Universidade Federal do ABC

FELIPE FUGITA

AVALIAÇÃO EDUCACIONAL: UM OLHAR MATEMÁTICO

Santo André
2018

UNIVERSIDADE FEDERAL DO ABC

CENTRO DE MATEMÁTICA, COMPUTAÇÃO E COGNIÇÃO

FELIPE FUGITA

AVALIAÇÃO EDUCACIONAL: UM OLHAR MATEMÁTICO

Orientador: Prof. Dr. Daniel Miranda Machado

Dissertação de mestrado apresentada ao Centro de Matemática, Computação e Cognição para obtenção do título de Mestre em Matemática

SANTO ANDRÉ
2018

Sistema de Bibliotecas da Universidade Federal do ABC
Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC
com os dados fornecidos pelo(a) autor(a).

Fugita, Felipe
Avaliação Educacional : Um Olhar Matemático / Felipe Fugita. — 2017.

87 fls. : il.

Orientador: Daniel Miranda Machado

Dissertação (Mestrado) — Universidade Federal do ABC, Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Santo André, 2017.

1. Teoria de Resposta ao Item. 2. Avaliação Educacional. 3. Probabilidade. 4. Estatística. 5. Estimativa de Máxima Verossimilhança. I. Miranda Machado, Daniel. II. Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, 2017. III. Título.

MINISTÉRIO DA EDUCAÇÃO
Fundação Universidade Federal do ABC
Programa de Pós-Graduação em Mestrado Profissional em Matemática em Rede Nacional
Avenida dos Estados, 5001 – Bairro Santa Terezinha – Santo André – SP
CEP 09210-580 · Fone: (11) 4996-0017
profmat@ufabc.edu.br

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Felipe Fugita, realizada em 8 de dezembro de 2017:

Daniel M. Machado

Prof.(a) Dr.(a) **Daniel Miranda Machado** (Universidade Federal do ABC) – Presidente

Viviana Giampaoli

Prof.(a) Dr.(a) **Viviana Giampaoli** (Universidade de São Paulo) – Membro Titular

Coletti; Cristian Favio.

Prof.(a) Dr.(a) **Cristian Favio Coletti** (Universidade Federal do ABC) – Membro Titular

Prof.(a) Dr.(a) **Rafael de Mattos Grisi** (Universidade Federal do ABC) – Membro Suplente

Prof.(a) Dr.(a) **Renato Alessandro Martins** (Universidade Federal de São Paulo) – Membro Suplente

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do autor e com a anuência de seu orientador.

Santo André, 20 de Agosto de 2018.

Assinatura do autor: Felipe Fugita

Assinatura do orientador: Daniel M. Machado

Agradecimentos

À minha esposa e companheira, Flávia Renata Pereira de Almeida Fugita, pela motivação e apoio incondicional que foram fundamentais durantes esses anos para realização desse sonho.

Ao meu orientador, Prof. Dr. Daniel Miranda Machado, pela paciência, atenção, dedicação, reflexões e diálogos que contribuíram para o meu enriquecimento pessoal e profissional.

Ao corpo docente do Programa de Mestrado Profissional em Matemática da UFABC que realizam seu trabalho com amor e dedicação.

Resumo

Um dos objetivos desse trabalho é explicar a Teoria de Resposta ao Item, conhecida como TRI, enfatizando o modelo logístico de três parâmetros e descrevendo suas principais características. Outro objetivo é mostrar como o professor pode utilizar ferramentas estatísticas, em uma planilha eletrônica, para: verificar a qualidade das questões que compõe sua prova; analisar se existe uma correlação entre dois instrumentos de avaliação; utilizar a média escolar de um aluno para inferir sobre o seu desempenho no vestibular; entre outras possibilidades.

Com a finalidade de explicar a TRI e seu método de estimação de parâmetros por Máxima Verossimilhança, são apresentados previamente os modelos Matemáticos, Probabilísticos e Estatísticos, pilares dessa teoria. Além disso, é descrito como os programas de avaliações educacionais em larga escala de diversos países utilizam a TRI para monitorar o desempenho de seus sistemas educacionais. Em seguida, são expostas algumas ferramentas Estatísticas, em específico, o coeficiente de correlação, o método de mínimos quadrados e o ponto bisserial que podem colaborar nos processos de avaliações educacionais que fazem parte da rotina escolar. São ilustrados também exemplos de planilhas eletrônicas com a descrição passo a passo de sua construção e dos comandos utilizados.

Desse modo, espera-se contribuir para compreensão da TRI e, consequentemente, dos indicativos educacionais produzidos pelos programas de avaliações em larga escala, bem como, para atuação e reflexão da prática docente em seus métodos de avaliação educacional.

Palavras-chave: Teoria de Resposta ao Item, TRI, Avaliação Educacional, Probabilidade, Estatística, Avaliação educacional em larga escala, Correlação, Método de Mínimos Quadrados, Ponto Bisserial, Estimativa de Máxima Verossimilhança

Abstract

One of the goals of this work is to explain Item Response Theory, known as IRT, emphasizing the Three-Parameter Logistic model and describing its main characteristics. Another objective is to demonstrate how educators can use statistical tools within a spreadsheet to: verify the quality and reliability of test questions; examine whether there is a correlation between two assessment tools; use the school average of a student to predict his or her performance in entrance examinations; among other possibilities. To explain IRT and its method of parameter estimation by maximum likelihood, this work presents the mathematical, probabilistic and statistical models that are the pillars of the theory. It also describes how the large-scale educational assessment programs of various countries use IRT to monitor the performance of their education systems. Then, this work presents a selection of statistical tools, specifically, the correlation coefficient, the least squares method and the point biserial correlation, which could contribute to the process of routine educational assessments. Also provided are illustrated examples of spreadsheets with step-by- step descriptions of their creation and the commands used. Thus, the work hopes to contribute to the understanding of IRT and, consequently, of the educational indicators produced by large-scale assessment programs, as well as benefit educators in their practice and reflection on methods of educational evaluation.

Palavras-chave: Item Response Theory, IRT, Educational Evaluation, Correlation, Least Square Method, Point-Biserial

Conteúdo

1 Introdução **1**

2 Probabilidade **7**

2.1 Experimentos Probabilísticos 8
2.2 Probabilidade e seus Axiomas 12
2.3 Probabilidade em Espaços Equiprováveis 15
2.4 Probabilidade Condicional e Independência 18
2.5 Variáveis Aleatórias 27
2.6 Valor Esperado 32

3 Estatística **37**

3.1 Estimação de parâmetros 41
3.2 Estimativas de Máxima Verossimilhança 45
3.3 Coeficiente de Correlação 47
3.4 Método dos Mínimos Quadrados 50

4 Teoria de Resposta ao Item **55**

4.1 Modelo Logístico de Três Parâmetros 58
4.1.1 Curva Característica do Item (CCI) 59
4.1.2 Acerto ao Acaso 60
4.1.3 Dificuldade do Item 61
4.1.4 Discriminação do Item 63
4.2 Escala de Proficiência 66
4.3 Estimação dos Parâmetros e da Proficiência 70

5 Uso de Estatística em Processos Avaliativos **75**

5.1 Correlação 75
5.2 Mínimos Quadrados 77
5.3 Ponto Bisserial 81

1 Introdução

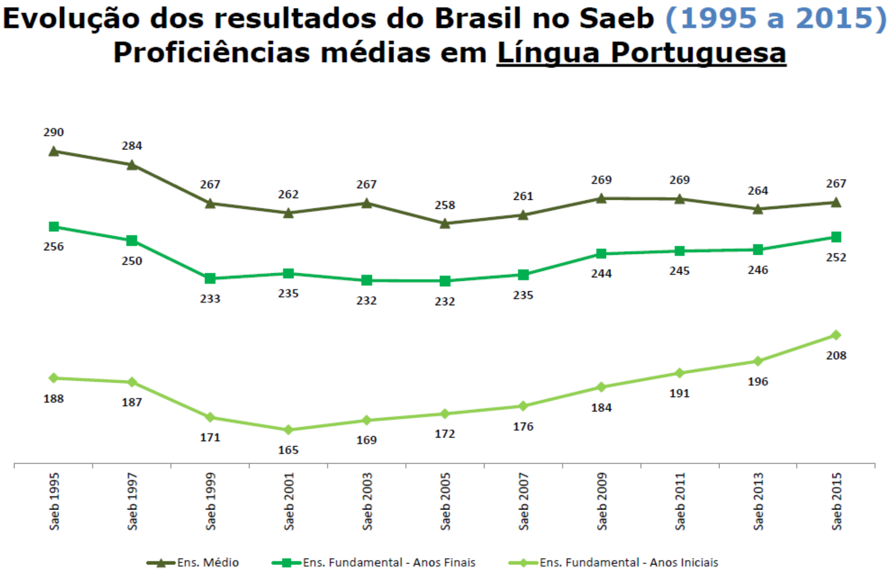
Os programas de avaliações educacionais em larga escala que são coordenados por órgãos externos às instituições escolares surgiram nas décadas de 1990 e 2000. A finalidade principal desses programas é a produção de indicadores que contribuam para diagnosticar a qualidade da educação, bem como, a obtenção de informações que sirvam para subsidiar a formulação, reformulação e o monitoramento de políticas educacionais.

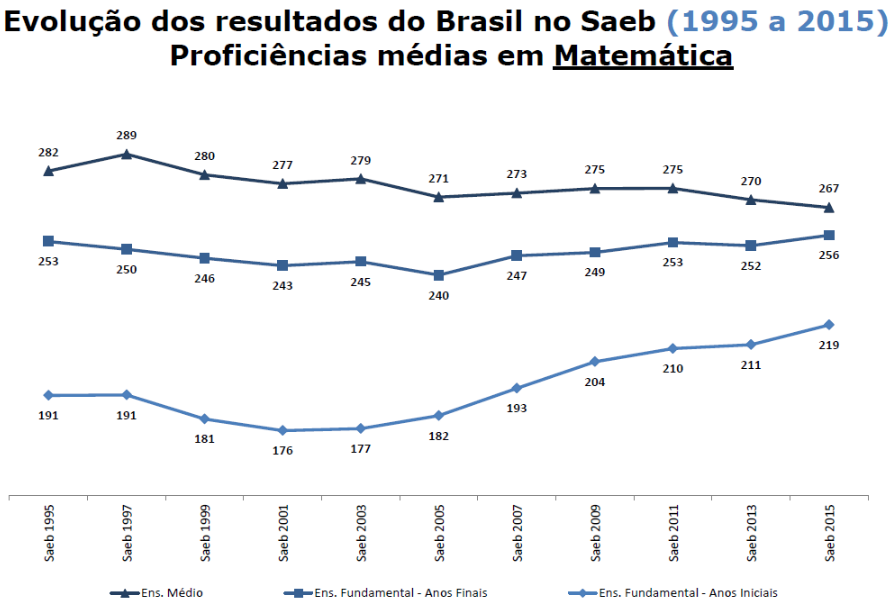
O PISA, Programme for International Student Assessment, é um programa internacional de avaliação educacional em larga escala, aplicado a cada três anos a estudantes na faixa dos 15 anos, idade em que se pressupõe o término da escolaridade básica obrigatória na maioria dos países. Esse programa é desenvolvido e coordenado internacionalmente pela Organização para Cooperação e Desenvolvimento Econômico (OCDE). Em 2015, mais de meio milhão de estudantes, que representam 28 milhões de jovens de 15 anos em 72 países e economias, fizeram o teste de duas horas e foram avaliados em ciências, matemática e leitura.

O SAEB, Sistema de Avaliação da Educação Básica, implantado no Brasil em 1990, é um conjunto de três avaliações em larga escala: Avaliação Nacional da Educação Básica (ANEB), Avaliação Nacional do Rendimento Escolar (conhecida como Prova Brasil) e a Avaliação Nacional da Alfabetização (ANA). O seu objetivo central é avaliar a qualidade da educação básica brasileira. Os gráficos apresentados a seguir foram extraídos do relatório de 2015 disponibilizado no endereço <http://download.inep.gov.br/educacao_ basica/saeb/aneb_ anresc/resultados/resumo_ dos_ resultados_ saeb_ 2015.pdf> e acessado em 17/10/2017. Eles mostram a evolução dos resultados dos estudantes brasileiros no SAEB de 1995 a 2015 em relação às proficiências médias em Língua Portuguesa e Matemática .

A simples análise gráfica dos resultados no SAEB de 1995 a 2015 fornece algumas pistas acerca da evolução do desempenho médio dos estudantes brasileiros em Língua Portuguesa e Matemática. Com exceção do Ensino Fundamental I, que apresentou um pequeno progresso em relação à qualidade de ensino em ambas as disciplinas, tanto o Ensino Fundamental II quanto o Ensino Médio demonstraram uma estagnação ou até regrediram ao longo das duas décadas consideradas.

O ENEM, Exame Nacional do Ensino Médio criado em 1998, nasceu com a finalidade de avaliar o desempenho escolar e acadêmico dos alunos brasileiros ao final do Ensino Médio. No entanto, com as transformações pelas quais passou, ganhou a posição de maior exame de seleção de ingressantes para o Ensino Superior. Em 2004, o Ministério da Educação implantou o Programa Universidade para Todos (ProUni) e vinculou a concessão de bolsas aos





resultados do programa e, em 2009, o ENEM se tornou o principal mecanismo de ingresso para as Universidades Federais do país. Desde então, esse programa de avaliação educacional em larga escala ganhou notoriedade, despertando o interesse dos profissionais de educação e da sociedade. Diferente dos vestibulares tradicionais, no qual a nota é o número de questões respondida corretamente, nas avaliações em larga escala, a nota depende de outras variáveis e é obtida por meio da Teoria de Resposta ao Item, conhecida como TRI. A dificuldade em explicar como dois candidatos com o mesmo número acertos podem ter resultados distintos foi uma das características relevante do ENEM que repercutiu nos meios de comunicação.

A Teoria de Resposta ao Item surgiu entre os anos de 1950 e 1960. Essa teoria é um conjunto de modelos Matemáticos, probabilísticos e estatísticos, que possibilita a criação de escalas para comparação e mensuração de traços latentes. Traços latentes são características implícitas do ser humano como, por exemplo, o nível de conhecimento em uma disciplina escolar ou o domínio de um determinado idioma. Nas décadas posteriores, a TRI foi implantada em programas de avaliações educacionais de larga escala por diversos países como, por exemplo, os Estados Unidos da América, a Holanda, a Coreia do Sul e os países participantes do PISA. No Brasil, o Exame Nacional do Ensino Médio (ENEM), o Sistema de Avaliação da Educação Básica (SAEB), o Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP), entre outros, também começaram usufruir dessa teoria.

Pela TRI, a aprendizagem e o desenvolvimento cognitivo em uma determinada área do conhecimento são mensurados por meio de uma régua denominada escala de proficiência. A escala de proficiência possibilita monitorar a evolução das habilidades desenvolvidas ao longo da trajetória escolar. Com isso, é possível acompanhar o progresso de um sistema educacional estabelecendo comparações entre grupos de alunos submetidos a provas diferentes e entre alunos em anos escolares distintos.

Por outro lado, nas escolas brasileiras, a avaliação educacional é tarefa do professor. Além da tradicional prova, trabalhos em grupos ou individuais, apresentação de seminários, exercícios avaliativos e outros dispositivos podem ser utilizados na execução dessa tarefa. Geralmente, de acordo com critérios predefinidos, o professor corrige e atribui notas aos instrumentos de verificação da aprendizagem para, no final, calcular a média escolar. Desse modo, todo esse processo de avaliação educacional é resumido e representado por um número, ou conceito, chamado de média escolar.

A média escolar é o indicativo fundamental para uma aprovação ou reprovação. Por isso, ela deveria fornecer informações sobre como está ocorrendo a aprendizagem dos estudantes em relação:

- à aquisição de saberes da disciplina;
- ao desenvolvimento de raciocínios;
- ao domínio de certas estratégias para resolução de problemas;
- à capacidade crítica sobre respostas obtidas;
- à interpretação e produção de diferentes gêneros textuais;
- à clareza na exposição de ideias e
- à capacidade de fazer análises e questionamentos.

Para que isso fosse possível, o professor teria que corrigir cada instrumento de avaliação extraindo a maior quantidade de dados que possibilitariam acompanhar e mapear o desenvolvimento dos processos cognitivos. Nesse sentido, seria importante conhecer as características e as limitações desses instrumentos, bem como, as relações que existem entre eles e as suas contribuições para a composição da média escolar.

De todos os dispositivos de verificação da aprendizagem, sem dúvida, a prova é o mais significativo. No entanto, elaborar e corrigir uma prova não é uma atividade simples. Um dos principais problemas é garantir que o resultado final seja fruto apenas do conhecimento do aluno. Para contextualizar melhor esse problema, vamos analisar duas situações.

Primeira situação. Imagine que um aluno resolveu, em um único dia, três provas diferentes, teoricamente similares, sobre o mesmo conteúdo e obteve os seguintes resultados:

- Prova 1 – nota 10,0 (dez)
- Prova 2 – nota 2,0 (dois)
- Prova 3 – nota 5,0 (cinco)
- *Observação:* as três provas continham dez questões, cada uma valia 1,0 ponto e a nota de cada prova foi a soma das pontuações obtidas em cada questão.

Se as provas são teoricamente similares, sobre o mesmo conteúdo e foram resolvidas no mesmo dia, como explicar esses resultados discrepantes? A explicação mais provável é que o grau de dificuldade de cada prova não foi o mesmo, logo o conhecimento do aluno não foi o único responsável por essas notas. Em outras palavras, o grau de dificuldade das avaliações interferiram nos resultados finais.

Segunda situação. Em uma prova com 10 questões, cada uma valendo 1,0 ponto cuja nota é a soma das pontuações de cada questão, imagine que o aluno A e o aluno B tiraram notas iguais a 6,0, porém pontuações diferentes em cada questão. Nesse caso, os alunos A e B possuem o mesmo nível de conhecimento? Uma prova geralmente contém questões com diferentes graus de dificuldade, por isso notas iguais não são suficientes para afirmar que os alunos possuem o mesmo nível de conhecimento. Nesse contexto, a nota depende dos critérios de correção e não apenas do conhecimento dos alunos.

As descrições e reflexões sobre os processos avaliativos externos e internos apresentadas até aqui motivaram o desenvolvimento desse trabalho que nasceu com o objetivo de explicar a Teoria de Resposta ao Item, enfatizando o modelo logístico unidimensional de três parâmetros e suas características. Para isso, foi necessário escrever o primeiro capítulo sobre Probabilidade e o segundo capítulo sobre Estatística, que são os modelos matemáticos pilares da TRI. Esses capítulos contém, além de definições e teoremas, problemas e experimentos que são modelados e estudados por essas teorias. O terceiro capítulo apresenta as ideias centrais da TRI e descreve como a Probabilidade e a Estatística auxiliam na construção de escalas de proficiências que possibilitam a mensuração de traços latentes.

Durante a execução dessa dissertação, a comparação entre os indicadores educacionais proporcionados pela TRI e os indicadores produzidos internamente nas instituições escolares levou a seguinte indagação: como o professor pode produzir indicadores educacionais confiáveis que retratam a aprendizagem e o desenvolvimento cognitivo dos seus alunos?

Para contribuir com a prática docente nesse aspecto, foi produzido o quarto e último capítulo, Uso de Estatística em Processos Avaliativos, cujo objetivo é mostrar como algumas ferramentas Estatísticas podem ser aplicadas, em uma planilha eletrônica, para analisar os dados produzidos pelos diversos instrumentos de avaliação utilizados pelo professor na sua rotina escolar.

Espera-se, assim, contribuir para a atuação dos profissionais de educação, para compreensão dos indicativos produzidos pelos programas de avaliações educacionais em larga escala e, consequentemente, para reflexão sobre os métodos de avaliação educacional.

2 Probabilidade

Atualmente, modelos probabilísticos são utilizados em diversas áreas da ciência como Medicina, Biologia e Física. Mas, o que motivou o desenvolvimento dessa teoria são problemas relacionados aos jogos de azar, jogos em que ganhar ou perder dependente unicamente da sorte e não da habilidade dos competidores.

Segundo Howard Eves [**eves2004**]:

"Embora os filósofos gregos da Antiguidade discutissem necessidade longa e detalhadamente, talvez seja correto dizer que não houve nenhum tratamento matemático da probabilidade até por volta do final do século XV e início do século XVI, quando alguns matemáticos italianos tentaram avaliar as possibilidades em alguns jogos de azar, como o de dados. Cardano escreveu um breve manual do jogador que envolvia alguns aspectos da probabilidade matemática. Mas em geral se concorda que a questão à qual está ligada a origem da ciência da probabilidade é o *problema dos pontos*. Esse problema pede que se determine a divisão das apostas de um jogo de azar interrompido, entre dois jogadores igualmente hábeis, supondo-se conhecida a contagem no momento da interrupção e o número necessário de pontos para se ganhar o jogo."

A Mega-Sena é um jogo de azar muito popular hoje no Brasil administrado pelo Governo Federal. Para apostar, é preciso escolher no mínimo seis e no máximo quinze números entre os inteiros de 1 a 60, chamados de dezenas. Sorteiam-se seis dezenas distintas e as apostas que acertaram quatro, cinco ou todas as dezenas sorteadas são premiadas financeiramente. Como é impossível prever o resultado desse jogo, é comum ouvirmos a indagação: qual é a chance de ganhar na loteria?

Outras indagações frequentes nas quais avaliam-se as chances de eventos futuros ocorrerem são: será que vai chover? Qual é a probabilidade do meu time ganhar? Será que ficarei parado no trânsito? Além dessas, existem muitas outras.

A probabilidade é um ramo da Matemática que pesquisa e desenvolve modelos que possibilitam quantificar as chances de eventos futuros acontecerem. Neste capítulo, apresentaremos como determinados tipos de fenômenos e experimentos de natureza aleatória são estudados por meio de um modelo probabilístico.

2.1 Experimentos Probabilísticos

Existem diversos tipos de fenômenos e experimentos que são estudados por meio de modelos Matemáticos. Um modelo matemático pode ser determinísticos ou probabilísticos. A função quadrática, a função exponencial e a função linear são exemplos de modelos determinísticos que descrevem, respectivamente, o movimento uniformemente variado, o decaimento radioativo e a dilatação térmica.

Os modelos determinísticos são utilizados para estudar fenômenos e experimentos de natureza não aleatória. Isto é, fenômenos e experimentos que sempre apresentam resultados iguais quando são executados sob as mesmas condições. A seguir, três fenômenos de natureza não aleatória.

- Movimento uniformemente variado: medir o espaço percorrido por um corpo, em queda livre no vácuo, 10 segundos após o início da sua queda.
- Decaimento radioativo: determinar em que instante uma partícula sofrerá decaimento.
- Dilatação térmica: determinar o volume ocupado por um gás a 10^oC e a 80^oC.

Os modelos probabilísticos são utilizados para descrever fenômenos e experimentos de natureza aleatória. Um fenômeno ou experimento é chamado de aleatório quando, repetido sob as mesmas condições, pode apresentar resultados diferentes. A seguir, quatro exemplos de experimentos aleatórios:

- Jogar uma moeda em uma superfície plana e anotar a face voltada para cima;
- Lançar um dado cúbico e registrar a face voltada para cima;
- Retirar ao acaso uma carta de um baralho;
- Sortear uma letra do alfabeto;

Observe que podemos lançar uma moeda de modo idêntico inúmeras vezes, mas não é possível garantir que a face voltada para cima será sempre a mesma a cada lançamento. Podemos apenas dizer que os resultados possíveis são cara ou coroa. Os demais exemplos também apresentam essas mesmas características: podem ser realizados indefinidamente de modo idêntico e, a cada repetição, pode apresentar um resultado diferente, por isso, podemos apenas descrever um conjunto de todos os resultados possíveis.

Definição 2.1. *Um **espaço amostral** Ω é um conjunto de todos os resultados possíveis de um experimento aleatório.*

Exemplos 2.2 O espaço amostral de cada experimento será indicado por Ω_i.

Vejamos alguns exemplos de espaços amostrais

1) Observar a face voltada para cima após o lançamento de um moeda.

$$\Omega_1 = \{\text{cara; coroa}\}$$

2) Anotar a quantidade de caras obtidas ao jogar uma moeda seis vezes.

$$\Omega_2 = \{0, 1, 2, 3, 4, 5, 6\}$$

3) Lançar e cronometrar o tempo que um avião de papel fica voando.

$$\Omega_3 = \{t; t \in \mathbb{R}^+\}$$

4) Retirar uma carta ao acaso de um baralho atual que contém 52 cartas, divididas em 4 naipes: copas, espadas, ouros e paus. As cartas de cada naipe são: A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3 e 2.

$$\Omega_4 = \{(x, y); x = A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2; y = \text{copas, espadas, ouros, paus}\}$$

5) Registrar os números das faces voltadas para cima após o laçamento simultâneo de dois dados cúbicos.

$$\Omega_5 = \{(x; y); x, y = 1, 2, 3, 4, 5, 6\}$$

6) Adicionar os números das faces voltadas para cima após o laçamento simultâneo de dois dados cúbicos.

$$\Omega_6 = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Nos dois últimos exemplos, o experimento é o mesmo, lançar dois dados cúbicos simultaneamente, no entanto os resultados observados são diferentes. Enquanto o espaço amostral Ω_5 descreve os possíveis pares de números que poderão sair, Ω_6 descreve as somas possíveis desses pares. É importante ter clareza do que se deseja analisar ao descrever um espaço amostral associado a um experimento aleatório.

