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, first¹:

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

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

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

$$\left| \prod_{k=1}^{n+1} u_k - \prod_{k=1}^{n+1} w_k \right| \leq \left| \prod_{k=1}^{n} u_k \right| |u_{n+1} - w_{n+1}| + \left| \prod_{k=1}^{n} w_k \right| |u_{n+1} - w_{n+1}|$$

$$\leq M^n |u_{n+1} - w_{n+1}| + M \times M^{n-1} \sum_{k=1}^{n} |u_k - w_k|$$

$$= M^{(n+1)-1} \sum_{k=1}^{n+1} |u_k - w_k|. \hspace{1cm} \square$$

**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 $u_k = 1 + \frac{z_k}{n}$ and $w_k = e^{z_n/n}$ for $k = 1, \ldots, n$, we get

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

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

$$L = \sup_{n \in \mathbb{N}} M_n^{n-1} < \infty.$$  \hspace{1cm} (9.3)

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

**Last Updated:** November 21, 2013
To estimate the last term in (9.2), we start with the Taylor expansion
\[ e^b = 1 + b + \sum_{k=2}^{\infty} \frac{b^k}{k!}, \]
which converges absolutely for all \( b \in \mathbb{C} \). Then, we use the fact that \( \frac{1}{k!} \leq 2^{-k+1} \), to obtain
\[
|e^b - 1 - b| \leq \sum_{k=2}^{\infty} \frac{|b|^k}{k!} \leq |b|^2 \sum_{k=2}^{\infty} 2^{-k+1} = |b|^2, \text{ for } |b| \leq 1. \quad (9.4)
\]
Since \( |z_n|/n \leq 1 \) for large-enough \( n \), it follows from (9.2), (9.4) and (9.3), that
\[
\limsup_n \left| (1 + \frac{z_n}{n})^n - e^{z_n} \right| \leq \limsup_n nL \left| \frac{z_n}{n} \right|^2 = 0.
\]

It remains to remember that \( e^{\bar{z}_n} \to e^z \) to finish the proof. \( \square \)

**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 \( \mu \) and the characteristic function \( \phi = \phi_\mu \) such that \( \phi'(0) \) exists. Then, \( c = -i\phi'(0) \) is a real number and
\[
\frac{1}{n} \sum_{k=1}^{n} X_k \to c \text{ in probability.}
\]

**Proof.** Since \( \phi(-s) = \overline{\phi(s)} \), we have
\[
\phi'(0) = \lim_{s \to 0} \frac{\phi(-s) - 1}{-s} = \lim_{s \to 0} \frac{\overline{\phi(s)} - 1}{s} = -\lim_{s \to 0} \frac{\phi(s) - 1}{s} = -\phi'(0).
\]

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

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

The iid property of \( \{X_n\}_{n \in \mathbb{N}} \) and the fact that \( \phi_{\alpha X}(t) = \phi_X(at) \) imply that
\[
\phi_{\frac{1}{n} S_n}(t) = \left( \phi\left( \frac{t}{n} \right) \right)^n = \left( 1 + \frac{z_n}{n} \right)^n,
\]
where \( z_n = n(\phi\left( \frac{t}{n} \right) - 1) \). By the assumption, we have \( \lim_{s \to 0} \frac{\phi(s) - 1}{s} = ic \), and so \( z_n \to t\phi'(0) \). Therefore, by Lemma 9.2 above, we have \( \phi_{\frac{1}{n} S_n}(t) \to e^{itc} \). \( \square \)

**Remark 9.4.**

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

2. \( X_1 \in L^1 \) implies \( \phi'(0) = \mathbb{E}[X_1] \), so that Theorem 9.3 covers the classical case. As we have seen in Problem 8.6, there are cases when \( \phi'(0) \) 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 \( \varphi_X \) is not differentiable at 0 and show that there is no constant \( c \) such that
\[
\frac{1}{n} S_n \xrightarrow{P} c,
\]
where \( S_n = \sum_{k=1}^n X_k \). Hint: What is the distribution of \( \frac{1}{n} S_n \)?

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
\[
\left| e^{i\xi} - \sum_{k=0}^n \frac{(i\xi)^k}{k!} \right| \leq \min\left( \frac{|\xi|^{n+1}}{(n+1)!}, 2 \frac{|\xi|^n}{n!} \right).
\]

