

MARIA JOSÉ CARVALHO, ADELAIDE FREITAS

## NÍVEL DE CONHECIMENTO EM PROBABILIDADE CONDICIONADA E INDEPENDÊNCIA: UM CASO DE ESTUDO NO ENSINO SECUNDÁRIO PORTUGUÊS

KNOWLEDGE LEVEL IN CONDITIONAL PROBABILITY AND INDEPENDENCE OF EVENTS:  
A STUDY CASE IN THE PORTUGUESE HIGH SCHOOL

### RESUMEN

Partiendo de dos situaciones problemas que refieren a la probabilidad condicional, la independencia y eventos mutuamente excluyentes, se cuantifica, en paralelo, el grado de desempeño y el grado de rigor científico de una respuesta usando dos medidas ordinales en la escala de Likert. Un análisis descriptivo de las medidas, en 43 estudiantes del año terminal del liceo, determinó el grado de conocimiento de los alumnos en los referidos conceptos y apuntó que no siempre una respuesta correcta está descrita con rigor. Del análisis interpretativo de las respuestas se concluye que persisten, en la enseñanza en Portugal, conflictos en la interpretación y en el cálculo de la probabilidad condicionada y en los conceptos de independencia y de eventos mutuamente excluyentes. El estudio recomienda más prácticas involucrando la formulación matemática de problemas de probabilidad condicional y más énfasis en el carácter probabilístico que está asociado al concepto de independencia.

### PALABRAS CLAVE:

- *Concepto de probabilidad condicional*
- *Concepto de eventos independientes*
- *Concepto de eventos mutuamente excluyentes*
- *Grado de desempeño*
- *Grado de rigor científico*

### ABSTRACT

Taking two problem-situations involving conditional probability, independence and incompatibility, we suggest the measurements of the performance and the scientific rigor of a response using, in parallel, two ordinal Likert scale measures. Based on the responses given by 43 students in 12th grade (age 17), a descriptive analysis of these measurements was executed in order to evaluate the levels of knowledge of the students in those concepts. The results showed that not always a correct written response comes with rigor. Besides, an interpretative analysis of the same responses confirmed the existence of conflicts in the teaching of those concepts in Portugal. The conflicts are

### KEY WORDS:

- *Concept of conditional probability*
- *Concept of independent events*
- *Concept of mutually exclusive events*
- *Degree of performance*
- *Degree of scientific rigor*



concerning with the interpretation and the calculation of conditional probabilities and the notions of independence and incompatibility. The present study recommends more practice on the mathematical formulation of statements involving conditional probability and more emphasis on the probabilistic feature of the notion of independence.

## RESUMO

Partindo de duas situações-problema envolvendo probabilidades condicionadas, independência e incompatibilidade, sugere-se quantificar, em paralelo, o grau de desempenho e o grau de rigor de uma resposta usando duas medidas ordinais, ambas em escala de Likert. Uma análise descritiva dessas medidas, num grupo de 43 alunos do 12.º ano, permitiu estabelecer o nível de conhecimento desses alunos naqueles conceitos e constatar que nem sempre uma resposta correta é acompanhada de rigor na sua elaboração. Uma análise interpretativa das mesmas respostas permitiu ainda constatar que, no ensino português, persistem conflitos na interpretação e cálculo da probabilidade condicionada e conflitos nas noções de independência e incompatibilidade. O estudo recomenda mais prática na formulação matemática de enunciados envolvendo probabilidade condicionada e mais ênfase no carácter probabilístico da noção de independência.

## PALAVRAS CHAVE:

- *Conceito de probabilidade condicionada*
- *Conceito de acontecimentos independentes*
- *Conceito de acontecimentos incompatíveis*
- *Grau de desempenho*
- *Grau de rigor*

## RÉSUMÉ

Partant de deux situations-problème impliquant des probabilités conditionnelles, indépendance et incompatibilité, on suggère de quantifier, en parallèle, le degré de performance et le degré de rigueur d'une réponse utilisant deux mesures ordinales, toutes deux sur l'échelle de Likert. Une analyse descriptive de ces mesures dans un groupe de 43 élèves de Terminal a permis d'établir le niveau de connaissance des élèves, en relation à ces concepts, et de constater qu'une réponse correcte n'est pas toujours accompagnée de rigueur dans son élaboration. Une analyse interprétative de ces réponses a encore permis de voir que, dans l'enseignement portugais, les conflits persistent dans l'interprétation et le calcul des probabilités conditionnelles ainsi que dans les notions d'indépendance et d'incompatibilité. L'étude recommande plus de pratique dans la formulation mathématique des énoncés de probabilités conditionnelles et de mettre davantage l'accent sur le caractère probabiliste associé à la notion d'indépendance.

## MOTS CLÉS:

- *Concept de probabilité conditionnelle*
- *Concept d'évènements indépendants*
- *Concept d'évènements incompatibles*
- *Degré de performance*
- *Degré de rigueur*

## 1. INTRODUÇÃO

Estudos realizados na área da Educação Matemática revelam a existência de raciocínios enviesados, erros de compreensão e aplicações intuitivas erradas nos conceitos de probabilidade, probabilidade condicionada e independência de acontecimentos (D'Amelio, 2009; Díaz, Batanero & Contreras, 2010; Díaz y de la Fuente, 2006; Fernandes, Nascimento, Cunha e Contreras, 2011; Kataoka, Trevethan & Borim da Silva, 2010; Lonjedo-Vicent, Huerta-Palau & Carles-Fariña, 2012; Sánchez, 1996; Sobreiro, 2011; entre outros). Muitas das falhas podem ser explicadas em termos semióticos (Batanero, 2005). E, embora tais falhas possam estar enraizadas pois, como Fischbein e Schnarch (1997) alerta, equívocos envolvendo raciocínio probabilístico podem permanecer com a idade do aluno, é essencial a identificação dos tipos de conflitos existentes para as superar (ex., Díaz et al., 2010).

A realização de atividades matemática envolvendo o cálculo de probabilidades representa um instrumento importante no processo de aprendizagem para promover o desenvolvimento do pensamento probabilístico (Way, 2003; Kataoka et al., 2008), proporcionar diversos tipos de conexões entre conceitos da Teoria da Probabilidade e, conseqüentemente, agilizar a capacidade de raciocinar e de compreender, em particular, os conceitos de probabilidade condicionada e de acontecimentos independentes. No modelo ontosemiótico do conhecimento e do Ensino da Matemática, desenvolvido por Godino e seus colaboradores nas duas últimas décadas (Godino y Batanero, 1994; Godino, 2002; Godino, Batanero & Font, 2007; Godino, Font y Wilhelmi, 2008; Godino, 2012; entre outros) visando analisar, conjuntamente, o pensamento matemático, os manifestos que o acompanham e os fatores que condicionam o seu desenvolvimento, é destacada a necessidade de uma análise detalhada da atividade matemática considerando seis tipos de entidades primárias: a linguagem (termos, expressões, notações, gráficos,...) nos seus diversos registos (escrito, oral, gestual,...); a situação-problema (aplicações extra matemáticas, exercícios, ...); os conceitos-definições (abordados através de definições ou descrições); as proposições (enunciados sobre conceitos); os procedimentos (algoritmos, operações, técnicas de cálculo,...); e os argumentos (enunciados utilizados para validar ou explicar as proposições e procedimentos, de natureza dedutiva ou outra) (Neto, 2009). Na abordagem ontosemiótica, a compreensão de qualquer conceito matemático está associado à competência do saber fazer, no sentido que cada sujeito compreende determinado objeto matemático quando o usa de maneira capaz em diferentes práticas (Godino et al., 2007). A presente investigação baseou-se nas ideias teóricas do modelo ontosemiótico como marco de referência, sendo a probabilidade condicionada e a noção de acontecimentos independentes os objectos matemáticos de interesse no modelo e de análise na perspectiva de prática matemática.

Face a situações-problema, a aplicação de diferentes processos de resolução propicia a revelação de diferentes vertentes de um mesmo conceito (Díaz y de la Fuente, 2006) e, conseqüentemente, de novas estratégias para fazer juízos válidos (Cunha, 2010). Do ponto de vista matemático, probabilidade condicionada e independência são definições que dependem de expressões algébricas simples, de fácil cálculo. Torna-se então fundamental conhecer as estratégias que os alunos elaboram na resolução de situações-problema (D'Amore, 2006) e as dificuldades que revelam na aplicação de definições e no estabelecimento de ligações entre elas, face a um enunciado. Recentemente, Díaz et al. (2010) sintetizou vários dos conflitos semióticos referenciados na literatura especializada dando exemplos de enunciados de problemas que poderão ajudar, alunos e professores, à tomada de consciência de raciocínios enviesados associados àqueles dois conceitos. Resumidamente, relativamente à probabilidade condicionada, podem mencionar-se:

- convicção que é mais provável a conjunção de dois acontecimentos do que cada um dos acontecimentos em separado (Tversky & Kahneman, 1983), conhecido por *falácia da conjunção*;
- dificuldade em distinguir os dois sentidos da probabilidade condicionada,  $P(A|B)$  e  $P(B|A)$ , conhecida por *falácia da condicional transposta* (Falk, 1986);
- dificuldade em compreender que se tome um acontecimento condicionante que ocorre depois do acontecimento condicionado, conhecida por *falácia do eixo temporal* (Falk, 1979; Falk, 1986);
- confusão entre probabilidade condicionada,  $P(A|B)$ , e probabilidade conjunta,  $P(A \cap B)$ , (Watson & Moritz, 2002) devida, entre outros, à não clareza da linguagem corrente utilizada no enunciado dos problemas, levantando dificuldades na sua interpretação (Falk, 1986). Por exemplo, a utilização da vírgula “,” num texto verbal para descrever uma conjunção (“e”) (Sobreiro, 2011);
- dificuldade em identificar o espaço amostral reduzido no cálculo de probabilidades de acontecimentos definidos por experiências compostas dadas por uma sequência temporal de experiências simples sucessivas, conhecida por *situação sincrónica* (Falk, 1986);
- cálculo do denominador na fórmula de Bayes (Díaz & Batanero, 2009).

