Math 711 Homework

Austin Mohr

June 16, 2012

Problem 1

Proposition 1. Convergence in probability of a sequence of random variables does not imply almost sure convergence of the sequence.

Proof. Let \(([0,1], \mathcal{B}([0,1]), \lambda)\) be the ambient probability space. Consider a process wherein, at the \(n\)th stage, we choose (independently at random) a closed subinterval \(F_n\) of \([0,1]\) of length \(\frac{1}{n}\).

Define now the random variable \(X_n = I(F_n)\). Evidently, \(X \xrightarrow{pr} 0\), since, for each \(n\), \(X_n\) is nonzero only on the interval of length \(\frac{1}{n}\). Thus, for any \(\epsilon > 0\),

\[
P(|X_n| > \epsilon) = \frac{1}{n} \to 0.
\]

To see that \(X_n\) does not enjoy almost sure convergence to zero, observe that, for each \(\omega \in [0,1]\),

\[
\sum_{n=1}^{\infty} P(X_n(\omega) \neq 0) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.
\]

Thus, by the Borel Zero-One Law, \(P(X_n(\omega) \neq 0 \text{ i.o.}) = 1\), and so it cannot be that \(X_n\) converges almost surely to zero. \(\square\)

Problem 2

Proposition 2. For any sequence \(\{X_n\}\) of random variables, there exists a sequence of constants \(\{a_n\}\) such that \(\frac{X_n}{a_n}\) converges almost surely to zero.

Proof. Let \(\epsilon > 0\) be given. Choose \(a_n\) such that \(P(|X_n| > \epsilon) \leq \frac{1}{2^n}\). Observe that

\[
\sum_{n=1}^{\infty} P(|X_n| > \epsilon) \leq \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2^n} < \infty,
\]
and so the Borel-Cantelli Lemma gives that \( P(\{|X_n| > \epsilon\} \text{ i.o.}) = 0 \). That is, \( X_n \overset{a.s.}{\rightarrow} 0 \).

**Problem 3**

**Proposition 3.** For a monotone sequence of random variables, convergence in probability implies almost sure convergence.

**Proof.** Let \( X_n \) be a monotone sequence of random variables with \( X_n \overset{\text{pr}}{\rightarrow} X \). It follows that, for all \( \epsilon > 0 \),

\[
0 = \lim_{n \to \infty} P(|X_n - X| > \epsilon)
\]

\[
= \lim_{N \to \infty} P \left( \bigcup_{n \geq N} [|X_n - X| > \epsilon] \right) \quad \text{(by monotonicity of the } X_n) 
\]

\[
= P \left( \limsup_{n \to \infty} [|X_n - X| > \epsilon] \right) 
\]

\[
= P(|X_n - X| > \epsilon \text{ i.o.}).
\]

Therefore, \( X_n \overset{a.s.}{\rightarrow} X \). \( \square \)

**Problem 4**

**Proposition 4.** Let \( \{X_n\} \) be a sequence of random variables and define, for each \( n \), \( Y_n = X_n I\{|X_n| \leq a_n\} \). There exists a sequence \( \{a_n\} \) of positive real numbers such that \( P\{|X_n \neq Y_n \text{ i.o.}\} = 0 \).

**Proof.** For each \( n \), choose \( a_n \) such that \( P\{|X_n| > a_n\} \leq \frac{1}{2^n} \). Thus,

\[
\sum_{n=1}^{\infty} P(X_n \neq Y_n) = \sum_{n=1}^{\infty} P(|X_n| > a_n)
\]

\[
\leq \sum_{n=1}^{\infty} \frac{1}{2^n}
\]

\[
< \infty.
\]

By the Borel-Cantelli Lemma, we conclude that \( P(|X_n \neq Y_n \text{ i.o.}) = 0 \). \( \square \)

**Problem 5**

**Proposition 5.** Suppose \( n \) points are chosen randomly on the unit circle. Define the random variable \( X_n \) to be the arc length of the largest arc not containing any of the chosen points. In this case, \( X_n \to 0 \) almost surely.
Proof. Let $\epsilon > 0$ be given. We have that

$$P([X_n > \epsilon] \ i.o.) = P\left(\limsup_{n \to \infty} [X_n > \epsilon]\right)$$

$$= \lim_{N \to \infty} P\left(\bigcup_{n \geq N} [X_n > \epsilon]\right)$$

$$= \lim_{n \to \infty} P(X_n > \epsilon) \quad \text{(by monotonicity of the } X_n).$$

We bound this last probability by breaking the unit circle into $4\pi/\epsilon$ disjoint intervals of length $\epsilon/2$. Thus, $P(X_n \leq \epsilon)$ is no larger than the probability of having a point contained in every interval. Thus,

$$P([X_n > \epsilon] \ i.o.) = \lim_{n \to \infty} P(X_n > \epsilon)$$

$$\leq \lim_{n \to \infty} \frac{4\pi}{\epsilon} \left(\frac{2\pi - \frac{\epsilon}{2}}{2\pi}\right)^n$$

$$= 0,$$

and so $X_n \overset{a.s.}{\to} 0.$

\begin{proof}{Problem 6}
Proposition 6. If, for all $a < b$,

$$P\{\{X_n < a\} \ i.o. \text{ and } \{X_n > b\} \ i.o.\} = 0,$$

then $\lim_{n \to \infty} X_n$ exists almost surely.

Proof. By taking complements, we have

$$P\{X_n \geq a \text{ or } X_n \leq b\} = 1,$$

for all sufficiently large $n$. If it is the case that the former holds for all $a$, then $\lim_{n \to \infty} X_n = \infty$ almost surely. Similarly, if the latter holds for all $b$, then $\lim_{n \to \infty} X_n = -\infty$ almost surely. Otherwise, there is some $c$ so that $P(X_n \leq b) = 1$ for all $b \geq c$ and $P(X_n \leq b) = 0$ for all $b < c$. By keeping $b = c$ fixed and letting $a \uparrow b$, we see that $\lim_{n \to \infty} X_n = c$ almost surely.
\end{proof}

\begin{proof}{Problem 7}
Proposition 7. If $X_n \to 0$ in probability, then, for any $\alpha > 0$,

$$\frac{|X_n|^\alpha}{1 + |X_n|^\alpha} \overset{pr}{\to} 0.$$

3
Proof. Let ϵ > 0 be given. Choose N such that

\[ P( |X_n| > \epsilon^{\frac{1}{\alpha}} ) < \epsilon, \]

for all n ≥ N. It follows that

\[ P( |X_n| > \epsilon ) < \epsilon \]

for all n ≥ N. Now,

\[ \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} \leq |X_n|. \]

Thus, for all n ≥ N,

\[ P \left( \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} > \epsilon \right) \leq P(|X_n|^\alpha > \epsilon) < \epsilon, \]

and so \[ \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} \overset{pr}{\to} 0. \]

Problem 8

Proposition 8. If, for some \( \alpha > 0 \),

\[ \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} \overset{pr}{\to} 0, \]

then \( X_n \overset{pr}{\to} 0 \).

Proof. Let ϵ > 0 be given. Choose N such that

\[ P \left( \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} > \epsilon \right) < \epsilon \]

for all n ≥ N. Now,

\[ \left( \frac{|X_n|}{1 + |X_n|} \right)^\alpha \leq \frac{|X_n|^\alpha}{1 + |X_n|^\alpha}. \]

Thus, for all n ≥ N,

\[ P \left( \left( \frac{|X_n|}{1 + |X_n|} \right)^\alpha > \epsilon \right) \leq P \left( \frac{|X_n|^\alpha}{1 + |X_n|^\alpha} \right) < \epsilon, \]

and so

\[ P \left( \frac{|X_n|}{1 + |X_n|} > \epsilon^{\frac{1}{\alpha}} \right) < \epsilon, \]

from which it follows that \( X_n \overset{pr}{\to} 0 \). \qed
Problem 9

Proposition 9. Let \( \{X_n\} \) be a collection of independent random variables with
\[
P(X_n = n^2) = \frac{1}{n^2} \quad \text{and} \quad P(X_n = -1) = 1 - \frac{1}{n^2}
\]
for all \( n \). In this case, \( \sum_{i=1}^n X_i \) converges almost surely to \(-\infty\) as \( n \to \infty \).

Proof. Observe first that, by definition of the \( X_n \), \( P(X_n \in \{n^2, -1\}) = 1 \). That is, except for a null set, \( X_n \) takes on only the values \( n^2 \) or \(-1\).

Now, we have
\[
\sum_{n=1}^{\infty} P(X_n = n^2) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,
\]
and so \( P([X_n = n^2] \text{ i.o.}) = 0 \) by the Borel-Cantelli Lemma.

Similarly,
\[
\sum_{n=1}^{\infty} P(X_n = -1) = \sum_{n=1}^{\infty} 1 - \frac{1}{n^2} = \infty,
\]
and so \( P([X_n = -1] \text{ i.o.}) = 1 \) by the Borel Zero-One Law (note here that we require the independence of the \( X_n \)).

Thus, for almost all \( \omega \in \Omega \), there exists \( N_\omega \) such that \( X_n(\omega) = -1 \) for all \( n \geq N_\omega \). It follows that
\[
\lim_{n \to \infty} S_n(\omega) = \lim_{n \to \infty} \sum_{i=1}^{n} X_i(\omega)
\]
\[
= \sum_{i=1}^{N_\omega} X_i(\omega) + \sum_{i>N_\omega} X_i(\omega)
\]
\[
\leq \sum_{i=1}^{N_\omega} n^2 + \sum_{i>N_\omega} -1
\]
\[
= -\infty,
\]
and so \( S_n \xrightarrow{a.s.} -\infty \). \( \square \)

Problem 1

Proposition 10. Let \( \mathbb{E}(X^2) = 1 \) and \( \mathbb{E}(|X|) \geq a > 0 \). For \( 0 \leq \lambda \leq 1 \),
\[
P(|X| \geq \lambda a) \geq (1 - \lambda)^2 a^2.
\]
Proof. Let $A$ denote the set \{\(|X| \geq \lambda a\)\}. We have
\[
\mathbb{E}(|X|) = \mathbb{E}(|X|_A) + \mathbb{E}(|X|_{A^c}).
\]
Now, applying Hölder’s inequality to the first term,
\[
\mathbb{E}(|X|_A) \leq \sqrt{\mathbb{E}(X^2) \mathbb{E}(1^2_A)} = \sqrt{P(A)}.
\]
We have also that
\[
\mathbb{E}(|X|_{A^c}) < \lambda a.
\]
Taken together, we see that
\[
a \leq \mathbb{E}(|X|) \leq \sqrt{P(A)} + \lambda a.
\]
Rearranging terms gives,
\[
\sqrt{P(A)} \geq a - \lambda a = a(1 - \lambda),
\]
and so $P(A) \geq a^2(1 - \lambda)^2$, as desired.

