

O USO DA SIMULAÇÃO NO CÁLCULO DE PROBABILIDADES

Relatório de projeto

Andrea Inês Gaspar Cravo Dias

Trabalho realizado sob a orientação de

Professor Doutor Rui Santos, Instituto Politécnico de Leiria

Leiria, março de 2015

Mestrado em Ciências da Educação – Especialização em Utilização Pedagógica das TIC

ESCOLA SUPERIOR DE EDUCAÇÃO E CIÊNCIAS SOCIAIS

INSTITUTO POLITÉCNICO DE LEIRIA

AGRADECIMENTOS

As minhas primeiras palavras de agradecimento são dirigidas ao meu orientador, o Professor Doutor Rui Santos, pela inestimável colaboração, disponibilidade e incentivo demonstrado desde o primeiro momento. Todas as suas sugestões e todo o apoio prestado na orientação foram essenciais para o desenvolvimento e conclusão do presente trabalho.

Agradeço a todos os alunos participantes da experiência realizada no âmbito do projeto pela disponibilidade, empenho e interesse.

Agradeço a todos os professores e colegas de mestrado pelos momentos enriquecedores de aprendizagem e partilha.

Por último, e não menos importante, um infinito reconhecimento à minha família e aos meus amigos pela força e pelas palavras de encorajamento para concretizar o objetivo a que me propus.

RESUMO

As orientações curriculares da disciplina de matemática salientam frequentemente a importância do uso da tecnologia, nomeadamente dos computadores, como recurso pedagógico na construção de conhecimento nos mais diversos temas. A temática das probabilidades não é exceção e vários estudos confirmam o valioso contributo dos simuladores computacionais, especialmente no desenvolvimento correto de intuições probabilísticas e na desconstrução de perceções erradas muitas vezes partilhadas por docentes e discentes. A História da Matemática é igualmente mencionada nos programa como tema transversal a ser trabalhado ao longo do ano, com o objetivo de proporcionar aos alunos uma visão da matemática como uma ciência viva e em evolução. É igualmente defendido na literatura o estabelecimento de conexões entre os conteúdos probabilísticos trabalhados e a realidade dos estudantes, de modo a que estes tenham uma perceção verdadeira da importância da matemática na vida diária, bem como do seu contributo no desenvolvimento de outras ciências. Neste sentido, este trabalho apresenta material didático baseado na resolução de problemas históricos e reais, a partir da exploração de simuladores construídos com o *software* R. Os resultados da experiência pedagógica realizada, com recurso aos materiais desenvolvidos, sugerem que o uso de simuladores facilita a compreensão dos problemas apresentados, proporciona aos alunos um papel mais ativo na descoberta da solução e possibilita a concretização de todas as tarefas propostas. Foi igualmente consensual que a utilização do simulador oferece maior confiança nas respostas aos problemas propostos. Há, deste modo, fortes evidências que o recurso a simuladores computacionais deve ser tido em conta como ferramenta pedagógica no ensino das probabilidades.

Palavras chave

Ensino secundário, história da matemática, materiais pedagógicos, probabilidades, simulação, *software* R.

ABSTRACT

The curriculum guidelines for the teaching of mathematics in high school often emphasize the importance of using technology, namely computers, as a pedagogical resource in the construction of knowledge in various subjects. The theme of probability is no exception and several studies acknowledge the valuable contribution of computer simulators, mainly in the correct development of probabilistic intuitions and in the deconstruction of misperceptions held by both teachers and students. The history of mathematics is also mentioned in the curriculum as a crosscutting theme to be worked throughout the year in order to provide students with a vision of mathematics as a dynamical science, and therefore in constant progress. It is also supported in the literature the establishment of connections between the probabilistic contents and the day-to-day reality of the students, so that they have a true understanding of the importance of mathematics in daily life as well as its contribution in the development of other sciences. Thus, this work presents teaching resources based on the resolution of historical and real problems, by making use of simulators programmed in the statistical software R. The results of the performed pedagogical experiment, using the developed resources, reveal that the use of simulators enables the understanding of the presented problems, provides students a more active role in the discovery of the solution, and allows the implementation of all these tasks. It was also agreed that the use of the simulator provides greater confidence in the answers to the proposed problems. Therefore, there is strong evidence that the use of computer simulators must be taken into account as a pedagogical tool in the teaching of probability.

Keywords

High school education, history of mathematics, probability, R software, simulation, teaching resources.

ÍNDICE GERAL

Agradecimentos	iii
Resumo	v
Abstract.....	vii
Índice Geral	ix
Índice de Figuras	xiii
Índice de Tabelas	xv
Capítulo 1 – Introdução	1
1.1. Contextualização e objetivos do estudo	1
1.2. Estrutura do trabalho	3
Capítulo 2 – O ensino de probabilidades.....	5
2.1. Probabilidades no ensino básico e secundário.....	5
2.2. Problemas no ensino das probabilidades	8
2.2.1. Troca de acontecimentos na probabilidade condicionada	10
2.2.2. Condicionamento, causalidade e temporalidade	11
2.2.3. Confusão entre probabilidade condicionada e probabilidade conjunta.....	12
2.2.4. Confusão entre independência e incompatibilidade	13
2.2.5. Restrição do espaço amostral na probabilidade condicionada	13
2.2.6. Enviesamento da equiprobabilidade.....	14
2.3. A utilização da simulação.....	14
2.3.1. A simulação como estratégia educacional no ensino de probabilidades.....	14
2.3.2. O software R.....	17
Capítulo 3 – Propostas de aplicação	19
3.1. Problemas históricos.....	19
3.1.1. O jogo passe-dix	20

3.1.2. O paradoxo da caixa de Bertrand	24
3.1.3. O dilema de Monty Hall	27
3.2. Análises clínicas	32
3.2.1. Sensibilidade versus especificidade da análise.....	32
3.2.2. Exploração do problema a partir de tabelas de contingência	36
3.2.3. Exploração do problema a partir do Teorema da Probabilidade Total e da regra de Bayes	39
3.2.4. Contextualização didática do problema.....	40
.....	42
Capítulo 4 – Experiência pedagógica	43
4.1. Metodologia e participantes do estudo	43
4.2. Fases do estudo.....	45
4.3. Apresentação dos dados e discussão dos resultados.....	47
4.3.1. Falácias e erros probabilísticos.....	49
4.3.1.1. O raciocínio proporcional.....	50
4.3.1.2. Extração com reposição.....	52
4.3.1.3. Falácia da condicional transposta	53
4.3.1.4. Falácia da inversão do eixo temporal	55
4.3.1.5. Acontecimentos independentes	57
4.3.1.6. Falácia da conjunção	59
4.3.1.7. Considerações finais sobre a adesão aos erros probabilísticos.....	61
4.3.2. Resolução de problemas históricos a partir da simulação	64
4.3.3. Breve análise dos inquéritos de satisfação	67
Capítulo 5 – Conclusões finais, limitações ao estudo e trabalho futuro.....	69
Bibliografia.....	73
Anexos.....	1
Anexo A1 – Programas construídos para o <i>software</i> R.....	2

<i>A.1.1.- O jogo passe-dix</i>	2
<i>A.1.2.- O paradoxo da caixa de Bertrand</i>	3
<i>A.1.3.- O dilema de Monty Hall</i>	4
<i>A.1.4.- Análises clínicas</i>	6
<i>Anexo A2 – Materiais utilizados na experiência pedagógica</i>	7
<i>A.2.1.- Inquérito - pré teste</i>	7
<i>A.2.2.- Inquérito - pós teste</i>	11
<i>A.2.3.- Inquérito de satisfação</i>	14
<i>A.2.4.- Resultados do inquérito de satisfação</i>	15
<i>A.2.5.- Ficha de apoio</i>	18
<i>A.2.6.- Apresentação em powerpoint</i>	21

ÍNDICE DE FIGURAS

Figura 1 – Observação da soma obtida em 50 e 50 000 experiências.	21
Figura 2 – Evolução da proporção de 9 e 10 pontos obtidos em 35 sequências de 10000 experiências.	22
Figura 3 – Análise das possibilidades de se obter os resultados $\{1,2,6\}$, $\{2,2,5\}$ e $\{3,3,3\}$	23
Figura 4 – Contabilização dos casos favoráveis à obtenção de 9 e 10 pontos.	23
Figura 5 – Observação do tipo de moeda obtida em 50 e 20000 experiências.....	25
Figura 6 – Evolução da proporção de moedas de ouro e prata em 50 sequências de 1000 experiências.	25
Figura 7 – Diagrama de árvore relativo ao Paradoxo da Caixa de Bertrand.	27
Figura 8 – Resultados obtidos em 50 e 10000 experiências.....	29
Figura 9 – Evolução da proporção de troca e não troca de porta em 20 sequências de 500 experiências.	29
Figura 10 – Diagrama de árvore relativo ao dilema de Monty Hall.....	31
Figura 11 – Diagnóstico e situação real.....	34
Figura 12 – Maior sensibilidade.	34
Figura 13 – Maior especificidade.	35
Figura 14 – Diagrama do formulário facultado no exame nacional de MACS.....	40
Figura 15 – Valores do VP, FP, FN e VN obtidos na aplicação de um teste a 50000 utentes.....	42
Figura 16 – Resolução do aluno com o código 58.	51
Figura 17 – Resolução do aluno com o código 8.	51
Figura 18 – Resolução do aluno com o código 8.	52
Figura 19 – Resolução do aluno com o código 22.	53
Figura 20 – Resolução do aluno com o código 23.	55
Figura 21 – Resolução do aluno com o código 6.	56
Figura 22 – Resolução do aluno com o código 58.	57
Figura 23 – Resolução do aluno com o código 85.	59
Figura 24 – Resolução do aluno com o código 58.	60
Figura 25 – Resolução do aluno com o código 22.	60
Figura 26 – Resolução do aluno com o código 69	65
Figura 27 – Resolução do aluno com o código 23.	66

ÍNDICE DE TABELAS

Tabela 1 – Tabelas de contingência com os dados obtidos nos dois questionários.....	61
Tabela 2 – Associação entre as variáveis desempenho escolar e raciocínio probabilístico.	63

CAPÍTULO 1 – INTRODUÇÃO

1.1. CONTEXTUALIZAÇÃO E OBJETIVOS DO ESTUDO

Quer no ensino básico quer no ensino secundário português, à semelhança do que acontece na generalidade dos países ocidentais bem como em quase todos os países desenvolvidos, a estatística e as probabilidades já ocupam um lugar de relevo no atual programa de matemática contribuindo, para essa realidade, a implementação de vários projetos, as suas publicações e a disponibilização de materiais dos quais destacamos, em Portugal, o projeto ALEA – Acção Local de Estatística Aplicada (www.alea.pt).

De entre as razões mencionadas por diversos autores, que levam a incluir o ensino do estocástico no currículo escolar, Batanero, Godino & Roa [7] destacam:

- A sua presença na sociedade e na vida das pessoas, patente na frequente ocorrência de fenómenos aleatórios;
- A utilidade na tomada de decisões ao nível económico, social e político em ambientes de incerteza e a conseqüente necessidade de formar profissionais que possuam conhecimentos e raciocínios estocásticos;
- O papel que desempenha como ferramenta de investigação nas mais diversas áreas do conhecimento;
- O seu contributo para a promoção do raciocínio crítico.

Em Portugal o estudo dos fenómenos aleatórios faz parte da unidade curricular do 9.º e 12.º anos da disciplina de matemática, nos temas Probabilidade e Probabilidade e Combinatória, respetivamente. Como mencionado por vários estudos, este é um tema onde docentes e discentes apresentam diversas dificuldades. Essa realidade pode ser atestada nos diversos fóruns, *sites* e blogues criados para a troca de ideias, materiais e colocação de dúvidas pela comunidade docente de matemática, onde é frequente solicitar-se ajuda para a resolução de exercícios específicos que envolvam o cálculo de probabilidades.

Algumas das razões que se podem apontar para as dificuldades manifestadas pelos estudantes estão relacionadas com a natureza metodológica do seu ensino, uma vez que a abordagem aos conteúdos é feita com recurso quase exclusivo a exercícios do manual que privilegiam experiências aleatórias abstratas, restringem o cálculo da probabilidade à definição clássica e, já no 12.º ano, à aplicação rotineira de fórmulas combinatórias

pouco intuitivas. As atividades experimentais com recurso às tecnologias, fundamentais para complementar o ensino do estocástico, apesar de recomendado nos programas escolares nacionais ainda não são, pela nossa experiência pessoal, uma realidade na generalidade das nossas escolas.

Sendo uma temática com forte aplicabilidade, quer na vida quotidiana quer na vida profissional dos cidadãos, é importante desde logo guarnecer os jovens de melhores conhecimentos e de ferramentas mais adequadas para o desenvolvimento das suas intuições probabilísticas. Nesta perspetiva, é nosso objetivo apresentar materiais pedagógicos que explorem e facilitem a compreensão de conceitos probabilísticos, com recurso às TIC (Tecnologias de Informática e Comunicação), consideradas por diversos autores, incluindo Jonassen [18], ferramentas poderosas no apoio do processo de aprendizagem na medida em que facilitam ao aluno a descoberta, a resolução de problemas e a aprendizagem de conceitos. Em sùmula, permitem o seu envolvimento ativo na construção do conhecimento.

Os materiais apresentados foram produzidos com recurso ao *software* gratuito *R* (cf. www.r-project.org), e compreendem a realização de simulações probabilísticas de três problemas históricos e um problema real, selecionados em linha de consideração com o programa e com os princípios defendidos pelas teorias de aprendizagem construtivistas que privilegiam os simuladores como metodologia de ensino, na medida que estes permitem ao utilizador manipular variáveis, formular e testar hipóteses, analisar os resultados, raciocinar e construir/reconstruir conceitos. Todos os recursos apresentados podem ser trabalhados ao nível do ensino secundário, sendo nosso propósito suscitar o interesse e o envolvimento dos estudantes na procura das respostas aos desafios propostos.

De modo a testar o impacto da sua utilização no estudo do comportamento de fenómenos aleatórios e na melhoria das intuições probabilísticas, os materiais elaborados foram ainda aplicados a um pequeno grupo de alunos do ensino secundário que frequentaram a disciplina de matemática A. Para além da utilização de alguns materiais expostos no presente estudo, na experiência pedagógica foram ainda avaliadas questões que estão na base das dificuldades já identificadas em investigações desenvolvidas nos últimos anos, no que diz respeito à compreensão e aplicação dos conceitos de probabilidade condicionada, probabilidade da interseção de acontecimentos e independência entre acontecimentos.

Deste modo, com o presente trabalho pretende-se disponibilizar recursos didáticos, quer em termos de identificação de potenciais problemas, quer em termos de proporcionar o acesso a programação para a realização de simulações com recurso ao *software* R, nomeadamente numa aplicação à fiabilidade dos resultados das análises clínicas, bem como, através da realização de uma breve experiência pedagógica, aferir o contributo da utilização de simuladores computacionais na compreensão do comportamento dos fenómenos aleatórios e certificar as principais dificuldades sentidas pelos estudantes nesta área (entre as usuais identificadas na literatura).

1.2. ESTRUTURA DO TRABALHO

Este estudo encontra-se organizado em cinco partes principais. O primeiro capítulo consiste na introdução onde, após um resumido enquadramento do estudo, são explicitados a problemática e objetivos orientadores que lhe estão subjacentes, bem como justificada a relevância e pertinência do trabalho desenvolvido. Por último, é apresentada a sua estrutura organizativa.

No segundo capítulo, intitulado *O ensino de probabilidades*, apresentam-se, por um lado, as orientações curriculares relativas à temática das probabilidades ao nível do ensino básico e secundário e, por outro lado, o seu enquadramento teórico nomeadamente ao nível de estudos já realizados sobre as dificuldades do seu ensino e aprendizagem. Ainda neste capítulo, em consonância com os dois pontos anteriores, é fundamentado o recurso à simulação, apresentada como metodologia de ensino-aprendizagem eficiente e eficaz no desenvolvimento de intuições probabilísticas.

No terceiro capítulo, *Propostas de aplicação*, são descritos os materiais pedagógicos elaborados com recurso ao *software* R, tendo por base as considerações reproduzidas no segundo capítulo. A partir de três problemas históricos e de um problema de contexto real, procurou-se evidenciar a importância que a tecnologia e as simulações têm para a resolução de problemas probabilísticos e para o estudo do comportamento de fenómenos aleatórios.

No quarto capítulo, descreve-se a experiência pedagógica realizada, fundamentando-se as opções metodológicas e definindo-se as tarefas concebidas no desenvolvimento da experiência, bem como caracterizando os participantes e referindo as técnicas e os

instrumentos de recolha de dados. Por fim, são apresentados, analisados e discutidos os resultados obtidos, tendo em consideração os objetivos traçados.

No quinto e último capítulo apresentam-se as principais conclusões e limitações do estudo desenvolvido, bem como algumas sugestões para trabalho futuro.

Por último, apresenta-se a bibliografia seguida dos anexos. Nos anexos, entre outros, são fornecidas as programações relativas às propostas de aplicação apresentadas no capítulo 3, algumas das quais utilizadas na experiência pedagógica descrita no capítulo 4.

CAPÍTULO 2 – O ENSINO DE PROBABILIDADES

2.1. PROBABILIDADES NO ENSINO BÁSICO E SECUNDÁRIO

No atual contexto educativo em Portugal, a escolaridade obrigatória abrange as crianças e jovens com idades compreendidas entre os seis e os dezoito anos. O primeiro contato dos alunos com o estudo de probabilidades e de fenómenos aleatórios ocorre no nono ano de escolaridade do ensino regular, numa altura em os alunos estão ainda inseridos na escolaridade obrigatória. No programa constam os conceitos básicos associados à linguagem da probabilidade e suas propriedades elementares, a definição de Laplace de probabilidade (definição clássica) e a relação entre a frequência relativa de um acontecimento numa experiência aleatória repetida um grande número de vezes e a sua probabilidade (definição frequentista) [29].

Esta temática volta a ser objeto de estudo no ensino secundário, no último ano curricular das disciplinas de Matemática A, Matemática B e Matemática Aplicada às Ciências Sociais. Entre outros novos conteúdos, é introduzida a noção de probabilidade condicionada e de acontecimentos independentes.

Ao nível de indicações metodológicas, os programas do ensino secundário recomendam que se deve partir de problemas e **situações experimentais** para que, com o apoio na **intuição**, o aluno aceda gradualmente à formalização dos conceitos basilares de probabilidades. As experiências realizadas devem tirar partido de **materiais lúdicos** e de **simulações**, contribuindo para o esclarecimento de conceitos, a dinamização de discussões de carácter científico, bem como para incentivar o trabalho cooperativo. O programa preconiza a utilização obrigatória de **tecnologia** (nomeadamente o uso de calculadoras e computadores) que, além de ferramenta, deve ser uma fonte de motivação para a investigação e promover a participação ativa do estudante na sua aprendizagem. Como tema transversal devem ser trabalhados aspetos da **História da Matemática**, mostrando-a como ciência em construção e em constante interação com outras ciências. É ainda referido a importância da simulação e do jogo na construção adequada do espaço dos resultados e na obtenção de valores experimentais para a probabilidade de acontecimentos que estão a ser estudados, cf. [23], [24], [25], [26] e [27].

Realçamos, neste ponto, a importância dada nos programas de Matemática ao uso de tecnologias e, em particular, da simulação para o desenvolvimento da compreensão da noção de probabilidade, que deve ser trabalhada tanto no seu aspeto teórico e formal, bem como no experimental.

No novo programa do ensino básico, apenas encontram-se elencados, nas Metas Curriculares, os objetivos gerais a trabalhar em cada domínio de conteúdos, competindo aos organismos encarregues da gestão pedagógica dos estabelecimentos de ensino definir o modo de as concretizar “... devem decidir quais as metodologias e os recursos mais adequados para auxiliar os seus alunos a alcançar os desempenhos definidos nas Metas Curriculares.” [29, p. 28]. Desta forma, o presente programa valoriza a autonomia e a experiência dos professores e das escolas. A liberdade pedagógica das práticas letivas torna, cada vez mais evidente, a necessidade do professor estar informado e atualizado sobre estudos e investigações desenvolvidas nas diferentes áreas focadas nos conteúdos programáticos, para que opte pela melhor estratégia de ensino em consonância com as características dos seus alunos.

Apesar de se iniciar o ensino das probabilidades no nono ano de escolaridade, numa altura em que a generalidade dos alunos terá 14 anos de idade, o NCTM (National Council of Teachers of Mathematics) [30] recomenda antecipar o seu ensino para a educação primária de forma a proporcionar aos alunos uma experiência estocástica mais direta desde a sua infância, bem como contribuir, desde cedo, para a sua formação ao nível das intuições probabilísticas. Tal ideia está em consonância com o currículo escolar de outros países, tais como o Canadá, a Espanha, o Reino Unido e os Estados Unidos da América, assim como com as conclusões dos estudos realizados por Fischbein e Gazit [16] sobre intuições infantis, nos quais concluíram existirem ideias formadas e parcialmente corretas de conceitos probabilísticos em crianças com 10 anos de idade, tendo destacado os efeitos da instrução na melhoria destas intuições.

No antigo programa de matemática do ensino básico (homologado a 28 de dezembro de 2007 [22] e em vigor até ao vigente programa homologado a 17 de junho de 2013 [29]), um dos tópicos em estudo, logo nos dois últimos anos do primeiro ciclo, diz respeito às situações aleatórias. Pretende-se que o aluno trabalhe informalmente a noção de acaso, com recurso à experimentação de diferentes tipos de jogos e exploração de situações relacionadas com o dia a dia, e adquira o vocabulário próprio para as descrever (certo, possível, impossível, provável e improvável).

No entanto, contrariando os estudos e programas de Matemática em vigor noutros países, este tópico foi retirado no novo programa escolar na sequência da revogação do documento “Currículo Nacional do Ensino Básico – Competências Essenciais” (Despacho n.º 17169/2011, de 23 de dezembro). Na elaboração das metas curriculares (homologadas a 3 de agosto de 2012), podemos ler que:

As questões relativas a processos aleatórios foram propositadamente deixadas de lado por se entender que apresentam um grau de complexidade demasiado elevado para este nível de ensino, por falta de critérios suficientemente simples que conduzam os alunos a utilizar adequadamente a linguagem associada à interpretação dos fenómenos regidos pelo acaso. [28, p. 2].

Contudo, Batanero [1] defende que os conceitos probabilísticos e estatísticos devem ser trabalhados a partir dos primeiros anos do ensino básico, desde que as ideias fundamentais estocásticas sejam apresentadas nestas idades com um nível de adequação apropriado. Segundo a autora, os alunos devem construir o seu conhecimento de uma forma gradual e contínua dado que as ideias fundamentais estocásticas não diferem na forma estrutural, nas diferentes etapas educativas, mas apenas na forma linguística e no seu nível de aprofundamento.

No que diz respeito aos conceitos de probabilidade condicionada e independência, estes são lecionados apenas no ensino secundário. Contudo, o NCTM recomenda o seu ensino formal a alunos dos 9 aos 12 anos de idade e são vários os estudos, citados por Fernandes [15], (e.g., Jones *et al.* [19]; Tarr & Lannin [39]; Tarr [38] e Watson [43]) que referem que os tópicos de probabilidade condicionada e independência são apropriados para o currículo de matemática do 3.º ciclo do ensino básico.

Uma vez que as investigações realizadas evidenciam que as crianças possuem intuições probabilísticas, mesmo antes do seu ensino formal, seria importante integrar o seu desenvolvimento como componente básica para o ensino das probabilidades. E existindo intuições contrárias aos conceitos probabilísticos, seria igualmente necessária a sua exploração, de modo a corrigir ideias incorretas e aumentar a eficácia das intuições, formando indivíduos mais preparados para a tomada de decisões com que frequentemente se deparam. Além disso, e ainda segundo Fernandes, relacionar as características matemáticas de conceitos com ideias intuitivas dos indivíduos “ajuda-os a compreender e a aceitar resultados teóricos” [15, p. 109] e, deste modo, aprofundar os conhecimentos formais.

2.2. PROBLEMAS NO ENSINO DAS PROBABILIDADES

A teoria das probabilidades centra-se no estudo de fenómenos aleatórios, isto é, fenómenos que se caracterizam pela impossibilidade de prever os resultados das suas realizações individuais, mas igualmente pela existência de uma regularidade de comportamento quando visualizados um conjunto razoável de resultados.

Desta forma, a teoria das probabilidades procura quantificar a incerteza, quantificar o grau de possibilidade de um acontecimento se realizar. Como tal, podemos afirmar que o pensamento probabilístico distingue-se do algébrico, onde dois mais dois tem como único resultado possível o valor quatro, e do raciocínio lógico que é caracterizado pela dicotomia verdade *versus* falsidade. Na realidade, o número de sucessos de ocorrência dos fenómenos aleatórios não é simplesmente uma quantidade fixa, mas uma quantidade que varia em experiências repetidas (diferentes resultados são obtidos de cada vez que é realizada a experiência aleatória) acrescentando, desta forma, uma nova categoria à abordagem lógica: a categoria do possível.

Já em 1989, Steinbring, afirmou que as questões centrais do ensino de Probabilidades são:

How is it possible to convey the concepts of randomness and indeterminacy, and use them with the help of deterministic mathematical concepts? How can we make predictions about uncertain, random situations in the form of mathematical statements and what is the specific character of these predictions? [37, p. 205].

Estas questões atuais, apesar de decorridas mais de duas décadas, destacam claramente que a estocástica é um conteúdo completamente diferente de outros conteúdos matemáticos. Em consequência, é essencial desenvolver perspetivas e interpretações diferentes quando se ensina questões estocásticas.

Nos últimos anos foram realizados diversos estudos e investigações com o objetivo central de identificar as principais dificuldades evidenciadas pelos alunos na resolução de problemas relacionados com as probabilidades. Vários foram os erros identificados quer ao nível do raciocínio, quer ao nível da compreensão e aplicação dos conceitos probabilísticos, sobretudo no que se refere a situações que envolvem o conceito de probabilidade condicionada.

Em termos formais define-se probabilidade de A condicionada a B, ou probabilidade de A dado B, ao quociente entre a probabilidade conjunta dos acontecimentos A e B e a probabilidade do acontecimento dado. Em termos simbólicos:

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

A partir da definição de probabilidade condicionada pode ser deduzida a probabilidade da interseção de dois acontecimentos, conhecida como regra da multiplicação ou Teorema da Probabilidade Composta, $P(A \cap B) = P(B) \times P(A|B)$. Na probabilidade condicionada, a existência de informação sobre o acontecimento B irá permitir atualizar a atribuição de probabilidade ao acontecimento A. A probabilidade de A é assim avaliada sob as condições de um novo espaço amostral, condicionado pela ocorrência do acontecimento B.

Se a aplicação da fórmula parece não envolver cálculos complexos, já do ponto de vista didático, ao nível da interpretação, parece existir uma complexidade no raciocínio subjacente ao conceito de probabilidade condicionada, devido às muitas dificuldades manifestadas pelos alunos quando é necessário aplicá-lo na resolução de problemas e na tomada de decisões. Reconhecendo a sua importância, a compreensão intuitiva de probabilidade condicionada tem sido, nos últimos anos, alvo de avaliação por parte dos Exames Nacionais de Matemática A. A partir da elaboração de composições é pedido ao aluno que, sem aplicação da fórmula, determine o valor da probabilidade condicionada apenas por interpretação do seu significado e explicação do raciocínio que realizou, decorrente da restrição do espaço amostral.

Um outro conceito relacionado com a probabilidade condicional, onde parece existir dificuldades ao nível da interpretação, é o de acontecimentos independentes. Dois acontecimentos dizem-se independentes se e só se a probabilidade conjunta é igual ao produto das probabilidades de cada um deles, isto é,

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

Assim, podemos afirmar que dois acontecimentos são independentes se a realização de um dos acontecimentos não interferir na probabilidade de realização do outro acontecimento. Definindo a independência de acontecimentos em função da probabilidade condicionada podemos afirmar que A e B são independentes se e só se

$$P(A|B) = P(A) \text{ ou } P(B|A) = P(B), \text{ com } P(A) \neq 0 \text{ e } P(B) \neq 0.$$

Devido à falta de perceção de independência, é frequente os alunos calcularem a probabilidade da interseção de dois acontecimentos como sendo o produto das suas

probabilidades (o que só é correto quando existe independência entre os acontecimentos) em vez de usarem o Teorema da Probabilidade Composta. Ainda no cálculo da probabilidade conjunta, é frequente os alunos confundirem-na com a probabilidade da união ou simplesmente adicionarem as probabilidades dos acontecimentos que a constituem.

