Math 639: Lecture 4

Convergence of random series and large deviations

Bob Hough

February 2, 2017

Renewal theory

Let $X_1, X_2, \ldots$ be i.i.d. with $0 < X_i < \infty$. Let $T_n = X_1 + \cdots + X_n$ and

$$N_t = \sup\{n : T_n \leq t\}.$$

Given a sequence of events which happen in succession with waiting time X_n on the nth event, we think of N_t as the number of events which have happened up to time t.

Theorem

If $\mathrm{E}[X_1] = \mu \leq \infty$, *then as* $t \to \infty$,

$$\frac{N_t}{t} \to \frac{1}{\mu} \text{ a.s.}.$$

Proof.

Since $T(N_t) \leq t < T(N_t + 1)$, dividing through by N_t gives

$$\frac{T(N_t)}{N_t} \leq \frac{t}{N_t} \leq \frac{T(N_t + 1)}{N_t + 1} \frac{N_t + 1}{N_t}.$$

We have $N_t \to \infty$ a.s.. Hence, by the strong law,

$$\frac{T_{N_t}}{N_t} \to \mu, \qquad \frac{N_t + 1}{N_t} \to 1.$$

□

Empirical distribution functions

Let $X_1, X_2, \ldots$ be i.i.d. with distribution F and let

$$F_n(x) = \frac{1}{n} \sum_{m=1}^{n} \mathbf{1}_{(X_m \leq x)}.$$

Theorem (Glivenko-Cantelli Theorem)

As $n \to \infty$,

$$\sup_x |F_n(x) - F(x)| \to 0 \text{ a.s.}.$$

Empirical distribution functions

Proof.

Note that F is increasing, but can have jumps.

- For $k = 1, 2, \ldots$, and $1 \le j \le k-1$, define $x_{j,k} = \inf\{x : F(x) \ge \frac{j}{k}\}$. Set $x_{0,k} = -\infty$, $x_{k,k} = \infty$.
- Write $F(x-) = \lim_{y \uparrow x} F(y)$.
- Since each of $F_n(x_{j,k}-)$ and $F_n(x_{j,k})$ converges by the strong law, and $F_n(x_{j,k}-) - F_n(x_{j-1,k}) \le \frac{1}{k}$, the uniform convergence follows.

□

Entropy

- Let $X_1, X_2, \ldots$ be i.i.d., taking values in $\{1, 2, \ldots, r\}$ with all possibilities of positive probability. Set $\mathrm{Prob}(X_i = k) = p(k) > 0$.
- Let $\pi_n(\omega) = p(X_1(\omega))p(X_2(\omega))\ldots p(X_n(\omega))$. By the strong law, a.s.

$$-\frac{1}{n}\log \pi_n \to H \equiv -\sum_{k=1}^{r} p(k)\log p(k).$$

The constant H is called the *entropy*.

The tail σ-algebra

Definition

Let $X_1, X_2, \ldots$ be a sequence of random variables. Their *tail* σ-algebra is

$$\mathscr{T} = \bigcap_{n=1}^{\infty} \sigma\left(X_n, X_{n+1}, \ldots\right).$$

The tail σ-algebra

Example

- If $\{B_n\}$ is a sequence from the Borel σ-algebra $\mathscr{B}$, then $\{X_n \in B_n \ i.o.\} \in \mathscr{T}$.
- Let $S_n = X_1 + X_2 + \cdots + X_n$. We have
  - $\{\lim_{n\to\infty} S_n \text{ exists}\} \in \mathscr{T}$,
  - $\{\limsup_{n\to\infty} S_n > 0\} \notin \mathscr{T}$
  - $\{\limsup_{n\to\infty} \frac{S_n}{c_n} > x\} \in \mathscr{T}$ if $c_n \to \infty$.

Kolmogorov's 0-1 law

Theorem

If $X_1, X_2, \ldots$ *are independent and* $A \in \mathscr{T}$ *then* $\mathrm{Prob}(A) = 0$ *or* $\mathrm{Prob}(A) = 1$.

Kolmogorov's 0-1 law

Proof.

