Universidade Virtual Africana

INFORMÁTICA APLICADA: MATH 1302

INTRODUÇÃO À ESTATÍSTICA E PROBABILIDADE

João Carlos Lopes Horta

Prefácio

A Universidade Virtual Africana (AVU) orgulha-se de participar do aumento do acesso à educação nos países africanos através da produção de materiais de aprendizagem de qualidade. Também estamos orgulhosos de contribuir com o conhecimento global, pois nossos Recursos Educacionais Abertos são acessados principalmente de fora do continente africano.

Este módulo foi desenvolvido como parte de um diploma e programa de graduação em Ciências da Computação Aplicada, em colaboração com 18 instituições parceiras africanas de 16 países. Um total de 156 módulos foram desenvolvidos ou traduzidos para garantir disponibilidade em inglês, francês e português. Esses módulos também foram disponibilizados como recursos de educação aberta (OER) em oer.avu.org.

Em nome da Universidade Virtual Africana e nosso patrono, nossas instituições parceiras, o Banco Africano de Desenvolvimento, convido você a usar este módulo em sua instituição, para sua própria educação, compartilhá-lo o mais amplamente possível e participar ativamente da AVU Comunidades de prática de seu interesse. Estamos empenhados em estar na linha de frente do desenvolvimento e compartilhamento de recursos educacionais abertos.

A Universidade Virtual Africana (UVA) é uma Organização Pan-Africana Intergovernamental criada por carta com o mandato de aumentar significativamente o acesso a educação e treinamento superior de qualidade através do uso inovador de tecnologias de comunicação de informação. Uma Carta, que estabelece a UVA como Organização Intergovernamental, foi assinada até agora por dezenove (19) Governos Africanos - Quênia, Senegal, Mauritânia, Mali, Costa do Marfim, Tanzânia, Moçambique, República Democrática do Congo, Benin, Gana, República da Guiné, Burkina Faso, Níger, Sudão do Sul, Sudão, Gâmbia, Guiné-Bissau, Etiópia e Cabo Verde.

As seguintes instituições participaram do Programa de Informática Aplicada: (1) Université d'Abomey Calavi em Benin; (2) Université de Ougagadougou em Burkina Faso; (3) Université Lumière de Bujumbura no Burundi; (4) Universidade de Douala nos Camarões; (5) Universidade de Nouakchott na Mauritânia; (6) Université Gaston Berger no Senegal; (7) Universidade das Ciências, Técnicas e Tecnologias de Bamako no Mali (8) Instituto de Administração e Administração Pública do Gana; (9) Universidade de Ciência e Tecnologia Kwame Nkrumah em Gana; (10) Universidade Kenyatta no Quênia; (11) Universidade Egerton no Quênia; (12) Universidade de Addis Abeba na Etiópia (13) Universidade do Ruanda; (14) Universidade de Dar es Salaam na Tanzânia; (15) Universite Abdou Moumouni de Niamey no Níger; (16) Université Cheikh Anta Diop no Senegal; (17) Universidade Pedagógica em Moçambique; E (18) A Universidade da Gâmbia na Gâmbia.

Bakary Diallo

O Reitor

Universidade Virtual Africana

Créditos de Produção

Autor

João Carlos Lopes Horta

Par revisor(a)

Arlindo Oliveira da Veiga

UVA - Coordenação Académica

Dr. Marilena Cabral

Coordenador Geral Programa de Informática Aplicada

Prof Tim Mwololo Waema

Coordenador do módulo

Florence Tushabe

Designers Instrucionais

Elizabeth Mbasu

Benta Ochola

Diana Tuel

Equipa Multimédia

Sidney McGregor

Barry Savala

Edwin Kiprono

Kelvin Muriithi

Victor Oluoch Otieno

Michal Abigael Koyier

Mercy Tabi Ojwang

Josiah Mutsogu

Kefa Murimi

Gerisson Mulongo

Direitos de Autor

Apoiado por

Projeto Multinacional II da UVA financiado pelo Banco Africano de Desenvolvimento.

Índice

Descrição Geral do Curso

Bem-vindo (a) a Introdução à Estatística e Probabilidade

A disciplina de Introdução à Estatística e Probabilidade fornece instrumentos teóricos para recolher, explorar, descrever e, finalmente, interpretar e analisar dados.

Na análise dos dados surgem algumas incertezas. Então, a Probabilidade é uma ferramenta para medir o grau dessa incerteza.

Esta disciplina tem aplicações em muitas áreas de conhecimentos, como por exemplo, Genética, Economia, Ciências Sociais, Engenharias, Ciências da Educação, Administração, Ciências da Computação, Medicina, Biologia, Psicologia, etc.

Pré-requisitos

Revisão de:

- Matemática Discreta:
- Conjuntos e suas operações;
- Técnicas de Contagem (combinatória).
- Cálculo Aplicado para Computação
- Limite e Continuidade em R
- Diferenciação e Integração em R

Materiais

Os materiais necessários para completar este curso incluem:

- Fernandes, S., & Pinto, M. (s.d.). Afinal o que são e como se calculam os quartis? Universidade de Algarve - Departamento de Matemática, Departamento de Matemática, Faro.
- Leithold, L. (1994). O Cálculo com Geometria Analítica (3ª ed., Vol. I). (C. D. Patarra, Trad.) Harbra Ltda.
- Hazzan, S. et al.. (1977). Fundamentos de Matemática Elementar. Combinatória e Probabilidade. (3ª ed.) São Paulo: Atual Editora.
- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill.
- Gersting, J. l. (1995). Fundamentos Matemáticos para Ciência de Computação. (L. L. Fialho & M. M. Filhos, Trad.). (3ª ed). Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A. (Obra original publicada em 1993).
- Farber, L. (2010). Estatística Aplica. (4ª ed.). São Paulo: Pearson Prentice Hall.
- Máquina Científica: TI-83 ou semelhante.
- Software: SPSS ou R.

Objetivos do Curso

Após concluir este curso, o(a) aluno(a) deve ser capaz de:

- Representar, analisar e resolver problemas de tratamentos de dados na Ciência de computação;
- Aplicar o conceito básico de Probabilidade, bem como algumas das suas propriedades;
- Explorar as situações aleatórias que envolvam os conceitos de fortuna e azar;
- Conceber e realizar estudos, utilizando métodos estatísticos descritivos, para fundamentar os resultados obtidos;
- Seleccionar e aplicar métodos estatísticos adequados na análise dos dados;
- Explorar, analisar, interpretar e explorar informações de natureza estatística.

Unidades

Unidade 0: Exploração de pré-requisitos. Teste Diagnóstico

Rever a teoria de conjuntos, a teoria de técnicas de contagem, a teoria de limite, continuidade, diferenciação e integração de uma função real de variável real.

Unidade 1: Dados estatísticos. Estatística Descritiva

Conceitos de população e amostra. Classificação dos dados estatísticos. Introdução das medidas de tendência central e de dispersão dos dados.

Unidade 2: Probabilidade e teoremas associados

Conceitos de Espaço de Resultados e Eventos. Definições de probabilidade, seus axiomas e suas consequências. Probabilidade Condicional e teoremas associados (Teorema de Bayes).

Unidade 3: Variável aleatória (unidimensional). Função de distribuição de uma variável

Introduz-se o conceito de variável aleatória (unidimensional) e faz-se a classificação da variável aleatória (discreta, contínua e mista). Introduz-se o conceito de função de distribuição, bem como as suas principais propriedades.

Unidade 4: Aplicação do Estudo de Estatística e Probabilidade ao curso de Ciência Informática Aplicada (doravante CIA)

Propor exercícios, em mais diversas áreas, no âmbito do curso em questão, para motivar os alunos. Apresentar alguns softwares utilizados na resolução de problemas estatísticos.

Avaliação

Em cada unidade encontram-se incluídos instrumentos de avaliação formativa a fim de verificar o progresso do(a)s aluno(a)s.

No final de cada módulo são apresentados instrumentos de avaliação sumativa, tais como testes e trabalhos finais, que compreendem os conhecimentos e as competências estudadas no módulo.

A implementação dos instrumentos de avaliação sumativa fica ao critério da instituição que oferece o curso. A estratégia de avaliação sugerida é a seguinte:

1	Avaliação formativa	20%
2	Avaliação sumativa I	40%
3	Avaliação sumativa II	40%

Calendarização

Unidade	Temas e Atividades	Estimativa do tempo
Unidade 1: Dados Estatísticos. Estatística Descritiva	Atividade 1 – Estatística. Classificação de dados e planeamento de experiência. (5 horas) Atividade 2 – Organização dos dados em uma experiência estatística. (5 horas) Atividade 3 – Medidas de tendência centra e de dispersão. (10 horas)	20 horas
Unidade 2: Probabilidade e teoremas associados	Atividade 1 – Espaço-amostra e Acontecimento. Probabilidade e seus axiomas. (10 horas) Atividade 2 – Probabilidade condicional e os teoremas relacionados; O Teorema de Bayes. (15 horas) Atividade 3 – Acontecimentos Independentes e suas consequências. (5 horas)	30 horas
Unidade 3: Variável aleatória (unidimensional). Função de distribuição de uma variável	Atividade 1 – Variável aleatória e Função de distribuição (10 horas) Atividade 2 – Distribuições discretas. (20 horas) Atividade 3 – Distribuições contínuas. (20 horas)	50 horas

Unidade 4: Aplicação do Estudo de Estatística e Probabilidade à CIA	Atividade 1 – (6 horas) Software utilizadas na resolução de problemas estatísticos. Atividade 2 – (7 horas) Estatística descritiva: utilização de software Atividade 3 – (7 horas) Cálculo de probabilidades: utilização de software	20 horas

Leituras e outros Recursos

As leituras e outros recursos deste curso são:

Unidade 0

Leituras e outros recursos obrigatórios:

- Hazzan, S. et al.. (1977). Fundamentos de Matemática Elementar. Combinatória e Probabilidade. (3ª ed.) São Paulo: Atual Editora. (Capítulos I e II)
- Gersting, J. I. (1995). Fundamentos Matemáticos para Ciência de Computação. (L. L. Fialho & M. M. Filhos, Trad.). (3ª ed). Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A. (Obra original publicada em 1993).
- Leithold, L. (1994). O Cálculo com Geometria Analítica (3ª ed., Vol. I). (C. D. Patarra, Trad.) Harbra Ltda.

Unidade 1

Leituras e outros recursos obrigatórios:

- Farber, L. (2010). Estatística Aplicada. (4ª ed.). São Paulo: Pearson Prentice Hall.
- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill de Portugal.
- Fernandes, S., & Pinto, M. (s.d.). Afinal o que são e como se calculam os quartis? Universidade de Algarve - Departamento de Matemática, Departamento de Matemática, Faro.

Leituras e outros recursos opcionais:

- Máquina Científica: TI-83 ou semelhante.
- Software: Geogebra, R e Excel

Unidade 2

Leituras e outros recursos obrigatórios:

- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill de Portugal.
- Neto, J., (2013). Introdução à Probabilidade (Relatório Técnico).

Leituras e outros recursos opcionais:

- Farber, L. (2010). Estatística Aplicada. (4ª ed.). São Paulo: Pearson Prentice Hall.
- O Software
- Máquina Científica: TI-83 ou semelhante.
- Software: Geogebra, R e Excel.

Unidade 3

Leituras e outros recursos obrigatórios:

- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill de Portugal.
- Farber, L. (2010). Estatística Aplicada. (4ª ed.). São Paulo: Pearson Prentice Hall.
- Software: Geogebra ou R.

Leituras e outros recursos opcionais:

- Máquina Científica: TI-83 ou semelhante.

Unidade 4

Leituras e outros recursos obrigatórios:

- Farber, L. (2010). Estatística Aplica. (4ª ed.). São Paulo: Pearson Prentice Hall.
- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill de Portugal.
- Software: Geogebra, R e Excel (folha de cálculo).

Leituras e outros recursos opcionais:

- Qualquer software estatístico (SPSS ou outro).
- Máquina Científica: TI-83 ou semelhante

Unidade 0. Diagnóstico

Introdução à Unidade

O propósito desta unidade é verificar a compreensão dos conhecimentos que possui relacionados com este curso.

Nesta unidade apresentam-se alguns exercícios de teoria de conjuntos e de teoria de contagem que sirvam de base para a compreensão e computação de probabilidade de um determinado acontecimento. Para mais detalhes, aconselha-se o leitor a consultar as bibliografias propostas para a unidade zero.

De uma forma muito ligeira, introduz-se os softwares SPSS – Statistical Package for Social Science – e R – language and environment for statistical computing and graphics – como ferramentas para realização de alguns exercícios ao longo deste curso e, também, na realização de futuras tarefas, fora do âmbito desta disciplina.

Objetivos da Unidade

Após a conclusão desta unidade, deverá ser capaz de:

- Definir e identificar conjuntos disjuntos dois a dois.
- Reunir e intersectar conjuntos.
- Determinar o complementar de um conjunto.
- Determinar as partes de um conjunto.
- Aplicar propriedades de operações com conjuntos, nomeadamente princípio aditivo para partes disjuntas e princípio da Inclusão-Exclusão, na resolução de problemas
- Identificar problemas onde se aplica princípio fundamental da contagem, arranjos, permutações e combinações.
- Resolver problemas aplicando princípio fundamental da contagem, arranjos, permutações e combinações.
- Determinar limite de uma função real de variável real.
- Analisar a continuidade e a diferenciabilidade de uma função real de variável real.
- Determinar integral de uma função real de variável real.

Termos-chave

Estatística: é a ciência que coleta, organiza, analisa e interpreta dados para a tomada de decisões.

Estatística Descritiva: ocupa da análise de dados apenas, mas não faz inferências.

População ou Universo: é o maior conjunto para um determinado estudo estatístico.

Amostra: é um subconjunto da população donde recai um estudo estatístico.

Amostragem: é o processo de determinação de uma amostra.

Censo: é um estudo estatístico feito com todos os elementos de uma população.

Leituras e Outros Recursos

As leituras e outros recursos desta unidade encontram-se na lista de "Leituras e Outros Recursos do curso".

Unidade 1. Dados estatísticos. Estatística Descritiva.

Introdução à Unidade

Nesta unidade vai introduzir-se alguns conceitos essenciais e indispensáveis para a realização e compreensão de uma experiência estatística, como classificação de dados e tipo de estudos estatísticos.

Vai introduzir-se algumas medidas de tendência central e de dispersão, no âmbito da estatística descritiva.

Pretende-se, também, mostrar a diferença entre estatística descritiva e inferencial, apesar de o estudo da estatística inferencial estar fora do âmbito deste módulo.

Objetivos da Unidade

Após a conclusão desta unidade, deverá ser capaz de:

- Definir Estatística e alguns conceitos a ela associados.
- Diferenciar amostra da população.
- Organizar os dados estatísticos.
- Diferenciar os diferentes tipos de amostragem.
- Identificar e diferenciar as medidas de tendência central e, as de dispersão.
- Identificar as técnicas utilizadas no planeamento de um estudo estatístico.
- Identificar as técnicas de coleta de dados numa experiência de dados.
- Analisar descritivamente um conjunto de dados.

Termos-chave

Sondagem: é um estudo estatístico feito com elementos de uma amostra.

Variável: é um atributo de um estudo estatístico.

Representação de dados: são meios para expor os dados estatísticos de uma forma resumida, para poder analisá-los com mais rigor e facilidade.

Medidas de tendência central: é um valor que representa uma entrada central de um conjunto de dados estatísticos.

Medidas de dispersão: São valores utilizados para avaliar o grau de variabilidade, ou dispersão, dos valores em torno da média.

Actividades de Aprendizagem

Actividade 1 – Estatística. Classificação de dados e planeamento de experiência.

Introdução

Nesta actividade vai introduzir-se alguns conceitos essenciais e indispensáveis para a realização e compreensão de uma experiência estatística, como classificação de dados e tipo de estudos estatísticos.

Aborda-se, também, os objetivos da estatística.

Detalhes da actividade

Definição (Dados estatísticos). Dados estatísticos consistem em informações que vêm de observações, contagens, medições ou respostas.

Definição (Estatística). É um conjunto de técnicas e métodos de pesquisa que entre outros tópicos envolve o planeamento do estudo a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações.

Num estudo estatístico há dois tipos de conjuntos de dados que são chamados população e amostra.

Uma população é o conjunto de todas as unidades estatísticas. Mas muitas vezes, por abuso de linguagem, chama-se população aos resultados, respostas, medições ou contagens que advém de atributos das unidades estatísticas. É o conjunto fundamental para efetuar estudo estatístico.

Uma amostra é um subgrupo finito de uma população.

A forma de seleção de uma amostra a partir da população, designada por processo de amostragem, é determinante para a qualidade das inferências que venham fazer-se. As técnicas corretamente para colher amostras deixam total ou parcialmente ao acaso a seleção dos elementos que virão integrá-las. Se intervém apenas o acaso e se todos os elementos de população têm igual possibilidade de ser escolhidos, a amostragem diz-se casual. Por outro lado, quando a escolha da amostra é feita apenas por critérios arbitrários do investigador, a amostragem diz-se dirigida.

Exemplo. No estudo sobre "idade de cada membro da equipa reitoral da UVA", a população é equipa reitoral da UVA, enquanto que no estudo "pesquisa com espectadores de um estádio com espectadores", a amostra é composta por aqueles espetadores (a população neste caso é composta pelos espetadores).

O termo "estatística" é utilizado para representar uma descrição numérica de uma característica de amostra.

O termo "parâmetro" é utilizado para representar uma descrição numérica de uma característica de população.

Exemplo. "A média (que se definirá mais adiante) dos salários anuais para dos contadores de uma empresa é de USD" é uma estatística.

"Sessenta e dois dos passageiros de um determinado aeronave sobreviveram a uma explosão" é um parâmetro.

Ramos de Estatística

O estudo de estatística tem duas ramificações consideráveis:

Estatística descritiva: é o ramo da estatística que envolve a organização, o resumo e a representação dos dados.

Estatística inferencial: é o ramo da estatística que envolve o uso de uma amostra para chegar a conclusão sobre a população.

Exemplo. A estatística descritiva envolve afirmações do tipo "para os que são solteiros ainda estavam vivos aos " e "para os casados, estavam vivos aos ". Uma inferência tirada do estudo é que "o facto de ser casado está associado a uma vida mais longa", por exemplo.

Tipo de dados estatísticos

Dados qualitativos: consistem de atributos, rótulos ou entradas não numéricas. Às vezes são representados numericamente, mas deve-se estabelecer a correspondência entre os números utilizados (quase sempre números inteiros) com as modalidades. Por exemplo, género
- {homem,mulher}, então homem≡0 e mulher≡1.Dados quantitativos: consistem de medidas numéricas ou contagens, que se apresentam-se em diferentes intensidades ou valores.

Exemplo. No recenseamento Geral Agrícola de 1989, cada unidade estatística – exploração agrícola – foi observada em relação a vários atributos qualitativos, entre os quais podem salientar-se:

- Exploração segundo a natureza jurídica do produtor – sete modalidades – produtor singular autónomo, produtor singular empresário, sociedades, cooperativas agrícolas, baldios, Estado e outros entes públicos, outas entidades;
- Explorações segundo a forma de exploração – quatro modalidades – conta própria, arrendamento fixo, arrendamento de campanha e arrendamento variável.

Entre os atributos quantitativos podem mencionar-se:

- Superfície agrícola utilizada;
- Número de blocos de exploração. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 9)

Definição (variável discreta e contínua). Em dado quantitativo, qualquer que seja o atributo considerado, o seu valor numérico pode variar de elemento para elemento. Para assinalar este facto representam-se esses valores por uma variável, seja . Se a amostra tiver dimensão n, ter-se-á x_1,x_2,…,x_n,

onde x_i (i=1,2,…,n) será o valor do atributo na i-ésima observação.

As variáveis, com que correntemente se depara, são de dois tipos: discretas e contínuas.

São do primeiro tipo as variáveis que podem tomar somente um número finito ou uma infinidade numerável de valores, como por exemplo, "número de acidentes por apólice de seguros".

São do segundo tipo as variáveis que podem assumir qualquer valor dentro de um intervalo de números reais, co por exemplo, "Altura de um indivíduo".

Os dados também podem ser classificados consoante o nível de mensuração:

- Dados no nível nominal de mensuração: são apenas qualitativos. Neste nível, os dados podem ser categorizados usando-se nomes, rótulos ou qualidades. Não são realizados cálculos matemáticos neste nível. Por exemplo, estado civil - {casado,solteiro,divorciado,separado}..

- Dados no nível ordinal de mensuração: são qualitativos ou quantitativos. Dados neste nível podem ser organizados em ordem ou posição, mas as diferenças entre as entradas de dados não são significantes. Por exemplo, Grau de especialização dos trabalhadores - {não especializado≡1,semi-especializado≡2, especializado≡3,muito especializado≡4 }. A ordem é feita respeitando a ordem das várias modalidades, mas isso não quer dizer que, por exemplo, o trabalhador "muito especializado" corresponde a quatro vezes mais do que trabalhador "não especializado".
- Dados no nível de mensuração intervalar: podem ser ordenados e pode-se calcular diferenças significativas entre as entradas de dados. No nível intervalar, um registro nulo simplesmente representa uma posição em uma escala; a entrada não é um zero inerente (zero que significa "nada", como zero dólar numa conta bancária).
- Dados no nível de mensuração racional: são similares aos dados no nível intervalar, mas com uma propriedade adicional – neste nível um registo nulo é um zero inerente. Uma razão de dois valores de dados pode ser formada de modo que um valor de dado possa ser significativamente expresso como múltiplo do outro.

Exemplo. "A temperatura corporal, em graus Celsius, durante uma sessão de exercícios" está no nível intervalar de mensuração.

"Os índices cardíacos, em batidas por minuto, de um atleta durante uma sessão de exercícios" estão no nível de mensuração racional.

"Uma coleção de números de telefone" está no nível de mensuração nominal.

"As posições finais de equipas de futebol num determinado campeonato" estão no nível de mensuração ordinal.

Nota. A tabela abaixo é utilizada para facilitar a determinação do nível de mensuração de um determinado tipo de dados:

Nível de mensuração	Categorizar os dados	Ordenar os dados	Subtrair os valores dos dados	Determinar se um valor de dado é múltiplo do outro
Nominal	Sim	Não	Não	Não
Ordinal	Sim	Sim	Não	Não
Intervalar	Sim	Sim	Sim	Não
Racional	Sim	Sim	Sim	Sim

Os dados são univariados se de cada unidade se mede uma variável (ou atributo). Os dados são bivariados se de cada unidade se medem duas variáveis (ou atributos). Os dados são multivariados se de cada unidade se medem mais do que duas variáveis (ou atributos). Os dados bivariados, ou multivariados, contêm mais informações do que os univariados.

Exemplo. A tabela que se segue mostra o número de graus académicos atribuídos num determinado país e num determinado ano, classificados por duas variáveis nominais: o nível do grau e o sexo do candidato.

Graus obtidos, por grau e sexo			
Sexos\Graus	**Bacharelato**	**Mestrado**	**Doutoramento**
Homens			
Mulheres			
Total			

Várias interpretações podem ser feitas a partir desta tabela. Por exemplo:

Qual é a percentagem dos graus de doutor que foram atribuídos a indivíduos do sexo feminino?

Qual a percentagem dos graus de doutor foram e atribuídos às senhoras?

Qual a percentagem dos graus, atribuídos às senhoras, foram de doutor?

Para responder às estas questões pode-se acrescentar mais uma coluna à tabela (total geral), no sentido de facilitar os cálculos, como se segue:

Graus obtidos, por grau e sexo				
Sexos\ Graus	**Bcharelato**	**Mestrado**	**Doutoramento**	**Total geral**
Homens				
Mulheres				
Total				

Sendo assim:

Alínea (a).

$$\frac{7\,803}{34\,076} \times 100 \cong 22.90\% .$$

Alínea (b).

$$\frac{7\,803}{1\,281\,520} \times 100 \cong 0.61\%$$

Alínea (c).

$$\frac{7\,803}{578\,953} \times 100 \cong 1.35\% .$$

Planeamento de um estudo estatístico

O objetivo de todo estudo estatístico é coletar dados e então usá-los para tomar decisão.

Qualquer decisão que seja tomada usando os resultados de um estudo estatístico será tão boa quanto ao processo utilizado para a sua obtenção desses dados.

Se o processo tiver falhas, então a decisão resultante será questionável.

Instruções para planeamento de um estudo estatístico:

1. Identifique a variável (ou variáveis) de interesse (foco) e a população do estudo.
2. Desenvolva um plano detalhado para a recolha de dados. Se usar uma amostra, tenha a certeza de que ela representa a população.
3. Colete os dados.
4. Descreva os dados utilizando as técnicas de estatística descritiva (vista na atividade 1.2)
5. Interprete os dados e tome as decisões sobre a população usando estatística inferencial.
6. Identifique quaisquer erros possíveis.

Método para coleta de dados.

Há várias maneiras de coletar os dados, mas é o foco de estudos que determina, ou orienta, a melhor maneira de coleta de dados.

Alguns métodos de coleta de dados:

1. Faça um estudo observacional. Neste estudo, um pesquisador observa e mede as caraterísticas de interesse de parte de uma população, mas não muda as condições existentes.
2. Realize uma experiência. Ao realizar uma experiência, um tratamento é aplicado em uma parte da população e as respostas são observadas. Outra parte da população pode ser utilizada como grupo de controle, na qual nenhum tratamento foi aplicado. As respostas do grupo de tratamento e do grupo de controle são comparadas estudadas.
3. Em muitos casos indivíduos de grupo de controle recebem placebos – um tratamento não medicamentoso e que não causa danos, feito para parecer tratamento real.
4. Use uma simulação. Uma simulação é o uso de um modelo matemático ou físico para reproduzir as condições de uma situação ou processo.
5. Use um levantamento ou pesquisa de mercado. É uma pesquisa de uma ou mais característica de uma determinada população.

Nota. A diferença de fundo entre uma experiência e um estudo observacional consiste no facto de que num estudo observacional o pesquisador não influencia as respostas, enquanto que numa experiência, um pesquisador deliberadamente aplica um tratamento antes de observar uma resposta.

Exemplo. Num "estudo dos efeitos da ingestão de farinha de aveia na redução de pressão arterial", deve-se realizar uma experiência na coleta de dados.

Num "estudo sobre como alunos do ensino primário resolve uma quebra-cabeça", deve-se realizar um estudo observacional na coleta de dados.

Os fabricantes de automóveis usam simulações com bonecos para "estudar os efeitos de batida de um automóvel em humanos".

Definição (Censo e Sondagem). Um estudo estatístico que utiliza todos os elementos da população denomina-se por censo. Enquanto que um estudo estatístico realizado numa amostra denomina-se por sondagem.

Conclusão

População e amostra são dois tipos de conjuntos de dados fundamentais em estudos estatísticos.

Os dados podem ser classificados em qualitativos e quantitativos. Os níveis de mensuração podem ser: nominal, ordinal, intervalar e racional.

O método (ou métodos) de coleta de dados num estudo estatístico depende em larga escala do tipo de variável (variáveis) em estudo.

Avaliação

1. Determine se o conjunto de dados é uma amostra ou uma população:
   a. "A altura de cada quarta pessoa que entra num parque de diversões"
   b. "O número de televisores em cada residência em Nairobi".
2. Determine, justificando, se o valor numérico, abaixo indicado, é uma estatística ou um parâmetro:
   a. "Em janeiro de 2007, 44% dos governadores de um determinado país eram do partido da oposição".
   b. "Num determinado ano, a nota média de Estatística e Probabilidade para todos os graduandos num determinado curso, numa determinada universidade, é 14 valores (na escala de 0 a 20).

3. Em um estudo, voluntários que dormiram 8 horas em uma noite eram três vezes mais capazes de responder corretamente as questões de um teste de Matemática, em relação àqueles que não tiveram horas de sono suficientes.

   a. Indique a amostra usada no estudo.

   b. Qual era a população em estudo?

   c. Que parte do estudo representa o ramo descritivo da estatística?

4. Determine se os dados, abaixo indicados, são qualitativos ou quantitativos. Classifica-os quanto ao nível de mensuração:

   a. "As temperaturas diárias mais altas para o mês de julho".

   b. "As durações de músicas em um MP3 player".

   c. "Os números de jogadores de um time de futebol".

   d. "Respostas de uma pesquisa de opinião".

   e. "Medidas de pressão arterial diastólica".

5. Qual o método de coleta de dados que deva usar em cada um dos estudos indicados abaixo:

   "Estudo sobre os efeitos dos exercícios no alívio de depressão".

   "Um estudo de sucesso de graduandos da UVA para encontrar um emprego no primeiro ano de graduação".

   a. Indique o foco de cada um destes estudos.

   b. Identifique a população do estudo.

   c. Escolha um método apropriado para a coleta de dados.

Atividade 2 – Organização dos dados em uma experiência estatística

Introdução

Nesta atividade aborda-se as diferentes formas de organizar os dados estatísticos, para melhor realização de estudos.

Detalhes da atividade

Gráficos Caule-e-folhas. É utilizada para organizar os dados de uma amostra pequena. Permite que o observador se torne mais sensível ao aspeto global dos dados, sem perda de informação, e comece a explorar a existência de padrões.

Exemplo. O quadro abaixo apresenta os dados referentes à taxa anual de nupcialidade, TAN, (número de casamento ocorrido durante o ano, referido à população média desse ano, por habitantes). Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 3)

Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN
1960	7.8	1969	9	1978	8.3	1987	7.2	1996	6.3
1961	8.8	1970	9.4	1979	8.1	1988	7.1	1997	6.5
1962	7.9	1971	9.7	1980	7.3	1989	7.4	1998	6.6
1963	7.9	1972	9.1	1981	7.7	1990	7.3	1999	6.8
1964	8.1	1973	9.9	1982	7.4	1991	7.3	2000	6.2
1965	8.4	1974	9.3	1983	7.5	1992	7	2001	5.7
1966	8.6	1975	10.9	1984	7	1993	6.8	2002	5.4
1967	8.9	1976	10.5	1985	6.8	1994	6.6		
1968	8.7	1977	9.4	1986	6.9	1995	6.6		

Quadro 1 - Taxa de nupcialidade em Portugal mo período 1960 a 2002

O diagrama caule-e-folha para estes dados é:

N.º de Folhas	Caule	Folhas
2	5	47
10	6	2356668889
14	7	00123334457899
8	8	11346789
7	9	0134479
2	10	59
N.º de dados: ; unidade:		

Quadro 2 Diagrama Caule-e-folhas de Taxa de Nupcialidade em Portugal, 1960-2002

Cada caule – dígito dominante (neste caso a parte inteira) – está numa linha e tem um número variável de folhas – dígitos dominados (no caso presente a parte decimal) –; cada observação é uma folha de um determinado caule. Por exemplo, a dado 9,4 corresponde a 4ª e a 5ª observações do quinto caule – 8l44. O caule 8 tem 8 folhas – 8l11346789 – que correspondem as oito observações: 8,1, 8,1, 8,3, 8,4, 8,6, 8,7, 8,8 e 8,9.

Com a construção do gráfico caule-e-folhas procura dar-se, através de uma análise exploratório, uma visão geral dos dados, salientando ou identificando propriedades, tais como:

- Existência de valores em torno dos quais se concentram os restantes;
- Maior ou menor concentração dos valores;
- Simetrias ou assimetrias;
- Presença de valores extremos.

Distribuições de frequências e histogramas

Quando o número de dados é muito grande, os gráficos caule-e-folhas perdem um pouco de interesse, por não ser possível analisar os dados com facilidade.

As distribuições de frequências nesses casos podem ser uma boa solução.

Definição (Frequência absoluta e relativa em variáveis discretas e qualitativas).). Seja uma coleção de dados {x_1,x_2,…,x_n}, onde x_i, i=_(1,n), representa o valor numérico do atributo observado para a unidade estatística i. Seja a_1,a_2,…,a_j,…, a sucessão disposta por ordem crescente dos valores possíveis que a variável x pode assumir. O conjunto dos valores da sucessão chama-se domínio da variável x, e representa-se por A.

Chama-se frequência absoluta de a_j na coleção, e representa-se por F_j, ao número de vezes que este valor é observado.

Deste modo, tem-se a sucessão das frequências absolutas, $F_1, F_2, \ldots, F_j, \ldots$, verificando-se, $\sum_j F_j = n$, onde n é o número total de observações da coleção de dados.

Chama-se frequência relativa do valor a_j, e representa-se por f_j, ao quociente entre F_j e n, ou seja $f_j = F_j/n$.

Da mesma maneira, obtém-se a sucessão das frequências relativas $f_1, f_2, \ldots, f_j, \ldots$, verificando-se $\sum_j f_j = 1$.

Tratando-se de uma variável discreta, basta fazer corresponder a cada valor observado um número natural, e, a partir, daí elaborar a tabela de frequências absoluta e relativa.

Exemplo. O quadro que se segue contém distribuição de frequências do "número de indivíduos por agregado doméstico privado" no Continente (Europeu), retirada do Inquérito aos orçamentos Familiares 1989/1990. A Amostra é composta por agregados. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 14).

Número de Indivíduos	Frequências absolutas	Frequências relativas
1	1138	0.118
2	2748	0.285
3	2304	0.239
4	2082	0.216
5	848	0.088
≥6	520	0.054
Total	9640	1.000

Quadro 3 Inquérito aos Orçamentos Familiares 1989/1990

Pode-se retirar algumas informações, de carácter descritivas, dos dados, como por exemplo, há "2340 agregados domésticos privados com três indivíduos", "5,4% de agregados domésticos privados é composto de 6 ou mais indivíduos", etc.

Exemplo. Perguntou-se a cada funcionário de uma determinada loja, qual o tipo de música de fundo que preferiam ouvir durante o expediente. Os resultados foram:

R	O	P	C	J	P	P
R	J	O	P	P	P	O
C	P	C	N	J	R	R
J	R	R	P			

C - CLÁSSICA	J - JAZ
P - POP	O - OUTRO TIPO DE MÚSICA
R - ROCK	N - NENHUM

Então, a tabela de frequências absolutas e relativas destes dados é a que se segue:

Preferência musical	Frquências Absolutas	Frequências Relativas
CLÁSSICA - C	3	0.12
JAZ -J	4	0.16
N - NENHUM	1	0.04
O -OUTRO TIPO DE MÚSICA	3	0.12
P - POP	8	0.32
R -ROCK	6	0.24
TOTAL	25	1

Quadro 4 Preferência Musical dos Funcionários de uma loja

Pode-se, também, acrescentar a tabela de frequências, a frequência absoluta acumulada e a frequência relativa acumulada.

Definição (Frequência acumulada absoluta (ou relativa)). A frequência absoluta (ou relativa) acumulada de uma observação de uma variável discreta (ou de uma variável qualitativa) é a soma da frequência absoluta (ou relativa) daquela observação com as frequências absolutas (ou relativas) das observações que a antecedem na tabela de frequências.

A frequência acumulada absoluta (ou relativa) da primeira observação na tabela de frequências é a respetiva frequência absoluta (ou relativa).

A frequência acumulada absoluta (ou relativa) da última observação na tabela de frequências é igual ao n.º total de observações (ou igual a).

Exemplo. A tabela que se segue ilustra as frequências acumuladas dos valores da variável "tipo de música de fundo preferido durante o expediente". Ver (Preferência Musical dos Funcionários de uma loja).

Preferência musical			Freq. Absol. Acumulada	Freq. Rel. Acumulada
CLÁSSICA - C	3	0.12	3	0.12
JAZ -J	4	0.16	7	0.28
N - NENHUM	1	0.04	8	0.32
O -OUTRO TIPO DE MÚSICA	3	0.12	11	0.44
P - POP	8	0.32	19	0.76
R -ROCK	6	0.24	25	1
TOTAL	25	1		

Quadro 5: Preferência Musical dos Funcionários de uma loja

Definição (Intervalo de classes em variáveis contínuas). Para classes de valores de uma variável contínua tomam-se, em regras intervalos, que são designados por intervalos de classe. A definição das classes consiste na fixação dos limites desses intervalos ou limites de classes. Admita-se que o número de classes é fixado em m, e represente-se o j-ésimo intervalo de classe por I_j. Para não haver ambiguidade é necessário que os intervalos não tenham pontos em comum, isto é, $I_j \cap I_k = \emptyset, j \neq k$.

Para não ficar valores por classificar, deve-se verificar $\bigcup_{j=1}^{m} I_j = A$, onde A é o domínio da variável.

Os m intervalos devem ser da forma:

$I_1 =]l_0, l_1], I_2 =]l_1, l_2], \ldots, I_j =]l_{j-1}, l_j], \ldots, I_m =]l_{m-1}, m]$,

onde $l_0 < l_1 < \cdots < l_j < \cdots < l_m$, $l_0 < \min(x_i)$ e $l_m \geq \max(x_i)$, e $\min(x_i)$ e $\max(x_i)$ são os valores mínimo e máximo, respetivamente, dos valores observados.

