UNIVERSIDADE DA BEIRA INTERIOR
Ciências

Simulações e Visualizações no Ensino das Probabilidades e Estatística:
Uma Aplicação Shiny

Nadab Ismael Cumbi Jorge

Dissertação para obtenção do Grau de Mestre em
Matemática Para Professores
(2^0 ciclo de estudos)

Orientadora: Professora Doutora Ana Paula André Martins Fernandes
Covilhã, Junho de 2018

Dedicatória

A meu Pai

Agradecimentos

A Deus, pela vida e pela saúde.

À Professora Doutora Ana Paula André Martins Fernandes, por todo conhecimento que me transmitiu na realização deste trabalho e pelos conselhos e palavras motivadoras.

A todos os meus professores da UBI, pelo conhecimento transmitido.

A toda a minha família, especialmente às minhas Mães, e aos meus irmãos e sobrinhos.

Aos colegas e amigos.

Por fim, a todos aqueles que direta ou indiretamente estiveram do meu lado e contribuíram para que eu pudesse chegar até aqui.

Resumo

Nos dias de hoje vivemos num mundo visual no qual a representação do conhecimento e a sua disseminação são quase impossíveis de se realizar sem recorrer a imagens. Com o avanço da tecnologia, a utilização de visualizações para promover a aprendizagem pode ser aprimorada através de simulações com as quais se podem criar representações simples e mais adequadas para ilustrar ideias complexas e abstratas contidas nos textos. De um modo particular, a realização de simulações em Probabilidades e Estatística é fundamental pois facilita a compreensão de conceitos de aleatoriedade e variação de amostras, permite observar em alguns minutos a evolução temporal de um fenómeno que levaria horas, dias ou anos em tempo real, além de permitir a repetição de experiências aleatórias nas mesmas condições e tantas vezes quantas se queira. Este facto incentivou-nos a desenvolver no presente trabalho, com o auxílio do *software* `R`, uma aplicação `Shiny` para simular e ilustrar alguns conceitos e resultados de probabilidades e estatística a um nível introdutório.

O `Shiny` é um pacote do `RStudio` que é uma interface do `R` que permite a criação de ferramentas, nomeadamente aplicações interativas, para tratamento de dados de forma flexível e acessível para usuários do `R`. A gratuidade do *software* `R`, e o aspeto apelativo e fácil utilização das aplicações `Shiny`, são uma vantagem para a utilização da aplicação que desenvolvemos por professores e alunos no processo de ensino e aprendizagem de probabilidades e estatística.

Palavras-chave

Simulações, visualizações, probabilidades e estatística, *software* `R`, aplicação `Shiny`.

Abstract

In our days, we live in a visual world in which the representation of knowledge and its dissemination are almost impossible to perform without resorting to images. With the development of technology, the use of visualizations to promote learning can be improved through simulations with which simple and more adequate representations can be created to to illustrate abstract and complex ideas contained in texts. In particular, simulations are fundamental in Probability and Statistics because they facilitate the undertanding of the concepts of randomness and variation of samples, allow the observation, in few minutes, of the temporal evolution of a phenomenon that would take hours, days or years in real time. Besides this, they allow the repetition of random experiences in the same conditions and as many times as we want. This fact encouraged us to develop in the present work, with the support of software `R`, a `Shiny` application to simulate and illustrate some concepts and results of probabilities and statistics at an introductory level.

`Shiny` is a package of `RStudio`, a `R` interface, that allows the development of tools, namely interactive applications, for data processing in a flexible and accessible way for the users of `R`. Being software `R` free of charge and the appealing aspect and easy use of `Shiny` applications, make the application that we developed advantageous for teachers and students to use in the teaching and learning process of probabilities and statistics.

Keywords

Simulations, visualizations, probability and statistics, *software* `R`, `Shiny` application.

Índice

Lista de Figuras

Lista de Tabelas

Lista de Acrónimos

v.a.r.	variável aleatória real
i.i.d.	independentes e identicamente distribuídas
TLC	Teorema Limite Central

Introdução

Nos dias de hoje vivemos num mundo visual no qual a representação do conhecimento e a sua disseminação são quase impossíveis de se realizar sem recorrer a imagens. Quando bem elaborada, uma imagem pode transmitir a informação pretendida de forma rápida, objetiva e direta, fazendo com que nos lembremos dela por um longo período.

No campo educativo, segundo Mayer (2003), ao associarem-se imagens às palavras torna-se mais fácil aprofundar o conhecimento dos estudantes, uma vez que a aquisição de conhecimentos se torna mais efetiva se este for transmitido a partir de imagens e palavras do que apenas de palavras isoladas. Se o professor se focar num sistema de verbalização entre ele e os alunos, utilizando apenas palavras, estes terão maior dificuldade em recordar o que foi dito pelo professor pouco tempo após a informação ter sido transmitida. Tendo em consideração que a informação se processa através dos canais verbal e visual, ao interligar as palavras a imagens proporciona-se ao estudante uma aprendizagem complementar, isto porque uma não substitui nem sobrepõe a outra. Pelo contrário, são duas formas de linguagem que coincidem e se diferenciam, sendo que uma pode ser intérprete da outra, tornando-as dependentes entre si.

Considerando as vantagens decorrentes da utilização simultânea de imagens e palavras, Magalhães (2005) refere que enquanto sistema de representação da palavra, a imagem implica um exercício estruturado de capacidades de codificação-descodificação, envolve diversos atos de compreensão oral e escrita, apela às capacidades sensoriais e cognitivas que estimulam a sensibilidade e promovem a aprendizagem do aluno, facilita a memorização, ajuda o aluno a pensar, a imaginar, a ativar a polissemia conotativa de conceitos, entre outros aspetos.

Apesar das suas potencialidades, esta metodologia não é por si só capaz de colmatar as dificuldades presentes na aquisição eficiente de conhecimentos por parte dos estudantes. Com o avanço da tecnologia, os meios digitais vêm desempenhando gradualmente o papel de ferramentas primordiais para o cumprimento desta tarefa. Recorrendo-se ao computador, esta metodologia pode ser aprimorada através de simulações com as quais se podem criar representações novas e mais adequadas para ilustrar ideias complexas contidas nos textos, que com o papel e o lápis seriam difíceis de concretizar.

As simulações permitem reproduzir virtualmente uma situação real, tornando possível experimentar os efeitos de um determinado procedimento sem a sua ocorrência. Elas podem ser vistas como representações ou modelações de objetos específicos, reais ou imaginários, de sistemas ou fenómenos. São particularmente úteis, quando a experiência original for impossível de ser reproduzida. Assim, este recurso permite ao estudante observar o comportamento de um determinado fenómeno através de um modelo do mesmo.

De acordo com Guillermo *et al.* (2005) , as simulações podem trazer muitos benefícios na aprendizagem, tais como: garantem uma aprendizagem mais profunda onde os estudantes simulam um problema complexo, desenvolvendo estratégias e integrando habilidades; proporcionam um ambiente de baixo risco onde o estudante adquire experiência, com situações difíceis sem consequências caras ou irreversíveis, de modelos de alto custo; os estudantes são emocionalmente mais envolvidos quando mergulham em uma experiência real do mundo que o cerca; fornecem uma experiência de aprendizagem que produz efeitos, porque servem como uma ponte do ambiente para a realidade.

De um modo geral, pode-se dizer que as simulações potenciam em grande medida a utilização de visualizações no ensino e aprendizagem. Esta visão é defendida por Hekcler *et al.* (2007) que destacam que as simulações podem ser consideradas como a solução de inúmeros problemas que os professores enfrentam ao tentar explicar aos seus alunos fenómenos demasiado abstratos para serem compreendidos exclusivamente através de uma descrição em palavras e demasiado complicados para serem representados através de uma única figura. Elas possibilitam também observar em alguns minutos a evolução temporal de um fenómeno que levaria horas, dias ou anos em tempo real, além de permitirem ao estudante repetir a observação sempre que o desejar. Por este motivo, existe uma crescente necessidade de se adotar esta ferramenta no ensino de probabilidades e estatística.

Na visão de Gordon e Gordon (2009) os conceitos de aleatoriedade e a variação entre as amostras podem ser melhor compreendidos com a realização de simulações de experiências aleatórias. Segundo Fernandes *et al.* (2007) , a simulação também permite experimentar novos fenómenos variando as condições e observando os resultados, proporcionando ao estudante uma experiência probabilística difícil de adquirir na vida real. Estes autores enumeram os seguintes aspetos positivos que a simulação apresenta para o ensino das probabilidades:

1. do ponto de vista representacional, os alunos podem pensar e formular modelos concretos em vez de trabalhar com modelos teóricos;

2. o cálculo vem muito facilitado ao passar do cálculo teórico de probabilidades ou do cálculo combinatório ao cálculo de frequências para estimar uma probabilidade;

3. a aprendizagem da modelação probabilística, ao implicar conceber uma experiência e pensar num modelo antes de efetuar cálculos sem sentido e intencionalidade.

Não podemos, no entanto, ver as simulações e visualizações como uma solução de todos os problemas de aprendizagem dos conceitos probabilísticos e estatísticos, por parte dos estudantes. Gordon e Gordon (2009) consideram que é essencial que, quando utilizadas na sala de aula, sejam apropriadas e pensadas para os estudantes a que se dirigem, devendo-se garantir que as mensagens transmitidas por estes métodos estejam de acordo com o processo cognitivo de cada um deles. Exige-se assim uma especial atenção em relação ao perigo que acarreta o fato dos alunos explorarem demasiado rápido os problemas, não lhes sendo dada a oportunidade para pensar cuidadosamente no trabalho realizado e passar com confiança de um nível de pensamento para outro. Quando cumpridos esses pressupostos, os benéficos da utilização destes métodos na sala de aula são de certo inquestionáveis.

Pensando nesses benefícios e aproveitando as potencialidades do *software* `R` e `RStudio`, mais concretamente do pacote `Shiny`, propomo-nos neste trabalho desenvolver uma aplicação de utilização fácil por professores e alunos para o ensino e aprendizagem de alguns conceitos e resultados de probabilidades e estatística. Para tal, dividimos este trabalho em dois capítulos que vamos passar a descrever em detalhe.

O primeiro capítulo é dedicado ao `RStudio` e ao pacote `Shiny`. Aqui fazemos uma breve introdução destas ferramentas, aproveitando para descrever as suas principais características, funções e potencialidades.

No segundo capítulo apresentamos detalhadamente a estrutura e funcionamento da aplicação `Shiny` desenvolvida. Aqui propomo-nos explorar as potencialidades de utilização da aplicação na sala de aula, para além de expor os conceitos e resultados teóricos que justificam as simulações apresentadas.

Destacamos o trabalho de Gordon e Gordon (2009) como impulsionador para o desenvolvimento da nossa aplicação, que pode ser disponibilizada pelo autor deste trabalho (nadab.jorge@outlook.pt), bem como todo o código `R` que lhe deu origem. Salientamos que o aspeto da aplicação pode variar com o *browser* utilizado.

Capítulo 1

Breve introdução ao **RStudio** e ao pacote **Shiny**

1.1 Introdução

O *software* R foi criado em 1993 por Ross Ihaka e Robert Gentleman (departamento de Estatística da Universidade de Auckland, Nova Zelândia) e tem sido continuamente desenvolvido através de contribuições vindas de todo o mundo. É uma linguagem de programação, similar à linguagem S, e um ambiente para computação estatística, que fornece uma variedade de técnicas estatísticas (modelação linear e não linear, testes estatísticos clássicos, análise de séries temporais, etc.) e gráficas.

Apesar do R ser muitas vezes visto como um programa estatístico, ele é um ambiente em que técnicas estatísticas são implementadas, podendo ser estendido facilmente através de pacotes. Existem oito pacotes que são fornecidos inicialmente com o R, mas existem muitos mais disponíveis através da família CRAN (*Comprehensive R Archive Network*) de sítios da internet, cobrindo uma vasta área da estatística moderna e do cálculo probabilístico.

Duas das grandes vantagens do R são a facilidade com que se podem produzir gráficos bem desenhados e com qualidade de publicação, incluindo símbolos matemáticos e formulas, quando necessário, e o fato de ser gratuito. Pode ser descarregado na versão em português para *Linux*, *Windows* e *Mac OS X*, em `https://www.r-project.org/`.

As formas de interagir com o R são através das interfaces, como a mais básica `RGui` fornecida com o R ou a `RStudio`. A interface `RStudio` (`https://www.rstudio.com/`) é um ambiente de desenvolvimento integrado para o R que facilita a utilização do R. Outra das vantagens do `RStudio` é a de permitir construir de forma mais simples aplicações *web* interativas com o pacote `Shiny` (`https://shiny.rstudio.com/`).

De seguida explicaremos mais detalhadamente em que consiste o `RStudio` e o pacote `Shiny`, fundamentais para correr a aplicação por nós desenvolvida e que abordaremos no próximo capítulo.

1.2 O RStudio

O `RStudio` é um ambiente de desenvolvimento integrado (*Integrated Development Environment - IDE*) para o `R` que possibilita de forma simples a construção de aplicações. Para o correr é necessária a versão 2.11.1 ou superior do `R`. Inclui um editor de texto que suporta a edição, execução e visualização direta de código, possuindo ainda ferramentas que permitem instalar e atualizar os pacotes (os mesmos do `R`) bem como detetar e eliminar possíveis erros no código da aplicação desenvolvida.

Lançado em 2011, o `RStudio` é atualmente quase tão popular quanto o próprio `R`. De acordo com o seu fundador e principal programador, J.J. Allaire, este é o melhor *software* de código aberto para análise de dados, livremente disponível para qualquer pessoa com acesso a um computador. O seu foco está na criação de ferramentas para tratamento de dados de forma flexível e acessível para usuários do `R` e também para quem realiza o seu trabalho recorrendo a aplicativos, relatórios, painéis gráficos entre outros.

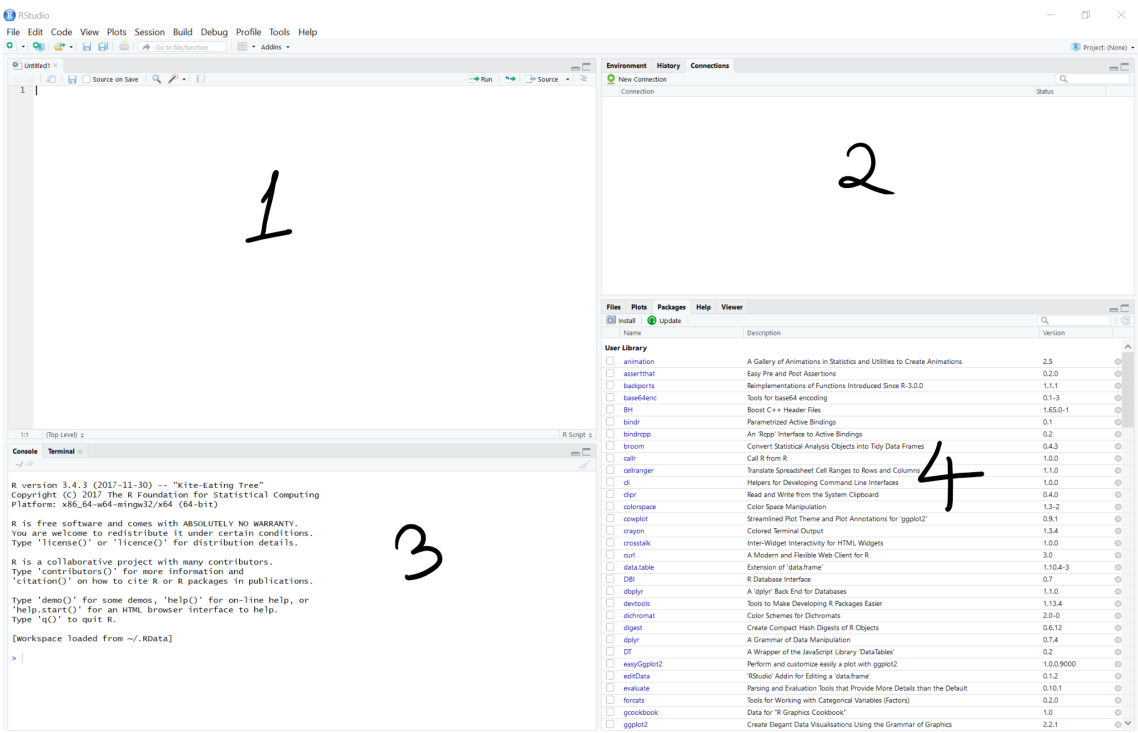

Figura 1.1: Interface RStudio.

