ME170b Lecture 4

2/2/24

Experimental Techniques

Last time:

General Formula for error propagation

$$q = f(x, \ldots, z)$$

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\delta x\right)^2 + \cdots + \left(\frac{\partial q}{\partial z}\delta z\right)^2}$$

errors are random & independent

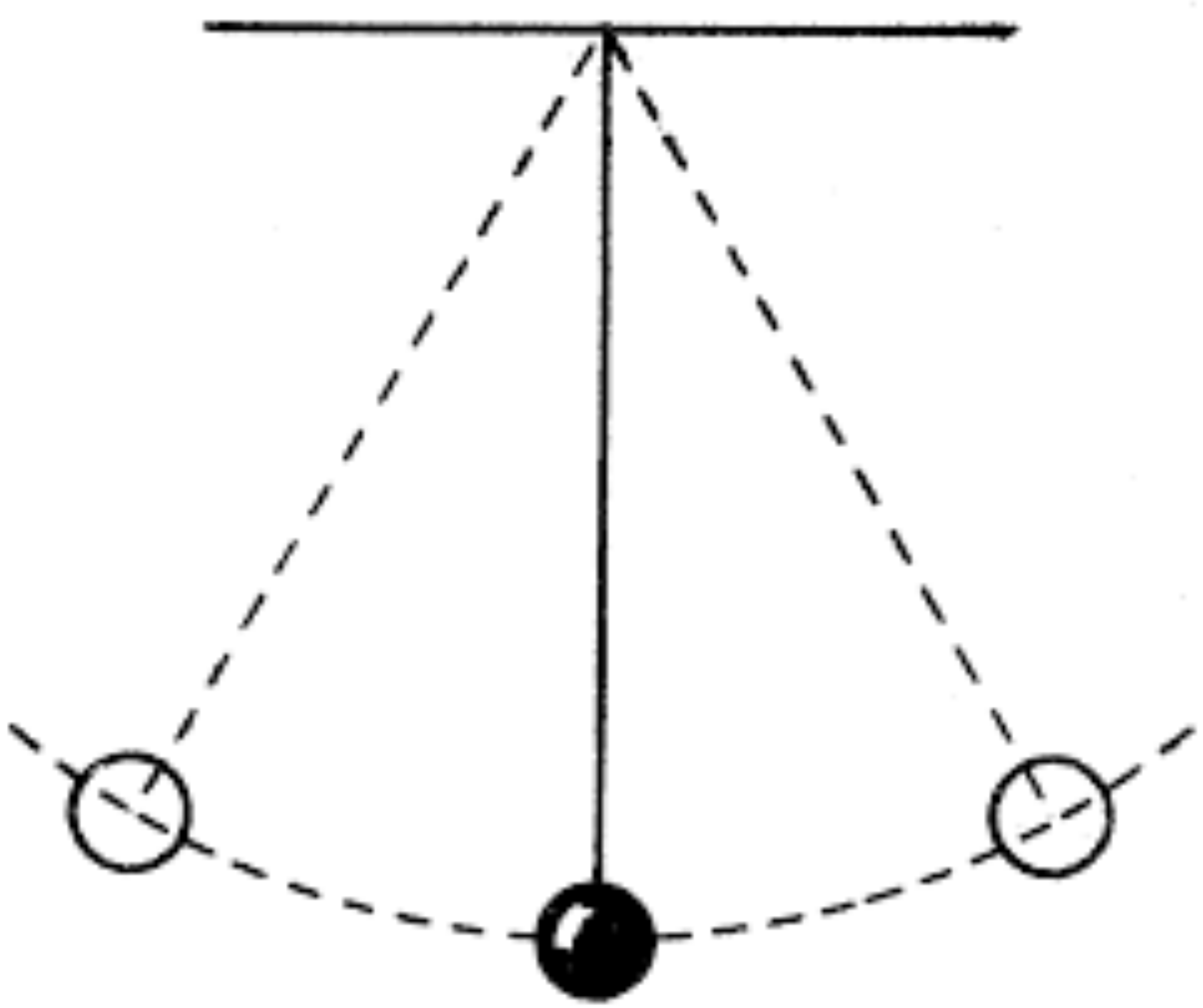
$$\delta q \leq \left|\frac{\partial q}{\partial x}\right|\delta x \cdots \left|\frac{\partial q}{\partial z}\right|\delta z$$

Today:

- statistical analysis (Ch. 4)
- Normal Distribution (Ch. 5)

Checking Understanding: Propagation of Uncertainties

Suppose we want to confirm the value of g, the acceleration due to gravity



Experiment: Use a pendulum, measure length & period → estimate g

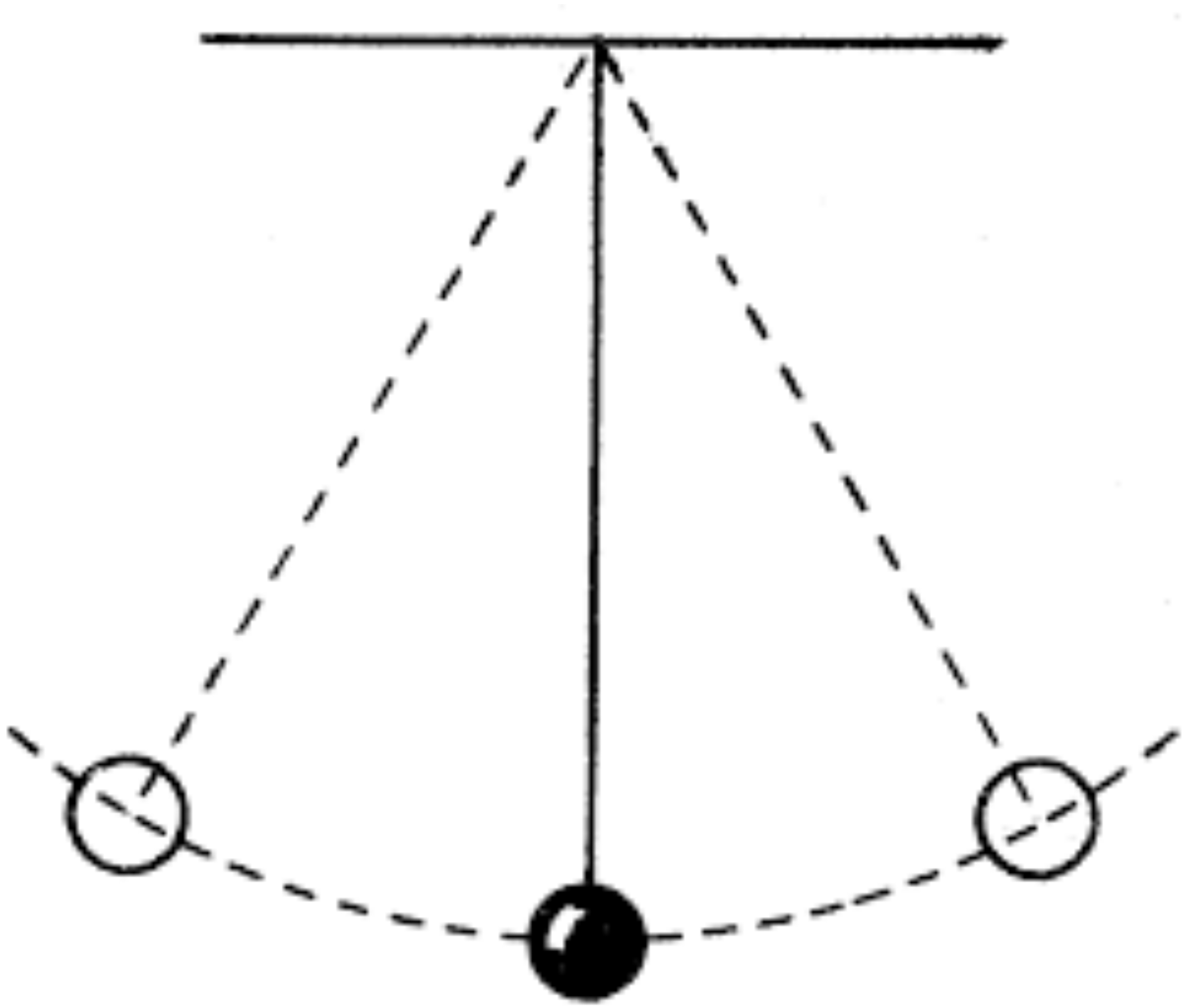
$$T = 2\pi\sqrt{\ell/g}$$

$$g = 4\pi^2 \ell / T^2$$

measure

measure

Checking Understanding: Propagation of Uncertainties



How can we track error propagation?

$$g = \underbrace{4\pi^2}_{\text{constant}} l / T^2$$

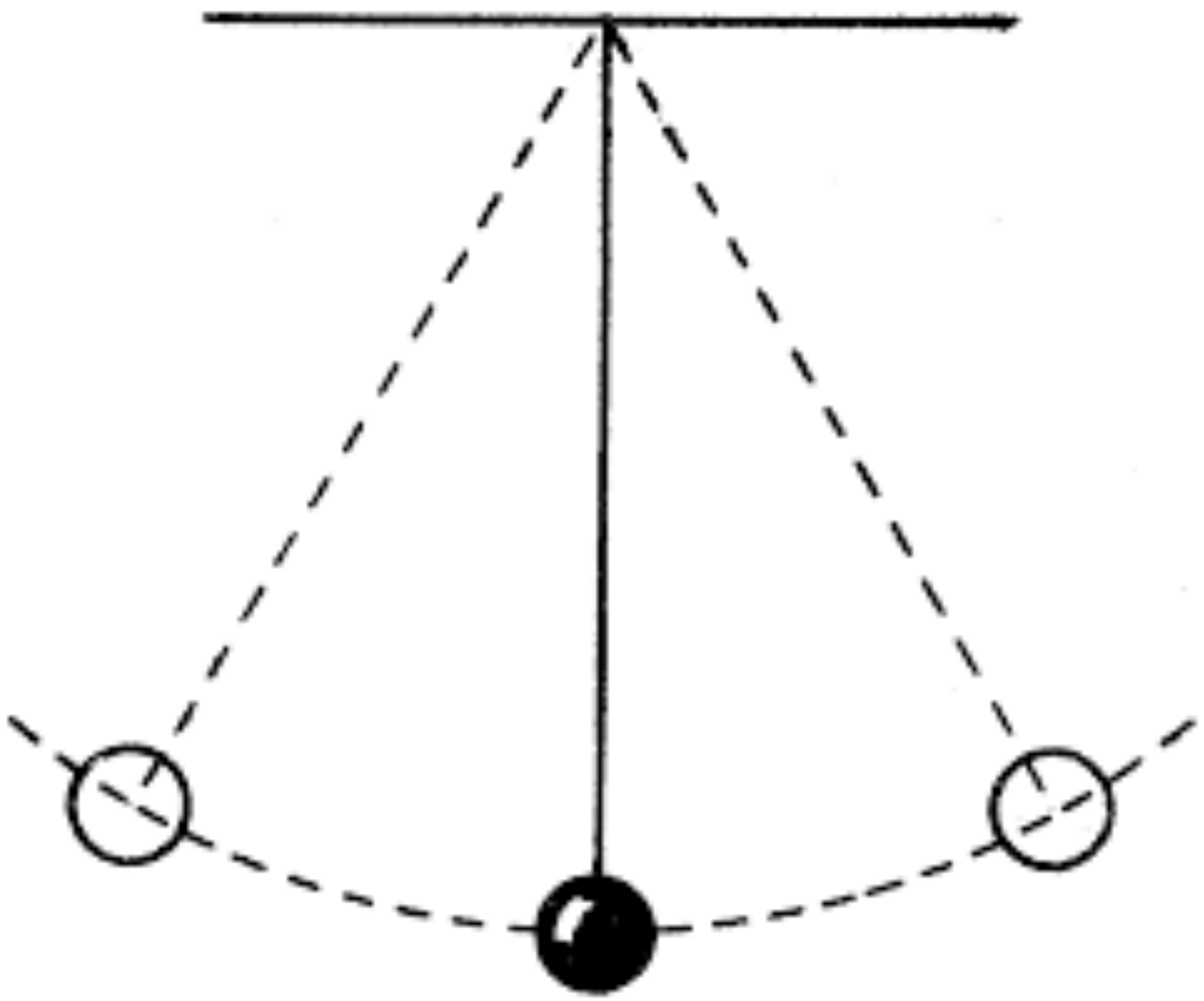
Provisional method:

