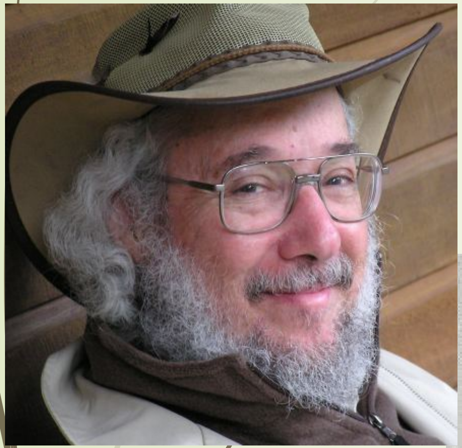
Fate of a Gambler:

A cautionary tale for cavalier applications of the central limit theorem

Joe Redish



Harvey Gould

Kevin Bassler

R.K.P. Zia

Center for Soft Matter and Biological Physics,
Physics Department, Virginia Tech
and
Department of Physics, University of Houston

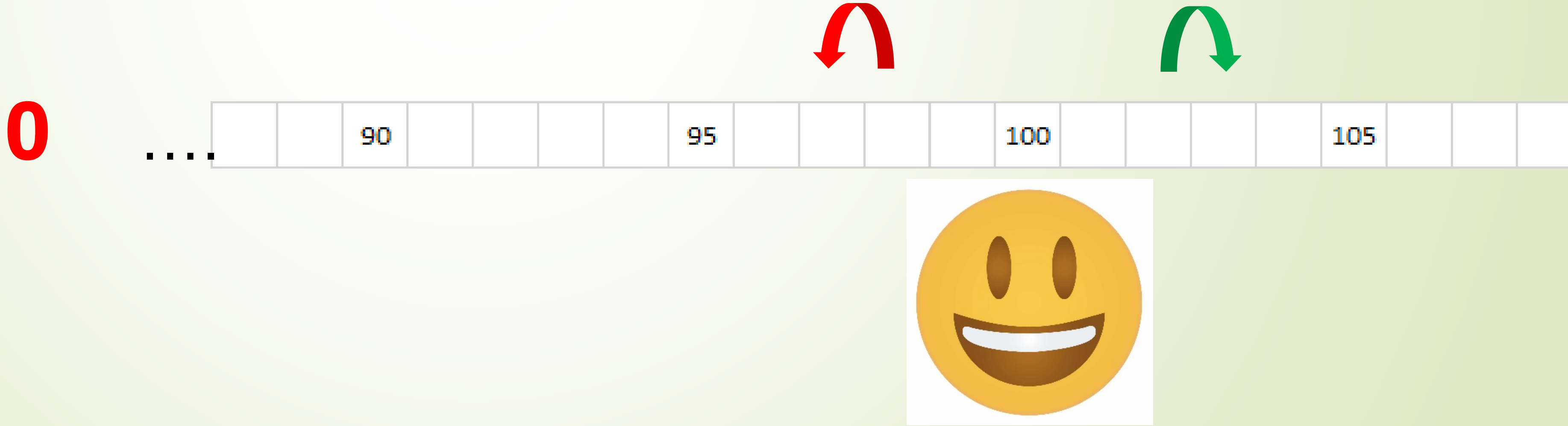
Standard Gambler's Ruin:

A fair coin is tossed. Each time a gambler wins/loses $1 against the house.
Starting with $100, say, he/she will
eventually ***lose it all***.

The earliest known mention of the gambler's ruin problem is a letter from Blaise Pascal to Pierre Fermat in **1656**
https://en.wikipedia.org/wiki/Gambler's_ruin

Standard Gambler’s Ruin:

~ simple Random Walk (unbiased) on a line



A *percentage* version:

A fair coin is tossed. Each time a gambler wins/loses a ***fixed percentage*** (e.g., 10%) of his/her wealth (at the time of the toss) .

He/she will *never* "lose it all", but what can we expect after N tosses?

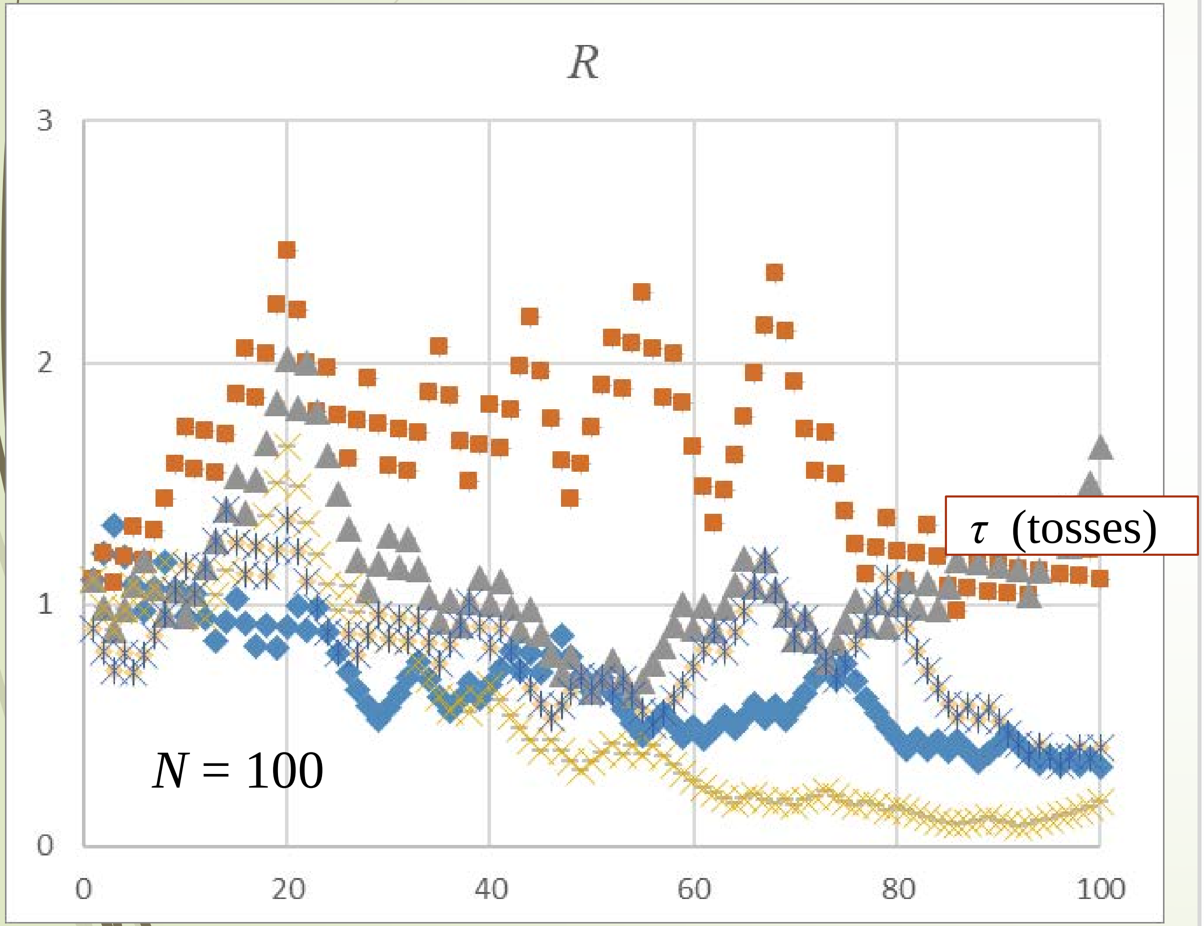
Intuitive guess:

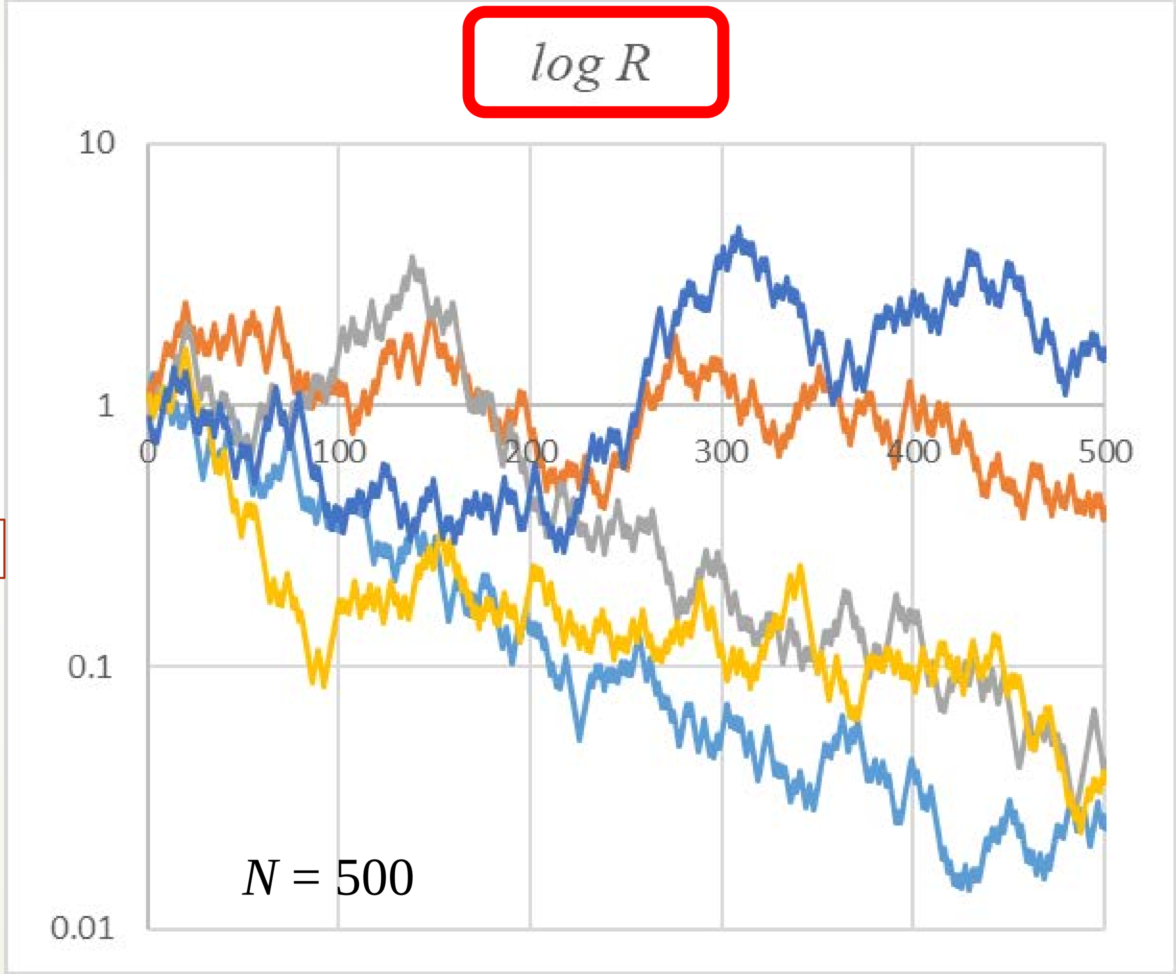
Since the coin is fair, the "return"

$$R \equiv \frac{\text{wealth (at the end)}}{\text{wealth (at start)}}$$

should be **1** (on *average*, for any N)!

