18.175: Lecture 11

Central limit theorem variants

Scott Sheffield

MIT

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Outline

Recall Fourier inversion formula

- If $f : \mathbb{R} \to \mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.

Recall Fourier inversion formula

- If $f : \mathbb{R} \to \mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.
- **Fourier inversion:** If f is nice: $f(x) = \frac{1}{2\pi}\int \hat{f}(t)e^{itx}dt$.

Recall Fourier inversion formula

- If $f:\mathbb{R}\to\mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.
- **Fourier inversion:** If f is nice: $f(x) = \frac{1}{2\pi}\int \hat{f}(t)e^{itx}dt$.
- Easy to check this when f is density function of a Gaussian. Use linearity of $f \to \hat{f}$ to extend to linear combinations of Gaussians, or to convolutions with Gaussians.

Recall Fourier inversion formula

- If $f:\mathbb{R}\to\mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.
- **Fourier inversion:** If f is nice: $f(x) = \frac{1}{2\pi}\int \hat{f}(t)e^{itx}dt$.
- Easy to check this when f is density function of a Gaussian. Use linearity of $f \to \hat{f}$ to extend to linear combinations of Gaussians, or to convolutions with Gaussians.
- Show $f \to \hat{f}$ is an isometry of Schwartz space (endowed with L^2 norm). Extend definition to L^2 completion.

Recall Fourier inversion formula

- If $f : \mathbb{R} \to \mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.
- **Fourier inversion:** If f is nice: $f(x) = \frac{1}{2\pi}\int \hat{f}(t)e^{itx}dt$.
- Easy to check this when f is density function of a Gaussian. Use linearity of $f \to \hat{f}$ to extend to linear combinations of Gaussians, or to convolutions with Gaussians.
- Show $f \to \hat{f}$ is an isometry of Schwartz space (endowed with L^2 norm). Extend definition to L^2 completion.
- **Convolution theorem:** If

$$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x-y)dy,$$

then

$$\hat{h}(t) = \hat{f}(t)\hat{g}(t).$$

Recall Fourier inversion formula

- If $f : \mathbb{R} \to \mathbb{C}$ is in L^1, write $\hat{f}(t) := \int_{-\infty}^{\infty} f(x)e^{-itx}dx$.
- **Fourier inversion:** If f is nice: $f(x) = \frac{1}{2\pi}\int \hat{f}(t)e^{itx}dt$.
- Easy to check this when f is density function of a Gaussian. Use linearity of $f \to \hat{f}$ to extend to linear combinations of Gaussians, or to convolutions with Gaussians.
- Show $f \to \hat{f}$ is an isometry of Schwartz space (endowed with L^2 norm). Extend definition to L^2 completion.
- **Convolution theorem:** If

  $$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x-y)dy,$$

  then

  $$\hat{h}(t) = \hat{f}(t)\hat{g}(t).$$

- **Observation:** can define Fourier transforms of generalized functions. Can interpret finite measure as generalized function.

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.
- Bochner's theorem: a continuous function from $\mathbb{R}$ to $\mathbb{C}$ with $\phi(0) = 1$ is a characteristic function of a some probability measure on $\mathbb{R}$ if and only if it is positive definite.

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.
- Bochner's theorem: a continuous function from $\mathbb{R}$ to $\mathbb{C}$ with $\phi(0) = 1$ is a characteristic function of a some probability measure on $\mathbb{R}$ if and only if it is positive definite.
- Why positive definite?

