The Danish PhD-programs in Statistics

And a personal perspective on being a statistical researcher

Niels Richard Hansen
Department of Mathematical Sciences

Brief bio

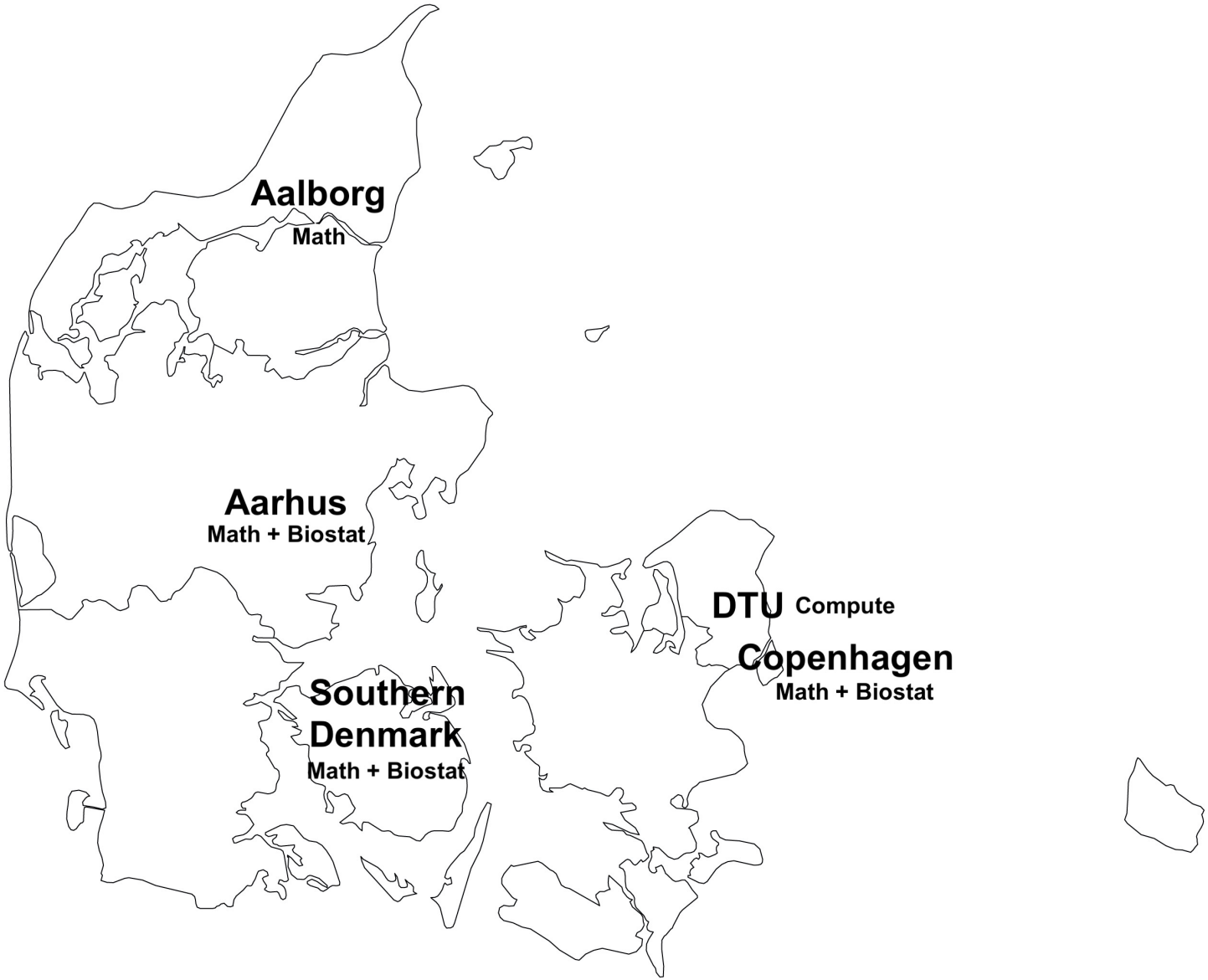
- Professor of Computational Statistics, University of Copenhagen. Head of section of Statistics and Probability Theory, Department of Mathematical Sciences.
- Has been editor of *Scandinavian Journal of Statistics* and Chair of the European Regional Committee of the *Bernoulli Society*.
- Has been a member of a number of hiring and assessment committees – in Copenhagen and internationally.
- Supervised 35 students for their master's theses.
- Main supervisor for 6 graduated PhD-students, and currently supervising 3 PhD-projects.

