Math 341: Probability
Eighteenth Lecture (11/12/09)

Steven J Miller
Williams College

Steven.J.Miller@williams.edu
http://www.williams.edu/go/math/sjmiller/public_html/341/

Bronfman Science Center
Williams College, November 12, 2009

Summary for the Day

Summary for the day

- Complex Analysis:
  - Review definitions / statements.
  - Accumulation point theorems.

- Clicker question:
  - Statement of the CLT.
  - Clicker question on rate of convergence.

- Central Limit Theorem:
  - Poisson example.
  - Proof with MGFs.
  - Proof with Fourier analysis.

Review

Accumulation points and functions

Theorem

Let f be an analytic function on an open set U, with infinitely many zeros $z_1, z_2, z_3, \dots$. If $\lim_{n\to\infty} z_n \in U$, then f is identically zero on U. In other words, if a function is zero along a sequence in U whose accumulation point is also in U, then that function is identically zero in U.

Schwartz Space

Schwartz space $\mathcal{S}(\mathbb{R})$: all infinitely differentiable functions f such that, for any non-negative integers m and n,

$$\sup_{x\in\mathbb{R}} \left| (1+x^2)^m \frac{d^n f}{dx^n} \right| < \infty.$$

Inversion Theorem for Fourier Transform: Let $f \in \mathcal{S}(\mathbb{R})$. Then

$$f(x) = \int_{-\infty}^{\infty} \widehat{f}(y) e^{2\pi i x y} dy.$$

$f, g \in \mathcal{S}(\mathbb{R})$ with $\widehat{f} = \widehat{g}$ then $f(x) = g(x)$.

Accumulation and Moments

Accumulation Points and Moments

Theorem

Assume the MGFs $M_X(t)$ and $M_Y(t)$ exist in a neighborhood of zero (i.e., there is some δ such that both functions exist for $|t| < \delta$). If $M_X(t) = M_Y(t)$ in this neighborhood, then $F_X(u) = F_Y(u)$ for all u. As the densities are the derivatives of the cumulative distribution functions, we have $f = g$.

Accumulation Points and Moments

Theorem

Let $\{X_i\}_{i \in I}$ be a sequence of random variables with MGFs $M_{X_i}(t)$. Assume there is a $\delta > 0$ such that when $|t| < \delta$ we have $\lim_{i \to \infty} M_{X_i}(t) = M_X(t)$ for some MGF $M_X(t)$, and all MGFs converge for $|t| < \delta$. Then there exists a unique cumulative distribution function F whose moments are determined from $M_X(t)$ and for all x where $F_X(x)$ is continuous, $\lim_{n \to \infty} F_{X_i}(x) = F_X(x)$.

Accumulation Points and Moments

Theorem: X and Y continuous random variables on $[0, \infty)$ with continuous densities f and g, all of whose moments are finite and agree, and

1. $\exists C > 0$ st $\forall c \leq C$, $e^{(c+1)t} f(e^t)$ and $e^{(c+1)t} g(e^t)$ are Schwartz functions.
2. The (not necessarily integral) moments

$$\mu'_{r_n}(f) = \int_0^\infty x^{r_n} f(x)dx \quad \text{and} \quad \mu'_{r_n}(g) = \int_0^\infty x^{r_n} g(x)dx$$

   agree for some sequence of non-negative real numbers $\{r_n\}_{n=0}^\infty$ which has a finite accumulation point (i.e., $\lim_{n\to\infty} r_n = r < \infty$).

Then $f = g$ (in other words, knowing all these moments uniquely determines the probability density).

Application to equal integral moments

Return to the two densities causing trouble:

$$\begin{aligned} f_1(x) &= \frac{1}{\sqrt{2\pi x^2}}\, e^{-(\log^2 x)/2} \\ f_2(x) &= f_1(x)\left[1 + \sin(2\pi \log x)\right]. \end{aligned}$$

Application to equal integral moments

Return to the two densities causing trouble:

$$
\begin{aligned}
f_1(x) &= \frac{1}{\sqrt{2\pi x^2}}\, e^{-(\log^2 x)/2} \\
f_2(x) &= f_1(x)\,[1+\sin(2\pi \log x)]\,.
\end{aligned}
$$

- Same integral moments: $e^{k^2/2}$.
- Have the correct decay.
- Using complex analysis (specifically, contour integration), we can calculate the $(a+ib)^{\text{th}}$moments:

$$\text{For } f_1 : \quad e^{(a+ib)^2/2}$$

$$\text{For } f_2 : \quad e^{(a+ib)^2/2} + \frac{i}{2}\left(e^{(a+i(b-2\pi))^2/2} - e^{(a+i(b+2\pi))^2/2}\right).$$