We show that A is independent of itself, so that $\mathrm{Prob}(A) = \mathrm{Prob}(A \cap A) = \mathrm{Prob}(A)^2$.

- Observe that for each k, $\sigma(X_1, ..., X_k)$ and $\sigma(X_{k+1}, X_{k+2}, ...)$ are independent. This follows, since $\sigma(X_{k+1}, X_{k+2}, ...)$ is generated by $\sigma(X_{k+1}, ..., X_{k+m})$ for $m = 1, 2, 3, ...$, whose union forms a π-system.
- Since $\mathscr{T} \subset \sigma(X_{k+1}, X_{k+1}, ...)$, $\mathscr{T}$ is independent of $\sigma(X_1, X_2, ..., X_k)$ for each k, and hence of $\sigma(X_1, X_2, ...)$.

□

Kolmogorov's 0-1 law

Example

If $A_1, A_2, \ldots$ are independent then

- $\text{Prob}(A_n \text{ i.o.})$ is 0 or 1
- $\text{Prob}(\lim_{n\to\infty} S_n \text{ exists})$ is 0 or 1.

Kolmogorov's maximal inequality

Theorem (Kolmogorov's maximal inequality)

Suppose $X_1, ..., X_n$ *are independent with* $\mathrm{E}[X_i] = 0$ *and* $\mathrm{Var}(X_i) < \infty$. *If* $S_k = X_1 + \cdots + X_k$, *then*

$$\mathrm{Prob}\left(\max_{1 \le k \le n} |S_k| \ge x\right) \le x^{-2}\,\mathrm{Var}(S_n).$$

Kolmogorov's maximal inequality

Proof.

- Let $A_k = \{|S_k| \geq x\} \setminus \bigcup_{j=1}^{k-1} \{|S_j| \geq x\}$ be those trials for which the sum first exceeds x at step k.
- We have

$$\begin{aligned} \mathrm{E}\left[S_n^2\right] &\geq \sum_{k=1}^n \int_{A_k} S_n^2 dP = \sum_{k=1}^n \int_{A_k} (S_k + (S_n - S_k))^2 dP \\ &\geq \sum_{k=1}^n \int_{A_k} S_k^2 dP + \sum_{k=1}^n \int 2 S_k \mathbf{1}_{A_k} (S_n - S_k) dP. \end{aligned}$$

- Since $S_k \mathbf{1}_{A_k} \in \sigma(X_1, ..., X_k)$, it is independent of $S_n - S_k$, so that the second sum above is 0. Since $|S_k| \geq x$ on A_k, it follows that $\mathrm{E}\left[S_n^2\right] \geq x^2 \,\mathrm{Prob}\left(\max_{1 \leq k \leq n} |S_k| \geq x\right)$.

$\square$

Convergence of series

Theorem

Let $X_1, X_2, \ldots$ *be independent, have* $\mathrm{E}[X_i] = 0$ *and*

$$\sum_{n=1}^{\infty} \mathrm{Var}(X_n) < \infty.$$

With probability 1, $\sum_{n=1}^{\infty} X_n$ *converges.*

Convergence of series

Proof.

- Let $S_N = \sum_{n=1}^N X_n$.
- By Kolmogorov's maximal theorem,

$$\mathrm{Prob}\left(\max_{M \leq m \leq N} |S_m - S_N| > \epsilon\right) \leq \epsilon^{-2} \sum_{n=M+1}^{N} \mathrm{Var}(X_n),$$

so

$$\mathrm{Prob}\left(\sup_{m \geq M} |S_m - S_M| > \epsilon\right) \leq \epsilon^{-2} \sum_{n=M+1}^{\infty} \mathrm{Var}(X_n).$$

□

Convergence of series

Proof.

- Let $w_M = \sup_{m,n \geq M} |S_m - S_n|$. We have

$$\text{Prob}\left(w_M > 2\epsilon\right) \leq \text{Prob}\left(\sup_{m \geq M} |S_m - S_M| > \epsilon\right) \to 0$$

  as $M \to \infty$, so $w_M \downarrow 0$ a.s..
- Hence $\sum X_n$ is a.s. Cauchy, hence convergent.

