Introduction to Bayesian Methods - 1

Edoardo Milotti

Università di Trieste and INFN-Sezione di Trieste

Course webpage:

http://wwwusers.ts.infn.it/~milotti/Didattica/Bayes/Bayes.html

Conditional probabilities and Bayes' Theorem

$$P(AB) = P(A|B)P(B) = P(B|A)P(A)$$

Joint probability and conditional probabilities

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem

Conditional probabilities and Bayes' Theorem

$$P(AB) = P(A|B)P(B) = P(B|A)P(A)$$

Joint probability and conditional probabilities

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Bayes' theorem: a purely logical statement

$$P(H|D) = \frac{P(D|H)}{P(D)}P(H)$$

Bayes' theorem again: now as an inferential statement

Likelihood

Prior distribution

$$P(H|D) = \frac{P(D|H)}{P(D)} P(H)$$

Posterior distribution

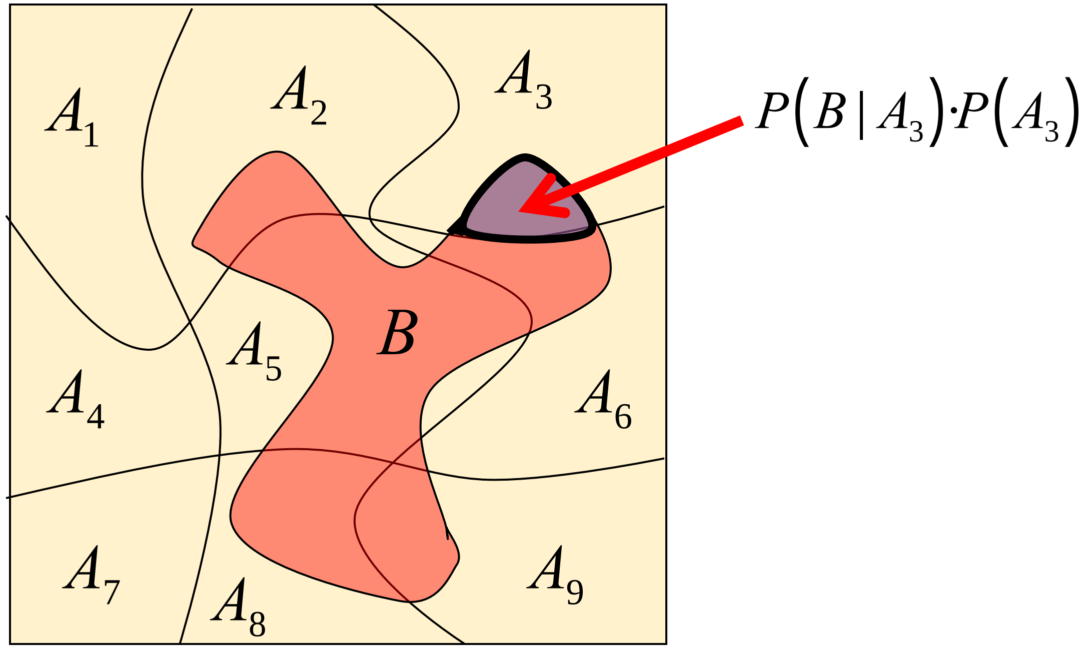
Evidence

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

$$P(A_k \mid B) = \frac{P(B \mid A_k) \cdot P(A_k)}{P(B)} \qquad k = 1, \dots, N$$

if the events A_k are mutually exclusive, and they fill the universe

$$P(B) = \sum_{k=1}^{N} P(B \mid A_k) \cdot P(A_k)$$



$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

+

$$P(B) = \sum_{k=1}^{N} P(B|A_k) \cdot P(A_k)$$

$$P(A_k|B) = \frac{P(B|A_k)}{\sum_{j=1}^{N} P(B|A_j)P(A_j)} P(A_k)$$

Likelihood

Prior distribution

$$P(H_k|D) = \frac{P(D|H_k)}{\sum_j P(D|H_j)P(H_j)} P(H_k)$$

Posterior distribution

Evidence

MAP estimates

A problem of male twins (Efron, 2003)

Pregnant with twins:
fraternal or identical?

Fraternal: 2/3 of all cases

Identical: 1/3 of all cases

What is the probability of identical twins IF both boys in sonogram?