Application to equal integral moments

Return to the two densities causing trouble:

$$\begin{aligned} f_1(x) &= \frac{1}{\sqrt{2\pi x^2}}\, e^{-(\log^2 x)/2} \\ f_2(x) &= f_1(x)\left[1+\sin(2\pi \log x)\right]. \end{aligned}$$

- No sequence of real moments having an accumulation point where they agree.
- a^{th}moment of f_2 is

$$e^{a^2/2} + e^{(a-2i\pi)^2/2}\left(1-e^{4ia\pi}\right),$$

  and this is never zero unless a is a half-integer.
- Only way this can vanish is if $1 = e^{4ia\pi}$.

Clicker Questions

Normalization of a random variable

Normalization (standardization) of a random variable

Let X be a random variable with mean μ and standard deviation σ, both of which are finite. The normalization, Y, is defined by

$$Y \;:=\; \frac{X - \mathbb{E}[X]}{\mathrm{StDev}(X)} \;=\; \frac{X - \mu}{\sigma}.$$

Note that

$$\mathbb{E}[Y] \;=\; 0 \quad \text{and} \quad \mathrm{StDev}(Y) \;=\; 1.$$

Statement of the Central Limit Theorem

Normal distribution

A random variable X is normally distributed (or has the normal distribution, or is a Gaussian random variable) with mean μ and variance σ^2 if the density of X is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

We often write $X \sim N(\mu, \sigma^2)$ to denote this. If $\mu = 0$ and $\sigma^2 = 1$, we say X has the standard normal distribution.

Statement of the Central Limit Theorem

Central Limit Theorem

Let $X_1, \dots, X_N$ be independent, identically distributed random variables whose moment generating functions converge for $|t| < \delta$ for some $\delta > 0$ (this implies all the moments exist and are finite). Denote the mean by μ and the variance by σ^2, let

$$\overline{X}_N = \frac{X_1 + \cdots + X_N}{N}$$

and set

$$Z_N = \frac{\overline{X}_N - \mu}{\sigma/\sqrt{N}}.$$

Then as $N \to \infty$, the distribution of Z_N converges to the standard normal.

Statement of the Central Limit Theorem

Why are there only tables of values of standard normal?

Statement of the Central Limit Theorem

Why are there only tables of values of standard normal?

Answer: normalization. Similar to log tables (only need one from change of base formula).

Alternative Statement of the Central Limit Theorem

Central Limit Theorem

Let $X_1, \dots, X_N$ be independent, identically distributed random variables whose moment generating functions converge for $|t| < \delta$ for some $\delta > 0$ (this implies all the moments exist and are finite). Denote the mean by μ and the variance by σ^2, let

$$S_N = X_1 + \cdots + X_N$$

and set

$$Z_N = \frac{S_N - N\mu}{\sqrt{N\sigma^2}}.$$

Then as $N \to \infty$, the distribution of Z_N converges to the standard normal.

Key Probabilities:

Key probabilities for $Z \sim N(0, 1)$ (i.e., Z has the standard normal distribution).

- $\text{Prob}(|Z| \le 1) \approx 68.2\%$.
- $\text{Prob}(|Z| \le 1.96) \approx 95\%$.
- $\text{Prob}(|Z| \le 2.575) \approx 99\%$.

Convergence to the standard normal

Question:

Let $X_1, X_2, \dots$ be iidrv with mean 0 and variance 1, and let $Z_N = \overline{X}_N/(1/\sqrt{N})$. By the CLT $Z_N \to N(0,1)$; which choice converges fastest? Slowest?

1. Uniform: $X \sim \mathrm{Unif}(-\sqrt{3}, \sqrt{3})$.
2. Laplace: $f_X(x) = e^{-\sqrt{2}|x|}/\sqrt{2}$.
3. Normal: $X \sim N(0,1)$.
4. Millered Cauchy: $f_X(x) = \frac{4a\sin(\pi/8)}{\pi} \frac{1}{1+(ax)^8}$, $a = \sqrt{\sqrt{2}-1}$.

Convergence to the standard normal

Question:

Let $X_1, X_2, \ldots$ be iidrv with mean 0 and variance 1, and let $Z_N = \overline{X}_N/(1/\sqrt{N})$. By the CLT $Z_N \to N(0,1)$; which choice converges fastest? Slowest?

