

Alguns perfis de praticantes de atividade física no concelho de Évora

Luísa Carvalho

Departamento de Matemática, ECT - Universidade de Évora, *carvalh.luisa@gmail.com*

Paulo Infante

Centro de Investigação em Matemática e Aplicações e Departamento de Matemática, ECT - Universidade de Évora, *pinfante@uevora.pt*

Anabela Afonso

Centro de Investigação em Matemática e Aplicações e Departamento de Matemática, ECT - Universidade de Évora, *aafonso@uevora.pt*

Palavras-chave: atividade física, curva ROC, razão das chances, regressão logística.

Resumo: Têm sido reconhecidos os efeitos benéficos da atividade física na qualidade da vida das populações, principalmente na área da saúde. Assim, tem-se verificado um aumento ao incentivo da prática de atividade física por parte dos profissionais da educação, saúde e responsáveis políticos. Este trabalho tem como objetivo definir os perfis mais prováveis do praticante de caminhada/pedestrianismo, do praticante de BTT/ciclismo/cicloturismo e do utilizador da ecopista. Com base nos dados de um inquérito realizado no concelho de Évora, construíram-se modelos de regressão logística. Para avaliar a capacidade discriminativa de cada um dos modelos, procedeu-se à análise da curva ROC, constatando-se que os modelos apresentam boa capacidade discriminativa.

1 Introdução

Nos últimos anos, tem-se verificado a existência de uma relação muito estreita entre a saúde e a atividade física. Desta forma, têm sido várias as organizações, como a Organização Mundial de Saúde (OMS) e União Europeia (UE), que recomendam a prática de atividade física. As recomendações dirigem-se aos decisores em todos os níveis (europeu, nacional, regional e local), quer do sector público quer do privado, visando o aumento da atividade física. Ao mesmo tempo que confirmam a abordagem definida pela OMS, procuram definir passos importantes para ajudar a traduzir os objetivos em ações. Neste sentido, o envolvimento e o apoio das autarquias locais são pré-requisitos essenciais para o desenvolvimento e a sustentabilidade das iniciativas e dos programas que levam a população ao aumento da prática

de atividade física [5, 7, 9].

Em parceria com a Câmara Municipal de Évora, elaborámos um estudo com o objetivo de caracterizar o comportamento dos habitantes do concelho de Évora perante a atividade física. Este estudo permitiu evidenciar as preferências dos residentes no concelho de Évora em termos de atividade física, realçando o contexto dessa atividade, seja ela praticada de forma não competitiva e informal, ou praticada mais formalmente ao nível competitivo. As caminhadas/pedestrianismo (41%), atividades *fitness* (18,2%) e BTT/ciclismo/cicloturismo (15%) são as atividades com maior preferência entre os inquiridos. Foram ainda identificadas as principais razões da não prática de atividade física. A ecopista é o equipamento desportivo municipal que apresenta a maior taxa de utilização. O perfil do praticante de atividade física no concelho de Évora é alguém do género masculino, jovem, trabalhador por conta própria, conhecedor de iniciativas desportivas do município e satisfeito com a oferta desportiva do concelho [1, 3].

Neste trabalho pretendemos traçar o perfil dos praticantes de duas das atividades mais praticadas no concelho de Évora e dos utilizadores do equipamento municipal mais utilizado. Estruturalmente, na secção 2 encontra-se a descrição da metodologia utilizada na análise estatística, e na secção 3 são apresentados os resultados dos perfis mais prováveis do praticante de caminhada/pedestrianismo, BTT/ciclismo/cicloturismo e do utilizador da ecopista. Finalmente, na secção 4 tecem-se algumas considerações finais.

2 Metodologia

Entre 24 de Junho e 11 de Agosto de 2011 foram realizadas entrevistas telefónicas junto da população residente no concelho de Évora, com 15 ou mais anos e possuidora de telefone fixo. Em Afonso *et al.* [1] e Carvalho *et al.* [3] podem ser consultados mais detalhes sobre o processo de amostragem. Para identificar os fatores que contribuem ou condicionam a prática de caminhada/pedestrianismo, a prática de BTT/ciclismo/cicloturismo, e a utilização da ecopista, construímos modelos de regressão logística, uma vez que as características que se pretendem explicar (variáveis resposta) são variáveis dicotómicas nominais (i.e. assumem apenas dois valores que indicam a verificação ou não de um dado evento).

