What is probability?

Marco Cattaneo
Department of Physics and Mathematics, University of Hull

Open Day
25 October 2014

Met Office

Settings | Email alerts

Search

Weather | Climate | Learning | Research | Products | News | Holiday weather | Get ready for winter

UK forecast | Severe weather warnings | Specialist forecasts | Last 24 hours
UK climate | World weather | Weather guide | Weather services

Hull

Nearby locations

Recent locations

Enter town, city or postcode to find forecast

Share

Hull last 24 hours

Thursday	Friday	Saturday	Sunday	Monday
15°C 13°C	14°C 7°C	14°C 11°C	14°C 11°C	16°C 13°C

Time	00:00	01:00	02:00	03:00	04:00	05:00	06:00	07:00	08:00	09:00	10:00	11:00
Temperature	14°	14°	14°	14°	14°	13°	13°	13°	13°	13°	13°	13°
Feels like temperature (°C)	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°
Precipitation probability	<5%	<5%	<5%	15%	<5%	10%	10%	35%	80%	75%	55%	45%
Wind direction, speed & gust (mph)	SSW 12	SSW 12	SSW 12	SSW 12	SSW 11	SSW 11	SSW 11	SSW 10	SSW 9	SSW 8	SW 8	WSW 7

Time	12:00	13:00	14:00	15:00	16:00	17:00	18:00	19:00	20:00	21:00	22:00	23:00
Temperature	13°	13°	13°	13°	13°	13°	12°	12°	11°	11°	10°	10°
Feels like temperature (°C)	12°	12°	12°	12°	12°	11°	10°	10°	9°	8°	8°	7°
Precipitation probability	15%	15%	50%	10%	<5%	<5%	50%	10%	10%	<5%	<5%	<5%
Wind direction, speed & gust (mph)	WSW 7	WSW 7	WSW 7	WSW 6	WSW 6	WSW 7	W 9	W 9	WSW 9	WSW 9	WSW 9	WSW 9

Hull

Nearby locations

Recent locations

Enter town, city or postcode to find forecast

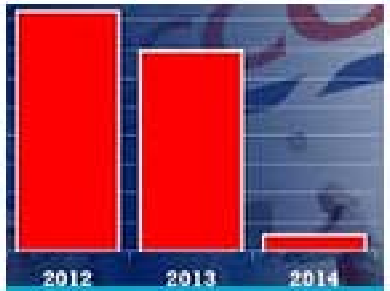
Share

Hull last 24 hours

Thursday	Friday	Saturday	Sunday	Monday
15°C 13°C	14°C 7°C	14°C 11°C	14°C 11°C	16°C 13°C

	00⁰⁰	01⁰⁰	02⁰⁰	03⁰⁰	04⁰⁰	05⁰⁰	06⁰⁰	07⁰⁰	08⁰⁰	09⁰⁰	10⁰⁰	11⁰⁰	12⁰⁰	13⁰⁰	14⁰⁰	15⁰⁰	16⁰⁰	17⁰⁰	18⁰⁰	19⁰⁰	20⁰⁰	21⁰⁰	22⁰⁰	23⁰⁰
Temperature	14°	14°	14°	14°	14°	13°	13°	13°	13°	13°	13°	13°	13°	13°	13°	13°	13°	13°	12°	12°	11°	11°	10°	10°
Feels like temperature (°C)	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	12°	11°	10°	10°	9°	8°	8°	7°
Precipitation probability	<5%	<5%	<5%	15%	<5%	10%	10%	35%	80%	75%	55%	45%	15%	15%	50%	10%	<5%	<5%	50%	10%	10%	<5%	<5%	<5%
Wind direction, speed & gust (mph)	SSW 12	SSW 12	SSW 12	SSW 12	SSW 11	SSW 11	SSW 11	SSW 10	SSW 9	SSW 8	SW 8	WSW 7	WSW 7	WSW 7	WSW 7	WSW 6	WSW 6	WSW 7	W 9	W 9	WSW 9	WSW 9	WSW 9	WSW 9

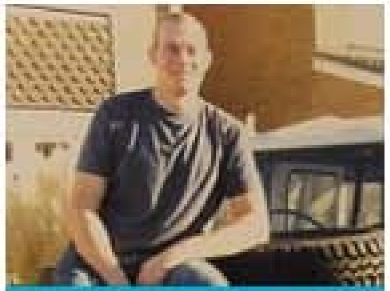
What a tragic waste of two lives: Lying



Tesco chairman to stand down as

Michael Schumacher's

Shaving half a leg and using fabric



Beekeeper who was allergic to bees

Church minister, 71, charged with sex

Taking the Pill 'raises the risk of breast cancer by 50 per cent'

- Women taking Pill have 50 per cent higher overall risk of breast cancer
- Some pills with high levels of oestrogen can raise risk threefold compared to women who have never taken pill
- Pills containing low-dose hormones carried no extra risks, US study finds

By JENNY HOPE MEDICAL CORRESPONDENT

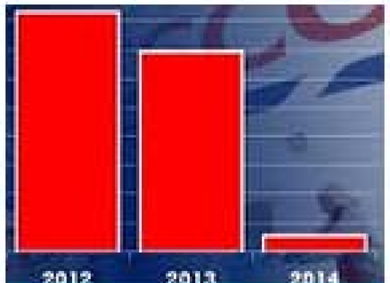
PUBLISHED: 07:08, 1 August 2014 | **UPDATED:** 09:15, 1 August 2014

Women taking contraceptive pills have a 50 per cent higher overall risk of developing breast cancer, a study has found.

