CSE 312

Foundations of Computing II

Lecture 18: The Central Limit Theorem

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Aleks Jovcic

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Anna Karlin, Alex Tsun, Rachel Lin, Hunter Schafer & myself ☺

Agenda

- Practice with Normals
- Closure of the Normal
- The Central Limit Theorem (CLT)

$\Phi(k)$

Example – Off by Standard Deviations

$Z \sim N(0,1)$

Let $X \sim \mathcal{N}(\mu, \sigma^2)$. $\quad E[X] = \mu, \quad Var(X) = \sigma^2$

$\mathbb{P}(|X - \mu| < k\sigma) = \mathbb{P}\left(\frac{|X-\mu|}{\sigma} < k\right) \qquad \frac{X-\mu}{\sigma}$

$= \mathbb{P}\left(-k < \frac{X-\mu}{\sigma} < k\right) = \Phi(k) - \Phi(-k)$

$= \Phi(k) - (1 - \Phi(k))$

$\boxed{2\Phi(k) - 1}$

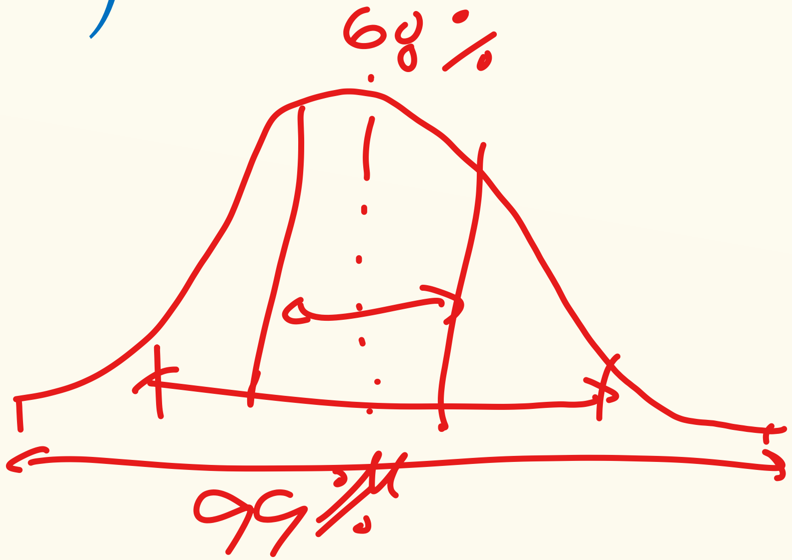
Example – Off by Standard Deviations

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.

$$\mathbb{P}(|X-\mu| < k\sigma) = \mathbb{P}\left(\frac{|X-\mu|}{\sigma} < k\right) =$$

$$= \mathbb{P}\left(-k < \frac{X-\mu}{\sigma} < k\right) = \Phi(k) - \Phi(-k)$$

e.g. $k = 1$: 68%, $k = 2$: 95%, $k = 3$: 99%



Summary of procedure for doing calculations with normal r.v.

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$

Therefore,

$$F_X(z) = \mathbb{P}(X \leq z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{z-\mu}{\sigma}\right) = \Phi\left(\frac{z-\mu}{\sigma}\right)$$

Agenda

- Practice with Normals
- Closure of the Normal
- The Central Limit Theorem (CLT)

CDF of normal distribution

Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Standard (unit) normal $\boldsymbol{Z} \sim \mathcal{N}(0, 1)$

CDF. $\Phi(z) = \mathbb{P}(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \mathrm{d}x$ for $Z \sim \mathcal{N}(0, 1)$

Note: $\Phi(z)$ has no closed form – generally given via tables

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $F_X(z) = \mathbb{P}(X \leq z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{z-\mu}{\sigma}\right) = \Phi(\frac{z-\mu}{\sigma})$

Closure of the normal -- under addition

Fact. If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ (both independent normal RV) then $aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_X^2 + b^2\sigma_Y^2)$

Note: The special thing is that the sum of normal **RVs is still a normal RV.**

The values of the expectation and variance is not surprising.