Yet, simulations show…





Yet, simulations seem

… to indicate

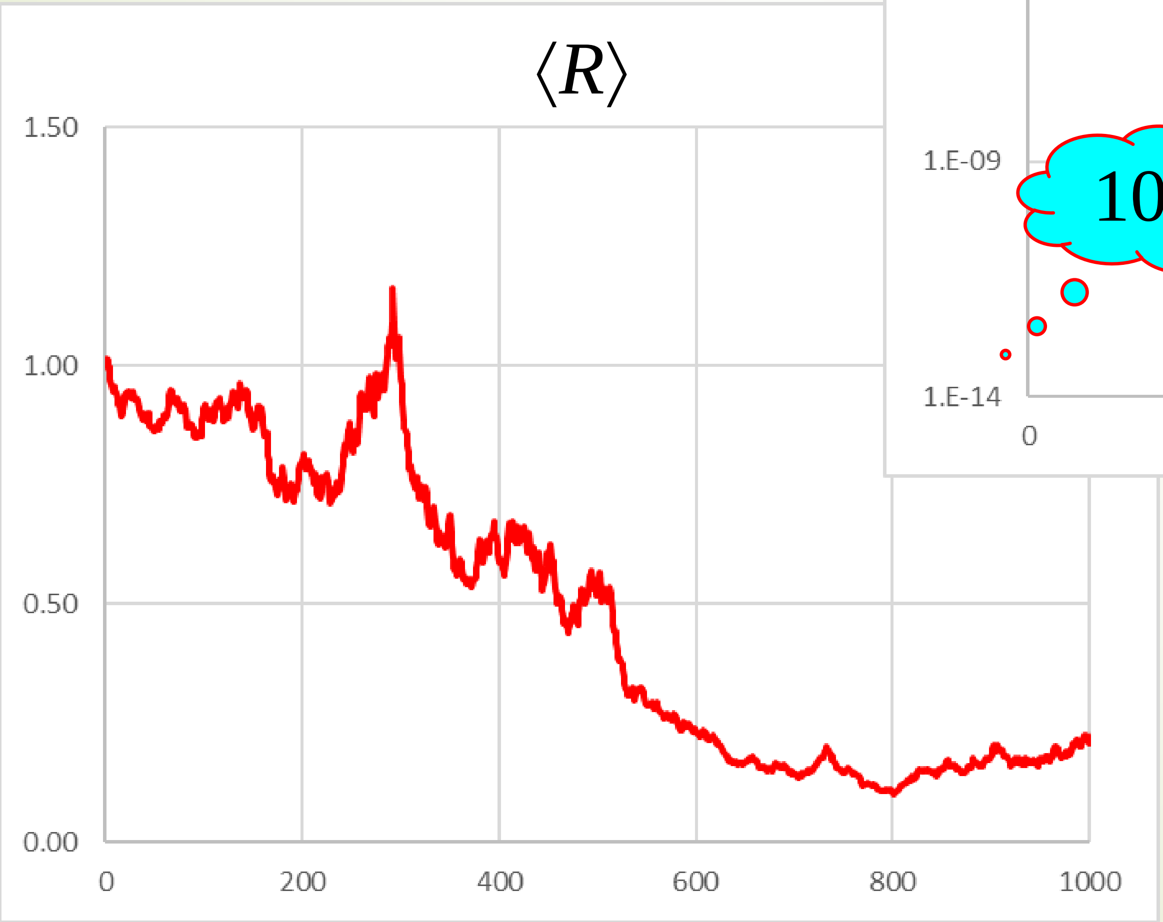
R dropping *precipitously* with N !

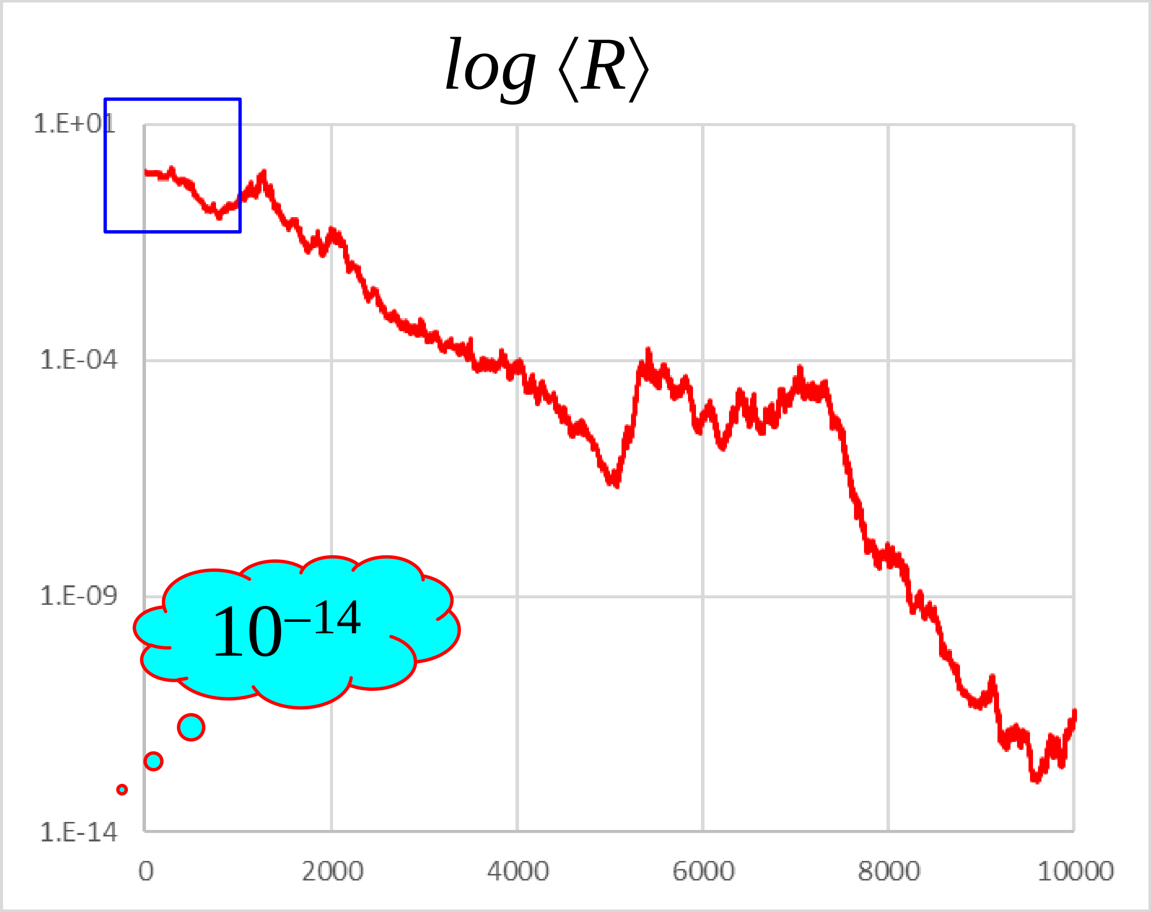
e.g., $\langle R \rangle$ - average over 100 runs,

for N= 1K and 10K:

⟨…⟩ average over 100 runs

$N = 1000$





$N = 10K$

OTOH, if we invoke the Central Limit Theorem,

$\langle R \rangle$ *increases* exponentially with N !

What's the CLT? and
Why should we invoked here?

Intuitive guess is right!

Exact computation shows

$$\langle R \rangle \equiv 1$$

for *any* N !!

Intuitive guess is right!

So, how do both
simulations and **CLT**
lead us astray
… so *badly* !?

For the curious:
It's not a hard calculation.
Ask me at the end if you want.

Here's an initial hint that this game is ***UNFAIR!***

An fair coin means we "win/lose" half the time:

$$W = N/2$$

That's no good enough, since we really need *R*!

