Course: Theory of Probability I
Term: Fall 2013
Instructor: Gordan Zitkovic

Lecture 9

The weak law of large numbers and the central limit theorem

The weak law of large numbers

We start with a definitive form of the weak law of large numbers. We need two lemmas, first[1]:

[1] Feel free to skip the proofs, but understand why the statement of Lemma 9.2 even needs one

Lemma 9.1. *Let $u_1, u_2, \dots, u_n$ and $w_1, w_2, \dots, w_n$ be complex numbers, all of modulus at most $M > 0$. Then*

$$\left|\prod_{k=1}^{n} u_k - \prod_{k=1}^{n} w_k\right| \leq M^{n-1} \sum_{k=1}^{n} |u_k - w_k| . \tag{9.1}$$

Proof. We proceed by induction. For $n = 1$, the claim is trivial. Suppose that (9.1) holds. Then

$$\begin{aligned}\left|\prod_{k=1}^{n+1} u_k - \prod_{k=1}^{n+1} w_k\right| &\leq \left|\prod_{k=1}^{n} u_k\right| |u_{n+1} - w_{n+1}| + |w_{n+1}| \left|\prod_{k=1}^{n} u_k - \prod_{k=1}^{n} w_k\right| \\ &\leq M^n |u_{n+1} - w_{n+1}| + M \times M^{n-1} \sum_{k=1}^{n} |u_k - w_k| \\ &= M^{(n+1)-1} \sum_{k=1}^{n+1} |u_k - w_k| .\end{aligned}$$

□

Lemma 9.2. *Let $\{z_n\}_{n\in\mathbb{N}}$ be a sequence of complex numbers with $z_n \to z \in \mathbb{C}$. Then $(1 + \frac{z_n}{n})^n \to e^z$.*

Proof. Using Lemma 9.1 with $u_k = 1 + \frac{z_n}{n}$ and $w_k = e^{z_n/n}$ for $k = 1, \dots, n$, we get

$$\left|(1 + \tfrac{z_n}{n})^n - e^{z_n}\right| \leq n M_n^{n-1} \left|1 + \tfrac{z_n}{n} - e^{z_n/n}\right| , \tag{9.2}$$

where $M_n = \max(\left|1 + \frac{z_n}{n}\right| , \left|e^{z_n/n}\right|)$. Let $K = \sup_{n\in\mathbb{N}} |z_n| < \infty$, so that $\left|e^{z_n/n}\right|^n \leq e^{|z_n|} \leq e^K$. Similarly, $\left|1 + \frac{z_n}{n}\right|^n \leq (1 + \frac{K}{n})^n \to e^K$. Therefore

$$L = \sup_{n\in\mathbb{N}} M_n^{n-1} < \infty. \tag{9.3}$$

To estimate the last term in (9.2), we start with the Taylor expansion $e^b = 1 + b + \sum_{k\geq 2} \frac{b^k}{k!}$, which converges absolutely for all $b \in \mathbb{C}$. Then, we use the fact that $\frac{1}{k!} \leq 2^{-k+1}$, to obtain

$$\left|e^b - 1 - b\right| \leq \sum_{k\geq 2} |b|^k \tfrac{1}{k!} \leq \left|b^2\right| \sum_{k\geq 2} 2^{-k+1} = |b|^2, \text{ for } |b| \leq 1. \tag{9.4}$$

Since $|z_n|/n \leq 1$ for large-enough n, it follows from (9.2), (9.4) and (9.3), that

$$\limsup_n \left|(1 + \tfrac{z_n}{n})^n - e^{z_n}\right| \leq \limsup_n nL \left|\tfrac{z_n}{n}\right|^2 = 0.$$

It remains to remember that $e^{z_n} \to e^z$ to finish the proof. □

Theorem 9.3 (Weak law of large numbers)**.** *Let $\{X_n\}_{n\in\mathbb{N}}$ be an iid sequence of random variables with the (common) distribution μ and the characteristic function $\varphi = \varphi_\mu$ such that $\varphi'(0)$ exists. Then, $c = -i\varphi'(0)$ is a real number and*

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to c \text{ in probability.}$$

Proof. Since $\varphi(-s) = \overline{\varphi(s)}$, we have

$$\varphi'(0) = \lim_{s\to 0} \tfrac{\varphi(-s)-1}{-s} = \lim_{s\to 0} \tfrac{\overline{\varphi(s)}-1}{-s} = -\lim_{s\to 0} \overline{\tfrac{\varphi(s)-1}{s}} = -\overline{\varphi'(0)}.$$

Therefore, $c = -i\varphi'(0) \in \mathbb{R}$.

