Lecture 1:

Probability and Distributions

- Context
- Probability traps
- Binomial & Poisson Distributions
- Expectation and Variance
- Estimators
- Gaussian Distribution

SCIENTIFIC AMERICAN™

Neutrino Experiment Replicates Faster-Than-Light Finding

Latest data show the subatomic particles continue to break the spe
limit.

November 18, 2011

nature

By Eugenie Samuel Reich of Nature magazine

Physicists have replicated the finding that the subatomic partic
called neutrinos seem to travel faster than light. It is a remarkab
confirmation of a stunning result, yet most in the field remain sc
that the ultimate cosmic speed limit has truly been broken.

THE Sun

Tuesday, 4 October 2011

Login | Register

Oxford | 16C

HOME
NEWS
Politics
Weather
VIDEO
SPORT
Football
Dream Team
Rugby World Cup
Cricket
SHOWBIZ
TV

News

GOT A STORY? EMAIL: TALKBACK@THE-SUN.CO

Einstein wrong (in theory, at least)

Boffins 'beat speed of light'

By DAN SALES

Published: 23 Sep 2011

Add a comment:

SCIENTISTS claimed yesterday that they have proved Albert Einstein got his famous theory WRONG.

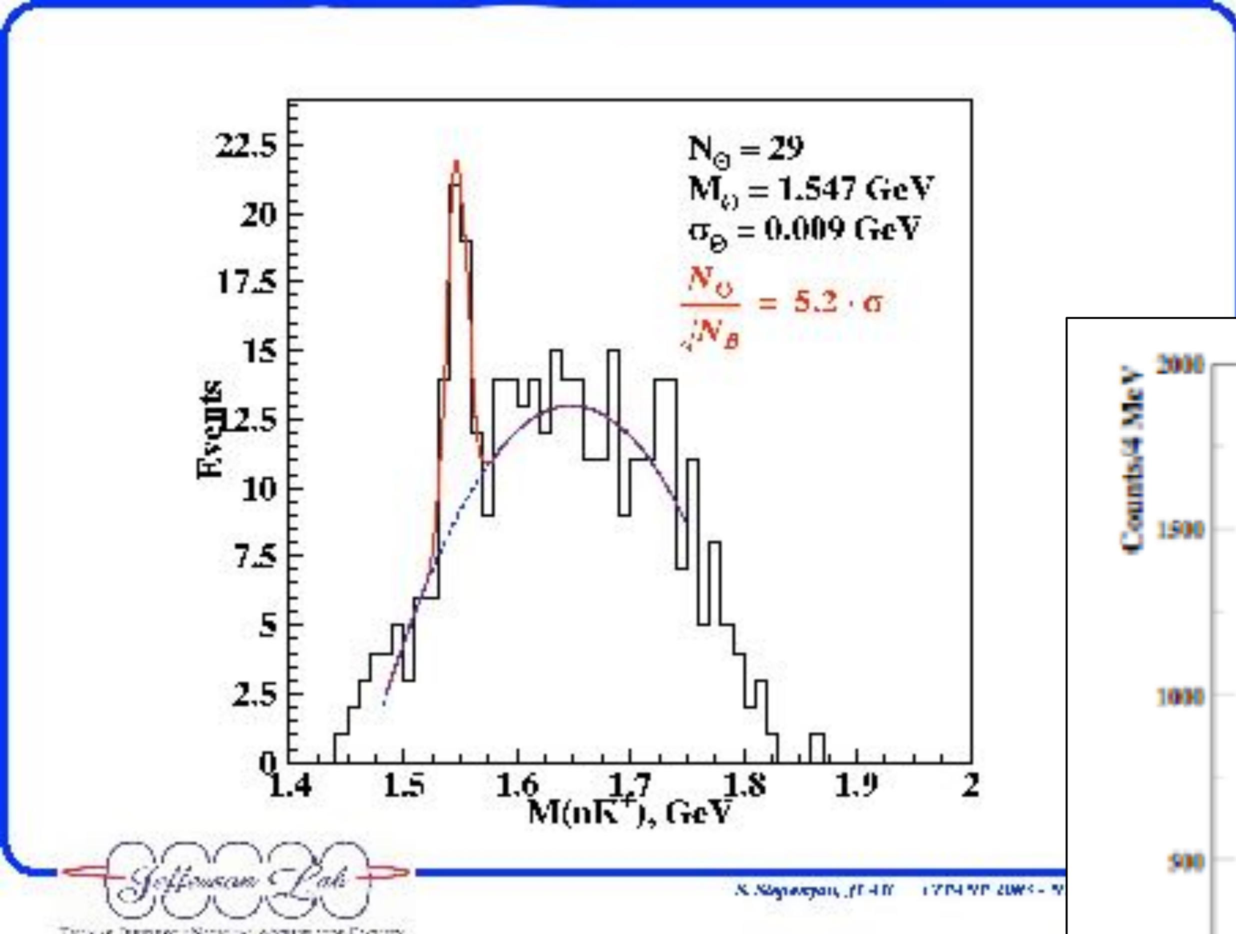
They said an experiment had shown the celebrated genius was mistaken when he declared that nothing can go faster than light.

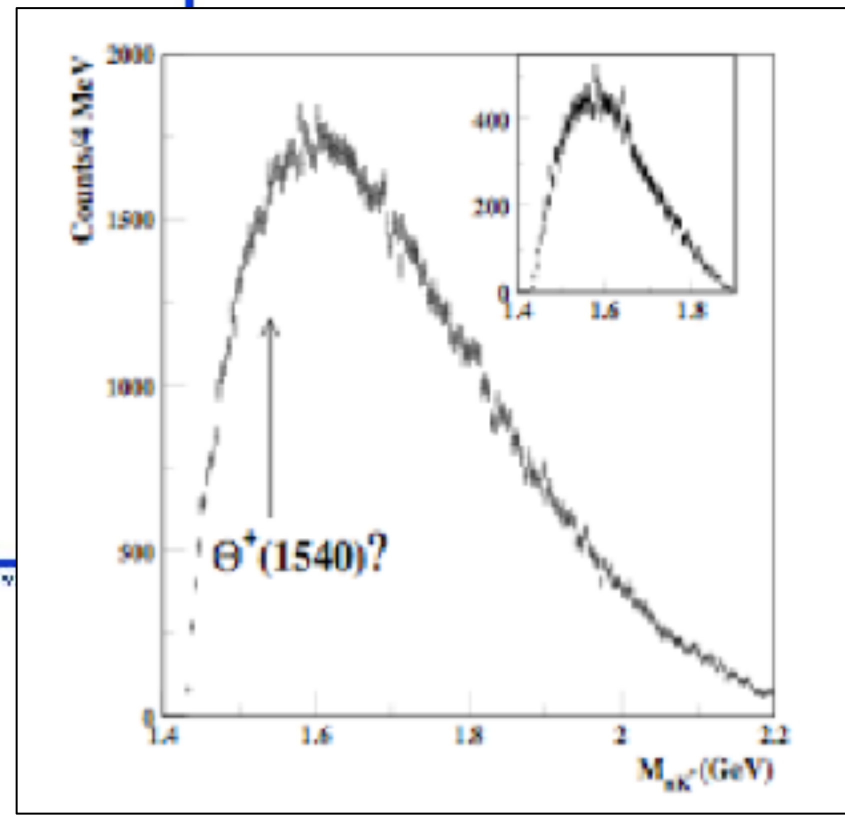
It will overturn one of the fundamental laws of the universe if they are right.

The international team of experts said they have recorded

Incorrect ... scientists say Einstein was mista

RELATED STORIES





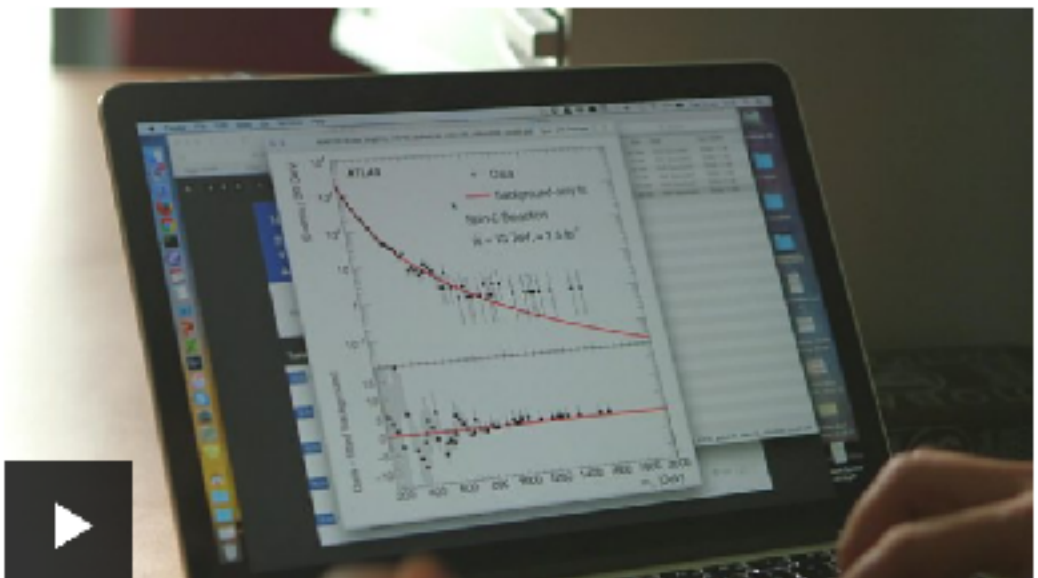
"There are two or three recent experiments that find weak evidence for signals near the nominal masses, but there is simply no point in tabulating them in view of the overwhelming evidence that the claimed pentaquarks do not exist... The whole story—the discoveries themselves, the tidal wave of papers by theorists and phenomenologists that followed, and the eventual "undiscovery"—is a curious episode in the history of science." **(2008 Review of Particle Physics)**

