Capítulo 2

Distribuições de Probabilidade

Estimativas de parâmetros e tempos-até-falha

Flávio Fogliatto

Ajustes de distribuições

- **Em estudos de confiabilidade, dados são amostrados a partir de uma população de unidades de interesse.**

- ***Exemplo*:**

  Para determinar a distribuição dos tempos-até-falha (e assim estimar a vida média) de lâmpadas elétricas, 1000 lâmpadas são colocadas em teste por um período de tempo e seus tempos até falha são registrados.

Amostras aleatórias

- **1000 tempos obtidos no teste c/ lâmpadas compõem uma amostra aleatória da população de interesse (lâmpadas de um determinado tipo, produzidas sob condições similares).**

- **Amostras aleatórias são coletadas c/ objetivo de obter informações sobre parâmetros populacionais desconhecidos.**

- ***Exemplo*: dados amostrados seguem uma distr. Exponencial; desejamos estimar o parâmetro λ.**

Distribuições de probabilidade

- Dados empíricos normalmente seguem uma distribuição de probabilidade c/ densidade conhecida.
- A partir da distribuição de probabilidade, demais informações podem ser derivadas:
  - Média
  - Dispersão

O que são distribuições de probabilidade?

Distribuições de probabilidade

- **Considere uma variável aleatória *X*:**
  - **por exemplo, valores de tempo até falha de lâmpadas**
- **Distribuições de probb. são definidas observando:**
  - **os valores que *X* pode assumir; e**
  - **a probabilidade de *X* assumir um determinado valor.**
- **Uma distrib. de probb. é totalmente definida por uma função denominada *função de densidade de probabilidade* (ou simplesmente *densidade*).**

Função de densidade de probabilidade

- **Designada por *f*(*x*).**

- **Utilizada p/ calcular uma *área* que representa a probabilidade de *X* assumir valores no intervalo [x_1, x_2].**

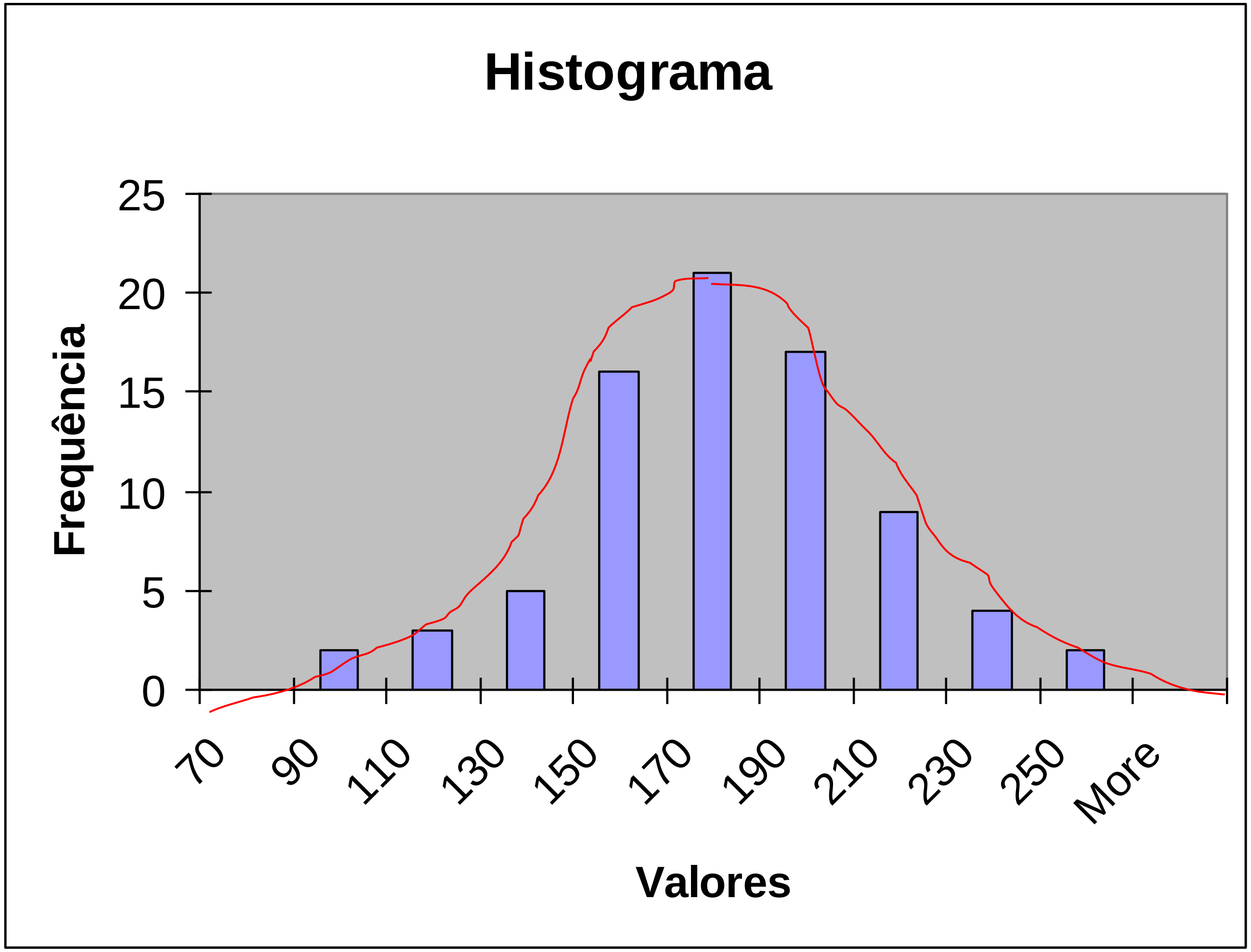
- **A probabilidade de *X* assumir valores no intervalo [x_1, x_2] é dada pela integral de *f*(*x*) avaliada no intervalo.**

Relação entre densidade e frequência

- **Considere os dados de tensão de compressão de cabos de alumínio:**

105	221	183	186	121	181	180	143
97	154	153	174	120	168	167	141
245	228	174	199	181	158	176	110
163	131	154	115	160	208	158	133
207	180	190	193	194	133	156	123
134	178	76	167	184	135	229	146
218	157	101	171	165	172	158	169
199	151	142	163	146	171	148	158
160	175	149	87	160	237	150	135
196	201	200	176	150	170	118	149

Em termos gráficos



Dados parecem se ajustar à uma distribuição normal

- **Informação no gráfico pode ser resumida em equação que descreva formato da curva que passa pelo topo das barras de freqüência.**

- **Equação da curva = função de densidade.**

- **No caso da Normal:**

$$f_X(x;\ \mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\ ,\ -\infty < x < +\infty$$

Parâmetros descrevem integralmente a função de densidade

$$f_X(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad , -\infty < x < +\infty$$

No caso da Normal, parâmetros correspondem à média e desvio-padrão da distribuição:

μ= média
σ = desvio

Distribuição acumulada

- ***Definição***: a função de distribuição acumulada de uma variável ***X*** é

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

para -∝ < *x* < +∝.

