UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL

A ESTATÍSTICA E AS PROBABILIDADES NO ENSINO SECUNDÁRIO:

ANÁLISE DOS PROGRAMAS DE MATEMÁTICA A E B NA PERSPECTIVA DO PROFESSOR E DOS ALUNOS

Sara Cristina Baião Caldeira

MESTRADO EM PROBABILIDADES E ESTATÍSTICA

Dissertação orientada pela Prof.ª Doutora
Maria Helena Mouriño Silva Nunes

2009

AGRADECIMENTOS

Apesar da dissertação de Mestrado ser um processo solitário ao qual o investigador está destinado, reúne contributos de várias pessoas. Por isso, este trabalho não ficaria completo sem agradecer a todos os que me ajudaram a concretizá-lo.

Em primeiro lugar, quero agradecer à Professora Doutora Helena Mouriño, orientadora da dissertação, todo o seu saber, a sua ajuda, os seus conselhos e o modo como sempre me apoiou e incentivou, bem como, pela forma interessada e disponível com que sempre acompanhou este trabalho.

Quero agradecer aos alunos que responderam aos inquéritos e que ajudaram a sustentar esta investigação.

Quero também agradecer aos professores e aos Presidentes dos Conselhos Executivos das escolas envolvidas por terem permitido a recolha de dados para este trabalho.

Aos meus colegas de Mestrado que me acompanharam ao longo deste percurso, pelo incentivo e apoio que sempre me deram.

Ao meu namorado, André, que tantas vezes abdicou da minha presença e que sempre me apoiou ao longo da elaboração desta dissertação.

Finalmente, agradeço do fundo do coração aos meus pais todo o amor e a motivação que sempre me deram. Foi graças à paciência que sempre tiveram para acolher os meus momentos de desalento e ao apoio e incentivo em alturas mais críticas que consegui chegar até aqui. É pelos meus pais que sinto a maior alegria ao concluir esta tese e é a eles que a dedico.

R E S U M O

Este trabalho de investigação está centrado em dois objectivos. O primeiro consiste em analisar as unidades de *Estatística*, de *Probabilidades e Combinatória* e de *Modelos de Probabilidade* nos programas de Matemática A e B do Ensino Secundário. O segundo objectivo pretende analisar a opinião dos alunos do 3.ºCiclo e do Ensino Secundário relativamente aos programas de Matemática, aos conteúdos de Estatística e de Probabilidades, bem como à forma como estes foram abordados nas aulas.

Para desenvolver a primeira parte do trabalho foram utilizados diversos documentos, entre eles os programas de Matemática A e B do Ensino Secundário e as brochuras de *Estatística* e de *Probabilidades e Combinatória* que são sugeridas como guias de apoio ao professor na leccionação destas componentes.

Em relação à segunda parte do trabalho, foram elaborados inquéritos como instrumento de recolha de dados. Esta parte do estudo contou com a participação de 1144 alunos do 3.ºCiclo e do Ensino Secundário de diversas escolas da região de Lisboa e Vale do Tejo.

O estudo realizado permitiu concluir que é necessário efectuar uma revisão das temáticas de Estatística e de Probabilidades nos programas de Matemática A e B do Ensino Secundário, bem como uma revisão nas brochuras de *Estatística* e de *Probabilidades e Combinatória*.

Da análise dos inquéritos concluiu-se que a maioria dos alunos do 3.ºCiclo e do Ensino Secundário reconhece que as temáticas de Estatística e de Probabilidades são interessantes, admite que os conteúdos destas temáticas nem são fáceis nem são difíceis e além disso, não considera que a forma como os professores abordaram estes conteúdos seja difícil.

Os resultados obtidos neste estudo evidenciaram a necessidade de ocorrer mudanças de forma a criar condições para o ensino e aprendizagem das temáticas de Estatística e de Probabilidades de acordo com a importância que actualmente lhes é reconhecida.

Palavras-chave: Estatística; Probabilidades; Ensino Secundário; Tabelas de contingência; Inquéritos aos alunos.

ABSTRACT

This work was centered around two objectives. The first consisted of analyzing the units of *Statistics*, *Probabilities and Combinatorics* and *Probability Models* in Mathematics programs A and B of Secondary Education. The second objective was to analyze the opinions of students of the 3rd Cycle and Secondary Education in relation to Mathematics programs, the contents of Statistics and Probabilities, as well as the way in which these themes were approached in the lessons.

To develop the first part of the work several documents were used, including Mathematics programs A and B of Secondary Education and the brochures of *Statistics* and *Probabilities and Combinatorics* which are suggested as guides to support teachers in teaching these components.

In relation to the second part of the work, inquiries were created as an instrument of data collection. In this part of the study participated 1144 students of the 3rd Cycle and Secondary Education of several schools of the region of Lisbon and Vale do Tejo.

This study has led to the conclusion that it is necessary to effect a revision of the thematics of Statistics and Probabilities in Mathematics programs A and B of Secondary Education, as well as a revision of the brochures of *Statistics* and *Probabilities and Combinatorics*.

The analysis of the inquiries led to the conclusion the majority of the students of 3rd Cycle and Secondary Education recognize the thematics of Statistics and Probabilities are interesting, and admit that, on the overall, the contents of these thematics are neither easy, nor difficult and does not consider the method used by teachers to approach this contents is difficult.

The results obtained show the need to for change in order to create better conditions for teaching and learning the thematics of Statistics and Probabilities according to the important role they both play nowadays.

Keywords: Statistics; Probabilities; Secondary Education; Contingency tables; Students inquiries.

ÍNDICE

ÍNDICE DE GRÁFICOS

INTRODUÇÃO

ANÁLISE DA TEMÁTICA DE ESTATÍSTICA

ANÁLISE DA TEMÁTICA DE PROBABILIDADES E COMBINATÓRIA

ANÁLISE DOS DADOS DO 7.º ANO DO ENSINO BÁSICO

ANÁLISE DOS DADOS DO 8.º ANO DO ENSINO BÁSICO

ANÁLISE DOS DADOS DO 9.º ANO DO ENSINO BÁSICO

ANÁLISE DOS DADOS DO 10.º ANO DO ENSINO SECUNDÁRIO

ANÁLISE DOS DADOS DO 12.º ANO DO ENSINO SECUNDÁRIO

ÍNDICE DE TABELAS

ANÁLISE DOS DADOS DO 7.º ANO DO ENSINO BÁSICO

ANÁLISE DOS DADOS DO 8.º ANO DO ENSINO BÁSICO

COMPARAÇÃO ENTRE O 7.º E O 8.ºANOS DO ENSINO BÁSICO

ANÁLISE DOS DADOS DO 9.º ANO DO ENSINO BÁSICO

COMPARAÇÃO ENTRE A ESTATÍSTICA DO 7.º E 8.º ANOS E AS PROBABILIDADES E ESTATÍSTICA DO 9.º ANO

ANÁLISE DOS DADOS DO 10.º ANO DO ENSINO SECUNDÁRIO

ANÁLISE DOS DADOS DO 12.º ANO DO ENSINO SECUNDÁRIO

C A P Í T U L O I

INTRODUÇÃO

1.1. A importância do estudo

As características do mundo actual fizeram com que "a Estatística se tenha tornado uma parte integral das nossas vidas" (Lightner, 1991, p.630). De facto, "basta apenas dar uma olhada nos jornais de hoje para ver até que modo é que a linguagem das Probabilidades e da Estatística se tornou uma parte integral das nossas vidas" (p.629). Desde os jornais, às revistas, às informações acerca do tempo, relatórios económicos e administrativos, sondagens de opinião, quer tenham um carácter político quer se trate de publicidade de produtos de consumo, todos ilustram as informações com gráficos e tabelas dos mais diversos tipos, cuja leitura e interpretação pressupõe alguns conhecimentos estatísticos. Mas não é só fora da escola que as informações são apresentadas utilizando os métodos e a linguagem estatística. Os manuais escolares das várias disciplinas recorrem frequentemente a conceitos estatísticos e utilizam gráficos e tabelas de vários tipos para apresentarem os seus conteúdos.

Atendendo a tudo isto, existe uma necessidade de formação estatística e probabilística para todos, no sentido de promover uma participação activa, crítica e esclarecida por parte de qualquer cidadão em relação a resultados que lhe são apresentados.

Alguns investigadores realçam o papel da educação estatística no desenvolvimento das capacidades de argumentar e tomar decisões. Neste sentido, Carvalho (2001) diz que:

> *"Numa sociedade onde a informação faz cada vez mais parte do dia-a-dia da maioria das crianças, onde grandes quantidades de dados fazem parte da realidade quotidiana das sociedades ocidentais, importa que as crianças, desde logo, consigam coligir, organizar, descrever dados de forma a saberem interpreta--las e, com base nelas, tomarem decisões."* (Carvalho, 2001, p.29-30)

Deste modo, os conhecimentos de Estatística e de Probabilidades, ao possibilitarem o desenvolvimento da capacidade de tomar decisões, "são essenciais quer no exercício da cidadania quer na vida profissional" (Scheaffer, 2000, p.158).

Para além do desenvolvimento das capacidades de argumentar e de tomar decisões, outra das finalidades da educação estatística é o desenvolvimento do espírito crítico, pois um cidadão "estatisticamente competente" não se deixará influenciar de modo errado por resultados divulgados de forma pouco ou nada correcta. Neste sentido Moore (2000) diz que:

> *"Não podemos escapar dos dados, assim como não podemos evitar o uso de palavras. Tal como palavras os dados não se interpretam a si mesmos, mas devem ser lidos com entendimento. Da mesma maneira que um escritor pode dispor as palavras em argumentos convincentes ou frases sem sentido, assim também os dados podem ser convincentes, enganosos ou simplesmente inócuos. A instrução numérica, a capacidade de acompanhar e compreender argumentos baseados em dados, é importante para qualquer um de nós. O estudo da estatística é parte essencial de uma formação sólida."* Moore (2000)

Apesar de cada vez mais se acentuar a importância da Estatística, das Probabilidades e das suas aplicações na sociedade em que vivemos, estas continuam a ser áreas pouco trabalhadas em termos de investigação, comparativamente com outros ramos da Matemática, e predominando as investigações em laboratório em vez de realizadas em situações escolares (Boaventura, 2003). Tornam-se emergentes os estudos que apontam no sentido de se rever a forma como os temas de Estatística e de Probabilidades têm sido trabalhados na sala de aula.

Branco (2000c) considera que "a realização de um inquérito aos professores e alunos parece essencial para identificar quais as maiores dificuldades da aprendizagem e do ensino" (p.16). O autor detalha, dizendo:

> *"(…)porque é que os alunos gostam, ou não gostam, das matérias ensinadas, qual o seu aproveitamento, (…), o que pensam do currículo e do interesse da estatística na formação dos alunos e em geral, são perguntas que podem fornecer respostas de grande utilidade para prosseguir no aperfeiçoamento do ensino"* (Branco, 2000c, p.16)

Dado o escasso número de estudos que incidem sobre estas temáticas de "relevância crescente e de interesse inquestionável" (Almeida, 2002, p.149), este estudo poderá constituir um contributo importante, principalmente para se compreender a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística e de Probabilidades, bem como à forma como estes foram abordados nas aulas.

1.2. Objectivos do estudo

O presente estudo está dividido em duas partes. A primeira parte consiste na análise dos programas de Matemática A e B do Ensino Secundário, centrando essa análise nas temáticas de *Estatística*, de *Probabilidades e Combinatória* e de *Modelos de Probabilidade*. A segunda parte do estudo foca-se na análise da informação recolhida através dos inquéritos aplicados aos alunos do 3.º Ciclo e do Ensino Secundário.

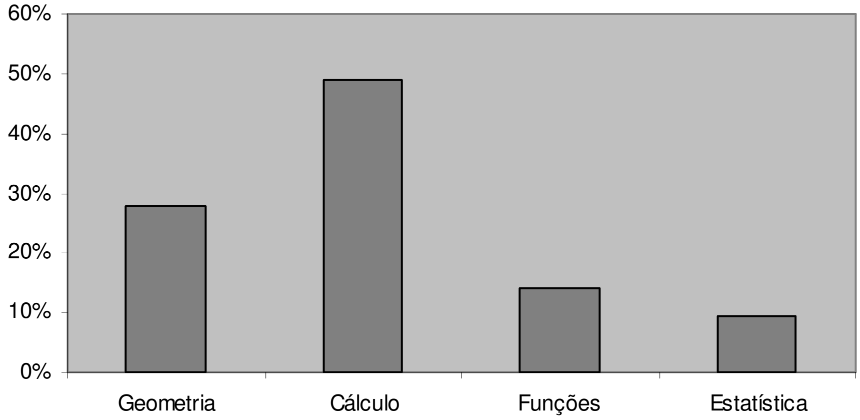
Para se proceder ao desenvolvimento da primeira parte do estudo foram utilizados os programas que se encontram em vigor nos anos lectivos 2007/08 e 2008/09, nomeadamente, o programa de Matemática do 3.º Ciclo do Ensino Básico e os programas de Matemática A e B do Ensino Secundário.

Os programas de Matemática do 3.º Ciclo e do Ensino Secundário regem-se segundo os mesmos princípios. Consideram "o aluno como agente da sua própria aprendizagem", sugerem a utilização da tecnologia (computadores e calculadoras), propõem que se trabalhe a partir de problemas e situações experimentais, que se estabeleça uma maior relação entre a Matemática e o real, que se diversifique as formas de trabalho dentro e fora da sala de aula e também os instrumentos de avaliação.

Ao analisar-se o tempo que os programas atribuem às temáticas de Estatística e de Probabilidades, verifica-se que estas são ainda marginais aos currículos e que são facilmente afastadas para segundo plano.

Os gráficos seguintes tornam evidente o peso relativo que os temas de Estatística e de Probabilidades têm nos programas de Matemática do 3.º Ciclo do Ensino Básico e nos programas de Matemática A e B do Ensino Secundário.

Gráfico 1: Peso relativo dos temas do 7.º ano



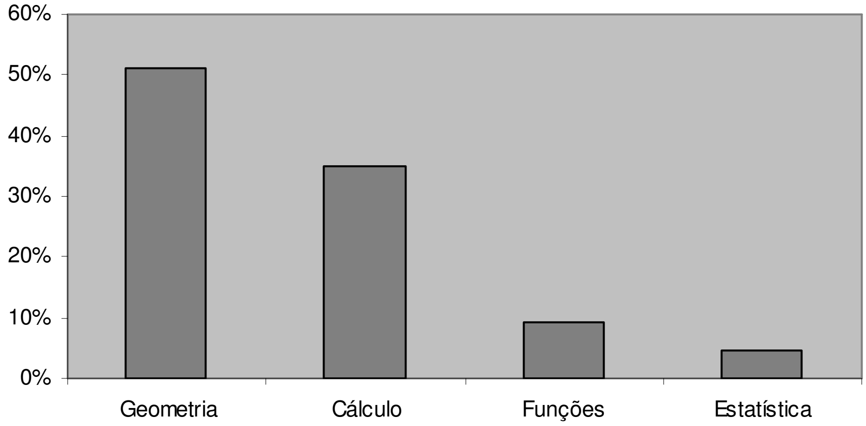
Número de aulas previstas pelo programa:

- Geometria – 24 aulas
- Cálculo – 42 aulas
- Funções – 12 aulas
- Estatística – 8 aulas

Total – 86 aulas

(Retirado do Programa de Matemática do 3.º Ciclo do Ensino Básico: Plano de Organização do Ensino-Aprendizagem, 1991)

Gráfico 2: Peso relativo dos temas do 8.º ano



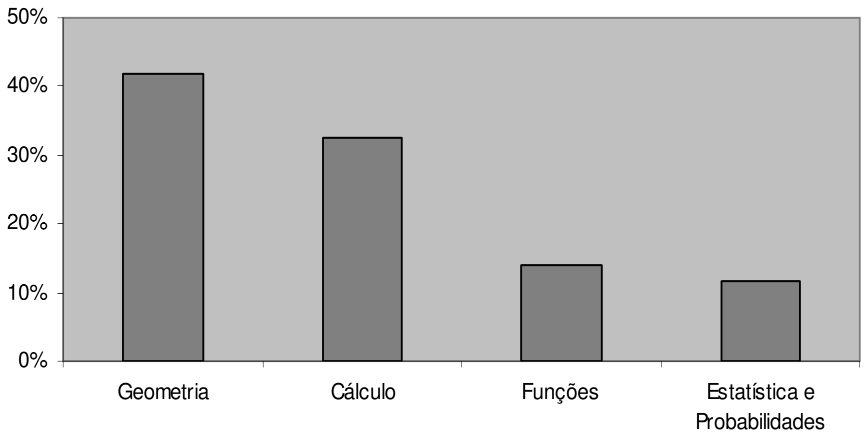
Número de aulas previstas pelo programa:

- Geometria – 44 aulas
- Cálculo – 30 aulas
- Funções – 8 aulas
- Estatística – 4 aulas

Total – 86 aulas

(Retirado do Programa de Matemática do 3.º Ciclo do Ensino Básico: Plano de Organização do Ensino-Aprendizagem, 1991)

Gráfico 3: Peso relativo dos temas do 9.º ano



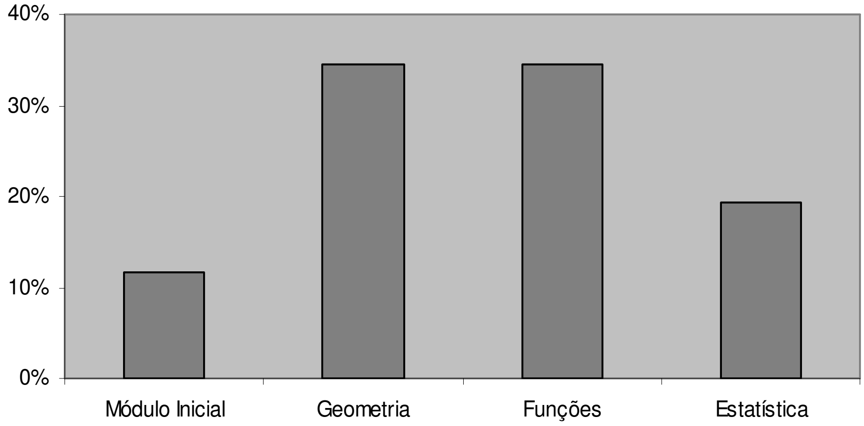
Número de aulas previstas pelo programa:

- Geometria – 36 aulas
- Cálculo – 28 aulas
- Funções – 12 aulas
- Estatística e Probabilidades – 10 aulas

Total – 86 aulas

(Retirado do Programa de Matemática do 3.º Ciclo do Ensino Básico: Plano de Organização do Ensino-Aprendizagem, 1991)

Gráfico 4: Peso relativo dos temas do 10.º ano de Matemática A



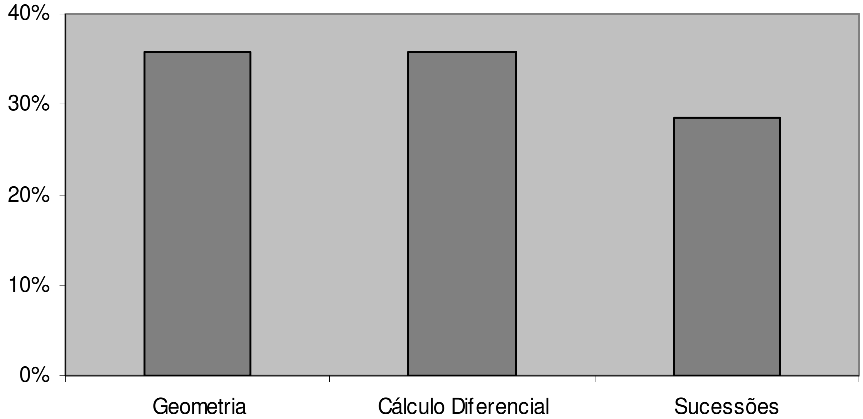
Número de aulas previstas pelo programa:

- Módulo Inicial – 9 aulas
- Geometria – 27 aulas
- Funções – 27 aulas
- Estatística – 15 aulas

Total – 78 aulas

(Retirado do Programa de Matemática A do 10.º ano, 2001a)

Gráfico 5: Peso relativo dos temas do 11.º ano de Matemática A



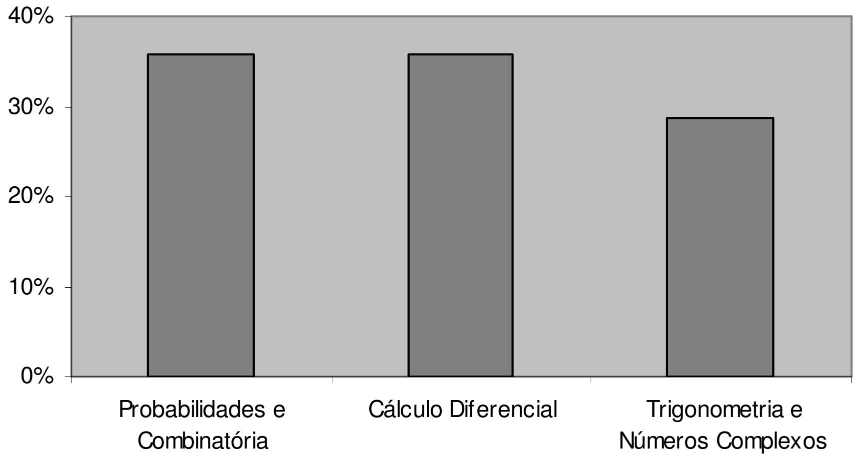
Número de aulas previstas pelo programa:

- Geometria – 30 aulas
- Cálculo Diferencial – 30 aulas
- Sucessões – 24 aulas

Total – 84 aulas

(Retirado do Programa de Matemática A do 11.º ano, 2002b)

Gráfico 6: Peso relativo dos temas do 12.º ano de Matemática A



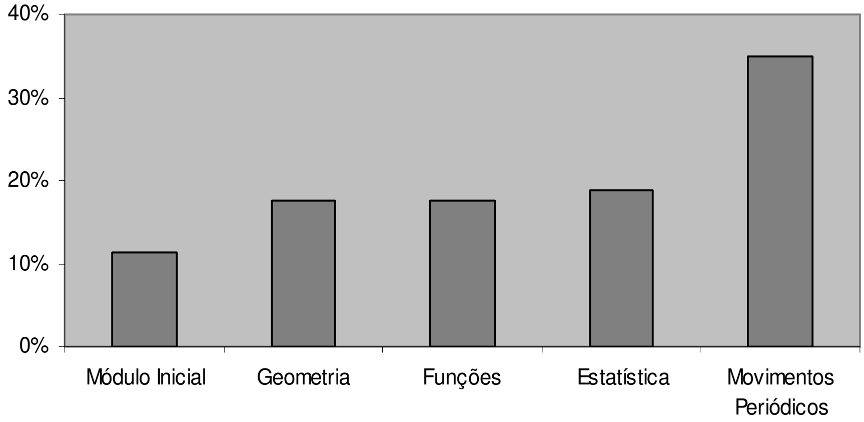
Número de aulas previstas pelo programa:

- Probabilidades e Combinatória – 30 aulas
- Cálculo Diferencial – 30 aulas
- Trigonometria e Números Complexos – 24 aulas

Total – 84 aulas

(Retirado do Programa de Matemática A do 12.º ano, 2002c)

Gráfico 7: Peso relativo dos temas do 10.º ou 11.º ano de Matemática B



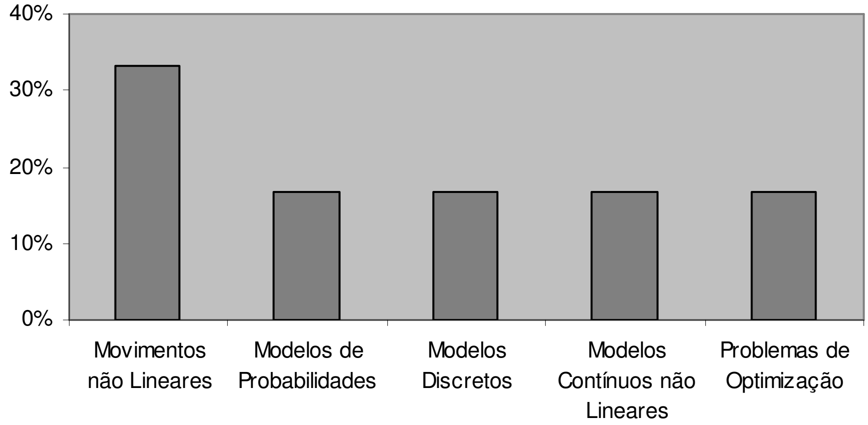
Número de aulas previstas pelo programa:

- Módulo Inicial – 9 aulas
- Geometria – 14 aulas
- Funções – 14 aulas
- Estatística – 15 aulas
- Movimentos Periódicos – 28 aulas

Total – 80 aulas

(Retirado do Programa de Matemática B do 10.º ou 11.º anos, 2001b)

Gráfico 8: Peso relativo dos temas do 11.º ou 12.º ano de Matemática B



Número de aulas previstas pelo programa:

- Movimentos não Lineares – 28 aulas
- Modelos de Probabilidades – 14 aulas
- Modelos Discretos – 14 aulas
- Modelos Contínuos não Lineares – 14 aulas
- Problemas de Optimização – 14 aulas

Total – 84 aulas

(Retirado do Programa de Matemática B do 11.º ou 12.º anos, 2002a)

Para além do tempo que os programas atribuem às temáticas de Estatística e de Probabilidades também alguns professores apontam estes temas como temas a excluir ou a simplificar dos programas de Matemática, tanto a nível do Ensino Básico como do Secundário. No entanto, os conteúdos de Estatística e de Probabilidades estão longe de serem retirados dos programas de Matemática, pois como afirma Carvalho (2001):

> *"(…) o facto de as sociedades regularem cada vez mais a vida dos cidadãos por indicadores numéricos cria a necessidade de que todos eles tenham algum conhecimento que os ajude a compreender o seu significado e, ainda, de como o processo é gerado. Ter conhecimentos de Estatística tornou-se uma inevitabilidade para exercer uma cidadania crítica, reflexiva e participativa, tanto em decisões individuais como colectivas e esta necessidade não é exclusiva dos*

> *adultos uma vez que também as crianças desde cedo estão expostos a dados estatísticos."* (Carvalho, 2001, p.19)

Apesar da extrema importância que os temas de Estatística e de Probabilidades assumem nos dias de hoje, ainda há muito que fazer dada a reduzida atenção que tem sido dispensada a estes tópicos. É necessário ocorrer mudanças, mudanças essas que passam por um lado por dar maior importância às componentes de Estatística e de Probabilidades e por outro lado por uma revisão curricular destas unidades nos programas de Matemática A e B do Ensino Secundário.

É de facto importante efectuar uma revisão curricular das unidades de Estatística e de Probabilidades, pois estas apresentam algumas lacunas que, em alguns casos não chegam a ser muito graves, mas que poderiam ser eliminadas.

Como tal, na primeira parte do estudo é efectuada uma análise crítica destas unidades temáticas onde são apresentados os aspectos mais positivos e menos positivos. Em alguns casos são apresentadas sugestões de modo a melhorar os aspectos menos positivos. Esta análise é também alargada às brochuras de *Estatística* e de *Probabilidades e Combinatória* uma vez que estas são sugeridas pelos programas como guias de apoio aos professores na leccionação destas componentes.

Já no que concerne à segunda parte do estudo é efectuado o tratamento da informação recolhida através dos inquéritos aplicados aos alunos do 3.º Ciclo e do Ensino Secundário. Estes inquéritos têm como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística e de Probabilidades bem como à forma como estes foram abordados nas aulas.

Face à inexistência de estudos que visam questionar os alunos do 3.º Ciclo e do Ensino Secundário com vista a analisar e perceber a opinião destes em relação a estas problemáticas, leva a que este estudo represente um contributo importante para a investigação sobre o ensino e aprendizagem da Estatística e das Probabilidades nas escolas básicas e secundárias.

1.3. Breve sumário do estudo

O presente estudo está estruturado em seis capítulos.

Neste primeiro capítulo – *Introdução* – que termina com a organização geral da tese, realizou-se um enquadramento do estudo, fez-se uma introdução do problema com a justificação da sua importância e apresentou-se os objectivos da investigação.

No segundo capítulo – *Análise da temática de Estatística nos programas de Matemática A e B do Ensino Secundário* – é efectuada uma análise da unidade de Estatística.

No terceiro capítulo – *Análise da temática de Probabilidades e Combinatória no programa de Matemática A do Ensino Secundário* – é apresentada uma análise da unidade de Probabilidades e Combinatória.

No quarto capítulo – *Análise da temática de Modelos de Probabilidade no programa de Matemática B do Ensino Secundário* – é efectuada uma análise da componente de Modelos de Probabilidade.

O quinto capítulo é dedicado aos inquéritos aplicados aos alunos do 3.º Ciclo e do Ensino Secundário. Como tal, é apresentado o modo como foi levada a cabo a operacionalização da investigação, nomeadamente, o método de recolha e o tratamento efectuado aos dados, é desenvolvida a análise dos dados e posteriormente são apresentadas as conclusões obtidas a partir do estudo que são acompanhadas por breves comentários de interesse que ampliam a informação puramente estatística.

Finalmente, no sexto capítulo – *Conclusão* – são apresentadas as principais conclusões do estudo e são feitas algumas recomendações para futuros estudos, tendo em vista contribuir para melhorar e/ou aprofundar o estudo da problemática em causa.

CAPÍTULO II

ANÁLISE DA TEMÁTICA DE ESTATÍSTICA NOS PROGRAMAS DE MATEMÁTICA A e B DO ENSINO SECUNDÁRIO

Esta análise é feita com base no programa de Matemática A do 10.º ano, Ministério da Educação, Departamento do Ensino Secundário, 22 de Fevereiro de 2001; no programa de Matemática B do 10.º ano, Ministério da Educação, Departamento do Ensino Secundário, 22 de Fevereiro de 2001; e na brochura de apoio ao professor de Matemática na leccionação da componente de Estatística do programa do 10.º ano: Estatística – 10.º ano de escolaridade, Ministério da Educação, Departamento do Ensino Secundário, Agosto de 1997.

Organização do Tema de Estatística nos Programas de Matemática A e B do Ensino Secundário

A temática de Estatística dos programas de Matemática A e B do 10.º ano propõe numa primeira parte – ***Estatística-Generalidades*** – o estudo dos objectivos da Estatística enquanto Ciência, da sua evolução histórica, da sua utilidade e dos seus campos de trabalho. A segunda parte – ***Organização e interpretação de caracteres estatísticos*** – está ligada à organização e interpretação de dados bem como às medidas de localização e de dispersão de uma amostra. Finalmente, na terceira parte – ***Referência a distribuições bidimensionais*** – propõe-se uma abordagem gráfica e intuitiva de correlação e de recta de regressão.

Uma vez que a temática de Estatística está estruturada em três partes, a análise que se segue está igualmente organizada nas mesmas três secções.

2.1. Estatística – Generalidades

Na primeira secção dos programas de Matemática A e B surgem os seguintes pontos:

> • Objecto da Estatística e breve nota histórica sobre a evolução desta Ciência; utilidade na vida moderna. Clarificação de quais os fenómenos que podem ser objecto de estudo estatístico; exemplificação de tais fenómenos com situações da vida real, salientando o papel relevante da Estatística na sua descrição.
>
> • Recenseamento e Sondagem. As noções de população e amostra. Compreensão do conceito de amostragem e reconhecimento do seu papel nas conclusões estatísticas; distinção entre os estudos e conclusões sobre a amostra e a correspondente análise sobre a população. Noções intuitivas sobre as escolhas de amostras, sobre a necessidade de serem aleatórias, representativas e livres de vícios de concepção.
>
> • Estatística Descritiva e Estatística Indutiva.
>
> (Silva *et al.*, 2001a, p. 29)
> (Silva *et al.*, 2001b, p. 23)

Ao consultar a brochura de Estatística que visa apoiar o professor de Matemática na leccionação da componente de Estatística (Martins, M. E. G. (coord.), Monteiro, C., Viana, J. P. e Turkman, M. A. A. (1997). *Estatística: matemática – 10.º ano de escolaridade*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário), constatei que os autores desta brochura não fizeram qualquer tipo de referência ao primeiro ponto dos programas. Mas, se esta brochura tem por objectivo apoiar o professor de Matemática na leccionação da componente de Estatística, não deveria focar todos os conteúdos referidos nos programas?

2.1.1. Recenseamento e Sondagem. População e Amostra

No que se refere ao recenseamento e à sondagem, os autores da brochura de Estatística referiram que estes conceitos são *"suficientemente motivadores para o Professor introduzir os conceitos mais gerais de população e amostra"* (p.11). Mas, se os conceitos de recenseamento e de sondagem motivam a introdução dos conceitos de população e amostra, por que motivo os autores recorreram aos termos de população e de amostra, ainda antes destes terem sido definidos para apresentarem o conceito de sondagem? Na brochura de Estatística encontra-se o seguinte:

> *"Sondagem - Estudo científico de uma parte de uma população com o objectivo de estudar atitudes, hábitos e preferências da população relativamente a acontecimentos, circunstâncias e assuntos de interesse comum."* (p. 12)

> *"(…) as sondagens inquirem ou analisam apenas uma parte da população em estudo, isto é, restringem-se a uma amostra dessa população, mas com o objectivo de extrapolar para todos os elementos da população os resultados observados na amostra."* (p. 13)

> ***"Sondagem*** *- Estudo estatístico de uma população, feito através de uma amostra, destinado a estudar uma ou mais das suas características tal como elas se apresentam nessa população.*" (p. 14)

Para além do conceito de sondagem ter sido definido à custa dos termos de população e de amostra, o conceito de recenseamento está pouco explícito. Inclusivamente, tenho algumas dúvidas de que um professor, com pouca formação em Estatística, consiga compreender e principalmente transmitir aos seus alunos a seguinte definição:

> *"**Recenseamento** - Estudo científico de um universo de pessoas, instituições ou objectos físicos com o propósito de adquirir conhecimentos, observando todos os seus elementos, e fazer juízos quantitativos acerca de características importantes desse universo."* (Brochura de Estatística, p. 11)

Depois de ter sido definido o conceito de recenseamento, mas antes ainda de ser apresentado o termo de população, os autores da brochura referiram o seguinte:

> *"É importante que fique claro que a palavra recenseamento está associada à análise de todos os elementos da população em causa (…). Não é contudo viável nem desejável, principalmente quando o número dos elementos da população é muito elevado, inquirir todos os elementos da população sempre que se quer estudar uma ou mais características particulares dessa população."* (p. 12)

Mas, afinal, o que é a população? O que são os elementos da população? Como é que sabemos se o número de elementos da população é muito elevado? A partir de quando é que esse número é elevado?

De facto, não é fácil definir os conceitos de sondagem e de recenseamento sem se recorrer aos termos de população e de amostra. Como tal, penso que teria sido mais correcto e menos confuso se os autores da brochura tivessem começado por abordar os conceitos de população e de amostra e só depois os termos de sondagem e de recenseamento. Mas, para isso, também seria necessário haver uma reestruturação nos programas de Matemática A e B de forma a indicar aos professores de Matemática que os conceitos de população e de amostra devem ser os primeiros conceitos a serem abordados.

Na brochura de Estatística encontra-se a seguinte definição para população:

> *"**População** - colecção de unidades individuais, que podem ser pessoas, animais, resultados experimentais, com uma ou mais características em comum, que se pretendem analisar."* (p. 15)

Como a "Estatística é a Ciência que trata dos "dados"" (Silva *et al.*, 2001a, p. 30; Silva *et al.*, 2001b, p. 24), penso que os autores da brochura não deveriam ter falado em pessoas nem em animais quando definiram o conceito de população. Como tal, penso que seria mais correcto apresentar o seguinte:

> Em Estatística, **população** designa um conjunto de elementos com uma ou mais características em comum.

Depois de ter sido apresentada a definição de população, os autores da brochura referiram que existem várias razões *"que podem levar a que não se possa observar exaustivamente todos os elementos de uma população"* (p.16). Entre elas, destaca-se "*o facto de algumas populações terem dimensão infinita."* (p. 16)

Mas, se *a priori* não for apresentado o conceito de dimensão de uma população, como é que um professor irá compreender e principalmente explicar aos seus alunos que um dos motivos que leva a que não sejam observados todos os elementos de uma população é o facto de essa população ter dimensão infinita?

Na minha opinião, seria importante que os autores da brochura referissem o seguinte:

> *"A população pode ser finita ou infinita, dependendo do número de elementos que a compõe ser finito ou infinito."* (Reis, 1996, p. 43)

Exemplo: O conjunto de indivíduos nascidos ontem na cidade de Setúbal constitui uma população finita.

Exemplo: O conjunto de valores da pressão atmosférica nos diversos pontos do planeta num determinado instante constitui uma população infinita.

No que diz respeito ao conceito de amostra, os autores da brochura de Estatística apresentaram a seguinte definição:

*"**Amostra** - subconjunto da população, que se observa com o objectivo de tirar conclusões para a população de onde foi recolhida."* (p.16)

2.1.2. Amostras Aleatórias e Representativas da População

Depois de ter sido apresentada a definição de amostra, que na minha opinião foi apresentada de uma forma bastante simples e clara, os autores referiram o seguinte:

"(…) a amostra deve ser tão representativa quanto possível da população." (p. 16)

Mas, quando é que uma amostra é representativa da população? O que significa ser representativa da população?

Antes dos autores da brochura apresentarem uma resposta para estas questões, lê-se na brochura que:

"Quando uma amostra não é representativa da população, diz-se que é enviesada." (p.17)

Mas, do que adianta a um professor ou a um aluno saber a designação de uma amostra que não seja representativa da população, se não sabe quando é que uma amostra é, ou não, representativa da população?

Posteriormente, na brochura de Estatística encontra-se a seguinte frase:

"A selecção de uma amostra representativa da população a estudar é um problema que nem sempre é simples de abordar, mas existe um princípio que deve estar presente que é o da aleatoriedade." (p. 17)

Mas, o que é o princípio da aleatoriedade? Será que todos os alunos que estão no 10.º ano sabem o conceito de aleatoriedade? Sabe-se que há alunos que estão no 10.º ano e que nunca tiveram contacto com conceitos de Probabilidades, pois é do conhecimento de todos que a unidade de Probabilidades e Estatística, que faz parte do currículo do 9.º ano, é geralmente remetida para o terceiro período, ficando muitas vezes por leccionar.

Na minha opinião, seria boa ideia recordar o significado do termo aleatório.

> *"O termo "aleatório" significa "dependente do acaso", e é antónimo do termo "determinista", que significa que há um conhecimento perfeito do fenómeno, no sentido em que conhecendo as suas condições iniciais conhecemos a sua evolução. Muitos fenómenos físicos macroscópicos são considerados, à escala a que os observamos, deterministas, e é esse determinismo que nos permite prever marés, horas do nascer e do pôr do Sol, eclipses da Lua, etc. (…) Mas a maior parte dos fenómenos tem um lado determinista e um lado aleatório (…). A hora a que acordamos, a quantidade de água que bebemos por dia, o dinheiro que gastamos por semana – de um modo geral são pequenas flutuações aleatórias em torno de valores não aleatórios."* (Pestana e Velosa, 2006, p. 84)

Imediatamente depois dos autores da brochura de Estatística terem referido o princípio da aleatoriedade, apresentaram a seguinte definição:

> *"Dada uma população, uma amostra aleatória é uma amostra tal que, qualquer outra amostra possível, da mesma dimensão, tem igual possibilidade de ser seleccionada."* (p. 17)

No entanto, será que um professor que tenha pouca formação em Estatística, quando consultar a brochura que foi feita especialmente para apoiar o professor, ficará esclarecido no que diz respeito ao conceito de amostra aleatória?

Na minha opinião, seria mais esclarecedor para o professor se o conceito de amostra aleatória tivesse sido apresentado da seguinte forma:

> Uma **amostra aleatória** de dimensão n de uma população X é um conjunto de n elementos tirados de X de tal modo que, em cada tiragem, os elementos presentes na população X têm a mesma probabilidade de serem escolhidos.

Note-se que o número de elementos que constituem a amostra designa-se por **dimensão da amostra**.

Além disso, note-se que "As amostras «representativas» são *amostras aleatórias*." (Mello, 1997, p. 14)

Os autores da brochura de Estatística, para além de terem apresentado um exemplo que envolve amostras aleatórias simples, apresentaram um exemplo em que a amostra era estratificada e deram, ainda, uma ideia intuitiva de amostragem sistemática.

Tanto a amostragem estratificada como a amostragem sistemática não fazem parte dos programas de Matemática do 10.º ano, no entanto, penso que é importante, na brochura de apoio ao professor, aprofundar os assuntos um pouco mais do que aquilo que os programas sugerem, de modo a que o professor possa planificar e desenvolver as actividades de aprendizagem com mais facilidade e flexibilidade.

Os autores da brochura apresentaram ainda uma actividade que explica como se obtém amostras aleatórias simples com o auxílio da calculadora. No entanto, é do conhecimento de todos que há muitos alunos a frequentarem o 10.º ano sem calculadoras (gráficas ou científicas) e que também há muitas escolas que não têm calculadoras (gráficas ou científicas) para facultar aos alunos.

Para finalizar esta secção, os autores da brochura levantaram a seguinte questão: *"**Qual a dimensão que se deve considerar para a amostra?**"* (p. 22)

Na minha opinião, os autores da brochura, antes de procurarem responder a esta questão, deveriam ter apresentado a definição de *dimensão de uma amostra.* No entanto, nunca o fizeram. Além disso, para responderem à questão mencionaram diversos conceitos, tais como, *variabilidade da população*, *estimativas dos parâmetros* e *variável*, sem que estes tenham sido previamente definidos. Desta forma, será que um professor com pouca formação em Estatística conseguirá responder à questão colocada pelos autores?

2.1.3. Estatística Descritiva e Estatística Indutiva

No que se refere à Estatística Descritiva e à Estatística Indutiva, os autores da brochura, para além de terem fornecido apenas uma ideia intuitiva destes termos, não foram totalmente ao encontro das Indicações Metodológicas sugeridas pelos programas de Matemática A e B do 10.º ano.

Nas Indicações Metodológicas sugeridas pelos programas de Matemática A e B, lê-se o seguinte:

> *"(…) num procedimento estatístico estão envolvidas, de um modo geral, duas fases: uma fase de organização dos dados recolhidos, em que se procura reduzir, de forma adequada, a informação neles contida –* ***Estatística Descritiva****, e uma segunda fase, em que se procura tirar conclusões e tomar decisões para um conjunto mais vasto, de onde se recolheram os dados -* ***Inferência Estatística****. Existe, no entanto, uma fase pioneira, que diz respeito à aquisição dos próprios "dados"."* (Silva *et al.*, 2001a, p. 30; Silva *et al.*, 2001b, p. 24)

Ao passo que na brochura de Estatística, lê-se o seguinte:

> *"Uma vez recolhida a amostra procede-se ao seu estudo. Este consiste em resumir a informação contida na amostra construindo tabelas, gráficos e calculando algumas características amostrais (estatísticas). Este estudo descritivo dos dados é o objectivo da Estatística Descritiva. No entanto, ao estudar a amostra tem-se, normalmente, como objectivo final inferir para a população as propriedades estudadas na amostra. Assim o objectivo do estudo estatístico pode ser o de estimar uma quantidade ou testar uma hipótese, utilizando-se técnicas estatísticas convenientes, as quais realçam toda a potencialidade da Estatística, na medida em que vão permitir tirar conclusões acerca de uma população, baseando-se numa pequena amostra, dando-nos ainda uma medida do erro cometido." (p. 26)*

As Indicações Metodológicas fornecidas pelos programas de Matemática A e B referem diversas vezes a palavra "dados", ao passo que, na brochura de apoio ao professor, esta palavra foi substituída por "amostra". Mas, se a brochura de Estatística foi elaborada para apoiar os professores, não teria sido importante que os autores da brochura tivessem tido o cuidado de utilizar os mesmos termos que os autores dos programas?

Para além do que acabei de referir, no excerto anterior pode-se ler que *"(…) o objectivo do estudo estatístico pode ser o de estimar uma quantidade ou testar uma hipótese, (…), na medida em que vão permitir tirar conclusões acerca de uma população, baseando-se numa pequena amostra"* (p. 26) [o sublinhado é meu]. No entanto, quando se recolhe uma amostra com o intuito de inferir para a população, a amostra não pode ser demasiado pequena, pois deixa de ser representativa da população. Portanto, na minha opinião, os autores da brochura de Estatística deveriam ter redigido a frase anterior da seguinte forma: *o objectivo do estudo estatístico pode ser o de estimar uma quantidade ou testar uma hipótese (…) na medida em que vão permitir tirar conclusões acerca de uma população, baseando-se numa amostra representativa da mesma.*

Depois de ter sido apresentada uma ideia intuitiva de Estatística Descritiva e de Estatística Indutiva, os autores da brochura apresentaram um exemplo para ilustrar melhor estes dois conceitos. Na minha opinião, a parte do exemplo que se refere à Estatística Descritiva está bastante clara e perceptível. No entanto, a parte que se refere à Estatística Indutiva está pouco explícita. Passo a citar um excerto do exemplo:

> *"Poderemos agora inferir que 63.5% dos eleitores da cidade do Porto pensam votar no Senhor X? A resposta a esta pergunta nem é sim, nem não, mas talvez." (p. 27)*

Será que esta é uma resposta desejável para os professores que procuram na brochura algum apoio? Se continuar a ler o excerto, anteriormente citado, encontro o seguinte:

> *"(…) o intervalo [60.5%, 66.5%] contém o valor exacto para a percentagem de eleitores da cidade que pensam votar no Senhor X, com uma confiança de 95%." (p. 27)*

Será que os professores que procuram na brochura algum apoio já ficaram esclarecidos? Suspeitando que existirão muitos professores de Matemática que não sabem o que significa "uma confiança de 95%", os autores da brochura de Estatística escreveram a seguinte nota:

> ***"Nota** - A confiança de 95% deve ser entendida no seguinte sentido: se se recolherem 100 amostras, cada uma de dimensão 1000, então poderemos construir 100 intervalos; destes 100 intervalos esperamos que 95 contenham o*

verdadeiro valor da percentagem (desconhecida) de eleitores da cidade do Porto, que pensam votar no candidato." (p. 27)

Uma vez que os intervalos de confiança podem ser construídos com uma confiança diferente de 95%, penso que os autores da brochura poderiam ter generalizado a nota anterior:

No Caso Geral: A confiança de 100x(1-α)% deve ser interpretada da seguinte forma: ao recolher 100 amostras da mesma dimensão e ao calcular os respectivos intervalos de confiança, espera-se que 100x(1-α) desses intervalos contenham o verdadeiro valor do parâmetro desconhecido.

Posteriormente, os autores da brochura de Estatística chamaram a atenção para o seguinte:

"Muitas das inferências feitas são imperfeitas, a maior parte das vezes por terem como base dados imperfeitos." (p. 28)

Mas, o que são dados imperfeitos? E se não tivermos dados imperfeitos, as inferências ficam perfeitas? Para além destas questões, por que motivo os autores passaram a falar em dados em vez de amostra?

2.1.4. Exemplo de Aplicação da Estatística

Os autores da brochura de Estatística apresentaram alguns exemplos de aplicação da Estatística. Um dos exemplos apresentados é referente ao campo da Medicina, cujo problema foi apresentado da seguinte forma:

"Problema - pretende-se, a partir dos resultados obtidos, realizar um teste de hipóteses para tomar uma decisão sobre qual dos medicamentos é melhor." (p.30)

Uma vez que os *testes de hipóteses* não fazem parte dos programas de Matemática do 10.º ano, penso que este problema deveria ter sido estruturado da seguinte forma:

Problema – pretende-se, a partir dos resultados obtidos, verificar se existe evidência estatística para afirmar que o novo medicamento é melhor.

2.2. Organização e Interpretação de Caracteres Estatísticos (Qualitativos e Quantitativos)

Na segunda secção dos programas de Matemática A e B surgem os seguintes pontos:

> • Análise gráfica de atributos qualitativos (gráficos circulares, diagramas de barras, pictogramas); determinação da moda.
> • Análise de atributos quantitativos: variável discreta e variável contínua. Dados agrupados em classes.
> • Variável discreta; função cumulativa.
> • Variável contínua: tabelas de frequências (absolutas, relativas e relativas acumuladas); gráficos (histograma, polígono de frequências); função cumulativa.
> • Medidas de localização de uma amostra: moda ou classe modal; média; mediana; quartis.
> • Medidas de dispersão de uma amostra: amplitude; variância; desvio padrão; amplitude interquartis.
> • Discussão das limitações destas estatísticas.
> • Diagramas de ”extremos e quartis”.
>
> (Silva *et al.*, 2001a, p. 30 e 31)
> (Silva *et al.,* 2001b, p. 24)

2.2.1. Atributos *vs.* Dados. Atributos Qualitativos e Quantitativos. Dados Qualitativos e Quantitativos

Os dois primeiros pontos dos programas fazem referência a atributos qualitativos e a atributos quantitativos. Mas o que são atributos qualitativos? E atributos quantitativos?

Ao consultar a brochura de Estatística, que tem como objectivo apoiar o professor de Matemática, não encontrei resposta a estas questões. Os autores não fizeram, ao longo de toda a brochura, qualquer tipo de referência a estes dois termos. Eles optaram por fazer referência a *dados qualitativos* e a *dados quantitativos*. Mas, será que *dados* e *atributos* têm o mesmo significado?

Na minha opinião, é importante clarificar estes dois conceitos:

> **Atributos** ou **características de uma população** é o que se pretende conhecer quando se estuda uma população ou uma amostra dessa população.
>
> Chama-se **dados** ao resultado da observação de certo(s) atributo(s).

Uma vez que os programas de Matemática A e B fazem a separação entre os atributos qualitativos e os atributos quantitativos, penso que teria sido importante que os autores da brochura de Estatística tivessem apresentado a definição destes dois conceitos:

> *"As características ou atributos dos elementos de uma população podem ser qualitativos ou quantitativos. Uma característica diz-se qualitativa se as suas diferentes modalidades não são mensuráveis, isto é, não podem ser expressas através de números. Se pelo contrário, as modalidades de um atributo são mensuráveis então trata-se de um atributo quantitativo."* (Reis, 1996, p. 45)

Depois de terem sido apresentados os conceitos de atributos qualitativos e quantitativos, é que os autores da brochura deveriam ter enunciado os conceitos que passo a citar:

> ***"Dados qualitativos** - Representam a informação que identifica alguma qualidade, categoria ou característica, não susceptível de medida, mas de classificação, assumindo várias modalidades."* (Brochura de Estatística, 1997, p. 32)
>
> ***"Dados quantitativos** - Representam a informação resultante de características susceptíveis de serem medidas, apresentando-se com diferentes intensidades, que podem ser de natureza **discreta** - dados discretos, ou **contínua** - dados contínuos."* (Brochura de Estatística, 1997, p. 34)

2.2.2. Variável Discreta e Variável Contínua

O segundo ponto dos programas, para além de fazer referência a atributos quantitativos, faz também referência a variáveis discretas e contínuas.

Ao consultar a brochura de Estatística encontrei a seguinte definição de **variável**:

> *"A uma característica comum que possa assumir valores ou modalidades diferentes, de indivíduo para indivíduo, chamamos variável."* (p. 32)

Mas, será que o conceito de variável só se aplica a indivíduos? Será que não se pode aplicar a objectos ou a animais? Na minha opinião, seria mais correcto apresentar o conceito de variável numa das seguintes formas:

> *"Uma variável estatística é uma característica que pode ser diferente nas diversas observações feitas, e pode ser de natureza qualitativa ou quantitativa" (Pestana e Velosa, 2006, p. 52)*

> **Variável** é todo o símbolo (X, Y, Z, …) que representa determinada característica de uma população ou amostra de uma população e que numa dada questão pode tomar qualquer valor ou categoria pertencente a um conjunto previamente fixado (ou conhecido).

Se continuar a consultar a brochura de Estatística encontro a distinção entre variável discreta e contínua.

> *"Uma variável é **discreta** se só pode tomar um nº finito (ou infinito numerável) de valores distintos."* (p. 34)

> *"No caso de uma variável **contínua**, esta pode tomar todos os valores numéricos, compreendidos no seu intervalo de variação."* (p. 35)

Se um professor de Matemática apresentar aos seus alunos a definição de variável discreta, exactamente igual à que é apresentada na brochura de apoio ao professor, os alunos ficam sem compreender o conceito, pois ainda não sabem o que é um número infinito numerável. Além disso, a definição de variável contínua está também pouco explícita para alunos que frequentam o 10.º ano de escolaridade.

Na minha opinião, seria mais correcto apresentar os conceitos de variável discreta e contínua da seguinte forma:

> *"Uma variável X diz-se discreta se os seus valores admissíveis se podem seriar na forma $x_1 < x_2 < \ldots < x_i < \ldots$ de tal modo que entre quaisquer dois valores consecutivos desta cadeia a variável não pode tomar nenhum outro valor."* (Mello, 1993, p. 16)
>
> Exemplo: Número de irmãos; número de acidentes, por dia, num determinado cruzamento.

> "*Uma variável X é contínua se, dados dois valores quaisquer do seu domínio, x' e x'', existem sempre valores admissíveis x tais que $x' < x < x''$.*" (Mello, 1993, p.17)
>
> Exemplo: peso; altura; tensão arterial; nível de colesterol; …

Os autores da brochura de Estatística, depois de terem apresentado a definição de variável discreta e de variável contínua, fizeram uma chamada de atenção para o facto de que a classificação de uma variável, em discreta ou contínua, ser, por vezes, susceptível de algumas dúvidas e, como tal, apresentaram o exemplo que passo a citar: "*(…) a variável idade, ao contrário do que possa parecer à primeira vista, já que só utilizamos números inteiros para a representar, é uma variável contínua (…)*" (p.35) [o sublinhado é meu]. Mas, será que há uma variável idade? Ou será que os autores da brochura quereriam dizer: *a variável que representa a idade*?

É muito importante saber distinguir as variáveis discretas das contínuas. Esta distinção é absolutamente necessária uma vez que não se podem aplicar os mesmos instrumentos estatísticos aos dois tipos de variáveis.

2.2.3. Tabela de Frequências. Frequência Absoluta e Relativa

Quando se tem um conjunto de dados, resultante da observação de todos os elementos da população ou de apenas alguns elementos da população, o primeiro procedimento estatístico que se deve fazer é organizar os dados numa **tabela de frequências**.

Os autores da brochura de Estatística referiram o seguinte:

> *"Dado um conjunto de dados, estes são organizados na forma de uma tabela de frequências, que apresenta o número de elementos - frequência absoluta (ou só frequência) de cada uma das modalidades ou classes, que os dados assumem."* (p. 32)

Na minha opinião, seria menos confuso, tanto para um professor com pouca formação em Estatística como para os alunos, se estes conceitos fossem apresentados do seguinte modo:

> A **tabela de frequências** é construída da seguinte forma: numa coluna colocam--se todos os valores ou modalidades que a variável apresenta e na segunda

coluna o número de ocorrências correspondentes a cada valor ou a cada modalidade da variável.

A **frequência absoluta** é o número de vezes que cada valor ou modalidade da variável se repete na amostra ou na população.

Depois de terem definido o conceito de frequência absoluta, os autores da brochura de Estatística referiram o seguinte:

"Numa tabela de frequências, além das frequências absolutas, também se apresentam as frequências relativas, onde

$$frequência \quad relativa = \frac{frequência \quad absolua}{dimensão \quad da \quad amostra}$$

entendendo-se por dimensão da amostra o número de elementos da amostra."

(p.32)

A fórmula anterior só é válida no caso de estarmos a estudar os elementos de uma amostra. Então, e se quisermos estudar todos os elementos de uma população, que fórmula devemos utilizar? Na minha opinião, os autores da brochura de Estatística deveriam ter referido que, quando se pretende estudar todos os elementos da população, a fórmula que nos permite calcular a frequência relativa é a seguinte:

$$frequência \quad relativa = \frac{frequência \quad absolua}{dimensão \quad da \quad população}$$

entendendo-se por *dimensão da população* o número de elementos da população.

Na minha opinião, os autores da brochura de Estatística, para além de terem apresentado a fórmula da frequência relativa, deveriam também ter apresentado o conceito de frequência relativa:

A **frequência relativa** de um valor ou de uma modalidade da variável é o número de vezes que esse valor ou que essa modalidade ocorre relativamente ao número total de elementos da amostra ou da população.

Posteriormente, os autores da brochura apresentaram um exemplo para ilustrar as tabelas de frequências. Acontece que a amostra seleccionada para este exemplo não é representativa da população, pois as pessoas que convivem diariamente e que têm a mesma cultura, que é o caso das pessoas que vivem numa determinada aldeia, acabam por ver os mesmos programas televisivos e, inclusivamente, as mesmas telenovelas. Passo então a citar o exemplo que acabei de referir:

"Perguntou-se a cada um dos 100 habitantes de uma determinada aldeia, qual a telenovela preferida, do seguinte conjunto:

***CI** – Cinzas* ***PP** - Pedra sobre Pedra* ***CA**- Corpo e Alma*

***MP** - Mico Preto* ***BA** - Barriga de Aluguer* ***PL** - Plumas e Lantejoulas*

Obtiveram-se os seguintes resultados (Obviamente que ninguém respondeu " Não gosto de nenhuma" …):

Classes	*Freq. abs.*	*Freq. rel.*
CI	*11*	*0.11*
PP	*31*	*0.31*
BA	*8*	*0.08*
CA	*21*	*0.21*
PL	*13*	*0.13*
MP	*16*	*0.16*
Total	*100*	*1.00*

A redução dos dados anteriores segundo uma tabela de frequências permite concluir imediatamente que:

A novela preferida por mais pessoas é a Pedra sobre Pedra

A novela preferida por menos pessoas é a Barriga de Aluguer"

(p. 32 e 33)

Ainda em relação ao exemplo anterior, penso que os autores não deveriam ter utilizado, na tabela de frequências, a palavra *"classes"*. Esta poderia ter sido substituída por *"categorias"* ou *"telenovela preferida"*.

2.2.4. Dados Agrupados em Classes

Ao consultar a brochura de Estatística encontro a seguinte citação:

"(…) no caso de dados discretos, a construção da tabela é análoga à que foi feita para os dados qualitativos, mas em vez das categorias consideram-se os valores distintos que surgem na amostra, os quais vão constituir as classes." (p.35)

Mas, será que faz sentido construir classes apenas com um único valor? Será que as classes não deveriam apenas resultar do agrupamento dos dados?

Posteriormente, os autores da brochura de Estatística referiram o seguinte:

> *"De um modo geral, as classes vão ser intervalos fechados à esquerda e abertos à direita, todos eles com a mesma amplitude."* (p. 37)

Será que um professor com pouca formação em Estatística, depois de ler os excertos anteriores, irá conseguir compreender a noção de classe?

2.2.5. Diagrama de Barras

No que se refere à representação gráfica de dados discretos, os autores da brochura de Estatística sugeriram que *"Uma representação gráfica adequada para estes dados, é o diagrama de barras."* (p. 41) e, como tal, apresentaram a seguinte definição:

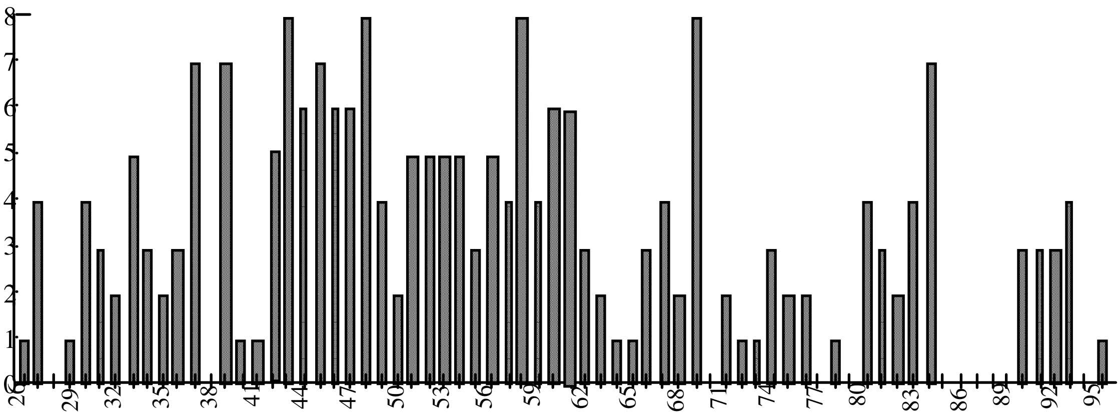
> *"**Diagrama de barras** - Representação gráfica, que consiste em marcar num sistema de eixos coordenados, no eixo dos xx, o valor das classes e nesses pontos barras verticais de altura igual à frequência absoluta ou à frequência relativa."* (p.41)

Na minha opinião, esta definição está um pouco confusa. Penso que seria mais correcto apresentar esta definição do seguinte modo:

> **Diagrama de barras** – Representação gráfica que consiste em marcar no eixo das abcissas os possíveis valores da variável e no eixo das ordenadas as frequências absolutas ou relativas. Por cima dos valores que estão marcados no eixo das abcissas, traçam-se linhas verticais com altura igual à respectiva frequência absoluta ou relativa.

Se um professor com pouca formação em Estatística for consultar a brochura de apoio ao professor para saber quais devem ser os *"passos a seguir na construção do diagrama de barras"* (p. 41), irá ficar com a ideia de que, quando se pretende comparar diagramas de barras de amostras diferentes, é conveniente construir diagramas de barras referentes à frequência relativa, pois temos a garantia de que a soma das barras é igual a 1. Mas, afinal, as amostras aleatórias não deveriam ser todas diferentes? Ou, será que os autores queriam dizer amostras de dimensões diferentes?

Na minha opinião, os autores da brochura de Estatística não deveriam ter apresentado o seguinte diagrama de barras:



(retirado da Brochura de Estatística, p. 43)

Este diagrama de barras, para além de ser de difícil leitura, é pouco apelativo e nem sequer tem legendas nem título. Além disso, por que motivo os autores da brochura construíram este diagrama de barras, em que no eixo das abcissas colocaram todos os valores da amostra, se anteriormente tinham apresentado uma tabela em que os dados estavam agrupados em classes?

2.2.6. Histograma e Polígono de Frequências

No que se refere ao histograma, os autores da brochura apresentaram a seguinte definição:

> *"**Histograma** - Para a representação gráfica de dados contínuos, usa-se um diagrama de áreas ou histograma, formado por uma sucessão de rectângulos adjacentes, tendo cada um por base um intervalo de classe e por área a frequência relativa (ou a frequência absoluta). Deste modo, a área total coberta pelo histograma é igual a 1 (respectivamente igual a n, a dimensão da amostra)."*
>
> (p. 43)

Mas, será que o histograma é apenas uma representação gráfica de dados contínuos? Se tiver uma amostra de dados discretos e agrupar os dados em classes, será que não posso representar estes dados através de um histograma?

Na minha opinião, seria mais correcto apresentar a definição de histograma da seguinte forma:

> O **histograma** é uma sucessão de rectângulos adjacentes. Quando os intervalos de classe têm amplitude constante, h, cada rectângulo tem base h e altura igual

> à respectiva frequência relativa [absoluta]. Assim, quando, por exemplo, se empregam frequências relativas, a área do rectângulo correspondente à classe $I_j\,(j=1,2,...,m)$ é hfr_j. Quando as classes têm amplitudes diferentes, h_j, os rectângulos têm base h_j e altura igual à respectiva frequência relativa [absoluta] dividida pela amplitude do intervalo de classe. Deste modo, a área do rectângulo correspondente à classe $I_j\,(j=1,2,...,m)$ é a frequência relativa [absoluta], e a soma das áreas é igual a 1 [número de observações, n].

Posteriormente, pode-se ler na brochura que:

> *"(…) se se pretender comparar várias amostras através de histogramas, deve-se ter o cuidado* de os construir *(…) de modo que a área total ocupada por cada um dos histogramas seja 1."* (p. 44)

Mas, será que devo ter este cuidado sempre que quero comparar amostras? Ou, será apenas quando quero comparar amostras com desigual número de observações?

Na minha opinião, foi bastante importante o cuidado que os autores da brochura de Estatística tiveram em ensinar a construir histogramas com a máquina de calcular. No entanto, volto a referir, que hoje em dia, ainda há muitos alunos que não têm calculadora gráfica e que também há muitas escolas que não têm calculadoras gráficas em número suficiente para emprestar aos alunos.

Dado que a brochura de Estatística tem como objectivo apoiar os professores de Matemática, penso que os autores deveriam ter tido o cuidado de abordar todos os pontos referidos nos programas de Matemática do 10.º ano. Por exemplo, o quarto ponto dos programas, para além de referir os histogramas como uma possível representação gráfica de variáveis contínuas, refere também os polígonos de frequências. No entanto, este último tipo de representação gráfica nunca foi abordado na brochura de apoio ao professor.

2.2.7. Função Cumulativa. Frequência Absoluta Acumulada e Frequência Relativa Acumulada

Os programas de Matemática fazem referência, tanto no terceiro como no quarto ponto, à *função cumulativa*. Um professor com pouca formação em Estatística e que não saiba o que é uma função cumulativa, se consultar a brochura de Estatística, irá encontrar o seguinte:

> "*Para representar graficamente as frequências acumuladas considera-se a função cumulativa*". (p.47)

Mas o que são as frequências acumuladas? Apesar dos programas de Matemática do 10.º ano referirem as frequências acumuladas, os autores da brochura de apoio ao professor não apresentaram nenhuma definição para este tipo de frequências, apenas deram uma ideia intuitiva de como se calculam estas frequências.

Na minha opinião, seria importante apresentar o conceito de frequência absoluta acumulada e de frequência relativa acumulada, antes de ser abordada a função cumulativa.

> **Frequência absoluta acumulada** de um valor, de uma classe ou de uma categoria é igual à soma das frequências absolutas anteriores com a frequência absoluta desse valor, dessa classe ou dessa categoria.
>
> **Frequência relativa acumulada** de um valor, de uma classe ou de uma categoria é igual à soma das frequências relativas anteriores com a frequência relativa desse valor, dessa classe ou dessa categoria.

Os programas de Matemática A e B fazem referência à função cumulativa quer no caso de variáveis discretas quer no caso de variáveis contínuas. No entanto, os autores da brochura de apoio aos programas só fizeram referência à função cumulativa para variáveis contínuas.

Uma vez que a função cumulativa varia muito quando se tem dados contínuos ou quando se tem dados discretos, penso que deveriam ter sido abordadas as duas situações.

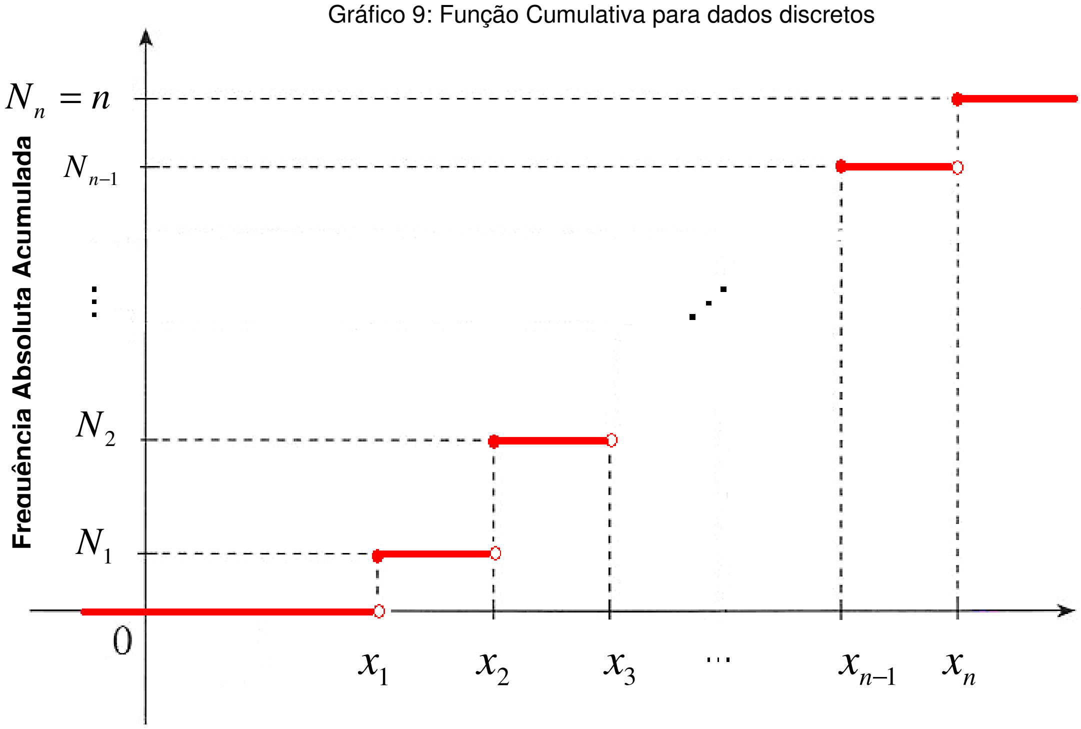
- <u>Função cumulativa para dados discretos utilizando as frequências absolutas acumuladas</u>

Num estudo estatístico em que a variável toma os valores $x_1, x_2, x_3,..., x_n$, com as frequências absolutas acumuladas $N_1, N_2, N_3,..., N_n$, chama-se **função cumulativa**, F, à função assim definida:

$$F(x)=\begin{cases}0 & se \quad x\in\,]-\infty;x_1[\\ N_1 & se \quad x\in[x_1;x_2[\\ N_2 & se \quad x\in[x_2;x_3[\\ \dots\dots\dots\dots & \\ N_{n-1} & se \quad x\in[x_{n-1};x_n[\\ N_n=n & se \quad x\in[x_n;+\infty[\end{cases}$$

A representação gráfica da função cumulativa é a seguinte:

Gráfico 9: Função Cumulativa para dados discretos



Neste caso, a função cumulativa apresenta as seguintes propriedades:

- Está definida em todo o x real;
- É sempre não decrescente;
- Só assume valores no conjunto $\{0, N_1, N_2, ..., N_{n-1}, N_n = n\}$.

