Universidade de Aveiro Departamento de Matemática,
2020

Alda Manuela da Costa Campos Reis

A Aprendizagem Ativa em Matemática com Recurso à Tecnologia

Universidade de Aveiro Departamento de Matemática,
2020

Alda Manuela da Costa Campos Reis

A Aprendizagem Ativa em Matemática com Recurso à Tecnologia

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática para Professores, realizada sob a orientação científica do Professor Doutor Luís Descalço e co-orientação científica da Professora Doutora Paula Carvalho, Professores Auxiliares do Departamento de Matemática da Universidade de Aveiro.

o júri / the jury

presidente / president

Doutora Maria Paula de Sousa Oliveira

Professora Auxiliar do Departamento de Matemática da Universidade de Aveiro

Doutora Maria Helena Coelho Ribeiro

Professora Coordenadora do Instituto Politécnico de Leiria – Escola Superior de Tecnologia e Gestão

Doutor Luís António Arsénio Descalço

Professor Auxiliar do Departamento de Matemática da Universidade de Aveiro (orientador)

agradecimentos

À Professora Doutora Paula Carvalho e ao Professor Doutor Luís Descalço, cuja orientação foi muito importante para a realização e conclusão deste trabalho. Obrigada pelo incentivo, paciência e disponibilidade que sempre demonstraram.

Aos colegas e auxiliares do Agrupamento de Escolas da Trofa, mais concretamente, da Escola Secundária da Trofa agradeço por toda a ajuda e disponibilidade na concretização deste projeto.

Aos meus alunos da turma 06 do 11º (2019/2020), muito obrigada pela vossa participação neste trabalho. Sem vocês não teria sido possível a conclusão desta dissertação.

Ao meu marido, Mário, aos meus filhos, Bárbara e Rodrigo e ao futuro genro Miguel, obrigado pelo vosso apoio.

Aos meus Pais, Arnaldo e Genoveva e restante família pela pelo incentivo, por acreditarem em mim e estarem sempre presentes.

Aos meus amigos que de uma forma ou de outra me ajudaram neste processo e em especial ao apoio do amigo António Magalhães, muito obrigada.

A todos os que, de uma forma ou outra, contribuíram para a concretização deste objetivo, os meus sinceros agradecimentos.

Palavras-chave Inferência estatística, população, amostra, parâmetro, valor médio, desvio padrão, intervalo de confiança, SIACUA

Resumo O Ensino quer, em Portugal, quer na Europa tem vindo a sofrer algumas alterações. A mudança está enquadrada legalmente por um conjunto de normativos que nos regem desde 2018 com os quais temos que trabalhar nas escolas. Estes diplomas tem como base toda uma estratégia internacional para a Educação até 2030: Projeto Educação 2030, OCDE, 2016; Repensar a Educação, UNESCO, 2016; Resumo de Políticas, UNESCO, 2017.

Desde os programas das disciplinas às práticas letivas, muitas têm sido as mudanças que a Escola tem sofrido ao longo dos anos. Uma das principais razões para estas modificações está relacionada com algum insucesso existente ainda nos alunos. A falta de motivação e de interesse dos alunos pela aprendizagem tem condicionado a obtenção de sucesso nas várias disciplinas. No sentido de melhorar os resultados escolares, os Professores têm, ao longo dos anos, vindo a alterar as suas práticas letivas, adequando-as aos interesses, motivações e objetivos dos alunos.

Este trabalho tem como objetivo a criação de recursos digitais para a Plataforma SIACUA no apoio ao Ensino da Inferência Estatística na disciplina de Matemática Aplicada às Ciências Sociais do décimo primeiro ano no ensino secundário.

Assim, os conceitos de população, amostra, valor médio, desvio padrão e intervalo de confiança têm um papel fundamental neste estudo.

Keywords Statistical inference, population, sample, parameter, mean value, standard deviation, confidence interval, SIACUA

Abstract Teaching both in Portugal and in Europe has undergone some changes. The change is legally framed by a set of rules that govern us since 2018 with which we must work in schools. These legal documents are based on an international strategy for Education by 2030: Education Project 2030, OECD, 2016; Rethinking Education, UNESCO, 2016; Policy Summary, UNESCO, 2017.

From subject programs to teaching practices, there have been many changes that the school has suffered over the years. One of the main reasons for these modifications is related to some student failure. Students' lack of motivation and interest in learning has conditioned their success in the various subjects. In order to improve school results, teachers have, over the years, been changing their teaching practices, adapting them to the interests, motivations and goals of students.

This work is based on the creation of digital resources for the SIACUA Platform to support the Teaching of Statistical Inference in the subject of Applied Mathematics to the Social Sciences of the eleventh year in secondary education.

Therefore, this study was based on the concepts of population, sample, mean value, standard deviation and confidence interval.

Conteúdo

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas

AE – Aprendizagens Essenciais

APA – Articulação com o Perfil dos Alunos

AE| APA – Aprendizagens Essenciais| Articulação com o Perfil dos Alunos

CEBS – CAEBPA – Currículo do Ensino Básico e do Ensino Secundário para a Construção de Aprendizagens Essenciais baseadas no Perfil do Aluno

MACS – Matemática Aplicada às Ciências Sociais

PE – Projeto Educativo

Capítulo 1: INTRODUÇÃO

"À medida que as civilizações e a tecnologia avançam, as nossas vidas tornam-se mais dependentes da Matemática". [1]

O Ensino, quer em Portugal quer na Europa, tem vindo a sofrer algumas alterações. A mudança está enquadrada legalmente por um conjunto de normativos que regem as escolas desde 2018, com os quais as mesmas têm que trabalhar. Estes diplomas têm como base toda uma estratégia internacional para a Educação até 2030, a saber : Projeto Educação 2030, OCDE, 2016; Repensar a Educação, UNESCO, 2016; Resumo de Políticas, UNESCO, 2017. [2], [3]

Desde os programas das disciplinas às práticas letivas, muitas têm sido as mudanças que a Escola tem sofrido ao longo dos anos. Uma das principais razões para estas modificações está relacionada com algum insucesso existente ainda nos alunos. A falta de motivação e de interesse dos alunos pela aprendizagem têm condicionado a obtenção de sucesso nas várias disciplinas. No sentido de melhorar os resultados escolares, os professores têm, ao longo dos anos, vindo a alterar as suas práticas letivas, adequando-as aos interesses, motivações e objetivos dos alunos.

Segundo Prensky, os atuais alunos são uns autênticos "nativos digitais", respiram tecnologia e habitualmente dominam as ferramentas.[4]

Assim, este facto levou o que o Ministério da Educação (ME), há uns anos (2009) criasse alguns programas que permitiram transformar as escolas. Esses programas foram tão importantes que permitiram substituir os velhos recursos por alguns novos recursos tecnológicos, na tentativa de aproveitar a motivação dos alunos. É pena que este investimento tenha parado e não tenha perpetuado no tempo até aos dias de hoje.

Este trabalho tem por base a criação de recursos digitais para a Plataforma SIACUA [5] no apoio ao Ensino da Inferência Estatística na disciplina de Matemática Aplicada às Ciências

Sociais do décimo primeiro ano do Ensino Secundário, sempre tendo em consideração as Aprendizagens Essenciais (AE) definidas pelo Ministério da Educação (ME) para a disciplina/ano de escolaridade. [6]

Assim, este estudo tem como base os conceitos de população, amostra, valor médio, desvio padrão e intervalo de confiança. [6]

A Aprendizagem ativa em Matemática recorrendo à Tecnologia, pretende fornecer aos alunos, uma forma diferente e atrativa de aprender Matemática, fornecendo-lhes a possibilidade de utilizar uma ferramenta desconhecida para eles até ao momento. Assim pretende-se que os alunos possam aprender Matemática de uma forma mais lúdica e mais autónoma, despertando-lhes a vontade de querer aprender mais, desenvolvendo o pensamento reflexivo, crítico e criativo e a possibilidade de utilizar uma plataforma que lhes permite corrigir os seus próprios erros por um lado e por outro ouvir novamente extra-aula, os conteúdos explorados em sala de aula, através da utilização de vídeos. Para os professores poderá ser uma forma alternativa de apresentação dos conteúdos, recorrendo às novas tecnologias e deixando de parte o ensino tradicional ou como complemento ao trabalho realizado em sala de aula.

Apesar de os professores, constantemente, por um lado, tentarem desmistificar o facto de a Matemática ser considerada por muitos uma disciplina difícil, talvez considerada a mais difícil da matriz curricular, por outro tentarem propor aos alunos atividades diferentes e atrativas, ainda há muito a fazer e esta “luta” não pode ser posta de lado. Muitos são os professores que tentam combater o insucesso da Matemática, trabalhando de forma árdua, alterando as suas práticas pedagógicas na tentativa de combater o estigma e pessimismo que paira sobre a disciplina.

Assim, na tentativa de contribuir para esta mudança, referida anteriormente proponho-me desenvolver as atividades que a seguir descrevo, utilizando e produzindo conteúdos para a plataforma desenvolvida pela Universidade de Aveiro, SIACUA. [5]

Esta apresentação contempla conteúdos a serem utilizados em cinco aulas de cem minutos, duas no início e três no final do capítulo, referentes ao tema “Inferência Estatística”. Este tema na planificação da minha escola, está previsto para dez aulas de cem minutos.

O capítulo 2, será dividido em duas partes, na primeira será abordado a contribuição da disciplina de Matemática Aplicada às Ciências Sociais para o perfil do aluno à saída da escolaridade obrigatória e numa segunda parte, como pode a utilização da plataforma SIACUA contribuir para o perfil do aluno à saída da escolaridade obrigatória.

No capítulo 3, é feito o enquadramento teórico de tema: “Inferência Estatística”, começando por uma alusão à importância da estatística ao longo do tempo, passando pelos fundadores da estatística matemática, pelo método estatístico de resolução de um problema e, seguidamente, a abordagem de conceitos, métodos e teoremas utilizados nesta temática.

No capítulo 4, apresento a explicação dos objetivos dos vídeos produzidos, o contexto em que foram aplicados, os resultados da sua aplicação. É também, apresentado um dos vários exercícios produzidos, o porquê da sua escolha, o seu objetivo, como foi construído e como foram elaboradas as opções de escolha múltipla, quer a certa, quer as erradas. Tanto para os vídeos como para os exercícios alojados na plataforma SIACUA são apresentados a forma de como aceder aos conteúdos e os links para a visualização dos vídeos. [5]

O capítulo 5, está dividido em três partes, na primeira parte é feita uma breve descrição do Agrupamento de Escolas e em particular da Escola Secundária onde leciono e do meio onde quer o Agrupamento, quer a Escola estão inseridos. Na segunda parte está caracterizada de uma forma abreviada a turma onde foi aplicado este projeto e por fim uma monitorização dos resultados.

No capítulo 6 é apresentada a conclusão do trabalho realizado.

Capítulo 2: ENQUADRAMENTO

2.1. A contribuição da disciplina de Matemática Aplicada às Ciências Sociais para o perfil do aluno à saída da escolaridade obrigatória

A disciplina de Matemática Aplicada às Ciências Sociais (MACS) faz parte do curriculum do Curso de Línguas e Humanidades dos Cursos Científico-Humanísticos do Ensino Secundário e é lecionada quer no décimo, quer no décimo primeiro ano. Os alunos de outros cursos que optem por um percurso formativo próprio, também podem frequentar a disciplina ou auto proporem-se a realizar o Exame Nacional, nos termos da legislação em vigor.

Com esta disciplina pretende-se desenvolver a capacidade de formular e resolver problemas, desenvolver a capacidade de comunicação de ideias, do ponto de vista da Matemática. Mais importante que os jovens consigam adquirir ferramentas técnicas, é que tenham experiências matemáticas significativas que lhes permitam saber apreciar devidamente a importância das abordagens matemáticas nas suas profissões futuras. Os alunos devem também adquirir a capacidades de intervenção social contribuindo para uma cidadania ativa e participativa. Assim, esta disciplina desempenha um papel incontornável para a formação dos alunos, contribuindo para uma abordagem tão completa quanto possível de situações reais.

De entre inúmeros assuntos interessantes que ligam a Matemática à vida de todos os dias, vários foram os temas selecionados para os conteúdos da disciplina. Um deles, a salientar, foi a Estatística. Este conteúdo é lecionado em duas partes, uma no décimo ano, e outra no décimo primeiro ano. No décimo primeiro ano é abordada a introdução à Inferência Estatística, mostrando como se podem tirar conclusões a partir do estudo dos dados. Nesta altura os alunos vão adquirir mais alguns conceitos de Estatística, pois os estudantes têm a possibilidade de aprender como se podem tirar conclusões, partindo do particular para o geral, enquanto se quantifica o erro cometido. [6]

Na disciplina de MACS, começa-se por explorar exemplos simples, nomeadamente os que tenham sido objeto de estudo na parte da Estatística Descritiva lecionada no décimo ano e seguidamente são estudados mais conceitos e aplicações até à construção de intervalos de confiança. Esses exemplos vão permitir mostrar como se pode fechar o ciclo de um procedimento estatístico, que se iniciou com o planeamento da experiência e uma consequente recolha de dados, com o objetivo de uma tomada de decisão.

Desta forma, os jovens estudantes são convidados a fortalecer os seus conhecimentos, a desenvolver as suas capacidades e atitudes e a aplicá-los, quer em contextos matemáticos quer em contextos não matemáticos, fomentando também trabalho autónomo, colaborativo e de caráter interdisciplinar. Desta forma, há a contribuição da Matemática para que os alunos consigam adquirir as competências nas áreas: A-Linguagem e textos, B-Informação e Comunicação, C-Raciocínio e resolução de problemas, D-Pensamento critico e pensamento criativo, F-Desenvolvimento pessoal e autonomia e I-Saber científico, técnico e tecnológico previstas no perfil do aluno à saída da escolaridade obrigatória.

Para as restantes áreas de competência, E-Relacionamento interpessoal, G-Bem-estar, saúde e ambiente, H-Sensibilidade estética e artística e J-Consciência e domínio do corpo, previstas no Perfil do Aluno à Saída da Escolaridade Obrigatória, a Matemática dá igualmente contributos essenciais, quer pelos temas escolhidos para a disciplina, quer pelos processos e métodos matemáticos utilizados ao longo da abordagem dos conteúdos.

Os temas/conteúdos a abordar em MACS facilmente podem ser trabalhados na forma de projetos da disciplina ou integrados em projetos interdisciplinares. [2] e [6]

2.2. Como pode a utilização da plataforma SIACUA contribuir para o perfil do aluno à saída da escolaridade obrigatória

A aprendizagem ativa em Matemática recorrendo à tecnologia, pretende fornecer aos alunos uma forma diferente e atrativa de aprender Matemática, fornecendo-lhes a possibilidade de utilizar uma ferramenta. A utilização da plataforma SIACUA pode contribuir para o perfil do aluno à saída da escolaridade obrigatória, uma vez que os alunos podem, de uma forma mais lúdica, autónoma e criativa, aprender Matemática. A plataforma é de acesso fácil e com conteúdos que versam quer os temas referentes ao Ensino Superior, quer ao Ensino Básico e Secundário. Assim, os estudantes podem de uma forma gratuita, registarem-se e acederem a vários itens, para aprender um conteúdo, pois existem vídeos explicativos dos conceitos com exemplos, simplesmente para resolverem exercícios de escolha múltipla. A plataforma SIACUA proporciona aos alunos apoio ao estudo autónomo, uma vez que a mesma permite aos estudantes a correção dos seus próprios erros, pois os exercícios têm uma explicação detalhada da sua resolução.[5]

Os professores poderão como complemento ao trabalho a realizar na sala de aula, utilizar a plataforma SIACUA, propondo aos alunos a resolução de exercícios de escolha múltipla quer na própria aula, quer extra-aula, ou socorrendo-se, a uma forma alternativa de apresentação dos conteúdos, com recurso às novas tecnologias em vez do ensino tradicional.

Qualquer que seja a abordagem que um professor faça desta ferramenta de aprendizagem está sem dúvida a contribuir para o desenvolvimento de várias das competências espelhadas no perfil do aluno à saída da escolaridade obrigatória, já atrás mencionadas.

Capítulo 3: ENQUADRAMENTO TEÓRICO- Inferência Estatística

"É um erro capital teorizar sem se ter dados. Insensivelmente, começamos a distorcer os factos para os adaptar às teorias, em vez de adaptar as teorias aos factos." [7]

Segundo Grima, a estatística é uma disciplina familiar para todos. Estamos habituados a ouvir falar dela nos meios de comunicação social: um estudo (estatístico, claro) assegura que o consumo de tal substância baixou entre os jovens, os resultados de um inquérito afirmam que aquele político é mais valorizado do que o outro, ou que, se se realizassem hoje as eleições, tal partido venceria com tantos pontos de vantagem. Até mesmo nos jogos de futebol comenta-se e infere-se a análise da estatística: dizem as estatísticas, por exemplo, que dada equipa marca mais golos na segunda parte (...).

Curiosamente, apesar de falarmos de estatística para nos referirmos a esta disciplina, «estatísticas» no plural significam simplesmente dados. Que informação pode ser retirada dos dados e qual o grau de confiança dessa informação (aqui, a «estatística» no singular) nem sempre são perguntas óbvias. Às vezes, a estatística também é considerada um campo pouco sério. Dizer algo não significa que isso aconteça, e é muito possível que essa equipa de futebol que marca sempre na segunda parte desta vez nem marque qualquer golo. Nesse aspecto, a estatística contrasta com a restante matemática, que goza de uma imagem muito mais séria. Se uma equipa é «matematicamente campeã», então será campeã, independentemente do que vier a acontecer. Esta imagem associada à incerteza é igualmente afectada pela perplexidade criada pela capacidade dos políticos em apresentarem dados e estatísticas que apoiam sempre as suas teorias ou os seus interesses. Mas há muito mais do que tudo isto. A estatística está presente em muitas áreas: na investigação médica (será que este novo medicamento é melhor?), na biologia (quantos exemplares de determinada espécie existem numa região? Estarão em perigo de extinção?), quando é necessário realizar previsões (quanta electricidade vai ser gasta amanhã?), em estudos de mercado (que tipo de recipiente o consumidor prefere?), nos estudos sociológicos (o que pensam os jovens sobre tal assunto?), na economia (quanto aumentaram os preços?), em estudos de fiabilidade industrial (com que periodicidade é necessário verificar as peças de um avião?), ou na gestão da qualidade nas empresas (em que problema se deve concentrar esforços?). Talvez esta lista tenha ficado demasiado longa, mas não é de todo exaustiva: há muitas áreas nas quais a estatística é fundamental para fazer avançar o conhecimento." As estatísticas são dados, dados e mais dados. Mas não restam dúvidas de que esta disciplina utiliza o método científico para a sua recolha e análise. Não tem resposta para todas as questões, mas oferece uma boa base de referência a quem sabe o que perguntar e como processar os dados obtidos. A estatística estuda como recolher dados (quantos? De que maneira?) e como os analisar para obter a informação que permite responder às questões que colocamos. Trata-se de produzir conhecimento através da observação e da análise da realidade, de uma forma inteligente e objectiva. É a essência do método científico. [8]

A estatística tem-se difundido de forma tão rápida, que se torna cada vez mais difícil listar os ramos da atividade humana em que a sua aplicação se tenha revelado fecunda ou mesmo indispensável. (…) A questão que imediatamente surge é a seguinte: qual a motivação para tão vasta aplicações? Para responder a esta pergunta deve começar por notar-se que quando alguém (indivíduo ou instituição) pretende estudar um determinado fenómeno - com a simples finalidade de o descrever, ou com objectivos mais ambiciosos, como sejam explicá-lo, interpretá-lo, fazer previsões ou tomar decisões - depara-se com uma situação de incerteza, que tem por consequência nunca ser possível conhecer o fenómeno de forma completamente rigorosa. Nestas circunstâncias, começa-se normalmente por recolher os factos que pareçam importantes sobre os atributos mais relevantes do fenómeno em estudo. Estes factos devem conter informação e ser compilados de forma organizada; por isso se designam por dados. Em muitos casos os dados são de natureza quantitativa, isto é, são dados numéricos. Muitas vezes , pela insuficiente qualidade dos dados ou pelo deficiente conhecimento das técnicas, ou, ainda, pela irrelevância para os objectivos do estudo, o tratamento estatístico não passa além da exploração e da descrição dos dados recolhidos. Neste caso, é quase sempre suficiente calcular alguns indicadores que os caracterizem sumariamente e construir alguns gráficos adequados. (…).Quando os objectivos são mais ambiciosos - explicar, interpretar, prever ou decidir torna-se imprescindível aprofundar a análise do fenómeno, uma vez que a incerteza que o caracteriza se revela através de um dos aspectos mais significativos para a análise estatística - a variabilidade dos dados. Procura-se então modelar esta variabilidade estabelecendo um padrão que capture os aspectos essenciais do fenómeno e consentindo que os dados observados apresentem desvios, maiores ou menores, em relação ao padrão. Como não é geralmente possível captar toda a variabilidade contida nos dados, o modelo que vai definir o padrão e assim dar respostas aos objetivos fixados - constitui sempre uma representação simplifica da do fenómeno em estudo. Apenas se procura estabelecer uma «relação razoável» entre o padrão e os respectivos desvios. Como quase sempre existem aspectos do padrão adoptado que são desconhecidos, deve então procurare estimar o modelo, reduzindo tanto quanto possível os desvios, e testar a sua adequação. [9]

Moore (1992), refere que a Estatística é uma ciência que estuda a variabilidade apresentada pelos dados. Permite-nos mesmos retirar conclusões e também exprimir o grau de confiança que devemos ter nessas conclusões. A potencialidade da Estatística manifesta-se nesta particularidade. Podemos considerar três grandes áreas nesta ciência dos dados: Aquisição de dados, Análise de dados e Inferência a partir dos dados. [10]

Os fundadores da estatística matemática: Na última metade do século XIX, os alemães Helmert (1843 - 1917) e Lexis (1837 - 1914), o dinamarquês Thiele (1838 - 1910) e O inglês Edgeworth (1845 - 1926), obtiveram resultados que, além de contribuirem para o desenvolvimento da inferência estatística, foram valiosas antecipações só mais tarde plenamente compreendidas. O impulso decisivo deu-se, porém, no início do século XX, com K. Pearson (1857-1936), W. S. Gosset (1876-1937) e, emparticular, R. A. Fisher (1890 - 1962). Tendo sido inicialmente físico matemático, Pearson, levado pelo entusiasmo que Darwin fomentara, dedicou-se ao estudo da evolução. Entre 1890 e 1900 desenvolveu notavelmente a ciência que trata os dados da observação: publicou numerosos trabalhos na revista *Biotnetrika* que fundou em parceriacom Weldon (1860-1906) e Galton (1822-1911). Pearson, além da significativa contribuição que deu para a teoria da regressão e da correlação, levou a que a estatística fosse reconhecida como disciplina autónoma. Um importante grupo de artigos encontra-se recolhido em *Karl Pearson Early Statistical Papers* (Ed. S. Pearson, Cambridge Univesity Press, 1948). Em 1908, Gosset, que escreveu sob o pseudónimo de *Student,* publicou um artigo na *Biometrika* tratando da interpretação de pequenas amostras, dando origem a uma nova e importante fase nos estudos estatísticos. Gosset usava pseudónimo devido ao facto de trabalhar para a *Guiness,* empresa produtora de cerveja, que não desejava revelar aos concorrentes o emprego que fazia da estatística, Os estudos de Gosset podem-se consultar em *Student Collected Papers* (Ed, por E,S, Pearson e J. Wishart, University College, Londres, 1942). A Contribuição de Fisher para a estatística moderna é, pode dizer-se, a mais importante e a mais decisiva de todas. Formado em astronomia pela Universidade de Cambridge foi o fundador do célebre *Statistical Laboratory* da prestigiosa Estação Agronômica de Rothamsted. No princípio dos anos 20 estabeleceu o que muitos aceitam como a estrutura da moderna estatística analítica. O seu livro, *Statistical Methods for Research Workers,* publicado pela primeira vez em 1925, fez mais do que qualquer outro estudo no sentido de familializar os investigadores com as aplicações práticas dos métodos estatísticos e de criar mentalidade estatística entre a nova geração de cientistas. As investigações de Fisher encontram-se dispersas por numerosas revistas, tendo sido mais importantes reunidas em volume: *Contributions to Mathematical Statistics* (J. Wiley & Sons, Inc., Nova Iorque, 1950). [9]

Método estatístico de resolução de um problema: Para que se obtenham resultados válidos, o investigador deve seguir todos os passos que definem o método estatístico de resolução de problemas, como se pode ler em [11]:

1. *Identificar correctamente* o *problema* em análise. Mesmo em estudos exploratórios cujo objectivo é identificar possíveis relações entre as características dos indivíduos sem que, à partida, se defina um modelo regulador dessas relações, é necessário identificar o problema para o qual se pretendem encontrar respostas.