Problem 2

Proposition 11. For $1 \leq s \leq t < \infty$ and $X \in L_t$, $\|X\|_s \leq \|X\|_t$ (where $\|X\|_p = (\mathbb{E}(X^p))^{\frac{1}{p}}$).

Proof. There is nothing to show for the case $s = t$, so we assume that $s < t$. Let $\alpha = \frac{s}{t}$ and $\beta = \frac{t}{t-s}$, so that $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. By Hölder’s inequality,
\[
\mathbb{E}(|X|^s) = \mathbb{E}(|X|^{s} \cdot 1) \leq \mathbb{E}(|X|^s) \cdot \mathbb{E}(1^\frac{s}{s}) \leq \mathbb{E}(|X|^s) \cdot \mathbb{E}(1^\frac{t}{t}) = \mathbb{E}(|X|^t)^\frac{t}{s}.
\]
Thus,
\[
\|X\|_s = \mathbb{E}(|X|^s)^\frac{1}{s} \leq \left( \mathbb{E}(|X|^t)^\frac{1}{t} \right)^\frac{t}{s} = \mathbb{E}(|X|^t)^\frac{t}{s} = \|X\|_t < \infty,
\]
as desired.
Problem 3

For a random variable $X$, define
$$||X||_{\infty} = \sup \{ M : P\{|X| > M\} > 0 \}$$
and also
$$L_{\infty} = \{ X : ||X||_{\infty} < \infty \}.$$

Proposition 12. For $X \in L_{\infty}$ and $1 < p < q < \infty$,
$$0 \leq ||X||_p \leq ||X||_q \leq ||X||_{\infty}.$$  

Proof. Since $X \in L_{\infty}$, $X$ is bounded, and so $X \in L_t$ for any $t$. Thus, by Problem 2, it remains only to show that $||X||_q \leq ||X||_{\infty}$. Now,
$$||X||_q = (E(|X|^q))^{\frac{1}{q}} \leq (M^q)^{\frac{1}{q}} = M = ||X||_{\infty}.$$ 

Proposition 13. For $1 < p < q < \infty$, $L_{\infty} \subset L_q \subset L_p \subset L_1$.

Proof. We show, equivalently, that $X \in L_q$ implies $X \in L_p$ whenever $1 \leq p \leq q \leq \infty$. To that end, let $1 \leq p \leq q \leq \infty$ and let $X \in L_q$. Applying the previous result, we have
$$||X||_p \leq ||X||_q < \infty,$$
and so $X \in L_p$. 

Proposition 14. For $X,Y \in L_{\infty}$,
$$E(|XY|) \leq ||X||_1 ||Y||_{\infty}.$$  

Proof. Note first that, since $L_{\infty} \subset L_1$, $X \in L_1$. Now,
$$E(|XY|) \leq E(|XM|) \leq E(|X|) \cdot M \leq ||X||_1 ||Y||_{\infty}.$$ 

Proposition 15. For $X,Y \in L_{\infty}$,
$$||X + Y||_{\infty} \leq ||X||_{\infty} + ||Y||_{\infty}.$$
Proof. It follows immediately that
\[
\|X + Y\|_\infty = \sup\{M : P\{|X + Y| > M\} > 0\} \\
\leq \sup\{M : P\{|X| + |Y| > M\} > 0\} \\
\leq \sup\{M : P\{|X| > M\} > 0\} + \sup\{M : P\{|Y| > M\} > 0\} \\
= \|X\|_\infty + \|Y\|_\infty.
\]

\[\square\]

Problem 4

Proposition 16. Let \(X \in L_1\), the map
\[
\chi : [1, \infty] \to [0, \infty]
\]
defined according to \(\chi(p) = \|X\|_p\) is continuous on \([1, p_0)\), where
\[
p_0 = \sup\{p \geq 1 : \|X\|_p < \infty\}.
\]
Furthermore, the continuous function \(\eta(p) = \log (\|X\|_p^\gamma)\) is convex on \([1, p_0)\).

Proof. For the continuity of \(\chi\), consider any sequence \(\{p_n\}\) converging to \(p \in [0, p_0)\) and set \(\xi = \|X\|_p + 1\). Note that \(\|X\|_p < \infty\). Since \(|X|^{p_n} \to |X|^p\) pointwise and \(|X|^{p_n} \leq \xi\) for all \(n\), the Lebesgue dominated convergence theorem gives
\[
E(X^{p_n}) \to E(X^p).
\]
Since the operation of exponentiation is continuous, it follows that
\[
(E(X^{p_n}))^{\frac{1}{p_n}} \to (E(X^p))^\frac{1}{p}.
\]
That is, \(\|X\|_{p_n} \to \|X\|_p\), and so \(\chi\) is continuous in \(p\).

For the convexity of \(\eta\), let \(u\) and \(v\) be arbitrary real numbers and let \(\alpha \in [0, 1]\).

We seek to establish
\[
\log \|X\|_{\alpha u + (1-\alpha)v}^\gamma \leq \alpha \log \|X\|_u^\gamma + (1-\alpha) \log \|X\|_v^\gamma,
\]
where \(\overline{\alpha} = 1 - \alpha\). By exponentiating both sides, it is equivalent to show
\[
\|X\|_{\alpha u + (1-\alpha)v} \leq \|X\|_u^\alpha \|X\|_v^{1-\alpha}.
\]
It follows that via Hölder’s inequality that
\[
\|X\|_{\alpha u + (1-\alpha)v} = E(|X|^{\alpha u + (1-\alpha)v}) \\
\leq \left(E\left(|X|^{\alpha u}\right)^{\frac{1}{\alpha}}\right)^{\alpha} \left(E\left(|X|^{(1-\alpha)v}\right)^{\frac{1}{1-\alpha}}\right)^{\overline{\alpha}} \\
\leq (E(|X|^u))^\alpha (E(|X|^v))^{\overline{\alpha}} \\
= \|X\|_u^\alpha \|X\|_v^{1-\alpha}.
\]

\[\square\]
Problem 5

For random variables $X$ and $Y$, define
\[
\rho(X, Y) = \inf \{ \delta > 0 : P(|X - Y| \geq \delta) \leq \delta \}.
\]

In the space $\mathcal{S}$ of random variables, form the space $\mathcal{E}$ of equivalence classes $\{[X] : X \in \mathcal{S}\}$ where $[X] = \{Y : X = Y \text{ a.s.}\}$.

**Proposition 17.** The function $\rho$ defined above is a metric on $\mathcal{E}$. Moreover, $X_n \xrightarrow{p} X$ if and only if $\rho(X_n, X) \to 0$.

**Proof.** We show first that $\rho$ is well-defined on $\mathcal{E}$. To that end, let $X$, $X'$, $Y$, and $Y'$ be random variables with $X = X'$ almost surely and $Y = Y'$ almost surely. It follows that
\[
\rho(X, Y) = \inf \{ \delta > 0 : P(|X - Y| \geq \delta) \leq \delta \} = \inf \{ \delta > 0 : P(|X' - Y'| \geq \delta) \leq \delta \} = \rho(X', Y').
\]

By virtue of this, we will henceforth let a random variable $X$ represent the entire equivalence class $[X]$.

In what follows, let $X$, $Y$, and $Z$ be random variables.

We see immediately that
\[
\rho(X, Y) = \inf \{ \delta > 0 : P(|X - Y| \geq \delta) \leq \delta \} = 0
\]
if and only if $X = Y$.

We have next that
\[
\rho(X, Y) = \inf \{ \delta > 0 : P(|X - Y| \geq \delta) \leq \delta \} = \inf \{ \delta > 0 : P(|Y - X| \geq \delta) \leq \delta \} = \rho(Y, X).
\]

Finally, we have
\[
\rho(X, Z) = \inf \{ \delta > 0 : P(|X - Z| \geq \delta) \leq \delta \} = \inf \{ \delta > 0 : P(|X - Y + Y - Z| \geq \delta) \leq \delta \} \leq \inf \{ \delta > 0 : P(|X - Y| + |Y - Z| \geq \delta) \leq \delta \} \leq \inf \{ \delta > 0 : P(|X - Y| \geq \delta) \leq \delta \} + \inf \{ \delta > 0 : P(|Y - Z| \geq \delta) \leq \delta \} = \rho(X, Y) + \rho(Y, Z).
\]

Taken together, we have shown that $\rho$ is indeed a metric on $\mathcal{E}$.

For the remaining claim, observe that $X_n \xrightarrow{p} X$ if and only if, for all $\delta > 0$,
\[
\lim_{n \to \infty} P(|X_n - X| > \delta) = 0.
\]
The above occurs if and only if, there is \(N\) such that, for all \(n \geq N\),

\[
P(|X_n - X| > \delta) < \delta,
\]

which is equivalent to

\[
\lim_{n \to \infty} \inf \{\delta : P(|X_n - X| > \delta) < \delta\} = 0,
\]

which is to say that \(\rho(X_n, X) \to 0\).

\[\square\]

**Problem 6**

**Proposition 18.** Let \(\{X_n\}\) be a sequence of independent, identically distributed random variables with \(E(X_n) = 0\) and \(\mathbb{V}(X_n) = \sigma^2\) for all \(n \geq 1\). Let \(\{a_n\}\) be a sequence of real numbers and define \(S_n = \sum_{i=1}^{n} a_i X_i\) for all \(n \geq 1\). The sequence \(\{S_n\}\) is \(L_2\)-convergent if and only if \(\sum_{i=1}^{\infty} a_i^2 < \infty\).