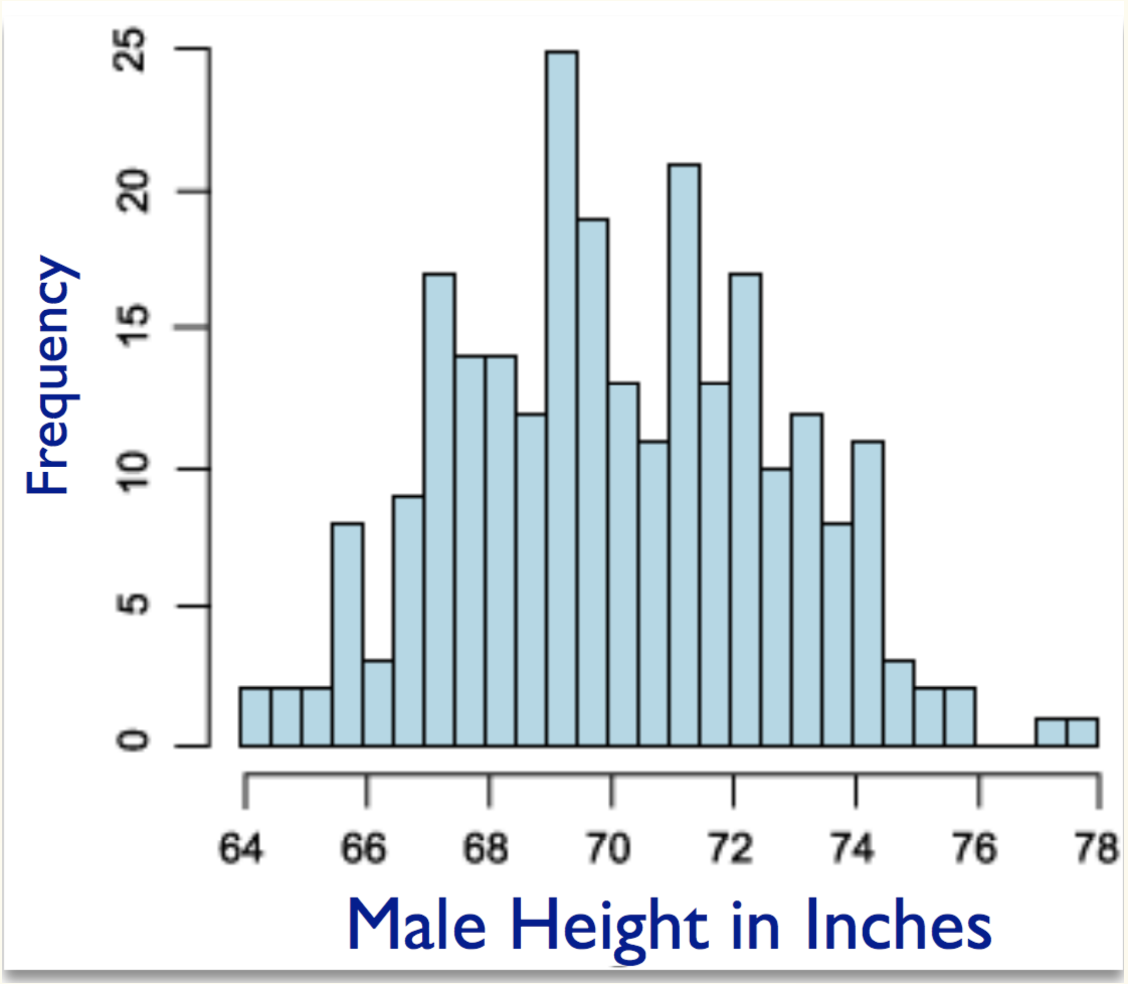
- Linearity of expectation (always true)
- When X and Y are independent, $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$

Agenda

- Practice with Normals
- Closure of the Normal
- The Central Limit Theorem (CLT)

Gaussian in Nature

Empirical distribution of collected data often resembles a Gaussian …



e.g. Height distribution resembles Gaussian.