De entre os vários erros descritos nas principais investigações realizadas, quer no campo da Psicologia como da Educação Matemática, sobre a compreensão da probabilidade condicionada destacam-se:

- dificuldade em resolver problemas em que o acontecimento condicionante ocorre depois do condicionado,
- confusão entre acontecimentos independentes e disjuntos,
- dificuldade em determinar o espaço amostral,
- troca dos acontecimentos envolvidos na probabilidade condicionada, e
- confusão entre esta e a probabilidade conjunta, sendo-lhe muitas vezes atribuída um valor maior do que a probabilidade simples.

Em sùmula, existem diversas situações nas quais se verifica a violação das regras lógicas do cálculo de probabilidades.

De seguida, passemos à análise mais detalhada de cada um dos problemas anteriormente mencionados.

2.2.1. TROCA DE ACONTECIMENTOS NA PROBABILIDADE CONDICIONADA

Segundo Falk [14], muitos alunos não diferenciam adequadamente $P(A|B)$ de $P(B|A)$, ou seja, é frequente confundirem uma probabilidade condicionada com a sua transposta, erro que denominou por falácia da condicional transposta. Este erro é muito observado em problemas de contexto médico, onde se confunde a probabilidade de se ter uma doença quando o teste de diagnóstico deu positivo com a probabilidade de ter um resultado positivo no teste de diagnóstico sabendo que a pessoa está doente. Uma possível explicação para a confusão entre os dois significados da probabilidade condicionada é a imprecisão da linguagem usada nos enunciados dos problemas que, por vezes, se torna ambígua e dificulta a distinção entre a probabilidade condicionada e a sua transposta. Em oposição, a notação matemática é explícita na identificação do

acontecimento condicionante e do condicionado e, por isso, Falk propõe que se dispense a linguagem corrente e apenas se utilize a linguagem simbólica das probabilidades.

2.2.2. *CONDICIONAMENTO, CAUSALIDADE E TEMPORALIDADE*

Nas palavras de Batanero, Contreras e Diaz [4], a causalidade é um conceito científico, filosófico e psicológico complexo, apesar de intuitivamente ser compreendido e aceite pela maioria das pessoas, uma vez que todo o nosso conhecimento é construído numa base de relações causa e efeito entre diversos acontecimentos. Com base na experiência do quotidiano em situações de relação causa e efeito, é frequente algumas pessoas identificarem intuitivamente a condicionalidade e a causalidade como sendo similares. Assim, no cálculo da probabilidade condicionada, $P(A|B)$, há tendência para avaliar o tipo de relação existente entre o acontecimento condicionado (A) e o acontecimento condicionante (B). Nas investigações realizadas por Tversky e Kahneman [41] foram identificados dois tipos diferentes de relações que, psicologicamente, os indivíduos estabelecem entre os acontecimentos A e B. Se dentro do contexto apresentado, o sujeito interpreta B como uma causa da ocorrência de A, estabelece entre os acontecimentos uma relação causal, como por exemplo, quando é pedida a probabilidade de uma menina ter olhos azuis sabendo que a sua mãe tem olhos azuis. Por outro lado, se interpreta A como uma possível causa de B, estabelece entre os acontecimentos uma relação diagnóstica, como por exemplo, quando é pedida a probabilidade de uma mãe ter olhos azuis se a sua filha tem olhos azuis.

No tratamento normativo da probabilidade condicionada, $P(A|B)$, não se considera o tipo de relação que possa existir entre A e B mas as informações fornecidas pelo acontecimento condicionante. No entanto, de acordo com Falk [14], devido à prevalência de relações de causalidade na nossa perceção do mundo a maioria das pessoas indica que as relações causais são mais fortes que as relações diagnósticas e, por isso, consideram (erradamente) que $P(A|B)$ é sempre superior a $P(B|A)$, nos casos em que B parece ser uma causa natural de A.

Associada à relação de causalidade temos a sequência temporal, isto é, dificuldade na resolução de problemas de probabilidade condicionada que envolvam a inversão do eixo do tempo na ocorrência dos acontecimentos. A ideia errada que consiste em considerar

que um acontecimento não pode condicionar outro que ocorra anteriormente, ou seja, que o acontecimento condicionante na probabilidade condicionada tem de preceder sempre temporalmente o acontecimento condicionado, é designado por falácia da inversão do eixo temporal.

Podemos então identificar a existência de duas concepções erróneas da probabilidade condicionada $P(A|B)$: uma concepção cronológica em que a probabilidade condicionada é vista como uma relação temporal entre A e B, sendo B necessariamente anterior a A; e uma concepção causal em que a probabilidade condicionada é vista como uma relação causa-efeito entre os acontecimentos B (a causa) e A (a consequência).

Segundo Batanero, Contreras e Diaz [4], é importante erradicar estas concepções erradas uma vez que a ideia de rever uma probabilidade tendo em consideração resultados que lhe são posteriores é fundamental na compreensão e aplicação do Teorema de Bayes, resultado capital no estudo do aleatório.

As concepções erradas baseadas no princípio da causalidade e na irreversibilidade do tempo são também conhecidas por ‘Fenómeno Falk’ [15].

2.2.3. CONFUSÃO ENTRE PROBABILIDADE CONDICIONADA E PROBABILIDADE CONJUNTA

Alguns estudos referem que é frequente os alunos interpretarem a interseção como condicionalidade, ou seja, confundem os significados das probabilidades conjunta e condicional. Mais uma vez, como defendido por Falk [14], esta situação pode dever-se a uma incorreta compreensão da linguagem presente nos enunciados dos problemas, que resulta em dificuldades de tradução do enunciado para uma simbologia adequada e numa troca de identificação das probabilidades $P(A|B)$ e $P(A \cap B)$. Mas esta não será a única justificação para a existência dessa troca uma vez que, também quando os dados são apresentados em tabelas de dupla entrada, é frequente ocorrer equívocos no cálculo das probabilidades conjunta e condicionada [11]. Relacionada com esta dificuldade temos a falácia da conjunção que consiste na crença de que é mais provável a interseção de dois acontecimentos do que a probabilidade de qualquer um dos seus acontecimentos constituintes. Esta apreciação vai contra o princípio da lei da conjunção que estabelece que se $A \subseteq B$ então $P(A) \leq P(B)$ e, consequentemente, $P(A \cap B) \leq P(A)$ e $P(A \cap B) \leq P(B)$.

Segundo Tversky e Kahneman [42], o erro é resultado de se considerar a conjunção como mais representativa do que cada acontecimento separado. Esta situação ocorre, nomeadamente, quando a probabilidade de ocorrência de um dos acontecimentos é muito alta em comparação com o outro, levando a que o aluno só tenha em consideração esse acontecimento e não se consciencialize de que a interseção com outro acontecimento reduz a sua probabilidade.

2.2.4. CONFUSÃO ENTRE INDEPENDÊNCIA E INCOMPATIBILIDADE

A noção de independência de acontecimentos é muitas vezes confundida com o conceito de acontecimentos mutuamente exclusivos (disjuntos ou incompatíveis). Dois acontecimentos A e B dizem-se disjuntos quando $A \cap B = \emptyset$, e portanto, $P(A \cap B) = 0$. Logo, quando dois acontecimentos são mutuamente exclusivos, a realização de um dos acontecimentos implica a não realização do outro, logo $P(A|B) = P(B|A) = 0$, considerando $P(A) \neq 0$ e $P(B) \neq 0$. Distintamente, quando dois acontecimentos são independentes, a realização de um dos acontecimentos em nada interfere na probabilidade de realização do outro acontecimento. Nas situações em que A e B são compatíveis, os alunos interpretam essa informação como sendo acontecimentos dependentes, o que nem sempre se verifica. E nas situações em que A e B são disjuntos, os alunos interpretam essa informação como sendo acontecimentos independentes o que é falso já que um dos acontecimentos não pode ocorrer quando ocorre o outro e, por isso, acontecimentos disjuntos (com probabilidade não nula) são dependentes. Naturalmente, excluímos desta análise os casos especiais $P(A) = 0$ ou $P(B) = 0$.

2.2.5. RESTRIÇÃO DO ESPAÇO AMOSTRAL NA PROBABILIDADE CONDICIONADA

No cálculo de uma probabilidade condicionada $P(A|B)$, é necessário restringir o espaço amostral em função do acontecimento condicionante (B). É frequente os alunos não terem em conta a informação dada pelo acontecimento B e simplesmente calcularem a probabilidade de ocorrer o acontecimento A ou fazerem uma incorreta enumeração do espaço amostral. Também em situações de experiências envolvendo extrações com e sem reposição, é mais frequente ocorrer erros na determinação do espaço amostral

quando não há reposição dos elementos extraídos, como demonstrado nos estudos desenvolvidos por Fischbein e Gazit [16] onde um dos equívocos que identificaram no raciocínio dos alunos provinha exatamente da não percepção de que é modificado o espaço amostral nas situações sem reposição.

2.2.6. ENVIESAMENTO DA EQUIPROBABILIDADE

Para finalizar, é frequente os alunos admitirem que acontecimentos com caráter aleatório são por natureza equiprováveis, isto é, assumirem a equiprobabilidade dos acontecimentos elementares de uma qualquer experiência aleatória, mesmo nas situações em que tal não se verifica. Esse erro, comum entre os discentes, foi também praticado pelo matemático Jean d'Alembert (1717-1783) na publicação de um artigo sobre probabilidades, na Enciclopédia Francesa de 1754. Nesse artigo apresenta alguns problemas e as respectivas soluções, que estão incorretas dado as combinações por ele descritas não serem igualmente prováveis. Um dos problemas apresentado consiste em determinar a probabilidade de se obter pelo menos uma cara em dois lançamentos de uma moeda [40]. Em vez de $3/4$, apresentou erradamente a resposta $2/3$, não considerando que os acontecimentos elementares por ele definidos (obter uma cara, obter duas caras e não obter qualquer cara) não tinham a mesma possibilidade de ocorrerem.

2.3. A UTILIZAÇÃO DA SIMULAÇÃO

2.3.1. A SIMULAÇÃO COMO ESTRATÉGIA EDUCACIONAL NO ENSINO DE PROBABILIDADES

Como já mencionado, os fenómenos aleatórios caracterizam-se pela existência de uma irregularidade individual mas estabilidade coletiva. Estas características encontram-se patentes na noção frequentista de probabilidade, sendo esta definida como o valor para o qual tende a estabilizar a frequência relativa da realização do acontecimento, à medida que se aumenta o número de repetições da experiência aleatória. Assim, com um maior

número de experiências repetidas, espera-se que os valores estimados sejam boas aproximações das probabilidades teóricas.

No entanto, apesar de na maioria das situações reais a quantificação de probabilidade de um determinado acontecimento só ser possível empiricamente com a observação dos resultados de um grande número de experiências, existe maioritariamente nos manuais escolares, quer do ensino básico quer do ensino secundário, uma abordagem *a priori* da probabilidade com aplicação da Lei de Laplace, que só pode ser utilizado quando os resultados são equiprováveis e, ao nível do secundário, com recurso ao cálculo combinatório, pouco intuitivo e frequentemente de difícil compreensão.

Alargando a visão restrita dada às probabilidades nos manuais, o recurso à simulação, entendida como um processo artificial que permite reproduzir o comportamento de um fenómeno aleatório, pode ser considerado como uma importante estratégia de ensino para o desenvolvimento de uma adequada e completa compreensão do conceito de probabilidade [35].

De modo a permitir observar o comportamento de fenómenos aleatórios (nomeadamente a estabilidade da evolução das frequências relativas associadas a um qualquer acontecimento) e se conseguir obter uma estimativa próxima da probabilidade pretendida, é necessário repetir a experiência um elevado número de vezes, razão pela qual é inviável a realização de experiências aleatórias em sala de aula, com recurso a materiais manipuláveis. Como forma de ultrapassar os constrangimentos físicos e de morosidade que limitam a natureza e o âmbito dos problemas que podem ser trabalhados manualmente, é útil o uso de ferramentas tecnológicas, nomeadamente do computador, de modo a visualizar de forma rápida e intuitiva os resultados obtidos na simulação de um grande número de experiências [31].

Como recurso didático, um simulador computacional permite modelar fenómenos aleatórios, obter uma descrição das suas características e, a partir da identificação de regularidades e padrões, fornecer uma solução aproximada do problema tendo por base a interpretação frequencista de probabilidade ou a Lei dos Grandes Números. Os programas de simulação fornecem ao aluno instrumentos de exploração que o conduzem, de modo dinâmico, à descoberta e experimentação de conceitos e princípios que de outro modo seriam muito mais abstratos, contribuindo inequivocamente para o desenvolvimento de uma correta intuição probabilística e, provavelmente, para o aumento da motivação na aprendizagem das leis combinatórias e de probabilidades necessárias à prova. Ao permitir aos alunos experienciar os conceitos probabilísticos de

modo dinâmico, é-lhes mais fácil atribuir uma interpretação significativa da probabilidade determinada. No entanto, uma vez que não fornecem justificações nem demonstrações sobre a validade da solução obtida, os resultados aproximados alcançados através da simulação devem ser comparados com os valores teóricos determinados com recurso ao cálculo formal. Esta abordagem comparativa é defendida por diversos autores, entre os quais Batanero *et al.* [7], Graça Martins e Pontes [21] e Borovcnik e Peard [10], podendo-se encontrar igualmente referida no programa “... é importante ir sempre chamando a atenção dos estudantes para a confrontação dos resultados obtidos com os conhecimentos teóricos” [23, p. 16].

Tal abordagem dá maior credibilidade aos valores obtidos via simulação e permite ultrapassar a resistência de alguns alunos em aceitar resultados de experiências que não foram realizadas pessoalmente por eles [12]. Desta forma, a simulação permite aos alunos reconhecerem as diferenças e complementaridade entre os dois conceitos de probabilidades ensinados até ao 12.º ano, cf. [7] e [6].

Realçamos ainda a importância do uso da definição frequencista, na perceção e interpretação do significado atribuído ao valor da probabilidade de um acontecimento que muitas vezes não se consegue obter a partir da razão obtida pela definição clássica. Ainda relativamente ao uso da simulação, apesar das limitações, é igualmente importante realizar experiências estocásticas com recurso a objetos manipuláveis (por exemplo, dados, cartas ou roletas) por forma a que os estudantes experienciem gerar e reunir fisicamente os seus próprios dados. Face a esta questão, Borovcnik e Peard [10] propõem que, numa primeira abordagem às probabilidades ou como primeira etapa de modelação da experiência, os alunos utilizem objetos na realização de experiências aleatórias concretas e que, seguidamente, estabeleçam comparações com simulações em computador. Deste modo, procura-se que sejam os próprios alunos a sentir a necessidade de recorrer a ferramentas tecnológicas para identificar regularidades de algumas características da experiência e obter conclusões válidas sobre o problema em estudo, que dificilmente conseguiriam alcançar com a realização de um número reduzido de experiências.

2.3.2. O SOFTWARE R

Atualmente, já se encontram muitos programas disponíveis na Internet que possibilitam explorar diversas experiências aleatórias, com recurso à simulação. Esses recursos podem ser utilizados para melhorar a intuição dos alunos e, conseqüentemente, superar alguns dos erros expressos pela literatura. No entanto, Shaughnessy [36], mencionado por Fernandes [15], destaca a importância do papel do docente na construção dos programas computacionais. Desta forma, ao realizar os seus próprios materiais, o professor conseguirá direcionar as tarefas para as principais ideias e conceitos que pretende transmitir e ir ao encontro das dificuldades sentidas pelos seus discentes.

Na elaboração do material didático apresentado, optámos por explorar o *software* R por ser gratuito e ser um dos programas mais utilizados por estatísticos. Por já existirem *packages* que permitem uma mais fácil utilização (consulte-se, por exemplo, referências disponíveis no site do projeto ALEA – www.alea.pt), é igualmente cada vez mais aplicado no ensino das Probabilidades e da Estatística. Por estes motivos, pode ser instalado em qualquer computador e ser utilizado facilmente por docentes e discentes, razão pela qual elegemos este *software* ainda pouco utilizado no ensino básico e secundário.

Fernandes [15] refere que, do ponto de vista do ensino, as representações gráficas são frequentemente usadas nas melhores simulações e que é importante a sua exploração por variação dos valores de certos parâmetros. Assim, de acordo com estes princípios, os materiais desenvolvidos pretendem ilustrar a evolução das frequências relativas para um número crescente de experiências, nomeadamente através da exploração de gráficos obtidos pelos alunos a partir da escolha do número de iterações bem como do número de seqüências a observar. Por um lado, pretende-se explorar a sensibilidade dos resultados para a escolha do número de experiências realizadas e, por outro, proporcionar imagens visuais da solução matemática das experiências em estudo.

CAPÍTULO 3 – PROPOSTAS DE APLICAÇÃO

Reconhecendo que os alunos possuem intuições probabilísticas, é importante identificar estratégias eficazes para integrá-las no processo de ensino-aprendizagem. Essas estratégias poderão passar pelo recurso a problemas históricos e a problemas reais de aplicação das probabilidades que, para além da motivação, permitem trabalhar diversos conceitos e propriedades matemáticas (como espaço amostral, probabilidade condicionada, probabilidade conjunta, independência, Lei dos Grandes Números), diferentes definições de Probabilidade (regra de Laplace e definição frequencista), bem como aplicar diversos procedimentos nas suas resoluções (como a simulação e os diagramas de árvore).

Neste capítulo serão apresentados exemplos de três problemas históricos (Subsecção 1) e um problema real (Subsecção 2), todos eles envolvendo o cálculo de probabilidades condicionadas e que facilmente poderão ser enquadrados no currículo de Matemática A do 12.º ano de escolaridade. Serão analisadas as questões didáticas e formais respeitantes a cada um dos problemas e apresentado, em Anexo 1, a programação que foi construída em R para ser usada pelos professores para trabalhar esses problemas com os seus alunos em termos intuitivos, com recurso à simulação. As situações que irão ser apresentadas foram cuidadosamente escolhidas tendo em conta os princípios defendidos pela NCTM [30] sobre o uso de tarefas matemáticas que devem ser significativas para desenvolver e desafiar intelectualmente os alunos, despertar a sua curiosidade e envolvê-los para as questões matemáticas. Os problemas apresentados são, no nosso entender, relevantes ao nível da captação do interesse dos estudantes para a importância que a tecnologia e as simulações probabilísticas têm na resolução de problemas e na exploração do comportamento de fenómenos aleatórios.

3.1. PROBLEMAS HISTÓRICOS

A História da Probabilidade e da Estatística está repleta de episódios e problemas desafiadores que mostram que a intuição estocástica frequentemente nos engana, dado que as suas soluções são, por vezes contraintuitivas [3] (como ilustram o famoso problema de Monty Hall [5], o paradoxo da caixa de Bertrand [17] ou o jogo passe-dix

[35]). Muitos desses paradoxos são de solução acessível para os estudantes, ao nível do ensino secundário, e permitem organizar atividades didáticas no ensino e aprendizagem de conceitos probabilísticos. Além disso, é frequente os estudantes encontrarem, ao longo da sua vivência e aprendizagem, os mesmos paradoxos e situações contra intuitivas que aparecem no desenvolvimento histórico do cálculo de probabilidades e relativamente às quais possui intuições muitas vezes incorretas e resistentes à perspectiva normativa.

Lesser [20] defende que o uso adequado e moderado de exemplos contraintuitivos em aula apoiam uma pedagogia construtivista, promovendo uma aprendizagem profunda a partir das crenças prévias dos alunos e dando ao professor um papel de facilitador da aprendizagem. Como benefícios adicionais para os estudantes refere o desenvolvimento da motivação, da metacognição, do pensamento crítico e da aprendizagem por descoberta, além do estabelecimento de conexões entre aplicações da vida real e a história, como preconizado nos programas nacionais.

Como referido por Fernandes [15], a exploração de situações que ilustrem concepções erradas e enviesamentos no cálculo das probabilidades não reúne o consenso de todos os autores. No entanto, tais dúvidas podem ser ultrapassadas desde que os docentes façam uma escolha seletiva das situações a abordar, em concordância com a idade e as características dos seus alunos, de forma a garantir que estes dispõem dos meios adequados para analisar os paradoxos de modo significativo.

3.1.1. O JOGO PASSE-DIX

A origem das probabilidades está, de certa forma, relacionada com os chamados “jogos de azar”. A correspondência trocada em 1654 entre Pierre de Fermat (1601-1665) e Blaise Pascal (1623-1662) contendo reflexões sobre a resolução de problemas ligados a jogos, propostos por Antoine Gombaud (1610-1685) (conhecido por Chevalier de Méré), é considerada o despontar do cálculo das probabilidades. Contudo, não são os primeiros matemáticos a resolver problemas sobre cálculo de probabilidades. De entre os problemas relacionados com jogos, que previamente motivaram a reflexão e a discussão em torno das probabilidades, temos o jogo “passe-dix”. Este jogo de dados era frequentemente praticado nas cortes francesas no século XVII e consistia no lançamento simultâneo de três dados ao ar. Se a soma dos pontos obtidos fosse maior

que 10, o jogador ficaria com o montante apostado. Caso contrário, a aposta seria aumentada e passaria a vez a outro jogador.

O “paradoxo” subjacente ao jogo, que despertou a curiosidade de alguns matemáticos famosos como Girolamo Cardano (1501-1576) e Galileo Galilei (1564-1642), é a comparação entre a probabilidade de se obter 9 pontos com a de se obter 10 pontos, que muitos consideravam serem iguais pelo facto de ambas terem o mesmo número de possibilidades de ocorrer [35]. Representando, dentro de chavetas, as pontuações obtidas nos três dados lançados, podemos verificar que são seis as formas distintas de se obter soma igual a 9 pontos ($\{1,2,6\}$, $\{1,3,5\}$, $\{1,4,4\}$, $\{2,2,5\}$, $\{2,3,4\}$ e $\{3,3,3\}$) bem como de se obter soma igual a 10 pontos ($\{1,3,6\}$, $\{1,4,5\}$, $\{2,2,6\}$, $\{2,3,5\}$, $\{2,4,4\}$ e $\{3,3,4\}$). Contrariamente a este resultado os jogadores mais experientes, por repetida observação no lançamento dos três dados, suspeitavam ser mais provável obter-se 10 pontos do que 9 pontos.

A solução para esta aparente contradição pode ser facilmente encontrada a partir da simulação da experiência em causa, com recurso ao *software* R. À medida que se aumenta o número de experiências realizadas, a partir de sucessivas atribuições de valores crescentes ao parâmetro n nas simulações de n lançamentos dos 3 dados, e se obtém as respetivas somas, o utilizador poderá constatar que a frequência relativa de se obter 10 pontos é superior à de se obter 9 pontos.

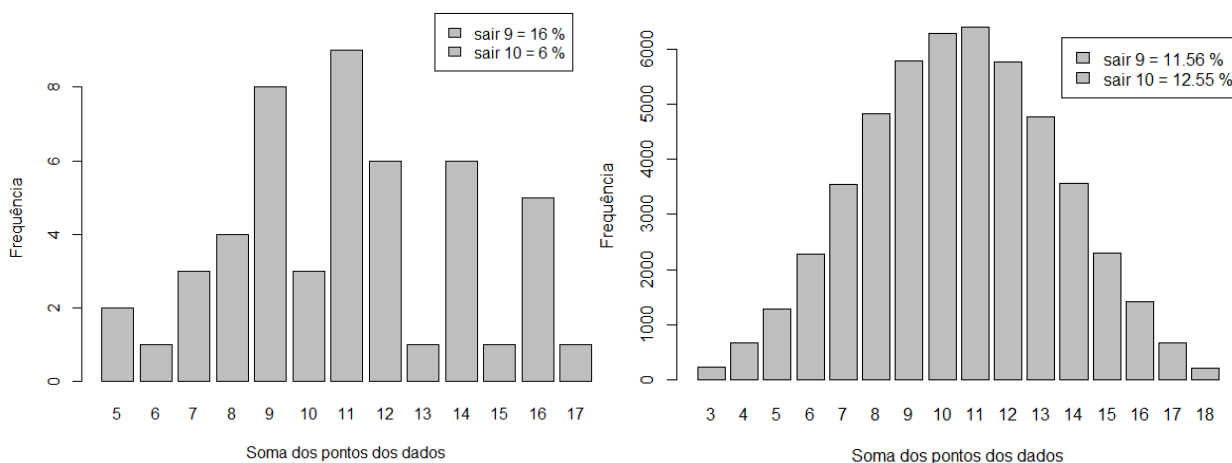


Figura 1 – Observação da soma obtida em 50 e 50 000 experiências.¹

¹ Consultar programação em Anexo A.1.1.

Note-se que, ao fim de 50 lançamentos dos três dados, de acordo com a Figura 1 obtém-se como estimativa para a probabilidade de se obter soma igual a 9 pontos 16%, percentagem muito superior à probabilidade de se obter soma igual a 10 pontos. No entanto, continuando o processo de simulação, com os sucessivos aumentos do número de experiências, as frequências obtidas para as diferentes somas tendem a estabilizar e, ao fim de 50 000 lançamentos, a probabilidade passa a ser de 11,56% no caso da soma igual a 9 pontos e de 12,55% no caso da soma igual a 10 pontos. Intuitivamente podemos concluir que os dois acontecimentos não são equiprováveis.

Sendo de extrema importância a relação entre a evolução da frequência relativa e o número de repetições da experiência, poderemos ainda a partir do programa R, construir gráficos como ilustrado na Figura 2. Com a simulação de várias sequências da experiência, poderemos observar e concluir que os valores para os quais tendem a estabilizar as frequências relativas da realização dos acontecimentos em estudo são próximos, sendo a probabilidade de se obter soma igual a 9 pontos um valor inferior a 12% (aproximadamente 0,115) e de se obter soma igual a 10 pontos um valor superior a 12% (aproximadamente 0,125).

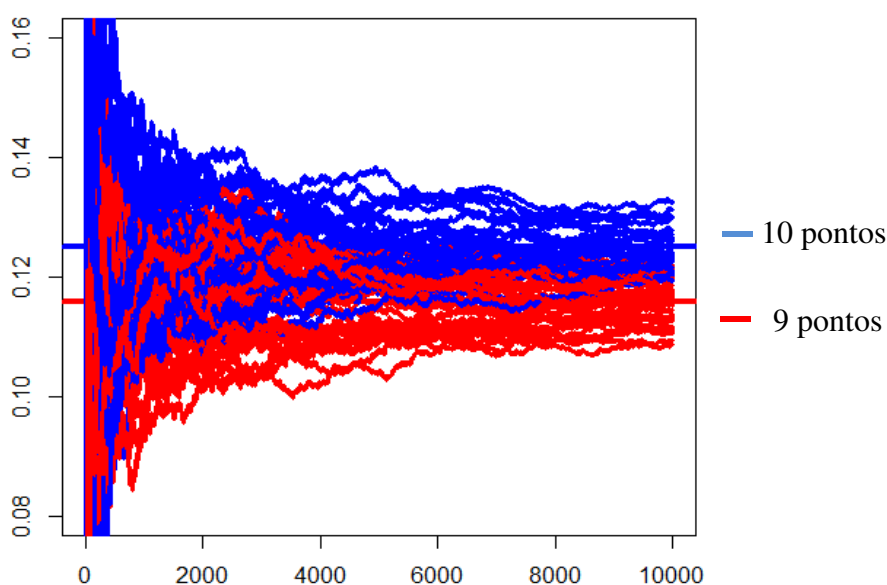


Figura 2 – Evolução da proporção de 9 e 10 pontos obtidos em 35 sequências de 10000 experiências.²

Na verdade, existe o mesmo número de formas distintas de alcançar somas iguais a 9 e 10 pontos. No entanto, os resultados apresentados têm diferentes possibilidades de ocorrerem e, por isso, cada um dos resultados deve ser analisado em termos de casos

² Consultar programação em Anexo A.1.1.

favoráveis à sua ocorrência. Tal como podemos verificar na Figura 3, quando os dados são lançados ao ar pode-se obter, por exemplo, o resultado {1,2,6} de seis formas distintas (tal como todos os resultados cujas pontuações dos dados sejam distintas). No entanto, o resultado {2,2,5} (e todos os resultados em que existam dois dados com pontuações iguais) pode ser obtido apenas de três formas diferentes e o resultado {3,3,3} (pontuações iguais nos três dados) tem uma única possibilidade de ocorrer.