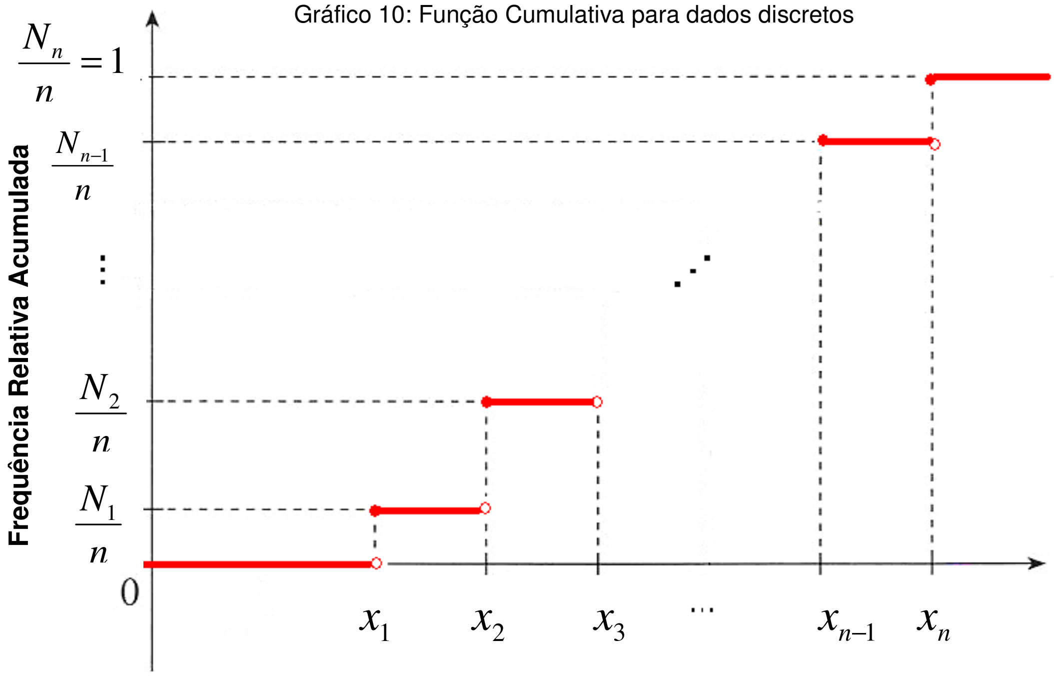
- <u>Função cumulativa para dados discretos utilizando as frequências relativas acumuladas</u>

Num estudo estatístico em que a variável toma os valores $x_1, x_2, x_3, ..., x_n$, com as frequências relativas acumuladas $\frac{N_1}{n}, \frac{N_2}{n}, \frac{N_3}{n}, ..., \frac{N_n}{n}$, chama-se **função cumulativa**, F*, à função assim definida:

$$F^*(x) = \begin{cases} 0 & se \quad x \in]-\infty; x_1[\\ \dfrac{N_1}{n} & se \quad x \in [x_1; x_2[\\ \dfrac{N_2}{n} & se \quad x \in [x_2; x_3[\\ \ldots\ldots\ldots\ldots & \\ \dfrac{N_{n-1}}{n} & se \quad x \in [x_{n-1}; x_n[\\ \dfrac{N_n}{n} = 1 & se \quad x \in [x_n; +\infty[\end{cases}$$

A representação gráfica da função cumulativa é a seguinte:

Gráfico 10: Função Cumulativa para dados discretos



Neste caso, a função cumulativa apresenta as seguintes propriedades:

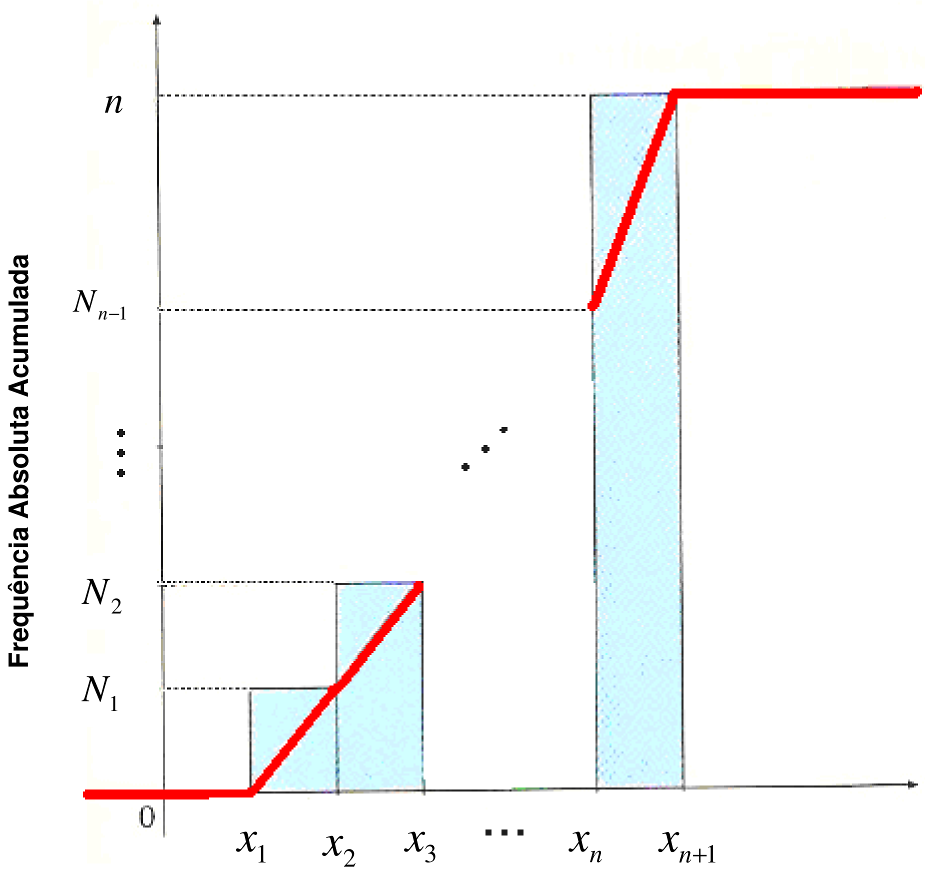
- Está definida em todo o x real;
- É sempre não decrescente;
- O contradomínio da função é o conjunto

$$\left\{0, \frac{N_1}{n}, \frac{N_2}{n}, \ldots, \frac{N_{n-1}}{n}, \frac{N_n}{n} = 1\right\} \subset [0;1].$$

- **<u>Função cumulativa para dados contínuos ou discretos agrupados em classes utilizando as frequências absolutas acumuladas</u>**

Num estudo estatístico em que os dados estão organizados em classes $[x_1;x_2[$, $[x_2;x_3[,...,[x_n;x_{n+1}[$, com as frequências absolutas acumuladas $N_1, N_2, N_3,..., N_n = n$, chama-se **função cumulativa**, F, à função cuja representação gráfica é a seguinte:

Gráfico 11: Função Cumulativa para dados contínuos e/ou agrupados em classes



O gráfico da função cumulativa obtém-se unindo os vértices superiores direitos dos rectângulos do histograma das frequências absolutas acumuladas.

Para definir analiticamente a função cumulativa é necessário definir cada um dos segmentos de recta e semi-rectas que a compõem.

Neste caso, a função cumulativa apresenta as seguintes propriedades:

- Está definida em todo o x real;
- É sempre não decrescente;
- É contínua;
- Só assume valores no intervalo $[0;n]$.

<u>Exemplo 1</u>: (Costa e Rodrigues, 2004)

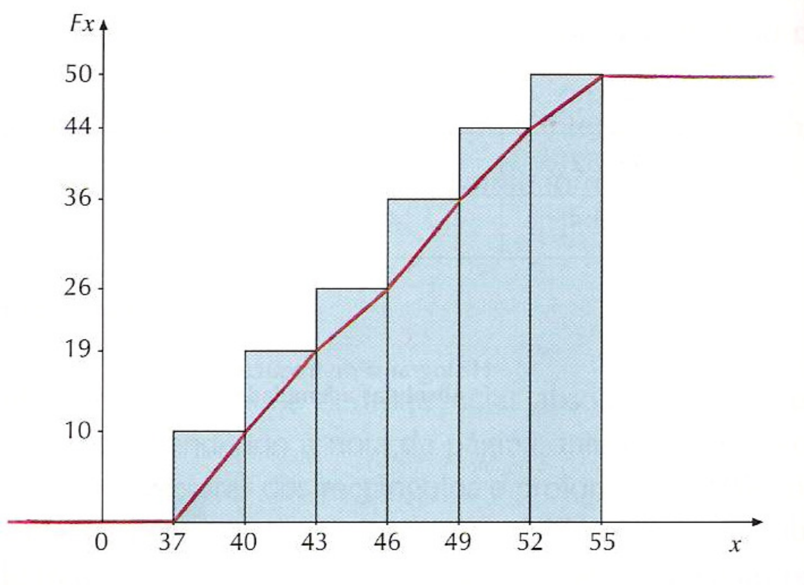
Num horto fez-se uma plantação de roseiras, todas com a mesma altura. Passados dois meses, recolheu-se uma amostra e registou-se o crescimento, em centímetros, das referidas roseiras.

Os dados encontram-se registados na seguinte tabela:

Crescimento (cm)	**Frequência Absoluta**	**Frequência Absoluta Acumulada**
[37; 40[	10	10
[40; 43[	9	19
[43; 46[	7	26
[46; 49[	10	36
[49; 52[	8	44
[52; 55[	6	50

O gráfico seguinte ilustra a representação gráfica da função cumulativa deste estudo.

Gráfico 12: Gráfico da Função Cumulativa relativo ao Exemplo 1



Para definir analiticamente a função cumulativa é necessário definir cada um dos segmentos de recta e semi-rectas que a compõem.

O segmento de recta de extremos *A* (37; 0) e *B* (40; 10) está contido numa recta definida por uma equação do tipo $y = mx + b$.

Como *A* e *B* pertencem a essa recta, tem-se:

$$\begin{cases} 37m+b=0 \\ 40m+b=10 \end{cases} \Leftrightarrow \begin{cases} b=-37m \\ 3m=10 \end{cases} \Leftrightarrow \begin{cases} b=-\dfrac{370}{3} \\ m=\dfrac{10}{3} \end{cases}$$

Assim a recta $\overline{AB}$ é definida por $y=\frac{10}{3}x-\frac{370}{3}$.

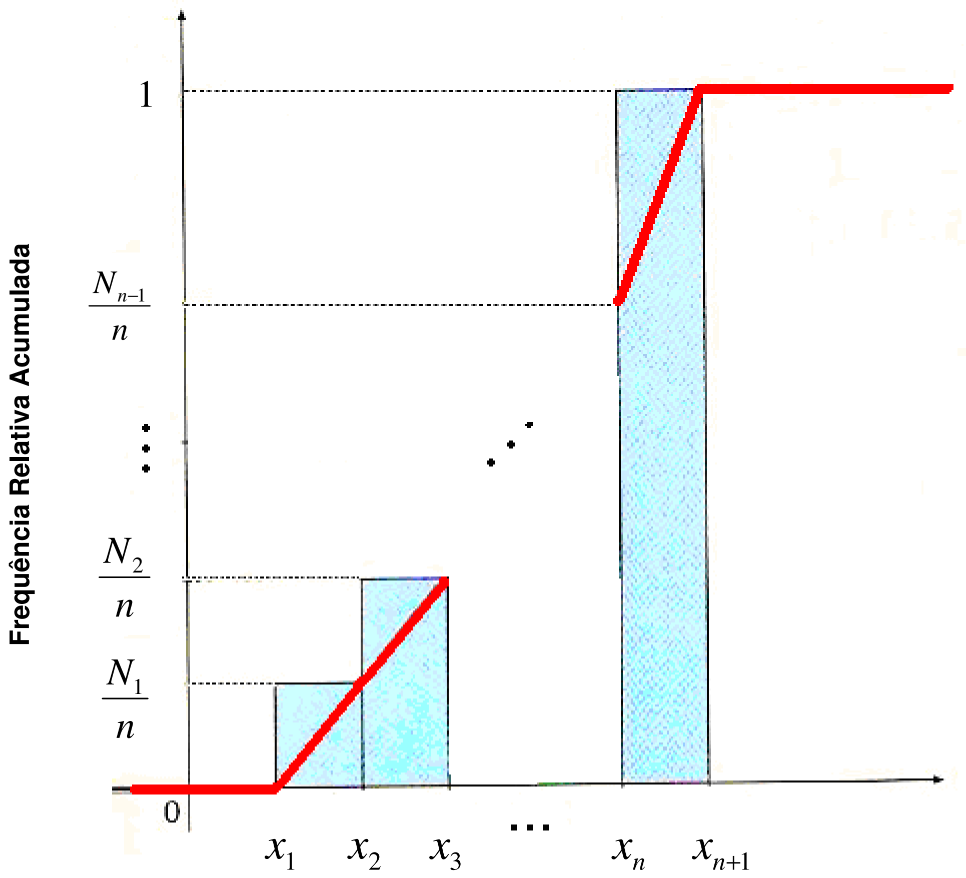
Procedendo do mesmo modo em relação aos outros segmentos, obtém-se:

$$F(x)=\begin{cases} 0 & se \quad x<37 \\ \dfrac{10}{3}x-\dfrac{370}{3} & se \quad 37\le x<40 \\ 3x-110 & se \quad 40\le x<43 \\ \dfrac{7}{3}x-\dfrac{244}{3} & se \quad 43\le x<46 \\ \dfrac{10}{3}x-\dfrac{382}{3} & se \quad 46\le x<49 \\ \dfrac{8}{3}x-\dfrac{284}{3} & se \quad 49\le x<52 \\ 2x-60 & se \quad 52\le x<55 \\ 50 & se \quad x\ge 55 \end{cases}$$

- <u>Função cumulativa para dados contínuos ou discretos agrupados em classes utilizando as frequências relativas acumuladas</u>

Num estudo estatístico em que os dados estão organizados em classes $[x_1;x_2[$, $[x_2;x_3[,\ldots,[x_n;x_{n+1}[$, com as frequências relativas acumuladas $\frac{N_1}{n},\frac{N_2}{n},\frac{N_3}{n},\ldots,\frac{N_n}{n}=1$, chama-se **função cumulativa**, F*, à função cuja representação gráfica é a seguinte:

Gráfico 13: Função Cumulativa para dados contínuos e/ou agrupados em classes



O gráfico da função cumulativa obtém-se unindo os vértices superiores direitos dos rectângulos do histograma das frequências relativas acumuladas.

Para definir analiticamente a função cumulativa é necessário definir cada um dos segmentos de recta e semi-rectas que a compõem.

Neste caso, a função cumulativa apresenta as seguintes propriedades:

- Está definida em todo o x real;
- É sempre não decrescente;
- É contínua;
- Só assume valores no intervalo $[0;1]$.

2.2.8. Diagrama Circular. Caule-e-Folhas. Diagrama de Extremos e Quartis. Caixa de Bigodes (*Box-Plot*)

Para além das representações gráficas que já foram referidas, os autores da brochura de Estatística abordaram outras representações gráficas, tais como, o diagrama circular, o caule-e-folhas, o diagrama de extremos e quartis e a caixa de bigodes (*Box-plot*).

No que se refere ao diagrama circular, os autores da brochura referiram o seguinte:

> "*Como o nome sugere esta representação é constituída por um círculo, em que se apresentam vários sectores circulares, tantos quantas as classes consideradas na tabela de frequências da amostra em estudo. Os ângulos dos sectores são proporcionais às frequências das classes.*" (p.50)

Como este tipo de representação é utilizada essencialmente para dados qualitativos e pode ser aplicada quer ao estudo dos elementos de uma amostra quer ao estudo de uma população, penso que seria mais correcto referir que:

> O **diagrama circular** é constituído por um círculo, em que se apresentam vários sectores circulares, tantos quantas as classes ou categorias consideradas na variável em estudo. A amplitude de cada sector é proporcional à frequência da respectiva classe ou categoria.

Posteriormente, os autores apresentaram o seguinte:

> "*Por exemplo uma classe com uma frequência relativa igual a 0.20, terá no diagrama circular um sector com um ângulo igual a 360x0.20=72 graus.*" (p. 50)

Na minha opinião, estes cálculos tornar-se-iam mais claros se fossem apresentados através de uma regra de três simples, tal como se segue:

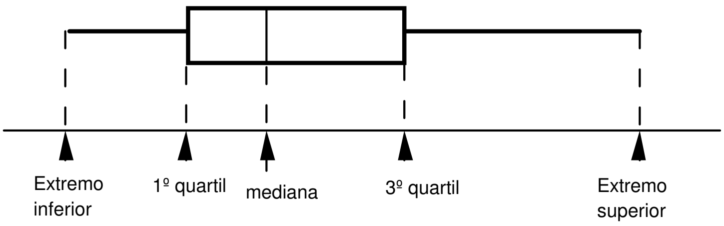
$$\begin{array}{lcl} 360^{\circ} & \text{———} & 1 \\ x & \text{———} & 0{,}2 \end{array} \qquad\qquad x = 0{,}2 \times 360 \Leftrightarrow x = 72^{\circ}$$

Esta forma de apresentar os cálculos, para além de ser mais perceptível, é uma boa sugestão metodológica para os professores utilizarem na sala de aula.

No que se refere ao diagrama de extremos e quartis, os autores da brochura de Estatística referiram o seguinte:

> "*É um tipo de representação gráfica, em que se realçam algumas características da amostra. O conjunto dos valores da amostra compreendidos entre o 1º e o 3º QUARTIS, que vamos representar por Q_1 e Q_3 é representado por um rectângulo (caixa) com a MEDIANA indicada por uma barra. A largura do rectângulo não dá qualquer informação, pelo que pode ser qualquer. Consideram-se seguidamente*

duas linhas que unem os meios dos lados do rectângulo com os extremos da amostra. Para obter esta representação, começa por se recolher da amostra, informação sobre 5 números, que são: os 2 extremos (mínimo e máximo), a mediana e o 1º e 3º quartis. A representação do diagrama de extremos e quartis tem o seguinte aspecto:



O extremo inferior é o mínimo da amostra, enquanto que o extremo superior é o máximo da amostra." (p.56 e 57)

Mas, afinal o que é a mediana? E o que são o 1º e o 3º QUARTIS?

Na minha opinião, faria mais sentido se as medidas de localização (moda ou classe modal; média; mediana; quartis) tivessem sido abordadas antes do diagrama de extremos e quartis. De facto, nos programas de Matemática do 10.º ano, o ponto referente às medidas de localização aparece antes do ponto que refere os diagramas de extremos e quartis. Desta forma, concluo que a brochura de Estatística, que serve de apoio ao professor de Matemática, se afasta mais uma vez do que é indicado nos programas de Matemática.

Se continuar a consultar a brochura de Estatística, verifico que os autores da brochura deram bastante importância à representação em diagrama de extremos e quartis, pois apresentaram o seguinte:

"*Realça informação importante sobre os dados, nomeadamente sobre o centro da amostra (mediana), variabilidade e simetria.*" (p.57)

Mas afinal o que é a variabilidade? E a simetria? De forma a poderem responder a estas questões, os autores da brochura referiram o seguinte:

"*Existem fundamentalmente três características da representação extremos e quartis, que nos dão ideia da simetria ou enviesamento dos dados e da sua maior ou menor concentração:*

- distância entre a linha indicadora da mediana e os lados do rectângulo;

- comprimento da caixa;

- comprimento das linhas que saem dos lados dos rectângulos."

(p.57 e 58)

Será que os professores com pouca formação em Estatística, depois de lerem estas três características, irão conseguir concluir acerca da simetria e da variabilidade dos dados a partir da representação do diagrama de extremos e quartis?

Como as representações em caule-e-folhas e em caixa de bigodes não fazem parte dos programas de Matemática do 10.º ano, penso que os autores da brochura de Estatística, em vez de terem abordado estas representações, deveriam ter dado mais atenção aos gráficos que de facto se encontram nos programas de Matemática do 10.º ano. Por exemplo, o primeiro ponto dos programas refere o seguinte:

> *"Análise gráfica de atributos qualitativos (gráficos circulares, diagramas de barras, pictogramas)"* (Silva *et al.*, 2001a, p. 30; Silva *et al.*, 2001b, p. 24)

Destas três representações gráficas, apenas foram abordados os gráficos circulares. Apesar dos diagramas de barras terem sido abordados, apenas foram apresentados no contexto de atributos quantitativos. Em relação aos pictogramas, nem sequer foi feita nenhuma referência ao longo de toda a brochura.

No que se refere à análise gráfica de variáveis contínuas, os programas de Matemática do 10.º ano fazem referência ao histograma e ao polígono de frequências. Este último, tal como os pictogramas, nem sequer foi referido ao longo de toda a brochura de Estatística. Portanto, penso que deveriam ter sido abordadas estas representações gráficas, em vez dos caules-e-folhas e das caixas de bigodes.

2.2.9. Tipos de Distribuições

Para além da atenção que os autores da brochura dedicaram às representações dos caules-e-folhas e das caixas de bigodes, dedicaram também alguma atenção aos diferentes tipos de distribuições. Apesar de sair mais uma vez fora do âmbito dos programas de Matemática do 10.º ano, penso que, de facto, é importante para o professor saber que há vários tipos de distribuições, que designação é atribuída a cada uma das distribuições e, principalmente, saber identificar as diferentes distribuições. Mas para isso, seria necessário que os autores da brochura tivessem apresentado a definição de distribuição, o que não aconteceu.

2.2.10. Medidas de Localização

2.2.10.1. Média

No que se refere às medidas de localização, a **média** é sem dúvida a medida de localização mais utilizada, no entanto, esta só pode ser calculada para dados do tipo quantitativo. Se for consultar os programas de Matemática A e B, leio o seguinte:

> *"A título de exemplo referimos o facto de não ter qualquer sentido calcular a média para dados de tipo qualitativo, mesmo que as diferentes categorias assumidas pela variável em estudo estejam representadas por números."* (Silva *et al.*, 2001a, p. 30; Silva *et al.*, 2001b, p. 24)

Se for consultar a brochura de Estatística, também leio o seguinte:

> *"Chama-se a atenção para que só tem sentido calcular a média para dados de tipo quantitativo."* (p.76)

Apesar de os autores da brochura terem feito esta observação, ela foi feita como uma espécie de conclusão ao estudo da média. Na minha opinião, dada a importância desta observação, ela deveria ter sido feita como uma introdução ao estudo da média e não como uma conclusão. Pois, não nos podemos esquecer, que certamente existirão muitos professores que não lêem as secções na íntegra.

No que diz respeito às expressões para calcular a média, penso que os autores da brochura preocuparam-se em apresentá-las de uma forma bastante perceptível e, inclusivamente, separaram todas as situações possíveis de uma forma clara e simples.

2.2.10.2. Mediana e Classe Mediana

Outra medida de localização do centro da amostra é a **mediana**. Se um professor com pouca formação em Estatística for consultar a brochura, irá encontrar o seguinte:

> *"(…) ordenados os elementos da amostra, a mediana é o valor (pertencente ou não à amostra) que a divide ao meio, isto é, 50% dos elementos da amostra são menores ou iguais à mediana e os outros 50% são maiores ou iguais à mediana.*
>
> *Para a determinação da mediana, utiliza-se a seguinte regra, depois de ordenada a amostra de n elementos:*
>
> *- Se **n** é **ímpar**, a mediana é o elemento médio.*

*- Se **n** é **par**, a mediana é a semi-soma dos dois elementos médios.*

Uma forma simples de aplicar a regra anterior é considerar o quociente $\frac{n+1}{2}$ *:*

- Se este quociente for um nº inteiro, considera-se para mediana o elemento nessa posição;

- Se este quociente terminar em 0.5, considera-se a sua parte inteira e faz-se a semi-soma do elemento a que corresponde essa ordem, com o seguinte." (p.79)

Uma vez que o quociente $\frac{n+1}{2}$ ou é um número inteiro ou é um número não inteiro, penso que os autores deveriam ter substituído a frase: *"Se este quociente terminar em 0.5, considera-se a sua parte inteira e faz-se a semi-soma do elemento a que corresponde essa ordem, com o seguinte."* (p.79) pela seguinte frase: *Se este quociente não for um número inteiro, considera-se a sua parte inteira e faz-se a semi-soma do elemento a que corresponde essa ordem, com o seguinte.*

Além disso, nestas frases pode-se ler: "*faz-se a semi-soma do elemento a que corresponde essa ordem, com o seguinte*". Mas, afinal o que é a ordem de um elemento? Será que um professor com pouca formação em Estatística sabe o que é a ordem de um elemento? De forma a colmatar mais esta dificuldade, os autores da brochura de Estatística deveriam ter apresentado o seguinte:

Quando o atributo é quantitativo ou qualitativo em escala ordinal (isto é, embora as categorias não sejam numéricas é possível estabelecer uma ordem ou uma hierarquia), os dados da colecção $\{x_1, x_2, x_3, ..., x_n\}$ podem ser ordenados da seguinte forma:

$$x_{(1)}, x_{(2)}, x_{(3)}, ..., x_{(n)} \quad \text{em que} \quad x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq ... \leq x_{(n)}$$

Assim, $x_{(i)}$ é o elemento de ordem i, com $i=1,2,3,\ldots,n$.

Os programas de Matemática do 10.º ano não fazem referência à **classe mediana**. No entanto, os autores da brochura tiveram o cuidado de ensinar a determinar a classe mediana e ainda tiveram o cuidado de ensinar a calcular um valor aproximado da mediana, a partir de dados agrupados.

Na minha opinião, é muito importante saber calcular tanto a mediana como a classe mediana, pois: "*Por vezes os dados apresentam-se agrupados, sendo necessário* *calcular a mediana a partir das tabelas ou das representações gráficas correspondentes.*" (Brochura de Estatística, 1997, p.82).

De forma a concluir o estudo da mediana, os autores da brochura de Estatística fizeram a seguinte chamada de atenção:

> "***Nota****: Deve-se chamar a atenção para que, com dados de tipo qualitativo, as únicas características amostrais que se podem calcular são a moda, categoria com maior frequência, e por vezes a mediana, quando for possível estabelecer uma hierarquia entre as diferentes categorias ou modalidades que a variável em estudo possa assumir. Por exemplo, numa grande empresa em que os trabalhadores podem assumir um de 5 postos possíveis, representados pelas letras A, B, C, D e E, em que o posto A é o mais importante e E o menos importante, recolheu-se uma amostra de 15 empregados, registando-se as respectivas categorias:*
>
> *A, E, E, E, B, C, E, D, D, E, B, D, D, E, E*
>
> *Ordenando a amostra anterior, por ordem crescente de importância do posto de trabalho, obtém-se:*
>
> E, E, E, E, E, E, E, D, D, D, D, C, B, B, A
>
> ↑
>
> **Mediana**
>
> *Se a amostra anterior não tivesse o elemento A, então a mediana seria o posto de trabalho E, pois 50% dos elementos da amostra têm categoria inferior ou igual a E."* (p.85)

Os autores da brochura cometeram um erro ao fazerem a última afirmação, pois a amostra anterior sem o elemento A fica constituída por 14 observações e se a ordenarmos por ordem crescente de importância do posto de trabalho obtemos:

E, E, E, E, E, E, $\boxed{\text{E, D,}}$ D, D, D, C, B, B

$$\text{Mediana} = \frac{E+D}{2} = ???$$

Como não é possível calcular a semi-soma entre diferentes postos de trabalho, o resultado da mediana não tem interpretação, portanto, não faz sentido calcularmos a mediana para esta última amostra.

2.2.10.3. Quartis

Outra medida de localização que também é utilizada com muita frequência é o quartil. Ao consultar a brochura de Estatística, constatei que os autores da brochura, para além de apresentarem a noção de 1º e 3º quartis, referiram que *"Há vários processos para a determinação dos quartis, que nem sempre conduzem aos mesmos resultados."* (p. 85). Face a esta afirmação, referiram que *"Um dos processos pode ser o de utilizar a mesma metodologia aplicada para a obtenção da mediana, isto é, consideram-se os quartis como as medianas das duas partes em que ficou dividida a amostra inicial pela mediana. A parte inferior é dividida pelo 1º quartil, enquanto que a parte superior é dividida pelo 3º quartil."* (p.85 e p.86).

Este deve ser sem dúvida um dos processos mais utilizados para determinar o 1º e o 3º quartis, no entanto, penso que os autores da brochura de Estatística deveriam ter apresentado as seguintes fórmulas para localizar os quartis num conjunto de dados:

$$Q_p = \begin{cases} \dfrac{x_{(np)} + x_{(np+1)}}{2}, & \text{se } np \text{ inteiro} \\ x_{([np]+1)} \quad , & \text{se } np \text{ não é inteiro} \end{cases}$$

2.2.10.4. Moda e Classe Modal

Ao consultar a brochura de Estatística, verifiquei que a **moda** é definida do seguinte modo:

> *"Para um conjunto de dados, define-se **moda** como sendo o valor que surge com mais frequência, se os dados são discretos, ou o intervalo de classe com maior frequência, se os dados são contínuos e estão agrupados."* (p. 87)

Será que o intervalo de classe com maior frequência designa-se por **moda** ou por **classe modal**? Além disso, será que só se tem intervalos de classe quando os dados são contínuos e estão agrupados? Será que não é possível agrupar em classes dados discretos?

Na minha opinião, seria mais correcto apresentar o seguinte:

> *"Quando estamos a trabalhar com dados discretos (...), a moda é o elemento mais frequente. Esta definição não faz sentido no caso de dados contínuos, pois nessa situação todos os dados são distintos. No entanto, quando agrupamos a*

> *amostra em classes, podemos chamar* ***classe modal*** *à classe em que observamos maior frequência. A moda agrupada é a abcissa do ponto de intersecção das duas rectas definidas pelos "vértices" do histograma na classe modal."* (Pestana e Velosa, 2006, p.137)

Uma característica importante da moda, à qual os autores da brochura de Estatística fizeram referência, é o facto de poder ser determinada para qualquer que seja a natureza do atributo observado (quantitativo ou qualitativo).

Outra característica da moda, que apesar de não ter sido referida na brochura de Estatística, penso que seria interessante referir:

> Para um conjunto de dados, pode existir mais do que uma moda ou até nem existir moda.
> - Se o conjunto de dados tiver duas modas, ele diz-se **bimodal**; no caso de ter mais do que duas modas, diz-se **multimodal**.
> - Se o conjunto de dados não tiver moda, ele diz-se **amodal**.

2.2.11. Medidas de Dispersão

2.2.11.1. Variância e Desvio Padrão

No que diz respeito às medidas de dispersão, penso que os autores da brochura de Estatística apresentaram de uma forma bastante correcta e clara os conceitos de variância e de desvio padrão. No entanto, penso que os autores também deveriam ter ensinado a calcular estas medidas de dispersão quando os dados se encontram organizados na forma de tabelas de frequências. Poderiam ter apresentado, por exemplo, o seguinte:

- <u>Os dados são discretos e estão organizados a partir dos diferentes valores que surgem na amostra</u>.

Para este caso, tem-se o seguinte:

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

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

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

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

$$s^2 = \frac{n}{n-1} \sum_{i=1}^{k} \left(x_i{}^2 - 2x_i\bar{x} + \bar{x}^2\right) fr_i$$

$$s^2 = \frac{n}{n-1} \sum_{i=1}^{k} \left(x_i{}^2 fr_i - 2x_i\bar{x} fr_i + \bar{x}^2 fr_i\right)$$

$$s^2 = \frac{n}{n-1} \left[\sum_{i=1}^{k} x_i{}^2 fr_i - 2\bar{x} \sum_{i=1}^{k} x_i fr_i + \bar{x}^2 \sum_{i=1}^{k} fr_i \right]$$

$$s^2 = \frac{n}{n-1} \left[\sum_{i=1}^{k} x_i{}^2 fr_i - 2\bar{x}\bar{x} + \bar{x}^2 .1 \right]$$

$$s^2 = \frac{n}{n-1} \left[\sum_{i=1}^{k} x_i{}^2 fr_i - \bar{x}^2 \right]$$

Portanto, a variância calcula-se a partir da seguinte expressão:

$$s^2 = \frac{n}{n-1} \left[\sum_{i=1}^{k} x_i{}^2 fr_i - \bar{x}^2 \right]$$

onde k é o número de valores diferentes que surgem na amostra; f_i é a frequência absoluta do i-ésimo valor da amostra que se encontra na tabela; fr_i é a frequência relativa do i-ésimo valor da amostra que se encontra na tabela; x_i corresponde ao i-ésimo valor da amostra que se encontra na tabela; $n = \sum_{i=1}^{k} f_i$, ou seja, n é a dimensão da amostra.

Neste caso, o desvio padrão é dado por: $s = \sqrt{\frac{n}{n-1} \left[\sum_{i=1}^{k} x_i{}^2 fr_i - \bar{x}^2 \right]}$

Ou simplesmente, por: $s = \sqrt{s^2}$

- **<u>Os dados são discretos ou contínuos e estão agrupados em classes</u>.**

Para os dados agrupados em classes, a variância também se calcula a partir da expressão:

$$s^2 = \frac{n}{n-1}\left[\sum_{i=1}^{k} x_i^{\,2} fr_i - \bar{x}^2\right]$$

mas neste caso, o k representa o número de classes do agrupamento; fr_i representa a frequência relativa da i-ésima classe; x_i representa o ponto médio da i-ésima classe, o qual é considerado como elemento representativo da classe; $n = \sum_{i=1}^{k} f_i$, ou seja, n é a dimensão da amostra.

Neste caso, o desvio padrão também é dado por: $s = \sqrt{\frac{n}{n-1}\left[\sum_{i=1}^{k} x_i^{\,2} fr_i - \bar{x}^2\right]}$

Ou simplesmente, por: $s = \sqrt{s^2}$

<u>Observação</u>: Quando se calcula a variância e o desvio padrão de dados agrupados em classes, não se obtêm os verdadeiros valores da variância e do desvio padrão, mas sim valores aproximados. Para se obter os valores exactos tem de se considerar os dados originais, caso estejam disponíveis.

<u>Observação</u>: Existem autores que dividem a variância por n. Neste caso, as expressões da variância e do desvio padrão serão, respectivamente, as seguintes:

$$s^2 = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 f_i}{n} \quad ; \quad s = \sqrt{\frac{\sum_{i=1}^{k} (x_i - \bar{x})^2 f_i}{n}}$$

Ou simplesmente,

$$s^2 = \sum_{i=1}^{k} x_i^{\,2} fr_i - \bar{x}^2 \quad ; \quad s = \sqrt{\sum_{i=1}^{k} x_i^{\,2} fr_i - \bar{x}^2}$$

Uma característica importante do desvio padrão é ser uma medida pouco resistente. Ao consultar a brochura de Estatística, verifiquei que os autores fizeram referência a esta característica através do seguinte excerto:

"O desvio padrão, da mesma forma que a média, é muito sensível à presença de outliers, sendo portanto uma medida de dispersão <u>pouco resistente</u>. Assim, um

valor elevado para o desvio padrão pode ser devido ou a uma grande variabilidade nos dados, ou então a uma pequena variabilidade com a existência de um ou mais outliers." (p. 96)

Mas o que são *outliers*? Será que um professor com pouca formação em Estatística sabe o conceito de *outlier*? Será que ele vai conseguir transmitir aos seus alunos esta característica do desvio padrão?

Uma vez que os *outliers* não fazem parte dos programas de Matemática do 10.º ano, penso que os autores da brochura de Estatística deveriam ter substituído o excerto anterior pelo seguinte:

> O desvio padrão, da mesma forma que a média, é muito sensível à presença de valores muito grandes ou muito pequenos na amostra, sendo portanto uma medida de dispersão <u>pouco resistente</u>. Assim, um valor elevado para o desvio padrão pode ser devido a uma grande variabilidade nos dados, ou então a uma pequena variabilidade com existência de valores muito grandes ou muito pequenos na amostra.

Outra característica importante da variância e do desvio padrão, à qual os autores da brochura de Estatística não fizeram referência, é o facto destas medidas de dispersão só poderem ser calculadas para dados quantitativos.

2.2.11.2. Amplitude e Amplitude Inter-Quartil

A variância e o desvio padrão são duas medidas que são muito utilizadas para medir a variabilidade dos dados. No entanto, estas não são as únicas medidas que permitem medir a dispersão dos dados. Há outras medidas, tais como, a **amplitude** e a **amplitude inter-quartil**.

A **amplitude** é a medida de dispersão mais simples de se calcular e define-se como sendo *"a diferença entre o máximo e o mínimo da amostra:*

R = máximo – mínimo".

(Brochura de Estatística, p. 96)

Mas o que é o máximo e o mínimo da amostra? Se um professor com pouca formação em Estatística for consultar a brochura para saber a definição de máximo e de mínimo não irá encontrar a definição de nenhum destes dois conceitos. Na minha opinião, será importante referir o seguinte:

> O elemento $x_{(1)}$ designa-se por **mínimo da amostra**, ou seja, o menor elemento da amostra e o elemento $x_{(n)}$ designa-se por **máximo da amostra**, ou seja, o maior elemento da amostra.

Apesar de se conseguir encontrar o mínimo e o máximo da amostra, quer para dados de natureza quantitativa quer para dados de natureza qualitativa em que a escala é ordinal, apenas se pode calcular a amplitude quando os dados são de natureza quantitativa.

A utilização da medida anterior tem como principal desvantagem o facto de ter apenas em conta os dois valores extremos que a variável toma, ou seja, é muito sensível à existência de uma observação muito grande ou muito pequena. Desta forma, define-se uma outra medida, a **amplitude inter-quartil**, que de um modo geral não é afectada pela existência de um reduzido número de observações muito grandes ou muito pequenas. Ao consultar a brochura de Estatística encontrei o seguinte:

> *"Esta medida é definida como sendo a diferença entre os 1º e 3º quartis:*
>
> *amplitude inter-quartil = 3º quartil - 1º quartil*
>
> *ou, utilizando a notação que introduzimos quando falamos nos quartis,*
>
> *amplitude inter-quartil =* $Q_3 - Q_1$
>
> *Do modo como se define a amplitude inter-quartil, concluimos que 50% dos elementos do meio da amostra estão contidos num intervalo com aquela amplitude."* (p.97)

2.2.11.3. Desvio Padrão *vs.* Amplitude Inter-Quartil

Depois de abordarem o conceito de desvio padrão e de amplitude inter-quartil, os autores da brochura de Estatística colocaram as seguintes questões:

"*Qual das medidas de dispersão utilizar? Desvio padrão ou amplitude inter-quartil?"* (p. 97).

De forma a responderem a estas questões, os autores da brochura de Estatística afirmaram o seguinte:

> *"- A amplitude inter-quartil é mais resistente, relativamente à presença de outliers, do que o desvio padrão, que é mais sensível aos dados. Por outro lado, a amplitude inter-quartil não reflecte o conjunto de todos os dados, como o desvio padrão.*
> *- Se a distribuição é enviesada pode acontecer que o desvio padrão seja muito superior à amplitude inter-quartil, sobretudo se se verificar a existência de "outliers"."* (p. 97)

Uma vez que o conceito de *outlier* não faz parte dos programas de Matemática do 10.º ano, penso que seria mais correcto se as afirmações anteriores tivessem sido estruturadas da seguinte forma:

> - A amplitude inter-quartil é mais resistente, relativamente à presença de observações muito grandes ou muito pequenas, do que o desvio padrão, que é mais sensível aos dados. Por outro lado a amplitude inter-quartil não reflecte o conjunto de todos os dados, como o desvio padrão.
> - Se se verificar a existência de observações muito grandes ou muito pequenas, pode acontecer que o desvio padrão seja muito superior à amplitude inter-quartil.

2.3. Referência a Distribuições Bidimensionais (Abordagem Gráfica e Intuitiva)

Na terceira secção dos programas de Matemática A e B surgem os seguintes pontos:

> • Diagrama de dispersão; dependência estatística; ideia intuitiva de correlação; exemplos gráficos de correlação positiva, negativa ou nula.
> • Coeficiente de correlação e sua variação em $[-1, 1]$.
> • Definição de centro de gravidade de um conjunto finito de pontos; sua interpretação física.
> • Ideia intuitiva de recta de regressão; sua interpretação e limitações.
>
> (Silva *et al.*, 2001a, p. 31)
> (Silva *et al.*, 2001b, p. 25)

2.3.1. Diagrama de Dispersão

Pode acontecer que sobre um elemento da população ou da amostra em estudo se recolha informação sobre duas características ou variáveis quantitativas, obtendo-se assim um conjunto de dados sobre a forma de pares de dados. Normalmente, o que se pretende neste caso é estudar a relação entre as duas variáveis que se supõe estarem relacionadas.

Ao consultar a brochura de Estatística, encontrei o seguinte:

> *"No caso de se pretender estudar duas características conjuntamente, os valores observados aparecem sob a forma de pares de valores, isto é, cada indivíduo ou resultado experimental contribui com um conjunto de dois valores."* (p.103)

Mas será que as populações ou as amostras apenas podem ser constituídas por indivíduos ou resultados experimentais? Penso que seria mais correcto dizer que: *cada elemento da população ou da amostra contribui com um conjunto de dois valores*.

O processo adequado para descrever a relação entre duas variáveis é começar pela representação gráfica que se designa por **nuvem de pontos** ou **diagrama de dispersão**.

> *"**Diagrama de dispersão** - É uma representação gráfica para os dados bivariados, em que cada par de dados (x_i,y_i) é representado por um ponto de coordenadas (x_i,y_i), num sistema de eixos coordenados."* (Brochura de Estatística, p. 103)

Este tipo de representação gráfica é muito útil, pois permite retirar informação acerca da forma, direcção e grau de associação entre as variáveis em estudo.

Apesar dos autores dos programas de Matemática A e B e dos autores da brochura de Estatística nunca referirem que existem diferentes tipos de relações entre as variáveis, penso que deveriam tê-lo feito, pois para além de ser importante saber identificar a relação que há entre duas variáveis, na análise de regressão é necessário que haja uma dependência funcional entre as variáveis (isto é, uma relação de causa--efeito).

De forma a combater esta lacuna, penso que teria sido importante referir o seguinte:

> *"A relação entre duas variáveis pode ser de dependência funcional (relação de causa-efeito) de uma em relação à outra, isto é, a magnitude de uma das variáveis (…) é função, ou é determinada pela magnitude da outra variável (…) sem que o recíproco seja também válido. Por exemplo, na relação entre a pressão sanguínea e a idade podemos assumir que a pressão sanguínea (…) é determinada pela idade (…). Porém, enquanto a pressão sanguínea pode ser função da idade, a idade não é, obviamente determinada, pela pressão sanguínea. Existem, porém, variáveis em que a relação entre elas não é de dependência de uma relativamente à outra, embora possam apresentar-se correlacionadas. Por exemplo, o número de gelados consumidos por uma amostra de cidadãos portugueses e por uma amostra de cidadãos argentinos durante o Verão. Frequentemente, como neste exemplo, a magnitude das duas variáveis está relacionada e a correlação observada deve-se ao facto de as variáveis estudadas terem influências comuns e alterarem simultaneamente os seus valores. Não é, porém, razoável admitir que existe uma relação de dependência funcional de uma variável relativamente à outra."* (Maroco e Bispo, 2005, p. 279 e 280)

Os programas de Matemática A e B sugerem que os professores de Matemática devem transmitir aos seus alunos uma *"ideia intuitiva de correlação"* e devem apresentar *"exemplos gráficos de correlação positiva, negativa ou nula"*. De facto, se for consultar a brochura de Estatística, verifico que os autores desta brochura tiveram o cuidado de exemplificar as diferentes situações que podem surgir, reflectindo os diferentes graus de correlação que se podem verificar entre as variáveis. No entanto, os autores da brochura não apresentaram o conceito de variáveis correlacionadas. Na minha opinião, seria importante referir o seguinte:

Quando representamos uma amostra de dados bivariados (x_i, y_i) através de um diagrama de dispersão e verificamos que existe uma certa relação entre as observações x_i e y_i que compõem os pares, dizemos que as **variáveis são correlacionadas** ou simplesmente, que **existe correlação entre as variáveis**.

Além disso, se os pontos se concentrarem à volta de uma linha recta, dizemos que é uma correlação linear. Se pelo contrário, os pontos estiverem concentrados à volta de uma curva, dizemos que é uma correlação não linear.

Um ponto que não vem referido nem no programa de Matemática A nem no de Matemática B é a identificação da **variável de interesse** (ou **dependente**) e da **variável explicativa** (ou **independente**). Apesar desta identificação ser bastante importante, os autores da brochura de apoio aos programas também não lhe deram muita importância, apenas fizeram uma breve referência a estes conceitos a partir do seguinte excerto:

"Para melhor compreendermos estes dados podemos fazer a representação gráfica adequada, obtendo uma nuvem de pontos, em que representamos nas ordenadas a variável de interesse (distância atingida no salto em comprimento) e em abcissa a variável explicativa (peso do estudante)." (p. 104)

Na minha opinião, seria importante referir o seguinte:

Denomina-se por **variável explicativa** (ou **independente**) aquela cuja modificação se supõe poder produzir uma modificação no valor da **variável de interesse** (ou **dependente**).

Inclusivamente, quando se pretende ilustrar a relação entre duas variáveis por meio de um diagrama de dispersão, por convenção, representa-se no eixo das ordenadas os valores da variável dependente e no eixo das abcissas os valores da variável independente.

A identificação correcta da variável independente e da variável dependente assume particular importância quando se estuda análise de regressão, pois os resultados da análise não são idênticos caso se escolha uma ou outra variável como independente/dependente.

2.3.2. Coeficiente de Correlação Linear

No que se refere ao coeficiente de correlação linear, se for consultar os programas de Matemática A e de Matemática B, verifico que ambos fazem referência ao coeficiente de correlação e à sua variação no intervalo [-1; 1], no entanto, nenhum deles faz referência às suas limitações.

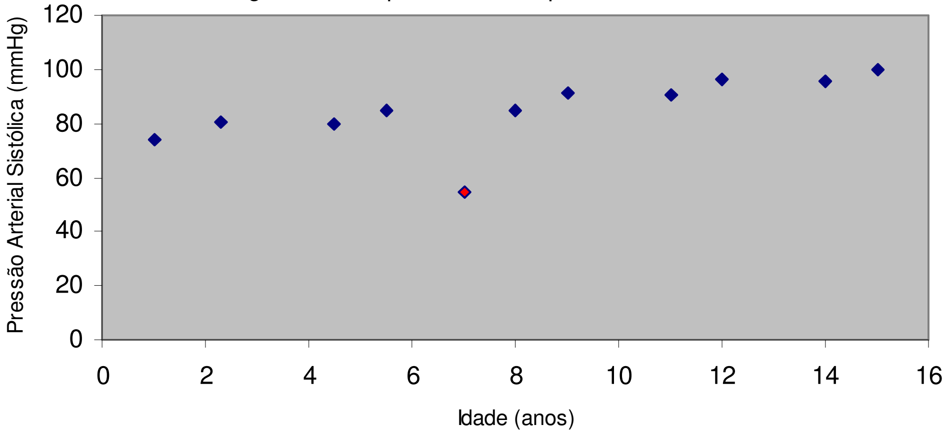
Na minha opinião, seria importante chamar a atenção dos alunos para o facto de que a existência de correlação elevada entre duas variáveis não significa necessariamente uma relação de causa-efeito.

Penso que também seria importante referir, nomeadamente dando exemplos, que o coeficiente de correlação é uma medida pouco resistente, sendo portanto influenciada por observações muito grandes ou muito pequenas.

Exemplo 2: (Maroco e Bispo, 2005)
Consideremos o estudo estatístico sobre pressão arterial sistólica *vs.* idade, em pessoas com idades inferiores a 20 anos. As observações e o diagrama de dispersão apresentam-se a seguir:

Idade Anos (X)	**Pressão arterial sistólica, mmHg (Y)**
1	74.3
2.3	80.4
4.5	80
5.5	84.8
7	**53.3**
8	84.7
9	91
11	90.8
12	96.5
14	95.8
15	100

Gráfico 14: Diagrama de dispersão entre a pressão arterial sistólica e a idade



Pela análise do diagrama de dispersão, verificamos que a observação (7 ; 53.3) [a vermelho] se encontra claramente fora da tendência linear das restantes observações.

Se calcularmos o coeficiente de correlação obtemos o valor 0.4213, mas se retirarmos a observação (7; 53.3) da amostra, o valor do coeficiente de correlação passa a ser 0.9421.

Portanto, concluímos que o coeficiente de correlação é de facto uma medida pouco resistente, ou seja, é influenciada por observações que caem fora da tendência linear das restantes observações.

Finalmente, penso que seria importante referir que o coeficiente de correlação apenas mede o grau de associação linear entre duas variáveis. Portanto, não será de estranhar se existir uma forte associação curvilínea entre duas variáveis e o valor do coeficiente de correlação ser aproximadamente zero.

Se for consultar a brochura de Estatística, constato que destas três limitações do coeficiente de correlação, os autores da brochura apenas fizeram referência à última. Uma vez que as duas primeiras não são menos importantes do que a última, penso que os autores da brochura de Estatística também as deveriam ter mencionado.

2.3.3. Recta de Regressão

No que diz respeito à recta de regressão, o programa de Matemática A e o programa de Matemática B sugerem apenas uma ideia intuitiva de recta de regressão.

Apesar dos professores de Matemática terem apenas de transmitir uma ideia intuitiva de recta de regressão aos seus alunos, penso que seria importante para os professores com pouca formação em Estatística conhecer o conceito de recta de regressão. Se for consultar a brochura de apoio ao professor, encontro a seguinte definição:

> *"Um dos critérios mais usados para definir esta recta, é o de tornar mínima a soma dos quadrados dos desvios dos pontos em relação à recta*[1]*. Essa recta é a chamada recta de regressão (dos mínimos quadrados)."* (p.108)

Para além do que os autores referiram penso que também seria importante acrescentar que:

A **recta de regressão** é a recta que melhor se ajusta aos pontos do diagrama de dispersão.

A **recta de regressão** tem de equação uma expressão do tipo: $y = ax + b$ onde

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}} \qquad b = \overline{y} - a\overline{x}$$

Ou simplesmente,

$$a = \frac{\sum_{i=1}^{n} x_i y_i - \overline{y}\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i} \qquad b = \overline{y} - a\overline{x}$$

Ou ainda,

$$a = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2} \qquad b = \overline{y} - a\overline{x}$$

---

[1] Designamos por desvio no ponto de abcissa x_i à diferença entre o valor observado y_i e o valor correspondente sobre a recta.

Ainda no que diz respeito à recta de regressão, os autores dos programas de Matemática A e B apresentaram a seguinte frase:

> *"Quando (…) se verificar uma tendência para a existência de uma associação linear entre as duas variáveis em estudo, identifica-se uma medida que quantifica o grau de associação - o coeficiente de correlação, assim como se apresenta um modelo matemático que permitirá, conhecido o valor de uma das variáveis, obter uma estimativa para o valor da outra variável."* (Silva *et al.*, 2001a, p. 31; Silva *et al.*, 2001b, p. 25)

Se for consultar a brochura de Estatística, verifico que os autores desta brochura apresentaram também a mesma ideia do que os autores dos programas, mas através da frase que passo a citar:

> *"Quando a correlação entre duas variáveis é elevada (quer seja positiva, quer seja negativa), isso significa que se conhecermos o valor de uma das variáveis então é possível ter uma ideia do valor que a outra variável irá tomar."* (p.108)

Será que se o diagrama de dispersão sugerir que há uma associação entre as duas variáveis, a recta que "melhor" se aproxima dos pontos que constituem a nuvem de pontos, irá permitir estimar o valor de uma das variáveis a partir do valor da outra variável? É importante que fique claro que uma recta de regressão permite conhecer como se reflectem na **variável dependente** as modificações processadas na **variável independente**, ou seja, a recta de regressão permite estimar o valor da **variável dependente** a partir do valor da **variável independente**, mas o recíproco não é válido. Por exemplo, na relação entre a pressão sanguínea e a idade, podemos assumir que a pressão sanguínea (variável dependente) é determinada pela idade (variável independente), porém, a idade não é obviamente determinada pela pressão sanguínea.

Quanto às limitações da recta de regressão, os autores da brochura de Estatística tiveram o cuidado de fazer referência ao facto de a recta de regressão ser pouco resistente. Como tal, apresentaram o seguinte exemplo:

> ***Exemplo****: Fez-se um estudo para averiguar a relação existente entre o número de veículos roubados por cada mil habitantes e a densidade populacional na cidade de Chicago. Seleccionaram-se, para o efeito, 18 distritos dessa cidade. Registou--se, para cada distrito, a sua densidade populacional (DP) e o número de veículos*

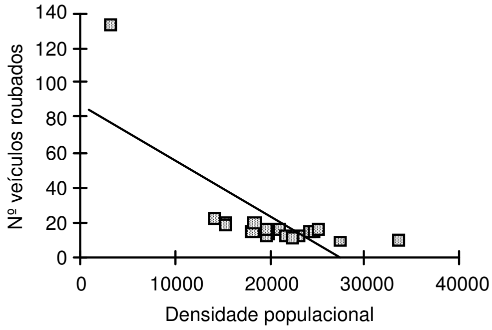
aí roubados (NVR) por cada mil habitantes, tendo-se obtido os seguintes resultados:

DP	NVR	DP	NVR	DP	NVR
3235	132.8	19581	16.5	21675	12.5
24182	14.9	14077	22.2	22315	11.8
20993	16.7	18137	15.8	18402	19.6
15401	20.0	22919	13.3	33445	10.5
19749	14.2	24534	15.1	27345	10.1
19487	13.5	24987	16.2	15358	19.0

O coeficiente de correlação é -0.74 e a recta de regressão de mínimos quadrados é

$$NVR = 88.195 - 0.00326\ DP$$

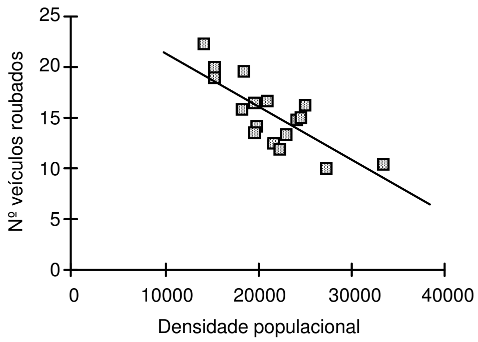
Se fizermos a representação gráfica destes dados vemos que há um distrito que tem um comportamento totalmente diferente dos outros.



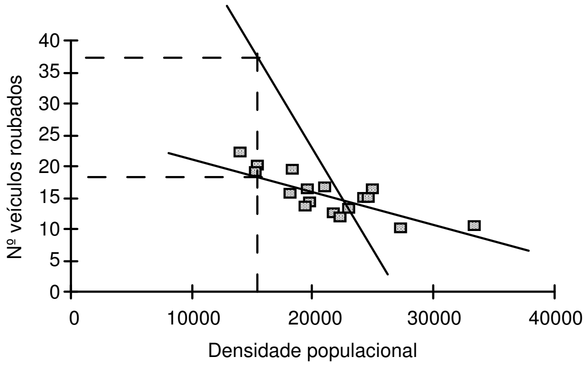
O 1º distrito que aparece na tabela tem uma densidade populacional muito baixa, mas um elevado número de veículos roubados. Uma averiguação mais cuidada levou à conclusão que aquele distrito correspondia ao Centro de Chicago, uma área essencialmente de comércio e de escritórios, e consequentemente uma área em que a densidade de veículos não tem a ver com a densidade populacional. Este distrito não deveria ter sido incluído na amostra. Assim, retirando este distrito, podemos construir uma nova recta de regressão. Obtém-se agora a recta

$$NVR = 27.36 - 0.00056\ DP$$

sendo o coeficiente de correlação -0.79.



Repare-se na alteração verificada. As conclusões extraídas de uma recta e de outra podem ser bem diferentes. Por exemplo, se considerarmos o valor de 15401 para a densidade populacional, que é um dos valores tabelados, o valor previsto para o número de carros roubados, utilizando a primeira recta de regressão é 38.0, enquanto que o previsto pela segunda recta de regressão é 18.7, bem mais próximo de 20 (valor observado).



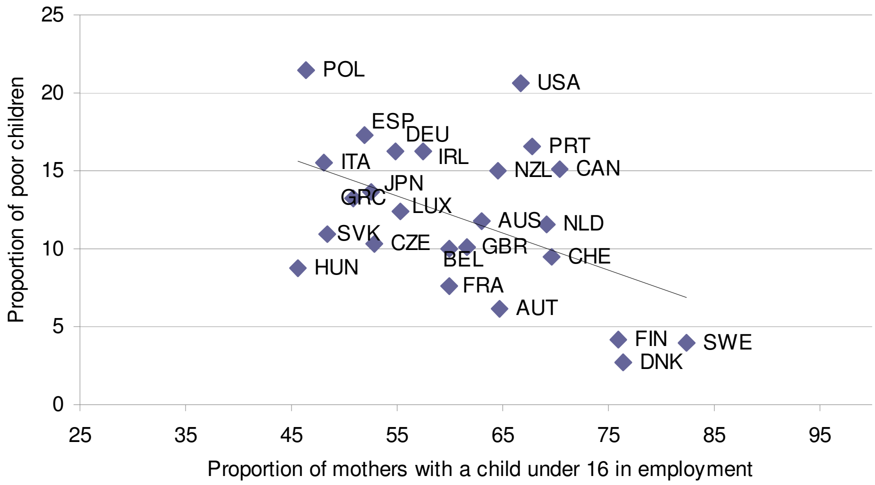
(Retirado da Brochura de Estatística 10.º ano, p. 111, 112 e 113)

Através deste exemplo, os autores da brochura deixaram transparecer um pouco a ideia de que quando há uma observação que se encontra claramente fora da tendência linear, o melhor procedimento será eliminar esta observação, pois desta forma, irá ser obtida uma recta que melhor se ajusta à nuvem de pontos. No entanto, é preciso deixar bem claro que a eliminação pura e simples de observações que caem fora da tendência linear das restantes observações não é um procedimento adequado. Antes de se proceder à eliminação de observações, que à partida parece que só irão prejudicar a qualidade de ajustamento da recta de regressão, é necessário fazer uma análise exploratória dos dados para se ter a certeza de que, de facto, se deve ou não eliminar observações, pois a eliminação de observações nem sempre conduz a uma recta que melhor se ajusta à nuvem de pontos.

<u>Exemplo 3:</u>

A meados da década de 2000, a OCDE realizou um estudo de forma a comparar a taxa de pobreza das crianças com a taxa de mães empregadas com um filho com menos de 16 anos.

Os dados obtidos foram apresentados através do seguinte gráfico:

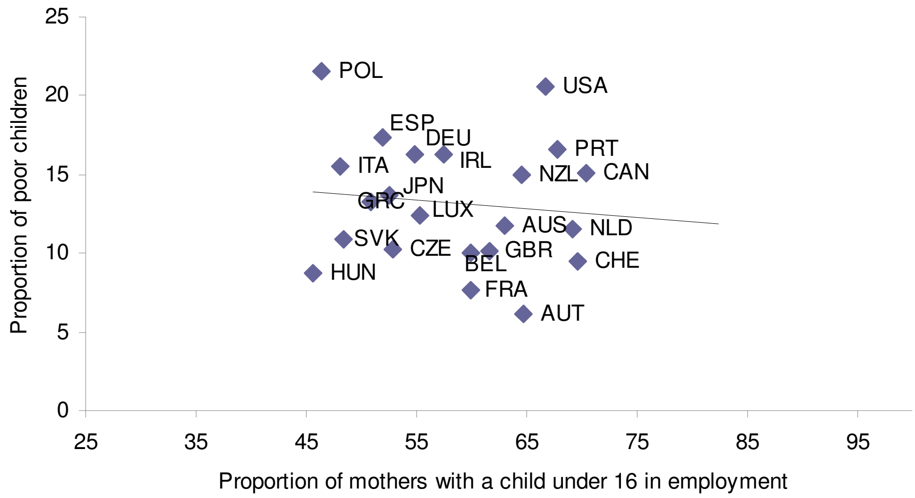


(Retirado do *site http://www.oecd.org/els/social/family/database*)

A recta de regressão de mínimos quadrados é $y = -0{,}2371x + 26{,}422$ e o coeficiente de correlação é -0,487.

Através da análise do gráfico anterior, verifica-se que a Finlândia, a Dinamarca e a Suécia estão afastadas das restantes observações formando um pequeno *cluster*. Assim, se estes três países forem eliminados, obtém-se o seguinte gráfico:

Gráfico 15: Diagrama de dispersão entre a taxa de pobreza das crianças e a taxa de mães empregadas com um filho com menos de 16 anos



Agora a recta de regressão dos mínimos quadrados é $y = -0{,}0561x + 16{,}446$ e o coeficiente de correlação é -0,11258.

> Neste caso, a eliminação de observações que estavam afastadas das restantes observações fez com que deixasse de haver regressão linear, ou seja, deixou de haver relação entre as variáveis em estudo.

Outra limitação da recta de regressão, à qual os autores da brochura de Estatística não fizeram qualquer menção, é o facto de a recta só poder ser construída quando há uma relação de causa-efeito entre as variáveis. Caso contrário, não fará qualquer sentido fazer extrapolações.

2.3.4. Centro de Gravidade de um Conjunto Finito de Pontos

No que se refere ao centro de gravidade de um conjunto finito de pontos e à sua interpretação física, os autores da brochura de Estatística deram pouca importância a este ponto, referindo apenas o seguinte:

> *"(…) a chamada recta de regressão (…) passa pelo centro de gravidade da distribuição, isto é, pelo ponto $(\bar{x}, \bar{y})$."* (p.108)

De facto, dado que o centro de gravidade pouco contribui para a obtenção da recta de regressão, penso que este ponto poderia ser eliminado quer do programa de Matemática A quer do programa de Matemática B, dando assim a possibilidade ao professor de aprofundar outros conteúdos de maior relevância.

A título de conclusão e perante todos os aspectos que foram referidos, é necessário que haja uma revisão curricular do programa de Matemática A e do programa de Matemática B no que se refere à unidade de Estatística bem como uma revisão global da brochura de Estatística, que tem por objectivo apoiar o professor de Matemática na leccionação da componente de Estatística do 10.º ano.

Para além destas revisões, é também necessário que haja "uma mudança de perspectiva, deixando de encarar a Estatística como um capítulo "pobre" e pouco interessante da Matemática, para passar a considerar como um elemento fundamental na formação básica da generalidade dos cidadãos" (Ponte e Fonseca, 2001, pág.112).

CAPÍTULO III

ANÁLISE DA TEMÁTICA DE PROBABILIDADES E COMBINATÓRIA NO PROGRAMA DE MATEMÁTICA A DO ENSINO SECUNDÁRIO

Esta análise é feita com base no programa de Matemática A do 12.º ano, Ministério da Educação, Departamento do Ensino Secundário, 17 de Maio de 2002 e na brochura de apoio ao professor de Matemática na leccionação da componente de Probabilidades e Combinatória do programa do 12.º ano: Probabilidades e Combinatória: matemática – 12.º ano de escolaridade, Ministério da Educação, Departamento do Ensino Secundário, Setembro de 1999.

Organização do Tema de Probabilidades e Combinatória no Programa de Matemática A do Ensino Secundário

A temática de Probabilidades e Combinatória do programa de Matemática A do 12.º ano, propõe numa primeira parte – ***Introdução ao cálculo de Probabilidades*** – o estudo das aproximações conceptuais para probabilidade, da definição axiomática de probabilidade, bem como da probabilidade condicionada e independência. Na segunda parte – ***Distribuição de frequências relativas e distribuição de Probabilidades*** – pretende-se abordar a relação entre as estatísticas e os parâmetros populacionais e entre a distribuição de frequências e a distribuição de probabilidades. Finalmente, na terceira parte – ***Análise Combinatória*** – pretende-se aplicar a análise Combinatória ao cálculo de Probabilidades.

Uma vez que a temática de Probabilidades e Combinatória está estruturada em três partes, a análise que se segue está igualmente organizada nas mesmas três secções.

3.1. Introdução ao Cálculo de Probabilidades

Na primeira secção do programa de Matemática A encontram-se os seguintes pontos:

- Experiência aleatória; conjunto de resultados; acontecimentos.
- Operações sobre acontecimentos.
- Aproximações conceptuais para Probabilidade:
  - aproximação frequencista de probabilidade;
  - definição clássica de probabilidade ou de Laplace.
  - definição axiomática de probabilidade (caso finito); propriedades da probabilidade.
- Probabilidade condicionada e independência; probabilidade da intersecção de acontecimentos. Acontecimentos independentes.

(Silva *et al.*, 2002c, p. 2)

Ao consultar a brochura de Probabilidades e Combinatória (Martins, M. E. G. (coord.), Monteiro, C., Viana, J. P. e Turkman, M. A. A. (1999). *Probabilidades e Combinatória: matemática – 12.º ano de escolaridade.* Lisboa: Ministério da Educação, Departamento do Ensino Secundário), verifiquei que os autores desta brochura antes de darem início à abordagem dos conceitos referidos no programa de Matemática A do 12.º ano, optaram por fazer uma breve introdução ao estudo das Probabilidades. Através desta introdução, os autores transmitiram uma ideia intuitiva do conceito de Probabilidades e apresentaram algumas situações do dia-a-dia em que as pessoas utilizam, até com uma certa perícia, o conceito de Probabilidades sem se aperceberem disso. Na minha opinião esta introdução é muito importante, não só para os professores com pouca formação universitária em Probabilidades, como para todos os professores que poderão tirar daqui várias ideias para introduzir, na sala de aula, o estudo das Probabilidades.

3.1.1. Experiência Aleatória. Espaço de Resultados. Acontecimentos

No que diz respeito à experiência aleatória, os autores da brochura apresentaram a seguinte definição:

> *"**Experiência aleatória** – processo que conduz à obtenção de uma observação ou resultado, de entre um conjunto de resultados possíveis (método utilizado para aquisição de dados)."* (p. 14)

Apesar de só se utilizarem experiências aleatórias no estudo das Probabilidades, penso que os autores da brochura também deveriam ter definido o conceito de experiência determinista. Assim, seria mais fácil conseguir compreender quais as experiências que se utilizam no estudo das Probabilidades. Como tal, penso que os autores da brochura deveriam ter referido o seguinte:

> *"Uma **experiência** é um procedimento que permite a obtenção de observações."* (Pestana e Velosa, 2006, p. 217)
>
> As **experiências deterministas** caracterizam-se por produzirem o mesmo resultado, desde que sejam repetidas sob as mesmas condições.
>
> As **experiências aleatórias** caracterizam-se pela impossibilidade de prever o resultado que se obterá, ainda que as experiências sejam realizadas sob as mesmas condições.
>
> <u>Exemplo</u>: (Neves *et al.*, 2006, p.16)
> Se atirarmos uma pedra a um lago já conhecemos o resultado mesmo antes de a atirarmos. A pedra vai ao fundo!
> A esta experiência chamamos <u>determinista</u>.
>
> <u>Exemplo</u>: (Neves *et al.*, 2006, p.16)
> Se lançarmos uma moeda ao ar não podemos afirmar qual das faces fica voltada para cima antes de termos efectuado o lançamento.
> A esta experiência chamamos <u>aleatória</u>.