The Danish PhD program

- 3 years program
- 1/2 years of coursework
- 1/2 years of teaching
- 2 years of research, including mandatory stay abroad
- Prerequisite is the equivalent of a 3 years BSc and 2 years MSc

Recently, 3 + 5 programs require only a BSc, but are not widely used.

Statistics research departments in DK



PhD degrees in Statistics in DK 2016-2020*

Department	# PhD	# Faculty
Aalborg Math	6	10
South DK Math	4	4
Aarhus Math	10	12
Copenhagen Math	14	13
DTU Compute	15	15
South DK Biostat	2	5
Aarhus Biostat	5	7
Copenhagen Biostat	6	14
Total	62	80

*Based on a survey among the departments. Numbers are ballpark figures.

Where did they go?

Position	Number
Postdoc	12
Tenure-track	8
Public institutions	5
Pharma	7
Energy	5
Finance/Insurance	4
Data Scientist	3
Consultancy	1
Other	2

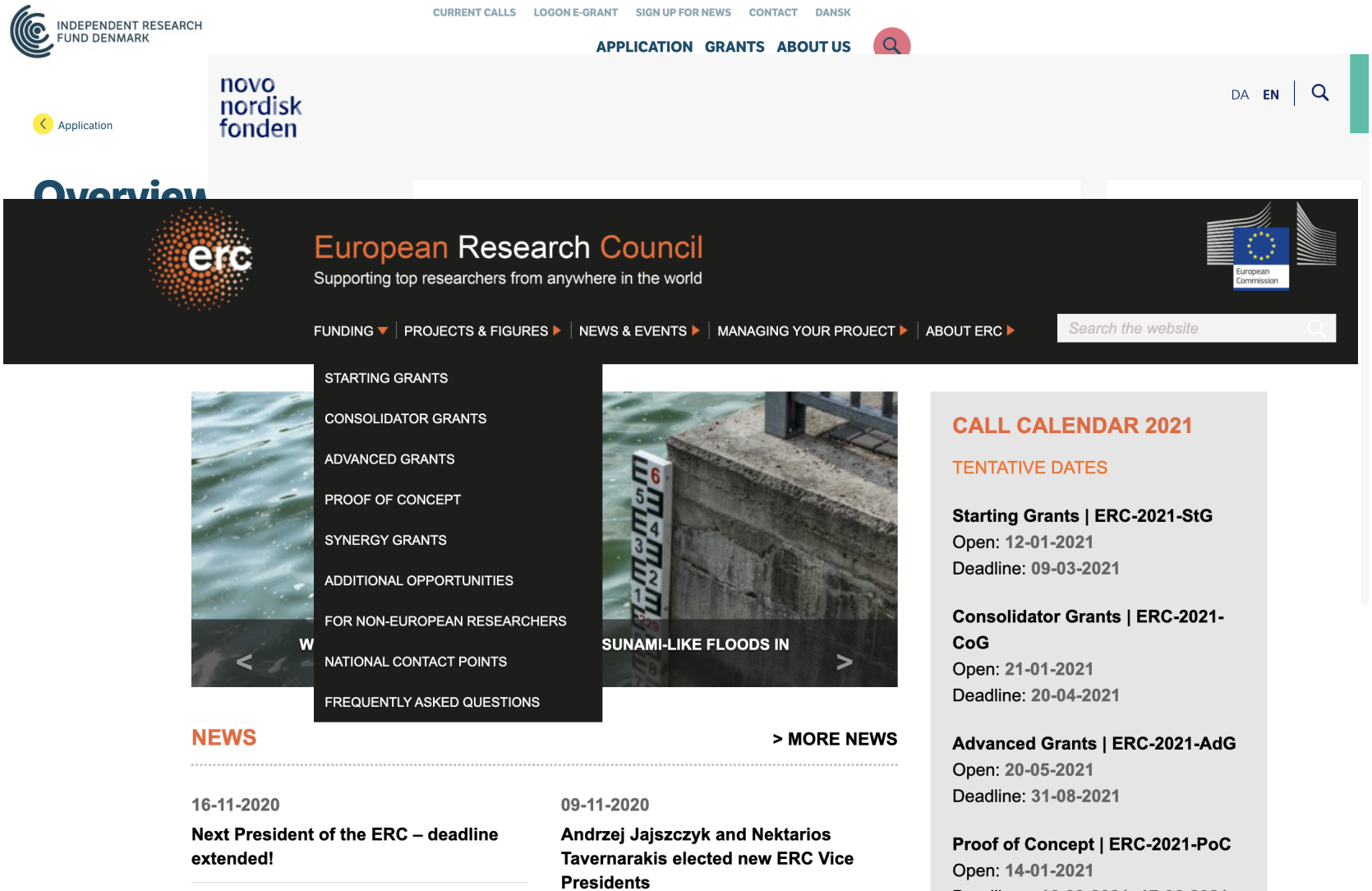
DTU Compute excluded. A verbal description of their students' career tracks matches the distribution above.

