# Fundamentals of Simulation Methods

Exercise Sheet 10 — Sampling techniques and Monte Carlo simulation

Daniel Rosenblüh, Janosh Riebesell · Lecturer: Volker Springel · January 22nd, 2016

## 1 Rejection method

We would like to produce a random sample $\{x_i\}$ drawn from the probability distribution function (PDF)

$$p(x) = \frac{p_0}{(x-2)^4 + \sin^8(x-3)}, \tag{1}$$

for $0 \leq x < 5$, with $p(x) = 0$ outside of this interval.

(a) Determine $p_0$ such that the function is normalized, i.e.

$$\int_0^5 p(x)\,\mathrm{d}x = 1. \tag{2}$$

(b) Use the rejection method with a uniform parent distribution over the interval $0 \leq x < 5$ to create a sample of $N = 10^6$ numbers from this distribution. What is the rejection rate? Plot a histogram of the distribution of your points, using 100 bins with a width $\Delta x = 0.05$, and compare it with the target distribution function on a common plot with logarithmic $y$ -axis.

(c) Now consider a function $f(x)$ meant to provide an envelope for $p(x)$. Confirm that the piecewise linear

$$f(x) = \frac{y_{n+1} - y_n}{x_{n+1} - x_n}(x - x_n) + y_n \qquad \text{for } x_n \leq x \leq x_{n+1}, \tag{3}$$

with $n \in \{0, 1, 2, 3, 4\}$ fulfills $p(x) \leq f(x)$ over the interval $[0, 5]$ for the points

| $n$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $x_n$ | 0 | 1.8 | 2.35 | 3 | 5 |
| $y_n$ | 0.01 | 0.15 | 2.5 | 0.1 | 0.002 |

Use $f(x)$ as an auxiliary function in the rejection method to more efficiently sample the function $p(x)$. What is the rejection rate now? Verify with another histogram plot that the obtained sample is correct.

(a) Using numerical integration, we find $\int_0^5 \left[(x-2)^4 + \sin^8(x-3)\right]^{-1} \mathrm{d}x \approx 12.810$, so that a normalized PDF requires approximately

$$p_0 \overset{!}{\approx} \frac{1}{12.810} \approx 0.0781. \tag{4}$$

(b) Using a parent distribution uniformly scattered across the domain $[0, 5] \times [0, 30.5]$ comprising the unnormalized $p(x)/p_0$ (whose maximum is $\approx 30.35$), we inferred a rejection rate of

$$R \approx 0.916 \approx 92\,\%. \tag{5}$$

A histogram of the distribution along the $x$-axis of the randomly generated points that passed the rejection test together with the target distribution function $p(x)$ is shown in Figure 1.

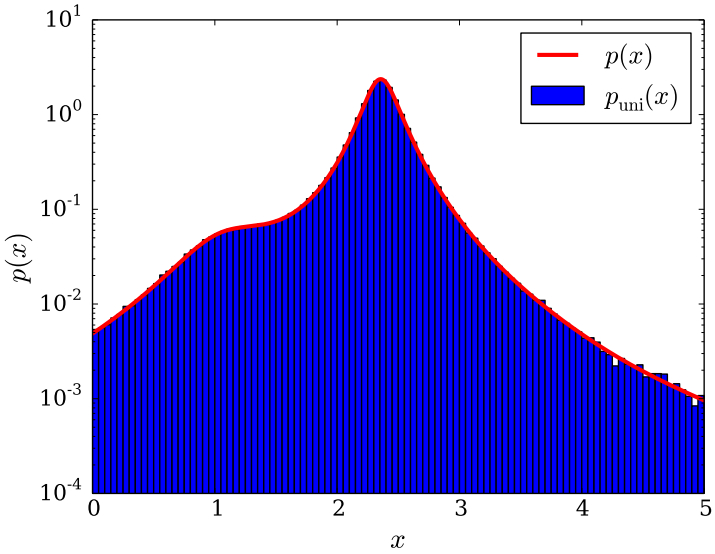



Figure 1: Histogram of accepted points, overplotted with the target distribution function $p(x)$

(c) Figure 2 shows how the piecewise linear function $f(x)$ envelopes the target distribution function $p(x)$. We clearly have $f(x) \geq p(x)\ \forall x \in [0, 5]$.

