Problem 1. Let $X_1, X_2, \ldots$ be independent and identically distributed exponential random variables with parameter $\lambda$. Prove

$$\limsup_{n \to \infty} \frac{X_n}{\log n} = \frac{1}{\lambda}$$

almost surely.

Solution 1. Fix an $\epsilon > 0$. Since $P(X_n > t) = e^{-\lambda t}$ and

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

we have by the first Borel-Cantelli lemma that

$$\limsup_{n \to \infty} \frac{X_n}{\log n} \leq \frac{1}{\lambda} + \epsilon$$

almost surely.

Also, since the random variables $X_1, X_2, \ldots$ are independent and

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

we have by the second Borel-Cantelli lemma that

$$\liminf_{n \to \infty} \frac{X_n}{\log n} \geq \frac{1}{\lambda}$$

almost surely.

By the sequential continuity of $P$ the event

$$\left\{ \limsup_{n \to \infty} \frac{X_n}{\log n} = \frac{1}{\lambda} \right\} = \bigcap_{k=1}^{\infty} \left\{ \frac{1}{\lambda} \leq \limsup_{n \to \infty} \frac{X_n}{\log n} \leq \frac{1}{\lambda} + \frac{1}{k} \right\}$$

has probability one.

Problem 2. Let $X_1, X_2, \ldots$ be independent absolutely continuous random variables such that $X_n$ has density function $f_n$ given by

$$f_n(x) = \frac{n}{\pi(1 + n^2 x^2)}.$$ 

With respect to which modes of convergence does $X_n$ converge to zero as $n \to \infty$?
Solution 2. Since $\mathbb{E}|X_n|^p = +\infty$ for all $p \geq 1$, the sequence $X_1, X_2, \ldots$ does not converge in any $L^p$. On the other hand for every $\epsilon > 0$ we have

$$\mathbb{P}(|X_n| > \epsilon) = \frac{2}{\pi} \tan^{-1}\left(\frac{1}{\epsilon n}\right) \approx \frac{2}{\pi n \epsilon}$$

for large $n$. Hence $X_n \to 0$ in probability and in distribution. But since $X_1, X_2, \ldots$ are independent and $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \epsilon) = +\infty$ then $|X_n| > \epsilon$ infinitely often almost surely by the second Borel-Cantelli lemma. Thus $X_n$ does not converge to zero almost surely.

Problem 3. Let $M(t) = \mathbb{E}(e^{tX})$ be the moment generating function of a random variable $X$. Prove

$$\mathbb{P}(X \geq \epsilon) \leq \inf_{t \geq 0} e^{-\epsilon t} M(t).$$

Solution 3. For every $t \geq 0$ we have

$$\mathbb{P}(X \geq \epsilon) = \mathbb{P}(e^{tX} \geq e^{\epsilon t}) \leq \frac{\mathbb{E}(e^{tX})}{e^{\epsilon t}}$$

by Markov’s inequality. Since the inequality holds for each $t \geq 0$, it holds for the infimum.

Problem 4. Let $X$ and $Y$ be jointly normal with zero means and unit variances and correlation $\rho$. Prove

$$\mathbb{P}(X > 0, Y > 0) = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \rho.$$ 

Solution 4. Let $Z = (1 - \rho^2)^{-1/2}(Y - \rho X)$. Since the random variables $X$ and $Y$ are jointly normal, the random variables $X$ and $Z$ are also jointly normal. Also, since $\text{Cov}(X, Z) = 0$, then $X$ and $Z$ are independent.

By changing from Cartesian to polar coordinates $(x, z) \mapsto (r, \theta)$ we have

$$\mathbb{P}(X > 0, Y > 0) = \mathbb{P}(X > 0, \rho X + (1 - \rho^2)^{1/2} Z > 0)$$

$$= \int_{x>0, \rho x + (1 - \rho^2)^{1/2} z > 0} \frac{1}{2\pi} e^{-(x^2 + z^2)/2} dx \, dz$$

$$= \frac{1}{2\pi} \int_{\theta = -\sin^{-1} \rho}^{\pi/2} \int_{r=0}^\infty r e^{-r^2/2} dr \, d\theta$$

$$= \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \rho.$$ 

Problem 5. If $X_n \to X$ in $L_1$ then prove $\mathbb{E}(X_n) \to \mathbb{E}(X)$, If $X_n \to X$ in $L_2$ then prove $\text{Var}(X_n) \to \text{Var}(X)$.

