Practical session 1

Recall the Chernoff-Hoeffding inequalities which were proved in the lecture.

Theorem 0.1 (Chernoff-Hoeffding inequalities)**.** *Let $X_1, \ldots, X_n$ be independent random variables such that $\Pr(X_i = 1) = \Pr(X_i = -1) = 1/2$ for every $1 \le i \le n$, and let $X = \sum_{i=1}^n X_i$. Then, for every $t > 0$ it holds that*

$$\Pr(X > t) < e^{-t^2/(2n)},$$
$$\Pr(X < -t) < e^{-t^2/(2n)}.$$

Exercise 1 A gambler plays the following game. In each round he flips two fair coins, independently of one another, and independently of the previous rounds. If both coins yield the same result (i.e., both are heads or both are tails), he gains 11 shekels, otherwise he loses 1 shekel. Let S_n be the total amount of shekels the gambler has after n rounds (assuming he started with 0 shekels). Let $p_n = \Pr(S_n \ge 0)$, be the probability that the gambler has a non-negative amount of shekels after n rounds.

1. Use Chebyshev's inequality to lower bound p_{1000}.
2. Use the Chernoff-Hoeffding inequalities to lower bound p_{1000}.

Solution

For every positive integer i, let X_i denote the amount of shekels the gambler gained in the ith round, that is,

$$X_i = \begin{cases} 11 & \text{if both coins yield the same result in the } i\text{th coin toss} \\ -1 & \text{otherwise.} \end{cases}$$

Then $S_n = \sum_{i=1}^n X_i$.

1. For every positive integer i it holds that

$$\mathbb{E}(X_i) = \frac{1}{2} \cdot 11 + \frac{1}{2} \cdot (-1) = 5,$$
$$\mathbb{E}\left(X_i^2\right) = \frac{1}{2} \cdot 11^2 + \frac{1}{2} \cdot (-1)^2 = 61,$$
$$\mathrm{Var}(X_i) = \mathbb{E}\left(X_i^2\right) - (\mathbb{E}(X_i))^2 = 36.$$

Therefore, by linearity of expectation it holds that

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1) = 5n.$$

Moreover, since all coin flips are mutually independent, it follows that

$$\mathrm{Var}(S_n) = n \cdot \mathrm{Var}(X_1) = 36n.$$

Therefore

$$\begin{aligned}
\Pr(S_{1000} \geq 0) &= 1 - \Pr(S_{1000} < 0) \\
&= 1 - \Pr(S_{1000} - \mathbb{E}(S_{1000}) < -\mathbb{E}(S_{1000})) \\
&\geq 1 - \Pr(|S_{1000} - \mathbb{E}(S_{1000})| > \mathbb{E}(S_{1000})) \\
&\geq 1 - \frac{\mathrm{Var}(S_{1000})}{(\mathbb{E}(S_{1000}))^2} \\
&= 1 - \frac{36000}{(5000)^2} \\
&= 0.99856,
\end{aligned}$$

where the first inequality is due to the fact that the event $\{|S_{1000} - \mathbb{E}(S_{1000})| > \mathbb{E}(S_{1000})\}$ contains the event $\{S_{1000} - \mathbb{E}(S_{1000}) < -\mathbb{E}(S_{1000})\}$, and hence its probability is not smaller, and the second inequality is due to Chebyshev's inequality.

2. For every i, let $Y_i = \frac{X_i - 5}{6}$, and let $Y = \sum_{i=1}^{1000} Y_i$. Then $\Pr(Y_i = 1) = \Pr(Y_i = -1) = 1/2$ for every i. Therefore

$$\begin{aligned}
\Pr(S_{1000} \geq 0) &= \Pr\left(\sum_{i=1}^{1000} X_i \geq 0\right) \\
&= \Pr\left(\sum_{i=1}^{1000} \frac{X_i - 5}{6} \geq -\frac{5000}{6}\right) \\
&= \Pr\left(\sum_{i=1}^{1000} Y_i \geq -\frac{5000}{6}\right) \\
&= 1 - \Pr\left(Y < -\frac{5000}{6}\right) \\
&\geq 1 - e^{-\frac{\left(\frac{5000}{6}\right)^2}{2000}} \\
&= 1 - e^{-\frac{3125}{9}},
\end{aligned}$$

where the first inequality is the Chernoff-Hoeffding inequality (Theorem 0.1).

Exercise 2 A fair coin is tossed n times, all coin tosses being mutually independent. Let X be the number of coin tosses whose outcome is heads. Prove that

$$\lim_{n \to \infty} \Pr\left(|X - n/2| \leq \sqrt{n \ln n}\right) = 1.$$

Solution

We will use the Chernoff-Hoeffding inequalities (Theorem 0.1). Let X_i be the indicator random variable for the event that the outcome of the ith coin toss is heads. Note that the X_i's are mutually independent and that $X = \sum_{i=1}^{n} X_i$. For every i let $Y_i = 2X_i - 1$, and let $Y = \sum_{i=1}^{n} Y_i$. Then the Y_i's are mutually independent and $\Pr(Y_i = 1) = \Pr(Y_i = -1) = 1/2$. It thus follows by Theorem 0.1 that

$$\begin{aligned}
\Pr\left(X > n/2 + \sqrt{n \ln n}\right) &= \Pr\left(\sum_{i=1}^{n} X_i > n/2 + \sqrt{n \ln n}\right) \\
&= \Pr\left(\sum_{i=1}^{n}(2X_i - 1) > 2 \cdot \sqrt{n \ln n}\right) \\
&= \Pr\left(Y > 2 \cdot \sqrt{n \ln n}\right) \\
&\leq e^{-(2\cdot\sqrt{n \ln n})^2/(2n)} \\
&= n^{-2}.
\end{aligned}$$

Similarly, it holds that

$$\Pr\left(X < n/2 - \sqrt{n \ln n}\right) \leq n^{-2}.$$

Since the events $\left\{X > n/2 + \sqrt{n \ln n}\right\}$ and $\left\{X < n/2 - \sqrt{n \ln n}\right\}$ are disjoint, it follows that

$$\Pr\left(|X - n/2| > \sqrt{n \ln n}\right) = \Pr\left(X > n/2 + \sqrt{n \ln n}\right) + \Pr\left(X < n/2 - \sqrt{n \ln n}\right) \leq 2n^{-2}.$$

We conclude that

$$1 \geq \lim_{n\to\infty} \Pr\left(|X - n/2| \leq \sqrt{n \ln n}\right) \geq 1 - \lim_{n\to\infty} 2n^{-2} = 1$$

and so

$$\lim_{n\to\infty} \Pr\left(|X - n/2| \leq \sqrt{n \ln n}\right) = 1$$

as claimed.