Relativamente a conflitos semióticos associados à noção de independência de acontecimentos, podem mencionar-se os seguintes:

- confusão intuitiva entre as palavras “independência” e “incompatibilidade” (Cordani & Wechsler, 2006);
- dificuldade em atender o significado somente probabilístico da noção de independência (D'Amelio, 2009);

- associação da noção de independência apenas a acontecimentos independentes no tempo (Karaota et al., 2010).

Este artigo apresenta uma análise interpretativa de um estudo de caso (Ponte, 1994) centrada na compreensão dos conceitos de probabilidade condicionada, acontecimentos independentes e acontecimentos incompatíveis por alunos do 12º ano de escolaridade do ensino português. Uma análise interpretativa de dados na metodologia de um estudo de caso visa uma investigação pormenorizada de todos os dados recolhidos com vista a organizá-los e classificá-los em categorias, de modo a explorar e a explicar o fenómeno em estudo (Tesch, 1990). Neste estudo, os participantes são os alunos e os dados são os documentos escritos produzidos pelos alunos, em contexto escolar, juntos dos seus docentes titulares de turma e em momento avaliativo.

Numa primeira instância, esta investigação, inserida no paradigma descritivo, procurou quantificar os processos desenvolvidos pelos alunos na resolução de questões, exploratórias e investigativas, e em situação de avaliação escrita. Definiram-se duas escalas ordinais para medir a qualidade de cada resposta. Tomando essas medidas, avaliou-se o nível de conhecimento que os alunos demonstraram nos conceitos de probabilidade condicionada e de independência de acontecimentos e nas suas interligações com outros conceitos. Analisou-se a conexão entre o nível de desempenho e o nível de rigor, sendo este último baseado na qualidade dos procedimentos e argumentos descritos e a linguagem usada. Estudou-se ainda a importância da capacidade de interpretar e de traduzir, em termos matemáticos, o enunciado de um problema no desenvolvimento da aprendizagem dos conceitos.

Numa segunda instância, a partir de uma análise interpretativa dos conteúdos das respostas expressas pelos alunos, procurou-se tipificar as fontes de conflitos semióticos detetados nas respostas dadas pelos alunos.

## 2. CONTEXTUALIZAÇÃO NO ENSINO PORTUGUÊS

Em Portugal, o tema *Probabilidades e Combinatória* consta do currículo atual do 12.º ano de escolaridade. Entre os tópicos inseridos nesse tema estão probabilidade condicionada e acontecimentos independentes. Nestes conceitos, os alunos “apresentam muitas dúvidas e dificuldades, apesar da grande importância que lhes é atribuída nos Exames Nacionais e nos Testes Intermédios, uma vez que são temas constantes em todas essas provas de avaliação externa, da responsabilidade do Ministério da Educação” (Cunha, 2010, p. 28).

Watson (1995) sugere que o estudo da probabilidade condicionada deve ser introduzido desde cedo, no Ensino Básico, e trabalhado de forma intuitiva, tal como é recomendado para o ensino da Matemática em Portugal, pelo Ministério da Educação, no documento *Programa de Matemática do Ensino Básico* (Ponte et al., 2001).

Nos manuais escolares de Matemática para o Ensino Secundário, o conceito de probabilidade condicionada é abordado, em geral, seguindo uma definição formal mais ou menos extensível em termos de complementaridade de informação. No caso da independência de acontecimentos, o conceito é menos trabalhado, sendo que a generalidade dos manuais não realça a importância do sentido probabilístico desse conceito.

### 3. ABORDAGEM METODOLÓGICA

A investigação em causa insere-se numa perspetiva qualitativa, com uma abordagem interpretativa (Tesch, 1990) das respostas dadas por alunos em ambiente natural, a sala de aula, com vista a obter uma descrição e explicação de padrões cognitivos subjacentes à cada resposta dada. Baseia-se num estudo de caso, com a seleção de uma amostra intencional, com vista a captar características, tão pormenorizadas quanto possível, da ocorrência de dificuldades por parte dos participantes (Ponte, 1994). Uma análise dos documentos produzidos pelos alunos contribui para detetar diferentes processos de resolução usados, reconhecer tipos de erros, observar estratégias e avaliar o nível de conhecimento na matéria.

Colaboraram neste estudo um total de 43 alunos pertencentes a duas turmas do Ensino Regular do Curso de Ciências e Tecnologias do 12.º ano (a frente designadas por turma A e turma C) e a uma turma do Ensino Profissional do 12.º ano (a frente designa por turma B) de duas escolas secundárias diferentes do distrito do Porto (a frente designadas por E1 e E2). Para a implementação do estudo realizaram-se reuniões individuais prévias com cada um dos docentes titulares das turmas A, B e C. Foi obtida uma caracterização das suas turmas, discutido o tipo de linguagem a ser utilizada no enunciado dos problemas a propor, o tempo e o período adequado à realização da prova escrita, e a forma de recolha das respostas elaboradas pelos alunos. Foi ainda abordado o desempenho expectável de cada turma, bem como a ênfase dada nas aulas aos tópicos probabilidade condicionada e acontecimentos independentes. Foi referido que a linguagem, procedimentos e argumentos considerados no processo de ensino daqueles conceitos está conforme o previsto no documento *Programa de Matemática do Ensino Básico*, Ponte e outros (2001).

Da turma A, constituída por 24 alunos, participaram nesta investigação 20 (os restantes não deram autorização para divulgar os seus resultados). Estes eram 9 do sexo feminino e 11 do sexo masculino, todos dentro da escolaridade obrigatória e com média de idade igual a 16,5 anos. Nenhum aluno desta turma usufruiu de apoio educativo no ano anterior. Todos os alunos revelaram gostar da escola. A docente de Matemática da turma A é efetiva na escola E1, com larga experiência na lecionação do 12.º ano, tendo acompanhado os seus alunos desde o Ensino Básico (7.º ano), pelo que conhece muito bem o trabalho individual destes. A turma B é constituída por 14 alunos, 6 do sexo feminino e 8 do sexo masculino. Todos os alunos estão dentro da escolaridade obrigatória à exceção de um aluno. A média de idades é de 16,8 anos. A turma C é constituída por 9 alunos, todos do sexo masculino. Apenas 5 alunos estão dentro da escolaridade obrigatória. A média de idades é de 17,6 anos. As docentes de Matemática das turmas B e C são efetivas na Escola E2 e com uma larga experiência de lecionação do 12.º ano. Estas docentes também têm acompanhado há muito tempo os seus alunos, não tendo sido referido na entrevista desde quando o fazem.

A turma A pode ser considerada como uma turma de “elite”, com alunos maioritariamente com interesse em prosseguir estudos a nível universitário, e distinta da turma C, com alguns alunos não interessados em seguir estudos a nível superior (politécnico ou universitário), e da turma B, com alunos que, não colocando de lado a prossecução de estudos superiores, optaram por um ensino secundário menos académico.

Foi elaborado o enunciado de dois problemas (designados a frente por P1 e P2), a implementar numa prova de avaliação, para os alunos das turmas A, B e C. Os problemas foram inspirados nos manuais escolares adotados nas escolas E1 e E2. Para a construção final do enunciado de P1 e P2 (em Apêndices 1 e 2) foram tidas em conta as ferramentas conceptuais do enfoque ontosemiótico, propostas por Godino et al. (2007) como sejam a linguagem, os conceitos, as propriedades, os procedimentos, as conjecturas e os argumentos ligados aos conceitos base de probabilidade condicionada e de independência, tendo como suporte uma análise sistemática de possíveis resoluções às situação-problemas propostas. Leitores interessados em detalhes dos objetos e relações primárias que intervêm na solução de cada uma das situação-problemas podem consultar Carvalho (2013).

Os professores de Matemática das turmas A, B e C analisaram os problemas e, após ligeiras adaptações, declararam que o enunciado, na sua versão final (ver Apêndices 1 e 2), estava adequado em termos de conteúdo programático, linguagem usada e nível de dificuldade. Porém, a docente da turma B solicitou a substituição da designação do acontecimento P por B no problema P2, já que os alunos estavam menos familiarizados com a letra P para designar um acontecimento. Esta alteração foi apenas considerada para a turma B.