- **Informações como média e desvio-padrão são derivadas diretamente da função de distribuição.**

Dos dados amostrados à distribuição de probabilidade

- **Coletam-se dados com o objetivo de determinar:**
  1. **Uma distribuição de probabilidade que os descreva;**
  2. **Os parâmetros que caracterizam essa distribuição.**

- Tarefa 1: **utilizam-se gráficos de freqüência e testes de aderência (gráficos e analíticos) p/ hipotetizar sobre distribuições candidatas.**

- Tarefa 2: **utilizam-se estimadores dos parâmetros das distribuições que usem os dados amostrados.**

Começando pela Tarefa 2: Estimadores

- **Um estimador de um parâmetro populacional é uma fórmula que usa informações obtidas na amostra aleatória p/ gerar uma estimativa do parâmetro de interesse.**

- ***Por exemplo:* o estimador do parâmetro λ da distr. exponencial é:**

tamanho da amostra

estimador do parâmetro (parâmetro real da população é designado por λ).

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

observações que compõem a amostra

Método da máxima verossimilhança

- **Considere uma amostra aleatória obtida de uma população c/ densidade *f*(*x*) e parâmetro θ.**

- **A função de verossimilhança é dada pelo produto da densidade avaliada em cada ponto da amostra:**

$$l(\mathbf{x},\boldsymbol{\theta}) = \prod_{i=1}^{n} f(x_i;\boldsymbol{\theta})$$

Máxima verossimilhança a partir de um exemplo

- **Processo é monitorado recolhendo periodicamente amostras de 15 unidades.**

- **Seja *p* a proporção de defeitos na produção.**

- **Probb. de ocorrência de *x* defeitos nas amostras de 15 produtos é binomial:**

$$P(X = x) = \binom{15}{x} p^x (1 - p)^{15-x} \text{ , } x = 0, \ldots, 15.$$

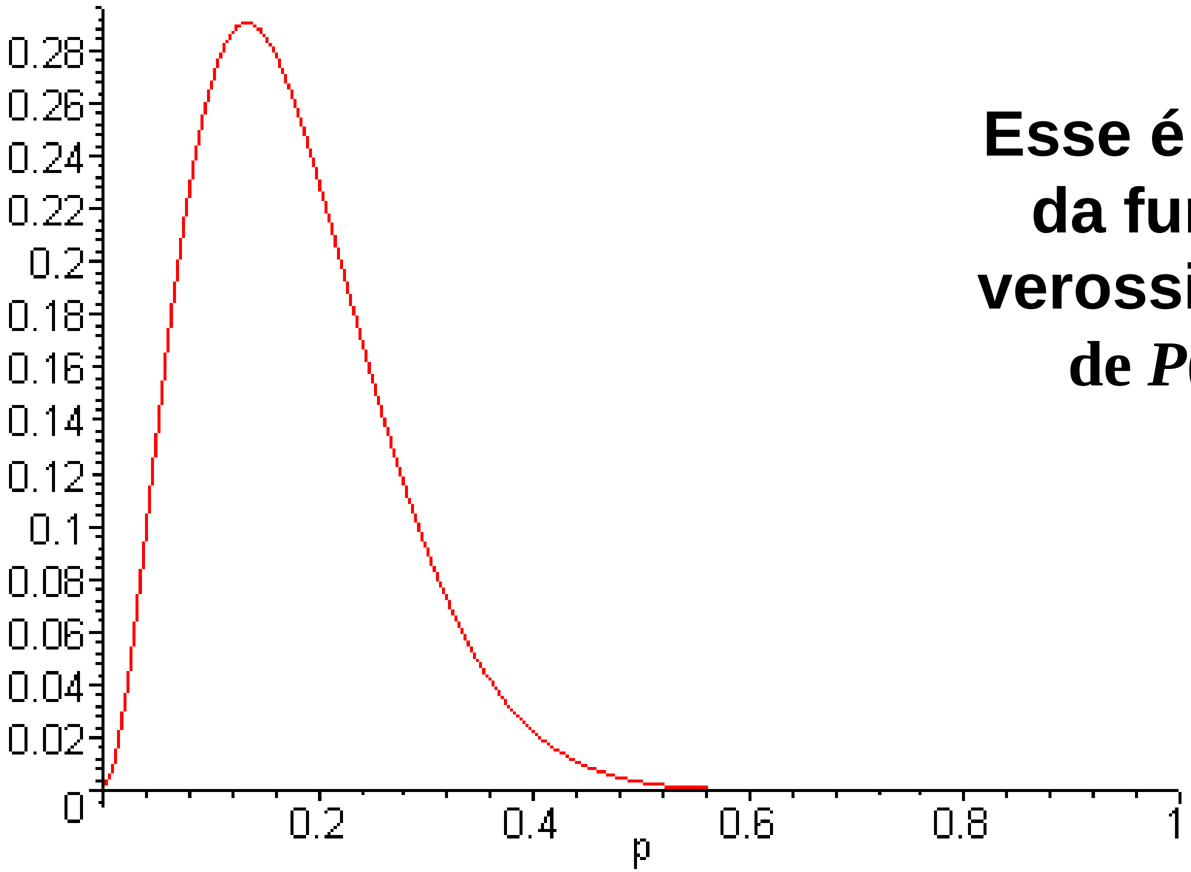
Considere a probb de ocorrência de 2 defeitos na amostra

- **Ou seja:**

$$P(X=2)=\binom{15}{2}p^2(1-p)^{13}$$

- **Como o valor de *p* é desconhecido, plotaremos esta função para diversos valores do parâmetro.**

Ocorrência de 2 defeitos *versus* diversos valores de *p*



Esse é o gráfico da função de verossimilhança de $P(X = 2)$.