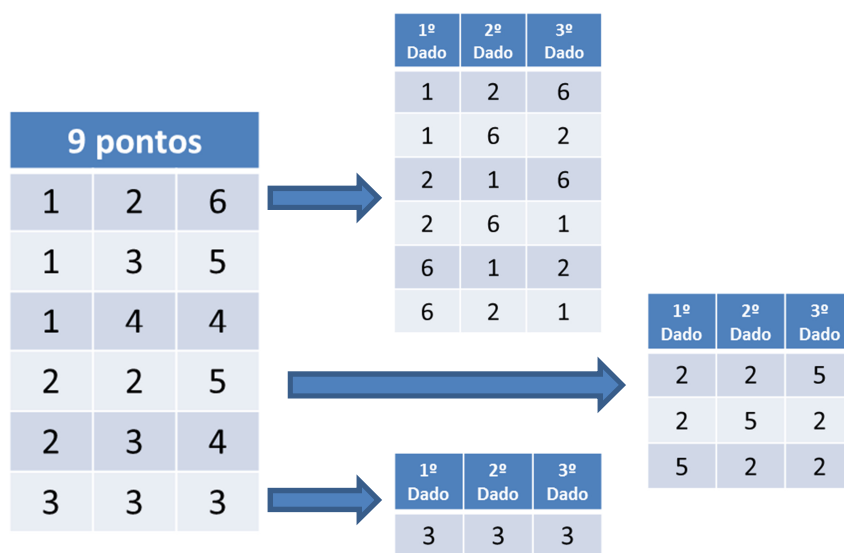


Figura 3 – Análise das possibilidades de se obter os resultados {1,2,6}, {2,2,5} e {3,3,3}.

Deste modo, dos 216 resultados distintos e equiprováveis do lançamento dos três dados ($6 \times 6 \times 6$), existem apenas 25 casos favoráveis à soma ser igual a 9 pontos e 27 favoráveis a ser igual a 10 pontos (conforme Figura 4), pelo que, a experiência dos jogadores está efetivamente de acordo com as verdadeiras probabilidades:

9 pontos	Possibilidades	10 pontos	Possibilidades
1 2 6	6	1 3 6	6
1 3 5	6	1 4 5	6
1 4 4	3	2 2 6	3
2 2 5	3	2 3 5	6
2 3 4	6	2 4 4	3
3 3 3	1	3 3 4	3
Total	25	Total	27

P (“obter soma igual a 9”) = 25/216

e

P (“obter soma igual a 10”) = 27/216.

Figura 4 – Contabilização dos casos favoráveis à obtenção de 9 e 10 pontos.

Note-se que os valores obtidos via simulação (ver Figuras 1 e 2) estão efetivamente próximos dos valores exatos das probabilidades de se obter soma igual a 9 e a 10 pontos.

Este problema histórico é um excelente exemplo para abordar os dois conceitos de probabilidades lecionados no ensino básico e secundário, na medida em que permite associar a observação dos jogadores experientes com a definição frequencista e enfatizar a restrição da aplicação da definição clássica às situações de equipossibilidade dos acontecimentos elementares (problemática analisada na subsecção 2.2.6.).

3.1.2. O PARADOXO DA CAIXA DE BERTRAND

O problema da caixa de Bertrand foi formulado por Joseph Bertrand (1822-1900), matemático francês do século XIX, no seu livro “Calcul des Probabilités”, publicado em 1888, no qual descreve o seguinte problema:

Temos três caixas e cada caixa tem duas gavetas, cada uma com uma moeda: uma caixa contém duas moedas de ouro, a outra caixa duas moedas de prata e a terceira caixa, uma moeda de cada tipo. Depois de se escolher uma caixa ao acaso, abre-se uma gaveta de forma aleatória verificando-se que esta contém uma moeda de ouro. Qual é a probabilidade de que a outra moeda dessa caixa seja também de ouro?

O problema descrito retrata uma experiência composta sujeita a uma condição (a primeira moeda é de ouro) e por isso estamos perante o cálculo de uma probabilidade condicionada.

Conduzidos por uma intuição errada, é expectável que alguns alunos façam o seguinte raciocínio: Depois de escolher uma caixa ao acaso e retirar uma moeda de ouro de umas das gavetas escolhidas aleatoriamente, resta duas opções: (a) ou teríamos escolhido a caixa que contém as duas moedas de ouro e neste caso a outra moeda que está na caixa será de ouro; (b) ou teríamos escolhido a caixa com uma moeda de ouro e uma moeda de prata e neste caso a outra moeda que está na caixa será de prata. Assim, a probabilidade de que a outra moeda que ficou na caixa também seja de ouro é igual a $\frac{1}{2}$ (o que está, evidentemente, errado).

Novamente, com recurso ao *software* R na simulação da experiência descrita, os alunos poderão rapidamente constatar que esta solução está incorreta uma vez que, à medida

que aumentam o número de experiências realizadas, a frequência relativa da moeda ser de ouro não é igual à de ser de prata. Intuitivamente concluem que a probabilidade pedida não é $1/2$.

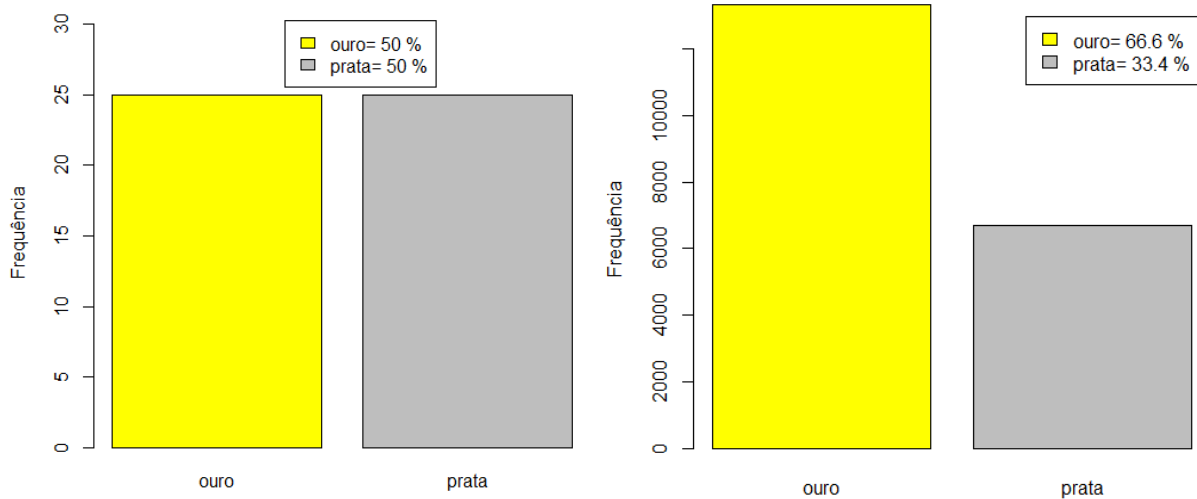


Figura 5 – Observação do tipo de moeda obtida em 50 e 20000 experiências.³

Poderão ainda graficamente encontrar o valor aproximado da probabilidade da segunda moeda ser de ouro, com a realização de várias sequências da experiência em causa, e chegar à solução correta de que a probabilidade é aproximadamente $2/3$.

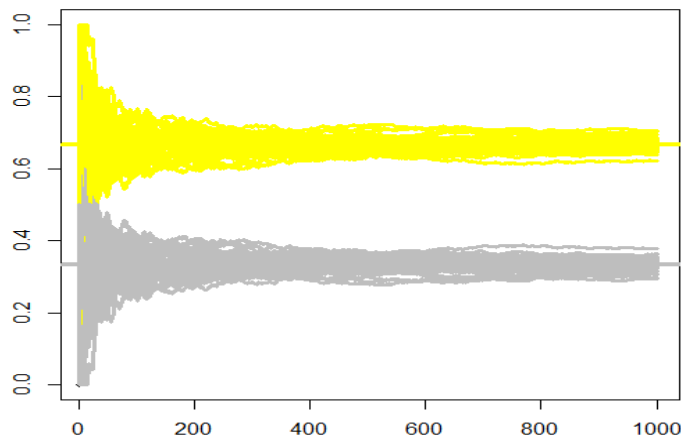


Figura 6 – Evolução da proporção de moedas de ouro e prata em 50 sequências de 1000 experiências.⁴

A incorreta percepção do espaço amostral e a incorreta percepção de independência dos acontecimentos são as causas apontadas pelos investigadores [17] para a existência de

³ Consultar programação em Anexo A.1.2.

⁴ Consultar programação em Anexo A.1.2.

soluções erradas, que foram descritas na resolução deste problema, e que também se refletem nos erros comuns dos alunos no cálculo de outras probabilidades condicionadas. É frequente estes não visualizarem a estrutura da experiência composta ou suporem erradamente a independência das sucessivas ocorrências (escolher uma caixa e escolher uma moeda). Este erro de raciocínio, conhecido como a falácia da inversão do eixo temporal e retratado na secção 2.2 do capítulo 2, consiste em supor que o acontecimento condicionante tem de preceder temporalmente o acontecimento condicionado, ou seja, que as informações atuais (a moeda mostrada) não podem afetar um evento que ocorreu antes da mesma (caixa selecionada).

Outra razão apontada consiste numa incorreta enumeração do espaço amostral. A intuição pode levar-nos a deduzir que, uma vez escolhida a caixa e obtendo-se a moeda de ouro, restam apenas dois casos equiprováveis (não se considera a caixa que contém as duas moedas de prata) e, portanto, existe 50% de possibilidade de que a moeda seja de ouro ou de prata. Não se tem, deste modo, em consideração que o primeiro universo (caixa com duas moedas de ouro) é mais provável do que o segundo (caixa com uma moeda de ouro e uma de prata), não diferenciando a ordem das gavetas na caixa que contém as duas moedas de ouro.

Como previamente referido, a simulação e a experimentação desempenham uma importante função na estabilização das intuições das crianças e na materialização dos problemas probabilísticos em estudo. No entanto devem ser complementadas por uma abordagem mais formal, a partir de cálculos combinatórios ou com recurso a outros esquemas auxiliares de contagem, de que os diagramas de árvore, os diagramas de Venn e as tabelas de contingência são exemplos, de modo a incentivar os alunos a resolver os problemas por vários processos. No caso do paradoxo em estudo, o recurso a um diagrama de árvore é essencial para que os alunos esquematizem e representem os diferentes acontecimentos da experiência, de modo a visualizar a sua natureza sequencial: escolher uma das três caixas, abrir uma das gavetas (segunda divisão no diagrama de árvores), e observar o tipo de moeda que fica na caixa (último ramo). Como se pode verificar na Figura 7, toda a informação do problema encontra-se sintetizada no diagrama de árvore.

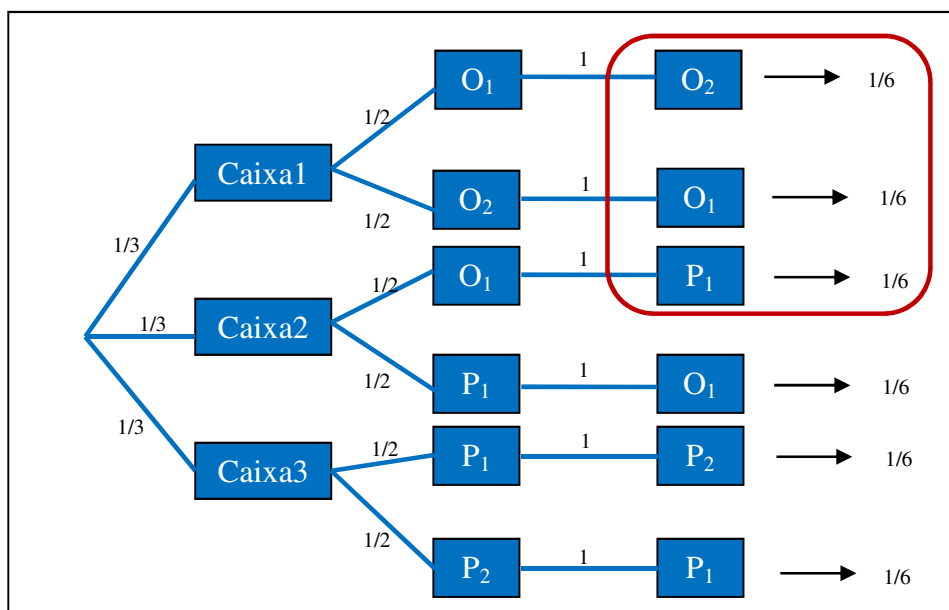


Figura 7 – Diagrama de árvore relativo ao Paradoxo da Caixa de Bertrand.

Desta forma, os alunos poderão observar que, uma vez escolhida a caixa e obtendo-se a moeda de ouro, restam três casos equiprováveis: a) ou teríamos aberto a gaveta da caixa que continha apenas uma moeda de ouro e, neste caso, a outra moeda será de prata; (b) ou teríamos aberto a primeira gaveta da caixa que contém as duas moedas de ouro e, neste caso, a outra moeda será de ouro; (c) ou então teríamos aberto a segunda gaveta da caixa que contém as duas moedas de ouro e, neste caso, a outra moeda será também de ouro. Ou seja, nos três casos possíveis, em dois verifica-se a situação de a outra moeda da caixa ser também de ouro. Por aplicação da definição clássica e da probabilidade condicionada, obtém-se o mesmo resultado:

$$P = \frac{\frac{1}{6} + \frac{1}{6}}{\frac{1}{6} + \frac{1}{6} + \frac{1}{6}} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}.$$

3.1.3. O DILEMA DE MONTY HALL

São muitas as variantes do paradoxo da Caixa de Bertrand e, por vezes, essas mudanças de formulação fazem com que não se reconheça o problema e apareçam novamente as mesmas dificuldades na sua resolução. Uma dessas variantes e, provavelmente, a mais conhecida, é o dilema de Monty Hall [5]. O problema de Monty Hall foi inspirado de

um concurso televisivo intitulado “Let’s Make a Deal”, emitido entre 1963 e 1986 na televisão americana e o seu nome provém do apresentador do concurso, Monty Hall. Também em Portugal, a partir de 1984, foi emitido uma versão do concurso com o nome “1, 2, 3”. O jogo consiste em apresentar três portas ao concorrente, sabendo-se que uma das portas esconde um carro enquanto as outras duas escondem uma cabra como prémio de consolação. Na primeira etapa, o concorrente escolhe uma porta (que não é aberta). De seguida, o apresentador abre uma das outras duas portas que o concorrente não escolheu, que contém uma cabra. Na segunda etapa, dá-se a possibilidade ao concorrente de trocar a sua escolha inicial pela porta ainda fechada. Colocam-se as seguintes questões: Qual é a estratégia que o jogador deve adotar? Ficar com a porta escolhida inicialmente ou mudar de porta? Será mais vantajoso trocar de porta, manter a escolha inicial ou será indiferente?

O referido jogo gerou polémica em relação às possíveis soluções do problema matemático que lhe é latente e que consiste em determinar que tipo de jogador tem maior probabilidade de ganhar o carro, o que nunca troca a porta escolhida inicialmente ou o que troca sempre. De modo similar ao problema da Caixa de Bertrand, anteriormente retratado, é expectável que alguns estudantes, conduzidos por uma intuição errada, pensem que é indiferente trocar a porta ou manter a escolha inicial. Este raciocínio incorreto resulta de assumir a equiprobabilidade em relação às duas portas que se mantêm fechadas, após ser retirada a porta que não tem prémio pelo apresentador. Assim, a porta escolhida inicialmente pelo jogador tem 50% de probabilidade de conter o carro e por isso é igual trocar de porta ou não fazê-lo.

O material apresentado, a partir do *software* R, proporciona uma experimentação empírica sobre os resultados que se obtém neste jogo com cada uma das distintas estratégias: trocar ou não trocar de porta. Da observação dos gráficos produzidos, o aluno facilmente obterá dados experimentais suficientes que o ajudarão a intuir a solução correta do problema: o jogador deve trocar de porta.

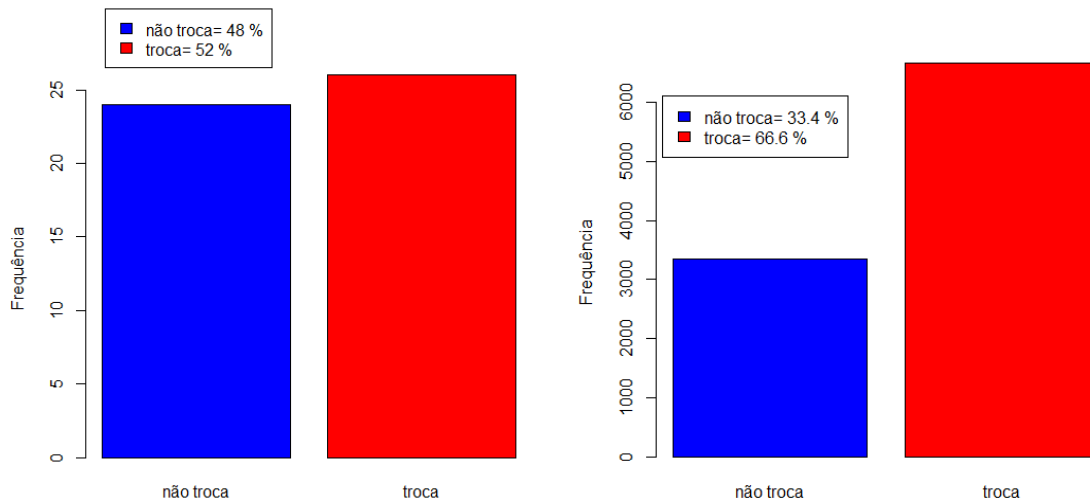


Figura 8 – Resultados obtidos em 50 e 10000 experiências.⁵

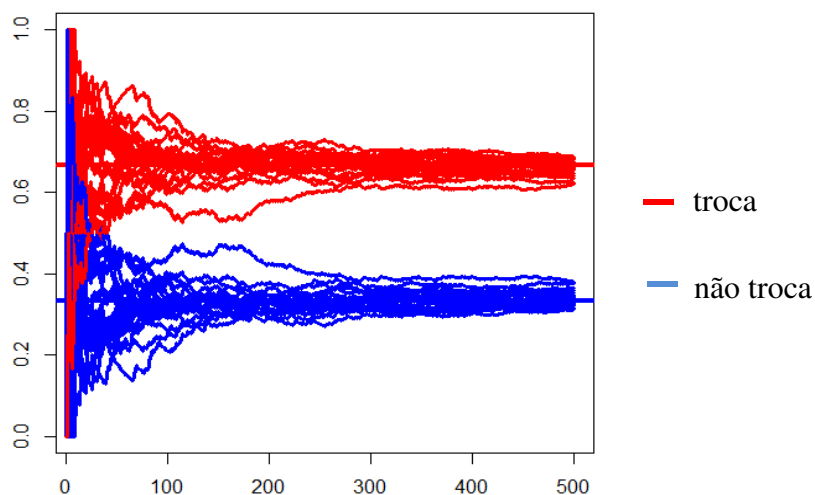


Figura 9 – Evolução da proporção de troca e não troca de porta em 20 seqüências de 500 experiências.⁶

Quando a solução encontrada é contraditória com a intuição inicial, gera-se um conflito cognitivo que, eventualmente, poderá encorajar os estudantes a examinarem criticamente as suas teorias intuitivas e a procurarem a prova formal do problema em estudo.

Para chegar à solução formal, e partindo da escolha feita inicialmente pelo jogador, podemos definir os acontecimentos:

A: “O jogador seleciona a porta que contém o carro.”

B: “O jogador seleciona a porta que contém uma cabra.”

G: “O jogador ganha o carro.”

⁵ Consultar programação em Anexo A.1.3.

⁶ Consultar programação em Anexo A.1.3.

Pretendemos determinar $P(G)$ para cada uma das estratégias possíveis. Uma vez que os acontecimentos A e B definem uma partição do espaço amostral (Ω) da nossa experiência, em que $A \cap B = \emptyset$ e $A \cup B = \Omega$, temos:

$$P(G) = P(G \cap \Omega) = P[G \cap (A \cup B)] = P[(G \cap A) \cup (G \cap B)].$$

Por aplicação do axioma da probabilidade da união de acontecimentos disjuntos (uma vez que de $A \cap B = \emptyset$ podemos concluir que $G \cap A \cap B = \emptyset$) e do Teorema da Probabilidade Composta obtemos:

$$P(G) = P(G \cap A) + P(G \cap B) = P(G|A) \times P(A) + P(G|B) \times P(B).$$

Por aplicação da Regra de Laplace, $P(A) = 1/3$ e $P(B) = 2/3$ pois, no universo em causa, existem duas cabras e um carro.

Finalmente, para calcular as probabilidades condicionadas, vamos considerar as duas estratégias distintas relativas à opção tomada pelo jogador.

Estratégia 1- No caso do jogador que nunca muda a sua escolha inicial, temos:

$$P(G|A) = 1 \text{ e } P(G|B) = 0.$$

$$\text{Neste caso, } P(G) = 1 \times \frac{1}{3} + 0 \times \frac{2}{3} = \frac{1}{3}.$$

Estratégia 2- No caso do jogador que troque sempre a porta escolhida inicialmente, temos: $P(G|A) = 0$ e $P(G|B) = 1$.

$$\text{Neste caso, } P(G) = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3}.$$

E por isso, sendo maior a probabilidade de ganhar na estratégia 2, é mais vantajoso o jogador trocar sempre a sua escolha inicial, tendo o dobro de hipóteses de ganhar relativamente à opção de não trocar de porta.

Tal como no paradoxo da Caixa de Bertrand, as principais causas apontadas pelos investigadores [5] para a existência de soluções erradas referem-se essencialmente à incorreta perceção do espaço amostral e à incorreta perceção de independência dos sucessivos acontecimentos. Um dos possíveis erros de raciocínio resulta de não se considerar ou não se compreender a forma como a informação proporcionada pelo apresentador (ao abrir uma das portas que contém uma cabra) afeta a probabilidade inicial de se ganhar o carro.

É também frequente não se considerar que a escolha do apresentador (que conhece, *a priori*, a porta que contém o carro) está condicionada à escolha inicial do jogador, influenciando deste modo o espaço amostral associado a cada acontecimento. Se o jogador escolhe, em primeira opção, a porta que contém o carro (com uma probabilidade de $1/3$), então o apresentador pode abrir qualquer uma das duas portas restantes. A escolha do apresentador restringe-se a duas possibilidades (cada uma das escolhas do apresentador tem probabilidade igual a $1/2$), sendo que o jogador perde o carro caso opte por trocar a porta escolhida inicialmente. Mas se o jogador inicialmente escolhe uma porta que contém uma das cabras (com uma probabilidade de $2/3$), só resta ao apresentador abrir uma porta (a única porta restante que esconde a outra cabra). Neste caso, o espaço amostral é reduzido a um só elemento (a probabilidade é 1) pelo que, se o jogador trocar a sua opção inicial, ganhará o carro.

Mesmo numa abordagem formal ao problema, o recurso a um diagrama de árvore, como apresentado na Figura 10, é essencial para que os alunos esquematizem a estrutura da experiência e os diferentes resultados, e clarifiquem os possíveis erros de raciocínios anteriormente descritos.

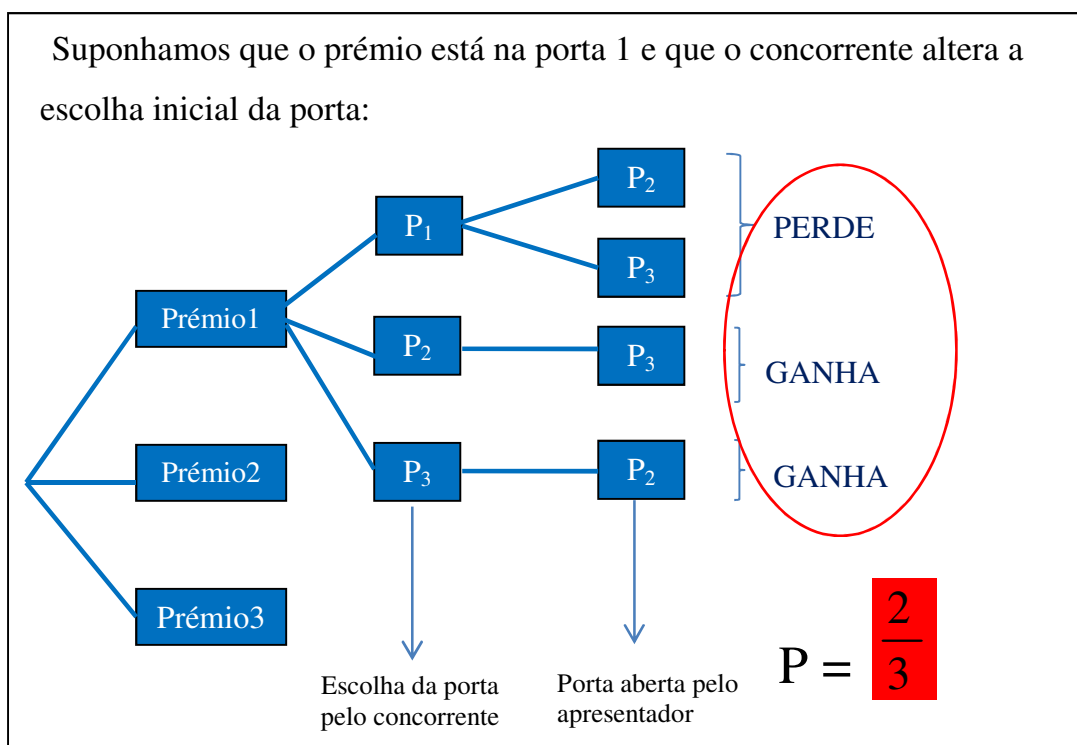


Figura 10 – Diagrama de árvore relativo ao dilema de Monty Hall

3.2. ANÁLISES CLÍNICAS

3.2.1. SENSIBILIDADE VERSUS ESPECIFICIDADE DA ANÁLISE

O conceito de probabilidade condicionada é fundamental nos estudos estocásticos porque permite atualizar a probabilidade associada a cada acontecimento, à medida que é adquirida nova informação (atualização do condicionante). O raciocínio condicionado também é aplicado no quotidiano e no campo profissional onde, frequentemente, é necessário tomar decisões em situações de incerteza como em testes médicos, em investimentos, em sentenças judiciais, entre outros. Este é um tema em que se observa a complexa inter-relação entre intuição e conhecimento formal no campo da probabilidade [8].

Tomemos, como exemplo da importância do raciocínio condicionado aplicado a situações quotidianas, os resultados de análises clínicas a que se recorre frequentemente, como meio auxiliar de diagnóstico de uma determinada doença e a sua correta interpretação [32].

Não sendo as análises clínicas infalíveis, poderá acontecer que alguns indivíduos, não tendo a doença, apresentem resultado positivo no teste clínico, vulgarmente designados por falsos positivos (FP). Ou ainda, poderá acontecer que indivíduos portadores da doença tenham resultado negativo, designados por falsos negativos (FN).

Assim, podemos dizer que no conjunto das pessoas portadoras de uma determinada doença (D), temos indivíduos cujo resultado da análise dá positivo (os verdadeiros positivos – VP) e os indivíduos cujo resultado dá negativo (FN). De forma análoga, de entre as pessoas que não têm a doença (\bar{D}), existem indivíduos cujo resultado clínico dá negativo (os verdadeiros negativos – VN) e os que têm resultado positivo (FP). Considerando que n representa o conjunto da população em estudo, a informação anterior pode ser registada numa tabela de dupla entrada como se apresenta de seguida:

	Doente	Não doente	Teste
Positivo	VP	FP	+
Negativo	FN	VN	-
Doença	D	\bar{D}	n

O ideal seria só ter verdadeiros positivos e negativos, ou seja, todos os doentes apresentassem resultado positivo e todas as pessoas sãs apresentassem resultado negativo nas análises. Em termos clínicos, designa-se por sensibilidade da análise (S) a percentagem de doentes que apresenta um resultado positivo num teste. Estamos perante o cálculo de uma probabilidade condicionada:

$$S = P(+ | D) = \frac{P(+ \cap D)}{P(D)} = \frac{VP}{VP+FN}.$$

Um teste com uma sensibilidade elevada será particularmente útil no rastreio de um determinado diagnóstico (baixa percentagem de falsos negativos), uma vez que a sensibilidade mede a capacidade de a análise identificar um doente.

