±

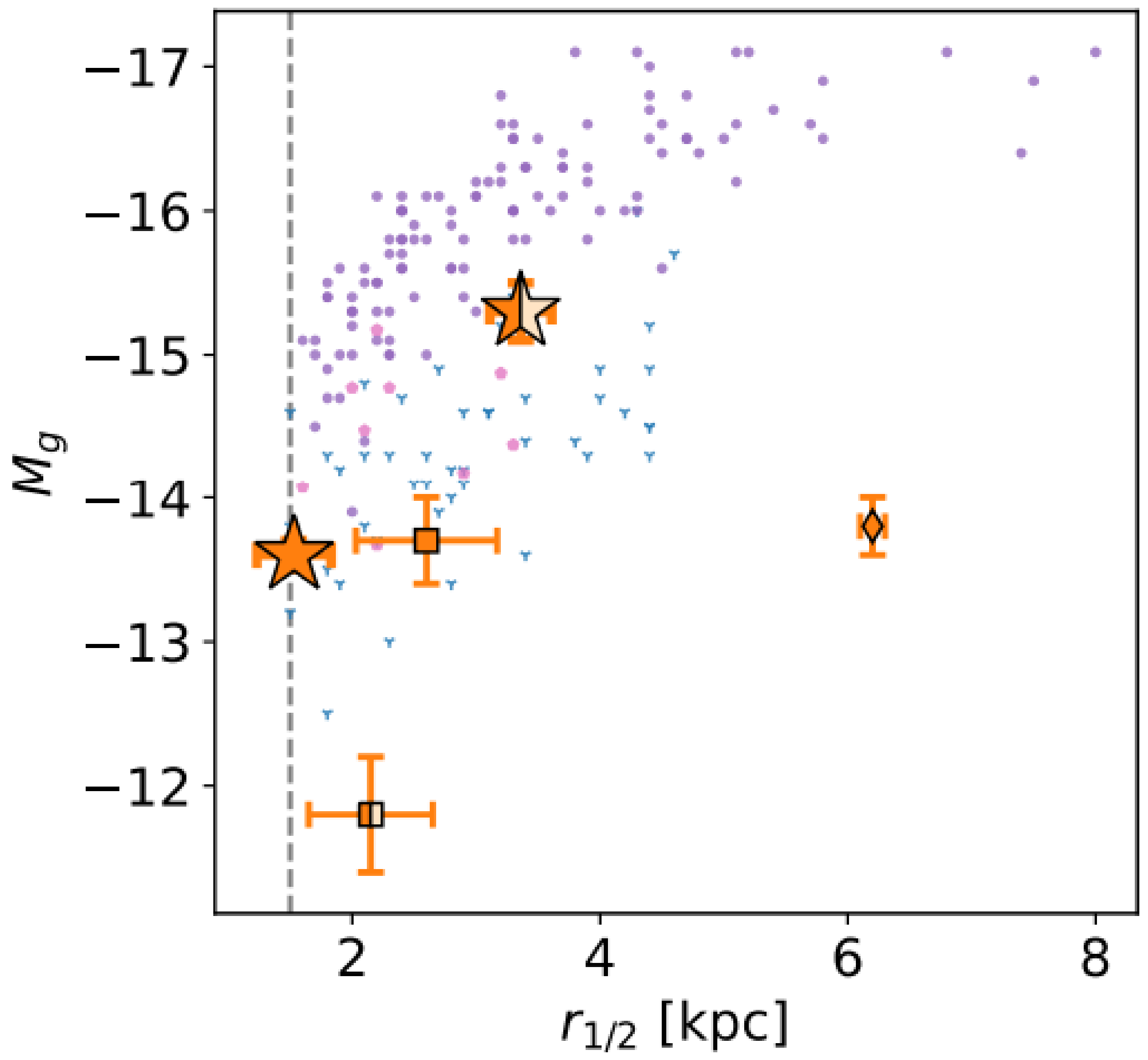
errors

$\pm$

uncertainties?

~~errors~~

what do error bars mean?

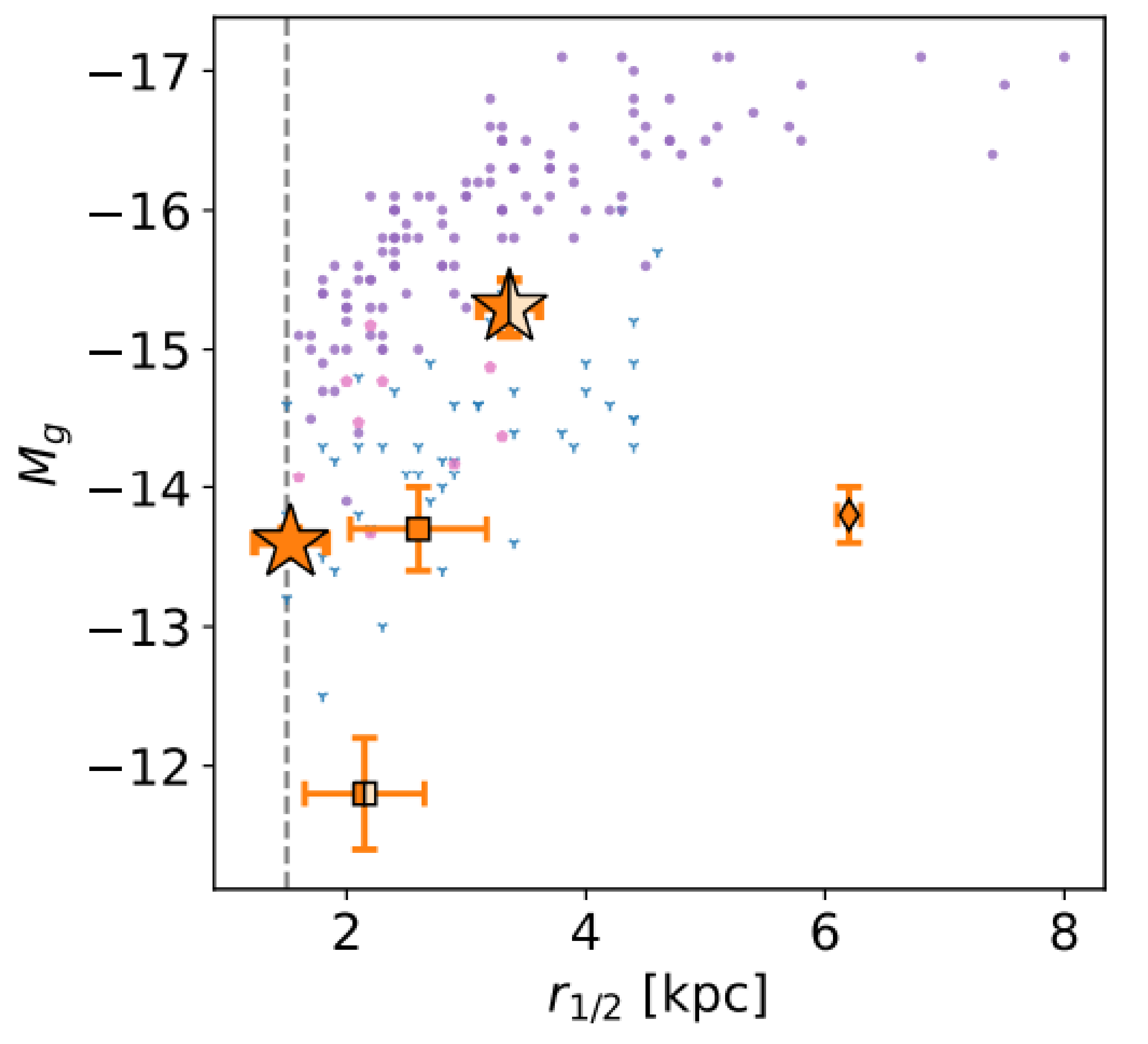


what do these errors mean?

SFR (NUV; $M_\odot$ yr^{-1} $\times 10^{-3}$)	6.2 ± 2.2
SFR (FUV; $M_\odot$ yr^{-1} $\times 10^{-6}$)	78 ± 41

plot and results from Fielder et al. 2024

what do error bars mean?