- fraction uncertainties

  → add

  → $\sqrt{(\ )^2 + (\ )^2}$

Checking Understanding: Propagation of Uncertainties



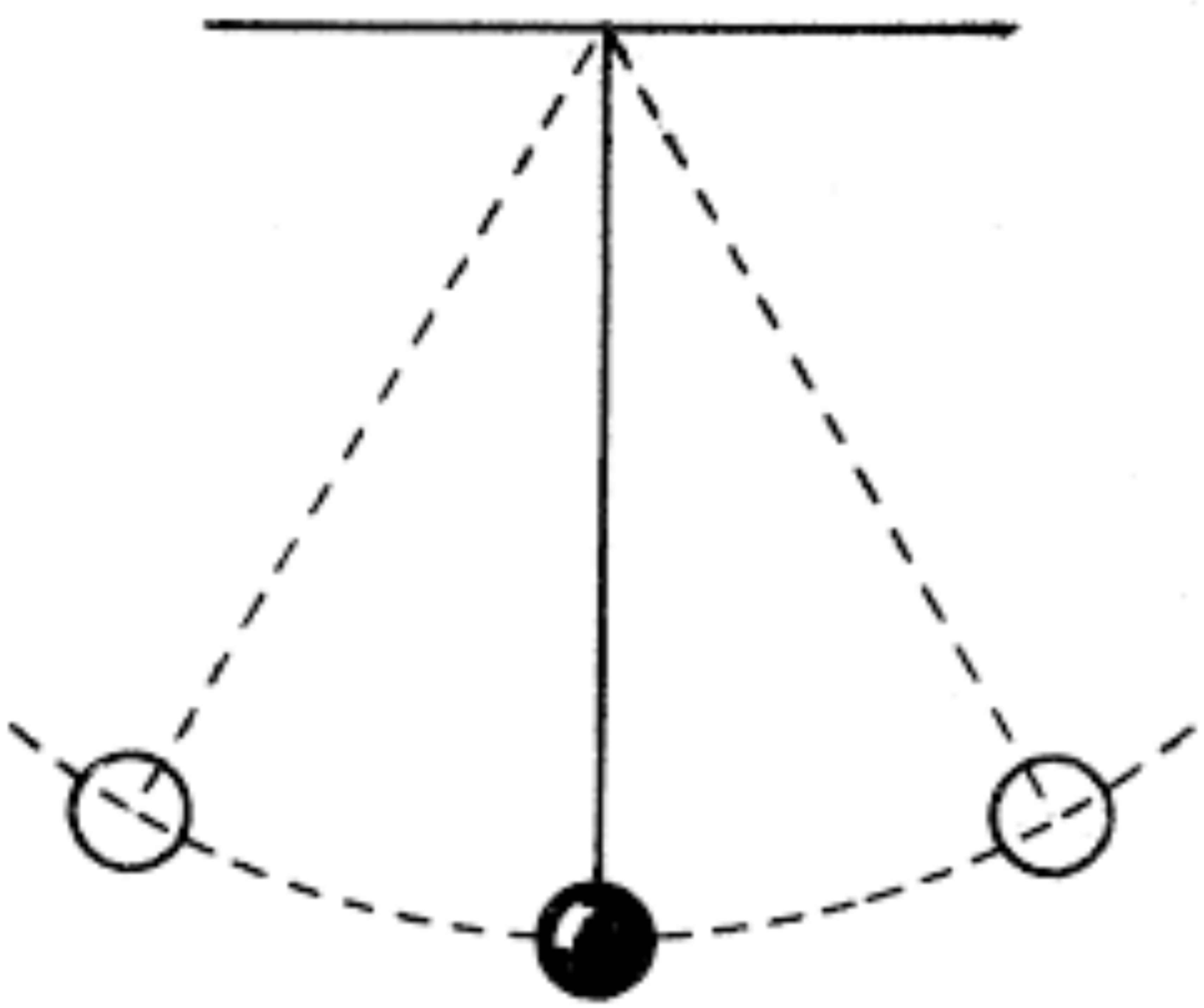
How can we track error propagation?

$$g = 4\pi^2 l/T^2$$

l (cm) all ± 0.1	T (sec) all ± 0.002	g (cm/s^2)	$\delta l/l$ (%)	$\delta T/T$ (%)	$\delta g/g$ (%)	answer $g \pm \delta g$
93.8	1.944	980	0.1	0.05	0.14	980 ± 1.4

error propogation

Checking Understanding: Propagation of Uncertainties



How can we track error propagation?

$$g = 4\pi^2 l/T^2.$$

General Formula for Error Propagation: If $q = q(x, \ldots, z)$ is any function of $x, \ldots, z$, then

$$\delta q = \sqrt{\left(\frac{\partial q}{\partial x}\,\delta x\right)^2 + \cdots + \left(\frac{\partial q}{\partial z}\,\delta z\right)^2}$$

(provided all errors are independent and random)

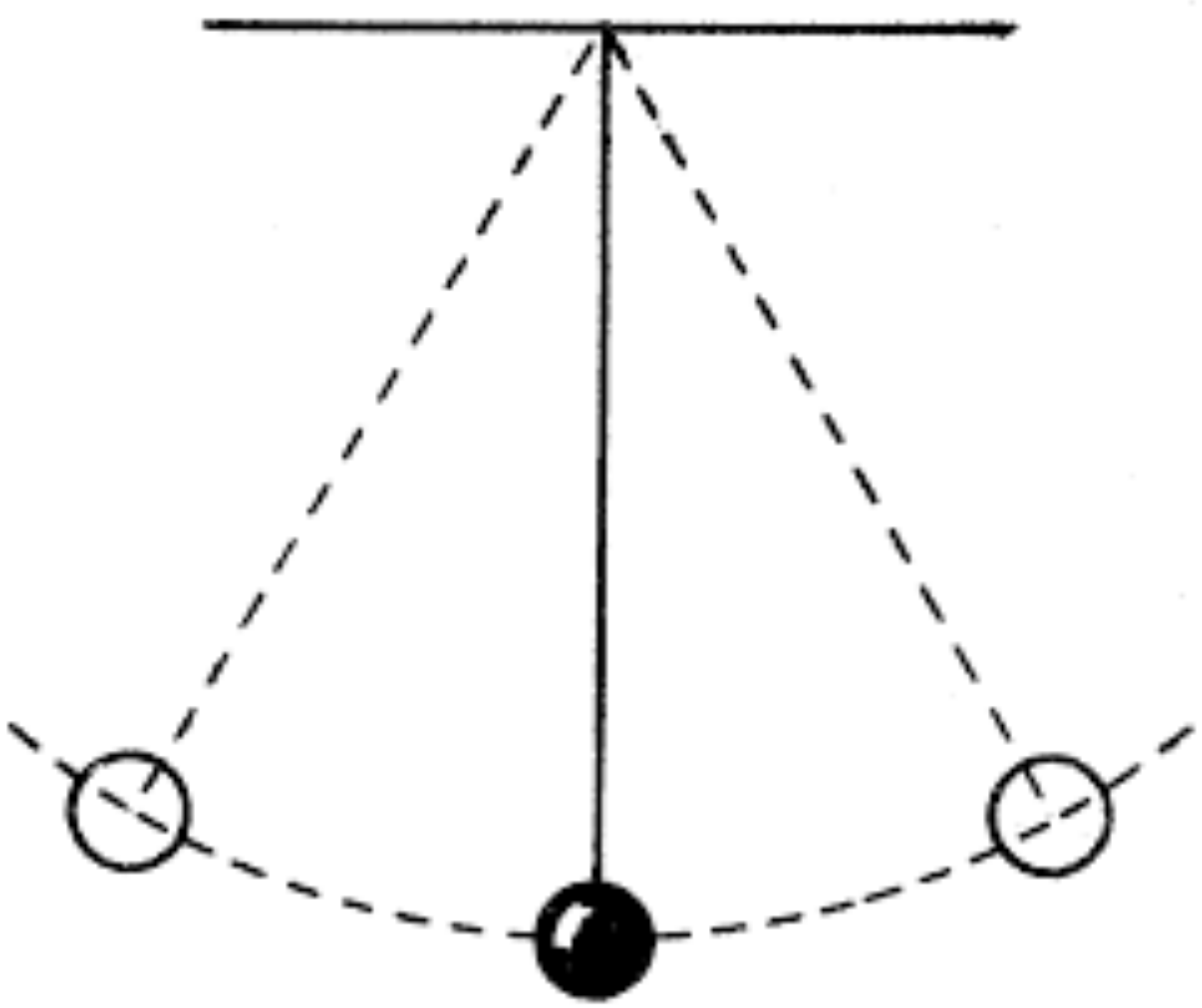
and

$$\delta q \leqslant \left|\frac{\partial q}{\partial x}\right| \delta x + \cdots + \left|\frac{\partial q}{\partial z}\right| \delta z$$

(always). [See (3.47) & (3.48)]

We can apply the general formula

Checking Understanding: Propagation of Uncertainties



How can we track error propagation?

$$g = 4\pi^2 l/T^2.$$

l (cm) all ± 0.1	T (sec) all ± 0.001	g (cm/s^2)
93.8	1.944	980
70.3	1.681	
45.7	1.358	
21.2	0.922	

How can we leverage repeated measures to estimate uncertainty directly from data? A: Statistics!

First, which types are errors can be estimated statistically?

Random Errors: uncertainties that can be revealed by <u>repeating the measurements</u>

Systematic Errors: errors that are not random are systematic

Consider timing the revolutions of a turntable using a stop watch

maybe we want to test whether a vintage record playing is at the correct RPM

Where do the errors come from?

- reaction time for stop watch (random, different every time)
- accuracy of the stop watch (systematic)

Conceptual example: ‘experiment’ is series of ‘shots’ at target accurate measurement == center of target. qualify systematic and random error for each (small or large)

random: small
systematic: small

random: small
systematic: large

random: large
systematic: small

large.