Na Figura 1.1 podemos ver a interface `RStudio`, composta por quatro janelas, que passamos a descrever. Em (1) temos a janela Source na qual podemos abrir os *scripts* (códigos de programação,do R, previamente escritos e salvos em arquivos com extensão .R) e criar novos *scripts*. Nesta janela ao editar o script é possível comentar ou não automaticamente (sem necessidade de digitação) determinada parte do código utilizando a opção Edit/Comment/Uncomment Lines, mantendo o cursor na linha a ser executada ou selecionando o conjunto de linhas desejado. Os comandos que compõem o script podem ser executados através do botão Run disponível no cimo desta janela.

A janela (2) é composta por três separadores, dos quais destacamos dois, o `Environment` e o `History`. No separador `Environment` ficam armazenados os objetos e bases de dados criadas ou importadas e no separador `History` fica o histórico de todos os *scripts*, funções e ações executadas, histórico esse que também podem ser arquivado.

Na janela (3) temos a Console que é uma janela igual à janela padrão disponibilizada com o R e por conseguinte tem as mesmas funcionalidades de execução dos comandos do R.

A última janela, a (4), tem quatro separadores dos quais ressaltamos o `Plots` onde são exibidos os gráficos gerados por códigos digitados na linha de comandos, o `Packages` onde se pode ver que pacotes estão instalados, atualizar esses pacotes ou instalar novos pacotes e o `Help` no qual o utilizador pode obter ajuda sobre o funcionamento do `RStudio`.

De seguida iremos explorar as potencialidades do pacote "Shiny"para a construção de aplicações interativas.

1.3 O pacote **Shiny**

O `Shiny`, é um pacote do R que facilita a construção de aplicações *web* interativos, diretamente a partir do R, sem que seja necessário o utilizador possuir conhecimentos sobre o desenvolvimento de aplicações. Possui as funcionalidades *Shiny Server* e *RStudio Connect*, com as quais os utilizadores podem facilmente compartilhar as suas aplicações com outras equipas e utilizadores, dando a estes a capacidade de aceder a sofisticados algoritmos e modelos através da interatividade familiar da *web*. Esta partilha de aplicações é feita via nuvem, intranet ou internet e podem ser executadas mesmo em navegadores de *smartphones*, tornando-as ideais para uso dentro e fora da sala de aula.

Este pacote, de código aberto, dispõe de oito exemplos que demonstram como funciona e que permitem ao utilizador criar facilmente uma aplicação "Shiny"através destas já existentes. Na Tabela podemos ver os exemplos disponíveis e as suas funcionalidades.

Exemplo	Funcionalidade
runExample(“01_hello”)	um histograma
runExample(“02_text”)	tabelas e quadros de dados
runExample(“03_reactivity”)	uma expressão reativa
runExample(“04_mpg”)	variáveis globais
runExample("05_sliders")	barras deslizantes
runExample("06_tabsets")	janelas com separadores
runExample("07_widgets")	texto de ajuda e botões de envio
runExample("08_html")	aplicação Shiny construída em HTML
runExample("09_upload")	assistente de carregamento de ficheiros
runExample("10_download")	assistente de descarregamento de ficheiros
runExample("11_timer")	um temporizador automatizado

Tabela 1.1: Exemplos de aplicações Shiny e suas funcionalidades.

Para ver as aplicações da Tabela 1.1 é preciso instalar o pacote `Shiny`. Pera tal, basta estar ligado à internet, abrir uma sessão do `R` e executar o comando install.packages("Shiny"). Seguidamente, para ver por exemplo a primeira aplicação referida na tabela, é so executar os seguintes comandos:

```
library(shiny)
runExample("01_hello")
```

Em geral, uma aplicação `Shiny` é constituída, por dois ficheiros:

1. o *script* de interface do utilizador `ui.R` - componente que lida com a experiência do utilizador, isto é, define os detalhes da página (a aparência da aplicação), lista as opções e define os formatos de saída (ver Figura 1.2);

2. o *script* do servidor `(server.R)` - componente que faz todo o trabalho do `R`, o que significa que manipula todos os dados de entrada e instruções dadas para a aplicação e devolve à saída os objetos a serem exibidos no navegador (ver Figura 1.3).

Uma aplicação `Shiny` funciona como um conjunto de ferramentas desenhado para realizar tarefas e trabalhos específicos no computador ou em outro meio.

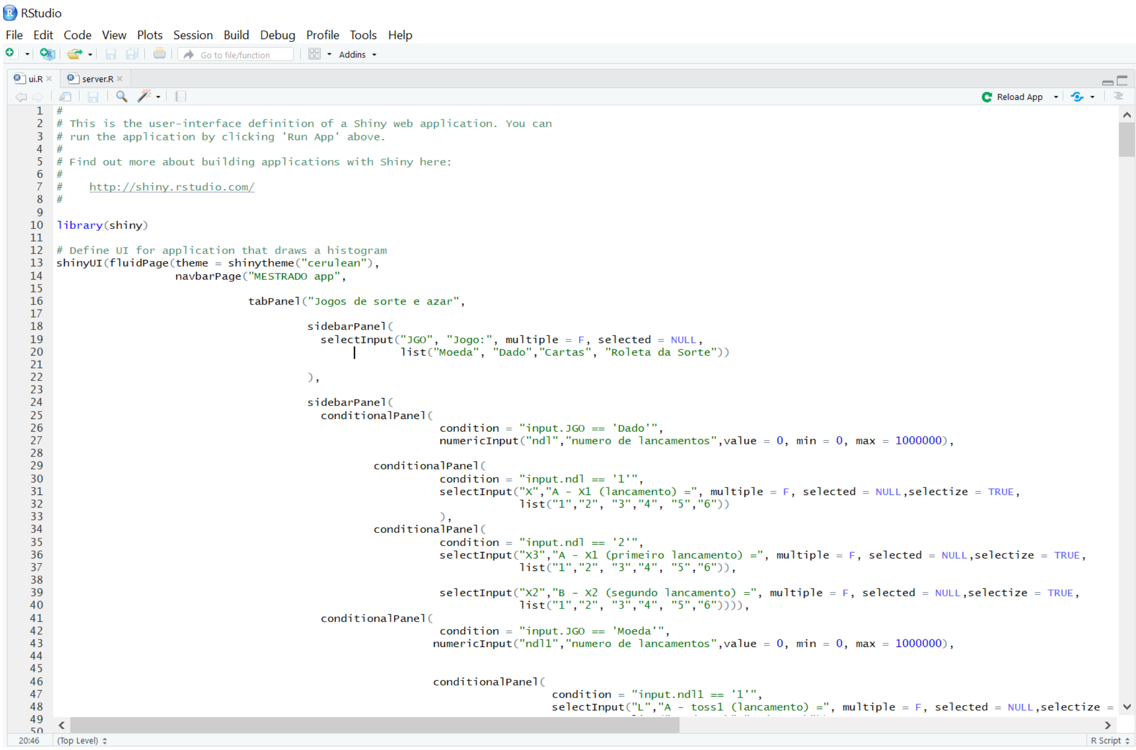

Figura 1.2: *Script* de interface do utilizador ui.R.

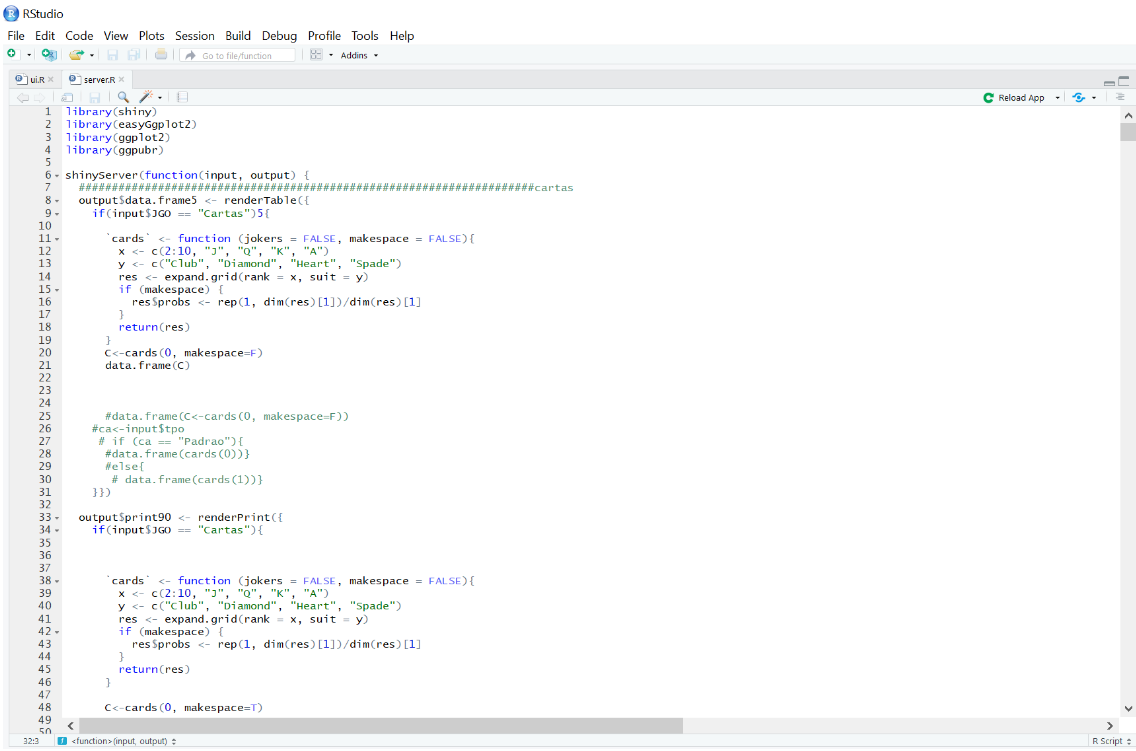

Figura 1.3: *Script* do servidor server.R.

Tem janelas e botões que possibilitam a entrada de dados que são interpretados em código `R` e, posteriormente, enviadas de volta para a janela do navegador. Os dados de entrada podem ser de diversos tipos, texto, números, etc., existindo suportes de saída em quase todas as formas familiares e necessárias para o tratamento de informação, incluindo figuras, tabelas, gráficos e resumos. Nos vários exemplos de aplicações `Shiny` que constam na Tabela 1.1 ilustram os vários suportes de saída disponíveis.

Em `https://shiny.rstudio.com/tutorial/` podemos encontrar um conjunto de tutoriais, na forma de vídeo e de texto, que nos conduzem na construção de uma aplicação `Shiny`.

Tendo em conta as potencialidades que as ferramentas do pacote `Shiny` apresentam, desenvolvemos uma aplicação para o ensino de probabilidade e estatística, ao nível do ensino secundário e de cursos introdutórios nesta área. O próximo capítulo será dedicado à apresentação da aplicação por nós desenvolvida. Iremos explorar todas as suas funcionalidades numa perspetiva de utilização no processo de ensino e aprendizagem. Ao longo deste processo iremos sempre apresentar os conceitos e resultados teóricos subjacentes a cada uma das funcionalidades.

Capítulo 2

Uma aplicação **Shiny** para ensino das probabilidades e estatística

2.1 Introdução

O ensino das probabilidades e estatística assume um papel cada vez mais importante por possibilitar o desenvolvimento da capacidade de tomar decisões e do espírito crítico. Tal irá promover uma participação activa, crítica e esclarecida por parte dos alunos em relação a resultados que lhe são apresentados no dia a dia, muitas vezes através dos meios de comunicação social.

Na perspetiva de Carvalho (2006), nas sociedades modernas um cidadão deve ter conhecimento dos conceitos probabilísticos e estatísticos de modo a tornar-se num ser crítico em relação à informação disponível, entender e comunicar com base nessa informação e tomar decisões individuais e coletivas, uma vez que grande parte da organização dessas mesmas sociedades é feita com base nesses conhecimentos. Por sua vez, numa perspectiva curricular, Abrantes *et al.* (1999) argumentam que os conceitos de estatística e probabilidades ajudam a compreender outros tópicos matemáticos ligados aos números, às medidas ou às representações gráficas e evidenciam diversas conexões matemáticas com noções relativas a frações, percentagens, proporções ou números decimais, entre outras.

Apesar de cada vez mais se acentuar a importância das probabilidades e estatística e das suas aplicações na sociedade e no desenvolvimento do raciocínio matemático, no sistema educativo Português os programas que se encontram em vigor no ano lectivo 2017/18, nomeadamente, o programa de Matemática do 3^o Ciclo do Ensino Básico[1] e os programas de Matemática A[2] e B[3] e do Ensino Secundário, parecem não lhes dar o devido valor. O tempo que os programas

[1] `http://www.dge.mec.pt/sites/default/files/Basico/Metas/Matematica/programa_matematica_basico.pdf`

[2] `http://www.dge.mec.pt/sites/default/files/Secundario/Documentos/Documentos_Disciplinas_novo/Cursos_Cientifico_Humanisticos/programa_metas_curriculares_matematica_a_secundario.pdf`

[3] `http://www.dge.mec.pt/sites/default/files/Secundario/Documentos/Documentos_Disciplinas_novo/Curso_de_Artes_Visuais/matematica_b_10_novo.pdf` e [4]

atribuem às temáticas de estatística e de probabilidades, torna-as ainda marginais aos currículos, afastando-as facilmente para segundo plano Caldeira (2009).

.

No 3^o Ciclo do Ensino Básico a temática de Estatística e Probabilidades temática é abordada no tópico de organização e tratamento de dados, sendo que no 7^o ano se faz referência à organização e tratamento de dados, no 8^o ano ao planeamento estatístico e no 9^o ano se introduz a noção de probabilidade. De acordo com o programa de matemática deste ciclo de estudo, ao longo deste período o aluno deve adquirir conhecimento de conceitos e representações de modo a compreender e ser capaz de produzir informação estatística e de a utilizar para resolver problemas e tomar decisões informadas. O mesmo refere também a necessidade do aluno desenvolver a compreensão da noção de probabilidade, tanto no seu aspeto teórico como experimental.

Já no ensino secundário, a abordagem da estatística e das probabilidades é feita com o intuito de complementar as aprendizagens básicas adquiridas no ensino básico, com algumas noções novas e ferramentas que não podem ser compreendidas nesse nível de ensino. No 10^o ano começa-se por introduzir as definições de variável estatística, amostra, média, variância, desvio-padrão e percentil, analisam-se as propriedades básicas destes conceitos e as respetivas interpretações em exemplos concretos. No 11^o ano estudam-se as retas de mínimos quadrados associadas a uma sequência de pontos do plano. As coordenadas destes pontos podem em particular representar os valores de uma amostra bivariada, o que permite a aplicação deste conceito ao estudo da correlação de duas variáveis estatísticas definidas numa mesma amostra. No 12^o ano, aborda-se de um modo mais geral a noção de probabilidade, começando por se introduzir a noção de função de probabilidade definida no conjunto das partes de um conjunto finito relacionado com situações de equiprobabilidade. É igualmente abordada a noção de probabilidade condicionada e de independência de acontecimentos, apresentando-se em particular a lei dos grandes números e o Teorema da probabilidade total.

Na maior parte dos casos, a transmissão destes conhecimentos na sala de aula é feita por intermédio de cálculos longos e trabalhosos que, além de exigirem um elevado esforço mental, podem tornar as aulas altamente desmotivadoras. Esta prática pode também limitar a capacidade dos alunos em desenvolver o raciocínio probabilístico e estatístico uma vez que, mesmo que consigam realizar os cálculos, estes podem apresentar grandes dificuldades em dar um significado aos resultados obtidos. De acordo com o programa de matemática do ensino secundário, é importante que o estudo da estatística e das probabilidades contribua para melhorar a capacidade dos estudantes para avaliar afirmações de carácter probabilístico e estatístico, fornecendo-lhe ferramentas apropriadas para rejeitar quer certos anúncios publicitários quer notícias ou outras informações em que a interpretação de dados ou a realização da amostragem não tenha sido correta. Abre-se aqui uma oportunidade para a realização de atividades interdisciplinares, individuais ou em grupo, incentivando-se veemente a utilização do computador.

Quer no programa de matemática do 3^o ciclo do ensino básico, quer no programa do ensino secundário, é feita uma alusão à utilização de meios tecnológicos no decorrer das atividades letivas relacionadas ao ensino da estatística e das probabilidades. A partir do 3^o ciclo do ensino básico, os professores são encorajados a fazer uso de simulações, quer seja com a máquina de calcular, quer seja com o computador. As simulações são um instrumento poderoso que com o desenvolvimento dos meios computacionais contribuem de forma decisiva para o estudo de leis da probabilidade e cálculo de probabilidades associadas a determinados acontecimentos. Com este instrumento torna-se possível estimar a probabilidade de um acontecimento, através repetição sucessiva da experiência aleatória e e posterior contabilização da proporção de vezes que o acontecimento se realiza nas sucessivas repetições, correspondendo isto à definição frequencista de probabilidade.

