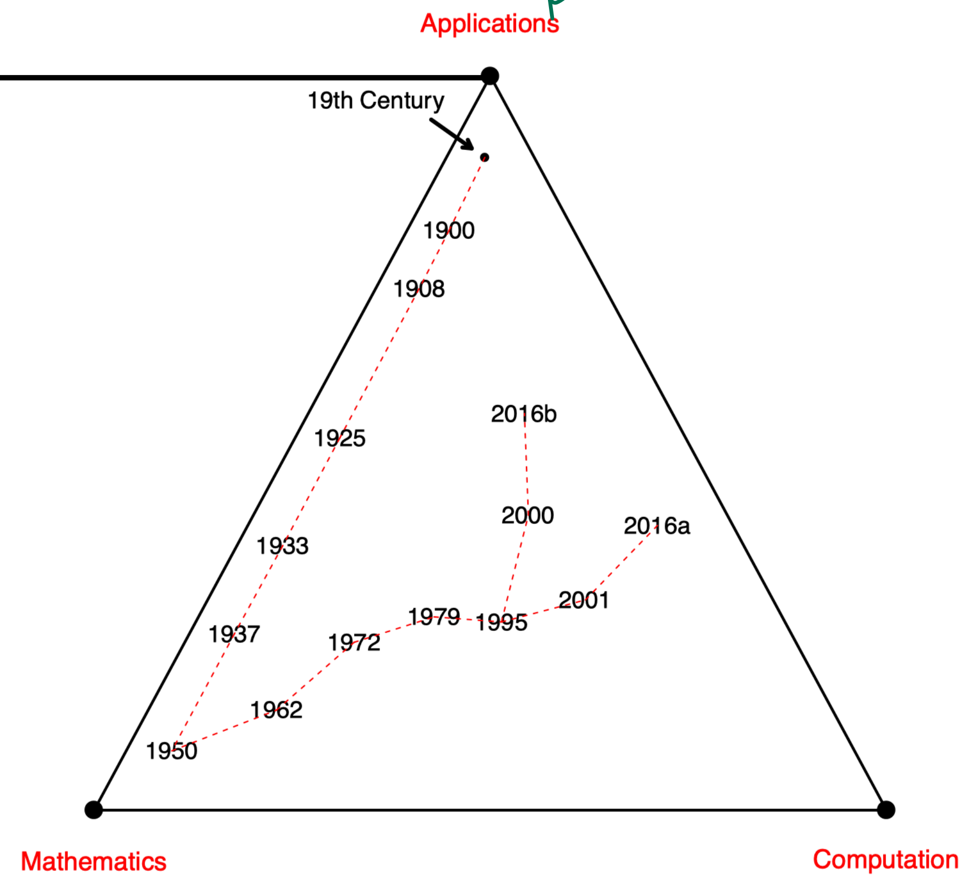
۶. احتمال و اندکی آمار

χ^2

قرن ۲۰ام



Pierre de Fermat (1607-1665)

Monsieur,

If I undertake to make a point with a single die in eight throws, and if we agree after the money is put at stake, that I shall not cast the first throw, it is necessary by my theory that I take 1/6 of the total sum to he impartial because of the aforesaid first throw.

Blaise Pascal (1623-1669)

Monsieur,

1. Impatience has seized me as well as it has you, and although I am still abed, I cannot refrain from telling you that I received your letter in regard to the problem of the points 1 yesterday evening from the hands of M. Carcavi, and that I admire it more than I can tell you. I do not have the leisure to write at length, but, in a word, you have found the two divisions of the points and of the dice with perfect justice. I am thoroughly satisfied as I can no longer doubt that I was wrong, seeing the admirable accord in which I find myself with you.

7. i have no time to send you the proof of a difficult point which astonished M. (de M´er´e) so greatly, for he has ability but he is not a geometer (which is, as you know, a great defect) and he does not even comprehend that a mathematical line is infinitely divisible and he is firmly convinced that it is composed of a finite number of points. I have never been able to get him out of it. If you could do so, it would make him perfect.

He tells me then that he has found an error in the numbers for this reason .

If one undertakes to throw a six with a die, the advantage of undertaking to do it in 4 is as 671 is to 625.

If one undertakes to throw double sixes with two dice the disadvantage of the undertaking is 24.

But nonetheless, 24 is to 36 (which is the number of faces of two dice) 2 as 4 is to 6 (which is the number of faces of one die).

This is what was his great scandal which made him say haughtily that the theorems were not consistent and that arithmetic was demented. But you will easily see the reason by the principles which you have.

Monsieur,

1. Our interchange of blows still continues, and I am well pleased that our thoughts are in such complete adjustment as it seems since they have taken the same direction and followed the same road.

3. For the rest, there is nothing that I will not write you in the future with all frankness. Meditate however, if you find it convenient, on this theorem: The squared powers of 2 augmented by unity 11 are always prime numbers. [That is,] The square of 2 augmented by unity makes 5 which is a prime number;
The square of the square makes 16 which, when unity is added makes 17, a prime number;
The square of 16 makes 256 which, when unity is added, makes 257, a prime number;
The square of 256 makes 65536 which, when unity is added, makes 65537, a prime number;
and so to infinity.
This is a property whose truth I will answer to you. The proof of it is very difficult and I assure you that I have not yet been able to find it fully. I shall not set it for you to find unless I come to the end of it.

Leonhard Euler (1707-1783)

No

A game of chance

اتفاق غیرقطعی

- A game of chance
- Sample space
- Event

A رخداد

رخ داد

$s \in A$

Outcome

- $\Omega = \{s_1, s_2 \ldots, s_n\}$
- Mutually exclusive and homogenous
- $\boxed{A \subset \Omega}$

پدیده سکه $\{$ خط , شیر $\}$

تاس $\{۱, ۲, ۳, ۴, ۵, ۶\}$

P(شیر آمدن) $= \frac{1}{2}$

naive def

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

Bayesian قدرت باور $\frac{1}{2}$

Frequentist احتمال تکرارهای بی‌نهایت $\lim_{n \to \infty} \frac{\text{اتفاق}}{\text{کل}}$

Axioms of probability

- non-zero
- Certainty
- Disjoint events

|A|

اشتراک ندارن

- $\mathbb{P}(A) \geq 0$
- $\boxed{\mathbb{P}(\Omega)} = 1$
- $\mathbb{P}(A) + \mathbb{P}(B)$

Results of axioms of probability

- Complement

$\mathbb{P}(A) = \mathbb{P}(A^c)$

$\underbrace{A + A^c}_{\text{disjoint}} = \Omega$

$\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega)$ 1

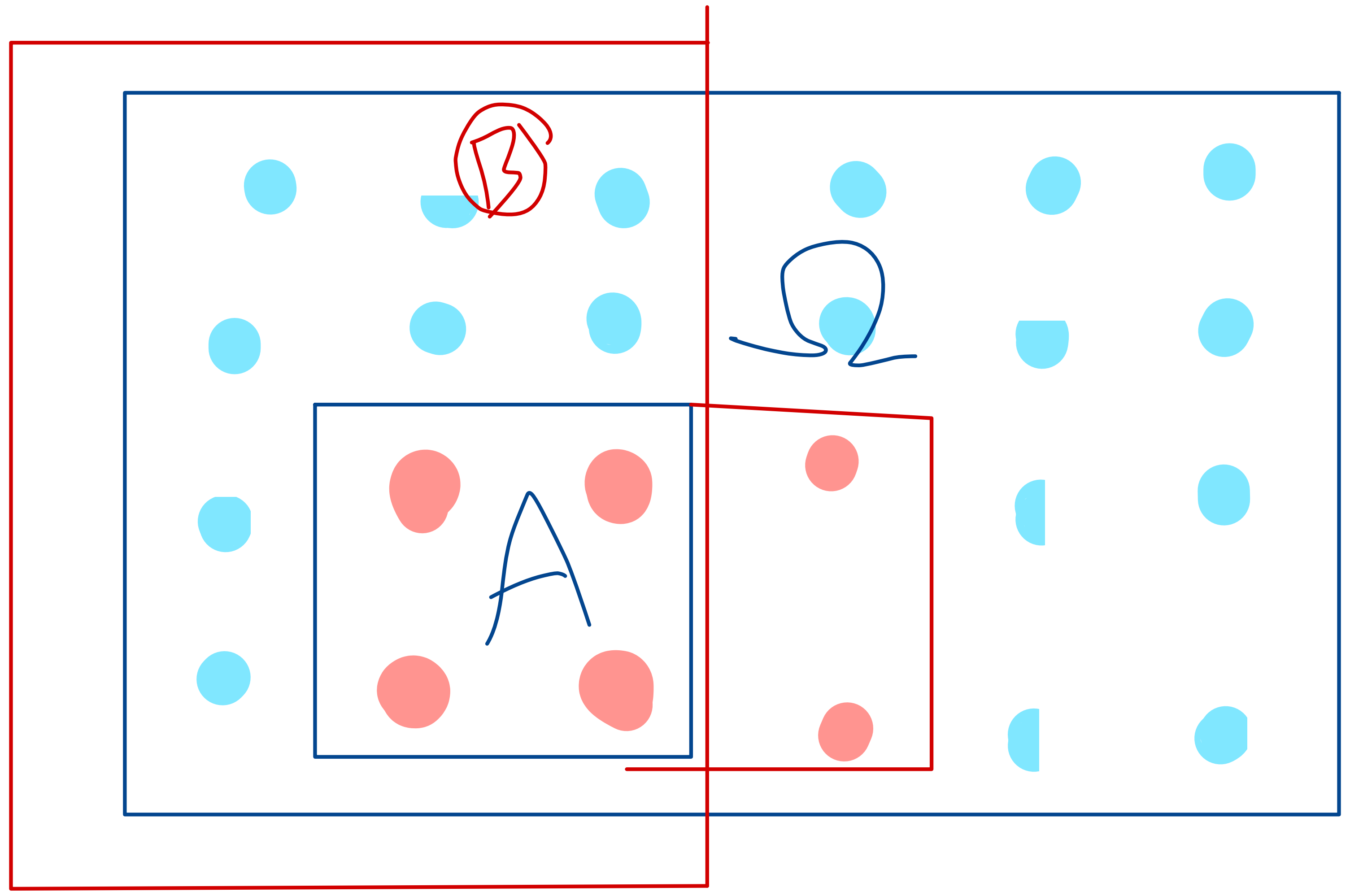
- $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(A) \leq 1$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