No que se refere ao espaço de resultados, os autores da brochura de Probabilidades e Combinatória apresentaram a seguinte definição:

> *"**Espaço de resultados S** – conjunto de resultados possíveis associados a uma experiência aleatória."* (p. 15)

Na minha opinião, esta definição estaria mais correcta se fosse apresentada da seguinte forma:

O **espaço de resultados** é o conjunto de todos os resultados possíveis, associados a uma experiência aleatória.

Penso que também seria importante referir que o espaço de resultados é muitas vezes designado por **espaço amostral** e é, habitualmente, representado por S ou Ω.

No que diz respeito aos acontecimentos, penso que os autores da brochura, para além de terem apresentado a definição de acontecimento, deveriam ter referido que os acontecimentos têm nomes específicos e que são classificados do seguinte modo:

*"**Acontecimento elementar***
Acontecimento associado a uma experiência aleatória que consta de um só elemento do espaço de resultados.

Acontecimento composto
Acontecimento associado a uma experiência aleatória que consta de dois ou mais elementos do espaço de resultados.

Acontecimento certo
É o acontecimento que consta de todos os elementos do espaço de resultados. O acontecimento certo verifica-se sempre.

Acontecimento impossível
É o acontecimento que não tem qualquer elemento do espaço de resultados. O acontecimento impossível nunca se verifica."

(Neves *et al.*, 2006, p. 17)

Exemplo: (Neves *et al.*, 2006, p.17)

No espaço de resultados $S = \{1,2,3,4,5,6\}$ sejam os acontecimentos:

A: "sai 2"	*A* = {2}
B: "sai um número par"	*B* = {2,4,6}
C: "sai um número positivo"	*C* = {1,2,3,4,5,6}
D: "sai 7"	*D* = { }

A é um acontecimento elementar;
B é um acontecimento composto;
C é um acontecimento certo;
D é um acontecimento impossível.

Depois de terem apresentado a definição de acontecimento, os autores da brochura de Probabilidades e Combinatória apresentaram alguns exemplos de experiências aleatórias. Passo a citar o enunciado de um desses exemplos:

> *"**Exemplo 4** - Considere a experiência aleatória que consiste em lançar dois dados e verificar as faces que ficam voltadas para cima. Identifique o espaço de resultados e os acontecimentos "o número de pintas é igual nos dois dados" e "a soma das pintas é 7"."* (p. 17)

Neste exemplo, os autores colocaram a seguinte nota de rodapé associada à palavra "dados":

> *"No texto, um dado é constituído por 6 faces, com 1, 2, 3, 4, 5 ou 6 pintas, a menos que seja explicitamente referido o contrário."* (p. 17)

Na minha opinião, os autores da brochura deveriam ter acrescentado, nesta nota de rodapé, que os dados são equilibrados, a menos que seja explicitamente referido o contrário.

3.1.2. Extracções Com Reposição e Sem Reposição

No que diz respeito às extracções com reposição e sem reposição, os autores da brochura de Probabilidades e Combinatória tiveram o cuidado de elaborar esquemas para se tornar mais fácil a visualização das diferentes situações. A construção de esquemas é, sem dúvida, de uma elevada importância quer na sala de aula quer nesta brochura que tem como objectivo apoiar os professores com pouca formação na área de Probabilidades e Combinatória. Passo, então, a citar um exemplo, no qual os autores da brochura tiveram o cuidado de apresentar esquemas:

> *"**Exemplo 6** - Considere a experiência aleatória que consiste em extrair 2 berlindes, de um saco com 3 berlindes vermelhos e 2 azuis. Qual é o espaço de resultados?*
>
> *Para já é necessário saber se a extracção se faz com reposição ou sem reposição. Vamos considerar as duas situações. Para identificar o espaço de resultados será mais fácil numerar os berlindes, pelo que vamos numerar os berlindes vermelhos com 1, 2 e 3 e os azuis com 4 e 5.*

Com reposição *- Quando se retira um berlinde verifica-se a cor e torna-se a repor o berlinde no saco antes de extrair o próximo. O espaço de resultados é constituído por todos os resultados, em número de 25, do esquema seguinte:*

	①	②	③	④	⑤
①	①①	①②	①③	①④	①⑤
②	②①	②②	②③	②④	②⑤
③	③①	③②	③③	③④	③⑤
④	④①	④②	④③	④④	④⑤
⑤	⑤①	⑤②	⑤③	⑤④	⑤⑤

Sem reposição *- Neste caso o espaço de resultados é constituído por todos os resultados do espaço do esquema anterior, exceptuando os pares constituídos pelo mesmo berlinde:*

	①	②	③	④	⑤
①		①②	①③	①④	①⑤
②	②①		②③	②④	②⑤
③	③①	③②		③④	③⑤
④	④①	④②	④③		④⑤
⑤	⑤①	⑤②	⑤③	⑤④	

O acontecimento "tirar 2 berlindes de cor diferente" é constituído pelos resultados {(1,4), (1,5), (2,4), (2,5), (3,4), (3,5), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3)} tanto no esquema com reposição, como sem reposição." (p. 21)

No exemplo que acabei de citar, os autores começaram por colocar a seguinte questão: *"Qual o espaço de resultados?"*. Para poderem responder a esta questão, os autores necessitavam de ter acrescentado o seguinte:

Os espaços de resultados S_C e S_S correspondentes às duas situações, com reposição e sem reposição, são respectivamente:

$$\begin{aligned} S_C = \{&(1,1),(1,2),(1,3),(1,4),(1,5),(2,1),(2,2),(2,3),(2,4),(2,5),(3,1),(3,2),\\ &(3,3),(3,4),(3,5),(4,1),(4,2),(4,3),(4,4),(4,5),(5,1),(5,2),(5,3),(5,4),(5,5)\} \end{aligned}$$

$$\begin{aligned} S_S = \{&(1,2),(1,3),(1,4),(1,5),(2,1),(2,3),(2,4),(2,5),(3,1),(3,2),(3,4),(3,5),\\ &(4,1),(4,2),(4,3),(4,5),(5,1),(5,2),(5,3),(5,4)\} \end{aligned}$$

3.1.3. Operações com Acontecimentos

Ao consultar a brochura de Probabilidades e Combinatória, verifiquei que os autores deste guia começaram por definir acontecimento complementar. No entanto, penso que esta não foi a melhor opção, pois como ainda não tinham definido acontecimento união nem acontecimento intersecção, os autores nunca chegaram a concluir que da definição de acontecimento complementar resulta que $\mathrm{A} \cap \overline{\mathrm{A}} = \{\ \}$ e $\mathrm{A} \cup \overline{\mathrm{A}} = \mathrm{S}$.

Além disso, verifiquei que os autores deste guia definiram acontecimento união e intersecção da seguinte forma:

> *"- Acontecimento **Intersecção***
> ***Intersecção** dos acontecimentos A e B, **A** $\cap$ **B** , ou (A **e** B) é o acontecimento que se realiza sse A e B se realizam simultaneamente."* (p.24)

> *"- Acontecimento **União***
> ***União** dos acontecimentos A e B, **A** $\cup$ **B** , ou (A **ou** B) é o acontecimento que se realiza sse A ou B se realizam."* (p. 25)

Na minha opinião, estas definições estão um pouco confusas no que diz respeito à notação. Como tal, penso que estariam mais perceptíveis se tivessem sido apresentadas da seguinte forma:

> *"A **intersecção** dos acontecimentos **A** e **B** representa-se por $A \cap B$ (lê-se "A e B") e é o acontecimento que se realiza se e só se **A** e **B** se realizam simultaneamente."* (Costa e Rodrigues, 2006, p. 25)

> *"A **união** dos acontecimentos **A** e **B** representa-se por $A \cup B$ (lê-se "A ou B") e é o acontecimento que se realiza se e só se **A** ou **B** se realizam."* (Costa e Rodrigues, 2006, p. 24)

Agora que já se definiu acontecimento união e acontecimento intersecção, pode-se definir acontecimento complementar. Ao consultar a brochura de Probabilidades e Combinatória, encontrei o seguinte:

> *"- Acontecimento **Complementar** do acontecimento **A**:*
> *O acontecimento **complementar** do acontecimento A, representa-se por $\overline{\mathrm{A}}$ ou A^C e é o acontecimento constituído por todos os resultados de S, que não estão em A."* (p. 23)

Na minha opinião, esta definição está bastante perceptível, no entanto, os autores deveriam ter referido que o acontecimento complementar é também designado por acontecimento contrário. Além disso, da definição de acontecimento complementar sai que $\mathrm{A} \cup \overline{\mathrm{A}} = \mathrm{S}$ e $\mathrm{A} \cap \overline{\mathrm{A}} = \{\ \}$, ou seja, a união dos acontecimentos A e $\overline{\mathrm{A}}$ é um acontecimento certo, enquanto que a intersecção dos acontecimentos A e $\overline{\mathrm{A}}$ é um acontecimento impossível.

Se continuar a consultar a brochura de Probabilidades e Combinatória, encontro a seguinte definição:

> *"- Acontecimento **A implica B***
> *O acontecimento A **implica** a realização do acontecimento B, quando todo o resultado de A é um resultado de B; indica-se este facto escrevendo **A**⊂ **B**."* (p.24)

Mas, será que o acontecimento *A* tem que estar estritamente contido em B? Ou, será que podemos escrever $A \subseteq B$?

No que diz respeito ao acontecimento diferença, os autores da brochura de apoio ao professor apresentaram a seguinte definição:

> *"- Acontecimento **Diferença***
> *Acontecimento diferença entre A e B, **A-B**, é o acontecimento que se realiza se A se realiza, sem que B se realize."* (p. 26)

Na minha opinião, a definição anterior é pouco intuitiva. Penso que seria mais perceptível para um professor com pouca formação em Probabilidades se o conceito fosse definido da seguinte forma:

> O acontecimento **diferença** entre *A* e *B* representa-se por *A-B* ou *A\B* ou $A \cap \overline{B}$ e é constituído pelos elementos que pertencem a *A* e não pertencem a *B*.

3.1.4. Modelos de Probabilidades

Ao consultar a brochura de Probabilidades e Combinatória com o intuito de saber o que são os modelos de probabilidades, encontrei o seguinte:

> *"Um dos primeiros passos na definição de um modelo de probabilidade que descreva uma experiência aleatória é precisamente a definição do espaço de resultados associado. Posteriormente teremos de associar probabilidades a cada um dos elementos do espaço de resultados."* (p. 29)

Será que um professor com pouca formação em Probabilidades irá ficar esclarecido no que diz respeito aos modelos de Probabilidades? Além disso, será que irá conseguir transmitir aos seus alunos a ideia correcta de modelos de probabilidades? Na minha opinião, os autores da brochura de Probabilidades e Combinatória deveriam ter referido o seguinte:

> *"(…) um **modelo de probabilidade** é encarado como uma representação simples que traduza o melhor possível uma experiência aleatória.*
>
> *Para encontrar um modelo que descreva uma certa experiência aleatória há duas fases a considerar:*
>
> - *definir o espaço de resultados associado à experiência aleatória;*
> - *associar a cada acontecimento elementar um número que quantifique a possibilidade de esse acontecimento ocorrer, ou seja, a probabilidade de acontecimento ocorrer."* (Costa e Rodrigues, 2006, p. 30)

Posteriormente, os autores da brochura de Probabilidades e Combinatória referiram o seguinte:

> *"Consideremos uma experiência aleatória que conduza a um espaço de resultados S discreto, isto é, que só assume um número finito ou infinito numerável de resultados distintos que representamos por E_1, E_2, E_3, ….. Então, qualquer que seja o modo de construir o modelo de probabilidade (isto é, obter as probabilidades associadas aos acontecimentos elementares que constituem o espaço de resultados), vamos fixar como regras básicas as seguintes:*
>
> *<u>Regra 1</u> - A probabilidade de qualquer acontecimento elementar E_i é um número entre 0 e 1*
>
> $$0 \leq P(E_i) \leq 1$$

Regra 2 - A soma das probabilidades dos acontecimentos elementares que compõem o espaço de resultados S = {E_1, E_2, E_3, …} é igual a 1

$$\sum P(E_i) = 1$$

Qualquer que seja o acontecimento A, associado ao espaço de resultados S, define-se:

*A **probabilidade** P(A) do acontecimento A, é a soma das probabilidades dos acontecimentos elementares que compõem A."* (p.32)

Uma vez que o conceito de infinito numerável não faz parte dos conteúdos programáticos do Ensino Secundário, penso que os autores da brochura poderiam ter estruturado o texto anterior da seguinte forma:

Qualquer modelo de probabilidades associado a uma experiência aleatória com *n* acontecimentos elementares $E_1, E_2, ..., E_n$ satisfaz as seguintes propriedades:

- A probabilidade de qualquer acontecimento elementar E_i com $i \in \{1,2,\ldots,n\}$ é um número entre 0 e 1.

$$0 \leq P(E_i) \leq 1 \quad \text{com} \quad i \in \{1,2,...,n\}$$

- A soma das probabilidades de todos os acontecimentos elementares é igual a 1.

$$P(E_1) + P(E_2) + ... + P(E_n) = 1$$
$$\left(ou, \quad \sum_{i=1}^{n} P(E_i) = 1 \right)$$

- Qualquer que seja o acontecimento composto *A*, a probabilidade do acontecimento *A* é igual à soma das probabilidades dos acontecimentos elementares que o constituem.

3.1.5. Aproximações Conceptuais para Probabilidade

3.1.5.1. Aproximação Frequencista de Probabilidade

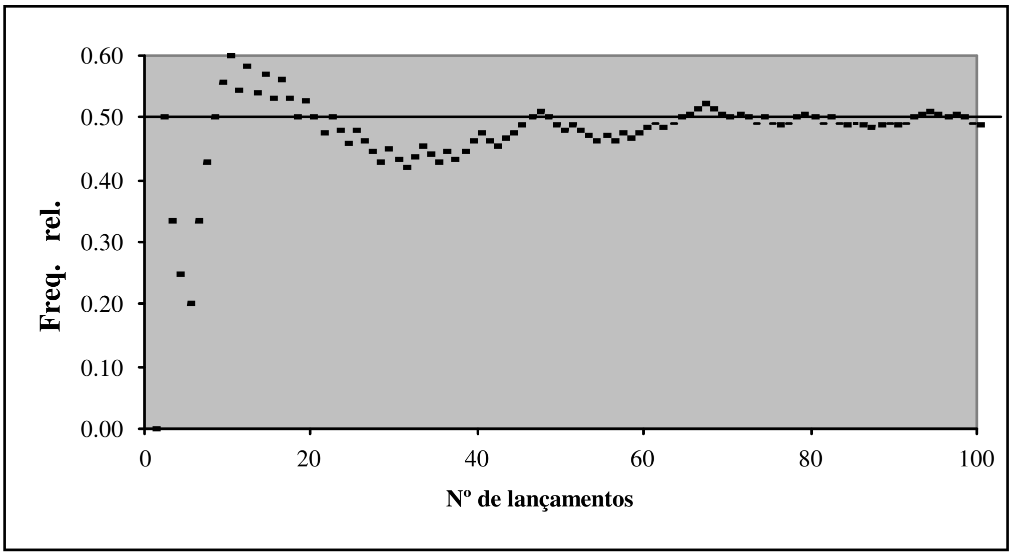
No que se refere à aproximação frequencista de probabilidade, os autores da brochura de Probabilidades e Combinatória começaram por transmitir uma ideia intuitiva de como se obtém a probabilidade de um acontecimento segundo a definição frequencista e inclusivamente, apresentaram o seguinte exemplo:

"Suponhamos, por exemplo, a experiência aleatória que consiste no lançamento de uma moeda ao ar e observar a face que fica virada para cima. Realizaram-se 100 lançamentos, tendo-se obtido os seguintes resultados:

1	*cara*	*21*	*cara*	*41*	*cara*	*61*	*coroa*	*81*	*Cara*
2	*coroa*	*22*	*coroa*	*42*	*cara*	*62*	*cara*	*82*	*Coroa*
3	*cara*	*23*	*cara*	*43*	*coroa*	*63*	*coroa*	*83*	*Cara*
4	*cara*	*24*	*cara*	*44*	*coroa*	*64*	*coroa*	*84*	*Cara*
5	*cara*	*25*	*coroa*	*45*	*coroa*	*65*	*coroa*	*85*	*Coroa*
6	*coroa*	*26*	*cara*	*46*	*coroa*	*66*	*coroa*	*86*	*Cara*
7	*coroa*	*27*	*cara*	*47*	*coroa*	*67*	*coroa*	*87*	*Cara*
8	*coroa*	*28*	*cara*	*48*	*cara*	*68*	*cara*	*88*	*Coroa*
9	*coroa*	*29*	*coroa*	*49*	*cara*	*69*	*cara*	*89*	*Coroa*
10	*coroa*	*30*	*cara*	*50*	*cara*	*70*	*cara*	*90*	*Cara*
11	*cara*	*31*	*cara*	*51*	*coroa*	*71*	*coroa*	*91*	*Coroa*
12	*coroa*	*32*	*coroa*	*52*	*cara*	*72*	*cara*	*92*	*Coroa*
13	*cara*	*33*	*coroa*	*53*	*cara*	*73*	*cara*	*93*	*Coroa*
14	*coroa*	*34*	*cara*	*54*	*cara*	*74*	*coroa*	*94*	*Coroa*
15	*cara*	*35*	*cara*	*55*	*coroa*	*75*	*cara*	*95*	*Cara*
16	*coroa*	*36*	*coroa*	*56*	*cara*	*76*	*cara*	*96*	*Cara*
17	*cara*	*37*	*cara*	*57*	*coroa*	*77*	*coroa*	*97*	*Coroa*
18	*cara*	*38*	*coroa*	*58*	*cara*	*78*	*coroa*	*98*	*Cara*
19	*coroa*	*39*	*coroa*	*59*	*coroa*	*79*	*coroa*	*99*	*Cara*
20	*cara*	*40*	*coroa*	*60*	*coroa*	*80*	*cara*	*100*	*Cara*

Se ao fim dos 100 lançamentos se verificaram 49 coroas, então a frequência relativa com que se verificou o acontecimento saída de coroa foi de 0.49. O valor para que tende a frequência relativa da saída de coroa, ao fim de um grande número de lançamentos, é interpretado como a probabilidade do acontecimento saída de coroa.

O gráfico obtido para a frequência relativa após cada lançamento, tem o seguinte aspecto:



> *A frequência relativa, à medida que o número de provas aumenta, tem tendência a estabilizar à volta do valor 0.5. Assim, dizemos que a probabilidade de sair coroa é 0.5.”* (p.33 e 34)

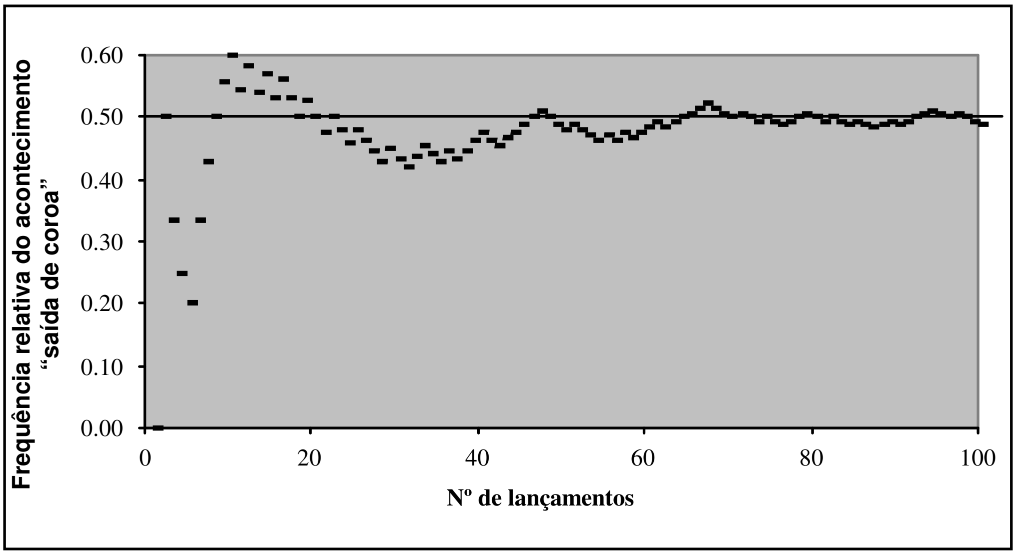
Na minha opinião, este exemplo está bastante perceptível, inclusivamente para os professores com pouca formação em Probabilidades. No entanto, gostava de referir que os autores da brochura poderiam ter sido mais rigorosos com a escrita, no sentido em que falam em frequência relativa e de um valor da frequência relativa, mas não mencionam a que acontecimento se refere a frequência relativa. Perante o que acabei de referir, penso que os autores da brochura deveriam ter apresentado o exemplo anterior da seguinte forma:

> Suponhamos, por exemplo, a experiência aleatória que consiste no lançamento de uma moeda ao ar e observar a face que fica virada para cima. Realizaram-se 100 lançamentos, tendo-se obtido os seguintes resultados:

1	cara	21	cara	41	cara	61	coroa	81	Cara
2	coroa	22	coroa	42	cara	62	cara	82	Coroa
3	cara	23	cara	43	coroa	63	coroa	83	Cara
4	cara	24	cara	44	coroa	64	coroa	84	Cara
5	cara	25	coroa	45	coroa	65	coroa	85	Coroa
6	coroa	26	cara	46	coroa	66	coroa	86	Cara
7	coroa	27	cara	47	coroa	67	coroa	87	Cara
8	coroa	28	cara	48	cara	68	cara	88	Coroa
9	coroa	29	coroa	49	cara	69	cara	89	Coroa
10	coroa	30	cara	50	cara	70	cara	90	Cara
11	cara	31	cara	51	coroa	71	coroa	91	Coroa
12	coroa	32	coroa	52	cara	72	cara	92	Coroa
13	cara	33	coroa	53	cara	73	cara	93	Coroa
14	coroa	34	cara	54	cara	74	coroa	94	Coroa
15	cara	35	cara	55	coroa	75	cara	95	Cara
16	coroa	36	coroa	56	cara	76	cara	96	Cara
17	cara	37	cara	57	coroa	77	coroa	97	Coroa
18	cara	38	coroa	58	cara	78	coroa	98	Cara
19	coroa	39	coroa	59	coroa	79	coroa	99	Cara
20	cara	40	coroa	60	coroa	80	cara	100	Cara

> Se ao fim dos 100 lançamentos se verificaram 49 coroas, então a frequência relativa com que se verificou o acontecimento saída de coroa foi de 0,49. O valor para que tende a frequência relativa da saída de coroa, ao fim de um grande número de lançamentos, é interpretado como a probabilidade do acontecimento saída de coroa.
>
> O gráfico da frequência relativa do acontecimento “saída de coroa” após cada lançamento tem o seguinte aspecto:

Gráfico 16: Gráfico da frequência relativa do acontecimento "saída de coroa" no lançamento de uma moeda ao ar



A frequência relativa do acontecimento "saída de coroa", à medida que o número de provas aumenta, tem tendência a estabilizar à volta do valor 0.5. Assim, dizemos que a probabilidade de sair coroa é 0.5.

Depois dos autores da brochura de Probabilidades e Combinatória terem apresentado a definição frequencista de probabilidade, apresentaram algumas sugestões de actividades para os professores desenvolverem na sala de aula. Na minha opinião, estas propostas de actividades foram bastante bem concebidas pois, para além de permitirem que o aluno apreenda integralmente a noção frequencista de probabilidade, poderão ser desenvolvidas através da calculadora ou através do simples uso de dados.

3.1.5.2. Definição Clássica de Probabilidade ou de Laplace

Quanto à definição clássica de Probabilidades ou de Laplace, os autores da brochura apresentaram a seguinte definição:

*"Dado o espaço de resultados S constituído por um número finito **n** de elementos, todos eles igualmente possíveis, define-se **Probabilidade de um acontecimento A** e representa-se por **P(A)**, como sendo a razão entre o número de resultados **favoráveis** a A (resultados que compõem A) – n_A e o número de resultados **possíveis** (resultados que constituem S) – **n**:*

$$P(A)=\frac{n_A}{n}\text{"}$$ (p. 43)

Na minha opinião, a forma como esta definição está redigida dá pouco ênfase ao facto de que os n elementos que constituem o espaço de resultados, S, têm de ser todos eles equiprováveis, ou seja, todos eles têm de ter igual probabilidade de ocorrer. Como tal, sugeria que a definição clássica de Probabilidade ou de Laplace fosse apresentada da seguinte forma:

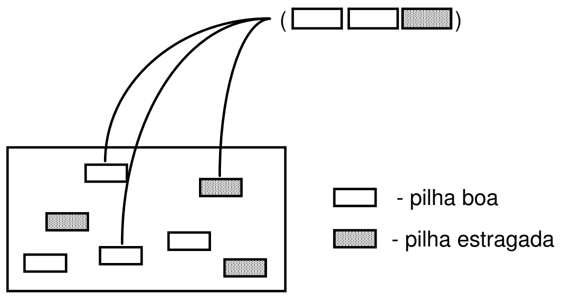
Seja E uma experiência aleatória em que o espaço amostral S é constituído por n elementos, sendo **equiprováveis** os n acontecimentos elementares. Se um acontecimento A é formado por m acontecimentos elementares, sendo $m \leq n$, a probabilidade de A é dada pelo quociente entre o número de casos favoráveis a A e o número de casos possíveis, isto é,

$$P(A) = \frac{número \quad de \quad casos \quad favoráveis}{número \quad de \quad casos \quad possíveis} = \frac{m}{n}$$

Posteriormente, os autores da brochura apresentaram exemplos e sugestões de actividades para os professores desenvolverem na sala de aula. Na minha opinião, tanto os exemplos como as actividades propostas foram bastante bem concebidos, uma vez que permitem que o aluno apreenda integralmente a noção clássica de Probabilidade. Contudo, um dos exemplos apresentados foi resolvido com base nas regras de Combinatória. Apesar dos autores da brochura terem referido que o capítulo de Combinatória iria ser objecto de estudo posterior, penso que não deveriam ter optado por esta resolução, uma vez que os professores com pouca formação em Análise Combinatória poderão sentir dificuldades ao consultar esta resolução. Passo, então, a citar o exemplo que acabei de referir:

*"**Exemplo 17** – Numa caixa estão 4 pilhas boas e 3 estragadas. Retiram-se 3 pilhas ao acaso. Qual a probabilidade de se obterem 2 pilhas boas e uma estragada?*

Resolução:



O número de modos possíveis de retirar 3 pilhas da caixa = combinações de 7, 3 a 3, que é igual a $\frac{7!}{3!\,4!}$.

Destes resultados possíveis, nem todos são favoráveis, pois só nos interessam os que tenham 2 boas e 1 estragada. Os resultados favoráveis são dados pelas combinações de 4 pilhas 2 a 2, (2 pilhas boas) vezes combinações de 3 pilhas, 1 a 1 (1 pilha estragada) : $\frac{4!}{2!2!}\times\frac{3!}{1!2!}$.

A probabilidade pretendida vem igual a $\dfrac{\frac{4!}{2!2!}\times\frac{3!}{1!2!}}{\frac{7!}{3!4!}} = 0.51$.*"* (p. 47)

De seguida sugiro uma resolução para este exemplo, sem recorrer à Análise Combinatória.

Vamos começar por identificar as quatro pilhas boas por B1, B2, B3 e B4 e as três pilhas estragadas por E1, E2 e E3. De seguida vamos construir o espaço de resultados associado à experiência aleatória que consiste em seleccionar três pilhas ao acaso de entre as sete pilhas. Um processo simples que nos permite obter todas as possibilidades é construir uma tabela tridimensional:

		B1	B2	B3	B4	E1	E2	E3
B1	B1							
	B2			B1B2B3	B1B2B4	B1B2E1	B1B2E2	B1B2E3
	B3		B1B3B2		B1B3B4	B1B3E1	B1B3E2	B1B3E3
	B4		B1B4B2	B1B4B3		B1B4E1	B1B4E2	B1B4E3
	E1		B1E1B2	B1E1B3	B1E1B4		B1E1E2	B1E1E3
	E2		B1E2B2	B1E2B3	B1E2B4	B1E2E1		B1E2E3
	E3		B1E3B2	B1E3B3	B1E3B4	B1E3E1	B1E3E2	
B2	B1			B2B1B3	B2B1B4	B2B1E1	B2B1E2	B2B1E3
	B2							
	B3	B2B3B1			B2B3B4	B2B3E1	B2B3E2	B2B3E3
	B4	B2B4B1		B2B4B3		B2B4E1	B2B4E2	B2B4E3
	E1	B2E1B1		B2E1B3	B2E1B4		B2E1E2	B2E1E3
	E2	B2E2B1		B2E2B3	B2E2B4	B2E2E1		B2E2E3
	E3	B2E3B1		B2E3B3	B2E3B4	B2E3E1	B2E3E2	
B3	B1		B3B1B2		B3B1B4	B3B1E1	B3B1E2	B3B1E3
	B2	B3B2B1			B3B2B4	B3B2E1	B3B2E2	B3B2E3
	B3							
	B4	B3B4B1	B3B4B2			B3B4E1	B3B4E2	B3B4E3
	E1	B3E1B1	B3E1B2		B3E1B4		B3E1E2	B3E1E3
	E2	B3E2B1	B3E2B2		B3E2B4	B3E2E1		B3E2E3
	E3	B3E3B1	B3E3B2		B3E3B4	B3E3E1	B3E3E2	
B4	B1		B4B1B2	B4B1B3		B4B1E1	B4B1E2	B4B1E3
	B2	B4B2B1		B4B2B3		B4B2E1	B4B2E2	B4B2E3
	B3	B4B3B1	B4B3B2			B4B3E1	B4B3E2	B4B3E3
	B4							
	E1	B4E1B1	B4E1B2	B4E1B3			B4E1E2	B4E1E3
	E2	B4E2B1	B4E2B2	B4E2B3		B4E2E1		B4E2E3
	E3	B4E3B1	B4E3B2	B4E3B3		B4E3E1	B4E3E2	

E1	B1		E1B1B2	E1B1B3	E1B1B4		E1B1E2	E1B1E3
	B2	E1B2B1		E1B2B3	E1B2B4		E1B2E2	E1B2E3
	B3	E1B3B1	E1B3B2		E1B3B4		E1B3E2	E1B3E3
	B4	E1B4B1	E1B4B2	E1B4B3			E1B4E2	E1B4E3
	E1							
	E2	E1E2B1	E1E2B2	E1E2B3	E1E2B4			E1E2E3
	E3	E1E3B1	E1E3B2	E1E3B3	E1E3B4		E1E3E2	
E2	B1		E2B1B2	E2B1B3	E2B1B4	E2B1E1		E2B1E3
	B2	E2B2B1		E2B2B3	E2B2B4	E2B2E1		E2B2E3
	B3	E2B3B1	E2B3B2		E2B3B4	E2B3E1		E2B3E3
	B4	E2B4B1	E2B4B2	E2B4B3		E2B4E1		E2B4E3
	E1	E2E1B1	E2E1B2	E2E1B3	E2E1B4			E2B4E3
	E2							
	E3	E2E3B1	E2E3B2	E2E3B3	E2E3B4	E2E3E1		
E3	B1		E3B1B2	E3B1B3	E3B1B4	E3B1E1	E3B1E2	
	B2	E3B2B1		E3B2B3	E3B2B4	E3B2E1	E3B2E2	
	B3	E3B3B1	E3B3B2		E3B3B4	E3B3E1	E3B3E2	
	B4	E3B4B1	E3B4B2	E3B4B3		E3B4E1	E3B4E2	
	E1	E3E1B1	E3E1B2	E3E1B3	E3E1B4		E3E1E2	
	E2	E3E2B1	E3E2B2	E3E2B3	E3E2B4	E3E2E1		
	E3							

O número de maneiras possíveis de retirar três pilhas da caixa é 210.

Destes resultados possíveis, apenas 108 são favoráveis, pois só nos interessam os que tenham duas pilhas boas e uma estragada.

Assim, a probabilidade pretendida é igual a $\frac{108}{210} = 0{,}51$

Este exemplo poderá ser utilizado na sala de aula com o intuito de mostrar aos alunos que Cálculo Combinatório se torna muito útil quando a contagem directa é muito morosa devido ao elevado número de possibilidades em causa.

3.1.5.3. Aproximação Subjectiva de Probabilidade

O programa de Matemática A do 12.º ano apenas faz referência à aproximação frequencista de probabilidade e à definição clássica de probabilidade. No entanto, os autores da brochura de Probabilidades e Combinatória também fizeram uma breve abordagem à aproximação subjectiva de probabilidade. Como esta brochura tem como objectivo apoiar os professores de Matemática, penso que, de facto, é importante alertá-los para a existência de uma teoria subjectiva de probabilidade, que é válida sempre que fazemos o nosso próprio julgamento sobre um determinado acontecimento.

3.1.6. Definição Axiomática de Probabilidade

Se consultar a brochura de Probabilidades e Combinatória para analisar a definição axiomática de Probabilidade, encontro o seguinte:

> *"Considere-se um espaço de resultados **S**, finito, e um conjunto W de subconjuntos de S (acontecimentos) que satisfaçam as seguintes condições:*
>
> *a) Se um acontecimento A está em W, então o seu complementar $\overline{A}$ também está em W.*
> *b) Se dois acontecimentos estão em W, então a sua união também está em W.*
>
> *Dado o par **(S, W**), a que chamamos espaço de acontecimentos, a cada elemento **A**$\in$ W, associa-se um número que se chama **Probabilidade** e se representa por **P(A)**. As probabilidades associadas aos elementos de um espaço de acontecimentos satisfazem as seguintes condições ou **axiomas**:*
>
> ***1º axioma** - A probabilidade de qualquer acontecimento é sempre maior ou igual a zero*
>
> $$P(A) \geq 0$$
>
> ***2º axioma** - A probabilidade do acontecimento certo - S, é 1*
>
> $$P(S) = 1$$
>
> ***3º axioma** - Dados dois acontecimentos disjuntos, a probabilidade da sua união é igual à soma das probabilidades de cada um*
>
> $$\text{Se } A \cap B = \varnothing \quad \Rightarrow \quad P(A \cup B) = P(A) + P(B)\text{"}$$
>
> (p.52 e p.53)

Será que um professor com pouca formação em Probabilidades sabe a definição de espaço de acontecimentos?

Uma vez que o conceito de espaço de acontecimentos não é abordado no Ensino Secundário e os autores da brochura não o definiram, penso que não o deveriam ter invocado. Como tal, poderiam ter apresentado o seguinte:

> Considere-se um espaço de resultados, S, finito e um conjunto W de subconjuntos de S (acontecimentos) que satisfaçam as seguintes condições:
>
> a) Se um acontecimento A está em W, então o seu complementar $\overline{A}$ também está em W.
>
> b) Se dois acontecimentos estão em W, então a sua união também está em W.

A cada elemento A de W associamos um número a que chamamos **probabilidade de A** (representa-se por P(A)) e que satisfaz as seguintes condições ou axiomas:

1º axioma - A probabilidade de qualquer acontecimento é sempre maior ou igual a zero

$$P(A) \geq 0$$

2º axioma - A probabilidade do acontecimento certo, S, é 1

$$P(S) = 1$$

3º axioma - Dados *dois acontecimentos disjuntos*, a probabilidade da sua união é igual à soma das probabilidades de cada um

Se $A \cap B = \varnothing$ => $P(A \cup B) = P(A) + P(B)$

3.1.6.1. Propriedades da Probabilidade

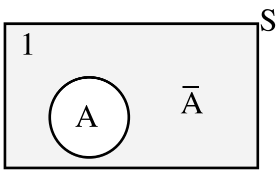
No que diz respeito às propriedades da Probabilidade, penso que os autores da brochura de apoio ao programa poderiam ter sido mais rigorosos nas demonstrações dos teoremas. Passo a citar o que os autores da brochura apresentaram:

"Com a ajuda de diagramas de Venn, e tendo em consideração os axiomas das Probabilidades, facilmente se mostram as seguintes propriedades para a Probabilidade:

***1** – $P(\varnothing) = 0$*

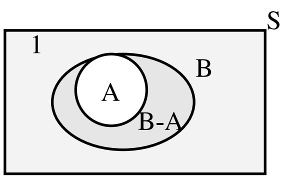
$$P(S \cup \phi) = P(S) + P(\phi)$$
$$= 1$$

***2** – $P(\overline{A}) = 1 - P(A)$*



$$P(A \cup \overline{A}) = P(A) + P(\overline{A})$$
$$= 1$$

***3** – Se $A \subset B$ então $P(A) \leq P(B)$*

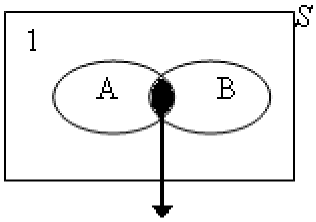


$$P(B) = P(A) + P(B - A)$$
$$\geq 0$$

4 – Qualquer que seja o acontecimento A, 0 ≤ P(A) ≤ 1

Corolário do resultado anterior.
***5** – Quaisquer que sejam os acontecimentos A e B,*

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



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

$A \cap B$

Para obter o resultado anterior basta ter em consideração que se pode exprimir a união dos acontecimentos A e B em termos de acontecimentos disjuntos, para então se aplicar o axioma 3:

$$\mathrm{A} \cup \mathrm{B} \ = \ (\mathrm{A} - \mathrm{B}) \cup (\mathrm{A} \cap \mathrm{B}) \cup (\mathrm{B} - \mathrm{A})$$

tendo ainda em conta que P(A-B)= P(A) – P(A∩ B).” (p. 54 e p. 55)

Na minha opinião, os teoremas anteriores poderiam ter sido demonstrados da seguinte forma:

Teorema: A probabilidade do acontecimento impossível é 0.

Demonstração: Como os acontecimentos $\varnothing$ e Ω são contrários, temos

$$\varnothing \cap \Omega = \varnothing \quad \text{e} \quad \varnothing \cup \Omega = \Omega$$

e, então, pelo terceiro axioma,

$$\mathrm{P}(\varnothing \cup \Omega) = \mathrm{P}(\varnothing) + \mathrm{P}(\Omega)$$

e, pelo segundo axioma, P(Ω) = 1, donde,

$$\mathrm{P}(\varnothing) + \mathrm{P}(\Omega) = \mathrm{P}(\Omega) \iff \mathrm{P}(\varnothing) + 1 = 1 \iff \mathrm{P}(\varnothing) = 0 \qquad \square$$

Teorema: Sendo *A* um acontecimento qualquer, a probabilidade do acontecimento contrário de *A* é $P(\overline{A}) = 1 - P(A)$

Demonstração:

Como, por hipótese, $\overline{A}$ é o acontecimento contrário de A, então, por definição A e $\overline{A}$ são acontecimentos disjuntos e $A \cup \overline{A} = \Omega$.

Então, como

$P(A \cup \overline{A}) = P(A) + P(\overline{A})$, pelo terceiro axioma,

$P(\Omega) = 1$, pelo segundo axioma

vem

$P(A \cup \overline{A}) = P(\Omega) \quad \Leftrightarrow \quad P(A) + P(\overline{A}) = 1 \quad \Leftrightarrow \quad P(\overline{A}) = 1 - P(A)$ □

Teorema: Se A_1, A_2, $A_3, \ldots, A_n$ (com $n \in IN$) são acontecimentos disjuntos dois a dois, então

$$P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n)$$

Demonstração:

$P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n) = P[A_1 \cup (A_2 \cup A_3 \cup \ldots \cup A_n)]$

$= P(A_1) + P(A_2 \cup A_3 \cup \ldots \cup A_n)$, por A_1 e $A_2 \cup A_3 \cup \ldots \cup A_n$ serem disjuntos

$= P(A_1) + P[A_2 \cup (A_3 \cup \ldots \cup A_n)]$

$= P(A_1) + P(A_2) + P[A_3 \cup (A_4 \cup \ldots \cup A_n)]$, por A_2 e $A_3 \cup \ldots \cup A_n$ serem disjuntos

E assim sucessivamente até se obter

$P(A_1 \cup A_2 \cup A_3 \cup \ldots \cup A_n) = P(A_1) + P(A_2) + P(A_3) + \ldots + P(A_n)$ □

Teorema: Dados dois acontecimentos *A* e *B* quaisquer,
Se $A \subset B$ então $P(A) \leq P(B)$

Demonstração:

Como $B = A \cup (B \cap \overline{A})$ e os acontecimentos A e $B \cap \overline{A}$ são disjuntos, vem

$$P(B) = P(A) + P(B \cap \overline{A})$$

e, pelo axioma 1, $P(B \cap \overline{A}) \geq 0$, portanto

$P(B) \geq P(A) \quad \Leftrightarrow \quad P(A) \leq P(B)$ □

Teorema: $0 \leq P(A) \leq 1$, qualquer que seja o acontecimento A.

Demonstração:

$P(A) \geq 0$, pelo primeiro axioma.

Como $A \subset \Omega$, $P(A) \leq P(\Omega)$, pelo teorema anterior.

Logo, pelo segundo axioma, $P(A) \leq 1$. □

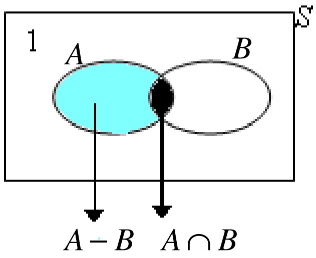
Teorema: Quaisquer que sejam os acontecimentos A e B ,

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

Demonstração:

Observando o diagrama de Venn, facilmente concluímos que

$A = (A-B) \cup (A \cap B)$



Como os acontecimentos $A-B$ e $A \cap B$ são disjuntos, conclui-se que $P(A) = P(A-B) + P(A \cap B)$.

Portanto, $P(A-B) = P(A) - P(A \cap B)$ □

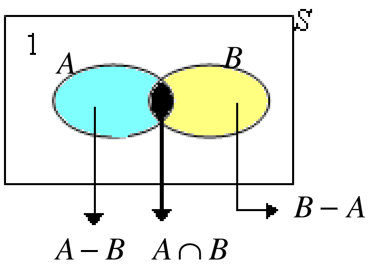
Teorema: Quaisquer que sejam os acontecimentos A e B ,

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

Demonstração:

Observado o diagrama de Venn, facilmente concluímos que

$A \cup B = (A-B) \cup (A \cap B) \cup (B-A)$



Como os acontecimentos $A-B$, $A \cap B$ e $B-A$ são disjuntos dois a dois, obtemos

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

Pelo teorema anterior, tem-se que

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

Substituindo em (*), vem:

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

Reduzindo os termos semelhantes, obtemos:

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

Depois de terem apresentado as propriedades da Probabilidade, os autores da brochura apresentaram algumas actividades sobre o cálculo de probabilidades e sobre a aplicação das suas propriedades. Passo a citar uma das actividades propostas sobre o cálculo de probabilidades:

"Uma caixa de disquetes tem 10 disquetes, das quais 3 são defeituosas. Retiram--se aleatoriamente 2 disquetes da caixa (extracção sem reposição). Calcule a probabilidade dos acontecimentos:

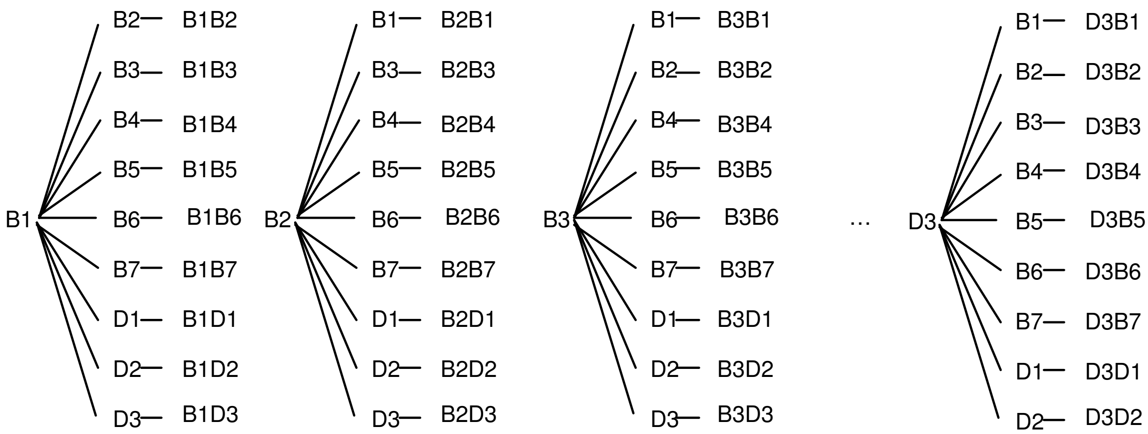
A - Obter uma só defeituosa;

B - Pelo menos uma defeituosa;

C - Nenhuma defeituosa.

Resolução:

Vamos começar por obter o espaço dos resultados, que é constituído por todas as extracções possíveis de 2 das 10 disquetes que tem a caixa. Para obter todas as maneiras possíveis vamos ter que arranjar cada possibilidade correspondente à 1ª extracção, com cada possibilidade correspondente à 2ª extracção, pelo que um processo simples é construir uma árvore de probabilidades:



No esquema anterior representámos por B1, B2, …, B7 as disquetes boas, e por D1, D2, D3 as defeituosas. Quando se extraem 2 disquetes simultaneamente, é o mesmo que retirar uma disquete e em seguida retirar outra, sem repor a primeira. Assim, a primeira disquete a sair pode ser qualquer das 10 existentes na caixa, enquanto que a segunda disquete pode ser uma qualquer, diferente da primeira. Temos 90 possibilidades diferentes de extrairmos as 2 disquetes da caixa. Qualquer destas possibilidades tem a mesma probabilidade de se verificar, pelo que o espaço de resultados associado a esta experiência é constituído por 90 resultados todos eles igualmente possíveis. Para calcular as probabilidades pretendidas, basta ver quantos são os resultados favoráveis aos acontecimentos em causa.

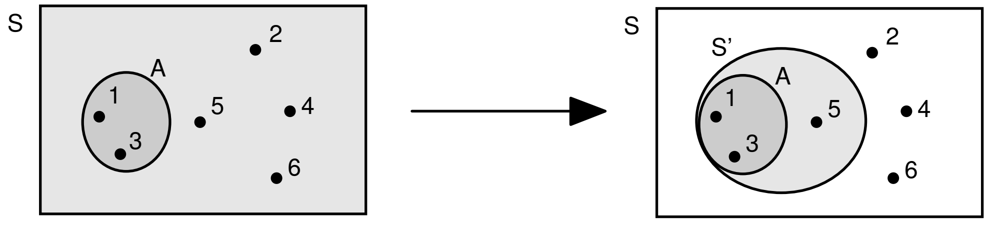
Seja A o acontecimento "obter uma só defeituosa". Este acontecimento é constituído por 42 resultados pelo que P(A)=42/90=7/15.
O número de resultados do acontecimento B, "pelo menos uma defeituosa" é 48, pelo que a probabilidade pretendida é P(B)=48/90=24/35.
Repare-se que o acontecimento C, "nenhuma defeituosa" é o complementar de "pelo menos uma defeituosa" pelo que P(C)= 1-24/35=11/35." (p. 57 e 58)

Neste exercício, os autores apresentaram um erro de cálculo pois, $P(B)=\frac{48}{90}=\frac{24}{45}$ e $P(C)=1-\frac{24}{45}=\frac{7}{15}$.

3.1.7. Probabilidade Condicional e Independência. Probabilidade da Intersecção de Acontecimentos. Acontecimentos Independentes

No que se refere à Probabilidade Condicional, os autores da brochura de Probabilidades e Combinatória introduziram este conceito através dos seguintes exemplos:

"Consideremos, por exemplo, a experiência aleatória que consiste em lançar um dado e verificar o número de pintas que sai. A probabilidade do acontecimento A, sair "1 ou 3 pintas" é 2/6, já que o nosso espaço de resultados S, é constituído por 6 casos igualmente possíveis, dos quais 2 são favoráveis à realização de A. Se no entanto pretendermos a probabilidade desse mesmo acontecimento, sabendo de antemão que saiu um número de pintas ímpar, neste momento já o espaço de resultados S', é constituído por 3 resultados, igualmente possíveis, dos quais 2 são favoráveis, pelo que a probabilidade pretendida é 2/3, o dobro da obtida anteriormente, quando não tínhamos nenhuma informação. Exemplificando com um diagrama de Venn



Vejamos ainda uma outra situação. Suponhamos, por exemplo, a experiência aleatória que consiste em retirar 2 bolas sem reposição, de uma caixa contendo 4 *bolas brancas B1, B2, B3 e B4 e 3 bolas pretas P1, P2, P3. Os **N** diferentes resultados obtidos na realização da experiência são:*

B1B2	*B1B3*	*B1B4*	*B1P1*	*B1P2*	*B1P3*
B2B1	*B2B3*	*B2B4*	*B2P1*	*B2P2*	*B2P3*
B3B1	*B3B2*	*B3B4*	*B3P1*	*B3P2*	*B3P3*
B4B1	*B4B2*	*B4B3*	*B4P1*	*B4P2*	*B4P3*
P1B1	*P1B2*	*P1B3*	*P1B4*	*P1P2*	*P1P3*
P2B1	*P2B2*	*P2B3*	*P2B4*	*P2P1*	*P2P3*
P3B1	*P3B2*	*P3B3*	*P3B4*	*P3P1*	*P3P2*

Representando por n(Branca1) e n(Branca2), respectivamente, o número de vezes em que se verificou o acontecimento Branca1 – "saiu bola branca na 1ª extracção" e o número de vezes que se realizou o acontecimento Branca2 – "saiu bola branca na 2ª extracção", e por n(Branca1 $\cap$ Branca2) o número de vezes que se realizou o acontecimento Branca1 $\cap$ Branca2 – "saiu branca na 1ª e 2ª extracções", temos,
P(Branca1) = 24/42,   P(Branca2) = 24/42,   P(Branca1 $\cap$ Branca2) = 12/42

Suponhamos, no entanto, que sabíamos que tinha saído branca na 1ª extracção, isto é, que se tinha verificado o acontecimento Branca1. Qual a probabilidade de sair branca na 2ª extracção, isto é de se verificar o acontecimento Branca2, tendo em conta esta informação adicional? Neste momento o espaço de resultados foi substancialmente reduzido, pois o número de resultados possíveis é 24 (ter saído branca na 1ª extracção), dos quais só 12 é que são favoráveis, pelo que

P(Branca2 sabendo que Branca1) = 12/24

À probabilidade anterior chamamos probabilidade condicional do acontecimento Branca2, sabendo que (ou dado que) se realizou o acontecimento Branca1, e representamos por P(Branca2|Branca1).
Repare-se que

$$P(Branca2|Branca1) = \frac{n(\text{Branca1} \cap \text{Branca2})}{n(\text{Branca1})}$$

$$= \frac{\frac{n(\text{Branca1} \cap \text{Branca2})}{N}}{\frac{n(\text{Branca1})}{N}}$$

$$= \frac{P(\text{Branca1} \cap \text{Branca2})}{P(\text{Branca1})}$$

ou seja

$$P(Branca2|Branca1) = \frac{P(\text{Branca1} \cap \text{Branca2})}{P(\text{Branca1})}$$

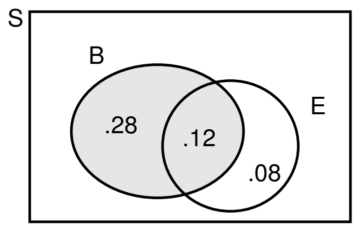
Assim, a probabilidade condicional de se realizar o acontecimento Branca2, sabendo que se realizou Branca1, é o quociente entre a probabilidade da realização de Branca1 e Branca2, e a probabilidade da realização de Branca1. Esta probabilidade condicional só tem sentido se P(Branca1) for superior a zero."

(p. 61 e 62)

Na minha opinião, estes exemplos para além de serem bastante simples, ajudam a compreender e a assimilar a definição de probabilidade condicional. Estes exemplos poderão ser uma boa sugestão metodológica para os professores utilizarem nas suas aulas como introdução ao conceito de Probabilidade Condicional.

Depois de ter sido apresentada a definição rigorosa de Probabilidade Condicional, os autores da brochura apresentaram diversos exemplos. Um dos exemplos apresentados foi o seguinte:

*"**Exemplo 19** (Siegel et al, 1988) –. Consideremos a experiência aleatória que consiste em observar, numa dada multinacional, a impressão causada (boa ou má) na entrevista dos candidatos a um emprego, assim como se conseguem ou não o emprego. Pensemos nos acontecimentos B – "o candidato causa boa impressão" e E – "o candidato consegue o emprego". Suponhamos que os acontecimentos anteriores estão representados num diagrama de Venn e que se conhecem as probabilidades assinaladas:*



No diagrama de Venn os números indicados representam:

$P(B–E) = 0.28$

$P(E–B) = 0.08$

$P(B \cap E) = 0.12$

A partir do diagrama anterior sabemos que

$$P(\text{“Conseguir emprego”}) = 0.12 + 0.08 = 0.20$$

o que significa que 20% dos candidatos, que vão à entrevista, conseguem o emprego. Será que o facto de causar boa impressão, aumenta as possibilidades de ser bem sucedido, na obtenção do emprego? Isto é, será que a informação adicional de que "um candidato causou boa impressão" tem efeito na probabilidade de obter o emprego? Para responder a esta questão, temos de nos cingir unicamente aos candidatos que causam boa impressão, em vez de considerarmos todos os candidatos. A dimensão deste grupo é 40% de todos os candidatos, já que P("Causar boa impressão") = 0.28 + 0.12 = 0.40

Dentro deste grupo, quantos conseguem o emprego? A resposta obtém-se restringindo este grupo aos que conseguem o emprego

*P("Causar boa impressão **e** Conseguir o emprego") = 0.12*

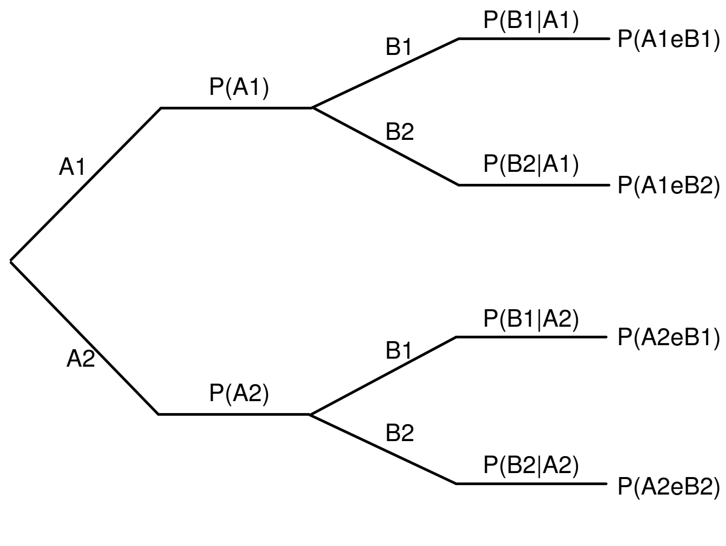
Finalmente podemos calcular a probabilidade de uma pessoa que causou boa impressão, conseguir o emprego. Esta probabilidade é dada pela resposta à seguinte questão "0.12 que percentagem é de 0.40"? resposta esta que se obtém dividindo 0.12 por 0.40, como aliás se deduz da definição anteriormente dada de probabilidade condicional:

$$P(\text{"Conseguir o emprego"} \mid \text{"Causou boa impressão"}) = \frac{0.12}{0.40} = 0.30$$

Vemos que a probabilidade de conseguir o emprego aumentou de 20% para 30%, com a informação adicional disponível. Isto significa que 30% dos candidatos que causam boa impressão, conseguem o emprego, comparados com unicamente 20% dos candidatos em geral (causando ou não boa impressão). Intuitivamente esperávamos que o facto de um candidato causar boa impressão, aumentasse as suas possibilidades de sucesso, e o que acabamos de medir foi precisamente quão grande é esse efeito." (p. 63, 64 e 65)

Posteriormente, os autores referiram o seguinte:

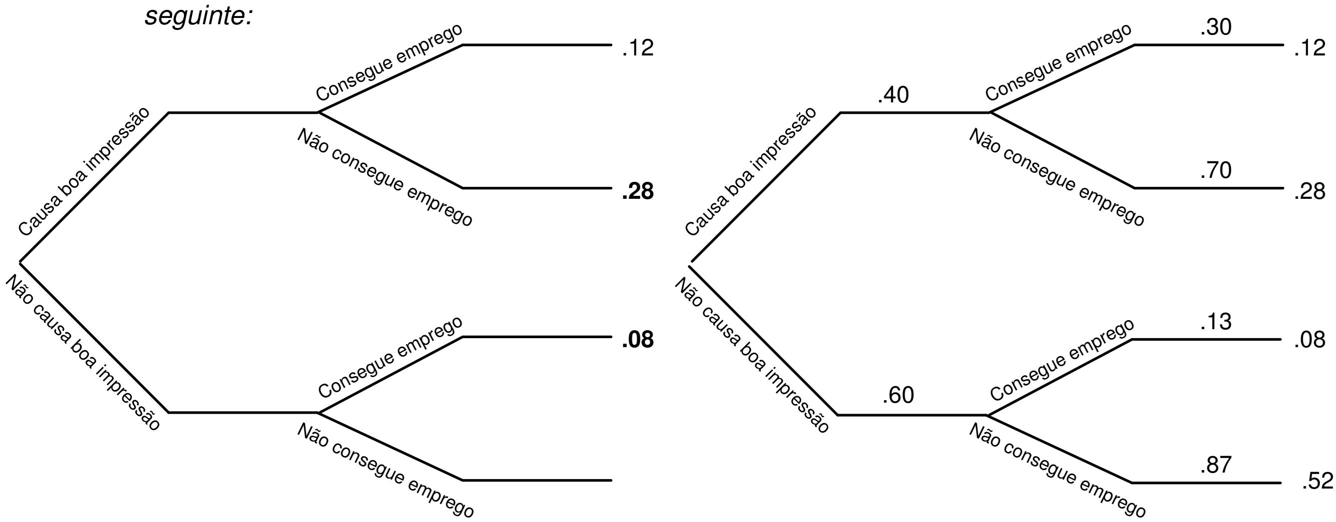
Um outro processo de visualizar as probabilidades condicionais é através da construção de uma árvore de probabilidades, que já vimos ser uma técnica útil quando pretendemos calcular probabilidades de acontecimentos associados a experiências aleatórias que envolvam vários passos. Por exemplo se num primeiro passo se puderem verificar um de dois acontecimentos A1 ou A2 e num segundo passo um dos dois acontecimentos B1 ou B2, podemos construir a seguinte árvore de probabilidades:



(Retirado da brochura de Probabilidades e Combinatória, 1999, p.65)

Com base no que acabei de citar, os autores da brochura desenvolveram o exemplo anterior:

> *"**Exemplo 19** (cont.) – Representando num diagrama em árvore as probabilidades apresentadas no diagrama de Venn, obtemos a árvore da esquerda no esquema seguinte:*



> *A partir da árvore da esquerda conseguimos facilmente completar a árvore da direita, utilizando a seguinte metodologia: inscrevemos na árvore o valor 0.40 que corresponde à probabilidade de "causar boa impressão", seguindo-se o valor 0.60, que corresponde à probabilidade de "não causar boa impressão". Uma vez colocado este valor, coloca-se o valor 0.52, que corresponde à probabilidade de "não causar boa impressão e não conseguir o emprego". Finalmente preenchem--se os ramos correspondentes às probabilidades condicionais, dividindo cada valor no extremo da árvore pelo valor do ramo anterior."* (p. 66)

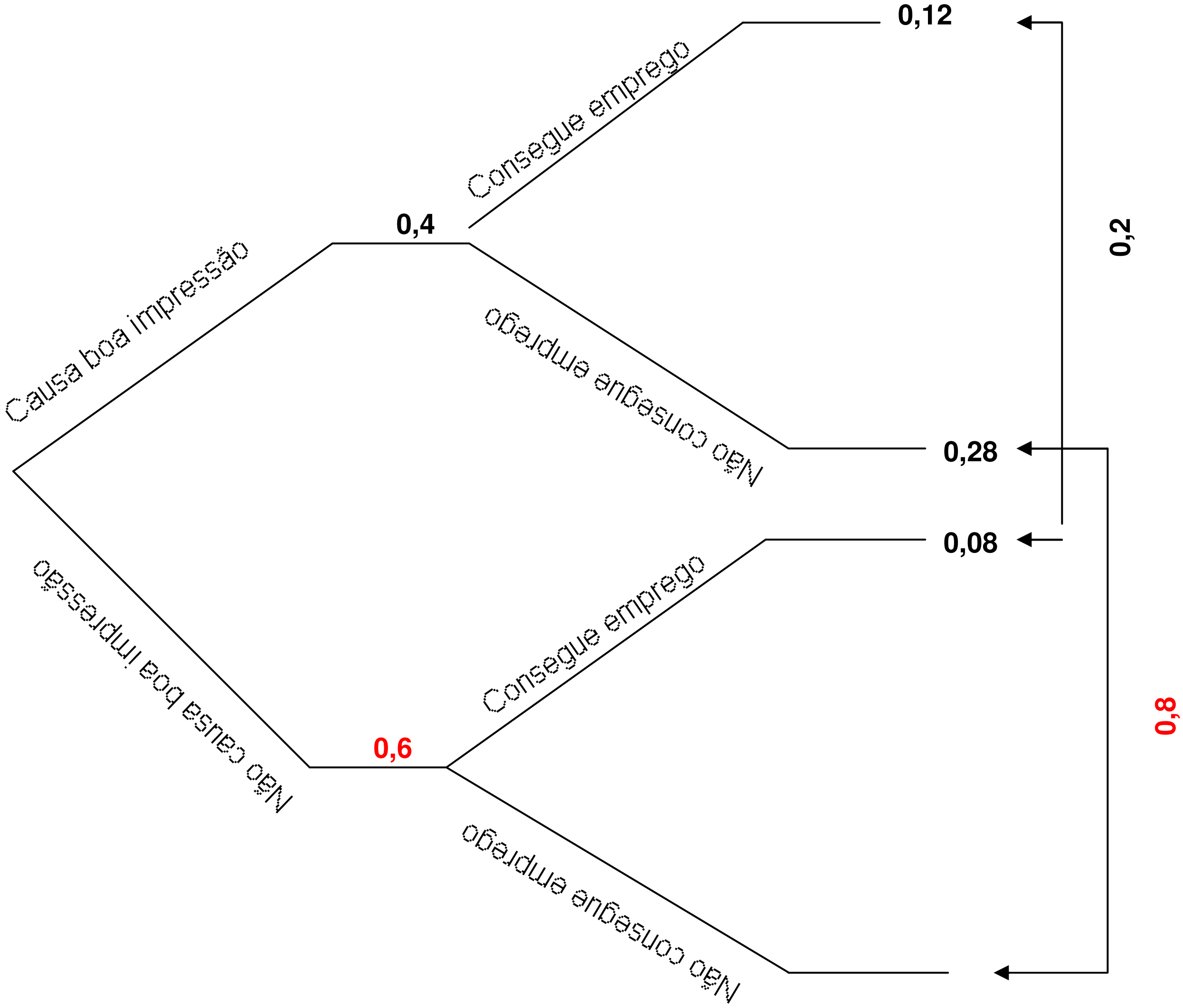
Será que um professor com pouca formação em Probabilidades irá considerar a construção do diagrama em árvore simples e óbvia tal como os autores da brochura referiram? Na minha opinião, os autores da brochura deveriam ter construído o diagrama em árvore da seguinte forma:

Comecemos por construir o diagrama em árvore com a informação dada inicialmente e ainda, com as informações seguintes:

$$\begin{aligned} P(\text{"}Não\ causar\ boa\ impressão\text{"}) &= 1 - P(\text{"}causar\ boa\ impressão\text{"}) \\ &= 1 - 0{,}40 \\ &= 0{,}6 \end{aligned}$$

e

$$P(\text{"Não consegue emprego"}) = 1 - P(\text{"consegue emprego"})$$
$$= 1 - 0{,}2$$
$$= 0{,}8$$



Na árvore temos dados suficientes para determinar as seguintes probabilidades:

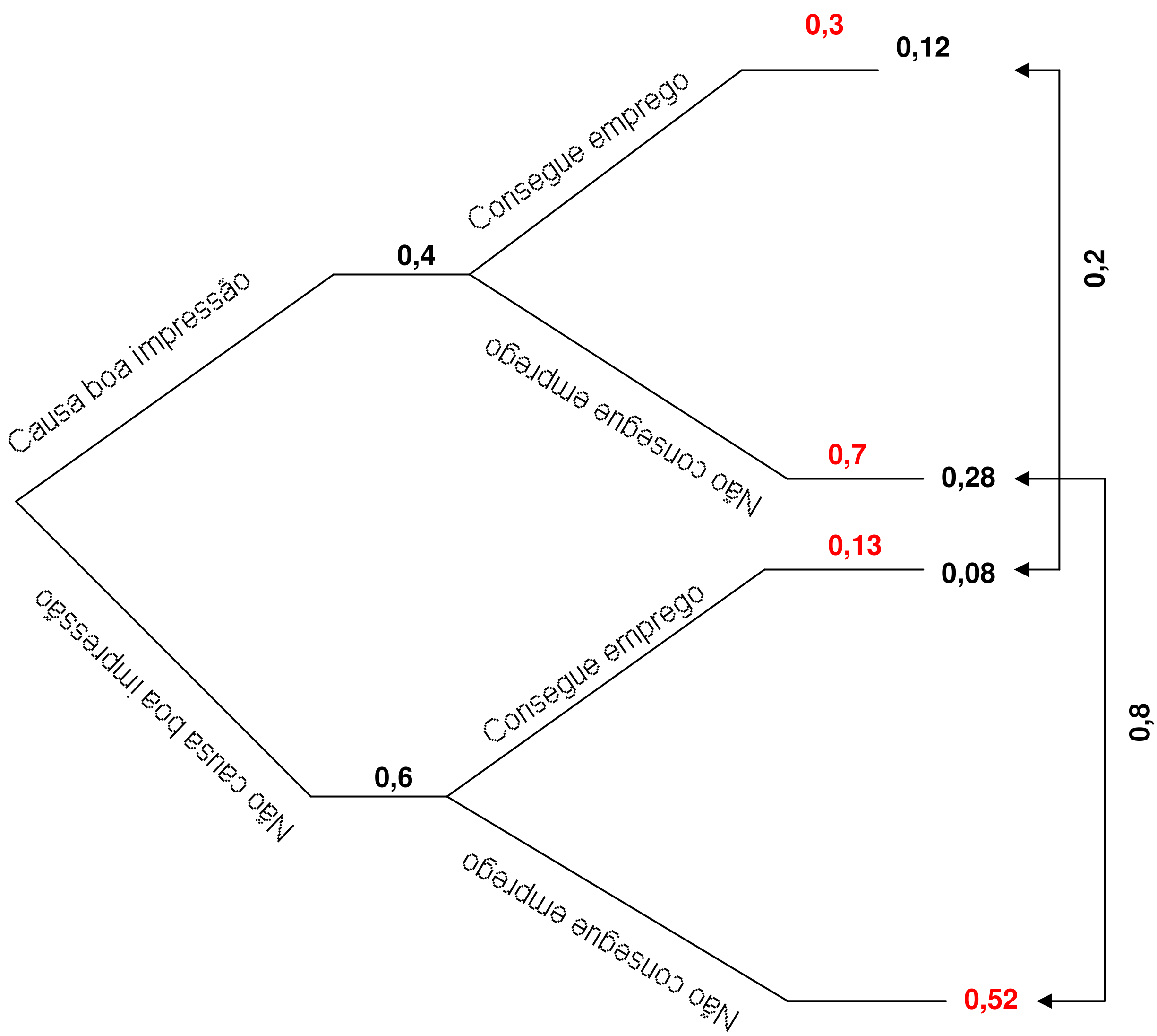
$$P(\text{"consegue emprego} \mid \text{causa boa impressão"}) =$$
$$= \frac{P(\text{"consegue emprego} \cap \text{causa boa impressão"})}{P(\text{"causa boa impressão"})}$$
$$= \frac{0{,}12}{0{,}4}$$
$$= 0{,}3$$

$$P(\text{"não consegue emprego} \mid \text{causa boa impressão"}) =$$
$$= 1 - P(\text{"consegue emprego} \mid \text{causa boa impressão"}) =$$
$$= 1 - 0{,}3$$
$$= 0{,}7$$

$$P(\text{"}consegue \quad emprego \mid não \quad causa \quad boa \quad impressão\text{"}) =$$
$$= \frac{P(\text{"}consegue \quad emprego \cap não \quad causa \quad boa \quad impressão\text{"})}{P(\text{"}não \quad causa \quad boa \quad impressão\text{"})}$$
$$= \frac{0{,}08}{0{,}6}$$
$$= 0{,}13$$

$$P(\text{"}não \quad consegue \quad emprego \cap não \quad causa \quad boa \quad impressão\text{"}) = 0{,}8 - 0{,}28$$
$$= 0{,}52$$

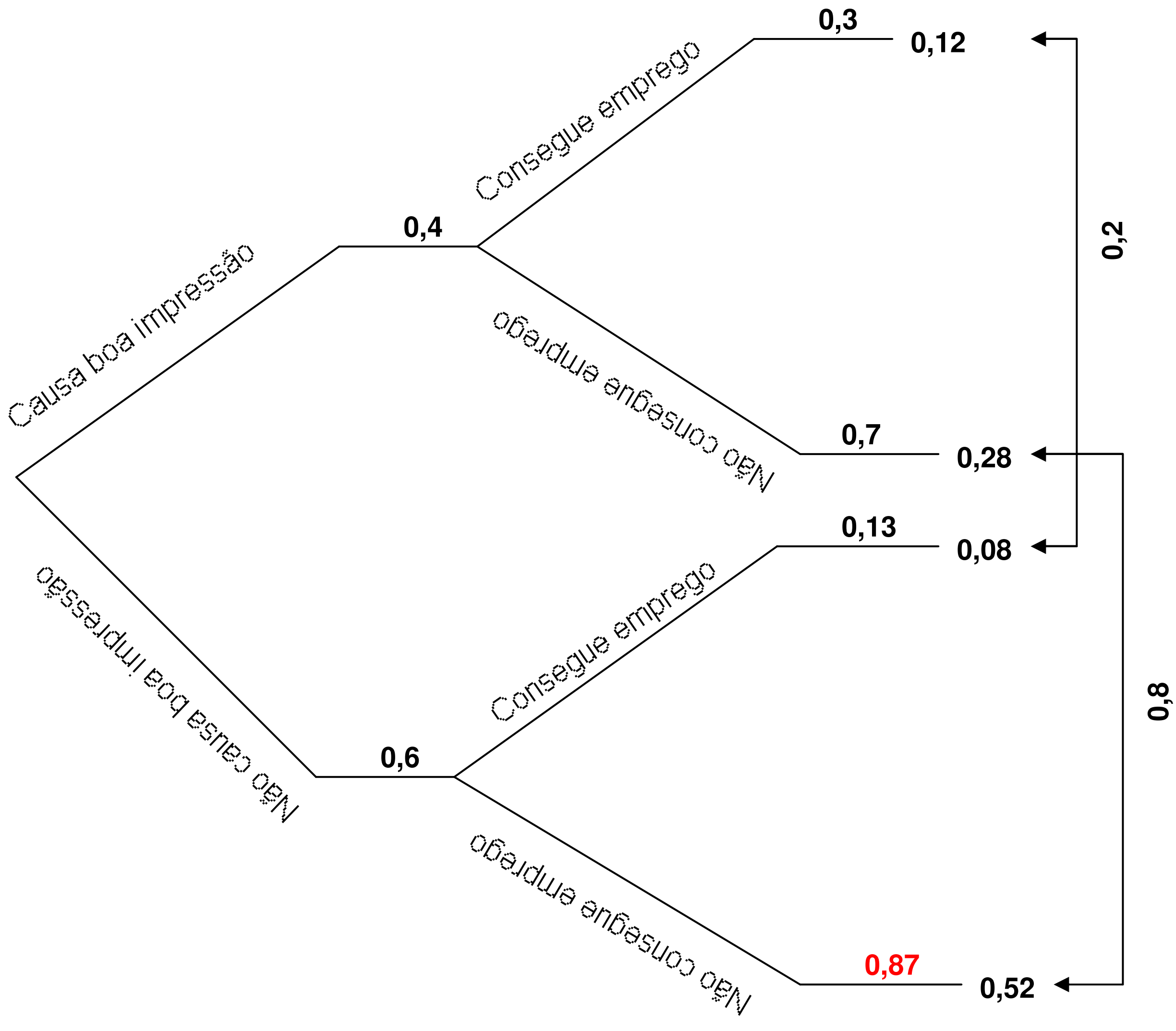
Coloquemos estas probabilidades no diagrama em árvore, tal como indica a figura seguinte:



Para completarmos o diagrama em árvores basta calcular a seguinte probabilidade condicional:

$$
\begin{aligned}
&P(\text{"}não \quad consegue \quad emprego \mid não \quad causa \quad boa \quad impressão\text{"}) = \\
&\quad = \frac{P(\text{"}não \quad consegue \quad emprego \cap não \quad causa \quad boa \quad impressão\text{"})}{P(\text{"}não \quad causa \quad boa \quad impressão\text{"})} \\
&\quad = \frac{0{,}52}{0{,}6} \\
&\quad = 0{,}87
\end{aligned}
$$

Desta forma completamos o diagrama em árvore:



Posteriormente, os autores da brochura de apoio aos professores chamaram a atenção para o seguinte:

"Na definição de probabilidade condicional, pode acontecer que o acontecimento que está a condicionar seja a intersecção de dois acontecimentos. Por exemplo

$$P(A|B \cap C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} = \frac{P(A \cap B \cap C)}{P(B)P(B|C)} \qquad (*)$$

donde

$$P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C) \text{ ou } P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B)$$

De um modo geral, dados n acontecimentos A_1, A_2, ..., A_n, com probabilidade positiva, tem-se

$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1)\, P(A_2|A_1)P(A_3|A_1 \cap A_2) \ldots P(A_n|A_1 \cap A_2 \cap \ldots \cap A_{n-1})$" (**)

(p. 70 e 71)

A expressão (*) não se encontra correcta, pois a definição de probabilidade condicional indica-nos que $P(B \cap C) = P(C)P(B \mid C)$. Assim, em vez da expressão (*) devia-se ter a expressão:

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

Além disso, penso que os autores da brochura poderiam ter referido que quando se tem a intersecção de *n* acontecimentos, geralmente, omite-se o sinal de intersecção, tal como nas expressões algébricas se omite o sinal de produto. Assim, a expressão (**) poderia ser substituída pela seguinte:

$$P(A_1\, A_2 \ldots A_n) = P(A_1)\, P(A_2|A_1)P(A_3|A_1A_2) \ldots P(A_n|A_1A_2 \ldots A_{n-1})$$

Esta expressão, geralmente, designa-se por ***regra da multiplicação*** ou ***regra da probabilidade de uma cadeia***.