Por vezes, convém deixar o primeiro intervalo fechado à esquerda e/ou deixar o último intervalo aberto a direita. A amplitude da j-ésima classe é dada por $h_j = l_j - l_{j-1}$, para $j = 1, 2, \ldots, m$.

Sempre que não haja inconveniência e tal seja possível, deve-se manter constante as amplitudes das classes.

Há que saber como determinar o número de classes (k) e a amplitude de classe (h). Para este efeito, algumas regras são dadas.

1.Tabela de Truman L. Kelley

n	5	10	25	50	100	200	500	1000
k	2	4	6	8	10	12	15	15

2.$k=5$, para $n \leq 25$ e $k \cong \sqrt{n}$ para $n>25$

3.Regra de Sturges

Para uma amostra de dimensão n, o número de classes é dado por k≅1+3,22 log(n).

Nota. Normalmente estipula-se que o número de classes deve estar entre 5 e 20, para não dificultar os estudos estatísticos.

Para determinar a amplitude de classes deve-se em primeiro lugar determinar a amplitude dos dados (ou amplitude total), dado por . Depois, divide-se esse número pelo número de classes, e arredonda-se para o próximo número que seja conveniente.

Definição (Frequências em Intervalos de classes). Feito o apuramento dos valores pertencentes a cada classe, a frequência absoluta da classe I_j, j=1,2,…,m, F_j, ou a frequência relativa da mesma, f_j, são obtidas pelos métodos descritos anteriormente, para variáveis discretas (ou qualitativa). O mesmo acontece para as frequências, absoluta acumulada e relativa acumulada.Nota. Pode-se organizar dados de uma variável discreta em intervalos de classe, desde que o número de dados é muito grande e muito distintos.

Exemplo. Os dados do quadro que se segue representam "o tempo de hemodiálise (HD) antes da transplantação renal" efetuada em doentes no Hospital Santa Cruz. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 16)

NB: embora a variável seja contínua (o tempo), as observações são medidas em meses.

HD (Mês)	HD (Mês)	HD (Mês)	HD (Mês)	HD (Mês)	HD (Mês)	HD (Mês)	HD (Mês)
51	35	7	17	60	50	36	64
24	60	35	24	12	55	54	70
55	42	79	18	44	96	28	19
75	34	28	19	36	36	36	54
24	60	16	32	3	60	42	10
27	65	26	52	22	26	32	21
22	48	67	37	76	69	5	30
23	144	97	49	24	19	38	55
48	84	59	103	18	43	21	30
18	50	26	18	10	22	18	72
96	28	39	136	36	58	10	108
24	36	47	84	37	10	15	30
26	93	135	16	52	21	32	30

Quadro 6 Tempo de HD (meses) em 104 doentes transplantados

Distribui-se estes dados em intervalos de classes, e calcula-se as frequências absolutas e relativas.

No conjunto dos 104 valores observados, tem-se: min(x_i)=3 e max(x_i)=144. Então, tomou-se 0 para o limite inferior da primeira classe e 150 para o limite superior da última classe. Forma-se, desta forma, dez classes de iguais amplitudes, h=15

Classes de HD (Mês)	**Frequências Absolutas**	**Frequências Relativas**
]0, 15]	9	0.087
]15, 30]	35	0.337
]30, 45]	20	0.192
]45, 60]	20	0.192
]60, 75]	7	0.067
]75, 90]	4	0.038
]90, 105]	5	0.048
]105, 120]	1	0.010
]120, 135]	1	0.010
]135, 150]	2	0.019
Total	104	1.000

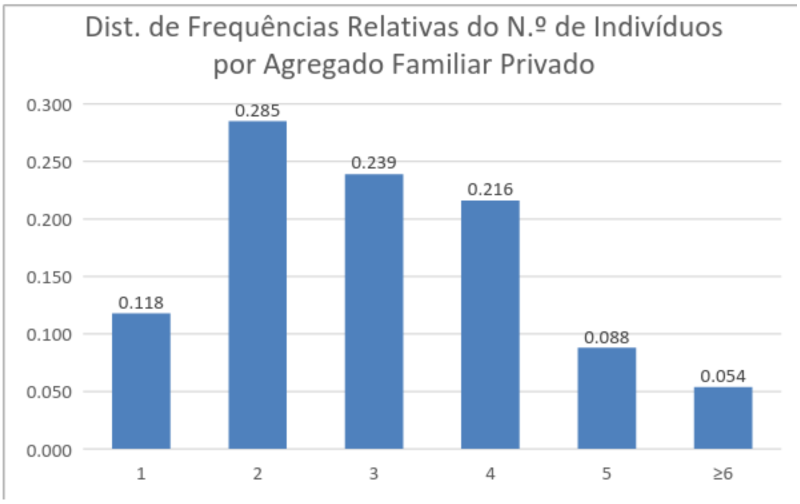
Quadro 7 Distribuição do tempo de HD (meses) em 104 doentes transplantados

Representação gráfica de dados: diagramas de barras e histogramas

Diagrama de barra. Utilizado para representar dados relativos à uma para variável discreta. No eixo das abcissas marca-se os valores das variáveis, e no eixo das ordenadas as respetivas frequências absolutas ou relativas.

As alturas das barras correspondem às frequências absolutas (ou relativas) dos respetivos dados.

Exemplo. A figura que se segue é o diagrama de barras da distribuição do "número de indivíduos por agregado doméstico privado". Ver (Inquérito aos Orçamentos Familiares 1989/1990) – Faça "ctrlclique" sobre a referência para ir aos dados.



Nota. Para organizar dados de uma variável quantitativa discreta numa tabela de frequências é necessário organizar os dados por ordem crescente.

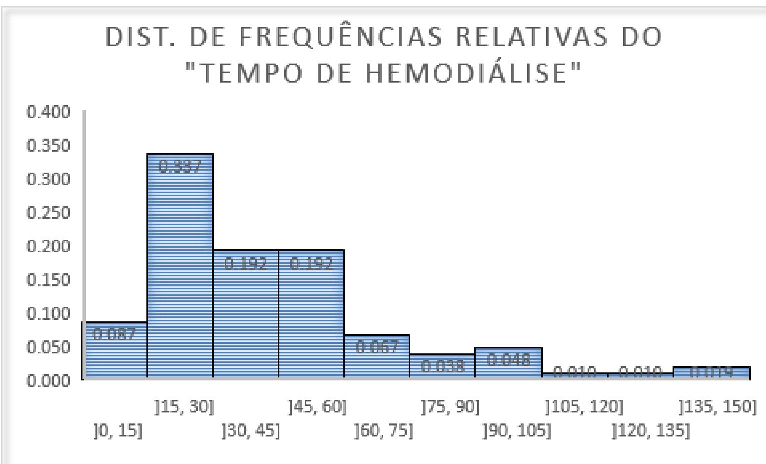
Histograma. É utilizado para representar dados agrupados em classes (geralmente para variáveis contínuas), e é formado por uma sucessão de retângulos adjacentes.

Quando as classes têm amplitude constante, h, cada retângulo tem base h e altura igual a respetiva frequência relativa (ou absoluta). Deste modo, a área do retângulo é igual a hf_j (ou hF_j), e a soma das áreas dos retângulos é igual a h (ou nh), onde n é o n.º de observações.

Como se pode supor sempre que h=1 (basta tomar uma escala conveniente), então a área ou a altura de cada retângulo é igual a respetiva frequência relativa (ou absoluta).

Se as amplitudes das classes são diferentes, cada retângulo tem base h_j e altura igual a respetiva frequência relativa (ou absoluta) dividida pela amplitude de classe, f_j/h_j (ou F_j/h_j). Deste modo, a área do retângulo correspondente à classe I_j é a frequência relativa (ou absoluta), e a soma das áreas dos retângulos é igual a 1 (ou n.º de observações).

Exemplo. A figura que se segue é o histograma de distribuição do "tempo de hemodiálise". Ver (Tempo de HD (meses) em 104 doentes transplantados) – Faça "ctrl+clique" sobre a referência para ir aos dados.



Histograma 1: Tempo de Hemodiálise

Definição (Marca da classe). É o ponto médio de uma classe, que se obtém adicionando os extremos da classe e dividindo a respetiva quantia por dois.

Polígono de frequência. É um gráfico linha, que se constrói a partir de um de um histograma, ou a partir de um gráfico de barras, e que enfatiza as mudanças contínuas nas frequências.

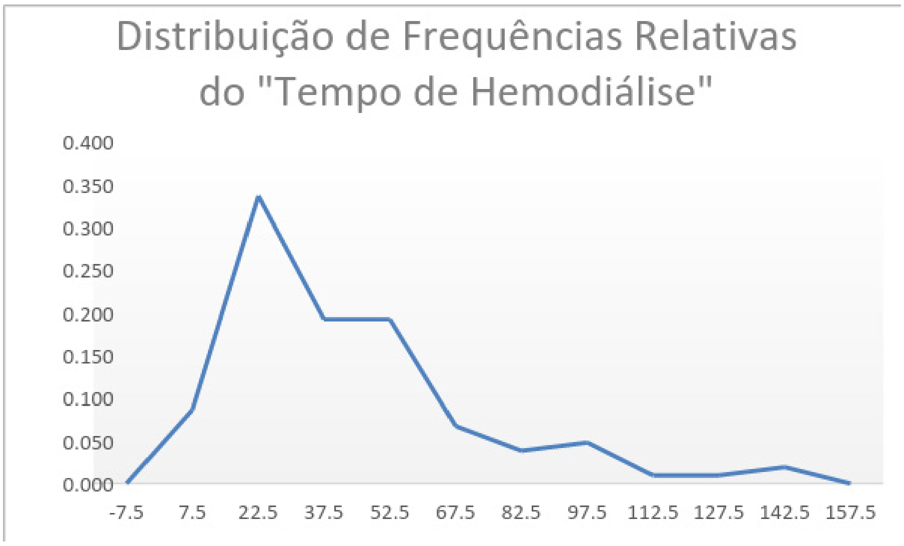
Pode-se construir um polígono de frequências relativamente às frequências absolutas (ou relativas) e às frequências (absolutas ou relativas) acumulada.

Polígono de frequências absolutas, ou relativas. Obtém-se, no caso de gráfico de barra, unindo os pontos mais altos dos retângulos; e no caso de um histograma, obtém-se unindo os pontos médios dos topos dos retângulos com os pontos que ficam à esquerda da primeira classe, que se obtém subtraindo do extremo inferior, dessa classe a sua amplitude, e o que fica à direita da última classe, que se obtém adicionando ao extremo superior, dessa classe, a sua amplitude.

O ponto que tiver mais altura num polígono de frequência representa a maior frequência, ao passo que a área abaixo da curva inclui a totalidade dos dados existentes.

Polígono de frequência (absoluta ou relativa) acumulada (ou ogiva). Mostra a frequência acumulada de cada classe em seu limite superior (às vezes utiliza-se o conceito de fronteira da classe na construção de uma ogiva, quando os limites da classe são obtidos de outra forma – Ver (Larson & Farber, 2010, p. 36)).

Exemplo. Veja abaixo o polígono de frequências relativas de distribuição do "tempo de hemodiálise".



Polígono de Frequência 1 (Relativa): Tempo de Hemodiálise

Nota. Existem mais tipos de gráficos que representam os dados estatísticos. Para obtenção de mais informações, pode-se consultar (Larson & Farber, 2010).

Tipos de distribuição de dados estatísticos. A partir do polígono de frequência pode-se classificar uma distribuição de dados estatísticos em:

Simétrica: quando uma linha vertical pode ser traçada do meio do gráfico da distribuição e as metades resultantes são aproximadamente imagens espelhadas. Neste caso, a média, a moda e a mediana ocupam, quase, a mesma posição.

Uniforme: quando todas as observações, ou classes, têm frequências iguais, ou aproximadamente iguais. Uma distribuição uniforme é, também, simétrica.

Assimétrica: quando a "cauda" do polígono de frequência se alonga mais em um dos lados.

Uma distribuição é assimétrica à esquerda (ou assimétrica negativa) se a cauda estender à esquerda.

Uma distribuição é assimétrica à direita (ou assimétrica positiva) se a cauda estender à direita.

Nota. A média situa-se sempre na direção em que a distribuição é assimétrica.

Algumas distribuições podem não classificar-se em simétrica, uniforme ou assimétrica.

Exemplo. A "Distribuição do Tempo de Hemodiálise", ver exemplo anterior, é assimétrica à direita.

Conclusão

Resumidamente, são duas as formas de representar os dados estatísticos: tabelas de distribuições de frequências e diagramas (gráfico de barras, histograma, gráficos caule-e-folhas, etc.).

Representação dos dados estatísticos depende muito do tipo de variável (atributo) em estudo e do número de observações.

As formas de representação de dados de uma variável discreta são, de uma certa forma, diferentes das formas para representar os dados de uma variável contínua.

Avaliação

1. O conjunto de dados amostrais, a seguir, lista o "número de minutos que usuários de internet gastam na rede durante sua mais recente sessão".

50	40	41	17	11	7	22	44	28
41	78	56	72	56	17	7	69	30
18	29	34	59	73	77	36	39	30
21	80	62	19	56	54	23	29	67
37	33	39	51	46	31	54	31	53
42	39	44	86	20				

a. Use uma das regras para determinação do número de classes, para calcular o número de classes desta distribuição.

b. Determine a amplitude de classe.

c. Elabore a tabela de frequências absoluta, relativa e acumulada.

d. Esboce o histograma.

2. Considere a seguinte tabela de frequências absolutas:

Classe	F_i
$]19, 30]$	19
$]30, 41]$	43
$]41, 52]$	68
$]52, 63]$	69
$]63, 74]$	74
$]74, 85]$	68
$]85, 96]$	24

a.Determine:

A amplitude da classe;

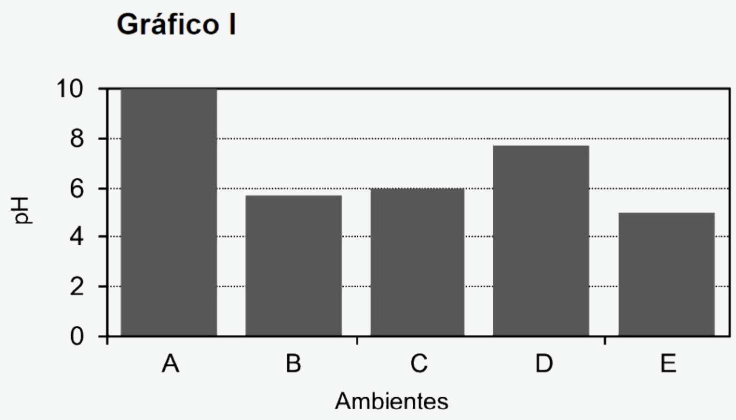
A amplitude dos dados;

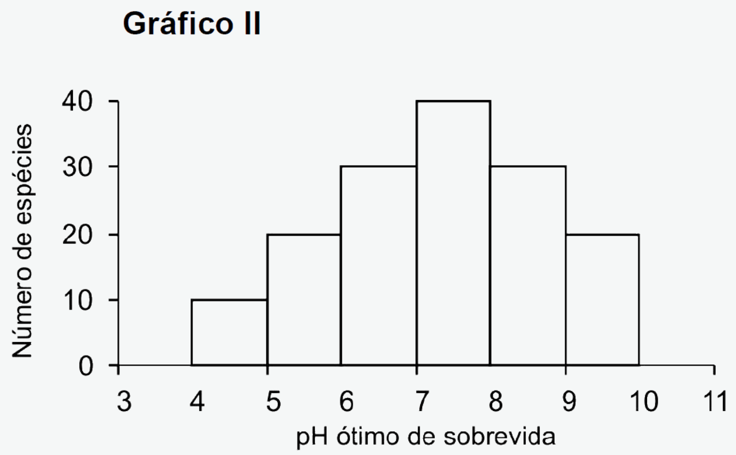
A marca de cada classe.

b.Esboce um histograma de frequências relativas.

c.Esboce o polígono de frequências relativas.

3. Um estudo caraterizou 5 ambientes aquáticos, nomeados de A a E, em uma região, medindo parâmetros físico-químicos de cada um deles, incluindo o pH nos ambientes. O gráfico I representa os valores de pH dos 5 ambientes. Utilizando o gráfico II, que representa a distribuição estatística de espécies em diferentes faixas de pH, pode-se esperar um maior número de espécies no ambiente A, B, C, D ou E?





4. Considere a seguinte tabela, de um conjunto de dados bivariáveis:

Milionários, por idade e sexo (1972, em milhares)			
Sexos\ Graus	**Abaixo de 50 anos**	**De 50 a 64 anos**	**65 anos e mais**
Homens	24	34	31
Mulheres	39	26	25

Quantos milionários existiam em 1972?

a. Quantos milionários existiam em 1972?

b. Que percentagem desses milionários era senhoras?

c. Que percentagem dos milionários tinham menos de 50 anos?

d. Que percentagem dos milionários do sexo feminino, tinham 65 ou mais anos?

e. Que percentagem dos milionários com menos de 50 anos, eram do sexo masculino?

Actividade 3 – Medidas de tendência central e de dispersão

Introdução

Neta atividade vai introduzir-se algumas medidas de tendência central e de dispersão, indispensáveis para identificar características importantes de um conjunto de dados estatísticos.

Detalhes da atividade

Medidas de tendência central. É um valor que representa uma entrada central de um conjunto de dados estatísticos.

As medidas de tendência central mais comuns são: a média, a moda e a mediana.

Nota. Embora se faça referência à população, uma ou outra vez, nesta atividade dá-se realce às medidas para dados amostrais.

Definição (Média aritmética). É a soma das entradas de um conjunto de dados dividida pelo número das entradas.

A média (para todos os efeitos, a média aritmética) indica um valor central da distribuição, entendido como o valor em torno qual se distribuem os valores do conjunto dos dados.

O cálculo da média de um conjunto de dados faz-se a partir de uma das formas:

$\mu=\frac{\sum_{i=1}^{N} x_i}{N}$ (média populacional) ou $\bar{x}=\frac{\sum_{i=1}^{n} x_i}{n}$ (média amostral),

onde N representa o número de dados em uma população, n o número de dados em uma amostra e x_i, para cada i, representa uma observação de dados.

Se os dados estiverem organizados numa tabela de frequências absolutas (F_i), pode-se calcular a média da seguinte forma:

$\mu=\frac{\sum_{j=1}^{N} F_j a_j}{N}$ (média populacional) ou $\bar{x}=\frac{\sum_{j=1}^{n} F_j a_j}{n}$ (média amostral),

onde $a_1, a_2, \ldots, a_j, \ldots$ é a sucessão disposta por ordem crescente dos valores possíveis que a variável x pode assumir, ou seja, são os diferentes valores da variável x.

Quando os dados estiverem agrupados em classes os valores de x_i representam os pontos médios das classes, ou marcas das classes.

Exemplo. O quadro abaixo apresenta os dados referentes à taxa anual de nupcialidade, TAN, (número de casamento ocorrido durante o ano, referido à população média desse ano, por habitantes), ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 3):

Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN
1960	7.8	1969	9	1978	8.3	1987	7.2	1996	6.3
1961	8.8	1970	9.4	1979	8.1	1988	7.1	1997	6.5
1962	7.9	1971	9.7	1980	7.3	1989	7.4	1998	6.6
1963	7.9	1972	9.1	1981	7.7	1990	7.3	1999	6.8
1964	8.1	1973	9.9	1982	7.4	1991	7.3	2000	6.2
1965	8.4	1974	9.3	1983	7.5	1992	7	2001	5.7
1966	8.6	1975	10.9	1984	7	1993	6.8	2002	5.4
1967	8.9	1976	10.5	1985	6.8	1994	6.6		
1968	8.7	1977	9.4	1986	6.9	1995	6.6		

Tabela 1 Taxa Anula de Nupcialidade

Calcula-se a média deste conjunto de dados:

$\sum x_i=336.1; n=43$, logo $\bar{x}=(\sum x_i)/n=336.1/43\cong 7.816$

Definição (Estatística de ordem). Quando um atributo é quantitativo ou qualitativo, com o nível de mensuração nominal, os dados da coleção $\{x_1,x_2,\ldots,x_n\}$ podem ser ordenados,

$x_{(1)},x_{(2)},\ldots,x_{(n)}$, com $x_{(1)}\le x_{(2)}\le\cdots\le x_{(n)}$,

e diz-se que se calcularam as estatísticas de ordem. Assim, $x_{(i)}$ é a estatística de ordem i, com $i=1,2,\ldots,n$. Duas estatísticas de ordem importantes são os extremos, $x_{(1)}=\min(x_i)$ e $x_{(n)}=\max(x_i)$.

Definição (Mediana). É um valor que ocupa a posição central quando os dados estão ordenados, e representa-se por $\tilde{x}$. Ou seja, a mediana é o valor do conjunto dos dados, que deixa 50% de observações inferiores e 50% de observações superiores.

Se $x_{(1)},x_{(2)},\ldots,x_{(n)}$ n forem estatísticas de ordem, a mediana será o elemento que ocupará a posição central. Em termos rigorosos, a mediana poder-se-á definir-se por:

$\tilde{x}=\begin{cases} x_{((n+1)/2)} & \text{se } n \text{ é ímpar} \\ (x_{(n/2)}+x_{(n/2+1)})/2 & \text{se } n \text{ é par} \end{cases}$.

Exemplo. A mediana sobre o conjunto de dados do exemplo anterior é $\tilde{x}=x_{(22)}=7.5$.

Se os dados estiverem a grupados em classes, para determinar a mediana proceder-se-á da seguinte forma:

1. Calcular-se-á a ordem n/2;
2. Pela tabela das frequências acumuladas, identificar-se-á a classe que contém a mediana – classe Md;
3. Utilizar-se-á a fórmula

$\tilde{x}=l_{Md}+\frac{(n/2-f_{(ac\_ca)})\cdot h}{F_{Md}}$,

onde:

- l_{Md} – Limite inferior da classe Md;
- n – Dimensão da amostra;
- $f_{(ac\_ca)}$ – Frequência acumulada da classe anterior à classe Md;
- h – Amplitude da classe Md;
- F_{Md} – Frequência absoluta da classe Md.

Exemplo. Calcula-se a média e a mediana do conjunto de dados sobre o tempo de hemodiálise, ver o “Tempo de HD (meses) em 104 doentes transplantados”.

Para isso, apresenta-se a seguinte tabela:

Classes de HD (Mês)				
]0, 15]	9	9	7.5	67.5
]15, 30]	35	44	22.5	787.5
]30, 45]	20	64	37.5	750
]45, 60]	20	84	52.5	1050
]60, 75]	7	91	67.5	472.5
]75, 90]	4	95	82.5	330
]90, 105]	5	100	97.5	487.5
]105, 120]	1	101	112.5	112.5
]120, 135]	1	102	127.5	127.5
]135, 150]	2	104	142.5	285
Total	104			4470

Então a média é:

$\bar{x}=\frac{\sum F_i a_i}{n}=\frac{4470}{104}\cong 42.981$.

Cálculo da mediana: $n/2=104/2=52$; então a classe mediana é Md=]30,45]. Logo,

$\tilde{x}=l_{Md}+\frac{(n/2-f_{(ac\_ca)})\cdot h}{F_{Md}}=30+\frac{(52-44)\cdot 15}{20}=36$.

Nota. Segundo (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 25), “a mediana está muito menos sujeita a observações extremas ou, em linguagem mais técnica, é mais resistente do que a média”, pois a média se altera sempre que se altere uma observação, enquanto que a mediana não, por se ocupar a posição central.

Definição (quantis). Um quantil de ordem α, $0<\alpha<1$, que se representa por q_α, é o valor do conjunto de dados que tem αn observações inferiores e $(1-\alpha)n$ observações superiores.

A mediana é um quantil de ordem $\alpha=0,5$.

Outros quantis importantes são os percentis, obtidos quando $\alpha=0,01;0,02;\ldots;0,99$, os decis, quando $\alpha=0,1;0,2;\ldots;0,9$, o primeiro quartil, $Q1=q_{0,25}$, e o terceiro quantil, $Q3=q_{0,75}$.

Estas medidas aplicam-se a dados quantitativos (discretos ou contínuos).

Apresenta-se, de seguida, as técnicas para determinação dos quantis, quando se têm o conjunto de dados ordenados, isto é, os dados estão na forma

$x_{(1)},x_{(2)},\ldots,x_{(n)}$.

Como existem n-1 intervalos unitários entre 1 e n, a ordem da observação correspondente ao quantil de ordem α, q_α, será dada por $1+(n-1)\alpha$ se este valor for inteiro. Assim, fazendo $r=1+(n-1)\alpha$, o quantil de ordem α é $q_\alpha=x_{(r)}$.

Se $1+(n-1)\alpha$ não for inteiro, tomar-se-á $1+(n-1)\alpha=r+\varepsilon$, onde r é um número inteiro e $0<\varepsilon<1$. Neste caso, o valor $r+\varepsilon$ pode ser considerado a ordem aproximada da observação correspondente ao quantil. Contudo, muitas vezes, o quantil de ordem α obtém-se por "interpolação linear", fazendo a seguinte média ponderada entre $x_{(r)}$ e $x_{(r+1)}$:

$q_\alpha=x_{(r)}+\varepsilon(x_{(r+1)}-x_{(r)})=(1-\varepsilon)x_{(r)}+\varepsilon x_{(r+1)}$.

Nota. Não se deve definir o quantil de ordem α quando $\alpha<1/(n+1)$ ou $\alpha>n/(n+1)$, ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 26).

Apresenta-se agora, um mecanismo para determinação dos quartis Q1 e Q3, pois mais adiante esses valores serão necessários no cálculo de medidas de dispersão.

Definição (Quartis). São os valores que dividem um conjunto de dados em quatro partes iguais.

Uma vez ordenado os dados, o segundo quartil Q2 – que também coincide com a mediana – é o valor que fica no meio dos valores dos elementos do conjunto de dados, isto é, o valor que divide o conjunto de dados em duas partes iguais. Depois o primeiro quartil, Q1, será o valor que fica no meio na primeira metade do conjunto de dados, e o terceiro quartil, Q3, será, analogamente, o valor que fica a meio da segunda metade do conjunto de dados.

Se os dados forem discretos, normalmente, a localização dos quartis far-se-á por intermédio de dois métodos: inclusivo e exclusivo. (para mais informações ver (Fernandes & Pinto)).

No entanto, a tabela que se segue baseia-se no método "CDF-Cumulative distribuition function" (ver (Fernandes & Pinto, p. 4)). Considera-se uma amostra de dimensão n (par ou ímpar). Ao dividir n por 4, o resto da divisão é 0, 1, 2 ou 3, ou seja, n=4m ou n=4m+2, caso n seja par, ou n=4m+1 ou n=4m+3, caso n seja ímpar.

		Localização de	Localização de
n par		k=(n+2)/4	k=(3n+2)/4
ímpar	n=4m+1	k=(n+3)/4	k=(n+1)/4
	n=4m+3	k=(3n+1)/4	k=(3n+3)/4

Conhecida a localização, o cálculo dos quartis faz-se baseando na seguinte tabela:

k inteiro	$Qp = x_k$
k não inteiro, $i < k < i + 1$	$Qp = \frac{x_i + x_{i+1}}{2}$

Exemplo. Os dados abaixo indicam a população ativa em percentagem da população residente para cada um dos concelhos de um determinado país:

C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
30.87	33.29	36.69	33.7	34.11	33	38.85	34.43	31.91	32.93	35.75	30.73	36.15	40.28

Quadro 8: População ativa em percentagem da população residente para cada um dos concelhos de um determinado país

Determina-se, assim, a mediana x ̃=Q2, o primeiro quartil Q1 e o terceiro quartil Q3.

Ora, é necessário ordenar os valores:

C12	C1	C9	C10	C6	C2	C4	C5	C8	C11	C13	C3	C7	C14
$x_{(1)}$	$x_{(2)}$	$x_{(3)}$	$x_{(4)}$	$x_{(5)}$	$x_{(6)}$	$x_{(7)}$	$x_{(8)}$	$x_{(9)}$	$x_{(10)}$	$x_{(11)}$	$x_{(12)}$	$x_{(13)}$	$x_{(14)}$
30.73	30.87	31.91	32.9	33	33	33.7	34.11	34.43	35.75	36.15	36.69	38.85	40.28

Sendo assim, tem-se:

x ̃=Q2=(x_7+x_8)/2=(33.7+34.11)/2≅33.91

Q1=x_((4))=32.9

Q3=x_((11))=36.15

Se os dados forem contínuos, a localização dos quartis far-se-á por intermédio das fórmulas que se seguem.

Determinação de Q1:

Calcular-se-á a ordem n/4;

Pela coluna da frequência acumulada, determinar-se-á a classe que conterá Q1;

Utilizar-se-á a fórmula:

Q1=l_Q1+((n/4-f_(ac_aQ1))×h)/F_Q1

Determinação de Q3

Calcular-se-á a ordem 3n/4;

Pela coluna da frequência acumulada, determinar-se-á a classe que conterá Q;

Utilizar-se-á a fórmula:

Q3=l_Q3+((3n/4-f_(ac_aQ3))×h)/F_Q3

Exemplo. Calcula-se a mediana e os quartis e do peso de cigarros de uma certa marca, cuja a distribuição de frequências apresenta-se de acordo com o quadro abaixo:

Peso (mg)	F_i	$F.Ac.$
]760,780]	4	4
]780,800]	43	47
]800,820]	118	165
]820,840]	168	333
]840,860]	117	450
]860,880]	39	489
]880,900]	11	500

Quadro 9: "Peso de cigarros de uma certa marca"

Ora, a classe mediana é , então a mediana é:

x ̃=l_Md+((n/2-f_(ac_ca))·h)/F_Md =820+((250-165))/168×20≅830.119 .

A classe que contém Q1 é]800,820], então o primeiro quartil é:

Q1=l_Q1+((n/4-f_(ac_aQ1))×h)/F_Q1 =800+((125-47))/118×20≅813.220 .

A classe que contém Q3 é]840,860], então o terceiro quartil é:

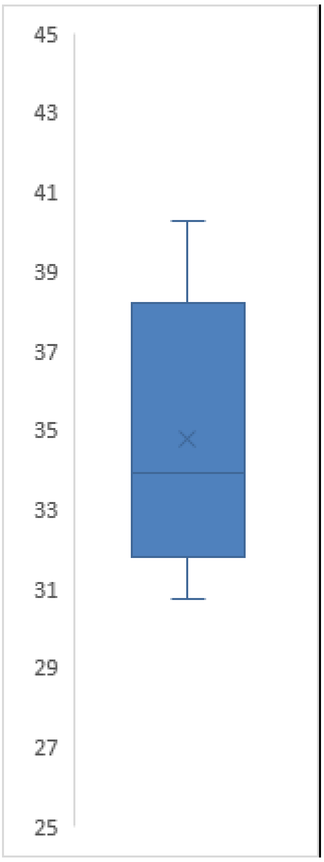
Q3=l_Q3+((3n/4-f_(ac_aQ3))×h)/F_Q3 =840+((375-333))/117×20≅847.179 .

Costuma-se utilizar uma representação gráfica, denominada caixa-de-bigodes, para representar os extremos, a mediana e os quartis Q1 e Q3, ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 27).

O retângulo (a "caixa") é desenhado de forma que os seus lados inferior e superior correspondam aos primeiro e terceiro quartis, respetivamente; o segmento no seu interior refere-se à mediana. O mínimo e o máximo (min(x_i) e max(x_i)) são representados pelos segmentos inferior e superior ("os bigodes"), desenhados no exterior do retângulo.

A caixa nos seus limites horizontais (entre Q1 e Q3) contém 50% das observações (ou dados).

Exemplo. A caixa-de-bigodes da "população ativa, em percentagem, da população residente para cada um dos 14 concelhos de um determinado país" (ver exemplo anterior) está representada abaixo.



Definição (Moda). Considerando os dados originais, a moda, que se passa a representar por , é a observação com maior frequência absoluta.

Se nenhuma observação é repetida, o conjunto de dados não tem moda.

Se duas observações ocorrem com a mesma frequência, cada uma dessas observações é uma moda, e o conjunto é chamado bimodal.

Há conjunto trimodal (3 modas) e plurimodal (com mais de 3 modas).

A moda pode ser determinada para qualquer tipo de atributo em causa (qualitativo, quantitativo discreto ou contínuo).

Se os dados estiverem organizados em classes de igual amplitude, a classe com maior frequência chamar-se-á classe modal.

Nesse caso, a moda pode ser calculada aplicando algumas fórmulas, das quais destaca-se, neste módulo, a fórmula de Czuber:

Identifica-se a classe modal, ou seja, a classe que tem maior frequência;

Aplica-se a fórmula

Mo=l_Mo+Δ_1/(Δ_1+Δ_2) h

onde:

– Limite inferior da classe modal;

– Diferença entre a frequência da classe modal e a da classe anterior a ela;

– Diferença entre a frequência da classe modal e a da classe posterior a ela.

– Amplitude das classes.

Exemplo. Calcula-se a moda do peso de cigarros de uma certa marca, cuja a distribuição de frequências apresenta-se de acordo com o quadro "Peso de cigarros de uma certa marca".

Ora, a classe modal é]820,840], então a moda é:

$Mo=l_{Mo}+\frac{\Delta_1}{\Delta_1+\Delta_2}h=820+\frac{50}{50+51}\times 20\cong 829.901$

Nota. A média e a mediana são medidas para dados quantitativos, enquanto que a moda pode aplicar-se a dados quantitativos e qualitativos.

Medidas de dispersão (ou variação). São valores utilizados para determinar a característica de variação de um conjunto de dados estatísticos. Deve definir-se tomando em conta a posição das observações em relação a uma referência fixa, que deve ser o valor escolhido para localizar a distribuição. São eles:

- Amplitude total;
- Desvio;
- Desvio médio, ou absoluto;
- Variância;
- Desvio padrão.

Amplitude total, definida anteriormente como $\max(x_i)-\min(x_i)$, é uma das medidas de dispersão, mas é muito limitada, pois depende apenas dos valores extremos, ou seja, não é afetada pela dispersão dos valores internos.

Definição (Desvio). É a diferença entre cada valor de uma observação (x_i), para todo i, e a média ($\bar{x}$,ou μ).

Quanto menor for os desvios, menos dispersos estão as observações em relação à média.

Da definição da média, obtém-se a seguinte propriedade:

$\sum_{i=1}^{n}(x_i-\bar{x})=0$ (dados amostral) e $\sum_{i=1}^{N}(x_i-\mu)=0$ (dados populacional)

propriedade que se anuncia da seguinte forma: a soma dos desvios em relação a média é igual a zero.

Definição (Desvio médio). É a média dos valores absolutos dos desvios, ou seja,

$\frac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n}$(dados amostral) e $\frac{\sum_{i=1}^{N}|x_i-\mu|}{N}$ (dados populacional).

Definição (Variância). É a média (aritmética) dos quadrados dos desvios, ou seja,

$s^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$ (Variância amostral) e $s^2=\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}$

(Variância populacional).

Definição (Desvio padrão (DP)). É a raíz quadrada positiva da variância, ou seja,

$s=\sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}}$ (DP amostral) e $s=\sqrt{\frac{\sum_{i=1}^{N}(x_i-\mu)^2}{N}}$ (DP populacional).

Nota. As fórmulas de variância e desvio padrão são utilizadas para dados quantitativos simples e agrupados em classes. No caso de dados em classes, os x_i representam as marcas das classes.

Estas fórmulas podem ser simplificadas, de uma certa forma; por exemplo:

$s^2=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}=\frac{1}{n}\sum_{i=1}^{n}(x_i^2-2x_i\bar{x}+\bar{x}^2)=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$,ou

$s^2=\frac{\sum_{i=1}^{n}F_i(a_i-\bar{x})^2}{n}=\sum_{i=1}^{m}f_j a_j^2-\bar{x}^2$ (com tabela de frequências).

Nota. Por vezes, utilizam-se a variância e o desvio padrão corrigidos, para dados amostrais específicos, ver " (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 22)":

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

Se se usar a tabela de frequências absolutas e a de frequências relativas, obter-se-á:

$s^2=\frac{\sum_{i=1}^{n}F_i(a_i-\bar{x})^2}{n}=\sum_{i=1}^{n}f_i(a_i-\bar{x})^2$ e $s^2=\frac{\sum_{i=1}^{N}F_i(a_i-\mu)^2}{N}=\sum_{i=1}^{N}f_i(a_i-\mu)^2$.

O desvio padrão é de grande utilidade na medição da dispersão em relação à média aritmética. Quanto mais dispersos são os dados em relação à média, maior é o desvio padrão.

Teorema (de Chebichev). A proporção de qualquer conjunto de dados, a menos de k, $k\in\mathbb{N}$ e $k>1$, desvios padrão a contar da média, é sempre ao menos

$1-\frac{1}{k^2}$.

Ou seja, o intervalo

$[\bar{x}-ks,\bar{x}+ks]$

tem ao menos

$\left(1-\frac{1}{k^2}\right)\times 100$

porcentos dos dados.

