ELEMENTS

OF

PROBABILITY CALCULUS

by

DIOGO PACHECO D'AMORIM

COIMBRA
Imprensa da Universidade
1914

ELEMENTOS

DE

CÁLCULO DAS PROBABILIDADES

POR

DIOGO PACHECO D'AMORIM

COIMBRA
Imprensa da Universidade
1914

ELEMENTS

OF

PROBABILITY CALCULUS

Editors' Foreword

In 1914, one year after graduating in Mathematics, Diogo Pacheco d' Amorim submitted his doctoral thesis *Elements of Probability Calculus* to the University of Coimbra. In the author's Preface it is said that "the title — *An Essay Towards Rationalizing Probability Calculus* — would perhaps be more appropriate".

In fact, Pacheco d' Amorim's endeavour seems to be an attempt to build axiomatic Probability Theory, as requested by Hilbert in his seminal address to the Paris 1900 International Congress of Mathematicians, avoiding embarrassing paradoxes, as exposed in Bertrand's *Calcul des Probabilités*.

To fulfill this task, he first defines the *standard model*: someone performs the selection of an element from a sample space which is qualitatively and quantitatively known (for instance, extraction of one ball from one urn where n_1 balls of colour $C_1, \ldots, n_r$ balls of colour C_r have been thoroughly mixed). This individual knows whether his extraction has been random or not (i.e., in the above example, whether any ball in the urn has or hasn't the same chance of being selected). Thus, in this standard model, contrarily to Poincaré's opinion, the concept of *randomly selecting an element from the sample space* has a clear meaning for the one performing the extraction, which can be used as a "primitive" concept in the construction of probability. This leads to the concept of *equipossibility*.

The consideration of chains of hierarchically dependent extractions is then used to build up a wise and elegant solution to the main problem of constructing stochastic models in which elementary events are no longer equiprobable. For Pacheco d'Amorim probability is always conditional probability, and in some aspects is construction anticipates Rényi's work on the foundations of Probability.

In 1909, Borel had published a remarkable paper on continuous probability, that surely influenced Pacheco d'Amorim's construction of "randomly throwing geometric objects" in continuous sample spaces. In chapter III of his thesis, he gives a solution to one of the celebrated Bertrand's paradoxes (a solution that in our view has a serious flaw, cf. the editorial note (13)), and in chapter IV the discussion of "image points" — an ingenious construction of the probability measures of functions of random variables, lacking the concepts of random variables and of distribution functions — effectively solves another class of Bertrand's paradoxes, namely questions arising from using equiprobability models both for choosing a number in $[0, 100]$ and in $\left[0^2, 100^2\right]$.

Pacheco d'Amorim's believed that he had solved Bertrand's paradoxes in the standard model, in which the subject performing the extraction knows whether this was or wasn't done at random. His next step is an anticipation of pseudo-randomness: he deduces Bernoulli's and de Moivre's limit results, and from them he judges whether or not a (long enough) sequence of trials performed by someone else, or even by a mechanical device, imitates closely randomness. In the wealth of ideas discussed in the closing chapter, the main ideas of significance and hypothesis testing are clearly shaped.

Pacheco d'Amorim thesis is not a mature work, and there are some blunders in the text, that we discuss or at least unveil at the appropriate places. The long and cumbersome discussion of "random figures" is the weak point of this thesis, and we have been unable to understand clearly what the author meant in the last section of Chapter III (if you think that our translation is difficult to understand, you are right: we couldn't agree on the original's meaning). But, on the other hand, it has many strong points, it anticipates some influential ideas in Probability and Statistics, and surely deserves a fair opportunity to have international recognition.

In this translation, we corrected obvious typos (and we hope we didn't introduce other typos); figures have been redrawn, and we adopted symbols that, in our view, improve the readability of the text.

We are thankful to Prof. José Pacheco d'Amorim, who authorized this edited translation of his father's thesis.

Sandra Mendonça, Dinis Pestana, Rui Santos

Lisbon, 2007 August 08

ELEMENTOS

DE

CÁLCULO DAS PROBABILIDADES

POR

DIOGO PACHECO D'AMORIM

COÍMBRA
IMPRENSA DA UNIVERSIDADE
1914

ELEMENTS

OF

PROBABILITY CALCULUS

by

DIOGO PACHECO D'AMORIM

COIMBRA
Imprensa da Universidade
1914

A meus Pais

To my Parents

Dissertação inaugural para o acto de doutoramento na Faculdade de Matemática, na Universidade de Coímbra.

Inaugural Dissertation for the act of Doctor of Philosophy in the Faculty of Mathematics, of the University of Coimbra.

PREFÁCIO

O presente volume que mais pròpriamente se poderia chamar — *Uma tentativa de racionalização do Cálculo das Probabilidades* — põe em especial relêvo uma proposição a que, até hoje, ninguem deu a importância devida — a proposição *tirar, à sorte, um elemento duma classe,* ou *lançar, à sorte, um ponto numa região.*

HENRIQUE POINCARÉ diz mesmo que ela, por si, não tem significação nenhuma [1]. Ora a verdade é que ela tem um sentido muito preciso e claro quando nós mesmos sômos os agentes das tiragens ou lançamentos e isso permite-nos construir a teoria das probabilidades com toda a clareza e rigor. Partindo dela, a teoria das probabilidades pode reduzir-se a uma sucessão de proposições e definições, como qualquer outro ramo das Matemáticas Puras; a probabilidade contínua e a probabilidade descontínua aparecem com feições em tudo idênticas; os paradoxos desaparecem.

---

[1] H. POINCARÉ, *La Science et l'Hypothèse*, pag. 226.

PREFACE

This volume, for which the title — *An Essay Towards Rationalizing Probability Calculus* — would perhaps be more appropriate, gives an outstanding rôle to a concept that, until now, never got the relevance it deserves — the concept of *extracting, at random, an element from a set* or of *throwing, at random, a point in a region.*

Henri Poincaré [1] goes as far as saying that such a statement has, by itself, no meaning. But the truth is that this proposition has a very clear and precise meaning for the agent of the random extraction or of the random throw, and this allows us to construct the theory of probability with clarity and rigor. Starting from this primitive concept, the theory of probability can be reduced to a systematic sequence of propositions and definitions, as any other branch of pure mathematics. In this approach, discontinuous and continuous probability are identical in all aspects, and paradoxes have no place in the ensuing theory.

[1] H. POINCARÉ, *La Science et l'Hypothèse*, p. 226.

Construida a teoria da probabilidade dos fenómenos correspondentes a tiragens ou lançamentos, feitos à sorte, por nós mesmos, é fácil de estender as suas conclusões aos fenómenos correspondentes a tiragens, ou lançamentos, feitos por agentes a nós semelhantes, caso essas tiragens sejam feitas em certas circunstâncias.

Para aplicar os resultados da teoria, assim construida, aos fénómenos naturais, teremos de regeitar, *à priori*, a hipótese determinista que é, aliás, incompativel com qualquer teoria das probabilidades; em seguida, a aplicação faz-se muito naturalmente.

*

O ponto de vista em que nos colocamos, levou-nos a alterar, em certos pontos, a fórma e mesmo a essência desta sciência. Tivemos de generalizar a definição de probabilidade, generalização essa já necessária na dedução da fórmula de Bayes, mas absolutamente indispensável na probabilidade contínua, como claramente mostra o prob. 3.º da pag. 48.

Tivemos de fazer a distinção entre a probabilidade dum ponto e a probabilidade dum outro que seja *sua imagem*, o que nos levou ao conceito da *lei da probabilidade*, etc.

A ordem das matérias também teve de sair das normas clássicas.

Once the theory of probability of random extractions and of random throws done by ourselves has been built, its extension to phenomena whose outcomes are similar to extractions or throws performed by agents similar to us is rather easy, in case the extractions are done under some rigid circumstances.

The theory thus constructed can be applied to the study of natural phenomena, insofar as we reject, *a priori*, the determinist hypothesis, that, in fact, is incompatible with probability theory; under this *proviso*, the application is easily done.

*

The perspective we have adopted led us to change the form and the essence of Probability. We had to generalize the definition of probability, a generalization needed to prove Bayes formula, and absolutely unavoidable in the study of continuous probability, as we can see in problem 3, page 48.

We had to distinguish the probability of one point from the probability of another point which is the *image* of the first one, and from this emerged the concept of *probability law*, etc.

The order of presentation couldn't, therefore, conform to the classical one.

A probabilidade contínua passou a ocupar o lugar e o desenvolvimento que lhe competia ao lado da probabilidade discontínua. Os Teorêmas de BERNOULLI seguem-se-lhes na exposição, porque se aplicam igualmente a ambas as probabilidades. A seguir ao 3.º teorêma de BERNOULLI colocamos umas generalizações do mesmo teorêma, feitas para dar à dedução da lei dos desvios que naturalmente se lhes segue, o rigôr que lhe faltava. A seguir colocâmos a teoria da *Esperança Matemática,* porque o valôr deste conceito se torna muito mais evidente com a aplicação do 3.º teorêma de BERNOULLI, do que com a propria definição de esperança matemática. Por fim, alargamos o campo de acção desta sciência, fazendo a sua aplicação a fenómenos de que nós não sômos os agentes. Reservamos também para o fim a classificação dos fenómenos de que se ocupa esta sciência, porque só depois do estudo dos fenómenos que directamente se ligam com o fenómeno tipo, apresentado na Introducção, se torna clara e racional essa classificação. Tencionavamos terminar com uma justificação da nossa maneira de considerar a teoria da probabilidade, seguida dum Apêndice onde se estudasse a probabilidade dos conjuntos *numeráveis,* depois de *numerados;* mas a já demasiada extensão desta dissertação impede-nos de o fazer.

Continuous probability in this book is presented in parallel with discontinuous probability, and with the development it deserves. Bernoulli's theorems are a natural follow up, since they can be applied to both discontinuous and continuous probability. After Bernoulli's 3rd theorem, we present some variants and extensions, necessary to establish the error law with the rigorous demonstration its usual presentation lacks. We next develop the theory of *Mathematical Expectation*, since the importance of this concept is more evident with the application of Bernoulli's 3rd theorem than with the definition of mathematical expectation itself. Finally, we broaden the scope of applications of Probability, dealing with phenomena of which we are not the agents. We also postpone until the end a classification of the phenomena that are the object of this science, since we believe that the classification is clear and rational after a deep understanding of how Probability deals with the standard phenomenon, discussed in the Introduction, and developed in the core of this thesis. I had in mind to finish with a justification of our concept of probability, and to add an Appendix developing the study of probability in denumerable sets; but the unusual extension of the present dissertation persuaded me to postpone the publication of these matters.

INTRODUÇÃO

INTRODUCTION

INTRODUÇÃO

O Cálculo das Probabilidades, como aliás toda e qualquer sciência, tem por fim relacionar factos que supomos conhecidos com outros desconhecidos mas susceptíveis de serem relacionados com os primeiros.

Vejamos, com um exemplo, os factos que supomos conhecidos nesta sciência e porque é que os consideramos como tais.

Suponhamos, para isso, uma urna contendo bolas, idênticas em tudo menos na côr.

Nestas condições, podem dar-se os seguintes casos:

1.º desconhecermos, por completo, as côres das bolas que a urna contêm e, portanto, as suas percentagens respectivas;

2.º conhecermos as côres das bolas e desconhecermos as suas percentagens;

3.º conhecermos as côres das bolas e as suas percentagens.

Tire-se, à sorte, uma bola dessa urna e suponhamos que somos forçados a apostar numa côr, à nossa escolha.

Em que côr apostar?

No primeiro caso, não o saberemos. Visto que desconhecemos, por completo, o conteúdo da urna, não teremos

INTRODUCTION

The aim of Probability Calculus, as of any other science, is to find associations relating known facts to other facts that, although being unknown, can be related to the former ones.

We begin by an example, illustrative of what we consider our [degree of] knowledge of the facts.

Suppose that one urn contains balls, identical in all aspects save, eventually, in their color.

There are three possible situations:

1. we do not know the colors of the balls in the urn, and therefore we do not know the percentage of each color, as well;
2. we know the colors of the balls [for instance, there are white balls and there are black balls], but we ignore the percentage of the balls of each color;
3. we know the colors, and the percentage of balls of each color in the urn.

A ball will be randomly extracted from the urn, and we have to bet on the color of the ball.

What bet should we choose?

In the first case, the question doesn't make sense. As we do not know anything about the colors present in the urn, there is no razão nenhuma para apostar na côr branca ou em qualquer outra.

No segundo caso, a nossa ignorância é um pouco mais restrita, visto que sabemos que a bola a tirar ou é branca ou é preta.

Mas como desconhecemos as percentagens das bolas brancas e pretas, ainda não temos razões que nos decidam por uma ou outra destas côres.

Suponhamo-nos, porêm, no terceiro caso e para fixar ideias, suponhamos que na urna há uma percentagem de nove décimos de bolas brancas, para um décimo de pretas.

Nestas condições não hesitaremos em apostar na côr branca.

Evidentemente que nada sabemos de certo acêrca da côr da bola que vai saír que poderá ser branca ou preta, mas nem por isso hesitaremos na aposta.

É isto que distingue essencialmente o terceiro caso que acabâmos de considerar, dos dois antecedentes: êle pode servir-nos de guia de conduta em certas circunstâncias.

Por isso, supô-lo-hemos *conhecido*.

*

Êle consiste numa tiragem, feita à sorte, numa urna de composição conhecida qualitativa e quantitativamente.

Nós suporemos, pois, que todo o fenómeno que pode identificar-se com uma tiragem, feita à sorte, numa urna de composição qualitativa e quantitativamente conhecida, fica explicado logo que essa identificação seja feita. Dum modo mais geral, nós suporemos explicado todo o fenóreason whatsoever to prefer any color to bet in.

In the second case, our ignorance has been moderated, since we know that the ball that will be extracted can be either white or black.

But as we still ignore the percentage of balls of each color, there are no grounds to decide which bet to take.

On the other hand, in the third case, assuming for instance that we know that 90% of the balls are white, we would surely decide to bet that a random extraction would produce white ball.

Obviously, we do not know for sure the color of the ball that will be extracted, it can be black or white, but we do not hesitate in choosing white as the sensible bet.

This distinguishes the third case from the former ones. It can serve as an example on how to take rational decisions with incomplete information.

For this reason we shall say that the third case describes a *known* urn.

*

The third case deals with random extractions from one urn whose composition is qualitatively and quantitatively known.

We shall assume that any phenomenon whose outcomes can be identified with random extractions of balls from one urn of qualitatively and quantitatively known composition is explained once that identification has been made. More generally, we consider explained any phenomenon meno que possa identificar-se com uma tiragem, feita à sorte, numa classe finita e qualquer, logo que essa classe seja conhecida qualitativa e quantitativamente.

*
* *

Como vimos, supunhamos conhecido o terceiro caso atrás considerado, por êle nos poder servir de guia de conduta em certas circunstâncias que nós sintetisámos numa aposta.

Vejamos, porêm, quais as razões que nesse exemplo nos levavam a apostar na côr branca.

A primeira era, sem dúvida, o conhecimento de que na urna havia mais bolas brancas do que pretas, ou, como atrás dissémos, que a percentagem das bolas brancas era superior à das pretas.

A segunda era o conhecimento de que a tiragem em questão era feita à sorte.

Sem estas duas condições não teriamas razões para apostar na côr branca.

A explicação da primeira destas condições é-nos dada pela Aritmética; a explicação da segunda, porêm, não nos é dada por sciência nenhuma até hoje constituida.

Ela leva-nos, por isso, a admitir como primitiva a seguinte proposição:

«Tirar, à sorte, um elemento duma classe finita de elementos».

É sôbre esta proposição que nós assentaremos as bases do Cálculo das Probabilidades.

which can be identified with a random selection of elements in a finite set, qualitatively and quantitatively known.

*

* *

As we have seen, we have distinguished the third case as known, because it can serve as a *standard model* on how to take decisions under uncertainty, i.e., under circumstances that we synthesized in the form of taking a bet.

Let us analyze in more detail the reasons that led us, in that example, to bet in white color.

The first reason was, indeed, the fact that we knew that more white balls than black balls existed in the urn, or, as stated in the example, the percentage of white balls was larger than the percentage of black balls.

The second one was the knowledge that the extraction was performed at random.

If one of these assumptions is withdrawn, there is no rationale for choosing to bet "white ball".

The reason why the first condition is an argument in favor of betting in white ball comes from Arithmetic; the explanation why the second condition is needed can be found only in the emerging science of Probability.

In that science we will therefore take the statement

"to extract an element, at random, from a finite set"

as a primitive concept in this branch of Mathematics.

We shall build Probability Calculus starting from this primitive concept.

*
* *

É conveniente notar que não tomâmos para base desta sciência a noção de *sorte* ou de *acaso*, noção esta demasiado vaga para sôbre ela assentar as bases de qualquer sciência, mas sim a proposição «tirar, à sorte, um elemento duma classe finita de elementos», o que é muito diferente.

Alguem dirá que esta proposição, visto ser formada com a palavra *sorte* e dela tirar todo o seu valor, fica tão vaga como a noção que essa palavra indica.

A isso responderemos dizendo que pouco interesse tem que a noção de *sorte* seja ou não vaga, logo que a proposição que com ela formamos seja entendida e com ela se exprima uma ideia capaz de nos orientar em dadas circunstâncias.

Tudo quanto acêrca dessa proposição se diga, não tem interesse nem debaixo do ponto de vista matemático nem debaixo do ponto de vista utilitário.

O mesmo acontece em Geometria com a noção de *espaço* e em Mecânica com a noção de *tempo*.

As questões levantadas à volta destas noções teem um mero interesse filosófico, nada mais. Pode mesmo dizer-se que estas noções nada teem que vêr com a Matemática.

O edifício matemático *podia* ser o mesmo se essas noções não existissem.

As verdades da Geometria, da Mecânica, ficariam latentes no simbolismo da Análise. Mas, embora invisíveis, essas verdades lá estavam.

As noções de espaço e de tempo servem apenas, como os reagentes córantes na microscópia, para dar às ver-

*

* *

It is worth observing that we didn't choose the concept "randomness" or "uncertainty" as the primitive concept upon which the theory of Probability would be constructed, since these concepts are vague, and as such inadequate to serve as the foundation for any science; our choice has been quite different, the concept of "extracting an element, at random, from a finite set".

Some could accuse us of using a foundation as vague as the concept of "randomness", since this concept is used in our primitive statement.

However, in the statement we choose as primitive, it is immaterial whether the formulation "*at random*" is or isn't vague, insofar as the proposition using it can be understood and expresses an idea that can guide our choices and decisions under precise circumstances.

Whatever we say about this proposition is irrelevant either from the mathematical viewpoint or in the perspective of applications.

The same could be said about the concepts of *space* in Geometry, or of *time* in Mechanics.

The discussion of these concepts is irrelevant in Mathematics, they are from the scope of Philosophy. Mathematics *would* be the same theoretical construct if these concepts didn't exist. The knowledge of what we consider Geometry and Mechanics would be latent in the symbolism of Mathematical Analysis, and no more, but this knowledge would still be valid, although less visible.

The usefulness of the concepts "space" and "time" can be compared to the usefulness of coloring reagents dades analíticas uma coloração que as torne visíveis nos vários campos em que podem ser aplicados.

*
* *

Uma pregunta há, porêm, que é legitímo de fazer-se: como distinguir uma tiragem feita à sorte, doutra que o não é?

Evidentemente que nem todas as tiragens feitas numa urna são feitas à sorte e forçoso é, por isso, dar um critério de distinção.

Para se dar êsse critério, admitiremos que todo o indivíduo sabe distinguir uma tiragem feita à sorte, doutra que o não é, logo que *êsse mesmo indivíduo seja o agente da tiragem.*

*

Dentro desta hipótese nós construiremos a teoria das probabilidades que ficará sendo uma sciência subjectiva, como toda a sciência pura. Essa teoria dar-nos-há depois o critério que nos deve guiar quando o agente da tiragem seja outrem que não nós.

Mas essa parte do Cálculo das Probabilidades está já no limiar das suas aplicações.

*
* *

A utilidade das sciências, em geral, provêm de nos ensinarem a prever os acontecimentos com uma aproximação práticamente suficiente.

in Chemistry: they enhance the visibility of the phenomena, but these exist independently of being or not being enhanced by the coloring reagent.

*

*   *

An important question must be raised at once: how can we distinguish between random and non-random extractions?

It is obvious that there are extractions that are non-random, and therefore we need a criterion to distinguish random from non-random extractions.

To construct such criterion, we shall assume that any individual knows whether an extraction has been made at random if *the extraction has been made by him.*

*

Under this assumption, we shall build up a theory of probability, which is a subjective science, as all pure science is. This theory will allow us to construct a criterion to distinguish between random and non-random extractions, when we are not, ourselves, the agent performing the extraction.

This area of Probability Calculus is at the onset of applications.

*

*   *

The usefulness of science is its general ability to forecast events with an approximation considered good enough in practical applications.

O Cálculo das Probabilidades, porêm, parece não atingir êste fim utilitário.

Com efeito, como prever a côr duma bola que se tire à sorte duma urna que contêm duas bolas brancas e uma preta ?

Evidentemente que o Cálculo das Probabilidades no-lo não diz.

Se em vez de duas bolas brancas e uma preta, a urna contivesse mil bolas brancas e uma preta, o Cálculo das Probabilidades continuava a nada prever, mas a nossa intuição principiava a prever a chegada duma bola branca.

É nessa previsão que a intuição nos sugére quando o número de bolas brancas é muito maior que o das pretas que reside o valor prático do Cálculo das Probabilidades.

Essa intuição ficará irredutívelmente separada da certeza, por mais que da unidade se apróxime a percentagem das bolas brancas. Mas nem por isso ela deixa de ter para nós um valor prático real.

*

O que se disse duma tiragem feita, à sorte, numa classe finita de elementos, dir-se-há também do lançamento *dum ponto,* feito à sorte, numa região finita, a um número qualquer de dimensões.

This pragmatism seems unfeasible in Probability Calculus.

In effect, how could we predict the color of the ball that will be extracted from one urn containing two white balls and one black ball?

It is obvious that Probability Calculus is unable to make an useful prediction, in this situation.

If instead of two white balls and one black ball, the urn composition was one thousand white balls and one black ball, prediction of the outcome of a random extraction would still be impossible, but to our intuition it would seem more plausible to forecast that a white ball would be extracted.

The practical usefulness of Probability Calculus lies in this evaluation of the degree of probability of a future event, and in the ensuing confidence that our intuition attaches to the plausibility of events whose probability approximates certitude.

Confidence based in probability will, in its essence, be different from certitude, no matter how nearly the percentage of white balls in the urn approximates 1. But this doesn't deface the real practical usefulness of Probability in decision making under incomplete or unreliable information.

*

What we have said about random extractions of elements from a finite set can also be said about randomly throwing *points* in a bounded region of space, in any number of dimensions.

CAPÍTULO I

CLASSES FINITAS

CHAPTER I

FINITE SETS

CAPÍTULO I

CLASSES FINITAS

Com o símbolo (A) representaremos uma classe com um número qualquer de elementos.

Se (A) e (B) forem classes, com o símbolo (A, B) representaremos a classe que se obtêm associando cada elemento de (A) com cada um dos elementos de (B).

O símbolo (A) representará tambem o número de elementos da classe (A).

*

Segundo esta notação, será

$$(A, B) = (A) . (B)$$

como é evidente.

*

Qualquer elemento de (A, B) diz-se *composto* de A e B e representa-se por (A B).

A classe (A, B) diz-se *composta* das classes (A) e (B). Uma classe qualquer de elementos compostos (que pode não ser uma classe composta) representá-la-hemos por $(A\ B)$.

CHAPTER I

FINITE SETS

We shall denote A, $B, \ldots$ sets with a finite number of elements.

The symbol $A \times B$ will denote the set of ordered pairs (a, b), obtained from the sets A and B, by associating each $a \in A$ with each $b \in B$.

The symbol $\#A$ denotes the cardinal of the set A.

$*$

With these notations, it is obvious that

$$\#A \times B = \#A \times \#B.$$

$*$

Each ordered pair $(a, b) \in A \times B$ is said to be *compound* of a and b.

The set $A \times B$ is *compound* from the sets A and B. Any set composed of compound elements (a, b) is not, necessarily, a compound set.

Proposição primitiva

a)

A proposição *tirar, à sorte, um elemento da classe* (A) consideramo-la como primitiva, isto é, como tendo um sentido próprio e não precisando, por isso, de explicação.

b)

A proposição A *é um elemento tirado, á sorte, da classe* (A) tem o mesmo sentido que a proposição *a*), mas presta-se melhor para o símbolismo da lógica matemática, enquanto que *a*) se presta melhor para o uso da linguagem vulgar.

*

Segundo a hipótese que acabamos de fazer, as proposições «tirar, à sorte, uma carta dum baralho», «lançar, à sorte, um dado» (tirar, à sorte, uma face dum dado), «tirar, à sorte, uma bola de uma urna», etc., não precisam de explicação.

DEFINIÇÃO 1.ª

Tirar, à sorte, um elemento de (A) *ou* (B) *ou* (C)..., significa, por definição, *tirar, à sorte, um elemento da classe* $(A+B+C\ldots)$, representando $(A+B+C+\ldots)$ a classe constituída pela totalidade dos elementos das classes (A), (B), (C)...

Primitive concept

a)

We consider the statement *to extract, at random* [or *to select*], *an element from the set* A as having a self evident meaning, and henceforth needing no further explanation; in other words, *to select, at random*, an element from a finite set is considered a primitive concept.

b)

The statement *a is a randomly chosen element from the set* A has the same meaning; *b*) is better suited to the formal symbolism of mathematical logic, while *a*) is more appropriate for the natural language.

$*$

From the above assertions, the propositions "randomly extracting a card from a card deck", "random throw of a die" (random selection of one die face), "randomly extracting a ball from an urn", etc., do not need further explanation.

DEFINITION 1

Randomly extracting an element from A*, or* B*, or* $C, \ldots$, is the same as *randomly taking an element from* $A \cup B \cup C \cup \cdots$, the set having all the elements from the sets A, B, C, ...

DEFINIÇÃO 2.ª

a)

Tirar, à sorte, um elemento de (A) *e outro de* (B) significa, por definição, *tirar, à sorte, um elemento de* (A, B).

b)

Tirar, à sorte, um elemento de (A), *outro de* (B) *e outro de* (C) significa, por definição, *tirar, à sorte, um elemento de* (A, B) *e outro de* (C), etc.

*

Segundo esta definição, *tirar, à sorte, um naipe e tirar à sorte um número* [1] significa, *tirar, à sorte, uma carta dum baralho* e recíprocamente.

DEFINIÇÃO 3.ª

a)

Seja (A) uma classe, a cada elemento da qual se faz corresponder outra classe (B), variável, em geral, de elemento para elemento.

Nestas condições, *tirar, à sorte, um elenento de* (A) *e outro na classe* (B) *correspondente ao elemento tirado de* (A) significa ainda, por definição, *tirar, à sorte, um elemento de* (A, B).

---

[1] Chamaremos *número* a qualquer carta, abstração feita do naipe.

DEFINITION 2

a)

Randomly extracting an element from A *and,* [*independently,*] *another from* B is, by definition, the same as *randomly extracting an element from* $A \times B$.

b)

Randomly extracting an element from A*, another from* B *and another from* C, [*the extractions being mutually independent*] is, by definition, the same as *randomly extracting an element from* $A \times B$ *and another from* C, etc.

*

According to this definition, *randomly choosing a suit and then randomly choosing a number*[(2)]*,* [*independently,*] is the same as *randomly choosing a card from the card deck.*

DEFINITION 3

a)

Let us associate to each $a \in A$ a set B_a, and denote $\{a\} \times B_a$ the set of ordered pairs $\{(a, b) : \; b \in B_a\}$.

Randomly extracting an element from A *and another element from the corresponding set* B_a is, by definition, the same as *randomly extracting an element* (a, b) *from* $A \times B_a$.

[(2)] In this context the *numbers* are 1 or ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, knave, queen, king, i.e. the card value, whichever the suit.

b)

Se a cada elemento de (B) se fizer corresponder uma nova classe (C), *tirar, à sorte, um elemento de* (A), *outro da classe* (B) *correspondente a* A *e outro da classe* (C) *correspondente a* B, significa, *tirar, à sorte, um elemento de* (A, B) *e outro da classe* (C) *correspondente ao elemento* B, *tirado em* (B), etc.

Possibilidade

1.º — Elementos possíveis

Segundo as definições que acabamos de dar, ou bem se trata de tiragens, feitas à sorte, numa só classe (proposição primitiva, def. 1.ª e 2.ª), ou se trata de tiragens feitas, à sorte, num complexo de classes [def. 3.ª, *a*) *b*)].

Tudo depende do sistema de tiragens que se considerar e das classes em que elas se efectuarem.

No caso das tiragens serem feitas numa só classe ou de tudo se passar como se tal se desse (prop. primit., def. 1.ª e 2.ª), dizem-se *possíveis* todos os elementos dessa classe.

No caso das tiragens serem feitas num complexo de classes, como na prop. 3.ª *a*), dizem-se *possíveis* todos os elementos que se obteem associando cada elemento A de (A) com os elementos da classe (B) que lhe corresponde.

No caso da prop. 3.ª *b*), dizem-se *possíveis* todos os elementos que se obteem associando cada elemento A B, *possível* em relação às duas primeiras tiragens, com cada um dos elementos da classe (C) correspondente a B, etc.

b)

If to each $b \in B_a$ we associate a set C_b, *randomly extracting an element from A, another element from the corresponding set B_a and another element from the corresponding set C_b* is, by definition, the same as *randomly extracting an element (a, b, c) from $A \times B_a \times C_b$*.

Possibility

1 — Possible elements

According to the above definitions, random extractions have a meaning either in a single set (primitive concept, definitions 1 and 2) or in a complex of sets [definition 3, *a*) and *b*)].

All depends on the extracting system, and on the sets from where the extractions are performed.

When the extractions are performed from a single set, or performed in such a way that they are equivalent to extractions from a single set (primitive concept, definitions 1 and 2), we say that all the elements from that set are *possible*.

When the extractions are sequentially performed from a complex of sets, as explained in definition 3 *a*), we say that the *possible* elements are those in $A \odot\!\!\rightarrow B = \bigcup_{a \in A} \{a\} \times B_a$.

On the other hand, in what concerns definition 3 *b*), the *possible* elements are those that can be sequentially extracted randomly choosing $a \in A$, and then randomly choosing one element $b \in B_a$, and next randomly choosing an element $c \in C_b$, i.e., the elements from the complex of sets $A \odot\!\!\rightarrow B \odot\!\!\rightarrow C = \left\{ (a, b, c) \in \bigcup_{(a,b) \in A \odot\!\!\rightarrow B} \{(a, b)\} \times C_b \right\}$, etc.

2.º — Classes possiveis

Uma classe diz-se *possível* quando é constituída por elementos *possíveis.*

Á classe constituída pela totalidade de elementos possíveis, chamaremos *classe total possível.*

DEFINIÇÃO 4.ª

Representando por (A) o número de elementos da classe (A), ao número

$$\pi_A = \frac{1}{(A)}$$

chamarei *possibilidade dum elemento* A qualquer, tirado à sorte, de (A), ou ainda, *possibilidade por unidade.*

*

Os elementos provenientes de tiragens efectadas numa mesma classe, ou a tal equivalentes, serão *egualmente possíveis.*

Proposição I

A possibilidade dum elemento composto é egual ao produto das possibilidades dos elementos componentes.

2 — Possible sets

The *total possible set* A [resp. B, $A \times B$, $A \odot\!\!\!\rightarrow B$, etc.] is the set with all possible elements.

Any $A' \subset A$ is a *possible set*, i.e. is a set whose elements are possible.

DEFINITION 4

The *possibility* of a randomly chosen element $a \in A$ (or in any of its possible subsets), or *unit possibility*, is

$$\pi_a = \frac{1}{\#A}\,.$$

*

Thus, all elements randomly chosen in the same set (or randomly chosen using an extracting system which is equivalent to random extraction from the same set) are equally possible.

Proposition I

The possibility of a compound element $(a, b) \in A \times B$ is the product of the possibilities of its components.

Com efeito, visto que

$$(A, B) = (A) \, . \, (B),$$

será

$$\frac{1}{(A, B)} = \frac{1}{(A)} \cdot \frac{1}{(B)}$$

e por isso,

$$\pi_{AB} = \pi_A \, . \, \pi_B$$

c. d. d.

DEFINIÇÃO 5.ª

Chama-se possibilidade duma classe possível (A), à soma das possibilidades dos seus elementos

$$\varpi_{(A)} = \sum_{(A)} \pi_{(A)}$$

Proposição II

Se (A) é uma classe possível, tal que

$$(A) = (A_1) + (A_2) + \ldots + (A_n),$$

será,

$$\varpi_{(A)} = \varpi_{(A_1)} + \varpi_{(A_2)} + \ldots + \varpi_{(A_n)},$$

como resulta imediatamente da definição de $\varpi_{(A)}$.

This is an obvious consequence of $\#A \times B = \#A \times \#B$:

$$\frac{1}{\#A \times B} = \frac{1}{\#A} \times \frac{1}{\#B}\,,$$

and thus

$$\pi_{(a,b)} = \pi_a \times \pi_b.$$

$\square$

DEFINITION 5

The possibility $\varpi_{A'}$ of a possible set A' is the sum of the possibility of its elements,

$$\varpi_{A'} = \sum_{a \in A'} \pi_a.$$

Proposition II

If A' is a possible set which may be partitioned into pairwise disjoint sets

$$A' = A'_1 \cup A'_2 \cup \cdots \cup A'_n$$

then

$$\varpi_{A'} = \varpi_{A'_1} + \varpi_{A'_2} + \cdots + \varpi_{A'_n}\,.$$

This is an immediate consequence of Definition 5.

Proposição III

A possibilidade da classe total possível é igual à unidade.

Com efeito:

a)

Se os elementos são todos provenientes de tiragens feitas numa só classe (A), a prop. é evidente:

$$\varpi_{(A)} = \sum_{(A)} \frac{1}{(A)} = \frac{(A)}{(A)} = 1.$$

b)

Consideremos, agora, o caso de os elementos serem provenientes de tiragens feitas num complexo de classes e, para fixar ideias, suponhamos que se trata do sistema de tiragens da def. 3.ª *a*).

Sejam A_1, A_2, ... A_n, os elementos de (A) e (B_1), (B_2), ... (B_n) as classes correspondentes. A possibilidade de qualquer elemento da classe que se obtêm associando A_1 com cada um dos elementos de (B_1) é (prop. I)

$$\pi_{A_1 B_1} = \frac{1}{(A)(B_1)}.$$

A possibilidade desta classe, será

$$\sum_{(B_1)} \pi_{A_1 B_1} = \frac{1}{(A)} \cdot \sum_{(B_1)} \frac{1}{(B_1)} = \frac{1}{(A)}.$$

Proposition III

The possibility of the total possible set is 1.

a)

If all the possible elements result from random extractions performed in the same set A, the proposition is obvious, since

$$\varpi_{A} = \sum_{a \in A} \frac{1}{\#A} = \frac{\#A}{\#A} = 1.$$

b)

Let us consider now sequential extractions from a complex of sets. Without loss of generality, consider the extraction system in definition 3 *a*).

Let $A = \{a_1, a_2, \ldots, a_n\}$, and denote $B_{a_k},\ k = 1, 2, \ldots, n$ the set associated with each element $a_k \in A$. From Proposition I, the possibility of any element resulting from pairing a_k with $b_j \in B_{a_k}$ is

$$\pi_{(a_k, b_j)} = \frac{1}{\#A \times \#B_{a_k}}\,,$$

and therefore the possibility of the set $a_k \times B_{a_k}$ is

$$\sum_{b_j \in B_{a_k}} \pi_{(a_k, b_j)} = \frac{1}{\#A} \sum_{b_j \in B_{a_k}} \frac{1}{\#B_{a_k}} = \frac{1}{\#A}\,.$$

A possibilidade da classe total será (prop. II)

$$\sum_{(A)} \frac{1}{(A)} = 1$$

c. d. d.

Esta demonstração estende-se fácilmente a qualquer sistema de tiragens.

Proposição IV

Se a classe (A, B) fôr composta das classes (A) e (B), será

$$\varpi_{(A,\,B)} = \varpi_{(A)} \cdot \varpi_{(B)}$$

visto que a possibilidade de cada elemento de (A, B) é o produto da possibilidade dum elemento de (A) pela possibilidade dum elemento de (B).

Probabilidade

DEFINIÇÃO 6.ª

Seja (A) uma classe possível e (A′) outra nela contida; chamaremos «probabilidade de (A′) em relação a (A)» ao número

$$P_{(A)}^{(A')} = \frac{\varpi_{(A')}}{\varpi_{(A)}}$$

sendo $\varpi_{(A')}$ e $\varpi_{(A)}$ as possibilidades respectivas de (A′) e de (A).

Thus, in view of Proposition II, the possibility of the total possible set is

$$\sum_{a\in A} \frac{1}{\#A} = 1.$$

$\square$

The above proof is easily extended for any complex extracting system.

Proposition IV

If the set $A \times B$ is compound from the sets A and B, then

$$\varpi_{A\times B} = \varpi_A \times \varpi_B,$$

since the possibility of each element (a, b) is the product of the possibility of an element of A by the possibility of an element of B.

Probability

DEFINITION 6

Let A' be a possible set and $A'' \subset A'$ another possible set(3). We shall call *probability of A'' relative to A'* the number

$$\mathbb{P}_{A'}(A'') = \frac{\varpi_{A''}}{\varpi_{A'}},$$

$\varpi_{A''}$ and $\varpi_{A'}$ denoting, as above, the possibilities of A'' and of A', respectively.

(3) We shall use, as a rule, $A'' \subset A' \subset A$.

Á classe (A′) costuma chamar-se *favorável* e à classe (A) — (A′), *classe contrária*.

Ás vezes emprega-se a palavra *caso* como sinónimo de *elemento*.

*

Se os elementos de (A) forem igualmente possíveis, será

$$\varpi_{(A)} = (A)\ \pi_A,\ \varpi_{(A')} = (A')\ \pi_A$$

e

$$P_{(A)}^{(A')} = \frac{(A')}{(A)}.$$

Logo: quando os elementos da classe possível forem igualmente possíveis, a probabilidade é dada pelo número de casos favoráveis divídido pelo número de casos possíveis.

*

Quando a classe possível fôr constituida pela totalidade dos elementos possíveis, será

$$\varpi_{(A)} = 1$$

e

$$P_{(A)}^{(A')} = \varpi_{(A')}\ (^1).$$

---

(1) A definição mais geral que Laplace deu da probabilidade, coìncide com êste caso particular ser $\varpi_{(A)} = 1$.

In the above context, the set A'' is said to be the *favorable* set, and $A' - A''$ is said to be the *unfavorable* or *contrary* set.

Sometimes we shall use the word *case* meaning *element*.

*

If the elements in A' are equally possible, it follows that

$$\varpi_{A'} = \#A' \times \pi_a , \qquad \varpi_{A''} = \#A'' \times \pi_a$$

and therefore

$$\mathbb{P}_{A'}(A'') = \frac{\#A''}{\#A'} .$$

In other words: When the elements in the possible set are equally possible, the probability is the number of favorable cases divided by the number of possible cases.

*

When the possible set A' is the total possible set A, from

$$\varpi_A = 1$$

it follows that

$$\mathbb{P}_A(A'') = \varpi_{A''}\ ^{(4)}.$$

(4) The most general definition of probability that can be found in Laplace is coincident with this particular case, of the reference set being the total possibility set A, with $\varpi_A = 1$.

Se a classe favorável fôr idêntica à classe possível, será

$$P = 1$$

e a probabilidade toma, neste caso, o nome de *certeza.*

Se a classe favorável fôr nula, diremos que

$$\varpi_{(A')} = 0$$

e, portanto,

$$P = 0.$$

A probabilidade toma, neste caso, o nome de *impossibilidade.*

A probabilidade varía, pois, entre os dois números limites, 0 e 1.

Postulado

Sejam S e S′ dois sistemas de tiragens, dando origem a elementos qualitativamente iguais. Nós diremos que estes dois sistemas são equivalentes, quando *classes* qualitativamente iguais tiverem em S e S′ as mesmas probabilidades.

*

A palavra *equivalente* empregada neste postulado, significa que tanto uma tiragem feita em S como uma feita èm S′, despertam em nós os mesmos móveis de acção.

Êste postulado reduz-nos todos os sistemas de tiragens que definimos, a tiragens feitas numa só classe.

If the favorable set A'' is the possible class A',

$$\mathbb{P}_{A'}(A') = 1$$

and in this case probability is *certitude*.

If the favorable set is empty, $A'' = \emptyset$,

$$\varpi_{\emptyset} = 0$$

and therefore

$$\mathbb{P}_{A'}(\emptyset) = 0.$$

In this case, probability is renamed *impossibility*.

Therefore, probability takes values between 0 and 1.

Postulate

Let S and S' be two extracting systems, originating qualitatively equal elements. We say that those two systems are equivalent if qualitatively equal sets have the same probability under S and S'.

*

The term *equivalent* in the above postulate means that similar extractions performed under S and under S' imply similar decisions.

This postulate reduces all extracting systems to extractions from a single set.

Proposição V

Da probabilidade total

Se a classe possível (A′) fôr constituída pelas classes parciais (A_1), (A_2), ... (A_n), de tal modo que

$$(A') = (A_1) + (A_2) + \ldots + (A_n),$$

será (prop. II)

$$\varpi_{(A')} = \varpi_{(A_1)} + \varpi_{(A_2)} + \ldots + \varpi_{(A_n)},$$

e, portanto

$$p_{(A)}^{(A')} = p_{(A)}^{(A_1)} + \ldots + p_{(A)}^{(A_n)}$$

isto é, *a probabilidade da classe total é igual à soma das probabilidades das classes parciais.*

Proposição VI

Da probabilidade composta

a)

Se (A, B) é uma classe possível, composta de (A) e (B) e (A′, B′) outra nela contida ([1]), será (prop. IV)

([1]) Em geral, com (A′) designaremos uma classe contida em (A).