□

Convergence of series

Example

Let $X_1, X_2, \ldots$ be i.i.d., taking values ± 1 with probability $\frac{1}{2}$. The series

$$\sum_{n=1}^{\infty} \frac{X_n}{n^\sigma}$$

converges a.s. if $\sigma > \frac{1}{2}$ and diverges a.s. if $\sigma \leq \frac{1}{2}$.

Kolmogorov's three-series theorem

Theorem

Let $X_1, X_2, \ldots$ be independent. Let $A > 0$ and let $Y_i = X_i \mathbf{1}_{(|X_i| \leq A)}$. In order that $\sum_{n=1}^{\infty} X_n$ converge a.s. it is necessary and sufficient that

1. $\sum_{n=1}^{\infty} \mathrm{Prob}(|X_n| > A) < \infty$
2. $\sum_{n=1}^{\infty} \mathrm{E}[Y_n]$ *converges*
3. $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n) < \infty$.

Kolmogorov's three-series theorem

Proof.

- We postpone the proof of necessity.
- To prove that the condition is sufficient, note that item 1 implies that $X_n \neq Y_n$ finitely often with probability 1.
- The a.s. convergence of $\sum_n Y_n$ is now guaranteed by the previous theorem.

□

Kronecker's lemma

Theorem

If $a_n \geq 0$ is an increasing sequence, $a_n \uparrow \infty$ and $\sum_{n=1}^{\infty} \frac{x_n}{a_n}$ converges then

$$\frac{1}{a_n} \sum_{m=1}^{n} x_m \to 0.$$

Kronecker's lemma

Proof.

Let $a_0 = b_0 = 0$ and $b_m = \sum_{k=1}^m \frac{x_k}{a_k}$. By summation by parts,

$$\frac{1}{a_n}\sum_{m=1}^n x_m = \frac{1}{a_n}\left\{\sum_{m=1}^n a_m(b_m - b_{m-1})\right\}$$
$$= b_n - \sum_{m=1}^n \frac{(a_m - a_{m-1})}{a_n} b_{m-1}.$$

Since b_n tends to a limit, and $a_n \uparrow \infty$,

$$\lim_{n\to\infty} b_n - \sum_{m=1}^n \frac{(a_m - a_{m-1})}{a_n} b_{m-1} = 0.$$

□

Rates of convergence

The following is a cheap version of the law of the iterated logarithm.

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d. random variables satisfying* $\mathrm{E}[X_i] = 0$ *and* $\mathrm{E}[X_i^2] = \sigma^2 < \infty$. *Let* $S_n = X_1 + \cdots + X_n$. *Then for* $\epsilon > 0$

$$\frac{S_n}{n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\epsilon}} \to 0 \qquad a.s..$$

Rates of convergence

Proof.

Let $a_n = n^{\frac{1}{2}}(\log n)^{\frac{1}{2}+\epsilon}$ for $n \geq 2$ and $a_1 = 1$. Then

$$\sum_{n=1}^{\infty} \mathrm{Var}\left(\frac{X_n}{a_n}\right) = \sigma^2 \left(1 + \sum_{n=2}^{\infty} \frac{1}{n(\log n)^{1+2\epsilon}}\right) < \infty,$$

so $\sum_{n=1}^{\infty} \frac{X_n}{a_n}$ converges a.s.. The claim now follows from Kronecker's lemma. $\square$

Random Dirichlet series

Example

- Form a 'random multiplicative function' by setting $X(1) = 1$, choosing $X(p)$, p prime to be i.i.d. ± 1 with equal probability, and declaring for all m, n, $X(mn) = X(m)X(n)$.
- For $\Re(s) > 1$, the Dirichlet series

$$L(s, X) = \sum_{n=1}^{\infty} \frac{X(n)}{n^s} = \prod_p \left(1 - \frac{X(p)}{p^s}\right)^{-1}$$

  converges absolutely and has an absolutely convergent Euler product.
- With probability 1, $\log L(s, X)$ has a holomorphic continuation to $\Re(s) > \frac{1}{2}$, so $L(s, X) \neq 0$ there.

