CAPÍTULO 4
CONCEITOS BÁSICOS DE ESTATÍSTICA E PROBABILIDADES

1. INTRODUÇÃO

- Conceito de população desconhecida π e proporção da amostra observada P.

π = P + pequeno erro

Perguntas: - Qual é o pequeno erro?

- Até que ponto estamos certos?

$$\boxed{\Pi = P \pm 1{,}96\sqrt{\frac{P(1-P)}{n}}}$$ **(com 95% de confiança)**

onde: n ... tamanho da amostra

P... proporção da amostra

Exemplo: Em 1972, pouco antes das eleições presidenciais nos Estados Unidos, uma pesquisa Gallup, feita junto a 2000 eleitores, acusou 760 favoráveis a McGovern e 1240 favoráveis a Nixon. Calcular o intervalo π de confiança de 95% em relação a proporção P, da população que votou em McGovern.

- Aumentando n, aumenta-se a confiança. Para 99% de confiança, tem-se:

$$\boxed{\Pi = P \pm 2{,}58\sqrt{\frac{P(1-P)}{n}}}$$

2. VARIABILIDADE

Duas peças nunca serão exatamente iguais: variações podem ser mínimas.

3. TIPOS DE VARIAÇÃO

A Dentro da própria peça

B. Entre peças produzidas no mesmo período

C. Entre peças produzidas em períodos diferentes

4. TIPOS DE CARACTERÍSTICAS DA QUALIDADE

- Variável: quando se faz um registro de uma característica medida, através de uma dimensão. Uma variável pode ter todos os valores possíveis entre dois valores limite**.**

- Atributo: quando se faz registros de um número de ítens que atendam ou não a qualquer requisito específico, dizemos que esses registros são feitos com base em atributos.

- Variáveis tratadas como atributos: a qualidade medida de ítens é chamada de variável, e o grupamento de itens em função de suas dimensões é chamado de atributo.

5. CLASSIFICAÇÃO DOS DADOS

- Discretos: assumem determinados valores sem valores intermediários.

- Contínuos: assumem qualquer valor dentro dos limites especificados.

6. DISTRIBUIÇÃO DE FREQUÊNCIAS

Nenhum processo pode fornecer ítens exatamente iguais, sempre existirá variação.

⇨ **Distribuição de freqüências é útil para entender e interpretar a natureza dos dados**

6.1. Diagrama de Freqüências

Exemplo de distribuição de freqüências para dados contínuos.

Peso de eixo de transmissão de aço (em Kg)				
2,559	2,556	2,566	2,546	2,561
2,570	2,546	2,565	2,543	2,538
2,560	2,560	2,545	2,551	2,568
2,546	2,555	2,551	2,554	2,574
2,568	2,572	2,550	2,556	2,551
2,561	2,560	2,564	2,567	2,560
2,551	2,562	2,542	2,559	2,556
2,556	2,550	2,561	2,559	2,561
2,559	2,557	2,532	2,575	2,551
2,550	2,559	2,565	2,552	2,560
2,534	2,547	2,569	2,559	2,549
2,544	2,550	2,552	2,536	2,570
2,564	2,553	2,558	2,538	2,564
2,552	2,543	2,562	2,571	2,553
2,539	2,569	2,552	2,536	2,537
2,532	2,552	2,575	2,545	2,551
2,547	2,537	2,547	2,533	2,538
2,571	2,545	2,545	2,556	2,543
2,551	2,569	2,559	2,534	2,561
2,567	2,572	2,558	2,542	2,574
2,570	2,542	2,552	2,551	2,553
2,546	2,531	2,563	2,554	2,554

6.2. <u>Cálculos para Distribuição de Freqüência</u>

A. *Montar a folha de contagem*

Folha de contagem de valores para pesos de eixos de ação (com origem em valores de 2,500 Kg).

PESO	TABULAÇÃO	PESO	TABULAÇÃO	PESO	TABULAÇÃO
31	1	46	1111	61	~~1111~~
32	11	47	111	62	11
33	1	48		63	1
34	11	49	1	64	111
35		50	1111	65	11
36	11	51	~~1111~~ 111	66	1
37	11	52	~~1111~~ 1	67	11
38	111	53	111	68	11
39	1	54	111	69	111
40		55	1	70	111
41		56	~~1111~~	71	11
42	111	57	1	72	11
43	111	58	11	73	
44	1	59	~~1111~~ 11	74	11
45	1111	60	~~1111~~	75	11

Números de células de 5 a 20.