BUT, we don't break even (R=1) when we "win" half the time!

$100

$90 $110

$81 $99 $121

lose

win

alf the time:

y need *R*!

1)

alf the time!

3/4 of the time, you ***LOSE***!

YET, if you average over all possible outcomes,

the game is *EVEN* !

–19 –1 –1 +21

$100

lose

$90 $110

win

$81 $99 $121

1/4 *1/2* *1/4*

From one toss to the next,

$$R(\tau+1) = \begin{matrix} [1+s] \\ [1-s] \end{matrix} R(\tau)$$

s is the "stake" e.g.,10%

From one toss to the next,

$$R(\tau+1) = [1+xs]\,R(\tau)$$

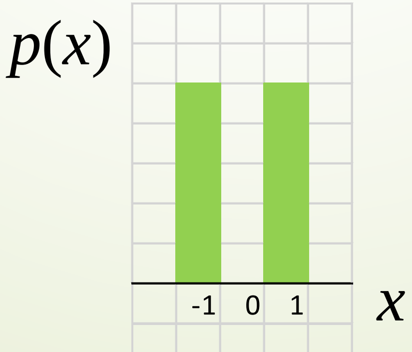
s is the "stake" e.g.,10%

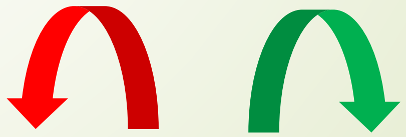
$x = \pm 1$ for win/loss of the toss.

So, a “history” of the game is given by a string of *x*’s:

$$x_1,\ x_2,\ \ldots\ x_\tau\ ,\ \ldots\ x_N$$

just like in the STANDARD game





So, a "history" of the game is given by a string of *x*'s, and at the end,

R is a *product* of these factors:

$$R = [1+x_N s] \ldots [1+x_1 s]$$

To change the ***product*** of a string

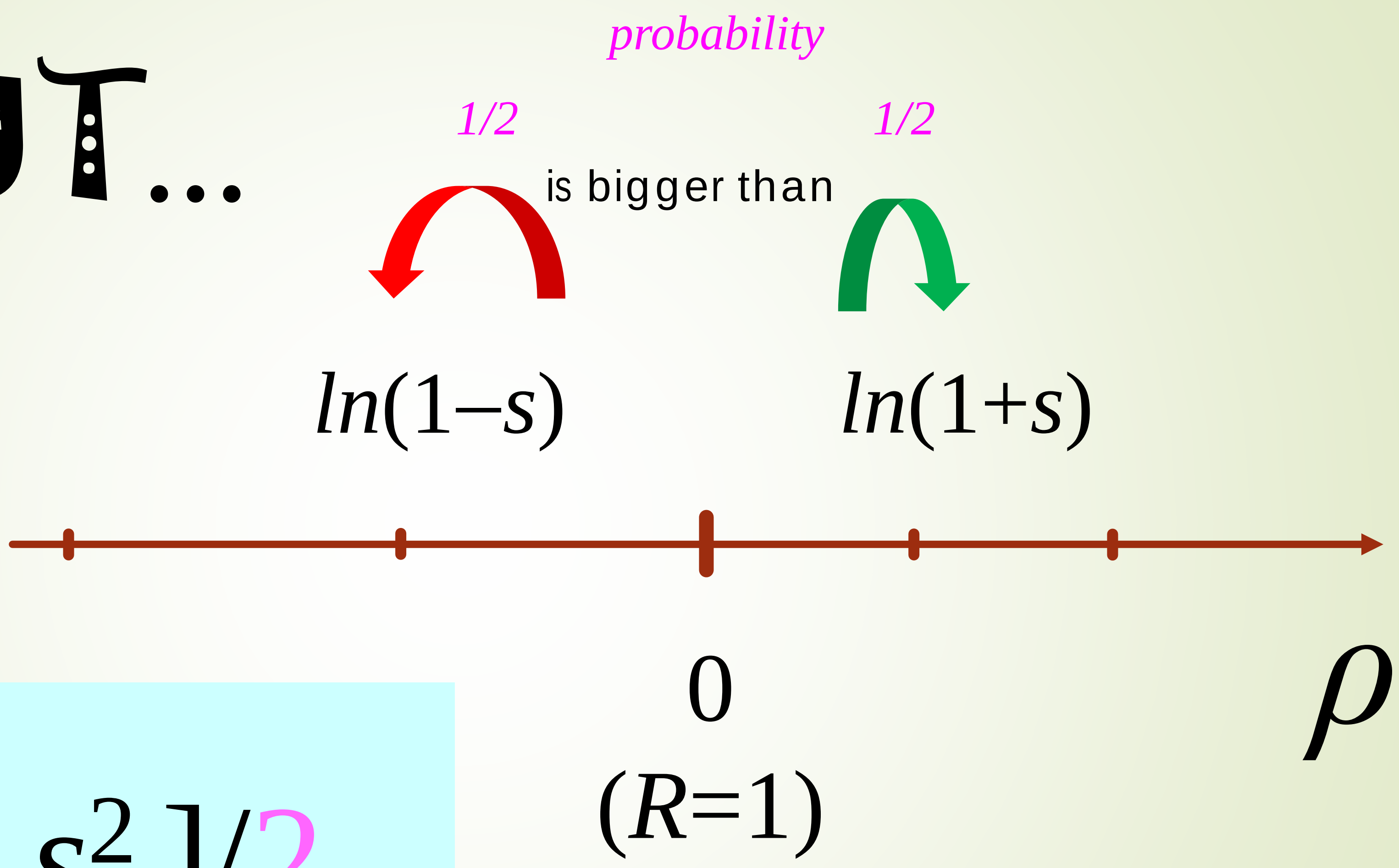
to the ***sum***

(so that we can use the mapping to a random walk and CLT),

just use *logarithm*!!

$$\rho_N \equiv \ln R(N) = \sum_{\tau=1}^{N} \ln\left[1 + x_\tau s\right]$$

BUT…



$$\mu = ln[1 - s^2]/2$$

$$\cong - s^2/2 < 0$$

BIASED Random Walk (in ρ space)!

…with <u>unequal steps</u>

On the average, you ***LOSE*** !

$$\mu = ln\,[1 - s^2\,]/2 \;\cong - s^2/2$$

$$\sigma^2 = -\,[ln(1{-}s)/ln(1{+}s)]^2/4 \;\cong\; s^2$$

To get an idea of how badly off we are with this

BIASED Random Walk,

…let's exploit the CLT

to see what $P(\rho_N)$ is like

(after N tosses).