A inclusão das tecnologias de informação no ensino da estatística e das probabilidades permite aos alunos concentrar a sua atenção nos aspetos mais relevantes como interpretar, organizar, discutir, argumentar e não sobre os mais mecânicos associados à sua realização.

É com o intuito de promover a inclusão de tecnologias como instrumentos facilitadores da aprendizagem no ensino da estatística e das probabilidades que a presente aplicação `Shiny` foi desenvolvida. As simulações propostas nessa aplicação podem auxiliar o professor na sala de aula e incentivar o trabalho colaborativo entre os alunos, despoletando uma troca construtiva de opiniões entre estes. Além de tornar a aula mais motivadora, levará os alunos a construir melhor o conhecimento sobre fenómenos demasiado abstratos que seriam mais difíceis de compreender apenas com os conceitos descritos durante as aulas teóricas. Ainda assim, defendemos que a aplicação deve ser utilizada apenas como uma ferramenta complementar devendo-se priorizar a realização de cálculos analíticos e construção de tabelas e gráficos utilizando os materiais de trabalho comuns.

No que se segue, iremos descrever com maior pormenor os separadores que constituem a aplicação `Shiny` e como as simulações que estes contêm podem ser aplicadas na sala de aula, tendo sempre como preocupação a revisão dos resultados teóricos fundamentais que estão por trás. Os conceitos e resultados aqui apresentados podem ser encontrados em Gonçalves e Lopes (2000) e Pestana e Velosa (2010).

2.2 A aplicação Shiny desenvolvida

A aplicação `Shiny` por nós desenvolvida pode ser um auxílio no processo de ensino e aprendizagem das probabilidades e estatística, uma vez que simula e ilustra alguns resultados referidos na literatura como de difícil percepção pelos alunos. Poderá também proporcionar um conhecimento mais abrangente e efetivo. Na Figura 2.1 podemos ver o aspeto da interface da aplicação, disponível a partir do momento a que a ela acedemos.

Figura 2.1: A interface da aplicação Shiny.

Como se pode observar, pela Figura 2.1, no topo da interface da aplicação existe uma barra de navegação na qual constam cinco separadores diferentes. Em cada um destes separadores ilustra-se um resultado de probabilidades e estatística: no primeiro simulam-se experiências relacionadas com jogos de sorte e azar e calculam-se probabilidade de acontecimentos relacionados com a experiência em causa; no segundo ilustra-se a definição frequencista de probabilidade; no terceiro ilustra-se a lei dos grandes números; no quarto ilustram-se graficamente as características de algumas distribuições de probabilidade; no último simulam-se as distribuições por amostragem da soma e da média de variáveis aleatórias reais independentes e identicamente distribuídas.

Ao clicar num dos separadores anteriores a aplicação apresenta alguns menus dinâmicos para entrada de dados, variando em função das características de cada simulação, visíveis na Figura 2.1.

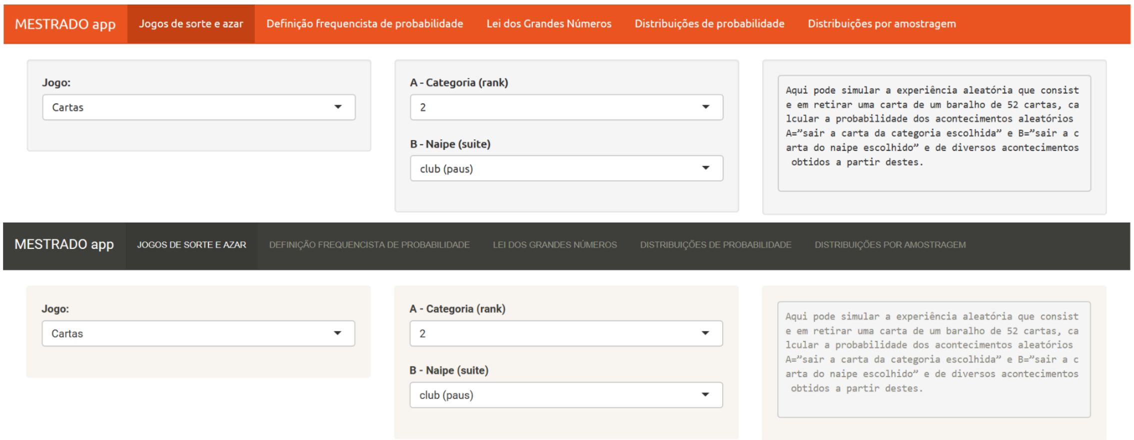

Figura 2.2: Alguns dos temas da interface da Aplicação Shiny.

O aspeto da interface pode ser alterado bastando para tal escolher um tema diferente no menu Select theme que aparece na interface, existindo 17 diferentes ao todo. Na Figura 2.2 podemos ver o aspeto com que fica a interface da aplicação ao selecionarmos alguns dos temas disponíveis.

Mais importante ainda que a diversidade de temas, é a linguagem `R` que está por trás desta interface. As simulações envolvidas em cada um dos separadores foram geradas com código `R` e respeitáram todas as condições exigidas nos correspondentes resultados teóricos. Descartamos assim, a existência de simulações viciadas ou com resultados pré-estabelecidos. No que se segue exploramos detalhadamente e individualmente cada um dos separadores.

2.2.1 Jogos de sorte e azar

O aparecimento da teoria das probabilidades teve sobretudo a ver com o objetivo de dar uma explicação racional a certos fenómenos observados em jogos de dados e cartas, genericamente designados por jogos de sorte e azar. Estes fenómenos são hoje em dia chamados "fenómenos em massa", isto é, propriedades que aparecem quando jogamos um número (suficientemente) grande de partidas.

As experiências aleatórias que estiveram na origem da Teoria das Probabilidades, ligadas como já referimos aos jogos de sorte e azar, apresentam duas características comuns:

- número finito de resultados possíveis;
- resultados individuais igualmente possíveis (hipótese de equiprobabilidade ou princípio de simetria).

Como experiências nestas condições temos, por exemplo, o lançamento de uma moeda ou de um dado equilibrado, a extração de uma carta de uma baralho com 52 cartas, etc. No separador Jogos de sorte e azar da aplicação `Shiny` podemos simular quatro experiências que verificam as condições anteriores, que incluem retirar uma carta de um baralho com 52 cartas, lançar duas vezes um dado equilibrado, lançar duas vezes uma moeda equilibrada e observar a secção onde cai a bola numa roleta. Na Figura 2.3 podemos ver este separador com o menu onde são disponibilizadas as quatro experiências aleatórias.

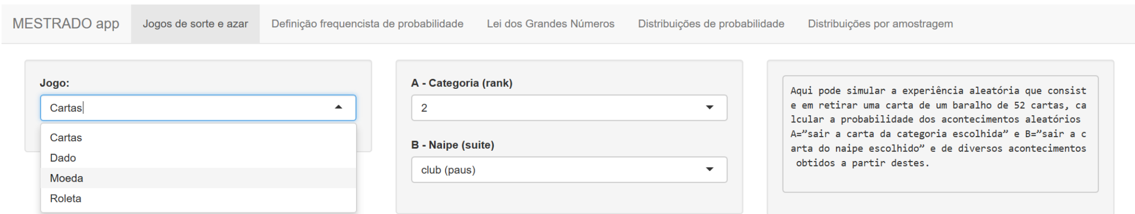

Figura 2.3: Experiências simuladas no separador Jogos de sorte e azar.

A partir destas simulações é possível determinar de forma imediata o espaço de resultados associado à experiência em causa. Recordamos que uma experiência aleatória é uma experiência que pode ser realizada infinitas vezes, mas o resultado obtido será sempre incerto desde que se realizem nas mesmas condições, contudo é possível conhecer todos os resultados possíveis de ocorrer antes da sua realização.

Para modelar matematicamente este tipo de experiências, a primeira etapa a ser cumprida deve ser a definição do conjunto de todos os resultados possíveis de obter com a sua realização, este conjunto é chamado espaço de resultados ou espaço fundamental e será denotado por Ω. Cada um dos seus elementos $\omega \in \Omega$ caracteriza completamente uma prova, ou seja, uma realização da experiência aleatória.

Determinar de forma correta o espaço de resultados associado a uma experiência aleatória nem sempre é trivial para os alunos, por um lado devido à dificuldade em idealizar a experiência e por outro pela enorme quantidade de cálculos analíticos que por vezes se precisa realizar. Este separador facilita essa tarefa, pois para as quatro experiências possíveis de simular, a aplicação apresenta o espaço de resultados bem como o cardinal deste conjunto. De seguida veremos um exercício que pode ser resolvido facilmente com a aplicação `Shiny`.

Exercício 2.2.1. *Fazem-se dois lançamentos de um dado perfeito, o qual apresenta as faces numeradas de 1 a 6.*

1. *Determine o espaço de resultados associado a esta experiência aleatória;*

2. *Defina o acontecimento:*

   (a) $A \equiv$*"saída do número 2 no primeiro lançamento";*

   (b) $B \equiv$*"saída do número 4 no segundo lançamento";*

   (c) *"saída do número 2 no primeiro lançamento ou saída do número 4 no segundo lançamento";*

   (d) *"saída do número 2 no primeiro lançamento e não saída do número 4 no segundo lançamento".*

Na Figura 2.4 podemos ver parte do espaço de resultados da experiência referida no exercício anterior bem como a indicação do número de elementos que o constituem. A partir da visualização do espaço de resultados por parte dos alunos, a definição dos acontecimentos pedida no exercício torna-se muito mais simples.

Após a definição do espaço de resultados, vimos que é possível o aluno facilmente identificar subconjuntos deste mesmo espaço, subconjuntos estes que representam acontecimentos associados à experiência aleatória em causa. Recordemos que todo acontecimento associado a uma experiência aleatória é identificado por um subconjunto de Ω. Deste modo, dois acontecimentos dir-se-ão equivalentes se determinam o mesmo subconjunto de Ω.

Denotemos por $\mathcal{P}(\Omega)$ o conjunto de partes de Ω, isto é, o conjunto de todos os subconjuntos de Ω, ou o conjunto de todos os acontecimentos possíveis de associar a uma experiência aleatória. Dado que perante uma certa experiência aleatória podem apenas interessar-nos certos acontecimentos, é fácil perceber que a classe de acontecimentos associada a uma experiência não é necessariamente $\mathcal{P}(\Omega)$. Contudo, é imprescindível garantir que uma tal classe de subconjuntos de

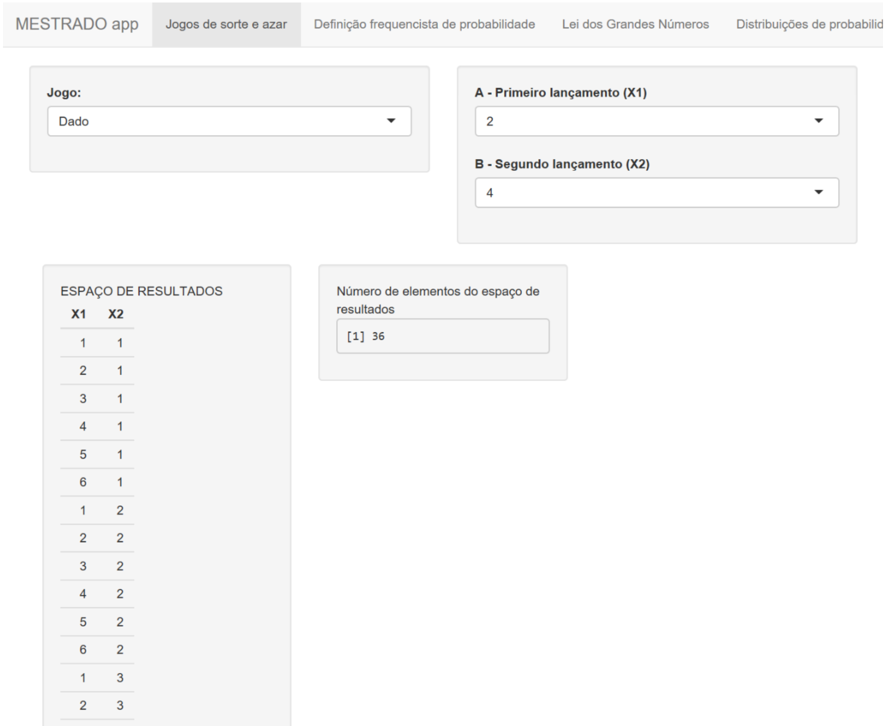

Figura 2.4: Espaço de resultados.

Ω seja estável para as operações usuais de conjuntos, para que todos os acontecimentos que nos interessam pertençam a essa classe. Um tal classe, que iremos denotar por $\mathcal{A}$, é uma σ-álgebra ou tribo de acontecimentos.

Definição 2.1. *Seja Ω um conjunto e $\mathcal{A}$ uma classe de subconjuntos de Ω. Dizemos que $\mathcal{A}$ é uma σ-álgebra ou tribo sobre Ω se verifica os seguintes axiomas:*

a) $\Omega \in \mathcal{A}$;

b) $\forall A \subset \Omega,\ A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$;

c) Se $\{A_n\}_{n \in \{N}$ *é uma sucessão qualquer de acontecimentos de* $\mathcal{A}$*, então* $\displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Assim, ressaltamos novamente que as experiências aleatórias simuladas na aplicação terão associado um Ω finito, logo nestes casos as reuniões numeráveis reduzem-se a uniões finitas, e por conseguinte a σ - álgebra $\mathcal{A}$, terá apenas de verificar os axiomas a) e b) da definição anterior e o axioma c') seguinte:

c') $\forall A, B \in \mathcal{A},\ A \cup B \in \mathcal{A}$

Notemos que no caso de Ω ser finito, podemos tomar $\mathcal{A} = \mathcal{P}(\Omega)$, embora em muitas circunstâncias baste tomar espaços de acontecimentos muito mais simples, tais como $\{\emptyset, \Omega, A, \overline{A}\}$.

Com a realização de um acontecimento torna-se necessário quantificar a possibilidade da sua realização e obter o que se designa por probabilidade do acontecimento. Tendo em conta as especificidades das experiências relacionadas com os jogos de sorte e azar Laplace em 1812 deu a seguinte definição de probabilidade.

Definição 2.2. *Se uma experiência aleatória tiver um espaço de resultados Ω finito e os seus resultados elementares forem igualmente e prováveis, então a probabilidade de A, um subconjunto de Ω, se realizar é igual ao quociente entre o número de casos favoráveis à ocorrência de A (número de elementos de A, $\#A$), e o número de casos possíveis (número de elementos do espaço de resultados Ω, $\#\Omega$). Assim,*

$$P(A) = \frac{\textit{número de casos favoraveis a } A}{\textit{número de casos possíveis}} = \frac{\#A}{\#\Omega}.$$

Esta definição de probabilidade está incluída na definição de probabilidade devida a Kolmogorov (1933) que apresentamos a seguir apenas para o caso de Ω finito.

Definição 2.3. *Seja Ω um espaço de resultados finito associados a uma experiência aleatória. Chama-se medida de probabilidade a uma função P, definida sobre $\mathcal{A}$, que satisfaça os seguintes axiomas:*

1. $P(\Omega) = 1$.

2. *Qualquer que seja o acontecimento A, existe $P(A) \geq 0$.*

3. *Se A, B são acontecimentos disjuntos, $A \cap B = \emptyset$, então*

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

Como consequências imediatas desta axiomática obtemos as propriedades que a seguir enunciamos e demonstramos.

Propriedade 2.2.1. *A probabilidade do acontecimento impossível é nula, isto é, $P(\emptyset) = 0$.*

Demonstração: Temos que $\Omega \cap \emptyset = \emptyset$ e $\Omega \cup \emptyset = \Omega$, logo pelo Axioma 3 vem

$$P(\Omega) = P(\Omega \cup \emptyset) = P(\Omega) + P(\emptyset)$$

e pelo Axioma 2 resulta $P(\emptyset) = 0$. □

Propriedade 2.2.2. *Se A e B são dois acontecimentos tais que $A \subseteq B$ então*

$$P(A) \leq P(B) \quad e \quad P(B - A) = P(B) - P(A).$$

Em particular, $P(A) = 1 - P(\overline{A})$.

