Problem 1  

**Strong Law of Large Numbers**

In this problem you will prove the strong law of large numbers, which can be stated as follows:

**Theorem 1 (The Strong Law of Large Numbers).** Let \( \{X_k\}_{k \geq 1} \) be a sequence of independent and identically distributed (IID) random variables that are integrable (i.e., \( E[|X_k|] < \infty \); usually this is written as \( X_k \in L_1 \)), and let \( \mu = E[X_k] \) be their expectation.\(^1\) Then almost surely

\[
\frac{1}{n} \sum_{k=1}^{n} X_k \rightarrow \mu \quad \text{as} \ n \rightarrow \infty.
\]

or, to put it more mathematically,

\[
\Pr \left[ \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n} X_k = \mu \right] = 1.
\]

Before we start with the proof we introduce some notation and review some mathematical facts. We will use the following notation: capital letters denote random quantities, whereas small letters and Greek letters denote deterministic quantities. This rule is broken when dealing with sets: sets are described using capital letters, however, of a different font, e.g., \( A \). Furthermore, we define

\[
X_* \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n} X_k,
\]

\[
X^* \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n} X_k.
\]

a) Review of \( \lim \) and \( \lim_{n \rightarrow \infty} \): Let \( Z_k = (-1)^k (1 - e^{-k}) \), for \( k \geq 0 \).

i) Roughly plot \( Z_k \).

ii) What is the value of \( \lim_{k \rightarrow \infty} Z_k \)?

iii) What is the value of \( \overline{\lim}_{k \rightarrow \infty} Z_k \)?

iv) What is the value of \( \lim_{k \rightarrow \infty} Z_k \)?

b) Review of Cesáro Mean: Prove the following lemma:

\(^1\)Since \( X_k \) is integrable, it follows that \( \mu \) is finite.
Lemma 2 (Cesáro Mean). Let \( a_n \) be a sequence that converges to \( a \):

\[
a_n \to a \quad \text{as } n \to \infty.
\]

Then

\[
b_n \triangleq \frac{1}{n} \sum_{k=1}^{n} a_k \to a \quad \text{as } n \to \infty.
\]

We are now ready to start with the proof. In the following you will be guided through the proof step-by-step.

c) Show that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n+\ell} X_k = X^* \tag{3}
\]

independently of the choice of \( \ell \geq 1 \).

This shows that \( X^* \) is independent of \( X_1, X_2, \ldots, X_\ell \), and since this holds for any \( \ell \), that \( X^* \) is in fact independent of itself. Consequently, \( X^* \) must be constant almost surely. A similar argument shows that also \( X_* \) must be constant almost surely. Hence, it only remains to show that these two constants are the same and are equal to \( \mu \).

d) We next try to reduce the problem. Justify the following statements:

i) We only need to consider nonnegative random variables \( X_k \).

ii) To establish \( X_* = X^* \) almost surely, it is sufficient to prove

\[
X_* \geq X^* \quad \text{a.s.} \tag{4}
\]

iii) We may assume that \( E[X^*] > 0 \).

iv) Argue that to establish (4) it is sufficient to prove

\[
E[X_*] \geq E[X^*]. \tag{5}
\]

e) Show that in order to prove (2), it is sufficient to prove that

\[
E[X_*] \geq \mu \geq E[X^*]. \tag{6}
\]

f) In this part, assuming that it is true that for any nonnegative sequence \( \{X_k\} \) we have \( E[X_1] \geq E[X^*] \), we prove that then it follows that \( E[X_*] \geq E[X_1] \).

To that goal, define

\[
Y_k \triangleq m - \min\{X_k, m\}
\]

for any given \( m > 0 \), and note that \( \{Y_k\} \) is nonnegative. By assumption, we thus have \( E[Y_1] \geq E[Y^*] \).

i) Show that \( Y^* = m - (\min\{X, m\})_* \).

ii) Relying on the assumption that \( E[Y_1] \geq E[Y^*] \) show that

\[
E[X_*] \geq E\left[ (\min\{X, m\})_* \right] \geq E[\min\{X_1, m\}].
\]

iii) Let \( m \to \infty \) and argue that—given that it is true that \( E[X_1] \geq E[X^*] \) for every nonnegative sequence \( \{X_k\} \)—we have \( E[X_*] \geq E[X_1] \).
Hence, we do not need to show (6), but may restrict ourselves to the proof of
\[
\mu = \mathbb{E}[X_1] \geq \mathbb{E}[X^*]
\] (7)
for every nonnegative sequence \(\{X_k\}\).

g) Argue that proving (7) is equivalent to showing that
\[
\lim_{n \to \infty} \mathbb{E} \left[ \frac{1}{n} \sum_{k=1}^{n} X_k \right] \geq \alpha, \quad \text{for every constant } 0 < \alpha < \mathbb{E}[X^*].
\] (8)

h) We now turn to the proof of (8). Using the fact that \(X^*\) is constant almost surely, show that
any \(\alpha < \mathbb{E}[X^*]\) almost surely gives a lower bound on \(\frac{1}{n} \sum_{k=1}^{n} X_k\) for infinitely many \(n\).