Generalizando o procedimento anterior

- **Suponha que *n* amostras de 15 unidds são coletadas. Conta-se os defeitos e repete-se o procedimento anterior, desta vez plotando:**

$$P(\#\text{ def. am. 1}) \times P(\#\text{ def. am. 2}) \times \ldots$$

- **O valor de *p* que maximiza o produto acima será o estimador de máxima verossimilhança de *p*.**

Conclusão

- **Função de verossimilhança tem um máximo em valores dos parâmetros da distribuição para os quais é mais provável que os valores amostrais venham a ser observados.**

- **Para determinar o estimador de máxima semelhança de um parâmetro θ, resolve-se a expressão:**

$$\left.\frac{\partial}{\partial\theta}\right|_{\hat{\theta}} l(\mathbf{x},\theta) = 0$$

Exemplo

- **Desejamos determinar o estimador de máxima verossimilhança do parâmetro λ da distribuição exponencial:**

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

Função de máxima verossimilhança:

$$l(x_1,\ldots,x_n;\ \lambda) = l(x;\ \lambda) = f(x_1;\lambda)\ldots f(x_n;\lambda)$$

$$= \prod_{i=1}^{n} f(x_i;\lambda)$$

$$= \lambda^n \prod_{i=1}^{n} e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum_{i=1}^{n} x_i}$$

Para obter a derivada da função é mais fácil tirar seu logaritmo

- **O logaritmo de *l* (*x*, λ) é:**

$$L(\mathbf{x};\,\lambda) = n\log\lambda - \lambda\sum_{i=1}^{n} x_i$$

- **A derivada é:**

Este é o estimador.

$$\left.\frac{\partial L(x;\,\lambda)}{\partial\lambda}\right|_{\hat{\lambda}} = \frac{n}{\hat{\lambda}} - \sum_{i=1}^{n} x_i = 0 \quad \rightarrow \quad \boxed{\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i}}$$

Propriedades desejadas em um estimador de θ

- *Não-tendencioso* - **não deve superestimar ou subestimar o valor real de θ.**
- *Consistente* - **estimador não-tendencioso que converge mais rapidamente p/ o valor real de θ à medida que o tamanho da amostra aumenta.**
- *Eficiente* - **estimador consistente com variância menor do que a variância de qualquer outro estimador.**
- *Suficiente* - **estimador que utiliza toda a informação sobre o parâmetro fornecida pela amostra.**

Exercício

3) Encontre o estimador para o parâmetro λ da distribuição de Rayleigh seguindo o método da máxima verossimilhança. A função de distribuição é dada por $f(x) = \lambda x e^{-\frac{\lambda x^2}{2}}$.

Solução

$$l(x_1, x_2, \ldots, x_n; \lambda) = \prod_{i=1}^{n} \lambda x_i e^{-\frac{\lambda x_i^2}{2}}$$

substituindo $\prod_{i=1}^{n} x_i = X$, temos:

$$l(x_1, x_2, \ldots, x_n; \lambda) = \lambda^n X e^{-\frac{\lambda}{2}\sum_{i=1}^{n} x_i^2}$$

O logaritmo da função acima é:

$$L(x_1, x_2, \ldots, x_n; \lambda) = n \log \lambda + \log X - \frac{\lambda}{2}\sum_{i=1}^{n} x_i^2$$

aplicando a derivada:

$$\frac{\partial L(x_1, x_2, \ldots, x_n; \lambda)}{\partial \lambda} = \frac{n}{\lambda} - \frac{1}{2}\sum_{i=1}^{n} x_i^2 = 0$$

Finalmente, isolando λ:

$$\hat{\lambda} = \frac{2n}{\sum_{i=1}^{n} x_i^2} .$$

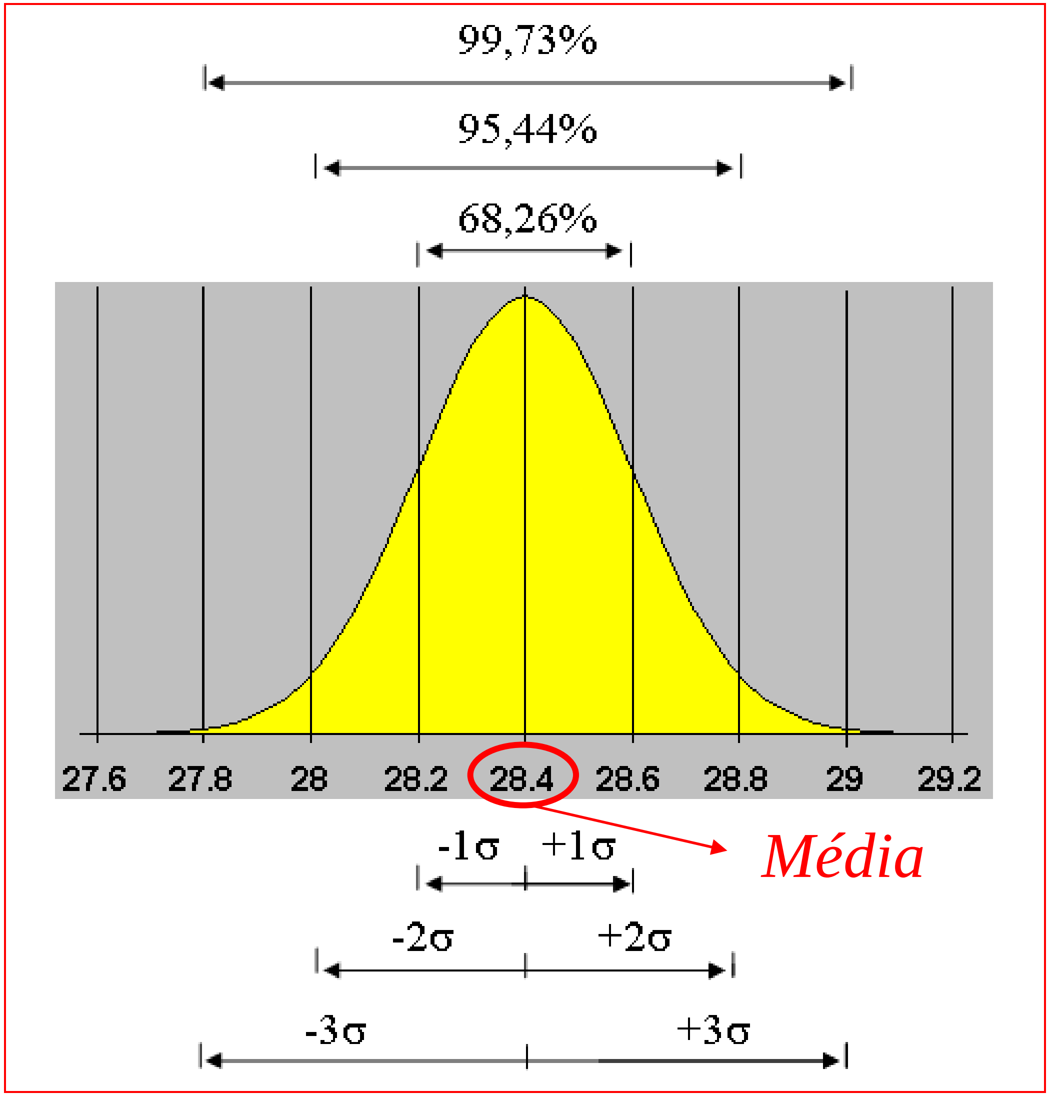
Estimação de parâmetros a partir de amostras completas