Solution 5. If $X_n \to X$ in $L^1$ then $|\mathbb{E}(X_n) - \mathbb{E}(X)| = |\mathbb{E}(X_n - X)| \leq \mathbb{E}|X_n - X| \to 0$ by Jensen’s inequality.
Now if \( X_n \to X \) in \( L^2 \), then \( \mathbb{E}|X_n - X| \leq \mathbb{E}[(X_n - X)^2]^{1/2} \) by Jensen’s inequality, so \( X_n \to X \) in \( L^1 \) and, by the first part, \( \mathbb{E}(X_n) \to \mathbb{E}(X) \). Furthermore,

\[
|\mathbb{E}(X_n^2) - \mathbb{E}(X^2)| \leq \mathbb{E}(|X_n - X||X_n + X|) \\
\leq \mathbb{E}[(X_n - X)^2]^{1/2}\mathbb{E}[(X_n + X)^2]^{1/2}.
\]

Since \( \mathbb{E}[(X_n - X)^2] \to 0 \) there exists an \( N \) such that \( \mathbb{E}[(X_n - X)^2] \leq 1 \) for all \( n \geq N \). Hence, for \( n \geq N \) we have

\[
\mathbb{E}[(X_n + X)^2] = \mathbb{E}[(X_n - X + 2X)^2] \\
\leq 2\mathbb{E}[(X_n - X)^2] + 8\mathbb{E}(X^2) \\
\leq 2 + 8\mathbb{E}(X^2)
\]

and we conclude that \( \mathbb{E}[(X_n + X)^2] \) is bounded uniformly in \( n \). Hence \( \mathbb{E}(X_n^2) \to \mathbb{E}(X^2) \) and thus \( \text{Var}(X_n) \to \text{Var}(X) \) as desired.

**Problem 6.** Let \( X_1, X_2, \ldots \) be a sequence of random variables such that \( \mathbb{E}\sup_{n \geq 1} |X_n| < \infty \). If \( X_n \to X \) in probability, then prove \( X_n \to X \) in \( L^1 \).

**Solution 6.** Let \( Z = \sup_{n \geq 1} |X_n| \). We first prove that \( |X| \leq Z \) almost surely. We could appeal to Problem 11 to assert the existence of a subsequence \( (X_{n_k})_{k \geq 1} \) such that \( X_{n_k} \to X \) almost surely. Hence \( |X| \leq \sup_{k \geq 1} |X_{n_k}| \leq Z \) almost surely.

Alternatively,

\[
\mathbb{P}(|X| > Z + \epsilon) = \mathbb{P}(|X| > Z + \epsilon, |X_n - X| \geq \epsilon) + \mathbb{P}(|X| > Z + \epsilon, |X_n - X| < \epsilon) \\
\leq \mathbb{P}(|X_n - X| \geq \epsilon) + \mathbb{P}(|X_n| > Z) \\
\to 0
\]

as \( n \uparrow \infty \) and hence \( X \leq Z + \epsilon \) almost surely. By either method, we can conclude that the event \( \{X \leq Z\} = \cap_{k=1}^{\infty} \{X \leq Z + 1/k\} \) has probability one. Note, in particular, the inequality \( |X_n - X| \leq |X_n| + |X| \leq 2Z \) holds almost surely.

Next we claim that if \( A_1, A_2, \ldots \) is a sequence of events with \( \mathbb{P}(A_n) \to 0 \) then \( \mathbb{E}(Z\mathbb{1}_{A_n}) \to 0 \). Pick a \( z > 0 \) and note that

\[
\mathbb{E}(Z\mathbb{1}_{A_n}) = \mathbb{E}(Z\mathbb{1}_{A_n \cap \{Z > z\}}) + \mathbb{E}(Z\mathbb{1}_{A_n \cap \{Z \leq z\}}) \\
\leq \mathbb{E}(Z\mathbb{1}_{\{Z > z\}}) + z\mathbb{P}(A_n).
\]

Letting \( n \uparrow \infty \) first and then \( z \uparrow \infty \) proves the claim.