Science & Environment

New particle hopes fade as LHC data 'bump' disappears

By Pallab Ghosh
Science correspondent, BBC News

5 August 2016 Science & Environment Share



Disappointment: Cern's Dr Marco Delmastro speaks to Horizon on BBC2

Hopes for the imminent discovery of a particle that might fundamentally change our understanding of the Universe have been put on hold.

PHYS ORG Nanotech

Diphoton

BBC Sign in
NEWS
Home UK World Busine

Science & Environment

Has the LHC d

By Paul Rincon
Science editor, BBC News website

5 July 2016 Science & Environm

A giant depiction of the CMS expe

After its much heralded re-st
the Large Hadron Collider, I

Other Sciences

tical papers

V edition

Jernej Kamenik

NewScientist

Surprise LHC blip hints at Higgs – again

22:49 22 July 2011

"...The combined statistical significance, taking all three types of excess reported by ATLAS into account, is 2.8 sigma, slightly below the 3 sigma threshold (equivalent to a 1-in-370 chance of being due to a fluke) that a measurement must pass to count as "evidence" for something new: only 5 sigma data, equivalent to a 1-in-1.7 million chance of being due to a fluke, gains "discovery" status.

The other main detector at the LHC, called CMS, has found an excess in a similar range, between 130 and 150 GeV, reports Nature. The size of that excess is roughly 2 sigma, writes physicist Adam Falkowski on the Resonances blog.

If all this sounds a tad familiar, rewind back to April, when four physicists claimed to have found hints of the Higgs in ATLAS data in a study abstract leaked online. A subsequent official analysis by the collaboration of 700 physicists who run ATLAS concluded that result was an error. Unlike that claim, the new excesses have been vetted by the ATLAS and CMS collaborations respectively."

BBC NEWS

BBC Home ▸ BBC News ▸ Sci/Environment

▾Menu

Higgs boson 'hints' also seen by US lab

24 July 11 15:57

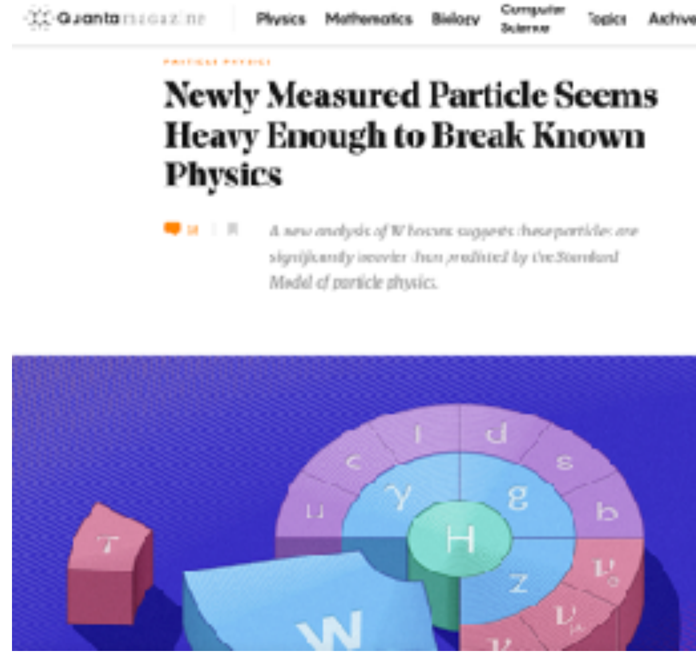
By Paul Rincon

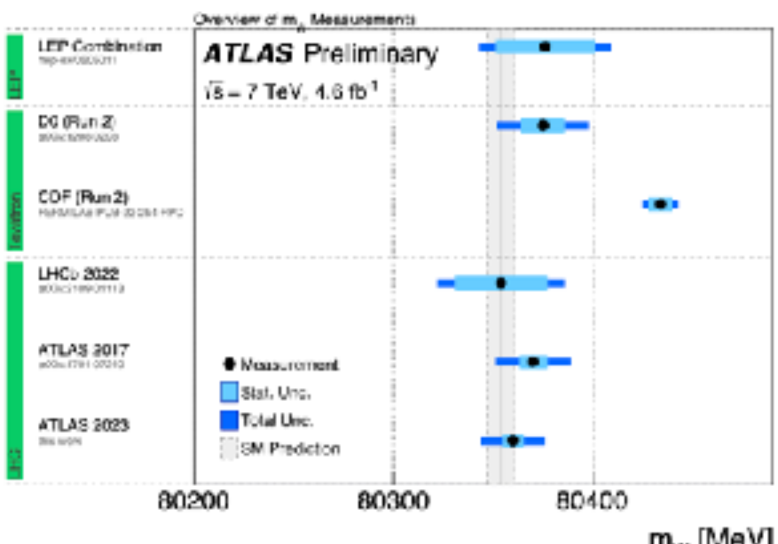
Science editor, BBC News Website, Grenoble

A US particle machine has seen possible hints of the Higgs boson, it has emerged, after reports this week of similar glimpses at Europe's Large Hadron Collider (LHC) laboratory.

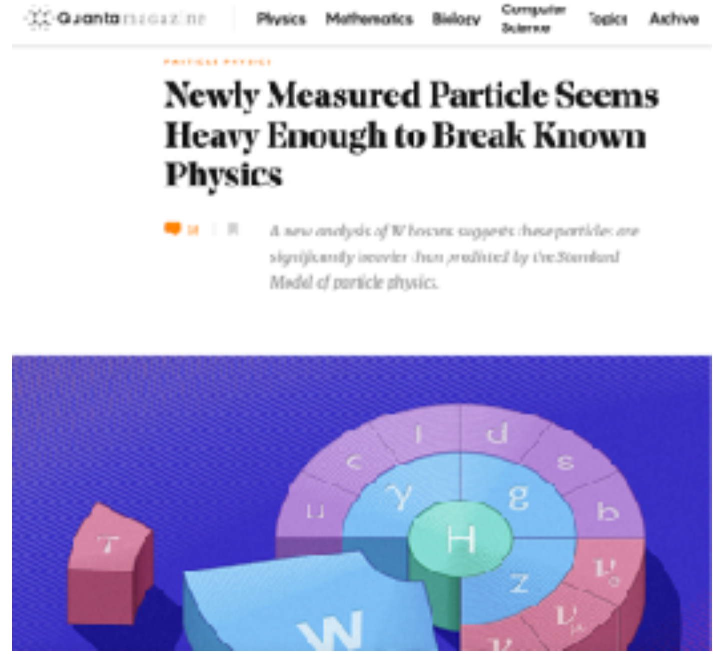
"...The hints seen at the Tevatron are weaker than those reported at the LHC, but occur in the same 'search region'."

