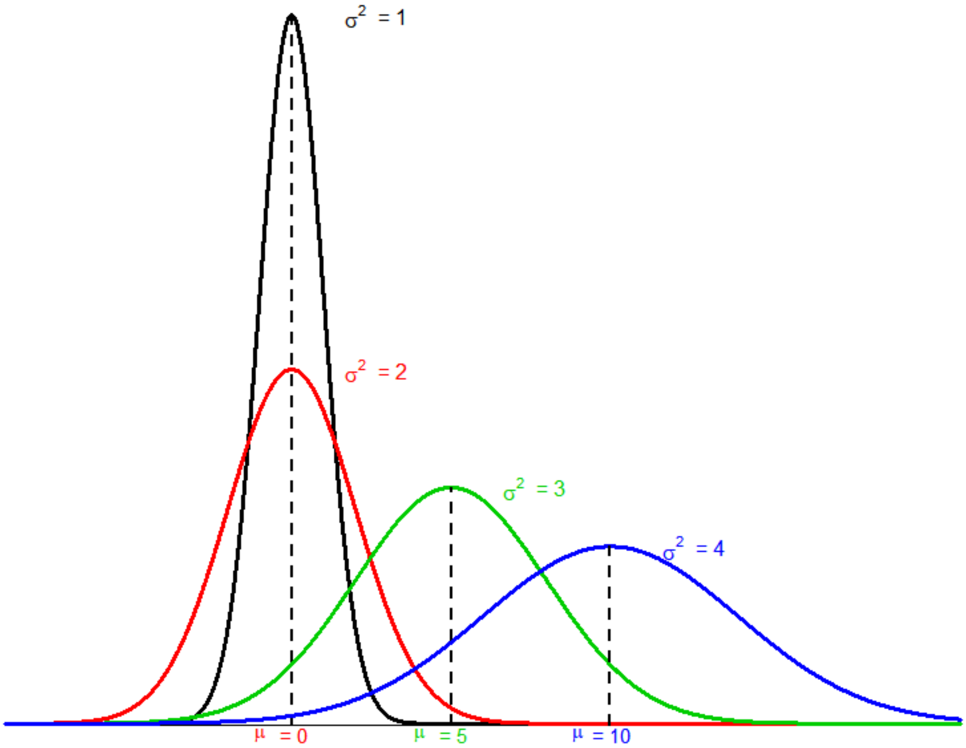
Universidade Federal do Ceará

Campus de Russas



Prof^{a}.: Rosineide F. da Paz

Russas - Ce
Maio–2019

Sumário

1 Introdução **4**

1.1 Conceitos Básicos 5

2 Estatística Descritiva **7**

2.1 Variável 7

2.2 Exercícios para a Seção 2.1 8

2.3 Estrutura dos Dados e Notação 9

2.3.1 Notação 10

2.4 Distribuição de Frequência 10

2.4.1 Tabelas de Frequência 10

2.4.2 Exercícios para a Seção 2.4.1 15

2.4.3 Gráficos de Frequência 16

2.4.4 Exercícios para a Seção 2.4.3 19

2.5 Medidas de Resumo 20

2.5.1 Medidas de Tendencia Central 20

2.5.2 Separatrizes 27

2.5.3 Boxplot 28

2.5.4 Exercícios para a Seção 2.5 30

2.6 Medidas de Dispersão 31

2.6.1 Variância 32

2.6.2 Desvio-padrão 33

2.6.3 Distância interquartil 33

2.6.4 Amplitude total 34

2.6.5 Coeficiente de Variação 34

2.6.6 Exercícios para a Seção 2.6 37

3 Análise Combinatória **40**

3.1 O princípio básico da contagem 41

3.2 Permutações 42

3.3 Arranjo 42

3.4 Combinação 43

3.5 Exercícios 45

4 Introdução as Teorias de Probabilidade **46**

4.1 Experimento Aleatório, Espaço Amostral e Eventos 47

4.1.1 Experimentos 47

4.1.2 Tipos de Espaço Amostral 48

4.1.3 Operações com Eventos 48

4.1.4 Partição do espaço amostral 50

4.1.5 Exercícios para Seção 4.1 51

4.2 Definições de probabilidade 52

4.2.1 Propriedades . . . 53
4.2.2 Probabilidade Condicional . . . 54
4.2.3 Eventos Independentes . . . 55
4.2.4 Exercícios para Seção 4.2 . . . 55
4.3 Lei da Probabilidade Total e Teorema de Bayes . . . 57
4.3.1 Exercícios para a Seção 4.3 . . . 59

5 Variáveis Aleatórias **61**
5.1 Variáveis Aleatórias Discretas . . . 62
5.1.1 Função de Probabilidade . . . 65
5.1.2 Esperança, Média ou Valor Esperado de uma v.a. Discreta . . . 67
5.1.3 Variância de um v.a. Discreta . . . 69
5.2 Exercícios para a Seção 5.1 . . . 70
5.3 Principais Modelos Probabilísticos Discretos . . . 71
5.3.1 Modelo Uniforme Discreto . . . 71
5.3.2 Modelo de Bernoulli . . . 72
5.3.3 Modelo Binomial . . . 73
5.3.4 Modelo Hipergeométrico . . . 74
5.3.5 Modelo de Poisson . . . 75
5.4 Exercícios para a Seção 5.3 . . . 77
5.5 Variáveis Aleatórias Contínuas . . . 79
5.5.1 Função de Distribuição Acumulada . . . 82
5.5.2 Esperança e Variância de um v.a. Contínuas . . . 82
5.5.3 Principais Modelos Contínuos . . . 83
5.5.4 Exercícios para a Seção 5.5 . . . 87

Capítulo 1

Introdução

A estatística consiste numa metodologia científica para obtenção, organização, redução, análise e interpretação de dados oriundos das mais variadas áreas das ciências experimentais, cujo objetivo principal é auxiliar a tomada de decisão em situações de incerteza.

Embora não se trate de ramos isolados, basicamente, podemos dividir a estatística em duas áreas:

- Estatística Descritiva: Conjunto de técnicas que objetivam, organizar, resumir, analisar e interpretar os dados sob consideração.
- Estatística Inferencial: Processo de obter informações sobre uma população a partir de resultados observados na amostra.

> Mas como surgem os dados?

Os dados são resultados de observações de ocorrências de variáveis, podendo surgir a partir de observações espontâneas ou com a realização de um experimento planejado.

- Dados oriundos de observações de fenômenos quaisquer:
  - observar o desempenho natural de um novo equipamento.
- Dados oriundos de experimentos planejados:
  - observar o desempenho de novos equipamentos, alterando de modo proposital somente a variável de interesse.

> Se o interesse é tirar conclusões sobre um conjunto maior que o conjunto observado, devemos considerar técnicas de amostragem para obter uma amostra que seja representativa desse conjunto maior.

A estatística é de grande utilidade quando o método científico é utilizado (quando são realizados os experimentos planejados) para testar teoria ou hipóteses em muitas áreas do conhecimento. Esse método pode ser resumido nos seguintes passos.

1. Um problema é formulado em que, muitas vezes, uma hipótese precisa ser testada.
2. Para solucionar o problema, deve-se coletar informações que sejam relevantes, para isso pode-se formular um experimento. Em muitas áreas do conhecimento o planejamento do experimento não é simples, ou até mesmo não é possível, e uma estratégia pode ser a observação de algum fenômeno (variável) de interesse.

3. Os resultados do experimento podem ser utilizados para se obter conclusões, definitivas ou não.

4. os passo 2 e 3 podem ser repetidos quanta vezes forem necessárias.

É notável que nos passos descritos acima a estatística seja uma ferramenta indispensável, podendo ser requerida em todas as etapas.

1.1 Conceitos Básicos

- **População:** consiste em um conjunto de elementos que compartilham de pelo menos uma característica comum.

- **Unidade amostral:** é a entidade (ou elemento) da população sobre a qual a característica de interesse poderá ser observada.
  As unidades amostrais podem ser os próprios elementos da população ou podem ser formadas por grupos de elementos, compondo o que é chamado de **conglomerado**.

  Exemplo 1.1.1. *Conglomerados podem ser formados por:*

  $\Rightarrow$ *quarteirões;*

  $\Rightarrow$ *ruas (face dos quarteirões);*

  $\Rightarrow$ *departamentos;*

  $\Rightarrow$ *prateleiras;*

  $\Rightarrow$ *caixas;*

  $\Rightarrow$ *lotes de produtos; etc.*

- **Amostra:** conjunto de elementos extraídos da população.

- **Censo:** É o processo utilizado para levantar as características observáveis, abordando todos os elementos de uma população.

  Exemplo 1.1.2. *Como exemplos de população, podemos citar:*

  - *Pesquisa de opinião pública:*
    - *a população é o total de habitantes de um local;*
    - *a amostra é uma parte dessa população.*
  - *Em sucessivos lançamentos de uma moeda:*
    - *a população é formada por todos os resultados possíveis, cara ou coroa em cada lançamento, aqui a população e infinita;*
    - *a amostra é formada pelos resultados obtidos em uma sequência finita de lançamentos.*
  - *Investigar a porcentagem de lajotas defeituosas fabricadas em uma indústria, durante 6 dias, examinando 20 peças por dia.*
    - *População: todos as lajotas fabricadas durante 6 dias.*
    - *Amostra: o subconjunto de* $6 \times 20 = 120$ *peças, selecionadas para estudo.*

> Na prática não podemos utilizar qualquer amostra para o propósito de inferência. A extração de uma amostra pode ser feita de várias maneiras e deve seguir algumas regras, dependendo do problema que se deseja tratar.

A Figura 1.1 mostra de forma esquemática e resumida as possíveis etapas de uma análise estatística. Note que existem casos em que apenas uma análise descritiva (exploratória) dos dados é suficiente para tiramos conclusões a respeito da população de interesse. No entanto, se a população é maior do que os dados analisados, sob determinadas condições, podemos fazer uso das teorias das probabilidades para fazer inferência sobre as características (parâmetros) da população de interesse.

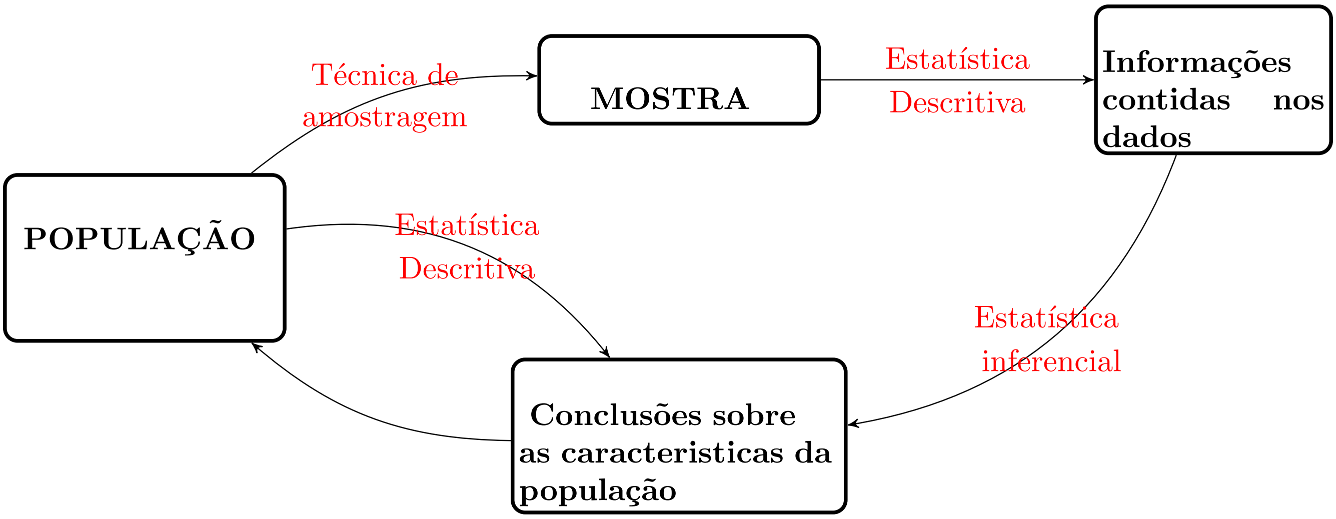


Figura 1.1: Análise estatística a partir da amostra ou da população.

Independentemente de estarmos diante de uma amostra ou de uma população, ao analisar um conjunto de dados devemos sempre fazer uma analise exploratória fazendo uso de ferramentas da **estatística descritiva**. Embora essa exploração ocorra de forma semelhante para população ou amostra, vamos utilizar notações diferentes para indicar se estamos diante de uma população ou de uma amostra.

Capítulo 2

Estatística Descritiva

Antes de qualquer análise estatística mais sofisticada, devemos realizar uma análise exploratória dos dados fazendo uso das ferramentas da **Estatística Descritiva**. Nessa etapa da análise, procuramos obter a maior quantidade possível de informações dos dados observados. Se estamos diante de uma amostra, é nessa fase que devemos obter informações sobre qual modelo probabilístico descreve o fenômeno investigado e pode ser utilizado em uma fase posterior denominada **Inerência Estatística**.

2.1 Variável

Os dados são observações de variáveis ou fenômeno de interesse. Uma variável é uma quantidade ou atributo cujo valor pode variar de uma unidade de investigação para outra. Por exemplo, as unidades podem ser pessoas portadoras de dor de cabeça e a variável o tempo entre tomar um remédio e sessar a dor. Uma observação, ou resposta, é o valor assumido por uma variável em uma das unidades investigadas. A observação da variável em várias unidades dá origem aos dados observados.

Exemplo 2.1.1. *Alguns exemplos de variáveis são:*

- *tempo de execução de um algoritmo;*
- *rendimento das famílias de uma grande cidade;*
- *número de erros em pacotes de dados enviados por um servidor;*
- *número de clientes com a mesma dúvida em um site de suporte durante um período de tempo;*
- *opinião dos consumidores de um determinado produto (péssimo, regular, ótimo etc).*

Existem vários tipos de variáveis, sendo que inicialmente podemos dividi-las em qualitativas e quantitativas. As variáveis são **qualitativa** quando seus valores forem expressos por atributos (não numéricas). Este grupo pode ser subdividido em: qualitativa nominal e ordinal. A variável é nominal se os atributos que ela representa não têm uma ordenação, por exemplo, cor de cabelo, sexo de indivíduos etc, enquanto que as ordinais exprimem alguma ordenação, como por exemplo, opinião sobre a qualidade de um produto (péssimo, regular, ótimo). As variáveis **quantitativas:** assumem valores numéricos. Essas variáveis também podem ser classificadas em dois grupos: contínua e discreta. As variáveis são discretas se assumem valores em um conjunto enumerável (contável), como por exemplo, número de carros que passam por um posto de pedágio em um intervalo de tempo. As variáveis contínuas assumem valores em um conjunto não-enumerável, ou seja, em um intervalo da reta, como por exemplo, alturas de pessoas de um determinado povoado. Veja um resumo das variáveis no esquema a seguir.

$$\text{Variável} \longmapsto \begin{cases} \text{Qualitativa} & \longmapsto \begin{cases} \text{nominal (Ex. Região de procedencia)} \\ \text{ordinal (Ex. Grau de instrução)} \end{cases} \\ \text{Quantitativa} & \longmapsto \begin{cases} \text{discreta (Ex. número de filhos)} \\ \text{contínua (Ex. Altura)} \end{cases} \end{cases}$$

Em geral, as medições dão origem às variáveis contínuas e as contagens ou enumerações às variáveis discretas.

2.2 Exercícios para a Seção 2.1

Exercício 2.2.1. *Classifique as seguintes variáveis em qualitativas (nominal/ordinal) ou quantitativa (discreta/contínua).*

a) *Classe social.*

b) *Número de clientes em um estabelecimento.*

c) *Salário mensal.*

d) *Cidade de nascimento.*

e) *Departamento que trabalha.*

f) *Número de filho.*

g) *Nível de escolaridade.*

h) *Número de processos analisados.*

i) *Opinião sobre a reforma agrária.*

j) *Opinião sobre atendimento de um estabelecimento.*

k) *Número de telefonemas recebidos.*

l) *Estado Civil.*

m) *Idade (anos).*

n) *Distância de sua casa na faculdade (marcado no velocímetro).*

o) *Número de idas ao cinema por semana.*

Exercício 2.2.2. *Declare se cada uma das seguintes variáveis é do tipo discreta ou contínua:*

1. *O número anual de suicídios no Brasil;*

2. *A concentração de chumbo em uma amostra de água;*

3. *A duração de tempo que um paciente sobrevive depois do diagnóstico de uma doença fatal;*

4. *O número de abortos prévios que uma mãe grávida teve.*

2.3 Estrutura dos Dados e Notação

Os dados, em geral, são dispostos em tabelas, ou planilhas, de modo que em cada coluna podem ser observados os valores observados de uma única variável. Assim, nas linhas da tabela estão os valores observados para cada variável. Essa estrutura simples permite que os dados possam ser analisados por meio de diversos software, tais como library calc, Rstudio, entre outros.

Exemplo 2.3.1. *Suponha, por exemplo, que um questionário foi aplicado aos alunos de um curso da UFC, fornecendo as seguintes informações: Idade em anos; Altura em metros; Peso em quilogramas; Estado Civil: Solteiro, casado, divorciado e viúvo, como mostra o Quadro a seguir.*

Estado civil	*Idade*	*Peso*	*Altura*
solteiro	*20*	*74*	*1,68*
solteiro	*18*	*46*	*1,6*
solteiro	*19*	*62*	*1,6*
solteiro	*19*	*64*	*1,7*
solteiro	*25*	*98*	*1,9*
solteiro	*24*	*68*	*1,72*
solteiro	*20*	*60*	*1,7*
solteiro	*35*	*71*	*1,68*
solteiro	*19*	*67*	*1,62*
solteiro	*20*	*79*	*1,87*
solteiro	*19*	*80*	*1,75*
solteiro	*20*	*65*	*1,74*
solteiro	*20*	*74*	*1,6*
solteiro	*20*	*65*	*1,7*
solteiro	*19*	*53*	*1,63*
solteiro	*19*	*60*	*1,67*
solteiro	*23*	*45*	*1,6*
divorciado	*26*	*70*	*1,7*
solteiro	*20*	*75*	*1,7*
solteiro	*21*	*75*	*1,7*
solteiro	*19*	*73*	*1,76*
casado	*46*	*70*	*1,7*
solteiro	*19*	*70*	*1,78*
solteiro	*28*	*58*	*1,75*
solteiro	*21*	*68*	*1,6*
solteiro	*23*	*62*	*1,7*
solteiro	*19*	*66*	*1,74*
solteiro	*20*	*74*	*1,8*
solteiro	*22*	*90*	*1,86*
casado	*58*	*98*	*1,8*
solteiro	*24*	*74*	*1,73*
solteiro	*20*	*70*	*1,7*
casado	*26*	*95*	*1,6*

...

solteiro	*20*	*46*	*1,54*
solteiro	*21*	*69*	*1,57*
solteiro	*19*	*57*	*1,57*
solteiro	*19*	*59*	*1,61*
solteiro	*17*	*58*	*1,49*
solteiro	*20*	*62*	*1,7*
solteiro	*20*	*60*	*1,65*
solteiro	*22*	*49*	*1,6*

O conjunto de informações disponíveis, após a tabulação do questionário ou pesquisa de campo, é denominado tabela de dados brutos e contém os dados da maneira que foram coletados inicialmente, após a crítica dos valores. Em nosso caso temos quatro variáveis envolvidas sendo uma qualitativa (estado civil) e as restantes quantitativas (idade, peso, altura).

2.3.1 Notação

Uma variável qualquer será denotada por uma letra maiúscula, como por exemplo X e uma sequência de valores observados dessa variável será denotada por letras minúscula, de modo que:

- $x_1, \cdots, x_n$ representa uma sequência de valores observados da variável X em uma amostra e
- $x_1, \cdots, x_N$ representa uma sequência de valores observados da variável X em uma população inteira,

em que n denota o tamanho da amostra e N denota o tamanho da população.

2.4 Distribuição de Frequência

O objetivo da Estatística Descritiva é resumir as principais características dos dados observados fazendo uso de tabelas, gráficos e resumos numéricos, para se ter uma ideia do comportamento da variável estudada. Pois, quando se estuda uma variável, o maior interesse é conhecer o seu comportamento, sua frequência, ou sua **distribuição de frequência**.

2.4.1 Tabelas de Frequência

Para se ter uma ideia dessa distribuição, podemos construir uma tabela de frequência para o conjunto de observações dessa variável.

Exemplo 2.4.1. *Suponha que observamos as notas finais (por conceito) de 30 alunos de um determinado curso e obtivemos os seguintes valores:*

C B B B B A C B B A D D B C B B C B C B C C C B B C B B A C

A variável de interesse é o conceito (A, B, C ou D, em que D significa a reprova do aluno). Será que, de um modo geral, a turma teve um bom desempenho?

Para responder essa questão, devemos observar a frequência da variável "conceito". Essa frequência fica evidente se os dados forem dispostos em uma tabela apropriada. Em particular, para esse conjunto de dados, podemos utilizar uma tabela de frequência simples. A Tabela 2.2 apresenta a frequência absoluta (n_i) e a frequência relativa (f_i) da variável conceito. A frequência relativa indica a proporção de vezes que um determinado valor da variável aparece, por exemplo, note que o conceito

B corresponde a 50% dos valores observados da variável. Em outras palavras, 50% dos estudantes obtiveram conceito B neste curso.

Tabela 2.2: Distribuição de frequência para a variável conceito.

Conceito	Frequência absoluta (n_i)	Frequência relativa (f_i)
A	3	$(3/30) = 0,1$
B	15	$(15/30) = 0,5$
C	10	$(10/30) \cong 0,33$
D	2	$(2/30) \cong 0,07$
Total (n)	30	1

Existem duas possibilidades para a construção de tabelas de frequência:

(i) tabelas de frequências simples;

(ii) e tabelas de frequências em intervalos de classes.

A tabela de frequência simples é usada para variáveis qualitativas e quantitativas discretas com poucos valores possíveis. A tabela em intervalos de classe é apropriada para variáveis quantitativas contínuas (ou discretas com muitos valores possíveis).

Tabela de Frequência Simples

Essa tabela é apropriada para para variáveis qualitativas ou quantitativas discretas com poucos valores possíveis. O formato geral para esse tipo de tabela pode ser visto na Tabela 2.3, em que:

- $x_1, \cdots, x_k$ representam os valores distintos e ordenados que podem ser encontrados no conjunto de dados.
- $n_1, \cdots, n_k$ representam a contagem das repetições de cada valor distinto, e são denominadas frequências absolutas, com $n_1 + n_2 + \cdots + n_k = n$.
- $f_1, \cdots, f_k$ são as frequências relativas (ou proporções).
- f_{ac} representa a frequência relativa acumulada, também podemos incluir uma coluna contendo os valores de frequências absolutas acumuladas (n_{ac}).
- n: número de observações (ou tamanho da amostra, caso sejam dados de uma população usamos N).
- k: número de classes na tabela de frequência.

Tabela 2.3: Formato geral para uma tabela de frequência simples.

Variável	(n_i)	(f_i)	f_{ac}
x_1	n_1	$f_1 = n_1/n$	n_1/n
x_2	n_2	$f_2 = n_2/n$	$(n_1 + n_2)/n$
$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_k	n_k	$f_k = n_k/n$	$(n_1 + n_2 + \cdots + n_k)/n.$
Total	n	1	

Exemplo 2.4.2. *: Considerando os dados da variável Estado Civil do Exemplo 2.3.1, temos uma tabela de frequência simples com $k = 4$ classes, como pode ser visto na Tabela 2.4. Nesta tabela, acrescentarmos a coluna de frequência relativa, definida por $f_i = n_i/n$, com isso podemos ver rapidamente que 90% dos estudantes entrevistados são solteiros.*

Tabela 2.4: Tabela de frequência para a variável Estado Civil.

Estado Civil	n_i	f_i
Solteiro	37	0,90
Casado	03	0,07
Divorciado	01	0,02
Viúvo	00	0,00

Considerando, novamente, os dados de estudantes apresentados na Tabela 2.3.1, vamos averiguar em torno de que valor tendem a se concentrar as estaturas obtidas, qual a menor ou qual a maior estatura, ou ainda, quantos alunos se acham abaixo ou acima de uma dada estatura. Observe que olhar diretamente para a tabela de dados brutos é difícil se ter uma ideia do comportamento da variável altura considerando o grupo inteiro de estudantes. Novamente podemos interpretar mais facilmente esse conjunto de dados fazendo uso de uma tabela de frequência. Considerando a Tabela de Frequência Simples 2.5, percebemos que esta se mostra muito longa e difere pouco da tabela de dados brutos. Assim, é conveniente montar outra estratégia para resumir esses dados em uma tabela de frequência.

Tabela 2.5: Tabela de frequência simples para a variável Altura.

Variável Altura	**Frequência absoluta (n_i)**
1,49	1
1,54	1
1,57	2
1,60	7
1,61	1
1,62	1
1,63	1
1,65	2
1,67	1
1,68	2
1,70	10
1,72	1
1,73	1
1,74	2
1,75	1
1,76	1
1,78	1
1,80	2
1,86	1
1,87	1
1,90	1

Observe que o processo empregado para variável Altura na Tabela 2.5 é inconveniente, pois exige muito espaço, mesmo quando o número de valores da variável (n) não é muito grande, e não nos esclarece muita coisa. Desta forma, o melhor seria formar agrupamentos. Assim, em vez de trabalharmos com os valores observados da variável, podemos formar intervalos que possam conter esses valores.

Tabelas de Frequências em Intervalos de Classes

Agora vamos construir uma tabela de frequência apropriada para a variável altura dos estudantes. Neste caso temos uma variável que não tem uma natureza discreta, então, é natural não haver muitas repetições no conjunto de dados. Assim, devemos construir uma tabela em intervalos de classes da seguinte procedendo da seguinte forma:

- Determina-se o número de classes, k, fazendo $k \approx \sqrt{n}$, se n é grande, podemos utilizar a regra de Sturges em que
$$k \approx 1 + 3,3 \times \log n,$$
ou ainda determinar esse valor conforme seja mais apropriado.
- definimos L_{inf} um valor menor ou igual ao valor mínimo ($L_{inf} \leq$ **valor mínimo**);
- definimos L_{sup} um valor maior o igual ao valor máximo ($L_{sup} \geq$ **valor máximo**);
- Obtemos a amplitude das classes $AT = L_{sup} - L_{inf}$;
- finalmente, obtemos $\Delta = AT/k$, amplitude de cada classe.

