ECE 534 SP26 HW4 Solutions

March 27, 2026

Problem 1. Prove the Kolmogorov 0-1 Law (you can look up the notes by Prof. Lévêque [1], but are responsible for submitting your own explanation of the proof). In the process, describe two more examples of tail events.

Solution.

We first state the Kolmogorov 0-1 Law:

Theorem 1.1 (Kolmogorov 0-1 Law)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of mutually independent random variables. For each n, let*

$$\mathcal{G}_n := \sigma(X_n, X_{n+1}, \ldots),$$

and then define the tail σ-algebra *as*

$$\mathcal{T} := \bigcap_{n\geq 1} \mathcal{G}_n. \tag{1.1}$$

Then, for any event $A \in \mathcal{T}$ (such an event is called a tail event*), we have either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.*

Proof of the Kolmogorov 0-1 Law. Let A be a tail event (i.e. $A \in \mathcal{T}$). We first show that A is independent of itself. Note that by the independence of $(X_n)_{n\geq 1}$, we have for each positive integer n that $\mathcal{G}_{n+1}$ is independent of the following σ-algebra

$$\mathcal{F}_n := \sigma(X_1, \ldots, X_n).$$

Then, from the definition of $\mathcal{T}$ in (1.1), we know that $A \in \mathcal{G}_{n+1}$ for every positive integer n, and thus A is independent of any event in $\mathcal{F}_n$ for every positive integer n. That is, A is independent of any event that depends on finitely many X_is. It then follows "by definition" that[1] A is independent of any event that depends on the whole sequence $(X_n)_{n\geq 1}$. In other words, A is independent of any event in $\sigma(X_1, X_2, \ldots)$. At the same time, A is clearly in $\sigma(X_1, X_2, \ldots)$ as well, since $A \in \mathcal{T} = \bigcap_{n\geq 1} \mathcal{G}_n \subseteq \mathcal{G}_1 = \sigma(X_1, X_2, \ldots)$. Therefore, A is independent of itself.

Now we have

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A \cap A) &(1.2)\\ &= \mathbb{P}(A)^2, &(1.3)\end{aligned}$$

where (1.2) followed from the identity that $A = A \cap A$ and in (1.3) we used the fact that A is independent of itself. This implies that either $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$. ☐

Examples of tail events: By the definition of $\mathcal{T}$ in (1.1), an event A is a tail event if and only if

$$A \in \mathcal{G}_n \text{ for all } n \geq 1.$$

That is, for any positive integer n, A only depends on $X_n, X_{n+1}, X_{n+2}, \ldots$. Such a condition is usually satisfied when we are discussing the limiting behavior of $(X_n)_{n\geq 1}$. In the following we give two examples of tail events:

Example 1. Let

$$A_1 := \left\{\omega \in \Omega \ : \ \lim_{n\to\infty} X_n(\omega) \text{ exists}\right\}.$$

[1] In fact, this step can be rigorously justified using Dynkin's π-λ Theorem: https://en.wikipedia.org/wiki/Dynkin_system#Sierpi%C5%84ski%E2%80%93Dynkin's_%CF%80-%CE%BB_theorem (also called Monotone Class Theorem or Monotone Class Lemma), which is out of the scope of this course. Here, you can extend the independence between A and any $\sigma(X_1, \ldots, X_n)$ to the independence between A and $\sigma(X_1, \ldots)$ and treat this as a definition. This footnote is made to avoid confusion in case you ever try to formally prove this step using only the definitions in the notes, such as [1, Definition 3.8] and [1, Definition 1.16].

It is known that for any real-valued sequence $(a_n)_{n\geq 1}$ and any positive integer k, the limit of $(a_1, a_2, a_3, \ldots)$ exists if and only if the limit of $(a_k, a_{k+1}, a_{k+2}, \ldots)$ exists. Therefore, for any positive integer k, we have

$$A_1 = \{\omega \in \Omega \ : \ \text{The limit of } (X_k(\omega), X_{k+1}(\omega), \ldots) \text{ exists}\},$$

which depends only on $X_k, X_{k+1}, \ldots$. Therefore, we have $A_1 \in \mathcal{G}_k$ for all $k \geq 1$, and thus $A_1 \in \bigcap_{k\geq 1} \mathcal{G}_k = \mathcal{T}$.

Example 2. Let

$$A_2 := \left\{\omega \in \Omega \ : \ \lim_{n\to\infty} \frac{X_1(\omega) + \cdots + X_n(\omega)}{n} \text{ exists}\right\}.$$

It is known that for any real-valued sequence $(a_n)_{n\geq 1}$ and any positive integer k,

$$\lim_{n\to\infty} \frac{a_1 + \ldots + a_n}{n} \text{ exists} \Leftrightarrow \lim_{n\to\infty} \frac{a_k + \ldots + a_n}{n} \text{ exists},$$

since $\frac{a_1+\cdots+a_{k-1}}{n}$ goes to 0 as n goes to ∞. Therefore, for any positive integer k, we have

$$A_2 = \left\{\omega \in \Omega \ : \ \lim_{n\to\infty} \frac{X_k(\omega) + \cdots + X_n(\omega)}{n} \text{ exists}\right\},$$

which implies $A_2 \in \mathcal{G}_k$ for all $k \geq 1$. It follows that $A_2 \in \bigcap_{k\geq 1} \mathcal{G}_k = \mathcal{T}$.

