# Central Theorems

Chris Piech
CS109, Stanford University


And now a moment of silence...

...before we present...

...a beautiful result of probability theory!

# Central Limit Theorem

# The sum of independent, identically distributed variables:

$$Y = \sum_{i=0}^{n} X_i$$

Is normally distributed:

$$Y \sim N(n\mu, n\sigma^2)$$

---

where $\mu = E[X_i]$

$$\sigma^2 = \mathrm{Var}(X_i)$$

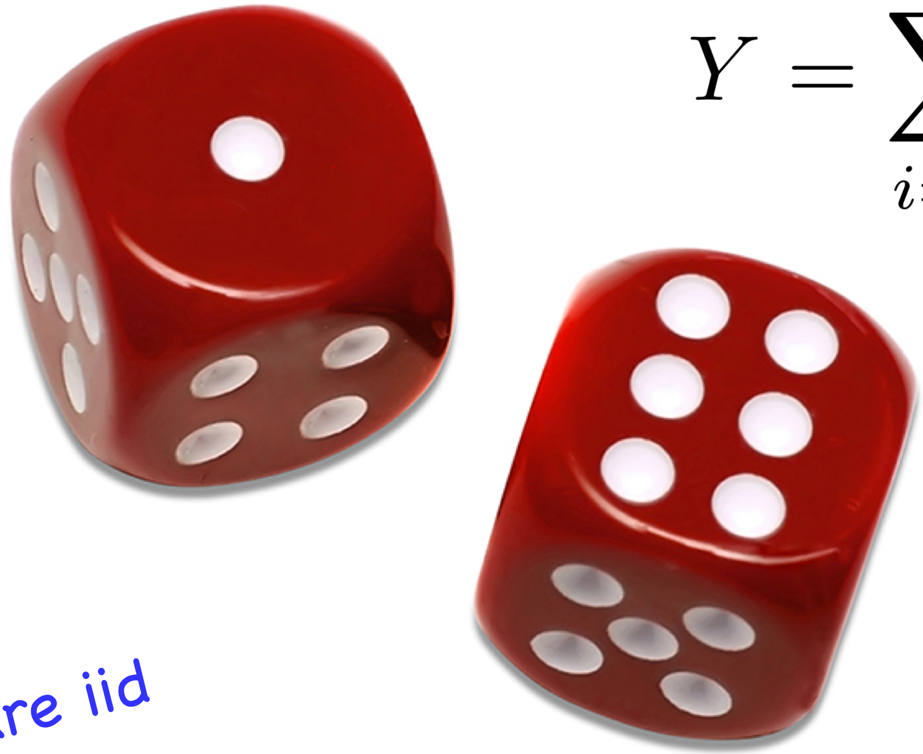
# IID Random Variables

- Consider $n$ random variables $X_1$, $X_2$, ... $X_n$
  - $X_i$ are all independently and identically distributed (I.I.D.)
  - All have the same PMF (if discrete) or PDF (if continuous)
  - All have the same expectation
  - All have the same variance