Previamente foi ainda solicitado a um estudante do 12.º ano, que não pertencia a nenhuma das escolas envolvidas, que realizasse a prova em regime experimental. Este estudante afirmou tratar-se de uma prova acessível e com linguagem clara e objetiva.

Ambos os problemas, P1 e P2, abordam os conceitos de probabilidade condicionada e de acontecimentos independentes, usando a linguagem verbal e simbólica. Em P1 há 7 alíneas e em P2 há 5 alíneas, de acordo com a distribuição sumária descrita na Tabela 1. A noção de independência é abordada sequencialmente e relacionada com a noção de incompatibilidade, para diferentes situações conforme é indicado na Tabela 2. Em P1, os dados do enunciado encontram-se previamente estruturados numa linguagem mais próxima da linguagem simbólica, sendo estes fornecidos na forma de uma tabela de dupla entrada. Em P2, os dados são descritos em linguagem apenas verbal.

TABELA I  
Distribuição das 12 questões da prova (constituída pelos problemas P1 e P2)  
por conteúdo primário e linguagem usada no enunciado.

<i>Conteúdo primário</i>	<i>Linguagem verbal</i>	<i>Linguagem simbólica</i>
Lei da Probabilidade Total	1.1	----
Probabilidade da intersecção	1.4.a)	2.1. i)
Probabilidade da reunião	1.4.a)	----
Probabilidade condicionada	1.2, 1.3, 2.4	2.1.ii)
Independência	1.4.a), 1.4.b), 1.4.d), 2.3	----
Incompatibilidade	1.4.c), 1.4.d), 2.2	----

TABELA II  
Distribuição das questões de independência e de incompatibilidade  
tendo em conta a validade, ou não, da independência e da incompatibilidade.  
Na 1.4.d) é pedido para relacionar os dois conceitos.

<i>Conteúdo primário</i>	<i>Condição válida</i>	<i>Condição não válida</i>	<i>Relacionar</i>
Independência	1.4.a)	1.4.b), 2.3	1.4. d)
Incompatibilidade	1.4 c)	2.2	1.4. d)

Os dados para análise corresponderam às respostas escritas dos 43 alunos das turmas A, B e C, em momento avaliativo, às 12 questões da prova. A recolha



dos dados foi realizada em dezembro de 2012, data em que os alunos concluíram o estudo do tema *Probabilidades e Combinatória*. O acesso à prova elaborada por cada aluno, autorizado pelos professores de Matemática das turmas A, B e C, foi antes destas terem sido cotadas pelos respectivos professores. A identificação do aluno foi eliminada e atribuído um numeral a cada prova. As cópias das provas foram guardadas em suporte digital, para posterior análise. Por questões legais e éticas, foi previamente solicitada autorização para a realização da experiência, a recolha dos dados e a publicitação dos resultados às direções das escolas participantes e aos encarregados de educação dos alunos envolvidos no estudo.

Numa primeira parte, a investigação teve um carácter essencialmente qualitativo com a análise das respostas recolhidas, de modo a caracterizar, em termos descritivos, o seu conteúdo. Para uma descrição de cada resposta, procedeu-se de seguida, a uma análise de conteúdo e de interpretação de raciocínios subjacente à resposta, tendo por base comparativa a resposta correta expectável. Essa análise comparativa permitiu a categorização de cada resposta e, consequentemente, o agrupamento de respostas por classes de respostas similares em termos de qualidade.

Díaz e Batanero (2009) propuseram estabelecer uma correspondência entre a qualidade das respostas dos alunos e um numeral entre 0 a 2 do seguinte modo: não resposta ou resposta errada é representada por 0; resposta parcialmente correta é representada por 1; solução correta é representada por 2. No presente estudo propõe-se estender aquela categorização e dividir a apreciação da qualidade das respostas segundo duas perspetivas qualitativas ordinais distintas: o grau de desempenho da resposta (GD) e grau de rigor (GR). O desempenho diz respeito à capacidade da resposta conter a solução e referência dos pontos chaves para chegar ao resultado pretendido na pergunta. O rigor diz respeito à capacidade argumentativa e à linguagem apresentada na resposta escrita.

Para incorporar a ausência de resposta e dar maior detalhe à qualidade da resposta dada, caso exista, o parâmetro GD foi distribuído segundo a seguinte escala ordinal: ausência de resposta é representada pelo numeral 0; resposta totalmente errada é representada por 1; resposta parcialmente errada é representada por 2; resposta com início de uma estratégia que é abandonada ou concluída de forma errada é representada por 3; resposta parcialmente certa ou certa mas com passos omissos é representada por 4; por fim, resposta totalmente certa com indicação dos passos chaves correto é representada pelo numeral 5.

No que concerne ao grau de rigor (GR), foi escolhida uma escala de Likert, variando de 0 a 3 e parâmetros de uniformização de critérios descritos na Tabela 3. Para cada resposta, o valor atribuído para GR corresponde àquele que cumpre o maior conjunto de indicadores. A opção de um número par de correspondências

alternativas obriga a atribuir uma opção definitivamente positiva ou negativa, já que não é possível optar por uma atitude neutra por esta não existir.

A necessidade de proceder a extensões das três categorias propostas em Díaz e Batanero (2009) para as escalas definidas pelo GD e GR está intimamente relacionada com a natureza diversa das turmas A, B e C (Ensino Regular e Ensino Profissional) consideradas neste estudo. Efectivamente, do conhecimento empírico da primeira autora, é previsível a existência de tendência dos alunos do Ensino Profissional não explicitarem / argumentarem o seu raciocínio nas respostas. Com esta divisão proposta de GD e GR, prevê-se que traços diferenciadores, ao nível do desempenho e do rigor, poderão ser detetadas.

TABELA III

Critérios de classificação das classes possíveis do grau de rigor (GR) de uma resposta.

<i>GR</i>	<i>Crítérios de qualificação</i>
0 (ausência de rigor)	Resposta sem evidenciar a origem dos dados; Justificação sem correção científica e sem coerência; Ausência de qualquer referência aos acontecimentos tratados; Existência de muitos erros grosseiros de cálculo; Aplicação de conceitos erradamente.
1 (pouco rigor)	Resposta onde evidência a origem de alguns dos dados; Justificação com pouca correção científica e com alguma coerência; Acontecimentos tratados de forma pouco explícita; Existência de erros de leitura de dados que condicionam a resolução, mas sem erros grosseiros de cálculo; Conceitos aplicados com algum conflito entre estes, mas aplica as fórmulas de forma correta, sem as explicitar.
2 (algum rigor)	Resposta onde evidência a origem de todos os dados; Justificação com alguma correção científica e com alguma coerência; Acontecimentos tratados de forma explícita; Arredondamentos de cálculo que condicionam a resposta; Conceitos aplicados com pouca clareza, mas sem conflito entre eles. Embora as justificações escritas sejam pouco claras, aplica as fórmulas de forma correta, podendo ou não explicitá-las.
3 (com rigor)	Resposta onde evidência a origem de todos os dados; Justificação com correção científica e com muita coerência; Acontecimentos tratados de forma explícita e bem estruturados; Sem erros em cálculos intermédios nem de leitura de dados; Conceitos aplicados com clareza e correção na sua utilização, sem qualquer conflito entre eles, acompanhados de justificação correta, com utilização e aplicação de fórmulas com exatidão, explicitação e correção.

#### 4. ANÁLISE DAS RESPOSTAS DADAS

Em termos de procedimento, observa-se que cerca de 50% dos alunos traduzem, em termos explícitos e antes de responder às questões, os dados do enunciado de um problema em linguagem simbólica, quer com recurso a um diagrama de árvore, tabela de dupla entrada ou outro esquema. Esta tendência de não estruturar ou formular, em termos matemáticos, a linguagem textual de um enunciado, é particularmente notória na turma B. Esta característica é relevante já que pode reflectir a incapacidade de interpretar e obter corretamente os dados do enunciado.

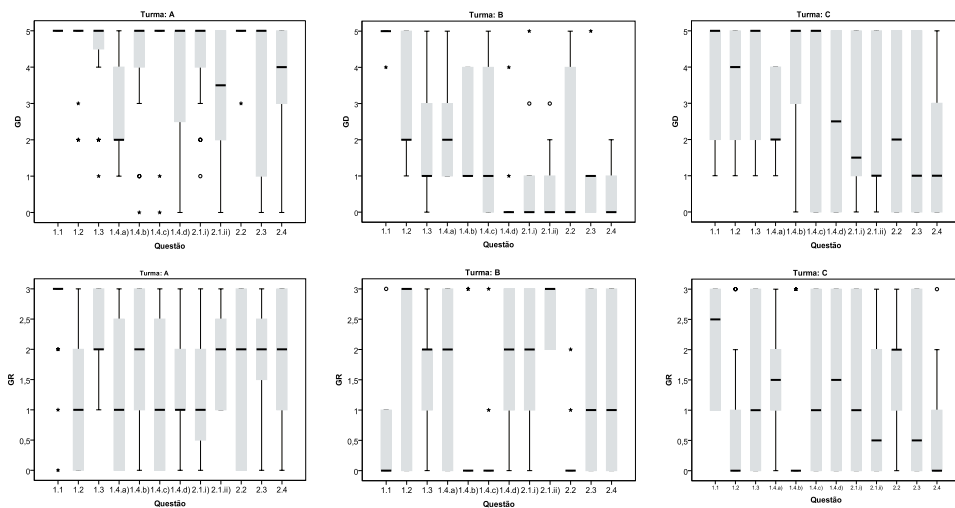
Em termos de justificações, os argumentos redigidos foram de natureza dedutiva baseados em síntese de cálculos realizados e propriedades dos conceitos. Várias vezes o rigor demonstrado nos cálculos nem sempre foi coerente com a explanação escrita e, outras vezes, foi traduzido não rigorosamente por palavras o que representavam os cálculos realizados. Em particular, os alunos do Ensino Profissional (turma B) apresentaram maior dificuldade em expressar um raciocínio argumentativo coeso e coerente com os resultados obtidos.