Remark. This example agrees with the two cases of the Strong Law of Large Numbers (see the beginning of [1, Section 5.7]).

Problem 2. Compute the characteristic function of a Gaussian $\mathcal{N}(\mu, \sigma^2)$ random variable.

Solution. Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Its characteristic function is

$$\varphi_X(t) \ = \ \exp\Big(i\mu t - \tfrac{1}{2}\sigma^2 t^2\Big).$$

Proof. By definition, the characteristic function of X is given by

$$\varphi_X(t) := \mathbb{E}\big[e^{itX}\big] = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\!\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx,$$

where $t \in \mathbb{R}$.

Step 1. Standardize. Let $z = \dfrac{x-\mu}{\sigma}$, so $x = \sigma z + \mu$ and $dx = \sigma\, dz$:

$$\begin{aligned}\varphi_X(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\!\big(it(\sigma z + \mu)\big)\, e^{-z^2/2}\, dz \\ &= e^{i\mu t} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{i\sigma t\, z}\, e^{-z^2/2}\, dz.\end{aligned}$$

Step 2. Complete the square. Combine the two exponentials in the integrand:

$$i\sigma t\, z - \frac{z^2}{2} = -\frac{1}{2}\big(z^2 - 2i\sigma t\, z\big) = -\frac{1}{2}\big(z - i\sigma t\big)^2 - \frac{\sigma^2 t^2}{2}.$$

Hence

$$\varphi_X(t) = e^{i\mu t}\, e^{-\sigma^2 t^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\!\left(-\frac{(z - i\sigma t)^2}{2}\right) dz.$$

Step 3. Evaluate the Gaussian integral. We use complex analysis to calculate

$$I := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\!\left(-\frac{(z - i\sigma t)^2}{2}\right) dz.$$

Using the substitution $w = z - i\sigma t$, we see that the integral I of interest is a contour integral of the function $f(w) := \frac{1}{\sqrt{2\pi}} e^{-w^2/2}$ along the horizontal line $\mathrm{Im}(w) = -\sigma t$ in the complex plane. We divide the derivation of this integral into the following sub-steps:

1) Interpret the integral of interest as the limit of the integral along the shifted horizontal path in the complex plane, moving left to right from $-R - i\sigma t$ to $R - i\sigma t$. That is, we have

$$I = \lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-R-i\sigma t}^{R-i\sigma t} e^{-\frac{w^2}{2}}\, dw$$

2) Since $f(w)$ is an entire function, Cauchy's Integral Theorem states that its integral around any closed contour is exactly zero ($\oint_\Gamma f(w)\, dw = 0$). Assuming $\sigma t > 0$ (the case $\sigma t < 0$ can be handled similarly), we define a rectangular closed contour Γ (which lies in the lower half-plane) traversed counterclockwise:
   - γ_1: the bottom horizontal segment from $-R - i\sigma t$ to $R - i\sigma t$ (moving right).
   - γ_2: the right vertical segment from $R - i\sigma t$ to R (moving up).
   - γ_3: the top horizontal segment from R to $-R$ along the real axis (moving left).
   - γ_4: the left vertical segment from $-R$ to $-R - i\sigma t$ (moving down).

   By the theorem, the sum of the integrals is zero: $\int_{\gamma_1} f(w)\, dw + \int_{\gamma_2} f(w)\, dw + \int_{\gamma_3} f(w)\, dw + \int_{\gamma_4} f(w)\, dw = 0$.
3) As $R \to \infty$, the integral over the top path γ_3 evaluates to the negative of the standard Gaussian integral, because it is traversed backwards from right to left:

$$\lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{\gamma_3} e^{-\frac{w^2}{2}}\, dw = \lim_{R\to\infty} \frac{1}{\sqrt{2\pi}} \int_{R}^{-R} e^{-\frac{x^2}{2}}\, dx = -1.$$

4) We can show the integral over γ_2 vanishes by parameterizing the upward path as $w = R - iu$, where u goes from σt down to 0, giving $dw = -i\, du$. Substituting this in, we find the magnitude is bounded:

$$\begin{aligned}
\left| \int_{\gamma_2} e^{-\frac{w^2}{2}}\, dw \right| &= \left| \int_{\sigma t}^{0} e^{-\frac{1}{2}(R-iu)^2} (-i)\, du \right| \\
&= \left| i \int_{0}^{\sigma t} e^{-\frac{1}{2}(R-iu)^2}\, du \right| \\
&\leq \int_{0}^{\sigma t} \left| e^{-\frac{R^2}{2} + \frac{u^2}{2} + iRu} \right| du \\
&= e^{-\frac{R^2}{2}} \int_{0}^{\sigma t} e^{\frac{u^2}{2}}\, du \qquad (2.1)
\end{aligned}$$

