Auriluci de Carvalho Figueiredo

PROBABILIDADE CONDICIONAL:

"Um enfoque de seu ensino-aprendizagem"

Mestrado em Educação Matemática

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de MESTRE em Educação Matemática, sob a orientação do Professor Doutor Benedito Antonio da Silva.

PUC/SP
São Paulo
2000

BANCA EXAMINADORA

RESUMO

Este trabalho tem por objetivo introduzir o conceito da Probabilidade Condicional em cursos de Estatística na Universidade. Para isso, elaboramos, aplicamos e analisamos os resultados de uma seqüência de ensino levando em consideração os princípios de uma Engenharia Didática.

Esta seqüência de ensino é composta de quatro atividades que foram criadas, baseando-se nas situações didáticas apresentadas por Carmen Batanero e outros autores, com o intuito de fazer o aluno refletir sobre circunstâncias que envolvam não só a Probabilidade Condicional, bem como os conceitos ligados ao Teorema da Probabilidade Total e o Teorema de Bayes.

Para trabalhar com tais conceitos, articulamos nas questões das atividades, diferentes registros de representação: linguagem natural, simbólica, diagrama de árvore e tabela de contingência, tomando como base a Teoria de Registros de Representação de Raymond Duval.

Aplicamos esta seqüência aos alunos dos cursos de Licenciatura de Matemática e Ciência da Computação e diante dos protocolos desses alunos, concluímos que nossa seqüência os auxiliou a minimizar as dificuldades levantadas por nós e pelos pesquisadores e ao mesmo tempo indicou temas para futuras pesquisas na área.

Dentre outras conclusões, ressaltamos que a maioria dos alunos diante de questões que envolvam a Probabilidade Condicional diferenciavam esta da Probabilidade da Interseção de Eventos e o Cálculo da P(A/B) do de P(B/A), desde que estes se apresentassem nas perguntas em linguagem natural. No entanto, quando questões análogas foram apresentadas na linguagem simbólica, muitos alunos mostraram dificuldades em resolvê-las.

Palavras – chave: probabilidade condicional, situações didáticas, registros de representação, diagrama de árvore, tabela de contingência.

ABSTRACT

This work has for objective to introduce the concept of the Conditional Probability in courses of Statistics in the University. For that we elaborated, we applied and we analyzed the results of a teaching sequence taking in consideration the beginnings of a Didactic Engineering.

This teaching sequence is composed of four activities that they were created, basing on Carmen Batanero's didactic situations and other, with the aim of doing the student to contemplate on circumstances that involve not only the Conditional Probability, as the concepts linked to the Theorem of the Total Probability and the Theorem of Bayes.

To work with such concepts we articulated in the subjects of the activities different representation registrations: natural, symbolic language, diagram of trees and contingency table, taking as base the Theory of Registrations of Representation of Raymond Duval.

We applied this sequence to students of the courses of Licentiate of Mathematics and Science of the Computation and before the those students' protocols, we concluded that our sequence aided them they decrease it the difficulties lifted by us and for the researchers and at the same time it indicated themes for future researches in the area.

Enter other conclusions stood out that most of the students before subjects that involve the Conditional Probability differentiated this of the Probability of the intersection of events and I calculate it of P(A/B) of the one of P(B/A), since these if they come in the questions in natural language. However when similar subjects were presented in the language symbolic many students they showed difficulties in solving them.

Words - key: conditional probability, didactic situations, representation registrations, tree diagram, contingency table

AGRADECIMENTOS

Ao meu orientador, Professor Doutor Benedito Antonio da Silva, por ter aceito o desafio de orientar este trabalho, mostrando sempre confiança e dedicação, incentivando-me em todo o nosso percurso.

À Professora Doutora Dione Lucchesi de Carvalho, por ter contribuído com bibliografias pertinentes, ter aceito fazer parte da banca examinadora e por sugerir modificações relevantes neste trabalho.

Ao professor Doutor Saddo Ag Almouloud, pelas importantes observações que muito melhoraram a apresentação deste texto.

Aos professores Cibele de Almeida Souza e Stefan Kabbach, que se mostraram cooperativos e interessados na elaboração desta pesquisa, e acompanhando o nosso trabalho durante algumas de suas aulas.

Aos alunos do 3º ano dos cursos de Licenciatura de Matemática e Ciência da Computação, da UNISANTOS, pelo entusiasmo com que responderam as questões propostas em classe nas atividades da nossa seqüência didática.

À direção da Universidade Católica de Santos, por ter auxiliado no nosso trabalho, permitindo a aplicação da nossa seqüência de ensino.

Ao Ten.-Cel.-Av. Paulo César da Silva Duarte, que me proporcionou condições para realizar esta obra, concedendo-me algum tempo de liberação das minhas atividades profissionais.

Ao meu marido, pelo seu incentivo constante para a execução desse trabalho.

Ao meu filho, pela ajuda nos recursos de informática, pela compreensão e apoio ao longo do caminho.

PROBABILIDADE CONDICIONAL:

“Um enfoque de seu ensino-aprendizagem”

ÍNDICE

Introdução.....07

Capítulo I: Estudos Estocásticos.....10

1.1 O Objeto Estocástico.....10

1.2 O Ensino da Probabilidade e da Estatística.....24

1.2.1 A Probabilidade e a Estatística nos Parâmetros Curriculares Nacionais.....24

1.2.2 Algumas Pesquisas na Área.....28

Capítulo II: Breve Histórico da Probabilidade Condicional.....34

Capítulo III: Da Problemática aos Procedimentos Metodológicos.....46

3.1 Problemática.....46

3.2 Fundamentação Teórica.....52

3.3 Metodologia da Pesquisa.....58

Capítulo IV: Pré-Experimentação.....61

Capítulo V: A Seqüência Didática.....76

5.1 Análise a Priori.....78

5.2 A Realização da Seqüência.....107

5.3 Análise dos Resultados.....112

Conclusões.....142

Bibliografia.....146

Anexos.....151

INTRODUÇÃO

A nossa primeira experiência com o ensino da Estatística se deu somente em 1995, quando lecionávamos no ILA (Instituto de Logística da Aeronáutica) uma disciplina considerada obrigatória para o ingresso no curso CELOG (Curso de Extensão em Logística), que tem como objetivo rever conceitos da Estatística Descritiva, embora nossa formação seja em Matemática e venha atuando há quinze anos no ensino Fundamental, bem como no Médio e Superior.

Considerávamos que os alunos com curso superior completo e oriundos de diversas áreas de Exatas, já tivessem tido contato com a Estatística. Porém ao longo das nossas quarenta aulas do curso preparatório para o CELOG, observamos que as dificuldades eram muitas. Essa disciplina se repetia a cada ano com novos alunos, e as dificuldades apresentadas por eles eram as mesmas. Resolvemos, então, procurar alguma bibliografia que tratasse das dificuldades que os alunos têm quando estudam Estatística.

Encontramos algumas pesquisas publicadas no Brasil e várias na França, EUA, Espanha e em outros países do mundo, algumas das quais descreveremos nos próximos capítulos.

Logo percebemos que não estávamos sozinhos diante de entraves relacionados ao ensino da Estatística, e que outras pessoas também tinham levantado algumas das dificuldades por nós encontradas e muitas outras.

Diante dessa situação, descobrimos que existe um campo de pesquisa muito grande no ensino da Estatística tanto para o Ensino Fundamental e Ensino Médio como na Universidade.

Com uma grande quantidade de temas possíveis para uma dissertação de mestrado, foi difícil, a princípio, escolher um deles. Elegemos trabalhar com o “Conceito de Probabilidade Condicional” na Universidade.

Nosso trabalho tem por objetivo introduzir o Conceito de Probabilidade Condicional nos cursos básicos de Estatística no Ensino Superior.

No Capítulo I, **“Estudos Estocásticos”**, consta **“O Objeto Estocástico”**, que contêm algumas sugestões de como trabalhar o **Conceito de Probabilidade Condicional,** incluindo algumas maneiras de sua **representação**. Neste mesmo capítulo, interessados em sabermos quando os alunos deveriam ter seus primeiro contatos com a Probabilidade e a Estatística no ensino, tomamos como base algumas pesquisas nesta área no Brasil, os Parâmetros Curriculares Nacionais, as Propostas Curriculares do Ensino Médio de São Paulo e as Matrizes Curriculares de Referência para o Sistema de Avaliação Escolar da Educação Básica

No Capítulo II, **“Breve Histórico da Probabilidade Condicional”**, estudamos a origem e a evolução do Conceito da Probabilidade Condicional. Esse estudo contribuiu para as escolhas das atividades propostas aos alunos, principalmente no que diz respeito às dificuldades de representação que foram encontradas na evolução desse conceito.

No Capítulo III, **“Da Problemática aos Procedimentos Metodológicos”**, nos referimos a algumas pesquisas sobre a Probabilidade e a Probabilidade Condicional, levantando algumas dificuldades apresentadas por alunos diante de situações que envolvam tal conceito, e à teoria sobre a qual vamos nos fundamentar, com o intuito de elaborarmos uma série de atividades que contribuam para que os alunos construam esse conceito. Como Metodologia de Pesquisa, tomaremos como base os princípios de Engenharia Didática para a elaboração de uma seqüência de ensino, que será aplicada aos alunos dos Cursos de Licenciatura em Matemática e Ciência da Computação.

No Capítulo IV, **“Pré-Experimentação”**, relatamos a aplicação de um teste para investigar a existência de algumas dificuldades dos alunos ao trabalharem com o Cálculo de Probabilidades que envolvem o Conceito da Probabilidade Condicional. Dificuldades estas que nos auxiliou em nossas escolhas para a elaboração das atividades da seqüência de ensino.

No Capítulo V, **“A Seqüência Didática”**, descrevemos as atividades juntamente com uma análise que faz a previsão dos possíveis comportamentos dos alunos frente a elas, como transcorreu sua aplicação e uma análise dos resultados.

Nas **Conclusões,** apresentamos os resultados dos testes que aplicamos aos alunos que efetuaram as atividades da seqüência didática. Esses resultados foram comparados com as questões que levantamos em nossa Problemática.

CAPÍTULO I: ESTUDOS ESTOCÁSTICOS

Neste capítulo mostraremos alguns estudos estocásticos que nos influenciaram no desenvolvimento deste trabalho. A palavra “estocástica” é um termo europeu comum para incluir “probabilidade e estatística”.

No “Objeto Estocástico” apresentaremos algumas abordagens da representação da Probabilidade Condicional, mostrando as articulações entre diferentes registros com o intuito de facilitar a leitura de nossas propostas.

Para a elaboração deste trabalho, partimos da indagação sobre quando a Estocástica e o Conceito de Probabilidade Condicional, em particular, deveriam ser trabalhados com o aluno. Para isso, tomamos como base algumas pesquisas nesta área no Brasil, os Parâmetros Curriculares Nacionais, as Propostas Curriculares do Ensino Médio de São Paulo e as Matrizes de Referência para o Sistema de Avaliação Escolar da Educação Básica.

1.1 O OBJETO ESTOCÁSTICO

Preliminarmente, elencaremos alguns pontos do objeto estocástico que intervêm de modo central em nosso estudo:

- Mostrar como o conceito da Probabilidade Condicional é introduzido em alguns livros didáticos;
- Apresentar as Probabilidades Condicionais “Cronológicas”, “Causais” e “Conjuntistas”, analisando as relações de dependência e independência dos eventos;
- Representar as soluções de problemas que mobilizam o conceito da Probabilidade Condicional através de “diagrama de árvore” e “tabela de contingência”; e
- Utilizar o “diagrama de árvore” e a “tabela de contingência” como ferramenta para resolver problemas que mobilizam os conceitos que envolvem o Teorema da Probabilidade Total e o Teorema de Bayes.

Mostraremos como se apresenta o conceito de Probabilidade Condicional no ensino superior brasileiro, na maioria dos livros didáticos que constam de nossa Bibliografia. Esses livros colocam um exemplo, e a partir dele, mostra-se a necessidade de introduzir tal conceito.

Vamos nos apropriar do conceito de condicional apresentado por Moretin e Bussab (1987), para em seguida fazermos algumas considerações:

“Suponha que o seguinte quadro represente uma possível divisão dos alunos matriculados em um dado instituto de matemática, num dado ano.

SEXO / ***CURSO***	***Homens (H)***	***Mulheres (F)***	***TOTAIS***
Matemática Pura (M)	***70***	***40***	***110***
Matemática Aplicada (A)	***15***	***15***	***30***
Estatística (E)	***10***	***20***	***30***
Computação (C)	***20***	***10***	***30***
Totais	***115***	***85***	***200”***

Segundo Moretin um estudante matriculado no curso de Estatística é escolhido ao acaso.

“ ... a probabilidade de ele ser mulher é de 20/30 = 2/3. Isto porque do total de 30 alunos que estudam Estatística, 20 são mulheres. Escrevemos:

P(mulher/Estatística) = 2/3”

Para dois eventos quaisquer A e B, sendo P(B) > 0, definimos a probabilidade condicional de A, dado B, P(A/B), como sendo:

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

(relação 1)"

Sendo que:

P(B) é a probabilidade de ocorrer o evento B;

P(A/B) é a probabilidade de ocorrer o evento A, sabendo-se que ocorreu B; e

P(A∩B) é a probabilidade de ocorrer A e ocorrer B.

"Para o exemplo mencionado, se B e A indicam, respectivamente, os eventos "aluno matriculado em Estatística" e "aluno é mulher", então

$P(A \cap B) = 20/200$, $P(B) = 30/200$, portanto

$$P(A/B) = \frac{20/200}{30/200} = \frac{2}{3},$$

como havíamos obtido.

Da relação (1) obtemos a chamada regra do produto de probabilidades,

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

Precisamos agora saber o que deve existir entre os eventos A e B para que possamos fazer uso da relação (1).

Considerando $A \subset \Omega$ e $B \subset \Omega$, onde Ω é o conjunto formado de todos os eventos, isto é, A e B são sub-conjuntos do espaço amostral Ω.

1) Se A e B são independentes, então $P(A/B) = P(A)$.
   Definimos nesse caso $P(A \cap B) = P(A).P(B)$.
   Podemos dizer que para o evento A ocorrer, independe de B.

2) Se A e B são dependentes então:

$$\mathbf{P(A \cap B) = P(A).P(A/B)\ ou\ P(A/B) = \frac{P(A \cap B)}{P(A)}}$$

Percebemos que, com os itens (1) e (2) não nos fornecem condições reais de analisarmos que relação deve existir entre A e B de modo que possamos classificá-las em independentes ou dependentes.

A ligação entre A e B pode ser de natureza lógica quando A é realizado, então B também é. Neste caso, existe simplesmente a inclusão $A \subset B$, o que podemos traduzir por: $P(A) \leq P(B)$.

Segundo Michel Henry:

“... ***esta ligação pode ser de natureza estocástica: quando se repete a experiência os elementos A e B são mais ou menos freqüentemente realizados conjuntamente, de tal sorte que desde que se saiba que um dos dois é realizado, isto dá uma idéia de natureza probabilista sobre a realização eventual do outro. Quando esta freqüência é elevada, fala-se de “forte correlação”, ou de dependência estocástica”.*** (Henry, 1997: 93)

Para melhor compreendermos esta questão, mostraremos algumas situações que envolvem dois eventos. Vamos analisar tais situações que revelam as probabilidades condicionais, em seguida mostraremos algumas maneiras possíveis de resolução e aproveitaremos para estudar dependência ou independência de dois eventos.

A) Situações cronológicas

A experiência é composta de dois ensaios sucessivos, há a ocorrência do primeiro evento e, em seguida, se desenrola o segundo.

Exemplificando:

Uma urna (I) contém 2 bolas brancas (B) e 3 vermelhas (V). Uma segunda urna (II) contém 4 bolas brancas e 2 vermelhas. Escolhe-se ao acaso uma urna e dela retira-se, também ao acaso, uma bola. Qual a probabilidade que ela seja branca?

A cor da bola obtida depende de dois acasos sucessivos, ela é condicionada à ocorrência do primeiro evento (escolha da urna). A probabilidade de ter uma bola branca é então composta das probabilidades: P(I), P(II) e P(B), podendo ser calculada por:

$$\mathbf{P(I).P(B/I) + P(II).P(B/II)},$$

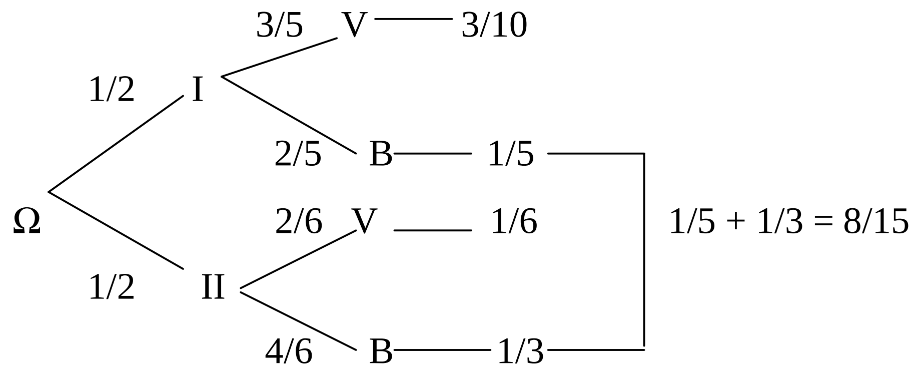
isto é, probabilidade de ser branca vindo da urna I mais a probabilidade de ser branca vindo da urna II.

- Como as urnas são escolhidas ao acaso, vamos considerar eqüiprováveis as suas escolhas, então a P(I) = P(II) = 1/2.
- P(B/I) = 2/5, usando o número de casos prováveis sobre número de casos possíveis, isto é, podemos escolher duas bolas brancas dentre cinco existentes na urna I.
- P(B/II) = 4/6, pela mesma explicação da anterior.

Portanto, a Probabilidade de tirarmos uma bola branca, será:

$$\mathbf{P(B) = \frac{1}{2} \cdot \frac{2}{5} + \frac{1}{2} \cdot \frac{4}{6} = \frac{8}{15}}$$

Podemos ver a resolução deste problema através do uso da **“árvore de probabilidades”**



Isto quer dizer que a probabilidade da bola ser branca é de 8/15.

Ao trabalharmos com a árvore de probabilidades, precisamos conhecer algumas regras que favorecerão o seu tratamento. Vamos nos utilizar de alguns dos termos usados por Bernard Parzysz [28] e [29], para melhor esclarecer o uso dessa ferramenta.

O ponto de partida dessa árvore é chamado de **raiz**, que deveremos pensar em todo o nosso espaço amostral. No desenho acima vem representada pelo símbolo Ω . Dessa **raiz** saem ramificações, que Parzysz chama de **galhos**. Os números romanos colocados no final desses primeiros **galhos** juntamente com estes, constituem o **primeiro nível** da árvore. Sobre esses **galhos** coloca-se a probabilidade de cada um desses eventos acontecerem.

O **segundo nível** da árvore começa logo em seguida dos números romanos do primeiro. A região entre os primeiros galhos e os seguintes, o autor chama de **nó.**

As probabilidades referentes aos segundos galhos, serão probabilidades calculadas, sabendo-se que já ocorreu o primeiro evento, portanto, **probabilidades condicionais**.

Parzysz (1987: 233,234,235) estabelece algumas regras, tais como:

- ♦ R_1 - A cada nó N da árvore, associamos o produto dos números figurantes sobre os ramos ligando o nó à raiz.
- ♦ R_2 - As letras colocadas nos nós da árvore representam os eventos; à raiz corresponde o conjunto fundamental (conjunto de todos os eventos possíveis de ocorrer).
- ♦ R_3 - O número atribuído ao galho $\mathrm{I} \rightarrow V$ da árvore é a probabilidades condicional de ocorrer V, sabendo-se que ocorreu I, o mesmo ocorre com os galhos $\mathrm{I} \rightarrow B$, $\mathrm{II} \rightarrow V$ e $\mathrm{II} \rightarrow B$ (observar a árvore da página anterior).
- ♦ R_4 - A cada nó N da árvore poderemos corresponder a interseção dos eventos associados aos nós da seqüência que liga este nó à Ω; e a probabilidade dessa interseção é o produto das probabilidades sobre os galhos associadas às probabilidades correspondentes a esses nós. Podemos verificar a utilização desta regra na árvore de probabilidades da página anterior que mostra 3/10 como sendo a $P(\mathrm{I} \cap V)$, e a probabilidade desta interseção é o produto 1/2.3/5.
- ♦ R_5 - A soma dos números associados a todos os galhos de mesma origem é igual a 1.
- ♦ R_6 - A soma dos números associados aos nós de um nível dado é igual a 1.

Observamos que, no registro de árvore, as Probabilidades Condicionais correspondem aos segundos galhos e a Probabilidade da Interseção dos Eventos à seqüência desses eventos desde a raiz, ocupando portanto, lugares diferentes nesse registro. Acreditamos que este fato favoreça aos alunos a visualização da diferença entre as Probabilidades Condicionais das suas Interseções, isto é, diferenciar $P(A/B)$ da $P(A \cap B)$. Sabemos que esta diferença muitas vezes é difícil para os alunos conseguirem identificar.

Outra maneira de ver a resolução desse problema, é apresentando uma **tabela de contingência** ou também chamada de **dupla entrada.**

Urnas / BOLAS	I	II	Total
B	**1/5**	**1/3**	**8/15**
V	**3/10**	**1/6**	**7/15**
Total	**5/10**	**3/6**	**1**

Observamos que:

$P(I \cap B) = 1/5$, isto é, a probabilidade da urna I ser a escolhida e uma bola branca ser retirada;

$P(II \cap B) = 1/3$, a probabilidade da urna II ser a escolhida e uma bola branca ser retirada;

$P(I \cap V) = 3/10$, a probabilidade da urna I ser escolhida e uma bola vermelha ser retirada;

$P(II \cap V) = 1/6$, a probabilidade da urna II ser escolhida e uma bola vermelha ser retirada;

$P(B) = 8/15$, a probabilidade da bola ser branca vindo da urna I ou II. Podemos observar que essa probabilidade é encontrada através da soma das probabilidades da linha correspondente;

$P(V) = 7/15$, a probabilidade da bola ser vermelha vindo da urna I ou II. É encontrada também pela soma das probabilidades da linha correspondente; e

$P(I) = 1/2 = P(II)$, que é igual a soma das suas colunas correspondentes.

Podemos ver que a soma tanto da última linha quanto da última coluna também é um.

Essa maneira de apresentar a probabilidade acreditamos favorecer a visualização de todos os dados do problema.

Vamos analisar as possíveis **dependências** ou **independências** do exemplo:

- ♦ P(B) = 8/15, P(B/I) = 2/5, portanto os eventos “bola branca” e “urna I” são dependentes.
- ♦ O mesmo ocorre com os eventos “bola vermelha” e “urna I”, como podemos observar abaixo:
  P(V) = 7/15, P(V/I) = 2/5, isto é, P(V) ≠ P(V/I).

B) Situações causais

Vamos analisar essa situação através do exemplo dado por M.Henry (1987: 94)

> ***“100 cobaias são “tratadas” por três produtos que provocam uma doença M.***
> ***50 são tratadas pelo produto P_1 que provoca M com a probabilidade 0,25.***
> ***25 são tratadas pelo produto P_2 que provoca M com a probabilidade 0,5.***
> ***25 são tratados pelo produto P_3 que provoca M com a probabilidade de 0,3.***

Uma cobaia tirada ao acaso tem a doença M. Qual é a probabilidade de ter sido tratada por P_1? ”

Encontramos nesse exemplo a relação entre os eventos, como sendo de **causa**, isto é, cada produto utilizado ocasiona a doença M com probabilidades diferentes. Podemos também pensar que M pode ser causada por diferentes produtos.

Considerando então:

$P(P_1/M)$ é a probabilidade de um indivíduo escolhido ao acaso, dentre os que contraíram a doença M, ter sido tratado pelo produto P_1 .

Podemos resolver este problema primeiro através da fórmula relacionada ao Teorema de Bayes, que colocamos a seguir:

$$P(P1/M)=\frac{P(M/P1)\cdot P(P1)}{P(M/P1)\cdot P(P1)+P(M/P2)\cdot P(P2)+P(M/P3)\cdot P(P3)}$$

Vamos extrair do enunciado do problema algumas informações:

$P(P_1) = 0,5$; existem 50 cobaias em 100 que serão tratadas pelo produto P_1 .
$P(P_2) = 0,25$; existem 25 cobaias em 100 que serão tratadas pelo produto P_2 .
$P(P_3) = 0,25$; existem 25 cobaias em 100 que serão tratadas pelo produto P_3.

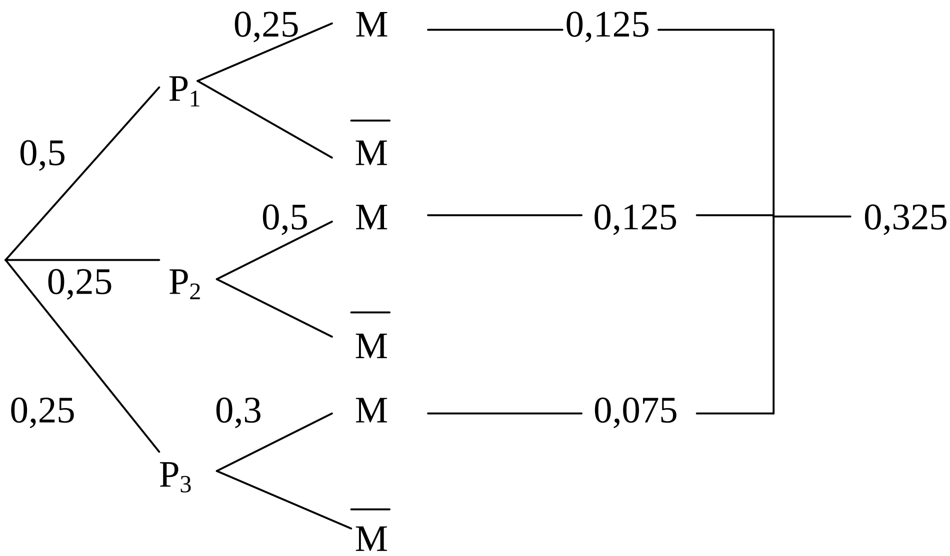
Sabemos que as cobaias tratadas por P_1, contraem a doença M com probabilidade 0,25, isto quer dizer que $P(M/P_1) = 0,25$ e da mesma forma, temos $P(M/P_2) = 0,5$ e $P(M/P_3) = 0,3$.

Substituindo na fórmula, obtemos:

$$P(P1/M)=\frac{0,25.0,5}{0,25.0,5+0,5.0,25+0,3.0,25}=\frac{0,125}{0,325}=0,384$$

Logo, a probabilidade de uma cobaia tratada pelo produto P_1 contrair a doença M é 0,384.

Através da "**árvore de probabilidades"** podemos também representar uma forma de solução.



$\overline{M}$ representa o evento “ser tratado pelos produtos respectivos, mas não contrair” a doença M.

Através da árvore temos que 0,325 é a probabilidade de uma cobaia, tratada por alguns dos produtos, ter contraído a doença M. Temos que calcular ainda a probabilidade do produto utilizado ter sido o P_1 .

Utilizando-se da divisão de probabilidade dos casos favoráveis pela probabilidade dos casos possíveis, temos: 0,125/0,325 = 0,384. Nessa resolução, não precisamos diretamente do uso da fórmula de Bayes, apesar de implícita nos cálculos feitos.

Podemos, também, como no outro exemplo, colocarmos os dados em uma tabela de **“dupla entrada”**:

P / Doença	P_1	P_2	P_3	Total
M	0,125	0,125	0,075	0,325
—	0,375	0,125	0,175	0,675
Total	0,5	0,25	0,25	1

Temos $P(M)= 0,325$ e , basta dividir 0,125 por 0,325 para obtermos a resposta 0,384.

Analisando a **dependência** ou não dos eventos “tratada por P_1 ” e “ter contraído M”, podemos dizer: como $P(P_1) = 0,25$ e $P(P_1/M) = 0,384$, os eventos P_1 e M são dependentes.

Verificamos nestas duas últimas representações da resolução do problema das cobaias, que é possível resolvermos questões que mobilizam conceitos que envolvem o Teorema de Bayes, sem precisar utilizar uma fórmula, como foi apresentada na nossa primeira resposta.

O nosso trabalho se apóia em resoluções de problemas que envolvem conceitos de Probabilidades Condicionais, Teorema da Probabilidade Total e Teorema de Bayes de maneira que, ao trabalharmos na montagem dos diagramas de árvore e das tabelas de contingência, tais conceitos são mobilizados quase que “intuitivamente”.

C) Situações “Conjuntistas”

Consideremos a tabela dos alunos matriculados em um dado instituto de Matemática, utilizado na introdução do conceito de Probabilidade Condicional, como um exemplo de situação “**conjuntista”** ou também chamada de “**cardinalista”.**

Tiramos ao acaso um aluno, observamos o “curso” e o “sexo” e destacamos os seguintes eventos: A, o aluno é do curso de estatística e B, o aluno é mulher. Podemos calcular a probabilidade de ocorrer o evento A, sabendo que o aluno é mulher, isto é, $P(A/B)$, e também calcular a probabilidade de ocorrer o evento B, sabendo que o aluno é do curso de Estatística, isto é, $P(B/A)$. Destacamos as Probabilidades Condicionais $P(A/B)$ e

P(B/A), pois sabemos que isto representa uma grande dificuldade para os estudantes.

Também nesse caso, podemos utilizar o diagrama de árvores ou a tabela de contingência, porém vamos mostrar uma maneira diferente de tratar esse problema, que é através da linguagem de conjunto, utilizando o diagrama de Venn, onde H representa os alunos homens , E representa os alunos de ambos os sexos de Estatística e F as alunas do Instituto.

