

## Accepted Manuscript

The future of statistics and data science

Sofia C. Olhede, Patrick J. Wolfe

PII: S0167-7152(18)30087-7  
DOI: <https://doi.org/10.1016/j.spl.2018.02.042>  
Reference: STAPRO 8164

To appear in: *Statistics and Probability Letters*



Please cite this article as: Olhede S.C., Wolfe P.J., The future of statistics and data science. *Statistics and Probability Letters* (2018), <https://doi.org/10.1016/j.spl.2018.02.042>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

# The Future of Statistics and Data Science

Sofia C. Olhede

*Department of Statistical Science, University College London*

Patrick J. Wolfe

*Department of Statistics, Purdue University*

---

## Abstract

The ubiquity of sensing devices, the low cost of data storage, and the commoditization of computing have together led to a big data revolution. We discuss the implication of this revolution for statistics, focusing on how our discipline can best contribute to the emerging field of data science.

*Keywords:* Algorithmic transparency, Data analysis, Data governance, Predictive analytics, Statistical inference, Structured and unstructured data  
*2010 MSC:* 00-01, 99-00

---

## 1. Introduction

The Danish physicist Niels Bohr is said to have remarked: “Prediction is very difficult, especially about the future.” Predicting the future of statistics in the era of big data is not so very different from prediction about anything else.  
5 Ever since we started to collect data to predict cycles of the moon, seasons, and hence future agriculture yields, humankind has worked to infer information from indirect observations for the purpose of making predictions.

Even while acknowledging the momentous difficulty in making predictions about the future, a few topics stand out clearly as lying at the current and future  
10 intersection of statistics and data science. Not all of these topics are of a strictly technical nature, but all have technical repercussions for our field. How might these repercussions shape the still relatively young field of statistics? And what

Download English Version:

<https://daneshyari.com/en/article/7548511>

Download Persian Version:

<https://daneshyari.com/article/7548511>

[Daneshyari.com](https://daneshyari.com)