   Since $e^{-R^2/2}$ goes to 0 as $R \to \infty$ and $\int_0^{\sigma t} e^{\frac{u^2}{2}}\, du$ does not depend on R, the quantity in (2.1) approaches 0 as $R \to \infty$, implying that the integral over γ_2 goes to 0 as $R \to \infty$. By a similar parameterization, the integral over γ_4 also vanishes.
5) Substituting these limits back into our closed-contour equation, we get $I + 0 + (-1) + 0 = 0$, where we used the fact that the integral along γ_1 is exactly our integral of interest I. We immediately conclude that $I - 1 = 0$, meaning the integral of interest I evaluates exactly to 1.

Step 4. Conclude. Substituting back we obtain

$$\boxed{\varphi_X(t) = \exp\Big(i\mu t - \tfrac{1}{2}\sigma^2 t^2\Big).}$$

□

Problem 3. Prove the CLT following the proof outlined in Prof. Lévêque's notes [1] (which does not use the characteristic function).

Solution. Here we only present the high-level outline of the proof. For details in each part please refer to the Prof. Lévêque's notes [1].

We first state the Central Limit Theorem:

Theorem 3.1 (Central Limit Theorem). *[1, Theorem 6.7] Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables such that both $\mathbb{E}[X_1] = \mu$ and $\mathrm{Var}(X_1) = \sigma^2$ exist and are finite. Define*

$$\begin{aligned}
S_n &\coloneqq X_1 + \cdots + X_n, \\
\widetilde{S}_n &\coloneqq \frac{S_n - n\mu}{\sqrt{n}\sigma}.
\end{aligned}$$

Then, we have

$$\widetilde{S}_n \xrightarrow[n\to\infty]{d} Z \sim \mathcal{N}(0,1).$$

We will prove the theorem under the additional assumption that $\mathbb{E}\left[|X_1|^3\right] < \infty$. To this end, we need the following lemmas:

Lemma 3.1. *[1, Remark 6.6] Let $(X_n)_{n\geq 1}$ be a sequence of random variables and let X be another random variable. Write $\mathcal{C}_b^3(\mathbb{R})$ for the space of all the functions g from $\mathbb{R}$ to $\mathbb{R}$ such that g, g', g'', g''' are bounded and continuous. If $\lim_{n\to\infty} \mathbb{E}\left[g(X_n)\right] = \mathbb{E}\left[g(X)\right]$ for all $g \in \mathcal{C}_b^3(\mathbb{R})$, then $X_n \xrightarrow[n\to\infty]{d} X$.*

Proof Sketch of Lemma 3.1. Fix a continuity point t of the CDF of X. Use cubic splines to approximate the function

$$\begin{aligned} h_t(x) &\coloneqq \mathbb{1}_{x\leq t} \\ &= \begin{cases} 1, & \text{if } x \leq t, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

See the proof of [1, Theorem 6.5] and [1, Remark 6.6] for more details. ☐

Lemma 3.2. *[1, Lemma 6.9] Let X, Y, Z be three random variables satisfying the following properties:*

1) X is independent of both Y and Z.
2) Both $\mathbb{E}\left[|Y|^3\right]$ and $\mathbb{E}\left[|Z|^3\right]$ are finite.
3) $\mathbb{E}\left[Y\right] = \mathbb{E}\left[Z\right]$ and $\mathbb{E}\left[Y^2\right] = \mathbb{E}\left[Z^2\right]$.

Then for all $g \in \mathcal{C}_b^3(\mathbb{R})$, we have

$$\left|\mathbb{E}\left[g(X+Y)\right] - \mathbb{E}\left[g(X+Z)\right]\right| \leq \frac{C}{6}\left(\mathbb{E}\left[|Y|^3\right] + \mathbb{E}\left[|Z|^3\right]\right),$$

where $C \coloneqq \sup_{x\in\mathbb{R}} |g'''(x)|$.

There was a tricky part in the original proof presented in [1] that needs to be handled more carefully, so here is a revised version of the proof.

Proof of Lemma 3.2. Let $R(x,y)$ be the exact remainder of the second-order Taylor expansion of g evaluated at $x+y$. We can define this function explicitly for any real numbers $x, y \in \mathbb{R}$:

$$R(x,y) \coloneqq g(x+y) - g(x) - yg'(x) - \frac{y^2}{2}g''(x). \tag{3.1}$$

By Taylor's theorem for real functions, we know that for any fixed x and y, there exists some real number u between x and $x+y$ such that $R(x,y) = \frac{y^3}{6}g'''(u)$.