Demonstração: Como P é uma função não negativa e

$$B = A \cup (B - a) \quad \text{e} \quad A \cap (B - A) = \emptyset,$$

logo pelo Axioma 3 vem

$$P(B) = P(A) + P(B_A),$$

o que prova o resultado. □

Em geral $B - A = B - (A \cap B)$, logo pelo resultado anterior temos que

$$\forall A, B \in \mathcal{A}, \; P(B - A) = P(B) - P(A \cap B).$$

Propriedade 2.2.3. *Se A e B são dos acontecimentos quaisquer*

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

Demonstração: Observa-se que $A \cup B = (A - B) \cup (B - A) \cup (A \cap B)$. Pelo facto de $(A - B)$, $(B - A)$ e $(A \cap B)$ serem disjuntos dois a dois (a interseção entre si é $\emptyset$), pelo o Axioma 3 temos

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

Mas tendo em conta a Propriedade 2.2.2, $P(A - B) = P(A) - P(A \cap B)$ e

$P(B-A) = P(B) - P(B \cap A)$ logo,

$$\begin{aligned} P(A \cup B) &= P(A-B) + P(B-A) + P(A \cap B) \\ &= P(A) - P(A \cap B) + P(B) - P(B \cap A) + P(A \cap B) \\ &= P(A) + P(B) - P(B \cap A). \end{aligned}$$

□

A definição de probabilidade de Laplace e as propriedades anteriores podem ser vistas em prática no separador Jogos de sorte e azar. Aqui, para as quatro experiências, podemos calcular probabilidades de dois acontecimentos A e B, bem com de outros acontecimentos compostos a partir destes. Para melhor visualizarmos este fato completamos o Exercício 2.2.1 com algumas questões sobre o cálculo de probabilidades de acontecimentos.

Exercício 2.2.2. *1. Calcule a probabilidade dos acontecimentos definidos no Exercício 2.2.1.*

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

A resposta às questões anteriores é facilmente obtida com a aplicação, como podemos ver pela Figura 2.5.

Os alunos podem usar agora este separador para confirmar os resultados obtidos. Além do mais, podem repetir o exercício para muitos outros acontecimentos diferentes e assim cimentar o seu conhecimento. Os professores podem aproveitar os resultados obtidos pela aplicação para questionar os alunos, por exemplo, sobre propriedades dos acontecimentos aleatórios, como a de independência ou incompatibilidade.

Se na experiência aleatória do Exercício 2.2.1 soubermos que no primeiro lançamento saiu o número 2, a probabilidade de sair o número 4 no segundo lançamento é diferente da anteriormente calculada, que era igual a 0.17, pois agora não precisamos de olhar para todos os resultados mas apenas para aqueles em que no primeiro lançamento aparece o número 2. Esta probabilidade é designada por probabilidade condicionada.

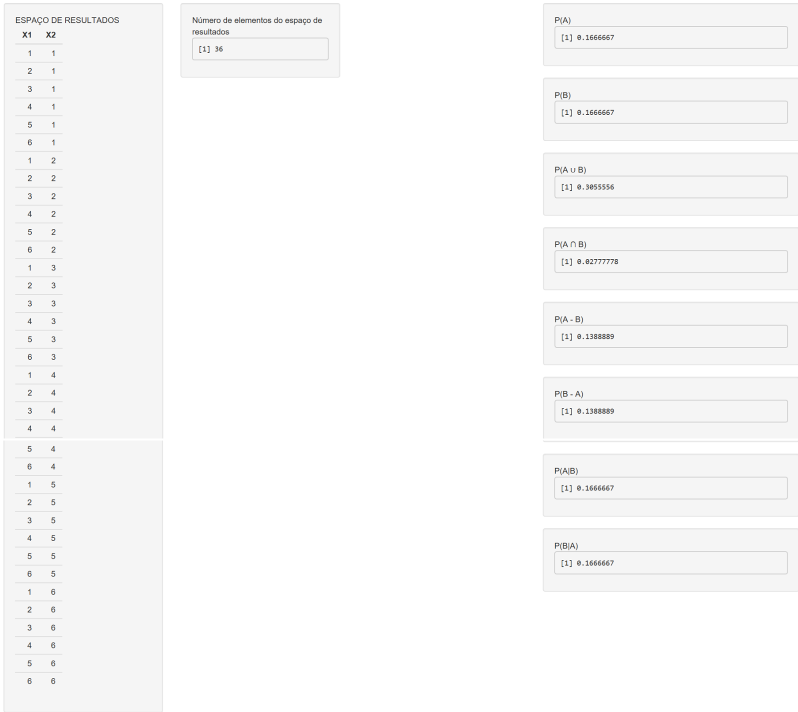

Figura 2.5: Cálculo de algumas probabilidades.

Definição 2.4. *Seja B um acontecimento de $\mathcal{A}$ de probabilidade não nula, chama-se probabilidade condicionada por B à função P_B, ou $P(.|B)$ definida sobre $\mathcal{A}$ por*

$$\forall A \in \mathcal{A}, P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)},$$

onde $P(A|B)$ lê-se probabilidade de A condicionada por B ou probabilidade de A dado B.

Obviamente que $P(A|B)$ não tem que ser necessariamente igual a $P(B|A)$, apesar de no Exercício 2.2.1 constatarmos que o valor coincide para os acontecimentos A e B, Figura 2.6.

.

Em conclusão, este separador pode ser utilizado quer por professores que por alunos para o ensino e aprendizagem dos conceitos básicos de probabilidades que abordámos anteriormente. O fato de haver quatro experiências aleatórias diferentes, para as quais é possível definir vários acontecimentos e calcular as suas probabilidades, aumenta a funcionalidade e abrangência deste separador.

Figura 2.6: Probabilidade condicionada.

2.2.2 Definição frequencista de probabilidade e Lei dos grandes números

A teoria elementar (ou clássica) das Probabilidades assenta fundamentalmente na constatação de que os fenómenos em causa apresentam as características seguintes:

- cada vez que se realiza a experiência, obtém-se um resultado individual que somos incapazes de prever com exatidão;
- repetindo a experiência em causa um número suficientemente grande de vezes nas mesmas condições, os resultados obtidos apresentam, quando analisados em conjunto, grande regularidade estatística.

Os fenómenos com as características anteriores são designados de fenómenos aleatórios e a regularidade estatística não é mais do que a frequência relativa de uma certo acontecimento estabilizar à volta de um valor constante, quando realizamos a experiência um número suficientemente grande de vezes nas mesmas condições.

A regularidade estatística associada aos fenómenos aleatórios conduz à interpretação intuitiva de probabilidade como frequência-limite. Tal levou à definição de probabilidade hoje designada por definição frequencista. O resultado que justifica tal aproximação foi obtido por Bernoulli e será enunciado de seguida.

Teorema 2.5. *Sendo $\mathcal{E}$ uma experiência aleatória e A um acontecimento associado a esta experiência, designemos por $f_n(A)$ a frequência relativa de A em n realizações da experiência sempre nas mesmas condições. Tem-se*

$$\forall \epsilon > 0, \lim_{n \to +\infty} P(|f_n(A) - p| \leq \epsilon) = 1,$$

onde $p \in [0,1]$ é a probabilidade de A.

Este resultado é a versão primária da Lei dos Grandes Números. Existe outra formulação a que habitualmente se dá o nome de Lei Fraca dos Grandes Números, que passamos a enunciar.

Teorema 2.6. *Seja* $\{X_n\}_{n\geq 1}$ *uma sucessão de v.a.r. independentes e identicamente distribuídas (i.i.d.) com* $E[X_1] = \mu$. *Então,*

$$\forall \epsilon > 0, \ \lim_{n\to+\infty} P(|\overline{X}_n - \mu| \leq \epsilon) = 1,$$

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

A Lei dos grandes números de Bernoulli mostra, entre outras coisas, que a sucessão de médias amostrais $\overline{X}_n$ converge para o valor médio populacional μ, sob a hipótese desse valor médio existir. O Teorema 2.5 é assim uma consequência do Teorema 2.6, uma vez que se considerarmos X a variável aleatória indicatriz do acontecimento A, de probabilidade $P(A)$,

$$X = \mathbb{1}_A = \left\{ \begin{array}{ccc} 0 & , & 1 \\ 1 - P(A) & , & P(A) \end{array} \right. ,$$

cujo valor médio é $P(A)$, temos,

$$f_n(A) = \frac{\mathbb{1}_{A_1} + \mathbb{1}_{A_2} + ... + \mathbb{1}_{A_n}}{n} = \overline{\mathbb{1}}_{A_n} \xrightarrow[n\to\infty]{P} E[\mathbb{1}_A] = P(A),$$

onde $\xrightarrow[n\to\infty]{P}$ denota a convergência em probabilidade.

A importância da Lei dos grandes números na teoria das probabilidades é incontestável, aliás Bernoulli chamava-lhe "o meu teorema de ouro". Assim, dedicámos na nossa aplicação dois separadores para este resultado. No primeiro destes separadores, chamado Definição frequencista de probabilidade, ilustramos o Teorema 2.5 com a simulação de sucessivas realizações da experiência aleatória "lançamento de n moedas k vezes e observação do número de caras obtido", com $n = 1, \ldots, 12$ e $k = 1, \ldots, 10000$. Na Figura 2.7 podemos ver este separador mais em pormenor.

Para além do utilizador poder escolher no separador Definição frequencista de probabilidade o número de moedas a lançar em cada lançamento, variando este

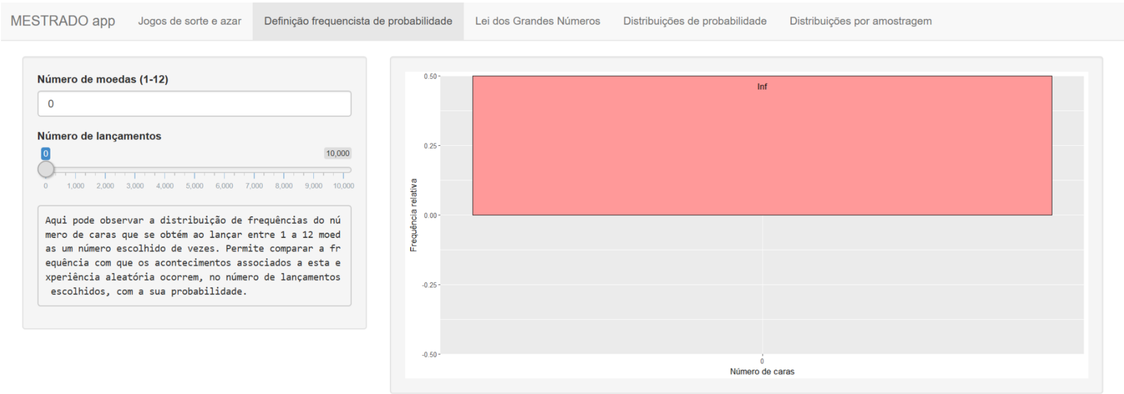

Figura 2.7: Separador Definição frequencista de probabilidade.

de uma a doze moedas, também pode escolher entre 1 a 10000 lançamentos das moedas. Após a sua escolha, a aplicação apresenta o gráfico de barras com as frequências relativas associadas ao número de caras obtidas. No gráfico podemos ver que à medida que o número de lançamentos aumenta o valor da frequência relativa do acontecimento "sair x caras", com $x = 0, 1, 2, \ldots,$ aproxima-se do valor da probabilidade com que este acontecimento ocorre. Tal, permite perceber o que o Teorema 2.5 significa em termos práticos o que se torna uma mais valia na sala de aula, uma vez que os alunos em geral têm dificuldade em perceber os resultados teóricos e as suas inúmeras aplicações. Aliás, perceber bem esta aproximação é fundamental quando fazemos estimação em Estatística Inferencial, pois neste caso a frequência relativa de um acontecimento é um estimador da probabilidade desse acontecimento.

Gordon e Gordon (2009) referem mesmo que é muitas vezes pedido a estudantes em cursos introdutórios de estatística para aceitarem muitas afirmações e procedimentos, uma vez que as justificações matemáticas são demasiado sofisticadas para eles perceberem. Tal só reforça a utilidade das visualizações de simulações, disponíveis numa aplicação como a nossa, para a melhor compreensão de conceitos e resultados probabilísticos e estatísticos difíceis de visualizar e compreender por parte dos alunos.

As potencialidades deste separador podem ser exploradas pelos alunos com exemplos ou exercícios propostos pelo professor. Enunciamos a seguir em exercício que poderá ser usado neste contexto, para conduzir o aluno até Lei dos Grandes Números.

Exercício 2.2.3. *Considere a experiência aleatória que consiste em lançar ao ar*

3 moedas equilibradas.

a *Determine a probabilidade de acontecimento A≡ "saída de uma cara".*

b *Utilize a aplicação* `Shiny`*, separador* Definição frequencista de probabilidade*, para repetir esta experiência aleatória 50 vezes. Qual é o valor da frequência relativa de sair uma cara?*

c *Repita a alínea anterior para 500, 1000, 5000 e 10000 lançamentos. Compare as frequências relativas obtidas com a probabilidade obtida na alínea a). O que conclui?*

Obviamente que a experiência do exercício anterior é uma experiência Binomial, por conseguinte, se os alunos já tiverem adquirido o conhecimento sobre a distribuição Binomial, a alínea a) pode ser resolvida através do cálculo da probabilidade de uma variável aleatória Binomial de parâmetros 3 e 0.5 tomar o valor um. Deste modo o aluno chegará mais facilmente ao valor de 0.375 pretendido. Na Figura 2.8 vê-se a resolução da alínea b) na aplicação.

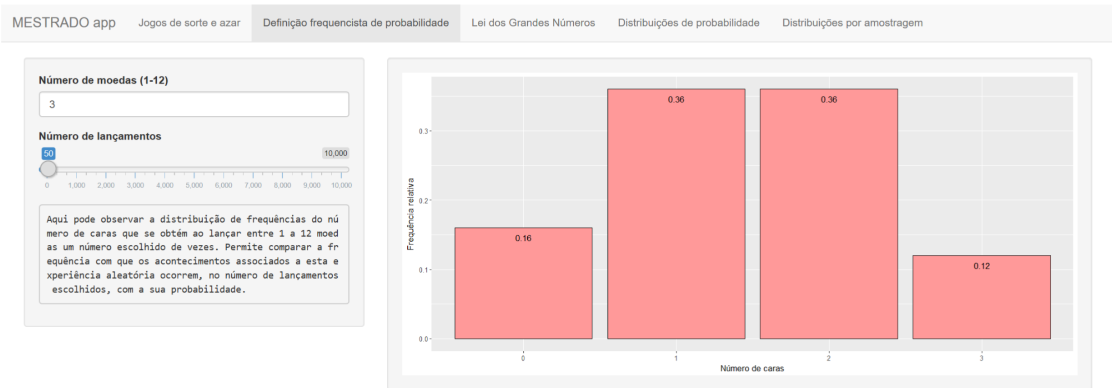

Figura 2.8: Frequências relativas do número de caras em 50 lançamentos de 3 moedas equilibradas.

Este separador pode ainda servir como introdução ao Teorema Limite Central, uma vez que, para um grande número de lançamentos, à medida que aumentamos o número de moedas começamos a ver que a distribuição de frequências começa a ter uma forma de sino semelhante ao modelo normal. Claro que o pequeno número de moedas lançadas não é suficiente para ilustrar a aproximação da Binomial à Normal, mas é sem dúvida um indicador do que estará para vir.

No segundo destes separadores, intitulado Lei dos Grandes Números, ilustramos o Teorema 2.6. A importância deste separador é justificada por Gordon e Gordon

(2009) , que afirmam que um problema comum para muitos alunos é a natureza não intuitiva de probabilidade que faz com que não consigam visualizar o que acontece a longo prazo numa situação aleatória. Muito provavelmente porque não entendem o que significa a aleatoriedade, contudo as simulações gráficas para a Lei dos Grandes Números que mostram a diferença entre o que acontece a curto e a longo prazo podem ajudar a resolver este problema. Na Figura podemos ver a estrutura deste separador.

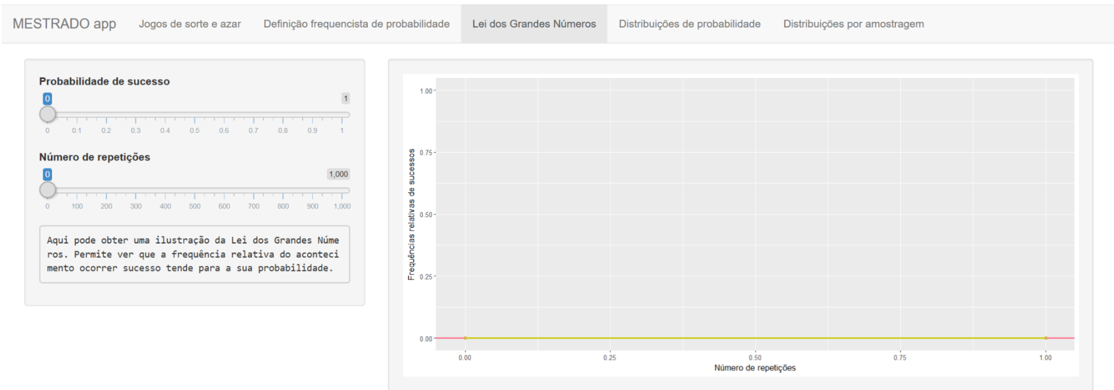

Figura 2.9: Separador Lei dos Grandes Números.