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.
- Bochner's theorem: a continuous function from $\mathbb{R}$ to $\mathbb{C}$ with $\phi(0) = 1$ is a characteristic function of a some probability measure on $\mathbb{R}$ if and only if it is positive definite.
- Why positive definite?
- Write $Y = \sum_{j=1}^n a_j e^{t_j X}$ and observe

$$Y\overline{Y} = \sum_{j=1}^n \sum_{k=1}^n a_j \overline{a_k} e^{(t_i - t_j)X},$$

$$\mathbb{E}Y\overline{Y} = \sum_{j=1}^n \sum_{k=1}^n a_j \overline{a_k} \phi(t_i - t_j).$$

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.
- Bochner's theorem: a continuous function from $\mathbb{R}$ to $\mathbb{C}$ with $\phi(0) = 1$ is a characteristic function of a some probability measure on $\mathbb{R}$ if and only if it is positive definite.
- Why positive definite?
- Write $Y = \sum_{j=1}^n a_j e^{t_j X}$ and observe

$$Y\overline{Y} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j \overline{a_k} e^{(t_i - t_j)X},$$

$$\mathbb{E} Y\overline{Y} = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j \overline{a_k} \phi(t_i - t_j).$$

Recall Bochner's theorem

- Given function ϕ and points $t_1, \ldots, t_n$, consider matrix with i, j entry given by $\phi(t_i - t_j)$. Call ϕ **positive definite** if this matrix is always positive semidefinite Hermitian.
- Bochner's theorem: a continuous function from $\mathbb{R}$ to $\mathbb{C}$ with $\phi(0) = 1$ is a characteristic function of a some probability measure on $\mathbb{R}$ if and only if it is positive definite.
- Why positive definite?
- Write $Y = \sum_{j=1}^n a_j e^{t_j X}$ and observe

$$Y\overline{Y} = \sum_{j=1}^n \sum_{k=1}^n a_j \overline{a_k} e^{(t_i - t_j)X},$$

$$\mathbb{E} Y\overline{Y} = \sum_{j=1}^n \sum_{k=1}^n a_j \overline{a_k} \phi(t_i - t_j).$$

- **Fourier transform:** natural one-to-one map from set of probability measures on $\mathbb{R}$ (describable by distribution functions F) to set of possible characteristic functions.

Recall continuity theorem

- **Strong continuity theorem:** If $\mu_n \implies \mu_\infty$ then $\phi_n(t) \to \phi_\infty(t)$ for all t. Conversely, if $\phi_n(t)$ converges to a limit that is continuous at 0, then the associated sequence of distributions μ_n is tight and converges weakly to a measure μ with characteristic function ϕ.

Recall CLT idea

- Let X be a random variable.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L_X'(0) = -\phi_X'(0)/\phi_X(0) = 0$ and $L_X'' = -(\phi_X''(0)\phi_X(0) - \phi_X'(0)^2)/\ \phi_X(0)^2 = 1$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L_X'(0) = -\phi_X'(0)/\phi_X(0) = 0$ and $L_X'' = -(\phi_X''(0)\phi_X(0) - \phi_X'(0)^2)/\ \phi_X(0)^2 = 1$.
- If $V_n = n^{-1/2}\sum_{i=1}^n X_i$ where X_i are i.i.d. with law of X, then $L_{V_n}(t) = nL_X(n^{-1/2}t)$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L_X'(0) = -\phi_X'(0)/\phi_X(0) = 0$ and $L_X'' = -(\phi_X''(0)\phi_X(0) - \phi_X'(0)^2)/\ \phi_X(0)^2 = 1$.
- If $V_n = n^{-1/2}\sum_{i=1}^n X_i$ where X_i are i.i.d. with law of X, then $L_{V_n}(t) = nL_X(n^{-1/2}t)$.
- When we zoom in on a twice differentiable function near zero (scaling vertically by n and horizontally by $\sqrt{n}$) the picture looks increasingly like a parabola.

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Lindeberg-Feller theorem

- CLT is pretty special. What other kinds of sums are approximately Gaussian?

Lindeberg-Feller theorem

- CLT is pretty special. What other kinds of sums are approximately Gaussian?
- **Triangular arrays:** Suppose $X_{n,m}$ are independent expectation-zero random variables when $1 \leq m \leq n$.

Lindeberg-Feller theorem

- CLT is pretty special. What other kinds of sums are approximately Gaussian?
- **Triangular arrays:** Suppose $X_{n,m}$ are independent expectation-zero random variables when $1 \leq m \leq n$.
- Suppose $\sum_{m=1}^{n} EX_{n,m}^2 \to \sigma^2 > 0$ and for all ϵ, $\lim_{n\to\infty} E(|X_{n,m}|^2; |X_{n,m}| > \epsilon) = 0$.