Therefore, we can establish a deterministic bound on the magnitude of the remainder function:

$$\begin{aligned} |R(x,y)| &\leq \frac{|y|^3}{6}\sup_{t\in\mathbb{R}} |g'''(t)| \\ &= \frac{C}{6}|y|^3. \end{aligned} \tag{3.2}$$

Because g, g', and g'' are continuous functions, $R(x,y)$ defined in (3.1) is a continuous (and thus Borel-measurable) function. Therefore, $R(X,Y)$ is a valid random variable (i.e. a measurable function), and we can write:

$$g(X+Y) = g(X) + Yg'(X) + \frac{Y^2}{2}g''(X) + R(X,Y).$$

Taking the expectation of both sides, and utilizing the fact that X and Y are independent, we obtain:

$$\mathbb{E}\left[g(X+Y)\right] = \mathbb{E}\left[g(X)\right] + \mathbb{E}\left[Y\right]\mathbb{E}\left[g'(X)\right] + \frac{1}{2}\mathbb{E}\left[Y^2\right]\mathbb{E}\left[g''(X)\right] + \mathbb{E}\left[R(X,Y)\right]. \tag{3.3}$$

Similarly, for the random variable Z, we obtain:

$$\mathbb{E}\left[g(X+Z)\right] = \mathbb{E}\left[g(X)\right] + \mathbb{E}\left[Z\right]\mathbb{E}\left[g'(X)\right] + \frac{1}{2}\mathbb{E}\left[Z^2\right]\mathbb{E}\left[g''(X)\right] + \mathbb{E}\left[R(X,Z)\right]. \tag{3.4}$$

Now, we subtract (3.4) from (3.3). By the assumptions of the lemma, $\mathbb{E}[Y] = \mathbb{E}[Z]$ and $\mathbb{E}\left[Y^2\right] = \mathbb{E}\left[Z^2\right]$. Because of this moment matching, the first three terms perfectly cancel out, leaving us with:

$$\mathbb{E}[g(X+Y)] - \mathbb{E}[g(X+Z)] = \mathbb{E}[R(X,Y)] - \mathbb{E}[R(X,Z)].$$

Applying the triangle inequality, we get:

$$|\mathbb{E}[g(X+Y)] - \mathbb{E}[g(X+Z)]| \leq |\mathbb{E}[R(X,Y)]| + |\mathbb{E}[R(X,Z)]|.$$

Since $|\mathbb{E}[V]| \leq \mathbb{E}[|V|]$ for any random variable V, we have:

$$|\mathbb{E}[g(X+Y)] - \mathbb{E}[g(X+Z)]| \leq \mathbb{E}[|R(X,Y)|] + \mathbb{E}[|R(X,Z)|].$$

Finally, we apply our bound in (3.2) to get

$$\begin{aligned}|\mathbb{E}[g(X+Y)] - \mathbb{E}[g(X+Z)]| &\leq \frac{C}{6}\mathbb{E}\left[|Y|^3\right] + \frac{C}{6}\mathbb{E}\left[|Z|^3\right] \\ &= \frac{C}{6}\left(\mathbb{E}\left[|Y|^3\right] + \mathbb{E}\left[|Z|^3\right]\right).\end{aligned}$$

This completes the proof. □

Remark. In the original proof of this lemma in [1], we actually do not know whether the "functions" U and V are measurable (and thus taking expectations over them or even just calling them "random variables" require more justifications). The reason is that the function U is defined via the Taylor's theorem, and we only know for each $\omega \in \Omega$ that $U(\omega)$ lies between $X(\omega)$ and $X(\omega) + Y(\omega)$. This property alone may not guarantee that U behaves in a measurable way. Therefore, a safer approach is to first derive a deterministic upper bound of $R(x,y)$ and then plug in X and Y, as shown in the proof above.

With the two lemma above, we can prove the Central Limit Theorem.

Proof sketch of the Central Limit Theorem for Finite-Third-Moment RV.
Define $Y_i := \frac{X_i - \mu}{\sqrt{n}\sigma}$. The random variables Y_i are i.i.d. and we can write our standardized sum as $\widetilde{S}_n = \sum_{i=1}^n Y_i$. We first calculate the first two moments of Y_1:

$$\begin{aligned}\mathbb{E}[Y_1] &= 0, \\ \mathbb{E}\left[Y_1^2\right] &= \frac{1}{n}.\end{aligned}$$

Then, we upper-bound $\mathbb{E}\left[|Y_1|^3\right]$ as follows: Note that

$$\mathbb{E}\left[|Y_1|^3\right] = \frac{\mathbb{E}\left[|X_1 - \mu|^3\right]}{n^{3/2}\sigma^3}.$$