Proposition V
Total probability

If the possible set A'' is partitioned pairwise disjoint partial sets $A''_1, A''_2, \dots, A''_n$,

$$A'' = A''_1 \cup A''_2 \cup \cdots \cup A''_n,$$

we have (Prop. II)

$$\varpi_{A''} = \varpi_{A''_1} + \varpi_{A''_2} + \cdots + \varpi_{A''_n}$$

and henceforth

$$\mathbb{P}_{A'}(A'') = \mathbb{P}_{A'}\left(A''_1\right) + \mathbb{P}_{A'}\left(A''_2\right) + \cdots + \mathbb{P}_{A'}\left(A''_n\right)$$

i.e., *the probability of the union of pairwise disjoint sets is the sum of the probabilities of the partial sets.*

Proposition VI
Compound probability

a)

If $A' \times B'$ is a possible set compound from A' and B', and $A'' \times B''$ is a possible subset of $A' \times B'$, we have (Prop. IV)

$$\varpi_{(A,B)} = \varpi_{(A)} \cdot \varpi_{(B)}$$

e

$$\varpi_{(A', B')} = \varpi_{(A')} \cdot \varpi_{(B')}$$

e, por isso,

$$P_{(A, B)}^{(A', B')} = P_{(A)}^{(A')} \cdot P_{(B)}^{(B')}.$$

Se as classes (A) e (B) forem independentes, esta proposição pode enunciar-se: *a probabilidade duma classe composta é igual ao produto das probabilidades das classes componentes.*

b)

A prop. VI ficou demonstrada para o caso das classes favorável e possível serem ambas compostas. Ela pode, porêm, generalizar-se para os casos seguintes:

1.º

Se

$$\varpi_{(A, B)} = 1,$$

isto é, se a classe possível fôr a classe total possível, será

$$\varpi_{(A)} = \varpi_{(B)} = \varpi_{(A)} \cdot \varpi_{(B)} = 1$$

e, portanto,

$$P_{(A, B)}^{(A', B')} = P_{(A)}^{(A')} \cdot P_{(B)}^{(B')}.$$

2.º

Se fôr

$$\varpi_{(A, B)} = \varpi_{(A)},$$

isto é, se a classe possível se obtiver da classe total pos-

$$\varpi_{A' \times B'} = \varpi_{A'} \times \varpi_{B'}$$

and

$$\varpi_{A'' \times B''} = \varpi_{A''} \times \varpi_{B''}$$

and therefore

$$\mathbb{P}_{A' \times B'}(A'' \times B'') = \mathbb{P}_{A'}(A'') \times \mathbb{P}_{B'}(B'').$$

In case the sets A' and B' are independent, this proposition may be stated as: *the probability of a compound set is the product of the probabilities of its components.*

b)

Proposition VI has been proved under the hypothesis that both the favorable and the possible sets are compound. It can, however, be generalized in the following ways:

1st

If $A' \times B' = A \times B$ is the total possible class, and thus

$$\varpi_{A' \times B'} = 1$$

we have

$$\varpi_{A'} = \varpi_{B'} = \varpi_{A' \times B'} = 1$$

and from this it follows that

$$\mathbb{P}_{A' \times B'}(A'' \times B'') = \mathbb{P}_{A'}(A'') \times \mathbb{P}_{B'}(B'').$$

2nd

If

$$\varpi_{A' \times B'} = \varpi_{A'},$$

i.e., if the possible set is obtained from the total possible set sível, pela exclusão de certos elementos de (A) com todos os elementos das classes (B) que lhes correspondem, será ainda

$$\varpi_{(B)} = 1,$$

e

$$P_{(A,\,B)}^{(A',\,B')} = P_{(A)}^{(A')} \cdot P_{(B)}^{(B')}.$$

Proposição VII

Sejam (A), (A′), (A″) classes tais que (A) contenha (A′) e (A′) contenha (A″). Teremos,

$$\frac{\varpi_{(A'')}}{\varpi_{(A)}} = \frac{\varpi_{(A')}}{\varpi_{(A)}} \cdot \frac{\varpi_{(A'')}}{\varpi_{(A')}};$$

logo,

$$P_{(A)}^{(A'')} = P_{(A)}^{(A')} \cdot P_{(A')}^{(A'')} \tag{1}$$

Corolário

De (1) tira-se

$$P_{(A')}^{(A'')} = \frac{P_{(A)}^{(A'')}}{P_{(A)}^{(A')}}.$$

Proposição VIII

Da probabilidade das causas

Quando as tiragens, à sorte, se efectuarem como ficou indicado na definição 3.ª, à classe (A) chama-se *causa* e às classes (B), *efeitos*.

by excluding some elements $a \in A$ together with all the elements from the corresponding sets B_a, from the fact that

$$\varpi_{B'} = 1$$

it follows that

$$\mathbb{P}_{A' \times B'}(A'' \times B'') = \mathbb{P}_{A'}(A'') \times \mathbb{P}_{B'}(B'').$$

Proposition VII

Let $A'' \subset A' \subset A$ be possible sets. As

$$\frac{\varpi_{A''}}{\varpi_A} = \frac{\varpi_{A'}}{\varpi_A} \times \frac{\varpi_{A''}}{\omega_{A'}}$$

it follows that

$$\mathbb{P}_A(A'') = \mathbb{P}_A(A') \times \mathbb{P}_{A'}(A''). \tag{1.1}$$

Corollary

From (1.1) it follows that

$$\mathbb{P}_{A'}(A'') = \frac{\mathbb{P}_A(A'')}{\mathbb{P}_A(A')}$$

Proposition VIII
On the probability of causes

When the random extractions are performed as described in Definition 3, the set A is the set of causes, and the sets $B = \bigcup_{a \in A} B_a$ is the set of effects.

O problema chamado *das probabilidades das causas,* tem por tipo o problema seguinte:

Consideremos um grupo de N urnas, das quais n_1, teem uma percentagem p_1 de bolas brancas, n_2, uma percentagem p_2, etc.

Tire-se, à sorte, uma urna de entre as N urnas, e da urna que saír, tire-se, à sorte, uma bola que, por hipótese, sáe branca.

Qual a probabilidade de que a bola tirada pertença a uma urna de percentagem p_i?

Resp.:

Visto que supomos que a bola saída é branca, a classe possível é constituída pela totalidade de elementos compostos de

urna — bola branca.

Designando por $\varpi_{(A)}$ a possibilidade desta classe, será (prop. II e IV)

$$\varpi_{(A)} = \frac{n_1}{N} p_1 + \frac{n_2}{N} p_2 + \ldots,$$

ou, pondo

$$\frac{n_i}{N} = \omega_i,$$

$$\infty_{(A)} = \Sigma\, \omega_i\, p_i.$$

A classe favorável é formada pela totalidade dos elementos compostos de

urna de ordem i — bola branca;

The problem of the probability of possible causes may be typified as follows:

Let us consider a set of N urns, n_1 of which have a fraction p_1 of white balls, n_2 of which have a fraction p_2 of white balls, etc.

Randomly choose one among the N urns, and from that urn randomly extract a ball; let's investigate the consequences of assuming that the extracted ball is white.

What is the probability that this ball has been extracted from an urn with percentage p_i of white balls?

The solution may be constructed as follows:

Under the hypothesis that the extracted ball is white, the elements of the possible set A are all the compound elements of the form

$$(\text{any urn}, \text{white ball}).$$

Denoting ϖ_A the possibility of this set, from Prop. II and IV we get that

$$\varpi_A = \frac{n_1}{N}\, p_1 + \frac{n_2}{N}\, p_2 + \cdots$$

or, denoting

$$\frac{n_k}{N} = \omega_k,$$

$$\varpi_A = \sum \omega_k\, p_k.$$

The elements of the favorable set A' are all the compound elements of the form

$$(\text{urn with } p_i \times 100\% \text{ white balls}, \text{white ball});$$

logo (prop. IV)

$$\varpi_{(A')} = \omega_i\, p_i.$$

Logo (def. 6.ª)

$$P_{(A)}^{(A')} = \frac{\omega_i p_i}{\Sigma\, \omega_i p^i} \tag{2}$$

que é a expressão conhecida pelo nome de fórmula de Bayes.

Nesta expressão, ω_i representa a probabilidade de tirar, dentre as N urnas, uma urna de ordem i e por isso chama-se-lhe *probabilidade à priori* da classe das urnas de ordem i ou duma urna de ordem i, mais simplesmente.

A probabilidade $P_{(A)}^{(A')}$ representa a probabilidade da mesma classe, *depois* de feita a primeira tiragem e de saír uma bola branca; por isso se lhe chama *probabilidade à posteriori.*

*
* *

Compreende-se que as *causas,* atraz consideradas, possam ser provenientes não duma tiragem feita numa só classe, mas dum sistema qualquer de tiragens.

Generalizemos a fórmula (1) para esse caso:

Sejam

$$\omega_1,\ \omega_2,\ \ldots,\ \omega_n$$

and therefore (Prop. IV)

$$\varpi_{A'} = \omega_i \, p_i .$$

Thus (Def. 6)

$$\mathbb{P}_A(A') = \frac{\omega_i \, p_i}{\sum \omega_k \, p_k}, \tag{1.2}$$

an expression known as Bayes formula.

In the above expression, ω_i is the probability of extracting, among the N urns, one with percentage p_i of white balls, and it is known as *a priori probability* of the urns with $p_i \times 100\%$ white balls.

The probability (1.2), $P_i = \mathbb{P}_A(A') = \frac{\omega_i \, p_i}{\sum \omega_k \, p_k}$, is the probability of extracting, among the N urns, one with percentage p_i of white balls, after performing the first extraction, resulting in white ball; for that reason, it is known as *a posteriori probability* of the urns with $p_i \times 100\%$ white balls.

*

*   *

It is obvious that the *causes* we are investigating may arise in any random extraction system, and that we cannot limit ourselves with extractions in a single set.

We now generalize formula (1.2) for sequential extractions from a complex of sets:

Denote

$$\omega_1, \omega_2, \ldots, \omega_n$$

as probabilidades *à priori* das *n causas* que podem dar lugar à saída de uma bola branca e sejam

$$p_1, p_2, \ldots, p_n$$

as probabilidades que cada uma dessas causas dá à saída duma bola branca.

Designemos por (A) a classe que se obtêm associando cada uma das causas a cada uma das *bolas* a que ela pode dar origem. Designemos por (A′) a classe que se obtêm associando cada uma das causas às *bolas brancas* a que ela pode dar origem. E designemos por (A″) a classe que se obtêm associando cada causa de ordem *i* com as *bolas brancas* a que cada uma dessas causas pode dar origem.

Teremos (prop. VI, *b*, 1.º)

$$P_{(A)}^{(A'')} = \omega_i p_i; \tag{3}$$

por outro lado (prop. VII),

$$P_{(A)}^{(A'')} = P_{(A)}^{(A')} \cdot P_{(A')}^{(A'')}, \tag{4}$$

sendo (prop. V)

$$P_{(A)}^{(A')} = \Sigma\, \omega_i p_i,$$

visto que

$$(A') = \Sigma (A'').$$

Logo

$$P_{(A')}^{(A'')} = \frac{\omega_i p_i}{\Sigma\, \omega^i p^i}.$$

the *a priori* probabilities of the n causes which may originate the extraction of white ball, and denote

$$p_1, p_2, \ldots, p_n$$

the probabilities that each of these causes confers to the event extraction of white ball.

Let us denote A the set that we obtain by associating each of the causes with each of the balls whose extraction it can originate. On the other hand, let us denote A' the set that we obtain by associating each of the causes with each of the *white balls* whose extraction it can originate. And let us denote $A''_k,\ k = 1, 2, \ldots, n$ the set that we obtain by associating each k-th cause with each of the *white balls* whose extraction each of these causes can originate.

From Prop. VI, b), 1st, we know that

$$\mathbb{P}_A(A''_i) = \omega_i\, p_i; \tag{1.3}$$

on the other hand (Prop. VII)

$$\mathbb{P}_A(A''_i) = \mathbb{P}_A(A') \times \mathbb{P}_{A'}(A''_i) \tag{1.4}$$

and (Prop. V)

$$\mathbb{P}_A(A') = \sum_{k=1}^{n} \omega_k\, p_k$$

since

$$A' = \bigcup_{k=1}^{n} A''_k.$$

Thus

$$\mathbb{P}_{A'}(A''_i) = \frac{\omega_i\, p_i}{\displaystyle\sum_{k=1}^{n} \omega_k\, p_k}\,.$$

*
* *

Esta dedução que acabamos de fazer, mostra bem como é erronea a dedução que desta fórmula se faz nos livros de probabilidades. Com efeito, no caso das urnas não terem o mesmo número de bolas, a dedução é feita por meio das igualdades (3) e (4), mas justificando-as a ambas com o princípio das probabilidades compostas. Ora a prop. VI é irredutível à proposição VII, porque a prop. VII não se aplica a elementos considerados como compostos.

Êste êrro passava encoberto debaixo da falta de precisão na significação do termo *fenómeno composto.*

De resto, sem êste sofisma, seria impossível deduzir a fórmula de BAYES com a definição de probabilidade dada nesses livros, pela simples razão de que a fórmula de BAYES se refere a um caso de probabilidades que não está previsto nessa definição: o caso de, não sendo os elementos igualmente possíveis, a *classe possível* ser menor do que a *classe total possível.*

Proposição IX

Fórmulas inversas da de Bayes

Representando a probabilidade à posteriori duma causa de ordem i por P_i, teremos, como vimos,

$$Pi = \frac{\omega_i p_i}{\Sigma\, \omega_i p_i}. \qquad (5)$$

*

* *

The above proof clearly shows that the usual argumentation that appears in other probability books is erroneous. In fact, when the urns do not have the same number of balls, the usual demonstration uses the formulas (1.3) and (1.4), justifying their use with the compound probability principle. But Proposition VI cannot be reduced to Proposition VII, since Prop. VII cannot be applied to compound elements.

This error was not evident due to lack of clarification of the meaning of *compound event* [and of *complex event*].

In fact, without this error it would have been impossible to establish Bayes formula with the definition of probability adopted is those books, since Bayes formula refers to a situation unforeseen in their definition: unequal probability of elementary events in a *possible set* which is a proper subset of the *total probability set*.

Proposition IX

Inverse formulas to Bayes' formula

Denoting the *a posteriori* probability of the i-th cause P_i, we have established that

$$P_i = \frac{\omega_i\, p_i}{\sum \omega_k\, p_k}. \tag{1.5}$$

Esta fórmula é simétrica em relação em relação a ω e p e nela supõem-se dados p e ω.

Consideremos, agora, o caso de serem dados P e p; neste caso teremos

$$\omega_i = \frac{\dfrac{P_i}{p_i}}{\Sigma \dfrac{P^i}{p_i}}.$$

Com efeito, de (5) tira-se

$$\frac{P_i}{p_i} = \frac{\omega_i}{\Sigma\, \omega_i p_i};$$

donde

$$\Sigma \frac{P_i}{p_i} = \frac{\Sigma\, \omega_i}{\Sigma\, \omega_i p_i} = \frac{1}{\Sigma\, \omega_i p_i}$$

e por isso

$$\omega_i = \frac{\dfrac{P_i}{p_i}}{\Sigma \dfrac{P_i}{p_i}}.$$

Pela simetria de (5), será ainda

$$p_i = \frac{\dfrac{P_i}{\omega_i}}{\Sigma \dfrac{P_i}{\omega_i}}.$$

This formula is symmetrical in what concerns the use of the ω_k and of p_k, which are given[(5)].

Let us now assume that the p_k and of P_k are given, and that our aim is to compute the *a priori* probabilities ω_i. We now prove that

$$\omega_i = \frac{\frac{P_i}{p_i}}{\sum \frac{P_k}{p_k}}.$$

In fact, from (1.5) we get that

$$\frac{P_i}{p_i} = \frac{\omega_i}{\sum \omega_k\, p_k};$$

therefore,

$$\sum \frac{P_i}{p_i} = \frac{\sum \omega_i}{\sum \omega_k\, p_k} = \frac{1}{\sum \omega_k\, p_k}$$

and thus $\omega_i = \dfrac{\frac{P_i}{p_i}}{\sum \frac{P_k}{p_k}}$. $\square$

Due to the symmetry of (1.5), we also have the inversion formula

$$p_i = \frac{\frac{P_i}{\omega_i}}{\sum \frac{P_k}{\omega_k}}.\ ^{(6)}$$

(5) *Editors' note*: this is not true: $\sum \omega_i = 1$, but $\sum p_i$ can be different from 1. For instance, in the classical Laplace's urn problem $\sum p_i = \sum_{k=0}^{N} \frac{k}{N} = \frac{N+1}{2}$.

(6) *Editors' note*: this is not true, unless $\sum p_i = 1$. The usefulness of the correct expression $\frac{p_i}{\sum p_k} = \frac{\frac{P_i}{\omega_i}}{\sum \frac{P_k}{\omega_k}}$ seems rather limited. On the other hand, $\omega_i = \frac{\frac{P_i}{p_i}}{\sum \frac{P_k}{p_k}}$ is true.

Proposição X

Consideremos o seguinte problema, em que se supõem as coisas dispostas como para o caso das probabilidades das causas.

«Tire-se, à sorte, uma urna e da urna tirada, uma bola que se verifica ser branca. Metida a bola na urna, pregunta-se: qual a probabilidade de que, feita outra tiragem na mesma urna, se obtenha uma bola branca» ?

1.ª solução

Resolvamos êste problema aplicando directamente a definição de probabilidade.

A classe possível, pode ser considerada como a totalidade dos elementos compostos de

urna qualquer — bola branca — bola qualquer;

será, pois,

$$\varpi_{(A)} = \Sigma\, \omega_i p_i \,.\, 1 = \Sigma\, \omega_i p_i.$$

A classe favorável será constituída pela totalidade dos elementos compostos de

urna qualquer — bola branca — bola branca;

logo

$$\varpi_{(A')} = \Sigma\, \omega_i p_i p_i = \Sigma\, \omega_i p_i^2.$$

Logo

$$P_{(A)}^{(A')} = \frac{\Sigma\, \omega_i p_i^2}{\Sigma\, \omega_i p_i}.$$

Proposition X

Let us now solve the problem that follows, where we assume the conditions stated for the problem of the probability of causes.

“From a randomly chosen urn, extract one ball; this ball is white, and after observation it is returned to the urn. What is the probability that a second extraction from this urn will result in white ball?”

1st solution

We shall solve this problem directly using the definition of probability.

The possible set is the set of all compound events of the form

(any urn, white ball, any ball)

and therefore

$$\omega_{A} = \sum \omega_k\, p_k \cdot 1 = \sum \omega_k\, p_k.$$

The favorable set is the set of all compound events of the form

(any urn, white ball, white ball)

and thus

$$\omega_{A'} = \sum \omega_k\, p_k\, p_k = \sum \omega_k\, p_k^2.$$

From the above, we get

$$\mathbb{P}_{A}(A') = \frac{\sum \omega_k\, p_k^2}{\sum \omega_k\, p_k}.$$

2.ª solução

Podemos ainda considerar as coisas do seguinte modo: a primeira tiragem equivale a substituir as urnas de probabilidade ω, por outras de probabilidade P, sendo ω as probabilidades à priori e P as probabilidades à posteriori. O problema em questão é, portanto, equivalente a êste outro:

Se n causas de probabilidades

$$P_1, P_2, \dots P_n,$$

dão a um determinado efeito, as probabilidades

$$p_1, p_2, \dots p_n,$$

respectivamente, qual a probabilidade dêsse efeito?

A probabilidade pedida será (prop. V)

$$P = \Sigma P_i p_i = \frac{\Sigma \omega_i p_i^2}{\Sigma \omega_i p_i},$$

como já tinhamos achado.

*

Dum modo mais geral:

Se numa urna, tirada à sorte, fizermos $m+n$ tiragens (metendo na urna cada bola tirada, antes de feita a tiragem imediata) e dessas tiragens nos resultarem m bolas

2nd solution

We may alternatively solve the problem in the following way: the effect of the observation of white ball in the first extraction is to change the *a priori* probabilities ω_k by the *a posteriori* probabilities P_k formerly computed. This problem is therefore equivalent to the following one:

If n causes with probabilities

$$P_1, P_2, \ldots, P_n$$

may result in a given effect with probabilities

$$p_1, p_2, \ldots, p_n,$$

respectively, what is the probability of that effect?

The desired probability is (Prop. V)

$$\mathbb{P} = \sum P_k\, p_k = \frac{\sum \omega_k\, p_k^2}{\sum \omega_k\, p_k}.$$

*

More generally:

If we perform $m+n$ extractions from a randomly chosen urn (returning each extracted ball to the urn after observation, before proceeding to the next extraction), resulting in m white balls and brancas e n pretas, a probabilidade de que a tiragem de ordem $m+n+1$ dê uma bola branca, é

$$P = \frac{\Sigma\, \omega_i p_i^{m+1} q_i^n}{\Sigma\, \omega_i p_i^m q_i^n},$$

resultado êste que se pode obter, ainda, pelos dois processos empregados para o caso particular de duas tiragens.

Corolário

Se as urnas tiverem a mesma probabilidade à priori, isto é, se

$$\omega_i = \text{const.},$$

será

$$P = \frac{\Sigma p_i^{m+1} q_i^n}{\Sigma p_i^m q_i^n}.$$

Problema

Dá-se uma urna contendo N bolas, brancas e pretas, de percentagens desconhecidas. Supondo que todas as percentagens são egualmente prováveis, qual a probabilidade de tirar uma bola branca na tiragem de ordem $m+n+1$, sabendo-se que nas primeiras $m+n$ tiragens se obtiveram m bolas brancas e n pretas ?

*

Êste problema é idêntico a êste outro.

n black balls, the probability of getting white ball in the $(m+n+1)$-th extraction is

$$\mathbb{P} = \frac{\sum \omega_k \, p_k^{m+1} \, q_k^{n}}{\sum \omega_k \, p_k^{m} \, q_k^{n}}. \qquad [q = 1 - p]$$

This result may be established by any of the two methods used in solving the former problem, which was the particular case of two extractions.

Corollary

If the urns have the same *a priori* probability, i.e., if

$$\omega_k = \text{constant}$$

then

$$\mathbb{P} = \frac{\sum p_k^{m+1} q_k^{n}}{\sum p_k^{m} q_k^{n}}.$$

Problem

One urn contains N balls, either white or black, in unknown proportions. Assuming that all the possible proportions $\left(0, \frac{1}{N}, \ldots, 1\right)$ of white balls are equiprobable, what is the probability of extracting white ball in the $(m+n+1)$-th extraction, if we know that the previous $m+n$ extractions resulted m times in white ball and n times in black ball?

*

The above problem is equivalent to the following one:

Dão-se $N+1$ urnas, contendo a primeira N bolas pretas, a segunda uma bola branca e $N-1$ pretas, a terceira 2 bolas brancas e $N-2$ pretas, etc., a última N bolas brancas.

Tira-se uma urna, à sorte, e fazem-se dela $m+n$ tiragens (tornando a pôr na urna cada bola, antes de tirar a seguinte) que dão m bolas brancas e n pretas. Pergunta-se: qual a probabilidade de que a tiragem de ordem $m+n+1$, dê uma bola branca?

*

O problema é resolvido, portanto, pela fórmula do corolário antecedente; pondo

$$p_i = \frac{N-\alpha}{N} \quad \text{e} \quad q_i = \frac{\alpha}{N},$$

virá

$$P = \frac{\sum\limits_{\alpha=0}^{N} \left(\frac{N-\alpha}{N}\right)^{m+1} \left(\frac{\alpha}{N}\right)^{n}}{\sum\limits_{\alpha=0}^{N} \left(\frac{N-\alpha}{N}\right)^{m} \left(\frac{\alpha}{N}\right)^{n}},$$

ou, aproximadamente,

$$P = \frac{\int_0^N \left(\frac{N-\alpha}{N}\right)^{m+1} \left(\frac{\alpha}{N}\right)^{n} d\alpha}{\int_0^N \left(\frac{N-\alpha}{N}\right)^{m} \left(\frac{\alpha}{N}\right)^{n} d\alpha}$$

ou, pondo

$$\alpha = N x,$$

There are $N+1$ urns, one of them with N black balls, another one with 1 white and $N-1$ black balls, another one with 2 white and $N-2$ black balls, etc., until the last urn, containing N white balls.

Performing $m+n$ extractions of one ball from a randomly chosen urn (always returning the extracted ball to the urn before proceeding to the next extraction), white ball is observed in m occasions, and black ball in n occasions. What is the probability of extracting white ball in the $(m+n+1)$-th extraction?

*

The solution is given in the corollary above, where we may use

$$p_k = \frac{k}{N} \qquad \text{and} \qquad q_k = \frac{N-k}{N}\,,$$

obtaining

$$\mathbb{P} = \frac{\sum\limits_{k=0}^{N} \left(\dfrac{k}{N}\right)^{m+1} \left(\dfrac{N-k}{N}\right)^{n}}{\sum\limits_{k=0}^{N} \left(\dfrac{k}{N}\right)^{m} \left(\dfrac{N-k}{N}\right)^{n}}\,,$$

which may be approximated by

$$\mathbb{P} \approx \frac{\displaystyle\int_0^N \left(\frac{\alpha}{N}\right)^{m+1} \left(\frac{N-\alpha}{N}\right)^{n} \mathrm{d}\alpha}{\displaystyle\int_0^N \left(\frac{\alpha}{N}\right)^{m} \left(\frac{N-\alpha}{N}\right)^{n} \mathrm{d}\alpha}\,.$$

Using the substitution

$$\alpha = N\,x,$$

$$P = \frac{\int_0^1 (1-x)^{m+1} x^n \, dx}{\int_0^1 (1-x)^m x^n \, dx}.$$

Ora

$$\int_0^1 x^n (1-x)^m \, dx = \frac{\Gamma(n+1) \, . \, \Gamma(m+1)}{\Gamma(m+n+2)}$$

e, sendo n inteiro,

$$\Gamma(n) = (n-1)\,!$$

Logo

$$P = \frac{\Gamma(m+2)\,\Gamma(n+1)}{\Gamma(m+n+3)} \cdot \frac{\Gamma(m+n+2)}{\Gamma(m+1)\,\Gamma(n+1)}$$

ou

$$P = \frac{(m+1)\,!}{(m+n+2)\,!} \cdot \frac{(m+n+1)\,!}{m\,!} = \frac{m+1}{m+n+2},$$

sendo esta fórmula tanto mais aproximada quanto maior fôr N.

$$\mathbb{P} \approx \frac{\int_0^1 x^{m+1} (1-x)^n \, \mathrm{d}x}{\int_0^1 x^m (1-x)^n \, \mathrm{d}x} \, .$$

As

$$\int_0^1 x^m (1-x)^n \, \mathrm{d}x = \frac{\Gamma(m+1)\,\Gamma(n+1)}{\Gamma(m+n+2)}$$

and, for natural n,

$$\Gamma(n) = (n-1)!$$

it follows that

$$\mathbb{P} \approx \frac{\Gamma(m+2)\,\Gamma(n+1)}{\Gamma(m+n+3)} \frac{\Gamma(m+n+2)}{\Gamma(m+1)\,\Gamma(n+1)}$$

or

$$\mathbb{P} \approx \frac{m+1}{m+n+2} \, ,$$

where the closeness of the approximation improves with the increase of N.

CAPÍTULO II

PROBABILIDADE CONTÍNUA

CHAPTER II

CONTINUOUS PROBABILITY

CAPITULO II

PROBABILIDADE CONTÍNUA

DEFINIÇÃO 1.ª

Se (A), (B), ... representam regiões limitadas quaisquer e a qualquer número de dimensões, com (A + B + ...) representaremos a região constituída pela totalidade dos pontos de (A), (B), ...

DEFINIÇÃO 2.ª

Se (A) e (B) representam regiões quaisquer, com (A, B) representaremos a totalidade dos pontos que se obtêm associando cada ponto de (A) com cada um dos pontos de (B).

*

Nas definições que acabamos de dar, a linguagem geométrica é simplesmente uma linguagem figurada em que a palavra *ponto*, num espaço a n dimensões, significa apenas um grupo qualquer de n números.

*

Demonstra-se na Pangeometria que se (A) e (B) repre-

CHAPTER II

CONTINUOUS PROBABILITY

If A, B, ... denote bounded regions in a space with any number of dimensions, $A \cup B \cup \cdots$ denotes the region with all the points from A, B, ...

If A, B, ... denote regions whatever, $A \times B$ will denote the set of ordered pairs (a, b), obtained from the sets A and B, by associating each point $a \in A$ with each point $b \in B$.

*

In the above definitions the use of geometric terminology is merely metaphoric, the word *point* meaning no more than any n-uple of numbers.

*

It has been proved in Pangeometry that if A and B are regions sentam regiões e os mesmos símbolos as suas *medidas* respectivas, se tem

$$(A, B) = (A) . (B),$$

ou melhor, a Pangeometria generaliza a noção de *medida* do hyper-espaço de modo tal que se mantêm a relação

$$(A, B) = (A) . (B).$$

(A, B) dir-se-há uma *região composta* de (A) e (B) e os seus pontos dir-se-hão *pontos compostos*, tal qual como na probabilidade discontínua.

Proposição primitiva

a)

Análogamente ao caso da probabilidade das classes discontínuas e finitas, consideraremos como primitiva a proposição: *Lançar, à sorte, um ponto numa região* (A) *limitada e a um número qualquer de dimensões.*

b)

A proposicão M *é um ponto lançado, à sorte, em* (A) tem uma significação análoga à da proposição precedente, mas presta-se melhor para o simbolismo da lógica.

DEFINIÇÃO 1.ª

Lançar, à sorte, um ponto em (A) *ou* (B) *ou* ... significa, por definição, *lançar, à sorte, um ponto na região* $(A + B + \ldots)$.

and $\mu(A)$, $\mu(B)$ the corresponding *measures*, then

$$\mu(A \times B) = \mu(A) \times \mu(B);$$

more precisely, Pangeometry has generalized the concept of *measure* in hyper-space in such a way that the relation

$$\mu(A \times B) = \mu(A) \times \mu(B);$$

is valid.

We shall say that $A \times B$ is a *compound region* from A and B, and its points (a, b) are referred to as *compound points* of a and b, similarly to the conventions we have adopted in the former chapter, dealing with discontinuous probability.

Primitive concept

a)

As in the case of probability of finite discontinuous sets, we consider as primitive the concept of *throwing* [or *selecting*, or *choosing*, or *extracting*] *a point, at random, in the bounded region* A *in any number of dimensions.*

b)

The statement X *is a point thrown, at random, in* A has the same meaning as *a*), *b*) being better suited to the formal symbolism of mathematical logic.

DEFINITION 1

Randomly extracting one point from A*, or* B*, or* $C, \ldots$, is the same as *randomly choosing one point in the region* $A \cup B \cup C \cup \cdots$

DEFINIÇÃO 2.ª

a)

Lançar, à sorte, um ponto em (A) *e outro em* (B), significa, por definição, *lançar, à sorte, um ponto em* (A, B).

b)

Lançar, à sorte, um ponto em (A) *e outro em* (B) *e outro em* (C), significa, por definição, *lançar, à sorte, um ponto em* (A, B) *e outro em* (C), etc.

Assim, lançar, à sorte, um ponto M no segmento AB e um ponto N no segmento AC, significa, por definição, lançar, à sorte, um ponto P no paralelogramo $ABDC$ *(fig. 1)*.

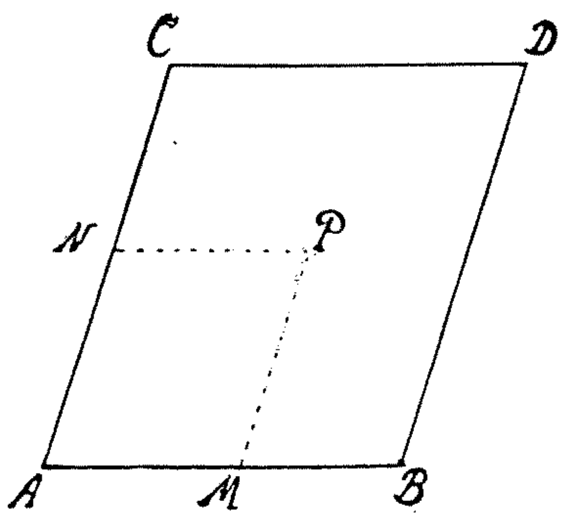

Figura 1

Lançar, à sorte, um ponto num arco de curva qualquer e outro num segmento de recta, sería o mesmo que lançar, à sorte, um ponto num segmento de superfície cilíndrica, etc.

DEFINIÇÃO 3.ª

a)

Consideremos, agora, o caso do lançamento à sorte ser feito, não em regiões quaisquer, mas em regiões sujeitas a certa dependência.

DEFINITION 2

a)

Randomly throwing one point from A and another[*, independently,*] *from B* is, by definition, the same as *randomly throwing one point from* $A \times B$.

b)

Randomly throwing one point from A, another from B and another from C is, by definition, the same as *randomly throwing one point from* $A \times B$ *and another from C,* [*independently,*] etc.

Thus, randomly choosing one point X in the segment $\overline{ab}$ and one point Y in the segment $\overline{ac}$ is the same as randomly choosing one point Z in the parallelogram $[abcd]$ (Fig. 1).

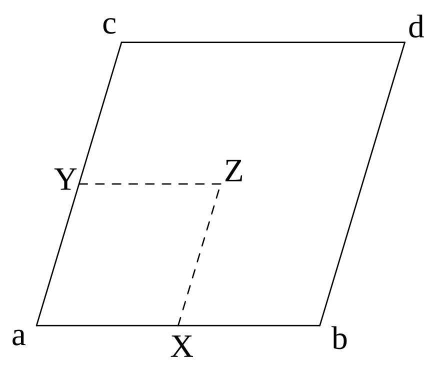

Figure 1

Randomly choosing one point in one arc and one point in a non coplanar line segment is the same as randomly choosing a point in the cylindrical surface generated by them, etc.

DEFINITION 3

a)

We now consider the case of constrained random selection, made in regions subject to some sort of mutual dependence.

Seja (A) uma região qualquer e suponhamos que a cada ponto de (A) se faz corresponder uma outra região (B), variável, em geral, com o ponto de (A) considerado.

Lançar, à sorte, um ponto em (A) *e outro na região* (B) *correspondente ao ponto de* (A) *coìncidente,* significa, por definição, *lançar, à sorte, um ponto na região* (A, B) sendo (B) a região correspondente ao ponto coìncidente de (A).

b)

Se a cada ponto de (B) fizermos corresponder uma nova região (C), *lançar, à sorte, um ponto em* (A), *outro na região* (B) *correspondente ao ponto coìncidente de* (A) *e*

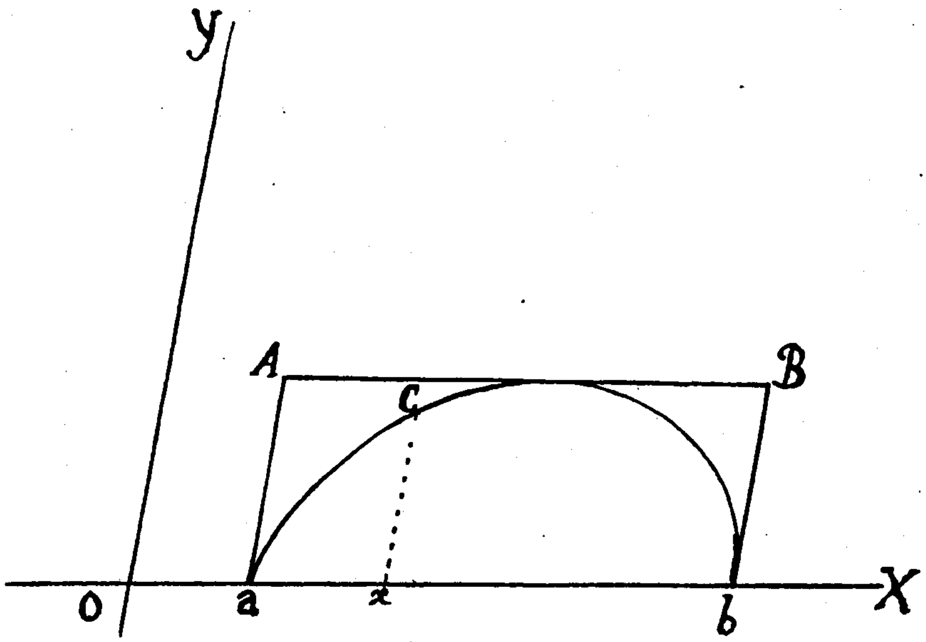

Figura 2

outro na região (C) *correspondente ao ponto coìncidente de* (B), significa por definição, *lançar, à sorte, um ponto em* (A, B) *e outro em* (C), etc.

*

Assim, suponhamos que (A) seja *(fig. 2)* o segmento *ab*

Let us associate with each $a \in A$ a region B_a, and denote $\{a\} \times B_a$ the set of ordered pairs $\{(a,b) : \ b \in B_a\}$.

Randomly selecting (or throwing) one point in A and another point in the corresponding region B_a is, by definition, the same as *randomly choosing one point (a,b) from $A \times B_a$.*

b)

If to each $b \in B_a$ we associate a region C_b, *randomly throwing one point in A, another point in the corresponding region B_a and another point in the*

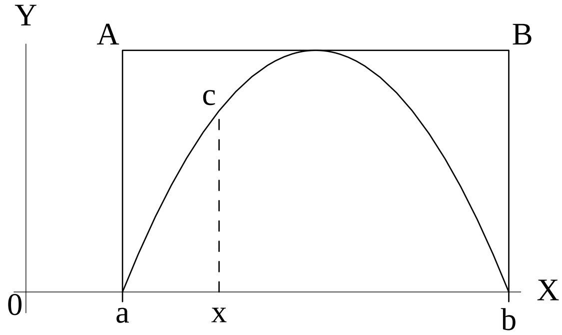

Figure 2

corresponding region C_b is, by definition, the same as *randomly throwing one point (a,b,c) in $A \times B_a \times C_b$*, etc.

*

As an example, let A be (Fig. 2) the line segment $\overline{ab}$ on the $\overrightarrow{OX}$ axis e que as regiões (B) são as ordenadas dos pontos da curva $a\,C\,b$, correspondendo a cada ponto x a ordenada $x\,C$.

Se o ponto lançado à sorte em ab coincidir com x, o segundo lançamento far-se-há em $x\,C$ e tudo se passará, segundo a definição, como se se fizesse um só lançamento num paralelogramo de base ab, com o lado superior passando por C e contido, portanto, no paralelograma $ab\,BA$.

Possibilidade

Segundo as definições que acabamos de dar, ou bem se trata de lançamentos feitos à sorte numa região ou a tais equivalentes (prop. prim. def. I e II), ou de lançamentos feitos num complexo de regiões (def. III *a* e *b*).

No primeiro caso, dizem-se *possíveis* todos os pontos da região em que os lançamentos são feitos. No caso da *fig. 1* seriam *possíveis* todos os pontos do paralelogramo $A\,B\,D\,C$.

No caso da def. III, *a*) dizem-se *possíveis* todos os pontos que resultam de associar cada ponto de (A) com cada um dos pontos da região (B) que lhe corresponde. Assim, no caso que a *fig. 2* representa, *possíveis* seriam todos os pontos da região $a\,b\,C$ limitada pela curva e pelo eixo dos XX, etc.

Possibilidade num ponto ou possibilidade por unidade

Seja (A) a grandeza duma região em que se faz um lançamento à sorte.

Ao número

$$\pi_A = \frac{1}{(A)}$$

and that for each x in $\overline{ab}$ the corresponding B_x is the vertical segment with endpoints in the $\overrightarrow{OX}$ axis and on the curve $\overset{\frown}{acb}$.

If in the first point randomly thrown in $\overline{ab}$ results x the second point will be randomly thrown in $\overline{xc}$ which is equivalent, according to the definition, of making only one random throw in the parallelogram with $\overline{ab}$ as base and the upper side passing from the point c, which is a subset of the parallelogram $[abAB]$.(7)

Possibility

According to the above definitions, we consider either randomly throwing one point in one region, single or compound (primitive concept. and Def. 1 and 2) or constrained randomly throwing one point in a complex region (Def. 3 a and b)

In the first situation, we shall consider *possible* all the points in the region where the random throws are done. For instance, in the example illustrated in Fig. 1, all the points in the parallelogram $[abcd]$ are possible points.

In the case of Def. 3 a) we shall say that the possible points are those that result from associating each point from region A with each point from the corresponding region B_a, i.e., the complex of regions $A \odot\!\!\rightarrow B = \bigcup_{a\in A} \{a\}\times B_a$. For instance, in the example illustrated in Fig. 2, the possible points are those lying in he region limited by the segment line $\overline{ab}$ and the curve $\overset{\frown}{acb}$, etc.

Point possibility or unit possibility

Let $\mu(A)$ be the measure of the region A where we are throwing points at random.

The number

$$\pi_a = \frac{1}{\mu(A)}$$

---

(7) *Editors' note*: In fact this is not true, and the observation where Pacheco d'Amorim says that all sampling schemes can be reduced to a single selection (or throw) are contradictory to the very detailed construction he builds to overpass the question of dependence; this is never explicitly stated, but it is evident that Pacheco d'Amorim tries to solve elegantly how to deal with joint probabilities. Observe also that his "reconstruction" of Fubinni's theorem is chapter IV clearly shows that this bold statement that hierarchical sampling can be reduced to single sampling cannot hold in dependent settings.

chama-se *possibilidade no ponto A* ou *possibilidade por unidade.*

Podemos dizer que neste caso todos os pontos de (A) são *igualmente possíveis,* para nos servirmos duma linguagem análoga à da probabilidade discontínua, querendo dizer com *pontos igualmente possíveis,* pontos *nos* quais a possibilidade é sempre a mesma.

*

Segundo esta definição, os pontos das regiões consideradas na *prop. prim.* e *def. I* e *II* são igualmente possíveis. Nos casos da *def. III* a possibilidade não será, em geral, a mesma para todos os pontos.

Em qualquer dos casos, porêm, ela é perfeitamente determinada, visto que, segundo as definições dadas, cada ponto possível faz parte duma região e duma só, em que se efectua um lançamento à sorte, do qual lhe resulta a propriedade de ser *possível.*

A possibilidade num ponto fica sendo uma função definida das coordenadas do mesmo.

Proposição I

Por razões perfeitamente idênticas às do caso análogo das probabilidades das classes finitas, *a probabilidade num ponto composto é igual ao produto das possibilidades nos pontos componentes.*

Possibilidade duma região

Uma região diz-se *possível* quando é formada por pontos possíveis.

will be called *possibility at point a* or *unit possibility*.

Similarly with what happens in the case of discontinuous probability, we may say that, in case all the random throwing of points is performed in the same [single or compound] region, all the sample points are *equally possible*, in the sense that the possibility is the same in each of those points.

*

According to this definition, all the points from regions as described in the primitive concept and in Def. 1 and 2 are equally possible. But in the cases addressed in Def. 3 the possibility will not be, in general, the same for all points.

However, the possibility is always well defined, since in all those definitions one point belongs to some uniquely defined region, where random throws are performed, its *possibility* resulting from the random throwing system adopted.

The possibility of each point is then a function of its coordinates.

Proposition I

Similarly to what happens in the case of probability in finite sets, *the possibility of a compound point* $(a, b) \in A \times B$ *is the product of the possibilities of its components.* The proof is in all points similar to the proof for the finite sets case.

Possibility of a region

One region is said to be possible if all its points are possible.

Chama-se possibilidade duma região (A), ao integral da possibilidade por unidade, estendido a toda essa região, caso êsse integral exista.

Representaremos pos $\varpi_{(A)}$ a possibilidade de (A).

Proposição II

Em virtude da definição antecedente, será ainda,

$$\varpi_{(A)} = \varpi_{(A_1)} + \varpi_{(A_2)} + \ldots + \varpi_{(A_n)}$$

todas as vezes que

$$(A) = (A_1) + (A_2) + \ldots + (A_n).$$

*

Chamaremos, ainda, *região total possível,* à totalidade dos pontos possíveis, em relação a determinado sistema de lançamentos.