Quanto aos acontecimentos independentes, penso que os autores da brochura de apoio ao professor poderiam ter mencionado o seguinte:

"A independência entre A e B implica a independência entre A e $\overline{B}$, entre $\overline{A}$ e B, e entre $\overline{A}$ e $\overline{B}$. Por exemplo,

$$P(A\overline{B}) = P(A - B) = P(A) - P(AB) = \quad \textit{(hipótese de independência entre A e B)}$$
$$= P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\overline{B})."$$

(Pestana e Velosa, 2006, p. 241)

Outro aspecto que também deveria ter sido referido na brochura de apoio, ainda relativamente aos acontecimentos independentes, é o seguinte:

Se cada um dos A_i for independente dos outros acontecimentos $A_1, A_2, ..., A_{i-1}$ então a expressão geral da probabilidade de uma cadeia de acontecimentos é simplificada da seguinte forma:

$$P(A_1A_2...A_{n-1}A_n) = P(A_1)P(A_2)...P(A_{n-1})P(A_n)$$

ou, de uma forma mais condensada: $P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$

Note-se que estas expressões só são válidas se os acontecimentos da cadeia forem independentes.

3.2. Distribuição de Frequências Relativas e Distribuição de Probabilidades

Na segunda secção do programa de Matemática A encontram-se os seguintes pontos:

- Variável aleatória; função massa de probabilidade:

  – distribuição de probabilidades de uma variável aleatória discreta; distribuição de frequências versus distribuição de probabilidades;

  – média versus valor médio;

  – desvio padrão amostral versus desvio padrão populacional.
- Modelo Binomial.
- Modelo Normal; histograma versus função densidade.

(Silva *et al.*, 2002c, p. 2)

3.2.1. Variável Aleatória. Função Massa de Probabilidade

Ao consultar a brochura de Probabilidades e Combinatória, encontrei a seguinte definição de variável aleatória:

> *"Variável aleatória – Uma variável aleatória, é uma variável cujo valor é um resultado numérico associado ao resultado de uma experiência aleatória."* (p. 78)

Se um professor de Matemática apresentar esta definição na aula, será que os alunos irão conseguir compreender o conceito de variável aleatória? Na minha opinião, seria mais fácil, tanto para os alunos como para os professores com pouca formação em Probabilidades, se o conceito de variável aleatória fosse apresentado da seguinte forma:

> *"Dada uma experiência aleatória à qual corresponde o espaço de resultados Ω, chama-se **variável aleatória X** a uma função que a cada elemento do espaço de resultados associa um número real.*
>
> $X:\Omega \longrightarrow IR$ " (Costa e Rodrigues, 2006, p.57)

As variáveis aleatórias podem ser de dois tipos: **discretas** ou **contínuas**.

No que diz respeito às variáveis aleatórias discretas, os autores da brochura referiram o seguinte:

> *"Dada uma variável aleatória discreta X, que assume um número finito de valores distintos* $x_1, x_2, ..., x_i, ..., x_N$ *, então as probabilidades* $p_i = P(X = x_i)$ *,* i *=1,...,N, devem satisfazer as seguintes condições:*
>
> *i)* $0 \le p_i \le 1$*,* i *=1,...,N*
>
> *ii)* $\sum_{i=1}^{N} p_i = 1$
>
> *Os valores (* xi, p_i *) constituem a distribuição de probabilidades de X"* (p.80)

Será que um professor com pouca formação em Probabilidades sabe o que é a distribuição de probabilidades de uma variável aleatória discreta? Na minha opinião, os autores da brochura de apoio aos professores deveriam ter referido o seguinte:

> *"Chama-se distribuição de probabilidades de uma variável aleatória* X *à aplicação que a cada valor* x_i *da variável* X *faz corresponder a respectiva probabilidade* p_i *."* (Neves *et al.*, 2006, p. 48)

A distribuição de probabilidades de uma variável aleatória discreta é também designada por função massa de probabilidades e geralmente, dispõe-se na forma:

$$X = \begin{cases} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{cases}$$

A representação gráfica da distribuição de probabilidades de uma variável aleatória discreta é o diagrama de barras em que cada uma das barras tem comprimento igual à probabilidade que a variável assume em cada um dos valores. A soma dos comprimentos de todas as barras é 1.

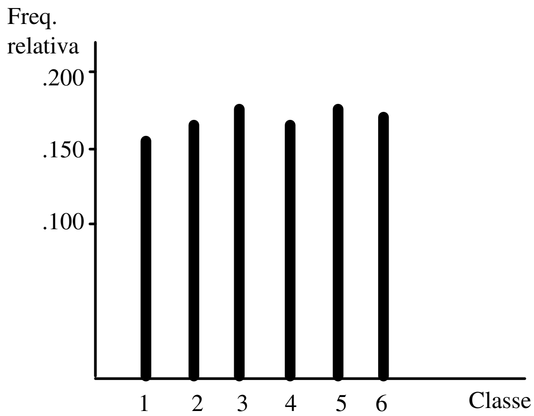
3.2.2. Distribuição de Probabilidade de uma Variável Aleatória Discreta

3.2.2.1. Distribuição de Frequências *vs.* Distribuição de Probabilidades

Para ilustrar a diferença entre a distribuição de frequências e a distribuição de probabilidades de uma variável aleatória discreta, os autores da brochura de Probabilidades e Combinatória apresentaram o seguinte exemplo:

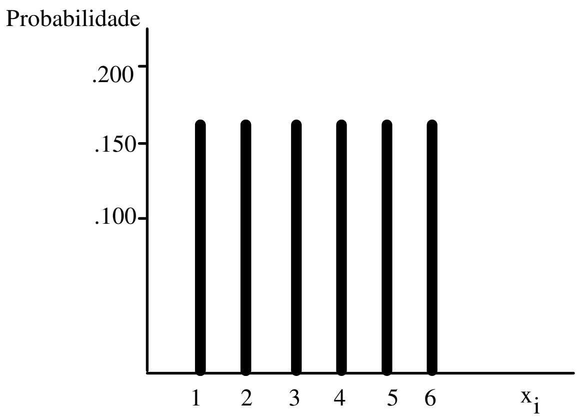
"(…) quando pretendemos resumir a informação contida num conjunto de dados discretos - observações de uma variável discreta, uma representação gráfica adequada é o diagrama de barras, também chamado distribuição de frequências. Suponhamos então que pretendíamos estudar a distribuição de probabilidades da variável aleatória que representa o número de pintas que se obtém no lançamento de um dado, que não temos a certeza de ser equilibrado. Para estudar esta População[1], constituída pelos valores que se podem obter no lançamento do dado, vamos recolher uma amostra de dimensão 1200, isto é vamos repetir a experiência de lançar o dado 1200 vezes. Suponhamos que os valores obtidos deram origem à seguinte tabela de frequências, onde representamos por n_i as frequências absolutas e por f_i as relativas, a partir da qual construímos o diagrama de barras:

Classe	n_i	f_i
1	185	0.154
2	198	0.165
3	208	0.173
4	195	0.163
5	209	0.174
6	205	0.171
Total	1200	1.00



Tendo em consideração a aproximação frequencista da probabilidade, em que vimos que se a dimensão da amostra for suficientemente grande, as frequências relativas podem ser interpretadas como valores aproximados para as probabilidades, o diagrama de barras sugere-nos que a hipótese do dado ser equilibrado é uma hipótese admissível e que poderemos admitir a seguinte distribuição de probabilidades para a variável aleatória X em estudo:

$X=x_i$	$P(X=x_i)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6
Total	1

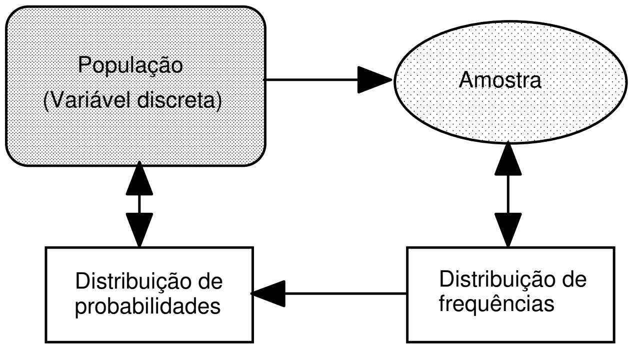


(p. 81 e 82)

[1] Passamos a identificar População com espaço de resultados

Na minha opinião, este exemplo é bastante perceptível, no entanto, a palavra "Classe" da tabela de frequências poderia ter sido substituída, por exemplo, por "Número de pintas".

De forma a resumir o processo descrito no exemplo anterior, os autores da brochura apresentaram o seguinte esquema:



(Retirado da Brochura de Probabilidades e Combinatória, 1999, p. 83)

3.2.2.2. Média *vs.* Valor Médio

Para introduzir o conceito de valor médio, os autores da brochura retomaram o exemplo anterior:

> *"A partir da tabela de frequências calculámos a média da amostra, utilizando a expressão da média para dados agrupados, que neste caso dá o valor exacto da média, por se tratar de uma amostra de dados discretos (as classes são os valores que surgem na amostra):*
>
> $$\bar{x} = 1 \times .154 + 2 \times .165 + 3 \times .173 + 4 \times .163 + 5 \times .174 + 6 \times .171 = 3.551$$
>
> *Se na expressão anterior substituirmos as frequências relativas pelos valores das probabilidades sugeridas para a distribuição de probabilidade da População de onde retirámos a amostra, vem*
>
> $$1 \times \overset{0.165}{\overset{\downarrow}{0.167}} + 2 \times \overset{0.154}{\overset{\downarrow}{0.167}} + 3 \times \overset{0.173}{\overset{\downarrow}{0.167}} + 4 \times \overset{0.163}{\overset{\downarrow}{0.167}} + 5 \times \overset{0.174}{\overset{\downarrow}{0.167}} + 6 \times \overset{0.171}{\overset{\downarrow}{0.167}} = 3.50$$
>
> *Ao substituirmos as frequências relativas pelas probabilidades obtemos uma característica análoga à média, mas agora relativa à variável aleatória, a que damos o nome de valor médio"* (p. 84)

Será que para calcularem o valor da média, os autores da brochura recorreram, de facto, à expressão da média para dados agrupados? Será que os dados estão, de facto, agrupados?

Posteriormente, os autores da brochura explicaram de uma forma bastante simples e clara a diferença entre média e valor médio. Além disso, a definição que os autores apresentaram de valor médio é também bastante perceptível, podendo esta ser utilizada nas aulas pelos professores.

Para finalizar esta secção, os autores apresentaram vários exemplos e várias sugestões de actividades que visam obter a distribuição de probabilidades e o valor médio de uma variável aleatória discreta. Na minha opinião, tanto os exemplos como as sugestões de actividades poderão ser aplicados em sala de aula.

3.2.2.3. Variância Amostral *vs.* Variância Populacional. Desvio Padrão Amostral *vs.* Desvio Padrão Populacional

Para introduzir o conceito de variância populacional e de desvio padrão populacional, os autores da brochura retomaram o exemplo que temos vindo a analisar:

"(…) voltando ao exemplo do dado que temos vindo a estudar, a variância da amostra obtida a partir da expressão considerada quando os dados se apresentam agrupados, é:

$s^2 = (1 - 3.551)^2 \text{x } 0.154 + (2 - 3.551)^2 \text{x } 0.165 + (3 - 3.551)^2 \text{x } 0.173 + (4 - 3.551)^2 \text{x } 0.163 + (5 - 3.551)^2 \text{x } 0.174 + (6 - 3.551)^2 \text{x } 0.171 = 2.875$

donde se obtém para o desvio padrão o valor

$$s = 1.696$$

Procedendo de forma análoga ao que fizemos anteriormente, em que substituímos as frequências relativas pelas probabilidades, e substituindo também a média pelo valor médio, obtemos o desvio padrão da população subjacente à amostra, a que chamamos desvio padrão populacional, e representamos por σ, para distinguir do desvio padrão amostral, calculado a partir da amostra:

$\sigma = 1.71$ " (p. 93)

Posteriormente, os autores da brochura apresentaram a seguinte definição:

"Define-se variância populacional, σ^2, de uma distribuição de probabilidades,

$(x_i, p_i), \quad i=1, 2, \ldots, N$

como sendo o valor que se obtém multiplicando cada resultado $(x_i - \mu)^2$ pela probabilidade $p_i = P(X=x_i)$, $i=1, 2, \ldots, N$ e adicionando os resultados obtidos:

$$\sigma^2 = \sum_{i=1}^{N} (x_i - \mu)^2 p_i \quad \text{"} \qquad \text{(p.93)}$$

Apesar dos autores da brochura terem apresentado a definição de variância populacional e de terem transmitido uma ideia de como se obtém o valor do desvio padrão populacional, penso que este último também deveria ter sido definido, pois trata-se de um conteúdo do programa de Matemática A do 12.º ano. Como tal, os autores da brochura de apoio aos professores deveriam ter apresentado o seguinte:

Define-se desvio padrão populacional, σ, de uma distribuição de probabilidades,

$(x_i, p_i), \quad i=1, 2, \ldots, N$

como sendo a raiz quadrada do valor da variância populacional, ou seja,

$$\sigma = \sqrt{\sigma^2}$$

3.2.3. Modelo Binomial

No que diz respeito ao Modelo Binomial, os autores da brochura de Probabilidades e Combinatória começaram por apresentar várias experiências aleatórias que gozam das seguintes propriedades:

"i) Considera-se à partida um número fixo n de observações, a que é usual chamar provas;

ii) As observações são independentes umas das outras;

iii) Em cada observação pode-se obter um de dois resultados possíveis, a que chamamos sucesso ou insucesso;

iv) A probabilidade de sucesso p, é constante de observação para observação.

À variável X, que representa o número de sucessos nas n provas chama-se variável aleatória com distribuição Binomial de parâmetros n e p." (p. 98)

Estas são sem dúvida as propriedades que caracterizam uma distribuição Binomial de parâmetros *n* e *p*. Contudo, penso que os autores poderiam ter referido que a distribuição Binomial de parâmetros *n* e *p* representa-se, geralmente, por B(n; p) e que cada realização (prova) é uma prova de Bernoulli.

De facto, se continuar a ler a brochura de apoio ao professor verifico que os autores da brochura fizeram referência à distribuição de Bernoulli. Passo, então, a citar a abordagem feita pelos autores:

> *"**Distribuição de Bernoulli***
> *Um caso particular da variável aleatória Binomial é o que se verifica quando n=1, obtendo-se a chamada variável de*

Nº sucessos – k	*0*	*1*
P(X=k)	*1-p*	*p*

> *Calcule o valor médio e a variância desta variável aleatória e verifique que são iguais, respectivamente a p e a p(1-p)."* (p.103)

Será que um professor com pouca formação em Probabilidades, ao ler este excerto, irá compreender que, se *X* for uma variável aleatória com distribuição Binomial de parâmetros *1* e *p* então *X* é uma variável aleatória com distribuição de Bernoulli de parâmetro *p*?

Na minha opinião, os autores da brochura de Probabilidades e Combinatória poderiam ter ilustrado a distribuição de Bernoulli através de um exemplo. Passo a apresentar um exemplo que, no meu ponto de vista, ilustra bem a distribuição de Bernoulli.

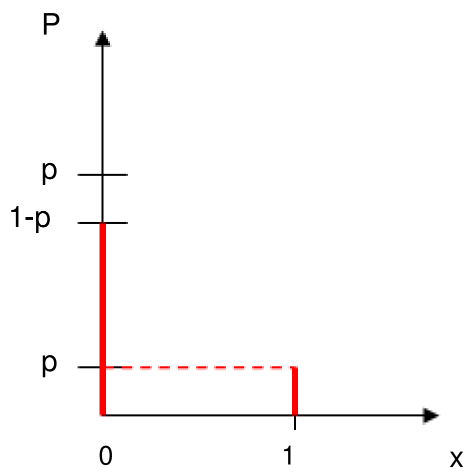
> *"Numa fábrica de parafusos estudou-se a qualidade do produto e concluiu-se que, com as actuais condições de fabrico, a probabilidade de encontrar um parafuso **com defeito** é de um em mil. Consideremos a experiência aleatória:*
>
> *"extrair da produção um parafuso, ao acaso, e verificar se é defeituoso."*
>
> *Esta experiência tem dois únicos resultados a que chamaremos:*
>
> - ***sucesso –*** S *: "o parafuso escolhido tem defeito"*
> - ***insucesso –*** $\overline{S}$ *: "o parafuso escolhido não tem defeito"*
>
> *Seja X a variável aleatória associada à experiência, tal que:*
> *X=1 quando "o parafuso escolhido tem defeito", ou seja, caso ocorra* S *;*
> *X=0 quando "o parafuso escolhido não tem defeito", ou seja, caso ocorra* $\overline{S}$ *.*

*A **distribuição de probabilidades** é definida pela tabela:*

$X = x_i$	0	1
$P(X = x_i)$	$1 - p = 0{,}999$	$p = 0{,}001$

Então, P(0) = 1 - p e P(1) = p.

O seu gráfico é:



*A qualquer variável aleatória X, com estas propriedades, chama-se **variável de Bernoulli**."* (Jorge *et al.*, 2006, p. 80 e 81)

"A variável aleatória X chama-se variável de Bernoulli de parâmetro p, se admite dois únicos valores (0, correspondente ao insucesso e 1, correspondente ao sucesso), sendo a distribuição de probabilidades assim definida:

P(1) = p

P(0) = 1 – p "

(Jorge *et al.*,2006, p. 81)

Na distribuição de Bernoulli, o valor médio, μ, e o desvio padrão populacional, σ, são dados por:

$$\mu = E(X) = 0 \times (1-p) + 1 \times p = p$$

$$\begin{aligned}\sigma &= \sqrt{(0-p)^2 \times (1-p) + (1-p)^2 \times p} \\ &= \sqrt{p^2(1-p) + (1-p)^2 p} \\ &= \sqrt{p(1-p)(p+1-p)} \\ &= \sqrt{p(1-p)}\end{aligned}$$

No final da secção referente à distribuição Binomial, os autores da brochura de apoio aos professores sugeriram uma actividade que ensina a utilizar a calculadora gráfica para obter a probabilidade de uma variável aleatória com distribuição Binomial. Na minha opinião, esta proposta de actividade é bastante importante, no entanto, há alunos a frequentar o 12.º ano sem calculadoras gráficas e também há muitas escolas que não têm calculadoras gráficas em número suficiente para facultar aos alunos.

3.2.4. O Modelo Normal

3.2.4.1. Histograma *vs.* Função Densidade

Nesta secção, os autores da brochura de Probabilidades e Combinatória, para além de terem procurado explicar a diferença entre histograma e função densidade de probabilidade, procuraram transmitir uma ideia de como se obtém a representação gráfica da função densidade a partir do histograma.

Uma vez que a brochura de Probabilidades e Combinatória tem como objectivo apoiar os professores com pouca formação na área das Probabilidades, penso que os autores da brochura também deveriam ter apresentado as propriedades da função densidade de probabilidade:

> Uma função $f(x)$ é uma função densidade de probabilidade de uma variável aleatória contínua X se:
>
> – $f(x) \geq 0$, para todo o x do intervalo em que está definida a variável aleatória;
>
> – a área abaixo da curva é 1;
>
> – a probabilidade de a variável tomar valores pertencentes ao intervalo $[x_i, x_j]$
>
> é igual à área abaixo da curva e corresponde ao intervalo $[x_i, x_j]$.

3.2.4.2. Modelo Normal

Se consultar a brochura de Probabilidades e Combinatória para estudar o modelo Normal, verifico que os autores da brochura introduziram o tema através do exemplo que se segue:

> *"Suponhamos, por exemplo, que estávamos interessados em estudar a característica altura da População constituída pelos indivíduos adultos, sexo masculino, de nacionalidade portuguesa. Identificando a População com os*

> *valores que a característica em estudo pode assumir, podemos dizer que estamos interessados em estudar a variável aleatória que representa a altura de um indivíduo escolhido ao acaso de entre os indivíduos adultos, sexo masculino, portugueses. Obviamente que esta variável aleatória já não é discreta, mas sim contínua."* (p.107) [O sublinhado é meu.]

Se um professor for consultar a brochura de Probabilidades e Combinatória para analisar o conceito de variável aleatória continua, não o irá encontrar. De facto, as indicações metodológicas do programa de Matemática A do 12.º ano referem que:

> *"Não é objectivo do programa entrar no estudo das variáveis aleatórias contínuas (….)"* (Silva *et al.*, 2002c, p. 2)

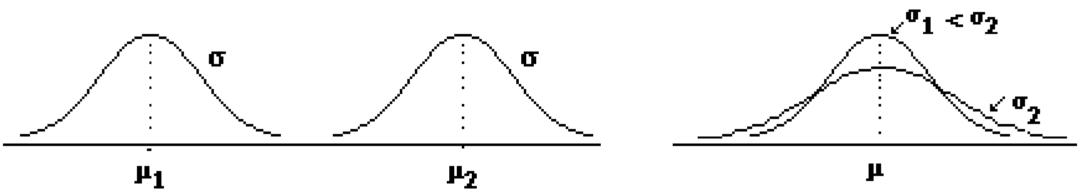
Mas, se um dos conteúdos do programa de Matemática A do 12.º ano é o modelo Normal, não fará sentido entrar no estudo das variáveis aleatórias contínuas? Na minha opinião, os autores da brochura de apoio aos professores deveriam ter referido o seguinte:

> *"As variáveis aleatórias que possam assumir todos os valores de um intervalo, sendo nula a probabilidade de assumirem valores isolados, dizem-se variáveis aleatórias **contínuas**. Enquanto que uma variável aleatória discreta se refere a qualquer tipo de contagem, uma v. a. contínua refere-se a uma medida, como por exemplo o peso, a altura, o tempo, etc.*
>
> *Exemplos:*
>
> - *tempo que um cliente espera numa "bicha" dum supermercado*
> - *peso de um bebé de 6 meses*
> - *tempo entre chegadas telefónicas consecutivas"*
>
> (Martins, 2000, p. 160)

A maior parte dos fenómenos do dia-a-dia que estão associados a variáveis aleatórias contínuas são traduzidas pela distribuição Normal. A representação gráfica da distribuição Normal chama-se **curva normal** ou **curva de Gauss** e tem propriedades muito importantes. Ao consultar a brochura de Probabilidades e Combinatória, encontrei o seguinte:

*“**Propriedades da curva normal**:*

– *É simétrica relativamente ao valor médio μ da variável, assumindo aí o valor máximo;*
– *Quanto maior for o desvio padrão σ, mais achatada é a curva;*



– *A área compreendida entre a curva e o eixo dos xx é igual a 1;*
– *A área compreendida entre a curva, o eixo dos xx e as rectas que passam pelos pontos μ-σ e μ+σ, é aproximadamente igual a 0.68;*
– *A área compreendida entre a curva, o eixo dos xx e as rectas que passam pelos pontos μ-2σ e μ+2σ, é aproximadamente igual a 0.95;*
– *A área compreendida entre a curva, o eixo dos xx e as rectas que passam pelos pontos μ-3σ e μ+3σ, é aproximadamente igual a 1.”*

(p. 111)

Para além das propriedades enunciadas em cima poderiam, também, ter sido apresentadas as seguintes:

– A curva é contínua.
– A concavidade da curva muda de sentido para $\mu-\sigma$ e $\mu+\sigma$, isto é, $\mu-\sigma$ e $\mu+\sigma$ são abcissas dos pontos de inflexão.
– O eixo das abcissas é assimptota da curva.
– A probabilidade de que a variável tome valores no intervalo [x_i ; x_j] é igual à área definida pelo eixo dos xx, pelo gráfico de função densidade e pelas recta $x = x_i$ e $x = x_j$.

No seguimento das propriedades da curva Normal, os autores da brochura de Probabilidades e Combinatória referiram o seguinte:

*“se **X** for Normal:*

$P(\mu - \sigma \leq X \leq \mu + \sigma)$=.683

$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$=.954

$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$=.997 ” (p. 111)

Uma vez que a probabilidade da variável aleatória contínua assumir valores isolados é nula, não há diferença para o cálculo da probabilidade de um intervalo, caso esse intervalo seja aberto ou fechado. Como tal,

$$P(\mu-\sigma \le X \le \mu+\sigma) = P(\mu-\sigma < X \le \mu+\sigma) = P(\mu-\sigma \le X < \mu+\sigma) =$$
$$= P(\mu-\sigma < X < \mu+\sigma) = 0{,}683$$

$$P(\mu-2\sigma \le X \le \mu+2\sigma) = P(\mu-2\sigma < X \le \mu+2\sigma) = P(\mu-2\sigma \le X < \mu+2\sigma) =$$
$$= P(\mu-2\sigma < X < \mu+2\sigma) = 0{,}954$$

$$P(\mu-3\sigma \le X \le \mu+3\sigma) = P(\mu-3\sigma < X \le \mu+3\sigma) = P(\mu-3\sigma \le X < \mu+3\sigma) =$$
$$= P(\mu-3\sigma < X < \mu+3\sigma) = 0{,}997$$

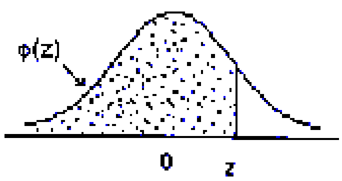
Apesar de ir além dos objectivos do programa de Matemática A do 12.º ano, penso que seria importante para o professor saber como se obtém os valores das probabilidades anteriores. Para isso, vou necessitar dos resultados que se seguem:

<u>Teorema:</u>

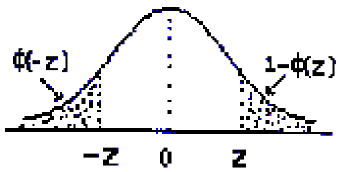
Se a variável aleatória X segue uma distribuição Normal com valor médio μ e desvio padrão σ então a variável aleatória $Z = \dfrac{X-\mu}{\sigma}$ segue uma distribuição Normal com valor médio 0 e desvio padrão 1.

<u>Nota:</u> À distribuição normal que tem valor médio 0 e desvio padrão 1 chamamos distribuição *standard* ou reduzida.

<u>Nota:</u> A função distribuição da normal reduzida tem uma notação especial. Assim, se Z for uma variável aleatória com distribuição normal reduzida, representamos $P(Z \le z) = \Phi(z)$.



<u>Propriedade:</u> Como a curva normal é simétrica, deduz-se facilmente que $\Phi(-z) = 1 - \Phi(z)$



Com base nos resultados anteriores e com base na tabela da distribuição normal reduzida facilmente se obtém os valores das probabilidades anteriores:

$$P(\mu - \sigma < X \leq \mu + \sigma) = P(-\sigma < X - \mu \leq \sigma) = P(-1 < \frac{X - \mu}{\sigma} \leq 1) = P(-1 < Z \leq 1) =$$

$$= \Phi(1) - \Phi(-1) = \Phi(1) - [1 - \Phi(1)] = \Phi(1) - 1 + \Phi(1) = 2\Phi(1) - 1 = 2 \times 0{,}8413 - 1 = 0{,}683$$

$$P(\mu - 2\sigma < X \leq \mu + 2\sigma) = P(-2 < Z \leq 2) = 2\Phi(2) - 1 = 2 \times 0{,}9772 - 1 = 0{,}954$$

$$P(\mu - 3\sigma < X \leq \mu + 3\sigma) = P(-3 < Z \leq 3) = 2\Phi(3) - 1 = 2 \times 0{,}9987 - 1 = 0{,}997$$

Posteriormente, os autores da brochura de Probabilidades e Combinatória ensinaram a analisar a distribuição normal através da calculadora gráfica. Na minha opinião, é bastante importante saber analisar as distribuições através do auxílio da calculadora gráfica ou de um *software* de Estatística. No entanto, é importante não esquecer que há alunos a frequentarem o 12.º ano sem calculadoras gráficas e que há muitas escolas que não têm calculadoras gráficas em número suficiente para facultar aos alunos nem estão equipadas com material informático e/ou *software* de Estatística.

No que diz respeito à função distribuição, os autores da brochura de Probabilidades e Combinatória apresentaram o seguinte:

> *"**Observação:** Dada uma variável aleatória X, define-se função distribuição ou função distribuição cumulativa de X, como sendo a função F(x), definida para todo o x real da seguinte forma:*
>
> *F(x) = P(X≤x)*

isto é, para cada x, a função distribuição dá-nos a probabilidade da variável aleatória assumir valores menores ou iguais a x. Quando pretendemos obter a probabilidade da variável aleatória pertencer ao intervalo (a, b), se tivermos a função distribuição F(x), então aquela probabilidade será

P(a<X≤b) = F(b) – F(a)” (p.115)

Uma vez que os autores da brochura de Probabilidades e Combinatória definiram a função distribuição dentro do capítulo do modelo Normal, será que os professores com pouca formação em Probabilidades ao consultarem a brochura não irão ficar com a ideia que a função distribuição só se aplica às variáveis aleatórias contínuas? Ou, apenas às variáveis que seguem uma distribuição Normal? De forma a evitar este tipo de pensamentos, os autores da brochura deveriam ter explicitado que a variável aleatória tanto pode ser discreta, como contínua. Assim, poderiam ter definido a função distribuição da seguinte forma:

Função distribuição de uma variável aleatória X (discreta ou contínua), é a função $F(x)$ tal que para todo o x real,

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

Além disso, os autores da brochura de apoio aos professores poderiam ter apresentado as propriedades da função distribuição.

*“**Propriedades da função distribuição de uma variável aleatória X (discreta ou contínua)***

1. $F(-\infty) = 0$ *(limite de* $F(x)$ *quando* $x \longrightarrow -\infty$ *) porque* $P(X \leq -\infty) = 0$

$F(+\infty) = 1$ *(limite de* $F(x)$ *quando* $x \longrightarrow +\infty$ *) porque* $P(X \leq +\infty) = 1$

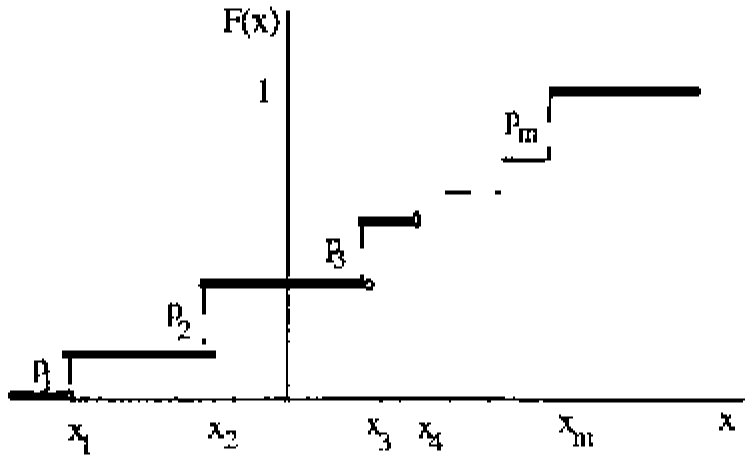
2. $F(x)$ *é uma função **não decrescente***

3. $F(x)$ *é **contínua à direita** (decorre da forma como foi definida)”*

(Martins, 2000, p. 163)

*“Se a v.a. **X** é **discreta**, a função distribuição é **descontínua** (só é contínua à direita) – é <u>uma função em escada</u>, com saltos nos pontos onde a v.a. assume valores com probabilidade diferente de zero.”* (Martins, 2000, p. 164)

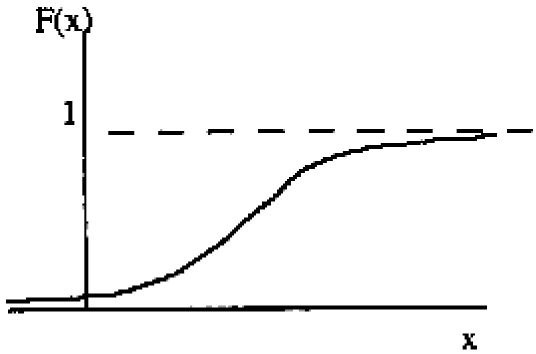
A função distribuição de uma variável aleatória discreta tem o seguinte aspecto:



(Martins, 2000, p. 163)

*"Se a v.a. **X** é **contínua**, a função distribuição é **contínua**, porque, qualquer que seja o ponto **a**, tem-se P(X=a)=0, pelo que a <u>função também é contínua à esquerda</u>."* (Martins, 2000, p. 164)

A função distribuição de uma variável aleatória contínua tem o seguinte aspecto:



(Martins, 2000, p. 164)

3.3. Análise Combinatória

Na terceira secção do programa de Matemática A encontram-se os seguintes pontos:

- Arranjos completos, arranjos simples, permutações e combinações.
- Triângulo de Pascal.
- Binómio de Newton.
- Aplicação ao cálculo de probabilidades.

(Silva *et al.*, 2002c, p. 3)

3.3.1. Arranjos Completos, Arranjos Simples, Permutações e Combinações

No que se refere aos arranjos completos e aos arranjos simples, os autores da brochura de apoio aos professores começaram por explicitar a diferença entre amostragem feita com reposição e amostragem feita sem reposição:

"Consideremos uma experiência que consiste em escolher ao acaso n elementos de S. Para facilidade de exposição podemos imaginar que os elementos são retirados um a um. Assim, de cada vez que um elemento é retirado da população, há dois modos possíveis de proceder: ou o elemento é reposto na população após se anotar o resultado (isto é qual o elemento retirado) ou não é. No primeiro caso diz-se que a amostragem é feita com reposição e no segundo caso diz-se que a amostragem é feita sem reposição.
Assim, a amostragem com reposição tem como consequência a possibilidade de conduzir a sequências em que os elementos se podem repetir e a amostragem sem reposição a sequências em que não há repetição de elementos." (p.128)

De forma a ilustrar melhor a diferença entre amostragem feita com reposição e amostragem feita sem reposição, os autores da brochura apresentaram o seguinte exemplo:

*"**Exemplo 3 -** Se tivermos uma urna com 6 bolas idênticas (no formato) numeradas de 1 a 6 e considerarmos uma experiência aleatória que consiste em retirar uma bola da urna e observar o número da bola saída, o espaço de resultados dessa experiência poderá escrever-se como: S = {1, 2, 3, 4, 5, 6}. Se repetirmos esta experiência 3 vezes, obtemos uma amostra de dimensão 3.*

Uma amostra ordenada de dimensão 3 possível é, por exemplo, (1, 2, 3). Isto significaria que a 1ª bola extraída teria o nº 1, a segunda o nº 2 e a terceira o nº 3. A ordenação corresponde pois à ordem de extracção. Note-se que esta amostra ordenada será distinta da amostra (3, 1, 2), embora contenham ambas os mesmos elementos. A ordem por que os elementos aparecem é pois importante.

(…)

Neste exemplo quantas serão as amostras possíveis de dimensão 3 (com reposição e sem reposição)?

Uma maneira fácil de "contar" é pensar do seguinte modo:

Com reposição

- *O 1º número saído pode ser qualquer. Há assim 6 hipóteses para o 1º número.*
- *Para cada número que sai na 1ª extracção há 6 números possíveis para o acompanhar na 2ª extracção. Temos assim um total de $6\times 6=36$ possibilidades após a 2ª extracção.*
- *Para cada um dos 36 pares possíveis que resultam das duas primeiras extracções há 6 números possíveis para a terceira extracção. Há assim um total de $36\times 6=216$ triplos possíveis.*

Generalizando para N – dimensão do espaço de resultados e para n – dimensão da amostra, verificamos então que o número possível de amostras de dimensão n, com reposição, que se pode extrair é $\underbrace{N\times N\times \ldots\times N}_{n \text{ vezes}} = N^n$.

(…) Com efeito, quando a amostragem é feita com reposição, uma amostra ordenada de dimensão n, não é mais do que um elemento do produto cartesiano $\underbrace{S\times S\times\ldots\times S}_{n \text{ vezes}}$

Sem reposição:

- *O 1º número saído pode ser qualquer. Há assim 6 hipóteses para o 1º número.*
- *Para cada número que sai na 1ª extracção há apenas 5 números possíveis para o acompanhar na 2ª extracção. Temos assim um total de 6x5=30 possibilidades após a 2ª extracção.*
- *Para cada um dos 30 pares possíveis que resultam das duas primeiras extracções há já só 4 números possíveis para a terceira extracção. Há assim um total de 30x4=120 triplos possíveis para as três extracções."* (p. 127, 128 e 129)

A partir daqui os autores concluíram o seguinte resultado:

"Resultado 1:

Para uma população de N elementos e um determinado valor n o número de amostras distintas de dimensão n que se pode obter numa extracção com reposição é igual a N^n *e sem reposição (com n ≤ N) é igual a* ${}^{N}A_n = N\times(N-1)\times\ldots\times(N-n+1)$*"* (p.129)

Será que um professor com pouca formação em Análise Combinatória sabe o significado de ${}^{N}A_n$? Além disso, será que sabe que este resultado só é válido para amostras ordenadas? Na minha opinião, o resultado anterior poderia ter sido apresentado da seguinte forma:

Para uma população de N elementos, o número de amostras <u>ordenadas</u> distintas de dimensão n que se pode obter numa extracção <u>com reposição</u> é igual a N^n e <u>sem reposição</u> (com $n \le N$) é igual a $N\times(N-1)\times\ldots\times(N-n+1)$.

Posteriormente, os autores da brochura de Probabilidades e Combinatória apresentaram as seguintes definições:

*"**Arranjos completos (Arranjos com repetição):***

Ao número de modos distintos de extrair ordenadamente e com reposição, n elementos de um conjunto com N elementos, dá-se o nome de arranjos completos de N, n a n e representa-se por ${}^{N}A'_n$. *Esse número é igual a* N^n*."* (p. 129)

*"**Arranjos simples***

Ao número de modos distintos de extrair ordenadamente e sem reposição, n elementos de um conjunto com N elementos, dá-se o nome de arranjos simples N, n a n e representa-se por ${}^{N}A_n$. *Esse número é igual a* $N\times(N-1)\times\ldots\times(N-n+1)$*."*

(p. 130)

Na minha opinião, estas definições estão bastante perceptíveis, no entanto, os autores da brochura poderiam ter referido que os arranjos completos e os arranjos simples são também designados por arranjos com reposição e arranjos sem reposição, respectivamente.

No que se refere às permutações, os autores da brochura de Probabilidades e Combinatória apresentaram o seguinte resultado:

"Resultado 2 (Permutações):

O número de amostras ordenadas distintas, de dimensão n que se pode obter, sem reposição, de um subconjunto de dimensão n de S é ${}^{n}A_{n} = n\times(n-1)\times\ldots\times 1$.

Este número costuma representar-se por $n!$ (lê-se factorial de n)" (p. 131)

Uma vez que as permutações de n elementos são geralmente representadas por P_n e não têm necessariamente que se obter de um subconjunto de S, penso que seria mais correcto se o resultado anterior fosse apresentado da seguinte forma:

> Chama-se permutações de n elementos a todas as amostras ordenadas distintas, de dimensão n, que se podem obter, sem reposição, de um conjunto de dimensão n. O número dessas sequências representa-se por P_n (lê-se "permutações de n") e é dado por:
>
> $$P_n = n\times(n-1)\times\ldots\times 2\times 1 = n!$$
>
> O símbolo $n!$ lê-se "n factorial" ou "factorial de n" e representa o produto dos n primeiros inteiros positivos.

No que diz respeito às combinações, os autores da brochura de Probabilidades e Combinatória começaram por apresentar o seguinte:

*"Em muitas situações pode não interessar a ordem por que aparecem os elementos na amostra. Por exemplo, a amostragem sem reposição pode ser feita retirando os n elementos todos de uma vez e consequentemente não podemos falar numa ordem, no sentido de podermos dizer qual o primeiro elemento retirado, qual o segundo, etc. Assim pode falar-se em **amostra não ordenada**.*

(…) Os resultados de análise combinatória que nos interessam nestas circunstâncias são então os respeitantes à enumeração de subconjuntos de um conjunto. Assim, por exemplo, as amostras ordenadas (3,1,2) ou (1,2,3) resultam na mesma amostra não ordenada que representaremos por {1,2,3}, ou seja o conjunto formado por aqueles três elementos." (p. 132)

Perante isto, os autores da brochura apresentaram o seguinte resultado:

"Resultado 3 (Combinações):

O número de subconjuntos de dimensão n que se podem formar de um conjunto S de dimensão N é dado por $\binom{N}{n} = \dfrac{{}^{N}A_{n}}{n!}$." (p. 134)

"Dito de outro modo, o número de amostras não ordenadas de dimensão n que se podem retirar, sem reposição, de uma população S de dimensão N é dado por $\binom{N}{n} = \frac{{}^{N}A_{n}}{n!}$. A este número dá-se o nome de ***combinações de N, n a n.****"*

(p. 134)

Uma vez que as combinações de N, n a n são muitas vezes representadas por ${}^{N}C_{n}$, penso que os autores da brochura de apoio aos professores deveriam ter feito referência a este tipo de notação.

3.3.2. Lei de Pascal e Lei da Simetria

Se um professor com pouca formação em Análise Combinatória recorrer à brochura de apoio ao professor para consultar a Lei de Pascal, irá encontrar o seguinte:

"É fácil de perceber que se tem a seguinte igualdade:

$$\binom{N}{n-1} + \binom{N}{n} = \binom{N+1}{n} \quad \text{com } n = 1,\ldots,N$$

Esta é a chamada Lei de Pascal, a qual é fácil de estabelecer.

Com efeito, para obtermos o número de subconjuntos com n elementos que podemos obter de um conjunto de N+1 elementos, podemos raciocinar do seguinte modo:

- *Seja x o elemento do conjunto com N+1 elementos que não pertence ao conjunto de N elementos.*
- *Podemos subdividir os subconjuntos de n elementos do conjunto com N+1 elementos em duas categorias: subconjuntos que contêm x e subconjuntos que não contêm x.*
- *Do conjunto com N elementos, (que não contém x), podemos formar $\binom{N}{n}$ subconjuntos de n elementos, dos quais nenhum contém obviamente o elemento retirado x.*
- *Para obter os subconjuntos de n elementos que contêm x, basta juntá-lo a todos os subconjuntos com n-1 elementos daquele conjunto com N elementos. Esses são em número de $\binom{N}{n-1}$.*

Assim se chega ao resultado pretendido." (p. 135)

Se um professor apresentar esta demonstração aos seus alunos, será que eles irão conseguir compreender? Na minha opinião, seria mais acessível para os alunos do Ensino Secundário, se a Lei de Pascal fosse demonstrada da seguinte forma:

Demonstração:

$$\binom{N}{n-1}+\binom{N}{n}=\frac{N!}{(n-1)![N-(n-1)]!}+\frac{N!}{n!(N-n)!}$$

$$=\frac{N!}{(n-1)!(N-n+1)(N-n)!}+\frac{N!}{(n-1)!n(N-n)!}$$

$$=\frac{nN!+(N-n+1)N!}{n!(N-n+1)!}$$

$$=\frac{nN!+(N+1)N!-nN!}{n!(N+1-n)!}$$

$$=\frac{(N+1)!}{n!(N+1-n)!}=\binom{N+1}{n}$$

Logo, $\binom{N}{n-1}+\binom{N}{n}=\binom{N+1}{n}$ $\square$

Outra lei apresentada pelos autores da brochura de apoio aos professores é a Lei da Simetria que é traduzida pela seguinte igualdade:

$$\binom{N}{n}=\binom{N}{N-n}$$

Apesar dos autores da brochura de apoio aos professores considerarem que a verificação desta igualdade é trivial, penso que poderia ter sido apresentada:

Demonstração:

Temos $\binom{N}{n}=\dfrac{N!}{n!(N-n)!}$

e temos

$$\binom{N}{N-n} = \frac{N!}{(N-n)![N-(N-n)]!} =$$

$$= \frac{N!}{(N-n)!n!} =$$

$$= \frac{N!}{n!(N-n)!}$$

Logo, $\binom{N}{n} = \binom{N}{N-n}$ □

3.3.3. Triângulo de Pascal

No que diz respeito ao triângulo de Pascal, os autores da brochura de Probabilidades e Combinatória apresentaram o seguinte:

> *"Suponhamos agora que vamos calcular todas as combinações de N elementos tomados n a n, fazendo variar N desde 0 e n desde 0 a N. Ao procedermos deste modo para vários valores de N, podemos fazer a sua representação numa forma tabular. Obtém-se aquilo que se costuma designar por triângulo de Pascal*

$$\binom{0}{0}=1$$

$$\binom{1}{0}=1 \quad \binom{1}{1}=1$$

$$\binom{2}{0}=1 \quad \binom{2}{1}=2 \quad \binom{2}{2}=1$$

$$\binom{3}{0}=1 \quad \binom{3}{1}=3 \quad \binom{3}{2}=3 \quad \binom{3}{3}=1$$

$$\binom{4}{0}=1 \quad \binom{4}{1}=4 \quad \binom{4}{2}=6 \quad \binom{4}{3}=4 \quad \binom{4}{4}=1$$

$$\binom{5}{0}=1 \quad \binom{5}{1}=5 \quad \binom{5}{2}=10 \quad \binom{5}{3}=10 \quad \binom{5}{4}=5 \quad \binom{5}{5}=1$$

....

> *Repare-se que cada termo do triângulo de Pascal se obtém como a soma dos dois termos que lhe estão acima."* (p. 135 e 136)

Para além da propriedade que os autores da brochura referiram, também poderiam ter mencionado o facto das linhas do triângulo de Pascal começarem e acabarem sempre em 1, ou seja,

$$\binom{N}{0} = \binom{N}{N} = 1 \qquad , \quad para \quad todo \quad o \quad N \in IN$$

Além disso, poderiam ainda ter apresentado a numeração que é atribuída a cada linha do triângulo de Pascal, ou seja,

$$\binom{0}{0}=1 \longrightarrow \text{Linha 0}$$

$$\binom{1}{0}=1 \quad \binom{1}{1}=1 \longrightarrow \text{Linha 1}$$

$$\binom{2}{0}=1 \quad \binom{2}{1}=2 \quad \binom{2}{2}=1 \longrightarrow \text{Linha 2}$$

$$\binom{3}{0}=1 \quad \binom{3}{1}=3 \quad \binom{3}{2}=3 \quad \binom{3}{3}=1 \longrightarrow \text{Linha 3}$$

$$\binom{4}{0}=1 \quad \binom{4}{1}=4 \quad \binom{4}{2}=6 \quad \binom{4}{3}=4 \quad \binom{4}{4}=1 \longrightarrow \text{Linha 4}$$

$$\binom{5}{0}=1 \quad \binom{5}{1}=5 \quad \binom{5}{2}=10 \quad \binom{5}{3}=10 \quad \binom{5}{4}=5 \quad \binom{5}{5}=1 \longrightarrow \text{Linha 5}$$

.... ……..

3.3.4. Binómio de Newton

Quanto ao Binómio de Newton, os autores da brochura de Probabilidades e Combinatória apresentaram o seguinte:

*"**Teorema Binomial**: Para quaisquer a e b reais e qualquer inteiro positivo N é válida a seguinte igualdade*

$$\begin{aligned}(a+b)^N &= \sum_{k=0}^{N}\binom{N}{k}a^{N-k}b^k \\ &= \binom{N}{0}a^N + \binom{N}{1}a^{N-1}b + ... + \binom{N}{k}a^{N-k}b^k + \\ &+ ... + \binom{N}{N-1}ab^{N-1} + \binom{N}{N}b^N\end{aligned}$$

Demonstração (por indução):

A igualdade é válida para N = 1, como é fácil de verificar .

Suponhamos que é válida para N. Vamos mostrar que continua válida para N+1.

Tem-se:

$$(a+b)^{N+1} = (a+b)^N(a+b) = \left(\sum_{k=0}^{N}\binom{N}{k}a^{N-k}b^k\right)(a+b) =$$

$$= \binom{N}{0}a^{N+1} + \binom{N}{0}a^N b + \binom{N}{1}a^N b + \binom{N}{1}a^{N-1}b^2 + \binom{N}{2}a^{N-1}b^2 + \binom{N}{2}a^{N-2}b^3$$

$$+ \ldots + \binom{N}{N-1}a^2 b^{N-1} + \binom{N}{N-1}ab^N + \binom{N}{N}ab^N + \binom{N}{N}b^{N+1}$$

O resultado pretendido segue imediatamente, usando o facto de que $\binom{N}{0} = 1$, $\binom{N}{N} = 1$ *para qualquer N e atendendo à igualdade* $\binom{N}{n-1} + \binom{N}{n} = \binom{N+1}{n}$ *. "*(p. 137)

Uma vez que a brochura de Probabilidades e Combinatória tem como objectivo apoiar os professores, penso que os autores da brochura poderiam ter apresentado algumas conclusões em relação ao binómio de Newton. Passo, então, a citar algumas conclusões que poderiam ter sido apresentadas pelos autores da brochura.

- O binómio de Newton tem $n+1$ parcelas;
- Os coeficientes de cada um dos monómios, assim ordenados, são os números da linha N do triângulo de Pascal;
- A soma dos expoentes de a e de b em cada monómio é N ;
- Os expoentes de a decrescem de N até zero e os de b crescem de zero até N .

3.3.5. Aplicação ao Cálculo de Probabilidades

Em relação ao último ponto do programa de Matemática A do 12.º ano – *Aplicação ao cálculo de probabilidades* – os autores da brochura de Probabilidades e Combinatória apresentaram diversas propostas de actividades, que na minha opinião são bastante simples e podem ser resolvidas na sala de aula com os alunos.

Para além das propostas de actividades, os autores da brochura apresentaram vários exemplos clássicos de cálculo de Probabilidades e alguns exercícios. No entanto, estes estão, de um modo geral, mais dirigidos para os professores.

Para finalizar o capítulo, os autores da brochura de Probabilidades e Combinatória apresentaram um problema e diversos métodos de resolução desse mesmo problema. Com isto, os autores pretendiam reforçar a ideia de que não existe um processo único para chegar à solução. É importante que os professores criem momentos de discussão na sala de aula para que os alunos possam comparar e discutir as várias formas de resolução um dado problema, pois só assim poderão apreciar as diferentes abordagens que se podem fazer para uma determinada situação.

CAPÍTULO IV

ANÁLISE DA TEMÁTICA DE MODELOS DE PROBABILIDADE NO PROGRAMA DE MATEMÁTICA B DO ENSINO SECUNDÁRIO

No programa de Matemática B encontram-se as seguintes indicações sobre a temática de *Modelos de Probabilidade*:

> "*Todos os dias somos confrontados com situações que nos conduzem a utilizar, intuitivamente, a noção de probabilidade. Nos mais variados aspectos da nossa vida está presente a incerteza: no jogo (totoloto, totobola,...), na probabilidade de chover ou fazer sol, na probabilidade de sucesso de um certo produto que se pretende lançar no mercado,... As noções de probabilidade devem fazer parte da formação de um cidadão para que ele se possa integrar nas sociedades actuais onde a incerteza e a aleatoriedade são noções usadas frequentemente. Os conteúdos a abordar neste tema devem ser apresentados aos estudantes depois de estes terem sido confrontados com actividades próprias do dia-a-dia de qualquer pessoa, associadas a este tema, como por exemplo: jogos, inquéritos, extracções com reposição e sem reposição de bolas de uma urna, lançamento de moedas, de dados,...*
> *Pretende-se que os estudantes sejam capazes de:*
>
> - *reconhecer as vantagens em encontrar modelos matemáticos apropriados para estudar fenómenos aleatórios;*
> - *compreender as aproximações conceptuais para a probabilidade:*
>   - *aproximação frequencista de probabilidade;*
>   - *definição clássica ou probabilidade de Laplace ;*
> - *construir modelos de probabilidade em situações simples e usá-los para calcular a probabilidade de alguns acontecimentos;*
> - *apreender as propriedades básicas das distribuições de probabilidade;*
> - *resolver problemas simples, recorrendo à calculadora gráfica ou computador, envolvendo distribuições de probabilidade, em particular envolvendo a distribuição normal.*

> *A base da aprendizagem deve estar na experimentação — recorrendo a materiais manipuláveis ou simulações — e na resolução de problemas. Ao modelarem situações, os estudantes são conduzidos a construírem o espaço de resultados de uma experiência aleatória e a definirem acontecimentos. Os estudantes poderão usar simulações para construir empiricamente distribuições de probabilidades e utilizar a noção frequencista de probabilidade comparando resultados de simulações para predizer valores da probabilidade de um acontecimento. Sugere--se a consulta da "Actividade – Exemplo de como organizar uma experiência na sala de aula" (Brochura de Probabilidades, pág 36).*
> *A definição de Laplace de probabilidade deve ser apresentada depois de serem criadas condições para se sentir a sua necessidade.*
> *Através da apresentação de uma tarefa (como por exemplo, "Jogo dos dois dados" da Brochura de probabilidades, pág. 44) e depois de ter sido abordada experimentalmente a noção de probabilidade de um acontecimento, os estudantes podem sentir dificuldades naturais. É esse o melhor modo de perceber o interesse de alguns resultados teóricos.*
> *Não se justifica, nesta disciplina, o estudo de modelos para situações que obriguem a utilizar técnicas de contagem que envolvam cálculo combinatório.*
> *É importante que os estudantes sejam capazes de estimar probabilidades de acontecimentos através da análise de um histograma. Recorrendo à calculadora ou ao computador, podem determinar a média e o desvio-padrão de uma distribuição.*"
>
> (Silva *et al.*, 2002a, p.1 e 2)

Ao ler as indicações referentes à temática de *Modelos de Probabilidade*, verifiquei que são sugeridas diversas metodologias e actividades que podem ser consideradas adequadas e suficientes, transmitindo ao professor uma perspectiva da natureza do trabalho que pode e deve ser desenvolvido na sala de aula. Estas sugestões são apresentadas da seguinte forma:

> *"Os conteúdos a abordar neste tema devem ser apresentados aos estudantes depois de estes terem sido confrontados com actividades próprias do dia-a-dia de qualquer pessoa, associadas a este tema, como por exemplo: jogos, inquéritos, extracções com reposição e sem reposição de bolas de uma urna, lançamento de moedas, de dados,..."* (p.1)

> *"Sugere-se a consulta da "Actividade – Exemplo de como organizar uma experiência na sala de aula" (Brochura de Probabilidades, pág 36)."* (p.2)

> *"Através da apresentação de uma tarefa (como por exemplo, "Jogo dos dois dados" da Brochura de probabilidades, pág 44)"* (p.2)

Porém, as indicações sobre os conteúdos programáticos a desenvolver são insuficientes. Estas indicações são apresentadas aos professores a partir das sugestões metodológicas. Mas, será que um professor com pouca formação em Probabilidades, depois de consultar o programa de Matemática B, consegue facilmente saber quais os conteúdos programáticos que devem ser desenvolvidos na sala de aula?

Na minha opinião, os autores do programa de Matemática B, para além de sugerirem actividades e metodologias, deveriam ter especificado os conteúdos programáticos, tal como o fizeram ao longo do programa de Matemática A. Desta forma, evitariam que alguns professores sentissem dificuldades em identificar os conteúdos que devem abordar na sala de aula.

O programa de Matemática B apela fortemente à utilização da tecnologia e de materiais manipuláveis. Lê-se no programa o seguinte:

> *"A base da aprendizagem deve estar na experimentação — recorrendo a materiais manipuláveis ou simulações (…)"* (p.2)

> "*Recorrendo à calculadora ou ao computador, podem determinar a média e o desvio-padrão de uma distribuição."* (p.2)

No entanto, em poucas escolas existe este tipo de recursos que sejam amplamente utilizados. É do conhecimento comum que a grande maioria das aulas de Matemática B não decorre em salas com computadores. Muitos alunos não adquirem calculadoras gráficas e as escolas não as têm em número suficiente para suprir esta falta e permitir o trabalho regular na aula. De uma forma geral, não há recursos físicos nem materiais adequados nas escolas, como por exemplo, uma sala de trabalho com computadores e outros materiais (manipuláveis ou tecnológicos) à disposição dos alunos. Neste sentido, falharam as escolas através dos Conselhos Executivos, mas sobretudo os organismos centrais do Ministério de Educação a quem competia prever as necessidades.

CAPÍTULO V

5.1. Metodologia

O presente estudo tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Probabilidades e Estatística assim como à forma como estes foram abordados nas aulas.

O estudo teve a participação de 1144 alunos (361 alunos do 7.º ano; 319 alunos do 8.º ano; 258 alunos do 9.º ano; 129 alunos do 10.º ano do Curso Científico-Humanístico; 77 alunos do 12.º ano do Curso Científico-Humanístico). Estes alunos pertencem a escolas da região de Lisboa e Vale do Tejo.

Os dados foram recolhidos no final do 3.º Período do ano lectivo 2007/2008 através da aplicação de inquéritos efectuados presencialmente na sala de aula, abrangendo todos os alunos presentes na aula.

Os inquéritos foram realizados em aulas de Matemática, não tendo os alunos gasto mais do que 10 minutos, o que não prejudicou a planificação prevista para aquelas aulas.

Os inquéritos foram compostos por questões maioritariamente de resposta fechada que pretendiam medir variáveis dicotómicas (por exemplo, *sexo*), variáveis categorizadas de nominais (por exemplo, *tema de Matemática que mais gosta*), variáveis categorizadas de ordinais (por exemplo, *rendimento a Matemática*) e variáveis categorizadas de razão (*idade*), no entanto, as últimas questões foram de resposta aberta.

Para se medir as variáveis ordinais, como por exemplo *a opinião que os alunos têm em relação à disciplina de Matemática* ou *a forma como os alunos avaliam os conteúdos de Probabilidades e Estatística*, utilizou-se uma escala de *Likert* de cinco categorias. Do ponto de vista estatístico, este tipo de escala é normalmente utilizada para medir aspectos subjectivos dos inquiridos, tais como, opiniões pessoais, experiências pessoais e grau de satisfação. Segundo Chisnall (2005), as escalas de

Likert são muito populares porque, para além de serem simples de construir, permitem recolher uma boa informação sobre o grau de sentimentos dos inquiridos.

Os inquéritos aplicados aos alunos do 7.º e 8.º anos foram compostos por 11 questões, os que foram aplicados aos alunos do 9.º ano foram constituídos por 10 questões e os que foram apresentados aos alunos do 10.º e 12.º anos foram compostos por 14 questões (ver ANEXO III).

Uma das questões que esteve presente em todos os inquéritos foi: *"Qual a sua opinião em relação à disciplina de Matemática?"* Com esta questão pretendia-se verificar a validade da ideia do senso comum de que os alunos não gostam de Matemática.

Outra das questões que também esteve presente em todos os inquéritos foi: *"Como avalia o seu rendimento?"* Esta questão foi colocada com o objectivo de saber se a opinião que os alunos têm em relação à Matemática está associada ao rendimento que têm nesta disciplina.

Também se procurou saber *como é que os alunos avaliam os conteúdos de Estatística* e *como avaliam a forma como estes conteúdos foram abordados nas aulas.* Estas questões foram colocadas porque muitos professores supõem que, para os alunos, os conteúdos de Estatística são muito fáceis e, como tal, não lhes dão a devida importância quando os estão a leccionar.

Nos inquéritos aplicados aos alunos do Ensino Secundário encontra-se a seguinte questão: *"No 3.º Ciclo do Ensino Básico, teve aulas de Probabilidades e Estatística?"* Esta questão foi colocada aos alunos com o intuito de confirmar a ideia de que os conteúdos de Estatística e de Probabilidades nem sempre são abordados no 3.º Ciclo.

Também se procurou saber, através dos inquéritos feitos aos alunos do Ensino Secundário, com que finalidade estes alunos utilizam as calculadoras gráficas. Pois, suspeita-se que nem todos os alunos procuram tirar o maior proveito desta tecnologia.

Para se poder aplicar os inquéritos em meio escolar, teve de ser solicitada a autorização do Director Geral de Inovação e Desenvolvimento Curricular (ver ANEXOI) bem como a dos Presidentes dos Conselhos Executivos das escolas (ver ANEXO II).

Uma vez que a resposta do Director Geral de Inovação e Desenvolvimento face ao pedido de autorização não foi imediata e como os inquéritos só podiam ser aplicados no final do ano lectivo não foi possível inquirir os alunos do 10.º e 11.º ano do Curso Científico-Humanístico de Artes Visuais. Visto que este é o único curso que tem a disciplina de Matemática B, não houve, desta forma, oportunidade para poder analisar a receptividade dos alunos face aos programas de Matemática B.

Para além de todos os pedidos de autorização, a aplicação dos inquéritos também dependeu da disponibilidade dos professores.

Em termos de tratamento de dados, inicialmente, analisou-se cada um dos inquéritos e transcreveu-se a informação contida nos mesmos para um ficheiro de *Excel*. Posteriormente, procedeu-se ao tratamento estatístico das variáveis através do *software SPSS (Statistical Package for the Social Sciences)* versão 17.0 e do *software R* (The R Project for Statistical Computing) versão 2.4.0.

Durante o tratamento dos dados foram construídas várias tabelas de contingência. Algumas destas apresentavam em determinadas categorias células nulas, que provavelmente terão resultado do facto da amostra ser demasiado pequena em relação ao número de células da tabela. Uma vez que não era possível aumentar a dimensão da amostra e para evitar adicionar meia unidade ao valor da frequência de cada célula da tabela (Bishop *et al.* (1975)), agruparam-se as categorias extremas adjacentes das variáveis que pretendiam medir a opinião dos alunos (Daniel, 2005). Desta forma, as tabelas de contingência deixaram de ter células nulas e os valores esperados das células também aumentaram. Este último facto é também muito importante, pois se os valores esperados das células forem pequenos, os resultados do teste do qui-quadrado serão muito pouco rigorosos, dado que se recorre a uma distribuição assimptótica (Everitt, 1992). Segundo Cochran (1954), a aproximação pelo qui-quadrado é válida em tabelas superiores a 2x2, desde que todos os valores esperados sejam maiores do que 1 e que apenas 20% das células tenham valor esperado inferior a 5. Também Lewontin e Felsenstein (1965) defendem que a aproximação pelo qui-quadrado é válida em tabelas $2 \times n$ se todos os valores esperados forem iguais ou superiores a 1.

