**Statistical Data Analyst (internship 2019)**

Here you find the details for the internship named "Statistical Data Analyst (internship 2019)" at the Royal Observatory of Belgium.

**Name: Statistical Data Analyst (internship 2019)**

**Company: Royal Observatory of Belgium**

**Description:**

The Royal Observatory of Belgium in Uccle is one of the main places in Belgium where research on astronomical topics is conducted. Its department of Solar Physics and Space Weather has as primary goal to advance knowledge on the Sun and its influence on the solar system, through **research** and **observations**. We provide this knowledge and expertise to the scientific community, to the society, the government and the industry through **operational services** and **dissemination** at both the national and the international level. The specific topics that we aim to research are inspired by our own operational activities, by our heritage in some sub-disciplines and – last but not least – by a passion for understanding the Sun and its effects at a more fundamental level.

We have a data set of 8274 solar flares detected in images of the AIA instrument onboard the satellite Solar Dynamics Observatory (SDO), and we want to investigate the statistical distribution of the brightness of these solar eruptions. A first analysis (performed with the poweRlaw package in R) showed that a lognormal distribution (<https://en.wikipedia.org/wiki/Log-normal_distribution>) provides a better fit than a power law distribution (<https://en.wikipedia.org/wiki/Power_law>), see Figure 1. This is interesting, since power laws are the prevalent model for many natural processes, including solar flares.

We want to investigate this further. In particular, we want to compare the power law and lognormal fits to other statistical models. We also want to compare the methods and numerical results obtained by a few existing statistical libraries, such as the poweRlaw and the fitdistrplus library in R and the powerlaw library in Python (see <http://tuvalu.santafe.edu/~aaronc/powerlaws/>).

The aim of the internship is to employ a few existing statistical libraries in R and Python to investigate solar flare data sets for several statistical models, under guidance from scientists at the Royal Observatory of Belgium.

### maximum_brightness_CCDF_fits

Figure 1: The black curve shows the Complementary Cumulative Distribution Function (CCDF) of the SDO/AIA peak flare intensity. The red and green curves are the CCDF of the power law respectively lognormal fit to the data. The vertical red and green lines, respectively, are the lower limits for which the power law and lognormal fit are valid.

What you will do:

* Get work experience in a scientific institute, with the opportunity to attend seminars and interact with researchers
* Get to know and work with a unique data set of solar flares
* Work with advanced statistical methods and software packages
* Write and debug your own code in Python and R and run it to conduct statistical analysis
* Interpret and compare results from statistical software packages
* Independent work in scientific research in collaboration with several scientists

If you are interested, please apply by sending both a cover letter and a resume to cis.verbeeck@oma.be.

**Target profiles:** Master student civil engineer, physics or mathematics

**In industries:** Scientific research

**Required special knowledge:** Knowledge of statistics and data analysis is recommended. Experience with Python and R are an asset. A great job for those with curiosity and a keen interest in science and mathematics.

**Duration:** 30 to 60 days

**Paid:** No

**Net wage:** -

**Foreign:** No

**Contact: Cis Verbeeck** (researcher in solar physics)

 **Royal Observatory of Belgium**

**Ringlaan 3, 1180 Brussels**

**cis.verbeeck@oma.be**

**Tel +32472739914**