For the experts:
The exact P is just a fancier binomial. But, getting the fraction of loss is not easy: No closed form for partial sums of binomials! Thus, CLT. Also, for a dice instead, exact P would be impossible!

What's the CLT? and

Suppose x is a random variable picked from some distribution, $p(x)$, with finite mean μ and variance σ^2

st. dev. σ

Generate N of them and ***add***:

$$X \equiv x_1 + \ldots + x_N$$

…and call the induced distribution for X $P(X)$.

For the experts and the curious: $\tilde{P} = [\tilde{p}]^N$

What's the CLT? and

An excellent approximation for P is $\mathcal{N}$, the ***normal*** (Gaussian) distribution, if N is large.

To be specific, $\mathcal{N}(X)$ has
mean $N\mu$ and variance $N\sigma^2$

All you need from p
are its mean and variance!

What's the CLT? and

An excellent approximation for P is $\mathcal{N}$,
the ***normal*** (Gaussian) distribution, if N is large.

s
ce $N\sigma^2$

All other properties of p
are " ***irrelevant*** " !!

All you need from p
are its mean and variance!

i.e., CLT assures us that $P(\rho_N) \cong$

$$\mathcal{N}(\rho_N) \propto exp\left(-\frac{(\rho_N - N\mu)^2}{2N\sigma^2}\right)$$

$\mu \cong -s^2/2$

$\sigma \cong s$

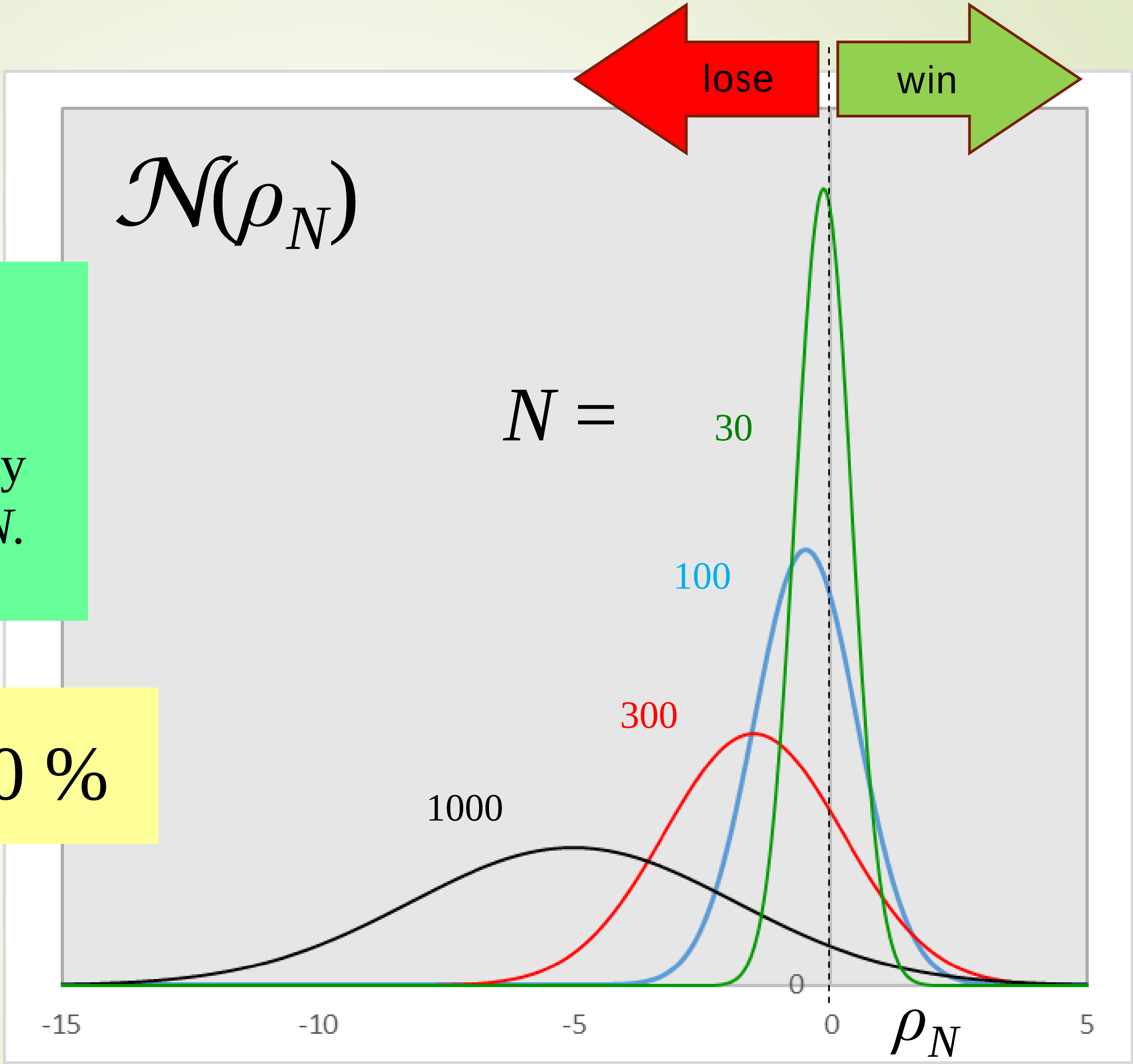
mean $= N\mu$ (*receding* as N !!)

and outpaces st. dev. $\sqrt{N}\sigma$ …

Thus, the fraction of times $\rho < 0$ (lose)
…increases dramatically with N.

$S = 10\ \%$

mean $\cong -Ns^2/2$
st.dev. $\cong \sqrt{N}s$



Thus, the fraction of times $\rho < 0$

…increases dramatically with N.

The CLT has helped us appreciate how simulations gave us the *terrible losses* !

For the experts: Fraction of winning histories ~ $\frac{e^{-N\Phi^2/2}}{\sqrt{2\pi N}\,\Phi}$

But then, the CLT can lead us *ASTRAY*, if we carelessly compute $\langle R \rangle$ with it…

So,

$$\langle R \rangle_{normal} = \langle e^{\rho} \rangle_{normal} =$$

$$\int e^{\rho} \, exp\left\{ - \frac{(\rho - N\mu)^2}{2N\sigma^2} \right\}$$

But then, the CLT can lead us *ASTRAY*, if we carelessly compute $\langle R \rangle$ with it…

can prove this is
POSITIVE!

