UNIVERSIDADE DE ÉVORA

Departamento de Matemática

MESTRADO EM MODELAÇÃO ESTATÍSTICA E ANÁLISE DE DADOS

Análise Estatística da Aptidão Física em Ambiente Escolar

Dissertação de Mestrado sob a orientação do Professor Doutor Paulo Infante

e co-orientação da Professora Doutora Dulce Pereira

Departamento de Matemática – Universidade de Évora

Cristina Alexandra Lopes Pereira

Évora, 2010

À Teresinha e à Joana pela ajuda na compilação dos dados,

à Patrícia pela companhia na caminhada da modelação estatística,

ao Paulo pela constante motivação e paciência…

Índice

Análise Estatística da Aptidão Física em Ambiente Escolar

Resumo

Mens sana in corpore san é uma conhecida expressão do latim que aponta para a importância de uma boa preparação física de modo a que a mente funcione bem. Numa escola básica e secundária encontram-se alunos com diferentes aptidões físicas, diferentes índices de massa corporal e diferentes hábitos alimentares.

Após uma avaliação da influência da idade, sexo e escola no índice de massa corporal, procuramos averiguar neste trabalho se existe alguma associação estatística entre a capacidade física dos alunos, medida pela bateria de testes *Fitnessgram* (bateria de testes que avalia o nível de aptidão física das crianças/adolescentes) e pelo seu desempenho a educação física, e o seu rendimento escolar, medido pelas notas finais nas disciplinas de Matemática e Língua Portuguesa.

Foram seleccionadas três escolas por facilidade na obtenção dos dados e de modo a tentarmos traduzir realidades diferentes, como sejam a cidade e a vila com uma grande envolvente rural, o norte e o sul. Recorreu-se ao ajustamento de dois modelos de regressão logística, um para Matemática e outro para Língua Portuguesa, com variável resposta binária correspondendo a se o aluno teve ou não aprovação na disciplina respectiva e ao ajustamento de dois modelos de regressão ordinal considerando como variável resposta as notas dos alunos.

Tratando-se de certo modo de algo que pode ser identificado como um caso de estudo apenas se pretende descrever e analisar algumas hipóteses colocadas *a priori* e propor novas conjecturas que resultam deste estudo e que poderão ser validadas ou declinadas em estudos futuros. Além de um desempenho excelente em Educação Física contribuir para um melhor desempenho em Matemática e em Língua Portuguesa, refira-se que em relação às variáveis da aptidão física, a flexibilidade esquerda revelou-se factor potenciador de uma melhor nota a Matemática e também os abdominais na Matemática e o Vai-vém na Língua Portuguesa se mostraram factores potenciadores de uma melhor nota.

Finalmente foi feita uma análise de "*missing values*", e comparou-se os resultados obtidos com e sem estimação dos valores em falta na base de dados. Observou-se uma boa concordância entre as duas abordagens.

Statistical Analysis of Physical Fitness in the School Environment

Abstract

Mens sana in corpore san is a famous latin expression that point out the importance of a good physical preparation so that the mind works well.

In primary and secondary schools there are students with different physical abilities, different body mass indexes and different dietary habits.

After an evaluation of the influence of age, sex and school in body mass index, we investigate if there is any association between the physical capability of students, measured by *Fitnessgram* test battery (battery of tests that evaluates the level of physical fitness in children / adolescents) and by their performance in physical education, and academic performance, measured by final grades in the courses of Mathematics and Portuguese Language.

Three schools were selected to collect the data and so we try to translate different realities, such as town and village with a large surrounding rural, north and south.

We adjust two logistic regression models, one for Mathematics and other for the Portuguese Language, with binary response variable corresponding to whether the student had or failure in the respective course and we also adjust two ordinal regression models considering as dependent variable students' grades.

This study can be identified as a case study, we only intended to describe and analyze a few *priori* assumptions and propose new conjectures that result from this study and that could be validated or declined in further studies. Excellent performances in Physical Education contribute to a better performance in Mathematics and Portuguese language. We point out that the variables of physical *fitness*, left flexibility are determinant of a better performance in Mathematics. The abdominal is a potentiating factor to have a better performance in Mathematics and the comings and goings in Portuguese Language.

We conclude that gender and education are significant factors in explaining the notes; girls tend to have higher marks than boys in Portuguese Language and Mathematics.

Finally, an analysis of "missing values" was made and we compared the results with and without estimation of missing values in the database. We concluded that there was a good agreement between the two approaches.

Capítulo I - Introdução

O presente estudo pretende ser mais uma fonte de partilha e de promoção da reflexão sobre temáticas actuais que preocupam os intervenientes das áreas da Educação Física, do Desporto e da Saúde. A perspectiva da Educação Física nas vertentes Saúde e Lazer, entre outras, representa um novo desafio.

Os estilos de vida saudáveis são promovidos por práticas activas que diminuam e combatam o sedentarismo que é responsável, entre outros, pelas designadas doenças da cidade, como a obesidade. Esta, afecta hoje não só os adultos mas também as crianças nas nossas escolas.

Vários investigadores de diferentes domínios das Ciências do Desporto têm tentado entender o alcance e a complexidade da noção de aptidão física. A saúde pública é um motivo crescente de preocupação ao nível da qualidade de vida das populações. Durante o último século as razões desta preocupação alteraram-se profundamente. Se há cinquenta anos atrás as principais causas de mortalidade eram provocadas por doenças infecto-contagiosas, à medida que a ciência e a tecnologia avançaram, estas causas, pelo menos nos países industrialmente desenvolvidos, passaram a dar lugar aos processos crónico-degenerativos, como doenças do coração, diabetes, cancro, entre outros (Bergmann *et al*. 2005). Neste sentido, a escola deve constituir-se como um contexto privilegiado de intervenção na Saúde Pública, de forma a prevenir a cada vez maior taxa de sedentarismo dos jovens portugueses. Duas das consequências da diminuição da actividade física são o aumento do tecido gorduroso e o aumento no risco de desenvolvimento de problemas cardiovasculares (Bouchard e Deprés, 1995). Como estrutura multidimensional físico-motora de cada um, a aptidão física (AptF), poderá ser considerada como um indicador do estado de saúde. Sobre o tema, Blair *et al.* (1989) verificaram um menor grau de mortalidade nos indivíduos com níveis elevados de AptF, quando comparados com os de baixo nível. É evidente cientificamente que a actividade física e o exercício têm efeitos saudáveis, pelo que aumentar a qualidade de vida, com um estilo de vida activo para prevenir doenças, pode ser uma das melhores alternativas da humanidade. Com a adopção de um estilo de vida mais activo, podem-se obter bons níveis de aptidão física, que por sua vez, podem proporcionar a sensação de bem-estar físico, mental e social.

Nos adultos evidencia-se que o nível de actividade física habitual influencia o estado de saúde físico-mental, gerando a necessidade de um estilo de vida fisicamente activo para ser uma pessoa mais saudável (FITNESSGRAM, 1987; Aahperd, 1988; Acsm, 2005; Who, 2007).

As crianças que não obtêm hábitos saudáveis como a prática de exercício físico, de forma sistematizada e orientada, têm uma maior probabilidade de ter uma vida adulta com os mesmos hábitos sedentários. Já o indivíduo que possui e/ou possuiu uma prática saudável ao longo da juventude tem maiores condições de perdurar tais costumes pela vida adulta e terceira idade (Nahas e Corbin, 1992; Marques e Gaya, 1999; Gaya, Guedes, Torres, *et a*l., 2002; Bergmann *et al*., 2005).

A Educação Física desempenha a importante e estratégica função de promover, de uma forma conhecedora e empenhada, a prática do exercício físico regular junto dos alunos.

Para este efeito é necessário que a Educação Física veicule os meios e os métodos de intervenção necessários. Assim, a escola porque é acessível às crianças de hoje em dia, deve ser também um espaço fundamental para a promoção da saúde e educação para a saúde. É também por excelência o local para promover hábitos de alimentação saudáveis e comportamentos de actividade física.

A avaliação da aptidão física é um elemento essencial a qualquer programa de actividade física que tenha como objectivo a educação para a Saúde.

O *Fitnessgram* é um método eficaz de avaliação da aptidão física por diversas razões:

a) permite que os alunos tenham um maior contacto com as várias áreas da aptidão física, ficando assim a saber identificar cada componente e a conhecer melhor a sua importância;
b) num tempo relativamente curto, fornece informações a partir das quais se podem avaliar as atitudes e políticas respeitantes à condição física dos alunos e, se for caso disso, proceder à sua alteração;
c) ajuda o aluno a tomar consciência da sua condição física, definir metas e assim motivar-se para melhorar a sua forma; por outro lado pode pôr em evidência problemas de saúde individuais;
d) compara os resultados dos testes com valores de referência associados a importantes indicadores da saúde (derivados de estudos cuidadosamente executados), utilizando para o efeito medidas internacionais já validadas por

milhões de alunos (a Educação Física é uma das raras disciplinas escolares praticada por todos os alunos em todo o espaço europeu).

A Aptidão Física é uma das três grandes áreas específicas de intervenção dos novos programas da disciplina de Educação Física.

A avaliação desta área é efectuada tendo como referência a Zona Saudável de Aptidão Física, dos testes do *Fitnessgram* que por sua vez são realizados nas aulas de Educação Física, sob orientação dos respectivos professores, seguindo os protocolos pré-estabelecidos.

Na análise estatística realizada neste trabalho pretende-se estudar factores que podem influenciar os diferentes valores obtidos para a aptidão física em ambiente escolar, comparando algumas variáveis que se pensa serem relevantes nos níveis de aptidão física, tais como a "idade", "sexo", "altura", "peso" e "IMC". Para além destes factores existem outros que vão entrar no estudo como possíveis variáveis explicativas, tais como os índices obtidos pelos alunos nos testes de *Fitnessgram* (vai-vém, abdominais e flexibilidade à esquerda e à direita). Foram, ainda, recolhidas as notas obtidas nas disciplinas de Matemática, Língua Portuguesa e Educação Física nas três escolas. Assim, pretende-se averiguar a influência/comportamento de todos estes factores nas notas obtidas nas disciplinas de Matemática e Língua Portuguesa, e averiguar diferenças entre duas Escolas do Alentejo e um agrupamento de Escolas do Norte.

No capítulo seguinte descrevem-se aspectos teóricos da metodologia estatística aplicada neste estudo, a análise univariada, o modelo de regressão logística, o modelo de regressão multinomial e o modelo de regressão ordinal. Expõem-se as técnicas de análise de variância multivariada. E, por último, a análise de *Missings values*.

No capítulo três será feita uma análise exploratória dos dados para cada variável utilizada. Posteriormente, averigua-se a existência de associações entre as diferentes variáveis. Para modelar o IMC (índice de massa corporal) será ajustado um modelo de Regressão Logística Binomial com o intuito de averiguar a relação entre o IMC e variáveis como o sexo, idade e escola a que pertencem os alunos. Para que este estudo fizesse sentido houve a necessidade de criar uma outra variável categórica que permite avaliar se os alunos estão dentro ou fora da zona aceitável para os valores de IMC. Também se criou uma outra variável categórica de forma a dar-nos a proporção de alunos que se encontram acima, dentro ou abaixo da Zona Saudável para o IMC, que foi

a variável resposta de um modelo de Regressão Multinomial, procurando analisar como esta é influenciada pelas covariáveis sexo, idade e Escola.

No capítulo quatro será feita uma análise ao desempenho escolar dos alunos. Para analisarmos diferenças significativas no desempenho escolar nas três disciplinas, aplicámos a MANOVA relativamente às Escolas e ao sexo.

Posteriormente foi ajustado um Modelo de Regressão Logística Binomial para as disciplinas de Matemática e Língua Portuguesa. No caso da disciplina de Educação Física, o número de alunos com nota negativa era pouco significativo tendo-se optado por não continuar o estudo para essa disciplina. Uma vez que as notas estão ordenadas por uma escala de 1 a 5 foi também ajustado um Modelo de Regressão Logística Ordinal, pois os modelos de regressão logística ordinal são os indicados para esta escala.

Nos dados fornecidos pelas escolas nas variáveis referentes aos testes de *Fitnessgram*, faltava muita informação, pois houve turmas inteiras para as quais não houve tempo suficiente nas aulas para a realização de todos os testes. Sendo que a ocorrência de um tão elevado número de valores omissos também se deve a várias situações, ou porque os alunos faltaram à aula em que foram feitos alguns dos testes, ou porque entretanto anularam a matrícula. Neste caso foi feita uma abordagem diferente, a análise de *Missings Values*. Para "completar" os dados em falta e possibilitar a análise com todos os indivíduos do estudo aplicou-se o método de Imputação Múltipla e ajustou-se um modelo para o desempenho escolar em Matemática, sendo comparados os resultados deste modelo com os obtidos pelo modelo sem a estimação dos *missings*.
No quarto capítulo apresentam-se as conclusões mais importantes das aplicações desenvolvidas no terceiro capítulo e terminamos com algumas referências.

Capítulo II - Metodologia Estatística (Alguns Aspectos Teóricos)

Neste capítulo descrevem-se os aspectos teóricos principais da metodologia estatística aplicada neste trabalho: o modelo de regressão logística, o modelo de regressão multinomial e o modelo de regressão ordinal. Faz-se também uma síntese de algumas técnicas de análise de variância multivariada e uma descrição do método de imputação múltipla usado na parte final aquando da análise dos valores em falta na base de dados.

2.1. O Modelo de Regressão Logística

O modelo de regressão linear assume que a variável dependente Y é contínua, sendo os resíduos aleatórios (e_i com $i=1,...,n$) independentes e normalmente distribuídos com a mesma variância (σ^2). O modelo de regressão linear simples (com uma só variável independente X) é dado por:

$$y_i = \beta_0 + \beta_1 x_i + e_i \text{ com } i=1,...,n$$

Os coeficientes do modelo β_0 e β_1 são estimados pelo Método dos Mínimos Quadrados que consiste em encontrar a ordenada na origem e o declive da recta de modo a minimizar a soma dos quadrados dos resíduos aleatórios.

Quando a variável dependente é dicotómica (assumindo os valores 0 e 1 como realização ou não de um evento em análise de acordo com um esquema de Bernoulli) e se pretende explicar a partir de um conjunto de variáveis independentes, o modelo de regressão linear não é adequado. Repare-se que, atendendo a que

$$\begin{cases} Y_i = 1 \rightarrow P(Y_i = 1) = \pi_i \\ Y_i = 0 \rightarrow P(Y_i = 0) = 1 - \pi_i \end{cases}$$

$$E(Y_i) = \beta_0 + \beta_1 X_i = \pi_i$$

e

$$\sigma^2(Y_i) = E\left[\left(Y_i - E(Y_i)\right)^2\right] = (1-\pi_i)^2 \pi_i + (0-\pi_i)^2 (1-\pi_i) = \pi_i (1-\pi_i) = E(Y_i)\left(1 - E(Y_i)\right)$$

Como

$\varepsilon_i = Y_i - \pi_i$ (com Y_i constante),

as variâncias dos erros são dadas por

$$\sigma^2(\varepsilon_i) = \pi_i(1-\pi_i) = (\beta_0 + \beta_1 X_i)(1-\beta_0 - \beta_1 X_i).$$

Portanto, a variância dos erros depende dos valores da variável explicativa, pelo que é heterogénea. Deste modo verificamos, o que era óbvio, que não é válido nenhum dos pressupostos inerentes ao modelo de regressão linear.
Por outro lado, observe-se que o valor médio da variável resposta está limitado entre 0 e 1 e, consequentemente, esta restrição é inapropriada para uma função de resposta linear.

Uma vez que a variável resposta é dicotómica, a solução passa por modelar usando a função logística, pois tem uma curva da resposta em forma de S quando a ocorrência de eventos aumenta com o valor da variável explicativa ou em forma de S invertido quando a ocorrência de eventos diminui com o valor da variável explicativa.

A resposta pode ser linearizada através da função logit (função probabilidade), pois

$$\text{logit}\,\pi(x) = \log\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x$$

e assim

$$E(Y) = \left[1 + \exp(-\beta_0 - \beta_1 X)\right]^{-1}$$

Esta análise é utilizada quando todas as variáveis explicativas são categóricas e a regressão logística é, muitas vezes, escolhida se as variáveis explicativas são uma mistura de variáveis contínuas e variáveis categóricas e / ou se as variáveis não estão bem distribuídas, pois a regressão logística não faz suposições sobre as distribuições das variáveis explicativas. Apenas exige que a variável contínua seja linear com o logit.

A grande popularidade da regressão logística assenta fundamentalmente na facilidade de interpretação dos parâmetros do modelo que a função logit permite, uma vez que o quociente entre a probabilidade de ocorrência do evento e a probabilidade de não ocorrência traduzem o odds (chance) de ocorrer o sucesso relativamente ao insucesso.

Atendendo a que

$$\text{logit}\,\pi\left(x_j\right)=\beta_0+\beta_j x$$

e

$$\text{logit}\,\pi\left(x_j+1\right)-\text{logit}\,\pi\left(x_j\right)=\beta_j$$

Tem-se que a exponencial do coeficiente da variável representa o *odds ratio* (OR) da ocorrência do evento relativamente à não ocorrência quando a variável independente x_j aumenta uma unidade (ou está num dada categoria, no caso de ser categórica)

$$OR=\frac{\pi}{1-\pi}$$

Ou seja, quando x_j varia uma unidade as chances de obter o sucesso aumentam (caso OR>1) ou diminuem (Caso OR<1) β_j unidades.

Para amostras grandes, os intervalos de confiança mais usuais para os pârametros de regressão são os de Wald, dados por (Agresti, 2007):

$\hat{\beta}_i \pm Z_{1-\alpha/2}\hat{\sigma}_{\hat{\beta}_i}$,

Em alternativa à linearização da função com a transformação *Logit*, surge outra função de ligação (muito usada nos modelos com variáveis dependentes qualitativas dicotómicas), a função de distribuição Normal. O modelo designa-se por modelo *Probit* e assume que a variável dependente dicotómica *Y* é operacionalizada por uma função latente contínua (η) tal que:

$$\begin{cases} Y=1, & \text{se}\,\eta+\varepsilon\geq\alpha \\ Y=0, & \text{se}\,\eta+\varepsilon<\alpha \end{cases},$$

com $\eta=\beta_0+\beta_1 X$ e α uma constante (ponte de corte). Neste caso o modelo ajustado é dado por

$\Phi^{-1}(\eta)=\beta_0+\beta_1 X+\varepsilon$, onde $\vec{\beta}=\beta_1,...,\beta_p$.

No caso de termos p variáveis independentes, o logit da regressão logística múltipla é dado pela expressão seguinte:

$\text{logit}\,\pi(x)=\beta_0+\beta_1 x_1+...+\beta_p x_p$ e assim $\text{E}(Y)=\left[1+\exp\left(-\beta_0-\beta_1 x_1-...-\beta_p x_p\right)\right]^{-1}$.

A expressão do modelo logístico múltiplo é dado por

$$E(Y)=\left(1+\exp\left(-\beta' X\right)\right)^{-1}$$

As variáveis independentes podem representar efeitos de interacção, ser quantitativas, ou ser qualitativas e representadas por variáveis indicadoras (caso em que as variáveis estão numa escala nominal tal como a etnia e o sexo).

A utilização de uma ou outra função de ligação e consequentemente a opção pelo modelo a utilizar depende da situação em causa. Contudo, a interpretação dos parâmetros do modelo probit é muito menos intuitiva e rica do que a interpretação dos parâmetros do modelo logit.

2.1.1. Ajustamento do modelo

Para avaliar a bondade do ajustamento e a qualidade do modelo utilizamos o teste da razão de verosimilhança entre o modelo nulo (só com a constante) e o modelo final completo (com todas as variáveis independentes significativas). As hipóteses a testar são:

$$\begin{cases} H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0, \left(\text{o modelo não é estatisticamente significativo}\right) \\ H_1 : \exists_i : \beta_i \neq 0 \left(i = 1,..., p\right), \left(\text{o modelo é estatisticamente significativo}\right) \end{cases}.$$

O cálculo da estatística de teste para avaliar a significância do modelo G^2 é a diferença entre o $-2LL$ dos dois modelos (diferença entre o valor da verosimilhança do modelo final e do modelo reduzido), e permite não rejeitar ou rejeitar a hipótese nula de que o modelo não é significativo. No entanto se concluirmos que o modelo é significativo, apenas podemos concluir que, pelo menos uma das variáveis independentes do modelo completo influencia a variável dependente do modelo. Ou seja, não quer dizer que o ajustamento realizado seja bom.

Para testar a significância do ajustamento do modelo completo, utiliza-se a estatística de teste do Qui-quadrado de Pearson (Hosmer & Lemeshow, 2000):

$$\chi_p^2 = \sum_{j=1}^{J} \frac{\left(O_j - E_j\right)^2}{E_j},$$

onde O_j e E_j são, respectivamente, o número de sucessos observados e de sucessos esperados para a célula *j*. Outros autores McCullag e Nelder (1989) utilizam a estatística *Deviance*:

$$D = -2Ln\left[\frac{L_c}{L_s}\right],$$

onde L_c é a verosimilhança do modelo ajustado e L_s é a verosimilhança do modelo com todas as variáveis independentes. A significância dos testes do Qui-quadrado e da *Deviance* permitem rejeitar ou não a hipótese de que "o modelo ajusta-se aos dados".

Depois de verificarmos se o modelo ajustado é significativo, importa identificar qual ou quais são as variáveis independentes que o influenciam significativamente. A forma mais usual de testar se uma dada variável independente é significativa para o modelo consiste em recorrer ao Teste de Wald, o qual está implementado em todos os softwares. A estatística de teste é dada pela expressão

$$T_{Wald_i} = \frac{\hat{\beta}_i}{\hat{\sigma}\left(\hat{\beta}_i\right)}$$

onde $\hat{\beta}_i$ é o estimador de β_i e $\hat{\sigma}\left(\hat{\beta}_i\right)$ é o estimador do desvio padrão de β_i. Esta estatística tem distribuição *t*-student, que se aproxima assimptoticamente da distribuição *normal padrão* para amostras grandes.

Tal como referem Hosmer e Lemeshow (2000), o teste de Wald é muito conservativo, sendo muitas vezes preferível usar o teste de razão de verosimilhanças para testar a significância de uma dada variável. Para tal, deve ser comparado o modelo com e sem a mesma, sendo o cálculo da estatística de teste para avaliar a significância do modelo a diferença entre a *deviance* dos dois modelos (diferença entre o valor da verosimilhança do modelo final e do modelo reduzido),

$$G=D\left(\text{modelo sem variavel}\right)\text{-}D(\text{modelo com variavel})=$$
$$=\text{-}2\ln\left(\frac{\text{verosomilhança sem variavel}}{\text{verosomilhança com variavel}}\right)$$

2.1.2. Selecção de variáveis

No ponto anterior foram focados procedimentos associados à inclusão de variáveis, significância e interpretação dos coeficientes dos modelos de regressão logística. Mas não deve ser indiferente o facto de que a inclusão de uma determinada variável pode depender do objectivo do estudo em causa, sem que o modelo final possa vir a ser o mais parcimonioso. Em estudos epidemiológicos é sugerida a inclusão de variáveis que sejam consideradas relevantes do ponto de vista clínico e/ou intuitivo, independentemente da sua significância estatística. Apesar dos *softwares* estatísticos terem implementados vários métodos automáticos de selecção de variáveis (backward, forward e stepwise, entre outros), estes devem servir como auxílio e não como a forma de definir o modelo final. Na parte de modelação deste trabalho formam seguidos os seguintes passos:

- Análise da influência individual de cada variável relativamente à variável resposta.
- Qualquer variável que, no ponto anterior, obtenha uma significância de 20% entra no modelo multivariado.
- Procede-se então à remoção das covariáveis não significativas, por ordem decrescente das suas significâncias e até que no modelo fiquem apenas variáveis significativas a um nível de 10%.
- Após a obtenção do modelo no qual é assumido que estão incluídas as variáveis essenciais, deve ser dado início a uma análise mais atenta às mesmas. Nesta fase está incluída a verificação de linearidade do logit com as variáveis contínuas, no caso de existirem. Usualmente faz-se uma representação gráfica dos pontos médios dos intervalos obtidos pela categorização da variável contínua usando os quartis como ponto de corte contra os coeficientes de cada categoria (a de referência naturalmente com o valor nulo) do modelo ajustado substituindo a variável contínua pela variável contínua categorizada. Naturalmente, espera-se que desta representação gráfica resulte aproximadamente uma recta.
- Depois de verificados os efeitos principais do modelo, devem ser testadas as interacções entre as variáveis do modelo, começando por verificar as interacções significativas uma a uma e posteriormente quais destas são significativas no modelo.

Antes que qualquer modelo seja utilizado para inferência dos resultados obtidos, devem ser verificadas a sua adequação e ajustamento. Uma avaliação completa do modelo ajustado envolve o cálculo das medidas da distância entre valores observados e estimados e uma análise gráfica de resíduos com o objectivo de procurar observações individuais influentes sobre a estimação dos parâmetros do modelo ou *outliers*.

2.1.3. Avaliação e diagnóstico do modelo

Uma avaliação completa do modelo ajustado envolve o cálculo das medidas sumárias da distância entre valores observados e estimados (bondade do ajustamento do modelo) e uma análise gráfica de resíduos com o objectivo de procurar observações individuais influentes sobre a estimação dos parâmetros do modelo ou *outliers*.

Embora as estimativas de máxima verosimilhança de parâmetros sejam as mesmas para qualquer forma de dados, as estatísticas Qui-quadrado de Pearson e a *Deviance* não o são. Estes testes de bondade de ajustamento só fazem sentido para dados agrupados. Quando calculadas para os modelos de regressão logística com variáveis explicativas contínuas estas estatísticas não têm distribuição aproximadamente Qui-quadrado (Agresti, 2007).

Esta questão é ultrapassada criando categorias para cada variável explicativa e aplicar essas estatísticas às contagens observadas e ajustadas em cada categoria. Lemeshow & Hosmer (1982) propõem obter os intervalos de categorização de acordo com as probabilidades estimadas, sendo formados g grupos (usualmente 10) de igual tamanho, em que o primeiro se refere ao grupo das observações com maiores probabilidades estimadas e assim sucessivamente. O teste de bondade de ajustamento de Hosmer-Lemeshow usa uma estatística de teste de Pearson para comparar as contagens observadas com as estimadas pelo modelo de regressão logística. A estatística de teste para testar o ajustamento do modelo aos dados é dada por

$$X^2_{HL} = \sum_{i=1}^{g} \frac{(O_i - E_i)^2}{E_i}$$

que segue, para amostras grandes uma distribuição qui-quadrado com g-2 graus de liberdade.

Os modelos logit não têm uma boa medida da qualidade do ajustamento intuitiva como o coeficiente de determinação para os modelos lineares. Várias medidas de pseudo-R^2 foram sugeridas na literatura entre as quais o R^2 de Nagelkerke e o R^2 de Cox-Snell. O R^2 de Nagelkerke é preferível, pois pode tomar o valor 1 enquanto o R^2 de Cox-Snell nunca atinge o valor 1 mesmo quando o ajustamento é perfeito. Contudo, os valores intermédios podem não ser interpretáveis como sendo a percentagem da variabilidade da variável dependente que é explicada pelo modelo (Mittlbock e Schemper, 1996). Hosmer e Lemeshow (2000) referem que os pseudo-R^2 não são medidas da variabilidade explicada pelo modelo, pois são baseados na comparação do modelo ajustado com o modelo nulo.