i) Let us for the moment fix \(n \geq 1\). The basic idea is to construct a (random) collection \(\mathcal{A}_1, \ldots, \mathcal{A}_J\) of disjoint intervals \(\mathcal{A}_j = \{K_j, K_j + 1, \ldots, K_j + N_j - 1\} \subseteq \{1, \ldots, n\}\) starting at \(K_j\) and having length \(N_j\) such that the mean \(\frac{1}{N_j} \sum_{k=K_j}^{K_j+N_j-1} X_k\) of the values observed during \(\mathcal{A}_j\) is no smaller than \(\alpha\). (This is possible due to the fact proven in Part h.) Moreover, we ask these intervals to be the shortest ones, i.e., if an interval \(\mathcal{A}\) starts at index \(k\), then its length shall equal
\[
L(k) \triangleq \min \left\{ r \geq 1 : \frac{1}{r} \sum_{i=1}^{r} X_{k+i-1} \geq \alpha \right\}.
\] (9)

Let’s call such intervals nice. To start with, let \(\mathcal{A}_1\) be the first, (i.e., leftmost) nice interval contained in \(\{1, \ldots, n\}\). If there is none, we have \(J = 0\). If \(\mathcal{A}_1, \ldots, \mathcal{A}_{j-1}\) have been found, let \(\mathcal{A}_j\) be the first nice interval to the right of \(\mathcal{A}_{j-1}\) which is contained in \(\{1, \ldots, n\}\). (Note that by construction the intervals \(\mathcal{A}_j\) do not necessarily need to be adjacent, but they try to be as close together as possible.) If there are no sets of this kind, we have \(J = j - 1\).

Study a simple example to understand the construction of these sets \(\mathcal{A}_j\): Let
\[(X_1, X_2, \ldots, X_n) = (5, 1, 3, 7, 2, 4, 3, 1),\]
i.e., we have \(n = 8\), and choose \(\alpha = 3.1\). Now list the sets \(\mathcal{A}_j\) and all indices that are not part of \(\mathcal{A}_1 \cup \cdots \cup \mathcal{A}_J\).

j) Let
\[
N \triangleq n - \sum_{j=1}^{J} N_j
\]
be the number of indices in \(\{1, \ldots, n\}\) not covered by the set \(\mathcal{A}_1 \cup \cdots \cup \mathcal{A}_J\). Argue that \(k \notin \mathcal{A}_1 \cup \cdots \cup \mathcal{A}_J\) implies \(L(k) > n - k + 1\).

k) Justify each step in the following chain of equalities:
\[
\mathbb{E}[N] \leq \mathbb{E} \left[ \sum_{k=1}^{n} \mathcal{I}\{L(k) > n - k + 1\} \right] \quad \text{(10)}
\]
\[
= \sum_{k=1}^{n} \mathbb{E}[\mathcal{I}\{L(k) > n - k + 1\}] \quad \text{(11)}
\]
\[
= \sum_{k=1}^{n} \Pr[L(k) > n - k + 1]. \quad \text{(12)}
\]

Note that \(\mathcal{I}\{\}\) denotes the indicator function whose value is 1 if the statement is correct and 0 otherwise.
l) Justify each step in the following chain of (in)equalities:

\[
\frac{1}{n} \sum_{k=1}^{n} X_k \geq \frac{1}{n} \sum_{k \in A_1 \cup \cdots \cup A_J} X_k \tag{13}
\]

\[
= \frac{1}{n} \sum_{j=1}^{J} N_j \cdot \frac{1}{N_j} \sum_{k \in A_j} X_k \tag{14}
\]

\[
\geq \frac{1}{n} \sum_{j=1}^{J} N_j \cdot \alpha \tag{15}
\]

\[
= \alpha - \frac{1}{n} \left( n - \sum_{j=1}^{J} N_j \right) \alpha \tag{16}
\]

\[
= \alpha - \frac{N \alpha}{n}. \tag{17}
\]

m) Argue that

\[
\lim_{n \to \infty} \Pr[L(1) > n] = 0. \tag{18}
\]

n) Finally, using stationarity, Lemma 2, (12), (17), and (18), justify each step in the following chain of (in)equalities:

\[
E \left[ \frac{1}{n} \sum_{k=1}^{n} X_k \right] \geq E \left[ \alpha - \frac{N \alpha}{n} \right] \tag{19}
\]

\[
= \alpha - \frac{E[N] \alpha}{n} \tag{20}
\]

\[
\geq \alpha - \frac{\alpha}{n} \sum_{k=1}^{n} \Pr[L(k) > n - k + 1] \tag{21}
\]

\[
= \alpha - \frac{\alpha}{n} \sum_{k=1}^{n} \Pr[L(1) > n - k + 1] \tag{22}
\]

\[
= \alpha \left( 1 - \frac{1}{n} \sum_{i=1}^{n} \Pr[L(1) > i] \right) \tag{23}
\]

\[
\to \alpha, \tag{24}
\]

which proves (8) as we have intended.