H E F

105 10 20 65

Como o nº de elementos de E é 30. E o nº de elementos de F sabendo que ocorreu E é 20 (número de elementos $E \cap F$).

Logo P(F/E) = 20/30 = 2/3.

Podemos calcular também P(E/F) observando o diagrama, isto é, sabemos que o nº de elementos de F é 85 e o nº de elementos de E sabendo que ocorreu F é 20, logo:

$$P(E/F) = \frac{20}{85} = \frac{4}{17}$$

Analisando **independência** ou **dependência** no exemplo, temos:

P(E) = 30/200 = 3/20 e P(E/F) = 4/17, logo os eventos E e F são dependentes.

As situações cronológicas, causais e “conjuntistas” apresentadas, tratam de probabilidades condicionais sempre com a dependência dos eventos destacados. Vamos mostrar um exemplo de condicional onde encontramos a independência dos eventos.

Uma urna contém 3 bolas brancas (B) e 9 bolas verdes (V). Escolhemos duas bolas ao acaso, com reposição, isto é, retiramos a primeira e observamos sua cor, em seguida recolocamos a bola na urna. Retiramos, então, a segunda bola. Qual a probabilidade das duas bolas serem brancas?

P(1ª bola B) = 3/12, P(2ª bola B/ 1ª bola B) = 3/12. Logo os eventos “1ª bola B” e “2ª bola B” são independentes.

O mesmo ocorre com eventos “1ª bola V” e “2ª bola B”, “1ª bola B” e “2ª bola V”, “1ª bola V” e “2ª bola V”.

Pelos exemplos dados, já podemos observar que é possível trabalharmos com os Teoremas de Probabilidade Total e de Bayes, de maneira implícita, através das resoluções por **diagramas de árvore** e **tabelas de contingência.** Acreditamos que, se o aluno compreender bem o conceito de **Probabilidade Condicional**, ele será capaz de aplicar ambos os teoremas quase intuitivamente.

1.2 O ENSINO DA PROBABILIDADE E DA ESTATÍSTICA

Com o objetivo de sabermos as recomendações do MEC para o Ensino da Probabilidade e da Estatística, tanto no Ensino Fundamental como no Ensino Médio, fizemos uma busca nos Parâmetros Curriculares Nacionais. Em seguida complementamos com alguns aspectos apresentados pela Proposta Curricular do Estado de São Paulo e com o documento que rege o Sistema de Avaliação Escolar (SAEB).

Procuramos relatar, também, pontos importantes encontrados por vários pesquisadores na área de Estatística e Probabilidade, assim como indicar alguns grupos que têm se dedicado a esta, com o intuito de facilitar o caminho para futuros pesquisadores.

1.2.1 A PROBABILIDADE E A ESTATÍSTICA NOS PARÂMETROS CURRICULARES NACIONAIS

Os Parâmetros Curriculares Nacionais (PCNs), elaborados pelo MEC, têm por finalidade fornecer elementos para ampliar o debate nacional sobre o ensino dos conceitos desta área de conhecimento, socializar informações e resultados de pesquisa, levando-as ao conjunto dos professores brasileiros, visando também à construção de um referencial que oriente a prática escolar.

Nos PCNs (1998) de Matemática do Ensino Fundamental, encontra-se o “Tratamento da Informação”. Nele consta que uma de suas finalidades do estudo da probabilidade é que: o aluno compreenda que muitos dos acontecimentos do cotidiano são de natureza aleatória e que se podem identificar possíveis resultados desses acontecimentos.

No 3º ciclo do ensino fundamental (5ª e 6ª séries), os conceitos e procedimentos esperados pelos alunos relativos à Probabilidade são:

- ♦ Representação e contagem de casos possíveis em situações combinatórias; e
- ♦ Construção do espaço amostral e indicação da possibilidade de sucesso de um evento pelo uso de uma razão.

No 4º ciclo do Ensino Fundamental (7ª e 8ª séries), os PCNs continuam enfatizando a necessidade de se trabalhar os conceitos e procedimentos ligados à Probabilidade, que são:

- ♦ Construção do espaço amostral, utilizando o princípio multiplicativo e a indicação da Probabilidade de um evento por meio de uma razão; e
- ♦ Elaboração de experimentos e simulações para estimar Probabilidades e verificar a Probabilidade prevista.

Nas Orientações Didáticas, apresentada nos PCNs de Matemática do Ensino Fundamental, no que se refere ao desenvolvimento da Probabilidade, são sugeridas simulações, como lançamento de moedas entre outras, para que os alunos compreendam o significado do espaço amostral e sua construção pela contagem dos casos possíveis. São enfatizados também as representações como da **tabela de dupla entrada** e o **diagrama de árvore** para que os alunos trabalhem na construção do espaço amostral.

Os objetivos do ensino da Matemática que constam nos PCNs do Ensino Médio, aparecem de forma mais genérica, como por exemplo:

> ***"Analisar e valorizar informações provenientes de diferentes fontes, utilizando ferramentas matemáticas para formar uma opinião própria que lhe permita expressar-se criticamente sobre problemas da Matemática, das outras áreas do conhecimento e da atualidade."*** (MEC, 1999 : 85)

Os procedimentos esperados pelos alunos diante do tratamento da informação, aparecem de maneira implícita, justificada pela necessidade, por serem aplicações matemáticas ligadas ao mundo real, que estão em grande crescimento, que são traduzidos por:

> ***"...descrever e analisar um grande número de dados, realizar inferências e fazer predições com base numa amostra de população, aplicar as idéias de probabilidade e combinatória a fenômenos naturais e do cotidiano..."*** (MEC,1999 : 91)

Segundo os PCNs, técnicas e raciocínios estatísticos e probabilísticos são, sem dúvida, instrumentos tanto das Ciências da Natureza, quanto das Ciências Humanas, sendo destacada, portanto, a importância de uma cuidadosa abordagem dos conteúdos de contagem, estatística e probabilidade no Ensino Médio.

O discurso referente ao ensino da Matemática, nos PCNs do Ensino Médio, parece-nos consciente da importância de que os professores trabalhem com seus alunos o "Tratamento da Informação", devido à sua aplicação no mundo de hoje, porém não fornece subsídios relativos ao possível conteúdo e nem sugestões de como trabalhar. Neste aspecto está diferindo do PCN do Ensino Fundamental.

Outros documentos que deverão orientar professores para o conteúdo a ser desenvolvido no Ensino Fundamental e Médio é o chamado Matrizes Curriculares de Referência para o Sistema de Avaliação Escolar da Educação Básica - SAEB. São apresentados nessa Matriz os "descritores", procurando indicar um conjunto de saberes significativos que privilegiam a manifestação da compreensão e do raciocínio dos alunos, a interpretação e produção de diferentes formas de representação, a diversidade de procedimentos, evitando a proposição de aspectos que possibilitem apenas a identificação de conhecimentos que devam ser memorizados.

Os conhecimentos e competências matemáticas indicadas nos descritores, foram retiradas de uma pesquisa elaborada pelos autores do documento tomando como base os Currículos dos Estados e Municípios.

Vamos destacar alguns pontos que são avaliados em Estatística e Probabilidade.

Até a 4ª série do Ensino Fundamental:

Na solução de situação – problema:

- Organizar, descrever e analisar dados;
- Construir representações gráficas, tais como: listas, tabelas simples e de dupla entrada, gráficos; e
- Comparar e interpretar dados apresentados graficamente.

Até a 8ª série do Ensino Fundamental:

Na solução de situação – problema:

- Fazer prognósticos a partir de dados apresentados em tabela ou gráficos; e
- Estimar probabilidades e verificar probabilidades previstas pela análise de dados apresentados em tabelas ou gráficos.

No Ensino Médio, na solução de situação – problema:

- Analisar dados organizados em tabelas, identificando padrões estatísticos;
- Identificar e interpretar o comportamento de dados a partir de uma representação;
- Comparar dados em diferentes representações gráficas; e
- Aplicar os conceitos relativos à probabilidade de ocorrência de um ou mais eventos.

Como podemos observar, tanto os PCNs do Ensino Médio como a Matriz Curricular de referência para o SAEB fornece pouco subsídio para orientação de trabalhos a se desenvolver com os alunos. Esperamos que os Estados e Municípios em suas Propostas Curriculares complementem as publicações citadas acima.

Dentre as Propostas Curriculares de Ensino Médio, vamos destacar as do Estado de São Paulo, mais particularmente no que se refere aos conteúdos relacionados à Estatística e à Probabilidade. São eles:

- Experimento aleatório;
- Regularidade estatística;
- Probabilidade de um evento: conceito de probabilidade de um evento, espaços eqüiprováveis, probabilidade de um evento em espaços eqüiprováveis e eventos complementares;
- Cálculo de probabilidades de eventos independentes;
- Cálculo de probabilidade de eventos dependentes: uso da probabilidade condicional; e
- Soma de probabilidades.

Nessas Propostas são apresentadas várias sugestões de problemas que abordam os conteúdos e colocam possíveis soluções para os mesmos. Mostram a ligação da solução de problemas de probabilidade, utilizando o **diagrama de árvore** como uma ferramenta somente para analisar o espaço amostral, não dando **tratamento** para o seu uso.

Os problemas que se referem à Probabilidade Condicional estão sempre relacionados às situações “conjuntistas”, mostrando a sua resolução através do Diagrama de Venn ou simplesmente de forma algébrica.

O cenário das probabilidades diante dos PCNs e da Proposta Curricular de São Paulo, apresenta a importância de seu ensino, mas parece estar longe de fornecer reais condições para professores trabalharem com este conteúdo.

1.2.2 ALGUMAS PESQUISAS NA ÁREA

Sabemos que hoje os meios de comunicação, jornais, revistas, televisão e outros, mostram diariamente muitas informações através de gráficos, de porcentagens, probabilidade de ocorrer ou não uma determinada situação, para dar algum tratamento a determinadas informações.

Esse é um dos fatos que nos fizeram refletir sobre a importância do ensino da Estatística tanto para o Ensino Fundamental e Médio como para o Superior. Percebe-se que existe grande interesse pelos métodos estatísticos nas mais variadas áreas, tais como: Medicina. Economia, Educação, Sociologia, Agronomia, Geografia e outras.

Hoje encontramos a disciplina Estatística na maioria dos currículos dos cursos superiores, seja ele de exatas, humanas ou biológicas. A expectativa seria que, em cada área, a abordagem para o seu ensino devesse ter algumas diferenças, pois há nesse cenário perfis de alunos bem diferentes.

Essa situação deveria ser percebida pelos professores de Estatística que atuam nessas várias realidades. Segundo Jacobini, muitos dos cursos de Estatística são ministrados para estudantes na área das ciências humanas e biológicas, em que geralmente os alunos possuem uma formação deficiente em Matemática.

> ***“Essa falta de conhecimento matemático é apontada pelos professores como sendo a maior dificuldade que eles enfrentam quando lecionam Estatística”*** (Jacobini, 1999: 01).

O autor enfatiza que o fato provoca muitas vezes uma insatisfação no estudante. Insatisfação essa que é transferida de turma para turma, criando assim uma certa aversão ao curso.

Paiano relata uma possível causa para a insatisfação dos alunos:

> “***Em grande parte das aulas de Matemática o professor propõe exercícios segundo o que foi explicado recentemente. Os problemas são propostos com o objetivo de aplicar as fórmulas e os algoritmos que foram estudados, não dando a oportunidade para que o aluno descubra novos caminhos e dando maior importância à resposta do que ao processo que se desencadeou para chegar até o resultado”.*** (Paiano, 1998: 42)

Barreto (1999), considera que apesar da Estatística possuir hoje uma grande aplicação no cotidiano, encontra resistência por parte dos estudantes, nas diversas áreas do seu ensino, na Universidade. Alguns alunos fazem indagações como: “Por que devo estudar isso?”, “Que benefício direto a Estatística me trará?”

Segundo esse autor:

> ***“Quanto a essa questão, julgamos que a solução se dará através da “conscientização”, que poderá ser facilitada através da contextualização dos conteúdos desenvolvidos sobre a realidade do aluno ”.*** (Barreto, 1999: 83)

Para caracterizar melhor essa contextualização, Barreto coloca como exemplo, no âmbito da saúde da população, que considera ser de interesse geral, são necessários dados como: número de casos de certas doenças, número de mortes por causa e por idade, números de habitantes servidos por água da rede pública e outros.

Com a preocupação de formar um cidadão que tenha capacidade de análise e de crítica, ao lidar com informações veiculadas em seu cotidiano, os currículos de Ensino Fundamental de vários estados brasileiros enfatizam ensino da Estatística e de Probabilidades. Lopes e Paiano destacam as Propostas Curriculares de três estados brasileiros: Minas Gerais, Santa Catarina e São Paulo:

- A proposta de **Minas Gerais** propõe a introdução desde a 3ª série da noção de combinatória e nas séries subseqüentes conteúdos como: experimentações aleatórias, conceitos de universo, idéia de probabilidade, leitura e interpretação de gráficos, construção de tabela de dados, apresentação pictorial, média, moda e mediana.
- Pela proposta de **São Paulo** somente na 8ª série se prevê explicitamente noções de Estatística, justificando esse fato apontando que é nesse

momento que o aluno conhece vários conteúdos como porcentagem, ângulos, circunferência, proporcionalidade, etc.

- Na proposta de **Santa Catarina,** as noções de Estatística aparecem desde a 2ª série do Ensino Fundamental, justificando a sua inclusão por ser uma disciplina que permite o desenvolvimento do pensamento lógico, o uso da criatividade, a construção de conceitos que podem estar ligados a realidade do aluno.

Encontramos uma preocupação nas propostas de outros países, em relação à inclusão de conteúdos de Estatística e Probabilidades em seus currículos de ensino desde a escola elementar. Alguns pontos dessas propostas são relatados por Lopes, dos quais destacaremos :

- Na Espanha: há uma preocupação com a realização de trabalhos que envolvam a realidade dos estudantes, dando ênfase às questões de caráter aleatório, às qualidades estéticas dos gráficos, à atitude crítica a ser desenvolvida no trabalho com informações.
- Na França: o ensino começa no 1º ano elementar (6-7 anos). Os educadores franceses consideram que:

  > "...***em Matemática, são necessários anos para se obter uma visão clara e para construir sistemas e não apenas elementos dispersos do conhecimento".*** (Lopes,1998: 58)

- No Japão: o Ensino Básico (shogakko) valoriza a expressão dos dados através de tabelas e gráficos, justificando o tema por sua relevância para uma sociedade informatizada.

Com a nossa investigação sobre o que está acontecendo com o ensino da Estatística, constatamos que em vários países já se formaram grupos de estudos para pesquisar o ensino da Probabilidade e Estatística, ou seja, da "Estocástica".

Um desses grupos se encontra na Espanha, na Faculdade de Ciências e da Educação na Universidade de Granada, com o qual se pode entrar em contato através da Internet pelo site **[batanero@urg.es.http://www.urg.es/local/ batanero/**]. Como um dos seus membros efetivos, a Dra Carmem Batanero Bernabeu, já esteve no Brasil ministrando palestras na Universidade de São Paulo (USP) e cursos na Universidade de Campinas (UNICAMP) em setembro de 1999.

No curso ministrado na UNICAMP, foram desenvolvidos alguns assuntos como: importância do estudo epistemológico, idéias estocásticas fundamentais, significados da probabilidade e aleatoriedade no Cálculo de Probabilidades, análise de uma investigação sobre o desenvolvimento cognitivo, análise exploratória de dados, análise de uma investigação sobre a compreensão da associação estatística, o ensino da Probabilidade na universidade, e outros.

O grupo de Batanero destaca a necessidade de se refletir sobre os principais objetivos do ensino da estocástica que para eles são:

- ***"...que os alunos cheguem a compreender e apreciar o papel da probabilidade e da estatística na sociedade, incluindo campos de aplicação e o modo como a estatística contribui para o seu desenvolvimento*;**
- ***que os alunos cheguem a compreender e a valorizar o método estatístico, isto é, a classe de perguntas que o uso inteligente da estatística pode responder, as formas básicas de raciocínio probabilístico e estatístico, seu potencial e limitações".*** (Batanero, 1999: 50)

Um outro grupo que destacaremos, encontra-se no Brasil, na Faculdade de Educação da Universidade de Campinas. O Círculo de Estudos Memória e Pesquisa em Educação Matemática (CEMPEM), dessa Faculdade, é responsável pelo ensino de pós-graduaçao para estudantes, principalmente de Matemática e Pedagogia.

Encontra-se no CEMPEM um subgrupo que trabalha com a Estatística, a Probabilidade e a Análise Combinatória. Destacamos alguns dos seus membros: Professores Anna Regina Lanner de Moura, Dario Fiorentine, Dione Luchessi de Carvalho da Faculdade de Educação, e João Frederico Meyer e Vera L. Figueiredo do Instituto de Estatística e Ciência da Computação.

Desde a formação desse grupo, um extensivo trabalho de organização tem sido devotado para uma revisão bibliográfica nacional e internacional nessa área. Mostra também interesse em tópicos, tais como: afinidades entre crença de estudantes e professores, concepções e representações sociais, e novas alternativas metodológicas.

CAPÍTULO II: BREVE HISTÓRICO DA PROBABILIDADE CONDICIONAL

Neste capítulo destacaremos alguns relatos históricos não só da Probabilidade como da Probabilidade Condicional, para analisarmos a evolução de tal conceito ao longo de três séculos, mas também verificar quais as dificuldades encontradas pelos matemáticos durante a sua criação e o seu desenvolvimento, nesse período.

Temos por objetivo destacar quais registros eram utilizados para definir conceitos ligados a Probabilidade Condicional e quais problemas alguns matemáticos se propunham a resolver.

Consideraremos também que alguns desses matemáticos apresentaram o conceito da Probabilidade, mostrando-o, a princípio, de uma maneira um tanto que empírica, baseando-se somente nas observações feitas sobre um evento. A seguir destacaremos alguns deles que tentaram sistematizar a Probabilidade Condicional, até receber a influência do formalismo matemático deste último século.

Podemos ter alguma dificuldade ao tentarmos descobrir historicamente a origem da Probabilidade. Sabemos que ela começa principalmente como uma ciência somente empírica e, com o passar dos anos, vão aparecendo várias concepções de probabilidade, como a clássica, a freqüentista ou empírica, a subjetiva e a formal, sendo que estas concepções não são necessariamente antagônicas, mas que em muitas situações podem se completar.

A origem reconhecida da noção de probabilidade como medida do grau de incerteza de um evento aleatório se deve à correspondência entre os dois matemáticos franceses Blaise Pascal e Pierre Fermat (1654), sobre a resolução de um problema célebre de divisão das partes que cabiam aos jogadores de um determinado jogo de azar. Apesar dessa origem reconhecida foram encontrados algumas considerações sobre as probabilidades feitas pelo matemático italiano Jérome Cardan (1547 – De Ludo Alea).

Anos depois, Jacques Bernouilli (1654–1705), em sua obra "L'Ars Conjectandi", reforça a concepção pascalina da eqüiprobabilidade dos eventos elementares. Para ele, a probabilidade dependia apenas do número de observações feitas sobre o evento estudado.

A primeira obra de que se tem conhecimento, que trata da Probabilidade Condicional, é "A doutrina das Chances" (1738) de Abraham De Moivre (1667-1754). Dela destacamos:

- definição clássica de probabilidade discreta e finita;
- noção de dependência e independência de eventos; e
- regra de cálculo da probabilidade conjunta.

Na rega de cálculo da probabilidade conjunta, com mais de dois eventos dependentes, já era praticada de maneira implícita a noção atual de probabilidade condicional, que podemos observar por um trecho dessa obra, transcrita por Totohasina :

> **"... a probabilidade do acontecimento de dois eventos dependentes, é o produto da Probabilidade do acontecimento de um deles pela Probabilidade de o outro ter acontecido, quando é considerado que o primeiro tenha ocorrido, e o mesmo se estenderá para o acontecimento de muitos eventos..."** *(Totohasina, 1992: 68)*

Para De Moivre:

> **"Dois eventos são independentes, quando eles não têm influência um sobre o outro, a probabilidade da realização de um não altera a realização do outro."** *(Totohasina, 1992: 68)*

A definição de dependência é dada pela negação da independência, mas o autor não escreve sobre que forma de influência pode ocorrer de um evento sobre o outro.

Thomas Bayes (1702-1761) em sua obra ”Tentativa no intuito de resolver um problema da Doutrina das Chances”, aprofunda a obra de seu mestre De Moivre. Nela contém sete definições, dez proposições e três regras, sendo que as definições são mais precisas e os teoremas são demonstrados de uma maneira mais rigorosa.

Retiraremos do trabalho de Totohasina trechos da obra de Bayes, como o da proposição três, que se refere à Probabilidade Condicional e diz:

> **“A probabilidade que dois eventos subsequentes se produzam todos os dois é em razão composta da probabilidade da primeira e da probabilidade da segunda, supondo que a primeira aconteceu”** *(Totohasina, 1992: 69)*

Dessa proposição decorre o seguinte corolário:

> **“...se de dois eventos subsequentes, a probabilidade do primeiro é a/N e que a probabilidade dos dois juntos é P/N, então a probabilidade do segundo (à condição que o primeiro aconteça) é P/a .”** *(Totohasina, 1992: 69)*

Observamos que tanto a proposição quanto o corolário lembra a definição de Probabilidade Condicional que temos hoje. Bayes, entretanto, não a relaciona com a noção de tempo nem de causa. Segundo Totohasina (1992,70):

> ***“A noção de “causa” aparece implicitamente em sua obra, mas não faz nenhuma alusão a ela”.***

Segundo Michel Henry, Bayes

> ***“ ordena a eqüiprobabilidade a priori dos valores possíveis para uma probabilidade desconhecida, parte para reajustá-la a posteriori, introduzindo assim duas noções de probabilidade: a primeira, a que procuramos estimar, é objetiva, a segunda, apreciando o valor possível da precedente, colocada a priori, é subjetiva.”*** (Henry, 1994:79)

Essa classificação considerada “subjetiva”, se deve ao fato de os métodos baysianos terem sua origem na idéia de atribuir uma probabilidade às causas de um evento observado a partir de um valor tomado a priori e recalculando-o em função dessa observação.

Esse valor, a priori da probabilidade subjetiva, pode ser muitas vezes determinado pela observação de várias ocorrências do objeto de estudo, mas também pode ser calculado baseado-se em determinados estudos de natureza, científica, como por exemplo: para se determinar se amanhã irá ou não chover, os meteorologistas não irão recorrer à observação do número de ocorrências de chuva desse dia ao longo dos anos, mas irão procurar tal Probabilidade dentro de técnicas mais sofisticadas que requerem uma compreensão específica de análises de fenômenos atmosféricos. Portanto, para se calcular esse valor a priori, muitas vezes deve-se tomar como base alguns estudos científicos, dependendo da área em que se quer determiná-lo.

Segundo Coutinho, a concepção subjetiva difere da concepção de probabilidade colocada por Jacques Bernoulli:

> ***“ ... dita “objetiva” , uma vez que dependia apenas do número de observações feitas sobre o evento estudado.”*** (Coutinho, 1994: 18)

Assinalamos que Bayes, através de sua “teoria”, foi um dos grandes pioneiros da Estatística. Sua obra deu origem a uma nova teoria que é a Estatística bayesiana, que não parou de evoluir. Podemos citar como exemplo dessa evolução a obra de B. Lecoutre (1984) “L’Analyse Bayesiene des Comparaisons”.

Jean Le Rond D’Alembert (1717-1783), motivado talvez pelas grandes navegações da época, escreve um artigo “Probabilidades”, na “Enciclopédia”, dando um exemplo de problema que envolve a probabilidade composta. Esse exemplo foi encontrado por nós em um Capítulo de Michel Henry (1997: 96), e vamos transcrevê-lo a seguir:

> ***“Um amigo partiu para as Índias numa esquadra de 12 barcos, dos quais sucumbiram três, e que um terço da tripulação dos barcos salvos morreu na viagem; a probabilidade que meu amigo esteja num dos barcos que chegaram bem ao porto, é 9/12, e que não seja um do terço da tripulação morto na rota, é 2/3. A “ probabilidade composta” que ele esteja ainda vivo, será entretanto 2/3 de 9/12 ou 6/12 ou uma metade de certeza. É entretanto para mim, estar entre a vida e a morte.”***

Percebemos nesse problema que D’Alembert utiliza o princípio da multiplicação das probabilidades, como faríamos com as proporções, sem desenvolver a noção de probabilidade condicional, nem de se preocupar com a independência dos eventos.

Pierre Simon Laplace (1749 –1827), um grande matemático de sua época, mostrou muito cedo o interesse pelo Cálculo de Probabilidade. Sua primeira obra publicada em 1774, “Memória sobre a probabilidade das causas pelos eventos”, tentava esclarecer as possibilidades, os limites e o cálculo das probabilidades no seu domínio, também dando uma interpretação “causalista” à teoria da probabilidade.

Na sua obra “Ensaio Filosófico sobre as probabilidades” (1825), juntando as definições e propriedades, enuncia dez princípios, que podem ser considerados uma primeira axiomática do cálculo de probabilidades. Vamos enunciar os seus sexto e sétimo princípios, que estão diretamente ligados às probabilidades condicionais, para, a seguir, analisar um exemplo de suas aplicações:

> ***“Sexto Princípio***
>
> ***Cada uma das causas, às quais um evento observado pode ser atribuído, é indicado como tanto mais de verossimilhança, que é mais provável que esta causa sendo suposta existir, o evento acontecerá; a probabilidade de existência de uma qualquer dessas causas é portanto uma fração cujo numerador é a probabilidade do evento, resultante desta causa, e cujo denominador é a soma das probabilidades semelhantes relativas a todas as causas: se estas diversas causas consideradas a priori são inegavelmente prováveis, é preciso no lugar da probabilidade, de um evento, resultante de cada causa, empregar o produto desta probabilidade, pela possibilidade da causa nela mesmo.”*** (Laplace, 1825 :14)

Laplace distingue nesse sexto princípio duas noções que são de “probabilidade de um evento” e de “possibilidade de uma causa”, estimada a priori, sem que esta causa seja claramente seguida de uma experiência aleatória. Parece que nessa interpretação de causa, lhe é difícil conceber a probabilidade do evento conjunto (interseção dos eventos). Neste sexto princípio percebemos claramente o Teorema de Bayes que conhecemos hoje, porém como essa distinção das noções de probabilidade de um evento e possibilidade de uma causa, parecem ser formas de cálculos distintas, dependo de uma análise bem subjetiva de quando devemos usar somente uma ou ambas.

Não sabemos porque, mas Laplace primeiro enuncia o que chamamos hoje de Teorema de Bayes, para depois exibir no seu sétimo princípio, o que conhecemos como Teorema das Probabilidades Totais, que colocaremos a seguir:

> ***"Sétimo Princípio***
> ***A probabilidade de um evento futuro é a soma dos produtos da probabilidade de cada causa, tirado o evento observado, pela probabilidade que, esta causa existindo, o evento futuro acontecerá."*** (Laplace, 1825:16)

Com esse sétimo princípio e o sexto juntos reconhecemos o teorema de Bayes. Para melhor analisarmos o uso de Laplace a esses dois princípios enunciaremos um exemplo dado por ele:

> ***"Imaginemos uma urna que não guarda mais que duas bolas, que sejam ou brancas ou pretas. Extraímos uma dessas bolas, que colocamos em seguida na urna, para proceder uma nova tiragem. Suponhamos que nas duas primeiras tiragens pegamos as bolas brancas; pergunta-se a probabilidade de tirar ainda uma bola branca na terceira tiragem."*** (Laplace, 1825 :17)

Vamos relatar como Laplace resolveu o problema:

> ***"Não se pode fazer aqui se não duas hipóteses, ou "uma das bolas é branca e a outra é negra", ou "todas as duas são brancas".***
> ***Na primeira hipótese, a probabilidade do evento observado é 1/4, ela é a unidade ou a certeza na segunda.***
> ***Assim, olhando estas hipóteses como de causas, nós teremos pelo sexto princípio, 1/5 e 4/5 para as suas probabilidades respectivas.***

> ***Ora, se a primeira hipótese acontece, a probabilidade de extrair uma bola branca na terceira tiragem é 1/2; ela é igual a unidade na segunda hipótese: multiplicando portanto estas últimas probabilidades, pelas hipóteses correspondentes, a soma dos produtos, ou 9/10, será a probabilidade de extrair uma bola branca na terceira tiragem."*** (Laplace, 1825:17)

Neste caso Laplace faz implicitamente as probabilidades das hipóteses da urna ter "uma bola preta e uma branca" e "as duas brancas", serem iguais, partindo do princípio que elas são eqüiprováveis. Mas sabemos que ao supormos eqüiprobabilidade podemos recair em situações, que a probabilidade final encontrada não tem sentido algum.