In reality we don't know the target!

random: small

systmatic: ?

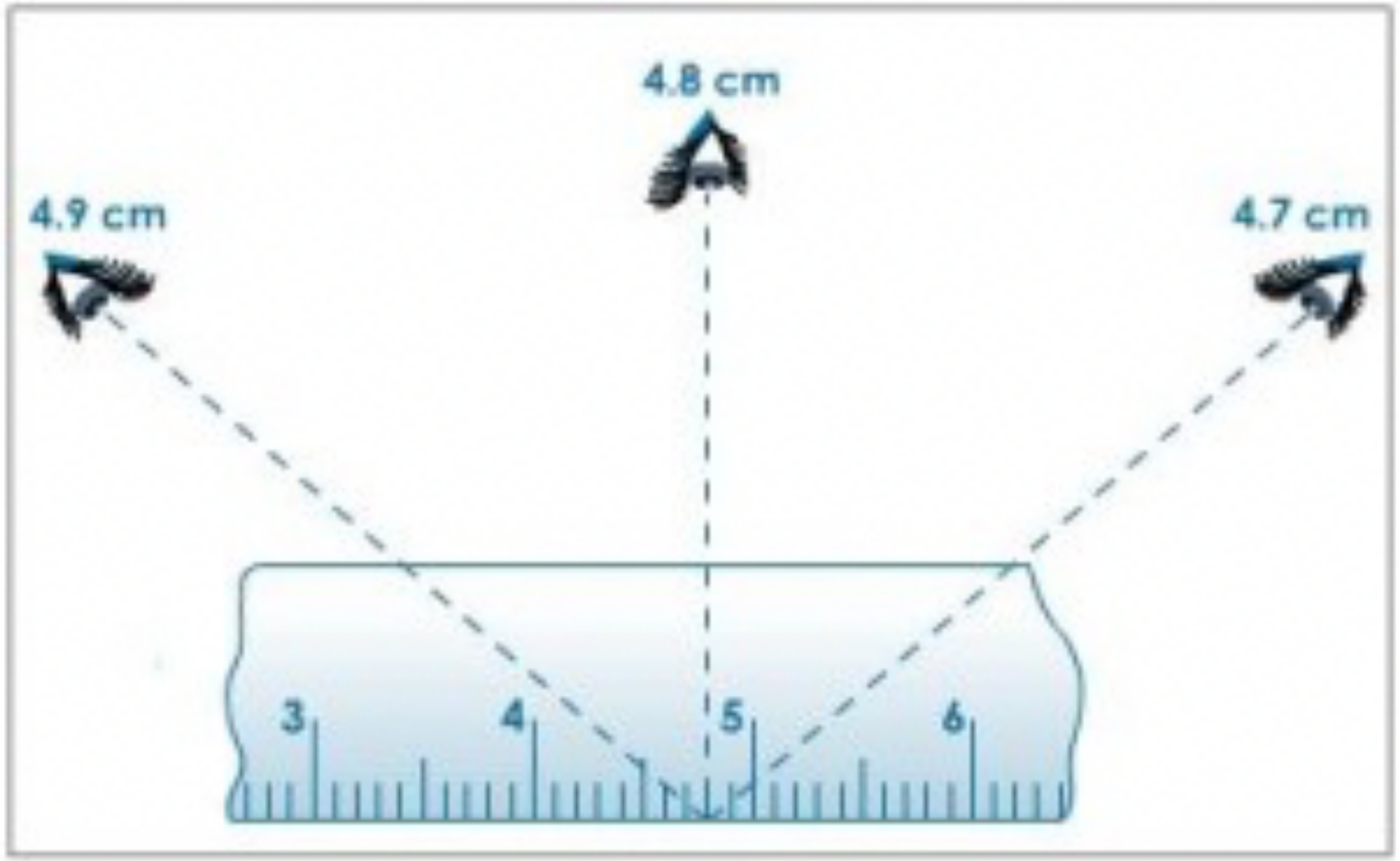
random: small

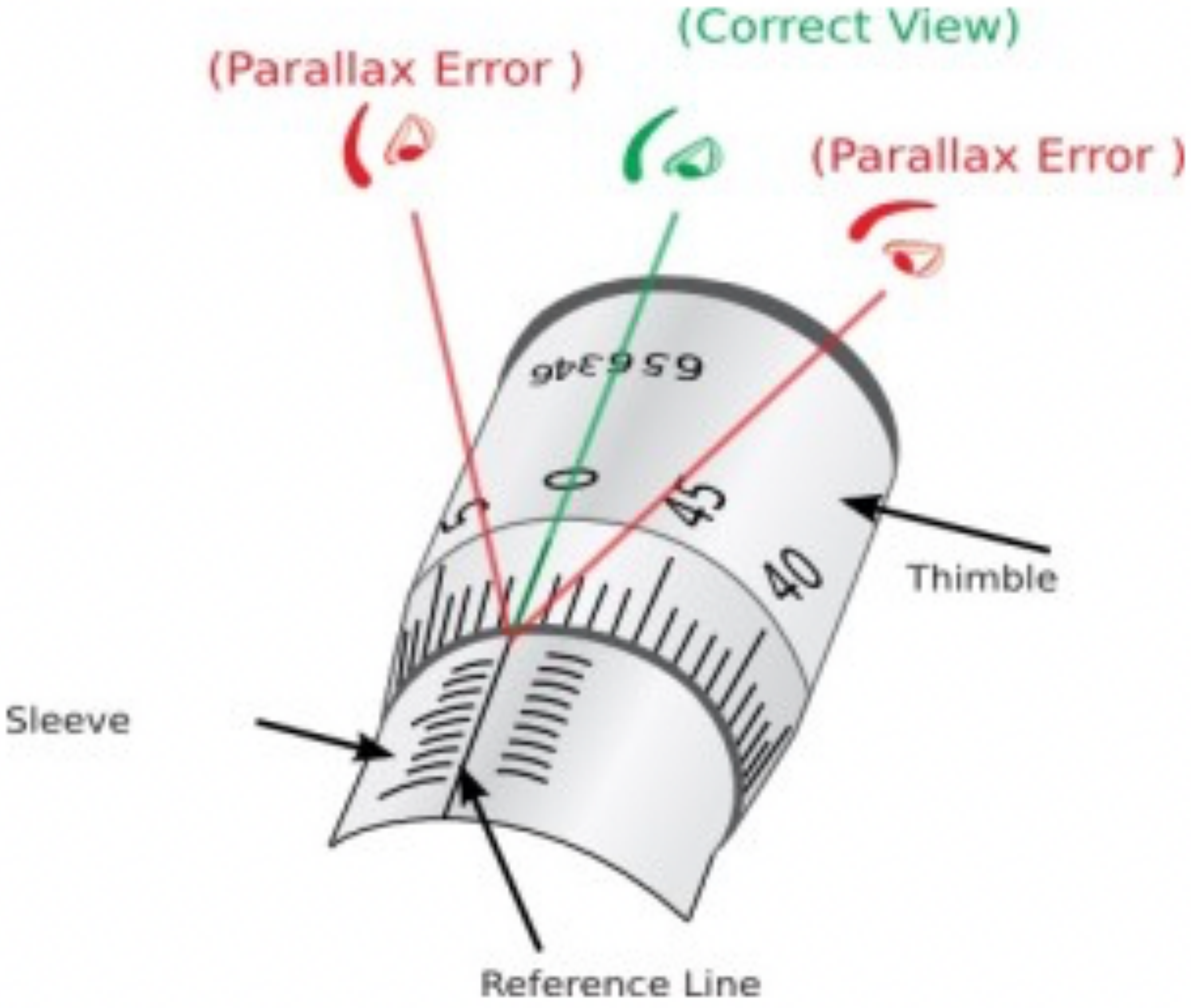
etc....

Key idea: we can identify random error even we don't know the target by repeating measures!

Errors are not always clear cut

random errors in one experiment may produce systematic errors in another





Dealing with errors

(1) Random errors

→ easier to deal with

→ we statistical tools to reliable estimate

(2) systematic errors

→ hard to evaluate & detect

→ Experimenters should carefully anticipate & eliminate source for systematic error.

→ calibrate device, control environment

Statistics: the very basics

> context: we want to minimize systematic error so that <u>all</u> of uncertainty comes from random variation

→ repeated measures

71, 72, 72, 73, 71

what should we use for our x_best?

$$x_{best} = \bar{x} = \frac{71 + 72 + 72 + 73 + 71}{5} = 71.8$$

Statistics: the very basics

71, 72, 72, 73, 71

How can we estimate uncertainty from our five measurements?

Standard Deviations?

What is a standard deviation?

71, 72, 72, 73, 71

x_1 x_2 x_3 x_4 x_5

$$d_i = x_i - \bar{x}$$

can we just take the average deviation for our uncertainty?

Trial number i	Measured value x_i	Deviation $d_i = x_i - \bar{x}$
1	71	-0.8
2	72	0.2
3	72	0.2
4	73	1.2
5	71	-0.8
	$\Sigma x_i = 359$	$\Sigma d_i = 0$

$$\text{mean, } \bar{x} = \Sigma x_i / N = 359/5 = 71.8$$

Definition of standard deviation

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(d_i)^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2}$$