A análise de resíduos tem um papel determinante no diagnóstico da regressão logística, uma vez que permite identificar *outliers* e casos influentes na estimação do modelo. No entanto é necessário ter em conta que as variâncias das observações não são constantes como acontece na regressão linear. Hosmer e Lemeshow (2000) estimam os resíduos ajustados (resíduos standardizados ou de Pearson). Para além de permitir a identificação de *outliers*, os resíduos podem ser utilizados para avaliar a influência de uma observação no ajustamento do modelo:

$$R^{*P}_i = \frac{(y_i - \hat{\mu}_i) w_i}{\sqrt{\hat{\phi} V(\hat{\mu}_i)(1 - h_{ii})}}$$

A desvantagem deste resíduo (Pearson) incide no facto da sua distribuição ser, geralmente, bastante assimétrica para modelos não normais.
Assim, muitas vezes utiliza-se o resíduo da função *Desviance*, onde o desvio residual corresponde à i-ésima observação para a função desvio:

$$R^D_i = \frac{\delta_i \sqrt{d_i}}{\sqrt{\hat{\phi}(1 - h_{ii})}}, \ \delta_i = sinal(y_i - \hat{\mu}_i)$$

Para termos a certeza de que não existem desvios isolados do modelo, isto é, uma ou mais observações mal ajustadas pelo modelo que não seguem o padrão das restantes observações, teremos que analisar a influência e a consistência dessas observações discordantes.

Em relação à influência das observações, temos que uma observação é influente quando uma ligeira modificação ou exclusão do modelo produz alterações significativas nas estimativas dos parâmetros do modelo. No entanto, as observações influentes não têm necessariamente resíduos elevados. Um indicador da influência da i-ésima observação no vector de parâmetros estimados pode ser calculado pela diferença das estimativas de máxima verosimilhança do vector parâmetro β obtidas da amostra sem a observação e da amostra com todas as observações. No caso dos modelos lineares normais é sugerida a aplicação da distância de Cook, a qual também se costuma usar nestes modelos.

A influência de cada observação na estimação de cada um dos coeficientes de regressão pode ser estimada pelos resíduos *DfBetas*

$$DfBeta_{ij} = \hat{\beta}_i - \hat{\beta}_{i(-j)}.$$

onde $\hat{\beta}_i$ representa a estimativa do coeficiente de regressão ajustado para todas as observações e $\hat{\beta}_{i(-j)}$ a estimativa do coeficiente de regressão ajustado sem a observação j.

Quando estamos perante uma observação com resíduo elevado significa que, geralmente, essa observação é inconsistente. Assim, deve-se adaptar o modelo sem uma dada observação e calcular os resíduos da observação eliminada em relação ao correspondente valor predito. Aos resíduos obtidos damos o nome de resíduos de eliminação, sendo os resíduos de eliminação de Pearson dados por:

$$R^{*P}_{(i)} = \frac{\left(y_i - \hat{\mu}_i\right) w_i}{\sqrt{\hat{\phi} V\left(\hat{\mu}_i\right)\left(1 + h_{(ii)}\right)}}, \qquad h_{(ii)} = z_i^T \left(Z_{(i)}^T W_{(i)} Z_{(i)}\right) z_i$$

onde h_{ii} representam os valores da diagonal da matriz Hessiana.

A Sensibilidade e Especificidade do modelo permitem avaliar se o modelo é eficiente na classificação dos indivíduos. A sensibilidade é a percentagem de classificações correctas na classe de referência "1-Sucesso" da variável dependente, e a especificidade é a percentagem de classificações correctas na classe "0-Insucesso". Reportando à Biomedicina, a sensibilidade pode ser definida como a probabilidade condicionada de um diagnóstico positivo sabendo que o paciente tem a doença. Por sua vez, a especificidade pode definir-se como a probabilidade condicionada de um diagnóstico negativo sabendo que o paciente não tem a doença. Isto é,

$$\begin{cases} Sensibilidade = P\left[\hat{Y} = 1 \mid Y = 1\right] \\ Especificidade = P\left[\hat{Y} = 0 \mid Y = 0\right] \end{cases}.$$

Um modelo com boas capacidades preditivas apresenta valores para a sensibilidade e especificidade superiores a 80%. No capítulo IV é feito o diagnóstico do Modelo de Regressão logística Binomial para o Desempenho a Matemática e a Língua Portuguesa, do qual faz parte a interpretação da área da curva ROC.

A área sob a Curva ROC é outra medida que permite avaliar a capacidade do modelo para discriminar os sujeitos com a característica de interesse *vs* sujeitos sem a característica de interesse. A área da Curva ROC varia entre 0 e 1, quanto mais próxima de 1, maior é a capacidade do modelo para discriminar os indivíduos com e sem a característica de interesse. A Curva ROC é o gráfico da probabilidade de se detectar os verdadeiros positivos (sensibilidade) e os verdadeiros negativos (1-especificidade) para diferentes pontos de corte. Embora o melhor ponto de corte dependa de um contexto específico, é muitas vezes usual tomar como ponto de corte o valor que torna máxima a sensibilidade e a especificidade.

De acordo com Hosmer e Lemeshow (2000), podemos interpretar o poder discriminante do modelo de regressão pelos seguintes valores da área ROC:

Área ROC	Poder discriminante do modelo
A = 0,5	Sem poder discriminativo
$0,5 \leq A \leq 0,7$	Discriminação fraca
$0,7 \leq A \leq 0,8$	Discriminação aceitável
$0,8 \leq A \leq 0,9$	Discriminação boa
$A \geq 0,9$	Discriminação excepcional

2.2. O Modelo de Regressão Multinomial

No modelo de regressão multinomial a variável resposta assume mais do que duas categorias. Enquanto, no modelo de regressão logística é utilizada uma variável binária parametrizada como $Y = 0$ ou $Y = 1$, para o modelo multinomial consideramos uma variável dependente com 3 ou mais classes, escolhendo uma delas como referência. O

que se faz usualmente é escolher $Y=0$ como referência e comparar com as funções logísticas $Y=1$, $Y=2$, ….

No caso de termos três categorias de resposta, no desenvolvimento do modelo assumimos que existem *p* covariáveis e um factor fixo, vector *X*, e definimos as duas funções *logit* da seguinte forma:

$$g_1(X)=\ln\left[\frac{P(Y=1|X)}{P(Y=0|X)}\right]=\beta_{10}+\beta_{11}X_1+\beta_{12}X_2+...+\beta_{1p}X_p=X'\beta_1$$

$$g_2(X)=\ln\left[\frac{P(Y=2|X)}{P(Y=0|X)}\right]=\beta_{20}+\beta_{21}X_1+\beta_{22}X_2+...+\beta_{2p}X_p=X'\beta_2$$

Como existe mais do que uma combinação de coeficientes que conduzem às mesmas probabilidades, torna-se necessário normalizar o sistema relativamente a uma das categorias da variável dependente. As probabilidades são dadas por:

$$P(Y=0|X)=\frac{1}{1+e^{g_1(X)}+e^{g_2(X)}} \quad ; \quad P(Y=1|X)=\frac{e^{g_1(X)}}{1+e^{g_1(X)}+e^{g_2(X)}} \quad \text{e}$$

$$P(Y=2|X)=\frac{e^{g_2(X)}}{1+e^{g_1(X)}+e^{g_2(X)}}.$$

A expressão geral da probabilidade para o modelo com três categorias é dada por $P(Y=j|X)=\frac{e^{g_j(X)}}{\sum_{k=0}^{2}e^{g_j(X)}}$, com $\beta_0=0$ e $g_0(X)=0$.

Os parâmetros são estimados recorrendo à função de máxima verosimilhança (Hosmer e Lemeshow, 2000). As medidas da qualidade do ajustamento utilizadas são análogas às do modelo Binomial, sendo a análise de resíduos análogas às da regressão logística.

Os *Odds Ratio* são calculados para cada uma das classes relativamente à classe de referência $Y=0$. O *Odds Ratio* para a classe $Y=j$ versus $Y=0$, para as covariáveis $X=a$ versus $X=b$ é dado por

$$OR_j(a,b)=\frac{P(Y=j|X=a)/P(Y=0|X=a)}{P(Y=j|X=b)/P(Y=0|X=b)}.$$

O *Odds Ratio* de duas classes não consideradas é igual ao Odds Ratio de cada uma dessas classes relativamente à classe de referência. Este raciocínio pode ser aplicado na comparação de *Odds Ratio* entre classes de variáveis independentes policotómicas (mais de duas classes, por exemplo, a opinião sobre um determinado bem: ("concordo", "discordo", "sem opinião").

O Intervalo de Confiança obtém-se, tal como no modelo logístico binomial, a partir da exponencial dos limites do intervalo de confiança dos coeficientes $\hat{\beta}_i \pm Z_{1-\alpha/2}\hat{\sigma}_{\hat{\beta}_i}$.

2.3. O Modelo de Regressão Ordinal

Os modelos de regressão logística ordinal permitem realizar uma análise de dados cuja variável resposta é apresentada em categorias com ordenação. A informação ordenada, na forma de escala tem sido cada vez mais utilizada em estudos epidemiológicos, tais como, qualidade de vida em escalas intervalares, indicadores de condição de saúde e mesmo de gravidade das doenças (Ananth CV, Kleinbaum DG, 1997). Estes modelos, dependendo do delineamento do estudo, permitem também calcular a estatística *odds ratio* (OR) ou a probabilidade de ocorrência de um evento. Existem vários modelos ordinais, tais como o modelo de *odds* proporcionais, modelo de razão contínua e modelo de categorias adjacentes. Apesar desta diversidade e da grande variedade de estudos sobre o assunto, a sua utilização ainda é escassa, devido à sua complexidade e dificuldade na validação dos pressupostos.

Tal como no modelo Logístico Binomial, antes de iniciar a modelação através dos modelos ordinais, esta deve sempre ser precedida pelo cruzamento de cada covariável com o evento de interesse. O teste do Qui-quadrado é um dos testes apropriados para a selecção dos efeitos principais, já que considera o carácter ordinal da variável resposta. Normalmente, utiliza-se um nível de significância conservador (geralmente entre 10% e 25%) para entrada das covariáveis no modelo. Além disso, pode-se estimar OR, considerando uma categoria da variável resposta como referência e comparando-a com as demais ou agrupando as categorias maiores e comparando-as às categorias menores.
No contexto da regressão ordinal, o modelo multinomial é visto como a referência do modelo *Logit*. Este termo surge, pois na regressão multinomial, o modelo encontra-se na forma parametrizada, comparando cada categoria (Y = k), com a categoria de referência (Y = 0). No entanto, nos modelos ordinais há que escolher quais as categorias da

variável resposta que se irão comparar. Suponhamos que se pretende comparar um determinado nível, com o nível imediatamente a seguir (ascendentemente na ordenação). Estamos então perante um *Modelo de Categorias Adjacentes*. O modelo pode ser considerado uma extensão do modelo de regressão multinomial e compara cada categoria da variável resposta com uma categoria de referência, normalmente a primeira categoria ou a última. Partimos do modelo logit

$$g_k(X) = \ln\left[\frac{\pi_k(X)}{\pi_0(X)}\right] = \beta_{k_0} + X'\beta_k,\ k = 1,2,...,k\ .$$

Assumindo que os odds não dependem da resposta e que estas podem ser escritas linearmente, então os logits para as categorias adjacentes escrevem-se da seguinte forma:

$$a_k(X) = \ln\left[\frac{P(Y=k\mid X)}{P(Y=k-1\mid X)}\right] = \ln\left[\frac{\phi_k(X)}{\phi_{k-1}(X)}\right] = \alpha_k + X'\beta,\ k = 1,2,...,k$$

com $X = \left[X_1\ X_2 \ldots X_p\right]$ e $\beta = \begin{bmatrix}\beta_1\\ \beta_2\\ \ldots\\ \beta_p\end{bmatrix}$.

Tomando um modelo apenas com uma variável explicativa e dois níveis, o Odds Ratio é dado por

$$OR_k = \frac{P(Y=k\mid X=1)/P(Y=k-1\mid X=1)}{P(Y=k\mid X=0)/P(Y=k-1\mid X=0)}.$$

Por outro lado, se quisermos comparar cada nível da variável resposta, com todos os níveis de resposta inferiores, isto é comparar $Y=k$ com $Y<k$, já estamos no âmbito dos *Modelos de Razão Contínua*. Este modelo permite comparar a probabilidade de uma resposta igual à categoria com determinado nível, digamos, *Y = j*, com a probabilidade de uma resposta maior, *Y > y*. O modelo logit é dado por

$$r_k(X) = \ln\left[\frac{P(Y=k\mid X)}{P(Y<k\mid X)}\right] = \ln\left[\frac{\phi_k(X)}{\phi_0(X)+\phi_1(X)+...+\phi_{k-1}(X)}\right] = \theta_k + X\beta_k, k = 1,2,...,k\ .$$

Este modelo possui diferentes interceptos e coeficientes para cada comparação e pode ser ajustado por *k* modelos de regressão logística binária. É mais apropriado quando há um interesse intrínseco numa categoria específica da variável resposta. Apesar de

possuir declives diferentes para cada categoria da variável resposta, o modelo pode ser aproximado por um vector constante de coeficientes para todas as categorias (tendo no entanto constantes diferentes), isto é, $r_k(X) = \theta_k + X\beta$.

Entre os modelos de regressão ordinal o modelo mais usado é o *Modelo de Odds Proporcionais*. Este modelo é mais indicado para situações em que a variável resposta original é contínua e foi posteriormente agrupada. O modelo é dado por:

$$o_k(X) = \ln\left[\frac{\sum_{j=0}^{k} P(Y=j \mid X)}{\sum_{j=k+1}^{K} P(Y=j \mid X)}\right] = \ln\left[\frac{\sum_{j=0}^{k} \phi_j}{\sum_{j=k+1}^{k} \phi_j}\right].$$

Neste modelo são considerados (k – 1) pontos de corte das categorias sendo que o j-ésimo ($j=1$, ..., $k-1$) ponto de corte é baseado na comparação de probabilidades acumuladas. Podemos escrever o modelo na forma (Abreu *et al*. 2009):

$$o_k(X) = \alpha_j + \left(\beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p\right),\ j = 1,...,k-1.$$

O termo α_j varia para cada uma das k categorias e cada β não depende do índice j, implicando que a relação entre o vector X e Y é independente da categoria. Logo, o modelo possui um pressuposto de *odds* proporcionais acerca dos (k-1) pontos de corte, também chamado pressuposto dos declives paralelos, que é assumido para cada covariável incluída no modelo. Este pressuposto deve ser testado para cada covariável separadamente e no modelo final.

No *software* R podemos ajustar este modelo através do pacote *Design* sendo possível obter os gráficos dos resíduos *score* e dos resíduos parciais para cada covariável. No gráfico do resíduo *score* se o pressuposto de *odds* proporcionais for válido é esperado que, para cada covariável, a tendência em torno das categorias da variável resposta tenha um comportamento horizontal constante. Já no gráfico do resíduo parcial, para um modelo bem ajustado espera-se que ocorra linearidade e que haja paralelismo entre as curvas para cada categoria da variável resposta.

Em todos os modelos ordinais mencionados a significância dos coeficientes pode ser testada através do teste de Wald. No entanto como o teste de Wald é conservativo pode-se testar a significância dos coeficientes pela razão de verosimilhanças.

A dada altura é importante avaliar a qualidade do ajustamento dos modelos de regressão logística ordinal, pois a falta de ajuste pode, por exemplo, levar a análises erradas na estimação de efeitos. A avaliação do ajuste pode detectar: covariáveis importantes;

interacções omitidas; casos em que a função de ligação (logit) não foi apropriada; casos em que a forma funcional da modelação das covariáveis não está correcta; e, finalmente, casos em que a suposição de *odds* proporcional foi violada.

Embora tenham sido desenvolvidos muitos métodos para avaliar o ajustamento de modelos de regressão logística binária, poucos foram estendidos para dados de resposta ordinal. Normalmente, a qualidade do ajustamento dos modelos ordinais é verificada usando os testes de Pearson ou *Deviance*. Estes testes envolvem a criação de uma tabela de contingência na qual as linhas contêm todas as possíveis configurações das covariáveis do modelo e as colunas são as categorias da resposta ordinal. As contagens esperadas dessa tabela são expressas por

$$E_{lj} = \sum_{l=1}^{N_l} \hat{p}_{lj} \ ,$$

onde N_l é o número total de indivíduos classificados na linha l e representa a probabilidade de um indivíduo na linha l ter a resposta j calculada a partir do modelo adoptado. O teste de Pearson avalia a adequação do ajuste comparando essas contagens esperadas com as observadas:

$$\chi^2 = \sum_{l=1}^{L} \sum_{j=1}^{K} \frac{\left(O_{lj} - E_{lj}\right)^2}{E_{lj}} \ .$$

Por sua vez, a *Deviance* fá-lo da forma:

$$D^2 = 2 \sum_{l=j}^{L} \sum_{j=1}^{K} O_{lj} \log \frac{O_{ij}}{E_{lj}} \ .$$

Um valor p significativo leva-nos a concluir a falta de ajuste do modelo aos dados estudados. Pulkstenis e Robinson (2004) relatam que as estatísticas de Pearson e da *Deviance*, não fornecem uma boa aproximação da distribuição Qui-quadrado quando são ajustadas covariáveis contínuas e propõem pequenas modificações neste caso. Hosmer e Lemeshow (2000) sugerem a utilização de regressões binárias separadas para cada ponto de corte, para que sejam criadas as estatísticas de diagnóstico para os modelos ordinais.

2.4. O Modelo MANOVA

A comparação de médias de duas ou mais populações (extracção de amostras aleatórias e independentes) pode fazer-se através da Análise de Variância (ANOVA), quando a distribuição da variável em estudo é Normal e as variâncias populacionais homogéneas. A ANOVA compara a variância dentro das amostras ou grupos (variância residual) com a variância entre as amostras ou grupos (variância do factor ou entre os grupos). Caso a variância residual seja significativamente inferior à variância entre os grupos ou amostras, então as médias populacionais estimadas a partir das amostras são significativamente diferentes.

Quando se rejeita H_0 na ANOVA, conclui-se que existe pelo menos uma média populacional que é significativamente diferente das restantes. No entanto a ANOVA não fornece informação sobre qual dos pares de médias são diferentes. Existem vários testes *post-hoc* de comparações múltiplas de médias, como por exemplo, Tukey, Bonferroni, Scheffé, etc. Neste trabalho não houve necessidade de utilizar nenhum dos testes, uma vez que os factores tinham apenas dois níveis (o factor escola só tem dois níveis por ter sido excluída a Escola Conde Vilalva por não ter nível secundário).

Na análise de variância multivariada (MANOVA), as variáveis dependentes são consideradas simultaneamente e organizadas de forma composta. Os efeitos são associados a cada variável ponderada pela correlação existente entre elas. A MANOVA permite identificar diferenças entre os grupos relativamente às variáveis.

A MANOVA pressupõe que os erros experimentais seguem uma distribuição normal multivariada (Normalidade); são independentes (Independência) e possuem uma matriz de variâncias-covariâncias homogénea (Homocedasticidade).
Podemos ter a MANOVA *"one-way"* (a um factor) onde consideramos um factor entre as amostras e queremos testar se as médias de *p* variáveis medidas em *m* grupos ou tratamentos replicados *n* vezes, diferem entre si. Para um vector aleatório Y_{ijk} com observações de p variáveis-resposta, o modelo da MANOVA pode escrever-se neste caso como

$$\mathbf{Y}_{ijk} = \mu_{ij} + \boldsymbol{\tau}_i + \boldsymbol{\mu}_{ijk},$$

onde μ_{ij} representa a média da variável *j* no grupo *i* e $k\left(k=1,...,n\right)$ é o número de repetições da variável *j* em cada grupo *i*; τ_i o vector dos efeitos fixos do factor e μ_{ijk} é o vector dos erros ou resíduos com distribuição normal multivariada, $N_p(\mathbf{0},\mathbf{\Sigma})$.

As hipóteses a testar são

$$\begin{cases} H_0 : \mu_{ij} = \mu_{lm} \ \left(\hat{\imath}, j \neq 1, m\right) \\ H_1 : \exists_{i,j,l,m} : \mu_{ij} \neq \mu_{lm} \end{cases}.$$

Enquanto, no caso da ANOVA, a estatística de teste F é a mais potente na comparação de duas ou mais médias, na MANOVA existem várias estatísticas de teste que podem ser utilizadas: Lambda de Wilks, o traço de Pillai, o traço de Hotteling-Lawley e o método de Roy. A estatística de teste de Lambda de Wilks é um dos métodos mais potentes quando estão validadas as condições de aplicação da MANOVA. Já o método de Pillai é mais potente no caso de amostras ou grupos de dimensões pequenas e diferentes e para covariâncias heterógeneas. Nos casos em que as variáveis estão fortemente intercorrelacionadas, o método de Roy é o mais potente.

Com dois factores passamos a ter um modelo da MANOVA "two-way". Considerando um vector aleatório Y_{ijk} com as observações de *p* variáveis, o modelo da MANOVA "two-way" com interacção escreve-se como:

$$\mathbf{Y}_{ijk} = \mu + \boldsymbol{\tau}_i + \boldsymbol{\beta}_j + \boldsymbol{\gamma}_{ij} + \boldsymbol{\varepsilon}_{ijk},$$

Neste modelo, μ representa a média geral para cada factor, $\boldsymbol{\tau}_i$ é o vector dos efeitos fixos do factor *A* para o grupo $i\left(i=1,...,a\right)$, $\boldsymbol{\beta}_j$ é o vector dos efeitos fixos do factor *B* para o grupo $j\left(j=1,...,b\right)$, $\boldsymbol{\gamma}_{ij}$ é o vector da interacção do grupo *i* com o grupo *j* e $\boldsymbol{\varepsilon}_{ijk}\left(k=1,...,n\right)$ é o vector dos erros com distribuição $N_p\left(\mathbf{0},\mathbf{\Sigma}\right)$. Pretendem-se testar as seguintes hipóteses:

$$\begin{cases} H_0{}^A : \mu_{\mathrm{ij}} = \mu_{lm} \ \ (\mathrm{i,j=1,...,p}) \, e\left(l,m=1,...,a\right) \\ H_1^A \, \exists i,j,l,m : \mu_{\mathrm{ij}} \neq \mu_{lm} \end{cases}, \text{ para o factor } A;$$

$$\begin{cases} H_0^{\,B} : \mu_{ij} = \mu_{lm} \quad (i,j=1,...,p)\, e\,(l,m=1,...,b) \\ H_1^{\,B}\ \exists i,j,l,m : \mu_{ij} \neq \mu_{lm} \end{cases}$$, para o factor *B*;

e $\begin{cases} H_0^{\,\gamma} : \gamma_{ij} = 0 \\ H_0^{\,\gamma} : \gamma_{ij} \neq 0 \end{cases} (i=1,...,a;\ j=1,...,b)$, para a interacção.

Tal como acontece na ANOVA, rejeita-se cada uma das hipóteses nulas se a respectiva soma de quadrados for superior à soma de quadrados dos resíduos, o problema está na comparação das matrizes. Começamos por testar o efeito de interacção, caso não seja significativo passamos para o teste dos factores (efeitos principais). Na MANOVA não se calculam os quadrados médios. Existem quatro critérios para testar cada uma das hipóteses, que podem ser aproximados pela distribuição da F-Snedecor. Para o teste *i*, as estatísticas de teste são:

Teste	Estatística de Teste
Lambda de Wilks	$\Lambda_A = \dfrac{\left\lvert \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij}\right)\left(y_{ijk}-\bar{y}_{ij}\right)' \right\rvert}{\left\lvert \sum_{i=1}^{a} bn\left(\bar{y}_i-\bar{y}\right)\left(\bar{y}_i-\bar{y}\right)' \right\rvert + \left\lvert \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij}\right)\left(y_{ijk}-\bar{y}_{ij}\right)' \right\rvert}$
Traço de Pillai	$V = tr\left[\dfrac{\sum_{i=1}^{a} bn\left(\bar{y}_i-\bar{y}\right)\left(\bar{y}_i-\bar{y}\right)}{\left(\sum_{i=1}^{a} bn\left(\bar{y}_i-\bar{y}\right)\left(\bar{y}_i-\bar{y}\right)' + \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij}\right)\left(y_{ijk}-\bar{y}_{ij}\right)' \right)} \right]$
Traço de Hotteling-Lawley	$U_A = tr\left(\dfrac{\sum_{i=1}^{a} bn\left(\bar{y}_i-\bar{y}\right)\left(\bar{y}_i-\bar{y}\right)'}{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij}\right)\left(y_{ijk}-\bar{y}_{ij}\right)'} \right)$
Raiz Máxima de Roy	R_A **o maior valor próprio da matriz** $\left(\dfrac{\sum_{i=1}^{a} bn\left(\bar{y}_i-\bar{y}\right)\left(\bar{y}_i-\bar{y}\right)'}{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(y_{ijk}-\bar{y}_{ij}\right)\left(y_{ijk}-\bar{y}_{ij}\right)'} \right)$

Deve-se escolher o critério cuja potência de teste "observed power" é mais elevada, pois quanto maior a potência do teste, maior o grau de confiança que se pode ter na conclusão obtida. Caso a MANOVA seja significativa, procede-se a múltiplas ANOVAS para identificar as variáveis e em seguida a testes de comparações múltiplas. O teste que deve ser analisado em primeiro lugar é o teste do efeito da interacção, pois caso se rejeite a hipótese nula H_0^{γ}, significa que existe interacção entre os factores e não se devem analisar os efeitos isolados. Neste caso deve-se proceder a testes de comparações múltiplas entre as médias das células.

Para que se possa aplicar o modelo MANOVA "two-way", os pressupostos a validar são a independência dos erros ε_{ijk} e distribuição normal multivariada de média zero e variâncias-covariâncias homogéneas (homocedasticidade). Geralmente, os métodos multivariados são robustos no que diz respeito à normalidade dos dados. Utilizando os testes de Kolmogorov-Smirnov, pode-se averiguar se cada uma das *p* variáveis $X_1, X_2, ..., X_p$ apresenta distribuição normal univariada de parâmetros μ_i e $\sigma_i, i = 1, 2, ..., p$. Se cada uma das variáveis, separadamente, seguir uma distribuição normal então o vector **X** também será normal multivariado. Embora, na prática, se possa aplicar esta regra, é possível demonstrar matematicamente que nem sempre é possível (Johnson, 1998). Como alternativa, Srivastava & Hui (1987) propõem um teste exacto com duas estatísticas de teste. Essas estatísticas de teste baseiam-se nas componentes principais presentes na matriz dos dados, e podem ser consideradas como uma generalização da estatística de Shapiro-Wilk. Outra forma passa pelo cálculo dos coeficientes de assimetria e achatamento para cada variável individualmente. Quando estes assumem valores próximos de zero, podemos então assumir a validação do pressuposto da multinormalidade.

No caso do pressuposto da homogeneidade das matrizes de variância-covariância, utiliza-se o teste M de Box, que é muito sensível à violação do pressuposto da normalidade. As hipóteses para o teste M de Box são:

$$\begin{cases} H_0 : \Sigma_1 = \Sigma_2 = ... = \Sigma_m \\ H_1 : \exists i, j : \Sigma_i = \Sigma_j \ (i \neq j; i, j = \{1, ..., m\}) \end{cases}.$$

As hipóteses são testadas através da seguinte estatística de teste

$$M = (N-m)\log|\mathbf{S}| - \sum_{i=1}^{m}(n-1)\log|\mathbf{S}|,$$

onde

$$\mathbf{S} = \frac{\sum_{i=1}^{m}(n-1)\,\mathbf{S_i}}{N-m}$$

é uma matriz de variâncias-covariâncias definida-positiva. A matriz $\mathbf{S_i}$ é a matriz de variâncias-covariâncias do grupo i e a regra de decisão é rejeitar H_0 se valor p $\leq \alpha$. No caso de termos as amostras com a mesma dimensão, a MANOVA é robusta à violação de homogeneidade das matrizes de variância-covariância. Se forem diferentes, então deve-se usar como estatística de teste o traço de Pillai para avaliar a significância (Tabacknick & Fidel, 1996). Analisamos o valor p em conjunto com a potência de teste. Por exemplo, se o valor p para o traço de Pillai for significativo e simultaneamente a estatística de teste para a Raiz de Roy também for, a escolha recai sobre o teste com maior potência.