2. *Recolher* a *informação necessária,* relevante para o problema em estudo, em tempo útil e tão completa quanto possível. Esta informação poderá consistir em dados primários, recolhidos através de um questionário, ou dados secundários, recolhidos e publicados através de outra fonte de informação.
3. *Classificar e organizar* os *dados,* por exemplo, através da codificação e criação de uma base de dados em suporte informático. Uma vez ultrapassada esta fase, é já possível reduzir a quantidade de informação, fazendo desaparecer os pormenores menos importantes através de medidas de estatística descritiva (medidas de tendência central, dispersão, concentração, etc), quadros e gráficos.
4. *Análise dos dados e apresentação dos resultados:* identificar relações, testar hipóteses, definir modelos com a ajuda de métodos estatísticos apropriados.
5. *Tomar* a *decisão* mais adequada, ponderando as possíveis opções face aos objetivos inicialmente propostos. A qualidade da informação recolhida e as capacidades do investigador determinam, em grande parte, a adequabilidade das opções propostas.

Na mesma fonte é também definido que [11]:
Os métodos de Inferência Estatística envolvem o cálculo de estatísticas, a partir das quais se infere sobre os parâmetros da população, isto é, permitem, com determinado grau de probabilidade, generalizar à população certas conclusões, por comparação com os resultados amostrais.

A figura seguinte ilustra o processo seguido [11]:

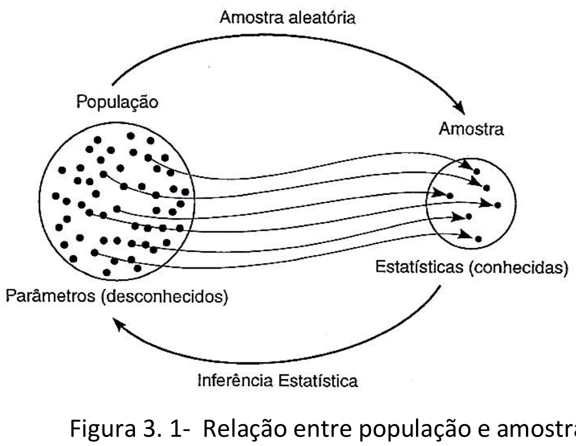


Figura 3. 1- Relação entre população e amostra

As definições, as propriedades, os teoremas e as demonstrações seguintes são baseados em [11], [12], [13] e [14].

Definição 3.1: População designa o conjunto dos elementos cujos atributos são objeto de um determinado estudo.

Definição 3.2: Amostra é um subconjunto finito da população.

Definição 3.3: Amostragem é a obtenção de informação sobre parte da população.

Definição 3.4: Processo de amostragem, é o processo seguido para escolher os elementos da população a incluir na amostra.

O processo de amostragem pode contribuir para obtermos várias amostras diferentes:

$$\text{População} \Rightarrow \begin{cases} \text{Amostra 1} \\ \text{Amostra 2} \\ \ldots\ldots \\ \text{Amostra n} \\ \ldots\ldots \end{cases}$$

Métodos de seleção da amostra

Existem dois grandes grupos de métodos para selecionar amostras: os métodos probabilísticos, também chamados de amostragem casual e os métodos não probabilísticos ou de amostragem dirigida. Será sobretudo analisado o primeiro daqueles grupos, pois a amostragem casual tem diversas vantagens sobre a amostragem dirigida, permitindo ao investigador:

1. Demonstrar a representatividade da amostra.
2. Medir explicitamente (em termos probabilísticos) o grau de incerteza com que se extrapola para a população/universo, isto é, o erro cometido por se usar uma amostra em vez da população.
3. Identificar explicitamente os potenciais enviesamentos.

Refira-se ainda que a precisão e o custo inerente ao processo de amostragem são fatores determinantes na escolha do tipo de método a utilizar.

Métodos de amostragem aleatória

Devido às suas bases teóricas, apoiadas na teoria das probabilidades, a amostragem aleatória tem sido adoptada pela pesquisa em muitas áreas científicas. O grau de confiança associado aos resultados obtidos, quando se utiliza um processo de amostragem aleatório, pode ser medido e controlado. Do mesmo modo, pode ser evitado qualquer desvirtuamento provocado por uma escolha dirigida dos respondentes, uma vez que o processo de seleção é casual e mecânico a partir de uma listagem de todos os indivíduos. Estes fatores podem ser considerados como as vantagens deste tipo de amostragem. No entanto, deverão ser também referidas as dificuldades em recolher uma amostra aleatória. A principal dificuldade consiste na obtenção de uma listagem completa da população a inquirir. Estas listagens são, na maioria dos casos, difíceis de conseguir, de custo elevado, demoradas na sua obtenção e nem sempre de fiabilidade aceitável. O segundo tipo de dificuldades relaciona-se com as não-respostas. Depois de definidos os respondentes, não poderão haver substituições, pelo que as não-respostas constituem uma importante fonte de enviesamento e terá de ser feito tudo para que a sua taxa seja minimizada. Todas as novas tentativas (por entrevista pessoal, telefone ou correio) para obter respostas bem sucedidas implicam aumento de custos e demora na obtenção dos resultados. A amostragem aleatória é, sem dúvida, o processo mais caro, mas os custos tendem a tornar-se pouco importantes face à fiabilidade dos resultados obtidos. De uma forma genérica podemos dizer que nos métodos de amostragem casual a probabilidade de selecionar determinado elemento da população é conhecida a *priori* e que tais métodos conduzem às chamadas amostras aleatórias. Importa caracterizar os métodos de amostragem casual mais frequentemente utilizados:

1. amostragem aleatória simples
2. amostragem sistemática
3. amostragem estratificada
4. amostragem por *clusters*
5. amostragem multi-etapas
6. amostragem multi-fásica.

Vamos agora ver cada um dos métodos:

1. *Amostragem aleatória simples*

Caracteriza-se por:

i) Cada elemento da população ter a mesma probabilidade de ser selecionado;

ii) Cada amostra de dimensão n ter a mesma probabilidade de ser escolhida.

Há duas formas de obter uma amostra:

i) a da lotaria;

ii) a dos números aleatórios.

As tabelas de números aleatórios são geradas por forma a garantir a natureza aleatória dos números que as compõem. Existem diferentes formas de obter números aleatórios, recorrerendo às tabelas já existentes, a calculadoras gráficas ou computadores . A grande dificuldade que os métodos de amostragem casual simples apresentam é a morosidade, sobretudo quando as amostras são de grande dimensão, a não ser que o processo de obtenção dos elementos que constituirão a amostra seja totalmente computorizada e se disponha de uma listagem dos elementos que constituem a população.

2. *Amostragem casual sistemática*

Este método é também chamado *quasi-aleatório* pois nem todas as amostras que se podem retirar de uma mesma população têm a mesma probabilidade de ocorrer. Para aplicação deste método é necessário calcular o valor $K = \frac{N}{n}$, N é o número de elementos da população e n é o número de elmentos da amostra. Em seguida, escolhe-se aleatoriamente um número, no intervalo $[1, K]$ que servirá como ponto de partida e primeiro elemento da amostra. Adicionando ao primeiro valor obtido o valor de K (arredondando o resultado por defeito), obtém-se o segundo elemento e a adição sucessiva do mesmo valor de K, permite encontrar os restantes elementos da amostra. Como se verifica, apenas o primeiro elemento é escolhido aleatoriamente enquanto que os restantes são determinados de modo sistemático pelo valor de K. Por exemplo, se $K = 2$, então a dimensão da amostra será constituída por metade (50%)da dimensão da População. Se $K = 20$, então a amostra será apenas 5% da População. Chama-se amostra sistemática a uma amostra obtida através deste procedimento. Em geral, o primeiro elemento a fazer parte da amostra é seleccionado aleatoriamente por um processo que se escolhe à partida.

3. *Amostragem estratificada*

Uma amostra estratificada obtém-se separando os elementos da população em grupos mutuamente exclusivos denominados estratos[1] e a partir destes a seleção de uma amostra aleatória simples dentro de cada estrato. Por mutuamente exclusivos pretende-se dizer que nenhum elemento da população pode estar simultaneamente presente em dois ou mais estratos. Este método permite, no caso de se conhecerem algumas características do universo ou população, obter resultados mais eficientes[2] com uma amostra de menor dimensão e igual representatividade. Essa eficiência será ainda mais importante se a variável a ser estratificada se encontrar correlacionada com várias outras variáveis como por exemplo idade, sexo, rendimento, status, área geográfica, etc., o que permitirá estratificar simultaneamente segundo várias variáveis, desde que se assegure uma adequada representatividade dos estratos existentes na população. Quando se utiliza um processo aleatório simples, o erro aleatório cometido resulta de dois erros diferentes: o

[1] Grupos homogéneos relativamente à característica ou características a estudar.
[2] Menor custo, menor tempo e menor possibilidade de erro.

erro dentro de cada estrato e o erro entre os diferentes estratos. Esta última componente é nula quando a amostra é estratificada, uma vez que se recolhem as opiniões dos diferentes estratos da população. A amostragem estratificada é ainda mais efetiva quando a diferença entre os vários estratos é mais acentuada, isto é, quando a dispersão dentro da população é elevada. Existem dois modos de obtenção de amostras estratificadas. No primeiro, cada estrato está representado na amostra proporcionalmente à sua importância (ou tamanho) na população total. No entanto, nos diferentes estratos, dimensões maiores poderão não estar associadas a uma maior dispersão ou variabilidade. Por essa razão, um modo de conseguir uma maior representatividade da amostra será representar os estratos na amostra tendo em conta a dispersão dentro de cada estrato da população. Este segundo modo de obtenção de uma amostra estratificada só pode ser aplicado nos casos em que se conhece a variabilidade dentro de cada estrato da população ou, no mínimo, quando existem estimativas dessa variabilidade retiradas de inquéritos feitos a populações semelhantes.

4. *Amostragem por* clusters

Este tipo de amostragem torna-se particularmente útil quando a população se encontra dividida num reduzido número de grupos ou *clusters,* caracterizados por terem uma dispersão idêntica à população total, isto é, os grupos deverão, tanto quanto possível, ser «microcosmos» da população a estudar. Primeiro, seleccionam-se aleatoriamente alguns dos grupos. Em seguida, incluem-se na amostra todos os indivíduos pertencentes aos grupos selecionados. Trata-se afinal de um processo de amostragem casual simples em que cada unidade é um *cluster.*

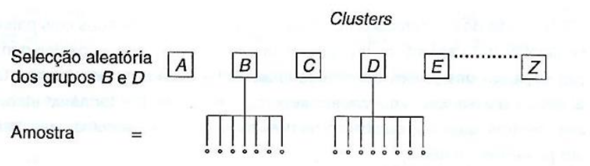


Figura 3. 2 - Clusters

Este tipo de amostragem é extremamente utilizado quando se torna impraticável ou até impossível construir uma lista de todos os elementos que constiuem determinada população sendo, no entanto, muito mais fácil listar grupos desses mesmos elementos.

5. *Amostragem multietapas*

O primeiro passo deste tipo de amostra é idêntico ao anterior. A população encontra-se dividida em vários grupos e selecionam-se aleatoriamente alguns desses grupos. No passo seguinte, também os elementos de cada grupo são aleatoriamente escolhidos. Este processo pode multiplicar-se por mais de duas etapas se os grupos estiverem divididos em sub-grupos.

Num estudo de mercados internacionais foram seleccionados dois países para se identificarem as táticas de posicionamento a seguir para as pastas dentífricas. Em cada um dos países escolhidos foram seleccionados cinco centros urbanos e, dentro destes, catorze estabelecimentos comerciais. Em todas as etapas (países, centros urbanos, estabelecimentos comerciais) as escolhas resultaram de um processo aleatório.

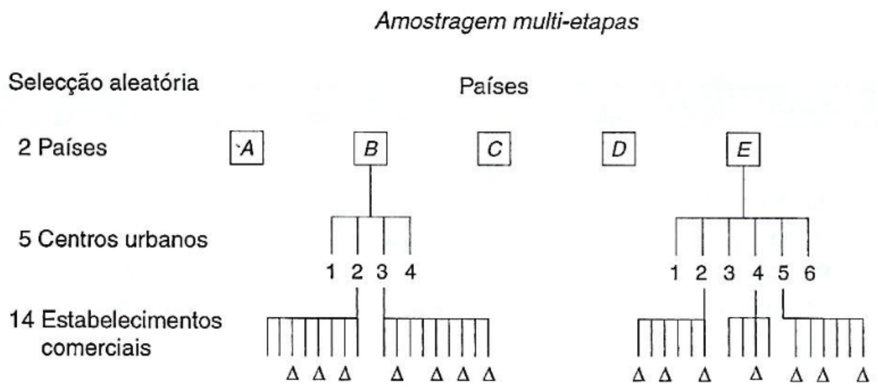


Figura 3. 3 - Amostragem multietapas

A desvantagem deste método adiante-se o facto de que os possíveis erros de amostragem se poderem multiplicar, dado que ao longo deste processo se vão o utilizando várias sub-amostras com a possibilidade de erros de amostragem em cada uma delas. A preocupação com a dimensão e precisão da amostra é aqui uma constante a nível de cada uma das etapas deste método.

6. *Amostragem multifásica*

Não deverão ser confundidos estes dois processos de amostragem: multietapas e multifásicas. No primeiro processo as unidades amostrais variam de uma etapa para outra. No exemplo referido no ponto anterior, as unidades amostrais eram, sucessivamente, os países, os centros urbanos e os estabecimentos comerciais, enquanto na amostragem multifásica define-se sempre mesma unidade amostral para todas as fases de extração da amostra. Na primeira fase, recolhem-se dados sobre determinadas características dos respondentes, por exemplo, o seu comportamento e frequência quanto ao consumo de determinado produto, variáveis demográficas, tamanho das empresas, a sua disponibilidade para responder novamente a um inquérito. Esta informação pode ser usada para a definição de uma listagem dos possíveis respondentes à segunda fase do inquérito. É então retirada desta listagem uma segunda amostra que responderá a um questionário com um nível de profundidade mais elevado.

Os diferentes tipos de métodos de amostragem aleatória abordados não são mutuamente exclusivos, podendo ser utilizados conjuntamente em fases diferentes do processo de amostragem. Por outro lado, uma amostra obtida por um método de amostragem do tipo aleatório não garante por si só uma resposta correta (a verdadeira, a que se obteria se se utilizasse o universo). No entanto, garante, isso sim, a capacidade de medir a probabilidade de obter a resposta errada. Existem outros processos de extrair amostras, sendo muitos deles combinações das técnicas anteriormente descritas com outras técnicas de amostragem não aleatória ou dirigida.

Definição 3.5: Parâmetro é uma característica da população, isto é, um valor caracterizador da população que, embora possa ser desconhecido, é fixo.

Definição 3.6: Estatística é uma característica da amostra, isto é, é uma função da amostra e, portanto, assume valores diferentes para diferentes amostras.

Definição 3.7: Variável aleatória, X , é uma função com domínio Ω , sendo Ω o espaço de resultados e com contradomínio $\mathcal{R}$. Assim $X\colon \omega \in \Omega \longrightarrow X(\omega) \in \mathcal{R}$.

Exemplo 3.1: Quando se lançam dois dados, o espaço de resultados pode ser dado pelo conjunto $\Omega = \{(i,j): i,j = 1,2,3,4,5,6\}$.

Quando se regista a soma dos pontos obtidos (Figura: 3.5), a variável aleatória é definida da seguinte maneira:

$X: (i,j) \in \Omega \longrightarrow i+j \in \mathcal{R}$ ou $X: \{(i,j)\} = i+j$.

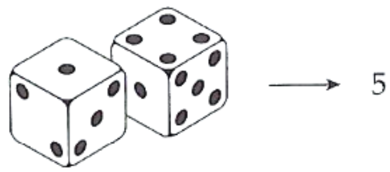


Figura 3. 4 - Variável aleatória

Exemplo 3.2:

Dos n apoiantes de um grande clube de futebol existem r que defendem a imediata substituição do treinador, Escolhem-se ao acaso, sucessivamente, três adeptos e regista-se se cada um deles concorda, C, ou discorda, D, com a referida substituição. O espaço de resultados é:

$$\Omega = \{CCC, CCD, CDC, DCC, CDD, DCD, DDC, DDD\}$$

Se interessar apenas o número de defensores da substituição na amostra de três apoiantes, sendo irrelevante a ordem por que apareçam, é natural introduzir a variável aleatória X definida pela correspondência:

$$X(CCC) = 3$$
$$X(CCD) = X(CDC) = X(DCC) = 2$$
$$X(CDD) = X(DCD) = X(DDC) = 1$$
$$X(DDD) = 0$$

Definição 3.8: Amostra aleatória de dimensão n é o vetor que se obtém quando se dispõe de uma amostra de n observações $x_1, x_2, \ldots, x_n$, supondo que tal amostra é a realização da variável n - dimensional ou vetor aleatório, $X_1, X_2, \ldots, X_n$.

Exemplo 3.3: Se para cada uma das amostras $X_1,\ X_2, \ldots, X_n$, referidas na definição anterior, se calcularmos a respetiva média para cada uma das amostras obtemos, $\bar{x}_1,\ \bar{x}_2, \ldots,\ \bar{x}_n$. Assim, podemos dizer que a média amostral, $\bar{X}$ é uma variável aleatória amostral, que assume um valor concreto $\bar{x}_n$, para cada amostra X_n.

O valor médio μ, e o desvio padrão σ, de uma população são parâmetros. Também podemos dizer que cada amostra aleatória retirada de uma população X, irá dar origem a estatísticas como valores diferentes. Deste modo, dizemos que as estatísticas são variáveis aleatórias e consequentemente têm uma certa distribuição de probabilidade.

A amostragem aleatória garante que todos os elementos da população têm a mesma probabilidade de ser incluídas na amostra. Assim considera-se uma população finita com m elementos e define-se, de acordo com essa população, uma variável aleatória Y com uma função de probabilidade associada p_y.

Como já referido anteriormente, existe uma diferença fundamental entre parâmetros e estatísticas: para uma dada população e uma dada variável aleatória sobre ela definida, os pârametros da distribuição correspondente (valor esperado, variância, desvio padrão, etc.) são fixos. As estatísticas (média amostral, variância amostral, desvio padrão amostral, etc.) variam de amostra para amostra. Dada esta variabilidade das estatísticas, interessará tentar definir a sua distribuição probabilística, através das respetivas funções de probabilidade ou densidade de probabilidade. As distribuições das estatísticas recebem o nome de distribuições de amostragem.

Este conceito será ilustrado no seguinte exemplo:

Exemplo 3.4:

Considere-se uma população com 4 elementos, que correspondem aos seguintes valores da variável aleatória $Y: \{2,4,6,6\}$. Na Tabela 3.1, apresenta-se a função de probabilidade desta variável, conjuntamente com o seu valor esperado e a sua variância.

$$\mu_Y = 2 \times \frac{1}{4} + 4 \times \frac{1}{4} + 6 \times \frac{1}{2} = 4.5$$

$$\sigma_Y^2 = (2 - 4.5)^2 \times \frac{1}{4} + (4 - 4.5)^2 \times \frac{1}{4} + (6 - 4.5)^2 \times \frac{1}{2} =$$

$$= 2.5^2 \times \frac{1}{4} + 0.5^2 \times \frac{1}{4} + 1.5^2 \times \frac{1}{2} = 2.75$$

y	p_y
2	$\frac{1}{4}$
4	$\frac{1}{4}$
6	$\frac{1}{2}$

Tabela 3.1 - Função de probabilidade da variável Y (Exemplo 3.4)

Pretende-se definir a distribuição por amostragem de $\bar{Y}$, a média amostral será calculada tendo como base amostras de dimensão dois, obtidas por um processo aleatório sem reposição. Assim, designou-se a média amostral por $\bar{Y}$, para significar que se trata de uma variável aleatória e não de um valor particular dessa variável, o qual se denota usualmente por $\bar{y}$.

Na Tabela 3.2 definem-se todas as amostras de dimensão 2, bem como as respetivas médias amostrais e as correspondentes probabilidades de ocorrência. A partir desta informação, definiu-se na Tabela 3.2 a função de probabilidade da média amostral, ou seja, definiu-se a distribuição de $\bar{Y}$.