Proof. If we write the remainder in the Taylor formula in the integral form (derived easily using integration by parts), we get
\[
e^{i\xi} - \sum_{k=0}^n \frac{(i\xi)^k}{k!} = R_n(\xi), \quad \text{where} \quad R_n(\xi) = i^{n+1} \int_0^\xi e^{iu}(\xi-u)^n \frac{du}{m}.
\]
The usual estimate of \( R_n \) gives:
\[
|R_n(\xi)| \leq \frac{1}{m!} \int_0^{|\xi|} (|\xi| - u)^n du = \frac{|\xi|^{n+1}}{(n+1)!}.
\]
We could also transform the expression for \( R_n \) by integrating it by parts:
\[
R_n(\xi) = \frac{i^{n+1}}{m!} \left( 1 - \frac{n}{7} \int_0^\xi e^{iu}(\xi-u)^{n-1} du \right)
= \frac{n}{(n-1)!} \left( \int_0^\xi (\xi-u)^{n-1} du - \int_0^\xi e^{iu}(\xi-u)^{n-1} du \right),
\]
since \( \xi^n = n \int_0^\xi (\xi-u)^{n-1} du. \) Therefore
\[
|R_n(\xi)| \leq \frac{1}{(n-1)!} \int_0^{|\xi|} (|\xi| - u)^{n-1} \left| e^{iu} - 1 \right| du
\leq \frac{2}{m!} \int_0^{|\xi|} n(|\xi| - u)^{n-1} du = \frac{2|\xi|^n}{m!}.
\]

\( \square \)

\( ^2 \) Remember, the density of the Cauchy distribution is given by \( f_X(x) = \frac{1}{\pi(1+x^2)} \), \( x \in \mathbb{R} \).
While the following result can be obtained as a direct consequence of twice-differentiability of the function $\varphi$ 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

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

Then

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

**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

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

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

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

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

**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 $\varphi$ 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

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

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

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

Corollary 9.6 guarantees that

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

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

$$ |\varphi(\frac{t}{\sqrt{n}}) - 1 + \frac{1}{2n}t^2| \leq \frac{t^2}{n} r(t/\sqrt{n}). $$
Lemma 9.1 with \( u_1 = \cdots = u_n = \varphi(\frac{r}{\sqrt{n}}) \) and \( w_1 = \cdots = w_n = (1 - \frac{t^2}{2n}) \) yields:
\[
\left| (\varphi(\frac{r}{\sqrt{n}}))^n - (1 - \frac{t^2}{2n})^n \right| \leq t^2 r(t/\sqrt{n}),
\]
for \( n \geq \frac{2}{t^2} \) (so that \( \max(|\varphi(\frac{r}{\sqrt{n}})|, |1 - \frac{t^2}{2n}|) \leq 1 \)). Since \( \lim_{n} r(t/\sqrt{n}) = 0 \), we have
\[
\lim_{n \to \infty} \left| (\varphi(\frac{r}{\sqrt{n}}))^n - (1 - \frac{t^2}{2n})^n \right| = 0,
\]
and (9.6) follows from the fact that \((1 - \frac{t^2}{2n})^n \to e^{-\frac{1}{2}t^2} \), for all \( t \). \qed

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
\[
\sum_{m=1}^{n} (\log(1 + c_{n,m}) - c_{n,m}) \to 0, \text{ as } n \to \infty. \quad (9.7)
\]
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 \( x \geq \log(1+x) \))
\[
|\log(1+x) - x| \leq x^2 \text{ for } x \geq -\frac{1}{2}.
\]
Since \( m_n < \frac{1}{2} \), we have \( |\log(1 + c_{n,m}) - c_{n,m}| \leq c_{n,m}^2 \) and so
\[
\sum_{m=1}^{n} (\log(1 + c_{n,m}) - c_{n,m}) \leq \sum_{m=1}^{n} |\log(1 + c_{n,m}) - c_{n,m}| \leq \sum_{m=1}^{n} c_{n,m}^2 \\
\leq m_n \sum_{m=1}^{n} |c_{n,m}| \to 0,
\]
because \( \sum_{n=1}^{\infty} |c_{n,m}| \) is bounded and \( m_n \to 0 \).

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

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

Then

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

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

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

where

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

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

\[
\sigma_{n,m}^2 = \mathbb{E}[X_{n,m}^2 \{ |X_{n,m}| \geq \epsilon \}] + \mathbb{E}[X_{n,m}^2 \{ |X_{n,m}| < \epsilon \}] \leq \epsilon^2 + \mathbb{E}[X_{n,m}^2 \{ |X_{n,m}| < \epsilon \}] \leq \epsilon^2 + s_n(\epsilon),
\]

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 \( t^2 \sigma_{n,m}^2 \leq 2 \) and \( M_n = 1 \).

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

\[
D_n(t) \leq t^2 \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \min(t \mid |X_{n,m}^2|, 1)]
\]

\[
\leq t^2 \sum_{m=1}^{n} \left( \mathbb{E}[X_{n,m}^2 \{ |X_{n,m}| \geq \epsilon \}] + \mathbb{E}[t \mid |X_{n,m}|^3 \{ |X_{n,m}| < \epsilon \}] \right)
\]

\[
\leq t^2 s_n(\epsilon) + t^3 \epsilon \sum_{m=1}^{n} \mathbb{E}[X_{n,m}^2 \{ |X_{n,m}| < \epsilon \}] \leq t^2 s_n(\epsilon) + 2 t^3 \epsilon \sigma^2.
\]

*Last Updated: November 21, 2013*
Therefore, \( \limsup_n D_n(t) \leq 2P \epsilon \sigma^2 \), and so, since \( \epsilon > 0 \) is arbitrary, we have \( \lim_n D_n(t) = 0 \).