Let $S_n = \sum_{k=1}^n X_k$. According to Proposition 7.10, it will be enough to show that $\frac{1}{n}S_n \xrightarrow{\mathcal{D}} c = -i\varphi'(0) \in \mathbb{R}$. Moreover, by Theorem 8.10, all we need to do is show that $\varphi_{\frac{1}{n}S_n}(t) \to e^{itc} = e^{t\varphi'(0)}$, for all $t \in \mathbb{R}$.

The iid property of $\{X_n\}_{n\in\mathbb{N}}$ and the fact that $\varphi_{\alpha X}(t) = \varphi_X(\alpha t)$ imply that

$$\varphi_{\frac{1}{n}S_n}(t) = (\varphi(\tfrac{t}{n}))^n = (1 + \tfrac{z_n}{n})^n,$$

where $z_n = n(\varphi(\frac{t}{n}) - 1)$. By the assumption, we have $\lim_{s\to 0} \frac{\varphi(s)-1}{s} = ic$, and so $z_n \to t\varphi'(0)$. Therefore, by Lemma 9.2 above, we have $\varphi_{\frac{1}{n}S_n}(t) \to e^{itc}$.

□

Remark 9.4.

1. It can be shown that the converse of Theorem 9.3 is true in the following sense: if $\frac{1}{n}S_n \xrightarrow{\mathbb{P}} c \in \mathbb{R}$, then $\varphi'(0)$ exists and $\varphi'(0) = ic$. That's why we call the result of Theorem 9.3 definitive.

2. $X_1 \in \mathbb{L}^1$ implies $\varphi'(0) = \mathbb{E}[X_1]$, so that Theorem 9.3 covers the classical case. As we have seen in Problem 8.6, there are cases when $\varphi'(0)$ exists but $\mathbb{E}[|X_1|] = \infty$.

Problem 9.1. Let $\{X_n\}_{n\in\mathbb{N}}$ be iid with the Cauchy distribution[2]. Show that φ_X is not differentiable at 0 and show that there is no constant c such that

$$\tfrac{1}{n}S_n \xrightarrow{\mathbb{P}} c,$$

where $S_n = \sum_{k=1}^n X_k$. *Hint:* What is the distribution of $\frac{1}{n}S_n$?

[2] Remember, the density of the Cauchy distribution is given by $f_X(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

An "iid"-central limit theorem

We continue with a central limit theorem for iid sequences. Unlike in the case of the (weak) law of large numbers, existence of the first moment will not be enough - we will need to assume that the second moment is finite, too. We will see how this assumption can be relaxed when we state and prove the Lindeberg-Feller theorem. We start with an estimate of the "error" term in the Taylor expansion of the exponential function of imaginary argument:

Lemma 9.5. *For $\xi \in \mathbb{R}$ we have*

$$\left| e^{i\xi} - \sum_{k=0}^{n} \tfrac{(i\xi)^k}{k!} \right| \le \min\left(\tfrac{|\xi|^{n+1}}{(n+1)!}, 2\tfrac{|\xi|^n}{n!}\right)$$

Proof. If we write the remainder in the Taylor formula in the integral form (derived easily using integration by parts), we get

$$e^{i\xi} - \sum_{k=0}^{n} \tfrac{(i\xi)^k}{k!} = R_n(\xi), \text{ where } R_n(\xi) = i^{n+1}\int_0^{\xi} e^{iu}\tfrac{(\xi-u)^n}{n!}\,du.$$

The usual estimate of R_n gives:

$$|R_n(\xi)| \le \tfrac{1}{n!}\int_0^{|\xi|}(|\xi|-u)^n\,du = \tfrac{|\xi|^{n+1}}{(n+1)!}.$$