Lindeberg-Feller theorem

- CLT is pretty special. What other kinds of sums are approximately Gaussian?
- **Triangular arrays:** Suppose $X_{n,m}$ are independent expectation-zero random variables when $1 \leq m \leq n$.
- Suppose $\sum_{m=1}^{n} EX_{n,m}^2 \to \sigma^2 > 0$ and for all ϵ, $\lim_{n\to\infty} E(|X_{n,m}|^2; |X_{n,m}| > \epsilon) = 0$.
- Then $S_n = X_{n,1} + X_{n,2} + \ldots + X_{n,n} \implies \sigma\chi$ (where χ is standard normal) as $n \to \infty$.

Lindeberg-Feller theorem

- CLT is pretty special. What other kinds of sums are approximately Gaussian?
- **Triangular arrays:** Suppose $X_{n,m}$ are independent expectation-zero random variables when $1 \leq m \leq n$.
- Suppose $\sum_{m=1}^{n} EX_{n,m}^2 \to \sigma^2 > 0$ and for all ϵ, $\lim_{n\to\infty} E(|X_{n,m}|^2; |X_{n,m}| > \epsilon) = 0$.
- Then $S_n = X_{n,1} + X_{n,2} + \ldots + X_{n,n} \implies \sigma\chi$ (where χ is standard normal) as $n \to \infty$.
- **Proof idea:** Use characteristic functions $\phi_{n,m} = \phi_{X_{n,m}}$. Try to get some uniform handle on how close they are to their quadratic approximations.

Berry-Esseen theorem

- If X_i are i.i.d. with mean zero, variance σ^2, and $E|X_i|^3 = \rho < \infty$, and $F_n(x)$ is distribution of $(X_1 + \ldots + X_n)/(\sigma\sqrt{n})$ and $\Phi(x)$ is standard normal distribution, then $|F_n(x) - \Phi(x)| \leq 3\rho/(\sigma^3\sqrt{n})$.

Berry-Esseen theorem

- If X_i are i.i.d. with mean zero, variance σ^2, and $E|X_i|^3 = \rho < \infty$, and $F_n(x)$ is distribution of $(X_1 + \ldots + X_n)/(\sigma\sqrt{n})$ and $\Phi(x)$ is standard normal distribution, then $|F_n(x) - \Phi(x)| \leq 3\rho/(\sigma^3\sqrt{n})$.
- Provided one has a third moment, CLT convergence is very quick.

Berry-Esseen theorem

- If X_i are i.i.d. with mean zero, variance σ^2, and $E|X_i|^3 = \rho < \infty$, and $F_n(x)$ is distribution of $(X_1 + \ldots + X_n)/(\sigma\sqrt{n})$ and $\Phi(x)$ is standard normal distribution, then $|F_n(x) - \Phi(x)| \leq 3\rho/(\sigma^3\sqrt{n})$.
- Provided one has a third moment, CLT convergence is very quick.
- **Proof idea:** You can convolve with something that has a characteristic function with compact support. Play around with Fubini, error estimates.

Local limit theorems for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.

Local limit theorems for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.

Local limit theorems for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.
- Write $p_n(x) = P(S_n/\sqrt{n} = x)$ for $x \in \mathcal{L}_n := (nb + h\mathbb{Z})/\sqrt{n}$ and $n(x) = (2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2)$.

