_{ki})]^2$

Condições de mínimo: Derivadas Parciais de F^2 em relação às constantes:
$a, b_1, \ldots e\ b_k = 0$ ➔ (k+1) equações e (k+1) incógnitas

Requisitos
Análise de Regressão

A mesma função regress (y, X); é usada para calcular o **AJUSTE**, os **ERROS**, os **INTERVALOS** de **CONFIANÇA** e outras informações estatísticas.

A matriz [X] contém na primeira coluna um vetor de n valores 1 e nas demais colunas os vetores das **ENTRADAS** $\{x_1\}$ ....$\{x_k\}$.

{y} é o vetor da **SAÍDA**.

O comando [b, bint, r, rint, stats] = regress (y, X); retorna:

- b – coeficientes a, b1, … bk
- bint – intervalos de confiança (95%) para os coeficientes.
- r – os resíduos do ajuste
- rint – os intervalos de confiança para os coeficientes
- stats – informações estatísticas adicionais

Ver no HELP do toolbox de estatística o item MULTIPLE LINEAR REGRESSION

Requisitos: Propagação de Erro

CASO A - Os **ERROS** de um processo de medição podem estar distribuídos em cada um dos componentes do **SM**: conhecidos o erro no sensor, o erro no condicionador e o erro no indicador, calcular o **ERRO** do **RESULTADO** da **MEDIÇÃO** feita com o **SM**.

CASO B - Várias medições independentes de grandezas físicas são combinadas para calcular o **RESULTADO** de um experimento: medição do diâmetro d com incerteza Δd e medição da altura h com incerteza Δh. Calcular a incerteza ΔV do volume V calculado a partir da expressão $V = \pi d^2 h/4$.

Estes dois problemas podem ser formulados considerando:

a) Se a incerteza de cada componente (ou medida) for conhecida, como calcular a **INCERTEZA GLOBAL** do **SM**?

b) Se a incerteza **TOTAL DESEJADA** for especificada, como determinar os **LIMITES** das **INCERTEZAS** de **CADA** componente do **SM** (ou de cada **MEDIDA**)?

Requisitos: Propagação de Erro

Seja uma grandeza $y = f(x_1, x_2, ..., x_n)$, sendo f uma função <u>conhecida</u> das n variáveis independentes x_i. Os x_i podem ser:

CASO A: medidas nas saídas de diferentes SM, ou

CASO B: saídas intermediárias dos componentes de um SM.

Conhecidos os erros Δx_i ➔ CALCULAR Δy

$$y \pm \Delta y = f(x_1 \pm \Delta x_1, x_2 \pm \Delta x_2, \ldots, x_n \pm \Delta x_n)$$

Expandindo f em uma série de TAYLOR e considerando Δx_i pequenos, de modo que os termos $(\Delta x_i)^n \approx 0$, para $n \geq 2$, RESULTA:

$$f(x_i + \Delta x_i) = f(x_i) + \sum \Delta x_i \frac{\partial f}{\partial x_i} \quad \Rightarrow \quad \Delta y_i = \sum \Delta x_i \frac{\partial f}{\partial x_i}$$

Erro ABSOLUTO

$$\boxed{\Delta y_i = \sum \left| \Delta x_i \frac{\partial f}{\partial x_i} \right|}$$

Erro RELATIVO

$$\boxed{\frac{\Delta y_i}{\bar{y}}}$$

Universidade Federal de Uberlândia

Requisitos:
Propagação de Erro

Se Δy for ESPECIFICADO, pode-se estimar os limites necessários para cada um dos **Δx_i**, considerando que suas **CONTRIBUIÇÕES** para o erro da saída são **IDÊNTICAS:**

$$\boxed{\Delta x_i \approx \frac{\Delta y}{n \dfrac{\partial f}{\partial x_i}}} \qquad \textbf{(1)}$$

Com os valores dos **Δx_{im}** obtidos nas medições das variáveis **x_i** é possível compara–los com os valores calculados por **(1)**.

Caso alguns dos $\Delta x_{im} > \Delta x_i \rightarrow$ REFAZER a MEDIÇÃO destas vaiáveis

Requisitos:
Propagação de Erro

Conhecidas as **MÉDIAS** $\overline{x}_i$ e **DESVIOS PADRÃO** S_{xi} das variáveis $\mathbf{x_i}$, o **DESVIO PADRÃO** S_y , da variável **y,** é calculado conforme as OPERAÇÕES ALGÉBRICAS efetuadas com as $\mathbf{x_i}$**:**

$$ADIÇÃO\ ou\ SUBTRAÇÃO: S_y = \sqrt{S_{x1}^{\ 2} + S_{x2}^{\ 2} + \ldots.. + S_{xn}^{\ 2}}$$

$$MULTIPLICAÇÃO: S_y = \overline{x}_1\overline{x}_2\ldots.\overline{x}_n\sqrt{\frac{S_{x1}^{\ 2}}{\overline{x}_1^2} + \frac{S_{x2}^{\ 2}}{\overline{x}_2^2} + \ldots.. + \frac{S_{xn}^{\ 2}}{\overline{x}_n^2}}$$

$$DIVISÃO \quad y = \frac{x_1}{x_2}: \quad S_y = \frac{\overline{x}_1}{\overline{x}_2}\sqrt{\frac{S_{x1}^{\ 2}}{\overline{x}_1^2} + \frac{S_{x2}^{\ 2}}{\overline{x}_2^2}}$$

$$EXPONENCIAÇÃO \quad y = x_1^k: \quad S = k\,\overline{x}_1^{\,k-1}S_{x1}$$

Universidade Federal de Uberlândia

Requisitos: Propagação de Erro

1. Calcular o volume de um cilindro: $V = \pi\ d^2\ h/4$ conhecendo os DESVIOS PADRÃO e as MÉDIAS das medidas do diâmetro (d) e da altura (h).

Valores ESTIMADOS: média $d_m = 1$, desvio padrão $S_d = 0.005$

média $h_m = 4$, desvio padrão $S_h = 0.040$

Usando os valores médios do diâmetro (d_m) e da altura (h_m) resulta:

$$V_m = \pi\, d_m^2\, h_m / 4 = \pi\, 1^2\, 4/4 = 3.14596$$

O desvio padrão do VOLUME resulta:

$$S_v = d_m^2\, h_m\, [2\, S_d^2 / d_m^2 + S_h^2 / h_m^2]^{1/2} = 0.049$$