We could also transform the expression for R_n by integrating it by parts:

$$\begin{aligned} R_n(\xi) &= \tfrac{i^{n+1}}{n!}\left(\tfrac{1}{i}\xi^n - \tfrac{n}{i}\int_0^{\xi} e^{iu}(\xi-u)^{n-1}\,du\right) \\ &= \tfrac{i^n}{(n-1)!}\left(\int_0^{\xi}(\xi-u)^{n-1}\,du - \int_0^{\xi} e^{iu}(\xi-u)^{n-1}\,du\right), \end{aligned}$$

since $\xi^n = n\int_0^{\xi}(\xi-u)^{n-1}\,du$. Therefore

$$\begin{aligned} |R_n(\xi)| &\le \tfrac{1}{(n-1)!}\int_0^{|\xi|}(|\xi|-u)^{n-1}\left|e^{iu}-1\right|\,du \\ &\le \tfrac{2}{n!}\int_0^{|\xi|} n(|\xi|-u)^{n-1}\,du = \tfrac{2|\xi|^n}{n!}. \end{aligned}$$

□

While the following result can be obtained as a direct consequence of twice-differentiability of the function φ at 0, we use the (otherwise useful) estimate based on Lemma 9.5 above:

Corollary 9.6. *Let X be a random variable with $\mathbb{E}[X] = \mu$ and $\mathbb{E}[X^2] = \nu < \infty$, and let the function $r : [0,\infty) \to [0,\infty)$ be defined by*

$$r(t) = \mathbb{E}[X^2 \min(t\,|X|\,,1)],\ t \geq 0. \tag{9.5}$$

Then

1. $\lim_{t\searrow 0} r(t) = 0$ *and*
2. $\left|\varphi_X(t) - 1 - it\mu + \frac{1}{2}\nu t^2\right| \leq t^2 r(|t|)$.

Proof. The inequality in 2. is a direct consequence of Lemma 9.5 (with the extra factor $\frac{1}{6}$ neglected). Part 1. follows from the dominated convergence theorem because

$$X^2 \min(1, t\,|X|) \leq X^2 \in \mathcal{L}^1 \text{ and } \lim_{t\to 0} X^2 \min(1, t\,|X|) = 0. \qquad \square$$

Theorem 9.7 (Central Limit Theorem - iid version)**.** *Let $\{X_n\}_{n\in\mathbb{N}}$ be an iid sequence of random variables with $0 < \mathrm{Var}[X_1] < \infty$. Then*

$$\tfrac{\sum_{k=1}^n (X_k - \mu)}{\sqrt{\sigma^2 n}} \xrightarrow{\mathcal{D}} \chi,$$

where $\chi \sim N(0,1)$, $\mu = \mathbb{E}[X_1]$ and $\sigma^2 = \mathrm{Var}[X_1]$.

Proof. By considering the sequence $\{(X_n - \mu)/\sqrt{\sigma^2}\}_{n\in\mathbb{N}}$, instead of $\{X_n\}_{n\in\mathbb{N}}$, we may assume that $\mu = 0$ and $\sigma = 1$. Let φ be the characteristic function of the common distribution of $\{X_n\}_{n\in\mathbb{N}}$ and set $S_n = \sum_{k=1}^n X_k$, so that

$$\varphi_{\frac{1}{\sqrt{n}} S_n}(t) = (\varphi(\tfrac{t}{\sqrt{n}}))^n.$$

By Theorem 8.10, the problem reduces to whether the following statement holds:

$$\lim_{n\to\infty} (\varphi(\tfrac{t}{\sqrt{n}}))^n = e^{-\frac{1}{2}t^2},\ \text{for each } t \in \mathbb{R}. \tag{9.6}$$

Corollary 9.6 guarantees that

$$\left|\varphi(t) - 1 + \tfrac{1}{2}t^2\right| \leq t^2 r(t) \text{ for all } t \in \mathbb{R},$$

where r is given by (9.5), i.e.,

$$\left|\varphi(\tfrac{t}{\sqrt{n}}) - 1 + \tfrac{t^2}{2n}\right| \leq \tfrac{t^2}{n} r(t/\sqrt{n}).$$

Lemma 9.1 with $u_1 = \cdots = u_n = \varphi(\frac{t}{\sqrt{n}})$ and $w_1 = \cdots = w_n = (1 - \frac{t^2}{2n})$ yields:

$$\left|(\varphi(\tfrac{t}{\sqrt{n}}))^n - (1 - \tfrac{t^2}{2n})^n\right| \leq t^2 r(t/\sqrt{n}),$$

for $n \geq \frac{2}{t^2}$ (so that $\max(|\varphi(\frac{t}{\sqrt{n}})|, |1 - \frac{t^2}{2n}|) \leq 1$). Since $\lim_n r(t/\sqrt{n}) = 0$, we have

$$\lim_{n\to\infty} \left|(\varphi(\tfrac{t}{\sqrt{n}}))^n - (1 - \tfrac{t^2}{2n})^n\right| = 0,$$

and (9.6) follows from the fact that $(1 - \frac{t^2}{2n})^n \to e^{-\frac{1}{2}t^2}$, for all t. □

