

PHY 517 / AST 443: Observational Techniques in Astronomy

Lecture 4:
Statistics, part I

Lab 2

- Request your transit dates ASAP!

Lab 1

- remember to do your weekly check-ins for Lab 1 with me or the TAs

General Information

Class time and place: MoWe 2:40 - 5:30pm, ESS 450

Credits: 3 (PHY 517) or 4 (AST 443)

Instructor: Anja von der Linden (anja.vonderlinden 'at' stony brook.edu, ESS 453)

Office hours: Wed 9:30-10:30am

TAs:

- Aaron Mueninghoff (aaron.mueninghoff 'at' stonybrook.edu), office hours Mo 12-1pm and Th 12-1pm, ESS 457A or by appointment
- Ben Levine (benjamin.c.levine 'at' stonybrook.edu), office hours Tu 12-1pm and Fr 1-2pm, ESS 457A or by appointment

STONY BROOK



8PM THURSDAYS
ESS 450

ALL MAJORS!
NO PRIOR KNOWLEDGE REQUIRED!
ROOFTOP OBSERVING!

MAILING LIST



DISCORD





Inaugural Colloquium of
the C.N. Yang Colloquium Series

Probing the Universe with Gravitational Waves

Dr. Barry Barish
Nobel Laureate 2017

September 6, 2022
Coffee/Tea: 4:00pm, Simons Center Cafe
Lecture: 4:30pm, Della Pietra Family
Auditorium, SCGP Room 103
Reception immediately following the talk



Department of Physics and Astronomy C.N. Yang Institute for Theoretical Physics Simons Center for Geometry and Physics

(A brief intro to)
Statistics

Statistics in Astronomy

- we are almost always working in the low signal-to-noise regime
- have to be very careful to make correct inferences from our data!
- robust (and advanced) statistical techniques play a very important role in astronomy

Measurements

- example: 99.123 ± 0.005
- what is 0.005 called?
 - (measurement) **uncertainty**
 - NOT “error” (inaccurate, though often used)
- what does this mean?
 - if we repeat the measurement many times, in 68% of the cases the true value would fall within the quoted uncertainty interval
 - not-quite-right interpretation: the quoted interval has a 68% chance of containing the true value

“Error”

- **error**: difference between *measured* and *true value*
- can be due to:
 - random fluctuations (statistical error)
 - instrumental / algorithmic limitations (systematic error)
 - mistakes (illegitimate error)
- measurements are meaningless if not accompanied by an estimate of the error
- but truth is unknown, have to estimate error indirectly

Accuracy vs. Precision

- **accuracy:** how close a measurement is to the truth
- **precision:** size of (statistical) measurement uncertainty

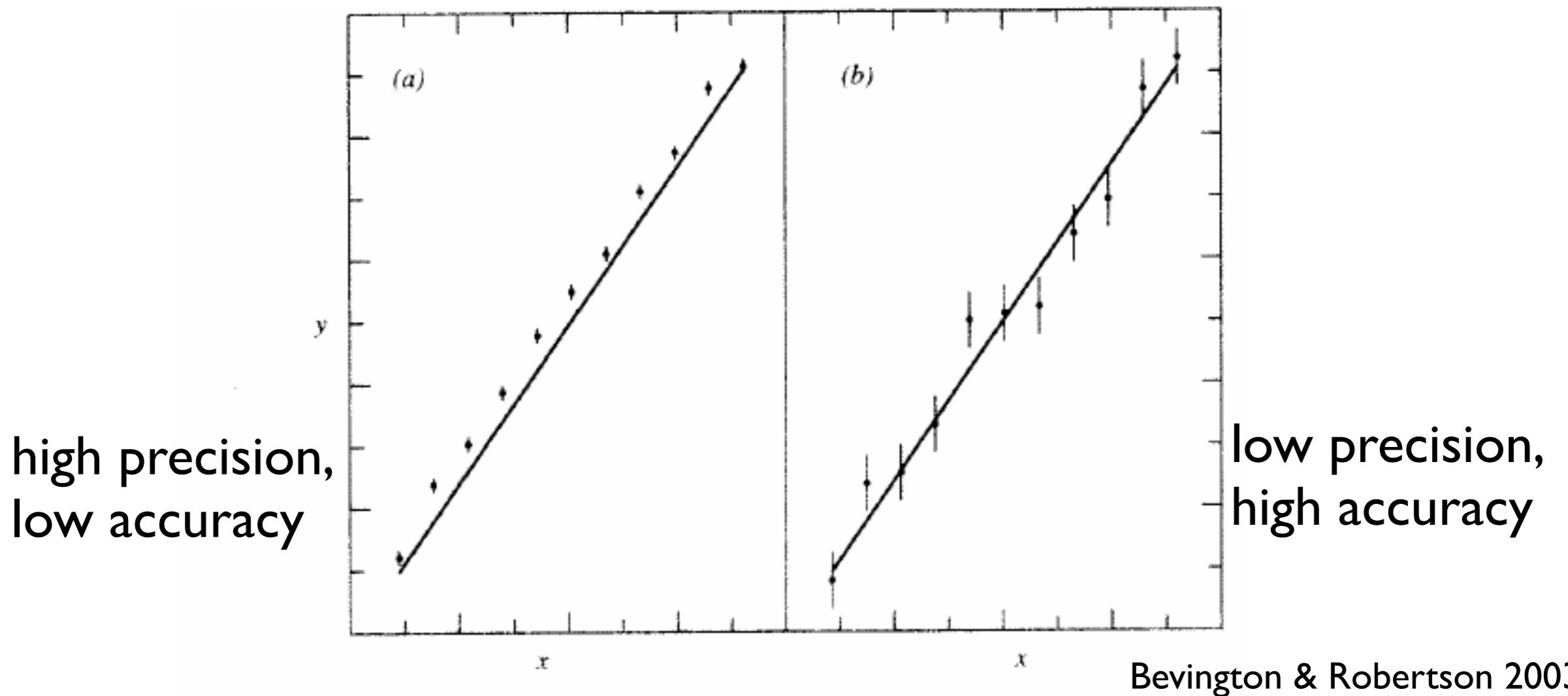
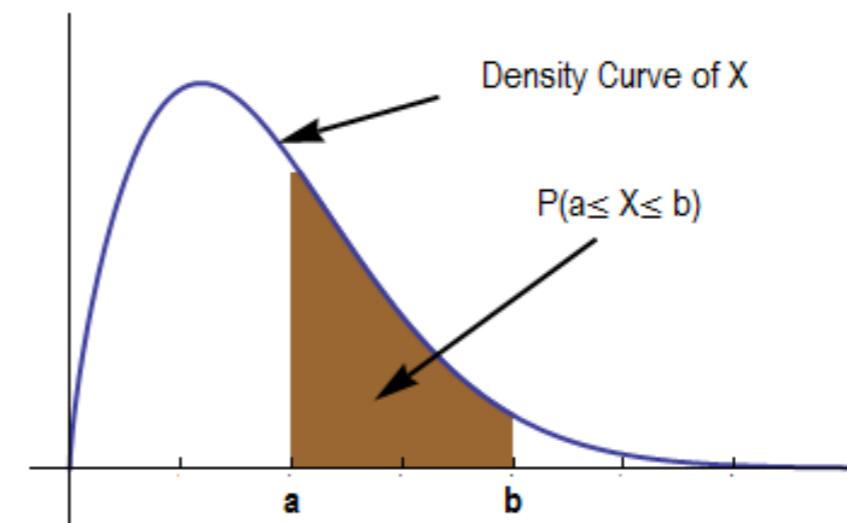
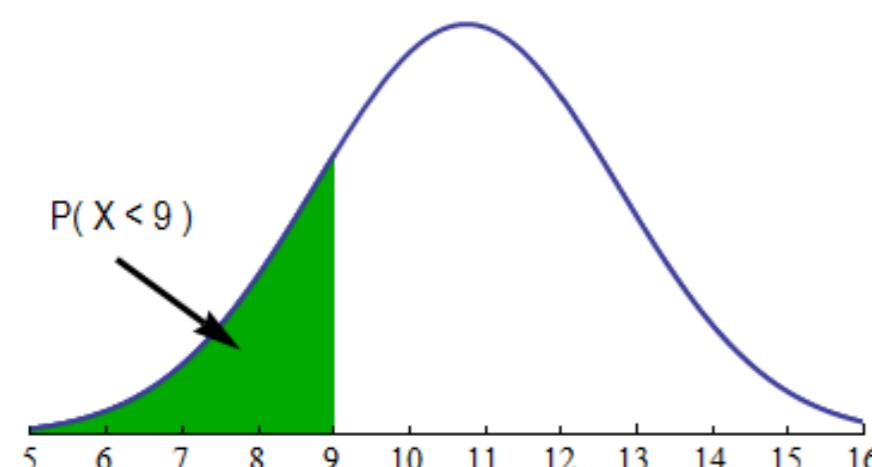
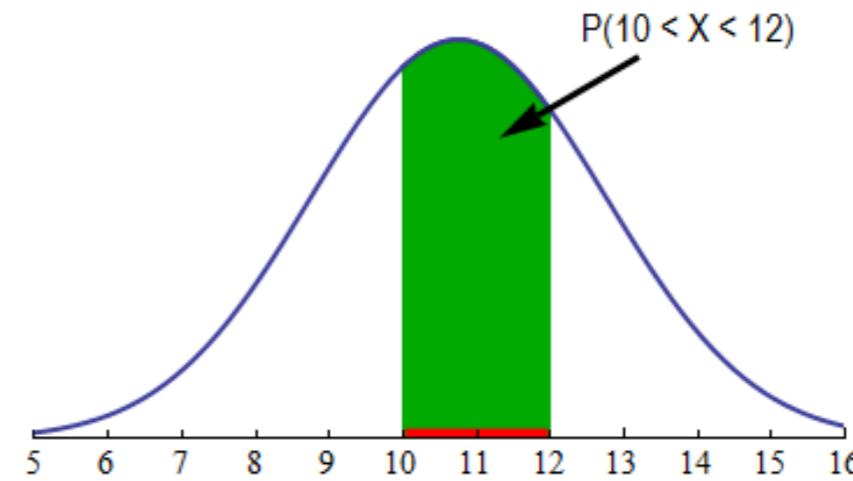
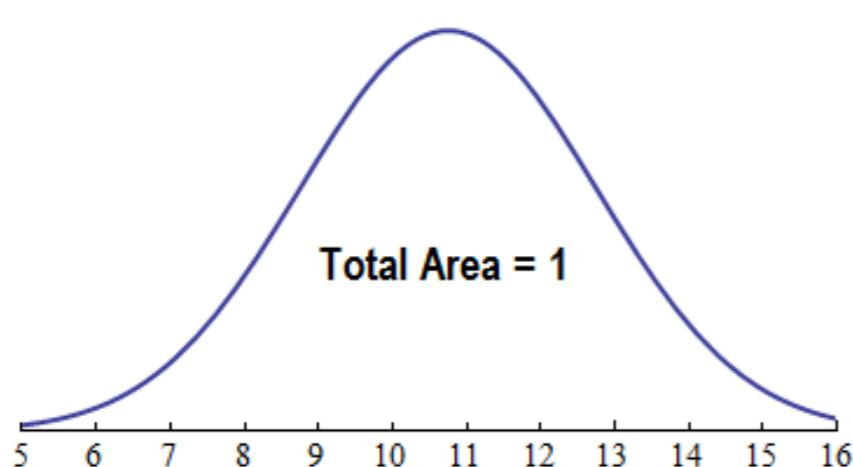