Os resultados de um experimento não são necessariamente números. No primeiro exemplo, os resultados possíveis são as faces cara e coroa de uma moeda.

O espaço amostral pode ser um conjunto finito ou infinito. Com exceção do experimento lançar e cronometrar o tempo de voo de um avião de papel, os outros exemplos apresentam espaços amostrais finitos, isto é, são conjuntos que podem ser escritos na forma $\Omega = \{a_1, a_2, a_3, ..., a_k\}$.

Neste capítulo, descreveremos apenas um modelo probabilístico para espaços amostrais finitos.

Definição 2.3. *Um **evento** E é um subconjunto de um espaço amostral Ω. Os subconjuntos com exatamente um elemento são chamados de **eventos elementares**.*

Exemplos 2.4 Em cada experimento, Ω_i representa o espaço amostral e E_i o evento considerado.

1) Sair um número par no laçamento de um dado cúbico.

$$\Omega_1 = \{1; 2; 3; 4; 5; 6\}$$

$$E_1 = \{2; 4; 6\}$$

2) Retirar uma vogal ao sortear uma letra do alfabeto.

$$\Omega_2 = \{\text{letras do alfabeto}\}$$

$$E_2 = \{a; e; i; o; u\}$$

3) Retirar uma carta de paus ao acaso de um baralho atual que contém 52 cartas, divididas em 4 naipes: copas, espadas, ouros e paus. As cartas de cada naipe são: A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3 e 2.

$$\Omega_3 = \{(x, y); x = A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2; y = \text{copas, espadas, ouros, paus}\}$$

$$E_3 = \{(x; \text{paus}); x = A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2\}$$

4) Obter cara após o lançamento de um moeda.

$$\Omega_4 = \{\text{cara; coroa}\}$$

$$E_4 = \{\text{cara}\}$$

5) Sortear a bola de número 5 de urna que contém dez bolas idênticas numeradas de 1 a 10.

$$\Omega_5 = \{1; 2; 3; 4; 5; 6; 7; 8; 9; 10\}$$

$$E_5 = \{5\}$$

O subconjuntos E_4 e E_5 são exemplos de eventos elementares, pois eles contém um único elemento.

Podemos definir operações com os eventos de um espaço amostral. Veja a seguir duas dessas operações.

Definição 2.5. *Definimos a* ***união*** *de dois eventos* A *e* B*, que denotaremos por* $A \cup B$*, como o evento formado pelos elementos que pertencem a pelo menos um dos conjuntos* A *ou* B.

Exemplo 2.6 Retira-se ao acaso uma bola de uma urna que contém dez bolas numeradas de 1 a 10. Se A é o evento sair um número ímpar e B é o evento sair um número primo. Descreva o evento $A \cup B$. ◂

Solução: O experimento possui o seguinte espaço amostral.

$$\Omega = \{1; 2; 3; 4; 5; 6; 7; 8; 9; 10\}$$

Os eventos A e B são:

$$A = \{1; 3; 5; 7; 9\}$$

$$B = \{2; 3; 5; 7; \}$$

Portanto, o evento $A \cup B$ é:

$$A \cup B = \{1; 2; 3; 5; 7; 9\}$$

□

Observe que todos os números contidos em $A \cup B$ pertencem a um dos eventos, A ou B.

Definição 2.7. *Definimos o* ***complementar de um evento*** *ou ainda o evento complementar de A, denotado A^C, como o conjunto dos elementos de Ω que não estão em A.*

$$A^C = \{x : x \in \Omega, x \notin A\}$$

Exemplo 2.8 No exemplo anterior, retira-se ao acaso uma bola de uma urna que contém dez bolas numeradas de 1 a 10, A é o evento sair um número ímpar. Qual é o evento complementar de A? ◂

Solução: O espaço amostral Ω e o evento A desse experimento são:

$$\Omega = \{1; 2; 3; 4; 5; 6; 7; 8; 9; 10\}$$

$$A = \{1; 3; 5; 7; 9\}$$

Logo, o evento complementar de A é sair um número par, ou seja:

$$A^C = \{2; 4; 6; 8; 10\}$$

□

Note que o complementar de A são os elementos do espaço amostral Ω que não pertencem ao evento A.

2.2 Probabilidade e seus Axiomas

O que caracteriza um experimento aleatório é não saber qual resultado particular ocorrerá quando ele for realizado, como descrito em experimentos probabilísticos, um fenômeno ou experimento é aleatório quando, repetido sob as mesmas condições, pode apresentar resultados diferentes.

Suponha que A seja um evento de um espaço amostral associado a um experimento aleatório, logo não podemos afirmar se A acontecerá ou não ao realizar esse experimento. Por isso, buscamos um número que indicará de algum modo quão verossímil é a chance de ocorrência do evento A. Buscar esse número nos leva à teoria da probabilidade.

O experimento aleatório a seguir tem a finalidade de ilustrar quais são as principais propriedades que o número que expressa a chance de um evento deve ter.

Imagine que uma pessoa lançou um dado cúbico 1.000 vezes e anotou.

Face que caiu voltada para cima	Número de ocorrências
1	156
2	174
3	167
4	196
5	163
6	144

Em seguida, concluiu que a chance de ocorrência de cada face pode ser expressa pela razão entre o número de ocorrência de cada face e o número de vezes que o dado foi lançado, ou seja:

Face do dado	Chance de ocorrência
1	$\frac{156}{1.000}$
2	$\frac{174}{1.000}$
3	$\frac{167}{1.000}$
4	$\frac{196}{1.000}$
5	$\frac{163}{1.000}$
6	$\frac{144}{1.000}$

Observe que as razões que indicam as chances de ocorrência de cada face desse dado apresentam as seguintes propriedades.

- Sempre serão números maiores ou iguais a 0 e menores ou iguais a 1.
- A soma de todas elas resulta em 1.

$$\frac{156}{1.000} + \frac{174}{1.000} + \frac{167}{1.000} + \frac{196}{1.000} + \frac{163}{1.000} + \frac{144}{1.000} = 1$$

- A chance de ocorrência da face 1 ou 2, por exemplo, é dado pela soma de suas razões. Portanto, a chance de ocorrência da face 1 ou 2 é:

$$\frac{156}{1.000} + \frac{174}{1.000} = \frac{330}{1.000}$$

Essas propriedades são uma das motivações da definição de espaço de probabilidade que descreve formalmente quais condições o número atribuído aos eventos de um espaço amostral deve satisfazer.

Definição 2.9. *Um **espaço de probabilidade** é um espaço amostral Ω munido com uma função $\mathbb{P}$ que atribui um valor entre 0 e 1 (que denominaremos probabilidade) aos eventos E e satisfazendo os seguintes axiomas:*

Axioma 1. $0 \leqslant \mathbb{P}(E) \leqslant 1$.

Axioma 2. $\mathbb{P}(\Omega) = 1$.

Axioma 3. *Se $E_1, E_2, \ldots E_n, \ldots$ forem eventos mutuamente excludentes dois a dois, i.e., $E_i \cap E_j = \varnothing$ para todo i e j, então:*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(E_i).$$

Propriedade 2.10. *Se $\varnothing$ é conjunto vazio (que denominaremos de evento impossível), então $\mathbb{P}(\varnothing) = 0$.*

Demonstração. Para qualquer evento A, temos que:

$$A = A \cup \varnothing$$

Como A e $\varnothing$ são mutuamente excludentes, decorre do axioma 3:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(A \cup \varnothing) \\ &= \mathbb{P}(A) + \mathbb{P}(\varnothing) \end{aligned}$$

Logo $\mathbb{P}(\varnothing) = 0$ □

Propriedade 2.11. *Se A^C é o evento complementar de A, então:*

$$\mathbb{P}\left(A^C\right) = 1 - \mathbb{P}(A)$$

Demonstração. Para qualquer evento A e o seu complementar A^C do espaço amostral Ω, temos que:

$$\Omega = A \cup A^C$$

Decorrem dos axiomas 2 e 3:

$$\begin{aligned}
\mathbb{P}\left(A \cup A^C\right) &= 1 \\
\mathbb{P}(A) + \mathbb{P}\left(A^C\right) &= 1 \\
\mathbb{P}\left(A^C\right) &= 1 - \mathbb{P}(A)
\end{aligned}$$

□

Propriedade 2.12. *Se* A *e* B *são dois eventos quaisquer, então:*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Demonstração. Podemos decompor os conjuntos $A \cup B$ e B em eventos mutuamente excludentes do seguinte modo:

$$A \cup B = A \cup (B \cap A^C)$$

$$B = (A \cap B) \cup (B \cap A^C)$$

Pelo axioma 3:

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}\left(B \cap A^C\right)$$

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}\left(B \cap A^C\right)$$

Subtraindo a segunda igualdade da primeira:

$$\mathbb{P}(A \cup B) - \mathbb{P}(B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

□

Propriedade 2.13. *Se* A *e* B *são dois eventos quaisquer e* $A \subset B$*, então:*

$$\mathbb{P}(A) \leqslant \mathbb{P}(B)$$

Como $A \subset B$, podemos decompor o conjunto B em eventos mutuamente excludentes do seguinte modo:

$$B = A \cup (B \cap A^C)$$

Pelo axioma 3:

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}\left(B \cap A^C\right)$$

Pelo axioma 1:

$$0 \leqslant \mathbb{P}\left(B \cap A^C\right) \leqslant 1$$

Portanto,

$$\mathbb{P}(A) \leqslant \mathbb{P}(B)$$

2.3 Probabilidade em Espaços Equiprováveis

Na seção anterior Probabilidades e seus Axiomas foi descrito um modo empírico de atribuir probabilidades aos eventos de um espaço amostral no qual uma pessoa lançou um dado cúbico 1.000 vezes e concluiu que a razão entre o número de ocorrência de cada face e o número de vezes que o dado foi lançado expressa a probabilidade de ocorrência de cada face. No entanto, esse modo empírico traz indagações como: quantas vezes tenho que repetir o experimento? Como garantir que o número encontrado dependa apenas da natureza do experimento aleatório e não de outros fatores externos? Além disso, como obter probabilidades sem recorrer a experimentação?

Para determinar as probabilidades $\mathbb{P}(E_i)$ de cada evento elementar E_i de um espaço amostral finito é necessário considerar, ou avaliar, alguma hipótese sobre os resultados possíveis do experimento.

Exemplo 2.14 Suponha que um experimento tenha apenas dois resultados possíveis, E_1 e E_2, e que a chance de E_1 ocorrer é 4 vezes maior que a chance de E_2, isto é:

$$\mathbb{P}(E_1) = 4 \cdot \mathbb{P}(E_2)$$

Calcule as probabilidades $\mathbb{P}(E_1)$ e $\mathbb{P}(E_2)$. ◂

Solução: Pelos axiomas 2 e 3, temos que:

$$E_1 \cup E_2 = \Omega$$

$$\mathbb{P}(E_1) + \mathbb{P}(E_2) = 1$$

Pelo enunciado, $\mathbb{P}(E_1) = 4 \cdot \mathbb{P}(E_2)$. Logo:

$$4 \cdot \mathbb{P}(E_2) + \mathbb{P}(E_2) = 1$$

Desse modo, as probabilidades são:

$$\mathbb{P}(E_2) = \frac{1}{5}$$

$$\mathbb{P}(E_1) = \frac{4}{5}$$

□

Note que os valores calculados, $\mathbb{P}(E_1) = \frac{4}{5}$ e $\mathbb{P}(E_2) = \frac{1}{5}$, são consequência da suposição de que a chance de E_1 ocorrer é 4 vezes maior que a chance de E_2.

A hipótese geralmente adotada para muitos experimentos com espaços amostrais finitos é a de que todos os resultados possíveis têm chances iguais de ocorrer. Por isso, definimos:

Definição 2.15. *Um espaço amostral Ω é **equiprovável** quando admitimos que os seus eventos elementares, $E_1, E_2, E_3, ..., E_n$, têm probabilidades iguais, ou seja:*

$$\mathbb{P}(E_1) = \mathbb{P}(E_2) = \mathbb{P}(E_3) = ... = \mathbb{P}(E_n)$$

Propriedade 2.16. *Se o espaço amostral equiprovável* $\Omega = \{E_1, E_2, E_3, ..., E_n\}$ *tem* n *eventos elementares, então a probabilidade:*

a) $\mathbb{P}(E_i)$ *de cada evento elementar é* $\frac{1}{n}$.

$$\mathbb{P}(E_i) = \frac{1}{n}$$

b) $\mathbb{P}(E)$ *de um evento* $E = \{E_1, E_2, E_3, ..., E_k\}$ *que contém* k *elementos é* $\frac{k}{n}$.

$$\mathbb{P}(E) = \frac{k}{n}$$

a) Demonstração.

$$E_1 + E_2 + E_3 + ... + E_n = \Omega$$
$$\mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(E_3) + ... + \mathbb{P}(E_n) = 1$$
$$\mathbb{P}(E_1) = \mathbb{P}(E_2) = \mathbb{P}(E_3) = ... = \mathbb{P}(E_n) = \mathbb{P}(E_i)$$
$$n \cdot \mathbb{P}(E_i) = 1$$
$$\mathbb{P}(E_i) = \frac{1}{n}$$

□

b) Demonstração.

$$\mathbb{P}(E) = \mathbb{P}(E_1) + \mathbb{P}(E_2) + \mathbb{P}(E_3) + ... + \mathbb{P}(E_k)$$
$$\mathbb{P}(E_1) = \mathbb{P}(E_2) = \mathbb{P}(E_3) = ... = \mathbb{P}(E_k) = \mathbb{P}(A)$$
$$\mathbb{P}(E) = k \cdot \mathbb{P}(A)$$
$$\mathbb{P}(E) = k \cdot \frac{1}{n}$$
$$\mathbb{P}(E) = \frac{k}{n}$$

□

Resumindo, em um espaço amostral equiprovável, a probabilidade de um evento E é a razão entre o número (k) de casos favoráveis ao evento e o número total (n) de casos possíveis.

$$\mathbb{P}(E) = \frac{\text{número de casos favoráveis ao evento (k)}}{\text{número total de casos possíveis (n)}}$$

Exemplo 2.17 Qual é a probabilidade de ocorrer cara ao lançar uma moeda balanceada? ◂

Solução: Há 1 caso favorável, $E_1 = \{\text{cara}\}$, em 2 resultados possíveis, $\Omega_2 = \{\text{cara, coroa}\}$.

Portanto, a probabilidade é:

$$\mathbb{P}(E_1) = \frac{1}{2}$$

□

Exemplo 2.18 Qual é a probabilidade de sair um número menor do que 3 ao lançar um dado cúbico simétrico? ◂

Solução: Existem 2 casos favoráveis, $E_2 = \{1, 2\}$, em 6 resultados possíveis, $\Omega_2 = \{1, 2, 3, 4, 5, 6\}$. Logo, a probabilidade é:

$$\mathbb{P}(E_2) = \frac{2}{6} = \frac{1}{3}$$

□

Admitir que o espaço amostral é equiprovável facilita calcular as probabilidades dos eventos. No entanto, essa hipótese deve ser justificada, assegurada. Nos exemplos, a moeda ser balanceada e o dado simétrico justificam a escolha de que os resultados possíveis têm chances iguais de ocorrer.

Além das características dos experimentos e dos fenômenos, é necessário analisar se os eventos elementares são igualmente prováveis. Suponha que um casal, ao planejar ter dois filhos, quer saber a probabilidade de nascer duas meninas, o espaço amostral $\Omega = \{0, 1, 2\}$ que descreve a quantidade de meninas que pode nascer não é equiprovável, a tabela a seguir mostra o pode acontecer em cada nascimento.

1º Nascimento	2º Nascimento
Menino	Menino
Menino	Menina
Menina	Menino
Menina	Menina

Dos quatro resultados possíveis, em apenas um nascem duas meninas, por isso os eventos elementares de $\Omega = \{0, 1, 2\}$ não são igualmente prováveis, logo é incorreto afirmar que a probabilidade de nascer 2 meninas é $\frac{1}{3}$. Nesse caso, para atribuir probabilidades, é necessário admitir que a chance do casal ter apenas uma menina é duas vezes maior do que ter duas ou nenhuma menina, isto é:

$$\mathbb{P}(1 \text{ menina}) = 2 \cdot \mathbb{P}(2 \text{ meninas}) = 2 \cdot \mathbb{P}(\text{Nenhuma menina})$$

Essa hipótese levaria a solução correta de que a $\mathbb{P}(2 \text{ meninas}) = \frac{1}{4}$.

Reduzir um problema a outro é uma maneira de obter espaços amostrais equiprováveis. Caso considerássemos o espaço amostral equiprovável $\Omega = \{(H, H), (H, M), (M, H), (M, M)\}$, onde H indica o nascimento de um menino e M o nascimento de uma menina, bastaria contar o número de casos favoráveis e o número total de casos para concluir que a probabilidade do casal ter 2 meninas é $\frac{1}{4}$.

2.4 Probabilidade Condicional e Independência

Sejam A e B dois eventos de um espaço amostral. Indicamos por $\mathbb{P}(A|B)$ a probabilidade condicional do evento A ocorrer, uma vez que o evento B já ocorreu e definimos:

Definição 2.19.

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \ \textit{com}\ \mathbb{P}(B) \neq 0$$

Vamos examinar a seguinte situação para compreender algumas ideias que motivaram essa definição.

Considere que uma indústria tem 65 funcionários distribuídos da seguinte forma.

Setor	Homens	Mulheres	Total
Vendas	7	18	25
Produção	23	17	40
Total	30	35	65

Considere também os seguintes eventos:

$$H = \{\text{escolher um funcionário homem ao acaso}\}$$

$$V = \{\text{escolher um funcionário do setor de vendas ao acaso}\}$$

Sabemos que a probabilidade do evento H ocorrer, sortear um homem, é:

$$\mathbb{P}(H) = \frac{30}{65} = \frac{6}{13}$$

Agora, imagine que o evento V aconteceu, ou seja, escolheram um funcionário do setor de vendas ao acaso. Pela tabela, sabemos que das 25 pessoas do setor de vendas, 7 são homens, logo a probabilidade do evento H ocorrer, sortear um homem, sabendo que o evento V aconteceu, foi sorteado um funcionário do setor de vendas, é:

$$\mathbb{P}(H|V) = \frac{7}{25}$$

Observe que 7 é o número de elementos do conjunto $H \cap V$, homens que trabalham em vendas, e 25 é número de elementos do conjunto V, total de funcionários do setor de vendas, ou seja:

$$\mathbb{P}(H|V) = \frac{7}{25} = \frac{n(H \cap V)}{n(V)}$$

Dividindo 7 e 25 pelo total de funcionários da empresa, $n(\Omega) = 65$, temos:

$$\mathbb{P}(H|V) = \frac{\frac{7}{65}}{\frac{25}{65}} = \frac{\frac{n(H \cap V)}{n(\Omega)}}{\frac{n(V)}{n(\Omega)}} = \frac{\mathbb{P}(H \cap V)}{\mathbb{P}(V)}$$

Cabe enfatizar que analisamos essa situação com o objetivo de apresentar algumas noções intuitivas que motivaram a definição de probabilidade condicional. Primeiro, calculamos $\mathbb{P}(H|V)$ consultando os dados da tabela, depois, verificamos que os dados utilizados foram $n(H \cap V)$ e $n(V)$ e, por último, dividimos $n(H \cap V)$ e $n(V)$ pelo total elementos do espaço amostral, $n(\Omega)$, para mostrar que $\mathbb{P}(H|V)$ é igual a razão entre as probabilidades $\mathbb{P}(H \cap V)$ e $\mathbb{P}(V)$.

Outro ponto importante foi que determinamos $\mathbb{P}(H|V) = \frac{7}{25}$ de um modo direto, sem utilizar a definição de probabilidade condicional. Isso é possível porque, ao saber que o evento V aconteceu, o espaço amostral é reduzido à V e o evento H é reduzido à $H \cap V$.

Vamos considerar agora, outros dois eventos M e P que poderiam ocorrer ao sortear um funcionário dessa indústria.

Setor	Homens	Mulheres	Total
Vendas	7	18	25
Produção	23	17	40
Total	30	35	65

$$M = \{\text{escolher uma funcionária mulher ao acaso}\}$$

$$P = \{\text{escolher um funcionário do setor de produção ao acaso}\}$$

Podemos calcular de duas maneiras $\mathbb{P}(M|P)$, a probabilidade condicional de sortear uma mulher, sabendo que o sorteado trabalha no setor de produção.

- Usando diretamente os dados da tabela, dos 40 funcionários que trabalham na produção, 17 são mulheres, logo $\mathbb{P}(M|P) = \frac{17}{40}$.
- Pela definição:

$$\mathbb{P}(M|P) = \frac{\mathbb{P}(M \cap P)}{\mathbb{P}(P)} = \frac{\frac{17}{65}}{\frac{40}{65}} = \frac{17}{40}$$

Da definição de probabilidade condicional obtemos uma consequência importante, conhecida como **fórmula da multiplicação de probabilidades**. Para isso, basta isolar $\mathbb{P}(A \cap B)$, veja:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \text{ com } \mathbb{P}(B) \neq 0$$

$$\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$$

A fórmula de multiplicação de probabilidade diz que a probabilidade de ocorrer simultaneamente os eventos A e B pode ser calculada multiplicando-se, nessa ordem, a probabilidade do evento B ocorrer pela probabilidade condicional de $(A|B)$ ocorrer.

Exemplo 2.20 Suponha que de grupo de 4 mulheres e 6 homens serão sorteados duas pessoas. Calcule a probabilidade de sortearem:

a) duas mulheres.

b) dois homens.

c) uma mulher e um homem nessa ordem.

d) um homem e uma mulher nessa ordem.

◂

Solução: Considere os eventos:

$$M = \{\text{sorteou uma mulher}\}$$

$$H = \{\text{sorteou um homem}\}$$

a) A probabilidade do primeiro sorteado ser uma mulher é:

$$\mathbb{P}(M) = \frac{4}{10}$$

A probabilidade do segundo sorteado também ser uma mulher, sabendo que o primeiro sorteado foi uma mulher, é:

$$\mathbb{P}(M|M) = \frac{3}{9}$$

Portanto, a probabilidade de sortearem duas mulheres é:

$$\mathbb{P}(M \cap M) = \mathbb{P}(M) \cdot \mathbb{P}(M|M) = \frac{4}{10} \cdot \frac{3}{9} = \frac{2}{15}$$

Repetindo o mesmo procedimento, calculamos as probabilidades dos demais itens.

b) A probabilidade de sortearem dois homens é:

$$\mathbb{P}(H \cap H) = \mathbb{P}(H) \cdot \mathbb{P}(H|H) = \frac{6}{10} \cdot \frac{5}{9} = \frac{1}{3}$$

c) A probabilidade do primeiro sorteado ser uma mulher e o segundo ser um homem é:

$$\mathbb{P}(M \cap H) = \mathbb{P}(M) \cdot \mathbb{P}(H|M) = \frac{4}{10} \cdot \frac{6}{9} = \frac{4}{15}$$

d) A probabilidade do primeiro sorteado ser um homem e o segundo ser uma mulher é:

$$\mathbb{P}(H \cap M) = \mathbb{P}(H) \cdot \mathbb{P}(M|H) = \frac{6}{10} \cdot \frac{4}{9} = \frac{4}{15}$$

□

A fórmula da multiplicação de probabilidades é generalizada para mais de dois eventos pela seguinte propriedade.