- **Quatro distribuições principais:**
  - **Normal**
  - **Exponencial**
  - **Weibull**
  - **Lognormal**

Distribuição Normal

- **Modela dados que apresentam variação aleatória e simétrica em torno de um valor central (média).**
- **A Normal é completamente descrita por dois parâmetros:**
  - **μ, que também corresponde à média da distribuição**
  - **σ, que também corresponde ao desvio-padrão da distribuição .**

Formato característico e percentuais da Normal

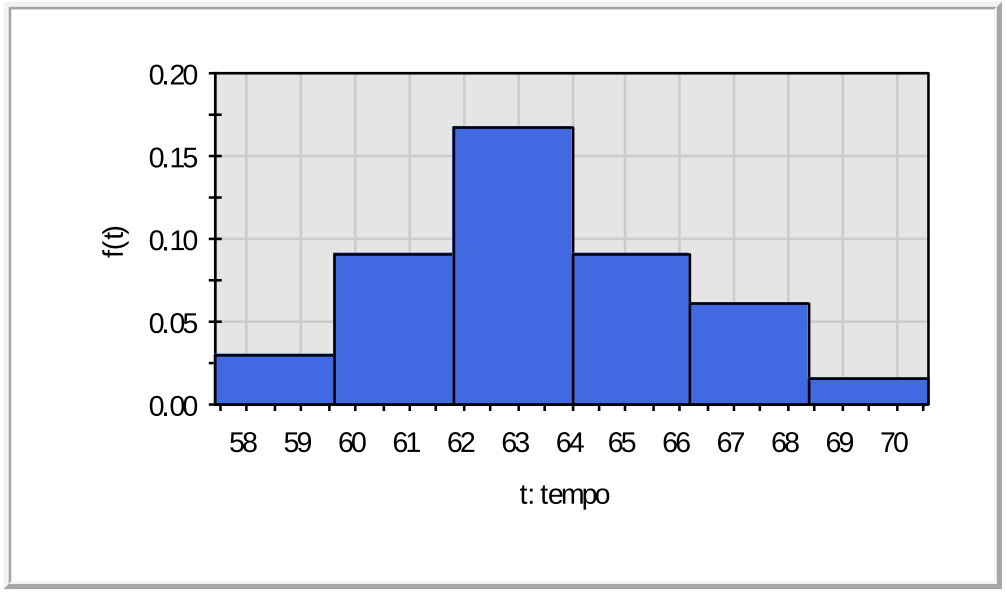


Observe a simetria em torno da média

Aplicações da Normal

- **A Normal modela bem uma grande diversidade de dados, tais como dados dimensionais, características de peso, altura, resistência.**

- *Por exemplo*: **resistência de isoladores cerâmicos**



Estimadores de máxima verossimilhança da Normal

- **Estimador de μ:**

$$\mu = \frac{\sum_{i=1}^{n} x_i}{n}$$

- **Estimador (não-tendencioso) de σ:**

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \mu)^2}{n-1}}$$

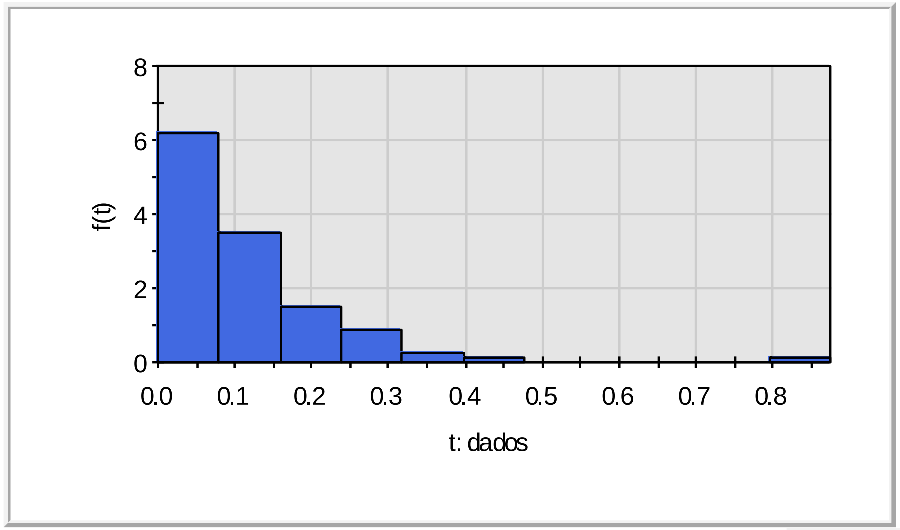
Distribuição Exponencial

- **Exponencial é assimétrica, com maiores valores de probabilidade associados a valores menores da amostra.**