**Proof.** Observe first that, for all \(n\),

\[
\mathbb{V}(X_n) = E(X_n^2) + E(X_n)^2
\]

\[= E(X_n^2),\]

so \(E(X_n^2) = \sigma^2\).

We show now, equivalently, that \(\{S_n\}\) is \(L_2\)-cauchy if and only if \(\{a_n^2\}\) is cauchy. To that end, let, without loss of generality, \(n > m\). We have

\[
E((S_n - S_m)^2) = E\left(\left( \sum_{i=m+1}^{n} a_i X_i \right)^2 \right)
\]

\[= E\left( \sum_{i=m+1}^{n} a_i^2 X_i^2 + \sum_{i=m+1}^{n} a_i a_j X_i X_j \right)
\]

\[= \sum_{i=m+1}^{n} a_i^2 E(X_i^2) + \sum_{i=m+1}^{n} a_i a_j E(X_i X_j)
\]

\[= \sum_{i=m+1}^{n} a_i^2 \mathbb{E}(X_i^2) + \sum_{i=m+1}^{n} a_i a_j \mathbb{E}(X_i) \mathbb{E}(X_j) \quad (\text{since } X_i \perp \perp X_j)
\]

\[= \sigma^2 \sum_{i=m+1}^{n} a_i^2.
\]

Therefore, \(\{S_n\}\) is \(L_2\)-cauchy if and only if \(\{a_n^2\}\) is cauchy. \[\square\]
Problem 7

Proposition 19. Let \( \{X_n\} \) be a sequence of random variables such that there exists an increasing function \( f : [0, \infty) \to [0, \infty) \) with \( \frac{f(x)}{x} \to \infty \) as \( x \to \infty \) and \( \sup_n \mathbb{E}(f(|X_n|)) < \infty \). The sequence \( \{X_n\} \) is uniformly integrable.

Proof. (Sketch) Since \( f \) is increasing and \( \frac{f(x)}{x} \to \infty \) as \( x \to \infty \), there exists \( a \in [0, \infty) \) such that \( f(x) > x \) whenever \( x > a \). Let \( A \) be any measurable subset of \( [0, \infty) \). It follows that,

\[
\sup_n \int_A |X_n| \, dP = \sup_n \left( \int_{A \{X_n > a\}} |X_n| \, dP + \int_{A \{X_n \leq a\}} |X_n| \, dP \right) \\
\leq \sup_n \left( \int_A f(|X_n|) \, dP + \int_A a \, dP \right) \\
\leq \sup_n \int_A f(|X_n|) \, dP + aP(A).
\]

The term \( aP(A) \) can be made arbitrarily small, but I do not see how to control the size of \( \sup_n \int_A f(|X_n|) \). Indeed, the situation where \( f = x^2 \) and \( X_n = \sqrt{n} I([0, \frac{1}{n}]) \) appears to be a counterexample to the theorem. I suspect my misunderstanding is that, although \( \mathbb{E}(f(|X_n|)) = 1 \) for all \( n \), \( \sup_n \mathbb{E}(f(|X_n|)) \) is not defined, as in the limit we have a single point at infinity and zero elsewhere. \( \square \)

Problem 8

Proposition 20. Suppose \( X_n \geq 0 \) for \( n \geq 0 \). Suppose further that \( X_n \overset{p.r.}{\to} X_0 \) and \( \mathbb{E}(X_n) \to \mathbb{E}(X_0) \). The sequence \( \{X_n : n \geq 1\} \) is \( L_1 \)-convergent to \( X_0 \).

Proof. (Sketch) We have

\[
\mathbb{E} (|X_n - X_0|) = \mathbb{E} (|(X_n - X_0)^+ - (X_n - X_0)^-|) \\
= \mathbb{E} (X_n - X_0)^+ + \mathbb{E} (X_n - X_0)^-.
\]

Now,

\[
\mathbb{E} (|X_n - X_0|) \leq \mathbb{E} (X_n - X_0) + 2 \mathbb{E} (X_n - X_0)^- \\
\to 2 \mathbb{E} (X_n - X_0)^-, 
\]

since \( \mathbb{E}(X_n) \to \mathbb{E}(X_0) \).

It remains to show that \( \mathbb{E}((X_n - X_0)^-) \to 0 \). To that end, we could use convergence in probability to obtain a subsequence \( \{X_{n_k}\} \) converging almost surely to \( X_0 \), but I am unable to make good use of this fact. \( \square \)
Problem 1

Proposition 21. Let \( \{X_n, n \geq 1\} \) be IID with
\[
\Pr\{X_n = \pm 1\} = 1/2, n = 1, 2, \ldots.
\]
The sum \( \sum_n \frac{1}{n} X_n \) converges almost surely.

Proof. Let \( Y_n = \frac{1}{n} X_n \) for each \( n \). We apply the Kolmogorov Three Series Theorem with \( c = 1 \) to the sequence \( \{Y_n\} \).

First, \( P[|Y_n| > 1] = 0 \) for all \( n \), and so \( \sum_n P[|Y_n| > 1] = 0 \).

Next, observe that \( Y_n 1_{|Y_n| \leq 1} = Y_n \) for each \( n \). We have \( \mathbb{E}(Y_n) = 0 \) and
\[
\mathbb{E}(Y_n^2) = \frac{1}{n^2},
\]

Thus,
\[
\mathbb{V}(Y_n 1_{|Y_n| \leq 1}) = \mathbb{V}(Y_n)
= \frac{1}{n^2},
\]
and so
\[
\sum_{n=1}^{\infty} \mathbb{V}(Y_n 1_{|Y_n| \leq 1}) = \sum_{n=1}^{\infty} \frac{1}{n^2}
< \infty.
\]
We have also
\[
\sum_{n=1}^{\infty} \mathbb{E}(Y_n 1_{|Y_n| \leq 1}) = \sum_{n=1}^{\infty} \mathbb{E}(Y_n)
= \sum_{n=1}^{\infty} 0
= 0.
\]
Therefore, \( \sum_n Y_n = \sum_n \frac{1}{n} X_n \) converges almost surely. \( \square \)

Problem 2

Randomly distribute \( r \) balls in \( n \) boxes so that the sample space \( \Omega \) consists of \( n^r \) equally likely elements. Define
\[
N_n = \sum_{i=1}^{n} I\{\text{ith box is empty}\}
\]
which counts the number of empty boxes.
Proposition 22. As $r/n \to c$, we have

\[ \mathbb{E}(N_n)/n \to \exp(-c); \]

\[ N_n/n \xrightarrow{p} \exp(-c). \]

Proof. Let $I_i$ be the event that the $i^{th}$ box is empty. Observe first that $P(I_i) = \left( \frac{n-1}{n} \right)^r = \left( 1 - \frac{1}{n} \right)^r$, and so $\mathbb{E}(N_n) = n \left( 1 - \frac{1}{n} \right)^r$. If we assume $\frac{r}{n} \to c$ constant, then

\[
\frac{1}{n} \mathbb{E}(N_n) = \frac{1}{n} n \left( 1 - \frac{1}{n} \right)^r = \left( \left( 1 - \frac{1}{n} \right)^n \right)^\frac{r}{n} \to \left( 1 - \frac{1}{n} \right)^c = e^{-c}.
\]

We show next that $\mathbb{V}(\frac{N_n}{n}) \to 0$, and so conclude that $\frac{N_n}{n} \to \mathbb{E}(\frac{N_n}{n}) = e^{-c}$. It follows that

\[
\mathbb{V} \left( \frac{N_n}{n} \right) = \frac{1}{n^2} \mathbb{V}(N_n)
\]

\[
= \frac{1}{n^2} \sum_{i=1}^{n} \left( 1 - \frac{1}{n} \right)^r \left( 1 - \left( 1 - \frac{1}{n} \right)^r \right)
\]

\[
\leq \frac{1}{n^2} \sum_{i=1}^{n} \left( 1 - \frac{1}{n} \right)^r
\]

\[
= \frac{1}{n} \left( 1 - \frac{1}{n} \right)^r
\]

\[
\to 0
\]

as $n \to \infty$. \qed

Problem 3

(Cesaro Sums-Type Results)

Proposition 23. Let $\{a_n, n = 1, 2, \ldots\}$ be a sequence of reals. If $a_n \to 0$ as $n \to \infty$, then $\frac{1}{n} \sum_{j=1}^{n} a_j \to 0$ as $n \to \infty$. 