Propriedade 2.21. *Sejam* $A_1, A_2, A_3, ..., A_k$ *eventos do espaço amostral* Ω*, onde está definida a probabilidade* P*, então:*

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_k) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1, A_2) \cdot .... \cdot \mathbb{P}(A_k|A_1, A_2, A_3, ..., A_{k-1})$$

Demonstração. Faremos a demonstração por indução. Para $k = 2$, temos:

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1)$$

Suponha que a fórmula da multiplicação de probabilidades seja verdadeira para $k - 1$, então:

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1}) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1, A_2) \cdot .... \cdot \mathbb{P}(A_{k-1}|A_1, A_2, A_3, ..., A_{k-2})$$

Considerando os eventos A_k e $A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1}$, pela probabilidade condicional, temos:

$$\mathbb{P}(A_k|A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1}) = \frac{\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_k)}{\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1})}$$

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_k) = \mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1}) \cdot \mathbb{P}(A_k|A_1 \cap A_2 \cap A_3 \cap ... \cap A_{k-1})$$

$$\mathbb{P}(A_1 \cap A_2 \cap A_3 \cap ... \cap A_k) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2|A_1) \cdot \mathbb{P}(A_3|A_1, A_2) \cdot .... \cdot \mathbb{P}(A_k|A_1, A_2, A_3, ..., A_{k-1})$$

□

Exemplo 2.22 De um lote de 20 peças, 5 são defeituosas. Ao escolher aleatoriamente sem reposição 5 peças desse lote, qual é a probabilidade de todas serem defeituosas? ◂

Solução: Seja $\mathbb{P}(D)$ a probabilidade das cincos peças escolhidas serem defeituosas, logo:

$$\mathbb{P}(D) = \frac{5}{20} \cdot \frac{4}{19} \cdot \frac{3}{18} \cdot \frac{2}{17} \cdot \frac{1}{16} = \frac{1}{15504}$$

□

Definição 2.23. *Dizemos que os eventos* $B_1, B_2, B_3, ..., B_k$ *são uma partição do espaço amostral* Ω *quando:*

- $B_1, B_2, B_3, ..., B_k$ *forem mutuamente excludentes dois a dois, isto é,* $B_i \cap B_j = \varnothing$ *para todo* $i \neq j$*;*
- $B_1 \cup B_2 \cup B_3 \cup ... \cup B_k = \Omega$*;*
- $\mathbb{P}(B_i) > 0$ *para todo* i.

Exemplo 2.24 Seja Ω o espaço amostral associado a retirada ao acaso de uma bola de uma urna que contém dez bolas numeradas de 1 a 10, então:

$$\Omega = \{1; 2; 3; 4; 5; 6; 7; 8; 9; 10\}$$

Os eventos B_1, B_2, B_3 e B_4, descritos a seguir, são uma partição do espaço amostral Ω porque são mutuamente excludentes dois a dois, a união de B_1, B_2, B_3 e B_4 resulta em Ω e todos eles têm probabilidade maior do que zero de ocorrência.

$$B_1 = \{1; 2\}$$

$$B_2 = \{3; 4; 5; 6\}$$

$$B_3 = \{7\}$$

$$B_4 = \{8; 9; 10\}$$

◂

Considere, no exemplo anterior, o evento A sortear uma bola par dessa urna, isto é:

$$A = \{2; 4; 6; 8; 10\}$$

Mostraremos que a probabilidade $\mathbb{P}(A)$ de sortear uma bola par dessa urna é dada em função da partição B_1, B_2, B_3 e B_4 do espaço amostral Ω do seguinte modo:

$$\mathbb{P}(A) = \mathbb{P}(B_1) \cdot \mathbb{P}(A|B_1) + \mathbb{P}(B_2) \cdot \mathbb{P}(A|B_2) + \mathbb{P}(B_3) \cdot \mathbb{P}(A|B_3) + \mathbb{P}(B_4) \cdot \mathbb{P}(A|B_4)$$

Para isso, observe que as intersecções do evento A com cada evento da partição B_1, B_2, B_3 e B_4 são:

$$A \cap B_1 = \{2\}$$

$$A \cap B_2 = \{4; 6\}$$

$$A \cap B_3 = \varnothing$$

$$A \cap B_4 = \{8; 10\}$$

Desse modo, o conjunto A pode ser descrito pela união das intersecções de A com cada uma das partições B_1, B_2, B_3 e B_4, veja.

$$A = \{2\} \cup \{4; 6\} \cup \varnothing \cup \{8; 10\}$$

$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup (A \cap B_4)$$

Como os eventos $(A \cap B_1), (A \cap B_2), (A \cap B_3)$ e $(A \cap B_4)$ são mutuamente excludentes dois a dois, a probabilidade do evento A ocorrer é dada em função da partição B_1, B_2, B_3 e B_4 por:

$$\mathbb{P}(A) = \mathbb{P}(A \cap B_1) + \mathbb{P}(A \cap B_2) + \mathbb{P}(A \cap B_3) + \mathbb{P}(A \cap B_4)$$

Além disso, $\mathbb{P}(A \cap B_i) = \mathbb{P}(B_i) \cdot \mathbb{P}(A|B_i)$, logo:

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}(B_1) \cdot \mathbb{P}(A|B_1) + \mathbb{P}(B_2) \cdot \mathbb{P}(A|B_2) + \mathbb{P}(B_3) \cdot \mathbb{P}(A|B_3) + \mathbb{P}(B_4) \cdot \mathbb{P}(A|B_4) \\ &= \frac{2}{10} \cdot \frac{1}{2} + \frac{4}{10} \cdot \frac{2}{4} + \frac{1}{10} \cdot \frac{0}{1} + \frac{3}{10} \cdot \frac{2}{3} \\ &= \frac{1}{2} \end{aligned}$$

A generalização da fórmula acima é conhecida como fórmula das probabilidades totais.

Propriedade 2.25. ***Fórmula das probabilidades totais***. *Sejam A um evento e $B_1, B_2, B_3, ..., B_k$ uma partição do espaço amostral Ω, onde está definida a probabilidade P, então:*

$$\mathbb{P}(A) = \mathbb{P}(B_1) \cdot \mathbb{P}(A|B_1) + \mathbb{P}(B_2) \cdot \mathbb{P}(A|B_2) + \cdots + \mathbb{P}(B_k) \cdot \mathbb{P}(A|B_k)$$

Exemplo 2.26 Uma comerciante trabalha com três marcas de camisas A, B e C. Ele constatou que a cada 10 camisas compradas da marca:

- A, duas são defeituosas;
- B, uma é defeituosa;
- C, quatro são defeituosa.

Certo dia, recebeu um lote com 20 camisas da marca A, 10 camisas da marca B e 20 camisas da marca C. Qual é a probabilidade dele vender aleatoriamente uma camisa defeituosa desse lote? ◂

Solução: Sejam o espaço amostral Ω o lote de 50 camisas, D o evento vender uma camisa defeituosa de Ω e A, B e C uma partição de Ω, em que A, B e C representam a venda de uma camisa das marcas A, B e C, respectivamente. Então, pela fórmula das probabilidades totais:

$$\begin{aligned} \mathbb{P}(D) &= \mathbb{P}(A) \cdot \mathbb{P}(D|A) + \mathbb{P}(B) \cdot \mathbb{P}(D|B) + \mathbb{P}(C) \cdot \mathbb{P}(D|C) \\ &= \frac{20}{50} \cdot \frac{2}{10} + \frac{10}{50} \cdot \frac{1}{10} + \frac{20}{50} \cdot \frac{4}{10} \\ &= \frac{13}{50} \end{aligned}$$

A probabilidade dele vender uma camisa defeituosa é $\frac{13}{50}$. □

No exemplo anterior, imagine agora que o comerciante escolheu aleatoriamente uma camisa defeituosa, qual é a probabilidade dela ser da marca A? Para responder essa pergunta, devemos calcular $\mathbb{P}(A|D)$. Para isso, vamos utilizar a definição de probabilidade condicional, veja:

$$\mathbb{P}(A|D) = \frac{\mathbb{P}(A \cap D)}{\mathbb{P}(D)}, \text{com } \mathbb{P}(D) \neq 0$$

Pelas fórmulas de multiplicação de probabilidades e das probabilidades totais, temos que:

$$\mathbb{P}(A \cap D) = \mathbb{P}(A) \cdot \mathbb{P}(D|A)$$

$$\mathbb{P}(D) = \mathbb{P}(A) \cdot \mathbb{P}(D|A) + \mathbb{P}(B) \cdot \mathbb{P}(D|B) + \mathbb{P}(C) \cdot \mathbb{P}(D|C)$$

Substituindo na definição de probabilidade condicional, concluímos que:

$$\begin{aligned}
\mathbb{P}(A|D) &= \frac{\mathbb{P}(A) \cdot \mathbb{P}(D|A)}{\mathbb{P}(A) \cdot \mathbb{P}(D|A) + \mathbb{P}(B) \cdot \mathbb{P}(D|B) + \mathbb{P}(C) \cdot \mathbb{P}(D|C)} \\
&= \frac{\frac{20}{50} \cdot \frac{2}{10}}{\frac{20}{50} \cdot \frac{2}{10} + \frac{10}{50} \cdot \frac{1}{10} + \frac{20}{50} \cdot \frac{4}{10}} \\
&= \frac{\frac{2}{25}}{\frac{13}{50}} \\
&= \frac{4}{13}
\end{aligned}$$

Portanto, a probabilidade da camisa defeituosa ser da marca A é $\frac{4}{13}$.
A generalização da fórmula utilizada para determinar $\mathbb{P}(A|D)$ é denominada fórmula de Bayes.

Propriedade 2.27. ***Fórmula de Bayes.*** *Sejam* A *um evento e* $B_1, B_2, ..., B_k$ *uma partição do espaço amostral* Ω*, onde está definida a probabilidade* P*, então:*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(B_i) \cdot \mathbb{P}(A|B_i)}{\mathbb{P}(B_1) \cdot \mathbb{P}(A|B_1) + \mathbb{P}(B_2) \cdot \mathbb{P}(A|B_2) + \cdots + \mathbb{P}(B_k) \cdot \mathbb{P}(A|B_k)}$$

Exemplo 2.28 Uma urna contém três dados cúbicos simétricos com as seguintes características:

- dado D_1 tem 1 face branca e 5 faces pretas;
- dado D_2 tem 2 faces brancas e 4 faces pretas;
- dado D_3 tem 3 faces brancas e 3 faces pretas.

Um menino escolhe aleatoriamente e lança um dado dessa urna. Se a face que caiu voltada para cima foi branca (B), qual é a probabilidade dele ter escolhido o dado D_3? ◂

Solução: Para utilizar a fórmula de Bayes, vamos considerar o espaço amostral Ω as 18 faces, pois são 3 dados com 6 faces cada um, o evento B caiu uma face branca voltada para cima e D_1, D_2 e D_3 uma partição de Ω que representa as 6 faces do dado que foi escolhido. Logo:

$$\begin{aligned}\mathbb{P}(D_3|B) &= \frac{\mathbb{P}(D_3)\cdot\mathbb{P}(B|D_3)}{\mathbb{P}(D_1)\cdot\mathbb{P}(B|D_1)+\mathbb{P}(D_2)\cdot\mathbb{P}(B|D_2)+\mathbb{P}(D_3)\cdot\mathbb{P}(B|D_3)} \\ &= \frac{\frac{1}{3}\cdot\frac{3}{6}}{\frac{1}{3}\cdot\frac{1}{6}+\frac{1}{3}\cdot\frac{2}{6}+\frac{1}{3}\cdot\frac{3}{6}} \\ &= \frac{1}{2}\end{aligned}$$

A probabilidade dele ter escolhido o dado D_3 é $\frac{1}{2}$. □

Dois eventos A e B são mutuamente excludentes se $A \cap B = \varnothing$, ou seja, se ocorrer um dos eventos o outro não ocorrerá, consequentemente, $\mathbb{P}(A|B) = 0$ e vice versa $\mathbb{P}(B|A) = 0$. No entanto, existem fenômenos e experimentos nos quais saber que A ocorreu não tem relação com B.

Exemplo 2.29 Suponha que duas moedas balanceadas são lançadas, uma depois da outra, e considere os eventos:

$$A = \{\text{saiu cara no primeiro lançamento}\}$$

$$B = \{\text{saiu cara no segundo lançamento}\}$$

Sabemos que $\mathbb{P}(A) = \frac{1}{2}$ e $\mathbb{P}(B) = \frac{1}{2}$. Além disso, como não existe nenhuma relação entre esses eventos, podemos afirmar que $\mathbb{P}(A|B) = \mathbb{P}(A) = \frac{1}{2}$ e concluir que:

$$\mathbb{P}(A \cap B) = \mathbb{P}(B)\cdot\mathbb{P}(A|B) = \mathbb{P}(B)\cdot\mathbb{P}(A) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$$

◂

O exemplo acima é uma das motivações da seguinte definição.

Definição 2.30. *Dois eventos A e B são independentes se, e somente se:*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\cdot\mathbb{P}(B)$$

Exemplo 2.31 Serão retiradas ao acaso duas bolas de urna que contém cinco bolas numeradas de 1 a 5, de modo que a primeira bola é recolocada na urna antes de retirar a segunda. Considere os eventos:

$$P = \{\text{saiu uma bola de número par}\}$$

$$I = \{\text{saiu uma bola de número ímpar}\}$$

Calcule a probabilidade de retirar:

a) duas bolas pares.

b) a primeira bola par e a segunda ímpar.

c) a primeira bola ímpar e a segunda par.

d) duas bolas ímpares.

◂

Solução: Os eventos P e I são independentes porque a primeira é recolocada na urna antes de retirar a segunda, logo:

a) a probabilidade de retirar duas bolas pares é:

$$\mathbb{P}\left(\mathrm{P}\cap\mathrm{P}\right)=\mathbb{P}\left(\mathrm{P}\right)\cdot\mathbb{P}\left(\mathrm{P}\right)=\frac{2}{5}\cdot\frac{2}{5}=\frac{4}{25}$$

b) a probabilidade de retirar a primeira bola par e a segunda ímpar é:

$$\mathbb{P}\left(\mathrm{P}\cap\mathrm{I}\right)=\mathbb{P}\left(\mathrm{P}\right)\cdot\mathbb{P}\left(\mathrm{I}\right)=\frac{2}{5}\cdot\frac{3}{5}=\frac{6}{25}$$

c) a probabilidade de retirar a primeira bola ímpar e a segunda par é:

$$\mathbb{P}\left(\mathrm{I}\cap\mathrm{P}\right)=\mathbb{P}\left(\mathrm{I}\right)\cdot\mathbb{P}\left(\mathrm{P}\right)=\frac{3}{5}\cdot\frac{2}{5}=\frac{6}{25}$$

d) a probabilidade de retirar duas bolas ímpares é:

$$\mathbb{P}\left(\mathrm{I}\cap\mathrm{I}\right)=\mathbb{P}\left(\mathrm{I}\right)\cdot\mathbb{P}\left(\mathrm{I}\right)=\frac{3}{5}\cdot\frac{3}{5}=\frac{9}{25}$$

□

Observe que para determinar as probabilidades $\mathbb{P}\left(\mathrm{P}\cap\mathrm{P}\right)$, $\mathbb{P}\left(\mathrm{P}\cap\mathrm{I}\right)$, $\mathbb{P}\left(\mathrm{I}\cap\mathrm{P}\right)$ e $\mathbb{P}\left(\mathrm{I}\cap\mathrm{I}\right)$ poderíamos adotar o seguinte espaço amostral equiprovável.

•	1	2	3	4	5
1	(1; 1)	(2; 1)	(3; 1)	(4; 1)	(5; 1)
2	(1; 2)	(2; 2)	(3; 2)	(4; 2)	(5; 2)
3	(1; 3)	(2; 3)	(3; 3)	(4; 3)	(5; 3)
4	(1; 4)	(2; 4)	(3; 4)	(4; 4)	(5; 4)
5	(1; 5)	(2; 5)	(3; 5)	(4; 5)	(5; 5)

Onde a primeira linha da tabela contém os números da primeira bola retirada e a primeira coluna da tabela contém os números da segunda bola retirada. Contando a quantidade de casos favoráveis para cada uma das probabilidades, obtemos os mesmos resultados, veja:

a) Os casos favoráveis a retirada de duas bolas pares são $(2,2)$, $(2,4)$, $(4,2)$ e $(4,4)$, logo:

$$\mathbb{P}\left(\mathrm{P} \cap \mathrm{P}\right) = \frac{4}{25}$$

b) Os casos favoráveis a retirada da primeira bola par e a segunda ímpar são $(2,1)$, $(2,3)$, $(2,5)$, $(4,1)$, $(4,3)$ e $(4,5)$, logo:

$$\mathbb{P}\left(\mathrm{P} \cap \mathrm{I}\right) = \frac{6}{25}$$

c) Os casos favoráveis a retirada da primeira bola ímpar e a segunda par são $(1,2)$, $(1,4)$, $(3,2)$, $(3,4)$, $(5,2)$ e $(5,4)$, logo:

$$\mathbb{P}\left(\mathrm{I} \cap \mathrm{P}\right) = \frac{6}{25}$$

d) Os casos favoráveis a retirada de duas bolas ímpares são $(1,1)$, $(1,3)$, $(1,5)$, $(3,1)$, $(3,3)$, $(3,5)$, $(5,1)$, $(5,3)$ e $(5,5)$, logo:

$$\mathbb{P}\left(\mathrm{I} \cap \mathrm{I}\right) = \frac{9}{25}$$

2.5 Variáveis Aleatórias

A descrição de um espaço amostral Ω não contém necessariamente números. Por exemplo, quando retiramos 3 peças de um lote e verificamos se cada uma delas é boa (B) ou defeituosa (D), descrevemos:

$$\Omega = \{\text{BBB; BBD; BDB; DBB; BDD; DBD; DDB; DDD}\}$$

No entanto, podemos estar interessado na quantidade de peças defeituosas e não na ordem em que elas são retiradas. Nesse caso, é comum definir uma função que associa a cada evento elementar do espaço amostral o número de peças defeituosas, ou seja:

Evento elementar	Quantidade de peças defeituosas
BBB	0
BBD	1
BDB	1
DBB	1
BDD	2
DBD	2
DDB	2
DDD	3

A função definida é chamada de variável aleatória, veja.

Definição 2.32. *Seja Ω um espaço amostral associado a um experimento ε. Uma função X, que associa a cada elemento $s \in \Omega$ um número real $X(s)$, é denominada de:*

- ***variável aleatória discreta** se o conjunto formado pelos possíveis valores de $X(s)$ for enumerável.*
- ***variável aleatória contínua** se o conjunto formado pelos possíveis valores de $X(s)$ for um intervalo de números reais.*

Neste capítulo, trataremos apenas de variáveis aleatórias discretas.

Exemplo 2.33 Considere que um aluno escolheu uma alternativa ao acaso em 3 testes de múltipla escolha de um vestibular, ou seja, "chutou"3 questões nesse vestibular. Podemos associar o espaço amostral abaixo, em que A indica um acerto e E um erro, a esse experimento aleatório.

$$\Omega = \{\text{AAA; AAE; AEA; EAA; AEE; EAE; EEA; EEE}\}$$

Se o aluno está interessado em saber o número de teste que ele acertou ao acaso, definimos então a variável aleatória X é o número de acertos ao acaso, logo:

$$X(\text{AAA}) = 3$$

$$X(\text{AAE; AEA; EAA}) = 2$$

$$X(\text{AEE; EAE; EEA}) = 1$$

$$X(\text{EEE}) = 0$$

◂

A probabilidade de cada valor da variável aleatória é dada pela função a seguir.

Definição 2.34. *A função que associa a cada valor $X(s_i)$ da variável aleatória sua probabilidade $\mathbb{P}(X(s_i)) = p_i$ é denominada função discreta de probabilidade ou função de probabilidade. Os valores de p_i devem satisfazer às seguintes condições:*

- $0 \leqslant p_i \leqslant 1$;
- $\sum p_i = 1$

Exemplo 2.35 Suponha que uma moeda viciada é lançada 3 vezes e, a cada lançamento, a probabilidade da face voltada para cima ser cara é 0,2 e ser coroa é 0,8. Qual é a probabilidade de nesse experimento observar 0, 1, 2 ou 3 coroas? ◂

Solução: Um espaço amostral associado ao experimento é:

$$\Omega = \{(KKK), (KKC), (KCK), (CKK), (KCC), (CKC), (CCK), (CCC)\}$$

Onde K indica cara e C coroa. A probabilidade de cada um desses eventos é:

- observou 3 caras.
$$\mathbb{P}(\mathsf{KKK}) = 0,2 \cdot 0,2 \cdot 0,2 = 0,2^3$$

- observou 2 caras e 1 coroa no último lançamento.
$$\mathbb{P}(\mathsf{KKC}) = 0,2 \cdot 0,2 \cdot 0,8 = 0,2^2 \cdot 0,8$$

- observou 2 caras e 1 coroa no segundo lançamento.
$$\mathbb{P}(\mathsf{KCK}) = 0,2 \cdot 0,8 \cdot 0,2 = 0,2^2 \cdot 0,8$$

- observou 2 caras e 1 coroa no primeiro lançamento.
$$\mathbb{P}(\mathsf{CKK}) = 0,8 \cdot 0,2 \cdot 0,2 = 0,2^2 \cdot 0,8$$

- observou 2 coroas e 1 cara no primeiro lançamento.
$$\mathbb{P}(\mathsf{KCC}) = 0,2 \cdot 0,8 \cdot 0,8 = 0,2 \cdot 0,8^2$$

- observou 2 coroas e 1 cara no segundo lançamento.
$$\mathbb{P}(\mathsf{CKC}) = 0,8 \cdot 0,2 \cdot 0,8 = 0,2 \cdot 0,8^2$$

- observou 2 coroas e 1 cara no último lançamento.
$$\mathbb{P}(\mathsf{CCK}) = 0,8 \cdot 0,8 \cdot 0,2 = 0,2 \cdot 0,8^2$$

- observou 3 coroas.
$$\mathbb{P}(\mathsf{CCC}) = 0,8 \cdot 0,8 \cdot 0,8 = 0,8^3$$

Considerando X a variável aleatória que faz as seguintes correspondências.

- Observou nenhuma coroa.
$$\mathsf{X}(\mathsf{KKK}) = 0$$

- Observou uma coroa.
$$\mathsf{X}(\mathsf{KKC}, \mathsf{KCK}, \mathsf{CKK}) = 1$$

- Observou duas coroas.
$$\mathsf{X}(\mathsf{KCC}, \mathsf{CKC}, \mathsf{CCK}) = 2$$

- Observou três coroas.
$$\mathsf{X}(\mathsf{CCC}) = 3$$

A função de probabilidade $\mathbb{P}(\mathsf{X})$ associada à variável aleatória X é dada por:

I) a probabilidade de observar nenhuma coroa.

$$\begin{aligned} \mathbb{P}(0) &= \mathbb{P}(\mathsf{KKK}) \\ &= 0,2^3 \end{aligned}$$

II) A probabilidade de observar 1 coroa.

$$\begin{aligned} \mathbb{P}(1) &= \mathbb{P}(\mathsf{KKC}) + \mathbb{P}(\mathsf{KCK}) + \mathbb{P}(\mathsf{CKK}) \\ &= 0,2^2 \cdot 0,8 + 0,2^2 \cdot 0,8 + 0,2^2 \cdot 0,8 \\ &= 3 \cdot 0,2^2 \cdot 0,8 \end{aligned}$$

III) A probabilidade de observar 2 coroas.

$$\begin{aligned} \mathbb{P}(2) &= \mathbb{P}(\mathsf{KCC}) + \mathbb{P}(\mathsf{CKC}) + \mathbb{P}(\mathsf{CCK}) \\ &= 0,2 \cdot 0,8^2 + 0,2 \cdot 0,8^2 + 0,2 \cdot 0,8^2 \\ &= 3 \cdot 0,2 \cdot 0,8^2 \end{aligned}$$

IV) A probabilidade de observar 3 coroas.

$$\begin{aligned} \mathbb{P}(3) &= \mathbb{P}(\mathsf{CCC}) \\ &= 0,8^3 \end{aligned}$$

□

Uma característica importante no exemplo anterior é que as probabilidades $\mathbb{P}(0)$, $\mathbb{P}(1)$, $\mathbb{P}(2)$ e $\mathbb{P}(3)$ são os resultados do desenvolvimento do binômio $(0,2+0,8)^3$, veja:

$$(0,2+0,8)^3 = \underbrace{0,2^3}_{\mathrm{I}} + \underbrace{3 \cdot 0,2^2 \cdot 0,8}_{\mathrm{II}} + \underbrace{3 \cdot 0,2 \cdot 0,8^2}_{\mathrm{III}} + \underbrace{0,8^3}_{\mathrm{IV}}$$

I) $\mathbb{P}(0) = 0,2^3$

II) $\mathbb{P}(1) = 3 \cdot 0,2^2 \cdot 0,8$

III) $\mathbb{P}(2) = 3 \cdot 0,2 \cdot 0,8^2$

IV) $\mathbb{P}(3) = 0,8^3$

Quando isso ocorre, dizemos que X é uma variável aleatória com distribuição binomial.