IID iid

# Sum of Two Dice

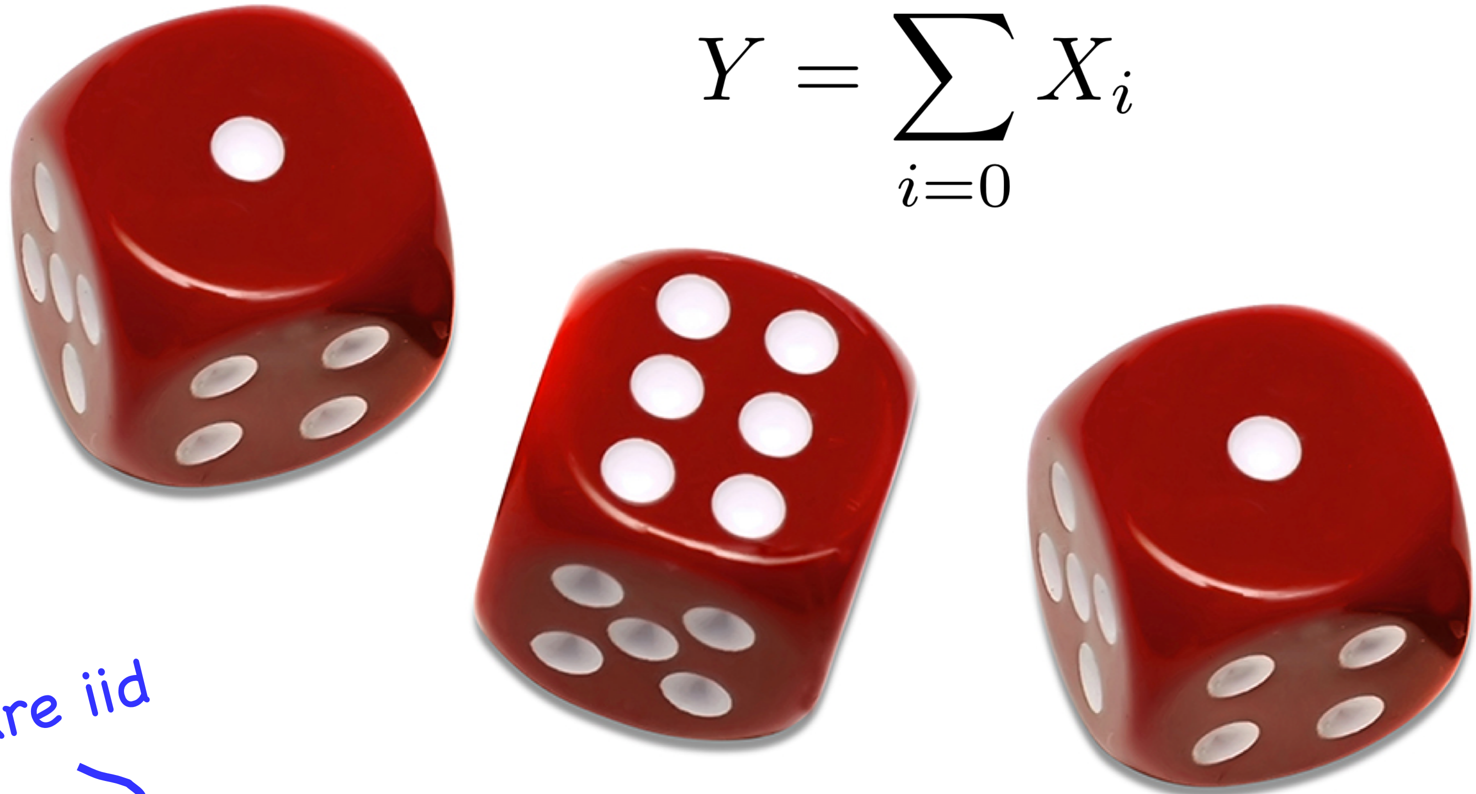
$$Y = \sum_{i=0}^{2} X_i$$



$X_i$ s are iid

$X_i$ is the outcome of dice roll $i$

# Sum of Three Dice



$$Y = \sum_{i=0}^{3} X_i$$

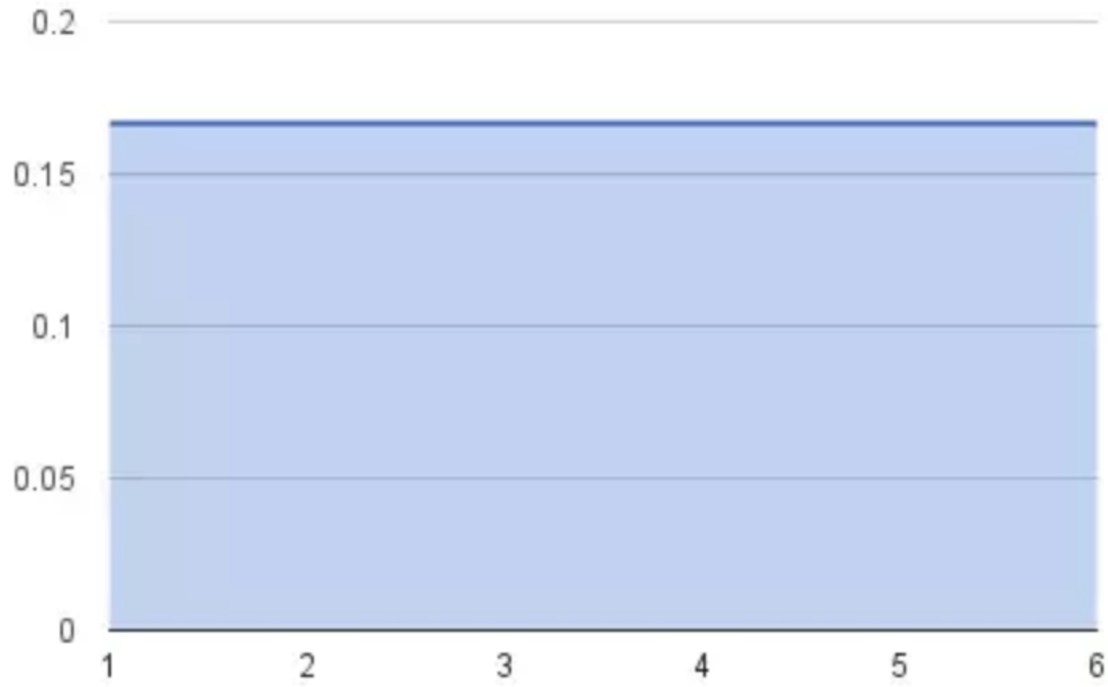
$X_i$ s are iid