Quando $k=2$, o intervalo $[\bar{x}-2s,\bar{x}+2s]$ terá ao menos 75% ddos dados.

Quando $k=3$, o intervalo $[\bar{x}-3s,\bar{x}+3s]$ terá ao menos 88,9% dos dados, aproximadamente.

Definição (Outliers). São observações "estranhas" que se localizam muito distantes da média.

As estatísticas descritivas são muito influenciadas por outliers.

O Teorema de Chebichev, dá uma boa contribuição na identificação dos outliers, ou seja, valores fora do intervalo $[\bar{x}-2s,\bar{x}+2s]$ – são ditos valores incomuns, ou muito incomuns, caso estejam fora do intervalo $[\bar{x}-3s,\bar{x}+3s]$ – devem ser analisados, pois podem ser outliers.

Definição (Amplitude interquartis). É a diferença entre Q3 e Q1, e representa-se por AIQ, isto é, AIQ=Q3-Q1.

A AIQ é utilizada para classificar os outliers. Uma observação x_i é um outlier severo quando $x_i<Q1-3(Q3-Q1)$ ou $x_i>Q3+3(Q3-Q1)$. Uma observação x_i é um outlier moderado quando $Q1-3(Q3-Q1)<x_i<Q1-1.5(Q3-Q1)$ ou $Q3+1.5(Q3-Q1)<x_i<Q3+3(Q3-Q1)$.

Definição ("Score" padronizado (z)). É o número de desvio padrão pelo qual uma observação se dista, para mais ou para menos, da média. Para todo i, score padronizado de x_i é dado por:

$z_i=(x_i-\bar{x})/(n)$ (dados amostra) e $z_i=(x_i-\mu)/(N)$ (dados populaconal) .

Nota. São considerados incomuns as observações com "scores" acima de dois desvios padrão da média.

Definição (Coeficiente de variação (CV)). É um valor, em percentagem, que indica a magnitude relativa do desvio padrão quando comparado com a média do conjunto de valores, ou seja,

$CV=s/\bar{x}\times 100$ (dados amostral) e $CV=s/\mu\times 100$ (dados populacional) .

Nota. Coeficiente de variação é útil para comparar a dispersão de dois, ou mais, conjuntos de dados, de ordem de grandezas diferentes. Quanto maior é a dispersão, maior é o coeficiente de variação.

Exemplo. Seja o conjunto de preços (em USA Dólares) de geladeiras em 7 lojas diferentes,

Loja	A	B	C	D	E	F
Preço	750	800	810	820	760	780

com media é $\bar{x}=787{,}14$ e com o desvio padrão $s=25{,}63$.

Seja também, o conjunto de preços de liquidificadores nas mesmas lojas acima,

Loja	A	B	C	D	E	F
Preço	50	45	55	43	52	54

com média $\bar{x}=49{,}14$ e com desvio padrão $s=4{,}81$.

Qual dos produtos tem maior variabilidade de preços, em relação à média?

Ora, o coeficiente de variação do primeiro conjunto de dados é

$CV=25{,}63/787{,}14\times 100\cong 3{,}3\%$.

O coeficiente de variação do segundo conjunto de dados é

$CV=4{,}81/49{,}14\times 100\cong 9{,}8\%$.

Conclui-se desta forma, que há menor dispersão de preços no conjunto dos preços de geladeiras.

Conclusão

As medidas de localização mais utilizadas são: a média, a mediana e a moda. A média é aplicada apenas aos dados quantitativos, a mediana aplica-se aos dados quantitativos e aos dados qualitativos com nível de mensuração ordinal, e a moda aplica-se aos dados quantitativos e qualitativos.

A variância e o desvio padrão são medidas de dispersão em relação à média.

Para muitos as medidas de dispersão mais usadas são variância e desvio padrão.

Avaliação

1. Os dados que se seguem referem-se à precipitação caída em Lisboa, em , no mês de janeiro dos anos de 1980 a 2006. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 49).

ANO	PREC.	ANO	PREC.	ANO	PREC.	ANO	PREC.	ANO	PREC.
1980	53.5	1986	52.2	1992	61.5	1998	62.3	2004	129.5
1981	11.6	1987	130	1993	19	1999	107.5	2005	0
1982	112.8	1988	107.7	1994	79.7	2000	17.7	2006	57.7
1983	6.2	1989	62.5	1995	57.3	2001	195.5		
1984	40.6	1990	63.1	1996	394	2002	72.9		
1985	224.6	1991	64.9	1997	136.6	2003	131		

a. Calcule a média, a variância e o coeficiente de variação para estas observações.

b. Calcule a variância corrigida e o desvio padrão corrigido.

c. Apresente o resumo dos “5 números” (os 2 extremos, a mediana e os quartis Q1 e Q3.

d. Construa uma caixa-de-bigodes para esta informação e identifique eventuais outliers.

e. Indique os scores dos eventuais outliers, apurados na alínea anterior.

2. Um estudo sobre os atrasos nos voos europeus durante o Verão de 2006, realizado em determinado aeroporto, conduziu aos seguintes resultados:

Atraso	]0,10]	]10,20]	]20,30]	]30,40]	]40,50]	]50,60]
N. º de voos	29	23	17	14	11	6

a. Construa um histograma.

b. Construa o polígono de frequência, e classifique a distribuição (simétrica, uniforme ou assimétrica).

c. Calcule a média, a variância e o desvio padrão referentes aos minutos de atraso de voos.

d. Calcule a moda e a mediana.

e. Determine o intervalo interquartis e a respetiva amplitude.

f. Determine a proporção de dados a menos de 2 e de 3 desvios padrões.

3. Abaixo estão listados os salários anuais (em milhares de dólares) para os funcionários municipais de duas, determinadas, cidades A (e) B de África:

Cidade *A*	20.2	26.1	20.9	32.1	35.9	23	28.2	31.6	18.3
Cidade *B*	20.9	18.2	20.8	21.1	26.5	26.9	24.2	25.1	22.2

a. Determine as médias e os desvios padrões das duas distribuições.

b. Diga qual dos dois conjuntos de dados apresenta maior variabilidade de dados, emrelação à média aritmética.

4. Numa sala estão 5 pessoas. A média das idades é de 30 anos. Entra na sala uma pessoa de 36 anos de idade. Agora, qual será a média das idades das pessoas naquela sala?

5. A tabela que se segue mostra as ligações de longa distância (em minutos) feitas por uma pessoa em um dia:

Duração	]0,5]	]5,10]	]10,15]	]15,20]	]20,25]	]25,30]	]30,35]	]35,40]	]40,45]
N. º de Chamadas	12	26	20	7	11	7	4	4	1

Determine a média, a moda e a mediana.

Determine o desvio padrão e o coeficiente de variação.

Determine Q1 e Q2.

Resumo da Unidade

As três principais características de um conjunto de dados estatísticos são um valor representativo de um conjunto de dados (medidas de tendência central), uma medida de dispersão ou variação e a natureza ou forma da distribuição de dados.

Para análise e interpretação dos dados é necessário organizar os dados em tabelas ou em gráficos ou pictogramas.

Leituras e outros Recursos

As leituras e outros recursos desta unidade encontram-se na lista de Leituras e Outros Recursos do curso.

Leituras.

- Fernandes, S., & Pinto, M. (s.d.). Afinal o que são e como se calculam os quartis? Universidade de Algarve - Departamento de Matemática, Departamento de Matemática, Faro.
- Larson, R., & Farber, B. (2010). Estatística Aplicada (4ª ed.). São Paulo: Pearson Prentice Hall.
- Murteira, B., Ribeiro, C. S., Silva, J. A., & Pimenta, C. (2002). Introdução à Estatística (2ª ed.). Lisboa: McGraw-Hill de Portugal.

Outros Recurso.

- O Software Excel.
- Máquina Científica: TI-83 ou semelhante.
- Software: Geogebra, R e Excel.

Unidade 2. Probabilidade e teoremas associados

Introdução à Unidade

Nesta unidade introduz-se os conceitos de experiência aleatória, espaço de resultados e de acontecimentos, a medida de probabilidade. Também, faz-se uma análise cuidadosa do conceito de probabilidade condicionada e do Teorema de Bayes.

Sempre que se menciona probabilidade, sem especificar o tipo, está-se a referir a probabilidade clássica (um dos tipos de medida de probabilidade).

Objetivos da Unidade

Após a conclusão desta unidade, deverá ser capaz de:

- Definir e identificar uma experiência aleatória.
- Determinar o espaço-amostra de uma experiência aleatória.
- Operar com eventos num espaço-amostra.
- Determinar probabilidade condicional.
- Determinar probabilidade com eventos independentes.
- Resolver problemas da vida quotidiana aplicando a teoria de probabilidade.

Termos-chave

Experiência Aleatória. É uma experiência estatística em que se não prevê o resultado de antemão.

Espaço-amostra (ou espaço de resultados). É o conjunto fundamental, e não vazio, formado por todos os resultados que hipoteticamente é possível obter quando se efetua determinada experiência aleatória.

Probabilidade condicional (ou condicionada). Probabilidade de um evento sob a condição de que se realiza um outro acontecimento.

Independência entre dois eventos. Acontece quando a probabilidade de interseção de dois eventos é igual ao produto das probabilidades desses eventos.

Atividades de Aprendizagem

Atividade 1 – Espaço-amostra (ou espaço de resultado) e Acontecimento. Probabilidade e seus axiomas.

Introdução

Nesta atividade vai introduzir-se os conceitos de experiência aleatória, espaço de amostra e de acontecimento (operações com acontecimentos). Introduz-se algumas definições de probabilidade, e anuncia-se os axiomas de probabilidade, bem como as suas consequências.

Detalhes da atividade

Definição (Experiência Aleatória). Diz-se que uma experiência é aleatória se satisfazer as seguintes características:

- O Conjunto de resultados possíveis é afixado anteriormente;
- O resultado da experiência nunca pode ser previsto de forma exata, mesmo que se desenvolvam todos os esforços para manter sob controlo as circunstâncias relevantes para o resultado.

Exemplo. Constituem exemplos de experiências aleatórias, os que se seguem:

- Lançamento de uma moeda e observação do lado que fica para cima;
- Lançamento de um dado e registo do número de pontos obtidos;
- Tiragem de uma carta de um baralho e anotação de um naipe;
- Registo do número de sinistros por apólice do ramo de automóvel durante uma anuidade

Def**inição (Espaço-amostra (Ou espaço de resultados))**. Denomina-se espaço de resultados ou (espaço-amostra), e representa-se por Ω, o conjunto fundamental, e não vazio, formado por todos os resultados que hipoteticamente é possível (resultado elementar e indivisível de uma experiência) obter quando se efetua determinada experiência aleatória. Cada elemento de Ω é representado por ω.

Exemplo. Considera-se uma experiência aleatória que consiste no lançamento de dois dados e observar os números obtidos nas faces voltadas para cima. Definir um espaço-amostra para essa experiência aleatória.

Com são dois dados, o espaço-amostra tem que albergar todos os resultados possíveis. Para cada lançamento ter-se-á dois números, pois são dois dados. Logo:

$$\Omega = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),(2,4),(2,5),(2,6) \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),(4,1),(4,2),(4,3),(4,4),(4,5),(4,6) \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),(6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}$$

$$= \{(i,j) : i,j = 1,2,3,4,5,6\}$$

Ou seja, o espaço-amostra tem 36 elementos.

Exemplo. Considera-se uma experiência aleatória que consiste em escolher aleatoriamente um ponto de um círculo, do plano euclidiano, com centro na origem e raio igual a 1.

Neste caso, o nosso espaço-amostra é um subconjunto de R^2:

Ω={(x,y)∈R^2 :x^2+y^2≤1} .

Definição (Acontecimento (ou evento)). Chama-se acontecimento (ou evento) a um subconjunto do espaço-amostra.

Exemplo. Ao lançar um dado e observar o resultado, podem ser considerados eventos os seguintes:

Observar um número par;

Observar um número maior ou igual a dois;

Observar o número quatro.

Definição (Realização de um acontecimento). Ao efetuar uma experiência aleatória associada a Ω, diz-se que o acontecimento $A\subset\Omega$ se realiza, se o resultado da experiência ω é um ponto pertencente a A: $\omega\in A$.

Os subconjuntos $\{\omega\}\subset\Omega$, Formados por um único elemento de Ω, dizem-se eventos (ou acontecimentos) elementares.

Como Ω é subconjunto de si próprio, tem-se que Ω é um acontecimento (ou evento), acontecimento certo.

Como ∅ é um subconjunto de Ω, então ∅ é um acontecimento, acontecimento impossível.

Eventos possivelmente realizáveis são chamados eventos aleatórios.

Nota. Como os eventos são subconjuntos de Ω, então a álgebra dos acontecimentos são muito parecidos com a álgebra dos conjuntos.

- Implicação de acontecimentos. O acontecimento A implica o acontecimento B se, e somente se, qualquer elemento de A é, também, um elemento de B, ou seja, $A\subset B$;
- Identidade de acontecimentos. Dois acontecimentos A e B são idênticos se, e somente se, possuem os mesmos elementos, ou seja, A=B.
- União de acontecimentos. União de dois acontecimentos A e B, $A\cup B$, é o acontecimento que se realiza quando pelo menos um deles se realiza.
- Intersecção de acontecimentos. A intersecção de dois acontecimentos A e B, $A\cap B$, é o acontecimento que se realiza quando ambos se realizam.

- Acontecimentos incompatíveis (ou mutuamente exclusivos). Dois acontecimentos A e B são incompatíveis ou (mutuamente exclusivos) se, e somente se, a realização de um implica a não realização do outro. Por outras palavras, os acontecimentos A e B são incompatíveis se, e somente se, $A \cap B$ é um acontecimento impossível, isto é, $A \cap B = \emptyset$.
- Diferença de acontecimentos. A diferença entre os acontecimentos A e B, A-B, é o acontecimento que se realiza se, e somente se, A se realiza sem que B se realize, isto é, $A - B = A \cap \overline{B}$.

Em particular, Ω-A, designa-se por acontecimento contrário (ou complementar) de A, e realiza-se se, e somente se, A não se realiza. Normalmente, representa-se Ω-A por $\overline{A}$, e $A \cap \overline{A} = \emptyset$ e $A \cup \overline{A} = \Omega$.

Algumas propriedades das operações definidas sobre acontecimentos.

$A \cup (B \cup C) = (A \cup B) \cup C$; $A \cap (B \cap C) = (A \cap B) \cap C$ – Associatividade.

$A \cup B = B \cup A$; $A \cap B = B \cap A$ – Comutatividade.

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$; $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ – Distributividade.

$\overline{(A \cup B)} = \overline{A} \cap \overline{B}$; $\overline{(A \cap B)} = \overline{A} \cup \overline{B}$ – Leis de De Morgan.

Exemplo. Designando por A_1, A_2 e A_3 acontecimentos quaisquer de Ω. Tem-se:

Ocorrência de A_1 e não ocorrência de A_2 e A_3: $A_1 \cap \overline{A_2} \cap \overline{A_3}$.

Ocorrência de pelo menos um dos A_j, $j=1,2,3$: $A_1 \cup A_2 \cup A_3$.

Ocorrência de um e um só dos $A_j, j=1,2,3$:

$(A_1 \cap \overline{A_2} \cap \overline{A_3}) \cup (A_2 \cap \overline{A_1} \cap \overline{A_3}) \cup (A_3 \cap \overline{A_1} \cap \overline{A_2})$.

Ocorrência de exatamente dois dos $A_j, j=1,2,3$:

$(A_1 \cap A_2 \cap \overline{A_3}) \cup (\overline{A_2} \cap A_1 \cap A_3) \cup (A_3 \cap \overline{A_1} \cap A_2)$.

Ocorrência de nenhum dos $A_j, j=1,2,3$: $\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} = \overline{(A_1 \cup A_2 \cup A_3)}$.

Definição (Medida de Probabilidade). A medida de probabilidade é uma função P que a cada acontecimento $A \subset \Omega$, faz corresponder um número real P(A), probabilidade do acontecimento A, que verifica os três axiomas seguintes:

$P(A) \geq 0$.

$P(\Omega) = 1$.

Se A e B são acontecimentos incompatíveis, então $P(A \cup B) = P(A) + P(B)$.

Nota. A propriedade (3) pode ser substituída por:

(3*) Se $A_1, A_2, A_3, \ldots, A_n, \ldots$ são acontecimentos em número infinito numerável, dois a dois incompatíveis, isto é, $A_i \cap A_j = \emptyset, i \neq j$, então

$$P\left(\bigcup_{j=1}^{+\infty} A_j\right) = \sum_{j=1}^{+\infty} P(A_j) .$$

Propriedades de medida de probabilidade.

$P(A)=1-P(\bar{A})$.

$P(\emptyset)=0$.

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

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

$A\subset B\Rightarrow P(A)\leq P(B)$.

$P(A)\leq 1$.

$P(A\cap \bar{B})=P(A)-P(A\cap B)$.

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

$+\,P(A\cap B\cap C)$.

Exemplo. Em determinada população, 9,8% das pessoas adquirem a revista A, 22,9%, a revista B, e 5,1%, ambas as revistas. Admite-se que a medida de probabilidade é a proporção dos indivíduos da população que adquirem as revistas. Deste modo, definem-se os acontecimentos:

A≡adquirir a revista A;

B≡adquirir a revista B.

A probabilidade de adquirir somente a revista A, acontecimento $A\cap\bar{B}=A-B$, é dada por:

$$\begin{aligned} P(A-B) &= P(A)-P(A\cap B) \\ &= 0{,}098-0{,}051 \\ &= 0{,}047 \end{aligned}$$

A probabilidade de uma pessoa escolhida ao acaso adquirir pelo menos uma das revistas, acontecimento $A\cup B$, é dada por:

$$\begin{aligned} P(A\cup B) &= P(A)+P(B)-P(A\cap B) \\ &= 0{,}098+0{,}229-0{,}051 \\ &= 0{,}276 \end{aligned}$$

A probabilidade de não adquirir nem a revista A nem, nem a revista B, acontecimento $\bar{A}\cap\bar{B}$, é dada por:

$$\begin{aligned} P(\bar{A}\cap\bar{B}) &= 1-P(\overline{(\bar{A}\cap\bar{B})})=1-P(A\cup B) \\ &= 1-0{,}276 \\ &= 0{,}724 \end{aligned}$$

Nota. Os exemplos anteriores não ensinam como calcular a probabilidade de um acontecimento. Aliás, essa parte ainda não foi abordada. O que se mostrou ali é como calcular probabilidades de determinados acontecimentos a partir de probabilidades de outros acontecimentos.

Interpretação do conceito de probabilidade

Definição (Clássica). Seja $\Omega=\{\omega_1,\omega_2,\ldots,\omega_n\}$ um espaço de resultado finito, e seja um acontecimento $A=\{\omega_{\alpha_1},\omega_{\alpha_2},\ldots,\omega_{\alpha_m}\}$. Então a probabilidade do acontecimento A é dada por:

$$P(A)=\frac{m}{n}=\frac{\#(A)}{\#(\Omega)}\ .$$

Nota (Importante!). Esta definição baseia-se em definição de resultados equiprováveis, ou de outro modo, em princípio de simetria, ou princípio da indiferença, isto é, todos os resultados são igualmente possíveis, como dois lados de uma moeda, as seis faces de um dado, 52 duas cartas de um baralho, etc.

O cálculo de probabilidade neste caso, resume-se à contagem de número de casos favoráveis e do número de casos possíveis. A probabilidade do acontecimento A, será dada então por:

$$P(A)=\frac{\text{n.º de casos favoráveis a A}}{\text{n.º de casos possíveis}}\ .$$

Exemplo. Seja a experiência aleatória que consiste no lançamento de duas moedas "perfeitas".

Qual é a probabilidade do acontecimento A, saída de exatamente uma face?

Os resultados desta experiência aleatória são equiprováveis, então, os casos possíveis são FF, FC, CF e CC, onde F representa a face e C a coroa de uma moeda. Os casos favoráveis são FC e CF. Então

$$P(A)=\frac{2}{4}=0{,}5\ .$$

Definição (Frequencista). Baseia-se na frequência relativa de um "número grande" de realizações de uma experiência. Ou seja, define-se a probabilidade P(A) de um acontecimento A usando o limite da frequência relativa de A em N repetições independentes da experiência, com N tendendo ao infinito, ou seja,

$$P(A)=\lim_{N\to\infty}\frac{1}{N}F_N(A)\ ,$$

onde $F_N(A)$ é o número de ocorrências de A em N realizações independentes da experiência (tipo frequência absoluta).

Nota. Um dos grandes constrangimentos da definição Frequencista é que uma experiência aleatória não pode ser realizada uma infinidade de vezes, logo não se pode avaliar a probabilidade de forma estrita.

A definição Frequencista é particularmente intuitiva se a experiência aleatória puder ser repetida um grande número de vezes, nas mesmas condições, ou em condições semelhantes.

Definição (Geométrica). Seja uma experiência aleatória que consiste na escolha de um ponto ao acaso numa região de $\Omega \subseteq \mathbb{R}^p$, $p \in \{1,2,3\}$. A probabilidade, P(A), de um evento $A \subseteq \Omega$ é:

P(A)=(Volume de A)/(Volume de Ω) ,

onde volume significa:

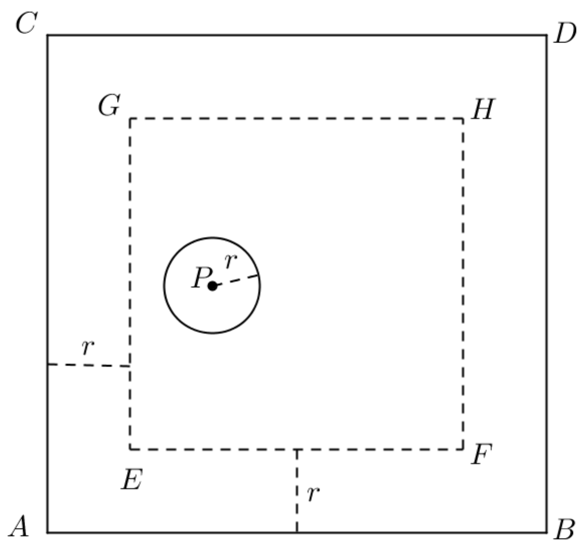
Volume geométrico se p=3;

Área geométrica se p=2;

Comprimento geométrico se p=1.

Exemplo (Jogo de franc-carreau). O jogo consiste em lançar uma moeda em um piso de azulejos de forma quadrada. Os jogadores então apostam se a moeda irá parar completamente sobre um único azulejo, posição chamada franc-carreau, ou sobrepor algum trecho de rejunte (junção dos azulejos). Em uma região de N azulejos de lado de comprimento igual a b centímetros, qual é a probabilidade de uma moeda, de raio r centímetros, parar na "posição franc-carreau"?

Considera-se a figura seguinte:



Onde:

O quadrado ABCD representa um azulejo;

O Círculo de centro P e raio r representa a moeda.

Portanto, a moeda estará na "posição de franc-carreau" se o centro P do círculo, de raio r, estará no interior do quadrado EFGH.

Então, utilizando a definição geométrica, a probabilidade procurada é:

(Soma das áreas dos quadrados EFGH)/(Soma das áreas dos quadrados ABCD)=(N(b-2r)^2)/(Nb^2)=(b-2r)^2/b^2 .

Definição (Subjectiva). Esta definição de probabilidade baseia-se em crenças e/ou informações do observador acerca do fenómeno em causa.

A probabilidade subjectiva constitui, assim, uma aproximação quantitativa à credibilidade que se atribui a determinado acontecimento.

Exemplo. Considere o evento A que se segue: "com este plano de marketing a empresa E pode alcançar dentro de dois anos um volume de vendas de y milhares de USA dólares.

O valor de y pode ser associado à uma probabilidade e, essa, depende das informações que o observador tem acerca da empresa e do plano de marketing em causa, por exemplo.

Definição (Amostra ordenada). Se k elementos são selecionados de um conjunto de n elementos e se a ordem da seleção é registada, o conjunto de k elementos forma uma amostra ordenada de dimensão k.

Definição (Amostra não ordenada). Se k elementos são seleccionados de um conjunto de n elementos e se a ordem da selecção não importa, o conjunto de k elementos forma uma amostra ordenada de dimensão k.

Definição (Amostragem sem reposição). Quando se selecciona um elemento e o mesmo não é posto no conjunto, diz-se que se seguiu um processo de amostragem sem reposição.

Definição (Amostra com reposição). Um processo de amostragem com reposição ocorre quando se selecciona um elemento e o mesmo é reposto no conjunto antes de se seleccionar o elemento seguinte.

Nota. Nestes tipos de amostragens, são fundamentais os conceitos de: princípio fundamental da contagem, Arranjos (com e sem repetição) – para amostragem ordenada com e sem reposição, permutações, combinações – para amostragem não ordenada sem reposição e permutações de elementos não todos distintos.

Conclusão

Uma experiência aleatória diferencia-se de qualquer outra, porque o resultado não pode ser previsto.

As regras de operações com eventos são semelhantes às de operações com conjuntos.

Cada uma das quatro definições de probabilidade, pode ser aplicado num contexto específico. No entanto, nesta atividade dá-se mais realce ao cálculo de probabilidade clássica e, de vez em quando, ao de probabilidade geométrica (no fundo um caso particular da definição clássica).

Avaliação

1. Sejam A e B dois eventos de um espaço de amostra Ω. Sabendo que $P(A)=0,3$, $P(B)=0,2$ e $P(A\cap B)=0,1$, calcule:

   a. $P(\bar{A})$;

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

   c. $P(A\cap \bar{B})$;

   d. $P(\overline{(A\cup B)})$.

2. Considere as duas principais variedades de óleos de cozinhas: os monoinsaturados e os polinsaturados. Duas matérias-primas para óleos de cozinha são, por exemplo, milho e canola. A tabela a seguir mostra o número de garrafas destes óleos em um supermercado:

		Tipo de óleo	
		Milho	Canola
Tipo de insaturação	Monoinsaturado	13	7
	Polinsaturado	77	93

   a. Se uma garrafa de óleo é escolhida aleatoriamente, qual é a probabilidade de serum óleo polinsaturado?

   b. Se uma garrafa de óleo é selecionada aleatoriamente, qual é a probabilidadede ser monoinsaturado e de canola?

3. Um certo tipo de motor elétrico falha se ocorrer uma das seguintes situações: emperramento dos mancais, queima dos rolamentos ou desgaste das escovas. Suponha que o emperramento seja duas vezes mais provável que a queima, esta sendo quatro vezes mais provável do que o desgaste das escovas. Qual é a probabilidade de que a falha seja devida a cada uma destas circunstâncias?

4. Suponha que A, B, C sejam eventos de um determinado espaço de amostra, tais que $P(A)=P(B)=P(C)=1/4$, $P(A\cap B)=P(B\cap C)=0$ e $P(A\cap C)=1/8$. Calcule a probabilidade que ao menos um dos eventos, A, B ou C, ocorra?

5. Qual é a probabilidade de observar quatro números diferentes, no lançamento de quatro dados?

6. Suponha que A, B, C sejam eventos, disjuntos dois a dois, de um determinado espaço de amostra. É possível ter $P(A)=0,3$, $P(B)=0,4$ e $P(C)=0,5$? Porquê?

7. Uma caixa tem 24 lâmpadas, dos quais 4 são defeituosas. Se uma pessoa seleciona 4 lâmpadas aleatoriamente, qual é a probabilidade das 4 lâmpadas serem defeituosas?

8. N pessoas irão se sentar aleatoriamente em N cadeiras alinhadas em fila. Qual é a probabilidade de duas pessoas em particular, A e B, se sentarem uma do lado da outra?

9. k pessoas irão se sentar aleatoriamente em N cadeiras alinhadas em fila. Qual é a probabilidade de k pessoas ocuparem lugares adjacentes?

10. Suponha que um comitê de 12 pessoas irá ser escolhido aleatoriamente dentre 100 pessoas. Qual é a probabilidade de duas pessoas em particular, A e B, serem escolhidas?

Atividade 2 - Probabilidade condicional e os teoremas relacionados.

Introdução

Nesta atividade introduz-se o conceito de probabilidade condicional e anuncia-se os principais teoremas relacionados com este conceito, em particular, o Teorema de Bayes. Esses conteúdos têm grande importância no cálculo de probabilidade.

Detalhes da atividade

Definição (Probabilidade condicional ou condicionada). Sejam dois acontecimentos A e B de uma experiência aleatória cujo espaço-amostra é Ω. A probabilidade de A se realizar sabendo que B se realizou – ou probabilidade de A condicionada por B, designada por P(A|B) – é definida por

$P(A|B)=P(A\cap B)/P(B)$,se $P(B)>0$.

Nota. A probabilidade condicionada pode interpretar-se como reavaliação da probabilidade de um acontecimento quando se tem a informação de que o outro acontecimento se realizou.

Se $P(B)=0$, $P(A|B)$ pode ser arbitrariamente definida. Mais tarde, ver-se-á que, por independência, é conveniente definir $P(A|B)=P(A)$.

Decorre da definição que $P(A\cap B)=P(A|B)\cdot P(B)$.

Exemplo. Supõe-se que uma fábrica possui 310 máquinas de soldar. Algumas dessas máquinas são elétricas (E), enquanto outras são manuais (M). Por outro lado, tem-se também que algumas são novas (N) e outras são usadas (U). A tabela abaixo mostra o número de máquinas de cada tipo:

	Elétricas	Manuais	Totais
Novas	10	60	70
Usadas	200	40	240
Totais	210	100	310

- Sabendo que uma determinada peça foi soldada usando uma máquina nova, qual éa probabilidade de ser soldada por uma máquina elétrica?
- Sabendo que uma determinada peça foi soldada com uma máquina elétrica, qual éa probabilidade de ter sido soldada por uma máquina nova?

Alínea (a).

$P(E|N)=P(E\cap N)/P(N)=10/70\cong 0{,}14285714$.

Alínea (b).

$P(N|E)=P(N\cap E)/P(E)=10/200=0{,}05$.

Teorema. A probabilidade condicional é uma medida de probabilidade.

Teorema (Teorema de multiplicação). Sejam $A_1,A_2,\dots,A_n$ acontecimentos de um espaço-amostra Ω. Então

$$
\begin{aligned}
P(A_1\cap A_2\cap\ldots\cap A_n) &= P(A_n|A_1\cap A_2\cap\ldots\cap A_{n-1}) \\
&\times P(A_{n-1}|A_1\cap A_2\cap\ldots\cap A_{n-2}) \\
&\times \ldots \\
&\times P(A_2|A_1)P(A_1)
\end{aligned}
$$

Em particular, se n=2, ter-se:

$$
\begin{aligned}
P(A_1\cap A_2) &= P(A_2|A_1)\times P(A_1) \\
&= P(A_1|A_2)\times P(A_2)
\end{aligned}
$$

Definição (Partição do espaço de resultado ou espaço-amostra). Diz-se que a classe de acontecimentos $\{A_1,A_2,\dots,A_m,\dots\}$ é uma partição do espaço de resultado Ω quando

$A_i\cap A_j=\emptyset\ (i\neq j)$ e $\bigcup_k A_k=\Omega$.

Nota. Pode-se considerar com número finito ou uma infinidade numerável de acontecimentos. A seguinte propriedade advém de propriedades gerais de medida de probabilidade:

$P(\bigcup_k A_k)=\sum_k P(A_k)=1$.

Exemplo. Considera-se um baralho de cinquenta e duas cartas. A sua divisão em quatro naipes (paus, ouros, copas e espadas), com treze cartas cada, constitui uma partição.

Teorema (Regra da probabilidade total). Seja $\{A_1,A_2,\dots,A_m,\dots\}$ uma partição de um espaço-amostra Ω. Suponha-se que $P(A_j)>0$, com $(j=1,2,\dots,m,\dots)$. Então, para qualquer acontecimento B tem-se:

$P(B)=\sum_j P(A_j)P(B|A_j)$.

Exemplo. Uma empresa produz circuitos em três fábricas, denotadas por I, II e III. A fábrica I produz 40% dos circuitos, enquanto a II e III produzem 30% cada uma. As probabilidades de que um circuito produzido por essas fábricas não funcione são 0,01, 0,04 e 0,03, respectivamente. Escolhido ao acaso um circuito da produção conjunta das três fábricas, qual é a probabilidade do circuito não funcionar?

Considera-se os seguintes acontecimentos:

A – "Circuito foi produzido pela fábrica I";

B – "Circuito foi produzido pela fábrica II";

C – "Circuito foi produzido pela fábrica III";

D – "O circuito não funciona".

Nota-se que os acontecimentos A, B e C formam uma partição do espaço-amostra. Então, pelo "Teorema de Probabilidade Total", tem-se:

$$
\begin{aligned}
P(D) &= P(D|A)P(A)+P(A|B)P(B)+P(D|C)P(C) \\
&= 0{,}01\times 0{,}4+0{,}04\times 0{,}3+0{,}03\times 0{,}3 \\
&= 0{,}025
\end{aligned}
$$

Teorema (Teorema de Bayes). Seja $\{A_1,A_2,\ldots,A_m,\ldots\}$ uma partição de um espaço-amostra Ω, e seja $P(A_j)>0$, $j=1,2,\ldots,m,\ldots$. Então, para qualquer acontecimento B tal que $P(B)>0$, tem-se:

$$P(A_j|B)=\frac{P(A_j)P(B|A_j)}{\sum_i P(A_i)P(B|A_i)},\quad j=1,2,\ldots,m,\ldots .$$

Exemplo. Uma empresa produz circuitos em três fábricas, denotadas por I, II e III. A fábrica I produz 40% dos circuitos, enquanto a II e III produzem 30% cada uma. As probabilidades de que um circuito produzido por essas fábricas não funcione são 0,01, 0,04 e 0,03, respectivamente. Um circuito é escolhido ao acaso da produção conjunta das três fábricas. Dado que o circuito escolhido não funciona, qual é a probabilidade do circuito ter sido produzido pela fábrica I?

Considera-se os seguintes acontecimentos:

A – "Circuito foi produzido pela fábrica I";

B – "Circuito foi produzido pela fábrica II";

C – "Circuito foi produzido pela fábrica III";

D – "O circuito não funciona".

Os acontecimentos A, B e C formam uma partição do espaço-amostra. Logo, aplicando o "Teorema de Bayes" tem-se:

$$P(A|D)=\frac{P(D|A)P(A)}{P(D|A)P(A)+P(D|B)P(B)+P(D|C)P(C)}=\frac{0{,}01\times 0{,}4}{0{,}025}=0{,}16 .$$

Conclusão

Os principais teoremas relacionados ao cálculo de probabilidades condicionais são "Teorema de Multiplicação", "Teorema de Probabilidade Total" e "Teorema de Bayes". É muito importante notar a diferença existente na aplicação dos dois últimos teoremas na resolução de problemas de cálculo de probabilidades, pois ambos são aplicados em situações de partições de espaço-amostra, mas em situações bem distintas.

Pode-se utilizar o conceito de probabilidade condicional para determinar a probabilidade da intersecção de dois acontecimentos.

Avaliação

1. Um empreiteiro apresentou orçamentos separados para execução da parte elétrica e da parte de encanamento de um edifício. Ele acha que a probabilidade de ganhar a concorrência da parte elétrica é de 50%. Caso ele ganhe a parte elétrica, a chance de ganhar a parte de encanamento é de 3/4, caso contrário, essa probabilidade é de 1/3.

   a. Qual é a probabilidade do empreiteiro ganhar os dois contratos?

   b. Qual é a probabilidade do empreiteiro ganhar apenas um contrato?

   c. Qual é a probabilidade do empreiteiro perder a parte elétrica e perder a parte deencanamento?

2. Uma caixa contém 24 lâmpadas, das quais 4 são defeituosas. Suponha que uma pessoa seleciona 10 lâmpadas aleatoriamente e, em seguida, uma outra pessoa seleciona as 14 lâmpadas restantes. Qual é a probabilidade das 4 lâmpadas defeituosas serem selecionadas pela mesma pessoa?

3. Um equipamento eletrónico e formado por 2 componentes, I e II. Suponha que:

   - A chance do componente I falhar é 0,20;

   - A chance de apenas o componente II falhar é 0,15.

   a. Calcule a probabilidade de apenas o componente I falhar.

   b. Calcule a probabilidade do componente I falhar dado que o componente II falhou.

Atividade 3 - Acontecimentos Independentes e suas consequências.

Introdução

Nesta atividade introduz-se o conceito de acontecimento independente, bem como as suas principais consequências no cálculo de probabilidades.

Detalhes da atividade

Definição (Independência entre dois eventos). Dois acontecimentos A e B, de um espaço-amostra, dizem-se independentes se, e somente se,

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

Nota. A definição anterior é válida para $P(A) \geq 0$ e $P(B) \geq 0$. Em particular, tem-se:

Se P(A) for igual a zero, então A será independente de qualquer outro acontecimento.

Qualquer acontecimento é independente dos acontecimentos $\emptyset$ e Ω.

