21-317 FUNCTIONAL INEQUALITIES: FROM ANALYSIS TO PROBABILITY
ROUND I ALTERNATE TOPIC PACKET 10
ENTROPY, LIKELIHOOD, AND GIBBS EQUILIBRIUM

MATTHEW ROSENZWEIG

Abstract. **Organizing question:** Why does relative entropy appear in both statistical inference and equilibrium statistical mechanics? This packet fixes the mathematical scope and normalization, identifies the required proof content, and provides a curated reading route. It is intended both for informed preference ranking and for preparation after assignment.

At a glance

Status	Alternate Round I topic
Setting	Finite probability spaces
Primary tools	Jensen, relative entropy, likelihood identities, variational principles
Background	Finite probability; logarithms; elementary optimization
Relative difficulty	2/4
Principal perspective	Statistical/information-theoretic and physical

1. Why this topic may appeal to you

This alternate is the most conceptual bridge between statistics and physics. It is not a broad historical survey: the talk is organized around two exact identities showing that the same relative entropy measures likelihood loss and free-energy excess.

2. Statistical identity

Let $x_1, \dots, x_n$ lie in a finite alphabet, let $\widehat{p}_n$ be the empirical law, and let q be positive on the observed symbols. Prove

$$\frac{1}{n} \log \prod_{j=1}^{n} q(x_j) = \sum_x \widehat{p}_n(x) \log q(x) = -H(\widehat{p}_n) - D(\widehat{p}_n \| q),$$

where $H(p) = -\sum_x p(x) \log p(x)$. Conclude that the maximum-likelihood estimate over all probability laws is $q = \widehat{p}_n$. Also prove

$$\mathbf{E}_p \left[\log \frac{q(X)}{p(X)} \right] = -D(p \| q).$$

Date: Fall 2026.

3. Physical identity

Let $E(x)$ be an energy and $\beta > 0$. Define

$$\mu_\beta(x) = \frac{e^{-\beta E(x)}}{Z_\beta}, \qquad \mathcal{F}_\beta(\nu) = \sum_x E(x)\nu(x) + \frac{1}{\beta}\sum_x \nu(x)\log\nu(x).$$

Prove

$$\mathcal{F}_\beta(\nu) - \mathcal{F}_\beta(\mu_\beta) = \frac{1}{\beta}D(\nu\|\mu_\beta),$$

and conclude that μ_β uniquely minimizes free energy.

4. Required proof content

Prove nonnegativity and the equality case for relative entropy using Jensen or the log-sum inequality. Derive both identities line by line. Explain why reversing the two arguments of relative entropy changes the interpretation.

5. Optional extension, exercises, and pitfalls

If time permits, use Stirling's formula to explain why the number of strings of empirical type p grows like $e^{nH(p)}$. Exercises may treat a two-level Gibbs system, Bernoulli maximum likelihood, or a three-state free-energy calculation. State that Boltzmann's constant has been set to one. Do not call relative entropy a metric, and distinguish thermodynamic entropy from relative entropy.

6. Suggested reading route and attribution

(1) D. Chafaï, *Mini-course on logarithmic Sobolev inequalities*, Sections 1–2, approximately pp. 1–3, for entropy in combinatorics, probability and statistics and the Boltzmann–Gibbs/free-energy identities (public mini-course notes).
(2) T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley, 2006, Chapter 2 for entropy and relative entropy and Chapter 11 for the method of types.
(3) R. van Handel, *Probability in High Dimension*, Section 4.1 for variational and information-theoretic context (public course notes).

7. Before the required consultation

Bring a completed one-page preparation document containing: the exact theorem statements; a dependency map of the proof; the portion you will prove in full; planned examples; two exercises with solutions; exact source locations; anticipated questions; and any point you do not yet understand. The consultation is for resolving mathematical and expository issues, not for choosing a topic from scratch.

8. AI-use disclosure

The preparation document must disclose any use of generative artificial-intelligence tools. Identify the tool, the task for which it was used, material incorporated into the presentation, and how mathematical assertions and references were independently verified. You remain responsible for every statement in the talk.

This is an instructor-authored assignment guide, not a substitute for reading the listed sources. All normalizations in the talk should agree with this packet unless a change is explicitly approved in advance.