Exercise Sheet 1

- ▷ Sheets should be submitted in teams of 3-4 students.
- ▷ Accepted format is a single PDF file. Photos of (readable) handwritten text and math are accepted inside the PDF. Please put the full names and student ID numbers of all team members on the title page of your submission.
- ▷ Submissions are made via a Sciebo dropoff link: `https://fz-juelich.sciebo.de/s/QSUUOgUHd1rW2iU`.
- ▷ Submissions have to follow the naming scheme *sheet01_USER1_USER2_USER3_USER4.pdf* using your university username ("Benutzername" in ILIAS/IDM). If a second version is uploaded use the ending *_v2.pdf* for the file. Only the latest version will be rated.
- ▷ At least 50 points are required to pass this sheet.

1. **Image generation (30 Points)**

Figure 1: Example image captured by the wildlife camera. The image shows "Justin Biber", a beaver that chose Forschungszentrum Jülich as its habitat.[2]

Suppose you have set up a wildlife camera near a lake to document the local wildlife. The camera triggers whenever it detects movement. It takes 8-bit unsigned integer gray-scale images with a resolution of 1000×800 pixels (see Fig. 1 for an example) and intensity values ranging from 0 (black) to 255 (white). This setup can be seen as a generative model of images.

(a) (5 Points) Calculate the dimensionality of the pixel space for the images.

(b) (5 Points) How many unique images can exist in this pixel space, assuming purely random pixel values?

[2]Source: `https://blogs.fz-juelich.de/zweikommazwei/2018/03/16/gestatten-justin-unser-biber/`

(c) (5 Points) Compared to the number of unique random images only very few pixel combinations represent plausible images taken by the camera. Explain why the space of plausible images lies on a much lower-dimensional manifold compared to the full pixel space.

(d) (5 Points) The current camera exhibits some noise due to poor lighting conditions at night. The noise follows an approximate Gaussian distribution. You plan to replace the camera with one that has reduced noise, while keeping all other aspects the same. How will this change the manifold of plausible images generated by this setup?

(e) (5 Points) To extend the collection of captured images, a colleague suggests linearly interpolating between two randomly captured images I_1 and I_2 in pixel space, by overlaying their pixel values to generate new images $I' = \frac{I_1+I_2}{2}$. Describe what the interpolated images will look like.

(f) (5 Points) Discuss whether linear interpolation between two randomly captured images in pixel space is a good generator or not.

2. **Inverse transform sampling (30 Points)**

Given a uniform sampler that generates N independent and identically distributed (i.i.d.) samples $y_i \sim U(0,1)$, we want to obtain samples from the exponential distribution with the following probability density function (PDF):

$$p(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0, \end{cases} \tag{1}$$

where $\lambda > 0$ is the rate parameter of the distribution. Using uniform samples y_i, apply inverse transform sampling to generate N i.i.d. samples $x_i \sim p(x;\lambda)$.

3. **Rejection sampling (40 Points)**

The PDF of a standard normal distribution is given by

$$p(x) = \mathcal{N}(x|0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \tag{2}$$

which does not have a closed-form solution for its cumulative distribution function. To generate samples $x \sim p(x)$, we can apply rejection sampling and use a zero-mean Laplace distribution as the source distribution with the following PDF:

$$q(x;b) = \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right). \tag{3}$$

Parameter $b > 0$ controls the variance of the Laplace distribution.

To perform rejection sampling, we perform the following steps:

- Sample $x \sim q(x;b)$ from the Laplace distribution.
- Sample $u \sim U(0,1)$ from the uniform distribution.

- Accept x as a sample from $p(x)$ if

$$u \leq \frac{p(x)}{M \cdot q(x; b)},$$

repeat the process otherwise.

Constant M ensures the acceptance ratio is properly scaled. On average, the process is repeated M times before a sample x is accepted.

a) (20 Points) The scaled source distribution $M \cdot q(x; b)$ must be greater than or equal to $p(x)$ for all values of x. Show that the minimal value for M that satisfies this condition is given by

$$M(b) = \sqrt{\frac{2}{\pi}}\, b \exp\left(\frac{1}{2b^2}\right)$$

by solving

$$M(b) = \max_x \frac{p(x)}{q(x; b)}.$$

b) (20 Points) To reduce computational effort, we want to minimize the average number of iterations M needed to receive an acceptance. Find a value for b that minimizes $M(b)$. What is the probability of acceptance in each iteration for this b?