Random Dirichlet series

Example

To check the last statement, write

$$\log L(s, X) = \sum_p \frac{X(p)}{p^s} + \text{absolutely convergent in } \Re(s) > \frac{1}{2}.$$

$$\sum_p \frac{X(p)}{p^s} = \int_0^\infty \frac{1}{x^s} d\left(\sum_{p \leq x} X(p)\right)$$
$$= s \int_0^\infty \frac{\sum_{p \leq x} X(p)}{x^{s+1}} dx.$$

With probability 1, for any $\epsilon > 0$ the numerator is $O_\epsilon(X^{\frac{1}{2}+\epsilon})$, so that the integral converges absolutely in $\Re(s) > \frac{1}{2}$, giving the holomorphic extension.

Rates of convergence

Theorem (Marcinkiewicz, Zygmund)

Let $X_1, X_2, \ldots$ be i.i.d. with $\mathrm{E}[X_1] = 0$ and $\mathrm{E}[|X_1|^p] < \infty$ where $1 < p < 2$. If $S_n = X_1 + \cdots + X_n$ then $\frac{S_n}{n^{\frac{1}{p}}} \to 0$ a.s.

Rates of convergence

Proof.

- Let $Y_k = X_k \mathbf{1}_{\left(|X_k| \leq k^{\frac{1}{p}}\right)}$ and $T_n = Y_1 + \cdots + Y_n$.
- We have

$$\sum_{k=1}^{\infty} \mathrm{Prob}(Y_k \neq X_k) = \sum_{k=1}^{\infty} \mathrm{Prob}\left(|X_k|^p > k\right) \leq \mathrm{E}\left[|X_k|^p\right] < \infty.$$

Thus $\mathrm{Prob}(Y_k \neq X_k \text{ i.o.}) = 0$ by Borel-Cantelli.

□

Rates of convergence

Proof.

- Calculate

$$\sum_{m=1}^{\infty} \mathrm{Var}\left(\frac{Y_m}{m^{\frac{1}{p}}}\right) \leq \sum_{m=1}^{\infty} \mathrm{E}\left[\frac{Y_m^2}{m^{\frac{2}{p}}}\right]$$

$$\leq \sum_{m=1}^{\infty} \int_0^{m^{\frac{1}{p}}} \frac{2y}{m^{\frac{2}{p}}} \mathrm{Prob}\left(|X_1| > y\right) dy$$

$$= \int_0^{\infty} \sum_{m > y^p} \frac{2y}{m^{\frac{2}{p}}} \mathrm{Prob}(|X_1| > y) dy$$

$$\ll \int_0^{\infty} y^{p-1} \mathrm{Prob}(|X_1| > y) dy = \mathrm{E}[|X_1|^p] < \infty.$$

□

Rates of convergence

Proof.

- Applying the theorem on convergence of series and Kronecker's lemma,

$$n^{-\frac{1}{p}} \sum_{k=1}^{n} (Y_m - \mathrm{E}[Y_m]) \to 0, \ a.s.$$

- It remains to verify that $n^{-\frac{1}{p}} \sum_{m=1}^{n} \mathrm{E}[Y_m] \to 0$. To check this, write $\mathrm{E}[Y_m] = -\mathrm{E}\left[X_1 \cdot \mathbf{1}\left(|X_1| > m^{\frac{1}{p}}\right)\right]$, so $|E[Y_m]|$ is bounded by

$$\mathrm{E}\left[|X_1| \cdot \mathbf{1}\left(|X_1| > m^{\frac{1}{p}}\right)\right] \leq m^{-1+\frac{1}{p}} \mathrm{E}\left[|X_1|^p \cdot \mathbf{1}\left(|X_1| > m^{\frac{1}{p}}\right)\right].$$

  Since $\sum_{m \leq n} m^{-1+\frac{1}{p}} \ll n^{\frac{1}{p}}$ and the expectation tends to 0 as $m \to \infty$, the claim follows.

□

Infinite mean

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d. with* $\mathrm{E}[|X_1|] = \infty$ *and let* $S_n = X_1 + \cdots + X_n$*. Let* a_n *be a sequence of positive numbers with* $\frac{a_n}{n}$ *increasing. Then* $\limsup_{n \to \infty} \frac{|S_n|}{a_n} = 0$ *or* ∞ *according as* $\sum_n \mathrm{Prob}(|X_1| \geq a_n) < \infty$ *or* $= \infty$*.*

Infinite mean

Proof.