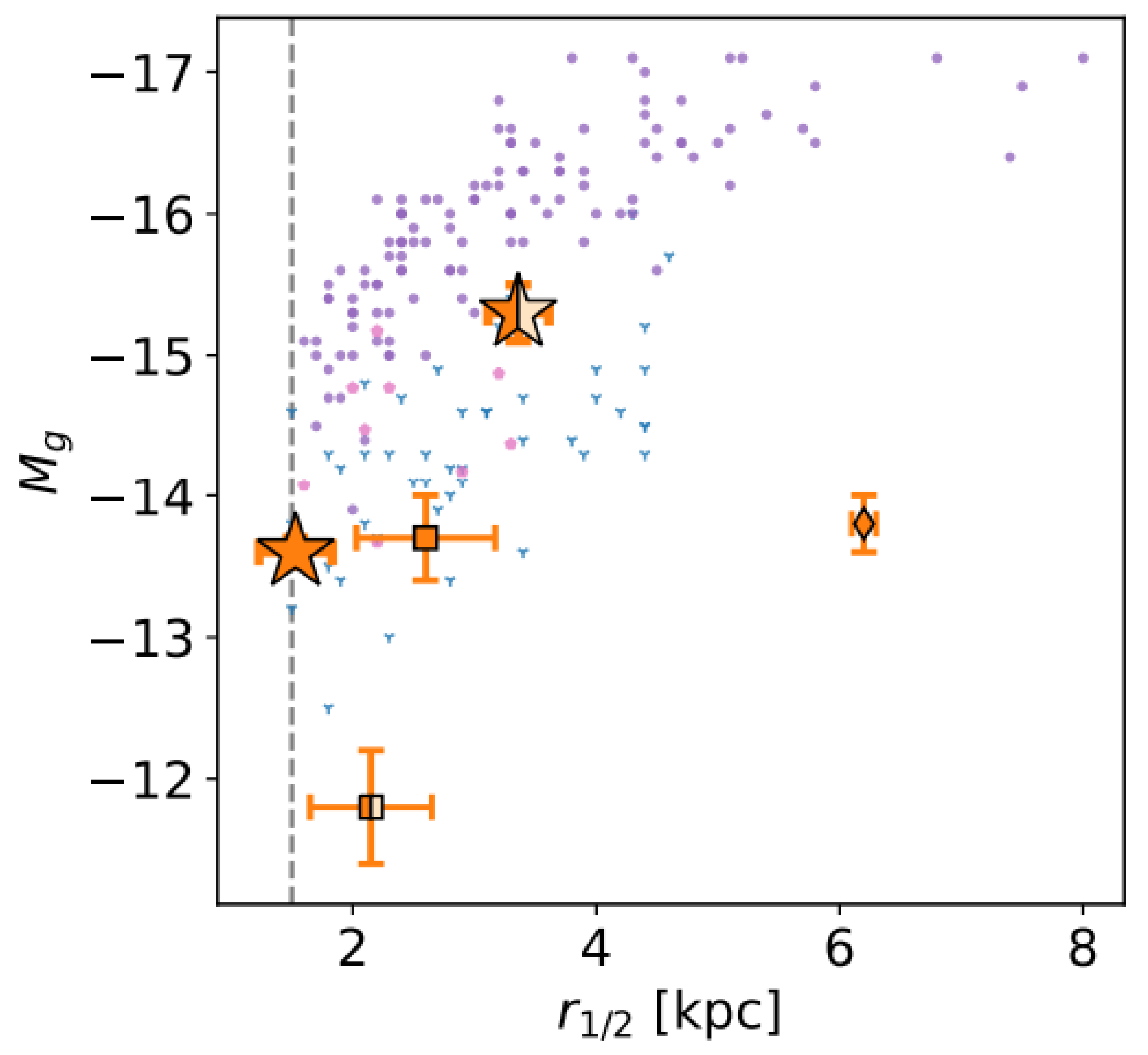
what do these errors mean?

SFR (NUV; $M_\odot$ yr^{-1} $\times 10^{-3}$)	6.2 ± 2.2
SFR (FUV; $M_\odot$ yr^{-1} $\times 10^{-6}$)	78 ± 41

we cannot really know unless we are told!

plot and results from Fielder et al. 2024

what do error bars mean?



what do these errors mean?

SFR (NUV; $M_\odot$ yr^{-1} $\times 10^{-3}$)	6.2 ± 2.2
SFR (FUV; $M_\odot$ yr^{-1} $\times 10^{-6}$)	78 ± 41

we cannot really know unless we are told!

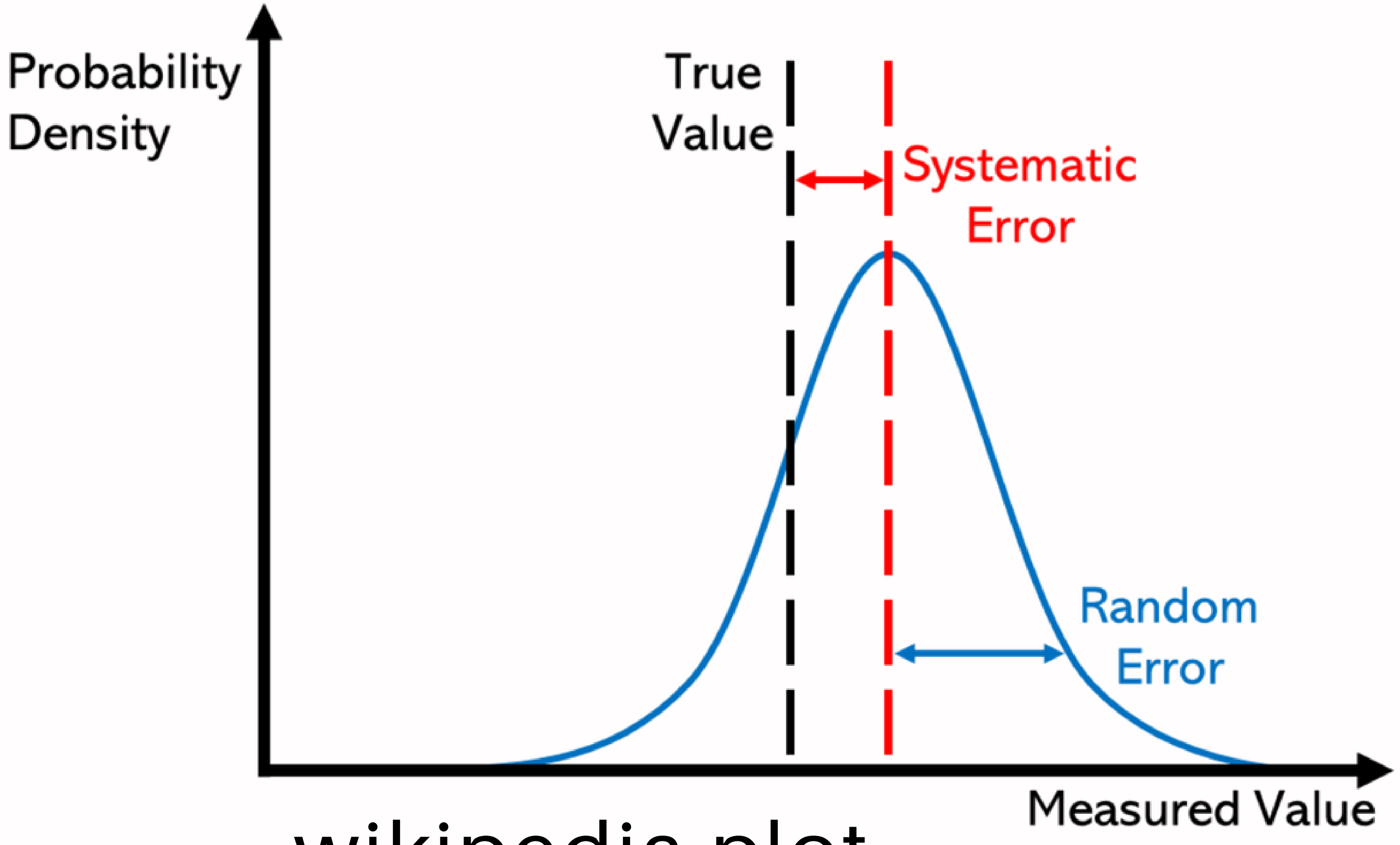
(we can imagine in most cases because astronomy works based on everyone agreeing on unspoken rules)

plot and results from Fielder et al. 2024

basic concept

real data measurements are not perfect for many reasons
(*read-out noise, sky, rounding, human imperfection, etc*)

there's random scatter (*will impact the standard deviation*)
and systematic scatter (*will impact the mean*)



wikipedia plot

“error on the parameter" tends to mean "the spread in the parameter values that are consistent with the data”