Some pills with high levels of oestrogen can raise the risk threefold, compared with that of women who have never taken the Pill or who have stopped using it, US scientists found.

Pills containing low-dose hormones carried no extra risk.


What a tragic waste of two lives: Lying



Tesco chairman to stand down as


Michael Schumacher's


Shaving half a leg and using fabric


Beekeeper who was allergic to bees


Church minister, 71, charged with sex

Taking the Pill 'raises the risk of breast cancer by 50 per cent'

- **Women taking Pill have 50 per cent higher overall risk of breast cancer**
- **Some pills with high levels of oestrogen can raise risk threefold compared to women who have never taken pill**
- **Pills containing low-dose hormones carried no extra risks, US study finds**

By JENNY HOPE MEDICAL CORRESPONDENT

PUBLISHED: 07:08, 1 August 2014 | **UPDATED:** 09:15, 1 August 2014

Women taking contraceptive pills have a 50 per cent higher overall risk of developing breast cancer, a study has found.

Some pills with high levels of oestrogen can raise the risk threefold, compared with that of women who have never taken the Pill or who have stopped using it, US scientists found.

Pills containing low-dose hormones carried no extra risk.

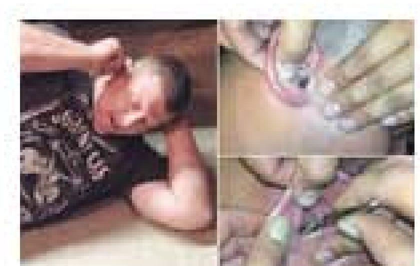
Ebola outbreak: Britain has 50% chance of importing deadly virus with the next three weeks

Oct 05, 2014 18:27 | By **Anthony Bond**

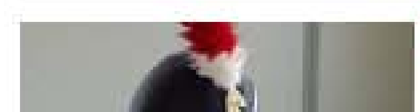
Experts used Ebola disease spread patterns and airline traffic data to predict the disease could reach the UK by October 24

Recommended In News

Improve these suggestions



VIRAL
You will never guess what was pulled out of this man's ear with a pair of tweezers



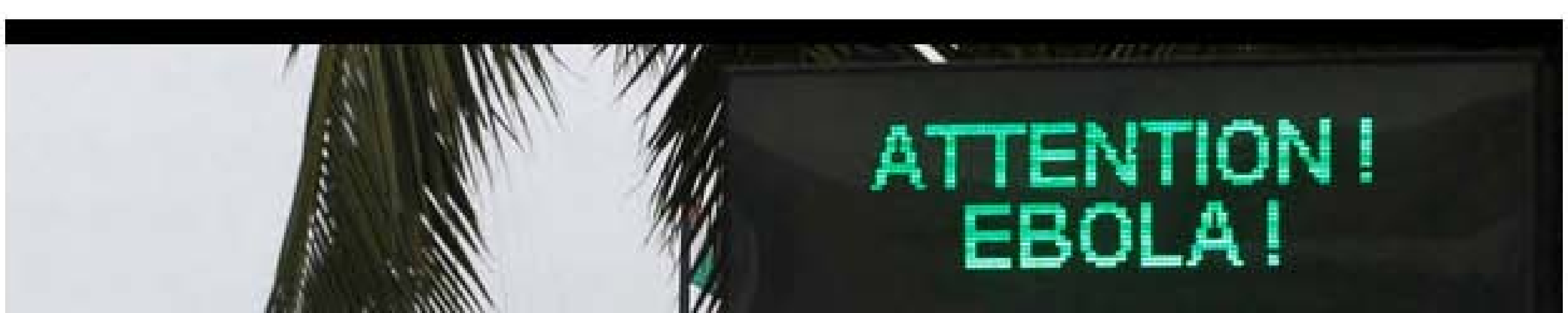
SOLDIERS
Devastated family's

Mirror

Ebola outbreak: Britain has 50% chance of importing deadly virus with the next three weeks

Oct 05, 2014 18:27 | By **Anthony Bond**

Experts used Ebola disease spread patterns and airline traffic data to predict the disease could reach the UK by October 24



statistical literacy

statistical literacy

- probability and statistics as high level numeracy skills

statistical literacy

- probability and statistics as high level numeracy skills
- expected of citizens in modern societies to understand news, product information, political debate, etc.

statistical literacy

- probability and statistics as high level numeracy skills
- expected of citizens in modern societies to understand news, product information, political debate, etc.
- increases employability:

> *"I keep saying the sexy job in the next ten years will be statisticians. People think I'm joking, but who would've guessed that computer engineers would've been the sexy job of the 1990s? The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades, not only at the professional level but even at the educational level for elementary school kids, for high school kids, for college kids. Because now we really do have essentially free and ubiquitous data. So the complimentary scarce factor is the ability to understand that data and extract value from it."*
>
> Hal Varian (professor at UC Berkeley, chief economist at Google), McKinsey, January 2009

statistical literacy

- probability and statistics as high level numeracy skills
- expected of citizens in modern societies to understand news, product information, political debate, etc.
- increases employability:

> *"I keep saying the sexy job in the next ten years will be statisticians. People think I'm joking, but who would've guessed that computer engineers would've been the sexy job of the 1990s? The ability to take data—to be able to understand it, to process it, to extract value from it, to visualize it, to communicate it—that's going to be a hugely important skill in the next decades, not only at the professional level but even at the educational level for elementary school kids, for high school kids, for college kids. Because now we really do have essentially free and ubiquitous data. So the complimentary scarce factor is the ability to understand that data and extract value from it."*
>
> Hal Varian (professor at UC Berkeley, chief economist at Google), McKinsey, January 2009

mathematical definition

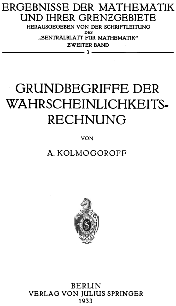
mathematical definition

probability is a normalized measure

mathematical definition

probability is a normalized measure

Andrey Kolmogorov (Tambov 1903 – Moscow 1987):
Grundbegriffe der Wahrscheinlichkeitsrechnung (1933)



ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK“
ZWEITER BAND
3

GRUNDBEGRIFFE DER
WAHRSCHEINLICHKEITS-
RECHNUNG

VON

A. KOLMOGOROFF

BERLIN
VERLAG VON JULIUS SPRINGER
1933

ERGEBNISSE DER MATHEMATIK
UND IHRER GRENZGEBIETE
HERAUSGEGEBEN VON DER SCHRIFTLEITUNG
DES
„ZENTRALBLATT FÜR MATHEMATIK"
ZWEITER BAND
3

GRUNDBEGRIFFE DER WAHRSCHEINLICHKEITS-RECHNUNG

VON

A. KOLMOGOROFF

BERLIN
VERLAG VON JULIUS SPRINGER
1933

FOUNDATIONS
OF THE
THEORY OF PROBABILITY

BY

A. N. KOLMOGOROV

Second English Edition

TRANSLATION EDITED BY
NATHAN MORRISON

WITH AN ADDED BIBLIOGRAPHY BY
A. T. BHARUCHA-REID
UNIVERSITY OF OREGON

CHELSEA PUBLISHING COMPANY
NEW YORK

1956

the system of axioms and in the further development of the theory, then the postulational concepts of a random event and its probability seem the most suitable. There are other postulational systems of the theory of probability, particularly those in which the concept of probability is not treated as one of the basic concepts, but is itself expressed by means of other concepts.[1] However, in that case, the aim is different, namely, to tie up as closely as possible the mathematical theory with the empirical development of the theory of probability.

§ 1. Axioms[2]

Let E be a collection of elements $\xi, \eta, \zeta, \ldots$, which we shall call *elementary events,* and $\mathfrak{F}$ a set of subsets of E; the elements of the set $\mathfrak{F}$ will be called *random events.*

I. *$\mathfrak{F}$ is a field[3] of sets.*

II. *$\mathfrak{F}$ contains the set E.*

III. *To each set A in $\mathfrak{F}$ is assigned a non-negative real number* $\mathsf{P}(A)$. *This number* $\mathsf{P}(A)$ *is called the probability of the event A.*

IV. $\mathsf{P}(E)$ *equals* 1.

V. *If A and B have no element in common, then*

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$$

A system of sets, $\mathfrak{F}$, together with a definite assignment of numbers $\mathsf{P}(A)$, satisfying Axioms I-V, is called a *field of probability.*

Our system of Axioms I-V is *consistent.* This is proved by the following example. Let E consist of the single element ξ and let $\mathfrak{F}$ consist of E and the null set 0. $\mathsf{P}(E)$ is then set equal to 1 and $\mathsf{P}(0)$ equals 0.

---

[1] For example, R. von Mises [1] and [2] and S. Bernstein [1].

[2] The reader who wishes from the outset to give a concrete meaning to the following axioms, is referred to § 2.

[3] Cf. HAUSDORFF, *Mengenlehre,* 1927, p. 78. A system of sets is called a field if the sum, product, and difference of two sets of the system also belong to the same system. Every non-empty field contains the null set 0. Using Hausdorff's notation, we designate the product of A and B by AB; the sum by $A+B$ in the case where $AB=0$; and in the general case by $A \dotplus B$; the difference of A and B by $A-B$. The set $E-A$, which is the complement of A, will be denoted by $\bar{A}$. We shall assume that the reader is familiar with the fundamental rules of operations of sets and their sums, products, and differences. All subsets of $\mathfrak{F}$ will be designated by Latin capitals.

the system of axioms and in the further development of the theory, then the postulational concepts of a random event and its probability seem the most suitable. There are other postulational systems of the theory of probability, particularly those in which the concept of probability is not treated as one of the basic concepts, but is itself expressed by means of other concepts.[1] However, in that case, the aim is different, namely, to tie up as closely as possible the mathematical theory with the empirical development of the theory of probability.

§ 1. Axioms[2]

Let E be a collection of elements $\xi, \eta, \zeta, \ldots$, which we shall call *elementary events*, and $\mathfrak{F}$ a set of subsets of E; the elements of the set $\mathfrak{F}$ will be called *random events*.