First suppose $\sum_n \text{Prob}(|X_1| \geq a_n) = \infty$.

- Since $\frac{a_n}{n}$ is increasing

$$\sum_{n=1}^{\infty} \text{Prob}(|X_1| \geq ka_n) \geq \sum_{n=1}^{\infty} \text{Prob}(|X_1| \geq a_{kn}) \geq \frac{1}{k} \sum_{m=k}^{\infty} \text{Prob}(|X_1| \geq a_m) = \infty.$$

  In particular, $\limsup_{n\to\infty} \frac{|X_n|}{a_n} = \infty$ with probability 1 by Borel-Cantelli.
- Since $\max(|S_{n-1}|, |S_n|) \geq \frac{|X_n|}{2}$, the claim follows.

□

Infinite mean

Proof.

Now suppose $\sum_n \mathrm{Prob}(|X_1| \geq a_n) < \infty$.

- Define $Y_n = X_n \mathbf{1}(|X_n| < a_n)$. Since $X_n \neq Y_n$ finitely often a.s., the proof consists in checking that $\sum \mathrm{Var}\left(\frac{Y_n}{a_n}\right) < \infty$ and $\frac{1}{a_n} \sum_{m=1}^n \mathrm{E}[Y_m] \to 0$.
- Calculate

$$\sum_{n=1}^{\infty} \mathrm{Var}\left(\frac{Y_n}{a_n}\right) \leq \sum_{n=1}^{\infty} \frac{\mathrm{E}[Y_n^2]}{a_n^2}$$
$$= \sum_{m=1}^{\infty} \int_{[a_{m-1}, a_m)} y^2 dF(y) \sum_{n=m}^{\infty} a_n^{-2}.$$

□

Infinite mean

Proof.

Now suppose $\sum_n \mathrm{Prob}(|X_1| \geq a_n) < \infty$.

- Since $a_n \geq \frac{n a_m}{m}$, $\sum_{n=m}^\infty a_n^{-2} \leq \frac{m^2}{a_m^2} \sum_{n=m}^\infty n^{-2} \ll \frac{m}{a_m^2}$.
- Thus

$$\sum_{m=1}^\infty \mathrm{Var}\left(\frac{Y_n}{a_n}\right) \ll \sum_{m=1}^\infty m \,\mathrm{Prob}(a_{m-1} \leq |X_i| < a_m)$$
$$= \sum_{m=1}^\infty \mathrm{Prob}(|X_1| > a_{m-1}) < \infty.$$

□

Infinite mean

Proof.

- To prove the mean condition, first, since $\mathrm{E}[|X_i|] = \infty$ and $\sum_{n=1}^{\infty} \mathrm{Prob}(|X_i| > a_n) < \infty$, we have $\frac{a_n}{n} \uparrow \infty$.
- Bound

$$\left|\frac{1}{a_n}\sum_{m=1}^{n} \mathrm{E}[Y_m]\right| \leq \frac{1}{a_n}\sum_{m=1}^{n} \mathrm{E}\left[|X_m| \cdot \mathbf{1}(|X_m| < a_m)\right]$$
$$\leq \frac{na_N}{a_n} + \frac{n}{a_n}\mathrm{E}\left[|X_1| \cdot \mathbf{1}\left(a_N \leq |X_1| \leq a_n\right)\right].$$

- If N grows sufficiently slowly, the first term tends to 0.

□

Infinite mean

Proof.

- Bound $\frac{n}{a_n}\mathrm{E}\left[|X_1|\cdot\mathbf{1}\left(a_N \leq |X_1| \leq a_n\right)\right]$ by

$$\sum_{m=N+1}^{n}\frac{m}{a_m}\mathrm{E}\left[|X_1|\cdot\mathbf{1}(a_{m-1}\leq |X_1| < a_m)\right]$$
$$\leq \sum_{m=N+1}^{\infty} m\,\mathrm{Prob}(a_{m-1}\leq |X_1| < a_m).$$

Since

$$\sum_{m=1}^{\infty} m\,\mathrm{Prob}(a_{m-1}\leq |X_1| < a_m) = \sum_{n=1}^{\infty}\mathrm{Prob}(|X_1|\geq a_{n-1}) < \infty,$$

the latter tends to 0 as $N\to\infty$.

