Math 639: Lecture 6

Rates of convergence, local limit theorem, Poisson approximation

Bob Hough

February 14, 2017

The Lindeberg-Feller Theorem

Recall the Lindeberg-Feller CLT.

Theorem

For each n let $X_{m,n}$, $1 \leqslant m \leqslant n$ be independent random variables with $\mathrm{E}[X_{n,m}] = 0$. *Suppose*

1. $\sum_{m=1}^{n} \mathrm{E}\left[X_{m,n}^2\right] \to \sigma^2 > 0$
2. *For all* $\epsilon > 0$, $\lim_{n\to\infty} \sum_{m=1}^{n} \mathrm{E}\left[|X_{n,m}|^2 \mathbf{1}(|X_{n,m}| > \epsilon)\right] = 0$.

Then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma\eta$ *as* $n \to \infty$.

Record values

- Let $Y_1, Y_2, \ldots$ be independent with $\mathrm{Prob}(Y_m = 1) = \frac{1}{m}$, $\mathrm{Prob}(Y_m = 0) = 1 - \frac{1}{m}$.
- Set $S_n = Y_1 + \cdots + Y_n$. Then $\mathrm{E}[S_n] \sim \log n$ and $\mathrm{Var}[S_n] \sim \log n$.
- For $n > 1$ set $X_{n,m} = \frac{Y_m - \frac{1}{m}}{(\log n)^{\frac{1}{2}}}$.
- We have $\mathrm{E}[X_{n,m}] = 0$ and $\sum_{m=1}^n \mathrm{E}[X_{n,m}^2] \to 1$, and for any $\epsilon > 0$

$$\sum_{m=1}^{n} \mathrm{E}\left[|X_{n,m}|^2 \mathbf{1}(|X_{n,m}| > \epsilon)\right] \to 0$$

  since $|X_{n,m}| \leqslant \epsilon$ once $\frac{1}{(\log n)^{\frac{1}{2}}} < \epsilon$.
- By the CLT, $(\log n)^{-\frac{1}{2}} \left(S_n - \sum_{m=1}^n \frac{1}{m}\right) \Rightarrow \eta$.

Kolmogorov's three series theorem

Recall the statement of Kolmogorov's three series theorem.

Theorem

Let $X_1, X_2, \ldots$ be independent, let $A > 0$, and let $Y_m = X_m \mathbf{1}(|X_m| \leqslant A)$. In order that $\sum_{n=1}^{\infty} X_n$ converges 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.

- The first condition is necessary since otherwise, $|X_n| > A$ i.o. with probability 1 by Borel-Cantelli.
- If 1 holds, but 3 does not, then consider

$$c_n = \sum_{m=1}^n \mathrm{Var}(Y_m), \qquad X_{n,m} = \frac{(Y_m - \mathrm{E}[Y_m])}{c_n^{\frac{1}{2}}}.$$

One has $\mathrm{E}[X_{n,m}] = 0$, $\sum_{m=1}^n \mathrm{E}[X_{n,m}^2] = 1$ and, for any $\epsilon > 0$

$$\sum_{m=1}^n \mathrm{E}\left[|X_{n,m}|^2 \mathbf{1}(|X_{n,m}| > \epsilon)\right] \to 0$$

since the sum is 0 once $\frac{2A}{c_n^{\frac{1}{2}}} < \epsilon$.

$\square$

Kolmogorov's three series theorem

Proof.

- The above conditions imply $S_n = X_{n,1} + \cdots + X_{n,n}$ satisfies $S_n \Rightarrow \eta$. But if $\sum_{m=1}^\infty X_m$ converges a.s. then $\sum_{m=1}^\infty Y_m$ exists, so

$$T_n = \frac{1}{c_n^{\frac{1}{2}}} \sum_{m=1}^n Y_m \Rightarrow 0.$$

  This implies that $S_n - T_n \Rightarrow \eta$, but this is impossible, since the difference is the sum of the means, hence deterministic.
- If 1 and 3 hold, then $\sum_n (Y_n - \mathrm{E}[Y_n])$ converges a.s. If $\sum_n X_n$ converges, then $\sum_n Y_n$ converges, whence $\sum_n \mathrm{E}[Y_n]$ converges.