1. Uniform: $X \sim \mathrm{Unif}(-\sqrt{3}, \sqrt{3})$. Kurtosis: 1.8.
2. Laplace: $f_X(x) = e^{-\sqrt{2}|x|}/\sqrt{2}$. Kurtosis: 6.
3. Normal: $X \sim N(0,1)$. Kurtosis: 3.
4. Millered Cauchy: $f_X(x) = \frac{4a\sin(\pi/8)}{\pi} \frac{1}{1+(ax)^8}$, $a = \sqrt{\sqrt{2}-1}$. Kurtosis: $1+\sqrt{2} \approx 2.414$.

Convergence to the standard normal

Question:

Let $X_1, X_2, \dots$ be iidrv with mean 0 and variance 1, and let $Z_N = \overline{X}_N/(1/\sqrt{N})$. By the CLT $Z_N \to N(0,1)$; which choice converges fastest? Slowest?

1. Uniform: $X \sim \mathrm{Unif}(-\sqrt{3}, \sqrt{3})$. Kurtosis: 1.8.
2. Laplace: $f_X(x) = e^{-\sqrt{2}|x|}/\sqrt{2}$. Kurtosis: 6.
3. Normal: $X \sim N(0,1)$. Kurtosis: 3.
4. Millered Cauchy: $f_X(x) = \frac{4a\sin(\pi/8)}{\pi}\,\frac{1}{1+(ax)^8}$, $a = \sqrt{\sqrt{2}-1}$. Kurtosis: $1+\sqrt{2} \approx 2.414$.

$$\log M_X(t) \;=\; \frac{t^2}{2} + \frac{(\mu_4-3)t^4}{4!} + O(t^6).$$

Convergence to the standard normal

Question:

Let $X_1, X_2, \dots$ be iidrv with mean 0 and variance 1, and let $Z_N = \overline{X}_N/(1/\sqrt{N})$. By the CLT $Z_N \to N(0,1)$; which choice converges fastest? Slowest?

1. Uniform: $X \sim \mathrm{Unif}(-\sqrt{3}, \sqrt{3})$. Excess Kurtosis: -1.2.
2. Laplace: $f_X(x) = e^{-\sqrt{2}|x|}/\sqrt{2}$. Excess Kurtosis: 3.
3. Normal: $X \sim N(0,1)$. Excess Kurtosis: 0.
4. Millered Cauchy: $f_X(x) = \frac{4a\sin(\pi/8)}{\pi}\frac{1}{1+(ax)^8}$, $a = \sqrt{\sqrt{2}-1}$. Excess Kurtosis: $1 + \sqrt{2} - 3 \approx -.586$.

$$\log M_X(t) = \frac{t^2}{2} + \frac{(\mu_4 - 3)t^4}{4!} + O(t^6).$$

Convergence to the standard normal

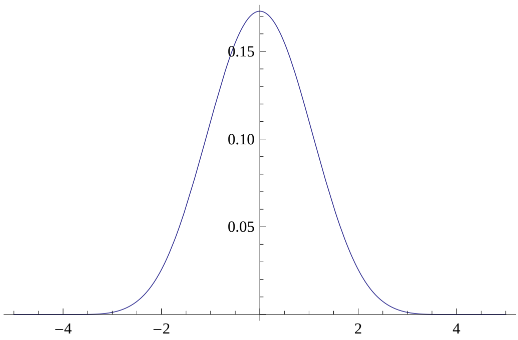

Figure: Convolutions of 5 Uniforms.

Convergence to the standard normal

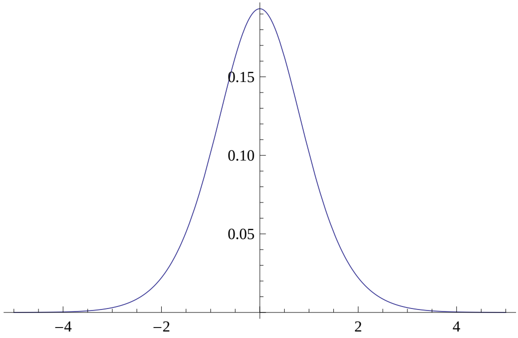

Figure: Convolutions of 5 Laplaces.