Proposição III

A possibilidade da região total possível é igual à unidade como se demonstra de modo perfeitamente análogo ao da probabilidade discontínua.

Proposição IV

Teremos ainda

$$\varpi_{(A,\,B)} = \varpi_{(A)} \cdot \varpi_{(B)}.$$

The possibility of a given possible region A' is the integral of the unit possibility over that region, in case this integral exists.

We shall denote $\varpi_{A'}$ the possibility of a possible region A'.

Proposition II

If A' is a possible region which may be partitioned into pairwise disjoint regions

$$A' = A'_1 \cup A'_2 \cup \cdots \cup A'_n$$

then

$$\varpi_{A'} = \varpi_{A'_1} + \varpi_{A'_2} + \cdots + \varpi_{A'_n}\,.$$

*

The possible regions A' are subsets of the *total possible region* A, the region of all possible points in the random throwing system considered.

Proposition III

The possibility of the total possible region is 1. (The proof has exactly the same steps detailed in the proof of the similar property in the case of finite sets.)

Proposition IV

If the region $A \times B$ is compound from the regions A and B, then

$$\varpi_{A\times B} = \varpi_A \times \varpi_B\,,$$

Com efeito, visto que a função $\frac{1}{(B)}$ é independente das coordenadas dos pontos de (A), e visto que

$$\frac{1}{(A, B)} = \frac{1}{(A)} \cdot \frac{1}{(B)},$$

será

$$\varpi_{(A, B)} = \int_{(A, B)} \frac{d(A, B)}{(A, B)} = \int_{(A)} \frac{d(A)}{(A)} \cdot \int_{(B)} \frac{d(B)}{(B)} = \varpi_{(A)} \cdot \varpi_{(B)}$$

c. d. d.

Probabilidade

Seja (A) uma região *possível* relativamente a um dado sistema de lançamentos à sorte, e (A′) outra nela contida.

Chama-se *probabilidade da região* (A′) *em relação à região* (A), *ao número*

$$P_{(A)}^{(A')} = \frac{\varpi_{(A')}}{\varpi_{(A)}},$$

sendo $\varpi_{(A')}$ e $\varpi_{(A)}$ as possibilidades de (A′) e de (A).

*

Se os pontos de (A) forem *igualmente possíveis*, será

$$P_{(A)}^{(A')} = \frac{(A')}{(A)}.$$

Quando a *classe possível* coincidir com a *classe total possível*, será

$$\varpi_{(A)} = 1$$

e

$$P_{(A)}^{(A')} = \varpi_{(A')}.$$

In fact, as the function $\dfrac{1}{\mu(B)}$ is independent of the coordinates of points from the region A, recalling that

$$\mu(A \times B) = \mu(A) \times \mu(B)$$

we have

$$\varpi_{A\times B} = \int\limits_{A\times B} \frac{\mathrm{d}(a,b)}{\mu(A \times B)} = \int\limits_{A} \frac{\mathrm{d}(a)}{\mu(A)} \times \int\limits_{B} \frac{\mathrm{d}(b)}{\mu(B)} = \omega_{A} \times \omega_{B}$$

$\square$

Probability

Let A' be a possible region in what concerns a given random throwing system, and $A'' \subset A'$ another possible region.

We shall call *probability of the region* A'' *relative to the region* A' the number

$$\mathbb{P}_{A'}(A'') = \frac{\varpi_{A''}}{\varpi_{A'}}\,,$$

$\varpi_{A''}$ and $\varpi_{A'}$ denoting, as above, the possibilities of the regions A'' and of A', respectively.

*

If the elements in A' are equally possible, it follows that

$$\mathbb{P}_{A'}(A'') = \frac{\mu(A'')}{\mu(A')}\,.$$

When the possible set A' is the total possible set A, from

$$\varpi_{A} = 1$$

it follows that

$$\mathbb{P}_{A}(A'') = \varpi_{A''}$$

Exemplos:

1.º

Parte-se, à sorte, um segmento em três partes: qual a probabilidade de que essas três partes formem um triângulo?

Resp.:

Partir um segmento em três partes é considerar sôbre êle dois pontos; parti-lo à sorte, será, por definição, lançar, à sorte, dois pontos sôbre êle, ou, o que é o mesmo (def. II), lançar, à sorte, um ponto num quadrado que tem êsse segmento de lado.

Seja $A\,B$ o segmento dado *(fig. 3)* e seja $A\,B\,C\,D$ o quadrado *(fig. 4)*.

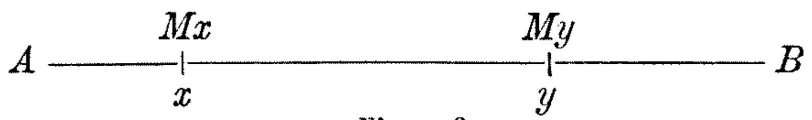

Figura 3

Sejam (x, y) as coordenadas do ponto M correspondente às posições Mx e My dos pontos lançados à sorte sôbre $A\,B$.

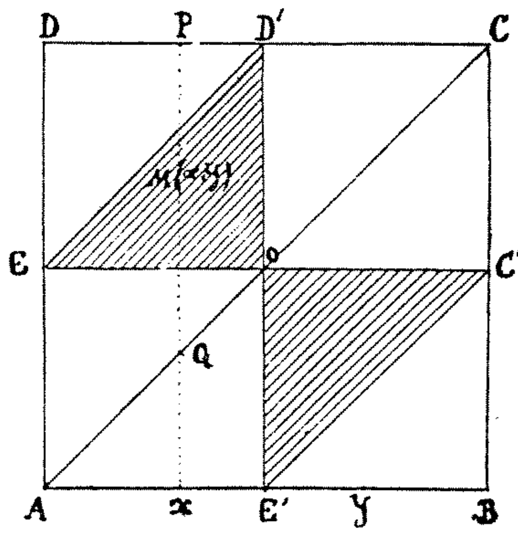

Figura 4

Supondo que os segmentos mencionados no enunciado do problema em questão são aditivos, teremos a determinar a probabilidade de que os segmentos Ax, xy (caso $x < y$) e yB formem um triân-

Examples:

1st

A line segment is randomly broken into three parts. What is the probability that the three resulting segments can be taken for sides of a triangle?

To break a segment into three parts, randomly, is the same as to throw randomly two points X and Y on it. By Def. 2, this is the same as to throw, randomly, one point in the square having the segment as one of its sides.

Let $\overline{ab}$ be the segment (Fig. 3) and $[abcd]$ be the associated square (Fig. 4).

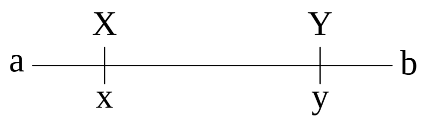

Figure 3

Let (X, Y) be the coordinates of the point Z corresponding to the positions X and Y of the two points randomly marked in the segment $\overline{ab}$.
Assuming that the segments mentioned in the problem are additive, we shall need to determine the probability that the segments $\overline{aX}$, $\overline{XY}$ and $\overline{Yb}$ can be taken as the sides of a triangle (case $X < Y$);

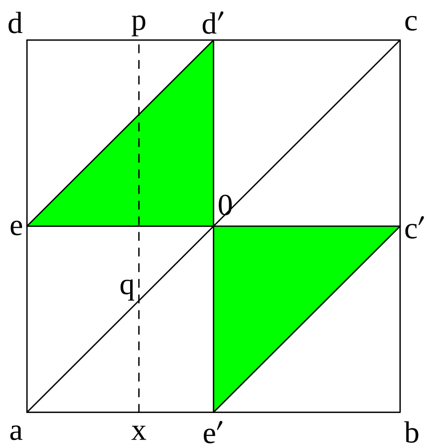

Figure 4

gulo; caso $y<x$, os segmentos a considerar serão Ay, yx, xB.

Seja, em primeiro lugar, $x<y$.

Para que os três segmentos possam formar triângulo será necessário e suficiente que

$$\begin{array}{l} x \leqslant (y-x)+(\alpha-y) \\ y-x \leqslant x+(\alpha-y) \\ \alpha-y \leqslant x+(y-x) \end{array}$$

representando por α o comprimento do segmento AB.

Estas condições são equivalentes a estas outras, como se vê imediatamente,

$$\begin{array}{l} 0 \leqslant y \geqslant \frac{\alpha}{2} \\ 0 \leqslant x \leqslant \frac{\alpha}{2} \\ 0 \leqslant y-x \leqslant \frac{\alpha}{2} \end{array} \qquad (1)$$

A região favorável (A′) será, pois, constituida pela totalidade dos pontos do quadrado $ABCD$ cujas coordenadas satisfaçam às condições (1).

A estas condições só satisfazem os pontos da região $OD'E$, como é óbvio.

No caso de $x \geqslant y$, a região favorável seria formada pelo triângulo simétrico em relação a AC.

E como todos os pontos sam igualmente possíveis, a probabilidade pedida será dada pelo quociente das áreas das regiões favorável e possível, isto é,

$$P = \frac{1}{4}.$$

or else, that the segments $\overline{aY}$, $\overline{YX}$ and $\overline{Xb}$ can be taken as the sides of a triangle (case $X > Y$).

Let us analyze first the case $X < Y$.

Denoting α the length of the segment $\overline{ab}$ [and assuming a=0, for simplicity], the three segments can be the sides of a triangle if and only if

$$0 < X < (Y - X) + (\alpha - Y),$$

$$0 < Y - X < X + (\alpha - Y),$$
$$0 < \alpha - Y < X + (Y - X).$$

These conditions are equivalent to

$$\begin{aligned} &0 < X < \frac{\alpha}{2}\,, \\ &0 < Y - X < \frac{\alpha}{2}\,, \\ &\frac{\alpha}{2} < Y < \alpha. \end{aligned} \tag{2.1}$$

The totality of points from the square [$abcd$] whose coordinates verify condition (2.1) is the favorable region.

From the analysis of Fig. 4, it is obvious that this region is [$od'e$].

In the case $X > Y$, a similar analysis shows that the favorable region is the triangle [$oc'e'$], symmetrical to [$od'e$] in reference to the line $\overline{ac}$.

As in this randomly throwing system all points are equally possible, the probability is given by the quotient of the area of the favorable region by the area of the possible region, i.e.

$$\mathbb{P} = \frac{1}{4}\,.$$

2.º

Suponhamos agora que partimos o mesmo segmento em dois e depois o segmento *maior* ainda em dois: qual a probabilidade de que os três segmentos possam formar triângulo?

Resp.:

O campo favorável é ainda o mesmo do problema antecedente; vejamos o campo possível.

Para $x \leqslant \frac{a}{2}$, será $Ax \leqslant xB$ e por isso $x < y < a$, isto é, quando o primeiro ponto cáe em x o segundo ponto cairá em qualquer ponto de xB e por isso todos os pontos de QP são possíveis e portanto são possíveis todos os pontos da região A O D′ D *(fig. 4)*.

Por razões idênticas é possível a região E′ O C B. Como a questão é simétrica em relação a A C, consideremos só a parte A O D′ D como região possível e O D′ E como região favorável.

No problema antecedente os elementos possíveis eram também igualmente possíveis, porque cada um dos pontos supunha-se lançado sôbre todo o segmento, o que neste caso se não dá. Chamando F à região favorável e P à região possível, será

$$\varpi_P = 1$$

e

$$\varpi_F = \int_{OD'E} \frac{2\,dx}{a} \cdot \frac{dy}{a-x} = \frac{2}{a} \int_0^{\frac{a}{2}} \frac{dx}{a-x} \cdot \int_{\frac{a}{2}}^{x+\frac{a}{2}} dy$$

2nd

Let us now assume that the segment is randomly broken into two segments, and then that the *bigger* subsegment is randomly broken into two. What is the probability that the three resulting segments can be the sides of a triangle?

The favorable region is obviously the same that we have constructed in the previous problem; let us now find the possible region.

When $X < \frac{\alpha}{2}$, we shall have $\overline{aX} < \overline{Xb}$ and therefore $X < Y < \alpha$, i.e., conditionally on the first point being $X < \frac{\alpha}{2}$ the second point is in $\overline{Xb}$, and thus (cf. Fig. 4) all the points in $B_x = \overline{qp}$ are possible. Hence, all the point in region $[aod'd]$ are possible.

For identical reasons, in the case $\frac{\alpha}{2} < X < \alpha$ the favorable region is $[oc'e']$ and the possible region is $[e'ocb]$; as all is symmetrical in reference to $\overline{ac}$, we shall make our computation for the case $X < \frac{\alpha}{2}$, the other one having the same numerical solution.

In the previous problem, all the possible points were equally possible, since each of the two points was randomly thrown into the segment $\overline{ab}$, without any restriction. In the present problem, this is not so[8]. Denoting F the favorable region and P the possible region, we have

$$\varpi_P = 1$$

and

$$\varpi_F = \iint_{[od'e]} \frac{2\,\mathrm{d}x}{\alpha} \cdot \frac{\mathrm{d}y}{\alpha - x} = \frac{2}{\alpha} \int_0^{\frac{\alpha}{2}} \left(\int_{\frac{\alpha}{2}}^{x+\frac{\alpha}{2}} \mathrm{d}y \right) \frac{\mathrm{d}x}{\alpha - x},$$

(8) *Editors' note*: Observe that to any $X \in \left(0, \frac{\alpha}{2}\right)$ we associate $B_X = (X, \alpha)$, and hence the possibility of any $(X, Y) \in [aod'd]$ is $\frac{2}{\alpha} \times \frac{1}{\alpha - X}$.

donde

$$\varpi_F = \frac{2}{a}\int_0^{\frac{a}{2}} \frac{x}{a-x}\,dx =$$

$$= \frac{2}{a}\int_0^{\frac{a}{2}}\left[-1+\frac{a}{a-x}\right]dx = -1+2\left[-\log.\,(a-x)\right]_0^{\frac{a}{2}} =$$

$$= 2\log.\,2 - 1 = 0{,}44\ \ldots\ldots$$

e, portanto,

$$P = 0{,}44\ \ldots\ldots$$

isto é, a probabilidade augmentou, como era de esperar.

3.º

Consideremos ainda outra variante do mesmo problema.

Lança-se à sorte um ponto M em A E′ *(fig. 4)* e outro M′ no segmento M B e supõe-se que o ponto M′ caíu entre E′ e B; pede-se a probabilidade de que os três segmentos A M, M M′ e M′ B possam formar triângulo.

Resp.:

As regiões *total* possível, possível e favorável são respectivamente AOD′D, EOD′D e EOD′.

Teremos pois:

$$\varpi_P = \iint_{EOD'D} \frac{dx}{\frac{a}{2}} \cdot \frac{dy}{a-x}$$

$$= \int_0^{\frac{a}{2}} \frac{dx}{a-x}$$

and so

$$\varpi_F = \frac{2}{\alpha}\int_0^{\frac{\alpha}{2}} \frac{x}{\alpha - x}\,\mathrm{d}x =$$

$$= \frac{2}{\alpha}\int_0^{\frac{\alpha}{2}} \left[-1 + \frac{\alpha}{\alpha - x}\right]\mathrm{d}x = -1 + 2\,\big(-\log(\alpha - x)\big]_0^{\frac{\alpha}{2}} =$$

$$= 2\,\log\,2 - 1 \approx 0.386$$

and therefore the probability we wanted to compute is

$$\mathbb{P} \approx 0.386$$

i.e., we get in this problem a bigger probability, as it should be expected by the extra conditions, which have increase the possibility of the three segments forming a triangle[9].

3rd

Now we consider a follow up of the above problem.

One point X is randomly thrown in $\overline{ae'}$ (Fig. 4), and another point X' is randomly thrown in the segment $\overline{Xb}$, and we further assume the condition that $X' \in \overline{e'b}$; what is the probability that the three segments $\overline{aX}$, $\overline{XX'}$ and $\overline{X'b}$ can be the sides of a triangle?

The *total possible region*, the possible region and the favorable region are $[aod'd]$, $[eod'd]$ and $[eod']$, respectively.

Therefore

$$\varpi_P = \iint_{[eod'd]} \frac{2\,\mathrm{d}x}{\alpha}\cdot\frac{\mathrm{d}y}{\alpha - x} =$$

$$= \int_0^{\frac{\alpha}{2}} \frac{\mathrm{d}x}{\alpha - x}$$

(9) *Editors' note*: We have corrected the final result given by the author who presents 0.44.

ou,

$$\varpi_P = \Big[- \log.\,(a - x) \Big]_0^{\frac{a}{2}}$$

$$= \log.\, a - \log.\, \frac{a}{2}.$$

$$= \log.\, 2$$

e (prob. 2.º)

$$\varpi_F = 2 \log.\, 2 - 1\ ;$$

logo

$$P = 2 - \frac{1}{\log.\, 2} = 0{,}6\ldots$$

Observação

Tudo o que se disse da probabilidade discontínua, diz-se da probabilidade contínua. Assim, as proposições V, VI, VII, demonstradas para a probabilidade discontinua, demonstram-se de modo idêntico para a probabilidade contínua e serão indicadas neste capítulo com os mesmos números. O «problema das probabilidades das causas» podia ser tratado aqui do mesmo modo que no capítulo antecedente.

No capítulo IV será esse problema tratado dum modo bastante diferente e mais geral.

Proposição VIII

Se um ponto M, variando numa dada região, se puder decompor em dois componentes, M_1 e M_2 tais que, a cada or

$$\varpi_{P} = (-\log(\alpha - x)\big]_0^{\frac{\alpha}{2}} =$$

$$= \log\,\alpha - \log\,\frac{\alpha}{2} =$$

$$= \log\,2,$$

and (probl. 2nd)

$$\varpi_{F} = 2\,\log\,2 - 1;$$

therefore

$$\mathbb{P} = 2 - \frac{1}{\log\,2} \approx 0,557.^{(10)}$$

Observation

All that has been said about discontinuous probability, is also valid for continuous probability. Therefore Propositions V, VI and VII from Chapter I may be established for continuous probability using the same arguments that have been used in the case of discontinuous probability, and we take them as Propositions V, VI and VII in this Chapter II, without explicitly rewriting them. The "problem of the probabilities of causes" could be dealt with here as we did in Chapter I.

But we postpone the investigation of that problem to Chapter IV, using a different and more general methodology.

Proposition VIII

If a variable point X in some given region can be decomposed in two components X_1 and X_2 in such a way that

---

(10) *Editors' note*: We have detailed the final result given by the author who presents 0.6.

posição de M_1 corresponda sempre para M_2 o mesmo campo favorável e o mesmo campo possível, M_1 será *independente* de M_2 e por isso as suas probabilidades podem calcular-se separadamente, o que, em geral, facilita muito a resolução do problema. No caso particular da probabilidade de M_1 ser igual à unidade, a probabilidade de M será independente dos parametros que definem a posição de M_1 que por isso poderemos supôr *fixo.*

Quando a região em que se faz o lançamento de M tiver um elemento de simetria, será, em geral, aplicável esta proposição.

Exemplo :

Lançam-se, à sorte, dois pontos sôbre a superfície duma esfera. Pergunta-se : qual a probabilidade de que o menor arco do círculo máximo que liga os dois pontos, seja inferior a α?

Resp. :

Qualquer que seja a posição dum dos pontos M, por exemplo, para o outro corresponderá sempre a mesma região favorável e a mesma possível.

Supondo, então, M fixo, a região favorável será dada pela calote tendo o ângulo 2α de abertura e M por vértice, a região possível, sendo toda a superfície da esfera. O problema é, pois, equivalente a êste outro : qual a probabilidade de que um ponto lançado à sorte na superfície duma esfera, cáia sôbre um segmento dessa superfície ? Problema êste de solução imediata.

whatever the position of X_1 the corresponding X_2 has always the same favorable and possible regions, X_1 and X_2 are *independent*, and henceforth their probabilities can be computed separately; this, as a rule, simplifies considerably the solution of problems. If, in particular, the probability of X_1 is 1, the probability of X is independent of the parameters defining the position of X_1, which as a consequence we may assume *fixed*.

When the region where points are randomly thrown has a symmetry element, this proposition can in general be used.

Example:

Two points are randomly thrown on a spherical surface. What is the probability that the smaller arc of the maximum circle defined by the two points is smaller than α?

Whatever the position of one of the points, say X, the favorable and possible regions for the other point Y are always the same.

In fact, given X, the favorable region is the spherical cap having vertex X and an angle 2α, and the possible region is the entire spherical surface. Therefore, the problem can be reformulated as follows: What is the probability that one point Y randomly thrown on a spherical surface lies in a given spherical cap? This problem has immediate solution.

Proposição IX

«Se o campo da variação de M se puder decompor em partes, de modo que a probabilidade de M seja a mesma em cada uma delas, teremos que a probabilidade no campo total será igual à probabilidade em qualquer das partes».

Com efeito, sejam (A), (B), ... (L), as partes do campo possível e (A′), (B′) ... (L′) as partes correspondentes do campo favorável; teremos que, por hipótese,

$$P = \frac{\varpi_{(A')}}{\varpi_{(A)}} = \frac{\varpi_{(B')}}{\varpi_{(B)}} = \ldots = \frac{\varpi_{(L')}}{\varpi_{(L)}}.$$

Logo

$$P = \frac{\Sigma\,\varpi_{(L')}}{\Sigma\,\varpi_{(L)}}$$

que é a probabilidade de M,

c. d. d.

DEFINIÇÃO 4.ª

A proposição *lançar, à sorte, um ponto em* (A) sendo (A) ilimitada, significa, *lançar, à sorte, um ponto em* (A′) sendo (A′) uma região contida em (A), limitada e arbitráriamente grande.

Se M é um ponto lançado à sorte em (A) e (A) é ilimitada, chama-se probabilidade de M na região (A) ao limite da probabilidade de M em (A′) quando (A) aumenta indefenidamente, isto é, ao número P tal que a todo o $\delta > 0$

Proposition IX

If the region of variation of X can be partitioned into subregions in such a way that the probability of X in each of them is always the same, the probability in the total region is the same as the probability in any of those regions.

In fact, let $A_1, A_2, \ldots, A_L$ be the pairwise disjoint subregions of the possible region A, and $A'_1, A'_2, \ldots, A'_L$ the corresponding favorable subregions of the favorable region A'; from the hypothesis

$$\mathbb{P} = \frac{\varpi_{A'_1}}{\varpi_{A_1}} = \frac{\varpi_{A'_2}}{\varpi_{A_2}} = \cdots = \frac{\varpi_{A'_L}}{\varpi_{A_L}}$$

we get that

$$\mathbb{P} = \frac{\sum_{k=1}^{L} \varpi_{A'_k}}{\sum_{k=1}^{L} \varpi_{A_k}},$$

which is the probability of X. □

DEFINITION 4

The proposition *randomly throw one point in the region* A, A being unbounded, has the same meaning as *randomly throw one point in the region* A', where $A' \subset A$ is an arbitrarily large bounded region.

If X is a randomly thrown point in A and A is unbounded, the probability of X in region A is the limit of the probability of X in A' when A' increases indefinitely, i.e. the probability of X is the number $\mathbb{P}$ such that for all $\delta > 0$

corresponda uma região (C) tal que

$$| P - P_{(B)} | < \delta$$

para todas as regiões (B) tais que

$$(B) > (C).$$

Trataremos mais desenvolvidamente dêste caso num *apêndice* em que estudaremos a probabilidade dos conjuntos numeráveis, depois de numerados.

there corresponds a region C such that

$$|\mathbb{P} - \mathbb{P}_{B}| < \delta$$

for all regions B such that

$$C \subset B.$$

This case will be dealt with in detail in an appendix, where we investigate the probability of denumerable sets[11].

[11] *Editors' note*: In the last sentence of the Preface, Pacheco d'Amorim says that he had conceived the intention of including an appendix on this subject, but that finally he has decided otherwise.

CAPÍTULO III

CHAPTER III

RANDOM FIGURES

CAPÍTULO III

PRIMEIRA PARTE

Todas as proposições em que entrem as frases *tirar à sorte, lançar à sorte* associadas a figuras rígidas ou variáveis, terão de ser definidas por meio de lançamentos ou tiragens à sorte, como está indicado nas definições 1.ª, 2.ª e 3.ª dos capítulos I e II.

Lançamentos de figuras rígidas

DEFINIÇÃO 1.ª

Tirar, à sorte, um sentido num espaço a n *dimensões,* significa, por definição, *lançar à sorte um ponto no espaço da figura cuja equação é*

$$(x_1-x'_1)^2+(x_2-x'_2)^2+\ldots+(x_n-x'_n)^2=1.$$

O sentido será o do vector definido pelo ponto de coordenadas $(x'_1, x'_2, \ldots, x'_n)$ como origem e pelo ponto lançado à sorte como extremidade.

O ponto $(x'_1, x'_2, \ldots, x_n)$ pode ser qualquer.

CHAPTER III

RANDOM FIGURES

FIRST PART

All the propositions where the terms *randomly extracting* or *randomly throwing*, or equivalents are used in the context of the construction of random figures (either rigid or variable) will be defined through the use of the concepts of random extractions from a finite set or random throws of points in a continuous region, as indicated in definitions 1, 2 and 3 of Chapters I and II.

Random rigid figures

DEFINITION 1

The random choice of an orientation in a space of dimension n is, by definition, the same as randomly throwing a point $(x_1, \ldots, x_n)$ in the set defined by the equation

$$\left(x_1 - x'_1\right)^2 + \left(x_2 - x'_2\right)^2 + \cdots + \left(x_n - x'_n\right)^2 = 1.$$

The orientation will be that of the vector having $({x'}_1, \ldots, {x'}_n)$ as origin and the randomly thrown point $(x_1, \ldots, x_n)$ as extremity.

Any point in the space may be taken as origin $({x'}_1, \ldots, {x'}_n)$.

No caso particular de $n = 2$, o lançamento faz-se sôbre uma circunferência e no de $n = 3$, sôbre uma esfera.

Análoga definição serve para a tiragem, à sorte, duma direção.

DEFINIÇÃO 2.ª

Sejam A B e A′ B′ dois segmentos de recta. Sobreponhamo-los de modo que o ponto A coìncida com o ponto A′ e façamos escorregar o segmento menor sôbre o maior até que o ponto B coìncida com o ponto B′. Dêste modo o segmento menor passa por todas as posições que êle pode ocupar em relação ao segmento maior e no final do percurso todos os pontos do segmento menor terão descrito segmentos de recta iguais.

Lançar, à sorte, o segmento menor sôbre o maior significa, *lançar, à sorte, um ponto qualquer do segmento menor sôbre o segmento que esse mesmo ponto descreve quando o segmento menor percorre o segmento maior.*

Lançar, à sorte, o segmento maior sôbre o menor, significa o mesmo que *lançar o menor sôbre o maior.*

*

Para que esta definição seja correcta é necessário e basta que ela não dependa de nenhum ponto particular, isto é, que seja a mesma qualquer que seja o ponto considerado sôbre o segmento. Ora de facto assim é, visto que todos os pontos do segmento descrevem segmentos iguais.

Como, em virtude do sistema das definições 1.ª, 2.ª e 3.ª, tudo se resume ao lançamento dum ponto numa

In the particular case $n = 2$, the random throw is done in a circumference, and in the case $n = 3$ the random throw is done in a spherical surface.

The definition of a random direction is done by analogy.

DEFINITION 2

Let $\overline{ab}$ and $\overline{a'b'}$ be two line segments. Superimpose the two segments in such a way that a and a' coincide, and then let the smaller one slide over the bigger one until points b and b' coincide. In other words, the smaller segment goes through all positions it may have upon the bigger one, and at the end of this procedure each point of the smaller segment will have defined segment trajectories of the same length.

Randomly throwing a smaller segment on a bigger one is the same as randomly throwing any given point of the smaller segment on the segment it defines when the smaller segment slides over the bigger one, as described above.

Randomly throwing the bigger segment on the smaller one is the same as randomly throwing the smaller segment on the bigger one.

*

The validity of the definition lies in the fact that all the segments defined by each point of the smaller segment when it slides over the bigger one are of equal length. Hence, it doesn't depend on a particular choice of the point, it has the same meaning whatever the point chosen in the segment.

As, in view of definitions 1, 2 and 3 of Chapters I and II, every random choice can be viewed either as a random choice in a single or compound região ou complexo de regiões, a cada figura lançada à sorte corresponderá, em geral, um ponto que lhe serve de definição no lançamento. A êsse ponto chamaremos ponto *equivalente* da figura em relação ao lançamento considerado. O ponto equivalente deve ser tal que, qualquer que seja a figura em questão, êle não dependa de nenhum dos pontos dessa figura.

Problema

Dois amigos passeiam todas as tardes, durante meia hora, num jardim público que está aberto das duas horas até às qùatro. Qual a probabilidade de que se encontrem em certo dia ?

Resp.:

Supõe-se casual a hora a que qualquer dos indíviduos se dirige para o jardim. Nestas condições, visto que o tempo é um *contínuo* a uma dimensão, podemos enunciar assim êste problema:

Dá-se um segmento de recta de comprimento a e mais dois outros de comprimentos b e c; lançam-se b e c, à sorte, sôbre a: qual a probabilidade de que se sobreponham ?

No caso de $b+c \geqslant a$, os segmentos sobrepõem-se sempre e por isso

$$p = 1.$$

No caso de $b+c < a$, vejamos como as coisas se passam.

Quando b percorrer o segmento a, qualquer dos seus pontos descreve segmentos iguais a $a-b$ e, análogamente, region or as a choice in a complex region, the choice of a random figure corresponds, as a rule, to randomly choosing a point which determines the figure in this random choice. We shall call such a point the *equivalent point* to the figure, in what concerns the random choice at hand. The equivalent point must be independent of all points from the figure, whichever the figure in question.

Problem

Each of two friends goes for a half hour walk to a public garden open from 2 p.m. till 4 p.m., separately. What is the probability that in a given day they meet during their walk in the public garden?

We assume that the time each of them starts his walk is random. Then, as time is continuous in one dimension, the problem may be reformulated as follows:

We randomly throw two segments of lengths b and c, respectively, over a given segment of length a. What is the probability that the two random segments have a nonempty intersection?

If $b + c > a$, the two segments always overlap, and therefore

$$\mathbb{P} = 1.$$

Now we analyze the case $b + c < a$.

When the segment of length b slides over the segment of length a, each of its points describes a segment of length $a - b$ and, similarly, qualquer ponto de c descreve segmentos iguais a $a-c$. Lançar à sorte os dois segmentos sôbre a, é lançar à sorte um ponto de b em $a-b$ e um ponto de c em $a-c$, ou, o que é o mesmo, lançar à sorte um ponto no rectângulo $(a-b)\,(a-c)$.

Escolhamos para pontos equivalentes de b e c as suas extremidades da direita e sejam x e y as distâncias dessas extremidades à origem 0 de a *(fig. 5)*.

$$\begin{array}{cccccccccc} & & c & & x & & b & & y & \\ 0 & \text{———} & | & \text{———} & | & \text{———} & | & \text{———} & | & \text{———}\; a \end{array}$$

Figura 5

Para que b e c se não sobreponham terá de ser

$$y-x \geqslant b \qquad \text{ou} \qquad x-y \geqslant c.$$

As rectas cujas equações são

$$y-x=b \qquad \text{e} \qquad x-y=c$$

determinam sôbre o rectângulo $(a-b)\,(a-c)$ dois semi-quadrados que são, para o problema proposto, a região *contrária*. E tem-se, como é fácil de vêr,

$$1-\mathrm{P} = \frac{\frac{1}{2}(a-b-c)^2 + \frac{1}{2}(a-c-b)^2}{(a-b)\,(a-c)} = \frac{(a-b-c)^2}{(a-b)\,(a-c)}\,.$$

Para

$$a=2,\ b=c=\frac{1}{2}\,,$$

vem

$$\mathrm{P} = 1-\frac{4}{9} = \frac{5}{9}\,.$$

each point of the segment of length c generates segments of length $a - c$. Randomly throwing the two segments over the segment of length a is the same as randomly throwing one point of the segment of length b in $a - b$ and one point of the segment of length c in $a - c$; and this is the same as randomly throwing one point in the rectangle $(a - b) \times (a - c)$.

Let's take as equivalent points of those segments, in what concerns the random throw described, their right extremities, and let x and y denote the distances of those right extremities to the origin 0 of the segment of length a (Fig. 5).

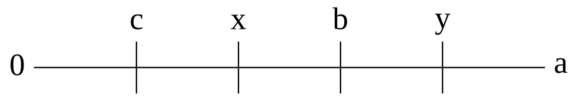

Figure 5

The two random segments don't overlap if and only if

$$y - x > b \qquad \text{or} \qquad x - y > c.$$

The lines with equations

$$y - x = b \qquad \text{and} \qquad x - y = c$$

determine on the rectangle $(a - b) \times (a - c)$ two half-squares which are, in what concerns this problem, the *contrary* region. It is then easy to compute

$$1 - \mathbb{P} = \frac{\frac{1}{2}(a-b-c)^2 + \frac{1}{2}(a-c-b)^2}{(a-b)(a-c)} = \frac{(a-b-c)^2}{(a-b)(a-c)}\,.$$

In the special case

$$a = 2, b = c = \frac{1}{2}\,,$$

we get

$$\mathbb{P} = 1 - \frac{4}{9} = \frac{5}{9}\,.$$

DEFINIÇÃO 3.ª

Lançar, à sorte, uma recta numa região (A), significa, por definição, *lançar, à sorte, um ponto em* (A) *e tirar, à sorte, uma direcção* sendo a recta definida por êstes dois elementos.

Exemplo:

Lança-se, à sorte, uma recta dentro dum círculo; qual a probabilidade de que ela intersecte uma corda inferior ao segmento *c*?

Para lançar, à sorte, a recta, principiamos por tirar, à sorte, a sua direcção. Em virtude da simetria do círculo, a qualquer direcção da recta corresponde sempre, para o ponto lançado no círculo, o mesmo campo favorável S′ (menor segmento de círculo limitado pela corda *c*) e o mesmo possível S (área total). Logo (prop. VIII) a probabilidade pedida será independente da direcção D *(fig. 6)* da recta que, por isso, suporemos fixa e dependerá apenas de M. Logo

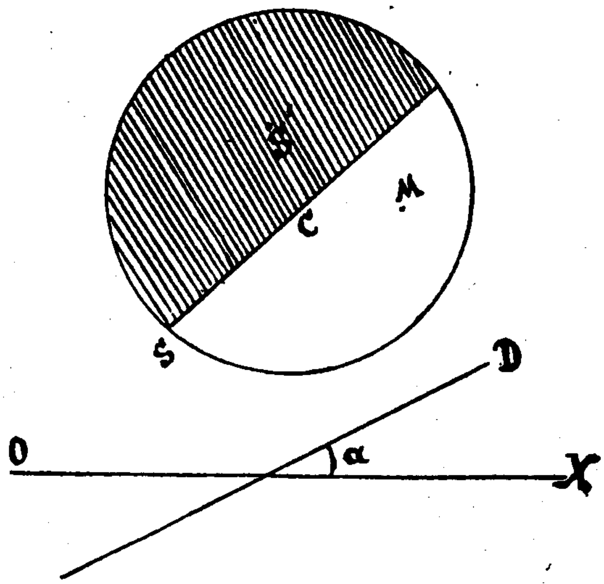

Figura 6

$$P = \frac{S'}{S} = \frac{\alpha - \operatorname{sen} \alpha}{\pi}$$

DEFINITION 3

To throw a straight line at random in a given region A means, by definition, *to throw a point, at random, in A, and to select at random one direction in the region A*, which determine the straight line.

Example[12]:

A straight line is randomly thrown in a circle. What is the probability that its intersection with the circle is a chord smaller than a given chord of length c?

Without loss of generality, we shall solve the problem in the unit circle S, i.e. with area π. Any chord of length c defines a smaller arc of amplitude α, say, and the area of the corresponding smaller circular segment is $\frac{\alpha-\sin\alpha}{2}$. Therefore (Prop. VIII of Chapter II) the probability wanted is independent of the direction D of the random straight line.

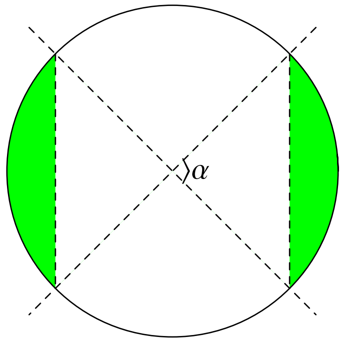

Figure 6

Whatever the direction of the random straight line, its intersection with the circle is a chord of length smaller than c if and only if the associated smaller circular segment has area less than $\frac{\alpha-\sin\alpha}{2}$.

Therefore when throwing a random chord in the circle, for any randomly chosen direction, any point in the circle is a possible point so that the straight line is thrown in the circle (Def. 3), and the favorable points M are those in the [two] circular segments S'_1 [and S'_2], with area $\frac{\alpha-\sin\alpha}{2}$ each, defined by the straight line[s] with the given direction whose intersection with the circle is a chord of length c.

Thus, the probability in question is

$$\mathbb{P} = \frac{2 \cdot \text{area of the circular segment } S'}{\text{area of the circle } S} = \frac{\alpha - \sin\alpha}{\pi},$$

(12) *Editors' note*: We have corrected the misprints in the formulation and drawn a figure more suited to follow the arguments in the solution given by Pacheco d'Amorim.

sendo α o ângulo subintendido pela corda c, visto que a corda será maior ou menor que c, conforme M estiver ou não sôbre S'.

*

A solução generaliza-se imediatamente ao caso duma recta lançada numa esfera.

A mesma solução se daria ao problema que dêste resulta substituindo a recta por um plano e a corda pela área de uma secção plana.

DEFINIÇÃO 4.ª

Lançar, à sorte, um segmento de recta numa região (A) significa, por definição, *lançar, à sorte, uma recta em* (A) (def. 3.ª) *e na parte dessa recta interior a* (A), *o segmento em questão* (def. 2.ª)

Exemplo:

Problema da agulha

Lança-se, à sorte, uma agulha (segmento da recta) sôbre uma folha de papel (plano ilimitado) cortada por linhas rectas paralelas e equidistantes; qual a probabilidade de que a agulha encontre uma dessas rectas?

Resp.:

Seja l o comprimento da agulha e a a distância das paralelas AB, A'B', ... *(fig. 7)*.

Principiemos por tirar, à sorte, a direcção α da recta que contêm o segmento; em seguida teremos de lançar à sorte um ponto M dentro dum segmento do plano arbitráriamente grande (cap. II, def. 4.ª). Por maior, porêm, since the chord which is the intersection of the random straight line with the circle will have length less or equal to c if and only if the point M lies on a segment in a region $S' = S'_1 \cup S'_2$ with area $\alpha - \sin\,\alpha$.[(13)]

*

The solution has immediate generalization for the case of a straight line thrown in a sphere.

The same solution applies, with the necessary modifications, for a similar problem, where the straight line is replaced by a plane and the chord by the area of a plane section.

DEFINITION 4

To throw, at random, one straight line segment in a region A means, by definition, *to throw a straight line at random in A* (Def. 3) *and to throw the segment in question* (Def. 2) *in the segment which is the intersection of A with the random straight line.*

Example:

The needle problem

A needle (straight line segment) is randomly thrown over a sheet of paper (unlimited plane) where parallel and equidistant straight lines have been drawn. What is the probability that the needle intersects one of those straight lines?

Let l denote the needle length, and δ denote the distance between the parallel straight lines $\overline{ab}$, $\overline{a'b'}$, … (Fig. 7).

We first randomly select the direction α of the straight line that contains the segment; next we throw a random point X in a portion arbitrarily large of the plane (Chapt. II, Def. 4). However big

(13) *Editors' note*: Although in the explanation Pacheco d'Amorim seems to overlook the fact that there are two circular segments, symmetrical in respect to the diameter with the chosen direction, whose points are favorable, the final expression he presents is correct. We have introduced the necessary corrections in his arguments.

Note however that, in our opinion, this ingenious solution he gives to the problem, of the class of the famous Bertrand's paradoxes, has a flaw. In fact, all the points that lie in a given chord with the given direction will correspond to the **same** randomly thrown straight line of the given direction, in Pacheco d'Amorim's definition, and it is obviously untrue that the two sets of points that lie in two chords of different lengths carry equal probability.

que seja êsse segmento do plano, poderemos sempre determinar um paralelogramo P Q R S com os lados paralelos à direcção achada e às rectas A B, A′ B′, paralelogramo êsse que contenha o segmento do plano considerado e fazer dentro dêle o lançamento de M. Suponhamos que estes dois lançamentos nos determinaram a recta D E. Determinada ela, resta lançar o segmento l sôbre o segmento D E (def. 4.ª). Ora, qualquer que seja a posição do ponto M, a probabilidade de l encontrar uma das rectas é sempre a mesma. Logo (cap. II, def. 8.ª) podemos supôr M fixo. Posto isto, façamos percorrer a l a recta D E. Quando a origem de l percorrer um segmento de D E compreendido entre duas paralelas consecutivas, o espaço durante o qual êsse segmento encontra uma das paralelas é sempre o mesmo.

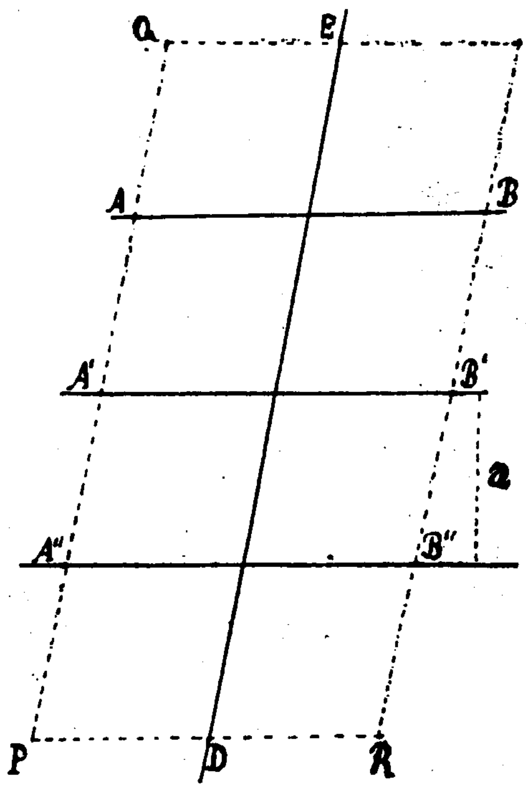

Figura 7

As paralelas dividem, pois, o campo possível do lançamento de l sôbre D E em partes de igual probabilidade (desprezando as duas partes extremas, o que é legítimo em virtude da arbitrariedade do paralelogramo); logo (def. 9.ª) bastará calcular a probabilidade numa dellas:

$$p = \frac{l}{BB'} = \frac{l \operatorname{sen} \alpha}{a},$$

this portion of the plane, we may always define in it a parallelogram $[pqsr]$ whose sides are parallel to the direction selected, and whose bases are parallel to the straight lines $\overline{ab}$, $\overline{a'b'}$, containing that portion of the plane, and throw the point X inside it. Assume that these two random throws have determined the straight line $\overline{de}$. Once this random straight line has been thrown, the next step is to throw the segment of length l over the segment $\overline{de}$ (Def. 4). Or, whatever the position of point X, the probability that the segment of length l intersects one of the parallel straight lines is always the same. Therefore (Chapt. II, Prop. 8) we may take X as fixed.