Amostra	$\bar{y}$	Probabilidade de ocorrência
2,4	3	$\frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$
2,6	4	$\frac{1}{4} \times \frac{2}{3} = \frac{2}{12}$
4,2	3	$\frac{1}{4} \times \frac{1}{3} = \frac{1}{12}$
4,6	5	$\frac{1}{4} \times \frac{2}{3} = \frac{2}{12}$
6,2	4	$\frac{2}{4} \times \frac{1}{3} = \frac{2}{12}$
6,4	5	$\frac{2}{4} \times \frac{1}{3} = \frac{2}{12}$
6,6	6	$\frac{2}{4} \times \frac{1}{3} = \frac{2}{12}$

Tabela 3.2 - Amostras de dimensão n = 2, valores correspondentes da média amostral e respetivas probabilidades de ocorrência (Exemplo 3.4)

Amostra com $\bar{y}$	Número de ocorrências	Probabilidade da ocorrência	Produto entre o número de ocorrências e a probabilidade da ocorrência $p_{\bar{Y}}$
3	2	$\frac{1}{12}$	$\frac{1}{12} \times 2 = \frac{2}{12} = \frac{1}{6}$
4	2	$\frac{2}{12}$	$\frac{2}{12} \times 2 = \frac{4}{12} = \frac{1}{3}$
5	2	$\frac{2}{12}$	$\frac{2}{12} \times 2 = \frac{4}{12} = \frac{1}{3}$
6	1	$\frac{2}{12}$	$\frac{2}{12} \times 1 = \frac{1}{6}$

Tabela 3.3 - Distribuição de amostragem da média amostral (Exemplo 3.4)

$$\mu_{\bar{Y}} = 3 \times \frac{1}{6} + 4 \times \frac{1}{3} + 5 \times \frac{1}{3} + 6 \times \frac{1}{6} = 4.5$$

$$\sigma_{\bar{Y}}^2 = (3-4.5)^2 \times \frac{1}{6} + (4-4.5)^2 \times \frac{1}{3} + (5-4.5)^2 \times \frac{1}{3} + \ (6-4.5)^2 \times \frac{1}{6} = 1.5^2 \times \frac{1}{6} + 0.5^2 \times \frac{1}{3} + 0.5^2 \times \frac{1}{3} + 1.5^2 \times \frac{1}{6} = 0.917$$

Na Tabela 3.3 apresenta-se o cálculo do valor esperado e da variância da distribuição por amostragem de $\bar{Y}$. O desvio padrão da distribuição $\sigma_{\bar{Y}}$ e, em geral, o desvio padrão da distribuição por amostragem de uma estatística é designado habitualmente por erro padrão (da estatística em causa). No exemplo considerado anteriormente, a especificação a distribuição de amostragem a partir da distribuição original (com base na qual a estatística foi definida) foi simples, porque a população era muito pequena.

Podemos ainda dizer que o valor esperado e a variância da varável Y são:

$\mu_Y = 4.5$

$\sigma_Y^2 = 2.75$

Enquanto que o valor esperado e da variância da distribuição por amostragem de $\bar{Y}$são:

$\mu_{\bar{Y}} = 4.5$

$\sigma_{\bar{Y}}^2 = 0.917$

Podemos dizer que o valor esperado quer da variável aleatória Y, quer da distribuição de amostragem $\overline{Y}$ são iguais, enquanto que a variância de $\overline{Y}$ é aproximadamente um terço da variância de Y.

Definição 3.9. O ***valor Esperado*** (média amostral), ***a variância*** e o ***desvio padrão*** das estatísticas são definidas por:

$$\bar{X} = \frac{1}{n}\sum_i X_i, \qquad S^2 = \frac{1}{n}\sum_i (X_i - \bar{X})^2 \qquad S = \sqrt{\frac{1}{n}\sum_i (X_i - \bar{X})^2}$$

são, a média, a variância e o desvio padrão amostrais, onde $X_1, X_2, \ldots, X_n$ são os valores observados e n a dimensão da amostra.

Existem vários métodos de seleção de amostras aleatórias, mais ou menos complexos, que condicionam a forma de fazer inferências a partir da amostra, como já visto anteriormente. Seguidamente o estudo será restringido apenas ao processo particular de amostragem, a amostragem casual.

Definição 3.10: Amostra casual é quando as n variáveis aleatórias, componentes do vetor, $X_1, X_2, \ldots, X_n$ são independentes e identicamente distribuídas.

Definição 3.11: Variáveis casuais identicamente distribuídas são aquelas que que os X_i $(i = 1, 2, \ldots, n)$ são «cópias» da variável aleatória X que representa o atributo da população em estudo. ***A independência*** é quando a função de distribuição de X, F é a função de distribuição conjunta das n variáveis X_i que compõem a amostra aleatória se define, pelo produto :

$$F_{x_1 x_2 \ldots x_n}(x_1 x_2 \ldots x_n) = F(x_1)F(x_2)\ldots F(x_n)$$

que se designa por distribuição da amostra. Esta distribuição traduz a estrutura da população de amostras de dimensão n (do espaço-amostra) obtidas da população representada pela variável aleatória X.

Teorema 3.1. Se $(X_1, X_2, \ldots, X_n)$ é uma amostra casual de população para a qual existem média $\mu = E(X_i)$, variância $\sigma^2 = Var(X_i)$ e desvio padrão $\sigma = \sqrt{Var(X_i)}$ $(i = 1, 2, \ldots, n)$, tem-se: $E(\bar{X}) = \mu$, $Var(\bar{X}) = \frac{\sigma^2}{n}$, $\sigma(\bar{X}) = \frac{\sigma}{\sqrt{n}}$.

Demonstração:

Assim,

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \mu$$

Do mesmo modo,

$$Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{\sigma^2}{n}$$

Como $\sigma(\bar{X}) = \sqrt{Var(\bar{X})} = \sqrt{Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)} = \sqrt{\frac{1}{n^2} Var(\sum_{i=1}^{n} X_i)} = \sqrt{\frac{1}{n^2}\sum_{i=1}^{n} Var(X_i)} =$

$= \frac{\sigma}{\sqrt{n}}$,

uma vez que se têm variáveis aleatórias independentes, e a variância de cada X_i é σ^2.

A partir do teorema podemos dizer:

1. Os resultados são válidos para qualquer distribuição que verifique as condições sobre a existência de μ e de σ^2.
2. A relação $E(\bar{X}) = \mu$ permite concluir que o valor esperado da média da amostra, $\bar{X}$, é igual à média da população, μ .
3. Como $Var(\bar{X}) = \frac{\sigma^2}{n}$, pode concluir-se que quanto maior é a dimensão da amostra menor é a variância de $\bar{X}$, ou seja, maior é a concentração dos valores observaveis de $\bar{X}$, em torno de μ. Esta proposição está, aliás, de acordo com a intuição de quem faz observações casuais de um fenómeno e tem a noção de que, por exemplo, as médias de grupos de 100 observações apresentam menor variabilidade (entenda-se variabilidade de grupo para grupo) do que as médias de grupos de 10 observações, e que, portanto, a média calculada a partir de 100 observações merece confiança (representa melhor a média teórica ou média da população) do que a média calculada a partir de 10 observações.

Teorema 3.2. Se $(X_1, X_2, \dots, X_n)$ é uma amostra casual de população para a qual existem média $\mu = E(X_i)$ e variância $\sigma^2 = Var(X_i)$ $(i = 1, 2, \dots, n)$, tem-se:

$$E(S^2) = \frac{n-1}{n}\sigma^2.$$

Demonstração:

Como,

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i{}^2 - 2\bar{X}X_i + \bar{X}^2\right) =$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i{}^2 - 2\bar{X}\,\frac{1}{n}\sum_{i=1}^{n}X_i + \frac{1}{n}\sum_{i=1}^{n}\bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}X_i{}^2 - 2\bar{X}\bar{X} + \frac{1}{n}\sum_{i=1}^{n}\bar{X}^2 =$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i{}^2 - 2\bar{X}^2 + \frac{1}{n}\times n\,\bar{X}^2 = \frac{1}{n}\sum_{i=1}^{n}X_i{}^2 - 2\bar{X}^2 + \bar{X}^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}X_i{}^2 - \bar{X}^2$$

E, utilizando as propriedades:

1. Seja X, uma variável aleatória, a variância pode ser calculada por:

   $$Var(X) = E(X^2) - [E(X)]^2$$

2. Seja $(X_1, X_2, \dots, X_n)$ é uma amostra de população para a qual existem média $\mu = E(X_i)$ e variância $\sigma^2 = Var(X_i)$ $(i = 1, 2, \dots, n)$, então:

   $$E\left(X_i{}^2\right) = \sigma^2 + \mu^2 \quad \text{e} \quad E(\bar{X}^2) = \frac{\sigma^2}{n} + \mu^2$$

vem:

$$E(S^2) = E\left(\tfrac{1}{n}\textstyle\sum_{i=1}^{n}X_i{}^2 - \bar{X}^2\right) = \tfrac{1}{n}\textstyle\sum_{i=1}^{n}E\left(X_i{}^2\right) - E(\bar{X}^2) =$$

$$= \left(\tfrac{1}{n}\textstyle\sum_{i=1}^{n}(\sigma^2 + \mu^2)\right) - \left(\tfrac{\sigma^2}{n} + \mu^2\right) =$$

$$= (\sigma^2 + \mu^2) - \left(\tfrac{\sigma^2}{n} + \mu^2\right) = \tfrac{n-1}{n}\sigma^2.$$

Fica assim justificado porque é que, relativamente ao Exemplo 3.4, o valor esperado obtido para a média amostral (Tabela 3.3) , é igual ao valor esperado para a variável original (Tabela 3.1)(Teorema 3.1). Para definir a variância, é necessário estabelecer a distinção entre amostras aleatórias simples (população finita com reposição ou população infinita) e amostras aleatórias que não são simples (população finita, sem reposição). No primeiro caso, as variáveis $(X_1, X_2, \ldots, X_n)$ são independentes, a variância da média amostral é obtida de acordo com o Teorema 3.1.

Pelo mesmo teorema, também podemos dizer que :

$$\sigma_{\bar{X}}^2 = \frac{1}{n}\sigma_X^2 \quad \text{(amostras aleatórias simples).}$$

Desta expressão resulta que a dispersão da média amostral é inferior á dispersão da variável original e tanto menor quanto maior for a dimensão da amostra. Como interpretação deste resultado pode-se dizer que no cálculo da média amostral, as observações que se desviem no sentido do valor esperado tenderão a ser compensadas por observações que se desviem no sentido contrário, e a compensação será tanto mais perfeita quanto maior for o número de termos da média amostral. No segundo caso (população finita, sem reposição), as variáveis $(X_1, X_2, \ldots, X_n)$, não são independentes. De facto, se uma delas toma um valor acima do valor esperado, as outras tenderão a tomar valores abaixo daquele parâmetro, existindo portanto entre elas uma correlação negativa. Neste caso, a variância da média amostral vem dada por :

$\sigma_{\bar{X}}^2 = \left(\frac{m-n}{m-1}\right)\frac{1}{n}\sigma_X^2$ (com $n \leq m$) (população finita, amostragem sem reposição), em que m e n representam as dimensões da população e da amostra, respetivamente.

Ao factor $\left(\frac{m-n}{m-1}\right)$ chama-se fator de correção (ou redução) para populações finitas.

Em geral (isto é, quando $1 < n < m$), este fator é inferior à unidade, o que significa que, para a mesma variância da variável original, X, a dispersão da média amostral é menor no caso de a amostragem se efetuar sem reposição a partir de uma população finita. Este decréscimo da dispersão da média amostral deve-se à correlação negativa entre as variáveis $(X_1, X_2, \ldots, X_n)$.

Voltando ao Exemplo 3.3., a variância amostral, calculada atrás, pode ser obtida recorrendo à expressão $\sigma_{\bar{Y}}^2 = \left(\frac{m-n}{m-1}\right)\frac{1}{n}\sigma_Y^2$, ou seja, $\sigma_{\bar{Y}}^2 = \left(\frac{4-2}{4-1}\right)\frac{1}{2} \times 2.75 = 0.917$

Note-se que o valor obtido coincide com aquele que tinha sido obtido diretamente a partir da distribuição de $\bar{Y}$.

Chama-se, à atenção para o valor que o fator de redução toma em algumas situações particulares:

1. $m \to \infty$: Quando a dimensão da população tende para infinito, a amostra (que se admite ter dimensão finita) é aleatória simples, haja ou não reposição, e o fator tende para 1.
2. $n = 1$: Neste caso, o fator de redução toma o valor 1, pois, havendo apenas um elemento na amostra, não há qualquer diferença entre os processos com e sem reposição (a média amostral coincide com a variável original, sendo portanto idênticas as suas variâncias).
3. $m = n$: Quando a amostra coincide com a população, a média amostral coincide sempre com o valor esperado da variável original, tornando-se portanto uma constante. Neste caso, o fator de redução e, portanto, a variância da média amostral anulam-se.

Definição 3.12: Distribuição Normal - Diz-se que a variável aleatória X tem uma distribuição normal com parâmetros μ e σ^2, quando a função de densidade é da forma:

$\mathrm{f}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} (-\infty < x < +\infty)$ onde $\mu \in \Re$ e $\sigma^2 > 0$.

Simbolicamente $X \sim N(\mu, \sigma^2)$.

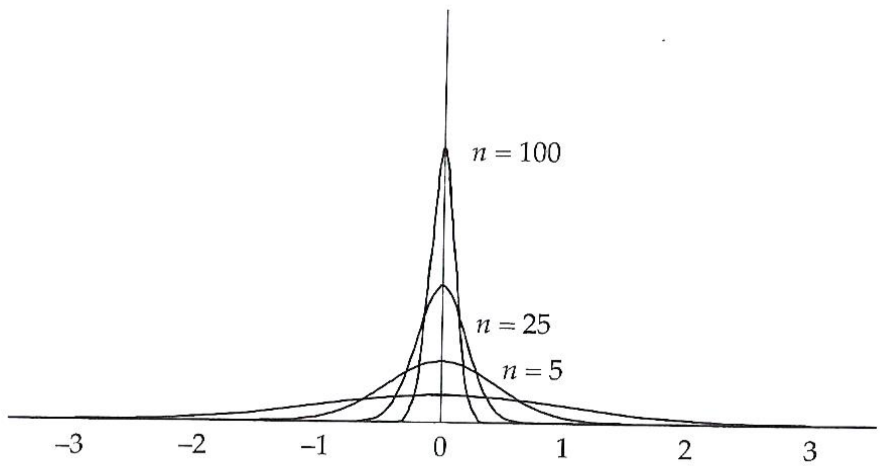


Figura 3. 5 - Distribuição X ~ N (0,1) e de $\bar{X}$ para n = 5,25 e 100.

Teorema 3.3:***Teorema do limite Central*** - Dada a sucessão de varíaveis aleatórias independentes e identicamente distribuídas $X_1, X_2, \dots, X_n$, com média μ e variância finita σ^2, então, quando $n \longrightarrow +\infty$, a função de distribuição da variável aleatória,

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma},$$

Tende para uma função de distribuição $N(0,1)$, ou seja, a distribuição aproximada de Z_n é $N(0,1)$. Simbolicamente, $Z_n \sim N(0,1)$.

O Teorema do limite central pode exprimir-se na forma alternativa:

$$\lim_{n \to +\infty} P(Z_n \leq x) = \lim_{n \to +\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x\right) = \Phi(x),$$

ou, ainda,

$$P(Z_n \leq x) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} \leq x\right) \approx \Phi(x) \text{ , (n grande).}$$

$\Phi(x) - A\ função\ acumulada\ da\ Distribuição\ Normal$

<u>**Considerações sobre o Teorema do limite central**</u> – no domínio da Estatística, este teorema constitui um dos desenvolvimentos teóricos mais notáveis, com inúmeras aplicações, pois permite, em particular, fazer progressos significativos na caracterização de distribuições por amostragem.

O que há de mais notável no teorema do limite central é o facto de não se impor nenhuma condição relativa à forma da distribuição original, ou seja, qualquer que seja esta forma, as variáveis soma ou média amostral têm uma distribuição aproximadamente Normal, se o número de termos for suficientemente elevado.

Exemplo 3.5: Analisando o mercado de certo produto, uma empresa concluiu que a procura diária a satisfazer (em centenas de quilogramas) é uma variável aleatória X com média 40 e variância 25. Sendo a produção anual planeada de 11500, pretende-se calcular a probabilidade de haver procura anual excedentária, considerando que um ano tem 289 dias úteis.

Representando por X_i $(i = 1, 2, 3, \dots, 289)$, a variável aleatória que exprime a procura no n-ésimo dia, tem-se $E(X_i) = \mu = 40$ e $Var(X_i\ ) = \sigma^2 = 25.$

Pretende-se calcular-se $P(\sum_{i=1}^{289} X_i > 11500).$

Pelo Teorema do limite central, tem-se:

$$P\left(\sum_{i=1}^{289} X_i > 11500\right) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} > \frac{11500 - 289 \times 40}{5\sqrt{289}}\right)$$
$$= P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} > -0.71\right) \approx 1 - \Phi(-0.71) = 0.7611$$

A procura excedentária é cerca de 0.76.

Definição 3.12: ***Estimador*** é uma variável aleatória, função da amostra casual, e representa-se por $T(X_1, X_2, \dots, X_n\ )$. A cada amostra observada faz corresponder uma ***estimativa***, uma avaliação, dos parâmetros desconhecidos[3].

Por exemplo, $T = \bar{X} = \frac{1}{n}\sum_i X_i$ e $T = S^2 = \frac{1}{n}\sum_i (X_i - \bar{X})^2$ são estimadores.

No âmbito da Inferência Estatística, ao calcular estatatísticas, há o objetivo de caracterizar a população a partir do qua a amostra foi retirada, estimando parâmetros dessa população. O exemplo seguinte ilustra a estimação pontual de um parâmetro.

Definição 3.13: ***Variável aleatória binomial*** - Seja X o número de sucessos obtidos na realização de n provas de Bernoulli independentes. Diremos que X tem distribuição binomial com parâmetros n e p, em que p é a probabilidade de sucesso em cada ensaio, se sua função de probabilidade for dada por:

$$p(k) = P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$$
$$X \sim B(n, p)$$

[3] Alguns autores têm uma abordagem muito pragmática: definem um estimador como uma regra que nos ensina como calcular o valor de uma estimativa com base numa amostra.

Exemplo 3.6:

Admita-se que uma determinada variável segue uma distribuição Binomial $X \sim B(N, p)$ em que os parâmetros N e p, são desconhecidos. Com base numa amostra aleatória simples qualquer, $(X_1, X_2, \ldots, X_n)$, pretende-se estimar:

a) o valor de cada um dos parâmetros;

b) de outros parâmetros definidos a partir daqueles, por exemplo, o valor esperado da distribuição $\mu = N.p$, ou da variância, $\sigma^2 = N.p.(1-p)$.

Assim, definem-se estatísticas cujos valores particulares constituam estimativas dos parâmetros em causa. Por exemplo, para estimar μ a partir da média amostral,

$$\bar{X} = \frac{1}{n}\sum_i X_i.$$

A estatística $\bar{X}$ é um estimador (pontual) do parâmetro em causa (o valor esperado), e uma realização particular $\bar{x}$ desta estatística (obtida a partir de um conjunto de valor observados $(x_1, x_2, \ldots, x_n)$, constitui uma estimativa daquele parâmetro.

Definição 3.14: ***Enviesamento*** - Seja θ um parâmetro qualquer de uma população e designa-se por $\hat{\theta}$ um estimador desse parâmetro. O enviesamento do estimador $\hat{\theta}$ define-se como a diferença entre o valor esperado do estimador, $E(\hat{\theta}) = \mu_{\hat{\theta}}$, e o valor do parâmetro, θ, isto é:

$$Enviesamento_{\hat{\theta}} = E(\hat{\theta}) - \theta = \mu_{\hat{\theta}} - \theta.$$

Definição 3.15: Um estimador diz-se **não-enviesado** quando o seu enviesamento for nulo, e **enviesado** no caso contrário. Na Figura 3.7 e na Figura 3.8 ilustram-se os conceitos de estimador não enviesado $(\hat{\Theta}_1)$ e estimador enviesado $(\hat{\Theta}_2)$.

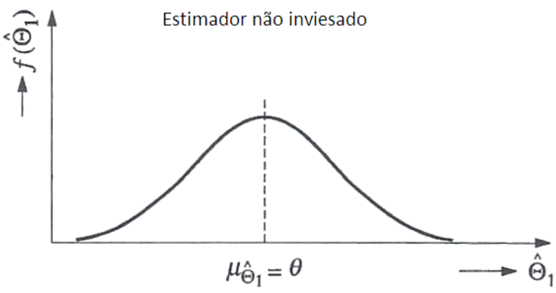


Figura 3.6 – Estimador Não Enviesado $(\hat{\theta}_1)$.

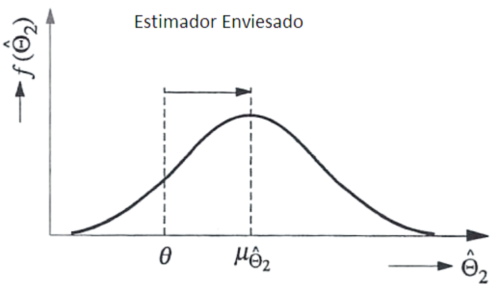


Figura 3.7 – Estimador Enviesado $(\hat{\theta}_2)$

Exemplo 3.7: Considera-se uma determinada população que é caracterizada pela varíavel aleatória X, a **estatística** média amostral: $E(\bar{X}) = E(X) = \mu_X$ é um estimador não enviesado do valor esperado, qualquer que seja a distribuição populacional, calculado com base em amostras aleatórias simples de dimensão N.

Conceito de Intervalo de Confiança:

Cada método de estimação pontual permite associar a cada parâmetro populacional um estimador. A cada estimador estarão associadas tantas estimativas diferentes quantas as amostras que forem utilizadas para o seu cálculo. Na generalidade dos casos, nenhuma dessas estimativas coincidirá com o valor do parâmetro. A grande limitação dos métodos de estimação pontual é a de não fornecerem qualquer informação ao rigor ou confiança das estimativas obtidas. Este constrangimento é ultrapassado recorrendo aos métodos de estimação por intervalos. Existem, no entanto situações em que é preferível a estimação por intervalos. Assim, em vez de propor apenas ums estimativa isolada $\hat{\theta}$, para θ, faz-se acompanhar esta de um certo intervalo $]t_1, t_2[$ para que o verdadeiro valor do parâmetro esteja possivelmente entre t_1 e t_2. Ao associar um intervalo à estimativa proposta atribui-se ao mesmo intervalo um grau de confiança. Em muitos casos, o intervalo é da forma $]\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon[$ em que o valor de ε pode ser considerado uma medida de precisão ou medida do erro inerente à estimativa $\hat{\theta}$.

Consideremos agora o seguinte exemplo:

Exemplo 3.8: O peso em (mg) de uma amostra de 20 palitos produzidos numa oficina de madeiras.