R.A.Fisher (1918) observed that the height is likely the outcome of the sum of many independent random parameters, i.e., can written as

$$X = X_1 + \cdots + X_n$$

Sum of Independent RVs

i.i.d. = independent and identically distributed

$X_1, \dots, X_n$ i.i.d. with expectation μ and variance σ^2

Define

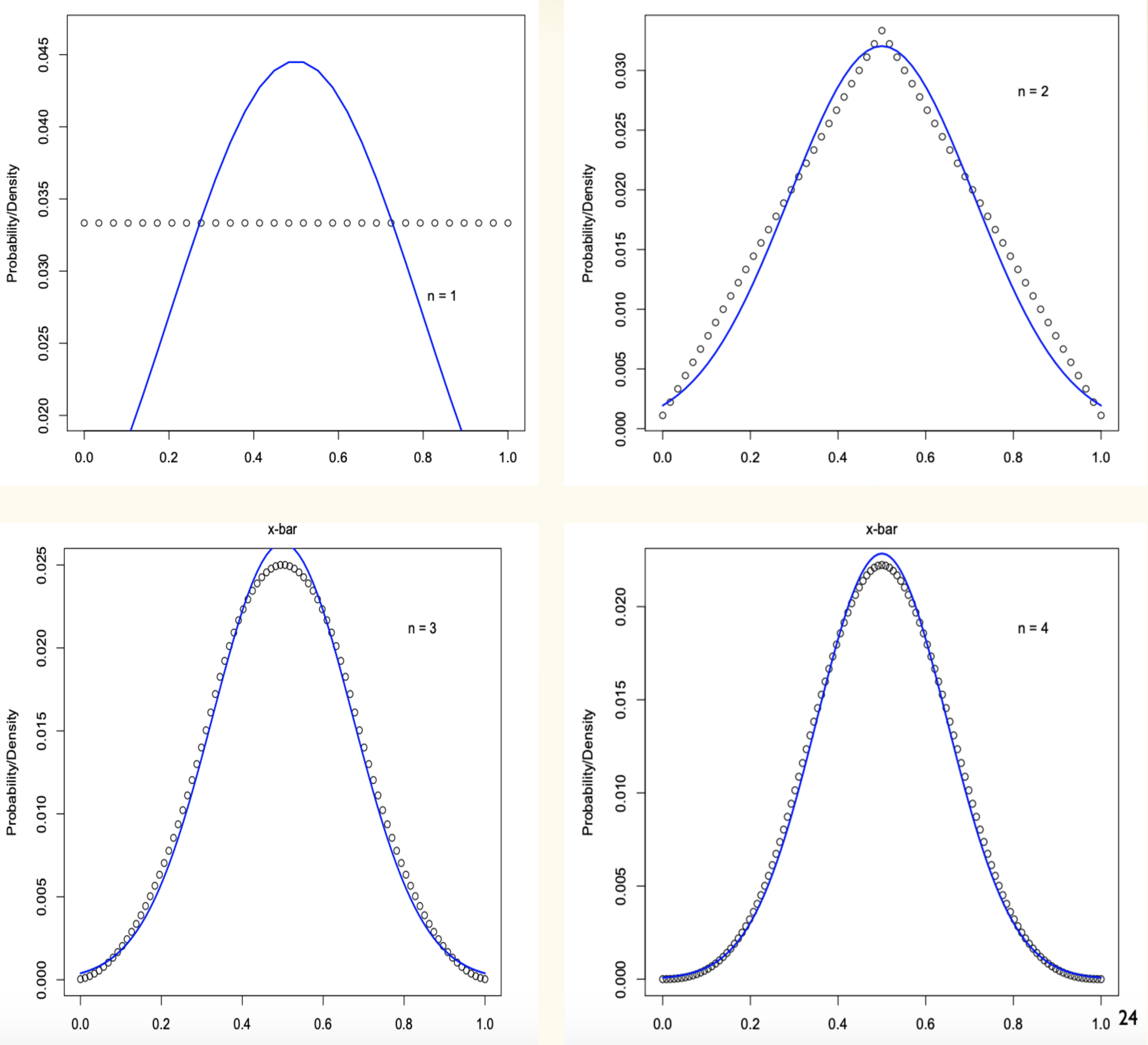
$$S_n = X_1 + \cdots + X_n$$

$$\mathbb{E}(S_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = n\mu$$

$$\mathrm{Var}(S_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = n\sigma^2$$

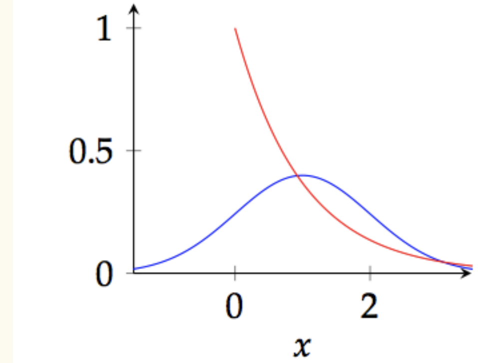
Empirical observation: S_n looks like a normal RV as n grows.