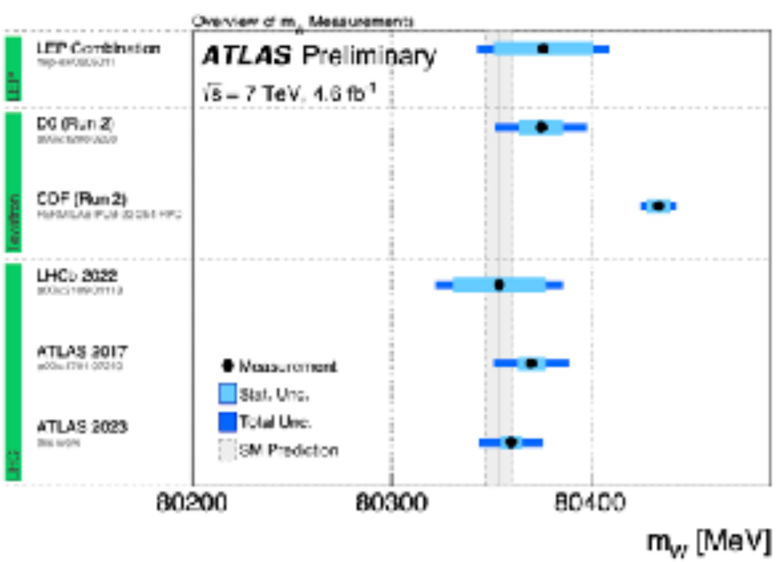

CERNCOURIER
Reporting on international high-energy physics
Physics
Technology
Community
In focus
Magazine
CDF sets W mass against the Standard Model
Credit: Fermilab

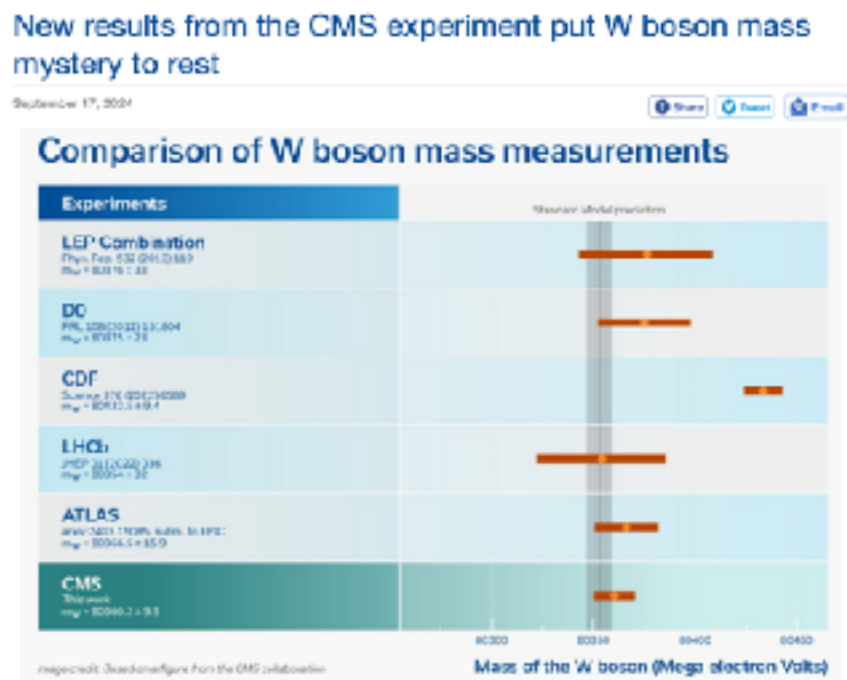

Physics
Mathematics
Biology
Computer Science
Topics
Archive
Newly Measured Particle Seems Heavy Enough to Break Known Physics
A new analysis of W bosons suggests these particles are significantly heavier than predicted by the Standard Model of particle physics.
H
W
Z
g
γ


Overview of m_W Measurements
ATLAS Preliminary
√s = 7 TeV, 4.6 fb⁻¹
LEP Combination
D0 (Run 2)
CDF (Run 2)
LHCb 2022
ATLAS 2017
ATLAS 2023
Measurement
Stat. Unc.
Total Unc.
SM Prediction
80200
80300
80400
m_W [MeV]


CERNCOURIER
Reporting on international high-energy physics
Physics
Technology
Community
In focus
Magazine
CDF sets W mass against the Standard Model


Physics
Mathematics
Biology
Topics
Archive
Newly Measured Particle Seems Heavy Enough to Break Known Physics
H
W


Overview of m_W Measurements
LEP Combination
D0 (Run 2)
CDF (Run 2)
LHCb 2022
ATLAS 2017
ATLAS 2023
ATLAS Preliminary
$\sqrt{s}$ = 7 TeV, 4.6 fb^{-1}
Measurement
Stat. Unc.
Total Unc.
SM Prediction
80200
80300
80400
m_W [MeV]


New results from the CMS experiment put W boson mass mystery to rest
Comparison of W boson mass measurements
Experiments
LEP Combination
D0
CDF
LHCb
ATLAS
CMS
Mass of the W boson (Mega electron Volts)

Competition!!

Competition winner from 2013: Mark Smith

Assumed "middle-aged" cod lived 13 years based on a quoted maximum cod age of 25 years. But that data was from Barents Sea - estimate for North Sea cod maximum age is more like 11 years. Inference of middle age from max age is dubious in any case.

actual best estimate:

436,900,000

The Telegraph

Home | News | World | Sport | Finance | Comment | Culture | Travel | Life | Women | Fashion

Politics | Obits | Education | Earth | Science | Defence | Health | Scotland | Royal | Celebrities

Earth News | Environment | Climate Change | Wildlife | Outdoors | Picture Galleries | Earth Vide

HOME » EARTH » ENVIRONMENT

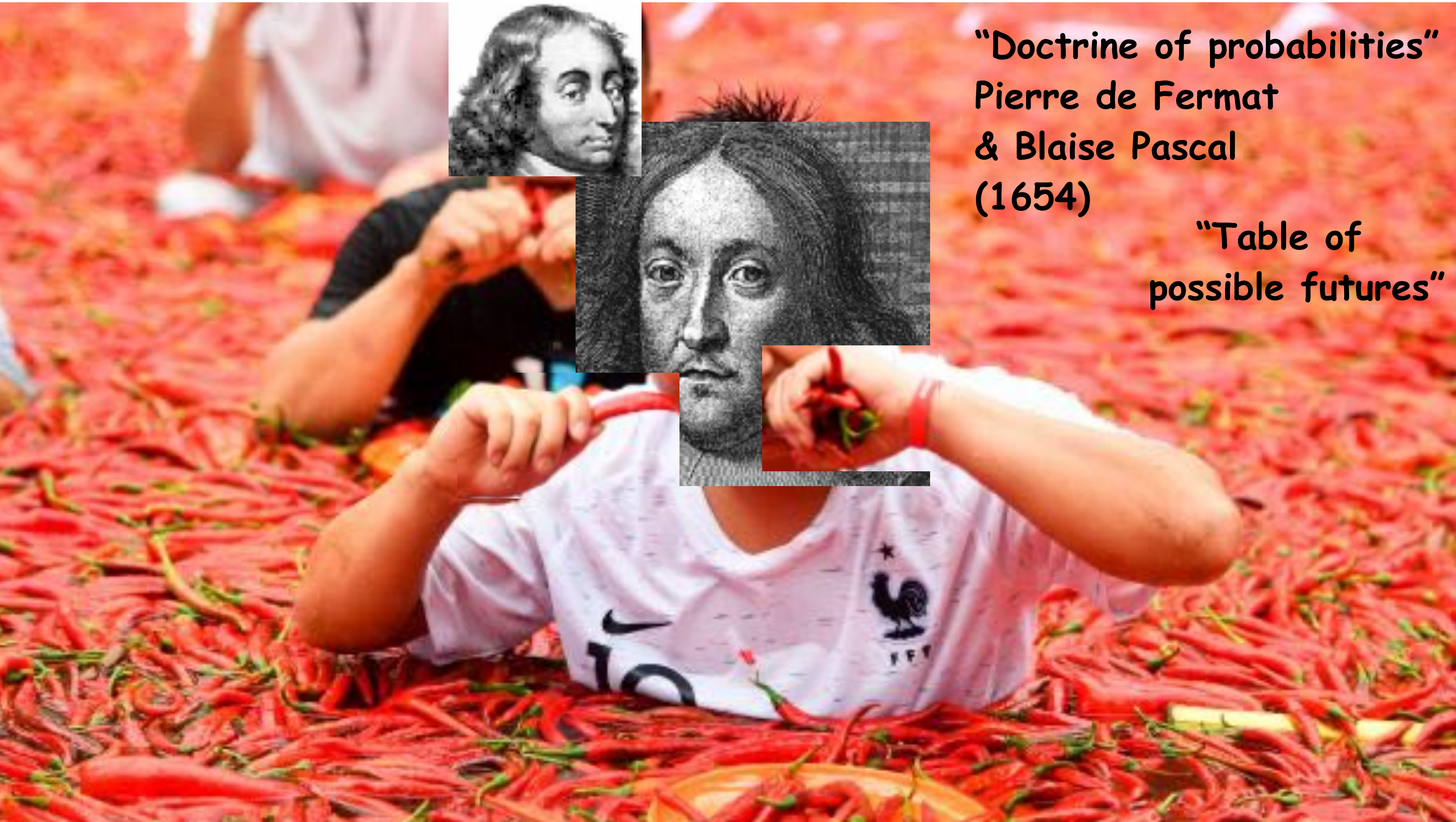
Just 100 cod left in North Sea

Overfishing has left fewer than 100 adult cod in the North Sea, it was reported.

Not a single cod aged over 13 was caught in the North Sea last year. Photo: GETTY

Competition!!