So,

$$\langle R \rangle_{normal} = \langle e^{\rho} \rangle_{normal} =$$

$$exp\ N\{\mu + \sigma^2/2\}$$

In other words, CLT lead us
to believe that return should…

diverge exponentially with N !!

…an advice that brings
assured ruin.

So, what went so wrong?

(with exploiting the CLT here)

What's the CLT? and Why should we invoked here?

An excellent approximation for P is $\mathcal{N}$, the ***normal*** (Gaussian) distribution, if N is large.

How excellent?

What's "large"?

What's the CLT? and

Uniform

convergence

How excellent?

excellent appro … N,

the ***normal*** (Gaussian) distri … N is large.

If you demand a maximum error of ε ,

you can find an N (that depends on ε) …

so that $\mathcal{N}$ differs from P by less than ε

for ALL X!!

For the experts/purists:
It's actually the CDF, not the PDF, that's uniformly convergent.

What’s the CLT? and

What’s “large”?

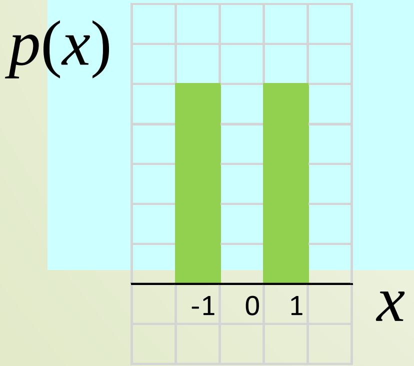
An excellent approximation for *P*

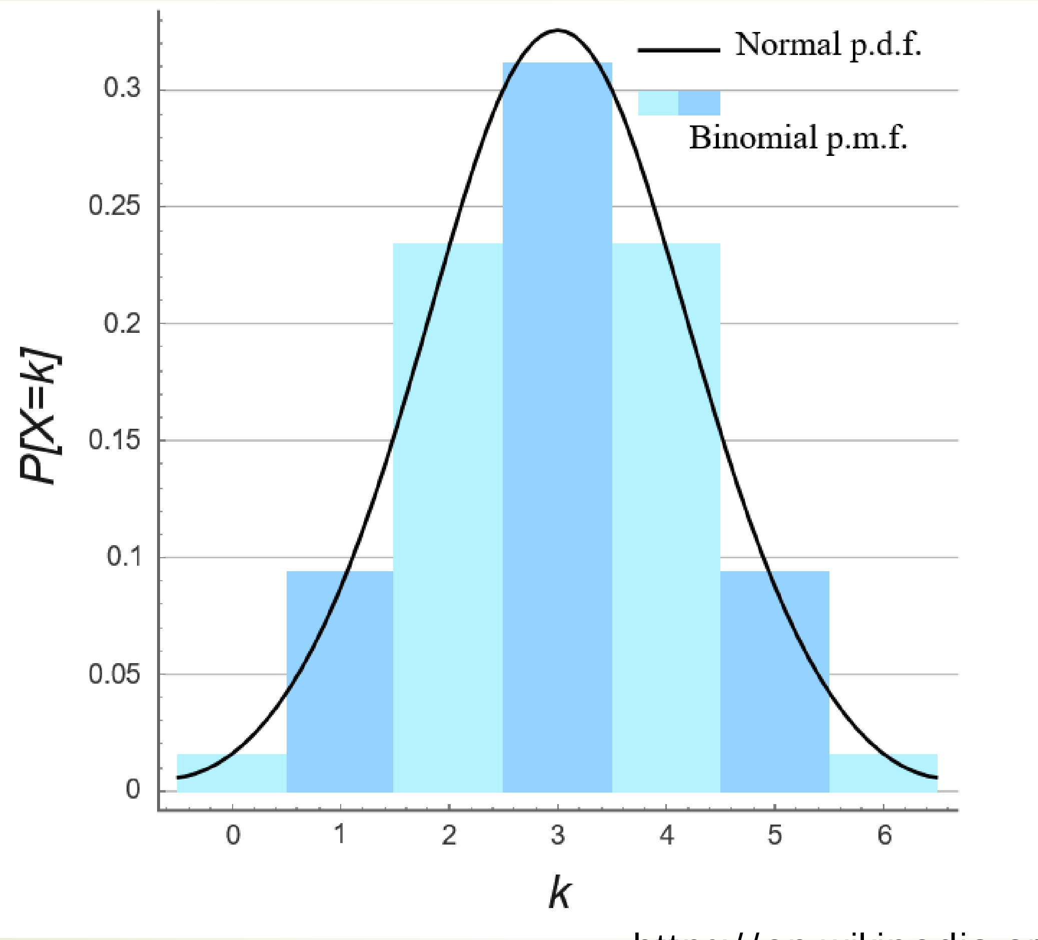
the ***normal*** (Gaussian) distribution, if N is large.

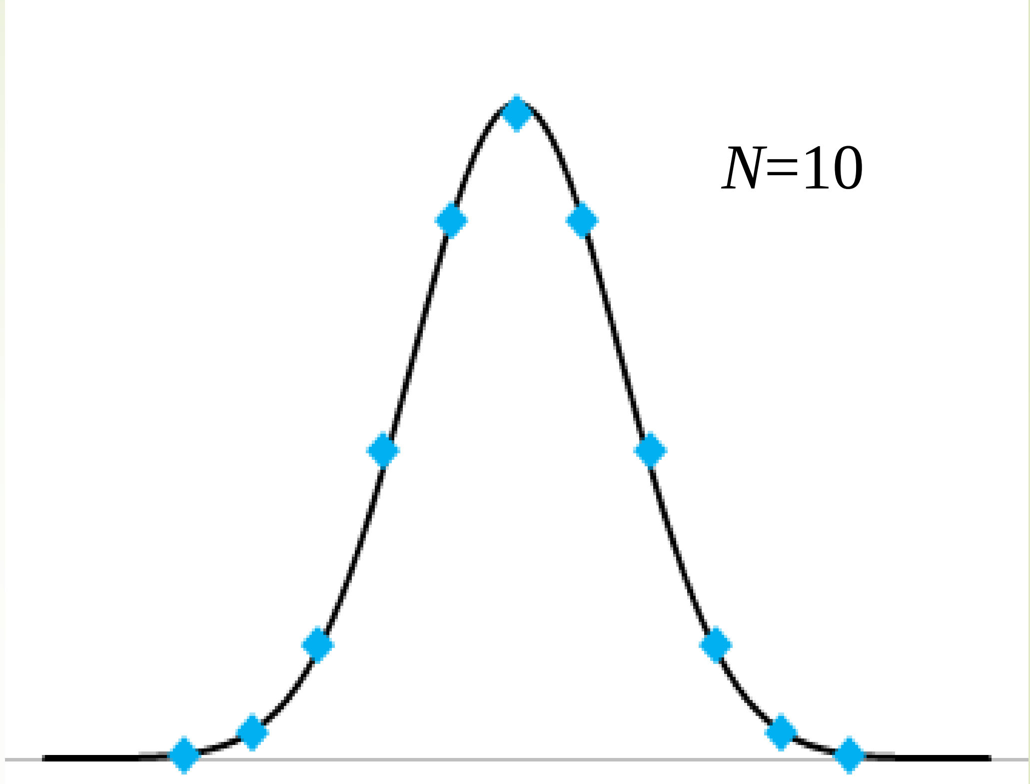
For an example of “how large,” consider the “fair coin” case … labeling $x = \pm 1$ for H/T, so that $\mu = 0$ and $\sigma = 1$.