- **Estimador de máxima verossimilhança para λ:**

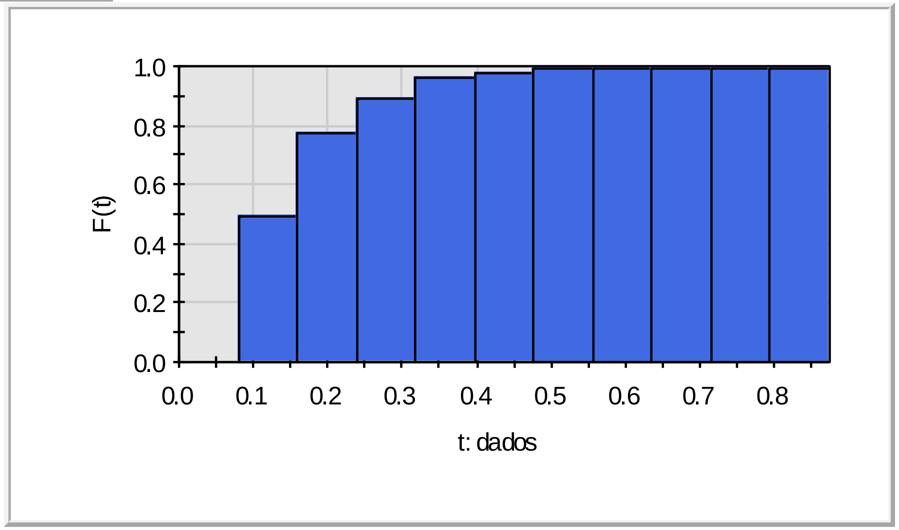
$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Densidade e Acumulada da Exponencial



Densidade

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



Acumulada

$$F(x) = 1 - e^{-\lambda x}$$

Aplicações da Exponencial

- **Usada freqüentemente para modelar a distribuição dos tempos de falha de componentes ou sistemas que apresentem uma taxa de falha constante (itens que não envelhecem no período de observação).**

- **Modela tempos de execução de algumas atividades como, por exemplo, de manutenção:**
  - **Alguns consertos são atipicamente longos, devido a incidência de defeitos raros.**

- **Outra utilização clássica da distribuição exponencial é a modelagem de tempos de fila**

Distribuição de Weibull

- **Os estimadores de máxima verossimilhança da distr. de Weibull não podem ser isolados, sendo dados pelas seguintes equações:**

$$\frac{n}{\hat{\gamma}} + \sum_{i=1}^{n} \ln x_i - \frac{1}{\hat{\theta}} \sum_{i=1}^{n} x_i^{\hat{\gamma}} \ln x_i = 0$$

$$-\frac{n}{\hat{\theta}} + \frac{1}{\hat{\theta}^2} \sum_{i=1}^{n} x_i^{\hat{\gamma}} = 0$$

Equações devem ser conjuntamente solucionadas p/ determinar estimadores.

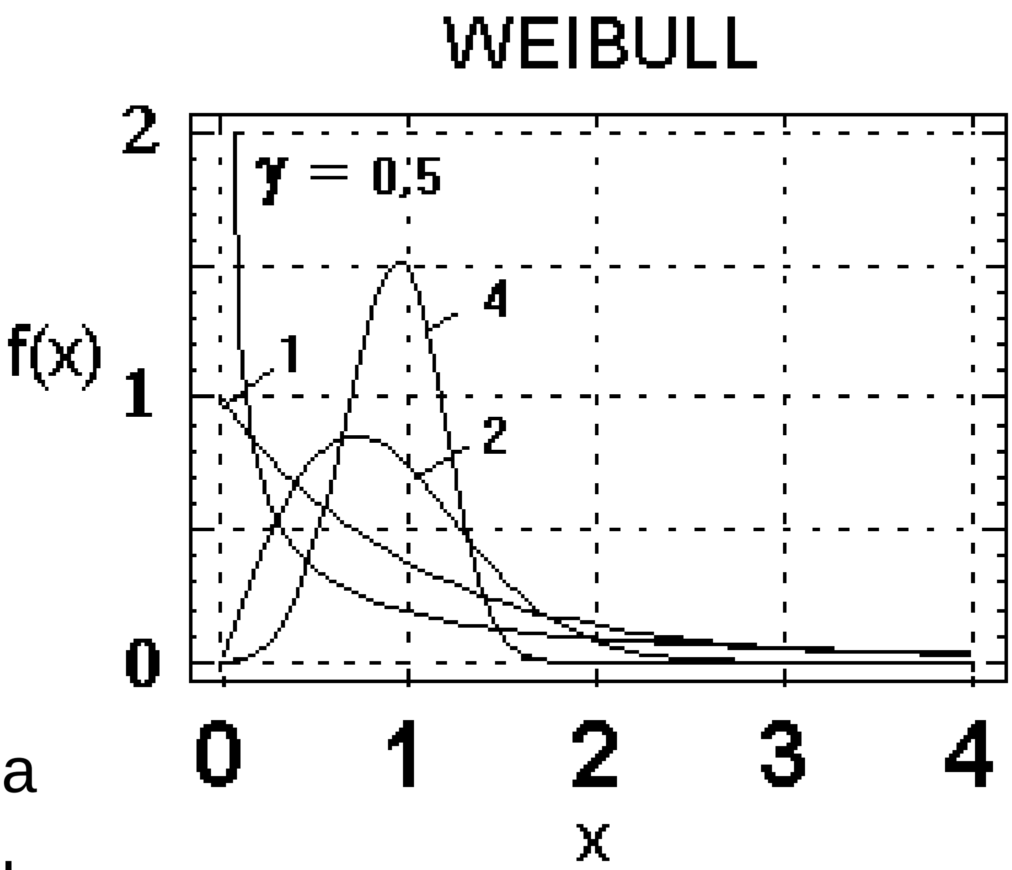
Weibull assume vários formatos conforme parâmetros

Gama	Distribuição
1	**Exponencial**
2	**Rayleigh**
3,26	**Normal**

$$f(x) = \frac{\gamma}{\theta}\left(\frac{x}{\theta}\right)^{\gamma-1} e^{-\left(\frac{x}{\theta}\right)^{\gamma}}$$

γ = parâmetro de forma

θ = parâmetro de escala



Aplicações da Weibull

- **Muito utilizada em estudos de Confiabilidade.**
- **Trata-se de uma distribuição coringa, que modela uma grande variedade de dados.**
- **No caso de escassez de dados, supor uma distribuição de Weibull como hipótese inicial pode ser uma boa política.**

