Statistical Inference 2 Summer Term 2025

Lecture: Prof. Dr. Göran Kauermann
Exercises: Sergio Buttazzo, Jan Anders
Tutorial: Eugen Gorich

Exercise Sheet 2

Some exercise parts may be flagged with **Optional**. We will cover these in the exercise sessions on Thursday. They should not be presented in the student presentations on Wednesday and also do not need to be included in your submissions on Moodle. We will not ask any questions on these parts after the presentations.

Exercise 2.1 - MCMC

In its simplest form, the weather can be modelled as states that have certain probabilities of staying the same or changing to a different condition. Let us assume three possible weather states, "sunny" (S), "cloudy" (C) and "rainy" (R). The probabilities of the weather on the next day, given today's weather, are denoted in the transition matrix below. For example, if it is sunny today, the probability of it being sunny tomorrow is 1/2, cloudy 1/3 and rainy 1/6. The probabilities have to sum up to one in each row.

	tomorrow	S	C	R
today	S	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$
	C	$\frac{1}{6}$	$\frac{2}{3}$	$\frac{1}{6}$
	R	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

(a) If the weather is *rainy* today, find the probability that it is rainy two days later.

(b) Write R-code that will simulate a Markov Chain of the weather on n days. Choose "sunny" as starting value and $n = 1000$. What is the estimated stationary (unconditional) distribution $\boldsymbol{\pi} = (\pi_S, \pi_C, \pi_R)$ of the three weather states from this chain, i.e. what is the probability of the weather being sunny, cloudy or rainy on any random given day?

(c) For the stationary distribution, $\boldsymbol{\pi P} = \boldsymbol{\pi}$ must hold, where $\boldsymbol{P}$ is the transition matrix given above. Set up a system of equations and determine the stationary distribution by hand, then compare with your results from (b).

Exercise 2.2 - MCMC, Gibbs Sampling

Assume two independent exponential random variables with mean 1, i.e. $X = (X_1, X_2)$ with $X_i \sim Exp(1)$, and define their sum as $S = X_1 + X_2$. In this exercise, we are interested in sampling from the distribution of X conditionally on $S > 10$, that is we want to sample from $Y = \{x \in X \mid S_x = \sum_{i=1}^{2} x_i > 10\}$.

(a) Calculate the (unconditional) probability $P(S > 10)$. Explain why using a traditional simulation-based approach (i.e. just simulating 2 exponentially distributed variables a lot of times and selecting results where the sum is larger than 10) for simulating a random vector (y_1, y_2) from Y is not very efficient. What happens to the applicability of this approach if we increase the threshold for the sum?

(b) Explain how to use the Metropolis-Hastings algorithm to generate a large number of samples from (X1, X2) the distribution of which approximately follows the distribution of Y defined above. Implement the procedure in `R`, generate a chain of 1000 samples and plot it. Explain what you see and compare it to the procedure from (a).

(c) **Optional:** We now want to sample j i.i.d. vectors (X_{j1}, X_{j2}) from the target distribution using our method from (b). What do we have to keep in mind when sampling from the simulated values?

(d) Instead of sampling from a two dimensional vector, we may also want to sample in a higher dimensional space with an adjusted condition, i.e. $(X_1, ..., X_n)$ with $\sum_{i=1}^{n} X_i > 5n$. What's an issue that can occur in high dimensional sampling spaces and how does Gibbs sampling solve this? What would we have to change to adjust the implementation from (b) to Gibbs sampling in this new setting? (No implementation necessary)

Exercise 2.3 - Sampling

Consider a discrete random variable Y with three different states (H), (R) and (W) with given probabilities:

$$\pi_H = P(H) = 0.3, \quad \pi_R = P(R) = 0.1, \quad \pi_W = P(W) = 0.6.$$

We will use the Metropolis (Hastings) algorithm to construct a Markov chain $\{Y_t^*,\ t \in \mathbb{N}_0\}$ with stationary distribution $\boldsymbol{\pi} = (\pi_H, \pi_R, \pi_W)$. We assume a symmetric proposal density.

The formula for the transition probability in the discrete case, i.e. the probability of going from state Y_t^* to Y_{t+1}^*, is given by:

$$p_{ij} = P(Y_{t+1}^* = j | Y_t^* = i) = P(Y_1^* = j | Y_0^* = i) = \begin{cases} q_{ij}\alpha_{ij} & i \neq j, \\ 1 - \sum_{k \in S \setminus \{i\}} q_{ik}\alpha_{ik} & i = j, \end{cases} \quad i, j \in S,$$

where α_{ij} is the probability that we accept for $Y_t^* = i$ a simulated value $Y_{t+1}^* = j$, and $q_{ij} = q(i, j)$ is the proposal probability for j given state i.

The following matrix of proposal densities Q is given:

$$\boldsymbol{Q} = \begin{pmatrix} 1-2q & q & q \\ q & 1-2q & q \\ q & q & 1-2q \end{pmatrix}, \quad q \in (0, 0.5].$$

(a) Compute the acceptance matrix $\boldsymbol{A}$, where each of the nine entries α_{ij} is the probability of the chain moving from state Y_t^* to Y_{t+1}^* given that the latter state Y_{t+1}^* was already proposed.

(b) Using the matrix A just obtained, compute the transition matrix $\boldsymbol{P}$, i.e. the matrix containing the nine probabilities of the chain moving from state Y_t^* to Y_{t+1}^* given that the latter state Y_{t+1}^* was not yet proposed.

(c) Show that $\boldsymbol{\pi}$ is an invariant distribution for P.

(d) **Optional:** Write down in pseudo-code a Metropolis (Hastings) algorithm that constructs a Markov chain with stationary distribution $\boldsymbol{\pi}$. Implement the algorithm in R. Generate a chain of length $K = 100$ for $q = 0.2$ and visualize the sampling path as well as the proposed values in each iteration. Compute the acceptance rate as well. Finally, check if the sample is approximately π distributed if we throw away the 50 first values as 'burn in'.