Now, fix an \( \epsilon > 0 \).

\[
\mathbb{E}|X_n - X| = \mathbb{E}(|X_n - X|\mathbb{1}_{\{|X_n - X| \leq \epsilon\}}) + \mathbb{E}(|X_n - X|\mathbb{1}_{\{|X_n - X| > \epsilon\}}) \\
\leq \epsilon + 2\mathbb{E}(Z\mathbb{1}_{\{|X_n - X| > \epsilon\}})
\]

Letting \( A_n = \{|X_n - X| > \epsilon\} \) in the above claim and letting \( \epsilon \downarrow 0 \) implies \( \mathbb{E}|X_n - X| \to 0 \) as desired.

[Note that this is a version of the dominated convergence theorem that holds under a weaker hypothesis than the one proved in the lecture.]
Problem 7. Let $X_2, X_3, \ldots$ be a sequence of independent random variables such that
\[
P(X_n = n) = P(X_n = -n) = \frac{1}{2n \log n}; \quad P(X_n = 0) = 1 - \frac{1}{n \log n}.
\]
Let $S_n = X_2 + \ldots + X_n$. Prove that $S_n/n \to 0$ in probability, but not almost surely. (That is, the weak law of large numbers holds, but not the strong law.)

Solution 7. Since the $X_n$’s are independent and $E(X_n) = 0$ for all $n \in \mathbb{N}$ we have
\[
E(S_n^2) = \sum_{k=2}^{n} \frac{n}{\log n}.
\]
Hence by Chebyshev’s inequality,
\[
P\left(\frac{|S_n|}{n} > \epsilon \right) < \frac{1}{n^2 \epsilon^2} \sum_{k=2}^{n} \frac{n}{\log n} < \frac{1}{\epsilon^2 \log n} \to 0
\]
where we have made use of the fact that $x \to x/\log x$ is increasing for $x > e$. Hence $S_n/n \to 0$ in probability.

On the other hand, since the random variables $X_2, X_3, \ldots$ are independent and
\[
\sum_{n=2}^{\infty} P(X_n = n) = \sum_{n=2}^{\infty} \frac{1}{2n \log n} = +\infty,
\]
the second Borel-Cantelli lemma asserts that there exists an event $E \subset \Omega$ with $P(E) = 1$ such that for all $\omega \in E$ the equation $X_n(\omega) = n$ is satisfied for an infinite number of $n$’s. But if $X_n(\omega) = n$ then
\[
\frac{S_n(\omega)}{n} - \frac{S_{n-1}(\omega)}{n-1} = \frac{X_n(\omega) + S_{n-1} - S_{n-1}(\omega)}{n} - \frac{S_{n-1}(\omega)}{n-1}
\]
\[
= 1 - \frac{S_{n-1}(\omega)}{n(n-1)}
\]
\[
\geq 1/2
\]
where we have used the inequality $S_{n-1} = X_2 + \ldots + X_{n-1} \leq 1 + \ldots + n - 1 = \frac{n(n-1)}{2}$. For each $\omega \in E$ the inequality $\frac{S_n(\omega)}{n} - \frac{S_{n-1}(\omega)}{n-1} \geq 1/2$ holds for infinitely many $n$’s. Thus the sequence $(S_n(\omega)/n)_n$ diverges for each $\omega \in E$.

Problem 8. Let $X$ be a random variable and let $K_X(t) = \log E(e^{tX})$ be the logarithm of the moment generating function. Prove that $K_X$ is convex. Suppose that $K_X$ has a Taylor series
\[
K_X(t) = \sum_{n=1}^{\infty} \frac{1}{n!} k_n(X)t^n.
\]
Compute $k_1(X)$, $k_2(X)$, and $k_3(X)$ in terms of the moments of $X$. If $X$ and $Y$ are independent, prove that $k_n(X + Y) = k_n(X) + k_n(Y)$ for all $n \geq 1$. 