Na construção dos modelos verificou-se a existência de um grupo de possíveis covariáveis explicativas comuns aos três perfis: idade, género (feminino ou masculino), zona de residência (freguesias do concelho de Évora), grau de instrução (1º ciclo ou inferior, 2º ciclo, 3º ciclo, secundário, superior) e situação perante o trabalho (reformado, estudante, desempregado, trabalhador por conta própria, trabalhador por conta de outrem).

No cálculo do número de variáveis explicativas a utilizar teve-se em consideração as recomendações de Agresti [2] e Peduzzi *et al.* [8], tendo-se verificado

cuidadosamente as frequências das categorias das variáveis envolvidas na regressão, bem como a existência de problemas de multicolineariedade.

A seleção das covariáveis a integrar os modelos foi realizada de acordo com a metodologia proposta por Hosmer e Lemeshow [4]: 1) o modelo inicial foi composto pelas covariáveis que individualmente apresentaram significância estatística inferior a 0,20; 2) de seguida foi aplicado o método de *backward*, no qual foram retiradas por ordem decrescente dos valores p as covariáveis que apresentaram significância estatística superior a 0,10; 3) de seguida incluíram-se as covariáveis que não entraram no modelo inicial, ficando no modelo as significativas; 4) foi analisado o pressuposto da linearidade das covariáveis contínuas com o *logit* e avaliada a possibilidade de junção de categorias; 5) foram testadas interações entre as covariáveis presentes no modelo e que faziam sentido neste contexto, tendo sido consideradas significativas as interações que atingiram a significância estatística inferior a 0,05; 6) verificou-se a existência de valores atípicos e observações influentes, através do cálculo e representação gráfica da desviância residual, pontos de alavanca, distância de Cook generalizada e os resíduos *dfBetas*; 7) avaliou-se a bondade do ajustamento do modelo pelo teste do Hosmer e Lemeshow [4] e a capacidade discriminativa do modelo através da curva ROC.

Na validação do modelo final, optámos por realizar uma validação cruzada por *bootstrap*, tendo sido ajustados modelos de 10 amostras aleatórias constituídas por 90% dos indivíduos da amostra original, registando-se os valores estimados por cada modelo para os 10% de indivíduos que ficaram de fora em cada modelo, e registando-se também os valores das estatísticas qui-quadrado do teste de ajustamento realizado a cada modelo. Para além desta, realizámos outros tipos de validação, nomeadamente a usual validação cruzada com 70% da amostra para treino e 30% para teste, tendo-se obtido resultados idênticos.

Todas as análises foram realizadas com o programa *R project*, versão 2.15.3.

3 Resultados

As características gerais da amostra estudada já foram descritas em estudos anteriores [1, 3]. Resumidamente, a amostra é constituída por 653 indivíduos, com idades compreendidas entre 15 e os 90 anos, na maioria são do género feminino, trabalhadores e possuem instrução ao nível do ensino secundário.

3.1 Perfil praticante BTT/ciclismo/cicloturismo

Para definir o perfil de um residente que pratica BTT/ciclismo/cicloturismo (referido daqui em diante apenas por ciclismo) foi considerada a totalidade da amostra recolhida ($n = 653$). A característica a explicar foi codificada como: 1 - pratica ciclismo; 0 - caso contrário (c.c.).

Além das covariáveis referidas na secção anterior foram ainda consideradas as variáveis dicotómicas (não ou sim): conhecimento dos equipamentos desportivos oferecidos pelos município, satisfação com a oferta desportiva, conhecimento das iniciativas da autarquia, utilização da ecopista e prática de duas atividades.

Este modelo revelou um bom ajustamento aos dados ($\chi^2_8 = 4,97$; valor $p = 0,76$), apresenta uma excelente capacidade discriminativa ($AUC = 0,86$; $IC_{95\%} =]0,82; 0,91[$), com uma sensibilidade de 75% e uma especificidade de 83% para um ponto de corte igual a 0,129. Na fase de validação, o valor médio do qui-quadrado para os 10 modelos foi igual a 4,1 com um mínimo de 2,3 e um máximo de 5,5, tendo-se obtido $AUC = 0,84$.