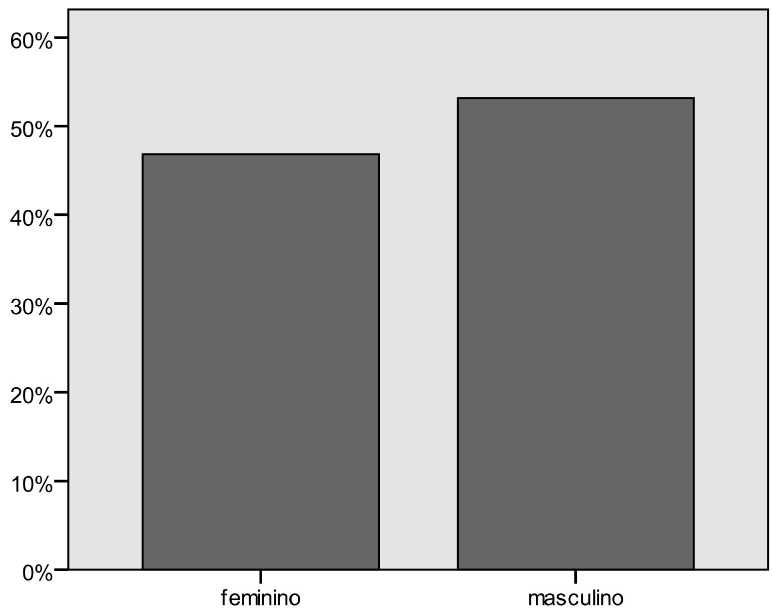
5.2. Análise de Dados

Os resultados obtidos no presente estudo são apresentados separadamente tendo em conta os diferentes anos de escolaridade.

5.2.1. 7.º Ano do Ensino Básico

Da amostra recolhida, constatei que 53,2% dos estudantes são do género masculino, enquanto que 46,8% são do género feminino (Gráfico 17).

Gráfico 17: Distribuição por género



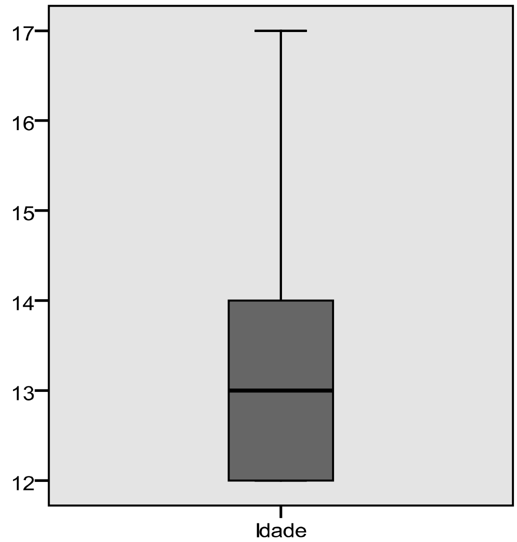
Pela análise da Tabela 1 e do Gráfico 18, concluí que das 361 idades disponíveis, a menor é de 12 anos e a maior é de 17 anos, ou seja, há uma grande variação das idades dos alunos que frequentam o 7.º ano de escolaridade; a moda das idades é de 13 anos; a média de idades dos inquiridos é de 13,2 anos e a dispersão das observações em relação à média – desvio padrão – é de 1,16 anos; 75% dos inquiridos têm idade entre os 12 e os 14 anos, como se pode depreender dos valores dos quartis e do mínimo da amostra; os restantes 25% dos inquiridos, que correspondem aos alunos mais velhos, têm idade entre os 14 e os 17 anos, como se depreende do valor do 3.º quartil e do máximo da amostra.

Analisando os valores anteriores, concluí que, de facto, faz sentido que a moda das idades seja de 13 anos, pois se um aluno fizer um percurso normal na escola quando chegar ao 7.º ano tem 12 ou 13 anos, consoante a data do seu aniversário. Portanto também faz sentido que a idade mínima observada seja de 12 anos. No entanto, é do conhecimento geral que nem todos os alunos apresentam um percurso normal na escola, daí haver muitos inquiridos com idade superior à moda (este facto é facilmente confirmado através do Gráfico 18).

Tabela 1: Estatísticas da variável idade

N	361
Média	13,2
Mediana	13
Moda	13
Variância	1,34
Desvio Padrão	1,16
Mínimo	12
Máximo	17
1.º Quartil	12
2.º Quartil	13
3.º Quartil	14

Gráfico 18: *Box-plot* das idades dos inquiridos



No que diz respeito à opinião que os alunos têm em relação à disciplina de Matemática, a maioria dos inquiridos (38,2%) afirma gostar de Matemática (Gráfico 19). No entanto, a maioria dos inquiridos (47,9%) auto-avalia o seu rendimento a esta disciplina de Suficiente, seguido de Insuficiente com 23,8% (Gráfico 20). De facto, ao consultar os Gráficos 21 e 22, concluí que a classificação que mais se verificou quer no final do 1.º Período quer no final do 2.º Período foi a classificação 3, seguida da classificação 2. Deste modo, constatei que apesar da maioria dos inquiridos gostar da disciplina de Matemática revelam poucos conhecimentos matemáticos.

Gráfico 19: Distribuição dos inquiridos segundo a sua opinião em relação à disciplina de Matemática

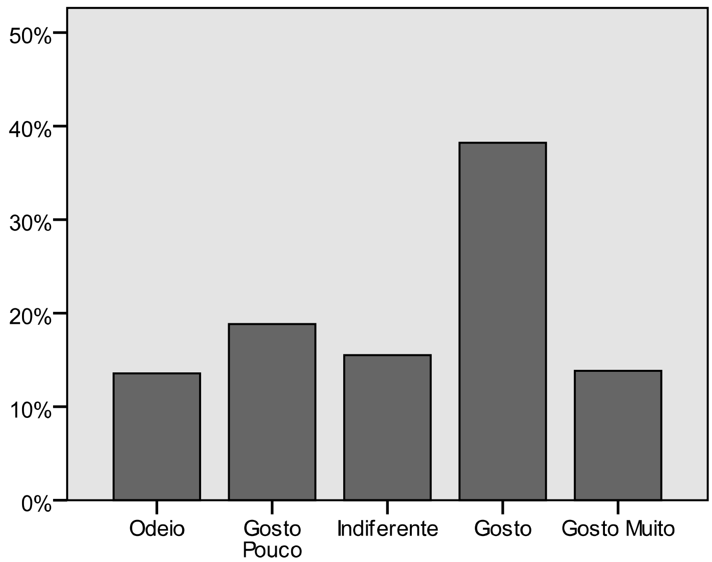


Gráfico 20: Distribuição dos inquiridos segundo a sua própria avaliação do seu rendimento a Matemática

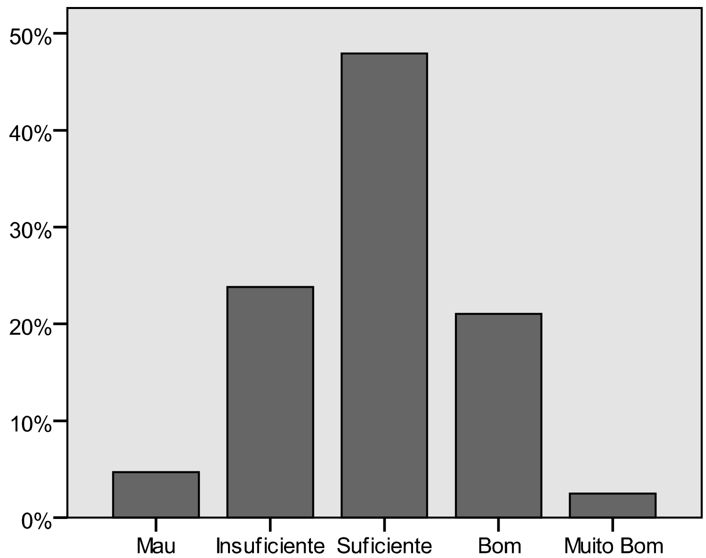


Gráfico 21: Classificação que os alunos obtiveram no 1.º Período

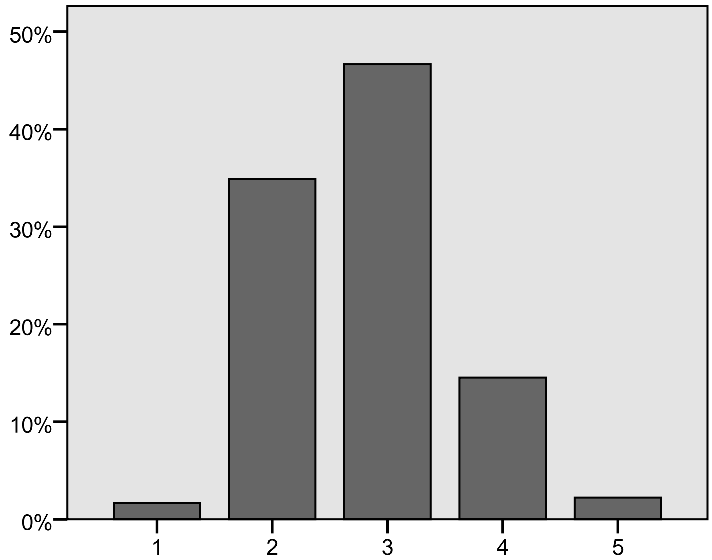
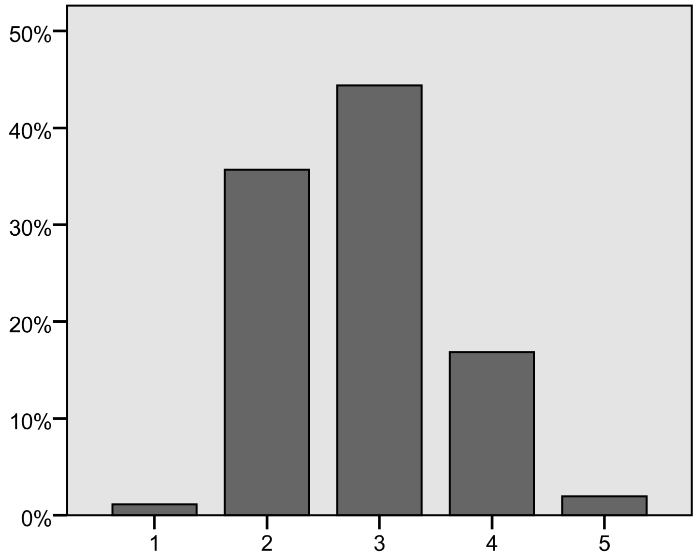


Gráfico 22: Classificação que os alunos obtiveram no 2.º Período



Para se conhecer o grau de associação (linear) entre duas variáveis, calcula-se o coeficiente de correlação.

Zar (1999) atribui a Galton as primeiras ideias sobre correlação, afirmando que começou por considerar a correlação entre as ordens ainda antes de desenvolver a correlação entre as próprias observações. No entanto, a formalização dos conceitos, tal como os usamos hoje, devem-se a Pearson e a Spearman.

Spearman (1904) desenvolveu um **coeficiente de correlação ordinal**, denotado por r_S, usando os *ranks* nas duas ordenações:

$$r_S = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n^3 - n}$$

onde d_i é a diferença entre os *ranks* de X e de Y, ou seja, $d_i = rankX_i - rankY_i$ [1].

No caso de haver dados repetidos são-lhes atribuídos os *ranks* médios e depois, geralmente, calcula-se o coeficiente de correlação através da fórmula anterior.

---

[1] Em vez de se utilizar a diferença entre *ranks* dos pares de X e Y, pode-se utilizar a soma dos *ranks*, ou seja, $S_i = rankX_i + rankY_i$ (Meddis, 1984; Thomas, 1989), e neste caso,

$$r_S = \frac{6\sum_{i=1}^{n} S_i^2}{n^3 - n} - \frac{7n + 5}{n - 1}$$

No entanto, alguns autores (por exemplo, Kendall (1945), Kendall (1970) e Thomas (1989)) consideram importante aplicar uma correcção no caso de haver observações repetidas e, neste caso, o coeficiente de correlação de Spearman é calculado da seguinte forma:

$$(r_S)_C = \frac{(n^3-n)/6 - \sum_{i=1}^{n} d_i^2 - \sum t_X - \sum t_Y}{\sqrt{\left[(n^3-n)/6 - 2\sum t_X\right]\left[(n^3-n)/6 - 2\sum t_Y\right]}}$$

onde $\sum t_X = \frac{\sum \left(t_i^3 - t_i\right)}{12}$, em que t_i é o número de observações de X iguais, em cada grupo de empates.

E, $\sum t_Y = \frac{\sum \left(t_i^3 - t_i\right)}{12}$, em que t_i é o número de observações de Y iguais, em cada grupo de empates.

Se $\sum t_X$ e $\sum t_Y$ forem zero, então o calculo de r_S é idêntico ao cálculo de $(r_S)_C$.

Face ao que acabei de referir, passo a analisar o grau de associação que há entre a opinião que os alunos têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina.

Por um lado, tenho 49 alunos que odeiam Matemática, 68 que gostam pouco, 56 consideram esta disciplina indiferente, 138 alunos gostam de Matemática e 50 gostam muito. Assim,

$$\sum t_X = \frac{(49^3-49)+(68^3-68)+(56^3-56)+(138^3-138)+(50^3-50)}{12} = 280034$$

Por outro lado, tenho 17 alunos que avaliam o seu rendimento a Matemática de Mau, 86 avaliam-no de Insuficiente, 173 avaliam-no de Suficiente, 76 consideram-no Bom e 9 avaliam-no de Muito Bom. Assim,

$$\sum t_Y = \frac{(17^3-17)+(86^3-86)+(173^3-173)+(76^3-76)+(9^3-9)}{12} = 521502{,}5$$

Além disso, $\sum_{i=1}^{361} d_i^{\ 2} = 3096795$

Portanto,

$$(r_S)_C = \frac{(361^3 - 361)/6 - 3096795 - 280034 - 521502{,}5}{\sqrt{[(361^3 - 361)/6 - 2 \times 280034] \cdot [(361^3 - 361)/6 - 2 \times 521502{,}5]}} = 0{,}5604$$

Depois de obter o valor do coeficiente de correlação é necessário avaliar a sua significância, ou seja, é necessário saber se, de facto, há correlação entre as variáveis na população. Como tal, pretendo testar:

$$H_0 : \rho_S = 0 \qquad vs. \qquad H_1 : \rho_S \neq 0$$

Se H_0 for verdadeira, a estatística de teste é dada por:

$$t = \frac{r_S \sqrt{n-2}}{\sqrt{1 - r_S^{\ 2}}} \sim t_{n-2}$$

em que a região de rejeição para a hipótese H_0, ao nível de significância α, é dada por:

$$|t| \geq t_{n-2;1-\frac{\alpha}{2}}$$

onde $t_{n-2;1-\frac{\alpha}{2}}$ é o quantil de probabilidade $1-\frac{\alpha}{2}$ da distribuição *t* de Student com $n-2$ graus de liberdade.

Com o auxílio do *software* R, concluí que o valor do *p-value* correspondente a esta estatística de teste é inferior a $2{,}2 \times 10^{-16}$ (ver ANEXO IV).

Como o valor do *p-value* é muito pequeno, rejeito H_0 para todos os níveis de significância usuais. Portanto, há evidência para afirmar que a correlação populacional entre as variáveis difere significativamente de zero. Desta forma, posso admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa desta disciplina.

De seguida, vou analisar o grau de associação que há entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º e 2.º Períodos, a esta disciplina.

Recorrendo ao *software* R, facilmente obtive a seguinte tabela:

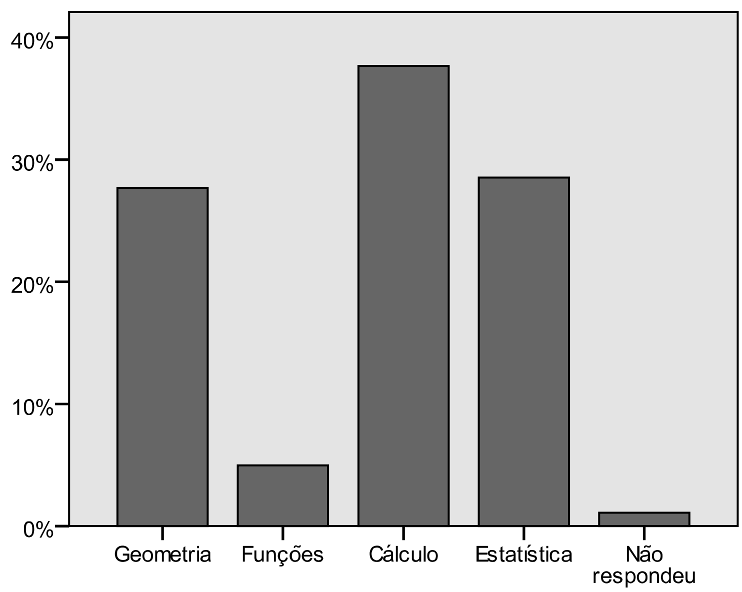
Tabela 2: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

	Como avalia o seu rendimento a Matemática?	
Qual a classificação que obteve no 1.º Período?	$(r_S)_C = 0{,}691$	$p < 2{,}2\times10^{-16}$
Qual a classificação que obteve no 2.º Período?	$(r_S)_C = 0{,}758$	$p < 2{,}2\times10^{-16}$

Pela leitura dos valores da Tabela 2, concluí que há uma associação positiva forte entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram no 1.º Período a esta disciplina. Esta associação torna-se ainda mais forte quando se relaciona a classificação que os alunos obtiveram no 2.º Período com a forma como avaliam o seu rendimento. O que significa que com o decorrer do ano lectivo, os alunos vão conseguindo auto-avaliar-se correctamente no que diz respeito ao seu rendimento na disciplina de Matemática.

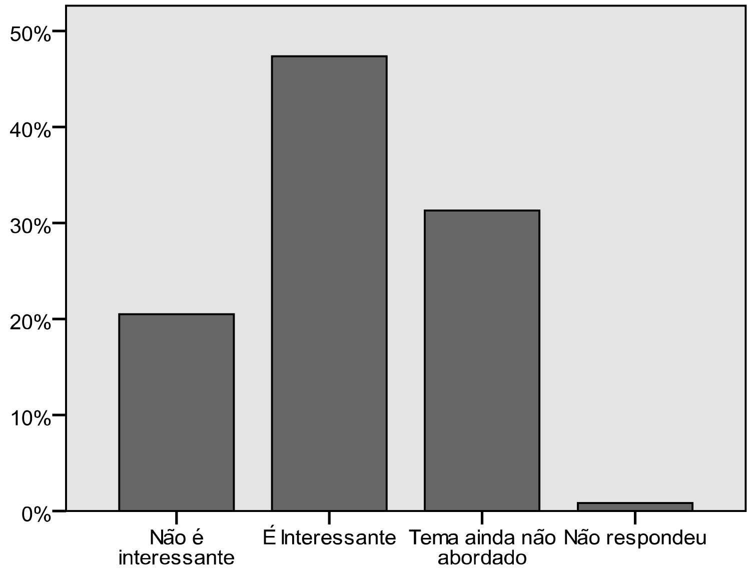
Relativamente aos temas leccionados, durante o ano lectivo, na disciplina de Matemática, o tema Cálculo foi o seleccionado pela maioria dos alunos (37,7%), seguido dos temas Estatística e Geometria com 28,5% e 27,7%, respectivamente (Gráfico 23).

Gráfico 23: Distribuição dos inquiridos de acordo com o tema preferido



Em relação ao capítulo da Estatística, a maioria dos inquiridos (47,4%) considera que o tema é interessante, no entanto, é importante realçar que apesar dos inquéritos terem sido realizados durante o 3.º Período, aproximadamente 31% dos inquiridos ainda não tinha tido, durante ano lectivo, contacto com o tema de Estatística (Gráfico 24). Será que os alunos que ainda não tinham abordado a temática de Estatística chegaram a ter a oportunidade de a abordar? Pois, é do conhecimento de todos que os conteúdos de Estatística nem sempre são leccionados, uma vez que fazem parte das matérias remetidas para o final do programa e, por isso mesmo, nem sempre são apresentadas aos alunos, tanto devido à falta de tempo, como devido à falta de convicção do seu real interesse (Branco, 2000a).

Gráfico 24: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação ao tema de Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 3: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística e a escola que frequentam

			Opinião em relação ao tema de Estatística		
			Não é interessante	É interessante	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	7	19	26
		Valor Esperado	7,9	18,1	26,0
	EB 2,3 Cardoso Lopes	Valor Observado	5	4	9
		Valor Esperado	2,7	6,3	9,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	9	24	33
		Valor Esperado	10,0	23,0	33,0
	EB 2,3/S Michel Giacometti	Valor Observado	30	71	101
		Valor Esperado	30,5	70,5	101,0
	Escola Secundária de Gil Vicente	Valor Observado	7	33	40
		Valor Esperado	12,1	27,9	40,0
	Escola Secundária Ferreira Dias	Valor Observado	16	20	36
		Valor Esperado	10,9	25,1	36,0
	Total	Valor Observado	74	171	245
		Valor Esperado	74,0	171,0	245,0

Segundo Sprent (1993), o teste de Wilcoxon-Mann-Whitney é o mais apropriado para testar a independência entre duas variáveis, quando uma delas é ordinal e a outra é nominal com apenas duas categorias. No caso da variável nominal apresentar mais do que duas categorias então o teste de Wilcoxon-Mann-Whitney é alargado ao teste de Kruskal-Wallis.

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação ao tema de Estatística" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Assim, se H_0 for verdadeira, a estatística de teste de Kruskal-Wallis é dada por:

$$H = \frac{12}{N(N+1)} \sum_{k=1}^{\upsilon} \frac{R_k^2}{n_k} - 3(N+1) \sim \chi^2_{(\upsilon-1)}$$

em que a região de rejeição para a hipótese H_0, ao nível de significância α, é dada por:

$$H > \chi^2_{(\upsilon-1);1-\alpha}$$

onde $\chi^2_{(\upsilon-1);1-\alpha}$ é o quantil de probabilidade $1-\alpha$ da distribuição χ^2 com $(\upsilon-1)$ graus de liberdade.

No caso de haver g grupos de observações empatadas, com τ_i observações no $i-$ésimo grupo de empates, usam-se *ranks* ajustados e a expressão da estatística de teste de Kruskal-Wallis, H, é multiplicada por um factor de correcção:

$$\frac{1}{1-\dfrac{\sum_{i=1}^{g}(\tau_i-1)\tau_i(\tau_i+1)}{(N-1)N(N+1)}}$$

obtendo-se

$$H^* = \frac{\dfrac{12}{N(N+1)}\sum_{k=1}^{\upsilon}\dfrac{R_k^{\;2}}{n_k}-3(N+1)}{1-\dfrac{\sum_{i=1}^{g}\left(\tau_i^{\;3}-\tau_i\right)}{N^3-N}}$$

Segundo Pestana e Velosa (2006) não vale a pena usar a correcção se o número de observações empatadas for baixo.

Sprent e Smeeton (2001), chamam a atenção para que se designarmos por $S=\sum_{k=1}^{\upsilon}\frac{R_k^{\;2}}{n_k}$ a soma dos quadrados dos *ranks* não ajustados, $T=\sum_{k=1}^{\upsilon}\sum_{j=1}^{n_k}\overline{r}_{kj}^{\;2}$ a soma de quadrados total dos *ranks* ajustados, se deve usar $C=\frac{N(N+1)^2}{4}$ para correcção dos empates. Assim,

$$H_C=\frac{(N-1)(S-C)}{T-C}$$

$$H_C=\frac{\sum_{k=1}^{\upsilon}\dfrac{R_k^{\;2}}{n_k}-\dfrac{N(N+1)^2}{4}}{\dfrac{1}{N-1}\left(\sum_{k=1}^{\upsilon}\sum_{j=1}^{n_k}\overline{r}_{kj}^{\;2}-\dfrac{N(N+1)^2}{4}\right)}$$

Com o auxílio do *SPSS* e dos dados da Tabela 3, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 9,50943 e que o *p-value* correspondente é 0,09039.

Com base no valor do *p-value*, concluí que, aos níveis de significância de 1 e 5%, não rejeito H_0. Portanto, não há evidência significativa para poder afirmar que a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística está associada à escola que frequentam.

Segundo Cochran (1954), a aproximação pelo qui-quadrado é válida em tabelas superiores a 2 x 2 desde que todos os valores esperados sejam maiores do que 1 e que apenas 20% das células tenham valor esperado inferior a 5.

Também Lewontin e Felsenstein (1965) defendem que a aproximação pelo qui--quadrado é válida em tabelas 2 x n se todos os valores esperados forem iguais ou superiores a 1.

Uma vez que a Tabela 3 apresenta apenas uma frequência esperada inferior a 5 e que é superior a 1, estou em condições de aplicar o teste do qui-quadrado para testar H_0[2].

Assim, se H_0 for verdadeira, a estatística de teste é dada por:

$$X^2 = \sum_{(i,j)} \frac{\left(n_{ij} - \frac{n_{i.}n_{.j}}{n}\right)^2}{\frac{n_{i.}n_{.j}}{n}} \sim \chi^2_{(r-1)(c-1)}$$

em que a região de rejeição para a hipótese H_0, ao nível de significância α, é dada por:

$$X^2 > \chi^2_{(r-1)(c-1);1-\alpha}$$

[2] Alguns autores (por exemplo, Schweigert (1994)) defendem que o teste do qui-quadrado não deve ser conduzido mesmo se apenas um dos valores esperados for inferior a 5. Pois, uma frequência esperada pequena tem um impacto maior no valor total do qui-quadrado do que uma frequência esperada muito grande, especialmente quando a frequência esperada é menor que cinco.

onde $\chi^2_{(r-1)(c-1);1-\alpha}$ é o quantil de probabilidade $1-\alpha$ da distribuição χ^2 com (r-1)(c-1) graus de liberdade.

Com o auxílio do *SPSS* e dos dados da Tabela 3, concluí que o valor da estatística de teste, X^2, é 9,54841 e que o *p-value* correspondente a esta estatística de teste é 0,08909.

Com base no valor do *p-value*, concluí que, aos níveis de significância de 1 e 5%, não rejeito H_0. Portanto, não há evidência significativa para poder afirmar que a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística está associada à escola que frequentam.

Uma estatística de teste alternativa à do qui-quadrado para testar H_0 é a estatística de Likelihood Ratio que é dada por:

$$G^2 = 2\sum_{i,j} n_{ij} \log\left(\frac{n_{ij}}{\frac{n_{i.}n_{.j}}{n}}\right) \sim \chi^2_{(r-1)(c-1)}$$

em que “log” é o logaritmo de base “10” ou o logaritmo de base “e”.

A região de rejeição para a hipótese H_0, ao nível de significância α, é dada por:

$$G^2 > \chi^2_{(r-1)(c-1);1-\alpha}$$

onde $\chi^2_{(r-1)(c-1);1-\alpha}$ é o quantil de probabilidade $1-\alpha$ da distribuição χ^2 com (r-1)(c-1) graus de liberdade.

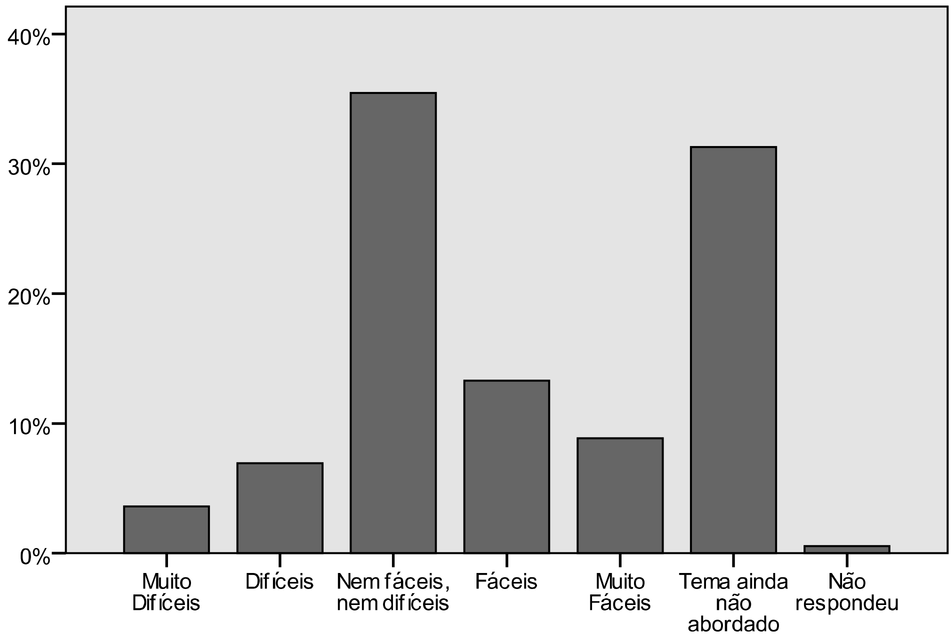
Segundo alguns autores (por exemplo, Sprent (1993)), o teste de Likelihood Ratio é melhor do que o teste do qui-quadrado porque a estatística de teste, G^2, geralmente converge mais rapidamente para a distribuição do χ^2, sob H_0, do que a estatística de teste X^2.

Recorrendo novamente ao *SPSS* e aos dados da Tabela 3, concluí que o valor da estatística de teste de Likelihood Ratio é 9,39687 e que o *p-value* correspondente a esta estatística de teste é 0,09424.

Com base no valor do *p-value*, concluí que, aos níveis de significância de 1 e 5%, não rejeito H_0. Portanto, não há evidência significativa para poder afirmar que a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística está associada à escola que frequentam.

Dos alunos que já tinham tido contacto com a temática de Estatística, cerca de 36% consideram que os conteúdos desta unidade não são fáceis nem são difíceis e, aproximadamente, 13% avaliam-nos como sendo fáceis (Gráfico 25).

Gráfico 25: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação aos conteúdos de Estatística, vou realizar um teste de independência para testar:

> *H_0: Há independência entre a opinião que os alunos do 7.º ano têm em relação aos conteúdos de Estatística e a escola que frequentam.*

Da amostra recolhida, obtive seguinte tabela:

Tabela 4: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística e a escola que frequentam

			Opinião em relação aos conteúdos de Estatística			
			Fáceis e Muito Fáceis	Nem fáceis nem difíceis	Difíceis e Muito Difíceis	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	13	9	4	26
		Valor Esperado	8,5	13,5	4,0	26,0
	EB 2,3 Cardoso Lopes	Valor Observado	1	4	4	9
		Valor Esperado	2,9	4,7	1,4	9,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	11	14	8	33
		Valor Esperado	10,7	17,2	5,1	33,0
	EB 2,3/S Michel Giacometti	Valor Observado	29	60	12	101
		Valor Esperado	32,8	52,6	15,6	101,0
	Escola Secundária de Gil Vicente	Valor Observado	19	20	2	41
		Valor Esperado	13,3	21,3	6,3	41,0
	Escola Secundária Ferreira Dias	Valor Observado	7	21	8	36
		Valor Esperado	11,7	18,7	5,6	36,0
	Total	Valor Observado	80	128	38	246
		Valor Esperado	80,0	128,0	38,0	246,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação aos conteúdos de Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 4, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 15,56382 e que o *p-value* correspondente é 0,00821.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para todos os níveis de significância usuais. Portanto, há evidência significativa para afirmar que a opinião que os alunos do 7.º ano têm em relação aos conteúdos de Estatística está associada à escola que frequentam.

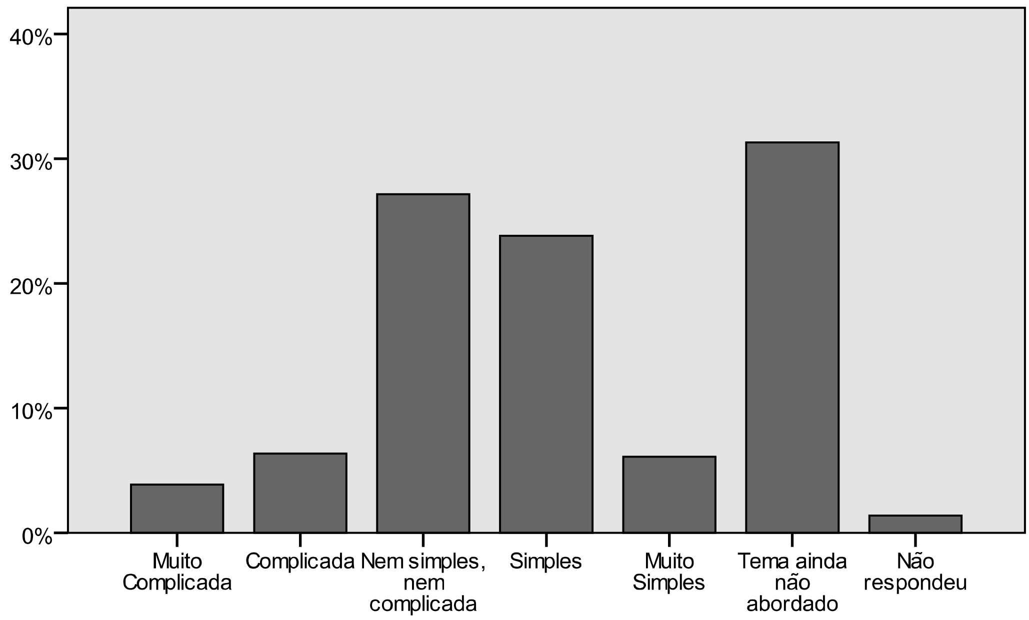
Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação aos conteúdos de Estatística” ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois mais de 20% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$ (proposto por Freeman e Halton (1951)).

Com o auxílio do *software* R e da Tabela 4, concluí que o valor do *p-value* associado ao teste exacto é 0,01098 (ver ANEXO V).

Com base no valor do *p-value*, concluí que, ao nível de significância de 1%, não rejeito H_0, portanto, não há evidência significativa para poder afirmar que a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística está associada à escola que frequentam. No entanto, ao nível de significância de 5%, rejeito H_0, portanto, há evidência significativa para poder afirmar que a opinião que os alunos do 7.º ano têm em relação ao tema de Estatística está associada à escola que frequentam. Uma vez que estes resultados são pouco conclusivos, no futuro, este estudo deveria ser alargado a mais escolas e a mais alunos de forma a conseguir-se obter um resultado conclusivo.

No que diz respeito à forma como os conteúdos de Estatística foram abordados, aproximadamente 27% dos inquiridos, que já tinham tido contacto com esta temática, consideram que a forma como os professores abordaram os conteúdos de Estatística não foi simples nem foi complicada, seguindo-se a forma simples com 23,8% de respostas (Gráfico 26).

Gráfico 26: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados



Para saber se há associação entre a escola e a opinião que os alunos têm em relação à forma como os conteúdos de Estatística foram abordados, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 7.º ano têm em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 5: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam

			Opinião em relação à forma como os conteúdos de Estatística foram abordados			
			Simples e Muito Simples	Nem simples Nem complicada	Complicada e Muito Complicada	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	11	12	3	26
		Valor Esperado	11,6	10,5	4,0	26,0
	EB 2,3 Cardoso Lopes	Valor Observado	2	4	3	9
		Valor Esperado	4,0	3,6	1,4	9,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	18	8	7	33
		Valor Esperado	14,7	13,3	5,0	33,0
	EB 2,3/S Michel Giacometti	Valor Observado	45	46	10	101
		Valor Esperado	44,9	40,7	15,4	101,0
	Escola Secundária de Gil Vicente	Valor Observado	23	14	1	38
		Valor Esperado	16,9	15,3	5,8	38,0
	Escola Secundária Ferreira Dias	Valor Observado	9	14	13	36
		Valor Esperado	16,0	14,5	5,5	36,0
	Total	Valor Observado	108	98	37	243
		Valor Esperado	108,0	98,0	37,0	243,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação à forma como os conteúdos de Estatística foram abordados" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 5, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 19,11298 e que o *p-value* correspondente é 0,00183.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para todos os níveis de significância usuais. Portanto, há evidência significativa para afirmar que a opinião que os alunos do 7.º ano têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação à forma como os conteúdos de Estatística foram abordados" ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois mais de 20% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$.

Com o auxílio do *software* R e da Tabela 5, concluí que o valor do *p-value* associado ao teste exacto é 0,001379.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para todos os níveis de significância usuais. Portanto, há evidência significativa para afirmar que a opinião que os alunos do 7.º ano têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

De seguida, vou analisar, por escola, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados; bem como, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados.

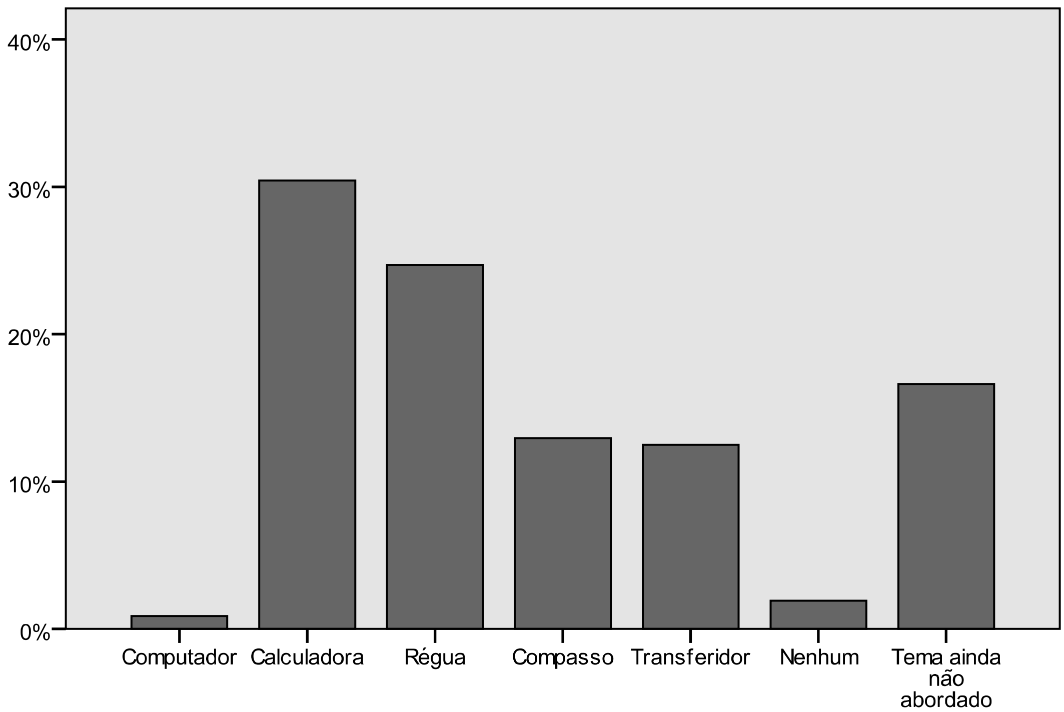
Tabela 6: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

		Opinião dos alunos em relação à forma como os conteúdos de Estatística foram abordados	
EB 2, 3 Navegador Rodrigues Soromenho	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}415$	$p = 0{,}035$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}811$	$p = 4{,}968 \times 10^{-7}$
EB 2,3 Cardoso Lopes	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = -0{,}046$	$p = 0{,}907$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}654$	$p = 0{,}056$
EB 2,3 Mestre Domingos Saraiva	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}433$	$p = 0{,}012$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}721$	$p = 2{,}199 \times 10^{-6}$
EB 2,3/S Michel Giacometti	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}261$	$p = 0{,}009$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}453$	$p = 1{,}937 \times 10^{-6}$
Escola Secundária de Gil Vicente	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}127$	$p = 0{,}447$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}722$	$p = 3{,}176 \times 10^{-7}$
Escola Secundária Ferreira Dias	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}534$	$p = 7{,}866 \times 10^{-4}$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}627$	$p = 4{,}302 \times 10^{-5}$

Pela análise da Tabela 6, concluí que, ao nível de significância de 5%, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados é, na maioria das escolas, positivo e moderado. Concluí ainda que o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados é, na maioria das escolas, positivo e forte. Sendo de admitir que, os alunos que estão satisfeitos com a forma como o docente abordou os conteúdos de Estatística são também aqueles que têm uma opinião mais positiva, quer em relação ao tema quer em relação aos conteúdos.

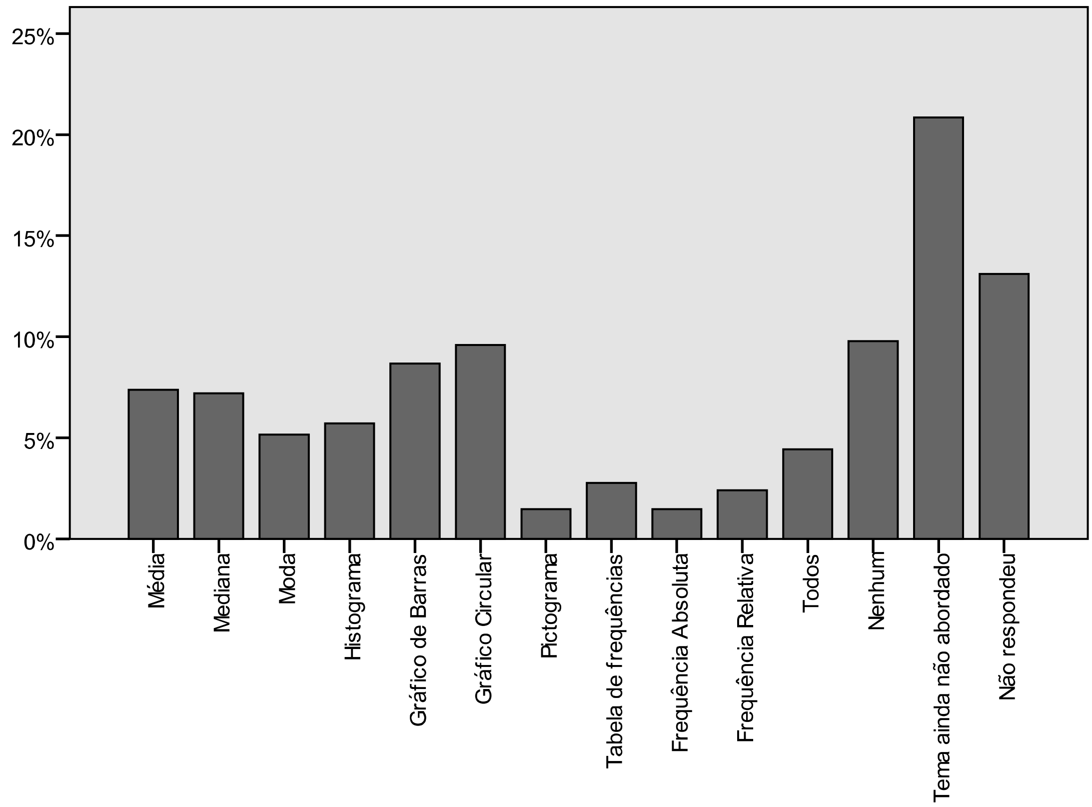
No que diz respeito aos materiais utilizados durante as aulas de Estatística, o Gráfico 27 permitiu-me concluir que a maioria dos inquiridos (30,4%) utilizou máquina de calcular e que 24,7% de alunos utilizaram régua. O compasso e o transferidor foram também utilizados durante as aulas de Estatística por, aproximadamente, 13% de alunos, cada um deles. O computador, por sua vez, foi utilizado durante as aulas de Estatística por um número muito reduzido de alunos (0,9%). Isto significa que, apesar de o programa de Matemática sugerir a utilização de computadores, são poucos os professores que planificam as suas aulas com vista à utilização desta tecnologia. Esta situação prende-se essencialmente com o facto de não haver recursos físicos, nem materiais adequados nas escolas, como por exemplo uma sala de trabalho com computadores à disposição dos alunos ou ainda com o facto de a maioria das aulas de Matemática não decorrer em salas com computadores.

Gráfico 27: Materiais usados durante as aulas de Estatística



Finalmente, no que diz respeito aos conteúdos de Estatística, grande parte dos alunos refere que sentiu dificuldades em pelo menos um dos conteúdos abordados. Pela leitura do Gráfico 28, constatei que o gráfico circular é o conteúdo em que os alunos sentiram mais dificuldades (9,6% dos inquiridos indicou este conteúdo). De seguida, vem o gráfico de barras assinalado por 8,7% dos inquiridos. Posteriormente, surge a média e a mediana como sendo conteúdos onde os alunos também sentiram dificuldades (cada um deles foi assinalado por, aproximadamente, 7% dos inquiridos). Para além destes valores, é importante realçar que 9,8% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos de Estatística e que 4,4% dos inquiridos afirmam ter sentido dificuldades em todos os conteúdos. Na minha opinião, a comunidade escolar deve ter em atenção estes últimos alunos, pois poderão estar a necessitar de algum tipo de apoio ou acompanhamento. É também importante destacar que, aproximadamente, 13% dos inquiridos não responderam a esta questão. Esta situação poderá ter estado relacionada com o facto de alguns alunos já não se lembrarem dos conteúdos de Estatística ou ainda com o facto de se tratar de uma questão de resposta aberta.

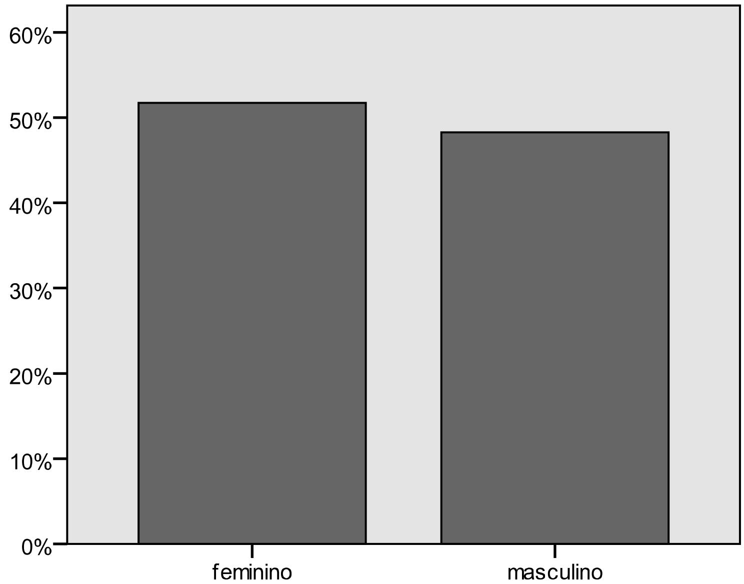
Gráfico 28: Conteúdos de Estatística em que os alunos sentiram dificuldades



5.2.2. 8.º Ano do Ensino Básico

Foram inquiridos 165 estudantes do género feminino, o que corresponde a 51,7% da amostra e do género masculino foram inquiridos 154 alunos, o que equivale a 48,3% da amostra (Gráfico 29).

Gráfico 29: Distribuição por género



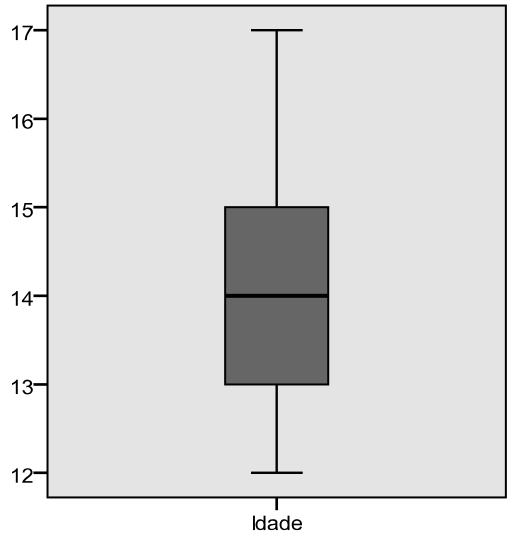
Pela análise da Tabela 7 e do Gráfico 30, concluí que dos 319 alunos inquiridos, o mais novo tem 12 anos e o mais velho tem 17 anos, ou seja, há uma grande variação das idades dos alunos que frequentam o 8.º ano de escolaridade. A moda e a mediana das idades são de 14 anos. A média das idades dos inquiridos é de 14,14 anos e a dispersão das observações em relação à média – desvio padrão – é de 1,04 anos. Com base nos valores dos quartis, sabe-se que 50% dos inquiridos têm idade compreendida entre os 13 e os 15 anos.

Analisando os valores anteriores, concluí que, de facto, faz sentido que a moda das idades seja de 14 anos, pois se um aluno fizer um percurso normal na escola quando chegar ao 8.º ano tem 13 ou 14 anos, consoante a data do seu aniversário. No entanto, verifica-se a existência de um inquirido com apenas 12 anos, o que significa que, provavelmente, entrou para a escola antes da idade prevista. Além disso, é do conhecimento geral que nem todos os alunos apresentam um percurso normal na escola, daí haver muitos inquiridos com idade superior à moda (este facto é facilmente confirmado através do Gráfico 30).

Tabela 7: Estatísticas da variável idade

N	319
Média	14,14
Mediana	14
Moda	14
Variância	1,08
Desvio Padrão	1,04
Mínimo	12
Máximo	17
1.º Quartil	13
2.º Quartil	14
3.º Quartil	15

Gráfico 30: *Box-plot* das idades dos inquiridos



No que diz respeito à opinião que os alunos do 8.º ano têm em relação à disciplina de Matemática, a maioria dos inquiridos (37%) afirma gostar de Matemática (Gráfico 31). No entanto, a maioria dos inquiridos (42,3%) auto-avalia de Suficiente o seu rendimento a esta disciplina, seguido de Insuficiente com 25,7% (Gráfico 32). De facto, ao consultar as classificações que os alunos obtiveram no 1.º e no 2.º Períodos, concluí que, no 1.º Período, a classificação que mais se verificou foi 2 com 42,8%, seguida da classificação 3 com 39% (Gráfico 33); no 2.º Período, a classificação que mais se verificou foi 3 com 41,6%, seguida da classificação 2 com 39,7% (Gráfico 34).

Daqui concluí, que apesar da maioria dos alunos do 8.º ano gostar de Matemática, o seu rendimento a esta disciplina não é equivalente à opinião que têm da mesma. Além disso, constata-se que o rendimento escolar dos alunos na disciplina de Matemática tende a diminuir do 7.º para o 8.º ano.

Gráfico 31: Distribuição dos inquiridos segundo a sua opinião em relação à disciplina de Matemática

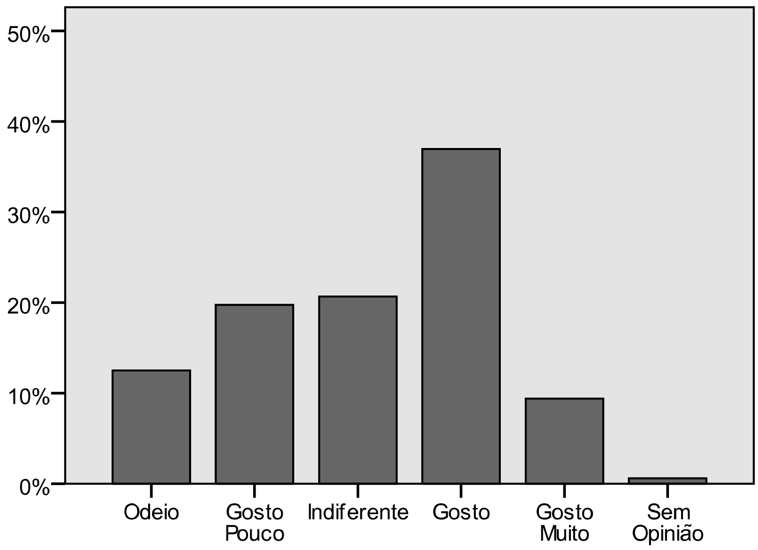


Gráfico 32: Distribuição dos inquiridos segundo a sua própria avaliação do seu rendimento a Matemática

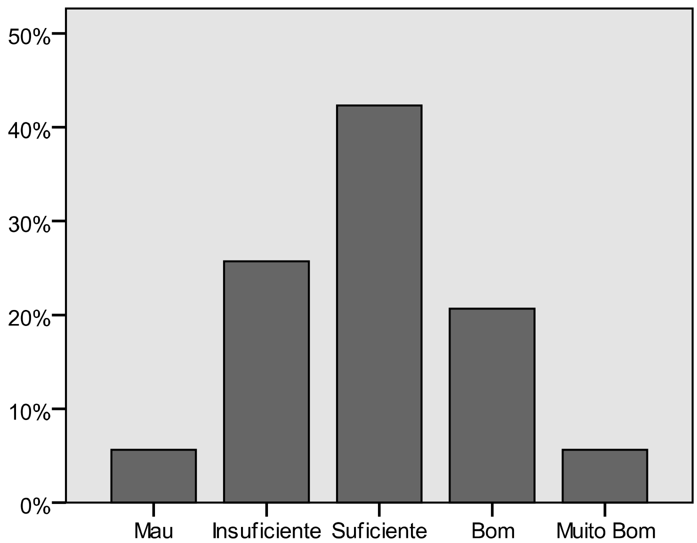


Gráfico 33: Classificação que os alunos obtiveram no 1.º Período

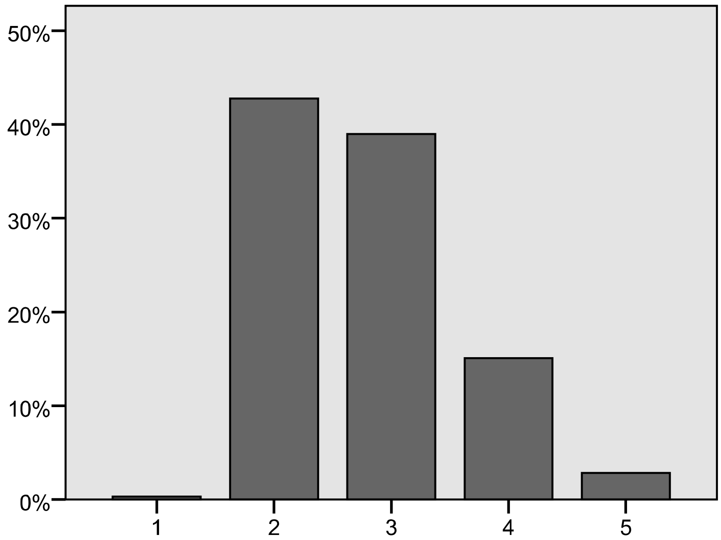
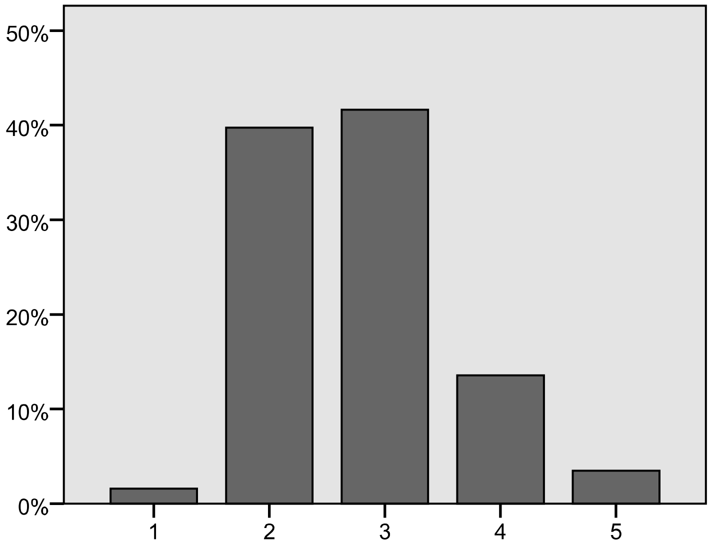


Gráfico 34: Classificação que os alunos obtiveram no 2.º Período



Ao recorrer ao *software* R para saber o tipo de associação que há entre a opinião que os alunos do 8.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina, concluí que o valor do coeficiente de correlação de Spearman é 0,611 e que o valor do *p-value* associado ao teste $H_0 : \rho_S = 0$ é inferior a $2{,}2 \times 10^{-16}$. Face a estes valores, concluí que há uma associação positiva moderada entre a opinião que os alunos do 8.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina. Sendo, portanto, de admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa desta disciplina.

De seguida, vou analisar o grau de associação que há entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º e 2.º Períodos, a esta disciplina.

Recorrendo ao *software* R, facilmente obtive a seguinte tabela:

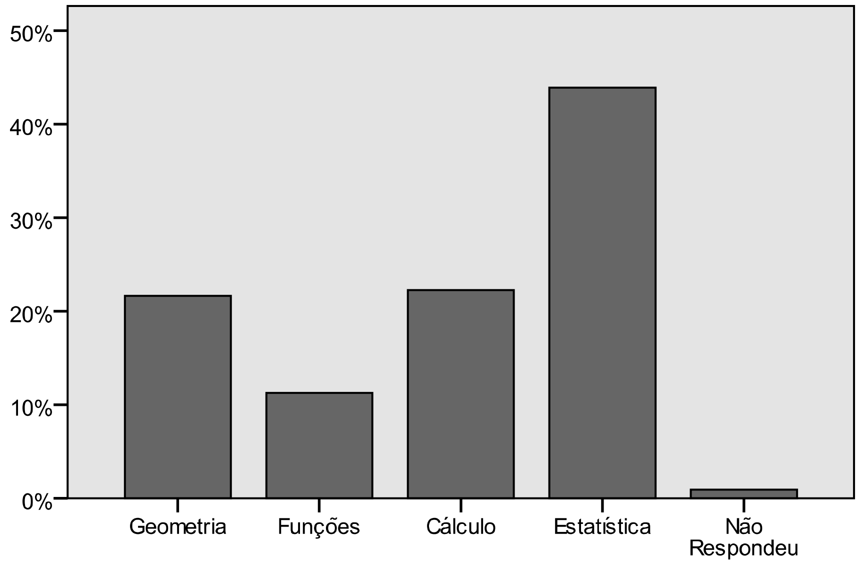
Tabela 8: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

	Como avalia o seu rendimento a Matemática?	
Qual a classificação que obteve no 1.º Período?	$(r_S)_C = 0{,}772$	$p < 2{,}2 \times 10^{-16}$
Qual a classificação que obteve no 2.º Período?	$(r_S)_C = 0{,}806$	$p < 2{,}2 \times 10^{-16}$

Pela leitura dos valores da Tabela 8, concluí que há uma associação positiva forte entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º Período, a esta disciplina. Esta associação torna--se ainda mais forte quando se relaciona a classificação que os alunos obtiveram no 2.º Período com a forma como avaliam o seu rendimento. O que significa que ao longo do ano lectivo, os alunos vão conseguindo auto-avaliar-se correctamente no que diz respeito ao seu rendimento na disciplina de Matemática.

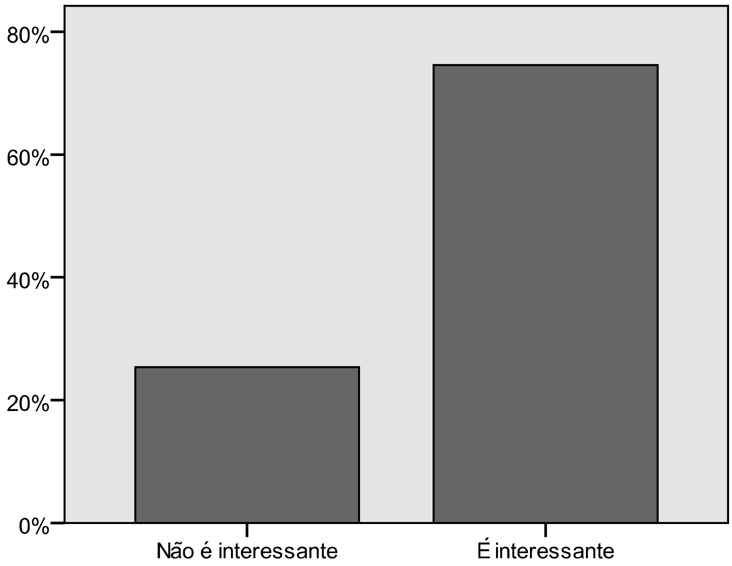
Quanto aos temas abordados ao longo do ano lectivo na disciplina de Matemática, a maioria dos alunos (43,9%) afirma preferir a unidade temática de Estatística, seguida das temáticas de Cálculo e de Geometria com, aproximadamente, 22% cada uma (Gráfico 35).

Gráfico 35: Distribuição dos inquiridos de acordo com o tema preferido



Em relação à temática de Estatística, a grande maioria dos alunos inquiridos (74,6%) considera o tema interessante. O contrário não seria de esperar, uma vez que o tema de Estatística é considerado o tema de eleição da maioria dos inquiridos. No entanto, os restantes alunos (25,4%) consideram que o tema não é interessante (Gráfico 36).

Gráfico 36: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação ao tema de Estatística, vou realizar um teste de independência para testar:

> *H_0: Há independência entre a opinião que os alunos do 8.º ano têm em relação ao tema de Estatística e a escola que frequentam.*

Da amostra recolhida, obtive a seguinte tabela:

Tabela 9: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística e a escola que frequentam

<table>
<tr><td colspan="3"></td><td colspan="3">Opinião em relação ao tema da Estatística</td></tr>
<tr><td colspan="3"></td><td>Não é interessante</td><td>É interessante</td><td>Total</td></tr>
<tr><td rowspan="14">Nome da Escola</td><td rowspan="2">EB 2, 3 Navegador Rodrigues Soromenho</td><td>Valor Observado</td><td>4</td><td>8</td><td>12</td></tr>
<tr><td>Valor Esperado</td><td>3,0</td><td>9,0</td><td>12,0</td></tr>
<tr><td rowspan="2">EB 2,3 Cardoso Lopes</td><td>Valor Observado</td><td>23</td><td>47</td><td>70</td></tr>
<tr><td>Valor Esperado</td><td>17,8</td><td>52,2</td><td>70,0</td></tr>
<tr><td rowspan="2">EB 2,3 Mestre Domingos Saraiva</td><td>Valor Observado</td><td>18</td><td>52</td><td>70</td></tr>
<tr><td>Valor Esperado</td><td>17,8</td><td>52,2</td><td>70,0</td></tr>
<tr><td rowspan="2">EB 2,3/S Michel Giacometti</td><td>Valor Observado</td><td>20</td><td>66</td><td>86</td></tr>
<tr><td>Valor Esperado</td><td>21,8</td><td>64,2</td><td>86,0</td></tr>
<tr><td rowspan="2">Escola Secundária de Gil Vicente</td><td>Valor Observado</td><td>12</td><td>37</td><td>49</td></tr>
<tr><td>Valor Esperado</td><td>12,4</td><td>36,6</td><td>49,0</td></tr>
<tr><td rowspan="2">Escola Secundária Da Baixa da Banheira</td><td>Valor Observado</td><td>4</td><td>28</td><td>32</td></tr>
<tr><td>Valor Esperado</td><td>8,1</td><td>23,9</td><td>32,0</td></tr>
<tr><td rowspan="2">Total</td><td>Valor Observado</td><td>81</td><td>238</td><td>319</td></tr>
<tr><td>Valor Esperado</td><td>81,0</td><td>238,0</td><td>319,0</td></tr>
</table>

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação ao tema de Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal--Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 9, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 5,48090 e que o *p-value* correspondente a esta estatística de teste é 0,36005.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação ao tema de Estatística está associada à escola que frequentam.

Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação ao tema de Estatística” ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas um dos valores esperados é inferior a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e dos dados da Tabela 9, obtive a seguinte tabela:

Tabela 10: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

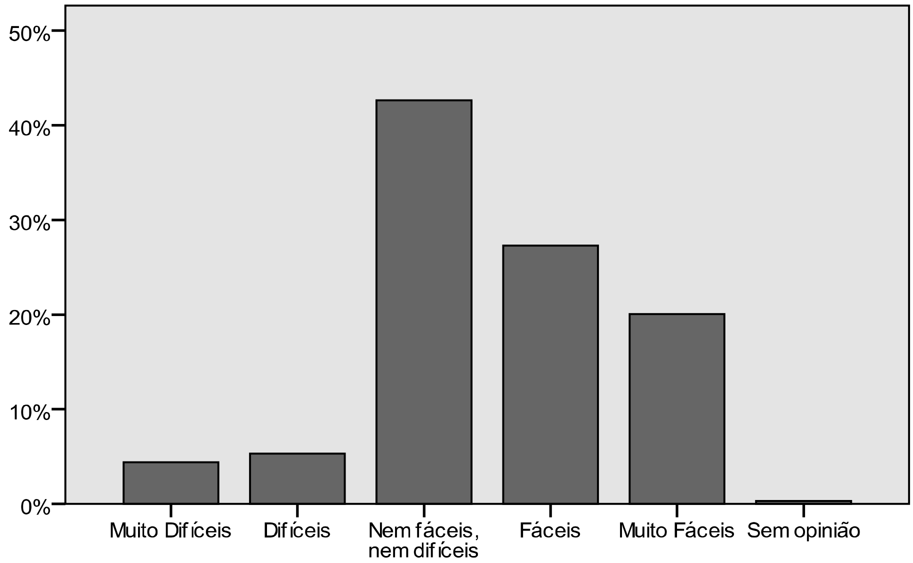
Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	5,49813	5	0,35815
Likelihood Ratio	5,81464	5	0,32468

Através da análise da Tabela 10, concluí que o valor da estatística de teste, X^2, é 5,49813 e que o *p-value* correspondente a esta estatística de teste é 0,35815. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 5,81464 e que o valor do *p-value* correspondente a esta estatística de teste é 0,32468.

Com base nos valores dos *p-values*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação ao tema de Estatística está associada à escola que frequentam.

No que diz respeito à opinião que os alunos têm em relação aos conteúdos de Estatística, a maioria dos inquiridos (42,6%) considera que os conteúdos nem são fáceis nem são difíceis, seguido de fáceis e de muito fáceis com 27,3% e 20,1%, respectivamente (Gráfico 37).

Gráfico 37: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação aos conteúdos de Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 8.º ano têm em relação aos conteúdos de Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 11: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística e a escola que frequentam

			Opinião em relação aos conteúdos de Estatística			
			Fáceis e Muito Fáceis	Nem fáceis nem difíceis	Difíceis e Muito Difíceis	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	5	4	3	12
		Valor Esperado	5,7	5,1	1,2	12,0
	EB 2,3 Cardoso Lopes	Valor Observado	28	35	7	70
		Valor Esperado	33,2	29,9	6,8	70,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	34	26	9	69
		Valor Esperado	32,8	29,5	6,7	69,0
	EB 2,3/S Michel Giacometti	Valor Observado	45	37	4	86
		Valor Esperado	40,8	36,8	8,4	86,0
	Escola Secundária de Gil Vicente	Valor Observado	22	22	5	49
		Valor Esperado	23,3	21,0	4,8	49,0
	Escola Secundária Da Baixa da Banheira	Valor Observado	17	12	3	32
		Valor Esperado	15,2	13,7	3,1	32,0
	Total	Valor Observado	151	136	31	318
		Valor Esperado	151,0	136,0	31,0	318,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação aos conteúdos de Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 11, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 4,01282 e que o *p-value* correspondente a esta estatística de teste é 0,54757.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação aos conteúdos de Estatística está associada à escola que frequentam.

Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação aos conteúdos de Estatística” ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas 16,7% dos valores esperados são inferiores a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e da Tabela 11, obtive a seguinte tabela:

Tabela 12: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

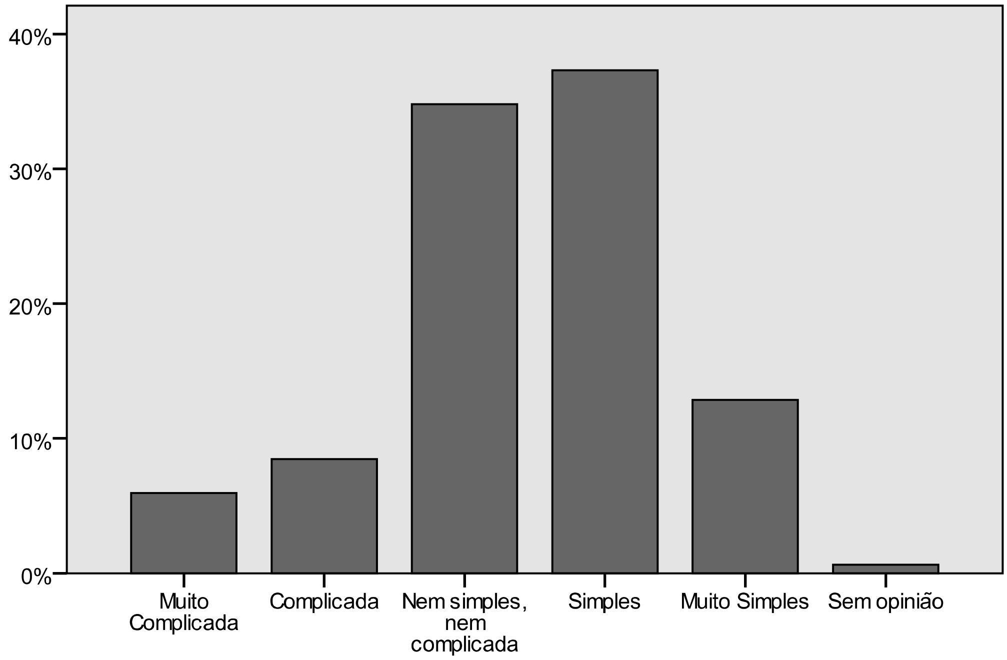
Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	9,39361	10	0,49520
Likelihood Ratio	9,03307	10	0,52897

Através da análise da Tabela 12, concluí que o valor da estatística de teste, X^2, é 9,39361 e que o *p-value* correspondente a esta estatística de teste é 0,4952. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 9,03307 e que o valor do *p-value* correspondente a esta estatística de teste é 0,52897.

