Applications of Stochastics — Exercise sheet 8

Gábor Pete
http://www.math.bme.hu/~gabor

April 24, 2018

First, some exercises on large deviations.

▷ **Exercise 1.** Let $\xi_1, \xi_2, \dots$ be i.i.d. random variables, and let $S_n = \xi_1 + \dots + \xi_n$. Show that

$$\lim_{n\to\infty} \frac{-\log \mathbf{P}[\, S_n > \alpha n \,]}{n} = \overline{I}_\xi(\alpha) \in [0,\infty]$$

and

$$\lim_{n\to\infty} \frac{-\log \mathbf{P}[\, S_n < \alpha n \,]}{n} = \underline{I}_\xi(\alpha) \in [0,\infty]$$

both exist for any $\alpha \in \mathbb{R}$. (Hint: use Fekete's subadditive convergence lemma.)

▷ **Exercise 2.** Prove

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} \frac{-\log \mathbf{P}[\, \mathsf{Binom}(n,p)/n \in (\alpha,\alpha+\epsilon) \,]}{n} = \alpha \log\frac{\alpha}{p} + (1-\alpha)\log\frac{1-\alpha}{1-p}$$

in two ways:

(a) Calculate by hands, using Stirling's formula.

(b) Just apply Cramér's large deviations theorem.

Compute from this the functions $\overline{I}_{\mathsf{Ber}(p)}(\alpha)$ and $\underline{I}_{\mathsf{Ber}(p)}(\alpha)$ of the previous exercise.

The formula in the previous exercise is the tip of an iceberg, a close relationship between large deviations and entropy theory. We will not discuss this relationship here, but at least, here is the definition and some basic properties of the **entropy** of a discrete random variable:

$$H(X) := -\sum_{x\in\Omega} \mathbf{P}[\, X = x \,] \log \mathbf{P}[\, X = x \,] \,.$$

If X and Y are defined on the same probability space, then $H(X,Y)$ is just the entropy of the variable (X,Y), while the **conditional entropy** $H(X \mid Y)$ is defined as the Y-average of the entropies of the conditional distributions $X \mid Y = y$:

$$H(X \mid Y) := \sum_{y\in\Omega} \left(-\sum_{x\in\Omega} \mathbf{P}[\, X = x \mid Y = y \,] \log \mathbf{P}[\, X = x \mid Y = y \,] \right) \mathbf{P}[\, Y = y \,] \,.$$

▷ **Exercise 3.**

(a) Show that if the probability space is finite, $|\Omega| = n$, then $H(X) \le \log n$, with equality iff X is uniform on Ω. (Hint: use the concavity of $-x\log x$ on $x \in [0,1]$.)

(b) Show that $H(X \mid Y) \le H(X)$, with equality iff X and Y are independent.

(c) Show that $H(X \mid Y) = H(X,Y) - H(Y)$. Deduce that $H(X,Y) \le H(X) + H(Y)$, with equality iff X and Y are independent.

As in class, the **Ising model** on a finite graph $G(V,E)$ is the random spin configuration $\sigma : V \longrightarrow \{\pm 1\}$ defined as follows. Given an external magnetic field $h \in \mathbb{R}$, the Hamiltonian is

$$H_h(\sigma) := -h \sum_{x\in V(G)} \sigma(x) - \sum_{(x,y)\in E(G)} \sigma(x)\sigma(y)\,,$$

and then the measure, at inverse temperature $\beta = 1/T \geq 0$, is

$$\mathbf{P}_{\beta,h}[\sigma] := \frac{\exp(-\beta H_h(\sigma))}{Z_{\beta,h}}\,, \quad \text{where} \quad Z_{\beta,h} := \sum_{\sigma} \exp(-\beta H_h(\sigma))\,.$$

▷ **Exercise 4.** The partition function $Z_{\beta,h}$ contains a lot of information about the model:

(a) Show that the **expected total energy** is

$$\mathbf{E}_{\beta,h}[\,H\,] = -\frac{\partial}{\partial\beta} \ln Z_{\beta,h}\,, \text{ with variance } \mathrm{Var}_{\beta,h}[H] = -\frac{\partial}{\partial\beta}\mathbf{E}_{\beta,h}[\,H\,]\,.$$

(b) The **average free energy** or **pressure** is defined by $f(\beta,h) := (\beta|V|)^{-1} \ln Z_{\beta,h}$. Show that for the **average total magnetization** $M(\sigma) := |V|^{-1}\sum_{x\in V}\sigma(x)$, we have

$$m(\beta,h) := \mathbf{E}_{\beta,h}[\,M\,] = \frac{\partial}{\partial h} f(\beta,h)\,.$$

(c) The **susceptibility** of the total magnetization to a change in the external magnetic field is

$$\chi(\beta,h) := \frac{1}{\beta}\frac{\partial}{\partial h} m(\beta,h) = \frac{1}{\beta}\frac{\partial^2}{\partial h^2} f(\beta,h)\,.$$

Relate this quantity to $\mathrm{Var}_{\beta,h}[M]$. Deduce that $f(\beta,h)$ is convex in h.