I. *$\mathfrak{F}$ is a field*[3] *of sets.*

II. *$\mathfrak{F}$ contains the set E.*

III. *To each set A in $\mathfrak{F}$ is assigned a non-negative real number* $\mathsf{P}(A)$. *This number $\mathsf{P}(A)$ is called the probability of the event A.*

IV. $\mathsf{P}(E)$ *equals* 1.

V. *If A and B have no element in common, then*

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$$

A system of sets, $\mathfrak{F}$, together with a definite assignment of numbers $\mathsf{P}(A)$, satisfying Axioms I-V, is called a *field of probability*.

Our system of Axioms I-V is *consistent*. This is proved by the following example. Let E consist of the single element ξ and let $\mathfrak{F}$ consist of E and the null set 0. $\mathsf{P}(E)$ is then set equal to 1 and $\mathsf{P}(0)$ equals 0.

---

[1] For example, R. von Mises [1] and [2] and S. Bernstein [1].

[2] The reader who wishes from the outset to give a concrete meaning to the following axioms, is referred to § 2.

[3] Cf. Hausdorff, *Mengenlehre*, 1927, p. 78. A system of sets is called a field if the sum, product, and difference of two sets of the system also belong to the same system. Every non-empty field contains the null set 0. Using Hausdorff's notation, we designate the product of A and B by AB; the sum by $A+B$ in the case where $AB = 0$; and in the general case by $A \dotplus B$; the difference of A and B by $A - B$. The set $E - A$, which is the complement of A, will be denoted by $\bar{A}$. We shall assume that the reader is familiar with the fundamental rules of operations of sets and their sums, products, and differences. All subsets of $\mathfrak{F}$ will be designated by Latin capitals.

Chapter I

ELEMENTARY THEORY OF PROBABILITY

We define as elementary theory of probability that part of the theory in which we have to deal with probabilities of only a finite number of events. The theorems which we derive here can be applied also to the problems connected with an infinite number of random events. However, when the latter are studied, essentially new principles are used. Therefore the only axiom of the mathematical theory of probability which deals particularly with the case of an infinite number of random events is not introduced until the beginning of Chapter II (Axiom VI).

The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra. This means that after we have defined the elements to be studied and their basic relations, and have stated the axioms by which these relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations.

In accordance with the above, in § 1 the concept of a *field of probabilities* is defined as a system of sets which satisfies certain conditions. What the elements of this set represent is of no importance in the purely mathematical development of the theory of probability (cf. the introduction of basic geometric concepts in the *Foundations of Geometry* by Hilbert, or the definitions of groups, rings and fields in abstract algebra).

Every axiomatic (abstract) theory admits, as is well known, of an unlimited number of concrete interpretations besides those from which it was derived. Thus we find applications in fields of science which have no relation to the concepts of random event and of probability in the precise meaning of these words.

The postulational basis of the theory of probability can be established by different methods in respect to the selection of axioms as well as in the selection of basic concepts and relations. However, if our aim is to achieve the utmost simplicity both in

Chapter I

ELEMENTARY THEORY OF PROBABILITY

We define as elementary theory of probability that part of the theory in which we have to deal with probabilities of only a finite number of events. The theorems which we derive here can be applied also to the problems connected with an infinite number of random events. However, when the latter are studied, essentially new principles are used. Therefore the only axiom of the mathematical theory of probability which deals particularly with the case of an infinite number of random events is not introduced until the beginning of Chapter II (Axiom VI).

The theory of probability, as a mathematical discipline, can and should be developed from axioms in exactly the same way as Geometry and Algebra. This means that after we have defined the elements to be studied and their basic relations, and have stated the axioms by which these relations are to be governed, all further exposition must be based exclusively on these axioms, independent of the usual concrete meaning of these elements and their relations.

In accordance with the above, in § 1 the concept of a *field of probabilities* is defined as a system of sets which satisfies certain conditions. What the elements of this set represent is of no importance in the purely mathematical development of the theory of probability (cf. the introduction of basic geometric concepts in the *Foundations of Geometry* by Hilbert, or the definitions of groups, rings and fields in abstract algebra).

Every axiomatic (abstract) theory admits, as is well known, of an unlimited number of concrete interpretations besides those from which it was derived. Thus we find applications in fields of science which have no relation to the concepts of random event and of probability in the precise meaning of these words.

