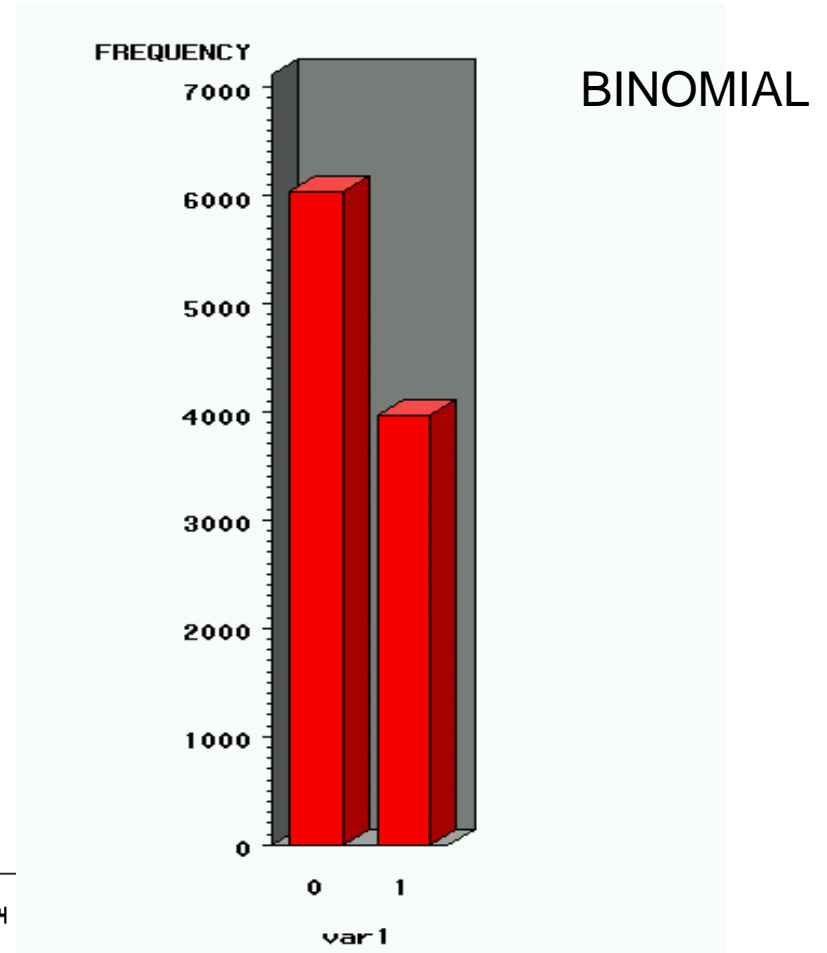
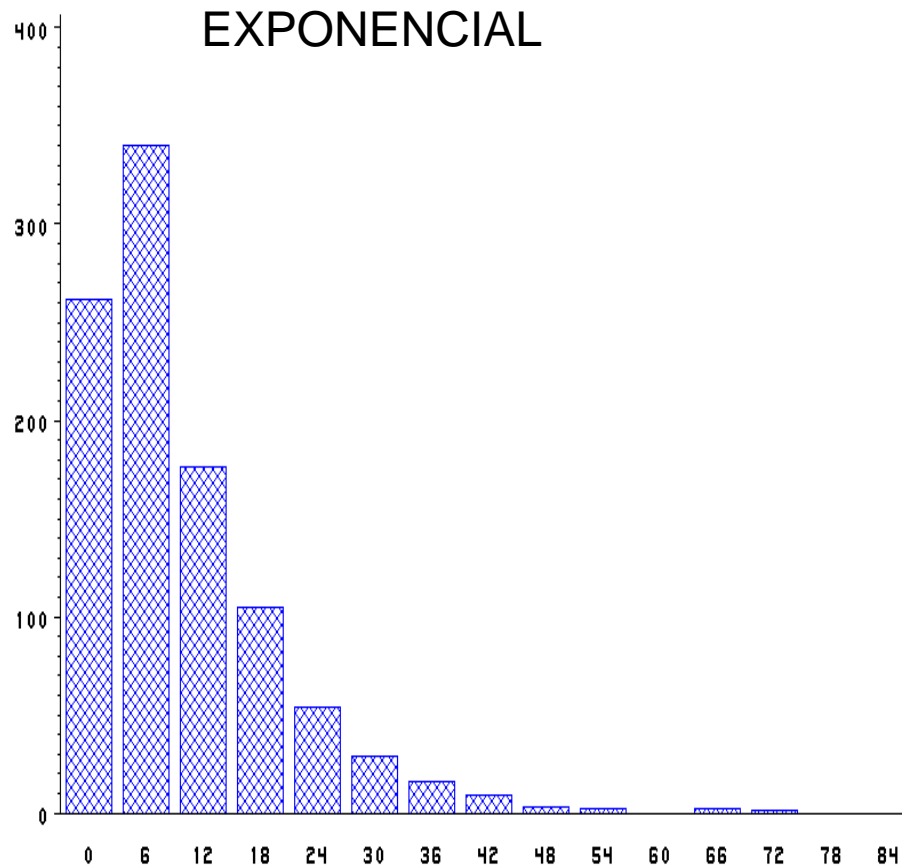

BIOESTATÍSTICA

DISTRIBUIÇÕES PROBABILÍSTICAS

DISTRIBUIÇÃO DE FREQUÊNCIA



DISTRIBUIÇÃO BINOMIAL

- PROVA DE BERNOULLI: SUCESSO OU FALHA
- SOMENTE DOIS POSSÍVEIS RESULTADOS PARA CADA INDIVÍDUO
- POPULAÇÃO DICOTÔMICA: DEFEITUOSA OU SEM DEFEITO, DOENTE OU SADIA, MACHO OU FÊMEA, MORTA OU VIVA, POSITIVA OU NEGATIVA.
- PROBABILIDADE DE SUCESSO: $P(S)=p$;
 $P(F)=1-p = q$

DISTRIBUIÇÃO BINOMIAL



JACOB BERNOULLI

**NASCEU EM 1654 E FALECEU
EM 1705 EM BASILÉIA NA
SUIÇA.**

**FOI PROFESSOR DE
MATEMÁTICA NA
UNIVERSIDADE DE BASILÉIA.
PUBLICOU DIVERSOS
TRABALHOS SOBRE CÁLCULO
DIFERENCIAL, GEOMETRIA E
PROBABILIDADE.**

