Markov chain Monte Carlo for dummies

- MCMC methods are used for approximate sampling from a given **target distribution** π on a **state space** X for which direct sampling is difficult.
- X is usually a subset of R^n. X can be both discrete and continuous.
- The main idea is to create an MC taking values at X such that **the limit distribution** of the MC coincides with π.
- Note that the MC generated by an MCMC method is typically composed of **dependent** random variables (or random vectors).
- While it is often simple to construct an MCMC method with the required limit distribution, it is usually hard to assess the **speed of convergence**, i.e., how long should the **burn-in phase** be, and how often should we sample from the MC, such that the **autocorrelation** of the sample is negligible.
- If possible, one should use **direct sampling** (mostly in smaller dimensions), e.g., the **rejection method**. In higher dimensions, direct sampling can suffer from the **curse of dimensionality** and the use of an MCMC method can be justified.
- The most common MCMC method is the so-called **Metropolis-Hastings** method. Another method is **Gibbs sampling** (which can be viewed as a variant of the MH method).

Metropolis-Hastings method

- In addition to X and π we need a system of **candidate distributions** indexed by the elements of X, i.e., Q(.|x) for all x in X. For the resulting method to have the limit distribution π, the system of candidate distributions must satisfy some theoretical properties, but we will not detail them here. We also need an initial state x_1.
- The MH algorithm proceeds as follows:
  1. Set i=1
  2. Generate the **candidate** y from Q(.|x_i)
  3. Generate u from U(0,1)
  4. If $u < \frac{\pi(y)Q(x_i|y)}{\pi(x_i)Q(y|x_i)}$ then **accept the candidate** (set x_{i+1}=y), otherwise **reject the candidate**, (set x_{i+1}=x_i)
  5. Set i=i+1 and proceed by step 2
- Note a major advantage of the MH algorithm: We do not need to know the normalizing constant of the distribution π.
- If Q is symmetric with respect to the two arguments, we obtain a simplified version of the MH algorithm called **Metropolis algorithm**. Notice the logic of the acceptance and rejection of the candidates in the Metropolis algorithm.
- If Q does not depend on the second argument, we say that the MH algorithm is an **independence sampler**.
- If Q(y|x) depends only on the distance between x and y we have the so-called **random walk sampler**.