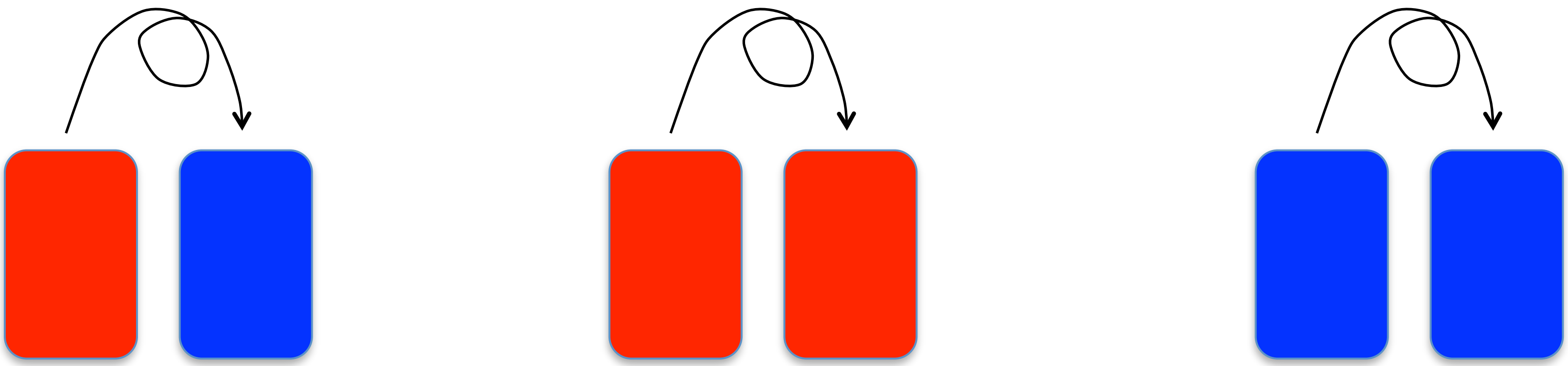
Competition!!



"Doctrine of probabilities"
Pierre de Fermat
& Blaise Pascal
(1654)

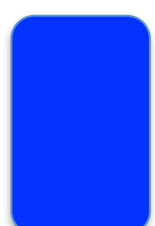
"Table of
possible futures"

“The Card Game”



"The Card Game"

I'll bet you £10 that the other side is blue.



side shown	other side
1 (R,R,B)	B
2 (B,R,B)	R,B

Chance for the other side to be blue is 2/3 !

“Prisoner’s Paradox”

One of you lucky boys
will **only** get life in prison. But
I have instructed the guard not to
inform you whether or not you will
hang until I announce to the press
tomorrow morning as a
last minute surprise!

Not good!

Survival probability = 1/3 !!

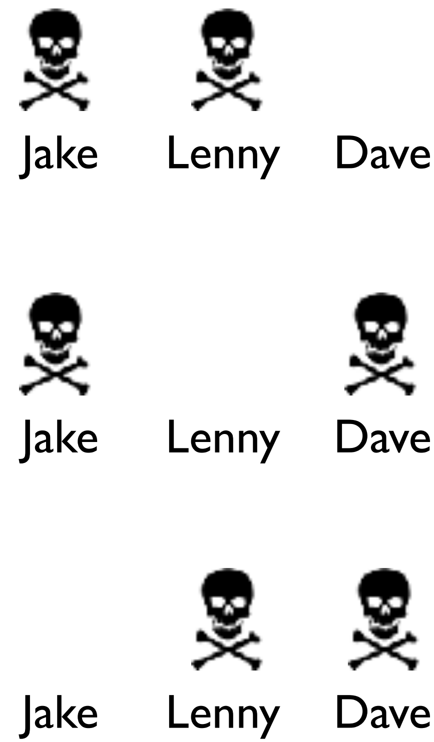
Jake Lenny Dave

“Prisoner’s Paradox”

“Prisoner’s Paradox”

So now it’s either me or Lenny... Great! My odds have improved to 1/2 !!

Dave



Mantra:

Ask the Right Question!

S S F S F S

“Bernoulli Trials”

Loaded Dice?

If not loaded (test to reject H0),

$$1/6 \times 1/6 \times 1/6 \times 1/6 \times 1/6 \times 1/6 = \frac{1}{46656}$$

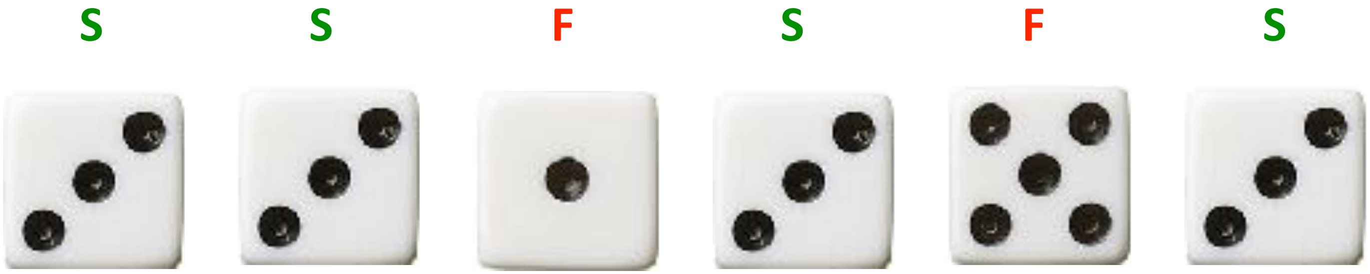
What’s the chance probability of getting **exactly** this sequence?

$$1/6 \times 1/6 \times 5/6 \times 1/6 \times 5/6 \times 1/6 = \frac{25}{46656}$$

or more generally: $p^k (1-p)^{n-k}$

where p is the probability of successes (getting a 3), k is the number of successes and n is the total number in the sequence.

What’s the chance probability of getting four 3’s in this order?



$$(1/6 \times 1/6 \times 5/6 \times 1/6 \times 5/6 \times 1/6)(\text{any 4 of 6})$$

distinct ways to order 6 things

$$(\text{any 4 of 6}) = \binom{6}{4} = \frac{6!}{4!2!} = 15 \quad \frac{375}{46656}$$

re-ordered of successes is not distinct

re-ordering of failures is not distinct

or more generally: $\binom{n}{k} p^k (1-p)^{n-k}$ Binomial Distribution ("two terms")

What's the chance probability of getting four 3's in any order?

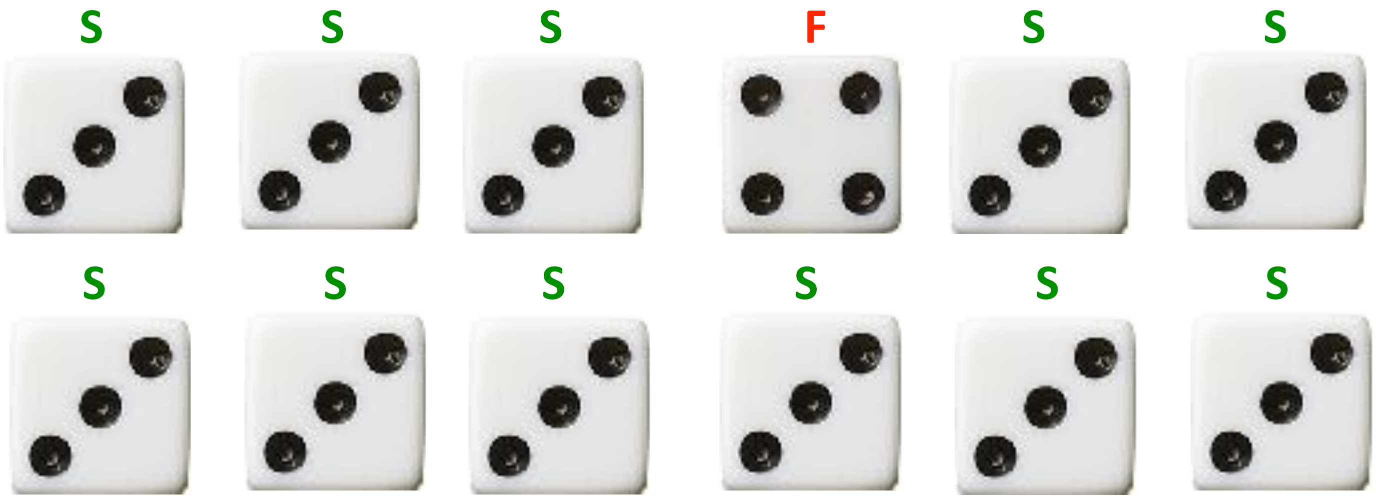
S S F S F S

what about this?

so we really want $6 \times \binom{6}{4} (1/6)^4 (1-1/6)^2$

$$= \frac{2250}{46656}$$

What's the chance probability of getting four of anything in any order?



