8. EXERCISE SHEET *Bayesian inference and data assimilation*

Exercise 1 Consider a Markov transition kernel $\pi(z'|z)$ and two PDFs π_1 and π_2. Define

$$\hat{\pi}_i(z') = \int_{\mathbb{R}} \pi(z'|z)\,\pi_i(z)\,\mathrm{d}z, \qquad i = 1, 2.$$

and show that

$$d_{\mathrm{TV}}(\hat{\pi}_1, \hat{\pi}_2) \leq d_{\mathrm{TV}}(\pi_1, \pi_2)$$

with the TV-distance $d_{\mathrm{TV}}(\nu, \mu)$ defined by

$$d_{\mathrm{TV}}(\nu, \mu) = \frac{1}{2} \sup_{|f|_\infty \leq 1} |\mathbb{E}_\mu[f] - \mathbb{E}_\nu[f]|.$$

Proof this directly using the definition given here on the sheet. Do not use the definition

$$d_{\mathrm{TV}}(\nu, \mu) = \frac{1}{2} \int |\mu(x) - \nu(x)|\mathrm{d}x.$$

Hint: The main step consists of finding another function g, that also has $|g|_\infty \leq 1$, s.t.

$$|\mathbb{E}_{\hat{\pi}_1}[f] - \mathbb{E}_{\hat{\pi}_2}[f]| \leq |\mathbb{E}_{\pi_1}[g] - \mathbb{E}_{\pi_2}[g]|.$$

for any f with $|f|_\infty = \sup_x |f(x)| \leq 1$.

Exercise 2 (6 Points). Let X be a random variable and F_X its CDF and π_X its PDF. Suppose we are given a guess F of the CDF of X. We define $T_x(y) = 1$ if $x \leq y$ and 0 otherwise. T can be interpreted as the CDF of the distribution that places unit mass at x. The continuous ranked probability score is defined as

$$S_{\mathrm{crps}}(F, x) = \int_{-\infty}^{\infty} (F(y) - T_x(y))^2 \mathrm{d}y.$$

Prove that,

$$\begin{aligned}
\mathbb{E}\left[S_{\mathrm{crps}}(F, X)\right] &= \int_{-\infty}^{\infty} S_{\mathrm{crps}}(F, x)\,\pi_X(x)\mathrm{d}x, \\
&= \int_{-\infty}^{\infty} (F(x) - F_X(x))^2 \mathrm{d}x \; + \int_{-\infty}^{\infty} F_X(x)(1 - F_X(x))\mathrm{d}x,
\end{aligned}$$

using the definition of the continuous ranked probability score.
Hint: First show that

$$\int_{-\infty}^{\infty} F(x) T_y(x) \pi_X(y) \mathrm{d}y = F(x) F_X(x).$$

Exercise 3. We now use the method developed in Exercise 3 of the last sheet to sample a more complicated distribution. Assume you are measuring the size of members of some animal species. The distribution of the height is given as

$$\pi(x) = \frac{1}{C} \exp(-V(x)),$$

with

$$V(x) = ((x-4)^2-2)^2,$$

and C chosen such that $\int_{-\infty}^{\infty} \pi_R(x)\mathrm{dx} = 1$. This distribution is not one of the basic distributions like a Gaussian. Nevertheless, we want to answer some questions about it. We do this, by generating samples from it. This can be done analogously to Exercise 3 from Sheet 7. The following SDE

$$\frac{\mathrm{d}}{\mathrm{d}t}X_t = -\nabla V(x) + \sqrt{2}\frac{\mathrm{d}B_t}{\mathrm{d}t}.$$

has invariant distribution π.

(i) Plot π for $C = 1$.

(ii) Repeat Exercise (*ii*) from the last sheet, but for the above SDE and only for $\delta t = 0.01$ and also run it up to time $T = 100$.

(iii) Repeat Exercise (*iii*) from the last sheet, but for the above SDE

(iv) Using the samples from (*ii*) at time $T = 100$, what do you expect to be the amount of animal who have a size greater than 6?