FIGURE 1.1

Illustration of the difference between precision and accuracy. (a) Precise but inaccurate data.
(b) Accurate but imprecise data. True values are represented by the straight lines.

Probability Distributions

- probability distributions: describe expected / measured distributions of measurements
- integrate over range of values to find probability to be in that range

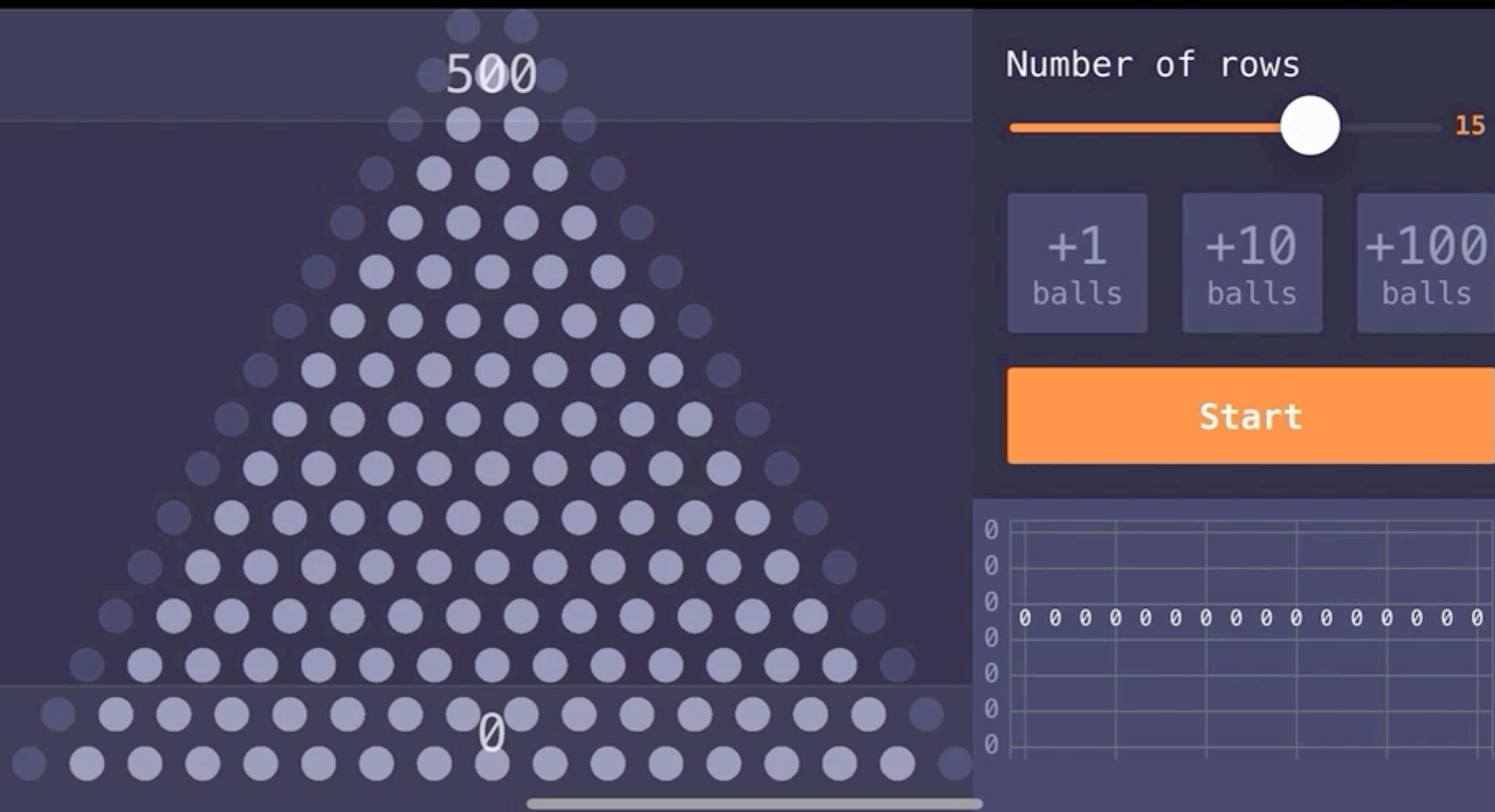


Sample vs. Parent Distribution

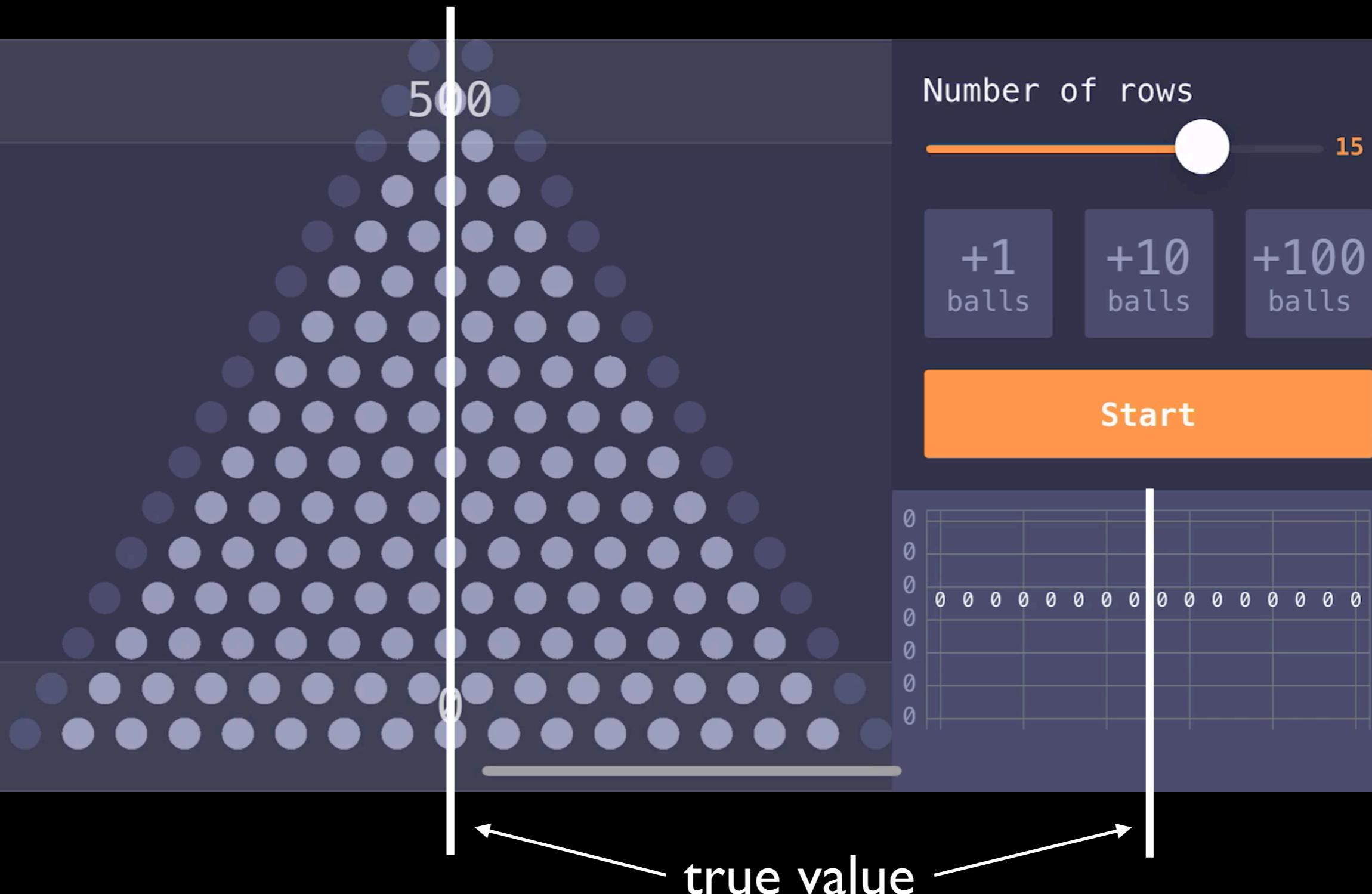