Also counts as achieving four 3's… and would draw even more suspicion!!

so, in this case we want

$$\sum_{j=k}^{n} 6x\binom{n}{j} p^j (1-p)^{n-j}$$

k=4, n=6

$$\frac{2250}{46656} + \frac{180}{46656} + \frac{6}{46656} = \frac{2436}{46656} = 5.2\%$$

What's the chance probability of getting four or more of any number in any order?

So, what do you think is the probability that the die is fair? Would you continue to play?

What if the die belongs to a guy named ‘Eddie,’ who seems to be making a lot of money on the street corner?

What if the die belongs to your mum, and looks like exactly the same one you have played with ever since you were very young?

But then what if your mum used to be a well known card shark, today is April 1st, and your brother tells you she has been pulling some funny tricks lately and he doesn’t trust her?

It seems like context ought to somehow enter into your final assessment of the probability that the die is fair. This is ‘Bayesian’ probability, as opposed to ‘statistical’ or ‘frequentist’ probability. Much more on this later!

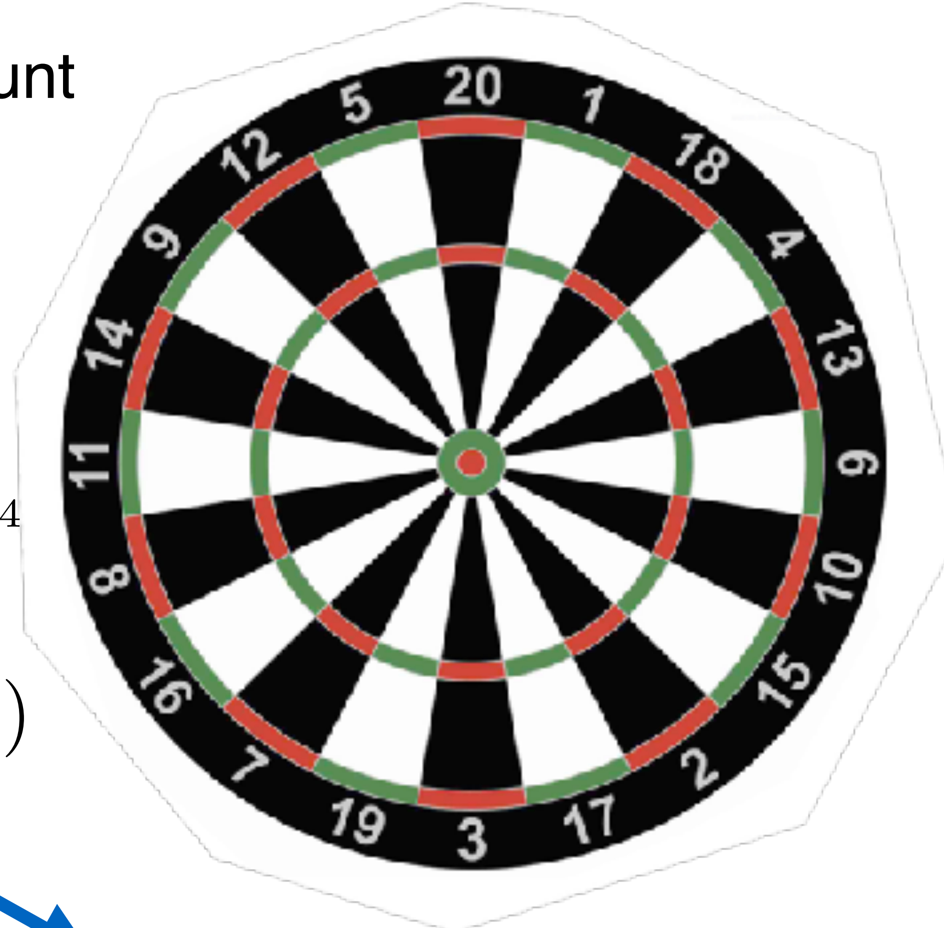
Statistical probability is basically the frequency with which a given "equivalent" outcome occurs if we were to repeat the same experiment over and over again.

What is the source of this statistical behaviour??

1) Hidden variations in initial conditions

2) Fundamental uncertainty (quantum mechanics)

Assume terrible aim, but only count throws that hit dart board. . .



What's the chance of hitting the bullseye given 100 throws?

$$p_s = (0.5in/17.75in)^2 = 7.93 \times 10^{-4}$$

$$P_{tot} = \sum_{k=1}^{100} P_{bin}(\text{k successes})$$

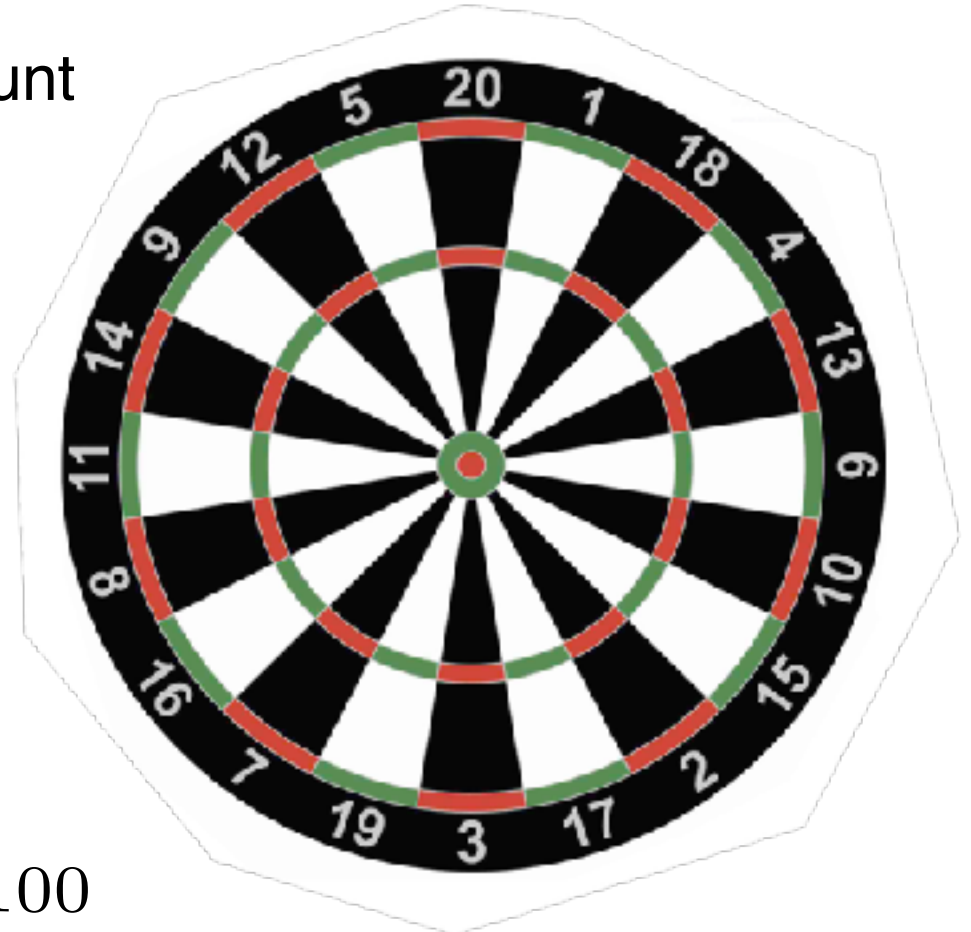
$$\binom{100!}{k!(100-k)!} p_s^k (1-p_s)^{100-k}$$

$$= 1 - P_{bin}(0 \text{ successes})$$

$$= 1 - (1-p_s)^{100}$$

$$= 7.63\% \qquad \sim 100 \times p_s$$

Assume terrible aim, but only count throws that hit dart board. . .



What's the chance of hitting the 20 given 100 throws?

$$p_s \sim 1/20 = 0.05$$

$$P_{tot} = 1 - (1 - 0.05)^{100}$$

$$= 99.4\% \qquad \neq 100 \times 0.05 \text{ !!!}$$

Binomial Distribution:

$$P(\text{k successes in n attempts}) = \left(\frac{n!}{k!(n-k)!}\right) p_s^k (1-p_s)^{n-k}$$

prob of each success

So, the expected (average) number of successes after summing over **n** identical Bernoulli trials is:

$$\mu = np$$

Now consider the case where the expected number of successes depends on the size of a **continuous variable** (*e.g.* length or time interval), which can be arbitrarily small.