Com base nos valores dos *p-values*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação aos conteúdos de Estatística está associada à escola que frequentam.

No que diz respeito à opinião que os alunos do 8.º ano têm em relação à forma como os conteúdos foram abordados, a maioria dos alunos (37,3%) considera que os professores de Matemática abordaram os conteúdos de uma forma simples, seguida de uma forma que nem foi simples nem foi complicada com 34,8% (Gráfico 38).

Gráfico 38: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados



Para saber se há associação entre a escola e a opinião que os alunos têm em relação à forma como os conteúdos de Estatística foram abordados, vou realizar um teste de independência para testar:

> *H_0: Há independência entre a opinião que os alunos do 8.º ano têm em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam.*

Da amostra recolhida, obtive a seguinte tabela:

Tabela 13: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam

			Opinião em relação à forma como os conteúdos de Estatística foram abordados			
			Simples e Muito Simples	Nem simples Nem complicada	Complicada e Muito Complicada	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	6	5	1	12
		Valor Esperado	6,1	4,2	1,7	12,0
	EB 2,3 Cardoso Lopes	Valor Observado	35	26	9	70
		Valor Esperado	35,3	24,5	10,2	70,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	36	23	10	69
		Valor Esperado	34,8	24,2	10,0	69,0
	EB 2,3/S Michel Giacometti	Valor Observado	40	32	13	85
		Valor Esperado	42,9	29,8	12,3	85,0
	Escola Secundária de Gil Vicente	Valor Observado	25	16	8	49
		Valor Esperado	24,7	17,2	7,1	49,0
	Escola Secundária da Baixa da Banheira	Valor Observado	18	9	5	32
		Valor Esperado	16,2	11,2	4,6	32,0
	Total	Valor Observado	160	111	46	317
		Valor Esperado	160,0	111,0	46,0	317,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação à forma como os conteúdos de Estatística foram abordados” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e da Tabela 13, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 0,65166 e que o *p-value* correspondente a esta estatística de teste é 0,98552.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação à forma como os conteúdos de Estatística foram abordados” ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas 16,7% dos valores esperados são inferiores a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e dos dados da Tabela 13, obtive a seguinte tabela:

Tabela 14: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	2,05408	10	0,99591
Likelihood Ratio	2,12775	10	0,99526

Através da análise da Tabela 14, concluí que o valor da estatística de teste, X^2, é 2,05408 e que o *p-value* correspondente a esta estatística de teste é 0,99591. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 2,12775 e que o valor do *p-value* correspondente a esta estatística de teste é 0,99526.

Com base nos valores dos *p-values*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência significativa para afirmar que a opinião que os alunos do 8.º ano têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

De seguida, vou analisar, por escola, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados; bem como, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como os mesmos foram leccionados.

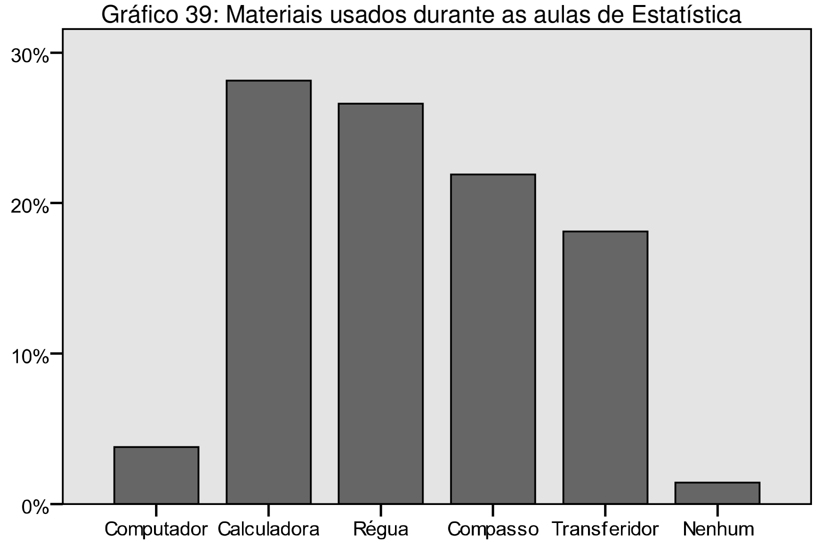
Tabela 15: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

		Opinião dos alunos em relação à forma como os conteúdos de Estatística foram abordados	
EB 2, 3 Navegador Rodrigues Soromenho	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}256$	$p = 0{,}421$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}569$	$p = 0{,}054$
EB 2,3 Cardoso Lopes	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}265$	$p = 0{,}027$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}678$	$p = 1{,}163 \times 10^{-10}$
EB 2,3 Mestre Domingos Saraiva	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}574$	$p = 3{,}181 \times 10^{-7}$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}512$	$p = 7{,}97 \times 10^{-6}$
EB 2,3/S Michel Giacometti	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}379$	$p = 3{,}451 \times 10^{-4}$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}627$	$p = 1{,}340 \times 10^{-10}$
Escola Secundária de Gil Vicente	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}603$	$p = 4{,}593 \times 10^{-6}$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}599$	$p = 5{,}328 \times 10^{-6}$
Escola Secundária Baixa da Banheira	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}520$	$p = 2{,}262 \times 10^{-3}$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}732$	$p = 1{,}962 \times 10^{-6}$

Pela análise da Tabela 15, concluí que, ao nível de significância de 5%, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados é, na maioria das escolas, moderado e positivo. Concluí ainda que o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram abordados é, na maioria das escolas, moderado e positivo, no entanto, chega a ser forte em algumas das escolas. Portanto, pode-se admitir que os alunos que estão satisfeitos com a forma como o docente abordou os conteúdos de Estatística são também aqueles que têm uma opinião satisfatória, quer em relação ao tema de Estatística quer em relação aos conteúdos deste tema.

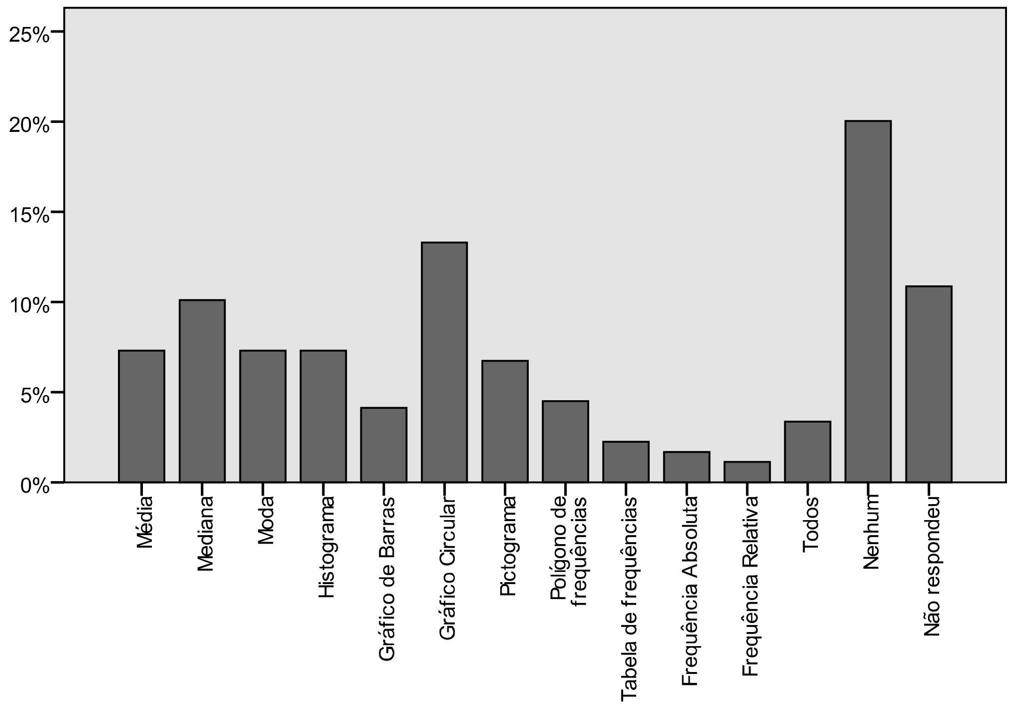
No que diz respeito aos materiais utilizados durante as aulas de Estatística, o Gráfico 39 permitiu-me concluir que a maioria dos alunos (28,1%) utilizou a máquina de calcular. A régua, o compasso e o transferidor foram também utilizados por 26,6%, 21,9% e 18,1% de alunos, respectivamente. O computador foi utilizado durante as aulas de Estatística por um número muito reduzido de alunos, apesar de ser um dos recursos sugeridos pelo programa de Matemática do 3.º Ciclo do Ensino Básico. Tal como já referi, esta situação deve-se, essencialmente, ao facto de a maioria das aulas de Matemática não decorrerem em salas com computadores e ainda, ao facto de não haver salas de trabalho com computadores à disposição dos alunos.

Gráfico 39: Materiais usados durante as aulas de Estatística



Finalmente, no que se refere aos conteúdos de Estatística, a maioria dos alunos afirma ter sentido dificuldades em pelo menos um dos conteúdos leccionados. Da análise do Gráfico 40, constatei que 13,3% de alunos afirmam ter sentido dificuldades no gráfico circular; 10,1% de alunos referem a mediana; a média, a moda e o histograma são mencionados por 7,3% de alunos, cada um; o pictograma é assinalado por 6,7% de alunos. Para além destes valores, é importante realçar que 20% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos de Estatística e que 3,4% dos inquiridos afirmam ter sentido dificuldades em todos os conteúdos. Na minha opinião, a comunidade escolar deve ter em atenção estes últimos alunos, pois poderão estar a necessitar de algum tipo de apoio ou acompanhamento. É também importante destacar que, aproximadamente, 11% dos inquiridos não responderam a esta questão. Esta situação poderá ter estado relacionada com o facto de alguns alunos já não se lembrarem dos conteúdos de Estatística ou ainda com o facto de se tratar de uma questão de resposta aberta.

Gráfico 40: Conteúdos de Estatística em que os alunos sentiram dificuldades



5.2.3. Comparação entre o 7.º e o 8.º Anos do Ensino Básico

Com base nas percentagens de alunos do 8.º ano que referem ter dificuldades nos conteúdos de Estatística e nas percentagens de alunos que referem os mesmos conteúdos, mas que estão a frequentar o 7.º ano, será que se pode afirmar que a proporção de alunos do 8.º ano com dificuldades nos conteúdos de Estatística é superior à de alunos do 7.º ano?

Para poder responder a esta questão e uma vez que as amostras são independentes, vou realizar, para cada escola, um teste de hipóteses sobre a diferença de proporções referente a cada um dos conteúdos abordados nas aulas de Estatística.

EB 2,3 Navegador Rodrigues Soromenho

Sejam X_1 e X_2 as variáveis aleatórias que representam o número de alunos do 7.º e 8.º anos, respectivamente, com dificuldades no cálculo da média.

Sejam p_1 e p_2 as proporções de alunos do 7.º e 8.º anos, respectivamente, com dificuldades no cálculo da média.

Pretende-se testar o seguinte:

$$H_0 : p_1 \geq p_2 \qquad vs. \qquad H_1 : p_1 < p_2$$

Sob H_0, a estatística de teste é dada por:

$$Z = \frac{|\hat{p}_1 - \hat{p}_2|}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}} \sim N(0,1)$$

Onde, $\overline{p} = \dfrac{x_1 + x_2}{n_1 + n_2}$ ou $\overline{p} = \dfrac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2}$; $\overline{q} = 1 - \overline{p}$

Alguns autores utilizam $\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$ como denominador da Estatística Z (ver Eberhardt e Fligner, 1977). Se $n_1 = n_2$ não existe diferença no resultado, no entanto, se $n_1 \neq n_2$ pode dar origem a um erro do tipo I (rejeitar H_0 quando H_0 é verdadeira) maior que o nível de significância α, o que não é desejável.

Muitos autores consideram importante o uso de uma “correcção de continuidade”, uma vez que se está a aproximar uma distribuição discreta pela distribuição da Normal, que é contínua (por exemplo, Zar, 1999). Desta forma, a estatística de teste que tem em conta a correcção de continuidade é dada por:

$$Z_C = \frac{\left|\hat{p}_1 - \hat{p}_2\right| - \frac{1}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}{\sqrt{\frac{\overline{p}\overline{q}}{n_1} + \frac{\overline{p}\overline{q}}{n_2}}} \underset{\text{Sob } H_0}{\sim} N(0,1)$$

em que a região de rejeição para a hipótese H_0, ao nível de significância α, é dada por:

$$Z_C(obs.) \leq -Z_{1-\alpha}$$

onde $Z_{1-\alpha}$ é o quantil de probabilidade $1-\alpha$ de uma $N(0,1)$.

Na escola EB 2,3 Navegador Rodrigues Soromenho foram inquiridos 74 alunos do 7.º ano, dos quais 11 afirmam ter sentido dificuldades no cálculo da média e foram inquiridos 20 alunos do 8.º ano, dos quais apenas 2 afirmam ter sentido dificuldades no cálculo da média.

Assim, $\hat{p}_1 = \frac{11}{74}$; $\hat{p}_2 = \frac{2}{20}$; $\overline{p} = \frac{13}{94}$; $\overline{q} = \frac{81}{94}$

$$Z_C(obs.) = \frac{\left|\frac{11}{74} - \frac{2}{20}\right| - \frac{1}{2}\left(\frac{1}{74} + \frac{1}{20}\right)}{\sqrt{\frac{\frac{13}{94}\times\frac{81}{94}}{74} + \frac{\frac{13}{94}\times\frac{81}{94}}{20}}} = 0{,}19$$

O *p-value* correspondente desta estatística de teste é dado por:

$$p = \mathrm{P}\left[Z_C \leq 0{,}19\right] = \Phi(0{,}19) = 0{,}5753$$

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência significativa para afirmar que a proporção de alunos do 8.º ano com dificuldades no cálculo da média é superior à proporção de alunos do 7.º ano com a mesma dificuldade.

Procedendo de modo análogo para os restantes conteúdos de Estatística e para as restantes escolas, facilmente obtive os valores da tabela seguinte:

Tabela 16: Valores do teste de hipóteses sobre a diferença de proporções e respectivos *p-values*

	EB 2, 3 Navegador Rodrigues Soromenho	**EB 2,3 Cardoso Lopes**	**EB 2,3 Mestre Domingos Saraiva**	**EB 2,3/S Michel Giacometti**	**Escola Secundária de Gil Vicente**
Média	$Z_C = 0{,}19$ $p = 0{,}5753$	$Z_C = 0{,}19$ $p = 0{,}5753$	$Z_C = 1{,}01$ $p = 0{,}8438$	$Z_C = 0{,}35$ $p = 0{,}6368$	$Z_C = -0{,}28$ $p = 0{,}3897$
Mediana	$Z_C = -0{,}18$ $p = 0{,}4286$	$Z_C = 0{,}44$ $p = 0{,}67$	$Z_C = 0{,}05$ $p = 0{,}5199$	$Z_C = 0{,}48$ $p = 0{,}6844$	$Z_C = -0{,}23$ $p = 0{,}409$
Moda	$Z_C = 0{,}49$ $p = 0{,}6879$	$Z_C = 0{,}20$ $p = 0{,}5793$	$Z_C = 0{,}22$ $p = 0{,}5871$	$Z_C = 0{,}29$ $p = 0{,}6141$	$Z_C = -0{,}28$ $p = 0{,}3897$
Histograma	$Z_C = 0{,}51$ $p = 0{,}6950$	$Z_C = 0{,}09$ $p = 0{,}5359$	$Z_C = 0{,}97$ $p = 0{,}8340$	$Z_C = 0{,}49$ $p = 0{,}6879$	$Z_C = 2{,}19$ $p = 0{,}9857$
Gráfico de Barras	$Z_C = 0{,}51$ $p = 0{,}6950$	$Z_C = 0{,}09$ $p = 0{,}5359$	$Z_C = 1{,}09$ $p = 0{,}8621$	$Z_C = 1{,}29$ $p = 0{,}9015$	$Z_C = -0{,}28$ $p = 0{,}3897$
Gráfico Circular	$Z_C = 0{,}51$ $p = 0{,}6950$	$Z_C = 0{,}98$ $p = 0{,}8365$	$Z_C = 0{,}70$ $p = 0{,}7580$	$Z_C = 0{,}36$ $p = 0{,}6406$	$Z_C = 1{,}09$ $p = 0{,}8621$
Pictograma	$Z_C = -0{,}44$ $p = 0{,}33$	$Z_C = 0{,}09$ $p = 0{,}5359$	$Z_C = 0{,}61$ $p = 0{,}7291$	$Z_C = 0{,}10$ $p = 0{,}5398$	$Z_C = 1{,}99$ $p = 0{,}9767$
Tabela de Frequências	$Z_C = -0{,}44$ $p = 0{,}33$	$Z_C = 0{,}03$ $p = 0{,}5120$	$Z_C = 0{,}02$ $p = 0{,}5080$	$Z_C = 2{,}10$ $p = 0{,}9821$	$Z_C = -0{,}28$ $p = 0{,}3897$
Frequência Absoluta	$Z_C = -0{,}44$ $p = 0{,}33$	$Z_C = 0{,}86$ $p = 0{,}8051$	$Z_C = -0{,}10$ $p = 0{,}4602$	$Z_C = 0{,}22$ $p = 0{,}5871$	$Z_C = 0{,}05$ $p = 0{,}5199$
Frequência Relativa	$Z_C = 0{,}51$ $p = 0{,}6950$	$Z_C = 1{,}48$ $p = 0{,}9306$	$Z_C = -0{,}10$ $p = 0{,}4602$	$Z_C = 1{,}04$ $p = 0{,}8508$	$Z_C = 0{,}36$ $p = 0{,}6406$

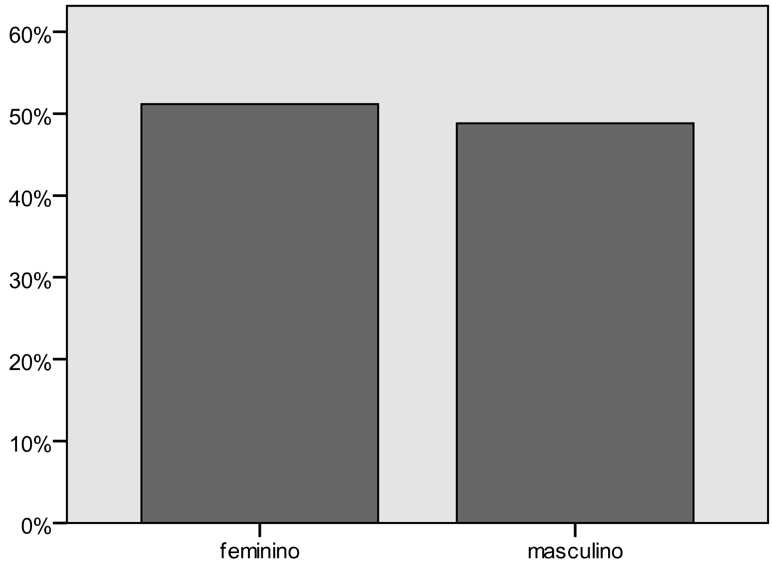
A tabela anterior apresenta *p-values* bastante elevados, o que significa que não há evidência significativa para afirmar que a proporção de alunos do 8.º ano com dificuldades nos conteúdos de Estatística é superior à de alunos que referem os mesmos conteúdos, mas que estão a frequentar o 7.º ano.

Será que se o inquérito tivesse sido alargado a mais alunos e a mais escolas chegaria aos mesmos resultados?

5.2.4. 9.º Ano do Ensino Básico

Da análise da amostra recolhida, concluí que 51,2% dos inquiridos são do género feminino, enquanto que 48,8% são do género masculino (Gráfico 41).

Gráfico 41: Distribuição por género



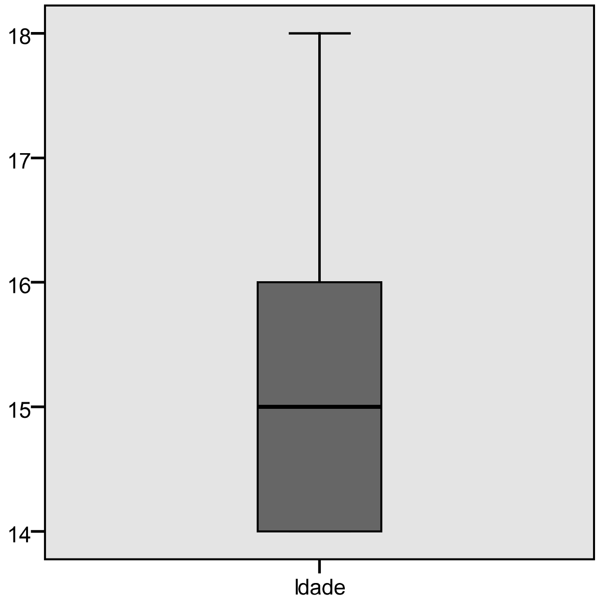
A partir da observação da Tabela 17 e do Gráfico 42, constatei que dos 258 alunos inquiridos, o mais novo tem 14 anos e o mais velho tem 18 anos, o que significa que há uma grande amplitude de variação das idades dos alunos que frequentam o 9.º ano; a moda e a mediana das idades são de 15 anos; a média das idades é de 15,03 anos e a dispersão das observações em relação à média – desvio padrão – é de 1,02 anos; 75% dos inquiridos têm idade compreendida entre os 14 e os 16 anos, como se pode depreender dos valores dos quartis e do mínimo da amostra; os restantes 25% dos inquiridos, que correspondem aos alunos mais velhos, têm idade compreendida entre os 16 e os 18 anos, como se pode depreender do valor do 3.º quartil e do máximo da amostra.

Analisando os valores anteriores, concluí que, de facto, faz sentido que a moda das idades seja de 15 anos, pois se um aluno fizer um percurso normal na escola quando chegar ao 9.º ano tem 14 ou 15 anos, consoante a data do seu aniversário. Portanto, também faz sentido que a idade mínima observada seja de 14 anos. No entanto, é do conhecimento geral que nem todos os alunos apresentam um percurso normal na escola, daí haver muitos inquiridos com idade superior à moda (este facto é facilmente confirmado através do Gráfico 42).

Tabela 17: Estatísticas da variável idade

N	258
Média	15,03
Mediana	15
Moda	15
Variância	1,05
Desvio Padrão	1,02
Mínimo	14
Máximo	18
1.º Quartil	14
2.º Quartil	15
3.º Quartil	16

Gráfico 42: *Box-plot* das idades dos inquiridos



No que diz respeito à opinião que os alunos do 9.º ano têm em relação à disciplina de Matemática, a maioria dos inquiridos (36,4%) afirma gostar da disciplina (Gráfico 43). No entanto, a maioria dos inquiridos (43%) auto-avalia de Suficiente o seu rendimento a esta disciplina, seguido de Insuficiente com 30,2% (Gráfico 44). De facto, ao consultar os Gráficos 45 e 46, verifiquei que, no 1.º Período, a classificação que ocorreu com maior frequência foi o 3 (com 43,6% de incidência), seguida da classificação 2 (com 37,7%); no 2.º Período, as classificações baixaram, passando a ser a classificação 2 a ocorrer com maior frequência (com 42,1% de incidência), seguida da classificação 3 (com 41,3%).

Gráfico 43: Distribuição dos inquiridos segundo a sua opinião em relação à disciplina de Matemática

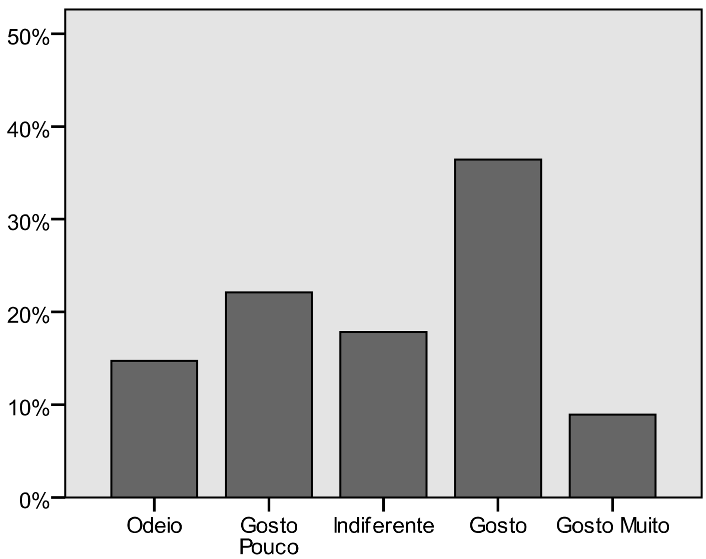


Gráfico 44: Distribuição dos inquiridos segundo a sua própria avaliação do seu rendimento a Matemática

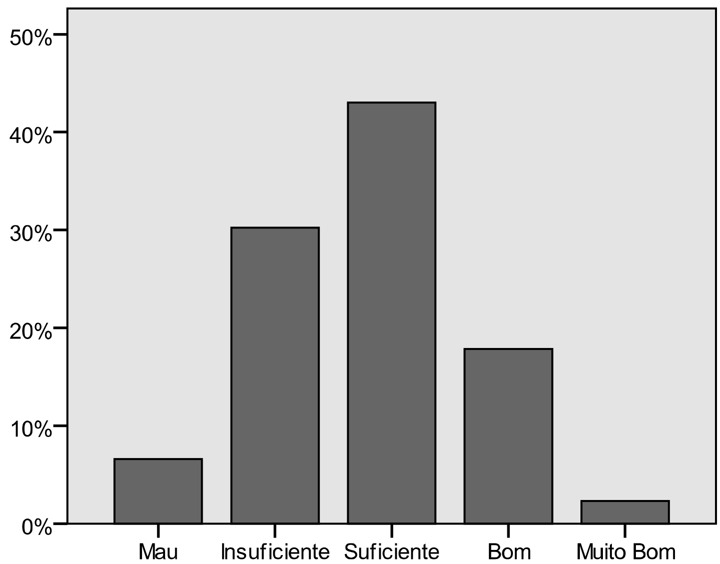


Gráfico 45: Classificação que os alunos obtiveram no 1.º Período

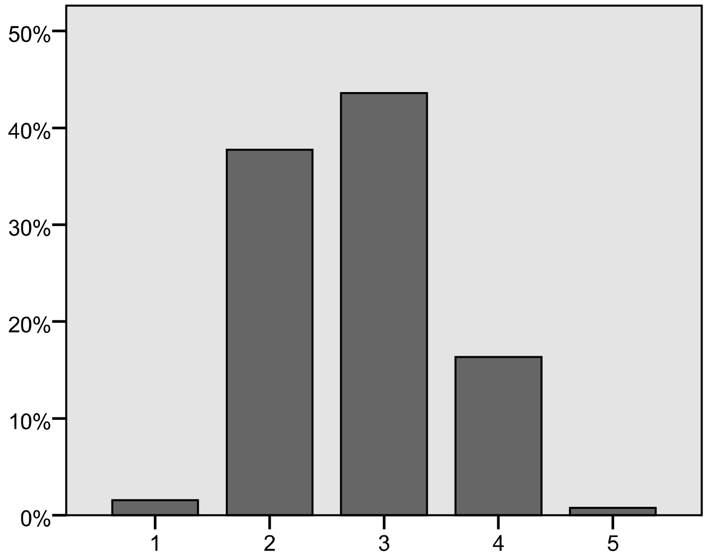
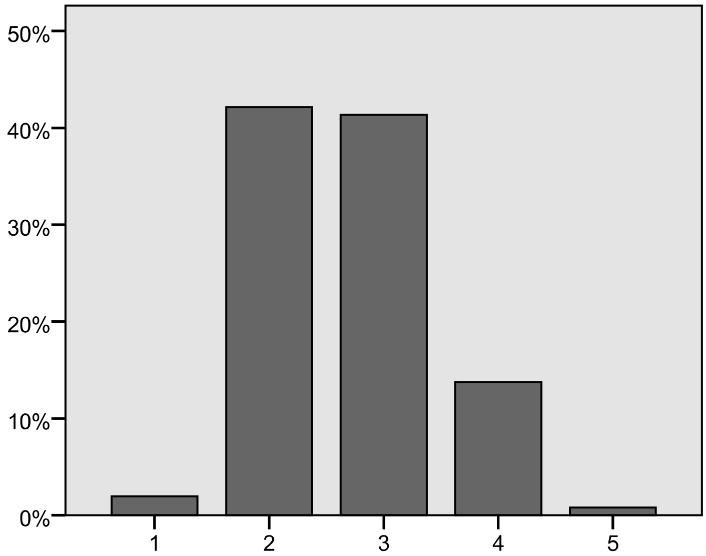


Gráfico 46: Classificação que os alunos obtiveram no 2.º Período



Ao recorrer ao *software* R para saber o tipo de associação que há entre a opinião que os alunos do 9.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina, concluí que o valor do coeficiente de correlação de Spearman é 0,587 e que o valor do *p-value* associado ao teste $H_0 : \rho_S = 0$ é inferior a $2{,}2 \times 10^{-16}$. Face a estes valores, concluí que há uma associação positiva moderada entre a opinião que os alunos do 9.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina. Sendo, portanto, de admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa desta disciplina.

De seguida, vou analisar o grau de associação que há entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º e 2.º Períodos, a esta disciplina.

Recorrendo novamente ao *software* R, obtive a seguinte tabela:

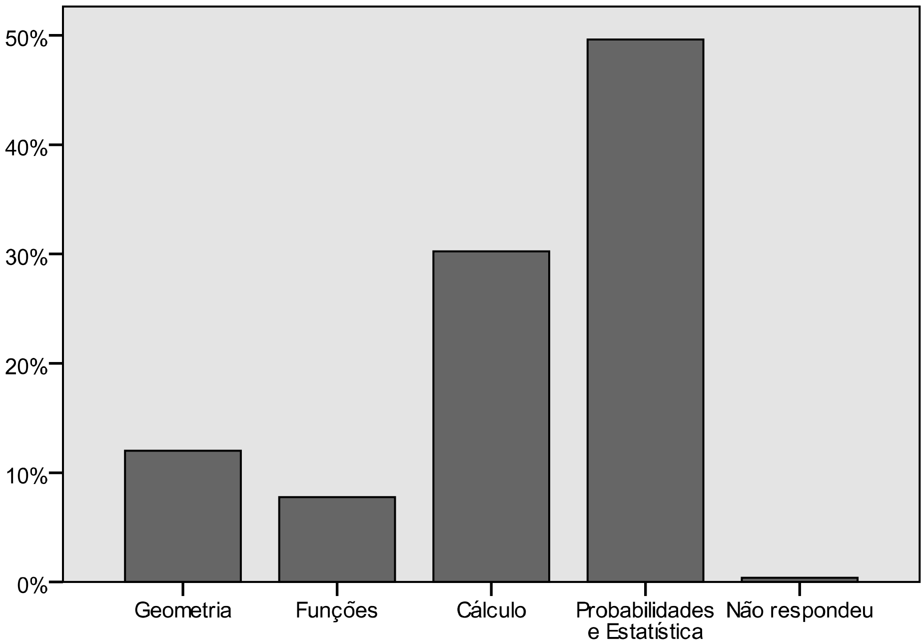
Tabela 18: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

	Como avalia o seu rendimento a Matemática?	
Qual a classificação que obteve no 1.º Período?	$(r_S)_C = 0{,}700$	$p < 2{,}2 \times 10^{-16}$
Qual a classificação que obteve no 2.º Período?	$(r_S)_C = 0{,}686$	$p < 2{,}2 \times 10^{-16}$

Pela leitura dos valores da Tabela 18, concluí que há uma associação positiva forte entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º Período, a esta disciplina. No entanto, esta associação torna-se ligeiramente mais fraca quando se relaciona a classificação que os alunos obtiveram no 2.º Período com a forma como avaliam o seu rendimento.

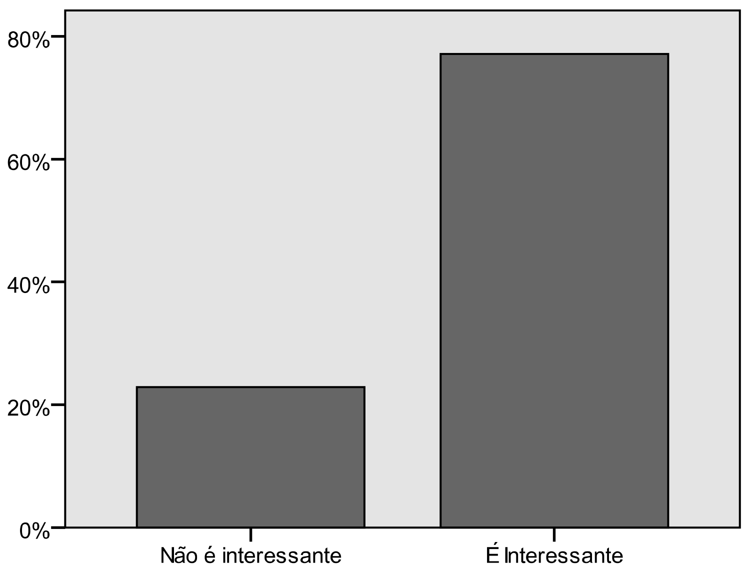
No que se refere aos temas abordados ao longo do ano lectivo na disciplina de Matemática, a maioria dos alunos (49,6%) afirma preferir a temática de Probabilidades e Estatística, seguida da temática de Cálculo com 30,2% (Gráfico 47).

Gráfico 47: Distribuição dos inquiridos de acordo com o tema preferido



Em relação à temática de Probabilidades e Estatística, tal como seria de esperar, a maioria dos inquiridos (77,1%) considera o tema interessante. No entanto, os restantes alunos (22,9%) consideram que o tema não é interessante (Gráfico 48).

Gráfico 48: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Probabilidades e Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação ao tema de Probabilidades e Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 9.º ano têm em relação ao tema de Probabilidades e Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 19: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Probabilidades e Estatística e a escola que frequentam

			Opinião em relação ao tema de Probabilidades e Estatística		
			Não é interessante	É Interessante	Total
Nome da Escola	EB 2, 3 Navegador Rodrigues Soromenho	Valor Observado	29	60	89
		Valor Esperado	20,4	68,6	89,0
	Escola Secundária de Gil Vicente	Valor Observado	20	81	101
		Valor Esperado	23,1	77,9	101,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	9	43	52
		Valor Esperado	11,9	40,1	52,0
	EB 2,3/S Michel Giacometti	Valor Observado	1	15	16
		Valor Esperado	3,7	12,3	16,0
	Total	Valor Observado	59	199	258
		Valor Esperado	59,0	199,0	258,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação ao tema de Probabilidades e Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 19, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 8,68443 e que o *p-value* correspondente a esta estatística de teste é 0,03379.

Com base no valor do *p-value*, concluí que rejeito H_0 ao nível de significância de 5% e não rejeito H_0 ao nível de significância de 1%. Portanto, ao nível de significância de 5%, há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação ao tema de Probabilidades e Estatística está associada à escola que frequentam. Mas, ao nível de significância de 1%, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação ao tema de Probabilidades e Estatística depende da escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação ao tema de Probabilidades e Estatística" ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas um dos valores esperados é inferior a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e dos dados da Tabela 19, obtive a seguinte tabela:

Tabela 20: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	8,71823	3	0,03328
Likelihood Ratio	9,16792	3	0,02714

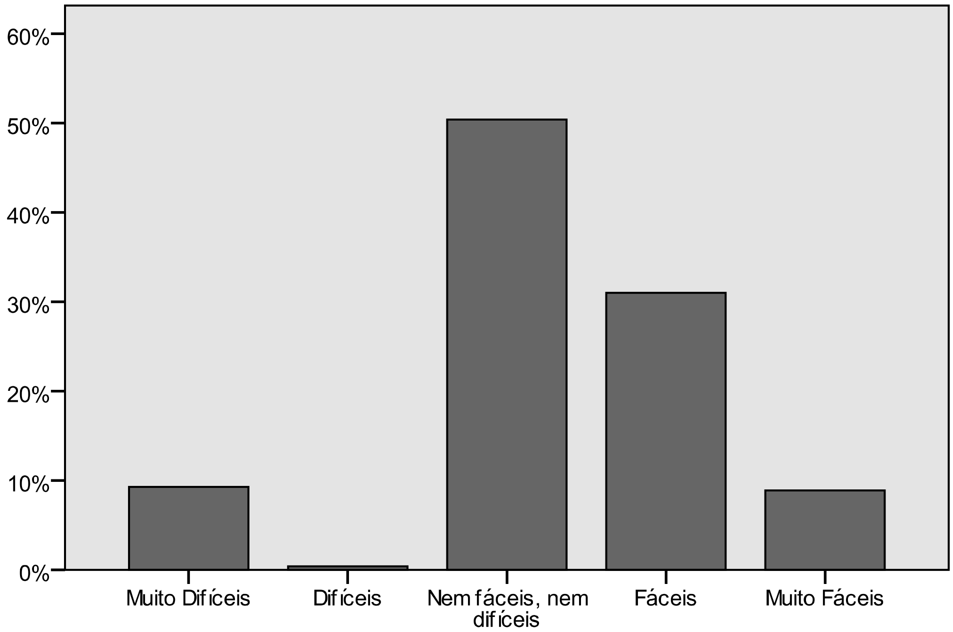
Através da análise da Tabela 20, concluí que o valor da estatística de teste, X^2, é 8,71823 e que o *p-value* correspondente a esta estatística de teste é 0,03328. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 9,16792 e que o valor do *p-value* correspondente a esta estatística de teste é 0,02714.

Com base nos valores dos *p-values*, concluí que rejeito H_0 ao nível de significância de 5% e não rejeito H_0 ao nível de significância de 1%. Portanto, ao nível de significância de 5%, há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação ao tema de Probabilidades e Estatística está associada à escola que frequentam. Mas, ao nível de significância de 1%, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação ao tema de Probabilidades e Estatística depende da escola que frequentam.

Apesar destes resultados serem um pouco inconclusivos, os três testes dão origem às mesmas conclusões, o que é bom. No futuro, este estudo deveria ser alargado a mais escolas e a mais alunos de forma a obter-se resultados conclusivos.

No que diz respeito à opinião que os alunos do 9.º ano têm em relação aos conteúdos de Probabilidades e Estatística, a maioria dos inquiridos (50,4%) considera que os conteúdos nem são fáceis nem são difíceis. No entanto, 31% dos inquiridos avaliam--nos como sendo fáceis (Gráfico 49).

Gráfico 49: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Probabilidades e Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 9.º ano têm em relação aos conteúdos de Probabilidades e Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 21: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Probabilidades e Estatística e a escola que frequentam

			Opinião em relação aos conteúdos de Probabilidades e Estatística			Total
			Fáceis e Muito Fáceis	Nem fáceis nem difíceis	Difíceis e Muito Difíceis	
Nome da Escola	EB 2,3 Navegador Rodrigues Soromenho	Valor Observado	25	51	9	85
		Valor Esperado	33,4	43,2	8,4	85,0
	Escola Secundária de Gil Vicente	Valor Observado	40	54	7	101
		Valor Esperado	39,7	51,3	10,0	101,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	24	18	8	50
		Valor Esperado	19,6	25,4	5,0	50,0
	EB 2,3/S Michel Giacometti	Valor Observado	10	5	1	16
		Valor Esperado	6,3	8,1	1,6	16,0
Total		Valor Observado	99	128	25	252
		Valor Esperado	99,0	128,0	25,0	252,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação aos conteúdos de Probabilidades e Estatística" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 21, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 6,37688 e que o *p-value* correspondente a esta estatística de teste é 0,09465.

Com base no valor do *p-value*, concluí que, aos níveis de significância de 1 e 5%, não rejeito H_0. Portanto, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação aos conteúdos de Probabilidades e Estatística depende da escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação aos conteúdos de Probabilidades e Estatística" ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas 16,7% dos valores esperados são inferior a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e dos dados da Tabela 21, obtive a seguinte tabela:

Tabela 22: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

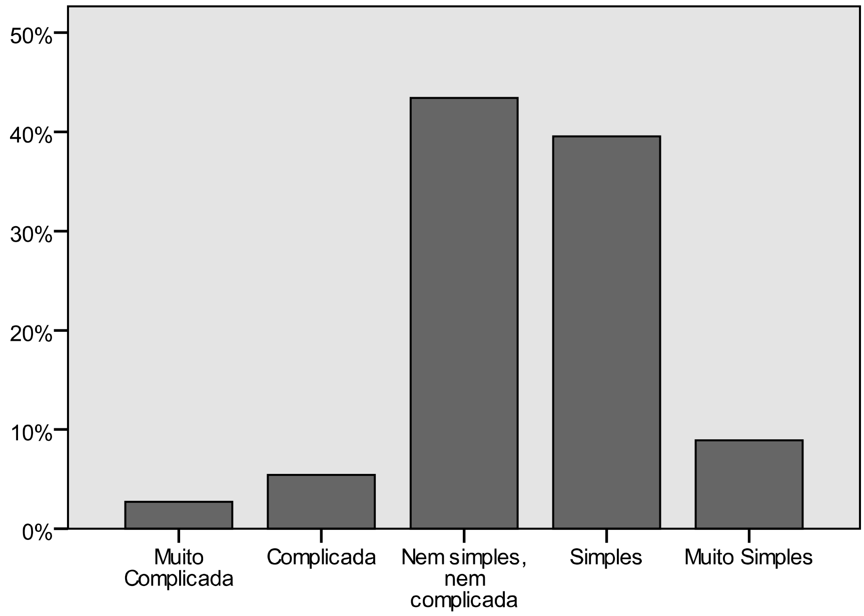
Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	13,21945	6	0,03968
Likelihood Ratio	13,22533	6	0,03959

Através da análise da Tabela 22, concluí que o valor da estatística de teste, X^2, é 13,21945 e que o *p-value* correspondente a esta estatística de teste é 0,03968. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 13,22533 que o valor do *p-value* correspondente a esta estatística de teste é 0,03959.

Com base nos valores dos *p-values*, concluí que rejeito H_0 ao nível de significância de 5% e não rejeito H_0 ao nível de significância de 1%. Portanto, ao nível de significância de 5%, há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação aos conteúdos de Probabilidades e Estatística está associada à escola que frequentam. Mas, ao nível de significância de 1%, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação aos conteúdos de Probabilidades e Estatística depende da escola que frequentam. Uma vez que estes resultados são pouco conclusivos, no futuro, este estudo deveria ser alargado a mais escolas e a mais alunos de forma a obter-se a um resultado conclusivo.

No que concerne à forma como os conteúdos de Probabilidades e Estatística foram abordados nas aulas, a maioria dos inquiridos (43,4%) considera que a forma utilizada pelos professores nem foi simples nem foi complicada, seguindo-se a forma simples com 39,5% de respostas (Gráfico 50).

Gráfico 50: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados



Para saber se há associação entre a escola e a opinião que os alunos têm em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados, vou realizar um teste de independência para testar:

> *H_0: Há independência entre a opinião que os alunos do 9.º ano têm em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados e a escola que frequentam.*

Da amostra recolhida, obtive a seguinte tabela:

Tabela 23: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados e a escola que frequentam

			Opinião em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados			Total
			Simples e Muito Simples	Nem simples nem complicada	Complicada e Muito Complicada	
Nome da Escola	EB 2,3 Navegador Rodrigues Soromenho	Valor Observado	34	45	10	89
		Valor Esperado	43,1	38,6	7,2	89,0
	Escola Secundária de Gil Vicente	Valor Observado	52	41	8	101
		Valor Esperado	48,9	43,8	8,2	101,0
	EB 2,3 Mestre Domingos Saraiva	Valor Observado	31	19	2	52
		Valor Esperado	25,2	22,6	4,2	52,0
	EB 2,3/S Michel Giacometti	Valor Observado	8	7	1	16
		Valor Esperado	7,8	6,9	1,3	16,0
Total		Valor Observado	125	112	21	258
		Valor Esperado	125,0	112,0	21,0	258,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 23, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 7,36608 e que o *p-value* correspondente a esta estatística de teste é 0,06110.

Com base no valor do *p-value*, concluí que não rejeito H_0 aos níveis de significância de 1 e 5%. Portanto, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados está associada à escola que frequentam.

Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados” ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois apenas 16,7% dos valores esperados são inferior a cinco, ou seja, estão satisfeitas as condições de Cochran.

Com o auxílio do *SPSS* e dos dados da Tabela 23, obtive a seguinte tabela:

Tabela 24: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	7,56834	6	0,27146
Likelihood Ratio	7,79066	6	0,25384

Através da análise da Tabela 24, concluí que o valor da estatística de teste, X^2, é 7,56834 e que o *p-value* correspondente a esta estatística de teste é 0,27146. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 7,79066 que o valor do *p-value* correspondente a esta estatística de teste é 0,25384.

Com base nos valores dos *p-values*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência para afirmar que a opinião que os alunos do 9.º ano têm em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados está associada à escola que frequentam.

De seguida, vou analisar, por escola, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Probabilidades e Estatística e a forma como os conteúdos foram leccionados; bem como, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Estatística e a forma como esses conteúdos foram leccionados.

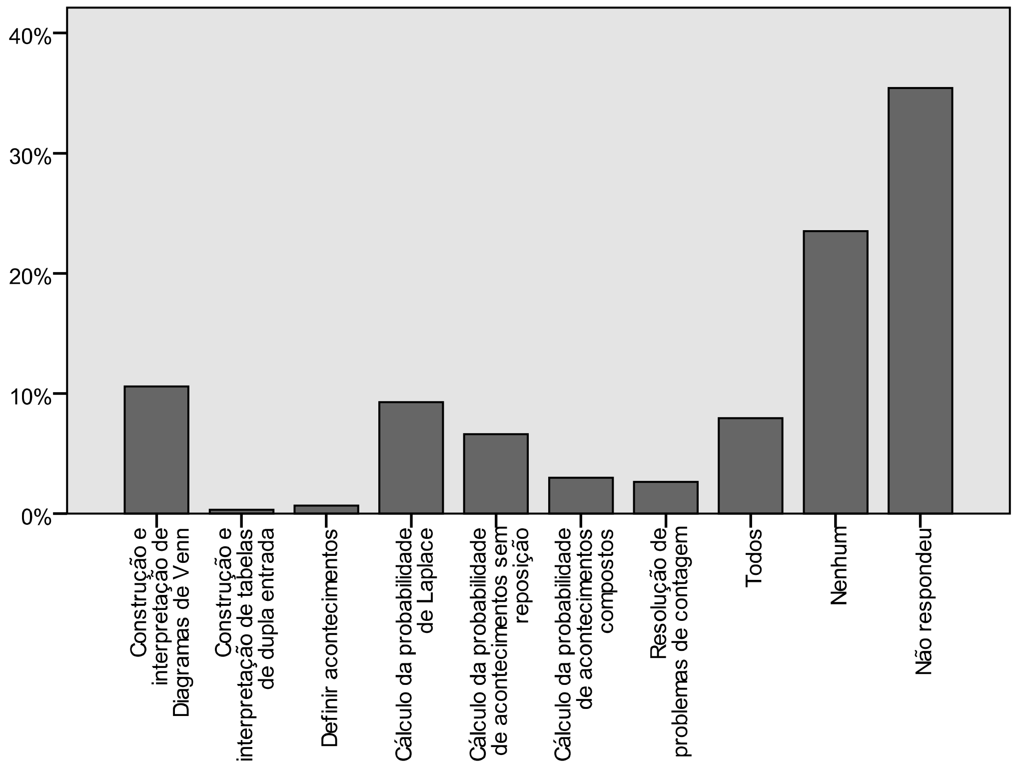
Tabela 25: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

		Opinião dos alunos em relação à forma como os conteúdos de Probabilidades e Estatística foram abordados	
EB 2, 3 Navegador Rodrigues Soromenho	Opinião dos alunos em relação ao tema de Probabilidades e Estatística	$(r_S)_C = 0{,}342$	$p = 1{,}051 \times 10^{-3}$
	Opinião dos alunos em relação aos conteúdos de Probabilidades e Estatística	$(r_S)_C = 0{,}403$	$p = 9{,}08 \times 10^{-5}$
EB 2,3 Mestre Domingos Saraiva	Opinião dos alunos em relação ao tema de Probabilidades e Estatística	$(r_S)_C = 0{,}067$	$p = 0{,}637$
	Opinião dos alunos em relação aos conteúdos de Probabilidades e Estatística	$(r_S)_C = 0{,}505$	$p = 1{,}353 \times 10^{-4}$
EB 2,3/S Michel Giacometti	Opinião dos alunos em relação ao tema de Probabilidades e Estatística	$(r_S)_C = 0{,}220$	$p = 0{,}413$
	Opinião dos alunos em relação aos conteúdos de Probabilidades e Estatística	$(r_S)_C = 0{,}324$	$p = 0{,}221$
Escola Secundária de Gil Vicente	Opinião dos alunos em relação ao tema de Probabilidades e Estatística	$(r_S)_C = 0{,}300$	$p = 2{,}303 \times 10^{-3}$
	Opinião dos alunos em relação aos conteúdos de Probabilidades e Estatística	$(r_S)_C = 0{,}483$	$p = 3{,}087 \times 10^{-7}$

Pela análise da Tabela 25, concluí que na Escola 2,3 Navegador Rodrigues Soromenho e na Escola Secundária Gil Vicente há associação positiva e moderada entre a opinião que os alunos têm em relação ao tema de Probabilidades e Estatística e a forma como os conteúdos foram abordados. No entanto, na Escola 2,3 Mestre Domingos Saraiva e na Escola 2,3/S Michel Giacometti não há associação entre as variáveis em questão. Concluí ainda que em todas as escolas, excepto na Escola 2,3/S Michel Giacometti, há associação positiva e moderada entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Estatística e a forma como os mesmos foram abordados nas aulas.

Finalmente, no que diz respeito aos conteúdos de Probabilidades e Estatística, grande parte dos alunos afirma ter sentido dificuldades em pelo menos um dos conteúdos. Pela leitura do Gráfico 51, constatei que 10,6% dos inquiridos admitem ter sentido dificuldades na construção e interpretação de Diagramas de *Venn*; 9,3% dos inquiridos referem ter sentido dificuldades no cálculo da probabilidade de *Laplace*; o cálculo da probabilidade de acontecimentos sem reposição é também assinalado por 6,6% dos inquiridos. Para além destes valores, é importante realçar que 23,5% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos de Probabilidades e Estatística e que 7,9% dos inquiridos revelam ter sentido dificuldades em todos os conteúdos. Na minha opinião, a comunidade escolar deve ter em atenção estes últimos alunos, pois poderão estar a necessitar de algum tipo de apoio ou acompanhamento. É também importante destacar que, aproximadamente, 35% dos inquiridos não responderam a esta questão. Esta situação poderá ter estado relacionada com o facto de alguns alunos já não se lembrarem dos conteúdos de Probabilidades e Estatística ou ainda com o facto de se tratar de uma questão de resposta aberta.

Gráfico 51: Conteúdos de Probabilidades e Estatística em que os alunos sentiram dificuldades



5.2.5. Comparação entre a Estatística do 7.º e 8.º Anos e as Probabilidades e Estatística do 9.º Ano

Supõe-se que, de um modo geral, os alunos preferem o tema de Estatística, que é leccionado no 7.º e 8.º anos, ao tema de Probabilidades e Estatística, que é leccionado no 9.º ano.

Para poder testar a validade desta suposição e uma vez que se tratam de amostras independentes, vou realizar, para cada escola, um teste de hipóteses sobre a diferença de proporções.

EB 2,3 Navegador Rodrigues Soromenho

Seja X_1 a variável aleatória que representa o número de alunos do 7.º e 8.º anos que consideram o tema de Estatística interessante.

Seja X_2 a variável aleatória que representa o número de alunos do 9.º ano que consideram o tema de Probabilidades e Estatística interessante.

Seja p_1 a proporção de alunos do 7.º e 8.º anos que consideram o tema de Estatística interessante.

Seja p_2 a proporção de alunos do 9.º ano que consideram o tema de Probabilidades e Estatística interessante.

Pretende-se testar o seguinte:

$$H_0 : p_1 = p_2 \qquad vs. \qquad H_1 : p_1 > p_2$$

Na escola EB 2,3 Navegador Rodrigues Soromenho foram inquiridos 38 alunos do 7.º e 8.º anos, dos quais 27 consideram o tema de Estatística interessante e foram inquiridos 89 alunos do 9.º ano, dos quais 60 consideram o tema de Probabilidades e Estatística interessante.

Assim, $\hat{p}_1 = \dfrac{27}{38}$; $\hat{p}_2 = \dfrac{60}{89}$; $\overline{p} = \dfrac{87}{127}$; $\overline{q} = \dfrac{40}{127}$

$$Z_C(obs.) = \frac{\left|\frac{27}{38} - \frac{60}{89}\right| - \frac{1}{2}\left(\frac{1}{38} + \frac{1}{89}\right)}{\sqrt{\frac{\frac{87}{127} \times \frac{40}{127}}{38} + \frac{\frac{87}{127} \times \frac{40}{127}}{89}}} = 0{,}20$$

O *p-value* correspondente desta estatística de teste é dado por:

$$p = \mathrm{P}[Z_C \geq 0{,}2] = 1 - \Phi(0{,}2) = 1 - 0{,}5793 = 0{,}4207$$

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência significativa para afirmar que a proporção de alunos que prefere o tema de Estatística do 7.º e 8.º anos é superior à proporção de alunos que prefere o tema de Probabilidades e Estatística do 9.º ano.

Procedendo de modo análogo para as restantes escolas, obtive facilmente os valores da tabela seguinte:

Tabela 26: Valores do teste de hipóteses sobre a diferença de proporções e respectivos *p-values*

	EB 2, 3 Navegador Rodrigues Soromenho	**EB 2,3 Mestre Domingos Saraiva**	**EB 2,3/S Michel Giacometti**	**Escola Secundária de Gil Vicente**
Estatística do 7.º e 8.º anos *vs.* Probabilidades e Estatística do 9.º ano	$Z_C = 0{,}20$ $p = 0{,}4207$	$Z_C = 1{,}04$ $p = 0{,}1492$	$Z_C = 1{,}51$ $p = 0{,}0655$	$Z_C = 0{,}08$ $p = 0{,}4681$

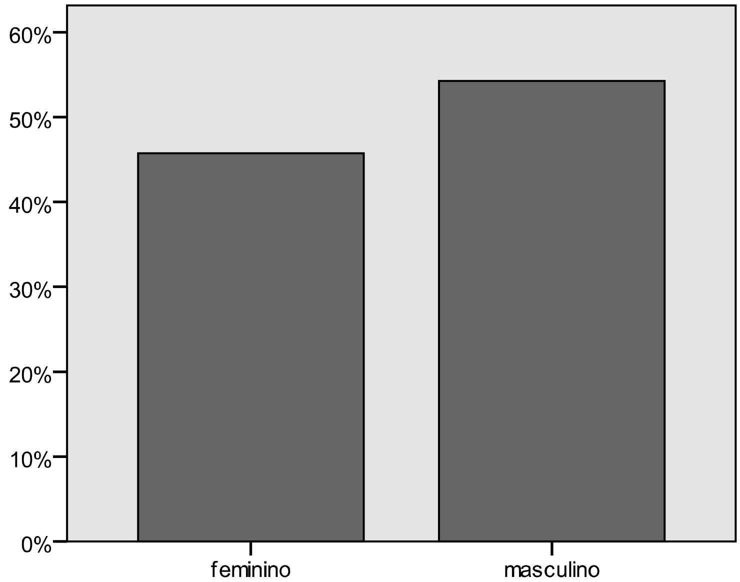
Com base nos valores dos *p-values* da tabela anterior, concluí que, aos níveis de significância de 1 e 5%, não rejeito H_0. Portanto, não há evidência significativa para afirmar que a proporção de alunos que prefere o tema de Estatística do 7.º e 8.º anos é superior à proporção de alunos que prefere o tema de Probabilidades e Estatística do 9.º ano.

Mas será que se o inquérito tivesse sido alargado a mais alunos e a mais escolas, chegaria aos mesmos resultados?

5.2.6. 10.º Ano do Ensino Secundário

Da análise da amostra recolhida, concluí que 54,3% dos inquiridos são do género masculino, enquanto que 45,7% são do género feminino (Gráfico 52).

Gráfico 52: Distribuição por género



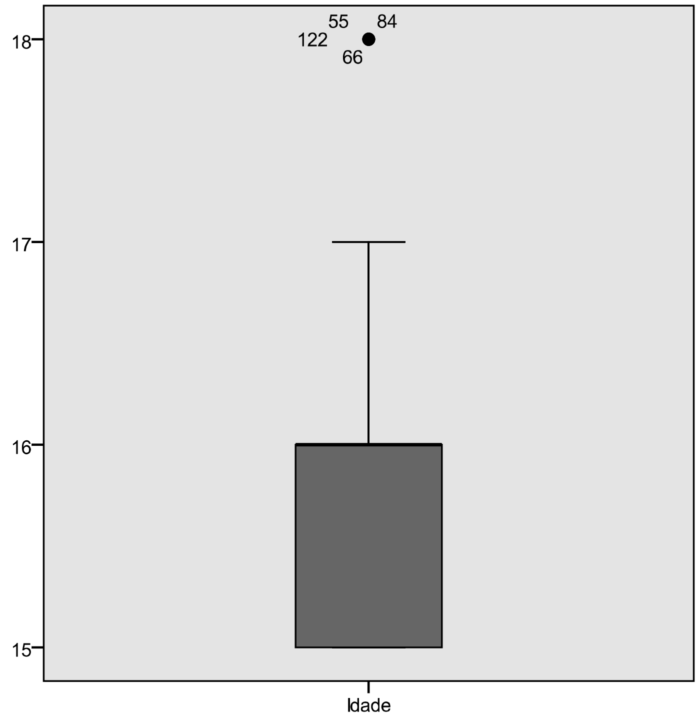
Pela análise da Tabela 27 e do Gráfico 53, concluí que dos 129 alunos inquiridos, o mais novo tem 15 anos e o mais velho tem 18 anos. A moda das idades é de 15 anos e a mediana das idades é de 16 anos. A média das idades dos inquiridos é de 15,77 anos e a dispersão das observações em relação à média – desvio padrão – é de 0,80 anos. Com base no Gráfico 53, verifiquei ainda a presença de quatro *outliers* moderados.

Analisando os valores anteriores, concluí que, de facto, faz sentido que a moda das idades seja de 15 anos, pois se um aluno fizer um percurso normal na escola quando chegar ao 10.º ano tem 15 ou 16 anos, consoante a data do seu aniversário. Portanto, também faz sentido que a idade mínima observada seja de 15 anos. No entanto, é do conhecimento geral que nem todos os alunos apresentam um percurso normal na escola, daí haver muitos inquiridos com idade superior à moda (este facto é facilmente confirmado através do Gráfico 53).

Tabela 27: Estatísticas da variável idade

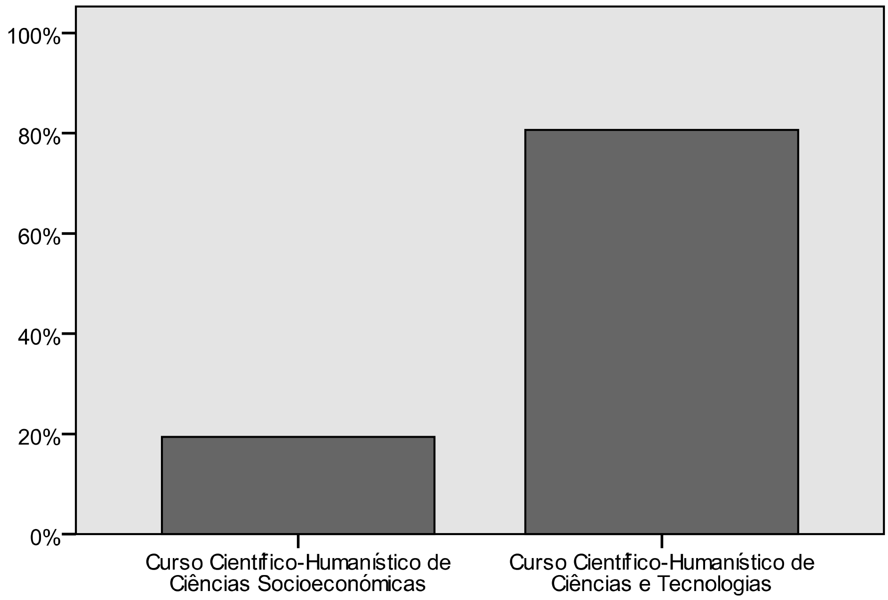
N	129
Média	15,77
Mediana	16
Moda	15
Variância	0,63
Desvio Padrão	0,80
Mínimo	15
Máximo	18
1.º Quartil	15
2.º Quartil	16
3.º Quartil	16

Gráfico 53: *Box-plot* das idades dos inquiridos



Em relação ao curso que frequentam, a grande maioria dos inquiridos (80,6%) frequenta o Curso Científico-Humanístico de Ciências e Tecnologias, enquanto que apenas 19,4% dos inquiridos frequentam o Curso Científico-Humanístico de Ciências Socioeconómicas (Gráfico 54).

Gráfico 54: Distribuição dos inquiridos de acordo com o curso que frequentam



No que diz respeito à opinião que os alunos do 10.º ano têm em relação à disciplina de Matemática, a maioria dos inquiridos (45%) afirma gostar de Matemática (Gráfico 55).

No entanto, aproximadamente, 41% dos alunos auto-avaliam de Suficiente o seu rendimento a esta disciplina e 31% dos inquiridos auto-avaliam-no de Insuficiente (Gráfico 56). De facto, ao consultar o Gráfico 57, concluí que o nível que mais se verificou no 1.º Período foi o 10 com 15,7%, seguido dos níveis 8 e 7 com 12,6% e 11,8%, respectivamente. No 2.º Período, o nível mais frequente foi o 9 com 15%, seguido do 8 e do 10 com 13,4% e 10,2%, respectivamente (Gráfico 58). Com base nos dados recolhidos, constatei que apesar da maioria dos alunos do 10.º ano gostar da disciplina de Matemática, eles não conseguem ter elevados rendimentos nesta disciplina.

Gráfico 55: Distribuição dos inquiridos segundo a sua opinião em relação à disciplina de Matemática

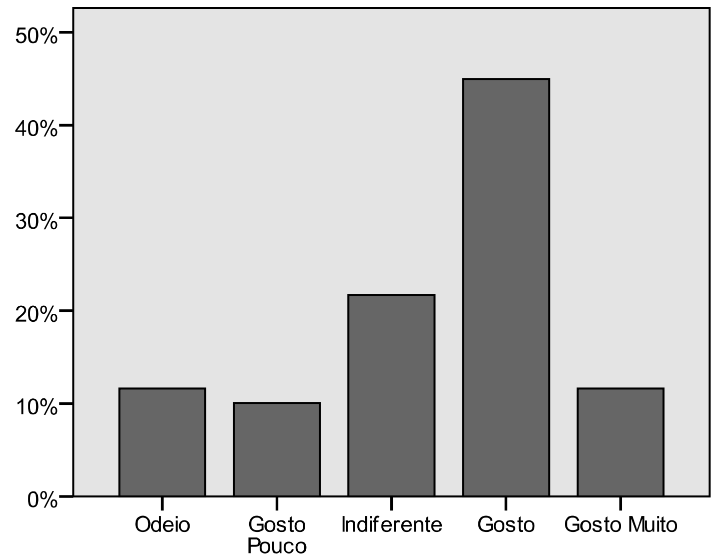


Gráfico 56: Distribuição dos inquiridos segundo a sua própria avaliação do seu rendimento a Matemática

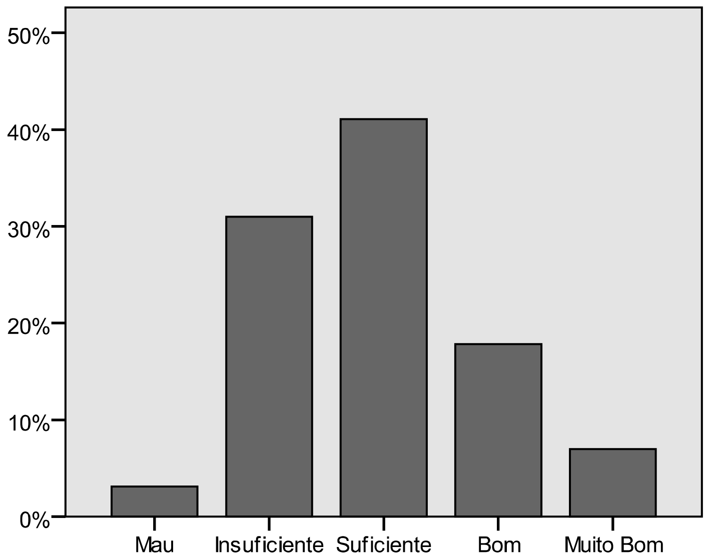


Gráfico 57: Classificação que os alunos obtiveram no 1.º Período

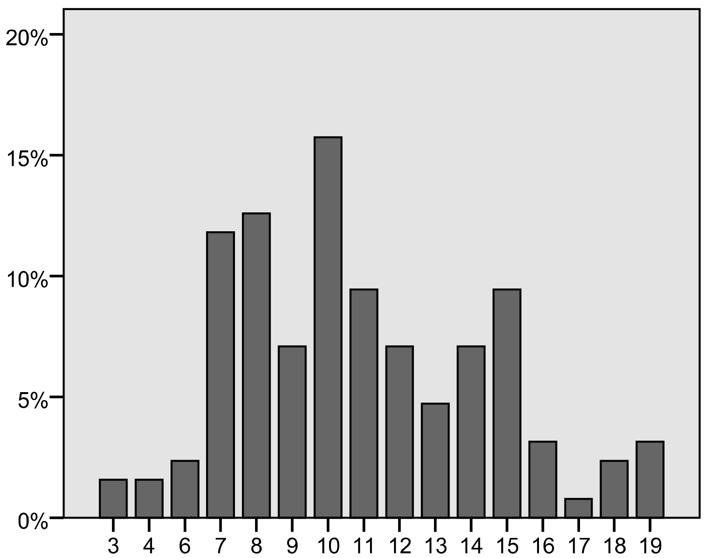
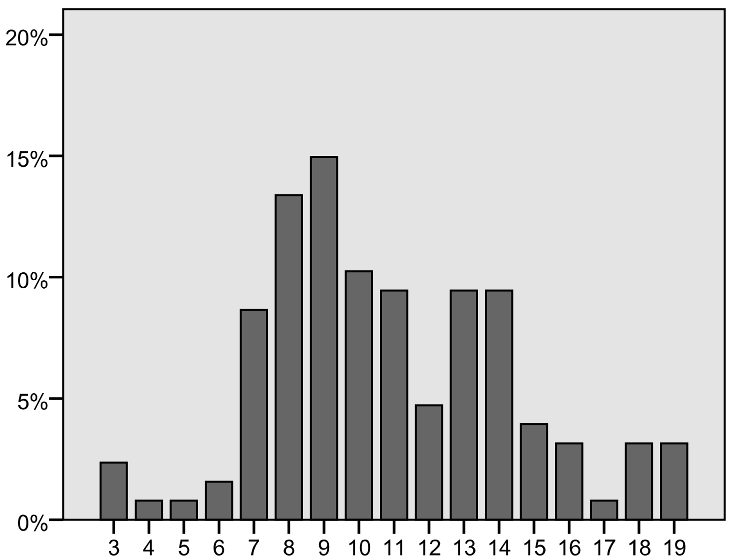


Gráfico 58: Classificação que os alunos obtiveram no 2.º Período



Ao recorrer ao *software* R para saber o tipo de associação que há entre a opinião que os alunos do 10.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina, concluí que o valor do coeficiente de correlação de Spearman é 0,462 e que o valor do *p-value* associado ao teste $H_0 : \rho_S = 0$ é $3{,}64 \times 10^{-8}$. Face a estes valores, concluí que há uma associação positiva moderada entre a opinião que os alunos do 10.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa desta disciplina.