- measurement x_i of a quantity x :
 - approximates x
 - not necessarily equal to x because of statistical uncertainty
- many measurements x_i :
 - expected to be distributed about true value
 - **sample distribution**
- **parent distribution**:
 - probability of particular result from single measurement
 - idealized outcome of infinite number of measurements

the sample distribution *samples* the parent distribution

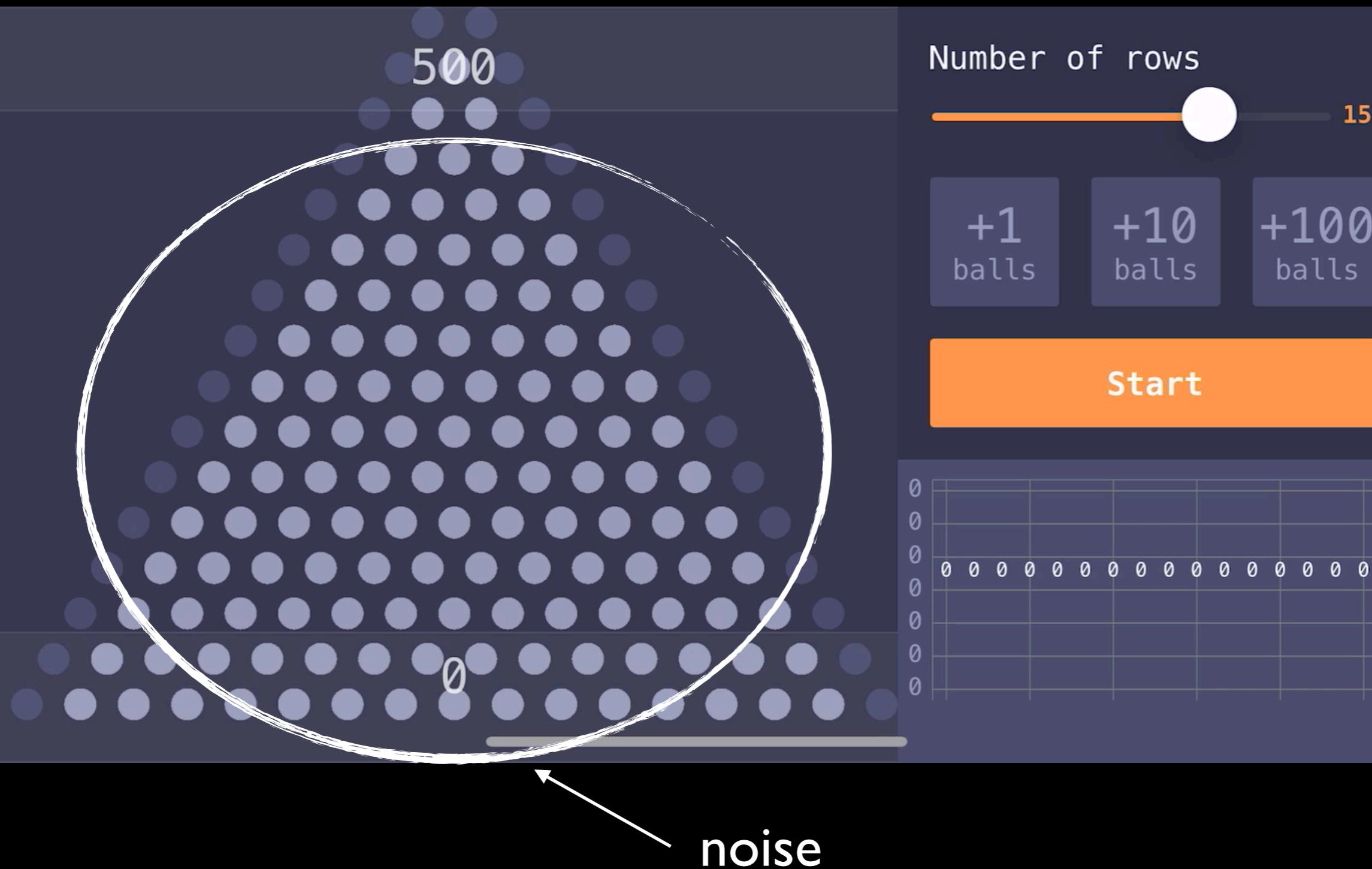
Galton Board:



Galton Board:

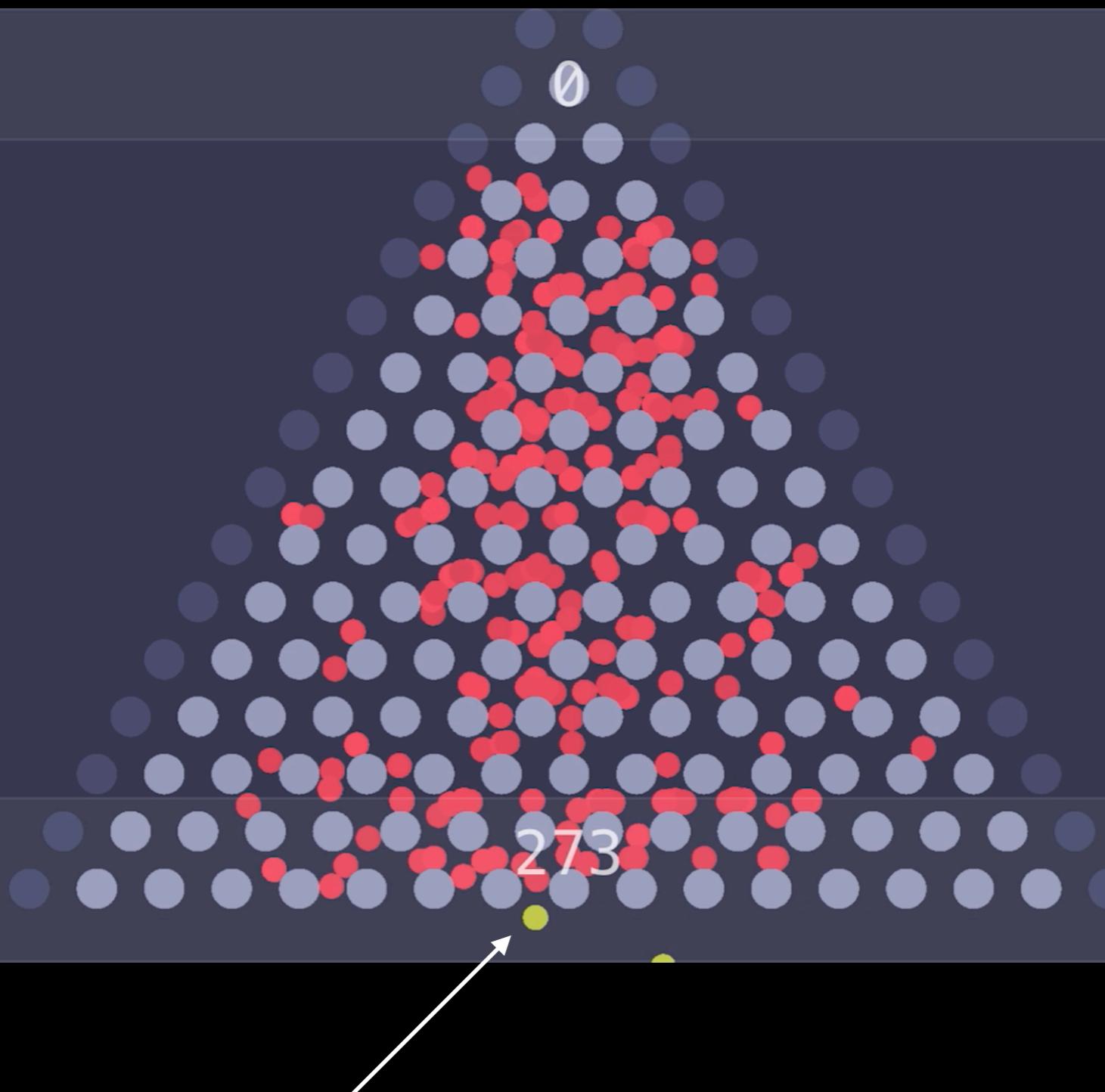


Galton Board:



Galton Board App for iOS, Edwin Veger

Galton Board:



one measurement

Number of rows

15

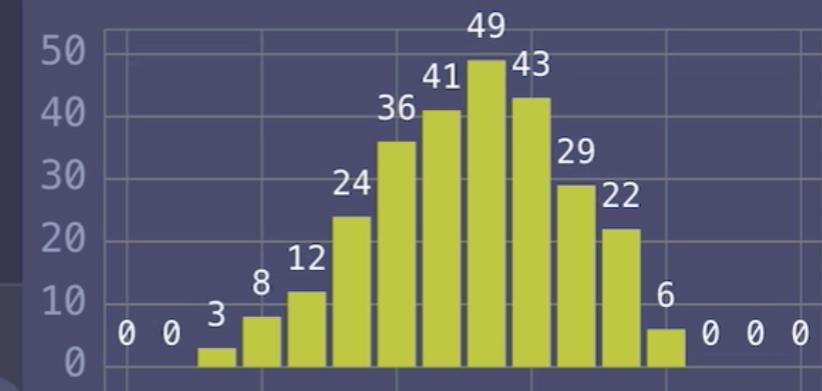
+1
balls

+10
balls

+100
balls

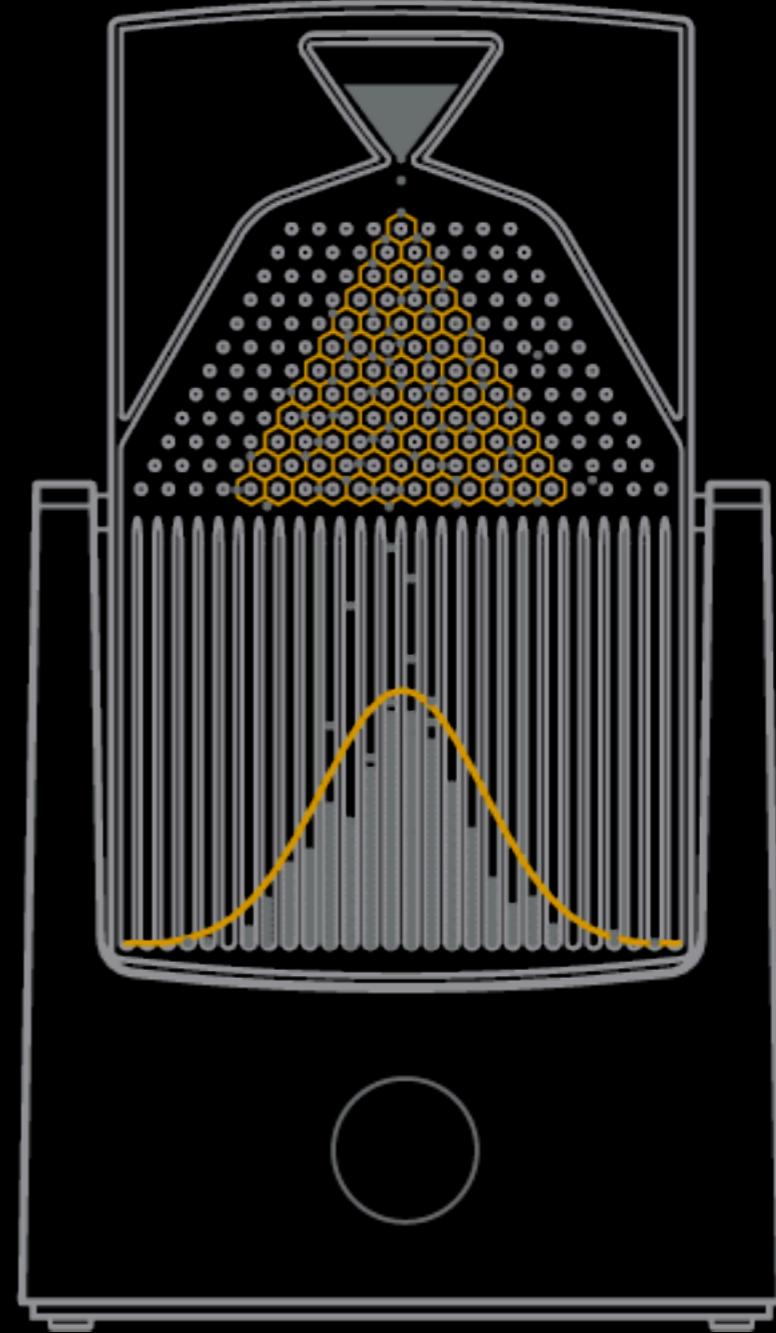
Pause

Reset



many measurements

Galton Board:



sample population: distribution of many measurements

parent population: Gaussian / normal distribution

Summary statistics

- only the full sample distribution is the full description of your data
- but usually, it is helpful to describe the sample distribution with a few numbers → summary statistics

Q: can you think of examples?

