Sample Energy: Solution

Let X be the continuous uniform distribution on $[0,1]$, with

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ t & 0 \leq t \leq 1 \\ 1 & 1 \leq t \end{cases}$$

Let Y be the discrete uniform distribution on $\{u, v, w\}$, with

$$F_Y(t) = \begin{cases} 0 & t < u \\ 1/3 & u \leq t < v \\ 2/3 & v \leq t < w \\ 1 & w \leq t \end{cases}$$

The integrand of the squared energy distance $\int_{-\infty}^{\infty} (F_X(t) - F_Y(t))^2 \, \mathrm{d}t$ is 0 outside of the interval $[0,1]$, so we can restrict the domain of integration, then split $[0,1]$ into four intervals and thus the integral into four:

$$\int_0^u t^2 \, \mathrm{d}t + \int_u^v (t - \tfrac{1}{3})^2 \, \mathrm{d}t + \int_v^w (t - \tfrac{2}{3})^2 \, \mathrm{d}t + \int_w^1 (t-1)^2 \, \mathrm{d}t$$

Instead of evaluating the integrals right away, let's expand the quadratics, then collect the t^2 terms from the four integrals into just one:

$$\int_0^1 t^2 \, \mathrm{d}t + \int_u^v -\tfrac{2}{3}t + \tfrac{1}{9} \, \mathrm{d}t + \int_v^w -\tfrac{4}{3}t + \tfrac{4}{9} \, \mathrm{d}t + \int_w^1 -2t + 1 \, \mathrm{d}t$$

Evaluating the integrals we then get a number of terms:

$$\begin{aligned} = \tfrac{1}{3} \; &- \; \tfrac{1}{3}(v^2 - u^2) \; - \; \tfrac{2}{3}(w^2 - v^2) \; - \; \tfrac{3}{3}(1 - w^2) \\ &+ \; \tfrac{1}{9}(v - u) \; + \; \tfrac{4}{9}(w - v) \; + \; \tfrac{9}{9}(1 - w) \end{aligned}$$

Combining like terms simplifies this to

$$\left(\tfrac{1}{3}u^2 + \tfrac{1}{3}v^2 + \tfrac{1}{3}w^2\right) - \left(\tfrac{1}{9}u + \tfrac{3}{9}v + \tfrac{5}{9}w\right) + \left(\tfrac{1}{3} - 1 + 1\right)$$

Way may complete the square to turn this into

$$\tfrac{1}{3}\left[(u - \tfrac{1}{6})^2 - \tfrac{1}{6^2} + (v - \tfrac{1}{2})^2 - \tfrac{1}{2^2} + (w - \tfrac{5}{6})^2 - \tfrac{5^2}{6^2}\right] + \tfrac{1}{3}$$

which simplifies to

$$\tfrac{1}{3}\Big[(u-\tfrac{1}{6})^2+(v-\tfrac{1}{2})^2+(w-\tfrac{5}{6})^2\Big]+\tfrac{1}{108}.$$

Thus, when $\{u,v,w\}=\{\frac{1}{6},\frac{1}{2},\frac{5}{6}\}$ the energy distance is minimized to $\frac{1}{\sqrt{108}}$.

In general the squared energy distance between the discrete uniform distribution on $\{u_1,\cdots,u_n\}$ and the continuous uniform distribution on $[0,1]$ is

$$E=\frac{1}{12n^2}+\frac{1}{n}\sum_{k=1}^{n}\left(u_k-\frac{2k-1}{2n}\right)^2.$$

In general, the squared energy distance $E=\mathrm{d}(X,U)^2$ between a random variable X and a discrete uniform random variable on $U=\{u_1,\cdots,u_n\}$ is

$$E=\sum_{k=0}^{n}\int_{u_n}^{u_{n+1}}\left(F_X(t)-\frac{k}{n}\right)^2\mathrm{d}t$$

where $u_1<\cdots<u_n$ and $u_0:=0, u_{n+1}:=1$. The partial derivatives are

$$\begin{aligned}\frac{\partial E}{\partial u_k} &\quad -\Big(F_X(u_k)-\frac{k}{n}\Big)^2+\Big(F_X(u_k)-\frac{k-1}{n}\Big)^2\\ &=\frac{1}{n}\Big(2F_X(u_k)-\frac{2k-1}{n}\Big)\end{aligned}$$

Setting $\nabla E=0$ and solving, we find the minimum $E_{\min}$ is attained when $U=\{F_X^{-1}(\frac{2k-1}{2n})\mid 1\le k\le n\}$ (intuitively, the narrower U is, the less it would approximate a continuous variable X, suggesting we needn't check boundary cases). Integrating ∇E from U to $V=\{v_1,\cdots,v_n\}$ yields

$$E=E_{\min}+\frac{2}{n}\int_{F_X^{-1}(\frac{2k-1}{2n})}^{v_k}(v_k-u)f_X(u)\,\mathrm{d}u.$$