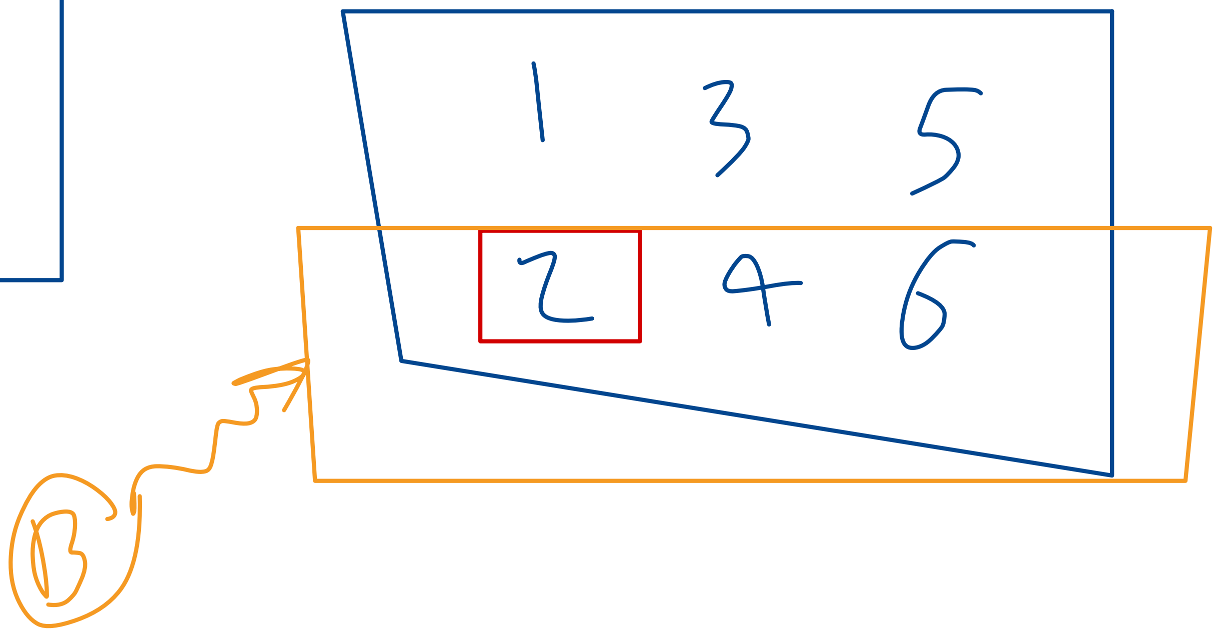
Conditional probability

- A certain event has happened
- What is the probability in this new space

- $\Omega_{new} = B$

$$A = \frac{|A \cap B|}{|B|} \qquad \frac{\text{حالات مطلوب}}{\text{حالات}}$$

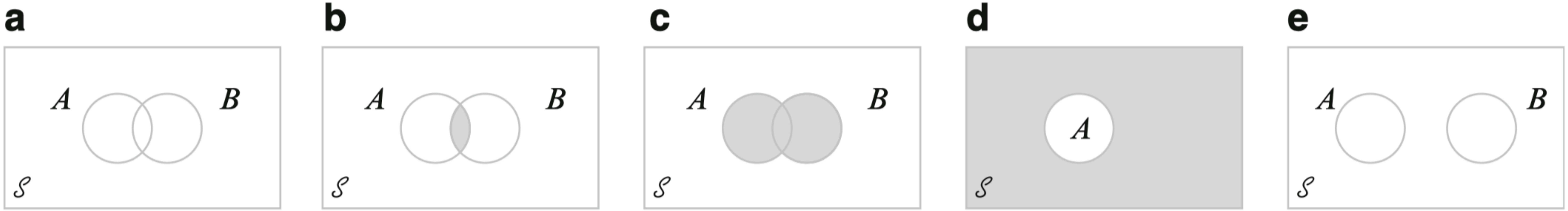




$$\mathbb{P}(2) = \frac{|A|}{|\Omega|} \qquad \frac{1}{6}$$

$$\mathbb{P}(2) = \frac{|A \cap B|}{|B|} \qquad \frac{1}{3}$$

Conditional probability



Venn diagram of events A and B

Shaded region is $A \cap B$

Shaded region is $A \cup B$

Shaded region is A'

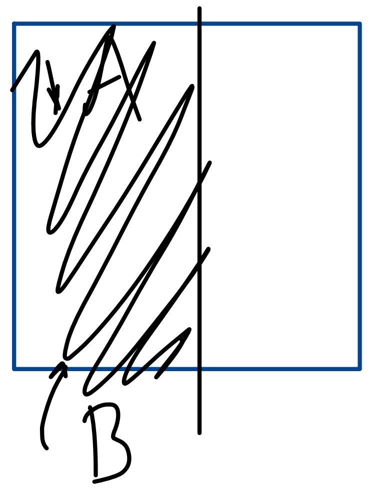
Mutually exclusive events

احتمال شرطی

مشروط اتفاق

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

Thomas Bayes (1701-1761)



$\mathbb{P}(A) = \frac{1}{2}$

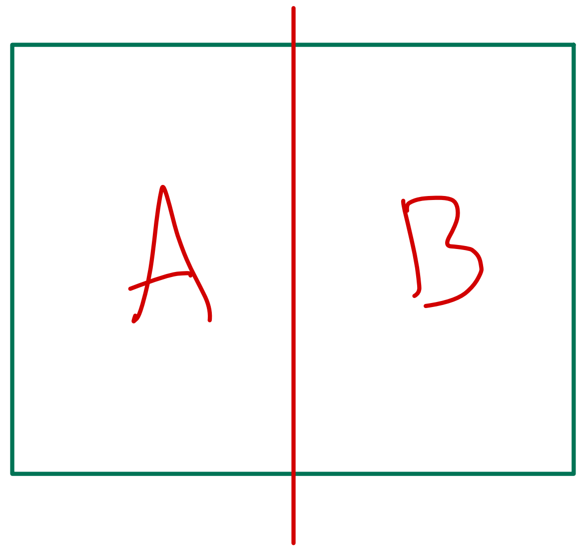
$\mathbb{P}(A|B) = 1$

تغییر

وابسته

شرطی

ناسازگار



$\mathbb{P}(A) = \frac{1}{2}$

$\mathbb{P}(A|B) = 0$

عوض

وابسته

استقلال A و B یعنی

دانستن اطلاعات در مورد B نظر من

در مورد A تغییر نکند

خب لحظه چی؟

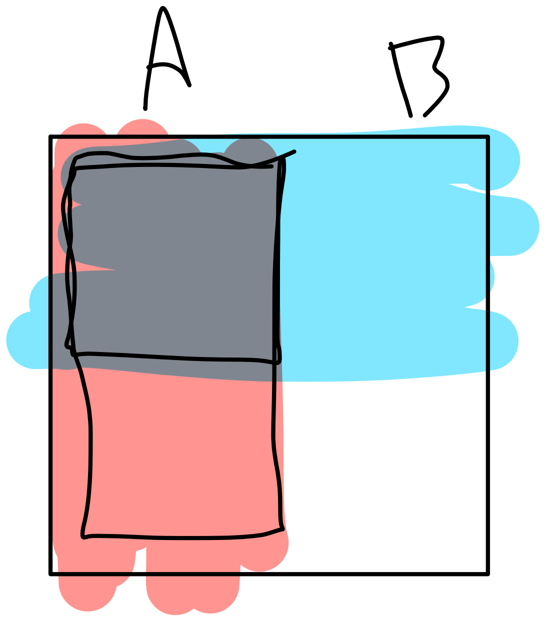
مهم!

مستقل

تغییر نکرد

$P(A) = \frac{1}{2}$

$\mathbb{P}(A|B) = \frac{\frac{1}{4}}{\frac{1}{2}}$



LII. *An Eſſay towards ſolving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Dear Sir,

Read Dec. 23, 1763.

I Now ſend you an eſſay which I have found among the papers of our deceaſed friend Mr. Bayes, and which, in my opinion, has great merit, and well deſerves to be preſerved. Experimental philoſophy, you will find, is nearly intereſted in the ſubject of it; and on this account there ſeems to be particular reaſon for thinking that a communication of it to the Royal Society cannot be improper.