Mean, median, and mode

- (unweighted) **mean** of the sample distribution:

$$\bar{x} = \frac{1}{N} \sum_i x_i$$

- IF there are no systematic errors, the mean of the parent population is:

$$\mu = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i x_i$$

Mean, median, and mode

- **median:** 50th percentile of distribution (half the measurements are smaller, half are greater)

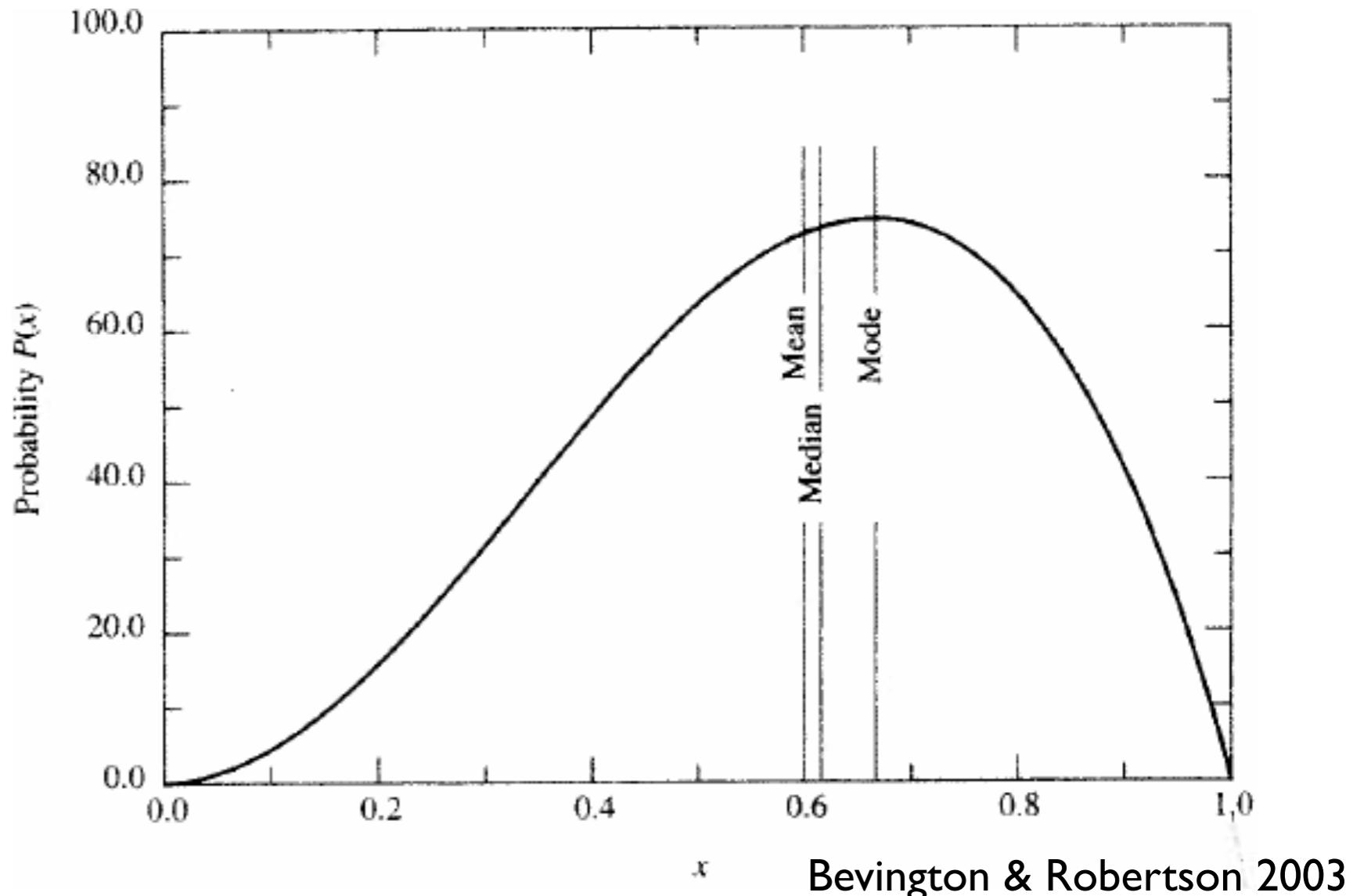
motivation: less susceptible to “outliers” than the mean

- **mode:** the most common measurement value

motivation: the most likely value

Mean, median, and mode

mean, median and mode for an example distribution:



- generally not equal to each other
- all 3 are useful; which to quote depends on the problem (and personal preference)

Deviation / variance / std. deviation

- **deviation** of one measurement: $d_i = x_i - \mu$
- sample **variance**: average of the squares of the deviations

$$\sigma^2 = \frac{1}{N} \sum_i (x_i - \mu)^2$$

- when computing from sample population:

$$s^2 = \frac{1}{N-1} \sum_i (x_i - \bar{x})^2$$

- **standard deviation**: $\sigma = \sqrt{\text{variance}}$

indicates how much the measurements typically deviate from the mean

How well can we measure the mean?



- width of sample distribution (std. dev.) remains the same
- the more measurements, the closer the mean to the true value

Uncertainty on the mean

- variance and std. deviation are measures of the *width* of the sample distribution
- with increasing number N of measurements, the typical deviation of **measured mean and true mean** decreases
- measurement uncertainty on the mean:
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$$
- width of distribution of repeated measurements of the mean
- σ : distribution of single measurements around the true value
- $\sigma_{\bar{x}}$: distribution of means of N measurements around the true value

Weighted mean

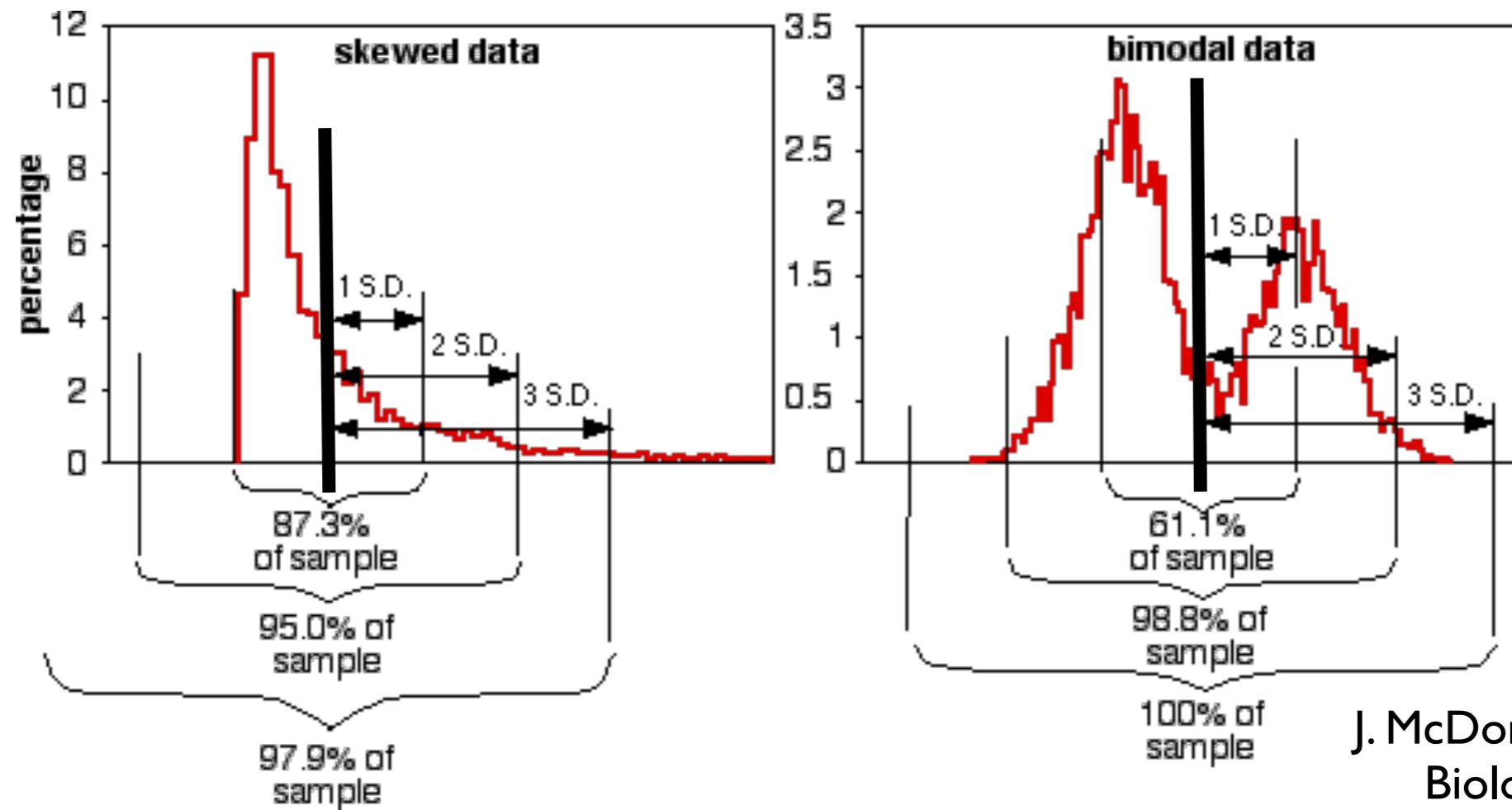
- previously, all measurements had equal weight
- some measurements are more precise than others; can assign weight w_i to each measurement x_i
- weighted mean:

$$\bar{x} = \frac{\sum_i w_i x_i}{\sum_i w_i} \quad \sigma_{\bar{x}}^2 = \frac{1}{\sum_i w_i}$$