Definição 2.36. *Seja A um evento do espaço amostral associado ao experimento ε, de modo que $\mathbb{P}(\mathsf{A}) = \mathsf{p}$ e $\mathbb{P}(\mathsf{A}^{\mathsf{C}}) = 1 - \mathsf{p}$. Considere n repetições de ε em que a probabilidade $\mathbb{P}(\mathsf{A}) = \mathsf{p}$ permaneça a mesma e a variável aleatória X indica o número k de vezes que o evento A ocorreu. Denominamos X de variável aleatória com **distribuição binomial**.*

Propriedade 2.37. *Seja X uma variável aleatória com distribuição binomial, baseada em n repetições. Então:*

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k};\ k = 0, 1, 2, 3, ..., n$$

Demonstração. Suponha que ao repetir n vezes um experimento ε, o evento A, que tem $\mathbb{P}(A) = p$ invariável a cada repetição, ocorreu nos k primeiros ensaios e não ocorreu nos $(n - k)$ últimos ensaios. Desse modo, a probabilidade deste caso particular é:

$$\underbrace{p \cdot p \cdot ... \cdot p}_{\text{k vezes}} \cdot \underbrace{(1-p) \cdot (1-p) \cdot ... \cdot (1-p)}_{\text{(n - k) vezes}} = p^k (1-p)^{n-k}$$

Observe que todos os casos particulares em que o evento A ocorre k vezes têm a mesma probabilidade $p^k(1-p)^{n-k}$, uma vez que só mudaria a ordem dos fatores na multiplicação.

O número de casos particulares de $(X = k)$ dentre as n repetições é igual ao número de combinações n elementos tomados k a k, ou seja, $\binom{n}{k}$. Logo:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k};\ k = 0, 1, 2, 3, ..., n$$

□

Exemplo 2.38 Um dado cúbico simétrico é lançado 10 vezes, a cada lançamento é anotado o número da face voltada para cima. Qual é a probabilidade de anotar 3 vezes a face de número 1? ◂

Solução: Sejam A o evento anotar a face número 1 e X a variável aleatória que indica o número de vezes que A ocorreu, então devemos calcular a probabilidade de $\mathbb{P}(X = 3)$.
O experimento ε lançar o dado é repetido 10 vezes, a cada repetição de ε, a probabilidade p de ocorrer o evento A é sempre a mesma, $p = \frac{1}{6}$ e, consequentemente, a probabilidade $1 - p$ também é sempre a mesma, $1 - p = \frac{5}{6}$. Logo, a variável aleatória X tem distribuição binomial e $\mathbb{P}(X = 3)$ é dada por:

$$\begin{aligned}
\mathbb{P}(X = 3) &= \binom{10}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^{10-3} \\
\mathbb{P}(X = 3) &= 120 \cdot \left(\frac{1}{216}\right) \cdot \left(\frac{78125}{279936}\right) \\
\mathbb{P}(X = 3) &\approx 0,155045
\end{aligned}$$

A probabilidade de anotar 3 vezes a face de número um é aproximadamente $0,155$. □

2.6 Valor Esperado

Uma avaliação escolar contém 4 questões em que o aluno deve assinalar uma única alternativa dentre 4 opções de respostas apresentadas. Suponha que um aluno não estudou e marcou todas as respostas ao acaso, qual é número de acertos mais provável desse aluno? Para responder essa questão, vamos considerar X a variável aleatória que indica o número de acertos desse aluno, $\frac{1}{4}$ a probabilidade do aluno marcar uma resposta correta ao acaso e $\frac{3}{4}$ a probabilidade dele marcar uma resposta incorreta, então X é uma variável aleatória com distribuição binomial e suas probabilidades são:

$$\mathbb{P}(\mathsf{X}=0)=\binom{4}{0}\cdot\left(\frac{1}{4}\right)^{0}\cdot\left(\frac{3}{4}\right)^{4-0}=\frac{81}{256}$$

$$\mathbb{P}(\mathsf{X}=1)=\binom{4}{1}\cdot\left(\frac{1}{4}\right)^{1}\cdot\left(\frac{3}{4}\right)^{4-1}=\frac{108}{256}$$

$$\mathbb{P}(\mathsf{X}=2)=\binom{4}{2}\cdot\left(\frac{1}{4}\right)^{2}\cdot\left(\frac{3}{4}\right)^{4-2}=\frac{54}{256}$$

$$\mathbb{P}(\mathsf{X}=3)=\binom{4}{3}\cdot\left(\frac{1}{4}\right)^{3}\cdot\left(\frac{3}{4}\right)^{4-3}=\frac{12}{256}$$

$$\mathbb{P}(\mathsf{X}=4)=\binom{4}{4}\cdot\left(\frac{1}{4}\right)^{4}\cdot\left(\frac{3}{4}\right)^{4-4}=\frac{1}{256}$$

Observe que a maior probabilidade é $\mathbb{P}(\mathsf{X}=1)=\frac{108}{256}$. Logo, o mais provável é o aluno acertar 1 questão. Observe também que se calcularmos a média M do número de acertos ponderando suas probabilidades, obtemos $\mathsf{M}=1$, veja.

$$\begin{aligned}\mathsf{M} &= 0\cdot\frac{81}{256}+1\cdot\frac{108}{256}+2\cdot\frac{54}{256}+3\cdot\frac{12}{256}+4\cdot\frac{1}{256}\\ &= \frac{108+108+36+4}{256}\\ &= 1\end{aligned}$$

Nesse contexto, a média ponderada M é chamada de valor esperado da variável aleatória X e definida do seguinte modo.

Definição 2.39. *Seja X uma variável aleatória discreta, o valor esperado de X, ou esperança de X, denotado por* $\mathsf{E}(\mathsf{X})$ *é:*

$$\mathsf{E}(\mathsf{X})=\sum_{i=1}^{k}x_ip_i$$

Exemplo 2.40 Considere uma dado cúbico viciado com a seguinte variável aleatória X.

x_i	1	2	3	4	5	6
p_i	0,1	0,1	0,1	0,1	0,1	0,5

O valor esperado de X é:

$$E(X) = 1 \cdot 0,1 + 2 \cdot 0,1 + 3 \cdot 0,1 + 4 \cdot 0,1 + 5 \cdot 0,1 + 6 \cdot 0,5 = 4,5$$

◂

No exemplo acima, o valor esperado não é um valor possível de X. A interpretação deste resultado, $E(X) = 4,5$, em um sentido probabilístico, é que esperamos que a média aritmética dos resultados, a cada vez que repetimos o experimento, se aproxime de 4,5.

Exemplo 2.41 Uma indústria constatou que a probabilidade de produzir uma peça defeituosa é 0,02 e não defeituosa é 0,98. Além disso, ela sabe que tem um prejuízo de 40 reais quando fabrica uma peça defeituosa e um lucro de 100 reais quando fabrica uma peça não defeituosa. Qual é o lucro médio esperado com a produção de uma peça? ◂

Solução: Seja X a variável aleatória que associa ao espaço amostral $\Omega = \{\text{defeituosa, não defeituosa}\}$ os seguintes números:

$$X\,(\text{defeituosa}) = -40$$
$$X\,(\text{não defeituosa}) = 100$$

Desse modo, a função de probabilidade associada à variável aleatória X é:

$$\mathbb{P}\,(-40) = 0,02$$

$$\mathbb{P}\,(100) = 0,98$$

O lucro médio esperado é o valor médio esperado $E(X)$ cujo valor é:

$$E(X) = -40 \cdot 0,02 + 100 \cdot 0,98 = 97,2$$

Portanto, o lucro médio esperado ao produzir uma peça é 97,2 reais. □

A interpretação de $E(X) = 97,2$ é que se a indústria fabrica ao longo de um ano 20 mil peças, então ela terá um lucro médio anual de $20.000 \cdot 97,2 = 1.944.000$ reais com essa produção.

Definição 2.42. *Seja X uma variável aleatória, definimos a variância de X, denotado por* $V(X)$ *ou* σ^2*, do seguinte modo:*

$$V(X) = E[X - E(X)]^2$$

A raiz quadrada de $V(X)$ *é chamada de* ***desvio padrão*** *de X e indicada por* σ.

Propriedade 2.43. *O cálculo de* $V(X)$ *pode ser efetuado do seguinte modo:*

$$V(X) = E(X^2) - E(X)^2$$

Demonstração.

$$\begin{aligned} V(X) &= E[X - E(X)]^2 \\ &= E[X^2 - 2XE(X) + E(X)^2] \\ &= E(X^2) - 2E(X)E(X) + E(X)^2 \\ &= E(X^2) - E(X)^2 \end{aligned}$$

□

Exemplo 2.44 Considere dois dados cúbicos viciados com as seguintes variáveis aleatórias.

X_1	1	2	3	4	5	6
p_i	0,05	0,05	0,4	0,4	0,05	0,05

X_2	1	2	3	4	5	6
p_i	0,4	0,05	0,05	0,05	0,05	0,4

Calcule a variância e o desvio padrão de X_1 e X_2. ◂

Solução: Os valores esperados das variáveis X_1 e X_2 são:

$$E(X_1) = 1 \cdot 0,05 + 2 \cdot 0,05 + 3 \cdot 0,4 + 4 \cdot 0,4 + 5 \cdot 0,05 + 6 \cdot 0,05 = 3,5$$

$$E(X_2) = 1 \cdot 0,4 + 2 \cdot 0,05 + 3 \cdot 0,05 + 4 \cdot 0,05 + 5 \cdot 0,05 + 6 \cdot 0,4 = 3,5$$

Para calcular as variâncias, precisamos dos seguintes valores:

$$E(X_1^2) = 1^2 \cdot 0,05 + 2^2 \cdot 0,05 + 3^2 \cdot 0,4 + 4^2 \cdot 0,4 + 5^2 \cdot 0,05 + 6^2 \cdot 0,05 = 13,3$$

$$E(X_2^2) = 1^2 \cdot 0,4 + 2^2 \cdot 0,05 + 3^2 \cdot 0,05 + 4^2 \cdot 0,05 + 5^2 \cdot 0,05 + 6^2 \cdot 0,4 = 17,5$$

As variâncias são:

$$V(X_1) = 13,3 - 3,5^2 = 1,05$$

$$V(X_2) = 17,5 - 3,5^2 = 5,25$$

Os desvios padrões são:

$$\sigma_1 = \sqrt{1,05} \approx 1,02$$

$$\sigma_2 = \sqrt{5,25} \approx 2,29$$

□

O exemplo ilustra que as variáveis aleatórias X_1 e X_2 têm o mesmo valor esperado, $E(X_1) = E(X_2) = 3,5$, e variâncias distintas, $V(X_1) = 1,05$ e $V(X_2) = 5,25$. Observe que o dado associado à variância $V(X_1) = 1,05$ apresenta 0,8 de probabilidade para ocorrência de uma face 3 ou 4, que são as faces mais próximas do valor esperado $E(X_1) = 3,5$, enquanto o dado associado à variância $V(X_1) = 5,25$ apresenta 0,8 de probabilidade para ocorrência de uma face 1 ou 6, que são as faces mais distantes do valor esperado $E(X_1) = 3,5$. Ou seja, a variância é um indicativo de dispersão em relação ao valor esperado, quanto maior a variância maior é a probabilidade de dispersão dos resultados em relação ao valor esperado.

3 Estatística

Uma pesquisa estatística consiste em selecionar e estudar uma quantidade de indivíduos para obter informações sobre o grupo ao qual eles pertencem. Veja as definições.

Definição 3.1. ***População*** *é o conjunto que contém todos os indivíduos, elementos, com a característica, ou propriedade, que se deseja estudar.*

Definição 3.2. ***Parâmetro da população*** *é a característica, ou propriedade, que se deseja estudar em uma população.*

Definição 3.3. ***Amostra*** *é um subconjunto da população.*

Definição 3.4. ***Estatística da amostra*** *é uma característica, ou propriedade, da amostra que possibilita estimar, ou analisar, o valor do parâmetro da população.*

Exemplo 3.5 Suponha que em uma cidade com 10 mil habitantes, um fabricante de roupas calculou a altura média de um grupo de 100 moradores selecionados aleatoriamente em um determinado dia para estimar a altura média dessa população.

- A população é o conjunto composto pelos 10 mil habitantes dessa cidade.
- O parâmetro da população é a altura média dos 10 mil habitantes.
- A amostra é o subconjunto da população composto pelo grupo de 100 moradores selecionados.
- A estatística da amostra é a altura média dos 100 moradores selecionados.

◂

Considerando ainda o exemplo, imagine que, um dia depois, o fabricante seleciona novamente 100 pessoas aleatoriamente. Essa nova amostragem será igual ou diferente a do dia anterior? E a nova altura média: será um valor próximo ou igual a do dia anterior? Como as amostras são obtidas de forma aleatória, elas devem receber um tratamento probabilístico, logo não podemos afirmar nem que as amostras e nem que as alturas médias serão iguais. A figura a seguir ilustra os mecanismos de uma pesquisa estatística.

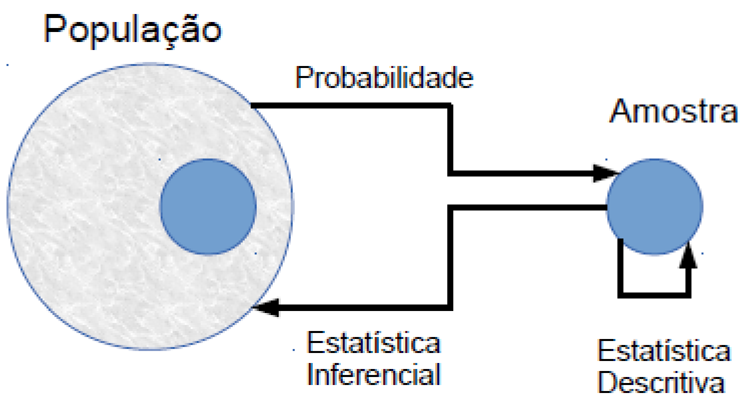


Figura 3.1: Pesquisa Estatística

Definição 3.6. *Seja X uma variável aleatória com uma função de probabilidade especificada. Sejam n variáveis aleatórias* $X_1, X_2, X_3, ..., X_n$ *independentes e tendo cada uma delas a mesma distribuição que X. Denominamos* $X_1, X_2, X_3, ..., X_n$ *ok a amostra aleatória da variável aleatória X.*

Em uma linguagem menos formal, uma amostra aleatória de tamanho n de uma variável aleatória X, corresponde a n mensurações repetidas de X, feitas sob condições essencialmente inalteradas.

Exemplos 3.7

- Um dado é lançado 10 vezes, os 10 resultados são uma amostra aleatória de tamanho 10 de uma população de todas as jogadas desse dado.
- Um diretor escolar escreve o nome de cada um dos seus 500 alunos em um pedaço de papel, todos de mesmo tamanho, coloca os papéis em um envelope, agita e sorteia 50 alunos para saber o interesse deles sobre um novo curso extracurricular. Os alunos sorteados são uma amostra aleatória de tamanho 50 da população de 500 alunos da escola.
- De um lote de 800 lâmpadas, um inspetor de qualidade verifica a vida útil de 8 lâmpadas escolhidas ao acaso. As lâmpadas escolhida são uma amostra aleatória de tamanho 8 da população de 800 lâmpadas.

Definição 3.8. *Seja* $(X_1, X_2, ..., X_n)$ *uma amostra aleatória de uma variável aleatória X. As seguintes estatísticas são importantes:*

1. *Média amostral* $\overline{X}$.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

2. *Variância amostral* S^2.

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

No capítulo anterior, definimos variância que neste capítulo, dependendo do contexto, será denominada de variância populacional para diferenciar da variância amostral.

Na seção estimação de parâmetros, explicaremos porque na definição de variância amostral dividimos por $(n-1)$ *e não por* n.

3. *Mínimo amostral Mín.*

$$\textit{Mín} = \textit{mínimo}(X_1, X_2, ..., X_n)$$

4. *Máximo amostral Máx.*

$$\textit{Máx} = \textit{máximo}(X_1, X_2, ..., X_n)$$

5. *Amplitude da amostra Amp.*

$$\textit{Amp} = \textit{Máx} - \textit{Mín}$$

O exemplo a seguir ilustra como calcular essas estatísticas.

Exemplo 3.9 Um fabricante de calçados selecionou ao acaso e mediu em centímetros o comprimento dos pés de 10 crianças de 8 anos. Veja os resultados na tabela abaixo.

X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
23	23	24	22	24	22	22,5	23,5	22,5	23,5

Em seguida, o fabricante calculou as estatísticas:

a) Média amostral $\overline{X}$.

$$\overline{X} = \frac{1}{10}\cdot(23 + 23 + 24 + 22 + 24 + 22 + 22,5 + 23,5 + 22,5 + 23,5) = 23$$

b) Variância amostral S^2.

X_i	$X_i - \overline{X}$	$(X_i - \overline{X})^2$
23	$23 - 23 = 0$	0
23	$23 - 23 = 0$	0
24	$24 - 23 = 1$	1
22	$22 - 23 = -1$	1
24	$24 - 23 = 1$	1
22	$22 - 23 = -1$	1
22,5	$22,5 - 23 = -0,5$	0,25
23,5	$23,5 - 23 = 0,5$	0,25
22,5	$22,5 - 23 = -0,5$	0,25
23,5	$23,5 - 23 = 0,5$	0,25

$$S^2 = \frac{1}{10-1} \cdot (0+0+1+1+1+1+0,25+0,25+0,25+0,25) = \frac{5}{9}$$

c) Mínimo amostral $Mín$.

$$Mín = 22$$

d) Máximo amostral $Máx$.

$$Máx = 24$$

e) Amplitude da amostra Amp.

$$Amp = 24 - 22 = 2$$

◂

O próximo teorema apresenta propriedades da média amostral.

Teorema 3.10. *Seja X uma variável aleatória, com valor esperado* $E(X) = \mu$ *e variância* $V(X) = \sigma^2$. *Seja* $\overline{X}$ *a média amostral de uma amostra aleatória de tamanho* n. *Então:*

a) $E(\overline{X}) = \mu$

b) $V(\overline{X}) = \frac{\sigma^2}{n}$

c) Para n *grande,* $\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$ *terá aproximadamente a distribuição* $N(0,1)$.

Vamos analisar o exemplo a seguir para compreender o teorema.

Exemplo 3.11 Considere que uma certa população tem a seguinte variável aleatória.

X_i	10	20
p_i	0,6	0,4

O valor esperado $E(X)$ é:

$$E(X) = 10 \cdot 0,6 + 20 \cdot 0,4 = 14$$

A variância $V(X)$ é:

$$V(X) = (10^2 \cdot 0,6 + 20^2 \cdot 0,4) - 14^2 = 24$$

Imagine que uma amostra aleatória de tamanho $n = 2$ é obtida dessa população. As possíveis amostras são:

$(X_1; X_2)$	$(10; 10)$	$(10; 20)$	$(20; 10)$	$(20; 20)$
$\overline{X}$	10	15	15	20
$p_{\overline{X}}$	$0,6 \cdot 0,6 = 0,36$	$0,6 \cdot 0,4 = 0,24$	$0,4 \cdot 0,6 = 0,24$	$0,4 \cdot 0,4 = 0,16$

O valor esperado da média amostral $E(\overline{X})$ é:

$$E(\overline{X}) = 10 \cdot 0,36 + 15 \cdot 0,24 + 15 \cdot 0,24 + 20 \cdot 0,16 = 14$$

A variância da média amostral $V(\overline{X})$ é:

$$V(\overline{X}) = (10^2 \cdot 0,36 + 15^2 \cdot 0,24 + 15^2 \cdot 0,24 + 20^2 \cdot 0,16) - 14^2 = 12$$

Observe que o valor esperado da média amostral é igual ao valor esperado da população e a variância da média amostral é igual a variância populacional dividido pelo tamanho da amostra, ou seja:

$$E(\overline{X}) = E(X) = 14$$

$$V(\overline{X}) = \frac{V(X)}{n} = \frac{24}{2} = 12$$

◂

3.1 Estimação de parâmetros

Nesta seção serão apresentados métodos que analisam o grau de confiança de uma estatística. Isto é, dada uma estatística de um parâmetro populacional, será possível classificá-la em tendenciosa ou não tendenciosa, coerente ou não coerente. Além disso, se duas ou mais estatísticas distintas têm como objetivo estimar o mesmo parâmetro populacional, veremos como escolher a "melhor"delas.

Definição 3.12. *Seja X uma variável aleatória com alguma distribuição de probabilidade que dependa de um parâmetro* θ *desconhecido. Seja* $X_1, X_2, ..., X_n$ *uma amostra de X e* $x_1, x_2, ..., x_n$ *os correspondentes valores amostrais, denominamos de:*

- ***Estimador*** *a função* $f(X_1, X_2, ..., X_n)$ *utilizada para estimação de* θ*, ou dizemos* f *é um estimador de* θ*.*
- ***Estimativa*** *de* θ*, indicado por* $\widehat{\theta}$*, o valor que a função* $f(x_1, x_2, ..., x_n)$ *assume, ou seja:*

$$\widehat{\theta} = f(x_1, x_2, ..., x_n)$$

As estatísticas média amostral, variância amostral, mínimo amostral, máximo amostral e amplitude amostral são exemplos de funções que denominamos de estimadores e os valores que elas assumem denominamos de estimativas.

Definição 3.13. *Seja* f *um estimador do parâmetro desconhecido* θ *associado à distribuição da variável aleatória X, tal que* $f(x_1, x_2, ..., x_n) = \widehat{\theta}$*. Dizemos que* f *é um estimador não tendencioso de* θ *se:*

$$E(\widehat{\theta}) = \theta$$

Um estimador é não tendencioso se o seu valor esperado é igual ao valor do parâmetro de interesse. A média amostral é um estimador não tendencioso.

Exemplo 3.14 Considere uma amostra aleatória $(X_1, X_2, ..., X_n)$ de uma população com média μ e variância σ^2, enquanto a variância populacional $V(X)$ é um estimador tendencioso, a variância amostral S^2 é um estimador não tendencioso, uma vez que:

$$
\begin{aligned}
\sum_{i=1}^{n}(X_i - \overline{X})^2 &= \sum_{i=1}^{n}(X_i - \mu + \mu - \overline{X})^2 \\
&= \sum_{i=1}^{n}[(X_i - \mu)^2 + 2(\mu - \overline{X})(X_i - \mu) + (\mu - \overline{X})^2] \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 + 2(\mu - \overline{X})\sum_{i=1}^{n}(X_i - \mu) + n(\mu - \overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 - 2n(\mu - \overline{X})^2 + n(\mu - \overline{X})^2 \\
&= \sum_{i=1}^{n}(X_i - \mu)^2 - n(\mu - \overline{X})^2
\end{aligned}
$$

Logo, o valor esperado da variância populacional $V(X)$ é dado por:

$$
\begin{aligned}
E(V(X)) &= E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) \\
&= \frac{1}{n}E\left(\sum_{i=1}^{n}(X_i - \mu)^2 - n(\mu - \overline{X})^2]\right) \\
&= \frac{1}{n}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) \\
&= \left(\frac{n-1}{n}\right)\sigma^2
\end{aligned}
$$

O quociente $\frac{n-1}{n}$ nunca será 1, exceto no limite quando n tende ao infinito, consequentemente, o valor esperado de $V(X)$ não será igual σ^2, por isso a variância populacional é um estimador tendencioso. Por outro lado, para a variância amostral, temos:

$$\begin{aligned} E(S^2) &= E\left(\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2\right) \\ &= \frac{1}{n-1}\left(n\sigma^2 - n\frac{\sigma^2}{n}\right) \\ &= \sigma^2 \end{aligned}$$

Portanto, a variância amostral é um estimador não tendencioso. ◂

Definição 3.15. *Seja* $\widehat{\theta}$ *uma estimativa não tendenciosa de* θ*. Dizemos que* $\widehat{\theta}$ *é uma estimativa não tendenciosa de variância mínima de* θ*, se para todas as estimativas de* θ^**, tais que* $E(\theta^*) = \theta$*, tivermos* $V(\widehat{\theta}) \leqslant V(\theta^*)$ *para todo* θ*. Isto é, dentre todas as estimativas não tendenciosas de* θ*,* $\widehat{\theta}$ *tem a variância menor de todas.*

Exemplo 3.16 Considere os seguintes estimadores, $\widehat{\theta_1}$ e $\widehat{\theta_2}$, para média populacional.
A média aritmética do máximo e do mínimo amostral $\widehat{\theta_1}$.

$$\widehat{\theta_1} = \frac{\text{Máx.} + \text{Mín.}}{2}$$

A média amostral $\widehat{\theta_2}$.

$$\widehat{\theta_2} = \overline{X}$$

A variância de $\widehat{\theta_1}$ não se altera com o tamanho da amostra, ou seja, sempre será um valor fixo, enquanto a variância de $\widehat{\theta_2}$, quanto maior o tamanho da amostra, menor será a sua variância, porque a média amostral depende do tamanho da amostra. Podemos observar esse fato de forma intuitiva, considere a seguinte lista de números.