The number of successes expected over a continuous interval of finite size can be viewed as resulting from the sum of an infinite number of Bernoulli trials carried out for arbitrarily small intervals such that:

$$\mu = \lim_{n \to \infty} np$$

So, set p= μ/n and evaluate

$$P(k) = \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k}$$

$$= \left(\frac{\mu^k}{k!}\right) \lim_{n\to\infty} \underbrace{\boxed{\frac{n!}{(n-k)!} \left(\frac{1}{n}\right)^k}} \left(1 - \frac{\mu}{n}\right)^n \left(1 - \frac{\mu}{n}\right)^{-k}$$

$$\lim_{n\to\infty} \frac{n!}{(n-k)!} \left(\frac{1}{n}\right)^k = \lim_{n\to\infty} \frac{n(n-1)(n-2)\ldots(n-k)(n-k-1)\ldots(1)}{(n-k)(n-k-1)\ldots(1)} \left(\frac{1}{n}\right)^k$$

$$= \lim_{n\to\infty} \frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k}$$

$$= \lim_{n\to\infty} \left(\frac{n}{n}\right) \left(\frac{n-1}{n}\right) \left(\frac{n-2}{n}\right) \ldots \left(\frac{n-k+1}{n}\right)$$

$$= 1$$

So, set p= μ/n and evaluate

$$P(k) = \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k}$$

$$= \left(\frac{\mu^k}{k!}\right) \lim_{n\to\infty} \boxed{\left(1 - \frac{\mu}{n}\right)^n} \left(1 - \frac{\mu}{n}\right)^{-k}$$

$$\lim_{n\to\infty} \left(1 - \frac{\mu}{n}\right)^n = \lim_{n\to\infty} \exp\left[\log\left(1 - \frac{\mu}{n}\right)^n\right]$$

$$= \lim_{n\to\infty} \exp\left[n \log\left(1 - \frac{\mu}{n}\right)\right]$$

$$= \exp\left[n\left(-\frac{\mu}{n}\right)\right]$$

$$= e^{-\mu}$$

So, set p= μ/n and evaluate

$$P(k) = \lim_{n\to\infty} \frac{n!}{k!(n-k)!} \left(\frac{\mu}{n}\right)^k \left(1 - \frac{\mu}{n}\right)^{n-k}$$

$$= \left(\frac{\mu^k}{k!}\right) \lim_{n\to\infty} \qquad e^{-\mu} \quad \boxed{\left(1 - \frac{\mu}{n}\right)^{-k}}$$

$$\lim_{n\to\infty} \left(1 - \frac{\mu}{n}\right)^{-k} = 1$$

So, set p= μ/n and evaluate

$$P(k) = \lim_{n\to\infty} \frac{n!}{k!(n-k)!}\left(\frac{\mu}{n}\right)^k \left(1-\frac{\mu}{n}\right)^{n-k}$$

$$= \left(\frac{\mu^k}{k!}\right) \lim_{n\to\infty} \qquad e^{-\mu}$$

$$= \boxed{\frac{\mu^k e^{-\mu}}{k!} \quad \text{Poisson Distribution}}$$

Counting statistics, decay processes… ➡ continuous variable is time

Interaction lengths ➡ continuous variable is distance

Radioactive Decay:

1 or more!

What's the probability of detecting a decay from a radioactive source after some time t ?

τ = average time for a decay to occur (mean lifetime)

μ = average # decays in time t, which must be t/τ

Probability for no decays (n=0) within time t

$$P_0 = \left(\frac{\mu^n e^{-\mu}}{n!}\right) \longrightarrow e^{-t/\tau}$$

$$\longrightarrow \quad P_{decay} = 1 - e^{-t/\tau}$$

(integrated over the time interval)

Differential Probability:

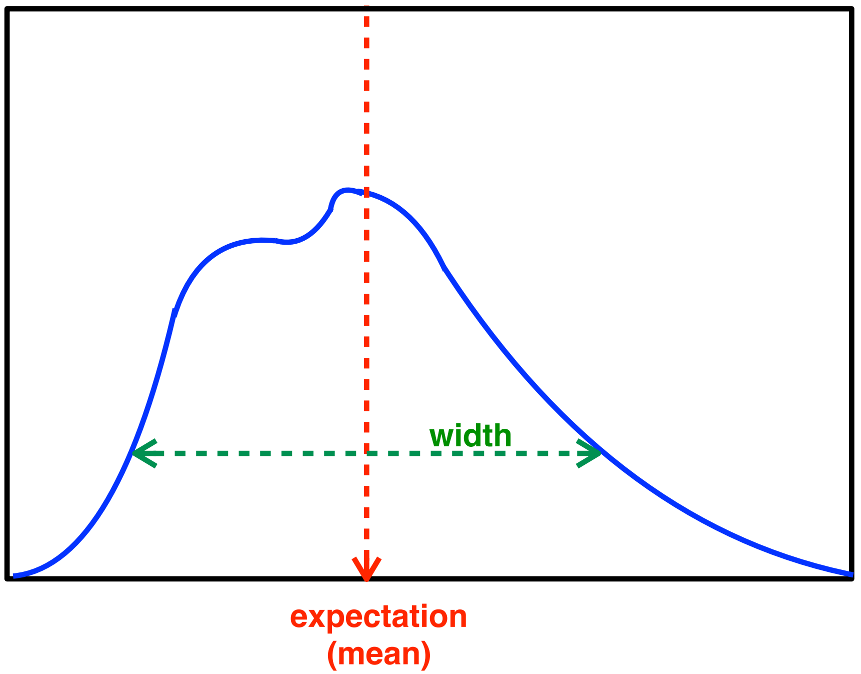
$$P'(t) = \frac{1}{\tau} e^{-t/\tau}$$

Note that this is now a probability for a continuous quantity!

Poisson distribution: the probability of success depends on <u>continuous</u> variable (μ), but the observation is a <u>discreet</u> number of successes (n).

But observations are not always of a discreet variable. For continuous random variables (*i.e.* time, length, *etc.*), the probability of obtaining a particular **exact** value is generally vanishingly small (no phase space!). But the relative probability of getting a value in this vicinity versus that vicinity is meaningful. That's when you talk about "probability densities".

But the terms "probability distribution" and "probability density function" are sometimes informally used interchangeably.


width
expectation
(mean)

Variance: "Average Squared Deviation from Mean"

note: $$\boxed{\langle (x-\mu)^2 \rangle} = \langle x^2 \rangle + \mu^2 - 2\mu\langle x \rangle \boxed{= \langle x^2 \rangle - \mu^2}$$

for Poisson:

$$\langle n^2 \rangle = \sum_{n=0}^{\infty} n^2 \frac{\mu^n}{n!} e^{-\mu} = e^{-\mu} \sum_{n=1}^{\infty} n \frac{\mu^n}{(n-1)!}$$

$$= e^{-\mu} \sum_{n=1}^{\infty} \left[(n-1) \frac{\mu^n}{(n-1)!} + \frac{\mu^n}{(n-1)!} \right] = e^{-\mu} \left[\sum_{n=2}^{\infty} \frac{\mu^n}{(n-2)!} + \sum_{n=1}^{\infty} \frac{\mu^n}{(n-1)!} \right]$$

$$= e^{-\mu} \left[\mu^2 \sum_{n=2}^{\infty} \frac{\mu^{n-2}}{(n-2)!} + \mu \sum_{n=1}^{\infty} \frac{\mu^{n-1}}{(n-1)!} \right] = e^{-\mu} \left[\mu^2 (e^{\mu}) + \mu (e^{\mu}) \right]$$

$$= \mu^2 + \mu$$

$$\boxed{\sigma^2 =} \langle (n-\mu)^2 \rangle = \langle n^2 \rangle - \mu^2 = \boxed{\mu}$$

variance = $\sigma^2 = \langle x^2 \rangle - \mu^2$

Units of σ are same as units of x (or μ)

But, for Poisson, $\sigma^2 = \mu$ How do units work?

Here, μ refers to the expected number of successes, which is unit-less (special case)

$$\sigma = \sqrt{\langle (x-\mu)^2 \rangle} = \sqrt{\langle x^2 \rangle - \mu^2}$$

= “RMS (Root Mean Squared) deviation”

universal

“Standard deviation”

when interpreted in the context of a Normal (Gaussian) distribution

Some Useful Consequences:

- The RMS deviation on a measured number of counts due to statistical fluctuations is the square root of the expected mean number of counts (sqrt of the measured number is often not a bad approximation)

- For a large numbers of events, the expected sensitivity for detecting a signal in a counting experiment in terms of the number of standard deviations above background fluctuations is $\sim S/\sqrt{B}$

- In a counting experiment, the number of signal and background events detected are proportional to the counting time. Thus, the signal sensitivity goes like $\sqrt{T}$ in the large n limit

Estimators

Often we don't know the true mean and variance of a distribution and want to estimate it from the data:

$$\boxed{\tilde{\mu} \simeq \frac{1}{n}\sum_{i=1}^{n} x_i}$$

We want this to be "unbiased," such that the expectation value is equal to the true value

$$\langle \tilde{\mu} \rangle = \left\langle \frac{1}{n}\sum_{i=1}^{n} x_i \right\rangle = \frac{1}{n}\sum_{i=1}^{n} \langle x_i \rangle = \frac{1}{n}(n\mu) = \mu$$

fair enough!