Teorema. Se dois eventos forem independentes, então $P(A|B)=P(A)$ se $P(B)>0$ e $P(B|A)=P(B)$ se $P(A)>0$.

Ou seja, se dois eventos forem independentes, a ocorrência de um deles não afetará a probabilidade de ocorrência da outra.

Exemplo. Considera-se uma experiência aleatória que consiste em escolher um ponto aleatoriamente no círculo de raio unitário, centrado na origem do plano cartesiano. Sejam A um evento formado pelos pontos que estão a menos de meia unidade de distância da origem e B um evento formado pelos pontos que possuem primeira coordenada maior que a segunda. Mostrar que os eventos A e B são independentes.

Ora,

$P(A \cap B)=1/8$ e $P(A)P(B)=1/4 \quad 1/2$.

Ou seja, $P(A \cap B)=P(A)P(B)$, significando que os eventos A e B são independentes.

Teorema. Se A e B forem acontecimentos independentes, também o serão:

A e $\bar{B}$;

$\bar{A}$ e B;

$\bar{A}$ e $\bar{B}$.

Nota. Não se deve confundir acontecimentos independentes com acontecimentos incompatíveis. Basta notar que se A e B forem dois acontecimentos incompatíveis e tais que $P(A)>0$ e $P(B)>0$, então

$P(A \cap B)=P(\emptyset)=0$,

que será diferente de $P(A)P(B)\neq 0$, pela hipótese levantada anteriormente.

A independência é uma propriedade da medida de probabilidade e não uma propriedade dos acontecimentos. Consequentemente, não é possível caracterizar a independência meramente a partir dos acontecimentos.

Exemplo. Suponha-se que com uma moeda não regular tenha-se.

$P(F)=p$ e $P(C)=1-p$, onde $0\leq p\leq 1$.

Considera-se os seguintes acontecimentos:

$A=\{FFF,FFC,FCF,CFF\}$ e $B=\{FFF,CCC\}$,

associados a uma experiência aleatória que consiste em três lançamentos da referida moeda.

Então,

$P(A)=p^3+3p^2(1-p)$ e $P(B)=p^3+(1-p)^3$.

Ora,

$P(A\cap B)=P\{FFF\}=p^3$,

enquanto que

$P(A)P(B)=p^6+p^3(1-p)^3+3p^5(1-p)+3p^2(1-p)^4$.

Portanto, $P(A\cap B)= P(A)P(B)$ verifica-se apenas quando $p=0$ ou $p=1$, e no caso simétrico, $p=1/2$. Assim os acontecimentos A e B podem ou não ser independentes, consoante a natureza da moeda.

Definição (Independência 2 a 2). Seja $\{A_i : i\in I\}$ uma coleção de eventos aleatórios indexada por um conjunto de índices I. Os eventos desta coleção são ditos independentes 2 a 2 se

$P(A_i\cap A_j)=P(A_i)P(A_j), \forall i,j\in I$ tais que $i\neq j$.

Definição (Independência mútua). Seja $B=\{A_i : i\in I\}$ uma coleção de eventos aleatórios indexada por um conjunto de índices I. Os eventos desta coleção são mutuamente independentes se, para toda subfamília finita $\{A_{i_1},A_{i_2},\dots,A_{i_n}\}$ de eventos de B, ter-se:

$P(\{A_{i_1},A_{i_2},\dots,A_{i_n}\})=P(A_{i_1})P(A_{i_2})\dots P(A_{i_n})$.

Nota. Eventos mutuamente independentes são necessariamente independentes 2 a 2, mas o recíproco nem sempre é verdadeiro.

Exemplo. Seja uma experiência aleatória que consiste no lançamento de dois dados. Sejam os eventos:

A - "O primeiro dado mostra um número par";

B - "O segundo dado mostra um número ímpar";

C - "Ambos os dados mostram números ímpares ou ambos mostram números pares".

Os eventos acima são independentes 2 a 2?

Os eventos acima são mutuamente independentes?

Alínea (a). Os eventos A, B e C são independentes 2 a 2 se.

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

Ora:

$P(A)=(3\times 6)/36=0{,}5$ e $P(B)=(6\times 3)/36=0{,}5$ e $P(C)=9/36+9/36=0{,}5$.

$P(A \cap B)=9/36=0{,}25$ e $P(A \cap C)=9/36=0{,}25$ e $P(B \cap C)=9/36=0{,}25$.

Logo,

$P(A \cap B)=P(A)P(B)$ e $P(A \cap C)=P(A)P(C)$ e $P(B \cap C)=P(B)P(C)$,

ou seja, os eventos A, B e C são independentes.

Alínea (b). Os eventos A, B e C são mutuamente independentes se, para além de,

$P(A \cap B)=P(A)P(B)$ e $P(A \cap C)=P(A)P(C)$ e $P(B \cap C)=P(B)P(C)$,

ter-se ainda,

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

Mas $P(A \cap B \cap C)=P(\emptyset)=0$, enquanto que $P(A)P(B)P(C)=0{,}125$, ou seja,

$P(A \cap B \cap C)\neq P(A)P(B)P(C)$.

Logo, os eventos A, B e C não são mutuamente independentes.

Conclusão

Acontecimentos independentes e acontecimentos incompatíveis são coisas distintas.

Eventos mutuamente independentes são necessariamente independentes 2 a 2, mas o recíproco nem sempre é verdadeiro.

Avaliação

1. Sejam A e B acontecimentos independentes, definidos no mesmo espaço de resultados, sendo $P(A)=1/3$ e $P(B)=3/4$. Calcule:

   a. $P(A \cup B)$.

   b. $P(B|A \cup B)$.

2. Suponha que vai passar um fim-de-semana, a uma das cidades maravilhosas dos países africanos, viajando de avião. Considere os seguintes acontecimentos:

   A – "A minha mala perde-se na viagem de ida";

   B – "A minha mala perde-se na viagem de regresso".

   a. Sabendo que os dois acontecimentos são independentes, que $P(A \cup B)=0{,}0298$, que$P(A \cap B)=0{,}0002$ e que $P(A)<P(B)$, determine $P(A)$ e $P(B)$.

3. Maria é uma promissora engenheira química, está interessada em saber se determinada substância está contaminada por uma impureza que conduz à sua inutilização. Existe um teste laboratorial que tem probabilidade 0,8 de assinalar a impureza quando a substância está efetivamente contaminada, e probabilidade de 0,1 de assinalar a presença de impureza quando esta não está contaminada. Sabe-se ainda que, antes de efetuar qualquer teste, a probabilidade da substância estar contaminada por essa impureza é de 0,4.

   a. Qual a probabilidade de um teste laboratorial assinalar a presença de impureza?

   b. Tendo o teste assinalado a impureza, calcule a probabilidade da substância estarefetivamente contaminada.

4. Por precaução, dada a sua inexperiência, a Maria, do exercício anterior, efetuou de uma forma independente duas vezes o teste laboratorial à mesma substância.

   a. Qual a probabilidade de os dois testes acusarem a presença de impureza, se asubstância estiver efetivamente contaminada? E se não estiver contaminada?

   b. Calcule a probabilidade de os dois testes acusarem a presença de impureza?

   c. Sabendo que os dois testes acusaram a presença de impureza, qual a probabilidadeda substância estar efetivamente contaminada?

Resumo da Unidade

No cálculo de probabilidade é fundamental a determinação do espaço de resultados. Das três definições de probabilidade apresentadas (clássica, frequencista e geométrica), a mais utilizada no cálculo de probabilidade é a clássica.

A probabilidade condicional deu lugar a teoremas fundamentais no cálculo de probabilidade, como o teorema de multiplicação, a regra de probabilidade total e o teorema de Bayes. Entre acontecimentos independentes, a probabilidade condicional não tem efeito.

Avaliação da Unidade

Verifique a sua compreensão!

Instruções

O Teste de avaliação tem sete questões, algumas com alíneas.

Responda cada uma das questões de uma forma clara e justifique cada passo de resolução.

Critérios de Avaliação

Cada ponto ou alínea vale 10 pontos. Considera-se aprovado o estudante que tiver pelo menos 50% da cotação total.

Avaliação

1. Se 12 bolas são colocadas aleatoriamente em 20 caixas, qual é a probabilidade de nenhuma caixa receber mais do que uma bola?

2. Uma caixa contém cinco bolas das quais duas são pretas. As bolas pretas estão numeradas de 1 a 2, e as outras bolas, de 3 a 5. Extraem-se duas bolas ao acaso (uma a seguir a outra e sem reposição), e observa-se o número escrito em cada uma delas.

   a. Enumere os elementos do espaço-amostra associado à experiência.

   b. Defina o espaço de resultados dos seguintes acontecimentos:

   - A_1 – " A primeira bola observada é preta";
   - A_2 – " A segunda bola observada é preta";
   - A_3 – "As duas bolas observadas são pretas";
   - A_4 – "Pelo menos uma das bolas observadas é preta";
   - A_5 – "Exatamente uma das bolas observadas é preta";

3. Um sistema eletrónico é formado por dois subsistemas, A e B. De ensaios anteriores sabe-se que: a probabilidade de A falhar é 0,2, a probabilidade de B falhar sozinho é 0,15, e a probabilidade de A e B falharem simultaneamente é 0,15. Determine a probabilidade de:

   a. B falhar.

   b. Falhar apenas A.

   c. Falhar pelo menos um deles, A ou B.

   d. Não falhar nem A nem B.

   e. A e B falharem simultaneamente.

4. Admita que num CD com 14 músicas, gosta de 8 delas. Utilizando o botão de escolha aleatória do leitor de CD, cada uma das 14 músicas tocará uma só vez, por ordem aleatória. Qual a probabilidade de, nas duas primeiras músicas, gostar apenas de uma delas?

5. Três lâmpadas em paralelo a um circuito elétrico, o que significa que para haver corrente no circuito pelo menos uma delas tem que estar acesa. A probabilidade de uma qualquer delas se queimar por excesso de tensão na rede é igual a 0,5.

   a. Qual a probabilidade de haver corrente no circuito em casa de tensão elevada?

   b. No caso de uma subida de tensão na rede, qual a probabilidade de se queimaremduas lâmpadas?

6. No trajeto de um avião de guerra há duas estações de radar inimigas equipadas com baterias antiaéreas, que são acionadas só se o avião for detetado. O avião tem uma probabilidade de 0,25 de ser detetado pela 1ª estação, e, sendo detetado, tem três hipóteses em cinco de não ser abatido por essa estação. Se o avião não for detetado pela primeira estação, ele aproxima-se da segunda nas mesmas condições que a primeira. Em contrapartida, se a 1ª estação o deteta sem o abater ele é de certeza detetado e abatido pela 2ª estação.

   a. Qual a probabilidade de o avião não ser abatido na 1ª estação?

   b. Qual a probabilidade de o avião não ser abatido?

   c. Sabendo que o avião foi abatido, qual a probabilidade de ter sido a 1ª estação aabate-lo?

Leituras e outros Recursos

As leituras e outros recursos desta unidade encontram-se na lista de Leituras e Outros Recursos do curso.

Leituras.

- Larson, R., & Farber, B. (2010). Estatística Aplicada (4ª ed.). São Paulo: Pearson Prentice Hall.
- Murteira, B., Ribeiro, C. S., Silva, J. A., & Pimenta, C. (2002). Introdução à Estatística (2ª ed.). Lisboa: McGraw-Hill de Portugal.
- Neto, J. H. (2013). Introdução à Probabilidade.

Outros recursos.

- Máquina Científica: TI-83 ou semelhante.
- Software: Geogebra, R e Excel.

Unidade 3. Variável aleatória (unidimensional). Função de distribuição de uma variável

Introdução à Unidade

Nesta unidade introduz-se o conceito de variável aleatória (unidimensional) e faz-se a classificação da variável aleatória (discreta, contínua e mista). Introduz-se o conceito de função de distribuição, bem como as suas principais propriedades. Analisa-se algumas medidas de tendência central e, outras, de dispersão, relativamente à variável aleatória discreta e contínua. Estuda-se algumas distribuições discretas e algumas contínuas.

Objetivos da Unidade

Após a conclusão desta unidade, deverá ser capaz de:

- Identificar uma variável aleatória.
- Determinar a probabilidade de um acontecimento de uma variável aleatória.

Termos-chave

Variável aleatória: é uma variável X associada à uma observação de uma experiência de probabilidade.

Variável aleatória discreta: toma um número finito ou infinito numerável de valores.

Variável aleatória contínua: toma um número infinito não numerável de valores, os quais podem ser associados com medidas numa escala contínua.

Valor esperado de uma variável aleatória: é a média dos valores que a variável aleatória X assume.

Função de distribuição: é uma função que fornece probabilidade de uma variável aleatória, sob determinadas condições.

Função probabilidade: é uma função que permite calcular a probabilidade de um acontecimento de uma variável aleatória discreta.

Função densidade: é uma função que permite calcular a probabilidade de um acontecimento de uma variável aleatória contínua.

Função de uma variável aleatória: é uma função cuja variável independente é uma variável aleatória.

Atividades de Aprendizagem

Atividade 1 - Variável aleatória e Função de distribuição

Introdução

Nesta atividade introduz-se o conceito de variável aleatória, e faz-se a respetiva classificação. Introduz-se também os conceitos de função de distribuição, função probabilidade e função densidade e analisa-se as suas principais propriedades.

O conceito de função de uma variável aleatória é também introduzido, como uma nova variável aleatória.

Detalhes da atividade

Definição (Variável aleatória (v.a.)). Uma variável aleatória X É uma variável que representa um valor numérico associado a cada resultado de um experimento de probabilidade, ou seja, é uma característica numérica do resultado de um experimento de probabilidade.

Matematicamente, uma variável aleatória é uma função com domínio Ω (espaço de resultados) e conjunto de chegada R, ou seja,

$X:\omega\in\Omega\longrightarrow X(\omega)\in R$.

Nota. A variável aleatória X permite passar da estrutura probabilística original para uma nova estrutura probabilística de mais fácil utilização, pois as imagens da variável aleatória são elementos do conjunto R.

É habitual chamar a variável aleatória X à sua imagem, ou variável dependente, X(ω).

Quando não houver confusão, utiliza-se a notação mais simples X, em vez de X(ω), para designar a variável dependente ou imagem da função.

Cada variável aleatória representa-se por uma letra maiúscula, enquanto os particulares valores que pode assumir são representados pela respetiva letra minúscula.

Exemplo. Quando se lança dois dados, o espaço de resultados é dado por

$\Omega=\{(i,j):i,j=1,2,3,4,5,6\}$,

Isto é, o espaço de resultados tem 36 acontecimentos elementares. Quando se regista apenas a soma de pontos obtida, a variável aleatória é definida da seguinte forma:

$X:(i,j)\in\Omega\longrightarrow i+j\in R$.

Neste exemplo, a característica numérica é a soma de pontos.

Exemplo. Suponha-se um experimento que consiste em escolher um ponto ao acaso no círculo unitário. Nesse caso, o espaço de resultados é

$\Omega=\{(x,y)\in R^2:x^2+y^2\leq 1\}$.

Pode-se definir uma v. a. X que associa o resultado do experimento à distância entre o ponto e a origem. Assim,

$X:(x,y)\in\Omega\longrightarrow\sqrt{x^2+y^2}\in R$.

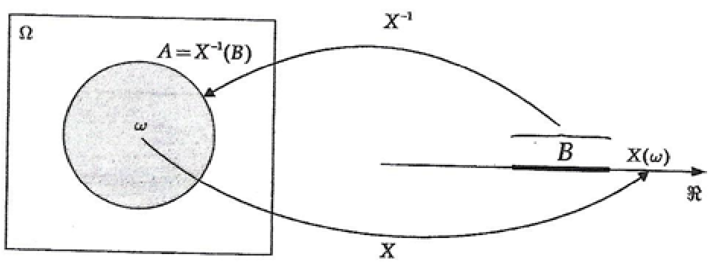
Com a variável aleatória, o conjunto R passa a ser novo espaço de resultados, os subconjuntos $B\subset R$ são os respetivos acontecimentos. Suponha-se que se quer determinar a probabilidade de B, utilizando a variável aleatória X, previamente definida, ou seja, que pretende calcular-se a probabilidade de X pertencer a B,

$P_X(B)=P_X(X(\omega)\in B)$.

Esta probabilidade é dada por:

$P_X(B)=P(A)$, onde $A=\{\omega\in\Omega:X(\omega)\in B\}$.

Ao conjunto B chama-se imagem de A, X(A)=B; ao conjunto A chama-se imagem inversa de B, $A=X^{-1}(B)$. Veja a figura abaixo, obtida em (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 102).



Nota. Só se pode calcular probabilidade probabilidades de acontecimentos $B\subset R$ cuja imagem inversa $X^{-1}(B)\in\Omega$ seja um acontecimento.

A medida de probabilidade da estrutura probabilística associada à variável aleatória X representa-se por P_X, mas pode-se usar-se apenas o símbolo P, embora a primeira ganha vantagem quando se tem mais do que uma variável aleatória em estudos.

Exemplo. Retome-se o "exemplo sobre lançamento de dois dados". No quadro abaixo, indicam-se as imagens, as imagens inversas e as probabilidades relevantes.

Imagens	Imagens inversas	Probabilidades
{2}	{(1,1)}	1/36
{3}	{(1,2),(2,1)}	1/18
{4}	{(1,3),(2,2),(3,1)}	1/12
{5}	{(1,4),(2,3),(3,2),(4,1)}	1/9
{6}	{(1,5),(2,4),(3,3),(4,2),(5,1)}	5/36
{7}	{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)}	1/6
{8}	{(2,6),(3,5),(4,4),(5,3),(6,2)}	5/36
{9}	{(3,6),(4,5),(5,4),(6,3)}	1/9
{10}	{(4,6),(5,5),(6,4)}	1/12
{11}	{(5,6),(6,5)}	1/18
{12}	{(6,6)}	1/36

Por exemplo,

$P_X(\{4\})=P_X(X=4)=P(\{(1,3),(2,2),(3,1)\})=1/12$.

Nota. No cálculo de probabilidade são utilizadas as seguintes notações:

$P(X\leq x)=P(\{\omega\in\Omega:X(\omega)\leq x\})$;

$P(X=x)=P(\{\omega\in\Omega:X(\omega)=x\})$;

$P(X<x)=P(\{\omega\in\Omega:X(\omega)<x\})$;

$P(X\geq x)=P(\{\omega\in\Omega:X(\omega)\geq x\})$;

$P(X>x)=P(\{\omega\in\Omega:X(\omega)>x\})$;

$P(X\in B)=P(\{\omega\in\Omega:X(\omega)\in B\})$.

Definição (Função de distribuição acumulada (ou função de distribuição)). A função de distribuição acumulada (f.d.a.), representada por F, de uma variável aleatória X é uma função real de variável real definida por

$F_X(x)=P(X\leq x)$.

Exemplo. Considera-se um experimento que consiste em escolher um número ao acaso em $\Omega=[-1,1]$. Sejam X uma variável aleatória que associa o número escolhido ao seu quadrado e F a função de distribuição de X.

a.Calcule F(0,25).

b.Determine F.

c.Construa o gráfico de F.

Alínea (a).

$$\begin{aligned} F(0{,}25) &= P(X\leq 0{,}25) \\ &= P(\{\omega\in\Omega:X(\omega)\leq 0{,}25\}) \\ &= P(\{\omega\in\Omega:\omega^2\leq 0{,}25\}) \end{aligned}$$

$= P(\{\omega\in\Omega: -0,5\le\omega\le 0,5\})$

$= (0,5-(-0,5))/(1-(-1)) = 1/2 = 0,5$

Alínea (b). Para $x<0$, tem-se

$F(x) = P(X\le x)$

$= P(\{\omega\in\Omega: X(\omega)\le x\})$

$= P(\{\omega\in\Omega: \omega^2\le x\})$

$= P(\emptyset)$

$= 0$

Para $0\le x\le 1$, tem-se:

$F(x) = P(X\le x)$

$= P(\{\omega\in\Omega: X(\omega)\le x\})$

$= P(\{\omega\in\Omega: \omega^2\le x\})$

$= P(\{\omega\in\Omega: -\sqrt{x}\le\omega\le\sqrt{x}\})$

$= (\sqrt{x}-(-\sqrt{x}))/(1-(-1)) = (2\sqrt{x})/2 = \sqrt{x}$

Para $x>1$, tem-se:

$F(x) = P(X\le x)$

$= P(\{\omega\in\Omega: X(\omega)\le x\})$

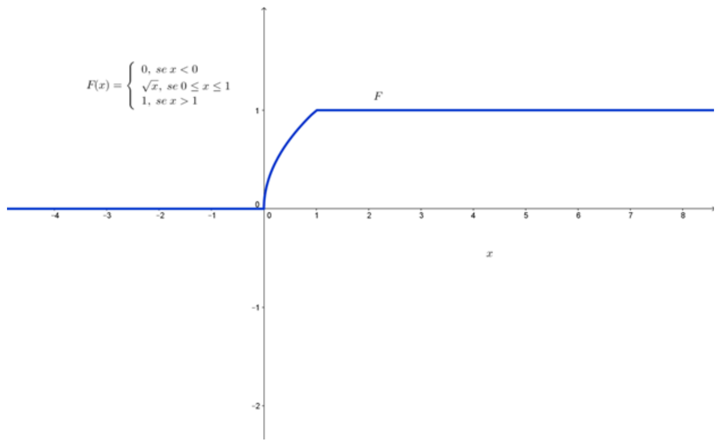
$= P(\{\omega\in\Omega: \omega^2\le x\})$

$= P(\Omega) = 1$

Então,

$$F(x)=\begin{cases}0, & \text{se } x<0\\ \sqrt{x}, & \text{se } 0\le x\le 1\\ 1, & \text{se } x>1\end{cases}.$$

Alínea (c). O gráfico de F é o seguinte:



Nota. Nota-se que $F_X(x)=P(X\le x)$ existe sempre por definição, e que $F_X(x)$ é a probabilidade do acontecimento $B=]-\infty,x]\subset\mathbb{R}$.

Sempre que não cause confusão, escreve-se, simplesmente F, em vez de $F_X(x)$, como aconteceu no "exemplo anterior, sobre função de distribuição".

Propriedades da função de distribuição.

Para qualquer função de distribuição, F, tem-se:

1. $0 \leq F(x) \leq 1$.

2. F é não decrescente, ou seja, dado $\Delta x > 0 \Rightarrow F(x) \leq F(x+\Delta x)$.

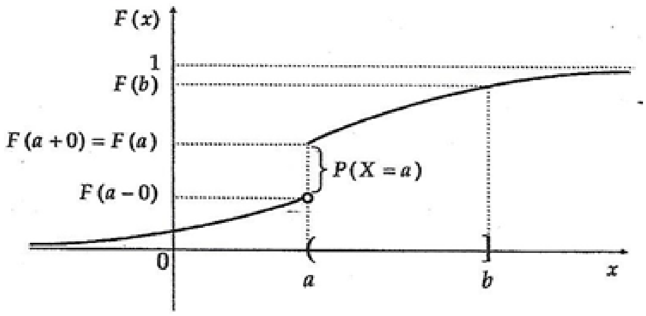
3. $F(-\infty)=\lim_{x \to -\infty} F(x)=0$ e $F(+\infty)=\lim_{x \to +\infty} F(x)=1$.

4. $P(a<X\leq b)=F(b)-F(a)$, para todos os números reias a e b tais que $b>a$.

5. F é contínua à direita, ou seja, $F(a^+)=\lim_{x \to a^+} F(x)=F(a)$, para todo numero real a.

6. $P(X=a)=F(a)-F(a^-)$, qualquer que seja o número real a.

A figura, que se segue, mostra de uma forma resumida as propriedades de uma função de distribuição, ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 106).



Nota. Se a função de distribuição é contínua num ponto, pela "propriedade 6", $P(X=x)=0$, ou seja, o acontecimento $\{x\}$ tem probabilidade zero; mas isto não significa que é um acontecimento impossível.

Teorema (Função de distribuição). A função real de variável real, F, é uma função de distribuição se, e só se, verifica as "propriedades 2, 3 e 5".

Exemplo. Prova-se que a função

$$F(x)=\begin{cases} 0, & x<0 \\ 1-e^{-x}, & x\geq 0 \end{cases}$$

define-se uma função de distribuição.

É necessário mostrar que esta função satisfaz as "propriedades 2, 3 e 5".

Propriedade 2. Para $x<0$, F é constante, logo não decrescente. Para $x\geq 0$, tem-se $F'(x)=e^{-x}>0$, ou seja F é estritamente crescente.

Propriedade 3. $\lim_{x\to -\infty} F(x)=\lim_{x\to -\infty} 0=0$ e $\lim_{x\to +\infty} F(x)=\lim_{x\to +\infty} (1-e^{-x})=1$.

Propriedade 5. Para todo $a\in R$, $F(a)=F(a^+)$, pois F é contínua em $R\backslash\{0\}$ e é contínua à direita de zero.

Outras propriedades da função de distribuição.

Para qualquer função de distribuição, F, tem-se:

a. $P(X<b)=F(b^-)$.

b. $P(X>a)=1-F(a)$.

c. $P(X\geq a)=1-F(a^-)$.

d. $P(a<X<b)=F(b^-)-F(a)$.

e. $P(a\leq X<b)=F(b^-)-F(a^-)$.

f. $P(a\leq X\leq b)=F(b)-F(a^-)$.

Teorema. Uma função F é uma função de distribuição se, e somente se, as condições abaixo forem satisfeitas:

1. $a\leq b$ implicar $F(a)\leq F(b)$, quaisquer que forem $a,b\in \mathbb{R}$, ou seja, F for função não decrescente;

2. $F(a^+)=F(a)$, ou seja, F é contínua à direita;

$$\lim_{x\to -\infty} F(x)=0;$$

$$\lim_{x\to +\infty} F(x)=1.$$

Classificação de variáveis aleatórias. Há variável aleatória discreta, contínua e mista.

Segundo (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 107) existem variáveis que nãos são nem discretas, nem contínuas e nem mistas. No entanto, nesta unidade ocupa-se apenas dos três tipos anunciados anteriormente.

Definição (Variável aleatória discreta). Uma variável aleatória X é discreta se toma um número finito ou enumerável de valores, ou seja, se existe um conjunto finito ou enumerável $D=\{x_1,x_2,\ldots\}\subset \mathbb{R}$ tal que $P(X\in\{x_1,x_2,\ldots\})=1$.

Se X for uma variável aleatória discreta, a função $f_X(x)=P(X=x)$ denomina-se por função de probabilidade de X.

Quando não houver ambiguidades costuma-se representar a função probabilidade apenas por $f(x)$.

Alguns autores representam a função de probabilidade por $p(x)$, ao invés de $f(x)$, pelo que neste módulo pode-se utilizar qualquer uma delas.

Teorema (Função probabilidade). Uma função $f(x)$ (ou $p(x)$) é função probabilidade de alguma variável aleatória discreta se, e somente se, existir um conjunto finito ou numerável $D=\{x_1,x_2,x_3,\ldots\}\subset \mathbb{R}$ tal que:

1. $f(x)>0$, para $x\in D$;

2. $f(x)=0$, caso contrário; e

3. $\sum_i f(x_i)=1$.

Definição (Suporte de uma distribuição de uma v. a. discreta). Seja X uma variável aleatória discreta. Chama-se suporte da distribuição de X ao conjunto

$\{x\in\mathbb{R}: f_X(x)>0\}$.

É prática corrente referir à função probabilidade apenas para os valores de x com probabilidade positiva, ou seja, os valores de x que pertencem a D. Este subdomínio da função f_X é o suporte da distribuição.

Por consequência da "definição da função probabilidade" pode-se calcular a probabilidade de qualquer acontecimento B exclusivamente a partir da função de probabilidade f; tem-se então:

$P(X\in B)=P(B)=\sum_{x\in B\cap D} f(x)$.

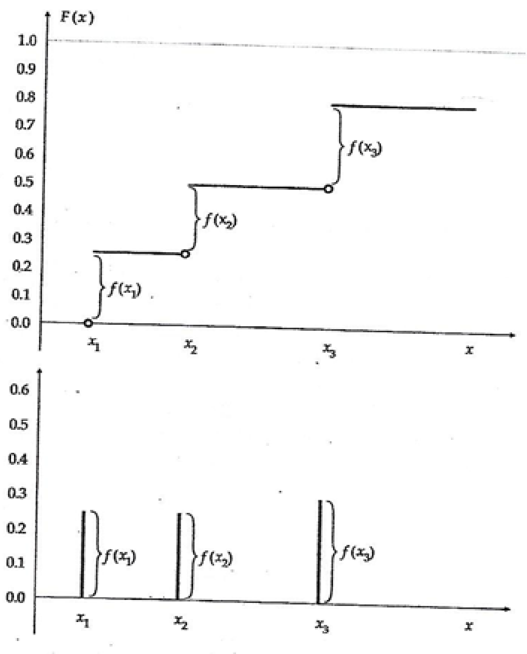
Fazendo $D=\{x_1,x_2,\ldots\}$, a função de distribuição pode exprimir-se facilmente em termos da respetiva função de probabilidade,

$F(x)=P(X\leq x)=\sum_{x_i\leq x} f(x_i), \forall x\in\mathbb{R}$.

É indiferente representar uma variável aleatória discreta com a respetiva função de distribuição ou com função probabilidade, pois cada $x\in D$ corresponde um salto da função distribuição.

O conjunto D é exatamente o conjunto de pontos de descontinuidade da função distribuição.

A figura que se segue representa graficamente a função distribuição e a respetiva função de probabilidade de uma variável aleatória discreta. A primeira é uma função em escada, não decrescente, descontínua à esquerda nos pontos de $D=\{x_1,x_2,\ldots\}$. A função probabilidade representa-se por um diagrama de barras. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 109).



Exemplo. Retoma-se o "exemplo sobre lançamento de dois dados". Nesse caso, a variável aleatória X associa cada resultado à soma dos números obtidos em cada dado.

a. Faça o gráfico da função de probabilidade de X.

b. Faça o gráfico da função de distribuição de X.

Para facilitar os esboços dos gráficos solicitados, apresenta-se a seguinte tabela (veja o exemplo acima):

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36
$F(x)$	1/36	1/12	1/6	5/18	5/12	7/12	13/18	5/6	11/12	35/36	1

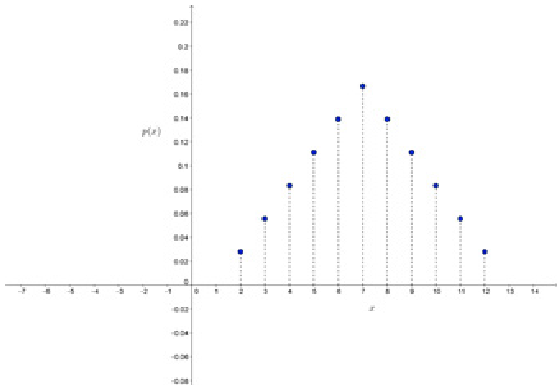


Figura 1 Gráfico de função de probabilidade p(x)

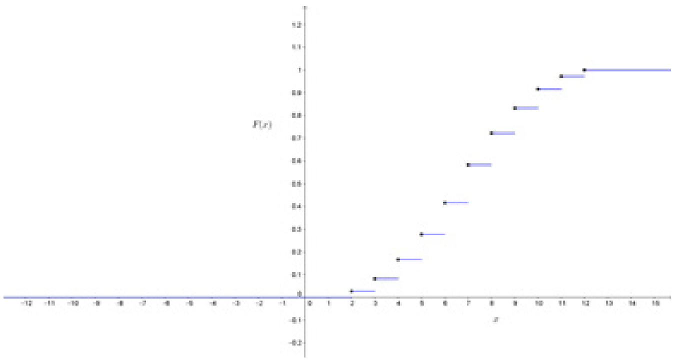


Figura 2 Gráfico de função de distribuição

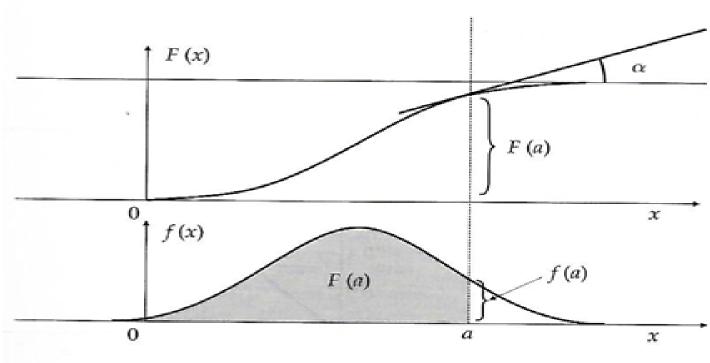
Definição (Variável aleatória contínua e densidade). Uma variável aleatória X é contínua se existe uma função integrável f tal que $f(x)\geq 0$ e

$F(x)=\int_{-\infty}^{x} f(t)\,dt,\ \forall x\in \mathbb{R}$.

Se X for contínua, a função f é chamada de função densidade de probabilidade de X, ou simplesmente densidade de X.

Na figura que se segue apresentam-se os gráficos da função distribuição e a respetiva função densidade.

Para todo o número real a, F(a) é a ordenada da função distribuição, e é também a área delimitada pela função densidade entre $-\infty$ e a; f(a) é a ordenada da função densidade, sendo também o declive da função de distribuição no ponto a. Tem-se, $tg(\alpha)=F'(a)=f(a)$. Ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 113).



Definição (Suporte de uma distribuição de uma v. a. contínua). Seja X uma variável aleatória contínua. Chama-se suporte da distribuição de X ao conjunto

$\{x\in\mathbb{R}: f_X(x)>0\}$,

onde f_X é a função de densidade.

Nota. Na prática corrente, a função densidade pode ser apenas considerada para os valores de x com densidade positiva, isto é para o subdomínio da função f igual ao suporte da distribuição.

Teorema. Se X é uma variável aleatória contínua, então:

a. $P(X=a)=0$, para todo $a\in\mathbb{R}$;

b. Para quaisquer números reais a, b tem-se

$$
\begin{aligned}
P(X\le a) &= P(X<a),\\
P(X\ge a) &= P(X>a) \text{ e}\\
P(a<X<b) &= P(a\le X\le b)\\
&= P(a<X\le b)\\
&= P(a\le X<b)
\end{aligned}
$$

Teorema. Se X é uma variável aleatória contínua com densidade f(x), então

a. Para todo acontecimento B,

b. Em particular, quando $B=a<X\le b$, tem-se

$P(a<X\le b)=\int_a^b f(x)\,dx=F(b)-F(a)\ ,\forall\ a,b\in\mathbb{R}$.

Teorema. Uma função f é densidade de alguma variável aleatória contínua se, e somente se,

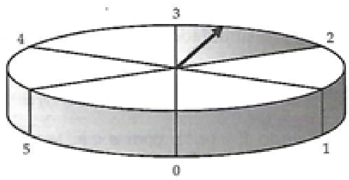
a. $f(x)\ge 0$, para todo $x\in\mathbb{R}$ e

b. $\int_{-\infty}^{+\infty} f(x)\,dx=1$.

Teorema. Para os pontos x onde F(x) é derivável, tem-se $f(x)=F'(x)$.

Nota. Relativamente ao "teorema anterior", deve-se dizer o seguinte: nos pontos onde não existe derivada da função distribuição, a função densidade não fica univocamente definida. Mas, normalmente, faz-se $f(x)=0$ nesses pontos.

Exemplo. Seja X uma variável aleatória contínua tomando valores no intervalo [0,6⊣[, e suponha-se que a probabilidade total se encontra uniformemente distribuída sobre esse intervalo. Admita-se que as observações experimentais da variável aleatória contínua X são obtidas por meio de uma tômbola, com a figura que se segue, "de construção suposta perfeita e de precisão infinita".



a.Determine P(X=a), para qualquer a∈[0,6⊣[.

b.Determine P([2,3⊣[)=P(X∈[2,3⊣[).

c.Determine a função de distribuição de probabilidade F de X.

d.Esboce o gráfico de F.

e.Determine a densidade f de X.

f.Esboce o gráfico de f.

g.Represente geometricamente P([2,3⊣[).

Alínea (a). Por ser X uma variável aleatória contínua P(X=a)=0.

Alínea (b). Como a probabilidade total se encontra uniformemente distribuída sobre o intervalo [0,6⊣[, então P([2,3⊣[)=P(X∈[2,3⊣[)=1/6.

Alínea (c). Para x<0, F(x)=0.

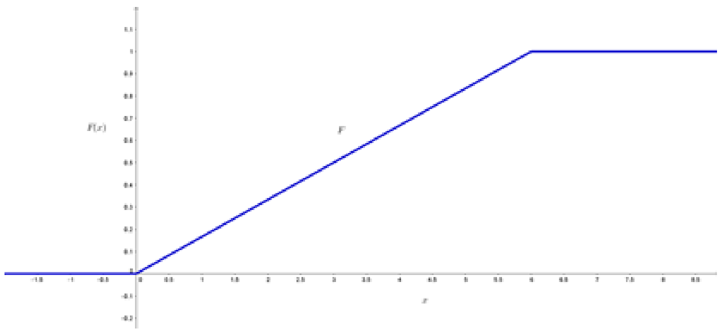
Para 0≤x<6, Como a probabilidade total se encontra uniformemente distribuída sobre o intervalo [0,6⊣[, então F(x)=P(X≤x)=x/6.