$X_i$ is the outcome of dice roll $i$

Demo

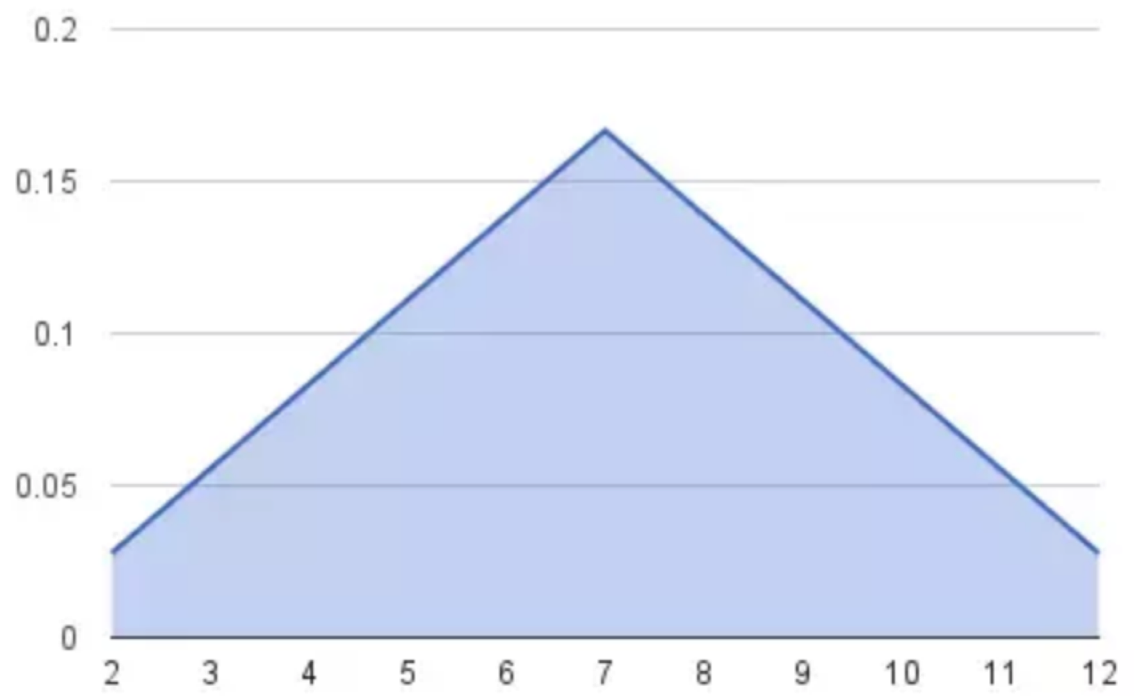
# C.L.T. Intuition

This is the PMF of the sum of one dice



# C.L.T. Intuition

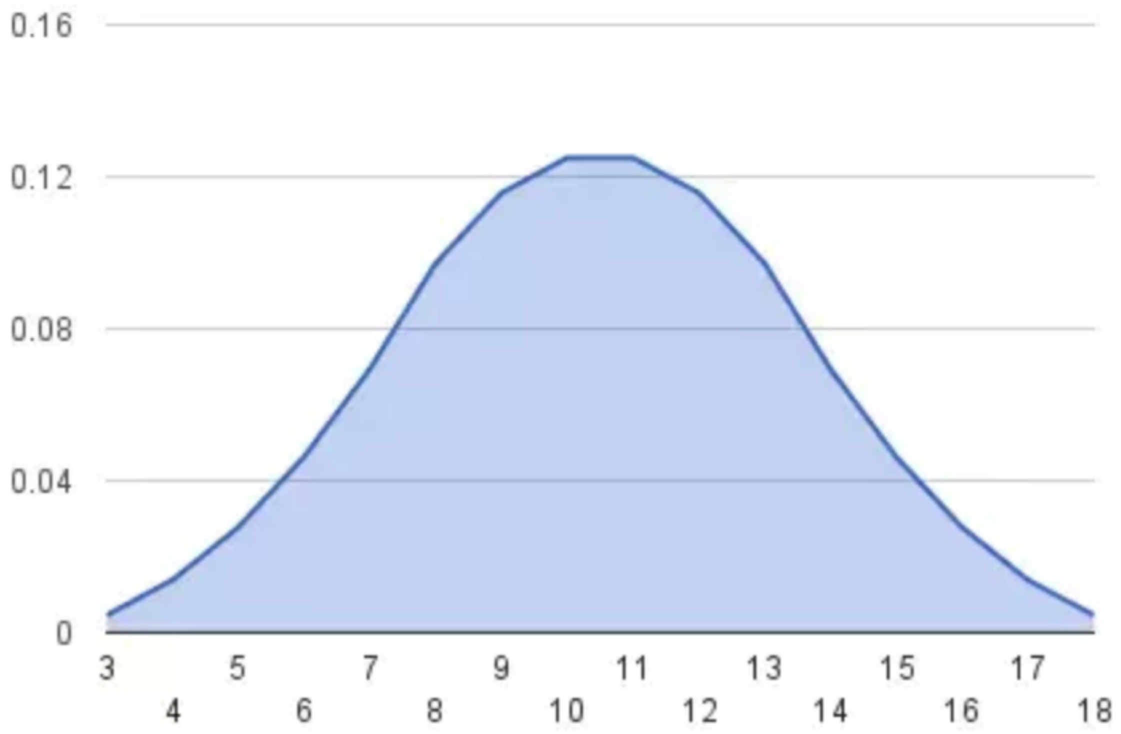
This is the PMF of the sum of two dice




Why is there more mass in the middle?