##### 4.1. *Nível de desempenho e Nível de rigor*

Analisadas as respostas e raciocínios apresentados pelos 43 alunos a cada uma das perguntas dos problemas P1 e P2, avaliou-se a qualidade de cada resposta em termos do grau de desempenho (GD) e do grau de rigor (GR). Considerando as 3 turmas em conjunto, 70% dos alunos envolvidos (todos os alunos da turma A, 1 na turma B e 9 na turma C) obtiveram desempenho positivo com soma dos GD, das 12 alíneas contidas nos dois problemas, igual ou superior a metade do valor máximo possível (60). No que concerne ao rigor das respostas corretas, constata-se que uma resposta certa não é geralmente acompanhada de uma linguagem e argumentação de nível elevado. A percentagem de alunos com soma dos GR, das 12 alíneas respondidas corretamente, igual ou superior a metade do valor máximo possível (36), baixa para 42% (14 da turma A, 0 na turma B e 4 na turma C).

Na Figura 1 encontram-se a distribuição dos valores correspondentes aos atributos (ordinais) das escalas GD e GR obtidos. Da Figura 1 conclui-se que as 3 turmas apresentam comportamentos diferentes em termos de desempenho e rigor nas respostas dadas pelos alunos. Tomando os valores medianos do GD, para cada uma das questões e por turma, observa-se que a turma A apresenta o melhor grau de desempenho, na maioria das questões, e a turma B os piores. Relativamente ao GR, em termos medianos, a turma A apresenta respostas com maior nível de

rigor e a turma B os mais baixos. Estes resultados vão de acordo com o sentir dos docentes, no terreno, ao afirmarem que os alunos do Ensino Profissional não apresentam os melhores resultados. Mas tal não significa a não existência de casos pontuais.



*Figura 1:* Caixas de bigodes dos valores obtidos para o GD (em cima) e para o GR (em baixo), para cada uma das 12 questões da prova, por turma. Todos os gráficos, à exceção do gráfico do GD para a turma C, apresentam observações atípicas (\* severas e ° moderadas).

Os valores atípicos assinalados na Figura 1, em maior quantidade no conjunto dos valores de GD, mostram a existência de alunos da turma B com níveis de desempenho superiores ao padrão da turma (definido, na Figura 1, pela caixa com os *bigodes*) em oposição à existência de alunos da turma A com níveis de desempenho abaixo do padrão distribucional da turma A. Analisando os valores mais centrais (amplitudes interquartis, Figura 1), observam-se, em geral, maiores amplitudes das caixas de bigodes para o GD na turma C comparativamente com a turma A (menos evidente no GR). Tal evidencia maior heterogeneidade de desempenho nas respostas dadas pelos alunos da turma C. Este facto leva a especular se esta constatação estará relacionada com a menor e maior diversidade de interesse académico futuro conhecido dos alunos das turmas A e C, respectivamente.

Centrando a atenção na linguagem, constata-se que nas resoluções apresentadas, os alunos privilegiaram o uso da linguagem verbal para representar os acontecimentos e a linguagem simbólica para identificar do tipo

de probabilidade pedida ou averiguar da independência e incompatibilidade de acontecimentos. Enfatizando ainda a diferença na linguagem dos enunciados dos dois problemas, linguagem simbólica (P1) versus linguagem verbal (P2), refira-se que cerca de metade dos alunos envolvidos traduzem em linguagem simbólica os dados do enunciado do problema P2 antes de resolvê-la.

Confrontado o número de questões não respondidas entre os problemas P1 e P2 (Figura 2), observa-se tendência para se registrar pior comportamento (i.e., não apresentar qualquer resposta) em questões do problema P2, em particular, nas turmas B e C, e menor tendência em deixar questões em branco por parte dos alunos da turma A.

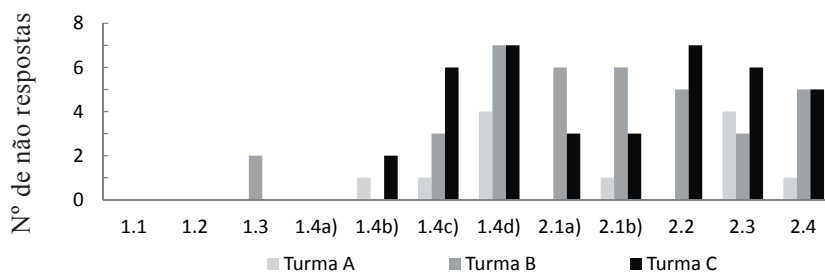


Figura 2. Frequência absoluta de não respostas a cada uma das 12 questões e para as turmas A, B e C.

Comparando as medianas dos valores observados para o GD e o GR entre as questões do problema P1 e P2 (Figura 1), observa-se valores medianos do GD máximos para as turmas do Ensino Regular (turmas A e C) e mais elevados para as respostas ao problema P1. Calculando a mediana dos valores de GD das respostas às questões do problema P1 e do P2 (Tabela 4), constata-se desempenho mediano melhor nas respostas dadas ao problema P1 do que ao problema P2, nas turmas B e C. Igual desempenho mediano nos problemas P1 e P2 foi observado na turma A. Estes factos sugerem que a turma A demonstra menor dificuldade em formular matematicamente o enunciado do problema P2, contrariamente às outras duas turmas. Avaliando globalmente o nível de desempenho dos alunos das três turmas em conjunto (Tabela 4, última linha), claramente no problema P2 o GD mediano é menor: pelo menos metade dos alunos apresentam em P2 um GD não superior a 2, em oposição ao problema P1 onde o nível máximo 5 é atingido em pelo menos 50% dos 43 alunos. Relativamente ao GR (Figura 1, em baixo, e Tabela 4), constata-se uma distribuição do nível de rigor mediano similar em ambos os problemas nas turmas A e B. Contudo, é a turma A que apresenta

maior grau de rigor, em termos medianos, o que significa uma melhor qualidade de argumentação e de linguagem na resposta correta dadas pelos alunos daquela turma nos dois problemas. Globalmente (Tabela 4, última linha), é em P2 onde se observa menor rigor mediano com pelo menos metade dos alunos a responder com um GR não superior a 1. Esta sumarização de comportamento denuncia que a utilização da linguagem verbal no enunciado dos dados de um problema pode influenciar negativamente a prestação global e, por inerência, a qualidade de respostas de alunos menos bem preparados. O facto da turma A demonstrar similar desempenho nos problemas P1 e P2 está em concordância com a indicação dada pela docente titular da turma A de estar ser uma turma com alunos, em geral, bastante empenhados.

TABELA IV

Valores medianos de GD / GR das respostas dadas nos problemas P1 e P2, por turma e no conjunto das 3 turmas. Por exemplo, 2 / 0 na tabela significa que, relativamente às respostas observadas às questões do problema P1 na turma B, pelo menos 50% delas apresentam GD inferior ou igual a 2 e pelo menos 50% (não necessariamente as mesmas) apresentam GR igual a 0.

<i>Turma</i>	<i>P1: Questões 1.1 - 1.4 d)</i>	<i>P2: Questões 2.1 - 2.4</i>
A	5 / 2	5 / 2
B	2 / 0	0 / 0
C	5 / 2	1 / 1
A, B e C, em conjunto	5 / 2	2 / 1

Para avaliar o nível de conhecimento nos conceitos de probabilidade condicionada e de independência (e incompatibilidade), calcularam-se os valores medianos do GD e do GR das respostas ao grupo das questões mais subjacentes ao conceito de probabilidade condicionada, questões 1.1, 1.2., 1.3, 2.1 e 2.4, e ao grupo das questões mais subjacentes ao conceito de independência e incompatibilidade, questões 1.4, 2.2 e 2.3 (Tabela 5). Relativamente ao GD, as 3 turmas apresentam, em termos medianos, a mesma qualidade de desempenho, nas respostas para as questões sobre probabilidade condicionada e para as questões sobre acontecimentos independentes e incompatíveis. Relativamente ao GR avaliado nas respostas corretas, nota-se, quer nas turmas A e C quer nas três turmas em conjunto e em termos medianos, um menor rigor nas respostas dadas às questões propostas sobre acontecimentos independentes e incompatíveis. Tal facto pode indiciar um menor à vontade dos alunos em explicar os resultados que envolvem daqueles dois conceitos.