2.5. Uma abordagem aos "*Missings Values*"

Quando o número de casos com dados omissos é pequeno (inferior a 5% nas amostras de maior dimensão), é comum eliminar esses casos da análise. Por outro lado, quando existem variáveis com mais de 5% de dados omissos devemos estimar valores para substituí-los.

Surgiram técnicas estatísticas que envolvem a imputação de dados omissos. Essas técnicas têm por objectivos "completar" os dados em falta e possibilitar a análise com todos os indivíduos do estudo. As primeiras técnicas de imputação desenvolvidas (Conversano e Cappelli, 2002) envolviam métodos relativamente simples, tais como, substituição dos dados omissos pela média, pela mediana, por interpolação ou até por regressão linear. Todas essas técnicas mencionadas permitem "preencher"os dados em falta por meio do que se chama de imputação única, ou seja, o dado ausente é preenchido uma única vez e então utiliza-se a base de dados completa para as análises. Entretanto, a incerteza associada à imputação deve ser levada em conta para que os resultados obtidos com os dados completos sejam válidos, pois os valores imputados

não são valores reais. Para solucionar essa questão foi desenvolvida a técnica de Imputação Múltipla (IM).

Podemos classificar os dados omissos em três grandes grupos:

a) Dados omissos completamente ao acaso (missing completely at random – MCAR) quando os valores em falta são distribuídos aleatoriamente em todas as observações;
b) dados omissos ao acaso (missing at random – MAR) quando os valores em falta não são distribuídos aleatoriamente em todas as observações, mas são distribuídos aleatoriamente dentro de uma ou mais sub-amostras);
c) perdas não-aleatórias (not missing at random – NMAR).

Para avaliarmos se os dados são MCAR, pode recorrer-se ao teste Little's MCAR, que é um teste do qui-quadrado. Se o valor p para o teste de Little's MCAR não é significativo, os dados podem ser assumidos como MCAR. Se os dados forem MCAR, devemos optar por supressão dos casos. Se os dados não são MCAR, os valores em falta devem ser imputados através de métodos de estimação. É mais comum os dados serem do tipo MAR do que MCAR.

Para testar se os dados omissos são MAR podemos fazê-lo usando o *software* SPSS, o qual gera uma tabela cujas linhas são todas as variáveis que têm 1% ou mais de dados em falta, e cujas colunas são todas as variáveis. Sendo o valor p associado inferior a 5%,, tal significa que os valores em falta na variável correspondente à linha são significativamente correlacionados com a variável da coluna e, portanto, não faltam ao acaso. Se os casos faltam ao acaso podem ser eliminados, mas de outra forma devemos proceder à imputação de valores.

As perdas não-aleatórias é a forma mais problemática. Surgem quando os valores ausentes não são distribuídos aleatoriamente em todas as observações, mas a probabilidade de omissão não pode ser prevista a partir das variáveis no modelo. Uma abordagem para as perdas não-aleatórias é imputar valores com base em dados externos à pesquisa.

Relativamente aos métodos de estimação, recorrendo ao SPSS dispomos de 4 métodos de estimação: *listwise, pairwise, regression,* e EM (*maximum likelihood estimation*). O método MLE (*Maximum likelihood estimation*) é implementado pelo algoritmo EM, que se baseia numa função de probabilidade que preenche os valores ausentes através de

uma distribuição que modela os dados (Dempster *et al*. 1977). Uma vez que este método exige menos dos dados em termos de pressupostos estatísticos e é geralmente considerado superior à imputação por regressão múltipla é um dos métodos mais comum de imputação. O método MLE assume que os valores em falta são MAR.

2.5.1. Imputação múltipla (MI)

A imputação múltipla (MI) é um método de gerar vários valores simulados para cada dado em falta. O objectivo é, indiscutivelmente, gerar estimativas que reflictam melhor a variabilidade verdadeira e a incerteza nos dados. Os valores simulados são gerados pelo método de Monte Carlo. O método de MI tem a vantagem da simplicidade relativamente ao MLE, tornando-o particularmente adequado para grandes conjuntos de dados. Nesta perspectiva, tem sido demonstrado que a eficiência da imputação de dados usando o MI é elevada, mesmo quando o número de conjuntos de dados imputados é baixo (na faixa de 3 a 10). A MI consiste em usar mais do que um valor para preencher os valores ausentes numa amostra, isto é, utiliza-se a média dos valores prováveis (Rubin, 1987). Além disso, os testes sugerem que a MI é bastante robusta, mesmo quando a simulação é baseada num modelo errado (ex., quando a normalidade é hipótese para fins de simulação, quando os dados subjacentes não são, de facto, normais).

Este método consiste em três passos:

1. Obter *m* bancos de dados completos por meio de técnicas adequadas de imputação;
2. Analisar separadamente os *m* bancos por um método estatístico tradicional, como se realmente fossem conjuntos completos de dados;
3. Os *m* resultados encontrados no passo anterior são combinados de forma a obter a chamada inferência da imputação repetida.

O primeiro passo é a parte fundamental da IM, pois as técnicas de imputação utilizadas têm de preservar a relação das observações omissas e presentes e ainda levar em conta o mecanismo de ausência e o padrão dos dados omissos. Os mecanismos dividem-se em: perdas completamente ao acaso (*missing completely at random* - MCAR), perdas ao

acaso (*missing at random* - MAR) e perdas não-aleatórias (*not missing at random* - NMAR).

A partir das *m* imputações realizadas, o passo 2 da IM pode ser realizado, ou seja, os *m* bancos de dados são analisados por métodos tradicionais de análise. Finalmente, os *m* resultados obtidos podem ser combinados usando-se as regras propostas por Rubin (1987).

As regras de Rubin (*Rubin rules*) estão amplamente divulgadas na literatura que trata de IM, pois são normas simples que resolvem o último passo da IM. Essas regras podem ser usadas independentemente do método utilizado para fazer a IM. A ideia é que a partir de cada análise sejam obtidas as estimativas para o parâmetro de interesse *Q*, ou seja, Q_j para j = 1, 2,...,m.

Segundo Harrell (2002) é possível definir linhas gerais para a escolha entre os métodos de imputação de acordo com a proporção de dados omissos em alguma das variáveis:

a) proporção $\leq$ 0,05 - neste caso pode ser usada a imputação única ou analisar somente os dados completos;
b) proporção entre 0,05 e 0,15 - a imputação única pode ser usada provavelmente sem problemas, entretanto o uso da imputação múltipla é indicado;
c) proporção $\geq$ 0,15 - a imputação múltipla é indicada na maior parte dos modelos.

No caso de haver muitas variáveis explicativas com dados omissos devem ser feitas as mesmas considerações acima, mas os efeitos das imputações serão mais pronunciados.

Capítulo III - Análise Exploratória de Dados e Modelação Estatística do IMC

Neste capítulo será feita uma análise exploratória dos dados para cada variável utilizada. Para modelar o IMC (índice de massa corporal) será ajustado um modelo de Regressão Logística Binomial e um modelo de Regressão Logística Multinomial com o objectivo de estudar a relação entre o IMC e variáveis como o sexo, idade e escola a que pertencem os alunos.

3.1. Análise Exploratória de Dados

A amostra utilizada no estudo provém de três escolas situadas em locais diferentes, tratando-se de uma amostragem dirigida. Neste caso o plano de amostragem não é aleatório. A amostra é constituída por 2234 crianças/adolescentes (1099 Raparigas e 1135 rapazes), com idades compreendidas entre 8 e 20 anos e que frequentam o ensino regular (5º ao 12º anos) e Cursos de Educação e Formação. As crianças/adolescentes do estudo pertencem a duas escolas do Alentejo, Escola EB 2,3/S Drº Hernâni Cidade de Redondo (451) e Escola EB 2,3 Conde Vilalva de Évora (561), e a um agrupamento de escolas do Norte, Agrupamento Vertical de Escolas de Castelo de Paiva (1222). Os dados para este estudo foram recolhidos pelos professores de Educação Física durante as aulas e foram gentilmente cedidos pelas escolas envolvidas.

Com o intuito de calcular o Índice de Massa Corporal (IMC) de todos os alunos, foram recolhidos o peso e a altura de cada um, bem como a idade e sexo dos alunos das três escolas consideradas. Como não existem valores tabelados do IMC para as idades de 8 e 9 anos, os alunos que apresentavam estas idades foram excluídos da amostra. Foram ainda recolhidos, pelos professores de Educação Física, os resultados dos testes de *Fitnessgram*: abdominais, vai-vém e flexibilidade à esquerda e à direita. Foi introduzida a variável ciclo, pois a amostra contém alunos do 2º ciclo (5º e 6º anos), 3º ciclo (7º, 8º e 9º anos) e secundário.

3.1.1. Escolas

As escolas envolvidas neste estudo encontram-se inseridas em meios envolventes diferentes: a Escola Conde Vilalva está situada na cidade de Évora e é frequentada por alunos que moram nessa cidade, a Escola de Redondo é frequentada por alunos que

moram na cidade de Redondo, vilas, aldeias e localidades pertencentes ao Conselho de Redondo, conselho este tradicionalmente rural. À semelhança da escola de Redondo, as escolas do Agrupamento Vertical de Castelo de Paiva estão inseridas num concelho também tradicionalmente rural e são frequentadas por alunos que moram na Cidade de Castelo de Paiva em vilas, aldeias e localidades pertencentes a este concelho. Tal como referido acima, o número de alunos difere de escola para escola: Escola EB 2,3 Conde Vilalva de Évora tem 561 alunos, a Escola EB 2,3/S Drº Hernâni Cidade de Redondo tem 451 alunos e o Agrupamento Vertical de Escolas de Castelo de Paiva 1222 alunos. Observando-se o gráfico circular (Figura 1), podemos ver a distribuição dos indivíduos por escola. Mais de 50% dos indivíduos da amostra pertencem ao agrupamento de escolas de Castelo de Paiva.

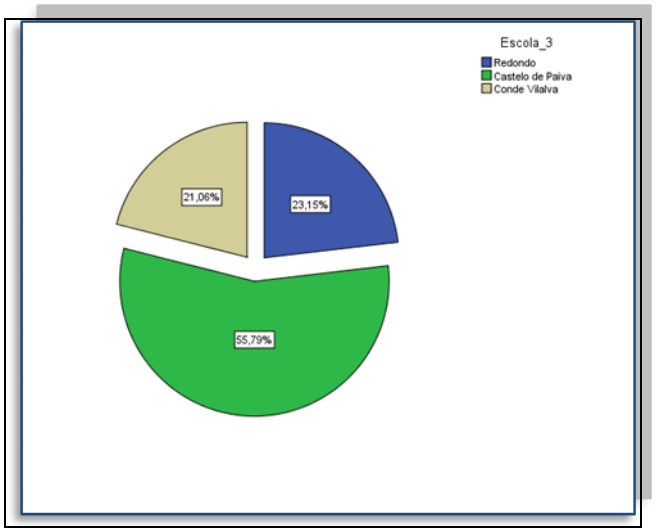

Figura 1: Distribuição de alunos por Escola

3.1.2. Sexo

Relativamente ao género, na Figura 2, observamos que o número de rapazes é ligeiramente superior ao número de raparigas em Redondo e Castelo de Paiva, enquanto -na escola Conde Vilalva o número de raparigas é ligeiramente superior.

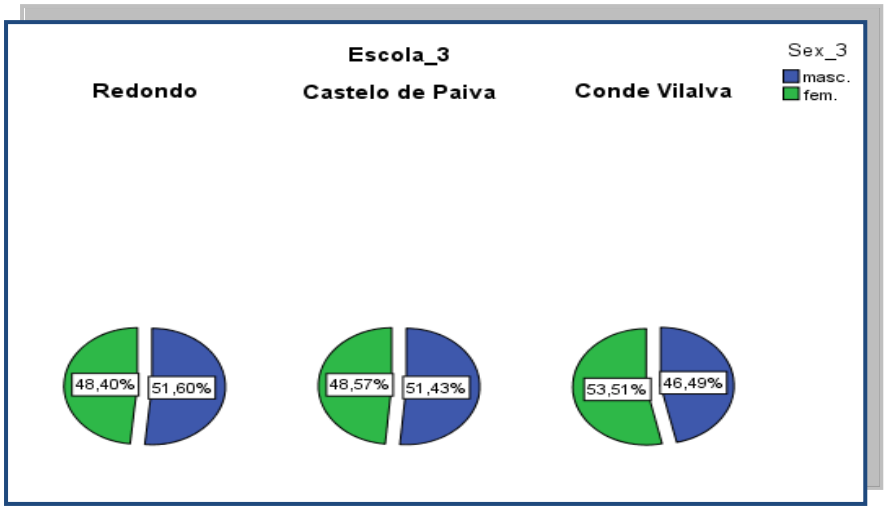

Figura 2: Distribuição por género e Escola

3.1.3. Idade

Como se pode observar na Figura 3, a distribuição das idades é muito semelhante para as escolas de Redondo e Castelo de Paiva, sendo aproximadamente simétricas. A distribuição das idades para a escola Conde Vilalva é assimétrica negativa. Os valores das médias e desvios-padrão das idades nas três escolas foram: Escola do Redondo $13,57 \pm 2,038$, Escola de Castelo de Paiva $13,77 \pm 2,092$ e Escola Conde de Vilalva $13,54 \pm 1,537$. Observe-se ainda a existência de alguns *outliers* para Castelo de Paiva e Conde Vilalva.

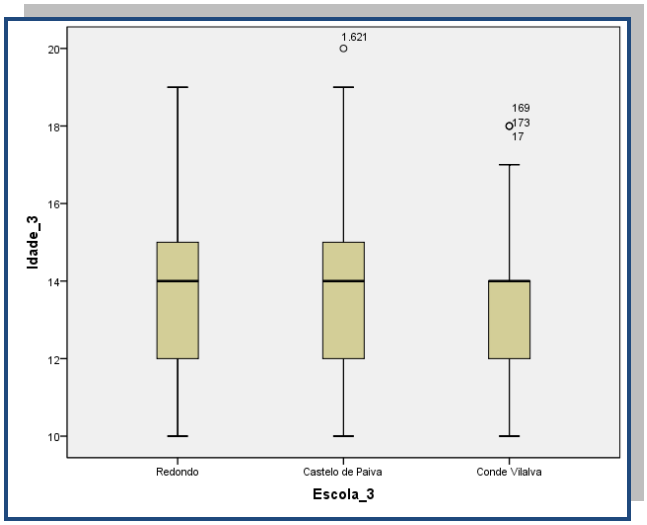

Figura 3: Caixa de bigodes para a Idade por Escola

Na Tabela 1 apresentam-se as medidas descritivas para a variável idade. Globalmente as idades variam entre os 10 e 20 anos, metade dos alunos tem mais de catorze anos, 25% têm até 12 anos e 75% tem menos de 15 anos. Globalmente esta distribuição é assimétrica positiva e platicúrtica.

	Idade
Média	13,68
Mediana	14,00
Moda	14,00
Desvio-Padrão	1,98
Mínimo	10,00
Máximo	20,00
Percentil 10	11,00
Primeiro Quartil	12,00
Terceiro Quartil	15,00
Percentil 90	16,00
Assimetria	0,36
Erro-Padrão (Assimetria)	0,06
Achatamento	-0,42
Erro-Padrão (Achatamento)	0,12

Tabela 1: Medidas descritivas para a variável idade

3.1.4. Peso e Altura

Como se pode observar na Figura 4, a distribuição das alturas é idêntica para as três escolas, sendo aproximadamente simétrica. Os valores observados das médias e desvios-padrão das alturas nas três escolas foram: Escola do Redondo $157 \pm 12,6$ cm, Escola de Castelo de Paiva $159 \pm 12,3$ cm e Escola Conde de Vilalva $159 \pm 10,9$ cm.

Em relação à variável "peso", observando-se a Figura 5, podemos verificar que as distribuições dos pesos são praticamente simétricas, aparecendo *outliers* em todas as escolas, correspondendo a alunos com excesso de peso. Os valores das médias e desvios-padrão dos pesos foram: Escola de Redondo $50,2 \pm 14,3$ kg, Escola de Castelo de Paiva $52,6 \pm 12,9$ kg e Escola de Conde Vilalva $49,1 \pm 13,4$ kg. O teste de Levene com um valor *p* igual a 0,3 permite-nos admitir a igualdade das dispersões do peso nas três escolas. Uma vez que não foi possível realizar uma ANOVA, para se testar a igualdade dos pesos médios, devido ao grande afastamento dos dados à normalidade, realizou-se o teste de Kruskal-Wallis tendo-se obtido um valor $p < 0,001$ que revela

uma diferença altamente significativa entre as medianas dos pesos dos alunos das três escolas.

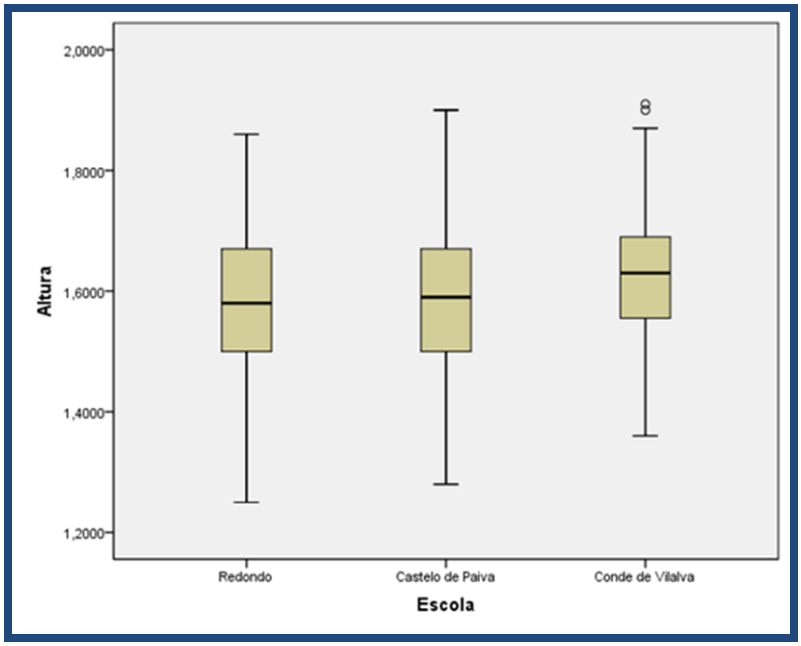

Figura 4: Caixa de bigodes para a altura dos alunos por escola.

Aplicando um teste não paramétrico de comparação múltipla podemos concluir, com uma significância de 5%, que os alunos da Escola de Castelo de Paiva são mais pesados, em termos medianos, que os alunos da Escola de Redondo. Ainda se verificou que os alunos de Castelo de Paiva são os mais pesados e os de Conde Vilalva os menos pesados.

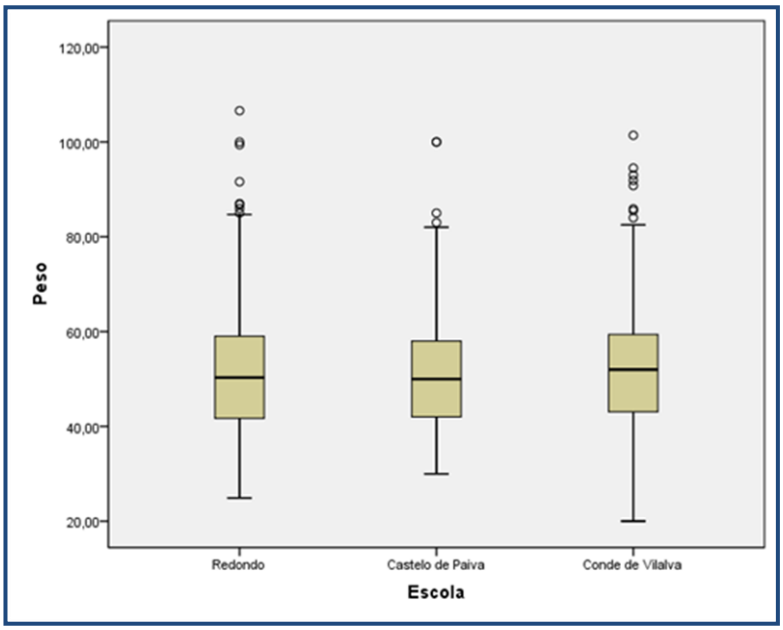

Figura 5: Caixa de bigodes para o peso dos alunos por escola.

Na Tabela 2 apresentam-se as medidas descritivas para as variáveis peso e altura da amostra. A altura média é aproximadamente 1,60m com desvio padrão de 0,12m. Para a variável peso, observa-se um peso médio de aproximadamente 51 kg com desvio padrão de 12,27 kg. É de notar que a variabilidade do peso é muito superior à da altura, pois o coeficiente de variação do peso é 24% enquanto para a altura é 7%. Na amostra encontram-se alunos com peso muito elevado para a idade que têm, o que para alguns casos pode conduzir a um índice de IMC elevado.

	Altura	Peso
Média	1,59	51,37
Mediana	1,60	50,20
Moda	1,52	50,00
Desvio-Padrão	0,12	12,27
Mínimo	1,25	20,00
Máximo	1,91	106,60
Percentil 10	1,44	36,00
Primeiro Quartil	1,51	42,00
Terceiro Quartil	1,67	58,77
Percentil 90	1,74	67,00
Assimetria	-0,06	0,61
Erro-Padrão (Assimetria)	0,06	0,06
Achatamento	-0,41	0,73
Erro-Padrão (Achatamento)	0,12	0,12

Tabela 2: Medidas amostrais quanto ao peso e altura.

3.1.5. Índice de Massa Corporal

Começamos por observar a Figura 6, a partir da qual podemos verificar uma distribuição do Índice de Massa Corporal (IMC) quase idêntica para as três escolas. No caso do IMC, pode-se verificar uma ligeira assimetria positiva, aparecendo vários *outliers* para as três escolas, todos correspondendo a um valor de IMC elevado. Os valores para as médias e desvios-padrão do IMC foram: Escola do Redondo $20,33 \pm 3,95$ cm, Escola de Castelo de Paiva $19,94 \pm 2,49$ cm e Escola Conde de Vilalva $19,48 \pm 4,16$. Devido à existência dos *outliers*, a média não é representativa.

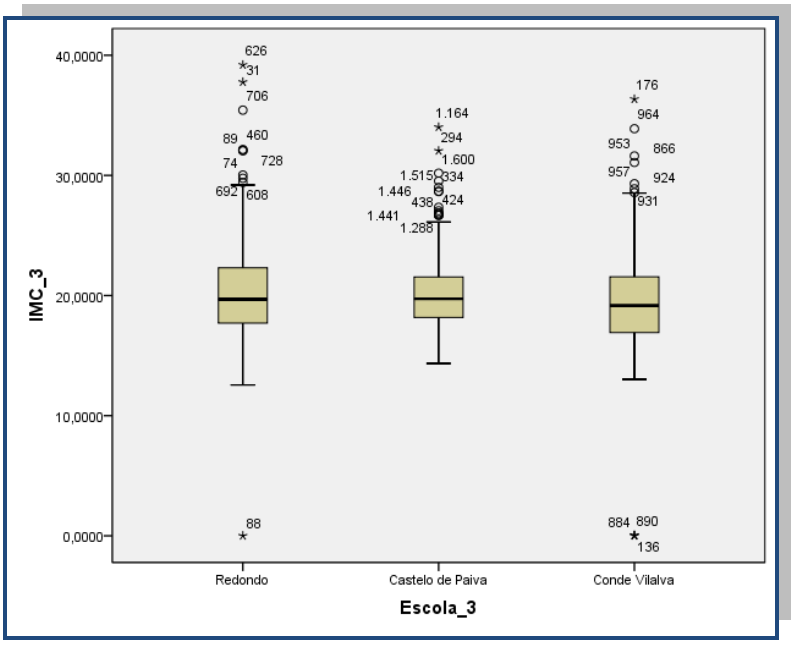

Figura 6: Caixa de bigodes para o IMC dos alunos por Escola

Na Tabela 3 apresentam-se as medidas descritivas para a variável IMC segundo o sexo dos alunos. Curiosamente, os rapazes (n=819) e as raparigas (n=805) apresentam valores muito próximos para todas as medidas descritivas de localização e de dispersão. Repare-se também que esta distribuição é simétrica, mas acentuadamente leptocúrtica, o que significa que as caudas são longas e pesadas, havendo uma grande probabilidade de ter valores extremos.

	Masculino	Feminino
Média	20,00	19,93
Mediana	19,72	19,49
Desvio-Padrão	3,17	3,40
Mínimo	<0,01	0,01
Máximo	35,43	39,20
Percentil 10	16,53	16,42
Primeiro Quartil	18,02	17,80
Terceiro Quartil	21,79	21,63
Percentil 90	23,81	23,82
Assimetria	0,12	0,70
Erro-Padrão (Assimetria)	0,08	0,09
Achatamento	4,68	6,00
Erro-Padrão (Achatamento)	0,17	0,17

Tabela 3: – Medidas descritivas por sexo quanto ao IMC.

Posteriormente iremos categorizar a variável IMC classificando os alunos em estudo como pertencendo ou não à zona considerada aceitável (Zona Saudável) para o respectivo índice de IMC segundo a idade e o peso de cada um. Será ainda criada uma nova variável que categoriza os alunos segundo o índice de IMC de cada aluno estar abaixo, dentro ou acima da Zona Saudável.

3.1.6. Ciclo

A variável ciclo permite categorizar os indivíduos pelo ciclo a que pertencem. Pela observação da Figura 7 verificamos que a maior parte dos alunos pertencem ao ensino básico (7º, 8º e 9º anos) e o menor número de alunos ao Secundário, o que já era de esperar uma vez que o secundário não era ensino obrigatório aquando da recolha de dados.

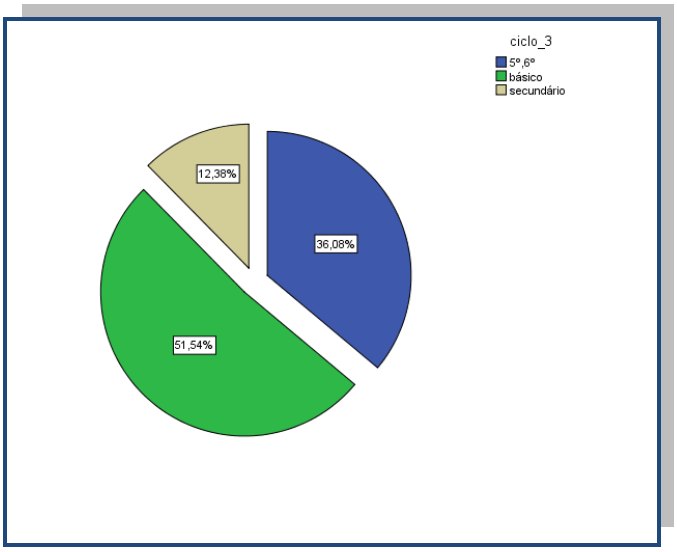

Figura 7: Gráfico circular para o ciclo dos alunos

3.1.7. Flexibilidade Esquerda

Para o caso dos resultados obtidos pelos alunos no teste da flexibilidade à esquerda, apenas são consideradas as escolas de Redondo e Conde Vilalva. A escola de Castelo de Paiva não forneceu os resultados reais, apenas dando informação sobre o aluno ter atingido ou não o índice desejado. Pela observação da Figura 8, verifica-se que na Escola Conde Vilalva a distribuição dos valores da flexibilidade à esquerda é

aproximadamente simétrica e tem a presença de um *outlier*. Já os resultados do teste na Escola de Redondo apresentam uma ligeira assimetria negativa. Os valores das médias e desvios-padrão para o teste da flexibilidade à esquerda foram: Escola do Redondo $20,69 \pm 7,57$, e Escola Conde de Vilalva $21,78 \pm 8,12$.