Our last task is to remember that \( \max_{1 \leq m \leq n} \sigma^2_{n,m} \to 0 \), note that \( \sum_{m=1}^{n} \sigma^2_{n,m} \to \sigma^2 \) (why?), and use Lemma 9.8 to conclude that
\[
\prod_{m=1}^{n} \left( 1 - \frac{1}{2} \sigma^2_{n,m} t^2 \right) - e^{-\frac{1}{2} \sigma^2 t^2}.
\]

\[\square\]

**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 \( \{1, \ldots, n\} \), i.e., the set of all bijections \( \pi : \{1, \ldots, n\} \to \{1, \ldots, n\} \).

One usually considers the probability measure on \( \Omega \) such that \( \Pi \) 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 \( (i_1 i_2 \ldots i_k) \) in \( \{1, \ldots, n\} \) such that \( \pi(i_l) = i_{l+1} \) for \( l = 1, \ldots, k-1 \) 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 \( (i_1 i_2 \ldots i_k) \) and pick \( j_1 \in \{1,2,\ldots,n\} \setminus \{i_1,\ldots,i_k\} \). If no such \( j_1 \) exists, \( \pi \) consist of a single cycle. If it does, we repeat the same procedure starting from \( j_1 \) to obtain another cycle \( (j_1 j_2 \ldots j_l) \), etc. In the end, we arrive at the decomposition

\[
(i_1 i_2 \ldots i_k)(j_1 j_2 \ldots j_l)\ldots
\]

of \( \pi \) 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 \( \{2,3,\ldots\} \) - 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,
\[
c(n, m) = \binom{n-1}{m-1} (m-1)! (n-m)! = (n-1)!, \text{ and so } p(n, m) = \frac{1}{n!}.
\]
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 \( \xi_2, \ldots, \xi_n \) such that \( \xi_i \) is uniformly distributed over the set \( \{0, 1, 2, \ldots, n - i + 1\} \). 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 \( \xi_2 \)-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 \( \xi_3 \)-th smallest element - let’s call it \( X_3 \) - in \( \{1, 2, \ldots, n\} \setminus \{X_1, X_2\} \) to the cycle. Once the cycle \( (X_1 X_2 \ldots X_k) \) is closed, we pick the smallest element in \( \{1, 2, \ldots, n\} \setminus \{X_1, X_2, \ldots, X_k\} \) - let’s call it \( X_{k+1} \) - and repeat the procedure starting from \( X_{k+1} \) and using \( \xi_{k+1}, \ldots, \xi_n \) as “sources of randomness”.

Let us now define the random variables (we stress the dependence on \( n \) here) \( Y_{n,1}, \ldots, Y_{n,n} \) by \( Y_{n,k} = 1_{\{\xi_k = 0\}} \). In words, \( Y_{n,k} \) is in indicator of the event when a cycle ends right after the position \( k \). It is clear that \( Y_{n,1}, \ldots, Y_{n,k} \) are independent (they are functions of independent variables \( \xi_1, \ldots, \xi_n \)). Also, \( p(n, k) = P[Y_{n,k} = 1] = \frac{1}{n-k+1} \). The number of cycles \( C_n \) is the same as the number of closing parentheses, 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 \( Y_{n,1}, \ldots, Y_{n,n} \)?)

It is easy to compute

\[
E[C_n] = \sum_{k=1}^{n} E[Y_{n,k} = 1] = \sum_{k=1}^{n} \frac{1}{n-k+1} = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \log(n) + \gamma + o(1),
\]

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:

\[
\text{Var}[C_n] = \sum_{k=1}^{n} \text{Var}[Y_{n,k}] = \sum_{k=1}^{n} \left( \frac{1}{(n-k+1)^2} \right) = \log(n) + \gamma - \frac{\pi^2}{6} + o(1).
\]

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

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

Last Updated: November 21, 2013
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

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

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

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

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:

![Histogram of number of cycles](image)

**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 $\alpha_n = \sqrt{\text{Var}[S_n]}$. Suppose that $\alpha_n > 0$ for all $n \in \mathbb{N}$ and that there exists a constant $\delta > 0$ such that

$$\lim_n \alpha_n^{-2(2+\delta)} \sum_{m=1}^{n} \mathbb{E}[|X_m - \mathbb{E}[X_m]|^{2+\delta}] = 0.$$
Show that
\[ \frac{S_n - \mathbb{E}[S_n]}{\alpha_n} \overset{D}{\to} \chi, \quad \text{where} \quad \chi \sim N(0,1). \]

**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
\[
Y_n = \frac{\sum_{k=1}^{n} X_k}{\sqrt{\sum_{k=1}^{n} X_k^2}}
\]
converges in distribution, and identify its limit.

*Hint:* Use Slutsky’s theorem (Problem 8.8)