B. ***Determinar a amplitude***

$$R = H - L$$

$$= 2{,}575 - 2{,}531 = 0{,}044 \text{ (Kg)}$$

C. Determinar o intervalo de células

$$C = R/\ i$$

onde: C = número de células

R = amplitude

i = intervalo de célula

D. Determinar os limites das células

Valores extremos, conhecidos como limite superior e inferior.

E. Determinar os pontos médios das células

Ponto central entre os limites superior e inferior.

F. Mostrar a freqüência após a célula

Quantidades de valores correspondentes a cada célula.

G. Freqüência acumulada Maior (ou Menor)

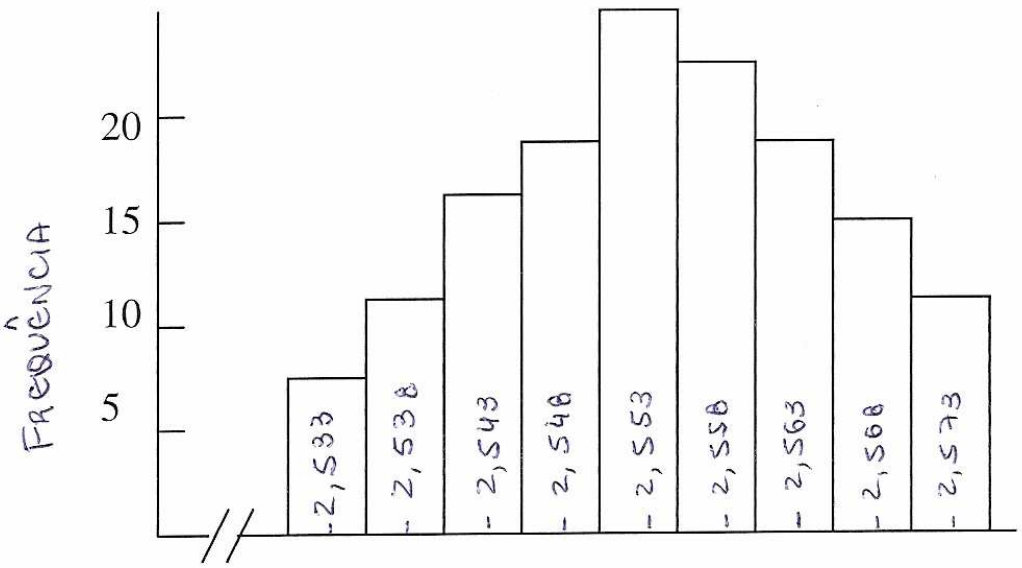
Quantidades de valores iguais ou maiores (ou menores) que a célula indicada.

Exemplo: Distribuição de freqüências de pesos de eixos de aço.