Answer provided by Bayes theorem

$$P(\text{Identical}|\text{Both boys}) = \frac{P(\text{Both boys}|\text{Identical})}{P(\text{Both boys})} P(\text{Identical})$$

$$P(\text{Identical}) = 1/3$$
$$P(\text{Fraternal}) = 2/3$$

Prior probabilities

$$P(\text{Both boys}|\text{Identical}) = 1/2$$
$$P(\text{Both boys}|\text{Fraternal}) = 1/4$$

Conditional probabilities from simple counting argument

$$\begin{aligned} P(\text{Both boys}) &= P(\text{Both boys}|\text{Identical})P(\text{Identical}) \\ &\quad + P(\text{Both boys}|\text{Fraternal})P(\text{Fraternal}) \\ &= (1/2)(1/3) + (1/4)(2/3) = 1/3 \end{aligned}$$

$$\begin{aligned} P(\text{Identical}|\text{Both boys}) &= \frac{P(\text{Both boys}|\text{Identical})}{P(\text{Both boys})} P(\text{Identical}) \\ &= \frac{(1/2)}{(1/3)}(1/3) = 1/2 \end{aligned}$$

A simple application to medical tests (HIV testing)

$$P(\text{positive}|\text{infected}) = 1; \qquad P(\text{positive}|\text{not infected}) = 0.015$$

what is the probability *P(infected|positive)* ?

A common answer is 98.5% ... and it is wrong!

Let's use Bayes' theorem ...

$$P(A_k \mid B) = \frac{P(B \mid A_k) \cdot P(A_k)}{\sum_{k=1}^{N} P(B \mid A_k) \cdot P(A_k)}$$

$$P(\text{infected}|\text{positive}) = \frac{P(\text{positive}|\text{infected}) \times P(\text{infected})}{P(\text{positive}|\text{infected}) \times P(\text{infected}) + P(\text{positive}|\text{not infected}) \times P(\text{not infected})}$$

$$= \left[\frac{P(\text{positive}|\text{infected})}{P(\text{positive}|\text{infected}) \times P(\text{infected}) + P(\text{positive}|\text{not infected}) \times P(\text{not infected})} \right] \times P(\text{infected})$$

The estimate depends on the size of the infected population

i.e., on the probabilities

P(infected) P(not infected)

$$P(\text{infected}|\text{positive}) = \left[\frac{P(\text{positive}|\text{infected})}{P(\text{positive}|\text{infected}) \times P(\text{infected}) + P(\text{positive}|\text{not infected}) \times P(\text{not infected})}\right] \times P(\text{infected})$$

The posterior estimate strongly depends on prior probability

Example: AIDS testing

(data from https://en.wikipedia.org/wiki/List_of_countries_by_HIV/AIDS_adult_prevalence_rate, accessed May 7th 2022)

$$P(\text{infected}|\text{positive}) = \left[\frac{P(\text{positive}|\text{infected})}{P(\text{positive}|\text{infected}) \times P(\text{infected}) + P(\text{positive}|\text{not infected}) \times P(\text{not infected})}\right] \times P(\text{infected})$$

$$P_{\text{Italy}}(\text{infected}|\text{positive}) = \frac{1}{1 \times 0.003 + 0.015 \times 0.997} \times 0.003 \approx 16.7\%$$

$$P_{\text{South Africa}}(\text{infected}|\text{positive}) = \frac{1}{1 \times 0.173 + 0.015 \times 0.827} \times 0.173 \approx 93.3\%$$

The large number of false positives and the small probability of finding a sick person mean that the probability of being infected if positive is not actually very high.

Repeating measurements changes the reference population.

We incorporate a new positive result in a repeated measurement by using the previous posterior as the new prior:

$$P_{\text{Italy}}(\text{infected}|\text{positive}, \text{positive}) = \frac{1}{1 \times 0.167 + 0.015 \times 0.833} \times 0.167 \approx 93.0\%$$

$$P_{\text{South Africa}}(\text{infected}|\text{positive}, \text{positive}) = \frac{1}{1 \times 0.933 + 0.015 \times 0.067} \times 0.933 \approx 99.9\%$$

The first test changes the reference population, and the second test, if positive, gives a significant result.