807.71	814.31	861.03	838.27
790.68	823.45	814.80	842.45
818.82	828.94	821.47	816.56
853.42	808.86	785.84	850.44
858.59	797.88	833.19	838.64

Tabela 3.4 - Amostra de pesos em mg de palitos

Admita-se que a variável aleatória que representa o peso dos palitos tem distribuição $N(\mu, \sigma^2)$, da qual se conhece o desvio padrão $\sigma = 22.5$ mg . O objetivo é propor um valor para o peso médio dos palitos. Assim, vamos utilizar a média da amostra,

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^{20} x_i}{20} = 825.27mg.$$

Noutra perspetiva, é possível dizer que, sendo $\bar{X}$ a variável aleatória que designa a média amostral da população, então $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\bar{X}-\mu}{\frac{22.5}{\sqrt{20}}} \sim N(0,1)$.

O facto da variável Z ter distribuição conhecida , independente do parâmetro a estimar, μ, permite calcular a probabilidade de Z assumir um valor compreendido entre quaisquer números reais. Por exemplo, usando a Tabela 3.5,

ε	.0005	.0010	.0050	.0100	.0200	.0250	.0500	.1000	.2000	.3000	.4000
z_ε	3.290	3.090	2.576	2.326	2.054	1.960	1.645	1.282	.842	.524	.253
$z_{\varepsilon/2}$	3.481	3.290	2.807	2.576	2.326	2.241	1.960	1.645	1.282	1.036	.842

$z_\varepsilon : P(Z > z_\varepsilon) = \varepsilon; z_{\varepsilon/2} : P(|Z| > z_{\varepsilon/2}) = \varepsilon.$

Tabela 3.5 - Distribuição Normal: $\Phi^{-1}(Z)$

$P(-1{,}96 < Z < 1{,}96) = P\left(-1{,}96 < \frac{\bar{X}-\mu}{5{,}03} < 1{,}96\right) = \Phi(1{,}96) - \Phi(-1{,}96) = 2\Phi(1{,}96) - 1 = 0{,}95$, o que, tendo em atenção que $\frac{22{,}5}{\sqrt{20}} \approx 5{,}03$, conduz a $P(-9{,}86 < \bar{X} - \mu < 9{,}86) = 0{,}95$. Resolvendo as desigualdades em ordem a μ, vem $P(\bar{X} - 9{,}86 < \mu < \bar{X} + 9{,}86) = 0{,}95$. Considerando uma amostra genérica, $\bar{X}$ é uma variável aleatória e $]\bar{X} - 9{,}86;\ \bar{X} + 9{,}86[$ um intervalo aleatório. A probabilidade de este intervalo aleatório conter o valor desconhecido da média da população, μ, é igual a 0,95. Assim, se se conceber a observação de um grande número de amostras de 20 elementos, a cada uma das quais corresponde um particular intervalo $]\bar{x} - 9{,}86;\ \bar{x} + 9{,}86[$, a proporção destes intervalos a que pertence μ é aproximadamente 95%.

Tem-se, assim, considerável confiança em que, para uma particular amostra de 20 observações, o intervalo correspondente $]\bar{x} - 9{,}86;\ \bar{x} + 9{,}86[$, abranja o verdadeiro valor de μ. Para medida dessa confiança toma-se a probabilidade de o intervalo aleatório que lhe está associado, $]\bar{X} - 9{,}86;\ \bar{X} + 9{,}86[$ conter μ, ou seja, 0,95. Diz-se então que o particular intervalo $]\bar{x} - 9{,}86;\ \bar{x} + 9{,}86[$, constitui um **intervalo de confiança a** 95% para μ. A probabilidade 0,95 é designada por **grau de confiança.**

Com a amostra considerada no exemplo anterior, o intervalo de confiança a 95% para μ é $]815{,}41; 835{,}13[$. Pode então afirmar-se, com elevado grau de confiança, que a média do peso dos palitos se situa entre $815{,}41$ e $835{,}13$mg. Procedendo de modo semelhante, obtêm-se intervalos de confiança com qualquer grau de confiança entre 0 e 1.

Quando a **variância é conhecida,** $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0{,}1)$. A dedução do intervalo de confiança a $(1-\alpha)100\%$, foi desenvolvida no Exemplo 3.7 e conduz a:

$\left]\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right[$, com $z\alpha_{/2}$ a verificar $\Phi\left(z_{\alpha/2}\right) = 1 - \frac{\alpha}{2}$.

Como Z tem distribuição simétrica em relação à origem, vem:

- A amplitude do intervalo, para um grau de confiança $1-\alpha$, é $\frac{2z_{\alpha/2}\sigma}{\sqrt{n}}$.

Exemplo 3.9: Seja X o tempo, em minutos, que certa tarefa leva a executar. Admita-se que X tem distribuição $N(\mu,\ \sigma^2\ )$, e que a experiência anterior permitiu determinar, com boa aproximação, que $\sigma = 2$ minutos.

a) Observada uma amostra de dimensão 16, com média $\bar{x} = 12{,}5$ construa-se um intervalo de confiança 95% a para μ.

Como $1-\alpha = 0{,}95$, tem-se $Z_{\alpha/2} = Z_{0{,}025} = 1{,}96$, donde o intervalo

$$\left]12{,}5 - 1{,}96\frac{2}{\sqrt{16}};\ 12{,}5 + 1{,}96\frac{2}{\sqrt{16}}\right[=]11{,}52; 13{,}48[.$$

b) Desejando estimar-se μ através de um intervalo de confiança a 95%, qual a dimensão da amostra de tempos que deve observar-se para garantir que a amplitude do intervalo obtido é inferior a 0,5 minutos?

Pretende-se determinar o menor valor de n tal que:

$$2z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < 0{,}5 \Leftrightarrow n > \left(\frac{2z_{\alpha/2\sigma}}{0{,}5}\right)^2 \Leftrightarrow n > \left(\frac{2\times1{,}96\times2}{0{,}5}\right)^2 \Leftrightarrow n > 15{,}68^2, \text{ logo } n = 246.$$

Exemplo 3.10: Admita-se que a altura dos alunos do 10º ano segue uma distribuição Normal cuja variância é conhecida: $\sigma^2 = (0{,}051m)^2$. Admita-se ainda que foi recolhida uma amostra aleatória com dimensão $n = 25$ alunos e calculada a respetiva média amostral, tendo-se obtido $\bar{x} = 1.70\ m$. Como definir um intervalo que, com uma probabilidade de 95%, contenha o valor esperado da altura, μ ?

Subsituindo na expressão geral :

$$\left]\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right[$$

Como $1-\alpha=0{,}95$, tem-se $Z_{\alpha/2}=Z_{0{,}025}=1{,}96$, donde o intervalo

$$\left]1{,}70-1{,}96\frac{0{,}051}{\sqrt{25}}\,;1{,}70+1{,}96\frac{0{,}051}{\sqrt{25}}\right[=]1{,}68;1{,}72[.$$

Este intervalo é designado por intervalo de confiança a 95%. A semiamplitude do intervalo que é $\frac{1{,}72-1{,}68}{2}=0{,}02$ é o erro máximo que com a confiança especificada, se comete na estimativa de μ .

Capítulo 4: RECURSOS: desenvolvidos na Plataforma SIACUA

Tendo em conta as turmas que me foram atribuídas e a análise dos modelos existentes na plataforma SIACUA, verifiquei que no domínio Estatística e Probabilidades, subdomínio Estatística não existia nenhum recurso. Assim, decidi construir modelos referentes a esta temática, mais propriamente na Inferência Estatística. [3]

Escolhi alguns temas deste capítulo a lecionar na turma de MACS do décimo primeiro ano. Assim, construí vídeos e exercícios que envolvessem os conceitos, que relacionassem as várias propriedades estudadas e que pudessem ser utilizados nas minhas aulas como estratégia de aprendizagem por parte dos alunos, na sala de aula ou extra-aula. Para construir os novos recursos, utilizei para a produção dos vídeos denominados: "Inferência Estatística-revisão de conceitos" e "Estimação de Parâmetros", o PowerPoint para o texto e para as imagens e o ScreenRec para captar o som. Para os vídeos intitulados: "Intervalos de Confiança para a Média" e "Intervalos de Confiança para a Proporção", utilizei o OneNote para o texto, o gravador de voz do Windows para captar o som e seguidamente OpenShot para editar o vídeo. Criei um canal no YouTube, para a publicação dos quatro vídeos e alojei-os também no SIACUA. [5]

Neste processo de alojamento produzi um pequeno texto para colocar junto a cada um dos vídeos, onde foi necessário recorrer à linguagem HTML e LaTeX. Para a produção dos exercícios, foi feita uma pesquisa nos enunciados dos exames nacionais e nos manuais de décimo primeiro ano.[15],[16] e [17]

Também foi necessário aprender um pouco de LaTeX e de HTML para escrever o "código" dos exercícios a alojar no SIACUA. [5]

Encontram-se nos apêndices, respetivamente 1 e 2, os diapositivos referentes aos vídeos e o código em HTML e LaTeX do texto que se encontra junto aos mesmos, assim com o código em LaTeX e HTML para os exercícios de escolha múltipla, que foram alojados no SIACUA.[5]

4.1. Vídeos

Atualmente, o uso das novas tecnologias é indispensável ao cotidiano de todo o cidadão, a informática, o *smartphone* e, em particular, os vídeos educacionais, entre outros, são recursos que vieram para ficar e servir a humanidade. Usar um vídeo educativo como ferramenta pedagógica é um poderoso recurso que auxilia o professor no seu processo de lecionação, e ao aluno na sua aprendizagem.

O primeiro vídeo tem como principal objetivo recordar alguns conceitos importantes lecionados no ano letivo anterior e que servem de base para os novos conteúdos a aprender este ano. Assim, antes de iniciar a lecionação do capítulo da Inferência Estatística, pedi aos alunos para verem, extra-aula, o vídeo colocado na plataforma SIACUA, explicando-lhes que seria para reverem conteúdos apreendidos no ano anterior. Esta estratégia, foi utilizada em vez de pedir para consultarem o manual do ano anterior. Desta forma, os alunos ficaram muito mais motivados para a revisão dos conceitos já estudados, contribuindo para uma aprendizagem mais interessante. Quando dei início ao estudo da inferência estatística na sala de aula, voltei a projetar o vídeo que tinha pedido aos alunos para verem em casa e apercebi-me que eles já tinham recordado os conceitos e estavam bem presentes. O vídeo dois, três e quatro têm como objetivo a abordagem do conceito de estimativa, parâmetro, estimação de parâmetros, intervalo de confiança para a média e para a proporção respetivamente. Em suma, a utilização deste tipo de recurso permitiu verificar que praticamente todos os alunos de uma forma autonomia e motivada fizeram a revisão, solicitada, dos conceitos do ano letivo anterior e que estiveram sempre muito interessados, atentos e empenhados. Posteriormente conseguiram colocar algumas questões de uma forma criativa. A utilização de um vídeo na sala de aula pode potencializar uma aprendizagem mais significativa através da visão.

Para aceder aos vídeos na plataforma SIACUA é necessário seguir os seguintes passos:

1. Utilizar o endereço https://siacua.web.ua.pt
2. Registar na plataforma
3. Entrar na plataforma – fazer Login

4. Escolher curso - Matemática Secundário
5. Clicar em Estatística, Cálculo Combinatório e Probabilidades – Estatística
6. Clicar à direita no lápis
7. Rever conceitos prévios – vídeo 1
8. Aprender um conceito – vídeo 2, 3 e 4

Os vídeos podem também ser vistos diretamente no YouTube. Os links são os seguintes:

Vídio 1: https://youtu.be/ZbVoddxmgms

Vídio 2: https://youtu.be/NqNijLA1qHk

Vídio 3: https://youtu.be/T4y_m4aSkSA

Vídio 4: https://youtu.be/s6y18ZnGWEo

4.2. Exercícios de Escolha Múltipla

Não é possível negar, as plataformas com exercícios interativos são atualmente e cada vez mais, muito atrativas e motivadoras para os alunos. O exercício seguinte foi um dos que construi e parametrizei para colocar na plataforma SIACUA [5].

Escolhi este para apresentar a seguir, como poderia ter escolhido qualquer outro, pois considero que construi alguns bastante interessantes por motivos diferentes. O motivo da escolha residiu no facto de ser um exercício de aplicação direta dos conceitos, não apela ao relacionamento de conteúdos, nem a raciocínios mais elaborados, nem à criatividade como seria necessário em alguns dos outros construídos. É um exercício de escolha múltipla, tem como objetivo o cálculo da proporção e a construção de um intervalo de confiança para a proporção, tendo em conta o valor da confiança. As opções de escolha múltipla contemplam a resposta certa em que o aluno calcula a proporção, utiliza bem o valor de z, escreve o intervalo com os dados e depois efetua os cálculos de uma forma totalmente correta. As opções erradas foram construídas pensando nos erros que os alunos por vezes cometem.

Exercício:

Recolheu-se uma amostra de 36 *smartphones* da produção de uma determinada fábrica. Dessas 36 unidades, verificou-se que 8 tinham uma avaria.

Com um nível de confiança de 95%, a proporção populacional de *smartphones* sem avarias está, aproximadamente, entre:

(A) 64% e 91%

(B) 9% e 36%

(C) 66% e 89%

(D) 76% e 80%

Resolução detalhada:

Com um intervalo de confiança de 95%, o valor de $z = 1{,}960$. O tamanho da amostra é $n = 36$

Como se pretende um intervalo de confiança para a proporção, então será necessário começar por calcular a proporção de *smartphones* sem avarias:

$$\hat{p} = \frac{28}{36} = \frac{7}{9}$$

O intervalo de confiança para a proporção, conforme referido no enunciado, obtém-se:

$$\left]\hat{p} - z\sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}}\right[$$

Agora substitui-se os valores da proporção $\hat{p}$, $\hat{p} = \frac{7}{9}$ calculada anteriormente e o valor de z, atrás mencionado, $z = 1{,}960$ e obtém-se:

$$\left]\frac{7}{9} - 1{,}960 \times \sqrt{\frac{\frac{7}{9}\times\left(1-\frac{7}{9}\right)}{36}}; \frac{7}{9} + 1{,}960 \times \sqrt{\frac{\frac{7}{9}\times\left(1-\frac{7}{9}\right)}{36}}\right[$$, ou seja, $]0{,}642; 0{,}914[$

Para encontrarmos a solução é necessário transformar os valores encontrados, em percentagem arredondada às unidades.

$$\left.\begin{array}{r} 0{,}642 \times 100 = 64{,}2\% \approx 64\% \\ 0{,}914 \times 100 = 91{,}4\% \approx 91\% \end{array}\right\} \rightarrow 64\% \text{ a } 91\%$$

Portanto a resposta correta é a (A)

<u>**Observação**</u>: A resposta correta é a (A), caso o enunciado seja exatamente este, pois na plataforma SIACUA as possíveis respostas não aparecem sempre pela mesma ordem.

<u>**As soluções erradas:**</u>

<u>Afirmação 1</u>: o aluno em vez de considerar o número de *smartphones* sem avarias, lê erradamente o enunciado e considera o número de *smartphones* com avarias e calcula o valor de $\hat{p}$ errado, ou seja,

$\hat{p} = \frac{8}{36} = \frac{2}{9}$

O tamanho da amostra é $n = 36$ e $z = 1{,}960$

O intervalo de confiança para a proporção, conforme referido no enunciado, obtém-se:

$$\left]\hat{p} - z\sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}\right[$$

Agora substitui-se os valores da proporção $\hat{p}$, $\hat{p} = \frac{2}{9}$ calculada anteriormente e o valor de z, atrás mencionado, $z = 1{,}960$ e obtém-se:

$$\left]\frac{2}{9} - 1{,}960 \times \sqrt{\frac{\frac{2}{9}\times\left(1-\frac{2}{9}\right)}{36}}; \frac{7}{9} + 1{,}960 \times \sqrt{\frac{\frac{2}{9}\times\left(1-\frac{2}{9}\right)}{36}}\right[=]0{,}086; 0{,}358[\quad \rightarrow 9\% \text{ a } 36\%$$

<u>Afirmação 2</u>: O aluno considera o valor de z incorreto, pois consulta erradamente a tabela que relaciona o valor de z com o nível de confiança.

O tamanho da amostra é $n = 36$ $z = 1{,}645$

Proporção de *smartphones* sem avarias:

$\hat{p} = \frac{28}{36} = \frac{7}{9}$

O intervalo de confiança para a proporção, conforme referido no enunciado, obtém-se:

$$\left]\hat{p} - z\sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}}\right[$$

Agora substitui-se os valores da proporção $\hat{p}$, $\hat{p} = \frac{7}{9}$ calculada anteriormente e o valor de z, atrás mencionado, $z = 1{,}645$ e obtém-se:

$$\left]\frac{7}{9} - 1{,}645 \times \sqrt{\frac{\frac{7}{9}\times\left(1-\frac{7}{9}\right)}{36}}; \frac{7}{9} + 1{,}645 \times \sqrt{\frac{\frac{7}{9}\times\left(1-\frac{7}{9}\right)}{36}}\right[=]0{,}664; 0{,}892[\rightarrow 66\% \text{ a } 89\%$$

Afirmação 3: Erro de cálculo, ao colocar os valores na calculadora, o número 36 fica fora da raiz quadrada:

Com um intervalo de confiança de 95%, o valor de $z = 1{,}960$. O tamanho da amostra é $n = 36$

Como se pretende um intervalo de confiança para a proporção, então será necessário começar por calcular a proporção de smartphones sem avarias:

$$\hat{p} = \frac{28}{36} = \frac{7}{9}$$

O intervalo de confiança para a proporção, conforme referido no enunciado, obtém-se:

$$\left]\hat{p} - z\sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}}\right[$$

Agora substitui-se os valores da proporção $\hat{p}$, $\hat{p} = \frac{7}{9}$ calculada anteriormente e o valor de z, atrás mencionado, $z = 1{,}960$ e obtém-se:

$$\left]\frac{7}{9} - 1{,}960 \times \sqrt{\frac{\frac{7}{9} \times \left(1 - \frac{7}{9}\right)}{36}}; \frac{7}{9} + 1{,}960 \times \sqrt{\frac{\frac{7}{9} \times \left(1 - \frac{7}{9}\right)}{36}}\right[$$

O aluno escreve corretamente o intervalo, mas ao colocar os dados na máquina coloca o 36 fora da raiz quadrada e chega ao seguinte intervalo:

$]0{,}755; 0{,}800[$ ➝ 76% a 80%

Para aceder aos exercícios na plataforma SIACUA é necessário seguir os seguintes passos:

1. Utilizar o endereço https://siacua.web.ua.pt
2. Registar na plataforma
3. Entrar na plataforma – fazer Login
4. Escolher curso - Matemática Secundário
5. Clicar em Estatística, Cálculo Combinatório e Probabilidades – Estatística
6. Escrever o ID do exercício.

Capítulo 5: Monitorização dos Resultados

5.1. Breve caracterização do Meio e do Agrupamento

Caraterização do Meio

O concelho da Trofa está situado na região de Entre-Douro-e-Minho, no extremo norte do distrito do Porto, confronta a Sul e a Poente com os municípios da Maia e de Vila do Conde, pertencentes à Área Metropolitana do Porto, e a Norte e a nascente com os concelhos de Vila Nova de Famalicão e Santo Tirso.

O concelho está em expansão demográfica e urbanística com uma base económica marcada pela construção civil e pela indústria (têxtil a metalomecânica). Paralelamente, existe uma agricultura de subsistência e uma agricultura intensiva de produção de leite e de estufas hortícolas, disseminadas por todo o Concelho. O comércio existente situa-se em estabelecimentos nas ruas ou em pequenos e médios centros comerciais, que se localizam na sua maioria no centro deste município.

Os setores de prestação de serviços desenvolver-se em estabelecimentos diversos, com o principal objetivo de prestação de serviços a empresas (médicos, gabinetes de apoio técnico e jurídico, de arquitetura, de contabilidade, de estilismo, de engenharia, de modelismo), bem como de comércio grossista e retalhista e de natureza social, várias entidades bancárias, entre outros.

Em suma, o Concelho da Trofa apresenta-se como uma região que possui imensos recursos, moderno e capaz de responder aos desafios a que se propõe, mas que apresenta fragilidades estruturais e sociais, que urge ultrapassar.[18]

O Agrupamento

As Escolas do Agrupamento pretendem ser espaços educativos e culturais onde se privilegie a formação integral dos alunos, locais de trabalho e de valorização de todos, veículos de progresso humano e social da comunidade em que se inserem.

O Agrupamento é frequentado, neste momento, por cerca de 2800 alunos divididos pelos estabelecimentos de ensino que o compõem, apresentando em algumas escolas um razoável número de alunos e turmas em conformidade com a sua capacidade. Os alunos da Escola Secundária são, em regra, provenientes de todo o Concelho, quer da zona urbana, quer da zona rural, são oriundos de famílias em que os pais são predominantemente trabalhadores operários, possuindo um nível de escolaridade médio.

Recentemente, um número crescente de alunos oriundos de países estrangeiros tem vindo a frequentar as escolas do Agrupamento, nomeadamente no Pré-escolar e 1º Ciclo.

No Agrupamento são lecionados o Pré-escolar, 1º, 2º e 3º Ciclos do Ensino Básico e Ensino Secundário (Regular e Profissional) em horário diurno. Existem ainda três turmas de Cursos Vocacionais (3º Ciclo), uma turma de Educação e Formação de Adultos (em horário noturno) e a Escola Secundária é sede de um Centro para a Qualificação e Ensino Profissional.

Pretende-se continuar a criação destes cursos de forma a prevenir o abandono escolar, aumentar o sucesso e proporcionar aos alunos uma via profissionalizante, em resposta ao crescente desemprego que existe na nossa região. [18]

5.2. Breve caracterização da turma

A turma, por mim escolhida para trabalhar neste projeto foi a turma seis do décimo primeiro ano. As razões da minha opção foram essencialmente três: o tema que eu gostava de abordar era na área da Estatística e das turmas que estou a lecionar este ano é a que mais se adequava. O facto de terem sido meus alunos no ano letivo anterior, o que permitiu um melhor conhecimento dos mesmos, das suas capacidades e/ou dificuldades e, em terceiro lugar, o facto de eu ser a diretora de turma. É uma turma constituída por 19 alunos, 13 raparigas e 6 rapazes (Gráfico 5.1), com idades compreendidas entre os 15 e os 18 anos (Gráfico 5.2). É de salientar que, neste grupo de alunos, apenas 8 beneficiam de Ação Social Escolar.

Em termos de aproveitamento, é uma turma heterogénea, tem um grupo de alunos com muito bom rendimento escolar e um grupo de alunos que apresentava muitas dificuldades, nomeadamente à disciplina de MACS. No letivo anterior, 2018/2019, todos obtiveram classificação positiva na disciplina de MACS. As classificações dos alunos no final do ano letivo anterior foram as seguintes (Gráfico 5.3):

- **i)** entre 10 e 13 - 7alunos;
- **ii)** entre 14 e 17 - 7 alunos;
- **iii)** entre 18 e 20 - 5 alunos.