Proof. Given $\epsilon > 0$, choose $N$ such that $a_n < \frac{\epsilon}{2}$ for all $n \geq N$. It follows that

$$\frac{1}{n} \sum_{j=1}^{n} a_j = \frac{1}{n} \sum_{j=1}^{N} a_j + \frac{1}{n} \sum_{j=N+1}^{n} a_j$$

$$< \frac{1}{n} \sum_{j=1}^{N} a_j + \frac{1}{n} \sum_{j=N+1}^{n} \frac{\epsilon}{2}$$

$$= \frac{1}{n} \sum_{j=1}^{N} a_j + \frac{n - N \epsilon}{2}$$

$$\leq \frac{1}{n} \sum_{j=1}^{N} a_j + \frac{\epsilon}{2}.$$

Now, $\sum_{j=1}^{N} a_j < \infty$, so for sufficiently large $n$, $\frac{1}{n} \sum_{j=1}^{N} a_j < \frac{\epsilon}{2}$. Therefore, $\frac{1}{n} \sum_{j=1}^{n} a_j \to 0$ as $n \to \infty$. \hfill \qed

**Proposition 24.** Let $f : [0, \infty) \to \mathbb{R}$. If $f(x) \to 0$ as $x \to \infty$, then

$$\frac{1}{t} \int_{0}^{t} f(x)dx \to 0 \quad \text{as} \quad t \to \infty.$$

**Proof.** For each $n$, define $a_n = \max\{|f(x)| : x \in [n, n+1]\}$. Since $f(x) \to 0$, we have $a_n \to 0$. Now,

$$\frac{1}{t} \int_{0}^{t} f(x)dx = \frac{1}{t} \sum_{j=0}^{t-1} \int_{j}^{j+1} f(x)dx$$

$$\leq \frac{1}{t-1} \sum_{j=0}^{t-1} \int_{j}^{j+1} a_j dx$$

$$= \frac{1}{t-1} \sum_{j=0}^{t-1} a_j$$

which converges almost surely to 0 with $t$ by the previous proposition. \hfill \qed

**Problem 4**

**Proposition 25.** Let $\{X_n, n \geq 1\}$ be independent rvs such that

$$\forall n : \mathbb{E}(X_n) = 0 \quad \text{and} \quad \mathbb{V}(X_n^2) < \infty.$$

With $S_n = \sum_{i=1}^{n} X_i$, we have, for every $\epsilon > 0$,

$$P \left\{ \max_{i \leq n} |S_i| > \epsilon \right\} \leq \frac{\mathbb{V}(S_n)}{\epsilon^2}.$$
Proof. Define for each $n$ the event $A_n$ that $|S_n| \geq \epsilon$ but $|S_i| < \epsilon$ for all $i < n$ (so $\omega \in A_n$ means that $n$ is the least index for which $|S_n(\omega)| \geq \epsilon$). The $A_n$ are disjoint with

$$\bigcup_{j=1}^{n} A_j = \left\{ \max_{j \leq n} |S_j| \geq \epsilon \right\}.$$ 

It follows that

$$\mathbb{V}(S_n) = \mathbb{E}(S_n^2)$$

$$\geq \sum_{j=1}^{n} \mathbb{E}(S_n^2 1_{A_j})$$

$$= \sum_{j=1}^{n} \mathbb{E}((S_j^2 + 2S_j(S_n - S_j) + (S_n - S_j)^2) 1_{A_j})$$

$$\geq \sum_{j=1}^{n} \mathbb{E}((S_j^2 + 2S_j(S_n - S_j)) 1_{A_j})$$

$$= \sum_{j=1}^{n} \mathbb{E}(S_j^2 1_{A_j}) + 2 \sum_{j=1}^{n} \mathbb{E}(S_n - S_j) \mathbb{E}(S_j 1_{A_j}) \quad \text{(by independence of the } S_j)$$

$$= \sum_{j=1}^{n} \mathbb{E}(S_j^2 1_{A_j})$$

$$\geq \epsilon^2 \sum_{j=1}^{n} \mathbb{E}(1_{A_j})$$

$$= \epsilon^2 \sum_{j=1}^{n} P(A_j)$$

$$= \epsilon^2 P\left( \bigcup_{j=1}^{n} A_j \right)$$

$$= \epsilon^2 P\left( \left\{ \max_{j \leq n} |S_j| \geq \epsilon \right\} \right).$$

Dividing by $\epsilon^2$ give the desired result. \qed

Problem 5

Let $\{Z_n, n = 1, 2, \ldots\}$ be independent standard normal random variables. With $S_n = \sum_{j=1}^{n} Z_j$, we can obtain via Kolmogorov's Inequality
\[
P \left\{ \max_{j \leq 10} |S_j| > 10 \right\} \leq \frac{\nu(S_{10})}{10^2} = \frac{10}{10^2} = \frac{1}{10}.
\]

The following R program (which performs 100,000 experiments) estimates the probability as 0.00211, which is significantly lower than the bound provided by Kolmogorov.

```r
loop <- 0
numLoop <- 100000
max <- 10
count <- 0
while (loop < numLoop) {
  v = rnorm(10,0,1)
  pv <- vector()
i <- 1
  while ( i <= 10 ) {
    col <- 1
    partial <- 0
    while ( col <= i ) {
      partial <- partial + v[col]
      col <- col + 1
    }
    pv[i] <- abs(partial)
i <- i + 1
  }
  if (max(pv) > max) {count <- count + 1}
  loop <- loop + 1
}
prob <- count / numLoop
print(prob)
```

**Problem 6**

**Proposition 26.** For a collection \( \{A_n, n \geq 1\} \) of events in some ambient probability space,

\[
P\left( \bigcup_{i=1}^{n} A_i \right) \geq \frac{\mathbb{E}(\sum_{i=1}^{n} f(A_i))^2}{\mathbb{E}\left( [\sum_{i=1}^{n} f(A_i)]^2 \right)}.
\]
Proof. Observe first that
\[
I \left( \bigcup_{j=1}^{n} A_j \right) \sum_{j=1}^{n} I(A_j) = \sum_{j=1}^{n} I(A_j)
\]
since the support of each \( I(A_j) \) is contained in \( \bigcup_{j=1}^{n} A_j \). Thus, we have
\[
E \left( I \left( \bigcup_{j=1}^{n} A_j \right) \sum_{j=1}^{n} I(A_j) \right) = E \left( \sum_{j=1}^{n} I(A_j) \right).
\]
Applying Cauchy-Schwarz yields
\[
\sqrt{E \left( I \left( \bigcup_{j=1}^{n} A_j \right)^2 \right) E \left( \sum_{j=1}^{n} I(A_j)^2 \right)} \geq E \left( \sum_{j=1}^{n} I(A_j) \right),
\]
which is equivalent to
\[
\sqrt{P \left( \bigcup_{j=1}^{n} A_j \right) E \left( \sum_{j=1}^{n} I(A_j)^2 \right)} \geq E \left( \sum_{j=1}^{n} I(A_j) \right).
\]
Rearranging terms gives
\[
P \left( \bigcup_{j=1}^{n} A_j \right) \geq \frac{E \left( \sum_{j=1}^{n} I(A_j)^2 \right)}{E \left( \sum_{j=1}^{n} I(A_j) \right)^2},
\]
as desired.

Problem 2

Proposition 27. If \( \{X_n, n \geq 1\} \) are IID with \( P(X_n = 0) = P(X_n = 2) = 1/2 \), then \( \sum_{n=1}^{\infty} X_n/3^n \) converges almost surely.

Proof. We invoke the Kolmogorov Convergence Criterion. In what follows, let \( Y_n = \frac{X_n}{3^n} \).

First,
\[
\mathbb{V}(Y_n) = \mathbb{E}(Y_n^2) - \mathbb{E}(Y_n)^2
\]
\[
= \frac{1}{2} \left( \frac{2}{3^n} \right)^2 - \left( \frac{1}{2} \cdot \frac{2}{3^n} \right)^2
\]
\[
= \frac{1}{3^{2n}}.
\]
for all \( n \). Thus,
\[
\sum_n V(Y_n) = \sum_n \frac{1}{3^{2n}} = \frac{1}{8} < \infty.
\]

By the convergence criterion, \( \sum_n (Y_n - \mathbb{E}(Y_n)) \) converges almost surely. Now,
\[
\mathbb{E}(Y_n) = \frac{1}{2} \cdot \frac{2}{3^n} = \frac{1}{3^n}.
\]
Since \( \sum_n \frac{1}{3^n} = \frac{1}{2} < \infty \), it follows that \( \sum_n Y_n \) converges almost surely. \( \Box \)

**Problem 3**

Let \( \{X_n, n \geq 1\} \) be IID random variables taking values in the set \( S = \{1, 2, \ldots, 17\} \). Define the discrete density (density with respect to counting measure) \( f_0(y) = P(X_1 = y), y \in S \) so that \( f_0(y) \geq 0 \) and \( \sum_{y \in S} f_0(y) = 1 \). Let \( f_1 \neq f_0 \) be another discrete density on \( S \) so that we also have \( f_1(y) \geq 0 \) and \( \sum_{y \in S} f_1(y) = 1 \). Set
\[
Z_n = \prod_{i=1}^n \frac{f_1(X_i)}{f_0(X_i)}, n \geq 1.
\]
The \( \{Z_n, n \geq 1\} \) is called the likelihood ratio process.

**Proposition 28.** Under \( f_0 \), \( Z_n \overset{a.s.}{\to} 0 \).

**Proof.** Define \( Y_n = \log(Z_n) \), so that
\[
Y_n = \sum_{i=1}^n \log \left( \frac{f_1(X_i) f_0(X_i)}{f_0(X_i) f_0(X_i)} \right).
\]
To show that \( Z_n \overset{a.s.}{\to} 0 \), we show equivalently that \( Y_n \overset{a.s.}{\to} -\infty \).
Applying Jensen’s Inequality, we have

\[
\mathbb{E}_{f_0} \left[ \log \left( \frac{f_1(X_i)}{f_0(X_i)} \right) \right] < \log \left( \mathbb{E}_{f_0} \left[ \frac{f_1(X_i)}{f_0(X_i)} \right] \right) \\
= \log \left( \sum_{y \in S} \frac{f_1(y)}{f_0(y)} f_0(y) \right) \\
= \log \left( \sum_{y \in S} f_1(y) \right) \\
= \log(1) \\
= 0.
\]

Now,

\[
\mathbb{E}_{f_0}(Y_n) = \sum_{i=1}^{n} \mathbb{E}_{f_0} \left[ \log \left( \frac{f_1(X_i)}{f_0(X_i)} \right) \right] \\
< 0.
\]

If the sum above diverges to \(-\infty\), then we are done. Suppose instead \(\mathbb{E}_{f_0}(Y_n)\) is equal to some finite, negative value \(\mu\) (and so \(\mathbb{E}_{f_0}|Y_n| = -\mu\) is also finite). By the Strong Law of Large Numbers,

\[
\frac{Y_n}{n} \xrightarrow{a.s.} \mu.
\]

Thus, \(Y_n\) decreases without bound, and so \(Y_n \xrightarrow{a.s.} -\infty\).