Funding

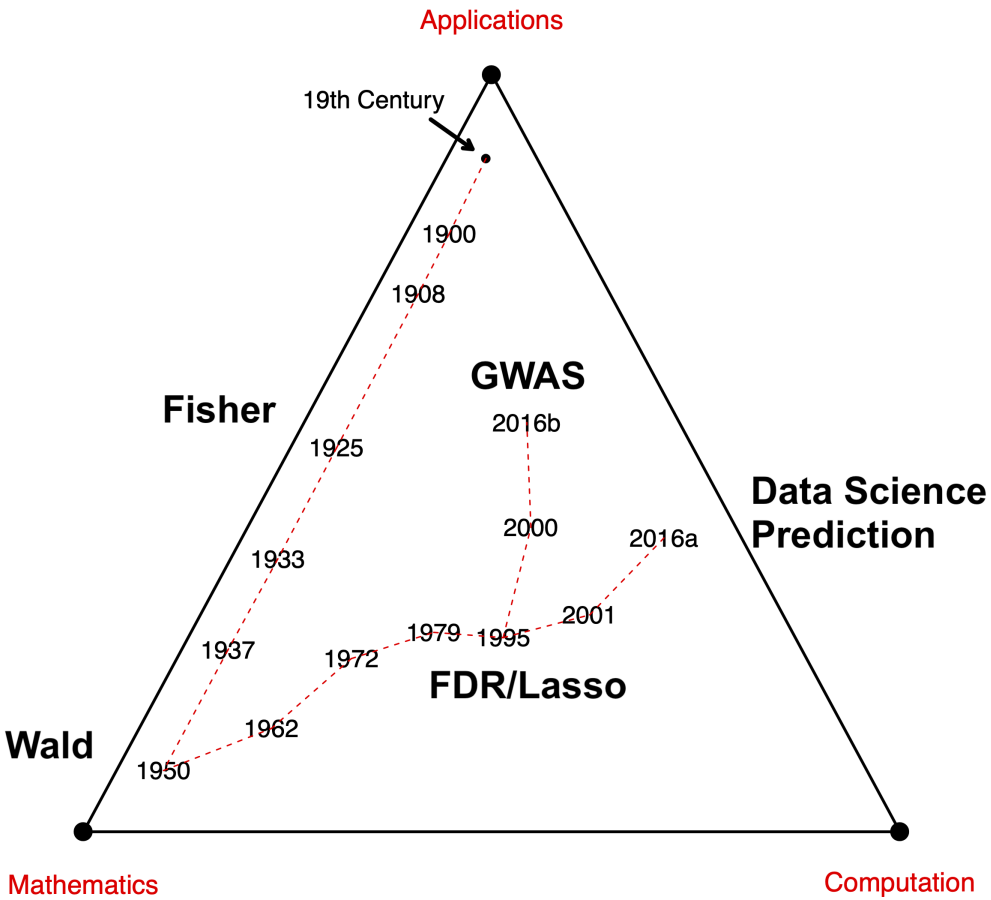
Source	Percentage
University	32%
Independent Research Fund DK	18%
Innovation Fund DK[1]	10%
Other DK	8%
Public international funding[2]	8%
EU	**5%**
VILLUM	16%
Novo Nordisk Foundation	3%

[1] Supports industrial PhD-students co-financed by companies.
[2] E.g. NIH and home country financed students.