TABELA V

Valores medianos de GD/GR das respostas dadas no grupo de questões envolvendo a noção de probabilidade condicionada e no grupo de questões envolvendo a noção de acontecimentos independentes por turma e no conjunto das 3 turmas.

Por exemplo, 3/1 na tabela significa que, relativamente às respostas dadas na turma C às questões sobre independência e incompatibilidade, pelo menos 50% delas apresentam GD inferior ou igual a 3 e pelo menos 50% (não necessariamente as mesmas) apresentam GR inferior ou igual a 1.

<i>Turma</i>	<i>Noção de probabilidade condicionada</i> <i>Questões: 1.1, 1.2, 1.3, 2.1, 2.4</i>	<i>Noção de indep. e incomp.</i> <i>Questões: 1.4, 2.2, 2.3</i>
A	5 / 3	5 / 2
B	1 / 0	1 / 0
C	3 / 3	3 / 1
A, B e C, em conjunto	4,5 / 1,5	5 / 1

Para averiguar o nível de associação entre as pontuações correspondentes aos atributos (ordinais) das escalas GD e GR obtidos em cada questão, considerou-se o coeficiente de correlação de Spearman (valores sumariados na Figura 3). A escolha deste coeficiente justifica-se por GD e GR serem medidas ordinais. O nível de correlação observado para cada questão é superior a 0,5, sendo o nível médio de correlação para o grupo de questões envolvendo probabilidade condicional e o grupo de questões envolvendo acontecimentos independentes igual a 0,77, para ambos os grupos. Tal significa que uma melhor qualidade de desempenho de uma resposta de ambos os tópicos, em média, vem acompanhada de um maior nível de rigor.

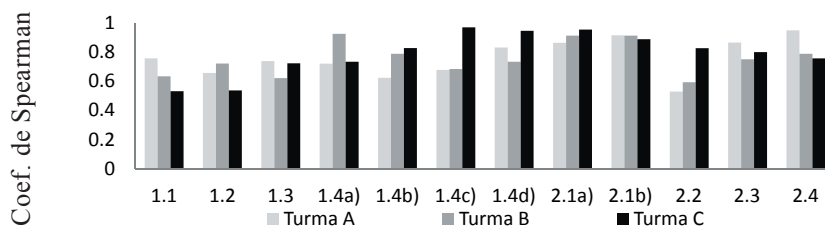


Figura 3. Valores do coeficiente de correlação de Spearman para cada questão e por turma.

## 4.2. Tipos de conflitos

Por observação de todas as resoluções com  $0 < GD < 5$  e  $0 < GR < 3$  entre os 43 alunos, conseguiu-se tipificar as fontes principais de erros que emergem das respostas incorretas em 3 grandes classes:

- C1 – conflitos na interpretação e cálculo da probabilidade condicionada;
- C2 – Propriedades incorrectas associadas à noção de independência;
- C3 – Propriedades incorrectas associadas à noção de incompatibilidade.

De seguida apresenta-se uma análise mais detalhada dos conflitos semióticos observados em cada classe.

### 4.2.1. Conflitos na interpretação e cálculo da probabilidade condicionada

Os conflitos inseridos na classe C1 correspondem a confusões de interpretação, num texto verbal, entre probabilidade condicionada e probabilidade conjunta e a dificuldades em delinear os passos essenciais corretos para o cálculo do denominador na fórmula da probabilidade condicionada (i. e., aplicar correctamente a lei da probabilidade total). Das respostas observadas evidencia-se ser a classe C1 a que origina mais respostas não corretas: 65% dos alunos apresentam pelo menos uma resposta com interpretação não correta dos dados do enunciado; 53% revela conflito no processo de cálculo de probabilidades condicionadas. O conflito atribuído à interpretação torna-se particularmente prejudicial pois origina a propagação de mais dificuldades como seja a incapacidade de estruturar, corretamente, em linguagem simbólica ou por tabela, os dados descritos em linguagem natural conduzindo assim a uma formulação matemática errada do enunciado.

Na Figura 4 são ilustrados alguns casos observados de conflitos de interpretação da probabilidade condicionada inseridos na classe C1. A Figura 4.a. contém um excerto da resposta do aluno 19 da turma A, à questão 1.3, que interpreta erradamente os textos “*Probabilidade de ser confirmado, no nascimento, que o bebé da Sra. Ana é do sexo feminino*” e “*Probabilidade de ser confirmado, no nascimento, que o bebé da Sra. Berta é do sexo masculino*” para uma probabilidade conjunta. As Figuras 4.b. e 4.c. mostram as respostas erradas do aluno 8 da turma B, à questão 1.2, e do aluno 1 da turma C, à questão 2.1, respetivamente, ao terem traduzido para probabilidade da intersecção um texto com enunciado tipicamente de probabilidade condicionada, no primeiro caso, da forma “*probabilidade de... [A]... se ...[B]*” e, no segundo caso, da forma “*probabilidade de...[A]... sabendo que ...[B]*”. Em ambos os textos está claramente explícito ser  $[B]$  o acontecimento



condicionante. Estes casos denunciam a inoperância do aluno de, face a um enunciado, averiguar se existe algum fenómeno ou acontecimento que restrinja o espaço amostral. Esta incapacidade sugere alguma imaturidade científica resultante da falta de informação ou de treino sobre enunciados tipicamente associados a probabilidades condicionadas, como os mencionados acima.

a.  $P(a \text{ ecografia fazer prever que é menino e, ao nascer, ser mesmo menino}) = \frac{127}{300} \approx 0,42$   
 $P(a \text{ ecografia fazer prever que é menino e, ao nascer, ser mesmo menino}) = \frac{68}{300} \approx 0,23$

« $P(a \text{ ecografia fazer prever que é menino e, ao nascer, ser mesmo menino}) = 127/300 \approx 0,42$ »  
 « $P(a \text{ ecografia fazer prever que é menino e, ao nascer, ser mesmo menino}) = 68/300 \approx 0,23$ »

b.  $1.2 - P(H \cap EF) = \frac{96}{300} = 0,32 = 3,2\%$   
 $P(M \cap EF) = 96/300 = 0,32 = 3,2\%$

c.  $2.1. P(A \cap B) = 99,8\%$   
 $P(A \cap B) = 99,8\%$

Figura 4. Respostas evidenciando a interpretação errada por probabilidade conjunta de textos relativos a probabilidades condicionadas.

Na Figura 5 estão transpostos alguns casos observados de conflitos de cálculo da probabilidade condicionada inseridos na classe C1. A Figura 5.a. ilustra uma resposta sem erros, do aluno 19 da turma A à questão 2.1, mas inacabada por falta do cálculo da probabilidade do acontecimento condicionante na fórmula da probabilidade condicionada. A Figura 5.b. mostra a estratégia da resposta do aluno 5 da turma C, na questão 2.1, de impor erradamente a independência de dois acontecimentos para obter um valor para o denominador na fórmula da probabilidade condicionada.

#### 4.2.2. *Propriedades incorrectas associadas à noção de independência*

Dentro da classe C2 incluem-se essencialmente dois tipos de conflitos, ambos derivados do estabelecimento de conexões erróneas, um com a noção de incompatibilidade e outro com a designação de acontecimentos incompatíveis. Enquanto o primeiro, mais grave, é fruto do aluno ainda não ter compreendido a noção probabilística associada ao conceito de independência confundindo-a com a noção de incompatibilidade, o segundo tipo realça a existência de similaridade nos termos incompatibilidade e independência, conduzindo à troca das designações em situação de stress, natural em momentos avaliativos. Para ambos os tipos registaram-se cerca de 30% dos alunos com pelo menos uma resposta nessas condições. Na Figura 6 apresentam-se dois exemplos do primeiro

tipo, com as respostas dadas pelo aluno 11, da turma C, e aluno 10, da turma A, à questão 1.4.b. Na Figura 7 apresentam-se dois exemplos do segundo tipo de erro com as respostas do aluno 17, da turma A, à questão 2.3 e do aluno 15, da turma A, à questão 1.4.b.

a. « 2. A: “Rato ser portador de bactérias”  
 $P(P|A) = 99,8\% = 0,998$   
 $P(P|\bar{A}) = 99,6\% = 0,996$   
 $P(A) = 6\% = 0,06$

2.1.  $P(A \cap P) = ?$   
 $P(P|A) = \frac{P(A \cap P)}{P(A)}$   
 $0,998 = \frac{P(A \cap P)}{0,06}$   
 $P(A \cap P) = 0,05988$

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

b. « 2.1.  $P(A \cap P) = 0,05988$   $P(A \cap \bar{P}) = 0,00012$   $P(A) = 0,06$   
 $P(B|A) = 0,998$   $P(B|\bar{A}) = 0,996$   $P(B) = 0,94$

$P(B|A) = \frac{P(B \cap A)}{P(A)}$   
 $0,998 = \frac{P(B \cap A)}{0,06} \Rightarrow P(B \cap A) = 0,05988$   
 $P(B \cap A) = P(A \cap B) = 0,05988$

$P(A) \times P(B) = P(A \cap B)$   
 $0,06 \times P(B) = 0,05988 \Rightarrow P(B) = 0,998$   
[erro de cálculos]

$P(A|B) = \frac{P(A \cap B)}{P(B)}$   
 $P(A|B) = 0,05988 / 0,0998 = 0,6$  »

Figura 5. Respostas evidenciando incapacidade em determinar corretamente o denominador na fórmula da probabilidade condicionada.

a. « 5) São os acontecimentos A e B não independentes, pois que o nascimento de um filho se interfere com o nascimento do outro. »

b. « 6) Não são independentes, porque no acontecimento A diz “O casal ter no máximo uma rapariga” e no B “O casal ter filhos de ambos os sexos”, logo nos 2 casos vai haver os 2 sexos, masculino e feminino, e os outros não são independentes. »

« Não são independentes, porque no acontecimento A diz “O casal ter no máximo uma rapariga” e no B “O casal ter filhos de ambos os sexos”, logo nos 2 casos vai haver os 2 sexos, masculino e feminino, então não são independentes. »

Figura 6. Respostas erradas evidenciando confusão entre a noção de independência e incompatibilidade.

a. 23-

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

$$0,05988 = 0,06 \times 0,06364$$

$$0,05988 = 0,0038184$$

Logo são incompatíveis.