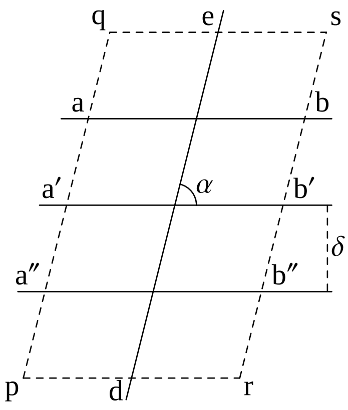

Figure 7

Let the segment of length l slides over the straight line $\overline{de}$; when its origin goes from one parallel to the next one, the segment it generates while the segment of length l intersects the next parallel straight line has always the same length.

Thus the parallel straight lines divide the possible region where the segment of length l is randomly thrown over $\overline{de}$ in subregions with equal probability (with the exception of the first and of the last ones, which may be discarded, in view of the arbitrary size of the parallelogram). Therefore, in view of Prop. IX of Chapt. II, it is enough to compute the probability in one of them:

$$\mathbb{P} = \frac{l}{\overline{bb'}} = \frac{l \sin \alpha}{\delta},$$

(sendo α o ângulo EDR), caso $a > l$; por outro lado, a probabilidade por unidade de α é $\frac{1}{\pi}$; logo

$$P = \int_0^{\pi} \frac{1}{\pi} \cdot \frac{l \operatorname{sen} \alpha}{a} \, d\alpha = \frac{l}{a\pi} \left(-\cos \alpha\right)_0^{\pi} = \frac{2\,l}{a\pi}.$$

No caso de $a < l$, decomponhamos o campo da variação de α em duas partes: a 1.ª constituida pela totalidade

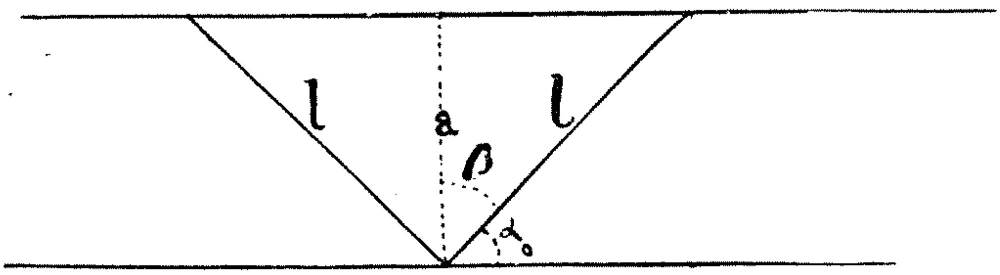

Figura 8

de valores de α para os quais pode não haver o encontro; a 2.ª constituida pelos outros valores. Aplicando os teoremas das probabilidades totais e compostas e pondo

$$a = l \cos B = l \operatorname{sen} \alpha_0,$$

virá

$$P = \frac{\pi - 2\alpha_0}{\pi} \cdot 1 + 2 \cdot \frac{\alpha_0}{\pi} \cdot \int_0^{\alpha_0} \frac{l \operatorname{sen} \alpha}{a} \cdot \frac{d\alpha}{\alpha_0}$$

$$= \frac{\pi - 2\alpha_0}{\pi} + \frac{2\,l}{a\pi} \left[1 - \cos \alpha_0\right]$$

$$= \frac{2B}{\pi} + \frac{2\,l}{a\pi} (1 - \operatorname{sen} B).$$

(where α is the angle $\langle [edr] \rangle$), in case $\delta \geq l$; on the other hand, the unit possibility of α is $\frac{1}{\pi}$; therefore

$$\mathbb{P} = \int_0^{\pi} \frac{1}{\pi} \cdot \frac{l \sin \alpha}{\delta} \, \mathrm{d}\alpha = \frac{l}{\delta \, \pi} \left(-\cos \alpha \right]_0^{\pi} = \frac{2l}{\delta \, \pi} \, .$$

In the case $\delta \leq l$, let us partition the field of variation of α in two parts: the first one the totality of values of α for which the intersection is void;

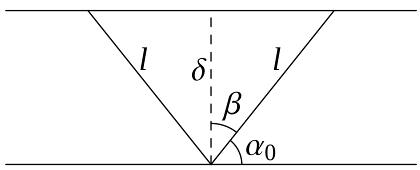

Figure 8

the second one with all the other values. Using the theorems of total probability and of compound probability, and writing

$$\delta = l \cos \beta = l \sin \alpha_0 ,$$

we get

$$P = \frac{\pi - 2\alpha_0}{\pi} \cdot 1 + 2 \, \frac{\alpha_0}{\pi} \int_0^{\alpha_0} \frac{l \sin \alpha}{\delta} \cdot \frac{\mathrm{d}\alpha}{\alpha_0} =$$

$$= \frac{\pi - 2\alpha_0}{\pi} + \frac{2l}{\delta \, \pi} \left[1 - \cos \alpha_0 \right] =$$

$$= \frac{2\beta}{\pi} + \frac{2l}{\delta \, \pi} \left(1 - \sin \beta \right) .$$

Nota

Convêm notar que a probabidade de encontro é proporcional ao comprimento da agulha logo que $l < a$.

DEFINIÇÃO 5.ª

Lançar, à sorte, um plano (1) *dentro duma região* (A) *também plana,* significa, por definição, *tirar, à sorte, em* (A) *um sentido.*

DEFINIÇÃO 6.ª

Lançar, à sorte, uma região plana dentro de outra também plana (A), significa, por definição, *lançar em* (A) *um plano* (def. 5.ª) *e em seguida lançar, à sorte, um ponto da região móvel ou a ela invariavelmente ligado, dentro da região que êsse mesmo ponto descreve quando a figura móvel ocupa dentro de* (A) *todas as posições compatíveis com a orientação que a sorte designa para o seu plano.*

Evidentemente que o ponto equivalente terá sempre a mesma possibilidade qualquer que seja o ponto escolhido no plano da figura móvel, visto que, deslocando-se o plano da figura móvel paralelamente a si mesmo, todos os seus pontos descrevem regiões iguais.

---

(1) Êste plano fica perfeitamente orientado logo que se dê dêle uma semi-recta.

Note

It is worth noting that when $l < \delta$ the probability that the needle intersects one of the parallel straight lines is directly proportional to the length of the needle.

DEFINITION 5

To throw a plane[14] *at random in a plane region A* means, by definition, *to choose randomly an orientation in A.*

DEFINITION 6

To throw, at random, a plane region in another plane region A means, by definition, *to throw a random plane in A* (Def. 5) *and then to throw, at random, a point of the mobile region (or a point invariably tied to that region) in the region that this point defines when the mobile plane region occupies inside A all the positions that are compatible with the orientation randomly chosen in the first step.*

It is obvious that an equivalent point will always have the same possibility, whichever the randomly chosen point in the plane of the mobile figure, since all the points of the mobile figure describe identical regions when the plane containing it moves taking on positions which are parallel to each other.

---

[14] The orientation of this plane is determined by a half-line.

DEFINIÇÃO 7.ª

Lançar, à sorte, um plano num espaço a n $>$ 3 *dimensões,* significa, por definição, *lançar, à sorte, um ponto nesse espaço e, á volta dele, dois sentidos.*

DEFINIÇÃO 8.ª

Lançar, à sorte, num espaço a n *dimensões, uma figura plana,* significa, por definição, *lançar, à sorte, um plano nessa região e, sôbre a parte do plano interior a essa região, a figura plana em questão.*

DEFINIÇÃO 9.ª

Lançar, à sorte, um espaço a três dimensões, numa região (A) *a três dimensões também,* significa, por definição, *lançar, à sorte, um ponto em* (A), *e à volta dele dois sentidos* (¹).

*

A generalização destas definições para espaços de um número superior de dimensões é intuitiva.

Nota

Convêm notar que, segundo estas definições, lançar, à, sorte, a região finita (B) sôbre a região também finita (A), é equivalente a lançar, à sorte, um ponto numa região que depende de (A) e (B). Desta dependência se conclue imediatamente basta analisar um caso particular, o do lan-

(¹) Bastam dois sentidos para orientar três eixos.

DEFINITION 7

To throw a plane at random in an n-dimensional space, $n \geq 3$ means, by definition, *to choose a random point and two directions emerging from it in that space.*

DEFINITION 8

To throw randomly a plane figure in an n-dimensional region means, by definition, *to throw a random plane in that region and then to throw at random that figure in the portion of the random plane inside that region.*

DEFINITION 9

To throw at random a three dimensional space in a three dimensional region A means, by definition, *to throw a point at random in A and to choose two random directions*[15] *emerging from it.*

*

The above definitions are easily extended for higher dimensional spaces.

Note

It is worth observing that, according to the above definitions, to throw at random a finite region B on another finite region A is equivalent to randomly throwing a point in a region that depends both from A and from B. From that dependence we may immediately conclude (it is enough to analyze one particular case, for instance throwing a straight line segment on a rectangular plane region)

[15] Two directions is sufficient to direct three axes.

çamento dum segmento de recta sôbre um segmento plano, por exemplo) que do lançamento, à sorte, do *todo,* se não conclue o lançamento, à sorte, de qualquer das *partes,* porque, no caso de regiões finitas, o ponto equivalente duma parte de (B) variará numa região diferente da do ponto equivalente de (B), isto é, lançando a parte junta ao todo, o campo de variação do seu ponto equivalente será um (campo do ponto de equivalência do todo), lançada separadamente, o seu campo de variação será outro.

No caso do lançamento ser feito em campo ilimitado, pode dar-se o caso de o lançamento à sorte do todo, arrastar o lançamento à sorte, da parte.

Assim, por exemplo, no problema da agulha (def. 4.ª), o lançamento à sorte da agulha arrasta o lançamento à sorte de qualquer das suas partes, porque, quer uma parte da agulha seja lançada, à sorte, isoladamente, quer fazendo parte da agulha, o seu ponto equivalente terà sempre o mesmo campo possível.

SEGUNDA PARTE

Lançamento, à sorte, de figuras variáveis

DEFINIÇÃO 10.ª

Lançar, à sorte, uma figura variável numa dada região (A), significa, por definição, *tirar, à sorte, a forma dêssa figura e, em seguida, lançá-la, á sorte, em* (A), *como se fosse rígida.*

A segunda proposição de que se compõe esta definição that from the random throw of all B we cannot conclude the random throw of any of its parts, because in the case of finite regions, the equivalent point of a part of B will vary in a different region of the equivalent point of all B, that is, if we throw together all region B the field of variation of its equivalent point (field of the equivalent point of the all) will be different from its field of variation if we throw the parts of B separately.

In the special case of random throws in an unbounded region, it may happen that the global random throw determines the partial random throws.

For instance, in the [Buffon's] needle problem (Def. 4), the random throw of the needle determines the random throw of any of its parts, since we get the same result conceptualizing the random throw of part of the needle either in isolation or as part of the needle, because its equivalent point would have the same field of variation in both cases.

SECOND PART

Random variable figures

DEFINITION 10

Randomly throwing a variable figure in a given region A is, by definition, *to select, at random, the form of the figure and, then, to throw it at random in A, as if it were a rigid figure.*

The discussion of the second statement in this definition já foi tratada na primeira parte dêste capítulo. Resta-nos apenas dizer o que entendemos por «tirar, à sorte, a forma duma figura variável».

Evidentemente que a questão não pode ser resolvida dum modo geral. Nós trataremos a questão das figuras poligonais (abertas ou fechadas) articuladas e, como caso limite, as curvas flexíveis e inextensíveis.

DEFINIÇÃO 11.ª

Figuras poligonais abertas

Uma figura poligonal articulada é, como se sabe, uma linha poligonal de ângulos variáveis.

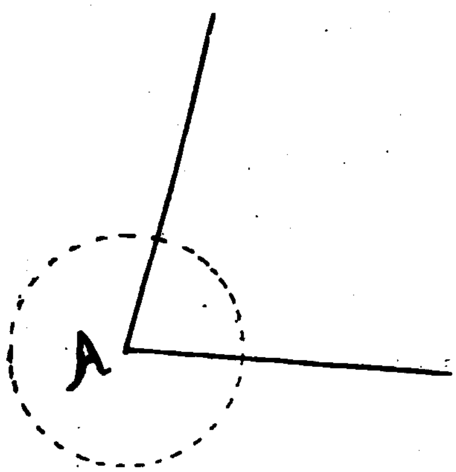

Figura 9

Fixar, à sorte, a sua forma, será, por definição, *tirar, à sorte, a forma de cada um dos seus vértices.*

Para definir esta última proposição, consideremos um dêsses vértices ou articulações A *(fig. 9)* que podemos supôr localizado no espaço a n dimensões e sejam $(x'_1, x'_2, \ldots x'_n)$ as coordenadas de A. Consideremos o espaço formado pelos pontos que satisfazem à equação

$$(x_1 - x'_1)^2 + \ldots + (x_n - x'_n)^2 = 1. \tag{1}$$

Suponhamos um dos ramos da articulação A fixo e o outro movendo-se de modo a ocupar em relação ao fixo has been done in the first part of the present Chapter. So, our present task is to discuss the meaning of randomly choosing the form of a variable figure.

Obviously this question cannot have an exhaustive treatment. We shall limit ourselves to articulated polygonal figures (either open or closed) and, as a limit case, of flexible inextensible curves.

DEFINITION 11

Open polygonal lines

An articulated polygonal figure is a polygonal line whose consecutive segments form variable angles.

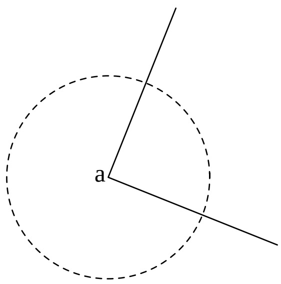

Figure 9

To choose, at random, its form is, by definition, *to choose, randomly, the form of each of its vertices.*
To explain the meaning of this last statement, let us consider one vertex or articulation a (Fig. 9), which we may assume to be an element of the n-dimensional space, with coordinates $(x'_1, x'_2, \ldots, x'_n)$. Let us consider the hyperspherical surface defined by

$$(x_1 - x'_1)^2 + \cdots + (x_n - x'_n)^2 = 1. \tag{3.1}$$

We assume that one of the sides of the articulation a is fixed, and that the other one can occupy any of the possible positions; therefore, at distance 1 from the vertex, it intersects the hyperspherical surface (3.1). Denote B the set of such intersection points [16].

[16] It is possible that B does not correspond to the total hyperspherical surface.
Editors' note: This is a mysterious footnote. How can the region B being constrained when we are choosing, at random, the form of one vertex of an open polygonal line? Perhaps this footnote is in the wrong place and it's relate about the random choose of the form of a closed polygonal line where it makes sense.

todas as posições possíveis. Seja (B) [1] a região por êle descrita sôbre o espaço (1). *Tirar, à sorte, a forma da articulação* (A) é, por definição, *lançar, à sorte, um ponto em* (B).

DEFINIÇÃO 12.ª

Figuras poligonais fechadas

a)

No plano

Principiaremos pelas figuras planas e passaremos em seguida para as figuras articuladas no espaço.

1.º

Suponhamos que se trata de tirar, à sorte, a forma dum

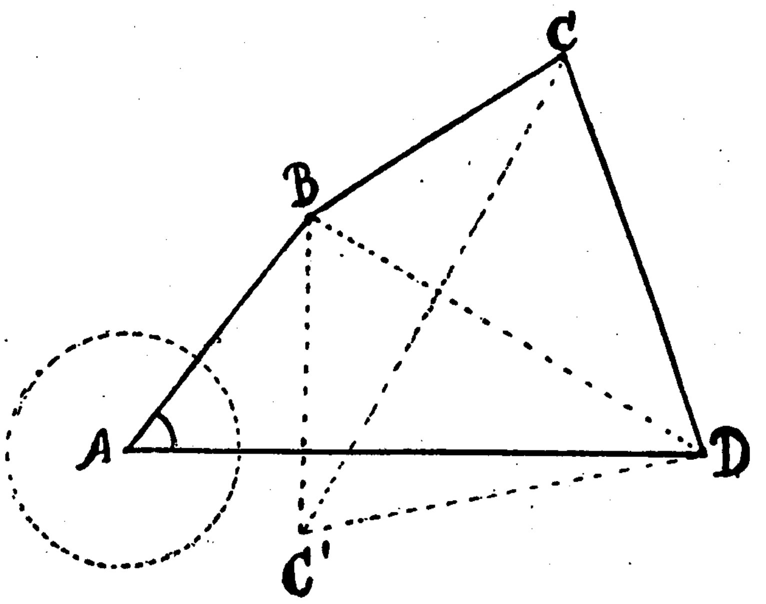

Figura 10

quadrilátero plano AB CD *(fig. 10)*. Quando êste quadrilá-

(1) (B) pode não ser a totalidade da superfície da hiper-esfera.

Thus, *to choose at random the form of the articulation a* is, by definition, *to choose a random point in the subset B of the hyperspherical surface.*

DEFINITION 12

Closed polygonal lines

a)

In the plane

We begin with plane figures, and next we shall discuss articulated figures in the higher dimensional spaces.

1st

Let's discuss, to start with, how to choose randomly the form of a four

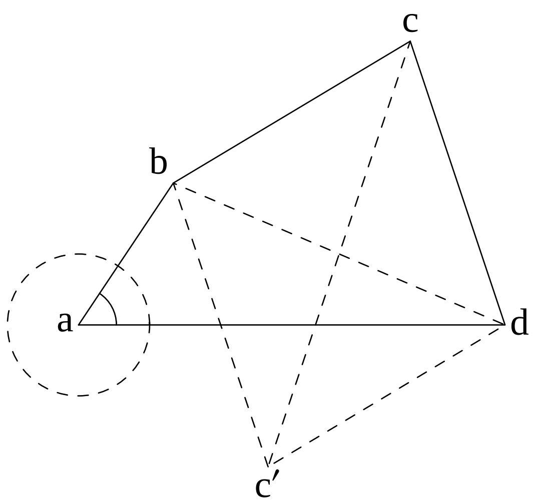

Figure 10

sided plane polygon [*abcd*] (Fig. 10). When tero passa por todas as formas possíveis, o ângulo A passa por duas espécies de valores: uns correspondentes à posição C do vértice que lhe é oposto; outros correspondentes à posição C′. Ou, o que é o mesmo, supondo que à volta de A se descrevera uma circunferência de raio unidade, e que AD se considera fixo, o ponto de intersecção do lado invariável com a circunferência, descreve duas regiões que podem sobrepôr-se no todo ou em parte [1] quando o quadrilátero passa por todas as posições possíveis. Mas nem por isso essas regiões deixam de ser distintas uma da outra. Representêmo-las por (A) e (A_1). Logo que seja dado um ponto qualquer destas duas regiões, a figura fica definida na forma. Consideremos as mesmas duas regiões em B, C, D, e sejam (B) e (B_1), (C) e (C_1), (D) e (D_1).

Tirar, à sorte, a forma do quadrilátero será, por definição, *tirar, à sorte, um ponto de* (A) *ou* (A_1) *ou* (B) *ou* (B_1) *ou* (C) *ou* (C_1) *ou* (D) *ou* (D_1) (def. I, cap. II).

Segundo esta definição o ponto equivalente não fica dependente de nenhum elemento da figura.

2.º

Consideremos, agora, um pentágono e vejamos como passar do lançamento dum quadrilátero para o desta figura.

Quando a articulação A *(fig. 11)* toma a forma particular A, o quadrilátero BCDE pode tomar uma infinidade

[1] As articulações podem estar sujeitas a ligações que não tornem possível a posição C′.

this polygon assumes all possible forms, its angle a can take values of one of two kinds: those corresponding to the position c of its non-adjacent vertex, and those corresponding to the position c' of its non-adjacent vertex. In other words, drawing a circumference with unit radius with center a, and considering the side $\overline{ad}$ fixed, the intersection point of the moving side with the circumference defines two regions, that can have non-empty intersection [17] when the polygon assumes all its possible forms. But we shall, in all cases, consider the two regions, which we denote A and A_1, as distinct. Once one point from one of those regions is given, the form of the figure has been determined. Consider similar regions with vertices b, c and d, and denote them B and B_1, C and C_1, D and D_1, respectively.

Randomly choosing the form of the plane four sided polygon is, by definition, *to choose randomly one point* (Chapter II, Def. 1) *from A or A_1, or from B or B_1, or from C or C_1, or from D or D_1.*

With this definition, the equivalent point doesn't depend on any element of the figure.

2nd

Let us now consider a pentagon, and investigate how to progress from the previous case to the random choice of a pentagon.

When the articulation a takes on the particular form shown in Fig. 11, the four sided plane polygon $[bcde]$ can take an infinity number

[17] The articulations can be subjected to restrictions such that the position c' is inadmissible.

de formas, umas no semi-plano BCE, outras no semi-plano BC'E.

A essas formas correspondem grupos (B) e (B'), (C) e (C'), (D) e (D'), (E) e (E'), para os ângulos de BCDE e grupos análogos, mas diferentes em geral, para o quadrilátero BC'D'E. Associemos os valores possíveis de A a cada um dos elementos dos grupos de BCDE e chamemos (A) ao

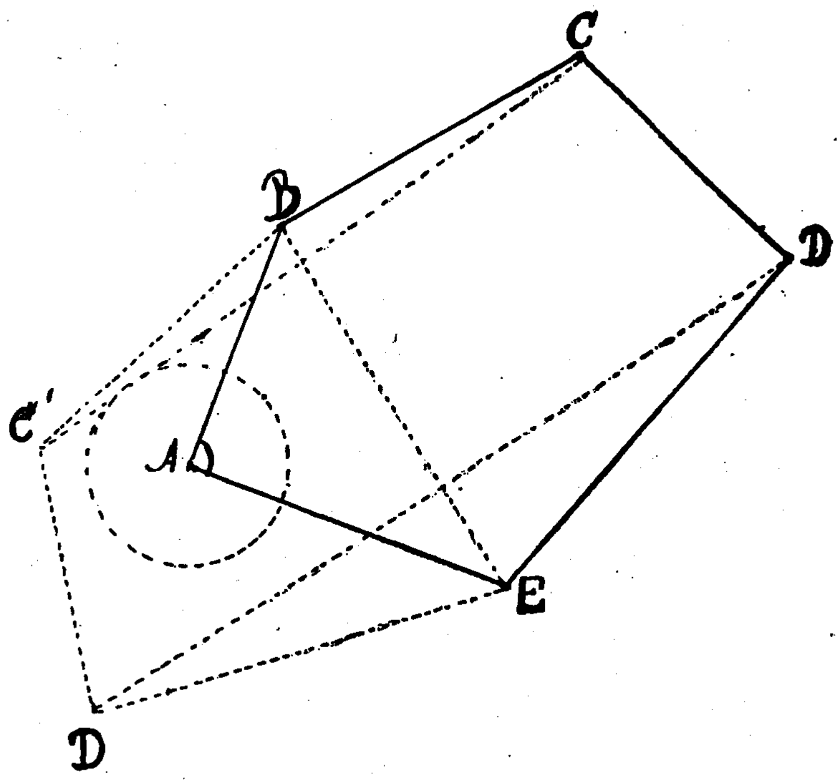

Fig. 11

conjunto obtido. Façamos o mesmo para o quadrilátero BC'D'E e chamemos ao conjunto obtido (A_1). Um elemento qualquer dêstes dois conjuntos basta para definir o pentágono. Façamos o mesmo para todos os outros vértices e sejam (B) e (B_1), (C) e (C_1), ..., os conjuntos obtidos.

Tirar, à sorte, a forma dum pentàgno, significa, por definição, *tirar, à sorte, um elemento de* (A) *ou* (A_1) *ou* (B) *ou* (B_1) ...

of forms, some in the half-plane $[bce]$, others in the half-plane$[bc'e]$.

These forms correspond to groups B and B', C and C', D and D', E and E', for the angles of $[bcde]$, and similar groups, but in general different ones, for the angles of $[bc'd'e]$. Let's associate the possible values of a with each of the elements of the groups $[bcde]$, and denote A the set thus obtained.

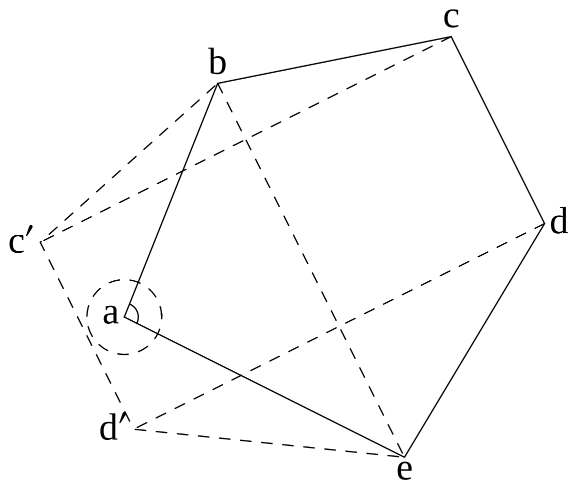

Figure 11

Let's do the same in what regards $[bc'd'e]$, and denote A_1 the set thus obtained. Any element of any of those sets will define the pentagon. Let's do the same with all the other vertices, and denote B and B_1, C and C_1, ..., the sets obtained as described.

Randomly choosing a pentagon is, by definition, *to choose, randomly, an element from A or A_1, or from B or B_1, ...*

Dum modo análogo se definia a tiragem, à sorte, da forma dum exágono, a dum eptágono, etc.

b)

No espaço

Para o lançamento duma figura poligonal fechada num espaço a um número qualquer de dimensões, não teremos mais do que substituir, nas definições precedentes, os pontos que se moviam sôbre circunferências, por pontos movendo-se sôbre hiper-esferas.

Lançamento de curvas flexíveis e inextensíveis

Por definição, *lançar, à sorte, uma curva flexível e inextensível, num espaço* (A) será, quer a curva seja aberta quer fechada, *lançar, à sorte, um polígono do mesmo comprimento e dum número de lados arbitráriamente grande.*

*

Todo o problema relativo a um polígono articulado, com um número arbitrário de lados, terá uma solução dependente do número e grandeza dêsses lados. Se essa solução tender para um limite quando os lados do polígono tenderem para zéro, dir-se-há que êsse limite é a solução do mesmo problema relativo a uma curva flexível e inextensível.

The definition of random choice of a hexagon, or of a heptagon, etc., is similar in all points.

b)

In the space

Randomly throwing a closed polygonal line in an n-dimensional space is in all points similar to what we have seen about randomly throwing a closed polygonal line in the plane, we only need to substitute, in the preceding definitions, the points varying in circumferences by points varying in hyperspherical surfaces[18].

Randomly throwing flexible inextensible curves

By definition, *randomly throwing a flexible inextensible curve*, open or closed *in the space A* is *to throw in that space, at random, a polygon with the same length and an arbitrary large number of sides.*

*

Any problem referring to an articulated polygon with an arbitrary number of sides will have a solution which depends on the number and length of those sides. If that solution converges for some limit when the supremum length of the polygons sides decreases to zero, we shall say that this limit is the solution of the same problem in the case of a flexible and inextensible curve.

(18) *Editors' note*: Cf. footnote 13.

CAPÍTULO IV

DO PONTO IMAGEM

CHAPTER IV

IMAGE POINT

CAPITULO IV

PONTO IMAGEM

Proposição I

Sejam (A) e (B) duas regiões tais que entre os seus pontos se possa estabelecer uma correspondência biunivoca, completa e contínua. Quando se lança, à sorte, um ponto M em (A), o ponto N, correspondente em (B), designa-se com o nome de *ponto imagem* do primeiro. Dum modo geral, qualquer elemento de (B) diz-se *imagem* do elemento correspondente de (A), logo que em (A) se efetue um lançamento à sorte.

A posição do ponto imagem N ficará dependente do acaso, visto que depende da posição do ponto M que é lançado à sorte em (A). Mas a dependência de N em (B) é muito diferente da dependência de M em (A), visto que M é lançado *directamente* à sorte em (A), enquanto que N varía em (B) como *imagem* de (A).

Por esta razão chamaremos a M *ponto livre* e a N ponto imagem ou *sujeito*.

Defeniremos a possibilidade no ponto sujeito, pela possibilidade no ponto livre correspondente. Possibilidade duma região (B′), imagem da região (A′), pela possibilidade de (A′). Dum modo geral, tudo que debaixo do ponto de vista da probabilidade se diz do ponto imagem N, va-

CHAPTER IV

IMAGE POINT

Proposition I

Let A and B be two regions such that it is possible to define a bijective, complete and continuous correspondence between their points. When a point M is randomly chosen in A, the corresponding point N in B is said to be the *image point* of M. Generally, each element in B is the *image* of a corresponding randomly chosen element in A.

The image point N is random, insofar as it depends on the original point M randomly chosen in A. But the random status of M in B is clearly different from the random status of M in A, since its random choice in A has been *direct*, while N randomly varies in B, but as an *image* of M.

For that reason we say that M is a *free point*, while N is an image or *dependent* point.

The possibility of a dependent point is, by definition, the possibility of the corresponding free point. The possibility of a region B' which is the image of a region A' is the possibility of A'. In a general way, all that can be said, in what concerns probability, about an image point N, riando em (B), define-se por meio do ponto livre correspondente variando em (A). As propriedades mencionadas para o ponto livre, são ainda verdadeiras para o ponto imagem, como será fácil de vêr.

Aplicação

Seja $f(x)$ uma função, contínua e crescente, de x, num intervalo (α, β). Tira-se, à sorte, um número dêste intervalo e pergunta-se: qual a probabilidade de que o valor de y correspondente tenha o número dígito d na casa decimal da ordem a?

Sejam ω e ω' os números inteiros que mais aproximadamente satisfazem às desigualdades

$$f(\alpha) < \frac{10\,\omega + d}{10^a}, \qquad f(\beta) > \frac{10\,\omega' + d + 1}{10^a},$$

e representemos por $f^{-1}(x)$ a função inversa de $f(x)$; os valores de $f(x)$ que satisfazem às condições do enunciado estão compreendidos nos intervalos

$$\left(\frac{10\,\omega + d}{10^a}, \frac{10\,\omega + d + 1}{10^a}\right),$$

$$\left(\frac{10\,(\omega + 1) + d}{10^a}, \frac{10\,(\omega + 1) + d + 1}{10^a}\right), \;\ldots\ldots$$

$$\left(\frac{10\,(\omega + i) + d}{10^a}, \frac{10\,(\omega + i) + d + 1}{10^a}\right), \;\ldots\ldots$$

varying in B, is defined *via* the corresponding free point varying in A. It is easily seen that the properties established for the free point [in the previous chapters] are also valid for the image point.

Application

Let f be a continuous increasing function defined in an interval $[\alpha, \beta]$. A number X is randomly chosen in that interval, and we want to know the probability that the corresponding Y has the digit d as its $a-$th decimal.

Let ω and ω' be the integers which most closely satisfy the inequalities

$$f(\alpha) < \frac{10\,\omega + d}{10^a} \qquad f(\beta) > \frac{10\,\omega' + d + 1}{10^a}\,,$$

and let us represent f^{-1} the inverse function of f; the values $f(x)$ which verify the conditions in our problem lie in the intervals

$$\left(\frac{10\,\omega + d}{10^a}, \frac{10\,\omega + d + 1}{10^a}\right),$$

$$\left(\frac{10\,(\omega + 1) + d}{10^a}, \frac{10\,(\omega + 1) + d + 1}{10^a}\right), \ldots$$

$$\left(\frac{10\,(\omega + i) + d}{10^a}, \frac{10\,(\omega + i) + d + 1}{10^a}\right), \ldots$$

$$\left(\frac{10\,\omega'+d}{10^a},\ \frac{10\,\omega'+d+1}{10^a}\right).$$

A probabilidade pedida será, pois, em virtude das definições antecedentes e do teorema das probabilidades totais,

$$P_{(d,\,a)}^{(\alpha,\,\beta)}=\frac{\Sigma\left\{f^{-1}\left[\frac{10\,(\omega+i)+d+1}{10^a}\right]-f^{-1}\left[\frac{10\,(\omega+i)+d}{10^a}\right]\right\}}{\beta-\alpha}$$

1.º

Apliquemos esta fórmula à função

$$y=\log_\alpha x$$

considerada no intervalo de variação de y

$$\left(0,\ \frac{10\,\omega+10}{10^a}\right).$$

Teremos, neste caso,

$$\omega=0,\quad \omega'=\omega,\quad \text{e}\quad P_{(d,\,a)}^{(0,\,\omega)}=\frac{\sum\limits_{i=0}^{\omega}\left[\alpha^{\frac{10\,i+d+1}{10^a}}-\alpha^{\frac{10\,i+d}{10^a}}\right]}{\alpha^{\frac{10\,\omega+10}{10^a}}-1}=$$

$$=\alpha^{\frac{d}{10^a}}\frac{\left[\alpha^{\frac{1}{10^a}}-1\right]\sum\limits_{i=0}^{\omega}\alpha^{\frac{i}{10^{a-1}}}}{\alpha^{\frac{\omega+1}{10^{a-1}}}-1}=\alpha^{\frac{d}{10^a}}\;\frac{\frac{1}{\alpha^{\frac{1}{10^a}}-1}}{\frac{1}{\alpha^{\frac{1}{10^{a-1}}}-1}}. \qquad (1)$$

$$\left(\frac{10\,\omega' + d}{10^a}\,,\,\frac{10\,\omega' + d + 1}{10^a}\right).$$

Taking into account the foregoing definitions and the theorem of total probability, the probability we wish to compute is therefore

$$\mathbb{P}_{(\alpha,\,\beta)}(d,\,a) = \frac{\sum\left\{f^{-1}\left(\frac{10\,(\omega+i)+d+1}{10^a}\right) - f^{-1}\left(\frac{10\,(\omega+i)+d}{10^a}\right)\right\}}{\beta-\alpha}\,,$$

1st

Let us apply the above formula to the function

$$y = \log_\alpha\, x$$

in the variation interval for y

$$\left(0,\,\frac{10\,\omega+10}{10^a}\right).$$

In this case,

$$\omega = 0, \quad \omega' = \omega, \qquad \text{and} \qquad \mathbb{P}_{(\alpha,\,\beta)}(d,\,a) = \frac{\sum\limits_{i=0}^{\omega}\left[\alpha^{\frac{10i+d+1}{10^a}} - \alpha^{\frac{10i+d}{10^a}}\right]}{\alpha^{\frac{10\,\omega+10}{10^a}}-1} =$$

$$= \alpha^{\frac{d}{10^a}}\,\frac{\left[\alpha^{\frac{1}{10^a}}-1\right]\sum\limits_{i=0}^{\omega}\alpha^{\frac{i}{10^{a-1}}}}{\alpha^{\frac{\omega+1}{10^{a-1}}}-1} = \alpha^{\frac{d}{10^a}}\,\frac{\alpha^{\frac{1}{10^a}}-1}{\alpha^{\frac{1}{10^{a-1}}}-1}\,. \tag{4.1}$$

A fórmula (1) mostra que a probabilidade pedida não depende de ω e portanto do intervalo em questão; podemos representá-la por $P_{(d,\,a)}$.

*

De (1) tira-se

$$\frac{P_{(d+1,\,a)}}{P_{(d,\,a)}} = \alpha^{\frac{1}{10^a}}.$$

relação esta independente de d.

Facilmente se vê que $P_{(d,\,a)}$ tende rapidamente para $\frac{1}{10}$, quando a aumenta.

2.º

Consideremos a função

$$y = \alpha^x, \quad \alpha > 1$$

no intervalo correspondente a (ω, ω'); será

$$P_{(d,\,a)}^{(\omega,\,\omega')} = \frac{\sum\limits_{n=\omega}^{\omega'} \left[\log_\alpha \frac{10\,n + d + 1}{10^a} - \log_\alpha \frac{10\,n + d}{10^a}\right]}{\log_\alpha \frac{10\,\omega' + 10}{10^a} - \log_\alpha \frac{10\,\omega}{10^a}}$$

$$= \frac{\sum\limits_{n=\omega}^{\omega'} \log_\alpha \left[1 + \frac{1}{10\,n + d}\right]}{\log_\alpha \frac{\omega' + 1}{\omega}}$$

Expression (4.1) shows that this probability doesn't depend on ω, and thus doesn't depend on the interval of variation of y; therefore, it can be represented $\mathbb{P}(d,\, a)$.

*

From (4.1) it is clear that

$$\frac{\mathbb{P}(d+1,\, a)}{\mathbb{P}(d,\, a)} = \alpha^{\frac{1}{10^a}},$$

independent of d.

It is easily seen that $\mathbb{P}(d,\, a)$ goes quickly to $\frac{1}{10}$ when a increases.

2nd

Let us now consider the function

$$y = \alpha^{x}, \quad \alpha > 1$$

in the interval corresponding to $(\omega,\, \omega')$; we get

$$\mathbb{P}_{(\omega,\,\omega')}(d,\, a) = \frac{\displaystyle\sum_{n=\omega}^{\omega'} \left(\log_\alpha \frac{10\,n + d + 1}{10^a} - \log_\alpha \frac{10\,n + d}{10^a} \right)}{\log_\alpha \frac{10\,\omega' + 10}{10^a} - \log_\alpha \frac{10\,\omega}{10^a}} =$$

$$= \frac{\displaystyle\sum_{n=\omega}^{\omega'} \log_\alpha \left(1 + \frac{1}{10\,n + d} \right)}{\log_\alpha \frac{\omega' + 1}{\omega}},$$

o que mostra que a probabilidade pedida é independente de a.

*

Facilmente se vê que

$$\lim_{\omega'=\infty} P_{(d,\,a)}^{(\omega,\,\omega')} = \frac{1}{10}.$$

Para o provar, mostraremos primeiro que a sucessão cujo termo geral é

$$U_n = \frac{\sum\limits_{n=\omega}^{n} \log_\alpha\left(1+\frac{1}{10\,n+d}\right)}{\sum\limits_{n=\omega}^{n} \log_\alpha\left(1+\frac{1}{n}\right)}$$

é crescente; em seguida que se mantêm inferior a 1.

a)

A sucessão U_n é crescende quando $d > 9$ (d pode ser um número qualquer, na expressão de U_n).

Com efeito, a função

$$f(n) = \frac{\log_\alpha\left[1+\frac{1}{10\,n+d}\right]}{\log_\alpha\left(1+\frac{1}{n}\right)}$$

showing that this probability is independent of a.

*

It is easily seen that

$$\lim_{\omega' \to \infty} \mathbb{P}_{(\omega,\, \omega')}(d,\, a) = \frac{1}{10}\,.$$

In order to establish this result, we shall first show that the sequence with general term

$$U_n(d) = \frac{\displaystyle\sum_{k=\omega}^{n} \log_\alpha\left(1 + \frac{1}{10\,k + d}\right)}{\displaystyle\sum_{k=\omega}^{n} \log_\alpha\left(1 + \frac{1}{k}\right)}$$

is increasing; then, we shall prove that it is upper bounded by 1.

a)

The sequence $U_n(d)$ is increasing when $d > 9$ (in the expression of $U_n(d)$, d may be any [integer] number).

In effect

$$f(k) = \frac{\log_\alpha\left(1 + \dfrac{1}{10\,k + d}\right)}{\log_\alpha\left(1 + \dfrac{1}{k}\right)}$$

que tem por derivada

$$f'(n)=\frac{\frac{-10}{(10n+d)\,(10n+d+1)}\log_\alpha\left(1+\frac{1}{n}\right)+\frac{1}{n\,(n+1)}\log_\alpha\left(1+\frac{1}{10n+d}\right)}{\log\alpha\,.\,\log^2_\alpha\left(1+\frac{1}{n}\right)}$$

é crescente, pois $f'(n)>0$. Com efeito, visto que

$$\left(1+\frac{1}{x}\right)^x$$

é uma função crescente, será

$$\left(1+\frac{1}{10n+d}\right)^{10n+d}>\left(1+\frac{1}{n}\right)^n$$

e também, se $d\geqslant 9$

$$\left\{\left(1+\frac{1}{10n+a}\right)^{10n+d}\right\}^{10n+d+1}>\left\{\left(1+\frac{1}{n}\right)^n\right\}^{10n+10};$$

tirando os logaritmos de base α a esta desigualdade, vê-se imediatamente que

$$f'(n)>0.$$

Ora, U_n é um quebrado que tem por numerador a soma dos numeradores de

$$f(1),\ f(2),\ \ldots\ f(n)$$

is an increasing function, since its derivative

$$f'(k) = \frac{-\dfrac{10}{(10\,k+d)(10\,k+d+1)}\log_{\alpha}\left(1+\dfrac{1}{k}\right)+\dfrac{1}{k\,(k+1)}\log_{\alpha}\left(1+\dfrac{1}{10\,k+d}\right)}{\log\,\alpha\,\left[\log_{\alpha}\left(1+\dfrac{1}{k}\right)\right]^{2}}$$

is such that $f'(k) > 0$. In effect, as

$$\left(1+\frac{1}{x}\right)^{x}$$

is an increasing function, we get

$$\left(1+\frac{1}{10\,k+d}\right)^{10\,k+d} > \left(1+\frac{1}{k}\right)^{k}$$

and, on the other hand, if $d \geq 9$,

$$\left[\left(1+\frac{1}{10\,k+d}\right)^{10\,k+d}\right]^{10\,k+d+1} > \left[\left(1+\frac{1}{k}\right)^{k}\right]^{10\,k+10};$$

taking basis α logarithms on both sides of the above inequality, it is immediate that

$$f'(k) > 0.$$

As the general term of the sequence U_n is a fraction whose numerator is the sum of the numerators of

$$f(1), f(2), \ldots, f(n)$$

e por denominador a soma dos denominadores e como $f(n)$ é crescente, U_n sê-lo-há também.

2.°

$$U_n < 1$$

qualquer que seja n.

Com efeito

$$\log_\alpha\left(1+\frac{1}{10\,n+d}\right) < \log_\alpha\left(1+\frac{1}{n}\right)$$

logo

$$\Sigma \log_\alpha\left(1+\frac{1}{10\,n+d}\right) < \Sigma \log_\alpha\left(1+\frac{1}{n}\right)$$

e por isso

$$U_n < 1.$$

Logo, U_n tende para um limite quando n tende para o infinito e $d \geqslant 9$.

*

Pela expressão de U_n se vê que

$$U_n(0) \gg U_n(d) \gg U_n(10)$$

se

$$0 \leqslant d \leqslant 10\,;$$

and whose denominator is the sum of their denominators, and as $f(k)$ is increasing, $U_n(d)$ is also increasing.

b)

$$U_n(d) < 1$$

for any n.

In effect

$$\log_\alpha\left(1+\frac{1}{10\,k+d}\right) < \log_\alpha\left(1+\frac{1}{k}\right)$$

and therefore

$$\sum \log_\alpha\left(1+\frac{1}{10\,k+d}\right) < \sum \log_\alpha\left(1+\frac{1}{k}\right)$$

from which we get

$$U_n(d) < 1.$$

As a consequence, we may state that $U_n(d)$ has a limit when $n \to \infty$ and $d \geq 9$.

