Non-Parametric Statistics Workshop 1

David Plazas Escudero
Juan Pablo Vidal
Juan Sebastián Cárdenas-Rodríguez
Mathematical Engineering, Universidad EAFIT

May 6, 2020

1 Workshop Exercises

Exercise 3

Calculate and plot the confidence bands for the empirical continuous distribution function (ECDF) of the coldest and hottest year in average with a confidence of 95 % . Are there any sectors that are not enclosed in the bands?

Proof. Let n be the size of the sample and $1-\alpha$ the desired confidence for the bands. To calculate the confidence bands for the ECDF, we first define ϵ_n by the following formula:

$$\epsilon_n = \sqrt{\frac{1}{2n}\ln\left(\frac{2}{\alpha}\right)}$$

Let $\hat{F}_n(x)$ be the ECDF. Then, for each x in the ECDF we define the lower $(L(\cdot))$ and upper $(U(\cdot))$ bound by:

$$L(x) = \max\{\hat{F}_n(x) - \epsilon_n, 0\}$$
$$U(x) = \min\{\hat{F}_n(x) + \epsilon_n, 1\}$$

The results obtained by using the temperatures of the coldest and hottest year are seen in Figure 1.

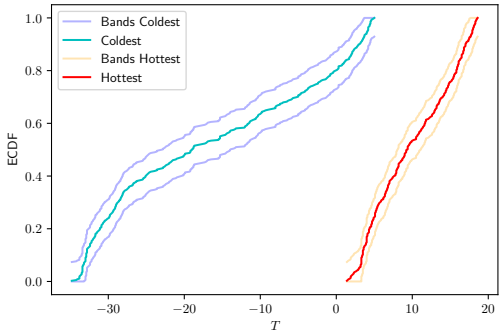

Figure 1: Bands for the coldest and hottest year.

It can be seen that the upper and lower bands fully enclose the ECDF. Nevertheless, there exists two points where the function and the bands meets. This happens in the lower band at the beginning and the upper band at the final point.
This phenomenon is due to the full certainty at those points in a sense that the lower bound, at the start, has to be the same point as it cannot go lower than 0. A similar reasoning can explain the upper bound and the final point. □

Exercise 4

Exercise 7

Exercise 11

Exercise 12

Exercise 14

2 Book Exercises

All exercises in this section are extracted from [1].

Exercise 16–10

References

[1] Larry Wasserman. *All of Nonparametric Statistics*. Springer Science & Business Media, 2006.