Computing G-star discrepancy

Daniel Miller

August 25, 2016

The basic idea is this. Suppose we have two probability measures μ and ν on $[0,\pi]$. These will be given by explicitly computable cumulative distribution functions cdf_μ, cdf_ν. For example, we might have

$$\mathrm{cdf}_\mu(x) = \frac{x - \sin(x)\cos(x)}{\pi}$$
$$\mathrm{cdf}_\nu(x) = \frac{\#\{n : x_n < x\}}{N},$$

where $(x_1, \ldots, x_N)$ is a sequence of points in $[0,\pi]$. The *star discrepancy* of ν with respect to μ is just the supremum of the difference between their CDFs:

$$\mathrm{disc}(\mu,\nu) = \sup_{x\in[0,\pi]} |\mathrm{cdf}_\mu(x) - \mathrm{cdf}_\nu(x)|.$$

The question I am interested in is: when cdf_μ is an explicit smooth function and cdf_ν is an explicit step function as above, how can we compute $\mathrm{disc}(\mu,\nu)$ quickly? Better yet, for $n \leqslant N$, write

$$\mathrm{cdf}_{\nu_n}(x) = \frac{\#\{i \leqslant n : x_i < x\}}{n}.$$

Is there a fast way of computing the sequence

$$\mathrm{disc}(\nu_1,\mu), \ldots, \mathrm{cdf}(\nu_N,\mu)?$$

A little bit of motivation. Sort the list of x_i's so that $x_1 \leqslant \cdots \leqslant x_N$. Then the star discrepancy can be proved to be equal to:

$$\max\left\{ \left|\frac{1}{N} - \mathrm{cdf}_\mu(x_1)\right|, \left|\frac{2}{N} - \mathrm{cdf}_\mu(x_1)\right|, \ldots, \left|\frac{N-1}{N} - \mathrm{cdf}_\mu(x_{N-1})\right|, |1 - \mathrm{cdf}_\mu(x_{N-1})| \right\}.$$

In particular, we need to maximize the function

$$(x,i) \mapsto \left|\frac{i}{N} - \mathrm{cdf}_\mu(x)\right|$$

over the set $\{(x_1,1),\ldots,(x_N,N)\}$. If $\mathrm{cdf}_\mu(x) = x$ is convex, then only need to maximize and minimize the function $(x,i) \mapsto \frac{i}{N} - x$ over a finite set. This comes down to finding the convex hull of the set $\{(x_i,i)\}_{i=1}^N$