*

From the expression of $U_n(d)$ we have that

$$0 \leq d \leq 10 \quad \Longrightarrow \quad U_n(0) \geq U_n(d) \geq U_n(10);$$

logo

$$\mathrm{U}_n(d) - \mathrm{U}_n(10) \leqslant \mathrm{U}_n(0) - \mathrm{U}_n(10) =$$

$$= \frac{\sum\limits_{n=\omega}^{n}\left[\log_\alpha\left(1+\frac{1}{10n}\right) - \log_\alpha\left(1+\frac{1}{10(n+1)}\right)\right]}{\log_\alpha(n+1) - \log_\alpha \omega} =$$

$$= \frac{\log_\alpha\left(1+\frac{1}{10\,\omega}\right) - \log_\alpha\left(1+\frac{1}{10(n+1)}\right)}{\log_\alpha(n+1) - \log_\alpha \omega} <$$

$$< \frac{\log_\alpha\left(1+\frac{1}{10\,\omega}\right)}{\log_\alpha(n+1)} < \delta$$

qualquer que seja $\delta > 0$, logo que n seja suficientemente grande. Logo, existe o

$$\lim_{n=\infty} \mathrm{U}_n(d)$$

qualquer que seja d e esse limite não depende de d; mas como

$$\lim_{n=\infty} \sum_{d=0}^{9} \mathrm{U}_n(d) = 1 = \sum_{d=0}^{9} \lim_{n=\infty} \mathrm{U}_n(d) = 10 \,.\, \lim_{n=\infty} \mathrm{U}_n(d)$$

segue-se que

$$\lim_{n=\infty} \mathrm{U}_n(d) = \frac{1}{10}\,.$$

Facilmente se vê que U_n converge muito rapidamente para $\frac{1}{10}$.

and therefore, for n large enough,

$$U_n(d) - U_n(10) \leq U_n(0) - U_n(10) =$$

$$= \frac{\sum_{k=\omega}^{n} \left[\log_\alpha \left(1 + \frac{1}{10\,k}\right) - \log_\alpha \left(1 + \frac{1}{10\,(k+1)}\right)\right]}{\log_\alpha (n+1) - \log_\alpha\, \omega} =$$

$$= \frac{\log_\alpha \left(1 + \frac{1}{10\,\omega}\right) - \log_\alpha \left(1 + \frac{1}{10\,(n+1)}\right)}{\log_\alpha (n+1) - \log_\alpha\, \omega} <$$

$$< \frac{\log_\alpha \left(1 + \frac{1}{10\,\omega}\right)}{\log_\alpha (n+1)} < \delta,$$

for any $\delta > 0$. Thus the limit

$$\lim_{n\to\infty} U_n(d)$$

exists for any d, and that limit doesn't depend on d; but as

$$\lim_{n\to\infty} \sum_{d=0}^{9} U_n(d) = 1 = \sum_{d=0}^{9} \lim_{n\to\infty} U_n(d)$$

it follows that

$$\lim_{n\to\infty} U_n(d) = \frac{1}{10}\,.$$

It is easily seen that U_n converges towards $\frac{1}{10}$ very quickly.

Nota

O problema geral que acabamos de resolver, dá-nos a distribuìção dos algarismos numa tábua ideal que contivesse todos os valores duma função num intervalo (α, β). Numa tábua qualquer, em que os valores da variável independente estejam em progressão arithmética, intervalos iguais, contidos em (α, β), compreendem, aproximadamente, o mesmo número de valores de x escritos na tábua, com êrro relativo tanto menor quanto menor fôr a razão da progressão dos valores de x. De modo que a probabilidade de que um valor de x, tirado à sorte em (α, β), pertença a um intervalo parcial, será, aproximadamente, proporcional à amplitude do intervalo, tal qual como acontecia na tábua ideal.

Esta tábua ideal será como que o limite duma sucessão de tábuas, em que a razão da progressão dos valores de x fosse decrescendo até zero. Portanto, a fórmula geral (1) dará tanto mais exactamente a distribuìção dos algarismos numa tábua de $f(x)$, quanto mais pequena fôr a razão da progressão dos valores de x. Assim, numa tábua de logarithmos decimais, visto que a mantissa se não altera com a divisão de x por uma potência de 10 (inteira), segue-se que a fórmula

$$\frac{P(d+1, a)}{P(d, a)} = 10^{\frac{1}{10^a}}$$

deve ser muito mais aproximada no fim da tábua do que no princípio. E assim é, realmente. Numa tábua pode vêr-se que desde 1289 a 1319, e de 1319 a 1349 se encontram respectivamente os algarismos 1 e 2 na segunda

Observation

The general problem that we have just solved gives the distribution of the digits in an ideal table containing all values of a regular function in an interval (α, β). In any real table with independent values in arithmetic progression, equal subsets of (α, β) contain approximately the same number of values of x written down in the table, with relative error decreasing with the step of the arithmetic progression of x values. From that, the probability that a value of x randomly chosen in (α, β) lies in a given subinterval is approximately proportional to the size of that subinterval, exactly as it happens in the ideal table.

This ideal table may be regarded as the limit of a sequence of real tables as described when the step of the arithmetic progression of the x's decreases towards 0. Thus, the smaller is the step of the arithmetic progression of the x's, the closer general formula (4.1) will be to the distribution of digits in a table of $f(x)$. Therefore, in a table of basis 10 logarithms, as the mantissa doesn't change when dividing x by an (integer) power of 10, we expect the formula

$$\frac{\mathbb{P}(d+1,a)}{\mathbb{P}(d,a)} = 10^{\frac{1}{10^{a}}}$$

to give much closer results at the end of the table than at its beginning.

This is in fact so. For instance, counting the number of digits 1 and 2 in the second decimal place in a table of basis 10 logarithms between 1289 and 1319, or between 1319 and 1349, we find casa decimal, o que dá

$$\frac{P_{(2,\,2)}}{P_{(1,\,2)}} = \frac{30}{30} = 1\,;$$

e desde 10232 a 10471, e dêste número a 10715 se encontram os mesmos algarismos na mesma casa decimal, o que dá

$$\frac{P_{(2,\,2)}}{P_{(1,\,2)}} = \frac{244}{239} = 1{,}0209\ \ldots$$

número êste muito mais próximo do número teórico

$$\sqrt[100]{10} = 1{,}0233$$

do que o primeiro.

3.º

É curioso determinar a relação

$$\frac{P_{d+1}}{P_d}$$

para as diferenças tabulares dos logarithmos. Estas diferenças podem considerar-se como valores da função

$$y = \log\,(1 + x) - \log\,(x)$$

$$= \log\left(1 + \frac{1}{x}\right)$$

30 of each of those digits, getting

$$\frac{\mathbb{P}(2,\,2)}{\mathbb{P}(1,\,2)} = \frac{30}{30} = 1;$$

on the other hand, the number of digits 1 and 2 in the second decimal place in a tables of basis 10 logarithms between 10232 and 10471, or between 10471 and 19715, are respectively 239 and 244, and thus

$$\frac{\mathbb{P}(2,\,2)}{\mathbb{P}(1,\,2)} = \frac{244}{239} \approx 1.0209,$$

much closer to the theoretical value

$$\sqrt[100]{10} \approx 1.0233$$

for the ideal table.

3rd

It has some independent interest to compute the ratio

$$\frac{\mathbb{P}(d+1)}{\mathbb{P}(d)}$$

for the tabular differences of logarithms. Those differences may be regarded as values of the function

$$y = \log\,(1+x) - \log\,x =$$
$$= \log\left(1 + \frac{1}{x}\right)$$

correspondentes a valores de x escritos em progressão arithmética. Ou, pondo

$$\log\left(1+\frac{1}{x}\right)=\frac{10\,\omega+d}{10^a},$$

$$\log\left(1+\frac{1}{x'}\right)=\frac{10\,\omega+d+1}{10^a},$$

$$\log\left(1+\frac{1}{x''}\right)=\frac{10\,\omega+d+2}{10^a},$$

será

$$\frac{P_{d+1}}{P_d}=\frac{x''-x'}{x'-x}=$$

$$=\frac{\dfrac{1}{10^{\frac{10\,\omega+d+2}{10^a}}-1}-\dfrac{1}{10^{\frac{10\,\omega+d+1}{10^a}}-1}}{\dfrac{1}{10^{\frac{10\,\omega+d+1}{10^a}}-1}-\dfrac{1}{10^{\frac{10\,\omega+d}{10^a}}-1}}$$

muito aproximadamente

$$=\frac{\dfrac{1}{\dfrac{10\,\omega+d+2}{10^a}\log 10}-\dfrac{1}{\dfrac{10\,\omega+d+1}{10^a}\log 10}}{\dfrac{1}{\dfrac{10\,\omega+d+1}{10^a}\log 10}-\dfrac{1}{\dfrac{10\,\omega+d}{10^a}\log 10}}$$

$$=\frac{(10\,\omega+d)\,(10\,\omega+d+1)}{(10\,\omega+d+1)\,(10\,\omega+d+2)}=\frac{10\,\omega+d}{10\,\omega+d+2}\,;$$

corresponding to x values in arithmetic progression. Rewriting

$$\log\left(1+\frac{1}{x}\right)=\frac{10\,\omega+d}{10^a}\,,$$

$$\log\left(1+\frac{1}{x'}\right)=\frac{10\,\omega+d+1}{10^a}\,,$$

$$\log\left(1+\frac{1}{x''}\right)=\frac{10\,\omega+d+2}{10^a}\,,$$

we get

$$\frac{\mathbb{P}(d+1)}{\mathbb{P}(d)}=\frac{x''-x'}{x'-x}=$$

$$=\frac{\dfrac{1}{10^{\frac{10\,\omega+d+2}{10^a}}-1}-\dfrac{1}{10^{\frac{10\,\omega+d+1}{10^a}}-1}}{\dfrac{1}{10^{\frac{10\,\omega+d+1}{10^a}}-1}-\dfrac{1}{10^{\frac{10\,\omega+d}{10^a}}-1}}$$

very approximately

$$\approx\frac{\dfrac{1}{\dfrac{10\,\omega+d+2}{10^a}\log 10}-\dfrac{1}{\dfrac{10\,\omega+d+1}{10^a}\log 10}}{\dfrac{1}{\dfrac{10\,\omega+d+1}{10^a}\log 10}-\dfrac{1}{\dfrac{10\,\omega+d}{10^a}\log 10}}=$$

$$=\frac{(10\,\omega+d)\,(10\,\omega+d+1)}{(10\,\omega+d+1)\,(10\,\omega+d+2)}=\frac{10\,\omega+d}{10\,\omega+d+2}\,;$$

logo

$$\frac{P_d}{P_{d+1}} = 1 + \frac{2}{10\,N + d} = 1 + \frac{2}{10^a\,D}, \tag{1}$$

sendo

$$D = \log\left(1 + \frac{1}{x}\right) = \frac{10\,N + d}{10^a}.$$

Visto que $10^a\,D$ é a parte inteira do produto duma qualquer diferença tabular, tendo o algarismo d na casa decimal de ordem a, por 10^a, segue-se que: dada uma diferença tabular D, se pode obter imediatamente por meio de (1) a relação

$$\frac{P_d}{P_{d+1}}$$

sendo d o número dígito que D tem na casa decimal de ordem a. Assim, dada a diferença tabular 0,0000524, deduz-se, para $a = 5$,

$$\frac{P_5}{P_6} = 1 + \frac{2}{5} = \frac{7}{5} = 1{,}400.$$

Procurando nas tábuas vê-se qne a diferença tabular 0,0000500 tem por número máximo, correspondente o número 8694 ; a diferença tabular 0,0000600, tem por correspondente o número 7243, a diferença 0,0000700 tem por número máximo correspondente, o número 6208 ; teremos, pois, segundo as tábuas,

$$\frac{P_5}{P_6} = \frac{1451}{1035} = 1{,}401.$$

therefore

$$\frac{\mathbb{P}(d)}{\mathbb{P}(d+1)} = 1 + \frac{2}{10\,N + d} = 1 + \frac{2}{10^a D}\,, \tag{4.2}$$

where

$$D = \log\left(1 + \frac{1}{x}\right) = \frac{10\,N + d}{10^a}\,.$$

As $10^a D$ is the integer part of the product of 10^a by any tabular difference, with the digit d in the $a-$th decimal place, we conclude that, given a tabular difference D, we can get from (4.2) the ratio

$$\frac{\mathbb{P}(d)}{\mathbb{P}(d+1)}\,,$$

where d is the digit in the $a-$th decimal place of D.

As an example, given the tabular difference 0.0000524, we get for $a = 5$,

$$\frac{\mathbb{P}(5)}{\mathbb{P}(6)} = 1 + \frac{2}{5} = 1.400.$$

Searching in the tables we find that the tabular difference 0.0000500 has a corresponding maximum at 8694; that the tabular difference 0.0000600 has a corresponding maximum at 7243; that the tabular difference 0.0000700 has a corresponding maximum at 6208; we therefore have, for those tables

$$\frac{\mathbb{P}(5)}{\mathbb{P}(6)} = \frac{1451}{1035} = 1.401.$$

Observação

Para as casas inteiras, isto é, para $\dot{a} = 0, -1, -2, \ldots,$ a fórmula

$$\frac{P_{d+1}}{P_d} = 10^{\frac{1}{10^a}}$$

é exacta, devido a que os números inteiros são logarithmos de valores de x escritos nas tábuas.

Proposição II

Lei da possibilidade

Seja (A) uma região contendo a região (A′) e sejam (B) e (B′) imagens respectivamente de (A) e (A′). Consideremos em (A) o ponto livre M e a sua visinhança ΔS e em (B) o ponto imagem N com a visinhança $\Delta S'$, imagem de ΔS. Representando por $\Delta \varpi$ a possibilidade de ΔS, será ainda $\Delta \varpi$ (prop. I) a possibilidade de $\Delta S'$. Consideremos o

$$\lim_{\Delta S' = 0} \frac{\Delta \varpi}{\Delta S'},$$

êste limite sendo tomado de modo que a maior dimensão de $\Delta S'$ tenda para zero conjuntamente com $\Delta S'$. *A totalidade de pontos* N *para os quais exista êste limite, forma o campo de existência duma função das coordenadas do ponto* N, *função cujo valor em cada ponto é dado por êsse limite e a que chamaremos lei da possibilidade.*

Observation

For the integer positions, i.e. for $a = 0, -1, -2, \ldots$, the formula

$$\frac{\mathbb{P}(d+1,a)}{\mathbb{P}(d,a)} = 10^{\frac{1}{10^a}}$$

is exact, since the integers are logarithms of values of x written in the tables.

Proposition II

Law of possibilities

Let $A' \subset A$, and let B and B' be the images of, respectively, A and A'. Let M denote a free point in A, ΔS a neighborhood of M, and let N be its image in B, its neighborhood $\Delta S'$ the image of ΔS. Denoting $\Delta\omega$ the possibility of ΔS, it will also be the possibility of $\Delta S'$. Let

$$\lim_{\mu(\Delta S')\to 0} \frac{\Delta\omega}{\mu\left(\Delta S'\right)},$$

where we assume that the limit is taken with the supremum of the projection of $\Delta S'$ on the coordinate axes goes to zero with $\Delta S'$. *The set of points N for which this limit exists is the domain of a function of the coordinates of N, whose value in each point is the above limit. That function is the law of possibility.*

Corolário

A possibilidade de (B') será dada por

$$\varpi_{(B')} = \int_{(B')} \left[\lim_{\Delta S' = 0} \frac{\Delta \varpi}{\Delta S'} \right] d S'.$$

Conhecida a lei da possibilidade poderemos, pois, determinar a possibilidade de qualquer região (B'), independentemente da consideração da região (A') de que a primeira é imagem.

Proposição III

Lei da probabilidade

De modo idêntico se define a lei da probabilidade como a função que tem em cada ponto N o valor dado pelo

$$\lim_{\Delta S' = 0} \frac{\Delta P}{\Delta S'} .$$

Proposição IV

Sendo (A) a região possível em relação à probabilidade ΔP e $\varpi(N)$ a lei da possibilidade de N, será

$$\lim_{\Delta S' = 0} \frac{\Delta P}{\Delta S'} = \lim_{\Delta S' = 0} \frac{\dfrac{\Delta \varpi}{\int_{(A)} \varpi(N)\, d\varpi}}{\Delta S'}$$

Corollary

The possibility of B' is

$$\omega_{B'} = \int\limits_{B'} \lim_{\mu(\Delta S')\to 0} \frac{\Delta\omega}{\mu\left(\Delta S'\right)}\ \mathrm{d}S'.$$

Thus, once the possibility law is known, we can compute the possibility of any region B' without any reference to the region A' of which B' is the image.

Proposition III

Law of probability

In analogy with the above definition, we define the probability law as the function whose value at each point N is given by

$$\lim_{\mu(\Delta S')\to 0} \frac{\Delta\mathbb{P}}{\mu\left(\Delta S'\right)}\,.$$

Proposition IV

If A is the possible region in what regards the probability $\Delta\mathbb{P}$ and $\omega(N)$ denotes the possibility of N, then

$$\lim_{\mu(\Delta S')\to 0} \frac{\Delta\mathbb{P}}{\mu\left(\Delta S'\right)} = \lim_{\mu(\Delta S')\to 0} \frac{\dfrac{\Delta\omega}{\int_A \omega(N)\ \mathrm{d}\omega}}{\mu\left(\Delta S'\right)}$$

ou

$$\lim_{\Delta S'=0} \frac{\dfrac{\Delta \varpi}{\Delta S'}}{\int_{(A)} \varpi(N)\, d\varpi} = \frac{\varpi(N)}{\int_{(A)} \varpi(N)\, d\varpi}$$

o que mostra que para uma dada região possível, a lei da probabilidade é proporcional à lei da possibilidade.

Leis à priori e leis à posteriori

No que vai seguir-se, suporemos que o ponto cuja posição depende do acaso, varía numa região plana, para facilitar a exposição. As demonstrações ficarão, aliás, com toda a generalidade.

Lei à priori

Seja $M(x, y)$ um ponto variando numa região plana *(fig. 12)* e $m(x)$ a sua projecção sôbre o eixo dos XX'.

Chamaremos *lei de probabilidade à priori do ponto* $M(x, y)$, *à lei da probabilidade da sua projecção* $m(x)$.

Proposição V

Sendo $\varphi(x, y)$ *a lei da probabilidade de* $M(x, y)$ *na região* (A), será

$$a(x) = \int \varphi(x, y)\, dy$$

a lei da probabilidade à priori do mesmo ponto.

or

$$\lim_{\mu(\Delta S')\to 0} \frac{\frac{\Delta\omega}{\mu(\Delta S')}}{\int_A \omega(N)\ \mathrm{d}\omega} = \frac{\omega(N)}{\int_A \omega(N)\ \mathrm{d}\omega}\,,$$

showing that in each possible region the probability law is proportional to the possibility law.

A priori **and** ***a posteriori*** **laws**

In what follows, we assume, without loss of generality, that the random point varies in a plane region, so that the arguments can be presented in an easy way.

A priori **law**

Let $M(x, y)$ denote a random point varying in a plane region (Fig. 12), and let $m(x)$ denote its projection in the OX axis.

The a priori law of the point $M(x, y)$ is the probability law of its projection $m(x)$.

Proposition V

If $\varphi(x, y)$ denotes the probability law of $M(x, y)$ in the region A, then its a priori probability law is

$$a(x) = \int \varphi(x, y)\ \mathrm{d}y.$$

Com efeito, consideremos um filete de espessura ΔS, paralelo ao eixo do $Y Y'$ e contendo os pontos de absissa x.

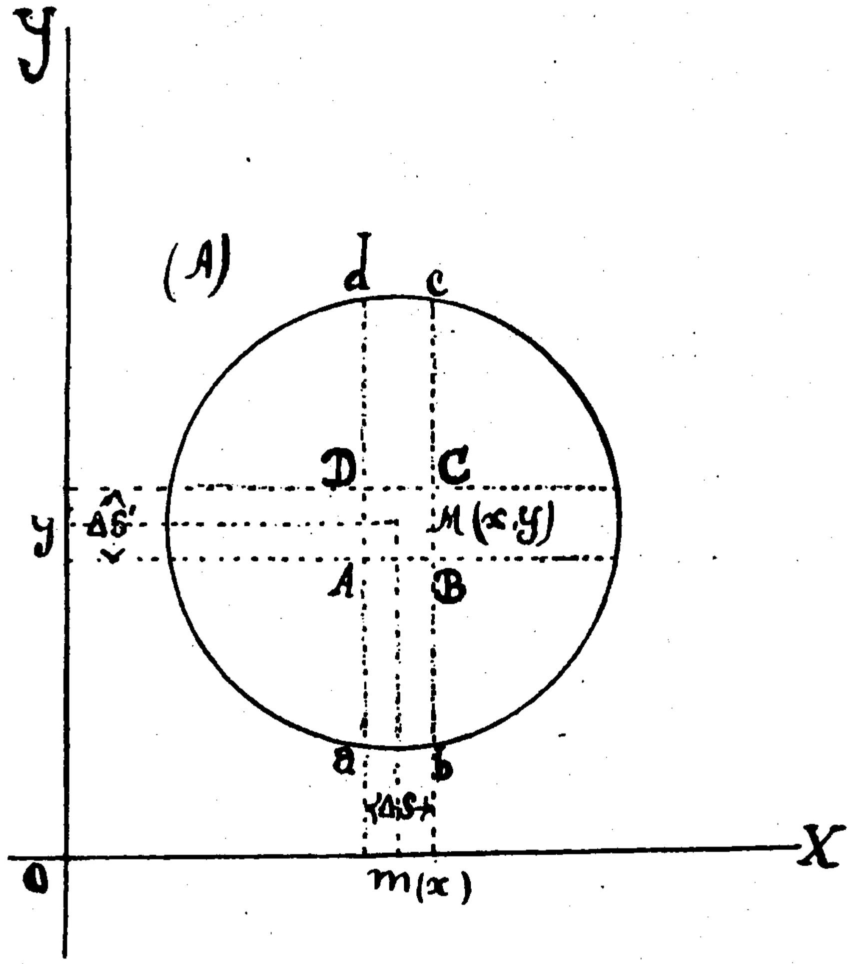

Figura 12

A probabilidade ΔP de que o ponto m caía na visinhança ΔS de x é dada pela probabilidade de que o ponto $M(x, y)$ cáia dentro da região $a\,b\,c\,d$; logo,

$$\Delta P = \iint_{a\,b\,c\,d} \varphi(x, y)\, dx\, dy = \Delta S \int_{y'_1}^{y'_2} \varphi(x_1, y)\, dy$$

sendo y'_1 e a y'_2 a menor e maior ordenada dos pontos de (A)

In effect, let us consider a vertical band with width ΔS, containing the points with abscissa x.

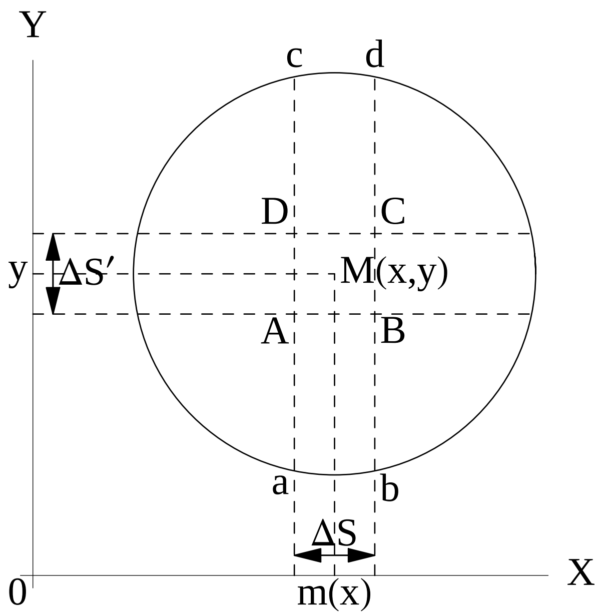

Figure 12

The probability $\Delta\mathbb{P}$ that the point m lies in a neighborhood ΔS of x is the probability that the point $M(x, y)$ lies in the region $[abcd]$; therefore

$$\Delta\mathbb{P} = \iint\limits_{[abcd]} \varphi(x, y)\ \mathrm{d}x\ \mathrm{d}y = \mu\,(\Delta S) \int_{y_1'}^{y_2'} \varphi(x_1, y)\ \mathrm{d}y,$$

where y_1' and y_2' stand for the smaller and the greater ordinates of the points in A

de absissa x_1. Ora, segundo a definição de lei *à priori*,

$$a(x) = \lim_{\Delta S = 0} \frac{\Delta P}{\Delta S} = \int_{y_1}^{y_2} \varphi(x, y)\, dy$$

c. d. d.

*

Por razões análogas teremos também

$$a(y) = \int \varphi(x, y)\, dx.$$

Com a letra a designaremos uma lei *à priori*; com a letra p as leis *à posteriori*.

Lei à posteriori

Consideremos na região (A) *(fig. 12)* dois filetes, um paralelo ao eixo dos XX' e contendo os pontos de ordenada y; o outro, paralelo ao eixo dos YY' e contendo os pontos da absissa x. Seja ΔS a espessura do segundo e $\Delta S'$ a do primeiro. A probabilidade da região ABCD em relação a *abcd* é (cap. II, prop. VII, observ.)

$$\Delta P = \frac{\displaystyle\iint_{ABCD} \varphi(x, y)\, dx\, dy}{\displaystyle\iint_{abcd} \varphi(x, y)\, dx\, dy} = \frac{\Delta S \,.\, \Delta S' \varphi(x', y')}{\Delta S \displaystyle\int_{y'_1}^{y'_2} \varphi(x'_1, y)\, dy} =$$

$$= \frac{\Delta S' \varphi(x', y')}{\displaystyle\int_{y'_1}^{y'_2} \varphi(x'_1, y)\, dy}$$

with abscissa x_1. Or, according to the definition of *a priori* law,

$$a(x) = \lim_{\mu(\Delta S)\to 0} \frac{\Delta \mathbb{P}}{\mu\,(\Delta S)} = \int_{y'_1}^{y'_2} \varphi(x,y)\ \mathrm{d}y,$$

$\square$

*

Similarly,

$$a(y) = \int \varphi(x,y)\ \mathrm{d}x.$$

We shall use a to denote an *a priori* law, and p to denote an *a posteriori* law.

***A posteriori* law**

Let us consider an horizontal and a vertical band in region A (Fig. 12), containing respectively the points with ordinate y and the points with abscissa x, and assume that the bandwidths are respectively $\Delta S'$ and ΔS. The probability of the region $[ABCD]$ in relation to the region $[abcd]$ (Chapter II, Prop. VII, Observation) is

$$\Delta \mathbb{P} = \frac{\displaystyle\iint_{[ABCD]} \varphi(x,y)\ \mathrm{d}x\ \mathrm{d}y}{\displaystyle\iint_{[abcd]} \varphi(x,y)\ \mathrm{d}x\ \mathrm{d}y} = \frac{\mu\,(\Delta S)\cdot\mu\,(\Delta S')\cdot\varphi(x',y')}{\mu\,(\Delta S)\displaystyle\int_{y'_1}^{y'_2} \varphi(x'_1,y)\ \mathrm{d}y} =$$

$$= \frac{\mu\,(\Delta S')\cdot\varphi(x',y')}{\displaystyle\int_{y'_1}^{y'_2} \varphi(x'_1,y)\ \mathrm{d}y}$$

sendo x' e x'_1 funções de ΔS que tendem para x quando ΔS tende para zero. Ao

$$\lim_{\Delta S = 0} \Delta P = \frac{\Delta S' \varphi(x, y')}{\int_{y'_1}^{y'_2} \varphi(x, y)\, dy}$$

chamaremos probabilidade *à posteriori* de $\Delta S'$.

Pode dizer-se que é a probabilidade de que y caia no intervalo $\Delta S'$, caso x tenha tomado o valor particular x.

Chamaremos lei *à posteriori* de y ao

$$\lim_{\Delta S' = 0} \frac{\Delta P}{\Delta S'} = \frac{\varphi(x, y)}{\int \varphi(x, y)\, dy} = p(y). \qquad (1)$$

Proposição VI

De (1) e da prop. V tira-se

$$\varphi(x, y) = a(x) \cdot p(y) = a(y) \cdot p(x).$$

*

As prop. V e VI são análogas às proposições relativas à probabilidade total e à probabilidade composta.

Destas duas proposições fácilmente se deduz uma fórmula análoga à

Fórmula de Bayes,

o que justifica as designações de lei *à priori* e lei *à posteriori* das leis atraz definidas.

where x' and x'_1 are functions of ΔS that converge to x when ΔS goes to zero. We shall say that

$$\lim_{\mu(\Delta S)\to 0} \Delta\mathbb{P} = \frac{\mu\,(\Delta S')\cdot\varphi(x,y')}{\displaystyle\int_{y'_1}^{y'_2} \varphi(x'_1,y)\ \mathrm{d}y}$$

is the *a posteriori* probability of $\Delta S'$.

We may interpret that a posteriori probability as the probability that y lies in the interval $\Delta S'$, given that x has taken the particular value x.

$$p(y) = \lim_{\mu(\Delta S')\to 0} \frac{\Delta\mathbb{P}}{\mu\,(\Delta S')} = \frac{\varphi(x,y)}{\int \varphi(x,y)\ \mathrm{d}y} \tag{4.3}$$

is the *a posteriori* law of y.

Proposition VI

From Prop. V and (4.3) we get

$$\varphi(x,y) = a(x)\cdot p(y) = a(y)\cdot p(x).$$

*

Prop. V and VI are similar to the propositions about total probability and compound probability.

From those two propositions it is easy to infer a formula similar to

Bayes' formula

and this justifies the terminology *a priori* and *a posteriori* laws that we have been using.

Com efeito, visto que (prop. VI)

$$\varphi(x, y) = a(x)p(y) = a(y)p(x)$$

e

$$a(y) = \int \varphi(x, y)\,dx = \int a(x)p(y)\,dx,$$

segue-se que

$$p(x) = \frac{\varphi(x, y)}{a(y)} = \frac{a(x)p(y)}{\int a(x)p(y)\,dx}$$

e, análogamente,

$$p(y) = \frac{a(y)p(x)}{\int a(y)p(x)\,dy},$$

fórmulas estas que nos dão as leis à *posteriori* duma das variáveis, logo que se conheça a sua lei à *priori*, a lei à *posteriori* da outra e o seu campo de variação.

*

Desta fórmula podem tirar-se outras a que pode chamar-se

Fórmulas inversas da de Bayes

Com efeito, de

$$p(x) = \frac{a(x)p(y)}{\int a(x)p(y)\,dx},$$

tira-se, tomando as derivadas parciais em ordem a x e atendendo a que

$$\int a(x)p(y)\,dx$$

In effect, from Prop. VI

$$\varphi(x,y) = a(x)\cdot p(y) = a(y)\cdot p(x).$$

and

$$a(y) = \int \varphi(x,y)\ \mathrm{d}x = \int a(x)\,p(y)\ \mathrm{d}x$$

it follows that

$$p(x) = \frac{\varphi(x,y)}{a(y)} = \frac{a(x)\,p(y)}{\int a(x)\,p(y)\ \mathrm{d}x}\,,$$

and, similarly,

$$p(y) = \frac{a(y)\,p(x)}{\int a(y)\,p(x)\ \mathrm{d}y}\,.$$

Those are the formulas for the *a posteriori* law for each of the variables. Hence the *a posteriori* law of one variable can be computed from its *a priori* law, if the *a posteriori* law of the other variable and the corresponding domain of variation are known.

*

From the above formulas we can infer other formulas, which we shall call

Inverses to Bayes' formula

From

$$p(x) = \frac{a(x)\,p(y)}{\int a(x)\,p(y)\ \mathrm{d}x}\,,$$

taking partial derivative in order to x and remembering that

$$\int a(x)\,p(y)\ \mathrm{d}x$$

não dependa de x,

$$\frac{\frac{\partial p(x)}{\partial x}}{p(x)} = \frac{a'(x)}{a(x)} + \frac{\frac{\partial p(y)}{\partial x}}{p(y)} ;$$

donde

$$\frac{a'(x)}{a(x)} = \frac{\frac{\partial p(x)}{\partial x}}{p(x)} - \frac{\frac{\partial p(y)}{\partial x}}{p(y)} ;$$

donde

$$a(x) = k(y) \frac{p(x)}{p(y)} ;$$

onde $k(y)$ é uma função arbitrária de y que se determina pela condição de

$$\int a(x)\, dx = k(y) \int \frac{p(x)}{p(y)}\, dx = 1 ;$$

donde,

$$k(y) = \frac{1}{\int \frac{p(x)}{p(y)}\, dx} ,$$

e

$$a(x) = \frac{\frac{p(x)}{p(y)}}{\int \frac{p(x)}{p(y)}\, dy}$$

e, análogamente,

$$a(y) = \frac{\frac{p(y)}{p(x)}}{\int \frac{p(y)}{p(x)}\, dy}$$

doesn't depend on x, we get

$$\frac{\dfrac{\partial p(x)}{\partial x}}{p(x)} = \frac{a'(x)}{a(x)} + \frac{\dfrac{\partial p(y)}{\partial x}}{p(y)}\,;$$

from this,

$$\frac{a'(x)}{a(x)} = \frac{\dfrac{\partial p(x)}{\partial x}}{p(x)} - \frac{\dfrac{\partial p(y)}{\partial x}}{p(y)}\,;$$

thus

$$a(x) = k(y)\,\frac{p(x)}{p(y)}\,,$$

where $k(y)$ is an arbitrary function of y which we determine using the condition

$$\int a(x)\ \mathrm{d}x = k(y)\int \frac{p(x)}{p(y)}\ \mathrm{d}x = 1\,;$$

therefore

$$k(y) = \frac{1}{\displaystyle\int \frac{p(x)}{p(y)}\ \mathrm{d}x}$$

and

$$a(x) = \frac{\dfrac{p(x)}{p(y)}}{\displaystyle\int \frac{p(x)}{p(y)}\ \mathrm{d}x}\,.$$

Similarly,

$$a(y) = \frac{\dfrac{p(y)}{p(x)}}{\displaystyle\int \frac{p(y)}{p(x)}\ \mathrm{d}y}\,.$$

fórmulas estas análogas às já achadas na probabilidade discontínua.

*

As definições e demonstrações que acabamos de fazer são duma generalização imediata. O que se disse para um ponto variando numa região plana estende-se imediatamente a um ponto variando numa região qualquer. Do que acabamos de dizer, devemos excetuar a dedução das fórmulas inversas da fórmula de BAYES, dedução essa que não é suscetivel de ser estendida imediatamente ao caso de muitas variáveis. A dedução pode fazer-se, dum modo geral, muito fácilmente.

Com efeito, ainda que x e y representem complexos de variáveis, teremos sempre

$$a(x) \,.\, p(y) = a(y)\, p(x) \tag{1}$$

sendo as funções $a(x)$ e $a(y)$ funções só dos complexos (x) e (y) respectivamente e $p(y)$ e $p(x)$ sendo funções dos dois complexos simultaneamente. De (1) tira-se

$$a(x) = a(y) \,.\, \frac{p(x)}{p(y)}$$

donde,

$$\int_{(x)} a(x)\, d(x) = a(y) \int_{(x)} \frac{p(x)}{p(y)}\, d(x) = 1$$

These formulas are analogous to those deduced in Chapter I.

*

The definitions and demonstrations we have presented have immediate generalization for higher dimensions. What we have established about a point varying in a plane region can be extended immediately for a point in any region, the only exception being these last formulas inverse to Bayes' formula, since the arguments are not usable in higher dimensions. But the extension to higher dimensions can also be made very easily.

In effect, if $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors, we still have

$$a(\boldsymbol{x}) \cdot p(\boldsymbol{y}) = a(\boldsymbol{y}) \cdot p(\boldsymbol{x}). \tag{4.4}$$

where $a(\boldsymbol{x})$ and $a(\boldsymbol{y})$ are functions of only $\boldsymbol{x}$ and of $\boldsymbol{y}$, respectively, and $p(\boldsymbol{y})$ and $p(\boldsymbol{x})$ are functions of, simultaneously, $\boldsymbol{x}$ and $\boldsymbol{y}$. From (4.4) we get

$$a(\boldsymbol{x}) = a(\boldsymbol{y}) \cdot \frac{p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

and from this it follows that

$$\int_{(\boldsymbol{x})} a(\boldsymbol{x})\ \mathrm{d}\boldsymbol{x} = a(\boldsymbol{y}) \int_{(\boldsymbol{x})} \frac{p(\boldsymbol{x})}{p(\boldsymbol{y})}\ \mathrm{d}\boldsymbol{x} = 1.$$

e portanto

$$a(x) = \frac{\dfrac{p(x)}{p(y)}}{\displaystyle\int_{(x)} \frac{p(x)}{p(y)}\, d(x)},$$

em todos os casos.

Therefore

$$a(\boldsymbol{x}) = \frac{\dfrac{p(\boldsymbol{x})}{p(\boldsymbol{y})}}{\displaystyle\int_{(\boldsymbol{x})} \frac{p(\boldsymbol{x})}{p(\boldsymbol{y})}\,\mathrm{d}\boldsymbol{x}}$$

in all cases.

CAPÍTULO V

TEOREMAS DE JACOB BERNOULLI

E

LEI DOS DESVIOS

CHAPTER V

JACOB BERNOULLI'S THEOREMS

AND THE

ERROR LAW

CAPITULO V

PRIMEIRA PARTE

Teoremas de Jacob Bernoulli

A uma tiragem, à sorte, numa classe finita de elementos ou a um lançamento feito, à sorte, numa região, chamaremos: um *caso,* um *acontecimento* ou um *fenómeno,* conforme está estabelecido pelo uso. Um fenómeno, caso ou acontecimento, dir-se-há *favorável* ou *contrário,* conforme o elemento obtido na tiragem ou lançamento pertencer à classe favorável ou não.

Posto isto, seja p a probabilidade do acontecimento favorável e q a do acontecimento contrário. Será, como é evidente,

$$p+q=1.$$

Se obrigarmos o acontecimento a produzir-se uma vez, dois casos poderão dar-se

$$p \quad \text{ou} \quad q$$

(representando os acontecimentos pelas suas probabilida-

FIRST PART

Jacob Bernoulli's theorems

As usual, we shall say that the *result* of the random selection of an element from a finite set, or of randomly throwing one point in a bounded region, is a *case*, or an *event*. A result, case or event is said to be *favorable* [or a *success*] if it is an element of the favorable set, and *contrary* [or a *failure*] if it is an element of the contrary set.

Let us denote p the probability of success, and q the probability of failure. Obviously,

$$p + q = 1.$$

Performing once a random experiment [trial] as described, two cases can occur, with probabilities

$$p \qquad \text{or} \qquad q$$

des). Se obrigarmos o fenómeno a produzir-se duas vezes, quatro casos podem dar-se:

$$pp \quad \text{ou} \quad pq \quad \text{ou} \quad qp \quad \text{ou} \quad qq.$$

Se repetirmos o fenómeno três vezes, os casos possíveis seriam dados pelos arranjos com repetição dos dois objectos p e q, três a três, etc. Isto, supondo que a produção dum fenómeno não alterava a probabilidade do seguinte.

Daqui se conclue que:

Proposição I

A probabilidade de que, produzindo m *vezes sucessivas o fenómeno cujas modalidades favorável e contrária são* p e q, $m - i$ *acontecimentos da probabilidade* p *e* i *de probabilidade* q *se produzam, segundo uma ordem préviamente fixada, é*

$$P = q^{m-i} q^{i},$$

como resulta imediatamente das proposições relativas á probabilidade composta.

Corolário

A probabilidade de que os fenómenos se sucedam segundo uma ordem préviamente fixada, tende para zero quando o número de experiências aumenta.

Com efeito, supondo, para fixar ideias, que $p < q$, será

$$P = p^{m-i} q^{i} < p^{m};$$

Performing this random experiment twice, there are four possible outcomes, the corresponding probabilities being

$$pp \qquad \text{or} \qquad pq \qquad \text{or} \qquad qp \qquad \text{or} \qquad qq$$

Three trials would produce eight possible outcomes, with probabilities given by arrangements with repetitions of the two elementary probabilities p and q, three by three, etc. In the above reasoning, we are assuming that any outcome doesn't change the probability of the outcomes in the following experiment.

We may therefore conclude the following:

Proposition I

Performing m times an experiment whose possible results are success, with probability p, or failure, with probability q, the probability of getting $m-i$ successes and i failures, in a given order, is

$$\mathbb{P} = p^{m-i}q^i;$$

this is a direct consequence of the propositions concerning compound probability.

Corollary

The probability of any given sequence of outcomes in pre-arranged order decreases to zero, when the number of trials increase.

As a matter of fact, and assuming, without loss of generality, that $q \leq p$, from

$$\mathbb{P} = p^{m-i}q^i \leq p^m$$

e como $p < 1$, segue-se que p^m tende para zero quando m aumente indefinidamente.

*

Se a probabilidade duma sucessão de m acontecimentos não depende da ordem por que se sucedem os seus elementos componentes, depende contudo do número de casos que nela entram duma ou outra categoria (favorável ou contrária). De que modo essa variação se faz, é o que vamos tratar neste capítulo.

Para tornar mais claro o enunciado dessa questão suporemos que os fenómenos considerados, são tiragens de bolas brancas e pretas, feitas à sorte numa urna que lhes dê as probabilidades respectivas p e q. É evidente que esta identificação só se poderá fazer quando p e q forem racionais; mas isso não se opõe a que nos sirvamos dela, como duma linguagem figurada, nos outros casos.

Proposição II

Fazendo numa urna que contêm bolas brancas e pretas, m *tiragens, a probabilidade de que se obtenham* n *bolas brancas e* m — n *pretas é, sendo* p *e* q *as probabilidades respectivas,*

$$P_{m,\,n} = \frac{m\,!}{(m-n)\,!\;n\,!}\,p^n\,q^{m-n}.$$

Com efeito, a probabilidade de qualquer sucessão de n

with $p < 1$, the result follows, and we see that p^m goes to zero when m increases to infinity.

*

When performing m trials as described, the probability of getting as outcome $m - i$ successes and i failures in a pre-arrange order is always the same, $\mathbb{P} = p^{m-i}q^i$; therefore, the probability of getting $m-i$ successes and i failures, whichever their order, depends on the number of possible sequences of m trials whose outcome consists exactly of $m-i$ successes and i failures. Our immediate goal will be to establish the appropriate formula.

For the sake of clarity, we solve the question in the context of the random extraction of white and black balls from an urn whose composition is such that the probability of extracting white ball is p and the probability of extracting black ball is q. Clearly this identification of the two problems is legitimate only when p and q are rational numbers; but, as a metaphor, we shall use this language in all cases.