- for reasonable (Gaussian) distributions, optimal weight is the inverse of the variance of each measurement:

$$\bar{x} = \frac{\sum_i x_i / \sigma_i^2}{\sum_i 1 / \sigma_i^2} \quad \sigma_{\bar{x}}^2 = \frac{1}{\sum_i 1 / \sigma_i^2}$$

- can calculate mean, variance, etc. for any set of data points
- *that does not guarantee that they are useful descriptions of the distribution !*



- if we know the shape of the parent distribution, we know which summary statistics to use

Common Probability Distributions

- many, many possible distributions have been quantified; here, consider 3 particularly important ones:
 - **Binomial distribution:** for experiments with only a small number of possible final states (e.g. coin toss)
 - **Poisson distribution:** counting experiments for discrete events (e.g. photon counts)
 - **Gaussian (or Normal) distribution:** distribution of events about the mean for a wide variety of processes; limiting case of binomial and poisson distributions

Binomial Distribution

- experiment with only two possible outcomes:
 - state 0: probability p
 - state 1: probability $q = (1-p)$
- n realizations
- probability that x of the n realizations are in state 0:

$$P_B(x|n,p) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

- x : positive integers; $0 \leq x \leq n$
- $0 < p < 1$
- $\sum_{x=0}^n P_B(x|n,p) = 1$

Binomial Distribution

- mean of the binomial distribution:

$$\mu = \sum_{x=0}^n (x \cdot P_B(x|n,p)) = np$$

(agrees with intuition!)

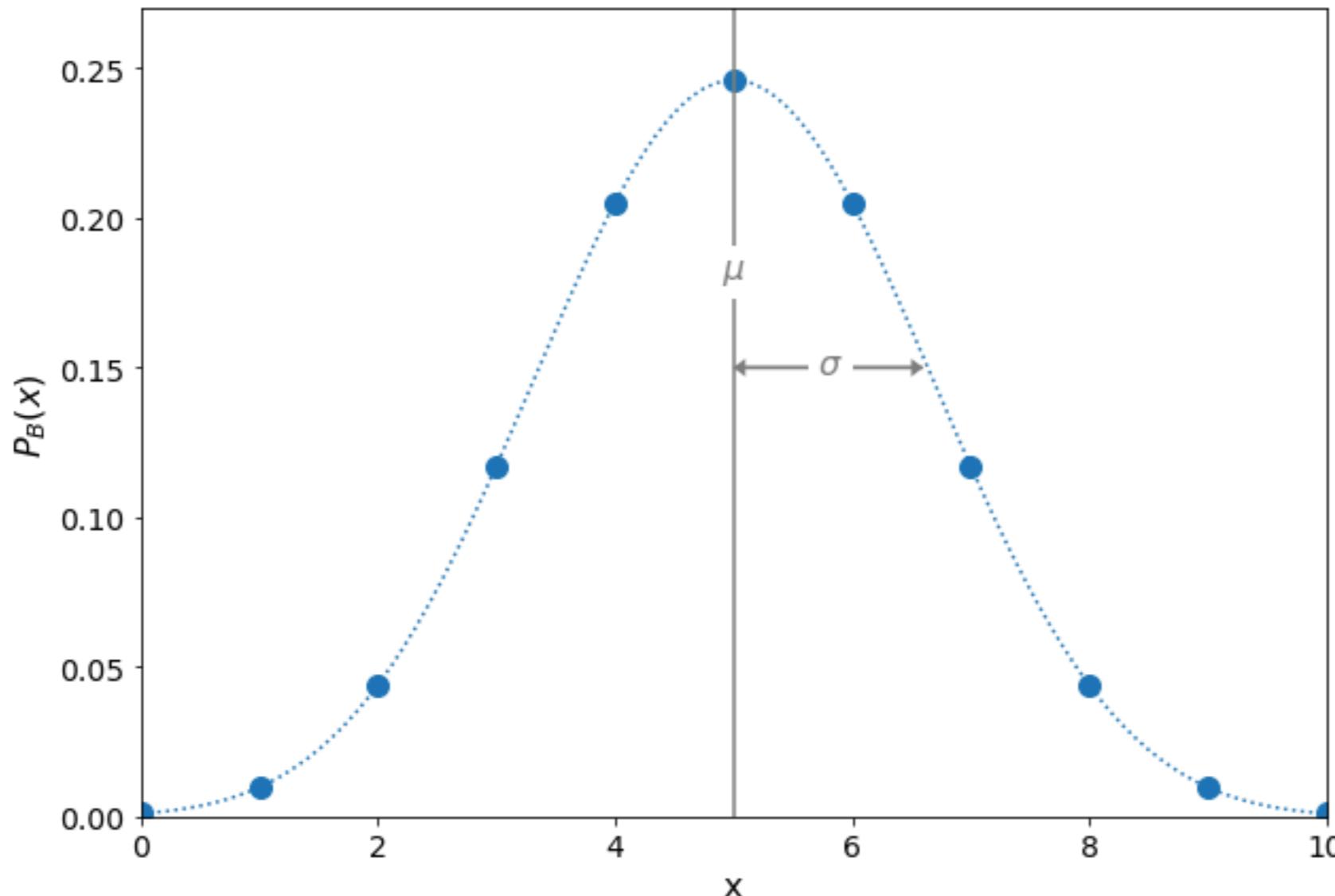
- variance of the binomial distribution:

$$\sigma^2 = \sum_{x=0}^n ((x - \mu)^2 \cdot P_B(x|n,p)) = np(1 - p)$$

Binomial Distribution - Example (I)