**Problem 4**

**Proposition 29.** Let \(\{X_n, n \geq 1\}\) be IID random variables with \(P(X_n > x) = \exp(-x), x \geq 0\). As \(n \to \infty\),

\[
\frac{\max_{i \leq n} X_i}{\log(n)} \xrightarrow{a.s.} 1.
\]

**Proof.** Let \(Y_n = \max_{i \leq n} X_i\). We show that \(\frac{Y_n}{\log(n)} \xrightarrow{a.s.} 1\).
Let $\epsilon > 0$ be given. We have

$$P\left(\frac{Y_n}{\log(n)} < 1 - \epsilon\right) = P(Y_n < (1 - \epsilon) \log(n))$$

$$= P\left(\bigcap_{i=1}^{n} [X_i < (1 - \epsilon) \log(n)]\right)$$

$$= \left(1 - e^{-(1-\epsilon) \log(n)}\right)^n$$

$$= \left(1 - \frac{1}{n^{1-\epsilon}}\right)^n$$

$$= \left(1 - \frac{1}{n^{1-\epsilon}}\right)^{n^{1-\epsilon}}.$$

Thus,

$$\sum_{n=1}^{\infty} P\left(\frac{Y_n}{\log(n)} < 1 - \epsilon\right) < \sum_{n=1}^{\infty} \left(1 - \frac{1}{n^{1-\epsilon}}\right)^{n^{1-\epsilon}}$$

$$\sim \sum_{n=1}^{\infty} (e^{-1})^{n^{1-\epsilon}}$$

$$= \sum_{n=1}^{\infty} \frac{1}{e^{n^{1-\epsilon}}}$$

$$< \infty.$$

Hence, by Borel-Cantelli I,

$$P\left(\frac{Y_n}{\log(n)} < 1 - \epsilon \text{ i.o.}\right) = 0.$$

By similar arguments, we have

$$P\left(\frac{X_n}{\log(n)} > 1 + \epsilon\right) = e^{-\log(n)(1+\epsilon)}$$

$$= \frac{1}{n^{1+\epsilon}}.$$

Thus,

$$\sum_{n=1}^{\infty} P\left(\frac{X_n}{\log(n)} > 1 + \epsilon\right) = \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}}$$

$$< \infty.$$

By Borel-Cantelli I,

$$P\left(\frac{X_n}{\log(n)} > 1 + \epsilon \text{ i.o.}\right) = 0.$$
Similarly,
\[
\sum_{n=1}^{\infty} P \left( \frac{X_n}{\log(n)} > 1 - \epsilon \right) = \sum_{n=1}^{\infty} \frac{1}{n^{1-\epsilon}} = \infty.
\]

By Borel-Cantelli II,
\[
P \left( \frac{X_n}{\log(n)} > 1 - \epsilon \text{ i.o.} \right) = 1.
\]

Taken together, we have
\[
\limsup_n \frac{X_n}{\log(n)} = 1 \text{ almost surely,}
\]
and so
\[
P \left( \frac{Y_n}{\log(n)} > 1 + \epsilon \text{ i.o.} \right) = 0,
\]
thus completing the proof.

**Problem 5**

Suppose \( \{X_n, n \geq 1\} \) are IID random variables taking values in the alphabet \( S = \{1, 2, \ldots, r\} \) with positive probabilities \( p_1, p_2, \ldots, p_r \). Define
\[
p_n(i_1, i_2, \ldots, i_n) = P\{X_1 = i_1, X_2 = i_2, \ldots, X_n = i_n\},
\]
and set
\[
\chi_n(\omega) \equiv p_n(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)).
\]

For interpretation, the \( \chi_n(\omega) \) is the probability that in a new sample of \( n \) observations, what is observed coincides with the original set of observations.

**Proposition 30.** With \( \chi_n(\omega) \) defined as above,
\[
-\frac{1}{n} \log \chi_n(\omega) \overset{a.s.}{\longrightarrow} -\sum_{j=1}^{r} p_j \log(p_j).
\]

**Proof.** Due to the independence of the \( X_n \),
\[
\chi_n(\omega) = P(X_1 = X_1(\omega), \ldots, X_n = X_n(\omega))
\]
\[
= \prod_{i=1}^{n} P(X_i = X_i(\omega))
\]
\[
= \prod_{i=1}^{n} p_{X_i(\omega)}.
\]
Now,
\[
\log(\chi_n(\omega)) = \sum_{i=1}^{n} \log(p_{X_i}(\omega)).
\]

Let \( S_n = \sum_{i=1}^{n} \log(p_{X_i}(\omega)) \). Since the \( X_i \) are IID, the \( \log(p_{X_i}(\omega)) \) terms are also IID. Now,
\[
E\left(\log(p_{X_i}(\omega))\right) = \sum_{j=1}^{r} p_j \log(p_j),
\]
which is finite. Thus, by the Strong Law of Large Numbers,
\[
-\frac{1}{n} \log \chi_n(\omega) = \frac{S_n}{n} \xrightarrow{a.s.} -E(S_n) \nrightarrow -\sum_{j=1}^{r} p_j \log(p_j),
\]
as desired. \( \square \)

**Problem 1**

Let \( S_n \) have a binomial distribution with parameters \( n \) and \( \theta \in (0, 1) \). By basic Central Limit Theorem it is known that
\[
\frac{S_n - n\theta}{\sqrt{n\theta(1-\theta)}} \Rightarrow N(0,1).
\]

**Proposition 31.** If \( Y_n = \log[S_n/(n-S_n)] \), then \( Y_n \sim AN\left(\frac{\theta}{1-\theta}, \frac{\theta}{n(1-\theta)\tau}\right) \).

**Proof.** Observe first that, from the given weak convergence of \( S_n \), we have
\[
\frac{\bar{S}_n - \theta}{\sqrt{\theta(1-\theta)n}} \Rightarrow N(0,1),
\]
so that \( \frac{\bar{S}_n}{n} \sim AN\left(\theta, \frac{\theta(1-\theta)}{n}\right) \).

Next, we write
\[
Y_n = \log\left(\frac{\bar{S}_n}{1 - \frac{\bar{S}_n}{n}}\right).
\]

Define \( g(x) = \log\left(\frac{x}{1-x}\right) \), so that \( Y_n = g\left(\frac{\bar{S}_n}{n}\right) \). Now, \( g'(x) = \frac{1}{x(1-x)} \), which is nonzero for all \( x \in (0,1) \). By the Delta Method, we have \( Y_n \sim AN\left(\frac{\theta}{1-\theta}, \frac{1}{n\theta(1-\theta)\tau}\right) \). \( \square \)
Problem 2

**Proposition 32.** Let $X_n, n = 1, 2, \ldots$ be independent random variables such that $P\{X_n = n\} = 1/n = 1 - P\{X_n = 0\}$. The sequence \{X_n, n = 1, 2, \ldots\} converges in probability, and thus weakly.

**Proof.** We show that $X_n \xrightarrow{pr} 0$. For any $\epsilon > 0$, we have

$$\lim_{n \to \infty} P(|X_n - 0| > \epsilon) = \lim_{n \to \infty} P(X_n = n)$$

$$= \lim_{n \to \infty} \frac{1}{n} = 0.$$

Now, convergence in probability always implies weak convergence, and so the claim is proven. \hfill \square

Problem 3

**Proposition 33.** Let \{U_n, n = 1, 2, \ldots\} be IID from a uniform distribution on [0, 1]. There exist sequences \{\mu_n\} and \{\sigma_n\} such that

$$\frac{\prod_{i=1}^{n} U_i^{1/n} - \mu_n}{\sigma_n}$$

converges weakly to a non-degenerate random variable.

**Proof.** Let $X_n = \log(U_i)$, so that $\log\left(\prod_{i=1}^{n} U_i^{\frac{1}{n}}\right) = \frac{1}{n} \sum_{i=1}^{n} \log(U_i) = \overline{X}_n$. Let $\mu_X = E(X_1)$ and $\sigma_X^2 = V(X_1)$ be the common mean and variance, respectively. By the basic Central Limit Theorem, we have

$$\sqrt{n} \left(\frac{\overline{X}_n - \mu_X}{\sigma_X}\right) \Rightarrow N(0, 1),$$

and so $\overline{X}_n \sim AN\left(\mu_X, \frac{\sigma_X^2}{n}\right)$. Applying the Delta Method with $g(x) = e^x$, we have

$$e^{\overline{X}_n} = \prod_{i=1}^{n} U_i^{\frac{1}{n}} \sim AN\left(e^{\mu_X}, e^{2\mu_X} \frac{\sigma_X^2}{n}\right).$$

The desired sequences are thus obtained by taking $\mu_n = e^{\mu_X}$ and $\sigma_n = e^{\mu_X} \frac{\sigma_X}{\sqrt{n}}$. \hfill \square

Problem 4

Let \{X_n, n = 1, 2, \ldots\} be IID from a unit exponential distribution. For a fixed $l$, denote by $X_{(l:n)}$ the $l$th smallest among $X_1, X_2, \ldots, X_n$ where $n \geq l$. 
Proposition 34. The sequence $nX_{(l:n)} \Rightarrow Y_l$ where $Y_l$ has a gamma distribution with shape parameter $l$ and scale parameter 1, that is,

$$P\{Y_l \leq y\} = \int_0^y \frac{1}{(l-1)!} w^{l-1} \exp(-w) \, dw.$$ 

Proof. Define, for $1 \leq i \leq n$, the spacings $D_i = (n-i+1)(X_{(i)} - X_{(i-1)})$ (where we take $D_1 = nX_{(1)}$). By Renyi’s representation, we have that the $D_i$ are IID unit exponential random variables. Observe that $X_{(l:n)} = \sum_{i=1}^l D_i$. Thus, $nX_{(l:n)} \rightarrow \sum_{i=1}^l D_i$. It is known that the sum of $l$ IID unit exponential random variables has a gamma distribution with shape parameter $l$ and scale parameter 1, thus completing the proof.