$\square$

Infinite variance

Example

- Let $X_1, X_2, \ldots$ be i.i.d. and have $\mathrm{Prob}(X_1 > x) = \mathrm{Prob}(X_1 < -x)$ and $\mathrm{Prob}(|X_1| > x) = x^{-2}$ for $x \geqslant 1$.
- Let $S_n = X_1 + \cdots + X_n$, and set

$$Y_{n,m} = X_m \mathbf{1}\left(|X_m| \leqslant n^{\frac{1}{2}} \log\log n\right).$$

- We have

$$\sum_{m=1}^{n} \mathrm{Prob}(Y_{n,m} \neq X_m) \leqslant n\,\mathrm{Prob}\left(|X_1| > n^{\frac{1}{2}} \log\log n\right) = \frac{1}{(\log\log n)^2}$$

tends to 0 as $n \to \infty$.

Infinite variance

Example

- Let $c_n = n^{\frac{1}{2}} \log\log n$. We have

$$\mathrm{E}\left[Y_{n,m}^2\right] = \int_1^\infty 2x\,\mathrm{Prob}(|Y_{n,m}| > x)dx$$
$$= \int_1^{c_n} 2x\left[\frac{1}{x^2} - \frac{1}{c_n^2}\right]dx$$
$$= \log n + 2\log\log\log n - 1.$$

  Thus $\sum_{m=1}^n \mathrm{E}\left[Y_{n,m}^2\right] \sim n\log n$.
- Since $\frac{Y_{n,m}}{\sqrt{n\log n}} \to 0$ in L^∞, the Lindeberg-Feller Theorem implies $\frac{1}{\sqrt{n\log n}}\sum_{m=1}^n Y_{n,m} \Rightarrow \eta$.

The Berry-Esseen Theorem

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d. with* $\mathrm{E}[X_i] = 0$, $\mathrm{E}[X_i^2] = \sigma^2$, *and* $\mathrm{E}\left[|X_i|^3\right] = \rho < \infty$. *If* $F_n(x)$ *is the distribution of* $\frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$ *and* N *the standard normal distribution function*

$$|F_n(x) - N(x)| \leqslant \frac{3\rho}{\sigma^3\sqrt{n}}.$$

The Berry-Esseen Theorem

Set $h_L(x) = \frac{1-\cos Lx}{\pi L x^2}$ with distribution H_L. This has characteristic function $\omega_L(\theta) = \left(1 - \left|\frac{\theta}{L}\right|\right)^+$.

Lemma (Smoothing lemma)

Let F and G be distribution functions, with $G'(x) \leqslant \lambda < \infty$. Let $\Delta(x) = F(x) - G(x)$, $\eta = \sup|\Delta(x)|$, $\Delta_L = \Delta * H_L$, and $\eta_L = \sup|\Delta_L(x)|$. Then

$$\eta_L \geqslant \frac{\eta}{2} - \frac{12\lambda}{\pi L}.$$

The Berry-Esseen Theorem

Proof.

- Δ goes to 0 at $\pm\infty$, G is continuous, F is a density function, so there is x_0 satisfying $\Delta(x_0) = \eta$ or $\Delta(x_0^-) = -\eta$. We'll treat the case $\Delta(x_0) = \eta$ as the other case may be handled similarly.
- The derivative condition implies in $s > 0$, $\Delta(x_0 + s) \geqslant \eta - \lambda s$.
- Let $\delta = \frac{\eta}{2\lambda}$, and $t = x_0 + \delta$

$$\Delta(t - x) \geqslant \begin{cases} \frac{\eta}{2} + \lambda x & |x| \leqslant \delta \\ -\eta & \text{otherwise} \end{cases}.$$

□

The Berry-Esseen Theorem

Proof.