Proposition II

In one urn there are white and black balls, the probability of getting white ball in a random extraction being p; performing m extractions, [with replacement of the extracted ball in the urn after each of them,] the probability of getting white ball in n of those extractions, and black ball in the remaining $m - n$ extractions, is

$$\mathbb{P}_{m,n} = \frac{m!}{(m-n)!\,n!}\,p^n\,q^{m-n}.$$

In effect, the probability of any of the sequences of n

bolas brancas e $m-n$ pretas, é (prop. I)

$$p^n q^{m-n}.$$

Ora, o número de sucessões *distintas* que se podem obter com n bolas brancas e $m-n$ pretas é dado pelo número de permutações de m objectos dos quais $m-n$ e n são repetidos, isto é, por

$$\frac{m!}{(m-n)!\,n!}.$$

Logo

$$P_{m,n} = \frac{m!}{(m-n)!\,n!} p^n q^{m-n},$$

c. d. d.

Corolário

A probabilidade de tirar m — n *bolas brancas e* n *pretas é dada pelo termo correspondente do desenvolvimento de*

$$(p+q)^m.$$

Proposição III

Supondo m fixo, tem-se que:

1.º *A probabilidade* $P_{m,n}$ *cresce com* n *desde zero até ao maior número inteiro contido em*

$$p(m+1);$$

white and $m-n$ black balls is (Prop. I)

$$p^n\, q^{m-n}.$$

On the other hand, the number of possible different sequences composed of n white and $m-n$ black balls is the number of combinations of m objects of two types, having n of the first and $m-n$ if the second type, i.e.,

$$\frac{m!}{(m-n)!\,n!}\,.$$

Therefore,

$$\mathbb{P}_{m,n} = \frac{m!}{(m-n)!\,n!}\,p^n\, q^{m-n}.$$

□

Corollary

The probability of getting n white and $m-n$ black balls is given by the corresponding term in the expansion of

$$(p+q)^m.$$

Proposition III

Assuming m fixed, we have:

1. *The probability $\mathbb{P}_{m,n}$ increases with n while n goes from zero until the biggest integer not greater than*

$$p\,(m+1).$$

2.º *Diminue desde o menor inteiro* n *que contêm*

$$p(m+1)-1$$

até $n = m$;

3.º *Torna-se máxima para o inteiro compreendido entre*

$$p(m+1)-1 \quad \text{e} \quad p(m+1).$$

Com efeito:

1.º

Se

$$n \leqq p(m+1), \tag{1}$$

será

$$\frac{1}{p} \leqq \frac{m+1}{n}$$

e

$$\frac{q}{p} \leqq \frac{m+1}{n} - 1 = \frac{m-n+1}{n}$$

e

$$1 \leqq \frac{m-n+1}{n} \cdot \frac{p}{q};$$

mas

$$\frac{m-n+1}{n} \frac{p}{q} = \frac{P_n}{P_{n-1}};$$

logo

$$P_{n-1} \leqq P_n,$$

c. d. d.

2.º

Se

$$n \geqq p(m+1)-1 \tag{2}$$

2 *The probability* $\mathbb{P}_{m,n}$ *decreases from the least integer greater than*

$$p\,(m+1)-1$$

until $n=m$.

3 *The probability* $\mathbb{P}_{m,n}$ *assumes its maximum in the integers in the interval* $[p\,(m+1)-1, p\,(m+1)]$.

In effect,

1st

If

$$n \leq p(m+1) \tag{5.1}$$

we have

$$\frac{1}{p} \leq \frac{m+1}{n},$$

$$\frac{q}{p} \leq \frac{m+1}{n} - 1 = \frac{m-n+1}{n}$$

and

$$1 \leq \frac{m-n+1}{n} \cdot \frac{p}{q}.$$

But

$$\frac{m-n+1}{n} \cdot \frac{p}{q} = \frac{\mathbb{P}_{m,n}}{\mathbb{P}_{m,n-1}},$$

and therefore

$$\mathbb{P}_{m,n-1} \leq \mathbb{P}_{m,n}.$$

□

2nd

If

$$n \geq p(m{+}1){-}1 \tag{5.2}$$

será

$$\frac{1}{p} \geqslant \frac{m+1}{1+n}$$

e

$$1+\frac{q}{p} \geqslant \frac{m+1}{1+n}$$

e

$$\frac{q}{p} \geqslant \frac{m-n}{n+1}$$

e

$$1 \geqslant \frac{m-n}{n+1} \cdot \frac{p}{q};$$

mas

$$\frac{m-n}{n+1} \frac{p}{q} = \frac{P_{n+1}}{P_n};$$

logo

$$P_n \geqslant P_{n+1},$$

c. d. d.

3.°

Da primeira e segunda parte desta proposição se deduz que P_n será máximo quando n satisfizer a (1) e (2), isto é, quando fôr o inteiro da forma

$$p(m+1)-r, \qquad (0 \leqslant r \leqslant 1).$$

Caso $p(m+1)$ seja um número inteiro, haverá ambiguidade na determinação de n. A indeterminação levanta-se atendendo a que, neste caso, haverá o sinal de igualdade na primeira e segunda parte da proposição antecewe have

$$\frac{1}{p} \geq \frac{m+1}{1+n},$$

$$1 + \frac{q}{p} \geq \frac{m+1}{1+n},$$

$$\frac{q}{p} \geq \frac{m-n}{n+1},$$

and

$$1 \geq \frac{m-n}{n+1} \cdot \frac{p}{q}.$$

But

$$\frac{m-n}{n+1} \cdot \frac{p}{q} = \frac{\mathbb{P}_{m,n+1}}{\mathbb{P}_{m,n}},$$

and therefore

$$\mathbb{P}_{m,n} \geq \mathbb{P}_{m,n+1}.$$

□

3rd

From the first and the second part of the present proposition, we can deduce that $\mathbb{P}_{m,n}$ will assume its maximum value when n verifies both (5.1) and (5.2), i.e., when it is an integer of the form

$$p(m+1) - r, \qquad (0 \leq r \leq 1).$$

When $p(m+1)$ is non integer, there is a unique value of n in $[p\,(m+1) - 1, p\,(m+1)]$ for which $\mathbb{P}_{m,n}$ assumes its maximum value. When $p(m+1)$ is an integer, dente e por isso qualquer dos termos de ordem

$$p(m+1) \quad \text{ou} \quad p(m+1)-1$$

será máximo, por serem iguais.

Proposição IV

Consideremos todos os arranjos com repetição que se podem obter com dois objectos, tomados m a m.

Sejam os objectos uma bola branca e outra preta, por exemplo. Á totalidade de arranjos com o mesmo número de bolas brancas e pretas, chamaremos uma *combinação*. Posto isto, teremos que:

A probabilidade da combinação mais provável tende para zero quando o número das tiragens aumenta indefinidamente.

Com efeito, segundo a prop. III, a probabilidade da combinação mais provável será dada por

$$\mathrm{P} = \frac{m!}{[p(m+1)-r]!\,[q(m+1)-1+r]!}\, p^{p(m+1)-r}\, q^{(m+1)-1+r}$$

visto que, representando por $(m+1)p-r$ o número mais provável de bolas brancas, será

$$m-(m+1)p+r = (m+1)q+r-1$$

o número correspondente de bolas pretas.

$\mathbb{P}_{m,n}$ assumes its maximum for any of the terms of order

$$p(m+1) \quad \text{or} \quad p(m+1)-1.$$

□

Proposition IV

Let us consider all the sequences that may be obtained by repeatedly extracting two objects, with replacement after each extraction.

For clarity, let the two objects be white ball and black ball [extracted from one urn such that the probability of extracting white ball is p and that of extracting black ball is $q = 1-p$]. We shall call a combination of outcomes the totality of sequences with the same number k of white and $m-k$ of black balls. With these assumptions, we have:

The probability of the most probable combination decreases to zero when the number of trials increases to infinity.

In effect, using Prop. III, the probability of the most probable combination is

$$\mathbb{P}_{m,p(m+1)-r} = \frac{m!}{[p(m+1)-r]!\,[q(m+1)-1+r]!}\, p^{p(m+1)-r} q^{q(m+1)-1+r},$$

since when the number of white balls is $(m+1)p-r$ (the most probable outcome in m extractions), the corresponding number of black balls is

$$m-(m+1)p+r = (m+1)q+r-1.$$

Ora,

$$m! = m^m e^{-m} \sqrt{2\pi m}\,(1 + \varepsilon_m),$$

sendo ε_m uma função que tende para zero quando m tende para infinito. Podemos, pois, escrever

$$P = \frac{m^m e^{-m} \sqrt{2\pi m}\,(1 + \varepsilon_m)}{(mp)^{mp} e^{-mp} \sqrt{2\pi mp}\,(1+\varepsilon_{mp})\,(mq)^{mq} e^{-mq} \sqrt{2\pi mq}\,(1+\varepsilon_{mq})} p^{mp} q^{mq}$$

expressão esta que se obtêm da antecedente de P, substituindo os factoriais pelas expressões equivalentes e desprezando r e $1 - r$, o que é ligítimo quando m crescer alêm de todo o limite. Efectuando as reduções em P, resulta, imediatamente

$$P = \frac{1 + \alpha_m}{\sqrt{2\pi m p q}}$$

sendo α_m uma função que tende para zero quando m tende para infinito. P tenderá, pois, para zero quando m aumente indefinidamente,

c. d. d.

Proposição V

(1.º Teorema de Jacob Bernoulli)

Seja p a probabilidade dum acontecimento favorável e q a do acontecimento contrário. Faça-se um determinado número de experiências e sejam (p) e (q) os números respectivos de acontecimentos, favoráveis e contrários, resul-

However,
$$m! = m^m \ \mathrm{e}^{-m} \sqrt{2\,\pi\,m}\,(1+\varepsilon_m),$$
where ε_m is a function going to zero when m goes to infinity. Therefore we may write

$$\mathbb{P}_{m,p(m+1)-r} =$$

$$= \frac{m^m \ \mathrm{e}^{-m} \sqrt{2\,\pi\,m}\,(1+\varepsilon_m)}{(mp)^{mp} \ \mathrm{e}^{-mp} \sqrt{2\,\pi\,m\,p}\,(1+\varepsilon_{mp})\,(mq)^{mq} \ \mathrm{e}^{-mq} \sqrt{2\,\pi\,m\,q}\,(1+\varepsilon_{mq})}\, p^{mp} q^{mq}\,,$$

an expression obtained using Stirling's approximation for the factorials in the previous expression, and cutting out the vanishingly small terms in $p-r$ and $q+r-1$ (a legitimate approximation when m becomes as large as we want). This may be rewritten
$$\mathbb{P}_{m,p(m+1)-r} = \frac{1+\alpha_m}{\sqrt{2\,\pi\,m\,p\,q}}$$
where α_m denotes a function going to zero when m goes to infinity. Therefore, $\mathbb{P}_{m,p(m+1)-r}$ goes to zero when m goes to infinity, as stated. $\square$

Proposition V

(Jacob Bernoulli's 1st Theorem)

Let p denote the probability of the favorable event or success, and q the probability of the contrary event or failure. Performing a certain number of trials, let us denote by (p) and by (q) the number of successes and the number of failures that occur [in the $m = (p) + (q)$ trials], respectively.

tantes. *A razão* $\frac{(p)}{(q)}$ *poderá tomar valores vários; mas o mais provável dêsses valores é aquele que menos se afasta de* $\frac{p}{q}$ *e quanto mais* $\frac{(p)}{(q)}$ *se afasta dêste número, menos provável se torna.*

A probabilidade de $\frac{(p)}{(q)}$ a que se refere a proposição que acabamos de enunciar, é a da combinação de (p) acontecimentos favoráveis e de (q) contrários.

Posto isto, entremos na demonstração da proposição enunciada.

Como vimos já (prop. III, 3.º ponto) o número mais provável de casos favoráveis é dado pelo maior inteiro da forma

$$p(m+1)-r, \qquad (0 \leqq r \leqq 1);$$

o número correspondente de acontecimentos contrários, será dado por

$$q(m+1)-1+r;$$

a razão dêstes dois números é

$$\frac{(p)}{(q)} = \frac{p(m+1)-r}{q(m+1)-1+r} = \frac{p}{q} + \frac{(1-r)p-qr}{(m+1)q-1+r} \cdot \frac{1}{q};$$

Supondo que o número de casos favoráveis cresce, a razão de probabilidade imediatamente inferior (prop. III, *The ratio* $\dfrac{(p)}{(q)}$ *can take different values; but the most probable among them is the one nearer to* $\dfrac{p}{q}$*; the more* $\dfrac{(p)}{(q)}$ *differs from* $\dfrac{p}{q}$*, the less probable it is.*

The probability of $\dfrac{(p)}{(q)}$ in the above proposition is the combination of (p) successes and (q) failures in any order.

The proposition may be established as follows:

As seen in Prop. III (3rd part), the most probable number of successes in m trials is the greatest integer that can be written in the form

$$p\,(m+1)-r, \qquad (0 \leq r \leq 1);$$

the corresponding number of failures is

$$q\,(m+1)-1+r;$$

the ratio of those numbers is

$$\frac{(p)}{(q)} = \frac{p\,(m+1)-r}{q\,(m+1)-1+r} = \frac{p}{q} + \frac{(1-r)\,p-qr}{(m+1)\,q-1+r} \cdot \frac{1}{q}\,.$$

Assuming that the number of successes increases, the ratio with immediately lower probability

2.ª parte) é

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + 1}{q\,(m+1) - 1 + r - 1} = \frac{p}{q} + \frac{p\,(1-r) - qr + 1}{q\,(m+1) - 1 + r - 1};$$

a imediata a esta, é

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + 2}{q\,(m+1) - 1 + r - 2} = \frac{p}{q} + \frac{q\,(1-r) - qr + 2}{q\,(m+1) - 1 + r - 2};$$

sendo o número mais provável de acontecimentos favoráveis excedido de α unidades, a razão em questão será,

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + \alpha}{q\,(m+1) - 1 + r - \alpha} = \frac{p}{q} + \frac{p\,(1-r) - qr + \alpha}{q\,(m+1) - 1 + r - \alpha}.$$

Como estas igualdades mostram, a diferença

$$\frac{(p)}{(q)} - \frac{p}{q}$$

vai crescendo com α; logo (prop. III, 2.ª parte), a sua probabilidade vai diminuindo,

c. d. d.

*

Se o número de acontecimentos favoráveis diminuisse, consideravamos a razão $\frac{(q)}{(p)}$ e ficavamos reduzidos ao caso antecedente.

(Prop. III, 2nd part) is

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + 1}{q\,(m+1) - 1 + r - 1} = \frac{p}{q} + \frac{(1-r)\,p - qr + 1}{(m+1)\,q - 1 + r - 1} \cdot \frac{1}{q};$$

and the immediate one is

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + 2}{q\,(m+1) - 1 + r - 2} = \frac{p}{q} + \frac{(1-r)\,p - qr + 2}{(m+1)\,q - 1 + r - 2} \cdot \frac{1}{q};$$

when the most probable number of successes is exceeded by α units, the ratio in question will be

$$\frac{(p)}{(q)} = \frac{p\,(m+1) - r + \alpha}{q\,(m+1) - 1 + r - \alpha} = \frac{p}{q} + \frac{(1-r)\,p - qr + \alpha}{(m+1)\,q - 1 + r - \alpha} \cdot \frac{1}{q}.$$

As those expressions show, the difference

$$\frac{(p)}{(q)} - \frac{p}{q}$$

increases with α; therefore (Prop. III, 2nd part), its probability decreases.

*

If the number of successes would decrease, we would work out similarly with the ratio $\frac{(q)}{(p)}$, and so this case could be dealt with as the above one.

□

Proposição VI

(2.º Teorema de Bernoulli)

Á medida que se multiplicam as experiências, a probabilidade de cada valor de $\frac{(p)}{(q)}$ *vai diminuindo, mas tanto mais rapidamente quanto maior fôr*

$$\left| \frac{p}{q} - \frac{(p)}{(q)} \right|.$$

Chamando afastamento ao número α que entra na proposição antecedente, a probabilidade de $\frac{(p)}{(q)}$ será máxima quando o afastamento fôr nulo (prop. ant.) e como neste caso ela tende para zero quando m aumenta indefinidamente (prop. IV), por maioria de razão ela tenderá também para zero nos outros casos. Além disso, tem-se

$$\frac{P_{\alpha-1}}{P_\alpha} = \frac{q(m+1)+r+\alpha}{p(m+1)-r-\alpha+1} \cdot \frac{p}{q} = \frac{A+p\alpha}{B-q\alpha},$$

relação esta que mostra que, quanto maior fôr α, maior é a razão

$$\frac{P_{\alpha-1}}{P_\alpha};$$

ou, o que é o mesmo, quanto maior fôr

$$\left| \frac{p}{q} - \frac{(p)}{(q)} \right|,$$

Proposition VI

(Jacob Bernoulli's 2nd Theorem)

As the number of trials increases, the probability of each ratio $\frac{(p)}{(q)}$ *decreases, and the greater is the absolute value of the difference*

$$\left|\frac{p}{q} - \frac{(p)}{(q)}\right|$$

the greater will be the rate of decrease.

The number α used in the proof of the previous proposition will be called deviation; the probability of $\frac{(p)}{(q)}$ attains its maximum when that deviation is zero (Prop. V), and since in that case it decreases to zero when m goes to infinity (Prop. IV), it will also decrease to zero in all the other less probable cases. On the other hand,

$$\frac{\mathbb{P}_{\alpha-1}}{\mathbb{P}_{\alpha}} = \frac{q\,(m+1)+r-\alpha}{p\,(m+1)-r-\alpha+1} \cdot \frac{p}{q} = \frac{A+p\,\alpha}{B-q\,\alpha}\,,$$

an expression that shows that the ratio

$$\frac{\mathbb{P}_{\alpha-1}}{\mathbb{P}_{\alpha}}$$

decreases with α. In other words, $\frac{\mathbb{P}_{\alpha-1}}{\mathbb{P}_{\alpha}}$ decreases with

$$\left|\frac{p}{q} - \frac{(p)}{(q)}\right|,$$

visto que o que se diz do afastamento relativo a p, se diz igualmente do afastamento relativo a q.

Proposição VII

(Lema de Vallée-Poussin)

Designemos por T_i *a probabilidade duma combinação que tem* i *casos favoráveis e* T_n *a combinação mais provável.*

Pondo

$$S = T_{n-\alpha} + T_{n-\alpha+1} + \ldots + T_n + \ldots + T_{n+\alpha},$$

teremos,

$$1 - S < \frac{m}{\left(1 + \frac{\alpha}{(m+1)pq}\right)^{\frac{\alpha}{2}}}.$$

Com efeito, tem-se

$$T_{n+1} = \frac{m-n}{n+1} \frac{p}{q},$$

onde

$$n = (m+1)p - r$$

e

$$n + 1 = (m+1)p - r + 1$$

e

$$m - n = (m+1)q - 1 + r;$$

since what is true for the deviation from p is also true for the deviation from q.

Proposition VII

(Vallée–Poussin's Lemma)

We now denote T_i the probability of a combination of outcomes with i successes, and by T_n the probability of the most probable combination of outcomes.

Denoting

$$S = T_{n-\alpha} + T_{n-\alpha+1} + \cdots + T_n + \cdots + T_{n+\alpha}$$

we have

$$1 - S < \frac{m}{\left(1 + \dfrac{\alpha}{(m+1)\,p\,q}\right)^{\frac{\alpha}{2}}}\,.$$

In effect, we have

$$\frac{T_{n+1}}{T_n} = \frac{m-n}{n+1} \cdot \frac{p}{q}\,,$$

where

$$n = (m+1)\,p - r$$

and

$$n + 1 = (m+1)\,p - r + 1$$

and

$$m - n = (m+1)\,q - 1 + r\,;$$

e, pondo

$$1 - r = \varepsilon, \quad (0 < \varepsilon < 1),$$

virá

$$\frac{T_{n+1}}{T_n} = \frac{(m+1)\,q - \varepsilon}{(m+1)\,p + \varepsilon} \cdot \frac{p}{q} = \frac{1 - \dfrac{\varepsilon}{q\,(m+1)}}{1 + \dfrac{\varepsilon}{p\,(m+1)}}.$$

Ora,

$$\frac{1 - \dfrac{\varepsilon}{q\,(m+1)}}{1 + \dfrac{\varepsilon}{p\,(m+1)}} < \frac{1}{1 + \dfrac{\varepsilon}{pq(m+1)}}$$

qualquer que seja $\varepsilon > 0$, como fácilmente se verifica; logo

$$\frac{T_{n+1}}{T_n} < \frac{1}{1 + \dfrac{\varepsilon}{pq\,(m+1)}}$$

Do mesmo modo se obteem as desigualdades seguintes, visto ε poder ser qualquer número positivo

$$\frac{T_{n+2}}{T_{n+1}} = \frac{m - n - 1}{n + 1 + 1}\ \frac{p}{q} = \frac{(m+1)\,q - \varepsilon - 1}{(m+1)\,p + 2 + 1}\ \frac{p}{q}$$

$$= \frac{(m+1)\,q - (1+\varepsilon)}{(m+1)\,q + (1+\varepsilon)} \cdot \frac{p}{q}$$

$$< \frac{1}{1 + \dfrac{\varepsilon + 1}{pq\,(m+1)}};$$

with

$$1 - r = \varepsilon, \quad (0 \leq \varepsilon \leq 1),$$

we get

$$\frac{T_{n+1}}{T_n} = \frac{(m+1)\,q - \varepsilon}{(m+1)\,p + \varepsilon} \cdot \frac{p}{q} = \frac{1 - \dfrac{\varepsilon}{q\,(m+1)}}{1 + \dfrac{\varepsilon}{p\,(m+1)}}.$$

It is easily established that *for any* $\varepsilon > 0$

$$\frac{1 - \dfrac{\varepsilon}{q\,(m+1)}}{1 + \dfrac{\varepsilon}{p\,(m+1)}} < \frac{1}{1 + \dfrac{\varepsilon}{pq\,(m+1)}};$$

therefore

$$\frac{T_{n+1}}{T_n} < \frac{1}{1 + \dfrac{\varepsilon}{pq\,(m+1)}}.$$

As ε can be any positive number, using similar arguments we get

$$\frac{T_{n+2}}{T_{n+1}} = \frac{m-n-1}{n+1+1} \cdot \frac{p}{q} = \frac{(m+1)\,q - \varepsilon - 1}{(m+1)\,p + \varepsilon - 1} \cdot \frac{p}{q} =$$

$$= \frac{(m+1)\,q - (1+\varepsilon)}{(m+1)\,p + (1+\varepsilon)} \cdot \frac{p}{q} <$$

$$< \frac{1}{1 + \dfrac{\varepsilon + 1}{pq\,(m+1)}};$$

dum modo geral

$$\frac{T_{n+\alpha+1}}{T_{n+\alpha}} < \frac{1}{1+\dfrac{\varepsilon+\alpha}{pq\,(m+1)}}\,.$$

Multiplicando membro a membro estas desigualdades, vem

$$\frac{T_{n+\alpha+1}}{T_n} < \frac{1}{1+\dfrac{\varepsilon}{pq\,(m+1)}}\cdot\frac{1}{1+\dfrac{\varepsilon+1}{pq\,(m+1)}}\cdots\frac{1}{1+\dfrac{\varepsilon+\alpha}{pq\,(m+1)}}$$

$$< \frac{1}{1+\dfrac{1}{pq\,(m+1)}}\cdots\frac{1}{1+\dfrac{\alpha}{pq\,(m+1)}}$$

$$= \prod_{x=1}^{\alpha}\frac{1}{1+\dfrac{x}{pq\,(m+1)}}\,;$$

invertendo a ordem a todos os factores dêste producto, vem

$$\frac{T_{n+\alpha+1}}{T_n} < \prod_{x=1}^{\alpha}\frac{1}{1+\dfrac{\alpha+1-x}{pq\,(m+1)}}\,;$$

multiplicando membro a membro estas duas desigualdades vem

$$\left(\frac{T_{n+\alpha+1}}{T_n}\right)^2 < \prod_{x=1}^{\alpha}\left[\frac{1}{1+\dfrac{x}{pq\,(m+1)}}\cdot\frac{1}{1+\dfrac{\alpha+1-x}{pq\,(m+1)}}\right];$$

and, in general

$$\frac{T_{n+\alpha+1}}{T_{n+\alpha}} < \frac{1}{1+\dfrac{\varepsilon+\alpha}{pq\,(m+1)}}\,.$$

Multiplying term by term these inequalities, we get

$$\frac{T_{n+\alpha+1}}{T_n} < \frac{1}{1+\dfrac{\varepsilon}{pq\,(m+1)}}\cdot\frac{1}{1+\dfrac{\varepsilon+1}{pq\,(m+1)}}\cdots\frac{1}{1+\dfrac{\varepsilon+\alpha}{pq\,(m+1)}} <$$

$$< \frac{1}{1+\dfrac{1}{pq\,(m+1)}}\cdots\frac{1}{1+\dfrac{\alpha}{pq\,(m+1)}} =$$

$$= \prod_{k=1}^{\alpha}\frac{1}{1+\dfrac{k}{pq\,(m+1)}}\,;$$

reversing the order of all the above factors

$$\frac{T_{n+\alpha+1}}{T_n} < \prod_{k=1}^{\alpha}\frac{1}{1+\dfrac{\alpha+1-k}{pq\,(m+1)}}\,;$$

multiplying term by term those two inequalities, we get

$$\left(\frac{T_{n+\alpha+1}}{T_n}\right)^2 < \prod_{k=1}^{\alpha}\left[\frac{1}{1+\dfrac{k}{pq\,(m+1)}}\cdot\frac{1}{1+\dfrac{\alpha+1-k}{pq\,(m+1)}}\right]\,;$$

mas

$$\left[1+\frac{x}{pq\,(m+1)}\right]\left[1+\frac{\alpha+1-x}{pq\,(m+1)}\right]=$$

$$=1+\frac{\alpha+1}{pq\,(m+1)}+\frac{x\,(\alpha+1-x)}{[pq\,(m+1)]^2}$$

$$>1+\frac{\alpha+1}{pq\,(m+1)}\;;$$

logo

$$\left(\frac{\mathrm{T}_{n+\alpha+1}}{\mathrm{T}_n}\right)^2<\prod_{x=1}^{\alpha}\left[1+\frac{\alpha+1}{pq\,(m+1)}\right]^{-1}$$

$$=\left[1+\frac{\alpha+1}{pq\,(m+1)}\right]^{-\alpha}$$

$$<\left[1+\frac{\alpha}{pq\,(m+1)}\right]^{-\alpha};$$

donde

$$\mathrm{T}_{n+\alpha+1}<\mathrm{T}_n\left[1+\frac{\alpha}{pq\,(m+1)}\right]^{-\frac{\alpha}{2}}$$

e, visto que

$$\mathrm{T}_n<1,$$

$$\mathrm{T}_{n+\alpha+1}<\left[1+\frac{\alpha}{pq\,(m+1)}\right]^{-\frac{\alpha}{2}}.$$

Por outro lado, as probabilidades das combinações que but

$$\left[1 + \frac{k}{pq\,(m+1)}\right]\left[1 + \frac{\alpha + 1 - k}{pq\,(m+1)}\right] =$$

$$= 1 + \frac{\alpha + 1}{pq\,(m+1)} + \frac{k(\alpha + 1 - k)}{[pq\,(m+1)]^2} >$$

$$> 1 + \frac{\alpha + 1}{pq\,(m+1)}\,;$$

therefore

$$\left(\frac{T_{n+\alpha+1}}{T_n}\right)^2 < \prod_{k=1}^{\alpha}\left[1 + \frac{\alpha + 1}{pq\,(m+1)}\right]^{-1} =$$

$$= \left[1 + \frac{\alpha + 1}{pq\,(m+1)}\right]^{-\alpha} <$$

$$< \left[1 + \frac{\alpha}{pq\,(m+1)}\right]^{-\alpha};$$

from the above inequality, it follows that

$$T_{n+\alpha+1} < T_n \left[1 + \frac{\alpha}{pq\,(m+1)}\right]^{-\frac{\alpha}{2}}$$

and, as

$$T_n < 1,$$

we get

$$T_{n+\alpha+1} < \left[1 + \frac{\alpha}{pq\,(m+1)}\right]^{-\frac{\alpha}{2}}$$

On the other hand, the probabilities of the combinations of outcomes não entram em S, são todas inferiores a $T_{n+\alpha+1}$ (prop. III); logo

$$1-S<(2\alpha+1)T_{n+\alpha+1}<m\,T_{n+\alpha+1}$$

$$<\frac{m}{\left[1+\frac{\alpha}{pq(m+1)}\right]^{\frac{\alpha}{2}}},$$

c. d. d.

Proposição VIII

(3.º Teorema de Bernoulli)

Tem-se uma probabilidade sempre crescente de que a razão do número de acontecimentos favoráveis para o número de acontecimentos contrários se não afastará da razão das suas probabilidades respectivas alèm de certos limites; e, por mais apertados que esses limites sejam, a probabilidade de que se trata, aproximar-se-há da unidade tanto quanto se queira, logo que o número de experiências aumente suficientemente.

Com efeito, como já vimos na proposição V, caso o afastamento α seja positivo, teremos

$$\frac{(p)}{(q)}=\frac{p}{q}+\frac{p-r+\alpha}{mq-p+r-\alpha}$$

e portanto

$$\left|\frac{(p)}{(q)}-\frac{p}{q}\right|=\frac{p-r+\alpha}{mq-p+r-\alpha}.$$

which are not in S are all smaller than $T_{n+\alpha+1}$ (Prop. III); in consequence,

$$1 - S \, < \, [m - (2\,\alpha + 1)]\, T_{n+\alpha+1} \, < \, m\, T_{n+\alpha+1} \, <$$

$$< \, \frac{m}{\left[1 + \frac{\alpha}{pq\,(m+1)}\right]^{\frac{\alpha}{2}}} \, .$$

□

Proposition VIII

(Jacob Bernoulli's 3rd Theorem)

The probability that the deviation of the ratio $\frac{(p)}{(q)}$ *between the number of successes and the number of failures from the odds ratio* $\frac{p}{q}$ *of the corresponding probabilities falls within given bounds is always increasing to* 1*, when the number of trials is large enough, however tight these bounds may be.*

In effect, in Prop. V we have seen that, assuming the deviation to be positive,

$$\frac{(p)}{(q)} = \frac{p}{q} + \frac{p - r + \alpha}{mq - p + r - \alpha} \cdot \frac{1}{q}\,,$$

and therefore

$$\left| \frac{(p)}{(q)} - \frac{p}{q} \right| = \frac{p - r + \alpha}{mq - p + r - \alpha} \cdot \frac{1}{q}\,.$$

Ora, para que

$$\left| \frac{(p)}{(q)} - \frac{p}{q} \right| > \varepsilon,$$

necessário se torna que

$$\frac{p - r + \alpha}{mq - p + r - \alpha} > \varepsilon$$

ou

$$p - r + \alpha > mq\,\varepsilon - p\,\varepsilon + r\,\varepsilon - \alpha\,\varepsilon$$

ou

$$\alpha(1 + \varepsilon) > mq\,\varepsilon + \ldots$$

ou

$$\alpha > A(m + 1) + B; \qquad (1)$$

onde A e B são constantes, sendo $A > 0$.

Mas, S representa na prop. VII a probabilidade de que o afastamento seja igual ou menor do que α; $1 - S$ representará a probabilidade de que o afastamento seja superior a α.

Logo, a probabilidade de que

$$\left| \frac{(p)}{(q)} - \frac{p}{q} \right| > \varepsilon$$

será

$$1 - S < \frac{m}{\left[1 + \frac{\alpha}{(m+1)pq}\right]^{\frac{\alpha}{2}}}$$

$$< \frac{m}{\left[1 + \frac{A(m+1) + B}{pq\,(m+1)}\right]^{\frac{A(m+1)+B}{2}}}$$

For

$$\left|\frac{(p)}{(q)} - \frac{p}{q}\right| > \varepsilon$$

it is necessary that

$$\frac{p-r+\alpha}{mq-p+r-\alpha} \cdot \frac{1}{q} > \varepsilon$$

or

$$p-r+\alpha > mq^2\varepsilon - pq\varepsilon + rq\varepsilon - q\alpha\varepsilon$$

i.e.

$$\alpha\,(1+q\varepsilon) > mq^2\varepsilon + \cdots$$

or

$$\alpha > A\,(m+1) + B, \tag{5.3}$$

where $A > 0$ and B are constants.

But, in Prop. VII, S denotes the probability that the deviation is less than or equal to α; in other words, $1-S$ is the probability that the deviation is greater than α.

In consequence, the probability that

$$\left|\frac{(p)}{(q)} - \frac{p}{q}\right| > \varepsilon$$

is

$$1-S \;<\; \frac{m}{\left[1+\dfrac{\alpha}{(m+1)\,pq}\right]^{\frac{\alpha}{2}}} \;<$$

$$< \frac{m}{\left(1+\dfrac{A\,(m+1)+B}{pq\,(m+1)}\right)^{\frac{A\,(m+1)+B}{2}}}$$

expressão esta que tende para zero quando m tende para infinito, visto que $A > 0$.

Este teorema é tambêm chamado *lei dos grandes numeros.*

Observação

Não consideramos o caso de $\alpha < 0$ porque para êle chegavamos a uma desigualdade análoga a (1), ficando por isso caídos no mesmo caso.

*

O 3.º teorema de Bernoulli pode enunciar-se do seguinte modo:

A probabilidade de que o afastamento seja da ordem do número de experiências, tende para zero quando o número de experiências tende para infinito.

Proposição IX

A probabilidade de que o afastamento α *seja tal que*

$$\frac{\alpha^{n+1}}{m^n} > \varepsilon, \tag{1}$$

tende para zero quando m *cresce, logo que* $n > 1$.

Com efeito, de (1) tira-se

$$\frac{m}{\left[1 + \dfrac{\alpha}{pq\,(m+1)}\right]^{\frac{\alpha}{2}}} < \frac{m}{\left[1 + \dfrac{\varepsilon\, m^{\frac{n}{n+1}}}{(m+1)\,pq}\right]^{\frac{m^{\frac{n}{n+1}}}{2}}}.$$

an expression that decreases to zero when m goes to infinity, since $A > 0$.

This theorem is also known as *the law of large numbers*.

Observation

We do not explicit the case $\alpha < 0$ since in that case an inequality similar to (5.3) holds, and therefore it reduces to the former situation $\alpha > 0$.

*

The 3rd Bernoulli's theorem may be rephrased as follows:

The probability that the deviation [*of* $\frac{(p)}{(q)}$ *from the odds ratio* $\frac{p}{q}$] *is of the order of the number of trials decreases to zero when the number of trials goes to infinity.*

Proposition IX

The probability that the deviation α *is such that*

$$\frac{\alpha^{n+1}}{m^n} > \varepsilon, \tag{5.4}$$

decreases to zero when m *increases, if* $n > 1$*.*

In effect, from (5.4) we get

$$\frac{m}{\left[1 + \dfrac{\alpha}{pq\,(m+1)}\right]^{\frac{\alpha}{2}}} < \frac{m}{\left[1 + \dfrac{\varepsilon'\, m^{\frac{n}{n+1}}}{pq\,(m+1)}\right]^{\frac{\varepsilon'\, m^{\frac{n}{n+1}}}{2}}}.$$

Ora, ao segundo membro desta desigualdade pode dar-se a forma

$$\frac{m}{\left[1 + \mathrm{A}\, m^{-\frac{1}{1+n}}\right]^{\mathrm{B}\, m^{\frac{n}{n+1}}}},$$

desprezando as parcelas finitas em presença das infinitas, ou ainda, pela mesma razão, a forma

$$\frac{m}{1 + \mathrm{A}_1\, m^{\frac{n}{n+1}} \,.\, m^{-\frac{1}{n+1}} + \mathrm{A}_2\, m^{2\frac{n}{n+1}} \,.\, m^{-\frac{2}{n+1}} \ldots}$$

$$= \frac{m}{1 + \mathrm{A}_1\, m^{\frac{n-1}{n+1}} + \mathrm{A}_2\, m^{2\frac{n-1}{n+1}} + \ldots}$$

$$= \frac{1}{m^{-1} + \mathrm{A}_1\, m^{\frac{n-1}{n-1}-1} + \mathrm{A}_2\, m^{2\frac{n-1}{n+1}-1} + \ldots}$$

expressão esta que tenderá para zero logo que $n > 1$, porque neste caso haverá sempre um inteiro i tal que

$$i\,\frac{n-1}{n+1} - 1 > 0.$$

Esta proposição pode enunciar-se:

Há uma probabilidade nula de que a ordem do número de experiencias em relação ao afastamento a, seja inferior à segunda.

Disregarding vanishing terms, the second member of the above inequality can be approximated by

$$\frac{m}{\left[1 + Am^{-\frac{1}{n+1}}\right]^{B\,m^{\frac{n}{n+1}}}},$$

which in turn may be expanded as

$$\frac{m}{1 + A_1 m^{\frac{n}{n+1}}\, m^{-\frac{1}{n+1}} + A_2 m^{2\,\frac{n}{n+1}}\, m^{-\frac{2}{n+1}} + \cdots} =$$

$$= \frac{m}{1 + A_1 m^{\frac{n-1}{n+1}} + A_2 m^{2\,\frac{n-1}{n+1}} + \cdots} =$$

$$= \frac{1}{\frac{1}{m} + A_1 m^{\frac{n-1}{n+1}-1} + A_2 m^{2\,\frac{n-1}{n+1}-1} + \cdots}$$

an expression that decreases to zero when the number of trials goes to infinity, provided $n > 1$, since there exists some integer i for which

$$i\,\frac{n-1}{n+1} - 1 > 0.$$

This proposition may be rephrased as follows:

The probability that the number of trials is of order less than two of the deviation α is zero.

Proposição X

A probabilidade de que

$$\frac{\alpha}{\sqrt[n]{m}} < \varepsilon$$

tende para zero quando m *aumenta indefinidamente, se* $n > 2$.

Com efeito, visto que

$$T_{n-\alpha} < T_n = \frac{1 + \varepsilon_m}{\sqrt{2\pi p q m}},$$

segue-se que

$$S < 2\alpha T_n = \frac{2\alpha}{\sqrt{2\pi p q m}}(1 + \varepsilon_m),$$

ou

$$S < \frac{\alpha}{\sqrt{m}} C_m,$$

onde C_m tende para uma constante quando m tende para infinito. Ora, se

$$\frac{\alpha}{\sqrt[n]{m}} < \varepsilon,$$

será

$$S < \frac{\sqrt[n]{m}}{\sqrt{m}} \varepsilon C_m$$

$$= m^{\frac{1}{n} - \frac{1}{2}} . \varepsilon C_m$$

Proposition X

The probability that

$$\frac{\alpha}{\sqrt[n]{m}} < \varepsilon$$

decreases to zero when the number of trials goes to infinity, if $n > 2$.

In effect, as

$$T_{n-\alpha} < T_n = \frac{1 + \alpha_m}{\sqrt{2\,\pi\,p\,q\,m}}\,,$$

it follows that

$$S < (2\,\alpha + 1)T_n = \frac{2\,\alpha + 1}{\sqrt{2\,\pi\,p\,q\,m}}\,(1 + \alpha_m)\,,$$

or

$$S < \frac{\alpha}{\sqrt{m}}\,C_m,$$

where C_m converges to a constant when m goes to infinity. Therefore, if

$$\frac{\alpha}{\sqrt[n]{m}} < \varepsilon$$

we get

$$S < \frac{\sqrt[n]{m}}{\sqrt{m}}\,\varepsilon\,C_m =$$

$$= m^{\frac{1}{n} - \frac{1}{2}} \cdot \varepsilon\,C_m$$

e

$$\lim_{m=\infty} S = 0$$

se

$$\frac{1}{n} - \frac{1}{2} < 0, \qquad \text{ou,} \qquad n > 2,$$

c. d. d.

Logo, *é nula a probabilidade de que o número de experiências seja de ordem superior à segunda, em relação* α.

Portanto

Proposição XI

A ordem do número de experiências relativa ao afastamento será a segunda como resulta das proposições IX e X.

*

* *

Todo o número se pode supôr escrito com a forma decimal e com um número infinito de casas. Assim, o número $\frac{1}{2}$ pode escrever-se debaixo da forma 0,5000...

Supondo todos os números com esta forma, consideremos o seguinte

Problema

Tira-se, à sorte, um número do intervalo (0, 1); *qual a probabilidade de que os seus algarismos se sucedam de modo a satisfazerem à lei de* BERNOULLI?

and thus

$$\lim_{m\to\infty} S = 0$$

if

$$\frac{1}{n} - \frac{1}{2} < 0, \qquad \text{or} \qquad n > 2.$$

$\square$

From this, it follows that *the probability that the number of trials is of order greater than two of the deviation α is zero.*

Henceforth,

Proposition XI

The number of trials is of order 2 in what regards the deviation α, an immediate consequence of Prop. IX and X.

*

*   *

Any real number can be written in decimal form with an infinite number of decimal places. For instance, the number $\frac{1}{2}$ can be written 0.50000...

Assuming that we are dealing with numbers in that form of representation, we shall examine the following

Problem

A number is randomly chosen in the interval $(0, 1)$; what is the probability that the sequence of digits in its decimal expansion satisfies Bernoulli's law?

Êste problema, assim posto, não encontra solução nas definições que demos da probabilidade.

Com efeito, êste problema pertence à probabilidade contínua, visto que a tiragem se faz no intervalo (0, 1) e nós, no capítulo correspondente dos presentes *Elementos,* definimos, apenas, probabilidade de *regiões* em relação a *regiões*. Portanto, para que êste problema lá encontrasse resposta, necessário seria que os números que constituem a classe favorável, constituissem intervalos completos, contidos no intervalo (0, 1), o que não é verdade; ou, talvez melhor, o que não pode afirmar-se *à priori*.

Modifiquemos por isso o enunciado do problema, do seguinte modo: *qual a probabilidade de que os primeiros* N *algarismos do numero achado satisfaçam à lei de* BERNOULLI ?

Êste problema já tem uma solução, porque os números cujos primeiros N algarismos são idênticos, formam um intervalo.

Esses intervalos uns serão favoráveis, outros não.

A soma dos intervalos favoráveis, dividida pelo intervalo total, dará a solução pedida.

Ora, sendo estes intervalos todos iguais, eles serão igualmente possíveis e a possibilidade de cada um deles será

$$\frac{1}{10^N}$$

visto que 10^N é o seu número.

Consideremos, agora, uma urna com 10 bolas, numeradas de 0 a 9. Qualquer número com N casos se pode

The problem as stated has no solution with the definition of probability that we have adopted.

In effect, we have a problem of continuous probability, since the random extraction is performed in the interval $(0, 1)$, and in Chapter II we have solely defined the probability of *regions* in reference to other possible *regions*. Therefore, in the context we have adopted, the problem would have a solution if the favorable class, i.e. the sequence of digits of a number, would be a complete interval contained in $(0, 1)$, which is not true, or at least *a priori* cannot be taken for granted.

In view of that, we shall instead consider a simpler problem, namely: *what is the probability that the N first digits in the decimal expression of a number randomly chosen in* $(0, 1)$ *satisfy Bernoulli's law?*

This problem has a solution, since the set of numbers which share the first N digits is an interval.

Some of these intervals are favorable, others aren't.