A regularidade estatística associada aos fenómenos aleatórios que referimos inicialmente pode assim ser ilustrada no separador Lei dos Grandes Números. Para percebemos melhor como funciona este separador consideremos o exemplo seguinte de Gonçalves e Lopes (2000).

Exemplo 2.2.1. Consideremos a experiência aleatória que consiste em efetuar o lançamento de uma moeda de 1 Euro.

Em cada lançamento, ou prova, pode ocorrer a face cara ou coroa. Se estivermos interessados em que sai face cara, então esta será o nosso "sucesso" que ocorre com probabilidade 0.5. No separador colocamos assim o botão Probabilidade de sucesso em 0.5. Podemos de seguida escolher o tamanho da sucessão de lançamentos, por exemplo 50, podendo este valor variar de 1 a 1000. O número médio de ocorrências de "face cara" ao longo das realizações da experiência pode agora ser visualizada graficamente na aplicação, como podemos ver na Figura 2.10.

Seguidamente podemos aumentar o número de lançamentos, por exemplo para 500 ou mesmo 1000 (ver Figura 2.11).

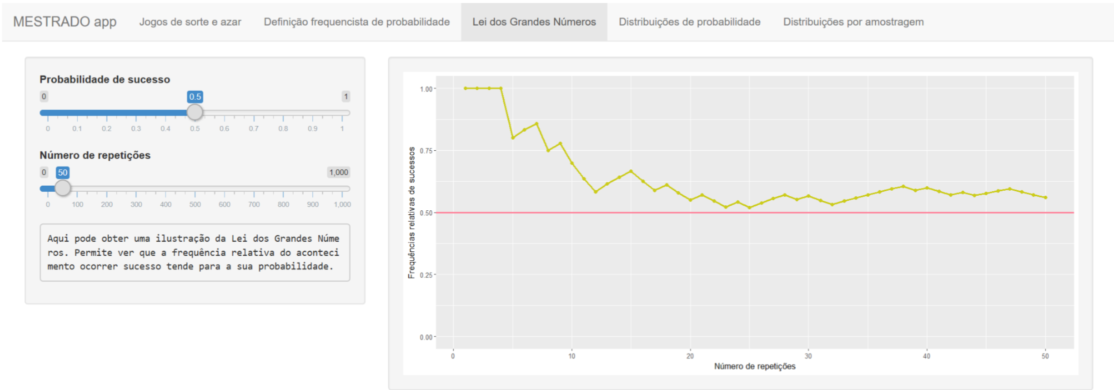

Figura 2.10: Simulação de 50 lançamentos de uma moeda.

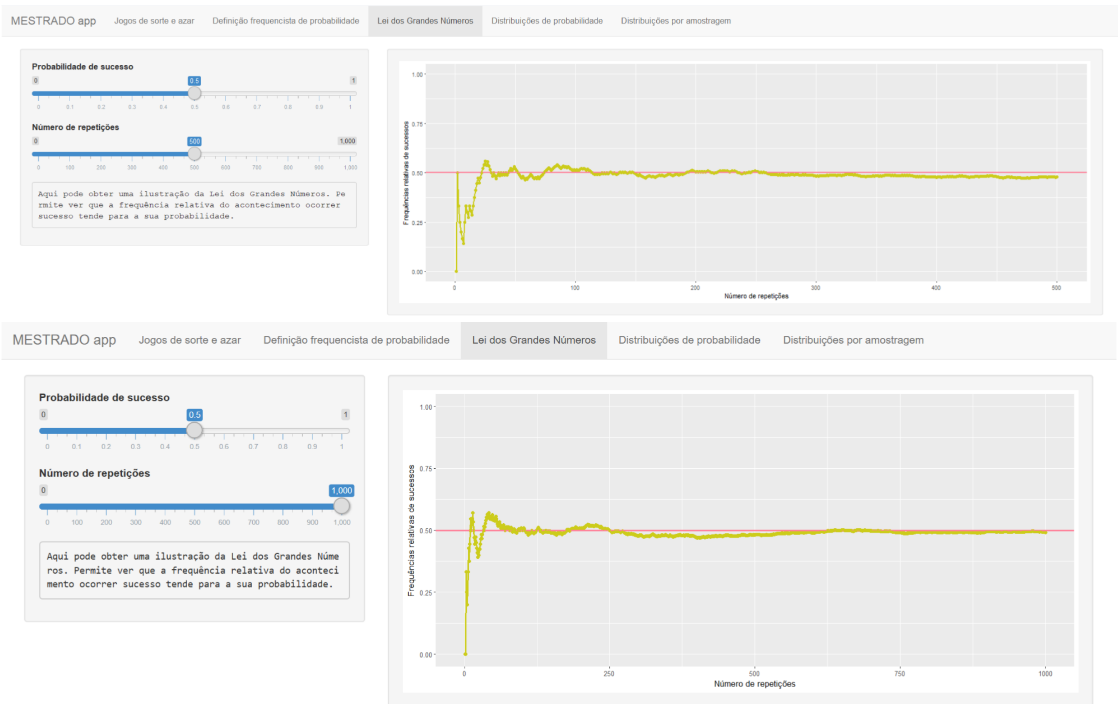

Figura 2.11: Simulação de 500 e 1000 lançamentos de uma moeda.

Como podemos constatar, prolongando bastante o número de lançamentos, o número médio de ocorrências de ocorrências "face cara" ao longo das realizações da experiência vai estabilizando à volta de um valor constante, neste caso 0.5 porque a moeda é equilibrada, que corresponde ao valor esperado de uma variável de Bernoulli, ou indicatriz como inicialmente referimos.

Pelo exemplo anterior pudemos ver que este separador para além de poder servir para ilustrar o Teorema 2.6, serve ainda para ilustrar a regularidade estatística associada aos fenómenos aleatórios. Muitos mais exemplos e exercícios podem ser explorados com os alunos a partir de experiências de Bernoulli com diferentes probabilidades de sucesso. Podemos, por exemplo, usar o gráfico deste separador para questionar os alunos sobre o fato de uma moeda ser equilibrada ou não.

2.2.3 Distribuições de probabilidade

O quarto separador da aplicação `Shiny`, Distribuições de probabilidade, como o próprio nome indica é dedicado às distribuições de probabilidade em $\mathbb{R}$. Neste separador é possível escolher de entre 5 distribuições discretas - Binomial, Geométrica, Hipergeométrica, Poisson, Uniforme - e sete distribuições contínuas - Exponencial, Fisher-Snedcor, Gama, Normal, qui-quadrado, t-Student, Uniforme. Os parâmetros destas distribuições podem ser escolhidos de entre um leque de valores possíveis, pelo utilizador. Após a escolha da distribuição e dos respetivos parâmetros a aplicação apresenta os gráficos da função de probabilidade (caso discreto) ou da função densidade de probabilidade (caso contínuo) e da função de distribuição.

Este separador permite deforma simples explorar as principais características das distribuições de probabilidade, características estas que vão desde a sua forma aos seus momentos, nomeadamente a média e o desvio padrão.

De seguida fazemos uma breve revisão das distribuições de probabilidade em $\mathbb{R}$ disponíveis no separador Distribuições de probabilidade e do conceito de função de distribuição, função de probabilidade e função densidade de probabilidade.

Definição 2.7. *A função de distribuição de uma variável aleatória real (v.a.r.) X é*

$$F(x) = P(X \leq x) = P(\omega : X(\omega) \in (-\infty, x]), \quad x \in \mathbb{R}.$$

Uma função de distribuição verifica as seguintes propriedades:

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

2. F(x) é uma função não decrescente;

3. $F(x)$ é contínua à direita para todo $x \in \mathbb{R}$.

É possível demonstrar que toda a função real com as propriedades anteriores é uma função de distribuição, ou seja dada uma função com as propriedades anteriores, existe sempre uma v.a.r. da qual ela é função de distribuição.

Teorema 2.8. *O conjunto de pontos de descontinuidade de uma função de distribuição, ou seja o conjunto dos pontos em que uma v.a.r. assume probabilidade positiva, é finito ou, quando muito, infinito numerável.*

O conjunto dos pontos de descontinuidade da função de distribuição permite classificar uma v.a.r. em discreta ou contínua.

Definição 2.9. *X é uma v.a.r. discreta se e só se existir um conjunto finito ou infinito numerável de pontos S, o suporte da variável aleatória, tal que*

$$P(X \in S) = \sum_{x\in S} P(X = x) = 1.$$

À distribuição de de uma v.a.r. discreta X chamamos função (massa) de probabilidade de X, denotamo-la por $f(x)$ e não é mais do que $P(X = x)$, $x \in S$.

A função de probabilidade tem as seguintes características:

1. $\forall x \in \mathbb{R}$, $f(x) \leq 0$;

2. $\sum_{x\in S} f(x) = 1$.

A função de distribuição de uma v.a.r. discreta é dada por

$$F(x) = \sum_{x\leq x} P(X = x') = \sum_{x\leq x} f(x'),$$

contínua, à direita para todo $x \in \mathbb{R}$.

À distribuição de uma v.a.r. contínua chamamos função densidade de probabilidade.

Definição 2.10. *Uma variável aleatória é (absolutamente) contínua se e só se existir uma função f, chamada função densidade (de probabilidade), tal que a*

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

Assim, a menos de um conjunto de medida nula, a função densidade de probabilidade é a derivada da função de distribuição.

As propriedades das funções densidade de probabilidade decorrem das propriedades das funções de distribuição. Se f for uma função densidade de probabilidade então:

1. $f(x) \geq 0$, para todo $x \in \mathbb{R}$;
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$.

Apresentamos de seguida a expressão analítica da função de probabilidade ou função densidade e da função de distribuição das distribuições incluídas no separador da aplicação. Aproveitamos também para apresentar os gráficos destas funções obtidos com a aplicação `Shiny`.

No que toca às distribuições discretas temos:

- **Distribuição uniforme discreta**

  Definição 2.11. *Dizemos que a v.a.r. discreta X de suporte $S = \{1, 2, ..., n\}$ segue uma distribuição uniforme discreta sobre $\{1, 2, ..., n\}, n \in \{N$, se a sua função de probabilidade é dada por*

  $$f(x) = P(X = x) = \frac{1}{n}\mathbb{I}_S(x).$$

  Simbolicamente, escrevemos $X \sim U\{1, 2, .., n\}$.

  Se X é uma v.a.r. discreta com distribuição uniforme tem-se que

  $$E[X] = \frac{n+1}{2}$$

  e

  $$Var[X] = \frac{n^2-1}{12}.$$

A sua função de distribuição é dada por

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \quad x < 1 \\ \sum_{x_i \leq} \frac{x_i}{n} & , \quad x_i \leq x < x_i + 1, \quad x_i = 1, 2, 3, ..., n-1 \\ 1 & , \quad x \geq n. \end{cases}$$

Ilustra-se na Figura 2.12 a função de probabilidade e a função de distribuição de uma v.a.r. uniforme discreta de suporte $S = \{1, 2, 3, 4\}$.

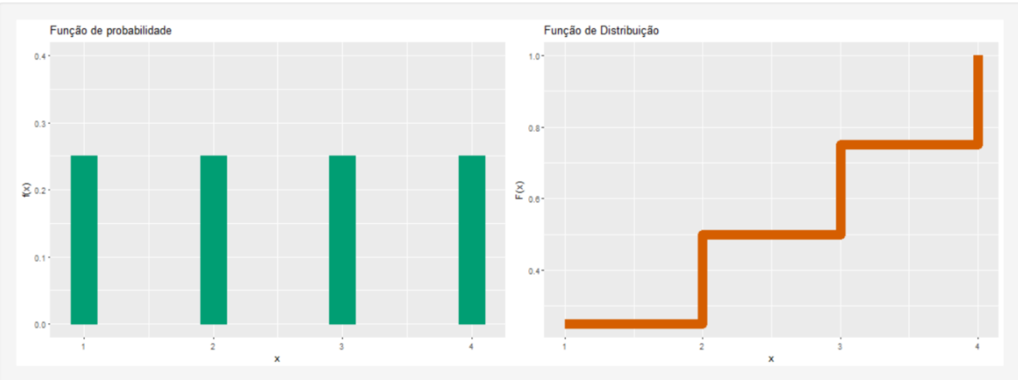

Figura 2.12: Função de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r $X \sim U\{1, 2, 3, 4\}$.

- **Distribuição de Bernoulli. Distribuição binomial**

  A distribuição de Bernoulli está associada à realização de uma experiência aleatória com dois resultados possíveis, denominados sucesso e insucesso, com probabilidade de sucesso p.

  Definição 2.12. *Dizemos que uma v.a.r. discreta X de suporte S={0,1}, segue uma distribuição de Bernoulli de parâmetro p, $p \in]0, 1[$, quando a sua função de probabilidade é*

  $$f(x) = P(X = x) = p^x (1-p)^{1-x} \mathbb{1}_S(x)$$

  e escrevemos $X \sim Bernoulli(p)$.

  A função de distribuição de X é dada por

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \quad x < 0 \\ 1-p & , \quad 0 \leq x < 1 \\ 1 & , \quad x \geq 1 \end{cases}$$

e tem-se

$$E[X] = p$$

e

$$Var[X] = p(1-p).$$

Quando se realizam n provas de *Bernoulli* independentes, estamos em presença de uma experiência binomial. Na experiência aleatória binomial interessa-nos a variável aleatória X que representa o número total de "sucessos" ocorridos nas n provas.

Definição 2.13. *Uma v.a.r. X de suporte $S = \{1, 2, ..., n\}$, número de sucessos em n provas de Bernoulli, segue uma distribuição binomial de parâmetros n e p, $n \in \mathbb{N}$, $p \in]0, 1[$, se a sua função de probabilidade é dada por*

$$f(x) = P(X = x) =^n C_x p^x (1-p)^{n-x} \mathbb{1}_S(x),$$

onde ${}^nC_x$ denota as combinações de n, x a x.

Simbolicamente, escrevemos $X \sim B(n, p)$.

A função de distribuição correspondente é dada por

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \quad x < 0 \\ \displaystyle\sum_{x_i=0}^{x} {}^nC_x p^{x_i} (1-p)^{n-x_i} & , \quad 0 \leq x \leq n, \\ 1 & , \quad x > n. \end{cases}$$

Apresentamos na Figura 2.13 o exemplo de uma distribuição binomial com parâmetros $n = 5$ e $p = 0.5$.

Notemos que a distribuição binomial de parâmetros 1 e p não é mais do que a distribuição de Bernoulli de parâmetro p, assim na nossa aplicação a distribuição de Bernoulli está definida com a distribuição binomial.

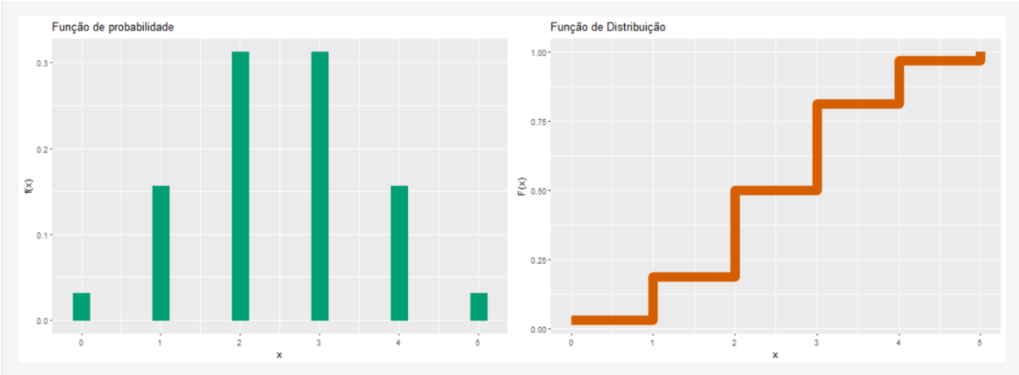

Figura 2.13: Função de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim B(5, 0.5)$.

- **Distribuição geométrica**

  A distribuição geométrica está associada a experiências aleatórias em que estamos interessados no número de provas de Bernoulli independentes a realizar até à obtenção do primeiro sucesso.

  Definição 2.14. *Dizemos que uma v.a.r. X de suporte $S = \{1, 2, ..., n\}$, número de provas de Bernoulli independnetes até à obtenção do primeiro sucesso, segue uma distribuição geométrica de parâmetro p, $p \in]0, 1[$, se tem função de probabilidade*

$$f(x) = P(X = x) = (1 - p)^{x-p}p.$$

  Simbolicamente, escrevemos $X \sim G(p)$.

  Se $X \sim G(p)$ então

$$E[X] = \frac{1}{p},$$
$$Var[X] = \frac{1-p}{p^2},$$

  e

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \quad x < 1 \\ 1 - (1-p)^{\lfloor x \rfloor} & , \quad x \geq 1, \end{cases}$$

  onde $\lfloor x \rfloor$ denota a parte inteira de x.