Then, since $\mathbb{E}\left[|X_1|^3\right]$ is finite, it can be shown that $\mathbb{E}\left[|X_1 - \mu|^3\right]$ is also finite (make sure you know how to prove this statement). It then follows that

$$\mathbb{E}\left[|Y_1|^3\right] = O(n^{-3/2}). \tag{3.5}$$

Next, we introduce a sequence of "target" random variables $Z_1, \ldots, Z_n$ that are i.i.d. $\mathcal{N}(0, \frac{1}{n})$ random variables independent of all X_i. These Gaussian variables perfectly match the first two moments of Y_1:

$$\begin{aligned}\mathbb{E}(Z_1) &= 0, \\ \mathbb{E}(Z_1^2) &= \frac{1}{n}, \\ \mathbb{E}(|Z_1|^3) &= O(n^{-3/2}).\end{aligned} \tag{3.6}$$

Note that $\sum_{i=1}^n Z_i \sim \mathcal{N}(0,1)$. Therefore, by Lemma 3.1, to show that $\widetilde{S}_n$ converges in distribution to $\mathcal{N}(0,1)$, it suffice to show for each $g \in \mathcal{C}_b^3(\mathbb{R})$ that

$$|\mathbb{E}[g(Y_1 + \cdots + Y_n)] - \mathbb{E}[g(Z_1 + \cdots + Z_n)]| \underset{n\to\infty}{\to} 0. \tag{3.7}$$

We use Lemma 3.2 multiple times to bound the LHS in (3.7). For demonstration purpose, we only show how the calculation works for $n = 3$, and you should be able to generalize the argument to an arbitrary n following the steps in the proof of [1, Lemma 6.10]. We bound $|\mathbb{E}[g(Y_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Z_3)]|$ in the following way:

$$\begin{aligned}
&|\mathbb{E}[g(Y_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Z_3)]| \\
&= |\mathbb{E}[g(Y_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Y_2+Y_3)] \\
&\qquad + \mathbb{E}[g(Z_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Y_3)] \\
&\qquad + \mathbb{E}[g(Z_1+Z_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Z_3)]| \qquad (3.8)\\
&\le |\mathbb{E}[g(Y_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Y_2+Y_3)]| \\
&\qquad + |\mathbb{E}[g(Z_1+Y_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Y_3)]| \\
&\qquad + |\mathbb{E}[g(Z_1+Z_2+Y_3)] - \mathbb{E}[g(Z_1+Z_2+Z_3)]| \qquad (3.9)\\
&\le \frac{C}{6}\left(\mathbb{E}\left[|Y_1|^3\right] + \mathbb{E}\left[|Z_1|^3\right]\right) + \frac{C}{6}\left(\mathbb{E}\left[|Y_2|^3\right] + \mathbb{E}\left[|Z_2|^3\right]\right) + \frac{C}{6}\left(\mathbb{E}\left[|Y_3|^3\right] + \mathbb{E}\left[|Z_3|^3\right]\right) \qquad (3.10)\\
&= \frac{3C}{6}\left(\mathbb{E}\left[|Y_1|^3\right] + \mathbb{E}\left[|Z_1|^3\right]\right), \qquad (3.11)
\end{aligned}$$

where $C := \sup_{t\in\mathbb{R}} |g'''(t)|$ and:

- (3.8) held since we were just subtracting some terms and then immediately adding them back.
- (3.9) followed from the triangle inequality.
- (3.10) is obtained by three uses of Lemma 3.2. To be more precise:
  1) The first term is obtained by setting $X = Y_2 + Y_3$, $Y = Y_1$, $Z = Z_1$ in Lemma 3.2.
  2) The second term is obtained by setting $X = Z_1 + Y_3$, $Y = Y_2$, $Z = Z_2$ in Lemma 3.2.
  3) The third term is obtained by setting $X = Z_1 + Z_2$, $Y = Y_3$, $Z = Z_3$ in Lemma 3.2.

  You can check that each choice of (X, Y, Z) above satisfies the requirement for Lemma 3.2.
- (3.11) followed from the fact that Y_i are iid and Z_i are iid.

We can generalize the argument to get for general n that

$$\begin{aligned}
|\mathbb{E}[g(Y_1+\cdots+Y_n)] - \mathbb{E}[g(Z_1+\cdots+Z_n)]| &\le \frac{nC}{6}\left(\mathbb{E}\left[|Y_1|^3\right] + \mathbb{E}\left[|Z_1|^3\right]\right) \\
&= O(n^{-1/2}), \qquad (3.12)
\end{aligned}$$

where we used the fact that $\mathbb{E}\left[|Y_1|^3\right] = O(n^{-3/2})$ and $\mathbb{E}\left[|Z_1|^3\right] = O(n^{-3/2})$ from (3.5) and (3.6), respectively. From (3.12) we see that (3.7) is true, which proves the convergence of distribution. □

Problem 4. Prove Markov's and Chebyshev's inequality.