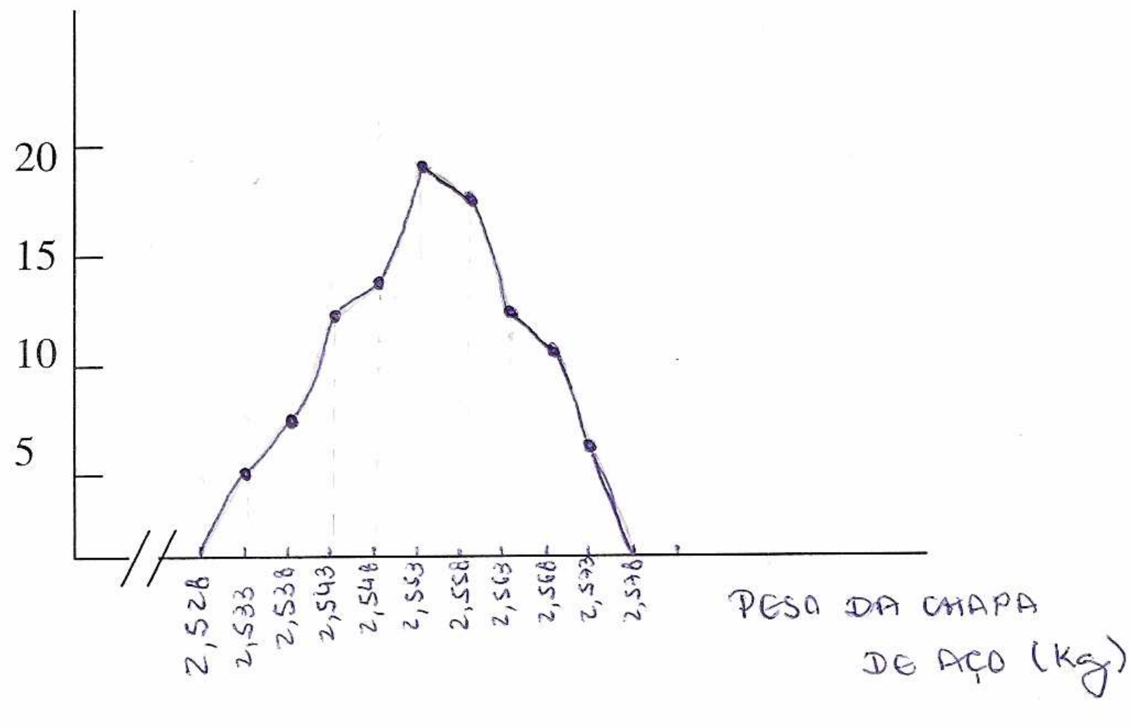
Limites das Células	**Ponto médio das células**	**Freqüência**	**Freqüência menor que**	**Acumulada maior que**
2,531 – 2,535	2,533	6	6	110
2,535 – 2,539	2,537	7	13	104
2,539 – 2,543	2,541	4	17	97
2,543 – 2,547	2,545	12	29	85
2,547 – 2,551	2,549	8	37	77
2,551 – 2,555	2,553	20	57	57
2,555 – 2,559	2,557	9	66	48
2,559 – 2,563	2,561	19	85	29
2,563 – 2,567	2,565	7	92	22
2,567 - 2,571	2,569	10	102	12
2,571 – 2,575	2,573	6	108	6
2,575 – 2,579	2,577	2	110	4