De seguida, vou analisar o grau de associação que há entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º e 2.º Períodos, a esta disciplina.

Recorrendo novamente ao *software* R, obtive a seguinte tabela:

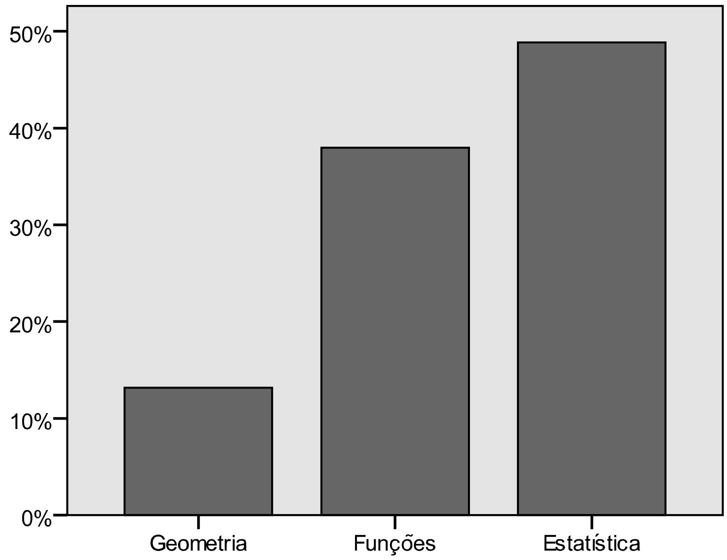
Tabela 28: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

	Como avalia o seu rendimento a Matemática?	
Qual a classificação que obteve no 1.º Período?	$(r_S)_C = 0{,}855$	$p < 2{,}2 \times 10^{-16}$
Qual a classificação que obteve no 2.º Período?	$(r_S)_C = 0{,}875$	$p < 2{,}2 \times 10^{-16}$

Pela análise dos valores da Tabela 28, concluí que há uma associação positiva forte entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º Período, a esta disciplina. Esta associação torna--se ainda mais forte quando se relaciona a classificação que os alunos obtiveram no 2.º Período com a forma como avaliam o seu rendimento. O que significa que ao longo do ano lectivo, os alunos vão conseguindo auto-avaliar-se correctamente no que diz respeito ao seu rendimento na disciplina de Matemática.

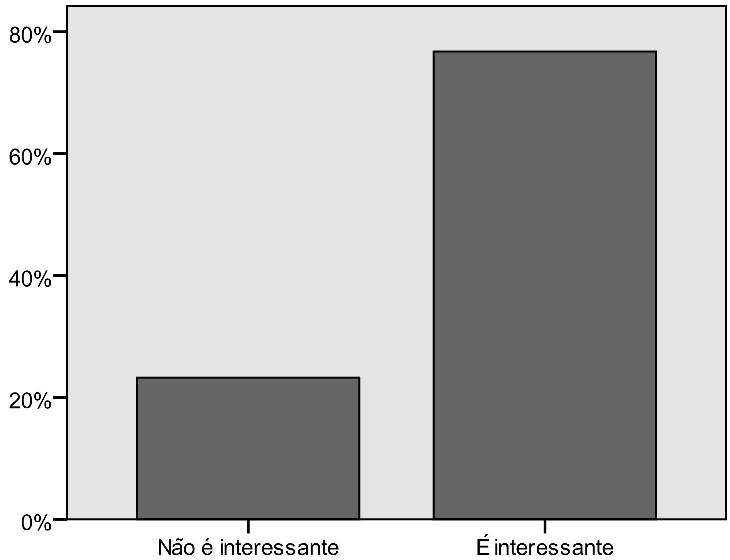
Relativamente aos temas leccionados durante o ano lectivo na disciplina de Matemática, a maioria dos alunos (aproximadamente 49%) afirma preferir a unidade temática de Estatística, seguida das temáticas de Funções e de Geometria com 38% e 13,2%, respectivamente (Gráfico 59).

Gráfico 59: Distribuição dos inquiridos de acordo com o tema preferido



Em relação à unidade temática de Estatística, a maioria dos inquiridos (76,7%) considera a temática interessante. O contrário não seria de esperar, uma vez que o tema de Estatística é considerado o tema de eleição da maioria dos inquiridos. Contudo, a percentagem de alunos que consideram o tema de Estatística interessante (76,7%) é bem superior à percentagem de alunos que o seleccionam como sendo o seu preferido (48,8%) (Gráfico 60).

Gráfico 60: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística



Para saber se há associação entre a escola e a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 29: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Estatística e a escola que frequentam

			Opinião em relação ao tema da Estatística		
			Não é interessante	É interessante	Total
Nome da Escola	Escola Secundária da Baixa da Banheira	Valor Observado	3	27	30
		Valor Esperado	7,0	23,0	30,0
	Escola Secundária Ferreira Dias	Valor Observado	10	13	23
		Valor Esperado	5,3	17,7	23,0
	Escola Secundária Fernão Mendes Pinto	Valor Observado	6	19	25
		Valor Esperado	5,8	19,2	25,0
	EB 2,3/S Michel Giacometti	Valor Observado	11	40	51
		Valor Esperado	11,9	39,1	51,0
	Total	Valor Observado	30	99	129
		Valor Esperado	30,0	99,0	129,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação ao tema de Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal--Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 29, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 8,24839 e que o *p-value* correspondente a esta estatística de teste é 0,04115.

Com base no valor do *p-value*, concluí que rejeito H_0 ao nível de significância de 5% e não rejeito H_0 ao nível de significância de 1%. Portanto, ao nível de significância de 5%, há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística está associada à escola que frequentam. Mas, ao nível de significância de 1%, não há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística depende da escola que frequentam.

Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação ao tema de Estatística” ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois nenhum dos valores esperados é inferior a cinco.

Com o auxílio do *SPSS* e dos dados da Tabela 29, obtive a seguinte tabela:

Tabela 30: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

Teste	**Valor da Estatística de Teste**	**Graus de Liberdade**	***p-value***
Qui-Quadrado	8,31283	3	0,03997
Likelihood Ratio	8,19254	3	0,04220

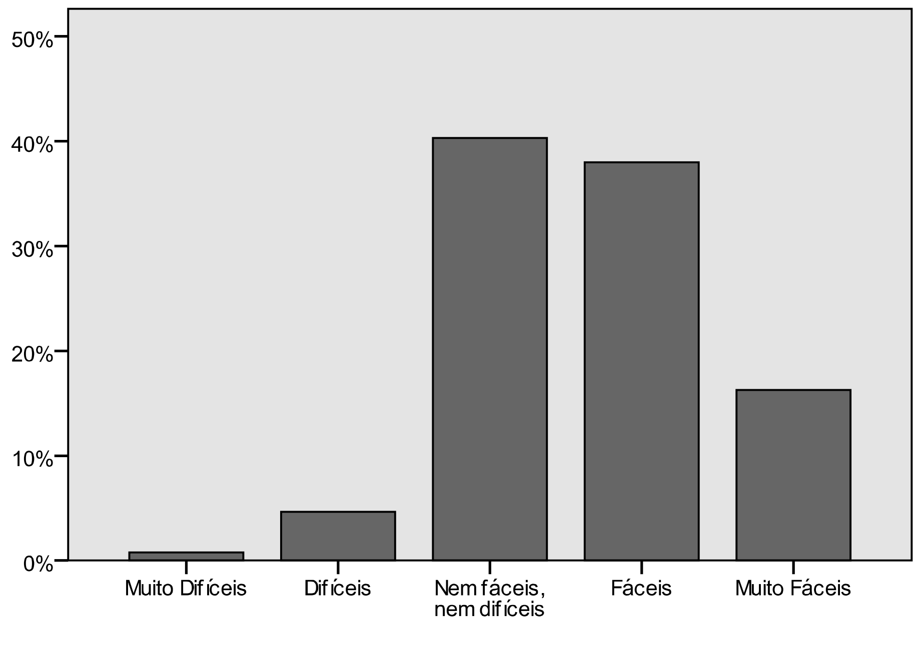
Através da análise da Tabela 30, concluí que o valor da estatística de teste, X^2, é 8,31283 e que o *p-value* correspondente a esta estatística de teste é 0,03997. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 8,19254 e que o valor do *p-value* correspondente a esta estatística de teste é 0,04220.

Com base nos valores dos *p-values*, concluí que rejeito H_0 ao nível de significância de 5% e não rejeito H_0 ao nível de significância de 1%. Portanto, ao nível de significância de 5%, há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística está associada à escola que frequentam. Mas, ao nível de significância de 1%, não há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação ao tema de Estatística depende da escola que frequentam.

Apesar destes resultados serem um pouco inconclusivos, os três testes dão origem às mesmas conclusões, o que é bom. No futuro, este estudo deveria ser alargado a mais escolas e a mais alunos de forma a obter-se resultados conclusivos.

No que concerne à opinião que os alunos têm em relação aos conteúdos de Estatística, a maioria dos inquiridos (40,3%) considera que os conteúdos nem são fáceis nem são difíceis, seguida da opinião de que são fáceis com 38% de respostas (Gráfico 61).

Gráfico 61: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística



Para saber se há associação entre a escola e a opinião que os alunos têm em relação aos conteúdos de Estatística, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 10.º ano têm em relação aos conteúdos de Estatística e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 31: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Estatística e a escola que frequentam

			Opinião em relação aos conteúdos de Estatística			
			Fáceis e Muito Fáceis	Nem fáceis nem difíceis	Difíceis e Muito Difíceis	Total
Nome da Escola	Escola Secundária da Baixa da Banheira	Valor Observado	15	14	1	30
		Valor Esperado	16,3	12,1	1,6	30,0
	Escola Secundária Ferreira Dias	Valor Observado	10	12	1	23
		Valor Esperado	12,5	9,3	1,2	23,0
	Escola Secundária Fernão Mendes Pinto	Valor Observado	15	9	1	25
		Valor Esperado	13,6	10,1	1,4	25,0
	EB 2,3/S Michel Giacometti	Valor Observado	30	17	4	51
		Valor Esperado	27,7	20,6	2,8	51,0
	Total	Valor Observado	70	52	7	129
		Valor Esperado	70,0	52,0	7,0	129,0

Uma vez que a variável “nome da escola” é do tipo nominal e a variável “opinião em relação aos conteúdos de Estatística” é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 31, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 1,49834 e que o *p-value* correspondente é 0,68265.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação aos conteúdos de Estatística está associada à escola que frequentam.

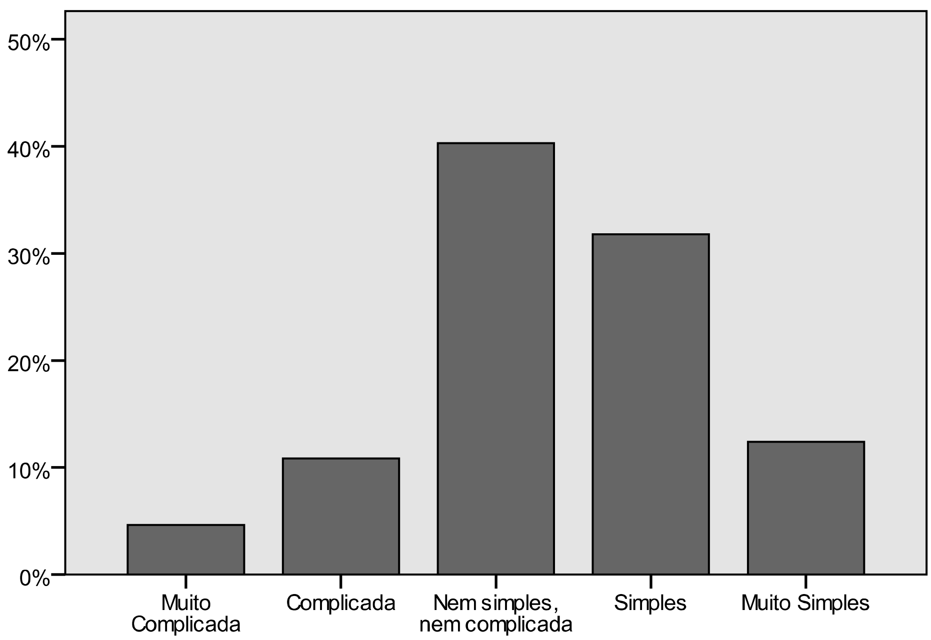
Se ignorar o facto de a variável “nome da escola” ser nominal e de a variável “opinião em relação aos conteúdos de Estatística” ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois mais de 20% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$.

Com o auxílio do *software* R, concluí que o valor do *p-value* associado ao teste exacto é 0,7551.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum dos níveis de significância usuais. Portanto, não há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação aos conteúdos de Estatística está associada à escola que frequentam.

No que diz respeito à forma como os conteúdos de Estatística foram abordados nas aulas de Estatística, a maioria dos inquiridos (40,3%) considera que a forma utilizada pelos docentes nem foi simples nem foi complicada, seguida da opinião de que foi simples com 31,8% de respostas (Gráfico 62).

Gráfico 62: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados



Para saber se há associação entre a escola e a opinião que os alunos têm em relação à forma como os conteúdos de Estatística foram abordados, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 10.º ano têm em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 32: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Estatística foram abordados e a escola que frequentam

			Opinião em relação à forma como os conteúdos de Estatística foram abordados			
			Simples e Muito Simples	Nem simples nem complicada	Complicada e Muito Complicada	Total
Nome da Escola	Escola Secundária da Baixa da Banheira	Valor Observado	17	12	1	30
		Valor Esperado	13,3	12,1	4,7	30,0
	Escola Secundária Ferreira Dias	Valor Observado	1	10	12	23
		Valor Esperado	10,2	9,3	3,6	23,0
	Escola Secundária Fernão Mendes Pinto	Valor Observado	13	11	1	25
		Valor Esperado	11,0	10,1	3,9	25,0
	EB 2,3/S Michel Giacometti	Valor Observado	26	19	6	51
		Valor Esperado	22,5	20,6	7,9	51,0
	Total	Valor Observado	57	52	20	129
		Valor Esperado	57,0	52,0	20,0	129,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação à forma como os conteúdos de Estatística foram abordados" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 32, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 29,52836 e que o *p-value* correspondente é 1,73413x10^{-6}.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para qualquer nível de significância. Portanto, há evidência para afirmar que a opinião que os alunos têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação à forma como os conteúdos de Estatística foram abordados" ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois 25% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$.

Com o auxílio do *software* R, concluí que o valor do *p-value* associado ao teste exacto é 7,26x10^{-6}.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para qualquer nível de significância. Portanto, há evidência para afirmar que a opinião que os alunos do 10.º ano têm em relação à forma como os conteúdos de Estatística foram abordados está associada à escola que frequentam.

De seguida, vou analisar por escola, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados; bem como, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados.

Tabela 33: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

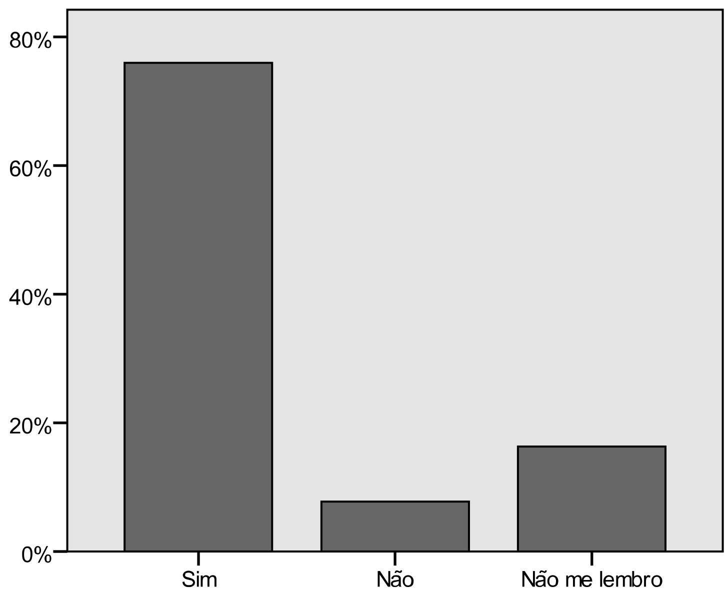
		Opinião dos alunos em relação à forma como os conteúdos de Estatística foram abordados	
EB 2,3/S Michel Giacometti	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}183$	$p = 0{,}199$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}612$	$p = 1{,}809 \times 10^{-6}$
Escola Secundária Ferreira Dias	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}257$	$p = 0{,}236$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}221$	$p = 0{,}312$
Escola Secundária Fernão Mendes Pinto	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = -0{,}145$	$p = 0{,}490$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}112$	$p = 0{,}593$
Escola Secundária Baixa da Banheira	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}150$	$p = 0{,}430$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}520$	$p = 3{,}2 \times 10^{-3}$

Pela análise da Tabela 33, concluí que não há em nenhuma das escolas associação entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos de Estatística foram abordados. Concluí ainda que na Escola 2,3/S Michel Giacometti e na Escola Secundária Baixa da Banheira há associação positiva e moderada entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como os mesmos foram abordados nas aulas. No entanto, na Escola Secundária Ferreira Dias e na Escola Secundária Fernão Mendes Pinto não existe associação entre as variáveis em questão.

Pela análise do Gráfico 63, concluí que nem todos os alunos abordaram o tema de Estatística no 3.º Ciclo do Ensino Básico. Da amostra recolhida, constatei que, aproximadamente, 8% dos inquiridos não tiveram aulas de Estatística no 3.º Ciclo, 16,3% não se recordam e apenas 76% tiveram aulas de Estatística no 3.º Ciclo do Ensino Básico.

Na realidade o que acontece é que, apesar do tema de Estatística estar incluído no programa de Matemática do 3.º Ciclo, nem sempre é leccionado, uma vez que faz parte das matérias remetidas para o final do programa e, por isso mesmo, nem sempre é apresentada aos alunos, tanto devido à falta de tempo, como devido à falta de convicção do seu real interesse (Branco, 2000a).

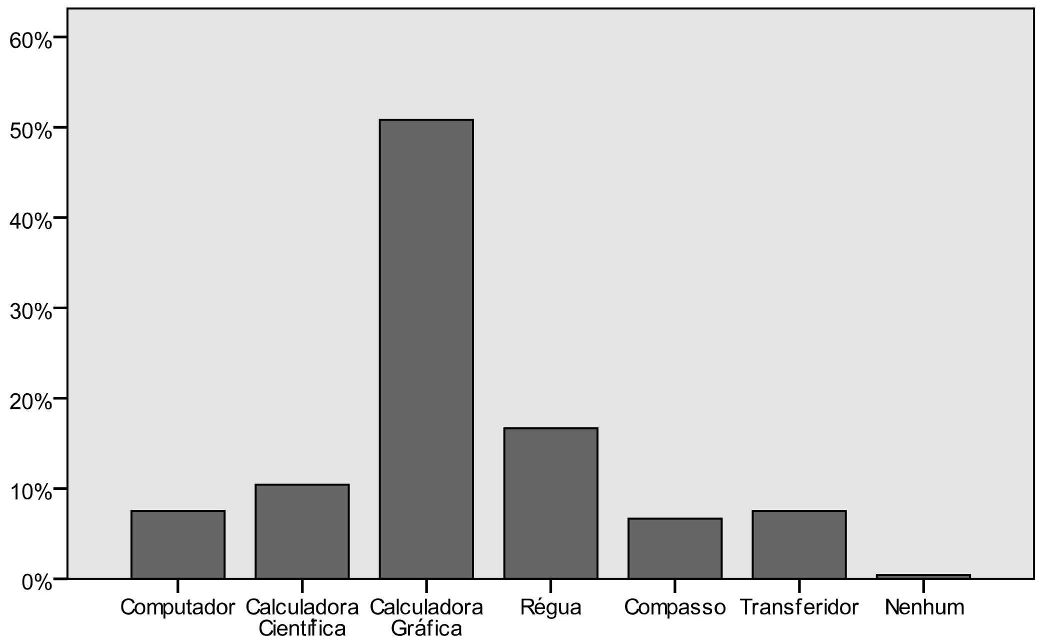
Gráfico 63: Alunos que tiveram aulas de Estatística no 3.º Ciclo do Ensino Básico



Quanto aos recursos utilizados durante as aulas de Estatística, concluí que a maioria dos estudantes (50,8%) utilizou a calculadora gráfica, seguida da régua e da calculadora científica com 16,7% e 10,4% de utilizadores, respectivamente. O computador foi apenas utilizado por 7,5% dos estudantes (Gráfico 64).

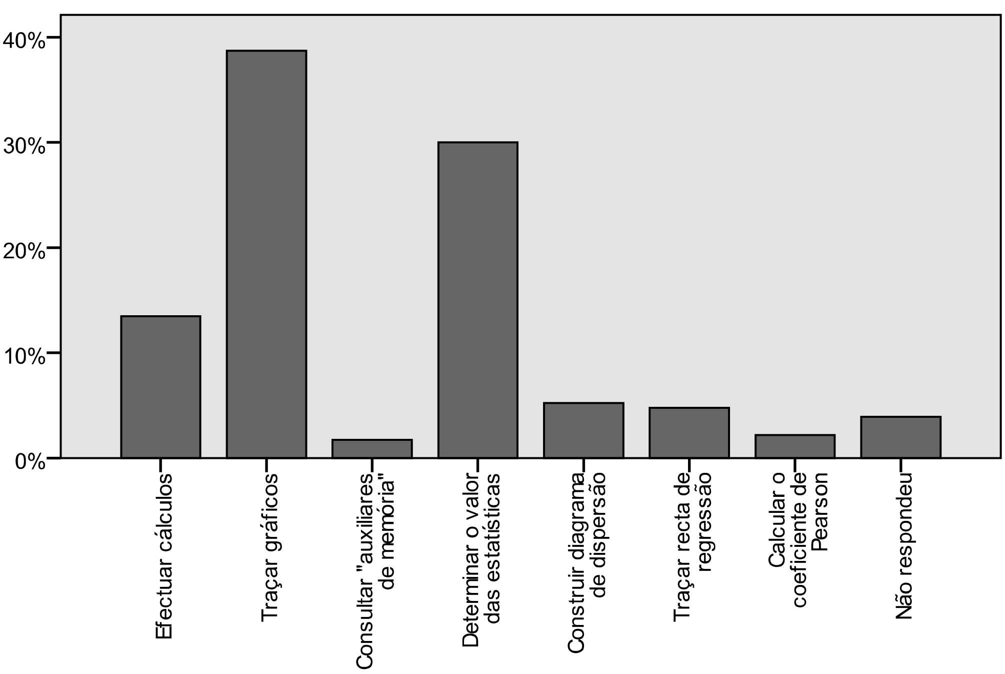
Apesar do programa de Matemática do 10.º ano sugerir a utilização de computadores e de calculadoras gráficas durante as aulas de Estatística, nem todos os alunos utilizam estes recursos. Muitos alunos não adquirem calculadoras gráficas e as escolas não as têm em número suficiente para suprir esta falta. De uma forma geral, não há recursos físicos, nem materiais adequados nas escolas, como por exemplo, uma sala de trabalho com computadores e outros materiais à disposição dos alunos.

Gráfico 64: Materiais usados durante as aulas de Estatística



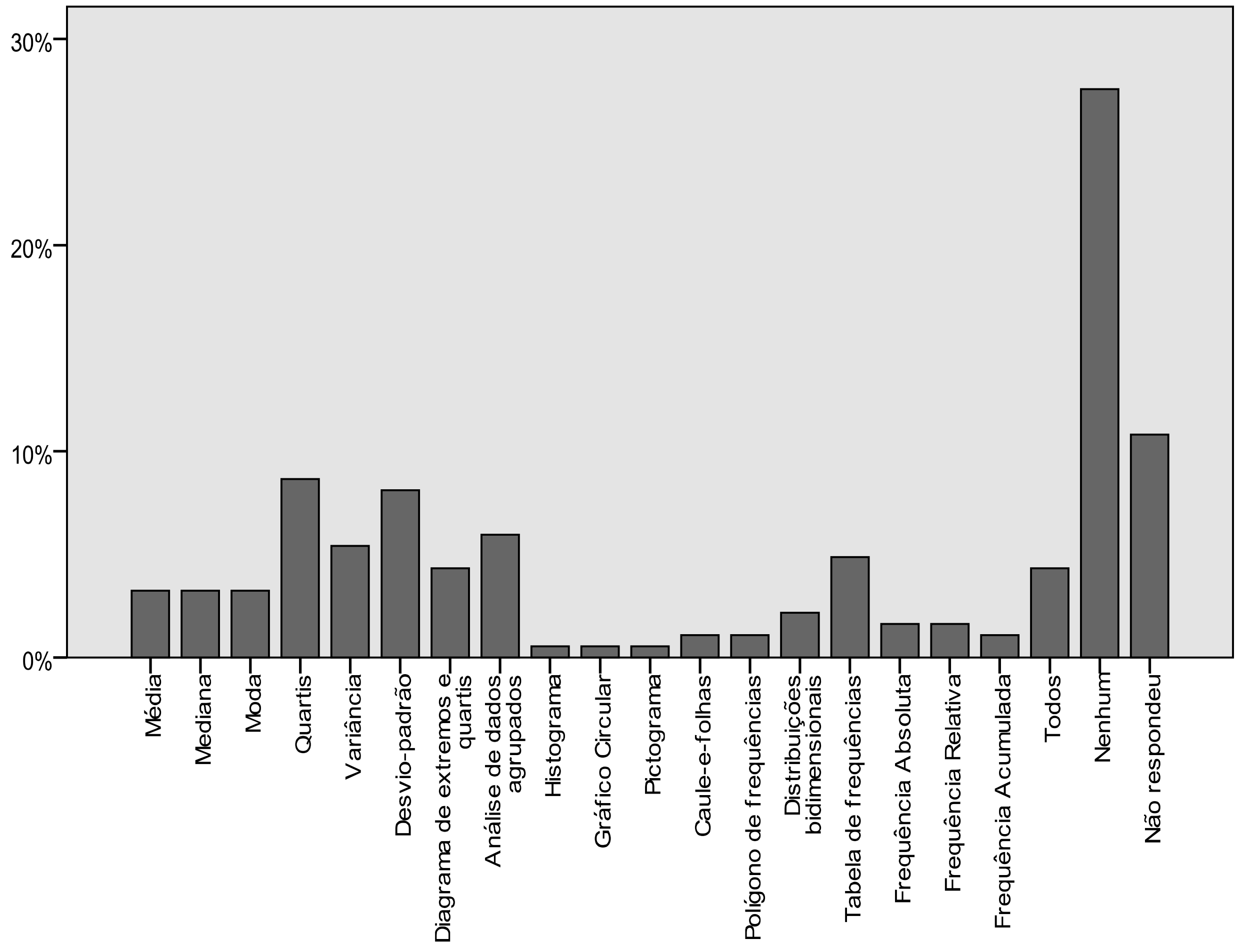
Da análise da amostra recolhida, concluí que a maioria dos inquiridos (38,7%) afirma que utilizou a calculadora gráfica durante as aulas de Estatística com a finalidade de traçar gráficos e 30% refere que a utilizou com o intuito de determinar o valor das estatísticas. No entanto, nem todos os alunos procuraram tirar o maior proveito da calculadora gráfica, pois 13,5% dos inquiridos afirmam ter utilizado a calculadora para efectuar cálculos elementares e aproximadamente, 2% dos inquiridos revelam ter utilizado a calculadora para consultar “auxiliares de memória” (Gráfico 65).

Gráfico 65: Finalidade da calculadora



Finalmente, no que se refere aos conteúdos de Estatística, a maioria dos alunos afirma ter sentido dificuldades em pelo menos um dos conteúdos abordados durante as aulas. Através da análise do Gráfico 66, constatei que os quartis e o desvio padrão são os conteúdos em que os alunos sentiram mais dificuldades (são mencionados por 8,6% e 8,1% dos inquiridos, respectivamente). Posteriormente, surge a análise de dados agrupados que é assinalada por 5,9% dos inquiridos. Para além destes valores, é importante realçar que 27,6% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos de Estatística e que 4,3% dos inquiridos afirmam ter sentido dificuldades em todos os conteúdos. Na minha opinião, a comunidade escolar deve ter em atenção estes últimos alunos, pois poderão estar a necessitar de algum tipo de apoio ou acompanhamento. É também importante destacar que, aproximadamente, 11% dos inquiridos não responderam a esta questão. Esta situação poderá ter estado relacionada com o facto de alguns alunos já não se lembrarem dos conteúdos de Estatística ou ainda com o facto de se tratar de uma questão de resposta aberta.

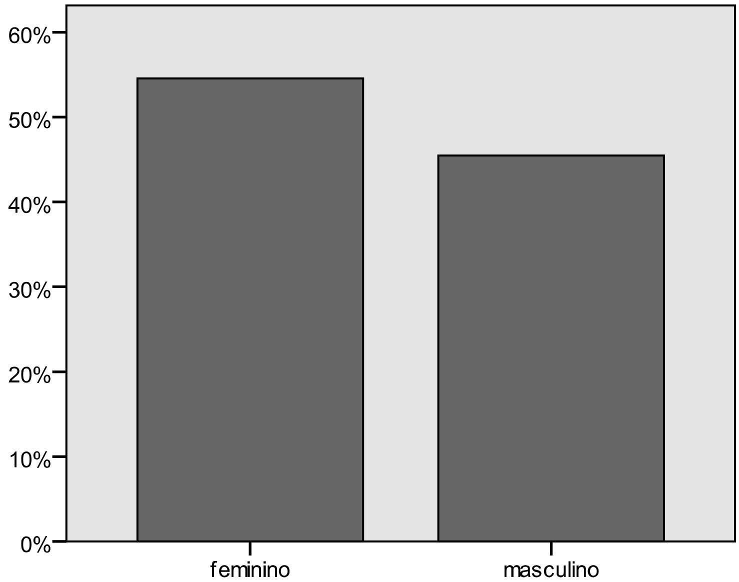
Gráfico 66: Conteúdos de Estatística em que os alunos sentiram dificuldades



5.2.7. 12.º Ano do Ensino Secundário

Da análise da amostra recolhida, constatei que 54,5% dos inquiridos são do género feminino, enquanto que 45,5% são do género masculino (Gráfico 67).

Gráfico 67: Distribuição por género



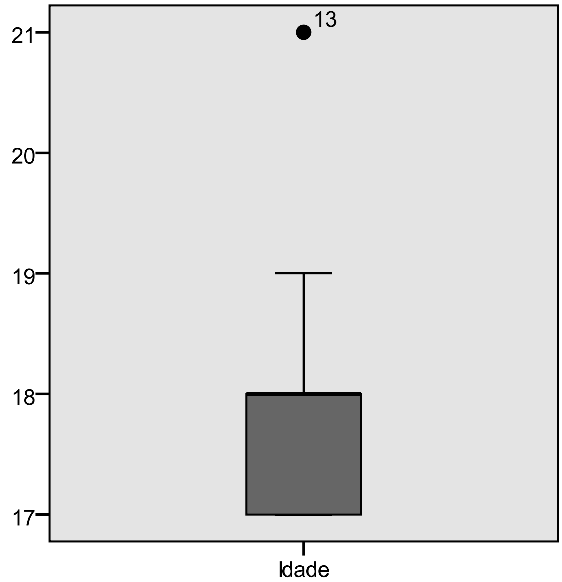
Através da análise da Tabela 34 e do Gráfico 68, concluí que dos 77 inquiridos, o mais novo tem 17 anos e o mais velho tem 21 anos. A moda e a mediana das idades são de 18 anos. A média das idades dos inquiridos é de 17,69 anos e a dispersão das observações em relação à média – desvio padrão – é de 0,7 anos, aproximadamente. Com base no Gráfico 68, verifiquei ainda a presença de um *outlier* moderado.

Analisando os valores anteriores, concluí que, de facto, faz sentido que a moda das idades seja de 18 anos, pois se um aluno fizer um percurso normal na escola quando chegar ao 12.º ano tem 17 ou 18 anos, consoante a data do seu aniversário. Portanto, também faz sentido que a idade mínima observada seja de 17 anos. No entanto, é do conhecimento geral que nem todos os alunos apresentam um percurso normal na escola, daí haver muitos inquiridos com idade superior à moda (este facto é facilmente confirmado através do Gráfico 68).

Tabela 34: Estatísticas da variável idade

N	77
Média	17,69
Mediana	18
Moda	18
Variância	0,428
Desvio Padrão	0,654
Mínimo	17
Máximo	21
1.º Quartil	17
2.º Quartil	18
3.º Quartil	18

Gráfico 68: *Box-plot* das idades dos inquiridos



Em relação ao curso que frequentam, a totalidade dos inquiridos frequenta o Curso Científico-Humanístico de Ciências e Tecnologias.

No que concerne à opinião que os alunos do 12.º ano têm em relação à disciplina de Matemática, a maioria dos inquiridos (58,4%) afirma gostar de Matemática e, aproximadamente, 30% dos inquiridos referem gostar muito desta disciplina (Gráfico 69). Apesar da opinião que os alunos têm em relação à disciplina de Matemática, aproximadamente, 50% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e 33,8% dos inquiridos auto-avaliam-no de Bom (Gráfico 70). De facto, ao consultar o Gráfico 71, constatei que o nível que mais se verificou no 1.º Período foi o 13 com 16%, seguido dos níveis 12 e 14 com 13,3% e 12%, respectivamente. No 2.º Período, os níveis mais frequentes foram o 10 e o 12 com 13,3% cada um, seguidos dos níveis 11 e 13 com 12% cada um (Gráfico 72). Apesar destas classificações, verifica-se que, aproximadamente, 7% dos alunos do Colégio São João de Brito tiveram classificação 20 no final do 2.º Período à disciplina de Matemática (Gráfico 72.2).

Com base nos dados recolhidos, concluí que apesar da maioria dos alunos gostar da disciplina de Matemática, eles não conseguem ter elevados rendimentos nesta disciplina.

Gráfico 69: Distribuição dos inquiridos segundo a sua opinião em relação à disciplina de Matemática

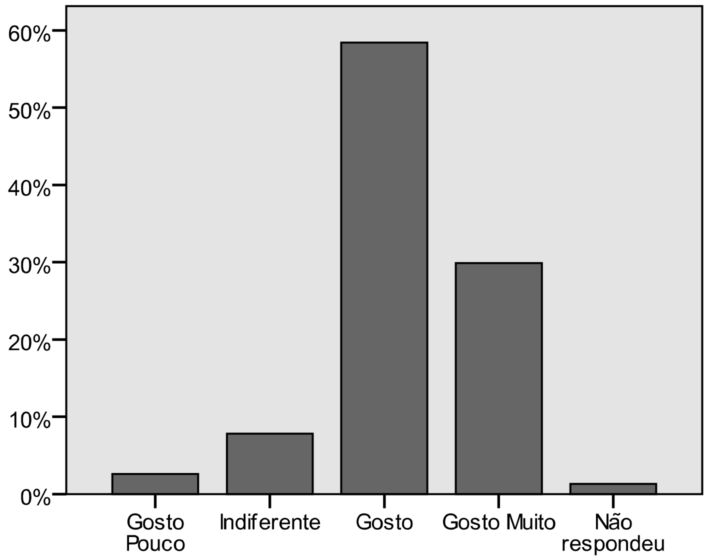


Gráfico 70: Distribuição dos inquiridos segundo a sua própria avaliação do seu rendimento a Matemática

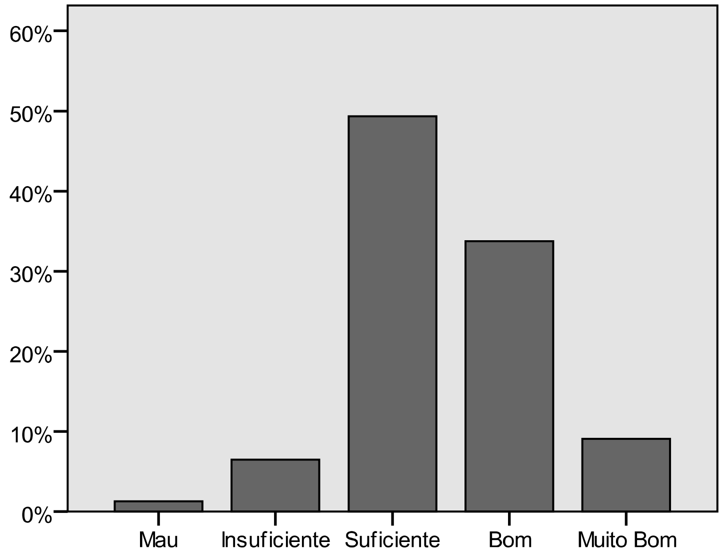


Gráfico 71: Classificação que os alunos obtiveram no 1.º Período

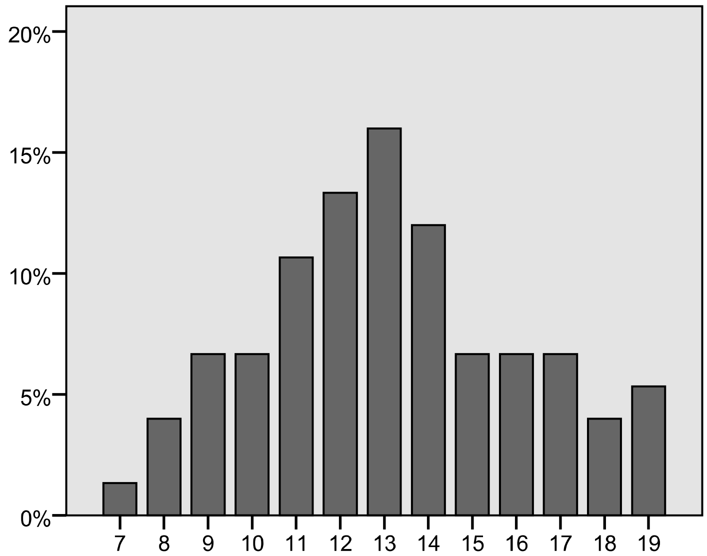


Gráfico 72: Classificação que os alunos obtiveram no 2.º Período

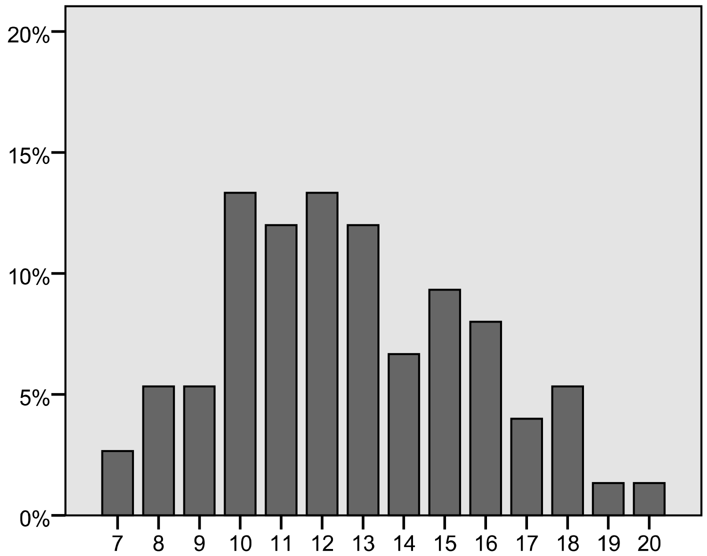


Gráfico 71.1: Classificação que os alunos das escolas públicas obtiveram no 1.º Período

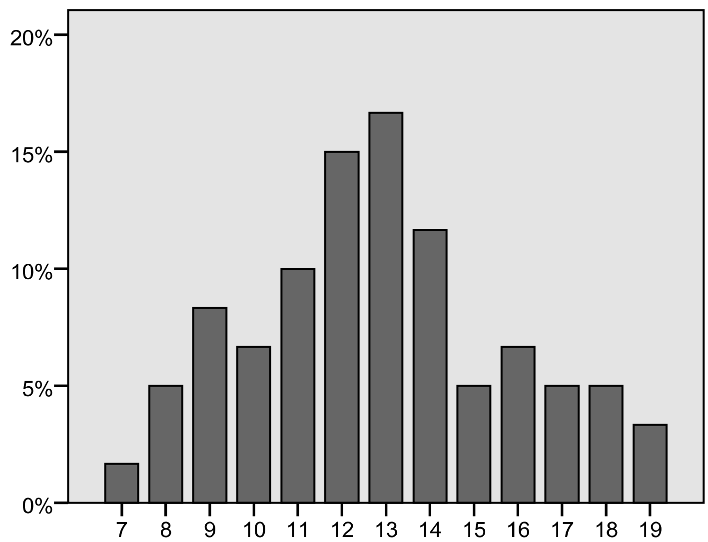


Gráfico 72.1: Classificação que os alunos das escolas públicas obtiveram no 2.º Período

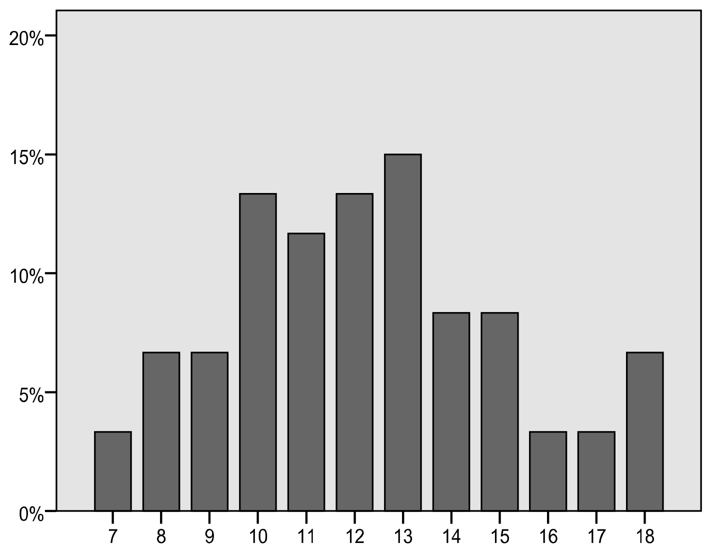


Gráfico 71.2: Classificação que os alunos do Colégio São João de Brito obtiveram no 1.º Período

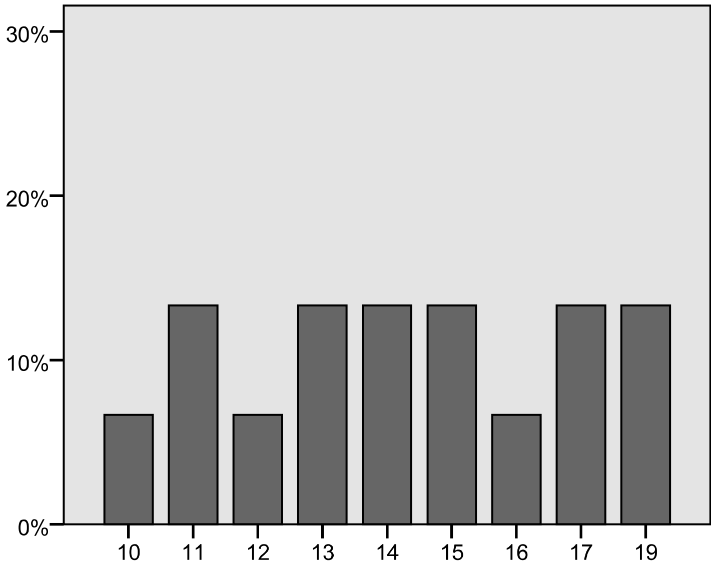
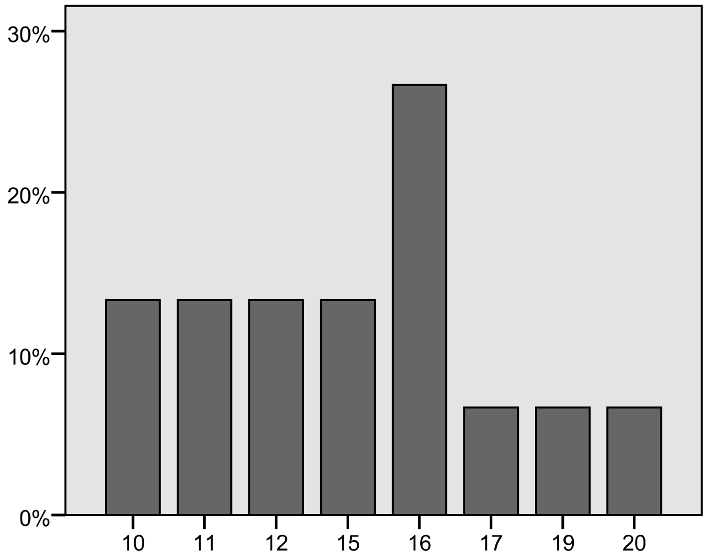


Gráfico 72.2: Classificação que os alunos do Colégio São João de Brito obtiveram no 2.º Período



Ao recorrer ao *software* R para saber o tipo de associação que há entre a opinião que os alunos do 12.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina, concluí que o valor do coeficiente de correlação de Spearman é 0,527 e que o valor do *p-value* associado ao teste $H_0 : \rho_S = 0$ é $1{,}027 \times 10^{-6}$. Face a estes valores, concluí que há uma associação positiva moderada entre a opinião que os alunos do 12.º ano têm em relação à Matemática e a avaliação que fazem do seu rendimento a esta disciplina. Sendo, portanto, de admitir que os alunos que avaliam o seu rendimento de forma negativa são também aqueles que têm uma opinião mais negativa desta disciplina.

De seguida, vou analisar o grau de associação que há entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º e 2.º Períodos, a esta disciplina.

Recorrendo novamente ao *software* R, obtive a seguinte tabela:

Tabela 35: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

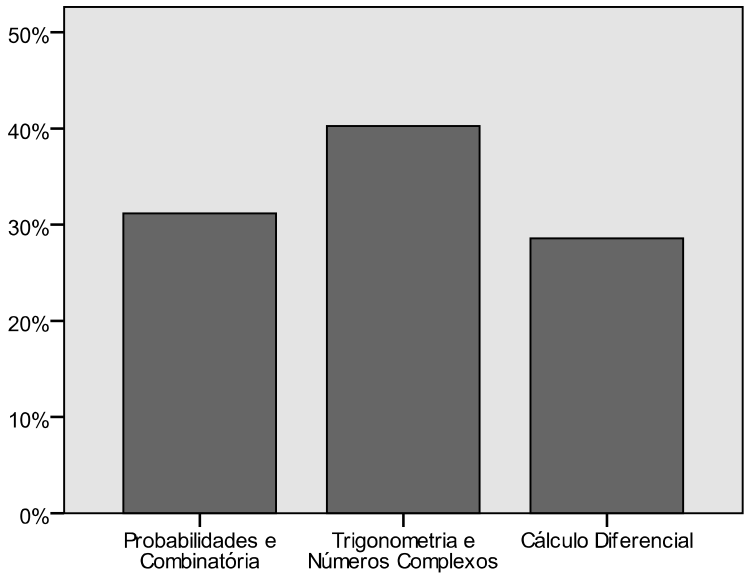
	Como avalia o seu rendimento a Matemática?	
Qual a classificação que obteve no 1.º Período?	$(r_S)_C = 0{,}745$	$p = 1{,}7 \times 10^{-14}$
Qual a classificação que obteve no 2.º Período?	$(r_S)_C = 0{,}845$	$p < 2{,}2 \times 10^{-16}$

Pela leitura dos valores da Tabela 35, concluí que há uma associação positiva forte entre a avaliação que os alunos fazem do seu rendimento a Matemática e a classificação que obtiveram, no 1.º Período, a esta disciplina. Esta associação torna-

-se ainda mais forte quando se relaciona a classificação que os alunos obtiveram no 2.º Período com a forma como avaliam o seu rendimento. O que significa que ao longo do ano lectivo, os alunos vão conseguindo auto-avaliar-se correctamente no que diz respeito ao seu rendimento na disciplina de Matemática.

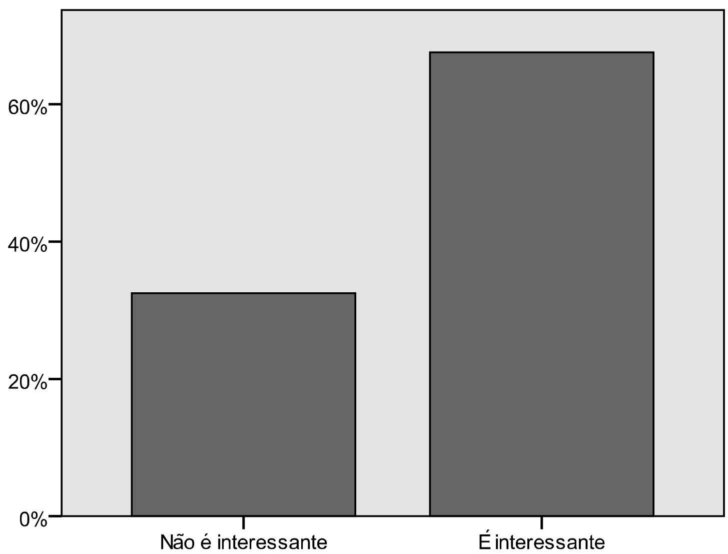
Quantos aos temas leccionados durante o ano lectivo na disciplina de Matemática, aproximadamente, 40% dos inquiridos seleccionam a temática de Trigonometria e Números Complexos como sendo a sua preferida, 31,2% dos inquiridos preferem a temática de Probabilidades e Combinatória e finalmente, surge o Cálculo Diferencial com 28,6% de votos (Gráfico 73).

Gráfico 73: Distribuição dos inquiridos de acordo com o tema preferido



Em relação à temática de Probabilidades e Combinatória, a maioria dos inquiridos (67,5%) considera esta temática interessante. No entanto, os restantes alunos (32,5%) consideram que o tema não é interessante (Gráfico 74).

Gráfico 74: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Probabilidades e Combinatória



Com o intuito de saber se há associação entre a escola e a opinião que os alunos têm em relação ao tema de Probabilidades e Combinatória, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 12.º ano têm em relação ao tema de Probabilidades e Combinatória e a escola onde estudam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 36: Distribuição dos inquiridos segundo a sua opinião em relação ao tema de Probabilidades e Combinatória e a escola que frequentam

			Opinião em relação ao tema de Probabilidades e Combinatória		
			Não é interessante	É interessante	Total
Nome da Escola	Escola Secundária de Gil Vicente	Valor Observado	3	18	21
		Valor Esperado	6,8	14,2	21,0
	EB 2,3/S Michel Giacometti	Valor Observado	21	19	40
		Valor Esperado	13,0	27,0	40,0
	Colégio São João de Brito	Valor Observado	1	15	16
		Valor Esperado	5,2	10,8	16,0
	Total	Valor Observado	25	52	77
		Valor Esperado t	25,0	52,0	77,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação ao tema de Probabilidades e Combinatória" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 36, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 15,30158 e que o *p-value* correspondente a esta estatística de teste é $4{,}75669 \times 10^{-4}$.

Como o valor do *p-value* é muito pequeno, rejeito H_0 para qualquer nível de significância. Portanto, há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação ao tema de Probabilidades e Combinatória depende da escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação ao tema de Probabilidades e Combinatória" ser ordinal, estou em condições de aplicar o teste do qui-quadrado para testar H_0, pois nenhum dos valores esperados é inferior a cinco.

Com o auxílio do *SPSS* e dos dados da Tabela 36, obtive a seguinte tabela:

Tabela 37: Valores das estatísticas de teste do Qui-Quadrado e de Likelihood Ratio e respectivos *p-values*

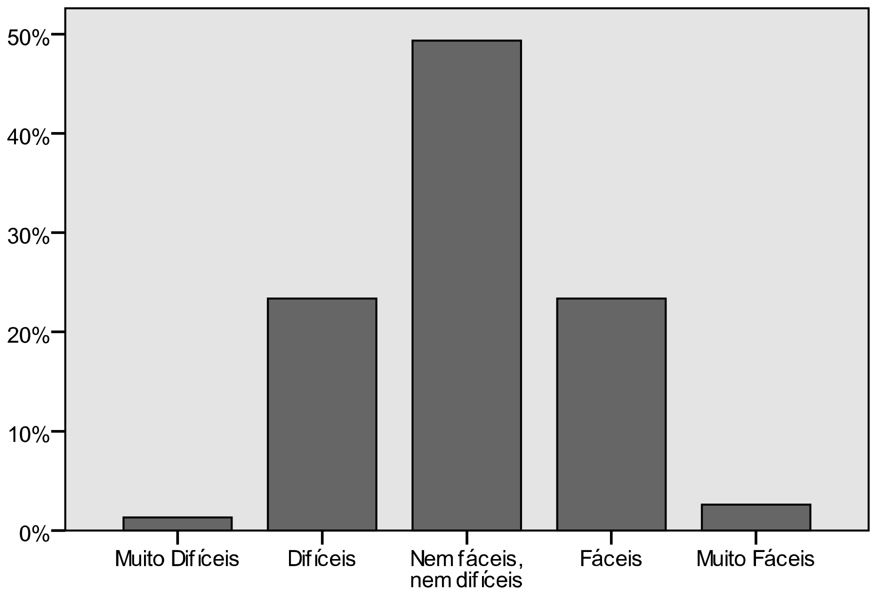
Teste	Valor da Estatística de Teste	Graus de Liberdade	*p-value*
Qui-Quadrado	15,50291	2	$4{,}30116\text{x}10^{-4}$
Likelihood Ratio	17,01495	2	$2{,}01954\text{x}10^{-4}$

Através da análise da Tabela 37, concluí que o valor da estatística de teste, X^2, é 15,50291 e que o *p-value* correspondente a esta estatística de teste é $4{,}30116\text{x}10^{-4}$. Concluí ainda que o valor da estatística de teste de Likelihood Ratio, G^2, é 17,01495 e que o valor do *p-value* correspondente a esta estatística de teste é $2{,}01954\text{x}10^{-4}$.

Como os valores dos *p-values* são praticamente nulos, rejeito H_0 para qualquer nível de significância. Portanto, há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação ao tema de Probabilidades e Combinatória depende da escola que frequentam.

No que diz respeito à opinião que os alunos do 12.º ano têm em relação aos conteúdos de Probabilidades e Combinatória, a maioria dos inquiridos (49,4%) considera que os conteúdos nem são fáceis nem são difíceis. No entanto, 23,4% dos inquiridos avaliam-nos como sendo fáceis e 23,4% avaliam-nos como sendo difíceis (Gráfico 75).

Gráfico 75: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Probabilidades e Combinatória



Com o objectivo de saber se há associação entre a escola e a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Estatística, vou realizar um teste de independência para testar:

> *H_0: Há independência entre a opinião que os alunos do 12.º ano têm em relação aos conteúdos de Probabilidades e Combinatória e a escola que frequentam.*

Da amostra recolhida, obtive a seguinte tabela:

Tabela 38: Distribuição dos inquiridos segundo a sua opinião em relação aos conteúdos de Probabilidades e Combinatória e a escola que frequentam

			Opinião em relação aos conteúdos de Probabilidades e Combinatória			
			Fáceis e Muito Fáceis	Nem fáceis nem difíceis	Difíceis e Muito Difíceis	Total
Nome da Escola	Escola Secundária de Gil Vicente	Valor Observado	8	11	2	21
		Valor Esperado	5,5	10,4	5,2	21,0
	EB 2,3/S Michel Giacometti	Valor Observado	8	19	13	40
		Valor Esperado	10,4	19,7	9,9	40,0
	Colégio São João de Brito	Valor Observado	4	8	4	16
		Valor Esperado	4,2	7,9	3,9	16,0
	Total	Valor Observado	20	38	19	77
		Valor Esperado	20,0	38,0	19,0	77,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação aos conteúdos de Probabilidades e Combinatória" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 38, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 4,52435 e que o *p-value* correspondente é 0,10412.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação aos conteúdos de Probabilidades e Combinatória está associada à escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação aos conteúdos de Probabilidades e Combinatória" ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois mais de

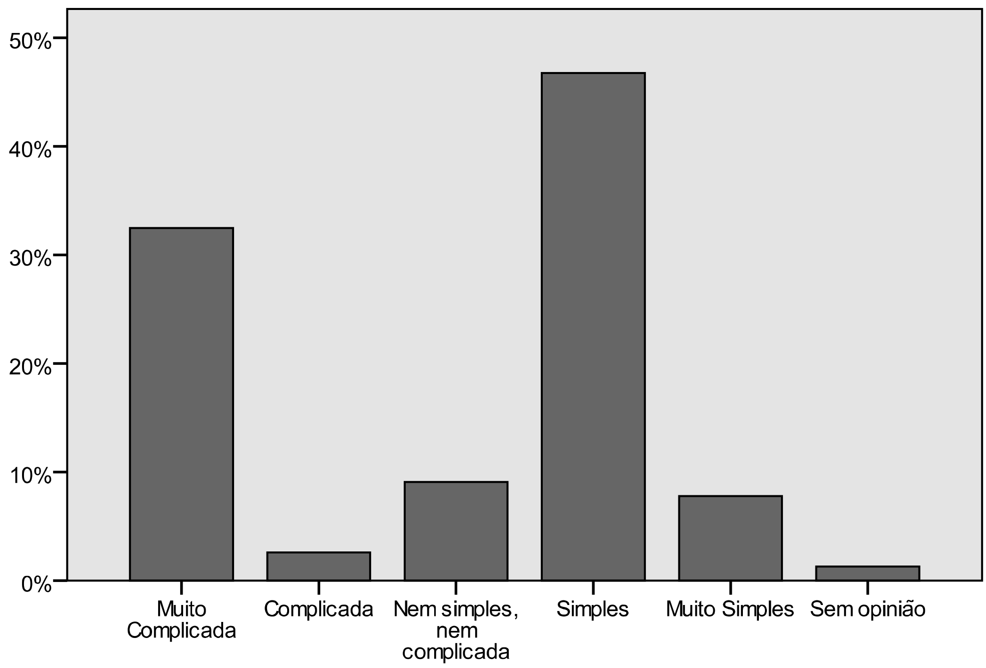
20% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$.

Com o auxílio do *software* R e da Tabela 38, concluí que o valor do *p-value* associado ao teste exacto é 0,3048.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação aos conteúdos de Probabilidades e Combinatória está associada à escola que frequentam.

No que se refere à forma como os conteúdos de Probabilidades e Combinatória foram abordados nas aulas, a maioria dos inquiridos (46,8%) considera que a forma utilizada foi simples, seguida da opinião de que foi muito complicada com 32,5% de respostas (Gráfico 76).

Gráfico 76: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados



Com o objectivo de saber se há associação entre a escola e a opinião que os alunos têm em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados, vou realizar um teste de independência para testar:

H_0: Há independência entre a opinião que os alunos do 12.º ano têm em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados e a escola que frequentam.

Da amostra recolhida, obtive a seguinte tabela:

Tabela 39: Distribuição dos inquiridos segundo a sua opinião em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados e a escola que frequentam

			Opinião em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados			
			Simples e Muito Simples	Nem simples Nem complicada	Complicada e Muito Complicada	Total
Nome da Escola	Escola Secundária de Gil Vicente	Valor Observado	15	1	5	21
		Valor Esperado	11,9	1,1	8,0	21,0
	EB 2,3/S Michel Giacometti	Valor Observado	19	2	18	39
		Valor Esperado	22,1	2,1	14,9	39,0
	Colégio São João de Brito	Valor Observado	9	1	6	16
		Valor Esperado	9,1	0,8	6,1	16,0
	Total	Valor Observado	43	4	29	76
		Valor Esperado	43,0	4,0	29,0	76,0

Uma vez que a variável "nome da escola" é do tipo nominal e a variável "opinião em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados" é do tipo ordinal, vou recorrer ao teste de Kruskal-Wallis para testar H_0.

Com o auxílio do *SPSS* e dos dados da Tabela 39, concluí que o valor da estatística de teste de Kruskal-Wallis com correcção, H_C, é 2,98605 e que o *p-value* correspondente a esta estatística de teste é 0,22469.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados está associada à escola que frequentam.

Se ignorar o facto de a variável "nome da escola" ser nominal e de a variável "opinião em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados" ser ordinal, não vou estar em condições de aplicar o teste do qui-quadrado para testar H_0, pois 33,3% dos valores esperados são inferiores a cinco, ou seja, não estão satisfeitas as condições de Cochran. Assim sendo, vou recorrer ao teste exacto para tabelas $r \times c$.

Com o auxílio do *software* R e da Tabela 39, concluí que o valor do *p-value* associado ao teste exacto é 0,4679.

Com base no valor do *p-value*, concluí que não rejeito H_0 para nenhum nível de significância usual. Portanto, não há evidência para afirmar que a opinião que os alunos do 12.º ano têm em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados está associada à escola que frequentam.

Acabo de concluir que a opinião que os alunos do 12.º ano têm em relação à forma como os conteúdos de Probabilidades e Combinatória foram abordados não depende da escola que frequentam, no entanto, tinha visto atrás que a opinião que os alunos do 12.º ano têm em relação ao tema de Probabilidades e Combinatória depende da escola que frequentam. Estes dois resultados, aparentemente contraditórios, levam-me, novamente, a reflectir sobre o seguinte: Será que se este estudo tivesse sido alargado a mais escolas e a mais alunos, as conclusões seriam as mesmas?

De seguida, vou analisar por escola, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Probabilidades e Combinatória e a forma como os conteúdos foram leccionados; bem como, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Combinatória e a forma como estes foram leccionados.

Tabela 40: Valores do Coeficiente de Correlação de Spearman e *p-values* associados ao teste $H_0 : \rho_S = 0$

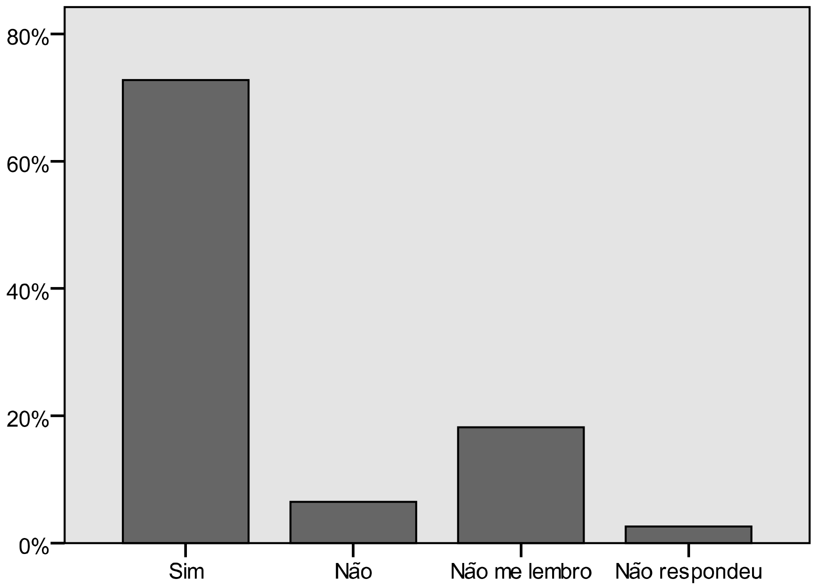
		Opinião dos alunos em relação à forma como os conteúdos de Estatística foram abordados	
EB 2,3/S Michel Giacometti	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}441$	$p = 0{,}005$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}444$	$p = 0{,}005$
Escola Secundária Gil Vicente	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}365$	$p = 0{,}104$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}110$	$p = 0{,}635$
Colégio São João de Brito	Opinião dos alunos em relação ao tema de Estatística	$(r_S)_C = 0{,}157$	$p = 0{,}561$
	Opinião dos alunos em relação aos conteúdos de Estatística	$(r_S)_C = 0{,}323$	$p = 0{,}223$

Pela análise da Tabela 40, concluí que na Escola Básica 2,3/S Michel Giacometti o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Probabilidades e Combinatória e a forma como os conteúdos foram leccionados é positivo e moderado. Verifica-se ainda, para esta escola, uma situação análoga no que diz respeito ao grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Combinatória e a forma como os mesmos foram abordados nas aulas. No entanto, na Escola Secundária Gil Vicente e no Colégio São João de Brito, não há associação linear entre as variáveis em questão.

Pela leitura do Gráfico 77, concluí que nem todos alunos abordaram o tema de Probabilidades e Estatística no 3.º Ciclo do Ensino Básico. Da amostra recolhida, constatei que, aproximadamente, 7% dos inquiridos não tiveram aulas de Probabilidades e Estatística no 3.º Ciclo do Ensino Básico, 18,2% dos inquiridos não se recordam e apenas 72,7% tiveram aulas de Probabilidades e Estatística.

Tal como já referi, o que de facto acontece é que, apesar deste tema fazer parte do Currículo do 3.º Ciclo do Ensino Básico, nem sempre é leccionado, uma vez que faz parte das matérias remetidas para o final do programa e, por isso mesmo, nem sempre é apresentado aos alunos, tanto devido à falta de tempo, como devido à falta de convicção do seu real interesse (Branco, 2000a).

Gráfico 77: Alunos que tiveram aulas de Probabilidades e Estatística no 3.º Ciclo do Ensino Básico

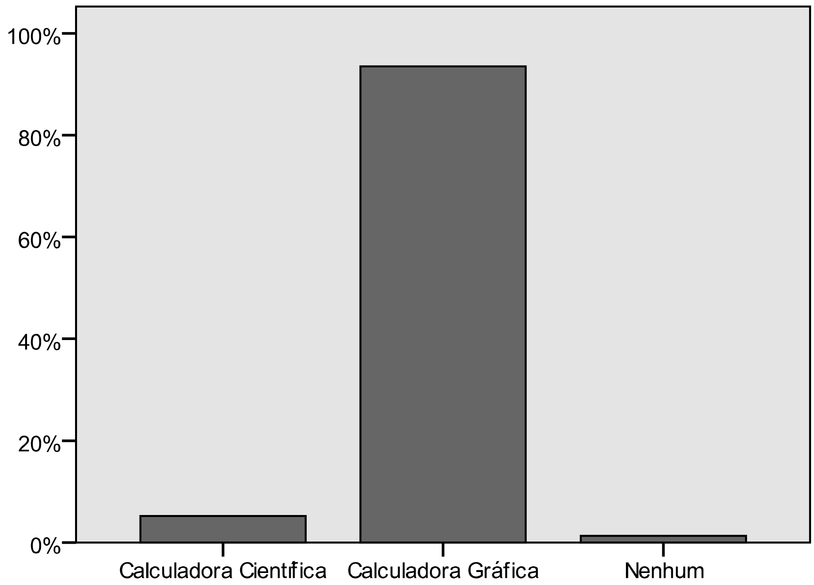


Quanto aos materiais utilizados durante as aulas de Probabilidades e Combinatória, a generalidade dos inquiridos (93,5%) utilizaram calculadora gráfica. No entanto, 5,2% dos inquiridos utilizaram calculadora científica e 1,3% dos inquiridos não utilizaram qualquer tipo de material (Gráfico 78).

Apesar do programa de Matemática do 12.º ano sugerir a utilização de calculadoras gráficas nas aulas de Probabilidades e Combinatória, nem todos os alunos as utilizam.