Meanwhile, $X = -N, -N+2, \ldots, N$

and $P(X)$ is just the binomial.

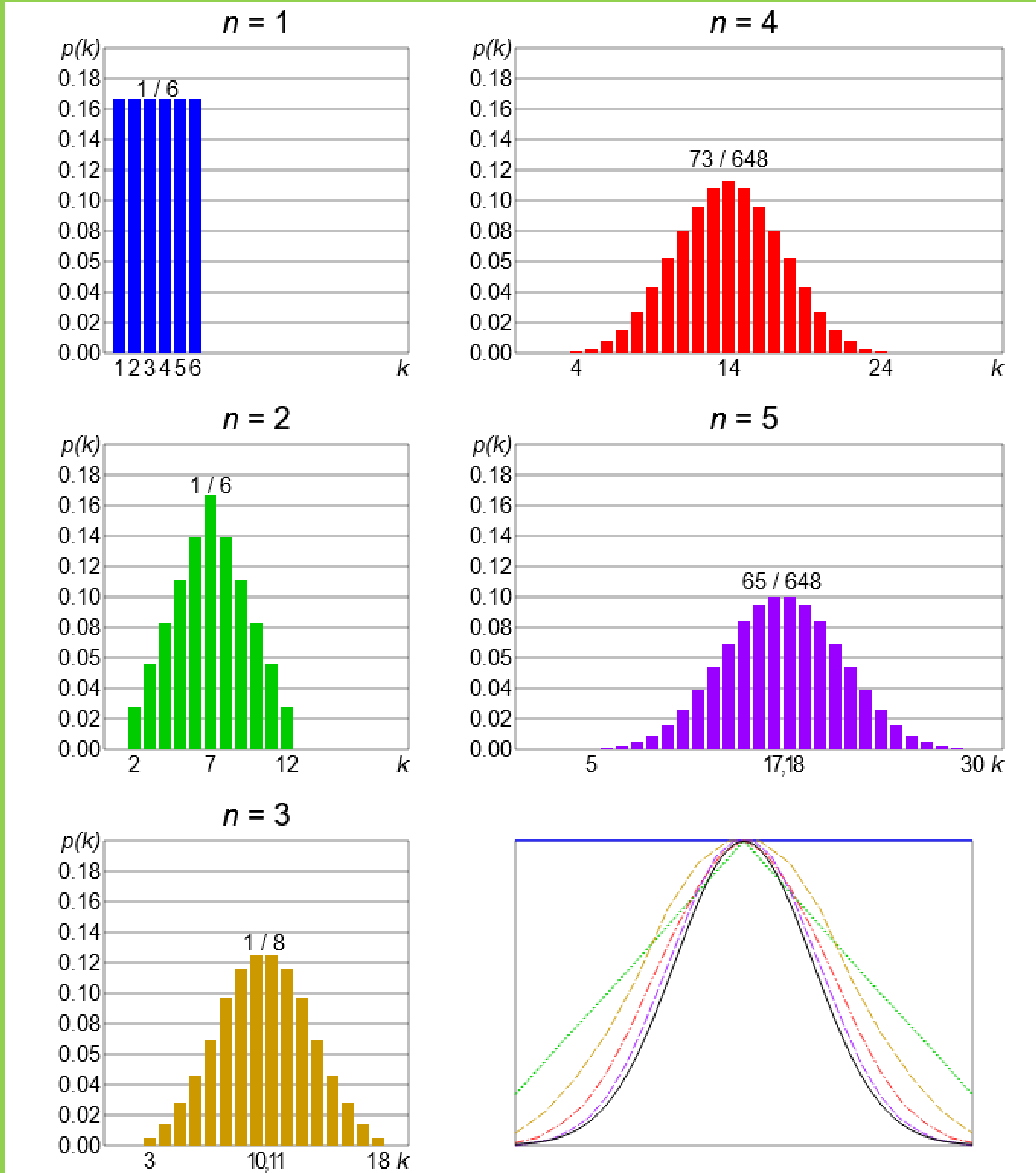






https://en.wikipedia.org/wiki/Binomial_distribution

https://en.wikipedia.org/wiki/Normal_distribution



What's the CLT? and Why should we invoked here?