$$\sigma_x^2 = \frac{1}{N}\sum d_i^2$$

variance

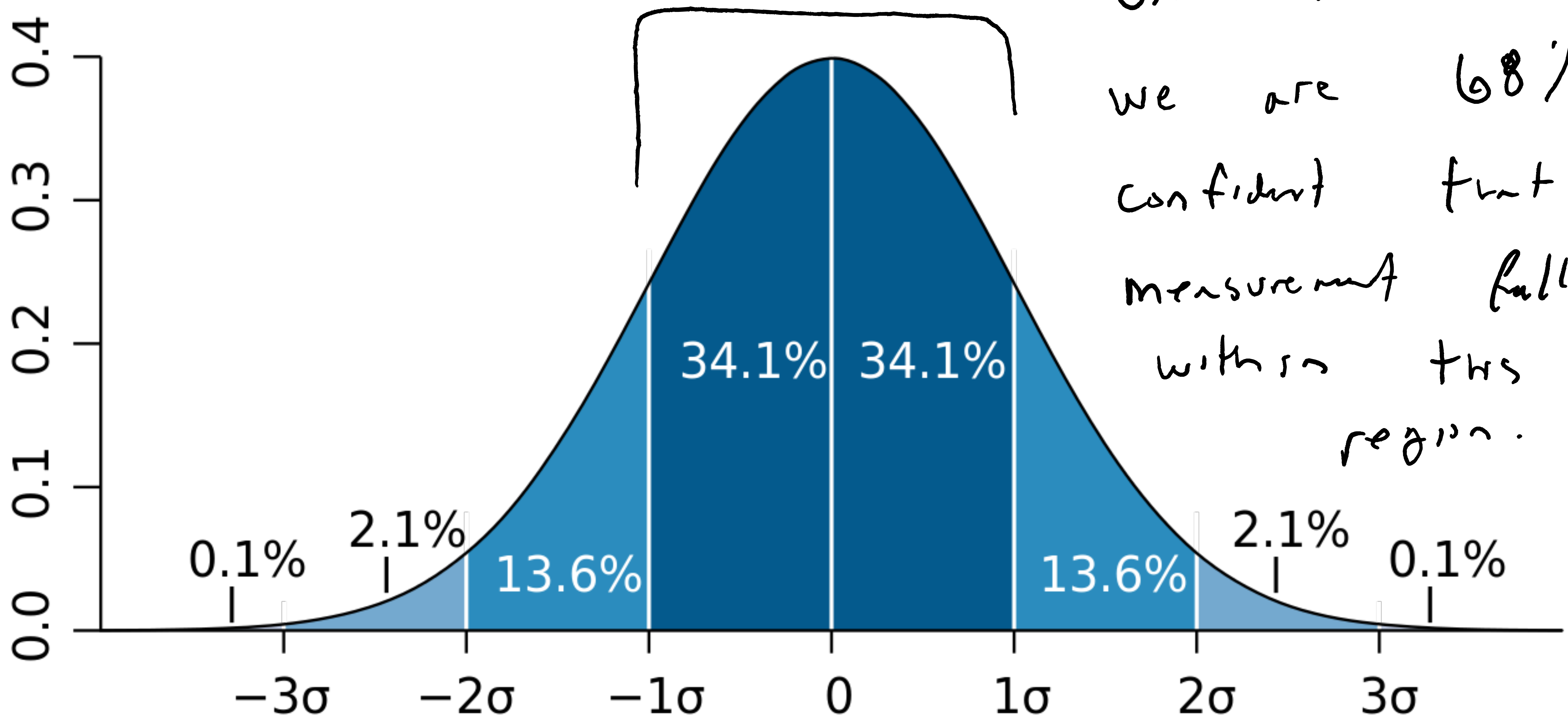
Bessels correction

$$\sigma_x = \sqrt{\frac{1}{N-1} \sum (x_i - \bar{x})^2}$$

Sample deviation - most appropriate for our purposes

Standard deviation as uncertainty of a single measurement

assumption — errors are random normally distributed



$\delta x = \sigma_x$

we are 68% confident that measurement falls within this region.

Standard deviation of the mean (standard error)

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{N}}$$

this is a good estimate for uncertainty using <u>all</u> <u>measurements</u>.

$$\boxed{x = \bar{x} \pm \sigma_{\bar{x}}}$$

* making more measurements doesn't really change SD.
* SE will **decrease** with increasing N.

CH.5 The Normal Distribution
Histograms and Distributions

By now, should be clear: serious uncertainty analysis, via statistics, requires taking many measurements.

→ tools to visualize these measurements.

$$\boxed{26, 24, 26, 28, 23, 24, 25, 24, 26, 25.}$$

written out, data doesn't convey much.

Histograms and Distributions

23, 24, 24, 24, 25, 25, 26, 26, 26, 28. we can order

Table 5.1. Measured lengths x and their numbers of occurrences.

Different values, x_k	23	24	25	26	27	28
Number of times found, n_k	1	3	2	3	0	1

$$F_k = \frac{n_k}{N}$$

Histograms and Distributions

Table 5.1. Measured lengths x and their numbers of occurrences.

Different values, x_k	23	24	25	26	27	28
Number of times found, n_k	1	3	2	3	0	1

$$\bar{x} = \sum_k x_k F_k$$: weighted sum of all different values x_k

where x_k is weighted by time fraction of time it occured.

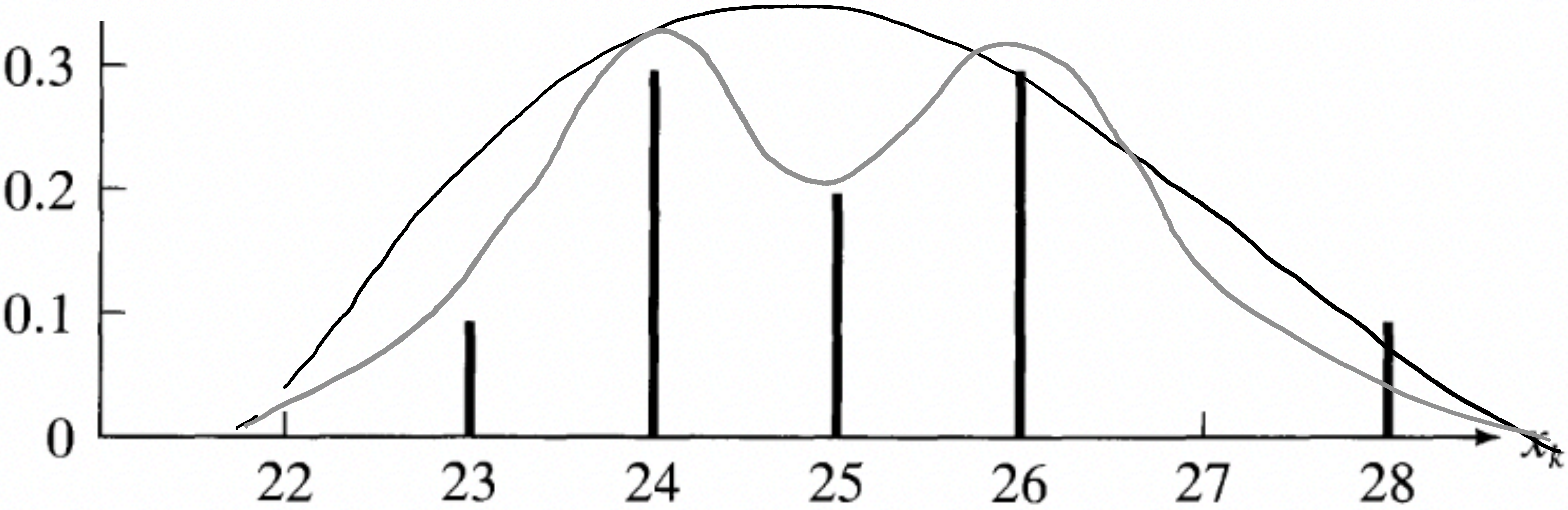
$$\sum F_k = 1$$

Now we are ready to display!

Histograms and Distributions

Table 5.1. Measured lengths x and their numbers of occurrences.

Different values, x_k	23	24	25	26	27	28
Number of times found, n_k	1	3	2	3	0	1



Histograms and Distributions

What about for 'untidy data'

26.4, 23.9, 25.1, 24.6, 22.7, 23.8, 25.1, 23.9, 25.3, 25.4.

Table 5.2. The 10 measurements (5.9) grouped in bins.

Bin	22 to 23	23 to 24	24 to 25	25 to 26	26 to 27	27 to 28
Observations in bin	1	3	1	4	1	0

Histograms and Distributions

What about for ‘untidy data’

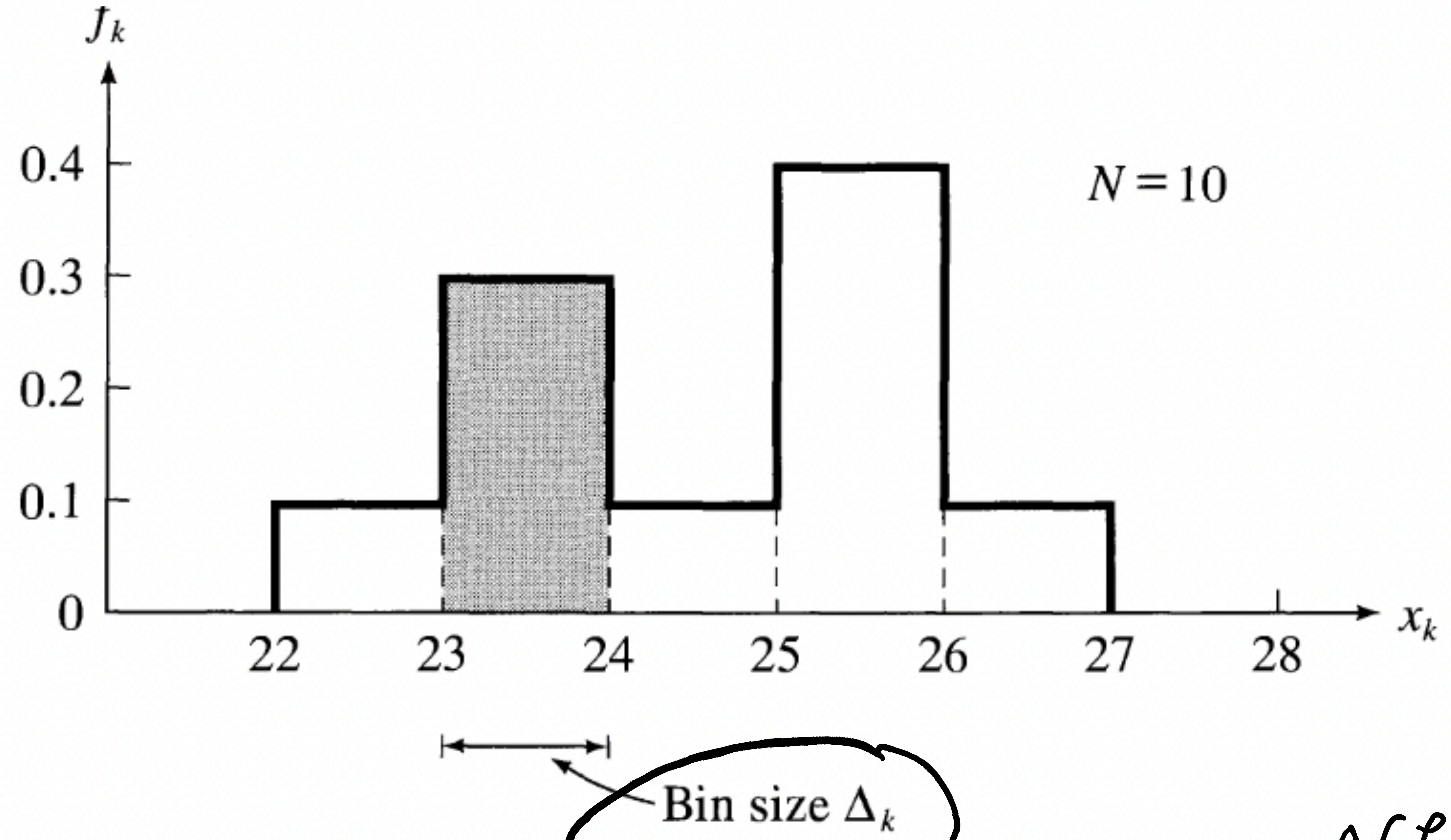
26.4, 23.9, 25.1, 24.6, 22.7, 23.8, 25.1, 23.9, 25.3, 25.4.

Histograms and Distributions

Table 5.2. The 10 measurements (5.9) grouped in bins.