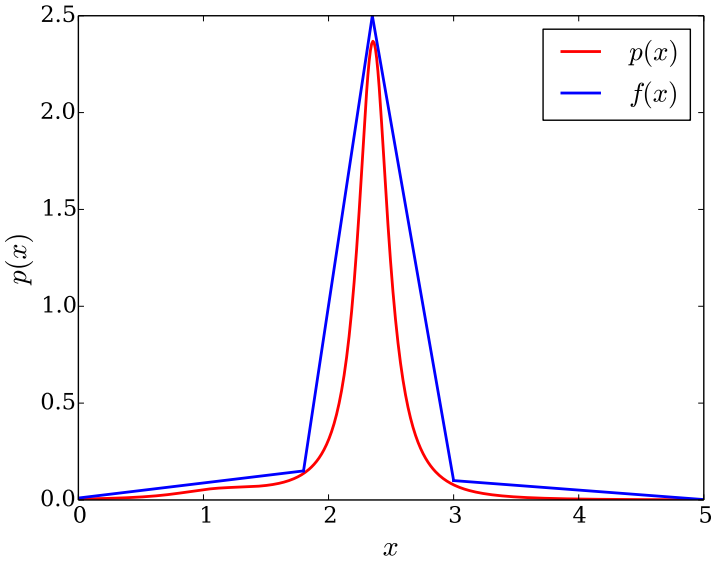



Figure 2: Comparison of $p(x)$ and $f(x)$ demonstrating that the latter is a majorant of the former

We now try to make the rejection method more efficient by no longer uniformly generating samples within a rectangle around $p(x)$, but only within the area enclosed by the envelope $f(x)$ and the $x$-axis. To generate samples with a distribution resembling our target distribution $p(x)$, we use *exact inversion*. We can easily transform one probability distribution $p_1$ into another one $p_2$ via

$$p_1(x_1)\,\mathrm{d}x_1 \stackrel{!}{=} p_2(x_2)\,\mathrm{d}x_2. \tag{6}$$

Equation (6) holds due to conservation of probability. $x_1$ and $x_2$ are not independent variables here. Rather, each is a function of the other,

$$x_1 = x_1(x_2), \qquad x_2 = x_2(x_1). \tag{7}$$

Suppose now that $p_1$ is the distribution we know how to produce, i.e. a uniform one, and $p_2$ is the distribution we would like to generate. To do so, we need to find exactly the second mapping

in eq. (7) that tells us how to compute the sample $x_2$ in the distribution $p_2$ corresponding to the sample $x_1$ in the distribution $p_1$. This can be done easily, at least in principle. $x_1$ and $x_2$ corresponding to each other means that the cumulative distribution functions (CDF) of $p_1$ and $p_2$ evaluated at these points respectively, must be equal. The CDF $P(x)$ of a distribution $p$ gives the probability that any randomly drawn sample of $p$ is smaller than $x$ and is hence defined as

$$P(x) \equiv \int_{-\infty}^{x} p(x')\,\mathrm{d}x'. \tag{8}$$

Thus, using that $P_1(x_1) = P_2(x_2)$, we obtain

$$P_1(x_1) = \int_{-\infty}^{x_1} p_1(x_1')\,\mathrm{d}x_1' = \int_{-\infty}^{x_2} p_2(x_2')\,\mathrm{d}x_2' = P_2(x_2). \tag{9}$$

Inserting for $p_1(x_1')$ a uniform distribution over the compact interval $[a, b]$, i.e.

$$p_1(x_1) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a, b], \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

with $x_1 \in [a, b]$, eq. (9) becomes

$$\frac{x_1 - a}{b - a} = \int_{-\infty}^{x_2} p_2(x_2')\,\mathrm{d}x_2'. \tag{11}$$

The $p_2$ in our case is given by $f$. So to calculate $f$-distributed samples $x_2$ from a variable $x_1$ that is uniformly distributed, we need to calculate the remaining integral in eq. (11) and then solve the whole expression for $x_2$. This last step, i.e. the actual inversion, is always well-defined since probability distributions are non-negative functions, making their integrals monotonically increasing and hence injective.