- Use $\int_{|x|>\delta} h_L(x)dx \leqslant 2\int_\delta^\infty \frac{2dx}{\pi L x^2} = \frac{4}{\pi L \delta}$.
- Use $\int_{|x|\leqslant\delta} x h_L(x)dx = 0$ to find $\eta_L \geqslant \Delta_L(t)$ and

$$\Delta_L(t) = \int \Delta(t-x)H_L(x)dx \geqslant \frac{\eta}{2}\left(1 - \frac{4}{\pi L\delta}\right) - \frac{4\eta}{\pi L \delta} = \frac{\eta}{2} - \frac{12\lambda}{\pi L}.$$

□

The Berry-Esseen Theorem

Lemma

Let K_1 and K_2 be distribution functions with mean 0, whose characteristic functions κ_i are integrable. Then,

$$K_1(x) - K_2(x) = \frac{1}{2\pi} \int e^{-itx} \frac{\kappa_1(t) - \kappa_2(t)}{it} dt.$$

The Berry-Esseen Theorem

Proof.

- By the integrability, the distributions have densities

$$k_i(y) = \frac{1}{2\pi}\int e^{-ity}\kappa_i(t)dt.$$

- Set $\Delta K = K_1 - K_2$ and integrate to find

$$\begin{aligned}\Delta K(x) - \Delta K(a) &= \frac{1}{2\pi}\int_a^x \int e^{-ity}(\kappa_1(t) - \kappa_2(t))dtdy \\ &= \frac{1}{2\pi}\int (e^{-ita} - e^{-itx})\frac{\kappa_1(t) - \kappa_2(t)}{it}dt.\end{aligned}$$

□

The Berry-Esseen Theorem

Proof.

- Since the distribution functions are mean 0, $\frac{1-\kappa_i(t)}{t} \to 0$ as $t \to 0$, so $\frac{\kappa_1(t)-\kappa_2(t)}{it}$ is bounded and continuous.
- Let $a \to -\infty$ and use Riemann-Lebesgue to conclude

$$\Delta K(x) = \frac{1}{2\pi} \int -e^{-itx} \frac{\kappa_1(t) - \kappa_2(t)}{it} dt.$$

□

The Berry-Esseen Theorem

Proof of the Berry-Esseen Theorem.

- Both sides of the inequality scale with σ, so assume $\sigma = 1$.
- Write F for F_n and G for the distribution function of the Gaussian.
- Let ϕ_F and ϕ_G be the characteristic functions of F and G. Write $F_L = F * H_L$ and $G_L = G * H_L$.
- By the previous lemma

$$\begin{aligned} |F_L(x) - G_L(x)| &\leqslant \frac{1}{2\pi} \int |\phi_F(t)\omega_L(t) - \phi_G(t)\omega_L(t)| \frac{dt}{|t|} \\ &\leqslant \frac{1}{2\pi} \int_{-L}^{L} |\phi_F(t) - \phi_G(t)| \frac{dt}{|t|}. \end{aligned}$$

$\square$

Proof of the Berry-Esseen Theorem.

- By the smoothing lemma,

$$|F(x) - G(x)| \leqslant \frac{1}{\pi}\int_{-L}^{L} |\phi_F(\theta) - \phi_G(\theta)| \frac{d\theta}{|\theta|} + \frac{24\lambda}{\pi L}.$$

  Here $\lambda = \sup_x G'(x) = G'(0) = (2\pi)^{-\frac{1}{2}} < \frac{2}{5}$.
- Use $\left|\phi(t) - 1 + \frac{t^2}{2}\right| \leqslant \rho \frac{|t|^3}{6}$ and

$$|\alpha^n - \beta^n| \leqslant n|\alpha - \beta| \max(|\alpha|, |\beta|)^{n-1}.$$

$\square$

The Berry-Esseen Theorem

Proof of the Berry-Esseen Theorem.

- Let $L = \frac{4\sqrt{n}}{3\rho}$. Then for $|\theta| \leqslant L$,

$$\left|\phi\left(\frac{\theta}{\sqrt{n}}\right)\right| \leqslant 1 - \frac{\theta^2}{2n} + \frac{\rho|\theta|^3}{6n^{\frac{3}{2}}} \leqslant \exp\left(-\frac{5\theta^2}{18n}\right).$$

- Bound $\left|\phi\left(\frac{\theta}{\sqrt{n}}\right)^n - \exp\left(-\frac{\theta^2}{2}\right)\right|$ by choosing $\alpha = \phi\left(\frac{\theta}{\sqrt{n}}\right)$, $\beta = \exp\left(-\frac{\theta^2}{2n}\right)$, $\gamma = \exp\left(-\frac{5\theta^2}{18n}\right)$.
- One can bound

$$n|\alpha - \beta| \leqslant \frac{\rho|\theta|^3}{6n^{\frac{1}{2}}} + \frac{\theta^4}{8n}, \qquad \gamma^{n-1} \leqslant \exp\left(-\frac{\theta^2}{4}\right) \ (n \geqslant 10).$$

□

Proof of the Berry-Esseen Theorem.