“measurement errors” tends to mean “the spread in the measurements you get repeating the experiment N times”

standard deviation as error

the standard deviation is often reported as the statistical uncertainty of a measurement.

if we have n observations of a random variable *x*, and a hypothesis for what the distribution function of f(*x*,m) looks like, where m is an unknown parameter, we can construct the likelihood m(x1,x2,...,xn). Using statistical methods (analytical ones, Monte Carlo, etc), we can estimate the standard deviation of this function.

if the function is a gaussian, this standard deviation is referred to as the error.

in an ideal world, we would repeat an identical measurement enough times to estimate the error distribution

this is hard to do in most areas but impossible in astronomy!

HOWEVER!!

if we understand the physics behind what we are measuring or what we are modelling then we can assume a distribution!

FOR EXAMPLE

counting photons is a Poisson process

why?

Poisson statistics are applicable when counting independent, random events which occur, when measured over a long period of time, at a constant rate.

this is cool because it means we can assume Poisson distributions not only for measurements but also for many of the noise factors since they also emit photons (e.g. dark current, sky)

with high exposition times, the mean becomes large enough that we can use a Gaussian distribution instead which is easier because it's symmetric.

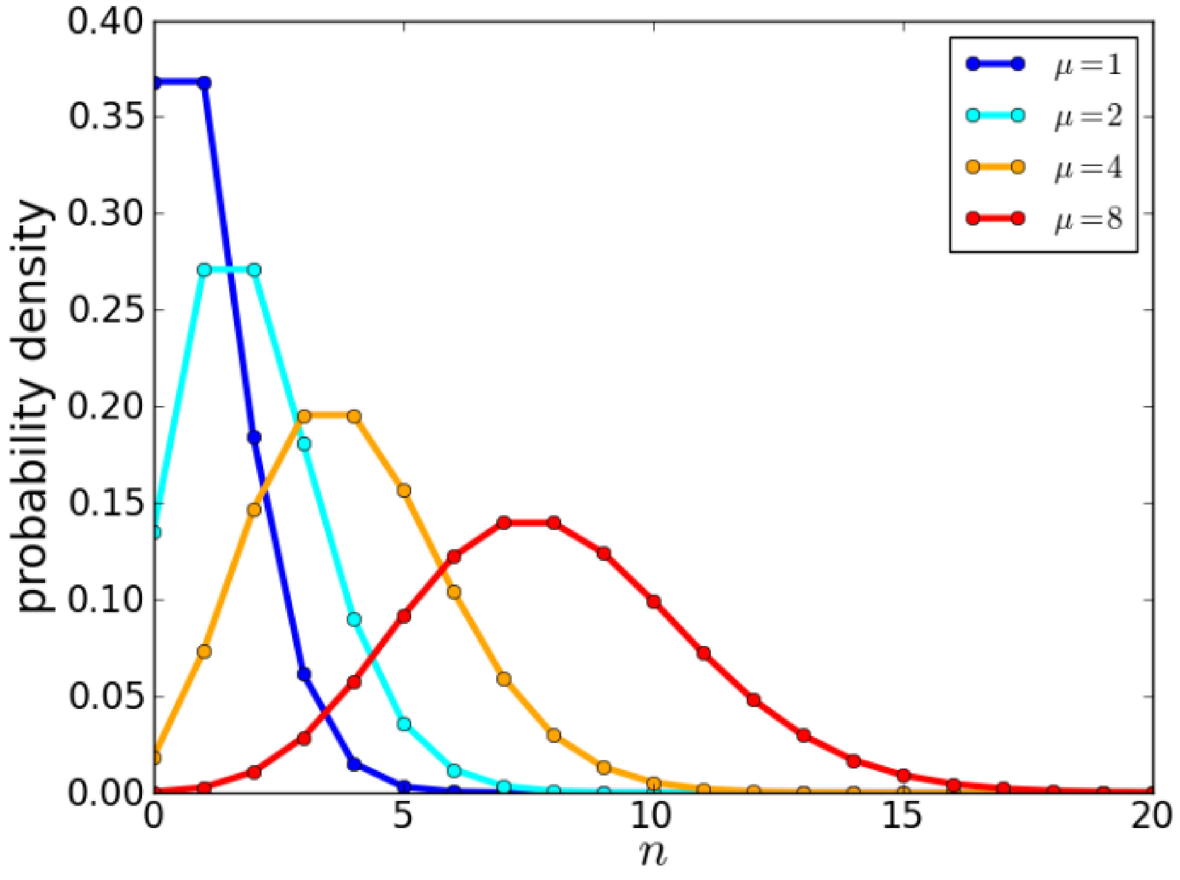


Figure 1: Examples of Poisson distributions with $\mu = 1$, $\mu = 2$, $\mu = 4$, and $\mu = 8$.