# C.L.T. Intuition

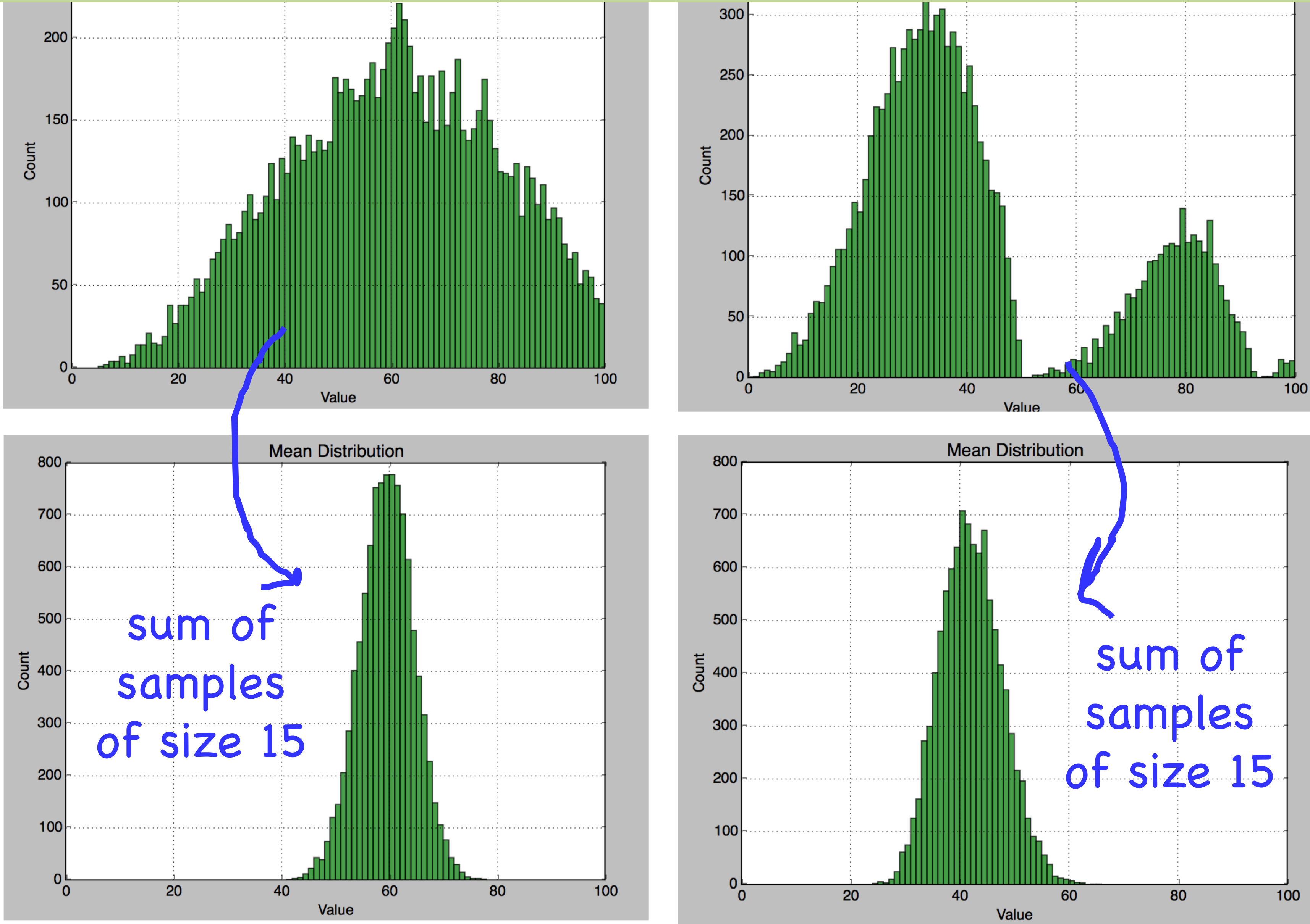
This is the PMF of the sum of three dice



