Stefan Walzer, Stefan Hermann
ITI Prof. Sanders
winter term 2025/2026

Exercise Sheet 9 – Coupling, Balls into Bins and Poissonisation

Probability and Computing

Exercise 1 – Coupling of a Random Walk

Let $X_1, X_2, \ldots \sim \mathcal{U}(\{-1, 1\})$ be independent random variables. For $n \in \mathbb{N}_0$, define $W_n := \sum_{i=1}^{n} X_i$. The sequence $(W_n)_{n\in\mathbb{N}_0}$ is called a random walk. We may also consider a shifted random walk $(V_n)_{n\in\mathbb{N}_0}$ defined by $V_n := W_n + 42$, which therefore has initial position $V_0 = 42$ instead of $W_0 = 0$. We aim to show that the choice of initial position typically does not matter in the long run. We will use without proof that the random walk visits every integer at least once with probability 1. In particular, $\lim_{n\to\infty} \Pr[\max\{W_1, \ldots, W_n\} < c] = 0$ for all $c \in \mathbb{N}$.

(i) Let $S_1, S_2, \ldots \subseteq \mathbb{Z}$ be arbitrary sets. Show that $\lim_{n\to\infty} |\Pr[W_n \in S_n] - \Pr[V_n \in S_n]| = 0$.
Hint: Construct a coupling $(W'_n, V'_n)_{n\in\mathbb{N}_0}$ of $(W_n)_{n\in\mathbb{N}_0}$ and $(V_n)_{n\in\mathbb{N}_0}$ such that $\lim_{n\to\infty} \Pr[W'_n = V'_n] = 1$.

(ii) Show that the result of part (i) does not hold in this form for a shift of 43 instead of 42.

Exercise 2 – Coupling and Total Variation Distance

Let X and Y be two random variables taking values in $\mathbb{N}$. The total variation distance (English: *total variation distance*) between X and Y (or their distributions) is defined as[1]

$$d(X, Y) = \frac{1}{2} \sum_{i\in\mathbb{N}} |\Pr[X = i] - \Pr[Y = i]|.$$

(i) Show: There exists a coupling (X', Y') of X and Y such that $\Pr[X' \neq Y'] = d(X, Y)$.

(ii) Show: No coupling (X', Y') of X and Y satisfies $\Pr[X' \neq Y'] < d(X, Y)$.

Exercise 3 – Properties of the Poisson Distribution

Let $X \sim \text{Pois}(\lambda)$. Show:

(i) $\mathbb{E}[X] = \lambda$.

(ii) $\text{Var}(X) = \lambda$.

[1] A general definition applicable also to continuous probability spaces can be found on Wikipedia.

(iii) For $Y \sim \text{Pois}(\rho)$ independent of X, we have $X + Y \sim \text{Pois}(\lambda + \rho)$.

(iv) For $X' \sim \text{Bin}(X, p)$, we have $X' \sim \text{Pois}(\lambda p)$.
Note: Here, a two-stage random experiment is performed. The outcome X of the first stage serves as a parameter of the second stage.

Exercise 4 – Poissonised Bloom Filters

We consider a Poisson model of Bloom filters, i.e., we assume that each position in the array independently appears as a hash value $\text{Pois}(\alpha k)$-many times.

(i) We again choose $\alpha k = \ln 2$. How can we show that the fraction $\frac{Z}{m}$ of zeros is with high probability close to $\frac{1}{2}$?

(ii) How could this result be transferred to a non-Poissonised model?