Conditional probability

- Multiplication

- $\mathbb{P}(A \cap B) = \mathbb{P}(A \mid B) \cdot \mathbb{P}(B)$

$$\mathbb{P}(B|A)\,\mathbb{P}(A) = \mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(A \cap B) \rightsquigarrow \mathbb{P}(A|B) = \frac{|A \cap B| / |\Omega|}{|B| / |\Omega|} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

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

Bayes' theorom

- Symmetry

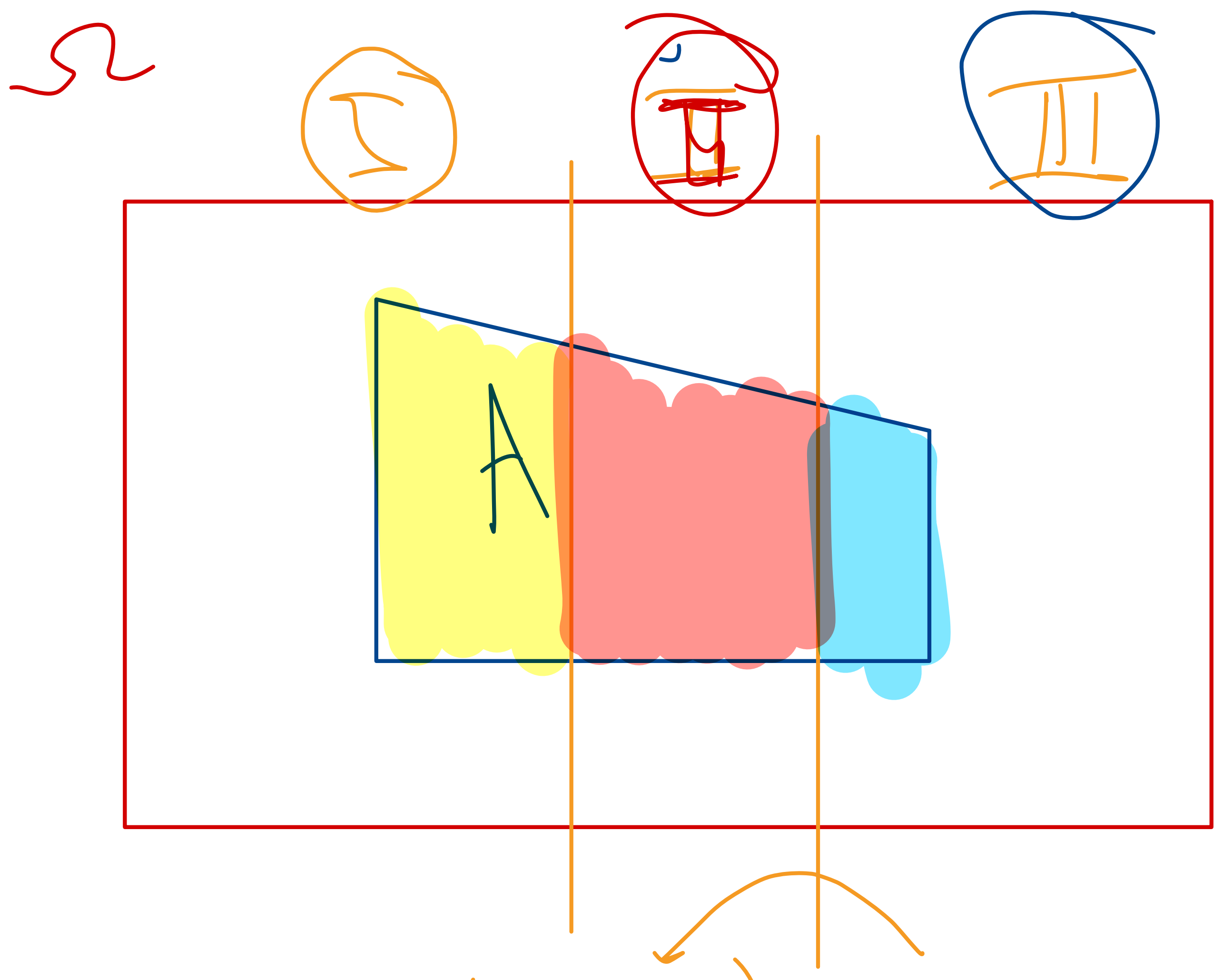
- $\mathbb{P}(A \cap B) = \mathbb{P}(B \cap A)$
- $\mathbb{P}(A \mid B) \cdot \mathbb{P}(B) = \mathbb{P}(B \mid A) \cdot \mathbb{P}(A)$

Bayes

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

Law of total probability

- Mutually exclusive and exhaustive chunks of the sample space



$\mathbb{P}(A) = \mathbb{P}(A \cap \mathrm{I}) + \mathbb{P}(A \cap \mathrm{II}) + \mathbb{P}(A \cap \mathrm{III})$

$\mathbb{P}(A \cap \mathrm{I}) = \mathbb{P}(A|\mathrm{I}) \times \mathbb{P}(\mathrm{I})$

$\mathbb{P}(A \cap \mathrm{II}) = \mathbb{P}(A|\mathrm{II}) \times \mathbb{P}(\mathrm{II})$

$\mathbb{P}(A \cap \mathrm{III}) = \mathbb{P}(A|\mathrm{III}) \times \mathbb{P}(\mathrm{III})$

- $A_1 \cup A_2 \cup \ldots \cup A_k = \Omega$
- $\mathbb{P}(B) = \mathbb{P}(B \cap A_1) + \ldots + \mathbb{P}(B \cap A_k)$

$\mathbb{P}(A)$

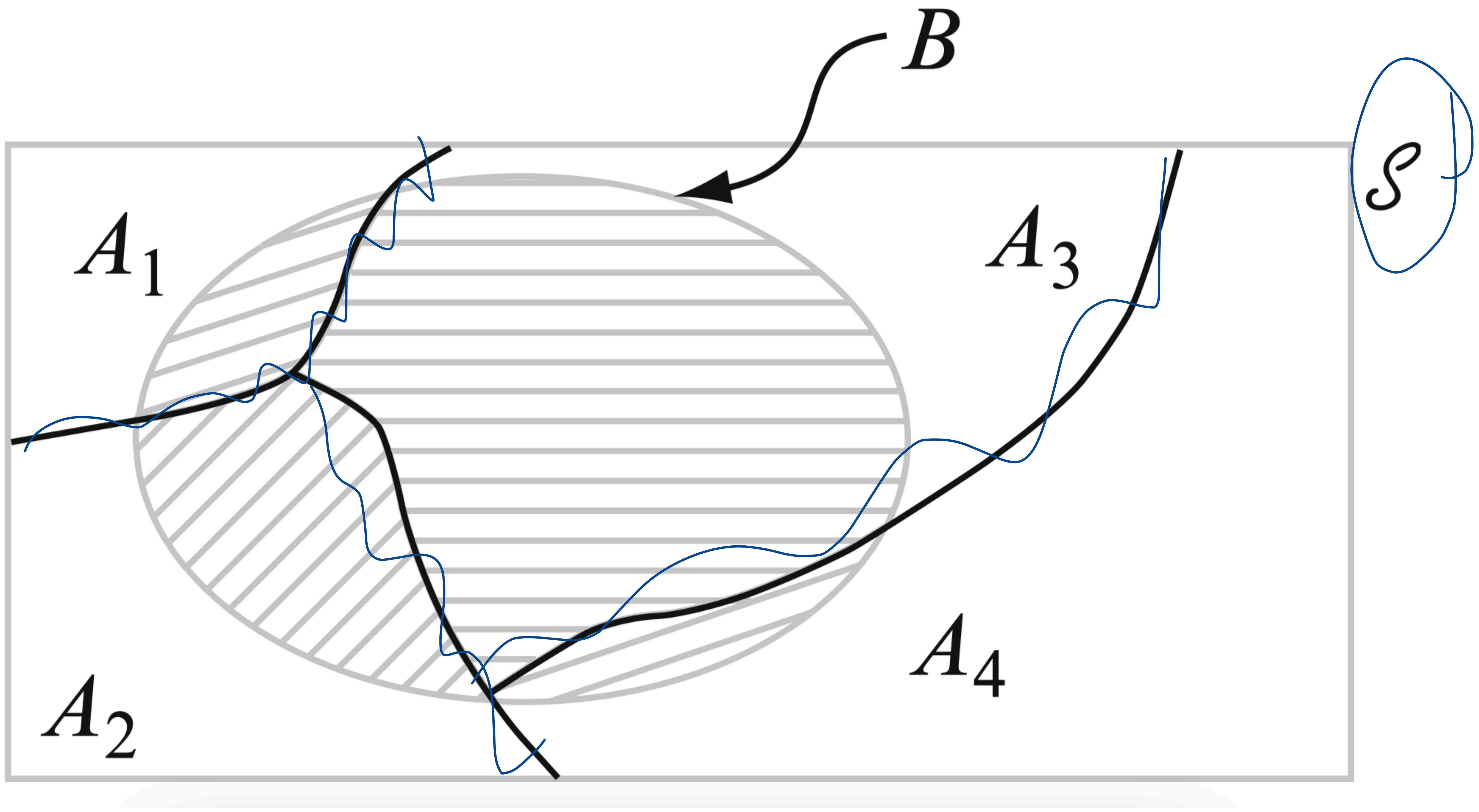
شطرنج

قوی	,1	,3
متوسط	,5	,6
ضعیف	,4	,9

(برنده)

,3 × ,1
+ ,6 × ,5
+ ,9 × ,4

Law of total probability



$$\mathbb{P}(B) = \sum_{1}^{k} \mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)$$

Independence

- Definition of independence
- Multiplication rule

- $\mathbb{P}(A \mid B) = \mathbb{P}(A)$
- $\boxed{\mathbb{P}(A \cap B)} = \mathbb{P}(A) \cdot \mathbb{P}(B)$