CLT (Idea)

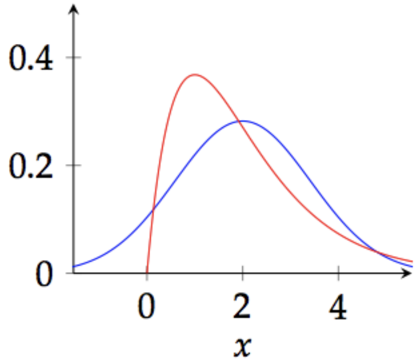


From: https://courses.cs.washington.edu/courses/cse312/17wi/slides/10limits.pdf

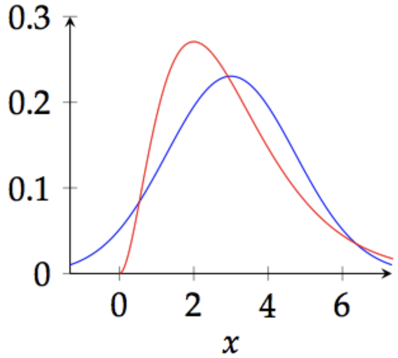
Sum of i.i.d. exponential random variables (param 1)



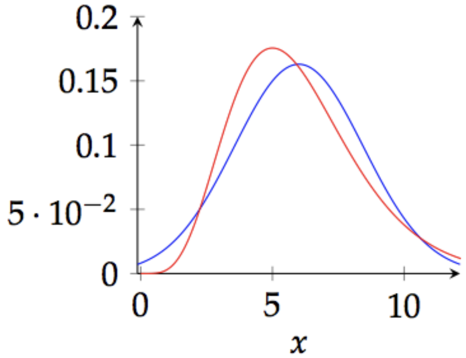
(a) $n = 1$