The Lindeberg-Feller Theorem

Unlike Theorem 9.7, the Lindeberg-Feller Theorem does not require summands to be equally distributed - it only prohibits any single term from dominating the sum. As usual, we start with a technical lemma:

Lemma 9.8. *Let $(c_{n,m})$, $n \in \mathbb{N}$, $m = 1, \ldots, n$ be a (triangular) array of real numbers with*

1. $\sum_{m=1}^n c_{n,m} \to c \in \mathbb{R}$, *and* $\sum_{m=1}^n |c_{n,m}|$ *is a bounded sequence,*
2. $m_n \to 0$, *as* $n \to \infty$, *where* $m_n = \max_{1\leq m\leq n} |c_{n,m}|$.

Then

$$\prod_{m=1}^n (1 + c_{n,m}) \to e^c \text{ as } n \to \infty.$$

Proof. Without loss of generality we assume that $m_n < \frac{1}{2}$ for all n, and note that the statement is equivalent to $\sum_{m=1}^n \log(1 + c_{n,m}) \to c$, as $n \to \infty$. Since $\sum_{m=1}^n c_{n,m} \to c$, this is also equivalent to

$$\sum_{m=1}^n (\log(1 + c_{n,m}) - c_{n,m}) \to 0, \text{ as } n \to \infty. \tag{9.7}$$

Consider the function $f(x) = \log(1 + x) + x^2 - x$, $x > -1$. It is straightforward to check that $f(0) = 0$ and that the derivative $f'(x) = \frac{1}{1+x} + 2x - 1$ satisfies $f'(x) > 0$ for $x > 0$ and $f'(x) < 0$ for $x \in (-1/2, \infty)$. It follows that $f(x) \geq 0$ for $x \in [-1/2, \infty)$ so that (the absolute value can be inserted since $x \geq \log(1 + x)$)

$$|\log(1 + x) - x| \leq x^2 \text{ for } x \geq -\tfrac{1}{2}.$$

Since $m_n < \frac{1}{2}$, we have $|\log(1 + c_{n,m}) - c_{n,m}| \leq c_{n,m}^2$, and so

$$\begin{aligned}\left|\sum_{m=1}^n (\log(1 + c_{n,m}) - c_{n,m})\right| &\leq \sum_{m=1}^n |\log(1 + c_{n,m}) - c_{n,m}| \leq \sum_{m=1}^n c_{n,m}^2 \\ &\leq m_n \sum_{m=1}^n |c_{n,m}| \to 0,\end{aligned}$$

because $\sum_{m=1}^{n} |c_{n,m}|$ is bounded and $m_n \to 0$. $\square$

Theorem 9.9 (Lindeberg-Feller). *Let $X_{n,m}$, $n \in \mathbb{N}$, $m = 1, \dots, n$ be a (triangular) array of random variables such that*

1. $\mathbb{E}[X_{n,m}] = 0$, *for all* $n \in \mathbb{N}$, $m = 1, \dots, n$,
2. $X_{n,1}, \dots, X_{n,n}$ *are independent,*
3. $\sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2] \to \sigma^2 > 0$, *as* $n \to \infty$,
4. *for each* $\varepsilon > 0$, $s_n(\varepsilon) \to 0$, *as* $n \to \infty$, *where* $s_n(\varepsilon) = \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{n,m}| \geq \varepsilon\}}]$.

Then

$$X_{n,1} + \cdots + X_{n,n} \xrightarrow{\mathcal{D}} \sigma\, \chi, \text{ as } n \to \infty,$$

where $\chi \sim N(0,1)$.

