TUTORIAL #2: STATISTICAL PHYSICS AND LARGE DEVIATIONS

CHRISTIAN BORGS AND HENRY COHN

We'll begin with some definitions from statistical physics. Let G be a graph. We will randomly color the vertices of G with q colors; i.e., we will consider random maps $\phi \colon V(G) \to [q] := \{1, 2, \dots, q\}$. We allow for all possible maps, not just proper colorings, and call such a map a *spin configuration*. To make the model nontrivial, different spin configurations get different weights, based on a symmetric $q \times q$ matrix J with entries $J_{ij} \in \mathbb{R}$ called the *coupling matrix*. Given G and J, a map $\phi \colon V(G) \to [q]$ then gets an *energy*

$$E_\phi(G, J) = -\frac{1}{|E(G)|} \sum_{\substack{u,v \in V(G) \\ (u,v) \in E(G)}} J_{\phi(u)\phi(v)}.$$

Given a vector $\mathbf{a} = (a_1, \dots, a_q)$ of nonnegative real numbers adding up to 1 (we denote the set of these vectors by $\triangle_q$), we consider configurations ϕ such that the (weighted) fraction of vertices mapped onto a particular color $i \in [q]$ is near to a_i. More precisely, we consider configurations ϕ in

$$\Omega_{\mathbf{a},\varepsilon}(G) = \left\{ \phi \colon [q] \to V(G) : \left| \frac{\#\{v \in V(G) : \phi(v) = i\}}{|V(G)|} - a_i \right| \le \varepsilon \text{ for all } i \in [q] \right\}.$$

On $\Omega_{\mathbf{a},\varepsilon}(G)$ we then define a probability distribution

$$\mu_{G,J}^{(\mathbf{a},\varepsilon)}(\phi) = \frac{1}{Z_{G,J}^{(\mathbf{a},\varepsilon)}} e^{-|V(G)| E_\phi(G,J)},$$

where $Z_{G,J}^{(\mathbf{a},\varepsilon)}$ is the normalization factor

$$Z_{G,J}^{(\mathbf{a},\varepsilon)} = \sum_{\phi \in \Omega_{\mathbf{a},\varepsilon}(G)} e^{-|V(G)| E_\phi(G,J)}.$$

The distribution $\mu_{G,J}^{(\mathbf{a},\varepsilon)}$ is usually called the *microcanonical Gibbs distribution of the model J on G*, and $Z_{G,J}^{(\mathbf{a},\varepsilon)}$ is called the *microcanonical partition function*.

We will not analyze the particular properties of the distribution $\mu_{G,J}^{(\mathbf{a},\varepsilon)}$, but we will be interested in the normalization factor, or more precisely its normalized logarithm

$$F_{\mathbf{a},\varepsilon}(G, J) = -\frac{1}{|V(G)|} \log Z_{G,J}^{(\mathbf{a},\varepsilon)},$$

which is called the *microcanonical free energy*. We will also be interested in the dominant term contributing to $Z_{G,J}^{(\mathbf{a},\varepsilon)}$, or more precisely its normalized logarithm, the *microcanonical ground state energy*

$$E_{\mathbf{a},\varepsilon}(G, J) = \min_{\phi \in \Omega_{\mathbf{a},\varepsilon}(G)} E_\phi(G, J).$$

Let $(G_n)_{n\geq 0}$ be a sequence of weighted graphs. We say that $(G_n)_{n\geq 0}$ has *convergent microcanonical ground state energies* if the limit

$$E_{\mathbf{a}}(J) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} E_{\mathbf{a},\varepsilon}(G, J) = \lim_{\varepsilon\to 0} \liminf_{n\to\infty} E_{\mathbf{a},\varepsilon}(G, J)$$

exists for all $q \in \mathbb{N}$, $\mathbf{a} \in \triangle_q$, and symmetric $J \in \mathbb{R}^{q\times q}$, and $(G_n)_{n\geq 0}$ has *convergent microcanonical free energies* if the limit

$$F_{\mathbf{a}}(J) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} F_{\mathbf{a},\varepsilon}(G, J) = \lim_{\varepsilon\to 0} \liminf_{n\to\infty} F_{\mathbf{a},\varepsilon}(G, J)$$

exists for all $q \in \mathbb{N}$, $\mathbf{a} \in \triangle_q$, and symmetric $J \in \mathbb{R}^{q\times q}$.