- Putting things together,

$$\frac{1}{|\theta|}\left|\phi^n\left(\frac{\theta}{\sqrt{n}}\right) - \exp\left(-\frac{\theta^2}{2}\right)\right| \leqslant \exp\left(-\frac{\theta^2}{4}\right)\left\{\frac{\rho\theta^2}{6n^{\frac{1}{2}}} + \frac{|\theta|^3}{8n}\right\}$$
$$\leqslant \frac{1}{L}\exp\left(-\frac{\theta^2}{4}\right)\left\{\frac{2\theta^2}{9} + \frac{|\theta|^3}{18}\right\}.$$

- Hence

$$\pi L\,|F_n(x) - \eta(x)| \leqslant \int \exp\left(-\frac{\theta^2}{4}\right)\left\{\frac{2\theta^2}{9} + \frac{|\theta|^3}{18}\right\} d\theta + 9.6.$$

- The remainder of the proof amounts to calculator work.

□

Lattice distributed measures

Definition

A random variable X has a *lattice distribution* if there are constants b and $h > 0$ so that

$$\mathrm{Prob}(X \in b + h\mathbb{Z}) = 1.$$

The largest h for which this holds is called the *span* of the distribution.

Lattice distributed measures

Theorem

Let $\phi(t) = \mathrm{E}[e^{itX}]$. One of the following possibilities holds.

1. $|\phi(t)| < 1$ *for all* $t \neq 0$.
2. *There is* $\lambda > 0$ *so that* $|\phi(\lambda)| = 1$ *and* $|\phi(t)| < 1$ *for* $0 < t < \lambda$. *In this case* X *has a lattice distribution with span* $\frac{2\pi}{\lambda}$.
3. $|\phi(t)| = 1$ *for all* t. *In this case,* $X = b$ *a.s. for some* b.

Lattice distributed measures

Proof.

- We checked several lectures ago that if there is $\lambda > 0$ such that $|\phi(\lambda)| = 1$ then X is supported in $b + \frac{2\pi}{\lambda}\mathbb{Z}$ for some b.
- Suppose there is a sequence $t_n \downarrow 0$ such that $|\phi(t_n)| = 1$. Choose $b_n \in \left(-\frac{\pi}{t_n}, \frac{\pi}{t_n}\right]$ such that $\mathrm{Prob}(X \in b_n + \frac{2\pi}{t_n}\mathbb{Z}) = 1$.
- It follows that $\mathrm{Prob}(X = b_n) \to 1$. This is possible only if $b_n = b$ and $\mathrm{Prob}(X = b) = 1$.

$\square$

Lattice distributed measures

Definition

A random variable X is *arithmetic* if there is $h > 0$ such that $\mathrm{Prob}(X \in h\mathbb{Z}) = 1$.

Local limit theorem, lattice case

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d.,* $\mathrm{E}[X_i] = 0$, $\mathrm{E}[X_i^2] = \sigma^2$, *and lattice distributed, satisfying* $\mathrm{Prob}(X_1 \in b + h\mathbb{Z}) = 1$ *for some span* $h > 0$. *Set* $S_n = X_1 + \cdots + X_n$. *We put*

$$p_n(x) = \mathrm{Prob}\left(\frac{S_n}{\sqrt{n}} = x\right), \qquad \eta(x) = \frac{\exp\left(-\frac{x^2}{2\sigma^2}\right)}{\sqrt{2\pi}\sigma}.$$

As $n \to \infty$,

$$\sup_{x \in \left\{\frac{nb+hz}{\sqrt{n}} : z \in \mathbb{Z}\right\}} \left|\frac{n^{\frac{1}{2}}}{h} p_n(x) - \eta(x)\right| \to 0.$$

Local limit theorem, lattice case

Proof.