Comparing hypotheses

$$P(H_k | D, I) = \frac{P(D | H_k, I)}{P(D | I)} \cdot P(H_k | I)$$

Odds ratio

$$\frac{P(H_k | D, I)}{P(H_n | D, I)} = \left(\frac{P(D | H_k, I)}{P(D | H_n, I)} \right) \cdot \left(\frac{P(H_k | I)}{P(H_n | I)} \right)$$

Bayes' factor

When prior probabilities are the same (equally probable hypotheses), the posterior probability ratio depends only on the Bayes' factor:

Odds ratio

$$\frac{P(H_k|D,I)}{P(H_n|D,I)} = \frac{P(D|H_k,I)}{P(D|H_n,I)} \times \frac{P(H_k|I)}{P(H_n|I)} = \frac{P(D|H_k,I)}{P(D|H_n,I)}$$

Bayes' factor

Bayes' factor

Uniform priors

From discrete sets of hypothesis to the continuum.
The Bayes' theorem in the context of parameter estimation.

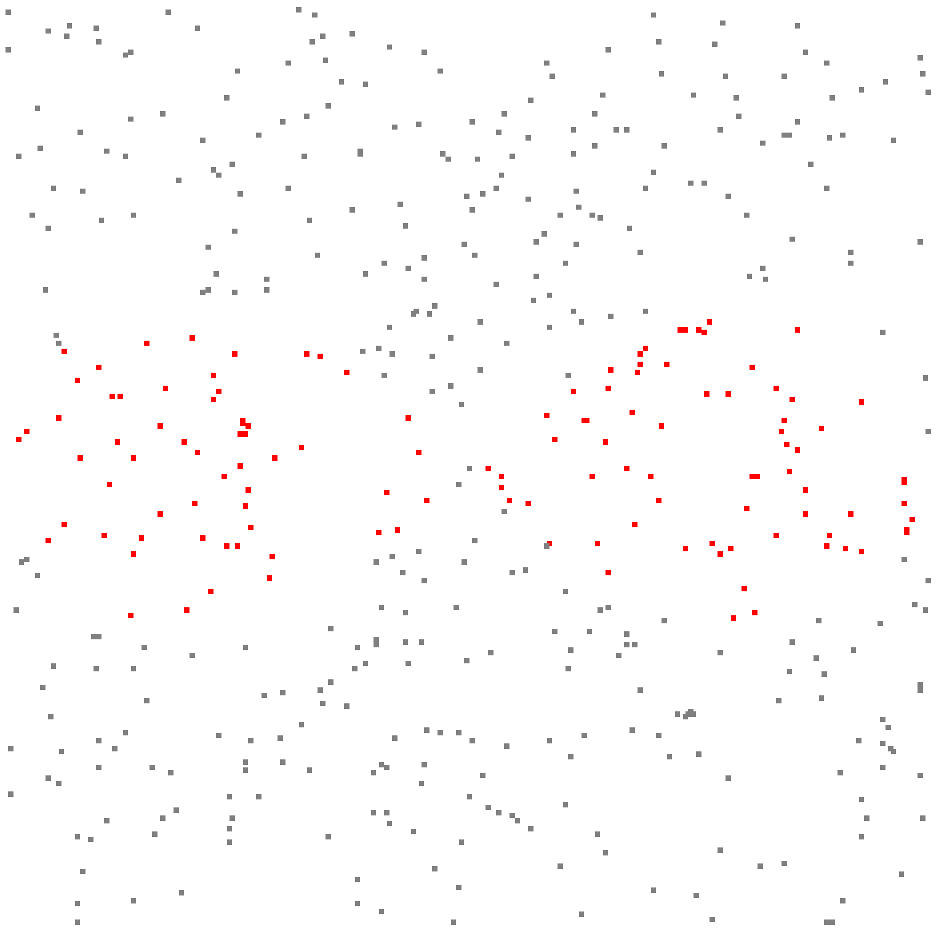
$$P(H_k \mid D, I) = \frac{P(D \mid H_k, I)}{P(D \mid I)} \cdot P(H_k \mid I) = \frac{P(D \mid H_k, I)}{\sum_{k=1}^{N} P(D \mid H_k, I) \cdot P(H_k \mid I)} \cdot P(H_k \mid I)$$

$$p(\boldsymbol{\theta} | D, I) = \frac{P(D | \boldsymbol{\theta}, I)}{\int_{\Theta} P(D | \boldsymbol{\theta}', I)\, p(\boldsymbol{\theta}' | I)\, d\boldsymbol{\theta}'} \times p(\boldsymbol{\theta} | I)$$

What if we “measure” a mathematical constant instead of a physical parameter?