Working through the above procedure, i.e. first calculating the CDF of our piecewise linear function which is a quadratic spline and can be inverted by the standard quadratic formula to yield $f$-distributed numbers, we were able to perform the rejection method with a rejection rate of

$$r \approx 0.451 \approx 45\,\%. \tag{12}$$

A histogram of the generated samples overlayed with the target distribution $p(x)$ is shown in Figure 3.

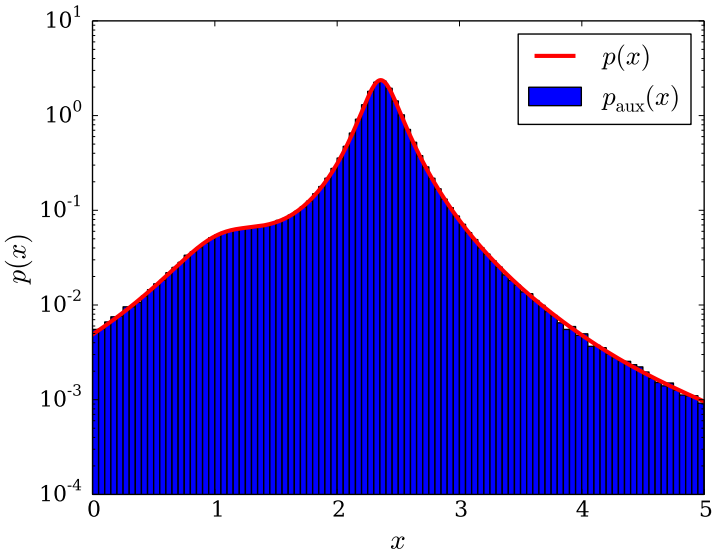



Figure 3: Histogram of samples generated from the auxiliary function $f(x)$ overlayed with the target distribution $p(x)$

## 2 Sampling a given distribution with a Monte Carlo Markov chain

Let's assume we want to generate random numbers from the distribution

$$p(x) \propto e^{-[x+2\cos^2(x)]^2}. \tag{13}$$

This simple case could also be done with the rejection method, but here we want to adopt a different approach, namely the use of a stochastic process constructed with the Metropolis algorithm.

(a) Start with some random guess $x_0$ for which $p(x)$ is not zero.

(b) Make a proposal for $x_i'$ in your chain by adding a random number drawn uniformly from the interval $[-1, 1]$ to $x_{i-1}$.

(c) Accept the proposal with probability

$$r = \min\left(1, \frac{p(x_i')}{p(x_{i-1})}\right), \tag{14}$$

i.e. in the case of acceptance, make it the entry $x_i$ in your Monte Carlo chain. Otherwise, adopt the unmodified $x_{i-1}$ as your element $x_i$. Then proceed with the next element $x_{i+1}$.

(d) Produce a chain with $N = 10^6$ elements, and make a histogram with bin size $\Delta x = 0.02$ of the entries in order to verify that they correctly sample the overplotted shape of $p(x)$. How many *different* points are in your chain?

For parts (a) to (c), see `markovchain.c`.

(d) A histogram of the elements in our Markov chain is shown in Figure 4. Sorting the array and incrementing a count every time two successive entries are unequal yields

$$n_{\mathrm{dist}} = 507\,049 \tag{15}$$

as the number of distinct entries in the Markov chain.