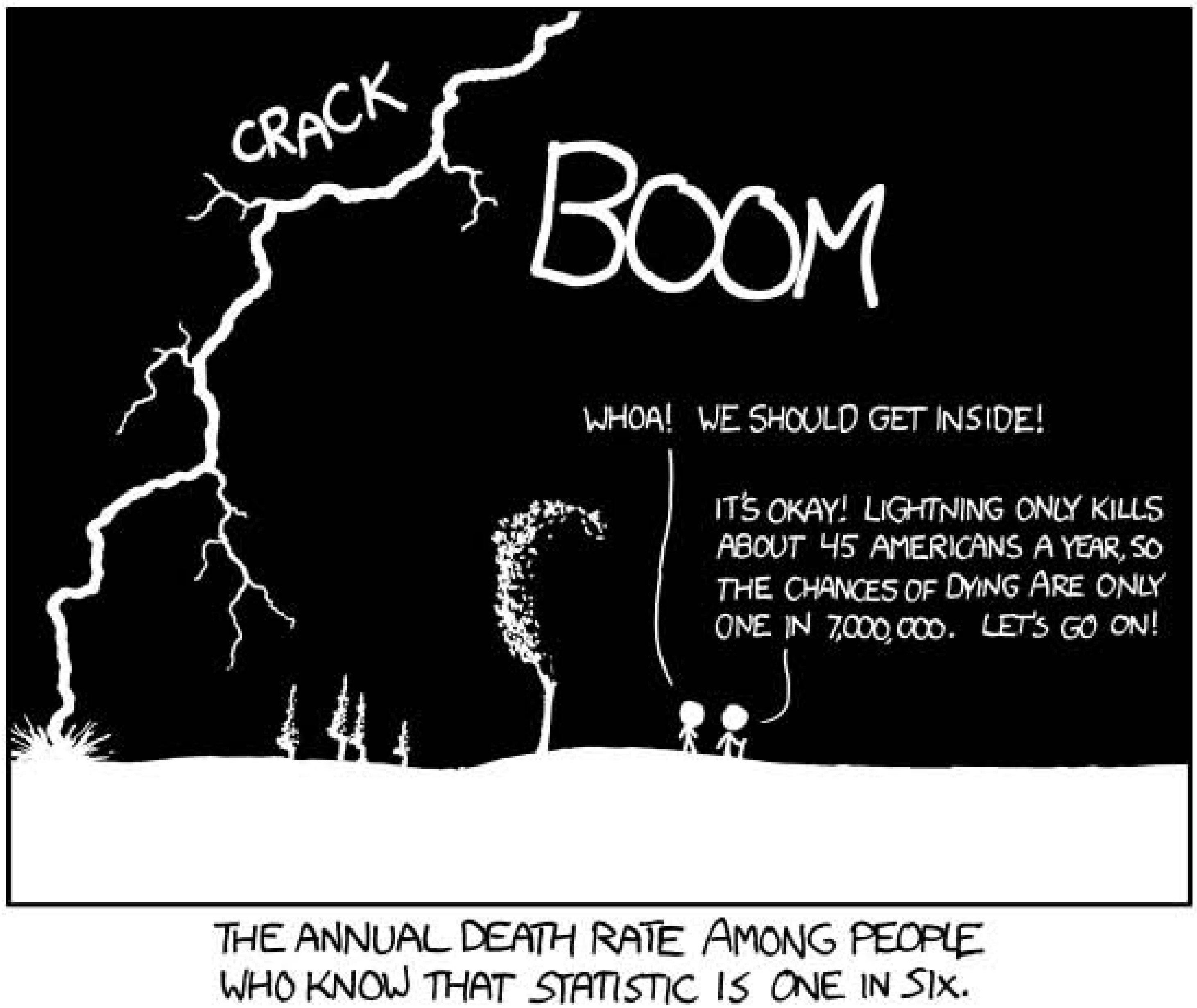
The postulational basis of the theory of probability can be established by different methods in respect to the selection of axioms as well as in the selection of basic concepts and relations. However, if our aim is to achieve the utmost simplicity both in xkcd

A webcomic of romance, sarcasm, math, and language.

My BOOK is out!
(#1 New York Times bestseller!)

Conditional Risk

|< | < Prev | Random | Next > | >|



example: probability of rain

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

- It will rain tomorrow in 30% of the Hull **region**.

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

- It will rain tomorrow in 30% of the Hull **region**.
- It will rain tomorrow for 30% of the **time**.

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

- It will rain tomorrow in 30% of the Hull **region**.
- It will rain tomorrow for 30% of the **time**.
- 30% of weather **forecasters** believe it will rain tomorrow.

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

- It will rain tomorrow in 30% of the Hull **region**.
- It will rain tomorrow for 30% of the **time**.
- 30% of weather **forecasters** believe it will rain tomorrow.
- It will rain on 30% of the **days** like tomorrow.

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**

▶ It will rain tomorrow in 30% of the Hull **region**.	16%
▶ It will rain tomorrow for 30% of the **time**.	10%
▶ 30% of weather **forecasters** believe it will rain tomorrow.	22%
▶ It will rain on 30% of the **days** like tomorrow.	19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. 16%
- It will rain tomorrow for 30% of the **time**. 10%
- 30% of weather **forecasters** believe it will rain tomorrow. 22%
- It will rain on 30% of the **days** like tomorrow. 19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. × 16%
- It will rain tomorrow for 30% of the **time**. 10%
- 30% of weather **forecasters** believe it will rain tomorrow. 22%
- It will rain on 30% of the **days** like tomorrow. 19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. × 16%
- It will rain tomorrow for 30% of the **time**. × 10%
- 30% of weather **forecasters** believe it will rain tomorrow. 22%
- It will rain on 30% of the **days** like tomorrow. 19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. × 16%
- It will rain tomorrow for 30% of the **time**. × 10%
- 30% of weather **forecasters** believe it will rain tomorrow. × 22%
- It will rain on 30% of the **days** like tomorrow. 19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. × 16%
- It will rain tomorrow for 30% of the **time**. × 10%
- 30% of weather **forecasters** believe it will rain tomorrow. × 22%
- It will rain on 30% of the **days** like tomorrow. (✓) 19%

Rebecca E. Morss, Julie L. Demuth, and Jeffrey K. Lazo: Communicating uncertainty in weather forecasts: A survey of the U.S. public, Weather and Forecasting 23 (2008)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. ×
- It will rain tomorrow for 30% of the **time**. ×
- 30% of weather **forecasters** believe it will rain tomorrow. ×
- It will rain on 30% of the **days** like tomorrow. (✓)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. ×
- It will rain tomorrow for 30% of the **time**. ×
- 30% of weather **forecasters** believe it will rain tomorrow. ×
- It will rain on 30% of the **days** like tomorrow. (✓)
- It will rain on 30% of the **days** for which the weather forecast says **30% probability of rain**.