Convergence to the standard normal

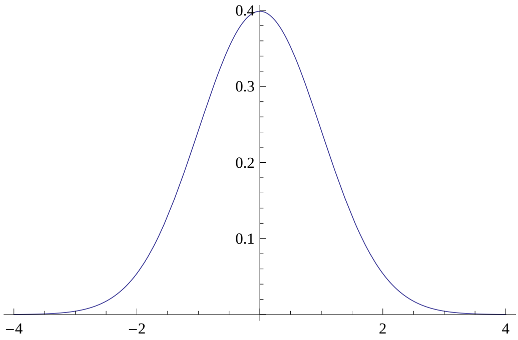

Figure: Convolutions of 5 Normals.

Convergence to the standard normal

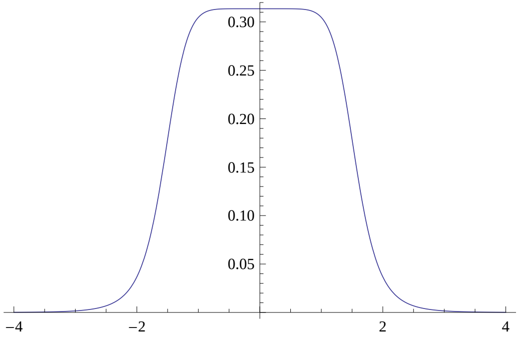

Figure: Convolutions of 1 Millered Cauchy.

Convergence to the standard normal

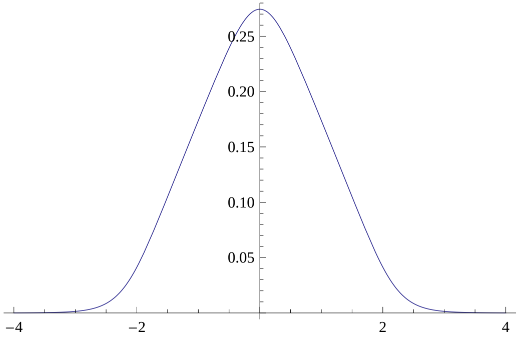

Figure: Convolutions of 2 Millered Cauchy.

Central Limit Theorem

MGF and the CLT

Moment generating function of normal distributions

Let X be a normal random variable with mean μ and variance σ^2. Its moment generating function satisfies

$$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

In particular, if Z has the standard normal distribution, its moment generating function is

$$M_Z(t) = e^{t^2/2}.$$

MGF and the CLT

Moment generating function of normal distributions

Let X be a normal random variable with mean μ and variance σ^2. Its moment generating function satisfies

$$M_X(t) \;=\; e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

In particular, if Z has the standard normal distribution, its moment generating function is

$$M_Z(t) \;=\; e^{t^2/2}.$$

Proof: Complete the square.

Poisson Example of the CLT

Example

Let $X, X_1, \ldots, X_N$ be Poisson random variables with parameter λ. Let

$$\overline{X}_N = \frac{X_1 + \cdots + X_N}{N}, \quad Y = \frac{\overline{X} - \mathbb{E}[\overline{X}]}{\mathrm{StDev}(\overline{X})}.$$

Then as $N \to \infty$, Y converges to having the standard normal distribution.

Poisson Example of the CLT

Example

Let $X, X_1, \dots, X_N$ be Poisson random variables with parameter λ. Let

$$\overline{X}_N = \frac{X_1 + \cdots + X_N}{N}, \quad Y = \frac{\overline{X} - \mathbb{E}[\overline{X}]}{\mathrm{StDev}(\overline{X})}.$$

Then as $N \to \infty$, Y converges to having the standard normal distribution.

Moment generating function: $M_X(t) = \exp(\lambda(e^t - 1))$.
Independent formula: $M_{X_1+X_2}(t) = M_{X_1}(t) M_{X_2}(t)$.
Shift formula: $M_{aX+b}(t) = e^{bt} M_X(at)$.

General proof via Moment Generating Functions

X_i's iidrv,

$$Z_N = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}} = \sum_{n=1}^{N} \frac{X_i - \mu}{\sigma\sqrt{N}}.$$

General proof via Moment Generating Functions

X_i's iidrv,

$$Z_N = \frac{\overline{X} - \mu}{\sigma/\sqrt{N}} = \sum_{n=1}^{N} \frac{X_i - \mu}{\sigma\sqrt{N}}.$$

Moment Generating Function is:

$$M_{Z_N}(t) = \prod_{n=1}^{N} e^{\frac{-\mu t}{\sigma\sqrt{N}}} M_X\left(\frac{t}{\sigma\sqrt{N}}\right) = e^{\frac{-\mu t\sqrt{N}}{\sigma}} M_X\left(\frac{t}{\sigma\sqrt{N}}\right)^N$$