Solution. Markov's Inequality: Let X be a non-negative random variable with finite mean. Then for every $a > 0$,

$$\mathbb{P}(X \ge a) \ \le \ \frac{\mathbb{E}[X]}{a}.$$

Proof. Define the indicator random variable $\mathbb{1}_{\{X\ge a\}}$. Note that we have the following:

$$X \ \ge \ a\,\mathbb{1}_{\{X\ge a\}},$$

which can be seen as follows:

- When $X \ge a$, the right-hand side equals a, which is less than or equal to X since we have assumed $X \ge a$.
- When $X < a$, the right-hand side equals 0, which is also less than or equal to X since we have assumed that X is non-negative.

Taking expectations on both sides (expectation preserves the inequality for integrable random variables):

$$\begin{aligned}
\mathbb{E}[X] &\ge a\mathbb{E}\left[\mathbb{1}_{\{X\ge a\}}\right] \\
&= a\mathbb{P}\left(X \ge a\right),
\end{aligned}$$

where we used the fact that the expectation of any indicator function $\mathbb{1}_A$ is simply the probability of the event A (i.e. $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}(A)$). Dividing both sides by $a > 0$ yields

$$\boxed{\mathbb{P}\left(X \ge a\right) \ \le \ \frac{\mathbb{E}\left[X\right]}{a}.}$$

□

Chebyshev's Inequality: Let X be a random variable with finite mean $\mu = \mathbb{E}[X]$ and finite variance $\sigma^2 = \mathrm{Var}(X)$. Then for every $k > 0$,

$$\mathbb{P}\big(|X-\mu| \geq k\big) \;\leq\; \frac{\sigma^2}{k^2}.$$

Proof. Consider the non-negative random variable $Y = (X-\mu)^2$. Note that

$$\mathbb{E}[Y] = \mathbb{E}\big[(X-\mu)^2\big] = \sigma^2.$$

Now observe that the event $\{|X-\mu| \geq k\}$ is the same as $\{(X-\mu)^2 \geq k^2\}$, i.e. $\{Y \geq k^2\}$.

Applying Markov's inequality to the non-negative random variable Y with threshold $a = k^2 > 0$:

$$\mathbb{P}\big(|X-\mu| \geq k\big) = \mathbb{P}\big(Y \geq k^2\big) \leq \frac{\mathbb{E}[Y]}{k^2} = \frac{\sigma^2}{k^2}.$$

Hence

$$\boxed{\mathbb{P}\big(|X-\mu| \geq k\big) \;\leq\; \frac{\sigma^2}{k^2}.}$$

□

Problem 5. Prof. Hajek's text [2], 2.23, 2.24, 2.29, 2.33.

Solution. The solution to [2, Problem 2.24] can be seen in [2, Chapter 12].

Solution to [2, Problem 2.23]:

Let $X_1, X_2, \ldots$ be i.i.d. Exp(1). For $n \geq 2$, let $Z_n = \dfrac{\max\{X_1, \ldots, X_n\}}{\ln n}$.

(a) Let $M_n = \max\{X_1, \ldots, X_n\}$. By independence,

$$F_{M_n}(x) = \mathbb{P}(M_n \leq x) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) = \big(1 - e^{-x}\big)^n, \quad x \geq 0.$$

Since $Z_n = M_n/\ln n$, for $z \geq 0$, we have

$$F_{Z_n}(z) = \mathbb{P}\left(\frac{M_n}{\ln n} \leq z\right) = \big(1 - e^{-z\ln n}\big)^n = \big(1 - n^{-z}\big)^n.$$

At the same time, Z_n is clearly non-negative, and thus $F_{Z_n}(z) = 0$ for $z < 0$. Altogether, the CDF of Z_n is given by

$$\boxed{F_{Z_n}(z) = \begin{cases} \big(1 - n^{-z}\big)^n, & \text{if } z \geq 0, \\ 0, & \text{otherwise.} \end{cases}}$$

(b) We first compute $\lim_{n\to\infty} F_{Z_n}(z)$ for $z > 0$. Taking the logarithm,

$$\ln F_{Z_n}(z) = n \ln\big(1 - n^{-z}\big).$$

Then note that we have

$$\begin{aligned}
\lim_{n\to\infty} \ln F_{Z_n}(z) &= \lim_{n\to\infty} n \ln\big(1 - n^{-z}\big) \\
&= \lim_{u\to 0^+} \frac{\ln\big(1 - u^z\big)}{u} &\quad (5.1)\\
&= \lim_{u\to 0^+} \frac{-zu^{z-1}}{1 - u^z} &\quad (5.2)\\
&= \begin{cases} 0, & \text{if } z > 1, \\ -1, & \text{if } z = 1, \\ -\infty, & \text{if } 0 < z < 1. \end{cases},
\end{aligned}$$

where in (5.1) we used the substitution $u = \frac{1}{n}$ and (5.2) followed from L'Hôpital's rule. Therefore, for $z > 0$ we have

$$\lim_{n\to\infty} F_{Z_n}(z) = \begin{cases} 0, & \text{if } 0 < z < 1, \\ \frac{1}{e}, & \text{if } z = 1 \\ 1, & \text{if } z > 1. \end{cases}$$