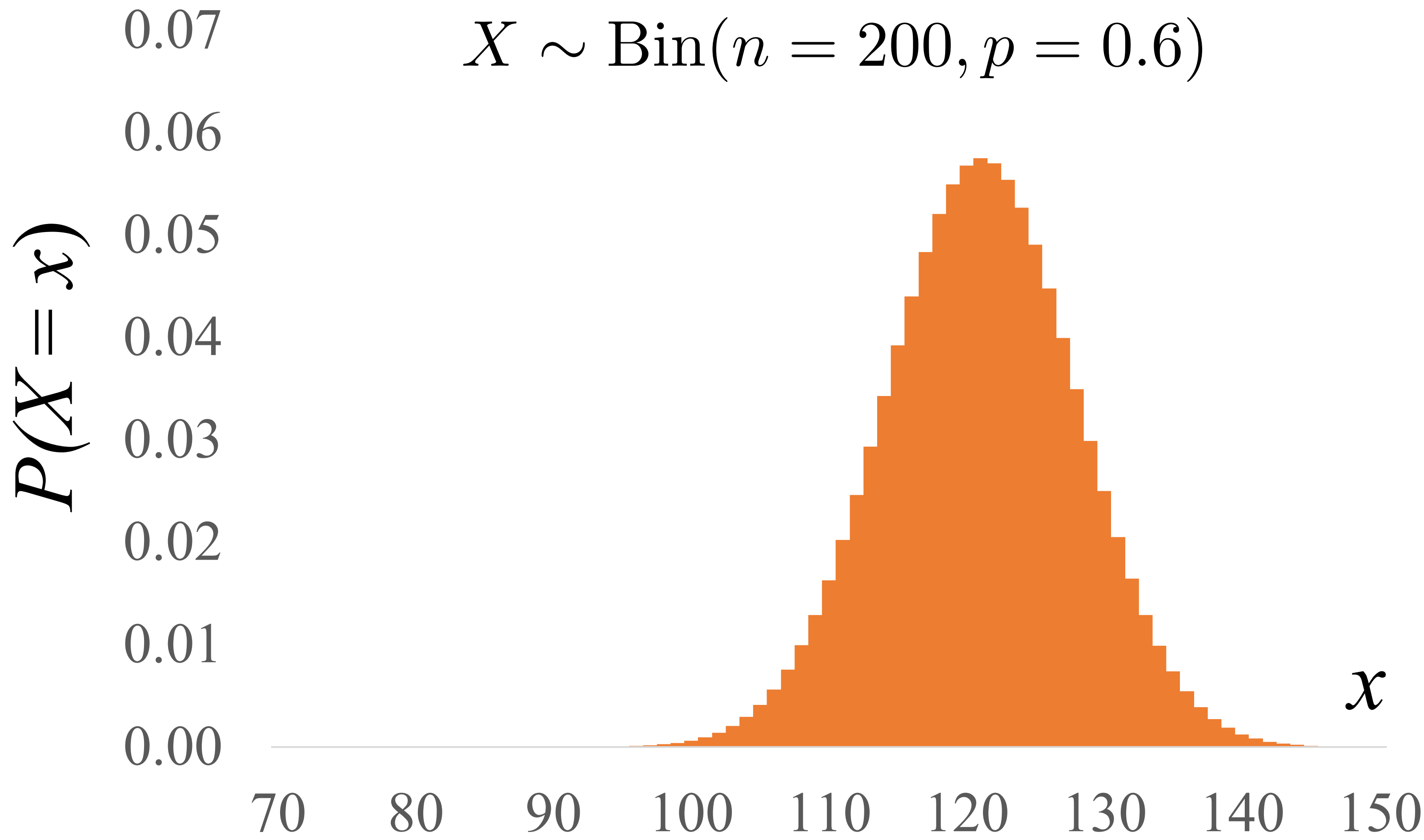

Why is there more mass in the middle?

# Other Functions?

# C.L.T. Explains This



# C.L.T. Explains This



# Binomial Approximation

- Consider I.I.D. Bernoulli variables $X_1$, $X_2$, ... With probability *p*
  - $X_i$ have $E[X_i] = p$ and $Var(X_i) = p(1-p)$