Para x≥6, F(x)=P(X≤x)=1.

Logo, a função distribuição de X é definida por:

$$F(x) = \begin{cases} 0 \ \ se \ \ x < 0 \\ x/6 \ \ se \ \ 0 \leq x < 6 \\ 1 \ \ se \ \ x \geq 6 \end{cases}$$

Alínea (d). O gráfico de F está de acordo com a figura que se segue:



Alínea (e). Como $F'(0^-)=0\neq F'(0^+)=1/6$, então $f(0)=0$.

Como, também, $F'(6^-)=1/6\neq F'(6^+)=0$, então $f(6)=0$.

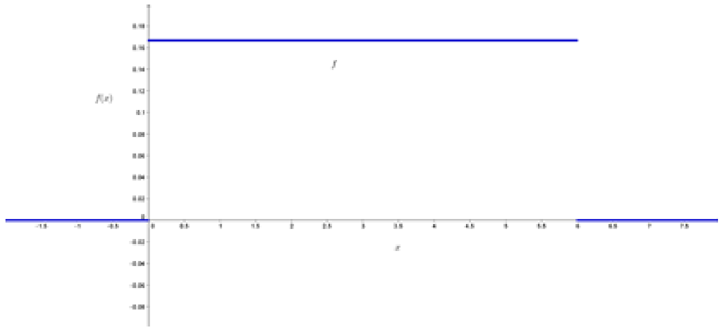
Para $x<0$, $F'(x)=0$.

Para $0<x<6$, $F'(x)=1/6$.

Logo,

$$f(x)=\begin{cases} 1/6 & 0<x<6 \\ 0, & \text{para outros valores de } x \end{cases}.$$

Alínea (f). O gráfico de f está de acordo com a figura que se segue:



Alínea (g). A representação geométrica de $P([2,3])$ é consoante a figura que se segue:

Recorda-se que $P([2,3])=\int_2^3 \frac{1}{6}\,dx=F(3)-F(2)=1/6$. Então, geometricamente $P([2,3])$ é a área do retângulo sombreado na figura.

Definição (Variável aleatória mista). Sejam F_{X_1} uma função de distribuição associada a uma variável aleatória discreta X_1, e F_{X_2}, outra função de distribuição caracterizadora de uma variável aleatória contínua, X_2. Diz-se que X é uma variável aleatória mista se, e só se, a respetiva função de distribuição puder ser escrita do seguinte modo:

$F_X(x)=\lambda F_{X_1}(x)+(1-\lambda)F_{X_2}$, onde $0<\lambda<1$.

Exemplo. Considere a variável aleatória X com função de distribuição dada por:

$$F_X(x)=\begin{cases} 0 \quad se \quad x<0 \\ (1/12)+(3/4)\left(1-e^{-x}\right) \quad se \quad 0\leq x<1 \\ (1/4)+(3/4)\left(1-e^{-x}\right) \quad se \quad x\geq 1 \end{cases}$$

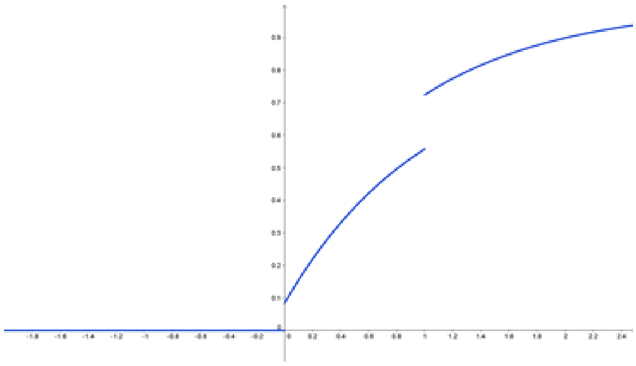
Bom, como esta função é descontínua em 0 – basta ver que $F(0^-)=0$, enquanto que $F(0^+)=F(0)=1/12$ – e em 1 – basta ver que $F(1^-)=1/12+3/4\times(e-1)/e$, enquanto que $F(1^+)=F(1)=1/4+3/4\times(e-1)/e$ – então:

$P(X=0)=F_X(0)-F_X(0^-)=1/12$.

$P(X=1)=F_X(1)-F_X(1^-)=2/12$.

$P(X=x)=0$, para outros valores reais de x .

O gráfico desta função distribuição está de acordo com a figura que se segue:



De acordo com a definição da variável aleatória mista uma função de distribuição de uma variável aleatória mista pode ser escrita na seguinte forma:

$F_X(x)=\lambda F_{X_1}(x)+(1-\lambda)F_{X_2}$,

onde F_{X_1} e F_{X_2} são funções de distribuições de uma variável aleatória discreta X_1 e de uma variável aleatória contínua X_2, respetivamente, e $0<\lambda<1$.

Então, esta função de distribuição – da variável aleatória mista X – pode ser escrita na forma referida acima, procedendo da seguinte forma:

Determinação de λ.

$$
\begin{aligned}
P(X=0) &= F_X(0)-F_X(0^-) \\
&= (\lambda F_{X_1}(0)+(1-\lambda)F_{X_2}(0))-(\lambda F_{X_1}(0^-)+(1-\lambda)F_{X_2}(0^-)) \\
&= \lambda(F_{X_1}(0)-F_{X_1}(0^-)) \\
&= \lambda f_{X_1}(0)=1/12\ ,
\end{aligned}
$$

uma vez que F_{X_2} é uma função de distribuição de uma variável contínua – X_2 – $F_{X_2}(0)=F_{X_2}(0^-)$, e f_{X_1} é a função probabilidade da variável discreta X_1.

Com as mesmas justificações apresentadas acima, tem-se:

$$
\begin{aligned}
P(X=1) &= F_X(1)-F_X(1^-) \\
&= (\lambda F_{X_1}(1)+(1-\lambda)F_{X_2}(1))-(\lambda F_{X_1}(1^-)+(1-\lambda)F_{X_2}(1^-)) \\
&= \lambda(F_{X_1}(1)-F_{X_1}(1^-)) \\
&= \lambda f_{X_1}(1)=2/12\ .
\end{aligned}
$$

Como só há dois pontos de descontinuidade – 0 e 1 – então a variável discreta X_1 assume probabilidade positiva apenas nestes pontos, logo:

$$f_{X_1}(0)+f_{X_1}(1)=1 \quad \Leftrightarrow \quad \lambda f_{X_1}(0)+\lambda f_{X_1}(0)=\lambda$$

$$\Leftrightarrow \quad 1/12+2/12=\lambda$$

$$\Leftrightarrow \quad \lambda=1/4$$

Determinação da função de distribuição da variável aleatória discreta X_1.

Ora:

$$f_{X_1}(0) = P(X=0)/\lambda$$

$$= 1/12\times 4=1/3$$

$$f_{X_1}(1) = P(X=1)/\lambda$$

$$= 2/12\times 4=2/3$$

Então:

Determinação da função de distribuição da variável aleatória contínua X_2.

Para $0\leq x<1$, tem-se por um lado

$$F_X(x)=\lambda F_{X_1}(x)+(1-\lambda)F_{X_2}(x) = \lambda f_{X_1}(0)+(1-\lambda)F_{X_2}(x)$$

$$= 1/12+3/4\, F_{X_2}(x),$$

e por outro lado

$$F_X(x)=1/12+3/4\,(1-e^{-x}) .$$

Então,

$$F_{X_2}(x)=1-e^{-x} .$$

Pode-se definir F_{X_2} da seguinte forma:

$$F_{X_2}=\begin{cases} 0 & se \quad x<0 \\ 1-e^{-x} & se \quad x\geq 0 \end{cases}$$

Função de uma variável aleatória.

Se X é uma variável aleatória, então qualquer função de X, diga-se $Y=g(X)$, é também uma variável aleatória que assume o valor $y=g(x)$, quando $X=x$.

Conhecida a função de distribuição de X, F_X, pretende-se determinar a função de distribuição de Y, F_Y.

Assim, pode-se considerar uma função $Y=g(X)$, com domínio e conjunto-chegada iguais a R.

Para cada conjunto $A \subset R$, considere-se um novo conjunto, denotado por g^(-1) (A), e definido por:

g^(-1) (A)={x∈R:g(x)∈A} .

A partir da distribuição de probabilidades da variável X, pode-se calcular as probabilidades para a variável Y, da seguinte forma:

P(Y∈A) = P(g(X)∈A)

= P(X∈g^(-1) (A)) .

Teorema. Seja X uma variável aleatória discreta, e seja Y=g(X) uma função de domínio e conjunto-chagada R. Então, Y também é uma variável discreta.

Exemplo. Considere-se a variável aleatória discreta com função probabilidade:

$$f(x) = \begin{cases} \frac{x^2}{15} & se \quad x = -1, 1, 0, -2, 2 \\ 0 \quad se & noutros\ casos \end{cases} .$$

Encontrar a distribuição de Y=X^2.

Considere-se a seguinte tabela:

x	-1	1	0	-2	2
y	1	1	0	4	4
$f_X(x)$	$\frac{1}{15}$	$\frac{1}{15}$	0	$\frac{4}{15}$	$\frac{4}{15}$

Então a variável aleatória Y pode assumir os valores 0, 1 e 4, com as seguintes probabilidades:

y	0	1	4
$f_Y(y)$	0	$\frac{2}{15}$	$\frac{8}{15}$

Teorema. Sejam X uma variável aleatória e Y=g(X) uma função da v. a. X. Sejam

D_X={x:f_X (x)>0} ,

isto é, o suporte da distribuição da v. a. X, e

D_Y={y:y=g(x),para algum x∈D_X }.

Então:

Se g for uma função crescente em D_X, F_Y (y)=F_X (g^(-1) (y)), para y∈D_Y.

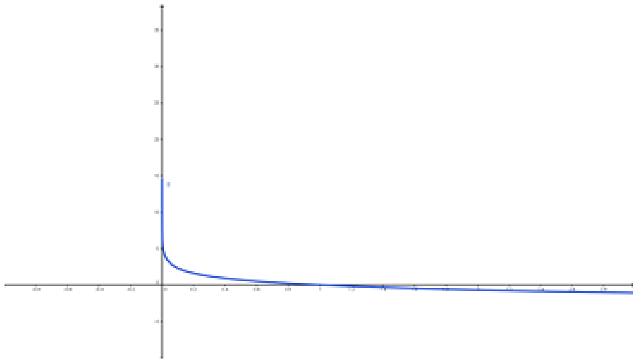
Se g for uma função decrescente em D_X e X for contínua, F_Y (y)=1-F_X (g^(-1) (y)), para y∈D_Y.

Exemplo. Seja X uma variável aleatória com densidade

$$f_X(x)=\begin{cases} 1 & se \quad 0<x<1 \\ 0 & se \quad noutros\ casos \end{cases} .$$

Encontrar a densidade da variável aleatória Y=g(X)=-ln(X).

Para começar, seja o gráfico da função g:



Da análise do gráfico da função g, pode-se constatar que ela é decrescente (poder-se-ia também calcular a derivada de g para concluir este facto).

Uma vez que D_X= ⊢]0,1⊣ [, então D_Y= ⊢]0,+∞⊣ [.

Nota-se também que g^(-1) (y)=e^(-y) e que

$$F_X(x)=\begin{cases} o & se \quad x\leq 0 \\ x & se \quad 0<x<1 \\ 1 & se \quad x\geq 1 \end{cases} .$$

Então, F_Y (y)=1-F_X (g^(-1) (y))=1-F_X (e^(-1))=1-e^(-y), para y∈ ⊢]0,+∞⊣ [.

Assim, tem-se:

$$F_Y(y)=\begin{cases} 1-e^{-y} & se \quad x\in\,]0,+\infty[\\ 0 \quad se & noutros\ casos \end{cases} .$$

Nota. A variável aleatória X pode ser contínua, a função Y=g(X) pode ser contínua, mas esses factos não garantem que Y seja uma variável aleatória contínua. No entanto, os dois teoremas, que se seguem, estipulam como determinar a função distribuição e a função densidade de uma variável aleatória contínua.

Teorema (Densidade de uma função). Sejam X uma v. a. E Y=g(X) uma função de X, onde g é uma função estritamente crescente ou estritamente decrescente. Suponha-se ainda que f_X (x) é contínua em D_X (o que garante a diferenciabilidade de F_X (x)) e que g^(-1) (y) é derivável em D_Y. A densidade da v. a. Y é dada por

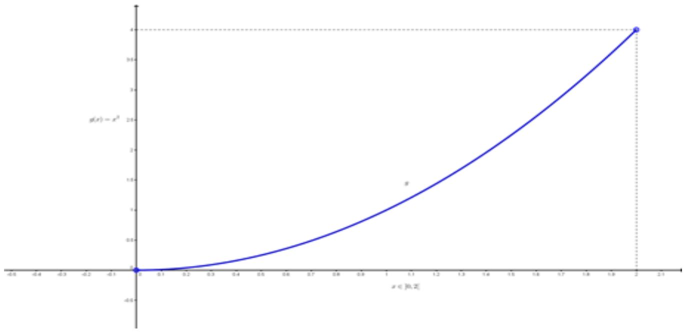
$$f_Y(y)=\begin{cases} f_X\left(g^{-1}(y)\right)\left|\dfrac{d}{dy}g^{-1}(y)\right| & se \quad y\in D_Y \\ 0 \quad se \quad noutros\ casos \end{cases} .$$

Exemplo. Seja X uma v. a. Com densidade

$$f_X(x) = \begin{cases} \dfrac{x}{2} & se \quad 0 < x < 2 \\ 0 & se \quad noutros\ casos \end{cases} .$$

Encontrara a densidade f_Y (y) de Y=X^2.

Neste caso, D_X=]0,2[e D_Y=]0,4[. A função g é estritamente crescente em D_X, como se pode ver pelo gráfico que se segue:



Ora, g^(-1) (y)=y^(1/2), para y∈]0,4[, logo tem-se:

$$\begin{aligned} f_Y(y) &= \begin{cases} f_X\left(g^{-1}(y)\right)\left|\dfrac{d}{dy}g^{-1}(y)\right| & se \quad y \in D_Y \\ 0 \quad se \quad noutros\ casos \end{cases} \\ &= \begin{cases} f_X\left(y^{\frac{1}{2}}\right)\left|\dfrac{1}{2}y^{-\frac{1}{2}}\right| & se \quad y \in]0,4[\\ 0 \quad se \quad noutros\ casos \end{cases} \\ &= \begin{cases} \dfrac{1}{4} & se \quad y \in]0,4[\\ 0 & se \quad noutros\ casos \end{cases} . \end{aligned}$$

Conclusão

A variável aleatória permite passar de uma estrutura probabilística original para uma outra de mais fácil tratamento.

São três as variáveis aleatórias mais comuns: discreta, contínua e mista.

A introdução de função de distribuição, função de probabilidade e função densidade permite determinar probabilidades de acontecimentos mais complexos.

O estudo de função de uma variável aleatória permite, de uma certa forma, a simplificação de um problema, mas também a análise mais profunda da mesma.

Avaliação

1. Suponha um experimento que consiste em arremessar um dado equilibrado. Seja X a variável aleatória associada ao número observado na face voltada para cima após o dado parar. Sejam f(x) e F(x) a função de probabilidade e a função de distribuição acumulada de X, respetivamente.

    a. Calcule f(2).

    b. Calcule f(3.5).

    c. Calcule F(2).

    d. Calcule F(3.7).

    e. Calcule F(500).

    f. Calcule F(-9).

    g. Construa o gráfico de f(x).

    h. Construa o gráfico de F(x).

2. Considere a variável aleatória associada com a contagem do número de faces obtido no lançamento de três moedas regulares. O número de faces em três lançamentos só pode ser 0, 1, 2 ou 3.

    a. Calcule F(x), para: $x<0$; $0\leq x<1$; $1\leq x<2$; $2\leq x<3$; $x\geq 3$.

    b. Determine F(x), para todo $x\in R$.

    c. Esboce o gráfico de F(x).

3. Joga-se um dado equilibrado. A variável aleatória que representa o número de lançamentos que se fazem até sair uma sena (6 pontos) pode assumir qualquer valor da sucessão dos números naturais x=1,2,3,….

    a. Calcule: P(X=1); P(X=2); P(X=3).

    b. Determine a função probabilidade f(x).

    c. Determine a função de distribuição de probabilidade F(x).

    d. Esboce o gráfico de f(x).

    e. Esboce o gráfico de F(x).

4. Considere a variável aleatória X que representa a duração da vida útil, em milhares de horas, de um componente de determinado tipo de radar (ou componentes eletrónicos, em geral. Em muitas situações, X tem função de distribuição dada por F(x)=1-e^(-θx), onde θ é um parâmetro que carateriza a qualidade do componente (quanto menor for θ, melhor é a qualidade). Responda as questões que se seguem para θ=0,03.

a. Esboce o gráfico de F(x).

b. Determine a função densidade f, e esboce o seu gráfico.

c. Determine a probabilidade de um acontecimento escolhido ao acaso: dure menosque 20 mil horas; dure entre 30 mil e 50 mil horas; dure mais que 60 mil horas.

5. Sabe-se que a probabilidade de determinada residência, segura contra roubo, ser vítima desta causalidade é de 0,01, e que, se esta ocorrência se der, o montante das perdas em relação ao capital seguro será dado pela variável aleatória X com função de densidade f_X (x)=12x^2 (1-x), para 0<x<1.

a. Represente o montante a pagar pela seguradora pela variável aleatória Y e determine F_Y (y).

b. Justifique que a variável aleatória Y é mista.

c. Determine λ∈]0,1[, F_(Y_1), função de distribuição de uma variável aleatóriadiscreta Y_1, e F_(Y_2), função de distribuição de uma variável aleatória contínuaY_2, de modo que

F_Y (y)=λF_(Y_1) (y)+(1-λ) F_(Y_2) (y) .

6. Determine o valor de real de k de modo que a seguinte função seja uma função de densidade de uma variável aleatória X:

$$f(x) = \begin{cases} kx & se \quad 0 < x < 1 \\ 2 - x & se \quad 1 < x < 2 \\ 0 & se \quad x \in]-\infty, 0] \cup \{1\} \cup [2, +\infty[\end{cases} .$$

7. O número de automóveis encomendados mensalmente num stand, é uma variável aleatória X com a seguinte função de probabilidade:

x	0	1	2	3	4
$f(x)$	0,3	0,3	0,2	0,1	0,1

a. Calcule a função de distribuição de X.

b. Determine o número mínimo de automóveis que o stand deve ter num mês para quea probabilidade de satisfazer todas as encomendas não seja inferior 0,75.

c. Num mês em que haja apenas dois automóveis em stock no stand, calcule aprobabilidade de serem todos vendidos, e especifique a distribuição da variávelaleatória que representa as vendas nesse mês.

8. Seja X uma variável aleatória discreta com função de distribuição dada por

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

   a. Determine a função probabilidade de X.

   b. Calcule P(X>1).

   c. Calcule P(X<0,5 | X≥0).

9. Supondo que a variável aleatória contínua X tem função de distribuição F_X, determine, para cada alínea, a função distribuição da variável aleatória correspondente:

   a. Y=|X|.

   b. Y=aX+b, onde a,b∈R e a≠0.

   c. Y=X^2.

10. Seja X uma variável aleatória com densidade

$$f_X(x) = \begin{cases} \frac{2}{9}(x+1) & se \quad -1 \leq x \leq 2 \\ 0 & se \quad outros\ casos \end{cases}.$$

   a. Encontre a função de distribuição da v. a. Y=X^2.

   b. Encontre a densidade f_Y da v. a. Y=X^2.

11. A quantidade de vinho (em dezenas de litros) que um produtor-engarrafador vende por dia é uma variável aleatória X com função de distribuição definida por:

$$F_X(x) = \begin{cases} 0 & se \quad x < 0 \\ \frac{x^2}{50} & se \quad 0 \leq x < 5 \\ \frac{20x - x^2}{50} - 1 & se \quad 5 \leq x < 10 \\ 1 & se \quad x \geq 10 \end{cases}.$$

   a. Determine a função densidade.

   b. Do conjunto de dias em que vende mais que 50 litros, qual a probabilidade devender menos que 90 litros?

   c. Se o ganho líquido diário for Y=2X-6, calcule a função de distribuição do ganholíquido e a proporção de dias em que há prejuízo.

Atividade 2 – Distribuições discretas

Introdução

Nesta atividade introduz-se os conceitos de valor esperado e momentos, etc., que são muito importantes na caracterização de uma variável aleatória.

Estuda-se também alguns modelos probabilísticos de muita importância nas aplicações da teoria de probabilidade e da estatística, relativamente à variável aleatória discreta.

Detalhes da atividade

Definição (Valor esperado de uma variável aleatória discreta). Seja X uma variável aleatória discreta, com função probabilidade f_X e com suporte $D_X=\{x_1,x_2,\ldots\}$. A expressão

$$E(X)=\sum_{x\in D_X} x f_X(x) ,$$

quando

$$\sum_{x\in D_X} |x| f_X(x) <+\infty ,$$

define o valor esperado, média ou esperança matemática de X.

Nota.

Se o conjunto D_X é finito a condição

$$\sum_{x\in D_X} |x| f_X(x) <+\infty$$

verifica sempre, logo existe o valor esperado da variável aleatória X.

Quando D_X é infinito numerável, o valor esperado de X é dado pela série

$$\sum_{x\in D_X} x f_X(x) ,$$

e esta, deve ser absolutamente convergente. Então será sempre necessário verificar a convergência desta série antes de avançar-se pelo cálculo do valor esperado.

No entanto, para o presente trabalho, parte-se de princípio que a convergência da série já está verificada, sempre que se está a calcular o valor esperado de uma variável aleatória discreta, com suporte infinito numerável.

Geralmente o valor esperado da v. a. X é representado por μ_X, ou simplesmente por μ, quando não houver confusão.

Do ponto de vista frequencista, o valor esperado de uma variável aleatória se aproxima da média aritmética dos valores obtidos, quando se repete a experiência aleatória um "grande número de vezes".

Exemplo. Seja X a variável aleatória que representa o número de pontos obtido com o lançamento de um dado regular.

Neste caso, o suporte é $D_X= \{1,2,3,4,5,6\}$, logo

$E(X)=\mu_X \quad = \quad \sum_{x\in D_X} x f_X$

$= \quad 1\times 1/6+2\times 1/6+3\times 1/6+4\times 1/6+5\times 1/6+6\times 1/6$

$= \quad 21/6$

$= \quad 3{,}5$

Definição (Valor esperado de uma função de variável discreta). Seja X uma variável aleatória discreta, com função probabilidade f_X e com suporte $D_X=\{x_1,x_2,\ldots\}$. A expressão

$E(\varphi(X))=\sum_{x\in D_X} \varphi(x)\, f_X(x)$

é valor esperado da variável aleatória $\varphi(X)$, onde φ é função da variável aleatória X.

Exemplo. Seja X uma variável aleatória discreta, com função probabilidade dada por

$$f_X(x)=\begin{cases} \frac{1}{3} & se \quad x=-1,0,1 \\ 0 & se \quad noutros\ casos \end{cases}.$$

Seja φ uma função da variável aleatória X definida por $\varphi(X)=X^2$.

Neste caso, pela definição,

$E(\varphi(X))=(-1)^2\times 1/3+0\times 1/3+1^2\times 1/3=2/3$.

Nota.

Pode-se calcular o valor esperado de $\varphi(X)$, a partir da "definição do valor esperado de uma variável aleatória discreta". Basta calcular o suporte da variável aleatória $\varphi(X)$ e a respetiva função probabilidade.

No "exemplo anterior", o suporte da variável aleatória $\varphi(X)$ é $D_{\varphi(X)}=\{0,1\}$ e a função probabilidade é

$$f_{\varphi(X)}(\varphi(x))=\begin{cases} \frac{1}{3} & se \quad \varphi(x)=0 \\ \frac{2}{3} & se \quad \varphi(x)=1 \\ 0 & se \quad noutros\ casos \end{cases}.$$

Logo,

$E(\varphi(X))=0\times 1/3+1\times 2/3=2/3$.

A existência de $E(X)$ não implica a existência de $E(\varphi(X))$, e inversamente.

Teorema (Propriedades do valor esperado). Seja X uma variável aleatória discreta, sejam ψ e φ funções da v. a. discreta X e seja c uma constante real. Então:

$E(c)=c$.

Se existe $E(\varphi(X))$, então $E(c\varphi(X))=cE(\varphi(X))$.

Se existem $E(\varphi(X))$ e $E(\psi(X))$, então $E(\varphi(X)+\psi(X))=E(\varphi(X))+E(\psi(X))$.

Definição (Momento de ordem k em relação à origem). Seja k um número inteiro positivo, e seja X uma variável aleatória discreta, com suporte $D_X=\{x_1,x_2,x_3,\ldots\}$. O valor esperado,

$$\mu_k'=E(X^k)=\sum_{x\in D_X} x^k f_X(x) \ ,$$

se existir, é o momento de ordem k em relação à origem, ou momento ordinário de ordem k da variável aleatória X.

Teorema. Seja k um número inteiro positivo, e seja X uma variável aleatória. Se existe o momento ordinário de ordem k da v. a. X, então existe o momento ordinário de ordem s, onde s é um inteiro positivo menor que k, da mesma v. a. X.

Nota.

Pode-se concluir, a partir da "definição de momento ordinário de ordem k ", que o momento de ordinário de ordem 1 de uma variável aleatória X, com suporte $D_X=\{x_1,x_2,x_3,\ldots\}$, é a média ou valor esperado dessa v. a. X, ou seja:

$$\mu_1=\mu_X=E(X)=\sum_{x\in D_X} x f_X(x) \ ,$$

que corresponde ao centro de gravidade da distribuição. Ou seja, assinalando a cada x, um ponto material de abcissa x e massa $f_X(x)=P(X=x)$, o conjunto de tais pontos forma um sistema material que tem por centro de gravidade o ponto de abcissa μ. Deste modo, pode encarar-se o valor esperado μ como um parâmetro de localização da distribuição de X.

Definição (Momento de ordem k em relação à média aritmética). Seja X uma variável aleatória discreta, com suporte $D_X=\{x_1,x_2,x_3,\ldots\}$, e seja k um inteiro positivo. O valor esperado

$$\mu_k=E((X-\mu)^k)=\sum_{x\in D_X} (x-\mu)^k f_X(x) \ ,$$

se existir, é o momento de ordem k em relação à média (ou valor esperado) μ, ou momento central de ordem k da variável aleatória discreta X.

Nota.

O primeiro momento em relação à média não tem qualquer interesse, pois é sempre nulo:

$$\mu_1=E(X-\mu)=E(X)-E(\mu)=\mu-\mu=0 \ .$$

Definição (Variância). A variância de uma variável aleatória discreta X, com suporte D_X, é o segundo momento em relação à média, e é representado pelos símbolos Var(X) ou σ_X^2, ou simplesmente σ se não houver ambiguidade, ou seja,

$$Var(X)=E((X-\mu))=\sum_{x\in D_X} (x-\mu)^2 f_X(x) \ .$$

Nota.

Para toda variável aleatória discreta X, $Var(X) \geq 0$, e dá-se a igualdade quando X tem toda a massa de probabilidade concentrada num ponto qualquer c, ou seja $P(X=c)=1$.

Como a variância de X é o valor esperado dos quadrados dos desvios entre X e a sua média μ, pode-se considerar a variância um parâmetro de dispersão da distribuição de X em torno da média.

Teorema (Propriedades da variância). Sejam c e d constantes reais, e seja X uma variável aleatória discreta. Então:

$Var(c)=0$.

$Var(cX)=c^2 Var(X)$.

$Var(cX+d)=c^2 Var(X)$.

Nota. Como $(X-\mu)^2=X^2-2X\mu+\mu^2$, verifica-se então que:

$$
\begin{aligned}
Var(X) &= E(X^2-2X\mu+\mu^2) \\
&= E(X^2)-2\mu E(X)+E(\mu^2) \\
&= E(X^2)-2\mu^2+\mu^2 \\
&= E(X^2)-\mu^2
\end{aligned}
$$

Definição (Desvio padrão). Seja X uma variável aleatória discreta, com suporte D_X. Ao valor positivo da raiz quadrada da variância,

$$\sigma_X=\sqrt{Var(X)},$$

chama-se desvio padrão da variável aleatória X.

Nota. Quando não existe ambiguidade o desvio padrão da v. a. discreta X pode representar-se simplesmente por σ.

É mais conveniente utilizar o desvio padrão para medir a dispersão dos elementos observados, do que a variância, pois o primeiro utiliza a unidade de medida original dos dados.

Definição (Coeficiente de variação). Seja X uma variável aleatória discreta, com suporte D_X contido no conjunto dos números reais positivos. O coeficiente de variação de X, se existir, é igual ao quociente entre desvio padrão e a média, ou seja,

$$CV=\sigma/\mu.$$

Nota.

O coeficiente de variação é uma grandeza que não depende de qualquer unidade de medida (como acontecem com o desvio padrão e a variância).

Essa grandeza (a de desvio padrão) é muito utilizada para comparar.

Definição (Coeficiente de assimetria). Seja X uma variável aleatória discreta, com suporte D_X. O quociente

$\gamma_1=\frac{E(X-\mu)^3}{\sigma^3}=\frac{\mu_3}{\sigma^3}$,

se existir, é designado por coeficiente de assimetria da variável aleatória discreta X.

Nota.

Uma distribuição simétrica conduz a um coeficiente de assimetria nulo.

Se $\gamma_1>0$, a distribuição será mais "abrupta do lado esquerdo" e os desvios positivos dominarão.

Inversamente, se $\gamma_1<0$, a distribuição será mais "abrupta do lado direito".

Exemplo. Sejam as variáveis aleatórias discretas X e Y, com funções probabilidades definidas nos suportes indicados por:

$f_X(x)=\frac{2-|x-4|}{4}$, em $D_X=\{3,4,5\}$.

$f_Y(y)=\frac{4-|y-4|}{16}$, em $D_Y=\{1,2,3,4,5,6,7\}$.

Calcula-se, em relação às estas duas variáveis aleatórias, as médias, desvios padrão, coeficientes de variação e coeficientes de assimetria.

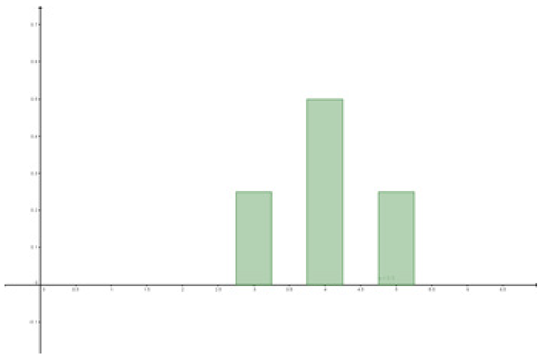
Alínea (a).

$\mu_X = \sum_{x\in D_X} x f_X(x) = 3\times 1/4+4\times 1/2+5\times 1/4=4$.

$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{\sum_{x\in D_X} (x-\mu_X)^2 f_X(x)} = \sqrt{1/4+1/4} = \sqrt{0,5} \cong 0,707$.

$CV = \sigma_X/\mu_X = \sqrt{0,5}/4 \cong 0,177$.

$\gamma_1 = \frac{E(X-\mu_X)^3}{\sigma^3} = \frac{-1/4+1/4}{(0,5)^3} = 0$



Alínea (b).

$\mu_Y = \sum_{y\in D_Y} y f_Y(y) = 1/16+2\times 1/8+3\times 3/16+4\times 1/4+5\times 3/16+6\times 1/8+7\times 1/16=4$.

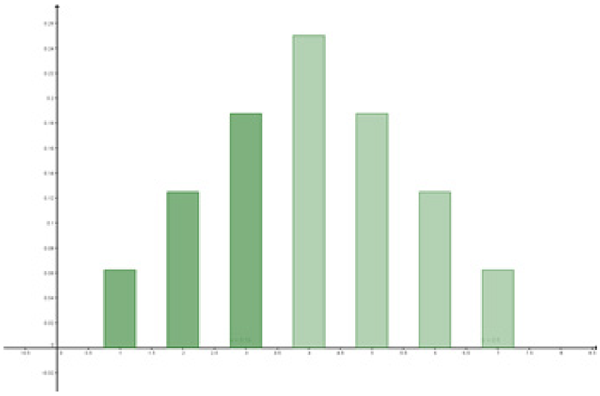
$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{\sum_{y\in D_Y} (y-\mu_Y)^2 f_Y(y)}$

$= \sqrt{9\times 1/16+4\times 1/8+3/16+3/16+4\times 1/8+9\times 1/16} = \sqrt{2,5} \cong 1,581$

$CV = \sigma_Y/\mu_Y = \sqrt{2,5}/4 \cong 0,395$.

$\gamma_1 = \frac{E(Y-\mu_Y)^3}{\sigma^3}$

$= \frac{-27\times 1/16-8\times 1/8-3/16+3/16+8\times 1/8+27\times 1/16}{(0,5)^3} = 0$



As duas distribuições do exercício anterior são simétricas, pois o coeficiente de assimetria, γ_1, é igual a zero.

Exemplo. Sejam as variáveis aleatórias discretas X e Y, com funções probabilidades definidas nos suportes indicados por:

a. $$f_X(x) = \begin{cases} \frac{2^{6-x}}{64} & se \quad x = 1,2,3,4,5,6 \\ \frac{1}{64} & se \quad x = 7 \end{cases} .$$

b. $$f_Y(y) = \begin{cases} \frac{1}{64} & se \quad y = 1 \\ \frac{2^{y-2}}{64} & se \quad y = 2,3,4,5,6,7 \end{cases} .$$

Calcula-se, em relação às estas duas variáveis aleatórias, as médias, desvios padrão, coeficientes de variação e coeficientes de assimetria.

Alínea (a).

μ_X = ∑_(x∈D_X)▒〖xf_X (x) 〗=1×1/2+2×1/4+3×1/8+4×1/16+5×1/32+6×1/64+7×1/64=127/64

≅ 1,984

σ_X^2 = ∑_(x∈D_X)▒〖(x-μ_X)^2 f_X (x) 〗

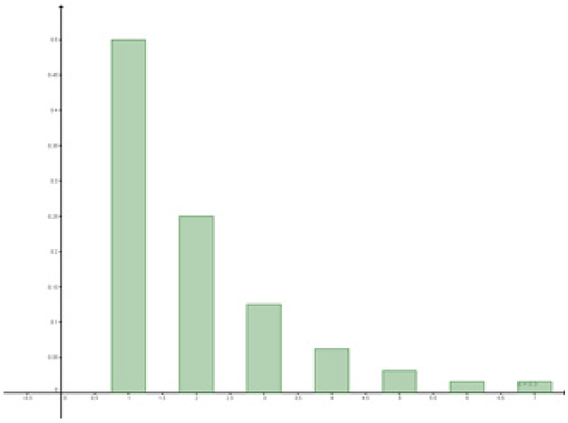
= (-63/64)^2×1/2+(1/64)^2×1/4+(65/64)^2×1/8+(129/64)^2×1/16+(193/64)^2×1/32+(257/64)^2×1/64

+(321/64)^2×1/64=1 662/831≅1,797

σ_X = √(σ_X^2)=√(1 662/831)=1 177/520≅1,340

CV = σ_X/μ_X =179/265≅0,675

γ_1 = E((X-μ_X)^3)/σ^3 =1 495/746≅1,664



Esta distribuição é assimétrica positiva, pois o coeficiente de assimetria, γ_1, é maior que zero.

Alínea (b).

$\mu_Y = \sum_{y\in D_Y} y f_Y(y) = 1\times 1/64+2\times 1/64+3\times 1/32+4\times 1/16+5\times 1/8+6\times 1/4+7\times 1/2 = 6\ 1/64$

$\cong 6{,}016$

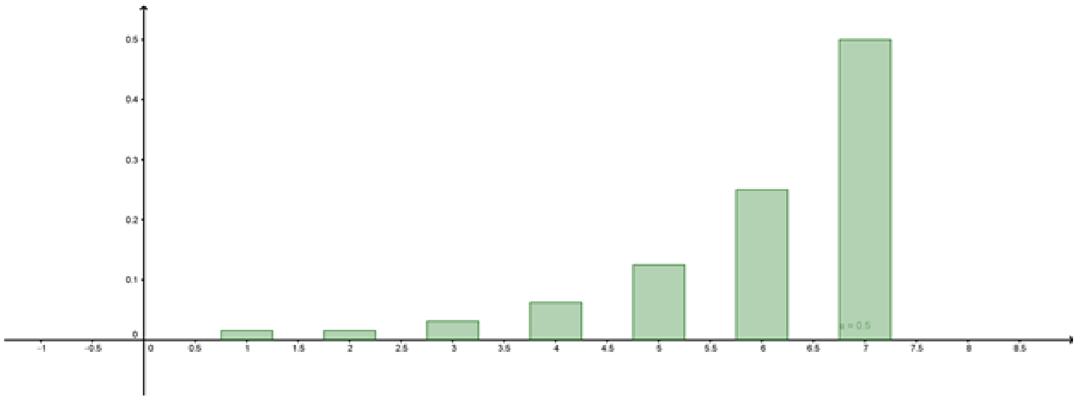
$\sigma_Y^2 = \sum_{y\in D_Y} (y-\mu_Y)^2 f_Y(y) = 1\ 662/831 \cong 1{,}797$

$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{1\ 662/831} = 1\ 177/520 \cong 1{,}340$

$CV = \sigma_Y/\mu_Y = 125/561 \cong 0{,}223$

$\gamma_1 = E((Y-\mu_Y)^3)/\sigma^3 = -1\ 859/997 \cong -1{,}862$

Esta distribuição é assimétrica negativa.