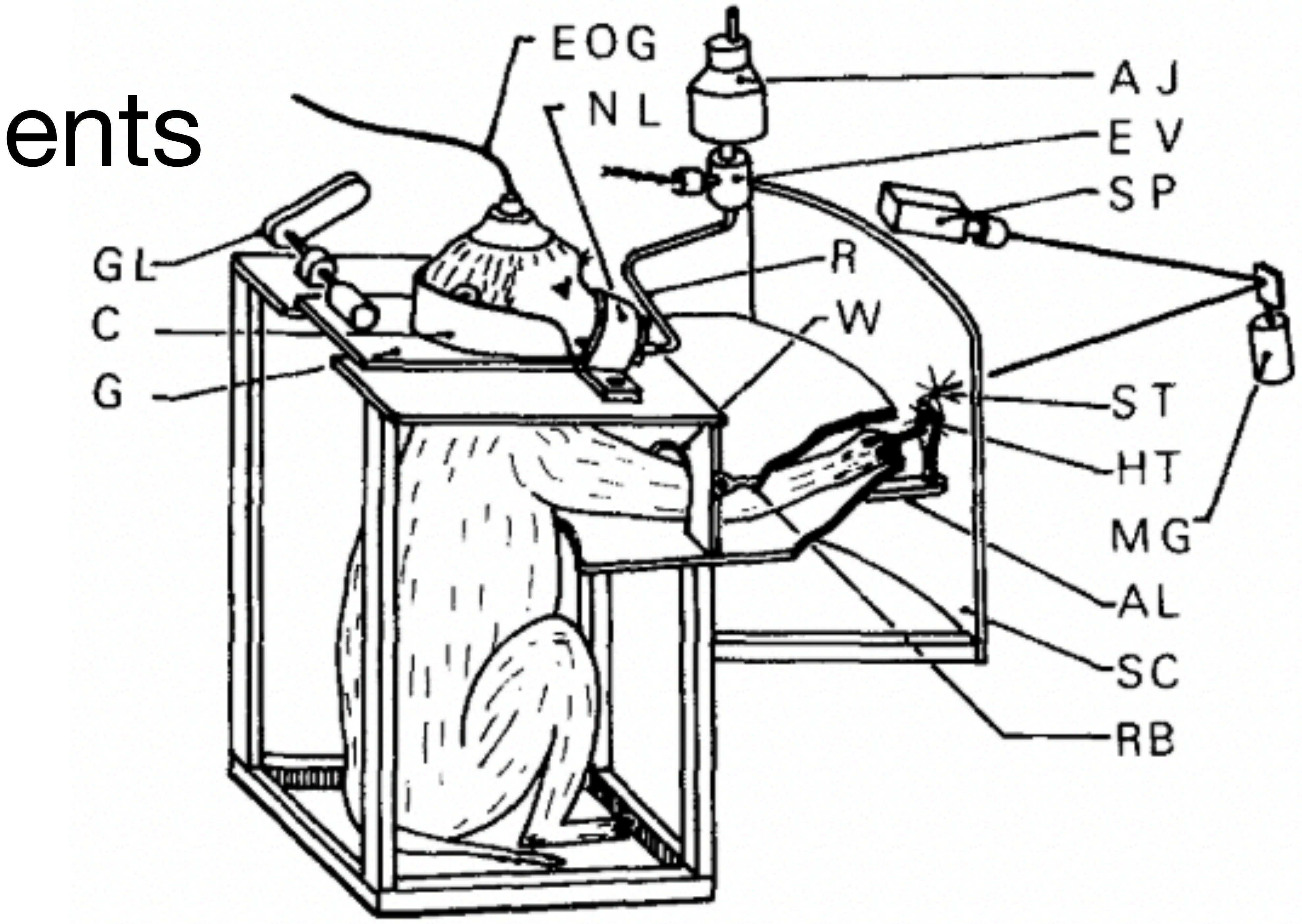
Bin	22 to 23	23 to 24	24 to 25	25 to 26	26 to 27	27 to 28
Observations in bin	1	3	1	4	1	0

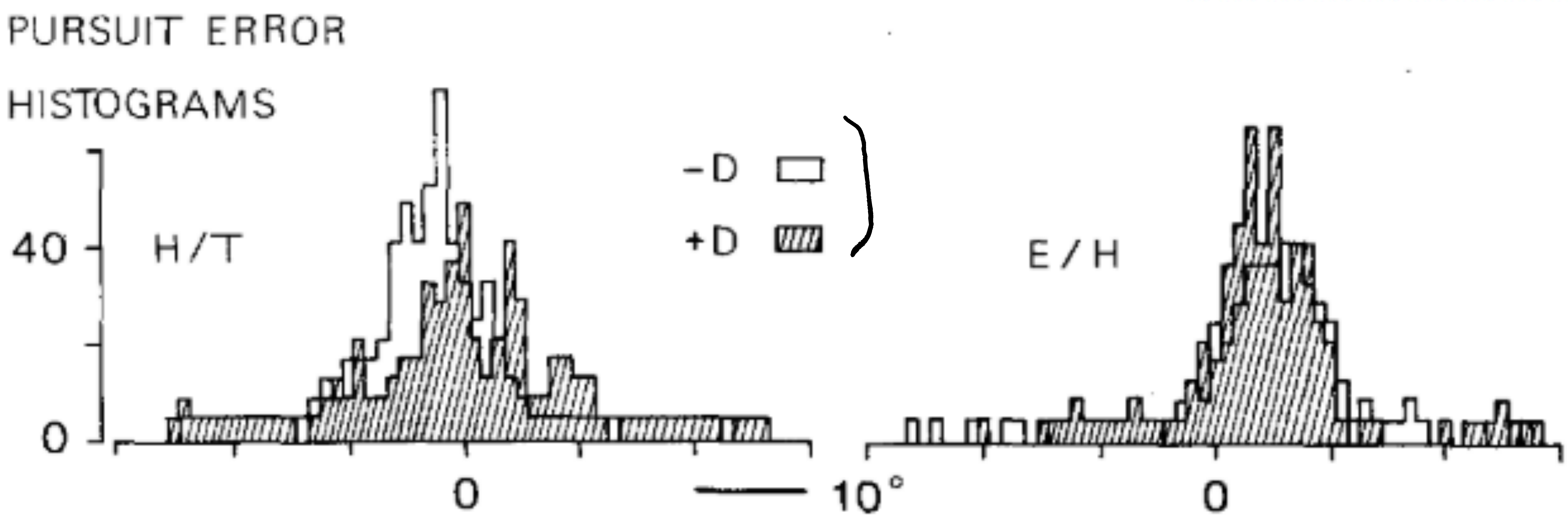
$f_K \Delta_K$ = fraction of measurements in the Kth bin

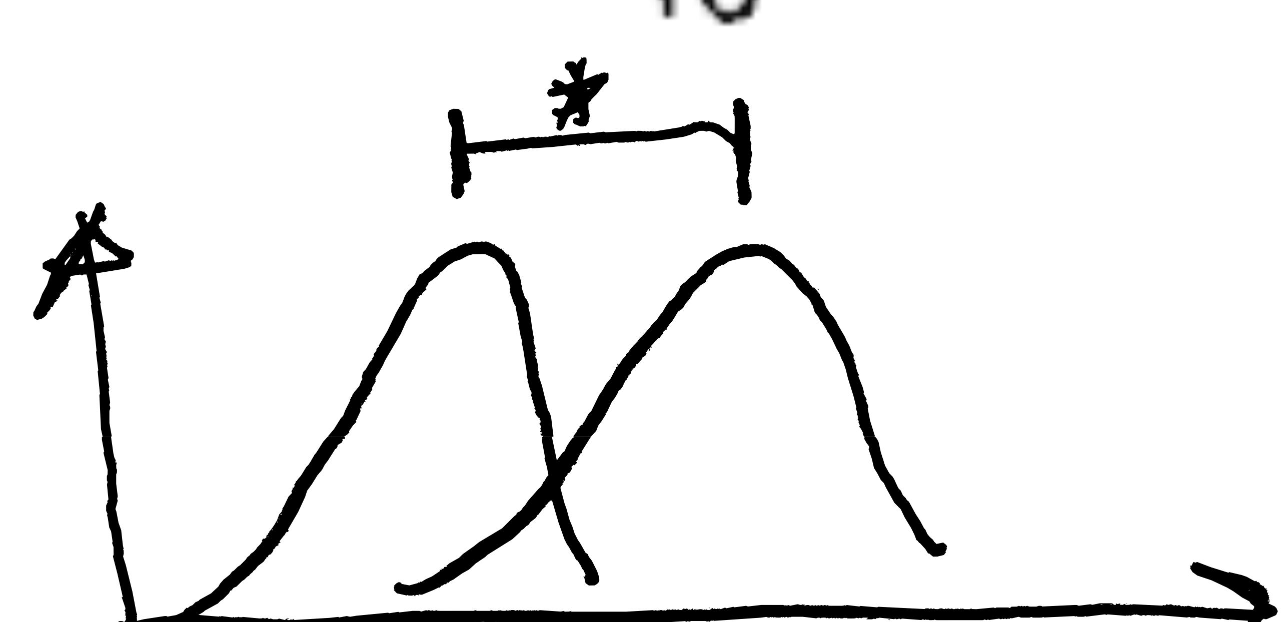


area $f_K \Delta_K$ corresponds to F_K in bar histogram

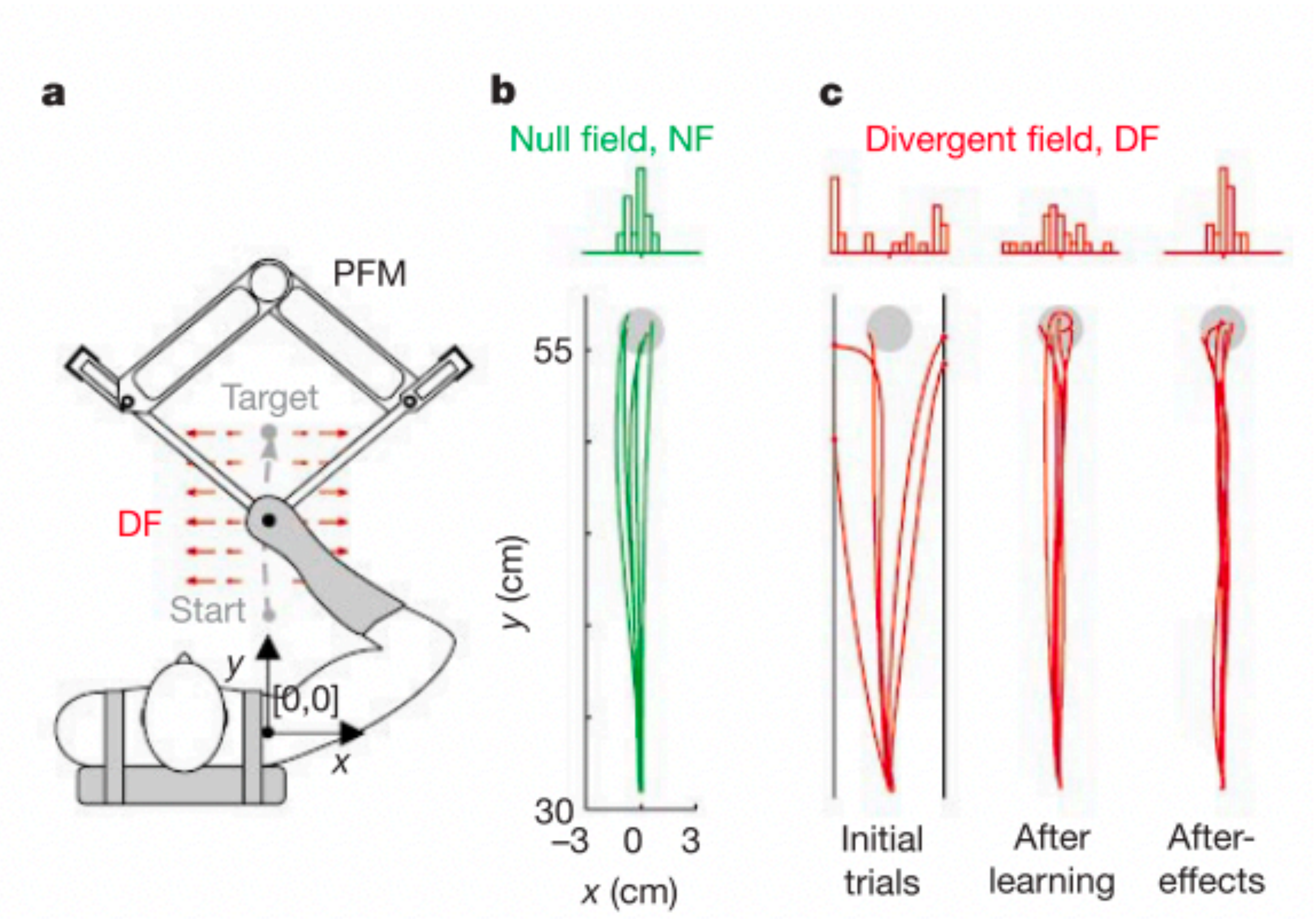
Example of histograms from real experiments