Local limit theorems for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.
- Write $p_n(x) = P(S_n/\sqrt{n} = x)$ for $x \in \mathcal{L}_n := (nb + h\mathbb{Z})/\sqrt{n}$ and $n(x) = (2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2)$.
- Assume X_i are i.i.d. lattice with $EX_i = 0$ and $EX_i^2 = \sigma^2 \in (0, \infty)$. **Theorem:** As $n \to \infty$,

$$\sup_{x \in \mathcal{L}^n} \left| \frac{n^{1/2}}{h} p_n(x) - n(x) \right| \to 0.$$

Local limit theorems for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.
- Write $p_n(x) = P(S_n/\sqrt{n} = x)$ for $x \in \mathcal{L}_n := (nb + h\mathbb{Z})/\sqrt{n}$ and $n(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2)$.
- Assume X_i are i.i.d. lattice with $EX_i = 0$ and $EX_i^2 = \sigma^2 \in (0,\infty)$. **Theorem:** As $n \to \infty$,

$$\sup_{x \in \mathcal{L}^n} \left| \frac{n^{1/2}}{h} p_n(x) - n(x) \right| \to 0.$$

- **Proof idea:** Use characteristic functions, reduce to periodic integral problem. Note that for Y supported on $a + \theta\mathbb{Z}$, we have $P(Y = x) = \frac{1}{2\pi/\theta}\int_{-\pi/\theta}^{\pi/\theta} e^{-itx}\phi_Y(t)dt$.

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Outline

Recall local CLT for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.

Recall local CLT for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.

Recall local CLT for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.
- Write $p_n(x) = P(S_n/\sqrt{n} = x)$ for $x \in \mathcal{L}_n := (nb + h\mathbb{Z})/\sqrt{n}$ and $n(x) = (2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2)$.

Recall local CLT for walks on $\mathbb{Z}$

- Suppose $X \in b + h\mathbb{Z}$ a.s. for some fixed constants b and h.
- Observe that if $\phi_X(\lambda) = 1$ for some $\lambda \neq 0$ then X is supported on (some translation of) $(2\pi/\lambda)\mathbb{Z}$. If this holds for all λ, then X is a.s. some constant. When the former holds but not the latter (i.e., ϕ_X is periodic but not identically 1) we call X a **lattice random variable**.
- Write $p_n(x) = P(S_n/\sqrt{n} = x)$ for $x \in \mathcal{L}_n := (nb + h\mathbb{Z})/\sqrt{n}$ and $n(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/2\sigma^2)$.
- Assume X_i are i.i.d. lattice with $EX_i = 0$ and $EX_i^2 = \sigma^2 \in (0,\infty)$. **Theorem:** As $n \to \infty$,

$$\Big| \sup_{x \in \mathcal{L}^n} |n^{1/2}/hp_n(x) - n(x)\Big| \to 0.$$

Recall local CLT for walks on $\mathbb{Z}$

- **Proof idea:** Use characteristic functions, reduce to periodic integral problem. Look up "Fourier series". Note that for Y supported on $a + \theta\mathbb{Z}$, we have $P(Y = x) = \frac{1}{2\pi/\theta} \int_{-\pi/\theta}^{\pi/\theta} e^{-itx}\phi_Y(t)dt$.

Extending this idea to higher dimensions

- Example: suppose we have random walk on $\mathbb{Z}$ that at each step tosses fair 4-sided coin to decide whether to go 1 unit left, 1 unit right, 2 units left, or 2 units right?

Extending this idea to higher dimensions

- Example: suppose we have random walk on $\mathbb{Z}$ that at each step tosses fair 4-sided coin to decide whether to go 1 unit left, 1 unit right, 2 units left, or 2 units right?
- What is the probability that the walk is back at the origin after one step? Two steps? Three steps?

Extending this idea to higher dimensions

- Example: suppose we have random walk on $\mathbb{Z}$ that at each step tosses fair 4-sided coin to decide whether to go 1 unit left, 1 unit right, 2 units left, or 2 units right?
- What is the probability that the walk is back at the origin after one step? Two steps? Three steps?
- One could compute this in Mathematica by writing out the characteristic function ϕ_X for one-step increment X and calculating $\int_0^{2\pi} \phi_X^k(t)dt/2\pi$.