General proof via Moment Generating Functions

X_i's iidrv,

$$Z_N = \frac{\overline{X}-\mu}{\sigma/\sqrt{N}} = \sum_{n=1}^{N} \frac{X_i-\mu}{\sigma\sqrt{N}}.$$

Moment Generating Function is:

$$M_{Z_N}(t) = \prod_{n=1}^{N} e^{\frac{-\mu t}{\sigma\sqrt{N}}} M_X\left(\frac{t}{\sigma\sqrt{N}}\right) = e^{\frac{-\mu t\sqrt{N}}{\sigma}} M_X\left(\frac{t}{\sigma\sqrt{N}}\right)^N$$

Taking logarithms:

$$\log M_{Z_N}(t) = -\frac{\mu t\sqrt{N}}{\sigma} + N \log M_X\left(\frac{t}{\sigma\sqrt{N}}\right).$$

General proof via Moment Generating Functions (cont)

Expansion of MGF:

$$M_X(t) \;=\; 1+\mu t+\frac{\mu_2' t^2}{2!}+\cdots \;=\; 1+t\left(\mu+\frac{\mu_2' t}{2}+\cdots\right).$$

General proof via Moment Generating Functions (cont)

Expansion of MGF:

$$M_X(t) = 1 + \mu t + \frac{\mu_2' t^2}{2!} + \cdots = 1 + t\left(\mu + \frac{\mu_2' t}{2} + \cdots\right).$$

Expansion for $\log(1+u)$ is

$$\log(1+u) = u - \frac{u^2}{2} + \frac{u^3}{3!} - \cdots.$$

General proof via Moment Generating Functions (cont)

Expansion of MGF:

$$M_X(t) \;=\; 1 + \mu t + \frac{\mu_2' t^2}{2!} + \cdots \;=\; 1 + t\left(\mu + \frac{\mu_2' t}{2} + \cdots\right).$$

Expansion for $\log(1+u)$ is

$$\log(1+u) \;=\; u - \frac{u^2}{2} + \frac{u^3}{3!} - \cdots.$$

Combining gives

$$\begin{aligned}\log M_X(t) \;&=\; t\left(\mu + \frac{\mu_2' t}{2} + \cdots\right) - \frac{t^2\left(\mu + \frac{\mu_2' t}{2} + \cdots\right)^2}{2} + \cdots \\ &=\; \mu t + \frac{\mu_2' - \mu^2}{2} t^2 + \text{terms in } t^3 \text{ or higher.}\end{aligned}$$

General proof via Moment Generating Functions (cont)

$$\begin{aligned}
&\log M_X\left(\frac{t}{\sigma\sqrt{N}}\right)\\
= \quad &\frac{\mu t}{\sigma\sqrt{N}} + \frac{\sigma^2}{2}\frac{t^2}{\sigma^2 N} + \text{terms in } t^3/N^{3/2} \text{ or lower in } N.
\end{aligned}$$

General proof via Moment Generating Functions (cont)

$$\begin{aligned}
&\log M_X\left(\frac{t}{\sigma\sqrt{N}}\right)\\
= &\quad \frac{\mu t}{\sigma\sqrt{N}} + \frac{\sigma^2}{2}\frac{t^2}{\sigma^2 N} + \text{terms in } t^3/N^{3/2} \text{ or lower in } N.
\end{aligned}$$

Denote lower order terms by $O(N^{-3/2})$. Collecting gives

$$\begin{aligned}
\log M_{Z_N}(t) &= -\frac{\mu t\sqrt{N}}{\sigma} + N\left(\frac{\mu t}{\sigma\sqrt{N}} + \frac{t^2}{2N} + O(N^{-3/2})\right)\\
&= -\frac{\mu t\sqrt{N}}{\sigma} + \frac{\mu t\sqrt{N}}{\sigma} + \frac{t^2}{2} + O(N^{-1/2})\\
&= \frac{t^2}{2} + O(N^{-1/2}).
\end{aligned}$$

Central Limit Theorem and Fourier Analysis

Convolutions

Convolution of f and g:

$$h(y) = (f * g)(y) = \int_{\mathbb{R}} f(x)g(y-x)\mathrm{d}x = \int_{\mathbb{R}} f(x-y)g(x)\mathrm{d}x.$$

Convolutions

Convolution of f and g:

$$h(y) = (f * g)(y) = \int_{\mathbb{R}} f(x)g(y-x)\mathrm{d}x = \int_{\mathbb{R}} f(x-y)g(x)\mathrm{d}x.$$