$\mathbb{P}(A \mid B)$

How to systematically update our beliefs

- Prior
- Likelihood
- Evidence
- Posterior

- $\mathbb{P}(H)$
- $\mathbb{P}(D \mid H)$
- $\mathbb{P}(D)$
- $\mathbb{P}(H \mid D)$

H $\quad \mathbb{P}(H)$

آزمایش

دیتا D

$$\mathbb{P}(H \mid D) = \frac{\mathbb{P}(D \mid H)\,\mathbb{P}(H)}{\mathbb{P}(D)}$$

$\mathbb{P}(D \mid H)\mathbb{P}(H)$ $\quad H$

$\mathbb{P}(D \mid \neg H)\mathbb{P}(\neg H)$ $\quad \neg H$

$$\mathbb{P}(H_1 \mid D) = \frac{\mathbb{P}(D \mid H_1) \cdot \mathbb{P}(H_1)}{\sum_1^k \mathbb{P}(D \mid H_i) \cdot \mathbb{P}(H_i)}$$

$H \qquad \neg H$

Random variable

تناظر

تابع

متغیر تصادفی

شیر 1

خط 0

سکه

- A mapping from sample space to real number

- $\Omega \to \mathbb{R}$

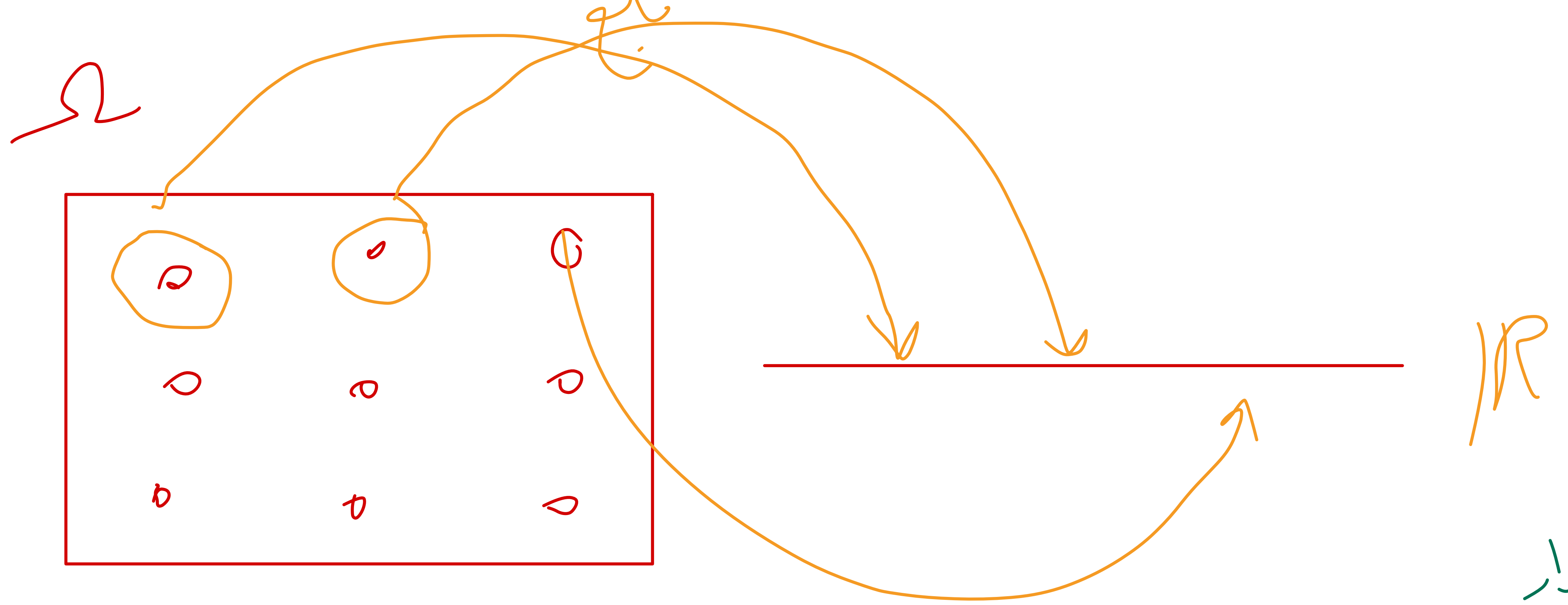
$\{ , , , \}$

X

$x =$ یک مقدار خاص

$X = 1$

Ω خط شیر



$\mathbb{P}(X=0) = \mathbb{P}($سکه خط$) = \frac{1}{2}$

$\mathbb{P}(X=1) = \mathbb{P}($سکه شیر$) = \frac{1}{2}$

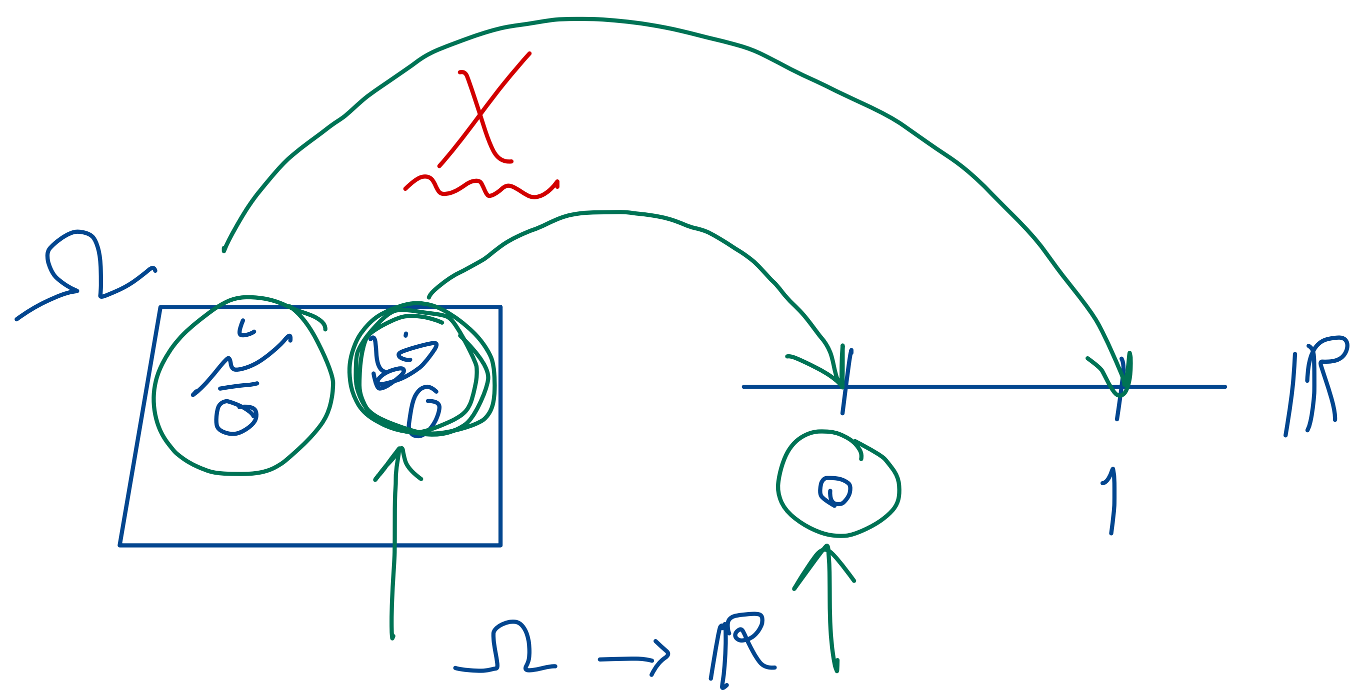
مقدار رخداد

$X \begin{cases} 0 & \text{خط آمدن} \\ 1 & \text{شیر آمدن} \end{cases}$

$\boxed{X} = 0$ = سکه خط امد

$X = 1$ = سکه شیر امد

$\mathcal{X} = \{0,1\}$

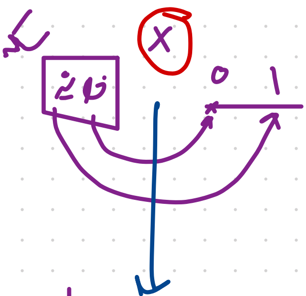


$x \rightarrow$ تابع $\rightarrow$ تابع توزیع احتمال

$\hookrightarrow \mathbb{P}(X = x)$

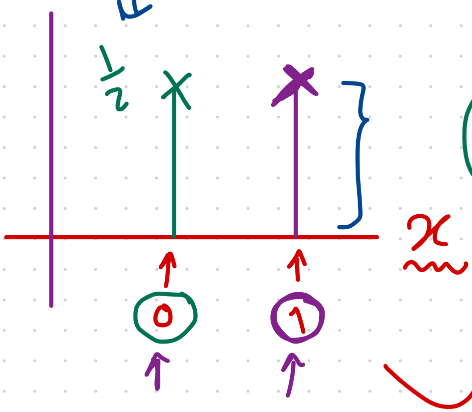
تابع توزیع احتمال $\rightarrow$ $f_X(x) = \mathbb{P}(X = x)$

سکه انداختن



$f_X(0) = \mathbb{P}(X=0) = \mathbb{P}(\text{خط}) = \frac{1}{2}$

$f_X(1) = \mathbb{P}(X=1) = \mathbb{P}(\text{شیر}) = \frac{1}{2}$



توزیع

$\mathbb{P}(\Omega) = 1$

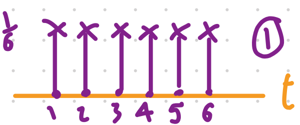
$\sum_x f_X(x) = 1$

$E_X = \sum_i \underbrace{x_i}_{\text{مقدار}} \underbrace{f_X(x_i)}_{\text{احتمال}}$