But what about $\tilde{\sigma}^2 \simeq \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$??

Quick Argument:
For every n data points, there are n-1 independent measures of the variance

$$\langle \tilde{\sigma}^2 \rangle = \left\langle \frac{1}{n}\sum (x_i - \overline{x})^2 \right\rangle \quad = \frac{1}{n}\sum \left\langle \left(x_i - \frac{1}{n}\sum x_j \right)^2 \right\rangle$$

$$= \frac{1}{n}\sum_i \left\langle x_i^2 - \frac{2}{n} x_i \sum_j x_j + \frac{1}{n^2}\sum_j \sum_k x_j x_k \right\rangle$$

$$= \frac{1}{n}\sum_i \left\langle \frac{n-2}{n} x_i^2 - \frac{2}{n}\sum_{j\neq i} x_i x_j + \frac{1}{n^2}\sum_j x_j^2 + \frac{1}{n^2}\sum_j \sum_{k\neq j} x_j x_k \right\rangle$$

$$= \frac{1}{n}\sum_i \left[\frac{n-2}{n}\langle x_i^2 \rangle - \frac{2}{n}\sum_{j\neq i}\langle x_i x_j \rangle + \frac{1}{n^2}\sum_j \langle x_j^2 \rangle + \frac{1}{n^2}\sum_j \sum_{k\neq j}\langle x_j x_k \rangle \right]$$

$$= \frac{1}{n}\sum_i \left[\frac{n-2}{n}(\sigma^2 + \mu^2) - \frac{2}{n}(n-1)\mu^2 + \frac{1}{n^2} n(\sigma^2 + \mu^2) + \frac{1}{n^2} n(n-1)\mu^2 \right]$$

$$= \sigma^2 \left(\frac{n-1}{n} \right)$$

Biased!!

So instead take:

$$\tilde{\sigma}^2 \simeq \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Thus cancelling the offending term!

Variance in the Estimated Mean

Note that: $\text{var}(\alpha x) = \left\langle (\alpha x)^2 \right\rangle - \langle \alpha x \rangle^2 \quad = \alpha^2 \left(\left\langle x^2 \right\rangle - \langle x \rangle^2 \right)$

$$= \alpha^2 \text{var}(x)$$

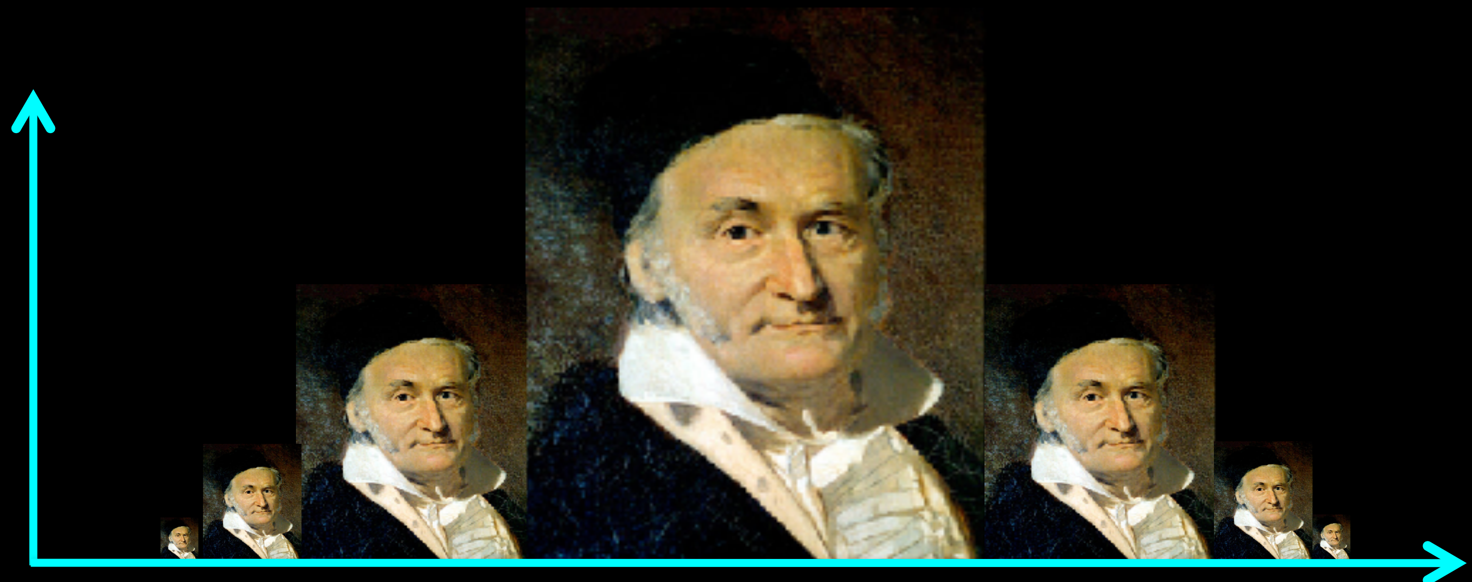
So, consider: $\sigma_m^2 = \text{var}\left(\frac{1}{n} \sum_{i=1}^{n} x_i \right) = \frac{1}{n^2} \text{var}\left(\sum_{i=1}^{n} x_i \right)$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \text{var}\left(x_i \right)$$

For independent variables (as will be shown in lecture 3)

$$= \frac{1}{n^2}\left(n\sigma^2\right) \quad = \frac{\sigma^2}{n} \quad \text{or} \quad \sigma_m = \frac{\sigma}{\sqrt{n}}$$

Gaussian Distributions

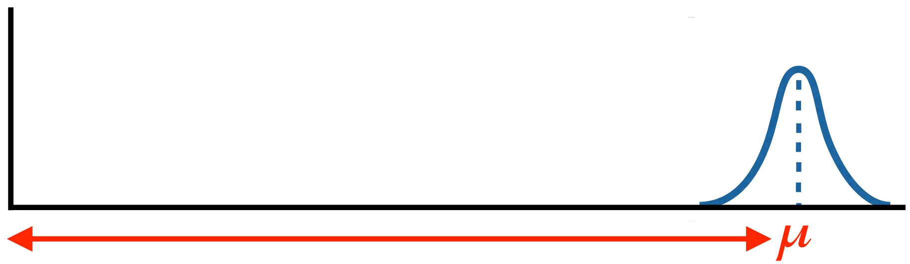


"Everybody believes in the exponential law of errors: the experimenters, because they think it can be proved by mathematics; and the mathematicians, because they believe it has been established by observation"

Gabriel Lippman (1845-1921), as quoted by Poincare

Gaussian (Normal) Distribution as a Limiting Case of Poisson Statistics

Assume μ and n large, with n ~ μ



Define n in terms of a perturbation about μ

$$n = \mu(1+\delta)$$

$$\delta << 1$$

Stirling's Approximation: $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ as $n \to \infty$

So, $$p(n\,|\,\mu) = \frac{\mu^n e^{-\mu}}{n!} \sim \frac{\mu^{\mu(1+\delta)} e^{-\mu}}{\sqrt{2\pi\mu(1+\delta)}\left(\frac{\mu(1+\delta)}{e}\right)^{\mu(1+\delta)}}$$

$$= \frac{\mu^{\mu(1+\delta)} e^{-\mu}}{\sqrt{2\pi\mu(1+\delta)} \left(\frac{\mu(1+\delta)}{e}\right)^{\mu(1+\delta)}} = \frac{\mu^{\mu(1+\delta)} e^{-\mu}}{\sqrt{2\pi\mu}\,\left[\mu^{\mu(1+\delta)}\right] \left[(1+\delta)^{\mu(1+\delta)+\frac{1}{2}}\right] \left[e^{-\mu(1+\delta)}\right]}$$

$$= \frac{e^{\mu\delta}}{\sqrt{2\pi\mu}(1+\delta)^{\mu(1+\delta)+\frac{1}{2}}} \equiv \frac{e^{\mu\delta}}{\sqrt{2\pi\mu}} \frac{1}{g}$$