Na Figura 2.14, estão representados os gráficos da função de probabilidade e da função de distribuição de uma distribuição geométrica de parâmetro $p = 0.6$.

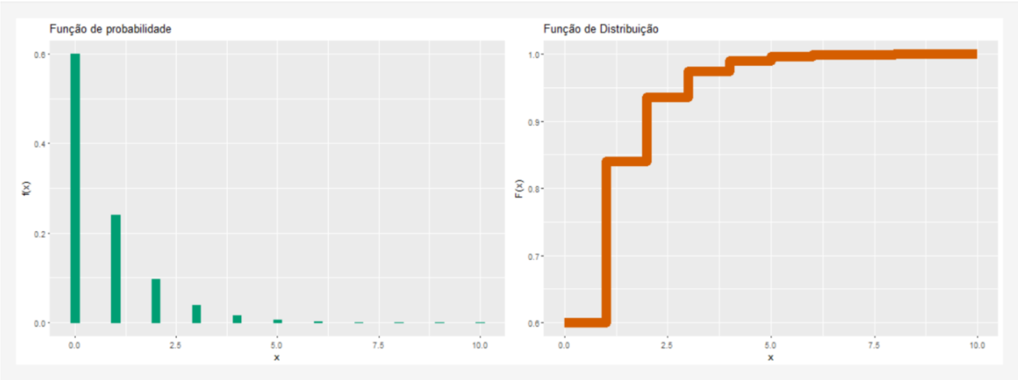

Figura 2.14: Função de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim G(0.6)$.

- **Distribuição hipergeométrica**

  A distribuição hipergeométrica é frequentemente utilizada quando se conta o número de elementos de um conjunto finito constituído por objetos com duas características diferentes (por exemplo, defeituosos e não defeituosos).

  Consideremos uma população com um número finito de N elementos dos quais K possuem uma característica C e $N - K$ do tipo $\overline{C}$. Escolhe-se, ao acaso, e sem reposição, n dos N elementos, com $n \leq N$. A v.a.r. X que representa o número de objetos escolhidos do tipo C tem uma distribuição hipergeométrica com, parâmetros N (tamanho da população), K (tamanho da amostra) e n (número de elementos com a característica pretendida, existente na população).

  Esta distribuição pode ser vista como uma binomial, pelo facto de representar o número de sucessos em n tentativas, cada uma com dois resultados possíveis, sucesso e insucesso. No entanto, utiliza-se a distribuição hipergeométrica quando o tamanho da população é finito e, as tentativas são dependentes com probabilidade de sucesso diferente a cada tentativa realizada.

Definição 2.15. *A v.a.r. X, número de sucessos contidos numa amostra de tamanho n retirada ao acaso, sem reposição, de uma população de N*

elementos, dos quais K $(K \leq N)$ constituem o sucesso, segue uma distribuição hipergeométrica de parâmetros N, K e n, e tem função de probabilidade dada por

$$f(x) = \frac{{}^{K}C_x \times^{N-K} C_{n-x}}{{}^{N}C_n}, \;\; x = \max\{0, n + K - N\}, ..., \min K, n.$$

Simbolicamente, escrevemos $X \sim H(N, K, n)$.

A distribuição hipergeométrica tem como função de distribuição

$$F(x) = P(X \leq x) = \begin{cases} 0 & , \;\; x < \max\{0, n - (N - K)\} \\ \sum_{i=0}^{\lfloor x \rfloor} \frac{{}^{K}C_i \times {}^{N-K}C_{n-i}}{{}^{N}C_n} & , \;\; \max 0, n - (N - K) \leq x \leq \min\{K, n\} \\ 1 & , \;\; x \geq \min\{K, n\}. \end{cases}$$

Com parâmetros $N = 20$, $K = 10$ e $n = 5$, ilustra-se na Figura 2.15 os gráficos da função de probabilidade e da função de distribuição da distribuição hipergeométrica.

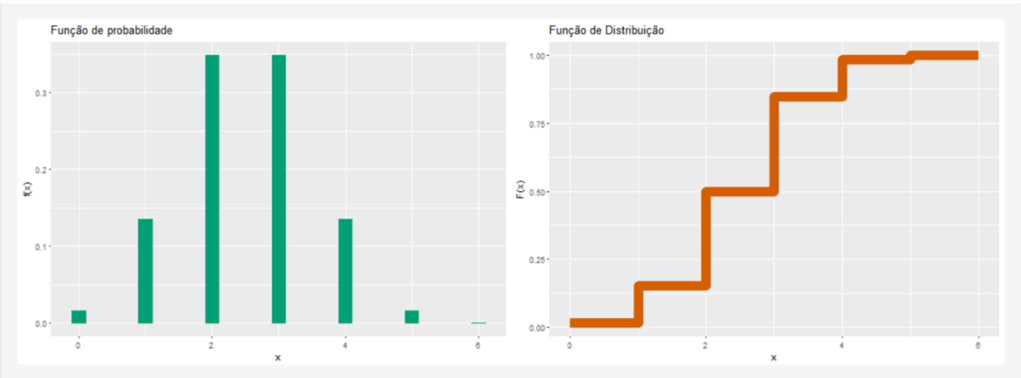

Figura 2.15: Função de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim Hipergeomtrica(20, 10, 5)$.

O valor esperado e a variância de uma v.a.r X com distribuição hipergeométrica de parâmetros N, K e n são dados por

$$E[X] = np, \;\; p = \frac{K}{N}$$

e

$$Var[X] = np(1-p)\frac{N-n}{N-1},\ p = \frac{K}{N}.$$

- **Distribuição Poisson**

  A distribuição de Poisson (nome dado em homenagem ao físico francês Simon Poisson (1781-1840)) está associada à contagem do número de acontecimentos que ocorrem num intervalo de tempo ou numa região espacial, sendo estes eventos independentes uns dos outros.

Definição 2.16. *Seja λ, $\lambda \in \mathbb{R}^+$, o número médio de acontecimentos que ocorrem num determinado intervalo de tempo, ou numa região espacial, diz-se que a v.a.r. X de suporte $\mathbb{N}_0$, segue uma distribuição de Poisson com parâmetro λ, quando a sua função de probabilidade é dada por*

$$f(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}\mathbb{I}_{\mathbb{N}_0}(x).$$

Simbolicamente, escrevemos $X \sim P(\lambda)$.

A expressão abaixo representa a função de distribuição associada a esta distribuição

$$F(x) = P(X \leq x) = \begin{cases} 0 & ,\quad x < 0 \\ e^{-\lambda}\displaystyle\sum_{i=0}^{\lfloor x \rfloor} \frac{\lambda^i}{i!} & ,\quad x \geq 0. \end{cases}$$

A Figura 2.16 ilustra as funções de probabilidade e de distribuição de uma distribuição de Poisson de parâmetro $\lambda = 4$.

Quando $X \sim P(\lambda)$ então

$$E[X] = Var[X] = \lambda.$$

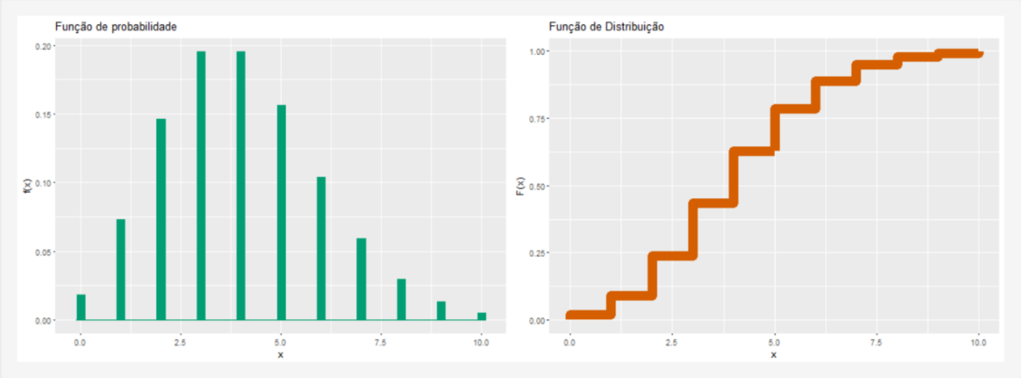

Figura 2.16: Função de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim P(4)$.

Em relação as distribuições continuas, apresentamos as seguintes:

- **Uniforme contínua**

  A distribuição uniforme é utilizada para representar uma quantidade que varia aleatoriamente num intervalo $[a, b]$ com probabilidade constante para quaisquer dois sub-intervalos de [a,b] de igual amplitude.

  Definição 2.17. *A v.a.r. contínua X segue uma distribuição Uniforme no intervalo $[a, b]$, $(a < b)$, se a sua função densidade de probabilidade se apresenta na forma*

  $$f(x) = \frac{1}{b-a} \mathbb{I}_{[a,b]}(x).$$

  Simbolicamente, escrevemos $X \sim U[a, b]$.

  Consequentemente, a sua função de distribuição é dada por

  $$F(x) = \begin{cases} 0 & , \quad x < a \\ \frac{x-a}{b-a} & , \quad a \leq x < b \\ 1 & , \quad x \geq b. \end{cases}$$

  Na Figura 2.17 podemos observar a função densidade de probabilidade e a função de distribuição de uma v.a.r. com distribuição uniforme contínua no intervalo [2,7].

  Para este tipo de distribuição temos que

  $$E[X] = \frac{a+b}{2}$$

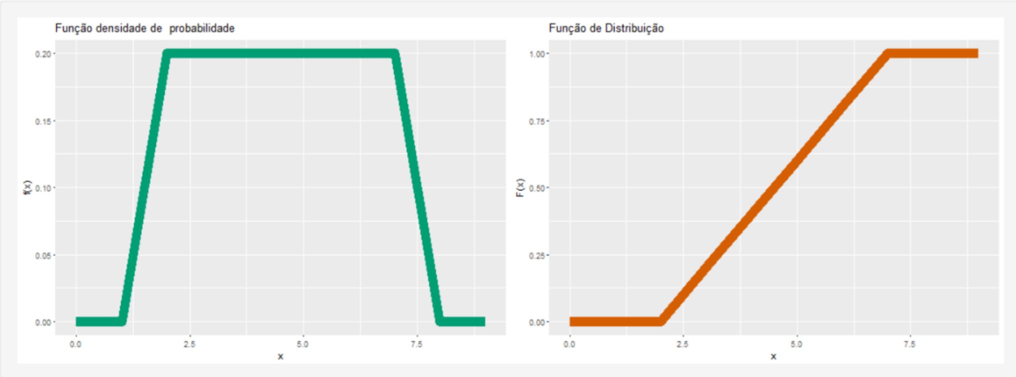

Figura 2.17: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim U[2,7]$.

e

$$Var[X] = \frac{(b-a)^2}{12}.$$

- **Normal**

É uma das distribuições mais utilizadas em Probabilidades e Estatística, pelo facto de inúmeras variáveis aleatórias que descrevem fenómenos, processos físicos ou características humanas seguirem este tipo de distribuição. Outro factor que coloca esta distribuição no grupo das mais importantes, é o facto de que algumas variáveis aleatórias puderem ser aproximadas por uma variável aleatória Normal.

Definição 2.18. *Uma v.a.r. contínua X segue uma distribuição normal com parâmetros μ e σ se a sua função densidade de probabilidade é*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}, \; x \in \mathbb{R}, \; \mu \in \mathbb{R}, \; \sigma \in \mathbb{R}^+.$$

Simbolicamente, escrevemos $X \sim N(\mu, \sigma)$.

A função de distribuição Normal é dada por

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} dt, \; x \in \mathbb{R}.$$

Através da Figura 2.18 observa-se que a função de probabilidade Normal tem a forma de um sino e é simétrica em relação a $x = \mu$.

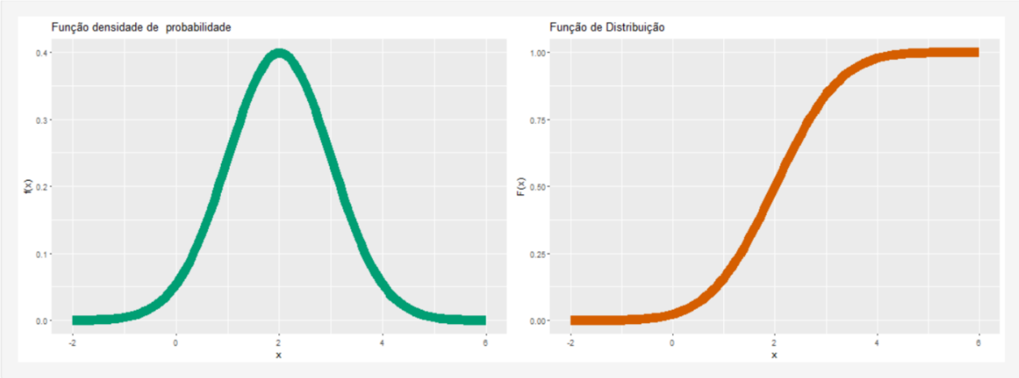

Figura 2.18: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim N(2,1)$.

Os parâmetros μ e σ representam, respetivamente, o valor esperado e o desvio padrão.

Pelo facto de μ e σ puderem assumir uma infinidade não numerável de valores ($\mu \in \mathbb{R}$ e $\sigma \in \mathbb{R}^+$), existe também uma infinidade não numerável de distribuições normais. Porém, para o cálculo de probabilidades, qualquer distribuição normal pode ser transformada em uma *normal padrão*, ou seja uma normal com parâmetros constantes $\mu = 0$ e $\sigma = 1$. Esta transformação recebe o nome de *padronização* e obtém-se subtraindo μ à variável e dividindo por σ,

$$Z = \frac{X - \mu}{\sigma}.$$

- **Exponencial**

A distribuição exponencial é um tipo de distribuição que se utiliza para representar o intervalo de tempo entre a ocorrência de dois acontecimentos.

Definição 2.19. *Dizemos que uma v.a.r. X, segue uma distribuição exponencial com parâmetro λ, $\lambda > 0$, quando a sua função densidade de probabilidade é da forma*

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}^+}(x).$$

Simbolicamente, escrevemos $X \sim Exp(\lambda)$.

Se X é uma v.a.r. tal que $X \sim Exp(\lambda)$, então

$$E[X] = \frac{1}{\lambda}, Var[X] = \frac{1}{\lambda^2}$$

e a sua função de distribuição vem

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

Na Figura 2.19 ilustra-se a função densidade e a função de distribuição de uma distribuição exponencial com parâmetro $\lambda = 1$.

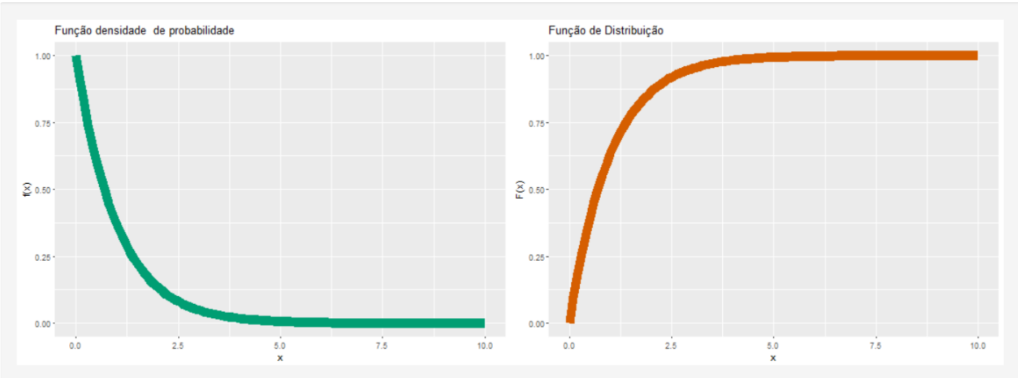

Figura 2.19: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a. $X \sim Exp(1)$.

- **Gama**

  A distribuição gama de parâmetros $\alpha > 0$ e $\lambda > 0$ pode ser vista como a distribuição da soma de α variáveis aleatórias exponenciais independentes e todas com o mesmo parâmetro λ.

  Definição 2.20. *A v.a.r. X segue uma distribuição gama de parâmetro α e λ, $\alpha > 0$, $\lambda > 0$ se a sua função densidade de probabilidade é da forma*

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} e^{-\lambda x} x^{\alpha - 1} \mathbb{I}_{\mathbf{R}^+}(x),$$

  onde $\Gamma(\alpha)$, é a função gama de Euler, i.e

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx.$$

Simbolicamente, escrevemos $X \sim Gama(\alpha, \lambda)$.

A função de distribuição da distribuição gama de parâmetros α e λ é dada por

$$F(x) = \int_0^x f(u)du = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)},$$

onde $\gamma(\alpha, \beta x)$ é a função gama incompleta.