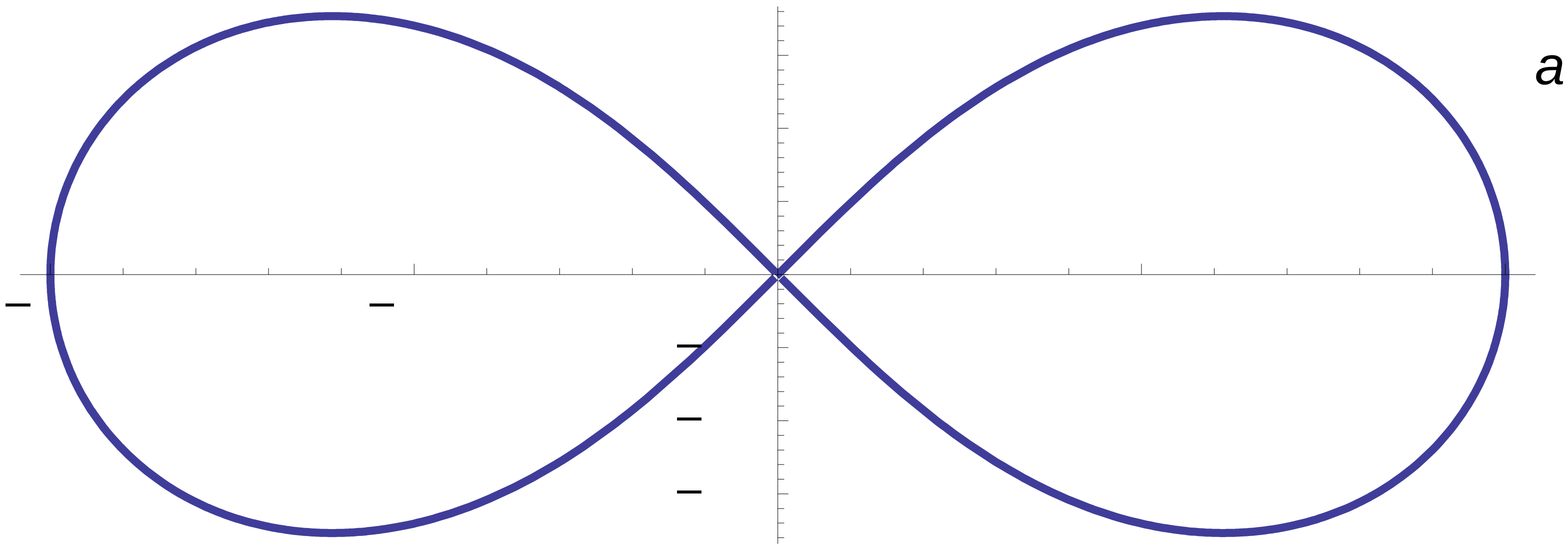
Example:

area of Bernoulli’s lemniscate obtained with a Monte Carlo simulation.



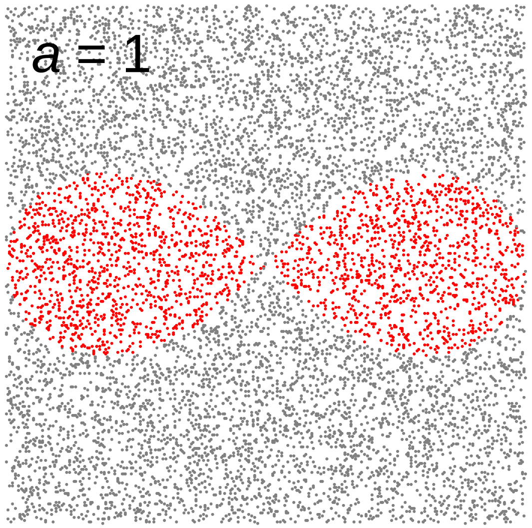
Parametric equation of Bernoulli's lemniscate