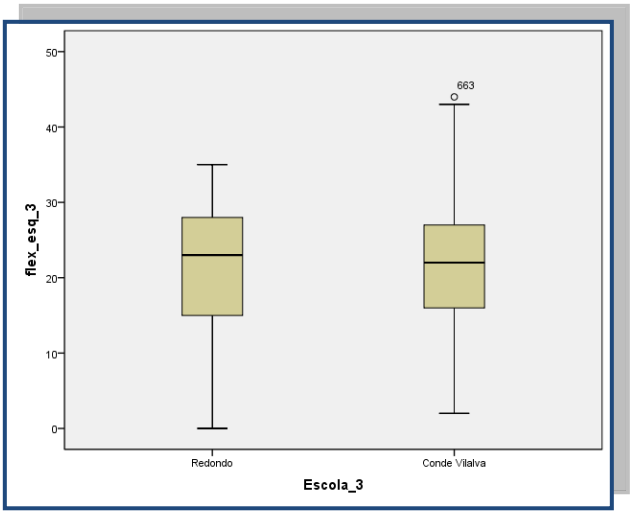

Figura 8: Caixa de bigodes para a flexibilidade à esquerda por Escola

Na Figura 9 podemos observar que o número de rapazes que não supera o valor de referência é inferior ao das raparigas.

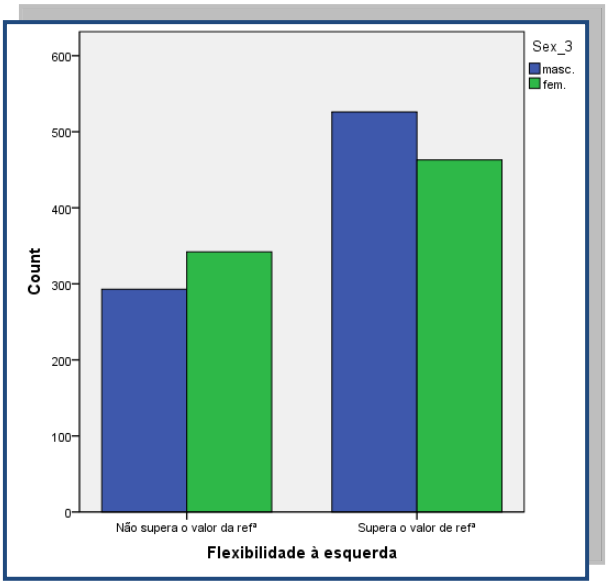

Figura 9: Gráfico de barras para a flexibilidade à esquerda

3.1.8. Flexibilidade Direita

Pelas razões referidas anteriormente, para o caso dos resultados obtidos pelos alunos no teste da flexibilidade à direita, apenas são consideradas as escolas de Redondo e Conde Vilalva. Pela observação da Figura 10 verifica-se que os resultados da Escola Conde Vilalva têm uma distribuição aproximadamente simétrica. Para Redondo surgem a presença de alguns *outliers* representando valores consideravelmente pequenos e os resultados do teste nesta escola apresentam uma ligeira assimetria negativa. Os valores das médias e desvios-padrão para o teste da flexibilidade à direita foram: Escola do Redondo $21,32 \pm 6,86$, e Escola Conde de Vilalva $22,94 \pm 7,47$.

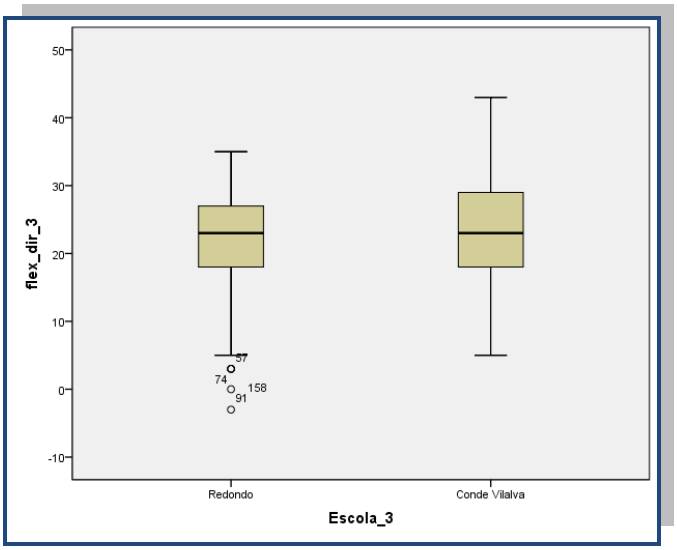

Figura 10: Caixa de bigodes para a flexibilidade à direita

Na Figura 11, podemos observar que o número de rapazes que não supera o valor de referência é inferior ao das raparigas.

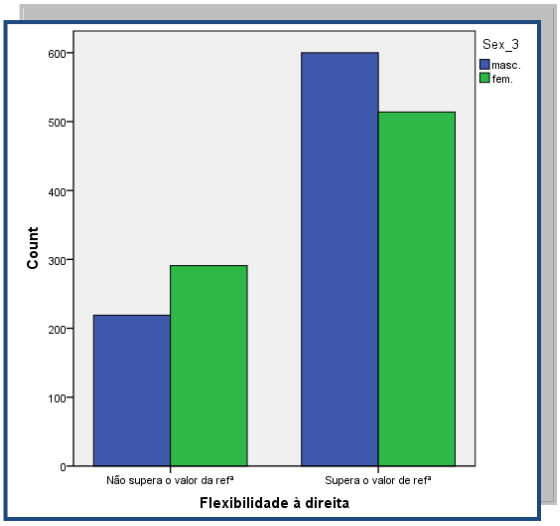

Figura 11: Gráfico de barras para a flexibilidade à direita

3.1.9. Vai-vém

Também para esta variável, apenas são consideradas as escolas de Redondo e Conde Vilalva. Pela observação da Figura 12 verifica-se que em Redondo e Conde Vilalva a distribuição dos valores é assimétrica positiva, sendo que os alunos de Redondo têm resultados superiores. Em ambas as escolas temos a presença de alguns *outliers*.

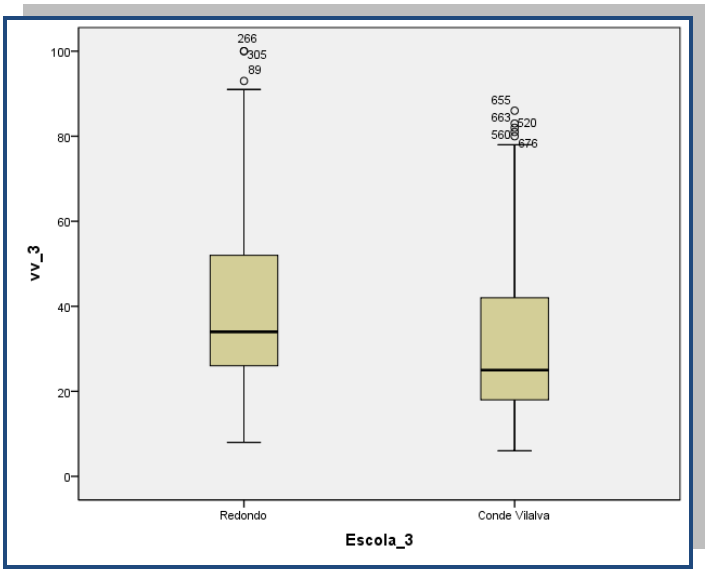

Figura 12: Caixa de bigodes para o vai-vém por Escola

Os valores das médias e desvios-padrão para o teste vai-vém foram: Escola do Redondo $40,47 \pm 19,26$, e Escola Conde de Vilalva $30,40 \pm 16,84$. Na Figura 13 podemos observar que o número de rapazes que não supera o valor de referência é inferior ao das raparigas.

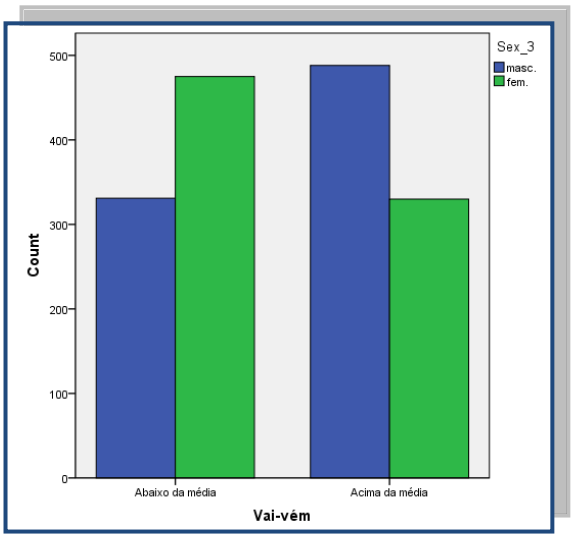

Figura 13: Gráfico de barras para o vai-vém

3.1.10. Abdominais

Também para esta variável apenas são consideradas as escolas de Redondo e Conde Vilalva. Pela observação da Figura 14 verifica-se que nas duas Escolas a distribuição dos valores é assimétrica positiva, havendo vários *outliers*, que representam valores consideravelmente grandes.

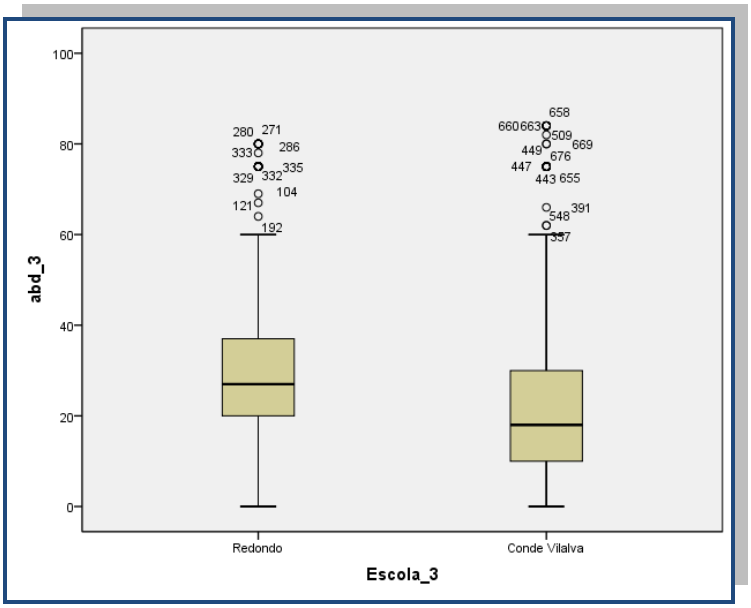

Figura 14: Caixa de bigodes para os abdominais por Escola

Os valores das médias e desvios-padrão para o teste de abdominais foram: Escola de Redondo $31,74 \pm 18,35$, e Escola Conde de Vilalva $24,35 \pm 21,75$. Na Figura 15, podemos observar que o número de rapazes que não supera o valor de referência é inferior ao das raparigas.

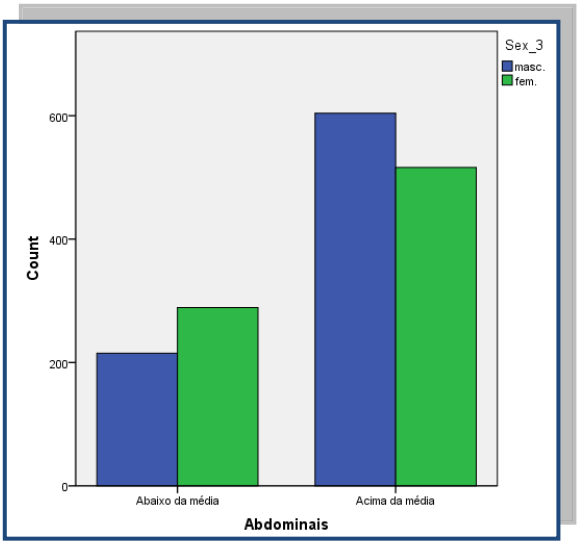

Figura 15: Gráfico de barras para os abdominais

3.1.11. Notas de Matemática

Como se pode observar na Figura 16 a distribuição das notas de Matemática regista uma forte assimetria para as escolas de Redondo e Castelo de Paiva, sendo simétrica na escola de Castelo de Paiva. Na escola de Redondo a distribuição das notas é fortemente assimétrica positiva, enquanto em Castelo de Paiva a distribuição é fortemente assimétrica negativa. As três escolas apresentam idêntico valor mediano para as notas de Matemática. Os valores das médias e desvios-padrão para as notas de Matemática foram: Escola de Redondo $3{,}17 \pm 0{,}87$, Escola de Castelo de Paiva $2{,}97 \pm 0{,}84$ e Escola de Conde Vilalva $3{,}15 \pm 0{,}99$.

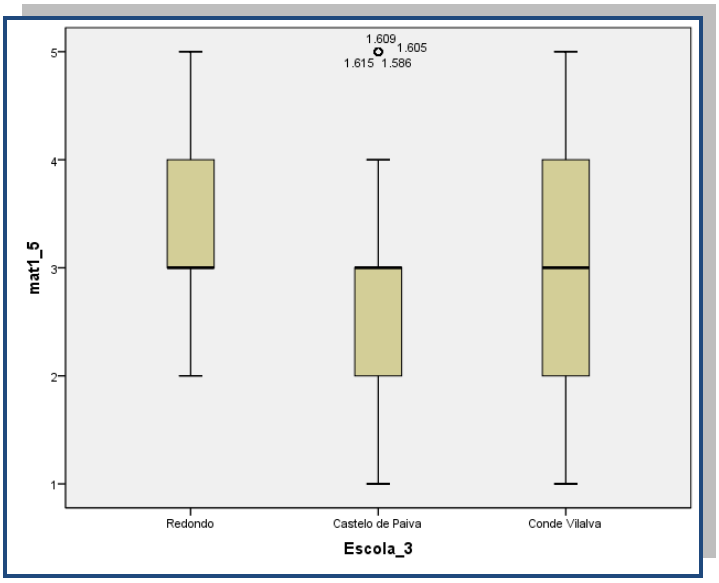

Figura 16: Caixa de bigodes para as notas de Matemática por Escola

3.1.12. Notas de Língua Portuguesa

Em relação à variável "nota de Língua Portuguesa", observando a Figura 17, podemos verificar uma forte assimetria positiva. Os valores das médias e desvios-padrão para as notas de Língua Portuguesa foram: Escola de Redondo $3{,}17 \pm 0{,}71$, Escola de Castelo de Paiva $3{,}15 \pm 0{,}74$ e Escola de Conde Vilalva $3{,}32 \pm 0{,}81$. O teste de Levene com valor *p* igual a 0, 52, permite admitir a igualdade das dispersões das notas nas três escolas.

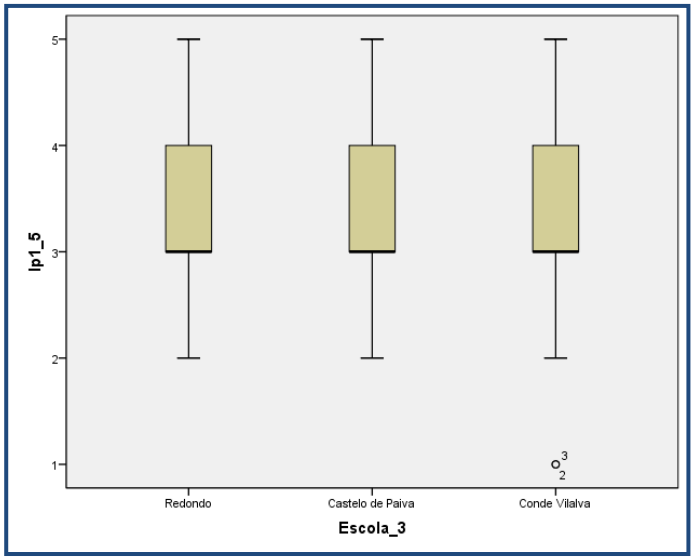

Figura 17: Caixa de bigodes para as notas de Língua Portuguesa por Escola

3.1.13. Notas de Educação Física

Em relação à variável "nota de Educação Física", observando a Figura 18 podemos verificar uma forte assimetria negativa. Os valores das médias e desvios-padrão para as notas de Educação Física foram: Escola de Redondo $3,72 \pm 0,74$, Escola de Castelo de Paiva $3,86 \pm 0,75$ e Escola de Conde Vilalva $3,65 \pm 0,72$.

Na amostra recolhida registámos 1,1% de negativas e 98,9% de positivas. Posteriormente houve necessidade de categorizar a variável referente às notas de Educação Física, como ter obtido nível cinco (17,8%) ou não (82,2%).

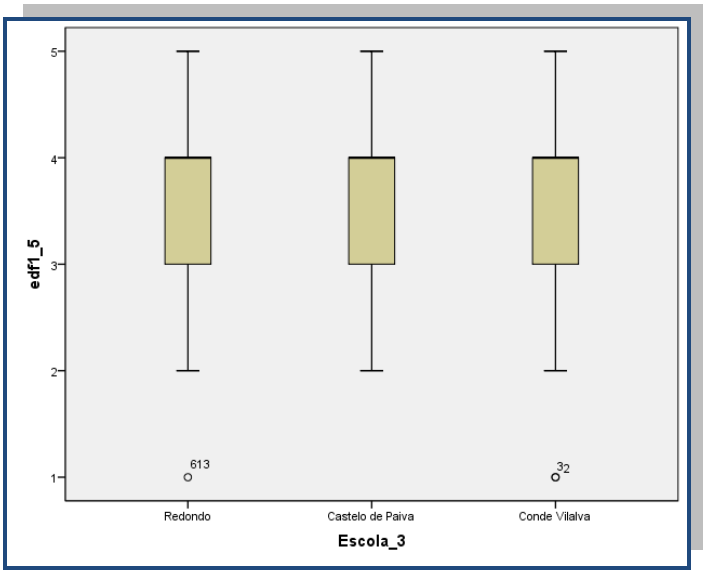

Figura 18: Caixa de bigodes para as notas de Educação Física por Escola

3.1.14. Zona Saudável de IMC

A variável IMC é uma variável contínua que nos dá o índice de IMC para cada aluno a partir do seu peso e altura. A partir das tabelas de referência do IMC para cada um dos sexos criou-se uma nova variável ("ZN_IMC") que classifica os alunos em estudo dentro ou fora da zona saudável. Pela Figura 19 verificamos que a maior parte dos indivíduos se encontram dentro da Zona Saudável para os índices de IMC, mas é de salientar que quase aproximadamente ¼ dos alunos está fora da zona saudável.

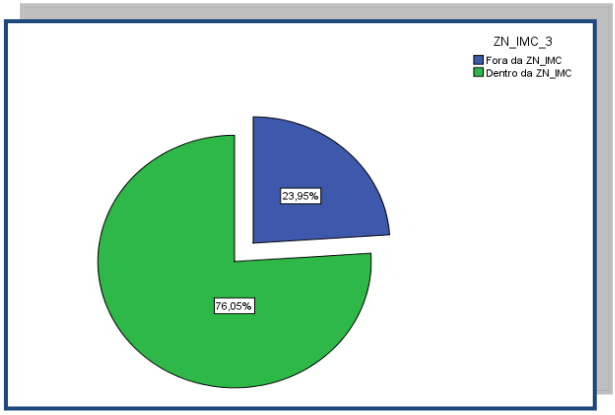

Figura 19: Gráfico circular para a Zona de IMC

3.2. Associações entre Variáveis

Com o intuito de se explicar a variável dependente "ZN_IMC" foram consideradas no estudo as variáveis "idade", "sexo", e ainda a variável "escola" que diferencia as escolas em estudo.

Começamos por analisar a possibilidade de existência de associação entre o sexo e o pertencer à Zona Saudável condicional à escola. Em ambos os períodos, não há evidência na nossa amostra para rejeitar a hipótese (valor p=0,18) da independência condicional, isto é, o sexo dos alunos é independente de pertencer ou não à Zona Saudável em termos de IMC.

Procedemos em seguida a uma análise idêntica colocando, desta vez, a idade como variável explicativa e continuando com a escola como variável de estratificação. Procuramos desta forma detectar diferenças entre as escolas na classificação "Dentro da Zona saudável" ou "Fora da Zona saudável" consoante as idades dos alunos. Para tal a variável "idade" foi categorizada em três classes: 8-12 anos, 13-15 anos, e 16 ou mais.

Observando a tabela 4, verificamos que nas escolas de Redondo e de Castelo de Paiva, o número de alunos da classe etária mais baixa "Dentro da Zona Saudável" é inferior ao que seria de esperar em caso de independência entre classe etária e classificação, e nas restantes classes são mais os que se encontram "Dentro da Zona Saudável" em relação ao valor esperado. Na escola Conde Vilalva, para as classes dos 8-12 anos e 16 ou mais, temos menos alunos "Dentro da Zona Saudável" do que seria esperado no caso de haver independência, sendo a classe dos 13-15 anos onde se encontram mais alunos "Dentro da Zona Saudável" do que era esperado. Estas diferenças, neste caso, são altamente significativas, pois aplicando o teste de Cochran-Mantel-Haenszel (HOSMER & LEMESHOW, 2000) generalizado obtemos um valor $p < 0,001$. Consequentemente rejeitamos a hipótese de independência entre classe etária e Zona de IMC condicional à escola.

Analisando a significância do teste do Qui-Quadrado para cada escola (Tabela 5), no primeiro período, verificamos que na Escola de Redondo não há associação entre a variável "idade" e o facto de os alunos pertencerem ou não à Zona Saudável, isto é, não há dependência entre a idade e a Zona Saudável estabelecida para cada intervalo do IMC correspondente à idade. Já nas escolas de Castelo de Paiva e Conde Vilalva verifica-se a existência de associação entre a "idade" dos alunos e o pertencer ou não à Zona Saudável associada ao IMC de cada um. Para o terceiro período, confirma-se a existência de associação entre a "idade" dos alunos e o pertencer ou não à Zona Saudável associada ao IMC de cada um, para as mesmas escolas.

Em seguida categorizou-se a variável "idade" em duas classes tendo como idade de corte os 14 anos, uma vez que é nesta idade em que os alunos adoptam comportamentos e hábitos diferentes devido ao despertar da sua sexualidade. Assim, consideram-se duas classes: 8-14 anos e 15 ou mais anos.

Escola					ZN_IMC		
					Fora da Zona Saudável	Dentro da zona Saudável	Total
1º P	Redondo	Idadecat1	08-Dez	Count	65	104	169
				Expected Count	58,6	110,4	169
			13-15	Count	63	117	180
				Expected Count	62,4	117,6	180
			>=16	Count	25	67	92
				Expected Count	31,9	60,1	92
			Total	Count	153	288	441
				Expected Count	153	288	441
	Castelo de Paiva	Idadecat1	08-Dez	Count	136	295	431
				Expected Count	87,7	343,3	431
			13-15	Count	81	393	474
				Expected Count	96,4	377,6	474
			>=16	Count	29	275	304
				Expected Count	61,9	242,1	304
			Total	Count	246	963	1209
				Expected Count	246	963	1209
	Conde Vilalva	Idadecat1	08-Dez	Count	103	102	205
				Expected Count	88,1	116,9	205
			13-15	Count	112	193	305
				Expected Count	131	174	305
			>=16	Count	26	25	51
				Expected Count	21,9	29,1	51
			Total	Count	241	320	561
				Expected Count	241	320	561
3º P	Redondo	Idadecat1	08-Dez	Count	53	102	155
				Expected Count	48,7	106,3	155
			13-15	Count	60	126	186
				Expected Count	58,5	127,5	186
			>=16	Count	19	60	79
				Expected Count	24,8	54,2	79
			Total	Count	132	288	420
				Expected Count	132	288	420
	Castelo de Paiva	Idadecat1	08-Dez	Count	56	229	285
				Expected Count	42,4	242,6	285
			13-15	Count	71	377	448
				Expected Count	66,7	381,3	448
			>=16	Count	26	269	295
				Expected Count	43,9	251,1	295
			Total	Count	153	875	1028
				Expected Count	153	875	1028
	Conde Vilalva	Idadecat1	08-Dez	Count	85	83	168
				Expected Count	65,9	102,1	168
			13-15	Count	92	181	273
				Expected Count	107,1	165,9	273
			>=16	Count	20	41	61
				Expected Count	23,9	37,1	61
			Total	Count	197	305	502
				Expected Count	197	305	502

Tabela 4: idadecat1 * ZN_IMC * Escola Crosstabulation

Chi-Square Tests					
	Escola		Value	df	Asymp. Sig. (2-sided)
1º Período	**Redondo**	Pearson Chi-Square	3,362	2	O,186
	Castelo de Paiva	Pearson Chi-Square	58,417	2	0,000
	Conde Vilalva	Pearson Chi-Square	10,622	2	0,005
3º Período	**Redondo**	Pearson Chi-Square	2,605	2	O,272
	Castelo de Paiva	Pearson Chi-Square	14,018	2	0,001
	Conde Vilalva	Pearson Chi-Square	13,665	2	0,001

Tabela 4: Teste Qui-quadrado para as escolas

Pela observação da Tabela 6, no primeiro período, verifica-se que tanto na escola de Redondo como na escola de Castelo de Paiva, os alunos mais novos que estão "Dentro da Zona Saudável" são menos do que seria de esperar no caso de haver independência. Já na escola Conde Vilalva há uma grande proximidade entre o que se observou e o que seria de esperar no caso de independência, o que é um forte indício de poder haver independência nesta escola. Aplicando o teste de Mantel-Haenszel (HOSMER & LEMESHOW, 2000), o valor $p < 0,001$ leva a rejeitar a hipótese de independência condicional à escola.

No caso do terceiro período, verifica-se para todas as escolas que, os alunos mais novos que estão "Dentro da Zona Saudável" são menos do que seria de esperar no caso de haver independência. Aplicando o teste de Mantel-Haenszel, o valor $p = 0,000$ leva a rejeitar a hipótese de independência condicional à escola.

Realizando um teste do Qui-Quadrado individualmente para cada escola Tabela 7, no primeiro período, podemos verificar que apenas a Escola de Castelo de Paiva apresenta um valor p significativo para a associação entre a "idade" e o pertencer ou não à Zona Saudável. Já no terceiro período, as três escolas apresentam um valor p significativo para a associação entre a "idade" e o pertencer ou não à Zona Saudável. O teste de Breslow-Day apresenta um valor p= 0,914, o que leva a assumir uma homogeneidade do Odds Ratio para todas as escolas no terceiro período.

