3 Exercises (due Wednesday 15.11.2023, 23:59pm)

There is a total of 7 points to achieve on this third exercise sheet. Please upload your solutions in typed or readable (!) handwritten form (scans, photographs), and also your code, on Moodle and don't forget to flag all the problems you were able to solve. **If the runtime of your simulations is very high, you may want to save your results beforehand so we don't have to wait at the homework session until your simulations are finished.** Good luck!

Exercise 3.1 (2P). *Let $X_1, \dots, X_n$ be iid with invertible cdf $F(x) = P(X_1 \leq x)$ and let $\hat{F}_n$ denote the corresponding empirical cdf. Show that for $\alpha \in (0,1)$,*

$$\hat{F}_n^\dagger(\alpha) \xrightarrow[n\to\infty]{i.p.} F^{-1}(\alpha).$$

Hint: *Note that it is enough to derive an upper bound on $P(|\hat{F}_n^\dagger(\alpha) - F^{-1}(\alpha)| > \varepsilon)$ which converges to zero as $n \to \infty$. Explain why $P(|\hat{F}_n^\dagger(\alpha) - F^{-1}(\alpha)| > \varepsilon) = P(\hat{F}_n^\dagger(\alpha) > F^{-1}(\alpha) + \varepsilon) + P(\hat{F}_n^\dagger(\alpha) < F^{-1}(\alpha) - \varepsilon)$. Use the property of the generalized inverse that $\{x \in \mathbb{R} : \hat{F}_n(x) \geq \alpha\} = [\hat{F}_n^\dagger(\alpha), \infty)$, to further upper bound these two probabilities. Convince yourself that $F(F^{-1}(\alpha) + \varepsilon) > \alpha$ and that $\mathbb{E}[\mathbb{1}_{(-\infty,x]}(X_i)] = F(x)$. Use the Law of Large Numbers on $\hat{F}_n(F^{-1}(\alpha) + \varepsilon)$.*

Exercise 3.2 (3P). *Write rejection sampling algorithms that generate bivariate random vectors from the pdf $f(x,y) \propto [\sin(xy)]^2[\cos(-xy)]^2 e^{-8x^2-2|y|^3}$. Think about ways to visualize the density f and the samples you are generating.*

a) ***(1P)*** *Write an exact algorithm by using a Gaussian proposal density*

$$g(x,y) = (2\pi)^{-1} \exp\left(-\frac{1}{2}(x^2 + y^2)\right).$$

b) ***(1P)*** *Write an approximate algorithm that uses a uniform proposal density with large support that covers 'most' of the distribution f.*

c) ***(1P)*** *Compare both approaches in terms of the number of iterations that are necessary to produce a sample of a given size N.*

Hint: *For Part a), convince yourself that $e^{-|y|^3} \leq e \cdot e^{-y^2}$ for all $y \in \mathbb{R}$. You can use a random number generator that produces standard normal random variables. For b) you can of course also use one that generates uniform random variables.*

Exercise 3.3 (2P). *Consider the iid Bernoulli model $p(x|\theta) = \prod_{i=1}^n \theta^{x_i}(1-\theta)^{1-x_i}$, $x = (x_1, \dots, x_n) \in \{0,1\}^n$, $\theta \in (0,1)$. Put a non-informative uniform prior on θ and compute and visualize the posterior distribution for different samples x of different sizes n. Don't do any Montecarlo simulation here. Use your computer only for the visualizations!*