Para os dados da variável Altura, vamos fixar $L_{sup} = 1,98$ e $L_{inf} = 1,48$ para determinar
$$AT = L_{sup} - L_{inf} = 1,98 - 1,48 = 0,50$$
Embora $n = 41$ e $\sqrt{41} \approx 6,4$, aqui vamos assumir (por conveniência) que $k = 5$ que fornece $\Delta = 0,50/5 = 0,1$. Logo, obtemos os seguintes **limites das classes**:

- $1,48 + 0,1 = 1,58 \quad \Rightarrow$ **classe**$_1$: $1,48 \dashv 1,58 = (1,48; 1,58]$
- $1,58 + 0,1 = 1,68 \quad \Rightarrow$ **classe**$_2$: $1,58 \dashv 1,68 = (1,58; 1,68]$
- $1,68 + 0,1 = 1,78 \quad \Rightarrow$ **classe**$_3$: $1,68 \dashv 1,78 = (1,68; 1,78]$
- $1,78 + 0,1 = 1,88 \quad \Rightarrow$ **classe**$_4$: $1,78 \dashv 1,88 = (1,78; 1,88]$
- $1,88 + 0,1 = 1,98 \quad \Rightarrow$ **classe**$_5$: $1,88 \dashv 1,98 = (1,88; 1,98]$

Observe que usamos duas notações para intervalos, sendo $1,48 \dashv 1,58$ um intervalo que vai desde 1,48 até 1,58 aberto em 1,48 e fechado em 1,58, assim como $(1,48; 1,58]$. Aqui vamos usar a notação com $\dashv$.

Para construir a coluna das frequências absolutas, a parti do hol (dados ordenados), contamos quantos elementos pertencem a cada intervalo de classe obtido. A distribuição de frequência para a variável Altura pode ser vista na Tabela 2.6 e os dados ordenados são:
(1,49; 1,54; 1,57; 1,57; 1,60; 1,60; 1,60; 1,60; 1,60; 1,60; 1,60; 1,61; 1,62; 1,63; 1,65; 1,67; 1,68; 1,68; 1,70; 1,70; 1,70; 1,70; 1,70; 1,70; 1,70; 1,70; 1,70; 1,70; 1,72; 1,73; 1,74; 1,74; 1,75; 1,75; 1,76; 1,78; 1,80; 1,80; 1,86; 1,87; 1,90).

Tabela 2.6: Frequência dos alunos segundo sua altura.

X	n_i	f_i	f_{ac}
$1,48 \dashv 1,58$	4	$4/41 \approx 0,10$	$4/41 \approx 0,10$
$1,58 \dashv 1,68$	14	$14/41 \approx 0,34$	$18/41 \approx 0,44$
$1,68 \dashv 1,78$	18	$18/41 \approx 0,44$	$36/41 \approx 0,88$
$1,78 \dashv 1,88$	4	$4/41 \approx 0,10$	$40/41 \approx 0,98$
$1,88 \dashv 1,98$	1	$1/41 \approx 0,02$	$1/41 = 1$
Total	41	1	

Note que, na Tabela 2.6, foi incluída uma coluna extra para uma sequência denominada **frequência relativa acumulada** (f_{ac}), essas frequências são importantes para qualquer tipo de variável que contém uma certa ordenação, pois a partir dessas frequências podemos tirar conclusões sobre quantos por cento dos valores estão acima ou abaixo de um determinado valor observado da variável. Como exemplo, podemos notar nesta tabela que aproximadamente 88% dos estudantes tem altura até 1,78 m, isso quer dizer que apenas 12% tem mais de 1,78 m.

Exceto pealas classes, a tabela de frequência em intervalos de classes é construída da mesma maneira da tabela de frequência simples. Sua interpretação também é similar, apenas devemos levar em consideração que os valores possíveis não são precisos.

Exemplo 2.4.3. *A Tabela 2.7 categoriza visitas ao consultório de doenças cardiovasculares por duração de cada visita (em minutos). Uma duração de 0 (zero) minuto implica que o paciente não teve contato direto com o especialista.*

Tabela 2.7: Tempo (minutos) de duração de visita ao cardiologista de um grupo de 10614 pessoas.

Duração (minutos)	Número de visitas (n_i)	f_i	f_{ac}
0	390	0,036	0,036
$1 \vdash 6$	227	0,02	0,056
$6 \vdash 11$	1023	0,09	0,146
$11 \vdash 16$	3390	0,31	0,456
$16 \vdash 21$	4431	0,41	0,866
$21 \vdash 26$	968	0,09	0,956
$26 \vdash 31$	390	0,03	0,986
Duração≥ 31	185	0,015	1
Total	10614		

Observe que essa é uma situação em que o limite superior da última classe não é determinado na tabela. Observando a Tabela 2.7 podemos procurar responder as seguintes questões:

1. *Qual o tamanho da amostra?*
2. *Pode-se fazer a afirmação de que as visitas a consultórios de especialistas de doenças cardiovasculares têm duração mais frequente entre 16 e 21 minutos?*
3. *É razoável que a secretária agende uma nova consulta a cada 10 minutos?*
4. *Se não for satisfatório agendar uma nova consulta a cada 10 minutos, qual seria o tempo entre o agendamento de duas consultas que você julga satisfatório?*

Como uma tabela deve aparecer em um texto?

1. Uma tabela deve sempre ser apresentada com suas laterais abertas, se as laterais são fechadas trata-se de um quadro.

2. O título da tabela deve sempre aparecer no topo, não em baixo.

3. Na parte de baixo, pode-se apresentar a fonte e outras informações.

4. Tabelas são elementos flutuantes em um texto, ela pode ser apresentada em qualquer lugar com sua devida numeração.

5. Ao se fazer referencia a uma tabela em um texto, deve-se usar sua numeração (ex. na Tabela 1 pode ser visto...).

2.4.2 Exercícios para a Seção 2.4.1

Exercício 2.4.1. *A massa (em quilogramas) de 17 trabalhadores de uma empresa com 100 funcionários está registrada a seguir:*

52	73	80	65	50	70	80	65	70	77	82	91	52	68	86	70	80

Com base nos dados obtidos, responda:

(a) *Qual é a variável nessa pesquisa? Ela é discreta ou contínua?*

(b) *Que frequências absolutas têm os valores 65 kg, 75 kg, 80 kg e 90 kg?*

Exercício 2.4.2. *A cantina de uma escola selecionou 50 alunos ao acaso e verificou o número de vezes por semana que eles compravam lanche., obtendo os seguintes resultados:*
0; 2; 2; 4; 3; 2; 2; 1; 1; 2; 1; 1; 0; 1; 1 ; 1; 1; 1; 1; 2; 2; 2; 3; 2; 2; 2; 0; 2; 2; 1; 1; 0; 2; 0; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 2; 1; 2; 5; 4.

(a) *Construa uma tabela de distribuição de frequências absolutas, relativas e frequências relativas acumuladas com esses dados.*

(b) *Qual é a proporção de alunos que compram pelo menos 2 lanches por semana?*

Exercício 2.4.3. *Um hospital tem o interesse em determinar a altura média dos pacientes de uma determinada área e relacioná-la com a incidência de determinada anomalia ortopédica. Foram selecionados 80 pacientes e as alturas (em m) podem ser encontradas abaixo.*

1,72	1,78	1,87	1,86	1,79	1,79	1,83	1,74	1,64	1,62
1,75	1,65	1,75	1,58	1,63	1,77	1,64	1,68	1,66	1,82
1,68	1,80	1,74	1,76	1,74	1,72	1,75	1,89	1,73	1,76
1,72	1,71	1,63	1,81	1,65	1,58	1,63	1,70	1,73	1,57
1,75	1,64	1,73	1,70	1,75	1,56	1,70	1,68	1,68	1,79
1,75	1,71	1,62	1,83	1,72	1,76	1,67	1,82	1,67	1,60
1,67	1,61	1,61	1,67	1,75	1,80	1,70	1,77	1,73	1,77
1,64	1,66	1,74	1,66	1,66	1,79	1,68	1,79	1,69	1,80

Construa a tabela de distribuição de frequências por intervalos de classes, apresentando também a coluna da distribuição de frequências acumulada.

2.4.3 Gráficos de Frequência

Vimos que a distribuição de frequências pode ser representada em tabelas de frequências. Outra opção para representar essas frequências é o uso de gráficos. Os principais gráficos para representação de distribuição de frequências são:

1. Gráficos em barras (apropriado para variáveis qualitativas);
2. Gráficos em setores (apropriado para variáveis qualitativas);
3. Histograma (apropriado para variáveis quantitativas) e
4. Polígono de frequências absolutas.

Muitos outros tipos de gráficos são encontrados, no entanto vários são versões diferentes dos tipos citados acima. Aqui, vamos nos limitar a discutir cada um desses.

Gráficos em barras

Os gráficos em barras são comumente usados para exibir distribuição de frequências para as variáveis qualitativas, como por exemplo a variável Estado Civil do Exemplo 2.3.1, como mostra a Figura 2.1. Podemos também representar a tabela de frequência simples apresentada na Figura 2.1 pelo Gráfico em Setores, como mostra a Figura 2.2 e 2.3. O gráfico em setores é comumente utilizado para representar parte de um todo, geralmente em porcentagens, e é, também, apropriado para variáveis qualitativas.

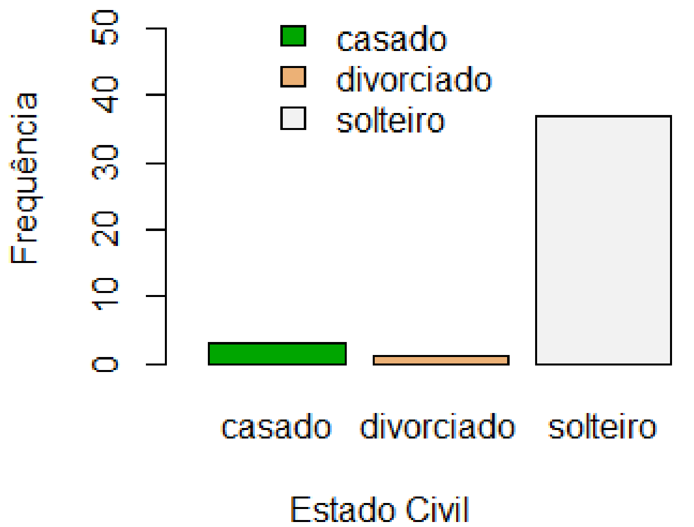


Tabela de frequência simples.

Estado Civil	n_i
Casado	3
Divorciado	1
Solteiro	37
Total	41

Figura 2.1: Gráfico (e tabela) de frequência de estudantes por estado civil.

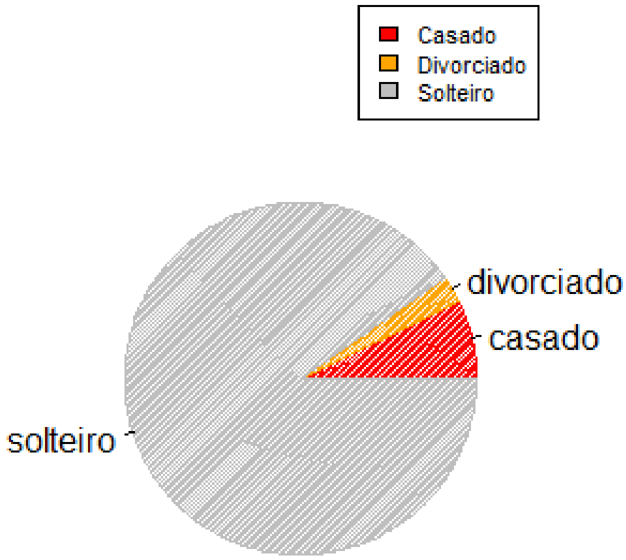


Figura 2.2: Frequência de estudantes por estado civil.

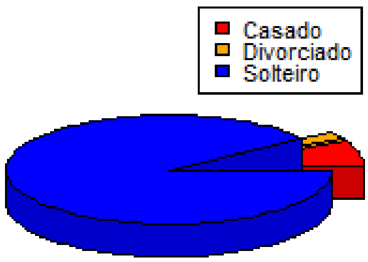


Figura 2.3: Frequência de estudantes por estado civil.

Histograma e o Polígono de Frequência

O histograma é um gráfico de barras contíguas apropriado para representar distribuições de frequências de variáveis quantitativas (contínuas ou discretas com muitos valores possíveis) e é um dos mais importantes gráficos no estudo de frequência de variáveis quantitativa. Sua construção é similar ao procedimento empregado para construir uma tabela de frequências em intervalos de classes.

As barras contíguas do histograma têm bases de mesma amplitude. Caso a tabela de frequência em intervalos de classes já tenha sido construída, as barras do histograma podem ser construídas de modo a serem proporcionais aos intervalos das classes dessa tabela. É importante que as áreas das barras sejam proporcionais as suas correspondentes frequência absoluta (n_i) ou relativa (f_i), sendo mais usual utilizar as frequências relativas (ou as porcentagens, $P_i = 100 \times f_i$).

Para construir o histograma de modo que as áreas das barras sejam dadas pelas frequências relativas, devemos obter a altura (h_i) de cada retângulo (ou barra) de modo que $h_i = f_i/\Delta$ (ou $h_i = n_i/\Delta$ caso seja utilizada a frequência absoluta), em que Δ é a amplitude das bases das barras, para $i = 1, \cdots, k$ com k sendo o número de barras (ou classes na tabela). Deste modo, quanto mais dados contiver a i-ésima classe, mais alto será o i-ésimo retângulo.

Como exemplo, vamos construir um histograma a partir da tabela de frequência para variável altura de um estudante, Tabela 2.7. Para isso, seguiremos os seguintes passos.

- vamos considerar k=5 classes com $L_{inf} = 1,48$ e $L_{sup} = 1,98$;
- obtemos $AT = 1,98 - 1,48 = 0,5$ é a amplitude total considerada;
- em seguida obtemos a largura das barras, ou amplitude, $\Delta = \frac{AT}{k} = \frac{0,5}{5} = 0,1$;
- com essa amplitude obtemos os limites das bases de cada barra de modo semelhante ao que foi feito na tabela de frequência;
- para cada amplitude, obtemos a altura da base de modo que $h_i = f_i/\Delta$.

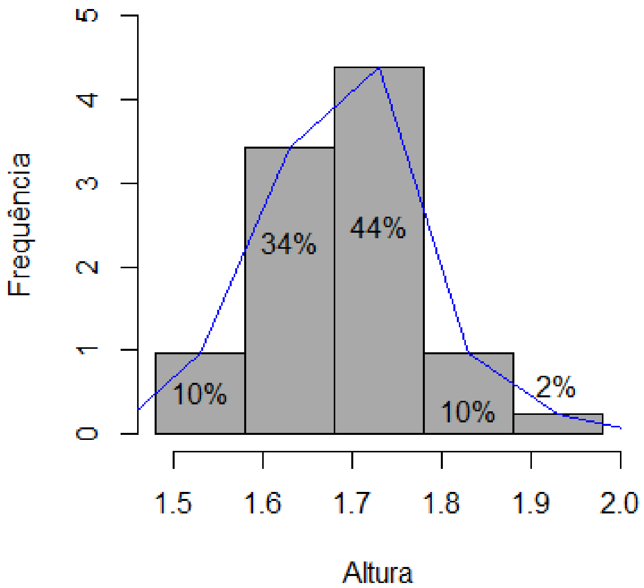


Figura 2.4: Frequência dos estudantes segundo altura.

Frequência dos alunos segundo altura.

X	n_i	f_i
$1,48 \dashv 1,58$	4	$4/41 \approx 0,10$
$1,58 \dashv 1,68$	14	$14/41 \approx 0,34$
$1,68 \dashv 1,78$	18	$18/41 \approx 0,44$
$1,78 \dashv 1,88$	4	$4/41 \approx 0,10$
$1,88 \dashv 1,98$	1	$1/41 \approx 0,02$
Total	41	1

A parti do histograma mostrado na Figura 2.4, podemos perceber que a grande maioria dos estudantes tem entre 1,60 e 1,80 metros de altura, correspondendo à 78% dos indivíduos desse grupo. Iss quer dizer que se sorteássemos ao acaso (aleatoriamente) um indivíduo desse grupo, teríamos 78% de chance de escolhermos alguém com a altura neste intervalo. Note que, pelo ponto médio do topo de cada barra, foi traçado uma linha (curva). Essa curva é denominada **polígono de frequência** e é uma ferramenta bastante útil para para dar uma ideia do comportamento da frequência de ocorrência da variável em questão. O histograma, assim como o polígono de frequência, é denominado densidade empírica da variável.

2.4.4 Exercícios para a Seção 2.4.3

Exercício 2.4.4. *Uma distribuição de frequências para os níveis séricos (em microgramas por decilitro) de zinco de 462 homens entre as idades de 15 a 17 anos é exibida na Tabela 2.8.*

Tabela 2.8: Nível sérico de zinco (mg/dl).

Nível	Número de homens
50 ⊢ 60	6
60 ⊢ 70	35
70 ⊢ 80	110
80 ⊢ 90	116
90 ⊢ 100	91
100 ⊢ 110	63
110 ⊢ 120	30
120 ⊢ 130	5
130 ⊢ 140	2
140 ⊢ 150	2
150 ⊢ 160	2
Total	462

1. *Complete a tabela de frequências. O que você pode concluir sobre essa distribuição de níveis séricos de zinco?*

2. *Elabore um histograma dos dados. Trace o polígono de frequência, observando a sua forma.*

Exercício 2.4.5. *A distribuição de frequência na Tabela 2.9 exibe os números de casos pediátricos de Aids registrados nos EUA entre 1983 e 1989.*

Tabela 2.9: Casos pediátricos de Aids registrados nos EUA.

Ano	Número de casos
1983	122
1984	250
1985	455
1986	848
1987	1412
1988	2811
1989	3098

Construa um gráfico de barras que mostre o número de casos por ano. O que o gráfico lhe conta sobre a Aids pediátrica nesse período?

Exercício 2.4.6. *Contou-se o número de erros de impressão da primeira página de um jornal durante 50 dias, obtendo-se os resultados: 8, 11, 8, 12, 14, 13, 11, 14, 14, 5, 6, 10, 14, 19, 6, 12, 7, 5, 8, 8, 10, 16, 10, 12, 12, 8, 11, 6, 7, 12, 7, 10, 14, 5, 12, 7, 9, 12, 11, 9, 14, 8, 14, 8, 12, 10, 12, 12, 7, 15.*

Represente os dados graficamente.

2.5 Medidas de Resumo

Muitas vezes é útil descrever numericamente as características dos dados observados a partir de variáveis quantitativas, uma vez que sua natureza permite o cálculo de algumas medidas que podem ser úteis para dar uma ideia do comportamento dessas variáveis.

- Dentre as medidas de resumo existentes, destacam-se:
  - as **medidas de posição,**
    - média,
    - moda e
    - separatrizes;
  - as **medidas de dispersão,**
    - variância,
    - desvio-padrão,
    - distância interquartílica e
    - coeficiente de variação.

As medidas de posição localizam a distribuição de frequência da variável no eixo das abcissas, enquanto as medidas de dispersão fornecem informações sobre o "espalhamento" dessa distribuição.

Vamos iniciar nossos estudos sobre as medidas de resumo pelas medidas de posição, as quais podem apresentar-se de várias formas, sendo que as mais importantes são:

- **medidas de tendência central**, são assim denominadas devido a tendência dos dados observados se agruparem em torno desses valores (média, moda e mediana);

- **separatrizes**, são assim chamadas porque separam, dividem um conjunto de dados ordenado em partes percentuais iguais (quartil, decil e percentil, genericamente quantís).

2.5.1 Medidas de Tendencia Central

A Figura 2.5 mostra o histograma para a variável altura dos estudantes construído a partir dos dados brutos, apresentado no Exemplo 2.3.1. Neste gráfico podemos observar geometricamente a média desse conjunto de dados, que se localiza próxima ao centro da distribuição de frequência representada pelo histograma.

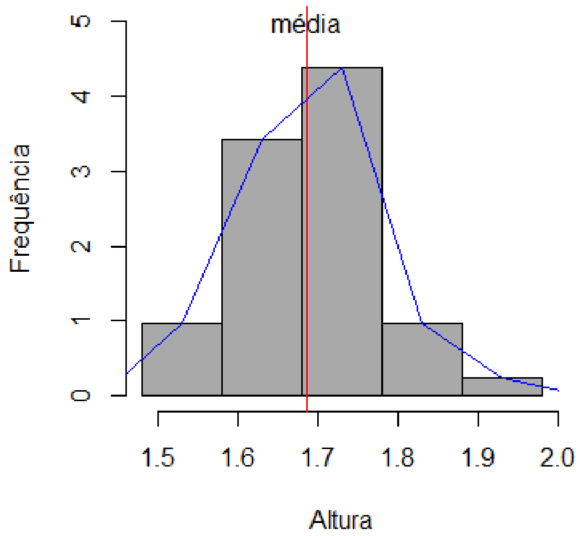


Figura 2.5: Frequência dos estudantes segundo a altura.

Média a partir da série de dados

Vamos considerar uma série de dados $(x_1, \ldots, x_n)$, ou $(x_1, \ldots, x_N)$, a partir desse conjunto podemos obter a média como:

- $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$, para uma amostra
- $\mu = \dfrac{\sum_{i=1}^{n} x_i}{N}$, para uma população inteira, sendo:
  - $\overline{x}$ a notação para média amostral,
  - μ a notação para média populacional,
  - n a quantidade de elementos na amostral e
  - N a quantidade de elementos na população.

Média a partir da tabela de frequência simples

Para exemplificar a obtenção da média a partir de uma tabela de frequência simples, consideremos a distribuição de frequência de uma amostra de 34 famílias de quatro filhos quanto ao número de filhos do sexo masculino apresentada na Tabela 2.10. Neste caso temos que o número médio de filhos homens por família é dado por:

Tabela 2.10: Frequência de famílias com quatro filhos segundo número de filhos do sexo masculino.

Número de meninos	n_j
0	2
1	6
2	10
3	12
4	4
Total	34

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i n_i}{n} = \frac{(0 \times 2) + (1 \times 6) + (2 \times 10) + (3 \times 12) + (4 \times 4)}{34} = \frac{78}{34} \cong 2,3$$

Assim a média desta amostra é de 2,3 meninos por família.

Genericamente, temos

- $\overline{x} = \dfrac{\sum_{j=1}^{k} n_j x_j}{n}$, para uma amostra,
- $\mu = \dfrac{\sum_{i=1}^{k} n_j x_j}{N}$, para uma população inteira,

  em que

  - x_j é o j-ésimo valor possível da variável (valores sem as repetições e ordenados),
  - n_j é a j-ésima frequência absoluta e
  - k é o número de classes da tabela de frequência.

Média a partir da tabela de frequência em intervalos de classe

A tabela a seguir representa a idade dos estudantes do curso de medicina veterinária da UFBA, ano/1993. A partir dessa tabela, vamos calcular a idade média desses alunos. Para isso, consideremos duas colunas extras na tabela: a coluna dos pontos médios das classes e a coluna dos pontos médios multiplicados pelas frequências absolutas ($x_j.n_j$).

Classe de Idade	n_j	Ponto médio das classes (x_j)	$x_j.n_j$
21 ⊢ 24	7	$\frac{21+24}{2}$=22,5	157,5
24 ⊢ 27	8	$\frac{24+27}{2}$ = 25,5	204
27 ⊢ 30	1	$\frac{27+30}{2}$=28,5	28,5
30 ⊢ 33	5	$\frac{30+33}{2}$=31,5	157,5
33 ⊢ 36	7	$\frac{33+36}{2}$=34,5	241,5
Total	28	142,5	789

- Neste caso, temos:

$$\overline{x} = \frac{\sum_{j=1}^{k} x_j.n_j}{n} = \frac{789}{28} \cong 28,18$$

  Logo a idade média dos alunos é de aproximadamente 28,2 anos.

Então, pra o cálculo da média, convencionamos que todos os valores incluídos em um determinado intervalo da classe coincidem com seu ponto médio, assim determinamos a média aritmética ponderada aproximada por

- $\overline{x} = \dfrac{\sum_{j=1}^{k} x_j n_j}{n}$ (para uma amostra),

- $\mu = \dfrac{\sum_{j=1}^{k} x_j n_j}{N}$ (para uma população),

  em que:

  - x_j representa o ponto médio da j-ésima classe,
  - n_j é a j-ésima frequência absoluta e
  - k é o número de classes da tabela de frequência.

Moda de uma série de dados

A moda é o valor que ocorre com maior frequência em uma série de dados.

Exemplo 2.5.1. *Vamos obter a moda nos seguintes casos:*

- *considerando a série: (7 , 8 , 9 , 10 , 10 , 10 , 11 , 12),*
  neste caso a moda é igual a 10;
- *considerando a série: (3 , 5 , 8 , 10 , 12), neste caso a série é* ***amodal;***
- *considerando a série: (2 , 3 , 4 , 4 , 4 , 5 , 6 , 7 , 7 , 7 , 8 , 9),*
  neste caso apresentam-se duas modas: 4 *e* 7*, a série é* ***bimodal.***

Em situações em que a série apresenta mais de duas modas, dizemos que a série é **multimodal**.

Moda a partir de uma tabela de frequência simples

Uma vez que os dados encontram-se agrupados em uma tabela de frequência simples, é possível obter imediatamente a moda: basta observar o valor da variável de maior frequência.

Exemplo 2.5.2. *Consideremos a distribuição relativa a 34 famílias de quatro filhos tomando para a variável o número de filhos do sexo masculino apresentada na Tabela 2.10 e repetida abaixo:*

Número de meninos	n_j
0	2
1	6
2	10
3	**12**
4	4
Total	34

Observe que a moda da variável número de meninos é $M_o = 3$*, conforme destacado acima. A classe destacada é denominada* ***classe modal.***

Moda a partir de uma tabela de frequência simples em intervalos de classe

Se os dados estão agrupados em uma tabela de frequência em intervalos de classe, devemos aproximar a moda dentro da classe modal. Aqui faremos essa aproximação de modo similar ao que foi feito no caso da média, ou seja, obter o ponto médio da classe.

Exemplo 2.5.3. *Veja a seguir os dados de idade dos alunos do curso de medicina veterinária da UFBA, ano/1993.*

Classe de Idade	n_i	x_i
$21 \vdash 24$	7	–
$24 \vdash 27$	**8**	**25,5**
$27 \vdash 30$	1	–
$30 \vdash 33$	5	–
$33 \vdash 36$	7	–

Para esse exemplo, a moda é $M_o = 25,5$, ou seja, há uma maior quantidade de alunos com idade de 25,5 anos, aproximadamente.