Distribuição Lognormal

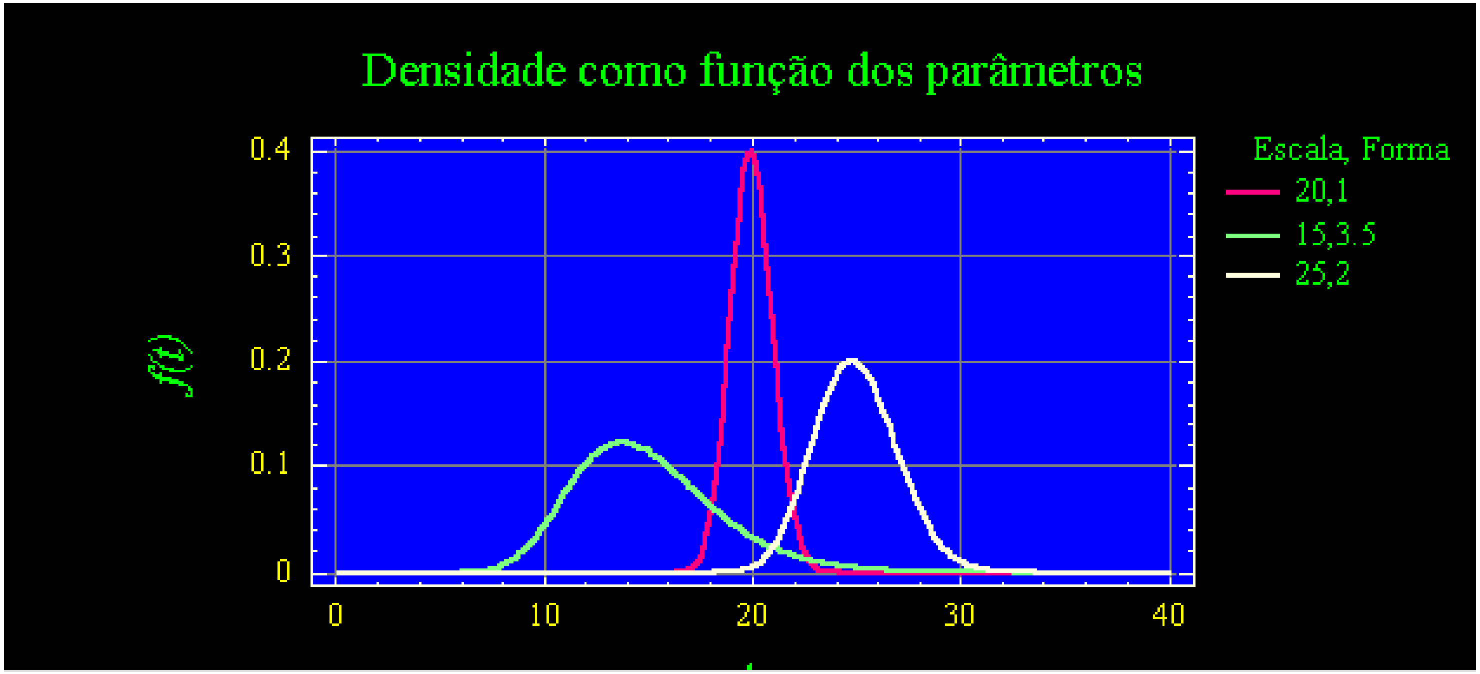
- **Pode modelar dados obtidos de:**
  - **testes com componentes eletrônicos sujeitos a um modo de falha dominante.**
  - **chegada de clientes ou serviços a servidores.**

- **Como a Normal, apresenta dois parâmetros:**
  - μ **= parâmetro de escala.**
  - σ **= parâmetro de forma.**

Densidade para diferentes valores de parâmetros

- **Função de densidade:**

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma} \right)^2 \right], -\infty < \mu < \infty,\ \ \sigma > 0,\ x > 0$$



Média da Lognormal

- **A média da lognormal é dada por:**

$$Media = \exp\left[\mu + \frac{\sigma^2}{2}\right]$$

- **Observe que, ao contrário da Normal, a média não é igual a μ; por outro lado:**

o que justifica o nome dado à distribuição.

$$Media[\ln x] = \mu$$

Recapitulando:
Dos dados amostrados à distribuição de probabilidade

- **Coletam-se dados com o objetivo de determinar:**
  - **Uma distribuição de probabilidade que os descreva;**
  - **Os parâmetros que caracterizam essa distribuição.**

- Tarefa 1: **utilizam-se gráficos de frequência e testes de aderência (gráficos e analíticos) p/ hipotetizar sobre distribuições candidatas.**

- Tarefa 2: **utilizam-se estimadores dos parâmetros das distribuições que usem os dados amostrados.**

Testes de aderência

- **Objetivo: distribuição de probabilidade dos dados é desconhecida; desejamos testar hipótese de uma determinada distribuição se ajustar aos dados.**

- **Duas famílias de testes podem ser usadas:**
  - **Gráficos – papéis de probabilidade;**
  - **Analíticos - testes do Qui-Quadrado e de Kolmogorov-Smirnov.**

Testes gráficos de aderência

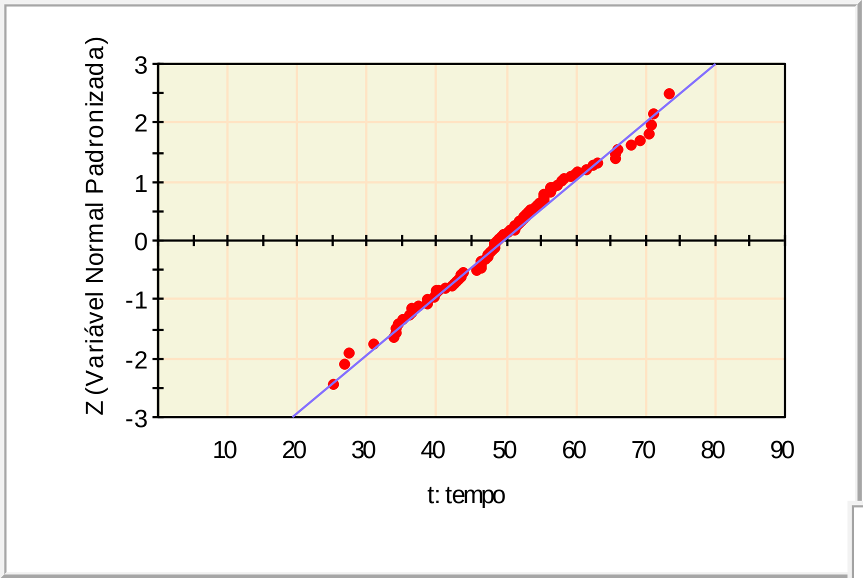
- **Papéis de probabilidade para as distribuições:**
  - **Normal**
  - **Lognormal**
  - **Exponencial**
  - **Weibull**