Lista	4	6	2	2	8	10	4	6	6	10	4	6

Se calcularmos a média dois a dois desses números na ordem apresentada, teremos as seguintes médias L2:

L2	5	2	9	5	8	5

Se calcularmos a média três a três dos números da lista na ordem apresentada, teremos as seguintes médias L3:

L3	4	6,666...	5,333...	6,666...

Se calcularmos a média quatro a quatro, teremos três médias L4:

L4	3,5	7	6,5

Se calcularmos a média seis a seis, teremos duas médias L6:

L6	5,333..	6

A variância da lista L6 é menor do que a variância da lista L4, a variância da lista L4 é menor do que a L3, e assim sucessivamente, ou seja:

$$V(L6) < V(L4) < V(L3) < V(L2)$$

Por isso, a média amostral é uma estimativa não tendenciosa de variância mínima. ◂

Definição 3.17. *Seja* $\widehat{\theta}$ *uma estimativa de* θ *baseada em uma amostra de tamanho* n. *A estimativa* $\widehat{\theta}$ *será coerente se:*

$$\lim_{n\to\infty} E(\widehat{\theta}) = \theta$$

$$\lim_{n\to\infty} V(\widehat{\theta}) = 0$$

Um estimador é coerente se, a medida que o tamanho da amostra aumenta, seu valor esperado converge para o parâmetro de interesse e sua variância converge para zero.

Exemplo 3.18 Suponha que de uma amostra aleatória de n lâmpadas de um lote, k lâmpadas estavam queimadas e que a variável aleatória $X = \sum_{i=1}^{n} = k$ tem distribuição binomial.
A estimativa mais natural para o parâmetro p desse lote, que indica a probabilidade de um lâmpada está queimada, é a média amostral $\widehat{p} = \frac{k}{n}$. Calculando o valor esperado e a variância da média amostral, obtemos:

$$E\left(\widehat{p}\right) = E\left(\frac{X}{n}\right) = \frac{1}{n}\left(np\right) = p$$

$$V\left(\widehat{p}\right) = V\left(\frac{X}{n}\right) = \frac{1}{n^2}\left(np\right)\left(1-p\right) = \frac{p\left(1-p\right)}{n}$$

$$\lim_{n\to\infty} V(\widehat{p}) = \lim_{n\to\infty} \frac{p\left(1-p\right)}{n} = 0$$

Portanto, a média amostral é uma estimativa coerente. ◂

Definição 3.19. *Seja X uma variável aleatória com alguma distribuição de probabilidade que dependa de um parâmetro* θ *desconhecido. Seja* $X_1, X_2, ..., X_n$ *uma amostra de X e* $\widehat{\theta}$ *uma função de* $(X_1, X_2, ..., X_n)$*, dizemos que* $\widehat{\theta}$ *é a melhor estimativa não tendenciosa linear de* θ *se:*

a) $E(\widehat{\theta}) = \theta$

b) $\widehat{\theta} = \sum_{i=1}^{n} a_i X_i$. *Isto é,* $\widehat{\theta}$ *é uma função linear da amostra.*

c) $V(\widehat{\theta}) \leqslant V(\theta^*)$*, onde* θ^* *é qualquer outra estimativa de* θ *que satisfaça os itens anteriores a) e b).*

3.2 Estimativas de Máxima Verossimilhança

Na seção anterior, estimação de parâmetros, foram apresentados critérios para classificar uma estimativa em tendenciosa ou não tendenciosa, coerente ou não coerente e um método de comparar variâncias quando são apresentadas duas ou mais estimativas para o mesmo parâmetro.

Nesta seção será apresentado um método, um procedimento geral, com o qual é possível encontrar estimativas denominadas de Máxima Verossimilhança.

Definição 3.20. *A estimativa de máxima verossimilhança de* θ*, ou seja,* $\widehat{\theta}$*, baseada em uma amostra aleatória* $X_1, X_2, ..., X_n$ *é aquele valor de* θ *que torna máxima* $L(X_1, X_2, ..., X_n)$*, considerada como uma função de* θ *para um dada amostra* $X_1, X_2, ..., X_n$*, onde* L *é definida por:*

$$L(X_1, X_2, ..., X_n; \theta) = f(X_1; \theta)f(X_2; \theta)...f(X_n; \theta)$$

- *se X for discreta,* $L(X_1, X_2, ..., X_n; \theta)$ *representará* $P[X_1 = x_1, X_2 = x_2, ..., X_n = x_n]$
- *se X for contínua,* $L(X_1, X_2, ..., X_n; \theta)$ *representará a função densidade de probabilidade conjunta de* $(X_1, X_2, ..., X_n)$

Para encontrar a estimativa de máxima verossimilhança, devemos determinar o valor máximo de uma função. Uma das técnicas usuais é considerar a função $\ln(x)$, uma vez que ela é crescente, isto é:

$$\ln(L(X_1, X_2, ..., X_n; \theta))$$

Admitindo que θ seja um número real e que $\ln(L(X_1, X_2, ..., X_n; \theta))$ seja uma função derivável de θ, então a estimativa de máxima verossimilhança de θ é dada pela equação abaixo conhecida como equação de verossimilhança.

$$\ln'(L(X_1, X_2, ..., X_n; \theta)) = 0$$

Exemplo 3.21 Seleciona-se uma amostra aleatória de 20 peças $(X_1, X_2, ..., X_{20})$ de um lote de 10.000. Suponha que as 10.000 peças tenham a mesma probabilidade p de serem defeituosas e uma peça defeituosa não interfere na probabilidade da outra. Estime pelo método de máxima verossimilhança a probabilidade p de uma peça ser defeituosa nesse lote se na amostra foram encontradas 4 peças defeituosas. ◂

Solução: Seja X a seguinte variável aleatória.

$$X(\text{peça defeituosa}) = 1$$
$$X(\text{peça não defeituosa}) = 0$$

Considerando k o número de peças defeituosas da amostra, então:

$$k = \sum_{1}^{20} X_i$$

As probabilidades $\mathbb{P}(X = k)$ das 21 amostras possíveis em função do parâmetro p é dada pelo modelo probabilístico binomial, logo temos as seguintes possibilidades:

$$\mathbb{P}(X = 0) = \frac{20!}{0!\,(20-0)!}(1-p)^{20}p^{0}$$

$$\mathbb{P}(X = 1) = \frac{20!}{1!\,(20-1)!}(1-p)^{19}p^{1}$$

$$\mathbb{P}(X = 2) = \frac{20!}{2!\,(20-2)!}(1-p)^{18}p^{2}$$

$$\mathbb{P}(X = 3) = \frac{20!}{3!\,(20-3)!}(1-p)^{17}p^{3}$$

$$\mathbb{P}(X = 4) = \frac{20!}{4!\,(20-4)!}(1-p)^{16}p^{4}$$

$$\ldots$$

$$\mathbb{P}(X = 20) = \frac{20!}{20!\,(20-20)!}(1-p)^{0}p^{20}$$

Porém, foram encontradas na amostra 4 peças defeituosas, logo X = 4, então pelo método de Máxima Verossimilhança:

$$L(X_1, X_2, X_3, ..., X_{20}, p) = \mathbb{P}(X = 4)$$

$$L(X_1, X_2, X_3, ..., X_{20}, p) = \frac{20!}{4!\,(20-4)!}(1-p)^{16}p^{4}$$

$$\ln L(X_1, X_2, X_3, ..., X_{20}, p) = \ln\left(\frac{20!}{4!\,(20-4)!}(1-p)^{16}p^{4}\right)$$

$$\ln L(X_1, X_2, X_3, ..., X_{20}, p) = \ln 4.845 + \ln(1-p)^{16} + \ln p^{4}$$

$$\ln L(X_1, X_2, X_3, ..., X_{20}, p) = \ln 4.845 + 16\ln(1-p) + 4\ln p$$

$$\ln L'(p) = \frac{16}{(1-p)}(-1) + \frac{4}{p}$$

$$\frac{16}{(1-p)}(-1) + \frac{4}{p} = 0$$

$$4 - 4p = 16p$$

$$p = \frac{4}{20} = 0,2$$

Logo, a estimativa de p é $0,2$. □

Observe que para $p = 0,2$, a probabilidade $\mathbb{P}(X = 4)$ terá o maior valor possível entre as 21 possibilidades.

$$\mathbb{P}(X = 0) = \frac{20!}{0!\,(20-0)!}\,(1-0,2)^{20}\,0,2^{0} \approx 0,01152922$$

$$\mathbb{P}(X = 1) = \frac{20!}{1!\,(20-1)!}\,(1-0,2)^{19}\,0,2^{1} \approx 0,05764608$$

$$\mathbb{P}(X = 2) = \frac{20!}{2!\,(20-2)!}\,(1-0,2)^{18}\,0,2^{2} \approx 0,13690943$$

$$\mathbb{P}(X = 3) = \frac{20!}{3!\,(20-3)!}\,(1-0,2)^{17}\,0,2^{3} \approx 0,20536414$$

$$\mathbb{P}(X = 4) = \frac{20!}{4!\,(20-4)!}\,(1-0,2)^{16}\,0,2^{4} \approx 0,21819940$$

$$\mathbb{P}(X = 5) = \frac{20!}{5!\,(20-5)!}\,(1-0,2)^{15}\,0,2^{4} \approx 0,17455952$$

$$\mathbb{P}(X = 6) = \frac{20!}{6!\,(20-6)!}\,(1-0,2)^{14}\,0,2^{4} \approx 0,10909970$$

3.3 Coeficiente de Correlação

Em alguns contextos é necessário estudar a relação, ou determinar o grau de correlação, entre duas ou mais variáveis aleatórias. Nesta seção, apresentaremos o coeficiente de correlação $\rho_{(X,Y)}$ entre variáveis aleatórias bidimensionais (X, Y).

Definição 3.22. *Dada uma amostra aleatória bidimensional* $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ *o seu coeficiente de correlação* $\rho_{(X,Y)}$ *é definido por:*

$$\rho_{(X,Y)} = \frac{\sum_1^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_1^n \left(X_i - \overline{X}\right)^2 \sum_1^n \left(Y_i - \overline{Y}\right)^2}}$$

O coeficiente de correlação mede o grau de dependência linear entre duas variáveis aleatórias. O seu valor é um número entre -1 e 1. Um resultado próximo de 1 indica um alto índice de dependência linear, próximo de 0 um baixo índice de dependência linear e próximo de -1 um índice alto de dependência linear inversa.

Representando os valores $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ de uma amostra em um gráfico de dispersão podemos interpretar geometricamente o grau de dependência linear entre as variáveis aleatórias (X, Y). A figura a seguir ilustra alguns gráficos de dispersão e os seus coeficientes de correlação.

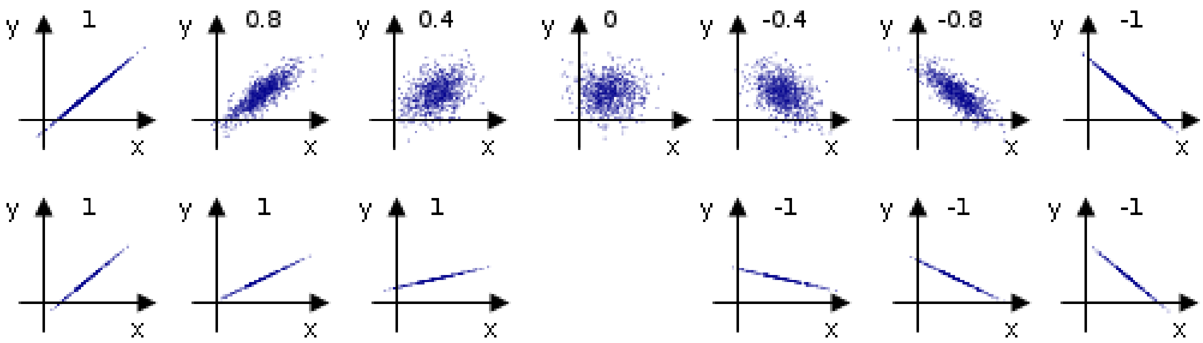


Figura 3.2: Gráficos de dispersão e seus coeficientes de correlação

Note que o gráfico de dispersão:

- é praticamente uma reta para coeficientes 1 e -1;
- se aproxima de uma reta para coeficientes $0,8$ e $-0,8$ e
- não lembra um reta para coeficientes $0,4$; $-0,4$ e 0.

Exemplo 3.23 A tabela a seguir mostra as notas de dez alunos nas disciplinas de Matemática e Física e as médias (valores esperados) de cada disciplina.

Aluno	Matemática (X)	Física (Y)
A	3,0	3,5
B	7,0	6,5
C	5,5	7,0
D	6,5	5,0
E	4,5	3,5
F	7,0	8,5
G	2,5	3,5
H	8,5	9,0
I	6,0	5,0
J	1,5	1,5
Valor esperado	$\overline{X} = 5,2$	$\overline{Y} = 5,3$

Considerando cada dupla de notas $(X; Y)$ como um par ordenado e representando esses pares ordenados em um mesmo plano cartesiano, temos um gráfico de dispersão que se aproxima de um reta, isso é um indicativo de que existe um bom nível de correlação entre as notas de Matemática e Física.

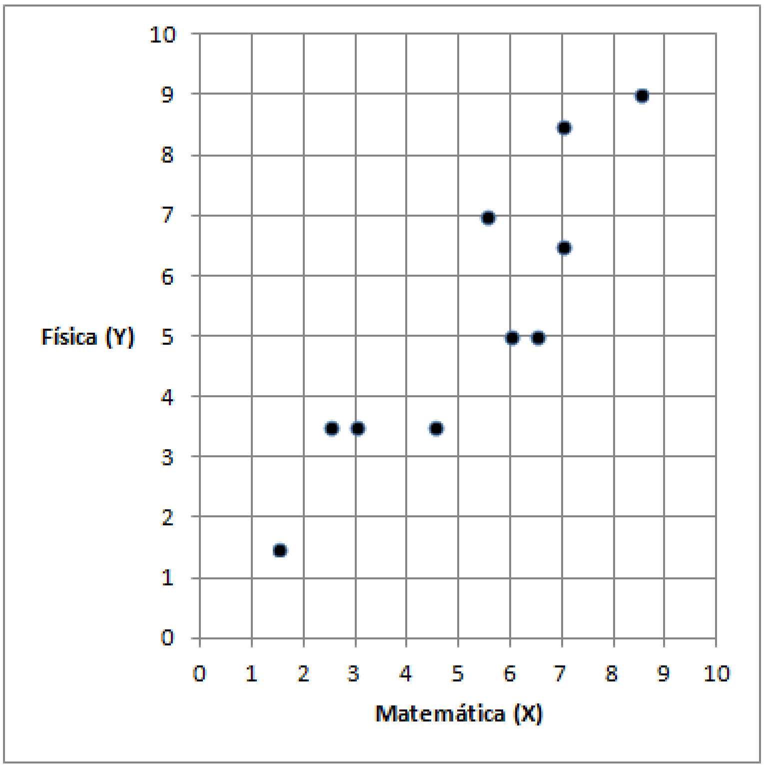


Figura 3.3: Gráfico de dispersão das notas

Para calcular o coeficiente de correlação entre essas notas são necessários os seguintes cálculos.

$X_i - \overline{X}$	$Y_i - \overline{Y}$	$(X_i - \overline{X})(Y_i - \overline{Y})$	$(X_i - \overline{X})^2$	$(Y_i - \overline{Y})^2$
-2,2	-1,8	3,96	4,84	3,24
1,8	1,2	2,16	3,24	1,44
0,3	1,7	0,51	0,09	2,89
1,3	-0,3	-0,39	1,69	0,09
-0,7	-1,8	1,26	0,49	3,24
1,8	3,2	5,76	3,24	10,24
-2,7	-1,8	4,86	7,29	3,24
3,3	3,7	12,21	10,89	13,69
0,8	-0,3	-0,24	0,64	0,09
-3,7	-3,8	14,06	13,69	14,44
	$\sum_1^n$	44,15	46,1	52,6

$$\sqrt{\sum_{1}^{n}\left(X_i-\overline{X}\right)^2\sum_{1}^{n}\left(Y_i-\overline{Y}\right)^2} = \sqrt{46,1\cdot 52,6} = \sqrt{2424,86} \approx 49,24$$

$$\rho_{(X,Y)} = \frac{44,15}{49,24} \approx 0,8966$$

O coeficiente de correlação é próximo de 1, então existe um alto índice de dependência linear entre as notas de Matemática e Física, isto é, se um aluno tem nota alta em uma das disciplinas, maior é a probabilidade dele ter uma nota alta na outra disciplina. Se ele tem uma nota baixa em uma das disciplinas, maior é a probabilidade dele ter uma nota baixa na outra disciplina.

◂

É importante ressaltar que correlação entre duas variáveis aleatórias não significa uma relação de casualidade. Em outras palavras, só porque dois eventos A e B ocorrem de modo similar, não podemos afirmar que A causa B, ou que B causa A, pois pode existir um terceiro fator C, por exemplo, que implica na ocorrência tanto de A quanto de B.

3.4 Método dos Mínimos Quadrados

O coeficiente de correlação, apresentado na seção anterior, é uma medida de dependência linear entre duas variáveis aleatórias (X, Y). Se existe uma forte correlação entre duas variáveis aleatórias (X, Y), o método de mínimos quadrados estabelece uma correspondência entre elas permitindo quantificar o valor de uma variável quando ocorre uma mudança na outra. Para compreender esse método, vamos analisar a seguinte situação.

Os dados a seguir são as notas de vinte alunos em duas avaliações diferentes de Matemática, a primeira avaliação ocorreu após um mês e meio de aula e a segunda avaliação após três meses de aula. As variáveis X e Y representam, respectivamente, as notas da primeira e da segunda avaliação.

X	Y	X	Y
5,5	2,7	8,0	6,2
4,8	2,8	6,6	4,8
8,6	8,1	5,1	3,0
9,4	7,2	3,1	2,4
9,8	8,2	8,9	7,8
8,2	7,0	8,0	5,2
9,8	9,7	7,2	5,8
8,7	8,3	8,8	8,1
70	6,1	7,4	4,8
5,8	3,0	9,8	8,8

O coeficiente de correção $\rho_{(X,Y)}$ entre as duas notas é:

$$\rho_{(X,Y)} = \frac{\sum_1^n \left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_1^n \left(X_i - \overline{X}\right)^2 \sum_1^n \left(Y_i - \overline{Y}\right)^2}} \approx 0,9378$$

Portanto, existe uma dependência linear alta entre as notas das duas avaliações, pois o valor do coeficiente é próximo de 1. Vamos imaginar que um aluno tirou 7,1 na primeira avaliação, $(X = 7,1)$, e não fez a segunda Y, como existe uma forte correlação, pelo método dos mínimos quadrados, é possível estimar qual seria o valor esperado da sua nota Y. Para isso, vamos admitir que a variável Y depende, entre outras coisas, do valor de X, e que:

$$Y = \alpha \cdot X + \beta + \epsilon$$

Onde α e β são constantes reais desconhecidas, X é a nota conhecida da primeira avaliação e ϵ é uma variável aleatória. Nesse modelo linear, estamos admitindo que o resultado da segunda avaliação Y é aleatório cujo valor pode ser decomposto em uma componente ϵ estritamente aleatória mais um termo $(\alpha \cdot X + \beta)$ que depende da nota da primeira avaliação.

Para que a variável aleatória ϵ não dependa da variável X, adotaremos a hipótese que, para todo X, o valor esperado de ϵ é zero, $E(\epsilon) = 0$, e sua variância é σ^2, $V(\epsilon) = \sigma^2$. Consequentemente, temos:

$$E(Y) = \alpha \cdot X + \beta$$

Definição 3.24. *Seja* $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$ *uma amostra aleatória e a relação* $E(Y) = \alpha \cdot X + \beta$*, em que* α *e* β *são constantes reais. As estimativas de mínimos quadrados dos parâmetros* α *e* β *são os valores que tornam mínima a expressão:*

$$\sum_1^n [Y_i - (\alpha \cdot X_i + \beta)]^2$$

Para estimar os valores de α e β, consideramos a seguinte função.

$$S(\alpha, \beta) = \sum_1^n [Y_i - (\alpha \cdot X_i + \beta)]^2$$

Logo, o valor mínimo de S em função α ocorre quando:

$$\text{I)}\frac{dS}{d\alpha} = 0$$

Do mesmo modo, o valor mínimo de S em função β ocorre quando:

$$\text{II)}\frac{dS}{d\beta} = 0$$

Derivando S em relação à α, obtemos:

$$\begin{aligned}\frac{dS}{d\alpha} &= \sum_1^n 2\left[Y_i - (\alpha \cdot X_i + \beta)\right](X_i)\\ &= -2\sum_1^n \left(X_i \cdot Y_i - \alpha \cdot X_i^2 - \beta \cdot X_i\right)\end{aligned}$$

Derivando S em relação à β, obtemos:

$$\begin{aligned}\frac{dS}{d\beta} &= \sum_1^n 2\left[Y_i - (\alpha \cdot X_i + \beta)\right](-1)\\ &= -2\sum_1^n (Y_i - \alpha \cdot X_i - \beta)\end{aligned}$$

Consequentemente,

$$\text{I})\frac{dS}{d\alpha} = 0 \Rightarrow \alpha\sum_1^n X_i^2 + \beta\sum_1^n X_i = \sum_1^n X_i Y_i$$

$$\text{II})\frac{dS}{d\beta} = 0 \Rightarrow \alpha\sum_1^n X_i + n\beta = \sum_1^n Y_i$$

As soluções do sistema linear composta pelas equações I e II são:

$$\widehat{\alpha} = \frac{\sum_1^n Y_i(X_i - \overline{X})}{\sum_1^n (X_i - \overline{X})^2}, \text{ onde: } \overline{X} = \frac{1}{n}\sum_1^n X_i$$

$$\widehat{\beta} = \overline{Y} - \widehat{\alpha}\overline{X}, \text{ onde: } \overline{Y} = \frac{1}{n}\sum_1^n Y_i$$

Retornando a situação inicial das duas avaliações X e Y de Matemática, para estimar os coeficientes α e β da relação $E(Y) = \alpha \cdot X + \beta$, efetuamos:

X_i	Y_i	$X_i - \overline{X}$	$(X_i - \overline{X})^2$	$Y_i(X_i - \overline{X})$
5,5	2,7	-2,025	4,100625	-5,4675
4,8	2,8	-2,725	7,425625	-7,63
8,6	8,1	1,075	1,155625	8,7075
9,4	7,2	1,875	3,515625	13,50
9,8	8,2	2,275	5,175625	18,655
8,2	7,0	0,675	0,455625	4,725
9,8	9,7	2,275	5,175625	22,0675
8,7	8,3	1,175	1,380625	9,7525
7,0	6,1	-0,525	0,275625	-3,2025
5,8	3,0	-1,725	2,975625	-5,175
8,0	6,2	0,475	0,225625	2,945
6,6	4,8	-0,925	0,855625	-4,44
5,1	3,0	-2,425	5,880625	-7,275
3,1	2,4	-4,425	19,580625	-10,62
8,9	7,8	1,375	1,890625	10,725
8,0	5,2	0,475	0,225625	2,47
7,2	5,8	-0,325	0,105625	-1,885
8,8	8,1	1,275	1,625625	10,3275
7,4	4,8	-0,125	0,015625	-0,60
9,8	8,8	2,275	5,175625	20,02
$\overline{X} = 7,525$	$\overline{Y} = 6,00$		$\sum_1^n = 67,2175$	$\sum_1^n = 77,60$

$$\widehat{\alpha} = \frac{\sum_1^n Y_i(X_i - \overline{X})}{\sum_1^n (X_i - \overline{X})^2} = \frac{77,60}{67,2175} \approx 1,1545$$

$$\widehat{\beta} = \overline{Y} - \widehat{\alpha}\overline{X} = 6,00 - 1,1545 \cdot 7,525 \approx -2,6873$$

Portanto, $E(Y) = 1,1545 \cdot X - 2,6873$. Desse modo, a estimativa do valor esperado da nota da segunda avaliação do aluno que tirou 7,1 na primeira é:

$$E(Y) = 1,1545 \cdot 7,1 - 2,6873 \approx 5,51$$

A figura apresenta em um mesmo plano cartesiano o gráfico de dispersão das notas das duas avaliações e o gráfico da função afim que foi determinado pelo método dos mínimos quadrados para relacionar essas duas notas.