Mediana de uma série de dados

A média, embora seja uma medida de tendência central muito utilizada, muitas vezes não descreve de maneira adequada um conjunto de dados, pois essa é uma medida que pode ser afetada por algumas características que os dados pode conter, como por exemplo a presença de assimetria acentuada na distribuição dos dados, ou presença de pontos que destoam dos demais, seja para cima ou para baixo. Nessas situações é necessário procurar medidas que não sejam afetadas por essas características. Uma medida que pode ser empregada nessas situações é a mediana, pois esta não é afetada por assimetria ou por pontos atípicos.

A mediana de um conjunto de valores ordenados $(x_{(1)}, x_{(2)}, \cdots, x_{(n)})$, é o valor situado de tal forma no conjunto que o separa em dois subconjuntos, de mesmo número de elementos. Aqui, $x_{(1)}$ corresponde ao **valor mínimo** da série e $x_{(n)}$ corresponde ao **valor máximo** da série de dados.

A mediana é considerada uma separatriz, por dividir a distribuição ou o conjunto de dados em duas partes iguais.

Para a obtenção da mediana de uma variável X, devemos considerar:

$$Med(X) = \begin{cases} x_{(\frac{n+1}{2})}, & \text{se n ímpar;} \\ \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2}, & \text{se n par.} \end{cases}$$

Exemplo 2.5.4. *Vamos considerar os exemplos a seguir.*

Para a série $X = (5, 2, 6, 13, 9, 15, 10)$, temos a seguinte ordenação:

$$(\underbrace{2, 5, 6,}_{3\ elementos} \boxed{9,} \underbrace{10, 13, 15}_{3\ elementos}) \Rightarrow Med(X) = 9$$

Para a série $Y = (1, 3, 0, 0, 2, 4, 1, 3, 5, 6)$, temos a ordenação:

$$(\underbrace{0, 0, 1, 1,}_{4\ elementos} \boxed{2,3,} \underbrace{3, 4, 5, 6}_{4\ elementos}) \Rightarrow Med(Y) = \frac{2+3}{2} = 2,5$$

Mediana a partir da tabela de frequência simples

Para obtenção da mediana a partir de uma tabela de frequência, vamos olhar, inicialmente para a coluna das frequências relativas acumuladas, pois a mediana é um valor que contém abaixo dele 50% dos dados, com isso podemos encontrar facilmente a classe mediana diretamente da tabela.

Exemplo 2.5.5. *Consideremos a distribuição relativa a 34 famílias de quatro filhos tomando para a variável o número de filhos do sexo masculino:*

Número de meninos	n_i	f_i	f_{ac}
0	*2*	*2/34*	*2/34 ≈ 0,06*
1	*6*	*6/34*	*8/34≈0,24*
2	**10**	**10/34**	**18/34≈0,53**
3	*12*	*12/34*	*30/34≈0,88*
4	*4*	*4/34*	*34/34=1*
Total	*34*	*1*	

- *Para se obter a mediana, observe que a até a terceira classe acumulam-se mais de 50% dos dados (53% dos dados), sendo assim, esta é a classe que contém a mediana com sobra, deste modo não importa se o total de elementos na série é par ou ímpar, a mediana é o valor que está nessa classe.*

- *Assim, a mediana é dada por $Med = 2$ meninos.*

Se n é par e existe uma classe que concentra até ela exatamente 50% dos dados, devemos obter a média aritmética simples entre o valor da classe mediana e o valor imediatamente posterior.

Exemplo 2.5.6. *Como exemplo, veja a distribuição de frequência abaixo:*

x_i	n_i	f_{ac}
12	1	1/8
14	2	3/8
15	1	**4/8=0,5**
16	**2**	**6/8=0,75**
17	1	7/8
20	1	8/8
Total	8	1

Na tabela acima, tem-se $f_{ac_3} = 0,5$ com n par, assim sabemos que existem dois valores diferente ocupando a posição central dos dados ordenados, logo:

$$Med = \frac{15 + 16}{2} = 15,5$$

Mediana a partir da tabela de frequência em intervalos de classe

A obtenção da mediana de dados agrupados em uma tabela de frequência em intervalos de classes é feira, primeiramente, localizando a classe que contém a mediana, assim como no caso de uma tabela de frequência simples. No entanto, o valor da mediana não pode ser obtido exatamente, exigindo, assim, uma aproximação dentro do intervalo que contém esse valor. Essa aproximação será feita aqui de modo a levar em consideração a distribuição de frequência por meio da relação:

$$Med = L_i + \left[\frac{\left(0,5 - f_{ac(ant)}\right)}{f_i} \right] \times \Delta$$

em que,

- L_i: limite inferior da classe mediana,
- $f_{ac(ant)}$: frequência relativa acumulada da classe anterior à classe mediana,
- Δ: amplitude da classe e
- n_i é a frequência absoluta (contagem) da classe mediana.

Exemplo 2.5.7. *A tabela a seguir representa a idade dos alunos do curso de medicina veterinária da UFBA, ano/1993.*

Classe de Idade	n_i	f_{ac}
$21 \vdash 24$	7	7/28
24 ⊢ 27	***8***	***15/28 ≈ 0,54***
$27 \vdash 30$	1	16/28
$30 \vdash 33$	5	21/28
$33 \vdash 36$	7	28/28
Total	28	1

Aqui, a classe que contém a mediana é a segunda, pois mais de 50% dos valores estão acumulados até essa classe. Para este exemplo, temos:

- $L_i = 24$,
- $f_{ac(ant)} = \frac{7}{28} = 0,25$,
- $\Delta = 27 - 24 = 3$,
- $n_i = 8$ *e*
- $n = 28$

Portanto,

$$\begin{aligned} Med &= L_i + \left[\frac{\left(0,5 - f_{ac(ant)}\right)}{f_i} \right] \times \Delta \\ &= 24 + \frac{(0,5 - 0,25)}{8/28} \times 3 \\ &= 24 + \frac{21}{8} = \boldsymbol{26{,}63} \end{aligned}$$

2.5.2 Separatrizes

Separatrizes (ou **quantis**) são os valores que dividem a série de dados ordenados $(x_{(1)}, x_{(2)}, \cdots, x_{(n)})$ em partes iguais. Vimos que a mediana divide a série dados ordenados em duas partes iguais, então a mediana é uma separatriz. No entanto, essa não é a única medida relevante que divide uma série de valores. Temos outras separatrizes também importantes. Como os **quartis**, os **decis** e os **percentis**.

- Os **quartis** são os valores da variável que dividem uma série de dados ordenados em quatro partes iguais, portanto são três medidas que são denotadas por Q_1 (primeiro quartil), Q_2(segundo quartil) e Q_3(terceiro quartil).
  - o primeiro quartil, Q_1, é o valor que divide aproximadamente, a quarta parte (25%) das observações abaixo dele, e os 75% restantes, acima dele.
  - O segundo quartil é exatamente a mediana ($Q_2 = Med$).
  - O terceiro quartil, Q_3, tem aproximadamente os três quartos (75%) das observações abaixo dele e os demais 25% acima.

A estratégia para a obtenção dos quartis é semelhante aquela empregada para se obter a mediana, ou seja, primeiramente temos que encontrar a classe que contém o quartil desejado. Para isso, basta observar as frequências relativas acumuladas. Com isso, podemos usar as equações:

$$p_j = \frac{j}{4}, \text{ para } j = 1, 2, 3 \tag{2.1}$$

$$Q_j = L_i + \left[\frac{\left(p_j - f_{ac(ant)}\right)}{f_i} \right] \times \Delta \tag{2.2}$$

com,

- L_i é o limite inferior da classe definida por p_j;
- $f_{ac(ant)}$ é a frequência absoluta acumulada da classe anterior à que contém o j-ésimo quartil;
- Δ é a amplitude da classe e
- n_i é a frequência da classe definida por p_j.

Obtenção dos quartis a partir do histograma

Exemplo 2.5.8. *Observe a Figura 2.6, em que é mostrada uma série de dados agrupados em intervalos de classe. Vamos determinar os quartis geometricamente, utilizando essa figura.*

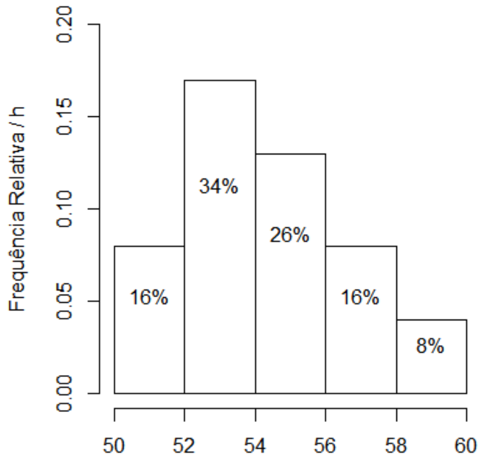


Figura 2.6: Intervalos de tempo (em minutos) de montagem de 50 equipamentos.

Em cada barra podemos observar a porcentagem de dados em cada base dos retângulos, então, é fácil concluir que o primeiro quartil (Q_1) está contido no intervalo $52 \vdash 55$, pois até esse intervalos acumulam-se 25% dos dados. A aproximação desse quartil dentro do intervalo pode, também, ser aproximado considerando que os dados estão dispostos uniformemente dentro do intervalo. Com isso, e levando em consideração que a área de cada retângulo corresponde a porcentagem de dos dados na classe, podemos aproximar facilmente o primeiro quartil da seguinte forma:

$$\begin{aligned} 25 - 16 &= 9\% \\ altura\ da\ barra &= 34/2 = 17 \\ \delta &= 9/17 \approx 0,53 \\ Q_1 &= 52 + \delta = 52 + 0,53 = 52,53 \end{aligned}$$

A obtenção dos demais quartis será deixada como exercício.

Decis e Percentis

- Se dividimos o conjunto em dez partes iguais temos os **decis**.
- Quando dividimos em cem partes, temos os **percentis**,

Para a obtenção dos decis(D_j) e percentis (P_j) podemos utilizar as seguintes expressões:

$$\begin{aligned} p_j &= \frac{j}{10}, \text{ para } j = 1,2,3,..,9 \text{ (para os decis)} \\ p_j &= \frac{j}{100}, \text{ para } j = 1,2,3,..,99 \text{ (para os percentis)} \\ D_j = P_j &= L_i + \left[\frac{\left(p_j - f_{ac(ant)}\right)}{f_i} \right] \times \Delta \end{aligned}$$

Observe que existem relações entre quartis, decis e percentis. $Q_1 = P_{25}$, $Q_2 = D_5 = P_{50}$, $Q_3 = P_{75}$, por exemplo.

2.5.3 Boxplot

O Boxplot, também conhecido como Desenho Esquemático, é um gráfico bastante útil na análise do comportamento de uma variável a partir de um conjunto de valores observados. Dentre as vantagens do boxplot, podemos destacar:

- a detecção rápida de uma possível assimetria na distribuição de frequência dos dados;
- a capacidade de fornecer uma ideia sobre a existência de possíveis pontos atípicos (muito além ou muito aquém dos demais pontos);
- a exibição dos quartis.

Para sua construção, vamos obter mais duas medidas para decidir quais são os pontos atípicos da série de dados. Vamos chamar essas medidas de **limite superior** (l_{sup}) e **limite inferior** (l_{inf}). Para obtê-los, fazemos:

$$l_{inf} = Q_1 - \frac{3}{2}(Q_3 - Q_1)$$
$$l_{sup} = Q_3 + \frac{3}{2}(Q_3 - Q_1).$$

Com essas medidas, podemos obter os valores que estão muito aquém de Q_1 ou muito além de Q_3. Tais pontos são chamados de **pontos discrepantes (ou aberrantes, ou ainda outliers)**.

Após a obtenção dos limites (l_{inf} e l_{sup}), podemos construir o boxplot da seguindo os seguintes passos:

1. No eixo cartesiano, constrói-se um retângulo na vertical de modo que:
   - A base no retângulo corresponde ao primeiro quartil (Q_1) e
   - o topo (lado superior) corresponde ao terceiro quartil (Q_3);
2. divide-se o retângulo em duas partes por meio de um segmento de reta orientado pela mediana;
3. Acima do retângulo traça-se um segmento orientado por l_{sup}
4. Abaixo do retângulo também é apresentado um traço orientado por l_{inf}
5. acima de l_{sup} e abaixo de l_{inf}, marca-se, geralmente com os pontos discrepantes.

A Figura 2.7, mostra o histograma para os dados de médias de notas da prova Brasil do ano de 2015 dos municípios do estado do Ceará. Neste gráfico podemos perceber que não existem assimetria nos dados, pois a mediana se encontra no centro do retângulo. Além disso, observamos que existem pontos acima e abaixo dos limites traçados, indicando presença de pontos atípicos, ou seja, existem municípios cujas médias estão muito acima, ou muito abaixo, das notas dos demais municípios.

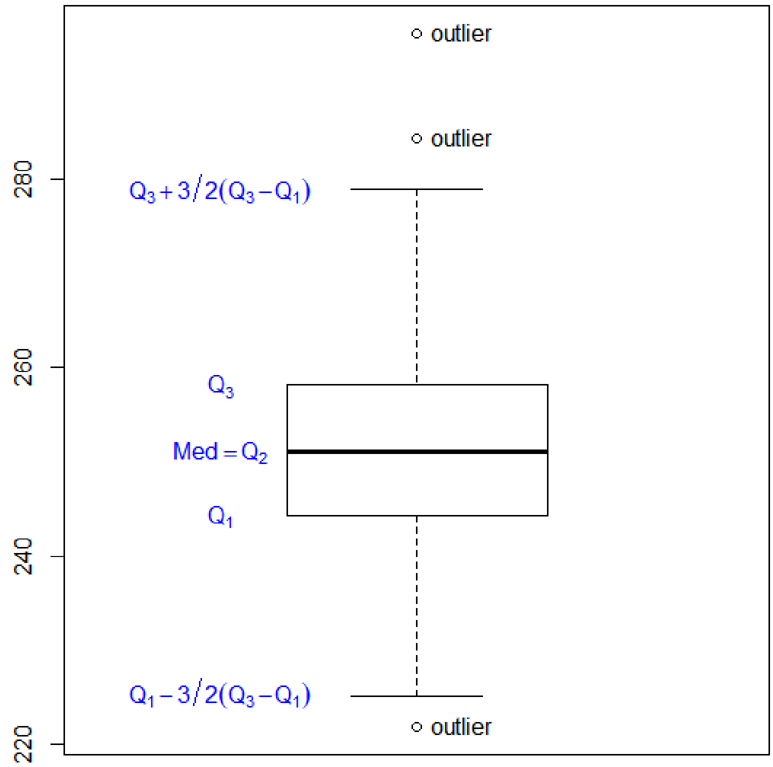


Figura 2.7: Boxplot para os dados de desempenho na prova de português dos estudantes do estado do Ceará na prova Brasil de 2015.

2.5.4 Exercícios para a Seção 2.5

Exercício 2.5.1. *Num determinado país a população feminina representa 53% da população total. Sabendo-se que a idade média (média aritmética das idades) da população feminina é de 38 anos e a da masculina é de 35 anos. Qual a idade média da população?*

Exercício 2.5.2. *Obtenha o segundo e o terceiro quartil dos dados de montagem de equipamentos representados pelo histograma apresentado na Figura 2.6.*

Exercício 2.5.3. *Os dados do quadro a seguir mostram os diâmetros abdominais em centímetros de 36 indivíduos adultos.*

59	60	60	60	62	63	63	63	63	64	66	66
66	68	69	69	69	70	71	74	75	75	77	78
81	85	86	86	87	88	88	91	95	101	107	120

Fonte: Academia Aquarius

a) Construa a tabela de distribuição de frequência em intervalos de classes, apresentando as frequências absolutas, as frequências relativas, as frequências acumuladas e o ponto médio de cada classe.

b) Obtenha a média a partir da tabela de frequência.

c) A média é uma boa medida de tendência central para este conjunto de dados? Justifique.

Exercício 2.5.4. *O gráfico apresentado na Figura 2.8 representa a variável nível de potássio no plasma em miliequivalentes (mEq), foram examinadas 20 pessoas.*

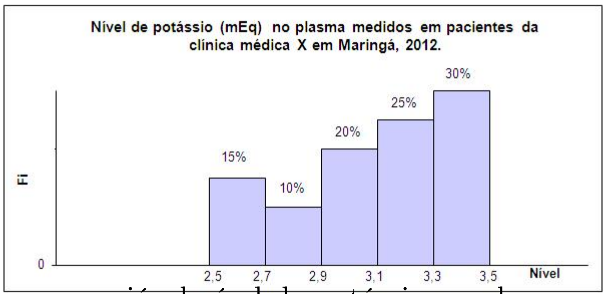


Figura 2.8: Histograma para a variável nível de potássio no plasma em miliequivalentes (mEq).

a) Calcule os quartis (Q_1, Q_2 e Q_3). Interprete-os.

b) Qual a porcentagem exata de nível de potássio acima de 3,15 mEq?

c) Faça o boxplot para esse conjunto de dados e interprete-o.

Exercício 2.5.5. *Os dados a seguir representam a fração de colesterol em miligramas por decilitro (mg/dl) fornecida pelo laboratório Vida e Saúde, de 9 indivíduos do sexo feminino.*

27,0	22,0	24,0	30,2	14,5	30,0	43,0	29,0	53,0

a) Qual o nível médio de colesterol deste grupo de indivíduos?

b) Qual o nível mediano e qual a moda?

Exercício 2.5.6. *Um hospital tem o interesse em determinar a altura média dos pacientes de uma determinada área e relacioná-la com a incidência de determinada anomalia ortopédica. Foram selecionados 80 pacientes e as alturas (em m) podem ser encontradas na tabela abaixo.*

Altura dos pacientes

1,72 1,78 1,87 1,86 1,79 1,79 1,83 1,74 1,64 1,62 1,75 1,65 1,75 1,58 1,63 1,77 1,64 1,68 1,66 1,82
1,68 1,80 1,74 1,76 1,74 1,72 1,75 1,89 1,73 1,76 1,72 1,71 1,63 1,81 1,65 1,58 1,63 1,70 1,73 1,57
1,75 1,64 1,73 1,70 1,75 1,56 1,70 1,68 1,68 1,79 1,75 1,71 1,62 1,83 1,72 1,76 1,67 1,82 1,67 1,60
1,67 1,61 1,61 1,67 1,75 1,80 1,70 1,77 1,73 1,77 1,64 1,66 1,74 1,66 1,66 1,79 1,68 1,79 1,69 1,80

Use o computador para resolver os itens abaixo.
a) (Feito na lista 1) Construa uma tabela de frequência agrupando os dados em intervalos de classe.
b) Construa o histograma.
c) Calcule o primeiro, segundo e o terceiro quartil e interprete-os.
c) Construa o boxplot.

2.6 Medidas de Dispersão

O resumo de um conjunto de dados por uma única medida de tendência central esconde toda a informação sobre a variabilidade do conjunto de observações. Por essa razão, precisamos empregar outro tipo de medida que informe sobre quão dispersos os dados estão.

Por exemplo, suponhamos que se deseja comparar a performance de dois empregados, com base na seguinte produção diária de determinada peça:

Funcionário	Variáveis	Total
A	70; 71; 69; 70; 70	350
B	60; 80; 70; 59; 81	350

Observe que $\overline{x}_A = 70$ e $\overline{x}_B = 70$, de acordo com as médias diríamos que a performance de B é igual a de A, no entanto se observarmos a variabilidade nos dados, observamos que a performance de A é bem mais uniforme.

Dependendo da medida de tendência central empregada, devemos adotar uma medida de dispersão apropriada. Aqui abordaremos as seguintes medidas de dispersão:

- variância e desvio-padrão,
- distância interquartil e
- coeficiente de variação.

2.6.1 Variância

A medida de tendência central mais comumente utilizada é a média. Uma vez que essa medida é adotada para descrever a posição da distribuição dos dados, faz-se necessário a escolha de uma medida da variabilidade em torno dessa média. Neste caso, a variância e o desvio padrão podem ser adotados.

Variância a partir de uma série de dados

A variância é a medida que fornece o grau de dispersão, ou variabilidade dos valores do conjunto de observações em torno da média. Ela é calculada tomando-se a média dos quadrados dos desvios em relação à média. Ou seja,

para uma população:

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

Em que,

- μ é a média da população.
- N é a quantidade de elementos na população;

no caso de uma amostra temos a variância amostral:

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

Em que

- n_j é a frequência da j-ésima classe.
- $\bar{x}$ é a **média da amostra** ou **média amostral**.

Variância a partir de uma tabela de frequência

Se os dados estão apresentados em uma tabela de frequência, a variância é obtida tomando-se a **média ponderada** dos quadrados dos desvios dos valores possíveis da variável (ou dos ponto médio das classes).

$$\sigma^2 = \frac{\sum_{j=1}^{K}(x_j - \mu)^2.n_j}{N} \qquad \text{ou} \qquad S^2 = \frac{\sum_{j=1}^{K}(x_j - \mu)^2.n_j}{n-1}$$

Em que

- n_j é a frequência da j-ésima classe.
- K é o número de classes na tabela.
- x_j é o j-ésimo valor possível da variável (ou ponto médio da classe).

2.6.2 Desvio-padrão

Como a variância é uma medida de dimensão igual ao quadrado da dimensão dos dados, pode-se causar problemas de interpretação. Então costuma-se usar o desvio padrão, que é definido como a raiz quadrada da variância, como segue:

$$\sigma = \sqrt{\sigma^2} \text{ (para uma população) ou } S = \sqrt{S^2} \text{ (para uma amostra).}$$

Propriedades do desvio padrão e da variância:

1. Somando (ou subtraindo) um valor constante e arbitrário, C a cada elemento de um conjunto de números, o desvio padrão desse conjunto não se altera, essa propriedade também vale para variância.

2. Multiplicando (ou dividindo) por um valor constante C, cada elemento de um conjunto de números, o desvio padrão fica multiplicado (ou dividido) pela constante C,

3. no caso da variância ela fica multiplicada pela constante elevada ao quadrado.

2.6.3 Distância interquartil

Se a mediana é usada como medida de tendencia central, a distância entre o primeiro e o terceiro quartil pode ser usada como uma medida da variabilidade dos dados em torno da mediana. Essa medida é chamada de distância "interquartil" e é dada por:

$$D = Q_3 - Q_1$$

Também é muito utilizado a "amplitude ou desvio semi-quartil", que seria o interquartil dividido por 2. Neste caso, essa é uma boa medida de dispersão, pois em um intervalo igual ao interquartil em torno da mediana estão 50% dos dados. Neste caso, o boxplot pode ser utilizado para visualizar o comportamento da variável que gerou os dos dados.

2.6.4 Amplitude total

Também podem ser usadas outras medidas para se ter uma ideia da dispersão dos dados. Um exemplo é a Amplitude Total (AT) que é a diferença entre o maior e o menor valor observado.

$$AT = x_{(\text{máx})} - x_{(\text{ mín})}$$

Dados agrupados em classes: Neste caso a AT é dada pela diferença entre o limite superior da última classe e o limite inferior da primeira classe.

$$AT = L_s - L_i$$

Obs: essa não é muito utilizada devido ser altamente afetada por pontos discrepantes, além de ser pouco informativa.

2.6.5 Coeficiente de Variação

O coeficiente de variação é uma medida de variabilidade relativa, que é definida como a razão entre o desvio padrão e a média. Assim, essa é uma medida expressa em percentual e é dada por:

$$CV\% = \frac{\sigma}{\mu} \times 100.$$

Note que o coeficiente de variação não tem unidade, pois o desvio-padrão e a média estão na mesma unidade, fazendo com que estas se cancelem.

Com o coeficiente de variação, podemos avaliar se a média é ou não uma boa medida de tendencia central.

Critérios para interpretação.

Quanto menor for o coeficiente de variação, mais representativa dos dados será a média. Coeficiente de variação acima de 50%, a média não é representativa e outra medida de tendencia central deve ser utilizada, como a mediana por exemplo.

Uma interpretação para o coeficiente de variação

- Se $0\% \leq CV\% < 30\%$, conclui-se pela baixa variabilidade dos dados e a média é uma ótima medida para representar os dados;
- Se $30\% \leq CV\% < 50\%$, conclui-se pela média variabilidade dos dados e a média é uma boa medida para representar os dados;
- Se $CV\% \geq 50\%$, conclui-se pela alta variabilidade dos dados e a média não é uma medida apropriada para representar os dados. Neste caso, deve-se pensar na mediana ou moda.

Por ser adimensional, o coeficiente de variação é uma boa opção para se comparar o grau de concentração dos dados em torno da média de séries de dados distintas, mesmo que tenham unidades diferentes.

Interpretação em Controle Estatístico de Processo

- Vários autores indicam diferentes métodos para se classificar o coeficiente de variação.
- Além disso, essa medida é intrínseca a cada processo, sendo muito utilizado na área agrícola, mais especificamente na experimentação agronômica.
- Em controle estatístico de processo, também pode-se utilizar a interpretação mostrada no quadro abaixo:

Faixa	CV %	Dispersão
menor ou igual a 15%	baixo	baixa dispersão dos dados
entre 15% e 30%	médio	média dispersão dos dados
maior que 30%	alto	alta dispersão dos dados

Exemplo 2.6.1. *Voltando ao exemplo da performance dos dois empregados, vamos calcular a variância, o desvio padrão e o coeficiente de variação dos dois conjuntos de valores de produção diaria dos empregados A e B:*

Empregado	*Variáveis*	Σ
A	*70; 71; 69; 70; 70*	*350*
B	*60; 80; 70; 59; 81*	*350*

Vimos que: $\bar{X}_A = 70$ *e* $\bar{X}_B = 70$. *Vamos agora obter medidas de dispersão para esse conjunto de valores.*

- *Variância de A.*

$$S_A^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} = \frac{(70-70)^2+(70-71)^2+(70-69)^2+(70-70)^2+(70-70)^2}{5-1} = \frac{1+1}{4} = \frac{2}{4} = 0,5$$

- *Variância de B.*

$$\begin{aligned} S_B^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n-1} &= \frac{(70-60)^2+(70-80)^2+(70-70)^2+(70-59)^2+(70-81)^2}{5-1} \\ &= \frac{100+100+121+121}{4} = \frac{442}{4} = 110,5 \end{aligned}$$

- *Desvio padrão e coeficiente de variação de A.*

$S_A = \sqrt{0,5} = 0,7$ *e* $CV_A\% = \frac{0,7}{70}.100 = 1\%$

- *Desvio padrão e coeficiente de variação de B.*

$S_B = \sqrt{110,5} = 10,51$ *e* $CV_B\% = \frac{10,51}{70}.100 = 15,01\%$

Conclusão: *as duas médias representam muito bem os dados, no entanto é fácil verificar que a dispersão dos valores de B é muito maior que a de A.*

Exemplo 2.6.2. *Considere a seguinte distribuição de frequências correspondente aos diferentes preços de um determinado produto em vinte lojas pesquisadas.*

Preços (R$)	*Nº de lojas*
50	2
51	5
52	6
53	6
54	1
Total	20

Vamos determinar a média, a variância, o desvio padrão e o coeficiente de variação dos preços. Adicionando as colunas complementares, a tabela completa fica:

Preços (R$)	*Nº de lojas*	$x_i.n_i$	$(x_i - \overline{x})$	$(x_i - \overline{x})^2$	$(x_i - \overline{x})^2.n_i$
50	2	100	-1,95	3,8025	7,605
51	5	255	-0,95	0,9025	4,5125
52	6	312	0,05	0,0025	0,015
53	6	318	1,05	1,1025	6,615
54	1	54	2,05	4,2025	4,2025
Total	20	1039			22,95

A partir dessa tabela, obtemos os valores desejados como segue:

- $\bar{x} = \dfrac{\sum_{i=1}^{n} x_i.n_i}{n} = \frac{1039}{20} = 51,95(R\$)$

- $S^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2.n_i}{n-1} = \frac{22,95}{19} = 1,21(R\$)^2$

- $S = \sqrt{1,21} = 1,1(R\$)$ *e* $CV\% = \frac{1,1}{51,95}.100 = 2,12\%$

A média, nesse caso, é uma ótima medida para representar os dados, pois existe uma baixa variabilidade em torno desse valor.