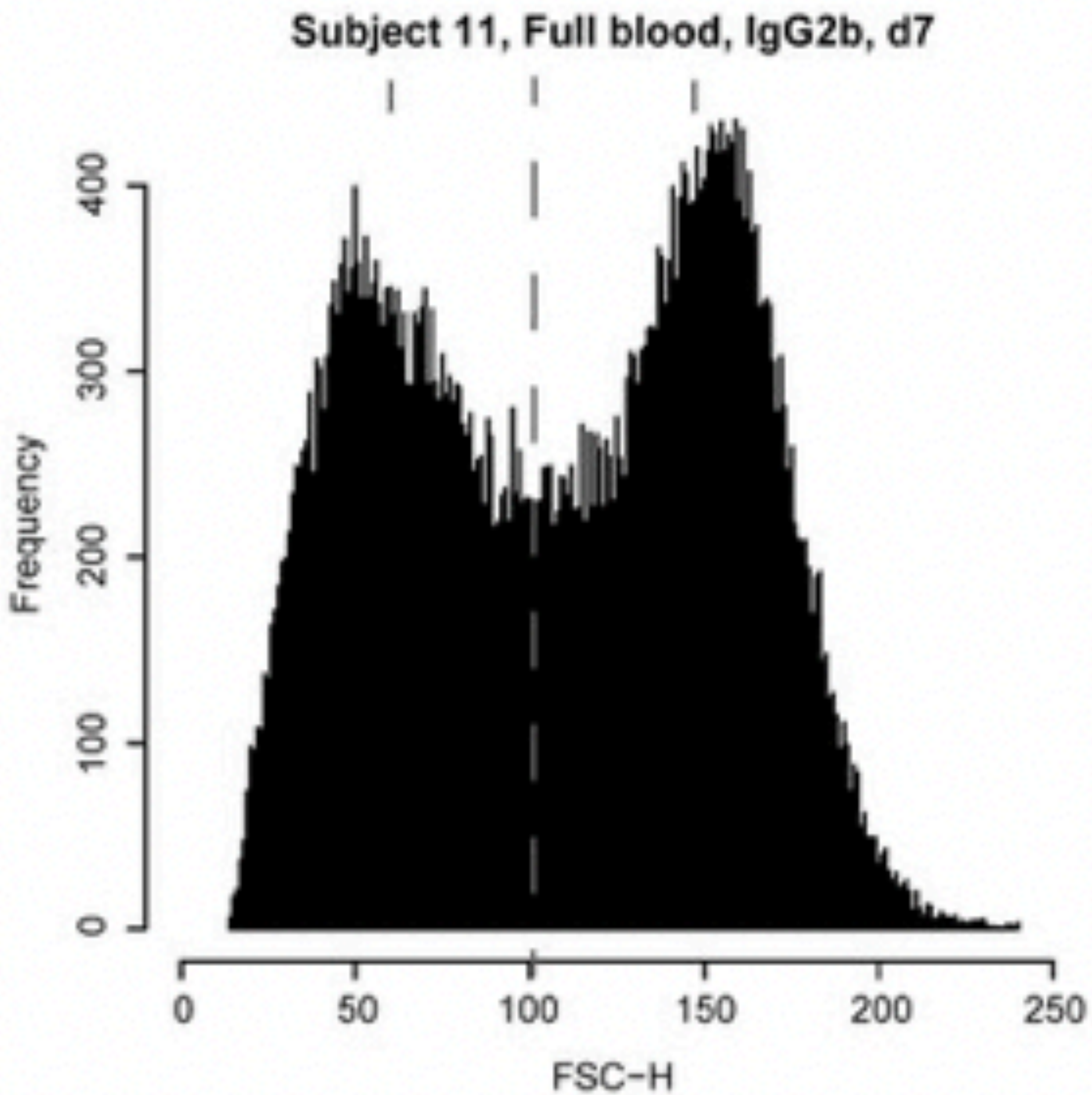


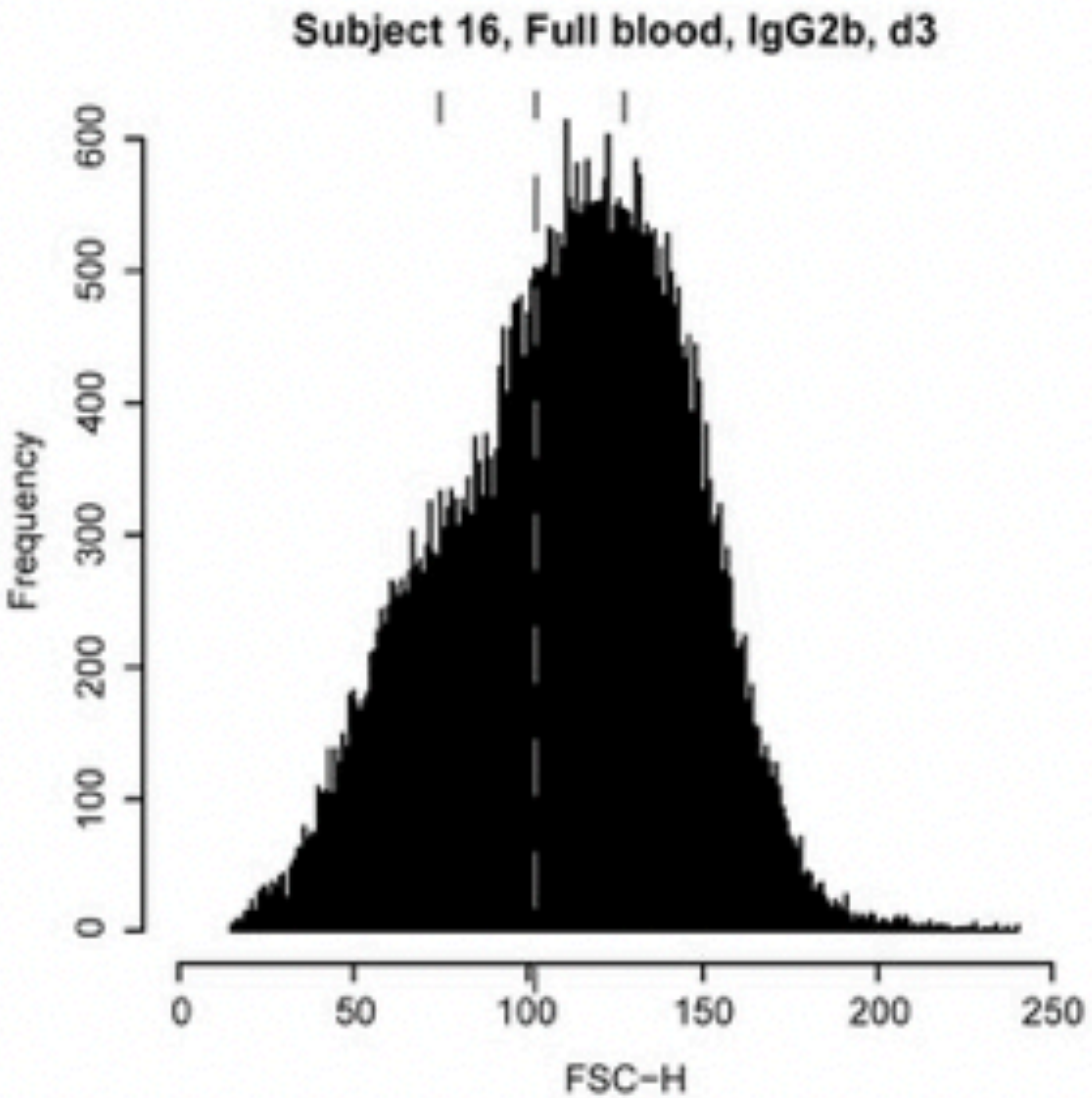


Example of histograms from real experiments

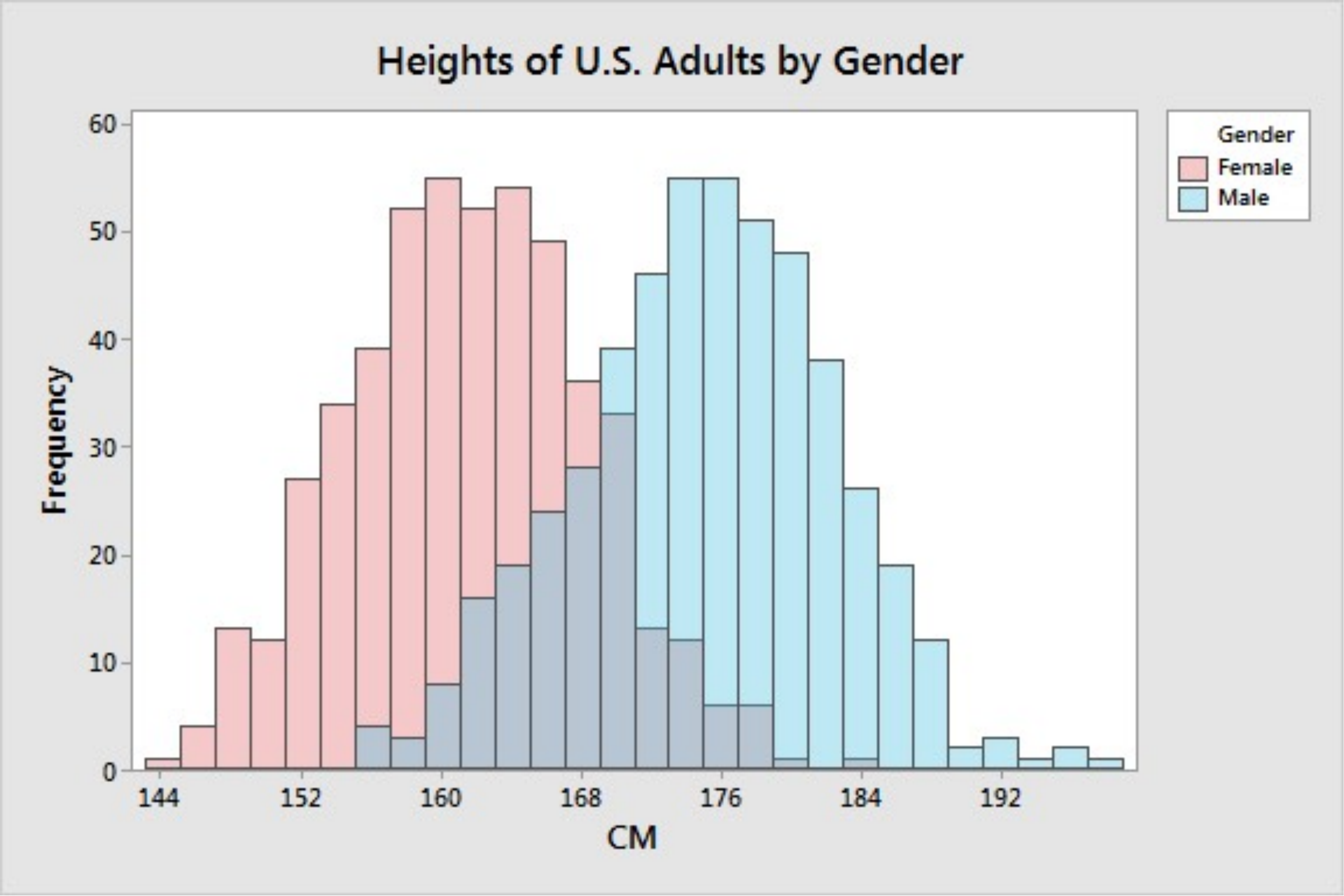


Example of histograms from real experiments





It’s really easy to visually test hypothesis with histograms



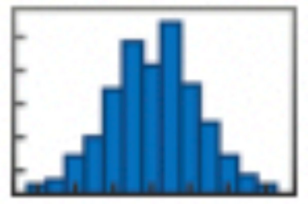
Histograms with matlab

histogram

Histogram plot

R2022b

expand all in page



Description

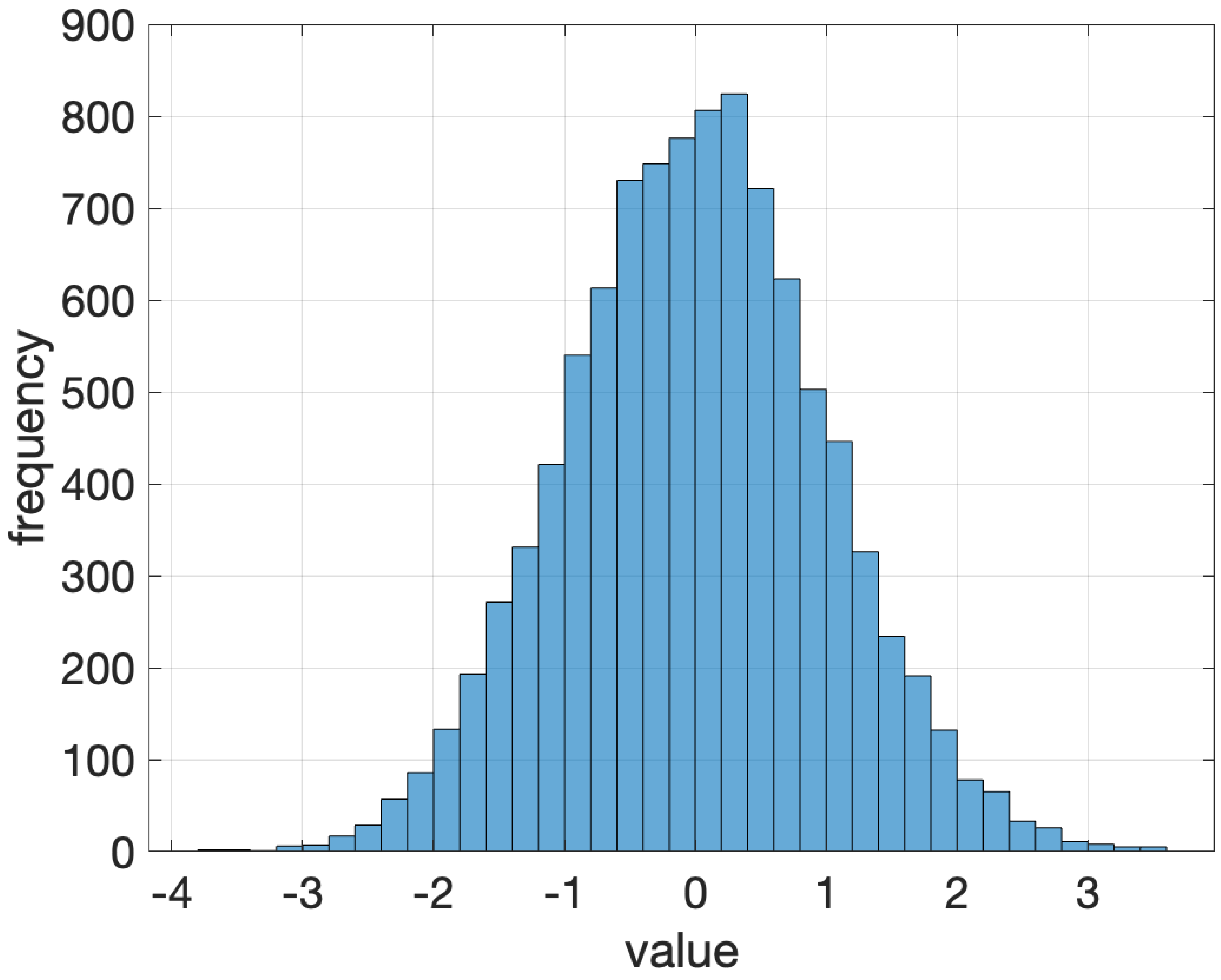