At the same time, for $z \leq 0$, we clearly have $\lim_{n\to\infty} F_{Z_n}(z) = \lim_{n\to\infty} 0 = 0$. In summary, we have for all $z \in \mathbb{R}$ that

$$\lim_{n\to\infty} F_{Z_n}(z) = \begin{cases} 0, & \text{if } z < 1, \\ \frac{1}{e}, & \text{if } z = 1 \\ 1, & \text{if } z > 1, \end{cases}$$

which is the CDF of the constant 1 at all continuity points ($z = 1$ is the sole discontinuity). By definition,

$$\boxed{Z_n \xrightarrow{d} 1.}$$

Solution to [2, Problem 2.29]:

Let $X_1, X_2, \ldots$ be mutually independent with $\mathbb{E}[X_i] = 0$ for all i. Let $S_n = X_1 + \cdots + X_n$. Show that S_n converges in the mean square sense if and only if $\sum_{i=1}^{\infty} \operatorname{Var}(X_i) < \infty$.

Proof. We apply the *Cauchy criterion for mean square convergence*: a sequence (S_n) converges in mean square if and only if it is Cauchy in L^2, i.e.,

$$\mathbb{E}\big[(S_m - S_n)^2\big] \longrightarrow 0 \quad \text{as } m, n \to \infty.$$

For $m > n$, we have

$$S_m - S_n = X_{n+1} + X_{n+2} + \cdots + X_m.$$

Since the X_i are mutually independent with mean zero, the cross terms vanish:

$$\mathbb{E}\big[(S_m - S_n)^2\big] = \mathbb{E}\left[\left(\sum_{i=n+1}^{m} X_i\right)^2\right] = \sum_{i=n+1}^{m} \mathbb{E}[X_i^2] + \sum_{\substack{i\neq j \\ n<i,j\leq m}} \underbrace{\mathbb{E}[X_i]\mathbb{E}[X_j]}_{=0} = \sum_{i=n+1}^{m} \operatorname{Var}(X_i).$$

Suppose $\sum_{i=1}^{\infty} \operatorname{Var}(X_i) < \infty$. Then the partial sums $V_n = \sum_{i=1}^{n} \operatorname{Var}(X_i)$ form a convergent series of non-negative reals, so for every $\varepsilon > 0$ there exists N such that for all $m > n \geq N$,

$$\mathbb{E}\big[(S_m - S_n)^2\big] = \sum_{i=n+1}^{m} \operatorname{Var}(X_i) = V_m - V_n < \varepsilon.$$

Hence (S_n) is Cauchy in L^2, and thus by the Cauchy criterion for mean square convergence, S_n converges in mean square.

Suppose S_n converges in mean square to some limit S. Then (S_n) is Cauchy in L^2, so

$$\sum_{i=n+1}^{m} \operatorname{Var}(X_i) = \mathbb{E}\big[(S_m - S_n)^2\big] \longrightarrow 0 \quad \text{as } m, n \to \infty.$$

This means the partial sums $V_n = \sum_{i=1}^{n} \operatorname{Var}(X_i)$ form a Cauchy sequence in $\mathbb{R}$, hence they converge, i.e., $\sum_{i=1}^{\infty} \operatorname{Var}(X_i) < \infty$.

Combining both directions,

$$\boxed{S_n = X_1 + \cdots + X_n \text{ converges in mean square} \iff \sum_{i=1}^{\infty} \operatorname{Var}(X_i) < \infty.}$$

■

Solution to [2, Problem 2.33]:

Let $X_1, X_2, \dots$ be i.i.d. standard Cauchy with $\Phi(u) = e^{-|u|}$. Let $S_n = X_1 + \cdots + X_n$.

(a) Find the characteristic function of S_n/n^θ for a constant θ.

By independence,

$$\varphi_{S_n}(u) = \big[\Phi(u)\big]^n = e^{-n|u|}.$$

Using the scaling property $\varphi_{aX}(u) = \varphi_X(au)$,

$$\boxed{\varphi_{S_n/n^\theta}(u) = \varphi_{S_n}\left(\frac{u}{n^\theta}\right) = \exp\Big(-n\left|\frac{u}{n^\theta}\right|\Big) = \exp\big(-n^{1-\theta}\,|u|\big).}$$

(b) Does S_n/n converge in distribution?