Proof. Set $\varphi_{n,m} = \varphi_{X_{n,m}}$, $\sigma_{n,m}^2 = \mathbb{E}[X_{n,m}^2]$. Just like in the proofs of Theorems 9.3 and 9.7, it will be enough to show that

$$\prod_{m=1}^{n} \varphi_{n,m}(t) \to e^{-\frac{1}{2}\sigma^2 t^2}, \text{ for all } t \in \mathbb{R}.$$

We fix $t \neq 0$ and use Lemma 9.1 with $u_{n,m} = \varphi_{n,m}(t)$ and $w_{n,m} = 1 - \frac{1}{2}\sigma_{n,m}^2 t^2$ to conclude that

$$D_n(t) \leq M_n^{n-1} \sum_{m=1}^{n} \left| \varphi_{n,m}(t) - 1 + \tfrac{1}{2}\sigma_{n,m}^2 t^2 \right|,$$

where

$$D_n(t) = \left| \prod_{m=1}^{n} \varphi_{n,m}(t) - \prod_{m=1}^{n} (1 - \tfrac{1}{2}\sigma_{n,m}^2 t^2) \right|$$

and $M_n = 1 \vee \max_{1 \leq m \leq n} (\left| 1 - \frac{1}{2} t^2 \sigma_{n,m}^2 \right|)$. Assumption (4) in the statement implies that

$$\begin{aligned} \sigma_{n,m}^2 &= \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{m,n}| \geq \varepsilon\}}] + \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{m,n}| < \varepsilon\}}] \leq \varepsilon^2 + \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{m,n}| < \varepsilon\}}] \\ &\leq \varepsilon^2 + s_n(\varepsilon), \end{aligned}$$

and so $\sup_{1 \leq m \leq n} \sigma_{n,m}^2 \to 0$, as $n \to \infty$. Therefore, for n large enough, we have $\frac{1}{2} t^2 \sigma_{n,m}^2 \leq 2$ and $M_n = 1$.

According to Corollary 9.6 we now have (for large-enough n)

$$\begin{aligned} D_n(t) &\leq t^2 \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \min(t\,|X_{n,m}|\,, 1)] \\ &\leq t^2 \sum_{m=1}^{n} \left(\mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{n,m}| \geq \varepsilon\}}] + \mathbb{E}[t\,|X_{n,m}|^3\, \mathbf{1}_{\{|X_{n,m}| < \varepsilon\}}] \right) \\ &\leq t^2 s_n(\varepsilon) + t^3 \varepsilon \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{n,m}| < \varepsilon\}}] \leq t^2 s_n(\varepsilon) + 2t^3 \varepsilon \sigma^2. \end{aligned}$$

Therefore, $\limsup_n D_n(t) \leq 2t^3\varepsilon\sigma^2$, and so, since $\varepsilon > 0$ is arbitrary, we have $\lim_n D_n(t) = 0$.

Our last task is to remember that $\max_{1\leq m\leq n} \sigma_{n,m}^2 \to 0$, note that $\sum_{m=1}^m \sigma_{n,m}^2 \to \sigma^2$ (why?), and use Lemma 9.8 to conclude that

$$\prod_{m=1}^n (1 - \tfrac{1}{2}\sigma_{n,m}^2 t^2) - e^{-\frac{1}{2}\sigma^2 t^2}. \qquad \square$$

Problem 9.2. Show how the iid central limit theorem follows from the Lindeberg-Feller theorem.

Example 9.10 (Cycles in a random permutation)**.** Let $\Pi : \Omega \to S_n$ be a random element taking values in the set S_n of all permutations of the set $\{1, \dots, n\}$, i.e., the set of all bijections $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$. One usually considers the probability measure on Ω such that Π is uniformly distributed over S_n, i.e. $\mathbb{P}[\Pi = \pi] = \frac{1}{n!}$, for each $\pi \in S_n$. A random element in S_n whose distribution is uniform over S_n is called a **random permutation**.

Remember that each permutation $\pi \in S_n$ be decomposed into cycles; a **cycle** is a collection $(i_1 i_2 \dots i_k)$ in $\{1, \dots, n\}$ such that $\pi(i_l) = i_{l+1}$ for $l = 1, \dots, k-1$ and $\pi(i_k) = i_1$. For example, the permutation $\pi : \{1,2,3,4\} \to \{1,2,3,4\}$, given by $\pi(1) = 3$, $\pi(2) = 1$, $\pi(3) = 2$, $\pi(4) = 4$ has two cycles: (132) and (4). More precisely, start from $i_1 = 1$ and follow the sequence $i_{k+1} = \pi(i_k)$, until the first time you return to $i_k = 1$. Write these number in order $(i_1 i_2 \dots i_k)$ and pick $j_1 \in \{1, 2, \dots, n\} \setminus \{i_1, \dots, i_k\}$. If no such j_1 exists, π consist of a single cycle. If it does, we repeat the same procedure starting from j_1 to obtain another cycle $(j_1 j_2 \dots j_l)$, etc. In the end, we arrive at the decomposition

$$(i_1 i_2 \dots i_k)(j_1 j_2 \dots j_l) \dots$$

of π into cycles.