Exemplo 2.6.3. *Um comerciante atacadista vende determinado produto em sacas que deveriam conter 16,50 kg. A pesagem de 40 sacas revelou os resultados representado na tabela:*

Classes de peso	n_i
14,55 ⊢ 15,05	1
15,05 ⊢ 15,55	3
15,55 ⊢ 16,05	8
16,05 ⊢ 16,55	9
16,55 ⊢ 17,05	10
17,05 ⊢ 17,55	6
17,55 ⊢ 18,05	3
Total	40

Determinar a média, a variância, o desvio padrão e o coeficiente de variação dos pesos.

Tabela completada.

Classe de peso	n_i	x_i	$x_i n_i$	$(x_i - \overline{X})$	$(x_i - \overline{X})^2$	$(x_i - \overline{X})^2.n_i$
14,55 ⊢ 15,05	1	14,8	14,8	-1,68	2,8224	2,8224
15,05 ⊢ 15,55	3	15,3	45,9	-1,18	1,3924	4,1772
15,55 ⊢ 16,05	8	15,8	126,4	-0,68	0,4624	3,6992
16,05 ⊢ 16,55	9	16,3	146,7	-0,18	0,0324	0,2916
16,55 ⊢ 17,05	10	16,8	168	0,32	0,1024	1,024
17,05 ⊢ 17,55	6	17,3	103,8	0,82	0,6724	4,0344
17,55 ⊢ 18,05	3	17,8	53,4	1,32	1,7424	5,2272
Total	40		659			21,276

A partir da última tabela, obtemos os valores desejados como segue:

- $\bar{x} = \frac{\sum_{i=1}^{n} x_i.n_i}{n} = \frac{659}{40} = 16,475kg$

- $S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2.n_i}{n-1} = \frac{21,276}{39} = 0,55kg^2$

- $S = \sqrt{0,55} = 0,74Kg$ *e* $CV\% = \frac{0,74}{16,475}.100 = 4,48\%$

O coeficiente de variação mostra a baixa variabilidade dos dados em torno da média, o que faz com que essa medida seja uma ótima medida para representar os dados.

2.6.6 Exercícios para a Seção 2.6

Exercício 2.6.1. *Uma distribuição de frequências para os níveis séricos (em microgramas por decilitro) de zinco de 462 homens entre as idades de 15 a 17 anos é exibida na Tabela 2.11.*

Tabela 2.11: Nível sérico de zinco (mg/dl).

Nível	Número de homens
$50 \vdash 60$	6
$60 \vdash 70$	35
$70 \vdash 80$	110
$80 \vdash 90$	116
$90 \vdash 100$	91
$100 \vdash 110$	63
$110 \vdash 120$	30
$120 \vdash 130$	5
$130 \vdash 140$	2
$140 \vdash 150$	2
$150 \vdash 160$	2
Total	462

1. *Obtenha a média, a moda e a mediana dos níveis de zinco.*

2. *Obtenha o desvio padrão e o coeficiente de variação. O que você pode concluir?*

3. *Apresente os quartins.*

4. *Exiba as separatrizes em um diagrama de caixa (boxplot).*

Exercício 2.6.2. *A seguir temos a distribuição de frequência dos pesos de uma amostra de 45 alunos:*

Pesos(Kg)	$40 \vdash 45$	$45 \vdash 50$	$50 \vdash 55$	$55 \vdash 60$	$60 \vdash 65$	$65 \vdash 70$
Número de alunos	04	10	15	08	05	03

a) Determinar a média, a moda e a mediana.
R: 53,5
b) Determinar o desvio padrão e o coeficiente de variação. O que você pode concluir? R: var= 45; d.p.= 6,708; C.V.=0,125

Exercício 2.6.3. *Entre os formados de 2012 de uma certa universidade no Paraná, 382 formados em ciências humanas receberam ofertas de emprego com salários anuais médios de 24 373 reais, 450 formados em ciências sociais receberam ofertas de emprego com salários anuais médios de 22 684 reais e 113 formados em ciência da computação receberam ofertas de emprego com salários anuais médios de de 31 329 reais. Qual foi a média de salário anual oferecida a esses 945 formados?*

Exercício 2.6.4. *Os dados a seguir referem-se aos rendimentos médios, em kg/ha, de 32 híbridos de milho recomendados para a Região Oeste catarinense*

Tabela 2.12: Rendimento médios, em kg/ha, de 32 híbridos de milho, Região Oeste, 1987/88.

3.973	4.550	4.770	4.980	5.117	5.403	6.166
4.500	4.680	4.778	4.993	5.166	5.513	6.388
4.550	4.685	4.849	5.056	5.172	5.823	
4.552	4.760	4.960	5.063	5.202	5.889	
4.614	4.769	4.975	5.110	5.230	6.047	

Fonte: Andrade e Ogliari

Use o computador para resolver os itens abaixo.
a) Construa uma tabela de frequência agrupando os dados em intervalos de classe.
b) Construa o histograma.
c) Calcule o primeiro, segundo e o terceiro quartil e interprete-os.
c) Construa o boxplot.

Capítulo 3

Análise Combinatória

A aplicação de ferramentas estatísticas muitas vezes faz uso das teorias de probabilidade. Isso se deve ao fato de não termos sempre acesso a toda população de interesse. Quando isso ocorre, devemos aplicar técnicas da **Inferência Estaística**, em que o comportamento da variável de interesse é modelada por meio de modelos probabilisticos. Por essa razão, para o bom entendimento as teorias estatísticas, devemos estudar um pouco das teorias de probabilidades. A definição mais intuitiva de probabilidade (chamda probabilidade clássica, que será abordada na seção posterior) envolve a contagem de possibilidades de fenômenos (eventos), com isso, muitas vezes, faz-se necessário a aplicação de outra ferramenta da matemática, a teoria matemática da contagem que é formalmente conhecida como **Análise Combinatória**. Para entender um pouco a relação entre contagem e probabilidade, vejamos um exemplo.

Exemplo 3.0.1. *Um sistema de comunicação formado por n antenas aparentemente idênticas devem ser alinhadas em sequência. O sistema resultante será capaz de receber qualquer sinal e será chamado de funcional desde que duas antenas consecutivas não apresentem defeito.*

> *Se exatamente m das n antenas apresentarem defeito, de quantas maneiras é possível um sistema ser funcional?*

Vamos denotar por 1 se a antena funciona e 0 se ela está com defeito, então para o caso especial em que $n = 4$ e $m = 2$, há seis configurações possíveis para o sistema. Ou seja,

1^a	2^a	3^a	4^a	5^a	6^a
1	0	0	1	0	1
1	1	0	0	1	0
0	1	1	1	0	0
0	0	1	0	1	1

Neste caso, o sistema funciona para três sequências. Quais são elas?

> *Qual será a probabilidade de que o sistema resultante seja funcional?*

Como o sistema funciona para as três primeiras sequências e não funciona nas outras três, parece razoável tomar $3/6 = 1/2$ como a probabilidade desejada. No caso de n e m quaisquer, a probabilidade poderia ser calculada de forma similar. Isto é, poderíamos contar o número de configurações que resultam no funcionamento do sistema e dividir esse número pelo número total de configurações possíveis.

> Existe um método que seja eficaz para contar o número de maneiras pelas quais as coisas podem ocorrer?

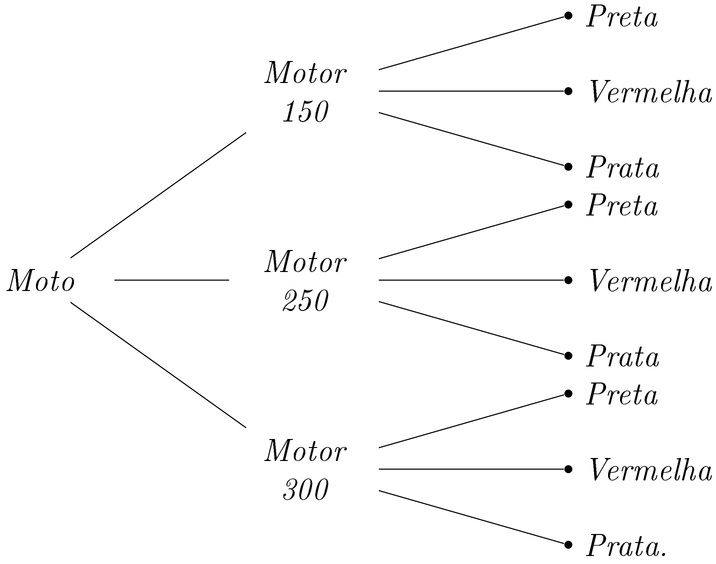
Assim, muitos problemas na teoria das probabilidades podem ser resolvidos simplesmente contando-se o número de diferentes maneiras pelas quais certo evento pode ocorrer. Para facilitar essa contagem, podemos utilizar a anlise combinatória.

3.1 O princípio básico da contagem

O princípio básico da contagem diz que se um experimento E_1 pode levar a qualquer um de m possíveis resultados e se outro experimento E_2 pode resultar em qualquer um de n possíveis resultados e E_1 for independente de E_2, então os dois experimentos em conjunto possuem $m \times n$ resultados possíveis.

Exemplo 3.1.1. *Vamos supor que uma fábrica produza motos com motores de 125, 250 ou 300 cilindradas de potência. O cliente ainda pode escolher as seguintes cores: preta, vermelha e prata. Quais são as possibilidades de venda que a empresa pode oferecer?*

Supondo a escolha do motor como o resultado do primeiro experimento, e a subsequente escolha de uma cor dentre as disponíveis como o resultado do segundo experimento, vemos a partir do princípio básico que há $3 \times 3 = 9$ *escolhas possíveis, como mostra o diagrama a seguir:*



> Quando há mais que dois experimentos a serem realizados, o princípio básico pode ser generalizado?

O principio fundamental da contagem pode ser generalizado para r experimentos, em que o primeiro pode levar a qualquer um de n_1 resultados possíveis, o segundo pode levar para cada um desses n_1 resultados a n_2 resultados possíveis, e assim por diante até o r-ésimo experimento, em que para cada um dos resultados do experimento anterior ocorre um dos n_r resultados possíveis. Ous seja, nesse caso, haverá um total de $n_1 \times n_2 \times ... \times n_r$ resultados possíveis para os r experimentos.

Exemplo 3.1.2. *O grêmio de uma faculdade é formado por 3 calouros, 4 estudantes do segundo ano, 5 estudantes do terceiro ano e 2 formandos. Um subcomitê de 4 pessoas, formado por uma pessoa de cada ano, deve ser escolhido. Quantos subcomitês diferentes são possíveis?*

Podemos fazer a escolha de um subcomitê como o resultado combinado dos quatro experimentos separados de escolha de um único representante de cada uma das classes. Assim, a partir da versão generalizada do princípio básico, temos que há $3 \times 4 \times 5 \times 2 = 120$ *subcomitês possíveis.*

3.2 Permutações

Temos o conjunto $A = \{3, 4, 6\}$, usando cada elemento de A apenas uma vez, quantos números com 3 algarismos podemos formar?

Temos que usar todos os elementos de A e formar agrupamentos que serão distinguidos apenas pela ordem em que aparecem. Estes agrupamentos são chamados permutações dos elementos de A. Assim, é possível perceber que um conjunto de 3 objetos permite 6 permutações possíveis. Esse resultado poderia ser obtido a partir do princípio básico, já que o primeiro objeto da permutação pode ser qualquer um dos três, o segundo objeto da permutação pode então ser escolhido a partir dos dois restantes e o terceiro objeto da permutação é o objeto restante. Assim, há $3 \times 2 \times 1 = 6$ permutações possíveis.

Considere um conjunto A com n elementos. Usando um raciocínio similar àquele que acabamos de utilizar leva a conclusão de que há

$$P(n) = n \times (n-1) \times (n-1) \times ... \times 3 \times 2 \times 1 = n!$$

permutação diferentes dos n elementos de A.

Temos o conjunto $B = \{2, 9, 6, 2, 2, 4, 6, 4\}$, usando cada elemento de A apenas uma vez, quantos números com 8 algarismos podemos formar?

Note que, se considerarmos os elementos 2 de B como distintos ($2_1, 2_2$ e 2_3), então permutando os 2's entre si teríamos 3! permutações diferentes, da mesma forma, para os números 4, 6 e 9 temos 2!, 2! e 1! permutações diferentes, respectivamente. Além disso, os números 2, 9, 6, 2, 2, 4, 6 e 4 permitem 8! permutações. Como permutando os números 2, 4 e 6 entre si os arranjos resultantes são os mesmos, então existem $\frac{8!}{3!2!2!1!}$ números formados pelos elementos de B.

Em geral, esse mesmo raciocínio pode ser usado para mostrar que há

$$P_n = \frac{n!}{n_1!n_2!...n_r!}$$

permutações diferentes de n objetos dos quais n_1 são da mesma espécie, n_2 são da mesma espécie, ...n_r são da mesma espécie.

Observação: Em permutações, a ordem dos elementos é importante e o número de posições é igual ao número de elementos que se deseja permutar.

3.3 Arranjo

Dado o conjunto A={2, 3, 4, 5, 6}, quantos números com algarismos diferentes se podem formar sabendo que tem 4 algarismos?

Para obter a quantidade de números, podemos usar o princípio básico da contagem. Para isso, raciocinamos da seguinte maneira:

- com os 5 elementos do conjunto A, queremos preencher 4 lugares

Lugar 1   Lugar 2   Lugar 3   Lugar 4

- o primeiro lugar pode ser preenchido de 5 maneiras distintas, o segundo de $5-1 = 4$ maneiras, o terceiro de $5-2 = 3$ maneiras e o quarto de $5-3 = 2$ maneiras, que pelo principio fundamental da contagem, resulta

$$\underline{\quad 5 \quad} \times \underline{\quad 4 \quad} \times \underline{\quad 3 \quad} \times \underline{\quad 2 \quad} = 120$$

Problemas em que se tem que escolher p entre n objetos, com $p < n$, e a ordem em que fazemos a escolha determina sequências diferentes, recebem o nome de arranjo simples de p elementos em n. Neste caso geral, temos que o primeiro lugar pode ser preenchido de n maneiras distintas, o segundo de $n-1$ maneiras, o terceiro de $n-2$ maneiras e assim sucessivamente até o p-ésimo lugar, que pode ser preenchido de $n-(p+1)$ modos diferentes, que pelo principio fundamental da contagem fornece

$$A(n,p) = n(n-1)(n-2)(n-3)\cdots(n-(p-1)) = \frac{n!}{(n-p)!}$$

Observações:

1. toda permutação é um arranjo (caso em que p = n);
2. para que a fórmula acima faça sentido, definimos $0! = 1$.

3.4 Combinação

Um comitê de três pessoas deve ser formado a partir de um grupo de 20 pessoas. Quantos comitês diferentes são possíveis?

Para resolver o problema acima, pense da seguinte forma: como há 20 maneiras diferentes de selecionar a primeira pessoa, 19 maneiras de selecionar a pessoa seguinte e 18 maneiras de selecionar a última pessoa, há portanto $20 \times 19 \times 18$ maneiras de selecionar o grupo de 3 se considerarmos a ordem de seleção das pessoas como relevante. Entretanto, como cada grupo de 3 é contado considerando a ordem como relevante, por exemplo, vamos denotar por P_1, P_2 e P_3 três pessoas quaisquer do grupo, assim o grupo formado pelas pessoas P_1, P_2 e P_3 será contado 6 vezes (isto é, todas as permutações $P_1P_2P_3$, $P_3P_2P_1$, $P_1P_3P_2$, $P_2P_1P_3$, $P_3P_1P_2$ e $P_2P_3P_1$ serão contadas quando a ordem da seleção for relevante). Então, como a ordem das pessoas para a formação dos comitês não é importante, o número total de grupos que podem ser formados é igual a

$$\frac{20 \times 19 \times 18}{3 \times 2 \times 1} = \frac{20 \times 19 \times 18}{6} = 1140$$

que corresponde a combinação de 20 pessoas tomadas 3 a 3.

Em geral, como $n(n-1)\cdots(n-r+1)$ representa o número de diferentes maneiras pelas quais um grupo de r itens pode ser selecionado a partir de n itens quando a ordem da seleção é relevante, e como cada grupo de r itens será contado $r!$ vezes, tem-se que o número de grupos diferentes de r itens que podem ser formados a partir de um conjunto de n itens é

$$C(n,r) = \frac{n(n-1)\cdots(n-(r-1))}{r!} = \frac{n!}{r!(n-r)!},$$

para $r \leq n$.

A combinação de n elementos tomados r a r é denotado por

$$\binom{n}{r} = C(n,r) = \frac{n!}{r!(n-r)!}.$$

Exemplo 3.4.1. *Considere um conjunto de n antenas das quais m apresentam defeito e $n-m$ funcionam, com $m \leq n-m+1$, e suponha que não seja possível distinguir as antenas defeituosas* *daquelas que funcionam. Quantos alinhamentos podem ser feitos sem que duas antenas com defeito sejam colocadas lado a lado?*

Se não for permitido que duas antenas com defeito sejam colocadas lado a lado, então os espaços entre as antenas que funcionam devem conter no máximo uma antena defeituosa. Ou seja, nas $n - m + 1$ *posições possíveis entre as* $n - m$ *antenas que funcionam, devemos selecionar* m *espaços onde colocar as antenas defeituosas. Portanto, há* $\binom{n-m+1}{m}$ *ordenações possíveis nas quais há pelo menos uma antena que funciona entre duas defeituosas.*

Como é definido $0! = 1$, temos que $\binom{n}{0} = \binom{n}{n} = 1$. Além disso, $\binom{n}{i} = 0$ para $i < 0$ ou $i > n$. Para $0 \leq i \leq n$, essa grandeza representa o número de diferentes subgrupos de tamanho i que podem ser formados em um conjunto de tamanho n. Ela é frequentemente chamada de coeficiente binomial por causa de seu destaque num teorema chamado teorema binomial o qual é enunciado mais adiante.

Para inteiros não negativos $n_1, ..., n_r$, cuja soma é n,

$$\binom{n}{n_1, n_2, ..., n_r} = \frac{n!}{n_1!n_2!...n_r!}$$

corresponde ao número de divisões de n itens em r subgrupos não superpostos de tamanhos $n_1, n_2, ..., n_r$ conhecidos como coeficientes multinomiais. Esse coeficiente aparece na generalização do teorema binomial, denominado teorema multinomial, que é enunciado a seguir.

Resultados

1. Teorema Binomial

$$(x + y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i}.$$

2. Teorema Multinomial

$$(x_1 + \cdots + x_r)^n = \sum_{n_1,...,n_r : x_1+\cdots+x_r=n} \binom{n}{n_1, n_2, ..., n_r} x_1^{n_1} \cdots x_r^{n_r}.$$

3. **Número de soluções inteiras positivas:** existem $\binom{n-1}{r-1}$ vetores distintos com valores inteiros positivos satisfazendo a equação

$$x_1 + x_2 + \cdots + x_r = n, x_i > 0 \text{ para todo} i = 1, \cdots, r.$$

4. **Número de soluções inteiras não negativas:** existem $\binom{n+r-1}{r-1}$ vetores distintos com valores inteiros não negativos satisfazendo a equação

$$x_1 + x_2 + \cdots + x_r = n, x_i \geq 0 \text{ para todo} i = 1, \cdots, r.$$

5. Identidade

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}, 1 \leq r \leq n.$$

A demonstração dos resultados apresentados nessa seção é omitida aqui, mas podem ser vistas, por exemplo, em Ross (2009).

Exemplo 3.4.2. *Quantas soluções com valores inteiros não negativos de* $x_1 + x_2 = 3$ *são possíveis?*

Existem $\binom{3+2-1}{2-1} = 4$ *soluções:* $(0, 3), (1, 2), (2, 1), (3, 0)$.

3.5 Exercícios

1. (Hazzan, 2013, pag. 10 ex. 5) Um edifício tem 8 portas. De quantas formas uma pessoa poderá entrar nesse edifício e sair por uma porta diferente da que usou para entrar?

2. (Ross, 2009) Vinte trabalhadores serão alocados em vinte tarefas diferentes, um em cada tarefa. Quantas alocações diferentes são possíveis?

3. (Ross, 2009) João, Júlio, Jonas e Jacques formaram uma banda com quatro instrumentos. Se cada um dos garotos é capaz de tocar todos os instrumentos, quantas diferentes combinações são possíveis? E se João e Júlio souberem tocar todos os quatro instrumentos, mas Jonas e Jacques souberem tocar cada um deles apenas o piano e a bateria?

4. **(a)** De quantas maneiras diferentes 3 garotos e 3 garotas podem sentar-se em fila?

   (b) De quantas maneiras diferentes 3 garotos e 3 garotas podem sentar-se em fila se os garotos e as garotas sentarem- se juntos?

   (c) E se apenas os garotos sentarem-se juntos?

   (d) E se duas pessoas do mesmo sexo não puderem se sentar juntas?

Capítulo 4

Introdução as Teorias de Probabilidade

Introdução

Os conceitos de chance e incerteza são tão antigos quanto a civilização. As pessoas sempre tiveram que lidar com a incerteza, seja sobre o tempo, seja sobre suplementos alimentícios necessários para sobrevivência, entre outras coisas. Assim, surgiu a necessidade de reduzir tais incertezas e seus efeitos. Uma grande motivação para a intensificação do estudo da incerteza foram os jogos de azar, que já existiam a muitos anos A.C. (3500 – 2000 A.C.). Acredita-se que os jogos de azar levaram a estudos mais aprofundados das medidas de chance, chegando ao inicio do desenvolvimento da teoria matemática de probabilidade na França pelos matemáticos Blaid Pascal (1623 - 1662) e Pierre de Fermat (1601 – 1665) .

No nosso cotidiano estamos sempre nos referindo a acontecimentos que sugerem essas teorias. Frequentemente ouvimos ou usamos expressões como: "é provável que vai chover hoje"; "exitem grandes chances do voo atrasar"; "é muito provável que será possível comparecer ao jantar de confraternização ". Cada uma dessas expressões são baseadas no conceito de probabilidade ou na possibilidade de que um evento específico irá ocorrer. Ou seja, a intenção é tentar quantificar a chance de um evento ocorrer.

As teoria de probabilidade tem sido desenvolvida a muito tempo e, também, amplamente aplicada em diversos campos de estudos. Na atualidade, a teoria de probabilidade é uma importante ferramenta na maioria das áreas das engenharias, ciências naturais e em gestão de modo geral. Muitos pesquisadores trabalham para estabelecer novas aplicações das teorias de probabilidade em campos como medicina, meteorologia, marketing, comportamento humano, predição de terremotos, projeto de sistemas para computadores, entre outros.

O conceito formal de probabilidade está associado aos experimentos, e pode ser apresentado de três formas que vamos denominar de clássica, frequentista e axiomática. Neste capítulo, vamos discutir esses conceitos e outros associados a essas definições, além de apresentar algumas propriedades importantes da probabilidade.

4.1 Experimento Aleatório, Espaço Amostral e Eventos

Exemplo 4.1.1. *Suponha que uma batelada (carga recarregável) contenha seis itens ($\{a, b, c, d, e, f\}$) e que dois itens sejam selecionados ao acaso, sem reposição. Suponha que o item f seja defeituoso, porém que os outros sejam bons. Qual é a probabilidade de que o item f apareça na amostra?*

- *Para responder essa questão, deve-se antes entender o experimento e responder outras questões.*

    1. *Qual é o espaço amostral do experimento?*
    2. *Qual é o evento de interesse?*
    3. *Qual a quantidade de eventos elementares no espaço amostral (casos possíveis)?*
    4. *Qual a quantidade de eventos elementares no evento de interesse (casos favoráveis)?*

4.1.1 Experimentos

O cálculo das probabilidades está associado aos experimentos, os quais podem ser classificados em dois tipos:

- **Experimento determinístico**: é aquele que repetido sob condições quase idênticas conduzem a um mesmo resultado.
- **Experimento aleatório**: é aquele que repetido sob condições quase idênticas produzem resultados em geral diferentes.

Aqui, o interesse recai sobre os experimentos aleatórios, como por exemplo:

a) ε_1 : Lançamento de um dado.

b) ε_2 : Selecionar um morador da cidade de Russas e verificar se ele já teve uma determinada doença.

c) ε_3 : Observar o tempo de vida de um equipamento eletrônico.

d) ε_4 : Observar a produção de um produto e contar quantos saem com defeito.

Nos experimentos aleatórios, não sabemos ao certo qual será o resultado do experimento antes de ele ser realizado, no entanto podemos conhecer quais resultados são possíveis de acontecer. Ao conjunto de todos os resultados possíveis do experimento, damos o nome de **Espaço Amostral**.

> O espaço amostral de um experimento aleatório é formado pelo conjunto de todos os resultados possíveis do experimento, e é representado aqui pela letra grega ômega, Ω.

Exemplo 4.1.2. *Exemplos de espaços amostrais.*

1. *Conta-se o número de veículos que passam por um posto de pedágio das 24 as 8 horas, então o espaço amostral é dado por $\Omega = \mathbb{N}$.*
2. *Se o experimento consiste em jogar duas moedas, então o espaço amostral é formado pelos quatro pontos a seguir:*
$$\Omega = \{CK, KC, CC, KK\},$$
em que C denota a face cara e K denota a face coroa das moedas.