Setting $\theta = 1$,

$$\varphi_{S_n/n}(u) = e^{-n^{1-1}|u|} = e^{-|u|} = \Phi(u) \quad \text{for all } n.$$

This is the characteristic function of a standard Cauchy for *every* n. In particular,

$$\lim_{n\to\infty} \varphi_{S_n/n}(u) = e^{-|u|},$$

which is continuous at $u = 0$. Since pointwise convergence of characteristic functions implies convergence in distribution (see [2, Proposition 2.6]), we have

$$\boxed{\frac{S_n}{n} \xrightarrow{d} \text{Cauchy}(0,1).}$$

(In fact $S_n/n \sim \text{Cauchy}(0,1)$ for every n, so convergence is trivial.)

(c) Does S_n/n^2 converge in distribution?

Setting $\theta = 2$,

$$\varphi_{S_n/n^2}(u) = e^{-n^{-1}|u|} \;\longrightarrow\; e^0 = 1 \quad \text{as } n \to \infty.$$

The limit $\varphi(u) \equiv 1$ is the characteristic function of the constant 0 (i.e., a point mass at 0), and it is continuous at $u = 0$. Similar to Part (b), we have

$$\boxed{\frac{S_n}{n^2} \xrightarrow{d} 0.}$$

(d) Does $S_n/\sqrt{n}$ converge in distribution?

Setting $\theta = 1/2$,

$$\varphi_{S_n/\sqrt{n}}(u) = e^{-\sqrt{n}\,|u|}.$$

For $u = 0$ this equals 1 for all n. For any $u \neq 0$,

$$e^{-\sqrt{n}\,|u|} \;\longrightarrow\; 0 \quad \text{as } n \to \infty.$$

Hence the pointwise limit is

$$\lim_{n\to\infty} \varphi_{S_n/\sqrt{n}}(u) = \begin{cases} 1, & u = 0, \\ 0, & u \neq 0, \end{cases}$$

which is *not continuous at* $u = 0$. Therefore, the pointwise limit of $\varphi_{S_n/\sqrt{n}}(u)$ is not a valid characteristic function. (You can take for granted the property that the characteristic function of any random variable is necessarily continuous.) We can then conclude that

$$\boxed{\frac{S_n}{\sqrt{n}} \text{ does \textit{not} converge in distribution.}}$$

Another approach. Here is another approach that does not use the characteristic function. We use CDF instead. From Part (b) we know that S_n/n and X_1 have the same distribution. We then have for each $c \in \mathbb{R}$ that

$$\begin{aligned}\mathbb{P}\left(\frac{S_n}{\sqrt{n}} \le c\right) &= \mathbb{P}\left(\frac{S_n}{n} \le \frac{c}{\sqrt{n}}\right) \\ &= \mathbb{P}\left(X_1 \le \frac{c}{\sqrt{n}}\right) \\ &= F_{X_1}\left(\frac{c}{\sqrt{n}}\right),\end{aligned}$$

where F_{X_1} is the CDF of X_1. Since the standard Cauchy distribution is a continuous distribution (we define it by describing its density), its CDF is a continuous function. Therefore, we can pass the limit directly inside F_{X_1}, which gives

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \le c\right) &= \lim_{n\to\infty} F_{X_1}\left(\frac{c}{\sqrt{n}}\right) \\ &= F_{X_1}\left(\lim_{n\to\infty} \frac{c}{\sqrt{n}}\right) \\ &= F_{X_1}(0) \\ &= \frac{1}{2}. \end{aligned} \tag{5.3}$$

Now assume for contradiction that $S_n/\sqrt{n}$ converges in distribution. Let Y be a random variable having the limiting distribution (i.e. $S_n/\sqrt{n} \xrightarrow[n\to\infty]{d} Y$). By (5.3) and the definition of convergence in distribution, the CDF of Y satisfies $F_Y(c) = \frac{1}{2}$ for all continuity points c of F_Y. Since there can be at most countably many discontinuity points of any CDF (a good exercise for you), we have $F_Y(c) = \frac{1}{2}$ except for at most countably many c. It then follows from the right continuity of F_Y that $F_Y(c) = \frac{1}{2}$ for all $c \in \mathbb{R}$. But the constant function $F_Y(c) = \frac{1}{2}$ is clearly a not valid CDF (for example, it does not go to 1 as c goes to infinity), and thus we have reached a contradiction.

This contradiction stems from the assumption that $S_n/\sqrt{n}$ converges in distribution, and therefore $S_n/\sqrt{n}$ does not converge in distribution.

Remark. Parts (b)–(d) illustrate a key feature of the Cauchy distribution: it has no finite mean, so the CLT (which would require $\sqrt{n}$ scaling) does not apply. Instead, the "right" normalization is n^1, not $n^{1/2}$, reflecting the heavy tails of the Cauchy distribution.

References

[1] O. Lévêque, "Advanced probability and applications (Part I)," July 2018. [Online]. Available: https://ipgold.epfl.ch/~leveque/Lecture_Notes/APA_lecture_notes_part1.pdf

[2] B. Hajek, *Random Processes for Engineers*. Cambridge university press, 2015. [Online]. Available: https://hajek.ece.illinois.edu/Papers/randomprocJuly14.pdf