Partial exam – Advanced probability

2026-10-20

Duration: 3 hours. No documents are permitted, and the use of any electronic device (mobile phone, etc.) is also forbidden. All answers must be carefully justified.

Exercise 1. Almost sure convergence vs. convergence in probability. Let $(X_n)_{n\in\mathbb{N}}$ and X be random variables defined on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and with values in some metric space $(\mathfrak{X}, d)$ (if needed, one can assume that $\mathfrak{X} = \mathbb{R}$).

(1) We assume that the sequence $(X_n)_{n\in\mathbb{N}}$ converges in probability to X: for any $\varepsilon > 0$, $\mathbb{P}[d(X_n, X) \geq \varepsilon] \to 0$. By using the Borel–Cantelli lemma, show that there exists a subsequence $(X_{n_k})_{k\in\mathbb{N}}$ that converges almost surely to X.

(2) Show that the following are equivalent:

   (i.) The sequence $(X_n)_{n\geq 0}$ converges in probability to X.

   (ii.) For any subsequence $(X_{n_k})_{k\geq 0}$, there exists a further subsequence $(X_{n_{k_l}})_{l\geq 0}$ which converges almost surely to X.

(3) Deduce from this criterion the following: if we have convergence in probability $X_n \to_{\mathbb{P}} X$ and if $f : (\mathfrak{X}, d) \to (\mathfrak{Y}, \delta)$ is a continuous map, then $f(X_n) \to_{\mathbb{P}} f(X)$.

Exercise 2. Conditional law of a Gaussian coordinate knowing the radius. In this exercise, $\mathbb{S}^{d-1}$ denotes the $(d-1)$-dimensional unit sphere in $\mathbb{R}^d$, with $d \geq 2$ fixed:

$$\mathbb{S}^{d-1} = \left\{(x_1, \dots, x_d) \in \mathbb{R}^d \,|\, (x_1)^2 + \cdots + (x_d)^2 = 1\right\}.$$

One is allowed to use without further justification the spherical change of variables formula: for any measurable function $f : \mathbb{R}^d \to \mathbb{R}$ which is integrable with respect to the Lebesgue measure $\mathrm{d}x = \mathrm{d}x_1 \cdots \mathrm{d}x_d$,

$$\int_{\mathbb{R}^d} f(x)\, \mathrm{d}x = \int_{\mathbb{R}_+ \times \mathbb{S}^{d-1}} f(rs)\, r^{d-1}\, \mathrm{d}r\, \mu_{\mathbb{S}^{d-1}}(\mathrm{d}s),$$

where $\mu_{\mathbb{S}^{d-1}}$ is the surface measure of the $(d-1)$-dimensional sphere (unique Borel measure up to multiplication by a scalar on $\mathbb{S}^{d-1}$ which is invariant by the rotation action of $\mathrm{SO}(d)$). The total mass of $\mu_{\mathbb{S}^{d-1}}$ will be computed during the exercise (it is determined by the identity above). To this purpose, we set for any $s > 0$

$$\Gamma(s) = \int_0^\infty \mathrm{e}^{-x}\, x^{s-1}\, \mathrm{d}x\,.$$

(1) We consider a vector $X = (X_1, \dots, X_d)$ of independent random variables, each X_i following a standard normal distribution $\mathscr{N}(0,1)$:

$$\mathbb{P}[X_i \in A] = \frac{1}{\sqrt{2\pi}} \int_A \mathrm{e}^{-\frac{x^2}{2}}\, \mathrm{d}x\,.$$

Compute the density with respect to the Lebesgue measure on $\mathbb{R}_+$ of the law of $R = \sqrt{(X_1)^2 + \cdots + (X_d)^2}$ (the ℓ^2-norm of the vector X), and at the same time the total mass of $\mu_{\mathbb{S}^{d-1}}$.

(2) Compute the joint density on $\mathbb{R} \times \mathbb{R}_+$ of the law of the vector (X_1, R).

(3) Compute the conditional law of the coordinate X_1 knowing the radius R.

(4) What is the conditional law of X_1 knowing R when $d = 3$?