(1) As motivation for the Gibbs distribution, prove the following characterization. Let $E_1, \dots, E_n$ be real numbers called *energies*, and let $\overline{E}$ satisfy $\min_i E_i < \overline{E} < \max_i E_i$. Suppose the probability distribution $p_1, \dots, p_n$ on $1, \dots, n$ maximizes the entropy

$$\sum_i -p_i \log p_i$$

(with $0 \log 0$ interpreted as 0) subject to $\sum_i p_i E_i = \overline{E}$. Prove that there exists a constant β such that

$$p_i = e^{-\beta E_i}/Z$$

for all i, where $Z = \sum_i e^{-\beta E_i}$. In other words, if we maximize entropy subject to constraining the expected energy of a system, we get a Gibbs distribution.

(2) Compute the limiting microcanonical ground state energies and free energies for the Ising model on the complete graph K_n as $n \to \infty$. For this model, $q = 2$ and $J_{ij} = (-1)^{i+j}$.

(3) Let X be a random variable whose moment generating function $M(t) = E(e^{tX})$ exists for all t. The *cumulant generating function* for X is $\log M(t)$. Prove that it is a convex and C^∞ function of t.

(4) Let X be a random variable whose moment generating function $M(t) = E(e^{tX})$ exists for all t. Let $f(t) = \log M(t)$ be the cumulant generating function. The n-th *cumulant* $\kappa_n(X)$ is $f^{(n)}(0)$; in other words, the Taylor series of $f(t)$ is

$$\sum_{n=1}^{\infty} \kappa_n(X) \frac{t^n}{n!}.$$

Prove that $\kappa_1(X) = E(X)$ and $\kappa_2(X) = \mathrm{Var}(X)$. Prove that if X and Y are independent, then $\kappa_n(X+Y) = \kappa_n(X) + \kappa_n(Y)$ for all n (a generalization the additivity of variance for independent random variables).

(5) Suppose $f\colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is a convex function (i.e., $\{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$ is a convex set) that is not identically equal to ∞. Its *Legendre transform* $f^*\colon \mathbb{R} \to \mathbb{R}\cup\{\infty\}$ is defined by

$$f^*(u) = \sup_v \big(uv - f(v)\big).$$

Prove that f^* is convex, that $f \leq g$ implies $g^* \leq f^*$, and that $f^{**} \leq f$. Under the assumption that

$$f(v) = \sup\{\ell(v) : \ell \text{ is an affine function satisfying } \ell \leq f \text{ everywhere}\},$$

prove that $f^{**} = f$. (This assumption is very mild; for example, it is equivalent to lower semicontinuity.)

[For intuition regarding the Legendre transform, note that the tangent line to f at v is $x \mapsto f'(v)(x-v)+f(v)$ when f is differentiable. If we choose v so $f'(v) = u$, then the tangent line is $x \mapsto ux - f^*(u)$. Thus, taking the Legendre transform amounts to describing the tangent lines of a convex function in terms of their slope, and we can reconstruct the function as the envelope of those lines.]

(6) Let X be a random variable whose moment generating function $M(t) = E(e^{tX})$ exists for all t, and let $X_1, X_2, \dots$ be i.i.d. copies of X. Prove that if $x \geq E(X)$, then

$$\mathrm{Prob}\left(\frac{\sum_{i=1}^n X_i}{n} \geq x\right) \leq e^{-nf^*(x)},$$

where f^* is the Legendre transform of the cumulant generating function $f(y) = \log M(y)$.

In fact, this bound is essentially sharp, in the sense that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathrm{Prob}\left(\frac{\sum_{i=1}^n X_i}{n} \geq x\right) = -f^*(x),$$

but you needn't prove that.