	Escola				ZN_IMC		
					Fora da Zona Saudável	Dentro da zona Saudável	Total
1º P	Redondo	Idadecat2	8-14	Count	114	193	307
				Expected Count	106,5	200,5	307
			>=15	Count	39	95	134
				Expected Count	46,5	87,5	134
			Total	Count	153	288	441
				Expected Count	153	288	441
	Castelo de Paiva	Idadecat1	8-14	Count	191	553	744
				Expected Count	151,4	592,6	744
			>=15	Count	55	410	465
				Expected Count	94,6	370,4	465
			Total	Count	246	963	1209
				Expected Count	246	963	1209
	Conde Vilalva	Idadecat1	8-14	Count	198	259	457
				Expected Count	196,3	260,7	457
			>=15	Count	43	61	104
				Expected Count	44,7	59,3	104
			Total	Count	241	320	561
				Expected Count	241	320	561
3º P	Redondo	Idadecat1	8-14	Count	101	182	283
				Expected Count	88,9	194,1	283
			>=15	Count	31	106	137
				Expected Count	43,1	93,9	137
			Total	Count	132	288	420
				Expected Count	132	288	420
	Castelo de Paiva	Idadecat1	8-14	Count	109	491	600
				Expected Count	89,3	510,7	600
			>=15	Count	44	384	428
				Expected Count	63,7	364,3	428
			Total	Count	153	875	1028
				Expected Count	153	875	1028
	Conde Vilalva	Idadecat1	8-14	Count	165	229	394
				Expected Count	154,6	239,4	394
			>=15	Count	32	76	108
				Expected Count	42,4	65,6	108
			Total	Count	197	305	502
				Expected Count	197	305	502

Tabela 5: Classificação para Dentro e Fora da Zona Saudável segundo a idade

Chi-Square Tests					
	Escola		Estatística	Gl	Valor p (2-sided)
1º Período	**Redondo**	Pearson Chi-Square	2,65	1	O,103
	Castelo de Paiva	Pearson Chi-Square	33,839	1	<0,001
	Conde Vilalva	Pearson Chi-Square	0,136	1	0,713
3º Período	**Redondo**	Pearson Chi-Square	7,307	1	0,007
	Castelo de Paiva	Pearson Chi-Square	12,264	1	<0,001
	Conde Vilalva	Pearson Chi-Square	5,334	1	0,021

Tabela 6: Teste Qui-quadrado para cada escola

Com uma estimativa global para Odds Ratio (Mantel-Haenszel) de 1,859, para os alunos que têm idade até 14 anos a possibilidade de estar Fora da Zona Saudável ($IC95\%:1,5-2,4$) é o dobro em relação aos mais velhos.

3.3. Modelo de Regressão Logística Binomial para o IMC

Para estudar se existe interacção entre as variáveis consideradas no estudo, realizámos uma abordagem diferente ajustando um modelo de regressão logística com as variáveis explicativas Escola, Idade e Sexo e a variável dependente pertencer ou não à Zona Saudável. Na Tabela 8, apresentam-se os valores dos coeficientes do modelo de regressão logística ajustado, bem como o desvio padrão e o valor p associado a cada coeficiente, para o primeiro período. A variável Sexo não se mostrou significativa, mas a interacção Escola*Idade mostrou-se significativa, tal como concluímos anteriormente aquando da estratificação por escolas. Consequentemente, podemos concluir que o efeito Escola depende da Idade e que o efeito Idade depende da Escola.

Variáveis	Coeficientes (β)	Desvio-padrão	Valor p
Castelo de Paiva	0,537	0,145	<0,001
Conde Vilalva	-0,258	0,151	0,088
Idade >15	0,364	0,224	0,104
Castelo de Paiva* Idade >=15	0,582	0,279	0,037
Conde Vilalva*Idade >=15	-0,283	0,314	0,368
Constante	0,526	0,118	<0,001

Tabela 7: Modelo de Regressão Logística ajustado

A partir do modelo obtido estimaram-se os Odds Ratio dos alunos mais velhos relativamente aos mais novos para cada Escola e estimaram-se os odds ratio da Escola de Castelo de Paiva e da Escola Conde Vilalva relativamente à Escola do Redondo para as duas categorias de idade. Os valores obtidos e os respectivos intervalos de confiança são apresentados na Tabela 9. Os valores para o Odds Ratio foram calculados a partir do modelo:

$$\pi = \left(1+\exp\left(-\beta' X\right)\right)^{-1} =$$

$$= \left(1+\exp\left(-0,526-0,537X_1+0,258X_2-0,364X_3-0,582X_1X_3+0,283X_2X_3\right)\right)^{-1}$$

A título de exemplo ficam os cálculos efectuados para o caso do Odds Ratio da interacção da escola de Castelo de Paiva e idade superior ou igual a 15:

$$\text{logit}\left(escola=1, idade=1\right)-\text{logit}\left(escola=1, idade=0\right)=$$
$$=\left(\beta_0+\beta_1+\beta_3+\beta_4\right)-\left(\beta_0+\beta_1\right)=$$
$$=\beta_3+\beta_4=0,364+0,582=0,946$$

Logo, $OR=\exp\left(0,946\right)=2,575$. Após o cálculo da variância, o intervalo de confiança é dado por: $\exp\left(\hat{\beta}_0+\hat{\beta}_1\pm 1,96\sigma_{\hat{\beta}_3+\hat{\beta}_4}\right)=\left(1,57;4,21\right)$. De forma análoga, foram realizados os cálculos para o Odds Ratio considerando a escola=0 (Redondo), escola=2 (Conde Vilalva) e a idade inferior ou igual a catorze anos.

Da sua observação, podemos concluir que apenas para a Escola de Castelo de Paiva os valores são significativos a 5% (a unidade não pertence ao intervalo de confiança a 95 %). Neste caso podemos concluir que para aquela Escola a possibilidade dos alunos mais velhos pertencerem à zona saudável é 2,5 vezes superior relativamente aos alunos mais novos, podendo ir de 1,5 vezes a 4,2 vezes com uma confiança de 95%.

		Odds Ratio	IC (95%)
Redondo	Idade <=14 Idade >=15	1,44	(0,93;2,23)
Castelo de Paiva	Idade <=14 Idade >=15	2,575	(1,57;4,21)
Conde Vilalva	Idade <=14 Idade >=15	1,084	(0,7;1,67)
Idade <=14	Redondo Castelo de Paiva	0,288	(1,29;2,27)
	Redondo Conde Vilalva	0,773	(0,57;1,04)
Idade >=15	Redondo Castelo de Paiva	3,062	(1,92;4,88)
	Redondo Conde Vilalva	0,582	(0,34;1)

Tabela 8: Odds Ratio dos alunos mais novos em relação aos mais velhos por escola

Comparando as escolas, verifica-se que para os alunos mais novos apenas há diferenças significativas entre a Escola de Redondo e a Escola de Castelo de Paiva. Neste caso os alunos mais velhos da Escola de Castelo de Paiva têm 3 vezes mais possibilidade de pertencer à Zona Saudável relativamente aos mais novos, estando esta razão de possibilidades aproximadamente entre 2 a 5 vezes para um intervalo de confiança de 95%. Finalmente, para os alunos mais velhos, há diferenças significativas entre a Escola de Redondo e as restantes escolas. Os alunos da Escola de Castelo de Paiva têm 3 vezes mais possibilidade de pertencerem à Zona Saudável do que os alunos da Escola de Redondo, enquanto os alunos da Escola de Redondo têm 1,7 vezes mais possibilidade de pertencer à Zona Saudável do que os alunos da Escola Conde Vilalva. Isto é, os alunos da Escola Conde Vilalva situada na cidade de Évora, apresentam menos possibilidades de pertencerem à Zona Saudável do que os alunos da Escola de Redondo que moram numa vila e aldeias pertencentes ao concelho. Os alunos da escola da região norte têm mais possibilidade de pertencerem à Zona Saudável do que os alunos da escola da região do Alentejo.

Para o terceiro período, a variável Sexo continua sem ser significativa e a interacção Escola*Idade deixou de ser significativa, uma vez que se pode estabelecer uma homogeneidade para o Odds Ratio das escolas (teste de Breslow-Day).

3.4. Modelo de Regressão Logística Multinomial para o IMC

Para se tentar perceber melhor o que se passava em cada escola criou-se uma variável que categorizou a variável "ZN_IMC" de forma a dar-nos a proporção de alunos que se encontram acima, dentro ou abaixo da Zona Saudável para o IMC. Pela observação da Figura 20, verificamos que a maior percentagem de alunos dentro da Zona Saudável pertence à Escola de Castelo de Paiva, a Escola de Redondo tem a maior percentagem de alunos acima da Zona Saudável, e é na Escola Conde Vilalva que se encontra a maior percentagem de alunos abaixo da Zona Saudável.

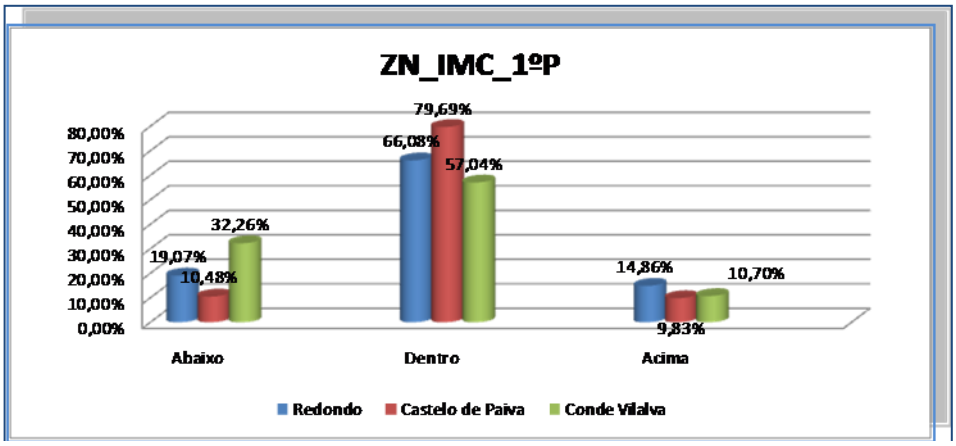

Figura 20: Distribuição dos alunos, por escola, abaixo, dentro ou acima da Zona Saudável, 1ºP

No terceiro período (Figura 21), verifica-se uma distribuição semelhante dos alunos em cada categoria.

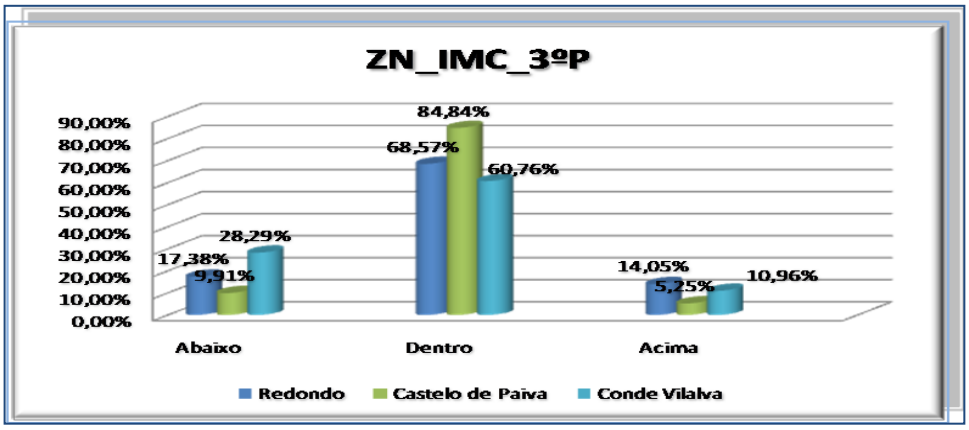

Figura 21: Distribuição dos alunos, por escola, abaixo, dentro ou acima da Zona Saudável, 3ºP

Procurámos de seguida analisar como esta nova variável é influenciada pelas covariáveis sexo, idade e Escola. Para tal ajustou-se um modelo de regressão logístico multinomial. Na Tabela 10, apresentam-se os valores dos coeficientes do modelo de regressão logística ajustado, bem como o desvio padrão e valor p associado a cada coeficiente.

ZN_IMC_int 1ºP	Variáveis	Coeficientes (β)	Desvio padrão	Valor p
Abaixo da ZN_IMC	Intercepto	-0,893	0,174	<0,001
	Sexo	-1,183	0,311	<0,001
	Conde Vilalva	-0,484	0,254	0,057
	Castelo de Paiva	-0,614	0,206	0,003
	Idade >=15	-0,833	0,212	<0,001
	Conde Vilalva *Sexo	0,727	0,412	0,078
	Castelo de Paiva*Sexo	0,102	0,364	0,779
	(Idade >=15)*Sexo	0,784	0,335	0,019
Acima da ZN_IMC	Intercepto	-1,364	0,194	<0,001
	Sexo	0,491	0,259	0,058
	Conde Vilalva	0,757	0,227	0,001
	Castelo de Paiva	-1,079	0,240	<0,001
	Idade >=15	0,094	0,190	0,622
	(Conde Vilalva) *Sexo	-0,294	0,312	0,346
	(Castelo de Paiva)*Sexo	0,573	0,318	0,071
	(Idade >=15)*Sexo	-1,434	0,300	<0,001

Tabela 9: Modelo de regressão logística multinomial com ZN_IMC categorizada, 1º P

Repare-se que abaixo da Zona Saudável temos as covariáveis sexo e idade e a sua interacção como significativas. Consequentemente estimámos os odds ratio entre as duas classes etárias fixando o sexo e estimámos os odds ratio entre os sexos fixando as classes etárias. Assim, podemos concluir que num aluno do sexo masculino a possibilidade de pertencer à zona abaixo reduz-se 56% no caso de o aluno ter idade superior a 15 anos comparativamente aos mais novos (IC95% 0,29-0,66). No caso dos

alunos do sexo feminino o odds ratio dos alunos mais velhos relativamente aos alunos mais novos não é significativo. Para os alunos mais novos, as possibilidades de pertencer à zona abaixo reduzem-se 60% no caso de ser do sexo feminino relativamente ao sexo masculino (IC95% 0,17-0,56). Para os alunos mais velhos o odds ratio do sexo feminino relativamente ao sexo masculino não é significativo. Acima da Zona Saudável, a Escola e a interacção entre sexo e idade são significativos. Assim, podemos concluir que um aluno do sexo masculino da Escola de Castelo de Paiva tem 2/3 menos de possibilidades de pertencer à zona acima do que um aluno do mesmo sexo da Escola de Redondo (IC95% 0,21-0,54). Já um aluno da Escola Conde de Vilalva tem aproximadamente o dobro das possibilidades de estar na zona acima do que um aluno da Escola de Redondo IC95% (1,37-3,33).

Podemos também concluir que num aluno do sexo feminino a possibilidade de pertencer à zona acima é 75% inferior se o aluno tiver idade superior a 15 anos comparativamente aos mais novos (IC95% 0,12-0,56). No caso dos alunos do sexo masculino não é significativa a diferença das idades. No caso dos alunos mais velhos a possibilidade de pertencer à zona acima reduz-se 60% no caso de ser do sexo feminino relativamente ao sexo masculino (IC95% 0,19-0,78). No caso dos alunos mais novos o odds ratio entre os dois sexos não é significativo.

Para o terceiro período (Tabela 11), também foi ajustado um modelo de regressão logístico multinomial, com as covariáveis sexo, idade e escola.

ZN_IMC_int 3ºP	Variáveis	Coeficientes (β)	Desvio padrão	Valor p
Abaixo da ZN_IMC	Intercepto	1,115	0,150	<0,000
	Sexo	0,166	0,128	0,196
	Redondo	-0,544	0,167	0,001
	Castelo de Paiva	0,734	0,169	0,000
	Idade >=15	0,582	0,149	0,000
Acima da ZN_IMC	Intercepto	-0,562	0,215	0,009
	Sexo	0,682	0,196	0,001
	Redondo	-0,721	0,238	0,002
	Castelo de Paiva	-0,428	0,244	0,08
	Idade >=15	-0,104	0,233	0,657

Tabela 10: Modelo de regressão logística multinomial com ZN_IMC categorizada, 3º P

No terceiro período, abaixo da Zona Saudável, a covariável sexo deixou de ser significativa. Apenas a idade se manteve como significativa. Acima da Zona Saudável, as covariáveis sexo e Escola são significativas.

Capítulo IV - Modelação do Desempenho Escolar

Neste capítulo começamos por analisar diferenças significativas no desempenho escolar nas três disciplinas, aplicando a MANOVA relativamente às Escolas e ao sexo. De seguida ajustam-se Modelos de Regressão Logística Binomial e Logística Ordinal para as disciplinas de Matemática e Língua Portuguesa. Por fim, aplicou-se o método de Imputação Múltipla para estimação dos dados omissos e ajustou-se um modelo para o desempenho escolar em Matemática comparando-se os resultados obtidos com a abordagem anteriormente realizada.

4.1. Análise de Variância Multivariada (MANOVA)

Na análise de variância multivariada (MANOVA), as variáveis dependentes consideradas simultaneamente foram as notas a Matemática, Língua Portuguesa e Educação Física. Iremos analisar diferenças significativas no desempenho escolar nas três disciplinas relativamente ao sexo e às escolas de Redondo e Castelo de Paiva. A escola Conde Vilalva não foi considerada por não ter ensino secundário.

Para validar os pressupostos da normalidade (Tabela 12) foi aplicado o teste de Kolmogorov-Smirnov com correcção de Lilliefors e Shapiro-Wilk e foram calculados os coeficientes de assimetria e curtose para as variáveis (Tabela 13).

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
LP_3	,103	201	,000	,969	201	,000
Mat_3	,112	201	,000	,972	201	,001
QdEdf	,152	201	,000	,966	201	,000

a. Lilliefors Significance Correction

Tabela 12: Teste à normalidade

Variáveis		Coeficiente	Desvio-padrão	
LP	Assimetria	-0,310	0,172	-1,8
	Curtose	-0,101	0,341	-0,3
Mat	Assimetria	0,321	0,172	1,87
	Curtose	-0,457	0,341	-1,34
Edf^2	Assimetria	-0,310	0,172	-1,8
	Curtose	-0,101	0,341	-0,3

Tabela 13: Coeficientes de assimetria e curtose

Como os quocientes pertencem ao intervalo $(-2;2)$, logo o afastamento à normalidade não é muito grande. No modelo foram utilizadas como variáveis dependentes, as variáveis LP, Mat e Edf^2, que validaram o pressuposto da multinormalidade. Na validação do pressuposto da homogeneidade das variâncias-covariâncias, foi utilizado o teste M de Box, que é muito sensível à violação do pressuposto da normalidade. Pelo teste M de Box (p-value=0,096>0,05), não rejeitamos a hipótese nula de que as covariâncias são homógeneas.

Uma vez que as condições de aplicação da Manova foram verificadas, podemos então testar as hipóteses relativas à Manova (Tabela 14). Neste caso estamos a testar a hipótese de igualdade entre os sexos, a hipótese de igualdade entre as escolas de Redondo e Castelo de Paiva, e de interacção entre os factores. Analisemos as conclusões obtidas com o traço de Pillai's e a maior raiz de Roy's para os factores sexo, escola e a sua interacção. Nos factores sexo e escola temos o valor p<0,001 e a potência do teste igual a um. A interacção entre os factores não tem um efeito estatisticamente significativo sobre os níveis atingidos pelos alunos nas disciplinas (Traço de Pillai=0,734 e Potência=0,134). O que quer dizer que os factores sexo e escola têm um efeito significativo.

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.	Noncent. Parameter	Observed Power[b]
Intercept	Pillai's Trace	,946	1127,814[a]	3,000	195,000	,000	3383,443	1,000
	Wilks' Lambda	,054	1127,814[a]	3,000	195,000	,000	3383,443	1,000
	Hotelling's Trace	17,351	1127,814[a]	3,000	195,000	,000	3383,443	1,000
	Roy's Largest Root	17,351	1127,814[a]	3,000	195,000	,000	3383,443	1,000
Sex_3	Pillai's Trace	,220	18,308[a]	3,000	195,000	,000	54,923	1,000
	Wilks' Lambda	,780	18,308[a]	3,000	195,000	,000	54,923	1,000
	Hotelling's Trace	,282	18,308[a]	3,000	195,000	,000	54,923	1,000
	Roy's Largest Root	,282	18,308[a]	3,000	195,000	,000	54,923	1,000
Escola_3	Pillai's Trace	,177	14,018[a]	3,000	195,000	,000	42,054	1,000
	Wilks' Lambda	,823	14,018[a]	3,000	195,000	,000	42,054	1,000
	Hotelling's Trace	,216	14,018[a]	3,000	195,000	,000	42,054	1,000
	Roy's Largest Root	,216	14,018[a]	3,000	195,000	,000	42,054	1,000
Sex_3 * Escola_3	Pillai's Trace	,007	,427[a]	3,000	195,000	,734	1,280	,134
	Wilks' Lambda	,993	,427[a]	3,000	195,000	,734	1,280	,134
	Hotelling's Trace	,007	,427[a]	3,000	195,000	,734	1,280	,134
	Roy's Largest Root	,007	,427[a]	3,000	195,000	,734	1,280	,134

a. Exact statistic

b. Computed using alpha = ,05

c. Design: Intercept + Sex_3 + Escola_3 + Sex_3 * Escola_3

Tabela 14: Testes à Manova

De seguida aplicou-se uma Anova univariada para cada uma das variáveis dependentes. Assim, analisando a Tabela 15 com os "*Tests for Between-Subjects Effects*", concluímos que a variável sexo possui um efeito significativo sobre os níveis das três disciplinas, enquanto a variável escola apenas tem um efeito significativo sobre as notas de Matemática e Educação Física. Nas escolas de Redondo e Castelo de Paiva, tanto para a média de Matemática como de Educação Física, os alunos de Redondo obtiveram uma média ligeiramente superior. No entanto não existe interacção significativa entre os factores sexo e escola para as variáveis em estudo.

Relativamente às médias das notas obtidas nas três disciplinas, em Língua Portuguesa e Matemática as raparigas têm médias mais elevadas que os rapazes em ambas as escolas. No caso da disciplina de Educação Física, são os rapazes que obtiveram médias mais elevadas em ambas as escolas. Portanto concluímos que o sexo e a escola são factores significativos na explicação das notas.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[b]
Corrected Model	LP_3	191,481[a]	3	63,827	10,270	,000	30,810	,998
	Mat_3	351,254[c]	3	117,085	11,104	,000	33,312	,999
	QdEdf	108817,143[d]	3	36272,381	7,865	,000	23,596	,989
Intercept	LP_3	19947,572	1	19947,572	3209,625	,000	3209,625	1,000
	Mat_3	17043,416	1	17043,416	1616,358	,000	1616,358	1,000
	QdEdf	7694167,056	1	7694167,056	1668,416	,000	1668,416	1,000
Sex_3	LP_3	114,456	1	114,456	18,416	,000	18,416	,990
	Mat_3	57,870	1	57,870	5,488	,020	5,488	,645
	QdEdf	43489,793	1	43489,793	9,430	,002	9,430	,863
Escola_3	LP_3	8,414	1	8,414	1,354	,246	1,354	,212
	Mat_3	230,740	1	230,740	21,883	,000	21,883	,996
	QdEdf	24073,864	1	24073,864	5,220	,023	5,220	,623
Sex_3 * Escola_3	LP_3	1,022	1	1,022	,164	,686	,164	,069
	Mat_3	11,535	1	11,535	1,094	,297	1,094	,180
	QdEdf	15,227	1	15,227	,003	,954	,003	,050
Error	LP_3	1224,340	197	6,215				
	Mat_3	2077,234	197	10,544				
	QdEdf	908496,827	197	4611,659				
Total	LP_3	35541,000	201					
	Mat_3	28110,000	201					
	QdEdf	1,508E7	201					
Corrected Total	LP_3	1415,821	200					
	Mat_3	2428,488	200					
	QdEdf	1017313,970	200					

a. R Squared = ,135 (Adjusted R Squared = ,122)

b. Computed using alpha = ,05

c. R Squared = ,145 (Adjusted R Squared = ,132)

d. R Squared = ,107 (Adjusted R Squared = ,093)

Tabela 15: Testes à significância multivariada.

4.2. Modelo de Regressão logística Binomial para o Desempenho a Matemática

Interessa agora verificar se os testes de *Fitnessgram*, sexo, idade, escola, IMC e o ciclo, influenciam as notas obtidas pelos alunos na disciplina de Matemática.

Posteriormente houve necessidade de categorizar a variável referente às notas de Matemática como positivas e negativas. Na amostra recolhida registámos 27,8% de negativas e 72,2% de positivas.

4.2.1. Análise Univariada

Numa primeira fase realizámos uma análise univariada para as notas do terceiro período na disciplina de Matemática. Na Tabela 16 apresentam-se os valores dos coeficientes de cada covariável, bem como o desvio padrão, o valor *p* associado a cada coeficiente, e IC(95%) para as notas de Matemática (análise individual).

Variáveis	Coeficientes (β)	Desvio-padrão	Valor p	Exp(β)	IC (95%)
Sexo	0,015	0,106	0,89	1,015	0,825-1,248
idade	β_0 3,669	0,387	<0,001	39,210	
	-0,196	0,027	<0,001	0,822	0,780-0,867
Idadecat2	β_0 1,152	0,066	<0,001	3,165	
	-0,552	0,111	<0,001	0,576	0,463-0,716
IMC	β_0 1,377	0,319	<0,001	3,961	
	-0,020	0,016	0,196	0,980	0,950-1,011
escola	β_0 1,275	0,121	<0,001	3,580	
Cpaiva	-0,473	0,140	0,001	0,623	0,473-0,821
Conde_vil	-0,206	0,159	0,194	0,814	0,596-1,111
ZN_IMC	β_0 1,059	0,106	<0,001	2,882	
	-0,118	0,123	0,335	0,889	0,699-1,130
Nota_lp	β_0 -0,383	0,127	0,002	0,682	
	1,645	0,141	<0,001	5,178	3,931-6,822
Nota_edf	β_0 -1,386	0,559	0,013	0,250	
	2,391	0,562	<0,001	10,921	3,633-32,828
Ciclo	β_0 1,588	0,1	<0,001	4,892	
7º,8º,9º	-0,972	0,122	<0,001	0,378	0,298-0,481
sec	-0,750	0,183	<0,001	0,473	0,330-0,677
abd	β_0 1,005	0,109	<0,001	2,733	
	-0,077	0,125	0,538	0,926	0,724-1,184
Flex_dir	β_0 0,978	0,066	<0,001	2,660	
	-0,060	0,114	0,601	0,942	0,753-1,178
Flex_esq	β_0 0,848	0,08	<0,001	2,325	
	0,198	0,109	0,069	1,219	0,985-1,508
vv	β_0 1,058	0,091	<0,001	2,882	
	-0,139	1,113	0,218	0,870	0,698-1,085

Tabela 16: Análise univariada para as notas de Matemática.

Individualmente, as variáveis que se mostram significativas são a idade e idade categorizada, a escola, as notas de Língua Portuguesa e Educação Física, o ciclo a que os alunos pertencem (5º-6º; 7º-9º; secundário) e a flexibilidade à esquerda (com $\alpha = 10\%$). Entre as variáveis significativas no modelo podemos concluir o seguinte:

- ✓ Por cada ano a mais a possibilidade de ter positiva a Matemática diminui 18%.
- ✓ Para os alunos com idade superior a 15 anos a possibilidade de ter positiva diminui 42%.
- ✓ Os alunos de Castelo de Paiva vêm a sua possibilidade de ter nível positivo a Matemática reduzida em cerca de 38% quando comparados com os alunos de Redondo.
- ✓ Para os alunos com positiva a Língua Portuguesa, a possibilidade de positiva a Matemática aumenta cerca de 5 vezes e com positiva a Educação Física aumenta até cerca de 11 vezes mais.
- ✓ Os alunos do ensino básico (7º, 8º e 9º anos) vêm a possibilidade de ter positiva a Matemática diminuída em 62% e os do secundário em cerca de 53% quando comparados com os de 2º ciclo.
- ✓ Os alunos que superam o valor médio de referência no teste da flexibilidade à esquerda têm 20% mais possibilidades de obter nível positivo a Matemática.

4.2.2. Análise Multivariada

Para estudar a interacção entre as variáveis consideradas no estudo foi ajustado um modelo de Regressão Logística, com as variáveis explicativas escola, idade categorizada, nota de Língua Portuguesa, a flexibilidade à esquerda e a nota de Educação Física, e com a variável dependente, nota a Matemática, seguindo o procedimento descrito no capítulo 2 referente à selecção de variáveis até chegar ao modelo final. Na Tabela 17 apresentam-se os valores dos coeficientes do modelo de regressão logística ajustado, bem como o desvio padrão e o valor *p* associado a cada coeficiente.