Exercise 3. The two-series theorem. Let $(X_n)_{n\geq 1}$ be a sequence of real random variables which are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, independent and square integrable:

$$\forall n \geq 1, \;\; X_n \in \mathscr{L}^2(\Omega, \mathscr{F}, \mathbb{P}).$$

We set $m_n = \mathbb{E}[X_n]$, $(\sigma_n)^2 = \mathrm{var}(X_n)$ and $S_n = \sum_{k=1}^n X_k$. The objective of the exercise is to establish the following implication:

$$\left(\sum_{n=1}^{\infty} m_n \text{ converges in } \mathbb{R}, \text{ and } \sum_{n=1}^{\infty} (\sigma_n)^2 < +\infty\right)$$
$$\Rightarrow ((S_n)_{n\in\mathbb{N}} \text{ converges almost surely to some real-valued variable } S)\,.$$

(1) Explain why one can assume without loss of generality that $m_n = 0$ for any $n \geq 1$. Thus, in the sequel, the variables X_k are supposed centered.

(2) By considering $S_N - S_n$, show that for any $n \geq 1$,

$$\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N \leq 2 \sup_{m\geq 0} \left| \sum_{k=n+1}^{n+m} X_k \right|$$

almost surely.

Let $(\xi_k)_{k\geq 1}$ be a sequence of independent and square integrable random variables with $\mathbb{E}[\xi_k] = 0$ for any k. We denote $\Sigma_n = \sum_{k=1}^n \xi_k$ and we fix $\varepsilon > 0$. Set

$$\tau = \inf\{n \geq 1 \text{ such that } |\Sigma_n| > \varepsilon\}\,.$$

By convention, $\tau = +\infty$ if all the sums Σ_n stay in $[-\varepsilon, \varepsilon]$.

(3) Show that for $k \leq n$, the two random variables $1_{(\tau=k)}\,\Sigma_k$ and $\Sigma_n - \Sigma_k$ are independent.

(4) Using $(\Sigma_n)^2 \geq 1_{(\tau\leq n)}(\Sigma_n)^2$, prove that:

$$\sum_{k=1}^{n} \mathbb{E}[(\xi_k)^2] \geq \varepsilon^2\, \mathbb{P}[\tau \leq n].$$

(5) Going back to the initial variables X_k, show that for any $n \geq 1$,

$$\mathbb{P}\left[\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N > \varepsilon\right] \leq \frac{4}{\varepsilon^2} \sum_{k=n+1}^{\infty} (\sigma_k)^2.$$

(6) Conclude the proof of the implication $\Rightarrow$.

Exercise 4. The first cycles of a random permutation. We recall that a *permutation* with size n is a bijection $\sigma : [\![1, n]\!] \to [\![1, n]\!]$. The set of permutations with size n is $\mathfrak{S}(n) = \mathfrak{S}([\![1, n]\!])$ and has cardinality $n!$. A permutation σ can be described by the sequence $\sigma(1), \sigma(2), \ldots, \sigma(n)$ of its values: for instance, $\sigma = 813542697$ is the element of $\mathfrak{S}(9)$ which sends 1 to 8, 2 to 1, 3 to 3, 4 to 5, etc. Any permutation can be written uniquely as a product of disjoint cycles; hence,

$$\sigma = 813542697 = (1, 8, 9, 7, 6, 2) \circ (3) \circ (4, 5)$$

is a permutation with size $n = 9$ and with 3 disjoint cycles (including the fixed point 3, which is a cycle with size 1). In the decomposition in disjoint cycles $c_1(\sigma) \circ c_2(\sigma) \circ \cdots \circ c_\ell(\sigma)$, one has unicity up to permutation of these cycles, but one can impose a canonical order as follows:

- $c_1(\sigma)$ is the cycle containing 1;
- $c_2(\sigma)$ is the cycle containing the smallest element of $[\![1, n]\!] \setminus c_1(\sigma)$;
- $c_3(\sigma)$ is the cycle containing the smallest element of $[\![1, n]\!] \setminus (c_1(\sigma) \sqcup c_2(\sigma))$;

etc. We then denote $s_k(\sigma) = |c_k(\sigma)|$ the size of the k-th cycle of σ. We convene that if $k > \ell$ with ℓ the number of disjoint cycles of σ, then $c_k(\sigma) = \mathrm{id}_{[\![1,n]\!]}$ and $s_k(\sigma) = 0$. In the previous example, $s_1(\sigma) = 6$, $s_2(\sigma) = 1$, $s_3(\sigma) = 2$, and $s_{k\geq 4}(\sigma) = 0$. With $k \geq 1$ fixed, the objective of the exercise is to describe the limit in law of the vector

$$\left(\frac{s_1(\sigma_n)}{n}, \frac{s_2(\sigma_n)}{n}, \dots, \frac{s_k(\sigma_n)}{n}\right),$$

where σ_n is a uniform random permutation in $\mathfrak{S}(n)$: $\mathbb{P}[\sigma_n = \sigma] = \frac{1}{n!}$ for any $\sigma \in \mathfrak{S}(n)$.

(1) Given $l \geq 1$, we decompose $\sigma = \sigma^{(l)} \circ \sigma_{(l)}$, where $\sigma^{(l)} = c_1(\sigma) \circ c_2(\sigma) \cdots \circ c_l(\sigma)$ is the product of the l first cycles of σ, and $\sigma_{(l)} = c_{l+1}(\sigma) \circ c_{l+2}(\sigma) \circ \cdots$ is the product of the remaining cycles. We also denote $A_l(\sigma) = c_1(\sigma) \sqcup \cdots \sqcup c_l(\sigma)$ the support of the l first cycles. Suppose that $\sigma_n \sim \mathrm{Unif}(\mathfrak{S}(n))$; then, $A_l(\sigma_n)$ is a random subset in $\mathfrak{P}([\![1,n]\!])$, and the pair $(A_l(\sigma_n), (\sigma_n)^{(l)})$ is a random element of

$$\mathfrak{X} = \bigsqcup_{A \subset [\![1,n]\!]} \{A\} \times \mathfrak{S}(A).$$

Show that the law of $(\sigma_n)_{(l)}$ conditionally to the random pair $(A_l(\sigma_n), (\sigma_n)^{(l)})$ is the uniform distribution on

$$\mathfrak{S}([\![1,n]\!] \setminus A_l(\sigma_n)).$$

(2) Given $k \in [\![1,n]\!]$, how many permutations in $\mathfrak{S}(n)$ satisfy $s_1(\sigma) = k$?

(3) Give the law of the random variable $s_1(\sigma_n)$, and show that one has the convergence in distribution

$$\frac{s_1(\sigma_n)}{n} \rightharpoonup U_1,$$

where the right-hand side is a random variable which follows a uniform distribution on the interval $[0,1]$.

(4) We would like to extend the result above to the random vectors $(\frac{s_1(\sigma_n)}{n}, \dots, \frac{s_k(\sigma_n)}{n})$, k being an arbitrary but fixed integer. We recall that the polynomials in $\mathbb{R}[x_1, \dots, x_k]$ form a dense subalgebra of the algebra $\mathscr{C}([0,1]^k)$ of continuous functions from $[0,1]^k$ to $\mathbb{R}$. Show that if $X_n = (X_{n,1}, \dots, X_{n,k})$ and $X = (X_1, \dots, X_k)$ are random vectors with values in $[0,1]^k$ almost surely, then the following assertions are equivalent:

- The sequence $(X_n)_{n\in\mathbb{N}}$ converges in law to X: $X_n \rightharpoonup X$.
- For any exponents $m_1, \dots, m_k$ in $\mathbb{N}$,

$$\mathbb{E}[(X_{n,1})^{m_1} \cdots (X_{n,k})^{m_k}] \to_{n\to\infty} \mathbb{E}[(X_1)^{m_1} \cdots (X_k)^{m_k}].$$

(5) Let $(U_k)_{k\geq 1}$ be a sequence of independent uniform random variables in $[0,1]$, and $(S_k)_{k\geq 1}$ be the corresponding *stick-breaking process*:

$$S_1 = U_1 \qquad ; \qquad S_{k\geq 2} = U_k(1 - S_1 - \cdots - S_{k-1}).$$

By using an induction and conditional laws, show that for any $k \geq 1$, we have the convergence in distribution

$$\left(\frac{s_1(\sigma_n)}{n}, \frac{s_2(\sigma_n)}{n}, \dots, \frac{s_k(\sigma_n)}{n}\right) \rightharpoonup (S_1, \dots, S_k).$$

Exercise 1.

(1) For any $k \geq 0$, there exists n_k such that

$$\mathbb{P}\left[d(X_{n_k}, X) \geq \frac{1}{2^k}\right] \leq \frac{1}{2^k}.$$

We can assume without loss of generality that the sequence of indices $(n_k)_{k\geq 0}$ is constructed to be increasing. Then, by Borel–Cantelli, the event

$$\text{there exists an infinity of indices } k \text{ such that } d(X_{n_k}, X) \geq \frac{1}{2^k}$$

has probability 0, since $\sum_{k\geq 0} \frac{1}{2^k} < +\infty$. So, with probability 1, for k large enough, $d(X_{n_k}, X) < \frac{1}{2^k}$, which proves that $X_{n_k} \to_{\text{a.s.}} X$.

(2) If $(X_n)_{n\geq 0}$ converges in probability to X, then obviously any subsequence $(X_{n_k})_{k\geq 0}$ also converges in probability to X. By the previous question, any subsequence admits a further subsequence $(X_{n_{k_l}})_{l\in\mathbb{N}}$ which converges almost surely to X. Conversely, if X_n does not converge in probability to X, then there is a subsequence $(X_{n_k})_{k\geq 0}$ and $\varepsilon, \eta > 0$ such that:

$$\mathbb{P}[d(X_{n_k}, X) \geq \varepsilon] \geq \eta$$

for any k. Then, any further subsequence has the same property, which is incompatible with the almost sure convergence towards X.

(3) If $X_n \to_{\mathbb{P}} X$, then subsequences of X_n admit further subsequences that converge almost surely to X, so subsequences of $f(X_n)$ admit further subsequences that converge almost surely to $f(X)$: indeed, continuous functions preserve the almost sure convergence. Therefore, $f(X_n) \to_{\mathbb{P}} f(X)$.

Exercise 2.

(1) By using the spherical change of variables, we get for any observable $F : \mathbb{R}_+ \to \mathbb{R}$ (bounded measurable):

$$\begin{aligned}\mathbb{E}[F(R)] &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} F\left(\sqrt{(x_1)^2 + \cdots + (x_d)^2}\right) \mathrm{e}^{-\frac{(x_1)^2+\cdots+(x_d)^2}{2}}\, \mathrm{d}x \\ &= \frac{\mathrm{vol}(\mathbb{S}^{d-1})}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}_+} F(r)\, \mathrm{e}^{-\frac{r^2}{2}}\, r^{d-1}\, \mathrm{d}r\,,\end{aligned}$$

where $\mathrm{vol}(\mathbb{S}^{d-1}) = \mu_{\mathbb{S}^{d-1}}(\mathbb{S}^{d-1})$. In particular, if $F = 1$, then one gets:

$$1 = \frac{\mathrm{vol}(\mathbb{S}^{d-1})}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}_+} \mathrm{e}^{-\frac{r^2}{2}}\, r^{d-1}\, \mathrm{d}r = \frac{\mathrm{vol}(\mathbb{S}^{d-1})}{2\,(\pi)^{\frac{d}{2}}} \int_{\mathbb{R}_+} \mathrm{e}^{-u}\, u^{\frac{d}{2}-1}\, \mathrm{d}u = \frac{\mathrm{vol}(\mathbb{S}^{d-1})}{2\,(\pi)^{\frac{d}{2}}}\, \Gamma\left(\frac{d}{2}\right).$$

Therefore, $\mathrm{vol}(\mathbb{S}^{d-1}) = \frac{2\,(\pi)^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$ (which gives indeed 2π and 4π in dimensions $d = 2$ and $d = 3$), and R has density

$$f_R(r) = 1_{(r\geq 0)}\, \frac{2^{1-\frac{d}{2}}}{\Gamma(\frac{d}{2})}\, \mathrm{e}^{-\frac{r^2}{2}}\, r^{d-1}.$$

(2) Setting $S = \sqrt{(X_2)^2 + \cdots + (X_d)^2} = \sqrt{R^2 - (X_1)^2}$, we compute similarly:

$$\begin{aligned}\mathbb{E}[F(X_1, R)] &= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} F\Big(x_1, \sqrt{(x_1)^2 + \cdots + (x_d)^2}\Big)\, \mathrm{e}^{-\frac{(x_1)^2+\cdots+(x_d)^2}{2}}\, \mathrm{d}x \\ &= \frac{\mathrm{vol}(\mathbb{S}^{d-2})}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}\times\mathbb{R}_+} F\Big(x_1, \sqrt{(x_1)^2 + s^2}\Big)\, \mathrm{e}^{-\frac{x_1^2+s^2}{2}}\, s^{d-2}\, \mathrm{d}x_1\, \mathrm{d}s \\ &= \frac{\mathrm{vol}(\mathbb{S}^{d-2})}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}\times\mathbb{R}_+} F(x_1, r)\, 1_{(|x_1|\le r)}\, \mathrm{e}^{-\frac{r^2}{2}}\, (r^2 - (x_1)^2)^{\frac{d-3}{2}}\, \mathrm{d}x_1\, r\, \mathrm{d}r\end{aligned}$$

by making the change of variables $r = \sqrt{(x_1)^2 + s^2}$. So, the joint density of the pair (X_1, R) is:

$$f_{(X_1,R)}(x_1, r) = 1_{(0\le|x_1|\le r)}\, 2^{1-\frac{d}{2}}\, \frac{(r^2 - (x_1)^2)^{\frac{d-3}{2}}\, r\, \mathrm{e}^{-\frac{r^2}{2}}}{\sqrt{\pi}\, \Gamma(\frac{d-1}{2})}.$$

(3) Dividing the results of the two previous questions yields the conditional density of X_1 knowing $R = r$:

$$f_{X_1\,|\,R=r}(x_1) = 1_{(0\le|x_1|\le r)}\, \frac{\Gamma(\frac{d}{2})}{\sqrt{\pi}\, \Gamma(\frac{d-1}{2})\, r}\, \left(1 - \frac{(x_1)^2}{r^2}\right)^{\frac{d-3}{2}}.$$

(4) If $d = 3$, then we get a constant in the conditional density:

$$f_{X_1\,|\,R=r}(x_1) = 1_{(0\le|x_1|\le r)}\, \frac{\Gamma(\frac{3}{2})}{\sqrt{\pi}\, r}$$

since $\Gamma(1) = 1$. Integrating over the possible values for x_1 yields:

$$1 = \frac{2\,\Gamma(\frac{3}{2})}{\sqrt{\pi}}.$$

So, we have in fact:

$$f_{X_1\,|\,R=r}(x_1) = \frac{1_{(-r\le x_1\le r)}}{2r}.$$

Conditionally to R, X_1 follows a uniform distribution on $[-R, R]$ in dimension 3.

Exercise 3.

(1) Suppose that the result is true for independent and centered random variables $\overline{X_n}$: if $\sum_{n=1}^{\infty} \mathrm{var}(\overline{X_n}) < +\infty$, then the sequence $(\overline{S_n})_{n\ge1}$ defined by $\overline{S_n} = \sum_{k=1}^{n} \overline{X_k}$ admits an almost sure limit. Consider then independent variables X_n with $m_n = \mathbb{E}[X_n]$ and $(\sigma_n)^2 = \mathrm{var}(X_n)$. We set $M_n = \sum_{k=1}^{n} m_k$ and we assume that $(M_n)_{n\ge1}$ converges in $\mathbb{R}$ and that $\sum_{n=1}^{\infty} (\sigma_n)^2 < +\infty$. Then, if $\overline{X_n} = X_n - m_n$, by the centered case, since $\mathrm{var}(\overline{X}_n) = \mathrm{var}(X_n)$,

$$\overline{S_n} = S_n - M_n$$

converges almost surely. By hypothesis, M_n also converges almost surely (to some constant), so by addition, S_n converges almost surely. Hence, it suffices to treat the case of centered random variables.

(2) We have

$$\begin{aligned}\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N &= \limsup_{N\to\infty}(S_N - S_n) - \liminf_{N\to\infty}(S_N - S_n)\\ &= \inf_{l\geq 1}\sup_{m\geq l}\sum_{k=n+1}^{n+m} X_k - \sup_{l\geq 1}\inf_{m\geq l}\sum_{k=n+1}^{n+m} X_k\\ &\leq \sup_{m\geq 1}\sum_{k=n+1}^{n+m} X_k + \sup_{m\geq 1}\sum_{k=n+1}^{n+m} -X_k \leq 2\sup_{m\geq 1}\left|\sum_{k=n+1}^{n+m} X_k\right|.\end{aligned}$$

(3) Notice that both variables $1_{(\tau=k)}$ and Σ_k are measurable with respect to the σ-algebra $\mathscr{F}_k = \sigma(\xi_1, \dots, \xi_k)$. Indeed, this is trivial for $\Sigma_k = \xi_1 + \cdots + \xi_k$, and for $1_{(\tau=k)}$, this is the indicator of the event:

$$\{|\xi_1| \leq \varepsilon\} \cap \{|\xi_1 + \xi_2| \leq \varepsilon\} \cap \cdots \cap \{|\xi_1 + \cdots + \xi_{k-1}| \leq \varepsilon\} \cap \{|\xi_1 + \cdots + \xi_k| > \varepsilon\},$$

and each event of the intersection belongs clearly to $\mathscr{F}_k$. So, $1_{(\tau=k)}\,\Sigma_k$ is $\sigma(\xi_1, \dots, \xi_k)$-measurable, whereas $\Sigma_n - \Sigma_k = \xi_{k+1} + \cdots + \xi_n$ is $\sigma(\xi_{k+1}, \dots, \xi_n)$-measurable. These two σ-algebras are independent by assumption on the variables ξ_n.

(4) Since the ξ_k are independent and centered,

$$\sum_{k=1}^{n} \mathbb{E}[(\xi_k)^2] = \operatorname{var}(\Sigma_n) = \mathbb{E}[(\Sigma_n)^2].$$

Then, we have:

$$\begin{aligned}\mathbb{E}[(\Sigma_n)^2] &\geq \mathbb{E}[1_{(\tau\leq n)}(\Sigma_n)^2] = \sum_{k=1}^{n} \mathbb{E}[1_{(\tau=k)}\,(\Sigma_k + \Sigma_n - \Sigma_k)^2]\\ &\geq \sum_{k=1}^{n} \mathbb{E}[1_{(\tau=k)}\,(\Sigma_k)^2] + 2\,\mathbb{E}[1_{(\tau=k)}\Sigma_k\,(\Sigma_n - \Sigma_k)] + \mathbb{E}[1_{\tau=k}\,(\Sigma_n - \Sigma_k)^2].\end{aligned}$$

On the second line, the first term is larger than $\varepsilon^2\,\mathbb{P}[\tau = k]$ by definition of τ; the second term vanishes because of the previous question, and because $\mathbb{E}[\Sigma_n - \Sigma_k] = \sum_{j=k+1}^{n} \mathbb{E}[\xi_j] = 0$; and the last term is positive. So,

$$\sum_{k=1}^{n} \mathbb{E}[(\xi_k)^2] \geq \varepsilon^2 \sum_{k=1}^{n} \mathbb{P}[\tau = k] = \varepsilon^2\,\mathbb{P}[\tau \leq n].$$

(5) Setting $\xi_k = X_{n+k}$ and combining questions (2) and (4), we get:

$$\begin{aligned}\mathbb{P}\left[\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N > \varepsilon\right] &\leq \mathbb{P}\left[\sup_{m\geq 1}|\Sigma_m| \geq \frac{\varepsilon}{2}\right]\\ &\leq \lim_{n\to\infty} \mathbb{P}\left[\sup_{1\leq m\leq n}|\Sigma_m| \geq \frac{\varepsilon}{2}\right] \leq \frac{4}{\varepsilon^2}\sum_{k=1}^{\infty}(\sigma_{n+k})^2.\end{aligned}$$

(6) Since the series of variances converges, making n go to infinity makes the right-hand side go to 0. Therefore, $\mathbb{P}[\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N > \varepsilon] = 0$ for any $\varepsilon > 0$. Taking the union of these events with parameters $\varepsilon = \frac{1}{l}$, $l \geq 1$, we get:

$$\mathbb{P}\left[\limsup_{N\to\infty} S_N - \liminf_{N\to\infty} S_N > 0\right] = 0.$$

Therefore, $\limsup_{N\to\infty} S_N = \liminf_{N\to\infty} S_N$ almost surely, so $(S_n)_{n\geq 1}$ converges almost surely in $\overline{\mathbb{R}} = \mathbb{R} \sqcup \{\pm\infty\}$. On the other hand, $(S_n)_{n\geq 1}$ is bounded in $\mathscr{L}^2(\Omega, \mathscr{F}, \mathbb{P})$,

because

$$\sup_{n\geq 1}\mathbb{E}[(S_n)^2]=\sum_{n=1}^{\infty}(\sigma_n)^2=K<+\infty.$$

Therefore, by the Bienaymé–Chebyshev inequality,

$$\mathbb{P}[|S_n|>M]\leq\frac{K}{M^2}\quad\text{for any } n\geq 1.$$

As $1_{(|S|>M)}=\lim_{n\to\infty}1_{(|S_n|>M)}$ almost surely, by dominated convergence,

$$\mathbb{P}[|S|>M]\leq\frac{K}{M^2}.$$

In particular, $\mathbb{P}[S\in\{\pm\infty\}]=0$, so the almost sure limit S is indeed a $\mathbb{R}$-valued random variable.

Exercise 4.

(1) We have to prove that for any triplet (A,ρ,τ) with $A\subset[\![1,n]\!]$, $\rho\in\mathfrak{S}(A)$ and $\tau\in\mathfrak{S}([\![1,n]\!]\setminus A)$,

$$\mathbb{P}[(\sigma_n)_{(l)}=\tau\,|\,(A_l(\sigma_n),(\sigma_n)^{(l)})=(A,\rho)]=\frac{1}{(n-|A|)!}.$$

If $A=[\![1,n]\!]$ and $\rho\in\mathfrak{S}(n)$, then the conditioning event imposes that $\sigma_n=(\sigma_n)^{(l)}$ and that $(\sigma_n)_{(l)}=\mathrm{id}$, the unique permutation in $\mathfrak{S}([\![1,n]\!]\setminus A)=\mathfrak{S}(\emptyset)$. Thus, the identity is clear in this case. Suppose now that $A\subsetneq[\![1,n]\!]$. It suffices to prove that

$$F(A,\rho,\tau)=\mathbb{P}[A_l(\sigma_n)=A\text{ and }(\sigma_n)^{(l)}=\rho\text{ and }(\sigma_n)_{(l)}=\tau]$$

only depends on (A,ρ), and not on τ. However, the function $F(A,\rho,\tau)$ can take two possible values:

- Set $\sigma=\rho\circ\tau$, and suppose that ρ has l disjoint cycles on A, and that any cycle c_i of ρ has its minimal element $\min c_i$ strictly smaller than $\min([\![1,n]\!]\setminus A)$. Then, ordering the cycles $c_1,\dots,c_l$ of ρ and the cycles $c_{l+1},\dots,c_\ell$ of τ according to their minimal elements, we have:

$$\rho=c_1\circ c_2\circ\cdots\circ c_l;$$
$$\sigma=c_1\circ c_2\circ\cdots\circ c_l\circ c_{l+1}\circ\cdots\circ c_\ell$$

  and this is the canonical decomposition of these two permutations. Therefore, $F(A,\rho,\tau)=\mathbb{P}[\sigma_n=\sigma]=\frac{1}{n!}$.
- Otherwise, ρ cannot be the product of the l first cycles of σ, so $F(A,\rho,\tau)=0$.

Thus,

$$F(A,\rho,\tau)=\frac{1}{n!}\,1_{(\rho\text{ has }l\text{ disjoint cycles, whose minimal elements are smaller than }\min A^c)}.$$

As wanted, $F(A,\rho,\tau)$ only depends on (A,ρ), and not on $\tau\in\mathfrak{S}(A^c)$.

(2) Fix $k\geq 1$. There are

$$(n-1)(n-2)\cdots(n-k+1)=\frac{(n-1)!}{(n-k)!}$$

cycles with length k of the form $c_1=(1,a_2,\dots,a_k)$: $n-1$ choices for a_2, $n-2$ choices for a_3, etc. Then, for any such cycle c_1, we have indeed $\min c_1=1<\min([\![1,n]\!]\setminus c_1)$, so we can complete c_1 by any of the $(n-k)!$ permutations $\tau\in\mathfrak{S}([\![1,n]\!]\setminus c_1)$ in order

to get a permutation $\sigma = c_1 \circ \tau$ such that $s_1(\sigma) = k$. So, the number of permutations σ such that $s_1(\sigma) = k$ is

$$\frac{(n-1)!}{(n-k)!}(n-k)! = (n-1)!$$

(3) Therefore,

$$\mathbb{P}[s_1(\sigma_n) = k] = \frac{1}{n!}(n-1)! = \frac{1}{n}.$$

So, $s_1(\sigma_n)$ follows the uniform distribution on $[\![1, n]\!]$. Then, for any bounded continuous function f,

$$\mathbb{E}\left[f\left(\frac{s_1(\sigma_n)}{n}\right)\right] = \frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right) \to_{n\to\infty} \int_0^1 f(x)\ \mathrm{d}x$$

by a Riemann sum argument. So, $\frac{s_1(\sigma_n)}{n}$ converges in law to a uniform random variable U_1 in $[0, 1]$.

(4) The implication $\Rightarrow$ is trivial since monomial functions $(x_1, \dots, x_k) \mapsto (x_1)^{m_1} \cdots (x_k)^{m_k}$ are continuous bounded functions on the space $\mathfrak{X} = [0, 1]^k$. Conversely, by Stone–Weierstrass, the polynomials form a dense subalgebra of $\mathscr{C}([0, 1]^k)$. Suppose the convergence of the expectation of any monomial function, and consider a continuous function f on $[0, 1]^k$, which is necessarily bounded by compactness. There exists a linear combination of monomial functions g such that $\|f - g\|_\infty \leq \varepsilon$, and by assumption,

$$\left|\mathbb{E}[g(X_{n,1}, \dots, X_{n,k})] - \mathbb{E}[g(X_1, \dots, X_k)]\right| \leq \varepsilon$$

for n large enough. Therefore, by the triangular inequality,

$$\left|\mathbb{E}[f(X_{n,1}, \dots, X_{n,k})] - \mathbb{E}[f(X_1, \dots, X_k)]\right| \leq 3\varepsilon$$

for n large enough. This implies the convergence in distribution $X_n \rightharpoonup X$.

(5) We have bounded random variables with values in $[0, 1]$, so by the previous question it suffices to prove that for any monomial $(x_1)^{m_1}(x_2)^{m_2} \cdots (x_k)^{m_k}$,

$$\mathbb{E}\left[\left(\frac{s_1(\sigma_n)}{n}\right)^{m_1} \cdots \left(\frac{s_k(\sigma_n)}{n}\right)^{m_k}\right] \to_{n\to\infty} \mathbb{E}[(S_1)^{m_1} \cdots (S_k)^{m_k}].$$

We already know that the result is true for $k = 1$. Suppose the result true up to order $k \geq 1$, and let us prove the case $k+1$ (for which we can assume the exponent $m_{k+1} \geq 1$). We compute the left-hand side by conditioning against $(A_k(\sigma_n), (\sigma_n)^{(k)})$:

$$\begin{aligned}\text{LHS} &= \mathbb{E}\left[\mathbb{E}\left[\left(\frac{s_1(\sigma_n)}{n}\right)^{m_1} \cdots \left(\frac{s_{k+1}(\sigma_n)}{n}\right)^{m_{k+1}} \mid (A_k(\sigma_n), (\sigma_n)^{(k)})\right]\right] \\ &= \mathbb{E}\left[\left(\frac{s_1(\sigma_n)}{n}\right)^{m_1} \cdots \left(\frac{s_k(\sigma_n)}{n}\right)^{m_k} \mathbb{E}\left[\left(\frac{s_{k+1}(\sigma_n)}{n}\right)^{m_{k+1}} \mid (A_k(\sigma_n), (\sigma_n)^{(k)})\right]\right].\end{aligned}$$

Indeed, knowing $(A_k(\sigma_n), (\sigma_n)^{(k)})$ allows one to determine $s_1(\sigma_n), \dots, s_k(\sigma_n)$, so these variables are measurable with respect to $\sigma(A_k(\sigma_n), (\sigma_n)^{(k)})$, and one can extract them from the conditional expectation. By the result from the first question, the conditional law of $s_{k+1}(\sigma_n) = s_1((\sigma_n)_{(k)})$ is the uniform law on $[\![1, n']\!]$ with

$$n' = |\,[\![1, n]\!] \setminus A_k(\sigma_n)| = n - s_1(\sigma_n) - \cdots - s_k(\sigma_n).$$

Thus, the quantity of interest rewrites as:

$$\mathbb{E}\left[\left(\frac{s_1(\sigma_n)}{n}\right)^{m_1} \cdots \left(\frac{s_k(\sigma_n)}{n}\right)^{m_k} \left(\frac{1}{n - s_1(\sigma_n) - \cdots - s_k(\sigma_n)} \sum_{l=1}^{n-s_1(\sigma_n)-\cdots-s_k(\sigma_n)} \left(\frac{l}{n}\right)^{m_{k+1}}\right)\right].$$

In the expectation, the last parenthesis is by convention 0 if $s_1(\sigma_n) + \cdots + s_k(\sigma_n) = n$. By using a comparison between sums and integrals, one can rewrite the sum as:

$$\frac{1}{m_{k+1}+1}\left(1 - \frac{s_1(\sigma_n) + \cdots + s_k(\sigma_n)}{n}\right)^{m_{k+1}} (1 + X_n)\,,$$

where the variable X_n is bounded in absolute value by $\frac{K}{n}$ for some constant K (which depends on m_{k+1}). We can now use the induction hypothesis:

$$\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{s_1(\sigma_n)}{n}\right)^{m_1} \cdots \left(\frac{s_k(\sigma_n)}{n}\right)^{m_k}\right] = \mathbb{E}\left[(S_1)^{m_1} \cdots (S_k)^{m_k}\, \frac{(1 - S_1 - \cdots - S_k)^{m_{k+1}}}{m_{k+1}+1}\right],$$

where $S_1, \ldots, S_k$ are the k first values of a stick-breaking process. However, if U_{k+1} is a uniform variable in $[0,1]$ independent from $S_1, \ldots, S_k$, then the right-hand side rewrites as:

$$\mathbb{E}\big[(S_1)^{m_1} \cdots (S_k)^{m_k}\, (U_{k+1}(1 - S_1 - \cdots - S_k))^{m_{k+1}}\big] = \mathbb{E}\big[(S_1)^{m_1} \cdots (S_{k+1})^{m_{k+1}}\big]$$

with $S_{k+1} = U_{k+1}(1 - S_1 - \cdots - S_k)$. This ends our proof by induction.