Distribuições discretas. Apresenta-se aqui alguns modelos probabilísticos de variáveis aleatórias discretas com que se depara frequentemente nas aplicações empíricas.

Definição (Distribuição uniforme). Seja X uma variável aleatória discreta, com suporte $D_X=\{x_1,x_2,\ldots,x_n\}$. Diz-se que X tem uma distribuição uniforme nos n pontos x_j, $j=1,2,\ldots,n$, se, e só se, a função probabilidade é dada por:

$$f_X(x) = \begin{cases} \frac{1}{n} & se \quad x \in D_X \\ 0 & se \quad x \notin D_X \end{cases}.$$

Algumas características de distribuição uniforme.

$$E(X)=\mu = \frac{1}{n}\sum_{j=1}^{n} x_j \;;$$

$$Var(X) = \frac{1}{n}\sum_{j=1}^{n}(x_j-\mu)^2 = \frac{1}{n}\sum_{j=1}^{n} x_j^2 - \mu^2 \;;$$

É simétrica.

Exemplo. O modelo probabilístico que descreve o lançamento de um dado perfeito é uma distribuição uniforme. Neste caso, $D=\{1,2,3,4,5,6\}$ e $P(X=x)=1/6$, onde $x=1,2,3,4,5,6$.

Neste caso, tem-se:

$$E(X)=\mu = \frac{1}{6}\times\sum_{j=1}^{6} x_j = \frac{1}{6}\times 21 = \frac{7}{2}\,.$$

$$Var(X) = \frac{1}{6}\times\sum_{j=1}^{6} x_j^2 - \mu^2 = \frac{1}{6}\times 91 - \frac{49}{4} = 2\,\frac{11}{12}$$

$$\sigma = \sqrt{Var(X)} = \sqrt{2\,\frac{11}{12}} \cong 1\,\frac{407}{575} \cong 1{,}709$$

Nota. Suponha-se que D_X seja igual a $D=\{1,2,3,\ldots,n\}$ e que $P(X=x)=1/n$, ou seja, que a variável aleatória discreta X segue uma distribuição uniforme.

Neste caso, ter-se-á:

$$E(X)=\mu = \frac{1}{n}\times\sum_{j=1}^{n} x_j = \frac{1}{n}\times\frac{n}{2}\times(n+1) = \frac{n+1}{2}\,.$$

$$\begin{aligned} Var(X) &= \frac{1}{n}\times\sum_{j=1}^{n} x_j^2 - \mu^2 = \frac{1}{n}\times\frac{n\times(n+1)\times(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\ &= (n+1)\times\left(\frac{2n+1}{6} - \frac{n+1}{4}\right) = (n+1)\times\frac{4n+2-3n-3}{12} \\ &= \frac{(n+1)\times(n-1)}{12} = \frac{n^2-1}{12}\,. \end{aligned}$$

$$\sigma = \sqrt{Var(X)} = \sqrt{\frac{n^2-1}{12}} = \frac{1}{2}\times\sqrt{\frac{n^2-1}{3}}$$

Definição (Distribuição de Bernoulli). Seja X uma variável aleatória discreta, com função probabilidade definida em $D_X=\{0,1\}$ por:

$$f_X(x\,|\,\theta) = \theta^x\cdot(1-\theta)^{1-x}, \text{para } 0<\theta<1\,.$$

Neste caso, diz-se que X tem uma distribuição de Bernoulli, e representa-se por $X \sim B(1;\theta)$.

Nota. A distribuição de Bernoulli aparece associada com experiência aleatória, designada por prova de Bernoulli, em que se observa a realização ou não de um acontecimento A, com probabilidade $P(A)=\theta$. A realização de A diz-se um "sucesso", e a sua não realização, que tem probabilidade $P(\bar{A})=1-\theta$, um "insucesso". Assim, a variável aleatória discreta X assume o valor 1 quando há sucesso, e o valor 0 quando há insucesso.

Algumas características da distribuição de Bernoulli.

$E(X) \quad = \quad 0\cdot(1-\theta)+1\cdot\theta=\theta$.

$\mathrm{Var}(X) \quad = \quad (1-\theta)\cdot\theta^2+\theta\cdot(1-\theta)^2=\theta\cdot(1-\theta)\cdot(\theta+1-\theta)=\theta\cdot(1-\theta)$.

$\gamma_1 \quad = \quad ((1-\theta)\cdot\theta^3+\theta\cdot(1-\theta)^3)/\sigma^3 =((1-\theta)\cdot\theta^3+\theta\cdot(1-\theta)^3)/(\theta\cdot(1-\theta)\cdot\sigma)=(\theta^2+(1-\theta)^2)/\sigma$

$= \quad (1-2\theta)/\sigma$.

Exemplo. Seja X uma variável aleatória associada ao lançamento de uma "moeda perfeita". X tem distribuição de Bernoulli.

Convencionando a saída de "face" como sendo "sucesso", e de "coroa", como um insucesso, tem-se:

$P(X=0)=1/2=\theta$ e $P(X=1)=1-\theta=1/2$.

Logo,

$f_X (x|1/2)=1/2$, para $x\in\{0,1\}$.

Definição (Distribuição binomial). Seja X uma variável aleatória discreta com função probabilidade dada por

$$f_X(x|\theta)=\begin{cases}\binom{a}{b}\ \theta^x\cdot(1-\theta)^{n-x} & se \quad x\in\{0,1,2,\ldots,n\}\\ 0 \quad se \quad outros\ casos\end{cases}\ ,\ para\ 0<\theta<1\,.$$

Neste caso, diz-se que X é uma distribuição binomial, e representa-se por $X\sim B(n;\theta)$.

Nota. A distribuição binomial aparece associada a uma sucessão de provas de Bernoulli independentes, isto é, uma sucessão de experiências aleatórias independentes, em cada uma das quais se observa a realização de determinado acontecimento A, com probabilidade $P(A)=\theta$, constante de prova a prova. Associado a essa sucessão de provas, aparece muitas vezes o seguinte problema: "qual a probabilidade para que, em n provas de Bernoulli independentes, se obtenham x sucessos, onde $x\in\{0,1,2,\ldots,n\}$, seja qual for a ordem em que esses sucessos são obtidos?"

Então, esta probabilidade calcula-se aplicando a função probabilidade da distribuição binomial.

Algumas características da distribuição binomial

$E(X) \quad = \quad n\cdot(n-1)\cdot\theta^2+n\theta$.

$\mathrm{Var}(X) \quad = \quad n\theta\cdot(1-\theta)$.

$\gamma_1 \quad = \quad (1-2\theta)/\sigma$.

Exemplo. Qual a probabilidade para, em cinco lançamentos de um "dado perfeito", se obter duas vezes a "sena"?

Nota-se que a variável aleatória associada ao lançamento de um "dado perfeito" é uma distribuição de Bernoulli, e que a experiência aleatória será realizada cinco vezes. Então, a probabilidade pedida pode ser determinada a partir da função probabilidade da distribuição binomial, tomando $\theta=1/6$, a probabilidade de sair uma "sena", no lançamento de um "dado perfeito", $n=5$, o número de provas, e $x=2$, o número de sucessos pretendidos:

$$P(X=2)=\binom{5}{2}\theta^x\cdot(1-\theta)^{(n-x)} = \frac{5!}{3!\cdot 2!}\cdot(1/6)^2\cdot(5/6)^3$$

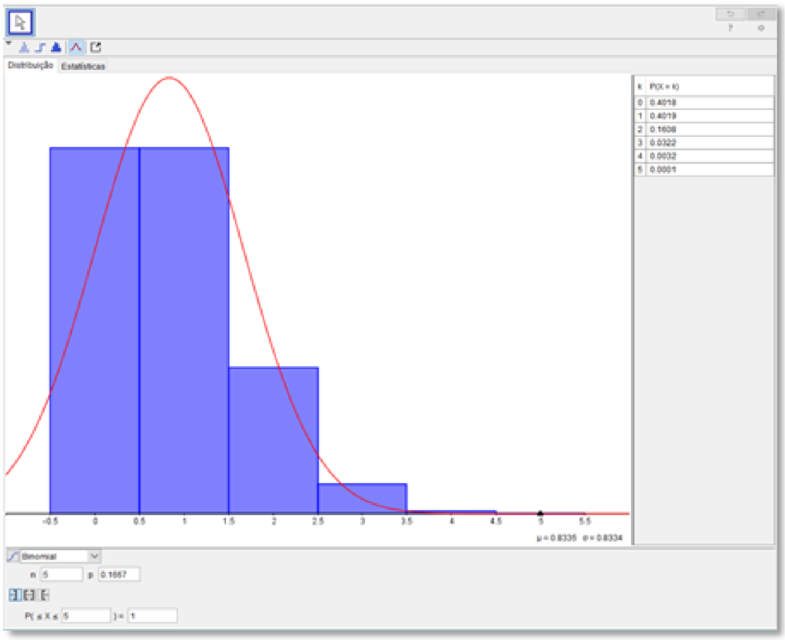
$$= 10\cdot 1/36\cdot 125/216$$

$$= 77/479 \cong 0{,}161$$

Nota.

A distribuição de Bernoulli pode obter-se a partir de uma distribuição binomial, fazendo $n=1$.

O cálculo das probabilidades binomiais torna-se laborioso mesmo para valores de n relativamente baixos. No entanto, recorrendo a meios computacionais adequados, pode-se obter essas probabilidades sem muitas dificuldades. Por exemplo, para o problema anterior, pode-se obter as probabilidades fazendo uso do software "Geogebra", como se pode ver na figura que se segue:



Há tabelas previamente preparadas donde se pode obter os valores de probabilidades, como por exemplo.

Para determinar probabilidades a partir de tabelas quando $\theta>0{,}50$, utiliza-se a seguinte relação:

$$X\sim B(n;\theta) \Leftrightarrow (n-X)\sim B(n;1-\theta)$$,

ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 172).

Exemplo. Num armazém estão preparadas para distribuição 500 embalagens de um produto, das quais exatamente 50 estão deterioradas. É efetuada uma inspeção sobre uma amostra de 10 embalagens escolhidas ao acaso com reposição.

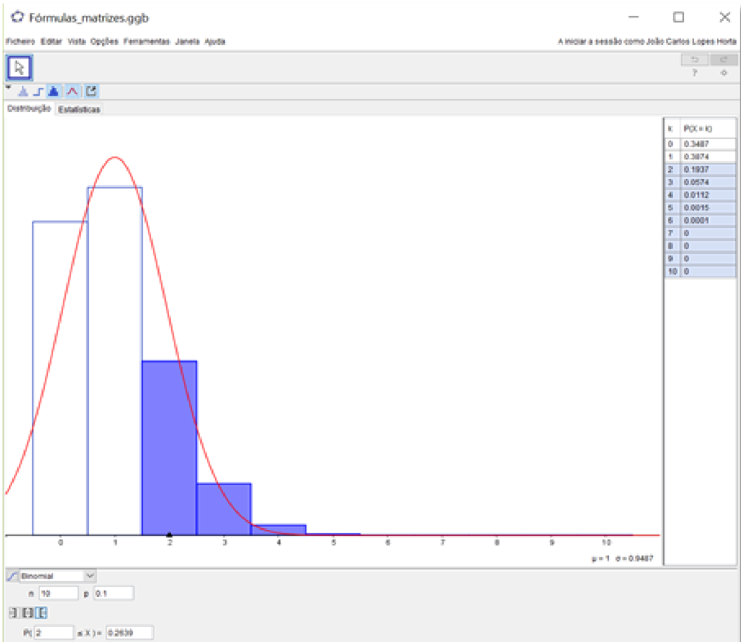
Qual a probabilidade de que a inspeção rejeite a comercialização do lote, quando só é admitida, no máximo, uma embalagem deteriorada na amostra?

Que dimensão mínima deve ter a amostra para que, adotado o mesmo critério, a probabilidade de rejeição seja superior a 0,5?

Alínea (a). Seja X o número de embalagens deterioradas numa amostra de 10. Então X~B(10;0,1), pois as extrações fazem-se com reposição e a probabilidade de um "sucesso", neste caso embalagem deteriorada, é θ=0,1. Então, a probabilidade pedida é:

$P(X>1)=\sum_{x=2}^{10}\binom{10}{x}(0,1)^x\cdot(0,9)^{10-x}=0.2639$.

Ver a figura que segue, que mostra a computação desta probabilidade utilizando o software Geogebra.



Alínea (b). Pretende-se saber o valor de n de modo que a P(X>1)>0,5, onde X é o número de embalagens deterioradas numa amostra com n embalagens. De acordo com os dados do problema, X~B(n;0,1).

$P(X>1)>1\Leftrightarrow\sum_{x=1}^{n}\binom{n}{x}(0,1)^x\cdot(0,9)^{n-x}>0,5\Leftrightarrow n=17$.

Definição (Distribuição geométrica). Seja X uma variável aleatória discreta, com funão probabilidade definida por

$$f_X(x|\theta)=\begin{cases}(1-\theta)^{x-1}\cdot\theta & se \quad x\in D_X=\{1,2,3,\ldots,\} \\ 0 \quad se \quad outros\ casos\end{cases} \quad , para\ 0<\theta<1 .$$

Neste caso, diz-se que X é uma distribuição geométrica, ou distribuição Pascal.

Nota. Uma variável aleatória discreta X segue uma distribuição geométrica quando ela está associada a observação do número provas de Bernoulli até que se obtenha um sucesso. Neste caso, o suporte da variável aleatória discreta X é D_X={1,2,3,…}.

Algumas características da distribuição geométrica

E(X) = 1/θ.

Var(X) = (1-θ)/θ^2 .

P(a≤X≤b) = (1-θ)^(a-1)-(1-θ)^b.

P(X≥a) = (1-θ)^(a-1).

P(X≤b) = 1-(1-θ)^b.

Nota. A partir da "característica 5.", acima referida, pode-se concluir que a função de distribuição de uma variável aleatória com distribuição geométrica é dada por

$$F_X(x) = \begin{cases} 0 \quad se \quad x < 1 \\ 1-(1-\theta)^k \quad se \quad k \leq x < k+1\,, \end{cases} \quad para\; k \in \{1,2,3,\ldots\}\;.$$

Exemplo. Um indivíduo que faz anos em 15 de janeiro dispõe-se a inquirir pessoas ao acaso até encontrar uma que faça anos também no mesmo mês.

Qual o número médio de inquirições que tem de fazer até ter um sucesso?

Qual a probabilidade de inquirir menos de 6 pessoas?

Alínea (a). Está-se a pedir E(X), onde X é a avariável aleatória discreta associada a observação do número de indivíduos inquiridos até que se obtenha um que tenha nascido no mês de janeiro, ou seja, X tem distribuição geométrica.

Suponha-se que o acontecimento A é "pessoa que nasceu no mês de janeiro", e que a probabilidade de uma pessoa nascer num determinado mês é 1/12, ou seja, todos os meses têm igual probabilidade.

Então, θ=1/12, logo E(X)=1/θ=12.

Alínea (b). Pretende-se determinar a probabilidade da variável aleatória X ser menor que 6, ou seja, X igual a 5, ou 4, ou 3, ou 2, ou 1.

P(X<6)=P(X≤5)=1-(1-1/12)^5=1-(11/12)^5=0,3528 .

Definição (Distribuição binomial negativa). Seja r um número natural, isto é, r∈N={1,2,3,…}. Seja X uma variável aleatória discreta com função probabilidade dada por

$$f_X(x|\theta) = \begin{cases} \binom{x-1}{r-1}\,\theta^r \cdot (1-\theta)^{x-r} \quad se \quad x \in \{r, r+1, r+2, \ldots\} \\ 0 \quad se \quad outros\ casos \end{cases} \;, \quad para\; 0 < \theta < 1\;.$$

Neste caso, diz-se que X tem distribuição binomial negativa, e representa-se por X~BN(r;θ).

Nota.

Para $r \in \mathbb{N}$, o suporte da v. a. discreta X é $D_X=\{r,r+1,r+2,\ldots\}$.

A distribuição geométrica obtém-se a partir da distribuição binomial negativa fazendo r=1. Assim, se X segue uma distribuição geométrica, escreve-se $X \sim BN(1;\theta)$.

Uma variável aleatória discreta X segue uma distribuição binomial negativa, $X \sim BN(r;\theta)$, $r \in \mathbb{N}$, quando ela está associada a observação do número provas de Bernoulli até que se obtenha r sucessos. Seja $x \in D_X=\{r,r+1,r+2,\ldots\}$. Se somente ao fim de x provas se completam r sucessos, então na x-ésima prova tem de obter-se um sucesso, o r-ésimo sucesso, e nas x-1 provas anteriores tem de obter-se r-1 sucessos e x-r insucessos. Logo

$$
\begin{aligned}
P(X=x) &= P(\text{r-1 sucessos em x-1 provas}) \cdot P(\text{1 sucesso na x-ésima prova}) \\
&= \binom{x-1}{r-1} \theta^{r-1} (1-\theta)^{x-1-(r-1)} \theta \\
&= \binom{x-1}{r-1} \theta^{r} (1-\theta)^{x-r}
\end{aligned}
$$

$BN(r,\theta)$ representa o número de insucessos que ocorrem em uma sequência de provas de Bernoulli antes de se atingir os r sucessos.

Algumas características da distribuição binomial negativa – $X \sim BN(r;\theta)$.

$$
\begin{aligned}
E(X) &= r/\theta. \\
Var(X) &= r(1-\theta)/\theta^2. \\
\gamma_1 &= r\theta/\sqrt{1-\theta}.
\end{aligned}
$$

Exemplo. Existe uma técnica de amostragem – bastante útil em algumas situações – que consiste em recolher observações independentes de determinada população até se observarem r unidades com determinada característica. Com esta técnica, o número de elementos a inquirir tem distribuição binomial negativa.

Suponha-se que se pretendia inquirir 100 apoiantes de determinado partido político, sobre a apreciação que fazem sobre a liderança do respetivo partido. A probabilidade de se ter de observar pelo menos 1000 pessoas será dada por

$$
\begin{aligned}
P(X \geq 1000) &= \sum_{x=1000}^{+\infty} \binom{x-1}{99} \cdot \theta^{100} \cdot (1-\theta)^{x-100} \\
&= 1-P(X<1000) \\
&= 1-\sum_{x=100}^{999} \binom{x-1}{99} \cdot \theta^{100} \cdot (1-\theta)^{x-100},
\end{aligned}
$$

onde θ representa a proporção de apoiantes desse partido na população. Se essa proporção for $\theta=0{,}1$, ter-se-á:

$$
\begin{aligned}
P(X \geq 1000) &= 1-\sum_{x=100}^{999} \binom{x-1}{99} \cdot (0{,}1)^{100} \cdot (1-0{,}1)^{x-100} \\
&= 1-\sum_{x=100}^{999} \binom{x-1}{99} \cdot (0{,}1)^{100} \cdot (0{,}9)^{x-100} \\
&= 1- 0.5154177 \\
&= 0.4845823
\end{aligned}
$$

Nota. Esta probabilidade tem que ser calculada com a ajuda de um computador. Por exemplo, este resultado foi obtido com a ajuda do software "R":

```
R Console
> pnbinom(q=900,size=100,prob=0.1)
[1] 0.5154177
> ProbXmaior1000 <- 1- (pnbinom(q=900,size=100,prob=0.1))
> ProbXmaior1000
[1] 0.4845823
> |
```

Alternativamente, pode-se proceder da seguinte forma, para obter o mesmo valor, utilizando o software "R":

```
RGui (64-bit) - [R Console]
File  Edit  View  Misc  Packages  Windows  Help
> prob <- numeric(0)
> for (i in 1:1000) {prob[i] <- if (i>=100) {choose(i-1,99)*((0.1)^100)*(0.9)^(i-100)} else {0}}
> sum(prob)
[1] 0.5154177
> probXmaior1000 <- (1-sum(prob))
> probXmaior1000
[1] 0.4845823
> |
```

Definição (Distribuição hipergeométrica). Sejam N, M e n números naturais arbitrários, com N maior que M e n. Seja X uma variável aleatória com suporte

$D_X=\{x\in N_0:\max\{0,n-(N-M)\}\leq x\leq \min\{n,M\}\}$.

Se a variável aleatória X tem função probabilidade dada por

$$f_X(x|N,M,n)=\begin{cases}\dfrac{\binom{M}{x}\cdot\binom{N-M}{n-x}}{\binom{N}{n}} & se \quad x\in D_X \\ 0 \quad se \quad x\notin D_X & \end{cases},$$

diz-se que tem distribuição hipergeométrica, e representa-se por X~H(N,M,n).

Nota. De uma certa forma, pode-se comparar a distribuição hipergeométrica com a distribuição binomial. Ambas consistem em determinar a probabilidade de numa amostra de n elementos, extraída de uma população com dimensão N, se obter x "sucessos" (um sucesso é a observação de um acontecimento A, com probabilidade θ, que, no presente contexto, representa cada um dos M elementos, da população de dimensão N, com determinados atributos). No entanto, quando se trata de uma distribuição binomial, as provas de Bernoulli são independentes, pois as extrações são feitas com reposição, ou seja, a probabilidade do acontecimento A não se altera de prova para prova; mas, quando se trata de uma distribuição hipergeométrica as provas de Bernoulli não são independentes, pois as extrações são feitas sem reposição, o que significa que a probabilidade do acontecimento A não se mantém constante de prova para prova.

Geralmente, na distribuição hipergeométrica, n é inferior a M e a N-M. Nesse caso, $D_X=\{0,1,2,3,\dots\}=N_0$.

Algumas características da distribuição hipergeométrica – $X\sim H(N,M,n)$.

$$E(X) = n\cdot M/N \, .$$

$$Var(X) = n\cdot M/N\cdot(N-M)/N\cdot(N-n)/(N-1) \, .$$

No presente caso, M/N é a probabilidade do acontecimento A – $P(A)=M/N=\theta$. Então:

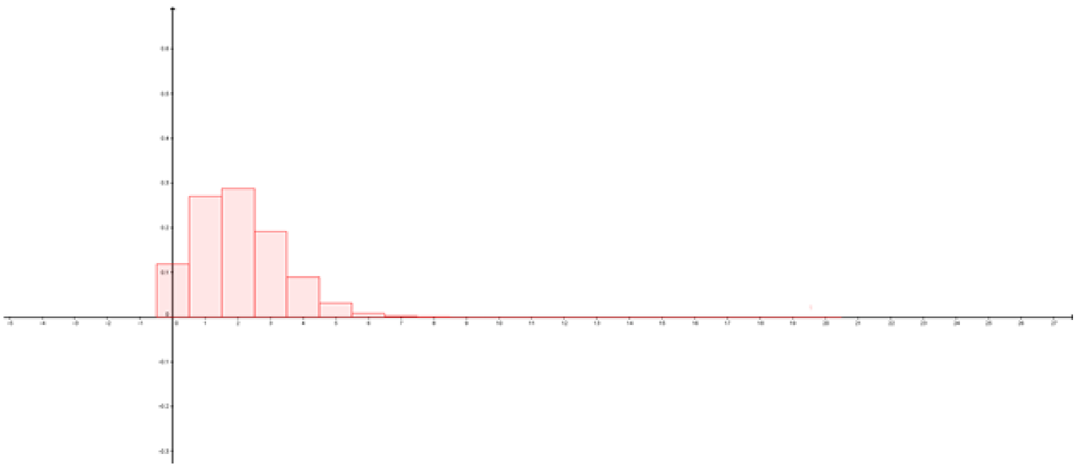
$E(X)=n\cdot\theta$ e $Var(X)=n\cdot\theta\cdot(1-\theta)\cdot(N-n)/(N-1)$.

Exemplo. De um grupo de 1000 explorações agrícolas, de uma certa região, há 20 que produzem espargos. Recolhida ao acaso uma amostra de 100 explorações, sem reposição, qual é a probabilidade de observarem 5 produtores de espargos?

Neste caso, a variável aleatória discreta tem suporte $D_X=\{0,1,2,\dots,18,19,20\}$. Então,

$P(X=9)=f_X(9\mid 1000,20,100)=\binom{20}{9}\cdot\binom{980}{91}/\binom{1000}{100}\cong 0{,}00004143$.

O gráfico que se segue representa a distribuição de probabilidades de X, em D_X.



Definição (Processo de Poisson). Suponha-se que se procede à contagem do número de acontecimentos (chegada de doentes, chegada de navios, etc.) ocorridos ao longo do tempo. Tem-se um processo de Poisson com parâmetro $\lambda>0$ quando se verificam as seguintes condições:

- O número de acontecimentos que ocorrem em dois intervalos, de tempo, disjuntos são independentes;
- A probabilidade de ocorrer exatamente um acontecimento em qualquer intervalo, de tempo, de amplitude Δt arbitrariamente "pequena" é aproximadamente $\lambda\Delta t$;
- A probabilidade de ocorrerem dois ou mais acontecimentos em qualquer intervalo, de tempo, de amplitude Δt arbitrariamente "pequena" é aproximadamente igual a zero.

Definição (Distribuição de Poisson). Seja X uma variável aleatória discreta com função probabilidade dada por

$$f_X(x|\lambda) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & se \quad x \in \{0,1,2,3,\ldots\} \\ 0 \quad se & outros\ casos \end{cases} \quad , \quad com\ \lambda > 0\,.$$

Neste caso, diz-se que X tem uma distribuição de Poisson, e representa-se por $X \sim Po(\lambda)$.

Algumas características da distribuição de Poisson – $X \sim Po(\lambda)$.

$E(X) = \lambda$.

$Var(X) = \lambda$.

$\gamma_1 = \lambda^{-1/2}$.

Se os acontecimentos num processo de Poisson ocorrem a uma taxa média de λ, por unidade de tempo, então o número de ocorrências num intervalo de amplitude t (positivo) representado pela variável aleatória X(t), tem distribuição de Poisson de parâmetro λt,

$$f(x|\lambda t) = P(X(t) = x|\lambda t) = \frac{e^{-\lambda t}\,(\lambda t)^x}{x!}\,, \quad x \in \{0,1,2,3,\ldots\}\,,\ \lambda > 0\,.$$

Nota. Pode então concluir-se que o número esperado de ocorrências num intervalo de amplitude t é λt.

O cálculo de probabilidades usando distribuição de Poisson deve ser feito utilizando meios computacionais adequados. Também pode-se auxiliar numa tabela de Poisson.

Exemplo. Numa fábrica de têxteis existem numerosos teares do mesmo tipo. Depois de muitas observações, chegou-se à conclusão de que o número de teares avariados segue um processo de Poisson com taxa média de 3 por mês. Logo, se X é a variável aleatória que representa o número de teares avariados por mês, tem-se $X \sim Po(3)$.

Qual é a probabilidade para que durante um mês se avariem sete ou mais teares?

Determine a capacidade mensal mínima disponível, C, da oficina de reparação, de modo a ser pelo menos 0.9 a probabilidade de não haver teares a aguardar reparação.

Alínea (a). Calcula-se:

$$P(X \geq 7) = \sum_{x=7}^{+\infty} \frac{e^{-3}\,3^x}{x!}$$

$$= 0.0335$$

Este cálculo foi realizado com o software "Geogebra" – Veja o Gráfico 1 - $P(X \geq 7)$ numa Po(3).

Pode-se também determinar a mesma probabilidade, procedendo da seguinte forma:

$$P(X \geq 7) = 1 - P(X \leq 6)$$

$$= 1 - 0.9665$$

$$= 0.0335$$

O Cálculo de $P(X \leq 6)$ foi realizado com o software "Geogebra" – Veja o Gráfico 2 - $P(X \leq 6)$ numa Po(3).

Alínea (b). Determina-se C, o menor inteiro a verificar:

$$P(X\geq C)=\sum_{x=0}^{C} \frac{e^{-3}\,3^x}{x!}\geq 0.9 \ .$$

$\Leftrightarrow \quad C=5$

Este cálculo foi realizado utilizando o software "R" – Veja o Gráfico 3 – Determinação do C em Po(3).

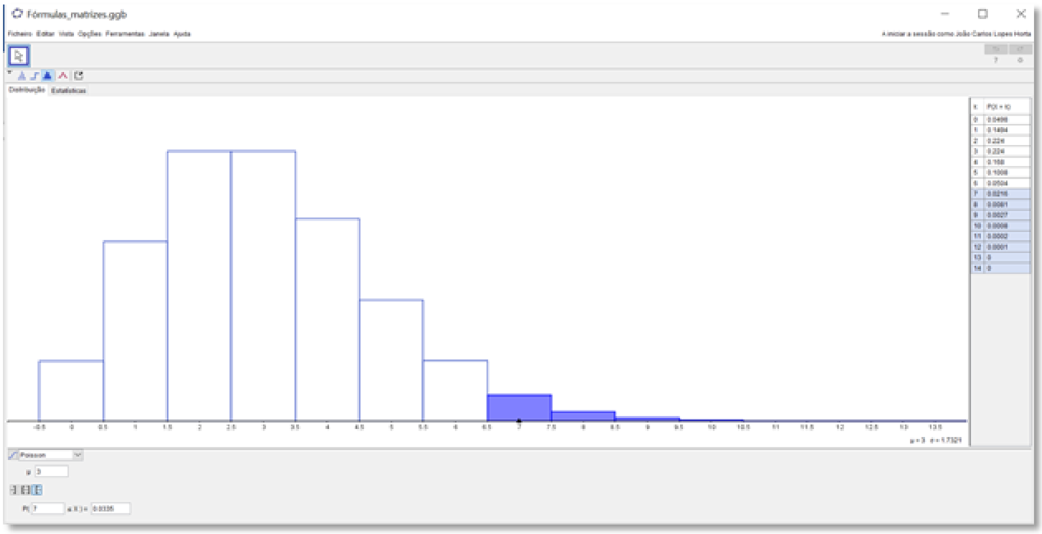


Gráfico 1 - P(X≥7) numa Po(3).

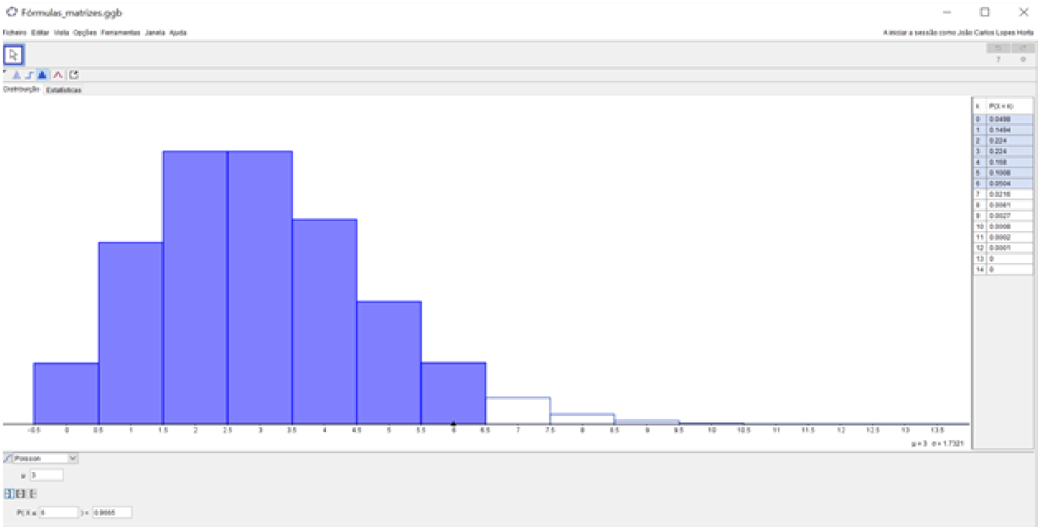


Gráfico 2 - P(X≤6) numa Po(3)

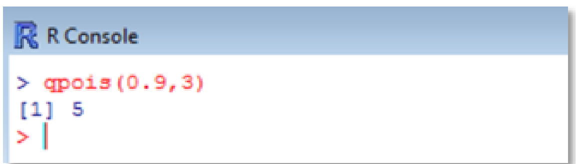


Gráfico 3 – Determinação do C em Po(3)

Conclusão

As distribuições de variáveis aleatórias discretas têm inúmeras aplicações na resolução de problemas de várias ordens.

Consoante o tipo de problema, pode-se ajustar a variável aleatória em estudo a uma das distribuições aqui estudadas (uniforme, Bernoulli, binomial ou hipergeométrica).

Os valores de valor esperado, variância e/ou desvio padrão, coeficiente de variação e coeficiente de assimetria permitem caracterizar a distribuição de uma variável aleatória (neste caso, v. a. Discreta).

Algumas distribuições são casos particulares de outras, mas por terem várias aplicações, os seus estudos fizeram-se em separados.

Avaliação

1. A variável aleatória X tem função probabilidade definida por f_X (x)=x/10, x∈D_X={0,1,2,3,4} .
   a. Calcule E(X) e Var(X).
   b. Seja Y=X^2. Determine Var(Y) e E(Y).
2. Sabe-se que o primeiro e o segundo momentos em relação à origem da v. a. discreta X são iguais a 6 e a 62, respectivamente. Sendo Y=X/2+3, determine a média, a variância e o desvio padrão de Y.
3. Um jogo consiste em baralhar e tirara duas cartas (com reposição) de um baralho vulgar (52 cartas); se saírem duas cartas de copa ganha-se 15 euros; em caso contrário, perde-se 1 euro.
   a. Qual a probabilidade de ganhar duas vezes seguidas o jogo?
   b. Em média, quando se ganha ou se perde de cada vez que se joga?
4. Num teste de resposta múltipla, cada pergunta tem cinco respostas possíveis, e apenas uma delas é correta. Admita que o aluno responda sequencialmente e ao acaso as perguntas.
   a. Qual a probabilidade de a primeira pergunta certa ser a quarta?
   b. Em média, a quantas perguntas tem de responder para acertar uma?
5. Qual a probabilidade de, no lançamento de uma moeda, se obter a terceira face no sétimo lançamento?

6. Um fornecedor de fechaduras para portas blindadas envia as peças em lotes de 100 fechaduras, para 50 fábricas de portas blindadas, independentes. Devido a falhas de produção resultou defeituosa uma série de fechaduras. Para evitar prejuízos decidiu incluir em cada um dos lotes 5 dessas fechaduras, na esperança de passarem despercebidas no controlo de qualidade dos clientes, que consiste em escolher duas fechaduras ao acaso, em cada lote, devolvendo o mesmo se alguma tiver defeito.

   a. Calcule a probabilidade de um lote ser aceite pela fábrica que o recebeu.

   b. Qual a probabilidade de no controlo de qualidade de um lote haja pelo menos umafechadura sem defeito?

7. Seja X uma variável aleatória com distribuição de Poisson de parâmetro $\lambda=5$. Calcule as seguintes probabilidades:

   a. $P(X=5)$.

   b. $P(X\geq 1)$.

   c. $P(4<X<8)$.

8. Uma fábrica produz azulejos que são embalados em caixas de 20 unidades. Sabe-se que o número de defeitos por azulejo segue um processo de Poisson com taxa média igual a 0.1.

   a. Qual a percentagem de azulejos com defeitos?

   b. Calcule a probabilidade de numa caixa existirem mais de dois azulejos com defeitos.

   c. Se a aceitação de um fornecimento de azulejos for feita na condição de o número total de defeitos encontrados numa amostra de 100 azulejos não exceder 15, com que probabilidade é rejeitado o fornecimento?

9. Suponha que 70% das consolas de PlayStation 3 não estão defeituosas. Encontre a probabilidade de encontrar a segunda consola sem defeito:

   a. Em três tiragem.

   b. Em no máximo seis tiragens.

   c. Em pelo menos cinco tiragens.

Atividade 3 – Distribuições contínuas

Introdução

Nesta atividade introduz-se algumas medidas de tendência central e, outras, de dispersão (valor esperado, momento, quantil, moda, etc.), relativamente à uma variável aleatória contínua. Estuda-se algumas distribuições de variável aleatória contínua.

Detalhes da atividade

Definição (Valor esperado de uma variável aleatória contínua). Seja X uma variável aleatória contínua, com função de densidade f_X. A expressão

$$E(X)=\int_{-\infty}^{+\infty} x f_X(x)\,dx \ ,$$

quando

$$\int_{-\infty}^{+\infty} |x|\, f_X(x)\,dx<+\infty \ ,$$

define o valor esperado, média ou esperança matemática de X.