**PEDIU QUE NO SEU TÚMULO
FOSSE ESCRITO:**

***“EU VOLTAREI O MESMO,
EMBORA MUDADO”***

BIOESTATÍSTICA

DISTRIBUIÇÃO DE POISSON

DISTRIBUIÇÃO DE POISSON

- $(p+q)^k$, k MUITO GRANDE
- $(0,001+0,999)^{1000} \rightarrow$ SOLUÇÃO MUITO COMPLEXA
- DISTRIBUIÇÃO DISCRETA DO NÚMERO DE VEZES QUE UM EVENTO RARO ACONTECE (NÚMERO DE VEZES QUE O EVENTO NÃO OCORRE É GRANDE)
- VARIÁVEL ESPACIAL (NÚMERO DE INDIVÍDUOS EM PARCELAS) OU TEMPORAL (NÚMERO DE ANIMAIS CAPTURADOS POR DIA)

DISTRIBUIÇÃO DE POISSON

Siméon Denis Poisson



NASCEU EM 1781 E FALECEU EM 1840 NA FRANÇA

FOI ALUNO DE DOIS GRANDES MATEMÁTICOS NA ESCOLA POLITÉCNICA, PARIS: *LAPLACE* E *LAGRANGE*

EM 1837 PUBLICOU IMPORTANTE TRABALHO SOBRE PROBABILIDADE: *Recherchés sur la probabilité des jugements*

FOI PROFESSOR DA ESCOLA POLITÉCNICA, ASTRÔNOMO DO SERVIÇO DE LONGITUDES E CATEDRÁTICO DE MATEMÁTICA PURA DA FACULDADE DE CIÊNCIAS (1809).

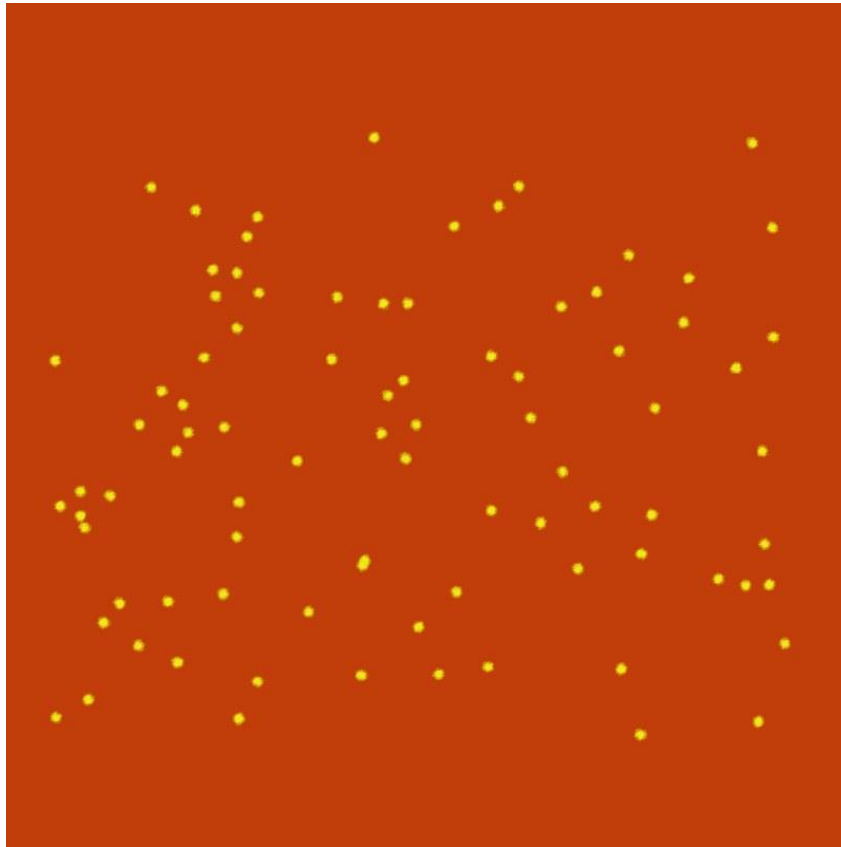
PROPRIEDADES

- MÉDIA DEVE SER PEQUENA EM RELAÇÃO AO NÚMERO MÁXIMO DE EVENTOS POSSÍVEIS, PORISSO O EVENTO É CONSIDERADO RARO.
- Ex.: NÚMERO DE INDIVÍDUOS DE MOGNO EM UMA PARCELA (MAIORIA DAS VEZES 0, 1 E 2).

$$P[X = x] = \frac{e^{-m} \cdot m^x}{x!}, x = 0, 1, 2, \dots,$$

$$m = \text{MÉDIA}$$

APLICAÇÃO:DISPERSÃO

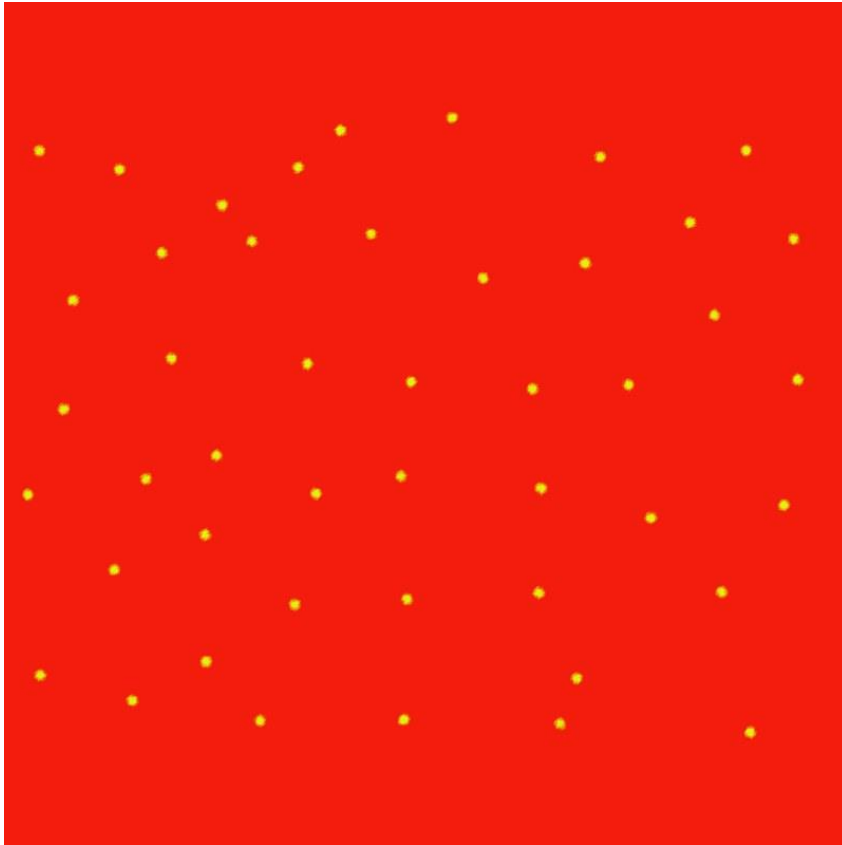


ALEATÓRIO



AGRUPADO (CONTÁGIO)