4
Solution 8. We give two proofs of the convexity of $K_X$. Proof 1:

\[ K''_X(t) = \frac{\mathbb{E}(X^2 e^{tX})\mathbb{E}(e^{tX}) - \mathbb{E}(X e^{tX})^2}{\mathbb{E}(e^{tX})^2} \geq 0 \]

by the Cauchy–Schwarz inequality $\mathbb{E}(YZ) \leq [\mathbb{E}(Y^2)]^{1/2}[\mathbb{E}(Z^2)]^{1/2}$ with $Y = e^{tX/2}$ and $Z = X e^{tX/2}$.

Proof 2: Let $0 < \lambda < 1$.

\[ K_X(\lambda s + (1 - \lambda)t) = \log \mathbb{E}[(e^{sX})^{\lambda}(e^{tX})^{1-\lambda}] \leq \log[\mathbb{E}(e^{sX})^{\lambda}\mathbb{E}(e^{tX})]^{1-\lambda} = \lambda K_X(s) + (1 - \lambda)K_X(t) \]

where we have used Hölder’s inequality $\mathbb{E}(YZ) \leq [\mathbb{E}(Y^p)]^{1/p}[\mathbb{E}(Z^q)]^{1/q}$ where $1/p + 1/q = 1$, with $Y = e^{sX}$, $Z = e^{tX}$, and $p = 1/\lambda$. [Note that the Cauchy–Schwarz inequality is the special case of Hölder’s inequality with $p = 2$.]

Now, if a function $f$ has a Taylor series

\[ f(t) = 1 + a_1t + a_2t^2 + a_3t^3 + \ldots \]

converging on some neighborhood of the origin, then

\[ \log f(t) = a_1t + (a_2 - \frac{1}{2}a_1^2)t^2 + (a_3 - a_1a_2 + \frac{1}{3}a_1^3)t^3 + \ldots \]

converging on some (possibly smaller) neighborhood of the origin.

Now, for a moment generating function $M_X(t) = \mathbb{E}(e^{tX})$, finite in some neighborhood of the origin, we have

\[ M_X(t) = 1 + \sum_{n=1}^{\infty} \frac{\mathbb{E}(X^n)}{n!} t^n. \]

Matching coefficients yields $k_1(X) = \mathbb{E}(X)$, $k_2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \text{Var}(X)$, and $k_3(X) = \mathbb{E}(X^3) - 3\mathbb{E}(X)\mathbb{E}(X^2) + 2\mathbb{E}(X)^3$.

[The number $k_n(X)$ is called the $n$-th cumulant of $X$, and the function $K_X$ is called the cumulant generating function. Note that a random variable is normal if and only if its cumulant generating function is quadratic.]

If $X$ and $Y$ are independent then

\[ \sum_{n=1}^{\infty} \frac{1}{n!} k_n(X + Y)t^n = \log \mathbb{E}(e^{t(X+Y)}) = \log \mathbb{E}(e^{tX}) + \log \mathbb{E}(e^{tY}) = \sum_{n=1}^{\infty} \frac{1}{n!} k_n(X)t^n + \sum_{n=1}^{\infty} \frac{1}{n!} k_n(Y)t^n \]

and $k_n(X + Y) = k_n(X) + k_n(Y)$ by the uniqueness of Taylor series.
**Problem 9.** Let $X_1, X_2, \ldots$ be a sequence of independent random variables each uniformly distributed on $[0, 1]$. Let $M_n = \min\{X_1, \ldots, X_n\}$. Prove that $nM_n$ converges in distribution to an exponential random variable with parameter one.

**Solution 9.** For $t \in [0, 1]$ we have
\[
\mathbb{P}(M_n > t) = \mathbb{P}(X_1 > t, \ldots, X_n > t) = (1 - t)^n
\]
Hence, the distribution function $F_n(t) = \mathbb{P}(nM_n \leq t)$ of $nM_n$ is given by
\[
F_n(t) = \begin{cases} 
0 & \text{if } t < 0 \\
1 - (1 - t/n)^n & \text{if } 0 \leq t < n \\
1 & \text{if } t \geq n
\end{cases}
\]
and $F_n(t) \to 1 - e^{-t}$ for all $t \geq 0$.

**Problem 10.** If a sequence of random variables $X_n \to c$ in distribution, where $c$ is a constant, then prove $X_n \to c$ in probability.