« 2.3.  $P(A \cap P) = P(A) \times P(P)$

$$0,05988 = 0,06 \times 0,06364$$

$$0,05988 = 0,0038184$$

Logo são incompatíveis »

b. b) Se forem incompatíveis:

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

$$\left(\frac{136}{300}\right) \times \left(\frac{164}{300}\right) = 0,30 \times 0,25$$

$$0,13 = 0,075$$

Como esta igualdade não se verifica, estes dois acontecimentos não são independentes.

« b) Se forem incompatíveis:

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

$$\left(\frac{136}{300}\right) \times \left(\frac{164}{300}\right) = 0,30 \times 0,25$$

Como esta igualdade não se verifica, estes dois acontecimentos não são independentes »

Figura 7. Respostas erradas evidenciando trocas nas designações entre acontecimentos independentes e acontecimentos incompatíveis.

#### 4.2.3. Propriedades incorrectas associadas à noção de incompatibilidade

As resoluções das questões envolvendo o conceito de incompatibilidade (alíneas 1.4.c e 2.2) revelam que, em geral, os alunos compreendem o conceito de acontecimentos incompatíveis e justificam a incompatibilidade recorrendo, não à definição, mas somente à condição necessária da probabilidade da intersecção de acontecimentos incompatíveis ser nula. A verificação, apenas, desta condição necessária não é garantia dos acontecimentos serem incompatíveis. Já a sua não verificação é garantia dos acontecimentos não serem incompatíveis. Este jogo de raciocínio lógico envolvendo a condição necessária, mas não suficiente, de incompatibilidade leva à existência de conflitos na noção de incompatibilidade que definem a classe C3. Contabilizaram-se cerca de 20% dos alunos com pelo menos uma resposta com este tipo de conflito. Na Figura 8 estão dois exemplos de respostas observadas com argumentos errados inseridos na classe C3. A Figura 8.a. ilustra a resposta do aluno 1, da turma A, à questão 1.4.c., que recorre ao cálculo da probabilidade da intersecção para concluir a incompatibilidade. A Figura 8.b. revela a resposta do aluno 14, da turma C, à questão 2.2, onde confunde o numeral zero com o conjunto vazio (acontecimento impossível).

a. c)  $P((A \cap B) \cap C) = P((A \cap B) \cap (B \cap C))$   
 $= P((A \cap B) \cap C) = \emptyset$   
 e incompatível, uma vez que a intersecção de C com A e B não faz sentido. logo é incompatível »

b. 2.2.  $A \cap B = 0$   
 ~~$P(A \cap B) = P(A) \cdot P(B)$~~   
 ~~$0,003 = 0,003 \cdot 0,003$~~   
 Não porque a probabilidade da  $A \cap B = 0,003$  pelo que deveria ser 0 »

« 2.2.  $A \cap B = 0$   
 Não porque a probabilidade da  $A \cap B = 0,003$  pelo que deveria ser 0 »

Figura 8. Respostas dadas com conflitos na noção de incompatibilidade.

## 5. CONCLUSÕES

A realização de uma situação-problema motiva a abordagem de conceitos, propriedades e procedimentos, quer o aluno tenha já trabalhado nessa temática ou não, propiciando um amaduramento científico dos conceitos e propriedades. O domínio cognitivo dos conceitos e propriedades é aperfeiçoado nos procedimentos aplicados facilitando a fundamentação e a argumentação. Por seu turno, os argumentos a usar numa resposta determinam o procedimento a adoptar. Esta espiral de crescimento cognitivo expectável entre domínio, procedimento e argumentação, incentiva a realização de atividades matemática baseadas em situações-problema.

Na presente investigação são analisadas em paralelo duas escalas ordinais, GD e GR, a respostas dadas a duas situações problemas, P1 e P2, para avaliar o nível de conhecimento sobre tópicos concretos da Teoria da Probabilidade leccionado actualmente no 12.º ano de escolaridade em Portugal. Ambas as situações-problema propostas, P1 e P2, envolviam os tópicos de probabilidade condicionada, acontecimentos independentes e acontecimentos incompatíveis. Contudo, em P1 os dados encontravam-se esquematizados numa tabela de dupla entrada, enquanto em P2 os dados estavam descritos na forma de texto, o que exigia a sua formulação matemática por parte do aluno.

Para um grupo investigado de 43 alunos do 12.º ano, verifica-se que o nível de conhecimento em probabilidade condicionada e independência de

acontecimentos, medido em termos de GD e GR está intrinsecamente associado à *qualidade* (alta, média ou baixa) de estudo do estudante. As respostas com melhor qualidade foram de alunos da turma A (turma de “elite”) e as piores foram de alunos da turma B (do Ensino Profissional). Na realidade, é na turma com os melhores alunos que se encontram respostas com melhor desempenho e, simultaneamente, mais rigorosas. Estes resultados levam a crer que o nível de conhecimento nos conceitos de probabilidade condicionada e de acontecimentos independentes está dependente da motivação e facilidade de aprendizagem do aluno, entendendo-se que, como atualmente se crê, os alunos do Ensino Profissional estão menos motivados para continuação de estudos a nível superior e apresentam mais dificuldade para o estudo.

Em termos medianos, constata-se que o nível de conhecimento entre os dois conceitos não difere quando avaliado em função da capacidade do aluno de obter o resultado certo (GD) para questões envolvendo probabilidade condicionada e acontecimentos independentes. Porém, as respostas são mais rigorosas (i.e., GR maior) quando aborda a probabilidade condicionada. Assim, os resultados obtidos levam a conjecturar que existe um menor à vontade por parte do aluno no conceito de acontecimentos independentes e incompatibilidade.

Da análise dos graus de desempenho e de rigor por tipo de problema (com e sem esquematização dos dados do enunciado), é evidente que a facilidade de interpretação do problema condiciona a qualidade da resposta. Os alunos com maior nível de conhecimento (turma A) não demonstraram dificuldade de interpretação do problema sendo capazes de formular matematicamente um enunciado, contrariamente aos outros alunos.

Relacionando o grau de desempenho e o grau de rigor de uma resposta, constatou-se que nem sempre uma resposta correta é acompanhada pelo rigor na sua elaboração; porém, uma resposta com maior qualidade de desempenho, em média, vem acompanhada de maior rigor em termos de procedimento, argumento e/ou linguagem usada. Tal característica é comum para questões sobre probabilidade condicionada e sobre independência.

A análise do grupo das respostas dadas pelos 43 alunos sugere a existência de 3 classes, C1 (conflitos na interpretação e cálculo da probabilidade condicionada), C2 (conflitos na noção de independência) e C3 (conflitos na noção de incompatibilidade), como fonte primária de conflitos semióticos em situações-problema envolvendo os conceitos de probabilidade condicionada e independência de acontecimentos. A existência destas classes de conflitos revela a necessidade de se aprofundar tais conceitos em contexto escolar. Em resumo, as recomendações emergentes do estudo apontam para uma maior prática na

formulação matemática de enunciados envolvendo probabilidade condicionada e um maior ênfase no carácter probabilístico associado à noção de independência em oposição à noção de incompatibilidade. Concretamente, o presente caso de estudo mostra a necessidade de se exercitar na tradução de problemas com linguagem verbal para linguagem simbólica, em particular, de textos tipicamente sobre probabilidade condicionada como sejam, por exemplo, “*probabilidade de... [A]... se ...[B]*”, “*probabilidade de...[A]... sabendo que ...[B]*” e “*probabilidade de em [B] ocorrer [A]*”. O objectivo é instruir os estudantes a captarem, num enunciado, se existe fenómeno que restrinja o espaço amostral e assim identificar o acontecimento condicionante. Mais ainda, esta investigação salienta a necessidade de se reforçar que o cálculo ou referência às probabilidades são fundamentais para justificar a independência de acontecimentos, podendo ser usados para justificar a não incompatibilidade mas não para justificar a incompatibilidade.

## 6. CONSIDERAÇÕES FINAIS

Este trabalho, que resulta da investigação realizada para a dissertação de Carvalho (2013), confirma resultados de outros estudos e fornece novos elementos para a investigação do Ensino da Matemática. O carácter inovador prende-se com a proposta de usar, em paralelo, duas medidas ordinais, em escala de Likert, GD e GR, para quantificar a qualidade de uma resposta em termos do seu desempenho e rigor.