Problem 5

For distribution functions $F$ and $G$ define

$$d(F,G) = \inf\{\delta > 0 : \forall x \in \mathbb{R}, F(x - \delta) - \delta \leq G(x) \leq F(x + \delta) + \delta\}.$$ 

Proposition 35. On the space of distribution functions, $d(\cdot, \cdot)$ is a metric.

Proof. Let $F$, $G$, and $H$ be distribution functions. We claim first that, whenever $d(F,G) < \alpha$, 

$$F(x - \alpha) - \alpha \leq G(x) \leq F(x + \alpha) + \alpha.$$ 

To see this, observe that, by definition of $d(\cdot, \cdot)$, there is $0 \leq \delta < \alpha$ such that 

$$F(x - \delta) - \delta \leq G(x) \leq F(x + \delta) + \delta.$$ 

As $F$ is an increasing function, the claim follows.

We show now that the function $d$ is symmetric. For any fixed $\alpha > d(F,G)$, we have 

$$F(x - \alpha) - \alpha \leq G(x) \leq F(x + \alpha) + \alpha,$$

from which is follows that 

$$G(x - \alpha) \leq F(x) + \alpha$$

and 

$$F(x) - \alpha \leq G(x + \alpha).$$

Equivalently, 

$$G(x - \alpha) \leq F(x) \leq G(x + \alpha) + \alpha.$$ 

Thus, $d(G,F) \leq \alpha$. Since alpha was arbitrary, we obtain that $d(G,F) \leq d(F,G)$. By the same argument, we also see that $d(F,G) \leq d(G,F)$, and so $d(F,G) = d(G,F)$.

We show next that $d(F,G) = 0$ if and only if $F = G$. If $F = G$, then 

$$F(x - 0) - 0 \leq G(x) \leq F(x + 0) + 0,$$
and so \( d(F, G) = 0 \). If \( d(F, G) = 0 \), then we have
\[
F(x - \alpha) - \alpha \leq G(x) \leq F(x + \alpha) + \alpha,
\]
for any \( \alpha > 0 \). Since \( F \) is right continuous, we have \( G \leq F \). Now, by symmetry, we have also \( d(G, F) = 0 \), and so the same argument gives \( F \leq G \). Thus, \( F = G \).

Finally, we show that \( d \) satisfies the triangle inequality. Choose \( \alpha > d(F, H) \) and \( \beta > d(H, G) \) and apply the initial claim to get
\[
F(x - \alpha) - \alpha \leq H(x) \leq F(x + \alpha) + \alpha
\]
and
\[
H(x - \beta) - \beta \leq G(x) \leq H(x + \beta) + \beta.
\]
It follows that
\[
G(x) \leq H(x + \beta) + \beta \leq F(x + \alpha + \beta) + \alpha + \beta
\]
and
\[
G(x) \geq H(x - \beta) - \beta \geq F(x - \alpha - \beta) - \alpha - \beta.
\]
Thus, \( d(F, G) \leq \alpha + \beta \). As \( \alpha \) and \( \beta \) are arbitrary, we have
\[
d(F, G) \leq d(F, H) + d(H, G),
\]
as desired. \( \square \)

**Proposition 36.** The metric \( d(\cdot, \cdot) \) metrizes convergence in distribution. That is, for distribution functions \( \{F_n \mid n \in \mathbb{N}\} \) and \( F \),
\[
F_n \Rightarrow F \quad \text{if and only if} \quad d(F_n, F) \to 0.
\]

**Proof.** (\( \Rightarrow \)) Let \( \epsilon > 0 \) be given. Since \( F \) is a distribution function, there are continuity points \( a \) and \( b \) such that
\[
0 \leq F(x) \leq \frac{\epsilon}{2} \quad \text{for all} \quad x \in (-\infty, a]
\]
and
\[
1 - \frac{\epsilon}{2} \leq F(x) \leq 1 \quad \text{for all} \quad x \in [b, \infty).
\]
Choose now a sequence \( \{x_n\} \) of continuity points of \( F \) such that
\[
a = x_0 < x_1 < \cdots < x_n = b
\]
and
\[
x_i - x_{i-1} < \epsilon \quad \text{for all} \quad i.
\]
Now,
\[
\lim_{n \to \infty} F_n(x_i) = F(x_i) \quad \text{for all} \quad i,
\]
so there exists $N_\epsilon$ such that

$$|F_n(x_i) - F(x_i)| \leq \frac{\epsilon}{2} \text{ for all } n \geq N_\epsilon.$$  

Let now $x \in [x_{i-1}, x_i] \subset [a, b]$. For $n \geq N_\epsilon$, we have

$$F_n(x) \leq F_n(x_i) \leq F(x_i) + \frac{\epsilon}{2} \leq F(x) + \epsilon.$$  

Similarly,

$$F_n(x) \geq F_n(x_{i-1}) \geq F(x_{i-1}) - \frac{\epsilon}{2} \geq F(x) - \epsilon.$$  

Thus, for all $x \in [a, b]$ and all $n \geq N_\epsilon$,

$$F(x) - \epsilon \leq F_n(x) \leq F(x) + \epsilon.$$  

For $x \in (-\infty, a)$, choose any continuity point $x^+$ such that $x < x^+$. We have for all $n \geq N_\epsilon$,

$$F_n(x) \leq F_n(x^+) \leq F(x^+) + \frac{\epsilon}{2} \leq \epsilon,$$

and so $|F_n(x) - F(x)| \leq \epsilon$.

For $x \in (b, \infty)$, choose any continuity point $x^-$ such that $x^- < x$. We have for all $n \geq N_\epsilon$,

$$F_n(x) \geq F_n(x^-) \geq F(x^-) - \frac{\epsilon}{2} \geq 1 - \epsilon,$$

and so $|F_n(x) - F(x)| \leq \epsilon$.

As $\epsilon$ is arbitrary, we have $d(F_n, F) \to 0$, as desired.

$(\Leftarrow)$ Let $\epsilon > 0$ be given. There is $N_\epsilon$ such that $d(F_n, F) < \epsilon$ for all $n \geq N_\epsilon$.

By earlier remarks, this implies

$$F(x - \epsilon) - \epsilon \leq F_n(x) \leq F(x) + \epsilon$$

for all $x \in \mathbb{R}$ and $n \geq N_\epsilon$. Hence,

$$F(x - \epsilon) - \epsilon \leq \liminf_{n \to \infty} F_n(x) \leq \limsup_{n \to \infty} F_n(x) \leq F(x) + \epsilon.$$  

Now, for any continuity point $x$, we have

$$\lim_{\epsilon \downarrow 0} F(x - \epsilon) - \epsilon = \lim_{\epsilon \downarrow 0} F(x) + \epsilon,$$

and so conclude that $\lim_{n \to \infty} F_n(x) = F(x)$ for all continuity points $x$. That is, $F_n \Rightarrow F$. \qed
Problem 6

Proposition 37. Let \( \{X_n, n = 1, 2, \ldots\} \) be IID from a standard uniform distribution. There exist sequences \( \{a_n\} \) and \( \{b_n\} \) such that \( (M_n - a_n)/b_n \) converges weakly to a nondegenerate distribution, where \( M_n = \bigvee_{i=1}^{n} X_i = \max\{X_1, X_2, \ldots, X_n\} \).

Proof. Observe first that

\[
P(M_n \leq x) = P\left(\left(\bigvee_{i=1}^{n} X_i \right) \leq x\right) = P((X_1 \leq x) \text{ and } \cdots \text{ and } (X_n \leq x)) = \prod_{i=1}^{n} P(X_i \leq x) = x^n.
\]

We next compute the mean and variance of \( M_n \) via the probability density function \( nx^{n-1} \):

\[
\mathbb{E}(M_n) = \int_{0}^{1} x \cdot x^{n-1} \, dx = \frac{n}{n+1}
\]

and

\[
\mathbb{V}(M_n) = \int_{0}^{1} x^2 \cdot x^{n-1} \, dx = \frac{n}{(n+1)^2(n+2)}.
\]

Considering only the order of the mean and variance, we claim that setting \( a_n = 1 \) and \( b_n = \frac{1}{n} \) achieves the desired result. To verify this, observe

\[
P\left(\frac{M_n - 1}{n} \leq x\right) = P\left(M_n \leq 1 + \frac{x}{n}\right) = \left(1 + \frac{x}{n}\right)^n \rightarrow e^x.
\]

\( \square \)

Problem 7

Proposition 38. If \( X_n \Rightarrow X_0 \) and for some \( \delta > 0 \), \( \sup_n \mathbb{E} \left[ |X_n|^{2+\delta} \right] < \infty \), then \( \mathbb{E}(X_n) \rightarrow \mathbb{E}(X_0) \) and \( \mathbb{V}(X_n) \rightarrow \mathbb{V}(X_0) \).
Proof. Observe first that, by the Baby Skorohod Theorem, there are random variables $X^*_n$ such that, for each $n$, $X^*_n \xrightarrow{d} X_n$ and $X^*_n \xrightarrow{a.s.} X_n$. Hence, it is equivalent to establish the desired results for $\{X^*_n\}$.

Since $\mathbb{E}(\|X^*_n\|^{2+\delta}) < \infty$, the Crystal Ball Condition guarantees the families $\{X^*_n\}$ and $\{(X^*_n)^2\}$ are uniformly integrable. Now, uniform integrability combined with almost sure convergence implies convergence of expectations. That is, $\mathbb{E}(X^*_n) \rightarrow \mathbb{E}(X_0^*)$ and

$$\forall (X^*_n) = \mathbb{E}\left((X^*_n)^2\right) + \mathbb{E}\left(X^*_n\right)^2$$

$$\rightarrow \mathbb{E}\left((X_0^*)^2\right) + \mathbb{E}\left(X_0^*\right)^2 \quad \text{(since both } \{X^*_n\} \text{ and } \{(X^*_n)^2\} \text{ are u.i.)}$$

$$= \mathbb{V}(X_0^*),$$

as desired. \hfill \Box

Problem 1

Let $X_i, i = 1, 2, \ldots, n$ be independent random variables with $X_i$ having a normal distribution with mean $\mu_i$ and variance $\sigma^2_i$.