Papel de probabilidade
Distribuição Normal

- **Papel de probabilidade permite testar graficamente o ajuste de diferentes modelos**

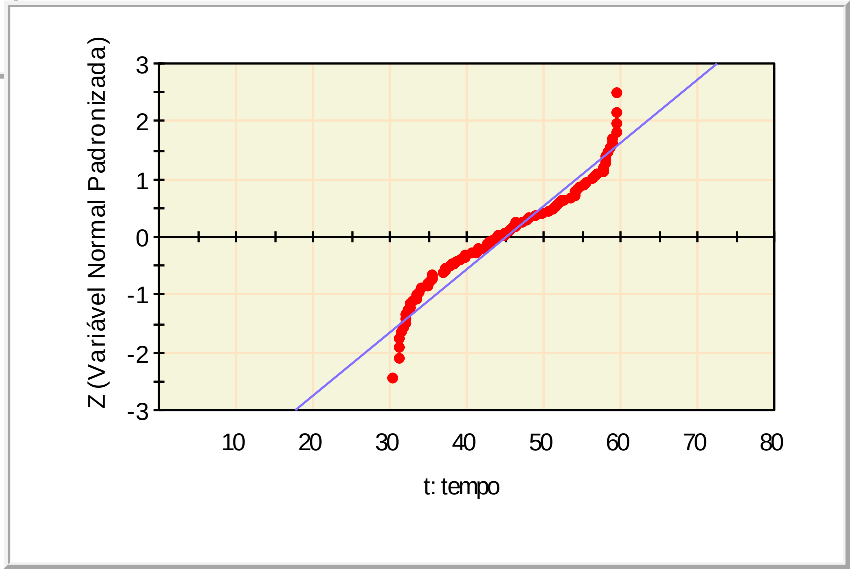
- **O papel de probabilidade da Normal possui uma escala vertical transformada:**
  - **se um conjunto de dados segue uma distribuição normal, suas freqüências acumuladas aparecerão dispostas ao longo de uma linha reta;**
  - **caso contrário, freqüências acumuladas irão apresentar curvatura no papel de probabilidade.**

Exemplos



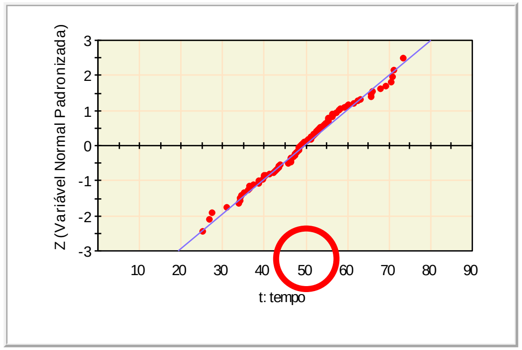
Dados bem ajustados à Normal

Dados mal ajustados à Normal



Papéis de probb permitem visualizar ajuste p/ amostras pequenas

- **Amostras c/ *n* = 10 não seriam suficientes p/ análise usando histograma:**
  - **Com ajuda do papel de probabilidade, é possível avaliar se os dados provém de uma população normal.**
- **Papel de probabilidade também pode ser usado para estimar os parâmetros do modelo (média e desvio, no caso da Normal).**
- **Por exemplo, a média é igual ao percentil de 50% (0, na escala vertical, descrita em termos de desvios da média).**



Outros papéis de probabilidade

- **Seguem a mesma lógica do papel da distribuição Normal:**
  - **Lognormal**
  - **Exponencial**
  - **Weibull**

- **Estimativa de parâmetros também é possível, entretanto utilizaremos aplicativos computacionais facilitar a tarefa.**

Testes analíticos de aderência

- **Teste do Qui-Quadrado**

- **Teste de Kolmogorov-Smirnov**

Teste do Qui-Quadrado

- **Temos uma amostra de *n* observações de uma população c/ distr. de probabilidade desconhecida.**

- **Organize os *n* pontos amostrais em um histograma de frequência com *k* classes.**

- **Seja O_i a frequência observada na $i^{ésima}$ classe.**

- **Seja E_i a frequência esperada caso a população amostrada siga uma distribuição de probabilidade hipotetizada.**

Teste do Qui-Quadrado

- **O teste compara O_i e E_i através da seguinte expressão:**

$$X_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

caso a amostra siga a distribuição hipotetizada, pode-se demonstrar que X_0^2 segue uma distribuição Qui-Quadrado, com *k* - *p* - 1 graus de liberdade.

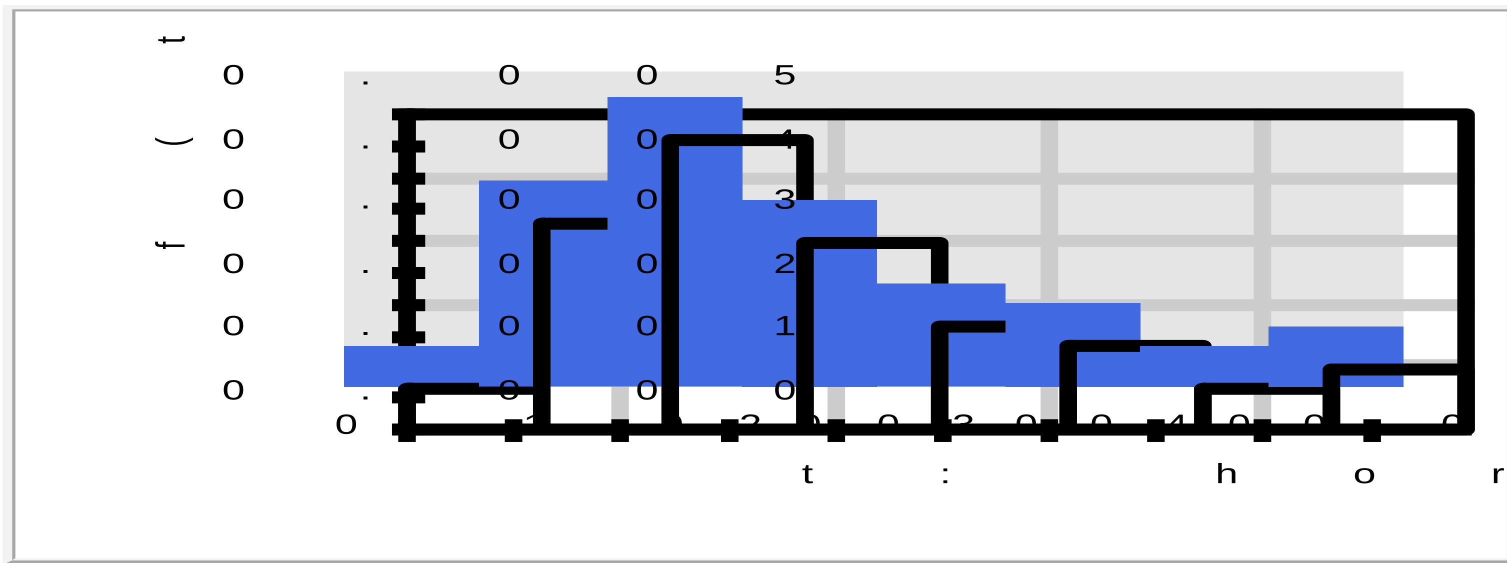
Exemplo: dados de tempo até falha de fontes de computador