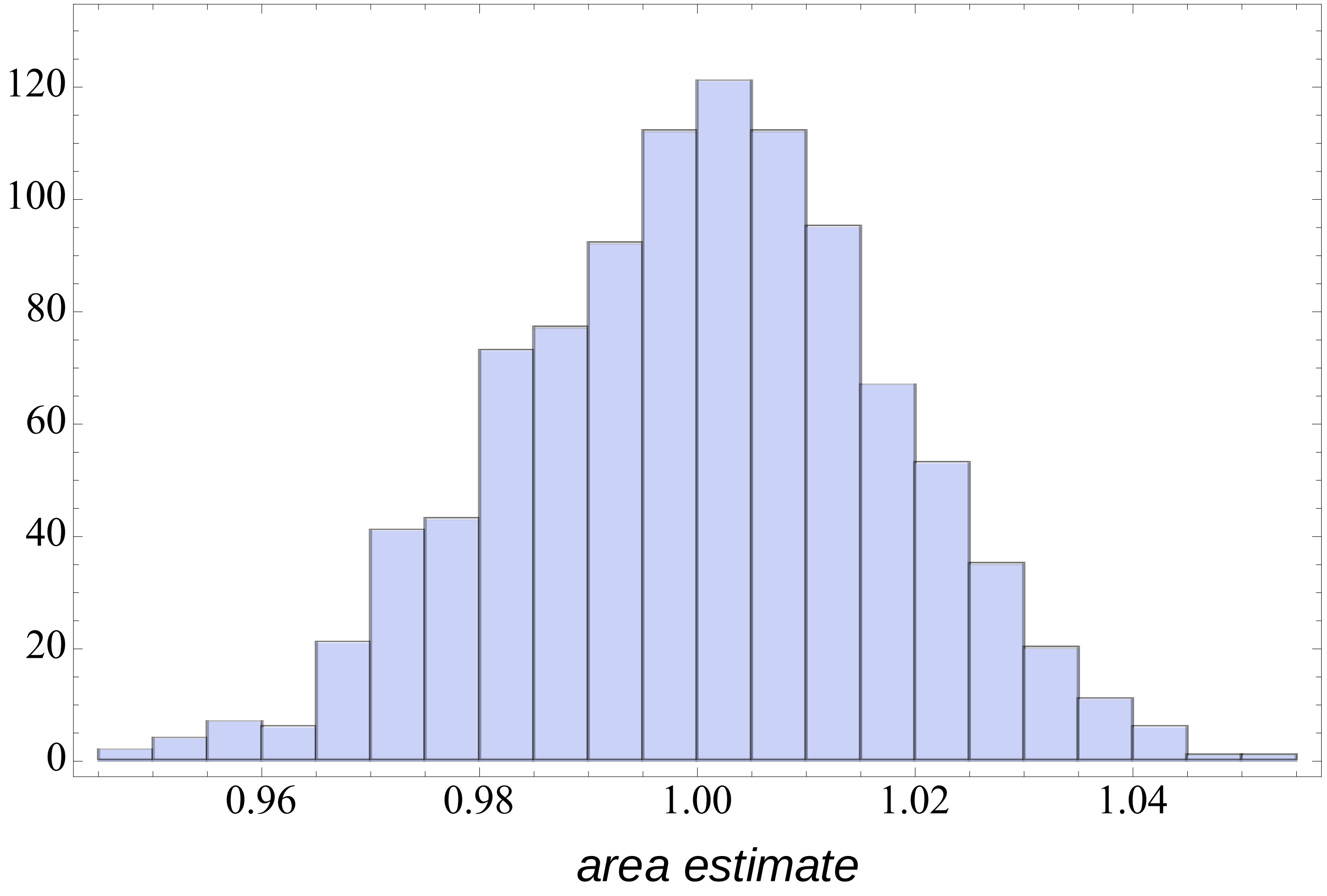
$$r = a\sqrt{\cos 2\theta}$$



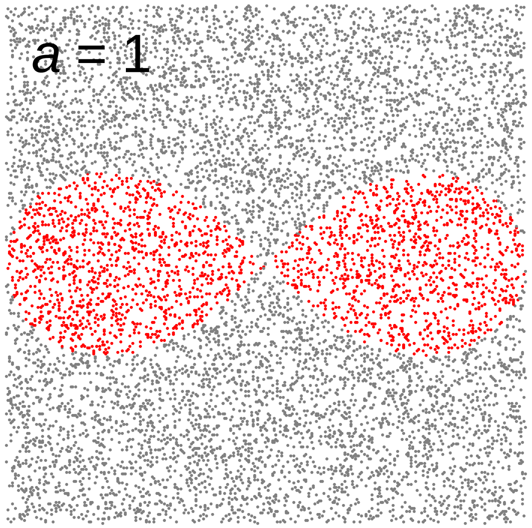
What is its area?