$$Y = \sum_{i=0}^{n} X_i$$

Y is the sum of the Bernoullis

$$Y \sim N(n\mu, n\sigma^2) \quad \text{as } n \to \infty$$

Central Limit Theorem

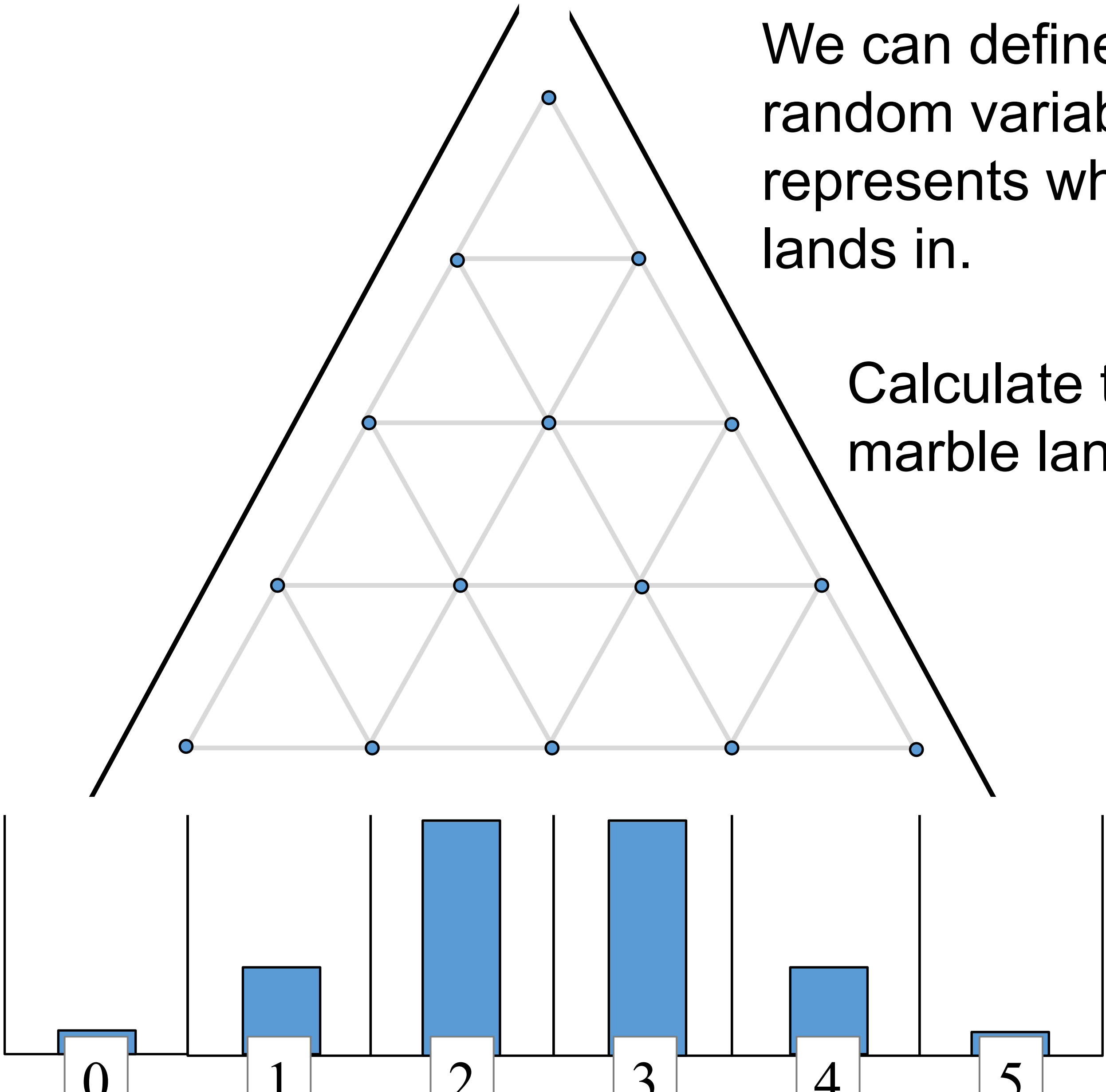
$$Y \sim N(np, np(1-p))$$

Substituting mean and variance of Bernoilli

# C.L.T. Explains This

We can define an indicator random variable (***B***) which represents what bucket a marble lands in.

Calculate the probability of a marble landing in a bucket.




PDF

The sum of independent, identically distributed variables:

$$Y = \sum_{i=0}^{n} X_i$$

Is normally distributed:

$$Y \sim N(n\mu, n\sigma^2)$$

---

where $\mu = E[X_i]$

$$\sigma^2 = \mathrm{Var}(X_i)$$

# On the Proof of the CLT

- The proof of the CLT uses the Fourier transform of the probability mass of the sample distance from the mean, divided by standard deviation, and shows that this approaches an exponential function in the limit:

$$f(x) = e^{-\frac{x^2}{2}}$$

- That exponential function is in turn the Fourier transform of the Standard Normal. The Fourier transform of a probability density function is called a *Characteristic Function*.

- The proof is beyond the scope of CS109.

# Central Limit Theorem in the Real World

- CLT is why some things in "real world" appear Normally distributed
  - Many quantities are sum of independent variables
  - Exams scores
    - Sum of individual problems on the SAT
    - Why does the CLT not apply to our midterm?
  - Election polling
    - Ask 100 people if they will vote for candidate X ($p_1$ = # "yes"/100)
    - Repeat this process with different groups to get $p_1, \ldots, p_n$
    - Will have a normal distribution over $p_i$
    - Can produce a "confidence interval"
      - How likely is it that estimate for true p is correct

# What about other functions?

Sum of iid? Normal

Average of iid?

Max of iid?

# The Central Limit Theorem

- Consider I.I.D. random variables $X_1$, $X_2$, ...
  - $X_i$ have same distribution with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$
  - Let: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
  - Central Limit Theorem:

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \quad \text{as} \quad n \to \infty$$

[Demo]

http://onlinestatbook.com/stat_sim/sampling_dist/

# But Wait! There is More

- Consider I.I.D. random variables $X_1$, $X_2$, ...
  - $X_i$ have distribution $F$ with $E[X_i] = \mu$ and $Var(X_i) = \sigma^2$

$$\bar{X} = \frac{1}{n}\sum_{i}^{n} X_i \qquad Y = \sum_{i}^{n} X_i \qquad \bar{X} = \frac{1}{n}Y$$

---

$$Y \sim N(n\mu, n\sigma^2)$$ By CLT

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$ Linear transform of a normal

By the Central Limit Theorem, the sample mean of IID variables are distributed normally.

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

# What about other functions?

Sum of iid? Normal

Average of iid? Normal

Max of iid?