X_1 and X_2 independent random variables with probability density p.

$$\mathrm{Prob}(X_i \in [x, x+\Delta x]) = \int_x^{x+\Delta x} p(t)\mathrm{d}t \approx p(x)\Delta x.$$

$$\mathrm{Prob}(X_1 + X_2) \in [x, x+\Delta x] = \int_{x_1=-\infty}^{\infty} \int_{x_2=x-x_1}^{x+\Delta x - x_1} p(x_1)p(x_2)\mathrm{d}x_2\mathrm{d}x_1.$$

Convolutions

Convolution of f and g:

$$h(y) \;=\; (f*g)(y) \;=\; \int_{\mathbb{R}} f(x)g(y-x)\mathrm{d}x \;=\; \int_{\mathbb{R}} f(x-y)g(x)\mathrm{d}x.$$

X_1 and X_2 independent random variables with probability density p.

$$\mathrm{Prob}(X_i \in [x, x+\Delta x]) \;=\; \int_x^{x+\Delta x} p(t)\mathrm{d}t \;\approx\; p(x)\Delta x.$$

$$\mathrm{Prob}(X_1+X_2) \in [x, x+\Delta x] \;=\; \int_{x_1=-\infty}^{\infty} \int_{x_2=x-x_1}^{x+\Delta x - x_1} p(x_1)p(x_2)\mathrm{d}x_2\mathrm{d}x_1.$$

As $\Delta x \to 0$ we obtain the convolution of p with itself:

$$\mathrm{Prob}(X_1+X_2 \in [a,b]) \;=\; \int_a^b (p*p)(z)\mathrm{d}z.$$

Exercise to show non-negative and integrates to 1.

Statement of Central Limit Theorem

- WLOG p has mean zero, variance one, finite third moment and decays rapidly so all convolution integrals converge: p infinitely differentiable function satisfying

$$\int_{-\infty}^{\infty} xp(x)\mathrm{d}x = 0, \quad \int_{-\infty}^{\infty} x^2 p(x)\mathrm{d}x = 1, \quad \int_{-\infty}^{\infty} |x|^3 p(x)\mathrm{d}x < \infty.$$

- $X_1, X_2, \dots$ are iidrv with density p.
- Define $S_N = \sum_{i=1}^{N} X_i$.
- Standard Gaussian (mean zero, variance one) is $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Statement of Central Limit Theorem

- WLOG p has mean zero, variance one, finite third moment and decays rapidly so all convolution integrals converge: p infinitely differentiable function satisfying

$$\int_{-\infty}^{\infty} xp(x)\mathrm{d}x \;=\; 0,\;\; \int_{-\infty}^{\infty} x^2 p(x)\mathrm{d}x \;=\; 1,\;\; \int_{-\infty}^{\infty} |x|^3 p(x)\mathrm{d}x \;<\; \infty.$$

- $X_1, X_2, \dots$ are iidrv with density p.
- Define $S_N = \sum_{i=1}^{N} X_i$.
- Standard Gaussian (mean zero, variance one) is $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Central Limit Theorem Let X_i, S_N be as above and assume the third moment of each X_i is finite. Then $S_N/\sqrt{N}$ converges in probability to the standard Gaussian:

$$\lim_{N\to\infty} \mathrm{Prob}\left(\frac{S_N}{\sqrt{N}} \;\in\; [a,b]\right) \;=\; \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2}\mathrm{d}x.$$

Proof of the Central Limit Theorem

- The Fourier transform: $\widehat{p}(y) = \int_{-\infty}^{\infty} p(x) e^{-2\pi i x y} \mathrm{d}x$.

Proof of the Central Limit Theorem

- The Fourier transform: $\widehat{p}(y) \ = \ \int_{-\infty}^{\infty} p(x) e^{-2\pi i x y} \mathrm{d}x$.
- Derivative of $\widehat{g}$ is the Fourier transform of $-2\pi i x g(x)$; differentiation (hard) is converted to multiplication (easy).

$$\widehat{g}'(y) \ = \ \int_{-\infty}^{\infty} -2\pi i x \cdot g(x) e^{-2\pi i x y} \mathrm{d}x;$$

g prob. density, $\widehat{g}'(0) = -2\pi i \mathbb{E}[x]$, $\widehat{g}''(0) = -4\pi^2 \mathbb{E}[x^2]$.