Let us first answer the following, warm-up, question: what is the probability $p(n,m)$ that 1 is a member of a cycle of length m? Equivalently, we can ask for the number $c(n,m)$ of permutations in which 1 is a member of a cycle of length m. The easiest way to solve this is to note that each such permutation corresponds to a choice of $(m-1)$ district numbers of $\{2, 3, \dots\}$ - these will serve as the remaining elements of the cycle containing 1. This can be done in $\binom{n-1}{m-1}$ ways. Furthermore, the $m-1$ elements to be in the same cycle with 1 can be ordered in $(m-1)!$ ways. Also, the remaining $n-m$ elements give rise to $(n-m)!$ distinct permutations. Therefore,

$$c(n,m) = \binom{n-1}{m-1}(m-1)!(n-m)! = (n-1)!, \text{ and so } p(n,m) = \tfrac{1}{n}.$$

This is a remarkable result - all cycle lengths are equally likely. Note, also, that 1 is not special in any way.

Our next goal is to say something about the number of cycles - a more difficult task. We start by describing a procedure for producing a random permutation by building it from cycles. The reader will easily convince his-/herself that the outcome is uniformly distributed over all permutations. We start with $n-1$ independent random variables $\xi_2, \dots, \xi_n$ such that ξ_i is uniformly distributed over the set $\{0, 1, 2, \dots, n-i+1\}$. Let the first cycle start from $X_1 = 1$. If $\xi_2 = 0$, then we declare (1) to be a full cycle and start building the next cycle from 2. If $\xi_2 \neq 0$, we pick the ξ_2-th smallest element - let us call it X_2 - from the set of remaining $n-1$ numbers to be the second element in the first cycle. After that, we close the cycle if $\xi_3 = 0$, or append the ξ_3-th smallest element - let's call it X_3 - in $\{1, 2, \dots, n\} \setminus \{X_1, X_2\}$ to the cycle. Once the cycle $(X_1\ X_2\ \dots\ X_k)$ is closed, we pick the smallest element in $\{1, 2, \dots, n\} \setminus \{X_1, X_2, \dots, X_k\}$ - let's call it X_{k+1} - and repeat the procedure starting from X_{k+1} and using $\xi_{k+1}, \dots, \xi_n$ as "sources of randomness".

Let us now define the random variables (we stress the dependence on n here) $Y_{n,1}, \dots, Y_{n,n}$ by $Y_{n,k} = \mathbf{1}_{\{\xi_k = 0\}}$. In words, $Y_{n,k}$ is in indicator of the event when a cycle ends right after the position k. It is clear that $Y_{n,1}, \dots, Y_{n,k}$ are independent (they are functions of independent variables $\xi_1, \dots, \xi_n$). Also, $p(n,k) = \mathbb{P}[Y_{n,k} = 1] = \frac{1}{n-k+1}$. The number of cycles C_n is the same as the number of closing parenthesis, so $C_n = \sum_{k=1}^n Y_{k,n}$. (Btw, can you derive the identity $p(n,m) = \frac{1}{n}$ by using random variables $Y_{n,1}, \dots, Y_{n,n}$?)

It is easy to compute

$$\begin{aligned}\mathbb{E}[C_n] &= \sum_{k=1}^n \mathbb{E}[Y_{n,k} = 1] = \sum_{k=1}^n \tfrac{1}{n-k+1} = 1 + \tfrac{1}{2} + \dots + \tfrac{1}{n} \\ &= \log(n) + \gamma + o(1),\end{aligned}$$

where $\gamma \approx 0.58$ is the Euler-Mascheroni constant, and $a_n = b_n + o(n)$ means that $|b_n - a_n| \to 0$, as $n \to \infty$.