Da obra de Henri Poincaré (1854–1912), "Cálculos das Probabilidades", destacaremos algumas linhas sobre probabilidades compostas:

> **"*O Cálculo das probabilidades repousa sobre dois teoremas: o teorema das probabilidades totais e o teorema das probabilidades compostas.***
> ***Ao objeto de dois eventos A e B, pode-se colocar diversos problemas de probabilidades, considerando que um desses eventos deve se produzir, ou todos os dois, ou nenhum.***
>
> - ***A soma das duas probabilidades, para que A se produza e para que B se produza é igual à soma das probabilidades para que um dos dois ao menos se produzam e para que todos os dois se produzam:***
>
> ***(A) + (B) = (A ou B) + (A e B)***
>
> - ***A probabilidade para que A e B se produzam todos os dois é igual a probabilidade para que B se produza, multiplicado pela probabilidade para que A se produza, quando sabemos que B se produziu.***
>
> ***(A e B) = (B) . (A se B)*** (Totohasina, 1992: 73,74)

Podemos observar na última igualdade, o uso da mesma notação utilizada nos conjuntos para identificar os eventos (letras maiúsculas do nosso alfabeto), porém Poincaré não utiliza letras para indicar as probabilidades como por exemplo: p(B) ou P(B). Nem utiliza os símbolos de $\cap$ (interseção) ou $\cup$ (união) dos eventos, colocando (A e B) ao invés de ($A\cap B$) e (A ou B) no lugar de ($A\cup B$), como fazemos hoje . Mas pode-se dizer que sua simbologia está bem próxima da notação atual. Parece-nos que essa expressão (A se B) já se referia ao conceito de "probabilidade condicional", mas Poincaré, não fazia explicitamente referência à **Probabilidade Condicional.**

Émile Borel (1871-1956), deu início realmente a um princípio axiomático. Sua preocupação era juntar em um tratado os resultados essenciais para o Cálculo das Probabilidades e de sua aplicações. Parecia consciente das dificuldades desse quadro. Mostrava preocupação com o formalismo e com o emprego das noções literais, que podemos observar com um trecho de sua obra "Tratado de Cálculo de Probabilidades e suas Aplicações".

B) Nessa obra, o autor destaca o Axioma geral do Princípio da Adição das Probabilidades, também chamado princípio das probabilidades totais, considerando que se os eventos E_1, E_2 , ... E_n , de probabilidades p_1, p_2, ..., p_n, se excluem mutuamente, então a probabilidade para que eles se produzam é $p_1 + p_2 + +...+ p_n$. Isto é, se E_1 , E_2 , ... E_n, são eventos mutuamente exclusivos, dizemos que a ocorrência de E_1 exclui a ocorrência dos demais e assim sucessivamente, portanto $E_1\cap E_2 \cap E_3 \cap ... \cap E_n = \&$: nesse caso, a probabilidade de $E_1\cup E_2\cup ... \cup E_n$ é $p_1 + p_2 + ...+ p_n$.

> ***A definição da probabilidade se completa pela propriedade que exprime um segundo princípio fundamental, o princípio da probabilidade composta:***
> ***Se p_1 é a probabilidade de um evento E_1 ,e p_2 é a probabilidade de um evento E_2 quando E_1, é produzido, a probabilidade para que se produza a sucessão E_1 E_2 , é $p_1 . p_2$.***

A aplicação deste princípio é de uma prática simples a partir de eventos que sejam independentes entre eles, sua ordem não intervém mais, e p_i é então, simplesmente, a probabilidade de E_i , i=1, ou 2.” (Totohasina, 1992: 73)

Borel indica aqui o cálculo da probabilidade da interseção de dois eventos, isto é, $P(E_1 \cap E_2) = p_1.p_2$, analisando E_1 e E_2 como se fossem eventos sucessivos, considerando, portanto, $P(E_1 \cap E_2) = P(E_1).P(E_2/E_1)$.

E se os eventos E_1 e E_2 são independentes a $P(E_1 \cap E_2) = P(E_1).P(E_2/E_1) = P(E_1).P(E_2)$ (se desejarmos mais detalhes sobre essa teoria axiomática de Borel, numa versão mais atual, podemos encontrá-la em [2] p 1-24).

Borel deixa claro sobre a utilização desses princípios em relação ao cálculo das probabilidades da interseção dos eventos, quando estes são independentes. Parece-nos que está ligando-os somente à noção de tempo, e não faz menção a outra condição entre eles.

Somente em 1933, com a teoria de Andrei Kolmogorov (1903 – 1987), é que se apresenta uma axiomática à teoria das Probabilidades, colocando-a no quadro da Teoria dos Conjuntos e tornando mais claras suas limitações.

Em 1956, Alfred Renyi demonstra na sua teoria de espaço de probabilidade condicional, a generalização da axiomática de Kolmogorov, que destacaremos a seguir:

“Se $P^*(A)$ é uma medida definida sobre Ω e $P^*(\Omega) = 1$ (i.e $(\Omega, \mathcal{A}, P^*)$ é um espaço de probabilidade de Kolmogorov, mais ainda se $\mathcal{B}^*$ designa a família de conjuntos $B \in \mathcal{A}$ tal que $P^*(B) > 0$, então o sistema $(\Omega, \mathcal{A}, \mathcal{B}^*, P^*(./B))$ é um espaço condicional de probabilidade, onde $P^*(./B)$ é definido por

$$P^*(A/B) = \frac{P^*(A \cap B)}{P^*(B)}$$

com $A \in \mathcal{A}$, $B \in \mathcal{B}^*$.

$(\Omega, \mathcal{A}, \mathcal{B}^*, P^*(./B))$ será chamado o espaço condicional de probabilidade de Kolmogorov $(\Omega, \mathcal{A}, P^*)$.” (Totohosina,1995:55)

Para a compreensão dessa definição devemos entender o espaço de probabilidade de Kolmogorov, que é associado à terna (Ω,A,P), onde:

- Ω é o espaço amostral associado a um experimento aleatório,
- A é uma σ-álgebra de subconjuntos de Ω também conhecida como álgebra de Borel (podemos encontrar em [1] e [2] dos livros didáticos da Bibliografia), e
- P é uma probabilidade em A.

Essa definição dada por Renyi sobre o espaço de probabilidade condicional, é uma ampliação da definição dada por Kolmogorov. Tanto a definição dada por Renyi como a de Kolmogorov, não constam na maioria dos livros didáticos utilizados nas Universidades brasileiras.

R. Falk (1980) é de opinião que essa teoria axiomática de Kolmogorov não deva ser ensinada no Ensino Médio, pois neste podemos nos restringir às variáveis discretas não sendo necessário o conceito de σ-álgebra, que envolve considerações sobre a Teoria dos Conjuntos.

Analisando a evolução histórica de De Moivre (1773) à Renyi (1956) do conceito da Probabilidade Condicional podemos destacar fatos importantes:

- ♦ Dificuldade em analisar o caráter subjetivo da Probabilidade Condicional;

- ♦ Dificuldade em definir a dependência e a independência dos eventos compostos;

- ♦ A ausência de uma notação adequada para expressar a Probabilidade Condicional;

- ♦ Dificuldade em expressar a interseção de dois ou mais eventos; e

- Dificuldade provocada pela falta de suporte matemático adequado nos estudos anteriores a Kolmogorov.

CAPÍTULO III: DA PROBLEMÁTICA AOS PROCEDIMENTOS METODOLÓGICOS

3.1 A PROBLEMÁTICA

Nossa prática docente tem nos mostrado que muitos alunos apresentam grande dificuldade na manipulação dos conceitos ligados à Probabilidade. Dentre os pontos que nos chamaram mais a atenção, pelo seu número de ocorrências, estão os questionamentos e as dificuldades que os alunos apresentam diante de situações que envolvem o conceito de Probabilidade Condicional.

Diante desse fato decidimos fazer uma investigação bibliográfica, com o intuito de conseguirmos descobrir se nossas observações eram compartilhadas com outros professores e/ou pesquisadores. Encontramos, no Brasil, a dissertação de Coutinho (1994), defendida na PUCSP. O objetivo da autora era investigar como se dava a aquisição dos primeiros conceitos de Probabilidade utilizando a visão freqüentista. Através de um questionário, buscou identificar as concepções sobre as probabilidades e, tomando como base os resultados apresentados , elaborou uma seqüência de ensino, ambos aplicados a alunos do primeiro ano do curso de Fonoaudiologia no Brasil e a alunos do Ensino Médio na França. Detectou, dentre outras coisas, que:

- existem concepções, como por exemplo, a crença da eqüiprobabilidade, que dificultam a aprendizagem desejada, devido à ausência de informações sobre o evento a ser observado; e

- concepções errôneas podem persistir mesmo após o aprendizado de noções básicas deste conteúdo (Conceito de Probabilidade através da visão freqüentista).

Constatamos que, em alguns países, há um número grande de pesquisas publicadas na área de probabilidade e algumas, em particular, tratando de situações que envolvem a Probabilidade Condicional. Destacaremos algumas salientando pontos que nos interessam mais diretamente.

David R. Green (1982) em sua tese de doutoramento “Conceitos de Probabilidade na escola com alunos em idade de 11 a 16 anos”- Universidade de Loughborough – Inglaterra, relata a aplicação de um teste a 3000 alunos para determinar seus níveis de desenvolvimento Piagetiano e descobrir o que eles sabiam sobre conceitos de probabilidade e a sua linguagem de incerteza. Green classificou os alunos em quatro níveis de rendimento. Ressaltamos algumas de suas conclusões:

- ♦ muitos estudantes não alcançaram ainda o estágio de operação formal por volta dos dezesseis anos;
- ♦ diagramas de árvores e o princípio multiplicativo eram pouco entendidos pelos alunos; e
- ♦ o conceito de aleatoriedade era quase totalmente ausente nesses estudantes.

O trabalho de Fischbein e Lecoutre (1998) publicado na revista “Recherche en Didactique des Mathemaques”, na França, descreve a aplicação de um teste a estudantes de 10 a 18 anos de idade, com o intuito de analisar comparativamente quanto às idades, como os estudantes colocam seu conhecimento nos diferentes modelos cognitivos que são ativados nas situações que envolvem probabilidade, e também fazer alguns estudos das justificativas apresentadas pelos alunos.

Uma das conclusões que podemos destacar neste trabalho está ligada ao conceito de probabilidade condicional, que Fischebein chama de “fenômeno Falk” (sobre o qual trataremos na página seguinte), que se manifesta nos alunos quando um evento condicionante acontece depois do condicionado, constatando que os estudantes apresentam uma quantidade grande de erros em questões dessa natureza.

Totohasina (1992), procura diagnosticar as dificuldades e os obstáculos encontrados pelos alunos que iniciam o estudo do conceito de Probabilidade Condicional, com o intuito de tentar compreender porque aparecem tais dificuldades. Dentre suas conclusões ressaltamos a constatação de dificuldades encontradas pelos estudantes, que confundem a probabilidade condicional com a probabilidade conjunta ($P(A\cap B)$), e de obstáculos (considerados por ele epistemológicos) que os alunos apresentam ao lidarem com probabilidades condicionais quando se inverte o que é condicionante com o que é condicionado nas noções de probabilidade condicional ligadas às situações cronológicas e causais.

Evidenciamos o importante trabalho de Shaughnessy (1992), que fez uma pesquisa bibliográfica tentando coletar as publicações existentes sobre o ensino da Estatística, para a qual contou com colaboradores de alguns países do mundo. A partir dessa pesquisa, o autor foi relatando o trabalho de vários pesquisadores, ressaltando as contribuições de cada um e indicando alguns campos de pesquisa aos quais podemos nos dedicar nessa área.

Dentre esses pesquisadores, destacamos Falk (1988), por este ter desenvolvido trabalhos que envolvem a Probabilidade Condicional e ser referenciado por vários outros estudiosos. Falk considera que situações que tratam da Probabilidade Condicional são difíceis para os alunos, pois estes muitas vezes não conseguem determinar o evento condicionado, confundem condicionalidade com casualidade, fazendo com que investigue $P(A/B)$ quando lhes é pedido $P(B/A)$.

A situaçao referida por Fischbein e Totohasina como “fenômeno FALK”, acontece quando se procura a probabilidade de um evento, quando este aconteceu no tempo, antes do evento que o condicionou. Falk sugere uma técnica pedagógica para ajudar a compreender essa situação, que consiste em: de uma urna que contenha duas bolas brancas e duas pretas retira-se uma bola e sem que ninguém a veja, coloca-se de lado, então retira-se a segunda bola. Foi mostrado aos alunos que a segunda bola é branca e então pergunta-se: qual a probabilidade de que a primeira bola (ainda oculta) seja também branca? Através dessa simulação física do problema, ajudaremos os estudantes a observarem que o segundo evento pode realmente ser um evento condicionado.

Dos trabalhos do grupo de Batanero, destacamos sua obra “Azar e Probabilidade” (1991) editada na Espanha, que mostra várias sugestões de situações didáticas[1] que o professor poderá utilizar em sua sala de aula. Dentre elas, podemos salientar atividades que mobilizam conceitos de Probabilidade Condicional, dependência de eventos, Teorema da Probabilidade Total e o Teorema de Bayes, valorizando a representação em **diagramas de árvore** e **tabelas de contingência**. Assim como Shaughnessy, Batanero dedica um capítulo do seu livro, para fazer referências a vários pesquisadores, seus trabalhos e suas conclusões sobre o ensino da Estocástica.

Levando em consideração as dificuldades que os pesquisadores levantaram acerca do ensino da Probabilidade Condicional, resolvemos fazer uma pré-experimentação com alunos da Universidade, no Brasil, para sabermos se eles apresentam algumas dessas dificuldades ao trabalharem com o conceito e a aplicação das Probabilidades Condicionais.

1 Situações Didáticas, segundo Carmen Batanero são documentos orientados aos alunos, em um estilo comunicativo direto, se apresentado como situações problemáticas.

Sabemos que através da situação do problema apresentado na nossa pré-experimentação, não daríamos conta de analisar todas as dificuldades apresentadas pelos pesquisadores, mas nos serviu de instrumento para percebermos algumas delas, como a dificuldade de encontrar o evento condicionado, confusão entre a probabilidade condicional e a conjunção (interseção de eventos), identificação do que se é dado e o que se está pedindo, e a falta de representação.

Diante desses resultados, levantamos a seguinte questão:

“Como introduzir o conceito de Probabilidade Condicional em cursos da Universidade, de maneira a minimizar essas dificuldades”?

Nossa preocupação não é só diagnosticar dificuldades como foi feito em alguns dos trabalhos citados, mas sim introduzir o referido conceito numa situação de sala de aula em cursos na Universidade, embora estejamos conscientes das dificuldades que possam se apresentar.

As pesquisas citadas que envolvem a Probabilidade Condicional aconteceram em países em que a realidade dos alunos é diferente da dos nossos. Como por exemplo, na França, os alunos têm contato com noções de probabilidade muito antes do aluno brasileiro.

Acreditamos que ajudaria em nosso trabalho se tivéssemos mais pesquisas sobre o perfil do aluno brasileiro em relação a conhecimentos probabilísticos, mas o que nos está interessando no momento, é estudar uma boa maneira de introduzir o conceito de Probabilidade Condicional para diminuir as dificuldades encontradas por nós e pelos pesquisadores.

Nossa **hipótese** é a seguinte: acreditamos que, a partir de sugestões dadas por Batanero, é possível desenvolvermos uma seqüência de ensino para que os alunos consigam construir o conceito de Probabilidade Condicional e trabalhar melhor com todos os conceitos que o cercam, utilizando vários registros de representação como a linguagem natural, o diagrama de árvore, a tabela de contingência e a linguagem simbólica.

Formulamos algumas questões que pretendemos responder com o nosso trabalho, que acreditamos sejam fundamentais não só para um bom entendimento do conteúdo por parte dos alunos, como para uma boa conclusão de nossa pesquisa. São estas:

- Os alunos, diante de um problema que envolva eventos, conseguem identificar os dependentes e os independentes?

- Diante de situações que envolvam condicional, será que eles diferenciarão da interseção de eventos?

- Os alunos conseguem fazer algumas representações ligadas à Probabilidade Condicional?

- Será que eles vão saber diferenciar o cálculo de P(A/B) de P(B/A)?

- Os alunos aplicarão o conceito da condicional para problemas que envolvam o Teorema das Probabilidades Totais e Teorema de Bayes?

3.2 FUNDAMENTAÇÃO TEÓRICA

Este trabalho foi desenvolvido com base em elementos da obra **“Azar e Probabilidade”** de Batanero, Godino e Cañizares, sobre Probabilidade e Estatística, assim como na **Teoria de Registro de Representação** de Raymond Duval e no capítulo **“Utilização das Árvores no Ensino de Probabilidades”** de Bernard Parzysz do livro “Ensinar as Probabilidades no Ensino Médio” [29].

Batanero, Godino e Cañizares nos fazem refletir, através de sua obra, sobre a importância da Probabilidade, lembrando que esta proporciona um modo de medir a incertude, em conseqüência, os modelos probabilísticos são o fundamento da maior parte da Teoria Estatística. Isto implica que é necessário o conhecimento da teoria da probabilidade para uma compreensão adequada dos métodos estatísticos, que são úteis e indispensáveis nos campos científicos, profissional e social.

Para se trabalhar com alguns desses modelos probabilísticos, os autores acreditam que:

> ***“... a probabilidade pode ser aplicada à realidade tão diretamente como a aritmética elementar, não sendo preciso o conhecimento de teorias físicas nem de técnicas matemáticas complicadas.”*** (Batanero, Godino e Cañizares, 1996: 12)

Os autores mostram modelos concretos que acreditam facilitar o tratamento deste tema. Apresentam uma coleção de unidades de aprendizagem de modo seqüencial, introduzindo as noções básicas sobre Probabilidade através de situações didáticas variadas. Para elaborar estas propostas, levaram em conta um estudo sobre fundamentos didáticos, revisão de algumas experiências realizadas em outros países, análises de pontos históricos e análises de alguns currículos do Ensino Fundamental e Médio.

As seqüências das unidades têm um crescimento de complexidade de acordo com os conceitos introduzidos, que vão desde a mera apresentação do caráter imprevisível do azar, à probabilidade de sucessos simples, probabilidades geométricas, probabilidade condicional, combinatória e concluindo com o Teorema de Bayes.

Dentre essas situações didáticas, encontramos algumas de experimentos compostos (mais de um evento), que estão diretamente ligadas ao tema do nosso trabalho, através das quais os alunos devem calcular a probabilidade do sucesso, utilizando a regra da multiplicação e se defrontando, em alguns casos, com o cálculo da Probabilidade Condicional P(A/B), tanto nos casos de sucessos dependentes como dos independentes, fazendo isso de um modo intuitivo, sem introduzir uma notação formal.

Dessa maneira intuitiva, através do uso de **tabelas de contingência** e de **diagramas de árvore,** os autores apresentam uma situação didática em que deverão conduzir o aluno a calcular diversas probabilidades fazendo uso dos Teoremas de Probabilidade Total e de Bayes, sem mesmo saber a fórmula explicitamente.

Batanero, Godino e Cañizares apresentam as situações, mas lembram que estas devem servir para auxiliar o trabalhos com tais temas, e sugerem que os professores façam a devida formalização matemática, levando em conta sempre o nível escolar do aluno com que se está trabalhando.

Diante das situações didáticas de Batanero, Godino e Cañizares, ficou bem marcado o uso das **tabelas de contingência** e de **diagramas de árvore** como forma de se trabalhar modelos probabilísticos. Isso nos fez tentar relacionar com a **Teoria de Registros de Representação Semiótica** de Raymond Duval.

Para Duval, não existe conhecimento matemático que possa ser mobilizado por um aluno sem que este utilize uma representação. Introduz a noção de registro para analisar a aquisição dos conhecimentos matemáticos, do lado dos sistemas produtores de representação e não do lado dos objetos. Ressalta que a atividade matemática mobiliza simultânea ou alternativamente vários registros de representação semiótica, alguns ligados ao funcionamento cognitivo comum como a língua materna, e outros criados pela necessidade do desenvolvimento da atividade matemática.

Podemos então considerar que tabelas de contingência sejam um registro de representação das probabilidades e os diagramas de árvore outro de um mesmo objeto. Segundo Duval, quanto maior for a flexibilidade entre diferentes registros de representação do mesmo objeto matemático, maior será a possibilidade de apreensão desse objeto.

Existem regras de tratamento próprias de cada registro de representação, como por exemplo, se estivermos trabalhando com a árvore de probabilidades como sendo uma representação da Probabilidade Condicional, teremos um tratamento interno referente somente a esse registro ou se optarmos pela tabela de contingência será outro tratamento. De acordo com Duval, o tratamento é interno ao registro, e a passagem de um registro a outro do mesmo objeto matemático é chamado de conversão.

No capítulo “Utilização de árvores no ensino das probabilidades” de Bernard Parzysy [29], o autor faz várias articulações envolvendo o uso das árvores de probabilidades e a Teoria de Registro de Representação de Duval. Através de um exemplo dado pelo próprio autor, faremos uma análise dessas articulações.

> ***“Num jogo de diversão, dispomos de três caixas aparentemente idênticas. Elas contêm respectivamente um, dois e três papéis, das quais somente um papel é marcado. Uma partida consiste em o jogador escolher ao acaso uma caixa e tirar, igualmente ao acaso, um papel dessa caixa.***

Calcular a probabilidade que o papel tirado seja o marcado.”

(Parzysz, 1997: 227)

Vamos analisar alguns procedimentos que devemos efetuar com o intuito de resolvermos a situação.

A experiência aleatória consiste em:

- escolher uma caixa ao acaso;
- tirar dessa caixa um papel ao acaso (lembrando que toda caixa tem um papel marcado)

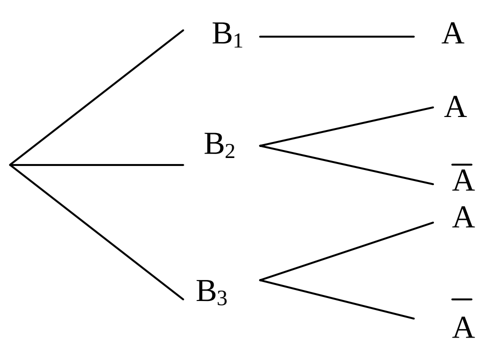
Uma estratégia possível de resolução pode consistir, por exemplo:

1º) Começamos a modelização: enumerar as caixas e definir os eventos.

A: tirar um papel marcado e

B_i (i = 1,2,3): escolher a caixa nº i.

2º) Visualizar a situação através de uma árvore de possibilidades, como na figura abaixo:



$\overline{A}$ designa o evento contrário de A

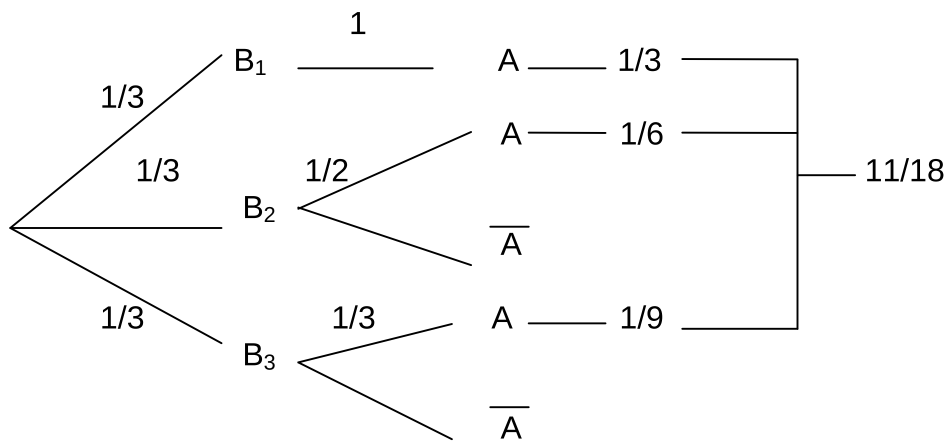
3º) Procurando a modelização, converter os dados do enunciado nos registros simbólicos, e fazer as hipóteses de eqüiprobabilidade sugeridas pela expressão “ao acaso” (também utilizaremos de forma mais ou menos implícita um certo número de “regras”, que podem ser regras de ação ou de teoremas).

Pode-se obter: $P(B_1) = P(B_2) = P(B_3) = 1/3$

$$P(A/B_1) = 1, \quad P(A/B_2) = 1/2 \text{ e } P(A/B_3) = 1/3$$

Colocamos, portanto, a probabilidade em um registro simbólico.

Vamos fazer uma conversão, isto é, passar do registro simbólico para a árvore de probabilidades e, a seguir, operar um tratamento nesse novo registro (ajudando nas fórmulas do tipo $P(A \cap B) = P(A/B).P(B)$ e Probabilidades Totais).



Portanto, a Probabilidade de que o papel tirado seja o marcado é de 11/18.

As dificuldades que os estudantes possam ter diante desta escolha de resolução, para o autor, podem recair no tratamento interno a um registro e na conversão de um registro a outro. Na tentativa de minimizar essas dificuldades, estabelece regras para tratamentos dentro do registro “árvore de probabilidades”, e regras também para a conversão entre os diferentes registros que se mobilizam para resolvermos problemas desta natureza.

Apresentaremos algumas regras de conversão e tratamento dadas por Parsysz.

Conversão:

- As letras colocadas nos “nós da árvore” representam os eventos e a raiz corresponde ao conjunto fundamental;

- O número colocado sobre o galho que liga um evento A ao B (A —— B) da árvore é a Probabilidade Condicional de B se ocorreu A; e

- A cada nó da árvore pode-se fazer corresponder a interseção dos eventos associados aos nós da seqüência que liga este nó à Ω, e a probabilidade dessa interseção é o número associado ao nó. Podemos visualizá-la, observando o local em que foram colocadas as probabilidades 1/3, 1/6 e 1/9 na árvore da página anterior. Nesta conversão não calculamos a probabilidade, e sim analisamos as correspondências entre os eventos colocados na árvore e as suas possíveis interseções.

Tratamento:

- A cada nó N da árvore associamos o produto dos números que figuram sobre os galhos que ligam N à raiz (como calcular as probabilidades da interseção dos eventos);
- A soma dos números associados a todos os galhos de mesma origem é igual a 1;
- A soma dos números associados aos nós de um nível dado é igual a 1;
- A soma de algumas probabilidades das interseções, como no exemplo dado nesta unidade, sugere o uso implícito do Teorema da Probabilidade Total.

Tomamos, a princípio, como base da nossa pesquisa as situações didáticas apresentadas por Batanero, Godino e Cañizares. Tais situações nos fizeram refletir que poderíamos apresentar as probabilidades aos alunos de diferentes maneiras através da linguagem natural, simbólica, diagrama de árvore e tabela de contingência levando-nos a associar com a Teoria de Registros de Representação de Duval, portanto essas diferentes maneiras poderíamos considerar como diferentes registros de representação. E finalmente o artigo de Parsysz nos mostrou um caminho de como articular alguns desses diferentes registros ligados à árvore de probabilidades.

3.3 METODOLOGIA DA PESQUISA

Nossa pesquisa teve como ponto de partida uma série de leituras sobre temas envolvendo Probabilidades e Estatística. Essas leituras englobaram: dissertações de Mestrado no Brasil, teses de Doutoramento de outros países, publicações de vários artigos sobre algumas pesquisas na área, livros didáticos nacionais e alguns outros livros não nacionais, alguns referentes ao objeto estocástico e outros como fonte de pesquisas.

Nossa pretensão inicial era trabalhar com a Probabilidade na introdução das variáveis contínuas em Estatística, tentando estabelecer uma relação mais "visível" para o aluno, entre a noção de área trabalhada no Cálculo Integral e a introdução do conceito da Probabilidade das variáveis contínuas através de área. Todavia, percebemos que seria um trabalho difícil, que exigiria uma correlação entre duas áreas do conhecimento matemático (Cálculo Integral e Estatística). Além disso, o âmbito de tal trabalho é bastante amplo e ultrapassaria os limites de uma dissertação de Mestrado, e por isso abandonamos temporariamente o possível tema.

Dentre os conceitos encontrados na bibliografia consultada, decidimos trabalhar com um dos mais iniciais da Probabilidade, o da Probabilidade Condicional, não somente pela falta de trabalhos brasileiros sobre o tema, como também pelas dificuldades na manipulação de tal assunto encontradas em nossa prática docente.

Resolvemos aplicar uma pré-experimentação a alunos de um curso de Publicidade, com o objetivo de saber como eles lidavam com a Probabilidade Condicional frente a um problema contextualizado, como o "Problema do Táxi", citado por pesquisadores como Batanero, Godino e Cañizares, Shaughnessy, Tversky e Kahneman, Stephen Tomlinson e Robert Quinn.

Por que a aplicação da Pré-Experimentação?

Além do motivo mencionado acima, nossa prática docente, até então, era com alunos que já haviam concluído o curso superior. Diante disso, procuramos saber como os alunos que acabavam de trabalhar com tal conteúdo, em um curso universitário brasileiro, reproduziam seus conhecimentos, pois sabíamos, através das leituras, quais eram as dificuldades que os alunos de outros países apresentavam, na sua maioria, no Ensino Médio.

De posse dos resultados das análises feitas dos protocolos dos alunos, do levantamento histórico, do objeto estocástico e de pesquisas realizadas, resolvemos elaborar uma seqüência didática na tentativa de contribuirmos para minimizar as dificuldades encontradas pelos estudantes. O objetivo é fazer com que os alunos trabalhem com o conceito de Probabilidade Condicional de maneira significativa, compreendendo este como um modo de resolver problemas ligados à incerteza.

A elaboração e a aplicação da seqüência, bem como a análise dos resultados baseiam-se nos princípios da Engenharia Didática como Metodologia de Pesquisa, que segundo Michèle Artigue, se caracteriza por um esquema experimental baseado nas realizações didáticas em sala de aula e trata das concepções, realizações, observações e análise de seqüência de ensino.

Este procedimento é composto de quatro fases:

- Primeira fase – **Análises preliminares**: levantamento das concepções envolvidas. Nessa fase, buscam-se os quadros teóricos orientadores do processo.

- Segunda fase – **Concepção e Análise a priori**: nessa fase o investigador decide por um determinado número de variáveis. São variáveis pertinentes ao problema estudado. Seu objetivo é determinar que seleções de variáveis melhor permitirá controlar o comportamento dos estudantes.

- Terceira fase – **Experimentação**: fase da realização da engenharia com uma certa população de alunos. Ela começa quando pesquisador, professor e observadores entram em contato com essa população de alunos, e é nessa fase que ocorre também a “Formalização” ou “Institucionalização” dos conceitos trabalhados na atividade aplicada.

- Quarta fase – **Análise Posteriori**: baseia-se num conjunto de dados recolhidos ao longo da experimentação, assim como nas observações realizadas durante a aplicação na seqüência de ensino.