São fatores potenciadores da prática de ciclismo: ser utilizador da ecopista, praticar outra atividade desportiva, não ser estudante, ter menos de 65 anos, ser do género masculino, ser habitante de uma freguesia rural e estar satisfeito com a oferta desportiva do concelho de Évora (Figura 1).

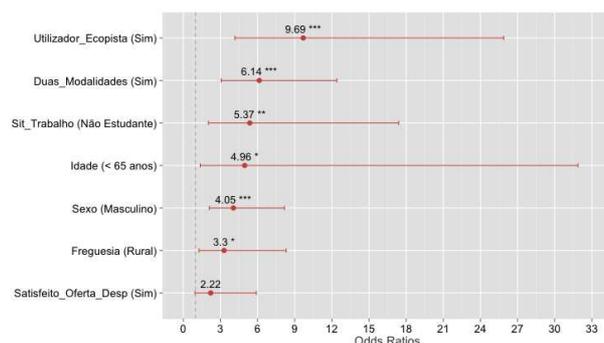


Figura 1: Razão das chances (RC: *Odds ratios*) e respetivos intervalos de confiança a 95%, para o modelo de regressão logística estimado para o praticante de ciclismo.

Admitindo fixas as restantes covariáveis do modelo podemos retirar as seguintes conclusões:

1. Um utilizador da ecopista tem quase 10 vezes mais chances de praticar ciclismo quando comparado com um não utilizador deste equipamento;
2. Para um indivíduo que pratique duas atividades desportivas, a chance de ser praticante de ciclismo aumenta 6 vezes relativamente a quem não pratique duas atividades;
3. Um indivíduo que não seja estudante tem quase 5,5 vezes mais chance de praticar ciclismo do que um estudante;
4. Um munícipe com menos de 65 anos tem 5 vezes mais chance de praticar ciclismo do que uma pessoa com 65 ou mais anos;

5. Um indivíduo do género masculino tem cerca de 4 vezes mais chance de ser praticante de ciclismo que um indivíduo do género feminino;
6. Um morador de uma freguesia rural tem 3 vezes mais chances de praticar ciclismo relativamente a um morador de uma freguesia urbana;
7. Um indivíduo que esteja satisfeito com a oferta desportiva do concelho tem 2 vezes mais chance de praticar ciclismo relativamente a alguém não satisfeito.

3.2 Perfil praticante caminhada/pedestrianismo

Na definição do perfil do praticante de caminhada/pedestrianismo (referido daqui em diante apenas por caminhada) foi considerada a amostra truncada aos praticantes de atividade física que praticam exclusivamente caminhada. A característica a explicar foi codificada como: 1 - pratica caminhada; 0 - c.c.

Além das covariáveis referidas anteriormente foram ainda consideradas as variáveis: pratica sozinho (sim ou não), utiliza a ecopista (sim ou não), tempo de prática, motivo da prática (saúde, diversão, controlo de *stress*, perder peso, socialização, estética, ou outros), quem incentivou a prática (próprio, amigos, família, médicos, professores, ou outros), espaço onde realiza a prática (equipamentos municipais, equipamentos privados, escola, ou outros), orçamento da prática, número de vezes que realiza a prática por semana, número de minutos por sessão e anos de prática desta atividade.

O modelo revelou um bom ajustamento ($\chi^2_8 = 4,12$; valor $p = 0,85$), podendo também concluir-se que tem uma muito boa capacidade discriminativa ($AUC = 0,81$; $IC_{95\%} =]0,77; 0,85[$) com uma sensibilidade de 67% e uma especificidade de 79% para um ponto de corte igual a 0,272. Na fase de validação, o valor médio do qui-quadrado para os 10 modelos foi igual a 6,8 com um mínimo de 3,2 e um máximo de 9,5, tendo-se obtido $AUC = 0,79$. São fatores potenciadores da prática de caminhada: não ser estudante nem trabalhador por conta própria, praticar mais de três vezes por semana e em sessões com duração inferior a 60 minutos, ser do sexo feminino, ser utilizador da ecopista e ser mais velho (Tabela 1). Note-se que as últimas três características referidas resultam da interpretação das interações que envolvem as variáveis género, idade e utilização da ecopista, como veremos em seguida.