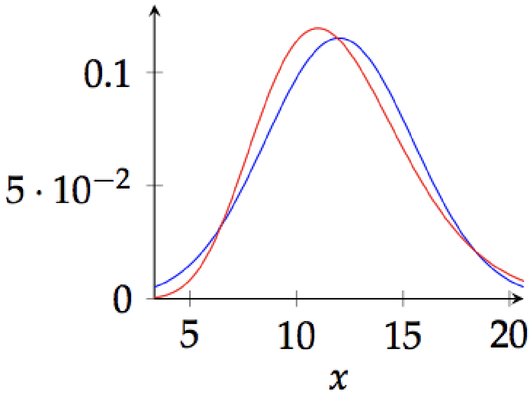
(b) $n = 2$



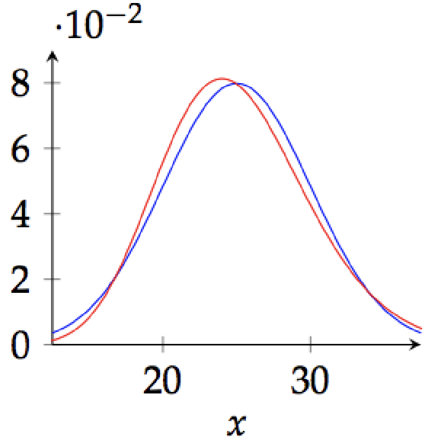
(c) $n = 3$



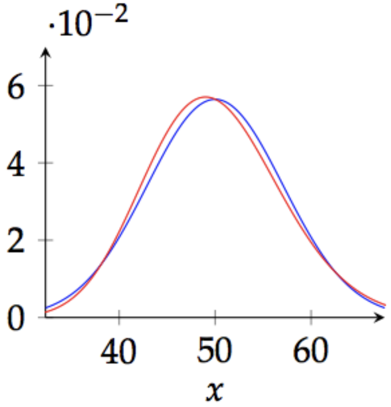
(d) $n = 6$



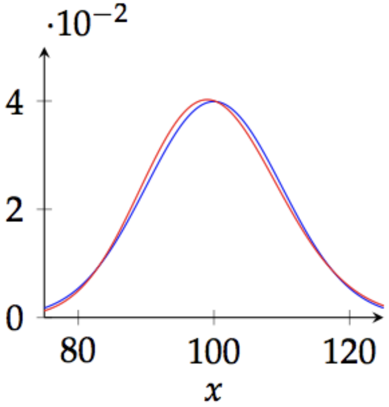
(e) $n = 12$



(f) $n = 25$

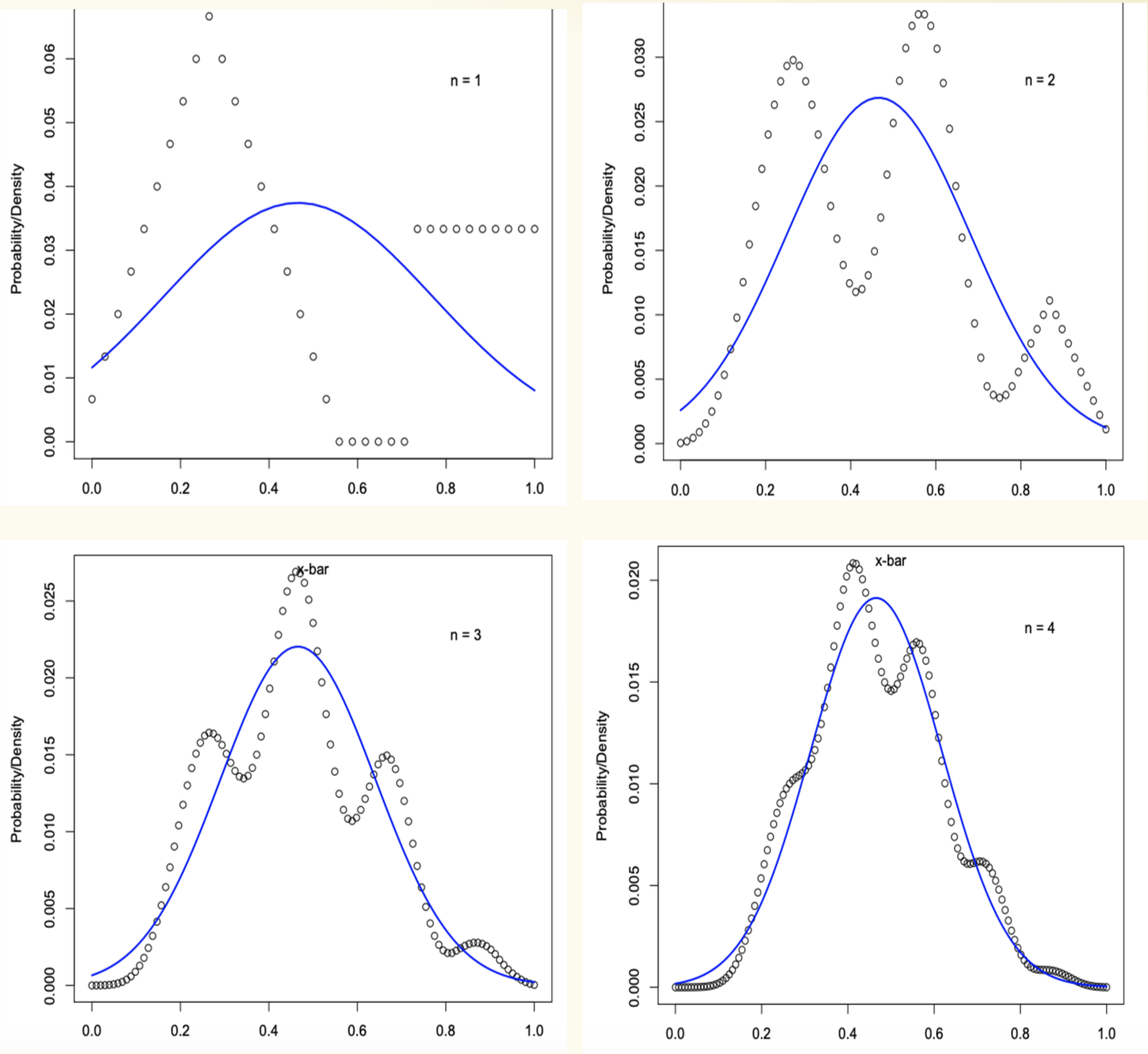


(g) $n = 50$



(h) $n = 100$

CLT (Idea)