6.3. Apresentação Gráfica de Distribuições de freqüências

A. Histograma

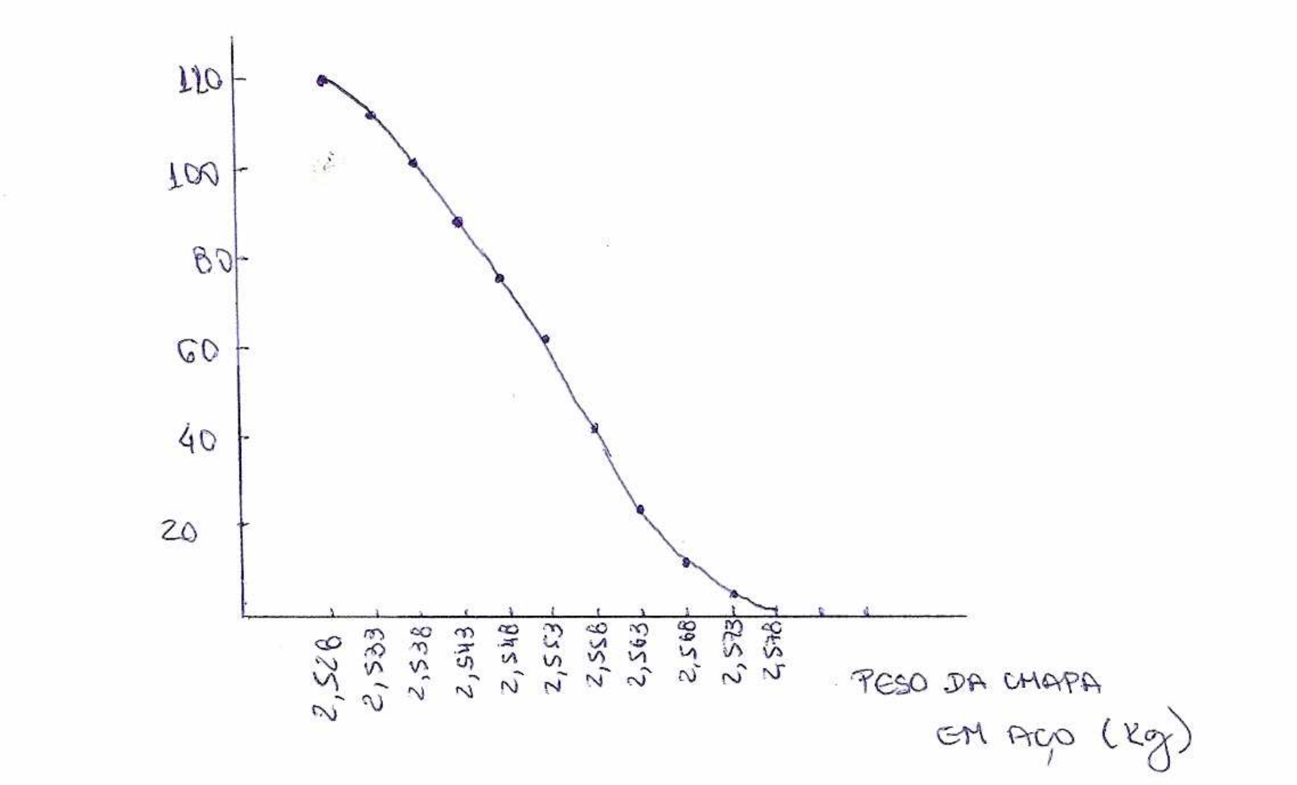


B. Polígono de Freqüências

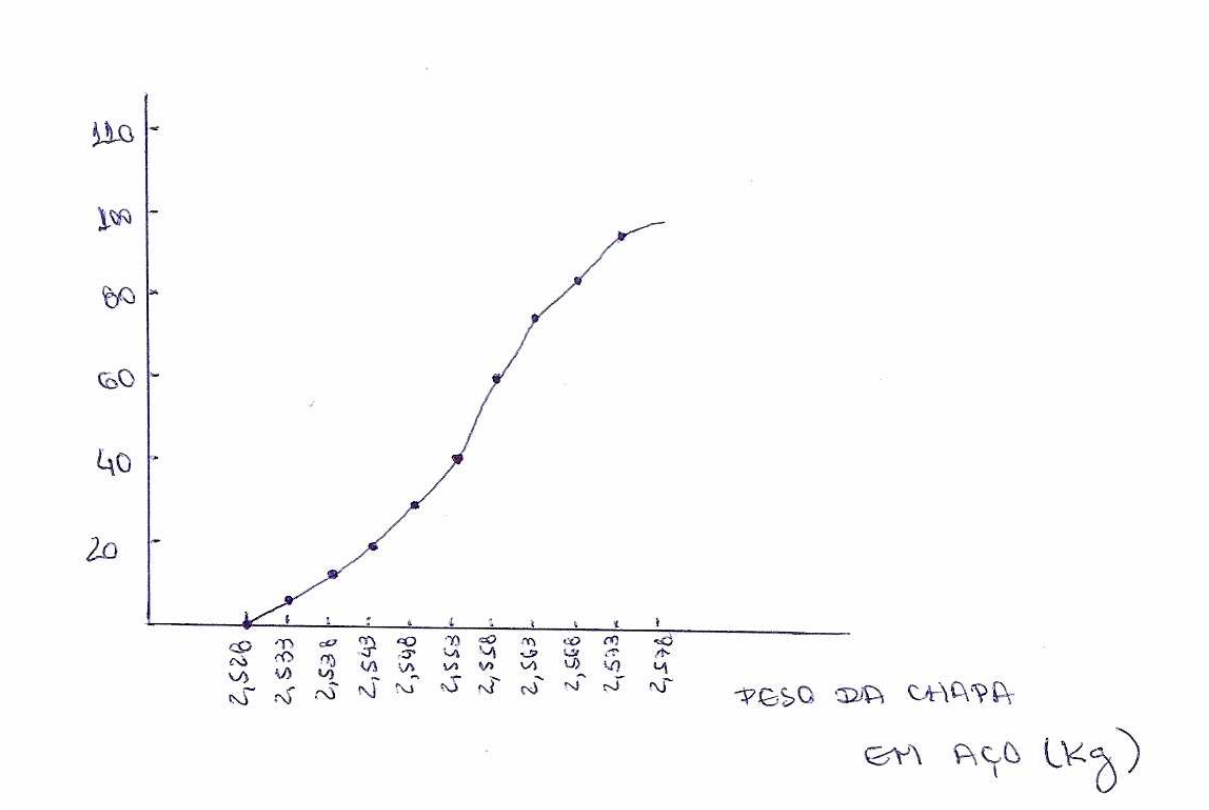


C. Gráfico de Freqüências Acumuladas ou Ogivas

- <u>Freqüência Acumulada Maior Que</u>



- <u>Freqüência Acumulada Menor Que</u>



8. MEDIDAS DE TENDÊNCIA CENTRAL

8.1. Moda

É o valor ou classe de maior frequência.

Ex: 2 5 5 7 9 9 12 11 5

8.2. Mediana

É o valor do centro da distribuição, após a mesma ter sido ordenada.

Ex: 2 5 5 5 7 9 9 11 12

E se o número de observações for par?

Ex: 1 3 5 7 8 9 11 12

Mediana: (7 + 8) / 2 = 7,5

8.3. Média X

Média aritmética das observações

$$x = \frac{x_1 + x_2 + \ldots + x}{n} = \frac{\sum^{n} x_i}{i = 1}$$

9. MEDIDAS DE DISPERSÃO OU VARIABILIDADE

9.1. Variância (S^2)

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - x)^2}{n-1}$$

9.2. Desvio Padrão (S):