**Solution 10.** We are given that $\mathbb{P}(X_n \leq t) \to 1_{[c, \infty)}(t)$ for all $t \neq c$. Fixing $\epsilon > 0$ we have
\[
\mathbb{P}(|X_n - c| \leq \epsilon) = \mathbb{P}(c - \epsilon \leq X_n \leq c + \epsilon) = \mathbb{P}(X_n \leq c + \epsilon) - \mathbb{P}(X_n < c - \epsilon) \geq \mathbb{P}(X_n \leq c + \epsilon) - \mathbb{P}(X_n \leq c - \epsilon) \to 1_{[c, \infty)}(c + \epsilon) - 1_{[c, \infty)}(c - \epsilon) = 1
\]
and hence $\mathbb{P}(|X_n - c| > \epsilon) \to 0$ as desired.

**Problem 11.** Let $X_n \to X$ in probability. Prove that there exists a subsequence $X_{n_1}, X_{n_2}, \ldots$ such that $X_{n_k} \to X$ almost surely.

**Solution 11.** Since $\mathbb{P}(|X_n - X| > \epsilon) \to 0$ for each $\epsilon > 0$, there exists a subsequence $X_{n_1}, X_{n_2}, \ldots$ such that $\mathbb{P}(|X_{n_k} - X| > 1/k) \leq 1/k^2$. For this subsequence we have
\[
\sum_{k=1}^{\infty} \mathbb{P}(|X_{n_k} - X| > 1/k) < \infty
\]
so that by the first Borel–Cantelli lemma, $\mathbb{P}(|X_{n_k} - X| > 1/k$ infinitely often) = 0. We are done since the events $\{|X_{n_k} - X| \leq 1/k$ eventually} and $\{X_n \to X\}$ are equal.

**Problem 12.** Let $U$ be uniformly distributed on $[0, 1]$. Let the conditional distribution of $X$ given $U$ be binomial with parameters $U$ and $n$. Find the distribution of $X$.

**Solution 12.** If $Y$ is binomial with parameters $n$ and $p$, then there are $n$ independent Bernoulli random variables with parameter $p$ such that $Y = Z_1 + \ldots + Z_n$. In particular, the probability generating function of $Y$ is given by $G_Y(s) = (G_Z(s))^n = (1 - p + ps)^n$. 

6
Now consider the probability generating function of $X$:

\[
G_X(s) = \mathbb{E}(s^X) = \mathbb{E}(\mathbb{E}(s^X|U)) = \mathbb{E}[(1 + (s - 1)U)^n] = \frac{1}{n+1} \frac{s^{n+1} - 1}{s - 1} = \frac{1}{n+1}(1 + s + s^2 + \ldots + s^n).
\]

Hence $X$ is uniformly distributed on $\{0, 1, \ldots, n\}$.

[Alternatively, one could proceed directly:

\[
\mathbb{P}(X = k) = \mathbb{E}[\mathbb{P}(X = k|U)] = \mathbb{E}\binom{n}{k} U^k (1 - U)^{n-k} = \frac{1}{n+1}.
\]

This approach requires the formula:

\[
\int_0^1 u^m (1-u)^n du = \frac{m!n!}{(m+n+1)!}.
\]

The above formula is true in more generality:

\[
\int_0^1 u^{s-1} (1-u)^{t-1} du = \frac{\Gamma(s) \Gamma(t)}{\Gamma(s+t)}
\]

for $s, t > 0$ where $\Gamma(q) = \int_0^\infty x^{q-1} e^{-x} dx$ is the gamma function. The above integral defines the beta function $B(s, t)$.

**Problem 13.** The random variables $X$ and $Y$ are distributed uniformly on the disk $\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ Find the density of the random variable $X/Y$.

**Solution 13.** The joint density $f$ of $(X, Y)$ is given by $f(x, y) = 1/\pi$ for $x^2 + y^2 \leq 1$.

\[
\mathbb{P}(X/Y \leq t) = \int \int_{x/y \leq t, x^2+y^2<1} \frac{1}{\pi} dx \, dy = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} t.
\]

Hence, the density $f$ of $X/Y$ is given by $f(t) = \frac{1}{\pi \left(1+t^2\right)}$. That is, the random variable $X/Y$ has the Cauchy distribution.