Apesar de o aproveitamento ser bom, alguns alunos demonstravam falta de atenção/concentração nas aulas, dificuldades na relacionação de conteúdos, dificuldades na compreensão oral e escrita, bem como no pensamento lógico/abstrato e no raciocínio matemático. Não há alunos repetentes.

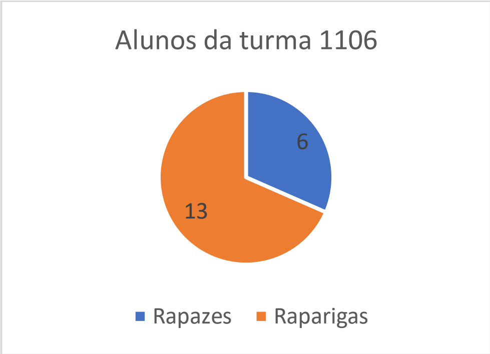


Gráfico 5.1 – Número de alunos

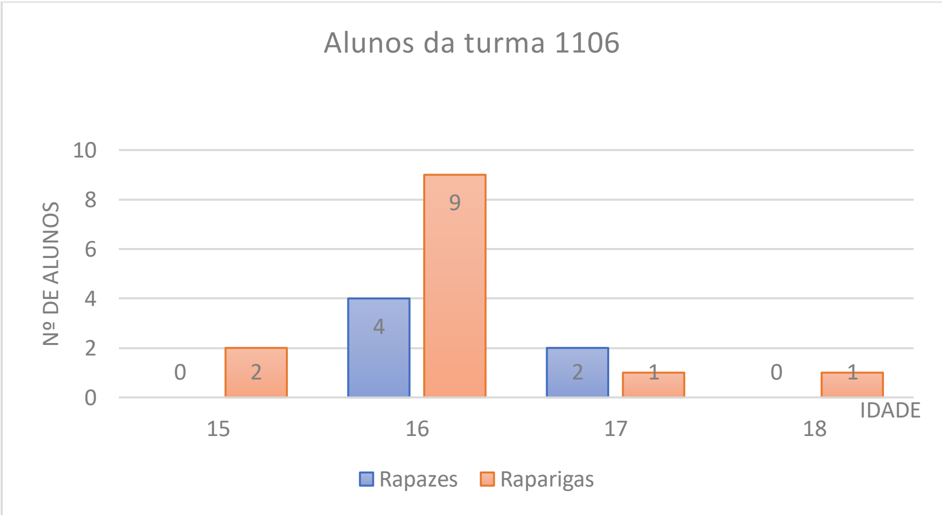


Gráfico 5.2 - Idade dos alunos

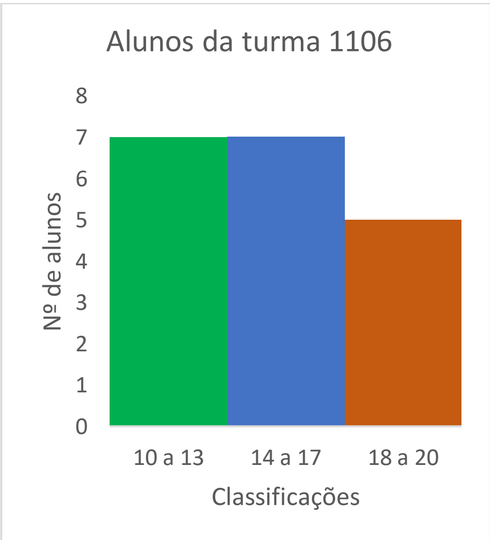


Gráfico 5.3 - Classificações dos alunos no ano letivo 2018/2019

5.3. Resultados do trabalho desenvolvido pela turma

No sentido de implementar o meu projeto na sala de aula, falei com o Diretor da escola explicando-lhe quais as dinâmicas e recursos que iria utilizar com os alunos.
Após a sua anuência, fui preparando os recursos (vídeos e exercícios de escolha múltipla) a alojar na plataforma SIACUA. Entretanto, falei com os alunos, perguntei-lhes se todos tinham *smartphone* com possibilidade de acederem à Internet, caso a da escola não estivesse a funcionar e expliquei-lhes a forma de como iriamos trabalhar alguns conteúdos do tema Inferência Estatística. Os alunos ficaram muito entusiamados, motivados, quiseram logo saber como poderiam aceder ao SIACUA. Expliquei-lhe também que o *smartphone* pode ser útil na sala de aula, quando bem utilizado. Os alunos foram inscritos na plataforma SIACUA, já na fase em que os recursos estavam quase prontos, para que pudessemos logo começar a utilizá-los, não deixando que os alunos perdessem o entusiamo e a curisosidade. Numa aula, ainda no tema anterior, comecei por verificar juntamente como os alunos, se todos conseguiam aceder à plataforma. Posteriormente, pedir-lhes que vissem o primeiro vídeo extra-aula, antes de começar o tema, pois esse vídeo tinha como objetivo rever os conceitos de Estatística lecionados no décimo ano. Na primeira aula que comecei esta temática da Inferência Estatística, projetei o vídeo para rever os conceitos e verifiquei que a maior parte dos alunos já o tinha visto. Os alunos acharam engraçado o facto de o vídeo conter a voz da professora. Nessa mesma aula resolveram, utilizando o *smartphone,* os três primeiros exercícios. Disse-lhes que deveriam em casa rever o que foi trabalhado na aula. Nas aulas seguintes foi utilizada a mesma estratégia, projetei o vídeo referente ao conteúdo a lecionar, trabalhamos os conteúdos e seguidamente os alunos resolviam o ou exercícios do SIACUA refentes a esse conteúdo, assim como os do manual escolar. Na quinta aula e última do tema resolveram unicamente exercícios do SIACUA, que tinham sido já propostos para serem realizados extra-aula. Não houve contatempos, todo correu como o previsto, embora em algumas das aulas os alunos tivessem que usar a sua Internet, pois nem sempre a mesma funciona na escola de uma forma aceitável. Salienta-se que, houve uns dias que os alunos entravam no SIACUA, conseguiam visualizar os vídeos, mas não conseguiam visualizar os exercícios. Passados uns

dias, voltei a pedir para acederam novamente aos exercícios e já tudo funcionava corretamente. [5]

Seguidamente apresenta-se duas tabelas com os acessos dos alunos à plataforma SIACUA.

2020-01-01 2020-02-24 macs11

VER

Mostrar 10 linhas por página Pesquisar:

Num	Número de ações	Número de logins PC	Número de logins Mobile	Escolheu OA	Escolheu MEGUA/SIACUA	Escolheu PmatE	Número de respostas certas	Número de respostas erradas	Número de visualizações de questões MEGUA/SIACUA	Número de visualizações de questões PmatE
1	85			4	17		9	1	18	
2	45			1	9		6	1	9	
3	42			3	10		3		10	
4	101			6	16		12	2	18	
5	51			4	6		4	1	8	
6	68			4	14		5		16	
7	111			4	24		14		24	
8	41			4	8		1		9	
9	113			2	23		17	6	23	
10	17			2	2		1		3	

Tabela 5.3 - 1 : Acessos ao SIACUA - primeira metade da turma 1106

2020-01-01 2020-02-24 macs11

VER

Mostrar 10 linhas por página Pesquisar:

Num	Número de ações	Número de logins PC	Número de logins Mobile	Escolheu OA	Escolheu MEGUA/SIACUA	Escolheu PmatE	Número de respostas certas	Número de respostas erradas	Número de visualizações de questões MEGUA/SIACUA	Número de visualizações de questões PmatE
11	4									
12										
13	83			4	15		11	3	15	
14	147			5	27		46	5	20	
15	35			4	4		2		4	
16	1									
17	79			6	10		8	3	12	
18	35			2	8		3		5	
19	134			4	26		15		27	
20	20			1	3		2		3	

Tabela 5.3 - 2 : Acessos ao SIACUA - segunda metade da turma 1106

Pela análise das Tabelas 5.3 – 1 e Tabelas 5.3 – 2, é possível afirmar que os alunos de uma forma geral acederam muitas vezes à plataforma SIACUA e tentaram resolverem os exercícios.

Os alunos realizaram a avaliação escrita destes conteúdos lecionados sobre Inferência Estatística no dia 09 de março. As escolas encerraram a 13 de março, devido à situação da pandemia e não foi realizado qualquer inquérito aos alunos sobre a utilização da Plataforma SIACUA nas aulas de MACS. No entanto, posso dizer que o teste de avaliação versava vários conteúdos lecionados e foi dedicado um grupo de três questões à Inferência Estatística, como se pode ver na Figura 5.3 – 1.

5. No dia de abertura do novo *resort* turístico *2MuchDreams*, na Vila da Cusquice, os clientes dividiram-se pelas seguintes atividades de animação: *karaoke*, voleibol de praia e aulas de aeróbica.
Na tabela que se segue apenas estão registados os números relativos às presenças no *karaoke* e no voleibol de praia.
Cada cliente só podia optar por uma das três atividades.

	Karaoke	**Voleibol de praia**
Homens	26	45
Mulheres	54	36

Tendo por referência apenas o número total de clientes apresentado na tabela, responda às questões que se seguem.

5.1. Construa um intervalo de confiança de 95% para estimar a proporção de homens presentes, por dia, no *karaoke*, no decurso da primeira semana de funcionamento do *resort*.
Relativamente aos valores dos extremos do intervalo, apresente o resultado em percentagem, arredondado às décimas.
Caso proceda a arredondamentos nos cálculos intermédios, conserve, no mínimo, quatro casas decimais.

5.2. Determine, com aproximação às centésimas, a margem de erro associada ao intervalo da alínea anterior.

5.3. Para um grau de confiança de 90%, qual deverá ser a dimensão da amostra, para que o erro cometido na estimação da proporção de homens presentes, por dia, no *karaoke*, não seja superior a 3 pontos percentuais?
Nos cálculos intermédios utilize duas casas decimais.

Figura 5.3 - 1 - Exercício do Teste de Avaliação MACS - 09 março 2020

Na pergunta 5.1. dos 19 alunos da turma 16 alunos tiveram a resposta completamente certa, ou seja, cerca de 84% dos alunos. Na pergunta 5.2, 17 alunos tiveram a resposta complemente certa, ou seja, cerca de 89% e na pergunta 5.3, 14 alunos tiveram a resposta completamente certa, ou seja, cerca de 74%. Considero que os resultados foram

francamente positivos e posso ainda a acrescentar que foi o grupo do teste onde os resultados foram mais altos em termos percentuais das respostas completamente certas.
A última vez que lecionei MACS de 11º ano foi há cinco anos, e do que me lembro, considero que estes resultados foram bem mais elevados do que os que habitualmente costumava ter.
O facto de ter utilizado uma estratégia mais motivadora, inovadora contribuiu bastante para que os alunos mostrassem mais empenho, interesse e dedicação às atividades propostas no SIACUA.[5]

Capítulo 6: CONCLUSÃO

A realização deste trabalho teve como principal objetivo a construção de recursos digitais, que pudessem ser utilizados quer pelos estudantes no seu estudo autónomo, quer como ferramenta a ser usada pelos Professores na lecionação das suas aulas, no tema Inferência Estatística. A Escola está em plena transformação e torna-se impensável continuar com um ensino puramente tradicional sem acesso a novas formas de aprender. A Escola tem de implementar mecanismos para motivar os alunos, para uma aprendizagem que promova o sucesso e cada Professor tem de contribuir na sua sala de aula para atingir tal meta. Como Professora, ao longo dos anos tenho tentado adquirir mais formação para poder adequar as minhas estratégias pedagógicas ao nível, realidade e interesse dos alunos. Só assim poderemos contribuir para que os jovens tenham uma formação que lhes dê a capacidade de enfrentar os desafios do século XXI. Os jovens de hoje têm de estar preparados não só para serem cidadãos portugueses, mas também cidadãos do mundo. A construção de materiais digitais, vídeos e exercícios de escolha múltipla com correção automática e com a possibilidade de verem a resolução adequada não pretende substituir de todo os Professores e/ou os manuais. O Professor continua e continuará a ser importante, na minha opinião, sobretudo na orientação do caminho a seguir pelos seus alunos, por muitas plataformas e vídeos explicativos que os mesmos utilizem. A construção e utilização de recursos digitais pode ser sem dúvida, um forte aliado e complemento aos manuais e à estratégia a utilizar em sala de aula.

No entanto, para que seja possível aos professores a utilização destas estratégias nas suas práticas pedagógicas é necessário que as escolas estejam mais bem apetrechadas de meios informáticos e, sobretudo, de uma boa rede para acesso à Internet, para que os alunos não tenham de recorrer à sua Internet particular. Por vezes a extensão dos programas pode ser também um entrave à utilização sistemática deste tipo de recursos digitais.

Todos os recursos, produzidos no âmbito deste projeto estão alojados na plataforma SIACUA, podendo ser utilizados e explorados, bastando fazer o registo gratuito nesta plataforma. [5]

Nesta dissertação foram descritos todas as tarefas e procedimentos executados para a obtenção dos vídeos e dos exercícios. No entanto, apesar de não refletir todo o trabalho desenvolvido nesta investigação, pode servir de base para a construção de mais materiais no âmbito deste tema, porque nem todos os conteúdos foram explorados com recurso a materiais digitais.

Ao longo dos meus quase trinta anos de carreira tenho feito muita formação, mas nada se compara com este projeto. Desenvolver este trabalho foi uma grande tarefa e ao mesmo tempo um grande desafio, pelo desconhecimento total sobre programação, pela necessidade de fazer muitas consultas e muitas pesquisas e de escrever um documento com a qualidade necessária. O enriquecimento pessoal e profissional, constitui uma forte e significativa valorização para o exercício da minha atividade como professora. Espero, conseguir continuar a motivar e a orientar os meus alunos para uma aprendizagem significativa em Matemática, mas também contribuir para a formação integral dos jovens como seres humanos capazes de enfrentar os desafios do futuro em qualquer parte do mundo.

Referências Bibliográficas

[1] R. Aharoni, *Aritmética para pais*, Lisboa: Gradiva, 2008.

[2] M. d. Educação, "Direção Geral da Educação," 2017. URL: https://www.dge.mec.pt/sites/default/files/Curriculo/Projeto_Autonomia_e_Flexibilidade/perfil_dos_alunos.pdf.

[3] M. d. Educação, "Direção Geral da Educação," 2017. URL:https://www.dge.mec.pt/sites/default/files/Curriculo/Projeto_Autonomia_e_Flexibilidade/ae_documento_enquadrador.pdf.

[4] M. Prensky, "*De On the Horizon (NCB University Press), Vol. 9",* 5 out 2001. URL:http://crisgorete.pbworks.com/w/file/fetch/58325978/Nativos.pdf

[5] "Plataforma Siacua," 2019. URL: www.siacua.web.ua.pt.

[6] M. d. Educação, "Direção Geral de Educação," 2017. URL:https://www.dge.mec.pt/sites/default/files/Curriculo/Projeto_Autonomia_e_Flexibilidade/11_macs.pdf.

[7] A. C. Doyle, *Um escândalo na Boémia, a liga dos "cabeça vermelho*", coleção Sherlock Holmes, Lisboa: Verso da História, 2006.

[8] P. Grima , *Os Segredos da Estatística,* National Geographic, 2018.

[9] B. Murteira e M. Antunes , *Probabilidades e Estatística*, Escolar Editora, 2013.

[10] D. Hoaglin e D. Moore*, Perspectives on Contemporary Statistics*, Mathematical Association of America, 1992.

[11] E. Reis, P. Melo, R. Andrade e T. Calapaz, *Estatística Aplicada,* Vol.2, Lisboa: Edições Sílabo, 1997.

[12] B. Murteira, C. Ribeiro, J. Silva, C. Pimenta e F. Pimenta, *Introdução à Estatística*, Lisboa: Escolar Editora, 2010,2015.

[13] D. D. Pestana e S. F. Velosa, *Introdução à Probabilidade e Estatística*, Lisboa: Fundação Calouste Gulbenkian, 2010.

[14] R. Campos Guimarães e J. A. Sarsfield Cabral, *Estatística,* Alfragide: Mc Graw Hill, 1998.

[15] E. Longo e I. Branco, *MACS 11º Ano,* Lisboa: Texto Editores, 2016.

[16] M. A. F. Neves, L. Faria e B. Ribeiro, *Máximo 11 - MACS*, Porto: Porto Editora, 2016.

[17] IAVE, "Instituto de Inovação Educativa," vários.
URL: http://iave.pt/index.php/avaliacao-de-alunos/arquivo-de-provas-exames/exemplo-arquivo.

[18] A. Escolas da Trofa, "Projeto Educativo," julho 2019.
URL: http://aetrofa.com/projeto-educativo/.

Apêndices

1-Vídeos

De seguida, apresentam-se os diapositivos relativamente aos vídeos, as programações dos textos que foram alojados com os vídeos no SIACUA.

Vídeo ①: Introdução à Inferência Estatística – Revisão de conceitos

Texto colocado junto ao vídeo:

"INTRODUÇÃO À INFERÊNCIA ESTATÍSTICA

REVISÃO DE CONCEITOS

Neste objeto de aprendizagem iremos rever alguns conceitos de estatística:

- Estatística Descritiva e Estatística Indutiva
- População, unidade estatística e dimensão da amostra
- Técnicas para a escolha de uma amostra

Segue-se um vídeo explicativo**:"**

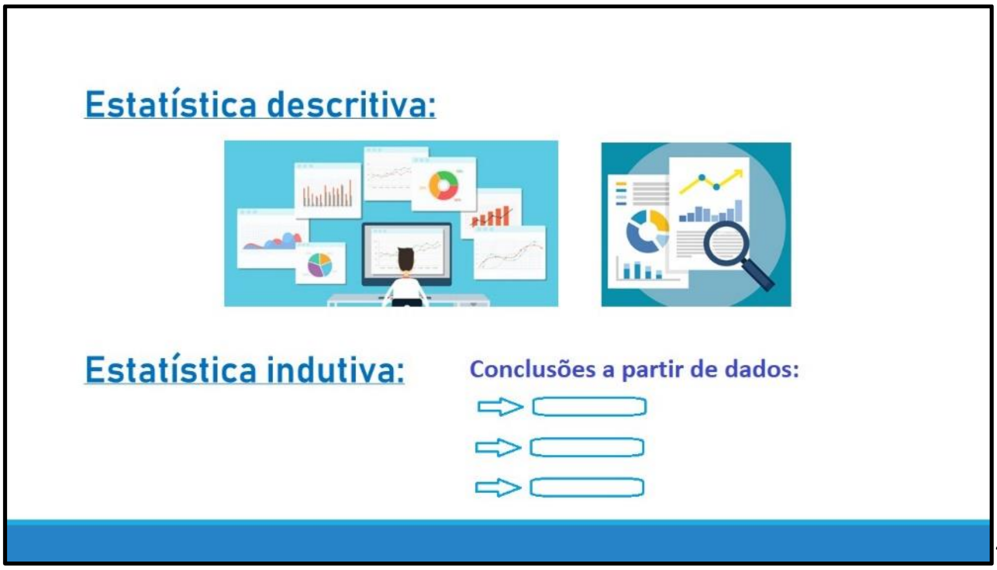
Programação:

```
<h1>Introdução à Inferência estatística</h1>
<br><br>
<h3>Revisão de Conceitos</h3>

<p>Neste objeto de aprendizagem iremos rever alguns <b>conceitos de
estatística:</b></p>

<ul>
<li>Estatística Descritiva e Estatística Indutiva</li>
<li>População, unidade estatística e dimensão da amostra</li>
```

```
<li>Técnicas para a escolha de uma amostra</li>
</ul>
 <p>Segue-se um <strong><u>vídeo explicativo</u></strong>:</p>
<iframe width="560" height="315"
src="https://www.youtube.com/embed/ZbVoddxmgms" frameborder="0"
allow="accelerometer; autoplay; encrypted-media; gyroscope; picture-in-
picture" allowfullscreen></iframe>
```

Diapositivos:

Figura A.1: Vídeo 1 - imagem 1



[4]

Figura A.2: Vídeo 1 - imagem 2

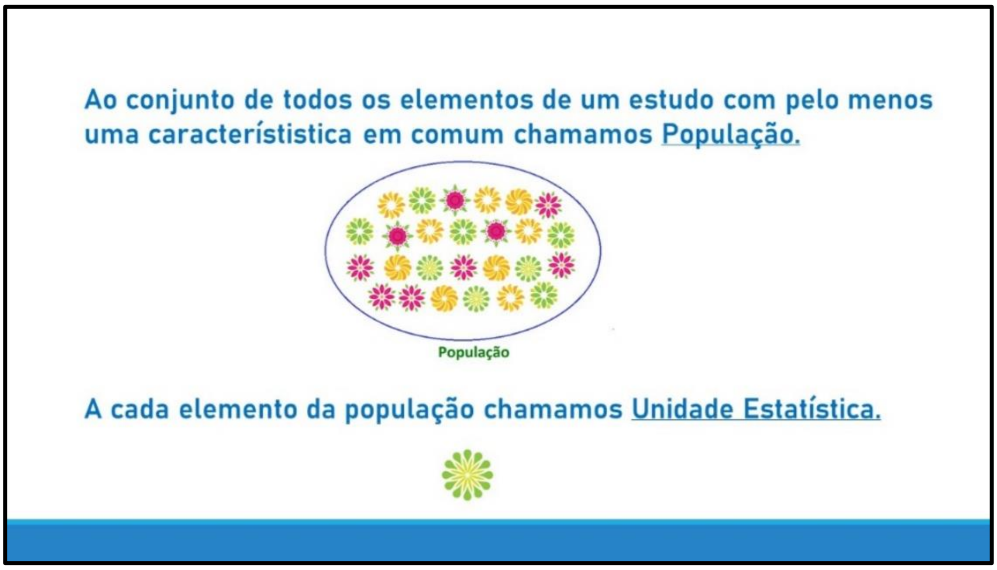
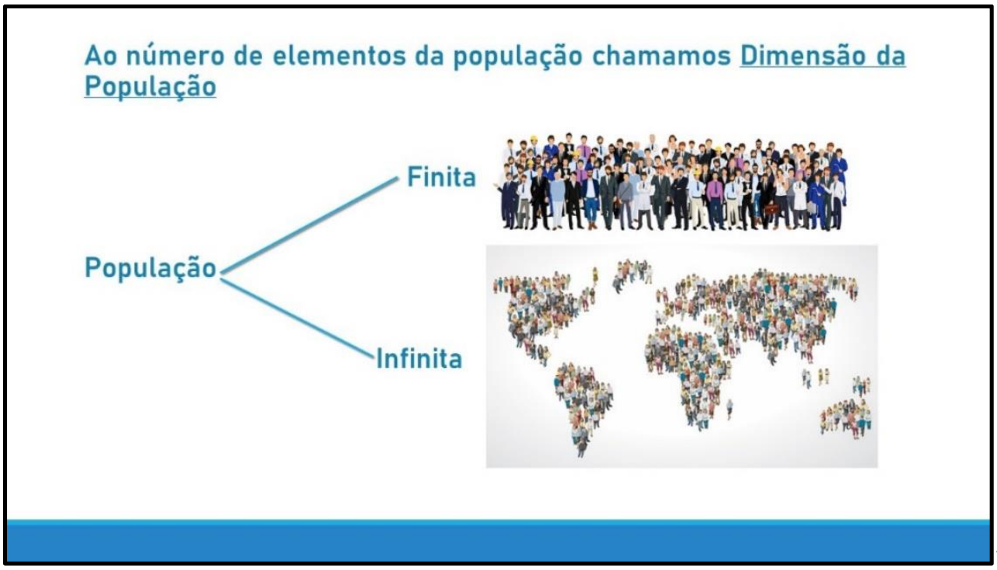