□

Large deviations

Let $X_1, X_2, \ldots$ be i.i.d., $\mathrm{E}[X_1] = \mu$, $|\mu| < \infty$. We are interested in the tail probability

$$\mathrm{Prob}(S_n > na)$$

for $a > \mu$. Define $\pi_n = \mathrm{Prob}(S_n > na)$ and $\gamma_n = \log \pi_n$.

Lemma

For $m, n \geq 1$,

$$\pi_{m+n} \geq \pi_m \pi_n.$$

Proof.

This follows from the independence, since

$$\pi_{m+n} \geq \mathrm{Prob}\left(S_m \geq ma, S_{n+m} - S_m \geq na\right) = \pi_m \pi_n.$$

□

Large deviations

Lemma

Let a_n be a sequence satisfying $a_{m+n} \geq a_m + a_n$. Then

$$\lim_{n\to\infty} \frac{a_n}{n} \to \sup_n \frac{a_n}{n}.$$

Proof.

- Since $\limsup \frac{a_n}{n} \leq \sup \frac{a_n}{n}$ it suffices to check $\liminf \frac{a_n}{n} \geq \frac{a_n}{n}$.
- Given $n > m$, write $n = km + \ell$, $0 \leq \ell < m$. We have

$$\frac{a_n}{n} \geq \left(\frac{km}{km+\ell}\right)\frac{a_m}{m} + \frac{a_\ell}{n}.$$

Letting $n \to \infty$ proves the claim.

□

Large deviations

Define $\gamma(a) = \lim_{n\to\infty} \frac{1}{n} \log \mathrm{Prob}\left(S_n \geq na\right)$. This exists by the previous lemma. Furthermore,

$$\mathrm{Prob}(S_n \geq na) \leq e^{n\gamma(a)}.$$

Moment generating function

Definition

The *moment generating function* for a random variable X is $\phi(\theta) = \mathrm{E}[e^{\theta X}]$.

Moment generating function

We assume that $\phi(\theta) < \infty$ for some $\theta > 0$. By Markov's inequality,

$$e^{\theta n a} \operatorname{Prob}(S_n \geq na) \leq \mathrm{E}[\exp(\theta S_n)] = \phi(\theta)^n,$$
$$\operatorname{Prob}(S_n \geq na) \leq \exp\left(-n(\theta a - \log \phi(\theta))\right).$$

Let $\theta_+ = \sup\{\theta : \phi(\theta) < \infty\}$.

Lemma

$\log \phi(\theta)$ *is continuous at 0, differentiable on* $(0, \theta_+)$ *and satisfies* $\lim_{\theta \downarrow 0} \frac{\phi'(\theta)}{\phi(\theta)} = \mathrm{E}[X]$.

Moment generating function

Proof.

Each of the statements follows by dominated convergence.

- For instance, to prove the continuity at 0, for $0 < \theta < \theta_0 < \theta_+$ use $e^{\theta x} \leq 1 + e^{\theta_0 x}$ to take the limit as $\theta \downarrow 0$.
- We proved in Lecture 2 that it's possible to differentiate under the expectation, from which the remaining two claims follow.

□

Weighted distribution

To prove the lower bound, consider the distribution function

$$F_\lambda(x) = \frac{1}{\phi(\lambda)} \int_{-\infty}^{x} e^{\lambda y} dF(y).$$

Note that this distribution has mean

$$\frac{1}{\phi(\lambda)} \int_{-\infty}^{\infty} y e^{\lambda y} dF(y) = \frac{\phi'(\lambda)}{\phi(\lambda)}.$$

Lemma

We have $\frac{dF^n}{dF_\lambda^n} = e^{-\lambda x} \phi(\lambda)^n$.

Weighted distribution

Proof.