$$S = \sqrt{s^2} \qquad\qquad S = \sqrt{\frac{\sum_{i=1}^{n} (x_i - x)^2}{n-1}}$$

9.3. Amplitude (R):

$$\boxed{\mathbf{R = X_{max} - X_{min}}}$$

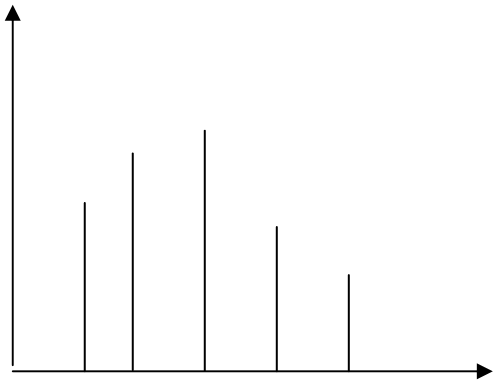
10. Distribuição de Probabilidades

$$P_r(e_1) = \lim \frac{n_1}{n} = \frac{possíveis}{todos}$$

Forma de associar os eventos possíveis com suas probabilidades.

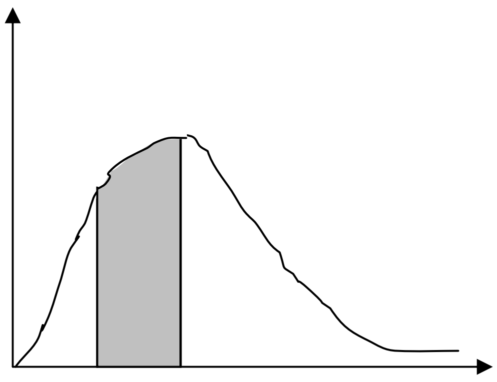
- Distribuições Discretas

  Variável discreta.



- Distribuições Contínuas

  Variáveis contínuas.



Conceitos Matemáticos de Apoio:

1) Arranjo: nº de grupos de K elementos que se pode ordenar de n elementos.

$$A_{n,K} = \frac{n!}{(n-K)!}$$

Exemplo: Quantas maneiras pode-se arranjar as letras A, B, C e D, duas a duas?

$$A_{4,2} = \frac{4!}{(4-2)!} = \frac{4.3.2.1}{2.1} = 12$$

2) Permutação: caso particular de arranjo de n elementos agrupados em formas diferentes.

P_n=n!

Exemplo: Quantas maneiram se permutam as letras A, B e C?