Extending this idea to higher dimensions

- Example: suppose we have random walk on $\mathbb{Z}$ that at each step tosses fair 4-sided coin to decide whether to go 1 unit left, 1 unit right, 2 units left, or 2 units right?
- What is the probability that the walk is back at the origin after one step? Two steps? Three steps?
- One could compute this in Mathematica by writing out the characteristic function ϕ_X for one-step increment X and calculating $\int_0^{2\pi} \phi_X^k(t)dt/2\pi$.
- How about a random walk on $\mathbb{Z}^2$?

Extending this idea to higher dimensions

- Example: suppose we have random walk on $\mathbb{Z}$ that at each step tosses fair 4-sided coin to decide whether to go 1 unit left, 1 unit right, 2 units left, or 2 units right?
- What is the probability that the walk is back at the origin after one step? Two steps? Three steps?
- One could compute this in Mathematica by writing out the characteristic function ϕ_X for one-step increment X and calculating $\int_0^{2\pi} \phi_X^k(t)dt/2\pi$.
- How about a random walk on $\mathbb{Z}^2$?
- Can one use this to establish when a random walk on $\mathbb{Z}^d$ is recurrent versus transient?

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?
- How many radioactive particles emitted during a time period in which the expected number emitted is 5?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?
- How many radioactive particles emitted during a time period in which the expected number emitted is 5?
- How many calls to call center during a given minute?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?
- How many radioactive particles emitted during a time period in which the expected number emitted is 5?
- How many calls to call center during a given minute?
- How many goals scored during a 90 minute soccer game?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?
- How many radioactive particles emitted during a time period in which the expected number emitted is 5?
- How many calls to call center during a given minute?
- How many goals scored during a 90 minute soccer game?
- How many notable gaffes during 90 minute debate?

Poisson random variables: motivating questions

- How many raindrops hit a given square inch of sidewalk during a ten minute period?
- How many people fall down the stairs in a major city on a given day?
- How many plane crashes in a given year?
- How many radioactive particles emitted during a time period in which the expected number emitted is 5?
- How many calls to call center during a given minute?
- How many goals scored during a 90 minute soccer game?
- How many notable gaffes during 90 minute debate?
- **Key idea for all these examples:** Divide time into large number of small increments. Assume that during each increment, there is some small probability of thing happening (independently of other increments).

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?
- Answer: $np = \lambda$.

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?
- Answer: $np = \lambda$.
- Let k be some moderate sized number (say $k = 4$). What is the probability that I see exactly k heads?

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?
- Answer: $np = \lambda$.
- Let k be some moderate sized number (say $k = 4$). What is the probability that I see exactly k heads?
- Binomial formula: $\binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)(n-2)\dots(n-k+1)}{k!} p^k (1-p)^{n-k}$.

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?
- Answer: $np = \lambda$.
- Let k be some moderate sized number (say $k = 4$). What is the probability that I see exactly k heads?
- Binomial formula: $\binom{n}{k}p^k(1-p)^{n-k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!}p^k(1-p)^{n-k}$.
- This is approximately $\frac{\lambda^k}{k!}(1-p)^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}$.

Bernoulli random variable with n large and $np = \lambda$

- Let λ be some moderate-sized number. Say $\lambda = 2$ or $\lambda = 3$. Let n be a huge number, say $n = 10^6$.
- Suppose I have a coin that comes up heads with probability λ/n and I toss it n times.
- How many heads do I expect to see?
- Answer: $np = \lambda$.
- Let k be some moderate sized number (say $k = 4$). What is the probability that I see exactly k heads?
- Binomial formula: $\binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} p^k (1-p)^{n-k}$.
- This is approximately $\frac{\lambda^k}{k!}(1-p)^{n-k} \approx \frac{\lambda^k}{k!} e^{-\lambda}$.
- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$ for integer $k \geq 0$.

Probabilities sum to one

- A **Poisson random variable** X with parameter λ satisfies $p(k) = P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.

Probabilities sum to one

- A **Poisson random variable** X with parameter λ satisfies $p(k) = P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$ for integer $k \geq 0$.
- How can we show that $\sum_{k=0}^{\infty} p(k) = 1$?

