PS: Advanced Probability Theory
Sheet 2 Solutions

lucas.teyssier@univie.ac.at

Exercise 1. (Moritz Dober, adapted)

a) We have by linearity $\mathbb{E}\left[(X-t)^2\right] = \mathbb{E}\left[X^2\right] - 2t\mathbb{E}[X] + t^2$ for all t, and therefore

$$\mathbb{E}[f(X)] = \mathbb{E}\left[\mathbb{E}\left[X^2\right] - 2X\mathbb{E}[X] + X^2\right] = 2(\mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2) = 2\mathrm{Var}(X).$$

b) Set $m = \mathbb{E}[X]$ to obtain

$$\begin{aligned} f(t) &= \mathbb{E}\left[(X-m+m-t)^2\right] = \mathbb{E}\left[(X-m)^2 + 2(X-m)(m-t) + (m-t)^2\right] \\ &= \mathbb{E}\left[(X-m)^2\right] + 2(m-t)\underbrace{\mathbb{E}[X-m]}_{=0} + (m-t)^2 = \mathrm{Var}(X) + \underbrace{(m-t)^2}_{\geq 0}. \end{aligned}$$

Hence f is minimal at $t = m = \mathbb{E}[X]$.

Remark. Differentiating f gives also a good solution.

c) As hinted, let X_1, X_2 be two independent copies of X and consider the r.v. $(f(X_1) - f(X_2))(g(X_1) - g(X_2))$. Note that $(f(X_1) - f(X_2))(g(X_1) - g(X_2)) \geq 0$ since f, g are both non-decreasing. Therefore, by monotonicity of the expectation we have

$$\begin{aligned} 0 &\leq \mathbb{E}[(f(X_1) - f(X_2))(g(X_1) - g(X_2))] \\ &= \mathbb{E}[f(X_1)g(X_1)] - \mathbb{E}[f(X_1)g(X_2)] - \mathbb{E}[f(X_2)g(X_1)] + \mathbb{E}[f(X_2)g(X_2)] \\ &\overset{*}{=} \mathbb{E}[f(X_1)g(X_1)] - \mathbb{E}[f(X_1)]\,\mathbb{E}[g(X_2)] - \mathbb{E}[f(X_2)]\,\mathbb{E}[g(X_1)] + \mathbb{E}[f(X_2)g(X_2)] \\ &\overset{**}{=} 2(\mathbb{E}[f(X)g(X)] - \mathbb{E}[f(X)]\,\mathbb{E}[g(X)]) = 2\,\mathrm{cov}(f(X), g(X)), \end{aligned}$$

where we used in $*$ that X_1, X_2 are independent (hence so are $f(X_i)$, $g(X_j)$ for $i \neq j$) and in $**$ that X has the same law as X_1, X_2.

Exercise 2. (M.D.)

a) As suggested in the hint, let Y_i be the time it takes to collect an ith new item. Then we have for $k \geq 1$

$$\mathbb{P}(Y_i = k) = \left(\frac{i-1}{n}\right)^{k-1}\left(\frac{n-i+1}{n}\right),$$

i.e. Y_i is geometrically distributed with parameter $p_i := \frac{n-i+1}{n}$.
Noting that $X_n = \sum_{i=1}^n Y_i$ we deduce

$$\mathbb{E}[X_n] = \sum_{i=1}^n \underbrace{\mathbb{E}[Y_i]}_{=1/p_i} = n \sum_{i=1}^n \frac{1}{n-i+1} = n \underbrace{\sum_{i=1}^n \frac{1}{i}}_{\substack{\sim \log n \\ (\text{as } n\to\infty)}} \sim n \log n \text{ (as } n \to \infty).$$

b) Since $Y_1, ..., Y_n$ are even independent they are uncorrelated, and hence

$$\text{Var}(X_n) = \sum_{i=1}^n \text{Var}(Y_i) = \sum_{i=1}^n \frac{1-p_i}{p_i^2} = \sum_{i=1}^n \frac{n(i-1)}{(n-i+1)^2} = n \sum_{i=1}^{n-1} \frac{i}{(n-i)^2}.$$

