Learning under constraints: inference, selection and validation with maximum entropy models

Giuffrida, F.

Stellingen

behorende bij het proefschrift

Learning Under Constraints: Inference, Selection and Validation with Maximum Entropy Models

1. In the context of NML description lengths, the enforcement of hard constraints comes at the cost of model complexity: microcanonical models are always more complex than their canonical counterparts.

[Chapter 2]

2. In complex systems, canonical and microcanonical distributions become inequivalent when an extensive number of constraints is enforced; this inequivalence is also visible at the model level, where it has direct inferential consequences.

[Squartini et al., Phys. Rev. Lett. **115**, 268701 (2015)]

3. In Bayesian inference, ensemble non-equivalence does not disappear: it is merely shifted from the choice between canonical and microcanonical models to the choice of prior, which therefore requires careful consideration.

[Peixoto et al, Phys. Rev. E **95**, 012317 (2017); Chapter 2 & 3]

4. E-values offer desirable theoretical properties over p-values, yet practical, ready-to-use implementations remain scarce, particularly outside mathematical statistics. Developing concrete e-value procedures for common inferential problems is essential to extend their reach.

[Grünwald et al., J. R. Stat. Soc. B **86**(5):1091–1128 (2024)]

5. Universal distributions underlie both MDL description lengths and optimal e-values. This connection opens the door to attaching measures of statistical significance to differences in description length.

[Chapter 2 & 3]

6. For e-value-based hypothesis testing, the appropriate construction depends critically on the data-collection regime.

[Chapter 4]

7. In null-model-based pattern validation, different null models lead to different results; rather than seeking a single optimal null model, results should be interpreted in light of the specific scientific question posed by each null model.

[Cimini et al., Commun. Phys. **5**, 76 (2022)]

8. In resting-state fMRI data, statistically validated topological structures align with information-theoretic measures of redundancy without any explicit assumption connecting the two, suggesting that topology and information theory reflect the same underlying organization.

[Chapter 5]

9. Genuine interdisciplinarity - as opposed to the mere application of tools from one field to another - requires curiosity, sustained effort, and patient dialogue with those who have spent years in a different discipline. The current pace of academic research, driven by publication pressure, leaves little room for any of these.

10. Human standards of intelligence and culture are poor measures of the richness of non-human minds and societies.

Francesca Giuffrida
Leiden, 10 september 2026