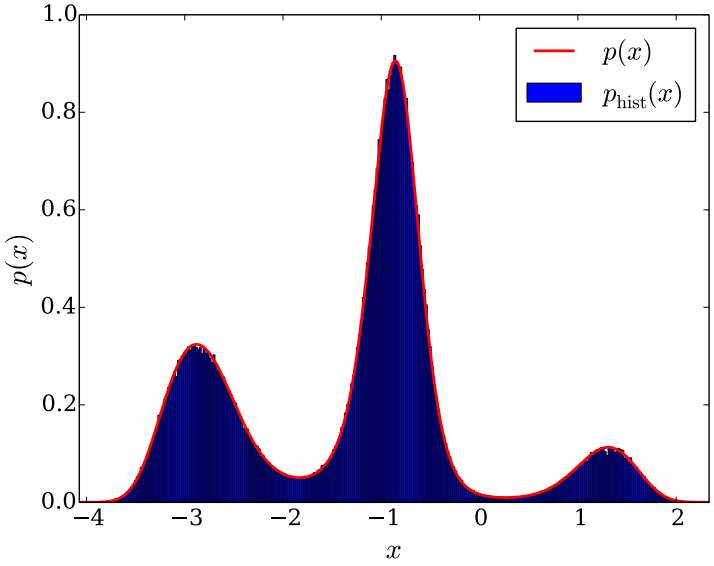



Figure 4: Histogram of points in a $10^6$-entry-long Markov chain, overplotted with the normalized target distribution function $p(x)$

## 3 Monte Carlo simulation of the 2D Ising model

We consider a two-dimensional Ising model with the partition function

$$Z = \sum_{\{s_x=\pm 1\}} \exp\left[-\frac{\beta}{2}\sum_{\langle x,y\rangle}(1 - s_x s_y)\right] \tag{16}$$

for spins with values $s_x = \pm 1$ on a regular lattice of Cartesian topology. For every lattice site $x$, the interaction term only involves the four nearest neighbours $y$ on adjacent lattice sites. Use an $N \cdot N = 32^2$-square lattice with periodic boundary conditions.
We have the goal to measure the average magnetization

$$\langle |M| \rangle = \left\langle \left| \frac{1}{N^2} \sum_x s_x \right| \right\rangle \tag{17}$$

in thermal equilibrium for different values of $\beta$, namely

$$\beta \in \{1.6, 1.3, 1.1, 1.0, 0.9, 0.8, 0.7, 0.6, 0.4, 0.1\}. \tag{18}$$

(a) For each value of $\beta$, go through the lattice in red-black order for at least 3000 times to establish thermalization. At each encountered lattice site $x$, calculate the interaction energies $E_+$ and $E_-$ for the spin up and the spin down directions of $x$, respectively. Then choose for the site the spin-up direction with the *heat bath* (Gibbs sampling) probability

$$p_+ = \frac{e^{-E_+}}{e^{-E_+} + e^{-E_-}}, \tag{19}$$

and the spin-down direction correspondingly with probability $p_- = 1 - p_+$.

(b) After the thermalization period, take measurements of $M$ after at least 1000 full mesh sweeps and calculate the average $\langle |M| \rangle$. Report this in a table for each $\beta$.

(c) Plot $\langle |M| \rangle$ as a function of $T \propto 1/\beta$. We note that this model has a phase transition at a critical $\beta_c = \ln\left(1 + \sqrt{2}\right) \approx 0.881$ (corresponding to a critical temperature $T_c \propto 1/\beta_c$). Below this temperature (i.e. for $\beta > \beta_c$), the system shows spontaneous magnetization.

The present two-dimensional model without external field has been solved analytically by Onsager. In three dimensions, no analytic solution is known but it can be readily obtained with Monte Carlo simulations.

(a) See `ising.c`.

(b) The average magnetizations we measured for the supplied values of $\beta$ are shown in Table 1.

| $\beta$ | 1.6 | 1.3 | 1.1 | 1.0 | 0.9 | 0.8 | 0.7 | 0.6 | 0.4 | 0.1 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\langle\lvert M\rvert\rangle$ | 0.996 | 0.984 | 0.955 | 0.910 | 0.719 | 0.206 | 0.103 | 0.062 | 0.041 | 0.028 |

Table 1: Average magnetization $\langle|M|\rangle$ at different temperatures

(c) The average magnetization $\langle|M|\rangle$ as a function of temperature $T$ is shown in Figure 5.