Qualquer subconjunto A do espaço amostral é conhecido como **evento**, ou seja, se A está contido em Ω, então A é um evento. Em outras palavras, um evento é um conjunto formado por um ou mais de um possível resultado do experimento.

Exemplo 4.1.3. *Considerando no Exemplo 4.1.2 o evento A= { (K,K) ,(K,C)}, então A é o evento em que a primeira moeda lançada dá coroa.*

Observações

1. Os eventos são, em geral, denotados por letras maiúsculas (A, B , C etc).
2. Um único ponto do espaço amostral, é dito ser um **evento elementar** e é denotado por ω.
3. **Evento Elementar** constitui um único resultado possível do experimento que não pode ser decomposto em eventos "menores".
4. Um ponto amostral ω que está "dentro" de um evento A é dito **pertencer** à A e denotamos $\omega \in A$.
5. Se o evento unitário ω **não pertence** ao evento A, denotamos $\omega \notin A$
6. Se um evento A contém outro evento B, denotamos $B \subset A$.
7. Ω é o evento certo do experimento e contém todos os outros eventos.
8. O **evento impossível** do experimento é o conjunto vazio, e é denotado por $\emptyset$.
9. O conjunto vazio pertence a todos os eventos, ou seja, $\emptyset \in A$ para todo $A \subset \Omega$.
10. Um evento de um experimento ocorre, se ocorre qualquer um de seus elementos.

Considere o experimento: lançar duas moedas e verificar a face superior. Quais são os eventos elementares desse experimento?

4.1.2 Tipos de Espaço Amostral

Se Ω for um conjunto finito ou infinito enumerável de todos os resultados possíveis de um experimento, então este espaço amostral é dito ser discreto. Mas se Ω é não enumerável, então dizemos que este espaço amostral é contínuo. Ou seja,

- espaço amostral discreto: finito ou infinito enumerável;

  exemplo: considerando o experimento "contar o número de peças com defeito num lote", $\Omega = \mathbb{N}$;

- espaço amostral contínuo: infinito não enumerável (um intervalo da reta);

  exemplo: considerando o experimento "medir o tempo de execução de um algoritmo", $\Omega = \mathbb{R}^+$.

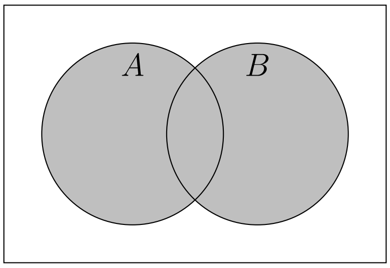
4.1.3 Operações com Eventos

As operações com eventos são realizadas seguindo as regras da teoria de conjuntos, ramo da matemática. Então, aqui, faremos uma revisão dessa teoria com foco em resultados de experimentos.

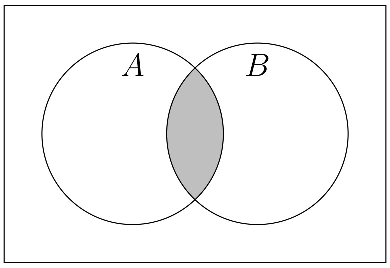
Definições básicas

Considere um espaço amostral Ω associado a um experimento e $\mathcal{P}(\Omega)$ denotando o conjunto das partes de Ω, ou seja $\mathcal{P}(\Omega)$ denota o conjunto que contém todos os subconjuntos de Ω. Agora, sejam A e B dois eventos, tais que $A \in \mathcal{P}(\Omega)$ e $B \in \mathcal{P}(\Omega)$, isto é, $A \subset \Omega$ e $B \subset \Omega$. Com isso, definimos:

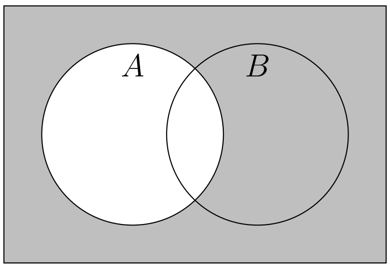
i) A **união** dos eventos A e B é dada por $\mathrm{A} \cup \mathrm{B} = \{\omega \in \Omega : \omega \in \mathrm{A} \text{ ou } \omega \in \mathrm{B}\}$, como mostra o diagrama de Venn:



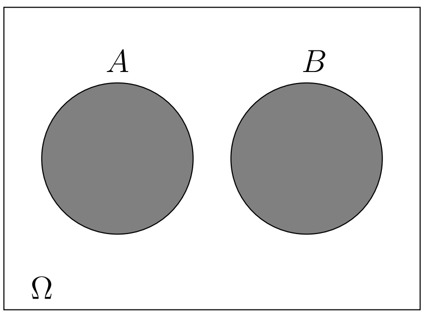
ii) A **interseção** dos eventos A e B é dada por: $\mathrm{A} \cap \mathrm{B} = \{\omega \in \Omega : \omega \in \mathrm{A} \text{ e } \omega \in \mathrm{B}\}$, como mostra o diagrama:



iii) O **complementar** de A é dado por: $\mathrm{A}^{\mathrm{C}} = \Omega - \mathrm{A} = \{\omega \in \Omega; \omega \notin \mathrm{A}\}$. Para essa situação temos o seguinte diagrama de Venn:



vi) Se $\mathrm{A} \cap \mathrm{B} = \emptyset$, então dizemos que A e B são eventos **disjuntos ou mutuamente excludente**.



Generalização das definições

As definições e propriedades apresentadas anteriormente podem ser generalizadas para qualquer quantidade de eventos $A_1, A_2, A_3, ..., A_n, n \geq 2$. Para união e interseção, temos:

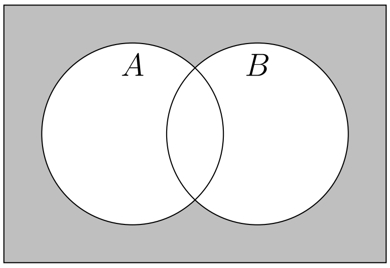
i) União: $\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup A_3 \cup ... \cup A_n = \{\omega \in A_i \text{ para algum } i\}$.

ii) Interseção: $\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap A_3 \cap ... \cap A_n = \{\omega \in A_i \text{ para todo } i\}$.

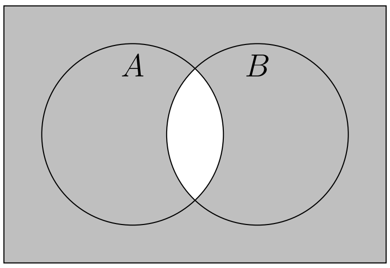
Propriedades

Leis de De Morgan.

i) $(A \cup B)^c = A^c \cap B^c = \Omega - A \cup B$



ii) $(A \cap B)^c = A^c \cup B^c = \Omega - A \cap B$



4.1.4 Partição do espaço amostral

Diremos que os eventos $A_1, A_2, \ldots, A_n$ formam uma partição de Ω, se são disjuntos dois a dois e sua união é Ω, ou seja, $\bigcup_{i=1}^{n} A_i = \Omega$ com $A_i \cap A_j = \emptyset$, para todo $i \neq j$, como ilustra a Figura 4.2, considerando $n = 6$.

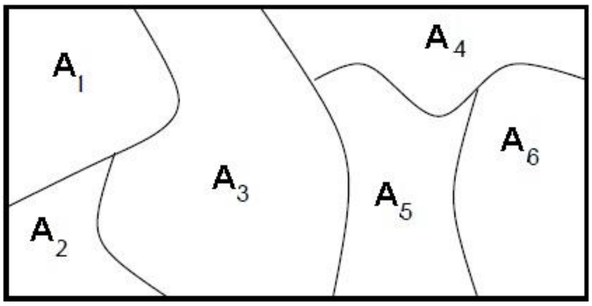


Figura 4.1: Partição do espaço amostral (n=6).

4.1.5 Exercícios para Seção 4.1

Exercício 4.1.1. *Mostre que:*

(i) $(A \cap B)^c = A^c \cup B^c$

(ii) $(A \cup B)^c = A^c \cap B^c$

Exercício 4.1.2. *Lançam-se três moedas. Enumerar o espaço amostral e os eventos:*

(a) *faces iguais;*

(b) *cara na 1ª moeda;*

(c) *coroa na 2ª e 3ª moedas.*

Exercício 4.1.3. *Considere a experiência que consiste em pesquisar frunílias com três crianças, em relação ao sexo delas, segundo a ordem do nascimento. Enumerar os eventos:*

(a) *ocorrência de dois filhos do sexo masculino;*

(b) *ocorrência de pelo menos um filho do sexo tnasculino;*

(c) *ocorrência de nomáximo duas crianças do sexo feminino.*

Exercício 4.1.4. *Sejam A, B e C três eventos de um espaço amostral. Exprimir os eventos abaixo usando as operações reunião, intersecção e complementação:*

(a) *somente A ocorre;*

(b) *A e C ocorrem, mas B não;*

(c) *A, B e C ocorrem;*

(d) *pelo menos um ocorre;*

(e) *exatamente um ocorre;*

(f) *nenhum ocorre;*

(g) *no máximo dois ocorrem;*

(h) *pelo menos dois ocorrem.*

Exercício 4.1.5. *Considere o conjunto universo (espaço amostral) dos inteiros de 1 a 10 ou* $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. *Sejam* $A = \{2, 3, 4\}$, $B = \{3, 4, 5\}$ *e* $C = \{5, 6, 7\}$. *Liste os elementos dos seguintes conjuntos:*
a) $A^C \cap B$;
b) $A^C \cup B$;
c) $(A^C \cap B^C)^C$;
d) $[A \cap (B \cap C)]^c$;
e) $[A \cap (B \cup C)]^c$.
Sugestão: *Use o diagrama de Venn para visualizar o evento de interesse.*

4.2 Definições de probabilidade

Quando falamos de Probabilidades existem algumas definições que é necessário se ter em mente. Sobre essas definições desenvolve-se toda a teoria das Probabilidades. Vamos iniciar os estudos da probabilidade dando uma definição em termos da frequência relativa dos eventos associados ao experimento (acontecimento) aleatório, em seguida vejamos a definição clássica de probabilidade (Fermat e Pascal, metade do século XVII) e finalmente a mais geral, a definição axiomática (Kolmogorov).

Definição 4.2.1. (Definição Frequentista de Probabilidade) *Considere que um experimento aleatório seja realizado n vezes e seja n_A o número de vezes que o evento A acorre. Definimos a frequência relativa de A por*

$$f_n = \frac{n_A}{n} = \frac{\textit{frequência do evento } A}{\textit{frequência total}}, \quad 0 \leq f_n \leq 1.$$

Dessa forma, pode ser mostrado que a probabilidade do evento A ocorrer é dada por:

$$P(A) = \lim_{n\to\infty} f_n.$$

Ou seja, se n for grande, f_n se aproxima da probabilidade de A ocorrer.

Exemplo 4.2.1. *Considere o problema em decidir se uma moeda é honesta. Para resolver esse problema, considere que a moeda seja lançada 100 vezes, caso a moeda seja honesta, qual o número aproximado de caras que esperamos obter?*

Observe que o exemplo anterior trata de um espaço amostral discreto e infinito, assim o que temos é uma aproximação para a probabilidade desejada. Agora vejamos uma definição de probabilidade um pouco mais restrita, em que consideramos espaços amostrais finitos, mas que fornece a probabilidade precisa de um certo evento acontecer.

Definição 4.2.2. (Definição Clássica de Probabilidade) *Vamos considerar um espaço amostral Ω finito em que todos os seus eventos elementares são igualmente prováveis. Assim, a probabilidade de um evento $A \subset \Omega$ é calculada como a razão entre o número de casos favoráveis ao evento A (eventos elementares de A) e o número de casos possíveis (número de eventos elementares de Ω). Ou seja:*

$$P(A) = \frac{n^o \textit{ casos favoráveis a } A}{n^o \textit{ de casos possíveis}} = \frac{\sharp A}{\sharp \Omega}$$

Exemplo 4.2.2. *Lança-se um dado honesto, qual a probabilidade de ocorrer a face 3?*
Sendo: $A=$ ocorrer a face 3, então $P(A) = \frac{1}{6}$

A definição clássica de probabilidade é bastante intuitiva e resolve muitos problemas práticos, no entanto não é suficiente para resolver todos os problemas possíveis. Assim, faz-se necessário uma definição mais geral que é dada a seguir.

Definição 4.2.3. (Definição Aximática de Probabilidade) *Seja ϵ um experimento e Ω o espaço amostral associado ao mesmo. A cada evento A desse espaço amostral associamos uma medida $P(A)$, denominada probabilidade de A, que satisfaz:*

i) $0 \leq P(A) \leq 1$;

ii) $P(\Omega) = 1$;

iii) se $A_1, A_2, \cdots, A_n \subset \Omega$ *forem disjuntos 2 a 2 (ou seja* $(A_i \cap A_j) = 0$, *para todo* $i \neq j$), *então* $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.

4.2.1 Propriedades

A partir dos três axiomas apresentados, podemos provar diversas propriedades da probabilidade. Vejamos algumas delas.

1) Se $\emptyset$ é o evento impossível, então $P(\emptyset) = 0$;

2) Se A e B são dois eventos quaisquer então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B);$$

3) Se $A \subset B$, então $P(A) \leq P(B)$;

4) $P(A^c) = 1 - P(A)$.

Exercício: Prove as propriedades acima.

Exemplo 4.2.3. *Considere um experimento e os eventos A e B associados, tais que* $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ *e* $P(A \cap B) = \frac{1}{4}$. *Encontre:*

a) $P(A^c)$ *e* $P(B^c)$;

b) $P(A \cup B)$;

c) $P(A^c \cap B^c)$;

d) $P(A^c \cup B^c)$;

e) $P(A^c \cap B)$.

Exemplo 4.2.4. *Suponha que um lote com 20 peças existam cinco defeituosas.*

- *Experimento: Escolher ao acaso quatro peças desse lote (uma amostra de 4 elementos) de modo que a ordem de retirada seja irrelevante.*
- *De quantas maneiras poderíamos obter essa amostra?*
- *Qual a probabilidade de ter duas peças defeituosas na amostra?*

4.2.2 Probabilidade Condicional

A **probabilidade condicional** é um dos conceitos mais importantes da teoria de probabilidade. Pois, frequentemente tem-se interesse em calcular probabilidades quando se tem alguma informação parcial a respeito do resultado de um experimento. Mesmo quando não se tem nenhuma informação parcial, as probabilidades condicionais podem ser frequentemente utilizadas para computar mais facilmente as probabilidades desejadas.

Exemplo 4.2.5. *Na Tabela abaixo temos dados referentes a alunos matriculados em quatro cursos de uma universidade em dado ano (Bussab, 2013, adaptado):*

Tabela 4.1: Distribuição de Alunos segundo o sexo e escolha de curso.

Sexo / Curso	Homens (H)	Mulheres (M)	Total
Matemática (M)	70	40	110
Engenharia de Software (S)	15	15	30
Estatística (E)	10	20	30
Computação (C)	20	10	30
Total	115	85	200

Seja o experimento: escolher um aluno ao acaso desses quatro cursos.

a) *Qual a probabilidade desse aluno ser do curso E (estatística)?*

b) *Qual a probabilidade desse aluno ser do curso de E e ser homem (H)?*

c) *Dado que o estudante esteja matriculado no curso E, qual a probabilidade de ser M?*

Resolução: a) 3/20; b) 1/20; c) 2/3.
Quando existe uma informação prévia de que um evento já ocorreu, como ocorre no item c), podemos levar em conta essa informação e usar a definição clássica de probabilidade para responder a questão.

Probabilidade Condicional e Teorema do Produto

Agora, vejamos como podemos obter formalmente a probabilidade condicional de um evento A dados que outro evento B ocorreu. Para isso, considere um espaço amostral Ω associado a um dado experimento. Sejam $A, B \subset \Omega$, então a probabilidade condicional do evento A dado que B ocorreu é dada por

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{com} \quad P(B) > 0, \text{ ou ainda}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{com} \quad P(A) > 0;$$

Isso implica que

$$P(A \cap B) = P(A|B) \cdot P(B) \quad \text{ou} \quad P(A \cap B) = P(B|A) \cdot P(A).$$

A relação acima é conhecida como **Teorema do Produto**.

Use as relações acima para responder a questão c) do Exercício 4.2.5.

Exemplo 4.2.6. *Uma urna contém 2 bolas brancas, 3 pretas e 4 verdes. Se duas bolas são retiradas ao acaso sem reposição, qual a probabilidade de que:*
a) ambas sejam verdes;
b) ambas sejam da mesma cor.
Resumo da resolução: (construir o diagrama em árvore): a) $P(V \cap V) = P(V|V)P(V)$*; b) basta fazer o mesmo procedimento para os demais e somar.*

Generalização do teorema do produto

$P(\bigcap_{i=1}^{n} A_i) = P(A_1)P(A_2|A_1)P(A_3|A_2 \cap A_1)\ldots P(A_n|A_1 \cap \ldots \cap A_{n-1})$.

Exemplo 4.2.7. *Uma urna contém 7 bolas brancas e 5 pretas. Retiramos três bolas da urna sem reposição. Assumindo que cada bola da urna é igualmente provável de ser retirada, qual a probabilidade de todas serem brancas?(construir o diagrama em árvore)*
Resumo da resolução: $P(B \cap B \cap B) = P(B|B \cap B)P(B|B)P(B)$

E se as retiradas forem feitas com reposição?

4.2.3 Eventos Independentes

Dois eventos são independentes quando a realização de um dos eventos não afeta a probabilidade de realização do outro e vice versa.

Definição 4.2.4. *A e B são eventos independentes se $P(A \cap B) = P(A) \cdot P(B)$.*

Exemplo 4.2.8. *Lançam-se três moedas. Verificar se são independentes os eventos:*
***A**: saída de cara na primeira moeda.*
***B**: Saída de coroa na segunda e cara na terceira moeda.*

Resolução: calcular $P(A)$; $P(B)$; $P(A \cap B)$; $P(A) \times P(B)$; verificar se $P(A \cap B) = P(A) \times P(B)$.

Exemplo 4.2.9. *Em uma caixa temos 10 objetos dos quais 4 são defeituosos. São retirados dois objetos ao acaso, um após o outro com reposição. Calcular a probabilidade de ambos serem bons.*
Resumo da resolução: $(6/10) \times (6/10) \approx 0,36$.

Exemplo 4.2.10. *Sejam A e B eventos tais que P(A)=0,2; P(B)= q e $P(A \cup B) = 0,6$. Calcular q considerando:*
a) mutuamente exclusivos;
b) independentes.

4.2.4 Exercícios para Seção 4.2

Exercício 4.2.1. *Prove as propriedades da probabilidade apresentadas anteriormente.*

Exercício 4.2.2. *O espaço amostral de um experimento aleatório é $\Omega = \{a, b, c, d, e\}$, em que cada evento elementar tem probabilidade 0,1; 0,1; 0,2; 0,4; e 0,2, respectivamente. Sejam os eventos $A = \{a, b, c\}$ e $B = \{c, d, e\}$. Determine:*
a) $P(A)$
b) $P(A^c)$
c) $P(A \cap B)$
d) $P(A \cup B)$
e) $P(B)$

Exercício 4.2.3. *Em uma bateria NiCd, uma célula completamente carregada é composta de hidróxido de níquel III. Níquel é um elemento que tem múltiplos estados de oxidação, sendo geralmente encontrado nos seguintes estados:*

Carga de Níquel	Proporção encontrada
0	0,17
+2	0,35
+3	0,33
+4	0,15

a) Qual a probabilidade de uma célula ter no mínimo uma das opções de níquel carregado positivamente? (R. 0,83)

b) Qual é a probabilidade de uma célula não ser composta de uma carga positiva de níquel maior que + 3? (R. 0,85)

Exercício 4.2.4. *Um lote de 100 chips semicondutores coném 20 itens defeituosos. Dois deles são selecionados ao acaso sem reposição.*

(a) *Qual a probabilidade de o primeiro chip selecionado seja defeituoso? (R: 20/100)*

(b) *Qual é a probabilidade de que o segundo chip seja defeituoso, dado que o primeiro já foi defeituoso? (R: 19/99)*

(c) *Qual é a probabilidade de que ambos sejam defeituosos? (R: 0,038)*

(d) *Como a resposta do item (b) mudaria se os chips selecionados fossem repostos antes da segunda seleção? (R: 0,2)*

Exercício 4.2.5. *A probabilidade de que um homem esteja vivo daqui a 30 anos é $\frac{2}{5}$; a de sua mulher é de $\frac{2}{3}$. Determinar a probabilidade de que daqui 30 anos:*

a) ambos estejam vivos; R:4/15

b) somente o homem esteja vivo; R:2/15

c) somente a mulher esteja viva; R:2/5

d) pelo menos um esteja vivo. R:4/5

Exercício 4.2.6. *Suponha que A e B sejam eventos mutuamente excludente (disjuntos). Construa um diagrama de Venn que contenha os três eventos A, B e C, de modo que:*
$P(A|C) = 1$ *e*
$P(B|C) = 0$.

Exercício 4.2.7. *Peças produzidas por uma máquina são classificadas como defeituosas (D), recuperáveis (R) e perfeitas (P) com probabilidades* $0,1$, $0,2$ *e* $0,7$, *respectivamente. De um grande lote foram sorteadas duas peças com reposição. Calcule:*
a)P(duas serem defeituosas).(R:0,01)
b)P(pelo menos uma ser perfeita). (R: 0,91)
c)P(uma ser perfeita e uma recuperável).(R:0,28)

Exercício 4.2.8. *Um time de futebol ganha com probabilidade* $0,7$ *se chove e com* $0,8$ *se não chove. Em setembro, a probabilidade de chuva é de* $0,3$. *O time de futebol ganhou uma partida em setembro, qual a probabilidade de ter chovido nesse dias? (R:* $0,27$*)*

4.3 Lei da Probabilidade Total e Teorema de Bayes

Uma das relações mais importantes envolvendo probabilidades condicionais é o **Teorema de Bayes**. Numa variedade de situações é útil ser capaz de derivar um conjunto de probabilidades condicionais a partir de um outro conjunto. O **Teorema de Bayes** relaciona probabilidades condicionais da forma $P(A|B)$ com probabilidades condicionais da forma $P(B|A)$, em que a ordem da condicionalidade é reversa.

Antes de apresentar o teorema de Bayes, convém lembrar a definição de probabilidade condicional, para registrar a diferença entre probabilidade condicional e o teorema de Bayes.

Vimos que a probabilidade de um evento A, dados que outro evento B ocorreu, é dada por:

$$P(A|B) = \frac{P(A \cap B)}{P(B)},$$

com $P(B) > 0$.

A partir da relação dada acima, obtemos a Regra ou Teorema do Produto:

$$P(A \cap B) = P(A|B)P(B).$$

Além disso, se A e B forem **independentes**, tem-se

$$P(A \cap B) = P(A)P(B).$$

Obs: essas relações podem ser generalizadas para um número finito de eventos.

Agora, considere uma sequência de eventos $A_1, A_2, \cdots, A_n$ tal que

$$A_i \cap A_j = \emptyset$$

para todo $i \neq j$. Ou seja, $A_1, A_2, \cdots, A_n$ são dois a dois disjuntos.

Vimos que se

$$\bigcup_{i=1}^{n} A_i = \Omega,$$

então $A_1, A_2, \cdots, A_n$ formam uma **Partição** do espaço amostral Ω, como mostra a Figura 4.2.

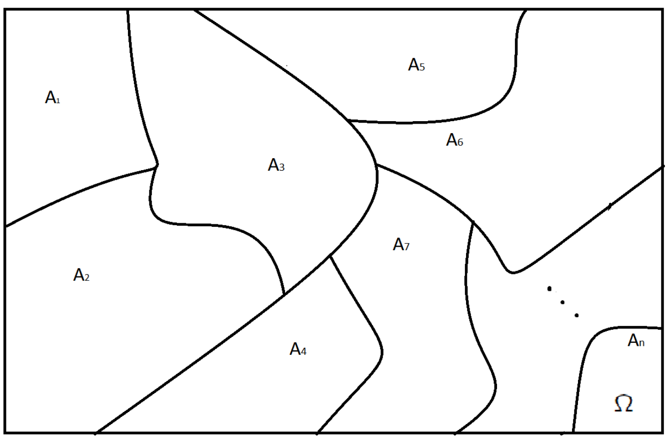


Figura 4.2: Representação geométrica de uma partição.

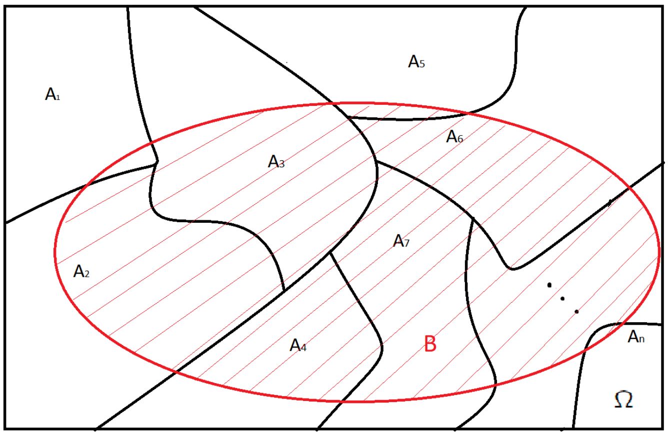


Figura 4.3: Representação geométrica de um e evento em uma partição.

Vamos considerar um evento B qualquer tal que $B \subset \Omega$, então B pode ser escrito como:

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \cdots \cup (B \cap A_n) \text{ (veja Figura 4.3.)}$$

Deste modo, a probabilidade total de B pode ser obtida pelo axioma III da probabilidade, como segue.