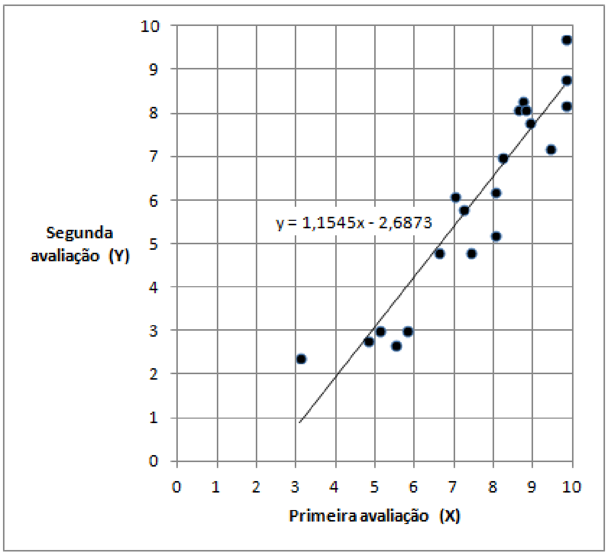


Figura 3.4: Método dos Mínimos Quadrados

4 Teoria de Resposta ao Item

Começaremos com algumas ideias gerais e explicando de modo geral o que é a Teoria de Resposta ao Item (TRI).

A altura de uma pessoa é uma característica explícita que quantificamos de forma direta com auxílio de um instrumento de medida, por exemplo, uma fita métrica, e uma escala com unidade de medida padronizada, como o metro ou o centímetro. Diferente da altura, a ansiedade, o nível de estresse, as habilidades cognitivas ou a proficiência em uma determinada área do conhecimento humano são características implícitas denominadas de traços latentes de um ser humano.

Não podemos comparar um traço latente entre dois seres humanos da mesma forma que comparamos suas alturas. Quando colocamos duas pessoas lado a lado, identificamos qual delas é a mais alta, mas não identificamos qual delas é a mais ansiosa. Além disso, colocar duas pessoas lado a lado não é suficiente para quantificar a diferença entre suas alturas. Para quantificar essa diferença, é necessário um sistema de comparação constituído de uma escala com uma unidade de medida adotada e um instrumento de medição.

A Teoria de Resposta ao Item é um conjunto de modelos matemáticos (probabilísticos e estatísticos) que possibilita a construção de um sistema de comparação de traços latentes. A construção desse sistema de comparação é feita por meio de variáveis observáveis relacionadas ao traço latente. Por exemplo, poderíamos construir um sistema de comparação e estimar a altura dos indivíduos de uma população solicitando para cada um responder apenas sim ou não às seguintes questões:

1. Uso roupas do tamanho pequeno?
2. Uso roupas do tamanho grande?
3. Sempre compro calçados de numeração abaixo de 36?
4. Sempre compro calçados de numeração acima de 42?
5. Numa fila por ordem de tamanho, você sempre é um dos últimos?
6. Numa fila por ordem de tamanho, você sempre é um dos primeiros?
7. Você acha que se daria bem em um time de vôlei?
8. Você tem dificuldade de se acomodar no ônibus?

9. Você consegue pegar um objeto no alto de um armário sem utilizar um banco ou uma escada?
10. Eu literalmente olho para meus colegas de baixo para cima?

De modo parecido, poderíamos criar um sistema de comparação e estimar o nível de ansiedade de cada indivíduo de uma população pedindo para cada um responder apenas sim ou não às seguintes perguntas:

1. Constantemente está preocupado ou tenso com alguma coisa?
2. Frequentemente tem cólicas intestinais?
3. Você tem facilidade para dormir?
4. Constantemente tem taquicardia ou palpitações?
5. Raramente fica nervoso?
6. Tem facilidade em permanecer concentrado?
7. Em uma conversa, você geralmente fala mais do que o outro?
8. Em uma conversa, você geralmente escuta mais do que o outro?
9. Você acredita que algo ruim vai acontecer se certas coisas não forem feitas de uma determinada maneira?
10. Frequentemente tem falta de ar ou fica ofegante?

Os dois questionários não têm valor científico, eles ilustram o que são variáveis observáveis de uma característica e, ilustram também, o pressuposto da unidimensionalidade da TRI. A unidimensionalidade é a hipótese de que existe uma habilidade ou uma proficiência dominante (traço latente) responsável pelas respostas ao conjunto de peguntas. No primeiro questionário, a altura seria a característica dominante responsável por cada uma das respostas e, no segundo, a ansiedade seria a aptidão dominante responsável pelas respostas.

Os modelos matemáticos da TRI relacionam as variáveis observáveis (as respostas dadas aos itens de um teste) com a aptidão não observável (o traço latente). De acordo com essa relação, cria-se um sistema de comparação e estima-se a aptidão de cada indivíduo. O exemplo a seguir mostra de modo intuitivo e simplificado como poderíamos prever qual pessoa é a mais alta analisando somente as variáveis observáveis.

Exemplo 4.1

Imagine que duas pessoas, A e B, responderam o questionário das alturas.

1. Uso roupas do tamanho pequeno? A Sim - B Não
2. Uso roupas do tamanho grande? A Não - B Sim
3. Sempre compro calçados de numeração abaixo de 36? A Sim - B Não
4. Sempre compro calçados de numeração acima de 42? A Não - B Sim
5. Numa fila por ordem de tamanho, você sempre é um dos últimos? A Não - B Sim
6. Numa fila por ordem de tamanho, você sempre é um dos primeiros? A Sim - B Não
7. Você acha que se daria bem em um time de vôlei? A Sim - B Sim
8. Você tem dificuldade de se acomodar no ônibus? A Não - B Não
9. Você consegue pegar um objeto no alto de um armário sem utilizar um banco ou uma escada? A Não - B Sim
10. Eu literalmente olho para meus colegas de baixo para cima? A Não - B Não

O sujeito A usa roupas do tamanho pequeno, compra calçados abaixo de 36, numa fila por ordem de tamanho sempre é um dos primeiros e não consegue pegar objetos no alto de um armário sem utilizar um banco ou uma escada, enquanto o sujeito B usa roupas do tamanho grande, compra calçados acima de 42, numa fila por ordem de tamanho sempre é um dos últimos e consegue pegar objetos no alto de um armário sem utilizar um banco ou uma escada, supondo que a altura foi a característica dominante responsável por essas respostas, o mais provável então é que o sujeito B seja mais alto do que o sujeito A.

◂

A palavra Item na TRI significa questão ou pergunta e não as alternativas de resposta. A TRI estuda cada item individualmente, ela fornece um modelo para representar a probabilidade de um indivíduo dar uma certa resposta a um item em função do traço latente em estudo e dos parâmetros do item. Por exemplo, para cada pergunta do questionário da altura, a TRI representaria a probabilidade de responder sim em função da altura e dos parâmetros da própria pergunta.

Uma vez definida essa função, dado um conjunto de respostas, são estimados os parâmetros de cada item, bem como, a aptidão em estudo de cada respondente. Por exemplo, suponha que no questionário da ansiedade foi definida para cada item a relação que representa a probabilidade de responder sim em função da ansiedade e dos parâmetros do item. Se um grupo de pessoas responder esse questionário, em um primeiro momento, os parâmetros de cada item serão estimados, depois, o grau de ansiedade de cada respondente.

Existem modelos diferentes da TRI que consideram o tipo de questão (falso ou verdadeiro, múltipla escolha, dissertativa) e os parâmetros dos itens que se deseja investigar. Neste capítulo será descrito o modelo logístico unidimensional de três parâmetros utilizado em avaliações educacionais como o Exame Nacional do Ensino Médio (Enem) e o Sistema de Avaliação da Educação Básica (SAEB).

4.1 Modelo Logístico de Três Parâmetros

O modelo matemático mais utilizado da TRI em avaliações educacionais de larga escala é o modelo logístico unidimensional de três parâmetros. Na maioria dessas avaliações, os itens que as compõem são de múltipla escolha com uma única alternativa correta, ou seja, entre as opções de resposta apresentadas para cada questão, o aluno deve selecionar apenas uma.

Um aluno pode acertar uma questão de múltipla escolha por sorte, por isso um dos três parâmetros considerados é o acerto ao acaso (chute). Os outros dois são a dificuldade e a discriminação do item. Desse modo, a TRI analisa separadamente cada item que compõem uma avaliação estimando:

- o seu poder de diferenciar, discriminar, alunos que possuem ou não uma proficiência;
- o seu grau de dificuldade e;
- a probabilidade de alunos sem proficiência, habilidade mínima necessária, acertarem ao acaso.

Nesse caso, o modelo logístico unidimensional de três parâmetros define a probabilidade de um indivíduo j responder corretamente um item i de múltipla escolha pela função:

$$\mathbb{P}\left(X_{ji} = 1|\theta_j\right) = c_i + \frac{(1 - c_i)}{1 + \exp\left[-Da_i\left(\theta_j - b_i\right)\right]}$$

no qual:

- $\mathbb{P}\left(X_{ji} = 1|\theta_j\right)$ é a probabilidade condicional do indivíduo j responder corretamente o item i sabendo que sua habilidade é θ_j. Ele também é identificado como a proporção de respostas corretas ao item i dos indivíduos com habilidade igual a θ_j.
- X_{ji} é a resposta do indivíduo j ao item i ($X_{ji} = 1$ se a resposta for correta e $X_{ji} = 0$ se for incorreta);
- θ_j é a proficiência, a habilidade (traço latente), do indivíduo j;
- c_i é o parâmetro de acerto ao acaso do item i;
- $\exp$ é a função exponencial;
- D é um fator de escala, que é igual a 1 na métrica logística e igual a 1,7 na métrica normal.
- a_i é o parâmetro de discriminação do item i;

- b_i é o parâmetro de dificuldade do item i.

Observe que a probabilidade $\mathbb{P}(X = 1|\theta)$ de um sujeito acertar o item está em função do seu traço latente θ e dos parâmetros a, b e c do item. Uma propriedade importante dos parâmetros a, b e c é que eles são invariantes, isto é, suas estimativas não dependem da amostra utilizada. Em outras palavras, eles são uma características intrínseca do próprio item.

A probabilidade $\mathbb{P}(X = 1|\theta)$ e o parâmetro c de acerto ao acaso são medidos em uma mesma escala de 0 a 1. Em uma outra escala cuja unidade de medida é 1 desvio padrão são medidos a proficiência θ e o parâmetro b de dificuldade do item.

4.1.1 Curva Característica do Item (CCI)

A representação gráfica da função definida pela TRI é denominada de Curva Característica do Item (CCI).

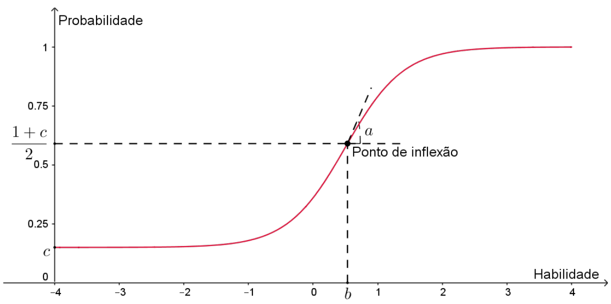


Figura 4.1: Curva Característica do Item (CCI)

A CCI é uma curva em forma de "S"com duas assíntotas horizontais e um ponto de inflexão (ponto onde a curva muda de concavidade), em que:

- o eixo horizontal indica os valores da habilidade θ na escala padronizada de média igual a 0 e desvio padrão 1.
- o eixo vertical indica a probabilidade $\mathbb{P}(\theta)$, em uma escala de 0 a 1, de um indivíduo com habilidade θ responder corretamente o item ;
- o parâmetro c de acerto ao acaso é o valor de $\mathbb{P}(\theta)$ dado pelo ponto em que assíntota inferior horizontal da curva intersecta o eixo vertical.
- o parâmetro b de dificuldade do item é o valor da habilidade θ correspondente a probabilidade de acerto igual a $\frac{1+c}{2}$, ou seja:

$$\mathbb{P}(\theta = b) = \frac{1+c}{2}$$

Desse modo, para localizar o valor de b na CCI, basta determinar no eixo horizontal o valor de θ correspondente a probabilidade $\frac{1+c}{2}$.

- o parâmetro a de discriminação do item é diretamente proporcional à inclinação da curva, em relação ao eixo horizontal, no seu ponto de inflexão.

Observe na curva característica do item (CCI) que, quanto maior é a habilidade θ, maior é a probabilidade $\mathbb{P}(\theta)$ de responder o item corretamente. Além disso, a forma como a curva se apresenta depende dos parâmetros a, b e c do item representado.

4.1.2 Acerto ao Acaso

Em uma questão de múltipla escolha, existe a possibilidade do respondente assinalar a alternativa correta por sorte (chute), principalmente em questões mais difíceis nas quais indivíduos de baixa aptidão arriscam qualquer uma. A TRI estima a probabilidade disso ocorrer por meio do parâmetro c de acerto ao acaso. Se fosse proibido "chutar", o valor de c seria zero ($c = 0$).

Para questões com cinco alternativas, o valor esperado de c deve ser no máximo $0,20$. Caso o valor seja muito superior, isso pode ser um indicativo de que há um equívoco na formulação do item, fazendo com que indivíduos de baixa habilidade assinalem a opção correta.

O valor de c é dado quando a habilidade θ em estudo tende a $-\infty$, ou seja:

$$\lim_{\theta \longrightarrow -\infty} \mathbb{P}(X = 1|\theta) = c$$

Na prática, o parâmetro c quantifica a probabilidade que um indivíduo, sem a habilidade mínima necessária, tem de acertar a questão.

Na curva característica do item, o parâmetro c corresponde a probabilidade indicada pelo ponto de intersecção entre a assintota inferior da curva e o eixo das probabilidades. A figura a seguir ilustra os gráficos das curvas características dos itens α (linha cheia) e β (linha tracejada). Comparando essas curvas, podemos observar que a assíntota inferior da curva α intersecta o eixo das probabilidades em um ponto de coordenada menor do que a assíntota de β, logo a probabilidade de acerto ao acaso do item α é menor do que β, ou seja:

$$c_\alpha < c_\beta$$

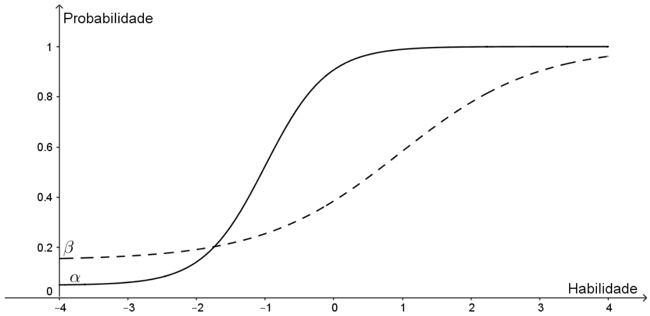


Figura 4.2: Curvas Características dos Itens α e β

4.1.3 Dificuldade do Item

A dificuldade do item é representado pelo parâmetro b. Ele indica o nível mínimo de proficiência θ que um sujeito precisa possuir para ter uma probabilidade de $\frac{1+c}{2}$ de solucionar corretamente a questão, onde c é o valor do parâmetro de acerto ao acaso, ou seja:

$$\mathbb{P}(b = \theta) = \frac{1 + c}{2}$$

Na curva característica do item, os valores de b e de $\frac{1+c}{2}$ são as coordenadas do ponto de inflexão.

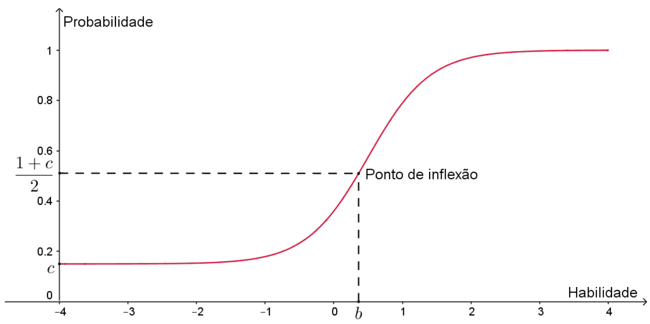


Figura 4.3: Parâmetro b de dificuldade do item

Observe que $\frac{1+c}{2}$ é a média das probabilidades c e 1, logo o grau de dificuldade da questão é definido como o valor da habilidade em que a probabilidade de acertar é igual a probabilidade de errar. Desse modo, quando a proficiência θ de um indivíduo coincidir com o valor do parâmetro b do item que ele está resolvendo ($\theta = b$), a probabilidade dele acertar o item é igual a probabilidade dele errar, veja:

$$\begin{aligned}
\mathbb{P}\left(X = 1|\theta = b\right) &= c + \frac{(1-c)}{1+\exp\left[-Da\left(\theta - b\right)\right]} \\
&= c + \frac{(1-c)}{1+\exp\left[-Da\left(b - b\right)\right]} \\
&= c + \frac{(1-c)}{1+\exp 0} \\
&= c + \frac{(1-c)}{2} \\
&= \frac{1+c}{2}
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}\left(X = 0|\theta = b\right) &= 1 - \mathbb{P}\left(X = 1|\theta = b\right) \\
&= 1 - \frac{1+c}{2} \\
&= \frac{1+c}{2}
\end{aligned}$$

Se fosse proibido "chutar", o valor de c seria zero ($c = 0$) e, consequentemente, o parâmetro b seria o valor da habilidade θ correspondente a probabilidade de 50%, veja:

$$\mathbb{P}\left(X = 1|\theta = b\right) = \frac{1+c}{2} = \frac{1+0}{2} = \frac{1}{2} = 50\%$$

Podemos comparar o grau de dificuldade de dois itens observando as coordenadas do ponto de inflexão em relação ao eixo das habilidades. Os gráficos a seguir são as curvas características dos itens α (linha cheia) e γ (linha pontilhada). Observe que, em relação ao eixo das habilidades, o ponto de inflexão da curva α tem coordenada menor do que o ponto de inflexão da curva γ, portanto o nível de dificuldade do item α é menor do que γ, isto é:

$$b_\alpha < b_\gamma$$

O parâmetro b é medido em uma escala padronizada em que o número 0, origem da escala, representa a proficiência média dos indivíduos submetidos ao teste e a unidade de medida é 1 desvio padrão. Teoricamente, a escala varia de $-\infty$ a $+\infty$. Na prática, varia de -3 a 3, intervalo onde estão cerca de $99,74\%$ dos casos. Valores de b próximos a -3 indicam questões fáceis e próximos de 3 difíceis. Um valor de b fora do intervalo de -3 a 3 sugere um item com problema de elaboração.

O nível de dificuldade das questões que compõem uma avaliação depende da sua finalidade. Em uma avaliação educacional em larga escala, é recomendada a distribuição de níveis de dificuldade dentro de uma curva normal. A tabela a seguir mostra a distribuição e a classificação adotada pela maioria dos autores.

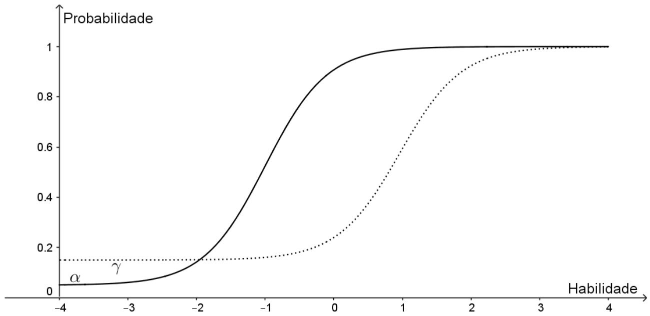


Figura 4.4: Curvas Características dos Itens α e γ

Classificação	Valores de b	$\%esperado$
Muito fáceis	até $-1,27$	10%
Fáceis	de $-1,27$ a $-0,51$	20%
Médias	de $-0,51$ a $0,51$	40%
Difíceis	de $0,51$ a $1,27$	20%
Muito difíceis	acima de $1,27$	10%

Nas provas de larga escala de Matemática, é difícil atender às recomendações da tabela porque as questões geralmente têm um índice alto de dificuldade.

A característica mais importante do parâmetro b é que tanto ele quanto a habilidade θ (traço latente) são medidos na mesma escala, isso permite uma interpretação pedagógica da proficiência em estudo. Esse assunto será explicado na seção Escala de Proficiência.

4.1.4 Discriminação do Item

A discriminação é a capacidade que o item tem de distinguir pequenas diferenças de habilidades. Um item com alto poder de discriminação pode diferenciar indivíduos com habilidades próximas do grau de dificuldade do item.

O nível de discriminação de um item é indicado pelo parâmetro a e o seu valor é proporcional a inclinação da curva característica desse item no seu ponto de inflexão em relação ao eixo horizontal.

Observe que quanto maior a inclinação da curva no seu ponto de inflexão, maior será o valor do parâmetro a e, quanto menor a inclinação da curva, menor o valor de a.

Não são esperados valores negativos para esse parâmetro, uma vez que eles indicariam que a probabilidade de acerto do item diminui com o aumento da habilidade do respondente. Quando

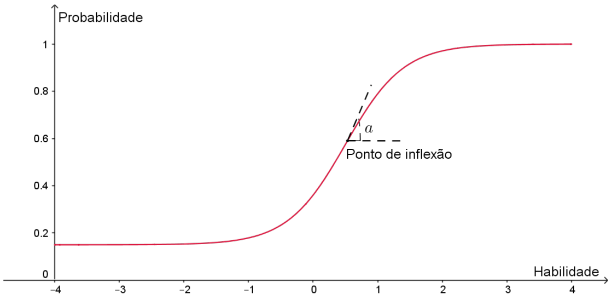


Figura 4.5: Parâmetro $\mathfrak{a}$ de discriminação do item

isso ocorre, o item é descartado. Valores de $\mathfrak{a}$ próximos de zero indicam um item com pouco poder de discriminação. Nesse caso, a CCI seria mais parecida com uma reta e não um S.

Exemplo 4.2 As figuras as seguir são as curvas características dos itens $\mathfrak{i}_1$ e $\mathfrak{i}_2$, $\mathfrak{i}_1$ com parâmetro de discriminação $\mathfrak{a}_1 = 0,3$ e $\mathfrak{i}_2$ com parâmetro de discriminação $\mathfrak{a}_2 = 1,7$.

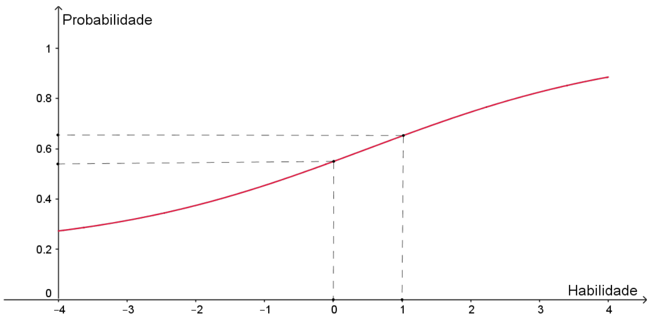


Figura 4.6: CCI com $\mathfrak{a}_1 = 0,3$

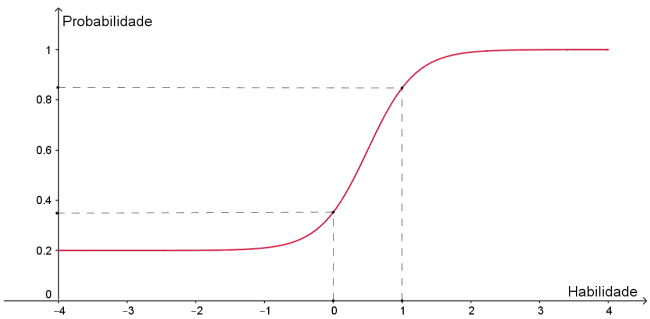


Figura 4.7: CCI com $\mathfrak{a}_2 = 1,7$

Suponha que dois alunos p e k, p com habilidade $\theta_p = 0$ e k com habilidade $\theta_k = 1$, responderam esses dois itens. Logo, pelas curvas características, as probabilidades aproximadas de cada um ter acertado são:

Item	Aluno p	Aluno k
i_1 $(a_1 = 0,3)$	$\mathbb{P}(\theta_p = 0) \approx 0,55$	$\mathbb{P}(\theta_k = 1) \approx 0,65$
i_2 $(a_2 = 1,7)$	$\mathbb{P}(\theta_p = 0) \approx 0,35$	$\mathbb{P}(\theta_k = 1) \approx 0,85$

Os dados da tabela ressaltam a capacidade que o item i_2 tem para diferenciar as habilidades dos alunos p e k. Enquanto no item i_1, p e k apresentam probabilidades de acerto acima de 50%, no item i_2, o mais provável é que o aluno p erre e o aluno k acerte a questão.

◂

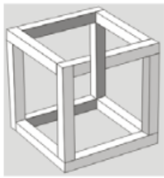
Geralmente, apenas itens que apresentam valores de discriminação superiores a $0,70$ são considerados. Alguns autores categorizam os itens de acordo com a seguinte tabela.