Destacamos que a confrontação das análises a priori e a posteriori se fundamenta a essência da “validação” das hipóteses formuladas na investigação.

CAPÍTULO IV : PRÉ-EXPERIMENTAÇÃO

Com o objetivo de confirmar a existência de algumas dificuldades dos alunos ao trabalharem com o Cálculo de Probabilidades, resolvemos aplicar um teste a uma classe de 28 alunos que cursavam Estatística II do 2º ano de um Curso de Publicidade, em uma Universidade particular na cidade de São Paulo. No semestre anterior em Estatística I, os alunos tiveram contato com Probabilidades envolvendo o cálculo da Probabilidade da interseção de dois ou mais eventos, Probabilidade Condicional, Teorema do Produto, Teorema da Probabilidade Total e o Teorema de Bayes.

O professor de Estatística II, ao começar o seu curso, fez uma revisão dos conteúdos citados acima e, depois de duas semanas, aplicamos o nosso teste.

A situação que consta do teste foi traduzida e adaptada por nós para uma melhor compreensão por parte dos alunos, sendo retirada de "Pesquisa em Probabilidade e Estatística: Reflexões e Direções", por J. Michael Shaughnessy. O problema apresentado aos alunos (ANEXO I), consta de nove questões organizadas de (a) a (j), sendo que as três primeiras são somente para que eles identifiquem e reconheçam os dados apresentados.

Para melhor compreensão das análises, reproduzimos o problema a seguir:

Um táxi foi envolvido em uma batida, à noite, e fugiu do local do acidente.

Há duas companhias que operam na cidade, a companhia de Táxi Azul e a companhia de Táxi Verde. É sabido que 85% dos táxis da cidade são verdes e 15% são azuis. Uma testemunha da cena identificou o táxi envolvido no acidente como sendo da companhia de Táxi Azul. Essa testemunha foi posta à prova sob condições semelhantes de visibilidade, e fez a identificação correta da cor do carro, tanto do táxi azul como do verde, em 70% dos testes.

Questão a) Qual a probabilidade de a testemunha identificar corretamente a cor de um carro qualquer?

Questão b) Qual a probabilidade de se encontrar um táxi azul na cidade?

Questão c) Qual a probabilidade de se encontrar um táxi verde na cidade?

A resposta para a questão a) é imediata, está escrito no enunciado do problema que a identificação correta, pelo motorista, da cor de um carro é de 70%. O mesmo ocorre com as questões b) e c), suas respostas são imediatas, isto é, estão escritas no enunciado P(A) = 15% (Probabilidade de se encontrar um táxi azul nesta cidade) e P(V) = 85% (Probabilidade se de encontrar um táxi verde nesta cidade. Esperamos que os alunos sejam capazes de identificar estas três respostas no enunciado do problema.

Questão d) Você estabelece alguma relação entre o acerto da testemunha e a porcentagem de carros azuis ou os verdes da cidade ?
Comente sua resposta.

Com a questão d) queremos saber se os alunos estabelecem relação entre carros da cor verde ou azul que poderão estar envolvidos no acidente e a cor do carro que foi identificado pela testemunha.

Essa questão tem o objetivo de verificar se o aluno reconhece a dependência do evento “a testemunha identifica corretamente a cor do carro” e o evento “cor do carro”. Sabemos, por nossas leituras, que os alunos têm dificuldades em diferenciar os eventos dependentes dos independentes. Eles poderiam reconhecer esta dependência de uma maneira quase intuitiva, através da leitura do enunciado, identificando que a probabilidade do acerto do motorista está diretamente relacionada à probabilidade da quantidade de carros azuis ou verdes da cidade.

Questão e) Você sabe dizer qual a probabilidade de a testemunha identificar um táxi como sendo verde, sabendo-se que o táxi é verde?

Com a questão e) pretendemos verificar se o aluno percebe uma condicional. É possível que ele chegue à conclusão que, pelo fato da cor do carro ser dada, a resposta seja 70%, independentemente da cor escolhida.

Os alunos podem expressar tal resposta através de uma linguagem formal, utilizando o símbolo que representa a Probabilidade Condicional: P(IV/V) = 70%. Podem também, responder na linguagem natural que a Probabilidade da testemunha identificar um táxi como sendo verde, sabendo-se que o táxi é verde é de 70%.

Tal questão pode ser resolvida, utilizando-se uma ferramenta muito útil que é a "árvore de probabilidades", como vem esboçada a seguir:

70% IV $P(V \cap IV) = 0{,}595 =$
59,5%
85% V
30% IA $P(V \cap IA) = 0{,}255 =$
25,5%
30% IV $P(A \cap IV) = 0{,}045 =$
4,5%
15% A
70% IA $P(A \cap IA) = 0{,}105 =$
10,5%

Contudo, não precisamos de toda esta árvore para responder a questão, basta somente observar a Probabilidade que está sobre o "galho" V IV, pois esta representa a Probabilidade Condicional pedida na questão. As demais Probabilidades que figuram nesta árvore servirão como referência para responder algumas das questões seguintes.

Questão f) Você sabe dizer qual a probabilidade do táxi ser verde, sabendo-se que a testemunha o identificou como verde?

Esta questão foi colocada com o intuito de observarmos se o aluno consegue diferenciar a P(IV/V) da P(V/IV), que na linguagem natural significa diferenciar a Probabilidade de a testemunha identificar a cor do carro como sendo verde, sabendo-se que ela é verde, com a Probabilidade do carro ser verde, sabendo-se que a testemunha o identificou como verde; estabelecer esta diferença é uma dificuldade apresentada pelos alunos, apontada por alguns pesquisadores.

Nesta questão, observaremos, também, se o aluno percebe que “a testemunha identificar como verde” depende da cor do carro, isto é, se o aluno consegue descrever todas as possibilidades para que isso ocorra, e também todas as condicionais que envolvem a questão.

Os alunos poderão utilizar-se da árvore de probabilidades para resolverem esta questão, e da tabela de contingência ou ainda da fórmula do Teorema de Bayes seguida da Probabilidade Condicional e depois do Teorema da Probabilidade Total.

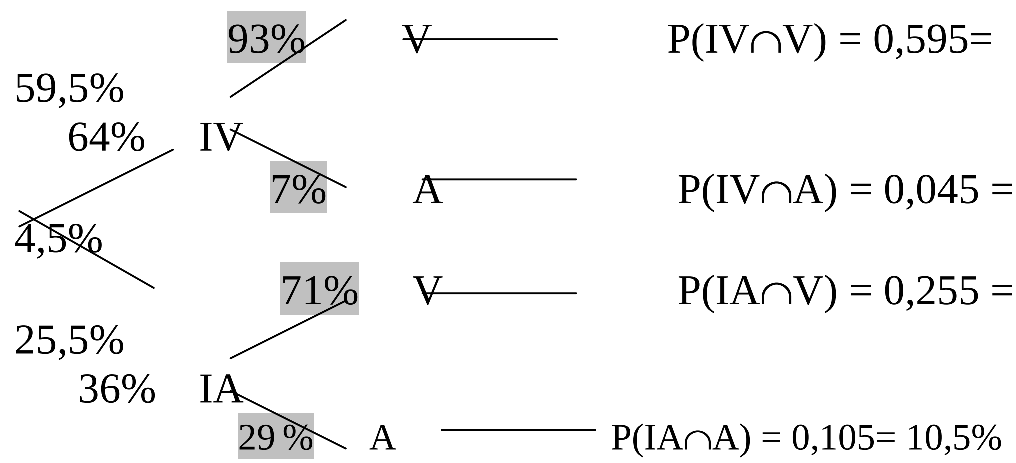
Para encontrarmos tal resposta, a princípio observaremos a árvore de probabilidade que construímos anteriormente, e percebemos, que ela não nos mostra resposta para essa situação, pois o que figura nos primeiros galhos é a cor do carro, e não a identificação feita pelo motorista, como queremos na questão. Portanto, precisamos de uma árvore que comece com a identificação do motorista, e isto faremos a seguir.

Tomando como base a nossa árvore já construída e mais uma tabela de contingência conseguiremos facilmente “montar” a nova árvore desejada. A tabela de contingência correspondente é a que figura abaixo:

Identificação / Cor do carro	IA	IV	Total
Azul	0,105	0,045	0,15

Verde	0,255	0,595	0,85
Total	0,36	0,64	1

Podemos observar na tabela todas as interseções e as Probabilidades Totais: a de o motorista identificar um carro como sendo azul é de 0,36 (P(IA)=0,36), e a de o motorista identificar um carro como sendo verde é de 0,64 (P(IV)=0,64). Este fato nos auxilia na construção de nossa nova árvore, começando pelo evento "identificação da cor", já calculado na tabela de contingência. Teremos, portanto:



Considerando que não temos as probabilidades condicionais dos segundos galhos, para encontrá-las basta dividirmos a Probabilidade das interseções pelas respectivas Probabilidades dos primeiros galhos, obtendo, assim, a probabilidade desejada, que assinalamos na árvore acima.

Resolvendo tal questão pela fórmula do Teorema de Bayes, teremos:

$$P(V / IV) = \frac{P(V \cap IV)}{P(IV / V).P(V) + P(IV / A).P(A)}$$

Lembrando que a $P(V \cap IV) = P(IV/V).P(V) = 0{,}7.0{,}85=0{,}595=59{,}5\%$.

E a P(IV/A) = 0,30, isto é, a probabilidade da testemunha identificar a cor do carro como sendo verde, quando sabemos que a cor é azul (probabilidade de erro da testemunha).

$$P(V / IV) = \frac{0{,}595}{0{,}595 + 0{,}3.0{,}15} \approx 0{,}93$$

Questão g) Qual a probabilidade de que o carro envolvido no acidente seja realmente o Táxi Azul e que a testemunha o identifique como azul?

A questão g) foi colocada logo depois de uma condicional, com o intuito de podermos analisar se o aluno a distingue de uma interseção de dois eventos. Com ela pretendemos explorar os conceitos que envolvem "interseção de eventos" e "condicional", que normalmente são confundidos.

Através do diagrama de árvore podemos obter esta resposta Basta observarmos a seqüência de A seguida de IA, ou vice-versa, e obteremos a resposta, isto é, P(A∩IA) = 10,5%

Questão h) Sabendo que a testemunha identificou o carro envolvido no acidente como sendo o azul. Qual a probabilidade de que o carro envolvido no acidente seja um Táxi Azul?

Questão i) Qual a probabilidade de que o carro envolvido no acidente foi realmente um Táxi Azul, sabendo-se que a testemunha o identificou como azul?

As questões h) e i) são iguais, porém apresentam simplesmente uma mudança de escrita dando primeiro a informação já ocorrida e depois perguntando qual a probabilidade e vice-versa. Elas foram colocadas com o objetivo de verificar se o aluno distingue o que é dado do que é pedido, isto é, se ele reconhece que foi mudada a ordem no enunciado mas não a pergunta da questão.

Da segunda árvore de probabilidades podemos retirar tal resposta. Basta observar a probabilidade que está em cima do galho IA ________ A, que é 29%. Pode-se recorrer para responder a questão a definição formal da Probabilidade Condicional e a seguir utilizar-se do Teorema da Probabilidade Total.

Questão j) Se você tivesse que decidir se realmente o táxi azul é o culpado. Que decisão tomaria? Comente sua resposta.

Na última questão queremos saber a opinião do aluno diante da situação apresentada e, além disso, saber em que se baseiam as suas escolhas.

Essas questões foram colocadas em uma determinada seqüência: em primeiro lugar para reconhecimento dos dados, a seguir estabelecer relações entre eles, interpretar as interseções e condicionais para calcular probabilidades e, finalmente, identificar de quais registros os alunos se utilizam para resolver problemas desta natureza. Todas as nossas questões se apresentam na linguagem natural.

ANÁLISE DAS RESPOSTAS

O teste foi aplicado a 28 alunos no dia 22 de março de 2000 no período matutino e teve a duração de uma hora, durante uma das aulas de Estatística II, para ser resolvido em dupla. A quantidade de erros, acertos e questões em branco estão na tabela a seguir. Foram colocadas na tabela as questões de a) a i). A questão j) é subjetiva e poderiam ser dadas várias interpretações na sua resposta.

Os alunos mostraram interesse em responder as questões.

QUESTÕES	ACERTOS	ERROS	EM BRANCO
A	13	01	-
B	14	-	-
C	14	-	-
D	07	06	01
E	08	06	-
F	-	13	01
G	08	06	-
H	-	12	02
I	-	13	01

Erros mais comuns apresentados pelos alunos.

Na questão d), as respostas apresentadas indicam que alguns dos erros cometidos decorrem do fato de os alunos não compreenderem quais eram os dados na questão, desprezando em alguns casos a identificação do motorista, ou por não acreditarem que é pertinente ou que 70% não representa ser uma porcentagem que garanta totalidade no acerto.

Dentre as respostas que consideramos erradas, vamos destacar algumas justificativas dadas pelos alunos:

- “Não, pois a porcentagem de carros não influi na visão da pessoa, senão seríamos todos cegos”.
- “Não estabelecemos nenhuma relação. O motorista pode ter “chutado” a resposta”.
- “Não, pois ele identificou corretamente só em 70% dos testes”.

Na questão e), das seis duplas que erraram, pudemos verificar que alguns calcularam a probabilidade da interseção dos eventos “Identificar a cor do carro” e o “táxi é verde”, portanto escreveram:

“P = 70.85 =59,5%”

Outras duplas consideraram pelo enunciado da questão que tudo o que se pedia era verde, adotando como resposta:

“ 85%, tudo que está sendo pedido é verde”.

Estas respostas indicam que essas duplas apresentam dificuldade de interpretação do enunciado.

Independente dos acertos ou erros, a maioria dos alunos responderam suas questões de forma bem lacônica, isto é, colocaram somente a porcentagem como resposta, não indicando o que significava tal resposta nem justificaram como a obtiveram.

Na questão f), das 13 duplas que responderam:

- 08 calcularam a probabilidade da interseção dos eventos, como se eles fossem independentes e não perceberam a condicional;
- 04 persistiram na idéia de que continuava havendo 70% de chance, prevalecendo o que pensava a testemunha, portanto também não reconhecendo a condicional; e
- 01 fez um cálculo que não tinha relação com a questão.

Na questão g),

- 04 alunos continuaram com a idéia de 70% do acerto da testemunha, desprezando a porcentagem de táxis azuis da cidade; e
- 01 aluno colocou as duas considerações como resposta, isto é, “15% de azuis e 70% do acerto da testemunha”.

Na questão h) e i),

- 05 alunos identificaram que as questões eram iguais;
- 02 desses interpretaram como sendo o cálculo da probabilidade da interseção dos eventos independentes “ser o táxi azul” e “a testemunha identificá-lo azul”; e
- os demais deram várias outras interpretações para o cálculo das probabilidades.

Observamos que na questão h), 06 duplas responderam como se o cálculo da probabilidade fosse uma interseção dos dois eventos novamente independentes; as outras duplas fizeram vários cálculos sem sentido ou só se baseando em um dos dados e desprezando o outro. Na questão i), apenas 04 alunos apresentaram o procedimento do cálculo da interseção dos eventos independentes e alguns só consideraram um dos dados.

Para responder a última questão, 07 dos alunos basearam-se somente nos 70% de probabilidade da testemunha acertar ser maior que 30% dele errar, portanto indicaram o carro azul baseado neste fato. Uma dupla baseou a sua resposta na probabilidade de interseção entre o "carro ser azul" e "a testemunha o identificar azul", ou o "táxi ser verde" e "a testemunha o identificar azul", todos os dois casos sendo tratados como se fossem eventos independentes, portanto "culparia" o táxi verde pelo fato de a segunda probabilidade ser maior. Alguns alunos responderam o que pensavam independente dos dados do problema.

Os alunos representaram suas respostas totalmente desprovidos de um formalismo matemático, não usaram "diagrama de árvore", ou qualquer outra representação sugerida anteriormente, dificultando muitas vezes nossas análises, por não sabermos como chegaram em suas respostas. Colocavam simplesmente as multiplicações com os fatores, sinal de igualdade e o produto. Mas o que eles estavam calculando, não se sabia.

Conclusões:

- Como era o esperado, não foi reconhecido pelos alunos nenhum caso de condicional. Em todas as questões que envolviam tais conhecimentos, observamos totalidade de erros.
- Foi unânime a presença do cálculo da probabilidade de interseção de eventos independentes no lugar do cálculo da condicional.
- Não foi considerada a informação dada como condicionante.

- ♦ Constatação da falta de formalismo matemático para indicar as operações e resultados dos problemas.
- ♦ Os alunos não mostraram saber representar situações que envolvessem probabilidade das interseções ou probabilidades condicionais, como diagramas de árvore, tabelas de contingência e tabelas das condicionais. Parece-nos que eles querem uma fórmula para resolverem o problema.
- ♦ Obtivemos total número de acertos pelos alunos nas três primeiras questões, o que era esperado, portanto eles identificam os dados explícitos.

ENTREVISTA COM UMA DAS DUPLAS

Durante a aplicação do teste foi observada uma dupla que discutia muito, e parecia não estar satisfeita com a sua atuação no teste.

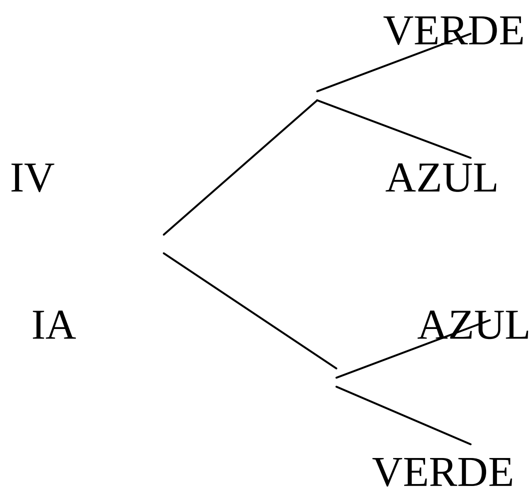
Como a professora deles acompanhou a aplicação, os dois alunos começaram a fazer várias perguntas a ela.

Resolvemos então voltar às perguntas do mesmo jeito que elas vinham, como por exemplo:

- professora, não é só fazer a multiplicação aqui na questão f)? Falta alguma coisa?
- Você acha que falta?
- Acho.
- O que você acha que falta?
- Será que a testemunha pode identificar o carro sendo verde e ele ser azul?
- O que você acha?
- Acho que sim.
- Mas o que eu faço com isso e com o fato de a testemunha identificar azul e ele ser azul? Somo? Multiplico?
- Você conhece a árvore de probabilidades?

- Conheço, só que tem muitos galhos, e eu me perco, mas quando faço, sempre acerto o problema.

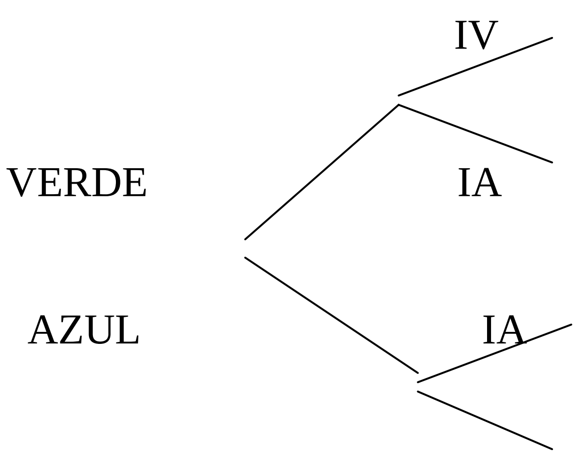
O aluno indica esta árvore;



(Parece que neste caso fica mais difícil para o aluno pensar na árvore, porque ele não reconheceu que a identificação da cor do carro é dependente da quantidade de carros de cada cor).

- Você sabe representar esta árvore de outra maneira?

- Sei, é só inverter a ordem das informações, e representou:



- Assim fica mais fácil, diz o aluno.

- E daí para frente ele consegue enxergar que existem mais possibilidades para a cor do carro, sendo que a testemunha o identificou como verde.

A partir daí, a dupla continuou a discutir, e seus componentes fizeram vários cálculos até chegarem onde eles queriam chegar: compreender o que estavam fazendo.

Várias outras perguntas apareceram, como por exemplo:

- Há alguma diferença quando temos <u>sabendo que</u> e <u>e</u>?

Eles sabem que há diferença, mas afirmam que sempre encontram dificuldades para lidar com esta situação. Sabem que com o <u>e</u> têm de fazer a multiplicação das probabilidades dos eventos, e no caso da condicional também, porém, dependendo da situação, têm algo mais para se fazer, mas não sabem o quê.

CONCLUSÃO DA ENTREVISTA

Conversando com a professora da classe, ela mostrou-se surpresa com as observações dessa dupla, pois considerava seus componentes, alunos desinteressados em sua aula e que apresentavam pouca produção.

Percebemos através dessa dupla, a confirmação das conclusões tiradas do teste, e também que muitas vezes o aluno não procura um outro caminho por considerá-lo difícil sem ao menos tentar, criando concepções como a da árvore de probabilidades ser “cheia de galhos”.

Como será que foi criada tal concepção? Será que foi pelo professor? Qual a concepção do professor sobre utilizar várias representações? Há ocasiões em que é melhor optar por uma ou outra representação e deve caber ao professor, saber trabalhá-la bem, de modo que não crie qualquer tipo de obstáculo didático[1] para o aluno, na aprendizagem, fazendo com que este não queira dela se utilizar. Como por exemplo, resolver um problema de probabilidade com muitos eventos utilizando a árvore de probabilidades como ferramenta. Esta situação levará à construção de uma árvore “cheia de galhos” como relata o aluno, que ele poderá facilmente se perder no meio de tantas ramificações, isto poderá vir reforçar tal concepção.

Algumas Considerações

As questões desta pré-experimentação foram elaboradas por nós, e se fôssemos realizá-la novamente, teríamos levado em consideração alguns outros aspectos, como por exemplo:

- ✓ Introduzir questões que envolvam os registros simbólico, árvore de probabilidades e tabela de contingência;

[1] Segundo Guy Brousseau os obstáculos de origem didática são aqueles que parecem depender apenas de uma escolha ou de um projeto do sistema educativo.

- ✓ Criar situações em que os alunos possam estabelecer comparações entre várias probabilidades condicionais; e

- ✓ Pedir justificativas de todas as questões, incluindo aquelas cujos cálculos parecem imediatos, pois só assim teríamos condições de analisar as escolhas do aluno.

Essas três considerações nos levaram a refletir sobre uma melhor maneira de elaborar questões que envolvessem o ensino da Probabilidade Condicional, que nos influenciou para as confecções das questões da nossa seqüência de ensino.

CAPÍTULO V : A SEQÜÊNCIA DIDÁTICA

Com o objetivo de introduzir o Conceito de Probabilidade Condicional, elaboramos uma Seqüência Didática composta de quatro atividades.

A primeira delas tem por objetivo introduzir o Conceito da Probabilidade Condicional e do **diagrama de árvore** através de uma situação lúdica, sem a preocupação com uma representação simbólica da Probabilidade Condicional.

Esta atividade foi inspirada nas situações didáticas de Batanero e outros. Acrescentamos em seu enunciado o nome de um Parque de diversões, que é muito comentado pelos jovens, o Hopi-Hary, por acreditarmos que isto tornaria a leitura do problema mais atraente. Elaboramos questões abertas nessa atividade e sempre pedindo aos alunos que justifiquem suas respostas, sejam na linguagem natural ou da maneira que eles preferirem.

Na segunda atividade, começamos a nos preocupar em trabalhar o **diagrama de árvore** e a **tabela de dupla entrada**, valorizando estas representações na resolução de problemas de probabilidade. Inspiradas também nas situações didáticas de Batanero, montamos a atividade baseada em dois partos que um médico irá realizar numa determinada cidade. Acreditamos que esta situação favorecerá uma boa discussão entre os alunos.

Com a terceira atividade, procuramos introduzir uma certa formalização na representação simbólica da Probabilidade Condicional, sempre articulando-a com o **diagrama de árvore** e a **tabela de dupla entrada,** procurando dar o devido tratamento a cada uma dessas representações. Para elaborarmos a situação desta atividade, pensamos na própria composição dos alunos da turma de Licenciatura em Matemática e Ciência da Computação, Homens e Mulheres desses cursos.

Consideramos a quarta atividade como se fosse uma “avaliação” do aprendizado dos conceitos trabalhados nas anteriores, pois nesta o aluno terá de fazer as articulações acima mencionadas para responder as questões aqui propostas. A situação foi inspirada num problema atual da nossa época: indivíduos portadores de HIV positivo. Criamos uma história com dados fictícios, mas que poderia representar muito bem qualquer cidade do Brasil, envolvendo esses portadores e a fidedignidade dos testes de laboratórios para diagnosticar tal doença.

Procuramos colocar nesta seqüência, situações bem diferentes das tradicionais. Criamos contextos que não estão ligados nem a “dados”, nem a “urnas”, nem a “caixas” e nem a “bolas”.

Estas atividades foram elaboradas com o intuito de fazer com que os alunos se motivem a discutir não somente as questões nelas envolvidas, mas que também se entusiasmem com a história de cada uma, criando um ambiente em sala de aula propício para o processo ensino-aprendizagem. Desejamos, também, que as situações que envolvem essas atividades, facilitem a compreensão das questões colocadas, pois diagnosticamos na Pré-Experimentação, que alguns alunos têm dificuldade de entender os enunciados por não compreenderem o contexto que os envolve.

Em todas essas atividades, elaboramos questões pouco convencionais, mas acreditamos que através delas estaremos provocando nos alunos uma maneira de refletir sobre conceitos que envolvem a Probabilidade Condicional e suas várias representações.

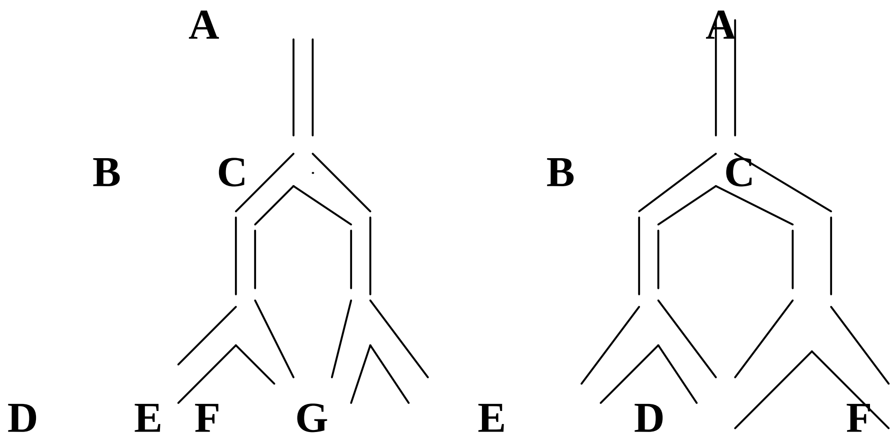
Esperamos, dessa forma, responder as perguntas formuladas na problemática. Acreditamos, assim, que estas articulações entre os diferentes registros auxiliem na apropriação do “saber” por parte dos alunos.

5.1 ANÁLISE A PRIORI

Faremos a análise a priori das atividades de nossa seqüência didática e alguns comentários sobre as suas escolhas.

ATIVIDADE 1

Beto e Meire foram ao Hopi-Hary. Encontraram lá um jogo que envolve duas máquinas. Colocamos o esquema delas abaixo:



Máquina 1 **Máquina 2**

Em ambas as máquinas, jogando-se uma bola em A, ganha-se um prêmio se a bola cair em D.

Esta atividade foi inspirada em uma situação didática encontrada no livro de Batanero, Godino e Cañizares. Adaptamos seu enunciado e acrescentamos alguns tópicos para tentarmos obter uma maior reflexão por parte dos alunos. A sua escolha se deve ao fato de, além de lúdica, ser uma atividade que recorda conceitos básicos de probabilidade. Acreditamos, também, que seja uma maneira de se introduzir com facilidade o conceito da Probabilidade Condicional.

Baseando-se nessa situação foram elaboradas oito perguntas:

Questão a) Em que máquina você pensa que Beto e Meire tenham mais chance de ganhar?

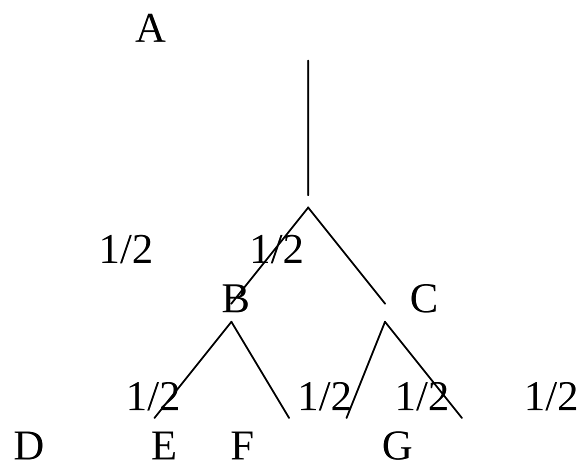
Com esta questão esperamos que os alunos identifiquem, somente através da visualização, que a maior chance de ganhar o prêmio ocorre quando se utiliza a máquina 2. Isso deverá acontecer, pois é visível que nela existem dois caminhos para chegar a D, enquanto que na 1 existe somente um caminho. Também queremos introduzir, através dessas máquinas, uma possível representação da árvore de probabilidades.

Questão b) Qual é a probabilidade de a bola cair em D na máquina 1? E na máquina 2?

Os alunos deverão calcular as probabilidades de cair em D tanto na máquina 1 como na 2. Poderão fazê-lo em ambas de duas maneiras:

Na máquina 1:

- Calculando o número de caminhos que levam a D, que é somente um caminho nesta máquina, e dividir pelo número total de caminhos que a bola possa percorrer (quatro caminhos), encontrando assim P(D) = 1/4, na máquina 1; ou
- Atribuir diretamente a probabilidade de cada trecho de caminho e usar o princípio multiplicativo. Isto ocorrerá caso o aluno consiga relacionar com eventos compostos. Vamos comentar esta resolução logo a seguir.