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \ldots P(B \cap A_n) \\ &= \sum_{i=1}^{n} P(B \cap A_i) \\ &= \sum_{i=1}^{n} P(B|A_i)P(A_i). \end{aligned}$$

Exemplo 4.3.1. *Suponha que você utilize peças de três fornecedores, que tem diferentes desempenhos quanto a sua qualidade. As peças são classificadas como conformes ou não conformes e você conhece a proporção de peças não conformes de cada fornecedor assim como a proporção recebida de cada fornecedor. Você recebe 20% de todas as peças que utiliza de F_1, 30% de outro fornecedor F_2 e 50% de F_3. Sendo que 20% das peças produzidas por F_1 são não conformes, enquanto F_2 e F_3, essa proporção é de 5% e 2%, respectivamente. Os lotes de peças são formados com peças dos três fornecedores. Se você selecionar, ao acaso, uma das peças do lote, qual é a probabilidade de ela ser não conforme?*

Resumo da resolução*: Se denotarmos B o evento "peças não conformes ", temos que $P(B|F_1) = 0,2$; $P(B|F_2) = 0,05$; $P(B|F_3) = 0,02$. Além disso, F_1, F_2 e F_3 formam uma partição do espaço amostral pois um lote de peças vem, necessariamente, de um dos três fornecedores. Desta forma, o evento A pode ser escrito em termos de interseções de A com os eventos F_1, F_2 e F_3, então:*

$$B = (B \cap F_1) \cup (B \cap F_2) \cup (B \cap F_3).$$

em que os eventos $(B \cap F_i)$ $(i = 1, 2, 3)$ são mutuamente exclusivos entre si. Logo:

$$\begin{aligned} P(B) &= P(B \cap F_1) + (B \cap F_2) + (B \cap F_3) \\ &= P(B|F_1)P(F_1) + (B|F_2)P(F_2) + (B|F_3)P(F_3). \end{aligned}$$

Exemplo 4.3.2. *Do exemplo anterior, dado que a peça é não conforme, qual a probabilidade de vir do fornecedor 1 (F_1)?*

$$P(F_1|B) = \frac{P(F_1 \cap B)}{P(B)}$$
$$= \frac{P(B|F_1)P(F_1)}{P(B|F_1)P(F_1) + P(B|F_2)P(F_2) + P(B|F_3)P(F_3)},$$

e, então

$$P(F_1|B) = \frac{0,2 \cdot 0,2}{0,2 \cdot 0,2 + 0,3 \cdot 0,05 + 0,5 \cdot 0,02} = 0,615$$

Portanto, a probabilidade de que a peça em questão tenha sido distribuído pelo F_1 é de 0,615 em contraste com as probabilidades 0,231 e 0,154 para os fornecedores F_2 e F_3, respectivamente.

O resultado usado para resolver o Exemplo 4.3.2 é conhecido como **Teorema de Bayes**.

Teorema de Bayes

Suponha que os eventos $A_1, A_2, \ldots A_n$ formem uma partição de Ω e que suas probabilidades sejam conhecidas. Suponha, ainda, que para um evento B, se conheçam as probabilidades $P(B|A_i)$ para todo $i = 1, 2, \ldots, n$. Então para qualquer j,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\displaystyle\sum_{i=1}^{n} P(B|A_i)P(A_i)}, j = 1, 2, \ldots, n.$$

Exemplo 4.3.3. *Um teste de laboratório detecta uma doença quando ela está presente em* 95% *dos casos. No entanto, o teste também fornece um resultado "falso positivo"para* 1% *das pessoas saudáveis testadas. (Isto é, se uma pessoa saudável faz o teste, então, com probabilidade* $0,01$*, o resultado do teste dirá que ela possui a doença.) Se* $0,5\%$ *da população tem a doença, qual é a probabilidade de uma pessoa ter a doença dado que o resultado do teste é positivo?*

4.3.1 Exercícios para a Seção 4.3

Exercício 4.3.1. *Um fabricante de impressoras obteve as seguintes probabilidades provenienes de um banco de dados de resultados de testes. Falhas nas impressoras estão associadas com três tipos de problemas: máquina (hardware), programa (software) e outros (tais como conectores), com probabilidades 0,1; 0,6 e 0,3, respectivamente. A probabilidade de uma falha na impressora devido a um problema de máquina é 0,9, devido a um problema de programa é 0,2 e devido a qualquer outro tipo de problema é 0,5. Se um consumidor entrar no site do fabricante para o diagnóstico da falha da impressora, qual será a causa mais provável do problema?*

(R.: a causa mais provável é "outros tipos".)

Exercício 4.3.2. *Uma rede local de computadores é composta por um servidor e cinco clientes (A, B, C, D e E). Registros anteriores indicam que dos pedidos de determinado tipo de processamento, realizados através de uma consulta, cerca de 10% vêm do cliente A, 15% do B, 15% do C, 40% do D e 20% do E. Se o pedido não for feito de forma adequada, o processamento apresentará erro. Usualmente, ocorrem os seguintes percentuais de pedidos inadequados: 1% do cliente A, 2% do cliente B, 1,5% do C, 20% do cliente D e 8% do cliente E.*

1. *Qual a probabilidade de o sistema apresentar erro?*

2. *Qual é a probabilidade de que o processo tenha sido pedido pelo cliente E, sabendo-se que apresentou erro?*

Exercício 4.3.3. *Em uma cidade em que os carros são testados para emissão de poluentes, 25% deles emitem quantidade considerada excessiva. O teste falha para 99% dos carros que emitem excesso de poluentes, mas resulta positivo para 17% dos carros que não emitem quantidade excessiva. Qual é a probabilidade de um carro que falha no teste realmente emitir quantidade excessiva de poluentes?*

Exercício 4.3.4. *Uma companhia de seguros acredita que as pessoas possam ser divididas em duas classes: aquelas que são propícias a sofrerem acidentes e as que não são. Suas estatísticas mostram que uma pessoa propícia a acidentes terá um acidente em algum momento dentro do período de um ano com probabilidade* $0,4$*, enquanto esta probabilidade diminui para* $0,2$ *para pessoas não propícias a acidentes. Supondo que* 30% *da população é propícia a sofrer acidentes, qual é a probabilidade de que um novo segurado sofra um acidente durante um ano em que comprou uma apólice?*

Capítulo 5

Variáveis Aleatórias

No Capitulo anterior consideramos um tipo de experimento que fornece um resultado que não pode ser previsto com certeza antes de sua realização, mas sabemos qual é o conjunto que contém todos os seus possíveis resultados. A esse tipo de experimento damos o nome de experimento aleatório e ao conjunto de todos os resultados possíveis do experimento aleatório damos o nome de espaço amostral do experimento, denotado aqui pela letra grega Ω. Qualquer subconjunto do espaço amostral, damos o nome de evento, aos quais associamos uma medida de probabilidade. Como o conjunto das partes do espaço amostral, $\mathcal{P}(\Omega)$, contém todos os eventos, ou seja todos os subconjuntos de Ω, essa medida de chance chamada de probabilidade é uma função definida em $\mathcal{P}(\Omega)$ e assume valores de 0 a 1. Formalmente, temos:

$$P : \mathcal{P}(\Omega) \longrightarrow [0, 1].$$

Nosso interesse ainda recai sobre atribuir medidas de chance aos eventos, no entanto, muitas vezes, para descrever um experimento aleatório é conveniente associar valores numéricos aos seus resultados, ou seja, aos eventos.

Para ilustrar uma situação como esta, considere que são inspecionadas 50 peças em uma linha de produção para verificar se cada uma delas tem ou não tem defeito. Se atribuirmos "D" se a peça inspecionada tem defeito e "ND" se a peça não tem defeito, o espaço amostral terá 2^{50} elementos, cada um deles sendo uma sequencia ordenada de D's e ND's de tamanho 50.

- É possível simplificar este espaço amostral?

Se o interesse for o número de peças que apresentam defeito (ou o número de peças boas) entre as 50, podemos definir:

X="número de D's (ou de ND's) registrados ao inspecionar as 50 peças".

O espaço amostral de X é o conjunto de números reais

$$S = \{0, 1, 2, ..., 50\},$$

que é muito mais fácil de lidar do que o espaço "amostral original" Ω além de manter a essência do problema.
Ao especificarmos a quantidade X, definimos uma transformação (uma função) a partir do espaço amostral original para um "novo espaço amostral" S, que é um conjunto de números reais. Essa função é chamada de **variável aleatória**.

Definição 5.0.1. (Variável aleatória, v.a.) *Uma variável aleatória X é uma função definida num espaço amostral Ω, que assume valores reais.*

- **Observação:** Em geral, variáveis aleatórias são denotadas com letras maiúsculas (X, Y etc) e os valores assumidos por elas são denotados por letras minúsculas correspondentes (x, y etc).

Quais valores podem assumir cada uma das seguintes v.a. ?

1. X_1: "número de conexões soldadas não conformes em uma placa de circuito impresso com 1000 conexões"

2. X_2: "Linhas em uso em um determinado período de tempo, em um sistema de comunicação por voz com 50 linhas "

3. X_3: "teor de umidade de um lote de matéria-prima, medido em percentual"

4. X_4: "número de falhas na superfície em uma serpentina de aço galvanizado"

Embora existam outros tipos de v.a's, aqui veremos as que mais surgem na prática, que são:

1. variáveis Aleatórias Discretas;

2. variáveis Aleatórias Contínuas.

As variáveis aleatórias discretas são aquelas que assumem valores em um conjunto finito ou infinito enumerável. As variáveis aleatórias contínuas são aquelas que assumem valores em um conjunto não enumerável (intervalos da reta). Inciaremos nossos estudos pelas variáveis aleatórias discretas, suas distribuições de probabilidades, variância e esperança. Em seguida, veremos as variáveis aleatórias contínuas com suas distribuições de probabilidades, esperança e variância.

5.1 Variáveis Aleatórias Discretas

Definição 5.1.1. *(Variável aleatória discreta) Uma* **variável aleatória discreta** *X é uma função definida num espaço amostral Ω, que assume valores reais,*

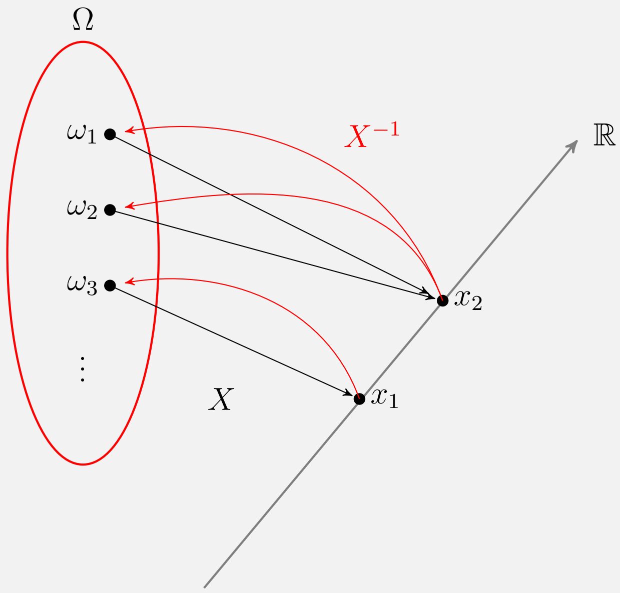
$$X : \Omega \mapsto \mathbb{R}$$

,

tal que:

$$[X = x_i] = X^{-1}(x_i) = \{\omega \in \Omega : X(\omega) = x_i\}$$

é um evento para todo $x_i \in \mathbb{R}$, ou seja para todo $i = 1, 2, 3, \cdots$.



Observe que:

- no diagrama acima
$$[X = x_1] = X^{-1}(x_1) = \{\omega_1\}$$
e
$$[X = x_2] = X^{-1}(x_2) = \{\omega_2, \omega_3\};$$
- $A = \{\omega_1, \omega_2\} = X^{-1}(x_2) = [X = x_2] \ \subset \ \Omega$, então se ω_1 ou ω_2 ocorre, podemos pensar que x_2 ocorre;
- de um modo geral $[X = x_i]$ representa a função inversa X^{-1} no ponto x_i e é um evento, ou seja
$$[X = x_i] = X^{-1}(x_i) = A \ \subset \ \Omega;$$
- como os eventos variam a cada realização do experimento, os valores numérico que lhes são atribuídos também variarão, fazendo sentido pensar em uma probabilidade associada ao valor numérico induzida pelo evento ao qual este valor está associado.

Exemplo 5.1.1. *Seja o experimento: lançar um dado e verificar a face superior.*

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Seja os eventos: A="a face superior é par"; B="a face superior é ímpar". Então,

$$A = \{2, 4, 6\} \quad e \quad B = \{1, 3, 5\}$$

Vamos definir a variável aleatória X tal que: $X(A) = 0$ e $X(B) = 1$.
Questões:

- *Como ficaria o diagrama de setas para essa v.a.?*
- *Qual é o espaço das amostras de X?*
- *Quem seria $X^{-1}(100) = [X = 100]$?*
- *Qual a probabilidade do evento $[X = 0]$?*
- *Qual a probabilidade do evento $[X = 1]$?*
- *Qual a probabilidade do evento $[X < 0]$?*

Exemplo 5.1.2. *Se o experimento consiste em retirar uma peça qualquer de uma linha de produção e verificar se ela tem defeito.*

- *Como podemos definir uma v.a. X para este experimento?*
- *Supondo que a probabilidade de sair uma peça com defeito é 0,1, como ficam as probabilidades em relação a X?*

Definição 5.1.2. *O conjunto formado pelos valores numéricos assumidos pela variável aleatória X é chamado de* **espaço das amostras de** X.

Nos dois exemplos anteriores podemos observar que existe uma variabilidade nos valores assumidos pela variável aleatória que resulta da variabilidade dos eventos. Assim, podemos pensar em uma probabilidade induzida pela transformação que associa cada evento a um número na reta. Ou seja, podemos pensar em probabilidades "induzidas" associadas a esse "novo espaço amostral", que consiste no conjunto de todos os resultados possíveis de X. Assim, o que nos interessa nas variáveis aleatórias são suas **distribuições de probabilidade** induzidas pelos eventos.

Distribuição de probabilidade de um v.a. discreta

Definição 5.1.3. *Chamamos de distribuição de probabilidades da v.a. discreta X a coleção de probabilidades:*

$$P(X = x_i), \text{ para todo } x_i \in \mathbb{R}, \tag{5.0}$$

tal que $[X = x_i] = X^{-1}(x_i) = \{\omega \in \Omega : X(\omega) = x_i\}$ é um evento para todo $x_i \in \mathbb{R}$.

Então, a distribuição de uma variável aleatória X é o conjunto de valores de probabilidades associados a cada possível valor de X.

Exemplo 5.1.3. ***Considere o experimento aleatório:*** *lançar um dado honesto e observar a sua face superior.*
Espaço amostral: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Defina a variável aleatória: *X:* $\Omega \to \mathbb{R}$,
tal que
X(1)=X(3)=X(5)=1,
X(2)=X(4)=2 e
X(6)=3.
Qual é o espaço das amostras de X?
Qual é a distribuição de probabilidades de X?

Se a variável aleatória discreta X tem poucos valores possíveis, ou seja, se o espaço das amostras de X tem poucos elementos, podemos representar sua distribuição de probabilidades em uma tabela.

Exemplo 5.1.4. *Considere que uma moeda é lançada duas vezes. Seja X a função definida no espaço amostral que é igual ao número de caras nos dois lançamentos (C - Cara e K - Coroa).*
A distribuição de probabilidade referente a variável aleatória X pode então ser representada como segue:

X	0	1	2
P(X=x)	1/4	1/2	1/4

5.1.1 Função de Probabilidade

Uma variável aleatória discreta pode ser caracterizada pela sua função massa de probabilidade (ou simplesmente função de probabilidade f.p.).

Definição 5.1.4. *Seja X uma variável aleatória discreta com valores em $\{x_1, x_2, ...\}$. A função massa de probabilidade (ou função de probabilidade f.p.) de X é definida por:*

$$p(x) = \begin{cases} P(X = x) & se\ x = x_i \\ 0 & se\ x \neq x_i\ \forall i. \end{cases}$$

A função de probabilidade de uma v.a. X atribui a cada elemento do espaço das amostras de X uma medida de probabilidade, caracterizando assim a distribuição de probabilidades de X. Em outras palavras, a função de probabilidade de uma v.a. discreta X é a função que atribui a cada valor do espaço das amostras de X, x_i para $i = 1, 2, \cdots$, sua probabilidade de ocorrência. Deste modo, essa função também pode ser referida como **função de frequência** da v.a. X.

Exemplo 5.1.5. *Considere novamente o Exemplo 5.1.3, neste caso podemos obter uma distribuição de probabilidade para X como segue.*
p(1)=P(X=1)=P({1,2,3})=P({1})+P({2})+P({3})=3/6=1/2,
p(2)= P(X=2)=P({4,5})=2/6=1/3,
p(3)=P(X=3)=P({6})=1/6. Esses valores podem ser organizados em uma tabela que vai representar a distribuição de probabilidades da v.a. X. Então, a funçao de probabilidade, ou função de frequência é dada por:

$$p(x) = \begin{cases} 1/2 & se \quad x = 1 \\ 1/3 & se \quad x = 2 \\ 1/6 & se \quad x = 3 \\ 0 & se \quad x \neq 1, 2, 3. \end{cases}$$

Propriedades da Função de Probabilidade

A função de probabilidade obedece várias propriedades, vejamos algumas delas.

1. $0 \leq p(x_i) \leq 1$ para todo $i = 1, 2, ...$
2. Se a sequência $x_1, x_2, ...$ inclui todos os possíveis valores de X, então $\sum_{i=1}^{\infty} p(x_i) = 1$.
3. Se X tem uma distribuição discreta, então

$$P(X \in I) = \sum_{x_i \in I} p(x_i), \text{ para qualquer intervalo } I \subset \mathbb{R}.$$

Exemplo 5.1.6. *As vezes bit transmitido por um canal de transmissão digital pode ser recebido com erro.*
Definimos:
X: "o número de bits com erro nos quatro próximos bits transmitidos"
Os possíveis valores para X são: $\{0, 1, 2, 3, 4\}$
Vamos supor a seguinte distribuição de probabilidade para X:

X	0	1	2	3	4
$p(x) = P(X = x)$	0,65	0,25	0,09	0,009	0,001

Ou seja, X tem função de probabilidade dada por:

$$p(x) = \begin{cases} 0,65 & se \quad x = 0 \\ 0,25 & se \quad x = 1 \\ 0,09 & se \quad x = 2 \\ 0,009 & se \quad x = 3 \\ 0,001 & se \quad x = 4 \\ 0 & se \quad x \neq 0, 1, 2, 3, 4. \end{cases}$$

Qual a probabilidade de X pertencer ao intervalo $[0, 2]$*?*
Qual é a probabilidade de X estar no intervalo $(-\infty, x]$ *para cada* $x = 0, 1, 2, 3, 4$*?*

Função de Distribuição Acumulada

Definição 5.1.5. *A Função de distribuição acumulada (f.d.a ou FDA) de uma v.a. discreta X é definida como*

$$F(x) = P(X \leqslant x) = \sum_{x_i \leqslant x} p(x_i),$$

para todo $x \in \mathbb{R}$.

Exemplo 5.1.7. *Suponha que uma produção diária de 850 peças fabricadas contenha 50 delas que não obedecem aos requerimentos do consumidor, duas peças são selecionadas ao acaso, sem reposição, da batelada. Seja a variável aleatória X o número de peças não conformes na amostra. Qual é a função de distribuição acumulada de X.*

Exemplo 5.1.8. *Determine a função de probabilidade de X a partir de sua f.d.a. dada a seguir:*

$$F(x) = \begin{cases} 0 \quad se \quad x < -2 \\ 0,2 \quad se \quad -2 \leq x < 0 \\ 0,7 \quad se \quad 0 \leq x < 2 \\ 1 \quad se \quad 2 \leq x \end{cases}$$

Construa, também, o gráfico da f.p. e da f.d.a.

5.1.2 Esperança, Média ou Valor Esperado de uma v.a. Discreta

Para motivar a definição de esperança de uma v.a., vamos supor uma urna com 2 bolas brancas, 2 vermelhas e 1 azul. Selecionamos ao acaso uma bola em que eu ganho:

$$X = \begin{cases} \text{R\$5,00}, & \text{se sair bola branca} \\ \text{R\$0,00}, & \text{se sair bola vermelha} \\ \text{R\$10,00}, & \text{se sair bola azul,} \end{cases}$$

Quanto devo pagar por partida para que o jogo seja honesto?

A variável aleatória assume valores no espaço de amostras $S = \{0, 5, 10\}$. Então podemos obter a seguinte função de probabilidade:

$$p(x) = \begin{cases} 0,4 & \text{se } x = 0 \\ 0,4 & \text{se } x = 5 \\ 0,2 & \text{se } x = 10 \end{cases}$$

Para que o jogo seja honesto (justo), o nosso investimento deve ser proporcional ao risco de perder. Ou seja, se existe 50% de chance de perdermos, devemos investir a metade do valor que poderíamos ganhar. No caso acima, devemos investir uma quantidade proporcional ao nosso risco. Como o risco é de 40% de perder tudo, 40% de ganhar 1frm-e do total e 20% de ganhar o total. Então, para que seja justo, devemos investir $0,4 \times 5 + 0,2 \times 10 = (2/5) \times 5 + (1/5) \times 10 = 4$ reais.

Definição 5.1.6. *A esperança (média ou valor esperado) de uma v.a. discreta X é a soma de todos os produtos possíveis da v.a. pela respectiva probabilidade.*

$$E(X) = \mu_X = \sum_{i=1}^{\infty} x_i \cdot P(X = x_i)$$

para $i = 1, 2, \cdots$.

Observações sobre esperança:

- Embora $E[X]$ seja denominada esperança da v.a. X, este valor depende da distribuição de X, podendo ser referida como esperança da distribuição da v.a. X.
- Duas variáveis aleatórias com mesma distribuição poderão ter a mesma esperança, mesmo que tenham significados diferentes.

Exemplo 5.1.9. *Uma seguradora paga R\$ 30.000,00 em caso de acidente de carro e cobra uma taxa de R\$ 1.000,00. Sabe-se que a probabilidade que um carro sofra acidente é de 3%. Quanto espera a seguradora ganhar por carro segurado?*

Propriedades da Esperança

Seja X e Y variáveis aleatórias discretas e $a, c =$constantes, então:

1. $E[c] = c$;
2. $E[c.X] = cE[X]$
3. $E[X \pm Y] = E[X] \pm E[Y]$
4. $E[aX \pm b] = aE[X] \pm b$
5. $E[X - E[X]] = 0$
6. $E[h(X)] = \sum_{i=1}^{\infty} h(x_i) \cdot P(X = x_i)$, em que $h(X)$ é uma função de X;
7. se $X_1, X_2, \cdots, X_n$ são variáveis aleatórias, então $E[\sum_{i=1}^{n} X_i] = \sum E[X_i]$.

Como exercício, prove essas propriedades.

Exemplo 5.1.10. *De acordo com conceitos administrativos, o lucro de uma empresa é dado pela expressão matemática:*

$$L = R - C,$$

em que L é o lucro, C o custo da produção e R a receita do produto (ganho bruto). A empresa mantém extensos registros das produções dos seus produtos. Para um determinado produto, com os dados coletados, foi construído a seguinte distribuição de probabilidade da variável aleatória

X=" número de itens produzidos por hora na empresa"

X	0	1	2	3	4	5
p(X)	0,04	0,2	0,35	0,30	0,1	0,01

a) *Calcule a média da produção diária* **(Y: quantidade de itens produzido POR DIA)** *desse produto, $\mu = \mathbb{E}[Y]$, sabendo que a empresa opera 24 horas por dia.*

b) *Dado que o custo do produto é dado pelo valor de produção de cada item vezes o número de itens fabricados, podendo ser descrito pela equação $C(Y) = Y^2 - 2Y$, e a receita é dada pelo número de itens produzidos no mês multiplicado pelo valor de venda, podendo ser descrito pela equação $R(Y) = 6Y - Y^2$. A partir dessas expressões, calcule o valor esperado do custo (C) e da receita (R).*

c) *Sabendo que $L = R - C$, qual é a expressão do lucro? Quanto a empresa espera receber de lucro (L) com as vendas dos itens produzidos em um único dia? Sabendo que $Var(Y) = \mathbb{E}[(24X)^2] - [\mathbb{E}(24X)]^2$, obtenha,também, a variância de Y.*

d) *Sabendo que $Var(L) = \mathbb{E}(L^2) - [\mathbb{E}(L)]^2$, obtenha a variância do lucro.*

O fato de conhecermos a média de uma distribuição de probabilidade já nos ajuda bastante em termos de resumir a distribuição, porém precisamos também de uma medida que nos dê o grau de dispersão de probabilidade em torno dessa média. Uma medida que fornece o grau da dispersão da distribuição da variável em torno da média é a **variância** da v.a.

5.1.3 Variância de um v.a. Discreta

A variância de uma variável aleatória fornece uma medida da dispersão de sua distribuição em torno de sua média. Assumindo que essa média e essa variância existam, definimos, para qualquer variável aleatória, a variância como segue.

Definição 5.1.7. *Definimos a variância de uma v.a. X qualquer como sendo a média dos quadrados dos desvios de X em torno da sua esperança, ou seja:*

$$Var(X) = \sigma^2 = E[(X - E[X])^2]$$

Da definição acima, decorre que a variância pode ser calcula por:

$$Var(X) = E[X^2] - (E[X])^2.$$

Se X é um v.a. discreta com valores em $\{x_1, x_2, \cdots\}$, então, na expressão acima, temos:

$$E[X^2] = \sum_{i=1}^{\infty} x_i^2 p(x_i).$$

5.2 Exercícios para a Seção 5.1

Exercício 5.2.1. *Suponha que X represente a diferença entre o número de caras e coroas obtido quando uma moeda é jogada 3 vezes.*

(a) *Quais são os possíveis valores de X?*

(b) *Se a moeda é honesta, quais são as probabilidades associadas aos valores que X pode assumir?*

Exercício 5.2.2. *Três homens e três mulheres são classificados de acordo com suas notas em uma prova. Suponha que não existam notas iguais e que todas as 6! classificações possíveis sejam igualmente prováveis. Faça X representar a melhor classificação obtida por uma mulher (por exemplo, $X = 1$ se a pessoa mais bem classificada for uma mulher, $X = 2$ se a pessoa classificada em segundo lugar for uma mulher etc). Apresente a distribuição de probabilidades de X em uma tabela.*

Exercício 5.2.3. *Seja X uma v.a. com FDA dada por:*

$$F(x) = \begin{cases} 0, \quad se \quad x < -1 \\ 0,25 \quad se \quad -1 \leq x < 3 \\ 0,75 \quad se \quad 3 \leq x < 5 \\ 1 \quad se \quad x \geq 5 \end{cases}$$

(a) *Construir gráfico de F(x).*

(b) *Determinar a função de probabilidade (p(x)) de X, E(X) e Var(X).*

(c) *Construir o gráfico da função de probabilidade $p(x)$.*

Exercício 5.2.4. *Seja X uma v.a. que assume valores em $S = \{0, 1, 2, 4, x\}$. Se cada valor em S é igualmente provável e $E(X) = 8$, qual é o valor de x?*

Exercício 5.2.5. *Considere que um produto pode estar perfeito (B), com defeito leve (DL) ou com defeito grave (DG). Seja a distribuição do lucro (em R$), por unidade vendida desse produto, dada pela Tabela 5.1.*

Tabela 5.1: Distribuição do lucro (R$).