Pela observação dos coeficientes do modelo, temos que apenas a escola Conde Vilalva não é significativa para explicar a nota de Matemática.

	B	S.E.	Wald	df	Sig.	Exp(B)	90% C.I.for EXP(B)	
							Lower	Upper
Escola_3			18,605	2	<,001			
Escola_3(1)	-,682	,163	17,438	1	<,001	,505	,386	,661
Escola_3(2)	-,305	,190	2,587	1	,108	,737	,539	1,007
flex_esq3_dic(1)	,305	,135	5,130	1	,024	1,357	1,087	1,693
lp_3_A(1)	1,610	,153	110,235	1	<,001	5,004	3,889	6,440
edfnovacat(1)	,765	,183	17,507	1	<,001	2,150	1,591	2,904
idadecat(1)	-,994	,129	59,123	1	<,001	,370	,299	,458
Constant	,374	,208	3,230	1	,072	1,454		

Tabela 17: Modelo de regressão logística multivariada

4.2.3. Diagnóstico do Modelo

No modelo apresentado na Tabela 17, a variável idade aparece categorizada com um ponte de corte nos 15 anos. Esta categorização deve-se ao facto da variável idade não ser linear com o logit. De facto, quando se realizou o último passo da modelação antes de se averiguar se alguma interacção era significativa, o modelo obtido tinha um coeficiente para a variável idade não categorizada igual a -0,149. Para averiguarmos se a variável idade era linear com o logit procedeu-se da forma descrita no capítulo 2, tendo-se categorizado a idade em quatro categorias sendo a primeira constituída pelos valores da idade mínima ao 1º quartil, a segunda do 1º ao 2º quartil, a terceira do 2º ao 3º quartil e a última do 3º quartil à idade máxima. De seguida substituiu-se a variável idade por esta nova variável categorizada, tendo-se registado os coeficientes para as três categorias. Representaram-se de seguida o valor destes coeficientes contra os pontos médios dos intervalos que constituíram as categorias, tendo obtido a Figura 22. Como se pode observar não obtemos um comportamento linear registando uma inversão nos 15 anos. Dos três coeficientes apenas o último se mostrou significativo, o que não justifica incluir a variável categorizada desta forma. Consequentemente categorizou-se a variável

como binária com ponte de corte na idade 15 anos. A partir deste ponto decidimos usar esta nova variável.

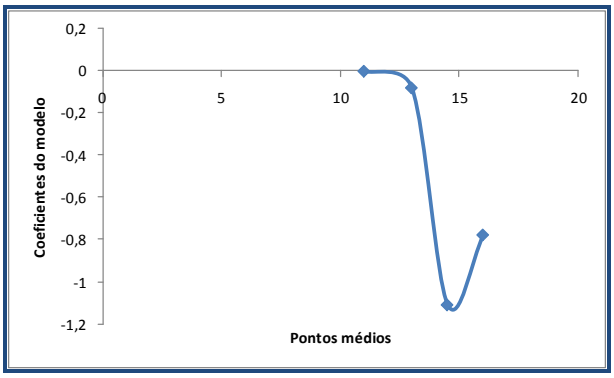

Figura 22: Gráfico de pontos médios e coeficientes das variáveis

No diagnóstico do modelo, começamos por realizar o teste de bondade de ajustamento de Hosmer-Lemeshow para o qual obtivemos o valor de estatística de teste χ^2=3,48 o que para 8 graus de liberdade corresponde a um valor *p* de 0,90. O que significa que aceitamos a hipótese de um bom ajustamento. A título de exemplo, observe-se a Tabela 18 onde figuram os valores estimados e observados em cada um dos 10 grupos constituídos no teste.

Contingency Table for Hosmer and Lemeshow Test

		mat_3 = 0		mat_3 = 1		
		Observed	Expected	Observed	Expected	Total
Step 1	1	95	97,593	47	44,407	142
	2	93	94,119	116	114,881	209
	3	70	63,506	110	116,494	180
	4	49	46,660	101	103,340	150
	5	40	41,161	121	119,839	161
	6	13	16,024	66	62,976	79
	7	41	44,405	225	221,595	266
	8	26	21,981	130	134,019	156
	9	16	17,258	142	140,742	158
	10	9	9,293	114	113,707	123

Tabela 18:Tabela de contingência para Hosmer & Lemeshow

O coeficiente de determinação, R^2 de Nagelkerke apresenta um valor de 0,198.

Na Figura 23 representa-se o valor da distância de Cook para os diferentes indivíduos. Da sua observação parecem destacar-se os indivíduos 19 e 30, embora o valor

correspondente para a distância de Cook seja inferior aos valores usuais para serem consideradas observações influentes (0,5 ou 1, consoante os autores).

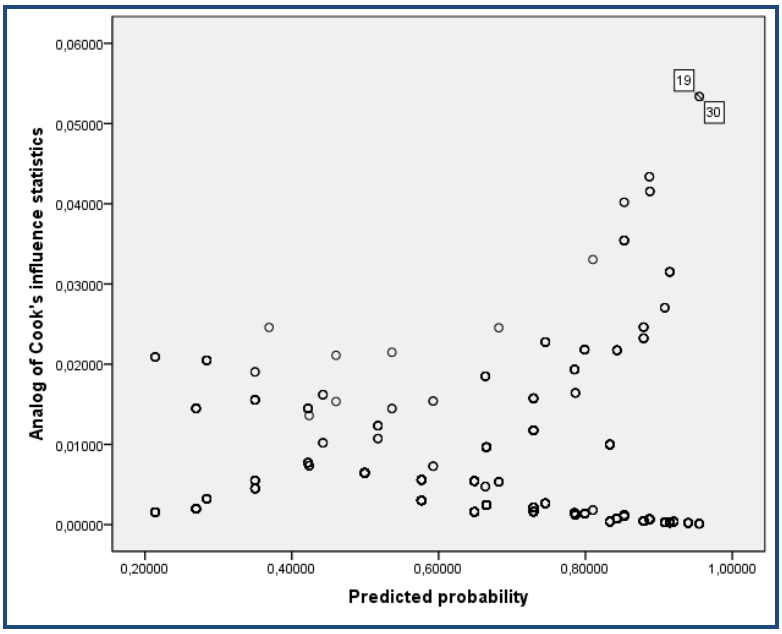

Figura 13: Distância de Cook

Na Figura 24 podem ser observados os resíduos Deviance onde são também assinalados os dois indivíduos referidos anteriormente. Embora haja outros indivíduos com um resíduo inferior a -2, são os que foram referidos anteriormente que se destacam por estarem associados a um valor maior de resíduo.

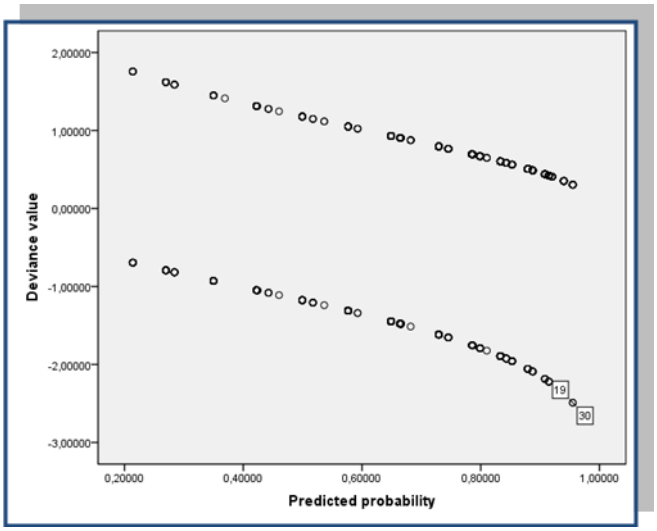

Figura 24: Resíduos Deviance

No modelo sem as observações 19 e 30 o valor da Deviance é idêntico ao obtido ao modelo anterior, pelo que estas observações não se afiguram ser *outliers*. Na Figura 25 representam-se alguns dos gráficos de resíduos *DfBetas* podendo verificar-se que não há nenhuma observação a destacar na estimação de cada um dos coeficientes de regressão.

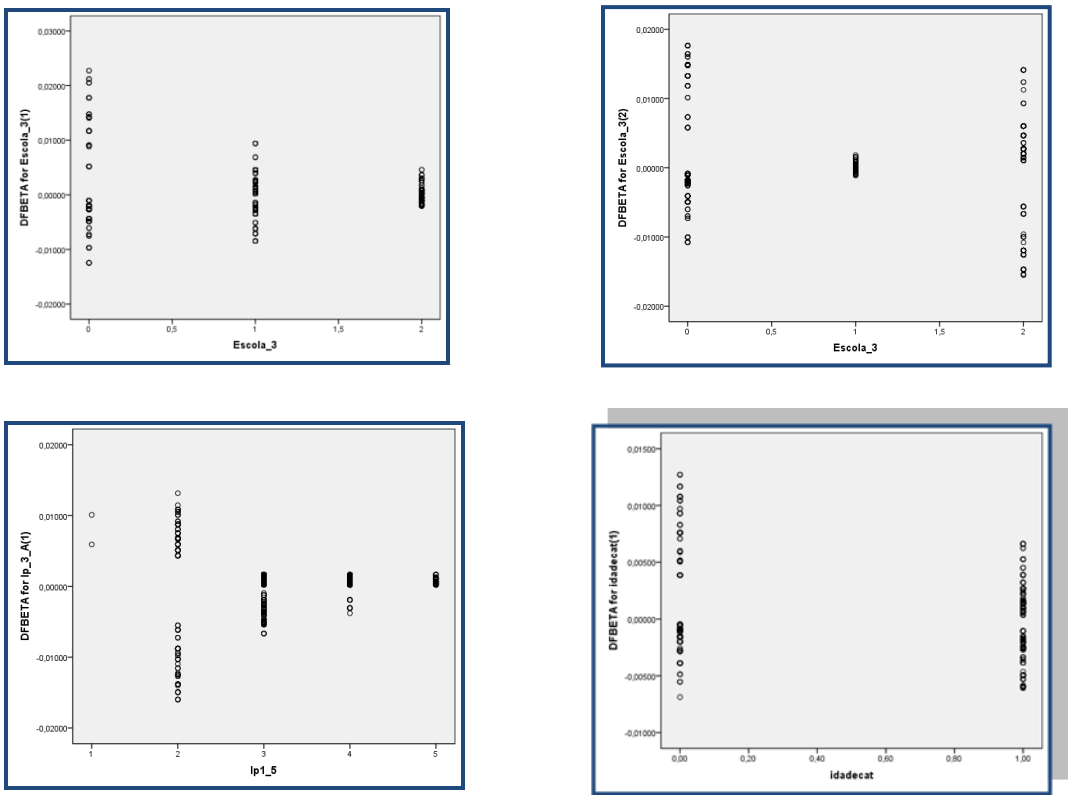

Figura 25: Resíduos DfBetas

A partir dos resíduos foi estimado o modelo sem os indivíduos19 e 30, com o objectivo de avaliar o comportamento do modelo. Na Tabela 19 apresentam-se os coeficientes das variáveis explicativas para o modelo com as observações 19 e 30, e o modelo sem essas observações. Uma vez que variação não ultrapassa os 14% então não se consideram observações influentes e vamos mantê-las no modelo.

Com o intuito de avaliar a capacidade do modelo para discriminar os sujeitos com a característica de interesse *vs* sujeitos sem a característica de interesse, recorreu-se à Curva ROC. Na Figura 26 apresentamos a Curva ROC cuja área é 0,734, ou seja a discriminação feita pelo modelo é aceitável.

	β com	β sem	Δβ
Escola_3			
Escola_3(1)	-,682	-,728	1%
Escola_3(2)	-,305	-,341	-14%
flex_esq3_dic(1)	,305	,321	13%
lp_3_A(1)	1,610	1,621	1%
edfnovacat(1)	,765	,823	-10%
idadecat(1)	-,994	-1,014	0%
Constant	,374	,404	-5%

Tabela 19: Variação dos coeficientes com e sem as observações 19 e 30

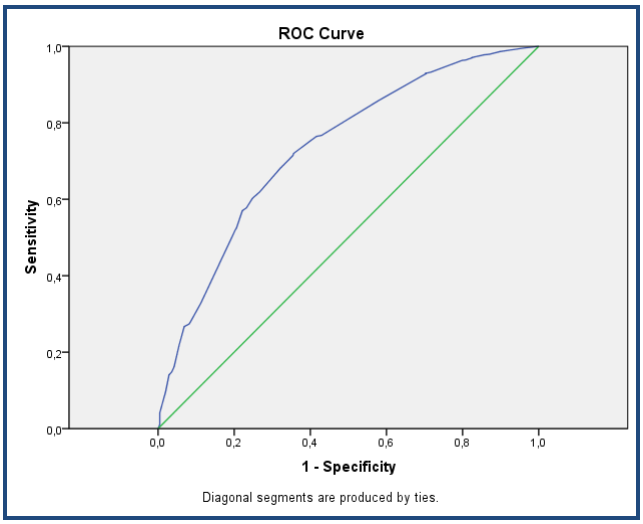

Figura 26: Curva ROC para as notas de Matemática

Na Tabela 20 apresentam-se os valores obtidos para a Sensibilidade e Especificidade do modelo. A Sensibilidade e Especificidade do modelo permitem avaliar se o modelo é eficiente na classificação dos indivíduos. Neste caso, para o ponto de corte que torna máxima a sua soma (0,656) o modelo apresenta uma Sensibilidade de 57,1% e uma Especificidade de 76,7%, com uma percentagem total de acertos de 71,2%, o que é razoável.

Classification Tablea

	Observed		Predicted		
			mat_3		Percentage Correct
			0	1	
Step 1	mat_3	0	258	194	57,1
		1	273	899	76,7
	Overall Percentage				71,2

Tabela 20: Sensibilidade e Especificidade do modelo

4.2.4. Interpretação dos Coeficientes

Na Tabela 17, apresentámos o modelo de regressão logística multinomial ajustado para as notas de Matemática. Após o diagnóstico ao mesmo e a verificação da bondade de ajustamento, podemos agora interpretar os coeficientes como sendo os do modelo final.

- ✓ Os alunos de Castelo de Paiva vêm a sua possibilidade de ter nível positivo a Matemática reduzida em cerca de 50% quando comparados com os alunos de Redondo.
- ✓ Para os alunos com idade superior a 15 anos a possibilidade de ter positiva diminui 2/3.
- ✓ Os alunos que superam o valor médio de referência no teste da flexibilidade à esquerda têm 40% mais possibilidades de obter nível positivo a Matemática.
- ✓ Para os alunos com positiva a Língua Portuguesa, a possibilidade de positiva a Matemática aumenta cerca de 5 vezes.
- ✓ Para os alunos com nível 5 a Educação Física, a possibilidade de positiva a Matemática aumenta aproximadamente para o dobro.

4.3. Modelo de Regressão logística Binomial para o Desempenho a Língua Portuguesa

Interessa agora verificar se os testes de *Fitnessgram*, sexo, idade, escola, IMC e o ciclo, influenciam as notas obtidas pelos alunos na disciplina de Língua Portuguesa.

Posteriormente houve necessidade de categorizar a variável referente às notas de Língua Portuguesa, como positivas e negativas. Na amostra recolhida registámos 14,8% de negativas e 85,2% de positivas.

4.3.1. Análise Univariada

À semelhança do que foi feito para as notas da disciplina de Matemática, realizámos uma análise univariada para as notas do terceiro período na disciplina de Língua Portuguesa. Na Tabela 21 apresentam-se os valores dos coeficientes de cada covariável, bem como o desvio padrão, o valor *p* associado a cada coeficiente, e IC(95%) para as notas de Língua Portuguesa.

Variável	Coeficientes (β)	Desvio-padrão	Valor p	Exp(β)	IC(95%)
Sexo	0,540	0,138	<0,001	1,715	1,31 – 2,246
idade	β_0 2,754	0,465	<0,001	15,709	
	-0,071	0,033	0,034	0,932	0,873-0,995
IMC	β_0 2,332	0,404	<0,001	0,932	
	-0,027	0,020	0,167	0,973	0,936-1,012
Escola	β_0 1,786	1,142	<0,001	5,966	
Cpaiva	-0,142	0,168	0,399	0,868	0,624-1,206
Conde_vil	0,310	0,202	0,126	1,363	0,917-2,026
ZN_IMC	β_0 0,1883	0,137	<0,001	6,574	
	-0,131	0,158	0,406	0,877	0,644-1,195
Nota_mat	β_0 0,795	0,097	<0,001	2,214	
	1,645	0,141	<0,001	5,178	3,931-6,822
Nota_edf	β_0 -0,847	0,488	0,082	0,429	
	2,685	0,493	<0,001	14,657	5,58-38,503
Ciclo	β_0 2,131	0,122	<0,001	8,427	
7º,8º,9º	-0,695	0,149	<0,001	0,499	0,373-0,668
sec	0,636	0,322	0,048	1,889	1,005-3,548
Idadecat2	β_0 1,846	0,082	<0,001	6,337	
	-0,190	0,143	0,184	0,827	0,625-1,094
abd	β_0 1,949	0,145	<0,001	7,019	
	-0,226	0,166	0,172	0,798	0,577-1,103
Flex_dir	β_0 1,810	0,121	<0,001	6,112	
	-0,049	0,147	0,740	0,953	0,715-1,270
Flex_esq	β_0 1,898	0,112	<0,001	6,674	
	-0,199	0,142	0,160	0,820	0,621-1,082
vv	β_0 1,835	0,116	<0,001	6,267	
	-0,121	0,143	0,396	0,886	0,670-1,172

Tabela 21: Análise univariada para as notas de Língua Portuguesa

Individualmente, as variáveis que se mostram significativas são o sexo, a idade, as notas de Matemática e Educação Física e o ciclo a que pertencem (5º-6º; 7º-9º; secundário).

Entre as variáveis significativas no modelo podemos concluir o seguinte:

- ✓ Por cada ano a mais a possibilidade de ter positiva a Língua Portuguesa diminui 7%.
- ✓ Para os alunos com idade superior a 15 anos a possibilidade de ter positiva diminui 18%.
- ✓ As raparigas têm 1,7 vezes mais possibilidades de obter positiva a Língua Portuguesa que os rapazes.
- ✓ Para os alunos com positiva a Matemática, a possibilidade de positiva a Língua Portuguesa aumenta cerca de 5 vezes e com positiva a Educação Física aumenta até cerca de 14 vezes mais.
- ✓ Os alunos do ensino básico (7º, 8º e 9º anos) vêm a possibilidade de ter positiva a Língua Portuguesa diminuída em 50% enquanto para os do secundário aumenta quase duas vezes quando comparados com os de 2º ciclo.
- ✓ Os alunos que superam o valor médio de referência no teste da flexibilidade à esquerda têm 20% mais possibilidades de obter nível positivo a Língua Portuguesa.

4.3.2. Análise Multivariada

Para estudar a associação entre as variáveis consideradas no estudo foi ajustado um modelo de Regressão Logística, com as variáveis explicativas escola, idade categorizada, nota de Língua Portuguesa, a flexibilidade à esquerda e a nota de Educação Física, e com a variável dependente, nota a Língua Portuguesa, seguindo o procedimento descrito no capítulo 2 referente à selecção de variáveis até chegar ao modelo final. Na Tabela 22 apresentam-se os valores dos coeficientes do modelo de regressão logística ajustado, bem como o desvio padrão e o valor *p* associado a cada coeficiente.

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Sex	0,636	0,153	17,185	1	<0,001	1,889	1,398	2,551
Mat	1,582	0,153	106,736	1	<0,001	4,866	3,604	6,570
Ciclo (7º,8º,9º)	-0,403	0,171	5,545	1	0,019	0,668	0,478	0,935
Ciclo (sec.)	0,907	0,339	7,138	1	0,008	2,477	1,273	4,818
Flex_esq.	-0,344	0,158	4,716	1	0,030	0,709	0,520	0,967
Edf	0,815	0,255	10,203	1	0,001	2,258	1,370	3,722
constante	0,645	0,256	6,373	1	0,012	1,907		

Tabela 22: Modelo de regressão logística multivariada

4.3.3. Diagnóstico do Modelo

No diagnóstico do modelo, começamos por realizar o teste de bondade de ajustamento de Hosmer-Lemeshow para o qual obtivemos o valor de estatística de teste χ^2 =8,308 o que para 8 graus de liberdade corresponde a um valor *p* de 0,404. O que significa que aceitamos a hipótese de um bom ajustamento.

O coeficiente de determinação, R^2 de Nagelkerke apresenta um valor de 0,188.

Na Figura 27 representa-se o valor da distância de Cook para os diferentes indivíduos, não existindo observações influentes a considerar.

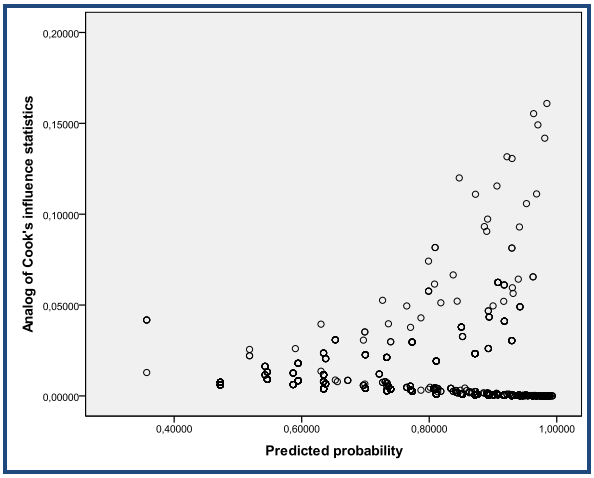

Figura 27: Distância de Cook

Na Figura 28 observam-se os resíduos Deviance e tal como na distância de Cook, não se destaca nenhuma observação.

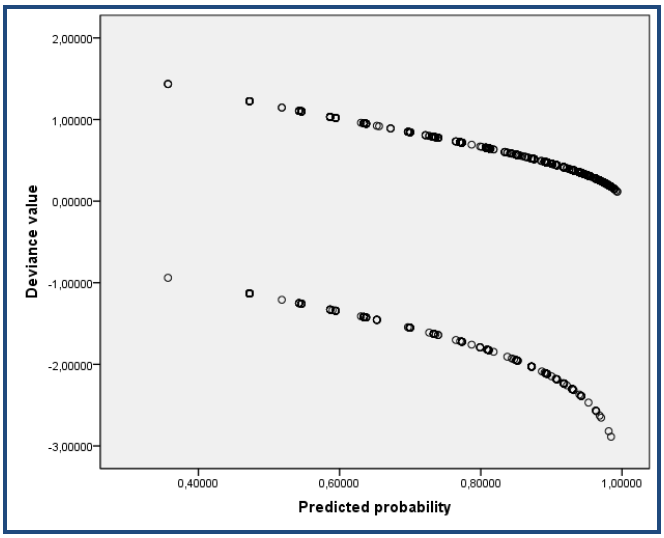

Figura 28: Resíduos Deviance

Pelos gráficos de resíduos *DfBetas*, alguns dos quais representados na Figura 29, verifica-se também que não há nenhuma observação a destacar na estimação de cada um dos coeficientes de regressão.

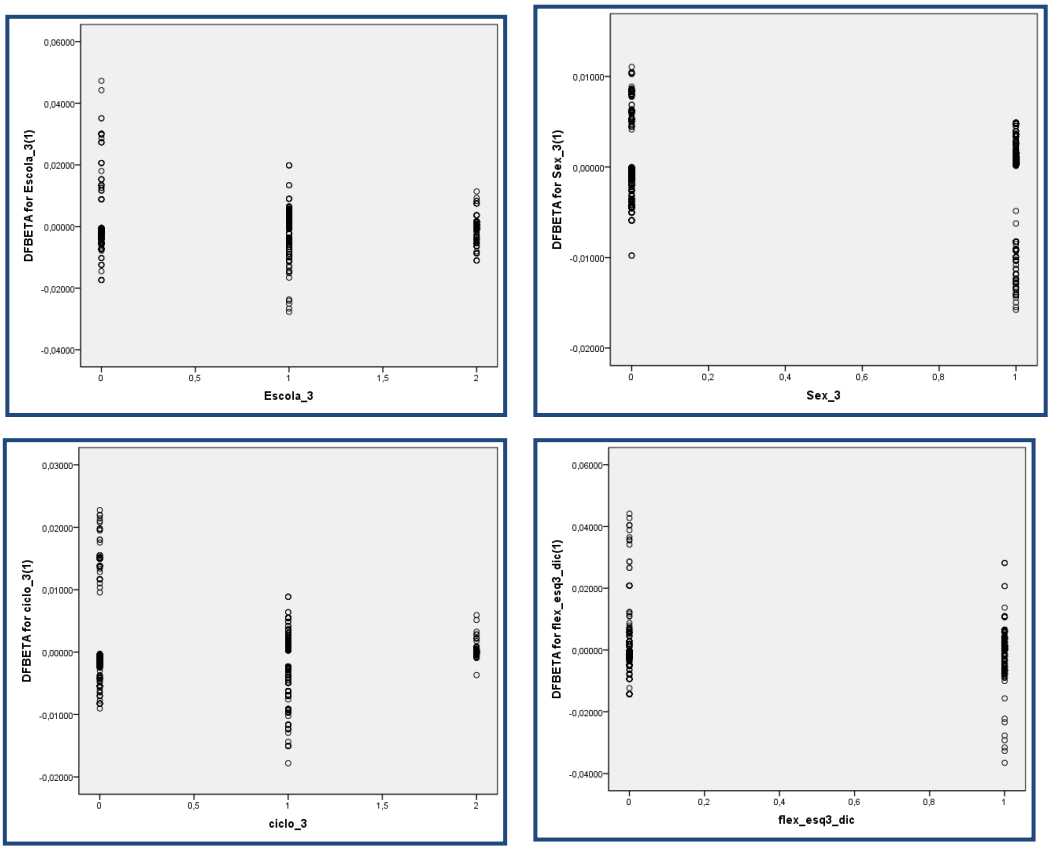

Figura 29: Resíduos DfBetas

Com o intuito de avaliar a capacidade do modelo para discriminar os sujeitos com a característica de interesse *vs* sujeitos sem a característica de interesse, recorreu-se à Curva ROC. Na Figura 30 apresentamos a Curva ROC cuja área é 0,753 (IC95%: 0,719; 0786), ou seja a discriminação feita pelo modelo é aceitável.

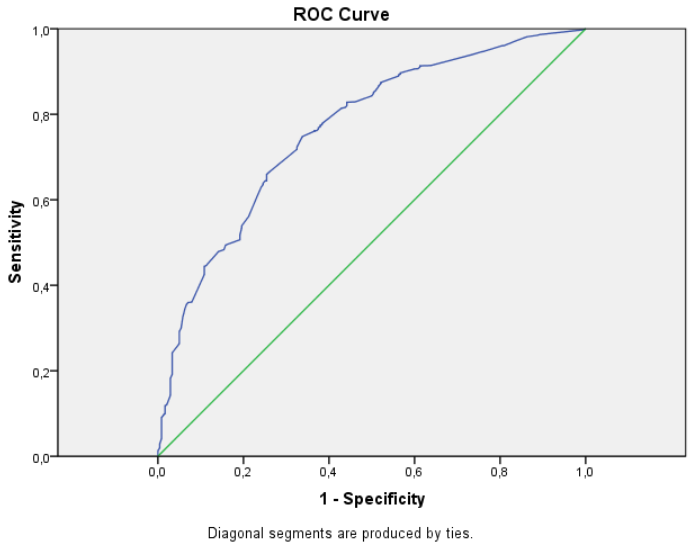

Figura 30: Curva ROC para as notas de Língua Portuguesa

A Sensibilidade e Especificidade do modelo permitem avaliar se o modelo é eficiente na classificação dos indivíduos. E neste caso o modelo apresenta uma Sensibilidade de 76,1% e uma Especificidade de 63,3%.