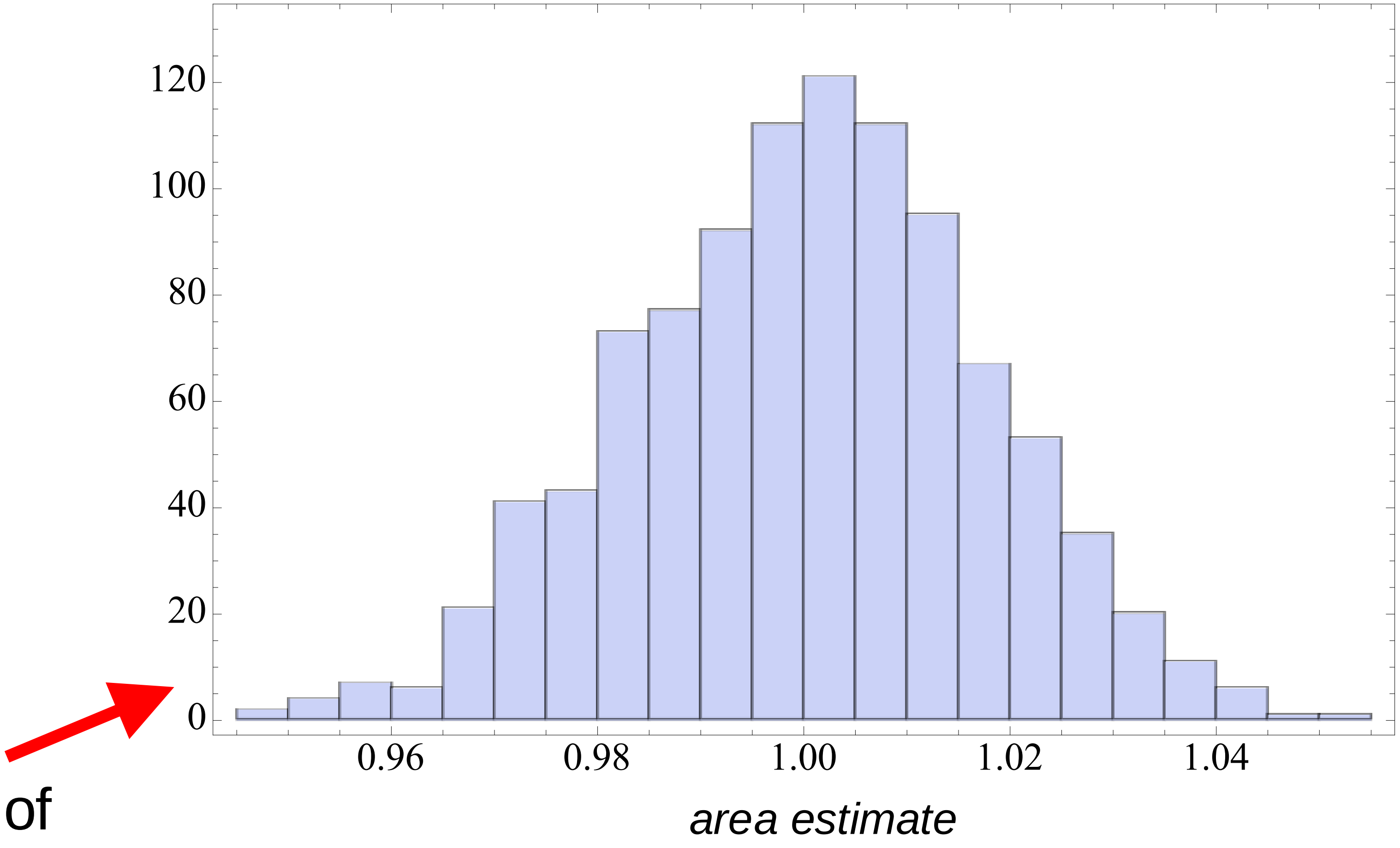
Empirical Monte Carlo distribution of the area estimate