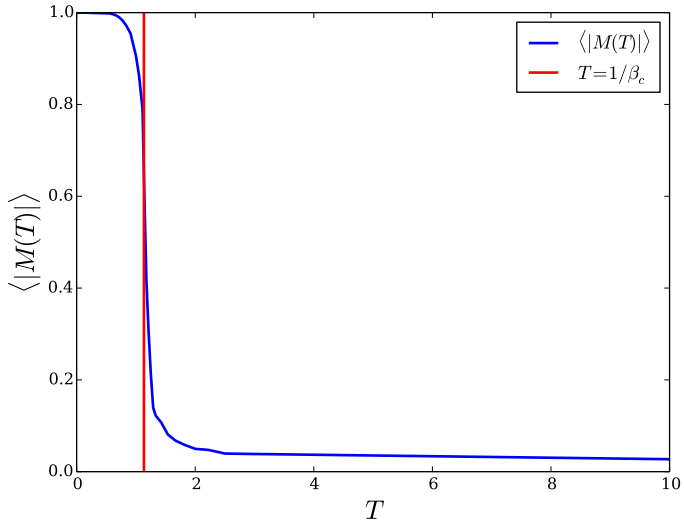
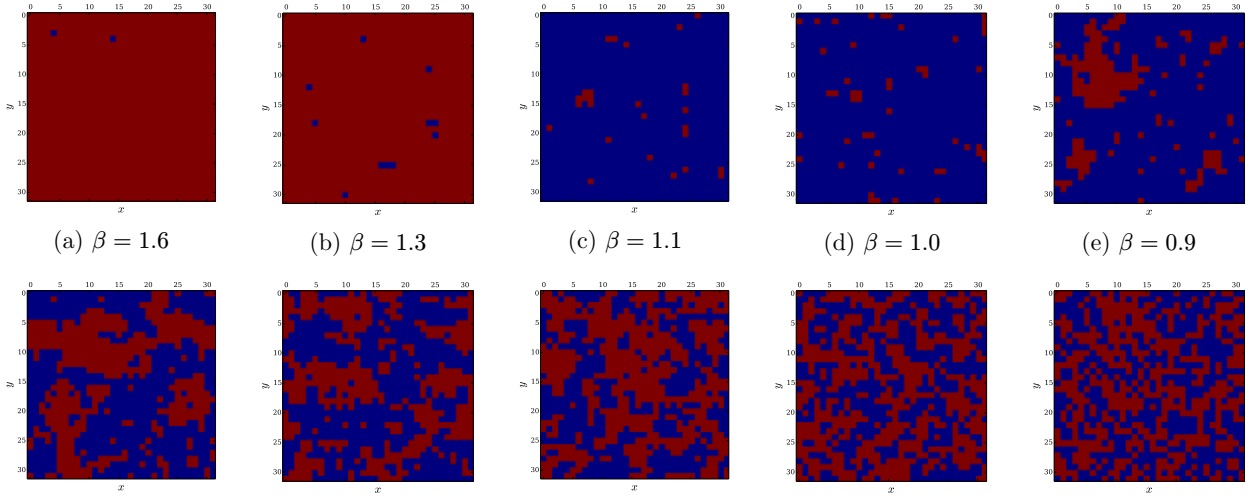



Figure 5: Average magnetization $\langle|M|\rangle$ as a function of $T$

**Note:** Figure 6 (a) to (j) show the spin states of the thermalized lattices at different $\beta$ as a color-coded surface. A red square indicates that the spin at the corresponding lattice site is pointing in the upward direction, whereas a blue square represents a spin down.



(a) $\beta = 1.6$ (b) $\beta = 1.3$ (c) $\beta = 1.1$ (d) $\beta = 1.0$ (e) $\beta = 0.9$

(f) $\beta = 0.8$ (g) $\beta = 0.7$ (h) $\beta = 0.6$ (i) $\beta = 0.4$ (j) $\beta = 0.1$

Figure 6: Color-coded (red: $s_x = 1$, blue: $s_x = -1$) lattice configurations at different temperatures after 3000 relaxation sweeps