Define: $f = \ln g = [\mu(1+\delta) + 1/2] \ln(1+\delta)$

Taylor Expand: $f' = \mu \ln(1+\delta) + [\mu(1+\delta) + 1/2]/(1+\delta)$

$(\delta \ll 1,\ \mu \gg 1)$

$$f'' = \frac{\mu}{1+\delta} + \frac{\mu}{1+\delta} - \frac{\mu(1+\delta)+1/2}{(1+\delta)^2}$$

$$f(0) = 0 \qquad f'(0) = \mu + \frac{1}{2} \simeq \mu \qquad f''(0) = \mu - \frac{1}{2} \simeq \mu$$

$$f \sim f(0) + f'(0)\delta + \frac{f''(0)}{2}\delta^2 = \mu\delta + \frac{\mu\delta^2}{2}$$

$$f \sim \mu\delta + \frac{\mu\delta^2}{2} \longrightarrow g \sim e^{\mu\delta + \mu\delta^2/2}$$

$$p(n\,|\,\mu) = \frac{e^{\mu\delta}}{\sqrt{2\pi\mu}}\frac{1}{g} \sim \frac{1}{\sqrt{2\pi\mu}}e^{\mu\delta - \mu\delta - \mu\delta^2/2}$$

$$= \frac{1}{\sqrt{2\pi\mu}}e^{-\mu\delta^2/2}$$

$$= \frac{1}{\sqrt{2\pi\mu}}e^{-(\mu\delta)^2/2\mu}$$

$$= \frac{1}{\sqrt{2\pi\mu}}e^{-(n-\mu)^2/2\mu}$$

recall:

$$n = \mu(1+\delta)$$
$$= \mu + \mu\delta$$

$$\mu \xrightarrow[\text{(Poisson)}]{} \sigma^2$$

$$\boxed{p(n\,|\,\mu) = \frac{1}{\sqrt{2\pi\sigma}}e^{-(x-\mu)^2/2\sigma^2}}$$

Central Limit Theorem

The distribution of sample averages, where samples are randomly drawn from an **arbitrary** distribution, will converge to that of a Normal distribution as the number of samples becomes large

Given any arbitrary probability distribution, $P(x)$, the Moment Generating Function is defined as:

$$M_x(t) \equiv \left\langle e^{tx} \right\rangle = \int e^{tx} P(x)\ dx$$

$$M_x(0) = \int P(x)\ dx = 1$$

$$M_x'(0) = \int xP(x)\ dx = \langle x \rangle$$

$$M_x''(0) = \int x^2 P(x)\ dx = \left\langle x^2 \right\rangle$$

etc.

Note that, for a Normal distribution with zero mean and unit variance:

$$M_N(t) = \int \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} e^{tx}\ dx$$

$$= \frac{1}{\sqrt{2\pi}} \int e^{-\frac{1}{2}(x-t)^2 + \frac{t^2}{2}}\ dx = e^{\frac{t^2}{2}}$$

Take n samples from any arbitrary distribution, and define:

$$Y_i \equiv \frac{x_i - \mu}{\sigma} \qquad Z \equiv \frac{1}{n} \sum \frac{(x_i - \mu)}{\sigma/\sqrt{n}} = \frac{1}{\sqrt{n}} \sum Y_i$$

$\longrightarrow$ Both Y and Z have zero mean & unit variance

$$M_Z(t) = \left\langle e^{tZ} \right\rangle = \left\langle e^{\frac{t}{\sqrt{n}} \sum Y_i} \right\rangle = \left\langle \prod_{i=1}^{n} e^{\frac{t}{\sqrt{n}} Y_i} \right\rangle = \left\langle e^{\frac{t}{\sqrt{n}} Y} \right\rangle^n = \left[M_Y\left(\frac{t}{\sqrt{n}} \right) \right]^n$$

for independent, identically distributed data (IID)

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = M_Y(0) + M_Y'(0)\frac{t}{\sqrt{n}} + M_Y''(0)\frac{t^2}{2n} + M_Y'''(0)\frac{t^2}{3!\, n^{3/2}} + \dots$$

$$= 1 + \langle Y \rangle \frac{t}{\sqrt{n}} + \left\langle Y^2 \right\rangle \frac{t^2}{2n} + \left\langle Y^3 \right\rangle \frac{t^2}{3!\, n^{3/2}} + \dots$$

$$= 1 + \quad 0 \quad + \quad \frac{t^2}{2n} \quad + \quad \left\langle Y^3 \right\rangle \frac{t^2}{3!\, n^{3/2}} + \dots$$

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = 1 + \quad 0 \quad + \quad \frac{t^2}{2n} \quad + \quad \langle Y^3 \rangle \frac{t^2}{3!\ n^{3/2}} + \ldots$$

$$M_Z(t) = \left[1 + \frac{t^2}{2n} + \langle Y^3 \rangle \frac{t^2}{3!\ n^{3/2}} + \ldots\right]^n$$

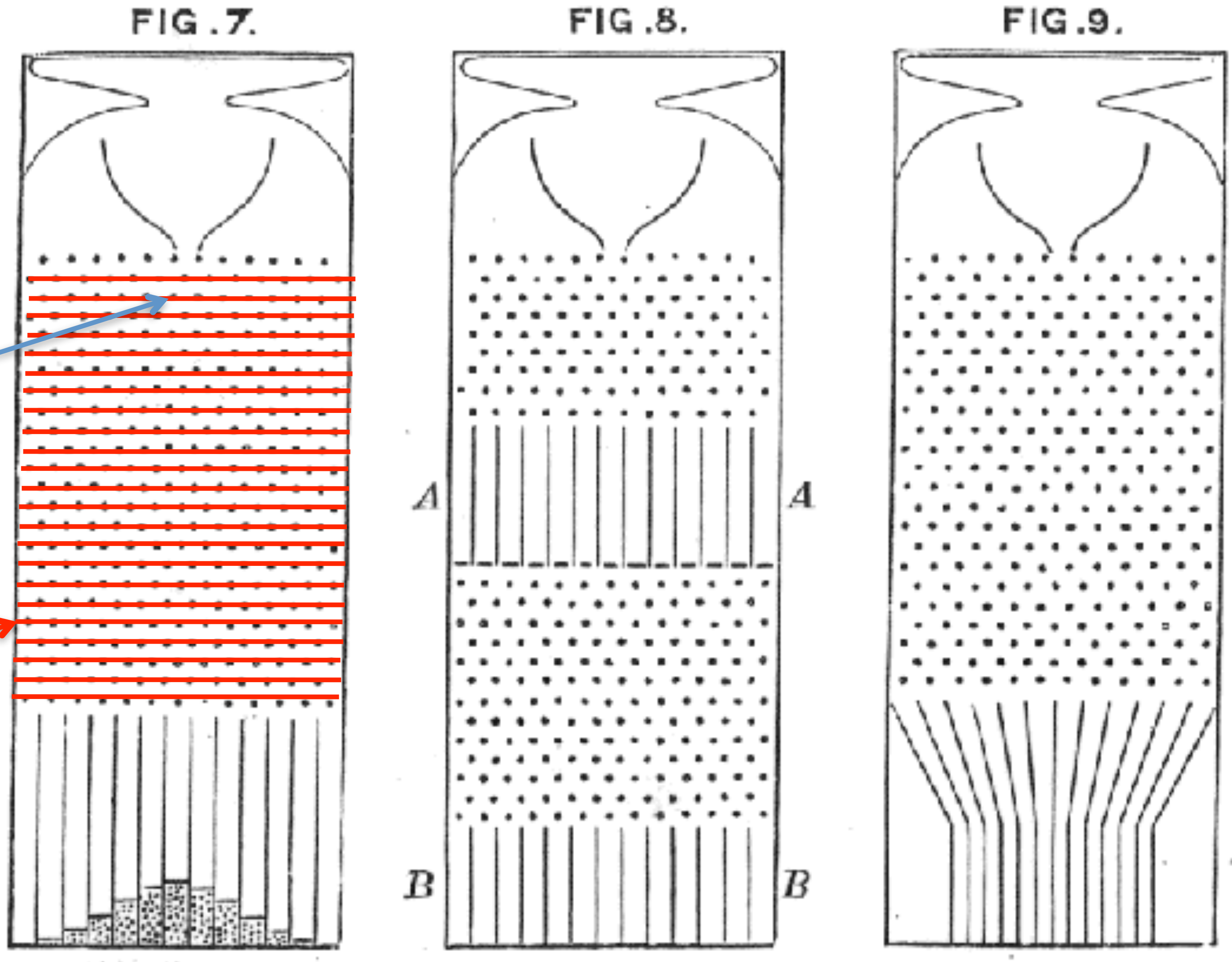
$$= \left[1 + \frac{\frac{t^2}{2} + \langle Y^3 \rangle \frac{t^2}{3!\ n^{1/2}} + \ldots}{n}\right]^n$$

$$\lim_{n\to\infty} M_Z(t) = \lim_{n\to\infty}\left[1 + \frac{\frac{t^2}{2}}{n}\right]^n = e^{\frac{t^2}{2}} = M_N(t)$$

So moments that would be generated from Z in this limit are identical to moments that would be generated from a Normal distribution!!

Galton Machine

Sir Francis Galton



ball has 50:50 chance of going right or left at each peg (underlying distribution)

"N" peg interactions (samples) get added to form the total deflection