Valores	Discriminação
$a = 0,00$	Nenhuma
$0,00 < a \leqslant 0,35$	Muito baixa
$0,35 < a \leqslant 0,65$	Baixa
$0,65 < a \leqslant 1,35$	Moderada
$1,35 < a \leqslant 1,70$	Alta
$a > 1,70$	Muito alta

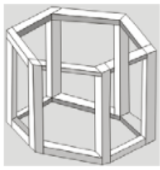
Exemplo 4.3 A questão deste exemplo foi aplicada no ENEM de 2007.

Representar objetos tridimensionais em uma folha de papel nem sempre é tarefa fácil. O artista holandês Escher (1898-1972) explorou essa dificuldade criando várias figuras planas impossíveis de serem construídas como objetos tridimensionais, a exemplo da litografia **Belvedere**, reproduzida ao lado.

Considere que um marceneiro tenha encontrado algumas figuras supostamente desenhadas por Escher e deseje construir uma delas com ripas rígidas de madeira que tenham o mesmo tamanho. Qual dos desenhos a seguir ele poderia reproduzir em um modelo tridimensional real?

A 

B 

C 

D 

E 

Matemática - Faixa: 461.7

O gráfico a seguir é a curva característica desta questão. Os valores dos seus parâmetros estimados pela TRI foram:

- discriminação do item: $a = 1,31$;
- dificuldade do item: $b = 0,96$ e
- acerto ao acaso: $c = 0,15$.

A questão apresentou uma boa capacidade de discriminação, um grau elevado de dificuldade e a alternativa correta não atraiu indivíduos com baixa proficiência. ◂

4.2 Escala de Proficiência

Construir uma escala exige a definição de uma origem e uma unidade de medida. Nas escalas de proficiências:

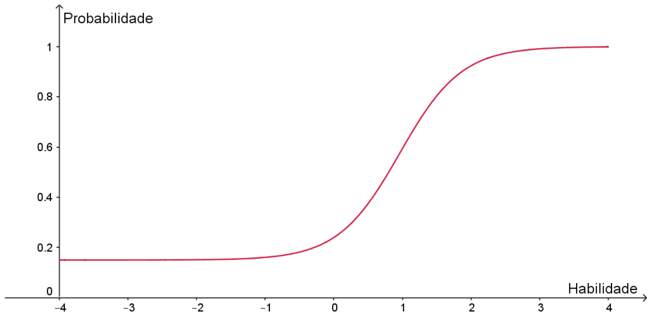


Figura 4.8: Curva característica do item

- a origem é a média das proficiências dos indivíduos submetidos ao teste e;
- a unidade de medida é o desvio padrão.

Os gráficos contidos neste capítulo apresentaram o eixo das habilidades em uma escala padronizada com origem no 0 e unidade de medida 1 desvio padrão, escala comumente representada por $(0,1)$. Logo, um indivíduo com habilidade $1,5$ está a $1,5$ desvio padrão acima da média e, do mesmo modo, um indivíduo com habilidade $-0,5$ está a $0,5$ desvio padrão abaixo da média.

Teoricamente, a escala $(0,1)$ varia de $-\infty$ a $+\infty$. Na prática, no intervalo de:

- $-0,5$ a $0,5$ estão cerca de 38% dos casos;
- $-1,0$ a $1,0$ estão cerca de 68% dos casos;
- $-1,5$ a $1,5$ estão cerca de 86% dos casos;
- $-2,0$ a $2,0$ estão cerca de 95% dos casos;
- $-2,5$ a $2,5$ estão cerca de 99% dos casos e;
- -3 a 3 estão cerca de $99,7\%$ dos casos.

Desse modo, uma habilidade $\theta = 2,1$ indica uma aptidão muito acima da média e, da mesma forma, uma habilidade $\theta = -2,1$ indica uma aptidão muito abaixo da média, pois apenas 5% dos casos aproximadamente apresentam habilidades fora do intervalo de -2 a 2.

Exemplo 4.4 Em 2009, o ENEM criou sua escala de proficiência com média igual a 500 e desvio padrão 100. Nesse caso, o valor 500 representa a proficiência média dos concluintes do ensino médio da rede pública de 2009 que realizaram esse exame e o valor 100 representa a unidade de medida de 1 desvio padrão nessa escala. Logo, uma nota 650 no ENEM indica uma habilidade de $1,5$ desvio padrão acima da média dos alunos da rede pública que prestaram o exame em 2009.

Observe também, na distribuição a seguir, que $98,57\%$ dos participantes de 2009 tiveram proficiências entre -2 e 2 desvios padrão em relação à média, ou seja, notas entre 300 e 700.

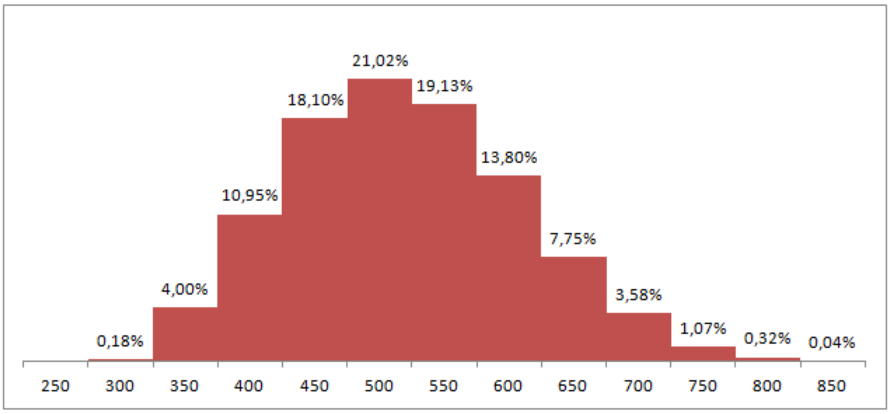


Figura 4.9: Distribuição do ENEM 2009

◂

Na TRI, tanto o parâmetro b de dificuldade de um item (questão) quanto a habilidade θ de um indivíduo (participante) são medidos na mesma escala de proficiência. Isso possibilita uma interpretação pedagógica da escala.

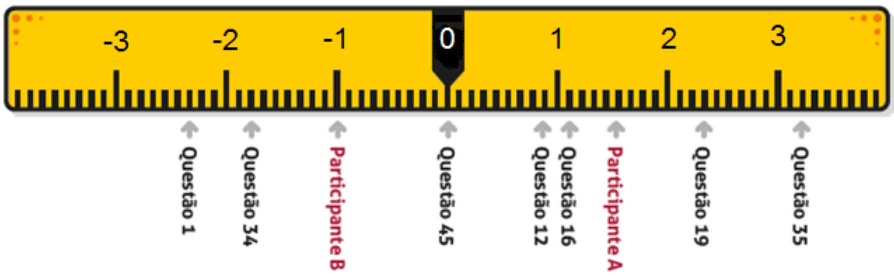


Figura 4.10: Escala de proficiência

A interpretação pedagógica é uma análise qualitativa que consiste na descrição de habilidades de cada faixa da escala. Por meio dessa descrição, é possível verificar quais habilidades um estudante demonstrou ter desenvolvido e quais precisa desenvolver.

A metodologia executada para interpretação pedagógica de uma escala de proficiência consiste em dois procedimentos básicos. O primeiro é a identificação de itens representativos em cada faixa da escala, denominados de itens âncoras. O segundo é a apresentação dos itens âncoras para um grupo de especialistas interpretarem pedagogicamente cada um desses itens.

O critério geralmente adotado para identificar um item âncora de um nível α é:

- 65% ou mais dos estudantes com habilidades em torno de α acertam o item;
- menos de 65% dos estudantes posicionados no nível anterior acertam o item;
- o ajuste da curva característica é bom.

Os especialistas recebem os itens âncoras separados por níveis, cada um com seu enunciado completo, suas estatísticas, sua curva característica e um gráfico da distribuição por alternativas da proporção de acertos por faixa da escala. De posses dessas informações, eles fazem a descrição de habilidades de cada nível da escala.

Exemplo 4.5 A título de ilustração, a descrição de alguns níveis das escalas de proficiências do ENEM das áreas: Matemática e suas Tecnologias; Linguagens, Códigos e suas Tecnologias; Ciências Humanas e suas Tecnologias e Ciências da Natureza e suas Tecnologias. Para descrição completa acesse www.inep.gov.br.

Matemática e suas Tecnologias

Valor	Descrição
586.1	Reconhecer representações de números naturais e informações de orientação em medidores de energia elétrica.
585.7	Determinar utilizando operações fundamentais com números naturais, o valor de um produto.
585.0	Interpretar medidas de posição e dispersão em uma tabela na classificação de corrida de regularidade.

Linguagens, Códigos e suas Tecnologias

Valor	Descrição
682.1	Reconhecer elementos estéticos do cubismo na pintura.
678.8	Reconhecer o uso do pronome de tratamento em situação de comunicação específica em uma crônica.
678.1	Identificar recursos linguísticos persuasivos em propaganda governamental.

Ciências Humanas e suas Tecnologias

Valor	Descrição
711.1	Analisar as implicações da produção de alimentos geneticamente modificados no espaço rural brasileiro contemporâneo.
710.7	Reconhecer as implicações dos movimentos sociais de 1968.
707.1	Identificar a apropriação de uma memória histórica do bandeirantismo pela propaganda da Revolução de 1932.

Ciências da Natureza e suas Tecnologias

Valor	Descrição
466.7	Relacionar a obesidade com a ocorrência de outras doenças em uma tirinha.
465.0	Reconhecer a energia potencial elástica em mecanismos que envolvem conversão de energia.
460.7	Identificar ações que reduzem o impacto ambiental de incineradores de lixo.

◂

Exemplo 4.6 Neste exemplo há quatro itens âncoras das escalas de proficiências do ENEM. O primeiro classificado na faixa 461.7 da escala de Matemática e suas Tecnologias, o segundo classificado na faixa 544.1 da escala de Linguagens, Códigos e suas Tecnologias, o terceiro classificado na faixa 750.0 da escala de Ciências Humanas e suas Tecnologias e o quarto classificado na faixa 630.4 da escala de Ciências da Natureza e suas Tecnologias. Todos os itens âncoras das escalas de proficiências do ENEM estão em www.inep.gov.br. ◂

ENEM 2012

Jogar baralho é uma atividade que estimula o raciocínio. Um jogo tradicional é a Paciência, que utiliza 52 cartas. Inicialmente são formadas sete colunas com as cartas. A primeira coluna tem uma carta, a segunda tem duas cartas, a terceira tem três cartas, a quarta tem quatro cartas, e assim sucessivamente até a sétima coluna, a qual tem sete cartas, e o que sobra forma o monte, que são as cartas não utilizadas nas colunas.

A quantidade de cartas que forma o monte é

(A) 21.

(B) 24.

(C) 26.

(D) 28.

(E) 31.

Figura 4.11: Matemática e suas Tecnologias - Faixa: 461.7

4.3 Estimação dos Parâmetros e da Proficiência

A TRI determina tanto os valores dos parâmetros dos itens quanto os valores das proficiências por meio de técnicas estatísticas de estimação. Existem maneiras diferentes para se fazer isso, ENEM 2012

Aqui é o país do futebol

Brasil está vazio na tarde de domingo, né?
Olha o sambão, aqui é o país do futebol
[...]
No fundo desse país
Ao longo das avenidas
Nos campos de terra e grama
Brasil só é futebol
Nesses noventa minutos
De emoção e alegria
Esqueço a casa e o trabalho
A vida fica lá fora
Dinheiro fica lá fora
A cama fica lá fora
A mesa fica lá fora
Salário fica lá fora
A fome fica lá fora
A comida fica lá fora
A vida fica lá fora
E tudo fica lá fora

SIMONAL, W. **Aqui é o país do futebol**. Disponível em: www.vagalume.com.br. Acesso em: 27 out. 2011 (fragmento).

Na letra da canção Aqui é o país do futebol, de Wilson Simonal, o futebol, como elemento da cultura corporal de movimento e expressão da tradição nacional, é apresentado de forma crítica e emancipada devido ao fato de

(A) reforçar a relação entre o esporte futebol e o samba.

(B) ser apresentado como uma atividade de lazer.

(C) ser identificado com a alegria da população brasileira.

(D) promover a reflexão sobre a alienação provocada pelo futebol.

(E) ser associado ao desenvolvimento do país.

Figura 4.12: Linguagens, Códigos e suas Tecnologias - Faixa: 544.1

dependendo da população pesquisada, da composição dos itens da avaliação e de outras variáveis. Nesta seção, será apresentado apenas as ideias básicas de um desses procedimentos que consiste em estimar primeiro os parâmetros dos itens, supondo conhecidas as habilidades dos respondentes, em seguida, estimar as habilidades dos respondentes.

> "Para ilustrar os passos usados pelo método de Máxima Verossimilhança, suponhamos que X sujeitos tenham respondido a Y itens que compõem um teste qualquer. São desconhecidos tanto os parâmetros dos itens como as habilidades dos

ENEM 2012

A experiência que tenho de lidar com aldeias de diversas nações me tem feito ver, que nunca índio fez grande confiança de branco e, se isto sucede com os que estão já civilizados, como não sucederá o mesmo com esses que estão ainda brutos.

NORONHA, M. Carta a J. Caldeira Brant. 2 jan.1751. Apud CHAIM, M. M. **Aldeamentos indígenas** (Goiás: 1749-1811). São Paulo: Nobel, Brasília: INL, 1983 (adaptado).

Em 1749, ao separar-se de São Paulo, a capitania de Goiás foi governada por D. Marcos de Noronha, que atendeu às diretrizes da política indigenista pombalina que incentivava a criação de aldeamentos em função

(A) das constantes rebeliões indígenas contra os brancos colonizadores, que ameaçavam a produção de ouro nas regiões mineradoras.

(B) da propagação de doenças originadas do contato com os colonizadores, que dizimaram boa parte da população indígena.

(C) do empenho das ordens religiosas em proteger o indígena da exploração, o que garantiu a sua supremacia na administração colonial.

(D) da política racista da Coroa Portuguesa, contrária à miscigenação, que organizava a sociedade em uma hierarquia dominada pelos brancos.

(E) da necessidade de controle dos brancos sobre a população indígena, objetivando sua adaptação às exigências do trabalho regular.

Figura 4.13: Ciências Humanas e suas Tecnologias - Faixa: 750.0

ENEM 2012

Em certos locais, larvas de moscas, criadas em arroz cozido, são utilizadas como iscas para pesca. Alguns criadores, no entanto, acreditam que essas larvas surgem espontaneamente do arroz cozido, tal como preconizado pela teoria da geração espontânea.

Essa teoria começou a ser refutada pelos cientistas ainda no século XVII, a partir dos estudos de Redi e Pasteur, que mostraram experimentalmente que

(A) seres vivos podem ser criados em laboratório.

(B) a vida se originou no planeta a partir de microrganismos.

(C) o ser vivo é oriundo da reprodução de outro ser vivo pré-existente.

(D) seres vermiformes e microrganismos são evolutivamente aparentados.

(E) vermes e microrganismos são gerados pela matéria existente nos cadáveres e nos caldos nutritivos, respectivamente.

Figura 4.14: Ciências da Natureza e suas Tecnologias - Faixa: 630.4

> respectivos sujeitos, então o primeiro passo consiste em separar os sujeitos em grupos ao longo de uma escala de habilidade hipotética, cada grupo tem Z sujeitos de habilidades iguais. A probabilidade de os sujeitos de cada grupo responderem adequadamente a um item específico será dada pelo quociente entre o número de sujeitos que realmente acertaram ao item e o número total de sujeitos daquele grupo. Dessa forma as probabilidades de acerto em cada nível de habilidade ao longo da escala podem ser calculadas, isto é, tem-se uma curva empírica para cada item. A partir disso tenta-se manipular os parâmetros do item, produzindo uma curva teórica que mais se aproxime da empírica. O processo de estimação dos parâmetros se encerra quando os valores estimados convergirem, ou seja, quando a partir de n interações não se consegue produzir mais melhorias na reprodução dos dados empíricos por meio das variações nos valores dos parâmetros dos itens" [3]

Depois desse processo, dizemos que os itens foram calibrados. Após a calibração, a proficiência de cada sujeito também é estimada pelo método de Máxima Verossimilhança, nesse caso, é determinada a proficiência que maximiza a probabilidade do conjunto de respostas dada pelo sujeito.

O modelo logístico de três parâmetros da TRI adota a hipótese de que a aquisição do conhecimento ocorre de forma cumulativa, de modo que habilidades mais complexas requerem o domínio de habilidades mais simples. Por isso, um dos fatores que influencia no cálculo da estimação da habilidade é a coerência das respostas corretas. Nesse sentido, os alunos que acertaram as questões com maior nível de dificuldade também devem acertar as questões dos níveis anteriores.

Exemplo 4.7 Imagine que em uma avaliação com 5 questões de múltipla escolha previamente calibradas, dois participantes A e B acertaram 3 questões cada um. O participante A acertou de acordo com o padrão pedagógico esperado os itens 1, 2 e 3, enquanto o participante B acertou os itens 1, 2 e o item 5 de maior dificuldade da prova. Nesse caso, o modelo irá considerar que o participante B acertou o item 5 ao acaso, logo sua proficiência será menor do que o participante A. A figura a seguir ilustra essa situação.

◂

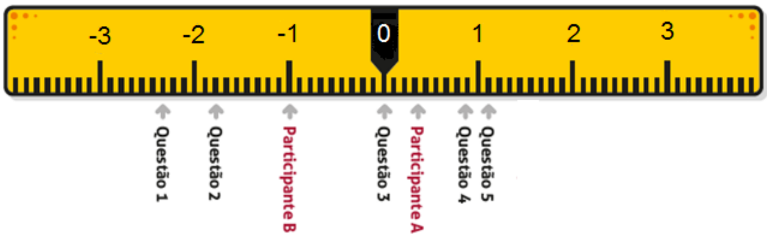


Figura 4.15: Escala de proficiência

5 Uso de Estatística em Processos Avaliativos

As disciplinas escolares utilizam diferentes instrumentos de avaliação para composição da média como trabalhos em grupos, exercícios avaliativos, acompanhamento de tarefas, relatórios e as tradicionais provas. Em geral, o professor corrige esses instrumentos de avaliação de acordo com uma pontuação predefinida e calcula a média de cada estudante.

Algumas indagações e reflexões que podemos fazer sobre os diferentes instrumentos de avaliação e todo esse processo para composição da média são:

1. Qual é a relação que existe entre os instrumentos de avaliação?
2. É possível, por exemplo, inferir a média final de um estudante analisando os resultados parciais de seus exercícios avaliativos?
3. Será que os critérios adotados por uma escola para correção das redações dos alunos do 3º do ensino médio são os mesmos de um determinado vestibular?
4. Qual é o grau de discriminação de cada questão que compõe uma prova, isto é, até que ponto cada questão consegue diferenciar estudantes com proficiências distintas?
5. Em uma prova, como identificar uma questão que tem problema de elaboração?

Vamos apresentar neste capítulo exemplos de como a Estatística pode colaborar na busca de respostas para essas indagações. Explicaremos também como o professor, ou a direção escolar, pode utilizar ferramentas estatísticas em uma planilha eletrônica para analisar os resultados de seus instrumentos de avaliação. Espera-se com isso, enriquecer o rol de indicativos escolares que podem auxiliar e direcionar as decisões pedagógicas.

5.1 Correlação

O coeficiente de correlação apresentado no capítulo de Estatística pode ser utilizado para quantificar o grau de correlação entre dois instrumentos de avaliação. Para calculá-lo, em uma planilha eletrônica, devemos:

1. Digitar em uma coluna os valores da primeira avaliação;

2. Digitar em outra coluna os valores da segunda avaliação;

3. Em uma célula vazia, digitar o comando de correlação.

Exemplo 5.1 Uma professora olhou cinco lições ao longo de um trimestre e aplicou uma prova trimestral valendo de 0 a 10. A planilha mostra o número de lições que cada aluno fez, sua nota na prova trimestral, o coeficiente de correlação e o gráfico de dispersão entre esses dois instrumentos.

	A	B	C
1	Aluno	Número de lições	Prova trimestral
2	A	5	9,0
3	B	2	6,0
4	C	3	6,0
5	D	4	7,0
6	E	3	8,0
7	F	2	4,5
8	G	0	3,5
9	H	1	3,5
10	I	2	4,0
11	J	5	7,5
12	K	2	4,5
13	L	2	3,5
14	M	3	5,0
15	N	4	6,5
16	O	5	8,0
17	P	1	4,0
18	Q	1	3,5
19	R	2	6,5
20	S	0	2,0
21	T	1	2,5
22			
23		Correlação	0,89

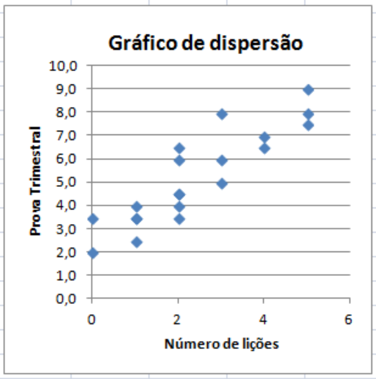


Figura 5.1: Planilha eletrônica

Para calcular o coeficiente de correlação, foi digitado na célula C23 dessa planilha o comando:

=correl{B2:B21;C2:C21}

Observe que os pontos do gráfico de dispersão lembram o formato de uma reta e que os dois alunos que não fizeram nenhuma das cinco lições tiraram na avaliação trimestral notas inferiores a 4,0, enquanto os três alunos que fizeram todas as lições tiraram notas superiores a 7,0. Além disso, o coeficiente de correlação entre os dois instrumentos foi 0,89, um alto índice de correlação linear, nesse caso, podemos dizer que um dos principais fatores que influenciou na nota da prova trimestral foi o número de lições que o aluno fez.

É importante lembrar que um alto índice correlação não significa uma relação de causa e consequência, ou seja, só porque um aluno faz lição não implica obter nota alta na prova trimestral. Porém, nesse contexto, uma hipótese provável é: o hábito de fazer lição demonstra

uma rotina de estudos o que faz obter uma nota alta na prova trimestral.
◂

Exemplo 5.2 A planilha apresenta a média final de Redação de uma turma de alunos formados no 3º ano do ensino médio de uma escola, as notas de Redação desses alunos no ENEM, o coeficiente de correlação e o gráfico de dispersão entre essas notas.

	A	B	C
1		Notas de Redação	
2	Aluno	Média final	ENEM
3	A	6,5	711,0
4	B	8,0	815,0
5	C	7,0	721,4
6	D	8,5	536,0
7	E	6,0	591,0
8	F	9,0	680,0
9	G	9,5	674,0
10	H	10,0	756,0
11	I	8,5	825,0
12	J	7,5	816,0
13	K	8,5	625,3
14	L	6,5	515,3
15	M	6,5	685,2
16	N	6,0	598,7
17	O	7,5	801,2
18	P	9,5	689,1
19	Q	10,0	714,2
20	R	6,0	647,5
21	S	7,0	746,5
22	T	5,0	623,5
23			
24		**Correlação**	**0,31**

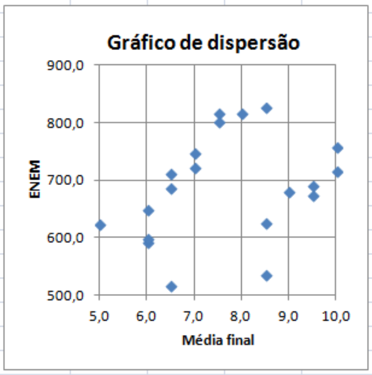


Figura 5.2: Planilha eletrônica

Para calcular o coeficiente de correlação, foi digitado o comando =correl{B2:B21;C2:C21}.

Observe que os pontos do gráfico de dispersão não lembram o formato de uma reta, isso é um indicativo de que não existe uma alto índice de correlação entre essas duas notas. O coeficiente de correlação foi 0,31, um índice baixo de correlação linear, nesse caso, podemos concluir que os critérios de correção de redação da escola são diferentes dos critérios de correção de redação do ENEM. Podemos verificar que há alunos com média final de Redação acima de 8,5 e nota no ENEM abaixo de 700, enquanto há alunos com média final abaixo de 8,5 e nota no ENEM acima de 700. ◂

5.2 Mínimos Quadrados

O Método de Mínimos Quadrados, descrito no capítulo de Estatística, é uma ferramenta que podemos utilizar para estabelecer uma relação linear entre dois instrumentos de avaliação. Sugerimos usar esse método apenas quando o índice de correlação entre os dois instrumentos for maior do que 0,7.

Para determinar, pelo método de mínimos quadrados, uma relação linear entre dois instrumentos de avaliação em uma planilha eletrônica, fizemos:

1. Digitamos em uma coluna os resultados da primeira avaliação e em outra coluna os resultados da segunda avaliação;

2. Selecionamos os resultados digitados e clicamos no comando "inserir gráfico". Das opções que apareceram, escolhemos o gráfico de dispersão que apresentou os resultados da primeira avaliação no eixo x e os resultados da segunda avaliação no eixo y;

3. Clicamos com o botão direito do mouse sobre um dos pontos do gráfico e selecionamos a opção "Adicionar linha de tendência". Na caixa de diálogo que abriu, marcamos "Exibir equação no gráfico".