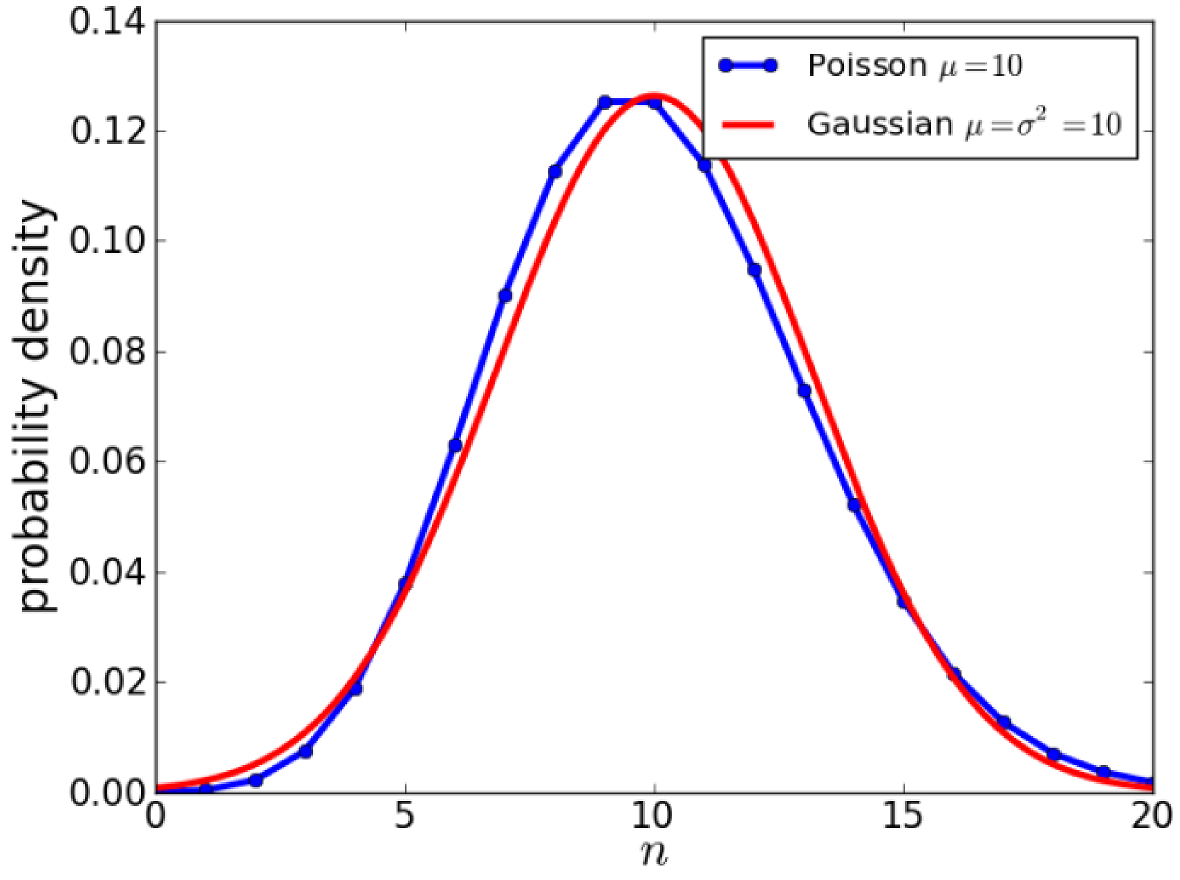
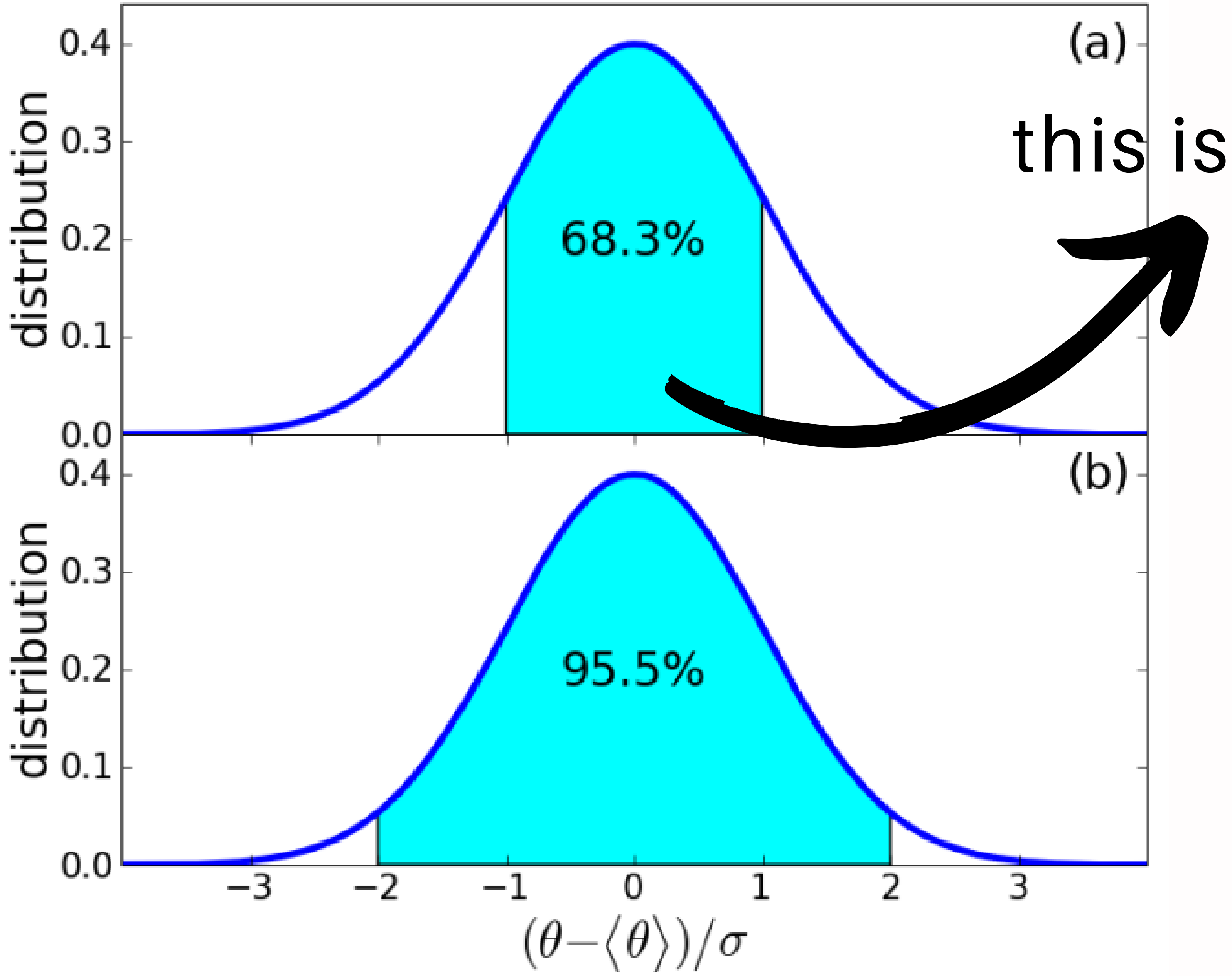


Figure 2: Poisson distribution with $\mu = 10$ and Gaussian distribution with $\sigma^2 = \mu = 10$. The Gaussian is a very good approximation to the Poisson distribution.

ANDRAE 2010


in surveys counts of objects
also behave in a poisson way
(are they though)

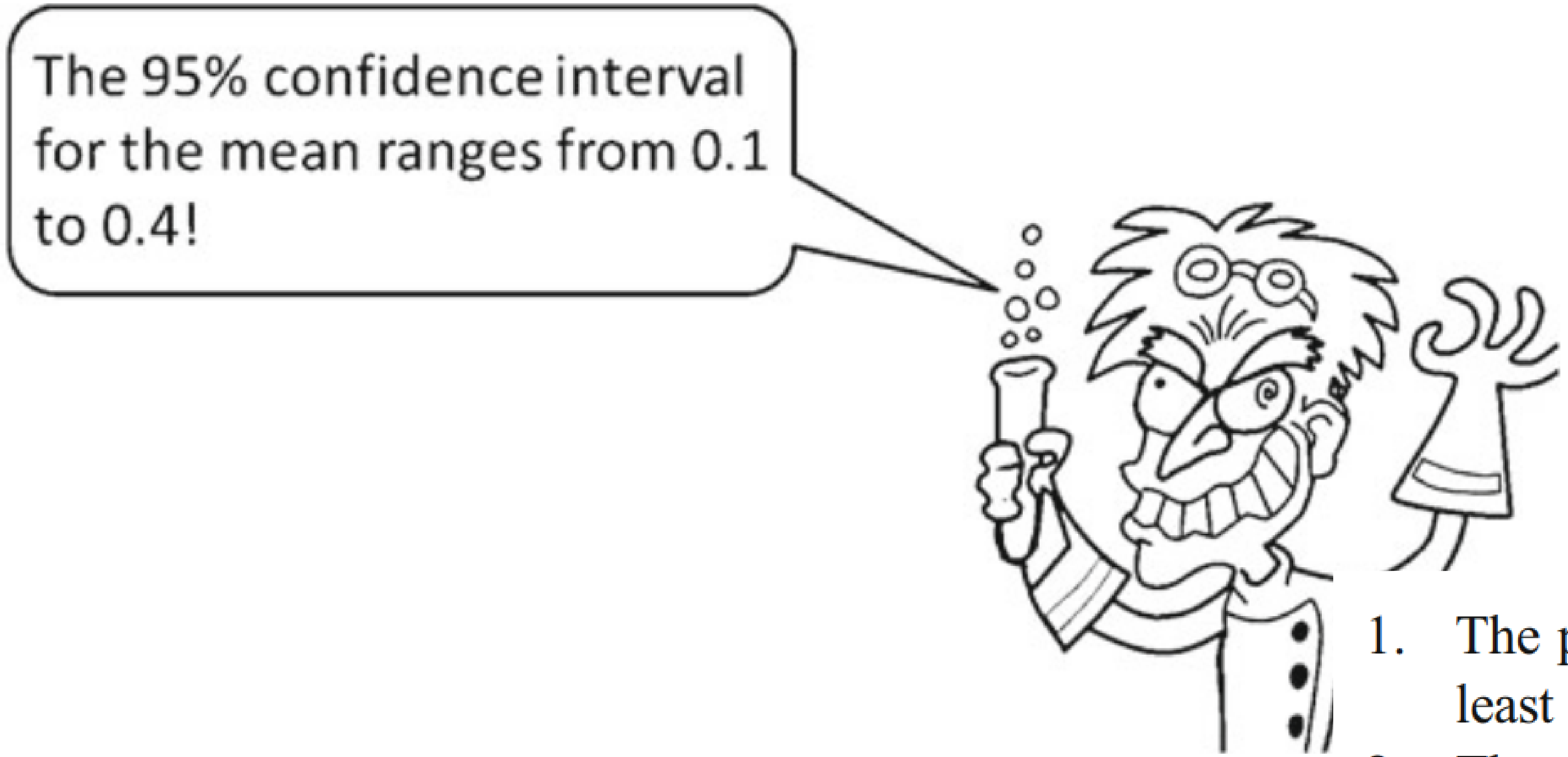
errors usually represent |1sigma| or |2sigma| uncertainty around a data point



this is a confidence interval

if we draw N values from this distribution , 68.3% will be within 1 sigma of the mean, and 95.5% will be within 2 sigma of the mean.