DISPERSÃO



UNIFORME

COEFICIENTE DE DISPERSÃO
(CD) = s^2 / m

CD > 1 AGRUPADO

CD = 1 ALEATÓRIO

CD < 1 UNIFORME

$$\chi^2 = \frac{(n-1).s^2}{\bar{x}}, \text{ com } (n-1) \text{ g.l.}$$

BIOESTATÍSTICA

DISTRIBUIÇÃO NORMAL

DISTRIBUIÇÃO NORMAL

- TAMBÉM CHAMADA DE DISTRIBUIÇÃO DE **GAUSS**
- O FRANCÊS **ABRAHAM DE MOIVRE** PUBLICOU A EQUAÇÃO EM 1733
- **KARL PEARSON** BATIZOU DE NORMAL PARA EVITAR UMA QUESTÃO INTERNACIONAL SOBRE AUTORIA .

DISTRIBUIÇÃO NORMAL



ABRAHAM DE MOIVRE
NASCEU EM 1667 NA FRANÇA E
FALECEU EM 1754 EM LONDRES

**PIONEIRO NO DESENVOLVIMENTO
DA GEOMETRIA ANALÍTICA E
TEORIA DE PROBABILIDADE
TENDO PUBLICADO, EM 1718, O
LIVRO: *The Doctrine of Chance*.
NESTE LIVRO ELE DEFINE
INDEPENDÊNCIA ESTATÍSTICA E
APRESENTA ALGUMAS
APLICAÇÕES PROBABILÍSTICAS
COM DADOS E CARTAS DE
BARALHO.**

**EM 1730 PUBLICA O TRABALHO
Miscellanea Analytica ONDE
APRESENTA A APROXIMAÇÃO
NORMAL DA DISTRIBUIÇÃO
BINOMIAL**

DISTRIBUIÇÃO NORMAL



Johann Carl Friedrich Gauss

NASCEU EM 1777 E FALECEU EM 1855, NA ALEMANHA

DURANTE O SEU CURSO SUPERIOR NA UNIVERSIDADE DE BRUNSWICK DESCOBRIU A APROXIMAÇÃO NORMAL DA DISTRIBUIÇÃO BINOMIAL

GAUSS AINDA PROPÔS O MÉTODO DOS QUADRADOS MÍNIMOS (BASE PARA ANÁLISE DE REGRESSÃO), QUANDO PUBLICOU O TRABALHO SOBRE A ÓRBITA MAIS PROVÁVEL DO ASTERÓIDE CERES

DISTRIBUIÇÃO NORMAL

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

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \left. \vphantom{f(z)} \right\} \begin{array}{l} \text{NORMAL} \\ \text{PADRONIZADA,} \\ \text{MÉDIA=0} \\ \text{VARIÂNCIA=1} \end{array}$$

RAZÕES PARA O ESTUDO DA NORMAL

- A MAIORIA DOS TESTES ESTATÍSTICOS PARTEM DO PRINCÍPIO QUE OS DADOS POSSUEM A DISTRIBUIÇÃO NORMAL (TESTES PARAMÉTRICOS)
- TABELAS DA DISTRIBUIÇÃO NORMAL FORAM EXTENSIVAMENTE PUBLICADAS, PRINCIPALMENTE A DE z ($N \sim 0,1$)
- A DISTRIBUIÇÃO DE MUITAS VARIÁVEIS BIOLÓGICAS É APROXIMADAMENTE NORMAL.
- VARIÁVEIS QUE NÃO SEGUEM A NORMAL PODEM SER TRANSFORMADAS PARA ATINGIR A NORMALIDADE. Ex.: LOG, RAIZ QUADRADA, ARCO SENOS, ETC.
- TEOREMA DO LIMITE CENTRAL