From: https://courses.cs.washington.edu/courses/cse312/17wi/slides/10limits.pdf

Central Limit Theorem

$\mathrm{Var}(X+a) = \mathrm{Var}(X)$

$X_1, \dots, X_n$ i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \cdots + X_n$ and

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$\mathbb{E}(Y_n) = \mathbb{E}\left[\frac{S_n - n\mu}{\sigma\sqrt{n}}\right] = \frac{1}{\sigma\sqrt{n}}\left(\mathbb{E}[S_n] - n\mu\right) = 0$

$\mathrm{Var}(Y_n) = \mathrm{Var}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}}\right) = \frac{1}{\sigma^2 n}\mathrm{Var}(S_n) = \frac{\sigma^2 n}{\sigma^2 n}$

$= 1$

Central Limit Theorem

$X_1, \ldots, X_n$ i.i.d., each with expectation μ and variance σ^2

Define $S_n = X_1 + \cdots + X_n$ and

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$$\mathbb{E}(Y_n) = \frac{1}{\sigma\sqrt{n}}(\mathbb{E}(S_n) - n\mu) = \frac{1}{\sigma\sqrt{n}}(n\mu - n\mu) = 0$$

$$\mathrm{Var}(Y_n) = \frac{1}{\sigma^2 n}\big(\mathrm{Var}(S_n - n\mu)\big) = \frac{\mathrm{Var}(S_n)}{\sigma^2 n} = \frac{\sigma^2 n}{\sigma^2 n} = 1$$

Central Limit Theorem

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

Theorem. (Central Limit Theorem) The CDF of Y_n converges to the CDF of the standard normal $\mathcal{N}(0,1)$, i.e.,

$$\lim_{n\to\infty} \mathbb{P}(Y_n \le y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} \mathrm{d}x$$

Central Limit Theorem

$F_X(x) = \mathbb{P}(X \leq x)$

$$Y_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}$$

Theorem. (Central Limit Theorem) The CDF of Y_n converges to the CDF of the standard normal $\mathcal{N}(0,1)$, i.e.,

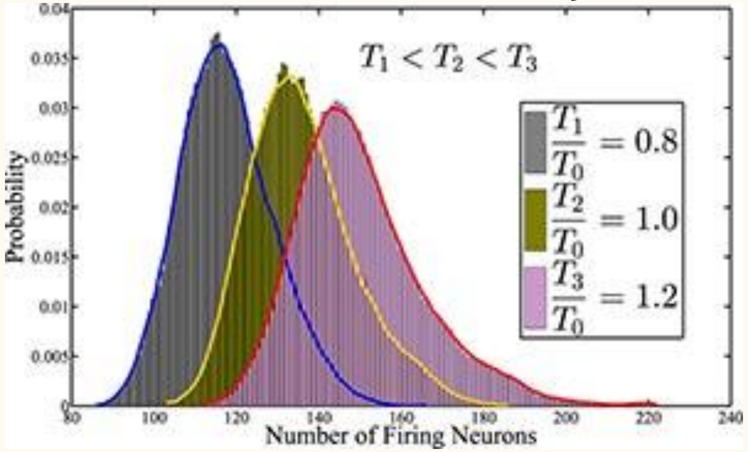
$$\lim_{n\to\infty} \mathbb{P}(Y_n \leq y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-x^2/2} \mathrm{d}x$$

Also stated as:

