ETH Zurich
Institute of Theoretical Computer Science
Dr. Karl Bringmann, Dr. Johannes Lengler, and Dr. Tsur Luria
Frank Mousset, Felix Weissenberger

FS 2015
Handout 2

Randomized Algorithms and Probabilistic Methods: Advanced Topics

Exercise 1

To get used to computing with conditional expectations, you should try to prove the following claims using only the definitions. For the conditional expectation of a discrete random variable, use the definition

$$\mathbb{E}[X \mid \mathcal{A}] = \sum_i i \Pr[X = i \mid \mathcal{A}].$$

If X and Y are both discrete variables, we also define $\mathbb{E}[X \mid Y] = g(Y)$ where $g(y) = \mathbb{E}[X \mid Y = y]$.

(a) Let $(X_t)_{t\geq 0}$ be a Markov Chain with state space $0 \in S \subseteq \mathbb{R}_{\geq 0}$. Let T be the first point in time where $X_T = 0$. Show that

$$\mathbb{E}[T \mid X_0 = x] = \sum_{y\in S} \Pr[X_1 = y \mid X_0 = x] \cdot \mathbb{E}[T \mid X_1 = y],$$

for $x \neq 0$.

(b) Let X and Y be discrete random variables. Prove that

$$\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X].$$

Exercise 2

Suppose that in the Clumsy Coupon Collector process, Alice loses *all* of her coupons with probability $1/(2n)$ each time that she gets a new coupon. Thus, at least after collecting $(1 - \varepsilon)n$ coupons, the expected number of coupons that she loses when she gets a new coupon is close to $\rho = 1/2$. Prove that, nevertheless, the expected number of rounds until she has all coupons is $(1 + o(1))n \ln n$.

Exercise 3

Prove that in the random decline process, the expected time until the process finishes is $\Omega(\ln n)$ for every constant $a > 0$ (where a is the parameter of the random decline, i.e., X_{t+1} is chosen u.a.r. from $\{0, 1, \ldots, \lfloor aX_t\rfloor\}$). *Hint:* Show that the case $0 < a \leq 1$ implies the other cases, and prove the claim only for this case.

Exercise 4

Let f be a function on the hypercube $H_n = \{0, 1\}^n$. Consider the following optimization algorithm:

(i) Choose $x_0 \in H_n$ uniformly at random.

(ii) **Forever Do**:

Generate $y = y_t$ by flipping each bit of x_t independently with probability $1/n$.
If $f(y) < f(x)$ **then** set $x_{t+1} := y$ **else** set $x_{t+1} := x_t$.
$t := t + 1$.

We call the *optimization time* of f the expected number of evaluations of loop (ii) before x is a minimal value of f, i.e., the smallest t such that x_t is minimal.
Now let f be a linear function on the hypercube $H_n := \{0, 1\}^n$, i.e., there are weights $w_1, \ldots, w_n \in \mathbb{R}$ such that $f(x_1 \ldots x_n) = \sum_{i=1}^n x_1 w_i$. We want to estimate the optimization time of f.

(a) Justify that you may restrict to the case where all weights are non-negative and sorted, so $w_1 \geq w_2 \geq \ldots \geq w_n \geq 0$.

(b) Prove that the probability p_0 that y_t differs from x_t in exactly one bit is in the interval $[1/e, 1]$.

(c) Consider the random variable $X_t := f(x_t)$. Using the previous statement, prove that for all a the drift of X_t at a is at least

$$\mathbb{E}[X_t - X_{t+1} \mid X_t = a] \geq \frac{a}{en}.$$

(d) Let $w_{\min} := \min_i\{w_i \mid w_i > 0\}$. Conclude that the optimization time is at most

$$en\left(1 + \log\left(\frac{\sum_{i=1}^n w_i}{w_{\min}}\right)\right).$$