4.3.4. Interpretação dos Coeficientes

Na Tabela 22, apresentámos o modelo de regressão logística multinomial ajustado para as notas de Língua Portuguesa. Após o diagnóstico ao mesmo e a verificação da bondade de ajustamento, podemos agora interpretar os coeficientes como sendo os do modelo final.

- ✓ As raparigas têm 1,9 vezes a mais de possibilidade de ter nota positiva a Língua Portuguesa.
- ✓ Os alunos que superam o valor médio de referência no teste da flexibilidade à esquerda têm 29% menos possibilidades de obter nível positivo a Língua

Portuguesa. O que não deixa de ser curioso, pois no caso da Matemática a possibilidade era superior.

- ✓ Para os alunos com positiva a Matemática, a possibilidade de positiva a Língua Portuguesa aumenta cerca de 5 vezes.
- ✓ Para os alunos com nível 5 a Educação Física, a possibilidade de positiva a Língua Portuguesa aumenta um pouco mais que o dobro.
- ✓ Os alunos do ensino básico (7º, 8º e 9º anos) vêm a possibilidade de ter positiva a Língua Portuguesa diminuída em 34% enquanto para os do secundário aumenta duas vezes e meia quando comparados com os de 2º ciclo.

4.4. Modelo de Regressão Ordinal para o Desempenho a Matemática

As notas das disciplinas envolvidas estão ordenadas por uma escala, no 2º e 3º ciclos a escala vai de 2 a 5 e no secundário de 0 a 20. E como os modelos de regressão logística ordinal permitem realizar uma análise de dados cuja variável resposta é apresentada em categorias com ordenação, foi ajustado um modelo de regressão ordinal para as notas de Matemática e Língua Portuguesa. Como o número de notas negativas em Educação Física era mínimo, não fazia sentido modelar uma regressão logística ordinal.

Na Tabela 23 apresentamos o valor *p* para o teste de Pearson e para a Deviance. O modelo ajustado é significativamente melhor que o modelo nulo (p=0,000). Em ambos os testes não se rejeita a hipótese nula (o modelo ajusta-se aos dados).

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	749,063	820	,963
Deviance	629,973	820	1,000

Link function: Logit.

Tabela 23: Teste de Pearson e Deviance

O coeficiente de determinação, R^2 de Nagelkerke apresenta um valor de 0,506.

Na Tabela 24 podemos observar os valores dos coeficientes do modelo de Regressão Ordinal ajustado, bem como o desvio padrão e o valor *p* associado a cada coeficiente:

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[matordinal = ,00]	-5,754	,338	289,907	1	,000	-6,416	-5,092
	[matordinal = 1,00]	-2,650	,322	67,627	1	,000	-3,282	-2,019
	[matordinal = 2,00]	-,267	,300	,787	1	,375	-,855	,322
Location	[Sex_3=0]	,304	,103	8,644	1	,003	,101	,506
	[Sex_3=1]	0[a]	.	.	0	.	.	.
	[adb3_dic=0]	-,225	,128	3,061	1	,080	-,476	,027
	[adb3_dic=1]	0[a]	.	.	0	.	.	.
	[Escola_3=0]	,297	,155	3,653	1	,056	-,008	,601
	[Escola_3=1]	-,440	,152	8,364	1	,004	-,739	-,142
	[Escola_3=2]	0[a]	.	.	0	.	.	.
	[idadecat=,00]	,981	,103	89,980	1	,000	,779	1,184
	[idadecat=1,00]	0[a]	.	.	0	.	.	.
	[ZN_IMC_int_3=0]	,312	,144	4,683	1	,030	,029	,594
	[ZN_IMC_int_3=1]	,367	,215	2,919	1	,088	-,054	,789
	[ZN_IMC_int_3=2]	0[a]	.	.	0	.	.	.
	[lpordinal=,00]	-6,336	,317	398,908	1	,000	-6,958	-5,715
	[lpordinal=1,00]	-5,102	,284	321,898	1	,000	-5,660	-4,545
	[lpordinal=2,00]	-2,370	,267	79,050	1	,000	-2,893	-1,848
	[lpordinal=3,00]	0[a]	.	.	0	.	.	.
	[edfnovacat=,00]	-,533	,137	15,161	1	,000	-,801	-,265
	[edfnovacat=1,00]	0[a]	.	.	0	.	.	.

Link function: Logit.

a. This parameter is set to zero because it is redundant.

Tabela 24: Modelo de Regressão Ordinal para as notas de Matemática

Relativamente à interpretação dos coeficientes deste modelo, verifica-se uma relação entre as notas de Matemática e as notas de Língua Portuguesa. Da observação da Tabela 24, evidenciam-se as seguintes considerações:

- ✓ Nos rapazes relativamente às raparigas a possibilidade de terem uma pior nota a Matemática diminui 25%.
- ✓ A possibilidade de ter pior nota aumenta 25% quando o valor do teste dos abdominais não supera a média.
- ✓ Para os alunos da escola de Redondo, a possibilidade de terem pior nota relativamente aos alunos da escola Conde Vilalva diminui 25%.
- ✓ Para os alunos de Castelo de Paiva, a possibilidade de terem pior nota relativamente aos alunos da escola Conde Vilalva aumenta 55%.
- ✓ Nos alunos mais novos, relativamente aos mais velhos, a possibilidade de terem pior nota diminui para 2/3.
- ✓ Para os alunos com um índice de IMC dentro ou acima da Zona Saudável, a possibilidade de tirarem piores notas diminui cerca de 30%.
- ✓ Quando piora a nota de Língua Portuguesa, maior é a possibilidade de ter pior nota em Matemática. Sendo que, relativamente aos alunos de 5, os alunos de 4

têm uma possibilidade que aumenta cerca de dez vezes, os alunos de 3 aumenta 164 vezes, e os alunos de 2 aumenta 565 vezes.

- ✓ Nos alunos com nota inferior a 5 em Educação Física, a possibilidade de tirar pior nota em Matemática aumenta 70%.

Do teste à homogeneidade dos declives obtém-se um qui-quadrado igual a 28,654, o que para 22 gruas de liberdade resulta um valor p igual a 0,155 que permite validar o pressuposto dos riscos serem proporcionais ao longo das categorias da variável resposta.

Nos gráficos de resíduos score (Figura 31), podemos observar que a tendência em torno das categorias da variável resposta tem um comportamento horizontal próximo de zero para cada covariável. Na Figura 32 podemos observar os gráficos de resíduos parciais, que validam o pressuposto dos riscos proporcionais, devido ao seu aspecto linear e à proximidade de paralelismo entre as rectas.

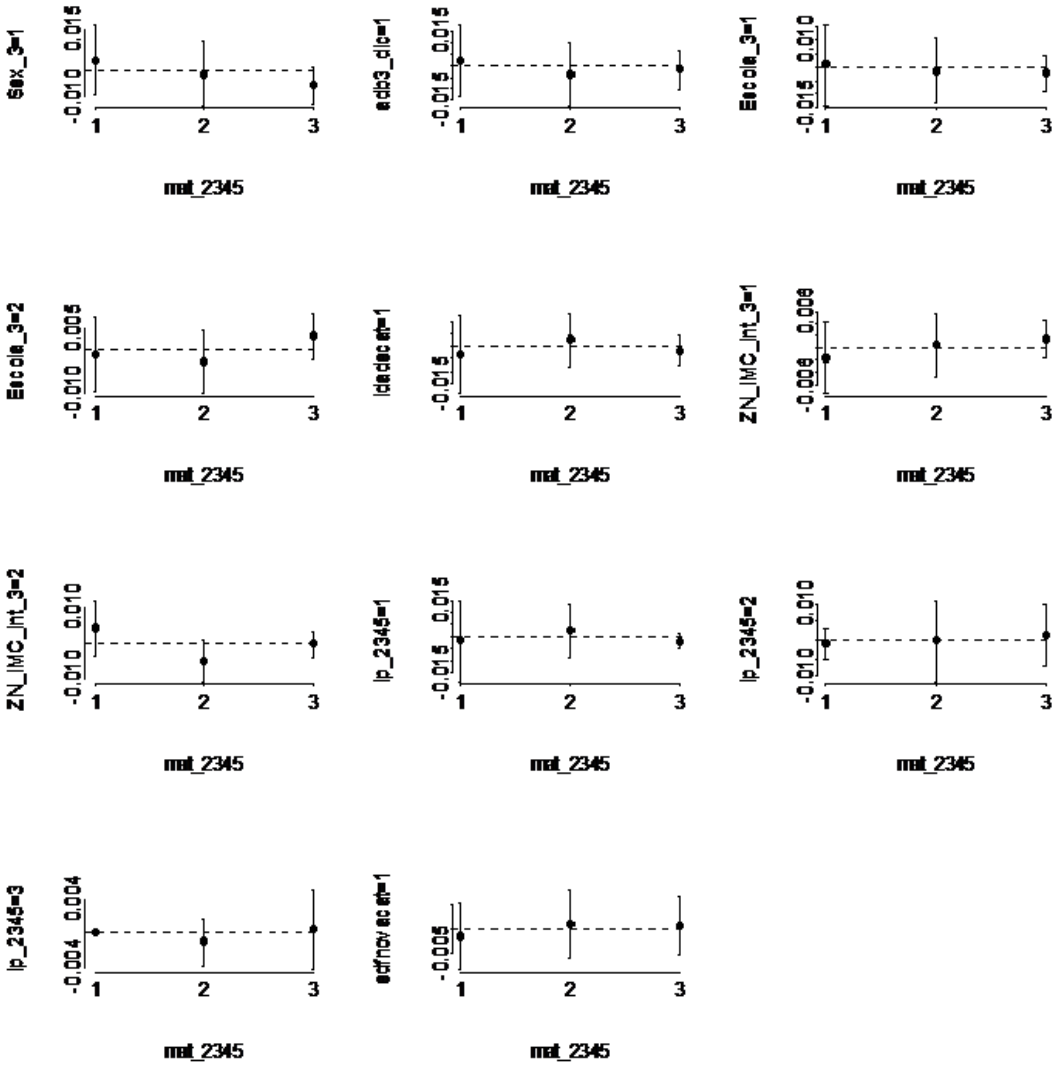

Figura 31: Gráficos de resíduos score para as covariáveis incluídas no modelo tendo como resposta as notas de Matemática

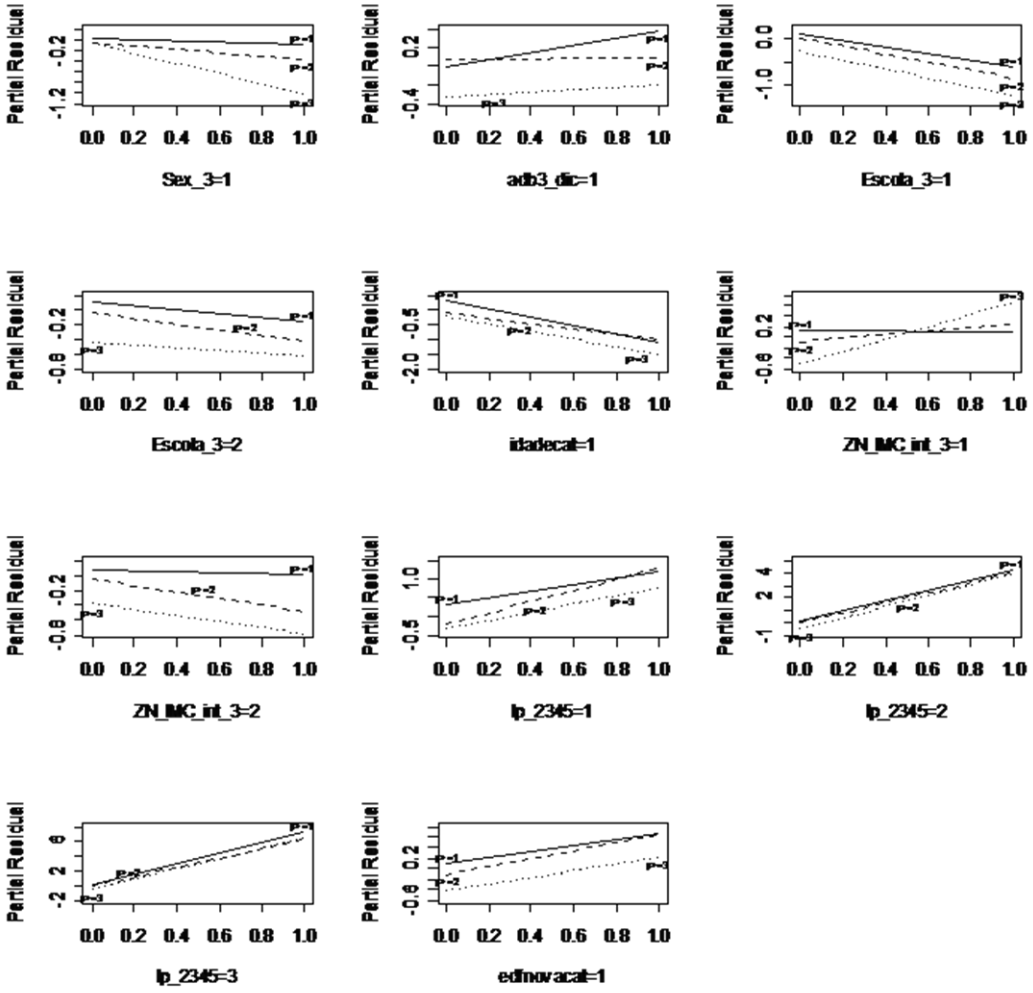

Figura 32: Gráficos de resíduos parciais para as covariáveis incluídas no modelo tendo como resposta as notas de Matemática

4.5. Modelo de Regressão Ordinal para o Desempenho a Língua Portuguesa

Tal como no caso das notas de Matemática, para as notas de Língua Portuguesa foi ajustado um Modelo de Regressão Ordinal.

Na Tabela 25 apresentamos o valor *p* para o teste de Pearson e para a Deviance. Em ambos os testes não se rejeita a hipótese nula (o modelo ajusta-se aos dados).

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	532,708	590	,956
Deviance	436,567	590	1,000

Link function: Logit.

Tabela 25:Teste de Pearson e Deviance

O coeficiente de determinação, R^2 de Nagelkerke apresenta um valor de 0,511. Curiosamente é um valor próximo do R^2 de Nagelkerke para o caso do modelo das notas de Matemática.

Na Tabela 26, podemos observar os valores dos coeficientes do modelo de Regressão Ordinal ajustado, bem como o desvio padrão e o valor *p* associado a cada coeficiente:

Da observação da Tabela 26, evidenciam-se as seguintes considerações:

- ✓ Nos rapazes relativamente às raparigas a possibilidade de terem uma pior nota a Língua Portuguesa aumenta aproximadamente para o dobro.
- ✓ A possibilidade de ter pior nota diminui 1/3 quando o valor do teste do vai-vém não supera a média.
- ✓ Para os alunos da escola de Redondo, a possibilidade de terem pior nota relativamente aos alunos da escola Conde Vilalva aumenta 85%. Para os alunos de Castelo de Paiva não há diferenças significativas.
- ✓ Relativamente aos alunos do secundário, nos alunos do 2º ciclo a possibilidade de tirarem piores notas aumenta mais 2 vezes. Enquanto que nos alunos do 3º ciclo essa possibilidade aumenta quase 3 vezes.
- ✓ Quando piora a nota de Matemática, maior é a possibilidade de ter pior nota em Língua Portuguesa. Sendo que, relativamente aos alunos de 5, os alunos de 4

têm uma possibilidade que aumenta cerca de dez vezes, os alunos de 3 aumenta 108 vezes, e os alunos de 2 aumenta 435 vezes.

- ✓ Nos alunos com nota inferior a 5 em Educação Física, a possibilidade de tirar pior nota em Língua Portuguesa aumenta quase para o dobro.

Parameter Estimates

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[lpordinal = ,00]	-8,695	,407	456,939	1	,000	-9,493	-7,898
	[lpordinal = 1,00]	-4,807	,373	166,006	1	,000	-5,539	-4,076
	[lpordinal = 2,00]	-1,364	,320	18,127	1	,000	-1,992	-,736
Location	[Sex_3=0]	-,750	,111	45,374	1	,000	-,968	-,532
	[Sex_3=1]	0[a]	.	.	0	.	.	.
	[Escola_3=0]	-,618	,162	14,508	1	,000	-,936	-,300
	[Escola_3=1]	-,232	,167	1,937	1	,164	-,559	,095
	[Escola_3=2]	0[a]	.	.	0	.	.	.
	[edfnovacat=,00]	-,622	,145	18,336	1	,000	-,907	-,337
	[edfnovacat=1,00]	0[a]	.	.	0	.	.	.
	[ciclo_3=0]	-,853	,183	21,725	1	,000	-1,212	-,494
	[ciclo_3=1]	-1,011	,178	32,291	1	,000	-1,360	-,662
	[ciclo_3=2]	0[a]	.	.	0	.	.	.
	[w3_dic=0]	,397	,135	8,603	1	,003	,132	,662
	[w3_dic=1]	0[a]	.	.	0	.	.	.
	[matordinal=,00]	-6,076	,288	444,282	1	,000	-6,641	-5,511
	[matordinal=1,00]	-4,685	,264	314,880	1	,000	-5,202	-4,167
	[matordinal=2,00]	-2,295	,251	83,664	1	,000	-2,787	-1,804
	[matordinal=3,00]	0[a]	.	.	0	.	.	.

Link function: Logit.

a. This parameter is set to zero because it is redundant.

Tabela 26: Modelo de Regressão Ordinal para as notas de Língua Portuguesa

Do teste à homogeneidade dos declives obtém-se um qui-quadrado igual a 32,394, o que para 20 gruas de liberdade resulta um valor p igual a 0,039 que permite validar o pressuposto dos riscos serem proporcionais ao longo das categorias da variável resposta ao nível de 1%, mas não ao nível usual de 5%.

Nos gráficos de resíduos score (Figura 33), podemos observar que a tendência em torno das categorias da variável resposta tem um comportamento horizontal próximo de zero para cada covariável. Na Figura 34 podemos observar os gráficos de resíduos parciais, que validam o pressuposto dos riscos proporcionais, devido ao seu aspecto linear e à proximidade de paralelismo entre as rectas.

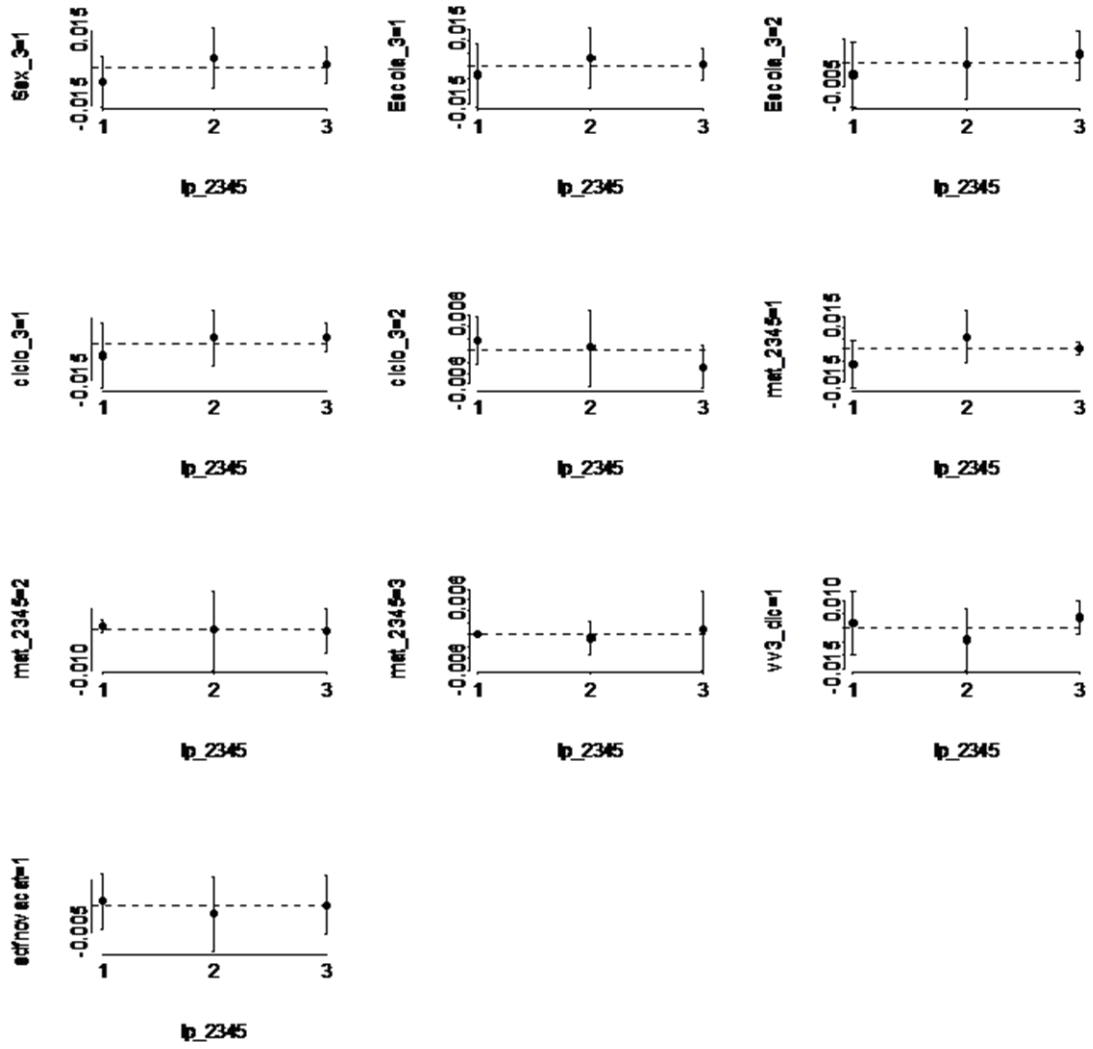

Figura 33: Gráficos de resíduos score para as covariáveis incluídas no modelo tendo como resposta as notas de Língua Portuguesa

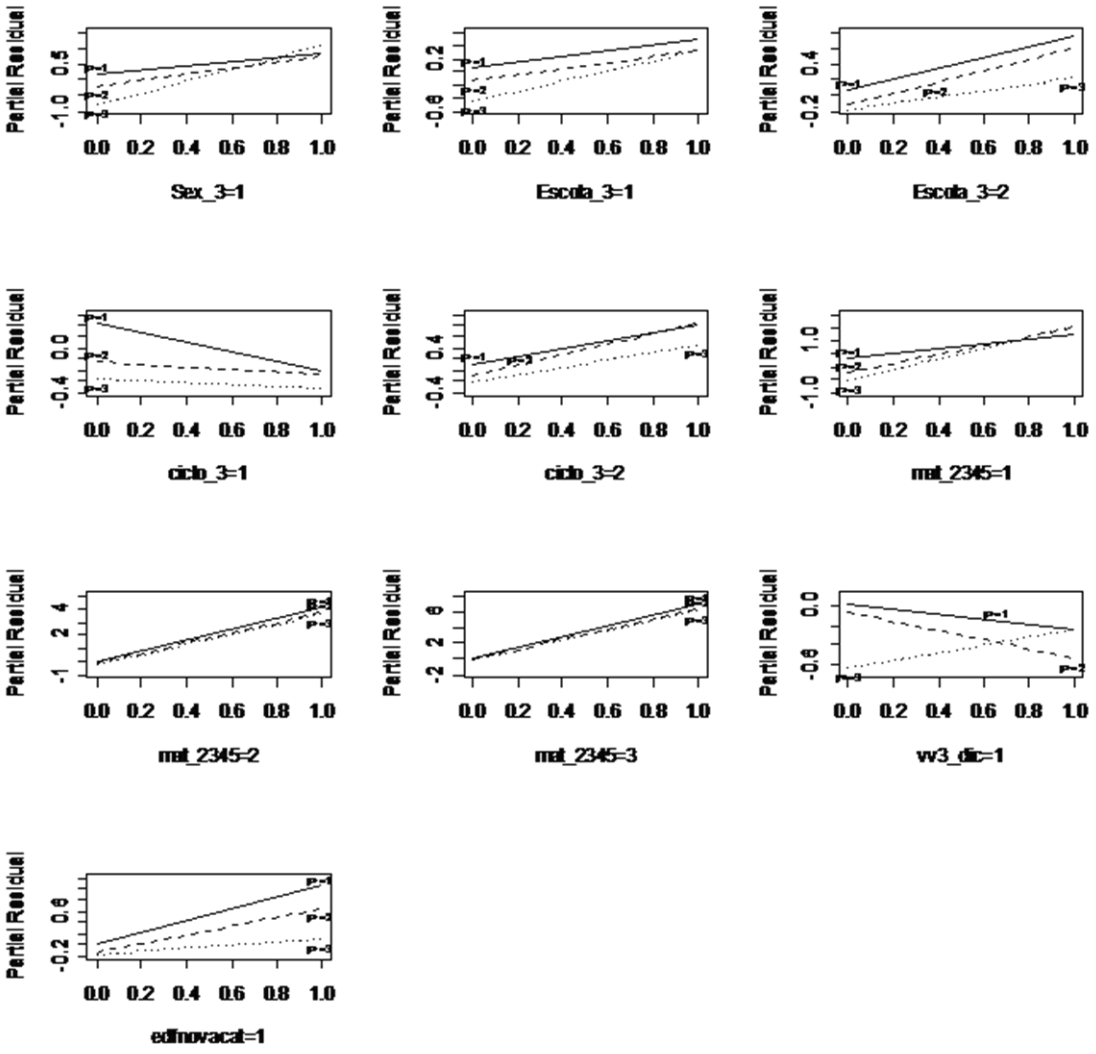

Figura 34: Gráficos de resíduos parciais para as covariáveis incluídas no modelo tendo como resposta as notas de Língua Portuguesa

4.6. Análise de *Missings*

A amostra utilizada neste estudo tem valores omissos nas variáveis dos testes de *Fitnessgram*, pelo que irá ser feita uma abordagem com e sem *missings*.

Para amostras em que o número de casos com dados omissos é pequeno (ex., <5% nas amostras de maior dimensão), a primeira abordagem consiste em eliminar esses casos da análise. Na Tabela 27, podemos observar os dados referentes às variáveis que continham *missings*.

Univariate Statistics							
	N	Média	Desvio-padrão	Missing		No. of Extremes[a]	
				Count	Percent	Low	High
abd_3	755	30.0013	21.47798	142	15.8	0	84
flex_dir_3	811	22.2019	7.16376	86	9.6	7	2
flex_esq_3	807	21.3872	7.62890	90	10.0	0	1
vv_3	842	34.7803	18.72367	55	6.1	0	14
a. Number of cases outside the range (Q1 - 1.5*IQR, Q3 + 1.5*IQR).							

Tabela 27: Análise às variáveis que continham *missings*

4.6.1. Tipos de dados *Missings*: MCAR, MAR ou NMAR

Quando existem variáveis com mais de 5% de dados omissos, devemos estimar valores para substituí-los. No entanto têm que ser realizados testes para determinar o tipo de dados omissos: MCAR, MAR ou NMAR. Para determinar se se tratavam de dados do tipo MCAR, foi realizado o teste Little's MCAR (Tabela 28). Caso o valor p não seja significativo, os dados podem ser assumidos como MCAR. Neste caso, rejeita-se H_0, concluindo-se que os dados não são MCAR.