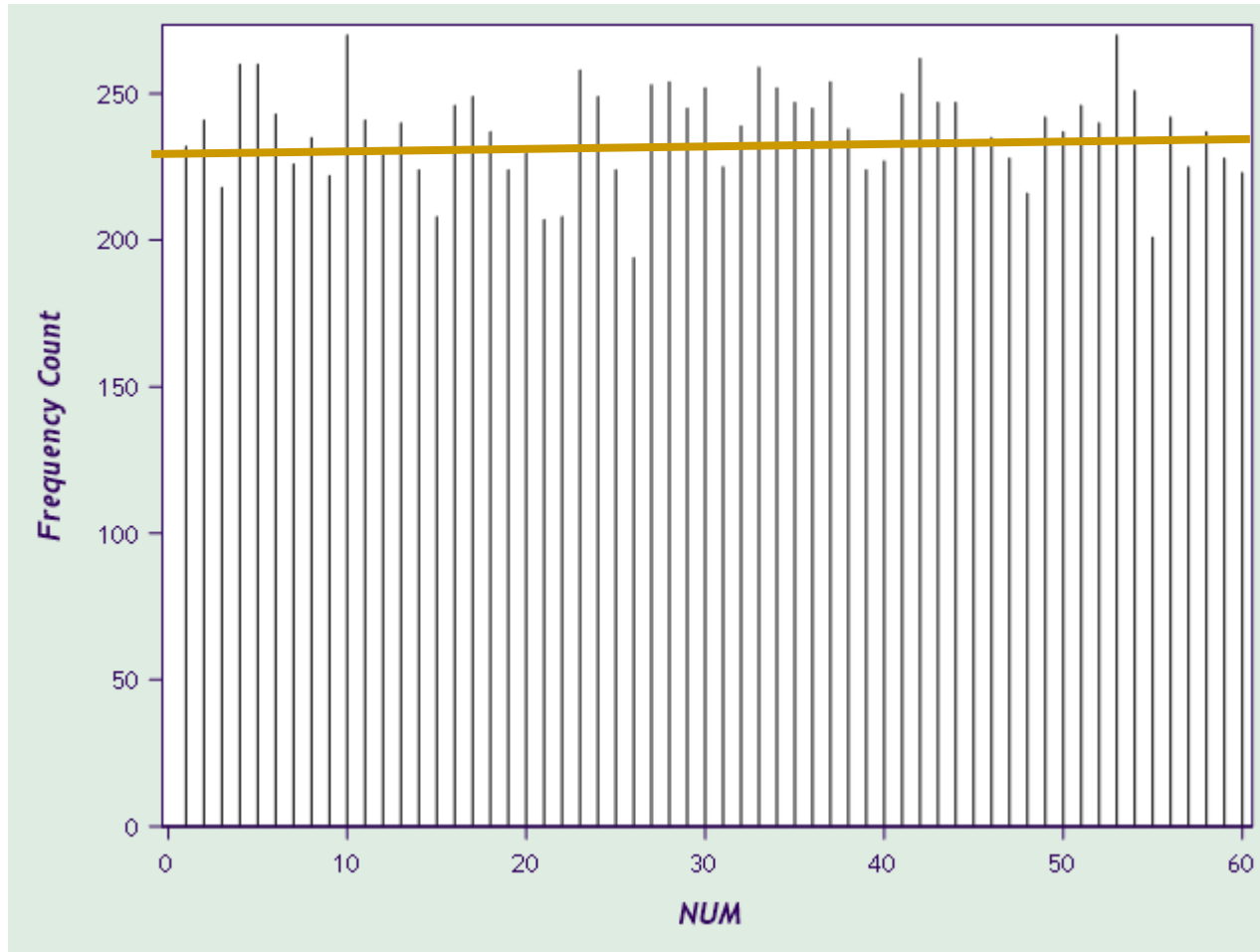
TEOREMA DO LIMITE CENTRAL

- SE X_i É UMA AMOSTRA ALEATÓRIA COM MÉDIA μ E VARIÂNCIA σ^2 , A DISTRIBUIÇÃO DA MÉDIA DA AMOSTRA TENDE A UMA DISTRIBUIÇÃO NORMAL COM MÉDIA μ E VARIÂNCIA σ^2/n , QUANDO n AUMENTA EM DIREÇÃO AO INFINITO.

Demonstração do Teorema do Limite Central

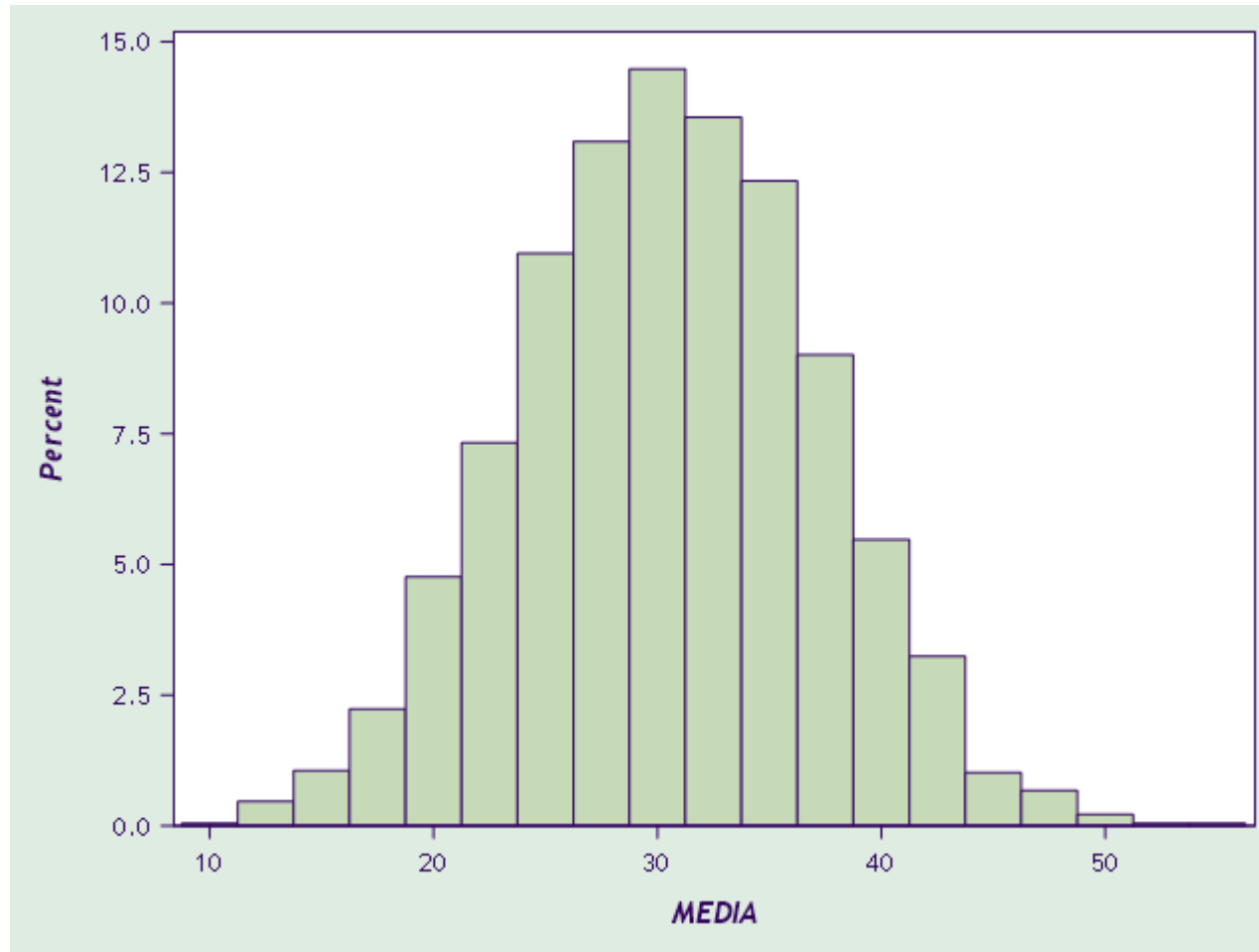
http://195.134.76.37/applets/AppletCentralLimit/Apppl_CentralLimit2.html

Exemplo da MegaSena (6 números)



