Problem 1

Let \((X_n)\) be a sequence of independent RVs such that \(X_1 = 0\), and for \(n \geq 2\),
\[
P(X_n = n) = P(X_n = -n) = f(n); \quad P(X_n = 0) = 1 - 2f(n),
\]
where \(f(n), 0 < f(n) < 1\) is some non-constant function of \(n\) (such as \(f(n) = \frac{1}{n}\) or similar). Further, let \(S_n = \frac{1}{n} \sum_{i=1}^{n} X_i\) be the empirical mean of the sequence.

(a) Give an example of \(f(n)\) such that \(S_n \xrightarrow{P} 0\) (converges to 0 in probability).
(b) Give an example of \(f(n)\) such that \(S_n\) converges almost surely.
(c) Give an example of \(f(n)\) such that \(S_n\) converges in probability and does not converge almost surely.

(a) We can take \(f(n) = \frac{1}{n}\), then \(E X_n = 0\)
\[
\text{Var}(X_n) = E X_n^2 = 2 n^2 \frac{1}{n} = \frac{2}{n}
\]
\[
\text{Var}(S_n) = \frac{1}{n^2} \cdot n \text{Var}(X_n) = \frac{2}{n^2}
\]
By Chebyshev's inequality,
\[
P(|S_n| > \varepsilon) \leq \frac{\text{Var}(S_n)}{\varepsilon^2} = \frac{2}{n^2 \varepsilon^2} \to 0 \quad n \to \infty
\]
i.e. \(S_n \xrightarrow{P} 0\)

(b) Note that for \(S_n\) to be small, \(X_n\) cannot be large i.o.
In other words, if \(P(|X_n| > \varepsilon \text{ i.o.}) = 0\), then
\[
P(S_n \to 0) = 1.
\]
Take \(f(n) = \frac{1}{2n^k}\), then
\[
\sum_{n=1}^{\infty} P(X_n \neq 0) = \sum_{n=1}^{\infty} \frac{1}{n^k} < \infty \quad \text{for any } k > 1
\]
providing the required examples.

(c) Take \(f(n) = \frac{1}{n \log n}\), then, arguing as in Part(a),
\[
P(|S_n| > \varepsilon) \leq \frac{\text{Var}(S_n)}{\varepsilon^2} = \frac{\sum_{i=1}^{n} \frac{2\varepsilon}{\log i}}{n^2 \varepsilon^2} \leq \frac{2n^2}{n^2 \log n} \to 0
\]
so \(S_n \xrightarrow{P} 0\)
At the same time, now \(X_n \neq 0\) i.o., implying that
\(|S_n| \geq 1\) i.o., i.e. \(S_n \xrightarrow{P} 0\) a.s. Since the limit, if it exists, must be 0 by uniqueness, this finishes the proof.
Problem 2. In our discussion of the monotone convergence theorem we proved that under some assumptions on RVs \( (X_n)_n \), it is true that \( E \sum_{n=1}^{\infty} X_n = \sum_{n=1}^{\infty} E X_n \). Give an example of an infinite sequence of RVs \( (X_n)_n \) for which this equality fails to hold (obviously, they do not satisfy the mentioned assumptions).

\[
\text{In fact, if } \; X_i \geq 0, \text{ then } \; E \sum_{i=1}^{\infty} X_i = \sum_{i=1}^{\infty} E X_i \; \text{(i.e., the series can be integrated term-wise). This means that our RVs have to take both positive and negative values.}
\]

Let \( \xi \) be the number of tosses of a fair coin till the 1st H, and let \( X_i = \begin{cases} 2^i & \text{if } i < \xi, \\ 2^i & \text{if } i = \xi, \\ 0 & \text{if } i > \xi. \end{cases} \)

Then
\[
p(\xi > i+1) = \frac{1}{2^i}; \quad p(\xi = i) = \frac{1}{2^i}; \quad p(\xi < i) = 1 - \frac{1}{2^i}
\]

\[
E X_i = 2^i \cdot \frac{1}{2^i} + 2^i \cdot \frac{1}{2^i} + 0 = 0
\]

\[
\sum_{i=1}^{\infty} E X_i = 0
\]

At the same time, for any value \( n \) s.t. \( \xi = n \),
\[
\sum_{i=1}^{n} X_i = (-2 + 2^2 - \ldots + 2^n) + 2^n = -2(2^n - 2) + 2^n = 2^n - 2
\]

and thus \( p(\sum_{i=1}^{n} X_i = 2) = 1 \), so
\[
E \sum_{i=1}^{\infty} X_i = 2.
\]

Problem 3. Let \( X_1, X_2, \ldots, X_n \) be independent, identically distributed (i.i.d.) RVs such that \( E[X_i^{-1}] \) exists, and let \( S_m = X_1 + \cdots + X_m \). Prove that if \( m \leq n \), then \( E[S_m] = \frac{m}{n} \).