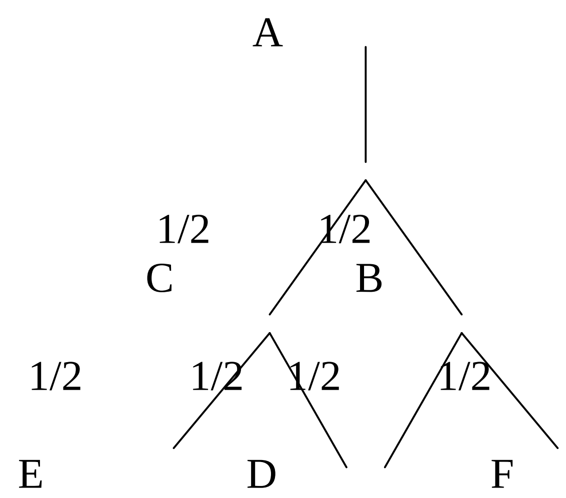
Observando o esquema acima da máquina 1, podemos usar o princípio multiplicativo para obtermos a probabilidade de chegar a D, isto é:

P(D)= 1/2 .1/2 =1/4, em que a probabilidade de a bola passar por B é igual a 1/2 e, em seguida cair em D, também é 1/2.

Na máquina 2:

- Calculando o número de caminhos que levam a D, que são dois nesta máquina, e dividindo pelo número total de caminhos que a bola possa percorrer (quatro caminhos), encontrando, portanto, P(D) = 2/4 =1/2 na máquina 2; ou

  - Atribuir diretamente a probabilidade de cada trecho de caminho e usar o princípio multiplicativo. Isto ocorrerá caso o aluno consiga relacionar com eventos compostos. Vamos comentar esta resolução logo a seguir.



Observando o esquema da máquina 2, podemos usar o princípio multiplicativo para obtermos a probabilidade de a bola chegar a D, passando por B e depois passando por C. Isto é, P(D passando por B)= 1/2 .1/2 =1/4, sendo que a probabilidade da bola passar por B é igual a 1/2 e, em seguida de cair em D também é 1/2; o mesmo ocorre se a bola passar por C, neste caso P(D) também é 1/4. Falta calcular a probabilidade total, isto é a soma da Probabilidade de cair em D, sabendo que ela passou por C, com a Probabilidade de cair em D, sabendo que ela passou por B. Para isso basta somar 1/4 +1/4 = 1/2.

Estes cálculos parecem ser fáceis e quase intuitivos, mas atrás deles existem conceitos, como os que envolvem o Teorema da Probabilidade Total, com os quais acreditamos que os alunos consigam trabalhar sem ser preciso formalizá-los.

Estes esquemas da máquina 1 e 2 sugerem, e muito, a árvore de probabilidade que pretendemos introduzir na atividade seguinte, e nosso intuito é exatamente este.

Questão c) Observe a máquina 1, e calcule a probabilidade de a bola chegar a F.

A questão foi colocada para uma futura comparação do seu resultado quando estabelecermos uma **condição** para que ocorra o mesmo evento. Na sua resolução os alunos não deverão ter dificuldade, pois eles poderão resolvê-la dos mesmos modos que resolveram a anterior.

Observando o esquema que colocamos da máquina 1, podemos obter a Probabilidade da bola chegar a F pelo princípio multiplicativo, mostrando-nos que essa probabilidade é igual a 1/4 , ou se preferirmos a divisão do número de caminhos que levam a F pelo número de caminhos totais que a bola possa percorrer, obteremos também a mesma resposta. Expressaremos isso através da linguagem simbólica a seguir:

$P(F_1) = 1/2.1/2 = 1/4$, ou

$P(F_1) = 1/4$, divisão de somente um caminho para chegar a F pelo total de caminhos que a bola possa percorrer (4).

Questão d) Ainda na máquina 1, calcule a probabilidade da bola chegar a F, sabendo-se que ela passou por C.

Nesta questão, os alunos já se encontram em uma situação de calcularem uma **probabilidade condicional**, pois estamos fornecendo nela um dado a mais que é, “ a bola passou por C”. Acreditamos que eles não terão dificuldade frente a essa situação, e atribuímos isso ao fato deles terem observado até aqui vários percursos da bola e já estarem, portanto, bem familiarizados com os caminhos percorridos pela bola nas duas máquinas.

Podemos responder a esta questão, observando o esquema da máquina 1 que fizemos anteriormente, pensando o seguinte: se a bola já passou por C, desprezamos esta probabilidade que é 1/2 , e vamos direto na probabilidade da bola cair em D que é 1/2. Indicaremos tal resposta utilizando:

- símbolo da condicional, P(F/C) = 1/2 ou,
- na linguagem natural, probabilidade da bola cair em D sabendo que passou por C é igual a 1/2.

Nosso objetivo nesta questão é fazer com que os alunos apliquem o conceito da Probabilidade Condicional através de uma linguagem simples sem que necessitem recorrer a fórmulas, preocupados mais com o seu conceito do que com a sua formalização.

Questão e) As probabilidades das letras c) e d) são iguais ou diferentes? Por que isto aconteceu?

O objetivo desta questão é instigar os alunos a refletirem sobre o porquê dessas duas situações darem probabilidades diferentes, para que na institucionalização possamos oficializar o conceito de **dependência e independência de eventos**. Esperamos que os alunos indiquem que as probabilidades sejam diferentes nas questões c) e d) e também que justifiquem suas respostas por termos diminuído o número de casos do nosso espaço amostral, mesmo que ainda não relacionem com as definições de **dependência dos eventos**.

Com essa questão pretendemos levar os alunos a refletirem sobre a ocorrência de dois eventos, mostrando que quando damos uma condição, poderemos estar mudando a probabilidade pedida, isto é, nos dois casos procuramos saber a Probabilidade da bola chegar a F, no primeiro sabemos que P(F) = 1/4, quando perguntamos qual a Probabilidade de chegar a F, sabendo que passou por C, obtemos P(F/C) = 1/2. Portanto, verificamos que a P(F) $\neq$ P(F/C) e estes dois eventos são dependentes.

Nesta institucionalização não pretendemos fazê-la de maneira formal, mas de uma maneira um tanto quanto intuitiva, mostrando que para chegar a F, se faz necessário passar por C, concluindo que estes dois eventos são dependentes.

Questão f) Observe agora a máquina 2, calcule a probabilidade da bola chegar a D.

Nesta questão repetimos uma parte da questão b), somente para ficar perto da próxima questão, com o objetivo de compará-las.

Questão g) Ainda na máquina 2, calcule a probabilidade da bola chegar a D, sabendo-se que ela passou por C. Em seguida calcule, a probabilidade dela chegar a D, sabendo-se que ela passou por B. Essas duas probabilidades são iguais ou diferentes? Justifique.

Estamos novamente diante de situações que envolvem probabilidades condicionais, desta vez tomando a máquina 2 como base. Quando os alunos calcularem a probabilidade da bola chegar a D passando por B ou por C, que acreditamos que eles o farão sem dificuldade, observarão que são iguais e deverão atribuir isto ao fato de que sempre existe caminho para chegar a D, independente de ser por B ou C, com mesma probabilidade.

Para que os alunos calculem tais probabilidades, basta observar o esquema da máquina 2 e verificar que a probabilidade da bola chegar a D, sabendo que passou por B é 1/2, o mesmo ocorrendo com a probabilidade da bola chegar a D, sabendo que passou por C, calculo este já foi justificado em questões anteriores.

Podemos concluir que :

$P(D/B) = P(D/C) = 1/2 = P(D)$ (questão f).

Se fosse de nosso interesse seria uma boa hora para introduzirmos a definição de eventos independentes usando a notação acima, mas continuamos a fazer com que o aluno pense na dependência ou independência de maneira intuitiva e não numérica.

Questão h) E as probabilidades das letras f) e g) são iguais ou diferentes? Por que você acredita que isso ocorra?

Temos como objetivo nesta questão propiciar aos alunos que reflitam diante das probabilidades de **eventos independentes**, mesmo que eles não saibam ainda identificá-los pelo nome, mas já dando condições de análise para posteriormente comentarmos tal conceito.

ATIVIDADE 2

O problema da natalidade mostrado nesta atividade e a apresentação desta através de um **diagrama de árvore** foi adaptada por nós, tomando como base as situações didáticas de Batanero e Godino.

Com esta atividade iniciamos uma certa formalização na **representação simbólica** e introduzimos a **tabela de contingência** e o **diagrama de árvore** como sendo outras maneiras de **representação** das probabilidades.

Destacamos que para resolver as questões dessa atividade, são necessários também conhecimentos básicos da aritmética, como: multiplicação e adição de frações ou, dependendo da escolha dos alunos, o cálculo com números decimais para operar com as porcentagens.

Reproduziremos abaixo essa situação adaptada, e em seguida, apresentaremos oito questões.

Ao estudar o número de natalidade de uma certa cidade do Brasil, observamos que a probabilidade de nascer um homem é de 40%, e de nascer uma mulher é de 60% .

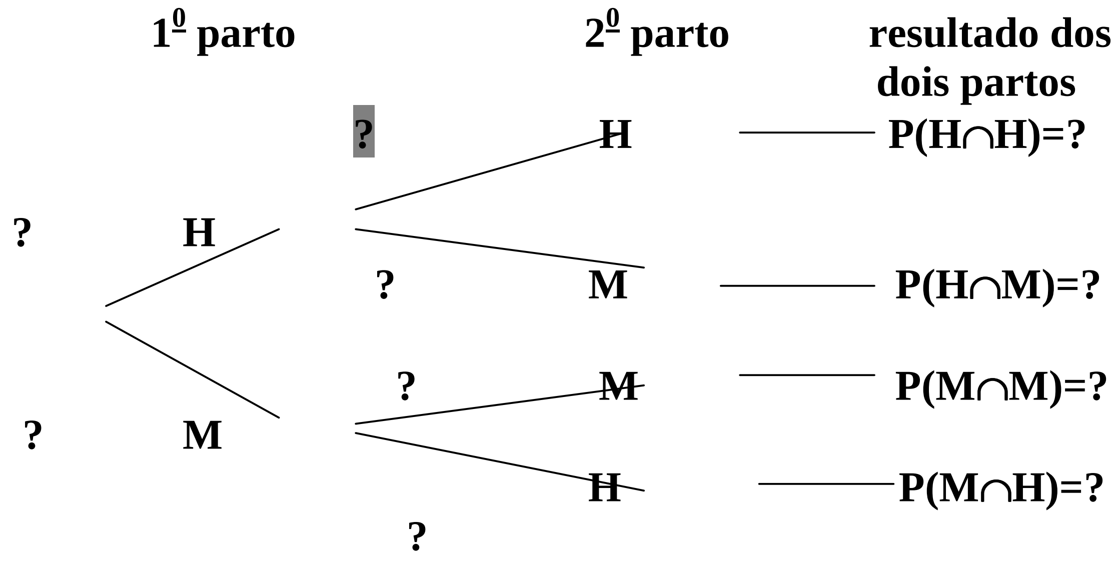
Questão a) Em uma determinada manhã um médico irá fazer dois partos. Qual dos eventos abaixo vocês consideram que é mais provável?

- **Os dois recém-nascidos são homens.**
- **Os dois recém-nascidos são mulheres.**
- **Um recém-nascido é homem e outro mulher.**

Esta questão foi colocada procurando propiciar uma reflexão sobre estes eventos, com o objetivo de visualizar posteriormente estas mesmas condições no **diagrama de árvore**.

Se os alunos conseguirem associar essa situação com as máquinas da atividade 1, poderão responder que o evento "um recém-nascido é homem e outro é mulher" tem maior probabilidade de ocorrer que "os dois recém-nascidos são mulheres" ou "os dois recém nascidos são homens". Porém, não acreditamos que isso irá ocorrer, uma vez que a questão está sendo apresentada apenas na linguagem natural, cabendo uma possível conversão (da linguagem natural para um diagrama de árvore ou para uma linguagem simbólica), e assim pensamos que poderão responder que os três eventos têm a mesma probabilidade.

Questão b) Utilize o diagrama abaixo para responder as questões a seguir:



(Na árvore acima assinalamos uma interrogação (?), que significa a probabilidade de nascer um homem no segundo parto, sabendo ter sido homem no primeiro).
Substitua as interrogações acima pelas suas respectivas probabilidades.

Nesta questão, apresentamos aos alunos a **árvore de possibilidades**, de maneira que eles compreendam o significado de alguns de seus componentes. Em seguida, pretendemos induzi-los a transformá-la em uma **árvore de probabilidades**, calculando suas respectivas probabilidades.

Esperamos que os alunos acertem essa questão, pois ao exibirmos a árvore de possibilidades, fizemos a primeira conversão (da linguagem natural para o diagrama de árvore). Cabe a eles somente colocar as probabilidades e, em seguida, fazer um cálculo elementar da multiplicação das referidas probabilidades. Para transformar a árvore de possibilidades do enunciado para a árvore de probabilidades são necessários alguns cálculos que mostraremos a seguir.

Para começarmos efetuar estes cálculos, basta ler no enunciado do problema qual a probabilidade de nascer homem e mulher, e colocar no lugar das respectivas interrogações, e em seguida calcular as interseções dos eventos, efetuando os produtos na ordem indicada na árvore, como se pode observar a seguir:

$1^{\underline{0}}$ parto $2^{\underline{0}}$ parto resultado dos dois partos

H $P(H \cap H) = 0{,}4.0{,}4 = 0{,}16$

0,4

H

0,4 0,6 M $P(H \cap M) = 0{,}4.0{,}60 = 0{,}24$

0,6 0,6 M $P(M \cap M) = 0{,}6.0{,}6 = 0{,}36$

M

0,4

H $P(M \cap H) = 0{,}6.0{,}4 = 0{,}24$

Sendo os eventos independentes, as probabilidades condicionais referentes aos segundos "galhos da árvore", são as mesmas probabilidades de ser homem ou mulher, isto é, P(H) = P(H/M) = P(H/H) = 60%, e acreditamos que o aluno compreenda isto sem necessariamente precisar escrever na linguagem simbólica. O alunos estarão, portanto, se familiarizando com algumas regras de tratamento no registro "árvore de probabilidades".

Introduzimos, também, a leitura da Probabilidade Condicional com a indicação **?** .

Questão c) Pensem novamente na questão a). Vocês mantêm sua resposta? Justifiquem sua resposta.

Ao pedirmos para voltar à primeira questão, após o contato com o esquema da árvore, esperamos que os alunos possam fazer agora uma análise melhor, e, com isso, reformularem ou não sua resposta.

Para responder esta questão baseando-se na árvore de probabilidades, basta efetuarmos :

$P(M \cap H) + P(H \cap M) = 24\% + 24\%$, e obter 48%, como resposta.

A árvore de probabilidades favorece os cálculos da probabilidade da soma das interseções (Teorema da Probabilidade Total), não sendo preciso mudar o registro, bastando compreender este tratamento na árvore.

Questão d) Coloquem os dados que vocês escreveram na árvore de probabilidades na tabela a seguir:

1º parto / 2º parto	H	M	Total
H	**$P(H \cap H)$=?**		
M			
Total			**1**

Nesta questão estamos mostrando uma tabela de **dupla entrada** e também provocando uma **conversão** para este novo **registro,** quando colocamos $P(H \cap H) = ?$

Acreditamos que os alunos não encontrarão dificuldade em responder, e esperamos que percebam esta forma de representar as interseções dos eventos, e as probabilidades totais de ser homem ou de ser mulher no primeiro e no segundo parto. Nosso objetivo é que o aluno identifique a **tabela de contingência** como uma ferramenta fácil de manipular e útil para a resolução de problemas que envolvam conceitos ligados à **Probabilidade Condicional**.

Os alunos poderão utilizar a árvore de probabilidade da questão b) para completar a tabela de contingência. Assim procedendo, estarão fazendo uma **conversão** do registro “árvore de probabilidades” para o registro “tabela de contingência”, ou se não relacionarem com a árvore, começarão a resolver a questão pelo enunciado do problema, e nesse caso, a **conversão** será do registro “linguagem natural” para o registro “tabela de contingência”. Em qualquer um dos casos, obterão os resultados:

1º parto 2º parto	H	M	Total
H	P(H∩H)=0,16	0,24	0,40
M	0,24	0,36	0,60
Total	0,40	0,60	1

Questão e) Suponhamos que vocês já saibam que no primeiro parto tenha nascido uma menina. Qual a probabilidade no segundo ser um menino?

Estamos diante de um problema em cuja solução se mobilizará o conceito de **Probabilidade Condicional.** Acreditamos que o aluno será capaz de resolvê-lo sem grandes dificuldades. Queremos que o aluno perceba que é fácil calcular as **probabilidades condicionais** através do **diagrama de árvore**.

Para responder esta questão, basta olharmos para a árvore de probabilidades da questão b) e identificarmos que tal probabilidade está colocada no “segundo galho” de M $\to$ H, que é 40%. Numa linguagem simbólica poderíamos colocar esta probabilidade da seguinte maneira:

P(H/M) = 40% .

Questão f) Qual a probabilidade de nascer um menino no segundo parto?

Nesta questão os alunos poderão recorrer tanto à árvore de probabilidades como à tabela de dupla entrada. Se eles recorrerem à árvore terão de efetuar P(H∩H) + P(M∩H), verificando que tal probabilidade é igual a 40%, o que significa estar utilizando implicitamente o Teorema da Probabilidade Total. Poderão recorrer também à tabela, observando que a soma da linha referente a H, é a probabilidade procurada.

Esperamos que a maioria dos alunos consiga estabelecer a relação entre a **árvore de probabilidades** e a **tabela de dupla entrada**, e que diante de problemas de **condicional,** mobilize esses diferentes **registros de representação** com o intuito de facilitar a resolução de questões que envolvam o conceito de **Probabilidade Condicional**.

Questão g) As probabilidades das questões e) e f) são iguais ou diferentes? Como você justifica o fato?

Os alunos possivelmente responderão que são iguais, e isto se deverá ao fato de que, no primeiro parto na questão e) já nasceu uma menina e, na questão f), mesmo sabendo que também poderá ser um menino ou uma menina que tenha nascido no primeiro parto, esse fato não alterará essa probabilidade. Em f) somaremos duas probabilidades de interseções de eventos e na e) consideraremos somente a probabilidade de nascer um menino, independente do nascimento no primeiro parto.

Desejamos que os alunos comecem a analisar a dependência e a independência de eventos, não através de uma regra, mas sim refletindo sobre a leitura do contexto em que os eventos estão inseridos.

Questão h) Suponhamos que vocês chegaram atrasados para conferir o sexo da criança que nasceu primeiro. Mas vocês verificaram que a segunda criança era uma menina. Qual a probabilidade do primeiro ter sido um menino?

Podemos resolver esta questão baseando-nos na independência dos eventos, isto é, como os eventos nascer homem e nascer mulher são independentes, esta probabilidade é igual a 40%, que é a mesma probabilidade de nascer homem em qualquer um dos partos.

Acreditamos que os alunos não conseguirão respondê-la, pois essa questão é reconhecida nas pesquisas feitas em Probabilidade Condicional, como sendo difícil para a maioria dos estudantes. A pergunta é formulada sabendo o que aconteceu no segundo parto, e questiona-se qual a probabilidade de ter ocorrido menino no primeiro parto. Para os alunos, tal pergunta não tem procedência, pois eles acreditam ser impossível não saber algo que já aconteceu, e portanto, não conseguem pensar em probabilidade se o fato é passado.

Simbolicamente podemos representar tal probabilidade por: $P(H_1/ M_2)$. E na linguagem natural seria a probabilidade de nascer homem no primeiro parto, sabendo que nasceu uma mulher no segundo.

Segundo os pesquisadores, os alunos respondem tal pergunta invertendo o que se pede no problema pelo que se é dado, isto é, calculando $P(M_2/H_1)$, ao invés de $P(H_1/M_2)$. Pesquisadores denominam este comportamento apresentado pelos estudantes, de “Fenômeno Falk”.

Acreditamos, portanto, que os alunos deverão ter dificuldade para resolver esta questão, mas a colocamos nesta atividade com o objetivo de gerar uma discussão entre os elementos das duplas, levando-os à reflexão sobre o assunto para trabalharmos com tal situação no momento da formalização.

ATIVIDADE 3

Nesta atividade, as três primeiras questões são pouco convencionais, e foram colocadas com o objetivo de fazer o aluno construir os registros “árvore de probabilidades” e “tabela de contingência”, para cada vez mais ir se familiarizando com esses registros, para que estes sirvam de ferramenta para responder as demais questões.

Essa atividade é composta de 10 questões, que são extraídas da seguinte situação:

Em uma Universidade, estão assistindo a uma aula de Estatística os alunos, conforme a tabela abaixo:

SEXO / CURSO	H	M	Total
LM	**5**	**3**	**8**
CC	**8**	**4**	**12**
Total	**13**	**7**	**20**

H: nº de Homens na sala de aula;

M: nº de mulheres na sala de aula;

LM: nº de alunos do curso de Licenciatura em Matemática; e

CC: nº de alunos do curso Ciência da Computação.

A escolha dessa atividade se deve ao fato da aplicação da mesma ser em uma sala de aula, em que estarão alunos de Matemática e de Ciência da Computação, portanto relatando no enunciado o que ocorre nesta sala.

Procuramos uma atividade ligada à noção “conjuntista” de Probabilidade Condicional, referenciada no capítulo que consta “O Objeto Estocástico”, que para Totohasina, é o enfoque em que o aluno apresenta menos dificuldade em manipular o referido conceito.

Apresentamos aos alunos uma situação que propicia efetuar algumas mudanças de registros (do enunciado para o diagrama de árvore, do diagrama para a tabela de contingência) e no mesmo registro, privilegiar alguns tratamentos para cada um deles, que explicaremos no decorrer de nossas análises. Para Duval, quanto maior for essa flexibilidade entre diferentes registros, maior será a possibilidade de apreensão do objeto estudado.

As questões são as seguintes:

Questão a) “Monte” uma árvore de possibilidades começando pelos eventos Homem (H) e Mulher (M), e calcule todas as probabilidades correspondentes.

Esperamos que os alunos acertem a questão, devido ao fato do **diagrama de árvore** ter sido manipulado por eles na atividade 2.

Nas atividades anteriores era dada a **árvore de possibilidades.** O objetivo desta questão é que os alunos comecem a construí-la a partir do seu enunciado, para, em seguida, transformá-la em uma **árvore de probabilidades** (**conversão**) e atribuir alguns tratamentos neste **registro.**

A resolução da questão, que figura abaixo, facilitará a compreensão desta conversão e dos tratamentos:

LM **P(H∩LM) = 13/20. 5/13=5/20=1/4**

5/13

H

13/20 **8/13** **CC** **P(H∩CC)= 13/20.8/13= 8/20= 2/5**

3/7 **LM** **P(M∩LM) = 7/20.3/7= 3/20**

7/20 **M**

4/7 **CC** **P(M∩CC) = 7/20.4/7 = 4/20 = 1/5**

Para “montarmos” a árvore de probabilidades acima, começamos pelo cálculo da probabilidade de escolhermos um homem da sala; para isso utilizaremos a divisão do número de homens da sala pelo total de alunos e encontraremos como resposta 13/20. O mesmo fazemos quando calculamos a probabilidade do evento “ser mulher”. Da leitura do enunciado para as probabilidades colocadas no diagrama de árvore, realizamos uma conversão sem necessariamente passar pelo registro simbólico.

Os cálculos das probabilidades dos “segundos galhos” devem ser efetuados considerando que já ocorreram os eventos “Homem” ou “Mulher”, portanto estarão sendo calculadas as Probabilidades Condicionais. Para isso, como estamos diante de uma situação “conjuntista” para calcular a probabilidade de os alunos serem do “Curso Licenciatura de Matemática”, sabendo que é Homem, basta dividirmos o número de Homens do curso de Licenciatura de Matemática pelo número de Homens da sala, resultando em 5/13. Procedemos da mesma maneira para obtermos o cálculo das outras três Probabilidades Condicionais. A passagem do enunciado do problema para a colocação das Probabilidades Condicionais nas árvores é uma conversão de registros.

O cálculo das probabilidades das interseções referentes à seqüência dos galhos, se efetua multiplicando-se as probabilidades encontradas nos galhos dessa seqüência, efetuando-se desse modo, um tratamento no registro, “diagrama de árvore”.

Questão b) “Monte” uma árvore de possibilidades começando pelos cursos, ao invés do sexo, e em seguida calcule as probabilidades correspondentes.

Esta questão tem por objetivo propiciar aos alunos a reflexão sobre uma outra maneira de construir a árvore de probabilidades, para que mais tarde eles possam perceber qual a melhor forma de utilizá-la. Acreditamos que eles consigam resolver a questão, pelo mesmo motivo apontado na anterior.

Vamos mostrar a montagem desta árvore:

H P(LM∩H) = 8/20.
5/8=5/20=1/4
5/8

LM

8/20 3/8 M P(LM∩M)= 8/20.3/8= 3/20

8/12 H P(CC∩H) = 12/20.8/12= 8/20
=2/5
12/20 CC

4/12 M P(CC∩M) = 12/20.4/12 =
4/20 = 1/5

Estamos diante de uma situação “conjuntista” que facilita os cálculos das interseções e das condicionais como foi mostrado no “O Objeto Estocástico” , portanto para a construção desta árvore basta seguir as etapas da montagem da anterior.

Questão c) Transforme a sua tabela de informações do enunciado em uma tabela de probabilidades.

O objetivo desta questão é fazer com que os alunos comecem a manipular as mudanças de registro e distinguir os **tratamentos** em cada um desses **registros.** Acreditamos que eles não apresentarão dificuldades em resolver tal questão. Eles têm as duas árvores a) e b), como ferramentas para respondê-la.

Vamos apresentar a tabela transformada para, a seguir, discutirmos alguns aspectos de sua montagem.

SEXO / CURSO	H	M	Total
LM	**5/20**	**3/20**	**8/20**
CC	**8/20**	**4/20**	**12/20**
Total	**13/20**	**7/20**	**1**

Se os alunos observarem a árvore de probabilidades e constatarem que as interseções calculadas são as que devem ser colocadas no quadro, estarão fazendo uma conversão do registro "diagrama de árvore" para a "tabela de contingência". A soma das linhas e das colunas nesse registro é um dos tratamentos próprios deste registro. Nas questões seguintes mostraremos outros tratamentos que podemos efetuar dentro da tabela de contingência.

Questão d) Sabendo-se que escolhemos um rapaz dessa sala, qual a probabilidade dele ser do curso de Licenciatura em Matemática?

Com esta questão pretendemos que os alunos percebam que a probabilidade solicitada já foi calculada, e está colocada na árvore de probabilidades da questão a), fazendo com que eles atentem para a utilidade do diagrama da árvore, na resolução de problemas que envolvem a condicional. Esperamos que eles respondam corretamente essa questão, pois acreditamos que já estão se familiarizando com a árvore de probabilidades. Para isto basta olhar a árvore montada na questão a) e observar que a probabilidade que se encontra no galho H —— LM, é 5/13; podemos escrever P(LM/H) = 5/13.

Também podemos responder essa questão utilizando alguns tratamentos na "tabela de contingência", isto é, observamos a probabilidade de ser Homem na tabela de contingência 13/20, em seguida a de ser Homem e do curso de Licenciatura de Matemática, 5/20, dividindo uma pela outra obteremos a probabilidade pedida, assim como representamos abaixo:

$$P(LM / H) = \frac{5/20}{13/20} = 5/13$$

Tomando alguns dados da tabela de contingência e aplicando o tratamento que demos acima, podemos calcular as probabilidades condicionais que envolvem estes eventos.

Apresentamos o enunciado da questão na linguagem natural, sem nos preocuparmos com a simbologia da Probabilidade Condicional, mas voltaremos a apresentá-la em algumas das questões que vêm a seguir.

Questão e) Sabendo-se que escolhemos um aluno do curso Licenciatura de Matemática, qual a probabilidade dele ser um Homem?

Nosso objetivo é propiciar aos alunos uma reflexão sobre qual árvore eles utilizarão como ferramenta para responderem à questão, e perceberem a importância de começar um **diagrama de árvore** pelos dados que são fornecidos pela condicional. Se houver tal reflexão, acreditamos que os alunos escolherão a árvore da questão b), e responderão corretamente à questão.

As perguntas d) e e) estão relacionadas com o "Fenômeno Falk", já comentado anteriormente. Na questão d) queremos calcular a Probabilidade do aluno ser do curso de Licenciatura em Matemática sabendo que é Homem, portanto tal probabilidade encontra-se na árvore da questão a) e na e) sabendo que o aluno é do curso de Licenciatura em Matemática, queremos calcular qual a probabilidade de ser Homem. Numa linguagem simbólica, temos:

- na questão d) queremos P(LM/H) = 5/13, que obtemos observando a árvore da questão a) , e

- na questão f) P(H/LM) = 5/8, observando a árvore da questão b).

Questão f) As probabilidades encontradas em d) e e) são iguais ou diferentes? Comente sua resposta.

Estamos apresentando uma situação em que os alunos possam observar que P(A/B) não é igual P(B/A), para que eles futuramente não cometam o erro de igualar os dois resultados. Esperamos que acertem a questão dizendo que as probabilidades são diferentes, e expliquem que, o que está sendo considerado dado em uma questão é diferente do que está sendo dado na outra. Como justificamos nas considerações da questão anterior.

Questão g) Substitua as interrogações abaixo pelas respectivas probabilidades:

I) P(H) = ?　　III) P(H/LM) = ?