Amostra composta de 49 pontos amostrais e obtidas simulando a partir da distribuição lognormal.

Tempos = × 1000

15	137	218	415
23	140	225	436
62	145	230	457
78	149	237	472
80	153	242	
85	158	255	
97	162	264	
105	167	273	
110	171	282	
112	175	301	
119	183	312	
121	189	330	
125	190	345	
128	197	360	
132	210	383	

Histograma de frequência



Histograma sugere duas possibilidades:

- **Weibull**
- **Lognormal**

Teste do Qui-Quadrado

		Freq.	Freq.
Limite Inferior	Limite Superior	Observada	Esperada
0	61,8	2	4,2
61,8	123,7	10	9,3
123,7	185,6	14	10,8
185,6	247,4	9	9,5
247,4	309,3	5	6,9
309,3	371,1	4	4,3
371,1	432,9	2	2,3
432,9	Mais	3	1,8

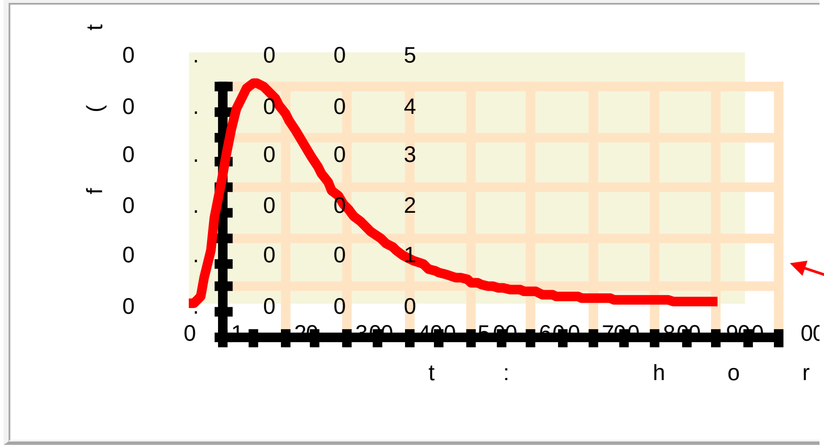
- **Hipótese de Weibull não pode ser rejeitada no teste do Qui-Quadrado.**
- **Significância = 0.62**

Teste do Qui-Quadrado

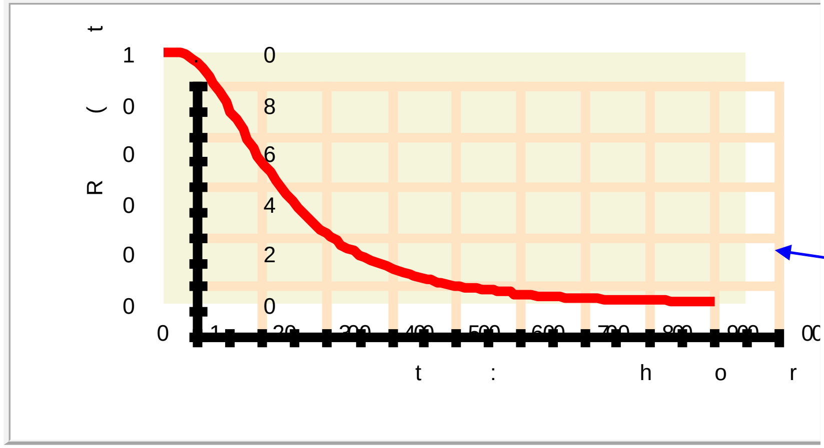
Limite Inferior	Limite Superior	Freq. Observada	Freq. Esperada
0	61,8	2	3,2
61,8	123,7	10	12,4
123,7	185,6	14	11,5
185,6	247,4	9	7,8
247,4	309,3	5	5
309,3	371,1	4	3,1
371,1	432,9	2	2
432,9	Mais	3	4

- **Hipótese de Lognormal não pode ser rejeitada no teste do Qui-Quadrado.**
- **Significância = 0.82 (maior que a de Weibull).**

Supondo distribuição lognormal



Densidade



Confiabilidade (complemento da acumulada)

Teste de Kolmogorov-Smirnov

- **Teste também trabalha com frequências observadas e esperadas, mas estatística de teste é diferente daquela utilizada no teste do Qui-Quadrado (utiliza frequências *acumuladas*).**

- **Teste K-S é não-paramétrico, ou seja, não é baseado em nenhuma distribuição de probabilidade (como a do Qui-Quadrado, p.ex.).**

Comparativo entre testes

- **Qui^2 apropriado p/ dados discretos; K-S apropriado p/ dados contínuos.**
- **Qui^2 é sensível ao agrupamento de dados em classes.**
- **K-S usa toda a informação na amostra ao usar probabilidades acumuladas.**
- **Qui^2 demanda tamanhos de amostra grandes; K-S funciona bem c/ qualquer tamanho de amostra.**

Exercícios

- **Resolver os exercícios 4, 5 e 6**