- Let $\phi(t) = \mathrm{E}[e^{itX}]$,

$$p_n(x) = \mathrm{Prob}\left(\frac{S_n}{\sqrt{n}} = x\right) = \frac{h}{2\pi\sqrt{n}} \int_{-\frac{\pi\sqrt{n}}{h}}^{\frac{\pi\sqrt{n}}{h}} e^{-itx}\phi^n\left(\frac{t}{\sqrt{n}}\right) dt.$$

- $\eta(x) = \frac{1}{2\pi}\int e^{-itx}\exp\left(-\frac{\sigma^2 t^2}{2}\right) dt.$
- We have

$$\left|\frac{n^{\frac{1}{2}}}{h} p_n(x) - \eta(x)\right| \leqslant \frac{1}{2\pi}\int_{-\frac{\pi\sqrt{n}}{h}}^{\frac{\pi\sqrt{n}}{h}} \left|\phi^n\left(\frac{t}{\sqrt{n}}\right) - \exp\left(-\frac{\sigma^2 t^2}{2}\right)\right| dt$$
$$+ \frac{1}{\pi}\int_{\frac{\pi\sqrt{n}}{h}}^{\infty} \exp\left(-\frac{\sigma^2 t^2}{2}\right) dt.$$

□

Local limit theorem, lattice case

Proof.

- For any fixed A,

$$\int_{-A}^{A} \left| \phi^n\left(\frac{t}{\sqrt{n}}\right) - \exp\left(-\frac{\sigma^2 t^2}{2}\right) \right| dt \to 0$$

  as $n \to \infty$ by bounded convergence.
- The remaining integral against $\exp\left(-\frac{\sigma^2 t^2}{2}\right)$ tends to 0 as a function of increasing A, so it remains to bound the integral against ϕ^n.

$\square$

Local limit theorem, lattice case

Proof.

- Use
$$|\phi(u)| \leqslant \left|1 - \frac{\sigma^2 u^2}{2}\right| + \frac{u^2}{2} \mathrm{E}\left[\min\left(|u||X|^3, 6|X|^2\right)\right].$$
Thus there is $\delta > 0$ such that for $|u| < \delta$, $|\phi(u)| < \exp\left(-\frac{\sigma^2 u^2}{4}\right)$, and so as $A \to \infty$,
$$\int_{A \leqslant |t| \leqslant \delta\sqrt{n}} \left|\phi^n\left(\frac{t}{\sqrt{n}}\right)\right| dt \to 0.$$
- For $\delta \leqslant |u| \leqslant \frac{\pi}{h}$, $|\phi|$ is bounded away from 1, so the remainder of the integral is exponentially small in n.

□

Non-lattice measures

Theorem

Let $X_1, X_2, \ldots$ *be i.i.d. with* $\mathrm{E}[X_i] = 0$, $\mathrm{E}[X_i^2] = \sigma^2 \in (0, \infty)$, *and having a common characteristic function* $\phi(t)$ *that has* $|\phi(t)| < 1$ *for all* $t \neq 0$. *Let* $S_n = X_1 + \cdots + X_n$ *and* $\eta(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$. *For* $a < b$, *as* $n \to \infty$, *if* $\frac{x_n}{\sqrt{n}} \to x$, *then*

$$\sqrt{n}\,\mathrm{Prob}\left(S_n \in (x_n + a, x_n + b)\right) \to (b - a)\eta(x).$$

Non-lattice measures

Proof.

- Let ψ be a Schwartz class function, and write $\hat{\psi}(t) = \int_{-\infty}^{\infty} \psi(x) e^{itx} dx$ for it's Fourier transform, which we assume to be of compact support, say in $[-T, T]$, $T > 0$. Such a function is said to be *band-limited*.
- Denote $\psi_{x_n}(x) = \psi(x - x_n)$ the translated function. We have

$$\hat{\psi}_{x_n}(t) = e^{itx_n} \hat{\psi}(t).$$

- Write $\phi(t) = \mathrm{E}[e^{itX_i}]$. By Plancherel,

$$\mathrm{E}\left[\psi_{x_n}(S_n)\right] = \frac{1}{2\pi} \int_{-T}^{T} \phi^n(t) e^{-itx_n} \overline{\hat{\psi}(t)} dt.$$

□

Non-lattice measures

Proof.