# What about other functions?

Sum of iid? Normal

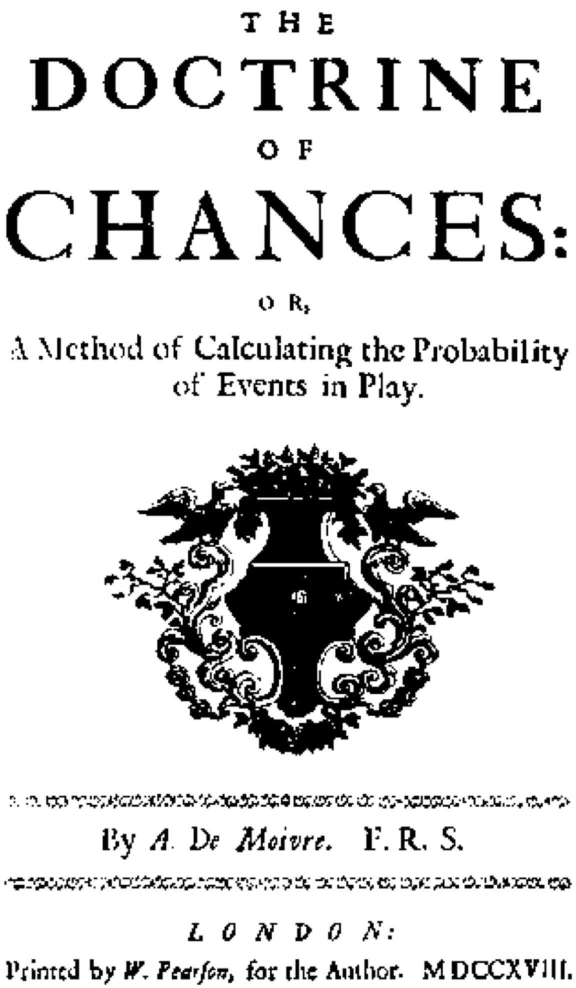
Average of iid? Normal

Max of iid? Gumbel

See Fisher Trippett Gnedenko Theorem

# Once Upon a Time…

Abraham De Moivre


THE

DOCTRINE

OF

CHANCES:

OR,

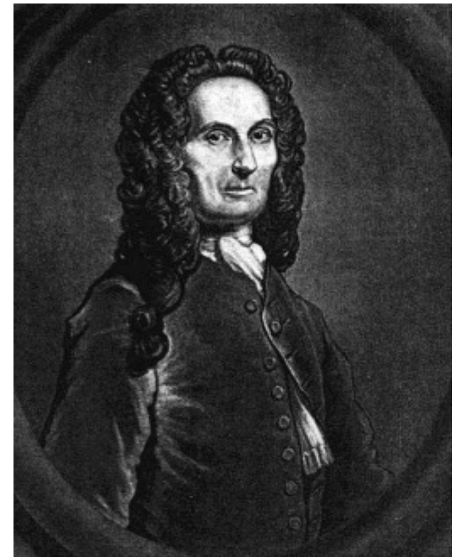
A Method of Calculating the Probability of Events in Play.

By *A. De Moivre.* F. R. S.

*LONDON:*

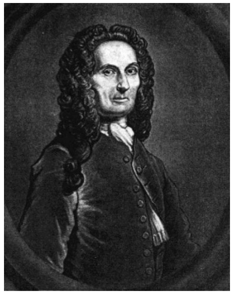
Printed by *W. Pearson*, for the Author. MDCCXVIII.

1733



# Once Upon a Time…

- History of the Central Limit Theorem
  - 1733: CLT for X ~ Ber(1/2) postulated by Abraham de Moivre



  - 1823: Pierre-Simon Laplace extends de Moivre's work to approximating Bin(n, p) with Normal

  - 1901: Aleksandr Lyapunov provides precise definition and rigorous proof of CLT
  - 2016: Beyonce releases Lemonade

    - It was her 6th album, bringing her total number of songs to 214
    - Mean quality of subsamples of songs is Normally distributed (thanks to the Central Limit Theorem)

# Estimating Clock Running Time

- Have new algorithm to test for running time
  - Mean (clock) running time: $\mu = t$ sec.
  - Variance of running time: $\sigma^2 = 4$ sec$^2$.
  - Run algorithm repeatedly (I.I.D. trials), measure time
    - How many trials s.t. estimated time = $t \pm 0.5$ with 95% certainty?
    - $X_i$ = running time of *i*-th run (for $1 \leq i \leq n$), $\bar{X}$ is the mean

---

$$0.95 = P(-0.5 < \bar{X} - t < 0.5)$$

$$\bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \sim N(t, \frac{4}{n}) \qquad \text{By CLT}$$

$$\bar{X} - t \sim N(0, \frac{4}{n}) \qquad \text{By linear transform of a normal}$$

$$0.95 = P(-0.5 < \bar{X} - t < 0.5) \qquad \bar{X} - t \sim N(0, \frac{4}{n})$$

---

$$0.95 = F_{\bar{X}-t}(0.5) - F_{\bar{X}-t}(-0.5)$$

$$= \Phi\Big(\frac{0.5 - 0}{\sqrt{4/n}}\Big) - \Phi\Big(\frac{-0.5 - 0}{\sqrt{4/n}}\Big)$$