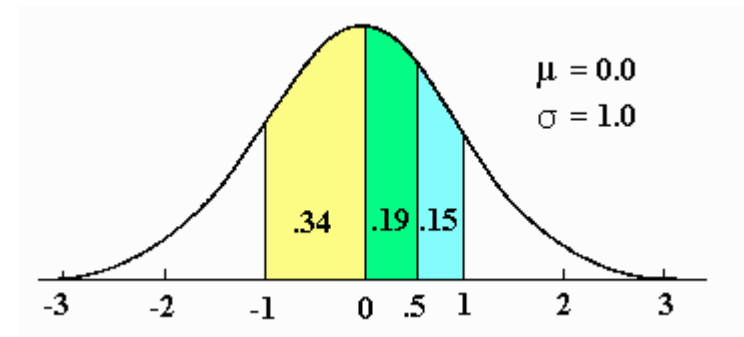
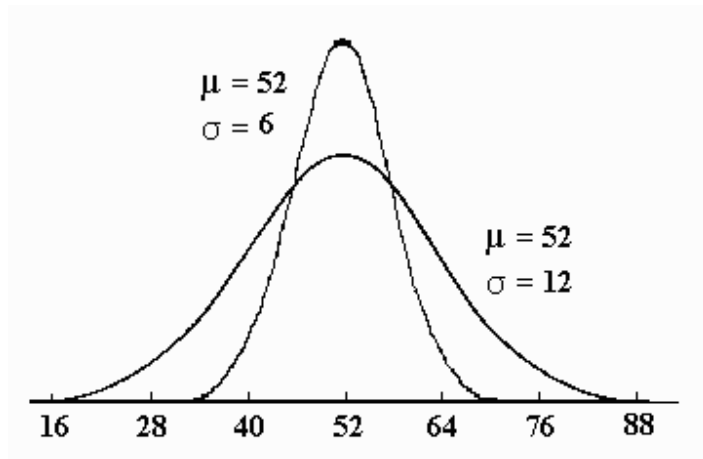
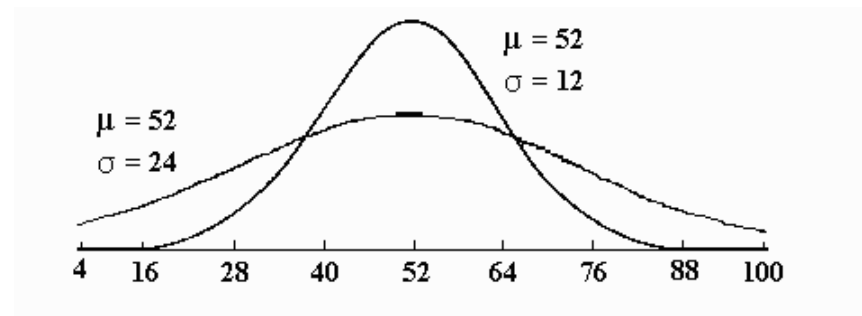
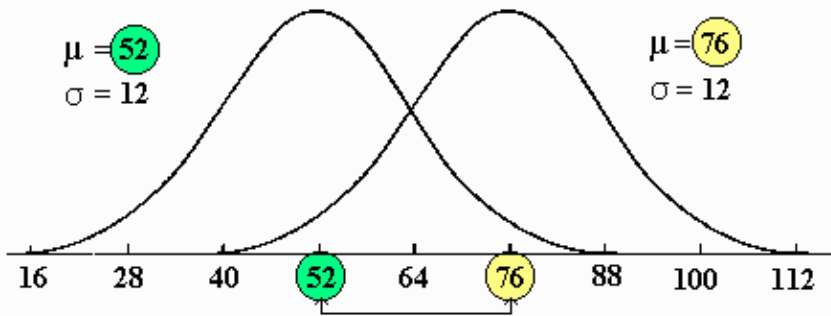
**Distribuição
Uniforme**