Denote \( T_i = E \left[ \frac{X_i}{\sum_{j=1}^{n} X_j} \right], \) \( i = 1, 2, \ldots, n \). By linearity of expectation, we need to compute \( \sum_{i=1}^{m} T_i \). Since \( X_i \) are identically distributed, \( T_i \) does not depend on \( i \). Now note that
\[
l = E \left[ \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i} \right] = \sum_{i=1}^{n} T_i, \quad \text{so } T_i = \frac{1}{n} \text{ for all } i, \text{ and } \sum_{i=1}^{m} T_i = \frac{m}{n}
\]

for any \( 1 \leq m \leq n \).
Problem 4. Let $X_1, X_2, X_3$ be i.i.d. RVs with pdf $f(x) = e^{-x}$ for $x \geq 0$ and $f(x) = 0$ o/w, i.e., $X_i \sim \text{Exp}(1)$ for all $i = 1, 2, 3$. Show that the RVs

$$Y_1 = \frac{X_1}{X_1 + X_2}, \quad Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3}, \quad \text{and} \quad Y_3 = X_1 + X_2 + X_3$$

are independent.

The joint pdf of $X_1, X_2, X_3$ is

$$f_{X_1,X_2,X_3}(x_1, x_2, x_3) = e^{-x_1-x_2-x_3}, \quad x_1, x_2, x_3 > 0$$

Now recall the transformation rule of pdf's:

$$y_1 = \frac{x_1}{x_1 + x_2}, \quad x_1 = y_1 y_2 y_3$$

$$y_2 = \frac{x_1 + x_2}{x_1 + x_2 + x_3}, \quad x_2 = (1 - y_1) y_2 y_3$$

$$y_3 = x_1 + x_2 + x_3, \quad x_3 = (1 - y_2) y_3$$

The Jacobian

$$J(y_1, y_2, y_3) = \left| \frac{\partial(x_1, x_2, x_3)}{\partial(y_1, y_2, y_3)} \right| = \begin{vmatrix} y_2 y_3 & -y_2 y_3 & 0 \\ y_1 y_3 & (-y_1) y_3 & -y_3 \\ y_1 y_2 & (1 - y_1) y_2 & 1 - y_2 \end{vmatrix}$$

$$= y_2 y_3 (1 - y_2)(1 - y_2) + y_1 y_2 y_3^2 + y_2 y_3^2 (-y_1) + y_1 y_2 y_3 (1 - y_2) = y_2 y_3^2$$

Then

$$f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) = f_{X_1,X_2,X_3}(x_1, x_2, x_3) |J(y_1, y_2, y_3)| = e^{-(x_1+x_2+x_3)} y_2 y_3^2$$

$$= 1 \cdot 2y_2 \cdot \frac{y_3^2 e^{-y_3}}{(y_3 e^{-y_3})^2}, \quad 0 \leq y_1, y_2 \leq 1; \quad 0 < y_3 < \infty$$

Each of the terms (a), (b), (c) integrates to one, so they form the marginal pdf's of $Y_1, Y_2,$ and $Y_3$, respectively. Since $f_{Y_1,Y_2,Y_3}$ factorizes into a product of the marginals, the RVs $Y_1, Y_2, Y_3$ are independent by definition.

Note: $Y_1$ is a uniform RV on [0,1], $Y_3$ follows the gamma distribution of order 3.
Problem 5. Let $\Omega = \mathbb{R}$ and $\mathcal{B}(\mathbb{R})$ be the Borel $\sigma$-algebra. Fix a number $\epsilon \in \mathbb{R}$ and for all $A \in \mathcal{B}(\mathbb{R})$,

$$\delta_{\epsilon}(A) = \begin{cases} 1 & \text{if } \epsilon \in A \\ 0 & \text{if } \epsilon \notin A \end{cases}.$$ 

(a) Show that $\delta_{\epsilon}$ is a measure, and that $(\Omega, \mathcal{B}(\mathbb{R}), \delta_{\epsilon})$ is a probability space. This measure is in fact called the delta function at $\epsilon$ (an impulse, as an engineer would say).

(b) Let $f : \mathbb{R} \to \mathbb{R}$ be a nonnegative measurable function. Show that $\int f \, d\delta_{\epsilon} = f(\epsilon)$.

(c) For $\epsilon \in \mathbb{R}$ define $\mu(\epsilon) = \sum_{n=1}^{\infty} \delta_{\epsilon}(A_n)$ to be the counting measure (a.k.a. an impulse train), $\mu(A)$ is simply the number of natural numbers in $A$, which can of course be finite or infinite. For a nonnegative measurable function $f$ show that $\int f \, d\mu = \sum_{n=1}^{\infty} f(n)$. This shows that a sum is simply an integral against a counting measure.

(a) $\delta_{\epsilon} \geq 0$; consider a collection of pairwise disjoint sets $(A_n)_n$. Then either \( \exists n : \epsilon \in A_n \), and $\delta_{\epsilon}(\bigcup A_n) = \sum_{n} \delta_{\epsilon}(A_n) = 1$ or

there is no such $n$, and $\delta_{\epsilon}(\bigcup A_n) = 0$.