Nas temáticas probabilidade condicionada e independência, os tipos de conflitos semióticos listados nas classes C1, C2 e C3 encontram-se bem documentados, com maior ou menos ênfase, em literatura especializada. Contudo, o presente caso de estudo permitiu confirmar que em Portugal esses tipos de conflitos persistem, demonstrando a necessidade de se criarem acções de formação que visem consciencializar atempadamente o professor da existência daqueles conflitos.

## AGRADECIMENTOS

Trabalho parcialmente subsidiado por fundos portugueses através do CIDMA (Centro de Investigação e Desenvolvimento em Matemática e Aplicações) da Universidade de Aveiro e FCT (Fundação para a Ciência e a Tecnologia), dentro do projecto UID/MAT/04106/2013.



## REFERÊNCIAS BIBLIOGRÁFICAS

- Batanero, C. (2005). Significados de la probabilidad en la educación secundaria. *Revista Latinoamericana de Investigación en Matemática Educativa*, 8(3), 247-263.
- Carvalho, M. J. (2013). *Ensino e aprendizagem de probabilidade condicionada e independência* (Dissertação de Mestrado não publicada). Universidade de Aveiro, Aveiro, Portugal. Recuperada de <http://hdl.handle.net/10773/12044>.
- Cordani, L. K. & Wechsler, S. (2006). Teaching independence and exchangeability. In A. Rossman & B. Chance (Eds.), *Proceedings of the 7<sup>th</sup> International Conference on Teaching Statistics*. Salvador, Brasil: International Association for Statistics Education.
- Cunha, M. C. (2010). *A influência do ensino nos raciocínios de alunos do 12.º ano de escolaridade em probabilidade* (Dissertação de Mestrado não publicada). Universidade do Minho, Braga, Portugal. Recuperada de: <http://hdl.handle.net/1822/10945>.
- D'Amelio, A. (2009). Undergraduate student difficulties with independent and mutually exclusive events concepts. *The Montana Mathematics Enthusiast*, 6(1-2), 47-56.
- D'Amore, B. (2006). Objetos, significados, representaciones semióticas y sentido. *Revista Latinoamericana de Investigación en Matemática Educativa*, 9(Número Especial), 177-196.
- Díaz, C. & Batanero, C. (2009). University students' knowledge and biases in conditional probability reasoning. *International Electronic Journal of Mathematics Education*, 4(3), 131-162.
- Díaz, C., Batanero, C., & Contreras, J. M. (2010). Teaching independence and conditional probability. *Boletín de Estadística e Investigación Operativa.*, 26(2), 149-162.
- Díaz, C. y de la Fuente, I. (2006). Enseñanza del teorema de Bayes con apoyo tecnológico. En P. Flores y J. Lupiáñez (Eds.), *Investigación en el aula de matemáticas. Estadística y Azar*. Granada, España: Sociedad de Educación Matemática Thales.
- Falk, R. (1979). Revision of probability and the time axis. *Proceedings of the 3<sup>rd</sup> International Conference for the Psychology of Mathematics Education* (pp. 64-66). Warwick, UK: Organizing Committee.
- Falk, R. (1986). Conditional probabilities: Insights and difficulties. In R. Davidson & J. Swift (Eds.), *Proceedings of the 2<sup>nd</sup> International Conference on Teaching Statistics* (pp. 292-297). British Columbia, Canada: University of Victoria.
- Fernandes, J. A., Nascimento, M. M., Cunha, M. C. e Contreras, J. M. (Junho, 2011). *Desenvolvimento do conceito de probabilidade condicionada em alunos do 12.º ano através do ensino*. Comunicação apresentada na 13ª Conferência Interamericana de Educação Matemática. Recife, Brasil: CIAEM. Recuperada de: <http://hdl.handle.net/1822/12924>.
- Fischbein, E. & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28(1), 96-105. doi: 10.2307/749665
- Godino, J. D. (2002). Un enfoque ontológico y semiótico de la cognición matemática. *Recherches en Didactique des Mathématiques*, 22 (2-3), 237-284.
- Godino, J. D. (2012). Origen y aportaciones de la perspectiva ontosemiótica de investigación en Didáctica de la Matemática. En A. Estepa, A. Contreras, J. Deulofeu, M. C. Penalva, F. J. García y L. Ordóñez (Eds.), *Investigación en Educación Matemática XVI* (pp. 49-68). Jaén, Espanha: SEIEM.
- Godino, J. D. y Batanero, C. (1994). Significado institucional y personal de los objetos matemáticos. *Recherches en Didactique des Mathématiques*, 14(3), 325-355.
- Godino, J. D., Batanero, C., & Font, V. (2007). The onto-semiotic approach to research in mathematics education. *ZDM. The International Journal on Mathematics Education*, 39 (1-2), 127-135. doi: 10.1007/s11858-006-0004-1

- Godino, J. D., Font, V. y Wilhelmi, M. R. (2008). Análisis didáctico de procesos de estudio matemático basado en el enfoque ontosemiótico. *Publicaciones*, 38, 25-48.
- Kataoka, V., Trevethan, H. e Borim da Silva, C. (2010). Independence of events: an analysis of knowledge level in different groups of students. In C. Reading (Ed.), *Proceedings of the 8<sup>th</sup> International Conference on Teaching Statistics*. Ljubljana, Eslovénia.
- Kataoka, V. Y., Souza, A. A., Oliveira, A. C. S., Fernandes, F. M. O., Paranaíba, P. F., & Oliveira, M. S. (July, 2008). *Probability Teaching in Brazilian Basic Education: Evaluation and Intervention*. Paper presented at the 11<sup>th</sup> International Congress on Mathematical Education, Monterrey, México: ICME. Recuperada de: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.214.9221>
- Lonjedo-Vicent, M. A., Huerta-Palau, M. P. e Carles-Fariña, M. (2012) Conditional probability problems in textbooks an example from Spain. *Revista Latinoamericana de Investigación en Matemática Educativa*, 15(3), 319-337.
- Neto, M. T. B. (2009). *O desenvolvimento do raciocínio dedutivo ao nível do Ensino Secundário: Recurso a geometrias planas* (Dissertação de Doutoramento não publicada). Universidade de Aveiro, Aveiro, Portugal.
- Ponte, J. P., Serrazina, L., Guimarães, H.M., Breda, A. Guimarães, F., Sousa, H., ... e Oliveira, P. A. (2001). *Programa de Matemática do Ensino Básico*. Lisboa, Portugal: Ministério da Educação.
- Ponte, J. P. (1994). O estudo de caso na investigação em educação matemática. *Quadrante*, 3(1), 3-18.
- Sánchez, E. (1996). Dificultades en la comprensión del concepto de eventos independientes. En F. Hitt (Ed.), *Investigaciones en Educación Matemática* (pp. 389-404). D.F., México: Grupo Editorial Iberoamericano.
- Sobreiro, D. (2011). *Probabilidade condicionada: um estudo com alunos do ensino secundário* (Dissertação de Mestrado não publicada). Universidade de Aveiro, Aveiro, Portugal.
- Tesch, R. (1990). *Qualitative research: Analysis Types and Software Tools*. New York, USA: Falmer.
- Tversky, A. & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293-315.
- Watson, J. (1995). Conditional probability: Its place in the mathematics curriculum. *Mathematics Teacher*, 88(1), 12-17.
- Watson, J. M. & Moritz, J. B. (2002). School students' reasoning about conjunction and conditional events. *International Journal of Mathematical Education in Science and Technology*, 33(1), 59-84. doi: 10.1080/00207390110087615.
- Way, J. (February, 2003). *The development of young children's notions of probability*. Paper presented at the European Research in Mathematics Education III, Bellaria, Itália: CERME3. Recuperada de: [http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5\\_way\\_cerme3.pdf](http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_way_cerme3.pdf)

## Autoras

---

**Maria José Carvalho.** Agrupamento de Escolas António Sérgio, Vila Nova de Gaia, Portugal. [maria.rodrigues.carvalho@gmail.com](mailto:maria.rodrigues.carvalho@gmail.com)

**Adelaide Freitas.** Departamento de Matemática & CIDMA, Universidade de Aveiro, Portugal. [adelaide@ua.pt](mailto:adelaide@ua.pt)



## APÊNDICE I

*Situação-problema P1: enunciado e resolução*

1. Um certo estudo numa maternidade revelou, acerca do sexo de 300 bebés nascidos naquela maternidade e das correspondentes previsões de sexo a partir da ecografia na 13ª semana de gestação, os seguintes resultados:

		<i>Sexo verdadeiro (ao nascer)</i>	
		Feminino	Masculino
<i>Sexo previsto na ecografia</i>	Feminino	127	96
	Masculino	9	68

Com base nestes resultados, e usando 3 casas decimais nos arredondamentos, responda às seguintes questões.