امید ریاضی

امید ریاضی تاس انداختن

$f_T(t) = \frac{1}{6}$



Discrete Random Variables

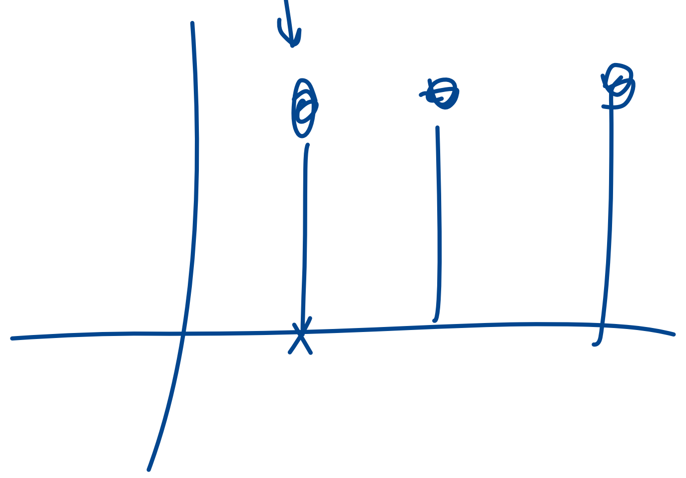
P

اسم — توزیع

$x \in \{\sim, \sim, \sim, \sim\}$

- Uniform
- Bernouli آزمایش ۲ حالته
- Binominal
- Poisson
- Negative Binominal
- Geometric
- Hypergeometric

اتفاقات تصادفی



تاس انداختن ~ Uniform (n=6)

یک بار سکه انداختن ~ Bernouli (p=0.6)

Probability mass function $f(n, k)$

PMF

60×3

20×1

$1 \times \frac{1}{20}$

Siméon Denis Poisson (1781-1840)

$X \sim \text{Binom}(n,k,p) = \binom{n}{k} \cdot p^k \cdot (1-p)^{(n-k)}$

$$\lim_{n\to\infty} \binom{n}{k} \cdot p^k \cdot (1-p)^{(n-k)}$$

$\frac{n!}{k!(n-k)!}$ $\quad \frac{n \times n-1 \times \cdots \times n-k+1 \times \cancel{n-k} \times \cdots \cancel{1}}{(\cancel{n-k}) \cdots \times 2 \cancel{\times 1} \times k!} \times p^k \times (1-p)^{(n-k)}$

$$n \to \infty$$
$$p \to 0$$
$$np \to \lambda$$

$\frac{n^k}{n^k} \frac{n \times (n-1) \times (n-2) \times \cdots (n-k+1)}{k!} \times p^k \times (1-p)^{n-k}$

$\lim_{\substack{n\to\infty \\ p\to 0 \\ np \to \lambda}} \frac{n}{n} \times \frac{(n-1)}{n} \times \frac{n-2}{n} \cdots \frac{(n-k+1)}{n} \quad \frac{1}{k!} n^k p^k \frac{(1-p)^n}{(1-p)^k}$

$1 \times 1 \times 1 \times \cdots \times 1 \qquad \times \frac{1}{k!} \times np^k \qquad 1$

$$\lim_{n\to\infty} (1 - \frac{a}{n})^n = e^{-a}$$

$e^{-\lambda} \frac{\lambda^k}{k!} = \frac{\lambda^k}{k!} \boxed{\lim_{\substack{p\to 0 \\ n\to\infty}} (1 - \frac{\lambda}{n})^n}$ $\quad e^{-\lambda}$

Continous Random Variables

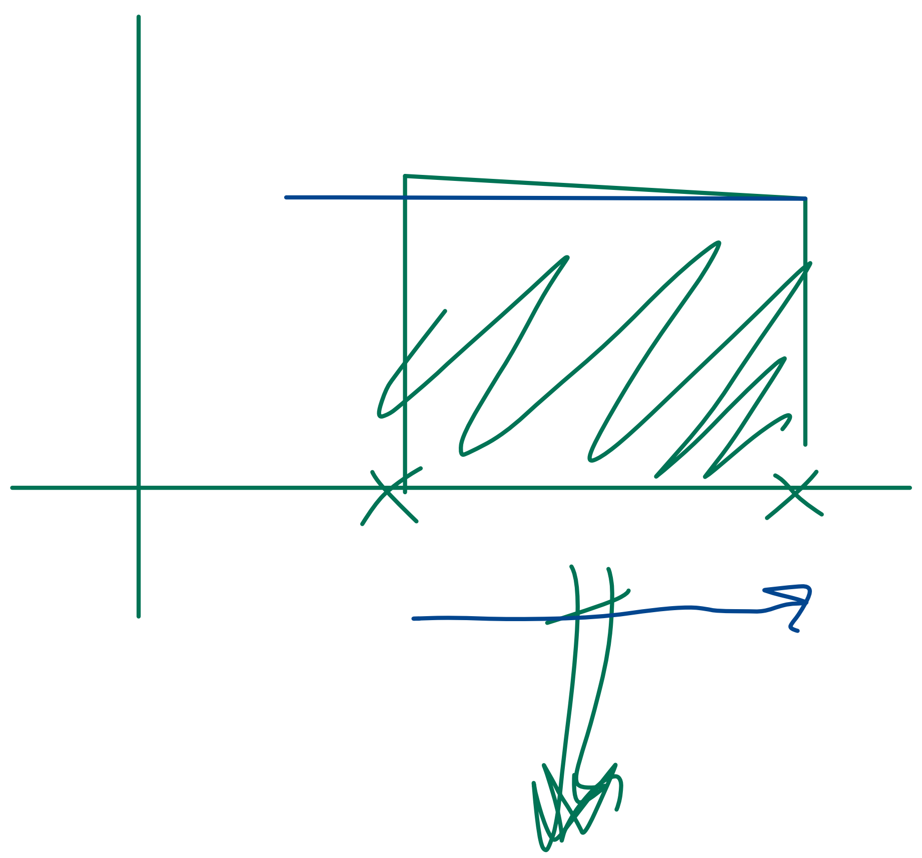
PDF

P Density

پیوسته

- Uniform
- Normal
- Geometric
- Gaussian

$x \in [1, 2]$



گوس

تقریب پیوسته از دوجمله

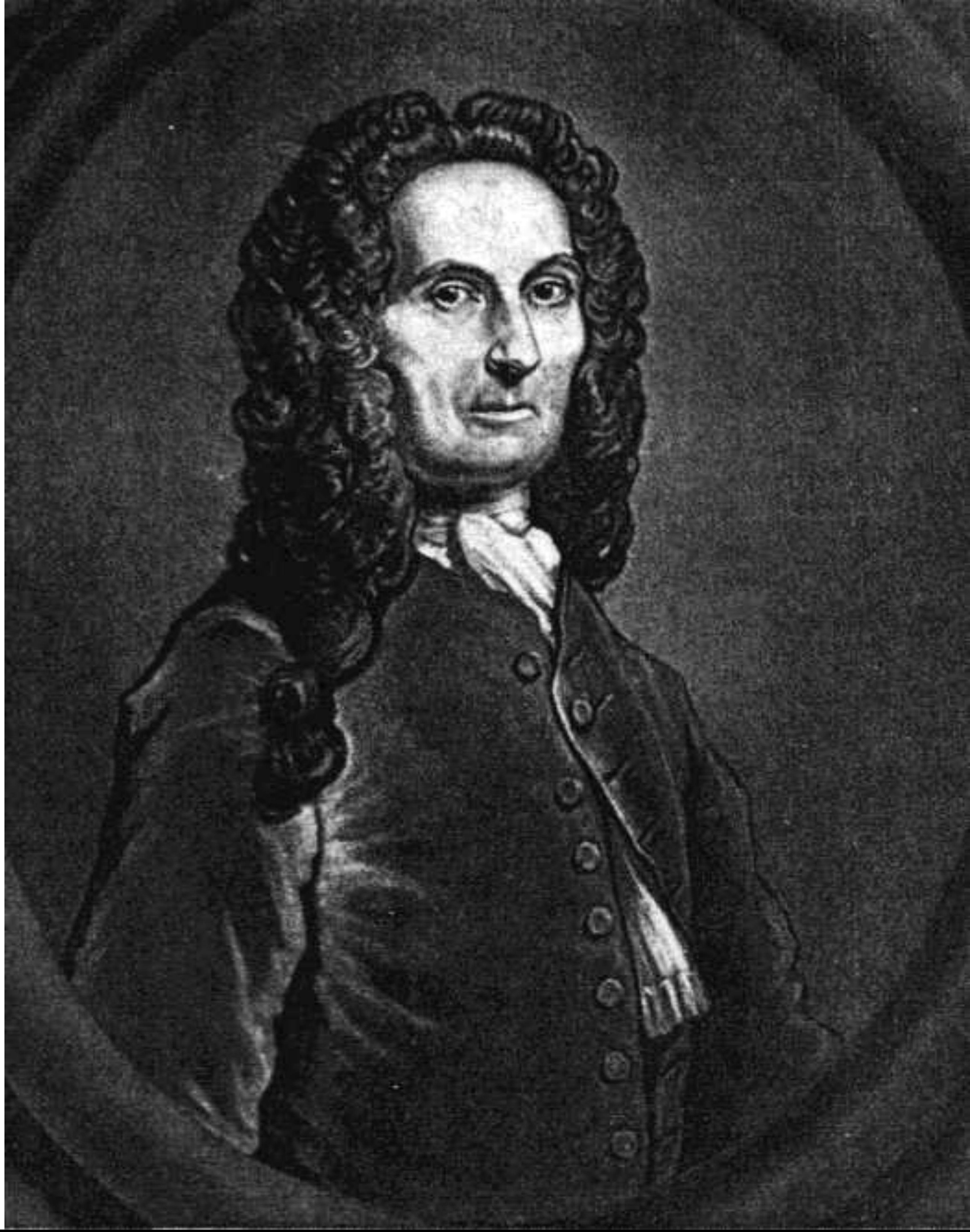
$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