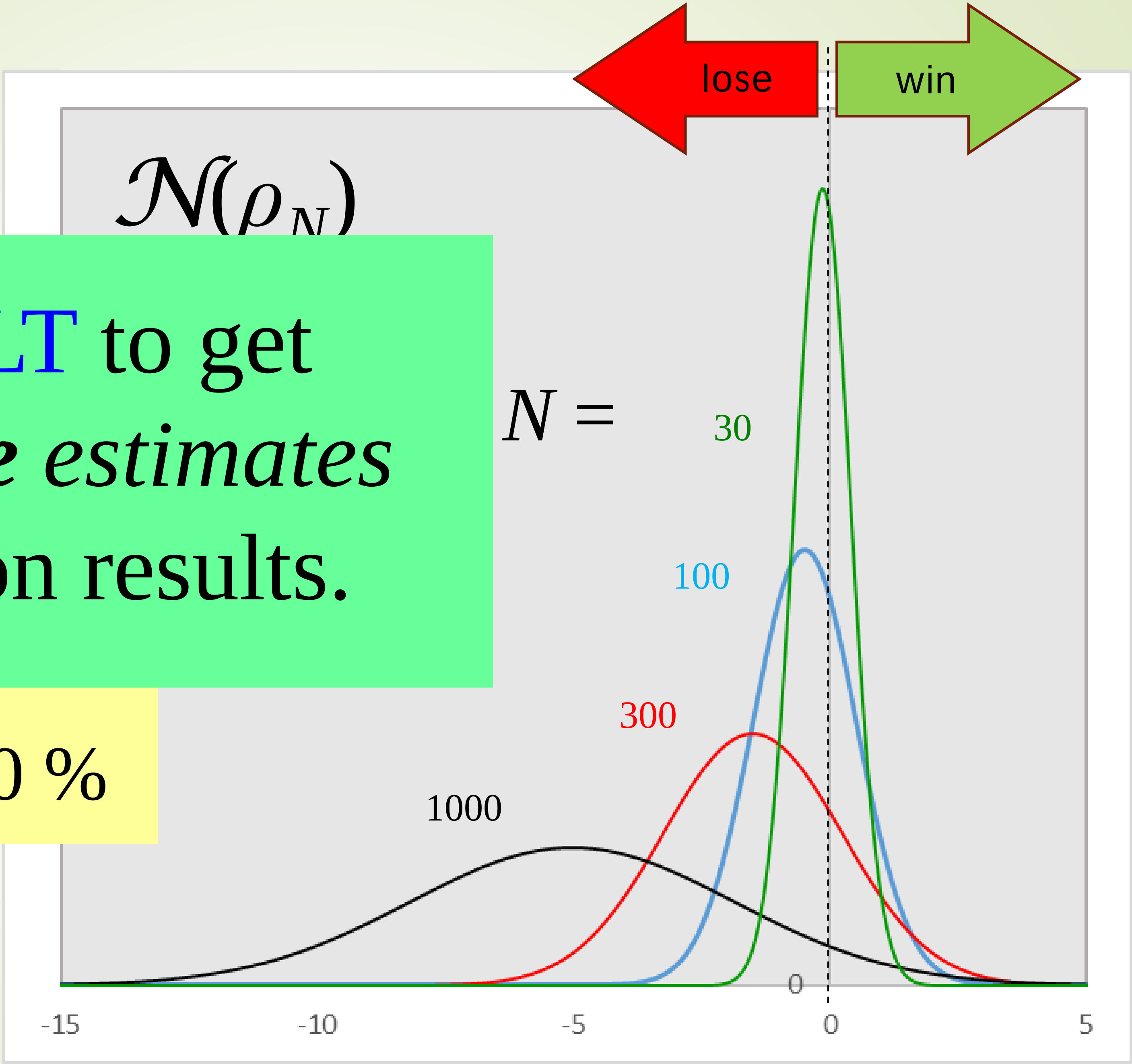
Add N x's of ± 1 and binomial $P(X) \cong \mathcal{N}(X\,;0,N)$

Add N ρ's of $\pm$ whatever and fancy binomial

$$P(\rho_N) \cong \mathcal{N}(\rho_N\,;\,N\mu,\sqrt{N}\sigma\,)$$

All we need are these!

We used the CLT to get *good* ***quantitative*** *estimates* for the simulation results.



$\boldsymbol{S}$ = 10 %

$\mu \cong -Ns^2/2$

$\sigma \cong \sqrt{N}s$

So, what went so wrong? using the CLT to find $\langle R \rangle$?

Though $P(\rho) \cong \mathcal{N}(\rho)$ is excellent for *all* ρ,

$e^{\rho} P(\rho)$ and $e^{\rho} \mathcal{N}(\rho)$

may be quite far apart for some ρ !!

“e^{ρ} has long/large tails !!”

…so that integrals over the above can be quite different.

So, what went so wrong? using the CLT to find $\langle R \rangle$?

Though $P(\rho) \cong \mathcal{N}(\rho)$ is excellent for *all* ρ

!!

For the experts:

Though $\int^X \mathcal{N}(\xi)$ converges *uniformly* to $\int^X P(\xi)$,

$\int^X e^{\xi} \mathcal{N}(\xi)$ does not c.u. to $\int^X e^{\xi} P(\xi)$!

…so that integrals over the above can be quite different.

Take-home messages:

What the CLT tells us is "impressive"!   <u>**BUT**</u>,
… if you want to find the average of anything (associated with *P*),
then you'd better look at the tails of that anything (before using the CLT blindly).

For the experts:
If you want to find the average of some function, f, of the "macroscopic variable" X, then …

- If your f is a function of $\xi \equiv X/N$ *<u>alone</u>*, then, the CLT assures you that $\langle f(\xi) \rangle \to f(\mu)$ for $N \to \infty$.
- If your f is a function of $\xi \equiv X/\sqrt{N}\sigma$ *<u>alone</u>*, AND you know μ=0, then, the CLT assures you that you can use the standard normal distribution to get $\langle f(\xi) \rangle$ for $N \to \infty$.
- In all other cases, study the tails of f *<u>carefully</u>* before relying on the CLT for $\langle f \rangle$!

For the aficionados:
ALL cumulants (beyond the first two, predicted by the CLT) are "***infinitely wrong***"!

Conclusion

If you teach/use the Central Limit Theorem,

please consider an extra warning label:

Do NOT blindly compute ⟨•⟩

with the Gaussian approximant!

Consider carefully …

…the tails of the • first.

Fate of a gambler:

A cautionary tale for cavalier applications of the central limit theorem.

Tossing a fair coin N times, a gambler wins/loses 10% of his/her holdings against the house if each toss is head/tail (H/T). Measuring his/her fortunes by R (ratio of final to initial wealth), then we may ask for $\langle R \rangle$ (the average over all possible 2^N histories). Since the game sounds like it's even, we may guess $\langle R \rangle=1$. When computed exactly, it is indeed so. Yet, when simulations are done, $\langle R \rangle$ drops exponentially with N, e.g., to $O(10^{-14})$ for $N=10^4$. A further puzzle is the following: It is tempting to apply the central limit theorem and replace the distribution of $H-T$ by a normal (since the exact one is just a binomial, in lnR). Replying on that leads to an $\langle R \rangle$ that *increases* exponentially with N! Along with resolutions to these paradoxes, I propose that we add a "warning label" when the central limit theorem is taught.