Therefore,

$$\frac{\text{Var}(X_n)}{\mathbb{E}[X_n]^2} \overset{a)}{\sim} \frac{\text{Var}(X_n)}{n^2 \log^2 n} = \frac{1}{n \log^2 n} \underbrace{\sum_{i=1}^{n-1} \frac{i}{(n-i)^2}}_{\leq (n-1) \sum_{i=1}^{n-1} \frac{1}{i^2}} \leq \underbrace{\frac{1}{\log^2 n}}_{\to 0} \underbrace{\sum_{i=1}^{n-1} \frac{1}{i^2}}_{<C} \to 0 \text{ as } n \to \infty.$$

c) By Chebyshev's inequality, we have

$$\mathbb{P}(|X_n - \mathbb{E}[X_n]| > \varepsilon \mathbb{E}[X_n]) \leq \frac{1}{\varepsilon^2} \underbrace{\frac{\text{Var}(X_n)}{\mathbb{E}[X_n]^2}}_{\to 0 \text{ by b)}} \xrightarrow[n\to\infty]{} 0.$$

This means by definition that $\frac{X_n}{\mathbb{E}[X_n]}$ converges in probability to the constant r.v. 1.

Exercise 3. (L.T.)

First observe that for all $i \geq 1$, $\mathbb{P}(X_i = 0) = \mathbb{P}(X_i = 1) = 1/2$ (because U and $\{U + 1/2^i\}$ have the same law, where $\{\cdot\}$ denotes the fractional part).
Then for all integer $n \geq 1$, and $(\epsilon_i)_{1 \leq i \leq n} \in \{0,1\}^n$,

$$\mathbb{P}\left(\bigcap_{i=1}^n \{X_i = \epsilon_i\}\right) = U\left(\left(\sum_{i=1}^n \frac{\epsilon_i}{2^i}, \sum_{i=1}^n \frac{\epsilon_i}{2^i} + \frac{1}{2^n}\right)\right) = \frac{1}{2^n} = \prod_{i=1}^n \mathbb{P}(X_i = \epsilon_i).$$

Taking unions of these elementary events, we deduce that for any finite $I \subset \mathbb{N}^*$,

$$\mathbb{P}\left(\bigcap_{i\in I} \{X_i = \epsilon_i\}\right) = \prod_{i \in I} \mathbb{P}(X_i = \epsilon_i),$$

which is the definition of the independance of the family $(X_i)_{i \geq 1}$.

Exercise 4. (M.D.)

The r.v. (X, Y) having a joint density function $f(x, y)$ means that

$$\mathbb{P}((X,Y) \in B) = \int_B f \, d\lambda^2 \quad \forall B \in \mathcal{B}(\mathbb{R}^2).$$

Let $A, B \in \mathcal{B}(\mathbb{R})$, then

$$\begin{aligned}\mathbb{P}\left(X \in A,\, Y \in B\right) = \mathbb{P}\left((X,Y) \in A \times B\right) = \int_{A\times B} f \, d\lambda^2 &= \int_{\mathbb{R}^2} \underbrace{1_{A\times B}(x,y)}_{=1_A(x)1_B(y)} g_1(x) g_2(y) \, d\lambda^2(x,y) \\ &= \int_{\mathbb{R}} 1_A(x) g_1(x) \left(\int_{\mathbb{R}} 1_B(y) g_2(y)\, d\lambda(y) \right) d\lambda(x) \\ &= \left(\int_A g_1(x)\, d\lambda(x) \right) \left(\int_B g_2(y)\, d\lambda(y) \right),\end{aligned}$$

where we used Fubini's theorem since $\int_{\mathbb{R}^2} |1_{A\times B}(x,y) g_1(x) g_2(y)|\, d\lambda^2(x,y) \le \int_{\mathbb{R}^2} f\, d\lambda^2 = 1$. Setting $B = \mathbb{R}$, we get

$$\mathbb{P}\left(X \in A\right) = \left(\int_A g_1 \, d\lambda \right) \underbrace{\left(\int_{\mathbb{R}} g_2 \, d\lambda \right)}_{=:C} = \int_A C g_1 \, d\lambda.$$

Hence $f_X := C g_1$ is a density for X. Analogously, by setting $A = \mathbb{R}$, we see that $f_Y := C' g_2$ is a density for Y, where $C' := \int_{\mathbb{R}} g_1 \, d\lambda$.
Now, set $A = B = \mathbb{R}$ to deduce $1 = \int_{\mathbb{R}^2} f\, d\lambda^2 = CC'$, and therefore $C' = \frac{1}{C}$.
Finally, we conclude

$$\begin{aligned}\mathbb{P}\left(X \in A,\, Y \in B\right) &= \left(\int_A g_1 \, d\lambda \right) \left(\int_B g_2 \, d\lambda \right) \\ &= \left(\int_A f_X \, d\lambda \int_B f_Y \, d\lambda \right) = \mathbb{P}\left(X \in A\right) \mathbb{P}\left(Y \in B\right).\end{aligned}$$