The measure of the union of the favorable intervals is the solution we search.

As all the intervals are identical, they are equally possible, and as there are 10^N of those intervals, the possibility of each of them is

$$\frac{1}{10^N}\,.$$

Let us consider one urn with 10 balls, numbered 0, 1, ..., 9. The first N digits of any number in $(0, 1)$ can be identified supôr obtido por meio de N tiragens nessa urna e a possibilidade de cada um dêsses números será

$$\frac{1}{10} \cdot \frac{1}{10} \cdots \frac{1}{10} = \frac{1}{10^N},$$

isto é, a mesma de qualquer dos intervalos parciais precedentes.

Portanto, no problema em questão, tanto faz lançar à sorte um ponto no intervalo (0, 1), como fazer N tiragens, à sorte, numa urna que contenha 10 bolas, numeradas de 0 a 9.

Á mesma conclusão podiamos chegar, servindo-nos da fórmula do problema tratado na página 74 dos presentes *Elementos,* onde teriamos de supôr $f(x) = x$, o que nos dava, imediatamente,

$$P = \frac{1}{10}$$

qualquer que fosse a e d.

Mas sendo assim, a relação do numero de combinações que satisfazem ao teorema de BERNOULLI, para o número total de combinações, tende para 1, à medida que N aumenta. É nisto mesmo que consiste o teorema de BERNOULLI.

Logo, a probabilidade pedida no segundo inunciado, tende para 1, à medida que N aumenta. A probabilidade pedida no primeiro que corresponde ao limite para $N = \infty$, será igual a 1.

with the sequence of digits obtained in N random extractions, with replacement, of balls from that urn. In effect, the possibility of each sequence is

$$\frac{1}{10} \cdot \frac{1}{10} \cdots \frac{1}{10} = \frac{1}{10^N},$$

the same that we had found previously for each of the partial intervals.

Therefore, in the problem at hand, randomly selecting a number in the interval $(0,1)$ is the same as performing N random extractions, with replacement, of balls of the urn described.

The same conclusion could be reached, also, using the expression obtained in the problem solved in page 74 of these *Elements*, assuming $f(x) = x$, that would immediately imply that

$$\mathbb{P} = \frac{1}{10}$$

whatever a and d.

From the identification of the modified problem with random extractions, it is immediate that the ratio between the number of combinations of outcomes that satisfy Bernoulli's law and the total number of combinations of outcomes increases towards 1, as N increases. This is the core of Bernoulli's theorem.

Thus, the probability asked for in the restated problem, increases to 1 when N increases. We may conclude that the probability asked for in the original problem, which corresponds to the limit when $N \to \infty$, is equal to 1.

*

Êste mesmo raciocínio prova que é igual a 1 a probabilidade de que sejam satisfeitas todas as outras leis análogas às de Bernoulli.

*

Daqui podemos concluir que a probabilidade de que o número achado seja racional, é igual a zero.

Com efeito, um número racional dá sempre logar a uma dizima periódica.

E, de duas uma: ou bem no período do número racional entram todos os algarismos de 0 a 9 e em proporções iguais; e nesse caso a distribuição dos algarismos satisfarà à lei de Bernoulli, mas não satisfará a nenhuma das outras, visto que o afastamento absoluto passará periodicamente pelos mesmos valores e portanto ficará inferior a certo máximo; ou o período do número racional não satisfaz à condição precedente e nesse caso não é satisfeita a lei de Beknoulli.

Os números racionais correspondem pois, a uma classe de combinações que, ou não satisfazem à lei de Bernoulli ou não satisfazem às leis análogas. A sua probabilidade será, pois, nula.

*

The argument above also shows that the probability that the sequence of digits of a randomly chosen number in $(0, 1)$ satisfies all laws similar to that of Bernoulli is 1.

*

From this we may deduce that *the probability that a randomly chosen number in* $(0, 1)$ *is rational is zero.*

In effect, a rational number has periodic decimal representation.

And the two possible cases are: either all digits from 0 to 9 appear, in the same proportion, in its period, or this is not so. In the second instance, the sequence of digits doesn't satisfy Bernoulli's law. On the other hand, in the first case, the distribution of the digits in the sequence satisfies Bernoulli's law, but none of the others, since the absolute deviation will periodically take on the same values, therefore it will never remain lower than a given bound.

Therefore, the set of rational numbers corresponds to a combination of outcomes that either doesn't satisfy Bernoulli's law or doesn't satisfy the analogous laws. Its probability is, therefore, zero.

SEGUNDA PARTE

Lei dos desvios

Demonstrado, dum modo rigoroso, o 3.º teorema de JACOB BERNOULLI e outros análogos, relativos à ordem de grandeza dos afastamentos ou desvios, vamos deduzir uma relação aproximada entre os desvios e as suas probabilidades.

Proposição XII

Representando por T_n a probabilidade da combinação de probabilidade máxima, combinação esta a que podemos chamar *normal*, a probabilidade de que o afastamento seja inferior, em valor absoluto, a k, será dada por

$$P(k) = T_{n-k} + \ldots + T_n + \ldots + T_{n+k} = \sum_{i=-k}^{k} T_{n+i}.$$

Ora,

$$T_{n+i} = \frac{m!}{(n+i)!\,(m-n-i)!} p^{n+i} q^{m-n-i},$$

onde n representa o número de bolas brancas da combinação normal e por isso é da forma

$$n = (m+1)p - r, \quad (0 \leqq r \leqq 1).$$

Supondo m bastante grande, poderemos, na fórmula aproximada que nos dá

$$(m-n+i)! \quad \text{e} \quad (n+i)!,$$

SECOND PART

LAW OF DEVIATIONS

(ERROR LAW)

After having presented a rigorous proof of Jacob Bernoulli's 3rd theorem and others similar theorems on the order of magnitude of deviations [from the most probable combination of outcomes], we shall now establish an approximate relation between the deviations and their probabilities.

Proposition XII

Denoting T_n the probability of the combination of outcomes of maximum probability, which we may call *normal* combination, the probability that the deviation has absolute value less than or equal to k is

$$\mathbb{P}(k) = T_{n-k} + \cdots + T_n + \cdots + T_{n+k} = \sum_{i=-k}^{k} T_{n+i},$$

where

$$T_{n+i} = \frac{m!}{(n+i)!\,(m-n-i)!}\,p^{n+i} q^{m-n-i},$$

where n denotes the number of white balls in the normal combination, and therefore is of the form

$$n = (m+1)\,p - r, \quad (0 \leq r \leq 1).$$

Assuming that m is large enough, we can use substituir

$$(m+1)p-r$$

por

$$mp$$

e T_{n+i} tomará a forma

$$T_{n+1}=\frac{m^m e^{-m}\sqrt{2\pi m}\;\; p^{mp+i}q^{mq-i}(1+\alpha_m)}{(mp+i)^{mp+i}e^{-mp-i}\sqrt{2\pi(mp+i)}(mq-i)^{mq-i}e^{-mq-i}\sqrt{2\pi(mq-i)}}$$

$$=\frac{\sqrt{m}\,p^{mp+i}\,q^{mq-i}}{\left(p+\frac{i}{m}\right)^{mp+i}\left(q-\frac{i}{m}\right)^{mq-i}\sqrt{2\pi(mp+i)\,(mq-i)}}(1+\alpha_m)$$

$$=\frac{p^{mp+i}\,q^{mq-i}}{\left(p+\frac{i}{m}\right)^{mp+i+\frac{1}{2}}\left(q-\frac{i}{m}\right)^{mq-i+\frac{1}{2}}\sqrt{2\pi m}}(1+\alpha_m)$$

$$=\frac{1}{\sqrt{2\pi mpq}}\cdot\frac{1}{\left(1+\frac{i}{mp}\right)^{mp+i+\frac{1}{2}}\left(1-\frac{i}{mq}\right)^{mq-i+\frac{1}{2}}}(1+\alpha_m)$$

$$=\frac{1}{\sqrt{2\pi mpq}}\,.\,H$$

sendo

$$H=\left(1+\frac{i}{mp}\right)^{-mp-i-\frac{1}{2}}\left(1+\frac{i}{mq}\right)^{-mq+i-\frac{1}{2}}(1+\alpha_m)^{-1}$$

$$mp$$

as an approximation for

$$(m+1)\,p - r;$$

in Stirling's approximation for

$$(m-n-i)! \qquad \text{and} \qquad (n+i)!.$$

T_{n+i} may then be approximated by

$$T_{n+i} = \frac{m^m\, \mathrm{e}^{-m}\, \sqrt{2\,\pi\, m} p^{mp+i}\, q^{mq-i} (1+\alpha_m)}{(mp+i)^{mp+i}\, \mathrm{e}^{-mp-i}\, \sqrt{2\,\pi\,(mp+i)}\,(mq-i)^{mq-i}\, \mathrm{e}^{-mq+i}\, \sqrt{2\,\pi\,(mq-i)}} =$$

$$= \frac{\sqrt{m}\, p^{mp+i}\, q^{mq-i}}{\left(p+\dfrac{i}{m}\right)^{mp+i} \left(q-\dfrac{i}{m}\right)^{mq-i} \sqrt{2\,\pi\,(mp+i)\,(mq-i)}}\,(1+\alpha_m) =$$

$$= \frac{p^{mp+i}\, q^{mq-i}}{\left(p+\dfrac{i}{m}\right)^{mp+i+\frac{1}{2}} \left(q-\dfrac{i}{m}\right)^{mq-i+\frac{1}{2}} \sqrt{2\,\pi\, m}}\,(1+\alpha_m) =$$

$$= \frac{1}{\sqrt{2\,\pi\, mpq}} \cdot \frac{1}{\left(1+\dfrac{i}{mp}\right)^{mp+i+\frac{1}{2}} \left(1-\dfrac{i}{mq}\right)^{mq-i+\frac{1}{2}}}\,(1+\alpha_m) =$$

$$= \frac{1}{\sqrt{2\,\pi\, mpq}} \cdot H$$

where

$$H = \left(1+\frac{i}{mp}\right)^{-mp-i-\frac{1}{2}} \left(1-\frac{i}{mq}\right)^{-mq+i-\frac{1}{2}} (1+\alpha_m)$$

e

$$\log \mathrm{H} = -\left(m\,p + i + \frac{1}{2}\right)\log\left(1 + \frac{i}{mp}\right)$$

$$-\left(mq - i + \frac{1}{2}\right)\log\left(1 - \frac{i}{m\,q}\right) - \log\left(1 + \alpha_{n1}\right);$$

mas (3.º teorema de Bernoulli) a probabilidade de que

$$\left|\frac{i}{m}\right| > \varepsilon$$

tende para zero quando m aumenta; haverá pois, uma probabilidade sempre crescente de que

$$\log \mathrm{H} = -\left(mp + i + \frac{1}{2}\right)\left[\frac{i}{mp} - \frac{i^2}{2m^2p^2} + \frac{i^3}{3m^3p^3} - \dots\right]$$

$$-\left(mq - i + \frac{1}{2}\right)\left[-\frac{i}{m\,q} - \frac{i^2}{2m^2q^2} - \frac{i^3}{3m^3q^3} - \dots\right]$$

$$-\log\left(1 + \alpha^m\right) =$$

$$= -\frac{i^2}{m}\left[\frac{1}{p} + \frac{1}{q}\right] - \frac{1}{2} \cdot \frac{i}{m}\left[\frac{1}{p} + \frac{1}{q}\right]$$

$$+ \frac{i^2}{2m}\left[\frac{1}{p} + \frac{1}{q}\right] + \frac{i^3}{2m^2}\left[\frac{1}{p^2} - \frac{1}{q^2}\right]$$

$$+ \frac{1}{2}\,\frac{i^2}{2m^2}\left[\frac{1}{p^2} - \frac{1}{q^2}\right] + \dots - \log\left(1 + \alpha_m\right)$$

$$= -\frac{i^2}{2mpq} - \frac{i}{2mpq} + \frac{i^3}{2\,m^2p^2q_2} + \dots$$

and

$$\log\, H = -\left(mp + i + \frac{1}{2}\right)\, \log\left(1 + \frac{i}{mp}\right) -$$

$$-\left(mq - i + \frac{1}{2}\right) \log\left(1 - \frac{i}{mq}\right) + \log(1 + \alpha_m)$$

But (Bernoulli's 3rd theorem) the probability that

$$\left|\frac{i}{m}\right| > \varepsilon$$

decreases to zero when m increases; therefore the probability that

$$\log\, H = -\left(mp + i + \frac{1}{2}\right)\, \left[\frac{i}{mp} - \frac{i^2}{2m^2p^2} + \frac{i^3}{3m^3p^3} - \cdots\right] -$$

$$-\left(mq - i + \frac{1}{2}\right) \left[-\frac{i}{mq} - \frac{i^2}{2m^2q^2} - \frac{i^3}{3m^3q^3} - \cdots\right] +$$

$$+ \log(1 + \alpha_m) =$$

$$= -\frac{i^2}{m}\,\left(\frac{1}{p} + \frac{1}{q}\right) - \frac{1}{2}\cdot\frac{i}{m}\,\left(\frac{1}{p} - \frac{1}{q}\right) +$$

$$+\frac{i^2}{2m}\,\left(\frac{1}{p} + \frac{1}{q}\right) + \frac{i^3}{2m^2}\,\left(\frac{1}{p^2} - \frac{1}{q^2}\right) +$$

$$+\frac{1}{2}\,\frac{i^2}{2m^2}\,\left(\frac{1}{p^2} + \frac{1}{q^2}\right) + \cdots + \log(1 + \alpha_m) =$$

$$= -\frac{i^2}{2mpq} - \frac{i(q-p)}{2mpq} + \frac{i^3(q^2 - p^2)}{2m^2p^2q^2} + \cdots$$

is always increasing.

Ora, o termo em α_m tende para zero, como se sabe da fórmula de aproximação; os termos em

$$\frac{i}{m}, \quad \frac{i^3}{m^2}, \ldots, \quad \frac{i^n}{m^n}, \quad \frac{i^{n+1}}{m^n},$$

tem uma probabilidade de se manterem superiores a ε, por menor que ε seja, que tende para zero quando o número m aumenta (prop. IX e X); logo, haverá uma probabilidade sempre crescente com m de que H seja tal que

$$\log \mathrm{H} = -\frac{i^2}{2mpq}$$

e portanto

$$\mathrm{H} = e^{-\frac{i^2}{2mpq}}$$

e

$$\mathrm{T}_{n+1} = \frac{1}{\sqrt{2\pi mpq}}\, e^{-\frac{i^2}{2mpq}}$$

e

$$\mathrm{P}(k) = \sum_{i=-k}^{k} \mathrm{T}_{n+i} = \frac{1}{\sqrt{2\pi mpq}} \sum_{i=-k}^{k} e^{-\frac{i^2}{2mpq}}$$

$$= \frac{2}{\sqrt{2\pi mpq}} \sum_{i=0}^{k} e^{-\frac{i^2}{2mpq}};$$

substituindo o Σ por um $\int$ tomado entre os mesmos limites, virá

$$\mathrm{P}(k) = \frac{2}{\sqrt{2\pi pqm}} \int_0^k e^{\frac{x^2}{2mpq}}\, dx,$$

On one hand, from what we know about Stirling's approximation, the term in α_m goes to zero; on the other hand, the probability that the summands

$$\frac{i}{m}\,,\ \frac{i^3}{m^2}\,,\ \ldots,\ \frac{i^n}{m^n}\,,\ \frac{i^{n+1}}{m^n}\,,\ \ldots$$

remain greater than ε, however small, also goes to zero when m increases (Prop. IX and X); therefore, the probability that

$$\log\, H \approx -\frac{i^2}{2mpq}$$

is always increasing, and thus

$$H \approx \mathrm{e}^{-\frac{i^2}{2mpq}}\,;$$

from this it follows that

$$T_{n+i} \approx \frac{1}{\sqrt{2\,\pi\, mpq}}\ \mathrm{e}^{-\frac{i^2}{2mpq}}$$

and

$$\mathbb{P}(k) = \sum_{i=-k}^{k} T_{n+i} \approx \frac{2}{\sqrt{2\,\pi\, mpq}}\ \sum_{i=0}^{k} \mathrm{e}^{-\frac{i^2}{2mpq}}\ :$$

substituting $\sum$ by $\int$ to be computed between the same limits, we have

$$\mathbb{P}(k) \approx \frac{2}{\sqrt{2\,\pi\, mpq}}\ \int_0^k \mathrm{e}^{-\frac{x^2}{2mpq}}\,\mathrm{d}x\,,$$

ou, fazendo

$$x = \lambda \sqrt{2mpq},$$

$$P(k) = \frac{2}{\sqrt{2\pi mpq}} \int_0^{\lambda_1} e^{-\lambda^2} d\lambda \,.\, \sqrt{2mpq}\,,$$

$$= \frac{1}{\sqrt{\pi}} \int_0^{\lambda_1} e^{-\lambda^2} d\lambda\,,$$

onde

$$\lambda_1 = \frac{k}{\sqrt{2mpq}}.$$

A λ_1 chama-se *afastamento relativo* para o distinguir do afastamento k, também chamado *afastamento absoluto*. Ao número

$$\sqrt{2mpq}$$

chama-se *unidade de afastamento*.

*

Costuma chamar-se *probabilidade de* λ, à probabilidade de que o afastamento relativo se mantenha, em valor absoluto, inferior a λ.

Portanto,

and with the substitution

$$x = \lambda \sqrt{2mpq},$$

$$\mathbb{P}(k) = \frac{2}{\sqrt{2\,\pi\, mpq}} \int_{0}^{\lambda_1} \mathrm{e}^{-\lambda^2}\, \mathrm{d}\lambda \sqrt{2mpq} =$$

$$= \frac{2}{\sqrt{\pi}} \int_{0}^{\lambda_1} \mathrm{e}^{-\lambda^2}\, \mathrm{d}\lambda,$$

where

$$\lambda_1 = \frac{k}{\sqrt{2mpq}}\,.$$

We shall say that λ_1 is the *relative deviation*, to distinguish it from the *absolute deviation* k. The value

$$\sqrt{2mpq}$$

is the *deviation unit*.

*

The probability that the absolute value of the relative deviation is smaller than λ_1 is the *probability of* λ_1.

Proposição XIII

Haverá uma probabilidade sempre crescente com m *de que seja*

$$\theta(\lambda) = \frac{1}{\sqrt{\pi}} \int_0^{\lambda} e^{-\lambda^2} d\lambda$$

a probabilidade do afastamento λ.

A $\theta(\lambda)$ costuma chamar-se *lei dos desvios, lei dos afastamentos* e tambêm *lei de Gauss*.

*

A lei que acabamos de deduzir é apenas uma lei *provável* e, alêm disso, aproximada. A sua probabilidade, porêm, tende muito rapidamente para 1, quando m aumenta, e os êrros cometidos na sua dedução tendem rapidamente para zero. A conveniência desta lei é tal que, em muitas aplicações, o resultado obtido é igual ao resultado verdadeiro. Na resolução dos problemas sôbre os afastamentos é ela sempre que se emprega.

*

A probabilidade de que a variável λ esteja compreendida entre 0 e ∞ será

$$P = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-\lambda^2} d\lambda = \frac{2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = 1,$$

Proposition XIII

When m increases, the probability that the probability of the absolute deviation is less than λ_1 is given by

$$\theta(\lambda_1) = \frac{2}{\sqrt{\pi}} \int_0^{\lambda_1} e^{-\lambda^2} \, d\lambda$$

also increases.

We shall refer to $\theta(\lambda_1)$ as the *error law*, *the deviations law*, or *Gauss' law*.

*

The law we have established is only a *probable* law and, in addition to that, an approximate result. Its probability, however, rapidly converges to 1 when m increases, and the errors incurred in the approximations used to deduce it vanish very quickly. The rate of convergence is so high that the approximation it gives is, in many practical applications, equal to the true result. It is always used in problems about deviations.

*

The probability that the variable $|\Lambda|$, the absolute value of the relative deviation, is between 0 and ∞ is

$$\mathbb{P} = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-\lambda^2} \, d\lambda = \frac{2}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2} = 1,$$

resultado êste, rigoroso, como era de esperar; quando λ tende para infinito, m tende também para o infinito e nestas condições a lei torna-se *exacta.*

*

As tabelas seguintes dão os valores de $\theta(\lambda)$ de centesima em centesima. Por elas se vê quão rapidamente $\theta(\lambda)$ tende para 1, quando λ cresce.

λ.	θ.	λ.	θ.	λ.	θ.
0,00....	0,0000000	0,24....	0,2657000	0,48...	0,5027498
0,01....	0,0112833	0,25....	0,2763263	0,49....	0,5116683
0,02...	0,0225644	0,26....	0,2868997	0,50....	0,5204999
0,03....	0,0338410	0,27....	0,2974182	0,51....	0,5292437
0,04....	0,0451109	0,28....	0,3078800	0,52....	0,5378987
0,05....	0,0563718	0,29....	0,3182834	0,53....	0,5464641
0,06 ...	0,0676215	0,30. ..	0,3286267	0,54....	0,5549392
0,07....	0,0788577	0,31....	0,3389081	0,55....	0,5633233
0,08....	0,0900781	0,32....	0,3491259	0,56....	0,5716157
0,09....	0,1012806	0,33....	0,3592785	0,57....	0,5798158
0,10....	0,1124630	0,34....	0,3693644	0,58....	0,5879229
0,11....	0,1236230	0,35 ...	0,3793819	0,59...	0,5959365
0,12....	0,1347584	0,36....	0,3893296	0,60....	0,6038561
0,13....	0,1458671	0,37....	0,3992059	0,61....	0,6116812
0,14....	0,1569470	0,38....	0,4090093	0,62....	0,6194114
0,15....	0,1679959	0,39....	0,4187380	0,63....	0,6270463
0,16....	0,1790117	0,40....	0,4283922	0,64 ...	0,6345857
0,17....	0,1899923	0,41....	0,4379690	0,65....	0,6420292
0,18....	0,2009357	0,42....	0,4474676	0,66....	0,6493765
0,19....	0,2118398	0,43....	0,4568867	0,67....	0,6566275
0,20....	0,2227025	0,44....	0,4662251	0,68....	0,6637820
0,21....	0,2335218	0,45....	0,4754818	0,69 ...	0,6708399
0,22....	0,2442958	0,46....	0,4846555	0,70....	0,6778010
0,23....	0,2550225	0,47....	0,4937452	0,71....	0,6846654

this being a rigorous result, as should be expected; when $\lambda_1 \to \infty$, $m \to \infty$ also, and under that condition the law is *exact*.

*

The table that follows[19] shows the value of $\theta(\lambda_1)$ for centesimal increases in the argument. They show how quickly $\theta(\lambda_1) \to 1$ when λ_1 increases.

λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$
0.00	0.0000000		0.24	0.2657001		0.48	0.5027497
0.01	0.0112834		0.25	0.2763264		0.49	0.5116683
0.02	0.0225646		0.26	0.2868997		0.50	0.5204999
0.03	0.0338412		0.27	0.2974182		0.51	0.5292436
0.04	0.0451111		0.28	0.3078801		0.52	0.5378986
0.05	0.0563720		0.29	0.3182835		0.53	0.5464641
0.06	0.0676216		0.30	0.3286268		0.54	0.5549393
0.07	0.0788577		0.31	0.3389082		0.55	0.5633234
0.08	0.0900781		0.32	0.3491260		0.56	0.5716158
0.09	0.1012806		0.33	0.3592787		0.57	0.5798158
0.10	0.1124629		0.34	0.3693645		0.58	0.5879229
0.11	0.1236229		0.35	0.3793821		0.59	0.5959365
0.12	0.1347584		0.36	0.3893297		0.60	0.6038561
0.13	0.1458671		0.37	0.3992060		0.61	0.6116812
0.14	0.1569470		0.38	0.4090095		0.62	0.6194115
0.15	0.1679960		0.39	0.4187387		0.63	0.6270464
0.16	0.1790118		0.40	0.4283924		0.64	0.6345858
0.17	0.1899925		0.41	0.4379691		0.65	0.6420293
0.18	0.2009358		0.42	0.4474676		0.66	0.6493767
0.19	0.2118399		0.43	0.4568867		0.67	0.6566277
0.20	0.2227026		0.44	0.4662251		0.68	0.6637822
0.21	0.2335219		0.45	0.4754817		0.69	0.6708401
0.22	0.2442959		0.46	0.4846554		0.70	0.6778012
0.23	0.2550226		0.47	0.4937451		0.71	0.6846656

(19) This table has been recalculated using Mathematica 5.1. Observe the accuracy of the computations in the original.

λ.	θ.	λ.	θ.	λ.	θ.
0,72...	0,6914330	1,08....	0,8733261	1,44....	0,9582966
0,73....	0,6981038	1,09....	0,8768030	1,45....	0,9596950
0,74....	0,7046780	1,10....	0,8802050	1,46....	0,9610535
0,75....	0,7111556	1,11....	0,8835330	1,47....	0,9623729
0,76....	0,7175367	1,12....	0,8867879	1,48....	0,9636541
0,77....	0,7238216	1,13....	0,8899707	1,49....	0,9648979
0,78....	0,7300104	1,14....	0,8930823	1,50....	0,9661052
0,79....	0,7361035	1,15 ...	0,8961238	1,51....	0,9672768
0,80. ..	0,7421010	1,16....	0,8990962	1,52....	0,9684135
0,81....	0,7480033	1,17....	0,9020004	1,53....	0,9695162
0,82....	0,7538108	1,18....	0,9048374	1,54....	0,9705857
0,83....	0,7595238	1,19....	0,9076083	1,55....	0,9716227
0,84....	0,7651427	1,20....	0,9103140	1,56....	0,9726281
0,85....	0,7706680	0,21....	0,9129555	1,57....	0,9736026
0,86....	0,7761002	0,22....	0,9155359	1,58....	0,9745470
0,87....	0,7814398	0,23....	0,9180501	1,59....	0,9754620
0,88....	0,7866873	1,24....	0,9205052	1,60....	0,9763484
0,89....	0,7918432	1,25....	0,9229001	1,61....	0,9772069
0,90....	0,7969082	1,26....	0,9252359	1,62....	0,9780381
0,91....	0,8018828	1,27....	0,9275136	1,63....	0,9788429
0,92....	0,8067677	1,28....	0,9297342	1,64....	0,9796218
0,93....	0,8115635	1,29....	0,9318987	1,65....	0,9803756
0,94....	0,8162710	1,30....	0,9340080	1,66....	0,9811049
0,95....	0,8208908	1,31....	0,9360632	1,67....	0,9818104
0,96....	0,8254236	1,32....	0,9380652	1,68....	0,9824928
0,97....	0,8298703	1,33....	0,9400150	1,69....	0,9831526
0,98....	0,8342315	1,34....	0,9419137	1,70....	0,9837904
0,99 ...	0,8385081	1,35....	0,9437622	1,71....	0,9844070
1,00 ...	0,8427008	1,36....	0,9455614	1,72....	0,9850028
1,01....	0,8468105	1,37....	0,9473124	1,73....	0,9855785
1,02....	0,8508380	1,38....	0,9490160	1,74. ..	0,9861346
1,03....	0,8547842	1,39....	0,9506733	1,75....	0,9866717
1,04....	0,8586499	1,40....	0,9522851	1,76....	0,9871903
1,05....	0,8624360	1,41....	0,9538524	1,77....	0,9876910
1,06....	0,8661435	1,42....	0,9553762	1,78....	0,9881742
1,07....	0,8697732	1,43....	0,9568573	1,79....	0,9886406

λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$
0.72	0.6914331		1.08	0.8733262		1.44	0.9582966
0.73	0.6981039		1.09	0.8768031		1.45	0.9596950
0.74	0.7046781		1.10	0.8802051		1.46	0.9610535
0.75	0.7111556		1.11	0.8835330		1.47	0.9623729
0.76	0.7175368		1.12	0.8867879		1.48	0.9636541
0.77	0.7238216		1.13	0.8899707		1.49	0.9648979
0.78	0.7300104		1.14	0.8930823		1.50	0.9661051
0.79	0.7361035		1.15	0.8961238		1.51	0.9672767
0.80	0.7421010		1.16	0.8990962		1.52	0.9684135
0.81	0.7480033		1.17	0.9020004		1.53	0.9695162
0.82	0.7538108		1.18	0.9048374		1.54	0.9705857
0.83	0.7595238		1.19	0.9076083		1.55	0.9716227
0.84	0.7651427		1.20	0.9103140		1.56	0.9726281
0.85	0.7706681		1.21	0.9129555		1.57	0.9736026
0.86	0.7761003		1.22	0.9155339		1.58	0.9745470
0.87	0.7814398		1.23	0.9180501		1.59	0.9754620
0.88	0.7866873		1.24	0.9205052		1.60	0.9763484
0.89	0.7918432		1.25	0.9229001		1.61	0.9772068
0.90	0.7969082		1.26	0.9252359		1.62	0.9780381
0.91	0.8018828		1.27	0.9275136		1.63	0.9788428
0.92	0.8067677		1.28	0.9297342		1.64	0.9796218
0.93	0.8115636		1.29	0.9318986		1.65	0.9803756
0.94	0.8162710		1.30	0.9340079		1.66	0.9811049
0.95	0.8208908		1.31	0.9360631		1.67	0.9818104
0.96	0.8254236		1.32	0.9380652		1.68	0.9824928
0.97	0.8298703		1.33	0.9400150		1.69	0.9831526
0.98	0.8342315		1.34	0.9419137		1.70	0.9837905
0.99	0.8385081		1.35	0.9437622		1.71	0.9844070
1.00	0.8427008		1.36	0.9455614		1.72	0.9850028
1.01	0.8468105		1.37	0.9473124		1.73	0.9855785
1.02	0.8508380		1.38	0.9490160		1.74	0.9861346
1.03	0.8547842		1.39	0.9506733		1.75	0.9866717
1.04	0.8586499		1.40	0.9522851		1.76	0.9871903
1.05	0.8624361		1.41	0.9538524		1.77	0.9876909
1.06	0.8661436		1.42	0.9553762		1.78	0.9881742
1.07	0.8697733		1.43	0.9568573		1.79	0.9886405

λ.	θ.	λ.	θ.	λ.	θ.
1,80....	0,9890905	2,16....	0,9977472	2,52....	0,9996345
1,81....	0,9895245	2,17....	0,9978511	2,53....	0,9996537
1,82....	0,9899431	2,18....	0,9979505	2,54....	0,9996720
1,83....	0,9903467	2,19....	0,9980459	2,55....	0,9996893
1,84....	0,9907359	2,20....	0,9981372	2,56....	0,9997058
1,85....	0,9911110	2,21....	0,9982244	2,57....	0,9997215
1,86....	0,9914725	2,22....	0,9983079	2,58 ...	0,9997364
1,87....	0,9918207	2,23 ...	0,9983878	2,59....	0,9997505
1,88....	0,9921562	2,24....	0,9984642	2,60 ...	0,9997640
1,89....	0,9924793	2,25....	0,9985373	2,61....	0,9997767
1,90....	0,9927904	2,26....	0,9986071	2,62....	0,9997888
1,91....	0,9930899	2,27....	0,9986739	2,63....	0,9998003
1,92....	0,9933782	2,28 ...	0,9987377	2,64....	0,9998112
1,93....	0,9936557	2,29....	0,9987986	2,65....	0,9998215
1,94....	0,9939226	2,30....	0,9988568	2,66....	0,9998313
1,95....	0,9941794	2,31....	0,9989124	2,67....	0,9998406
1,96....	0,9944263	2,32....	0,9989655	2,68....	0,9998494
1,97....	0,9946637	2,33 ...	0,9990162	2,69....	0,9998578
1,98....	0,9948920	2,34....	0,9990646	2,70....	0,9998657
1,99....	0,9951114	2,35....	0,9991107	2,71....	0,9998732
2,00....	0,9953223	2,36 ...	0,9991548	2,72 ...	0,9998803
2,01....	0,9955248	2,37 ...	0,9991968	2,73....	0,9998870
2,02....	0,9957195	2,38....	0,9992369	2,74....	0,9998933
2,03....	0,9959063	2,39....	0,9992751	2,75....	0,9998994
2,04....	0,9960858	2,40....	0,9993115	2,76....	0,9999051
2,05....	0,9962581	2,41....	0,9993462	2,77 ...	0,9999105
2,06....	0,9964235	2,42....	0,9993793	2,78....	0,9999156
2,07....	0,9965822	2,43....	0,9994108	2,79 ...	0,9999204
2,08....	0,9967344	2,44....	0,9994408	2,80....	0,9999250
2,09 ...	0,9968805	2,45....	0,9994694	2,81....	0,9999293
2,10....	0,9970205	2,46....	0,9994966	2,82....	0,9999334
2,11....	0,9971548	2,47....	0,9995226	2,83....	0,9999372
2,12....	0,9972836	2,48....	0,9995472	2,84....	0,9999409
2,13....	0,9974070	2,49....	0,9995707	2,85....	0,9999443
2,14....	0,9975253	2,50....	0,9995930	2,86....	0,9999476
2,15....	0,9976386	2,51....	0,9996143	2,87....	0,9999507

λ_1	$\theta(\lambda_1)$	λ_1	$\theta(\lambda_1)$	λ_1	$\theta(\lambda_1)$
1.80	0.9890905	2.16	0.9977472	2.52	0.9996345
1.81	0.9895245	2.17	0.9978511	2.53	0.9996537
1.82	0.9899432	2.18	0.9979506	2.54	0.9996720
1.83	0.9903468	2.19	0.9980459	2.55	0.9996893
1.84	0.9907359	2.20	0.9981372	2.56	0.9997058
1.85	0.9911110	2.21	0.9982244	2.57	0.9997215
1.86	0.9914725	2.22	0.9983079	2.58	0.9997364
1.87	0.9918207	2.23	0.9983878	2.59	0.9997505
1.88	0.9921562	2.24	0.9984642	2.60	0.9997640
1.89	0.9924793	2.25	0.9985373	2.61	0.9997767
1.90	0.9927904	2.26	0.9986071	2.62	0.9997888
1.91	0.9930899	2.27	0.9986739	2.63	0.9998003
1.92	0.9933782	2.28	0.9987377	2.64	0.9998112
1.93	0.9936557	2.29	0.9987986	2.65	0.9998215
1.94	0.9939226	2.30	0.9988568	2.66	0.9998313
1.95	0.9941793	2.31	0.9989124	2.67	0.9998406
1.96	0.9944263	2.32	0.9989655	2.68	0.9998494
1.97	0.9946637	2.33	0.9990162	2.69	0.9998578
1.98	0.9948920	2.34	0.9990646	2.70	0.9998657
1.99	0.9951114	2.35	0.9991107	2.71	0.9998732
2.00	0.9953223	2.36	0.9991548	2.72	0.9998803
2.01	0.9955248	2.37	0.9991968	2.73	0.9998870
2.02	0.9957195	2.38	0.9992369	2.74	0.9998934
2.03	0.9959063	2.39	0.9992751	2.75	0.9998994
2.04	0.9960858	2.40	0.9993115	2.76	0.9999051
2.05	0.9962581	2.41	0.9993462	2.77	0.9999105
2.06	0.9964235	2.42	0.9993793	2.78	0.9999156
2.07	0.9965822	2.43	0.9994108	2.79	0.9999204
2.08	0.9967344	2.44	0.9994408	2.80	0.9999250
2.09	0.9968805	2.45	0.9994694	2.81	0.9999293
2.10	0.9970205	2.46	0.9994966	2.82	0.9999334
2.11	0.9971548	2.47	0.9995226	2.83	0.9999373
2.12	0.9972836	2.48	0.9995472	2.84	0.9999409
2.13	0.9974070	2.49	0.9995707	2.85	0.9999443
2.14	0.9975253	2.50	0.9995930	2.86	0.9999476
2.15	0.9976386	2.51	0.9996143	2.87	0.9999507

λ.	θ.	λ.	θ.	λ.	θ.
2,88...	0,9999536	3,24 ..	0,9999954	3,60...	0,99999964414
2,89...	0,9999563	3,25...	0,9999957	3,61...	0,99999966975
2,90 ..	0,9999589	3,26...	0,9999960	3,62...	0,99999969358
2,91...	0,9999613	3,27...	0,9999962	3,63...	0,99999971574
2,92...	0,9999636	3,28...	0,9999965	3,64...	0,99999973636
2,93...	0,9999658	3,29...	0,9999967	3,65...	0,99999975551
2,94 ..	0,9999679	3,30...	0,9999996	3,66...	0,99999977333
2,95...	0,9999698	3,31...	0,9999971	3,67...	0,99999978990
2,96...	0,9999716	3,32...	0,9999973	3,68...	0,99999980528
2,97...	0,9999733	3,33...	0,9999975	3,69...	0,99999981957
2,98...	0,9999750	3,34...	0,9999977	3,70...	0,99999983285
2,99...	0,9999765	3,35...	0,9999978	3,71 ..	0,99999984517
3,00 ..	0,9999779	3,36...	0,9999980	3,72 ..	0,99999985663
3,01...	0,9999793	3,37...	0,9999981	3,73...	0,99999986726
3,02...	0,9999805	3,38...	0,9999982	3,74...	0,99999987712
3,03...	0,9999817	3,39...	0,9999984	3,75...	0,99999988629
3,04...	0,9999829	3,40 ..	0,9999985	3,76...	0,99999989477
3,05...	0,9999839	3,41...	0,9999986	3,77...	0,99999990265
3,06...	0,9999849	3,42...	0,9999987	3,78...	0,99999990995
3,07...	0,9999859	3,43...	0,9999988	3,79...	0,99999991672
3,08...	0,9999867	3,44...	0,9999989	3,80...	0,99999992200
3,09...	0,9999876	3,45...	0,9999989	3,81...	0,99999992881
3,10...	0,9999884	3,46...	0,99999900780	3,82...	0,99999993421
3,11...	0,9999891	3,47..	0,99999907672	3,83...	0,99999993921
3,12...	0,9999898	3,48...	0,99999914101	3,84...	0,99999994383
3,13...	0,9999904	3,49...	0,99999920097	3,85...	0,99999994812
3,14...	0,9999910	3,50...	0,99999925691	3,86...	0,99999995208
3,15...	0,9999916	3,51 ..	0,99999930905	3,87...	0,99999995575
3,16...	0,9999921	3,52...	0,99999935766	3,88...	0,99999995915
3,17...	0,9999926	3,53 ..	0,99999940296	3,89...	0,99999996230
3,18...	0,9999931	3,54...	0,99999944519	3,90...	0,99999996522
3,19...	0,9999936	3,55...	0,99999948452	3,91...	0,99999996790
3,20 ..	0,9999940	3,56...	0,99999952115	3,92...	0,99999997039
3,21...	0,9999944	3,57 ..	0,99999955527	3,93 ..	0,99999997260
3,22...	0,9999947	3,58...	0,99999958703	3,94 ..	0,99999997482
3,23...	0,9999951	3,59...	0,99999961661	3,95. .	0,99999997678

λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$
2.88	0.9999536		3.24	0.9999954		3.60	0.99999964414
2.89	0.9999563		3.25	0.9999957		3.61	0.99999966975
2.90	0.9999589		3.26	0.9999960		3.62	0.99999969358
2.91	0.9999613		3.27	0.9999962		3.63	0.99999971574
2.92	0.9999636		3.28	0.9999965		3.64	0.99999973635
2.93	0.9999658		3.29	0.9999967		3.65	0.99999975552
2.94	0.9999679		3.30	0.9999969		3.66	0.99999977333
2.95	0.9999698		3.31	0.9999971		3.67	0.99999978989
2.96	0.9999716		3.32	0.9999973		3.68	0.99999980528
2.97	0.9999733		3.33	0.9999975		3.69	0.99999981957
2.98	0.9999750		3.34	0.9999977		3.70	0.99999983285
2.99	0.9999765		3.35	0.9999978		3.71	0.99999984518
3.00	0.9999779		3.36	0.9999980		3.72	0.99999985663
3.01	0.9999793		3.37	0.9999981		3.73	0.99999986726
3.02	0.9999805		3.38	0.9999982		3.74	0.99999987712
3.03	0.9999817		3.39	0.9999984		3.75	0.99999988627
3.04	0.9999829		3.40	0.9999985		3.76	0.99999989476
3.05	0.9999839		3.41	0.9999986		3.77	0.99999990264
3.06	0.9999849		3.42	0.9999987		3.78	0.99999990995
3.07	0.9999859		3.43	0.9999988		3.79	0.99999991672
3.08	0.9999867		3.44	0.9999989		3.80	0.99999992300
3.09	0.9999876		3.45	0.9999989		3.81	0.99999992881
3.10	0.9999884		3.46	0.99999900780		3.82	0.99999993421
3.11	0.9999891		3.47	0.99999907671		3.83	0.99999993920
3.12	0.9999898		3.48	0.99999914100		3.84	0.99999994383
3.13	0.9999904		3.49	0.99999920097		3.85	0.99999994811
3.14	0.9999910		3.50	0.99999925690		3.86	0.99999995208
3.15	0.9999916		3.51	0.99999930905		3.87	0.99999995575
3.16	0.9999921		3.52	0.99999935766		3.88	0.99999995915
3.17	0.9999926		3.53	0.99999940297		3.89	0.99999996230
3.18	0.9999931		3.54	0.99999944518		3.90	0.99999996521
3.19	0.9999936		3.55	0.99999948452		3.91	0.99999996790
3.20	0.9999940		3.56	0.99999952115		3.92	0.99999997039
3.21	0.9999944		3.57	0.99999955527		3.93	0.99999997269
3.22	0.9999947		3.58	0.99999958704		3.94	0.99999997482
3.23	0.9999951		3.59	0.99999961661		3.95	0.99999997678

λ.	θ.	λ.	θ.	λ.	θ.
3,96...	0,99999997860	4,10 ..	0,99999999330	4,50..	0,99999999981
3,97...	0,99999998028	4,20...	0,99999999714	4,60...	0,99999999992
3,98 ..	0,99999998183	4,30 .	0,99999999880	4,70...	0,99999999997
3,99...	0,99999998327	4,40 .	0,99999999951	4,80...	0,99999999999
4,00...	0,99999998459				

Proposição XIV

A um afastamento absoluto

$$k = \lambda\sqrt{2mpq}$$

corresponde para $\frac{(p)}{(q)}$ uma expressão da forma

$$\frac{(p)}{(q)} = \frac{mp + \lambda\sqrt{2mpq}}{mq - \lambda\sqrt{2mpq}},$$

onde supomos o afastamento no sentído do acontecimento da probabilidade p; donde,

$$\frac{(p)}{(q)} - \frac{p}{q} = \frac{mp + \lambda\sqrt{2mpq}}{mq - \lambda\sqrt{2mpq}} - \frac{p}{q}$$

$$= \frac{\lambda\sqrt{2mpq}}{mq^2 - \lambda q\sqrt{2mpq}} = \frac{\lambda\sqrt{2pq}}{q^2\sqrt{m} - \lambda q\sqrt{2pq}}.$$

Se o afastametto fosse considerado no sentido do q, não

λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$		λ_1	$\theta(\lambda_1)$
3.96	0.99999997860		4.10	0.99999999330		4.60	0.99999999992
3.97	0.99999998028		4.20	0.99999999714		4.70	0.99999999997
3.98	0.99999998183		4.30	0.99999999881		4.80	0.99999999999
3.99	0.99999998326		4.40	0.99999999951		4.90	1.00000000000
4.00	0.99999998458		4.50	0.99999999980		5.00	1.00000000000

Proposition XIV

The expression of $\frac{(p)}{(q)}$ that corresponds to an absolute deviation

$$k = \lambda_1 \sqrt{2mpq}$$

is

$$\frac{(p)}{(q)} = \frac{mp + \lambda_1\sqrt{2mpq}}{mq - \lambda_1\sqrt{2mpq}}\,.$$

where we assume, without loss of generality, that the deviation is the consequence of an excess of successes; therefore

$$\frac{(p)}{(q)} - \frac{p}{q} = \frac{mp + \lambda_1\sqrt{2mpq}}{mq - \lambda_1\sqrt{2mpq}} - \frac{p}{q} =$$

$$\frac{\lambda_1\sqrt{2mpq}}{mq^2 - \lambda_1 q\sqrt{2mpq}} = \frac{\lambda_1\sqrt{2pq}}{q^2\sqrt{m} - \lambda_1 q\sqrt{2pq}}\,.$$

In case the deviation would be a consequence of an excess of failures, tinhamos mais do que mudar o sinal ao λ. Em qualquer caso teremos que

$$\left|\frac{(p)}{(q)}-\frac{p}{q}\right|=\left|\frac{\lambda\sqrt{2pq}}{q^2\sqrt{m}-\lambda q\sqrt{2pq}}\right|>\varepsilon,$$

se

$$\left|\lambda\right|>\frac{\varepsilon q^2\sqrt{m}}{\sqrt{2pq}\,(1\pm\varepsilon q)} \tag{1}$$

A probabilidade de que (1) seja satisfeita será (prop. XII)

$$P=1-\theta\left[\frac{\varepsilon q^2\sqrt{m}}{\sqrt{2pq}\,(1\pm\varepsilon q)}\right].$$

Por mais pequeno que ε seja, P tenderá rapidamente para zero, por causa do factor $\sqrt{m}$.