Nota. A condição

$$\int_{-\infty}^{+\infty} |x|\, f_X(x)\,dx<+\infty$$

destina-se a garantir, no caso do valor esperado E(X) ser uma integral impróprio, que o mesmo seja absolutamente convergente, verificando, desta feita, a unicidade do valor esperado E(X).

No entanto, para o presente trabalho, parte-se de princípio que a convergência do integral já está verificada, sempre que se está a calcular o valor esperado de uma variável aleatória contínua.

Exemplo. Seja X uma variável aleatória contínua, com função densidade

$$f_X(x)=3x^{-4} \ , \quad \text{para } x>1 \ .$$

Então, o valor esperado de X é dado por

$$\begin{aligned} E(X) &= \int_{-\infty}^{+\infty} x f_X(x)\,dx \\ &= \int_{-\infty}^{1} (x\cdot 0)\,dx+\int_{1}^{+\infty} x\cdot(3x^{-4})\,dx \\ &= 3/2 \ . \end{aligned}$$

Definição (Valor esperado de uma função de variável contínua)**. Seja X uma variável aleatória** contínua, e seja φ uma função real de variável real, com domínio R. A expressão

$$E(\varphi(X))=\int_{-\infty}^{+\infty} \varphi(x)\, f_X(x)\,dx \ ,$$

É o valor esperado da variável aleatória φ(X) (que pode ser contínua ou não), sob a condição de E(φ(X)) ser absolutamente convergente, em caso de ser um integral impróprio.

Exemplo. Seja a variável aleatória contínua X considerada no “exemplo anterior”. Sejam as funções reais de variáveis reais φ_1 e φ_2, definidas por

$$\varphi_1(x)=x^2$$

e

$$\varphi_2(x)=f(x)=\begin{cases}0, & x<2\\ 1, & x\geq 2\end{cases}.$$

Sejam as variáveis aleatórias $Y=\varphi_1(X)$ e $Z=\varphi_2(X)$.

Os valores esperados das variáveis aleatórias Y e Z são:

$$\begin{aligned} E(Y)=E(\varphi_1(X)) &= \int_{-\infty}^{+\infty} \varphi_1(x)\, f_X(x)\,dx \\ &= \int_1^{+\infty} x^2\cdot(3x^{-4})\,dx \\ &= \left[3\cdot x^{-1}\right]_1^{+\infty} \\ &= 3 \end{aligned}$$

e

$$\begin{aligned} E(Z)=E(\varphi_2(X)) &= \int_{-\infty}^{+\infty} \varphi_2(x)\, f_X(x)\,dx \\ &= \int_1^{2} 0\cdot(3x^{-4})\,dx+\int_2^{+\infty} 1\cdot(3x^{-4})\,dx \\ &= \left[-x^{-3}\right]_2^{+\infty} \\ &= 1/8\ . \end{aligned}$$

Nota. As “propriedades dos valores esperados” vistas na “Atividade 1.2” são válidas também para variáveis aleatórias contínuas.

Definição (Momento de ondem k em relação à origem). Seja k um número natural arbitrário, ou seja $k\in N=\{1,2,3,\ldots\}$, e seja X uma variável aleatória contínua. O valor esperado,

$$\mu_k'=E(X^k)=\int_{-\infty}^{+\infty} x^k\cdot f_X(x)\,dx\ ,$$

se existir, define o momento de ordem k em relação à origem, ou momento ordinário, de ordem k da variável aleatória contínua X.

Definição (Momento de ordem k em relação à média). Seja k um número natural arbitrário, ou seja $k\in N=\{1,2,3,\ldots\}$, e seja X uma variável aleatória contínua. O valor esperado,

$$\mu_k=E(X-\mu)^k=\int_{-\infty}^{+\infty} (x-\mu)^k\cdot f_X(x)\,dx\ ,$$

se existir, é o momento de ordem k em relação à média, ou momento central de ordem k da variável aleatória contínua X.

Definição (Variância de uma v. a. contínua). A variância de uma variável aleatória contínua X é dada por

$$Var(X)=\sigma_X^2=E((X-\mu)^2)=\int_{-\infty}^{+\infty} (x-\mu)^2\cdot f_X(x)\,dx\ ,$$

se existir o valor esperado.

Nota. As “propriedades de variância” vistas na “Atividade 1.2” são válidas para as variáveis aleatórias contínuas.

Os conceitos de "desvio padrão" (raíz quadrada da variância), "coeficiente de variação" e "coeficiente de assimetria" definem-se da mesma maneira que na "Atividade 1.2".

Definição (Quantil). Seja X uma variável aleatória contínua, com função densidade $f_X(x)$, e seja α um número real a verificar $0<\alpha<1$. Quantil de ordem α, ξ_α, é um valor de X que satisfaz a condição

$$\int_{-\infty}^{\xi_\alpha} f_X(x)\,dx=\alpha \Leftrightarrow F_X(\xi_\alpha)=\alpha .$$

Nota. O quantil de ordem α existe sempre, embora a equação $F_X(\xi_\alpha)=\alpha$ possa ter mais que uma solução; nesse caso, convenciona-se que o quantil de ordem α é a menor das soluções da referida equação.

A definição do quantil pode apresentar-se para qualquer variável aleatória, nomeadamente para variáveis aleatórias discretas. Mas, do ponto de vista prático, os quantis para variáveis aleatórias contínuas têm mais interesses.

Os quantis são conhecidos por parâmetros de ordem.

O conceito de quantil apresentado aquando do estudo da estatística descritiva, refere-se a coleções de dados ou amostras, logo trata-se de uma noção empírica.

Doravante, o que se quer é determinar um quantil de uma variável aleatória, ou distribuição, contínua, isto é, no âmbito do modelo probabilístico.

Definição (Mediana). Seja X uma variável aleatória contínua, com função densidade $f_X(x)$. Mediana de X, representada por μ_e ou med(X), é o quantil de ordem $\alpha=0{,}5$, $\xi_{0,5}$, ou seja, é a solução ("conveniente") da equação (na varável $\xi_{0,5}$)

$$\int_{-\infty}^{\xi_{0,5}} f_X(x)\,dx=0{,}5 .$$

Nota. A mediana é o valor de X tal que metade da massa de probabilidade se encontra à sua esquerda e a outra metade à sua direita, ela pode ser utilizada como medida de localização da distribuição, à semelhança da média (ou valor esperado).

Definição (Moda). Seja X uma variável aleatória contínua, com função densidade $f_X(x)$. Moda, representado por μ_*, é um valor de X que verifica a condição

$$f_X(x)\le f_X(\mu_*) \ , \quad \forall x\in R .$$

Nota. A moda não é um parâmetro de ordem. Contudo, trata-se de uma importante medida de localização, que deve ser considerada juntamente com a média e com a mediana.

No caso de distribuições simétricas unimodais, a média, a mediana e a moda são iguais.

Definição (Quartis). Seja X uma variável aleatória contínua, com função densidade $f_X(x)$. O primeiro quartil de X é o quantil de ordem $\alpha=0{,}25$, $\xi_{0,25}$, e delimita superiormente o primeiro quarto da massa de probabilidade; o terceiro quartil de X é o quantil de ordem $\alpha=0{,}75$, $\xi_{0,75}$, e delimita inferiormente o último quarto de massa de probabilidade; a mediana é o segundo quartil.

Definição (Amplitude interquartis). A amplitude interquartis de uma variável aleatória contínua X, representado por AIQ, é uma medida de dispersão absoluta, que se define por:

$AIQ=\xi_{0,75}-\xi_{0,25}$.

Pode-se obter uma medida de dispersão relativa, calculando

$(\xi_{0,75}-\xi_{0,25})/\mu_e$.

Nota. O intervalo $]\xi_{0,25},\xi_{0,75}[$, denominado intervalo interquartis, abrange 50% da massa de probabilidade.

Quando $\alpha=0,1\ ,0,2\ ,\ldots,0,9$ obtêm-se os decis; quando $\alpha=0,01\ ,0,02\ ,\ldots,0,99$ têm-se os percentis.

Nota. O quadro que se segue, apresenta o resumo das medidas de localização e dispersão, estudadas aqui neste texto.

Medidas	μ	σ	CV	μ_*	μ_e	AIQ	$\frac{AIQ}{\mu_e}$
Localização	X			X	X		
Dispersão Absoluta		X				X	
Dispersão relativa			X				X

Exemplo. Seja X uma variável aleatória contínua, com função densidade

$f_X(x)=3x^{-4}$,   para $x>1$.

(Ver o "exemplo acima").

O valor esperado de X é

$E(X)=\mu=3/2$,

pelos cálculos realizados anteriormente.

A variância obtém-se a plicando a definição:

$$\begin{aligned} Var(X)=\sigma^2 &= \int_{-\infty}^{+\infty}(x-\mu)^2\cdot f_X(x)\,dx \\ &= \int_1^{+\infty}(x-3/2)^2\cdot(3x^{-4})\,dx \\ &= 3/4 \end{aligned}$$

O desvio padrão, raíz quadrada positiva da variância, é:

$\sigma=\sqrt{3}/2$.

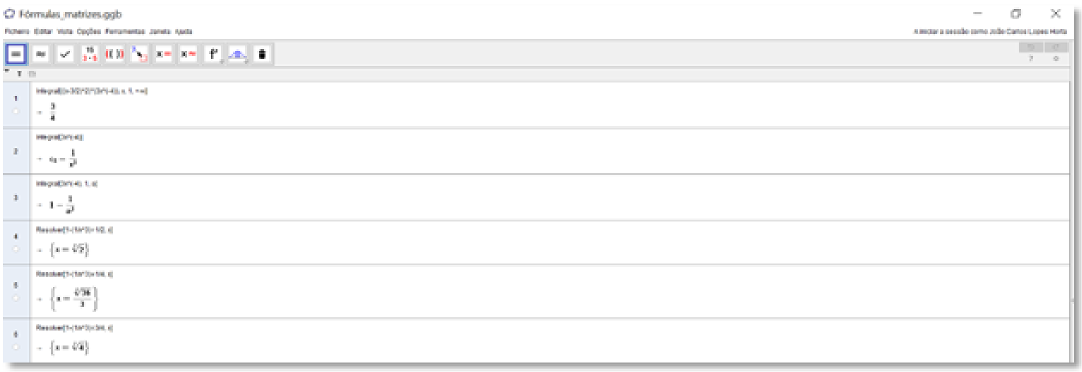
$$\int_{-\infty}^{\xi_{0,5}} f_X(x)\,dx=0,5 \quad\Leftrightarrow\quad \int_1^{\xi_{0,5}} 3x^{-4}\,dx=0,5$$

$$\Leftrightarrow \quad \xi_{0,5}=\sqrt[3]{2}=\mu_e$$

O primeiro e o terceiro quartis são, respetivamente,

$Q_1=\sqrt[3]{36}/3$ e $Q_3=\sqrt[3]{4}$.

Nota. Os cálculos realizados acima foram feitos com a ajuda do software "Geogebra – Folha CAS".



O intervalo interquartis é

$AIQ=Q_3-Q_1=\sqrt[3]{4}-\sqrt[3]{36}/3=(3\sqrt[3]{4}-\sqrt[3]{36})/3$.

As outras medidas podem ser calculadas a partir destas.

Definição (Distribuição uniforme). Seja X uma variável aleatória contínua, e sejam α, β números reais. Diz-se que X tem distribuição uniforme no seu suporte $D_X=\,]\alpha,\beta[$, com $\alpha<\beta$, e representa-se por $X\sim U(\alpha,\beta)$, quando a função densidade é dada por

$$f_X(x|\alpha,\beta)=\begin{cases}\dfrac{1}{\beta-\alpha} & se \quad x\in\,]\alpha,\beta[\\ 0 \quad se & outros\ casos\end{cases} .$$

A correspondente função de distribuição é definida por

$$F_X(x|\alpha,\beta)=\begin{cases}0 \quad se \quad x<\alpha\\ \dfrac{x-\alpha}{\beta-\alpha} \quad se \quad x\le x<\beta\\ 1 \quad se \quad x\ge\beta\end{cases} .$$

Algumas características da distribuição hipergeométrica – $X\sim H(N,M,n)$.

$E(X) = (\alpha+\beta)/2$.

$Var(X) = (\beta-\alpha)^2/12$.

Nota. O caso particular em que $\alpha=0$ e $\beta=1$ – isto é, em que $X\sim U(0,1)$ – tem especial interesse a nível teórico-prático (para mais informações, ver (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 225)).

Teorema (Transformação uniformizante).

Seja X uma variável aleatória contínua com função de distribuição F_X estritamente crescente no intervalo aberto $]a,b[$, onde X tem densidade positiva – este intervalo pode ser ilimitado, isto é, pode acontecer $a=-\infty$ ou $b=+\infty$ – e $F_X(a)=0$ e $F_X(b)=1$. Então, a variável aleatória $Y=F_X(X)$ tem distribuição $U(0,1)$.

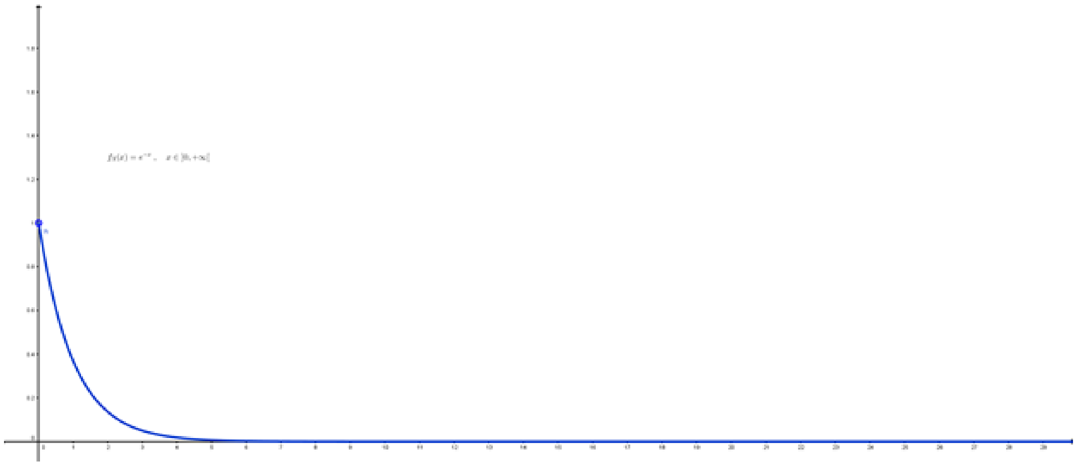
Seja Y uma variável aleatória com distribuição $U(0,1)$, e F_X a função de distribuição de uma variável aleatória contínua X. Supõe-se que F_X é estritamente crescente no intervalo aberto $]a,b[$, onde X tem densidade positiva – o intervalo pode ser ilimitado, isto é, $a=-\infty$ ou $b=\infty$ – e $F_X(a)=0$ e $F_X(b)=1$. Então, a variável aleatória $X=F_X^{-1}(Y)$ é contínua com função distribuição F_X.

A mudança de variável $Y=F_X(X)$ é chamada transformação uniformizante ou transformação de probabilidade.

Exemplo. Seja X uma variável aleatória com função densidade, definida no seu suporte $D_X=\,]0,+\infty[$ por

$f_X(x)=e^{-x}$.

Ora, $f_X(x)$ é positiva em D_X, ou seja, X tem densidade positiva em D_X.



Então, X satisfaz as condições da alínea (a) do "teorema anterior".

A função de distribuição de X é

$$F_X(x)=\begin{cases} -e^{-x}+1 & se \quad x>0 \\ 0 \quad se \quad x\le 0 \end{cases} .$$

Então, pode determinar-se a função densidade da variável aleatória $Y=1-e^{-X}$, obtendo-se $Y\sim(0,1)$, pelo referido teorema.

Definição (Distribuição normal). Seja X uma variável aleatória. Diz-se que X tem distribuição normal com parâmetros μ e σ^2, e representa-se por $X\sim N(\mu,\sigma^2)$, quando a função densidade é da forma

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

onde $\mu\in\mathbb{R}$ e $\sigma^2>0$.

Nota. Muitos dos atributos observáveis de determinada população são bem representados por variáveis aleatórias que seguem uma distribuição normal.

Os parâmetros da distribuição normal representam-se por μ e σ^2 porque correspondem, respetivamente, à média e a variância da variável aleatória em causa.

Segundo (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 227), a função de distribuição de uma variável aleatória que segue uma normal N(μ,σ^2) é dada por

$$F_X\left(x|\mu,\sigma^2\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(t-\mu)^2}\, dt\,,$$

para o qual não se conhece solução analítica.

Um caso importante da distribuição normal obtém-se quando μ=0 e σ^2=1. Nesta situação, a função densidade e a função distribuição são, respetivamente, dadas por

$$\phi\left(z\right) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}\,, \quad \Phi\left(z\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2}\, dt\,.$$

Teorema. Seja X uma variável aleatória contínua. Se X~N(μ,σ^2), a variável estandardizada

Z=(X-μ)/σ

tem distribuição N(0,1).

Nota. A distribuição N(0,1) é habitualmente designada por distribuição normal estandardizada.

Os símbolos ϕ e Φ são reservados para representar, respetivamente, funções densidade e de distribuição de uma variável aleatória que segue uma N(0,1).

Os cálculos envolvendo distribuição normal podem fazer-se com meios computacionais adequados (O software "R" dá resposta a qualquer problema proposto neste texto, mas quando os cálculos são acessíveis pode-se usar o software "Geogebra"). Quando não for possível utilizar os meios computacionais, pode-se utilizar a uma tabela normal – que normalmente refere-se a N(0,1).

A distribuição normal é simétrica em relação à reta x=μ. Em particular, tratando-se da distribuição normal estandardizada, tem-se:

$$\phi(x) = \phi(-x)\,, \quad \Phi(x) = 1 - \Phi(-x)\,.$$

Utilização da tabela da normal. Esta tabela diz respeito à função Φ, para valores não-negativos de z. Permite calcular para cada z a respetiva ordenada da função de distribuição Φ da normal estandardizada, N(0,1). Por exemplo, se z=0,53, então

$$\Phi(z) = \Phi(0,53) = P(z < 0,53) = P(z \leq 0,53) = 0,70194\,.$$

Este valor está na interseção da linha "0.5" e da coluna "0.03", ou seja, z=0,53=0,5+0,03 .

A simetria da distribuição normal estandardizada (é simétrica em relação à reta x=μ=0) garante que

$$\Phi(z) = 1 - \Phi(-z)\,.$$

Sendo assim, pode-se sempre determinar $\Phi(z)$ para valores negativos de z.

Quando $X \sim N(\mu,\sigma^2)$ faz-se a mudança de variável para variável estandardizada

$Z=(X-\mu)/\sigma$.

Assim sendo, tem-se

$$P(x_0<X<x_1) = P\left(\frac{x_0-\mu}{\sigma}<Z<\frac{x_1-\mu}{\sigma}\right)=\Phi\left(\frac{x_1-\mu}{\sigma}\right)-\Phi\left(\frac{x_0-\mu}{\sigma}\right)$$

$$P(X<x_1) = P\left(Z<\frac{x_1-\mu}{\sigma}\right)=\Phi\left(\frac{x_1-\mu}{\sigma}\right)$$

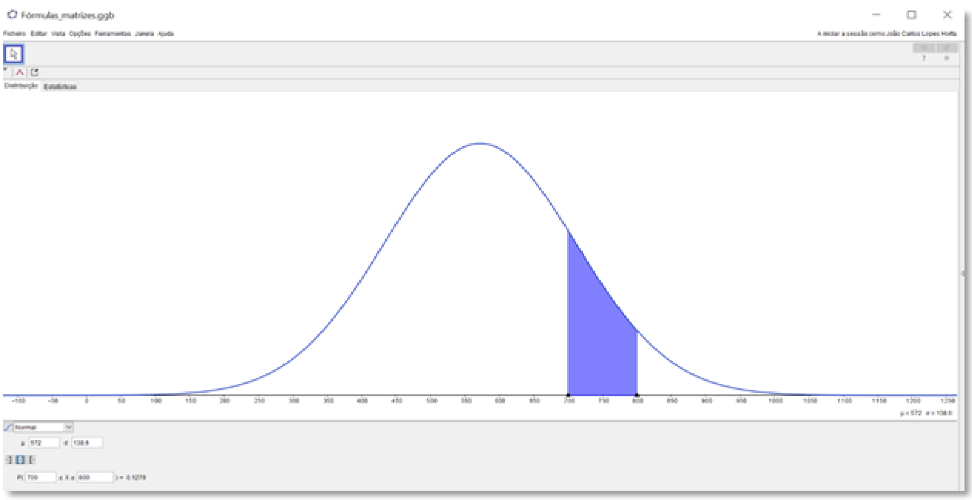
$$P(X>x_0) = P\left(Z>\frac{x_0-\mu}{\sigma}\right)=1-\Phi\left(\frac{x_0-\mu}{\sigma}\right)$$

Exemplo. A precipitação anual (em mm) num determinado lugar é bem modelada por uma distribuição normal com $\mu=572$ mm e $\sigma=138{,}6$ mm. Pretende-se calcular a probabilidade de a precipitação anual se situar entre 700 e 800 mm.

Seja X a variável aleatória que representa a situação em causa. Então, $X \sim N(572, 138.6^2)$. Como a normal não é estandardizada, deve fazer-se a mudança de variável para variável estandardizada, e depois determinar a probabilidade pedida. Assim,

$$\begin{aligned}
P(700<X<800) &= P\left(\frac{700-572}{138.6}<Z<\frac{800-572}{138.6}\right)\\
&= \Phi\left(\frac{800-572}{138.6}\right)-\Phi\left(\frac{700-572}{138.6}\right)\\
&= \Phi(1.645022)-\Phi(0.9235209)\\
&= 0.1279
\end{aligned}$$

Esta probabilidade pode ser obtida utilizando o software "Geogebra" (por exemplo):



Definição (Distribuição exponencial). Seja X uma variável aleatória com função densidade definida por

$$f_X(x|\lambda) = \begin{cases} 0 & se \quad x \leq 0 \\ \lambda e^{-\lambda x} & se \quad x > 0 \end{cases}, \quad \lambda > 0.$$

Neste caso, diz-se que X tem distribuição exponencial, ou exponencial negativa, e representa-se por $X \sim \mathrm{Ex}(\lambda)$.

Algumas características da distribuição exponencial – $X \sim \mathrm{Ex}(\lambda)$.

$$E(X) = 1/\lambda \, .$$

$$Var(X) = 1/\lambda^2 \, .$$

$$\gamma_1 = 2 \, .$$

$$CV = 1 \, .$$

Segundo (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 237), a distribuição exponencial, ou exponencial negativa, tem a sua origem associada ao processo de Poisson, embora a sua utilização na modelação estatística seja bastante mais ampla, aplicando-se, para além do tempo de espera entre acontecimentos originados por um processo de Poisson, a outros fenómenos.

Se a variável aleatória X representa a contagem do número de acontecimentos num intervalo unitário de tempo, em certas condições (condições do processo de Poisson) X tem distribuição de Poisson com parâmetro λ (o valor esperado de X – que pode ser interpretado com ritmo médio a que afluem os acontecimentos por intervalo unitário). Se as condições se mantêm e, o intervalo considerado, em vez de ser unitário, tem amplitude h, a vaiável aleatória Y, que representa a contagem no novo intervalo, tem distribuição de Poisson, com parâmetro λh.

Se a sucessão de acontecimentos constitui um processo de Poisson e a contagem é iniciada no momento 0 (zero), o tempo de espera pela chegada do primeiro acontecimento é uma variável aleatória T, cuja função distribuição F_T é dada por

$$F_T(t)=P(T \le t)=1-P(T>t) \quad , \quad \text{com } t>0 \, .$$

Como $P(T>t)$ é a probabilidade de não haver qualquer chegada no intervalo de $]0,t]$ tem-se:

$$P(T>t)=P(Z=0)=e^{-\lambda t} \, ,$$

onde $Z \sim \mathrm{Po}(\lambda t)$ é a variável aleatória que representa a contagem no intervalo $]0,t]$. Assim,

$$F_T(t)=\begin{cases} 0, & t<0 \\ 1-e^{-\lambda t}, & t \ge 0 \end{cases} \quad , \quad f_T(t)=\begin{cases} 0, & t \le 0 \\ \lambda e^{\lambda t}, & t>0 \end{cases} \quad ,$$

surgindo, desta forma, a distribuição exponencial.

Nota. O valor esperado de X de uma distribuição exponencial, mostra que, num processo de Poisson com parâmetro λ ($\lambda>0$), a média do tempo de espera por um acontecimento é o inverso da do ritmo de afluência dos acontecimentos. Por exemplo, se em média surgem 6 acontecimentos por hora, o tempo médio de espera por um acontecimento é um sexto de hora, ou seja, 10 minutos.

O coeficiente de variação e o de assimetria não dependem do parâmetro λ.

A distribuição exponencial não tem moda, uma vez que a função densidade é decrescente com x e o domínio é aberto. A mediana é dada por

$\mu_e=\ln(2)/\lambda$,

e, consequentemente, $\mu_e<\mu$.

Exemplo. A chegada de clientes a uma loja segue um processo de Poisson em que o ritmo médio de afluência é de 20 clientes por hora. Após abrir a loja, qual é a probabilidade de o comerciante ter de esperar mais que 5 minutos pela chegada do primeiro cliente?

Seja X a variável aleatória associada ao tempo de espera pelo primeiro cliente, em uma hora. Então, segundo os dados do problema, $X\sim Ex(20)$. Logo a probabilidade pedida obtém-se a partir de:

$$P(X>5/60)=P(X>1/12) = \int_{1/12}^{+\infty} 20e^{-20x}\,dx$$

$$= \left[-e^{-20x}\right]_{1/12}^{+\infty}$$

$$= e^{-5/3}\cong -0.19$$

Este cálculo pode ser realizado utilizando o software "Geogebra". Veja o Gráfico 4 - P(X>5/60) em Ex(20).

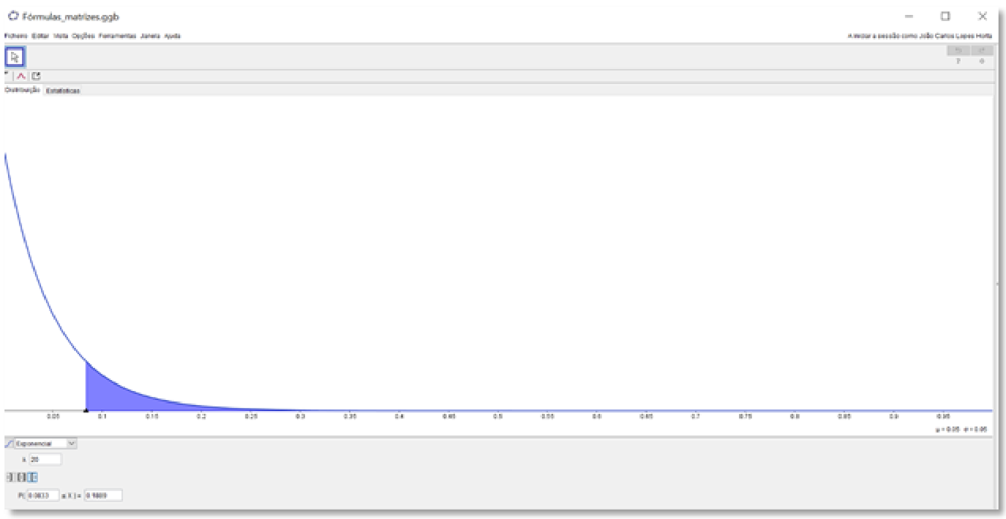


Gráfico 4 - P(X>5/60) em Ex(20)

Teorema. Seja $X\sim Ex(\lambda)$. Então:

$P(X>x+h\,|\,X>x)=P(X>h)$.

Nota. A propriedade enunciada no teorema anterior, designa-se pela "falta de memória" da distribuição exponencial. Com efeito, se qualquer objeto tem tempo de vida com distribuição exponencial, então, qualquer que seja a sua idade, o tempo residual de vida não é afetado pelo tempo já vivido, ou seja, o objeto "não envelhece"; e se chegou "vivo", por exemplo, ao fim de dez anos, o tempo de sobrevivência tem a mesma distribuição que o tempo de vida à nascença.

Definição (Função gama, ou fatorial generalizado). Seja α um número real positivo, ou seja, $\alpha>0$. A função gama, que se representa por $\Gamma(\alpha)$, define-se, por intermédio de um integral impróprio, da seguinte forma:

$\Gamma(\alpha)=\int_0^{+\infty} e^{-x} x^{\alpha-1}\,dx$.

Ou seja, trata-se de uma função real de variável real que faz corresponder a cada número real positivo, $\alpha>0$, o número real $\Gamma(\alpha)$.

Teorema (Propriedades da função gama). A função gama verifica as seguintes propriedades:

Quando $\alpha>1$, tem-se

$\Gamma(\alpha)=(\alpha-1)\Gamma(\alpha-1)$;

Se $\alpha=n$, $n\in N=\{1,2,3,\ldots\}$, então

$\Gamma(\alpha)=\Gamma(n)=(n-1)!$;

$\Gamma(1/2)=\sqrt{\pi}$.

Seja λ um número real positivo, ou seja, $\lambda>0$. Substituindo x por λx na expressão que define a função gama,

$\Gamma(\alpha)=\int_0^{+\infty} e^{-x} x^{\alpha-1}\,dx$,

e fazendo a respetiva mudança de variável, obtém-se

$\Gamma(\alpha)=\lambda^\alpha \int_0^{+\infty} e^{-\lambda x} x^{\alpha-1}\,dx$.

Definição (Distribuição gama). Seja X uma variável aleatória com função densidade definida por

$$f_X(x|\alpha,\lambda)=\begin{cases} 0 & se \quad x\leq 0 \\ \dfrac{\lambda^\alpha e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} & se \quad x>0 \end{cases} , \quad com\ \alpha>0\ e\ \lambda>0\,.$$

Nestas condições, diz-se que X tem distribuição gama de parâmetros α e λ, e representa-se por $X\sim G(\alpha,\lambda)$.

Nota. Segundo (Murteira, Ribeiro, Silva, & Pimenta, 2002, p. 240), a distribuição gama pode ser considerada uma generalização da distribuição exponencial e, tal com essa, também pode ser introduzida a partir do tempo de espera associado a um processo de Poisson.

Um caso particular da distribuição gama é a distribuição exponencial, que se verifica quando $\alpha=1$.

Algumas características da distribuição gama – $X\sim G(\alpha,\lambda)$.

E(X)	=	α/λ .
Var(X)	=	α/λ^2 .
CV	=	$1/\sqrt{\alpha}$
γ_1	=	$2/\sqrt{\alpha}$.
μ_*	=	$(\alpha-1)/\lambda$,quando $\alpha>1$.

Nota. Verifica-se, das caraterísticas indicadas da distribuição gama que o coeficiente de variação e o de assimetria dependem apenas de α, e decrescem com este parâmetro.

A distribuição gama não tem moda para $\alpha \leq 1$.

A mediana existe sempre, embora seja difícil obter a sua expressão analítica.

Sempre que existam μ_*, μ_e e μ, ou seja, moda, mediana e média, verifica-se $\mu_* < \mu_e < \mu$.

Os cálculos com distribuição gama fazem-se normalmente com recursos computacionais.

Exemplo. Admita-se que as indemnizações (em recursos) referentes a determinado risco, pagas por uma seguradora e representadas pela variável aleatória X, seguem uma distribuição gama de parâmetros $\alpha=2$ e $\lambda=0.008$. Determina-se a probabilidade de a seguradora pagar uma indemnização superior a 500 euros.

Calcula-se então a probabilidade de X ser superior a 500.

$$\begin{aligned} P(X>500) &= \int_{500}^{+\infty} \frac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)} \, dx \\ &= \int_{500}^{+\infty} 0.008^2 e^{-0.008x} x \, dx \\ &= 5/e^4 \\ &= 5e^{-4} \cong 0.091578194 \,. \end{aligned}$$

Os cálculos foram realizados com o auxílio do software "Geogebra":

Definição (Distribuição do qui-quadrado). Seja X uma variável aleatória, e seja n um número natural, ou seja, $n \in \mathbb{N}$. Diz-se que X tem distribuição do qui-quadrado com n graus de liberdade, e representa-se por $X \sim \chi^2(n)$, quando a respetiva função densidade é dada por

$$f(x|n) = \begin{cases} 0 & se \quad x \leq 0 \\ \dfrac{e^{-\frac{x}{2}} x^{-\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} & se \quad x > 0 \end{cases}, \quad n \in \mathbb{N}.$$

Nota. A distribuição do qui-quadrado, é um caso particular da distribuição gama, ou seja,

$X \sim \chi^2 \Leftrightarrow X \sim G(n/2, 1/2)$.

Os cálculos com a distribuição do qui-quadrado fazem-se, geralmente, com recursos computacionais, embora possa existir uma ou outra tabela para o efeito.

Entende-se por grau de liberdade o número de valores de um conjunto de dados que podem variar após terem sido impostas certas restrições a todos os valores; por exemplo: suponha que 10 estudantes obtiveram em um teste média de 8.0; assim a soma das 10 notas deve ser 80 (restrição); neste caso, temos um grau de liberdade 9=10-1, pois as nove primeiras notas podem ser escolhidas aleatoriamente, contudo a 10ª nota deve ser igual a

80-(a soma das 9 primeiras).

A distribuição qui-quadrado pode ser interpretado como a soma dos quadrados de normais padronizadas, ou seja:

$X\sim N(0,1)$, então $\sum_{i=1}^{n} X_i^2 = \chi^2(n)$.

A grande aplicação do qui-quadrado encontra-se nas inferências estatísticas (que está fora do escopo deste módulo).

Algumas características da distribuição gama – $X\sim G(\alpha,\lambda)$.

$E(X) = n$.

$Var(X) = 2n$.

$CV = \sqrt{2/n}$

$\gamma_1 = \sqrt{8/n}$.

$\mu_* = n-2$, quando $n>2$.

Algumas relações entre as distribuições: qui-quadrado, gama e normal

Seja X uma variável aleatória contínua.

Se $X\sim N(0,1)$, então $Y=X^2\sim\chi^2(1)$.

$X\sim G(\alpha,\lambda) \Leftrightarrow Y=cX\sim G(\alpha,\lambda/c)$, $c\neq 0$

$X\sim G(n,\lambda) \Leftrightarrow Y=2\lambda X\sim\chi^2(2n)$.

Exemplo. Suponha que a variável aleatória X segue uma distribuição qui-quadrado com 7 graus de liberdade.

Determine a $P(X>9)$.

Determine o valor de x tal que $P(X\leq x)=0.95$.

Determine o valor de x tal que $P(X>x)=0.95$.

Nota. Pode-se utilizar uma tabela do qui-quadrado para resolver as questões das alíneas (b) e (c). Ou seja, na tabela obtém-se os valores que correspondem aos pontos x tais que $P(\chi^2(n)\leq x)$.

Os cálculos podem ser sempre realizados utilizando os meios computacionais.

Definição (Função Beta). A função beta é uma função real de duas variáveis reais, que faz corresponder a cada par de números reais positivos, (α,β), um número real, $B(\alpha,\beta)$, definido por

$B(\alpha,\beta)=\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,du$.

Teorema (Propriedades da função beta). A função beta verifica as seguintes propriedades:

$B(\alpha,\beta)=B(\beta,\alpha)$.

$B(1,1)=1$.

As funções beta e gama estão relacionadas da seguinte forma

$B(\alpha,\beta)=\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$.

$B(1/2,1/2)=\pi$.

Definição (Distribuição Beta). Seja X uma variável aleatória, e sejam os números reais positivos α e β. Diz-se que X tem distribuição beta, e representa-se por $X\sim Be(\alpha,\beta)$, se X tem função densidade dada por

$$f_X(x|\alpha,\beta)=\begin{cases}\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1} & se \quad 0<x<1\\ 0 \quad se \quad outros\ casos\end{cases} \qquad para\ \alpha>0\ e\ \beta>0\,.$$

Algumas características da distribuição gama – $X\sim G(\alpha,\lambda)$.

E(X)	=	$\frac{\alpha}{\alpha+\beta}$.
Var(X)	=	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.
CV	=	$\sqrt{\frac{\beta}{\alpha(\alpha+\beta+1)}}$.
γ_1	=	$\frac{2(\beta-\alpha)\sqrt{\alpha+\beta+1}}{(\alpha+\beta+2)\sqrt{\alpha\beta}}$.

Nota. A distribuição beta é simétrica se $\alpha=\beta$; se $\alpha<\beta$, a distribuição é assimétrica positiva; se $\alpha>\beta$, a distribuição é assimétrica negativa.

Se $\alpha>1$ e $\beta>1$, a distribuição é unimodal e a moda é igual a

$\frac{\alpha-1}{\alpha+\beta-1}$.