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. ×
- It will rain tomorrow for 30% of the **time**. ×
- 30% of weather **forecasters** believe it will rain tomorrow. ×
- It will rain on 30% of the **days** like tomorrow. (✓)
- It will rain on 30% of the **days** for which
  the weather forecast says **30% probability of rain**. (✓)

example: probability of rain

Hull weather forecast: **30% probability of rain for tomorrow**
(at least 0.1 mm at the weather station)

- It will rain tomorrow in 30% of the Hull **region**. ×
- It will rain tomorrow for 30% of the **time**. ×
- 30% of weather **forecasters** believe it will rain tomorrow. ×
- It will rain on 30% of the **days** like tomorrow. (✓)
- It will rain on **approximately** 30% of the **days** for which the weather forecast says **30% probability of rain**. ✓

games of chance

games of chance

$$\textbf{probability of an event} = \frac{\textbf{number of favorable outcomes}}{\textbf{number of possible outcomes}}$$

games of chance

$$\textbf{probability of an event} = \frac{\textbf{number of favorable outcomes}}{\textbf{number of possible outcomes}}$$

- **Girolamo Cardano** (Pavia 1501 – Rome 1576):
  Liber de ludo aleae (1663, written around 1564)

games of chance

$$\textbf{probability of an event} = \frac{\textbf{number of favorable outcomes}}{\textbf{number of possible outcomes}}$$

- **Girolamo Cardano** (Pavia 1501 – Rome 1576):
  Liber de ludo aleae (1663, written around 1564)
- **Blaise Pascal** (Clermont-Ferrand 1623 – Paris 1662) and

games of chance

$$\textbf{probability of an event} = \frac{\textbf{number of favorable outcomes}}{\textbf{number of possible outcomes}}$$

- **Girolamo Cardano** (Pavia 1501 – Rome 1576):
  Liber de ludo aleae (1663, written around 1564)
- **Blaise Pascal** (Clermont-Ferrand 1623 – Paris 1662) and
  Pierre de Fermat (Beaumont-de-Lomagne 1601/1607 – Castres 1665):
  correspondence on the problem of division of the stakes (1654)



games of chance

$$\textbf{probability of an event} = \frac{\textbf{number of favorable outcomes}}{\textbf{number of possible outcomes}}$$

- **Girolamo Cardano** (Pavia 1501 – Rome 1576):
  Liber de ludo aleae (1663, written around 1564)
- **Blaise Pascal** (Clermont-Ferrand 1623 – Paris 1662) and
  Pierre de Fermat (Beaumont-de-Lomagne 1601/1607 – Castres 1665):
  correspondence on the problem of division of the stakes (1654)
- **Christiaan Huygens** (The Hague 1629 – The Hague 1695):
  De ratiociniis in ludo aleae (1657)



PSYCHOMETRIKA—VOL. 36, NO. 1
MARCH, 1971

BIAS AND RUNS IN DICE THROWING AND RECORDING: A FEW MILLION THROWS*

GUDMUND R. IVERSEN, WILLARD H. LONGCOR†, FREDERICK MOSTELLER, JOHN P. GILBERT, AND CLEO YOUTZ

An experimenter threw individually 219 different dice of four different brands and recorded even and odd outcomes for one block of 20,000 trials for each die—4,380,000 throws in all. The resulting data on runs offer a basis for comparing the observed properties of such a physical randomizing process with theory and with simulations based on pseudo-random numbers and RAND Corporation random numbers. Although generally the results are close to those forecast by theory, some notable exceptions raise questions about the surprise value that should be associated with occurrences two standard deviations from the mean. These data suggest that the usual significance level may well actually be running from 7 to 15 percent instead of the theoretical 5 percent.

The data base is the largest of its kind. A set generated by one brand of dice contains 2,000,000 bits and is the first handmade empirical data of such size to fail to show a significant departure from ideal theory in either location or scale.

1. *Introduction*

How well do the laws of chance actually work? When a die is repeatedly thrown and its outcomes recorded, do imperfections in the die, in the throwing, in the perception of the outcome, and in recording appear? What sorts of deviations from chance do we find?

Weldon's dice data [Fry, 1965] and Kerrich's coin tossing monograph [Kerrich, 1946] both give us some experience with large bodies of data produced by humanly run physical randomizing devices whose idealized probabilities and properties are known to a good approximation. In a sense, such experiments are controls on other experiments where probability plays an important role. For example, such dice and coin experiments give us an idea of how seriously we should take small departures from mathematically predicted results in investigations where we search for small departures from a standard. They do this by showing the sizes and kinds of departures observed in an experiment with *no planned* human or material effects. They are placebo experiments. If one does not believe in extra-sensory perception, then many ESP investigations also would be judged to qualify, but if one

* The analysis was facilitated by a National Science Foundation grant GS-341 and and its continuation GS-2044X. It forms part of a larger study of data analysis.