Se $X \sim Gama(\alpha, \lambda)$ temos $E[X] = \frac{\alpha}{\lambda}$ e $Var[X] = \frac{\alpha}{\lambda^2}$.

A função densidade de probabilidade e a função de distribuição de uma distribuição gama com parâmetros $\alpha = 4$ e $\lambda = 1$, têm a forma ilustrada na Figura 2.20.

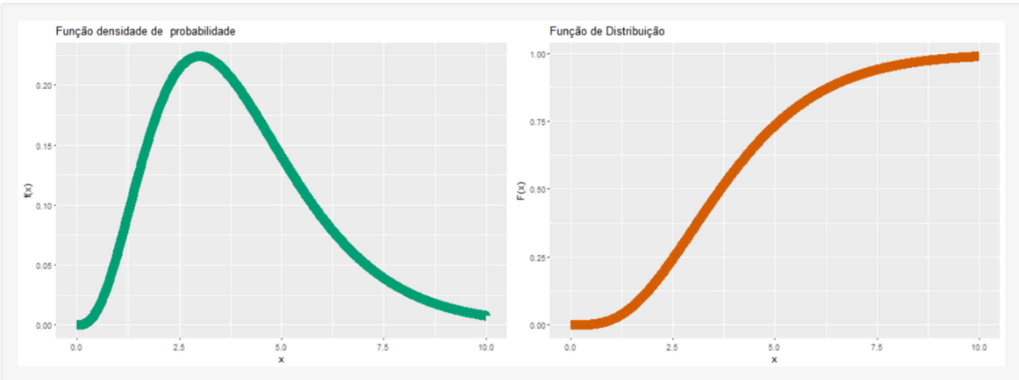

Figura 2.20: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $X \sim Gama(4, 1)$.

Se considerarmos $\alpha = 1$ na distribuição gama de parâmetros α e λ a distribuição obtida será exponencial de parâmetro $\lambda > 0$, isto é, $X \sim Exp(\lambda)$.

Se $\alpha = \frac{n}{2}$, $n \in \mathbb{N}$, e $\lambda = \frac{1}{2}$, a distribuição gama coincide com a distribuição qui-quadrado que é definida a seguir.

- **Qui-quadrado**

A soma de variáveis aleatórias normais padrão ao quadrado tem uma distribuição qui-quadrado.

Definição 2.21. *Dizemos que a variável aleatória* X*, segue uma distribuição qui-quadrado com* n *graus de liberdade se tem função densidade de* *probabilidade dada por*

$$f(x) = \frac{1}{2\Gamma(\frac{n}{2})} \left(\frac{x}{2}\right)^{\frac{n}{2}-1} e^{-\frac{x}{2}} \mathbb{I}_{]0,+\infty[}(x),$$

onde Γ *é a função gama de Euler.*

Simbolicamente, escrevemos $Y \sim \chi^2_n$.

Se $X \sim \chi^2_n$ tem-se $E[X] = n$ e $Var[X] = 2n$.

Se n for um número par positivo, a função de distribuição tem a forma

$$F(x) = 1 - e^{-\frac{x}{2}} \sum_{i=0}^{\frac{n}{2}-1} \frac{1}{i!} \left(\frac{x}{2}\right)^i, x > 0.$$

Para valores de n ímpares, a função de distribuição não está definida.

Na Figura 2.21 ilustra-se a função de densidade de probabilidade e a função de distribuição da distribuição qui-quadrado com 5 graus de liberdade.

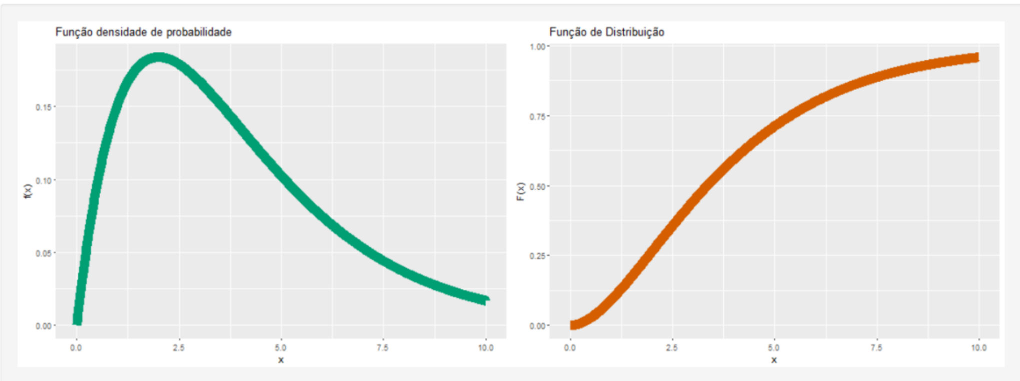

Figura 2.21: Função densidade de probabilidade (esquerda) e função de distribuição (direita) da distribuição qui-quadrado com 5 graus de liberdade.

- **t-Student**

  O nome desta distribuição deve-se aos trabalhos do estatístico inglês William Sealy Gosset "Student". Tal como a distribuição qui-quadrado, a distribuição t-Student possui um único parâmetro n, número de graus de liberdade, e a sua origem decorre da relação entre as distribuições Normal e qui-quadrado. Uma vez que, se $Z \sim N(0,1)$ e $Y \sim X_n^2$ são v.a.r. independentes, então $\frac{Z}{\sqrt{\frac{Y}{n}}}$ segue uma distribuição t-Student com n graus de liberdade.

Definição 2.22. *Dizemos que a v.a.r. X segue uma distribuição t-Student com n graus de liberdade se a sua função densidade é da forma*

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \quad \Gamma\left(\frac{n}{2}\right)} \left(\frac{1+x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in \mathbb{R},$$

Simbolicamente, escrevemos $X \sim t_n$

Na Figura 2.22 podemos ver os gráficos da função densidade de probabilidade e da função de distribuição da distribuição t-Student com $n = 4$ graus de liberdade.

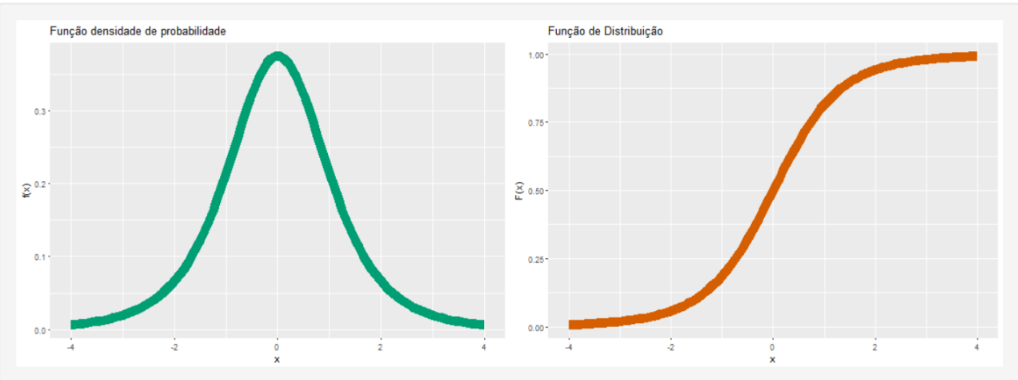

Figura 2.22: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $Y \sim t_4$.

Quando n tende para infinito, a função densidade de probabilidade da distribuição t-Student tende para a distribuição normal padrão.

Para $n = 1$ o valor esperado da distribuição t-Student não existe, contudo, se $n > 1$, $E[X] = 0$. Quanto a variância, não existe para valores de $n \leq 2$, mas se $n > 2$, $Var[X] = \frac{n}{n-2}$.

- **Fisher - Snedecor**

  O nome desta distribuição foi atribuído em honra ao estatístico Ronal Fisher. É usada em estatística inferencial quando se comparam variâncias populacionais.

Definição 2.23. *Sejam* $X \sim X_m^2$ *e* $Y \sim X_n^2$ *duas variáveis Qui-quadrado independentes com* m *e* n *graus de liberdade, a variável aleatória* $W = \frac{X/m}{Y/n}$ *segue uma distribuição de Fisher-Snedecor com* m *e* n *graus de liberdade, e tem como função de densidade*

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{x^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}x\right)^{\frac{m+n}{2}}} \mathbb{1}_{]0,+\infty[}(x).$$

Simbolicamente, escrevemos $W \sim F_{m,n}$.

Na Figura 2.23 vemos os gráficos da função densidade de probabilidade e da função de distribuição da distribuição Fisher-Snedecor com 5 e 4 graus de liberdade.

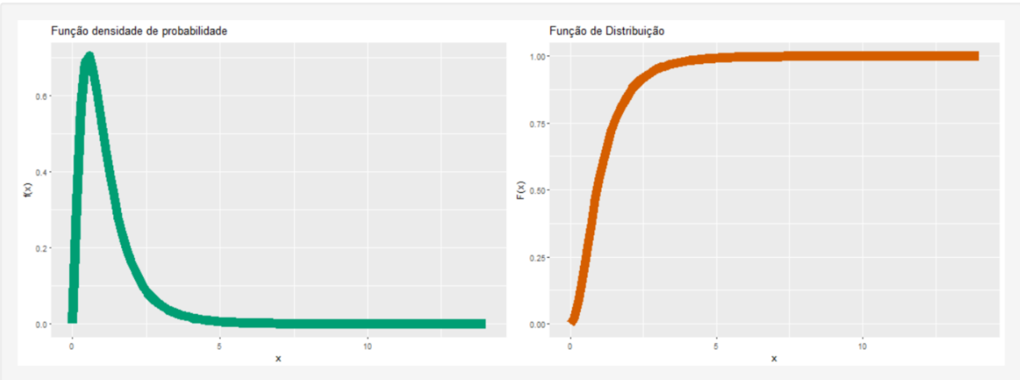

Figura 2.23: Função densidade de probabilidade (esquerda) e função de distribuição (direita) de uma v.a.r. $W \sim F_{5,4}$.

O valor esperado da distribuição de Fisher-Snedecor não está definido para valores de $n \leq 2$, no entanto, se $n > 2$ $E[X] = \frac{n}{n-2}$.

Para valores de $n \leq 4$ a variância da distribuição de Fisher-Snedecor também não está definida. Para $n > 4$ tem-se $Var[X] = \frac{2n^2(m+n-2)}{n(n-2)^2(n-4)}$.

Feita a caracterização das distribuições de probabilidade disponíveis na aplicação, podemos agora explorar um pouco mais as funcionalidades do separador Distribuições de probabilidade, que poderão ajudar os alunos na compreensão das principais características de algumas distribuições. Restringimo-nos às distribuições Normal e Binomial pelo fato de serem as distribuições mais abordadas no ensino básico e secundário.

Como a aplicação permite obter representações gráficas das distribuições de probabilidade de forma simples e rápida, podemos pedir aos alunos que obtenham os gráficos para três distribuições normais diferentes, como as da Figura 2.24, e que indiquem quais as diferenças que observam.

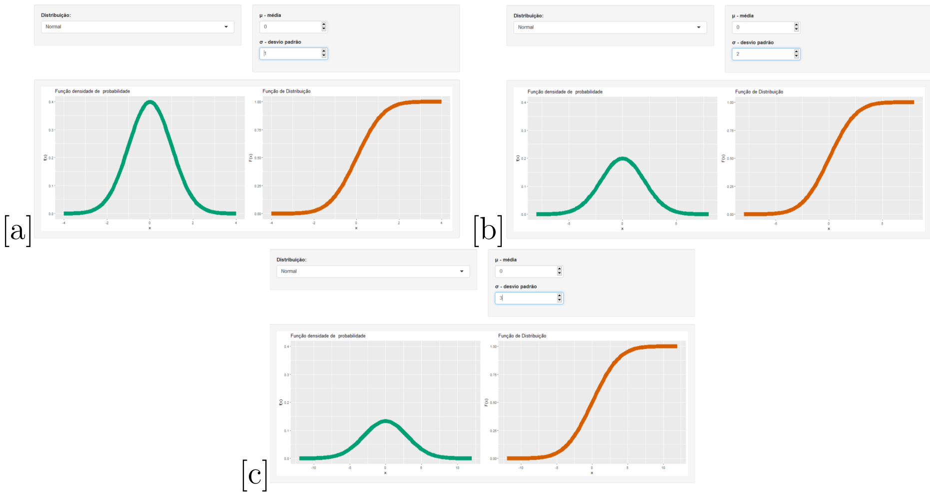

Figura 2.24: Função densidade (esquerda) e função de distribuição (direita) de uma v.a.r. [a] - $X \sim N(0,1)$,[b] - $X \sim N(0,2)$,[c] - $X \sim N(0,3)$.

Como as distribuições normais da Figura 2.24 têm a mesma média e desvios padrão diferentes, ao compará-las, `[a]`, `[b]` e `[c]`, é possível concluir que em numa distribuição normal, quanto maior for o desvio padrão maior será a variabilidade da variável em torno da média e, por esse motivo o sino da distribuição `[c]` é mais achatado que a `[b]`, que por sua vez também é mais achatado que a distribuição `[a]`. Tal permite ainda perceber que σ é uma parâmetro de dispersão ou de escala.

Pedindo de seguida ao aluno para considerar três distribuições normais com médias diferentes e igual desvio-padrão, como na Figura 2.25, ele poderá constatar que o achamento da distribuição não se altera, porém, vai se movimentando ao longo do das abcissas. Deste modo, o aluno perceberá que o parâmetro μ é uma parâmetro de localização da distribuição e que a distribuição é simétrica em torno do eixo vertical por ele definido.

A partir de distribuições binomiais com parâmetros distintos, como na Figura 2.26, o aluno perceberá que o valor do parâmetro p altera a simetria da distribuição, sendo está tão mais simétrica quanto mais próximo este parâmetro estiver de 0.5. Mais concretamente, nos gráficos [a] e [b] poderá observar que, dependendo

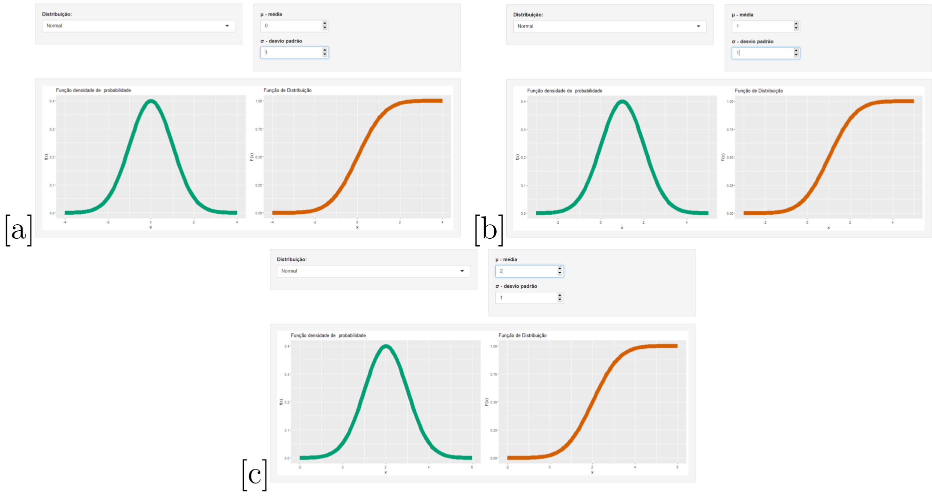

Figura 2.25: Função densidade (esquerda) e função de distribuição (direita) de uma v.a.r. [a] - $X \sim N(0,1)$, [b] - $X \sim N(1,1)$, [c] - $X \sim N(2,1)$.

do valor do parâmetro p, a distribuição pode ser assimétrica positiva (enviesada à esquerda) ou assimétrica negativa (enviesada à direita). No caso de [a] em que p é menor que 0.5, a distribuição é assimétrica positiva, enquanto que para p maior que 0.5 a distribuição é assimétrica negativa como ilustra o gráfico [b]. Quanto mais afastado estiver de 0.5 (quer seja a esquerda ou a direita), mais enviesada será a distribuição, que se pode confirmar com mais gráficos. A partir de [c] e [d] o aluno poderá ainda observar que, qualquer que seja o valor do parâmetro n, quando p for igual a 0.5, a distribuição binomial será sempre simétrica.

Os gráficos que resultam de manter o valor do parâmetro p constante e aumentar o valor do parâmetro n, permitirão aos alunos perceber que a função de probabilidade binomial começa a ser parecida com a distribuição normal, verificando mesmo que a parecença é tanto melhor quanto mais próximo p estiver de 0.5. Este exemplo simples poderá servir de introdução ao Teorema Limite Central que abordaremos na secção seguinte.