Se $\alpha=\beta=1$, a distribuição é uma distribuição uniforme, e nos outros casos, a distribuição não tem moda.

Para realizar cálculos de probabilidades de distribuições beta faz-se uso de meios computacionais.

Exemplo. Considere-se uma apólice de seguro contra roubo com um capital de 4000 euros. Admita-se que a variável aleatória X traduz o montante de uma indemnização com proporção do capital seguro, e que $X\sim Be(2,1.5)$. Qual a probabilidade de uma indemnização ultrapassar 3600 euros?

A probabilidade pedida obtém-se a partir de:

$$P(X>3600/4000) = \int_{0.9}^{1} \frac{1}{B(2,1.5)}\, x(1-x)^{0.5}\, dx$$

$$= 47\cdot\sqrt{10}/2000$$

$$\cong 0.07431353$$

O cálculo do valor do integral foi realizado com o software "Geogebra":

Conclusão

As distribuições de variáveis aleatórias contínuas têm inúmeras aplicações na resolução de problemas de várias ordens.

Consoante o tipo de problema, pode-se ajustar a variável aleatória em estudo a uma das distribuições aqui estudadas (uniforme, normal, exponencial, gama, etc.).

Excetuando as distribuições uniforme e beta, as outras distribuições de variáveis aleatórias contínuas servem para modelar fenómenos, para os quais não existam limitações severas nos domínios dessas variáveis.

Os valores de valor esperado, variância e/ou desvio padrão, coeficiente de variação e coeficiente de assimetria permitem caracterizar a distribuição de uma variável aleatória – neste caso variável aleatória contínua – por isso, de acordo com os problemas propostos pode ser necessário a determinação dos seus valores.

Algumas distribuições são casos particulares de outras, mas por terem várias aplicações, os seus estudos fizeram-se em separados.

Avaliação

1. Seja X uma variável aleatória com função densidade definida, no seu suporte D_X=├]0,+∞┤[, por f_X (x│θ)=θ·e^(-θ·x) ,   para θ>0 .

   Determine o valor esperado de X.

2. A resistência à compressão de amostras de cimento de um certo tipo é uma variável aleatória que pode ser modelada por uma distribuição normal com média 600 kg/cm^2 e desvio padrão 100 kg/cm^2. Determine:

   a. A probabilidade para que uma amostra de cimento tenha resistência superior a 6150 kg/cm^2;

   b. A probabilidade de que uma amostra de cimento tenha resistência entre 5900 kg/cm^2 e 5950 kg/cm^2;

   c. A resistência R que é excedida por 90 % das amostras de cimento.

3. Uma empresa, monopolista no mercado de determinado produto, tem produção constante de 90 toneladas por mês, e tem conhecimento de que a procura é uma variável aleatória X~N(80,100).

   a. Determine a probabilidade de haver procura excedentária.

   b. Para satisfazer a eventual procura excedentária, a empresa tem possibilidade deimportar no princípio do mês um stock de segurança. Qual deve ser esse stock, S,por forma que a probabilidade de haver procura insatisfeita baixe para 0.025?

4. Suponha que o tempo de vida útil de uma componente eletrónica é uma variável aleatória X, com distribuição exponencial, com média igual a 600 horas.

   a. Determine a probabilidade de a componente durar mais 700 horas.

   b. Suponha que a componente já tinha durado 400 horas; calcule a probabilidade dedurar ainda mais 700 horas.

5. Seja X uma variável aleatória com função densidade definida, no seu suporte, por f_X (x)={■(x,&0<x<1@1/2,&1<x<2)┤ .

   a. Calcule a média e a variância de X.

   b. Determine o primeiro e o terceiro quartis.

   c. Utilizando as propriedades do valor esperado, obtenha a média e a variância davariável aleatória Y=4X-2.

   d. Calcule a média das seguintes variáveis aleatórias: Z=1/X;   U={■(-1,&X<0.5@1,&x≥0.5)┤ .

6. A duração de pequenos anúncios (entre 5 e 12 segundos) numa cadeia de televisão pode considerar-se uma variável aleatória com distribuição uniforme.

    a. Ndique a função densidade.

    b. Qual a probabilidade de um pequeno anúncio ter duração superior a 7 segundos?

    c. Calcule e interprete $P(X>6 \mid X \leq 10)$.

    d. Calcule a média e o desvio padrão da duração dos pequenos anúncios.

7. O tempo médio gasto na estafeta 4×100 metros por cada um dos atletas da equipa A é de, respetivamente, 10.6, 10.8, 10.5 e 10.7 segundos. Admita que os tempos gastos tem uma distribuição normal com desvio padrão de, respetivamente, 0.2, 0.5, 0.4 e 0.6 segundos.

    a. Estabeleça um limite máximo para o tempo gasto pela equipa em pelo menos 90 % dos casos.

    b. Se o tempo gasto pela equipa B tiver distribuição normal com média 42.8 e variância 0.8, qual a probabilidade de esta vencer a equipa A?

8. Os eixos produzidos por uma máquina consideram-se não defeituosos se o desvio do diâmetro do eixo para as dimensões projetadas não é maior, em valor absoluto, do que dois milímetros. Os desvios aleatórios do diâmetro dos eixos obedecem a uma distribuição normal de média nula e desvio padrão 1.6 mm. Determine a percentagem dos eixos não defeituosos produzidos.

9. Um banco atende em média dois clientes em cada 3 minutos. Considere que o número de clientes atendidos é um processo de Poisson.

    a. Qual a probabilidade de decorrerem 3 minutos sem qualquer cliente atendido?

    b. Qual a probabilidade de o atendimento de um cliente demorar mais do que 3 minutos?

    c. Compare o resultado das duas alíneas anteriores.

    d. Qual a probabilidade de atendimento de um cliente demorar entre 3 e 6 minutos?

Resumo da Unidade

As variáveis aleatórias classificam-se em discreta, contínua e mista.

As funções probabilidades (densidades) e as de distribuições são fundamentais na caracterização das variáveis aleatórias.

Na resolução de um problema há que verificar que tipo de variável aleatória está em estudos e que tipo de função de distribuição se adequa à ela.

Os usos de recursos computacionais revelam-se de grande importância – às vezes indispensáveis – na realização de cálculos relacionados com cálculos de probabilidades e outros.

Avaliação da Unidade

Verifique a sua compreensão!

Variável aleatória (unidimensional). Função de distribuição de uma variável

Instruções

Responda cada uma das questões de uma forma clara e justifique cada passo de resolução.

Critérios de Avaliação

Cada ponto ou alínea vale 10 pontos. Considera-se aprovado o estudante que tiver pelo menos 50% da cotação total.

Avaliação

1. Sejam k∈R e X uma variável aleatória com função distribuição

$$F_X(x) = \begin{cases} \frac{k}{2}\left(1 - e^{-x}\right) & se \quad x > 0 \\ 0 \quad se \quad x \leq 0 \end{cases}.$$

   a. Calcule o valor de k.

   b. Encontre a densidade de X para o valor de k calculado acima.

2. Uma caixa contém cinco bolas pretas (P), três azuis (A) e sete vermelhas (V). A experiência aleatória consiste em retirar ao acaso duas bolas sem reposição. Suponha que se atribuiu a seguinte pontuação: bola preta – 1 ponto; bola azul – 2 pontos; bola vermelha – 3 pontos. Considere a variável aleatória X, "soma dos pontos obtidos".

a. Qual acontecimento que é imagem inversa do intervalo [3,5⊣[.

b. Determine a função de distribuição.

c. Classifique a variável aleatória X.

d. Calcule P(X>3│X<6).

3. Seja

$$F(x) = \begin{cases} 0 & se \quad x < 1 \\ \frac{x}{3} & se \quad 1 \leq x < 2 \\ 1 & se \quad x \geq 2 \end{cases}.$$

a. Verifique se se trata de uma função de distribuição.

b. Em caso afirmativo, determine a variável aleatória em causa.

4. Seja a variável aleatória X com função de distribuição dada por:

$$F(x) = \begin{cases} 0 & se \quad x < 0 \\ 1 - \frac{2}{3}e^{-x} & se \quad x \geq 0 \end{cases}.$$

a. Determine P(X>0).

b. Represente graficamente a função, e classifique a variável aleatória em causa.

5. Seja

$$F(x) = \begin{cases} 0 & se \quad x < 0 \\ x^2 & se \quad 0 \leq x < \frac{1}{2} \\ -3x^2 + 6x - 2 & se \quad \frac{1}{2} \leq x < 1 \\ 1 & se \quad x \geq 1 \end{cases}.$$

a. Prove que F é a função de distribuição de uma variável aleatória X.

b. Determine a função densidade correspondente.

c. Determine a P(X<3/2).

d. Calcule o valor de k, tal que P(X<k)=1/2.

6. Seja Xuma variável aleatória, com função densidade definida, no suporte D_X=⊢]0,2⊣[, da seguinte forma:

$$f_X(x) = \begin{cases} x & se \quad 0 < x < 1 \\ \frac{1}{2} & se \quad 1 \leq x < 2 \end{cases}.$$

a. Determine a função de distribuição de X.

b. Determine a função densidade da variável aleatória Y=4X-2.

c. Determine a função de distribuição, e classifique a seguinte variável aleatória:

$$U = \begin{cases} -1 & se \quad X < 0.5 \\ 1 & se \quad X \geq 0.5 \end{cases} .$$

7. Um produtor de refrigerantes resolveu lançar uma campanha publicitária, oferecendo prémios impressos nas cápsulas das garrafas. Durante a campanha, 5% das garrafas distribuídas para a venda tinham prémio. Ao adquirir 15 garrafa, qual a probabilidade de se receber pelo menos um prémio?

8. Um lote de 500 peças existem 50 defeituosas. Para efeitos de aceitação do lote retira-se ao acaso uma amostra, rejeitando-se o lote se nessa amostra existirem mais que duas peças defeituosas.

   a. Se a amostra for de dimensão 10, qual a probabilidade de se rejeitar o lote?

   b. Supondo que se mantém o critério de rejeição de um lote se numa amostraexistirem mais que duas peças defeituosas, determine a dimensão máxima deamostra de forma que a probabilidade de rejeição do lote seja inferior a 0,05.

9. Sabe-se que 1% dos parafusos de determinado fabricante são defeituosos. Os parafusos são vendidos em caixas de 12 unidades com garantia de devolução do valor pago caso existam dois ou mais parafusos defeituosos.

   a. Qual a probabilidade de ocorrer uma devolução?

   b. Se se comprar dez caixas, qual a probabilidade de haver devoluções?

10. O número de erros de ortografia que um aluno dá por página, numa prova escrita de Estatística, segue um processo de Poisson com taxa média de 1.5 erros.

   a. Qual a percentagem de provas de duas páginas sem erros de ortografia?

   b. Se um aluno escreveu quatro páginas, qual a probabilidade de ter cometido maisde 8 erros?

   c. Escolhidas ao acaso cinco provas de quatro páginas cada, qual a probabilidade deapenas uma delas não ter erros de ortografia?

11. Seja X uma variável aleatória com função densidade definida, no seu suporte, por

$f_X(x)=\begin{cases} x+2, & -2<x<-1 \\ -x, & -1<x<0 \end{cases}$.

   a. Obtenha a média, a mediana e a variância de X.

   b. A variável aleatória X tem distribuição simétrica?

   c. Qual o coeficiente de assimetria de X?

   d. Sendo Z=2X-2, calcule a respetiva média e variância.

12. Considere que o tempo gasto numa visita à feira do livro é uma variável aleatória com distribuição normal, de média igual a duas horas. Suponha que apenas 25 % dos visitantes permanecem mais de 3 horas.

    a. Qual o desvio padrão da variável?

    b. Sabendo que um visitante já chegou à uma hora, qual a probabilidade de se irembora nos próximos 30 minutos?

    c. Calcule a mediana e o intervalo interquartil de X.

    d. Calcule a probabilidade de em 20 visitantes selecionados ao acaso haver no máximoum que permaneça mais de três horas.

13. O número de partículas radioativas observadas num intervalo de 10 minutos segue um processo de Poisson.

    a. Supondo que a probabilidade de não se observar qualquer partícula no intervalo de 10 minutos é de 0.15, determine o ritmo do processo, isto é, em média quantotempo medeia entre duas observações consecutivas?

    b. Qual a probabilidade de ser necessário esperar mais que 30 minutos até se observara terceira partícula?

Leituras e outros Recursos

As leituras e outros recursos desta unidade encontram-se na lista de Leituras e Outros Recursos do curso.

Leituras e outros recursos obrigatórios:

- Murteira, B. et al. (2002). Introdução à Estatística. (2ª Ed.). Lisboa: McGraw-Hill de Portugal.
- Farber, L. (2010). Estatística Aplicada. (4ª ed.). São Paulo: Pearson Prentice Hall.
- Software: Geogebra, R.

Leituras e outros recursos opcionais:

- Máquina Científica: TI-83 ou semelhante.

Unidade 4. Aplicação do Estudo de Estatística e Probabilidade à CIA

Introdução à Unidade

Esta unidade destina-se a resolução de exercícios no âmbito de Estatística e Probabilidade, com recurso a softwares adequados.

Inicia-se com a apresentação e manuseio dos softwares que serão utilizados ao longo da unidade, nomeadamente "Excel", "R" e "Geogebra".

Seguidamente, apresenta-se problemas de Estatística Descritiva e de Cálculo de Probabilidades para serem analisados e resolvidos.

Objetivos da Unidade

Após a conclusão desta unidade, deverá ser capaz de:

- Operar com os seguintes softwares: "Excel", "R" e "Geogebra";
- Resolver problemas de Estatísticas Descritivas e de Probabilidade e Estatística com recurso a softwares

Termos-chave

Problema matemático: é um tipo de exercício matemático pelo qual não existe um algoritmo para a sua resolução.

Excel: é uma planilha (folha de cálculo) desenvolvida pela Microsoft, que realiza inúmeros cálculos e construções de gráficos em diversas áreas, nomeadamente na área de estatística e probabilidades.

R: é um software livre específico para cálculos e construções de gráficos na área de estatística e probabilidade.

Geogebra: é um software livre com aplicações (cálculos e construções de gráficos) em diversas áreas de Matemática, como Geometria, Estatística e probabilidades, Cálculo, etc.

Atividades de Aprendizagem

Atividade 1 – Software utilizadas na resolução de problemas estatísticos

Introdução

Nesta atividade vai apresentar-se os softwares "Excel", "R" e "Geogebra" no que tange as suas utilizações ao longo desta unidade.

Embora o Excel seja uma ferramenta muito forte na área de estatística e probabilidade, neste trabalho o Excel é utilizado mais como um meio para introdução de dados, a serem utilizados pelo "R", uma vez que este último é completamente virado para cálculos e construções de gráficos na área referida acima.

O "Geogebra" será utilizada uma ou outra vez, para calcular algumas probabilidades e para construir alguns gráficos.

Detalhes da atividade

Microsoft Excel. É um editor de planilhas produzido pela Microsoft, com boas ferramentas para efetuar cálculos e para construção de gráficos.

Nesta unidade vai utilizar-se o Excel 2016. Contudo, pode usar-se o Excel 2007, 2010, 2013 ou qualquer outra versão. Para o fim desejado neste curso, não há muita diferença entre essas versões, ou seja, qualquer uma delas serve de suporte para resolução de problemas em Estatística e Probabilidades.

Com o "Excel" podemos organizar os dados de um determinado problema e, a partir daí, calcular as medidas de centralização e de dispersão em relação a esses dados. Pode-se representar um conjunto de dados graficamente—através de gráfico de barra, histograma, gráfico de setores, etc. Também, pode-se fazer operações que envolvem cálculos de probabilidades com variáveis aleatórias discretas e contínuas.

A figura que se segue mostra a Interface do "Excel 2016", quando se escolhe o "livro em branco" —o que se deve escolher quando se inicia o tratamento de um conjunto de dados— na abertura desta aplicação.

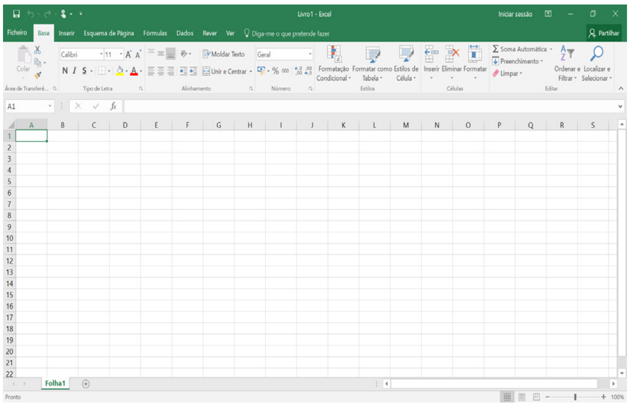


Figura 1 - Interface do Excel 2016

Para mais informações sobre a utilização dessa poderosa ferramenta de apoio em Estatística e Probabilidade, aconselha-se a consulta de alguns tutoriais sobre o Microsoft Excel, como por exemplo: doc. 1 ou doc. 2.

Geogebra. É um software gratuito e multiplataforma de matemática dinâmica, desenvolvido para o ensino e aprendizagem da matemática nos vários níveis de ensino. Possui ferramentas de cálculo e construção de gráfico na área de geometria, álgebra, cálculo, probabilidade e estatística, etc.

Nesta unidade utiliza-se o software Geogebra para auxiliar na construção de gráficos, no cálculo de probabilidades e na determinação de medidas de tendência central e de dispersão.

A figura que se segue mostra a interface do Geogebra, ao iniciá-lo:

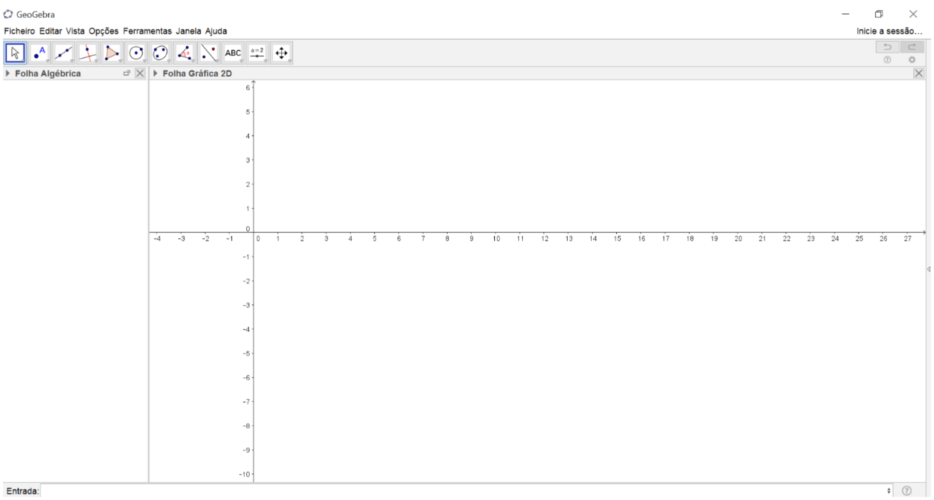
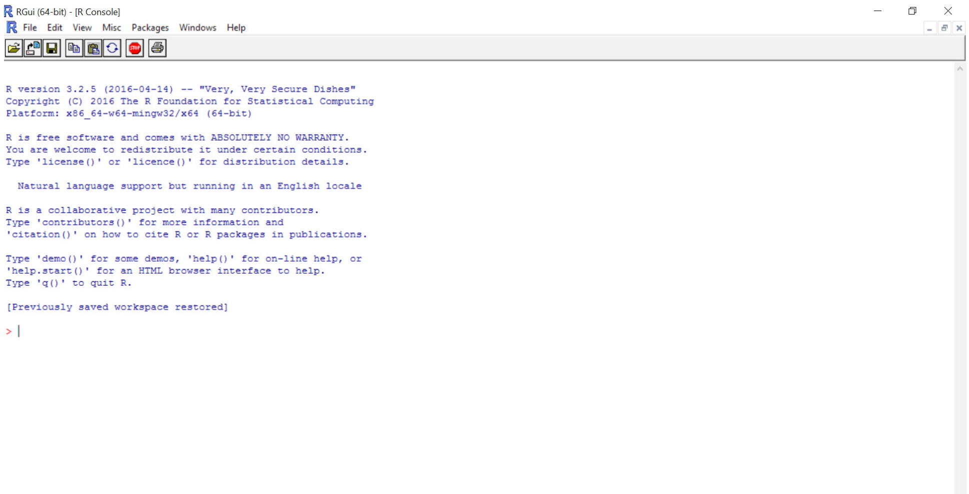


Figura 2 Interface do Geogebra 5.0.253.0-3D

Para mais informação sobre Geogebra pode, por exemplo, consultar os seguintes documentos: doc. 1, doc. 2.

R. É uma linguagem de programação e um ambiente de desenvolvimento integrado, gratuito, para análise estatística e construção de gráficos.

A figura que se segue mostra a interface de R ao iniciá-lo:



Dos três softwares, R é mais indicado para estatística e probabilidade, por isso aconselha-se especial atenção a esse software.

A medida que iremos prosseguir no desenvolvimento desta unidade, conforme for o tipo de problema a ser resolvido numa determinada ocasião, introduzir-se-á as ferramentas de "Excel", "Geogebra" e "R" para o efeito.

Exemplo. Considera-se os dados referentes à taxa anual de nupcialidade, TAN—já apresentados "num dos exemplos, na Unidade 1":

Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN	Ano	TAN
1960	7.8	1969	9	1978	8.3	1987	7.2	1996	6.3
1961	8.8	1970	9.4	1979	8.1	1988	7.1	1997	6.5
1962	7.9	1971	9.7	1980	7.3	1989	7.4	1998	6.6
1963	7.9	1972	9.1	1981	7.7	1990	7.3	1999	6.8
1964	8.1	1973	9.9	1982	7.4	1991	7.3	2000	6.2
1965	8.4	1974	9.3	1983	7.5	1992	7	2001	5.7
1966	8.6	1975	10.9	1984	7	1993	6.8	2002	5.4
1967	8.9	1976	10.5	1985	6.8	1994	6.6		
1968	8.7	1977	9.4	1986	6.9	1995	6.6		

Tabela 1 Taxa Anual de Nupcialidade

Leituras e outros Recursos

As leituras e outros recursos desta unidade encontram-se na lista de Leituras e Outros Recursos do curso.

Avaliação do Curso

Avaliação sumativa de Introdução à Estatística e Probabilidade

Instruções

O Teste de avaliação tem sete questões, algumas com alíneas.

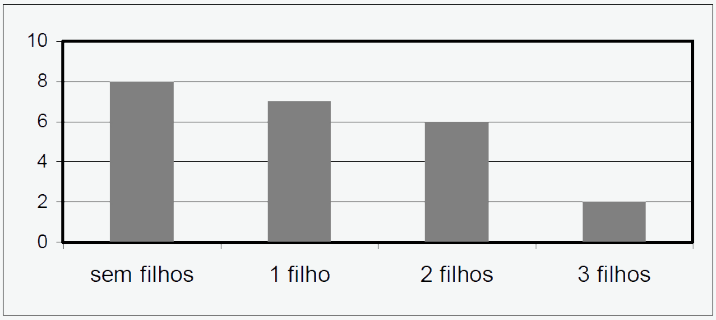
Responda cada uma das questões de uma forma clara e justifique cada passo de resolução.

Critérios de avaliação

Cada ponto ou alínea vale 10 pontos. Considera-se aprovado o estudante que tiver pelo menos 50% da cotação total.

Avaliação

1. As ex-alunas de uma turma que completou o Ensino Médio há anos se encontraram em uma reunião comemorativa. Várias delas haviam se casado e tido filhos. A distribuição das mulheres, de acordo com a quantidade de filhos, é mostrada no seguinte gráfico:



   a. Um prémio foi sorteado entre todos os filhos dessas ex-alunas. Qual é a probabilidade de que a criança premiada tenha sido um(a) filho(a) único(a)?

2. Considere a amostra que mostra o número de dias de férias de funcionários de uma determinada empresa, num determinado ano:

3	9	2	1	7	5	3	2	2	6
4	0	10	0	3	5	7	8	6	5

   a. Determine a mediana e os quartis e .

   b. Esboce a respetiva caixa-de-bigodes.

   c. Aproximadamente dos funcionários não gozaram mais do que quantos dias deférias?

3. As suas compras mensais com cartão de créditos (arredondadas para o dólar mais próximo) nos últimos dois anos e, as do seu amigo estão listadas a seguir:

Você:

60	95	102	110	130	130	162	200	215	120	124	28
58	40	102	105	141	160	130	210	145	90	46	76

Seu Amigo:

100	125	132	90	85	75	140	160	180	190	160	105
145	150	151	82	78	115	170	158	140	130	165	125

a. Determine a média e a moda de cada um destes conjuntos de dados.

b. Determine o desvio padrão de cada um dos conjuntos de dados.

c. Diga em qual dos dois conjuntos de dados, as observações estão menos dispersasem relação à média aritmética.

4. Considere a distribuição de frequências, que mostra o número de eleitores de um determinado país , de acordo com a idade:

Idade dos eleitores (em anos)	18 a 20	21 a 24	25 a 34	35 a 44	45 a 64	Acima dos 65
F_i (em milhões)	5.8	8.5	21.7	27.7	51.7	26.7

Encontre a probabilidade de um eleitor escolhido, aleatoriamente,

a. Esteja entre e anos.

b. Não esteja entre e anos.

5. Uma moeda é viciada de tal modo que sair cara é duas vezes mais provável do que sair coroa. Calcule a probabilidade de:

a. Ocorrer cara no lançamento desta moeda.

b. Ocorrer coroa no lançamento desta moeda.

6. Um grupo é constituído de homens e mulheres. Três pessoas são selecionadas ao acaso, sem reposição; qual a probabilidade de que ao menos duas sejam homens?

7. Uma urna tem bolas vermelhas e pretas. Outra urna tem bolas vermelhas e pretas. Uma urna é escolhida ao acaso e dela é escolhida uma bola também ao acaso. Qual probabilidade de observar-se:

a. Urna e bola vermelha?

b. Urna e bola preta?

c. Urna e bola vermelha?

d. Urna e bola preta?

8. Considere a função:

$$f(x) = \begin{cases} \dfrac{x^2}{14} & se \quad x = 1, 2, 3 \\ 0 \quad se & noutros\ casos \end{cases} .$$

a. Mostre que esta função é uma função de probabilidade.

b. Seja uma variável aleatória com função probabilidade . Qual é a função de distribuição de .

c. Seja uma variável aleatória com função de probabilidade . Determine .

9. Considere-se a variável aleatória que representa o número de cartas de copas em quatro retiradas aleatoriamente sem reposição de um baralho vulgar de cartas.

a. Encontre a função de probabilidade da varável aleatória .

b. Calcule a média e o desvio padrão de .

c. Qual a probabilidade de serem todas de copa?

d. Sabendo que saiu pelo menos uma carta de copas, nas quatro escolhidas ao acaso,qual a probabilidade de serem exatamente duas cartas de copas?

10. Uma região está dividida em distritos de superfícies semelhantes, pelo que se admite que cada distrito tem igual probabilidade de ser bombardeado. Calcula-se que o número de bombas recebidas por esta região foi de . Se um distrito for selecionado ao acaso, qual é a probabilidade de ser bombardeado com duas bombas?

Avaliação sumativa 1

Instruções

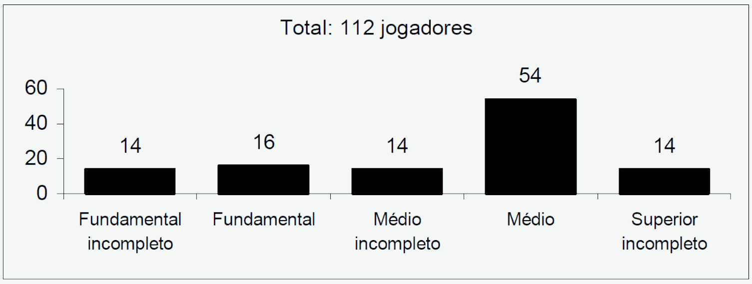
Responda cada uma das questões de uma forma clara e justifique cada passo de resolução.

Critérios de avaliação

Cada ponto ou alínea vale 10 pontos. Considera-se aprovado o estudante que tiver pelo menos 50% da cotação total.

Avaliação

1. A escolaridade dos jogadores de futebol nos grandes centros é maior do que se imagina, como mostra a pesquisa abaixo, realizada com os jogadores profissionais dos quatro principais clubes de futebol do Rio de Janeiro.



(*O Globo*, 24/7/2005.)

   a. Determine a percentagem dos jogadores dos quatro clubes que concluíram oEnsino Médio.

2. Considere a seguinte amostra, que mostra o número de horas que pessoas assistem televisão diariamente:

2	4	1	5	7	2	5	4	4	2	3	6	4	3
5	2	0	3	5	9	4	5	2	1	3	6	7	2

   a. Determine a mediana e os quartis e .

   b. Esboce a respetiva caixa-de-bigodes.

   c. Qual é a percentagem de pessoas que assistem mais do que horas de televisão pordia?

   d. Se selecionarmos uma pessoa aleatoriamente a partir da amostra, qual será aprobabilidade dessa pessoa assistir menos do que horas de televisão por dia?

3. Em uma amostra aleatória de aviões, lista-se o número de defeitos encontrados em suas fuselagens, conforme a seguinte tabela:

N. º de defeitos	0	1	2	3	4	5	6
N. º de aviões	4	5	2	9	1	3	1

   a. Determine a média e a moda.

   b. Determine o desvio padrão e o coeficiente de variação.

   c. Qual é a percentagem de defeitos que estão a menos desvios padrões da média?

4. Um estudante joga uma moeda e seleciona aleatoriamente um número de a . Qual é a probabilidade de obter coroa e número ?

5. Um dado é viciado, de modo que a probabilidade de se observar um número na face de cima é proporcional a esse número. Calcule a probabilidade de:

   a. Ocorrer número par.

   b. Ocorrer número maior ou igual a .

6. Uma urna contém bolas vermelhas e brancas. Duas bolas são extraídas ao acaso, com reposição, qual a probabilidade de:

   a. Ambas serem vermelhas?

   b. Ambas serem ser preta?

7. Em uma população, o número de homens é igual ao de mulheres. dos homens são daltônicos e das mulheres são daltônicas. Uma pessoa é selecionada ao acaso e verifica-se que é daltônica. Qual é a probabilidade de que ela seja mulher?

8. Seja uma variável aleatória com função densidade definida, no seu suporte , por:

   a. Mostre que 0.5.

   b. Sabendo que é positivo, e utilizando a função de distribuição, calcule a probabilidade de ser maior que .

   c. Obtenha a distribuição da variável aleatória que concentra os valores negativos deno ponto zero e mantém, sem alteração, os positivos. Classifique-a.

   d. Se , verifique que a respetiva função densidade é , para .

9. Sabe-se, por experiência, que a probabilidade de uma máquina necessitar de ser afinada em cada período de trabalho de minutos é de (períodos independentes). Determine:

    a. A média do número de afinações numa semana em que a máquina trabalhe horas?

    b. A probabilidade de em horas de trabalho se verificar pelo menos uma afinação, e ade se verificarem a afinações.

10. Determine a média e o desvio padrão para o número de raparigas em nascimentos.

    a. Diga se é usual ou não, em nascimentos, serem raparigas?

Avaliação sumativa 2

Instruções

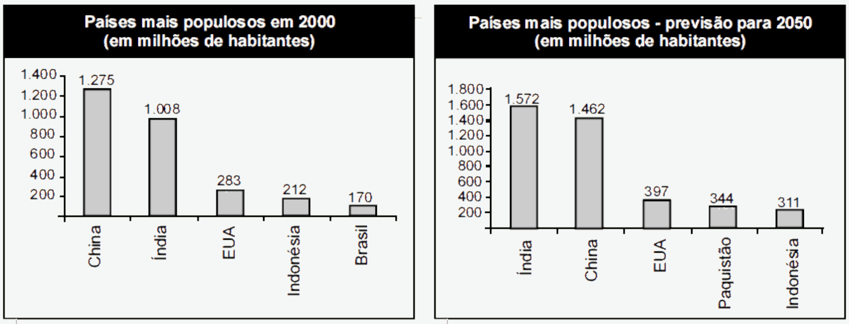
Responda cada uma das questões de uma forma clara e justifique cada passo de resolução.

Critérios de avaliação

Cada ponto ou alínea vale 10 pontos. Considera-se aprovado o estudante que tiver pelo menos 50% da cotação total.

Avaliação

1. 1.Nos últimos anos, ocorreu redução gradativa da taxa de crescimento populacional em quase todos os continentes. A seguir, são apresentados dados relativos aos países mais populosos em e também as projeções para .



Internet: <www.ibge.gov.br>.

Com base nas informações acima, é correto afirmar que, no período de a ,

a. A taxa de crescimento populacional da China será negativa.

b. A população do Brasil duplicará.

c. A taxa de crescimento da população da Indonésia será menor que a dos EUA.

d. A população de Paquistão crescerá mais de .

e. A China será o país com maior taxa de crescimento populacional do mundo.

2. Considere a seguinte amostra, que mostra as distâncias de um determinado aeroporto, depois de viagens:

2.8	2	3	3	3.2	5.9	3.5	3.6	4	4.1	4.6
1.8	5.5	3.7	5.2	3.8	3.9	6	2.5	5	5.5	6

a. Determine a mediana e os quartis e .

b. Determine a média e a moda.

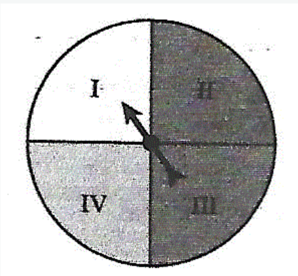
c. Determine o desvio padrão.

d. Esboce a respetiva caixa-de-bigodes.

e. Aproximadamente de distâncias não ultrapassaram quantas milhas?

3. Um site de internet compara as tacadas por rodadas de dois jogadores de golfe profissionais. Qual jogador é mais consistente, sabendo que a média e o desvio padrão do jogador são, respetivamente e tacadas, e os do jogador são, respetivamente, e tacadas?

4. Uma experiência de probabilidade consiste em rolar um dado de seis lados e girar a roleta, abaixo apresentada. A roleta tem a mesma probabilidade de parar em cada um dos quadrantes . Qual é a probabilidade de rolar um número maior que e a roleta parar no quadrante ?



5. Um número é escolhido ao acaso entre os inteiros de a . Qual é a probabilidade do número:

    a. Ser múltiplo de ?

    b. Ser múltiplo de e de ?

    c. Ser múltiplo de ou de ?

6. Um grupo é constituído de pessoas, entre elas Jonas e Cesar. O grupo é disposto ao acaso em uma fila. Qual e probabilidade de que haja exatamente pessoas entre Jonas e Cesar?

7. A probabilidade de um certo homem sobreviver mais anos, a partir de uma certa data é , e de que sua esposa sobreviva mais anos a partir da mesma data é . Qual a probabilidade de:

    a. Ambos sobreviverem mais anos a partir daquela data?

    b. Ao menos um deles sobreviver mais anos, a partir daquela data?

8. Considere a variável aleatória com função densidade definida, no seu suporte , por:

$$f_X(x) = \begin{cases} 0.5e^x & se \quad x < 0 \\ \dfrac{x}{4} & se \quad 0 < x < 2 \end{cases} .$$

    a. Determine a função distribuição de .

    b. Calcule .

9. Um jovem casal deseja ter, exatamente, duas filhas. Considere que a probabilidade de ser rapaz ou rapariga é igual, e que o casal terá filhos (rapazes ou raparigas) até realizar o desejo.

    a. Qual a função probabilidade do número de rapazes?

    b. Qual a probabilidade de ter quatro filhos?

    c. Qual a probabilidade de ter no máximo quatro cianças?

Referências do Curso

Fernandes, S., & Pinto, M. (s.d.). Afinal o que são e como se calculam os quartis? Universidade de Algarve - Departamento de Matemática, Departamento de Matemática, Faro.

Farber, L. (2010). Estatística Aplicada. (4ª ed.). São Paulo: Pearson Prentice Hall.

Software: Geogebra, R.

Larson, R., & Farber, B. (2010). Estatística Aplicada (4ª ed.). São Paulo: Pearson Prentice Hall.

Murteira, B., Ribeiro, C. S., Silva, J. A., & Pimenta, C. (2002). Introdução à Estatística (2ª ed.). Lisboa: McGraw-Hill de Portugal.

Sede da Universidade Virtual africana

The African Virtual University
Headquarters

Cape Office Park

Ring Road Kilimani

PO Box 25405-00603

Nairobi, Kenya

Tel: +254 20 25283333

contact@avu.org

oer@avu.org

Escritório Regional da Universidade Virtual Africana em Dakar

Université Virtuelle Africaine

Bureau Régional de l'Afrique de l'Ouest

Sicap Liberté VI Extension

Villa No.8 VDN

B.P. 50609 Dakar, Sénégal

Tel: +221 338670324

bureauregional@avu.org