Histograms are a type of bar plot for numeric data that group the data into bins. After you create a `Histogram` object, you can modify aspects of the histogram by changing its property values. This is particularly useful for quickly modifying the properties of the bins or changing the display.

Creation

Syntax

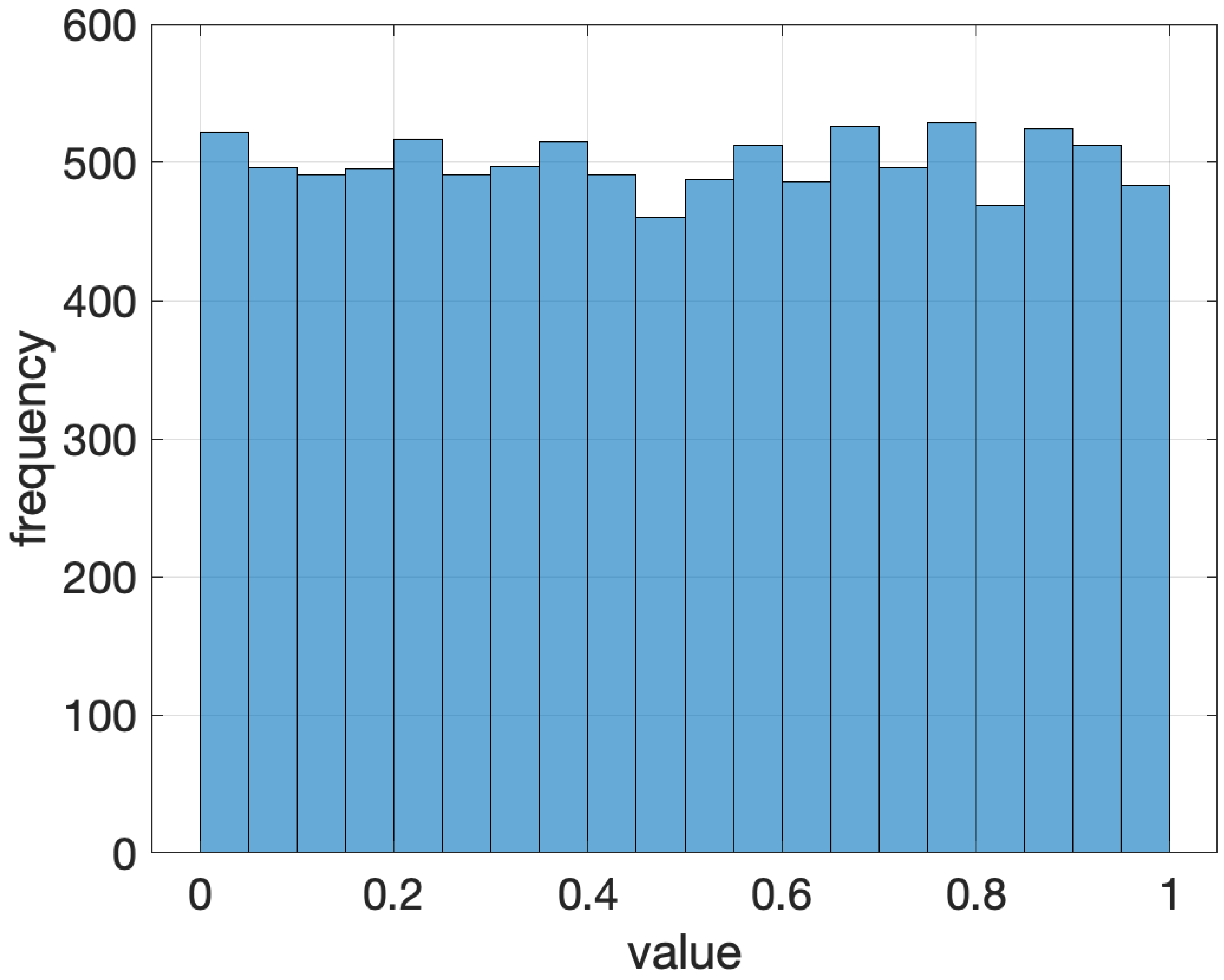
```
histogram(X)
histogram(X,nbins)
histogram(X,edges)
histogram('BinEdges',edges,'BinCounts',counts)
```

Histograms with matlab



data = randn(10000,1);