$$= 2\phi(\frac{\sqrt{n}}{4}) - 1$$

$$0.95 = 2\phi(\frac{\sqrt{n}}{4}) - 1$$

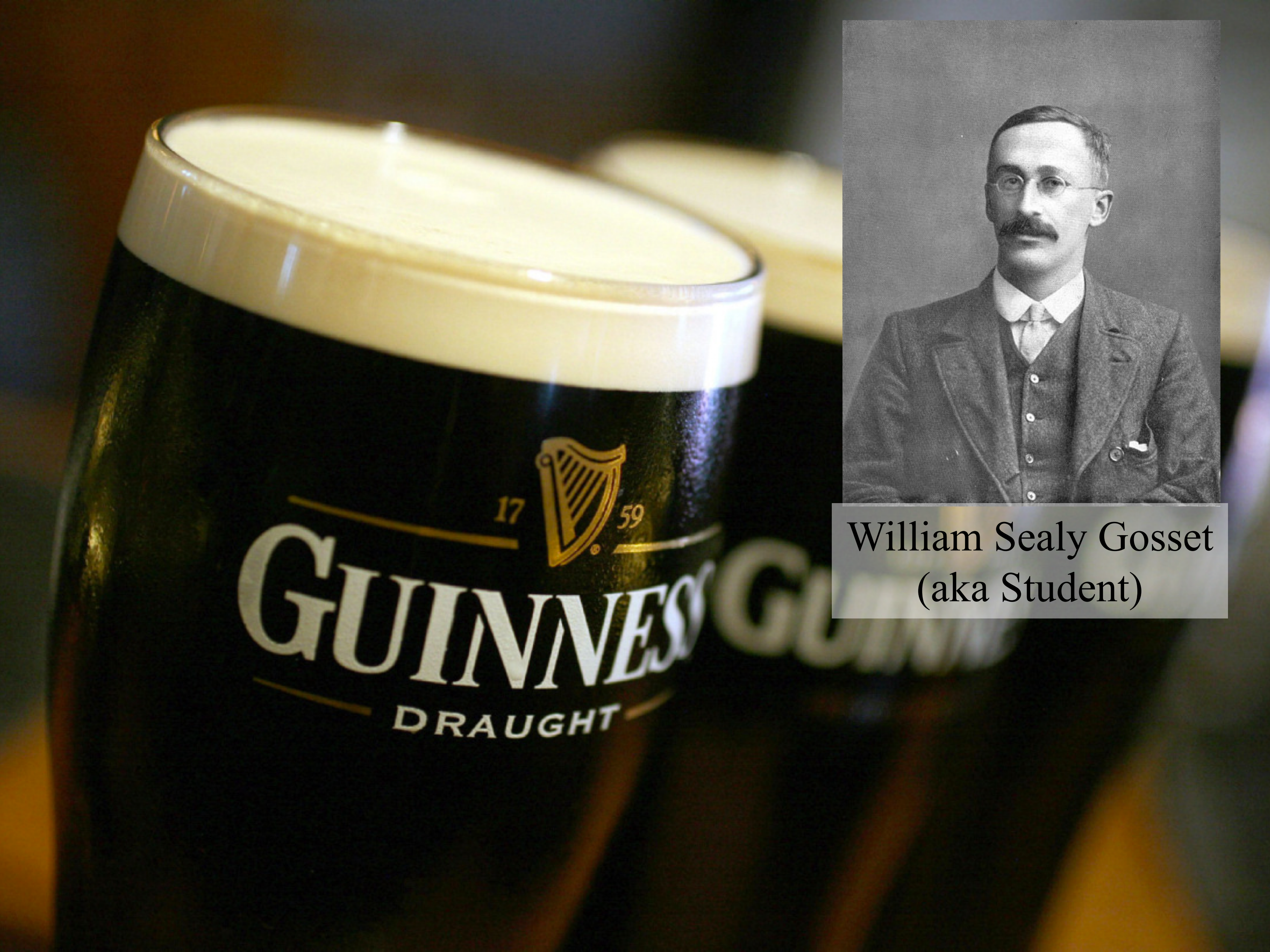
---

$$0.975 = \phi(\frac{\sqrt{n}}{4})$$

$$\phi^{-1}(0.975) = \frac{\sqrt{n}}{4}$$

$$1.96 = \frac{\sqrt{n}}{4}$$

$$n = 61.4$$




William Sealy Gosset
(aka Student)

# It’s play time!

# Sum of Dice

- You will roll 10 6-sided dice ($X_1$, $X_2$, …, $X_{10}$)
  - $X$ = total value of all 10 dice = $X_1 + X_2 + \ldots + X_{10}$
  - Win if: $X \leq 25$ or $X \geq 45$
  - Roll!
- And now the truth (according to the CLT)…

# Sum of Dice

- You will roll 10 6-sided dice ($X_1$, $X_2$, …, $X_{10}$)
  - X = total value of all 10 dice = $X_1 + X_2 + \ldots + X_{10}$
  - Win if: $X \leq 25$ or $X \geq 45$

---

- Recall CLT: $X = \sum_{i}^{n} X_i \rightarrow N(n\mu, n\sigma^2)$ As $n \rightarrow \infty$
  - Determine P($X \leq 25$ or $X \geq 45$) using CLT:

$\mu = E[X_i] = 3.5$ $\sigma^2 = \mathrm{Var}(X_i) = \frac{35}{12}$ $X \approx N(35, 29.2)$

$$1 - P(25.5 < X < 44.5) = 1 - P(\frac{25.5 - 35}{\sqrt{29.2}} < Z < \frac{44.5 - 35}{\sqrt{29.2}})$$

$$\approx 1 - (2\Phi(1.76) - 1) \approx 2(1 - 0.9608) = 0.0784$$

# Wonderful Form of Cosmic Order

I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the ”[Central limit theorem]". The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement, amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along.

-Sir Francis Galton