Proof of the Central Limit Theorem

- The Fourier transform: $\widehat{p}(y) \ = \ \int_{-\infty}^{\infty} p(x) e^{-2\pi ixy} \mathrm{d}x.$
- Derivative of $\widehat{g}$ is the Fourier transform of $-2\pi ixg(x)$; differentiation (hard) is converted to multiplication (easy).

$$\widehat{g}'(y) \ = \ \int_{-\infty}^{\infty} -2\pi ix \cdot g(x) e^{-2\pi ixy} \mathrm{d}x;$$

  g prob. density, $\widehat{g}'(0) = -2\pi i\mathbb{E}[x]$, $\widehat{g}''(0) = -4\pi^2\mathbb{E}[x^2]$.
- Natural: mean and variance simple multiples of derivatives of $\widehat{p}$ at zero: $\widehat{p}'(0) = 0$, $\widehat{p}''(0) = -4\pi^2$.

Proof of the Central Limit Theorem

- The Fourier transform: $\widehat{p}(y) \ = \ \int_{-\infty}^{\infty} p(x)e^{-2\pi ixy}\mathrm{d}x$.
- Derivative of $\widehat{g}$ is the Fourier transform of $-2\pi ixg(x)$; differentiation (hard) is converted to multiplication (easy).

$$\widehat{g}'(y) \ = \ \int_{-\infty}^{\infty} -2\pi ix \cdot g(x)e^{-2\pi ixy}\mathrm{d}x;$$

  g prob. density, $\widehat{g}'(0) = -2\pi i\mathbb{E}[x]$, $\widehat{g}''(0) = -4\pi^2\mathbb{E}[x^2]$.

- Natural: mean and variance simple multiples of derivatives of $\widehat{p}$ at zero: $\widehat{p}'(0) = 0$, $\widehat{p}''(0) = -4\pi^2$.
- We Taylor expand $\widehat{p}$ (need technical conditions on p):

$$\widehat{p}(y) \ = \ 1 + \frac{p''(0)}{2}y^2 + \cdots \ = \ 1 - 2\pi^2y^2 + O(y^3).$$

  Near origin, $\widehat{p}$ a concave down parabola.

Proof of the Central Limit Theorem (cont)

- $\text{Prob}(X_1 + \cdots + X_N \in [a, b]) \;=\; \int_a^b (p * \cdots * p)(z)\mathrm{d}z.$

Proof of the Central Limit Theorem (cont)

- $\text{Prob}(X_1 + \cdots + X_N \in [a, b]) = \int_a^b (p * \cdots * p)(z)\mathrm{d}z$.
- The Fourier transform converts convolution to multiplication. If $\text{FT}[f](y)$ denotes the Fourier transform of f evaluated at y:

$$\text{FT}[p * \cdots * p](y) = \widehat{p}(y) \cdots \widehat{p}(y).$$

Proof of the Central Limit Theorem (cont)

- $\text{Prob}(X_1 + \cdots + X_N \in [a, b]) = \int_a^b (p * \cdots * p)(z)\text{d}z$.
- The Fourier transform converts convolution to multiplication. If $\text{FT}[f](y)$ denotes the Fourier transform of f evaluated at y:

$$\text{FT}[p * \cdots * p](y) = \widehat{p}(y) \cdots \widehat{p}(y).$$

- Do not want the distribution of $X_1 + \cdots + X_N = x$, but rather $S_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}} = x$.

Proof of the Central Limit Theorem (cont)

- $\text{Prob}(X_1 + \cdots + X_N \in [a, b]) \;=\; \int_a^b (p * \cdots * p)(z)\mathrm{d}z$.
- The Fourier transform converts convolution to multiplication. If $\text{FT}[f](y)$ denotes the Fourier transform of f evaluated at y:

$$\text{FT}[p * \cdots * p](y) \;=\; \widehat{p}(y) \cdots \widehat{p}(y).$$

- Do not want the distribution of $X_1 + \cdots + X_N = x$, but rather $S_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}} = x$.
- If $B(x) = A(cx)$ for some fixed $c \neq 0$, then $\widehat{B}(y) = \frac{1}{c}\widehat{A}\left(\frac{y}{c}\right)$.