Empirical Monte Carlo distribution of the area estimate

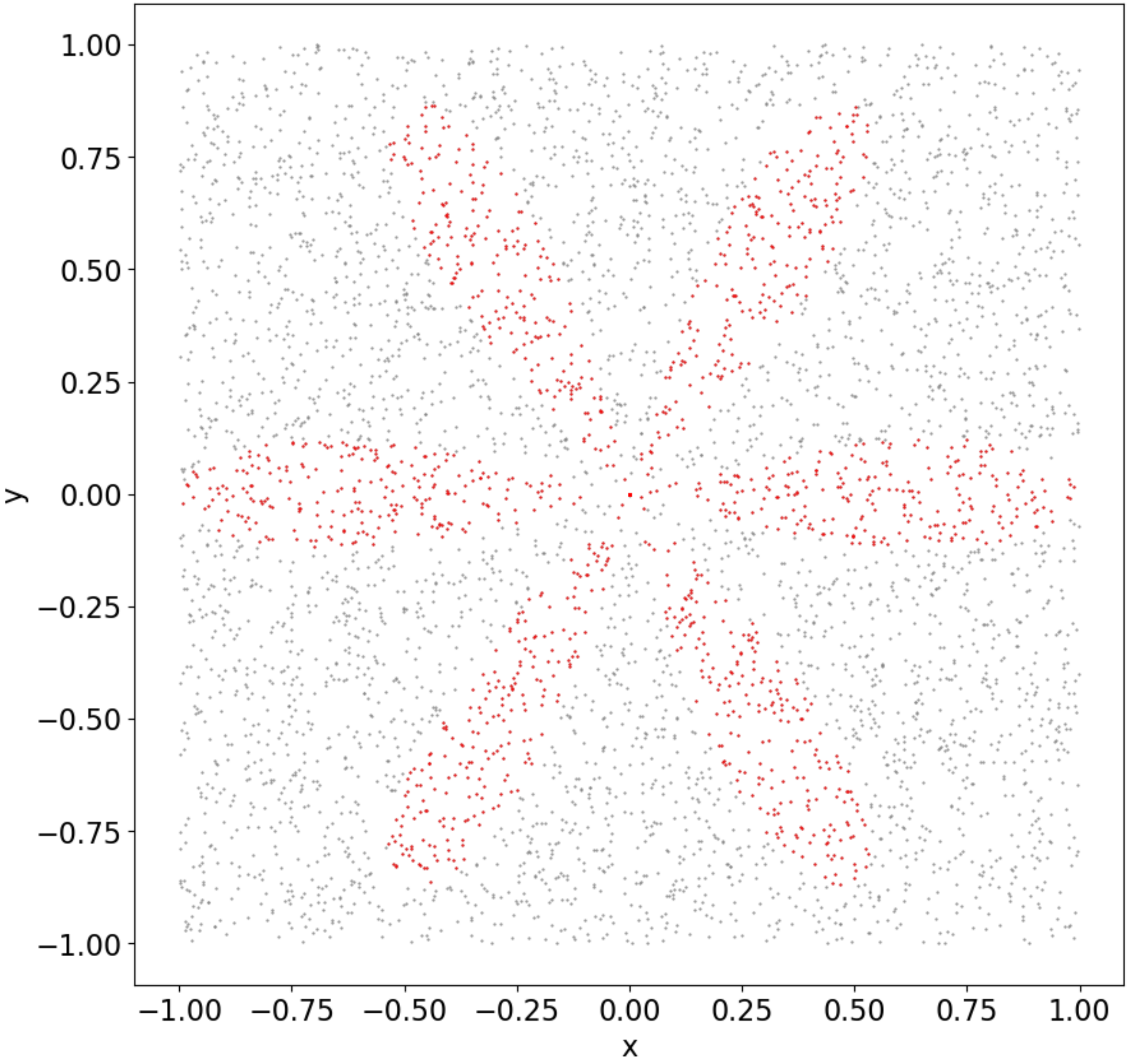




a probability distribution of
a mathematical constant???