- Write $\phi(t) = \mathrm{E}[e^{itX_i}]$. By Plancherel,

$$\mathrm{E}\left[\psi_{x_n}(S_n)\right] = \frac{1}{2\pi}\int_{-T}^{T} \exp\left(-\frac{n\sigma^2 t^2}{2}\right) e^{-itx_n}\overline{\hat{\psi}(t)}dt$$
$$+ O\left(\int_{-T}^{T}\left|\phi^n(t) - \exp\left(-\frac{n\sigma^2 t^2}{2}\right)\right| dt\right).$$

- The error term is $o\left(\frac{1}{\sqrt{n}}\right)$ by splitting the integral into three pieces as in the previous theorem.

□

Non-lattice measures

Proof.

- The main term is
$$\int \frac{e^{-\frac{x^2}{2n\sigma^2}}}{\sqrt{2\pi n\sigma^2}} \psi_{x_n}(x)\,dx.$$
- This suffices for the theorem, since the main term of the theorem is asymptotic to
$$\int_{x_n+a}^{x_n+b} \frac{e^{-\frac{x^2}{2n\sigma^2}}}{\sqrt{2\pi n\sigma^2}}\,dx$$
and the indicator function of $[a, b]$ can be approximated in L^1 from above and below by Schwartz functions whose Fourier Transform has compact support.

□

Poisson convergence

Recall the Poisson(λ) distribution has $\mathrm{Prob}(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}$.

Theorem

For each n, let $X_{n,m}$, $1 \leqslant m \leqslant n$ be independent random variables with $\mathrm{Prob}(X_{n,m} = 1) = p_{n,m}$, $\mathrm{Prob}(X_{n,m} = 0) = 1 - p_{n,m}$. *Suppose*

1. $\sum_{m=1}^{n} p_{n,m} \to \lambda \in (0, \infty)$.
2. $\max_{1 \leqslant m \leqslant n} p_{n,m} \to 0$.

If $S_n = X_{n,1} + \cdots + X_{n,n}$ then $S_n \Rightarrow Z$ where Z is Poisson(λ).

Poisson convergence

Proof.

- $\phi_{n,m}(t) = \mathrm{E}\left[\exp(itX_{n,m})\right] = (1 - p_{n,m}) + p_{n,m}e^{it}$.
- $\mathrm{E}\left[e^{itS_n}\right] = \prod_{m=1}^n \left(1 + p_{n,m}(e^{it} - 1)\right)$.
- Note $|\exp(p(e^{it} - 1))| = \exp(p(\Re(e^{it} - 1))) \leqslant 1$, $|1 + p(e^{it} - 1)| \leqslant 1$. Thus

$$\begin{aligned}
&\left|\exp\left(\sum_{m=1}^n p_{n,m}(e^{it} - 1)\right) - \prod_{m=1}^n (1 + p_{n,m}(e^{it} - 1))\right| \\
&\leqslant \sum_{m=1}^n \left|\exp(p_{n,m}(e^{it} - 1)) - (1 + p_{n,m}(e^{it} - 1))\right| \\
&\leqslant \sum_{m=1}^n p_{n,m}^2 \left|e^{it} - 1\right|^2 \leqslant 4\left(\max_{1 \leqslant m \leqslant n} p_{n,m}\right) \sum_{m=1}^n p_{n,m} \to 0.
\end{aligned}$$

□

Poisson convergence

Since $\sum_{m=1}^{n} p_{n,m} \to \lambda$, $\mathrm{E}\left[\exp(itS_n)\right] \to \exp(\lambda(e^{it} - 1))$. Since the characteristic function converges pointwise to the characteristic function of Poisson(λ), the convergence in distribution follows.

Poisson convergence

Example

- Suppose we roll two dice 36 times. The number of times that 'snake eyes' (two ones) occurs has distribution which is approximately Poisson(1).
- Let $\xi_{n,1}, \xi_{n,2}, ..., \xi_{n,n}$ be independent and uniformly distributed over $[-n, n]$. Let $X_{n,m}$ indicate the event that $\xi_{n,m} \in (a, b)$, which has probability $\frac{b-a}{2n}$. The number of events, $S_n = \sum_m X_{n,m}$ converges to a Poisson distribution of parameter $\frac{b-a}{2}$.