Exemplo 5.3 Retomando o exemplo 5.1, em que o coeficiente de correlação entre o número de lições e a prova trimestral foi 0,89, determinaremos pelo método de mínimos quadrados uma relação linear entre esses dois instrumentos de avaliação. As imagens a seguir ilustram as principais etapas desse procedimento.

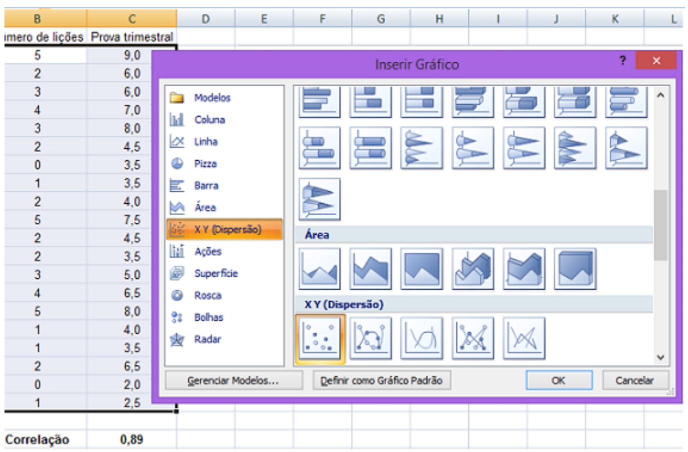


Figura 5.3: Comando inserir gráfico

Na última imagem aparece no gráfico a relação:

$$y = 1,1325x + 2,5321$$

Como y representa as notas da prova trimestral e x o número de lições, então se um aluno fez

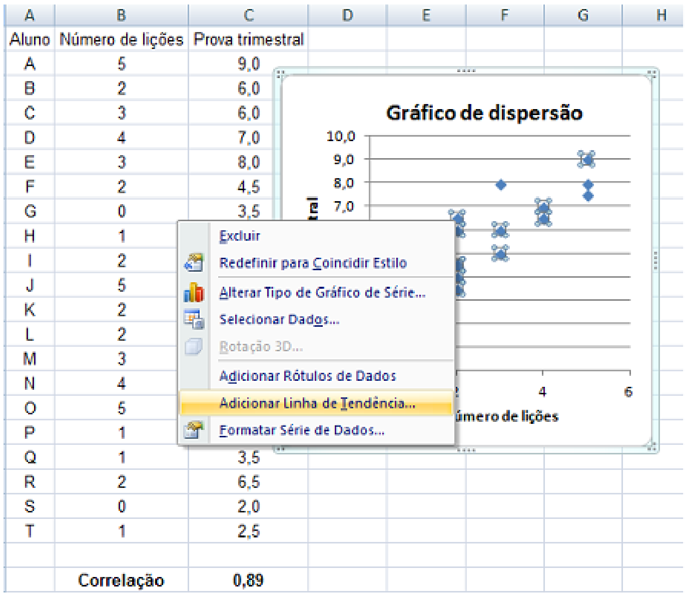


Figura 5.4: Comando adicionar linha de tendência

4 lições $(x = 4)$, o valor esperado da sua nota na prova trimestral é:

$$
\begin{aligned}
y &= 1,1325x + 2,5321 \\
&= 1,1325 \cdot 4 + 2,5321 \\
&= 7,0621
\end{aligned}
$$

Note que os dois alunos que fizeram 4 lições tiraram 6,5 e 7,0 na trimestral. Por meio da relação $y = 1,1325x + 2,5321$ podemos determinar o valor esperado da nota na prova trimestral em função do número de lições que o aluno fez. ◂

Exemplo 5.4 Uma professora de Português digitou em uma planilha a média final de redação de seus alunos do 3º ano do ensino médio e as notas de redação desses alunos no ENEM. Em seguida, calculou o coeficiente de correlação, esboçou o gráfico de dispersão e determinou uma relação linear entre essas notas.

O índice de correlação encontrado foi 0,87. Logo, a professora concluiu que seus critérios de correção de redação são parecidos com os critérios do ENEM porque os alunos com as maiores médias finais em sua disciplina conseguiram as maiores notas no ENEM e vice e versa.

A relação obtida pelo método de mínimos quadrados entre essas duas avaliações foi $y = 70,371x + 181,18$, onde y representa as notas no ENEM e x as médias finais. A professora, por meio dessa relação, construiu uma tabela com os valores esperados da nota no ENEM.

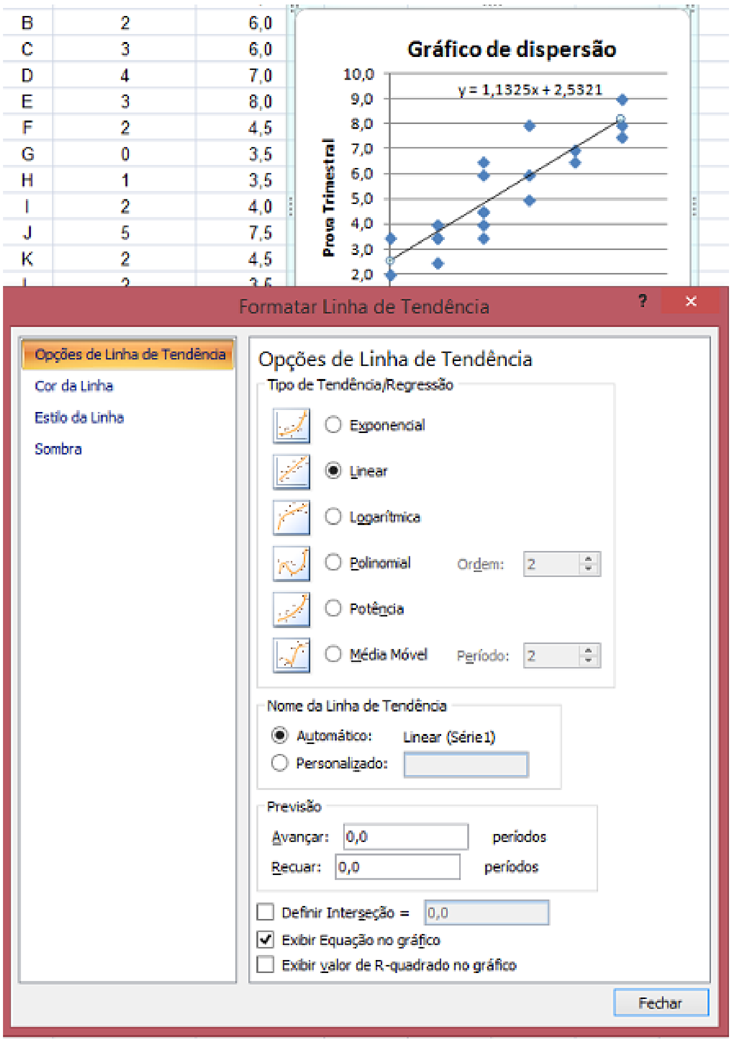


Figura 5.5: Comando exibir equação no gráfico

x	$y = 70,371x + 181,18$	y
6	$y = 70,371 \cdot 6 + 181,18$	603,4
7	$y = 70,371 \cdot 7 + 181,18$	673,8
8	$y = 70,371 \cdot 8 + 181,18$	744,1
9	$y = 70,371 \cdot 9 + 181,18$	814,5
10	$y = 70,371 \cdot 10 + 181,18$	884,9

◂

	Notas de Redação	
Aluno	Média final	ENEM
A	6,5	514,7
B	8,0	889,2
C	7,0	721,4
D	8,5	682,1
E	6,0	545,2
F	9,0	787,5
G	9,5	811,2
H	10,0	856,2
I	8,5	762,1
J	7,5	700,1
K	8,5	758,2
L	6,5	648,2
M	6,5	660,7
N	6,0	605,4
O	7,5	801,2
P	9,5	908,2
Q	10,0	898,1
R	6,0	598,4
S	7,0	689,2
T	5,0	553,1
	Correlação	**0,87**

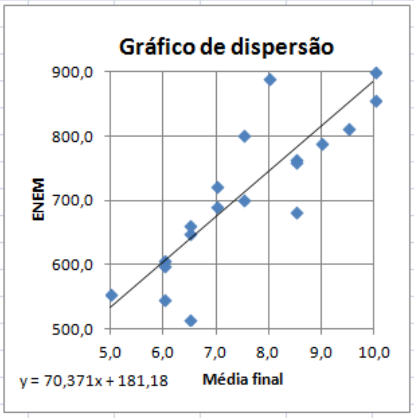


Figura 5.6: Planilha de notas

5.3 Ponto Bisserial

A prova é um instrumento de avaliação escolar que tem a finalidade de identificar o nível de proficiência ou habilidade de cada aluno. Por isso, ela deve conter questões:

- com diferentes graus de dificuldade;
- bem elaborada com alto poder de discriminação, ou seja, capazes de diferenciar estudantes com habilidades distintas.

Um modo de determinar o grau de dificuldade d_i de uma questão é calculando o seu percentual de acertos, ou seja, d_i é igual ao número n_i de alunos que acertaram a questão dividido pelo total t de alunos que fizeram a prova.

$$d_i = \frac{n_i}{t}$$

Logo, o parâmetro d_i varia de 0 a 1 de modo que as questões mais difíceis apresentarão um baixo percentual de acerto $(d_i < 0,3)$ e as questões mais fáceis apresentarão um alto percentual de acerto $(d_i > 0,7)$.

Exemplo 5.5 Vamos imaginar uma prova que contém 10 questões de múltipla escolha, cada questão valendo 1,0 ponto e grau de dificuldade d_i de acordo com a tabela.

Questão	1	2	3	4	5	6	7	8	9	10
d_i	0,41	0,28	0,76	0,93	0,35	0,49	0,56	0,62	0,18	0,86

Nessa prova, a questão 9 foi a mais difícil, com grau de dificuldade $d_i = 0,18$, e a questão 4 foi a mais fácil, com $d_i = 0,93$, ou seja, do total de alunos que fizeram a prova, 18% acertaram a questão 9 e 93% acertaram questão 4.
◂

Classificando as questões do exemplo anterior por grau de dificuldade da mais fácil para a mais difícil, obtemos a ordenação 4, 10, 3, 8, 7, 6, 1, 5, 2 e 9, veja.

Questão	4	10	3	8	7	6	1	5	2	9
d_i	0,93	0,86	0,76	0,62	0,56	0,49	0,41	0,35	0,28	0,18

Se cada questão dessa prova tem alto potencial de discriminação, então um aluno que tirou nota 6,0, deveria acertar as 6 questões mais fáceis (as questões 4, 10, 3, 8, 7 e 6) e errar as 4 questões mais difíceis (as questões 1, 5, 2 e 9). Do mesmo modo, se um aluno tirou nota 3,0, deveria acertar as 3 questões mais fáceis e errar as demais. Não é esperado, por exemplo, alunos obtendo notas acima de 6,0 errando a questão mais fácil, de número 4 com $d_i = 0,93$, e nem alunos tirando notas abaixo de 4,0 acertando a questão a mais difícil, de número 9 com $d_i = 0,18$.

Para quantificar o potencial de discriminação (P_d) das questões de uma prova, podemos separar em 3 grupos todos os indivíduos que resolveram a prova. Grupo superior com 27% dos indivíduos de maior desempenho na prova, grupo intermediário com 46% dos indivíduos com desempenho mediano e grupo inferior com 27% dos indivíduos de menor desempenho. Em seguida, para cada questão, determinar os percentuais de acerto de cada um desses grupos. A tabela a seguir ilustra como esquematizar esse procedimento.

Grupo	Número de indivíduos	Percentual de acertos
Superior	27% de maior desempenho	P_{superior}
Intermediário	46% de desempenho mediano	$P_{\text{intermediário}}$
Inferior	27% de menor desempenho	P_{inferior}

É esperado, para todas as questões, que o percentual de acertos do grupo superior seja maior do que o percentual de acertos do grupo intermediário e este seja maior do que o percentual de acertos do grupo inferior, isto é:

$$P_{\text{superior}} > P_{\text{intermediário}} > P_{\text{inferior}}$$

A diferença entre os percentuais de acerto do grupo superior e inferior é adotada como o potencial de discriminação (P_d) da questão.

$$P_d = P_{superior} - P_{inferior}$$

Quanto maior o valor do P_d, maior é o potencial de discriminação da questão. Na maioria das vezes, as questões são classificadas da seguinte maneira.

P_d	Classificação
$P_d \geqslant 40\%$	Boa
$30\% \leqslant P_d < 40\%$	Boa, mas sujeita a aprimoramento
$20\% \leqslant P_d < 30\%$	Precisa ser reelaborada
$P_d < 20\%$	Descartada

Exemplo 5.6 A planilha mostra a correção de uma prova com 10 questões de múltipla escolha em ordem crescente de notas. Q1 indica questão 1, Q2 questão 2 e assim por diante. Quando o aluno acertou a questão foi digitado o número 1 e 0 quando ele errou. A última linha apresenta o potencial de discriminação P_d, a linha **Pinferior** apresenta o percentual de acertos do grupo inferior, **Pintermediário** o percentual de acertos do grupo intermediário e **Psuperior** o percentual de acertos do grupo superior.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Correção da prova												
2	Alunos	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Nota	
3	A	0	1	0	0	0	1	0	0	0	0	2	Grupo inferior
4	B	1	1	0	0	0	0	0	0	0	0	2	
5	C	1	1	1	0	0	0	0	0	0	0	3	
6	D	1	1	1	0	0	0	0	0	0	0	3	
7	E	0	1	0	0	1	0	0	1	0	0	3	
8	F	1	1	0	0	0	0	0	0	1	0	3	
9	G	1	1	1	1	0	0	0	0	0	0	4	Grupo intermediário
10	H	1	1	1	1	0	0	0	0	0	0	4	
11	I	0	1	1	0	1	0	0	0	0	1	4	
12	J	1	1	1	1	1	0	0	0	0	0	5	
13	K	1	1	1	1	0	1	0	0	0	0	5	
14	L	1	1	0	1	0	1	1	0	0	0	5	
15	M	0	1	1	1	1	1	0	0	0	0	5	
16	N	1	1	0	1	1	0	1	0	1	0	6	
17	O	0	1	1	1	0	1	1	0	0	1	6	
18	P	1	1	1	1	1	1	1	0	0	0	7	
19	Q	0	1	1	1	1	1	1	1	0	1	8	Grupo superior
20	R	1	1	1	1	1	1	1	0	1	0	8	
21	S	0	1	1	1	1	1	1	1	0	1	8	
22	T	1	1	1	1	1	1	1	1	1	0	9	
23	U	0	1	1	1	1	1	1	1	1	1	9	
24	V	1	1	1	1	1	1	1	1	1	1	10	
25	Pinferior	0,67	1,00	0,33	0,00	0,17	0,17	0,00	0,17	0,17	0,00		
26	Pintermediário	0,70	1,00	0,80	0,90	0,50	0,50	0,40	0,00	0,10	0,20		
27	Psuperior	0,50	1,00	1,00	1,00	1,00	1,00	1,00	0,83	0,67	0,67		
28	Potencial de discriminação	-0,17	0,00	0,67	1,00	0,83	0,83	1,00	0,67	0,50	0,67		

Figura 5.7: Potencial de discriminação P_d

A primeira questão dessa prova tem potencial de discriminação negativo ($P_d = -0,17$), porque o percentual de acertos do grupo inferior foi maior do que o percentual de acertos do grupo superior, logo essa questão tem problemas de elaboração e deve ser examinada. Todos os alunos acertaram a questão 2, por isso o seu potencial de discriminação foi 0. As questões 4 e 7 têm potencial de discriminação 1, a diferença entre elas é o percentual de acertos do grupo intermediário, enquanto 90% dos alunos do grupo intermediário acertaram a questão 4, 40% apenas dos alunos do grupo intermediário acertaram a questão 7. Observe também que as questões 8 e 9 não apresentaram percentuais de acertos na seguinte ordem:

$$P_{\text{superior}} > P_{\text{intermediário}} > P_{\text{inferior}}$$

Tanto a questão 8 quanto a 9 tiveram percentuais de acertos do grupo inferior maior que do grupo intermediário. Logo, elas devem ser analisadas. ◂

O coeficiente de correlação ponto bisserial é a ferramenta Estatística recomendada para mensurar o potencial de discriminação de uma questão.

Definição 5.7. *O coeficiente de correlação ponto bisserial ρ_{pb} é definido por:*

$$\rho_{pb} = \frac{M_p - M}{\sigma} \cdot \sqrt{\frac{p}{1-p}}$$

Onde:

- M_p *é a média dos indivíduos que acertaram a questão;*
- M *é a média de todos que resolveram a prova;*
- σ *é o desvio padrão da média de todos que resolveram a prova;*
- p *é o grau de dificuldade* d_i *da questão, ou seja, é o percentual de acertos da questão.*

O ponto-bisserial ρ_{pb} varia de -1 a 1, valores próximos de zero ou negativo indicam questões com baixo índice de discriminação que precisam ser reelaboradas. Geralmente, itens com coeficiente ponto bisseral inferiores a $0,30$ são considerados de baixa discriminação.

O exemplo a seguir ilustra como calcular e analisar os indicativos estatísticos média, desvio padrão, grau de dificuldade e potencial de discriminação (coeficiente de correlação ponto bisseral) em uma planilha eletrônica.

Exemplo 5.8 A planilha ilustra a correção de uma prova com 10 questões de múltipla escolha, onde: o número 1 indica que o aluno acertou a questão e 0 que ele errou; Q1 representa a primeira questão, Q2 a segunda questão e assim por diante; Mp é a média dos alunos que acertaram a questão; M é a média de todos os alunos; σ é o desvio padrão da média de todos os alunos; p é o grau de dificuldade da questão e ρ_{Pb} é o potencial de discriminação da questão (coeficiente de correlação ponto bisseral).

Para calcular a média Mp dos alunos que acertaram a questão 1 foi digitado na célula B25 a fórmula:

	A	B	C	D	E	F	G	H	I	J	K	L
1	Correção da prova											
2	Alunos	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Nota
3	A	1	0	0	0	0	0	0	0	0	0	1
4	B	0	1	0	0	0	0	0	0	0	0	1
5	C	0	0	1	0	0	0	0	0	0	0	1
6	D	0	1	1	0	0	0	0	0	0	0	2
7	E	1	1	1	0	0	0	0	0	0	0	3
8	F	1	1	1	0	0	0	0	0	0	0	3
9	G	1	1	1	0	0	0	0	0	0	0	3
10	H	0	1	1	1	0	0	0	0	0	0	3
11	I	0	1	1	1	0	0	0	0	0	0	3
12	J	1	1	1	1	0	0	0	0	0	0	4
13	K	1	1	1	1	0	0	0	0	0	0	4
14	L	0	1	1	1	1	0	0	0	0	0	4
15	M	0	0	1	1	1	1	1	0	0	0	5
16	N	1	1	1	1	1	1	0	0	0	0	6
17	O	1	1	1	1	1	1	0	0	0	0	6
18	P	1	1	1	1	1	1	1	0	0	0	7
19	Q	0	1	1	1	1	1	1	1	0	0	7
20	R	0	1	1	1	1	1	1	1	0	0	7
21	S	1	0	1	1	1	1	1	1	1	0	8
22	T	0	1	1	1	1	1	1	1	1	0	8
23	U	0	1	1	1	1	1	1	1	1	1	9
24	V	0	1	1	1	1	1	1	1	1	1	9
25	Mp	4,5	4,9	5,1	6,0	6,9	7,2	7,5	8,0	8,5	9,0	
26	M	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7	4,7	
27	σ	2,60	2,60	2,60	2,60	2,60	2,60	2,60	2,60	2,60	2,60	
28	p	0,45	0,82	0,91	0,68	0,50	0,45	0,36	0,27	0,18	0,09	
29	ρ_{pb}	-0,08	0,18	0,45	0,72	0,84	0,87	0,81	0,77	0,68	0,52	

Figura 5.8: Ponto bisserial ρ_{Pb}

=médiase(B3:B24;1;L3:L24)

Para as demais questões, trocou o intervalo B3:B24 pelo intervalo correspondente à questão.

A fórmula =média(L3:L24) foi digitada na linha 26 de cada questão para determinar a média M de todos os alunos, do mesmo modo, digitou a fórmula =desvpad(L3:L24) na linha 27 de cada questão para determinar o desvio padrão σ.

Ao todo, 22 alunos resolveram essa prova, então para calcular o grau de dificuldade p da questão 1 digitou na célula B28:

=soma(B3:B24)/22

Para as demais questões, trocou o intervalo B3:B24 pelo intervalo correspondente à questão.

Na célula B29 digitou o seguinte comando para determinar o coeficiente de correlação ponto bisserial ρ_{pb} da questão 1:

=((B25-B26)/B27*raiz(B28/(1-B28)))

Para as demais questões, trocou a letra B pela letra correspondente à coluna de cada questão.

A seguir, a análise dos indicativos estatísticos.

As questões 1 e 6 tiveram o mesmo grau de dificuldade, ambas com 45% de acertos. A diferença entre elas foi o potencial de discriminação ρ_{pb}, enquanto a questão 6 apresentou o maior potencial de discriminação $\left(\rho_{pb} = 0,87\right)$ das questões dessa prova, a questão 1 apresentou o menor potencial de discriminação com um valor negativo $\left(\rho_{pb} = -0,08\right)$. Observe que a nota média M dos alunos foi 4,7, quem tirou acima de 4,7 acertou a questão 6 e quem tirou abaixo de 4,7 errou a questão 6, em outras palavras, a questão 6 separou os alunos em dois grupos, um grupo com notas acima da média e outro abaixo da média. O valor negativo $\left(\rho_{pb} = -0,08\right)$ da questão 1 indica problemas na elaboração dessa questão, observe na planilha que os alunos com as melhores notas estão errando a questão 1, enquanto alunos com notas medianas ou baixas estão acertando.

A questão 3 foi a mais fácil com 91% de acertos, apenas os alunos A e B erraram essa questão. Ela apresentou bom potencial de discriminação $\left(\rho_{pb} = 0,45\right)$, porque os dois alunos que erraram tiraram as menores notas dessa prova.

A segunda questão mais fácil foi a número 2, os alunos A, C, M e S foram os únicos que erraram essa questão. Observe que os alunos A e C tiraram nota 1 na prova, M tirou nota 5 e S tirou 8, por isso a questão 2 apresentou baixo potencial de discriminação $\left(\rho_{pb} = 0,18\right)$ e deve ser examinada.

A questão 10 foi a mais difícil com 9% de acertos, apenas os dois alunos que tiraram nota 9 acertaram essa questão. Por isso, ela apresentou bom potencial de discriminação $\left(\rho_{pb} = 0,52\right)$, uma vez que identificou os alunos com as maiores proficiências.
◂

Bibliografia

1 ANDRADE, D. F. de; TAVARES, H. R.; CUNHA VALLE, R. da. Teoria da Resposta ao Item: conceitos e aplicações. **ABE, Sao Paulo**, 2000.

2 CHUNG, K. L. **A course in probability theory**. [Sine loco]: Academic press, 2001.

3 COUTO, G.; PRIMI, R. Teoria de resposta ao item (TRI): conceitos elementares dos modelos para itens dicotômicos. **Boletim de Psicologia**, Associação de Psicologia de São Paulo, volume 61, número 134, páginas 1–15, 2011. Citado na página 73.

4 DANTAS, C. A. B. **Probabilidade: Um Curso Introdutório**. [Sine loco]: Edusp, 2013.

5 DE AYALA, R. J. **The theory and practice of item response theory**. [Sine loco]: Guilford Publications, 2013.

6 EVES, H. W. **Introdução à história da matemática**. [Sine loco]: Unicamp, 1995.

7 MAGALHÃES, M. N.; LIMA, A. C. P. de. **Noções de probabilidade e estatıstica**. [Sine loco]: IME-USP São Paulo: 2000.

8 MEYER, P. L. Introductory probability and statistical applications. Addison-Wesley, 1965.

9 PASQUALI, L.; PRIMI, R. Fundamentos da teoria da resposta ao item: TRI. **Avaliação Psicológica**, Instituto Brasileiro de Avaliação Psicológica. UFRGS, volume 2, número 2, páginas 99–110, 2003.

10 RABELO, M. Avaliação Educacional: fundamentos, metodologia e aplicações no contexto brasileiro. **Rio de janeiro: SBM**, volume 29, páginas 30–31, 2013.

11 ROSS, S. **Probabilidade: um curso moderno com aplicações**. [Sine loco]: Bookman, 2010.

12 RUMSEY, D. **Estatıstica II Para Leigos**. [Sine loco]: Alta Books Editora, 2014.