tossing 10 coins, x = number of tails

$$n = 10, p = 0.5 \rightarrow P(x) = P_B(x|10, 0.5)$$



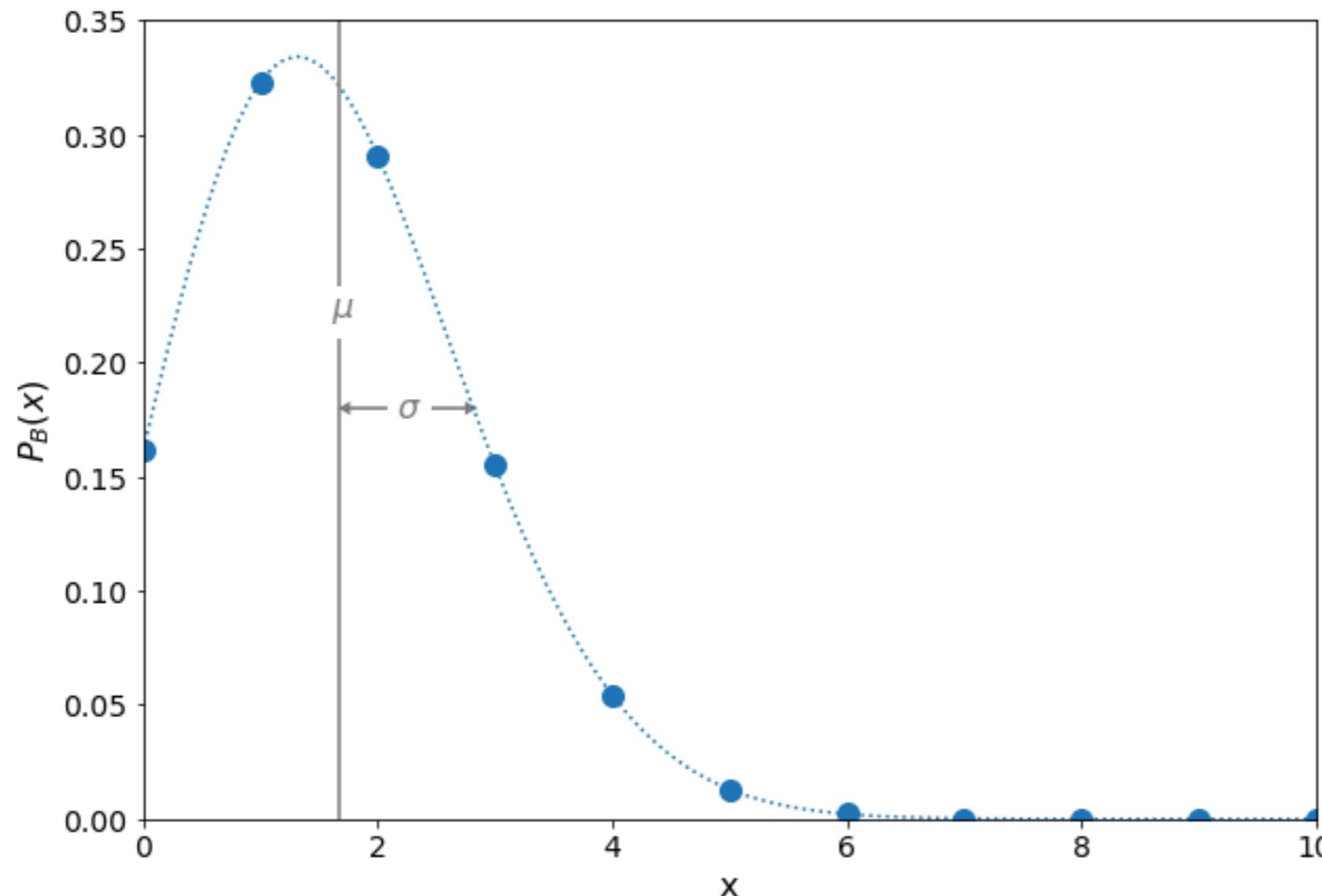
$$\begin{aligned}\mu &= np = 5 \\ \sigma^2 &= np(1-p) \\ &= 2.5\end{aligned}$$

since $q=p$: distribution is symmetric

Binomial Distribution - Example (2)

roll 10 dice, x = number of rolls with 6 eyes

$$n = 10, p = 1/6 \rightarrow P(x) = P_B(x|10, 1/6)$$



$$\begin{aligned}\mu &= np = 5/3 \\ \sigma^2 &= np(1-p) \\ &= 1.39\end{aligned}$$

note: mean is
not the mode

since $q \neq p$: distribution is not symmetric

Poisson Distribution

- limit of binomial distribution if number of trials is large, and probability of “success” in a given trial is small, while the mean $\mu = np$ remains finite

$$P_P(x|\mu) = \lim_{\substack{n \rightarrow \infty \\ p \rightarrow 0}} P_B(x|n,p) = \frac{\mu^x}{x!} e^{-\mu}$$

- for experiments where a mean number of events, μ , can be measured
- neither the number of realizations n , nor the probability of “success” p need to be known
- example: “counting” experiments, e.g. flux measurements

Poisson Distribution

- mean of the Poisson distribution: μ
- variance of the Poisson distribution: $\sigma^2 = \mu$
- standard deviation: $\sigma = \sqrt{\mu}$
- a flux measurement typically consists of measuring a number of events, N , per time interval Δt , with $\mu = N/\Delta t$
- assuming that the time interval is precisely known, we can assign an uncertainty to a single measurement from \sqrt{N}

Poisson Distribution - Example

On your CCD, you count a flux of 100 photons from a star.

Q: Report an estimate of the flux (in number of photons) and the uncertainty.

Only 1 measurement → flux estimate (μ) is 100 photons.

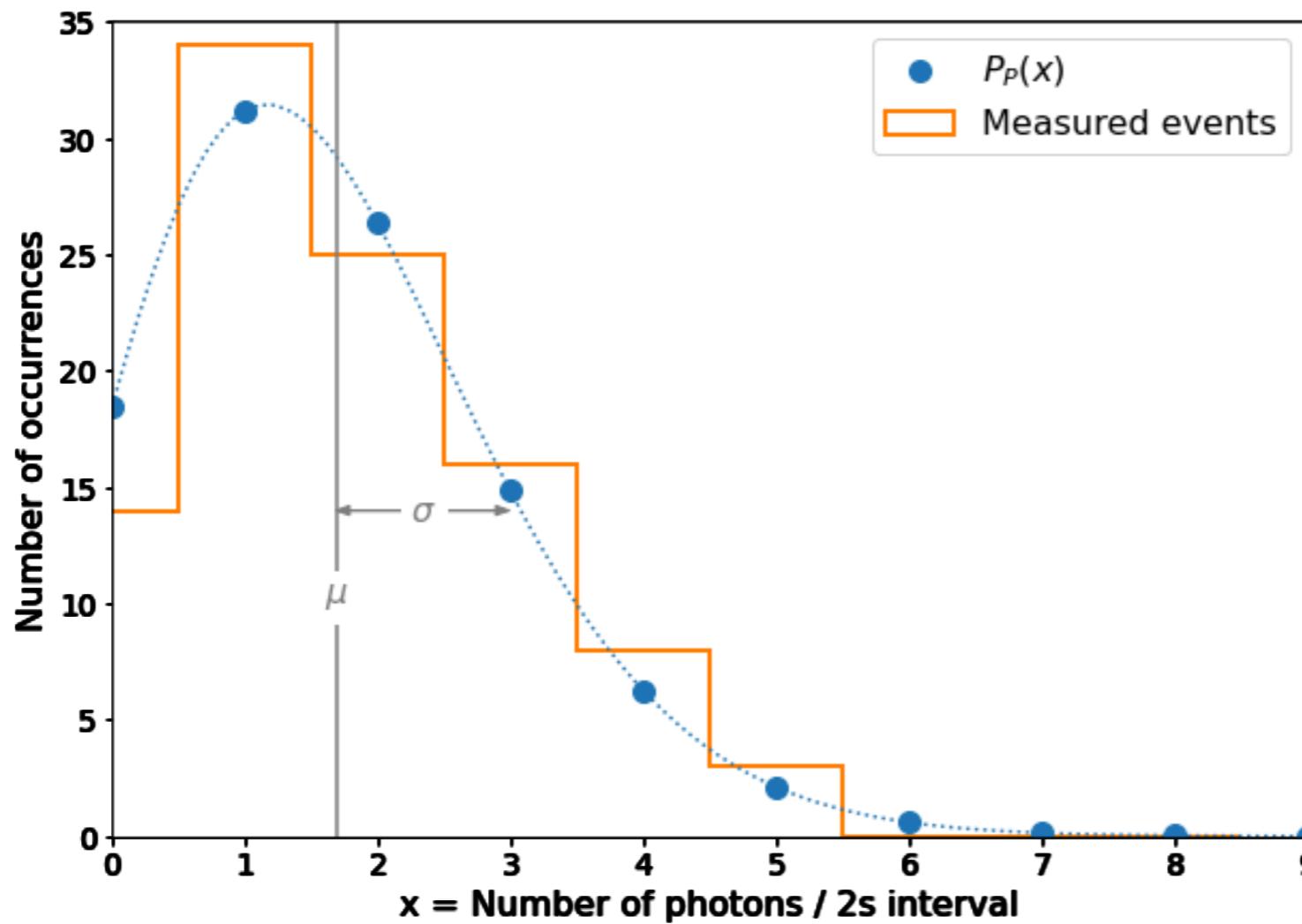
Counting experiment: standard deviation is $\sqrt{\mu}$.

Measurement: $F = 100 \pm 10$

Q: How can multiple measurements improve on this?