P_3 = 3! = 3 . 2. 1 = 6

3**) Combinação:** nº de grupos K elementos que pode obter de n elementos.

$$C_{n,K} = \frac{n!}{K!(n-K)!}$$

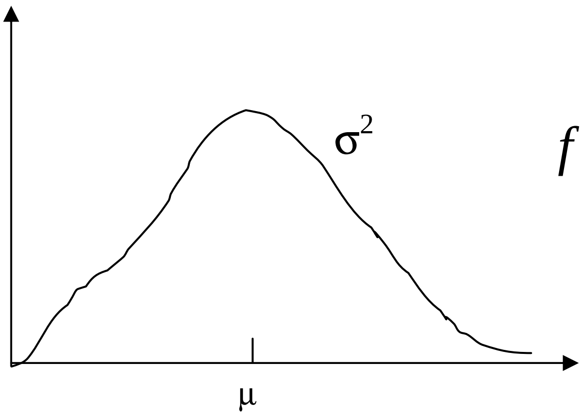
Exemplo: Quantos resultados são possíveis ao se retirar 2 bolas, com reposição de uma urna contendo uma bola vermelha, uma azul, uma amarela e uma preta?

$$C_{4,2} = \frac{4!}{2!(4-2)!} = \frac{4.3.2.1}{2.1.2.1} = 6$$

A) Distribuições Contínuas

A.1) Distribuição Normal

Uma das mais utilizadas:



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

Padronização: distribuição padrão normal N (μ, σ^2) $\rightarrow$ N (0,1)

$$z = \frac{x - \mu}{\sigma}$$

onde: x ... amostra

μ ... média populacional ($-\infty < \mu < \infty$)

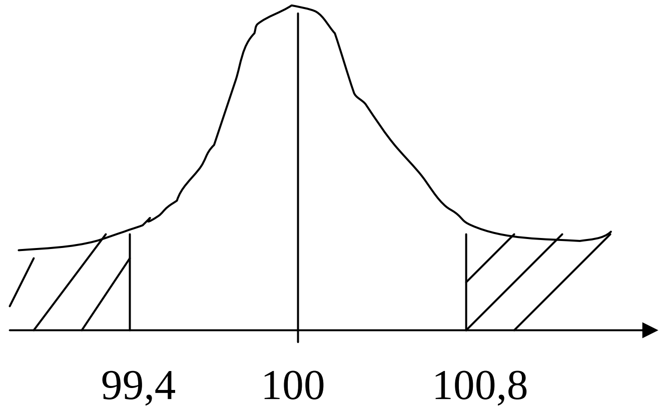
σ^2 ... variância da população ($\sigma^2 > 0$)

A padronização transforma uma variável aleatória com distribuição N (μ,σ^2) em uma distribuição padrão N (0,1).

Exemplo: A distribuição dos diâmetros de uma peça segue uma distribuição normal com μ = 100 mm e desvio padrão de 0,3 mm. Qual a porcentagem de peças que se encontra fora da especificação 100,1 $\pm$ 0,7?

Solução: Xmin = 100,1 – 0,7 = 99,4

Xmax = 100,1 + 0,7 = 100,8



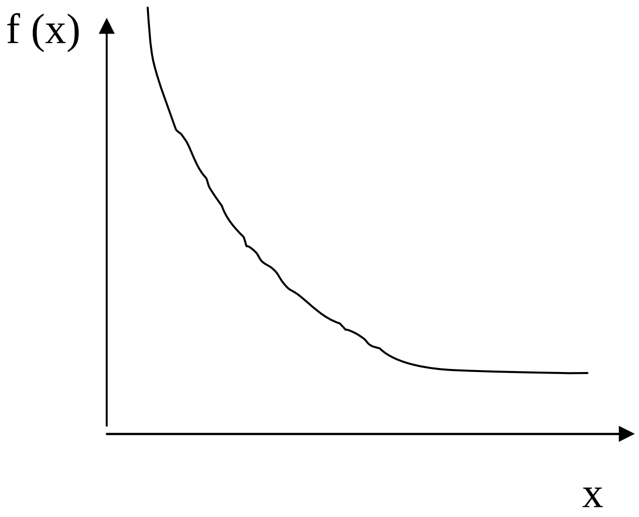
$$z = \frac{99,4 - 100}{0,3} = -2,00 \xrightarrow{\text{tabela}} 0,0228$$