Poisson convergence

Definition

The *total variation distance* between two probability measures μ and ν on a countable set S is

$$\|\mu - \nu\| = \frac{1}{2}\sum_z |\mu(z) - \nu(z)| = \sup_{A \subset S} |\mu(A) - \nu(A)|.$$

Poisson convergence

Lemma

If $\mu_1 \times \mu_2$ *denotes the product measure on* $\mathbb{Z} \times \mathbb{Z}$ *that has* $(\mu_1 \times \mu_2)(x, y) = \mu_1(x)\mu_2(y)$, *then*

$$\|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\| \leqslant \|\mu_1 - \nu_1\| + \|\mu_2 - \nu_2\|.$$

Poisson convergence

Proof.

$$
\begin{aligned}
2\|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\| &= \sum_{x,y} |\mu_1(x)\mu_2(y) - \nu_1(x)\nu_2(y)| \\
&\leqslant \sum_{x,y} |\mu_1(x)\mu_2(y) - \nu_1(x)\mu_2(y)| + \sum_{x,y} |\nu_1(x)\mu_2(y) - \nu_1(x)\nu_2(y)| \\
&= \sum_y \mu_2(y) \sum_x |\mu_1(x) - \nu_1(x)| + \sum_x \nu_1(x) \sum_y |\mu_2(y) - \nu_2(y)| \\
&= 2\|\mu_1 - \nu_1\| + 2\|\mu_2 - \nu_2\|.
\end{aligned}
$$

$\square$

Poisson convergence

Lemma

If $\mu_1 * \mu_2$ *denotes the convolution of* μ_1 *and* μ_2*, that is,*

$$\mu_1 * \mu_2(x) = \sum_y \mu_1(x-y)\mu_2(y)$$

then $\|\mu_1 * \mu_2 - \nu_1 * \nu_2\| \leqslant \|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\|$.

Poisson convergence

Proof.

$$\begin{aligned}2\|\mu_1 * \mu_2 - \nu_1 * \nu_2\| &= \sum_x \left| \sum_y \mu_1(x-y)\mu_2(y) - \sum_y \nu_1(x-y)\nu_2(y) \right| \\ &\leqslant \sum_x \sum_y |\mu_1(x-y)\mu_2(y) - \nu_1(x-y)\nu_2(y)| \\ &= 2\|\mu_1 \times \mu_2 - \nu_1 \times \nu_2\|.\end{aligned}$$

$\square$

Poisson convergence

Lemma

Let μ be the measure with $\mu(1) = p$ and $\mu(0) = 1 - p$. Let ν be a Poisson distribution with mean p. Then $\|\mu - \nu\| \leqslant p^2$.

Poisson convergence

Proof.

$$
\begin{aligned}
2\|\mu-\nu\| &= |\mu(0)-\nu(0)| + |\mu(1)-\nu(1)| + \sum_{n\geqslant 2}\nu(n) \\
&= |1-p-e^{-p}| + |p-pe^{-p}| + 1 - e^{-p}(1+p).
\end{aligned}
$$

Since $1-x \leqslant e^{-x} \leqslant 1$ for $x \geqslant 0$, one obtains

$$
= 2p(1-e^{-p}) \leqslant 2p^2.
$$

$\square$

Poisson convergence

Theorem

For each n, let $X_{n,m}$, $1 \leqslant m \leqslant n$ be independent random variables with $\mathrm{Prob}(X_{n,m} = 1) = p_{n,m}$, $\mathrm{Prob}(X_{n,m} = 0) = 1 - p_{n,m}$. *Suppose*

- $\sum_{m=1}^n p_{n,m} = \lambda$
- $\max_{1 \leqslant m \leqslant n} p_{n,m} \to 0$.

Let $S_n = X_{n,1} + \cdots + X_{n,n}$ have distribution μ_n, and let ν have distribution $\mathrm{Poisson}(\lambda)$. *Then*

$$\|\mu_n - \nu\| \leqslant \sum_{m=1}^n p_{n,m}^2.$$

Poisson convergence

Proof.