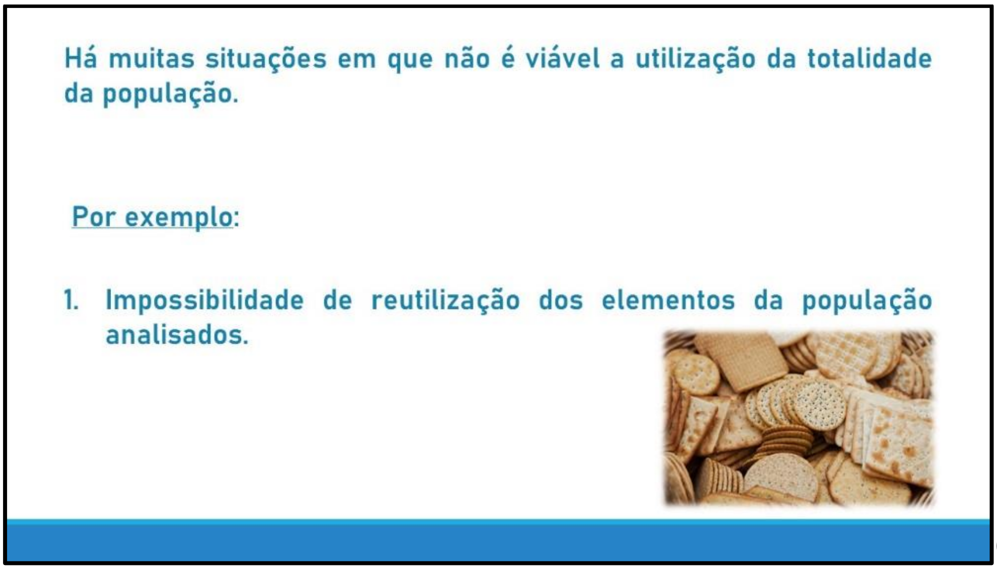
Figura A.3: Vídeo 1 - imagem 3

[4] Imagens retiradas:
https://seputarilmu.com/wp-content/uploads/2020/03/sim-scaled.jpg
https://portalmultempreendedor.com.br/blog/wp-content/uploads/2017/06/4-As-do-marketing-an%C3%A1lise-300x206.jpg



[5]

Figura A.4: Vídeo 1 - imagem 4



[6]

Figura A.5: Vídeo 1 - imagem 5

[5] Imagens retiradas:
https://lh3.googleusercontent.com/proxy/mRGfrYqe0CjUM6nxMk9B6giJ_lOILFmZVE1wpCChbRFg33k_5EPB31qiUjlP7pLzHSzVD4yyncsuC4kfJ6s_xjrQa708S3T0BGe9wyDyvT6AT_VzViRXJ-vjWnvHU2Yp-xlkNoGeY_cjKX3_pw6Qh9IL2w
https://api.ndla.no/image-api/raw/id/44057

[6] Imagem retirada:
https://media.gettyimages.com/photos/close-up-of-a-large-selection-of-cream-crackers-picture-id731743953?s=612x612

[7]

Figura A.6: Vídeo 1 - imagem 6

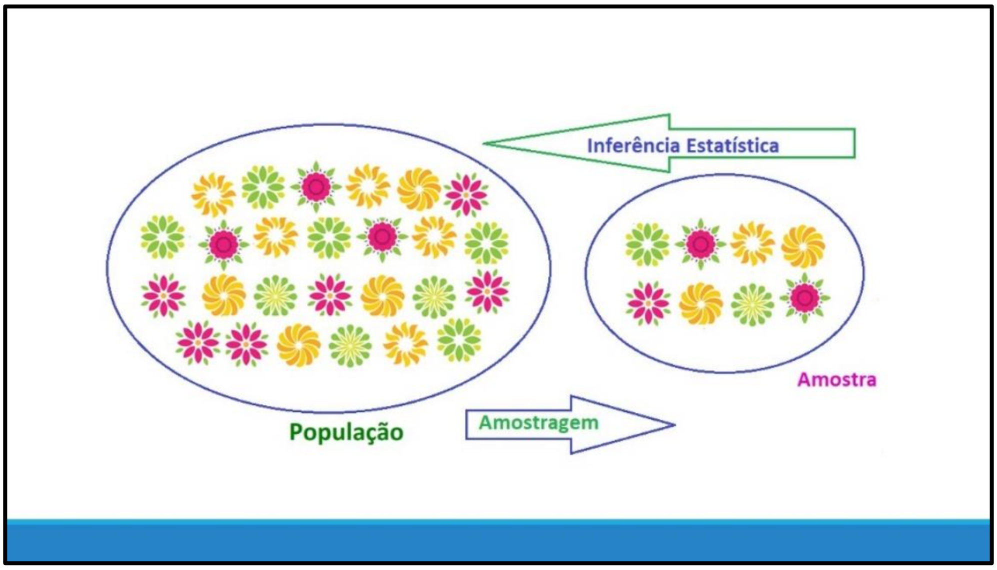
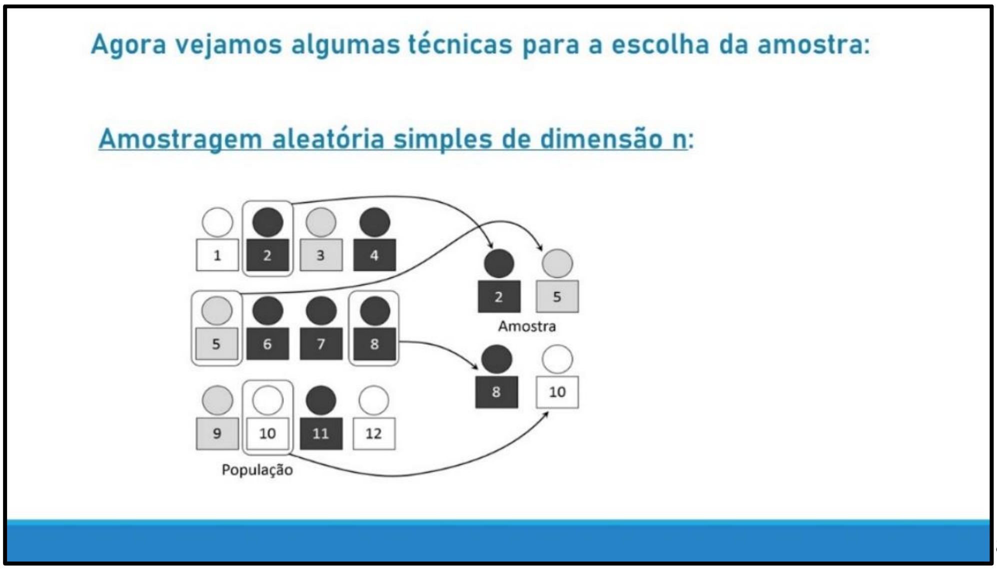


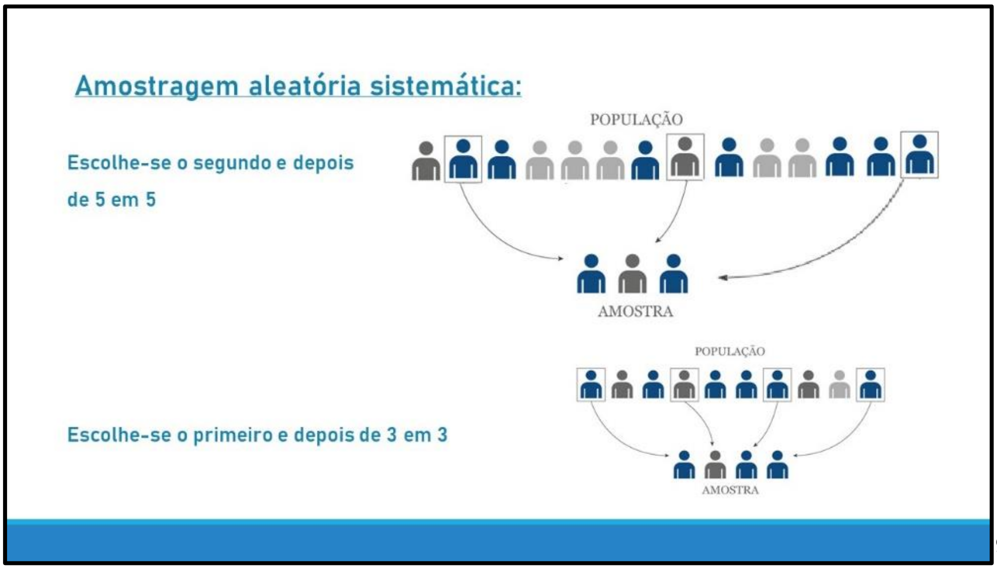
Figura A.7: Vídeo 1 - imagem 7

[7] Imagens retiradas:
https://a-static.mlcdn.com.br/618x463/gravador-usb-flash-eprom-spi-bios-series-24xx-25xx-atmega/tecinformaticavirtual/6073191258/9f2084bca196e51adbe0470cb7c0041b.jpg
https://encrypted-tbn0.gstatic.com/images?q=tbn%3AANd9GcQVidQPgz5IM6-cn0IxEC-zPU155tJ8Rj93LOGP5TCy-q4N2Kee&usqp=CAU



[8]

Figura A.8: Vídeo 1 - imagem 8

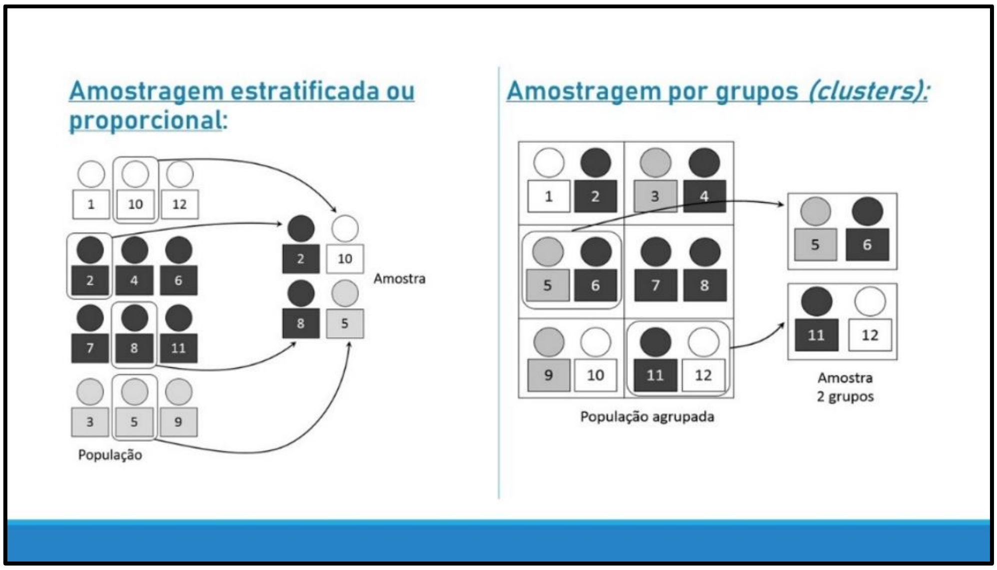


[9]

Figura A.9: Vídeo 1 - imagem 9

[8] Imagem retirada:
https://docplayer.com.br/docs-images/104/165030973/images/8-2.jpg

[9] Imagens retiradas:
https://miro.medium.com/max/1400/1*YAt01yr9Bnba_yg3KORCHw.gif



[10]

Figura A.10: Vídeo 1 - imagem 10

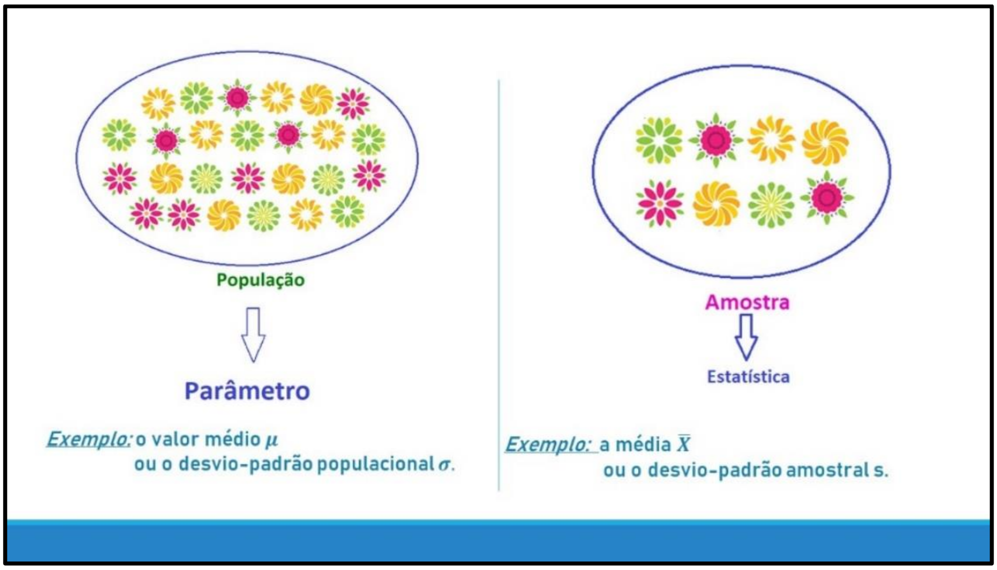
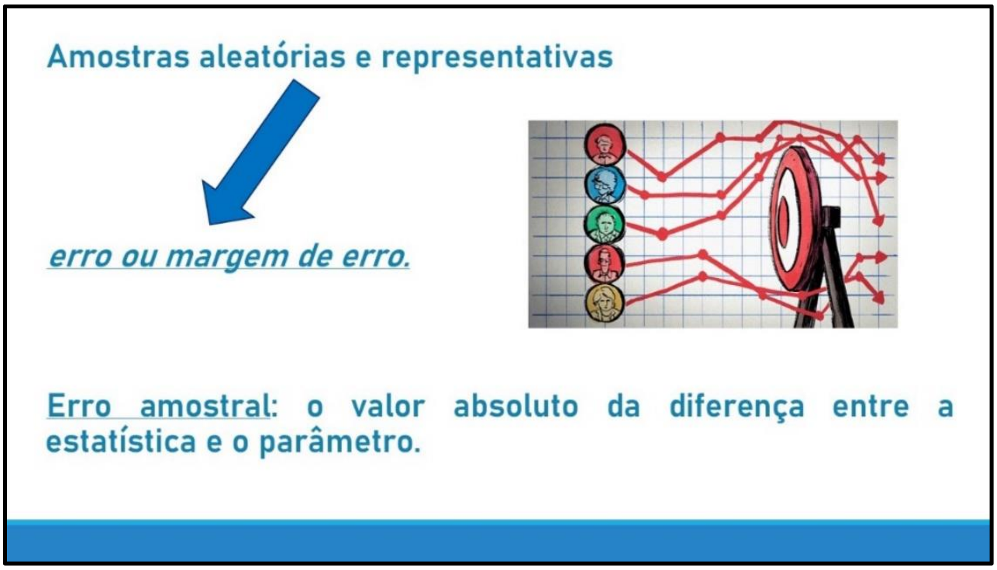


Figura A.11: Vídeo 1 - imagem 11

[10] Imagens retiradas:
https://docplayer.com.br/docs-images/104/165030973/images/8-2.jpg
e
https://www.netquest.com/hubfs/Imported_Blog_Media/amostra-2.png



[11]

Figura A.12: Vídeo 1 - imagem 12

[11] Imagem retirada: https://video-images.vice.com/articles/5bc75b67e752c20006d9909e/lede/1539890999553-pesquisas-eleicoes.jpeg?crop=0.562xw:1xh;0.182xw,0xh

Vídeo ②: Estimação de parâmetros

<u>Texto colocado junto ao vídeo</u>:

"INTRODUÇÃO À INFERÊNCIA ESTATÍSTICA

ESTIMAÇÃO DE UM PARÂMETRO

Neste objeto de aprendizagem, iremos introduzir o conceito de **parâmetro** e o que é estimar parâmetros.
Segue-se um **<u>vídeo explicativo</u>**, com dois exemplos ilustrativos:"

<u>Programação</u>:

```
<h1>Introdução à Inferência Estatística</h1>

<br><br>


<h2>Estimação de um parâmetro</h2>


<p>Neste objeto de aprendizagem, iremos introduzir o conceito de
<b>parâmetro</b> e o que é estimar parâmetros.</p>


 <p>Segue-se um <strong><u>vídeo explicativo</u></strong>, com dois
exemplos ilustrativos:</p>


<br>

<iframe width="560" height="315"
src="https://www.youtube.com/embed/NqNijLA1qHk" frameborder="0"
allow="accelerometer; autoplay; encrypted-media; gyroscope; picture-in-
picture" allowfullscreen></iframe>


<br><br><br><br>
```

Diapositivos:

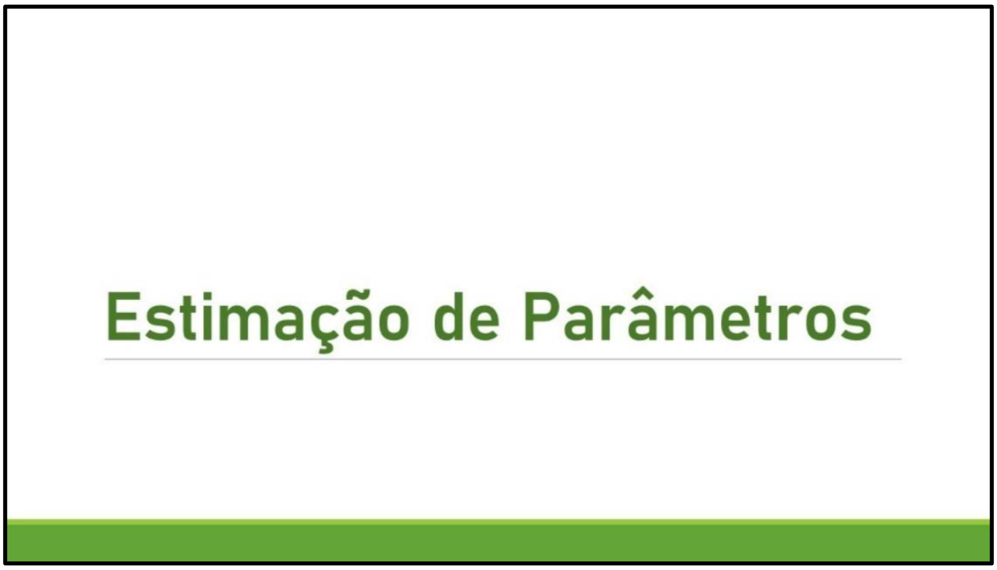


Figura B.1- Vídeo 2 - imagem 1

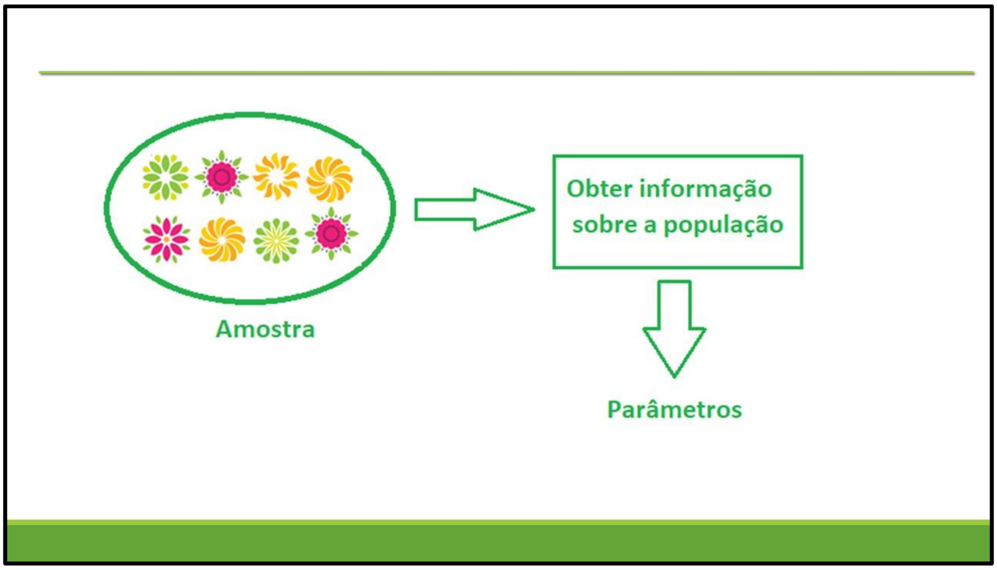
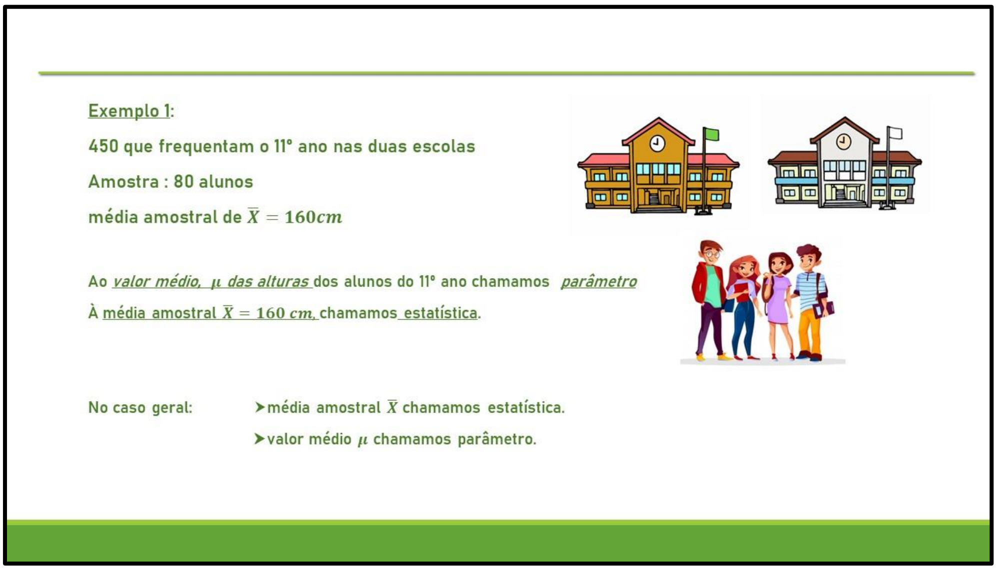


