Constraint choice and model selection in the generalized maximum entropy principle (MEP)

Andrea Somazzi, Diego Garlaschelli

Scuola Normale Superiore di Pisa, IMT School for Advanced Studies Lucca

SCUOLA NORMALE SUPERIORE

IMT SCUOLA ALTI STUDI LUCCA

Introduction: A wide class of entropies can be used in the MEP

- Jos Uffink [1], generalizing the Shore and Johnson's axioms, proved that the functionals which are suitable to be used in the Maximum Entropy Principle (MEP) belong to a one-parameter family, which the Shannon entropy is a member of.
- Such functionals are monotonically increasing functions f of

$$U_q(P) = \left(\sum_G p^q(G)\right)^{1/(1-q)}$$

- For $q \to 1$, $U_q(P) \to$ Shannon entropy
- P. Jizba and J. Korbel discussed [2] that this generalized approach is suitable to study systems which do not respect standard hypothesis such as ergodicity, short-range interactions or exponential growth of the sample space: the resulting probability distributions take into account correlations that may not have been observed.
- Rényi entropy

$$S_R(P;q) = \log U_q(P)$$

- Constraint choice: mean vs normalized q-mean

$$\underline{\langle C\rangle_q} = \sum_G \frac{C(G)p^q(G)}{\sum_G p^q(G)}$$

Methods: Maximum entropy and maximum likelihood

- Constrained maximization of entropy through Lagrange multiplier technique.
- Parameter estimation through **Maximum likelihood** estimator: maximize the joint probability of the data in order to fix the parameters.
- Compare results for both constraint choices.

Mathematical formulation: Mean vs q-mean constraints

- Constrained maximization
  - $$\max_P \left[S_R(P;q) - \alpha'\left(\sum_G p(G) - 1\right) - \theta'\left(\langle C\rangle - C^*\right)\right]$$
  - $$\max_P \left[S_R(P;q) - \alpha\left(\sum_G p(G) - 1\right) - \theta\left(\underline{\langle C\rangle_q} - C^*\right)\right]$$
- Solutions
  - $$p(G) = \frac{1}{Z'}(1-(q-1)\,\hat{\theta}'\,C(G))^{\frac{1}{q-1}}$$
  - $$p(G) = \frac{1}{Z}(1-(1-q)\,\hat{\theta}\,C(G))^{\frac{1}{1-q}}$$
- Maximum likelihood
  - $$\sum_G C(G)p^{2-q}(G|\,\hat{\theta}'_{ML}) = \sum_G C(G)p^{1-q}(G|\,\hat{\theta}'_{ML})freq(G)$$
  - $$\sum_G C(G)p^{q}(G|\,\hat{\theta}_{ML}) = \sum_G f(G)p^{q-1}(G|\,\hat{\theta}_{ML})freq(G)$$

Discussion: Generaized mean is a better constraint in the MEP

- The solutions of the MEP show a power-law behavior: the mean could diverge, so it could be a bad constraint choice.
- If q is such that the distribution is normalizable, then all its q-moments converge.
- The q-mean allows us to characterize the power law for every possible value of the exponent.
- The q-mean, unlike the standard one, makes the Maximum entropy and Maximum likelihood consistent each other.
- Model selection, i.e. choice of the preferred parameters (q, θ) given the data, can be performed through the maximum likelihood approach.