Por outro lado, designa-se por especificidade da análise (E) a probabilidade condicionada que corresponde à percentagem de saudáveis que apresenta um resultado negativo no teste:

$$E = P(- | \bar{D}) = \frac{P(- \cap \bar{D})}{P(\bar{D})} = \frac{VN}{FP+VN}.$$

Um teste com uma especificidade elevada será particularmente útil na confirmação de um determinado diagnóstico (baixa percentagem de falsos positivos), pois a especificidade mede a capacidade de a análise identificar um indivíduo saudável.

Se o ideal seria só ter verdadeiros positivos e negativos, o mesmo será dizer que a percentagem da sensibilidade e da especificidade da análise deveriam ser de 100%. Mas tal situação não é possível de ocorrer uma vez que qualquer alteração de S implica uma alteração inversa de E.

A maioria das análises são quantitativas, isto é, o resultado da análise é uma quantidade através da qual será possível distinguir entre doentes e saudáveis. Deste modo, na Figura 11 está representada uma curva correspondente aos doentes (D) e uma curva maior correspondente aos que não são doentes (\bar{D}), as quais representam as quantidades observadas nas análises nos indivíduos doentes e nos indivíduos saudáveis. A reta vertical representada é obtida a partir do ponto de corte que não é mais do que o valor da análise que distingue os infetados dos não infetados. Por outras palavras, é o valor da análise clínica que estabelece o limite entre doente e são: resultados inferiores ou superiores ao ponto de corte vão permitir classificar os indivíduos como sendo saudáveis ou doentes. A reta vertical define, deste modo, a fronteira entre os resultados que são considerados negativos e os resultados considerados positivos. No entanto, em ambas as situações podemos ter resultados falsos ou verdadeiros. A sombreado azul é

representado a percentagem de indivíduos sãos com diagnóstico positivo e a sombreado vermelho a percentagem de indivíduos doentes com diagnóstico negativo, respetivamente à direita e à esquerda da reta vertical.

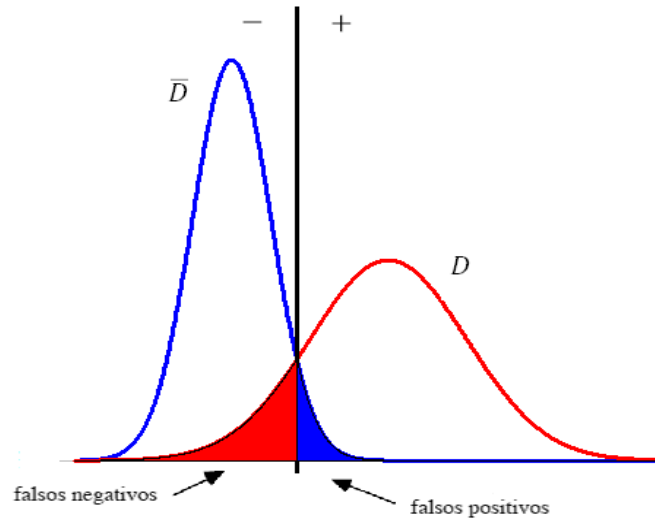


Figura 11 – Diagnóstico e situação real.

O aumento da percentagem da sensibilidade da análise (isto é, mais doentes a serem classificados como positivos) acarreta como contrapartida o aumento de falsos positivos (e por isso será maior a probabilidade de se diagnosticar erradamente a doença a um indivíduo são). Podemos verificar essa situação na Figura 12, a qual foi obtida deslocando a reta vertical para a esquerda em relação à Figura 11 (isto é, diminuindo o valor do ponto de corte).

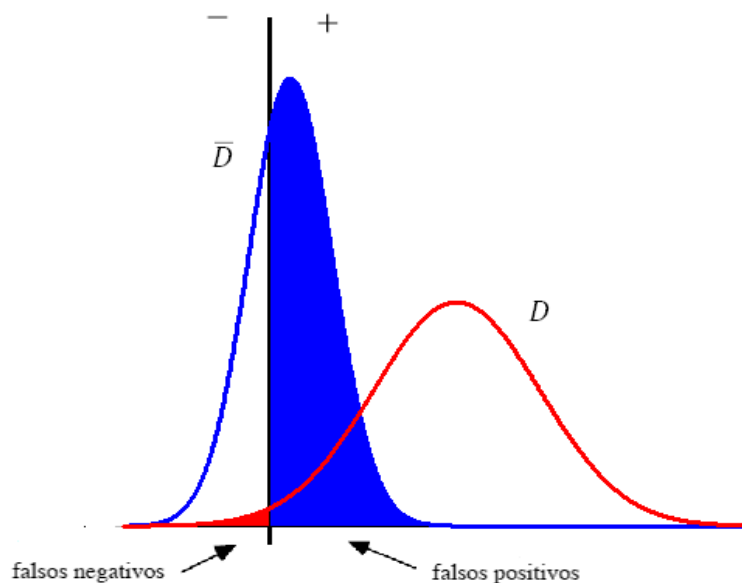


Figura 12 – Maior sensibilidade.

Por outro lado, se deslocarmos a reta vertical para a direita (Figura 13), o que corresponde a aumentar o ponto de corte, a especificidade da análise aumenta e, por isso, será menor a probabilidade de se diagnosticar erradamente a doença a um indivíduo sã. Mas, mais uma vez, como contrapartida, a sensibilidade da análise diminui e maior será a probabilidade de um doente escapar à deteção através desta análise.

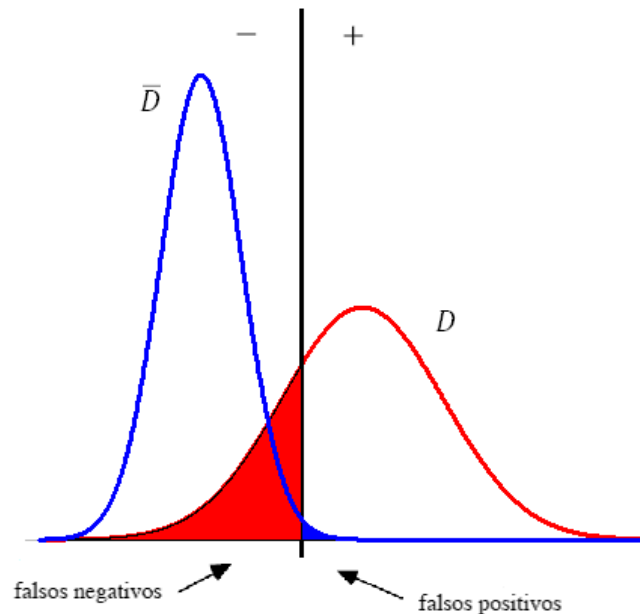


Figura 13 – Maior especificidade.

Para definir a sensibilidade e a especificidade da análise, ou seja, para estabelecer onde colocar a fasquia que separa negativos de positivos, é necessário ter em conta o equilíbrio entre os riscos mencionados, determinado a partir do tratamento estatístico dos dados disponíveis pelos diversos estudos realizados.

Muitas doenças são de difícil diagnóstico devido à diversidade dos sintomas clínicos e, por vezes, à sua semelhança com outras doenças. Na maior parte das vezes, os resultados das análises realizadas não permitem, só por si, um diagnóstico da doença, mas são fundamentais no cálculo das probabilidades condicionais transpostas da sensibilidade e especificidade da análise, $P(D|+)$ e $P(\bar{D}|-)$, uma vez que essa informação permite rever/atualizar a probabilidade inicial de existência de doença para os indivíduos que são testados e, com base nessa nova estimativa, permite ao clínico escolher a ação a realizar (efetuar outros testes clínicos ou iniciar tratamento terapêutico).

Em termos clínicos, $P(D|+)$ designa-se por valor preditivo positivo (VPP) e define a probabilidade de que um resultado positivo de um teste signifique presença de doença, isto é,

$$VPP = P(D|+) = \frac{P(D \cap +)}{P(+)} = \frac{VP}{VP+FP}.$$

De forma semelhante, $P(\bar{D}|-)$ designa-se por valor preditivo negativo (VPN) e define a probabilidade de que um resultado negativo de um teste signifique ausência de doença, ou seja, em termos de cálculo numérico,

$$VPN = P(\bar{D}|-) = \frac{P(\bar{D} \cap -)}{P(-)} = \frac{VN}{FN+VN}.$$

No caso de o resultado ser positivo, um teste com um VPP elevado permite fundamentar razoavelmente o diagnóstico inicial de doença e, no caso de ser negativo, um teste com um VPN elevado permite eliminar razoavelmente o diagnóstico de doença.

Ainda poderá ter interesse determinar a acurácia do teste (AT) que nos fornece o seu grau de precisão e que se obtém a partir do rácio entre os resultados verdadeiros obtidos na aplicação do referido teste e o número de testes realizados, isto é,

$$AT = (VP+VN) / n,$$

que corresponde à proporção de resultados corretos.

3.2.2. EXPLORAÇÃO DO PROBLEMA A PARTIR DE TABELAS DE CONTINGÊNCIA

Para exemplificar a importância dos resultados de uma análise e que interpretação se deve fazer dos mesmos, consideremos a *Lúpus Eritematoso Sistémico* (vulgarmente designada apenas por Lúpus). É uma doença reumática sistémica, autoimune, que tem diagnóstico precoce difícil, por possuir uma grande variedade de sintomas, por vezes intermitentes, que se assemelham aos de várias outras doenças e com grandes variações de manifestações de um paciente para outro⁷. A doença atinge cerca de 15 mil portugueses sendo as mulheres as mais afetadas, na proporção de dez a quinze para um homem⁸. Um dos exames utilizados, nos casos de suspeita da doença, é o “FAN (fator anti-núcleo), que está presente em cerca de 99% dos casos de Lúpus mas também em

⁷ Informação retirada do site www.spreumatologia.pt/doencas/lupus-eritematoso-sistemico no dia 10/07/2014.

⁸ Informação retirada do site <http://www.cmjornal.xl.pt/detalhe/noticias/nacional/saude/15-mil-sofrem-com-lupus-em-portugal> no dia 10/07/2014.

até 30% das pessoas normais”⁹. A partir destas informações, caso classifiquemos como doentes todos os indivíduos que tenham presente o FAN, podemos afirmar que a sensibilidade da análise é de 99% e que a sua especificidade é de 70%.

Sendo a prevalência de uma doença avaliada pela proporção de doentes, isto é, a fração $p = D/n$, no caso da doença do Lúpus estima-se que, na população portuguesa, seja $p = 15000/10000000 = 3/2000$, ou seja, 3 em cada 2 000 portugueses sofrem desta doença.

Suponhamos que um médico, devido ao quadro clínico apresentado, suspeita que uma das suas pacientes, a Maria, sofre de Lúpus e que em estudos anteriores, 40% dos indivíduos que apresentam os mesmos sintomas são portadores da doença. Como interpretar os resultados de uma análise ao FAN? Começemos por elaborar uma tabela de dupla entrada que relacione as variáveis doentes/não doentes com o resultado positivo/negativo obtido na análise. Suponha-se, apenas para facilitar os cálculos, que se tem informação de análises a 1000 indivíduos ($n = 1000$) na situação da Maria, isto é, com os mesmos sintomas. Uma vez que a prevalência da doença para esta subpopulação é de 40%, podemos dizer que $D = 1000 \times 0,4 = 400$ e que $\bar{D} = 1000 - D = 600$.

Se a sensibilidade da análise é 99% então:

$$VP = D \times S = 400 \times 0,99 = 396 \quad \text{e} \quad FN = D - VP = 400 - 396 = 4.$$

Se a especificidade da análise é 70% então:

$$VN = \bar{D} \times E = 600 \times 0,70 = 420 \quad \text{e} \quad FP = \bar{D} - VN = 600 - 420 = 180.$$

Temos assim:

	Doente	Não doente	
Positivo	396	180	576
Negativo	4	420	424
	400	600	1000

Aplicando a fórmula do cálculo da probabilidade condicionada, o VPP da análise é dado por:

$$P(D|+) = \frac{P(D \cap +)}{P(+)} = \frac{VP}{VP+FP} = \frac{396}{576} = 68,75\%.$$

⁹ Informação retirada do site <http://www.reumatologiaavancada.com.br/doencas-reumaticas/lupus/> no dia 10/07/2014.

E o VPN da análise é dado por:

$$P(\bar{D}|-) = \frac{P(\bar{D} \cap -)}{P(-)} = \frac{VN}{FN+VN} = \frac{420}{424} = 99,1\%.$$

Caso o resultado da análise dê positivo, os 40% iniciais da prevalência da doença são transformados numa probabilidade *a posteriori* de 68,75%. Apesar de superior, é ainda assim pouco conclusivo para diagnosticar a doença (possivelmente, o médico iria optar por realizar outros testes clínicos).

Caso o resultado da análise dê negativo, os 60% iniciais de a Maria não ter Lúpus são transformados numa probabilidade *a posteriori* de 99,1%, valor razoável para despistar a existência de doença.

Suponhamos agora uma nova situação: A Maria não apresenta nenhum sintoma da doença mas mesmo assim convence o seu médico de que deveria fazer a análise do FAN. Sendo o resultado da análise positivo, Maria ficou muito assustada. Qual é então a probabilidade de a Maria sofrer de facto de Lúpus?

Mais uma vez, para facilitar os cálculos, suponhamos que durante um rastreio da população faz-se a análise a 400 000 indivíduos. Como a prevalência da doença na população portuguesa é de $3/2000 = 0,0015$, podemos preencher a linha de totais da tabela com os valores:

$$D = \frac{3}{2\,000} \times 400\,000 = 600 \quad \text{e} \quad \bar{D} = 400\,000 - 600 = 399\,400.$$

De forma análoga à situação anterior, e usando os conhecimentos sobre a sensibilidade e a especificidade da análise, obtém-se a seguinte tabela de contingência:

	Doente	Não doente	
Positivo	594	119820	120414
Negativo	6	279580	279586
	600	399400	400000

Nesta situação, o valor preditivo positivo da análise é dado por:

$$P(D|+) = \frac{594}{120414} = 0,49\%.$$

E o valor preditivo negativo da análise é dado por:

$$P(\bar{D}|-) = \frac{279580}{279586} = 99,998\%.$$

Comparativamente à situação anterior, o valor preditivo positivo diminuiu mas, em contrapartida, aumentou o valor preditivo negativo, tal como maior sensibilidade arrasta menor especificidade.

Apesar do resultado positivo na análise, a probabilidade de a Maria ser portadora de Lúpus é de apenas 0,49%, valor bastante irrisório. Sendo tão baixa a probabilidade de um positivo ter de facto a doença, será inútil um paciente que não apresente sintomas realizar a análise ao FAN. No entanto, podemos analisar este resultado na perspetiva de que, a probabilidade de ter a doença sobe de 15 em 10 000 para 49 em 10 000 nos casos positivos da análise.

3.2.3. EXPLORAÇÃO DO PROBLEMA A PARTIR DO TEOREMA DA PROBABILIDADE TOTAL E DA REGRA DE BAYES

Na análise do estudo apresentado, recorreu-se às tabelas de dupla entrada. Mas o mesmo problema poderia ser abordado por aplicação do Teorema da Probabilidade Total e da regra de Bayes, conteúdos lecionados na disciplina de Matemática Aplicada às Ciências Sociais (MACS).

A partir da definição de probabilidade condicional, por um lado temos que:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A \cap B) = P(A|B) \times P(B).$$

Por outro lado, verifica-se igualmente que:

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Leftrightarrow P(A \cap B) = P(B|A) \times P(A).$$

Logo podemos concluir que:

$$P(B|A) \times P(A) = P(A|B) \times P(B) \Leftrightarrow P(B|A) = \frac{P(A|B) \times P(B)}{P(A)}.$$

Esta é a fórmula utilizada para calcular a probabilidade inversa da condicional, a qual se designa por regra de Bayes.

Se considerarmos a partição $\{B, \bar{B}\}$ do espaço amostral Ω (como ilustrado na Figura 14), a probabilidade do acontecimento A pode ser obtido a partir de:

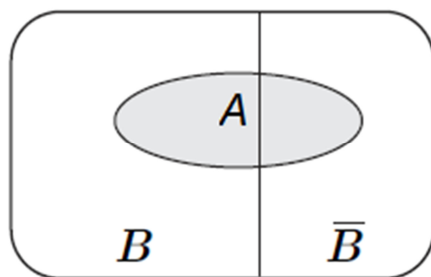


Figura 14 – Diagrama do formulário facultado no exame nacional de MACS.

$P(A) = P(A \cap \Omega) = P[A \cap (B \cup \bar{B})] = P[(A \cap B) \cup (A \cap \bar{B})] = P(A \cap B) + P(A \cap \bar{B})$,
 uma vez que os acontecimentos são disjuntos. Logo, podemos concluir que:

$$P(A) = P(A|B) \times P(B) + P(A|\bar{B}) \times P(\bar{B}),$$

designado por Teorema da Probabilidade Total.

Aplicando o Teorema da Probabilidade Total e a regra de Bayes no cálculo do VPP e VPN do primeiro caso apresentado, em que a prevalência da doença era de 40%, e considerando uma partição do universo em doentes (D) e não doentes (\bar{D}) teremos:

$$\text{VPP: } P(D|+) = \frac{P(+|D) \times P(D)}{P(+|D) \times P(D) + P(+|\bar{D}) \times P(\bar{D})} = \frac{0,99 \times 0,4}{0,99 \times 0,4 + (1 - 0,7) \times 0,6} = 0,6875$$

$$\text{VPN: } P(\bar{D}|-) = \frac{P(-|\bar{D}) \times P(\bar{D})}{P(-|\bar{D}) \times P(\bar{D}) + P(-|D) \times P(D)} = \frac{0,7 \times 0,6}{0,7 \times 0,6 + (1 - 0,99) \times 0,4} = 0,991.$$

Obtemos, deste modo, os mesmos valores que anteriormente recorrendo às tabelas de contingência.

3.2.4. CONTEXTUALIZAÇÃO DIDÁTICA DO PROBLEMA

Torna-se evidente a importância dos conhecimentos probabilísticos mesmo numa “simples” e correta interpretação das análises clínicas, onde é frequente confundir-se a probabilidade de ter uma doença quando o teste de diagnóstico deu positivo, com a probabilidade de obter um resultado positivo no teste de diagnóstico sabendo que se tem a doença. Por outras palavras, confusão entre uma probabilidade condicionada e sua transposta. Esse erro é comum no cálculo de probabilidades condicionadas e é designada por *Falácia da Condicional Transposta* (ver secção 2.2.1 do capítulo 2). Também as investigações na didática da Probabilidade mostram a existência, por parte

dos alunos, de intuições incorretas, erros de raciocínio e erros de compreensão e aplicação do conceito de probabilidade condicionada, erros bastante enraizados “que não são corrigidos simplesmente com uma aprendizagem formal da probabilidade” [2, p. 5], e por isso é necessário abordar os problemas de probabilidade condicional utilizando ferramentas adequadas [13].

O estudo anteriormente descrito, sobre a interpretação dos resultados das análises clínicas, é adequado para ser trabalhado no ensino secundário, principalmente por alunos do curso Ciências e Tecnologias onde, de forma mais realista, podem aplicar conhecimentos probabilísticos de acordo com o seu plano de estudos. De forma análoga, poderia ser trabalhado um estudo sobre investimentos, adequado a alunos do curso de Ciências Socioeconómicas. Seria uma boa forma de dar respostas aos alunos que procuram sempre uma justificação para o estudo das matérias curriculares, além de lhes permitir resolver problemas distintos dos jogos de azar relacionados com cartas, dados ou extrações de bolas, muito frequente nos manuais escolares.

No anexo A.1.4. podemos encontrar a programação referente ao problema das análises. Diferentes dos restantes problemas apresentados, em que o utilizador apenas varia os parâmetros n e $series$ a utilizar nas simulações de acordo com o número de experiências ou o número de séries que pretende realizar, a programação das análises clínicas pode ser aplicada a qualquer tipo de exame clínico, desde que sejam conhecidos os parâmetros prevalência da doença (p), sensibilidade da análise (s) e sua especificidade (e). Com esta programação o utilizador obtém os valores VP, FP, FN e VN da análise em estudo, os quais permitirão construir a tabela de contingência e complementar o estudo apresentado na subsecção 3.2.2. Poderão ainda alterar os parâmetros S e E (valores da sensibilidade e da especificidade) de forma a analisar as consequências de alterações dos valores como referenciado na subsecção 3.2.1. Este problema é adequado para ser utilizado na elaboração de trabalhos de exploração em grupo.

Na Figura 15 podemos observar os resultados obtidos na simulação de um teste a 50000 utentes, para uma determinada doença cuja prevalência é de 5%, com sensibilidade e especificidade de 95% e 75%, respetivamente.

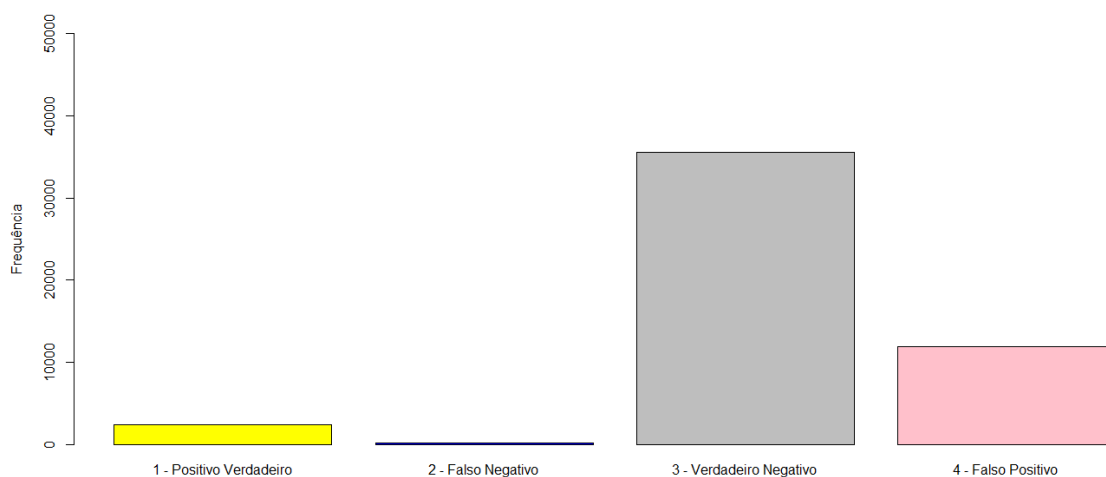


Figura 15 – Valores do VP, FP, FN e VN obtidos na aplicação de um teste a 50000 utentes.

Para além do gráfico, a programação em R apresenta como output os seguintes valores:

```
> sim4(50000,5,95,75)
"1 - Positivo Verdadeiro" "2427"
"2 - Falso Negativo=" "122"
"3 - Verdadeiro Negativo" "2427"
"4 - Falso Positivo" "11923"
"Total de utentes observados" "50000"
```

De forma a uma melhor exploração do conceito de probabilidade condicionada poderão ser determinadas as percentagens da prevalência, da sensibilidade e da especificidade observadas, bem como do valor preditivo positivo e do valor preditivo negativo observado e da acurácia do teste realizado. A compreensão destes valores condicionados permitem, por um lado, desmistificar a problemática referida na subsecção 2.2.5. sobre a restrição do espaço amostral no cálculo de probabilidades condicionadas e, por outro, concretizar os conceitos abstratos associados ao cálculo de probabilidades.

CAPÍTULO 4 – EXPERIÊNCIA PEDAGÓGICA

Elaborados os materiais em consonância com os princípios defendidos por investigadores da área, o projeto apresentado culminou com uma aplicação dos mesmos de modo a contribuir para uma melhor avaliação do impacto do uso da simulação no ensino das probabilidades, especificamente, no que se refere ao desenvolvimento das intuições probabilísticas.

Na realização da experiência pedagógica procurou-se, em primeiro lugar, avaliar se o uso de simuladores computacionais contribui para o desenvolvimento da correta percepção do comportamento de fenómenos aleatórios e, por outro lado, identificar quais os erros descritos na revisão da literatura que são mais comumente revelados pelos elementos da amostra do nosso estudo.

Neste capítulo serão apresentadas as opções metodológicas e a respetiva fundamentação bem como descritas as fases do estudo, atendendo ao contexto em que decorreu a experiência pedagógica e à natureza dos objetivos que foram definidos. São ainda apresentados e discutidos os resultados obtidos.

4.1. METODOLOGIA E PARTICIPANTES DO ESTUDO

Segundo Pontes [33], o conceito de investigação sobre a prática caracteriza-se pela existência de dois tipos principais de objetivos: alteração de algum aspeto da prática, quando existe reconhecimento da necessidade de mudança; e compreensão sobre a natureza dos problemas que afetam essa mesma prática, com vista ao desenvolvimento posterior de uma estratégia de ação. Ainda Richardson [34], citado por Pontes, frisa que a investigação sobre a prática **“não é conduzida para desenvolver leis gerais relacionadas com a prática educacional, e não tem como propósito fornecer a resposta a um problema. Em vez disso, os resultados sugerem novas formas de olhar o contexto e o problema e/ou possibilidades de mudanças na prática”** [33, p.7].¹⁰ Deste modo, pretendendo-se conhecer de que forma uma nova abordagem pedagógica (uso de simuladores) ajudará a solucionar um problema da própria prática enquanto

¹⁰ Negrito nosso.

docente de matemática (desenvolvimento de intuições probabilísticas), podemos afirmar que o estudo desenvolvido insere-se numa investigação sobre a prática, tendo-se optado por uma metodologia de natureza qualitativa. Essa opção é justificada pelo facto do estudo realizado revelar características que Bogdan e Biklen [9] identificam como sendo uma investigação qualitativa: o investigador é o instrumento principal (apesar de não ter sido realizada em aula, os participantes do estudo são alunos do ensino secundário e os dados foram recolhidos pelo investigador numa situação de contexto escolar); e é uma investigação descritiva onde se atribui mais importância ao processo e às perspectivas dos participantes do que simplesmente aos resultados ou produtos (os documentos produzidos pelos alunos foram analisados e a interpretação das justificações, que foram escritas nas resoluções dos problemas, constitui o instrumento primordial de análise). Não se pretende intervir, mas unicamente dar a conhecer a situação em estudo, tal como surge. Para tal, é realizada uma descrição precisa, sistemática e completa dos factos ocorridos durante a experiência pedagógica.

Os materiais apresentados são relativos a conteúdos programáticos do secundário. Deste modo, uma vez que não se pretende ensinar novos conceitos, mas unicamente explorar noções previamente aprendidas, apesar de, quase certamente, não terem sido exploradas com recurso à simulação computacional, optou-se por concretizar a experimentação pedagógica dos materiais após a realização do Exame Nacional de Matemática A. Segundo Bogdan e Biklen [9], para um primeiro trabalho de investigação a escolha de fontes de dados deve ser de fácil acesso do investigador. Assim, devido ao facto de a autora neste momento não se encontrar a lecionar em nenhuma escola, o presente estudo foi realizado, após prévia autorização, num único dia com duração aproximada de três horas, na Escola Superior de Tecnologias e Gestão do Instituto Politécnico de Leiria, a um conjunto de estudantes do 12.º ano que, no presente ano letivo, frequentaram a disciplina de Matemática A. Foi constituída uma amostra de conveniência, a partir do convite a alunos que no 10.º ano tinham sido seus discentes, tendo resultado desse convite a comparência voluntária de oito indivíduos, que manifestaram disponibilidade e vontade em participar no estudo.

4.2. FASES DO ESTUDO

Tendo em consideração os objetivos estabelecidos na investigação, as atividades desenvolvidas foram repartidas em três fases distintas. Para preservar o anonimato dos participantes e, simultaneamente, conseguir identificar os materiais elaborados nas diferentes fases como sendo do mesmo indivíduo, iniciou-se a sessão por se atribuir, aleatoriamente, um número a cada aluno (através da escolha, por parte de cada estudante, de uma bola de entre bolas numeradas de 1 a 90), que se designou por código de aluno.