Probabilities sum to one

- A **Poisson random variable** X with parameter λ satisfies $p(k) = P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- How can we show that $\sum_{k=0}^{\infty} p(k) = 1$?
- Use Taylor expansion $e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$.

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- What is $E[X]$?

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- What is $E[X]$?
- We think of a Poisson random variable as being (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- What is $E[X]$?
- We think of a Poisson random variable as being (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This would suggest $E[X] = \lambda$. Can we show this directly from the formula for $P\{X = k\}$?

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- What is $E[X]$?
- We think of a Poisson random variable as being (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This would suggest $E[X] = \lambda$. Can we show this directly from the formula for $P\{X = k\}$?
- By definition of expectation

$$E[X] = \sum_{k=0}^{\infty} P\{X = k\}k = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda}.$$

Expectation

- A **Poisson random variable** X with parameter λ satisfies $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$.
- What is $E[X]$?
- We think of a Poisson random variable as being (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This would suggest $E[X] = \lambda$. Can we show this directly from the formula for $P\{X = k\}$?
- By definition of expectation

$$E[X] = \sum_{k=0}^{\infty} P\{X = k\}k = \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda}.$$

- Setting $j = k - 1$, this is $\lambda \sum_{j=0}^{\infty} \frac{\lambda^j}{j!}e^{-\lambda} = \lambda$.

Variance

- Given $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?

Variance

- Given $P\{X = k\} = \frac{\lambda^k}{k!} e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.

Variance

- Given $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This suggests $\mathrm{Var}[X] \approx npq \approx \lambda$ (since $np \approx \lambda$ and $q = 1 - p \approx 1$). Can we show directly that $\mathrm{Var}[X] = \lambda$?

Variance

- Given $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This suggests $\mathrm{Var}[X] \approx npq \approx \lambda$ (since $np \approx \lambda$ and $q = 1 - p \approx 1$). Can we show directly that $\mathrm{Var}[X] = \lambda$?

Variance

- Given $P\{X = k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n, p) random variable with n very large and $p = \lambda/n$.
- This suggests $\mathrm{Var}[X] \approx npq \approx \lambda$ (since $np \approx \lambda$ and $q = 1 - p \approx 1$). Can we show directly that $\mathrm{Var}[X] = \lambda$?
- Compute

$$E[X^2] = \sum_{k=0}^{\infty} P\{X = k\}k^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda}.$$

Variance

- Given $P\{X=k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n,p) random variable with n very large and $p = \lambda/n$.
- This suggests $\mathrm{Var}[X] \approx npq \approx \lambda$ (since $np \approx \lambda$ and $q = 1-p \approx 1$). Can we show directly that $\mathrm{Var}[X] = \lambda$?
- Compute

$$E[X^2] = \sum_{k=0}^{\infty} P\{X=k\}k^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!}e^{-\lambda} = \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}.$$

- Setting $j = k-1$, this is

$$\lambda \left(\sum_{j=0}^{\infty} (j+1) \frac{\lambda^j}{j!} e^{-\lambda} \right) = \lambda E[X+1] = \lambda(\lambda+1).$$

Variance

- Given $P\{X=k\} = \frac{\lambda^k}{k!}e^{-\lambda}$ for integer $k \geq 0$, what is $\mathrm{Var}[X]$?
- Think of X as (roughly) a Bernoulli (n,p) random variable with n very large and $p = \lambda/n$.
- This suggests $\mathrm{Var}[X] \approx npq \approx \lambda$ (since $np \approx \lambda$ and $q = 1-p \approx 1$). Can we show directly that $\mathrm{Var}[X] = \lambda$?
- Compute

$$E[X^2] = \sum_{k=0}^{\infty} P\{X=k\}k^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!}e^{-\lambda} = \lambda \sum_{k=1}^{\infty} k \frac{\lambda^{k-1}}{(k-1)!}e^{-\lambda}.$$

- Setting $j = k-1$, this is

$$\lambda \left(\sum_{j=0}^{\infty} (j+1)\frac{\lambda^j}{j!}e^{-\lambda} \right) = \lambda E[X+1] = \lambda(\lambda+1).$$

- Then $\mathrm{Var}[X] = E[X^2] - E[X]^2 = \lambda(\lambda+1) - \lambda^2 = \lambda$.