Proposition 39. The sum $S = \sum_{i=1}^{n} X_i$ is normally distributed with mean $\sum_{i=1}^{n} \mu_i$ and variance $\sum_{i=1}^{n} \sigma^2_i$.

Proof. (via convolutions) We demonstrate the result for the case where $n = 2$ and appeal to induction for the general result.

Let $f(x) = c_1 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ and $g(x) = c_2 \exp\left(-\frac{(x-\nu)^2}{2\tau^2}\right)$ be the probability density functions of independent, normally distributed random variables $X$ and $Y$, respectively. We have

$$(f * g)(x) = \int_{\mathbb{R}} f(u)g(x-u)du$$

$$= c_1 \int_{\mathbb{R}} \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \exp\left(-\frac{(x-u-\nu)^2}{2\tau^2}\right) du \quad \text{(where } c = c_1c_2)$$

$$= c_1 \int_{\mathbb{R}} \exp\left(-\frac{\tau^2(u-\mu)^2 + \sigma^2(x-u-\nu)^2}{2\sigma^2\tau^2}\right).$$

Using a computer algebra system, one can rewrite the above as

$$c \exp\left(-\frac{(x-(\mu+\nu))^2}{2(\sigma^2 + \tau^2)}\right) \int_{\mathbb{R}} h(u)du,$$

where $h(u)$ does not depend on $x$. Hence, we obtain

$$(f * g)(x) = c' \exp\left(-\frac{(x-(\mu+\nu))^2}{2(\sigma^2 + \tau^2)}\right).$$
Therefore, $X + Y$ is normally distributed with mean $\mu + \nu$ and variance $\sigma^2 + \tau^2$.

**Proof.** (via characteristic functions) We demonstrate the result for the case where $n = 2$ and appeal to induction for the general result.

Let $\phi_X(t)$ and $\phi_Y(t)$ be the characteristic functions of independent, normally distributed random variables $X$ and $Y$, respectively. We have

$$\phi_{X+Y}(t) = \mathbb{E}(\exp(it(X + Y))) = \mathbb{E}(\exp(itX)) + \mathbb{E}(\exp(itY)) = \exp \left( it\mu - \frac{\sigma^2 t^2}{2} \right) + \exp \left( it\nu - \frac{\tau^2 t^2}{2} \right) = \exp(it(\mu + \nu) - \frac{(\sigma^2 + \tau^2)t^2}{2}).$$

Therefore, $X + Y$ is normally distributed with mean $\mu + \nu$ and variance $\sigma^2 + \tau^2$. \qed

**Problem 2**

A rv $X$ has a Poisson distribution with rate $\lambda$ if $P\{X = k\} = \exp(-\lambda)\lambda^k/k!$, $k = 0, 1, 2, \ldots$.

**Proposition 40.** The characteristic function of $X$ is $\exp(\lambda(\exp(iu) - 1))$.

**Proof.** Let $\phi(t)$ be the characteristic function of $X$. We have

$$\phi(t) = \mathbb{E}(\exp(itX)) = \sum_{k=0}^{\infty} \frac{\exp(-\lambda)\lambda^k}{k!} \exp(iuk) = \exp(-\lambda) \sum_{k=0}^{\infty} \frac{(\lambda \exp(iu))^k}{k!} = \exp(\lambda(\exp(iu) - 1)).$$

\qed

**Proposition 41.** Let $S = \sum_{i=1}^{n} X_i$ where $X_i \sim \text{POI}(\lambda_i)$. The random variable $S$ has a Poisson distribution with parameter $\sum_{i=1}^{n} \lambda_i$.

**Proof.** We demonstrate the result for the case where $n = 2$ and appeal to induction for the general result.
Let $X \sim POI(\lambda_1)$ and $Y \sim POI(\lambda_2)$ and let $f_X$ and $f_Y$ be their respective probability density functions. We have

$$P(X + Y = k) = \sum_{j=0}^{k} f_X(j)f_Y(k-j)$$

$$= \sum_{j=0}^{k} \frac{\lambda_1^j}{j!} e^{-\lambda_1} \frac{\lambda_2^{k-j}}{(k-j)!} e^{-\lambda_2}$$

$$= e^{-(\lambda_1+\lambda_2)} \sum_{j=0}^{k} \frac{\lambda_1^j}{j!} \frac{\lambda_2^{k-j}}{(k-j)!}$$

$$= e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!},$$

where the last equality results from an application of the binomial theorem.

### Problem 3

A rv $X$ is said to have a gamma distribution with shape parameter $\alpha (\alpha > 0)$ and scale parameter $\lambda (\lambda > 0)$ if its density function (wrt Lebesgue measure) is

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\lambda x\}, x \geq 0.$$  

We shall write in this case $X \sim \Gamma(\alpha, \lambda)$.

**Proposition 42.** The characteristic function of $X$ is

$$\left(1 - \frac{it}{\lambda}\right)^{-\alpha}.$$

**Proof.** We have

$$\mathbb{E}(e^{itX}) = \int_{0}^{\infty} e^{ity} \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} dy$$

$$= \int_{0}^{\infty} \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-(\lambda+it)y} dy$$

$$= \frac{\lambda^\alpha}{(\lambda - it)^\alpha} \int_{0}^{\infty} (\lambda - it)^{\alpha-1} e^{-(\lambda - it)y} dy$$

$$= \frac{\lambda^\alpha}{(\lambda - it)^\alpha}$$

$$= \left(1 - \frac{it}{\lambda}\right)^{-\alpha}.$$
Proposition 43. Let $X_i \sim \Gamma(\alpha_i, \lambda), i = 1, 2, \ldots, n$ and assume that they are independent. Let also $S = \sum_{i=1}^{n} X_i$. The characteristic function of $S$ is

$$\prod_{i=1}^{n} \left(1 - \frac{it}{\lambda}\right)^{-\alpha_i}.$$ 

Proof. We have

$$\mathbb{E} \left( e^{itS} \right) = \prod_{i=1}^{n} e^{itX_i} = \prod_{i=1}^{n} \left(1 - \frac{it}{\lambda}\right)^{-\alpha_i}.$$ 

Remark 1. To determine the distribution of $S$, we first obtain the density $f$ via

$$f(x) = \frac{1}{2\pi} \int_{0}^{\infty} e^{-iyx} \phi_S(y) dy$$

$$= \frac{1}{2\pi} \int_{0}^{\infty} e^{-iyx} \prod_{i=1}^{n} \left(1 - \frac{iy}{\lambda}\right)^{-\alpha_i} dy.$$ 

The distribution $F(u)$ is thus given by

$$F(u) = \frac{1}{2\pi} \int_{0}^{u} \int_{0}^{\infty} e^{-iyx} \prod_{i=1}^{n} \left(1 - \frac{iy}{\lambda}\right)^{-\alpha_i} dy dx$$

$$= \frac{1}{2\pi} \int_{0}^{\infty} \int_{0}^{u} e^{-iyx} \prod_{i=1}^{n} \left(1 - \frac{iy}{\lambda}\right)^{-\alpha_i} dx dy$$

$$= \frac{1}{2\pi} \int_{0}^{\infty} \prod_{i=1}^{n} \left(1 - \frac{iy}{\lambda}\right)^{-\alpha_i} e^{-iyu} dx$$

$$= \frac{1}{2\pi} \int_{0}^{\infty} \prod_{i=1}^{n} \left(1 - \frac{iy}{\lambda}\right)^{-\alpha_i} \left(\frac{e^{-iyu} - 1}{-iy}\right) dy.$$ 

I do not see how to further simplify the expression above.

Problem 4

Let $X_n, n = 1, 2, 3, \ldots$ be IID rvs with $P(X_n = -1) = P(X_n = +1) = 1/2, n = 1, 2, \ldots$. The Cantor random variable on $[-1/2, 1/2]$ is defined to be

$$C = \sum_{n=1}^{\infty} \frac{X_n}{3^n}.$$ 

31
Proposition 44. The characteristic function of $C$ is $\prod_{n=1}^{\infty} \cos \left( \frac{t}{3^n} \right)$.

Proof. We have
\[
\mathbb{E} \left( \exp \left( it \sum_{n=1}^{\infty} \frac{X_n}{3^n} \right) \right) = \mathbb{E} \left( \prod_{n=1}^{\infty} \exp \left( it \frac{X_n}{3^n} \right) \right) \\
= \prod_{n=1}^{\infty} \left[ \cos \left( \frac{X_n t}{3^n} \right) + i \sin \left( \frac{X_n t}{3^n} \right) \right] \\
= \prod_{n=1}^{\infty} \cos \left( \frac{t}{3^n} \right),
\]
where the last equality makes use of the fact that $P(X_n = 1) = P(X_n = -1) = \frac{1}{2}$.

Proposition 45. For all $t \in \mathbb{R}$,
\[
\frac{\sin t}{t} = \prod_{n=1}^{\infty} \cos \left( \frac{t}{2^n} \right).
\]

Proof. Let $Z = \sum_{n=1}^{\infty} \frac{X_n}{2^n}$. By the previous result, we have
\[
\mathbb{E} \left( \exp \left( it \sum_{n=1}^{\infty} \frac{X_n}{2^n} \right) \right) = \prod_{n=1}^{\infty} \cos \left( \frac{t}{2^n} \right).
\]
On the other hand,
\[
\mathbb{E} \left( \exp \left( it \sum_{n=1}^{\infty} \frac{X_n}{2^n} \right) \right) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathbb{E} \left( Z^n \right) \\
= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathbb{E} \left( Z^n \right) \\
= \sum_{j=0}^{\infty} \frac{(-1)^j t^{2j}}{(2j)!} \mathbb{E} \left( Z^{2j} \right) \\
= \frac{1}{t} \sum_{j=0}^{\infty} \frac{(-1)^j t^{2j+1}}{(2j)!} \frac{1}{2j + 1} \\
= \frac{\sin t}{t}.
\]

Problem 5
Let $X_1, X_2, \ldots, X_n$ be IID with a standard Cauchy distribution, that is, the common density with respect to Lebesgue measure is $f(x) = 1/[\pi(1 + x^2)]$ for $x \in \mathbb{R}$. As we have seen in class the chf of $X$ is $\exp \{-|t|\}$. 