Numa primeira fase, os participantes acederam a um questionário *online* elaborado a partir da ferramenta *Google Drive* (conforme anexo A.2.1.), constituído por quatro grupos. No primeiro grupo pretendeu-se obter algumas informações de carácter pessoal, que permitissem fazer uma pequena caracterização da amostra e correlacionar algumas variáveis como o desempenho escolar e o raciocínio probabilístico. No segundo grupo, pretendeu-se aferir sobre os conhecimentos dos participantes relativamente a alguns conceitos importantes de probabilidades, que se supõe estarem adquiridos, e que iriam ser trabalhados ao longo da sessão. Além disso, pretendeu-se averiguar se o uso de simuladores já fazia parte da realidade de ensino destes alunos. No terceiro grupo, a partir de um mesmo contexto (extração de cartas de um baralho), foram apresentadas diferentes experiências e acontecimentos aleatórios com o objetivo de avaliar de que forma as falácias e erros de raciocínios, descritos na literatura, estavam contempladas nas repostas dadas pelos participantes. Optou-se por um contexto familiar ao aluno (frequente em manuais escolares e em exames nacionais) de modo a garantir que as possíveis dificuldades reveladas não seriam provenientes da incorreta interpretação ou contextualização das situações apresentadas. Cada um dos itens foi elaborado para avaliar um certo erro ou conhecimento probabilístico e escolhido de modo a que não envolvessem a necessidade de usar fórmulas combinatórias. Apesar das perguntas elaboradas serem em formato de escolha múltipla, foi pedido aos alunos que justificassem as escolhas feitas, em suporte papel, podendo a mesma ser dada por palavras, esquemas ou cálculos. Procurou-se, desta forma, que os alunos não ficassem condicionados a um só método de resolução e, simultaneamente, obter informações para uma análise mais detalhada sobre o tipo de pensamento e abordagem usada em cada questão, por cada indivíduo. Ainda segundo Fernandes [15], as justificações das repostas em itens de escolha múltipla revelam-se meios eficazes para identificar e

compreender concepções erradas e para clarificar a compreensão dos alunos que selecionaram a alternativa correta.

Por fim, no último grupo, foram apresentados dois problemas históricos (o Paradoxo da Caixa de Bertrand e o dilema de Monty Hall, retratados no capítulo 3). De forma análoga ao grupo três, também neste grupo os alunos deverão justificar as suas respostas. Esta primeira fase, sem recurso à simulação, corresponde a aplicação de um pré-teste cujo principal objetivo é avaliar as intuições probabilísticas dos alunos.

Na segunda fase da sessão pretendeu-se, essencialmente, dar a conhecer simuladores construídos a partir do *software* R e os conhecimentos necessários à sua correta utilização na resolução de problemas sobre probabilidades. Iniciou-se com uma breve apresentação em *PowerPoint* (ver anexo A.2.6.) contendo, entre outros assuntos, a origem da probabilidade e alguns conceitos úteis à sua compreensão. De seguida, apresentou-se o jogo “Passe-dix” que serviu de mote para iniciar o uso de simuladores computacionais. De modo a realçar a necessidade de utilização deste recurso, apresentou-se a solução incorreta que na altura foi apresentada por alguns matemáticos conhecidos e discutiu-se o erro que lhe estava subjacente (a consideração de que os acontecimentos seriam equiprováveis). Por fim, recorreu-se ao simulador. Os alunos foram incentivados a realizar variações dos parâmetros para obter os diferentes gráficos e, a partir da participação coletiva, chegar à resposta correta do problema.

Numa terceira fase, os participantes voltam a aceder a um questionário (ver anexo A.2.2.), constituído por dois grupos. No primeiro grupo, pretende-se que os alunos utilizem os simuladores construídos para analisar os dois problemas históricos, apresentados no questionário da primeira fase. O objetivo é comparar as respostas dadas nos testes antes e após o uso de simuladores e, assim, compreender qual o papel dos simuladores utilizados na concretização das tarefas propostas e na obtenção da solução correta.

No segundo grupo, envolvendo as mesmas situações probabilísticas retratadas no terceiro grupo da primeira fase (primeiro inquérito) foram apresentadas diferentes questões, diferindo-se apenas no contexto das experiências aleatórias trabalhadas (extração de bolas), ainda que tal contexto seja igualmente familiar aos alunos. Os itens, abordando os erros e falácias do primeiro questionário, foram reordenados de modo a não viciar as suas respostas. O principal objetivo é observar se as respostas dadas pelos alunos indiciam os mesmos erros cometidos no primeiro questionário, após a revisão dos conteúdos realizada na segunda fase da sessão. Pretende-se aferir se a proveniência

dos erros são resultado do esquecimento ou falta de conhecimento ou se os erros estão intrínsecos ao raciocínio probabilístico dos alunos, sendo o ensino formal insuficiente para superá-los, tal como defendido por Batanero [5].

Por último, foi solicitado que posteriormente os alunos acessem a um breve questionário relativamente à sua satisfação sobre a experiência pedagógica realizada (ver Anexos A.2.3. e A.2.4.).

De forma a clarificar quaisquer dúvidas no preenchimento dos inquéritos durante as várias fases, foi disponibilizada uma ficha de apoio (ver Anexo A.2.5.), distribuída no início da experiência pedagógica, que serviu de guião e permitiu um melhor acompanhamento por parte dos estudantes.

4.3. APRESENTAÇÃO DOS DADOS E DISCUSSÃO DOS RESULTADOS

A amostra deste estudo é constituída por oito alunos (sendo seis elementos do género feminino e dois elementos masculinos), com idades compreendidas entre os 17 e os 18 anos, que frequentaram o décimo segundo ano, pela primeira vez, no ano letivo 2013/2014 no curso de Ciências e Tecnologias e, portanto, cujo desenho curricular contém a disciplina de Matemática A. Apesar de frequentarem todos a mesma escola, são alunos provenientes de quatro turmas diferentes. Os resultados obtidos na avaliação final da disciplina variaram entre os 9 e os 18 valores, com uma média de 12,5 valores e um desvio padrão de 3,1.

Dos inquiridos, apenas três afirmaram já terem realizado, em aula, simulações de experiências aleatórias com recurso à calculadora e a materiais manipuláveis (como por exemplo cartas, moedas, dados ou bolas) e, desses, apenas um referiu o uso do computador. Apesar das indicações metodológicas dos programas nacionais e dos estudos já realizados mencionarem a importância do uso das tecnologias no ensino, e especificamente no estudo das probabilidades, essa realidade ainda não faz parte do ensino de, pelo menos, algumas escolas ou turmas. Como já mencionado, o recurso à simulação é importante para abordar a definição frequencista de probabilidade, o que vem justificar algumas incorreções nas respostas que foram dadas por estes alunos quando solicitados a definir esse conceito. Atendendo ao termo frequencista, quatro alunos procuraram com as suas respostas relacionar a probabilidade com a frequência de ocorrência dos acontecimentos, mas nenhum fez qualquer tipo de referência ao número

de repetições na realização da experiência em causa. Curiosamente, o único aluno que afirmou ter usado em aula o computador para simular experiências aleatórias, aquando da definição frequencista de probabilidade respondeu “*Acho que nunca fui abordado sobre esta definição*”. Não sabendo nós qual o enfoque dado ao uso do simulador, podemos concluir, pelo menos, que não foi suficiente para o aluno reter o conceito subjacente à simulação. Pelo contrário, no que diz respeito à noção clássica de probabilidade, esta foi corretamente definida por todos os alunos embora só dois tenham referido que a sua aplicação exige a equiprobabilidade dos acontecimentos elementares. Mais corretamente, a aplicabilidade da Lei de Laplace ao cálculo da probabilidade de um acontecimento requer que o espaço de resultados associado à experiência aleatória seja constituído por um número finito de elementos, todos eles com igual possibilidade de se realizar. O facto de nenhum aluno ter referido a necessidade de os acontecimentos elementares serem em número finito não é de estranhar uma vez que, ao nível de critérios de correção dos exames nacionais, quando é solicitado aos alunos que enunciem a regra de Laplace esta condição não é referida/considerada.

“enunciar a regra de Laplace: a probabilidade de um acontecimento é o quociente entre o número de casos favoráveis e o número de casos possíveis, quando estes são todos equiprováveis (ou equivalente);” – critérios específicos de classificação do exame nacional de matemática da 1ª fase de 2010.

No que concerne à definição de acontecimentos independentes, cinco dos oito alunos responderam corretamente. Das respostas erradas, existem dois alunos que trocam a definição de acontecimentos independentes por acontecimentos disjuntos, erro mencionado na subsecção 2.4. do capítulo 2: “*Significa que dois acontecimentos não se intersejam entre si, ou seja, não existe nada em comum entre estes dois acontecimentos, daí a designação de independentes.*”

De seguida, são apresentadas e analisadas as respostas obtidas através dos dois questionários aplicados, respetivamente, na primeira e terceira parte da sessão. Ambos os instrumentos, o pré e o pós-teste, são constituídos por sete questões em formato de escolha múltipla, tendo sido pedido, em cada uma delas, que fosse apresentado uma justificação para a resposta escolhida. Os dois questionários envolvem as mesmas situações probabilísticas recorrendo a dois contextos diferentes mas igualmente familiares aos alunos: extração de cartas e extração de bolas.

4.3.1. FALÁCIAS E ERROS PROBABILÍSTICOS

No grupo III do primeiro questionário, cada um dos sete itens apresentados foi elaborado tendo em conta alguns dos erros probabilísticos que são frequentemente identificados nos raciocínios dos discentes. O contexto retratado corresponde à extração de cartas de um baralho:

Um baralho de cartas completo é constituído por 52 cartas, repartidas em 4 naipes (Espadas, Copas, Ouros e Paus). Em cada naipe há 13 cartas: um Ás, três figuras (Rei, Dama e Valete) e mais 9 cartas (do Dois ao Dez).

No grupo II do segundo questionário, a partir de um contexto de extração de bolas, são abordados os mesmos erros probabilísticos:

Uma caixa contém 30 bolas: dez bolas vermelhas numeradas de 1 a 10 e vinte bolas pretas numeradas de 1 a 20.

De seguida, apresentam-se as respostas dadas em cada item, expressas em percentagem, encontrando-se a negrito a alínea correspondente à resposta correta. Essa informação é complementada com uma análise dos raciocínios evidenciados pelos alunos na resolução/justificação apresentada nos diferentes itens, procurando-se efetuar o seu enquadramento com os erros probabilísticos descritos na secção 2.2. Sempre que se justifique, por se tratar de exemplos significativos tanto ao nível de respostas corretas como incorretas, serão reproduzidas as estratégias de resolução seguidas pelos alunos.

4.3.1.1. O RACIOCÍNIO PROPORCIONAL

Inquérito I

Item 8. Considera que as 52 cartas foram divididas em 2 grupos distintos: um grupo A constituído apenas pelas 13 cartas do naipe de copas e um grupo B constituído pelas 39 cartas restantes. Retira-se uma carta de cada um dos grupos constituídos anteriormente. Em qual dos 2 grupos, A ou B, é mais provável obter uma carta Ás?

a)	Grupo A.	0%
b)	Grupo B.	13%
c)	São igualmente prováveis.	88%

(Respostas corretas com justificações válidas: 7 em 7)

Inquérito II

Item 9. Considere que as 30 bolas da caixa foram repartidas por cor, em 2 grupos distintos: um grupo A contendo as bolas vermelhas numeradas de 1 a 10 e um grupo B contendo as bolas pretas numeradas de 1 a 20. Retira-se uma bola, ao acaso, de cada um dos grupos A e B. Em qual dos 2 grupos é mais provável obter uma bola com número par?

a)	Grupo A.	0%
b)	Grupo B.	25%
c)	São igualmente prováveis.	75%

(Respostas corretas com justificações válidas: 6 em 6)

Desde o 9.º ano de escolaridade que os alunos estão familiarizados com o uso da Lei de Laplace no cálculo de probabilidades que estabelece, sob a condição dos resultados serem igualmente possíveis e em número finito, que a probabilidade de um acontecimento é dada pelo quociente entre o número de casos favoráveis à sua ocorrência e o número de casos possíveis, estando por isso estritamente ligada ao raciocínio proporcional.

Nos itens 8 e 9 é necessário comparar em qual dos dois grupos, A e B, o acontecimento pretendido é mais frequente ocorrer, atendendo à proporção de resultados favoráveis na população em causa. Uma vez que a pergunta envolve um conhecimento considerado elementar no estudo das probabilidades e dada a nossa amostra ser constituída por finalistas do ensino secundário, seria expectável que a percentagem de respostas erradas fosse muito baixa, o que de facto se verificou.

A Lei de Laplace foi a estratégia utilizada para responder à questão: foi escolhido o grupo com a maior razão entre o número de cartas de “Ás” existentes e o número total de cartas. Para o apuramento do número de casos favoráveis e o número de casos possíveis, apenas um aluno (Figura 16) recorreu à análise combinatória tendo os restantes alunos (Figura 17) apresentado o valor absoluto da contagem.

8. Justificação

52 cartas \rightarrow 2 grupos $\left\{ \begin{array}{l} A: 13 \text{ cartas de copas} \\ B: 39 \text{ cartas restantes} \end{array} \right.$

A: Sair um ás.

Para A: $\frac{{}^1C_1}{{}^{13}C_1} = \frac{1}{13}$

Para B: $\frac{{}^3C_1}{{}^{39}C_1} = \frac{3}{39} = \frac{1}{13}$

Como $\frac{1}{13} = \frac{3}{39}$ e $\frac{1}{13}$ é mais provável obter um ás

$\frac{1}{13} = \frac{3}{39} \Rightarrow$ igualmente provável

Figura 16 – Resolução do aluno com o código 58.

8. Justificação

$P(A) = \frac{1}{13}$ $P(B) = \frac{3}{39} = \frac{1}{13}$

$P(A) = P(B)$

Os acontecimentos são igualmente prováveis

Figura 17 – Resolução do aluno com o código 8.

O aluno que deu a resposta incorreta (aluno com o código 85) considerou que, pelo facto dos cardinais dos casos favoráveis e possíveis do Grupo B serem maiores do que o Grupo A, existiria maior probabilidade de sair uma carta Ás no Grupo B. Não considerou, por isso, à existência de proporcionalidade, que torna as duas frações equivalentes.

No segundo questionário, o número de respostas erradas aumentou. O aluno com o código 85 manteve o mesmo tipo de erro (escolha do grupo que contém mais bolas números pares, sem atender ao número de elementos do espaço amostral) e o aluno com o código 8, que no primeiro questionário apresentou uma justificação válida, revelou no segundo questionário os mesmos erros do aluno 85.

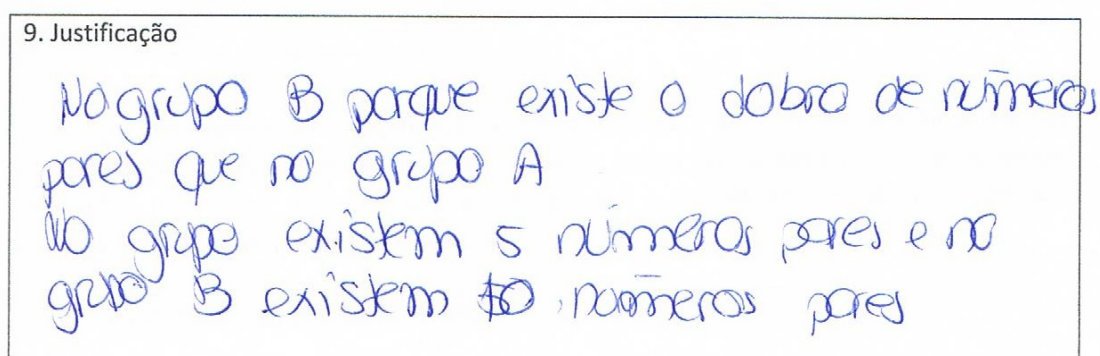


Figura 18 – Resolução do aluno com o código 8.

4.3.1.2. EXTRAÇÃO COM REPOSIÇÃO

Inquérito I

Item 9. Considera a experiência aleatória que consiste em retirar 2 cartas do baralho, com reposição. Qual dos acontecimentos tem maior probabilidade de ocorrer?

a) Tirar 2 cartas de copas.	0%
b) Tirar uma carta de copas e depois uma carta de ouros.	38%
c) São igualmente prováveis.	63%

(Respostas corretas com justificações válidas: 3 em 5)

Inquérito II

Item 4. Considera a experiência aleatória que consiste em retirar 2 bolas da caixa, com reposição. Qual dos acontecimentos tem maior probabilidade de ocorrer?

a) Tirar 2 bolas pares.	0%
b) Tirar uma bola par e depois uma bola ímpar.	38%
c) São igualmente prováveis.	63%

(Respostas corretas com justificações válidas: 4 em 5)

As respostas dadas e os erros revelados pelos alunos, quer no primeiro quer no segundo questionário, são similares e a maior fonte dessa imprecisão, que os levou a afirmar a equiprobabilidade dos acontecimentos, foi resultado de se considerar a existência de ordem na saída das cartas/bolas.

9. Justificação

$$P(\text{tirar 2 copas}) = \frac{13}{52} \times \frac{13}{52} = \frac{169}{2704} = \frac{1}{16}$$
$$P(\text{tirar 1 copas e 1 ouros}) = \frac{13}{52} \times \frac{13}{52} \times 2 = \frac{1}{8}$$
$$\frac{1}{8} > \frac{1}{16} \Rightarrow P(\text{1 copas e 1 ouros}) > P(\text{2 copas})$$

Figura 19 – Resolução do aluno com o código 22.

4.3.1.3. FALÁCIA DA CONDICIONAL TRANSPOSTA

Inquérito I

Item 10. Considera a experiência aleatória que consiste em retirar uma carta do baralho. Qual dos acontecimentos é mais provável?

a) Sair Rei sabendo que a carta retirada é de Ouros.	13%
b) Sair Ouros sabendo que a carta retirada é um Rei.	75%
c) Ambos os acontecimentos são equiprováveis	13%

(Respostas corretas com justificações válidas: 6 em 6)

Item 14. Extrai-se uma carta ao acaso do baralho. Seja A o acontecimento "Sai uma carta de Espadas" e B o acontecimento "Sai uma Dama". Qual dos acontecimentos é mais provável?

a) P(A B).	88%
b) P(B A).	0%
c) São equiprováveis.	13%

(Respostas corretas com justificações válidas: 6 em 7)

Inquérito II

Item 7. Considera a experiência aleatória que consiste em retirar uma bola da caixa. Qual dos acontecimentos é mais provável?

a) Sair uma bola com o n.º 2, sabendo que a bola retirada é vermelha.	25%
b) Sair uma bola vermelha, sabendo que a bola retirada tem o n.º 2.	63%
c) Ambos os acontecimentos são equiprováveis.	13%

(Respostas corretas com justificações válidas: 4 em 5)

Item 3. Extrai-se uma bola ao acaso da caixa. Seja A o acontecimento "Sai uma bola com um n.º 10" e B o acontecimento "Sai uma bola preta". Qual dos acontecimentos é mais provável?

a) P(A B).	13%
b) P(B A).	88%
c) São equiprováveis.	0%

(Respostas corretas com justificações válidas: 5 em 7)

Os itens 10 e 14 do primeiro questionário, retratam a falácia da condicional transposta descrita na subsecção 2.2.1. do Capítulo 2. Na elaboração das duas questões, de modo a aferir se os alunos aderem mais a este erro quando é utilizada a linguagem corrente em vez da linguagem matemática, foram utilizadas os dois tipos de notações e, estrategicamente, não foram colocadas em itens consecutivos.

Pelo que podemos constatar foram poucos os alunos que não responderam corretamente ao problema apresentado, quer na forma de linguagem corrente quer na forma de

linguagem matemática. No entanto, apesar de ter havido uma ligeira melhoria da percentagem de respostas corretas nos itens onde é utilizada a notação simbólica, continuam a ser apresentados cálculos incorretos na determinação das probabilidades condicionadas. Pela análise das justificações que não estavam corretas, os erros apresentados são resultado de uma ineficiente determinação/restrição do espaço amostral no cálculo da probabilidade condicionada e não de uma troca dos acontecimentos condicionado e condicionante. No nosso estudo, não se verificou a adesão à falácia da condicional transposta. Tal como no primeiro inquérito, os erros revelados no segundo questionário são resultado de uma incorreta determinação das probabilidades condicionadas, principalmente em virtude de uma ineficiente determinação/restrição do espaço amostral.

3. Justificação

$$P(A) = \frac{2}{30}$$

A: "Sai uma bola com um nº 10"

$$P(B) = \frac{1}{20}$$

B: "Sai uma bola preta"

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{20}}{\frac{1}{20}} = \frac{20}{21} \leftarrow \text{Acontecimento mais provável}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{21}}{\frac{2}{30}} = \frac{5}{7}$$

Inquérito I

Item 11. Retiram-se, sucessivamente e sem reposição, 2 cartas do baralho. Qual dos acontecimentos é mais provável?

- | | |
|--|------------|
| a) Sair figura na 2. ^a extração, sabendo que a 1. ^a carta extraída foi uma figura. | 25% |
| b) Sair figura na 1. ^a extração, sabendo que a 2. ^a carta extraída foi uma figura. | 25% |
| c) Ambos os acontecimentos têm a mesma probabilidade de ocorrer. | 50% |

(Respostas corretas com justificações válidas: 3 em 4)

Inquérito II

Item 6. Retiram-se, sucessivamente e sem reposição, 2 bolas da caixa. Qual dos acontecimentos é mais provável?

- | | |
|--|------------|
| a) Sair número par na 2. ^a extração, sabendo que a 1. ^a bola extraída foi par. | 25% |
| b) Sair número par na 1. ^a extração, sabendo que a 2. ^a bola extraída foi par. | 25% |
| c) Ambos os acontecimentos têm a mesma probabilidade de ocorrer. | 50% |

(Respostas corretas com justificações válidas: 1 em 4)

No item 11, apenas metade dos inquiridos conseguiram dar a resposta correta e, entre eles, houve um aluno que não justificou convenientemente a sua opção. Tal como referido na subsecção 2.2.2., os alunos da nossa amostra também revelaram muitas dificuldades na resolução de problemas de probabilidade condicionada que envolvam a inversão do eixo do tempo na ocorrência dos acontecimentos. O erro mais frequente foi relativo ao cálculo da probabilidade da alínea b). A alínea a) é uma situação natural e compatível com o eixo temporal e, por isso, menos suscetível de criar as dificuldades sentidas na alínea b) que apresenta uma situação contrária à ordem temporal. Neste caso, os alunos basearam as suas respostas apenas na composição da urna no início da experiência (determinaram simplesmente a probabilidade de sair figura na 1.^a extração), sem ter em consideração a informação relevante fornecida pelo acontecimento condicionante acerca da segunda extração efetuada.

11. Justificação

$$P(\text{sair figura na 2.ª extração sabendo que foi retirada figura na 1ª extração}) = \frac{11}{51}$$
$$P(\text{sair figura na 1ª extração sabendo que foi retirada figura na 2ª extração}) = \frac{12}{52} = \frac{3}{13}$$

Figura 21 – Resolução do aluno com o código 6.

No segundo questionário, apesar de se manter inalterável a percentagem de respostas corretas, os erros apresentados são em maior número e apenas um aluno conseguiu determinar e justificar convenientemente a sua resposta. Destacamos a resposta do aluno

com o código 58 onde é evidente a recusa em aceitar que um acontecimento condicione outro que ocorra anteriormente, argumentando que a extração da segunda bola, cuja realização é posterior, não pode influenciar a extração da primeira bola.

6. Justificação
 A: Sair por 1ª extração
 B: Sair por na 2ª extração
 → A bola que saiu na 2ª extração não influencia a da 1ª extração

Figura 22 – Resolução do aluno com o código 58.

Salienta-se ainda que dois dos alunos que apresentaram justificação válida para o item 11 do primeiro questionário, no segundo questionário selecionaram a opção correta mas não apresentaram qualquer justificação. Possivelmente, essa situação é resultado de um cansaço por parte dos alunos dado encontrarem-se na reta final da experiência pedagógica, iniciada algumas horas antes.

4.3.1.5. ACONTECIMENTOS INDEPENDENTES

Inquérito I

Item 12. Extrai-se uma carta ao acaso do baralho. Seja A o acontecimento "Sai uma carta de Ouros" e B o acontecimento "Sai um Ás". Os acontecimentos A e B são independentes?

a) Não são independentes porque no baralho existe um Ás de Ouros.	50%
b) Não são independentes porque o acontecimento A é mais provável que o acontecimento B.	13%
c) São independentes porque $P(A B) = P(A)$.	38%
d) São independentes porque $P(A B) = P(B)$.	0%

(Respostas corretas com justificações válidas: 2 em 3)

Inquérito II

Item 8. Extrai-se uma bola ao acaso da caixa. Seja A o acontecimento "Sai uma bola par" e B o acontecimento "Sai uma bola vermelha". Os acontecimentos A e B são independentes?

- | | |
|--|------------|
| a) Não são independentes porque na caixa existem bolas vermelhas pares. | 25% |
| b) Não são independentes porque o acontecimento A é mais provável que o acontecimento B. | 0% |
| c) São independentes porque $P(A B) = P(B)$. | 0% |
| d) São independentes porque $P(A B) = P(A)$. | 75% |

(Respostas corretas com justificações válidas: 3 em 6)

O item 12 retrata o problema descrito na subsecção 2.2.4. do capítulo 2 relativa à associação que é feita pelos alunos entre acontecimentos independentes e acontecimentos disjuntos.

A partir da análise dos resultados apresentados relativos ao primeiro questionário, este erro é praticado por metade dos alunos da nossa amostra, que consideraram que os acontecimentos A e B não são independentes por serem compatíveis.

No segundo questionário podemos verificar que a percentagem de alunos que selecionaram a resposta correta aumentou substancialmente. No entanto, analisando as suas justificações, podemos verificar que os três alunos que alteraram corretamente a sua resposta do questionário I para o questionário II, limitaram-se a provar o resultado $P(A|B) = P(A)$, no caso de A e B serem independentes. Não apresentaram cálculos para a probabilidade condicionada $P(A|B)$ nem para $P(A)$, no contexto da experiência apresentada. Esta situação pode dever-se ao facto de, antes de ser aplicado o segundo questionário, numa segunda fase, ter-se feito uma breve revisão do que são acontecimentos independentes e apresentado o resultado anteriormente referido.

8. Justificação

Para os acontecimentos serem independentes

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

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

Figura 23 – Resolução do aluno com o código 85.

4.3.1.6. FALÁCIA DA CONJUNÇÃO

Inquérito I

Item 13. Considera que o João vai jogar um jogo que consiste em retirar, sucessivamente e sem reposição, 5 cartas do baralho. Se no final das 5 cartas retiradas saírem 3 ases, o João ganha o jogo. Qual dos seguintes acontecimentos é mais provável ocorrer?

- | | |
|--|-----|
| a) A primeira carta retirada é um ás. | 50% |
| b) A primeira carta retirada é um ás mas o João perde o jogo. | 38% |
| c) Ambos os acontecimentos têm a mesma probabilidade de ocorrer. | 13% |

(Respostas corretas com justificações válidas: 1 em 4)

Inquérito II

Item 5. Considera que o João vai jogar um jogo que consiste em retirar, sucessivamente e sem reposição, 5 bolas da caixa. Se no final das 5 bolas retiradas saírem 3 bolas vermelhas, o João ganha o jogo. Qual dos seguintes acontecimentos é mais provável ocorrer?

- | | |
|--|-----|
| a) A primeira bola retirada é vermelha. | 50% |
| b) A primeira bola retirada é vermelha mas o João perde o jogo. | 50% |
| c) Ambos os acontecimentos têm a mesma probabilidade de ocorrer. | 0% |

(Respostas corretas com justificações válidas: 0 em 4)

Quer no primeiro quer no segundo questionário, a adesão à falácia da conjunção é evidente devido à percentagem de respostas obtidas na alínea b). Esse erro é resultante do facto do acontecimento “O João perde o jogo” ter uma grande probabilidade de ocorrer comparativamente ao acontecimento “A primeira carta retirada é um ás” e, por isso, a interseção dos dois acontecimentos é considerada mais provável que o último acontecimento. Fazendo uma análise mais detalhada das respostas, alguns alunos omitiram a interseção e no cálculo da probabilidade relativa à alínea b) apenas consideraram o acontecimento “O João perde o jogo”.

