Problem Set 1 (Graded) — Due Friday, October 1, before class starts
For the Exercise Sessions on September 24

<table>
<thead>
<tr>
<th>Last name</th>
<th>First name</th>
<th>SCIPER Nr</th>
<th>Points</th>
</tr>
</thead>
</table>

Problem 1: Review of Random Variables

Let $X$ and $Y$ be discrete random variables defined on some probability space with a joint pmf $p_{XY}(x,y)$. Let $a, b \in \mathbb{R}$ be fixed.

(a) Prove that $E[aX + bY] = aE[X] + bE[Y]$. Do not assume independence.

(b) Prove that if $X$ and $Y$ are independent random variables, then $E[X \cdot Y] = E[X] \cdot E[Y]$.

(c) Assume that $X$ and $Y$ are not independent. Find an example where $E[X \cdot Y] \neq E[X] \cdot E[Y]$, and another example where $E[X \cdot Y] = E[X] \cdot E[Y]$.

(d) Prove that if $X$ and $Y$ are independent, then they are also uncorrelated, i.e.,

$$\text{Cov}(X,Y) := E[(X - E[X])(Y - E[Y])] = 0. \quad (1)$$

(e) Find an example where $X$ and $Y$ are uncorrelated but dependent.

(f) Assume that $X$ and $Y$ are uncorrelated and let $\sigma^2_X$ and $\sigma^2_Y$ be the variances of $X$ and $Y$, respectively. Find the variance of $aX + bY$ and express it in terms of $\sigma^2_X, \sigma^2_Y, a, b$.

Hint: First show that $\text{Cov}(X,Y) = E[X \cdot Y] - E[X] \cdot E[Y]$.

Problem 2: Review of Gaussian Random Variables

A random variable $X$ with probability density function

$$p_X(x) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (2)$$

is called a Gaussian random variable.

(a) Explicitly calculate the mean $E[X]$, the second moment $E[X^2]$, and the variance $\text{Var}[X]$ of the random variable $X$.

(b) Let us now consider events of the following kind:

$$\mathbb{P}(X < \alpha). \quad (3)$$

Unfortunately for Gaussian random variables this cannot be calculated in closed form. Instead, we will rewrite it in terms of the standard Q-function:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \, du \quad (4)$$
Express $\mathbb{P}(X < a)$ in terms of the Q-function and the parameters $m$ and $\sigma^2$ of the Gaussian pdf.

Like we said, the Q-function cannot be calculated in closed form. Therefore, it is important to have bounds on the Q-function. In the next 3 subproblems, you derive the most important of these bounds, learning some very general and powerful tools along the way:

(c) Derive the Markov inequality, which says that for any non-negative random variable $X$ and positive $a$, we have

$$
\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}[X]}{a}.
$$

(5)

(d) Use the Markov inequality to derive the Chernoff bound: the probability that a real random variable $Z$ exceeds $b$ is given by

$$
\mathbb{P}(Z \geq b) \leq \mathbb{E}[e^{s(Z-b)}], \quad s \geq 0.
$$

(6)

(e) Use the Chernoff bound to show that

$$
Q(x) \leq e^{-\frac{x^2}{2}} \quad \text{for } x \geq 0.
$$

(7)

Problem 3: Moment Generating Function

Let $X$ be a real-valued random variable taking values on a finite set. The logarithmic moment generating function is defined as follows.

$$
\phi(s) := \ln \mathbb{E}[\exp(sX)] = \ln \sum_x p(x) \exp(sx)
$$

(a) Show that $p_s(x) := p(x) \exp(sx) \exp(-\phi(s))$ is a probability mass function.

(b) Let $X_s$ be a random variable taking the same value as $X$ but with probabilities $p_s(x)$, show that

$$
\phi'(s) = E[X_s].
$$

(c) Show that

$$
\phi''(s) = \text{Var}(X_s) := E[X_s^2] - E[X_s]^2
$$

and conclude that $\phi''(s) \geq 0$ and the inequality is strict except when $X$ is deterministic.

(d) Let $x_{\min} := \min\{x : p(x) > 0\}$ and $x_{\max} := \max\{x : p(x) > 0\}$ be the smallest and largest values $X$ takes. Show that

$$
\lim_{s \to -\infty} \phi'(s) = x_{\min}, \quad \text{and} \quad \lim_{s \to \infty} \phi'(s) = x_{\max}.
$$

Problem 4: Hoeffding’s Lemma

Prove Lemma 2.3 in the lecture notes. In other words, prove that if $X$ is a zero-mean random variable taking values in $[a, b]$ then

$$
\mathbb{E}[e^{\lambda X}] \leq e^{\frac{\lambda^2}{2}(a-b)^2/4}.
$$

Expressed differently, $X$ is $[(a-b)^2/4]$-subgaussian.

Hint: You can use the following steps to prove the lemma:
1. Let $\lambda > 0$. Let $X$ be a random variable such that $a \leq X \leq b$ and $\mathbb{E}[X] = 0$. By considering the convex function $x \to e^{\lambda x}$, show that

$$\mathbb{E}[e^{\lambda X}] \leq \frac{b}{b-a}e^{\lambda a} - \frac{a}{b-a}e^{\lambda b}.$$  \hfill (8)

2. Let $p = -a/(b-a)$ and $h = \lambda(b-a)$. Verify that the right-hand side of (8) equals $e^{L(h)}$ where

$$L(h) = -hp + \log(1 - p + pe^h).$$

3. By Taylor’s theorem, there exists $\xi \in (0,h)$ such that

$$L(h) = L(0) + hl'(0) + \frac{h^2}{2}L''(\xi).$$

Show that $L(h) \leq h^2/8$ and hence $\mathbb{E}[e^{\lambda X}] \leq e^{\lambda^2(b-a)^2/8}$.

**Problem 5: Expected Maximum of Subgaussians**

Let $\{X_i\}_{i=1}^n$ be a collection of $n$ $\sigma^2$-subgaussian random variables, not necessarily independent of each other. Let $Y = \max_{i \in \{1, 2, \ldots, n\}} X_i$. Prove that $\mathbb{E}[Y] \leq \sqrt{2\sigma^2 \log n}$. Hint: Recall that by Jensen, $e^{\lambda \mathbb{E}[X]} \leq \mathbb{E}[e^{\lambda X}]$. 
