



Universiteit
Leiden
The Netherlands

Topics in mathematical and applied statistics

Pas, S.L. van der

Citation

Pas, S. L. van der. (2017, February 28). *Topics in mathematical and applied statistics*. Retrieved from <https://hdl.handle.net/1887/46454>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/46454>

Note: To cite this publication please use the final published version (if applicable).

Stellingen
behorende bij het proefschrift
Topics in Mathematical and Applied Statistics
STÉPHANIE VAN DER PAS

1. De hoofijzer-*a-priori*-verdeling geeft de beste resultaten voor het *sparse-normal-means*-probleem wanneer de parameter τ van de orde $(p_n/n)\sqrt{\log(n/p_n)}$ is, waar n het aantal waarnemingen is en p_n het aantal signalen. Dit geldt zowel voor het schatten van de onderliggende parameters als voor het kwantificeren van de onzekerheid in de schatting. *Hoofdstukken 1 en 3.*
2. De hoofijzerverdeling heeft het het moeilijkst voor parameters die ongeveer ter grootte van $\sqrt{2 \log n}$ zijn. *Hoofdstukken 1 en 3.*
3. Het leveren van een minimax schatter van de onderliggende parameters in het *sparse normal means*-probleem is niet aan de hoofijzerverdeling voorbehouden. *Hoofdstuk 2.*
4. In het stochastisch blokmodel leiden een Beta-*a-priori*-verdeling op de kans op een connectie, een gegeneraliseerde-Bernoulli-*a-priori*-verdeling op de groepslabels en een Dirichlet-*a-priori*-verdeling op de proportie die elke groep van het geheel uitmaakt, tot een sterk consistente *maximum-a-posteriori*-schatter van de groepslabels, mits de verwachte graad minstens van orde $\log^2 n$ is, met n het aantal knopen in het netwerk. *Hoofdstuk 4.*
5. Voor de keuze tussen twee geneste modellen, waarbij de modellen leden van een exponentiële familie bevatten, is het switch-criterium consistent en minimax optimaal, op een factor $\log \log n$ na, met n het aantal waarnemingen. Als het kleinste model nuldimensionaal is, is het switch-criterium tevens bestend tegen optioneel stoppen. *Hoofdstuk 5.*
6. Uit statistisch oogpunt is het heel onpraktisch dat mensen twee heupen hebben. *Hoofdstuk 6.*
7. Voor wie statistiek alleen consumeert, zijn formules van ondergeschikt belang. Een kritische houding daarentegen is essentieel.
8. De problemen rondom hypothesetoetsen zijn onoplosbaar, omdat mensen verlangen naar zekerheid waar dat niet mogelijk is.
9. Namen van evenementen als *Pump Your Career* of *Biostat's Better Half* zijn goedbedoeld, maar werken vooroordelen in de hand.
10. In moderne termen kan *De Elementen* van Euclides het best vergeleken worden met een collegedictaat, terwijl de *Arithmetika* van Diophantus is geschreven als studieboek.

Propositions
accompanying the thesis
Topics in Mathematical and Applied Statistics
STÉPHANIE VAN DER PAS

1. The horseshoe prior performs best in the sparse normal means problem when the parameter τ is of the order $(p_n/n)\sqrt{\log(n/p_n)}$, where n is the number of observations and p_n the number of signals. This is the case when the aim is to recover the underlying parameters or to quantify uncertainty in the estimate. *Chapters 1 and 3.*
2. The horseshoe prior is challenged most by parameters that are approximately equal to $\sqrt{2 \log n}$. *Chapters 1 and 3.*
3. Yielding minimax estimators of the underlying parameters in the sparse normal means problem is not a unique property of the horseshoe prior. *Chapter 2.*
4. In the stochastic block model, a Beta prior on the connection probabilities, a generalized Bernoulli prior on the class labels and a Dirichlet prior on the class proportions lead to a strongly consistent maximum a posteriori estimator of the class labels, if the expected degree is at least of order $\log^2 n$, where n is the number of nodes in the network. *Chapter 4.*
5. The switch criterion is consistent and minimax optimal, up to a factor $\log \log n$, for selection between two nested models containing exponential family distributions, where n is the number of observations. Moreover, the switch criterion is robust to optional stopping, provided that the smallest model is a singleton. *Chapter 5.*
6. It is highly inconvenient, from a statistical point of view, that people have two hips. *Chapter 6.*
7. Formulas are not of great importance for consumers of statistics, but a critical attitude is essential.
8. The problems surrounding hypothesis testing cannot be solved, because people desire certainty when that cannot possibly be offered.
9. Event names like *Pump Your Career* or *Biostat's Better Balf* are created with the best of intentions, but only serve to confirm stereotypes.
10. In modern terminology, Euclid's *Elements* are best compared to lecture notes, while Diophantus' *Arithmetika* was written as a textbook.