- Let $\mu_{n,m}$ be the distribution of $X_{n,m}$, and let $\nu_{n,m}$ be Poisson$(p_{n,m})$.
- Thus $\nu_n = *_{m=1}^n \nu_{n,m} \sim \text{Poisson}(\lambda)$ and $\sum_{m=1}^n X_{n,m}$ has distribution $*_{m=1}^n \mu_{n,m}$.
- $\|\mu_n - \nu_n\| \leqslant \sum_{m=1}^n \|\mu_{n,m} - \nu_{n,m}\| \leqslant \sum_{m=1}^n p_{n,m}^2$.

□

Fixed points

Example

Let π be a random permutation of $\{1, 2, ..., n\}$, let $X_m = 1$ if $\pi(m) = m$ and 0 otherwise, and let $S_n = X_1 + \cdots + X_n$ be the number of fixed points. Let $A_m = \{X_m = 1\}$. By inclusion-exclusion,

$$\begin{aligned}\mathrm{Prob}\left(\bigcup_{m=1}^{n} A_m\right) &= \sum_m \mathrm{Prob}(A_m) - \sum_{\ell<m} \mathrm{Prob}(A_\ell \cap A_m) \\ &\quad + \sum_{k<\ell<m} \mathrm{Prob}(A_k \cap A_\ell \cap A_m) - \cdots \\ &= n \cdot \frac{1}{n} - \binom{n}{2}\frac{(n-2)!}{n!} + \binom{n}{3}\frac{(n-3)!}{n!} - \cdots .\end{aligned}$$

Fixed points

Example

We have $\text{Prob}(S_n = 0) = \sum_{m=0}^{n} \frac{(-1)^m}{m!}$ so

$$\left|\text{Prob}(S_n = 0) - e^{-1}\right| = \left|\sum_{m=n+1}^{\infty} \frac{(-1)^m}{m!}\right|$$

$$\leqslant \frac{1}{(n+1)!}\sum_{k=0}^{\infty}(n+2)^{-k} = \frac{1}{(n+1)!}\left(1 - \frac{1}{n+2}\right)^{-1}.$$

We can now compute

$$\text{Prob}(S_n = k) = \binom{n}{k}\frac{\text{Prob}(S_{n-k} = 0)}{n(n-1)\cdots(n-k+1)}$$

$$= \frac{\text{Prob}(S_{n-k} = 0)}{k!} \to \frac{e^{-1}}{k!}.$$

Occupancy problem

Theorem

Suppose r balls are placed at random in n boxes. If $ne^{-\frac{r}{n}} \to \lambda \in [0, \infty)$ the number of empty boxes approaches $\mathrm{Poisson}(\lambda)$ *as $n \to \infty$.*

Occupancy problem

Proof.

- Set $p_m(r,n)$ for the probability of m empty boxes on r tosses into n boxes.
- Since $\mathrm{Prob}(\text{boxes } i_1, i_2, ..., i_k \text{ empty}) = \left(1 - \frac{k}{n}\right)^r$, by inclusion-exclusion

$$p_0(r,n) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \left(1 - \frac{k}{n}\right)^r.$$

One obtains $p_0(r,n) \sim e^{-\lambda}$ by using $\left(1 - \frac{k}{n}\right)^r \sim \frac{\lambda^k}{n^k}$ for $k \leqslant K$, a large fixed constant, and $\left(1 - \frac{k}{n}\right)^r \lesssim \frac{\lambda^k}{n^k}$, $k > K$.

□

Occupancy problem

Proof.

- By choosing the boxes to be empty

$$p_m(r,n) = \binom{n}{m}\left(1 - \frac{m}{n}\right)^r p_0(r, n-m) \sim \frac{\lambda^m}{m!} p_0(r, n-m) \sim e^{-\lambda}\frac{\lambda^m}{m!}.$$

□

Coupon collector's problem

Example

Let $X_1, X_2, \ldots$ be i.i.d. uniform on $\{1, 2, \ldots, n\}$ and $T_n = \inf\{m : \{X_1, \ldots, X_m\} = \{1, 2, \ldots, n\}\}$. Since $T_n \leqslant m$ if and only if m balls fill up all n boxes, it follows

$$\mathrm{Prob}(T_n - n\log n \leqslant nx) = p_0(n\log n + nx, n) \to \exp(-e^{-x}).$$

This follows from the previous discussion, since if $r = n\log n + nx$ then $ne^{-\frac{r}{n}} \to e^{-x}$.