We check this by induction. For $n = 1$ the claim holds by definition. Write

$$
\begin{aligned}
F^n(z) = F^{n-1} * F(z) &= \int_{-\infty}^{\infty} dF^{n-1}(x) \int_{-\infty}^{z-x} dF(y) \\
&= \phi(\lambda)^n \int_{-\infty}^{\infty} dF_\lambda^{n-1}(x) \int_{-\infty}^{\infty} \mathbf{1}_{(x+y\leq z)} e^{-\lambda(x+y)} dF_\lambda(y) \\
&= \phi(\lambda)^n \,\mathrm{E}\left(\mathbf{1}_{(S_{n-1}^\lambda + X^\lambda)\leq z} e^{-\lambda(S_{n-1}^\lambda + X_n^\lambda)} \right) \\
&= \phi(\lambda)^n \int_{-\infty}^{z} e^{-\lambda u} dF_\lambda^n(u).
\end{aligned}
$$

□

Tail probability

Suppose that the distribution of X is not a point mass at μ. It follows that $\frac{\phi'(\theta)}{\phi(\theta)}$ is strictly increasing by convexity. If $a > \mu$ then there is at most one solution to the 'saddle point' equation

$$a = \frac{\phi'(\theta_a)}{\phi(\theta_a)}.$$

Theorem

Suppose there is $\theta_a \in (0, \theta_+)$ such that $a = \frac{\phi'(\theta_a)}{\phi(\theta_a)}$. Then, as $n \to \infty$,

$$\frac{1}{n} \log \mathrm{Prob}(S_n \geq na) \to -a\theta_a + \log \phi(\theta_a).$$

Tail probability

Proof.

The upper bound in the limit follows from taking $\theta = \theta_a$ in the inequality

$$\text{Prob}(S_n \geq na) \leq \exp(-n(a\theta - \log \phi(\theta))).$$

□

Tail probability

Proof.

To prove the lower bound, use the weighted distribution F_λ,

$$\begin{aligned}\mathrm{Prob}(S_n \geq na) &\geq \int_{na}^{nb} e^{-\lambda x}\phi(\lambda)^n dF_\lambda^n(x) \\ &\geq \phi(\lambda)^n e^{-\lambda nb}(F_\lambda^n(nb) - F_\lambda^n(na)).\end{aligned}$$

- Choose λ such that $a < \frac{\phi'(\lambda)}{\phi(\lambda)} < b$.
- By the weak law of large numbers, $F_\lambda^n(nb) - F_\lambda^n(na) \to 1$ as $n \to \infty$. Letting $b \downarrow a$ proves the claim.

□

Examples

Example (Normal distribution)

The standard normal distribution has exponential generating function

$$\begin{aligned}\phi(\theta) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\theta x} \exp\left(-\frac{x^2}{2}\right) dx \\ &= \frac{e^{\frac{\theta^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\theta)^2}{2}} dx = e^{\frac{\theta^2}{2}}.\end{aligned}$$

Hence $\theta_a = a$ and $\gamma_a = -a\theta_a + \log \phi(\theta_a) = -\frac{a^2}{2}$.

Examples

Example (Exponential distribution)

The exponential distribution with parameter 1 has exponential generating function

$$\phi(\theta) = \int_0^\infty e^{\theta x} e^{-x} dx = \frac{1}{1-\theta}, \qquad \theta < 1.$$

Hence $\frac{\phi'}{\phi}(\theta) = \frac{1}{1-\theta}$, so $\theta_a = 1 - \frac{1}{a}$ and

$$\gamma(a) = -a\theta_a - \log(1-\theta_a) = -a + 1 + \log a.$$

Characteristic functions

Definition

The *characteristic function* of random variable X is

$$\phi(t) = \mathrm{E}\left[e^{itX}\right].$$

For real valued random variables, the characteristic function exists for all real t, which gives the characteristic function an advantage over the moment generating function.

Note the following easy properties of characteristic functions.

