

Centro de Investigação em Matemática e Aplicações
Departamento de Matemática
Programa de Doutoramento em Matemática

Seminário online

06 de janeiro de 2021

Statistics of Extremes in Sports –
the POT methodology in action

Lígia Henriques-Rodrigues

(Professora Auxiliar, ligiahr@uevora.pt)

Departamento de Matemática, Escola de Ciências e Tecnologia, Universidade de
Évora

Centro de Investigação em Matemática e Aplicações, Instituto de Investigação e
Formação Avançada, Universidade de Évora

Abstract The main objective of the Extreme Value Theory (EVT) is the prediction of rare events and in sports these events are the ones related with *gold medals* and/or *records*. Therefore the statistics more appealing to champions and coaches are the extreme order statistics, and in particular maximum (or minimum) values and records. When dealing with extreme events we are interested in the tails of the distribution, where there are usually only few observations available and inference plays an important role. The distribution of the maximum of n independent and identically distributed observations, after proper normalization, converges to the designated distribution of Extreme Values, where the shape parameter of the distribution ξ is called the *tail index* or *extreme value index* (EVI), the primary parameter in EVT. The estimation of the EVI is one of the basis for the estimation of other parameters of rare events, like the right endpoint of the model underlying the data, a high quantile, the return period and the probability of exceedance of a high level. In this seminar we first sketch some of the most important results in EVT and then we address two applications of the Peaks over Threshold (POT) methodology in Sports. Under certain conditions, the probability distribution of the exceedances, i.e., of

the observations that exceed a certain threshold u , follows a generalized Pareto distribution (GPD). In the first one we use the POT-methodology in a classic framework to develop a new swimming performance index that surpasses the points system developed by the International Swimming Federation (FINA). In the second application we use a dynamic Bayesian model for the shape and scale parameters of the GPD distribution and MCMC methods to forecast the athletics records.

Keywords: Extreme Value Theory, Parametric estimation, Generalised Pareto Distribution, Peaks over threshold, Dynamic Bayesian model, MCMC methods, athletics records.

Acknowledgements This talk has been partially supported by Centro de Investigação em Matemática e Aplicações (CIMA), through the Project UIDB/04674/2020 of FCT-Fundação para a Ciência e a Tecnologia, Portugal.