Média dos seis números por concurso



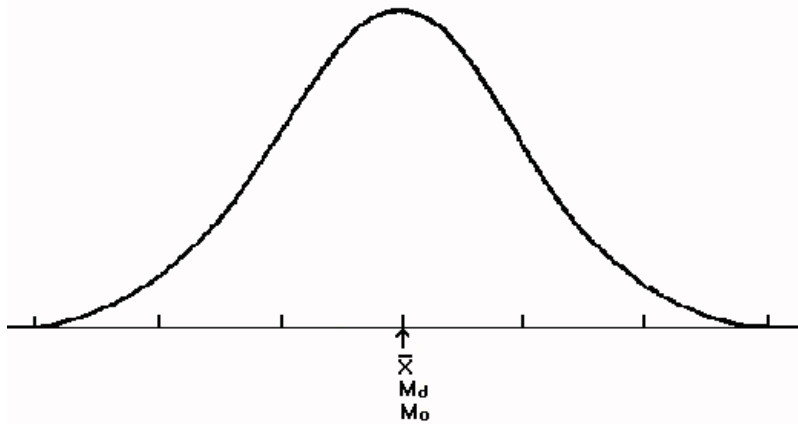
Distribuição Normal

DISTRIBUIÇÃO NORMAL

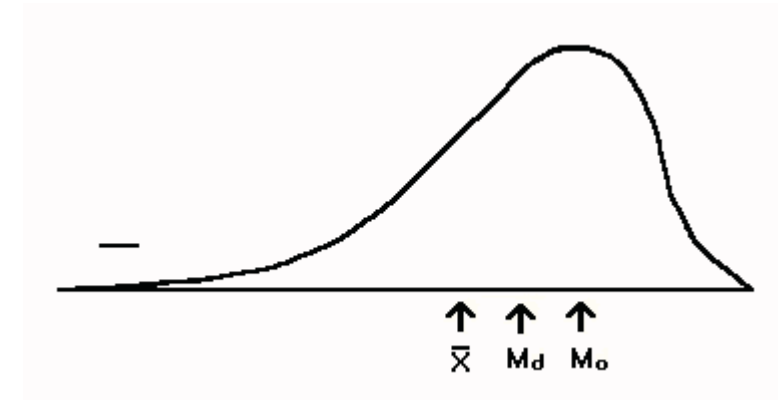
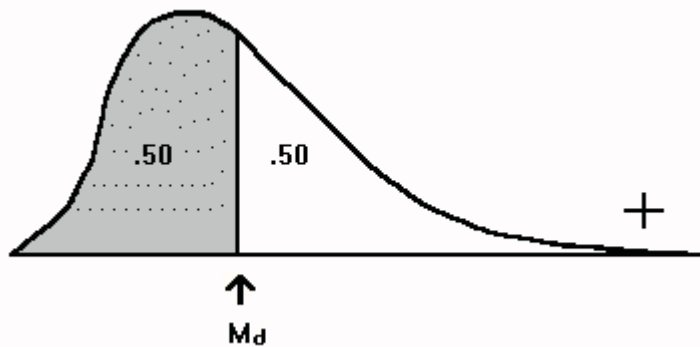
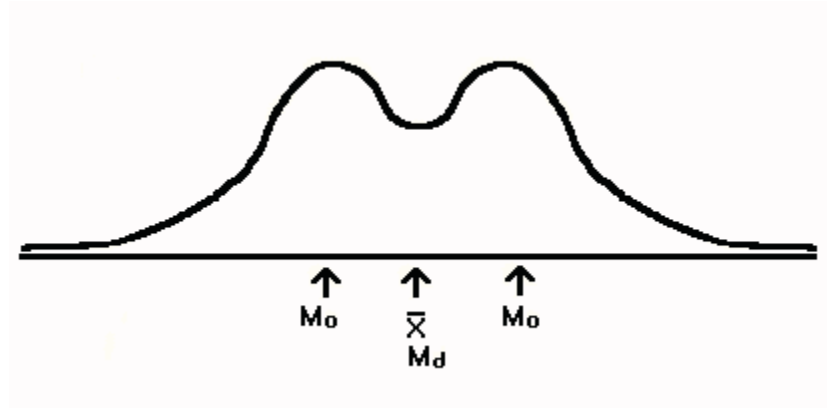


DISTRIBUIÇÃO NORMAL

NORMAL

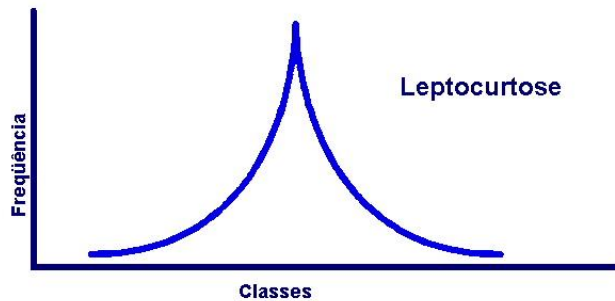
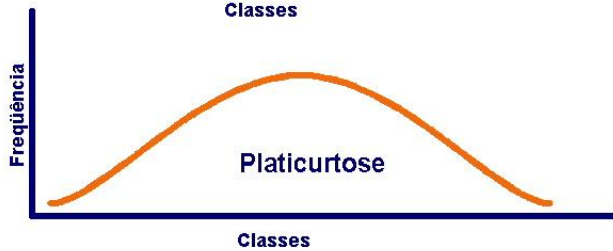
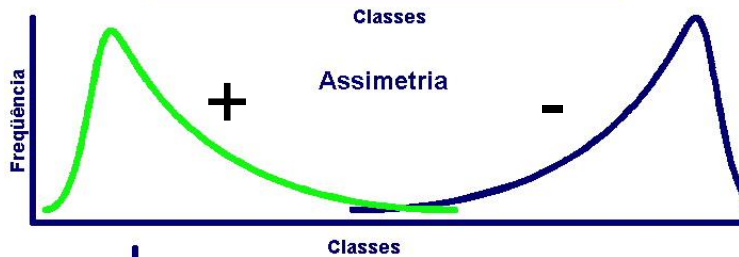
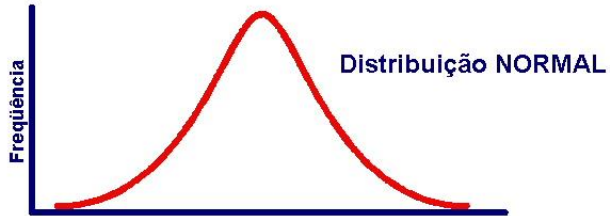


BIMODAL

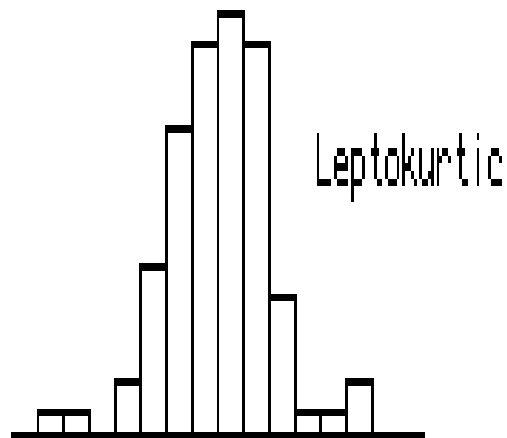


ASSIMETRIA

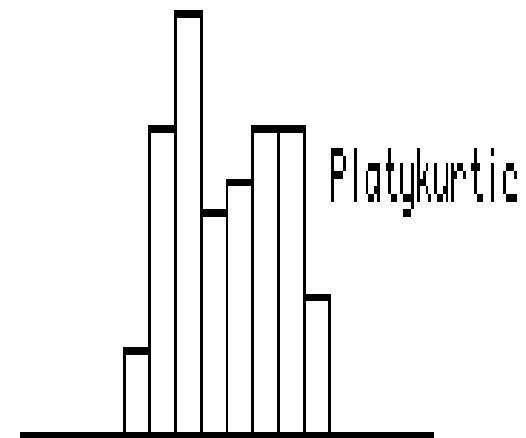
DESVIOS DA DISTRIBUIÇÃO NORMAL



Kurtosis = 1.25

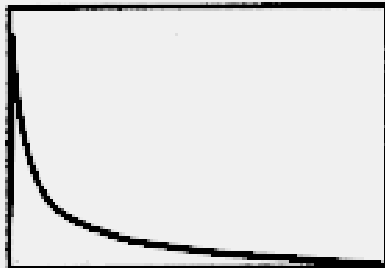
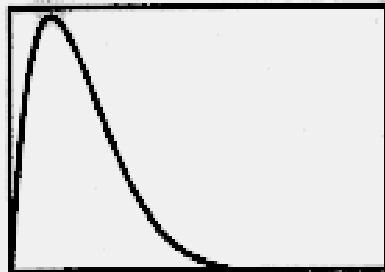
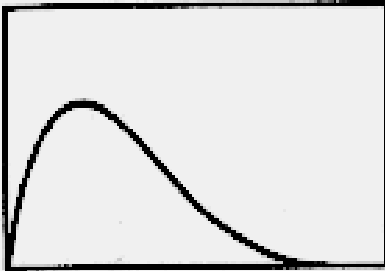