Thus, $\delta_{\epsilon}$ is nonnegative and countably additive \equiv measure. The distribution function associated with $\delta_{\epsilon}$ is

$$F(y) = 0, y < \epsilon; \quad F(y) = 1, y \geq \epsilon,$$

which is a valid CDF.

(b) Let $g$ be a constant function s.t. $g(y) = f(\epsilon)$ for all $y \in \mathbb{R}$.

Then \( g = f \) a.e. with respect to $\delta_{\epsilon}:

$$\delta_{\epsilon}(\{ g = f \}) = \delta_{\epsilon}(\{ y \in \mathbb{R} : g(y) = f(\epsilon) \}) = 1$$

since this set contains $\epsilon$.

Functions that are identical a.e., have the same integral:

$$f = g \text{ almost everywhere with respect to the measure } \delta_{\epsilon}.$$ 

\( \Rightarrow \int f \, d\delta_{\epsilon} = \int g \, d\delta_{\epsilon} = f(\epsilon) \int 1 \, d\delta_{\epsilon} = f(\epsilon) \delta_{\epsilon}(\mathbb{R}) = f(\epsilon). \)

Another way to show this: Take a simple function $g = \sum_{i} c_i 1_{A_i}$, then

$$\int g \, d\delta_{\epsilon} = \sum_{i} c_i \delta_{\epsilon}(A_i) = \sum_{i} c_i 1_{A_i}(\epsilon) = g(\epsilon).$$

From here, we can extend this claim to all nonnegative functions and then to all measurable functions.

(c) Again, start with a simple function $g = \sum_{i} c_i 1_{A_i}$, then our claim is obvious (for every $A_i$, $c_i$ will appear $\delta_{\epsilon}(A_i)$ times).
Problem 6. (a) Given a CDF \( F \), we denote \( F(x^-) = \lim_{y \to x^-} F(y) \) and similarly for \( F(x^+) \). From the definition of \( F \) we know that \( F(x^-) \neq F(x^+) \) for all \( x \in \mathbb{R} \). If \( F(x^-) \neq F(x) \), then \( F(x) \) has a discontinuity at \( x \). Your task is to show that the number of points at which \( F(x) \) has a discontinuity is at most countably infinite. One option is to use the relation \( \{ x : F(x) \neq F(x^-) \} = \bigcup_{n=1}^\infty \{ x : F(x) - F(x^-) \geq \frac{1}{n} \} \).

(b) Given a CDF \( F \) and a number \( a \geq 0 \), show that
\[
\int_{-\infty}^\infty (F(x+a) - F(x))dx = a.
\]

(a) It suffices to prove the statement for the interval \([0,1]\) because \( \mathbb{R} \) can be written as a countable union of unit intervals.
First, note that \( S_n = \{ x \in [0,1] : \frac{1}{n} \leq F(x) - F(x^-) \} \) is finite.
This is obvious because \( \sum_{x \in S_n} (F(x) - F(x^-)) = \infty \), but \( F(x) \leq 1 \) for all \( x \),
a contradiction.
Next, the set \( \bigcup_{n=1}^\infty S_n \) contains all points of discontinuity of \( F(x) \),
and is countable because it is a countable union of finite sets.

(b) For a fixed \( a > 0 \), write \( \int_{-\infty}^\infty \left[ F(x+a) - F(x) \right]dx \)
as a limit of an integral over a finite interval. Namely,
\[
\int_{-\infty}^{z_2} (F(x+a) - F(x))dx = \int_{-\infty}^{z_1} F(x+a)dx - \int_{-\infty}^{z_1} F(x)dx = \int_{-\infty}^{z_1} F(x)dx - \int_{-\infty}^{z_1} F(x)dx
\]
\[
= \int_{z_2}^{z_2+a} F(x)dx - \int_{z_1}^{z_1+a} F(x)dx = \int_{0}^{a} F(x+z_2)dx - \int_{0}^{a} F(x-z_1)dx.
\]

Now let \( z_1, z_2 \to \infty \):
\[
\lim_{z_2 \to \infty} \int_{z_1}^{a} F(x+z_2)dx = \int_{0}^{a} F(x+z_2)dx \quad \text{by bounded convergence}
\]
\[
= a \quad \text{by bounded convergence}
\]
\[
\lim_{z_1 \to \infty} \int_{0}^{a} F(x-z_1)dx = \int_{0}^{a} \lim_{z_1 \to \infty} F(x-z_1)dx = 0.
\]
and thus \( \int_{-\infty}^{\infty} (F(x+a) - F(x)) \, dx = \lim_{z_1 \to \infty} \lim_{z_2 \to \infty} \int_{-z_1}^{z_2} (F(x+a) - F(x)) \, dx = a. \)