Most important pattern in Pascal's Δ is **Pascal's identity**:

<u>Thm</u> $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

<u>e.g.</u>

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \end{array}$$

$$\begin{tikzcd}[column sep=small] \binom{n-1}{k-1} \arrow[dr] & & \binom{n-1}{k} \arrow[dl] \\ & \binom{n}{k} & \end{tikzcd}$$

<u>Note</u>: Let's you easily **fill in** Pascal's Δ!

<u>Pf</u>: Let's define a **bijection**

$$f : \left\{ \begin{matrix} k\text{-subsets} \\ \text{of } [n] \end{matrix} \right\} \to \left\{ \begin{matrix} k\text{-subsets} \\ \text{of } [n-1] \end{matrix} \right\} \cup \left\{ \begin{matrix} k-1\text{-subsets} \\ \text{of } [n-1] \end{matrix} \right\}$$

by $f(A) = \begin{cases} A & \text{if } n \notin A \quad (\text{a } k\text{-subset of } [n-1]) \\ A \setminus \{n\} & \text{if } n \in A \quad (\text{a } k-1\text{-subset of } [n-1]) \end{cases}$

<u>This</u> exactly corresponds to Pascal's identity. $\square$

<u>Rmk</u>: We have a similar identity

$$S(n,k) = S(n-1,k-1) + k \cdot S(n-1,k)$$

for Stirling #'s of 2^{nd} kind, w/ a very similar bijective proof. //

Math 4707: More Pascal's triangle and probability

2/3

Ch.'s 3+5 of LPV

Reminder: • HW #1 is due today!

Last class we introduced Pascal's triangle of $\binom{n}{k}$:

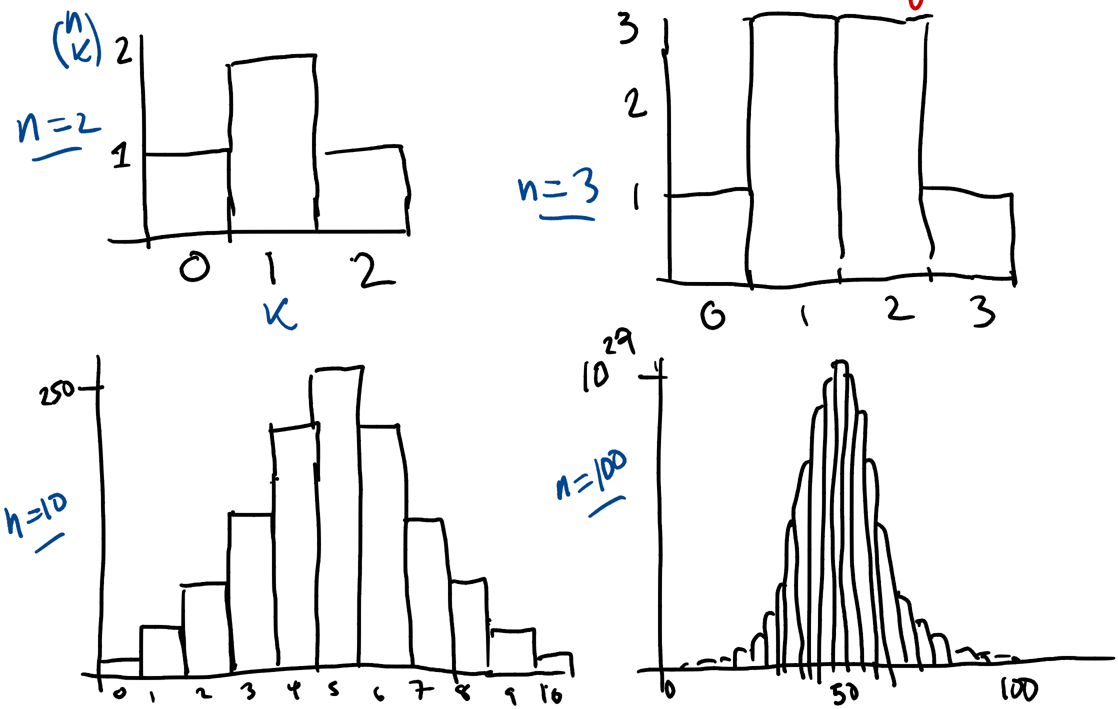
$$
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \vdots & & & &
\end{array}
$$

and discussed various patterns in it, like symmetry, the sum/alternating sum of a row, and, most important, Pascal's identity $\begin{array}{ccc} a & & b \\ & a+b & \end{array}$. We will discuss more patterns like this on the worksheet for today.