II) P(M) = ?　　IV) P(M/CC)= ?

Esperamos que os alunos observem a árvore de probabilidades da questão b) e respondam sem dificuldade a questão. Nosso objetivo é destacar essas probabilidades para analisar a sua dependência na questão seguinte.

Este é o primeiro enunciado que fizemos da Probabilidade Condicional na linguagem simbólica, embora tenhamos, durante a discussão das atividades anteriores, nos utilizados dela sempre na lousa, de modo que esperamos que os alunos já tenham adquirido alguma familiarização com a mesma.

Mostraremos a seguir as respostas esperadas:

I) P(H) = 13/20　　III) P(H/LM) = 5/8

II) P(M) = 7/20　　IV) P(M/CC)= 4/12

As respostas I e II aparecem na árvore de probabilidades da questão a), nos seus primeiros galhos, ou na soma das próprias colunas da tabela de contingência, as III e IV aparecem na árvore de probabilidades da questão b) sobre os segundos galhos.

Questão h) Tomando como base o item g).

As probabilidades I) e III) são iguais ou diferentes? Justifique.

As probabilidades II) e IV) são iguais ou diferentes? Justifique.

Nos dois casos esperamos que eles digam que são diferentes, justificando que os eventos são dependentes. O objetivo é fazer com que os alunos reflitam em cada atividade sobre a dependência ou independência dos eventos.

Questão i) Observando a árvore de probabilidade do item a), responda:

$P(H) + P(M) =$

$P(LM/H) + P(CC/H) =$

$P(LM/M) + P(CC/M) =$

$P(H \cap LM) + P(H \cap CC) + P(M \cap LM) + P(M \cap CC) =$

O que estas somas têm em comum? Justifique.

O objetivo desta questão é provocar uma reflexão sobre algumas regras de tratamento no registro “diagrama de árvore”, como por exemplo identificar que a soma das probabilidades sobre os galhos que saem de um mesmo nó é igual a 1, e que a soma de todas as interseções também é 1.

Esperamos que os alunos respondam em todos estes itens, que a soma é igual a 1, e consigam justificar dizendo que os eventos, em cada caso, são complementares.

Questão j) Para calcular P(M/CC) você pensa que seria melhor utilizar a árvore do item a) ou do item b)? Justifique o porquê da sua escolha.

Acreditamos que os alunos respondam que o item b) seja a melhor árvore para responder a essa questão, pois eles já refletiram sobre as duas árvores várias vezes.

Nosso objetivo é fazer com que os alunos percebam a importância de começar a árvore de probabilidades pelo dado que é fornecido na condicional.

ATIVIDADE 4

Essa atividade é composta de 9 questões que são extraídas da seguinte situação:

Numa determinada cidade, a probabilidade de encontrarmos um indivíduo portador do vírus HIV_+ é de 6%. Em um teste feito num determinado laboratório da cidade, ocorreram os seguintes resultados:

- **A probabilidade de um indivíduo ter seu teste positivo (T_+), sabendo-se que é portador do vírus HIV_+ , é de 99,8%.**
- **A probabilidade de um indivíduo ter seu teste negativo (T_-) , sabendo-se que não é portador do vírus HIV_+ , é de 99,6%.**

São utilizadas as notações:

D para representar o evento do indivíduo ser HIV_+ .

$\overline{D}$ para representar o evento do indivíduo não ser HIV_+ .

Esta atividade foi elaborada com o intuito de analisarmos como os alunos apresentam os seus conhecimentos após terem manipulado os registros: simbólico, árvore de probabilidades e tabela de contingência, nas anteriores.

Não colocamos nenhum dos dois registros nem na situação problema nem no enunciado das questões, simplesmente pedimos que os alunos os representem e respondam algumas questões envolvendo a probabilidade condicional.

Embora essa atividade envolva a probabilidade condicional “causalista”, que para Totohasina é aquela que apresenta maior dificuldade para os alunos, acreditamos que irão resolvê-la, pois já estarão bem familiarizados com todas as ferramentas que os ajudarão a fazê-lo.

Os dados fornecidos no problema sobre resultados de exames em indivíduos portadores de soro positivo, são fictícios. O que nos levou a escolher tal tema, foi o desejo de lembrar uma realidade mundial sobre a AIDS.

Questão a) Construa a árvore de probabilidade partindo dos eventos D e $\overline{D}$.

Para resolverem esta questão, os alunos deverão usar uma mudança de registro do enunciado do problema na **linguagem natural** para a **árvore de probabilidades**. Nesta questão, diferentemente das outras, eles terão que calcular as probabilidades complementares, pois nesse enunciado elas estão implícitas. Poderão fazê-lo lembrando das regras de tratamento que discutimos na atividade anterior, quando verificamos que a soma das probabilidades ligadas ao mesmo nó é igual a 1. Nosso objetivo, com essa questão, é fazer com que os alunos reflitam sobre as situações anteriores e comecem a apresentar uma certa autonomia na resolução de problemas dessa natureza. Acreditamos que não terão dificuldades em resolvê-la, por estarem bem familiarizados com essa representação.

Colocaremos a árvore a seguir para melhor ilustrar as observações feitas.

T_+ $P(D \cap T_+) = 6\%.99,8\% = 5,988\%$

99,8%

D

6% 0,2% T_- $P(D \cap T_-) = 6\%.0,2\% = 0,012\%$

0,4% T_+ $P(\overline{D} \cap T_+) = 94\%.0,4\% = 0,376\%$

94% $\overline{D}$

99,6% T_- $P(\overline{D} \cap T_-) = 94\%.99,6\% = 93,624\%$

O tratamento neste registro permite encontrar as probabilidades assinaladas na árvore, nos primeiros e segundos galhos, isto é, a soma das probabilidades ligadas ao mesmo nó é igual a 1.

Questão b) "Monte" a tabela de dupla entrada com as respectivas probabilidades.

Nosso objetivo é verificar se os alunos conseguem passar de um **registro** a outro (conversão), como explicaremos a seguir**.** Esperamos que eles não apresentem dificuldades em fazê-lo, pela familiarização adquirida até então.

Apresentamos a seguir esta tabela:

Portador / Teste	**D**	**$\overline{D}$**	**Total**
T_+	**5,988%**	**0,376%**	**6,364%**
T_-	**0,012%**	**93,624%**	**93,636%**
Total	**6%**	**94%**	**100%**

Os alunos deverão observar a árvore de probabilidades da questão anterior, e convertê-la.

Questão c) Calcule a probabilidade de se encontrar T_+ nos exames desse laboratório.

Com esta questão, pretendemos fazer os alunos manipularem a ferramenta **tabela de contingência**, além de verificarem a sua utilidade. A resposta resulta diretamente da **tabela de contingência** por eles construída. Ela representa um tratamento nesse registro, a soma da linha T_+; acreditamos por isso, que não haverá dificuldades para resolver essa questão. Aqui os alunos deverão utilizar o Teorema da Probabilidade Total, soma de decimais, e apresentar 6,364% como resposta ou, se estiverem familiarizados com as notações simbólicas, escreverão: $P(T_+) = 6{,}364\%$.

Questão d) Calcule a probabilidade de se encontrar T_- nos exames desse laboratório.

O objetivo e as expectativas desta questão são praticamente os mesmos da anterior. A diferença estará no cálculo, que, ao invés de somar-se a linha do T_+ , soma-se a do T_- . Deverá ser obtido 93,636% como resposta.

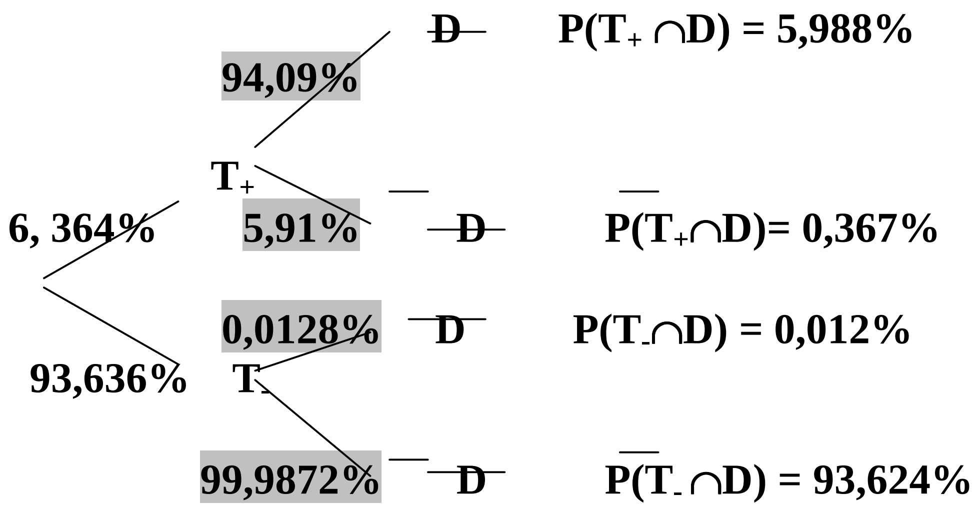
Questão e) Construa a árvore de probabilidades partindo dos eventos T_+ e T_- .

Colocamos os alunos novamente diante de uma **conversão**: eles poderão dispor de uma **árvore de probabilidades** começando por D e de uma **tabela de contingência** que podem ser utilizadas como ferramentas. Nessa questão mobilizam-se conceitos ligados à **Probabilidade Condicional** como, o **Teorema da Probabilidade Total** e a **Interseção de Eventos Dependentes.**

Os alunos não deverão estar sabendo, que ao passarem de uma árvore a outra já estarão utilizando o Teorema de Bayes. Acreditamos que esse tipo de conversão dá condições de resolver problemas que mobilizem tal teorema.

Esperamos que, após algumas reflexões, eles consigam resolver a questão. Sabemos que não é uma questão fácil, mas acreditamos que as atividades anteriores tenham dado aos alunos condições para resolvê-la.

Vamos apresentar a seguir a árvore em questão:



As questões c) e d) ajudarão a montar esta árvore, e foram colocadas nesta ordem com este objetivo. Os alunos têm, portanto, as probabilidades do início desta árvore calculadas nas duas questões anteriores. As interseções são as mesmas calculadas na árvore da questão a), bastando somente que os alunos calculem as condicionais e, para isto, deverão utilizar a divisão da probabilidade dos “primeiros galhos” pelas probabilidades das respectivas interseções (tratamento).

Questão f) Qual a probabilidade de escolhermos um indivíduo ao acaso desta cidade, sabendo-se que o seu teste deu negativo, que seja portador do vírus HIV_+?

Acreditamos que, se os alunos conseguirem construir a árvore da questão e), eles não apresentarão dificuldades em responder esta questão. Queremos que eles percebam a importância de manipular esse registro, para quando estiverem diante de um enunciado de eventos condicionais os representem de várias maneiras.

A probabilidade pode ser vista sobre os “segundos galhos” da questão e), que ligam T_{-} ——— D, que é 0,01281 e poderá ser escrito:

$$P(D/ T_{-})=0,01281\% .$$

Questão g) Os eventos D e T_{+} são dependentes ou independes? Comente sua resposta.

Deverá ser fácil para os alunos responderem que os eventos são **dependentes**, pela reflexão feita em outras atividades e em questões desta mesma como, por exemplo, a e). Como sabemos que reconhecer a **dependência** dos eventos é motivo de dificuldade para os alunos, colocamos essa questão para reforçar tal conceito.

Questão h) Responda os itens abaixo justificando como encontraram as suas respostas.

I) $P(T_{-} /D) =$
II) $P(D/T_{+}) =$
III) $P(T_{+}/ D)=$
IV) $P(D/T_{-}) =$

Esta questão foi formulada após a aplicação da atividade 3, pois foi diagnosticado naquela que a maior parte dos alunos não compreendem corretamente este símbolo da condicional, portanto salientamos a sua leitura durante as discussões das respostas da atividade 3, e colocamos aqui essa questão para observarmos quais serão as atitudes dos alunos diante desta simbologia.

Esperamos que eles consigam diferenciar o item I do IV e o II do III. Essa diferença os alunos a detectaram anteriormente através de perguntas feitas na linguagem natural, e desejamos verificar na linguagem simbólica.

Mostraremos a seguir as respostas:

I) $P(T_- /D) = 0,2\%$
II) $P(D/T_+) = 94,09\%$
III) $P(T_+/ D) = 99,8\%$
IV) $P(D/T_-) = 0,01281\%$

Podemos justificar tais respostas por serem probabilidades calculadas nas árvores das questões a) e e). As probabilidades dos itens I) e III) figuram nos segundos galhos que saem do Nó “D”, da árvore da questão a) e, dos itens II) e IV) dos segundos galhos que saem do Nó “T_-**”**, da árvore da questão e).

Questão i) Como você escreveria na linguagem corrente essas duas probabilidades: $P(D \cap T_+)$ e $P(D/T_+)$. As probabilidades são iguais ou diferentes? Justifique.

Esperamos que os alunos façam a **conversão do registro simbólico** para a **linguagem corrente**, que é o contrário do que em geral aparece nos enunciados dos problemas e acreditamos que eles não terão dificuldades em resolver. Além disso, desejamos propiciar aos alunos uma reflexão sobre duas situações **($P(D \cap T_+)$ e $P(D/T_+)$).** Esperamos, também, que os alunos respondam que essas probabilidades são diferentes, pelos eventos serem **dependentes**.

5. 2 REALIZAÇÃO DA SEQÜÊNCIA

A seqüência didática é composta de quatro atividades e foi aplicada aos alunos do 3º ano de Licenciatura em Matemática e Ciência da Computação, da Universidade Católica de Santos (UNISANTOS), que é oferecido somente no período noturno. Os alunos desses dois cursos assistem algumas disciplinas juntos, e Estatística é uma delas.

A disciplina Estatística tem uma carga horária 80 horas anuais, e seu conteúdo programático está no ANEXO VI.

A aplicação da seqüência ocorreu após os alunos terem contato com os conceitos básicos de Estatística Descritiva e de Probabilidade com um evento, isto é, o conceito de Probabilidade "Objetiva" (comentada no Capítulo III). O professor desta sala confirmou que não havia trabalhado com registro de "árvore" nem a "tabela de contingência" , com seus alunos.

O trabalho se realizou em três sessões. A primeira ocorreu no dia 28/08 do corrente ano com duração prevista de noventa minutos. Compareceram 32 alunos que foram divididos em duplas, de acordo com suas próprias escolhas.

Nessa primeira etapa foram trabalhadas as duas primeiras atividades. Entregamos a primeira e estabelecemos o prazo de vinte minutos para o seu término; em seguida fizemos a institucionalização da regra do produto e dos conceitos de dependência e independência de eventos, e ainda introduzimos a notação de Probabilidade Condicional e mudança dos esquema das “Máquinas 1 e 2” em **Diagramas de Árvore.** Utilizamos a correção das questões para introduzir tais conceitos, tentando aproveitar as próprias respostas dos alunos dadas oralmente às questões e estabelecendo uma relação com o que eles falavam e os conceitos mencionados acima. Fazendo uma formalização sempre através de um diálogo entre professor e alunos, e nunca através de um monólogo, no qual somente o professor fala. Esta postura foi tomada por nós em todas as formalizações.

A seguir, entregamos a segunda atividade para que, com essas institucionalizações, os alunos pudessem realizá-la tendo uma maior preocupação maior com a notação adequada, utilizando a nova ferramenta **diagrama de árvore**. No final dessa atividade estava prevista uma nova discussão, com a finalidade de reforçar os conceitos citados acima, além de introduzir uma nova ferramenta para trabalhar com o conceito **da Probabilidade Condicional**, que é a **tabela de dupla entrada**. Isso porém, não foi possível porque, embora a duração da aula fosse de noventa minutos, os alunos já mostravam sinais de cansaço, o que é bastante compreensível, uma vez que a aula de Estatística acontece das 21h05min às 22h35min. Resolvemos assim fazer essa institucionalização no início da próxima etapa.

Durante a aplicação dessas duas atividades foram muitos os questionamentos por parte dos alunos sobre como operar com as porcentagens. Eles preferiram manter suas respostas utilizando as porcentagens ao invés dos registros fracionários, porém mostraram uma certa dificuldade em efetuar operações com elas. Fizemos uma breve discussão de como trabalhar com operações que envolvem porcentagens, para podermos desenvolver melhor nossas atividades.

Acreditamos que suas preferências pelo uso da porcentagem devam ser devido aos meios de comunicação, que sempre tratam situações que envolvem incertezas através de porcentagem.

O professor de Estatística dos alunos permaneceu todo o tempo da aplicação das atividades na sala de aula, não interferindo em nenhum momento.

Os alunos mostraram muito interesse no desenvolvimento das atividades, e isso foi percebido pelas inúmeras perguntas feitas durante a sua execução, e também através das respostas que eles elaboraram, escrevendo quase sempre as justificativas das suas escolhas, e muitas vezes nos chamando para verificarmos se o que foi escrito dava para ser compreendido.

A segunda sessão foi realizada no dia 03/09, na semana seguinte, e contamos com a participação de 21 alunos. A falta de muitos deles foi justificada pelo professor por ser uma semana com feriados e muitos dos alunos acabam "emendando", mas mesmo assim, ele considerou que o número de presenças estava somente um pouco abaixo da média.

Como prevíamos, comentamos a atividade anterior e ratificamos o uso da **árvore de probabilidades** e da **tabela de contingência** para problemas que envolvessem a Probabilidade Condicional. Nesta oficialização da segunda atividade, os alunos deram oralmente algumas conclusões sobre os dois registros utilizados: a **árvore de probabilidades** e a **tabela de contingência.** Conclusões estas que destacamos a seguir:

- A soma de uma das colunas ou linhas da tabela de contingência nos dá as probabilidades totais.
- Os "segundos traços" da "árvore de probabilidades" refere-se à Probabilidade Condicional.
- Os dados que precisamos saber da tabela de contingência estão todos calculados na "árvore de probabilidades".

- Os dois eventos em questão são independentes, diferente da atividade 1, na máquina 2.

Os alunos organizaram-se em duplas, e como eram 21 alunos, um deles juntou-se a uma das duplas. Distribuímos a Atividade 3, e eles começaram a trabalhar com o mesmo interesse e participação da anterior. Sua duração foi de trinta e cinco minutos. Como o ocorrido na atividade anterior, não foi possível fazer comentários gerais sobre esta atividade e deixamos este acontecimento para a próxima sessão.

Durante a aplicação desta atividade, os alunos também fizeram perguntas de como trabalhar com as porcentagens; alguns deles optaram por dar a resposta no registro fracionário e outros no decimal, por acharem mais fácil realizar os cálculos, mas admitiram que compreendem melhor se a resposta estiver em porcentagem. Este fato nos faz refletir sobre:

"Que concepções os alunos têm da representação da probabilidade através do número decimal ou fracionário?"

Os alunos mostram que não têm problema de operar frações ou decimais, mas nos parece que têm alguma dificuldade em trabalhar com os seus respectivos significados no contexto da probabilidade.

A atividade 4 foi aplicada no dia 24/09, três semanas após a realização da terceira sessão. Isto ocorreu devido a dois fatores: na primeira semana houve uma série de palestras para o curso de Ciência da Computação, e na seguinte, os alunos tinham prova de Estatística. Portanto, depois dessas três semanas, fomos levados a fazer uma discussão da atividade 3 um pouco mais detalhada, tentando retomar não só a sua problemática como todos os conceitos envolvidos. Aproveitamos o intervalo, que ocorreu antes do início desta aula, para colocarmos na lousa a situação.

Na discussão da atividade 3, montamos as árvores de probabilidades e a tabela de contingência na lousa levando sempre em consideração as respostas dadas oralmente pelos alunos e procurando fazer com que eles as justificassem.

A esta quarta sessão, compareceram 29 alunos, sendo que somente 20 destes participaram da terceira, embora todos tenham participado da primeira. Isto criou uma certa dificuldade para corrigir a atividade anterior, mas mesmo os alunos que não haviam participado dela, fizeram algumas perguntas, mostrando interesse em compreender o que se estava falando.

Para a formação das duplas sugerimos aos alunos faltosos na atividade 3, que se agrupassem com quem estivera presente no dia em que fora trabalhada. A maioria deles aceitou a sugestão, mas uma dupla preferiu desenvolver esta atividade 4 sem que nenhum dos dois alunos tivesse participado da anterior. Ficamos em sala com 13 duplas e um trio.

Desta atividade 4 os alunos resolveram as quatro primeiras questões rapidamente, e sem nos fazer nenhuma pergunta. Percebemos que eles discutiam entre si, mas não nos apresentaram dúvidas. Estas só começaram a acontecer na questão e), pois precisavam calcular a Probabilidade Condicional que não foi dada no enunciado, e para isso fizeram muitas perguntas, que serão comentadas no capítulo seguinte.

Esta sessão durou cerca de trinta e cinco minutos, e os alunos mostraram que não tiveram dificuldades de compreender nem a situação e nem os dados que colocamos no enunciado.

A classe sempre se mostrou participativa e interessada em desenvolver as atividades, e percebemos que nesta última os alunos faziam com mais convicção as argumentações e os questionamentos entre eles, e também aqueles que se dirigiam a nós.

5.3 ANÁLISE DOS RESULTADOS

RESULTADOS OBTIDOS NA ATIVIDADE 1

Os alunos mostraram bastante interesse em responder as questões desta atividade, como era esperado. Eles acreditavam que realmente estavam diante de um jogo.

Como previamos, o número de acertos das 16 duplas que realizaram esta atividade foi grande, como podemos observar na tabela abaixo:

Questões	Acertos	Erros	Em Branco
a)	16	-	-
b)	16	-	-
c)	15	01	-
d)	14	02	-
e)	15	01	-
f)	15	01	-

g)	15	01	-
h)	15	01	-

Vamos analisar individualmente cada questão.

Questão a) Em que máquina você pensa que Beto e Meire tenham mais chance de ganhar?

Como era esperado, os alunos confirmaram que a maior chance de ganhar o prêmio estaria na escolha da máquina 2, e nos parece que tal resposta foi baseada somente na visualização das máquinas, pois todos não apresentaram nenhum cálculo.

Questão b) Qual a probabilidade de a bola cair em D na máquina 1? E na máquina 2?

Embora os alunos tenham acertado esta questão, ficou difícil analisarmos como obtiveram o resultado, pois somente 04 duplas indicaram como fizeram o cálculo. Dentre elas, 03 utilizaram a divisão do número de caminhos que levam a D, pelo número total de caminhos que a bola possa percorrer, e a outra dupla utilizou o princípio multiplicativo.

As demais duplas indicavam somente as máquinas 1 ou 2 seguida do seu resultado. Percebemos com este fato que os alunos mostravam uma falta tanto de representação em relação à probabilidade como uma falta de formalismo matemático. Fato este, observado pela ausência do sinal de igual, indicação de resposta através de flechas ou traços e ausência da letra P para indicar a probabilidade solicitada. Vamos colocar a seguir como se apresentaram algumas destas respostas:

- 1 → 1/4 e 2→ 1/2.
- Máquina 1 – 25%
  Máquina 2 – 50%

Questão c) Observe a máquina 1, calcule a probabilidade da bola chegar a F.

Os alunos não tiveram dificuldade diante da questão, como era esperado. Os mesmos problemas relatados na questão anterior aparecem nos resultados desta.

A dupla que errou a questão, desprezou a possibilidade da bola passar por B, considerando, portanto, somente a possibilidade da bola cair em C, admitindo a probabilidade da bola chegar a F ser de 50% .

Questão d) Ainda na Máquina 1, calcule a probabilidade da bola chegar a F, sabendo-se que ela passou por C.

Os alunos parecem ter percebido que a **condição** dada aumentava a probabilidade de chegar a F, pois resolveram uma questão cujo conceito de Probabilidade Condicional já estava nela embutido.

Embora o número de acertos tenha sido alto, tivemos alguma dificuldade de saber como as duplas efetuaram tal resposta, pela falta novamente de representação de sua resolução. Das 16 duplas somente quatro indicaram como foi feito o cálculo, não só com a indicação da Probabilidade que deverá ser calculada, como também através da linguagem natural, explicando os passos que deram para resolver a questão.

Das duas duplas que erraram, podemos supor que elas tenham desprezado a informação dada (sabendo que passou por C), pois colocaram como resposta, 25% .

Questão e) As probabilidades das letras c) e d) são iguais ou diferentes? Por que isto aconteceu?

Como era esperado, os alunos escreveram que eram diferentes e justificaram pelo fato de terem diminuído as possibilidades de chegar a F na questão d). Uma dupla errou. Esse fato pode ter ocorrido por ela ter errado a questão c) e d).

Questão f) Observe agora a máquina 2, calcule a probabilidade da bola chegar a D.

A maioria dos alunos percebeu que esta questão é igual à b). A dupla que errou, havia acertado na questão b), mas resolveu fazer um cálculo dessa probabilidade, considerando um só caminho para chegar a D. Portanto, encontrou tal probabilidade sendo 1/3.

Questão g) Ainda na máquina 2, calcule a probabilidade da bola chegar a D, sabendo-se que ela passou por C. Em seguida calcule a probabilidade dela chegar a D, sabendo-se que ela passou por B. Essas duas probabilidades são iguais ou diferentes? Justifique.

Como observamos no quadro de acertos dos alunos, houve uma grande maioria de acertos. As duplas mostraram ter compreendido bem a condição dada na questão, isto foi observado não só pelo acerto desta, como pelas justificativas que deram. Justificativas estas, que foram dadas na própria linguagem natural como também através de alguns símbolos que eles criaram para indicar a "condição", como por exemplo:

$$P/C = 50\%$$

$$P/B = 50\%$$

Duas duplas justificaram a resposta utilizando o termo "independente" para essas probabilidades serem iguais.

Questão h) E as probabilidades das letras f) e g) são iguais ou diferentes? Por que você acredita que isso ocorra?

Acreditamos que, com esta questão, conseguimos atingir o nosso objetivo, que era o de propiciar aos alunos reflexão diante das probabilidades de eventos independentes. As duplas, de um modo geral, se empenharam em justificar o fato das probabilidades serem iguais.

Nessas duas últimas questões, os alunos escreveram bastante seus pensamentos na linguagem natural, e quatro duplas referiram-se à igualdade das letras utilizando a palavra “independência”.

Vamos mostrar a seguir algumas justificativas dadas pelos alunos:

- “Iguais, porque não importa por onde a bola caia (B ou C), sempre há uma chance de 50% de cair em D”.
- “Iguais, porque a probabilidade da bola chegar a D independe do caminho que ela passe, B ou C, sempre terá probabilidade de 50%”.
- “Iguais, isto ocorre porque ambos os caminhos B e C podem terminar em D, com as mesmas probabilidades”.

RESULTADOS OBTIDOS NA ATIVIDADE 2

Os alunos fizeram muitas perguntas durante a aplicação desta segunda atividade, principalmente como efetuar a multiplicação de porcentagens, se podiam expressar na forma decimal ou fracionária.

Como fizemos entre as atividades algumas institucionalizações, os alunos, por sua vez, tentaram expressar-se utilizando alguns símbolos próprios da probabilidade e perguntavam se estavam usando-os corretamente. Os alunos demonstraram mais dificuldade nesta atividade que na anterior. Pudemos perceber isso não só pelos questionamentos durante a sua aplicação como pela quantidade de respostas erradas. Observamos esse fato na tabela a seguir:

Questões	Acertos	Erros	Em Branco
a)	04	12	-
b)	16	-	-
c)	06	10	-
d)	16	-	-
e)	09	07	-
f)	12	04	-

g)	09	06	01
h)	06	08	02

Vamos analisar individualmente cada questão.

Questão a) Em uma determinada manhã um médico irá fazer dois partos. Qual dos eventos abaixo vocês consideram que é mais provável?

- **Os dois recém-nascidos são homens.**
- **Os dois recém-nascidos são meninas.**
- **Um recém-nascido é homem e outro mulher.**

A maioria dos alunos errou a questão, como era esperado. As duplas consideraram o evento “os dois recém-nascidos são meninas” o mais provável de ocorrer. Acreditamos que isso tenha acontecido porque na questão, a probabilidade de nascer mulher é maior que a probabilidade de nascer homem. Os alunos simplesmente marcaram a resposta e não registraram nenhum cálculo que nos pudesse indicar em que dados foram baseadas as suas escolhas.

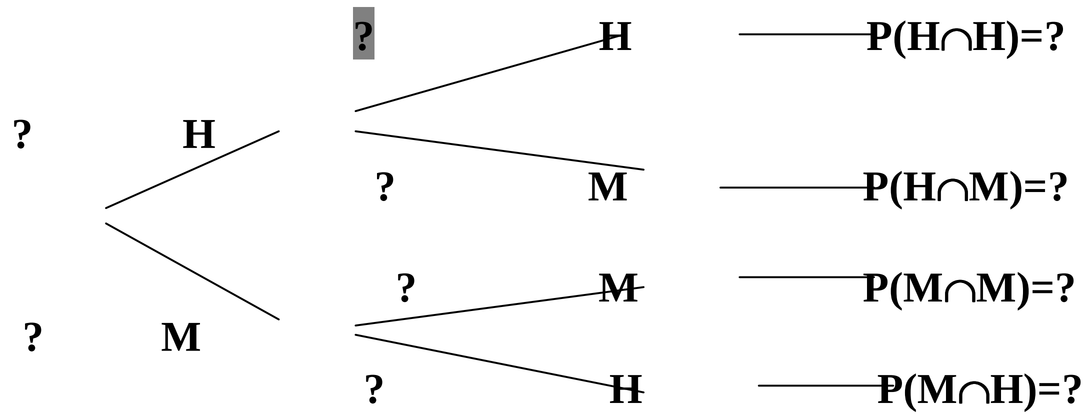
Dentre as duplas que acertaram a questão, observamos uma delas que registrou o seguinte ao lado da questão:

“AA $\boxed{\text{Aa Aa}}$ aa ”

Parece-nos que com este registro os alunos recorreram aos conhecimentos que possuíam sobre genética (Leis de Mendel), considerando <u>**A**</u> como sendo Homem e <u>**a**</u> Mulher ou o contrário. Tais conhecimentos os ajudaram a concluir que a probabilidade maior estava na alternativa “um recém-nascido é homem e outro mulher”

Questão b) Utilize o seguinte diagrama para responder as questões a seguir:

1º parto **2º parto** **resultado dos dois partos**



(Na árvore acima assinalamos uma interrogação (?), que significa a probabilidade de nascer um homem no segundo parto, sabendo-se ter sido homem no primeiro).