Theorem

The characteristic function $\phi(t)$ of X satisfies

- $\phi(0) = 1$.
- $\phi(-t) = \overline{\phi(t)}$.
- $|\phi(t)| \leq 1$, *with equality if and only if* $t = 0$ *or* $\mathrm{supp}(X) \subset \frac{2\pi}{t}\mathbb{Z} + c$.
- $|\phi(t+h) - \phi(t)| \leq \mathrm{E}[|e^{ihX} - 1|]$.
- $\mathrm{E}\left[e^{it(aX+b)}\right] = e^{itb}\phi(at)$.

Proof.

- The first two items are immediate.
- For the third, if $t \neq 0$ and $\mathrm{supp}(X) \in \frac{2\pi}{t}\mathbb{Z} + c$ then $e^{itX} = e^{itc}$ a.s.
- Going in the reverse direction, suppose $t \neq 0$ and $\phi(t) = e^{i\theta}$. Then $\tilde{X} = X - \frac{\theta}{t}$ has $\tilde{\phi}(t) = 1$, from which it follows that $\tilde{X} \in \frac{2\pi}{t}\mathbb{Z}$ a.s.
- $$\begin{aligned} |\phi(t+h) - \phi(t)| &= |\,\mathrm{E}[e^{i(t+h)X} - e^{itX}]| \\ &\leq \mathrm{E}[|e^{i(t+h)X} - e^{itX}|] = \mathrm{E}[|e^{ihX} - 1|] \end{aligned}$$
- $\mathrm{E}[e^{it(aX+b)}] = e^{itb}\,\mathrm{E}[e^{itaX}] = e^{itb}\phi(at)$.

□

Characteristic functions

Theorem

If X_1 and X_2 are independent and have characteristic functions ϕ_1 and ϕ_2, then $X_1 + X_2$ has characteristic function $\phi_1(t)\phi_2(t)$.

Proof.

We have

$$\mathrm{E}\left[e^{it(X_1+X_2)}\right] = \mathrm{E}\left[e^{itX_1}e^{itX_2}\right] = \mathrm{E}\left[e^{itX_1}\right]\mathrm{E}\left[e^{itX_2}\right].$$

□

Characteristic functions

Example (Classical characteristic functions)

- (Coin flips) If $\mathrm{Prob}(X=1)=\mathrm{Prob}(X=-1)=\frac{1}{2}$, then
$$\mathrm{E}[e^{itX}]=\frac{e^{it}+e^{-it}}{2}=\cos t.$$
- (Poisson distribution) If $\mathrm{Prob}(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}$ for $k=0,1,2,...,$ then
$$\mathrm{E}\left[e^{itX}\right]=\sum_{k=0}^{\infty}e^{-\lambda}\frac{\lambda^k e^{itk}}{k!}=\exp\left(\lambda(e^{it}-1)\right).$$

Characteristic functions

Example (Classical characteristic functions)

- (Normal distribution) The standard normal with density $\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ has $\phi(t) = e^{-\frac{t^2}{2}}$. To check this,

$$\phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}+itx} dx$$
$$= e^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} dx.$$

The last integral may be treated as a complex contour to complete the evaluation.

Characteristic functions

Example (Classical characteristic functions)

- (Uniform distribution on (a,b)) The density $\frac{1}{b-a}$ on (a,b) has $\phi(t) = \frac{e^{itb}-e^{ita}}{it(b-a)}$.
- (Triangular distribution, or Féjer kernel) The density $1-|x|$ on $(-1,1)$ has characteristic function

$$\phi(t) = \left(\frac{2\sin\frac{t}{2}}{t}\right)^2.$$

To check this, note that the density is the sum of two independent variables which are uniform on $\left(-\frac{1}{2}, \frac{1}{2}\right)$.

Characteristic functions

Example (Classical characteristic functions)

- (Exponential distribution) The density e^{-x} on $(0, \infty)$ has
$$\phi(t) = \int_0^\infty e^{itx-x} dx = \frac{1}{1-it}.$$
- (Bilateral exponential) The density $\frac{1}{2}e^{-|x|}$ on $\mathbb{R}$ has $\phi(t) = \frac{1}{1+t^2}$.
- (Cauchy distribution) The density $\frac{1}{\pi(1+x^2)}$ has $\phi(t) = \exp(-|t|)$. This follows from the previous calculation and the fact that the Fourier transform is an involution $L^2 \to L^2$.