2.2.4 Distribuições por amostragem

Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de uma população X com distribuição G. As v.a.r $X_1, \ldots, X_n$ são assim independentes e identicamente distribuídas (i.i.d.) com distribuição G, e as funções obtidas a partir delas são denominadas estatísticas. Destacamos a estatística mais importante na estatística inferencial, a média amostral $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. O conhecimento da distribuição desta estatís-

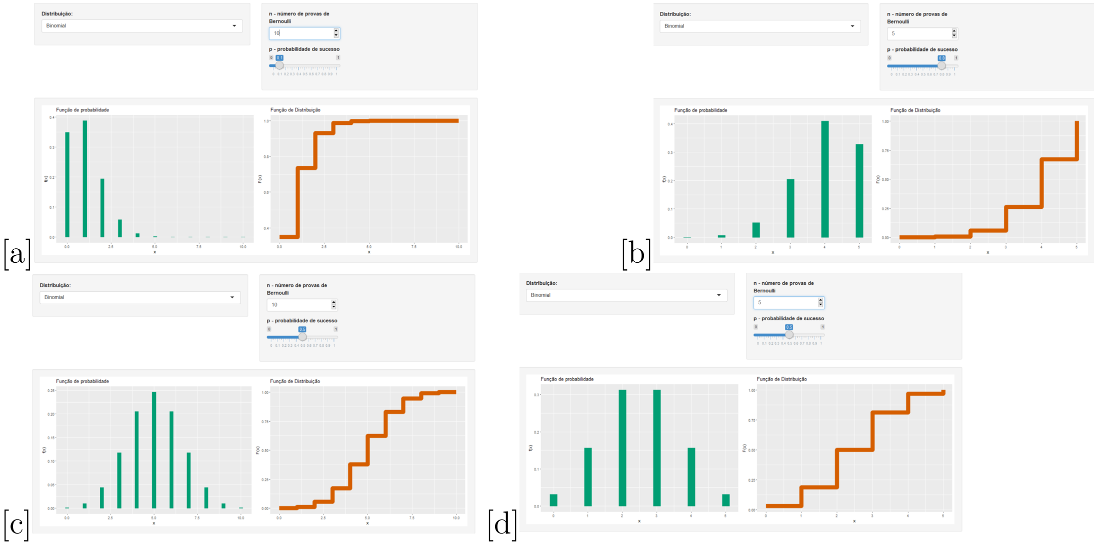

Figura 2.26: Função densidade (esquerda) e função de distribuição (direita) de uma v.a.r. [a] - $X \sim B(10, 0.1)$, [b] - $X \sim B(5, 0.8)$, [c] - $X \sim B(10, 0.5)$, [d] - $X \sim B(5, 0.5)$.

tica, denominada distribuição por amostragem ou distribuição amostral, é assim de extrema importância.

A distribuição exata da média amostral, $\overline{X}_n$, é conhecida quando G coincide com algumas das distribuições abordadas anteriormente. A sua obtenção resulta sobretudo da propriedade de estabilidade para somas que algumas distribuições têm, como a Normal. Tendo em conta que na maioria dos estudos as amostras são de grande dimensão, o que em geral mais interessa não é a distribuição exata mas sim a distribuição assintótica, quando $n \to +\infty$. O resultado fundamental neste contexto é o Teorema Limite Central, que a seguir recordamos.

Teorema 2.24. *Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de uma população X, com valor esperado $E[X] = \mu$ e variância $Var(X) = \sigma^2$. Então, a sucessão da v.a.r. de termo geral $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$ converge em distribuição, quando $n \to +\infty$, para uma variável aleatória de lei $N(0, 1)$, isto é,*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to+\infty]{d} Z \sim N(0,1)$$

onde $S_n = \sum_{i=1}^{n} X_i$, ou equivalentemente

$$\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n\to+\infty]{d} Z \sim N(0,1),$$

onde $\xrightarrow[n\to+\infty]{d}$ *denota a convergencia em distribuição ou em lei.*

O Teorema Limite Central, um dos mais célebres resultados da Teoria da Probabilidade, diz-nos que, para n suficientemente grande, a lei da soma S_n (média $\overline{X}_n$) é bem aproximada pela mesma lei normal, independentemente do tipo de lei comum às variáveis da amostra aleatória.

Apesar da importância e grande aplicabilidade do Teorema Limite Central ele continua a constituir um ponto fraca da maioria dos alunos. A justificação para isto advem sobretudo da dificuldade que têm em compreender as distribuições por amostragem. Diversos estudos têm analisado estas dificuldades por parte dos alunos, (ver Chance *et al.* (2004), Jacob e Doerr (2014), Ko (2016) e referências aí encontradas).

Gordon e Gordon (2009) referem que se o Teorema Limite Central for meramente apresentado aos alunos sem alguma corroboração, ele terá muito pouco impacto neles e eles não desenvolverão qualquer entendimento sobre ele. Por sua vez, no estudo de Ko (2016) é referido que a compreensão das distribuições por amostragem se desenvolve ao longo dos seguintes 5 passos:

1. os alunos não percebem a variabilidade amostral;

2. os alunos percebem a variabilidade amostral, mas confundem os dados numa amostra e as estatísticas como dados;

3. os alunos prestam atenção à dispersão dos valores da estatística mas não à sua tendência central;

4. prestam atenção à tendência central e à dispersão dos valores da estatística;

5. prestam atenção à tendência central e à dispersão dos valores da estatística e compreendem a relação entre o tamanho da amostra e a variabilidade amostral.

O uso de simulações gráficas apropriadas, por sua vez, permite aos alunos verem os resultados de repetidas amostras de uma população e a correspondente distribuição das médias amostrais e portanto aumenta a sua compreensão dos conceitos e métodos estatísticos (Gordon e Gordon (2009)). Esta foi a motivação para as simulações incluídas no separador Distribuições por amostragem.

No separador Distribuições por amostragem podemos simular dados de entre cinco populações, Binomial, Exponencial, Gama, Poisson, Normal, e ver a sua representação gráfica. Suponhamos que escolhemos uma população exponencial de parâmetro 4. O histograma desta população é exibido como podemos ver pela Figura 2.27, bem como a média e o desvio padrão, $\mu = 0,25$ e $\sigma = 0,0625$.

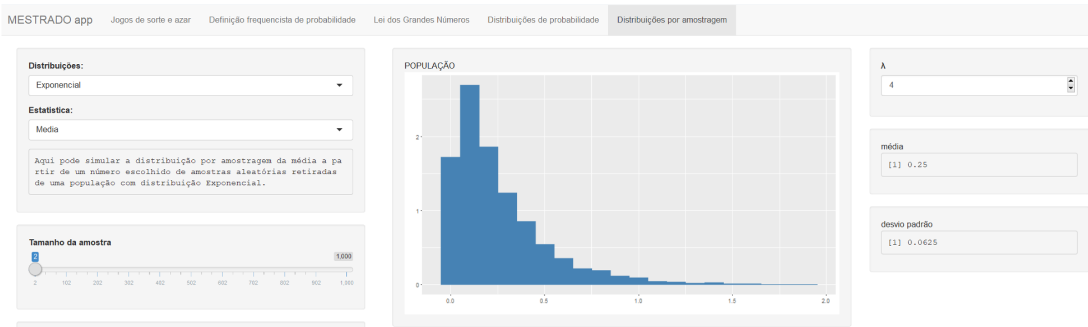

Figura 2.27: População exponencial de parâmetro 4.

As cinco populações que se podem escolher são bastante distintas, desde as simétricas às mais assimétricas. De seguida podemos escolher a estatística de entre duas disponíveis, média e soma, a dimensão e o número de amostras aleatórias simples a serem extraídas da população. Na Figura 2.28 podemos ver a distribuição por amostragem da média para 5000 amostras de dimensões 5, 20 e 500, respetivamente.

Pode-se constatar na Figura 2.28 que a forma da distribuição da média claramente não é normal para amostras de dimensão muito pequena. No entanto, quando o tamanho da amostra aumenta, a forma parece-se mais com a da normal, ilustrando-se assim o Teorema Limite Central. Pode-se aproveitar aqui para explorar com os alunos as relações entre os parâmetros da população e os da estatística média, uma vez que a aplicação disponibiliza a média e o desvio padrão das médias das várias amostras. Podemos ver pela Figura 2.28, para as amostras de dimensão 500, que a média das médias amostrais é aproximadamente igual a 0.25, o valor médio da população, e o desvio padrão aproximadamente igual a 0.011, ou seja, ao desvio padrão da população a dividir por $\sqrt{n}$, onde n é a dimensão da amostra. Tal permite aos alunos verificar duas importantes carac-

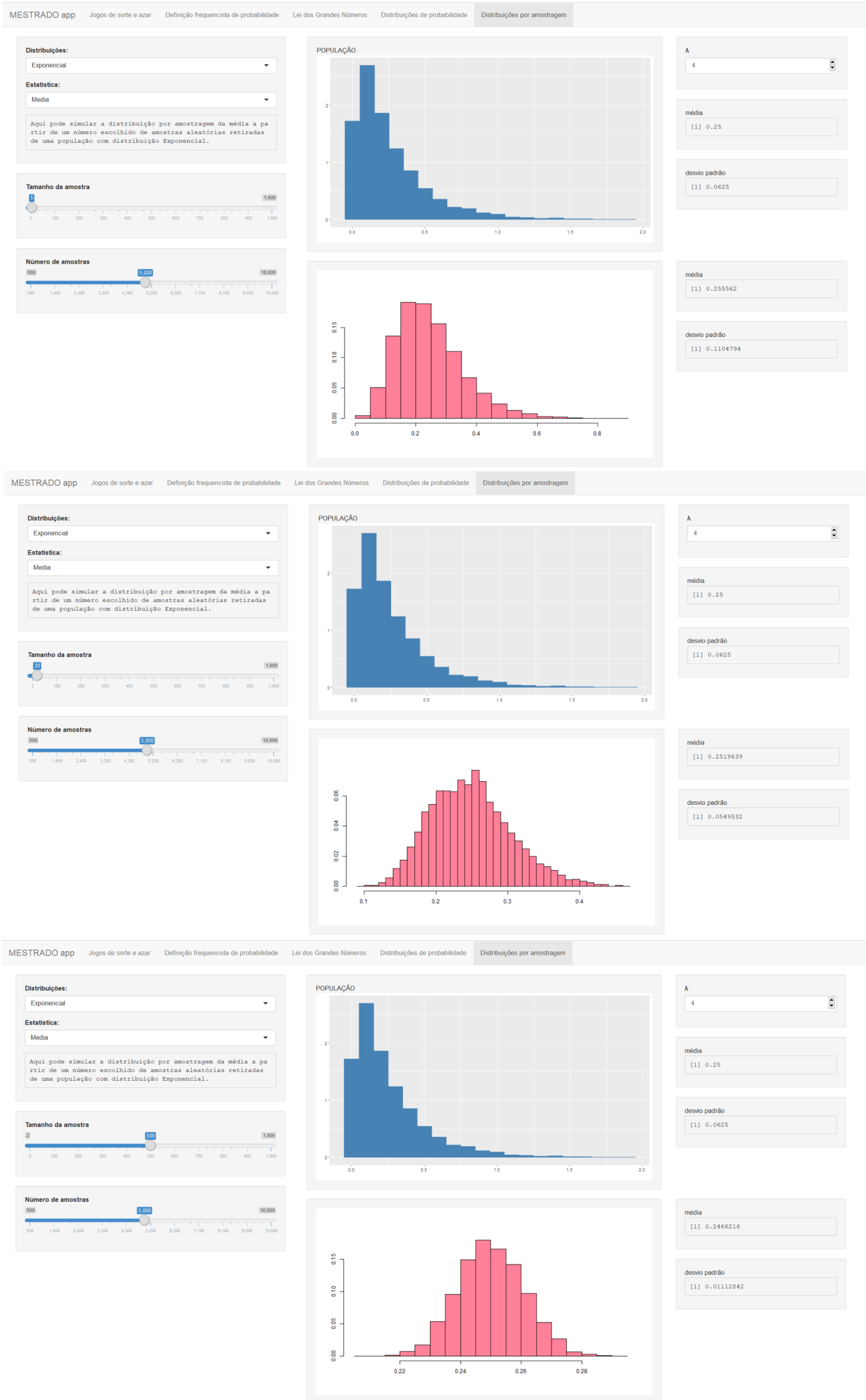

Figura 2.28: Distribuição por amostragem da média, com uma população subjacente Exponencial de parâmetro 4 e amostras de dimensão 5, 20 e 500, respetivamente.

terísticas da distribuição por amostragem da média $\overline{X}_n$, que são $E[\overline{X}_n] = \mu$ e $Var[\overline{X}_n] = \sigma^2/n$, ou eqivalentemente, o desvio padrão de $\overline{X}_n$ é igual a $\sigma/\sqrt{n}$, com μ a média da população e σ o desvio padrão da população.

Conclusão

Neste trabalho, desenvolvemos uma aplicação com o pacote `Shiny` do `Rstudio`, interface do *software* `R`, para o ensino de alguns temas de probabilidades e estatística abordados desde o ensino secundário ao ensino superior. Esta aplicação permite verificar no processo de ensino e aprendizagem que a teoria concorda com a realidade, tornado este processo a aprendizagem mais interessante e efetiva.

Estamos conscientes que existem muitas aplicações, facilmente encontradas na internet, com simulações semelhantes às que apresentámos na nossa aplicação, contudo não conhecemos nenhuma que agregue simultaneamente tantos conceitos e resultados de probabilidades e estatística em língua portuguesa. Pensamos assim que a nossa aplicação poderá ser uma mais valia no ensino em Angola, onde o autor deste trabalho leciona.

A aplicação desenvolvida está longe de estar terminada e será um processo em constante construção.

Referências

Abrantes, P., Serrazina, L., e Oliveira, I. (1999). *A Matemática na educação básica.* Ministério da Educação. Departamento da Educação Básica, 1^a Edição. Lisboa.

Caldeira, S. C. B. (2009). *A estatística e as probabilidades no ensino secundário: análise dos programas de Matemática A e B na perspetiva do professor e dos alunos,* (Dissertação de Mestrado não publicada). Universidade de Lisboa, Faculdade de Ciências, Departamento de Estatística e Ivestigação Operacional.

Carvalho, C. (2006). Olhares sobre a educação estatística em portugal. *Simpósio internacional de pesquisa em educação Matemática–SIPEMAT*.

Chance, B., del Mas, R., e Garfield, J. (2004). Reasoning about sampling distribitions. the challenge of developing statistical literacy, reasoning and thinking. *Springer*, 295–323.

Fernandes, J. A., Bernabeu, C. B., García, J. M. C., e Batanero, C. D. (2009). A simulação em probabilidades e estatística: potencialidades e limitações, vol.1. *Quadrante, XVIII*, 161–183.

Gonçalves, E., e Lopes, N. M. (2000). *Probabilidades: Princípios Teóricos.* Escolar Editora, Lisboa, 1^a Edição.

Gordon, S. P., e Gordon, F. S. (2009). Visualizing and understanding probability and statistics: graphical simulations using excel, vol.19. *PRIMUS*(4), 346–369.

Guillermo, O. E. P., Tarouco, L. M. R., e Endres, L. A. M. a. (2005). O poder das simulações no ensino de hidráulica,vol.3. *RENOTE: revista novas tecnologias na educação. Porto Alegre*(1).

Heckler, V., Saraiva, M. d. F. O., e Filho, K. d. S. O. (2007). Uso de simuladores, imagens e animações como ferramentas auxiliares no ensino/aprendizagem de óptica, vol.29. *Revista Brasileira de Ensino de Física*(2), 267–273.

Jacob, B. L., e Doerr, H. M. (2014). Statistical reasoning with the sampling distribuition. *ICOTS9 Invited Paper - Refereed,* New York USA.

Kerns, G. J. (2010). *Introduction to Probability and Statistics Using R.* ulu.com.

Ko, E.-S. (2016). Development of an understanding of a sampling distribution. in: Ben-zvi d., makar k. (eds) the teaching and learning of statistics. *Springer*, 63–70.

Magalhães, C. (2005). *A utilização das imagens em contexto de ensino-aprendizagem: um estudo de caso no* 1^{o} *ciclo do ensino básico,* (Tese de doutoramento não publicada). Universidade do Minho, Instituto de Estudos da Criança.

Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media, vol.13. *Learning and instruction*(2), 125–139.

Pestana, D. D., e Velosa, S. F. (2010). *Introdução à Probabilidades e à Estatística.* Fundação Calouste Gulbenkian, 4^{a} Edição.

Varma, J. R., Virmani, V., et al. (2017). Shiny alternative for finance in the classroom. *Indian Institute of Management Ahmedabad, Research and Publication Department*.

Vijayakumar, C. (2016). Rstudio: The futu(r)e of data science, vol.5. *Open Sourse Special: CIO Review - The Navigator for Enterprise Solutions*, 10–14.