Poisson convergence

- Idea: if we have lots of independent random events, each with very small probability to occur, and expected number to occur is λ, then total number that occur is roughly Poisson λ.

Poisson convergence

- Idea: if we have lots of independent random events, each with very small probability to occur, and expected number to occur is λ, then total number that occur is roughly Poisson λ.
- **Theorem:** Let $X_{n,m}$ be independent $\{0,1\}$-valued random variables with $P(X_{n,m}=1)=p_{n,m}$. Suppose $\sum_{m=1}^{n} p_{n,m} \to \lambda$ and $\max_{1 \le m \le n} p_{n,m} \to 0$. Then $S_n = X_{n,1} + \ldots + X_{n,n} \implies Z$ were Z is Poisson(λ).

Poisson convergence

- Idea: if we have lots of independent random events, each with very small probability to occur, and expected number to occur is λ, then total number that occur is roughly Poisson λ.
- **Theorem:** Let $X_{n,m}$ be independent $\{0,1\}$-valued random variables with $P(X_{n,m} = 1) = p_{n,m}$. Suppose $\sum_{m=1}^{n} p_{n,m} \to \lambda$ and $\max_{1 \leq m \leq n} p_{n,m} \to 0$. Then $S_n = X_{n,1} + \ldots + X_{n,n} \implies Z$ were Z is Poisson(λ).
- **Proof idea:** Just write down the log characteristic functions for Bernoulli and Poisson random variables. Check the conditions of the continuity theorem.

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Outline

CLT idea

CLT variants

More on random walks and local CLT

Poisson random variable convergence

Extend CLT idea to stable random variables

Recall continuity theorem

- **Strong continuity theorem:** If $\mu_n \implies \mu_\infty$ then $\phi_n(t) \to \phi_\infty(t)$ for all t. Conversely, if $\phi_n(t)$ converges to a limit that is continuous at 0, then the associated sequence of distributions μ_n is tight and converges weakly to a measure μ with characteristic function ϕ.

Recall CLT idea

- Let X be a random variable.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi'_X(0) = 0$ and $\phi''_X(0) = -1$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi'_X(0) = 0$ and $\phi''_X(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L'_X(0) = -\phi'_X(0)/\phi_X(0) = 0$ and $L''_X = -(\phi''_X(0)\phi_X(0) - \phi'_X(0)^2)/\ \phi_X(0)^2 = 1$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L_X'(0) = -\phi_X'(0)/\phi_X(0) = 0$ and $L_X'' = -(\phi_X''(0)\phi_X(0) - \phi_X'(0)^2)/\ \phi_X(0)^2 = 1$.
- If $V_n = n^{-1/2}\sum_{i=1}^n X_i$ where X_i are i.i.d. with law of X, then $L_{V_n}(t) = nL_X(n^{-1/2}t)$.

Recall CLT idea

- Let X be a random variable.
- The **characteristic function** of X is defined by $\phi(t) = \phi_X(t) := E[e^{itX}]$.
- And if X has an mth moment then $E[X^m] = i^m \phi_X^{(m)}(0)$.
- In particular, if $E[X] = 0$ and $E[X^2] = 1$ then $\phi_X(0) = 1$ and $\phi_X'(0) = 0$ and $\phi_X''(0) = -1$.
- Write $L_X := -\log \phi_X$. Then $L_X(0) = 0$ and $L_X'(0) = -\phi_X'(0)/\phi_X(0) = 0$ and $L_X'' = -(\phi_X''(0)\phi_X(0) - \phi_X'(0)^2)/\ \phi_X(0)^2 = 1$.
- If $V_n = n^{-1/2}\sum_{i=1}^n X_i$ where X_i are i.i.d. with law of X, then $L_{V_n}(t) = nL_X(n^{-1/2}t)$.
- When we zoom in on a twice differentiable function near zero (scaling vertically by n and horizontally by $\sqrt{n}$) the picture looks increasingly like a parabola.