Professor Bumbledorf conducts an experiment, analyzes the data, and reports:



1. The probability that the true mean is greater than 0 is at least 95 %.
2. The probability that the true mean equals 0 is smaller than 5 %.
3. The "null hypothesis" that the true mean equals 0 is likely to be incorrect.
4. There is a 95 % probability that the true mean lies between 0.1 and 0.4.
5. We can be 95 % confident that the true mean lies between 0.1 and 0.4.
6. If we were to repeat the experiment over and over, then 95 % of the time the true mean falls between 0.1 and 0.4.

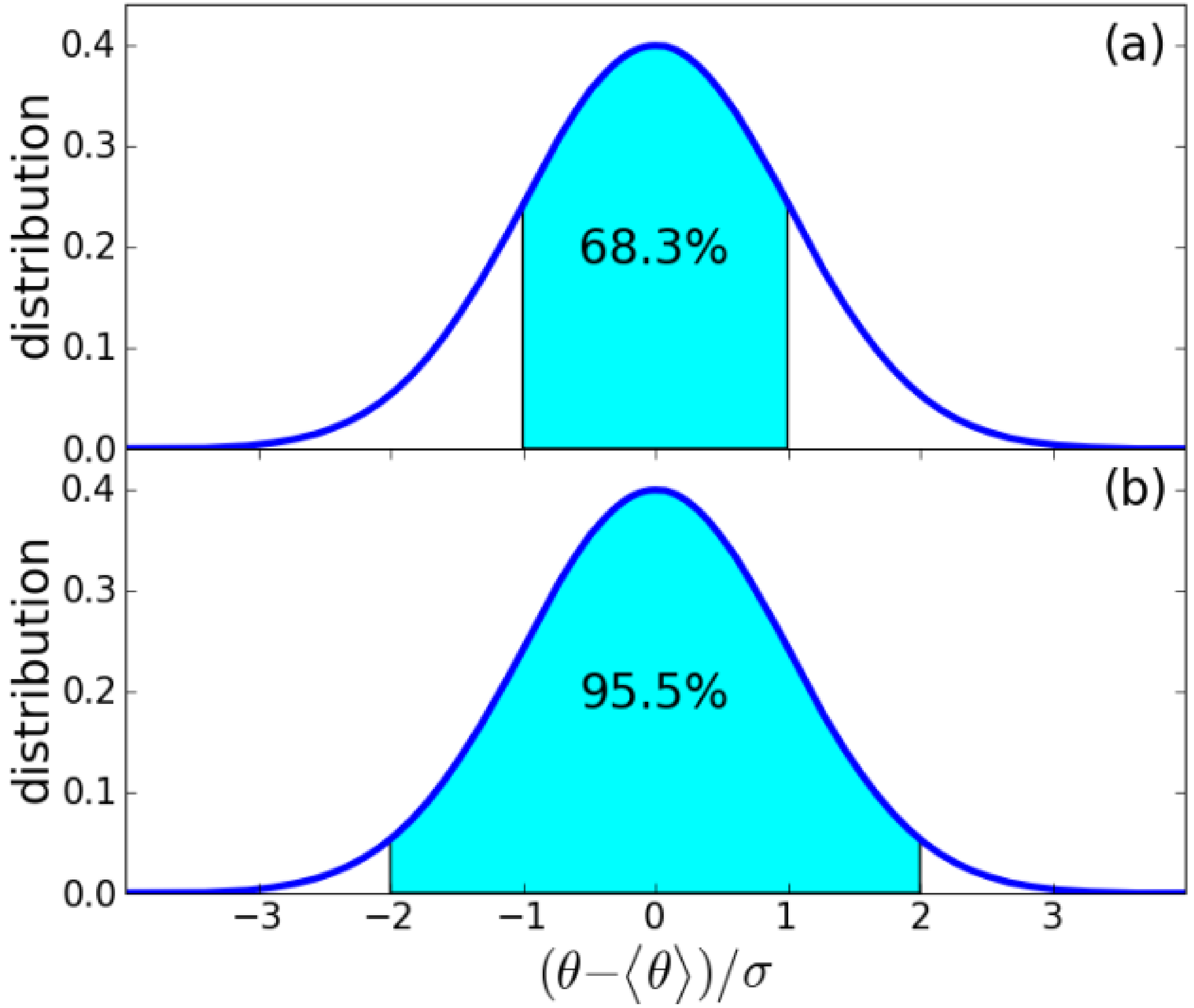
From Hoesktra et al. 2014

“If we were to repeat the experiment over and over, then 95 % of the time the confidence intervals contain the true mean.”

But what does that really mean? Some people will say it means if you did the experiment 100 times, the average (mean) you got would be within that 2.2–6.9 range in about 95 of the trials. Or that the true average for the whole population would fall within that range with 95 percent probability. Wrong wrong wrong. Think about it. Suppose you did the experiment a second time, and got an average of 5.9 pounds with a confidence interval of 4.1 to 8.8 pounds. Would you be 95 percent confident in both of the reported ranges? And you'd get different confidence intervals every time you did the experiment. How can they all give you the right range of 95 percent confidence?

In actual statistical fact, a confidence interval tells you ***not how confident to be in the answer, but how confident to be in your sampling.*** *In other words,* ***if you repeated the experiment*** *(****on different samples from your population****)* ***a gazillion times, your confidence interval will reliably contain the true value in 95 percent of the trials****. That merely tells you how often your confidence range will be valid over the course of many repetitions of the experiment. As statistician and political scientist Andrew Gelman expresses it, "****Under repeated sampling, there is a 95 percent probability that the true mean lies between the lower and upper bounds of the interval.****"*

"Scientists' grasp of confidence intervals doesn't inspire confidence" by Tom Siegfried, Science News errors usually represent |1sigma| or |2sigma| uncertainty around a data point



if we draw N values from this distribution , 68.3% will be within 1 sigma of the mean, and 95.5% will be within 2 sigma of the mean.

$$Pr(\mu - n\sigma \leq X \leq \mu + n\sigma) = \int_{\mu-n\sigma}^{\mu+n\sigma} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

variable change $u = \dfrac{x-\mu}{\sigma}$

$$\frac{1}{\sqrt{2\pi}} \int_{-n}^{n} e^{-\frac{u^2}{2}} du$$

this integral is independent of everything

$$\Pr(\mu - 1\sigma \leq X \leq \mu + 1\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{1} e^{-\frac{u^2}{2}} du \approx 0.6827$$

$$\Pr(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-2}^{2} e^{-\frac{u^2}{2}} du \approx 0.9545$$

$$\Pr(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = \frac{1}{\sqrt{2\pi}} \int_{-3}^{3} e^{-\frac{u^2}{2}} du \approx 0.9973.$$

this is all very frequentist
responding to the concept that we can just
run the same experiment N times

"*In an ideal Bayesian universe, error bars don't exist*"

Scott Oser in the slides for his Stats course

PRIOR FULL PDF → **POSTERIOR PDF**

this contains all the information we are trying to sum up in an error bar

an error bar is the result of assuming a shape for the PDF and then reporting its mean and standard deviation.