$$z = \frac{100,8 - 100}{0,3} = 2,67 \xrightarrow{\text{tabela}} 0,9962$$

P (fora da especificação) = (1 – 0,9962) + 0,0228

P (fora da especificação) = 2,66%

A.2) Distribuição Exponencial



f (x) = $\lambda\, e^{-\lambda x}$ para $x \geq 0$

$$P(x > x_o) = \int_{X_o}^{x} \lambda e^{-\lambda x} dx = e^{-\lambda x_o}$$

onde: λ ... taxa de acontecimentos de sucesso

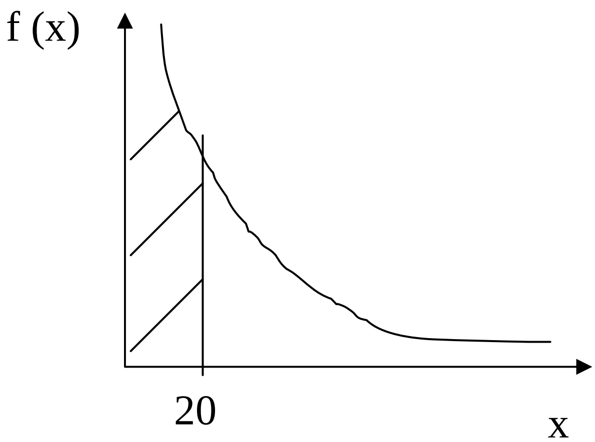
A média e a variância da distribuição exponencial são:

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

$$\sigma^2 = \frac{1}{\lambda^2}$$

Exemplo: Uma fábrica produz lâmpadas com uma duração de vida que pode ser considerada com uma distribuição exponencial com média 200 horas.

a) Qual a probabilidade de uma lâmpada queimar antes de 20 horas de uso?



Solução: μ = 200 h

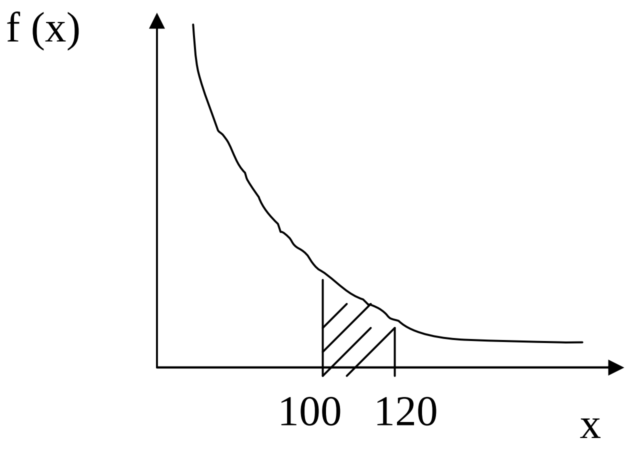
x = 20 h

$$\lambda = \frac{1}{\mu} = \frac{1}{200} \, queimas/h$$

$$P(x < 20) = 1 - e^{-\frac{1}{200}.20} = 0{,}0952$$

b) Qual a probabilidade de uma lâmpada queimar entre 100 e 120 h

Solução:



$$P(100 < x < 20) = e^{-\frac{1}{200}.100} - e^{\frac{-1}{200}.120} = 0{,}6065 - 05488 = 0{,}0577$$

c) Das lâmpadas que duram mais de 100 horas, qual a porcentagem que queimam entre 100 e 120 horas?

Solução:

$$P\left(\frac{100 < x < 20}{x > 100}\right) = \frac{0,0577}{0,6065} = 0,0952$$

B) Distribuições Discretas

B.1) Distribuição Hipergeométrica

A variável aleatória da distribuição hipergeométrica é:

X= nº de peças defeituosas numa amostra, retirada sem reposição de uma população finita.

N = nº de peças de um lote

D = nº de peças defeituosas deste lote

n = tamanho da amostra

A probabilidade de saírem K peças defeituosas na amostra é:

$$P(X = K) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-K}}{\binom{N}{n}}$$

onde:

$\binom{N}{n}$ ... nº de amostras distintas de tamanho n, que podem ser retiradas de uma população de tamanho N

$\binom{D}{K} \cdot \binom{N - D}{n - K}$ ... nº de amostras com exatamente K peças defeituosas.