Figura B.2: Vídeo 2 - imagem 2



12

Figura B.3 :Vídeo 2 - imagem 3

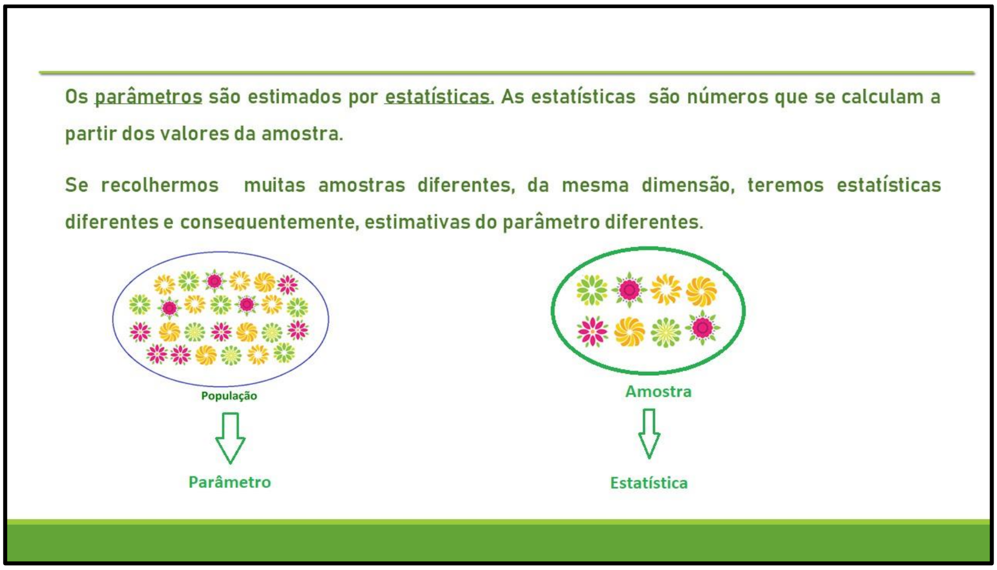


Figura B.4: Vídeo 2 - imagem 4

[12] Imagens retiradas:
https://cdn5.colorir.com/desenhos/color/201916/escola-edificios-outros-edificios-1523570.jpg
e
https://cdn5.colorir.com/desenhos/color/201523/escola-edificios-outros-edificios-pintado-por-vitorcely-1091354.jpg
e
https://www.clipartkey.com/mpngs/m/202-2026458_college-student-png-vector.png

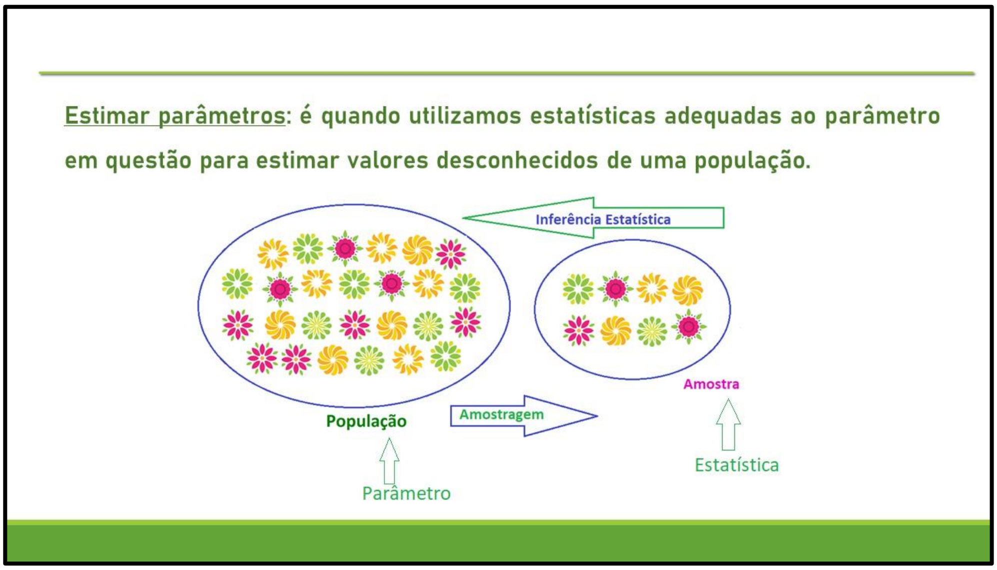


Figura B.5: Vídeo 2 - imagem 5

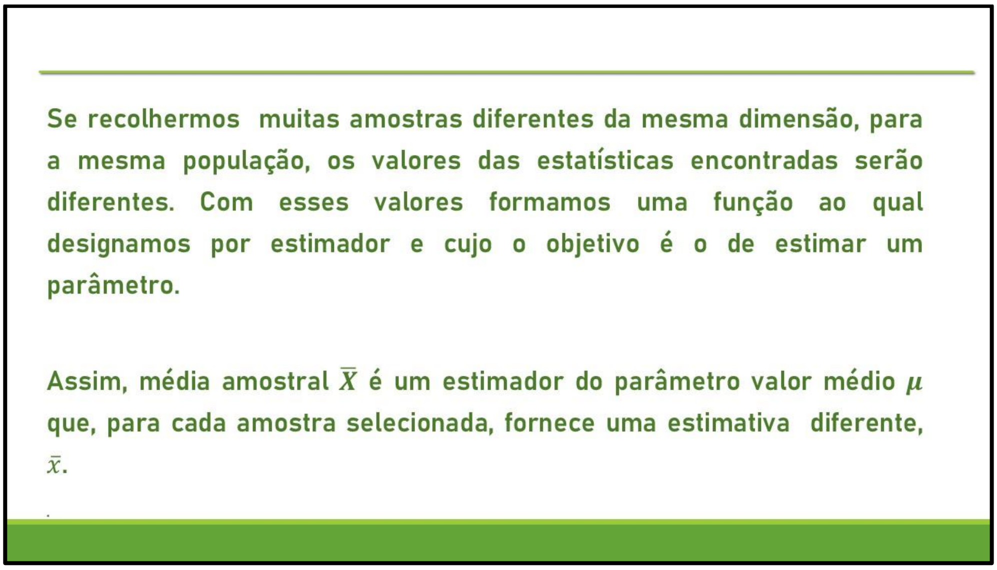
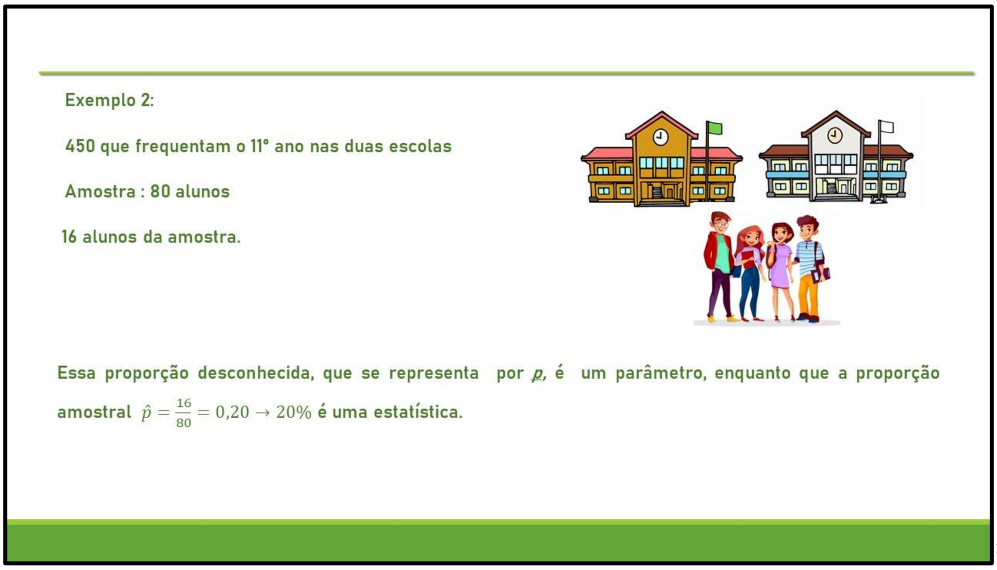


Figura B.6: Vídeo 2 - imagem 6



[13]

Figura B.7: Vídeo 2 - imagem 7

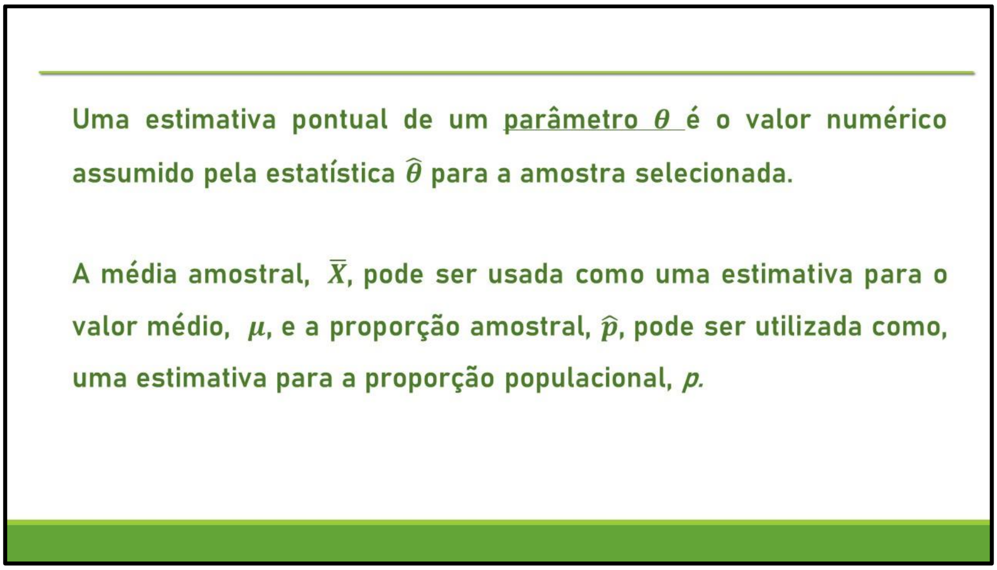


Figura B.8: Vídeo 2 - imagem 8

[13] Imagens retiradas:
https://cdn5.colorir.com/desenhos/color/201916/escola-edificios-outros-edificios-1523570.jpg
e
https://cdn5.colorir.com/desenhos/color/201523/escola-edificios-outros-edificios-pintado-por-vitorcely-1091354.jpg
e
https://www.clipartkey.com/mpngs/m/202-2026458_college-student-png-vector.png

Vídeo ③: Intervalo de Confiança para o Valor Médio

Texto colocado junto ao vídeo:

“INTERVALO DE CONFIANÇA PARA A MÉDIA:

Agora iremos abordar o conceito de intervalo de confiança, neste caso, para a média.

O intervalo de confiança é dado pela seguinte expressão:

$$\left]\bar{X} - z\frac{\sigma}{\sqrt{n}}, \bar{X} + z\frac{\sigma}{\sqrt{n}}\right[$$

Os graus de confiança mais utilizados são:

Nível de confiança	90%	95%	99%
z	1,645	1,960	2,576

Tabela 4.2 - 1: Níveis de Confiança

Segue-se um **vídeo explicativo**:”

Programação:

```
<h2>Intervalo de Confiança para a média:</h2>

<p>Agora iremos abordar o conceito de <b>intervalo de confiança</b>,
neste caso, para a média. </p>

<p><b>O intervalo de confiança</b> é dado pela seguinte expressão:</p>

\[\left]\overline{X}-z\times\frac{\sigma}{\sqrt{n}},
\overline{X}+z\times\frac{\sigma}{\sqrt{n}}\right[\]
<p>Os graus de confiança mais utilizados são: </p>
<Table border=1>
<tr>
      <td align="center" width=100 bgcolor=White><b>Nível de
confiança</b></td>
      <td align="center" width=60 bgcolor=White><b>90&percnt; </b></td>
      <td align="center" width=60 bgcolor=White><b>95&percnt; </b></td>
```

```
      <td align="center" width=60 bgcolor=White><b>99&percnt; </b></td>
</tr>
<tr>
       <td align="center" width=100 bgcolor=White><b>z</b></td>
       <td align="center" width=60 bgcolor=White>1,645</td>
       <td align="center" width=60 bgcolor=White>1,960</td>
       <td align="center" width=60 bgcolor=White>2,576</td>
      </tr>
</table>
<br>
<br>

 <p>Segue-se um <strong><u>vídeo explicativo</u></strong>:</p>
<iframe width="560" height="315"
src="https://www.youtube.com/embed/p0ERib71v4g" frameborder="0"
allow="accelerometer; autoplay; encrypted-media; gyroscope; picture-in-
picture" allowfullscreen></iframe>

<br><br><br><br>
```

Diapositivos:

[14]

Figura C.1: Vídeo 3 - imagem 1

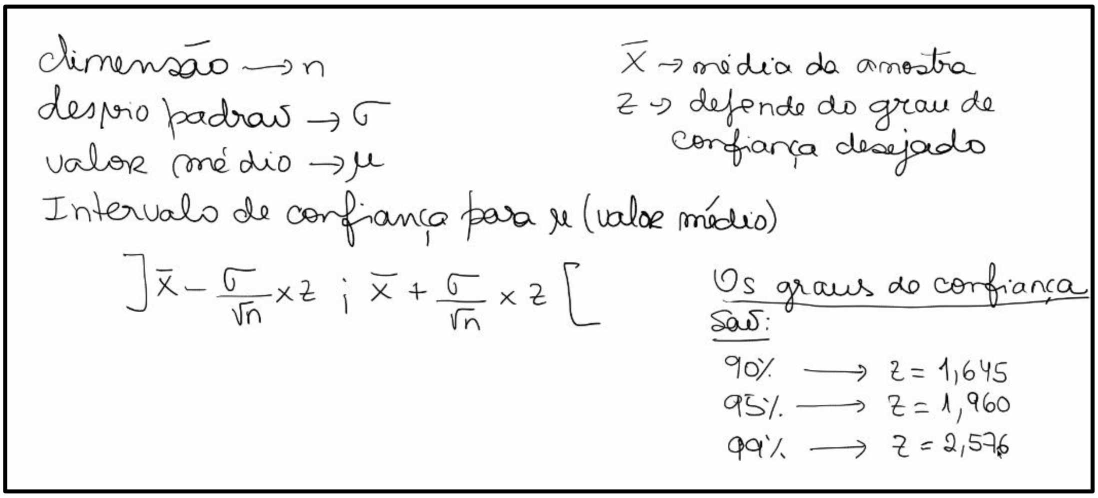
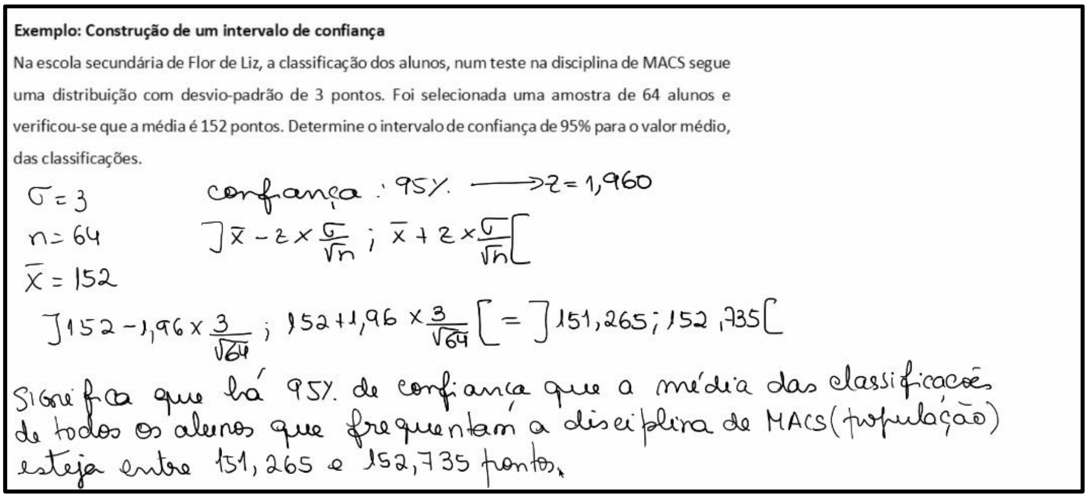


Figura C.2: Vídeo 3 - Imagem 2

[14] Imagens retiradas:
https://image.freepik.com/vetores-gratis/professor-ou-cientista-personagem-de-desenho-animado-com-chapeu-de-pos-graduacao-segurando-um-ponteiro_20412-147.jpg
e
https://lh3.googleusercontent.com/proxy/qX4N1nA3POiVeSJiTrSQX8enOHvehz2ppCTuwAUbZBrMzr1SxyXiWHk3ARDkem0ULrNoJUkh_HwU_ixc-1JS5zGwVg0t4hG09cSToxjiyWZYMoMqX4KwzlaMqWWh6Zuy0nLLe0Nbz0bp



Exemplo: Construção de um intervalo de confiança

Na escola secundária de Flor de Liz, a classificação dos alunos, num teste na disciplina de MACS segue uma distribuição com desvio-padrão de 3 pontos. Foi selecionada uma amostra de 64 alunos e verificou-se que a média é 152 pontos. Determine o intervalo de confiança de 95% para o valor médio, das classificações.

$\sigma = 3$

$n = 64$

$\bar{x} = 152$

confiança: 95% $\longrightarrow z = 1,960$

$\left]\bar{x} - z \times \frac{\sigma}{\sqrt{n}} ; \bar{x} + z \times \frac{\sigma}{\sqrt{n}}\right[$

$\left]152 - 1,96 \times \frac{3}{\sqrt{64}} ; 152 + 1,96 \times \frac{3}{\sqrt{64}}\right[= \left]151,265 ; 152,735\right[$

Significa que há 95% de confiança que a média das classificações de todos os alunos que frequentam a disciplina de MACS (população) esteja entre 151,265 e 152,735 pontos.

Figura C.3: Vídeo 3 - Imagem 3

Vídeo ④: Intervalo de Confiança para a Proporção

<u>Texto colocado junto ao vídeo</u>:

"INTERVALO DE CONFIANÇA PARA A PROPORÇÃO:

Também podemos determinar o intervalo de confiança para a proporção.

O intervalo de confiança é dado pela seguinte expressão:

$$\left]\hat{p}-z\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}};\hat{p}+z\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\right[$$

Os graus de confiança mais utilizados são:

Nível de confiança	90%	95%	99%
z	1,645	1,960	2,576

Tabela 4.2 – 2: Níveis de Confiança

Segue-se um <u>vídeo explicativo</u>**:**"

<u>Programação</u>:

```
<h2>Intervalo de confiança para a proporção</h2>

<p>Também podemos determinar o <b>intervalo de confiança</b> para a
proporção. </p>

<p><b>O intervalo de confiança</b> é dado pela seguinte expressão:</p>

\[\left]\hat{p}-z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}};
 \hat{p}+z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}\right[\]

<p>Os graus de confiança mais utilizados são: </p>
<Table border=1>
```

```
<tr>
      <td align="center" width=100 bgcolor=White><b>Nível de
confiança</b></td>
      <td align="center" width=60 bgcolor=White><b>90&percnt; </b></td>
      <td align="center" width=60 bgcolor=White><b>95&percnt; </b></td>
      <td align="center" width=60 bgcolor=White><b>99&percnt; </b></td>
</tr>
<tr>
       <td align="center" width=100 bgcolor=White><b>z</b></td>
       <td align="center" width=60 bgcolor=White>1,645</td>
       <td align="center" width=60 bgcolor=White>1,960</td>
       <td align="center" width=60 bgcolor=White>2,576</td>
      </tr>
</table>
<br>
<br>

 <p>Segue-se um <strong><u>vídeo explicativo</u></strong>:</p>

<iframe width="560" height="315"
src="https://www.youtube.com/embed/hri9S0EQcjg" frameborder="0"
allow="accelerometer; autoplay; encrypted-media; gyroscope; picture-in-
picture" allowfullscreen></iframe>
```

Diapositivos:

[15]

Figura D.1: Vídeo 4 - imagem 1

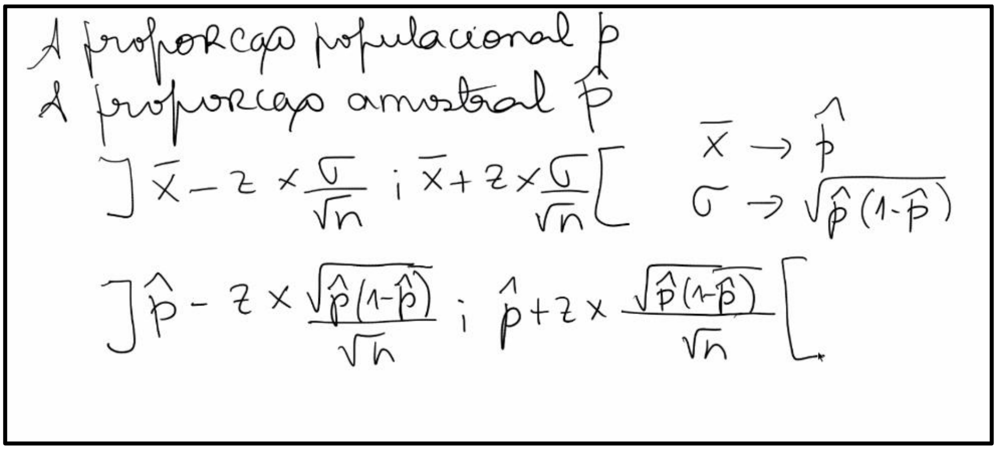
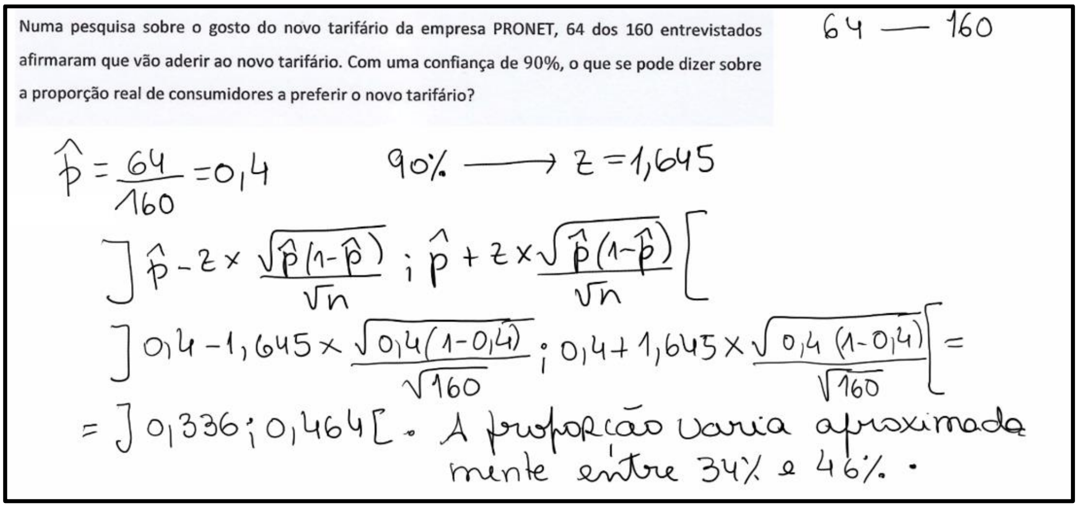


Figura D.2: Vídeo 4 - imagem 2

[15] Imagens retiradas:
https://i.pinimg.com/736x/04/01/7e/04017ee1ee1301e183d65021f45d81a2.jpg
e
https://www.onoticiasdatrofa.pt/wp-content/uploads/2013/01/esta%C3%A7%C3%A3o-Nova.jpg



Numa pesquisa sobre o gosto do novo tarifário da empresa PRONET, 64 dos 160 entrevistados afirmaram que vão aderir ao novo tarifário. Com uma confiança de 90%, o que se pode dizer sobre a proporção real de consumidores a preferir o novo tarifário?