If we want to know more about the variability of C_n, we can also compute its variance:

$$\begin{aligned}\mathrm{Var}[C_n] &= \sum_{k=1}^n \mathrm{Var}[Y_{n,k}] = \sum_{k=1}^n \left(\tfrac{1}{n-k+1} - \tfrac{1}{(n-k+1)^2}\right) \\ &= \log(n) + \gamma - \tfrac{\pi^2}{6} + o(1).\end{aligned}$$

The Lindeberg-Feller theorem will give us the precise asymptotic behavior of C_n. For $m = 1, \dots, n$, we define

$$X_{n,m} = \frac{Y_{n,m} - \mathbb{E}[Y_{n,m}]}{\sqrt{\log(n)}},$$

so that $X_{n,m}$, $m = 1, \ldots, n$ are independent and of mean 0. Furthermore, we have

$$\lim_n \sum_{m=1}^n \mathbb{E}[X_{n,m}^2] = \lim_n \frac{\log(n)+\gamma-\frac{\pi^2}{6}+o(1)}{\log(n)} = 1.$$

Finally, for $\varepsilon > 0$ and $\log(n) > 2/\varepsilon$, we have $\mathbb{P}[|X_{n,m}| > \varepsilon] = 0$, so

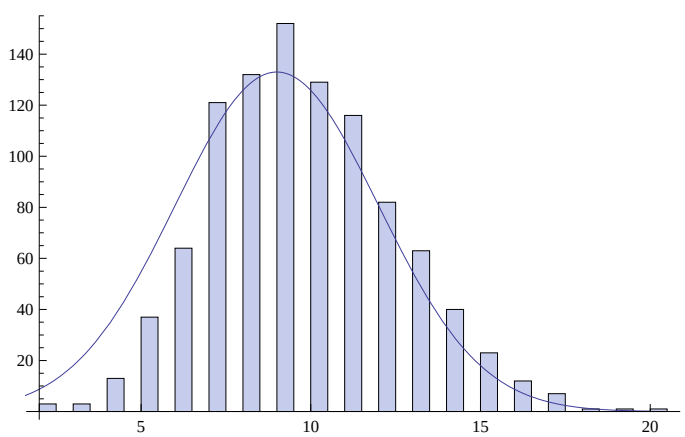
$$\sum_{m=1}^n \mathbb{E}[X_{n,m}^2 \mathbf{1}_{\{|X_{n,m}| \geq \varepsilon\}}] = 0.$$

Having checked that all the assumptions of the Lindeberg-Feller theorem are satisfied, we conclude that

$$\frac{C_n - \log(n)}{\sqrt{\log(n)}} \xrightarrow{\mathcal{D}} \chi, \text{ where } \chi \sim N(0,1).$$

It follows that (if we believe that the approximation is good) the number of cycles in a random permutation with $n = 8100$ is at most 18 with probability 99%.

How about variability? Here is histogram of the number of cycles from 1000 simulations for $n = 8100$, together with the appropriately-scaled density of the normal distribution with mean $\log(8100)$ and standard deviation $\sqrt{\log(8100)}$. The quality of approximation leaves something to be desired, but it seems to already work well in the tails: only 3 of 1000 had more than 17 cycles:



Additional Problems

Problem 9.3 (Lyapunov's theorem)**.** Let $\{X_n\}_{n\in\mathbb{N}}$ be an independent sequence, let $S_n = \sum_{m=1}^n X_m$, and let $\alpha_n = \sqrt{\mathrm{Var}[S_n]}$. Suppose that $\alpha_n > 0$ for all $n \in \mathbb{N}$ and that there exists a constant $\delta > 0$ such that

$$\lim_n \alpha_n^{-(2+\delta)} \sum_{m=1}^n \mathbb{E}[|X_m - \mathbb{E}[X_m]|^{2+\delta}] = 0.$$

Show that

$$\frac{S_n - \mathbb{E}[S_n]}{\alpha_n} \xrightarrow{\mathcal{D}} \chi, \text{ where } \chi \sim N(0,1).$$

Problem 9.4 (Self-normalized sums)**.** Let $\{X_n\}_{n\in\mathbb{N}}$ be iid random variables with $\mathbb{E}[X_1] = 0$, $\sigma = \sqrt{\mathbb{E}[X_1^2]} > 0$ and $\mathbb{P}[X_1 = 0] = 0$. Show that the sequence $\{Y_n\}_{n\in\mathbb{N}}$ given by

Hint: Use Slutsky's theorem (Problem 8.8)

$$Y_n = \frac{\sum_{k=1}^n X_k}{\sqrt{\sum_{k=1}^n X_k^2}},$$

converges in distribution, and identify its limit.