13. Justificação

Apesar de todas as cartas terem igual probabilidade de sair como há um menor nº de ases e é mais provável retirar uma carta não ás, logo é mais provável que o jogo perca o jogo

Figura 24 – Resolução do aluno com o código 58.

No que se refere às justificações dadas pelos alunos que optaram pela alínea a), a maioria das respostas continham incorreções no cálculo da probabilidade conjunta. Um dos erros apresentados está relacionado com a falta de consideração de ordem na saída das cartas de ases e, por consequência, não consideraram que a ordem de realização dos acontecimentos conduz a possibilidades distintas.

5. Justificação

$$P(19 \text{ vermelha}) = \frac{10}{30} = \frac{1}{3}$$

$$P(19 \text{ vermelha e João perde}) = \frac{10}{30} \times \frac{9}{29} \times \frac{20}{28} \times \frac{19}{27} \times \frac{18}{26}$$

$$+ \frac{10}{30} \times \frac{20}{29} \times \frac{19}{28} \times \frac{18}{27} \times \frac{17}{26}$$

$$\approx 0,104$$

Figura 25 – Resolução do aluno com o código 22.

4.3.1.7. CONSIDERAÇÕES FINAIS SOBRE A ADESÃO AOS ERROS PROBABILÍSTICOS

A partir da análise do conteúdo dos argumentos dados pelos estudantes que sustentam as suas decisões relativamente às questões do primeiro e segundo questionário, foram elaboradas as tabelas de contingência que relacionam as perguntas similares dos dois questionários. A cada justificação, por pergunta, foi atribuída um valor de 0 a 2: zero às perguntas em que os alunos não justificam ou fazem-no de forma incorreta, um às perguntas em que os alunos apresentam uma justificação parcialmente válida contendo, por exemplo, erros de cálculo e finalmente dois às perguntas que contêm uma justificação correta para o problema apresentado. Apresentam-se de seguida as tabelas resultantes dessa análise em que se representou por P o número da pergunta e por Q I e Q II os questionários 1 e 2, respetivamente.

		P 9 - Q II					P 4 - Q II					P 7 - Q II			
		0	1	2			0	1	2			0	1	2	
P 8-Q I	0	1	--	--	P 9-Q I	0	4	--	1	P 10-Q I	0	2	--	--	
	1	--	--	--		1	--	--	--		1	--	--	--	--
	2	1	--	6		2	--	--	3		2	--	2	4	
		P 6 - Q II					P 8 - Q II					P 5 - Q II			
		0	1	2			0	1	2			0	1	2	
P 11-Q I	0	5	--	--	P 12-Q I	0	5	--	1	P 13-Q I	0	6	--	--	
	1	--	--	--		1	--	--	--		1	--	1	--	
	2	2	--	1		2	--	--	2		2	--	1	--	
		P 3 - Q II					P 3 - Q II								
		0	1	2			0	1	2						
P 14-Q I	0	2	--	--											
	1	--	--	--											
	2	1	--	5											

Tabela 1 – Tabelas de contingência com os dados obtidos nos dois questionários

Da observação dos dados registados nas tabelas, poderemos rapidamente verificar se houve ou não melhorias do primeiro para o segundo inquérito: os valores da diagonal principal são relativos a resultados iguais nos dois questionários; os valores acima da diagonal principal são relativos a resultados nos quais se verifica uma melhoria

relativamente ao primeiro questionário; e, finalmente, os valores abaixo da diagonal principal são relativos a resultados nos quais se verifica ter-se piorado relativamente ao primeiro questionário.

Podemos então concluir que, maioritariamente, houve coerência nas respostas dadas pelos alunos nos dois questionários, evidenciando que as respostas corretas e os erros persistiram no pré e no pós teste. No entanto, comparando os resultados obtidos nos dois testes, e uma vez que encontramos mais valores abaixo da diagonal principal do que acima dela, os alunos revelaram melhor performance no primeiro questionário devido principalmente a não apresentarem qualquer tipo de justificação em várias perguntas no segundo questionário.

Esta situação pode ser justificada pelo facto de os alunos revelarem algum cansaço ao responder ao segundo questionário, dado já terem decorrido duas horas e meia desde o iniciar da sessão. Além disso, podemos realçar que, entre os alunos que escolheram a resposta correta, um número considerável apresentou justificações incorretas ou não apresentou justificações (situação mais frequente no segundo questionário). Como exceção temos as perguntas relativas ao raciocínio proporcional (P8 – QI e P9 – QII) e à falácia da condicional transposta com recurso à linguagem simbólica (P14 – QI e P3 – QII).

De seguida, foi elaborada a Tabela 2 com o objetivo de apurar a existência de correlação entre as variáveis desempenho escolar e raciocínio probabilístico. Na tabela constam as notas obtidas pelos alunos à disciplina de Matemática no final do 3.º período do corrente ano bem como o registo de respostas corretas (identificadas com o símbolo \surd) e erradas (identificadas com o símbolo \times) por cada item respondido nos dois questionários.

Da análise da tabela, os dados sugerem que existe uma associação positiva entre o desempenho escolar na disciplina de Matemática e o número de repostas corretas, na medida em que se pode associar melhores resultados nos inquéritos a níveis escolares superiores. A partir do *package RCommander* do *software R* podemos calcular o coeficiente de associação que é aproximadamente 0,79.

```
> cor(Exp1[,c("nível","número")], use="complete")
      nível número
nível 1.0000000 0.7882408
número 0.7882408 1.0000000
```

	9val.	9val.	11val	11val	13val	14val	15val	18val	N.º respostas corretas
Item 8/ 9	x / x	√ / √	√ / √	√ / x	√ / √	√ / √	√ / √	√ / √	7 / 6
Item 9/ 4	x / x	√ / √	x / x	√ / √	√ / √	√ / √	√ / √	x / x	5 / 5
Item 10/ 7	x / x	√ / x	√ / √	x / x	√ / √	√ / √	√ / √	√ / √	6 / 5
Item 11/ 6	x / x	x / x	x / x	√ / √	√ / √	√ / √	x / x	√ / √	4 / 4
Item 12/ 8	x / x	x / √	x / x	√ / √	x / √	√ / √	x / √	√ / √	3 / 6
Item 13/ 5	√ / √	x / x	x / x	x / x	√ / √	x / x	√ / √	√ / √	4 / 4
Item 14/ 3	√ / √	√ / √	√ / √	x / √	√ / x	√ / √	√ / √	√ / √	7 / 7
respostas corretas	2 / 2	4 / 4	3 / 3	4 / 4	6 / 6	6 / 6	5 / 6	6 / 6	

Tabela 2 – Associação entre as variáveis desempenho escolar e raciocínio probabilístico.

De entre as questões de probabilidade condicionada, destacam-se dificuldades mais acentuadas nos itens relativos à falácia de inversão do eixo temporal e à falácia da conjunção, não se tendo verificado a adesão à falácia da condicional transposta. Destaca-se também uma considerável melhoria do primeiro para o segundo questionário no item relativo aos acontecimentos independentes. Devido à justificação apresentada por alguns alunos se limitar a provar o resultado $P(A|B) = P(A)$ (no caso de A e B serem independentes), sem confirmação que tal igualdade é verificada no problema retratado, sugere que essa melhoria é resultado de se ter mencionado a propriedade antes da aplicação do pós teste, não tendo havido uma interiorização da sua aplicação.

4.3.2. RESOLUÇÃO DE PROBLEMAS HISTÓRICOS A PARTIR DA SIMULAÇÃO

Passemos a análise das respostas dadas pelos alunos aos dois problemas históricos (o Paradoxo da Caixa de Bertrand e o Paradoxo de Monty Hall) e comparação das respostas obtidas no primeiro questionário, onde os alunos não acederam à simulação, e no segundo questionário, com utilização do programa R, através do qual os alunos visualizaram a simulação dos problemas em análise.

O Paradoxo da Caixa de Bertrand: Temos três caixas e cada caixa tem duas gavetas, cada uma com uma moeda: uma caixa contém duas moedas de ouro, a outra caixa duas moedas de prata e a terceira caixa uma moeda de cada tipo. Depois de se escolher uma caixa ao acaso, abre-se uma gaveta de forma aleatória verificando-se que esta contém uma moeda de ouro. Qual é a probabilidade de que a moeda que está na outra gaveta da mesma caixa também seja de ouro?

- a) É a mesma de se obter uma moeda de prata.
- b) É maior do que se obter uma moeda de prata.
- c) É menor do que se obter uma moeda de prata.

Podemos verificar que relativamente ao Paradoxo da Caixa de Bertrand todos consideraram, erradamente, que a probabilidade da moeda, que está na outra gaveta da mesma caixa, seja também de ouro é a mesma de se obter uma moeda de prata. Dos oito alunos, seis mencionaram que essa probabilidade seria de $\frac{1}{2}$, não diferenciando a ordem das gavetas na caixa que contém as duas moedas de ouro e, portanto, considerando que depois de retirada uma moeda de ouro só restaria duas hipóteses equiprováveis. Houve mesmo alguns alunos que consideraram especificamente que a primeira moeda retirada correspondia à gaveta “de cima”, quando nada é referido no enunciado: “*Se saiu moeda de ouro, na gaveta de baixo só há duas hipóteses: ou saiu moeda de ouro ou sai moeda de prata.*”; “*A probabilidade de na segunda gaveta estar uma moeda...*”.

Na sua justificação, dois alunos recorreram ao diagrama de árvore mas não representaram corretamente a estrutura da experiência composta nem representaram as possibilidades que resultam da informação fornecida pelo acontecimento condicionante, de que a primeira moeda retirada é de ouro.

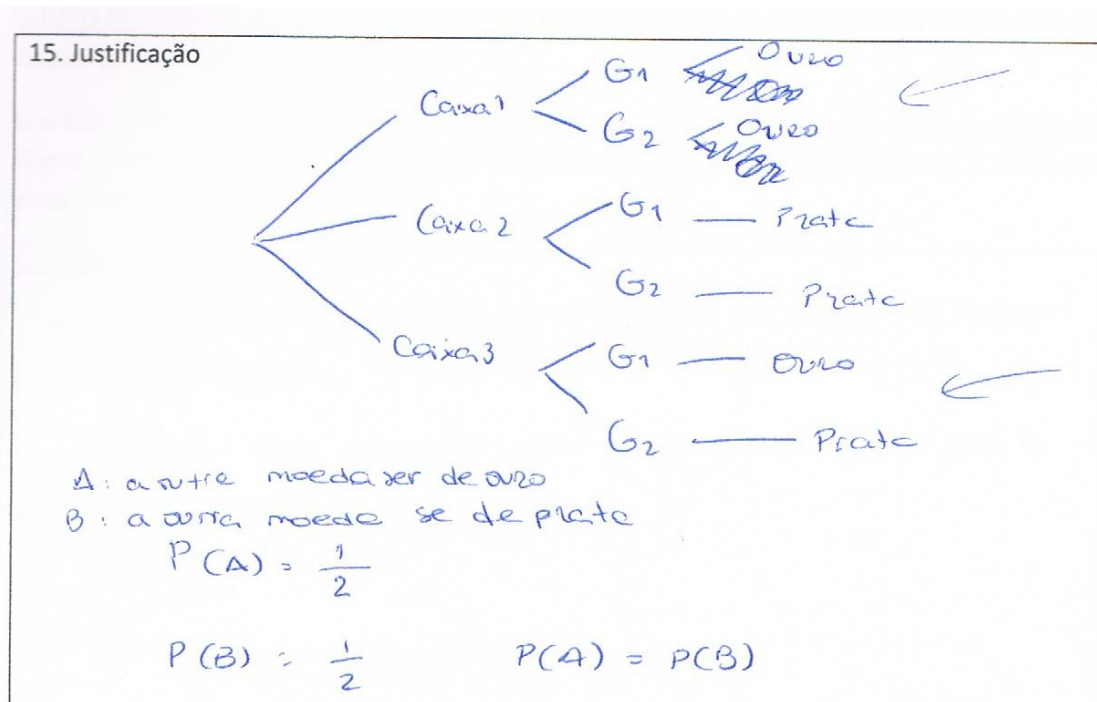


Figura 26 – Resolução do aluno com o código 69

Os restantes dois alunos que seleccionaram a opção a), referiram que a probabilidade de “calhar a caixa de ouros é $1/3$ assim como a probabilidade de sair a caixa com ouro e prata”.

Quando foi novamente colocado o mesmo problema, mas permitindo que os alunos recorrem-se aos simuladores criados com o programa R, todos os alunos alteraram a sua resposta para a opção correta de que que a probabilidade da moeda, que está na outra gaveta da mesma caixa, seja também de ouro é maior do que se obter uma moeda de prata. Os parâmetros utilizados pelos estudantes, nas simulações que realizaram individualmente e que foram apresentados como fundamentos às respostas dadas, são todos adequados no sentido de se verificar a estabilidade das frequências relativas e, consequentemente, permitirem obter estimativas “próximas” dos valores exatos das probabilidades em causa. O recurso ao simulador permitiu que os alunos constatassem que a sua intuição probabilística não estava correta conduzindo dois desses alunos à apresentação de uma justificação válida para a solução obtida a partir do computador. Tal facto leva-nos a afirmar que além do desenvolvimento de intuições probabilísticas corretas, o simulador pode desempenhar um papel importante na construção formal do conhecimento.

1. O Paradoxo da Caixa de Bertrand:

- sim2 (250 000)
- seq2 (20 , 5000)

A probabilidade de que a moeda que está na outra gaveta da mesma caixa também seja de ouro é maior do que se obter uma moeda de prata.

Se a moeda de ouro estiver na mesma gaveta da moeda de prata, a probabilidade de retirar uma moeda de prata é igual a $\frac{1}{3}$

A probabilidade de retirar uma moeda de ouro ^{da mesma caixa} é igual a $\frac{2}{3}$, porque ~~se~~ existem 2 moedas de ouro na mesma ~~caixa~~ ^{caixa} logo se retirar a moeda de ouro 1, fica na outra gaveta a moeda de ouro 2, ou se retirar a moeda 2, fica na outra gaveta a moeda 1, logo existem 2 casos possíveis na caixa que contém as duas moedas de ouro

Figura 27 – Resolução do aluno com o código 23.

O Paradoxo de Monty Hall: O jogo consiste em apresentar 3 portas ao concorrente, sabendo-se que uma das portas esconde um carro enquanto que as outras duas escondem uma cabra como prémio de consolação. Na 1.^a etapa o concorrente escolhe uma porta (que não é aberta). De seguida, o apresentador abre sempre uma porta que contém uma cabra, que é escolhida de entre as duas portas que o concorrente não escolheu. Na 2.^a etapa, dá-se a possibilidade ao concorrente de trocar a sua escolha inicial pela porta ainda fechada. Qual é a estratégia mais vantajosa para o jogador?

- Trocar de porta.
- Manter a porta escolhida inicialmente.
- É indiferente.

No que diz respeito ao dilema de Monty Hall, as repostas inicialmente dadas pelos alunos (sem recurso à simulação), foram um pouco mais dispersas: trocar de porta

(25%); manter a porta escolhida inicialmente (13%) e é indiferente (63%). Os dois alunos que optaram pela resposta correta de que se deve trocar de porta não apresentaram, no entanto, uma justificação válida para a sua resposta.

Tal como no problema da Caixa de Bertrand, quando foi permitido aos alunos recorrerem ao simulador, a conclusão obtida por todos é que efetivamente é mais vantajoso trocar de porta. No entanto, nenhum conseguiu apresentar uma justificação válida para a resposta apresentada.

Deste modo, estes dois problemas demonstram claramente que apesar de o estudante considerar inicialmente uma resposta, a simulação poderá convencê-lo que efetivamente a resposta correta é outra e, esta alteração, pode até permitir que o estudante chegue à justificação correta (após saber, via simulação, qual é a resposta verdadeira para o problema em estudo).

Apesar do cariz particular desta experiência não permitir fazer generalizações, os resultados obtidos sugerem que a resolução de problemas de probabilidades a partir da exploração de simuladores é uma estratégia pedagógica adequada no processo de aprendizagem desta temática. Podemos ainda realçar que, no inquérito sem simulação, mesmo os alunos que revelaram uma boa performance ao nível de escolha da opção correta, não apresentaram justificações adequadas e válidas para as suas respostas.

4.3.3. BREVE ANÁLISE DOS INQUÉRITOS DE SATISFAÇÃO

Procurando conhecer a opinião dos alunos em relação à experiência realizada com recurso ao software R e à sua motivação para a utilização das TIC como ferramenta de aprendizagem, foi solicitado que posteriormente respondessem a um inquérito de satisfação (cf. Anexo A.2.3.). As perguntas foram elaboradas de modo a conhecer a perceção dos alunos em relação a três aspetos fundamentais: o grau de facilidade na utilização do software R, a aprendizagem com a utilização do simulador e a motivação para a utilização das TIC em sala de aula. Da análise das respostas dadas pelos oito elementos da amostra (ver o resumo das respostas em anexo A.2.4.) podemos verificar que, à exceção de um aluno que não sentiu nem maior nem menor facilidade em interpretar os gráficos do simulador R, todos os outros consideraram que este *software* é de fácil utilização e interpretação. Todos os alunos concordam ou concordam totalmente que a utilização do simulador lhes deu maior confiança para responder aos problemas

propostos. No entanto essa percentagem diminuiu para 51% e 63% relativamente à compreensão dos conceitos subjacentes à simulação, como a definição frequencista e o comportamento dos fenómenos aleatórios respetivamente. Mesmo assim, é de realçar que as respostas foram maioritariamente favoráveis. No que concerne à motivação dos alunos para a aprendizagem quando se recorre a computadores, 63% dos inquiridos concordam ou concordam totalmente sentirem-se mais motivados quando utilizam computadores em sala de aula, 25% não concordam nem discordam e apenas um dos inquiridos (13%) diz discordar totalmente.

Relativamente à utilização de simuladores computacionais em sala de aula, em concordância com as respostas dadas sobre a motivação, embora com diferente distribuição, mantiveram-se as cinco opiniões favoráveis. De referir ainda que, relativamente a questão 1.7 do inquérito, tendo em conta que as respostas já observadas bem como uma nova leitura do enunciado da questão levantaram-nos dúvidas acerca da correta interpretação por parte dos estudantes, esta questão não foi tida em consideração na presente análise. Contudo, numa futura aplicação, esta questão será reformulada.

No campo deixado em aberto para possíveis comentários, as duas observações registadas foram positivas no que diz respeito à utilização do *software R*: “*Acho que deviam fazer uma maior divulgação do programa simulador em R.*” “*Foi bastante interessante usar o software R e acho que seria muito mais motivante para os alunos aprenderem por esse software.*”

CAPITULO 5 – CONCLUSÕES FINAIS, LIMITAÇÕES AO ESTUDO E TRABALHO FUTURO

As dificuldades partilhadas pelos alunos e docentes no que diz respeito ao ensino das probabilidades são, muitas vezes, originadas pela divergência existente entre a intuição e o verdadeiro significado concetual do que se está a estudar [6], sendo esta disparidade ainda mais acentuada em situações envolvendo o cálculo de probabilidades condicionadas, para as quais a informação conhecida (condicionante) implica uma atualização das probabilidades associadas a cada acontecimento, como ilustra o famoso problema de Monty Hall [5] ou o paradoxo da caixa de Bertrand [17].

Com este projeto pretendeu-se elaborar material didático para o ensino das probabilidades, ao nível do secundário, que abordasse as suas principais dificuldades, analisadas anteriormente no Capítulo 2 e, maioritariamente, provenientes de interpretações e intuições erradas. Integrando o contributo de vários estudos e investigações sobre o tema, os materiais produzidos foram concebidos para serem utilizados pelos alunos na investigação, exploração e melhor compreensão de diversos conceitos probabilísticos, nomeadamente, a noção frequencista de probabilidade, a lei dos grandes números, a independência de acontecimentos e a probabilidade condicionada. Tratando-se de simuladores computacionais, privilegiou-se uma metodologia de ensino assente nas novas tecnologias, em linha de consideração com as indicações expressas nos programas escolares relativas à temática das probabilidades. A construção do material teve por base três problemas da história da probabilidade e um problema de contexto real, todos eles passíveis de serem trabalhados por estudantes do ensino secundário, e selecionados pelas suas potencialidades ao nível dos conteúdos probabilísticos envolvidos e ao nível da captação do interesse pelos desafios colocados. Foi ainda objetivo deste trabalho a aplicação dos recursos elaborados, bem como a análise do impacto da sua utilização no que diz respeito a melhorias das intuições probabilísticas. Na aplicação pedagógica foi utilizado como instrumento de recolha de dados uma metodologia de natureza qualitativa com recurso à aplicação de dois questionários (o primeiro sem recurso ao *software* R e o segundo com recurso a simuladores), integrada num desenho do tipo pré-teste versus pós-teste, onde participaram oito alunos do 12.º ano. Com maior ou menor incidência, os alunos deste

estudo manifestaram os mesmos erros descritos na literatura, nomeadamente, as falácias de inversão do eixo temporal e da conjunção.

Apesar das conclusões obtidas a partir deste estudo, inserido num contexto particular, não se poderem generalizar, permitem ampliar e reforçar o conhecimento sobre o assunto em análise. Deste modo, os resultados evidenciam que o uso de simuladores facilitou a compreensão dos problemas apresentados, proporcionou aos alunos um papel mais ativo na aprendizagem e na descoberta da solução, e possibilitou a concretização de todas as tarefas propostas. A generalidade dos alunos concorda ou concorda totalmente sobre a facilidade na utilização do simulador R e a sua importância como ferramenta de apoio na resolução de problemas. As opiniões são um pouco mais divergentes, mas ainda assim maioritariamente positivas, no que diz respeito a melhor compreensão da definição frequencista e do comportamento dos fenómenos aleatórios aquando da utilização de simuladores. Em suma, as respostas dos alunos aos questionários após a realização de simulações, as respostas dadas no inquérito de satisfação, mas sobretudo as reações e a interação dos alunos durante a realização da experiência pedagógica perante a descoberta do raciocínio estocástico, revelam claramente (mesmo tendo em consideração todas as limitações da experiência realizada) que o recurso a simuladores computacionais contribui para o desenvolvimento da correta perceção do comportamento dos fenómenos aleatórios, razão pela qual consideramos fundamental a sua utilização no ensino dos conceitos basilares de probabilidade.

No entanto, tendo noção que a realização da experiência pedagógica apresenta algumas limitações, iremos tecer algumas considerações que podem sugerir futuras investigações. Em primeiro lugar, é de referir que a experiência pedagógica seria mais enriquecida caso fosse aplicada a um maior número de elementos e não existisse o constrangimento de esta ter de ser realizada num único dia. Apesar das circunstâncias não terem permitido realizar a triangulação dos dados recolhidos através de outros instrumentos, como por exemplo a realização de entrevistas aos elementos da amostra, esta seria uma mais-valia para aprofundar o conhecimento sobre as dificuldades sentidas pelos alunos e conhecer, através do discurso direto, de que forma sentiram que o uso do simulador lhes permitiu ou não vencer conceções probabilísticas erradas. Além disso, e uma vez que os processos escritos de justificação às respostas dadas foram escassos ou mesmo incorretos, as entrevistas iriam permitir que os alunos complementassem e/ou corrigissem verbalmente os processos de resolução adotados e as justificações dadas. O

tempo que tínhamos disponível forçou que a aplicação dos dois inquéritos tivesse sido feita num curto período de intervenção (um após o outro). Seria importante que decorresse, entre a aplicação dos dois inquéritos, um lapso de tempo de alguns dias de modo a permitir aos alunos a maturação/interiorização das ideias.

A investigação realizada teve como princípio de base a dimensão individual do contributo dos simuladores na construção do conhecimento probabilístico. Como trabalho futuro, poderia ser pertinente investigar o impacto da sua utilização numa dimensão colaborativa, no desenvolvimento de competências relacionados com o trabalho de grupo, e analisar até que ponto a interação entre pares (a exploração conjunta das tarefas, a formulação de conjeturas e discussão dos resultados), com recurso a simuladores, potencia o desenvolvimento correto de intuições probabilísticas.

BIBLIOGRAFIA

- [1] Batanero, C. (2004). Ideas estocásticas fundamentales: ¿Qué contenidos se debe enseñar en la clase de probabilidad? In J.A. Fernandes, M.V. Sousa, S.A. Ribeiro (Ed.), *Ensino e Aprendizagem de Probabilidades e Estatística, Actas I Encontro de Probabilidades e Estatística na Escola*. Braga: Centro de Investigação em Educação da Universidade do Minho, 9-30.
- [2] Batanero, C. (2006). Razonamiento probabilístico en la vida cotidiana: un desafío educativo. In P. Flores y J. Lupiáñez (Ed.), *Investigación en el aula de Matemáticas. Estadística y Azar*. Granada: Sociedad de Educación Matemática Thales.
- [3] Batanero, C., Contreras, J.M., Cañadas, Gustavo & Gea, M.M. (2012). Valor de las paradojas en la enseñanza de las matemáticas. Un ejemplo de probabilidad. *Novedades Educativas*, 261, 78-84.
- [4] Batanero, C., Contreras, J.M. & Díaz, C. (2012). Sesgos en el Razonamiento Sobre Probabilidad Condicional e Implicaciones Para la Enseñanza. *Revista digital Matemática, Educación e Internet*, 12(2).
- [5] Batanero, C., Fernandes, J.A. & Garcia, J.M.C. (2009). Un análisis semiótico del problema de Monty Hall e implicaciones didácticas. *Suma* 62, 11-18.
- [6] Batanero, C., Godino, J.D. & Cañizares, M.J. (2005). Simulation as a tool to train Pre-service School Teachers. *Proceedings of First ICMI African Regional Conference*. ICMI, Johannesburg.
- [7] Batanero, C., Godino, J.D. & Roa, R. (2004). Training Teachers To Teach Probability. *Journal of Statistics Education*, 12(1) (disponible em <http://www.amstat.org/publications/jse/v12n1/batanero.html>).
- [8] Batanero, C., Henry, M. & Parzysz, B. (2005). The nature of chance and probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning*. New York: Springer, 16-42.
- [9] Bogdan, R. & Biklen, S.K. (1994). *Investigação Qualitativa em Educação: Uma Introdução à Teoria e aos Métodos*. Porto: Porto Editora.

- [10] Borovcnik, M. & Peard, R. (1996). Probability. In A.J. Bishop *et al.* (Ed.), *International handbook of mathematics education*. Dordrecht: Kluwer Academic Publishers, 239-287.
- [11] Correia, P.F., Fernandes, J.A., & Contreras, J.M. (2011). Intuições de alunos do 9º ano de escolaridade sobre probabilidade condicionada. In C. Nunes, A. Henriques, A. Caseiro, A. Silvestre, H. Pinto, H. Jacinto, & J. Ponte (Ed.), *Actas do XXII Seminário de Investigação em Educação Matemática*. Lisboa: Associação de Professores de Matemática.
- [12] Coutinho, C. (2001). *Introduction aus situations aléatoires dés le Collège: de la modélisation à la simulation d'expériences de Bernoulli dans l'environnement informatique Cabri-géomètre-II*. Thèse de Doctorat. Université Joseph Fourier - Grenoble.
- [13] Díaz, C. (2009). Sessos en probabilidad condicional e implicaciones para la enseñanza. In J.A. Fernandes, M.V. Sousa & S.A. Ribeiro (Ed.), *Ensino e Aprendizagem de Probabilidades e Estatística*. Actas do II Encontro de Probabilidades e Estatística na Escola. Braga: Centro de Investigação em Educação, 100-116.
- [14] Falk, R. (1986). Conditional probabilities: insights and difficulties. In R. Davidson & J. Swift (Ed.), *Proceedings of Second International Conference on Teaching Statistics*. Victoria, Canada: International Statistical Institute, 291–297.
- [15] Fernandes, J.A. (1999). *Intuições e Aprendizagem de Probabilidades: Uma Proposta de Ensino de Probabilidades no 9º Ano de Escolaridade*. Doutoramento em Educação. Universidade do Minho, Braga.
- [16] Fischbein, E. & Gazit, A. (1984). Does the Teaching of Probability Improve Probabilistic Intuitions? An Exploratory Research Study. *Educational Studies in Mathematics*, 15(1), 1-24.
- [17] García, J., Batanero, C., Cezón, P. & Fuente, G. (2011). La paradoja de la caja de Bertrand: algunas formulaciones y cuestiones didáctica. *Epsilon - Revista de Educación Matemática*, 28(2).