تخمین و خطای آزمایش

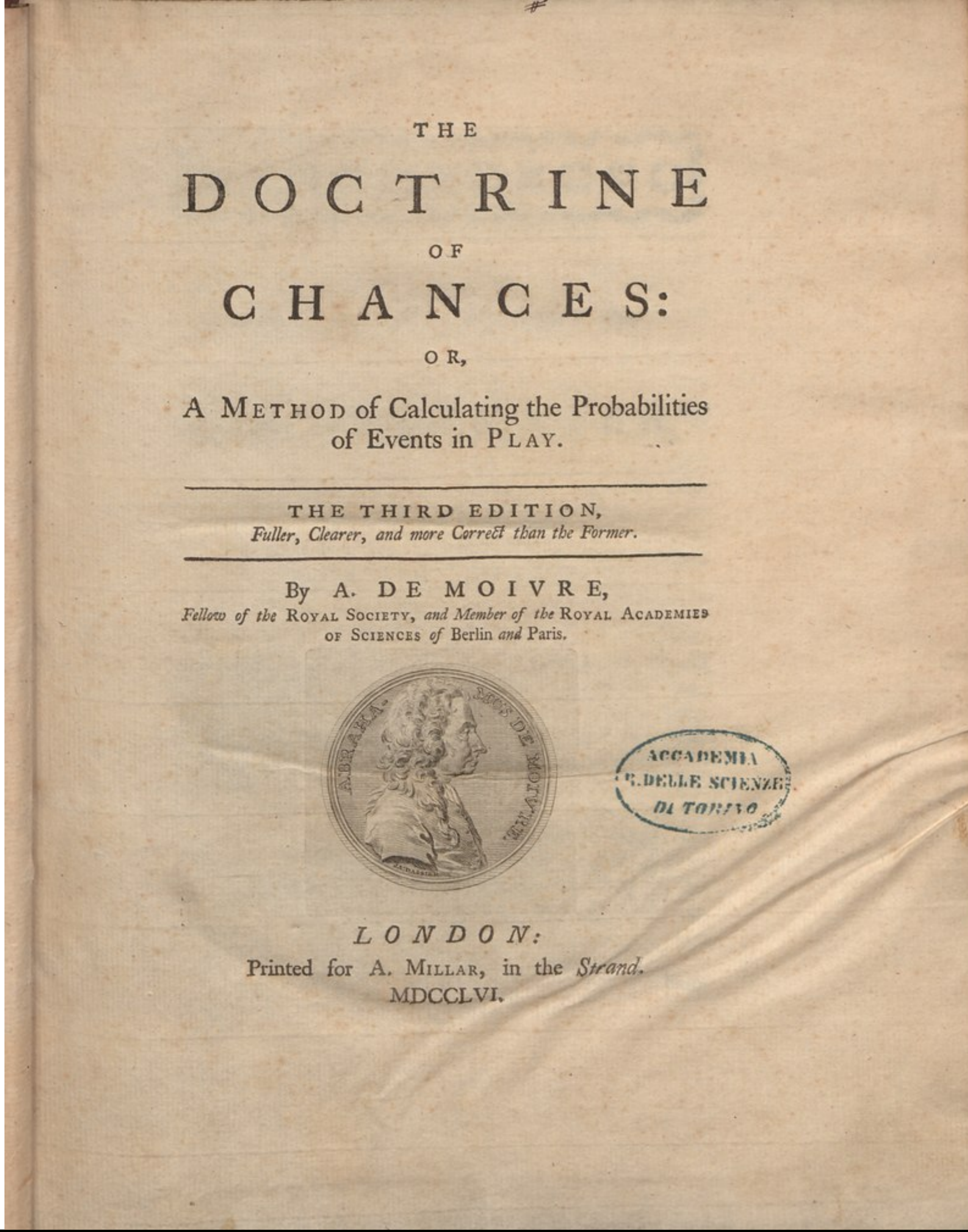
error curve

Abraham de Moivre (1667-1754)

$$\binom{n}{\frac{n}{2}+d}\left(\frac{1}{2}\right)^n \approx \frac{2}{\sqrt{2\pi n}}e^{\frac{-2d^2}{n}}$$



Doctorine of chances (1718)



THE

DOCTRINE

OF

CHANCES:

OR,

A METHOD of Calculating the Probabilities of Events in PLAY.

THE THIRD EDITION,

Fuller, Clearer, and more Correct than the Former.

By A. DE MOIVRE,

Fellow of the ROYAL SOCIETY, *and Member of the* ROYAL ACADEMIES OF SCIENCES *of* Berlin *and* Paris.

ABRAHAMUS DE MOIVRE

LONDON:

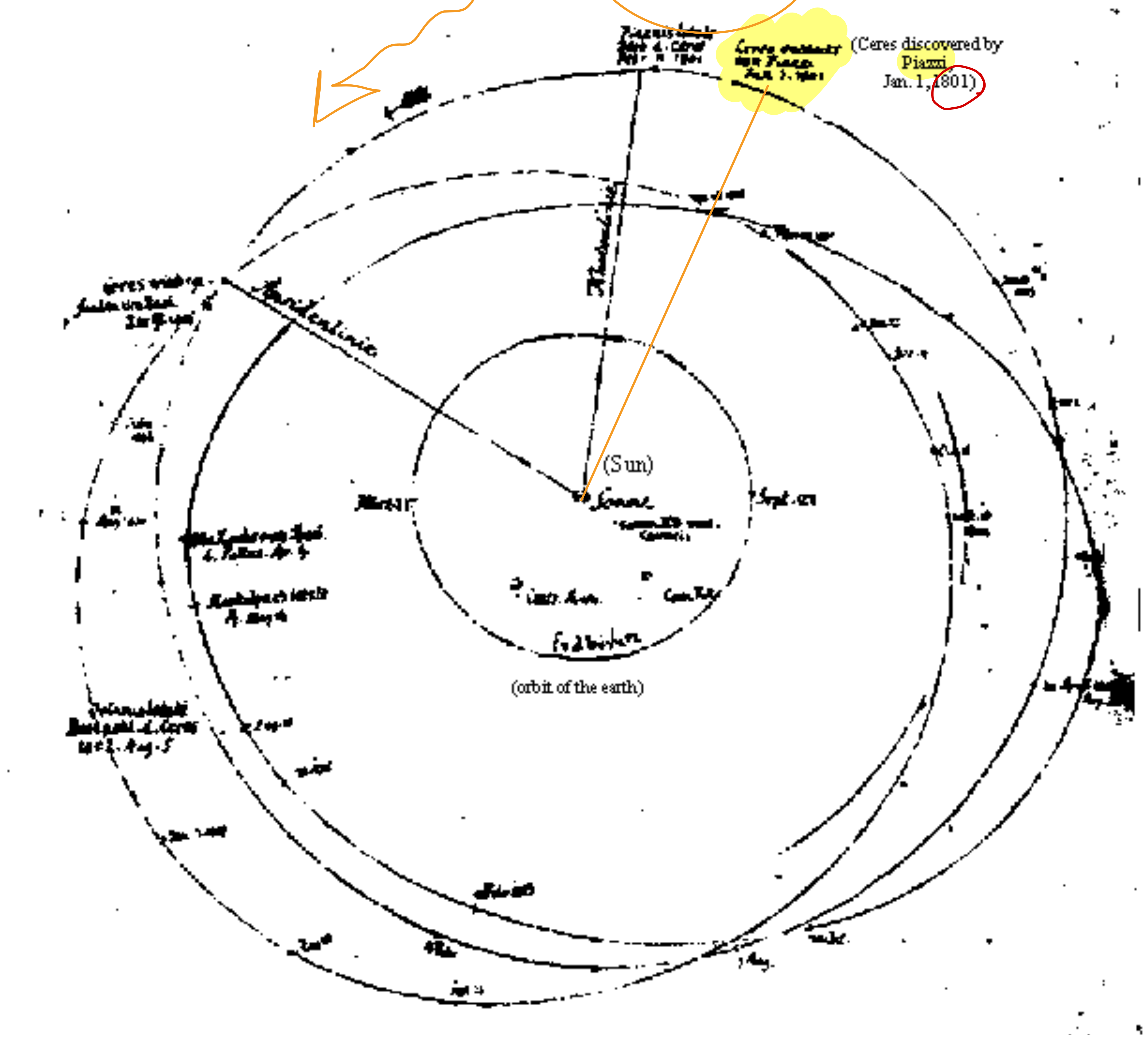
Printed for A. MILLAR, in the *Strand*.

MDCCLVI.