64 — 160

$\hat{p} = \frac{64}{160} = 0{,}4$ $\qquad 90\% \longrightarrow z = 1{,}645$

$\left]\hat{p} - z \times \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}} ; \hat{p} + z \times \frac{\sqrt{\hat{p}(1-\hat{p})}}{\sqrt{n}}\right[$

$\left]0{,}4 - 1{,}645 \times \frac{\sqrt{0{,}4(1-0{,}4)}}{\sqrt{160}} ; 0{,}4 + 1{,}645 \times \frac{\sqrt{0{,}4(1-0{,}4)}}{\sqrt{160}}\right[=$

$= \left]0{,}336 ; 0{,}464\right[$. A proporção varia aproximadamente entre 34% e 46%.

Figura D.3: Vídeo 4 - imagem 3

2. Exercícios de Escolha Múltipla

De seguida, apresentam-se as programações dos exercícios que foram alojados no SIACUA.

Exercício 1.1:

```
De entre os $3000$ alunos de uma escola selecionaram-se alguns para responder a um inquérito sobre o programa de televisão preferido.
Os resultados obtidos foram os seguintes: telejornal: $20$ pessoas, cinema: $16$ pessoas e programas desportivos: $24$ pessoas.
Qual é o número de elementos da população?
<br>
<br>
```

<u>Solução correta</u>:

```
A população é o número total de alunos, 3000.
<br>
<br>
```

<u>Soluções erradas</u>:

```
20
16
60
```

Exercício 1.2:

```
De entre os $3000$ alunos de uma escola selecionaram-se alguns para responder a um inquérito sobre o programa de televisão preferido.
Os resultados obtidos foram os seguintes: telejornal: $20$ pessoas, cinema: $16$ pessoas e programas desportivos: $24$ pessoas.
Qual é o número de elementos da amostra?
<br>
<br>
```

<u>Solução correta</u>:

```
A amostra é uma parte da população, 60.
<br>
```


Soluções erradas:

3000

20

16

Exercício 2.1:

Seja X a variável aleatória que representa o número de golos sofridos em cada jogo por um determinado guarda-redes.

Suponhamos que a distribuição de probabilidade de X é dada pela seguinte tabela:

<Table border=1>

tr>

<td align="center" width=100 bgcolor=White >x_i </td>

<td align="center" width=60 bgcolor=White>0</td>

<td align="center" width=60 bgcolor=White>1</td>

<td align="center" width=60 bgcolor=White>2</td>

<td align="center" width=60 bgcolor=White>3</td>

<td align="center" width=60 bgcolor=White>4</td>

</tr>

<tr>

<td align="center" width=100 bgcolor=White>p_i </td>

<td align="center" width=60 bgcolor=White>0,25</td>

<td align="center" width=60 bgcolor=White>0,40</td>

<td align="center" width=60 bgcolor=White>0,28</td>

<td align="center" width=60 bgcolor=White>0,05</td>

<td align="center" width=60 bgcolor=White>0,02</td>

</tr>

</table>

O valor médio da distribuição de probabilidades X é:

Solução correta: A expressão que nos permite chegar ao valor médio é a seguinte:

$$\mu=x_ {1}p_{1} +x_{2}p_{2}+x_{3}p_{3}+x_{4}p_{4}+x_{5}p_{5}$$

em que:

μ: valor médio

$x_i $: valores da variável estatística que constam na tabela do enunciado

$p_i $: valores da probabilidade que constam na tabela do enunciado

Substituindo na expressão do valor médio os valores de $x_i $ e $p_i $ obtemos:

$$\mu=0\times0,25+ 1\times0,40+2\times0,28+3\times0,05+4\times0,02\Leftrightarrow$$

$$\Leftrightarrow\mu=0+0,40+0,56+0,15+0,08=1,19\Leftrightarrow$$

$$\Leftrightarrow\mu=1,19$$

Soluções erradas:

1,44
0,238
3,26

Exercício 2.2:

```
Seja X a variável aleatória que representa o número de golos sofridos em
cada jogo por um determinado guarda-redes.
Suponhamos que a distribuição de probabilidade de X é dada pela seguinte
tabela:
<br>
<br>
<Table border=1>
<tr>
       <td  align="center"  width=100  bgcolor=White><b>x<sub>i</sub>  </b>
</td>
      <td align="center" width=60 bgcolor=White><b>0 </b></td>
      <td align="center" width=60 bgcolor=White><b>1</b></td>
      <td align="center" width=60 bgcolor=White><b>2</b></td>
      <td align="center" width=60 bgcolor=White><b>3</b></td>
      <td align="center" width=60 bgcolor=White><b>4</b></td>
</tr>
<tr>
      <td align="center" width=100 bgcolor=White><b>p<sub>i</sub></b></td>
      <td align="center" width=60 bgcolor=White>0,25</td>
      <td align="center" width=60 bgcolor=White>0,40</td>
      <td align="center" width=60 bgcolor=White>0,28</td>
      <td align="center" width=60 bgcolor=White>0,05</td>
      <td align="center" width=60 bgcolor=White>0,02</td>
</tr>
</table>
<br>
<br>

Considere que $\mu=1,19$
<br>
<br>
Admitindo que a época tem $50% jogos, em quantos desses jogos espera que
o guarda-redes sofra um número de golos pertencentes ao intervalo     $]
\mu-\sigma, \mu+\sigma [$:
<br>
<br>
```

Solução correta:

A expressão que nos permite chegar ao valor do desvio padrão é a seguinte:

$$\sigma=\sqrt{(x_1-\mu)^2p_1+(x_2-\mu)^2p_2+(x_3-\mu)^2p_3+(x_4-\mu)^2p_4+(x_5-\mu)^2p_5}$$.

μ: valor médio.

$x_i $: valores da variável estatística que constam na tabela do enunciado.

$p_i $: valores da probabilidade que constam na tabela do enunciado.

Substituindo na expressão do valor médio os valores de $x_i $ e ,$p_i $ obtém-se:

$$\sigma=\sqrt{(0-1,19)^2\times0,25+(1-1,19)^2\times0,40+(2-1,19)^2\times0,28+(3-1,19)^2\times0,05+(4-1,19)^2\times0,02} \simeq {0,9348}$$.

Consideremos agora o intervalo $] \mu-\sigma, \mu+\sigma [$, substituindo o valor de μ por $1,19$ e o de σ por $0,9348$ obtém-se:

$$]1,19-0,9348;1,19+0,9348[=]0,2552;2,1248[$$.

Então, agora vamos ver os valores de X da tabela do enunciado que pertencem ao intervalo que acabamos de determinar.
Os valores de X são $x_1 $ e $x_2 $.

Consultando a tabela, vemos que a probabilidade para $x_1 $ e $x_2 $ é $p_1 $ e $p_2 $ que são respetivamente $0,40$ e $0,28$, ou seja,

$$P(0,2552<{X}<2,1248) = p_1+p_2= 0,40+0,28=0,68$$.

Admitindo que a época tem 50 jogos, vamos ver quantos desses correspondem à probabilidade encontrada:

$$0,68\times50 = 34$$.

Soluções erradas:

47
46
20

Exercício 3.1.:

O clube de sondagens da Escola Secundária da Fonte Velha realizou um estudo sobre o grau de satisfação dos alunos da Escola relativamente às condições da biblioteca escolar.
Selecionaram, ao acaso, 120 alunos. Destes, 72 mostraram-se muito satisfeitos com as condições da biblioteca.
A escola tem um total de 500 alunos.

Qual é a população em estudo?

Solução correta:

A população é o número total de alunos, 500.

Soluções erradas:

72

120

192

Exercício 3.2.:

O clube de sondagens da Escola Secundária da Fonte Velha realizou um estudo sobre o grau de satisfação dos alunos da Escola relativamente às condições da biblioteca escolar.
Selecionaram, ao acaso, 120 alunos. Destes, 72 mostraram-se muito satisfeitos com as condições da biblioteca.
A escola tem um total de 500 alunos.

Qual o valor da estatística que utilizaram para tirar conclusões acerca da população?

Solução correta:

Se o pretendido é o grau de satisfação dos alunos da Escola relativamente às condições da biblioteca escolar e se foi escolhida uma amostra de 120 alunos e desses 72 manifestaram-se satisfeitos com as condições da biblioteca, então o valor da estatística para tirar conclusões sobre a população relativamente a esta questão é:

$\frac{72}{{120}}= 0,6 = 60\%$

Soluções erradas:

14,4%

24%

38,4%

Exercício 3.3.:

O clube de sondagens da Escola Secundária da Fonte Velha realizou um estudo sobre o grau de satisfação dos alunos da Escola relativamente às condições da biblioteca escolar.
Selecionaram, ao acaso, 120 alunos. Destes, 72 mostraram-se muito satisfeitos com as condições da biblioteca.
A escola tem um total de 500 alunos.

O valor da estatística que permite tirar conclusões sobre o grau de satisfação, por parte dos alunos, das condições da biblioteca são 60%.

Qual é previsão do número de alunos da escola que estão muito satisfeitos com as condições da biblioteca?

Solução correta:

Se a estatística é 60% e $60\% =0,60$ e o número total de alunos é 500, então:

$0,60\times500=300$

Soluções erradas:

72
120
192

Exercício 4.1.:

O clube de sondagens da Escola Secundária da Fonte Velha realizou um estudo sobre o grau de satisfação dos alunos da Escola relativamente às condições da biblioteca escolar.
Selecionaram, ao acaso, 120 alunos. Destes, 72 mostraram-se muito satisfeitos com as condições da biblioteca.
A escola tem um total de 500 alunos.

Qual é previsão do número de alunos da escola que estão muito satisfeitos com as condições da biblioteca?

Solução correta:

Segundo o teorema do Limite Central, sabemos que numa amostra de dimensão n, $n\geq30$, de uma população de valor médio $\mu=1200$ e desvio padrão $\sigma=a$,
pode ser aproximada por uma distribuição normal com o valor médio igual ao da população, 1200, e desvio padrão igual a:
$$\frac{a}{\sqrt{100}}$$.

Sabe-se que:
\[\frac{a}{\sqrt{100}}=8\Leftrightarrow\]
\[\Leftrightarrow a=8\times\sqrt{100}\]
\[\Leftrightarrow a=8\times10=80\].

O valor de a é 80.

Soluções erradas:

8
0,8
8

Exercício 5.1:

O diretor da companhia de um teatro solicitou ao seu contabilista um estudo sobre o valor médio das receitas de bilheteira por cada sessão realizada. O contabilista recolheu os dados das receitas de bilheteira de uma amostra de 50 sessões.
Com base nessa amostra, obteve, para o valor médio das receitas de bilheteira por sessão, o intervalo de confiança a 95$\%$ seguinte: $]4449,691;5214,309[$

A média e o desvio padrão das receitas de bilheteira obtidos pelo contabilista nessas 50 sessões é:

Solução correta:

br>
Com um intervalo de confiança de 95% , o valor de $z=1,960$ .Utilizando o intervalo de confiança para o valor médio, um vez que é este que é referido no enunciado:

$$\left]\overline{X}-z\times\frac{\sigma}{\sqrt{n}}, \overline{X}+z\times\frac{\sigma}{\sqrt{n}}\right[$$.

Soma-se quer os extremos do intervalo dado, quer os do intervalo no caso geral, pois permite que fiquemos só com a média.

$$\overline{X}-z\times\frac{\sigma}{\sqrt{n}}+\overline{X}+z\times\frac{\sigma}{\sqrt{n}}=2\times\overline{X}$$.

Seguidamente igualamos as somas obtidas e assim é possível determiner a média:

```
\[2\times\overline{X}=4449,691+5214,309\]
\[\Leftrightarrow 2\times\overline{X}=9664\]
\[\Leftrightarrow\overline{X}=4832\].

A partir do momento que temos o valor da média, recorremos a um dos extremos
no intervalo no caso geral para determinarmos o valor do desvio padrão.
<br>
$95\% \longrightarrow z=1,960$.
\[\overline{X}+z\times\frac{\sigma}{\sqrt{n}}=4832+1,960\times\frac{\sigm
a}{\sqrt{50}}\].
<br>
<br>
\[4832+1,960\times\frac{\sigma}{\sqrt{50}}=5214,309\]
\[\Leftrightarrow1,960\times\frac{\sigma}{\sqrt{50}}=5214,309-4832\]
\[\Leftrightarrow \frac{\sigma}{\sqrt{50}}=\frac{382,309}{\sqrt{1,960}}\]
\[\Leftrightarrow1,960\times\sigma=382,309\times\sqrt{50}\]
\[\Leftrightarrow\sigma=\frac{382,309\times\sqrt{50}}{1,960}\]
\[\Leftrightarrow\sigma\simeq1379\].
```

<u>Soluções erradas</u>:

$\bar{X} = 382, \sigma = 17433$

```
$\overline{X}=382,\space\sigma=17433$
```

$\bar{X} = 4832, \sigma = 18812$

```
$\overline{X}=4832,\space\sigma=18812$
```

$\bar{X} = 4832, \sigma = 8$

```
$\overline{X}=4832,\space\sigma=8$
```

Exercício 6.1:

```
Considere a população constituída pelos trabalhadores de uma empresa.
Pretende-se estudar o parâmetro vencimento médio, em euros.
Sabe-se que <b>$\sigma=250$ euros</b>.
Para um grau de <b>confiança de 99$\%$</b>, qual deverá ser a dimensão da
amostra para que o erro que se comete <b>não seja superior a 50€</b>?
<br>
<br>
```

Solução correta:

Com um intervalo de confiança de 99% , o valor de $z=2,576$.

$\sigma=250$

Utilizando o intervalo de confiança para o valor médio, conforme referido no enunciado:
$$\left]\overline{X}-z\times\frac{\sigma}{\sqrt{n}}, \overline{X}+z\times\frac{\sigma}{\sqrt{n}}\right[$$.

Subtrai - se os extremos do intervalo no caso geral, pois permite que fiquemos só com o número de elementos da amostra, como incógnita.
Seguidamente divide-se o resultado por 2.
Assim o resultado obtido será a amplitude do intervalo, ou seja, o erro cometido.

$$\left(\overline{X}+z\times\frac{\sigma}{\sqrt{n}}- \overline{X}+z\times\frac{\sigma}{\sqrt{n}}\right):2$$.
$$=\left(2z\frac{\sigma}{\sqrt{n}}\right):2=z\times\frac{\sigma}{\sqrt{n} }=z\frac{\sigma}{\sqrt{n}}$$.

Como se pretende a dimensão da amostra para que o erro que se comete não seja superior a 50 então:

$$z\frac{\sigma}{\sqrt{n}}\le50$$.

Agora substitui-se o valor de z e o do desvio padrão e calcula-se o valor de n:

$99\% \longrightarrow z=2,576$

\[z\frac{\sigma}{\sqrt{n}}\leq50\]
\[2,576\times\frac{250}{\sqrt{n}}\le50\]
\[\Leftrightarrow 2,576\times250\leq50\times\sqrt{n}\]
\[\Leftrightarrow\sqrt{n}\geq\frac{2,576\times250}{50}\]
\[\Leftrightarrow\sqrt{n}\geq12,88\]
\[\Leftrightarrow n\geq166\].

Assim, o valor de n será o menor valor que satisfaz a condição: $n\geq166$, ou seja, $n=166$.

Soluções erradas:

664
68
97

Exercício 7.1:

A Juliana trabalhou numa agência de viagens, na qual analisava os gastos mensais dos clientes. Num estudo realizado em abril de 2018, Juliana constituiu uma amostra de 100 clientes,de modo a estudar quantos tinham gasto mais de 1000 euros. Verificou que isso sucedeu com 58 clientes.
Com base na informação recolhida, o intervalo de confiança a 95$\%$, para a proporção dos clientes da agência que gastaram, nesse mês, mais de 1000 euros é:

Solução correta:

Com um intervalo de confiança de 95% , o valor de $z=1,960$.

$n=100$

Como se pretende um intervalo de confiança para a proporção, então será necessário começar por calcular a proporção dos clientes da agência que gastaram,
nesse mês, mais de 1000 euros.

$\hat{p}=\frac{58}{100}=0,58$.

Utilizando o intervalo de confiança para a proporção, conforme referido no enunciado:

$$\left]\hat{p}-z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}; \hat{p}+z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}\right[$$.

Agora substitui-se os valores da proporção $\hat{p}$ por $0,58$ e o valor de z por $1,960$.

$$\left]0,58-1,960\sqrt{\frac{0,58\times(1-0,58)}{100}}; 0,58+1,960\sqrt{ \frac{0,58\times(1-0,58)}{100}}\right[=]0,48;0,68[$$.

Soluções erradas:

$]0,45;0,70[$
$]0,48;0,67[$
$]0,50;0,66[$

Exercício 8.1:

Recolheu-se uma amostra de 36 smartphones da produção de uma determinada fábrica. Dessas 36 unidades, verificou-se que 8 tinham uma avaria.
Com um nível de confiança de 95$\%$, a proporção populacional de smartphones sem avarias está, aproximadamente, entre:

```
<br>
<br>
```

Solução correta:

```
<br>
<br>
Com um intervalo de confiança de $95\%$ , o valor de $z=1,960$ .
<br>
<br>
$n=36$.
<br>
<br>
Como se pretende um intervalo de confiança para a proporção, então será
necessário começar por calcular a proporção de smartphone sem avarias:
<br>
<br>
$$\hat{p}=\frac{28}{36}= \frac{7}{9}$$.
<br>
<br>
Utilizando o intervalo de confiança para a proporção, conforme referido no
enunciado:
<br>
<br>
$$\left]\hat{p}-z\sqrt{   \frac{\hat{p}\times\left(1-\hat{p}\right)}{n}};
\hat{p}+z\sqrt{ \frac{\hat{p}\times\left(1-\hat{p}\right)}{n}}\right[$$.
<br>
<br>
Agora substitui-se os valores da proporção: $\hat{p}$ por $\frac{7}{9}$ e
o valor de $z por 1,960$.
<br>
<br>
$$\left]\frac{7}{9}-1,960\sqrt{          \frac{\frac{7}{9}\times\left(1-
\frac{7}{9}\right)}{36}};\frac{7}{9}+1,960\sqrt{
\frac{\frac{7}{9}\times\left(1-
\frac{7}{9}\right)}{36}}\right[=]0,642;0,914[$$.
<br>
<br>
Para encontrarmos a solução é necessário transformar os valores
encontrados, em percentagem arredondados às unidades:
```


$0,642\times100=64,2\%\simeq64\%$.

$0,914\times100=91,4\%\simeq91\%$.

64% a 91%.

Soluções erradas:

9% e 36%
66% e 89%
76% e 80%

Exercício 9.1:

O número de utilizadores de uma das diversões de um parque, nas duas primeiras semanas do mês de agosto de 2017, são: domingo – 1ª semana:312; domingo – 2ª semana:288; sábado – 1ª semana:328; sábado – 2ª semana:344, num universo de 3640 utilizadores.
A amplitude de um intervalo de confiança para a proporção do número de utilizadores dessa diversão nos sábados e nos domingos,
face ao total do número de utilizadores no período de tempo registado é 0,0407301.
O nível de confiança desse intervalo é:

Solução Correta:

O número total de utilizadores dos sábados e domingos é :
$328+312+344+288=1272$.

Como se pretende um intervalo de confiança para a proporção, então será necessário começar por calcular a proporção do número de utilizadores dessa diversão nos sábados e nos domingos:

$$\hat{p}=\frac{1272}{3640}\simeq0,35$$.

Utilizando o intervalo de confiança para a proporção, conforme referido no enunciado:

$$\left]\hat{p}-z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}; \hat{p}+z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}\right[$$.

Subtrai-se os extremos do intervalo no caso geral e o resultado obtido será a amplitude do intervalo:

$$\hat{p}+z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}-\hat{p}+z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}=2z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}$$.

Se $0,0407301$ é a amplitude do intervalo de confiança, então igualamos a expressão encontada à amplitude dada.

Seguidamente substitui-se o valor da proporção $\hat{p}$ por $0,35$ e o valor n por 3640 (número de utilizadores) .
Resolvendo a equação formada, determina -se o valor de z.
\[2z\sqrt{ \frac{\hat{p}\times(1-\hat{p})}{n}}=0,0407301\]
\[\Leftrightarrow z\sqrt{ \frac{0,35\times(1-0,35)}{3640}}=\frac{0,0407301}{2}\]
\[\Leftrightarrow z\sqrt{ \frac{0,35\times0,65}{3640}}=0,02036505\]
\[\Leftrightarrow z=\frac{0,02036505}{\sqrt{\frac{0,35\times0,65}{3640}}}\]
\[\Leftrightarrow z\simeq2,576\]

```
<br>
Atendendo à tabela seguinte:
<br>
<Table border=1>
<tr>
       <td align="center" width=100 bgcolor=White><b>Nível de
confiança</b></td>
       <td align="center" width=60 bgcolor=White><b>90&percnt; </b></td>
            <td align="center" width=60 bgcolor=White><b>95&percnt;
</b></td>
        <td align="center" width=60 bgcolor=White><b>99&percnt; </b></td>
</tr>
<tr>
             <td align="center" width=100 bgcolor=White><b>z</b></td>
       <td align="center" width=60 bgcolor=White>1,645</td>
            <td align="center" width=60 bgcolor=White>1,960</td>
            <td align="center" width=60 bgcolor=White>2,576</td>
      </tr>
</table>
<br>
O valor de $z$  encontrado corresponde a um nível de confiança de $99\%$.
<br>
```