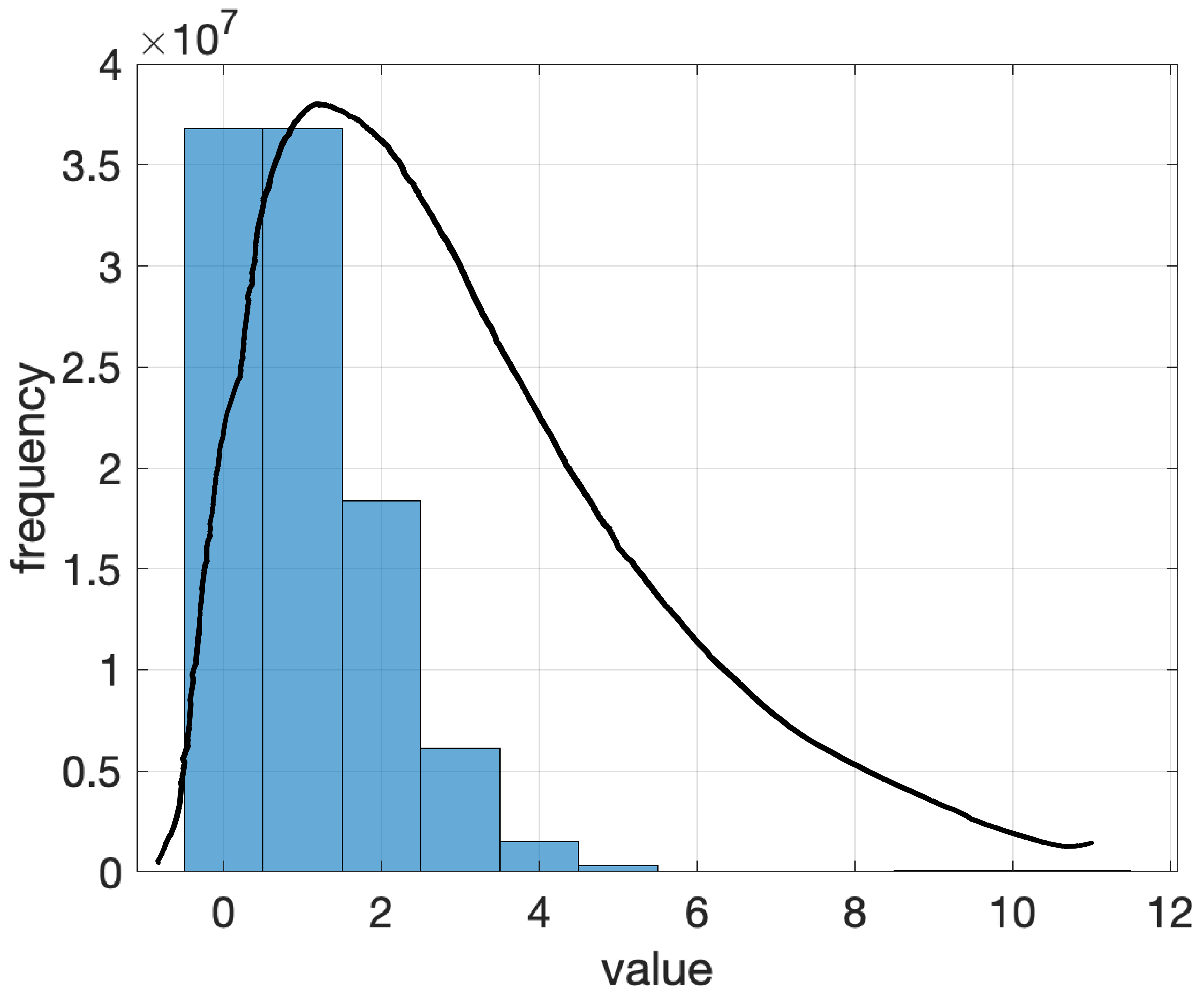
figure; histogram(data);
xlabel(‘value');
ylabel('frequency');



data = rand(10000,1);

figure; histogram(data);
xlabel(‘value');
ylabel('frequency');

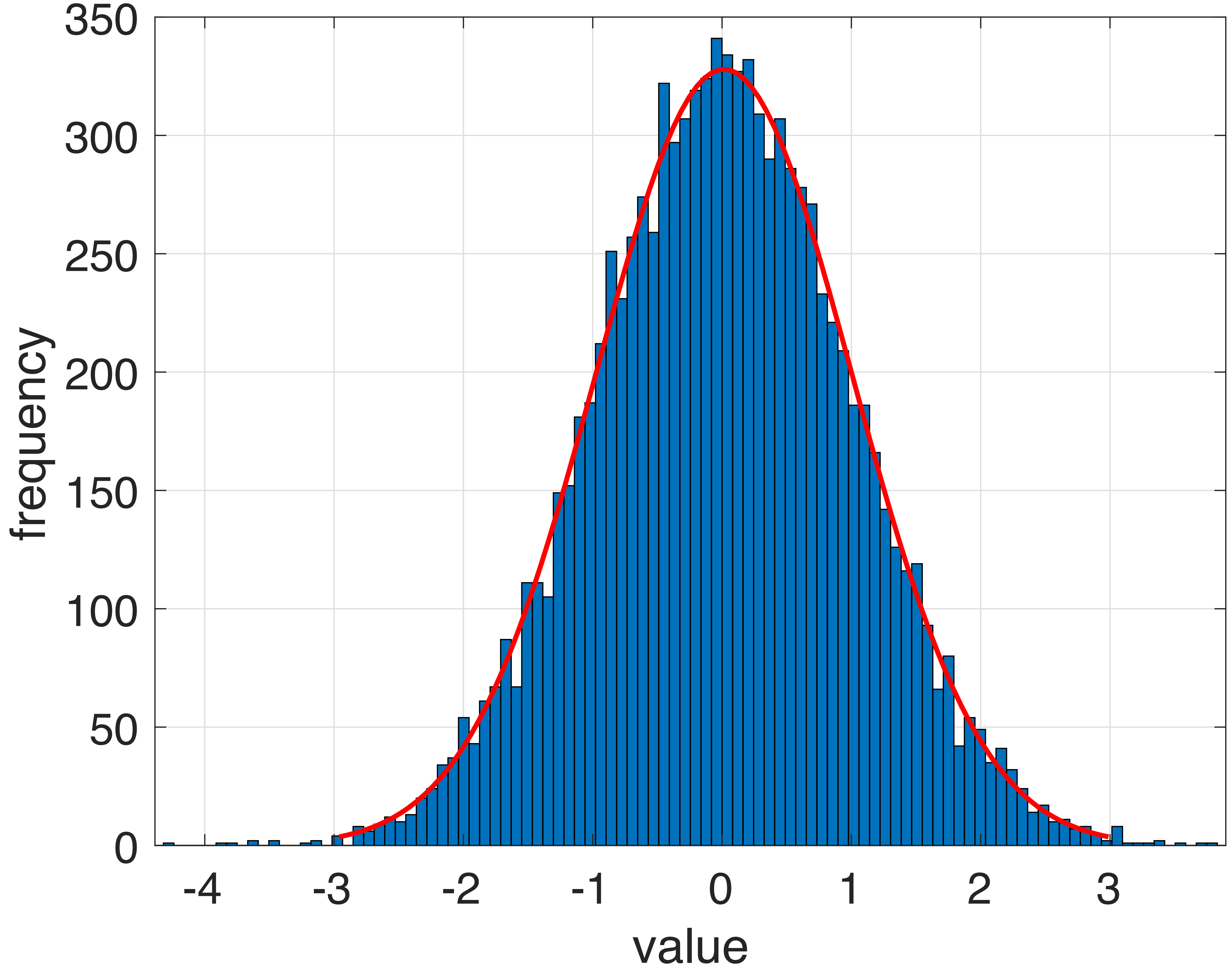
Histograms with matlab



data = poissrnd(1,10000)

figure; histogram(data);
xlabel(‘value');
ylabel('frequency');

Histograms with matlab



data = randn(10000,1);

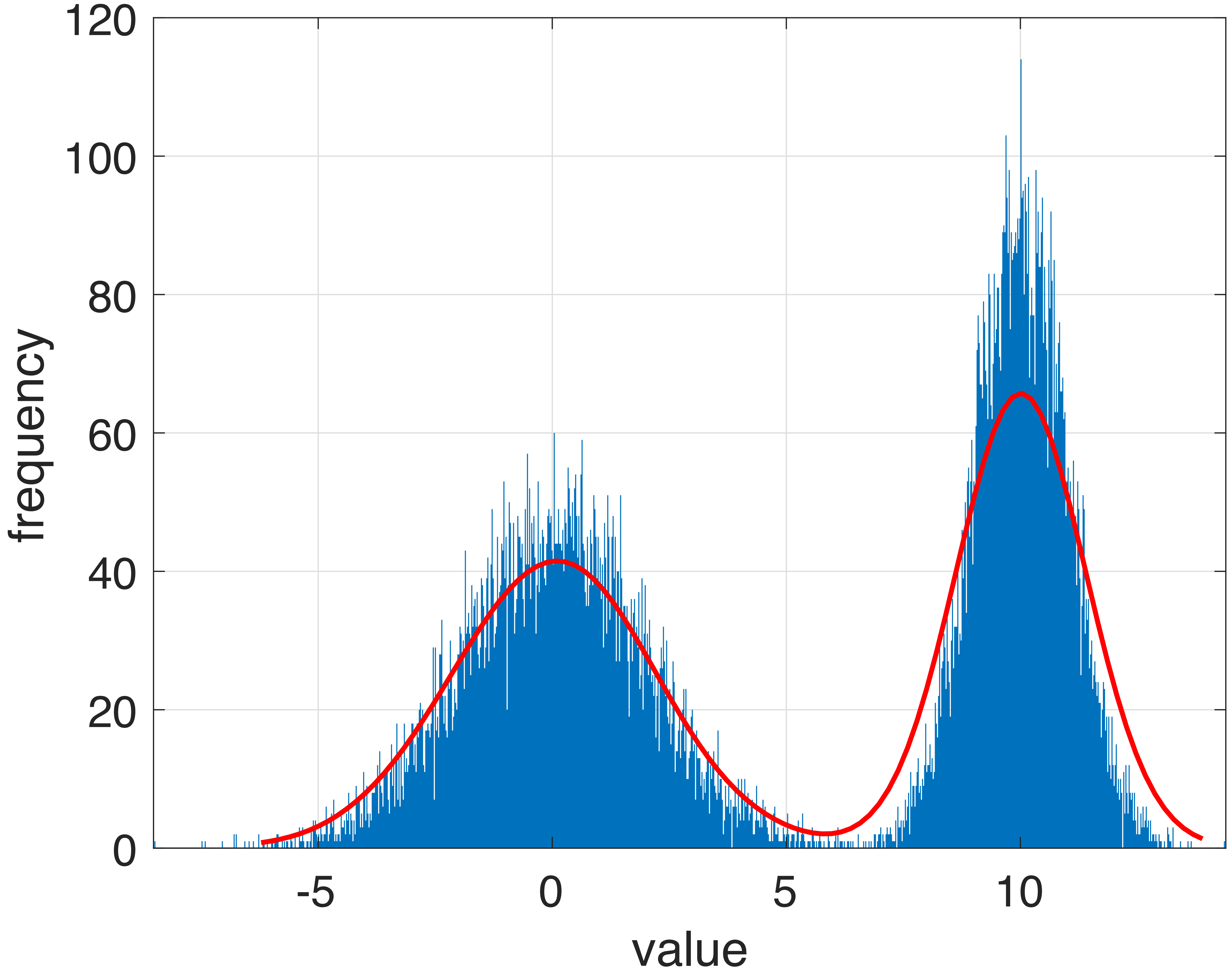
figure; histfit(data);

xlabel(‘value');

ylabel('frequency');

Histograms with matlab

data = [(randn(10000,1) + 10); 2*randn(10000,1)]



figure; histfit(data,1000,'kernel');
xlabel('value');
ylabel('frequency');