مریخ و مشتری

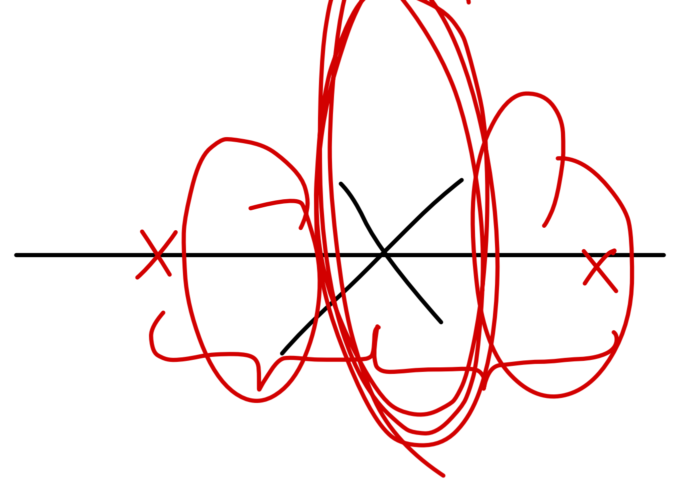
6 هفته

Ceres



The story of "the missing planet" and the celestial police lead to the development of error curves.

error curve



Galileo di Vincenzo Bonaiuti de' Galilei (1564-1642)

1. There is only one number which gives the distance of the star from the center of the earth, the true distance.
2. All observations are encumbered with errors, due to the observer, the instruments, and the other observational conditions.
3. The observations are distributed symmetrically about the true value; that is the errors are distributed symmetrically about zero.
4. Small errors occur more frequently than large errors.
5. The calculated distance is a function of the direct angular observations such that small adjustments of the observations may result in a large adjustment of the distance.

Pierre-Simon Laplace (1749-1827)

$$\phi(x) = \frac{m}{2} e^{-m|x|}$$

$$\phi(x) = \frac{1}{2a} \ln(\frac{a}{|x|})$$

Data Scientist

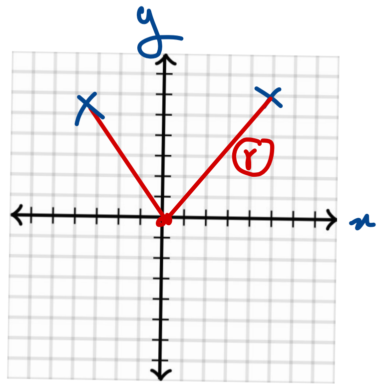
Johann Carl Friedrich Gauss (1777-1855)

$$g(\sqrt{x^2+y^2}) = g(x)g(y)$$

$$\phi(x) \propto e^{-h^2x^2}$$

$$\int_{\mathbb{R}} e^{-x^2} = \sqrt{\pi}$$

$$\phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2x^2}$$



$\phi(x,y) = \phi(R)$

$\phi(x)\,\phi(y) = \phi(\sqrt{x^2+y^2})$

$b^x b^y = b^{x+y}$

$b^{x^2} b^{y^2} = b^{x^2+y^2}$

e^{-x^2}

$V \approx \sum e^{-r^2}(2\pi r\Delta r + \cancel{\pi\Delta r^2})$

$V = \int_{r=0}^{\infty} e^{-r^2}\, 2\pi r\, dr$

$\int_{-\infty}^{\infty} e^{-x^2} dx = ? = A = \sqrt{\pi}$

$\iint_{\mathbb{R}\,\mathbb{R}} e^{-x^2} \cdot e^{-y^2} dx\, dy = A^2 = \pi$

$\pi \int_0^\infty 2re^{-r^2} dr$

$\left[-e^{-r^2}\right]_0^\infty = -e^{-\infty} - (-e^{0}) = 1$

$\pi(r+\Delta r)^2 - \pi r^2$

$\frac{d}{dx} e^{2x}$

Gauss had to estimate 6 parameters from only 19 data points. With scarce data, measurement error can substantially alter the prediction accuracy.

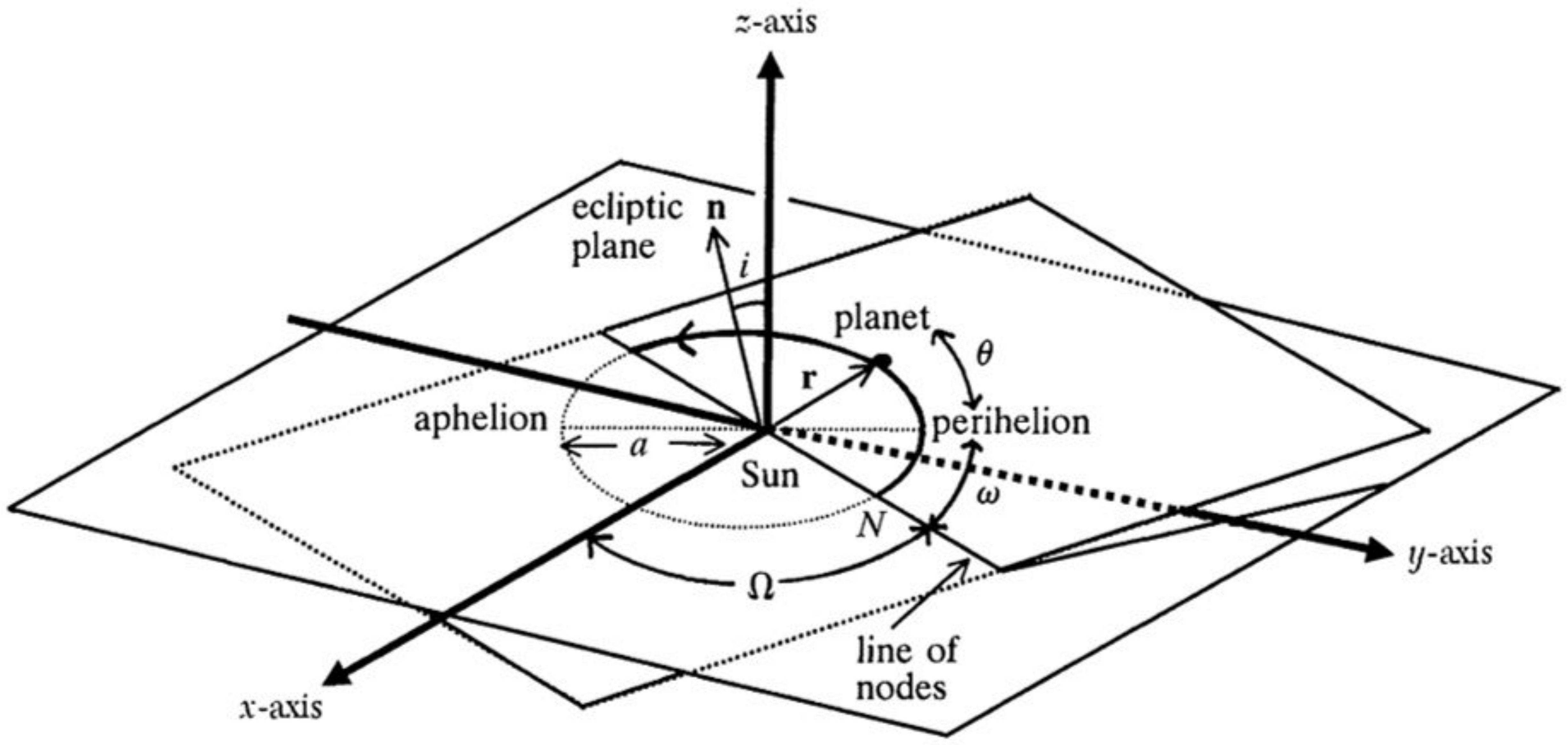
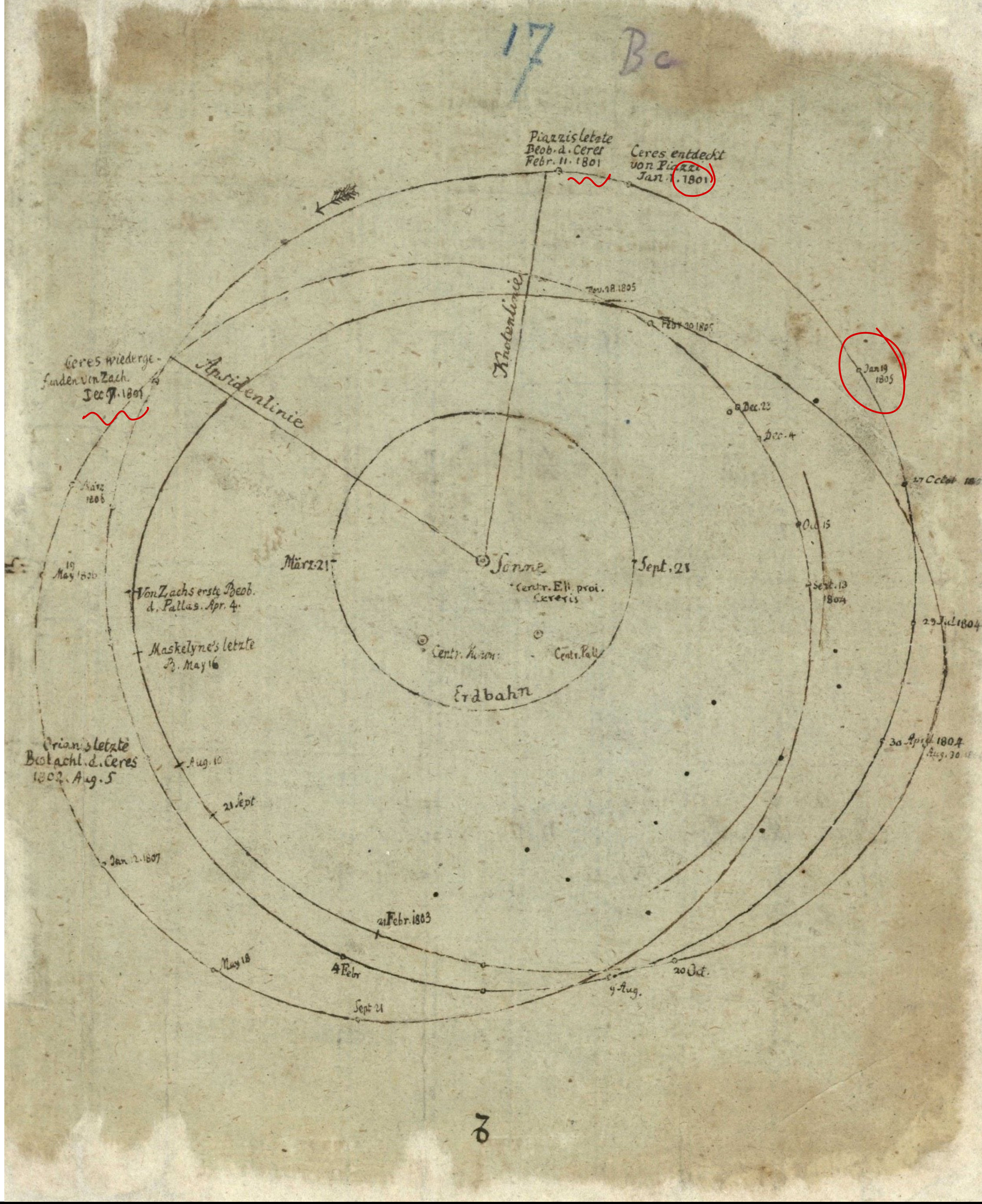