The **Curie-Weiss model** is the Ising model on the complete graph K_n, with edge weights $1/n$, so that the Hamiltonian is

$$H_{n,h}(\sigma) := -h\sum_{i=1}^{n}\sigma_i - \frac{1}{2n}\sum_{i,j=1}^{n}\sigma_i\sigma_j\,.$$

(The 1/2 factor is to make up for having each pair $\{i,j\}$ with $i \neq j$ twice in the sum. The appearance of the terms $i = j$ causes just a shift of H by a constant, which is not visible in $\mathbf{P}_{\beta,h}$.) In terms of the average magnetization $M(\sigma) = \sum_i \sigma_i/n$, note that we can write

$$H_{n,h}(\sigma) = -\big(hM(\sigma) + M(\sigma)^2/2\big)n\,,$$

and the number of σ's with $M(\sigma) = x \in \{-1, \frac{-n+2}{n}, \dots, \frac{n-2}{n}, 1\}$ is $\binom{n}{n(1+x)/2}$. Thus,

$$Z_{n,\beta,h} = \sum_x c_{n,\beta,h}(x)\,, \quad \text{where} \quad c_{n,\beta,h}(x) := \binom{n}{n(1+x)/2} \exp\Big(\beta n\big(hx + x^2/2\big)\Big)\,.$$

▷ **Exercise 5.**

(a) Show that $f(\beta,h) := \lim_{n\to\infty} f_n(\beta,h) = \lim_{n\to\infty} \frac{\max_x \ln c_{n,\beta,h}(x)}{\beta n}$.

(b) Similarly to Exercise 2 (a), show that $\ln c_{n,\beta,h}(x) = n\big(\beta hx - \Phi_\beta(x)\big) + o(n)$, where

$$\Phi_\beta(x) = \frac{1-x}{2}\ln\frac{1-x}{2} + \frac{1+x}{2}\ln\frac{1+x}{2} - \frac{\beta x^2}{2} \qquad \text{for } x\in[-1,1]\,.$$

(c) Sketch the curves $\Phi_\beta(x)$ and $\Phi'_\beta(x)$ on $x \in [-1,1]$, for some parameters $\beta < 1$, $\beta = 1$, and $\beta > 1$.

(d) By choosing the appropriate root $x = x_0(\beta, h)$ of $\Phi'_\beta(x) = \beta h$, find $\max_x \ln c_{n,\beta,h}(x)$. Note that part (a) gives

$$\frac{\partial}{\partial h} f(\beta, h) = \frac{\partial}{\partial h}\left(h x_0(\beta, h) - \frac{\Phi_\beta(x_0(\beta, h))}{\beta}\right) = x_0(\beta, h)\,.$$

(e) By Exercise 4 (b), $m_n(\beta, h) = \frac{\partial}{\partial h} f_n(\beta, h)$. Assuming that $m(\beta, h) := \lim_{n\to\infty} m_n(\beta, h) = \frac{\partial}{\partial h} f(\beta, h)$ holds for $h \neq 0$ (which is indeed the case), deduce from the above that

$$\lim_{h\to 0+} m(\beta, h) > 0 \qquad \text{and} \qquad \lim_{h\to 0-} m(\beta, h) < 0 \qquad \text{for } \beta > 1\,,$$

while the limits equal 0 for $\beta \leq 1$. Hence $m(\beta, h)$ is discontinuous at $h = 0$ iff $\beta > 1$.

(f) Show that

$$\frac{1}{\beta}\frac{\partial^2}{\partial h^2} f(\beta, h) = \frac{1}{\beta}\frac{\partial}{\partial h} x_0(\beta, h) = \frac{1 - x_0(\beta, h)^2}{1 - \beta(1 - x_0(\beta, h)^2)}\,.$$

For $\beta = 1$, deduce that $\frac{\partial}{\partial h} x_0(\beta, h) = \infty$. That is, $m(1, h)$ is continuous but not analytic at $h = 0$. Assuming that the limiting susceptibility $\chi(\beta, h) := \lim_{n\to\infty} \chi_n(\beta, h)$ equals $\frac{1}{\beta}\frac{\partial^2}{\partial h^2} f(\beta, h)$, we get that the limiting susceptibility is $\chi(1, 0) = \infty$. What does that mean for the variance of the average magnetization?

(g)* Show that $\frac{\partial}{\partial h} x_0(\beta, 0+) < \infty$ for $\beta > 1$, so that the limiting susceptibility is finite.