Poisson Distribution - Example

a detector measures the number of gamma-ray photons per 2 second interval, making 100 measurements:



measured mean:

$$\bar{x} = 1.69$$

blue points:

$$P_P(x|1.69)$$

note: Poisson distribution is defined for positive, integer values of x

Gaussian / Normal Distribution

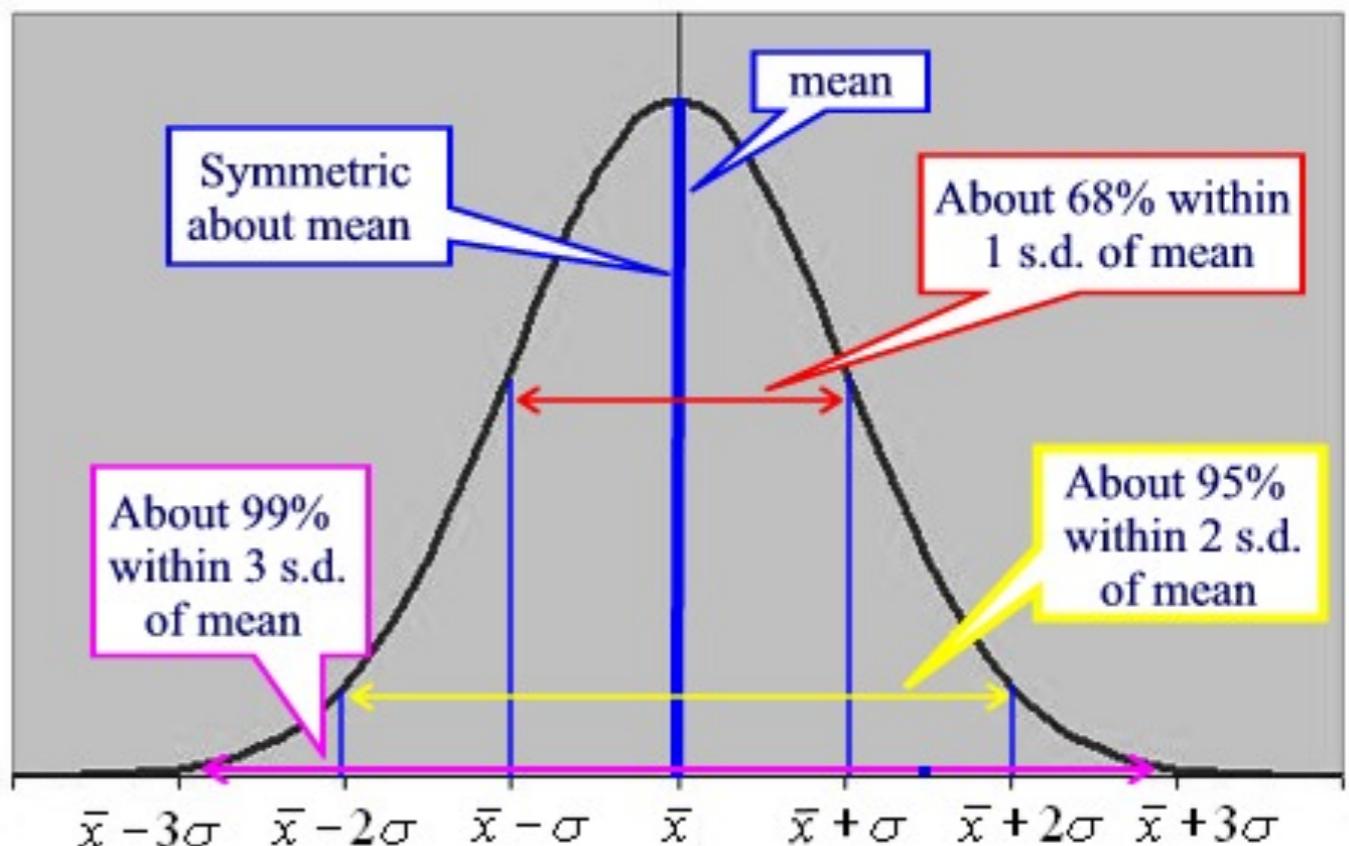
- the most commonly encountered probability distribution

$$P_G(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

- mean: μ , standard deviation: σ

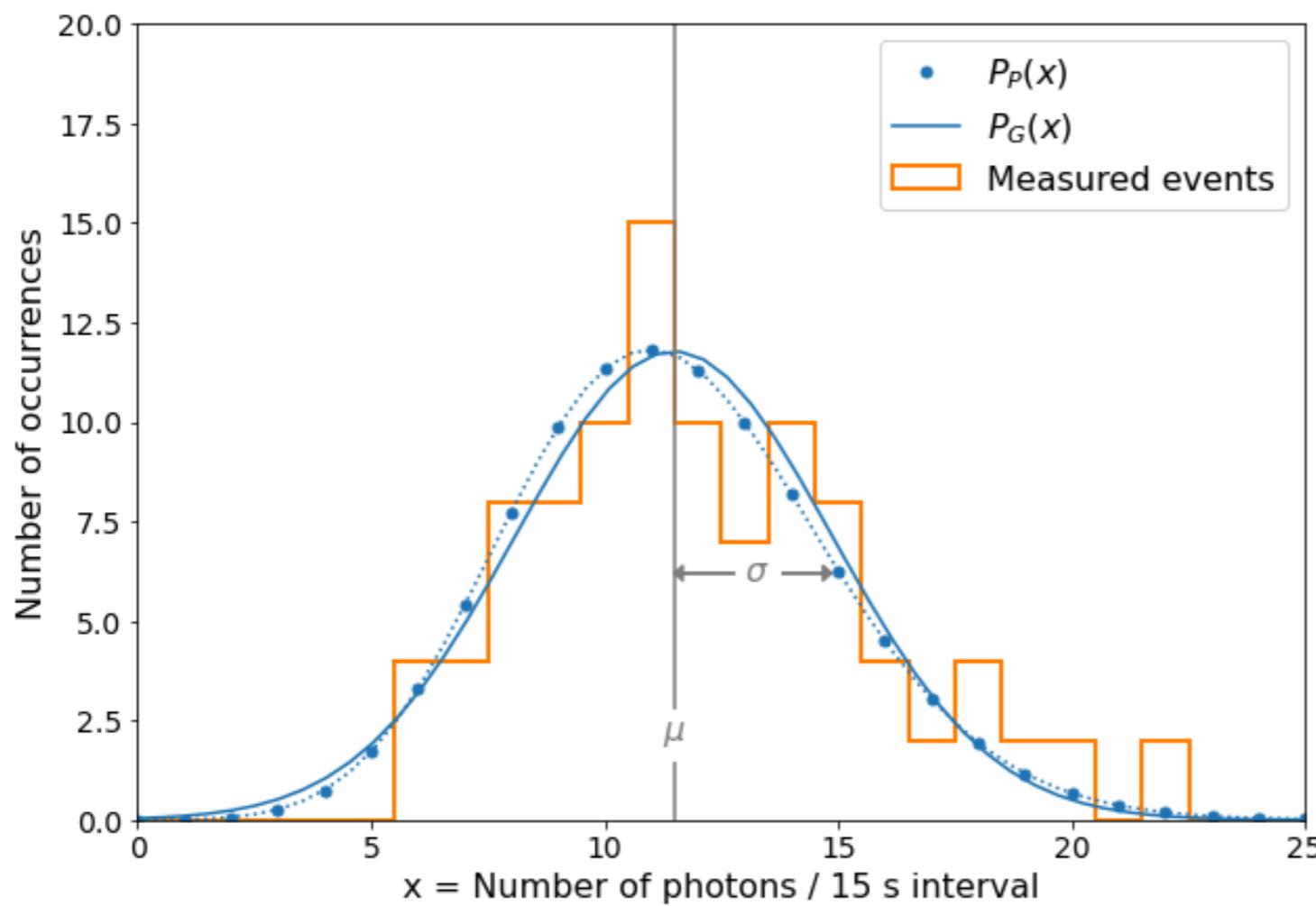
- can be derived as limit of the Poisson distribution for large values of the mean, $\mu \geq 30$

- can also be derived as limit of many other distributions



Gauss Distribution - Example

a detector measures the number of gamma-ray photons per 15 second interval, making 60 measurements:



measured mean:

$$\bar{x} = 11.48$$

blue points:

$$P_P(x|11.48)$$

blue curve:

$$P_G(x|\bar{x}, \sqrt{\bar{x}})$$

note: unlike Poisson distribution, Gaussian is continuous and defined for all x