Proof of the Central Limit Theorem (cont)

- $\mathrm{Prob}(X_1 + \cdots + X_N \in [a,b]) \ = \ \int_a^b (p * \cdots * p)(z)\mathrm{d}z$.
- The Fourier transform converts convolution to multiplication. If $\mathrm{FT}[f](y)$ denotes the Fourier transform of f evaluated at y:

$$\mathrm{FT}[p * \cdots * p](y) \ = \ \widehat{p}(y) \cdots \widehat{p}(y).$$

- Do not want the distribution of $X_1 + \cdots + X_N = x$, but rather $S_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}} = x$.
- If $B(x) = A(cx)$ for some fixed $c \neq 0$, then $\widehat{B}(y) = \frac{1}{c}\widehat{A}\left(\frac{y}{c}\right)$.
- $\mathrm{Prob}\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}} = x\right) \ = \ (\sqrt{N}p * \cdots * \sqrt{N}p)(x\sqrt{N})$.

Proof of the Central Limit Theorem (cont)

- $\text{Prob}(X_1 + \cdots + X_N \in [a, b]) \ = \ \int_a^b (p * \cdots * p)(z)\text{d}z.$
- The Fourier transform converts convolution to multiplication. If $\text{FT}[f](y)$ denotes the Fourier transform of f evaluated at y:

$$\text{FT}[p * \cdots * p](y) \ = \ \widehat{p}(y) \cdots \widehat{p}(y).$$

- Do not want the distribution of $X_1 + \cdots + X_N = x$, but rather $S_N = \frac{X_1 + \cdots + X_N}{\sqrt{N}} = x$.
- If $B(x) = A(cx)$ for some fixed $c \neq 0$, then $\widehat{B}(y) = \frac{1}{c}\widehat{A}\left(\frac{y}{c}\right)$.
- $\text{Prob}\left(\frac{X_1 + \cdots + X_N}{\sqrt{N}} = x\right) \ = \ (\sqrt{N}p * \cdots * \sqrt{N}p)(x\sqrt{N}).$
- $\text{FT}\left[(\sqrt{N}p * \cdots * \sqrt{N}p)(x\sqrt{N})\right](y) \ = \ \left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$

Proof of the Central Limit Theorem (cont)

- Can find the Fourier transform of the distribution of S_N:

$$\left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$$

Proof of the Central Limit Theorem (cont)

- Can find the Fourier transform of the distribution of S_N:

$$\left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$$

- Take the limit as $N \to \infty$ for **fixed** y.

Proof of the Central Limit Theorem (cont)

- Can find the Fourier transform of the distribution of S_N:

$$\left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$$

- Take the limit as $N \to \infty$ for **fixed** y.
- Know $\widehat{p}(y) = 1 - 2\pi^2 y^2 + O(y^3)$. Thus study

$$\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N.$$

Proof of the Central Limit Theorem (cont)

- Can find the Fourier transform of the distribution of S_N:

$$\left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$$

- Take the limit as $N \to \infty$ for **fixed** y.
- Know $\widehat{p}(y) = 1 - 2\pi^2 y^2 + O(y^3)$. Thus study

$$\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N.$$

- For any **fixed** y,

$$\lim_{N\to\infty}\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N = e^{-2\pi^2 y^2}.$$

Proof of the Central Limit Theorem (cont)

- Can find the Fourier transform of the distribution of S_N:

$$\left[\widehat{p}\left(\frac{y}{\sqrt{N}}\right)\right]^N.$$

- Take the limit as $N \to \infty$ for **fixed** y.
- Know $\widehat{p}(y) = 1 - 2\pi^2 y^2 + O(y^3)$. Thus study

$$\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N.$$

- For any **fixed** y,

$$\lim_{N\to\infty}\left[1 - \frac{2\pi^2 y^2}{N} + O\left(\frac{y^3}{N^{3/2}}\right)\right]^N = e^{-2\pi^2 y^2}.$$

- Fourier transform of $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ at y is $e^{-2\pi^2 y^2}$.

Proof of the Central Limit Theorem (cont)

We have shown:

- the Fourier transform of the distribution of S_N converges to $e^{-2\pi^2 y^2}$;
- the Fourier transform of $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ at y is $e^{-2\pi^2 y^2}$.

Therefore the distribution of S_N equalling x converges to $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

Proof of the Central Limit Theorem (cont)

We have shown:

- the Fourier transform of the distribution of S_N converges to $e^{-2\pi^2 y^2}$;
- the Fourier transform of $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ at y is $e^{-2\pi^2 y^2}$.

Therefore the distribution of S_N equalling x converges to $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.
We need complex analysis to justify this inversion. Must be careful: Consider

$$g(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

All the Taylor coefficients about $x = 0$ are zero, but the function is not identically zero in a neighborhood of $x = 0$.