- [18] Jonassen, D.H. (2000) *Computers as mindtools for schools: engaging critical thinking*, Prentice Hall.
<http://www.sciencedirect.com/science/article/pii/S0360131510003052?np=y>
- [19] Jones, G.A., Langrall, C.W., Thornton, C.A. y Mogill, A.T. (1999). Student's probabilistic thinking in instruction. *Journal for Research in Mathematics Education*, 30(5), 487-519.
- [20] Lesser, L. (1998). Countering indifference – Using counterintuitive examples. *Teaching Statistics*, 20(1), 10-12.
- [21] Martins, M.E.G. & Ponte, J.P. (2010). *Organização e Tratamento de Dados*. Ministério da Educação, Lisboa.
- [22] ME (2007). *Programa de Matemática do ensino básico*. Ministério da Educação. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinobasico/index.php?s=directorio&pid=71>.
- [23] ME (2001). *Programa de Matemática A do 10º ano*. Ministério da Educação, Departamento do Ensino Secundário. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinosecundario/index.php?s=directorio&pid=2>.
- [24] ME (2001). *Programa de Matemática Aplicada às Ciências Sociais*. Ministério da Educação, Departamento do Ensino Secundário. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinosecundario/index.php?s=directorio&pid=2>.
- [25] ME (2001). *Programa de Matemática B do 10º ou 11º ano*. Ministério da Educação, Departamento do Ensino Secundário. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinosecundario/index.php?s=directorio&pid=2>.
- [26] ME (2002). *Programa de Matemática A do 12º ano*. Ministério da Educação, Departamento do Ensino Secundário. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinosecundario/index.php?s=directorio&pid=2>.
- [27] ME (2002). *Programa de Matemática B do 12º ano*. Ministério da Educação, Departamento do Ensino Secundário. Lisboa. Acedido a 12-09-2013 em <http://www.dgicd.min-edu.pt/ensinosecundario/index.php?s=directorio&pid=2>.

- [28] MEC (2012). *Proposta de Metas Curriculares do Ensino Básico para a disciplina de Matemática*. Ministério da Educação e Ciência. Lisboa. Acedido a 12-09-2013 em <http://www.portugal.gov.pt/pt/os-ministerios/ministerio-da-educacao-e-ciencia/mantenha-se-atualizado/20120628-mec-metas-curriculares-ensbasico.aspx>.
- [29] MEC (2012). *Programa e Metas Curriculares de Matemática do Ensino Básico – 1.º, 2.º e 3.º Ciclos*. Ministério da Educação e Ciência. Lisboa. Acedido a 12-09-2013 em <http://dge.mec.pt/metascurriculares/index.php?s=directorio&pid=17>.
- [30] NCTM (2007). *Princípios e normas para a Matemática Escolar*. Lisboa: Associação de Professores de Matemática.
- [31] Pedro, S. (2012). *Modelação e Simulação - Uma Aplicação ao Problema da Ruína do Jogador*. Mestrado em Educação e Tecnologia em Matemática, Instituto Politécnico de Leiria.
- [32] Pestana, D. (2010). Devemos acreditar em análises clínicas? *Salutis Scientia*, 2.
- [33] Ponte, J.P. (2002). Investigar sobre a própria prática. In GTI (Org), *Reflectir e Investigar sobre a Prática Profissional*. APM: Lisboa, 5-28. Acedido a 22-09-2014 em [http://www.educ.fc.ul.pt/docentes/jponte/docs-pt/02-Ponte%20\(GTI\).pdf](http://www.educ.fc.ul.pt/docentes/jponte/docs-pt/02-Ponte%20(GTI).pdf).
- [34] Richardson, V. (1994). Conducting research on practice. *Educational Researcher*, 23(5), 5-10.
- [35] Santos, R. & Ribeiro, M.H. (2012). A ordem assintótica dos fenómenos aleatórios: uma ilustração via simulação. *Conferência Internacional de Investigação, Práticas e Contextos em Educação*. Escola Superior de Educação e Ciências Sociais de Leiria, 339-407.
- [36] Shaughnessy, J.M. (1992). Research in probability and statistics: reflections and directions. In Douglas A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning*. New York: Macmillan Publishing Company, 465-494.
- [37] Steinbring, H. (1989). The interaction between teaching practice and theoretical conceptions: a co-operative model of in-service training in stochastics for

mathematics teachers. In *Studies in mathematics education - Teaching statistics in schools*, 7. Paris: UNESCO, 202-214.

- [38] Tarr, J.E. (1997). *Using middle school students' thinking in conditional probability and independence to inform instruction*. Doctoral dissertation, Illinois State University.
- [39] Tarr, J.E., & Lannin, J.K. (2005). How can teachers build notions of conditional probability and independence? In G.A. Jones (Ed.), *Exploring probability in school: challenges for teaching and learning*. New York: Springer, 215-238.
- [40] Tenreiro, C. (2004). *Paradoxos no Cálculo das Probabilidades*. Acedido a 12-08-2014 em <http://www.mat.uc.pt/~tenreiro/divulgacao/Paradoxos.pps>
- [41] Tversky, A. & Kahneman, D. (1982). Causal schemas in judgment under uncertainty. In D. Kahneman, P. Slovic & A. Tversky (Ed.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press, 117-128.
- [42] Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293-315.
- [43] Watson, J. M. (1995). Conditional probability: Its place in the mathematics curriculum. *Mathematics Teacher*, 88, 12-17.

ANEXOS

ANEXO A1 – PROGRAMAS CONSTRUÍDOS PARA O *SOFTWARE R*

Nos programas apresentados, o utilizador deve escolher os valores dos parâmetros n e $series$ a utilizar nas simulações de acordo com o número de experiências ou o número de séries que pretende, respetivamente, realizar. Na simulação relativa às análises clínicas é ainda necessário definir os parâmetros prevalência da doença (p), sensibilidade da análise (s) e especificidade da análise (e).

A.1.1.- O JOGO PASSE-DIX

```
#####SIMULAÇÃO DO LANÇAMENTO DE 3 DADOS#####
```

```
sim1<-function(n)
{dados=3
  DM=rep(0,n)
  for (srun in 1:n) {DM[srun]=sum(ceiling(runif(dados,0,1)*6))}
  TDM=table(DM);
  barplot(TDM,      xlab="Soma      dos      pontos      dos      dados",
  ylab="Frequência",legend.text=c(paste("sair 9  =",round((TDM["9"])*100/n,2),"%"),
  paste("sair 10 =",round((TDM["10"])*100/n,2),"%")),args.legend = list(x = 20))}

#Para obter as somas dos pontos dos 3 dados, em 50 lançamentos, redigir sim1(50)#

#----- // -----
```

```
#####Evolução da proporção de 9 e 10 pontos#####
```

```
seq1<-function(series,n)
{ plot(0,0,type="b",xlab="", ylab="",pch=5,cex=0.25,main = "",xlim=c(10,n),
ylim=c(0.08,0.16))
  abline(h=25/216,col="red",lwd=4);abline(h=27/216,col="blue",lwd=4)

  for (srun in 1:series) {DN=rep(0,n);countN<-0;DD=rep(0,n);countD<-
0;dados=3;
  for (irun in 1:n) {
```

```

soma<-sum(ceiling(runif(dados,0,1)*6))
if(soma==9){countN<-countN+1} else if(soma==10){countD<-countD+1}
else {}
DN[irun]=round(countN/irun,4)
DD[irun]=round(countD/irun,4)}
lines(DN,col="red",lwd=3)
lines(DD,col="blue",lwd=3)}}
#Para obter as proporções de 9 e 10 pontos em 20 sequências de 10 000 lançamentos,
redigir seq1(20,10000).

```

#A vermelho representa-se as frequências dos 9 pontos e a azul as frequências dos 10 pontos.

A.1.2.- O PARADOXO DA CAIXA DE BERTRAND

```

#####Simulação da obtenção de moedas de ouro e prata#####

sim2<- function(n)
{DM=rep(0,n)
  caixas <- array(c("prata","prata","ouro","prata","ouro","ouro"),c(3,2))
  for (srun in 1:n) {
    caixa_escolhida=(ceiling(runif(1,0,1)*3))
    gaveta_escolhida=(ceiling(runif(1,0,1)*2))
    while(caixas[caixa_escolhida,gaveta_escolhida]=="prata")
    {caixa_escolhida=(ceiling(runif(1,0,1)*3));gaveta_escolhida=(ceiling(runif(1,0,1)*2))}
    gaveta_escolhida2=3-gaveta_escolhida
    DM[srun] = caixas[caixa_escolhida,gaveta_escolhida2]}
  TDM = table (DM)
  barplot (TDM, xlab="",ylab="Frequência",col=c("yellow","grey"),
  ylim=c(0,n*0.61), legend.text=c(paste("ouro=",round((TDM["ouro"])*100/n,1),"%"),
  paste("prata=",round((TDM["prata"])*100/n,1),"%")), args.legend = list(x =
  1.75,y=n*0.61-0.25))}
  #Para obter as quantidades de moedas de ouro e prata obtidas, na realização de
  50 experiências, redigir sim2(50)#

```

```

#----- // -----

#####Evolução da proporção de moedas de ouro e prata#####

seq2<-function(series,n)
  {caixas<-array(c("prata","prata","ouro","prata","ouro","ouro"),c(3,2))
  plot(0,0,xlab="", ylab="",pch=5,cex=0.25,main = "",xlim=c(10,n),
ylim=c(0,1))
  abline(h=1/3,col="grey",lwd=4);abline(h=2/3,col="yellow",lwd=4)
  for (srun in 1:series) {DM=rep(0,n);countM<-0;
    for (irun in 1:n) {
      caixa_escolhida=(ceiling(runif(1,0,1)*3))
      gaveta_escolhida=(ceiling(runif(1,0,1)*2))
      while(caixas[caixa_escolhida,gaveta_escolhida]=="prata")
{caixa_escolhida=(ceiling(runif(1,0,1)*3));gaveta_escolhida=(ceiling(runif(1,0,1)*2))}
      gaveta_escolhida2=3-gaveta_escolhida
      if(caixas[caixa_escolhida,gaveta_escolhida2]=="ouro"){countM<-
countM+1} else { }
      DM[irun]=round(countM/irun,4)
      lines(DM,col="yellow",lwd=3)
      lines(1-DM,col="grey",lwd=3)} }
#Para obter as proporções de moedas de ouro e prata obtidas em 20 sequências de 500
experiências realizadas, redigir seq2(20,500).

#A amarelo representa-se as frequências das moedas de ouro e a cinzento as frequências
das moedas de prata.

```

A.1.3.- O DILEMA DE MONTY HALL

```

#####Simulação de obtenção de prémio com a troca e não troca de porta#####

sim3<-function(n)
  {DM = rep (0,n)
  for (srun in 1:n) {

```

```

porta_premiada=(ceiling(runif(1,0,1)*3))
porta_apostada=(ceiling(runif(1,0,1)*3))
porta_retirada=(ceiling(runif(1,0,1)*3))
while(porta_retirada == porta_apostada | porta_retirada == porta_premiada)
{porta_retirada=(ceiling(runif(1,0,1)*3))}
porta_apostada=6/(porta_apostada*porta_retirada)
if(porta_apostada == porta_premiada){DM[srun] = "troca"} else {DM
[srun] = "não troca"}
TDM = table (DM)
barplot (TDM, xlab="",ylab="Frequência",col=c("blue","red"),
legend.text=c(paste("não troca=",round((n-TDM["troca"])*100/n,1,"%"),
paste("troca=",round((TDM["troca"])*100/n,1,"%")), args.legend = list(x =
1.1,y=n*0.61))}
#Para obter a percentagem de obtenção de prémio com a troca e a não troca de porta, na
realização de 50 experiências, redigir sim3(50)#

#----- // -----

#####Evolução da proporção da troca e não troca da porta#####

seq3<-function(series,n)
{ plot(0,0,type="b",xlab="", ylab="",pch=5,cex=0.25,main = "",xlim=c(10,n),
ylim=c(0,1))
abline(h=1/3,col="red",lwd=4);abline(h=2/3,col="blue",lwd=4)
for (srun in 1:series) {DM=rep(0,n);DT=rep(0,n);countM<-0;countT<-
0;A=0;B=0
for (irun in 1:n) {
porta_premiada=(ceiling(runif(1,0,1)*3))
porta_apostada=(ceiling(runif(1,0,1)*3))
porta_retirada=(ceiling(runif(1,0,1)*3))
while(porta_retirada == porta_apostada | porta_retirada == porta_premiada)
{porta_retirada=(ceiling(runif(1,0,1)*3))}
porta_apostada2=6/(porta_apostada*porta_retirada)
if(porta_apostada == porta_premiada){countM<-countM+1} else { }
if(porta_apostada2 == porta_premiada){countT<-countT+1} else { }

```

```

DM[irun]=round(countM/irun,4); A1=A; A=round(countM/irun,4)
segments(irun-1, A1, irun, A,col="red",lwd =3)
DT[irun]=round(countT/irun,4); B1=B; B=round(countT/irun,4)
segments(irun-1, B1, irun, B,col="blue",lwd =3)}}

```

#Para obter as proporções de prêmios obtidos com a troca e a não troca da porta em 50 sequências de 1 000 lançamentos, redigir seq3(50,1000).

##A vermelho representa-se as frequências de obtenção de prémio com a troca da porta e a azul as frequências da não troca.

A.1.4. - ANÁLISES CLÍNICAS

```

#####Simulação para obtenção dos valores VP, FP, FN e VN#####

sim4<-function(n,p,s,e)
{ DM=rep(0,n);VP=0;VN=0;FP=0;FN=0
  for (srun in 1:n) {u<-runif(1,0,1)*100;t<-runif(1,0,1)*100
    if (u<=p) {if (t<=s) {DM[srun]="1 - Positivo Verdadeiro";VP=VP+1}
      else {DM[srun]="2 - Falso Negativo";FN=FN+1}}
    else {if (t<=e) {DM[srun]="3 - Verdadeiro Negativo";VN=VN+1}
      else {DM[srun]="4 - Falso Positivo";FP=FP+1}}
  }
  TDM = table (DM)
  barplot(TDM, xlab="", ylab="Frequência", col=c("yellow","blue","gray","pink"),
  ylim=c(0,n))
  return(c("1 - Positivo Verdadeiro",VP,"2 - Falso Negativo= ",FN,"3 - Verdadeiro
  Negativo",VP,"4 - Falso Positivo",FP,"Total de utentes observados",n))}

#Para obter os valores VP, FP, FN e VN na aplicação de um teste a 50000 utentes cuja
doença tem uma prevalência de 5% e cuja análise tem uma sensibilidade de 95% e
especificidade de 75%, redigir sim4(50000,5,95,75).

```

A.2.1. - INQUÉRITO - PRÉ TESTE

Inquérito sobre Probabilidades - Parte 1

Caro aluno,

O presente questionário enquadra-se no âmbito de um trabalho de mestrado em Ciências da Educação. O seu principal objetivo é estudar as intuições probabilísticas dos alunos e identificar as principais dificuldades no cálculo de probabilidades.

O questionário é anónimo e confidencial pelo que se solicita a maior sinceridade nas respostas às questões apresentadas.

Obrigada pelo contributo!

*Obrigatório



1. Código de aluno *

Grupo I - Dados Gerais

Caraterização dos participantes

2. 1. Género *

Marcar apenas uma oval.

- Feminino
 Masculino

3. 2. Idade *

4. 3. Nota à disciplina de Matemática A no 12.º ano *

Marcar apenas uma oval.

- Sem avaliação
 1
 2
 3
 4
 5
 6
 7

- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20

Grupo II

No 9.º ano e no 12.º ano, no âmbito do tema de probabilidades, aprendeste a definição clássica de Probabilidade (ou Lei de Laplace) e a definição frequencista de Probabilidade.

5. **4. Por palavras tuas, diz como calcular a probabilidade de um acontecimento pela Lei de Laplace. ***

.....

6. **5. Por palavras tuas, diz o que entendes por definição frequencista de Probabilidade. ***

.....

7. **6. O que significa dois acontecimentos serem independentes? ***

.....

8. **7. Alguma vez, em aula, simulaste experiências aleatórias no cálculo de probabilidades? ***

Por exemplo, lançar dados ou moedas ao ar, extrair cartas de um baralho ou utilizar a calculadora/computador para o fazer.

Marcar apenas uma oval.

- Sim
- Não *Passe para a pergunta 11.*

9. **7.1. Em que ano escolar realizaste as simulações? ***

Podes seleccionar mais que uma opção.

Marcar tudo o que for aplicável.

- 9.º ano
- 12.º ano

10. **7.2. Que tipo de simulação realizaste? ***

Podes seleccionar mais que uma opção.

Marcar tudo o que for aplicável.

- Materiais manipuláveis (cartas, moedas, dados, bolas, etc)
- Calculadora
- Computador

Grupo III

Neste terceiro grupo, deves JUSTIFICAR a opção escolhida. Podes fazê-lo por palavras, através de esquemas ou de cálculos, em suporte papel (anexo I).

Um baralho de cartas completo é constituído por 52 cartas, repartidas em 4 naipes (Espadas, Copas, Ouros e Paus). Em cada naipe há 13 cartas: um Ás, três figuras (Rei, Dama e Valete) e mais 9 cartas (do Dois ao Dez).

11. **8. Considera que as 52 cartas foram divididas em 2 grupos distintos: um grupo A constituído apenas pelas 13 cartas do naipe de copas e um grupo B constituído pelas 39 cartas restantes. Retira-se uma carta de cada um dos grupos constituídos anteriormente. Em qual dos 2 grupos, A ou B, é mais provável obter uma carta Ás? ***

Marcar apenas uma oval.

- Grupo A.
- Grupo B.
- São igualmente prováveis.

12. **9. Considera a experiência aleatória que consiste em retirar 2 cartas do baralho, com reposição. Qual dos acontecimentos tem maior probabilidade de ocorrer? ***

Marcar apenas uma oval.

- Tirar 2 cartas de copas.
- Tirar uma carta de copas e uma carta de ouros.
- São igualmente prováveis.

13. **10. Considera a experiência aleatória que consiste em retirar uma carta do baralho. Qual dos acontecimentos é mais provável? ***

Marcar apenas uma oval.

- Sair Rei sabendo que a carta retirada é de Ouros.
- Sair Ouros sabendo que a carta retirada é um Rei.
- Ambos os acontecimentos são equiprováveis.

14. **11. Retiram-se, sucessivamente e sem reposição, 2 cartas do baralho. Qual dos acontecimentos é mais provável? ***

Marcar apenas uma oval.

- Sair figura na 2.^a extração, sabendo que a 1.^a carta extraída foi uma figura.
- Sair figura na 1.^a extração, sabendo que a 2.^a carta extraída foi uma figura.
- Ambos os acontecimentos têm a mesma probabilidade de ocorrer.

15. **12. Extrai-se uma carta ao acaso do baralho. Seja A o acontecimento "Sai uma carta de Ouros" e B o acontecimento "Sai um Ás". Os acontecimentos A e B são independentes? ***

Marcar apenas uma oval.

- Não são independentes porque no baralho existe um Ás de Ouros.
- Não são independentes porque o acontecimento A é mais provável que o acontecimento B.
- São independentes porque $P(A|B) = P(A)$.
- São independentes porque $P(A|B) = P(B)$.

16. **13. Considera que o João vai jogar um jogo que consiste em retirar, sucessivamente e sem reposição, 5 cartas do baralho. Se no final das 5 cartas retiradas saírem 3 ases, o João ganha o jogo. Qual dos seguintes acontecimentos é mais provável ocorrer? ***

Marcar apenas uma oval.

- A primeira carta retirada é um ás.
- A primeira carta retirada é um ás mas o João perde o jogo.
- Ambos os acontecimentos têm a mesma probabilidade de ocorrer.

17. **14. Extrai-se uma carta ao acaso do baralho. Seja A o acontecimento "Sai uma carta de Espadas" e B o acontecimento "Sai uma Dama". Qual dos acontecimentos é mais provável?**

Marcar apenas uma oval.

- $P(A|B)$.
- $P(B|A)$.
- São equiprováveis.

Grupo IV

Neste quarto grupo, deves JUSTIFICAR a opção escolhida. Podes fazê-lo por palavras, através de esquemas ou de cálculos, em suporte papel (anexo I).

18. **15. O Paradoxo da Caixa de Bertrand: Temos três caixas e cada caixa tem duas gavetas, cada uma com uma moeda: uma caixa contém duas moedas de ouro, a outra caixa duas moedas de prata e a terceira caixa uma moeda de cada tipo. Depois de se escolher uma caixa ao acaso, abre-se uma gaveta de forma aleatória verificando-se que esta contém uma moeda de ouro. Qual é a probabilidade de que a moeda que está na outra gaveta da mesma caixa também seja de ouro? ***

Marcar apenas uma oval.

- É a mesma de se obter uma moeda de prata.
- É maior do que se obter uma moeda de prata.
- É menor do que se obter uma moeda de prata.

19. **16. O Paradoxo de Monty Hall:** O jogo consiste em apresentar 3 portas ao concorrente, sabendo-se que uma das portas esconde um carro enquanto que as outras duas escondem uma cabra como prémio de consolação. Na 1.^a etapa o concorrente escolhe uma porta (que não é aberta). De seguida, o apresentador abre sempre uma porta que contém uma cabra, que é escolhida de entre as duas portas que o concorrente não escolheu. Na 2.^a etapa, dá-se a possibilidade ao concorrente de trocar a sua escolha inicial pela porta ainda fechada. Qual é a estratégia mais vantajosa para o jogador? *

Marcar apenas uma oval.

- Trocar de porta.
- Manter a porta escolhida inicialmente.
- É indiferente.

A.2.2.- INQUÉRITO - PÓS TESTE

Inquérito sobre Probabilidades - Parte 2

*Obrigatório



1. Código de aluno *

.....

Grupo I

Neste grupo, deves seleccionar a opção que consideras correta e indicar, no anexo II, os parâmetros utilizados nas simulações que foram decisivos para a tua resposta. Deves ainda JUSTIFICAR a opção escolhida por palavras, através de esquemas ou de cálculos, em suporte papel (anexo II).

2. **1. O Paradoxo da Caixa de Bertrand:** Temos três caixas e cada caixa tem duas gavetas, cada uma com uma moeda: uma caixa contém duas moedas de ouro, a outra caixa duas moedas de prata e a terceira caixa uma moeda de cada tipo. Depois de se escolher uma caixa ao acaso, abre-se uma gaveta de forma aleatória verificando-se que esta contém uma moeda de ouro. Qual é a probabilidade de que a moeda que está na outra gaveta da mesma caixa também seja de ouro? *

Marcar apenas uma oval.

- É a mesma de se obter uma moeda de prata.
- É maior do que se obter uma moeda de prata.
- É menor do que se obter uma moeda de prata.

3. **2. O Paradoxo de Monty Hall:** O jogo consiste em apresentar 3 portas ao concorrente, sabendo-se que uma das portas esconde um carro enquanto que as outras duas escondem uma cabra como prémio de consolação. Na 1.^a etapa o concorrente escolhe uma porta (que não é aberta). De seguida, o apresentador abre sempre uma porta que contém uma cabra, que é escolhida de entre as duas portas que o concorrente não escolheu. Na 2.^a etapa, dá-se a possibilidade ao concorrente de trocar a sua escolha inicial pela porta ainda fechada. Qual é a estratégia mais vantajosa para o jogador? *

Marcar apenas uma oval.

- Trocar de porta.
- Manter a porta escolhida inicialmente.
- É indiferente.

Grupo II

Neste grupo, deves JUSTIFICAR a opção escolhida. Podes fazê-lo por palavras, através de esquemas ou de cálculos, em suporte papel (anexo II).

Uma caixa contém 30 bolas: dez bolas vermelhas numeradas de 1 a 10 e vinte bolas pretas numeradas de 1 a 20.

4. **3. Extrai-se uma bola ao acaso da caixa. Seja A o acontecimento "Sai uma bola com um n.º 10" e B o acontecimento "Sai uma bola preta". Qual dos acontecimentos é mais provável? ***

Marcar apenas uma oval.

- $P(A|B)$
- $P(B|A)$
- São equiprováveis.

5. **4. Considera a experiência aleatória que consiste em retirar 2 bolas da caixa, com reposição. Qual dos acontecimentos tem maior probabilidade de ocorrer? ***

Marcar apenas uma oval.

- Tirar 2 bolas pares.
- Tirar uma bola par e uma bola ímpar.
- São igualmente prováveis.

6. **5. Considera que o João vai jogar um jogo que consiste em retirar, sucessivamente e sem reposição, 5 bolas da caixa. Se no final das 5 bolas retiradas saírem 3 bolas vermelhas, o João ganha o jogo. Qual dos seguintes acontecimentos é mais provável ocorrer? ***

Marcar apenas uma oval.

- A primeira bola retirada é vermelha.
- A primeira bola retirada é vermelha mas o João perde o jogo.
- Ambos os acontecimentos têm a mesma probabilidade de ocorrer.

7. **6. Retiram-se, sucessivamente e sem reposição, 2 bolas da caixa. Qual dos acontecimentos é mais provável? ***

Marcar apenas uma oval.

- Sair número par na 2.^a extração, sabendo que a 1.^a bola extraída foi par.
- Sair número par na 1.^a extração, sabendo que a 2.^a bola extraída foi par.
- Ambos os acontecimentos têm a mesma probabilidade de ocorrer.

8. **7. Considera a experiência aleatória que consiste em retirar uma bola da caixa. Qual dos acontecimentos é mais provável? ***

Marcar apenas uma oval.

- Sair uma bola com o n.º 2, sabendo que a bola retirada é vermelha.
- Sair uma bola vermelha, sabendo que a bola retirada tem o n.º 2.
- Ambos os acontecimentos são equiprováveis.

9. **8. Extrai-se uma bola ao acaso da caixa. Seja A o acontecimento "Sai uma bola par" e B o acontecimento "Sai uma bola vermelha". Os acontecimentos A e B são independentes? ***

Marcar apenas uma oval.

- Não são independentes porque na caixa existem bolas vermelhas pares.
- Não são independentes porque o acontecimento A é mais provável que o acontecimento B.
- São independentes porque $P(A|B) = P(B)$.
- São independentes porque $P(A|B) = P(A)$.

10. **9. Considera que as 30 bolas da caixa foram repartidas por cor, em 2 grupos distintos: um grupo A contendo as bolas vermelhas numeradas de 1 a 10 e um grupo B contendo as bolas pretas numeradas de 1 a 20. Retira-se uma bola, ao acaso, de cada um dos grupos A e B. Em qual dos 2 grupos é mais provável obter uma bola com número par? ***

Marcar apenas uma oval.

- Grupo A.
- Grupo B.
- São igualmente prováveis.

A.2.3.- INQUÉRITO DE SATISFAÇÃO

Inquérito sobre a experiência da utilização do simulador em R

O presente questionário tem como objetivo saber a opinião sobre a tua experiência com o uso do simulador R.

O questionário é anónimo e confidencial.

*Obrigatório

Marcar apenas uma oval por linha.

