

## Seminário

Grupo de Probabilidades e Estatística

16 de dezembro de 2020 15:15

Zoom meeting

---

### Previsão do número de consultas: uma comparação de modelos estatísticos preditivos por partições aleatórias de um ficheiro de utentes numa USF

**Pedro Rebelo**

*USF Aveiro Aradas, ACeS Baixo Vouga*

#### Resumo

**Objectivo:** Proceder a uma análise exploratória multivariada de preditores do número de consultas num ficheiro médico, e com base nessa análise criar modelos estatísticos de previsão.

**Justificação:** A estimativa das necessidades em número de consultas por conjunto de utentes (ficheiro médico ou uma partição do mesmo) é importante para se poder alocar recursos (nomeadamente no desenho de horários). Os modelos tradicionais de estimação têm pressupostos teóricos como a cobertura e o valor médio de consultas ou unidades ponderadas (UP) mas não têm em conta as particularidades reais observadas na distribuição conjunta do nº de consultas com os seus preditores.

**Material e Métodos:** Estudo retrospectivo. Procedeu-se a uma análise exploratória de dados multivariada de uma amostra anonimizada de 10000 utentes de uma USF do ACeS BV referente ao ano de 2016. Criaram-se modelos estatísticos, com diferentes combinações de regressores utilizando técnicas de stepwise. Estes, foram comparados entre si e com modelos pré-existentes, na totalidade da amostra e por partições da mesma, recorrendo a técnicas de reamostragem. Procedeu-se a uma validação cruzada dos diferentes modelos, em partições de dados que não foram utilizados na sua criação.

This seminar is supported by CIDMA – Center for Research and Development in Mathematics and Applications through FCT – Fundação para a Ciência e a Tecnologia, within projects UIDB/04106/2020 and UIDP/04106/2020.

## Seminário

### Grupo de Probabilidades e Estatística

16 de dezembro de 2020 15:15

Zoom meeting

**Variáveis em estudo:** variável dependente: número de consultas. Preditores base: Sexo; Idade; HTA (ICPC-2 K86 e K87); Diabetes (ICPC-2 T89 e T90); perturbação depressiva (ICPC-2 P76); patologia osteoarticular (ICPC-2 Lxx). Preditores sintéticos por transformação ou tratamento dos anteriores nomeadamente da Idade (UP,UP\_SIM).

**Resultados:** Ocorrência de contagens zero (32.6% dos casos) elevada e de outliers (3.76% dos casos) justifica um index de dispersão de Fisher  $ID = 4.215$  elevado. Modelos de Poisson mostraram-se desadequados. Modelos binomial negativo, zero inflacionado e Hurdle (com distribuição binomial negativa) mostraram-se mais adequados a descrever a relação entre as variáveis. O modelo que se mostrou mais adequado a explicar a relação entre as variáveis e a prever foi o de Hurdle. Em termos de capacidade preditiva, todos os modelos se revelaram úteis.

**Conclusões:** Preditores base sem nenhuma transformação, facilitaram a interpretabilidade dos modelos. Os sintéticos, resultado da transformação dos primeiros, aumentaram o poder preditor dos modelos sem perder interpretabilidade dos mesmos.

Link para aceder ao Meeting via Zoom:

<https://videoconf-colibri.zoom.us/j/85756037867?pwd=QVBZWJXUGxaS0hHOGRqQ3c4aWVCUT09>

ID da reunião: 857 5603 7867

Senha da reunião: **046504**

This seminar is supported by CIDMA – Center for Research and Development in Mathematics and Applications through FCT – Fundação para a Ciência e a Tecnologia, within projects UIDB/04106/2020 and UIDP/04106/2020.