Admitindo fixas as restantes covariáveis do modelo podemos retirar as seguintes conclusões:

1. Um indivíduo que realize prática desportiva mais do que 3 vezes por semana tem cerca de 6 vezes mais chances de ser praticante de caminhada que um indivíduo que realize menos vezes por semana;

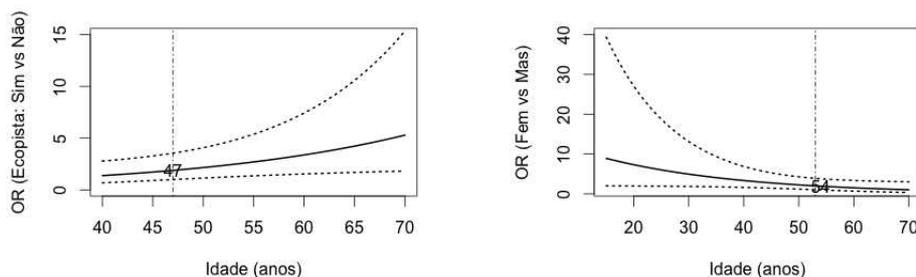
Tabela 1: Estimativas dos parâmetros ($\hat{\beta}$) e respectivos desvios-padrão ($SE(\hat{\beta})$), valores p associados, razão das chances (RC) e respectivos intervalos de confiança a 95% ($IC_{95\%}$), do modelo de regressão logística para o praticante de caminhada/peDESTRIANISMO de entre os praticantes de atividade física.

Variável	$\hat{\beta}$	$SE(\hat{\beta})$	valor p	$RC (IC_{95\%})$
Sem interação				
Número de minutos por sessão:				
< 60 minutos				1,00
≥ 60 minutos	-1,41	0,36	0,000	0,24 (0,12-0,49)
Situação perante o trabalho:				
Outras situações				1,00
Estudante ou trabalhador por conta própria	-1,57	0,49	0,002	0,21 (0,08-0,55)
Número de vezes que realiza a prática por semana:				
Até 3 vezes				1,00
Mais de 3 vezes	1,78	0,35	< 0,001	5,92 (3,01-11,65)
Com interação				
Idade:	-0,01	0,02	0,859	-
Gênero:				
Feminino				1,00
Masculino	-2,78	1,04	0,008	-
Utiliza ecopista:				
Não				1,00
Sim	-1,46	1,00	0,146	-
Gênero*Idade	0,04	0,02	0,057	-
Utiliza ecopista*Idade	0,05	0,02	0,024	-
Constante	-0,19	0,83	0,822	

- Um indivíduo que não seja estudante nem trabalhador por conta própria tem aproximadamente 5 vezes mais chances de praticar caminhada do que um estudante ou um trabalhador por conta própria;
- Um indivíduo que realize atividade física durante menos de 1 hora tem 4 vezes mais chances de ser praticante de caminhada que um indivíduo que pratique atividade física durante mais tempo.
- Ser utilizador da ecopista aumenta as chances de ser praticante de caminhadas, acentuando-se a importância desta variável com o aumento da idade. Por exemplo, para pessoas com 50 anos, caso seja utilizador da infraestrutura tem 2 vezes mais chances enquanto que aos 70 anos as chances aumentam para cerca de 5,5 vezes. De observar que até aos 47 anos as diferenças entre utilizadores e não utilizadores não são

significativas (Figura 2a);

5. Um praticante de 20 anos do género feminino tem cerca de 7 vezes mais chances de praticar caminhada do que um praticante do género masculino com a mesma idade, aos 40 anos as chances diminuem para um pouco mais de 3 vezes e a partir dos 54 anos as diferenças entre os géneros deixam de ser significativas (Figura 2b).
6. De entre os utilizadores da ecopista, quanto maior a diferença de idades maior a chance do mais velho ser praticante de caminhadas relativamente ao mais novo. Apesar da diferença entre idades ser significativa para ambos os géneros, a sua influência é muito maior no género masculino do que no feminino. Por exemplo, comparando duas pessoas com diferenças de 15 anos de idade, a mais nova tem 3,5 vezes mais chances de praticar caminhada do que a mais velha se ambas forem do género masculino, diminuindo as chances para menos de 2 vezes se ambas forem do género feminino (Figura 3). Para os não utilizadores da ecopista o aumento da idade não é significativo.