Substitua as interrogações acima pelas suas respectivas probabilidades.

Como era previsto, os alunos acertaram a questão. Eles perceberam quais as probabilidades que estavam sendo pedidas, e não mostraram dificuldade de interpretação da questão, porém foi uma surpresa observarmos a dificuldade do aluno diante de uma multiplicação de porcentagens. Além de não saberem multiplicá-las, eles não conseguiam visualizar numa porcentagem uma fração ou um número decimal. Como 14 das duplas fizeram a mesma pergunta, tivemos que fazer um esclarecimento geral sobre como se multiplicam as porcentagens, para que o trabalho continuasse.

Questão c) Pensem novamente na questão a). Vocês mantêm sua resposta? Justifiquem sua resposta.

Somente duas das duplas que erraram a questão a) reformularam a sua resposta. E a justificativa foi baseada na observação da árvore de probabilidade; elas perceberam que existia mais possibilidade de nascerem sexos diferentes do que iguais.

A maioria continuou dizendo que a probabilidade maior estava no nascimento das duas meninas, justificando por esta probabilidade ser 36%, portanto maior que qualquer outra. Essa justificativa mostrou-nos que eles não observaram que para calcular as probabilidades do nascimento de sexos diferentes, na árvore, teriam que efetuar a soma das probabilidades das duas interseções, que seria $P(H\cap M)$ e $P(M\cap H)$.

Questão d) Coloquem os dados que vocês escreveram na árvore de probabilidades na tabela a seguir:

$1^{\underline{0}}$ parto / $2^{\underline{0}}$ parto	**H**	**M**	**Total**
H	**$P(H\cap H)=$?**		
M			
Total			**1**

Os alunos não tiveram dificuldade em responder a esta questão, como nós já havíamos previsto. Eles estabeleceram as correspondências entre as probabilidades calculadas na **árvore** e as da **tabela**.

Como foi indicado um total igual a 1 na tabela, muitos alunos perceberam que teriam que escrever as probabilidades das interseções na forma fracionária ou em decimal. Porém, quatro das duplas persistiram em continuar com a porcentagem; duas delas colocaram uma interrogação sobre o número 1 da tabela, como se indagassem o porquê de ser 1 e não 100%.

Nessa questão nos parece que a maior dificuldade foi a de transformar a porcentagem em fração ou em decimal. Esta foi percebida pela quantidade de perguntas das duplas durante a elaboração das suas respostas Tal dificuldade não foi prevista por nós anteriormente.

Questão e) Suponhamos que vocês já saibam que o primeiro parto tenha nascido uma menina. Qual a probabilidade no segundo ser um menino?

Como podemos observar na tabela, a maioria dos alunos acertou a questão. Das duplas que acertaram, quatro delas justificaram suas respostas pelo fato de ter nascido um menino no segundo parto “independe” do que ocorreu no primeiro parto, justificativa esta feita através da linguagem natural.

Somente uma dupla utilizou o símbolo da Probabilidade Condicional, P(H/M) = 40%, justificando sua resposta da seguinte maneira:

“Se o primeiro é menina, olhando na árvore, vimos a probabilidade de ser homem e isso dá 40%”.

Com este fato, essa dupla mostrou que compreendeu bem o que significa o segundo galho da árvore: “ a Condicional”.

Entre as duplas que erraram, foi unânime a resposta 24%. Elas justificaram o fato por considerarem como sendo uma interseção dos eventos “nascer uma menina no primeiro parto” **e** “nascer um menino no segundo parto”. Estamos diante, portanto, de uma dificuldade já apresentada por outros pesquisadores que é: considerar a “interseção dos eventos” ao invés da “Condicional”.

Questão f) Qual a probabilidade de nascer um menino no segundo parto?

As respostas dos alunos confirmaram as nossas expectativas e houve um grande número de acertos nesta questão.

Dentre as justificativas dadas pelos alunos que acertaram, destacamos:

- 05 duplas consideraram a soma das seguintes probabilidades :

$$P(M \cap H) + P(H \cap H).$$

- 04 duplas consideraram a independência entre os eventos.
- 01 dupla considerou a soma da coluna H da tabela de contingência.

Duas duplas não justificaram como encontraram tal resposta.

Das duplas que acertaram, somente uma utilizou o símbolo de Probabilidade, isto é, indicou P(H), englobando duas linhas da tabela de contingência.

Das duplas que erraram, somente uma nos possibilitou perceber como foi efetuado o seu cálculo. Esta tomou como tendo nascido homem no primeiro parto fazendo, portanto, a interseção “homem no 1º parto” e “homem no 2º parto”. Observamos novamente a confusão da interseção de dois eventos com a condicional entre eles.

Questão g) As probabilidades das letras e) e f) são iguais ou diferentes? Como vocês justificam o fato?

Com esta questão a maioria dos alunos mostra uma melhora na maneira de referir-se à probabilidade. Começam a utilizar, além da linguagem natural, alguns símbolos para justificar a questão, como por exemplo: interseção dos eventos, existência da condição no enunciado da questão, as letras H e M para indicar os eventos, o símbolo P para indicar probabilidade, todas as possibilidades de uma probabilidade e a independência dos eventos.

Percebemos, dessa maneira, um crescimento não só da compreensão da Probabilidade e da Probabilidade Condicional, como na forma de representá-las. Porém observamos que algumas das duplas ainda estão tendo algumas dificuldades de trabalhar com esses conceitos. Os erros encontrados nessa questão se devem ao fato dessas duplas terem errado uma das anteriores.

Questão h) Suponhamos que vocês chegaram atrasados para conferir o sexo da criança que nasceu primeiro. Mas vocês verificaram que a segunda criança era uma menina. Qual a probabilidade do primeiro ter sido um menino?
Justifiquem sua resposta.

Como esperávamos, os alunos mostraram mais dificuldade na resolução desta questão que nas demais, o que confirma as mesmas dificuldades encontradas pelos sujeitos nas investigações realizadas por Falk e outros pesquisadores, como foi explicado na análise a priori.

Esta questão envolve a “reversibilidade” do tempo (o aluno não consegue imaginar que ele não saiba o que já aconteceu). Totohasina considera que as probabilidades condicionais cronológicas, que trabalham com essa “reversibilidade”, são difíceis para a compreensão do aluno.

Das duplas que acertaram, três delas justificaram suas respostas pela independência dos eventos e as outras três utilizaram a soma das probabilidades das interseções. Essas seis duplas mostraram gradativamente, através de suas respostas, um crescimento em relação aos conceitos ligados à probabilidade e à probabilidade condicional.

Das duplas que erraram a questão cinco deram como resposta a probabilidade de ser **“homem no primeiro parto”**, desconsiderando a possibilidade de ser **“mulher no primeiro parto”,** encontrando assim como resposta pela probabilidade da interseção dos eventos, **“homem no primeiro parto”** <u>e</u> **“mulher no segundo parto”.** Com essa resposta, percebemos a dificuldade do aluno em compreender o que está sendo dado no enunciado e o que está sendo pedido.

Uma das duplas que errou, considerou o evento ser “mulher no primeiro parto” e ser “mulher no segundo”, desprezando a pergunta da questão. E uma outra dupla considerou os dois eventos como sendo “homem no 1º parto” e “homem no 2º parto”.

Com o resultado da segunda atividade, pudemos observar um crescimento por parte dos alunos na escrita dos conceitos ligados à probabilidade, na linguagem natural. Fato este observado também nos alunos que erraram algumas questões. Isto nos deixou muito confiantes para as atividades futuras. Esperamos que tal crescimento continue acontecendo para o êxito não só dos alunos envolvidos neste processo, como para o da nossa pesquisa.

RESULTADOS OBTIDOS NA ATIVIDADE 3

Os alunos mostraram bastante interesse em realizar a atividade, fato este observado tanto pela quantidade de perguntas feitas por eles durante a realização desta como pela maneira como responderam às questões, justificando sempre as suas escolhas.

Nas respostas dos alunos ainda aparecem dificuldades de operar com as porcentagens. Acreditamos que das questões que os alunos erraram, algumas tenham sido pela falta de habilidade de manipular porcentagem, fração e número decimal. O depoimento dado por muitos dos alunos foi: “não consigo enxergar a probabilidade em uma fração nem em um número decimal me parece que a porcentagem diz mais sobre isto”.

Destacamos a situação da Atividade 3:

Em uma Universidade, estão assistindo a uma aula de Estatística os alunos, conforme a tabela abaixo:

SEXO	H	M	Tota

CURSO			l
LM	**5**	**3**	**8**
CC	**8**	**4**	**12**
Total	**13**	**7**	**20**

H: nº de Homens na sala de aula;

M: nº de Mulheres na sala de aula;

LM: nº de alunos do curso de Licenciatura em Matemática; e

CC: nº de alunos do curso Ciência da Computação.

Foi alto o número de acertos das questões, das 10 duplas de alunos que participaram desta atividade, podemos observar isto na tabela:

Questões	Acertos	Erros	Em Branco
a)	09	01	-
b)	09	01	-
c)	10	-	-
d)	08	02	-
e)	08	02	-
f)	08	02	-
g)	04	06	-
h)	10	-	-
i)	07	03	-
j)	06	04	-

Vamos analisar individualmente cada questão.

Questão a) "Monte" uma árvore de possibilidades começando pelos eventos Homem (H) e Mulher (M), e calcule todas as probabilidades correspondentes.

O número de acertos nesta questão confirma as nossas expectativas. Os alunos fizeram a conversão do enunciado do problema para o “diagrama de árvore”, mostrando que já estão adquirindo uma certa familiarização com este novo registro.

Consideramos errada a questão de uma dupla porque ficou faltando calcular as probabilidades dos “segundos galhos”, que eram as mais importantes, pois nelas está embutido o conceito da Probabilidade Condicional. Acreditamos que isto ocorreu porque os alunos dessa dupla não conseguiram reconhecer a dependência dos eventos, conseqüentemente não sabiam como calcular essa probabilidade. Efetuaram as interseções no final da árvore sem terem calculado a Probabilidade Condicional, realizando tal cálculo baseando-se na interseção dos conjuntos dados no enunciado. Esta dupla mostrou que não reconheceu a relação existente entre os eventos estabelecidos na questão.

Embora os alunos tenham tido um grande número de acertos, esta questão foi motivo de muitos questionamentos sobre como trabalhar com as porcentagens. Surgiram perguntas como: “Pode-se arredondar a porcentagem?”, “Se arredondar, como eu faço para dar 100% no total?”, “Como uma porcentagem pode não ser exata?” Mesmo com todas essas dificuldades, os alunos persistiam em continuar representando a probabilidade em porcentagem, mostrando uma certa “aversão” a representá-la na forma de fração.

Questão b) “Monte” uma árvore de possibilidades começando pelos cursos, ao invés do sexo, em seguida calcule as probabilidades correspondentes.

A maioria das duplas acertou a questão, como o esperado. Os questionamentos sobre o cálculo das porcentagens continuaram durante a execução dessa questão. Algumas duplas começam a refazer os seus cálculos e optaram pela forma fracionária para representar as probabilidades na “árvore”, justificando tal opção por ser mais fácil de realizar os cálculos.

A mesma dupla que errou a questão anterior errou esta, e acreditamos que isto tenha ocorrido pelo mesmo motivo.

Questão c) Transforme a sua tabela de informações do enunciado em uma tabela de probabilidades.

O número de acertos foi total, mas não conseguimos analisar como todas as duplas resolveram a questão. Talvez tenham tomado como base uma das árvores já construídas na questão a) ou na b), ou então partiram somente do enunciado do problema e fizeram novamente os cálculos. Em algumas das duplas foi possível diagnosticar como foram feitas as suas escolhas, pois durante o desenvolvimento dessa questão, várias duplas nos chamaram, perguntando-nos se não era só transcrever os resultados obtidos na “árvore de probabilidade” e em seguida somar as linhas e colunas correspondentes.

Se as outras duplas também resolveram a questão como descrevemos acima, elas conseguiram estabelecer relações entre os dois registros: “árvore de probabilidade” e “tabela de contingência”.

Questão d) Sabendo-se que escolhemos um rapaz dessa sala, qual a probabilidade dele ser do curso de Licenciatura em Matemática?

Como podemos observar na tabela, os alunos obtiveram um grande número de acertos nesta questão. Vamos colocar a seguir as justificativas dadas pelos alunos que acertaram a questão:

- 02 duplas indicaram a probabilidade condicional pedida como sendo P(LM/H) = 5/13, porque olharam na árvore de probabilidades da letra a).
- 01 dupla indicou como sendo P(H/LM) = 38%, justificando o mesmo que as duplas anteriores. Percebemos, então, um erro de notação e não de conceito. Para essa dupla, o que está sendo dado foi colocado em primeiro lugar na notação, e o que está sendo pedido em segundo, como aparece na árvore de probabilidades.

- 03 duplas colocaram a resposta certa, porém sem representar o P de probabilidade nem a condicional, justificando que esse resultado já foi calculado na árvore de probabilidades da letra a).
- 02 duplas fizeram o cálculo novamente, considerando tal probabilidade como sendo a divisão do número de homens do curso de Licenciatura de Matemática pelo número total de homens, não estabelecendo, portanto, correspondência entre a árvore de probabilidades da questão a) e esta questão.

Verificamos pela resposta de uma das duplas que errou, que esta considerou a interseção dos eventos "homens" **e** "alunos de Licenciatura em Matemática", não considerando a condicional. Essa dupla justificou a questão da seguinte maneira:

"25%, de acordo com o resultado final da árvore da questão a)".

Esse resultado, que é colocado no final da árvore, é de interseção dos eventos e não da probabilidade condicional.

A outra dupla que errou colocou 25% como resposta e não a justificou. O que talvez tenha considerado o mesmo que a outra dupla que errou, isto é, substituir a probabilidade da condicional pela probabilidade da interseção de eventos.

Questão e) Sabendo-se que escolhemos um aluno do curso de Licenciatura em Matemática, qual a probabilidade dele ser um Homem?

As mesmas duplas que erraram a questão anterior, erraram esta e pelo mesmo motivo; consideraram a interseção dos eventos, desprezando a condicional.

Das duplas que acertaram, destacaremos algumas observações:

- 02 continuaram a utilizar a representação da condicional corretamente e justificaram sua resposta através da árvore de probabilidades da questão b).
- 01 dupla, embora tenha acertado, continuou utilizando a representação da condicional invertendo a ordem na notação, justificando a resposta pela árvore de probabilidades da questão b).
- 03 colocaram somente a porcentagem como resposta, indicando que esta havia sido encontrada na árvore da questão b).
- 02 continuaram refazendo os seus cálculos, desprezando os dados já encontrados pelos mesmos, nas questões anteriores.

Como era esperado, a maioria dos alunos percebeu qual seria a “árvore de probabilidades” escolhida para responder a questão. Observamos através das respostas que eles não só estavam compreendendo a utilidade da árvore de probabilidades, como reconhecendo uma questão que envolve a Probabilidade Condicional.

Questão f) As probabilidades encontradas em d) e e) são iguais ou diferentes? Comente sua resposta.

Como podemos observar na tabela, a maioria dos alunos confirmou as nossas expectativas ao acertar a questão.

Das 08 duplas que acertaram, todas justificaram corretamente a sua resposta, embora tenham usado argumentos diferentes, como descreveremos a seguir:

- 04 duplas responderam que as probabilidades são diferentes pelo fato de na questão d) o que foi dado é ser “homem” e na letra e) ser do “curso de Licenciatura em Matemática”.
- 02 duplas responderam que as condições dadas indicam “caminhos diferentes”. Acreditamos que essas duas duplas tenham baseado sua resposta nas árvores de probabilidades, por usarem a expressão “caminhos”.
- 02 duplas consideraram que o “Universo da amostra em d) era limitado pelo sexo enquanto em e) era limitado pelo curso”.

Das justificativas dadas pelas duplas, verificamos que elas souberam, além de identificar cada uma das condições dadas, e também as suas diferenças. Percebemos com isto que os alunos, a partir do enunciado, diferenciaram a P(A/B) de P(B/A), diferença esta que é uma das questões levantadas na Problemática, que pretendíamos responder com essas atividades que era:

“Será que os alunos vão diferenciar o cálculo de P(A/B) de P(B/A)?”

As 02 duplas que erraram as questões d) e e), também erraram a f). Percebemos novamente que os alunos identificaram a interseção dos eventos, ao invés da condicional, e uma das duplas afirmou que o fato das duas letras terem a mesma probabilidade é que:

“a ordem dos fatores não altera o produto”.

Concluímos com esta resposta, que as duas duplas calcularam a $P(H \cap LM)$ se utilizando da tabela de contingência e a calcularam pelo produto das probabilidades dos eventos “ser homem” e “ser do curso de Licenciatura em Matemática”, mostrando com isto que não reconhecem a condicional e nem a dependência dos eventos.

Questão g) Substitua as interrogações abaixo pelas respectivas probabilidades:

I) P(H) = ? **III) P(H/LM) = ?**

II) P(M) = ? **IV) P(M/LM) = ?**

Somente quatro duplas acertaram a questão. Das duplas que erraram, todas acertaram os itens I) e II). A maioria das respostas dadas por elas para os outros dois itens, foi 25% para o item III) e 20% para o IV). Verificamos, com isto, que os alunos calcularam a probabilidade da interseção dos eventos ao invés da condicional.

Como os alunos não justificaram suas respostas, não podemos dizer se a dificuldade está na notação do que foi pedido, isto é, o que significa a **P(H/LM)** e **P(H/LM)** ou se a dificuldade está na compreensão do conceito da Probabilidade Condicional.

Como na questão anterior, a mesma pergunta foi feita na linguagem natural, apontando, portanto, que a dificuldade maior esteja com a notação da Probabilidade Condicional e não com o seu conceito, pois houve maior número de acertos.

Talvez fosse melhor antes dessa questão, nós termos colocado questões que articulassem os dois registros: o da linguagem natural com a representação simbólica da Probabilidade Condicional, para minimizar as dificuldades de notação apresentadas pelos alunos. Dificuldades estas que não havíamos previsto, mas nos fizeram introduzir na próxima atividade uma questão dessa natureza.

Questão h) Tomando como base a questão g).

As probabilidades I) e III) são iguais ou diferentes? Justifique.

As probabilidades II) e IV) são iguais ou diferentes? Justifique.

Todos acertaram a questão afirmando que as probabilidades eram diferentes, como o esperado, porém nenhuma dupla justificou utilizando a dependência dos eventos. Elas afirmaram que as probabilidades eram diferentes, porque o espaço amostral nos itens III) e IV) foi reduzido por não ser o total de alunos da classe, mas por estar se referindo também aos respectivos cursos.

Dos alunos que acertaram a questão anterior, pudemos constatar através das suas justificativas, que eles estabeleceram a relação de ter sido dado uma “condição” a mais para as probabilidades III) e IV), isto é, ser do curso de Licenciatura em Matemática na III) e ser do curso de Ciência da Computação na IV). Observamos que quatro das duplas que realizaram esta atividade, parecem estar aplicando tanto o conceito da Probabilidade Condicional como compreendendo a sua notação corretamente.

Questão i) Observando a árvore de probabilidades do item a), responda:

P(H) + P(M) =

P(LM/H) + P(CC/H) =

P(LM/M) + P(CC/M) =

P(H∩LM) + P(H∩CC) + P(M∩LM) + P(M∩CC) =

O que estas somas têm em comum? Justifique.

A maioria das duplas acertou a questão como era esperado. Das duplas que acertaram, a justificativa mais encontrada foi:

"Todos deram 100%. O primeiro porque a Probabilidade é calculada considerando o espaço total de alunos da sala, o segundo considerando o espaço total de homens da sala, o terceiro o total de mulheres da sala e o último o total de alunos da sala."

As sete duplas que acertaram a questão, parecem ter percebido a diferença entre a condicional e a interseção, pela justificativa que mencionamos acima. Só não sabemos o porquê deles algumas vezes confundirem estes dois conceitos.

Embora os alunos tenham justificado com coerência suas respostas, não sabemos de quais registros eles fizeram uso: se do diagrama de árvore, se da tabela de contingência ou de ambos.

Dos alunos que erraram a questão, observamos pelas nas suas respostas que consideraram em **P(LM/M) + P(CC/M)** sendo igual a 35% porque analisaram como sendo esta soma a Probabilidade de Mulheres da sala**,** identificando, portanto, como interseção e não como condicional, fazendo o mesmo com **P(LM/H) + P(CC/H) =65%**.

Todas as duplas acertaram o primeiro e o último itens da questão.

Questão j) Para calcular P(M/CC), você pensa que seria melhor utilizar a árvore do item a) ou a do item b)? Justifique o porquê da sua resposta.

A maioria dos alunos respondeu que seria melhor utilizar a árvore do item b), pois a resposta seria imediata, considerando que o aluno é do curso de Ciência da Computação, a resposta está na própria árvore, na probabilidade seguinte a do curso. Nas justificativas das duplas que acertaram, podemos perceber que os alunos, além de compreenderem a notação e o conceito da Probabilidade Condicional, reconhecem os elementos da “árvore de probabilidades”.

Os alunos que erraram identificaram o evento “ser mulher” como sendo dado, e o evento “aluno do curso de Ciência da Computação”, o que queríamos saber, isto é, calcularam a probabilidade de ser do curso de Ciência da Computação, sabendo-se que o aluno é mulher, trocando, portanto, o significado de P(M/CC) por P(CC/M). Acreditamos que os alunos trocaram essas probabilidades por encontrarem dificuldades com a notação da Probabilidade Condicional.

Com os resultados desta terceira atividade, podemos observar que os alunos começam a familiarizar-se com situações que envolvam a Probabilidade Condicional e as suas representações. Podemos verificar isto, não só pelo número de acertos, mas também porque os alunos cada vez mais escrevem suas respostas e as justificativas não somente na linguagem natural e, por já estarem adquirindo uma outra maneira de representar tal conceito através da representação simbólica da Probabilidade e da Condicional, além da compreensão do “diagrama de árvore” e da “tabela de contingência” como uma ferramenta para resolver problemas de Probabilidade.

RESULTADOS OBTIDOS NA ATIVIDADE 4

Os alunos, nos seus protocolos, mostraram estar familiarizados com os registros “diagrama de árvore” e “tabela de contingência”. Observamos no desenvolvimento desta atividade que eles estabeleceramm as ligações entre o enunciado do problema e esses dois registros.

Destacamos a situação da Atividade 4:

Numa determinada cidade, a probabilidade de encontrarmos um indivíduo portador de vírus HIV_+ é de 6%. Em um teste feito num determinado laboratório da cidade, ocorreram os seguintes resultados:

- **A probabilidade de um indivíduo ter seu teste positivo (T_+), sabendo-se que é portador do vírus HIV_+ , é de 99,8%.**

- **A probabilidade de um indivíduo ter seu teste negativo (T_-), sabendo-se que não é portador do vírus HIV_+ , é de 99,6%.**

São utilizadas as notações:

D representa o evento do indivíduo ser HIV_+ .

$\overline{D}$ representa o evento do indivíduo não ser HIV_+ .

Podemos observar, na tabela, que o número de acertos das 14 duplas que realizaram esta atividade foi grande, como era esperado.

Questões	Acertos	Erros	Em Branco
a)	13	01	-
b)	13	01	-
c)	13	01	-
d)	13	01	-
e)	11	03	-
f)	08	04	-
g)	12	-	02
h)	07	06	01
i)	10	02	02

Vamos analisar individualmente cada questão.

Questão a) Construa a árvore de probabilidade partindo dos eventos D e $\overline{D}$.

Os alunos fizeram a conversão do enunciado do problema para o “diagrama de árvore”, mostrando que já têm domínio sobre este registro, como era esperado. Pudemos constatar o fato não só pelo número de acertos, como também pela rapidez com que resolveram a questão e por não terem feito perguntas durante a realização desta.

As probabilidades complementares foram resolvidas, e os alunos não mostraram dificuldade neste cálculo. Acreditamos que isto tenha ocorrido por eles terem percebido as regras (tratamentos), já mencionadas anteriormente para a montagem da árvore de probabilidade.

Somente uma dupla montou a árvore sem calcular as probabilidades das interseções dos eventos, e por isso, consideramos a questão errada. Os dois alunos dessa dupla haviam faltado na aplicação da atividade 3, e não quiseram se separar para trabalharem com outros colegas. Notamos que eles tiveram muitas dificuldades não só na resolução desta questão, mas como em toda esta atividade.

Questão b) “Monte” a tabela de dupla entrada com as respectivas probabilidades.

Os alunos fizeram a mudança de registro do “diagrama de árvore” para a “tabela de contingência” com bastante facilidade, notamos isto pelo número de acertos e pela rapidez com que fizeram.

Pudemos observar, durante a resolução desta questão, que os alunos perceberam as ligações existentes entre estes dois registros e o enunciado do problema, e os diálogos entre eles sobre o que significava cada dado calculado mostravam isto. Destacamos algumas frases:

- “Devemos retirar da “árvore” as interseções”.
- “Quando somamos todas as interseções, sempre tem que dar 100%, se não der, erramos em algum lugar”.
- “Temos que fazer a soma das linhas e das colunas”.

- “Eu sei que a soma dessas linhas e dessas colunas serve para responder algumas perguntas de probabilidade”.

A dupla que errou esta questão, trocou a $P(D \cap T_-)$ pela $P(D \cap T_+)$, na tabela. Acreditamos que isto tenha ocorrido por falta de atenção e não por falta de conhecimento.

Questão c) Calcule a probabilidade de se encontrar T_+ nos exames desse laboratório.

Como era esperado, a maioria dos alunos acertou a questão, e destes, todos justificaram o seu acerto escrevendo algumas frases como:

- “É a soma da coluna do **T_+ ,** que pode ser observada na tabela da questão b).”
- “Se somarmos $P(D\cap T_+)$ e $P(\overline{D}\cap T_+)$ encontraremos esta probabilidade.”

Verificamos com estas duas respostas que alguns alunos recorreram à “tabela de contingência”, como na primeira justificativa, e outros à “árvore de probabilidades”, como na segunda justificativa.

Embora o número de acertos tenha sido grande, somente seis duplas indicaram a probabilidade utilizando o símbolo $P(T_+)$; os demais indicaram o valor em porcentagem e em seguida justificaram o seu cálculo sem se utilizar de qualquer simbologia, somente a linguagem natural.

Verificamos que os alunos utilizaram o Teorema da Probabilidade Total sem que este tivesse sido anteriormente formalizado para eles. O que significa que tal Teorema já está implícito na compreensão dos tratamentos tanto no “diagrama de árvore” como na “tabela de contingência”.

Os alunos continuam a optar pela porcentagem tanto para efetuarem os cálculos como para responderem às questões, e mostram que fazem esses cálculos com mais facilidade que na Atividade 3, pois os questionamentos sobre tais operações não ocorreram mais.

Questão d) Calcule a probabilidade de se encontrar T_- nos exames desse laboratório.

Como era esperado, os alunos acertaram a questão e perceberam que a diferença da questão anterior era a coluna escolhida na tabela de contingência. Afirmaram em suas justificativas que deveríamos somar as probabilidades da coluna do T_- ao invés do T_+ , como na anterior.

Somente seis duplas indicaram simbolicamente o que estavam calculando, já os outros utilizaram-se da linguagem natural.

Questão e) Construa a árvore de probabilidade partindo dos eventos T_+ e T_- .

Durante o desenvolvimento desta questão, os alunos começaram a fazer perguntas sobre a questão. Todos perceberam que deveriam começar a construir a árvore com as probabilidades que encontraram nas questões d) e e), mas como efetuar os cálculos das probabilidades dos "galhos" seguintes (as condicionais), foi motivo de muitas discussões entre eles. E estas geraram questionamentos (por parte dos alunos) dirigidos a nós, como:

- "As probabilidades dos segundos galhos não estão no enunciado do problema?"
- "Podemos utilizar tanto a árvore da questão a) como a tabela para responder a pergunta?"
- "As interseções desses dois eventos (D e T_+) são iguais aos da "árvore" da questão a) ?"
- "Devemos dividir a probabilidade da interseção dos eventos (D e T_+) pela probabilidade do teste ser positivo?"

Todas estas perguntas foram remetidas aos alunos que as formulavam, a fim de que eles respondessem aos seus próprios questionamentos.

Os alunos, além de perceberem as diferenças das Probabilidades Condicionais $P(D/T_{+})$ e $P(T_{+}/D)$, utilizaram os conceitos que evolvem o Teorema da Probabilidade Total e o Teorema de Bayes. Verificamos isto tanto pela quantidade de acertos como pelos questionamentos citados acima.

Das três duplas que erraram a questão, duas delas se perderam nos cálculos com as porcentagens, indicaram divisão correta, mas apresentaram a resposta errada, ou na posição da vírgula ou nos próprios cálculos. A outra dupla considerou $P(T_{+}/D)$ como sendo também $P(D/T_{+})$, errando totalmente a questão.

Esperávamos com esta questão que os alunos refletissem sobre alguns conceitos ligados à Probabilidade Condicional. Podemos perceber que isto aconteceu realmente e que uma questão considerada difícil foi resolvida para maioria dos alunos, como era esperado.