FIGURE 1
Parameters describing the planetary orbit

Gauss’ original sketch of the orbit of Ceres



جهان آمار

تصادفی

اتفاق سمت کائنات سمت داده

کائنات ∞

پارامتر — فرآیند تصادفی

θ, P — Binomial — Bayesian

از ۱۰ سکه ۸ شیر

فرض اولیه من

$H_0 = X \sim \text{Binomial}$

$\theta = \frac{1}{2}$ خودش یعنی سکه درستکاری

$P(\text{data} \mid H_0)$

$\binom{10}{8} \, 1/2^{10}$

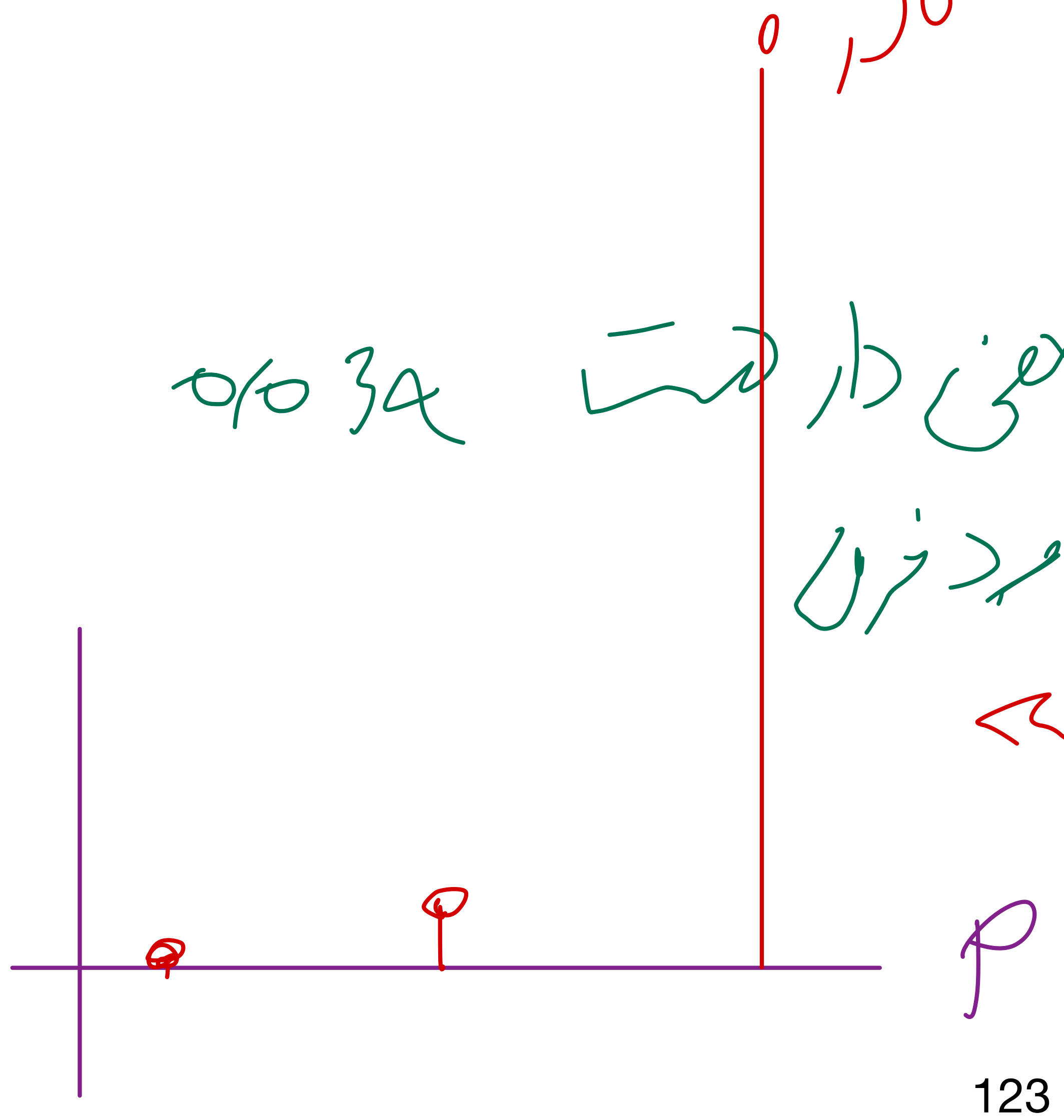
Pvalue $= 0.054$ معنی دار نیست

فرض نفی

۱۱ سکه ۹ شیر

Pvalue 0.034 معنی دار هست

رد فرض



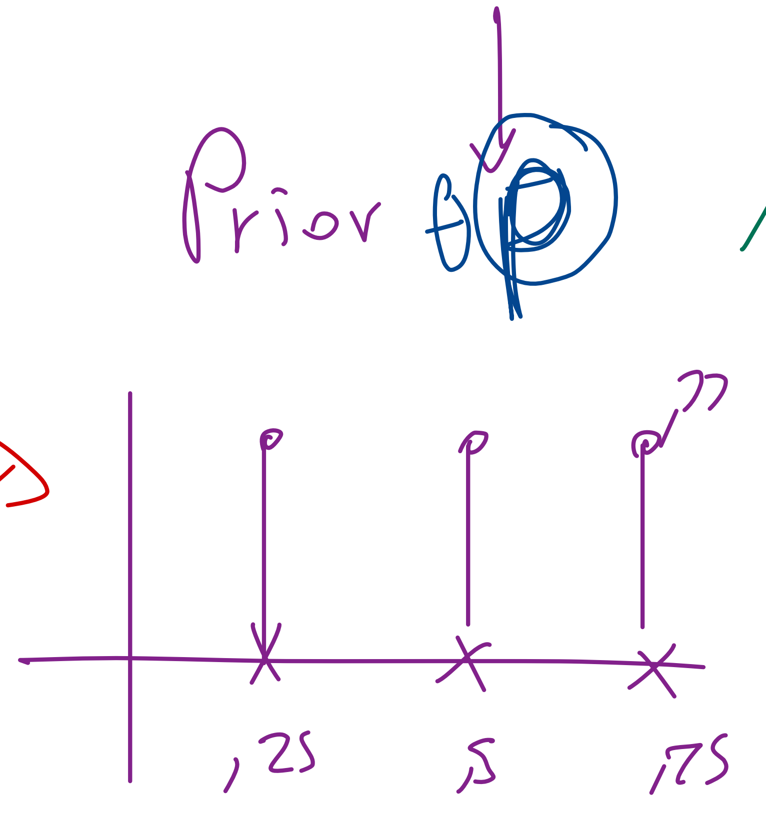
این شیاد و سه سکه

P $.25$

P $.5$

P $.75$

Prior θ P



Data

$P(\theta \mid \text{data}) =$

$$P(\theta \mid \text{data}) = \frac{P(\text{data} \mid \theta)\, P(\theta)}{P(\text{data})}$$

0.0042, 0.054, 0.152

$P(\text{data}) = P(\text{data} \mid .75)P(.75) + P(\text{data} \mid .5)P(.5) + P(\text{data} \mid .25)P(.25)$

$P(\text{data} \mid .25)\, P(.25)$ — 0.0042, $\frac{1}{3}$

$P(\text{data} \mid .5)\, P(.5)$ — 0.054, $\frac{1}{3}$

$P(\text{data} \mid .75)\, P(.75)$ — $.152$, $\frac{1}{3}$

$.57$, $.25$, $.9$, 1%

$P \sim 4\%$ $P \sim 5\%$ $P \sim 6\%$ $P \sim 7\%$

6% تالاسمی

$P =$ فراوانی تالاسمی در جمعیت

پارامتر مدل واقعی

$P = 5\%$

$\mathcal{L}(p) = \mathbb{P}(data \mid p)$

Likelihood P

4% 5%

متغیر تصادفی

برآورد $P = 4\%$

Maximum