(a) Razão das chances (OR) do utilizador da ecopista relativamente ao não utilizador em função da idade.

(b) Razão das chances (OR) do género masculino relativamente ao feminino em função da idade.

Figura 2: Razão das chances (linha contínua) e intervalos de confiança a 95% (linha descontínua), para o modelo de regressão logística estimado para o praticante de caminhada/pedestrianismo. A linha vertical marca a zona de significância.

3.3 Perfil praticante utilizador da ecopista

Para definir o perfil de um utilizador da ecopista foi considerada a totalidade da amostra recolhida ($n = 653$). A variável resposta foi definida como: 1 - utiliza a ecopista; 0 - c.c.

Além das covariáveis referidas na secção anterior foram ainda consideradas as variáveis dicotómicas (não ou sim): conhecimento de equipamentos desportivos oferecidos pelo município e satisfação com a oferta desportiva.

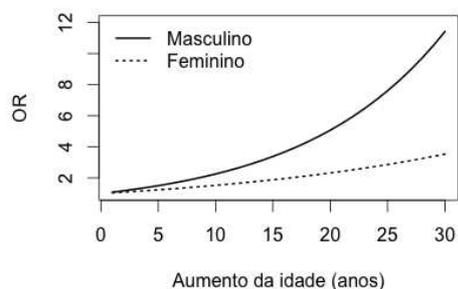


Figura 3: Razão das chances (OR) para várias diferenças de idade para cada género e utilizadores da ecopista, para o modelo de regressão logística estimado para o praticante de caminhada/peDESTRIANISMO.

O modelo revelou um bom ajuste aos dados ($\chi^2_8 = 5,3$; valor $p = 0,73$), apresenta uma boa capacidade discriminativa ($AUC = 0,75$; $IC_{95\%} =]0,71; 0,79[$) com uma sensibilidade de 67% e uma especificidade de 70% para um ponto de corte igual a 0,499. Na fase de validação, o valor médio do qui-quadrado para os 10 modelos foi igual a 5,6 com um mínimo de 2,7 e um máximo de 9,1, tendo-se obtido $AUC = 0,73$.

Foram identificados como fatores potenciadores para um indivíduo ser utilizador da ecopista: ter idade inferior a 65 anos, ter instrução ao nível do 3º ciclo, ser trabalhador por conta própria, por conta de outrem ou estudante, estar satisfeito com a oferta desportiva do concelho e viver nas freguesias junto à ecopista (Figura 4).

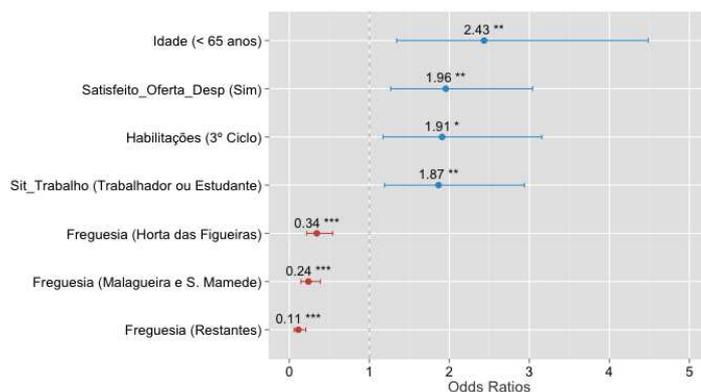


Figura 4: Razão das chances (RC: *Odds ratios*) e respetivos intervalos de confiança a 95%, para o modelo de regressão logística estimado para o utilizador da ecopista.