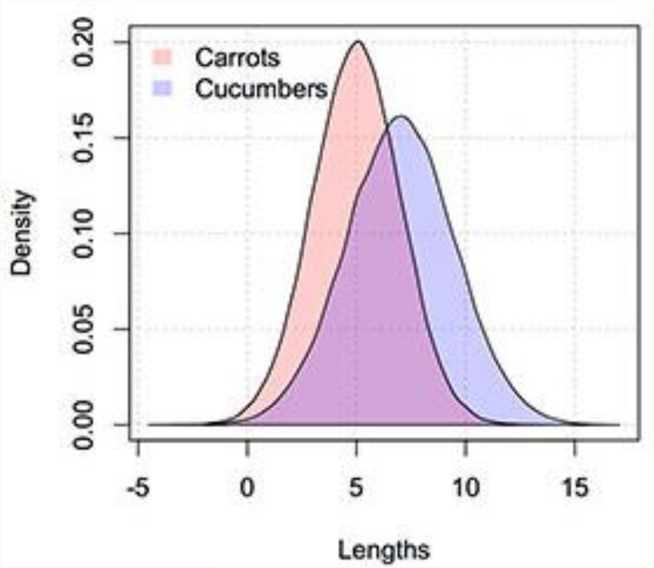
- $\lim_{n\to\infty} Y_n \to \mathcal{N}(0,1)$
- $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i \to \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$ where $\mu = E[X_i]$ and $\sigma^2 = Var(X_i)$

CLT → Normal Distribution EVERYWHERE

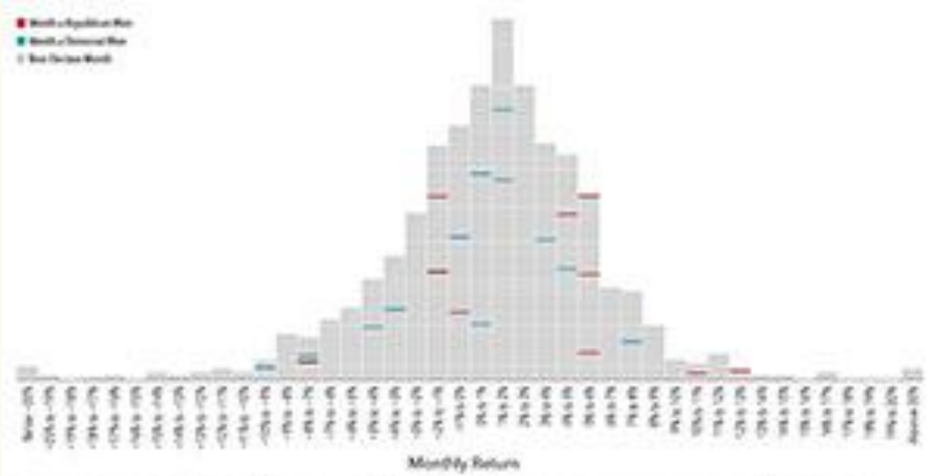
Neuron Activity



Vegetables



S&P 500 Returns after Elections



Examples from:
https://galtonboard.com/probabilityexamplesinlife

Example

$X \sim \text{Bin}(n, 0.75)$ $\quad X \approx N(7.5, 1.875)$

We flip n independent coins, heads with probability $p = 0.75$.

X = # heads $\qquad \mu = \mathbb{E}(X) = 0.75n \qquad \sigma^2 = \text{Var}(X) = 0.1875n$

$n = 10$ $\qquad 7.5 \qquad 1.875$

$\mathbb{P}(X \leq 7) = \mathbb{P}(0 \leq X \leq 7)$

$\mathbb{P}(X \leq 0.7n)$

$= \mathbb{P}\left(\frac{0 - 7.5}{\sqrt{1.875}} \leq \frac{X - 7.5}{\sqrt{1.875}} \leq \frac{-0.5}{\sqrt{1.875}}\right) = \Phi - \Phi$

Example

We flip n independent coins, heads with probability $p = 0.75$.

X = # heads $\quad \mu = \mathbb{E}(X) = 0.75n \quad \sigma^2 = \mathrm{Var}(X) = 0.1875n$

$\mathbb{P}(X \leq 0.7n)$

n	exact	$\mathcal{N}(\mu, \sigma^2)$ approx
10	0.4744072	0.357500327
20	0.38282735	0.302788308
50	0.25191886	0.207108089
100	0.14954105	0.124106539
200	0.06247223	0.051235217
1000	0.00019359	0.000130365