

High-Dim Post-Select Inference



I am a statistics PhD student and am considering to work on high-dimensional post-selection inference. Can you give me some papers to read that will introduce the topic to me, as well the key theorems and ideas in this field?



 Memory updated

High-dimensional post-selection inference is a challenging and active area of research in statistics. Here are some foundational papers and key resources that will introduce you to the topic and its key theorems and ideas:

1. "Selective Inference: Tools for Post-Selection Inference" by Jonathan Taylor and Robert Tibshirani (2015):

- This paper provides a comprehensive introduction to selective inference, focusing on tools and techniques for making valid inferences after model selection.
- [Link to paper](#)

2. "Post-Selection Inference for Lasso with High-Dimensional Data" by Jason Lee, Dennis Sun, Yuekai Sun, and Jonathan Taylor (2016):

- This paper discusses methods for making valid inferences in the context of the Lasso regression, which is particularly relevant for high-dimensional data.
- [Link to paper](#)

3. "Selective Inference for Sparse High-Dimensional Models" by Richard J. Samworth and Michail Tsagris (2019):

- This review paper covers various methods for selective inference in high-dimensional settings, providing a broad overview of the field.
- [Link to paper](#)

4. "A Unified Framework for Post-Selection Inference" by Emmanuel Candès, Yingying Fan, Lucas Janson, and Jinchi Lv (2018):

- This paper proposes a unified framework for post-selection inference that is applicable to a wide range of model selection procedures.
- [Link to paper](#)

5. "Exact Post-Selection Inference with the Lasso" by Jonathan Taylor and Robert Tibshirani (2018):

- This paper provides a detailed treatment of exact post-selection inference methods for the Lasso, including theoretical results and practical algorithms.
- [Link to paper](#)

6. "Post-Selection Inference for Linear Models" by Chiara Sabatti, Lucas Janson, and Emmanuel Candès (2014):

- This paper introduces techniques for performing inference after model selection in linear regression models.
- [Link to paper](#)

Key Theorems and Ideas

1. Selective Inference Framework:

- The framework involves conditioning on the selection event and making inferences about the selected model's parameters. This conditioning ensures that the inferences account for the selection process and remain valid.

2. Polyhedral Lemma:

- A key result in selective inference, especially for Lasso, stating that the selection event can often be expressed as a set of linear inequalities (a polyhedron), which simplifies the calculation of post-selection probabilities.

3. Post-Selection p-Values:

- Methods to compute p-values that remain valid after the model selection process. These p-values adjust for the selection process and are typically larger than naive p-values, reflecting the additional uncertainty introduced by selection.

4. High-Dimensional Asymptotics:

- Asymptotic results that hold when the number of parameters is large compared to the sample size. These results often involve assumptions about sparsity or other structural properties of the high-dimensional parameter space.

5. Conditional Likelihood:

- Techniques based on the conditional likelihood of the observed data given the selection event, which allows for valid inference by appropriately adjusting for the selection process.

These papers and ideas should give you a solid foundation to start exploring high-dimensional post-selection inference. If you have specific questions or need further resources, feel free to ask!

Is this conversation helpful so far?

ChatGPT can make mistakes. Check important info.