† Mr. Longcor is from Waukegan, Illinois; the other authors are from Harvard University. Dr. Iversen has moved to the University of Michigan.

PSYCHOMETRIKA—VOL. 36, NO. 1
MARCH, 1971

BIAS AND RUNS IN DICE THROWING AND RECORDING: A FEW MILLION THROWS*

GUDMUND R. IVERSEN, WILLARD H. LONGCOR†, FREDERICK MOSTELLER, JOHN P. GILBERT, AND CLEO YOUTZ

An experimenter threw individually 219 different dice of four different brands and recorded even and odd outcomes for one block of 20,000 trials for each die—4,380,000 throws in all. The resulting data on runs offer a basis for comparing the observed properties of such a physical randomizing process with theory and with simulations based on pseudo-random numbers and RAND Corporation random numbers. Although generally the results are close to those forecast by theory, some notable exceptions raise questions about the surprise value that should be associated with occurrences two standard deviations from the mean. These data suggest that the usual significance level may well actually be running from 7 to 15 percent instead of the theoretical 5 percent.

The data base is the largest of its kind. A set generated by one brand of dice contains 2,000,000 bits and is the first handmade empirical data of such size to fail to show a significant departure from ideal theory in either location or scale.

1. Introduction

How well do the laws of chance actually work? When a die is repeatedly thrown and its outcomes recorded, do imperfections in the die, in the throwing, in the perception of the outcome, and in recording appear? What sorts of deviations from chance do we find?

Weldon's dice data [Fry, 1965] and Kerrich's coin tossing monograph [Kerrich, 1946] both give us some experience with large bodies of data produced by humanly run physical randomizing devices whose idealized probabilities and properties are known to a good approximation. In a sense, such experiments are controls on other experiments where probability plays an important role. For example, such dice and coin experiments give us an idea of how seriously we should take small departures from mathematically predicted results in investigations where we search for small departures from a standard. They do this by showing the sizes and kinds of departures observed in an experiment with *no planned* human or material effects. They are placebo experiments. If one does not believe in extra-sensory perception, then many ESP investigations also would be judged to qualify, but if one

* The analysis was facilitated by a National Science Foundation grant GS-341 and and its continuation GS-2044X. It forms part of a larger study of data analysis.

† Mr. Longcor is from Waukegan, Illinois; the other authors are from Harvard University. Dr. Iversen has moved to the University of Michigan.

law of large numbers

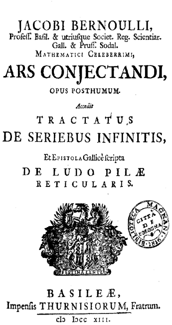
law of large numbers

probability of an event = long-run relative frequency of its occurrence

law of large numbers

probability of an event = long-run relative frequency of its occurrence

Jacob Bernoulli (Basel 1655 – Basel 1705):
Ars conjectandi (1713)



JACOBI BERNOULLI,
Profess. Basil. & utriusque Societ. Reg. Scientiar.
Gall. & Pruss. Sodal.
MATHEMATICI CELEBERRIMI,
ARS CONJECTANDI,
OPUS POSTHUMUM.
Accedit
TRACTATUS
DE SERIEBUS INFINITIS,
Et EPISTOLA Gallicè scripta
DE LUDO PILÆ
RETICULARIS.

BASILEÆ,
Impensis THURNISIORUM, Fratrum.
cIↃ IↃcc XIII.

JACOBI BERNOULLI,

Profeff. Bafil. & utriufque Societ. Reg. Scientiar. Gall. & Pruff. Sodal.

MATHEMATICI CELEBERRIMI,

ARS CONJECTANDI,

OPUS POSTHUMUM.

Accedit

TRACTATUS

DE SERIEBUS INFINITIS,

Et EPISTOLA Gallicè fcripta

DE LUDO PILÆ RETICULARIS.

BASILEÆ,

Impenfis THURNISIORUM, Fratrum.

cIↄ Iↄcc XIII.

JACOB BERNOULLI

THE ART OF CONJECTURING

together with

LETTER TO A FRIEND ON SETS IN COURT TENNIS

Translated with an introduction and notes by

Edith Dudley Sylla

THE JOHNS HOPKINS UNIVERSITY PRESS

Baltimore

between the maximum M and the bound L will exceed more than $c(s-1)$ times the same number of terms starting from the largest terms outside this bound. Similarly, they will exceed more than c times that many terms taken $s-1$ times. Therefore, even more obviously, they will exceed more than c times all the terms outside the bound L, of which there are only $s-1$ times as many more.