Produto	X	p(x)
B	6	0,7
DL	0	0,2
DG	-2	0,1

(a) Calcule o valor esperado e a variância do lucro, ou seja E(X) e Var(X).

(b) Com a redução de desperdícios, foi possível aumentar uma unidade no lucro de cada unidade do produto, qual é o novo valor esperado e a nova variância do lucro por unidade?

(c) Se o lucro duplicar, qual é o valor esperado e a variância do novo lucro por unidade?

Exercício 5.2.6. *Você tem R$1000,00, e certa mercadoria é vendida atualmente por R$2,00 o quilo. Suponha que uma semana depois a mercadoria passe a ser vendida por R$1,00 ou R$4,00 o quilo, com essas duas possibilidades sendo igualmente prováveis ao longo da semana.*

(a) *Se o seu objetivo é maximizar a quantidade esperada de dinheiro que você possuirá no final da semana, que estratégia você deve empregar?*

(b) *Se o seu objetivo é maximizar a quantidade esperada de mercadoria que você possuirá no final da semana, que estratégia você deve empregar?*

Dica: analise as esperanças de ganho em dinheiro e mercadorias.

Exercício 5.2.7. *Na produção de uma peça são empregadas duas máquinas. A primeira é utilizada para efetivamente produzir as peças, e o custo de produção é de R$ 50,00 por unidade. Das peças produzidas nessa máquina, 90% são perfeitas. As peças defeituosas (produzidas na primeira máquina) são colocadas na segunda máquina para a tentativa de recuperação (torná-las perfeitas). Nessa segunda máquina o custo por peça é de R$ 25,00, mas apenas 60% das peças são de fato recuperadas. Sabendo que cada peça perfeita é vendida por R$ 90,00, e que cada peça defeituosa é vendida por R$ 20,00, calcule o lucro por peça esperado pelo fabricante.*

5.3 Principais Modelos Probabilísticos Discretos

Os modelos probabilísticos discretos são dados por distribuições de probabilidades de variáveis aleatórias discretas. Frequentemente nos referimos a essas distribuições como "distribuições discretas de probabilidades". Existem muitos modelos probabilísticos discretos (ou modelo para variáveis aleatórias discretas). Aqui veremos apenas os mais utilizados na prática.

5.3.1 Modelo Uniforme Discreto

Uma variável aleatória discreta segue um modelo uniforme discreto se todos os seus possíveis valores tem a mesma chance de ocorrer.

Considere o experimento que consiste no lançamento de um dado honesto, e estamos interessados na v.a.:

X: número da face superior.

Neste caso todos os possíveis resultados ocorrem com a mesma probabilidade e, assim, podemos dizer que a probabilidade se distribui **uniformemente** entre os diversos resultados:

X	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Definição 5.3.1. *Uma v.a. X assumindo valores em um conjunto finito,* $S = \{x_1, x_2, ..., x_n\}$*, segue um modelo uniforme discreto se sua f.p. é dada por:*

$$p(x) = \begin{cases} 1/n, & se \;\; x = x_i \; para \; algum \; i = 1, 2, \cdots, n \\ 0, & se \;\; x \neq x_i \;\; para \; todo \;\; i = 1, 2, \cdots, n. \end{cases}$$

Notação: $X \sim U^d(x_1, x_2, \cdots, x_n)$.

A esperança e a variância da v.a. discreta X seguindo um modelo uniforme são dadas por:

- $E[X] = \frac{\sum_{i=1}^{n} x_i}{n}$,
- $Var(X) = \frac{\sum_{i=1}^{n} x_i^2}{n} - (E[X])^2$.

Na Figura 5.1 mostra um gráfico de uma variável aleatória seguindo um modelo uniforme discreto.

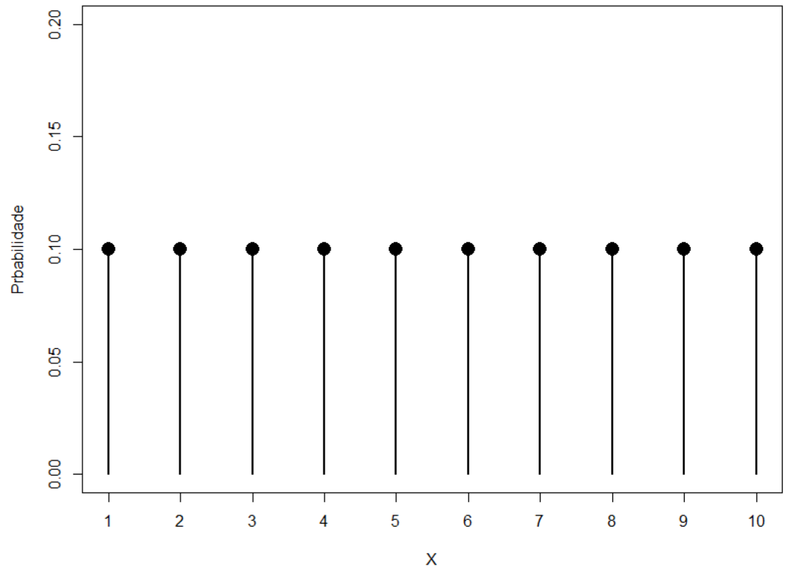


Figura 5.1: Modelo Uniforme Discreto para $X \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

5.3.2 Modelo de Bernoulli

Ensaios ou Experimentos de Bernoulli

Alguns experimentos têm alternativas dicotômicas, ou seja, possuem apenas dois resultados possíveis. Esses resultados podem ser representados por resposta do tipo "sucesso" ou "'fracasso", por exemplo:

- Uma peça é classificada como defeituosa ou não defeituosa;
  - não defeituosa: sucesso;
  - defeituosa: fracasso.
- O resultado de um exame médico para detecção de uma doença é positivo ou negativo;
  - positivo: sucesso;
  - negativo: fracasso.
- Um entrevistado concorda ou não com uma afirmação feita;
  - concorda: sucesso;
  - não concorda: fracasso.
- A direção que segue um veículo em uma bifurcação (A ou B);
  - direção A: sucesso;
  - direção B: fracasso.

Em homenagem ao matemático suíço James Bernoulli, esse tipo de experimento é denominado ensaios de Bernoulli. A partir desses ensaios, pode ser definida uma variável aleatória X que assume 0 se ocorre fracasso e 1 se ocorre o sucesso, em que o sucesso é o fenômeno que se deseja investigar e, geralmente, tem probabilidade pequena de ocorrer denotada por p. Ou seja,

$$X = \begin{cases} 0, & \text{se corre fracasso,} \\ 1, & \text{se ocorre sucesso.} \end{cases}$$

Uma v.a. discreta definida dessa forma é dita seguir uma **distribuição de Bernoulli**. Formalmente tem-se a definição a seguir.

Definição 5.3.2. *Uma variável aleatória X segue um modelo de Bernoulli com parâmetro p se sua função de probabilidade é dada por:*

$$p(x) = \begin{cases} p^x q^{1-x}, & se \;\; x = 0, 1 \\ 0, & se \;\; x \neq 0, 1, \end{cases} \qquad com \; q = 1 - p \; e \;\; 0 \leq p \leq 1.$$

Notação: $X \sim Bernoulli(p)$.

A esperança e a variância da variável aleatória de Bernoulli são dadas por:

- $E[X] = \sum_{i=1}^{2} x_i p(x_i) = 0 \times (1-p) + 1 \times p = p$,
- $Var(X) = E[X^2] - E[X]^2 = 0^2 \times (1-p) + 1^2 \times p - p^2 = p - p^2 = p(1-p)$.

5.3.3 Modelo Binomial

Considerando que n ensaios de um experimento de Bernoulli seja realizado de forma independente, a seguinte variável aleatória pode ser definida:

X: "N^o de sucessos em n ensaios de Bernoulli, independentes."

Neste caso X assume valores em $\{0, 1, 2, \cdots, n\}$. O espaço amostral desse experimento é composto por sequências de sucessos (S) e fracassos (F). Uma sequência com zero sucesso, pode ser escrita como:

$$FFFF \cdots FF$$

e uma sequência somente com sucessos pode ser escrita como:

$$SSSS \cdots SS.$$

Assim, sendo p, $0 \leq p \leq 1$, a probabilidade de sucesso, para essa sequência pode ser associado o seguinte probabilidade:

$p(0) = P(FFF \cdots F) = (1-p) \times (1-p) \times (1-p) \cdots (1-p) = (1-p)^n$

Seguindo o mesmo raciocínio, tem-se:

$p(1) = P(\text{"sequências com um sucesso"}) = np \times (1-p) \times (1-p) \cdots (1-p) = np(1-p)^{n-1}$

$p(2) = P(\text{"sequências com dois sucessos"}) = \binom{n}{2} p^2 (1-p)^{n-2}$

$p(3) = P(\text{"sequências com três sucessos"}) = \binom{n}{3} p^3 (1-p)^{n-3}$

$\vdots$

$p(n) = P(SSS \cdots SS) = p \times p \times p \cdots p = p^n$.

Então, a função de probabilidade da v.a. X pode ser escrita como:

$$p(x) = \left\{ \binom{n}{x} p^x (1-p)^{1-x}, \right.$$

para todo x no espaço das amostras de X. Neste caso X segue uma distribuição binomial. A Figura 5.2, mostra a distribuição de probabilidades para uma variável aleatória seguindo uma distribuição binomial com vetor de parâmetros $(10; 0,3)$ e a caracterização da variável aleatória é dada pela definição a seguir.

Definição 5.3.3. *Uma variável aleatória X assumindo valores em* $\Lambda = (x_1, x_2, \cdots, n)$ *segue um modelo binomial com parâmetros* n *e* p *se sua função de probabilidade é dada por:*

$$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{1-x}, & se \;\; x \in \Lambda, \\ 0 & se \;\; x \notin \Lambda. \end{cases}$$

Notação: $X \sim Binomial(n, p)$.

Pode ser mostrado que a esperança e a variância da distribuição binomial são dadas, respectivamente, por:

- E[X]=np,
- Var(X)=np(1-p).

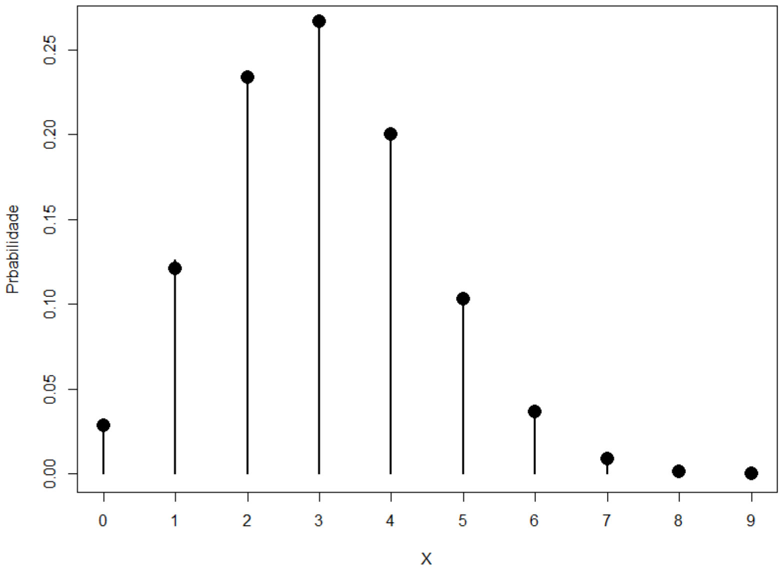


Figura 5.2: Gráfico das probabilidades da Binomial n=10 e p=0,3.

Exemplo 5.3.1. *Sabe-se que os parafusos produzidos por certa empresa têm probabilidade de 0,01 de apresentarem defeitos, independentemente uns dos outros. A empresa vende os parafusos em pacotes com 10 e oferece uma garantia de devolução de dinheiro se mais de 1 parafuso em 10 apresentar defeito. Que proporção de pacotes vendidos a empresa deve trocar?*

5.3.4 Modelo Hipergeométrico

Suponha que em um dia de produção de determinado produto 20 peças são fabricadas em um setor, das quais 5 não satisfazem os requerimentos dos consumidores.

Agora, suponha que duas peças são retiradas **sem reposição** dessa produção. Sejam A e B os eventos: a primeira e a segunda peças são não conformes, respectivamente.

Foi visto que,

$$P(B|A) = 4/19 = 0,21 \text{ e}$$

$$P(B) = P(B \cap A) + P(B \cap A^c) = 20/380 + 75/380 = 0,25$$

Nesta situação, pode ser visto que o conhecimento de que a primeira peça é não conforme torna menos provável que a segunda seja não conforme. Seja a variável aleatória definida como:

X:"N^o de peças não conforme na amostra".

Obtenha:
$P(X = 0)$
$P(X = 1)$
$P(X = 2)$

Essa distribuição é adequada quando consideramos extrações casuais feitas sem reposição de uma população dividida segundo dois atributos.

Suponha n objetos extraídos de forma aleatória de uma população de tamanho N, em que r objetos são do tipo 1 e N-r objetos são do tipo 2.

Se as retiradas são feitas **com reposição**, a v.a.

Y:"N^o de objetos do tipo 1 na amostra" $\sim$ Binomial(n,p=r/N)

Se as retiradas são feitas **sem reposição**, Y tem distribuição hipergeométrica com parâmetro (N,n,r), como definido a seguir.

Definição 5.3.4. *Uma v.a. X segue uma distribuição hipergeométrica com parâmetros N,n,r se sua função de probabilidade é dada por:*

$$p(y) = \begin{cases} \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}, & se \;\; y = 1, 2, \cdots, M, \qquad \text{em que } M = \min\{r, n\}, \\ 0, & se \;\; y \neq 1, 2, \cdots, M. \end{cases}$$

Aqui, tem-se que:

- *N é o total de elementos na população;*
- *n é o tamanho da amostras extraída de forma aleatória sem reposição;*
- *r é o número de elementos na população que são do tipo 1;*
- *Y é a v.a. que conta o número de elementos do tipo 1 na amostra.*

Notação: $Y \sim$ *Hipergeométrica*(N, n, r).

Se X é uma variável aleatória hipergeométrica sua esperança e variância são dadas por:

- $E(X) = np$ em que $p = \frac{r}{N}$;
- $Var(X) = np(1-p)(\frac{N-n}{N-1})$.

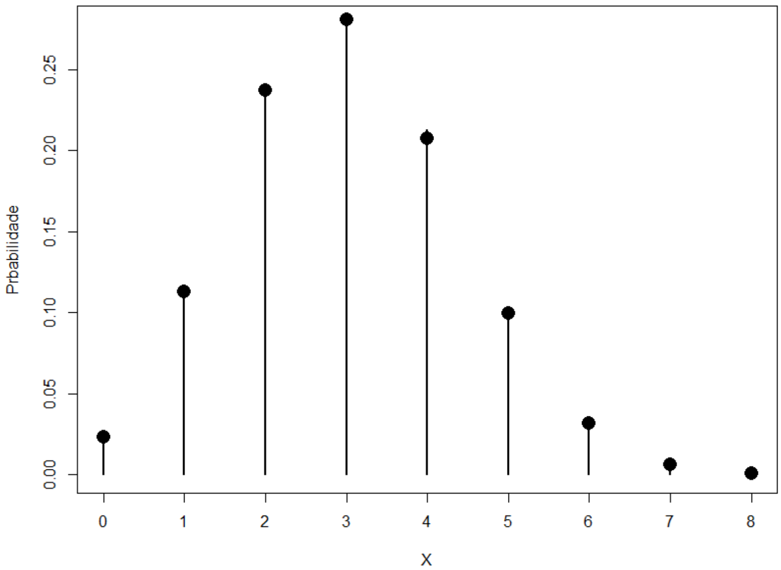


Figura 5.3: Modelo hipergeométrico com N=100, r=30 e n=10.

Exemplo 5.3.2. *Em problemas de controle de qualidade, suponha que num lote de N = 100 peças, r = 10 sejam defeituosas. Escolhendo n = 5 peças sem reposição.*

1. *Qual a probabilidade de não se obter peças defeituosas?*
2. *Qual a probabilidade de se obter no máximo uma com defeito?*
3. *E se as peças forem retiradas com reposição, como ficam essas probabilidades?*

5.3.5 Modelo de Poisson

Considere situações em que se avalia o número de ocorrências de um tipo de evento por unidade de tempo, de comprimento, de área, ou de volume, como por exemplo:

- O número de erros de impressão em uma página (ou em um grupo de páginas de um livro);

- O número de pessoas em uma comunidade que vivem mais de 100 anos;
- O número de telefone discados incorretamente em um dia;
- o número de falhas de um computador em um dia de operação;
- O número de partículas a descarregar por um material radioativo em um período de tempo fixo.

Esses são exemplos de um grupo de variáveis aleatórias discretas que assumem valores em um espaço qualquer. Esse tipo de variável muitas vezes pode ser modelada por uma distribuição de Poisson, que é definida a seguir.

Definição 5.3.5. *Seja X uma v.a. que pode assumir qualquer valor em $0, 1, 2, \cdots$. Então, X segue uma distribuição de Poisson com parâmetro $\lambda > 0$ se sua f.p. é dada por:*

$$p(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & se \;\; x = 0, 1, 2, \cdots \\ 0, & se \;\; x \neq 0, 1, 2, \cdots . \end{cases}$$

Notação: $X \sim Poisson(\lambda)$.

Pode ser mostrado que a variância e a esperança da distribuição de Poisson coincidem e são dadas por:

$$E(X) = Var(X) = \lambda.$$

Aqui, λ é uma taxa de ocorrência da variável em um período qualquer.

Portanto, a variável aleatória de Poisson conta o número de "sucesso" em um intervalo qualquer (tempo, espaço etc).

Aproximação da Binomial pela Poisson

A distribuição Binomial pode ser aproximada pela distribuição de Poisson.

Se $X \sim Bin(n, p)$, n é grande e p (probabilidade de Sucesso) é pequena, então podemos obter $E(X) = n \cdot p \approx \lambda$. Neste caso, podemos mostrar que:

$$p(x) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \approx \frac{e^{-\lambda}\lambda^k}{k!}.$$

Exemplo 5.3.3. *Considere novamente o exemplo da variável aleatória com distribuição binomial:*

Y:"N^o de parafusos com defeito em uma amostra de 10 parafusos"

em que

$$p = P(sucesso) = P(parafuso \ com \ defeito \ em \ uma \ retirada) = 0,01.$$

Vamos obter a probabilidade de mais de um parafuso com defeito usando a distribuição de Poisson e comparar o resultado obtido com o uso da distribuição binomial.

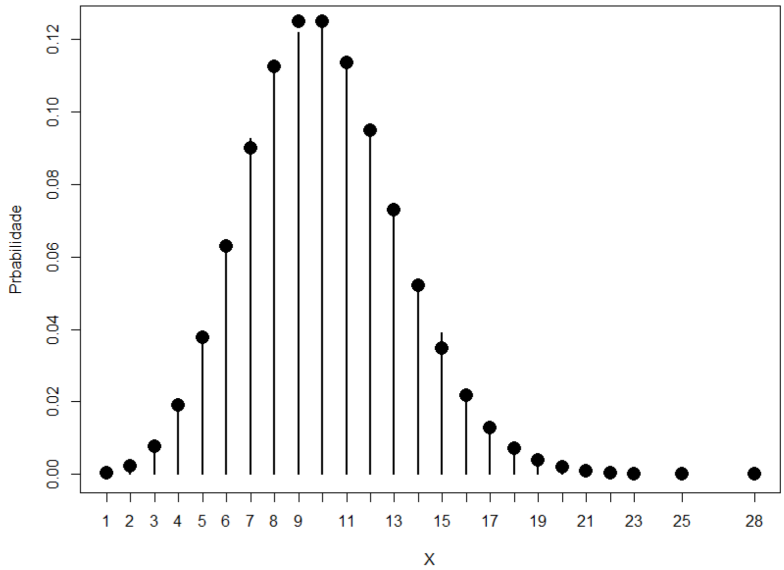


Figura 5.4: Grafico das probabilidades do modelo Poisson com $\lambda = 10$.

Exemplo 5.3.4. *A taxa de infecção por uma determinada doença infecciosa em um certo estado é de 1 por 100.000 habitantes cada mês.*

(a) *Determine a probabilidade de que em uma cidade de 400.000 habitantes, deste estado, ocorram 3 infecções ou mais em um dado mês, por essa doença.*

(b) *Qual é a probabilidade de que em pelo menos 2 meses durante o ano ocorram 3 infecções ou mais?*

5.4 Exercícios para a Seção 5.3

Exercício 5.4.1. *(Barbetta* et al.*, 2004, p. 136): Mensagens chegam ao servidor de acordo com uma distribuição de Poisson, com taxa média de cinco chegadas por minuto.*

a) *Qual é a probabilidade de que duas chegadas ocorram em um minuto?*

b) *Qual é a probabilidade de que uma chegada ocorra em 30 segundos?*

Exercício 5.4.2. *Uma entrevistadora tem consigo uma lista de pessoas que pode entrevistar. Se ela precisa entrevistar 5 pessoas, e se cada pessoa concorda (independentemente) em ser entrevistada com probabilidade $\frac{2}{3}$, qual é a probabilidade de que a sua lista permita que ela consiga obter o número necessário de entrevistas se ela contiver (a) 5 pessoas e (b) 8 pessoas? Na letra (b), qual é a probabilidade de que a entrevistadora fale com exatamente 6 pessoas?*

Exercício 5.4.3. *(Morettin pg. 120 adaptado) Uma remessa de 800 estabilizadores de tensão é recebida pelo controle de qualidade de uma empresa. São inspecionados 20 aparelhos da remessa, que será aceita se ocorrer no máximo um defeituoso. Há 100 defeituosos no lote. Qual a probabilidade de o lote ser aceito?*

Exercício 5.4.4. *(Montgomery pg. 58) Um produto eletrônico contém 40 circuitos integrados. A probabilidade de que qualquer circuito integrado seja defeituoso é de 0,01. Os circuitos operam de forma independente. O produto opera somente se não houver circuitos integrados defeituosos. Qual é a probabilidade de que o circuito opere?*

Exercício 5.4.5. *(Morettin Pg. 123 adaptado) O CRH de uma firma entrevista 150 candidatos a emprego por hora. Qual a probabilidade de entrevistar:*

a) *no máximo 4 candidatos em 3 minutos?*

b) *exatamente 3 candidatos em 4 minutos?*

Exercício 5.4.6. *(Morettin Pg. 122 adaptado) De acordo com a Divisão de Estatística Vital do Departamento de Saúde dos Estados Unidos, a média anual de afogamentos acidentais neste país é de 3 por 100.000 indivíduos. Determinar a probabilidade de que em uma cidade com 300.000 habitantes se verifiquem:*

a) *Nenhum afogamento.*

b) *No máximo 2 afogamentos.*

c) *Pelo menos um afogamento.*

Exercício 5.4.7. *(Bussab p. 157): Num certo tipo de fabricação de fita magnética ocorrem cortes a uma taxa de 1 por 2000 pés. Qual é a probabilidade de um rolo com 2000 pés de fita magnética tenha:*

a) *nenhum corte;*

b) *no máximo 2 cortes e*

d) *pelo menos 2 cortes?*

Exercício 5.4.8. *(Montgomery P. 65 adaptado) Cartões de circuito integrado são verificados em um teste funcional depois de serem preenchidos com chips semicondutores. Um lote contém 140 cartões, e 20 são selecionados sem reposição para o teste funcional.*

(a) *Se 20 cartões forem defeituosos, qual é a probabilidade de no mínimo um cartão defeituoso estar na amostra?*

(b) *Responda o item (a) usando a distribuição binomial e compare com o resultado obtido em (a) pela hipergeométrica.*

5.5 Variáveis Aleatórias Contínuas

Foi visto anteriormente que uma variável aleatória discreta assume valores em um conjunto contável, $x_1, x_2, \cdots$. No entanto existem outros tipos de espaço de amostras, como por exemplo, os espaços contínuos. Assim, uma variável aleatória (absolutamente) contínua X é uma função definida num espaço amostral Ω, que assume valores em um intervalo da reta (conjunto não enumerável de valores).

São exemplos de variáveis aleatórias contínuas:

- o tempo de duração de uma chamada telefônica;
- a altura da água em uma represa;
- comprimento de parafusos em uma linha de produção;
- pesos de bebês ao nascer;
- o tempo de vida de uma lampada.

Assim como no caso discreto, aqui o interesse recai na distribuição de probabilidades da variável aleatória. No entanto, não é possível atribuir probabilidades aos valores de uma variável aleatória contínua da mesma maneira que fazemos para as variáveis aleatórias discretas. Pois a soma de uma quantidade não enumerável de números positivos não poderia ser igual a 1. Portanto, em vez de atribuir probabilidades aos valores da variável, podemos atribuir probabilidades a intervalos de valores dessa variável. Formalmente, tem-se a seguinte definição de variável aleatória contínua.

Definição 5.5.1. *Dizemos que X tem uma distribuição contínua ou é uma variável aleatória contínua se existe uma função não-negativa $f : \mathbb{R} \to [0, \infty)$ que tenha a propriedade de que, para qualquer conjunto I de números reais,*

$$P(X \in I) = \int_I f(x)dx. \tag{5.0}$$

A função f é chamada de função densidade de probabilidade (f.d.p.) da variável aleatória X.

Ou seja, a probabilidade de que X esteja em I pode ser obtida integrando a função densidade de probabilidade ao longo do conjunto $I \in \mathbb{R}$.

Observação: a definição acima é válida somente para os conjuntos denominados mensuráveis de $\mathbb{R}$, no entanto, esses conjuntos incluem todos os conjuntos de interesse prático. Portanto, aqui, a preocupação com essas questões técnicas não é necessária.

A região sombreada no gráfico apresentado na Figura 5.5 corresponde a probabilidade de $X \in I$ quando $I = [a, b]$.

Figura 5.5: $P(a \leq X \leq b)$ =área sombreada.

Densidade empírica na análise estatística

Na análise de dados proveniente de variáveis aleatórias contínuas, não é possível saber com certeza qual é a função de densidade que gera os dados, assim como ocorre em alguns casos discretos. Por essa razão, deve-se buscar meios de se ter uma ideia de qual densidade seria mais apropriada para descrever um conjunto de dados particular. Em muitos casos, pode-se construir um histograma e, a partir deste, um **polígono de frequência**. Esses gráficos podem ser pensados como a densidade empírica desse conjunto de dados e podem fornecer uma ideia da forma da função densidade de probabilidade da variável aleatória que gerou esses dados. Isso torna esses gráficos de extrema importância em muitas análise estatística.