- 1.1 Calcule a probabilidade de um bebé, ao nascer, ser do sexo masculino

*Resolução:* Considere os seguintes acontecimentos:

M: “Ser menino”

R: “previsto ser menino”  $M^C$  “Ser menina”

$R^C$ : “previsto ser menina”

Com estas designações, a tabela dada e completa ficaria:

		<i>Ser</i>		<i>Total</i>
		$M^C$	M	
<i>Previsto ser</i>	$R^C$	127	96	223
	R	9	68	77
<i>Total</i>		136	164	300

Probabilidade pedida:  $P(M) = 164/300 = 0,547$

- 1.2 Determine a probabilidade de ser menino se a ecografia faz prever ser uma menina.

*Resolução:* Probabilidade pedida:  $P(M|R^C) = P(M \cap R^C)/P(R^C) = 96/223 = 0,430$

- 1.3 A Sra. Ana e a Sra. Berta estão a ser seguidas naquela maternidade. Ambas realizaram uma ecografia na 13ª gestação. À Sra. Ana a ecografia previu uma menina e à Sra. Berta um menino. Em que caso é mais provável ser

confirmado, no nascimento, a previsão do sexo indicada pela ecografia? Numa pequena composição fundamente a sua resposta.

Tópicos a referenciar:

- Probabilidade de ser confirmado, no nascimento, que o bebé da Sra. Ana é do sexo feminino.
- Probabilidade de ser confirmado, no nascimento, que o bebé da Sra. Berta é do sexo masculino.

*Tópicos da Resolução:* Na composição deve estar o cálculo das probabilidades condicionadas em causa,  $P(M^c|R^c)$  e  $P(M|R)$ , e indicado qual das duas probabilidades tem maior valor.

$$\text{Sra Ana : } P(M^c|R^c) = 1 - P(M|R^c) = 1 - 0,430 = 0,570$$

$$\text{Sra Berta: } P(M|R) = P(M \cap R)/P(R) = 68/77 = 0,883$$

*Para a Sra Berta é mais provável ser confirmado, no nascimento, o sexo masculino do bebé previsto na ecografia da 13ª semana de gestação.*

- 1.4 Um casal com três filhos, todos nascidos naquela maternidade, é selecionado ao acaso. Considere que o sexo de uma criança é independente do sexo dos irmãos e que a probabilidade de qualquer filho do casal ser do sexo masculino é igual à probabilidade calculada na alínea 1.1.

Considere os acontecimentos:

A: “O casal ter no máximo uma rapariga”

B: “O casal ter filhos de ambos os sexos”

C: “O casal só ter rapazes”

- a) Calcule a probabilidade de cada um dos acontecimentos.

*Resolução:* Do enunciado, resulta:

$$P(\text{“Um filho do casal ser do sexo masculino”}) = P(M) = 0,547$$

Dado um casal com 3 filhos, o espaço de resultados possíveis é:

$$\Omega = \{MMM, MMM^c, MM^cM, MM^cM^c, M^cMM, M^cMM^c, M^cM^cM, M^cM^cM^c\}$$

*Uma vez que  $P(M) = 0,547$  e, consequentemente,  $P(M^c) = 0,453$ , os acontecimentos  $M$  e  $M^c$  não são equiprováveis. Logo, os 8 acontecimentos elementares em  $\Omega$  não são equiprováveis. As probabilidades pedidas, porque existe independência do sexo entre irmão, são então calculadas do seguinte modo:*

$$\begin{aligned}
* P(A) &= P(MMM \cup MMM^c \cup MM^cM \cup M^cMM) = \\
&= P(MMM) + P(MMM^c) + P(MM^cM) + P(M^cMM) = \\
&= P(M)P(M)P(M) + P(M)P(M)P(M^c) + P(M)P(M^c)P(M) + P(M^c)P(M)P(M) = \\
&= 0,547^3 + 3 \times 0,547^2 \times 0,453 = 0,570 \\
* P(B) &= P(MMM^c \cup MM^cM \cup M^cMM \cup MM^cM^c \cup M^cMM^c \cup M^cM^cM) = \\
&= 1 - (P(MMM) + P(M^cM^cM^c)) = 1 - (P(M)P(M)P(M) + P(M^c)P(M^c)P(M^c)) = \\
&= 1 - (0,547^3 + 0,453^3) = 0,743 \\
* P(C) &= P(MMM) = 0,547^3 = 0,164
\end{aligned}$$

- b) Os acontecimentos A e B são independentes? Justifique.

*Resolução:*

$$\begin{aligned}
P(A \cap B) &= P(MMM^c \cup MM^cM \cup M^cMM) = P(MMM^c) + P(MM^cM) + P(M^cMM) = \\
&= 3 \times (0,547^2 \times 0,453) = 0,407
\end{aligned}$$

$$P(A)P(B) = 0,570 \times 0,743 = 0,424$$

Como,  $P(A \cap B) \neq P(A)P(B)$ , A e B não são independentes.

- c) Mostre que os acontecimentos  $A \cap B$  e C são incompatíveis.

*Resolução:* Uma vez que  $(A \cap B) \cap C = \{MMM^c, MM^cM, M^cMM\} \cap \{MMM\} = \{\}$ , então os acontecimentos  $A \cap B$  e C são incompatíveis

*Outra resolução:*  $(A \cap B)$  e C são incompatíveis, uma vez que é impossível um casal de 3 filhos ter filhos de ambos os sexos, com apenas uma menina, e simultaneamente serem todos rapazes.

- d) Tendo em conta as alíneas anteriores, justifique que os acontecimentos  $A \cap B$  e C não podem ser independentes e apresente uma situação (com outras condições de enunciado) em que o poderiam ser.

*Resolução:* Como  $A \cap B$  e C são incompatíveis então,  $P((A \cap B) \cap C) = 0$ .

Mas,  $P(A \cap B)P(C) = 0,407 \times 0,164 \neq P(A \cap B \cap C)$ . Logo,  $A \cap B$  e C não são independentes.

*Uma situação hipotética em que poderiam ser independentes era admitir que um dos acontecimentos,  $A \cap B$  ou C, tivesse probabilidade nula de ocorrência. Por exemplo, admitir que só se consideram casais com 3 filhos tendo pelo menos 1 menina. Assim,  $P(C) = 0$ .*

## APÊNDICE II

*Situação-problema P1: enunciado e resolução*

2. A leptospirose é também conhecida como doença de Weil. Esta doença é causada por duas espécies de bactérias. Numa população de ratos, a probabilidade de encontrar um rato portador destas bactérias é 6%. Um novo teste de diagnóstico da doença foi proposto. Para avaliar a qualidade deste novo teste, foram efetuados vários testes àquela população de ratos, no sentido de detetar a existência destas bactérias. Do estudo efetuado resultaram as seguintes conclusões:
- A probabilidade de um rato ter teste positivo (P), sabendo que é portador das bactérias, é 99,8%;
  - A probabilidade de um rato ter um teste negativo, sabendo que não é portador destas bactérias, é de 99,6%

Seja A o acontecimento “Rato ser portador das bactérias”.

- 2.1 Calcule o valor das seguintes probabilidades:

i)  $P(A \cap P)$

ii)  $P(A|P)$

*Resolução:* Do enunciado dado, resulta:

$$P(A) = 0,06$$

$$P(P|A) = 0,998$$

$$P(P^c|A^c) = 0,996$$

onde  $P^c$  e  $A^c$  representam o acontecimento complementar de P e A, respetivamente.

Probabilidades pedidas:

$$i) P(A \cap P) = P(A) P(P|A) = 0,06 \times 0,998 = 0,05988$$

$$ii) P(A|P) = P(A \cap P) / P(P) = 0,05988 / 0,06364 = 0,940918$$

*uma vez que*

$$P(P) = P(A \cap P \cup A^c \cap P) = P(A \cap P) + P(A^c \cap P) = 0,05988 + P(A^c) P(P|A^c) = 0,05988 + (1 - 0,06)(1 - 0,996) = 0,06364$$

- 2.2 Os acontecimentos A e P são incompatíveis? Justifique convenientemente a sua resposta.

*Resolução:* Não, porque  $P(A \cap P) \neq 0$ . Consequentemente,  $A \cap P$  nunca pode ser o acontecimento impossível (i.e., o acontecimento  $\{\}$ )

- 2.3 Os acontecimentos A e P são independentes? Justifique convenientemente a sua resposta.

*Resolução:* Não, porque a probabilidade de ocorrência do acontecimento A toma valores diferentes consoante se conheça a ocorrência de P ou não:

$$P(A) = 0,06 \text{ e } P(A|P) = 0,940918$$

*Outra resolução:* Não, porque  $P(A)P(P) = 0,06 \times 0,06364 = 0,003818$ , que é diferente de  $P(A \cap P) = 0,05988$ .

- 2.4 Escolhendo ao acaso um rato dessa população, qual é a probabilidade de ele ser portador das bactérias, sabendo que o teste efetuado deu negativo?

*Resolução:* Probabilidade pedida:

$$P(A|P^c) = P(A \cap P^c) / P(P^c) = 0,00012 / 0,93636 = 0,000128$$

pois

$$P(A \cap P^c) = P(A)P(P^c|A) = P(A)(1 - P(P|A)) = 0,06 \times (1 - 0,998) = 0,00012$$

$$P(P^c) = 1 - P(P) = 1 - 0,06364 = 0,93636.$$