Exemplo: Um lote de 50 peças contém exatamente 20% de peças defeituosas. Retirada uma amostra de 5 peças, qual a probabilidade de mais de 20% das peças da amostra serem defeituosas?

Solução:

N = 50

D = 0,2 x 50 = 10

n = 5

x = ?

P (X > 0,2 x 5) = (X > 1) = 1- P (X = 0) – P (X =1)

$$P(X = 0) = \frac{\binom{D}{k} \cdot \binom{N-D}{n-K}}{\binom{N}{n}} = \frac{\binom{10}{0}\binom{40}{5}}{\binom{50}{5}} = 0{,}3106$$

$$P(X=1)=\frac{\binom{10}{1}.\binom{40}{4}}{\binom{50}{5}}=0,4313$$

$$P(X>1)=1-0,3106-0,4313=0,2581$$

B.2) Distribuição Binomial

Condições:

a) **n** provas independentes

b) cada prova admite sucesso ou fracasso

c) a probabilidade de sucesso é **p**

d) X = nº de sucessos nas **n** provas

Função probabilidade:

$$P(X=K)=\binom{n}{k}p^{K}(1-p)^{n-K}$$

onde: K ... nº de ordens que pode-se ter

$p^k(1-p)^{n-K}$ ... probabilidade de nas **n** provas acontecer K sucessos e n-K fracassos numa particular ordem

$\binom{n}{K}$ ... combinação, n, K

Exemplo: Da produção mensal de uma máquina foi retirada uma amostra de 5 peças. Sabe-se que esta máquina apresenta uma porcentagem de peças defeituosas constante ao longo do tempo e igual a 15%.

a) Qual a probabilidade de dentre as 5 peças exatamente 2 serem defeituosas?

b) Qual a probabilidade de mais de 2 serem defeituosas?

Solução:

a)

$$P(X=2)=\binom{n}{k}p^K(1-p)^{n-K}=\binom{5}{2}0{,}15^2(1-0{,}15)^{5-2}=$$

$$=10.0{,}15^2(0{,}85)^3=0{,}1382$$

b) P (X > 2) = P (X = 3) + P (X = 4) + (X = 5)

= 1 – P (X = 0) – P (X = 1) – P (X = 2)

$$P(X = 0) = \begin{pmatrix} 5 \\ 0 \end{pmatrix} .0{,}15^0 (0{,}85)^5 = 0{,}4437$$

$$P(X = 1) = \begin{pmatrix} 5 \\ 1 \end{pmatrix} .0{,}15^1 .(0.85)^4 = 0{,}3915$$

$$P(X > 2) = 1 - 0{,}4437 - 0{,}3915 - 0{,}1382 = 0{,}0266$$

B.3) Distribuição Poisson

Condições:

a) sucessos acontecem independentemente

b) taxa de acontecimento de sucesso (λ) é constante em todo o intervalo t

c) X = nº de sucessos no intervalo t

d) t = intervalo contínuo de tempo, área, volume, etc.

Função probabilidade é:

$$P(X = K) = \frac{e^{-\lambda t} (\lambda t)^K}{K!} = \frac{e^{-\mu} \mu^K}{K!}$$

onde: μ ... média (μ = λ.t)

Exemplo: Na inspeção de chapas de aço são retiradas amostras de $2m^2$ e examinadas quanto ao número de imperfeições encontradas.

As chapas com duas ou mais imperfeições na amostra são rejeitadas. Se a média do processo de fabricação for de 0,75 imperfeições por metro quadrado, qual a probabilidade de uma chapa ser rejeitada?

Solução: $\lambda = 0,75$ imperfeições/m^2

$t = 2\ m^2$

$\mu = 2 \times 0,75 = 1,5$ imperfeições por amostra

X = nº de imperfeições (sucessos) numa amostra de $2\ m^2$

rejeição = 2 ou mais imperfeições

P (rejeitar) = P (X ≥ 2) = 1- P (X = 0) – P (X = 1)

$$P(X = 0) = \frac{e^{-1,5}1,5^{0}}{0!} = 0,2231$$

$$P(X = 1) = \frac{e^{-1,5}1,5^{1}}{1!} = 0,3347$$

$$P(X \geq 2) = 1 - 0,2231 - 0,3347 = 0,4422$$