EM Means[a]			
abd_3	flex_dir_3	flex_esq_3	vv_3
29.4275	22.1939	21.3013	34.7212
a. Little's MCAR test: Chi-Square = 68.898, DF = 17, Sig. = .000			

Tabela 28: Teste Little's MCAR

Quando os dados não são MCAR, os valores em falta devem ser imputados através de métodos de estimação.

De seguida, foi realizado um teste para determinar se os dados são do tipo MAR (Separate Variance t Tests). Como P (2-tail) <=. 05 para cada variável , significa que os valores em falta na variável correspondente à linha são significativamente correlacionados com a variável da coluna e, portanto, não faltam ao acaso. O que vem reforçar a utilização da Imputação de valores.

Para determinar se os dados são do tipo NMAR, procedeu-se em primeiro lugar à aplicação de *Maximum likelihood estimation (MLE)* e de seguida a *Multiple imputation* (MI). O método MLE assume que os valores em falta são MAR (ao contrário de MCAR). Após a aplicação do teste à normalidade, verificou-se que os dados sem valores omissos não são normais. Pelo que foi necessário avançar para o método de Imputação Múltipla. A imputação múltipla (MI) é um método de gerar vários valores simulados para cada dado em falta.

4.6.2. Estimação de *Missings* para o Desempenho a Matemática

Antes de introduzir a estimação dos valores omissos, foi construído um modelo cuja variável resposta é a nota obtida na disciplina de matemática. Pretende-se analisar a influência dos resultados obtidos nos testes de *Fitnessgram* (flexibilidade à direita e à esquerda, abdominais e o teste de vai-vém). Tendo em conta que o comportamento das variáveis é semelhante para a influência na nota obtida nas disciplinas de Português e Educação Física, será apresentado e analisado o caso das notas a Matemática. No entanto, não puderam ser considerados os dados do Agrupamento de Escolas de Castelo de Paiva, uma vez que vinham codificados como “ter alcançado” ou “não ter alcançado” o limite aceitável do teste. Os dados da escola de Redondo e Escola Conde Vilalva foram considerados por serem os valores reais obtidos em cada um dos testes considerados. Foram ainda excluídos os dados do secundário da Escola de Redondo, por serem em número insignificante e para poder ser feita a comparação com a Escola Conde Vilalva que só tem alunos de 2º ciclo e básico.

Inicialmente consideraram-se os valores reais obtidos para as variáveis da aptidão física. Após a estimação dos valores omissos, ajustou-se o modelo com as variáveis significativas tendo-se obtido os resultados apresentados na Tabela 29.

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
flex_dir_3	,077	,029	7,255	1	,007	1,080	1,021	1,142
flex_esq_3	-,055	,027	4,273	1	,039	,947	,899	,997
Lp_3	1,705	,179	91,195	1	,000	5,500	3,876	7,803
Edf_3	,290	,140	4,291	1	,038	1,336	1,016	1,757
ciclo_3(1)	-1,056	,212	24,738	1	,000	,348	,230	,528
Escola_3(1)	-,478	,191	6,251	1	,012	,620	,427	,902
Constant	-4,663	,779	35,852	1	,000	,009		

Tabela 29:Modelo com a estimação de missings

De seguida ajustou-se o modelo com as mesmas variáveis, mas para os dados sem a estimação dos valores omissos, podendo observar-se os resultados obtidos na Tabela 30. Neste modelo temos como variáveis significativas a escola, a nota de Língua Portuguesa e o ciclo a que os alunos pertencem.

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Escola_3(1)	-,359	,204	3,079	1	,079	,699	,468	1,043
LP_3	1,741	,193	81,566	1	<,001	5,701	3,907	8,317
ciclo_3(1)	-,819	,240	11,669	1	,001	,441	,276	,705
flex_dir_3	,004	,030	,022	1	,883	1,004	,946	1,066
flex_esq_3	,010	,027	,145	1	,704	1,010	,958	1,066
Edf_3	,109	,150	,536	1	,464	1,116	,832	1,496
Constant	-4,113	,802	26,305	1	<,001	,016		

Tabela 30: Modelo sem a estimação de *missings*

Apenas algumas variáveis se mantiveram significativas. Em relação ao modelo sem a estimação de *missings* houve diferenças no *Odds Ratio* de algumas variáveis. Assim, para a nota de Língua Portuguesa houve um decréscimo de 4%, o ciclo diminui 27%, a escola diminui 13% e a flexibilidade à esquerda também diminui em 7% o valor para o *Odds Ratio*. Para a flexibilidade à direita houve um aumento de 7% e a nota de Educação Física teve um aumento de 16%.

De seguida considera-se uma abordagem diferente de forma a poder incorporar os dados das três escolas. Para tal, consideram-se as variáveis da aptidão física categorizadas, tal como foi feito na regressão logística. Começou-se por analisar o modelo obtido anteriormente (sem a estimação de *missings*) para os dados com os *missings* estimados para as variáveis da aptidão física que posteriormente foram codificadas. Na Tabela 31, podemos comparar os coeficientes dos dois modelos e observar que não há grandes diferenças a registar entre eles. Até mesmo os valores para o Odds Ratio de cada variável são próximos nos modelos sem e com *missings*.

Variáveis	β		Desvio-padrão		OR	
	Sem est.	Com est.	Sem est.	Com est.	Sem est.	Com est.
Escola_3						
Escola_3(1)	-,682	-,676	,163	,159	,505	,509
Escola_3(2)	-,305	-,349	,190	,173	,737	,706
flex_esq3_dic(1)	,305	,266	,135	,127	1,357	1,305
lp_3_A(1)	1,610	1,602	,153	,147	5,004	4,963
edf_novac(1)	,765	,846	,183	,181	2,150	2,330
idadecat(1)	-,994	-,998	,129	,122	,370	,369
Constant	,374	,394	,208	,198	1,454	1,483

Tabela 31: Comparação entre as variáveis no modelo sem *missings* e modelo com *missings*

Comparando algumas medidas de desempenho para o modelo sem *missings* e para o modelo com *missings* (Tabela 32), verificamos que não há diferenças relevantes entre elas.

Medidas	Sem missings	Com missings
N	1624	1794
Deviance	1681,4	1841,9
Hosmer-Lemeshow	3,483	9,474
AUC	0,734	0,736
Sensibilidade	57,1	68,0
Especificidade	76,7	69,1
AIC	1695,4	1855,9

Tabela 32: Medidas de desempenho para os dois modelos

Finalmente ajustou-se o novo modelo (Tabela 33) considerando os dados com os missings estimados como sendo a nossa amostra real.

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Castelo de Paiva	-,792	,166	22,617	1	<,001	,453	,345	,596
Conde Vilalva	-,266	,179	2,204	1	,138	,767	,571	1,029
Flex_esq.	,226	,129	3,041	1	,081	1,253	1,013	1,550
Lp	1,563	,149	110,314	1	<,001	4,774	3,738	6,099
Edf	,795	,183	18,922	1	<,001	2,214	1,639	2,989
Idade>=15	-,858	,158	29,556	1	<,001	,424	,327	,550
Ciclo (7º,8º,9º)	-,328	,163	4,028	1	,045	,721	,551	,943
Abd	,277	,147	3,552	1	,059	1,319	1,036	1,681
Ciclo (sec.)	,004	,241	,000	1	,988	1,004	,675	1,493
constante	,410	,224	3,344	1	,067	1,506		

Tabela 33: Modelo final

No diagnóstico do modelo, começamos por realizar o teste de bondade de ajustamento de Hosmer-Lemeshow para o qual obtivemos o valor de estatística de teste χ^2=10,667 o que para 8 graus de liberdade corresponde a um valor *p* de 0,221. O que significa que aceitamos a hipótese de um bom ajustamento.

O coeficiente de determinação, R^2 de Nagelkerke apresenta um valor de 0,204. Para este modelo a Curva ROC apresenta uma área de 0,737, ou seja a discriminação feita pelo modelo é aceitável. Realizada uma análise de resíduos não se detectou nenhuma observação anómala.

Em relação às variáveis significativas comuns aos dois modelos, temos a escola, o teste de flexibilidade à esquerda, a nota de Língua Portuguesa e de Educação Física e a idade categorizada. No novo modelo os alunos do Agrupamento de Escolas de Castelo de Paiva vêm a sua possibilidade de ter nota positiva a Matemática diminuída em 55%, enquanto os da Escola Conde Vilalva em 27% quando comparados com os alunos da Escola de Redondo. Já os alunos com idade superior ou igual a 15 anos, têm menos 58% de possibilidade de ter positiva a Matemática. Os alunos com positiva a Língua Portuguesa têm cerca de 5 vezes a mais de ter positiva a Matemática e os alunos com nível cinco a Educação Física cerca de 2,2 vezes a mais.

Neste modelo tem-se ainda como variáveis significativas o ciclo a que os alunos pertencem e o teste de abdominais. Assim, os alunos do ensino básico têm a sua possibilidade de obter nível positivo a Matemática diminuída em cerca de 28%, e os do ensino secundário em 1%. Os alunos que obtiveram resultados acima da média para o teste de abdominais têm até 1,3 vezes a mais de ter positiva a Matemática.

Capítulo V - Considerações Finais

Sabe-se que em jogos com elevado desgaste mental, como é o caso do xadrez, é indispensável uma óptima condição física para aguentar longos jogos, pois o *stress* do jogo associado ao enorme esforço intelectual leva ao consumo de muitos hidratos de carbono, sofrendo os jogadores do mais alto nível desgastes muito grandes.

Na literatura disponível sobre a aptidão física e motora das crianças e adolescentes, refere-se a dificuldade em destacar a contribuição individual de cada um dos múltiplos factores envolvidos como, por exemplo, factores culturais, geográficos, ambientais, maturacionais e genéticos para o desempenho físico e motor (SILVA, 2002; OKANO, 2001). Neste estudo é de considerar o factor localização geográfica, uma vez que a escola de Castelo de Paiva apresenta um menor número de alunos Fora da Zona Saudável do que as duas escolas da região do Alentejo.

Em relação ao sexo dos alunos considerados na amostra, há mais alunos do sexo masculino Fora da Zona Saudável do que o que seria de esperar.

Categorizando a idade em três classes verificou-se que é nas classes etárias mais baixas que se encontra o maior número de alunos Fora da Zona Saudável. O que não deixa de ser curioso e preocupante, uma vez que se esperava que fossem esses que tivessem um peso mais equilibrado para a respectiva altura e consequentemente um IMC dentro dos parâmetros para a sua idade e sexo. Foi então estabelecido um ponto de corte na idade, tendo sido escolhida a idade de 14 anos para dividir em duas classes etárias. Esta escolha deve-se ao facto de ser nesta idade que ocorrem as maiores transformações hormonais nos alunos e por ser a transição para a adolescência. É nesta altura que começam a adquirir novos hábitos e regras alimentares e preocupações com o corpo e sexualidade. Talvez como consequência da época de informatização em que vivemos e da crescente preocupação dos adolescentes com o seu corpo e hábitos alimentares, verificou-se que são os alunos com 15 ou mais anos que estão mais dentro da Zona Saudável. Seria de esperar que fossem os mais novos que apresentassem o IMC dentro da Zona Saudável, tendo por princípio que ainda podem ter implícitos hábitos e regras alimentares mais rígidos. No Agrupamento de Escolas de Castelo de Paiva pertencente à região norte, a possibilidade de estar Fora da Zona Saudável para os alunos mais novos é cerca de 2,5 vezes superior em relação aos mais velhos.

Em relação à interacção entre as escolas e as idades dos alunos, concluímos que a possibilidade dos alunos mais velhos pertencerem à zona saudável é 2,5 vezes superior relativamente aos alunos mais novos, para a Escola de Castelo de Paiva. Fazendo uma comparação entre escolas, verifica-se que para os alunos mais novos apenas há diferenças significativas entre a Escola de Redondo e a Escola de Castelo de Paiva, os alunos da Escola de Castelo de Paiva têm 3 vezes mais possibilidade de pertencer à Zona Saudável. Dentro do grupo de alunos mais velhos, os alunos da Escola de Castelo de Paiva têm 3 vezes mais possibilidade de pertencerem à Zona Saudável do que os alunos da Escola de Redondo. E por sua vez os alunos da Escola de Redondo têm 1,7 vezes mais possibilidade de pertencer à Zona Saudável do que os alunos da Escola Conde Vilalva.

É também no Agrupamento de Escolas de Castelo de Paiva que podemos encontrar a maior percentagem de alunos dentro da Zona Saudável para o IMC. Para os alunos do sexo masculino, os mais velhos têm menos possibilidades de pertencer à zona abaixo da saudável. Para os alunos do sexo feminino, os mais velhos têm menos possibilidades de pertencer à zona acima da saudável. De entre os mais novos, são os do sexo feminino que têm menos possibilidades de estarem na zona abaixo, enquanto que, entre os mais velhos são também os do sexo feminino que têm menos possibilidades de estarem acima da Zona Saudável.

Finalmente, um aluno pertencente ao Agrupamento de Escolas de Castelo de Paiva tem menos ¾ de possibilidades de estar acima da Zona Saudável relativamente a um aluno da Escola de Redondo. Por sua vez, os alunos da Escola de Redondo têm metade das possibilidades de estarem acima da Zona Saudável que os alunos da Escola Conde Vilalva.

Este é ainda um estudo preliminar que permitiu desde já identificar as variáveis Sexo, idade e Escola como factores importantes nos níveis de IMC dos alunos, e deixa várias possibilidades em aberto para um estudo mais completo. Estudo esse que poderá envolver mais escolas e variáveis que possam determinar e analisar a razão para as diferenças entre os níveis de IMC entre escolas situadas em regiões diferentes ou com características similares.

Em relação à possibilidade de associação estatística entre a capacidade física dos alunos, medida pela bateria de testes *Fitnessgram* (bateria de testes que avalia o nível de

aptidão física das crianças/adolescentes) e pelo seu desempenho a educação física, e o seu rendimento escolar, medido pelas notas finais nas disciplinas de Matemática e Língua Portuguesa, também há considerações interessantes a registar. Assim, o sexo e a escola são factores significativos na explicação das notas; neste caso as raparigas têm tendência a terem notas mais elevadas que os rapazes em Língua Portuguesa e em Matemática. A variável escola tem um efeito significativo sobre as notas de Matemática e Educação Física. Por outro lado, os alunos do ensino básico (7º, 8º e 9º anos) e do Secundário têm uma possibilidade de tirar piores notas que os alunos do 2º ciclo (5º e 6º anos).

O facto de os alunos terem positiva a Língua Portuguesa aumenta a possibilidade de terem positiva a Matemática. Os alunos do Agrupamento de Escolas de Castelo de Paiva apresentam uma possibilidade inferior de ter positiva a Matemática, em relação aos alunos das outras duas escolas. No caso da Língua Portuguesa, os alunos que têm positiva a Matemática, também têm maior possibilidade de ter positiva a Língua Portuguesa. Além de um desempenho excelente em Educação Física contribuir para um melhor desempenho em Matemática e em Língua Portuguesa, refira-se que em relação às variáveis da aptidão física, a flexibilidade esquerda revelou-se factor potenciador de uma melhor nota a Matemática e também os abdominais na Matemática e o Vai-vém na Língua Portuguesa se mostraram factores potenciadores de uma melhor nota.

Na análise que envolvia os resultados dos testes de aptidão física não puderam ser considerados, inicialmente, os dados do Agrupamento de Escolas de Castelo de Paiva, uma vez que vinham codificados como “ter alcançado” ou “não ter alcançado” o limite aceitável do teste. Nos modelos estimados só com os valores reais dos testes de aptidão física, registaram-se algumas diferenças no modelo sem os missings comparativamente ao modelo com os valores omissos estimados (utilizando o método de Imputação Múltipla). Assim, na abordagem sem valores omissos a influência da nota de Língua Portuguesa, o ciclo, a escola e o teste de flexibilidade à esquerda diminuíram. O teste de flexibilidade à direita e a nota de Educação Física aumentaram a sua significância no nível de Matemática.

Como as notas das disciplinas envolvidas estão ordenadas por uma escala de 1 a 5, ajustou-se um modelo de regressão ordinal para Matemática e Língua Portuguesa. Em ambos os casos os rapazes têm mais possibilidade de obter piores resultados, e os alunos da Escola de Redondo têm mais possibilidades de ter melhores resultados em Matemática do que em Língua Portuguesa. Também para as duas disciplinas, os alunos

com nível 5 a Educação Física têm mais possibilidades de ter melhores resultados, o que vem confirmar considerações atrás referidas sobre a relação entre excelentes alunos a Educação Física e as notas de Matemática e Língua Portuguesa. O teste de abdominais vê a sua influência positiva nos resultados a Matemática reforçada, enquanto para os alunos com bons resultados no teste de vai-vém há mais possibilidade de terem positiva a Língua Portuguesa.

Para considerar as três escolas nesta abordagem teve que ser feita uma codificação das variáveis da aptidão física. Comparando o modelo obtido sem a estimação de missings com o modelo com os missings estimados, não houve grande diferença nos resultados. Aquando da estimação do modelo final, considerando os dados com os missings estimados como sendo a nossa amostra real, registaram-se alguns resultados curiosos. A variável escola, o teste de flexibilidade à esquerda, a nota de Língua Portuguesa e de Educação Física e a idade categorizada mostraram ser significativas nos dois casos (modelos com e sem missings estimados). No entanto surgiram novas variáveis significativas. Assim, os alunos do ensino básico e secundário têm a sua possibilidade de obter nível positivo a Matemática diminuída. Os alunos com resultados acima da média para o teste de aptidão física abdominais, têm mais possibilidades de ter positiva a Matemática.

É de referir que ao longo das várias abordagens em termos de ajustamento de modelos aos dados da amostra, os testes de aptidão física revelaram influenciar e explicar de forma positiva os resultados positivos nas disciplinas de Matemática e Língua Portuguesa. O que leva a considerar mais hipóteses de estudo entre a aptidão física dos alunos e o seu rendimento escolar, pois há inúmeras variáveis que poderiam vir a integrar este estudo e torná-lo mais rico. E até quem sabe, levar os mais cépticos em relacionar a aptidão física e o rendimento escolar, a aguçarem a curiosidade e integrar o espírito estatístico na condição física dos nossos alunos.

Mens sana in corpore san.

Bibliografia

AAHPERD. Physical Best. Reston, *Virginia: American Alliance for Health, Physical Education and Recreation and Dance*, (1988).

ABREU, M.; SIQUEIRA, A., *Regressão logística ordinal em estudos epidemiológicos*; Revista Saúde (2009); 43(1);183-194.

ACSM. *American College of Sports Medicine: Resource Manual for Guidelines for Exercise Testing and Prescription.* 5th ed. USA: Lippincott, Williams & Wilkins, (2005)

AGRESTI, A. *An Introduction to Categorical Data Analysis*, Second Edition, Wiley-Interscience (2007)

ANANTH CV, KLEINBAUM DG., *Regression models for ordinal responses: a review of methods and applications*. *Int J Epidemiol. (*1997);26(6):1323-33. DOI: 10.1093/ije/26.6.1323

BERGMANN et al , Rev. Bras. Cineantropom. *Desempenho. Hum.* (2005);7(2):55-61

BERGMANN, G.G.; ARAÚJO, M.L.B.; GARLIPP, D.C.; LORENZI, T.D.C.; GAYA A, *Alteração anual no crescimento e na aptidão física relacionada à saúde de escolares*, Revista Brasileira de Cineantropometria e Desempenho Humano. **(**2005); 7(2): 55-61.

BLAIR SN, Kohl HW; PAFFENBARGER RS, Clark DG, COOPER KH, GIBBONS LW (1989). *Physical fitness and all-cause mortality: A prospespective study of healthy men and women.* JAMA 262: 2395-2401

BOUCHARD C, DESPRÉS, JP. *Physical Activity and Health : Atherosclerotic, Metabolic, and Hypertensive Diseases. Res Q Exer Sport* (1995);66(4):268-275.

CONVERSANO, C., CAPPELLI, C., *Missing data incremental imputacion through tree-based methods*, Hardle, W. e Ronz, B. editors, Compstat (2002):14th Conference on Computacional Statistics.

DEMPSTER, A.P., LAIRD, N.M., e RUBIN, D.B., *Maximum likelihood from incomplete data via the EM algoritm*, Journal of the Royal Statistical Society, Series B (Methodological), (1977), 39:1-38.

FITNESSGRAM - Institute For Aerobics Research. User`s Manual. Texas: Institute For Aerobics Research, (1987).

GAYA, A.; GUEDES, D.P.G.; TORRES, L.; CARDOSO, M.; POLETTO, A.; SILVA M.; et al., *Aptidão Física Relacionada à Saúde: um estudo piloto sobre o perfil de escolares de 7 a 17 anos da Região Sul do Brasil*. Revista Perfil. **(**2002); 6(6): 50-60.

HOSMER, D. W. & LEMESHOW, S., *Applied Logistic Regression (Second Edition)*, John Wiley & Sons, New York, (2000)

JOHNSON, D. E., *Applied Multivariate Methods for Data Analysts*, Brooks/Cole Publishing Company, Pacific Grove. (1998)

LEMESHOW, S., HOSMER, D., *A review of goodness of fit statistics for use in the development of logistic regression models.*, American Journal of Epidemiology and Statistics. (1982).

MAROCO, J., *Análise Estatística com utilização do SPSS*, 3ª Edição,Edições Sílabo, Lisboa (2007).

MARQUES, A.T.; GAYA, A., *Atividade Física, Aptidão Física e Educação Para a Saúde: estudos na área pedagógica em Portugal e no Brasil*. Revista Paulista de Educação Física. **(**1999); 13(1): 83-103.

MITTLBOCK, M., SCHEMPER, M., Explained Variation for Logistic Regression, Statistics in Medicine (1996).

NAHAS, M.V.; CORBIN, C.B., *Aptidão Física e Saúde nos Programas de Educação Física: desenvolvimentos recentes e tendências internacionais*. Revista Brasileira de Ciência e Movimento. **(**1992); 6(2): 47-58.

OKANO, A. H.; ALTIMARI, L. R.; COELHO, C.F.; ALMEIDA, P. B. L.; CYRINO, E. S., *Comparação entre o desempenho motor de crianças de diferentes sexos e grupos étnicos*. Rev. Bras. Ciên. E Mov. Brasília 9(3): 39-44, 2001.

POWERS, D.; XIE, Y., *Statistical Methods for Categorical Data Analysis*, Academic Press Inc., (1999)

PULKSTENIS E, ROBINSON TJ., *Goodness-of-fit tests for ordinal response regression models*. *Stat Med.* (2004);23(6):999-1014. DOI: 10.1002/sim.1659

RUBIN, D. B., *Multiple imputation for non reponses in surveys*, New York: Wiley-Interscience (1987).

SILVA, R. J. S. Características de Crescimento, Composição Corporal e Desempenho Físico relacionado à Saúde em Crianças e Adolescentes de 07 a 14 Anos da Região do Cotinguiba (SE), (2002).

SRIVASTAVA, M. S. & KHATRI, E. M., *An Introduction to Applied Multivariate Statistics*, Elsevier Science Publishing, New York (1983).

WHO - WORLD HEALT ORGANIZATION. *Social Determinants of Health: The solid facts.*(2003)

Anexo I

Variáveis e Covariáveis utilizadas

Sigla	Significado
IMC	Índice de Massa Corporal
Sex	Sexo dos alunos
ZN_IMC	Zona Saudável para o IMC
Idadecat1	Idade dos alunos categorizada em três classes: 8-12 anos, 13-15 anos, e 16 ou mais.
Idadecat2	Idade dos alunos categorizada em duas classes: 8-14 anos, e 15 ou mais.
Redondo	Escola de Redondo
Castelo de Paiva	Agrupamento de Escolas de Castelo de Paiva
Conde Vilalva	Escola de Conde Vilalva
ZN_IMC_int 1ºP	Zona Saudável para o IMC (1ºPeríodo) categorizada em três classes: Abaixo da ZN_IMC, Dentro da ZN_IMC, Acima da ZN_IMC.
ZN_IMC_int 3ºP	Zona Saudável para o IMC (3ºPeríodo) categorizada em três classes: Abaixo da ZN_IMC, Dentro da ZN_IMC, Acima da ZN_IMC.
Sex_3	Sexo dos alunos do 3ºPeríodo
Escola_3	Variável para as escolas no 3ºP
Sex_3*Escola_3	Interacção entre as variáveis sexo e escola
Nota_lp	Nível obtido a Língua Portuguesa
Nota_edf	Nível obtido a Educação Física
Nota_mat	Nível obtido a Matemática
Ciclo	Ciclo a que os alunos pertencem: 5º e 6º; 7º, 8º,9º; e secundário
ciclo_3	Ciclo a que os alunos pertencem (3ºP): 5º e 6º; 7º, 8º,9º; e secundário

Sigla	Significado
abd	Resultado obtido no teste “abdominais”
Flex_dir	Resultado obtido no teste “flexibilidade à direita”
Flex_esq	Resultado obtido no teste “flexibilidade à esquerda”
vv	Resultado obtido no teste “vai-vém”
abd_3	Resultado obtido no teste “abdominais” no 3ºP
flex_dir_3	Resultado obtido no teste “flexibilidade à direita” no 3ºP
flex_esq_3	Resultado obtido no teste “flexibilidade à esquerda” no 3ºP
vv_3	Resultado obtido no teste “vai-vém” no 3ºP
Idadecat(1) Idadecat(2)	Idade categorizada com ponto de corte nos 15 anos
Flex_esq_dic	Variável dicotómica para os resultados obtidos no teste “flexibilidade à esquerda”: 1- supera o valor médio de referência do teste; 0- não atinge o valor médio de referência do teste
Flex_dir_dic	Variável dicotómica para os resultados obtidos no teste “flexibilidade à direita”: 1- supera o valor médio de referência do teste; 0- não atinge o valor médio de referência do teste
lpordinal	Nível de Língua Portuguesa categorizado de 2 a 5
edfordinal	Nível de Educação Física categorizado de 2 a 5
matordinal	Nível de Matemática categorizado de 2 a 5

Anexo II

Valores Fitnessgram para a Zona Saudável de Aptidão Física, segundo a idade e sexo dos alunos

Raparigas

Idade	Vaivém		IMC		Abdominais	
	Percursos		(kg/m2)			
10	15	41	23,5	16,6	12	26
11	15	41	24	16,9	15	29
12	23	41	24,5	16,9	18	32
13	23	51	24,5	17,5	18	32
14	23	51	25	17,5	18	32
15	32	51	25	17,5	18	35
16	41	51	25	17,5	18	35
17	41	51	26	17,5	18	35
17+	41	51	27,3	18	18	35

Extensão do Tronco		Extensão de Braços		Senta e Alcança
(cm)		Nº Execuções		(cm)
23	30	7	15	23
23	30	7	15	25,5
23	30	7	15	25,5
23	30	7	15	25,5
23	30	7	15	25,5
23	30	7	15	30,5
23	30	7	15	30,5
23	30	7	15	30,5
23	30	7	15	30,5

Rapazes

Idade	Vaivém		IMC		Abdominais	
	Percursos		(kg/m2)			
10	23	61	21	15,3	12	24
11	23	72	21	15,8	15	28
12	32	72	22	16	18	36
13	41	72	23	16,6	21	40
14	41	83	24,5	17,5	24	45
15	51	94	25	18,1	24	47
16	61	94	26,5	18,5	24	47
17	61	94	27	18,8	24	47
17+	61	94	27,8	19	24	47

Extensão do Tronco		Extensão de Braços		Senta e Alcança
(cm)		Nº Execuções		(cm)
23	30	7	20	20
23	30	8	20	20
23	30	10	20	20
23	30	12	25	20
23	30	14	30	20
23	30	16	35	20
23	30	18	35	20
23	30	18	35	20
23	30	18	35	20