Stable laws

- Question? Is it possible for something like a CLT to hold if X has infinite variance? Say we write $V_n = n^{-a} \sum_{i=1}^{n} X_i$ for some a. Could the law of these guys converge to something non-Gaussian?

Stable laws

- Question? Is it possible for something like a CLT to hold if X has infinite variance? Say we write $V_n = n^{-a} \sum_{i=1}^n X_i$ for some a. Could the law of these guys converge to something non-Gaussian?
- What if the L_{V_n} converge to something else as we increase n, maybe to some other power of $|t|$ instead of $|t|^2$?

Stable laws

- Question? Is it possible for something like a CLT to hold if X has infinite variance? Say we write $V_n = n^{-a} \sum_{i=1}^n X_i$ for some a. Could the law of these guys converge to something non-Gaussian?
- What if the L_{V_n} converge to something else as we increase n, maybe to some other power of $|t|$ instead of $|t|^2$?
- The the appropriately normalized sum should be converge in law to something with characteristic function $e^{-|t|^\alpha}$ instead of $e^{-|t|^2}$.

Stable laws

- Question? Is it possible for something like a CLT to hold if X has infinite variance? Say we write $V_n = n^{-a} \sum_{i=1}^n X_i$ for some a. Could the law of these guys converge to something non-Gaussian?
- What if the L_{V_n} converge to something else as we increase n, maybe to some other power of $|t|$ instead of $|t|^2$?
- The the appropriately normalized sum should be converge in law to something with characteristic function $e^{-|t|^\alpha}$ instead of $e^{-|t|^2}$.
- We already saw that this should work for Cauchy random variables. What's the characteristic function in that case?

Stable laws

- Question? Is it possible for something like a CLT to hold if X has infinite variance? Say we write $V_n = n^{-a} \sum_{i=1}^n X_i$ for some a. Could the law of these guys converge to something non-Gaussian?
- What if the L_{V_n} converge to something else as we increase n, maybe to some other power of $|t|$ instead of $|t|^2$?
- The the appropriately normalized sum should be converge in law to something with characteristic function $e^{-|t|^\alpha}$ instead of $e^{-|t|^2}$.
- We already saw that this should work for Cauchy random variables. What's the characteristic function in that case?
- Let's look up stable distributions.

Infinitely divisible laws

- Say a random variable X is **infinitely divisible**, for each n, there is a random variable Y such that X has the same law as the sum of n i.i.d. copies of Y.

Infinitely divisible laws

- Say a random variable X is **infinitely divisible**, for each n, there is a random variable Y such that X has the same law as the sum of n i.i.d. copies of Y.
- What random variables are infinitely divisible?

Infinitely divisible laws

- Say a random variable X is **infinitely divisible**, for each n, there is a random variable Y such that X has the same law as the sum of n i.i.d. copies of Y.
- What random variables are infinitely divisible?
- Poisson, Cauchy, normal, stable, etc.

Infinitely divisible laws

- Say a random variable X is **infinitely divisible**, for each n, there is a random variable Y such that X has the same law as the sum of n i.i.d. copies of Y.
- What random variables are infinitely divisible?
- Poisson, Cauchy, normal, stable, etc.
- Let's look at the characteristic functions of these objects. What about compound Poisson random variables (linear combinations of Poisson random variables)? What are their characteristic functions like?

Infinitely divisible laws

- Say a random variable X is **infinitely divisible**, for each n, there is a random variable Y such that X has the same law as the sum of n i.i.d. copies of Y.
- What random variables are infinitely divisible?
- Poisson, Cauchy, normal, stable, etc.
- Let's look at the characteristic functions of these objects. What about compound Poisson random variables (linear combinations of Poisson random variables)? What are their characteristic functions like?
- More general constructions are possible via Lévy Khintchine representation.