Kurtosis = -1.23



TRANSFORMAÇÃO

FORMA



TRANSFORMAÇÃO

RAIZ QUADRADA

LOGARITMO

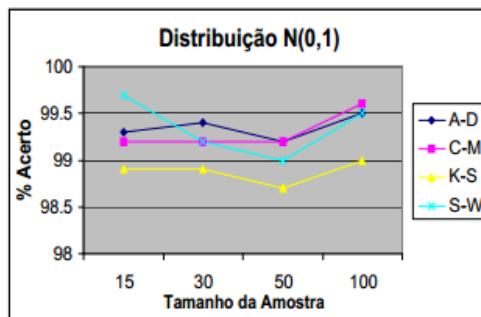
INVERSO

TESTE DE NORMALIDADE

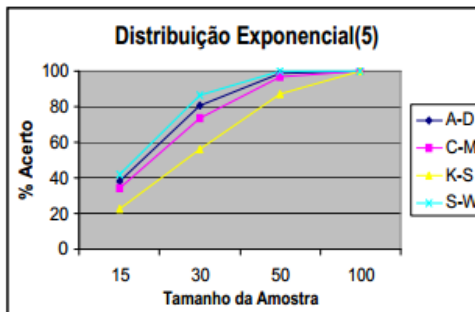
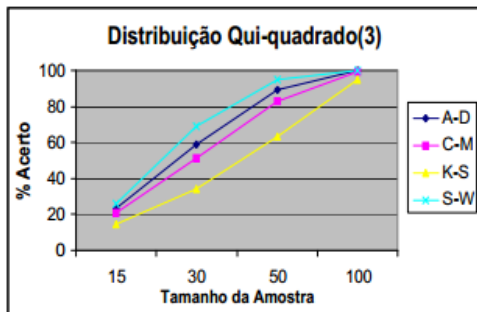
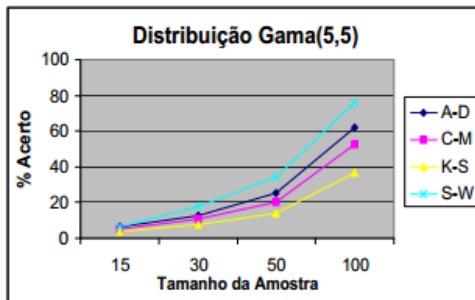
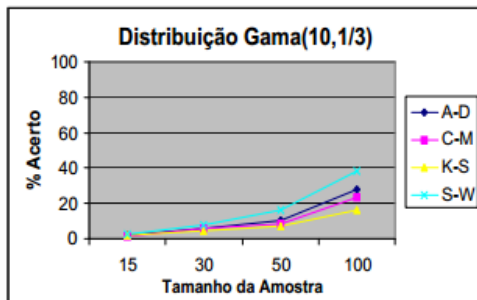
- SAS PROPÕE USO DO TESTE DE **SHAPIRO-WILK** PARA “PEQUENAS” (<1000) AMOSTRAS E DE **KOLMOGOROV-SMIRNOV** PARA “GRANDES” AMOSTRAS.
- TESTE DE SHAPIRO-WILK BASEIA-SE NO AJUSTE DA DISTRIBUIÇÃO ACUMULADA EM ESCALA PROBABILÍSTICA.

Estudo da eficiência de diferentes testes de Normalidade

Simulação com 1000 amostras de tamanhos 15 a 100, de diferentes distribuições.



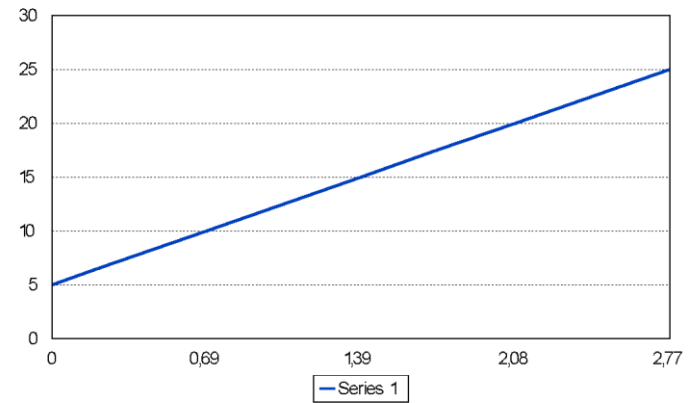
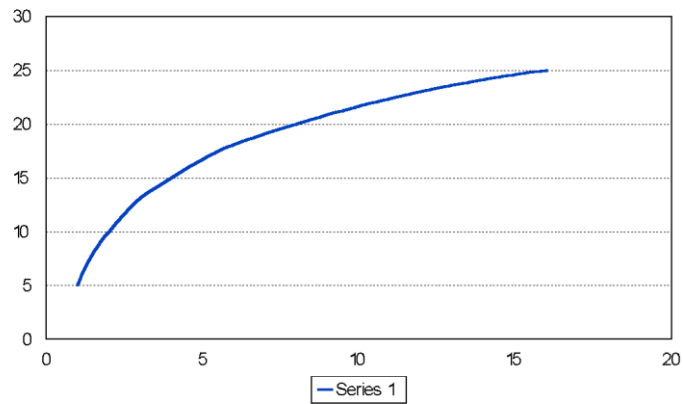
A-D = Anderson-Darling
C-M = Cramer von Mises
K-S = Kolmogorov-Smirnov
S-W = Shapiro-Wilk