32
Proposition 46. Let $\overline{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}$. The distribution of $\overline{X}_n$ is given by

$$\phi_{\overline{X}_n}(t) = \mathbb{E} \left( \exp(it\overline{X}_n) \right) = \mathbb{E} \left( \exp \left( \frac{tS_n}{n} \right) \right) = e^{-n\mathbb{E} \left( \exp(itS_n) \right)} = e^{-n\phi_X(t)^n} = e^{-n e^{-|t|}} = \phi_X(t).$$

Thus, $\overline{X}_n$ is distributed identically to each of the $X_i$. \qed

Problem 1

Let $\{X_n, n \geq 1\}$ be independent random variables with $\mathbb{E}(X_n) = 0$ and $\mathbb{V}(X_n) = \sigma_n^2 < \infty$. Let $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$.

Proposition 47. If there exists a $\delta > 0$ such that, as $n \to \infty$,

$$\sum_{i=1}^{n} \mathbb{E}(|X_i|^2 + \delta) \to 0,$$

then the Lindeberg condition below holds:

$$\forall \epsilon > 0 : \lim_{n \to \infty} \frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}(X_i^2 I\{|X_i| > \epsilon s_n\}) \to 0.$$

Proof. Let $\epsilon > 0$ be given. It follows directly that

$$\frac{1}{s_n^2} \sum_{i=1}^{n} \mathbb{E}(X_i^2 I\{|X_i| > \epsilon s_n\}) = \sum_{i=1}^{n} \mathbb{E} \left( \frac{X_i}{s_n} \right)^2 \mathbb{E}(X_i I\{|X_i| > \epsilon s_n\})$$

$$\leq \sum_{i=1}^{n} \mathbb{E} \left( \frac{X_i}{s_n} \right)^2 \mathbb{E}(\frac{|X_i|}{\epsilon s_n} \delta \mathbb{I}\{|X_i| > \epsilon s_n\})$$

$$\leq \frac{1}{\epsilon^\delta} \sum_{i=1}^{n} \mathbb{E}(X_i^2 + \delta) \frac{s_n^2}{s_n^2}$$

$$\to 0. \qed$$

Remark 2. Define $S_n = \sum_{i=1}^{n} X_i$. Since the Lindeberg condition holds for $\{X_n\}$, the Lindeberg-Feller central limit theorem implies that $\frac{S_n}{\sigma_n^2} \Rightarrow N(0,1)$. 33
Problem 2

Let \( \{X_n, n \geq 1\} \) be a sequence of independent random variables with mean zero and variance \( \sigma_n^2 \).

**Proposition 48.** Even if there exists a \( B > 0 \) such that \( \forall n : \frac{1}{B} \leq \sigma_n^2 \leq B \), it does not follow that

\[
\frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}} \Rightarrow N(0,1).
\]

**Proof.** For a counterexample, define \( X_n = nI([0, \frac{1}{n^2}]) - nI([1 - \frac{1}{n^2}, 1]) \). We have, for all \( n \),

\[
E(X_n) = n \cdot \frac{1}{n^2} - n \cdot \frac{1}{n^2} = 0
\]

and

\[
V(X_n) = E(X_n^2) + (E(X_n))^2 = n^2 \cdot \frac{1}{n^2} + n^2 \cdot \frac{1}{n^2} = 2.
\]

Let now \( Y_n = \sum_{i=1}^{n} X_i \). For all \( n \), \( Y_n \) is equal to zero on the interval \([\frac{1}{n^2}, \frac{3}{n^2}]\). Thus, the distribution of the limiting random variable is constant on this interval, and so in particular is not normal.

**Proposition 49.** If, in addition, for some \( B < \infty \), \( \forall n : P\{|X_n| \leq B\} = 1 \), ie., the rvs are bounded, then (1) holds.

**Proof.** We establish the condition in Problem 1 and then appeal to the Lindeberg-Feller central limit theorem.

We have

\[
\sum_{i=1}^{n} \frac{E(|X_i|^{2+\delta})}{s_n^{2+\delta}} \leq \sum_{i=1}^{n} \frac{E(B^{2+\delta})}{(\frac{n}{B})^{2+\delta}} = \frac{nB^{2+\delta}}{n^{2+\delta}} \to 0.
\]
Problem 3

Suppose that $Y_s$ has a Poisson distribution with parameter $s$, not necessarily an integer, so that

$$P\{Y_s = k\} = \frac{\exp(-s)s^k}{k!}, k = 0, 1, \ldots.$$ 

Proposition 50. As $s \to \infty$,

$$\frac{Y_s - s}{\sqrt{s}} \Rightarrow N(0, 1).$$

Proof. Let $Z_s = \frac{Y_s - s}{\sqrt{s}}$. By the Uniqueness Theorem, it suffices to show that the characteristic function of $Z_s$ converges to the characteristic function of a standard normal random variable.

We have

$$\phi_{Z_s}(t) = \phi_{Y_s}\left(\frac{t}{\sqrt{s}}\right) \exp(-it\sqrt{s})$$

$$= \exp\left(s(e^{\frac{it}{\sqrt{s}}} - 1)\right) \exp(-it\sqrt{s})$$

$$= \exp\left(se^{\frac{it}{\sqrt{s}}} - s - it\sqrt{s}\right).$$

Taking logarithms, we proceed by showing the exponent converges to $-\frac{t^2}{2}$. 

$$se^{\frac{it}{\sqrt{s}}} - s - it\sqrt{s} = s \sum_{k=0}^{\infty} \frac{(it)^k}{(\sqrt{s})^k k!} - s - it\sqrt{s}$$

$$= s \left(1 + \frac{it}{\sqrt{s}} - \frac{t^2}{2s} - \frac{it^3}{6s^{\frac{3}{2}}} + \cdots\right) - s - it\sqrt{s}$$

$$= \left(s + it\sqrt{s} - \frac{t^2}{2} - \frac{it^3}{6s^{\frac{3}{2}}} + \cdots\right) - s - it\sqrt{s}$$

$$= -\frac{t^2}{2} - \frac{it^3}{6s^{\frac{3}{2}}} + \cdots.$$ 

Since $t$ is constant, the above converges to $-\frac{t^2}{2}$ as $s \to \infty$, as desired. \qed

Problem 4

Let $\{X_n, n \geq 1\}$ be independent random variables with $X_n \sim N(0, \sigma_n^2)$. Let $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$ and $S_n = \sum_{i=1}^{n} X_i$.

Proposition 51. If the $\sigma_n^2$ are chosen such that $\max_{i \leq n} \sigma_i^2/s_n^2$ does not converge to zero as $n \to \infty$, then the Lindeberg condition does not hold.
Proof. Define $\sigma_k = 2^{-(k-1)}$. We have, for all $n$,
\[
\max_{i \leq n} \frac{\sigma_i^2}{s_n^2} = \frac{1}{\sum_{k=1}^{n} 2^{-(k-1)}}
\]
and so $\max_{i \leq n} \frac{\sigma_i^2}{s_n^2}$ does not converge to zero.

As for the Lindeberg condition, observe first that $s_n \to \sqrt{\frac{2}{2}}$. Choose $\epsilon_0$ such that $P\{|X_1| > \epsilon_0 \sqrt{\frac{2}{2}}\} > 0$ (such $\epsilon_0$ exists since $X_1$ is not identically zero). We have, for all $n$,
\[
\sum_{i=1}^{n} E[X_i^2 I\{|X_i| > \epsilon_0 s_n\}] \geq E[X_1^2 I\{|X_1| > \epsilon_0 s_n\}] \\
\to E[X_1^2 I\{|X_1| > \epsilon_0 \sqrt{\frac{2}{2}}\}].
\]
By our choice of $\epsilon_0$, the expectation above is equal to some positive constant not depending on $n$. Therefore, the Lindeberg condition does not hold.

**Proposition 52.** For the choice of $\sigma_n^2$ in the proposition above, we have $S_n/s_n \Rightarrow N(0,1)$. [Remark: This shows that the Lindeberg condition is not necessary for the CLT to hold.]

**Proof.** We know that $S_n \sim N(0, s_n^2)$, and so the characteristic function $\phi_{S_n}(t) = \exp \left(-\frac{1}{2} t^2 s_n^2\right)$. Let now $Y_n = \frac{S_n}{s_n}$. We have
\[
\phi_{Y_n}(t) = \phi_{S_n} \left(\frac{t}{s_n}\right) \\
= \exp \left(-\frac{1}{2} \left(\frac{t}{s_n}\right)^2 s_n^2\right) \\
= \exp \left(-\frac{t^2}{2}\right).
\]
Therefore, by the Uniqueness Theorem, $\frac{S_n}{s_n}$ converges to a standard normal random variable.

**Proposition 53.** For the choice of $\sigma_n^2$ in the above proposition, the following condition does not hold:
\[
\forall \epsilon > 0, \max_{i \leq n} P\{|X_i| > \epsilon s_n\} \to 0.
\]
Proof. Choose $\epsilon_0$ such that $P\{|X_1| > \epsilon_0 \sqrt{2}/2\} > 0$ (such $\epsilon_0$ exists since $X_1$ is not identically zero). We have, for all $n$,

$$\max_{i \leq n} P\{|X_i| > \epsilon_0 s_n\} \geq P\{|X_1| > \epsilon_0 s_n\}$$

$$\rightarrow P \left\{ \frac{|X_1| > \epsilon_0 \sqrt{2}}{2} \right\}.$$  

By our choice of $\epsilon_0$, the probability above is equal to some positive constant not depending on $n$. Therefore, the condition does not hold. \qed