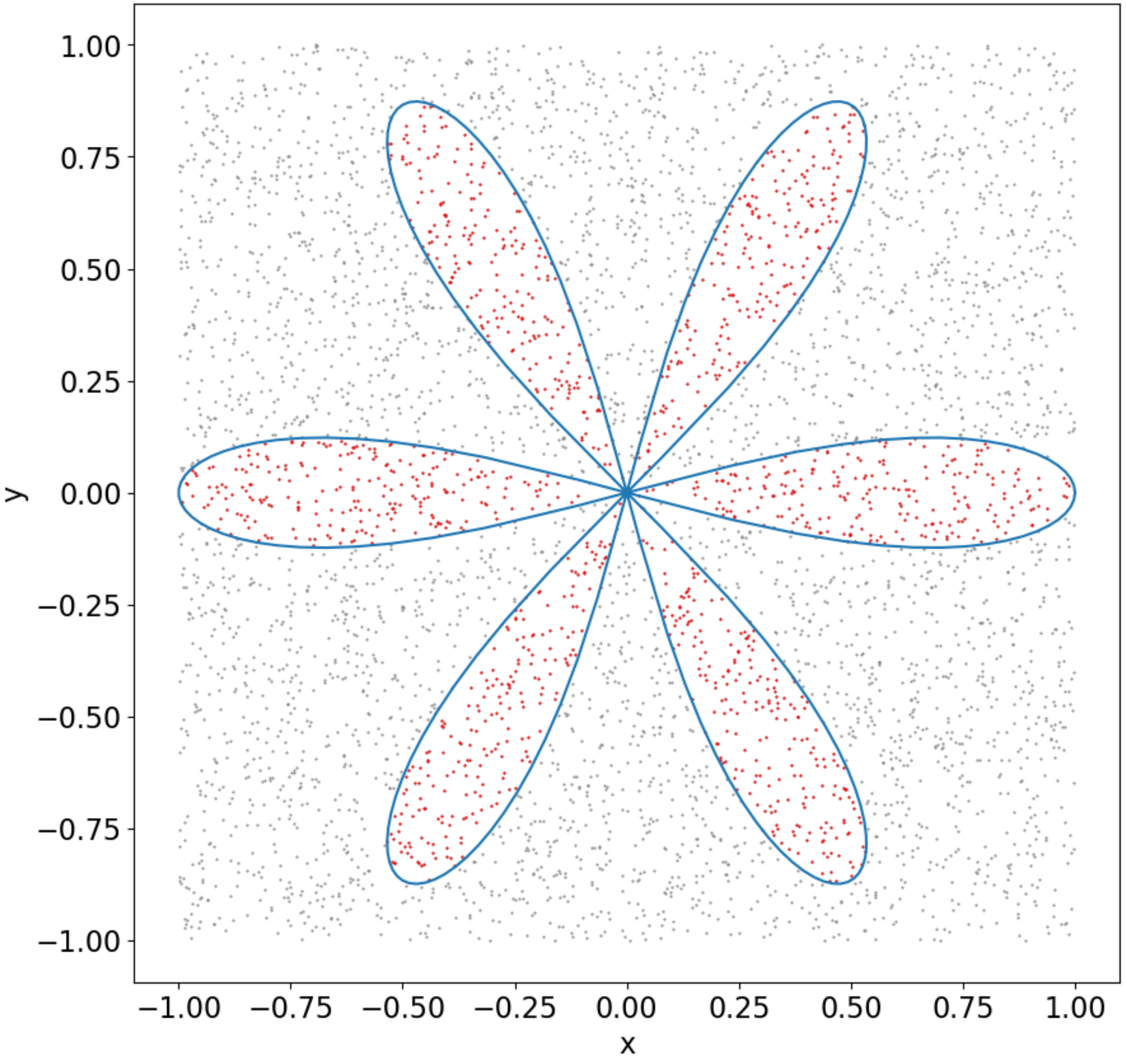
Now, try yourself with a rhodonea

$$r^2(\theta) = a^2 \cos(6\theta)$$



Now, try yourself with a rhodonea

$$r^2(\theta) = a^2 \cos(6\theta)$$



AMERICAN JOURNAL of PHYSICS

A Journal Devoted to the Instructional and Cultural Aspects of Physical Science

VOLUME 14, NUMBER 1 JANUARY–FEBRUARY, 1946

Probability, Frequency and Reasonable Expectation

R. T. Cox
The Johns Hopkins University, Baltimore 18, Maryland

See also R T Cox, *The Algebra of Probable Inference,*
The John Hopkins Press (Baltimore, 1961)
https://bayes.wustl.edu/Manual/cox-algebra.pdf

Boolean algebra (symbolic logic)

a, **b**, **c** ... propositions (true or false)

Basic operations

OR: **a** v **b**

AND: **a**·**b**

NOT: ~**a**

Truth tables

a	b	a v b
T	T	T
T	F	T
F	T	T
F	F	F

a	b	a · b
T	T	T
T	F	F
F	T	F
F	F	F

a	~a
T	F
F	T

see handout