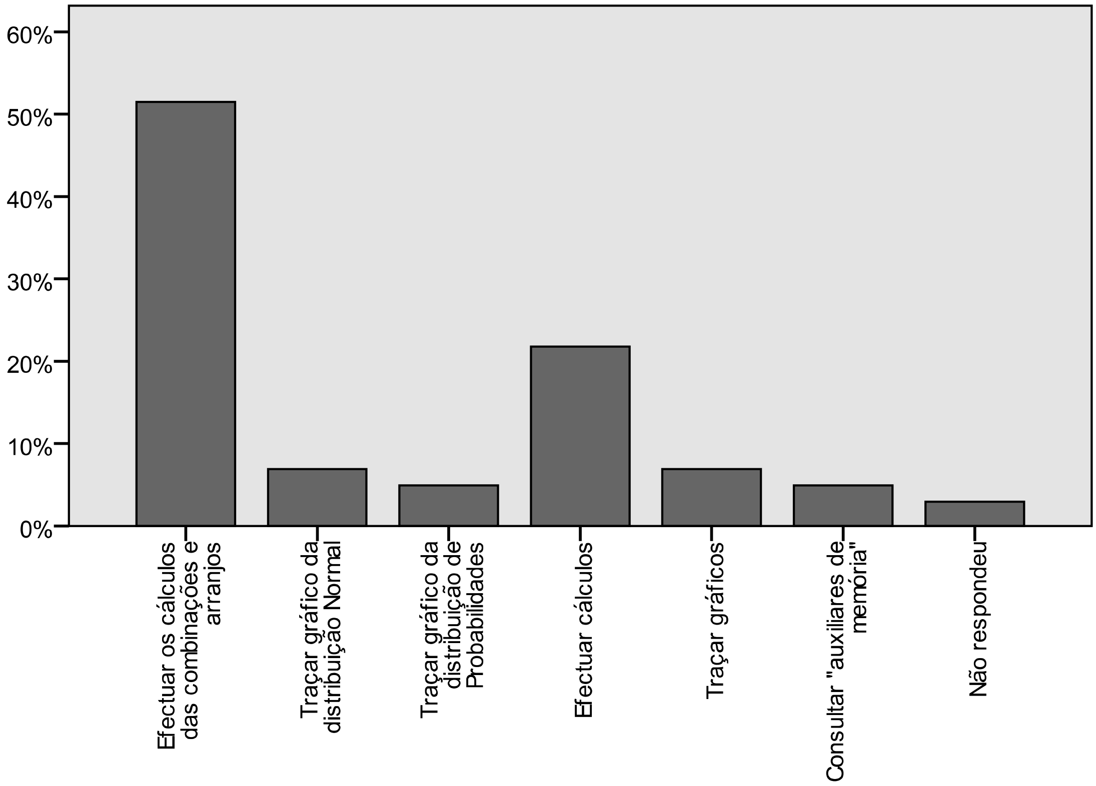
Na realidade, muitos alunos não adquirem calculadoras gráficas e as escolas não as têm em número suficiente para suprir esta falta.

Gráfico 78: Materiais usados durante as aulas de Probabilidades e Combinatória



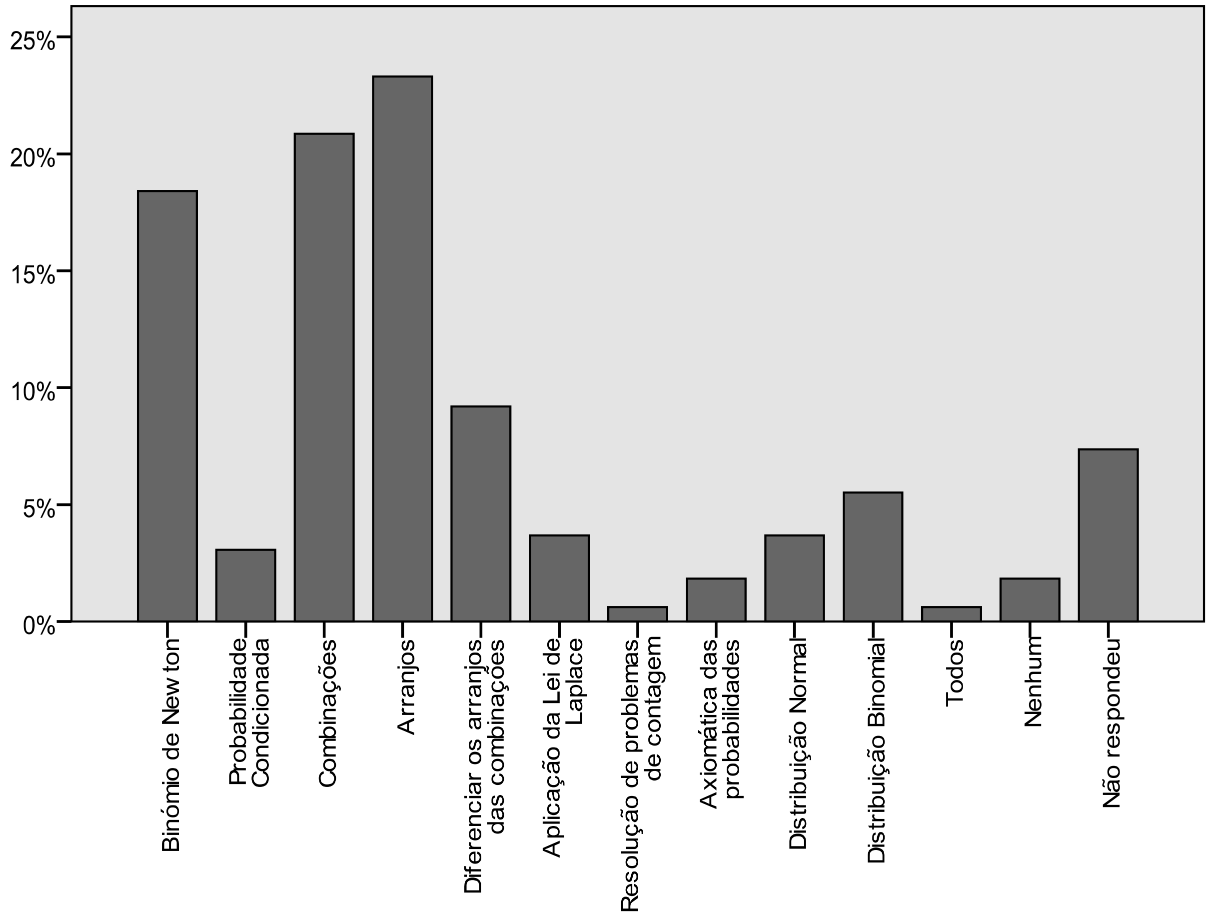
Pela análise do Gráfico 79, concluí que a maioria dos inquiridos (51,5%) utilizou a calculadora gráfica, durante as aulas de Probabilidades e Combinatória, com o objectivo de efectuar os cálculos das combinações e dos arranjos. No entanto, nem todos os alunos procuraram tirar o maior proveito das calculadoras gráficas, aproximadamente, 22% dos inquiridos utilizaram a calculadora para efectuar cálculos elementares e, aproximadamente, 5% revelaram ter utilizado a calculadora para consultar "auxiliares de memória".

Gráfico 79: Finalidade da calculadora



Finalmente, no que diz respeito aos conteúdos de Probabilidades e Combinatória, a generalidade dos inquiridos afirma ter sentido dificuldades em pelo menos um dos conteúdos abordados. Pela leitura do Gráfico 80, constatei que os arranjos e as combinações são os conteúdos em que os alunos sentiram mais dificuldades (são mencionados por 23,3% e 20,9% dos inquiridos, respectivamente). Posteriormente, surge o binómio de Newton que é referido por 18,4% dos inquiridos. Para além destes valores, é importante realçar que 1,8% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos de Probabilidades e Combinatória e que 0,6% dos inquiridos revelam ter sentido dificuldades em todos os conteúdos. Na minha opinião, a comunidade escolar deve ter em atenção estes últimos alunos, pois poderão estar a necessitar de algum tipo de apoio ou acompanhamento. É também importante destacar que, aproximadamente, 7% dos inquiridos não respondeu a esta questão. Esta situação poderá ter estado relacionada com o facto de alguns alunos já não se lembrarem dos conteúdos de Probabilidades e Combinatória ou ainda com o facto de se tratar de uma questão de resposta aberta.

Gráfico 80: Conteúdos de Probabilidades e Combinatória em que os alunos sentiram dificuldades



5.2.8. Comparação entre a Estatística do 10.º Ano e as Probabilidades e Combinatória do 12.º Ano

Supõe-se que, de um modo geral, os alunos preferem o tema de Estatística, que é leccionado no 10.º ano, ao tema de Probabilidades e Combinatória, que é leccionado no 12.º ano.

Para poder testar a validade desta suposição e uma vez que se tratam de amostras independentes, vou realizar, para a Escola Básica 2,3/S Michel Giacometti, um teste de hipóteses sobre a diferença de proporções.

EB 2,3/S Michel Giacometti

Seja X_1 a variável aleatória que representa o número de alunos do 10.º ano que consideram o tema de Estatística interessante.

Seja X_2 a variável aleatória que representa o número de alunos do 12.º ano que consideram o tema de Probabilidades e Combinatória interessante.

Seja p_1 a proporção de alunos do 10.º ano que consideram o tema de Estatística interessante.

Seja p_2 a proporção de alunos do 12.º ano que consideram o tema de Probabilidades e Combinatória interessante.

Pretende-se testar o seguinte:

$$H_0 : p_1 = p_2 \qquad vs. \qquad H_1 : p_1 > p_2$$

Na escola EB 2,3/S Michel Giacometti foram inquiridos 51 alunos do 10.º ano, dos quais 40 consideram o tema de Estatística interessante e foram inquiridos 40 alunos do 12.º ano, dos quais 19 consideram o tema de Probabilidades e Combinatória interessante.

Assim, $\hat{p}_1 = \dfrac{40}{51}$; $\hat{p}_2 = \dfrac{19}{40}$; $\overline{p} = \dfrac{59}{91}$; $\overline{q} = \dfrac{32}{91}$

$$Z_C\,(obs.) = \frac{\left|\dfrac{40}{51} - \dfrac{19}{40}\right| - \dfrac{1}{2}\left(\dfrac{1}{51} + \dfrac{1}{40}\right)}{\sqrt{\dfrac{\dfrac{59}{91}\times\dfrac{32}{91}}{51} + \dfrac{\dfrac{59}{91}\times\dfrac{32}{91}}{40}}} = 2{,}85$$

O *p-value* correspondente desta estatística de teste é dado por:

$$p = \mathrm{P}\left[Z_C \geq 2{,}85\right] = 1 - \Phi(2{,}85) = 1 - 0{,}9978 = 0{,}0022$$

Com base no valor do *p-value*, concluí que rejeito H_0 para todos os níveis de significância usuais. Portanto, há evidência significativa para afirmar que a proporção de alunos que prefere o tema de Estatística do 10.º ano é superior à proporção de alunos que prefere o tema de Probabilidades e Combinatória do 12.º ano.

5.3. Conclusões da Análise de Dados

5.3.1. 7.º Ano do Ensino Básico

- A amostra recolhida é constituída por 361 alunos, sendo 192 deles do sexo masculino e 169 do feminino.

- As idades dos alunos inquiridos variam entre os 12 e os 17 anos, inclusive.

- A maioria dos inquiridos (38,2%) afirma gostar da disciplina de Matemática. No entanto, aproximadamente, 48% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e cerca de 24% de Insuficiente.

- Há uma associação positiva e moderada entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

- Tanto no final do 1.º como do 2.º Períodos, a classificação que mais se verificou na disciplina de Matemática foi 3, seguida da classificação 2. Apesar de a maioria dos alunos do 7.º ano gostar de Matemática, o seu rendimento a esta disciplina não é equivalente à opinião que têm da mesma.

- A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte. Além disso, esta associação torna-se ainda mais forte no 2.º Período. O que significa que com o decorrer do ano lectivo os alunos vão conseguindo auto-avaliar-se correctamente.

- Dos temas leccionados durante o ano lectivo, o tema de Cálculo foi o eleito pela maioria dos alunos (37,7%), seguido dos temas de Estatística, de Geometria e de Funções com 28,5%, 27,7% e 5% de votos, respectivamente.

- Cerca de 31% dos inquiridos ainda não tinham abordado o tema de Estatística quando os inquéritos foram realizados.

• Dos alunos que já tinham tido contacto com o tema de Estatística, cerca de 47% consideram que o tema é interessante e, aproximadamente, 21% dos inquiridos não o consideram interessante. Sabe-se ainda que a opinião que os alunos têm em relação ao tema de Estatística é independente da escola que frequentam.

• Dos alunos que já tinham abordado os conteúdos de Estatística, cerca de 36% dos alunos não os consideram fáceis nem difíceis; 13,3% avaliam-nos como sendo fáceis; 8,9% como sendo muito fáceis; 6,9% como sendo difíceis e apenas 3,6% como sendo muito difíceis.

• No que diz respeito à forma como os professores abordaram os conteúdos de Estatística, cerca de 27% dos inquiridos, que já tinham tido contacto com esta temática, consideram que a forma utilizada nem foi simples nem foi complicada; 23,8% consideram-na simples; 6,4% complicada; 6,1% muito simples e apenas 3,9% muito complicada. Sabe-se ainda que a opinião que os alunos têm em relação à forma como os conteúdos foram abordados depende da escola que frequentam.

• Ao nível de significância de 5%, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados é, na maioria das escolas, positivo e moderado. Portanto, pode-se admitir que os alunos que estão mais satisfeitos com a forma como os conteúdos de Estatística foram abordados são também aqueles que têm uma opinião mais positiva em relação a este tema.

• A associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados é, na maioria das escolas, positiva e forte. Portanto, pode-se admitir que os alunos que estão mais satisfeitos com a forma como os conteúdos de Estatística foram abordados são também aqueles que têm uma opinião mais positiva em relação a estes conteúdos.

• Em relação aos materiais utilizados nas aulas de Estatística, aproximadamente, 30% dos inquiridos utilizaram máquina de calcular; 24,7% utilizaram régua; 12,9% utilizaram compasso; 12,5% utilizaram transferidor e apenas 0,9% dos alunos utilizaram computador. Apesar do programa de Matemática sugerir a utilização de computadores, são poucos os professores que planificam as suas aulas com vista à utilização desta tecnologia. Esta situação prende-se, essencialmente, com o facto de não haver recursos físicos, nem materiais adequados nas escolas, como por exemplo, uma sala de trabalho com computadores à disposição dos alunos ou ainda com o facto de a maioria das aulas de Matemática não decorrer em salas com computadores.

• No que se refere aos conteúdos de Estatística, a maioria dos alunos refere ter sentido dificuldades em pelo menos um dos conteúdos abordados; 9,8% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos; 4,4% sentiram dificuldades em todos os conteúdos.

5.3.2. 8.º Ano do Ensino Básico

• A amostra recolhida é constituída por 319 alunos, sendo 154 deles do sexo masculino e 165 do feminino.

• As idades dos alunos inquiridos variam entre os 12 e os 17 anos, inclusive.

• A maioria dos inquiridos (37%) afirma gostar da disciplina de Matemática. No entanto, aproximadamente, 42% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e cerca de 26% de Insuficiente.

• Há uma associação positiva e moderada entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

• No 1.º Período, a classificação que mais se verificou foi 2 com 42,8%, seguida da classificação 3 com 39%. No 2.º Período, a classificação que mais se verificou foi 3 com 41,6%, seguida da classificação 2 com 39,7%. Apesar de a maioria dos alunos do 8.º ano gostar de Matemática, o seu rendimento a esta disciplina não é equivalente à opinião que têm da mesma.

• A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte. Além disso, esta associação torna-se ainda mais forte no 2.º Período. O que significa que com o decorrer do ano lectivo os alunos vão conseguindo auto-avaliar-se correctamente.

• Dos temas leccionados durante o ano lectivo, o tema de Estatística foi o eleito pela maioria dos alunos (43,9%), seguido dos temas de Cálculo, de Geometria e de Funções com 22,3%, 21,6% e 11,3% de votos, respectivamente.

• Quanto ao tema de Estatística, cerca de 75% consideram que o tema é interessante e, aproximadamente, 25% dos inquiridos não o consideram interessante. Sabe-se ainda que a opinião que os alunos têm em relação ao tema de Estatística é independente da escola que frequentam.

• No que se refere aos conteúdos de Estatística, cerca de 43% dos alunos não os consideram fáceis nem difíceis; 27,3% avaliam-nos como sendo fáceis; 20,1% como sendo muito fáceis; 5,3% como sendo difíceis e apenas 4,4% como sendo muito difíceis. Sabe-se ainda que a opinião que os alunos têm em relação aos conteúdos de Estatística é independente da escola que frequentam.

• No que diz respeito à forma como os professores abordaram os conteúdos de Estatística, cerca de 37% dos inquiridos consideram que a forma utilizada foi simples; 34,8% não a consideram simples nem complicada; 12,9% muito simples; 8,5% complicada e apenas 6% muito complicada. Sabe-se ainda que a opinião que os alunos têm em relação à forma como os conteúdos foram abordados é independente da escola que frequentam.

• Ao nível de significância de 5%, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram leccionados é, na maioria das escolas, moderado e positivo. Portanto, pode-se admitir que os alunos que estão mais satisfeitos com a forma como os conteúdos de Estatística foram abordados são também aqueles que têm uma opinião mais positiva em relação a este tema.

• A associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados é, na maioria das escolas, positiva e moderada, no entanto, chega a ser forte em algumas das escolas. Portanto, pode-se admitir que os alunos que estão mais satisfeitos com a forma como os conteúdos de Estatística foram abordados são também aqueles que têm uma opinião mais positiva em relação a estes conteúdos.

• Em relação aos materiais utilizados nas aulas de Estatística, aproximadamente, 28% dos inquiridos utilizaram máquina de calcular; 26,6% utilizaram régua; 21,9%

utilizaram compasso; 18,1% utilizaram transferidor; 3,8% utilizaram computador; 1,4% dos alunos não utilizaram nenhum material durante as aulas de Estatística.

• No que se refere aos conteúdos de Estatística, a maioria dos alunos refere ter sentido dificuldades em pelo menos um dos conteúdos abordados; 20% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos; 3,4% sentiram dificuldades em todos os conteúdos.

5.3.3. Comparação entre o 7.º e o 8.º Anos do Ensino Básico

Apesar de se supor que a proporção de alunos do 8.º ano que referem ter tido dificuldades nos conteúdos de Estatística ser superior à proporção de alunos que referem os mesmos conteúdos, mas que estão a frequentar o 7.º ano, não há, para nenhuma das escolas, evidência significativa para afirmar que a proporção de alunos do 8.º ano com dificuldades nos conteúdos de Estatística é superior à de alunos que referem os mesmos conteúdos, mas que estão a frequentar o 7.º ano.

5.3.4. 9.º Ano do Ensino Básico

• A amostra recolhida é constituída por 258 alunos, sendo 126 deles do sexo masculino e 132 do feminino.

• As idades dos alunos inquiridos variam entre os 14 e os 18 anos, inclusive.

• A maioria dos inquiridos (36,4%) afirma gostar da disciplina de Matemática. No entanto, 43% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e cerca de 30% de Insuficiente.

• Há uma associação positiva e moderada entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

• No 1.º Período, a classificação que mais se verificou foi 3 com 43,6%, seguida da classificação 2 com 37,7%. No 2.º Período, a classificação que mais se verificou foi 2 com 42,1%, seguida da classificação 3 com 41,3%. Apesar da maioria dos alunos do 9.º ano gostar de Matemática, o seu rendimento a esta disciplina não é equivalente à opinião que têm da mesma.

• A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte. No entanto, esta associação torna-se ligeiramente mais fraca no 2.º Período.

• Dos temas leccionados durante o ano lectivo, o tema de Probabilidades e Estatística foi o eleito pela maioria dos alunos (49,6%), seguido dos temas de Cálculo, de Geometria e de Funções com 30,2%, 12% e 7,8% de votos, respectivamente.

• Quanto ao tema de Probabilidades e Estatística, cerca de 77% consideram que o tema é interessante e, aproximadamente, 23% dos inquiridos não o consideram interessante.

• No que se refere aos conteúdos de Probabilidades e Estatística, cerca de 50% dos alunos não os consideram fáceis nem difíceis; 31% avaliam-nos como sendo fáceis; 9,3% como sendo muito difíceis; 8,9% como sendo muito fáceis e apenas 0,4% como sendo difíceis.

• No que diz respeito à forma como os professores abordaram os conteúdos de Probabilidades e Estatística, cerca de 43% dos inquiridos consideram que a forma utilizada nem foi simples nem foi complicada; 39,5% consideram-na simples; 8,9% muito simples; 5,4% complicada e apenas 2,7% muito complicada. Sabe-se ainda que a opinião que os alunos têm em relação à forma como os conteúdos foram abordados é independente da escola que frequentam.

• Na Escola 2,3 Navegador Rodrigues Soromenho e na Escola Secundária Gil Vicente há associação positiva e moderada entre a opinião que os alunos têm em relação ao tema de Probabilidades e Estatística e a forma como os conteúdos foram abordados. No entanto, na Escola 2,3 Mestre Domingos Saraiva e na Escola 2,3/S Michel Giacometti não há associação entre as variáveis em questão.

• Em todas as escolas, excepto na Escola 2,3/S Michel Giacometti, há associação positiva e moderada entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Estatística e a forma como os mesmos foram abordados nas aulas.

• No que se refere aos conteúdos de Probabilidades e Estatística, grande parte dos alunos refere ter sentido dificuldades em pelo menos um dos conteúdos abordados; 23,5% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos; 7,9% sentiram dificuldades em todos os conteúdos.

5.3.5. Comparação entre a Estatística do 7.º e 8.º Anos e as Probabilidades e Estatística do 9.ºAno

Apesar de se supor que alunos preferem o tema de Estatística, que é leccionado no 7.º e 8.º anos, ao tema de Probabilidades e Estatística, que é leccionado no 9.º ano, não há, em nenhuma das escolas, evidência significativa para afirmar que a proporção de alunos que prefere o tema de Estatística do 7.º e 8.º anos é superior à proporção de alunos que prefere o tema de Probabilidades e Estatística do 9.º ano.

5.3.6. 10.º Ano do Ensino Secundário

• A amostra recolhida é constituída por 129 alunos, sendo 70 deles do sexo masculino e 59 do feminino.

• As idades dos alunos inquiridos variam entre os 15 e os 18 anos, inclusive.

• A maioria dos inquiridos (80,6%) frequenta o Curso Científico-Humanístico de Ciências e Tecnologias, enquanto que 19,4% dos inquiridos frequentam o Curso Científico-
-Humanístico de Ciências Socioeconómicas.

• A maioria dos inquiridos (45%) afirma gostar da disciplina de Matemática. No entanto, aproximadamente, 41% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e 31% de Insuficiente.

• Há uma associação positiva e moderada entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

• No 1.º Período, o nível que mais se verificou foi o 10 com 15,7%, seguido dos níveis 8 e 7 com 12,6% e 11,8%, respectivamente. No 2.º Período, o nível mais frequente foi o 9 com 15%, seguido do 8 e do 10 com 13,4% e 10,2%, respectivamente. Apesar da maioria dos alunos do 10.º ano gostar de Matemática, o seu rendimento a esta disciplina não é equivalente à opinião que têm da mesma.

• A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte. Além disso, esta associação torna-se ainda mais forte no 2.º Período. O que significa que com o decorrer do ano lectivo os alunos vão conseguindo auto-avaliar-se correctamente.

• Dos temas leccionados durante o ano lectivo, o tema de Estatística foi o eleito pela maioria dos alunos (48,8%), seguido dos temas de Funções e de Geometria com 38% e 13,2% de votos, respectivamente.

• Quanto ao tema de Estatística, cerca de 77% consideram que o tema é interessante e, aproximadamente, 23% dos inquiridos não o consideram interessante.

• No que se refere aos conteúdos de Estatística, cerca de 40% dos alunos não os consideram fáceis nem difíceis; 38% avaliam-nos como sendo fáceis; 16,3% como sendo muito fáceis; 4,7% como sendo difíceis e apenas 0,8% como sendo muito difíceis. Sabe-se ainda que a opinião que os alunos têm em relação aos conteúdos de Estatística é independente da escola que frequentam.

• No que diz respeito à forma como os professores abordaram os conteúdos de Estatística, cerca de 40% dos inquiridos consideram que a forma utilizada nem foi simples nem foi complicada; 31,8% consideram-na simples; 12,4% muito simples; 10,9% complicada e apenas 4,7% muito complicada. Sabe-se ainda que a opinião que os alunos têm em relação à forma como os conteúdos foram abordados depende da escola que frequentam.

• Em nenhuma das escolas há associação entre a opinião que os alunos têm em relação ao tema de Estatística e a forma como os conteúdos foram abordados.

• Na Escola 2,3/S Michel Giacometti e na Escola Secundária da Baixa da Banheira há uma associação positiva e moderada entre a opinião que os alunos têm em relação aos conteúdos de Estatística e a forma como estes foram leccionados. No entanto, na

Escola Secundária Ferreira Dias e na Escola Secundária Fernão Mendes Pinto não há associação entre as variáveis em questão.

• Aproximadamente, 8% dos inquiridos afirmam que não tiveram aulas de Estatística no 3.º Ciclo do Ensino Básico, 16,3% não se recordam e apenas 76% tiveram aulas de Estatística no 3.º Ciclo do Ensino Básico.

• Em relação aos materiais utilizados nas aulas de Estatística, aproximadamente, 51% dos inquiridos utilizaram calculadora gráfica; 16,7% utilizaram régua; 10,4% utilizaram calculadora científica; 7,5% utilizaram computador; 7,5% utilizaram transferidor; 6,7% utilizaram compasso; 0,4% dos alunos não utilizaram nenhum material durante as aulas de Estatística.

• A maioria dos inquiridos (38,7%) afirma ter utilizado a calculadora gráfica, durante as aulas de Estatística, com a finalidade de traçar gráficos e 30% dos inquiridos afirmam ter utilizado a calculadora para determinar o valor das estatísticas. No entanto, nem todos os alunos procuram tirar o maior proveito da calculadora gráfica, 13,5% dos inquiridos afirmam ter utilizado a calculadora para efectuar cálculos elementares e aproximadamente 2% dos inquiridos revelam ter utilizado a calculadora para consultar "auxiliares de memória".

• No que se refere aos conteúdos de Estatística, a maioria dos alunos refere ter sentido dificuldades em pelo menos um dos conteúdos abordados; 27,6% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos; 4,3% sentiram dificuldades em todos os conteúdos.

5.3.7. 12.º Ano do Ensino Secundário

• A amostra recolhida é constituída por 77 alunos, sendo 35 deles do sexo masculino e 42 do feminino.

• As idades dos alunos inquiridos variam entre os 17 e os 21 anos, inclusive.

• A totalidade dos alunos inquiridos frequenta o Curso Científico-Humanístico de Ciências e Tecnologias.

• A maioria dos inquiridos (58,4%) afirma gostar da disciplina de Matemática. No entanto, aproximadamente, 50% dos inquiridos auto-avaliam o seu rendimento a esta disciplina de Suficiente e apenas 33,8% de Bom.

• Há uma associação positiva e moderada entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

• No 1.º Período, o nível que mais se verificou foi o 13 com 16%, seguido dos níveis 12 e 14 com 13,3% e 12%, respectivamente. No 2.º Período, os níveis mais frequentes foram o 10 e o 12 com 13,3% cada um, seguidos dos níveis 11 e 13 com 12% cada um.

• A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte. Além disso, esta associação torna-se ainda mais forte no 2.º Período. O que significa que com o decorrer do ano lectivo os alunos vão conseguindo auto-avaliar-se correctamente.

• Quantos aos temas leccionados durante o ano lectivo na disciplina de Matemática, aproximadamente, 40% dos inquiridos seleccionam a temática de Trigonometria e Números Complexos como sendo a sua preferida, 31,2% dos inquiridos preferem a temática de Probabilidades e Combinatória e finalmente, surge o Cálculo Diferencial com 28,6% de votos.

• Em relação à temática de Probabilidades e Combinatória, a maioria dos inquiridos (67,5%) considera esta temática interessante. No entanto, os restantes alunos (32,5%) consideram que o tema não é interessante. Sabe-se ainda que a opinião que os alunos têm em relação ao tema de Probabilidades e Combinatória depende da escola que frequentam.

• No que se refere aos conteúdos de Probabilidades e Combinatória, cerca de 49% dos alunos não os consideram fáceis nem difíceis; 23,4% avaliam-nos como sendo fáceis; 23,4% como sendo difíceis; 2,6% como sendo muito fáceis e apenas 1,3% como sendo muito difíceis. Sabe-se ainda que a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Combinatória é independente da escola que frequentam.

• No que diz respeito à forma como os professores abordaram os conteúdos de Probabilidades e Combinatória, cerca de 47% dos inquiridos consideram que a forma utilizada foi simples; 32,5% consideram-na muito complicada; 9,1% nem simples nem complicada; 7,8% muito simples e cerca de 3% complicada. Sabe-se ainda que a opinião que os alunos têm em relação à forma como os conteúdos foram abordados é independente da escola que frequentam.

• Na Escola Básica 2,3/S Michel Giacometti, o grau de associação que há entre a opinião que os alunos têm em relação ao tema de Probabilidades e Combinatória e a forma como os conteúdos foram leccionados é positivo e moderado. No entanto, na Escola Secundária Gil Vicente e no Colégio São João de Brito não há associação entre as variáveis em questão.

• Na Escola Básica 2,3/S Michel Giacometti, o grau de associação que há entre a opinião que os alunos têm em relação aos conteúdos de Probabilidades e Combinatória e a forma como estes foram abordados é positivo e moderado. No entanto, na Escola Secundária Gil Vicente e no Colégio São João de Brito não há associação linear entre as variáveis em questão.

• Aproximadamente, 7% dos inquiridos afirmam que não tiveram aulas de Probabilidades e Estatística no 3.º Ciclo do Ensino Básico, 18,2% não se recordam e apenas 72,7% tiveram aulas de Probabilidades e Estatística no 3.º Ciclo do Ensino Básico.

• Em relação aos materiais utilizados nas aulas de Probabilidades e Combinatória, a generalidade dos inquiridos (93,5%) utilizou calculadora gráfica. No entanto, 5,2% dos inquiridos utilizaram calculadora científica e 1,3% dos inquiridos não utilizaram qualquer tipo de material. Apesar do programa de Matemática do 12.º ano sugerir a utilização de calculadoras gráficas nas aulas de Probabilidades e Combinatória, nem todos os alunos as utilizam. Na realidade, muitos alunos não adquirem calculadoras gráficas e as escolas não as têm em número suficiente para suprir esta falta.

• A maioria dos inquiridos (51,5%) utilizou a calculadora gráfica, durante as aulas de Probabilidades e Combinatória, com o objectivo de efectuar os cálculos das combinações e dos arranjos. No entanto, nem todos os alunos procuraram tirar o maior proveito das calculadoras gráficas, aproximadamente, 22% dos inquiridos afirmaram ter utilizado a calculadora para efectuar cálculos elementares e aproximadamente 5% dos inquiridos revelaram ter utilizado a calculadora para consultar “auxiliares de memória”.

- No que se refere aos conteúdos de Probabilidades e Combinatória, a maioria dos alunos refere ter sentido dificuldades em pelo menos um dos conteúdos abordados; 1,8% dos inquiridos afirmam não ter sentido dificuldades em nenhum dos conteúdos; 0,6% sentiram dificuldades em todos os conteúdos.

5.3.8. Comparação entre a Estatística do 10.º Ano e as Probabilidades e Combinatória do 12.º Ano

Na Escola Básica 2,3/S Michel Giacometti há evidência significativa para afirmar que a proporção de alunos que prefere o tema de Estatística do 10.º ano é superior à proporção de alunos que prefere o tema de Probabilidades e Combinatória do 12.º ano.

CAPÍTULO VI

CONCLUSÃO

Com este estudo pretendeu-se analisar as componentes de Estatística e de Probabilidades nos programas de Matemática A e B do Ensino Secundário. Uma vez que estes programas sugerem as brochuras de *Estatística* e de *Probabilidades e Combinatória* como guias de apoio ao professor na leccionação destas temáticas, a presente análise foi também alargada a estes guias de apoio.

A análise efectuada mostra que todos estes documentos apresentam lacunas, que em alguns casos não chegam a ser muito graves, mas que poderiam ser eliminadas.

Existem diversas investigações que revelam que a maioria dos professores de Matemática possui uma reduzida formação em Estatística e em Probabilidades (Almeida, 2002; Batanero, 2000; Branco, 2000a; Branco, 2000b; Branco, 2000c; Carvalho, 2001; Ribeiro, 2005). Como tal, estes docentes ao consultarem os programas de Matemática ou os guias de apoio ao professor podem ser facilmente induzidos em erro devido às lacunas que estes documentos apresentam e posteriormente, conduzir em erro os alunos durante a leccionação dos conteúdos. De facto, são muitos os estudos que evidenciam que os alunos, dos diferentes níveis de ensino, revelam dificuldades nos mais diversos conceitos de Estatística e de Probabilidades, incluindo os mais elementares (Batanero, 2000; Batanero *et al.*, 1994; Barros e Fernandes, 2005; Boaventura, 2003; Cai, 1995; Carvalho, 1996; Carvalho e César, 2000; Dreyfus e Levy, 1996; Fernandes, 1999; Fischbein *et al.*, 1991; Green, 1983; Li e Schen, 1994; Sousa, 2002).

Portanto, é importante que se efectue uma revisão curricular das temáticas de Estatística e de Probabilidades nos programas de Matemática A e B do Ensino Secundário, bem como uma revisão nas brochuras de *Estatística* e de *Probabilidades e Combinatória*.

Igualmente importante será o Ministério da Educação ou as instituições de ensino superior criarem programas de formação contínua especialmente direccionada para os professores com pouca formação na área da Estatística e das Probabilidades.

Este estudo, para além de pretender analisar as componentes de Estatística e de Probabilidades nos programas de Matemática A e B do Ensino Secundário, pretendeu também analisar os dados resultantes dos inquéritos aplicados aos alunos do 3.º Ciclo e do Ensino Secundário com o objectivo de analisar a opinião destes em relação aos programas de Matemática, aos conteúdos de Probabilidades e de Estatística, bem como à forma como estes foram abordados nas aulas.

Como nunca tinha sido desenvolvido um estudo desta natureza, a nível de inquéritos aplicados aos alunos do 3.º Ciclo e do Ensino Secundário, o presente estudo assume um carácter muito importante, pois trata-se de um estudo pioneiro que revelou resultados bastante interessantes.

Da análise dos inquéritos concluiu-se que, apesar do senso comum considerar que a maioria dos alunos não gosta de Matemática, a verdade é que a maioria inquiridos afirma gostar da disciplina de Matemática. Além disso, a maioria também avalia o seu rendimento a Matemática de Suficiente.

A associação que há entre a opinião que os alunos têm da disciplina de Matemática e a forma como avaliam o seu rendimento nesta disciplina é, para todos os anos de escolaridade, positiva e moderada. Portanto, pode-se admitir que os alunos que avaliam o seu rendimento a Matemática de forma negativa são também aqueles que têm uma opinião mais negativa da disciplina.

No final do 1.º Período, os níveis que mais se verificaram no 3.º Ciclo à disciplina de Matemática foram os níveis 2 e 3; no 10.º ano, foram o 7, 8 e 10; no 12.º ano, foram o 12, 13 e 14. No final do 2.º Período, os níveis que mais se verificaram no 3.º Ciclo continuaram a ser o 2 e o 3; no 10.º ano, passaram a ser o 8, 9 e 10; no 12.º ano, passaram a ser o 10, 11, 12 e 13. Os níveis do 12.º ano são ligeiramente superiores aos dos restantes anos de escolaridade, talvez pelo facto do 12.º ano ser o ano de transição para a universidade ou talvez pelo facto da disciplina de Matemática ser uma das disciplinas especificas para entrada de muitos cursos universitários.

A associação que há entre a forma como os alunos avaliam o seu rendimento a Matemática e a classificação que obtiveram nesta disciplina, quer no 1.º quer no 2.º Períodos, é positiva e forte para todos os anos de escolaridade. Além disso, essa associação tende, de um modo geral, a ser ainda mais forte no 2.º Período. O que significa que com o decorrer do ano lectivo os alunos vão conseguindo auto-avaliar-se correctamente.

No que diz respeito à opinião que os alunos têm em relação aos temas de Probabilidades e de Estatística, concluiu-se que a maioria dos alunos do 3.º Ciclo e do Ensino Secundário avalia estas unidades como sendo interessantes. Além disso, ao nível de significância de 1%, concluiu-se que a opinião que os alunos do 3.º Ciclo e do 10.º ano têm em relação a estas temáticas é independente da escola que frequentam. No entanto, a opinião que os alunos do 12.º ano têm em relação à temática de Probabilidades e Combinatória já não é independente da escola que frequentam. Uma vez que estes resultados são pouco conclusivos, seria importante que no futuro este estudo fosse alargado a mais escolas e a mais alunos de forma a obter-se resultados conclusivos.

Concluiu-se ainda que, de acordo com o senso comum, os alunos preferem o tema de Estatística do 10.º ano ao tema de Probabilidades e Combinatória do 12.º ano.

No que se refere à opinião que os alunos têm em relação aos conteúdos de Estatística e de Probabilidades, concluiu-se que a maioria dos alunos do 3.º Ciclo e do Ensino Secundário nem os considera fáceis nem difíceis. Além disso, concluiu-se que, ao nível de significância de 1%, a opinião que os alunos do 8.º, 10.º e 12.º anos têm em relação a estes conteúdos é independente da escola que frequentam. Em relação ao 7.º e ao 9.º ano, os testes deram resultados inconclusivos, donde se conclui que o estudo deveria ser alargado a mais escolas e a mais alunos.

Em relação à forma como os docentes abordaram os conteúdos de Estatística e de Probabilidades, concluiu-se que a maioria dos alunos avalia a forma utilizada pelos docentes em uma das duas categorias: simples ou nem simples nem complicada. Concluiu-se ainda que, ao nível de significância de 1%, a opinião que os alunos do 3.º Ciclo e do Ensino Secundário têm em relação à forma como os conteúdos foram abordados é independente da escola que frequentam, excepto no caso dos alunos do 7.º e 10.º anos em que a opinião destes está associada à escola. Uma vez que estes resultados são pouco conclusivos, seria importante que no futuro este estudo fosse alargado a mais escolas e a mais alunos de forma a obter-se resultados conclusivos.

A associação que há entre a opinião que os alunos têm em relação aos temas de Estatística e de Probabilidades e a opinião que têm em relação à forma como os conteúdos foram abordados é positiva e moderada para a maioria das escolas do 7.º e 8.º anos, para a Escola 2,3 Navegador Rodrigues Soromenho no 9.º ano, para a Escola Secundária Gil Vicente no 9.º ano e finalmente, para a Escola 2,3/S Michel Giacometti no 12.º ano. Nas restantes escolas não há associação entre as variáveis em questão. Portanto, para as escolas em que há associação pode-se admitir que os alunos que estão satisfeitos com a forma como os docentes abordaram os conteúdos de Estatística e de Probabilidades são também aqueles que têm uma opinião mais satisfatória em relação a estes temas.

A associação que há entre a opinião que os alunos têm em relação aos conteúdos de Estatística e de Probabilidades e a opinião que têm em relação à forma como esses conteúdos foram abordados é positiva e forte na maioria das escolas do 7.º ano e em algumas das escolas do 8.º ano. Esta associação é, também, positiva e moderada para algumas das escolas do 8.º ano, para a maioria das escolas do 9.º ano, para a Escola Secundária Baixa da Banheira no 10.º ano e finalmente, para a Escola 2,3/S Michel Giacometti no 10.º e 12.º anos. Nas restantes escolas não há associação entre as variáveis em questão. Portanto, para as escolas em que há associação pode-se admitir que os alunos que estão satisfeitos com a forma como os docentes abordaram os conteúdos de Estatística e de Probabilidades são também aqueles que têm uma opinião mais satisfatória em relação a estes conteúdos.

A partir da análise dos inquéritos concluiu-se que alguns dos alunos que se encontram no Ensino Secundário não tiveram aulas de Estatística nem de Probabilidades no 3.º Ciclo do Ensino Básico. Inclusivamente, verificou-se que 31% dos alunos inquiridos do 7.º ano ainda não tinham abordado o tema de Estatística quando responderam ao inquérito. Portanto, pode-se concluir que as temáticas de Probabilidades e de Estatística nem sempre são leccionadas, uma vez que são, geralmente, remetidas para o final do ano lectivo e por isso mesmo, nem sempre são apresentadas aos alunos.

No que se refere aos materiais utilizados, concluiu-se que apesar dos programas de Matemática do 3.º Ciclo e do Ensino Secundário sugerirem a utilização de computadores nas aulas, estes foram pouco utilizados quer pelos alunos do 3.º Ciclo quer pelos alunos do Ensino Secundário. De facto, são poucos os professores que planificam as suas aulas com vista à utilização desta tecnologia. Esta situação deve--se, essencialmente, ao facto de as escolas não terem uma sala de trabalho com computadores à disposição dos alunos e ao facto da maioria das aulas de Matemática não decorrer em salas com computadores.

Para além dos computadores, os programas de Matemática do Ensino Secundário também sugerem a utilização de calculadoras gráficas. No entanto, pela análise dos inquéritos, concluiu-se que apenas 50,8% dos alunos do 10.º ano utilizaram calculadora gráfica durante as aulas de Estatística e que aproximadamente, 7% dos alunos do 12.º ano não utilizaram esta tecnologia durante as aulas de Probabilidades e Combinatória. De facto, existem muitos alunos que nunca adquirem calculadoras gráficas e as escolas também não as têm em número suficiente para suprir esta falta.

Ainda relativamente às calculadoras gráficas, verificou-se que, dos alunos que tiveram acesso às calculadoras gráficas, nem todos procuraram desfrutar das potencialidades que a máquina lhes oferece. Alguns alunos, quer do 10.º quer do 12.º ano, afirmam ter utilizado a calculadora gráfica apenas para efectuar cálculos elementares ou, simplesmente, para consultar "auxiliares de memória".

No que se refere aos conteúdos de Estatística e de Probabilidades, a maioria dos alunos do 3.º Ciclo e do Ensino Secundário afirma ter sentido dificuldades em pelo menos um dos conteúdos abordados durante as aulas. De facto, tal como já foi referido anteriormente, existem diversos estudos efectuados que revelam que a maioria dos alunos do 3.º Ciclo e do Ensino Secundário apresenta dificuldades nos mais diversos conceitos de Estatística e de Probabilidades, incluindo os mais elementares. É importante realçar que alguns dos inquiridos, quer do 3.º Ciclo quer do Ensino Secundário, afirmam sentir dificuldades em todos os conteúdos de Estatística e de Probabilidades. Como tal, as comunidades escolares devem ter em atenção estes alunos, pois poderão estar a necessitar de algum tipo de apoio ou de acompanhamento.

Os resultados apresentados neste trabalho não poderão, obviamente, generalizar-se à população de estudantes do 3.º Ciclo e do Ensino Secundário em Portugal, uma vez que a amostra obtida é pouco representativa. No entanto, não se pode deixar de realçar a identificação de algumas pistas para estudos posteriores que poderão contribuir para o aprofundamento da compreensão do ensino e aprendizagem das temáticas de Estatística e de Probabilidades.

BIBLIOGRAFIA

Almeida, M. R. (2002). *Imagens sobre o ensino e a aprendizagem da Estatística*. Lisboa: Instituto de Inovação Educacional.

Barros, P. M. e Fernandes, J. A. (2005). Dificuldades em estocastica de uma futura professora do 1º e 2ºciclos do Ensino Básico. *Revista Portuguesa de Educação, 18(1)*. pp.117-150. Universidade do Minho. Braga.

Batanero, C., Godino, J. D., Green, D. R., Holmes, P. e Vallecillos, A. (1994). Errors and difficulties in understanding statistical concepts. *International Journal of Mathematics Education in Science and Technology. 25*(4). pp. 527-547.

Batanero, C. (2000). Dificultades de los estudiantes en los conceptos estadísticos elementares: el caso de las medidas de posición central. In C. Loureiro, F. Oliveira e L. Brunheira (orgs.), *Ensino e Aprendizagem da Estatística*. Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamentos de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa, pp. 31-48.

Bishop, Y. M. M., Fienberg, S. E. e Holland, P. W. (1975). *Discrete Multivariate Analysis: Theory and practice.* Cambridge, MA: MIT Press.

Boaventura, M. G. (2003). *Dificuldades de Alunos do Ensino Secundário em Conceitos Estatísticos*. Dissertação de mestrado não publicada. Universidade do Minho. Braga.

Branco, J. (2000a). Estatística no secundário: O ensino e seus problemas. In C. Loureiro, F. Oliveira e L. Brunheira (Orgs.). *Ensino e Aprendizagem da Estatística.* Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamentos de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa. pp. 11-30.

Branco, J. (2000b). Estatística no secundário: O ensino e os seus problemas. In S. M. Marques (Ed.). *Jornal de Matemática Elementar.* 190. pp. 10-18.

Branco, J. (2000c). Estatística no secundário: O ensino e os seus problemas. In S. M. Marques (Ed.). *Jornal de Matemática Elementar.* 191. pp. 10-17.

Cai, J. (1995). Beyond the computational algorithm: student's understanding of the arithmetic average concept. In L. Meira e D. Carraher (eds.). *Proceedings of the 19th PME Conference.* Vol. 3. Universidade Federal de Pernambuco. pp.144-151.

Carvalho, C. (1996). Algumas questões em torno de tarefas estatísticas com alunos do 7º ano. In A. Roque e M. J. Lagarto (orgs.), *Actas do ProfMat 96*. Lisboa: Associação de Professores de Matemática, pp. 165-171.

Carvalho, C. e César, M. (2000). As aparências iludem: reflexões em torno do ensino da estatística no ensino básico. In C. Loureiro, F. Oliveira e L. Brunheira (orgs.). *Ensino e Aprendizagem da Estatística*. Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamentos de Educação e de Estatística e Investigação operacional da Faculdade de Ciências da Universidade de Lisboa. pp. 212-225.

Carvalho, C. (2001). *Interacções entre pares: Contributos para a promoção do desenvolvimento lógico e do desempenho estatístico no 7º ano de escolaridade*. Lisboa: Associação de Professores de Matemática.

Chisnall, P. (2005). *Marketing Research.* 7th Edition. Mc Graw-Hill. New York.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics.* Vol. 10. pp. 417-451.

Costa, B. e Rodrigues, E. (2004). *Espaço B: Ensino Secundário 10.º ou 11.ºAno. 2ª*Edição. Porto: Edições ASA.

Costa, B. e Rodrigues, E. (2006). *Espaço B: Ensino Secundário 12.ºano.* 1ªEdição. Porto: Edições ASA.

Daniel, W. W. (2005). *Biostatistics: A Foundation for Analysis in the Health Sciences.* 8th Edition. Wiley. New York.

Dreyfus, A. e Levy, O. (1996). Are the notion of mean and related concepts too difficult for 6th and 7th grade biology students? *European Journal of Teacher Education*. 19(2). pp. 137-152.

Eberhardt, K. R. e Fligner, M. A. (1977). A comparison of two tests for equality of proportions. *Amer. Statist.* 31. pp.151-155.

Everitt, B. S. (1992). *The Analysis of Contingency tables*. Chapman and Hall. London.

Fernandes, J. A. (1999). *Intuições e Aprendizagem de Probabilidades: Uma Proposta de Ensino de Probabilidades no 9º Ano de Escolaridade*. Tese de doutoramento não publicada, Universidade do Minho, Braga.

Ficheiro da OCDE em http://www.oecd.org/els/social/family/database

Fischbein, E., Nello, M. S. e Marino, M. S. (1991). Factors affecting probabilistic judgments in children and adolescents. *Educational Studies in Mathematics*. 22. pp. 523-549.

Freeman, G. H. e Halton, J. H. (1951). Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika.* Vol. 38. pp.141-149.

Green, D. R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D. R. Grey, P. Holmes, V. Barnett e G. M. Constable (eds.). *Proceedings of the First International Conference on Teaching Statistics.* Vol. 2. Sheffield, UK: Teaching Statistics Trust. pp. 766-783.

Jorge, A. M. B., Alves, C. B., Fonseca, G. e Barbedo, J. (2006). *Infinito 12 B*. Porto: Areal Editores.

Kendall, M. G. (1945). The treatment of ties in ranking problems. *Biometrika*. Vol. 33. pp. 239-251.

Kendall, M. G. (1970). *Rank Correlation Methods*. 4th Edition. Charles Griffin & Co. London. England.

Lewontin, R. C. e Felsenstein, J. (1965). The robustness of homogeneity tests in 2 x N tables. *Biometrics.* Vol. 21. pp. 19-33.

Li, K. e Shen, S. (1994). Students' weaknesses in statistical projects. In D. Green (Ed.). *Teaching Statistics at its Best*. Sheffield, UK: Teaching Statistics Trust. pp.42-48.

Lightner, J. E. (1991). A brief look at history of probability and statistics. *The Mathematics Teacher, 84*(8), pp. 623-630.

Maroco, J. e Bispo, R. (2005). *Estatística aplicada às ciências sociais e humanas.* 2ª Edição. Lisboa: Climepsi Editores.

Martins, M. E. G. (coord.), Monteiro, C., Viana, J. P. e Turkman, M. A. A. (1997). *Estatística: matemática – 10.º ano de escolaridade.* Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Martins, M. E. G. (coord.), Monteiro, C., Viana, J. P. e Turkman, M. A. A. (1999). *Probabilidades e Combinatória: matemática – 12.º ano de escolaridade.* Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Martins, M. E. G. (2000). *Introdução às Probabilidades e à Estatística*. 2ªEdição. Lisboa: Sociedade Portuguesa de Estatística, Departamento de Estatística e Investigação Operacional da FCUL.

Meddis, R. (1984). *Statistics Using Ranks: A Unified Approach.* Basil Blackwell, Oxford, England. pp. 449.

Mello, F. G. (1993). *Probabilidades e Estatística, Conceitos e Métodos Fundamentais*. Lisboa: Escolar Editora.

Mello, F. G. (1997). *Probabilidades e Estatística, Conceitos e Métodos Fundamentais.* (Vol. II). Lisboa: Escolar Editora.

Ministério da Educação (1991). *Programa de Matemática: Plano de Organização do Ensino-Aprendizagem (3.º Ciclo do Ensino Básico).* Vol. II. Lisboa: Imprensa Nacional Casa da Moeda.

Moore, D. (2000). *A Estatística Básica e sua prática*. Rio de Janeiro: Ed. LTC.

Neves, M. A. F., Silva, M. C., Guerreiro, L. e Pereira, A. (2006). *Matemática B 12.ºano*. Porto: Porto Editora.

Pestana, D. D. e Velosa, S. F. (2006). *Introdução à Probabilidade e à Estatística.* (Vol. I). 2ª Edição. Lisboa: Fundação Calouste Gulbenkian.

Ponte, J. P. e Fonseca, H. (2001). *Orientações curriculares para o ensino da Estatística: Análise comparativa de três países.* Quadrante, 10(1), pp. 93-132.

Reis, E. (1996). *Estatística Descritiva*. 3ª Edição. Lisboa: Edições Sílabo.

Ribeiro, S. A. L. (2005). *O Ensino da Estatística no 7ºano de Escolaridade. Caracterização e dificuldades sentidas pelos professores.* Dissertação de Mestrado. Universidade do Minho.

Scheaffer, R. (2000). Statistics for a new century. In M. J. Burke & F. R. Curcio (Orgs.), *Learning mathematics for a new century* (pp. 158-173). Reston, VA: National Council of Teachers of Mathematics.

Schweigert, W. A. (1994). *Research Methods and Statistics for Psychology*. Brooks/Cole Publi. Co.. Pacific Grove. California.

Silva, J. C. (coord.), Fonseca, M. G., Martins, A. A., Fonseca, C. M. C. e Lopes, I. M. C. (2001a). *Matemática A – 10º ano*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Silva, J. C. (coord.), Fonseca, M. G., Martins, A. A., Fonseca, C. M. C. e Lopes, I. M. C. (2001b). *Matemática B – 10º ou 11º anos*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Silva, J. C. (coord.), Fonseca, M. G., Martins, A. A., Fonseca, C. M. C. e Lopes, I. M. C. (2002a). *Matemática B – 11º ou 12º anos*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Silva, J. C. (coord.), Fonseca, M. G., Martins, A. A., Fonseca, C. M. C. e Lopes, I. M. C. (2002b). *Matemática A – 11º ano*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Silva, J. C. (coord.), Fonseca, M. G., Martins, A. A., Fonseca, C. M. C. e Lopes, I. M. C. (2002c). *Matemática A – 12º ano*. Lisboa: Ministério da Educação, Departamento do Ensino Secundário.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*. Vol.15. pp. 72-101

Sprent, P. (1993). *Applied Nonparametric Statistical Methods.* 2nd Edition. Chapman & Hall. London.

Sprent, P. e Smeeton, N. C. (2001). Applied Nonparametric Statistical Methods. 3rd Edition. Chapamn and Hall/CRC. Boca Raton.

Sousa, O. (2002). Investigações estatísticas no 6.º ano. In GTI - Grupo de Trabalho de Investigação (Org.), *Reflectir e Investigar sobre a Prática Profissional.* Lisboa: Associação de Professores de Matemática. pp. 75-97.

Thomas, G. E. (1989). A note on correcting for ties with Spearman's ρ. *J. Statist. Computa. Simula.* 31. pp. 37 – 40.

Zar, J. H. (1999). *Biostatistical Analysis.* 4th edition. Prentice Hall, Upper Saddle River, New Jersey.

ANEXO I

(Autorização do Director Geral de Inovação e Desenvolvimento Curricular para aplicação de inquéritos em meio escolar)

FW: Inquéritos

De: **Luís Capucha (DGIDC)** (luis.capucha@dgidc.min-edu.pt)
Enviada: sexta-feira, 27 de junho de 2008 17:24:33
Para: caldeira_sara@hotmail.com
Anexos: image001.jpg (1,6 KB), Cristina Caldeira.tif (333,8 KB)

Exma. Senhora
Dra. Sara Cristina Caldeira,

Venho por este meio informar que o pedido de aplicação do inquérito em meio escolar, no âmbito da realização de uma Tese de Mestrado em Probabilidade e Estatística, é autorizado. Gostaria, no entanto, de sugerir que colocasse uma breve introdução no início dos questionários que explicitasse o âmbito do estudo.

Com os melhores cumprimentos e votos de um bom trabalho

Luís Capucha
(Director-Geral)

Ana do Vale

---

dgidc
Direcção-Geral de Inovação e Desenvolvimento Curricular
Av. 24 de Julho, 140
1399-026 Lisboa
Tel: 21 393 4630

A N E X O I I

(Requerimento aos Conselhos Executivos)

Lisboa, 20 de Maio de 2008

Exmo. Sr. ou Sr.ª:
Presidente do Conselho Executivo

Eu, Sara Cristina Baião Caldeira, professora de Matemática do 3.º Ciclo e do Ensino Secundário, do grupo de recrutamento 500, aluna do Curso de Mestrado em Probabilidades e Estatística, da Faculdade de Ciências da Universidade de Lisboa, venho por este meio, solicitar a sua autorização para fazer um inquérito aos alunos do 3.º Ciclo do Ensino Básico e aos alunos do Ensino Secundário sobre os temas de Estatística e de Probabilidades, no âmbito de uma investigação individual que culminará com a Dissertação de Mestrado.

A Dissertação em causa visa analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Probabilidades e de Estatística, assim como à forma como estes foram abordados nas aulas.

Tal como defendem vários investigadores, a grande importância que os temas de Estatística e de Probabilidades assumem, hoje em dia, na educação matemática, resultante do facto de se tratarem de áreas com uma enorme expressão na actividade social e em muitos domínios do conhecimento, por si só justifica que se investigue e se reflicta sobre a forma como estes temas estão a ser apresentados nas escolas.

Fico à inteira disposição de V. Ex.ª para esclarecer toda a informação que julgue oportuna.

Agradecendo desde já a sua colaboração, subscrevo-me com os melhores cumprimentos,

Atenciosamente

(Sara Caldeira)

ANEXO III

(Inquérito aplicado aos alunos do 7.º ano, do 8.º ano, do 9.º ano,
do 10.º e 12.º ano dos Cursos Científico-Humanísticos e
do 10.º e 11.º ano dos Cursos Cientifico-Humanisticos de Artes Visuais)

Inquérito para os Alunos do 7.º ano

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

4. Como avalia o seu rendimento a Matemática?

○ Muito Bom (Nível 5)
○ Bom (Nível 4)
○ Suficiente (Nível 3)
○ Insuficiente (Nível 2)
○ Mau (Nível 1)

5. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: ___________ 2.º Período: __________

6. Qual o tema de Matemática que mais gosta?

○ Geometria ○ Funções ○ Cálculo ○ Estatística

7. Considera que o tema de Estatística:

○ É interessante ○ Não é interessante

8. Como avalia os conteúdos de Estatística?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

9. Na sua opinião, os conteúdos de Estatística são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

V.S.F.F.

10. Este ano, que materiais utilizou durante as aulas de Estatística?

- ○ Computador
- ○ Calculadora
- ○ Régua
- ○ Compasso
- ○ Transferidor
- ○ Outro. Qual? ____

11. Nas aulas de Estatística foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____
____

MUITO OBRIGADO!

Inquérito para os Alunos do 8.º ano

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

4. Como avalia o seu rendimento a Matemática?

○ Muito Bom (Nível 5)
○ Bom (Nível 4)
○ Suficiente (Nível 3)
○ Insuficiente (Nível 2)
○ Mau (Nível 1)

5. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: ___________ 2.º Período: _________

6. Qual o tema de Matemática que mais gosta?

○ Geometria ○ Funções ○ Cálculo ○ Estatística

7. Considera que o tema de Estatística:

○ É interessante ○ Não é interessante

8. Como avalia os conteúdos de Estatística?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

9. Na sua opinião, os conteúdos de Estatística são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

V.S.F.F.

10. Este ano, que materiais utilizou durante as aulas de Estatística?

- ○ Computador
- ○ Calculadora
- ○ Régua
- ○ Compasso
- ○ Transferidor
- ○ Outro. Qual? ____

11. Nas aulas de Estatística foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____
____

MUITO OBRIGADO!

Inquérito para os Alunos do 9.º ano

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Probabilidades e Estatística, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

4. Como avalia o seu rendimento a Matemática?

○ Muito Bom (Nível 5)
○ Bom (Nível 4)
○ Suficiente (Nível 3)
○ Insuficiente (Nível 2)
○ Mau (Nível 1)

5. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: ___________ 2.º Período: __________

6. Qual o tema de Matemática que mais gosta?

○ Geometria ○ Funções ○ Cálculo ○ Probabilidades e Estatística

7. Considera que o tema de Probabilidades e Estatística:

○ É interessante ○ Não é interessante

8. Como avalia os conteúdos de Probabilidades e Estatística?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

9. Na sua opinião, os conteúdos de Probabilidades e Estatística são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

V.S.F.F.

10. Nas aulas de Probabilidades e Estatística foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

MUITO OBRIGADO!

Inquérito para os Alunos do 10.º ano
de Cursos Científico-Humanísticos

[Matemática A]

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual o curso que frequenta?

Curso Científico-Humanístico de ______________________________________

4. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

5. Como avalia o seu rendimento a Matemática?

○ Muito Bom (18 – 20 valores)
○ Bom (14 – 17 valores)
○ Suficiente (10 – 13 valores)
○ Insuficiente (5 – 9 valores)
○ Mau (0 – 4 valores)

6. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: __________ 2.º Período: _________

7. Qual o tema de Matemática que mais gosta?

○ Geometria ○ Funções ○ Estatística

8. Considera que o tema de Estatística:

○ É interessante ○ Não é interessante

V.S.F.F.

9. Como avalia os conteúdos de Estatística?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

10. Na sua opinião, os conteúdos de Estatística são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

11. No 3.º Ciclo do Ensino Básico, teve aulas de Estatística?

○ Sim ○ Não ○ Não me lembro

12. Este ano, que materiais utilizou durante as aulas de Estatística?

○ Computador
○ Calculadora científica
○ Calculadora gráfica
○ Régua
○ Compasso
○ Transferidor
○ Outro. Qual? ____

13. Se na pergunta anterior indicou a calculadora, indique com que finalidade a utilizou.

____
____
____
____

14. Nas aulas de Estatística foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____
____

MUITO OBRIGADO!

Inquérito para os Alunos do 12.º ano

de Cursos Científico-Humanísticos

[Matemática A]

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Probabilidades e Combinatória, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual o curso que frequenta?

Curso Científico-Humanístico de ______________________________________

4. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

5. Como avalia o seu rendimento a Matemática?

○ Muito Bom (18 – 20 valores)
○ Bom (14 – 17 valores)
○ Suficiente (10 – 13 valores)
○ Insuficiente (5 – 9 valores)
○ Mau (0 – 4 valores)

6. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: __________ 2.º Período: _________

7. Qual o tema de Matemática que mais gosta?

○ Probabilidades e Combinatória
○ Trigonometria e Números Complexos
○ Cálculo Diferencial

8. Considera que os conteúdos de Probabilidades e Combinatória:

○ São interessantes ○ Não são interessantes

V.S.F.F.

9. Como avalia os conteúdos de Probabilidades e Combinatória?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

10. Na sua opinião, os conteúdos de Probabilidades e Combinatória são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

11. No 3.º Ciclo do Ensino Básico, teve aulas de Probabilidades e Estatística?

○ Sim ○ Não ○ Não me lembro

12. Este ano, que materiais utilizou durante as aulas de Probabilidades e Combinatória?

○ Computador
○ Calculadora científica
○ Calculadora gráfica
○ Outro. Qual? ____

13. Se na pergunta anterior indicou a calculadora, indique com que finalidade a utilizou.

____
____
____
____
____

14. Nas aulas de Probabilidades e Combinatória foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____

MUITO OBRIGADO!

Inquérito para os Alunos do 10.º ano

do Curso Científico-Humanístico de Artes Visuais

[Matemática B]

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Estatística, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

4. Como avalia o seu rendimento a Matemática?

○ Muito Bom (18 – 20 valores)
○ Bom (14 – 17 valores)
○ Suficiente (10 – 13 valores)
○ Insuficiente (5 – 9 valores)
○ Mau (0 – 4 valores)

5. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: __________ 2.º Período: _________

6. Qual o tema de Matemática que mais gosta?

○ Geometria ○ Funções ○ Estatística ○ Movimentos Periódicos

7. Considera que o tema de Estatística:

○ É interessante ○ Não é interessante

8. Como avalia os conteúdos de Estatística?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

V.S.F.F.

9. Na sua opinião, os conteúdos de Estatística são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

10. No 3.º Ciclo do Ensino Básico, teve aulas de Estatística?

○ Sim ○ Não ○ Não me lembro

11. Este ano, que materiais utilizou durante as aulas de Estatística?

○ Computador
○ Calculadora científica
○ Calculadora gráfica
○ Régua
○ Compasso
○ Transferidor
○ Outro. Qual? ____

12. Se na pergunta anterior indicou a calculadora, indique com que finalidade a utilizou.

____
____
____
____

13. Nas aulas de Estatística foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____
____

MUITO OBRIGADO!

Inquérito para os Alunos do 11.º ano

do Curso Científico-Humanístico de Artes Visuais

[Matemática B]

O presente inquérito tem como objectivo analisar a receptividade dos alunos face aos programas de Matemática, aos conteúdos de Modelos de Probabilidades, assim como à forma como estes foram abordados nas aulas.

Esta análise vem no seguimento de uma investigação individual que culminará com a Dissertação de Mestrado.

1. Sexo: ○ Masculino ○ Feminino

2. Idade: ________ anos

3. Qual a sua opinião em relação à disciplina de Matemática?

○ Gosto Muito ○ Gosto ○ Indiferente ○ Gosto Pouco ○ Odeio

4. Como avalia o seu rendimento a Matemática?

○ Muito Bom (18 – 20 valores)
○ Bom (14 – 17 valores)
○ Suficiente (10 – 13 valores)
○ Insuficiente (5 – 9 valores)
○ Mau (0 – 4 valores)

5. Que classificação obteve a Matemática nos períodos anteriores?

1.º Período: ___________ 2.º Período: _________

6. Qual o tema de Matemática que mais gosta?

○ Movimentos não Lineares
○ Modelos de Probabilidades
○ Modelos Discretos
○ Modelos Contínuos não Lineares
○ Problemas de Optimização

7. Considera que os conteúdos de Modelos de Probabilidades:

○ São interessantes ○ Não são interessantes

V.S.F.F.

8. Como avalia os conteúdos de Modelos de Probabilidades?

○ Muito Fáceis ○ Fáceis ○ Nem fáceis nem difíceis ○ Difíceis ○ Muito Difíceis

9. Na sua opinião, os conteúdos de Modelos de Probabilidades são abordados, nas aulas, de uma forma:

○ Muito Simples ○ Simples ○ Nem simples nem complicada ○ Complicada ○ Muito Complicada

10. No 3.º Ciclo do Ensino Básico, teve aulas de Probabilidades e Estatística?

○ Sim ○ Não ○ Não me lembro

11. Este ano, que materiais utilizou durante as aulas de Modelos de Probabilidades?

○ Computador
○ Calculadora científica
○ Calculadora gráfica
○ Outro. Qual? ____

12. Se na pergunta anterior indicou a calculadora, indique com que finalidade a utilizou.

____
____
____
____
____

13. Nas aulas de Modelos de Probabilidades foram abordados vários conteúdos. Indique os três conteúdos que considerou mais difíceis.

____
____
____
____
____

MUITO OBRIGADO!

A N E X O I V

(Comando do *software* R)

```
>
> #ANÁLISE DO GRAU DE ASSOCIAÇÃO QUE EXISTE ENTRE A OPINIÃO QUE OS ALUNOS DO 7°ANO TÊM EM RELAÇÃO À MATEMÁTICA
> #E A AVALIAÇÃO QUE FAZEM DO SEU RENDIMENTO A ESTA DISCIPLINA.
>
>
> # Leitura dos dados a partir do ficheiro 7ano.OpiniãoMatemática.Rendimento.txt. Cria um data.frame
>
> ficheiro.dados<-read.table(file=
+ "C:\\Documents and Settings\\Proprietário\\Os meus documentos\\Tese da Sara10\\R\\7ano.OpiniãoMatemática.Rendimento.txt",
+ sep="\t")
>
> # Transforma o data.frame em matriz ou vector
> tab<-mat.or.vec(361,2)
> tab<-ficheiro.dados
>
> OpiniãoMatemática<-tab[,1]
> Rendimento<-tab[,2]
>
> # Coeficiente de Correlação de Spearman
> cor.test(OpiniãoMatemática,Rendimento,alternative="two.sided",method="spearman",exact=NULL,conf.level=0.95)

        Spearman's rank correlation rho

data:  OpiniãoMatemática and Rendimento
S = 3446824, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho
0.5604056

Warning message:
Cannot compute exact p-values with ties in: cor.test.default(OpiniãoMatemática, Rendimento, alternative = "two.sided",
>
>
>
```

A N E X O V

(Comando do *software* R)

```
> #ANÁLISE DOS DADOS REFERENTES À OPINIÃO QUE OS ALUNOS DO 7°ANO TÊM EM RELAÇÃO AOS CONTEÚDOS DE ESTATÍSTICA
>
>
> # Leitura dos dados a partir do ficheiro 7ano.conteudos.txt. Cria um data.frame
>
> ficheiro.dados<-read.table(file="C:\\Documents and Settings\\Proprietário\\Os meus documentos\\Tese da Sara6\\R\\7ano.conteudos.txt",sep="\t")
> ficheiro.dados
  V1 V2 V3
1 13  9  4
2  1  4  4
3 11 14  8
4 29 60 12
5 19 20  2
6  7 21  8
>
> # Transforma o data.frame em matriz
>
> dados<-data.matrix(ficheiro.dados)
> dados
  V1 V2 V3
1 13  9  4
2  1  4  4
3 11 14  8
4 29 60 12
5 19 20  2
6  7 21  8
>
> # Dá nomes às linhas e colunas da matriz
> dimnames(dados) = list(c("EB 2,3 Navegador Rodrigues Soromenho", "EB 2,3 Cardoso Lopes","EB 2,3 Mestre Domínguos Saraiva",
+ "EB 2,3/S Michel Giacometti","Escola Secundária de Gil Vicente","Escola Secundária Ferreira Dias"),
+ c("Fáceis e Muito Fáceis","Nem fáceis, nem difíceis", "Difíceis e Muito Difíceis"))
> dados
                                     Fáceis e Muito Fáceis Nem fáceis, nem difíceis Difíceis e Muito Difíceis
EB 2,3 Navegador Rodrigues Soromenho                    13                        9                         4
EB 2,3 Cardoso Lopes                                     1                        4                         4
EB 2,3 Mestre Domínguos Saraiva                         11                       14                         8
EB 2,3/S Michel Giacometti                              29                       60                        12
Escola Secundária de Gil Vicente                        19                       20                         2
Escola Secundária Ferreira Dias                          7                       21                         8
> # Teste de Fisher
> fisher.test(dados,hybrid=TRUE)

        Fisher's Exact Test for Count Data

data:  dados
p-value = 0.01098
alternative hypothesis: two.sided

Warning message:
'hybrid' is ignored for a 2 x 2 table in: fisher.test(dados, hybrid = TRUE)
>
```