Funding landscape in DK



In which direction is statistics moving?



From *Computer Age Statistical Inference* by Efron and Hastie.

From *Computer Age Statistical Inference*

Quoting Efron and Hastie:

> *A great amount of ingenuity and experimentation has gone into the development of modern prediction algorithms, with statisticians playing an important but not dominant role. There is no shortage of impressive success stories. In the absence of optimality criteria,..., the prediction community grades algorithmic excellence on performance.*
>
> *"Optimal" is the key word here. Before Fisher, statisticians didn't really understand estimation. The same can be said now about prediction.*

Open problems?

A MESSAGE FROM THE PRESIDENT

WHAT ARE THE OPEN PROBLEMS IN BAYESIAN STATISTICS?

- Michael I. Jordan -
ISBA President, 2011

que, I am afraid. In particu
sembled are a highly non-
are a set of people who h
being in the intersection
respected senior statistici
my email address book.

The question that I ask
iew as the top two or th

JOURNAL OF COMPUTATIONAL AND GRAPHICAL STATISTICS
2017, VOL. 26, NO. 4, 745–766
https://doi.org/10.1080/10618600.2017.1384734

50 Years of Data Scien

David Donoho

Department of Statistics, Stanford Uni

ABSTRACT
More than 50 years ago, John Tu
Analysis," he pointed to the exi
learning from data, or "data an
Leo Breiman independently on
classical domain of theoretical
presentation rather than statisti
inference. Cleveland and Wu ev

HDSR
HARVARD DATA SCIENCE REVIEW

A Telescopic, Microscopic, and Kaleidoscopic View of Data Science

HOME ABOUT MASTHEAD CONTENT FEATURES

As an open access platform of the Harvard Data Science Initiative, the *Harvard Data Science Review* features **foundational thinking, research milestones, educational innovations, and major applications,** with a primary emphasis on **reproducibility, replicability, and readability.** It aims to publish contents that help to define and shape data science as a scientifically rigorous and globally impactful multidisciplinary field based on the principled and purposed production, processing, parsing and analysis of data. By disseminating **inspiring, informative, and intriguing** articles and media materials, HDSR aspires to be a global forum on "***everything data science and data science for everyone.***"

Important topics to think about for statisticians

- The *Science about Data Science*; empirical research; meta-analysis of how the data analysis is done and how data scientists work.
- Models of data pipelines with well understood statistical properties; post-model selection inference.
- Causality; generalizability; transportability; robustness; explanability; fairness.
- The interplay between computational complexity and inferential power.
- Optimality; nonparametric lower bounds.

We cannot justify theoretical research focusing only on isolated steps like estimating a mathematically nice parameter in an apriori specified model.

Important roles for me as a statistician

- Researcher, collaborator, data analyst, software developer.
- Teacher, educator and supervisor.
  - How we teach applied statistics to non-statisticians is pivotal for how statistical methodology will be applied
  - Influencing which role statisticians will play in the future
  - Shaping the next generation of statisticians

A role I don't like to see a statistician in: Policeman.

Conclusion

My best advice today to PhD students is to be good

1. scientists
2. statisticians
3. and then mathematicians

in that order.

> **The most important ingredients are honesty, desire, clear thinking, confidence and hard work.** *If you aren't willing to work long, hard hours and sacrifice in pursuit of this goal, then you are not willing to pay the price and maybe you should move over and give someone else a chance.*
>
> The Real Final Exam, Donald S. Coffey

UCPH MATH has postdoc calls mid November, and PhD calls mid November and mid April, www.math.ku.dk/english/about/jobs/