life isn’t as gaussian as we think it is
gaussian tails get small far too quickly

the central limit theorem tells us gaussians are usually kinda good but also

$$P(\epsilon_i) = \frac{1}{\sqrt{2\pi}} e^{-\epsilon_i^2 / 2\sigma_i^2}$$

$$\mathcal{L}(\epsilon_1, \epsilon_2, \ldots, \epsilon_N) = \left(\frac{1}{\sqrt{2\pi}}\right)^N \cdot \frac{1}{\sigma_1 \cdot \sigma_2 \cdot \ldots \sigma_N} \cdot e^{-\frac{1}{2}} \sum_{i=1}^{N} \frac{\epsilon_i^2}{\sigma_i^2}$$

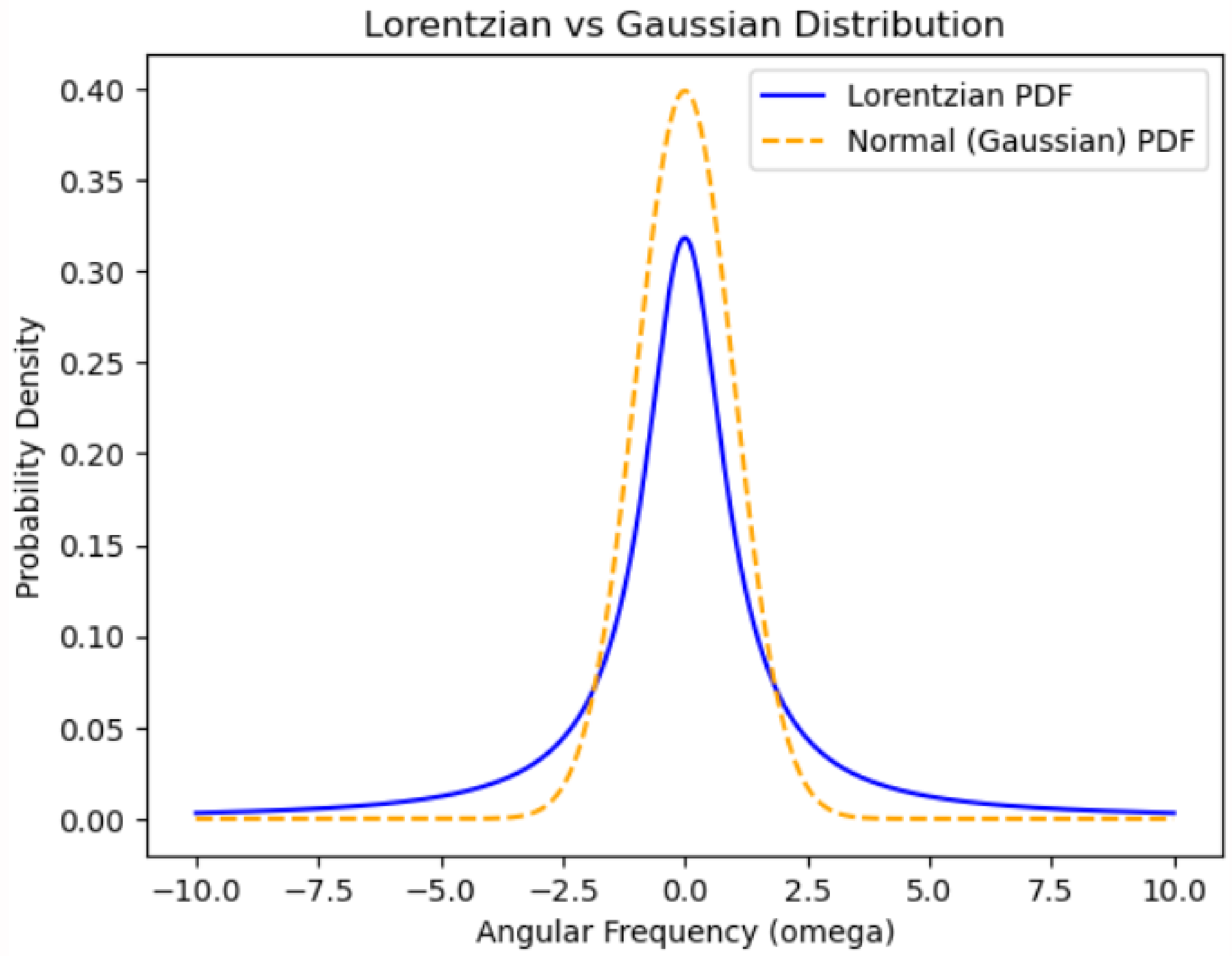
this is a very easy function to work with
if we want to differentiate it
it’s just easy

on the other hand

this looks like an innocent function

when we model things that have a high probability of extreme values we need functions with heavier tails like this one

this models things like resonances



"If one stands in front of a line and kicks a ball with a direction (more precisely, an angle) uniformly at random towards the line, then the distribution of the point where the ball hits the line is a Cauchy distribution." - Wikipedia example

and yet

$$p_x = \frac{1}{\pi(1+x^2)}$$

the characteristic function of this is

$$\mathbb{E}[e^{i\omega X}] = \int_{-\infty}^{\infty} \frac{e^{i\omega x}}{\pi(1+x^2)} dx = e^{-|\omega|}$$

the derivative exists
but it's not differentiable
at w=0

$$\frac{1}{i}\frac{d}{d\omega}e^{-|\omega|}$$

$$\int_{-\infty}^{\infty} \frac{x}{\pi(1+x^2)} \mathrm{d}x$$

or, if we try to calculate the mean, we get
an improper integral

this is cool because it means we can assume Poisson distributions not only for measurements but also for many of the noise factors since they also emit photons (e.g. dark current, sky)

with high exposition times, the mean becomes large enough that we can use a Gaussian distribution instead which is easier because it’s symmetric.

remember this?

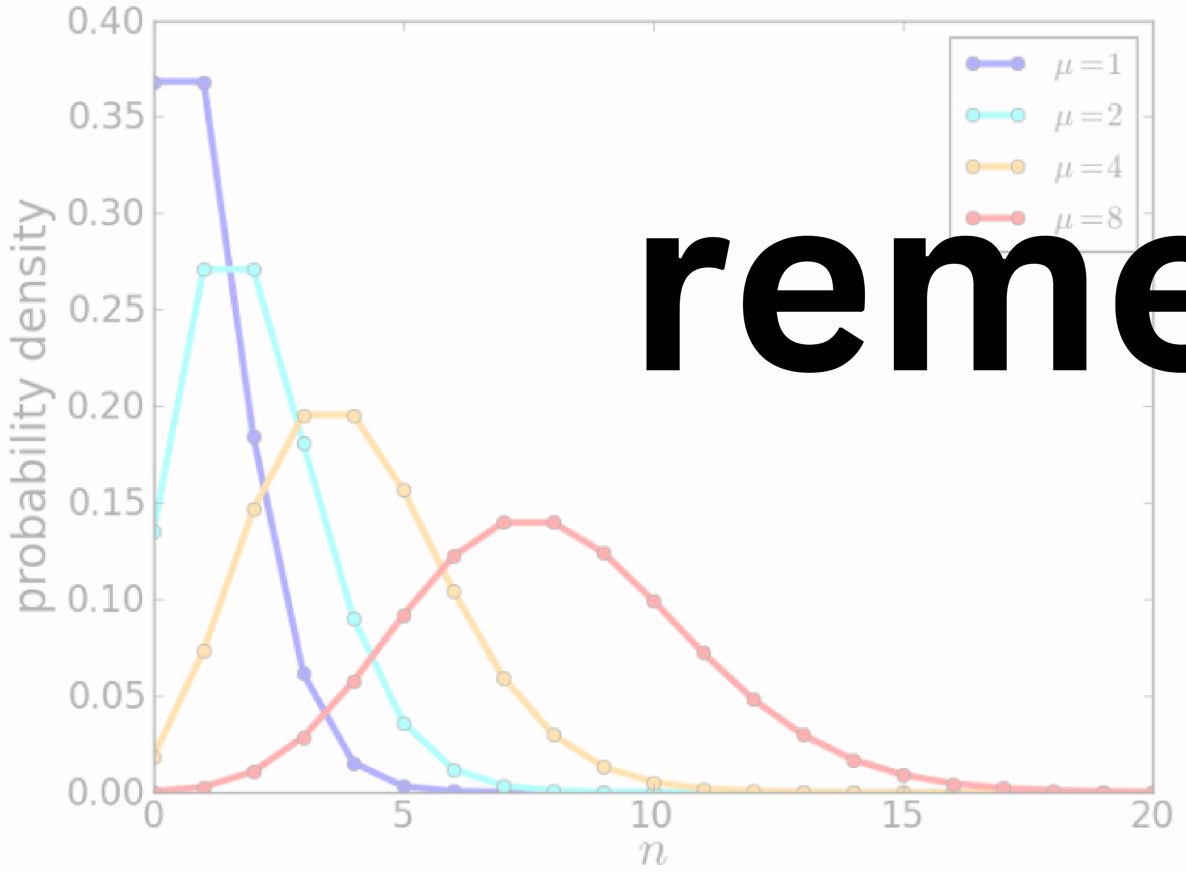


Figure 1: Examples of Poisson distributions with $\mu = 1$, $\mu = 2$, $\mu = 4$, and $\mu = 8$.