Relembrando a construção do histograma.

- O histograma é um gráfico de barras contíguas de bases iguais (intervalos de mesmo comprimento) e com áreas proporcionais a proporção de valores dos dados em cada intervalo (f_i).
- No histograma a altura de cada retângulo é proporcional a frequência de ocorrência da variável no intervalo que constituí a base, e, por conveniência, deve ser obtida dividindo a frequência relativa de ocorrência pela amplitude da base (ou f_i/Δ), Δ o comprimento das bases dos retângulos.
- Quanto mais dados contiver o intervalo de uma determinada base, mais alto será o retângulo correspondente.

Como exemplo, vejamos a frequência de um grupo de estudantes segundo a variável altura.

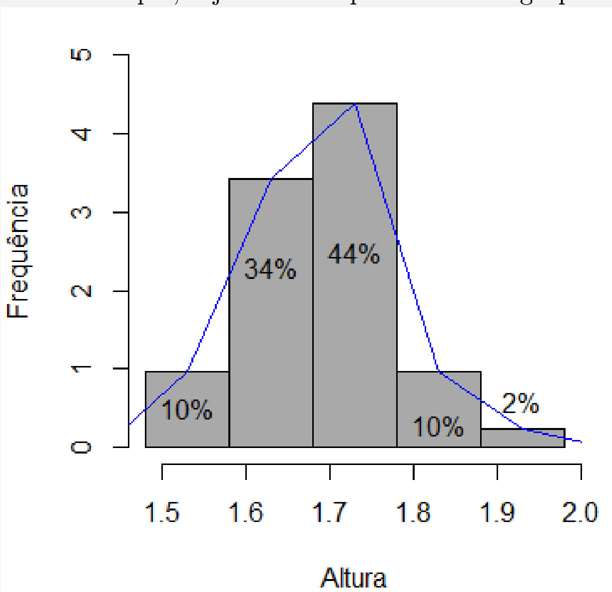


Tabela de frequência em intervalos de classes.

X	n_i	f_i
$1,48 \dashv 1,58$	4	$4/41 \approx 0,10$
$1,58 \dashv 1,68$	14	$14/41 \approx 0,34$
$1,68 \dashv 1,78$	18	$18/41 \approx 0,44$
$1,78 \dashv 1,88$	4	$4/41 \approx 0,10$
$1,88 \dashv 1,98$	1	$1/41 \approx 0,02$
Total	41	1

Em que n_i é a frequência e f_i é a proporção de dados no intervalo.

Se os dados foram coletados de forma aleatória (ao acaso) de um grupo maior, que denominamos população, podemos definir a variável aleatória "altura dos indivíduos da população" e utilizar o polígono de frequência para se ter uma ideia da forma da densidade que melhor descreve esses dados. A partir dessa ideia, pode-se realizar testes estatísticos que vão confirmar ou não essa suposição inicial. Se esses testes não confirmam essa suposição inicial, todos os resultados alcançados a partir disto perdem sua validade é não devem ser levados em consideração na tomada de decisão.

Propriedades da função densidade de probabilidade

Se X é uma variável aleatória contínua com função densidade de probabilidade $f(x)$, então para qualquer intervalo I da reta tem-se:

$$P(X \in I) = \int_I f(x)dx$$

em que $f(x)$ é sua função densidade de probabilidade e tem as seguintes propriedades:

(i) $f(x) \geqslant 0$ para todo $x \in \mathbb{R}$;

(ii) $P\left(X \in (-\infty, \infty)\right) = P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)dx = 1$.

(iii) Se $I = [a, b] \in \mathbb{R}$, então $P(X \in [a, b]) = P(a < X < B) = \int_a^b f(x)dx$.

(iv) Se $a = b$, então $P(X \in (a, b)) = P(a < X < a) = \int_a^a f(x)dx = 0$

Note que, pela propriedade (iv), a probabilidade de uma variável aleatória contínua assumir um valor qualquer da reta, digamos k, é sempre zero, pois $P(X = k) = \int_k^k f(x)dx = 0$. Deste modo, a inclusão ou não dos extremos a e b do intervalo $I = [a, b]$ na obtenção de $P(X \in [a, b])$ é irrelevante, ou seja:

$$P(X \in [a, b]) = P(a \leq X \leq b) = P(a < X < b) = \int_a^b f(x)dx.$$

Fato: Uma função real f que satisfaz (i) e (ii) é função densidade de probabilidade de alguma variável aleatória contínua X. Isso significa que para se ter uma função densidade de probabilidade, basta que se tenha uma função real positiva e obter a constante normalizadora que forneça a área abaixo da curva igual a 1.

Exemplo 5.5.1. *Consideremos a função abaixo:*

$$f(x) = \begin{cases} c(x^2 + x), & se\ 0 \leq x \leq 1; \\ 0, & caso\ contrário. \end{cases}$$

Encontre c de modo que $f(x)$ seja uma função densidade de probabilidade de alguma variável aleatória X.
Resolução:

- *Como $f(x) > 0$ para $0 \leq x \leq 1$ e $c > 0$, para a verificação, basta integrar a função f(x) em todo o seu domínio e verificar se esta integral tem valor 1.*
- *Ou seja,*

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} 0dx + \int_0^1 c[x^2+x]dx + \int_1^{\infty} 0dx = c\int_0^1 (x^2 + x)\, dx = c\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_0^1 = c\left(\frac{1}{3} + \frac{1}{2}\right) = \frac{5c}{6} = 1 \Rightarrow c = \frac{6}{5}.$$

Neste caso, $c = 6/5$ é a constante normalizadora que torna essa função positiva em uma f.d.p..

Exemplo 5.5.2. *Suponha que quantidade de tempo em que um computador funciona sem estragar é uma variável aleatória X contínua com função densidade de probabilidade*

$$f(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & se\ x > 0; \\ 0 & caso\ contrário. \end{cases}$$

(a) *Obtenha λ de modo que f(x) seja uma f.d.p.*
(λ é a constante normalizadora).

(b) *Usando a f.d.p., obtenha a probabilidade de que o computador funcione entre 50 e 150 horas.*

Resolução: *Como*

$$1 = \int_{-\infty}^{\infty} f(x)dx = \lambda \int_0^{\infty} e^{-\frac{x}{100}} dx,$$

obtemos

$$1 = -\lambda \times 100 \times e^{-\frac{x}{100}}|_0^{\infty} = 100 \times \lambda \Rightarrow \lambda = \frac{1}{100}.$$

$P(50 < X < 150) = \int_{50}^{150} \frac{1}{100} e^{-\frac{x}{100}} dx = -e^{-\frac{x}{100}}|_{50}^{150} = e^{-1/2} - e^{-3/2} \cong 0,384.$

Exemplo 5.5.3. *Suponha que X seja uma variável aleatória contínua cuja função densidade de probabilidade é apresentada como:*

$$f(x) = \begin{cases} \frac{3}{8}\left(4x - 2x^2\right) & se\ 0 < x < 2; \\ 0 & caso\ contrário. \end{cases}$$

(a) *$f(x)$ é de fato uma função densidade de probabilidade? Justifique.*

(b) *Se sim, determine $P(X > 1)$.*

Exemplo 5.5.4. *Seja X uma variável aleatória tal que sua função densidade de probabilidade seja $f(x)$ definida abaixo, com c sendo uma constante.*

$$f(x) = \begin{cases} c, & se \ -1 \ < \ x \ < \ 5; \\ 0, & caso \ contrário. \end{cases}$$

Qual deve ser o valor da constante c para que $f(x)$ cumpra os requisitos para ser uma função densidade de probabilidade?

5.5.1 Função de Distribuição Acumulada

Assim como no caso discreto, o interesse no estudo das variáveis aleatórias contínuas recai sobre sua distribuição de probabilidades. Esse distribuição é definida formalmente a seguir.

Definição 5.5.2. *A distribuição da variável aleatória X é a coleção de probabilidades $P(X \in I)$ para todo $I \subset \mathbb{R}$.*

Vimos que a função densidade de probabilidade fornece a coleção de probabilidades requerida, caracterizando, assim, a distribuição da variável aleatória. De modo análogo ao visto no caso discreto, existem também outras funções que caracterizam essa distribuição, aqui, como antes, apresentamos a Função de Distribuição Acumulada.

Definição 5.5.3. *(Função de Distribuição Acumulada, FDA) Na Definição 5.5.2, suponha $I = (-\infty, x)$, neste caso, supondo que $f(x)$ seja a f.d.p. de X, a função de distribuição acumulada de X é dada por:*

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt,$$

para todo $x \in \mathbb{R}$.

Note que, para obtermos a probabilidade da v.a. X estar em um intervalo [a,b], basta calcularmos sua f.d.a e obter:

$$P(a < x < b) = P(a \leq x \leq b) = F(b) - F(a).$$

Além disso, derivando a função de distribuição de uma v.a. X, obtemos a sua f.d.p..

Exemplo 5.5.5. *Suponha que a quantidade de tempo que um computador funciona sem estragar é uma variável aleatória X contínua com função densidade de probabilidade*

$$f(x) = \begin{cases} \frac{1}{100}e^{-\frac{x}{100}} & se \ x > 0; \\ 0 & caso \ contrário. \end{cases}$$

(a) *Obtenha a FDA de X.*

(b) *Usando a FDA, obtenha a probabilidade de que o computador funcione entre 50 e 150 horas.*

5.5.2 Esperança e Variância de um v.a. Contínuas

Uma variável aleatória contínua X, com função de probabilidade $f(x)$, tem esperança (média ou valor esperado) dada por:

$$E(X) = \mu = \int_{-\infty}^{\infty} xf(x)dx;$$

Assim como visto no caso discreto, a variância de uma v.a. X fornece uma medida do grau de dispersão da distribuição de X em torno de sua média. Assumindo a existência da esperança de X e X^2, obtemos a variância de X como segue:

$$Var(X) = \sigma^2 = E(X^2) - E(X)^2.$$

As interpretações dessas medidas podem ser feitas de forma análoga ao caso discreto. Além disso, todas as propriedades enunciadas para o caso discreto continuam válidas para o caso contínuo.

Exemplo 5.5.6. *Seja X uma variável aleatória contínua com função densidade de probabilidade dada por:*

$$f(x) = \begin{cases} 2x & se\ 0 \leq x \leq 1; \\ 0 & caso\ contrário. \end{cases}$$

a) *Determine a esperança, $E(X)$, e a variância, $Var(X)$, de X.*

b) *Se $Y = 2X + 10$, qual é a esperança e a variância de Y?*

5.5.3 Principais Modelos Contínuos

Os modelos probabilísticos sã expressos pelas distribuições de probabilidade das variáveis aleatórias. Então, um modelo probabilístico pode ser caracterizado pela sua função densidade de probabilidade. Aqui são apresentados os modelos probabilísticos contínuos mais utilizados na prática e esta apresentação se dá pela descrição de sua função densidade de probabilidade.

Modelo Exponencial

Definição 5.5.4. *Uma v.a. X segue uma distribuição exponencial de parâmetro λ, com $\lambda > 0$, se sua função densidade é dada por:*

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & se \quad x \geq 0 \\ 0, & se \quad x < 0. \end{cases}$$

- Notação: $X \sim Exp(\lambda)$.
- Pode ser mostrado que se $X \sim Exp(\lambda)$, então sua esperança e sua variância são dadas, respectivamente, por:

$$E(X) = \frac{1}{\lambda} \quad \text{e} \quad Var(X) = \frac{1}{\lambda^2}.$$

A Figura 5.6 mostra algumas formas da densidade da distribuição exponencial, considerando valores de $\lambda \in 0,2; 0,5; 1; 5$.

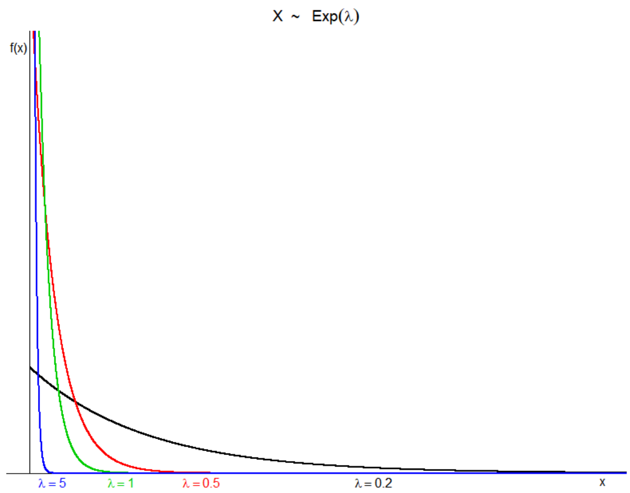


Figura 5.6: Função densidade de probabilidade da distribuição exponencial para alguns valores de λ.

A função de distribuição acumulada de $X \sim Exp(\lambda)$ pode ser obtida a partir de sua f.d.p. e é dada por:

$$F(x) = \int_0^x \lambda e^{-\lambda x} ds = 1 - e^{-\lambda x}.$$

A distribuição exponencial tem forte ligação com a distribuição discreta Poisson. Enquanto esta pode ser usada para modelar o número de ocorrências em um período (de tempo ou comprimento), a distribuição exponencial pode modelar a variável aleatória contínua que representa o intervalo (de tempo ou comprimento) entre as ocorrências. Exemplos:

a) tempo (em minutos) até a próxima inspeção a uma obra;

b) tempo (em segundos) entre pedidos a um servidor;

c) distância (em metros) entre defeitos de uma fita.

A distribuição exponencial pode ser usada quando as suposições de Poisson (independência entre as ocorrências e taxa média de ocorrência constante no intervalo considerado) estiverem satisfeitas.

Na prática, a distribuição exponencial surge como a distribuição da quantidade de tempo até que ocorra algum evento específico. Por exemplo, a quantidade de tempo (a partir deste momento) até a ocorrência de uma chuva.

Exemplo 5.5.7. *O tempo de vida de um tipo de fusível segue uma distribuição exponencial com vida média de 100 horas. Cada fusível tem um custo de R$10,00 e se durar menos de 200 horas há um custo adicional de R$8,00.*

(b) *Qual é a probabilidade de um fusível durar mais de 200 horas?*

(c) *Determinar o custo esperado.*

Modelo Normal

O modelo normal é um dos mais importantes modelos probabilístico, sendo aplicado em inúmeros fenômenos e constantemente utilizado para desenvolvimento teórico da inferência estatística. O modelo normal é definido pela **Distribuição Normal** dada a seguir.

Definição 5.5.5. *Uma variável aleatória X segue uma distribuição normal com parâmetros μ e σ^2, se sua função densidade é dada por:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \text{ para todo } x \in \mathbb{R},$$

em que $-\infty < \mu < \infty$ e $\sigma > 0$.

Notação: $X \sim N(\mu, \sigma^2)$.

Como pode ser visto nas Figuras 5.7, 5.8, 5.9 e 5.10, a curva da densidade da distribuição normal, $f(x)$, tem forma de sino, de modo que:

- $f(x)$ atinge seu máximo em μ;
- a curva de $f(x)$ é simétrica em relação a μ;
- $\mu - \sigma$ e $\mu + \sigma$ são pontos de inflexão de $f(x)$.

Disto decorre que a esperança e a variância da variável seguindo distribuição normal são dadas, respectivamente, por:

- $E(X) = \mu$;
- $Var(X) = \sigma^2$.

Logo, o desvio-padrão é $DP = \sigma$. Portanto, os parâmetros da distribuição normal são diretamente dados pela média e pelo desvio-padrão da variável aleatória.

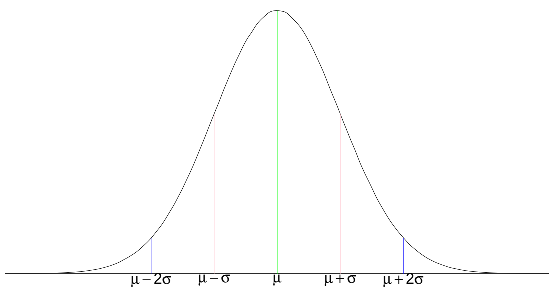


Figura 5.7: Curva de densidade da normal com alguns quantis úteis.

A curva da densidade da distribuição normal é simétrica (com forma de sino) como mostra a Figura 5.7, em que podem ser observados os seguintes intervalos:

- O intervalo $\mu \pm \sigma$ contém 68,27 % dos valores da variável.
- O intervalo $\mu \pm 2 \times \sigma$ contém 95,45 % dos valores da variável.

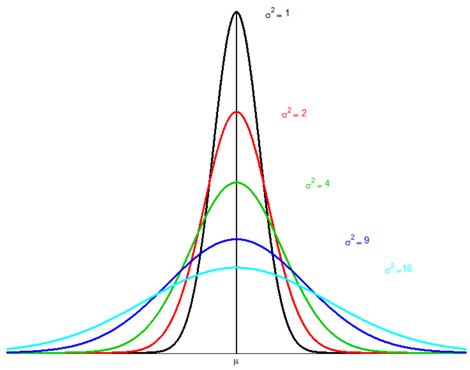


Figura 5.8: Curva da densidade da normal para valores variados da variância e mesma média.

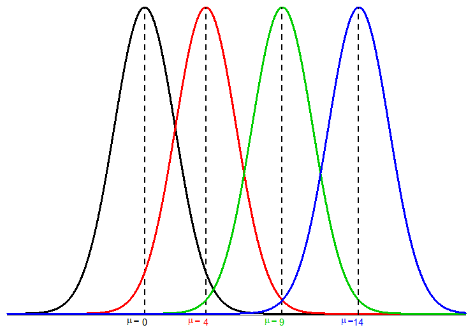


Figura 5.9: Curva de densidade da normal para valores variados da média e mesma variância.

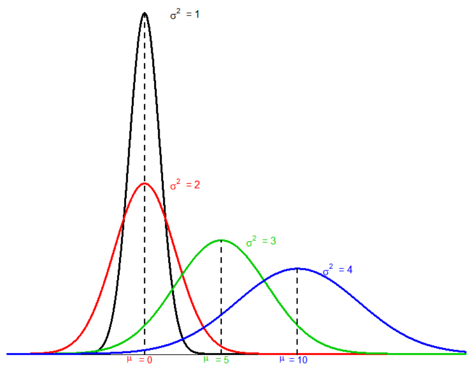


Figura 5.10: Curva de densidade da normal para valores variados de média e variâncias.

Para calcular a probabilidade da variável com distribuição normal assumir valores em um intervalo, faz-se necessário o uso de integração da densidade, como vimos anteriormente. No entanto, essa integração não é possível de ser realizada analiticamente. Então, faz-se necessário o uso de integração numérica. Para facilitar o obtenção dos valores de probabilidades que são muito utilizados na prática, foi feita a construção de tabelas que contêm esses valores de probabilidades de forma organizada e de fácil utilização. Como construir uma tabela para cada combinação de μ e σ não é viável, nem mesmo possível, a ideia é criar uma padronização, de modo que qualquer distribuição normal seja obtida a partir de uma única distribuição. Assim, pode-se criar uma tabela com os principais valores de probabilidade para essa distribuição, que é conhecida como **"distribuição normal padrão"**.

Distribuição normal padrão

Definição 5.5.6. *A função de distribuição acumulada da normal é dada pela integral*

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt, \text{ para todo } x \in \mathbb{R},$$

No entanto, como já mencionado, essa integral não tem solução analítica. Então, para obter uma distribuição padrão que possa ter os seus principais valores de probabilidades tabelados, define-se a seguinte transformação:

$$Z = \frac{X - \mu}{\sigma}.$$

Note que, pelas propriedades da esperança e da variância

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{E(X) - \mu}{\sigma} = 0 \text{ e}$$

$$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{Var(X)}{\sigma^2} = 1$$

Como uma transformação linear de uma variável aleatória com distribuição normal leva a outra distribuição normal, tem-se:

$$Z \sim N(0, 1).$$

Neste caso, Z é dita ter distribuição Normal Padrão.

Como já foi dito, os principais valores de probabilidades da distribuição normal padrão pode ser encontrados em tabelas. Usando essas tabelas, podemos obter os valores de probabilidades para qualquer distribuição normal, como segue:

$$Z = \frac{X-\mu}{\sigma} \Longrightarrow \quad P(X \leq a) = P\left(\frac{X-\mu}{\sigma} \leq \frac{a-\mu}{\sigma}\right) = P(Z \leq \frac{a-\mu}{\sigma}),$$

essa última igualdade pode ser resolvida por aproximação usando uma tabela de probabilidades normais.

Exemplo 5.5.8. *Seja $X \sim N(100; 25)$. Usando uma tabela de probabilidades normais, calcule:*

(a) $P(100 \leq X \leq 106)$

(b) $P(89 \leq X \leq 107)$

(c) $P(112 \leq X \leq 116)$

(d) $P(X \leq 108)$

(e) *em $P(X \leq q) = 0,8$, o valor de q.*

Exemplo 5.5.9. *Suponha que a medida da corrente em pedaços de fios segue distribuição normal com média 10 miliampéres e variância de 4 (miliampéres)^2.*

(a) *Qual é a probabilidade da medida exceder 13 miliampéres?*

(b) *Qual é a probabilidade da medida estar entre 9 e 11 miliampéres?*

(c) *Determine o valor para o qual a probabilidade de uma medida da corrente estar abaixo desse valor seja 0,98.*

5.5.4 Exercícios para a Seção 5.5

Exercício 5.5.1. *Suponha que uma variável aleatória contínua X tenha função densidade de probabilidade dada por:*

$$f(x) = \begin{cases} \frac{1}{6}x + k, & 0 < x < 3 \\ 0, & c.c. \end{cases}$$

a) Qual o valor de k para que f seja de fato uma densidade?
b) Determine $P(X \leq 1)$.
c) Determine $P(1 \leq X \leq 2)$.

Exercício 5.5.2. *O tempo adequado de troca de um amortecedor de certa marca em automóveis, sujeitos a uso contínuo e severo, pode ser considerado como uma variável contínua X, medida em anos. Suponha que a função densidade é dada pela seguinte expressão:*

$$f(x) = \begin{cases} \frac{1}{6}x, & 0 \leq x \leq 2; \\ \frac{1}{12}, & 2 < x \leq 6; \\ 0, & c.c. \end{cases}$$

a) Verifique que a função acima é, de fato, uma densidade.
b) Qual é a probabilidade de um automóvel, sujeito às condições descritas acima, necessitar de troca de amortecedores antes de 1 ano de uso? e entre 1 a 3 anos?

Exercício 5.5.3. *Suponha que a FDA da variável aleatória X seja:*

$$F(x) = \begin{cases} 0, & se\ x \leq 0; \\ 0,2x, & se\ 0 < x \leq 5; \\ 1, & se\ 5 \leq x \end{cases}$$

Obtenha:
a) $P(X < 1,8)$;
b) $P(X < -2)$;
c)$P(X > -1,5)$.

Exercício 5.5.4. *Seja a v.a. X com FDA. dada por:*

$$F(x) = \begin{cases} 0 & se\ x < -2; \\ 0,2x + 0,5 & se\ -2 \leq x < 2; \\ 1 & se\ 2 \leq x. \end{cases}$$

Obtenha:

a) $P(-1 < X < 1)$

b) $P(X < -2)$

d) $P(X > -1,5)$

Exercício 5.5.5. *Seja a v.a. X com f.d.p. dada por:*

$$f(x) = \begin{cases} x & se\ 0 \leq x < 1; \\ 2 - x & se\ 1 \leq x < 2; \\ 0 & se\ x \notin [0,2). \end{cases}$$

Obtenha a função de distribuição acumulada F(x) de X e calcule:

a) $P(0 < X < 5)$

b) $P(0 < X < 1)$

d) *E(X) e* $Var(X)$.

Exercício 5.5.6. *O tempo adequado de troca de um amortecedor de certa marca em automóveis, sujeitos a uso contínuo e severo, pode ser considerado como uma variável contínua X, medida em anos. Suponha que a função densidade é dada pela seguinte expressão:*

$$f(x) = \begin{cases} \frac{1}{6}x, & 0 \leq x \leq 2; \\ \frac{1}{12}, & 2 < x \leq 6; \\ 0, & c.c. \end{cases}$$

a) Verifique que a função acima é, de fato, uma densidade.
b) Obtenha a f.d.a. de X.
c) Qual é a probabilidade de um automóvel, sujeito às condições descritas acima, necessitar de troca de amortecedores antes de 1 ano de uso? e entre 1 a 3 anos?

Exercício 5.5.7. *Suponha que a espessura média de arruelas produzidas em uma fábrica tenha distribuição normal com média* $11,15mm$ *e desvio padrão* $2,238mm$. *Qual a porcentagem de arruelas que tem espessura entre* $8,70mm$ *e* $14,70mm$*?*

Exercício 5.5.8. *Suponha que o peso médio de* 800 *porcos de uma certa fazenda é de* $64kg$*, e o desvio padrão é de* $15kg$*. Supondo que este peso seja distribuído de forma normal, quantos porcos pesarão entre* $42kg$ *e* $73kg$*.*

Exercício 5.5.9. *A velocidade de transferência de um arquivo de um servidor da universidade para um computador pessoal na casa de um estudante, em uma noite de dia de semana, é distribuída normalmente, com média de 60kbits por segundo e um desvio padrão de 4kbits por segundo.*

(a) Qual é a probabilidade de o arquivo ser transferido a uma velocidade de 70kbits por segundo ou mais? (R: 0,0062)

(b) Qual é a probabilidade de o arquivo ser transferido a uma velocidade menor que 58kbits por segundo? (R: 0,3085)

Exercício 5.5.10.

Uma enchedora automática de refrigerantes está regulada para que o volume médio de líquido em cada garrafa seja de 1000 cm3 e desvio padrão de 10 m3. Admita que o volume siga uma distribuição normal.

(a) Qual é a porcentagem de garrafas em que o volume de líquido é menor que 990 cm3?

(b) Qual é a porcentagem de garrafas em que o volume de líquido não se desvia da média em mais do que dois desvios padrões?

(c) Se 10 garrafas são selecionadas ao acaso, qual é a probabilidade de que, no máximo, 4 tenham volume de líquido superior a 1002 cm^3*?*

Referências

Barbetta, P. A., Reis, M. M. & Bornia, A. C. (2004). *Estatística: para cursos de engenharia e informática*, volume 3. Atlas São Paulo.

Hazzan, S. (2013). *Fundamentos de matemática elementar, 5, Combinatória, Probabilidade*. Atual, São Paulo - SP.

Morettin, L. (2010). *Estatistica basica: probabilidade e inferência : volume único*. MAKRON.

Ross, S. (2009). *Probabilidade: Um Curso Moderno com Aplicações*. Bookman.

WILLIAM, J. S. (1981). Estatística aplicada à administração.