For the terms to the right, I proceed in the same way. I take the ratio $(s+1)/s < (rs+r)/(rs-s)$, I set $(s+1)^m/s^m \geq c(r-1)$, and I find that $m \geq \log[c(r-1)]/[\log(s+1) - \log s]$. Next, in the series of fractions $(nrs+nr)/(nrs-ns+s) \cdot (nrs+nr-r)/(nrs-ns+2s) \cdot (nrs+nr-2r)/(nrs-ns+3s) \ldots (nrs+r)/nrs$, which signify the ratio M/Λ, I suppose that the fraction in the mth position, namely $(nrs+nr-mr+r)/(nrs-ns+ms)$, is equal to $(s+1)/s$, and from this I find that $n = m + (mr-r)/(s+1)$, and hence that $nt = mt + (mrt-rt)/(s+1)$. This having been done, it is similarly shown, as before, that, when the binomial $r+s$ is taken to this power, its maximum term M exceeds the bound Λ more than $c(r-1)$ times. Consequently also, all the terms included between the maximum M and the bound Λ exceed by more than c times all the terms outside this bound, of which there are only $r-1$ times more. Thus finally, in the end, we conclude that when the binomial $r+s$ is raised to the power of which the index is equal to the larger of these two quantities, **[236]** $mt + (mst-st)/(r+1)$ and $mt + (mrt-rt)/(s+1)$, then the sum of the terms included between the two bounds L and Λ exceeds by much more than c times the sum of the terms beyond the bounds on both sides. Therefore a finite power has been found that has the desired property. Q.E.D.

Principal Proposition. Finally, there follows the proposition for the sake of which all this has been said, but whose demonstration can now be given with only the application of the foregoing lemmas. To avoid tedious circumlocution, I will call the cases in which a certain event can happen *fecund* or *fertile.* I will call *sterile* those cases in which the event can not happen. I will also call experiments *fecund* or *fertile* in which one of the fertile cases is discovered to occur; and I will call *nonfecund* or *sterile* those in which one of the sterile cases is observed to happen. Let the number of fertile cases and the number of sterile cases have exactly or approximately the ratio r/s, and let the number of fertile cases to all the cases be in the ratio $r/(r+s)$ or r/t, which ratio is bounded by the limits $(r+1)/t$ and $(r-1)/t$. It is to be shown that so many experiments can be taken that it becomes any given number of times (say c times) more likely [*verisimilius*] that the number of fertile observations will fall between these bounds than outside them, that is, that the ratio of the number of fertile to the number of all the observations will have a ratio that is neither more than $(r+1)/t$ nor less than $(r-1)/t$.

between the maximum M and the bound L will exceed more than $c(s-1)$ times the same number of terms starting from the largest terms outside this bound. Similarly, they will exceed more than c times that many terms taken $s-1$ times. Therefore, even more obviously, they will exceed more than c times all the terms outside the bound L, of which there are only $s-1$ times as many more.

For the terms to the right, I proceed in the same way. I take the ratio $(s+1)/s < (rs+r)/(rs-s)$, I set $(s+1)^m/s^m \geq c(r-1)$, and I find that $m \geq \log [c(r-1)]/[\log(s+1) - \log s]$. Next, in the series of fractions $(nrs+nr)/(nrs-ns+s) \cdot (nrs+nr-r)/(nrs-ns+2s) \cdot (nrs+nr-2r)/(nrs-ns+3s) \ldots (nrs+r)/nrs$, which signify the ratio M/Λ, I suppose that the fraction in the mth position, namely $(nrs+nr-mr+r)/(nrs-ns+ms)$, is equal to $(s+1)/s$, and from this I find that $n = m + (mr-r)/(s+1)$, and hence that $nt = mt + (mrt-rt)/(s+1)$. This having been done, it is similarly shown, as before, that, when the binomial $r+s$ is taken to this power, its maximum term M exceeds the bound Λ more than $c(r-1)$ times. Consequently also, all the terms included between the maximum M and the bound Λ exceed by more than c times all the terms outside this bound, of which there are only $r-1$ times more. Thus finally, in the end, we conclude that when the binomial $r+s$ is raised to the power of which the index is equal to the larger of these two quantities, **[236]** $mt + (mst-st)/(r+1)$ and $mt + (mrt-rt)/(s+1)$, then the sum of the terms included between the two bounds L and Λ exceeds by much more than c times the sum of the terms beyond the bounds on both sides. Therefore a finite power has been found that has the desired property. Q.E.D.

Principal Proposition. Finally, there follows the proposition for the sake of which all this has been said, but whose demonstration can now be given with only the application of the foregoing lemmas. To avoid tedious circumlocution, I will call the cases in which a certain event can happen *fecund* or *fertile.* I will call *sterile* those cases in which the event can not happen. I will also call experiments *fecund* or *fertile* in which one of the fertile cases is discovered to occur; and I will call *nonfecund* or *sterile* those in which one of the sterile cases is observed to happen. Let the number of fertile cases and the number of sterile cases have exactly or approximately the ratio r/s, and let the number of fertile cases to all the cases be in the ratio $r/(r+s)$ or r/t, which ratio is bounded by the limits $(r+1)/t$ and $(r-1)/t$. It is to be shown that so many experiments can be taken that it becomes any given number of times (say c times) more likely [*verisimilius*] that the number of fertile observations will fall between these bounds than outside them, that is, that the ratio of the number of fertile to the number of all the observations will have a ratio that is neither more than $(r+1)/t$ nor less than $(r-1)/t$.

conclusion

"Il est remarquable qu'une science qui a commencé par la considération des jeux, se soit élevée aux plus importans objets des connaissances humaines."

Pierre-Simon de Laplace:
Essai philosophique sur les probabilités (1814)

["It is remarkable that a science which began with the consideration of games of chance should have raised itself to the most important objects of human knowledge."]