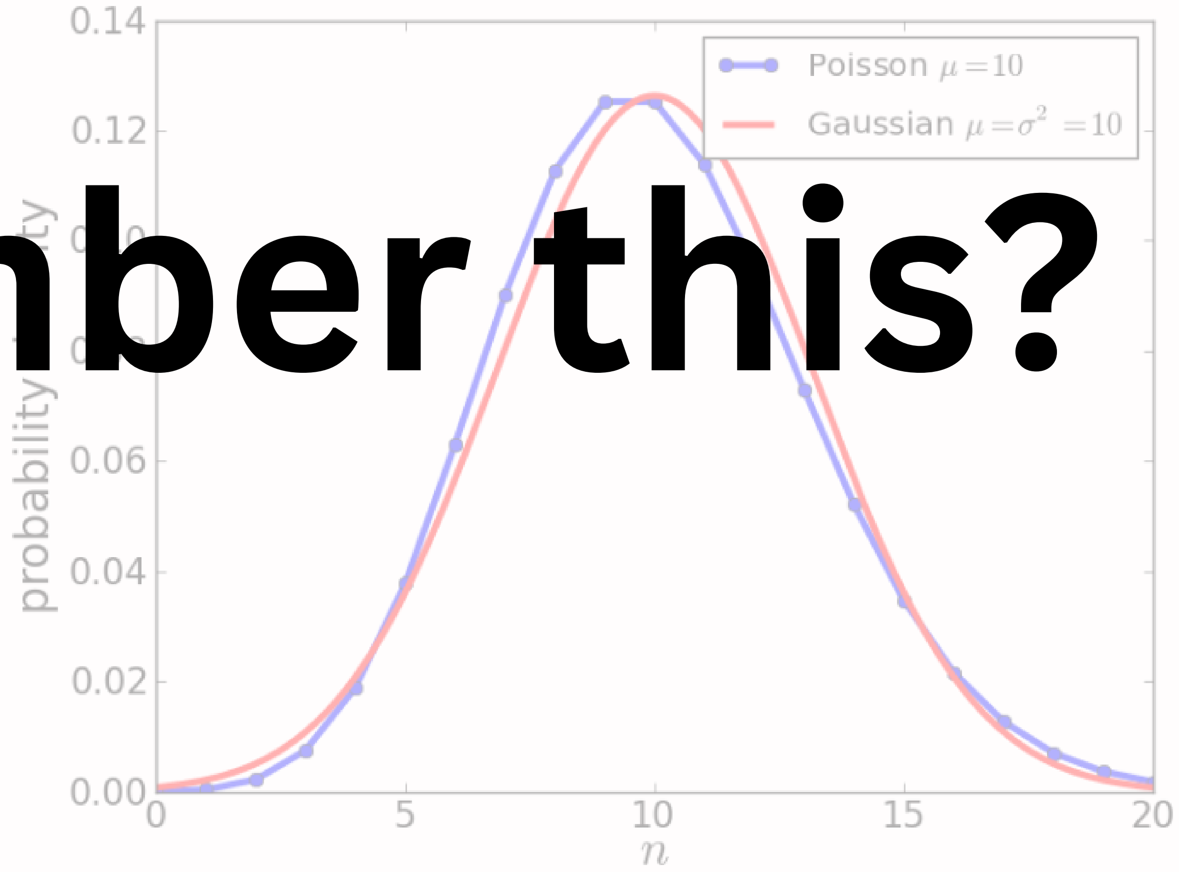


Figure 2: Poisson distribution with $\mu = 10$ and Gaussian distribution with $\sigma^2 = \mu = 10$. The Gaussian is a very good approximation to the Poisson distribution.

ANDRAE 2010

what if there’s a flock of birds passing by
what if there’s a plane
what if there’s a cloud
what if the telescope isn’t tracking perfectly
what if there’s a cosmic ray
what if a chip is slightly warmer than the others
what if your aperture isn’t perfect
what if the operator fell asleep for just a second

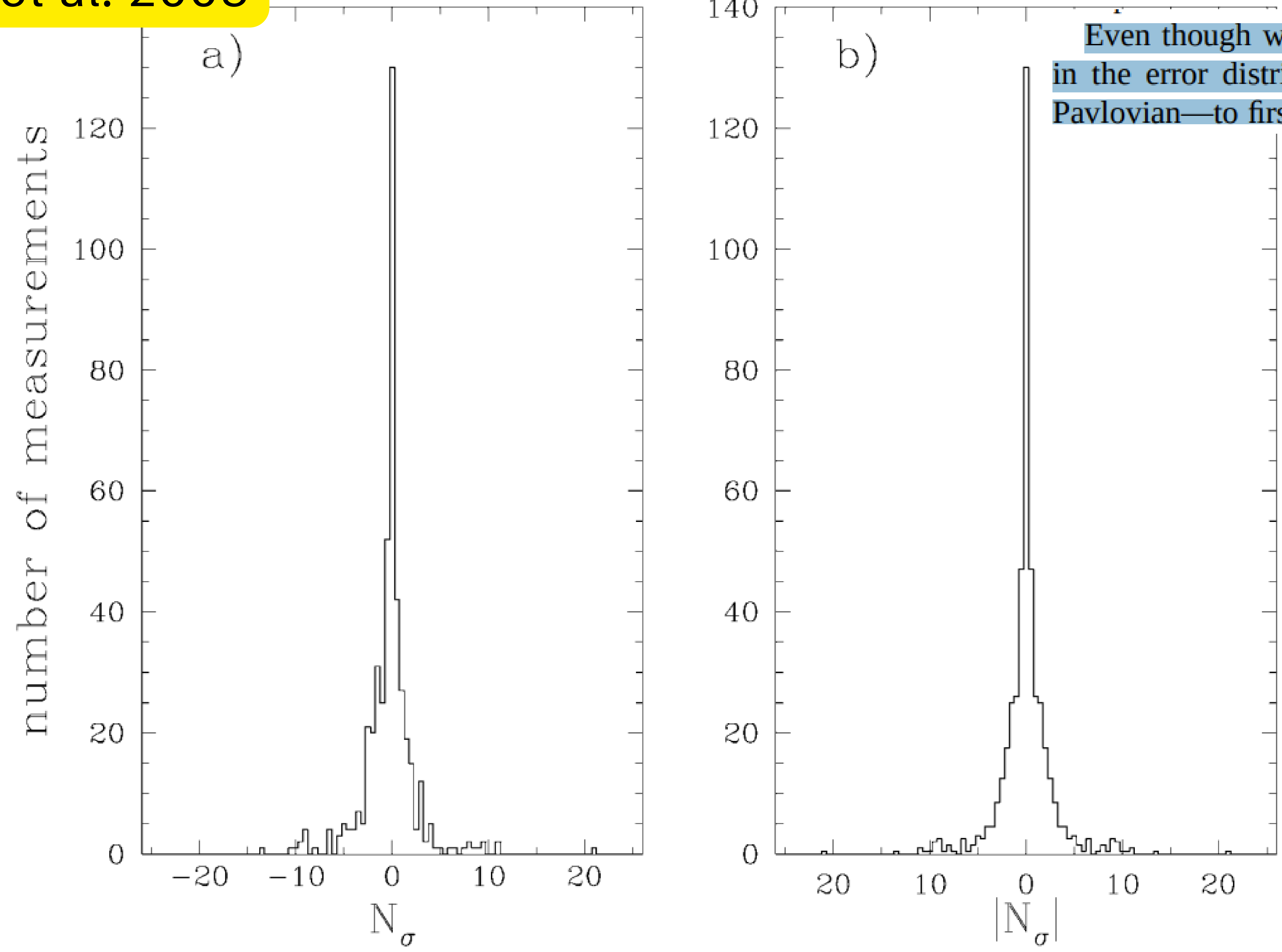
this is not just not gaussian
it’s not even symmetric!