	1- Discordo Totalmente	2- Discordo	3- Nem concordo nem discordo	4- Concordo	5- Concordo Totalmente
1.1.- Foi fácil utilizar o simulador R.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.2.- Foi fácil interpretar os gráficos resultantes do simulador R.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.3.- Com o uso do simulador R senti maior confiança nas escolhas das opções para os problemas propostos (Monty Hall e Caixa de Bertrand).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.4.- A simulação permitiu-me compreender melhor o comportamento dos fenómenos aleatórios.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.5.- A simulação permitiu-me compreender melhor a definição frequencista de probabilidade.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.6.- Os simuladores computacionais poderiam ser usados, em sala de aula, no ensino das probabilidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.7.- Não existe qualquer utilidade em usar computadores quando é possível chegar aos resultados a partir de aplicação de fórmulas e cálculos numéricas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
1.8.- Sinto-me mais motivado para aprender a matéria quando utilizo computadores em sala de aula.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. **Utilize o seguinte espaço para escrever qualquer comentário que considere pertinente ou propor melhorias ao trabalho realizado.**

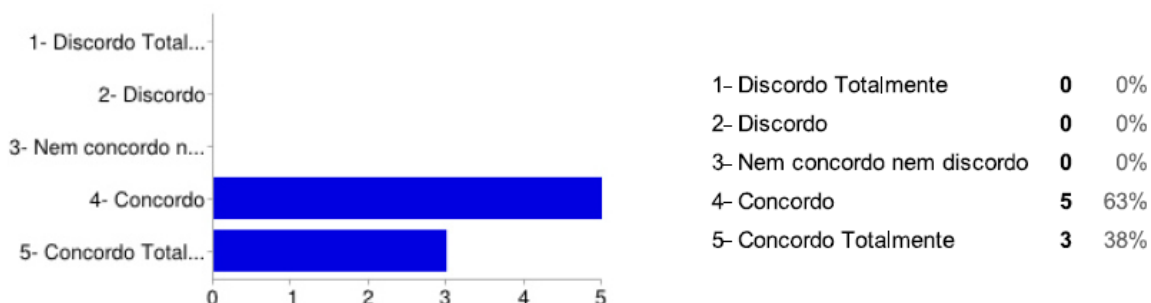
Muito obrigada pela vossa colaboração, que será fundamental para o sucesso deste estudo.



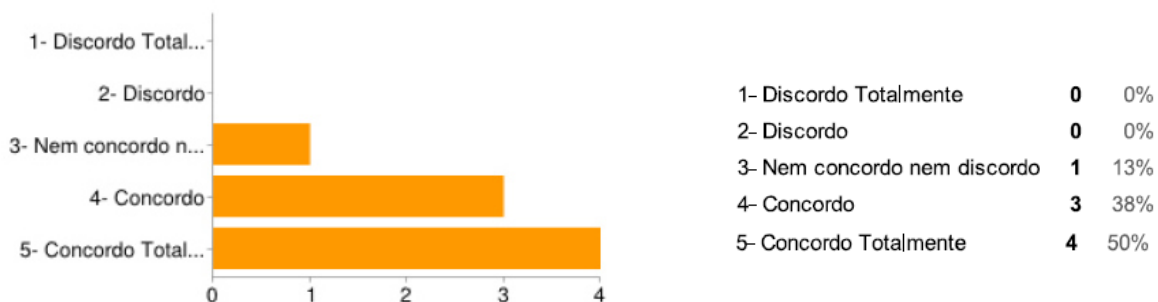
A.2.4.- RESULTADOS DO INQUÉRITO DE SATISFAÇÃO

Resumo

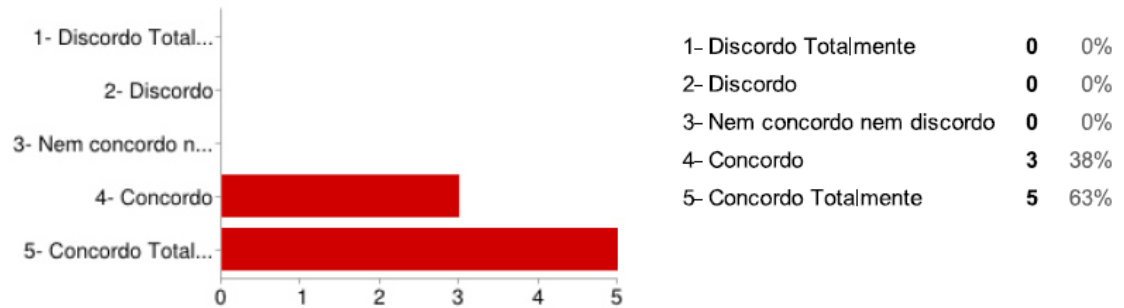
1.1.- Foi fácil utilizar o simulador R. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



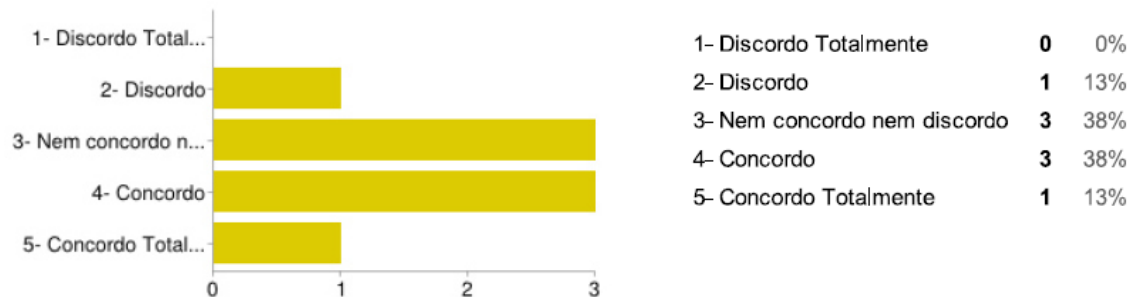
1.2.- Foi fácil interpretar os gráficos resultantes do simulador R. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



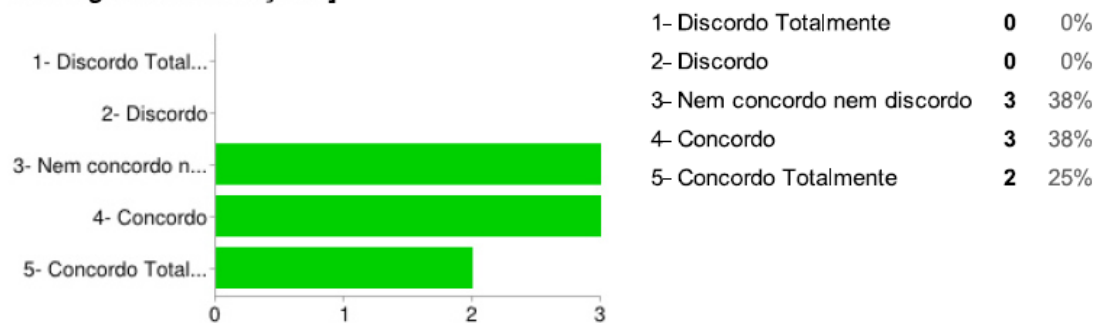
1.3- Com o uso do simulador R senti maior confiança nas escolhas das opções para os problemas propostos (Monty Hall e Caixa de Bertrand). [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



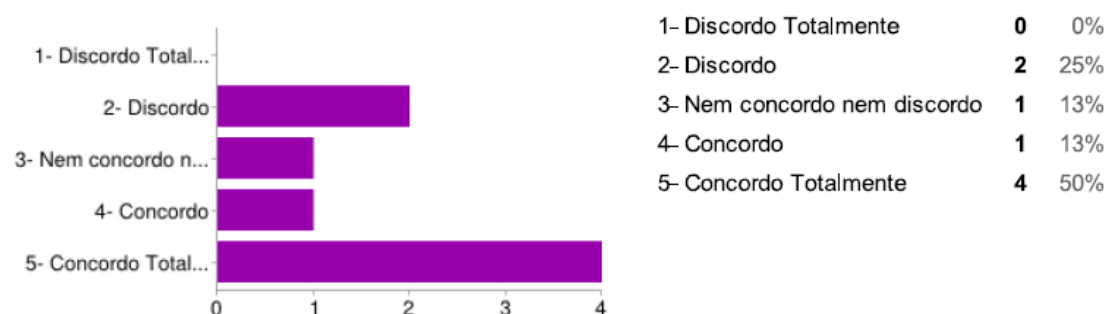
1.4.- A simulação permitiu-me compreender melhor o comportamento dos fenómenos aleatórios. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



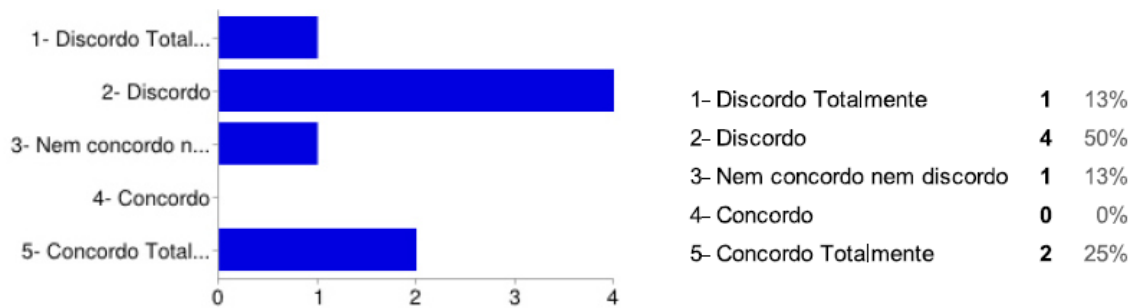
1.5.- A simulação permitiu-me compreender melhor a definição frequencista de probabilidade. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



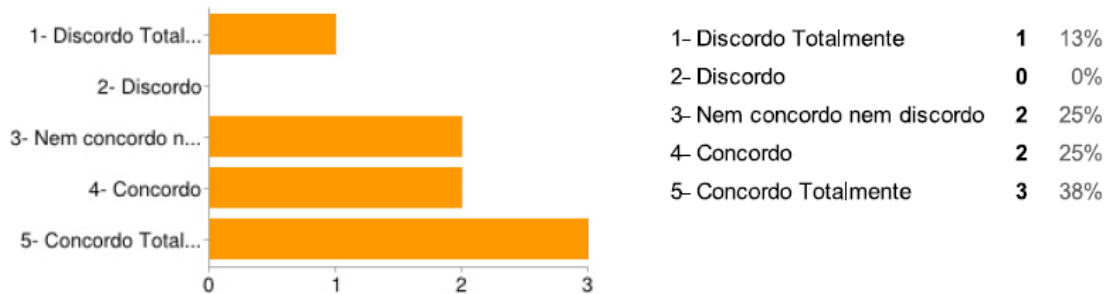
1.6.- Os simuladores computacionais poderiam ser usados, em sala de aula, no ensino das probabilidades. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



1.7.- Não existe qualquer utilidade em usar computadores quando é possível chegar aos resultados a partir de aplicação de fórmulas e cálculos numéricas. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



1.8.- Sinto-me mais motivado para aprender a matéria quando utilizo computadores em sala de aula. [1- Diz qual é o teu grau de concordância com cada uma das seguintes afirmações:]



2- Utilize o seguinte espaço para escrever qualquer comentário que considere pertinente ou propor melhorias ao trabalho realizado.

Acho que deviam fazer uma maior divulgação do programa simulador em R
 Foi bastante interessante usar o software R e acho que seria muito mais motivante para os alunos aprenderem por esse software.

FICHA DE APOIO

- 1- Durante o decorrer de toda a sessão irás ser representado por um número, atribuído de forma aleatória, e que servirá de “código de aluno”. Esse código de aluno, que deverás preservar durante todo o tempo, irá permitir o anonimato das tuas respostas e identificar os diferentes materiais que irás preencher como sendo do mesmo indivíduo.

- 2- A primeira parte da sessão consiste no preenchimento de um questionário sobre probabilidades. Para isso deves aceder ao link:
https://docs.google.com/forms/d/18bKAdBULf13RdTzqY-lkTGHVD_h73ov1ezBguwV4-ps/viewform?usp=send_form

- 2.1.- No campo 3 do grupo I, deverás indicar a nota final atribuída no 12.^o ano.
Caso tenhas anulado a disciplina ou apenas frequentado sem estares matriculado devido à falta de aprovação em anos anteriores, deverás seleccionar “Sem avaliação”.

- 2.2.- No grupo II, pretende-se que definas, por palavras tuas, alguns termos usados em probabilidades. Caso não te recordes ou nunca ouvistes falar, deverás preencher os campos com essa indicação.

- 2.3.- No grupo III e IV, para cada item, pretende-se que seleções a opção que consideras correta e justifiques a tua escolha, preenchendo o anexo I e indicando o código de aluno que te foi atribuído no início da sessão. A justificação pode ser feita a partir de esquemas, diagramas de árvores, tabelas, por palavras ou cálculos. Caso o espaço reservado para a justificação dos itens seja insuficiente, poderás completá-la no final da folha indicado o item a que se refere.

- 3- Na segunda parte da sessão, irá ser feita uma breve apresentação que, além da revisão alguns termos usados em probabilidades, contém os procedimentos a efetuar para a realização de simulações no programa R.

3.1.- Procedimentos para simular em R:

- Abrir o programa R;
- No menu *File*, seleccionar o comando *Open script...*;
- Aceder ao ambiente de trabalho, seleccionar e abrir o ficheiro *simulacao.R*;
- No menu *Edit*, seleccionar o comando *Run all*;
- Fechar a janela do editor de R.

3.2.- Simulação do jogo passe-dix:

- Para obter os vários resultados das somas dos pontos dos 3 dados, basta escolher o número de lançamentos que se pretende realizar.
Assim, para obter os resultados de 50 lançamentos basta redigir no programa R: `sim1(50)`;
- Para obter as proporções obtidas de 9 e 10 pontos em 20 sequências de 1000 lançamentos digitar: `seq1(20,1000)`.
A vermelho visualizamos as frequências referentes à obtenção de 9 pontos e a azul as frequências de 10 pontos.

- 4- Na terceira parte da sessão, irás novamente preencher um questionário acedendo ao link:

<https://docs.google.com/forms/d/1S-B1LrAMXjWL->

[MWYk0P4HwLbqcDubEvukdjDgUw0fk/viewform?usp=send_form](https://docs.google.com/forms/d/1S-B1LrAMXjWL-MWYk0P4HwLbqcDubEvukdjDgUw0fk/viewform?usp=send_form)

Muito importante, deverás indicar o código de aluno, igual ao preenchido no primeiro questionário.

4.1.- No grupo I, deverás seleccionar a opção que consideras correta, a partir da informação obtida das simulações por ti realizadas. No anexo II, deverás indicar os parâmetros utilizados nessas simulações e que foram decisivos para a tua resposta. Justifica a tua escolha a partir de esquemas, diagramas de árvores, tabelas, por palavras ou cálculos.

- Simulação do paradoxo da Caixa de Bertrand:

- Para obter as quantidades de moedas de ouro e prata obtidas, na realização de 50 experiências, redigir `sim2(50)`;

- Para obter as proporções de moedas de ouro e prata obtidas em 20 sequências de 500 experiências realizadas, redigir seq2(20,500).

A amarelo representa-se as frequências das moedas de ouro e a cinzento as frequências das moedas de prata.

- Simulação do paradoxo de Monty Hall:


- Para obter a percentagem de prémios obtidos com a troca e a não troca de porta, na realização de 50 experiências, redigir sim3(50);

- Para obter as proporções de prémios obtidos com a troca e a não troca da porta em 50 sequências de 1 000 lançamentos, redigir seq3(50,1000).

A azul representa-se as frequências para os prémios obtidos com a troca da porta e a vermelho as frequências da não troca.


4.2.- No grupo II, pretende-se que selecciones a opção que consideras correta e justifiques a tua escolha, preenchendo o anexo II. À semelhança do primeiro questionário, a justificação pode ser feita a partir de esquemas, diagramas de árvores, tabelas, por palavras ou cálculos. Caso o espaço reservado para a justificação dos itens seja insuficiente, poderás completá-la no final da folha indicado o item a que se refere.

A.2.6.- APRESENTAÇÃO EM POWERPOINT

 Mestrado em Ciências da Educação,
Especialização em Utilização Pedagógica das TIC


O Uso da Simulação no Cálculo de Probabilidades

Andrea Dias



Professor orientador: Doutor Rui Santos

julho de 2014

 O Uso da Simulação no Cálculo de Probabilidades

Introdução

A Teoria das Probabilidades centra-se no estudo de

Fenómenos Aleatórios

que se caracterizam:

- pela impossibilidade de prever os resultados das suas realizações individuais,
- mas pela existência de uma regularidade de comportamento quando visualizados um conjunto razoável de resultados.

Introdução

- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

2/24

Introdução

• Introdução

• Origem das Probabilidades

• Conceitos de Probabilidade

• Jogo "passe-dix"

• Simulação

• Caixa de Bertrand

• Monty Hall

• Demonstração

• O termo **Probabilidade** é utilizado todos os dias de forma mais ou menos intuitiva.

• Atualmente os estudos estocásticos desempenham um importante papel como ferramenta de investigação nas mais diversas áreas do conhecimento e na tomada de decisões ao nível económico, social e político em ambientes de incerteza.

3/24

Origem das Probabilidades

• Introdução

• Origem das Probabilidades

• Conceitos de Probabilidade

• Jogo "passe-dix"

• Simulação

• Caixa de Bertrand

• Monty Hall

• Demonstração

• A sua origem está relacionada, de certa forma, com os chamados "jogos de azar".

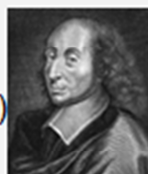
• A correspondência trocada entre



Pascal (1623-1662)

e

Fermat (1601-1665)



contendo reflexões sobre a resolução de problemas ligados aos jogos de dados, é considerada o despojar do Cálculo das Probabilidades.

4/24

Conceitos de Probabilidade

- Introdução
- Origem das Probabilidades
- **Conceitos de Probabilidade**
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

Definição Clássica ou Lei de Laplace

Quando os acontecimentos elementares são equiprováveis e em número finito:

$$\text{Probabilidade} = \frac{\text{n}^\circ \text{ de casos favoráveis}}{\text{n}^\circ \text{ de casos possíveis}}$$



A definição anterior é considerada a primeira definição de probabilidade, sendo atribuída a Laplace (1749-1827).



Conceitos de Probabilidade

- Introdução
- Origem das Probabilidades
- **Conceitos de Probabilidade**
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

Definição frequencista (ou empírica)

Repetindo a experiência muitas vezes:

$$\text{Probabilidade} = \text{valor para o qual tende a estabilizar a frequência relativa da realização do acontecimento}$$

Esta definição está associada à "Lei dos grandes números", formulada por Jacques Bernoulli (1645-1705).



Conceitos de Probabilidade**Probabilidade Condicionada**

- Introdução
- Origem das Probabilidades
- **Conceitos de Probabilidade**
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

Chama-se probabilidade de A condicionada a B ou probabilidade de A dado B:

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

7/24

Conceitos de Probabilidade**Acontecimento independentes**

- Introdução
- Origem das Probabilidades
- **Conceitos de Probabilidade**
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

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

Uma vez que a realização de um dos acontecimentos não interfere na probabilidade de realização do outro acontecimento, então:

$$P(A|B) = P(A) \text{ e } P(B|A) = P(B)$$

8/24

Jogo “passe-dix”

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo “passe-dix”
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

- De entre os problemas relacionados com jogos, que motivaram a reflexão e a discussão em torno das probabilidades, analisaremos o jogo “passe-dix”, frequentemente jogado nas cortes francesas.

Jogo “passe-dix”

- Lançamento simultâneo de 3 dados.
- O jogador ganha se a soma dos pontos obtidos for maior que 10. Caso contrário, a aposta é aumentada e passa a vez a outro jogador.

9/24

Jogo “passe-dix”

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo “passe-dix”
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

- O “paradoxo” subjacente ao jogo, que despertou a curiosidade de alguns matemáticos famosos como



Galileo Galilei (1564-1642)

e

Cardano (1501-1576)



- é o de confrontar a probabilidade de se obter 9 pontos com a de se obter 10 pontos.

10/24

Jogo “passe-dix”

• Introdução

• Origem das Probabilidades

• Conceitos de Probabilidade

• Jogo “passe-dix”

• Simulação

• Caixa de Bertrand

• Monty Hall

• Demonstração

Que muitos consideravam serem iguais pelo facto de ambas terem 6 formas de ocorrer.

9 pontos		
1	2	6
1	3	5
1	4	4
2	2	5
2	3	4
3	3	3

10 pontos		
1	3	6
1	4	5
2	2	6
2	3	5
2	4	4
3	3	4

11/24

Jogo “passe-dix”

• Introdução

• Origem das Probabilidades

• Conceitos de Probabilidade

• Jogo “passe-dix”

• Simulação

• Caixa de Bertrand

• Monty Hall

• Demonstração

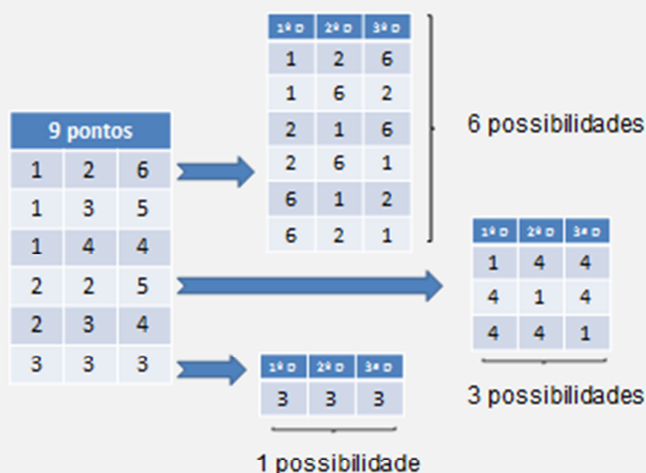
O que contradizia a experiência dos jogadores que, por observação repetida do lançamento de três dados, suspeitavam que é mais provável obter-se 10 pontos do que 9 pontos.



12/24

Jogo "passe-dix"

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passe-dix"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração



13/24

Jogo "passe-dix"

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passe-dix"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

Resumindo:

9 pontos	Possibilidades	10 pontos	Possibilidades
1 2 6	6	1 3 6	6
1 3 5	6	1 4 5	6
1 4 4	3	2 2 6	3
2 2 5	3	2 3 5	6
2 3 4	6	2 4 4	3
3 3 3	1	3 3 4	3
Total	25	Total	27

- Existem 27 maneiras igualmente prováveis de obter 10 pontos e apenas 25 maneiras de obter 9 pontos.

14/24

Jogo “passe-dix”

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo “passe-dix”
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

$$P(\text{“soma igual a 9”}) = \frac{25}{6 \times 6 \times 6}$$

$$P(\text{“soma igual a 10”}) = \frac{27}{6 \times 6 \times 6}$$

- Tal como a experiência dos jogadores fazia supor, é maior a probabilidade de se obter 10 pontos.

Simulação

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo “passe-dix”
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

- Um simulador computacional permite:
 - modelar fenómenos aleatórios,
 - fornecer solução aproximada a partir da visualização rápida e intuitiva dos resultados de um grande número de experiências.

Para simular os problemas apresentados, como o jogo “passe-dix”, vamos utilizar o software livre

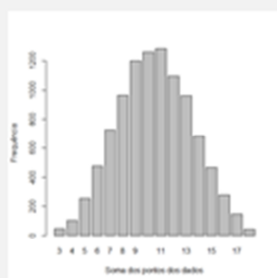
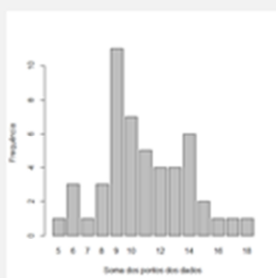


Simulação

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- **Simulação**
- Caixa de Bertrand
- Monty Hall
- Demonstração

➤ Para obter os vários resultados das somas dos pontos dos 3 dados, basta escolher o número de lançamentos que se pretende realizar.

Assim, para obter os resultados de 50 e 10 000 lançamentos basta redigir no programa R: `sim1(50)` e `sim1(10000)`, respetivamente.



17/24

Simulação

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- **Simulação**
- Caixa de Bertrand
- Monty Hall
- Demonstração

Ainda, graficamente, poderemos encontrar o valor aproximado da probabilidade de se obter 9 e 10 pontos, realizando seqüências de um número elevado de experiências.

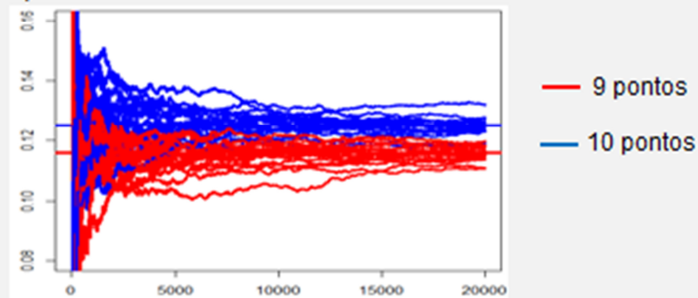


Figura: Evolução da proporção de 9 e 10 pontos em 20 seqüências de 20000 lançamentos

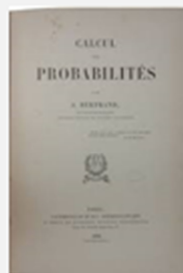
18/24

Caixa de Bertrand

O Paradoxo da Caixa de Bertrand:

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

- Problema formulado por Joseph Bertrand, matemático francês do Século XIX, no seu livro "Calcul des Probabilités" (1888).



Caixa de Bertrand

O Paradoxo da Caixa de Bertrand:

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-din"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

Temos três caixas e cada caixa tem duas gavetas, cada uma com uma moeda: uma caixa contém duas moedas de ouro, a outra caixa duas moedas de prata e a terceira caixa, uma moeda de cada tipo. Depois de se escolher uma caixa ao acaso, abre-se uma gaveta de forma aleatória verificando-se que esta contém uma moeda de ouro. Qual é a probabilidade de que a moeda que está na outra gaveta da mesma caixa também seja de ouro?

Monty Hall

O Problema de Monty Hall:

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

O problema foi inspirado de um concurso televisivo "Let's Make a Deal", emitido entre 1963 e 1986 na televisão americana e o seu nome provém do apresentador do concurso, Monty Hall.



Em Portugal, uma versão do concurso também foi apresentado com o nome "1,2,3", a partir de 1984.



21/24

Monty Hall

O Paradoxo de Monty Hall

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração

O jogo consiste em apresentar 3 portas ao concorrente, sabendo-se que uma das portas esconde um carro enquanto que as outras duas escondem uma cabra como prémio de consolação.

- Na 1ª etapa o concorrente escolhe uma porta (que não é aberta).
- De seguida, o apresentador abre uma das outras duas portas que o concorrente não escolheu, que contem uma cabra.
- Na 2ª etapa, dá-se a possibilidade ao concorrente de trocar a sua escolha inicial pela porta ainda fechada.

Qual é a estratégia que o jogador deve adotar?
Ficar com a porta escolhida inicialmente ou mudar de porta?
Será mais vantajoso trocar de porta, manter a escolha inicial ou será indiferente?

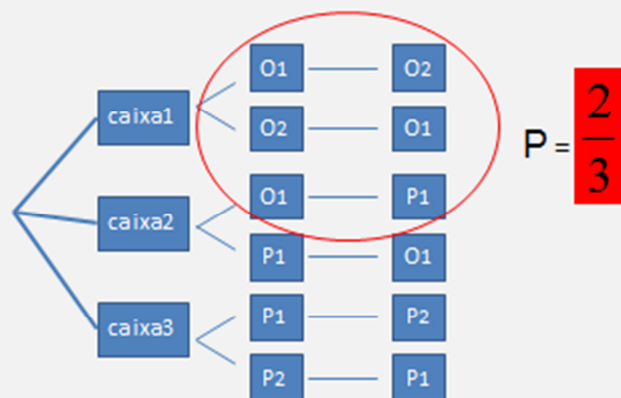


22/24

Demonstração Formal

O Paradoxo da Caixa de Bertrand:

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração



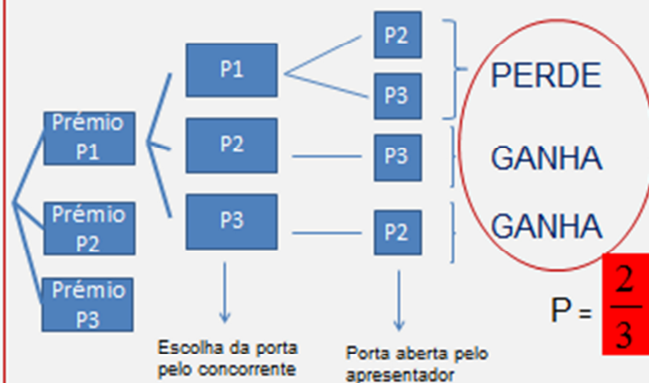
23/24

Demonstração Formal

O Paradoxo de Monty Hall:

Suponhamos que o prêmio está na porta 1 e que o concorrente altera a escolha inicial da porta:

- Introdução
- Origem das Probabilidades
- Conceitos de Probabilidade
- Jogo "passa-diz"
- Simulação
- Caixa de Bertrand
- Monty Hall
- Demonstração



24/24