Questão f) Qual a probabilidade de escolhermos um indivíduo ao acaso dessa cidade, sabendo-se que o seu teste deu negativo, que seja portador do vírus HIV_{+} ?

Como era esperado, a maioria dos alunos acertou a questão, percebendo que poderiam utilizar a “árvore” da questão anterior para responder a pergunta, justificando, dessa forma, o seu resultado.

Somente cinco das duplas que acertaram, indicaram que $P(D/T_{-})$ era a probabilidade que estava sendo pedida, as demais colocavam a resposta em porcentagem seguida da justificativa.

Destacaremos abaixo as respostas das quatro duplas que erraram esta questão:

- Coloca a resposta correta, mas indica que é P ($T_- \cap D$) é a probabilidade calculada.
- Indica corretamente qual probabilidade que é calculada P(D/T_-), mas multiplica o seu resultado pelo número 6. Acreditamos que os alunos tenham multiplicado pela probabilidade do indivíduo ser portador do vírus.
- Afirma que a probabilidade pedida é a mesma do enunciado do problema, confundindo, portanto, P(D/T_-) com P(T_-/D).
- Indica que a probabilidade pedida é P($\overline{D}$/T_-) e calcula corretamente esta probabilidade. Acreditamos que estes alunos confundiram a notação, mas não o conceito da Probabilidade Condicional.

Questão g) Os eventos D e T_+ são dependentes ou independentes? Comente sua resposta.

Os alunos responderam que estes dois eventos são dependentes, como era o esperado. Destacaremos, a seguir, algumas das respostas colocadas por eles em seus protocolos:

- "Para haver um teste positivo foi necessário existir um indivíduo portador ou não."
- "São dependentes, pois a porcentagem do teste ser positivo depende do indivíduo. Basta analisar: se a ordem de início da árvore de probabilidades for D, dará um resultado, se for T, dará outro."

Os alunos compreendem que existe uma relação de dependência entre os eventos, embora ainda não coloquem isto na forma simbólica, mas também este não era o nosso objetivo.

Questão h) Responda os itens abaixo justificando como encontraram as suas respostas.

I) P(T_- /D) =

II) $P(D/T_+)$ =
III) $P(T_+/D)$ =
IV) $P(D/T_-)$ =

Os alunos tiveram muita dificuldade de diferenciar o item I) do III) e o II) do IV), e verificamos que esta questão foi a que mais eles erraram.

Os alunos que acertaram a questão justificaram as suas respostas no item I) e III) como sendo as probabilidades encontradas sobre os segundos galhos da árvore da questão a), e dos outros dois itens como sendo as probabilidades encontradas na árvore da questão e).

Dos alunos que erraram, verificamos em seus relatos a grande confusão que eles fizeram com a notação da Probabilidade Condicional. Eles não diferenciaram o que era dado do que era pedido trocando, portanto, $P(D/T_+)$ por $P(T_+/D)$, até aqueles alunos que acertaram a mesma pergunta feita na linguagem natural na questão f).

Os alunos de um modo geral têm dificuldade de relacionar a linguagem natural com a notação da Probabilidade Condicional.

Questão i) Como você escreveria na linguagem corrente essas duas probabilidades: $P(D \cap T_+)$ e $P(D/T_+)$? As probabilidades são iguais ou diferentes? Justifique.

Como era esperado, os alunos fizeram a conversão do registro simbólico para a linguagem natural. Portanto, o fato mostra que eles estabelecem a diferença entre probabilidade da interseção e a Probabilidade Condicional destes dois eventos. E ainda, que eles conseguem ler corretamente a notação da Probabilidade Condicional.

Talvez fosse o caso de explorarmos mais questões que envolvessem a notação da Probabilidade Condicional, de modo a articular com mais ênfase a passagem da linguagem natural para o símbolo da condicional e vice-versa, para melhor compreender o que acontece com os alunos frente a essas notações.

Conclusões

Nosso trabalho teve como objetivo introduzir o conceito de Probabilidade Condicional em cursos de Estatística da Universidade. Levamos em consideração, para isso, dificuldades levantadas por nós e por alguns pesquisadores sobre o ensino da Probabilidade Condicional. Construímos, então, uma seqüência didática, composta de quatro atividades, para tentarmos minimizar essas dificuldades. Os protocolos dos alunos e as discussões durante as correções das atividades da seqüência de ensino nos permitiram levantar algumas conclusões que apresentaremos a seguir.

O início da aplicação da seqüência provocou uma ruptura do contrato didático[1] existente até o momento de iniciar a seqüência. Os alunos durante a aplicação foram participativos, mostraram interesse fazendo indagações e buscando orientação sobre o encaminhamento de algumas questões, nunca solicitando diretamente as respostas. Também ao trabalharem em dupla, eles participavam ativamente na formação do conceito de Probabilidade Condicional, quando discutiam a realização de cada atividade proposta. Esta forma de trabalho até então não era usual. Embora não seja o nosso objetivo analisar os efeitos do contrato didático na dinâmica da aplicação da seqüência, acreditamos que esta ruptura contribuiu para o desenvolvimento de nossas atividades.

Como nossas atividades eram compostas, em sua maioria, de questões abertas e ao pedirmos sempre que os alunos as justificassem, tornou possível sabermos como foram feitas as suas escolhas, facilitando assim nossas análises.

---

1 – Segundo Guy Brousseau chama-se contrato didático o conjunto de comportamentos do professor que são esperados pelos alunos e o conjunto de comportamentos do aluno que são esperados pelo professor.

Constatamos através dessas justificativas que, em cada questão, elas foram se aprimorando no desenrolar das atividades, mostrando que os alunos não só adquiriram um vocabulário próprio da probabilidade como também sabiam usá-lo, principalmente na linguagem natural.

Com a nossa seqüência procurávamos fazer com que os alunos articulassem vários Registros de Representação, e verificamos através das suas respostas que a maioria deles mostrou saber "montar" e utilizar os registros, "diagrama de árvore" e "tabela de contingência", como ferramenta para responder questionamentos que envolvam a Probabilidade Condicional.

Verificamos no final das aplicações dessas atividades, que estas proporcionaram-nos condições de respondermos às questões formuladas na Problemática. Tais conclusões mostraremos a seguir:

- A maioria dos alunos diferenciava os eventos dependentes dos independentes, tomando como base para isto a interpretação do enunciado e a montagem da "árvore de probabilidades" e a "tabela de contingência";

- ➢ Os alunos, quase que na sua totalidade, aplicavam o conceito da Condicional para problemas que envolvessem o Teorema de Probabilidade Total e o Teorema de Bayes de maneira implícita, sem precisar formalizá-los;

- ➢ Diante das situações que envolviam a condicional, a maioria dos alunos a diferenciava da interseção de eventos, desde que as situações se apresentassem na linguagem natural; e

- ➢ A maioria dos alunos diferenciava o cálculo da probabilidade condicional P(A/B) de P(B/A) desde que esta se apresentasse nas perguntas em linguagem natural.

Estes dois últimos itens, podem ser complementados com alguns aspectos que a nossa pesquisa constatou, são eles:

- ✓ A maioria dos alunos confundem a notação P(A/B) com P(B/A).

- ✓ Existem alunos que confundem a notação P(A∩B) com P(A/B).

No nosso Breve Histórico referido no Capítulo II, registramos que os matemáticos que trabalharam com o conceito da Probabilidade Condicional de 1738 até a nossa época, levaram cerca de três séculos para encontrarem uma simbologia “adequada” para a Condicional, que foi tratada até 1933 na linguagem natural, e parece que os nossos alunos também a preferem. Acreditamos que para afirmarmos qualquer ligação entre as dificuldades apresentadas pelos alunos em interpretar a notação simbólica da Condicional e o desenvolvimento histórico desta representação, é preciso fazer mais investigações sobre o uso de tal notação.

Constatamos, também, que os alunos compreendem melhor uma Probabilidade quando esta se apresenta no registro de porcentagem do que no de fração, embora consigam operar melhor em fração do que em porcentagem.

Acreditamos que este fato deva ser motivo de mais investigações para sabermos quais são as concepções que os alunos têm da representação da Probabilidade através do número na forma decimal ou fracionária.

Outro aspecto que queremos abordar refere-se ao tempo de aplicação da seqüência. Utilizamos seis aulas de quarenta e cinco minutos e este tempo pode parecer muito num primeiro instante para um professor de Estatística, preocupado em vencer o conteúdo programático; porém destacamos que não trabalhamos apenas com o Conceito da Probabilidade Condicional, mas também com a introdução e manipulação dos registros “diagrama de árvore” e “tabela de contingência”. Acreditamos que se os alunos conhecessem melhor esses registros o tempo poderia ser menor.

Como a nossa abordagem facilitou a aprendizagem desses alunos, nos encorajamos a sugerir que as escolas de Ensino Fundamental e Médio sigam as sugestões dos PCNs e das Propostas Curriculares dos Estados de trabalhar com as probabilidades, utilizando a ”árvore de probabilidades” e a “tabela de dupla entrada” no Ensino Fundamental e a Probabilidade Condicional através das “situações conjuntistas” nos “diagramas de árvore” no Ensino Médio. Porém, acreditamos que se fazem necessárias mais pesquisas de como articular os registros utilizados na nossa seqüência e os conteúdos do Ensino Básico, para se trabalhar com outros conceitos ligados à Probabilidade.

O resultado do nosso trabalho confirma que as sugestões dadas por Batanero contribuíram para o desenvolvimento de uma seqüência de ensino a fim de que os alunos conseguissem construir o conceito de Probabilidade Condicional e melhor trabalhassem com os conceitos que o envolvem. Queremos ressaltar que não estamos dando, com isto, uma receita de como introduzir tal conceito, no entanto estamos apontando aspectos que devem ser considerados quando trabalharmos não só com a Probabilidade Condicional, mas com a Estocástica de um modo geral.

Referências Bibliográficas

[1] ARTIGUE, Michèle. Ingeniería Didáctica. Ingeniería Didáctica em Educación Matemática, Grupo Editorial Iberoamérica, p.33 – 59, Bogotá, 1995.

[2] BARRETO, Antonio Rodolfo. Uma Abordagem Histórica do Desenvolvimento da Estatística no Estado de São Paulo . Dissertação de Mestrado. UNESP – Rio Claro, 1999.

[3] BATANERO, Carmen Bernabeu. Didáctica de la Probabilidad y Estadística. Universidade de Campinas, Brasil, 1999.

[4] ______________________________. Teaching Statistcs, v.21, nº 01, Section of de IASE, p. 1- 4, Summer, 1999.

[5] BROUSSEAU, Guy. Fondements et méthodes de la didactique des mathématiques, RDM, vol.7, nº 2 – Ed. La Pensée Sauvage, Grenoble, 1986.

[6] COUTINHO, Cileda Q. S. Introdução ao Conceito de Probabilidade por uma Visão Freqüentista – Estudo Epistemológico e Didático. Dissertação de Mestrado. PUC/SP, São Paulo, 1994.

[7] DAMACENO, José Antonio Elias. Estudo Exploratório das Concepções Probabilísticas Correspondentes aos Níveis de Green. Bolema, Ano 10, nº 11, p. 43 – 61, 1995.

[8] DUVAL, Raymond. L'analyse cognitive du Fonctionnement de la pensée et De l'activité mathématique, PUC/SP, Février, 1999.

[9] FALK, Ruma. A Classic Probability Puzzle, Teaching Statistcs, v. 18, nº 01, p. 17 – 19, Spring 1996.

[10] FISCHBEIN, Efraïm et LECOUTRE, Marie-Paule. Évolution avec l'âge de <<MISCONCEPTIONS>> dans les Intuitions Probabilistes en France et en Israël. Recherches en Didactique des Mathématiques, Vol. 18, nº 3, p. 311 – 332, 1998.

[11] FISCHBEIN, Efraïm, NELLO, Maria Sainati, MARINO, Maria Sciolis. Factors Affecting Probabilistic Judgements in Children and Adolescents. Educational Studies in Mathematics, vol. 22, p. 523 – 549, 1991.

[12] GODINO, J. Dias, BATANERO, M. C., CAÑIZARES, M. J. Azar y Probabilidad. Madrid: Sinteses, 1996.

[13] GRANGÉ, Jean-Pierre. Probabilité conditionnele et indépendance. HENRY, Michel, CHAPUT, Brigitte (coord). Enseigner les probabilités au lycée. IREM de Reims, p.339 – 374, juin,1997.

[14] GRAS, Regis and TOTOHASINA, André. Conceptions d'élèves sur la notion de Probabilité Conditionnelle révélées par une méthode d'Analyse des données: Implication – Similarité – Corrélation. Educational Studies in Mathematics, número28,1995.

[15] HENRY, Michel. L'enseignement du Calcul des Probabilités dans le second degré – Perpective Histories, épistemologiques et didactiques. REPERES – IREM de Besançon, número 14, p. 60 – 104, janvier, 1994.

[16] ________________________. Modélisation en probabilités conditionnelles. HENRY, Michel, CHAPUT, Brigitte (coord). Enseigner les probabilités au lycée. IREM de Reims, p. 93 – 102, juin,1997.

[17] JACOBINI, Otávio Roberto. A Modelação Matemática Aplicada no Ensino de Estatística em Cursos de Graduação. Dissertação de Mestrado. Rio Claro: UNESP, 1999.

[18] LAPLACE, Pierre-Simon. Essai Philosophique sur les Probabilités. (1825) Gauthier-Villais Editeur. Paris, 1921

[19] LOPES, Celi A. E. A Probabilidade e a Estatística no Ensino Fundamental: Uma Análise Curricular. Dissertação de Mestrado. Universidade Estadual de Campinas – Faculdade de Educação, 1998.

[20] MADSEN, Richard W. Secondary Students'Concepts of Probability. USA, Teaching Statistcs, v. 17, nº 03, p. 90 – 93, Autumn 1995.

[21] Matrizes Curriculares de Referência para o Sistema Avaliação da Educação Básica. Brasília: MEC – Instituto Nacional de Estudos e Pesquisa Educacionais, 2ª edição, 1999.

[22] NEWSLETTER OF THE INTERNATIONAL STUDY GROUP FOR RESEARCH ON LEARNING PROBABILITY AND STATISTCS. Spain, v. 12, nº 03, september, 1999. (http://www.urg.es/~batanero/v12se99.htm)

[23] PANAINO, Robinson. Estatística no Ensino Fundamental: Uma Proposta. de Inclusão de Conteúdos Matemáticos. Dissertação de Mestrado. UNESP - Rio Claro, 1998.

[24] Parâmetros Curriculares Nacionais – Matemática (1º e 2º ciclos). Brasília: MEC – Secretaria do Ensino Fundamental, 1998.

[25] Parâmetros Curriculares Nacionais – Matemática (3º e 4º ciclos). Brasília: MEC – Secretaria do Ensino Fundamental, 1998.

[26] Parâmetros Curriculares Nacionais: ensino médio: ciências da natureza, matemática e suas tecnologias. Brasília: MEC – Secretaria de Educação Média e Tecnológica, 1999.

[27] Proposta Curricular Para o Ensino de Matemática 2º grau. São Paulo: Secretaria de Estado da Educação – Coordenadoria de Estudos e Normas Pedagógicas, 1992.

[28] PARZYSZ, Bernard. Des Statistiques aux Probabilites – Exploitons les Arbres. REPERES – IREM de Paris, número 10, p. 91 – 104, janvier. 1993.

[29] ____________________. Utilisation des arbres dans l'enseignement des probabilités. HENRY, Michel, CHAPUT, Brigitte (coord). Enseigner les probabilités au lycée. IREM de Reims, p. 224 – 238, juin,1997.

[30] SHAUGHNESSY, J. M. Research in probability and statistics: reflections and directions. In: GROUWS, D. A. Handbook of Research on Mathematics Teaching and Learning. USA: NCTM, 1992.

[31] TOTOHASINA, André. Méthode implicative en analyse de données et application à l'analyse de conceptions d'étudiants sur la notion de probabilité conditionnelle. Thèse de Doctorat –Rennes I, 1992.

[32] ____________________. L'Introduction du Concept de Probabilite Avantages et Incovenients de Lárborescence. Repéres-IREM, número15. Publicação dos IREM de Rennes, avril 1994.

[33]________________________. Chronologie et Causalité, Conceptions Sources d'obstacles épistémologiques á la notion de Probabilité Conditionnelle. Recherches en Didactique des Mathématiques, Vol. 15, número 1, p.49 – 95,1995.

[34] TOMLINSON, Stephen, QUINN, Robert. Understanding Conditional Probability. USA, Teaching Statistcs, v. 19, nº 01, p. 02 – 07, Spring 1997.

LIVROS DIDÁTICOS

[1] DEGROOT, Morris H. Probability and Statistics. Massachusetts: Addison-Wesley Publishing Company, 1975.

[2] JAMES,Barry R. Probabilidade: um curso em nível intermediário. Rio de Janeiro: Instituto de Matemática Pura e Aplicada, 1981.

[3] MEYER, Paul L. Probabilidade Aplicações à Estatística. Rio de Janeiro: Livros Técnicos e Científicos, 1976.

[4] MORETTIN, Pedro A., BUSSAB, Wilton O. Estatística Básica. 4ª ed. São Paulo: Atual, 1987.

[5] MORETTIN, Luiz Gonzaga. Estatística Básica. São Paulo: MAKRON Books do Brasil, 1999.

[6] SILVA, Elio Medeiros, GONÇALVES, Valter, SILVA, Ermes Medeiros et al. Estatística. São Paulo: Atlas, 1995.

[7] SPIEGEL, Murray Ralph. Estatística. 2ª ed. São Paulo: McGraw-Hill, 1985.

ANEXO I

Leia o problema e responda as questões abaixo.

Um táxi foi envolvido em uma batida, à noite, e fugiu do local do acidente.

Há duas companhias que operam na cidade, a companhia de Táxi Azul e a companhia de Táxi Verde. É sabido que 85% dos táxis da cidade são verdes e 15% são azuis. Uma testemunha da cena identificou o táxi envolvido no acidente como sendo da companhia de Táxi Azul. Esta testemunha foi posta à prova sob condições semelhantes de visibilidade, e fez a identificação correta da cor do carro, tanto do táxi azul como do verde, em 70% dos testes.

a) Qual a probabilidade de a testemunha identificar corretamente a cor de um carro qualquer?

b) Qual a probabilidade de se encontrar um táxi azul na cidade?

c) Qual a probabilidade de se encontrar um táxi verde na cidade?

d) Você estabelece alguma relação entre o acerto da testemunha e a porcentagem de carros azuis ou os verdes da cidade ?

Comente sua resposta.

e) Você sabe dizer qual a probabilidade de a testemunha identificar um táxi como sendo verde, sabendo-se que o táxi é verde?

f) Você sabe dizer qual a probabilidade do táxi ser verde, sabendo-se que a testemunha o identificou como verde?

g) Qual a probabilidade de que o carro envolvido no acidente seja realmente o Táxi Azul e que a testemunha o identifique como azul?

h) Sabendo que a testemunha identificou o carro envolvido no acidente como sendo o azul. Qual a probabilidade de que o carro envolvido no acidente seja um Táxi Azul?

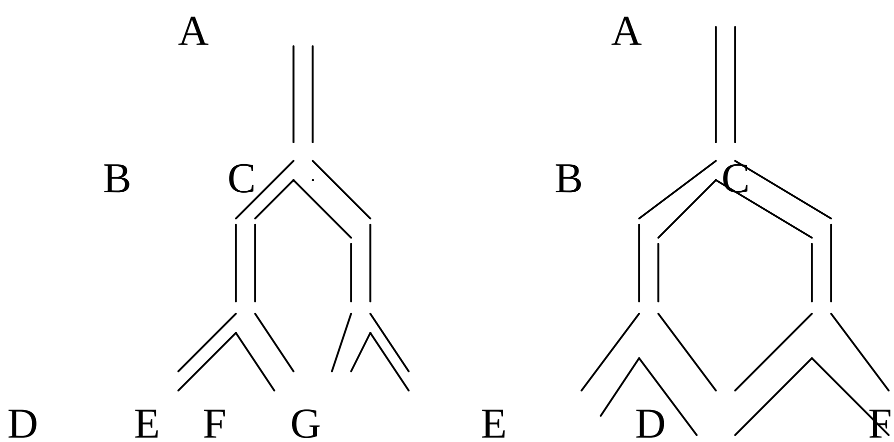
i) Qual a probabilidade de que o carro envolvido no acidente foi realmente um Táxi Azul, sabendo-se que a testemunha o identificou como azul?

j) Se você tivesse que decidir se realmente o táxi azul é o culpado. Que decisão tomaria? Comente sua resposta.

ANEXO II

ATIVIDADE 1

Beto e Meire foram ao Hopi-Hary. Encontraram lá um jogo que envolve duas máquinas. Colocamos o esquema delas abaixo:



Máquina 1 Máquina 2

Em ambas as máquinas, jogando-se uma bola em A, ganha-se um prêmio se a bola cair em D.

a) Em que máquina você pensa que Beto e Meire tenham mais chance de ganhar?
b) Qual e a probabilidade de cair em D na máquina1? E na máquina 2?
c) Observe a máquina 1, e calcule a probabilidade de a bola chegar a F.
d) Ainda na máquina 1, calcule a probabilidade da bola chegar a F, sabendo que ela passou por C.
e) As probabilidades das letras c) e d) são iguais ou diferentes? Por que isto aconteceu?
f) Observe agora a máquina 2, calcule a probabilidade da bola chegar a D.
g) Ainda na máquina 2, calcule a probabilidade da bola chegar a D, sabendo-se que ela passou por C. Em seguida calcule a probabilidade dela chegar a D, sabendo-se que ela passou por B. Essas duas probabilidades são iguais ou diferentes? Justifique.
h) E as probabilidades das letras f) e g) são iguais ou diferentes? Por que você acredita que isso ocorra?

ANEXO III

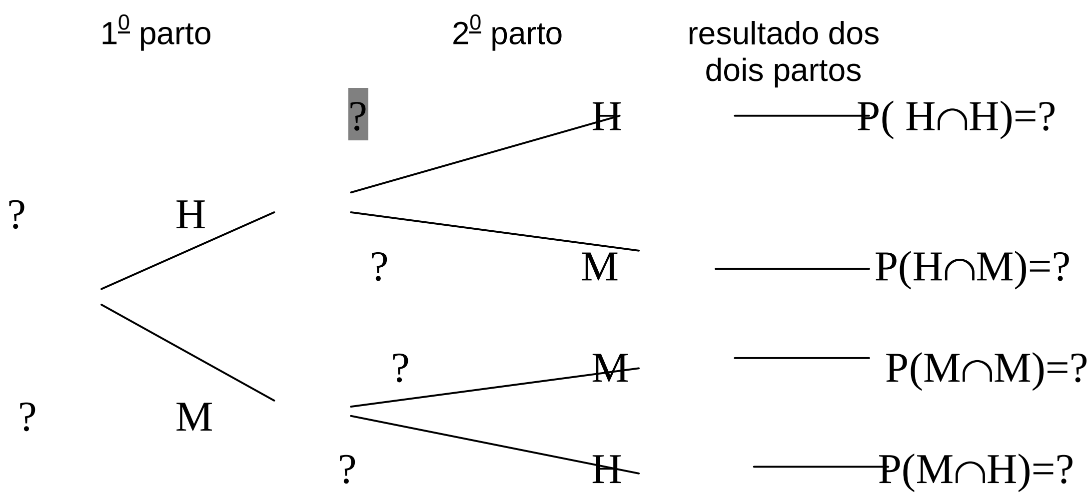
ATIVIDADE 2

Ao estudar o número de natalidade de um certa cidade do Brasil, observamos que a probabilidade de nascer um homem é de 40% e de nascer uma mulher é de 60% .

a) Em uma determinada manhã um médico irá fazer dois partos.
Qual dos eventos abaixo você considera que é o mais provável?

- Os dois recém-nascidos são homens.
- Os dois recém-nascidos são meninas.
- Um recém-nascido é homem e outro mulher.

b) Utilize o seguinte diagrama para responder as questões a seguir:



(Na árvore acima assinalamos uma interrogação (?), que significa a probabilidade de nascer um homem no segundo parto, sabendo ter sido homem no primeiro).

Substitua as interrogações acima pelas suas respectivas probabilidades.

c) Pense novamente na questão a). Você mantêm sua resposta? Justifique sua resposta.

d) Coloque os dados que você escreveu na árvore de probabilidades na tabela a seguir:

1º parto / 2º parto	**H**	**M**	**Total**
H	P(H∩H)=?		
M			
Total			1

e) Suponha que você já saiba que no primeiro parto tenha nascido uma menina. Qual a probabilidade do segundo ser um menino? Explique sua resposta.

f) Qual a probabilidade de nascer uma menino no segundo parto? Como você encontrou essa probabilidade?

g) As probabilidades das letras e e f) são iguais ou diferentes? Como você justifica o fato?

h) Suponha que você chegou atrasado para conferir o sexo da criança que nasceu primeiro. Mas você verificou que a segunda criança era uma menina. Qual a probabilidade do primeiro ter sido um menino?

Justifique a sua resposta.

ANEXO IV

ATIVIDADE 3

Em uma Universidade, estão assistindo uma aula de Estatística os alunos, conforme a tabela abaixo:

CURSO \ SEXO	H	M	Total
LM	5	3	8
CC	8	4	12
Total	13	7	20

H: nº de Homens na sala de aula;
M: nº de mulheres na sala de aula;
LM: nº de alunos do curso de Licenciatura em Matemática; e
CC: nº de alunos do curso Ciência da Computação.

a)"Monte" uma árvore de possibilidades começando pelos eventos Homem (H) e Mulher (M), e calcule todas as probabilidades correspondentes.

b)"Monte" uma árvore de possibilidades começando pelos cursos, ao invés do sexo, em seguida calcule as probabilidades correspondentes.

c)Transforme a sua tabela de informações do enunciado em uma tabela de probabilidades.

d)Sabendo-se que escolhemos um rapaz dessa sala, qual a probabilidade dele ser do curso de Licenciatura em Matemática?

e)Sabendo-se que escolhemos um aluno do curso Licenciatura de Matemática, qual a probabilidade dele ser um Homem?

f) As probabilidades encontradas em d) e e) são iguais ou diferentes? Comente sua resposta.

g) Substitua as interrogações abaixo pelas suas respectivas probabilidades:

III) P(H) = ?
IV) P(M) = ?
V) P(H/LM) = ?
VI) P(M/CC) = ?

h) Tomando como base o item g).
As probabilidades I) e III) são iguais ou diferente? Justifique
As probabilidades II) e IV) são iguais ou diferentes? Justifique

i) Observando a árvore de probabilidade do item a) responda:

- P(H) + P(M) =
- P(LM/H) + P(CC/H) =
- P(LM/M) + P(CC/M) =

- $P(H \cap LM) + P(H \cap CC) + P(M \cap LM) + P(M \cap CC) =$

O que estas somas tem em comum? Justifique.

j) Para calcular P(M/CC) você pensa que seria melhor utilizar a árvore do item a) ou do item b)? Justifique o porquê da sua escolha.

ANEXO V

ATIVIDADE 4

Numa determinada cidade a probabilidade e encontrarmos um indivíduo portador de vírus HIV_+ é de 6%. Em um teste feito num determinado laboratório da cidade, ocorreram os seguintes resultados:

- A probabilidade de um indivíduo ter seu teste positivo (T_+), sabendo-se que é portator do virus HIV_+, é de99,8%.
- A probabilidade de um indivíduo ter seu teste negativo(T_-), sabendo-se que não é portador do virus HIV_+ , é de 99,6%.

São utilizadas as notações::
D para representar o evento do indivíduo ser HIV_+ .
$\overline{D}$ para representar o evento do indivíduo não ser HIV_+ .
A seguir:

a) Construa a árvore de probabilidades partindo dos eventos D e $\overline{D}$.

b)“Monte” a tabela de dupla entrada com as respectivas probabilidades.

c)Calcule a probabilidade de se encontrar T_+ nos exames desse laboratório.

d) Calcule a probabilidade de se encontrar T_- nos exames desse laboratório.

e) Construa a árvore de probabilidades partindo dos eventos T_+ e T_- .

f) Qual a probabilidade de escolhermos um indivíduo ao acaso dessa cidade, sabendo-se que o seu teste deu negativo, que seja portador do vírus HIV_+ .

g) Os eventos D e T_+ são dependentes ou independentes? Comente sua resposta.

h) Responda os itens abaixo justificando como encontraram as suas respostas

I) $P(T_- /D)$ =
II) $P(D/T_+)$ =
III) $P(T_+/D)$=
IV) $P(D/T_-)$ =

i) Como você escreveria na linguagem corrente essas duas probabilidades: $P(D \cap T_+)$ e $P(D/T_+)$. As probabilidades são iguais ou diferentes? Justifique.

C) ANEXO VI

Conteúdo Programático da Disciplina Estatística

01. Estatística Descritiva
1.1. Resumo de Dados
1.2. Tipos de Variáveis
1.3. Distribuição de freqüências
1.4. Representação gráfica das variáveis Quantitativas
1.5. Medidas de posição e dispersão
1.6. Variáveis bidimensionais
1.7. Diagrama de dispersão
1.8. Coeficiente de correlação

02. Probabilidades
2.1. Conceito – Definição
2.2 Probabilidade Condicional e Independência
2.3. Teorema de Bayes

03. Variáveis Aleatórias Discretas
3.1. Esperança Matemática
3.2. Propriedades
3.3. Função de Distribuição Acumulada
3.4. Modelos Probabilísticos (Distribuição de Bernoulli e Binomial)

04. Variáveis Aleatórias Contínuas
4.1. Esperança Matemática
4.2. Função de Distribuição Acumulada
4.3. Modelos Probabilísticos (Distribuição Normal)

05. Variáveis Aleatórias Bidimensionais
5.1. Distribuição conjunta
5.2. Marginal e condicional
5.3. Covariância