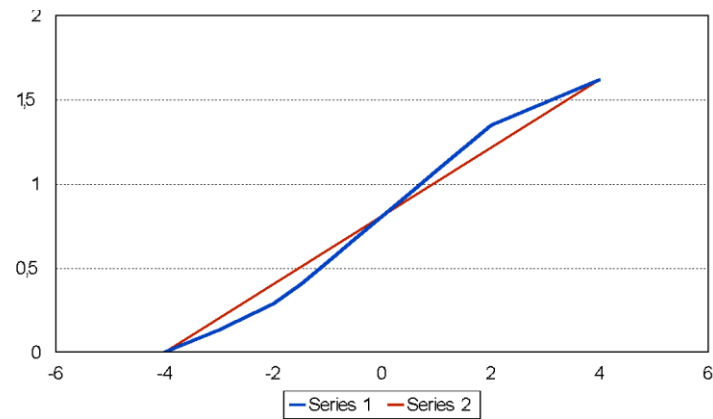
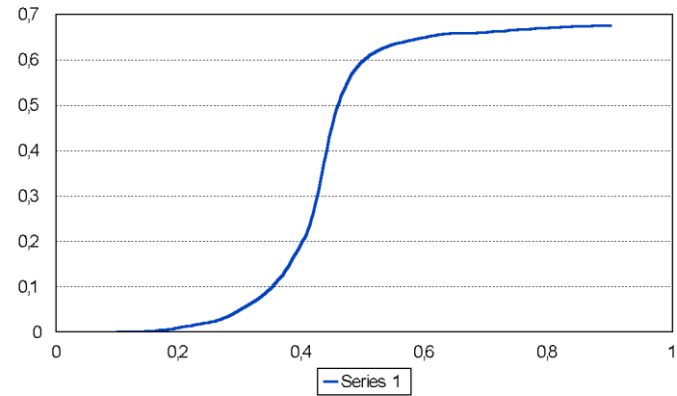
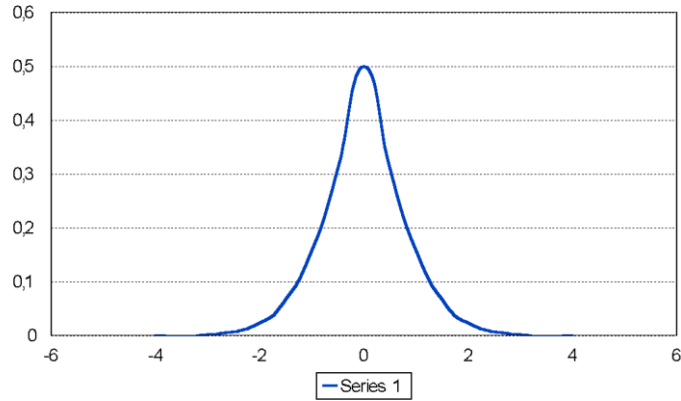
O Teste de Shapiro-Wilk foi o que apresentou maior número de acertos para não normalidade.

Leotti et al (2005)

SHAPIRO-WILK



SHAPIRO-WILK



```
TITLE2'**** B I O E S T A T Í S T I C A ****';
TITLE4'**** Análise de normalidade de dados ****';
TITLE6'**** DAP de árvores na Floresta Amazônica ****';
DATA AMAZONIA;
INPUT ARV CAP;
DAP=CAP/(ARCOS(-1));
DATALINES;
  1 34
  2 22.2
  3 19
  4 27.5
  ....
;;
ODS PDF FILE='C:\Bioestatística2014\RESULTADO_10.PDF';
PROC UNIVARIATE DATA=AMAZONIA NORMAL;
VAR DAP;
RUN;
ODS PDF CLOSE;
```

Programa SAS para teste de
Normalidade

The SAS System
****** BIOESTATÍSTICA ******

****** *Análise de dados para testar normalidade* ******

****** *DAP de uma área na Floresta Amazônica* ******

Moments			
N	60	Sum Weights	60
Mean	9.98697268	Sum Observations	599.218361
Std Deviation	5.46530964	Variance	29.8696094
Skewness	3.51480389	Kurtosis	16.4993525
Uncorrected SS	7746.68435	Corrected SS	1762.30696
Coeff Variation	54.7243876	Std Error Mean	0.70556844

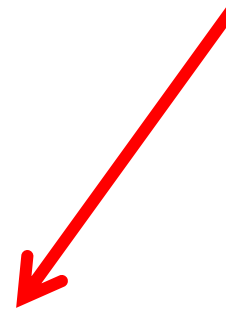
**Resultado da análise do
PROC UNIVARIATE**

Basic Statistical Measures			
Location		Variability	
Mean	9.986973	Std Deviation	5.46531
Median	8.833099	Variance	29.86961
Mode	5.156620	Range	35.01409
		Interquartile Range	4.45634

Note: The mode displayed is the smallest of 14 modes with a count of 2.

Tests for Location: $\mu_0=0$				
Test	Statistic		p Value	
Student's t	t	14.15451	Pr > t	<.0001
Sign	M	30	Pr >= M	<.0001
Signed Rank	S	915	Pr >= S	<.0001

Teste de Normalidade de SHAPIRO-WILK



Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.665817	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.200251	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.687756	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	4.337918	Pr > A-Sq	<0.0050

Quantiles (Definition 5)	
Quantile	Estimate
100% Max	40.10705
99%	40.10705
95%	17.07733
90%	13.11437
75% Q3	11.45916
50% Median	8.83310
25% Q1	7.00282
10%	5.57042
5%	5.20437
1%	5.09296
0% Min	5.09296

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5.09296	18	15.1197	37
5.15662	55	15.3744	43
5.15662	52	18.7803	23
5.25211	45	27.6930	33
5.41127	35	40.1070	32
5.25211	45	27.6930	33
5.41127	35	40.1070	32

Transformação de dados

- Uma pesquisadora coletou um conjunto de dados de biomassa em uma área restaurada com espécies nativas há mais de 50 anos.
- Ela vai realizar uma análise estatística e deseja saber se os dados apresentam distribuição Normal.
- Deseja também saber se a transformação logarítmica pode normalizar os dados.

Dados coletados (Biomassa Mg/ha)

998,68	247,34	91,09
198,88	336,33	376,98
271,73	75,12	470,96
161,48	303,53	122,06
165,89	621,75	714,73
883,22	923,67	908,45

Fazer agora.

EXERCÍCIO

(data de entrega a ser definida pelo prof.)

- Com os dados fornecidos para cada aluno/grupo produzir um relatório científico mostrando a influência do tamanho da amostra (5, 10, 15 e TODOS os animais - frangos) na distribuição dos dados. Os dados são de um estudo sobre pesagem automática de frangos em granjas (peso em kg). Arquivo DADOS_PESO10.xlsx. Semanas 4 e 5 e DIV, 1, 2, 3 e 4.

OBRIGADO !!!

*ATÉ A
PRÓXIMA !!!*