what if we are not doing photometry

what if we are not doing photometry

Chen et al. 2003

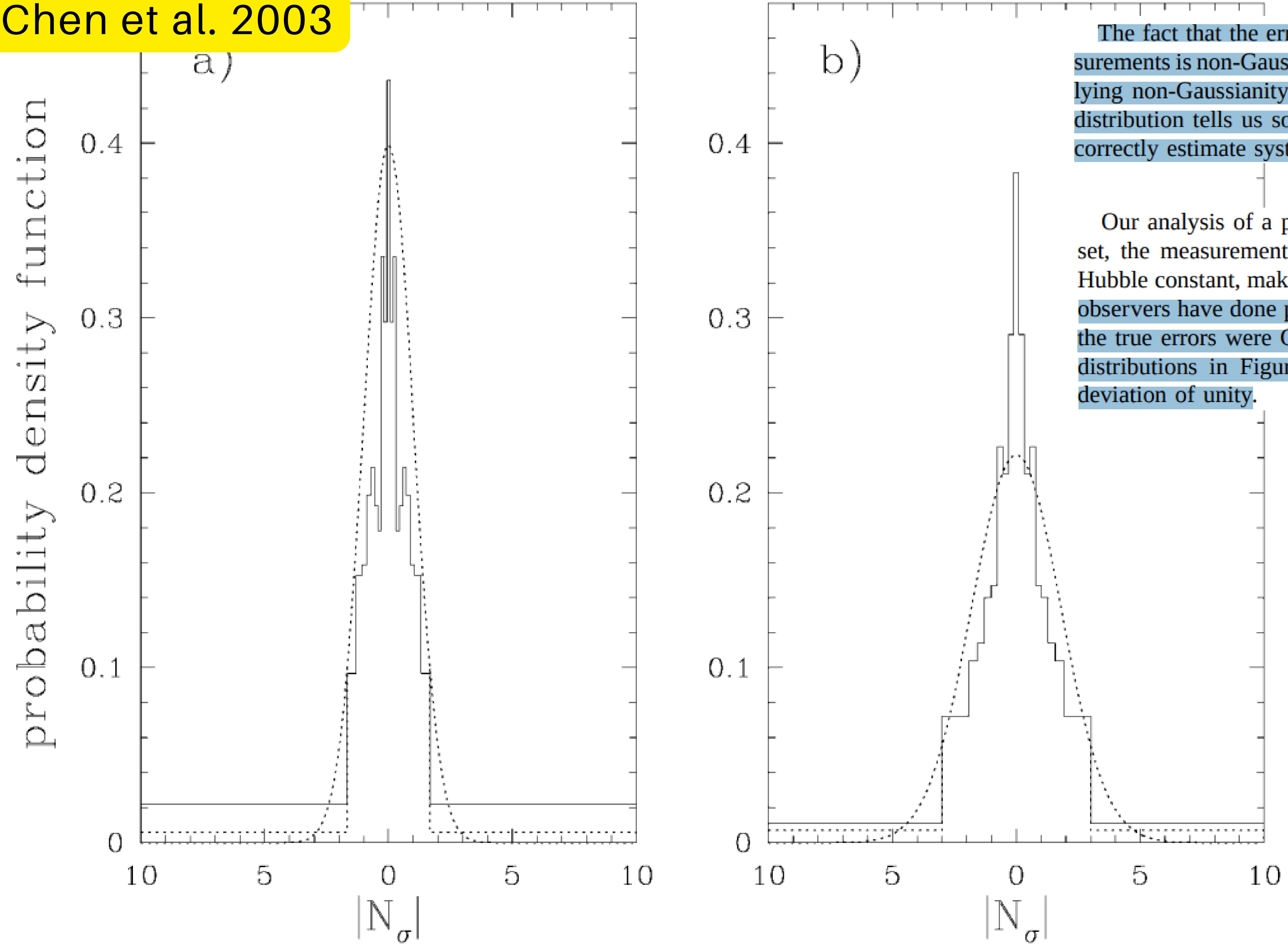
Chen et al. 2003



Even though we have noted the existence of extended tails in the error distributions of Figure 1, it is natural—perhaps Pavlovian—to first consider the Gaussian distribution, initially

FIG. 1.—Number of measurements (in half standard deviation bins) away from the central value of $H_0 = 71$ km s^{-1} Mpc^{-1} estimated by the *WMAP* Collaboration. (*a*) Sign of the deviation; (*b*) only the magnitude of the deviation. In panel (*a*), bins with positive (negative) N_σ correspond to measurements where H_0 is measured to be higher (lower) than 71 km s^{-1} Mpc^{-1}.

Chen et al. 2003



The fact that the error distribution of Hubble constant measurements is non-Gaussian does not necessarily imply an underlying non-Gaussianity in the measurement errors. Rather, the distribution tells us something about the observers' ability to correctly estimate systematic and statistical uncertainties.

Our analysis of a perhaps unique (because of its size) data set, the measurement errors of all available estimates of the Hubble constant, makes for some interesting conclusions. If all observers have done perfect jobs at estimating their errors and the true errors were Gaussian, as might be expected, then the distributions in Figure 1 should be Gaussian with standard deviation of unity.

Chen et al. 2003

deviation of their result. Overoptimism would produce error bars that were too small, while overconservatism would produce error bars that were too large. Which occurs in practice? The results here suggest that astronomers were overoptimistic by almost a factor of 2. Why? In some case there were systematic errors of which the observers were simply unaware (such as mistaking H II regions for bright stars). In other cases, standard candles were not as standard as imagined, leaving some steps in the distance ladder wrong by more than people thought. Also, using self-consistency in the data as a check on the errors can lead to large tails because it occasionally induces one to be overoptimistic (the Student *t* effect). And the real data may have non-Gaussian tails (say, in the luminosity of standard candles). In general, overconservatism (the urge to be right) always competes with overoptimism (the urge to have more interesting limits). In the case of the Hubble constant, astronomers were overoptimistic. In a history of science context, it might be of interest to more closely examine the most systematics

frequentist approach struggles with systematics

if systematics fluctuate then it’s fine because
we can still think of experiments as “random”
but some systematic uncertainties are constant

example

measuring a temperature

systematic fluctuating uncertainty: the temperature of the room in which you conduct the experiment

systematic fixed uncertainty: the thermometer is off by a degree *astronomical example*

measuring a temperature

systematic fluctuating uncertainty: the atmosphere

systematic fixed uncertainty: calibration error in bayesianland, we can always just marginalize over uncertainties or include them in our prior

if we measure a quantity X, we have a prior $P(X|I)$, some observed data D, and a likelihood $P(D|X,I)$

if the likelihood depends on a systematic parameter m, we can treat X and m as unkown parameters, assign a prior to each, and then we get $P(X,m|D,I)$

since we don’t care about m, we can integrate over it and then get $P(X|I)$

to determine the value of our systematic or model the prior for it, we can either

1) take calibration measurements
2) fit it from the main data
3) use “theory” (aka what are measurement uncertainties in theory parameters, what is the spread between different theory estimates)
4) use monte carlo but we won’t discuss that here

BIBLIOGRAPHY

The techniques of least squares and stellar photometry with CCD by Stetson

Error estimation in Astronomy: a guide (Andrae 2010)

All Puffed Up: Tidal Heating as an Ultra Diffuse Galaxy Formation Pathway

Scott Oser's lectures (10 and 11)

This website

Hoekstra et al.

Statistical Data Analysis by Cowan

Non-Gaussian Error Distribution of Hubble Constant Measurements (Chen et al 2003)