Numerical simulations: Explanation

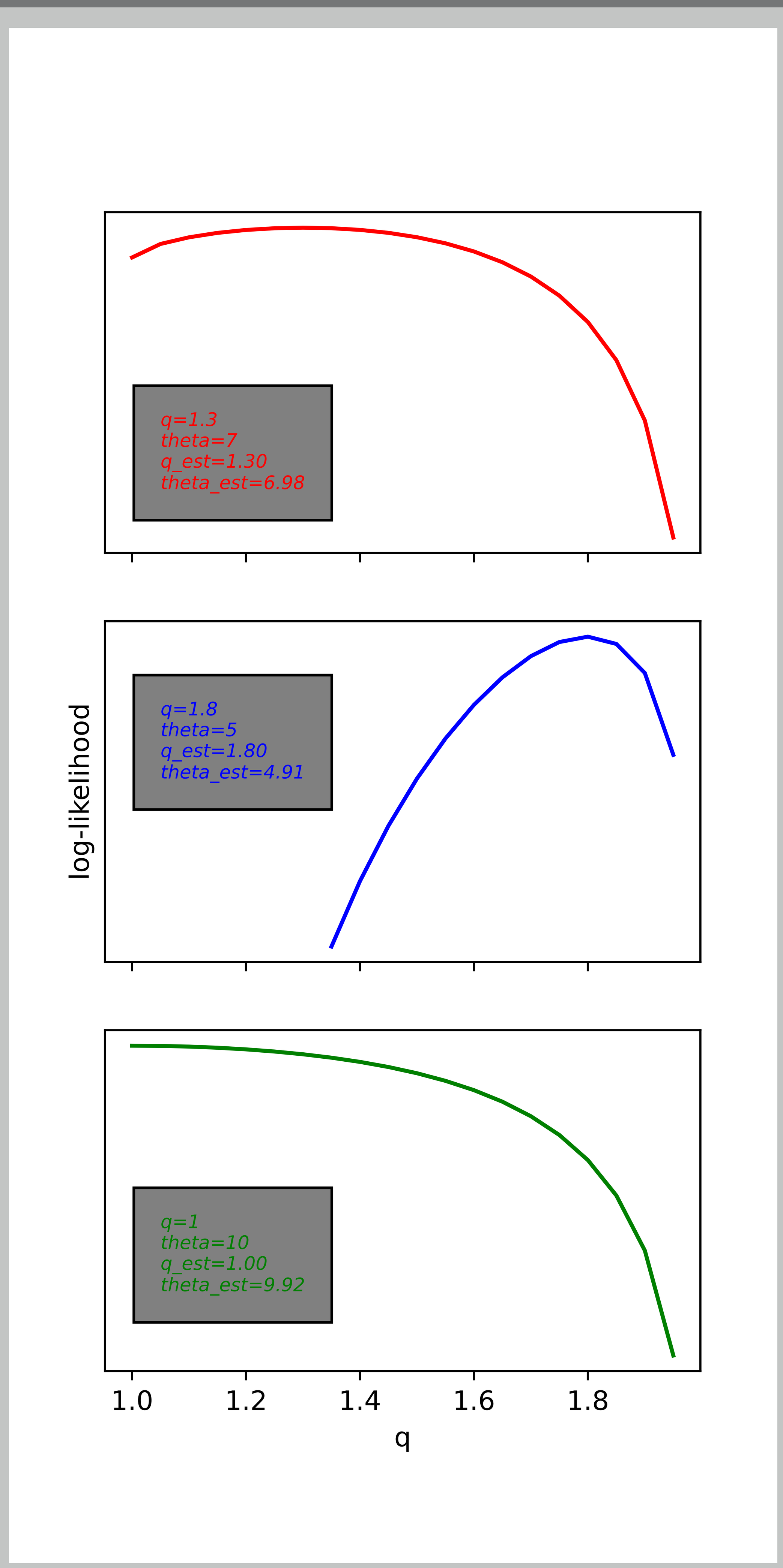
- Data sampled from p(G) in the simple case where $C(G) \in [0,+\infty)$ and $g(C) = 1$, for different values of q and θ ($q = 1$ being the exponential distribution)
- $P(C) = (2-q)\theta(1-(1-q)\theta C)^{\frac{1}{1-q}}$
- Parameters estimated through maximum likelihood model selection

Numerical results: Parameter estimation

- Number of points sampled $N = 10^5$...
- ...for each of three PDFs (r, b, g) with different parameters

Probability density	q	θ	q_{est}	θ_{est}
(r), 1st moment converge	1.3	7	1.30	6.98
(b), 1st moment diverge	1.8	5	1.80	4.91
(g), exponential PDF	1	10	1.00	9.92

Numerical results: Log-likelihood vs q plot



Log-likelihood vs q plot. The log-likelihood, for each q, is evaluated in the point $(q,\theta) = (q, \hat{\theta}_{ML}(q))$

References

[1] Jos Uffink.
Can the maximum entropy principle be explained as a consistency requirement?
Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 26(3):223–261, 1995.

[2] Petr Jizba and Jan Korbel.
Maximum entropy principle in statistical inference: Case for non-shannonian entropies.
Physical review letters, 122(12):120601, 2019.

Contact Information

- Linkedin: Andrea Somazzi
- Email: andrea.somazzi@snsit

andrea.somazzi@sns.it