Admitindo fixas as restantes covariáveis do modelo podemos retirar as seguintes conclusões:

1. Um munícipe com menos de 65 anos tem 2,5 vezes mais chances de utilizar a ecopista do que uma pessoa com 65 ou mais anos;
2. Para alguém satisfeito com a oferta desportiva do concelho as chances de utilizar a ecopista são 2 vezes superior relativamente a quem não está satisfeito;
3. Para os que têm uma formação ao nível do 3^o ciclo, as chances de ser utilizador da ecopista são aproximadamente 2 vezes superior a quem tem outra formação;
4. Um trabalhador por conta própria, um trabalhador por conta de outrem ou um estudante tem quase 2 vezes mais chances de utilizar a ecopista do que um indivíduo desempregado, reformado ou com outra situação de trabalho;
5. Um indivíduo que viva nas freguesias junto à ecopista tem cerca de 9 vezes mais chances de ser utilizador da ecopista do que um habitante numa freguesia rural ou em Santo Antão, tem cerca de 4 vezes mais chances de ser utilizador da ecopista do que um morador da Horta das Figueiras, e tem cerca de 3 vezes mais chances de ser utilizador da ecopista do que um habitante da Malagueira ou de S. Mamede.

4 Considerações Finais

Com o presente trabalho identificámos as principais características dos praticantes das duas atividades físicas mais praticadas no concelho de Évora e, ainda, dos utilizadores do equipamento municipal mais utilizado, a ecopista. O perfil mais provável do praticante de caminhada/peDESTRIANISMO, de entre os praticantes, distingue-se por ser do género feminino, não estudante nem trabalhador por conta própria, que pratica mais de 3 vezes por semana em sessões inferiores a uma hora e utilizador da ecopista no caso de ser mais idoso.

O perfil mais provável do praticante de BTT/ciclismo/cicloturismo, de entre os residentes, é alguém do género masculino, morador numa freguesia rural, com menos de 65 anos, utilizador da ecopista, não estudante, satisfeito com a oferta desportiva do concelho e que pratica uma outra atividade desportiva. O perfil mais provável do utilizador da ecopista é alguém que vive numa das freguesias junto à ecopista, com menos de 65 anos, empregado ou estudante, com uma formação ao nível do 9^o ano e satisfeito com a oferta desportiva do concelho.

As limitações deste estudo são inerentes aos métodos utilizados, sendo de salientar de entre as mesmas o facto de as entrevistas abrangerem apenas a população residente em Évora, à data dos contactos e com telefone fixo,

ainda que, segundo os dados do INE de 2011, a taxa de cobertura de telefone fixo no concelho seja elevada (82% da população).

Agradecimentos

Os autores agradecem à Divisão de Desporto da Câmara Municipal de Évora, em especial à Dr.^a Teresa Engana, toda a disponibilidade e dedicação ao longo deste trabalho.

Anabela Afonso e Paulo Infante são membros do Centro de Investigação em Matemática e Aplicações (CIMA-UE), centro de investigação apoiado financeiramente pela Fundação para a Ciência e Tecnologia (FCT).

Referências

- [1] Afonso, A., Infante, P., Carvalho, L., Engana, T.. *A Atividade física no concelho de Évora*. Câmara Municipal de Évora, Évora.
- [2] Agresti, A. (2007). *An Introduction to Categorical Data Analysis*, 2nd Edition. Wiley.
- [3] Carvalho, L., Infante, P., Afonso, A. (aceite). Évora residents and sports activity. *Actas do XIX Congresso da Sociedade Portuguesa de Estatística*.
- [4] Hosmer, D., Lemeshow, S. (2000). *Applied Logistic Regression*, 2nd Edition. Wiley.
- [5] IDP (2009). Orientações Europeias para a Actividade Física - Políticas para a Promoção da Saúde e Bem-Estar. Instituto do Desporto de Portugal. Acedido em 30 de março de 2013. Disponível em http://www.idesporto.pt/ficheiros/File/Livro_IDPfinalJan09.pdf.
- [6] McCullagh, P., Nelder, J.A. (1989). *Generalized Linear Models*, 2nd Edition. Chapman & Hall.
- [7] Observatório Nacional da Actividade Física e do Desporto (2011). *Livro Verde da Actividade Física*. Instituto Desporto Portugal. Acedido em 30 de março de 2013. Disponível em http://observatorio.idesporto.pt/Multimedia/Livros/Actividade/LVerdeActividadeFisica_GERAL.pdf.
- [8] Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology* 49:1373-1379.
- [9] WHO (2010). *Global recommendations on physical activity for health*. World Health Organization. Acedido em 30 de março de 2013. Disponível em http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf.