Week 12: convergence in distribution

Submission of solutions. Feedback can be given on Exercise 1 and any other exercise from the Training exercises. If you want to hand in, do it so by Monday 11/12/2023 17:00 (online) following the instructions on the course website

https://metaphor.ethz.ch/x/2023/hs/401-3601-00L/

Please pay attention to the quality, the precision and the presentation of your mathematical writing.

***

1 Exercise covered during the exercise class

The following exercise will be covered during the exercise class.

Exercise 1. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables following the uniform distribution on $[0,1]$.

(1) Show that $n\min(X_1,\ldots,X_n)$ converges in distribution to a random variable Z when $n\to\infty$ and give the law of Z.

(2) Show that

$$(X_1+\cdots+X_n)\min(X_1,\ldots,X_n) \quad \xrightarrow[n\to\infty]{(d)} \quad Z/2.$$

Solution:

(1) We calculate the limit of the cumulative distribution function of $n\min(X_1,\ldots,X_n)$. For $y<0$, we have $\mathbb{P}(n\min(X_1,\ldots,X_n)\leq y)=0$. For $y\geq 0$ and $n>y$ we write

$$\mathbb{P}(n\min(X_1,\ldots,X_n)\leq y)=1-\mathbb{P}\left(\min(X_1,\ldots,X_n)>\frac{y}{n}\right)=1-\mathbb{P}\left(X_1>\frac{y}{n},\ldots,X_n>\frac{y}{n}\right)=1-\left(1-\frac{y}{n}\right)^n,$$

where we have used independence for the last equality. But

$$\left(1-\frac{y}{n}\right)^n=\exp\left(n\ln\left(1-\frac{y}{n}\right)\right)=\exp\left(n\left(-\frac{y}{n}+o\left(\frac{1}{n}\right)\right)\right) \quad \xrightarrow[n\to\infty]{} \quad e^{-y}.$$

We conclude that

$$\mathbb{P}(n\min(X_1,\ldots,X_n)\leq y) \quad \xrightarrow[n\to\infty]{} \quad 1-e^{-y}.$$

So the cdf of $n\min(X_1,\ldots,X_n)$ converges pointwise to the cdf of an exponential distribution with parameter 1, which is continuous. We conclude that $n\min(X_1,\ldots,X_n)$ converges in law to an exponential random variable of parameter 1.

(2) Set $Z_n = n\min(X_1,\dots,X_n)$ and $Y_n = \frac{X_1+\cdots+X_n}{n}$. Then $(X_1+\cdots+X_n)\min(X_1,\dots,X_n) = Y_nZ_n$. By (1), Z_n converges in distribution to Z. By the strong law of large numbers, Y_n converges a.s. and thus in probability to $\mathbb{E}[X_1] = 1/2$. By Slutsky's theorem (Y_n, Z_n) converges in distribution to $(1/2, Z)$.

By using the continuity of $f(y,z) = yz$, this implies that Y_nZ_n converges in distribution to $Z/2$.

□

2 Training exercises

Exercise 2. Let $(X_n)_{n\geq 1}$ be a sequence of real-valued random variables such that X_n has density p_n. Assume that there is a measurable function p such that $p_n(x) \to p(x)$ for λ almost all x (where λ is the Lebesgue measure).

(1) Is p always the density of some random variable? Justify your answer.

(2) Assume that there is an integrable (with respect to λ) measurable function $q : \mathbb{R} \to \mathbb{R}_+$ such that for every $n \geq 1$, $p_n(x) \leq q(x)$ for λ-almost all x. Show that p is the density of some random variable X and that X_n converges in distribution to X.

Solution:

(1) No, for example if X_n is uniform on $[n, n+1]$ then p_n converges pointwise to 0, which is not the density of some random variable.

(2) First $p \geq 0$ almost everywhere since $p_n \geq 0$. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous bound function. Using the transfer theorem, write:

$$\mathbb{E}[f(X_n)] = \int_{\mathbb{R}} f(x)p_n(x)\mathrm{d}x.$$

Observe that $f(x)p_n(x) \to f(x)p(x)$ for almost all x and that $|f(x)p_n(x)| \leq q(x)\|f\|_\infty$, which is integrable by assumption. By dominated convergence we thus get that

$$\int_{\mathbb{R}} f(x)p_n(x)\mathrm{d}x \quad \underset{n\to\infty}{\longrightarrow} \quad \int_{\mathbb{R}} f(x)p(x)\mathrm{d}x.$$

By taking f to be the constant function equal to 1, we get $\int_{\mathbb{R}} p(x)\mathrm{d}x = 1$, so that p is the density of some random variable X. In addition, by the transfer theorem again

$$\int_{\mathbb{R}} f(x)p(x)\mathrm{d}x = \mathbb{E}[f(X)],$$

so $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$, which establishes convergence in distribution.

□

Exercise 3. Let $(X_n)_{n\geq 1}$ and X be real-valued random variables such that $\mathbb{P}(X=t)=0$ for every $t\in\mathbb{R}$. Show that X_n converges in distribution to X if and only if $\mathbb{P}(X_n<t)\to\mathbb{P}(X<t)=\mathbb{P}(X\leq t)$ for every $t\in\mathbb{R}$.

Solution:

For every $t\in\mathbb{R}$ since $\mathbb{P}(X=t)=0$, we have $\mathbb{P}(X\leq t)=\mathbb{P}(X<t)$.

For the implication, we use the Portemanteau theorem with the set $B=(-\infty,t)$: since $\overline{B}\backslash\mathring{B}=\{t\}$ and $\mathbb{P}(X\in\{t\})=0$, this implies that $\mathbb{P}(X_n<t)\to\mathbb{P}(X<t)$.

For the converse, assume that $\mathbb{P}(X_n<t)\to\mathbb{P}(X\leq t)$ for every $t\in\mathbb{R}$. We show that $\mathbb{P}(X_n\leq t)\to\mathbb{P}(X\leq t)$ for every $t\in\mathbb{R}$ by adapting a proof seen in the lecture.

First, clearly $\mathbb{P}(X_n<t)\leq\mathbb{P}(X_n\leq t)$, so

$$\mathbb{P}(X\leq t)=\mathbb{P}(X\leq t)=\lim_{n\to\infty}\mathbb{P}(X_n<t)\leq\liminf_{n\to\infty}\mathbb{P}(X_n\leq t)\leq\limsup_{n\to\infty}\mathbb{P}(X_n\leq t).$$

Next, take $u>t$ and write

$$\limsup_{n\to\infty}\mathbb{P}(X_n\leq t)\leq\limsup_{n\to\infty}\mathbb{P}(X_n<u)=\mathbb{P}(X\leq u).$$

Thus, for every $u>t$,

$$\mathbb{P}(X\leq t)\leq\liminf_{n\to\infty}\mathbb{P}(X_n\leq t)\leq\limsup_{n\to\infty}\mathbb{P}(X_n\leq t)\leq\mathbb{P}(X\leq u).$$

But as $u\to t$ with $u>t$ we have $\mathbb{P}(X\leq u)\downarrow\mathbb{P}(X\leq t)$ by right-continuity of the cdf at t. This shows that $\mathbb{P}(X_n\leq t)\to\mathbb{P}(X\leq t)$ and completes the proof. □

Exercise 4. Let $f:\mathbb{R}_+\to\mathbb{R}_+$ be a C^1 (continuously differentiable) weakly increasing function with $f(0)=0$.

(1) Let X be a non-negative real-valued random variable. Show that

$$\mathbb{E}[f(X)]=\int_0^\infty f'(x)\mathbb{P}(X>x)\mathrm{d}x.$$

(2) Let $(X_n)_{n\geq 1}$ be a sequence of non-negative real valued random variables converging in distribution to X.

(a) Show that $\mathbb{P}(X\geq 0)=1$.

(b) Show that $\mathbb{E}[f(X)]\leq\liminf_{n\to\infty}\mathbb{E}[f(X_n)]$.

Solution:

(1) Using Fubini-Tonelli (observe that $f' \geq 0$), write

$$\int_0^\infty f'(x)\mathbb{P}(X > x)\mathrm{d}x = \int_0^\infty f'(x)\mathbb{E}[\mathbb{1}_{X>x}]\mathrm{d}x = \mathbb{E}\left[\int_0^\infty f'(x)\mathbb{1}_{X>x}\mathrm{d}x\right] = \mathbb{E}\left[\int_0^X f'(x)\mathrm{d}x\right] = \mathbb{E}[f(X)].$$

(2) (a) Since $(-\infty, 0)$ is open, by the Portemanteau Theorem we have

$$0 = \liminf_{n\to\infty} \mathbb{P}(X_n < 0) \geq \mathbb{P}(X < 0),$$

so $\mathbb{P}(X < 0) = 0$, which implies $\mathbb{P}(X \geq 0) = 1$.

(b) Using (1) write

$$\mathbb{E}[f(X_n)] = \int_0^\infty f'(x)\mathbb{P}(X_n > x)\mathrm{d}x.$$

Since $X_n \to X$ in distribution, we know that $\mathbb{P}(X_n > x) \to \mathbb{P}(X > x)$ for every x where the cdf of X is continuous. Since the cdf of X has at most a countable amount of discontinuity points, it follows that $\mathbb{P}(X_n > x) \to \mathbb{P}(X > x)$ for almost Lebesgue-all x. Thus, by Fatou's Lemma:

$$\begin{aligned}
\mathbb{E}[f(X)] &= \int_0^\infty f'(x)\mathbb{P}(X > x)\mathrm{d}x \\
&= \int_0^\infty \liminf_{n\to\infty} f'(x)\mathbb{P}(X_n > x)\mathrm{d}x \\
&\leq \liminf_{n\to\infty} \int_0^\infty f'(x)\mathbb{P}(X_n > x)\mathrm{d}x \\
&= \mathbb{E}[f(X_n)].
\end{aligned}$$

□

Exercise 5. Let $(X_n)_{n\geq 1}$ be a sequence of real-valued random variables converging in distribution to a uniform random variable on $[0,1]$. Let $(Y_n)_{n\geq 1}$ be a sequence of real-valued random variables converging in probability to 0. Show that $\mathbb{P}(X_n < Y_n) \to 0$ as $n \to \infty$.

Solution:

Fix $\epsilon \in (0,1)$ and write

$$\begin{aligned}
\mathbb{P}(X_n < Y_n) &= \mathbb{P}(X_n < Y_n, Y_n > \varepsilon) + \mathbb{P}(X_n < Y_n, Y_n \leq \varepsilon) \\
&\leq \mathbb{P}(Y_n > \varepsilon) + \mathbb{P}(X_n < \varepsilon)
\end{aligned}$$

Denote by X a uniform random variable on $[0,1]$. But $\mathbb{P}(Y_n > \varepsilon) \to 0$ because $Y_n \to 0$ in probability, and by Portemanteau's theorem $\limsup_{n\to\infty} \mathbb{P}(X_n < \varepsilon) \leq \mathbb{P}(X < \varepsilon) = \varepsilon$ because $X_n \to X$ in distribution and $\mathbb{P}(X = \varepsilon) = 0$. We conclude that

$$\limsup_{n\to\infty} \mathbb{P}(X_n < Y_n) \le \varepsilon,$$

which entails the desired result. □

3 More involved exercise (optional, will not be covered in the exercise class)

Exercise 6. A stick of length 1 is broken at n points chosen uniformly and independently at random. Let L_n be the length of the longest of the $n+1$ pieces obtained. How does L_n behave when $n \to \infty$?

The aim of this exercise is to show that $(n+1)L_n - \ln(n+1)$ converges in distribution to a real-valued random variable whose cdf is $x \mapsto e^{e^{-x}}$ on $\mathbb{R}$ (called a Gumbel distribution).

Part 1. To model the problem, let $(U_i)_{1\le i\le n}$ be i.i.d. uniform random variables on $[0,1]$ representing the locations where the stick is broken.

(1) Show that $\mathbb{P}\left(\exists i,j \in \{1,2,\dots,n\} : i \neq j \text{ and } U_i = U_j\right) = 0$.

(2) Show that there exists a random permutation σ such that $\mathbb{P}\left(U_{\sigma(1)} < \cdots < U_{\sigma(n)}\right) = 1$

Thus if $(\Delta_1,\dots,\Delta_{n+1})$ denote the lenghts of the pieces, we have $\Delta_i = U_{\sigma_i} - U_{\sigma_{i-1}}$ for $1 \le i \le n+1$ (with the convention $U_{\sigma_{n+1}} = 1$ and $U_{\sigma_0} = 0$).

(3) Show that $(U_{\sigma(1)},\dots,U_{\sigma(n)})$ has density

$$n!\, \mathbb{1}_{\{0\le x_1<\ldots<x_n\le 1\}}\, dx_1 \ldots dx_n.$$

Part 2. Let $(X_i)_{1\le i\le n+1}$ be exponential i.i.d. random variables with parameter 1. For $1 \le i \le n+1$, set

$$S_i = X_1 + \cdots + X_i, \qquad Y_i = \frac{X_i}{S_{n+1}}.$$

(4) Determine the joint law of $(X_1,\dots,X_n,S_{n+1})$ and deduce that of $(Y_1,\dots,Y_n)$.

(5) Show that $(\Delta_1,\dots,\Delta_n)$ and $(Y_1,\dots,Y_n)$ have the same distribution. Deduce that $\max(Y_1,\dots,Y_{n+1})$ has the same law as L_n.

(6) Show that for $x \in \mathbb{R}$, $(x + \ln(n+1))\left(\frac{S_{n+1}}{n+1} - 1\right)$ converges in probability to 0.

(7) Deduce the desired result.

Solution:

(1) We have

$$\mathbb{P}\left(\exists i,j \in \{1,2,\dots,n\} : i \neq j \text{ et } U_i = U_j\right) \le \sum_{i,j\in\{1,2,\dots,n\}, i\neq j} \mathbb{P}\left(U_i = U_j\right).$$

But using the transfer theorem

$$\begin{aligned} \mathbb{P}\left(U_i = U_j\right) &= \mathbb{E}\left[\mathbb{1}_{U_i=U_j}\right] \\ &= \int_{[0,1]^n} dx_1 dx_2 \cdots dx_n \mathbb{1}_{x_i=x_j} \\ &= \int_{[0,1]^2} dx dy \mathbb{1}_{x=y} \\ &= \int_{[0,1]} dx \int_x^x dy \\ &= 0, \end{aligned}$$

and the result follows.

(2) Let A be the event $\{\forall i, j \in \{1,2,\ldots,n\}:$ we have $U_i \neq U_j$ if $i \neq j\}$, so that $\mathbb{P}(A) = 1$ by (1). On event A (i.e. if $\omega \in A$), the numbers $U_1, \ldots, U_n$ can be arranged in (strictly) increasing order. We can therefore define σ so that $U_{\sigma(1)} < \cdots < U_{\sigma(n)}$ when $\omega \in A$. If $\omega \neq A$, we define σ to be equal to the identity, so that σ is well defined on the whole set Ω. Since $\mathbb{P}(A^c) = 0$, we have

$$\mathbb{P}\left(U_{\sigma(1)} < \cdots < U_{\sigma(n)}\right) = \mathbb{P}\left(U_{\sigma(1)} < \cdots < U_{\sigma(n)} \cap A\right) = \mathbb{P}(A) = 1$$

because by construction the events $\{U_{\sigma(1)} < \cdots < U_{\sigma(n)}\} \cap A$ and A are equal.

Remark. σ depends on ω, but as usual in probability theory, this dependence is not explicitly written down. Another tricky point is that we need to define σ on the whole domain Ω, since a random variable is by definition an application defined on Ω. That's why we had to define σ on A on the one hand, and on the complementary of A on the other.

Now let τ be a fixed permutation of $\{1,2,\ldots,n\}$. Then by the transfer theorem

$$\mathbb{P}\left(U_{\tau(1)} < \cdots < U_{\tau(n)}\right) = \int_{[0,1]^n} dx_1 \cdots dx_n \mathbb{1}_{x_{\tau(1)} < x_{\tau(2)} < \cdots < x_{\tau(n)}} = \int_{[0,1]^n} dx_1 \cdots dx_n \mathbb{1}_{x_1 < x_2 < \cdots < x_n}$$

by the change of variables $x_i' = x_{\tau(i)}$. The last quantity is $\mathbb{P}(U_1 < \cdots < U_n)$ and thus, the probability

$$\mathbb{P}(\sigma = \tau) = \mathbb{P}\left(U_{\tau(1)} < \cdots < U_{\tau(n)}\right)$$

does not depend on the permutation τ. The random permutation σ therefore follows the uniform law on permutations of $\{1,2,\ldots,n\}$.

(3) Let $f : [0,1]^n \to \mathbb{R}_+$ be measurable. Denote by $\mathcal{S}_n$ the set of all permuations of $\{1,2,\ldots,n\}$. As in

the computations of (2), write

$$\begin{aligned}
\mathbb{E}\Big[f(U_{\sigma(1)},\dots,U_{\sigma(n)})\Big] &= \sum_{\tau\in\mathcal{S}_n}\mathbb{E}\Big[f(U_{\sigma(1)},\dots,U_{\sigma(n)})\mathbb{1}_{\sigma=\tau}\Big]\\
&= \sum_{\tau\in\mathcal{S}_n}\mathbb{E}\Big[f(U_{\tau(1)},\dots,U_{\tau(n)})\mathbb{1}_{U_{\tau(1)}<\dots<U_{\tau(n)}}\Big]\\
&= \sum_{\tau\in\mathcal{S}_n}\int_{[0,1]^n} dx_1dx_2\cdots dx_n\mathbb{1}_{x_{\tau(1)}<x_{\tau(2)}<\dots<x_{\tau(n)}}f(x_{\tau(1)},\dots,x_{\tau(n)})\\
&= \sum_{\tau\in\mathcal{S}_n}\int_{[0,1]^n} dx_1dx_2\cdots dx_n\mathbb{1}_{x_1<x_2<\dots<x_n}f(x_1,\dots,x_n)\\
&= n!\int_{[0,1]^n} dx_1dx_2\cdots dx_n\mathbb{1}_{x_1<x_2<\dots<x_n}f(x_1,\dots,x_n),
\end{aligned}$$

which gives the desired result (we can change $<$ into $\leq$ by (1)).

(4) We start by determining the joint distribution $(X_1,\dots,X_n,S_{n+1})$ using the dummy function method. Let $f:\mathbb{R}^{n+1}\to\mathbb{R}$ be a continuous bounded function. According to the transfer theorem

$$\mathbb{E}[f(X_1,\dots,X_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(x_1,x_2,\dots,x_1+\dots+x_{n+1})e^{-(x_1+\dots+x_{n+1})}dx_1\cdots d_{x_{n+1}}.$$

By making the change of variable $u_1=x_1,\dots,u_n=x_n,u_{n+1}=x_1+\dots+x_n+1$ of Jacobian 1, we obtain

$$\mathbb{E}[f(X_1,\dots,X_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(u_1,u_2,\dots,u_{n+1})e^{-u_{n+1}}\mathbb{1}_{u_{n+1}\geq u_1+\dots+u_n}du_1\cdots du_{n+1}$$

which determines the joint distribution $(X_1,\dots,X_n,S_{n+1})$.

We now determine the joint distribution $(Y_1,\dots,Y_n,S_{n+1})$. Let $f:\mathbb{R}^{n+1}\to\mathbb{R}$ be a continuous bounded function. According to the transfer theorem applied to $(X_1,\dots,X_n,S_{n+1})$:

$$\mathbb{E}[f(Y_1,\dots,Y_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f\left(\frac{x_1}{x_{n+1}},\dots\frac{x_n}{x_{n+1}},x_{n+1}\right)\mathbb{1}_{x_{n+1}\geq x_1+\dots+x_n}e^{-x_{n+1}}dx_1\cdots d_{x_{n+1}}.$$

By making the change of variables

$$u_1=\frac{x_1}{x_{n+1}},u_2=\frac{x_2}{x_{n+1}},\dots,u_n=\frac{x_n}{x_{n+1}},u_{n+1}=x_{n+1},$$

avec $0\leq u_1+\dots+u_n\leq 1$ et $x_i=u_iu_{n+1}$, with Jacobian u_{n+1}^n we get

$$\mathbb{E}[f(Y_1,\dots,Y_n,S_{n+1})] = \int_{]0,\infty[^{n+1}} f(u_1,u_2,\dots,u_{n+1})u_{n+1}^n e^{-u_{n+1}}\mathbb{1}_{u_1+\dots+u_n<1}du_1\cdots du_{n+1}.$$

We deduce that for any function $f : \mathbb{R}^n \to \mathbb{R}$ continuous bounded we have

$$\mathbb{E}[f(Y_1,\dots,Y_n)] = n! \int_{]0,\infty[^n} f(u_1,u_2,\dots,u_n)\mathbb{1}_{u_1+\dots+u_n<1} du_1 \cdots du_n, \qquad (1)$$

which determines the law of $(Y_1,\dots,Y_n)$.

(5) Recall that $\Delta_1 = U_{\sigma(1)}$ and $\Delta_i = Z_{\sigma(i)} - Z_{\sigma(i-1)}$ for $2 \le i \le n$. By a change of variable, we deduce that

$$\mathbb{E}[f(\Delta_1,\Delta_2,\dots,\Delta_n)] = n! \int_{]0,\infty[^n} f(u_1,u_2,\dots,u_n)\mathbb{1}_{u_1+\dots+u_n<1} du_1 \cdots du_n.$$

Given (1), we deduce that $(Y_1,\dots,Y_n)$ and $(\Delta_1,\dots,\Delta_n)$ have the same distribution. Since $Y_{n+1} = 1 - Y_1 - \dots - Y_n$ and $\Delta_{n+1} = 1 - \Delta_1 - \dots - \Delta_n$, we conclude that $(Y_1,\dots,Y_{n+1})$ and $(\Delta_1,\dots,\Delta_{n+1})$ have the same law. The fact that $\max(Y_1,\dots,Y_{n+1})$ has the same law as L_n follows immediately from this, since $L_n = \max(\Delta_1,\dots,\Delta_{n+1})$.

(6) Let $W_n = (x+\ln(n+1))\left(\frac{S_{n+1}}{n+1} - 1\right)$. We have $\mathbb{E}[W] = 0$ and $\mathrm{Var}(W) \sim \frac{\ln(n)^2}{n} \to 0$. From the Bienaymé-Chebyshev inequality, we deduce that $W_n \to 0$ in probability.

(7) Let $Z_n = \max(X_1,\dots,X_{n+1}) - \log(n+1)$. From question (5), we have

$$\begin{aligned}
\mathbb{P}((n+1)L_n - \ln(n+1) \le x) &= \mathbb{P}((n+1)\max(Y_1,\dots,Y_{n+1}) - \ln(n+1) \le x) \\
&= \mathbb{P}\left((n+1)\frac{\max(X_1,\dots,X_{n+1})}{S_{n+1}} - \ln(n+1) \le x\right) \\
&= \mathbb{P}\left(\max(X_1,\dots,X_{n+1}) - \ln(n+1)\frac{S_{n+1}}{n+1} \le x\frac{S_{n+1}}{n+1}\right) \\
&= \mathbb{P}\left(\max(X_1,\dots,X_{n+1}) - \ln(n+1) \le x + (x+\ln(n+1))\left(\frac{S_{n+1}}{n+1} - 1\right)\right) \\
&= \mathbb{P}(Z_n - W_n \le x)
\end{aligned}$$

Now a little calculation shows that $\mathbb{P}(\max(X_1,\dots,X_{n+1}) - \log(n+1) \le x) \to e^{e^{-x}}$ for all $x \in \mathbb{R}$. So Z_n converges to $\mathcal{G}$, where $\mathcal{G}$ is the Gumbel distribution function $x \mapsto e^{e^{-x}}$. Now $W_n \to 0$ in probability, so $Z_n + W_n$ converges in law to $\mathcal{G}$ according to Slutsky's lemma. So $\mathbb{P}(Z_n + W_n \le x) \to e^{e^{-x}}$. We conclude that for all $x \in \mathbb{R}$ we have

$$\mathbb{P}((n+1)L_n - \ln(n+1) \le x) \underset{n\to\infty}{\longrightarrow} e^{e^{-x}},$$

which was the desired result.

□

4 Fun exercise (optional, will not be covered in the exercise class)

Exercise 7. Let $n \ge 1$ be an integer. An urn contains n white balls and n colored balls. The balls are drawn successively and without replacement until there are only balls of one color left in the urn. As $n \to \infty$, what is the behavior of the number of remaining balls?

Solution:
Let H_n be the random variable equal to the number of remaining balls. We show that (H_n) converges in distribution to a G(1/2) distribution, that is for every $k \geq 1$

$$\mathbb{P}(H_n = k) \quad \underset{n\to\infty}{\longrightarrow} \quad \frac{1}{2^k}. \tag{2}$$

Let $k \in \{1, 2, \dots, n\}$. There are k white balls left in the urn at the end if, and only if, the $(2n-k)^{\text{th}}$ draw results in a colored ball, and all $(n-1)$ colored balls have already been drawn in the first $(2n-k-1)$ draws. There are therefore $\binom{2n-k-1}{n-1}$ such configurations. By symmetry, the probability of k white balls remaining is the same as the probability of k colored balls remaining. Thus,

$$\mathbb{P}(H_n = k) = 2\frac{\binom{2n-k-1}{n-1}}{\binom{2n}{n}}.$$

According to Stirling's formula,

$$\binom{2n}{n} \sim \frac{4^n}{\sqrt{\pi n}}, \qquad \binom{2n-k-1}{n-1} \sim \frac{1}{2^{k+1}}\frac{4^n}{\sqrt{\pi n}},$$

and (2) follows. □