But in today's lecture, instead we're going to talk about the large scale behavior of Pascal's Δ, and its connections to basic probability. The material for today is mostly "cultural", i.e., I will not assess you on it. However, it is still very interesting + important.

Q: What does the n^{th} row of Pascal's Δ roughly "look like," for big n?

To answer this, helpful to draw a histogram:



What do we see in these pictures?

- symmetry $\binom{n}{k} = \binom{n}{n-k}$ ✓
- numbers get bigger towards the middle

Indeed,

$$\binom{n}{k} < \binom{n}{k+1} \iff \frac{n!}{k!(n-k)!} < \frac{n!}{(k+1)!(n-k-1)!}$$

$$\iff 1 < \frac{n-k}{k+1}$$

$$\iff k < \frac{n-1}{2},$$

So for first half of k, have $\binom{n}{k} < \binom{n}{k+1}$ ✓

- middle number is pretty big

Recall sum of $\binom{n}{k} = 2^n$, and there are $n+1$ k's

$$\Rightarrow \underset{\text{over } k}{\text{average}} \binom{n}{k} = \frac{1}{n+1} 2^n$$

$$\Rightarrow \text{biggest } \binom{n}{k} \left(= \binom{n}{n/2} \right) \geq \frac{1}{n+1} 2^n$$ ✓

(In fact, from Stirling's approx. $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, can show $\binom{n}{n/2} \sim \sqrt{2/\pi n}\; 2^n$.)

- histogram looks like it approaches a curve

In fact, letting $n=2m$ for convenience, have

$$\binom{2m}{m-t} \Big/ \binom{2m}{m} \approx e^{-t^2/m}$$

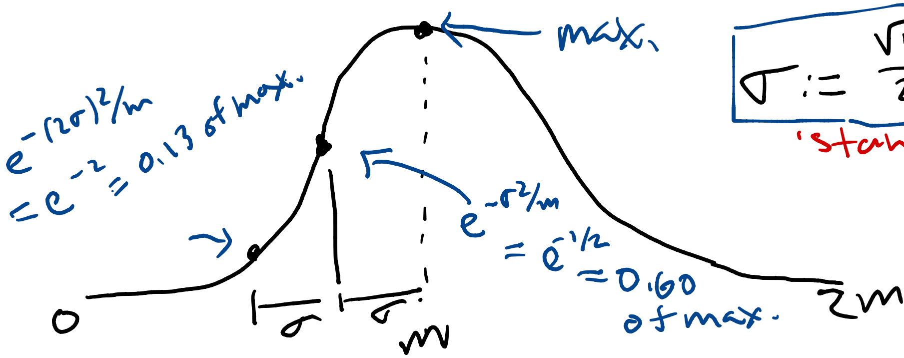
'Gaussian curve'

A.K.A. 'Normal curve'

A.K.A. 'Bell curve'

What does this mean?

Let's draw the curve to see...

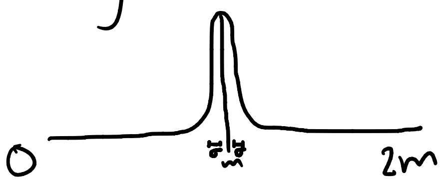


$$\boxed{\sigma := \frac{\sqrt{n}}{2} = \frac{\sqrt{m}}{\sqrt{2}}}$$

'standard deviation'

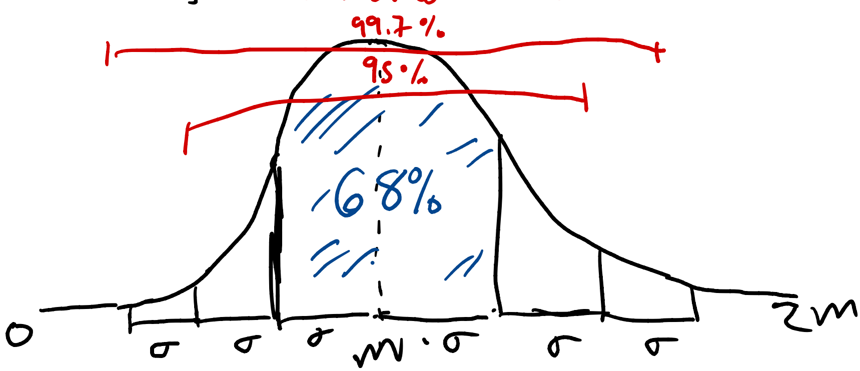
Note: picture is misleading since $\sigma \ll 2m$

'Real picture' =



Upshot: values $\binom{n}{k}$ drop off rapidly from middle

Also, most of area under curve is in middle:



Precise lemmas from book are:

Lemma For $0 \le t \le m$,

$$e^{-t^2/(m-t+1)} \le \binom{2m}{m-t} \Big/ \binom{2m}{m} \le e^{-t^2/(m+t)}$$

Lemma For $0 \le k \le m$, and $c := \binom{2m}{k} \Big/ \binom{2m}{m}$,

$$\binom{2m}{0} + \binom{2m}{1} + \cdots + \binom{2m}{k-1} < \frac{c}{2} \cdot 2^{2m}.$$

total area under curve = sum of $2m^{th}$ row of Δ

e.g., $m = 500$ then $\binom{1000}{448} \Big/ \binom{1000}{500} < 0.01$

Thus sum of 1^{st} 447 $\binom{1000}{k} < 0.5\%$ of total sum

By symmetry, last 447 $\binom{1000}{k}$ also $< 0.5\%$ total

So middle 107 terms account for $> 99\%$ of sum of the 1000^{th} row of Pascal's Δ!

Pf of these lemmas: Skipped. Based on $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ + manipulating inequalities, taking logarithms, Stirling's approx., etc. $\blacksquare$

Q: Why are we interested in these facts?

A: Basic probability theory!

S = finite set = "sample space"

e.g. $S = \{1,2,3,4,5,6\}$ = outcomes of rolling a die

$S = \{H, T\}$ = flipping a coin

$S = \{HH, HT, TH, TT\}$ = flipping two coins

An event is any subset of S.

e.g. roll ≥ 3 $= \{3,4,5,6\} \subseteq \{1,2,3,4,5,6\}$

get one heads in two flips $= \{HT, TH\} \subseteq \{HH, HT, TH, TT\}$

A probability distribution on $S = \{s_1, s_2, \ldots, s_n\}$ is a way of assigning nonnegative real numbers $p(s_1), p(s_2), \ldots, p(s_n)$ s.t. $\boxed{p(s_1) + \cdots + p(s_n) = 1.}$

Commonly: uniform distribution $p(s_i) = \frac{1}{n}$ $\forall i$.

(but can also allow a weighted die, etc...)

The probability of event $A \subseteq S$ is $\sum_{s_i \in A} p(s_i)$.

If we have uniform distr., this is $\boxed{\frac{\# A}{\# S}}$.

e.g. $Pr(\text{roll} \geq 3) = \frac{\#\{3,4,5,6\}}{\#\{1,2,3,4,5,6\}} = \frac{4}{6}$. ✓

Independence: Two events $A, B \subseteq S$ are independent if $Pr(A \cap B) = Pr(A) \cdot Pr(B)$.

Roughly, A + B are unrelated...

Do we see why uniform distr. on $\{HH, HT, TH, TT\}$ = "flipping two independent fair coins"?

Q: What is the probability of getting exactly k heads when flipping n coins?

A: $\frac{\binom{n}{k}}{2^n}$ = fraction of area under n^{th} row Pascal's Δ curve at position k

Let's see this in action w/ Galton board...

Q: If we do 1000 coinflips, what fraction of heads should we expect?

Thm (Law of Large numbers)

For any $\epsilon > 0$,

$$\Pr\left(\begin{array}{l}\text{fraction of heads in } n \text{ coinflips} \\ \text{is between } \tfrac{1}{2}-\epsilon \text{ and } \tfrac{1}{2}+\epsilon\end{array}\right) \to 1$$

as $n \to \infty$.

I.e., in 1000 coinflips, should expect very close to 50% heads!

Pf: Recall picture of n^{th} row of Pascal's Δ:

All the mass is very close to middle

= 50% heads $\blacksquare$

A more precise result called the **Central limit theorem** says that as $n \to \infty$, histogram of

$$\sqrt{n}\left(\frac{1}{n}\left(\begin{array}{c}\#\text{ of heads}\\ \text{in } n \text{ flips}\end{array}\right) - \frac{1}{2}\right) \to \frac{2}{\sqrt{2\pi}} e^{-2x^2}$$

rescale by $\sqrt{n}$ to see 'fluctuations' from average

This just repeats what we saw earlier w/ Pascal's Δ.

So what? Significance of LLN + CLT is that they apply not just to coinflips, but any time we take average of independent random variables (e.g., dice rolling, error of scientific measurement, etc.) They explain:

- why the scientific method works
- why polling works, etc.

and why (in 'fairytale land' at least) the Gaussian curve emerges as a universal limit (e.g., human height distributions, etc.).

Now let's take a 5 min. break,

and when we come back we

can work on a worksheet on more

combinatorial patterns in Pascal's Δ

in our breakout groups

(this worksheet is not really

related to LLN/CLT...)