Exemplo:

Jogam-se 200 jogos de *cara* ou X, a tostão cada jogo. Qual a probabilidade de ganhar ou perder uma quantia superior a 10 tostões?

Resp.:

Neste caso é

$$p=q=\frac{1}{2}:\quad m=200;\quad k>10;$$

logo, se

$$k=\lambda\sqrt{2\,.\,200\,.\,\frac{1}{2}\,.\,\frac{1}{2}}=10\,.\,\lambda>10,$$

the only alteration would be to change the sign of λ_1 [in the numerator]. In any case, we have that

$$\left| \frac{(p)}{(q)} - \frac{p}{q} \right| = \left| \frac{\lambda_1 \sqrt{2pq}}{q^2 \sqrt{m} - \lambda_1 q \sqrt{2pq}} \right| > \varepsilon$$

if

$$|\, \lambda_1 \,| > \frac{\varepsilon\, q^2 \sqrt{m}}{\sqrt{2pq}\,(1 \pm \varepsilon\, q)} \tag{5.5}$$

The probability that (5.5) holds is (Prop. XII)

$$\mathbb{P} = 1 - \theta \left(\frac{\varepsilon\, q^2 \sqrt{m}}{\sqrt{2pq}\,(1 \pm \varepsilon\, q)} \right) .$$

Even for very small values of ε, $\mathbb{P}$ decreases very quickly towards 0, because of the factor $\sqrt{m}$.

Example:

What is the probability of winning or loosing more than 10 cents in a sequence of 200 bets on the result of coin throwing, when the money at stake in each trial is 1 cent?

Assuming that

$$p = q = \frac{1}{2}; \quad m = 200; \quad k > 10;$$

therefore, if

$$k = \lambda_1 \sqrt{2 \cdot 200 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 10 \cdot \lambda_1 ,$$

será

$$\lambda > 1 \quad \text{e} \quad 1 - \theta(1) = \frac{16}{100} = \frac{4}{25},$$

aproximadamente.

Proposição XV

BOREL [1] generalizou a lei dos desvios

$$\theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_0^\lambda e^{-\lambda^2} d\lambda$$

para o caso das tiragens serem feitas em urnas de composições diferentes.

Assim, sejam

$$p_1 \text{ e } q_1, \quad p_2 \text{ e } q_2, \quad \ldots \quad p_n \text{ e } q_n$$

as composições respectivas de n urnas contendo bolas brancas e bolas pretas.

Façamos m_1 tiragens na primeira urna, m_2 na segunda... m_n na última. O número mais provável de bolas brancas será

$$m_1 p_1 + m_2 p_2 + \ldots + m_n p_n .$$

Mas, em geral, não será êste número, a que poderiamos

---

(1) E. BOREL, *Éléments de la Théorie des Probabilités*, deuxième édition, pág. 77.

we get

$$\lambda_1 > 1 \quad \text{and} \quad 1 - \theta(1) = \frac{16}{100} = \frac{4}{25},$$

approximately.

Proposition XV

Borel[20] proved a generalization of the law of deviations

$$\theta(\lambda_1) = \frac{2}{\sqrt{\pi}} \int_0^{\lambda_1} e^{-\lambda^2} \, d\lambda$$

assuming that the extractions are done from urns with different compositions.

Let

$$p_1 \text{ and } q_1, \quad p_2 \text{ and } q_2, \quad \ldots \quad p_n \text{ and } q_n$$

be the compositions of n urns with white and black balls.

Assume that we extract m_1 balls from the first urn, m_2 balls from the second urn, …, m_n balls from the last urn. The most probable number of white balls in the lot will be

$$m_1 p_1 + m_2 p_2 + \cdots + m_n p_n .$$

But in general this number of white balls, which we could

[20] E. BOREL, *Eléments de la Théorie des Probabilités*, deuxième édition, p. 77.

chamar *normal*, o obtido. Haverá, em geral, um afastamento absoluto h, soma algebrica de n afastamentos, h_1, h_2, ..., h_n, provenientes das urnas respectivas.

Ora Borel demonstrou, dum modo muito fácil aliás, que a lei dos desvios ainda terá a mesma forma, caso se tome para unidade do afastamento h a raiz quadrada da soma dos quadrados dos afastamentos correspondentes às diversas urnas.

call *normal*, is not the number of white balls that we fetch in a real experiment. We shall in general get an absolute deviation h, which is the addition of deviations $h_1, h_2, \ldots, h_n$ of the white balls extracted from each urn when compared to the corresponding normal number.

Borel has proved, and using very simple arguments, that the law of deviations still holds true, provided we use as unit deviation h the square root of the sum of squares of the deviations corresponding to the different urns.

CAPÍTULO VI

ESPERANÇA MATEMÁTICA E VALOR MÉDIO

CHAPTER VI

MATHEMATICAL EXPECTATION

AND

MEAN VALUE

CAPITULO VI

ESPERANÇA MATEMÁTICA E VALOR MÉDIO

DEFINIÇÃO 1.ª

Seja (A) uma classe possível contendo a classe (A′). A cada elemento de (A′) façamos corresponder um número. Obtemos assim uma função que representaremos por f(A′). Multiplicando a probabilidade de cada elemento de (A′) pelo valor correspondente da função f(A′) e somando, obtemos um número que é de costume chamar-se *esperança matemática da classe* (A′), *relativa à função* f(A′).

Supondo que a probabilidade dos elementos de (A′) é relativa à classe (A), à esperança matemática da classe (A′) quando ela coìncide com (A), chama-se *valor médio* ou *valor provável* da função f(A) [1].

Representaremos a esperança matemática com o símbolo

$$E\,[f(A)]$$

---

(1) Embora esta distinção entre esperança matemática e valor médio não costume vir explicitamente feita nos livros de Probabilidades, todos os autores dão a êstes termos a significação que acabamos de atribuir-lhes.

CHAPTER VI

MATHEMATICAL EXPECTATION AND MEAN VALUE

DEFINITION 1

Let A' be a subset of the possible class A. We associate to each element $a \in A'$ one value, obtaining therefore one function that we shall denote f. The sum of products of the probability of each $a \in A'$ by $f(a)$ is *the mathematical expectation of the class* A'*, in what concerns the function* f*.*

Assuming that the probability of the elements of A' is relative to A, the mathematical expectation of the class $A' = A$ is called the *mean value* or *probable value* of the function f. [21]

We denote the mathematical expectation of A by

$$\mathbb{E}_{A}\left[f\right]$$

[21] Although this distinction between mathematical expectation and mean value isn't explicitly stated in most Probability books, all authors attach to these terms the concepts we state.

e o valor médio ou provável com o símbolo

$$M\,[f(A)].$$

Assim, se a cada face dum dado fizermos corresponder o respectivo número de pontos, construimos uma função cujo valor médio é

$$M = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \ldots + \frac{1}{6} \cdot 6 = 3{,}5.$$

A esperança matemática relativa às faces 1, 2, será

$$E = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 = 0{,}5.$$

Proposição I

A esperança matemática duma classe é igual à soma das esperanças matemáticas das suas partes, como resulta imediatamente da def. 1.ª

Proposição II

É igualmente evidente que *a esperança matemàtica de uma soma é igual à soma das esperanças matemáticas das parcelas.*

and the mean or probable value

$$\mathbb{M}_A [f] .$$

As an example, in dice throwing, consider the function that associates to each face the number of dots in it. The mean value of that function is

$$\mathbb{M} = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3.5.$$

The mathematical expectation relative to the faces 1 and 2 is

$$\mathbb{E} = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 = 0.5.$$

Proposition I

The mathematical expectation of one class is the sum of the mathematical expectations of its parts, an immediate consequence of Def. 1.

Proposition II

It is also obvious that *the mathematical expectation of a sum of functions is the sum of the mathematical expectations of the summands.*

Proposição III

Sejam (A) e (B) duas classes possíveis e (A, B) a classe composta das duas e suponhamos que fazemos

$$f(A, B) = f(A) \cdot f(B).$$

Nestas condições teremos

$$E[f(A, B)] = E[f(A)] \cdot E[f(B)],$$

isto é, *a esperança matemática da classe composta é igual ao produto das esperanças matemáticas das classes componentes.*

Com efeito, representando por P_x a probabilidade do elemento x, teremos

$$E[f(A)] = \Sigma f(A) P_A ; \quad E[f(B)] = \Sigma f(B) P_B ;$$
$$E[f(A, B)] = \Sigma f(A, B) P_{AB} ;$$

ora

$$\begin{aligned} \Sigma f(A, B) P_{AB} &= \Sigma f(A) \cdot f(B) \cdot P_A \cdot P_B \\ &= \Sigma f(A) \cdot P_A \cdot \Sigma f(B) \cdot P_B ; \end{aligned}$$

logo

$$E[f(A, B)] = E[f(A)] \cdot E[f(B)],$$

c. d. d.

Proposition III

Let A and B be two possible classes, and $A \times B$ the compound class of the two. Define

$$f(a,b) = f_1(a) \cdot f_2(b).$$

Under these assumptions, we get

$$\mathbb{E}_{A\times B}[f] = \mathbb{E}_A[f_1] \cdot \mathbb{E}_B[f_2],$$

i.e., *the mathematical expectation of the compound class is the product of the expectations of its components.*

In effect, denoting $\mathbb{P}_x$ the probability of x, we have

$$\mathbb{E}_A[f_1] = \sum_{a\in A} f_1(a)\,\mathbb{P}_a, \qquad \mathbb{E}_B[f_2] = \sum_{b\in B} f_2(b)\,\mathbb{P}_b$$

$$\mathbb{E}_{A\times B}[f] = \sum_{(a,b)\in A\times B} f(a,b)\,\mathbb{P}_{(a,b)}$$

As

$$\sum_{(a,b)\in A\times B} f(a,b)\,\mathbb{P}_{(a,b)} = \sum_{(a,b)\in A\times B} f_1(a)\cdot f_2(b)\,\mathbb{P}_a\,\mathbb{P}_b =$$

$$= \sum_{a\in A} f_1(a)\,\mathbb{P}_a \cdot \sum_{b\in B} f_2(b)\,\mathbb{P}_b$$

we conclude that

$$\mathbb{E}_{A\times B}[f] = \mathbb{E}_A[f_1] \cdot \mathbb{E}_B[f_2].$$

□

DEFINIÇÃO 2.ª

Seja M um ponto livre ou sujeito, variando numa região que contêm a região (A) e representemos por P (M) a sua lei de probabilidade em relação a (A). Seja, alêm disso, φ (M) uma função das coordenadas de M, definida em (A). Sendo (A′) uma região contida em (A), ao número

$$E_{(A')}[\varphi(M)] = \int_{(A')} P(M)\,\varphi(M)\,d(A)$$

chamaremos *esperança matemática da região* (A′) *relativa à função* φ (M). No caso particular de (A′) coincidir com (A), chamaremos a êsse número *valor médio* ou *valor provável* de φ (M).

Proposição IV

É evidente que, se

$$(A') = (A_1) + (A_2) + \dots + (A_n),$$

será

$$E(A') = E(A_1) + E(A_2) + \dots + E(A_n)$$

Proposição V

É igualmente evidente que

$$E[\varphi_1(M) + \varphi_2(M) + \dots] = E[\varphi_1(M)] + E[\varphi_2(M)] + \dots$$

DEFINITION 2

Let X denote a free point or an image point varying in some region containing the possible region A; and $\mathbb{P}_X$ denote its probability law relative to A. Consider a function $\varphi(X)$, defined in A, of the coordinates of point X. If $A' \subset A$, the *mathematical expectation of the region A' relative to the function $\varphi(X)$* is

$$\mathbb{E}_{A'}\left[\varphi(X)\right] = \int\limits_{A'} \mathbb{P}_X \, \varphi(X) \, \mathrm{d}a.$$

In the particular case $A' = A$, that number is the *mean value* or *probable value* of $\varphi(X)$.

Proposition IV

It is evident that if A' can be partitioned in pairwise disjoint sets

$$A' = A_1 \cup A_2 \cup \cdots \cup A_n$$

then

$$\mathbb{E}_{A'}\left[X\right] = \mathbb{E}_{A_1}\left[X\right] + \mathbb{E}_{A_2}\left[X\right] + \cdots + \mathbb{E}_{A_n}\left[X\right].$$

Proposition V

It is also obvious that

$$\mathbb{E}\left[\varphi_1(X) + \varphi_2(X) + \cdots\right] = \mathbb{E}\left[\varphi_1(X)\right] + \mathbb{E}\left[\varphi_2(X)\right] + \cdots.$$

Proposição IV

Se M fôr um ponto variando em (A) e N outro variando em (B), será, análogamente ao caso da prop. III,

$$E_{(A,\,B)}[\varphi(M)\,.\,\psi(N)] = \int_{(A,\,B)} P(M)\,P(N)\,\varphi(M)\,\psi(N)\,d(A,B)$$

$$= \int_{(A)} P(M)\,\varphi(M)\,d(A) \int_{(B)} P(N)\,\psi(N)\,d(B)$$

$$= E_{(A)}[\varphi(M)]\,.\,E_{(B)}[\psi(N)].$$

Proposição VII

Seja M um ponto variando numa certa região que contêm (A), $f(M)$ uma dada função das suas coordenadas, e $P(M)$ a sua lei de probabilidade relativamente a (A). Já vimos que, por definição, é

$$E_{(A)}[f(M)] = \int_{(A)} P(M)\,f(M)\,d(A).$$

Ora, pondo $f(M) = z$, teremos

$$E = \int_{(A')} z\,P(M)\,d(A) = \int_{z_0}^{z} z \int P(M)\,d(A)$$

sendo o segundo integral estendido à região de (A) em

Proposition VI

Let X denote a varying point in A, Y a varying point in B. Similarly to what we stated in Prop. III [22],

$$\mathbb{E}_{A\times B}\left[\varphi(X)\cdot\psi(Y)\right] = \int\limits_{A\times B} \mathbb{P}_X\,\mathbb{P}_Y\,\varphi(X)\,\psi(Y)\,\mathrm{d}(a,b) =$$

$$= \int\limits_A \mathbb{P}_X\,\varphi(X)\,\mathrm{d}a\cdot\int\limits_B \mathbb{P}_Y\,\psi(Y)\,\mathrm{d}b =$$

$$= \mathbb{E}_A\left[\varphi(X)\right]\cdot\mathbb{E}_B\left[\psi(Y)\right].$$

Proposition VII

Let X denote a varying point in some region containing A, φ denote some function of its coordinates, and $\mathbb{P}_X$ the probability law of X relative to A.

We have defined

$$\mathbb{E}_A\left[\varphi(X)\right] = \int\limits_A \mathbb{P}_X\,\varphi(X)\,\mathrm{d}a.$$

Writing $\varphi(X) = Z$, we get

$$\mathbb{E} = \int\limits_A Z\,\mathbb{P}_X\,\mathrm{d}a = \int\limits_{z_0}^{z_1} Z\int \mathbb{P}_X\,\mathrm{d}a$$

where the second integral is to be computed for the values of A for which

[22] *Editors' note*: Perhaps this reference to Proposition III is to include the independence hypotheses in this Proposition, because otherwise it would not be correct.

que z toma valores compreendidos entre z e $z+dz$. Mas êste integral é, por definição de P (M), a probabilidade de que z esteja compreendido entre z e $z+dz$ e o seu valor pode representar-se por

$$P(z)\,dz,$$

sendo $P(z)$ a lei da probabilidade de z. Logo,

$$E_{(A)}[\varphi(M)] = \int_{z_0}^{z} z\,P(z)\,dz = E_{(z)}(z).$$

Por um exemplo que adiante daremos, se verá a utilidade desta proposição na determinação das esperanças matemáticas.

Proposição VIII

Sendo dada a lei da probabilidade da variável z, *a esperança matemática de qualquer função de* z, $\varphi(z)$, *será dada por*

$$E_{(z)}[\varphi(z)] = \int_{z_0}^{z} \varphi(z)\,P(z)\,dz.$$

Para provar isto, bastaria fazer

$$f(M) = \varphi(z)$$

na proposição antecedente e tomar ainda o segundo inte-

Z is between z and $z + \mathrm{d}z$. But this integral is, by definition, $\mathbb{P}_Z$, the probability that Z lies between z and $z+\mathrm{d}z$, and its value may be denoted

$$\mathbb{P}_Z(z)\ \mathrm{d}z,$$

where $\mathbb{P}_Z(z)$ is the probability law of Z. Therefore,

$$\mathbb{E}_A\left[\varphi(X)\right] = \int_{z_0}^{z_1} Z\,\mathbb{P}_Z(z)\ \mathrm{d}z = \mathbb{E}_Z(Z).$$

Later on, we shall present an example showing this proposition usefulness in the computation of mathematical expectations.

Proposition VIII

Given the probability law of the variable Z, *the mathematical expectation of any function* $\varphi(Z)$ *of* Z *is*

$$\mathbb{E}_Z\left[\varphi(Z)\right] = \int_{z_0}^{z_1} \varphi(z)\,\mathbb{P}_Z(z)\mathrm{d}z.$$

To prove it let

$$f(X) = \varphi(Z)$$

in the previous proposition, computing the second integral

gral estendido à região em que z *toma valores compreendidos entre* z e $z + dz$.

Proposição IX

A esperança matemática duma constante é igual à própria constante multiplicada pela probabilidade da região isto é

$$E_{(A)}(C) = C \, . \, P_{(A)}$$

como resulta imediatamente da definição.

Para o valor médio será

$$P_{(A)} = 1 \quad \text{e} \quad M(C) = C.$$

Proposição X

Se uma função $f(M)$ *é sempre positiva e o seu valor médio se pode tornar inferior a qualquer número* δ, *por mais pequeno que* δ *seja, a probabilidade de que* $f(M)$ *se mantenha superior a certo numero* m, *numa região* (A'), *é menor do que* $\dfrac{\delta}{m}$.

Com efeito, se, ao longo de (A'), é

$$f(M) \geqslant m,$$

in the region where Z takes values between z and $z + dz$.

Proposition IX

The mathematical expectation of a region A', in what concerns a constant function c, is the product of that constant by the probability of the region, i.e.,

$$\mathbb{E}_{A'}(c) = c \cdot \mathbb{P}(A'),$$

a result that follows immediately from the definition.

In particular, the mean value of a constant is that constant, since

$$\mathbb{P}(A) = 1 \quad \Longrightarrow \quad \mathbb{M}(c) = c.$$

Proposition X

If $f(X)$ is a positive function and its mean value is smaller than a number δ, however small this δ may be, the probability, that $f(X)$ is greater than a given m, is lower than $\dfrac{\delta}{m}$.

In effect, if A' is the region where

$$f(X) \geq m,$$

segue-se que

$$\int_{(A)} f(M)\,P(M)\,d(A) = \int_{(A-A')} f(M)P(M)d(A) + \int_{(A')} f(M)P(M)d(A)$$

$$\geqslant \int_{(A')} f(M)\,P(M)\,d(A)$$

$$\geqslant m\,.\,P_{(A')}\,;$$

logo

$$m\,.\,P_{(A')} \leqslant \delta$$

e

$$P_{(A')} \leqslant \frac{\delta}{m}$$

c. d. d.

Problema

Consideremos um polígono articulado aberto, de lados

$$l_1, \quad l_2, \quad \ldots, \quad l_n\,,$$

e sejam A e B os seus pontos extremos. Lança-se, à sorte, esse polígono sôbre um plnao: determinar

$$M_n\,(d^2),$$

sendo

$$d = A\,B.$$

it follows that

$$\int\limits_A f(X)\,\mathbb{P}_X \mathrm{d}a = \int\limits_{A-A'} f(X)\,\mathbb{P}_X \mathrm{d}a + \int\limits_{A'} f(X)\,\mathbb{P}_X \mathrm{d}a \geq$$

$$\geq \int\limits_{A'} f(X)\,\mathbb{P}_X \mathrm{d}a \geq$$

$$\geq m \cdot \mathbb{P}(A');$$

therefore

$$m \cdot \mathbb{P}(A') \leq \delta$$

and

$$\mathbb{P}(A') \leq \frac{\delta}{m}.$$

□

Problem

Consider a random plane open polygonal line with sides of lengths

$$l_1,\, l_2,\, \ldots,\, l_n,$$

denote A and B its endpoints, and d the length of $\overline{AB}$.

Compute

$$\mathbb{M}_n(d^2).$$

Solução

No cálculo de $M_n(d^2)$ nós podemos (prop. VIII) entrar com a lei da probabilidade de d. Além disso, à lei da probabilidade de d, nós podemos substituir a lei de probabilidade de qualquer ponto M, a d convenientemente ligado (Prop. VII); o ponto equivalente do polígono, por exemplo. Posto isto, consideremos o caso do polígono ter um lado só, l_1; será (prop. IX)

$$M_1(d^2) = M_1(l_1^2) = l_1^2 .$$

Consideremos agora o caso do polígono ter dois lados; teremos (prop. VIII)

$$M_2(d^2) = \int_0^{l_1+l_2} P(d) \,.\, d^2 \,.\, d(d) ;$$

$$= \int_0^{\pi} \frac{d\alpha}{\pi} (l_1^2 + l_2^2 - 2\, l_1\, l_2 \cos \alpha)$$

$$= l_1^2 + l_2^2 ,$$

sendo α o ângulo de l_1 com l_2.

Suponhamos que no caso de i lados ainda tinhamos

$$M_i(d^2) = l_1^2 + l_2^2 + \ldots + l_i^2$$

Solution

To compute $\mathbb{M}_n(d^2)$ we can use the probability law of d (Prop. VIII). But, by Prop. VII, instead of the probability law of d we may use the probability law of any convenient point X tied to it, for instance the equivalent point to the polygonal line. Consider first the simplest case of a polygonal line with only one side of length l_1; from Prop. IX, it follows that

$$\mathbb{M}_1(d^2) = l_1^2.$$

Let us now consider the case of a random polygonal line with two sides; from Prop. VIII it follows that

$$\mathbb{M}_2(d^2) = \int_0^{l_1+l_2} \mathbb{P}(d)\, d^2 \,\mathrm{d}d =$$

$$= \int_0^{\pi} \frac{1}{\pi}\,(l_1^2 + l_2^2 - 2\,l_1\,l_2 \cos\alpha)\,\mathrm{d}\alpha =$$

$$= l_1^2 + l_2^2,$$

where in the above computation α denotes the angle between the two sides of the polygonal line.

Let us now assume the induction hypothesis that in the case of a polygonal line with i sides we have

$$\mathbb{M}_i(d^2) = l_1^2 + l_2^2 + \cdots + l_i^2$$

e provemos que nesse caso ainda

$$M_{i+1}\,(d^2)$$

nos é dado pela mesma lei.

Representando por δ o segmento que une a origem de l_1 com a extremidade de l_i, teremos, por hipótese,

$$M_i(d^2) = l_1^2 + l_2^2 + \ldots + l_i^2.$$

Ora, representando por $\alpha_1, \alpha_2, \ldots \alpha_i$, os ângulos das articulações, d^2 será uma função $f(\alpha_1, \alpha_2, \ldots \alpha_i)$ dêsses ângulos e o valor médio procurado será de forma

$$M_{i+1} = \int f(\alpha_1, \alpha_2, \ldots \alpha_i)\frac{d\alpha_1}{2\pi}\cdot\frac{d\alpha_2}{2\pi}\cdots\frac{d\alpha_i}{2\pi}$$

$$= \int\frac{d\alpha_1}{2\pi}\cdot\frac{d\alpha_2}{2\pi}\cdots\frac{d\alpha_{i-1}}{2\pi}\int_0^{2\pi} f(\alpha_1, \alpha_2, \ldots \alpha_i)\,\frac{d\alpha_i}{2\pi}$$

$$= \int\frac{d\alpha_1}{2\pi}\cdot\frac{d\alpha_2}{2\pi}\cdots\frac{d\alpha_{i-1}}{2\pi}\int_0^{2\pi}\frac{1}{2\pi}(\delta^2 + l^2{}_{i-1} + 2\delta l_{i+1}\cos\alpha_i)\,d\alpha_i$$

$$= \int\frac{d\alpha_1}{2\pi}\cdot\frac{d\alpha_2}{2\pi}\cdots\frac{d\alpha_{i-1}}{2\pi}\,(\delta^2 + l^2{}_{i+1})$$

$$= M_i\,(\delta^2 + l_{i+1}{}^2)$$

$$= M_i\,(\delta^2) + M_i\,(l^2{}_{i+1})$$

$$= l_1^2 + l_2^2 + \ldots + l_i^2 + l^2{}_{i+1},$$

c. d. d.

in order to prove that we also get

$$\mathbb{M}_{i+1}(d^2) = l_1^2 + l_2^2 + \cdots + l_i^2 + l_{i+1}^2.$$

Denoting δ the length of the segment from the origin of l_1 with the endpoint of l_i, from the induction hypothesis

$$\mathbb{M}_i(\delta^2) = l_1^2 + l_2^2 + \cdots + l_i^2.$$

Denoting $\alpha_1, \alpha_2, \ldots, \alpha_i$ the angles in the articulations of the polygonal line, d^2 is a function $f(\alpha_1, \alpha_2, \ldots, \alpha_i)$ of those angles, and the mean value we want to compute is of the form

$$\mathbb{M}_{i+1}(d^2) = \int f(\alpha_1, \alpha_2, \ldots, \alpha_i)\frac{\mathrm{d}\alpha_1}{2\pi} \cdot \frac{\mathrm{d}\alpha_2}{2\pi} \cdots \frac{\mathrm{d}\alpha_i}{2\pi} =$$

$$= \int \frac{\mathrm{d}\alpha_1}{2\pi} \cdot \frac{\mathrm{d}\alpha_2}{2\pi} \cdots \frac{\mathrm{d}\alpha_{i-1}}{2\pi} \int_0^{2\pi} f(\alpha_1, \alpha_2, \ldots, \alpha_i)\frac{\mathrm{d}\alpha_i}{2\pi} =$$

$$= \int \frac{\mathrm{d}\alpha_1}{2\pi} \cdot \frac{\mathrm{d}\alpha_2}{2\pi} \cdots \frac{\mathrm{d}\alpha_{i-1}}{2\pi} \int_0^{2\pi} \frac{1}{2\pi}(\delta^2 + l_{i+1}^2 - 2\,\delta l_{i+1}\cos\alpha_i)\,\mathrm{d}\alpha_i =$$

$$= \int \frac{\mathrm{d}\alpha_1}{2\pi} \cdot \frac{\mathrm{d}\alpha_2}{2\pi} \cdots \frac{\mathrm{d}\alpha_{i-1}}{2\pi}(\delta^2 + l_{i+1}^2) =$$

$$= \mathbb{M}_i(\delta^2 + l_{i+1}^2) =$$

$$= \mathbb{M}_i(\delta^2) + \mathbb{M}_i(l_{i+1}^2) =$$

$$= l_1^2 + l_2^2 + \cdots + l_i^2 + l_{i+1}^2.$$

□

Logo,

$$M_n(d^2) = l_1^2 + l_2^2 + \ldots + l_n^2.$$

Se os lados forem todos iguais e representarmos o perímetro por L, será

$$M_n = n\, l_1^2 = \frac{L^2}{n}.$$

Daqui resulta, para $n = \infty$, que

O valor provável do quadrado da distância que separa os pontos extremos duma curva flexível, lançada á sorte sôbre um plano, é nulo, qualquer que seja o comprimento da curva, logo que seja finito.

Proposição XI

A esperança matemática pode, às vezes, calcular-se sem a determinação prévia das parcelas que a constituem ou sem a determinação da lei da probabilidade. Determina-se a soma duma só vez. Vamos dar disso um exemplo curioso servindo-nos do problema da agulha. Como vimos na nota da página 64, dividindo a agulha em partes iguais, cada parte ficava com probabilidades iguais. Se as partes em vez de se conservarem topo a topo, e sôbre a mesma recta, tomarem outra posição relativa, a sua probabilidade ainda será a mesma, como fácilmente se pode vêr; se essas partes, em vez de se conservarem invariavelmente ligadas, formarem um sistema articulado, a probabilidade de cada parte a mesma será ainda, como resulta da defi-

From this it follows that

$$\mathbb{M}_n(d^2) = l_1^2 + l_2^2 + \cdots + l_n^2.$$

If the sides of the random polygonal line are all of equal length, denoting L its perimeter we have

$$\mathbb{M}_n(d^2) = n\, l_1^2 = \frac{L^2}{n}\,.$$

From this, letting $n \to \infty$, we get that:

The probable value of the square of the distance between the endpoints of a random plane flexible curve is zero, whatever the length of the curve, provided this is finite.

Proposition XI

The mathematical expectation can be computed, in some cases, without previous computation of the summands involved in its definition, or of the probability law. We exemplify using a curious example, an alternative way of solving Buffon's needle problem. As remarked in the observation in page 64, dividing the needle into equal parts, each of those would have equal probability of intersection one of the separation lines. If those parts, instead of being collinear, have different relative positions, forming a polygonal line, each part will still have the same probability, an immediate consequence nição de lançamento, à sorte, duma figura de forma variável.

Suponhamos, agora, que a cada encontro de qualquer lado do sistema que estamos considerando com uma das paralelas, fazemos corresponder um mesmo número, a unidade. A esperança matemática de cada parte será igual à sua probabilidade. A soma de todas estas esperanças, será proporcional ao número dessas partes que supomos de igual comprimento e, por isso, proporcional ao perímetro do sistema. Isto qualquer que seja o sistema, rígido ou articulado, qualquer que seja o comprimento dos seus lados. Passando para o limite, podemos ainda dizer que o integral das esperanças elementares duma curva flexível ou rígida é proporcional ao seu comprimento:

$$E(l) = K\,l,$$

sendo K independente da forma, natureza e perímetro da figura considerada.

Para determinar K, consideremos uma circunferência de diâmetro igual à distância de duas paralelas consecutivas.

Lançando, à sorte, esta circunferência sôbre o plano das paralelas, ela encontrará sempre uma paralela e uma só, em dois pontos; logo

$$E(\pi a) = K \pi a = 2$$

e

$$K = \frac{2}{\pi a}.$$

of the definition of randomly throwing a variable form figure.

Let us assume, now, that we associate the same number, one, to each intersection of one side of the polygonal line with one of the parallel lines. The mathematical expectation of each part will be the probability that it hits one parallel. The sum of all these expectations is proportional to the number of parts, and thus proportional to the perimeter of the polygonal line. And this is so, whatever the polygonal line, rigid or articulated, and whatever the length of its sides. In the limit, we can still say that the integral of the elementary expectations of a rigid or flexible curve is proportional to its length:

$$\mathbb{E}(l) = K\,l,$$

where K is independent of the form, nature and perimeter of the figure.

To determine K, let us consider a needle which is a circumference whose diameter is the distance between two consecutive parallels.

Randomly throwing this circumference in the plane of the parallels, it will always have two points in common with the system of parallels, [either] because it intersects one parallel in two points, [or because it is tangent to two consecutive parallels][(23)]; therefore

$$\mathbb{E}(\pi\,a) = K\,\pi\,a = 2$$

and

$$K = \frac{2}{\pi\,a}\,.$$

(23) *Editors' note*: Pacheco d'Amorim forgets the second possibility.

O valor achado para K dá-nos

$$E(l) = \frac{2l}{\pi a}.$$

Êste valor coíncide com a probabilidade de encontro achado, no problema da agulha, para o caso de $l \leqslant a$. E assim deve ser, porque, caso $l \leqslant a$, a agulha só pode ter um encontro ou nenhum e por isso a sua esperança matemática coìncidirá com a probabilidade. Podiamos aproveitar este facto para resolver o problema da agulha neste caso.

*

Se as rectas paralelas fossem substituidas por círculos concêntricos e equidistantes, a esperança matemática seria a mesma, mas nada podiamos concluir acêrca da probabilidade do encontro dum segmento rectílinio com as circunferências dêsses círculos.

Proposição XII

Consideremos um fenómeno de duas modalidades e sejam p e q as suas respectivas probabilidades; à modalidade de probabilidade p façamos corresponder o número a e à modalidade q, o número b.

O valor médio da função assim construida, será

$$M = ap + bq.$$

Provoquemos o fenómeno em questão um grande nú-

From that we get

$$\mathbb{E}(l) = \frac{2\,l}{\pi\,a}\,.$$

This is the value we had obtained for the probability that the needle intersects one of the parallels, when solving Buffon's needle problem, when $l \leq a$. In fact, when $l \leq a$, the needle either intersects one of the parallels in one point, or it doesn't, and therefore the mathematical expectation is the probability of the event that it intersects one of the parallels. So, another way of solving Buffon's needle problem is via the exploitation of the concept of mathematical expectation.

*

If the parallel lines are substituted by equidistant circumferences with the same center, the mathematical expectation would still be the same, but we couldn't state anything about the probability of intersection of a linear segment with one of the circumferences.

Proposition XII

Let us consider an experiment with two possible outcomes, [success or failure,] and denote p and q the corresponding probabilities. Consider the function $f(\text{success}) = a$, $f(\text{failure}) = b$.

The mean value of this function is

$$\mathbb{M} = ap + bq.$$

Repeat the experiment a large number of times, mero de vezes e suponhamos que a modalidade p se repete (p) vezes e a modalidade q, (q) vezes.

Seja

$$M' = \frac{(p)a + (q)b}{(p) + (q)}$$

a média arimética dos valores obtidos para a dita função.

Consideremos a expressão

$$\left| M - M' \right| = \left| ap + bq - \frac{a(p) + b(q)}{(p) + (q)} \right|$$

$$\leqslant | a | \cdot \left| p - \frac{(p)}{(p) + (q)} \right| + | b | \cdot \left| q - \frac{(q)}{(p) + (q)} \right|;$$

ela tende para *zero*, à medida que $(p) + (q)$ aumenta, ou melhor, a probabilidade de que $| M - M' |$ se mantenha inferior a ε, por menor que ε seja, tendera para *zero*, à medida que $(p) + (q)$ tenda para o infinito (3.° teorema de BERNOULLI).

*

O que acaba de dizer-se dêsse fenómeno de duas modalidades diz-se dum fenómeno dum número qualquer de modalidades.

*

O que se disse do valor médio, diz-se igualmente da esperança matemática de qualquer classe. Daí a seguinte proposição:

A esperança matemática duma classe finita de elementos and let's assume that (p) times we get success and (q) times we get failure.

Denote

$$\overline{x} = \frac{(p)\,a + (q)\,b}{(p) + (q)}$$

the arithmetic mean of the observed values of the function we have defined.

The expression

$$|\,\mathbb{M} - \overline{x}\,| = \left| ap + bq - \frac{(p)\,a + (q)\,b}{(p) + (q)} \right| \leq$$

$$\leq |a| \cdot \left| p - \frac{(p)}{(p) + (q)} \right| + |b| \cdot \left| q - \frac{(q)}{(p) + (q)} \right|;$$

goes to zero when $(p) + (q)$ increases; more precisely, the probability that $|\,\mathbb{M} - \overline{x}|$ is smaller than ε, however small ε is, goes to one when $(p)+(q) \to \infty$ (3rd Bernoulli's theorem).

*

What we just established for a random experiment with two possible outcomes is valid for an experiment with any number of possible outcomes.

*

What we have proved about mean values is valid, with the necessary adaptations, for the mathematical expectation of any class. Therefore:

The mathematical expectation of a finite class of numerical elements *numéricos, isto é, de elementos a que se fazem corresponder números, é igual ao limite da soma dos números dessa classe, obtidos numa série de experiências, dividida pelo número total de experiências, quando êste último número tende para o infinito.*

É a esta proposição que a esperança matemática deve toda a sua importância.

No caso da classe considerada coincidir com a *classe possível*, a *esperança matemática* confunde-se com o *valor médio* e a proposição antecedente transforma-se nesta outra:

O valor médio duma função susceptível de passar por um numero finito de valores, é igual ao limite da média arithmética dos valores achados para a função, numa série de experiências cujo numero aumenta indefinidamente.

Daí toda a importância de que gosa a média arithmética nas aplicações do Cálculo das Probabilidades.

*

As proposições antecedentes podem ainda generalizar-se para uma função variando dum modo contínuo, numa dada região.

Generalizemos a proposição para o valor médio que só difere da esperança matemática, na forma.

Seja (A) a região, f(M) a função, P(M) a lei da probabilidade.

i.e., of elements to which we associated a number, is the limit, when the number of experiments goes to infinity, of the sum of the observed numbers in that class when we perform repeated experiments, divided by the number of experiments.

The practical importance of the mathematical expectation comes from the above statement.

If the class considered in the above statement is the *total possible class*, the *mathematical expectation* is the *mean value*, and the above proposition becomes:

The mean value of a function that can assume a finite number of values is the limit of the averages of the observed values of that function, in repeated experiments, when the number of experiments goes to infinity.

This is the reason why averages are so important in the applications of Probability.

*

All those propositions can be generalized for functions continuously varying in some region.

For instance, considering the mean value, that only formally differs from the mathematical expectation:

Let A be the region, $f(X)$ the function, $\mathbb{P}_X$ the probability law.

O valor médio da função será

$$M = \int_{(A)} f(M) \,.\, P(M) \,.\, d(A) \,.$$

Decomponhamos (A) em n partes e numeremo-las de 1 a n. A esperança matemática da região parcial (A_i), será

$$E_i = \int_{(Ai)} f(M)\, P(M)\, d(A)$$

$$= f(M_i) \int_{(Ai)} P(M)\, d(A)$$

$$= f(M_i) \,.\, P_{(Ai)} \qquad (1)$$

visto que P(M) é uma função sempre positiva e portanto podemos aplicar o 1.º teorema da média; em (1), $f(M_i)$ representará, portanto, o valor de $f(M)$ num ponto de (A_i) e $P_{(Ai)}$ a probabilidade da região (A_i).

Por outro lado, teremos que

$$M = \Sigma\, E_i = \Sigma f(M_i)\, P_{(Ai)} \,.$$

Suponhamos agora que grupamos os valores obtidos para a fumção $f(M)$ em classes correspondentes às regiões parciais (A_i) e, dentro de cada uma dessas regiões (A_i), substituamos $f(M)$ por $f(M_i) + \varepsilon_i$ onde ε_i, em virtude da suposta continuidade de $f(M)$, será infinitamente pequeno em relação a $f(M_i)$ e, portanto, tenderá para *zero*, à medida que (A_i) tenda para *zero*.

The mean value of the function is

$$\mathbb{M} = \int_A f(X)\,\mathbb{P}_X \mathrm{d}a.$$

Let's partition A in n parts $A_i,\ i = 1, \ldots, n$. The mathematical expectation of each of the parts is

$$\mathbb{E}_i = \int_{A_i} f(X)\,\mathbb{P}_X \mathrm{d}a =$$

$$= f(X_i) \int_{A_i} \mathbb{P}_X \mathrm{d}a =$$

$$= f(X_i) \cdot \mathbb{P}(A_i), \tag{6.1}$$

since $\mathbb{P}_X$ is always a positive function, and thus we can use the mean value 1st theorem. In (6.1), X_i denotes the value of the function f in a point $X_i \in A_i$, and $\mathbb{P}(A_i)$ the probability of the region A_i.

On the other hand, we have

$$\mathbb{M} = \sum \mathbb{E}_i = \sum f(X_i)\,\mathbb{P}(A_i).$$

Let us now assume that we group the observed values of the function f in classes corresponding to the partial regions A_i and that, inside each of those regions, we substitute $f(X)$ by $f(X_i) + \varepsilon_i$, where $\varepsilon_i = \mathrm{o}\,(f\,(X_i))$, because of the assumed continuity of f. Thus $\varepsilon_i \to 0$ when A_i decreases to $\emptyset$.

Consideremos a média dos valores de $f(M)$ decomposta em duas parcelas, a primeira correspondente aos valores $f(M_i)$ e a segunda correspondente aos valores de ε_i. Quando o número de experiências aumente indefinidamente, a primeira parcela tende (prop. antecedente) para $\Sigma f(M_i)\,P_{(A_i)}$ e portanto para M, qualquer que seja o modo da divisão de (A); a segunda parcela tenderá para um número cujo valor absoluto é inferior a outro número positivo, tão pequeno quanto se queira, visto que nós podemos fazer a divisão de (A) em partes tão pequenas quanto queiramos.

A segunda parcela tenderá, pois, para *zero*.

O limite da média dos valores de $f(M)$ *existirá*, portanto, e será igual ao *valor médio* ou *provável* de $f(M)$,

c. d. d.

*

* *

Como vimos no capítulo precedente, a probabilidade de que o afastamento relativo λ se mantenha, em valor absoluto, inferior a λ é dada por

$$\theta(\lambda) = \frac{2}{\sqrt{\pi}} \int_0^{\lambda} e^{-\lambda^2}\, d\lambda$$

A lei de probabilidade da variável λ será, pois,

$$P(\lambda) = \frac{1}{\sqrt{\pi}}\, e^{-\lambda^2}.$$

Let us consider the average of the observed values $f(x_k)$ of $f(X)$ decomposed into two summands, the first one corresponding to the values to the values $f(X_i)$ and the second one corresponding to the values ε_i. When the number of experiments goes to infinity, the first summand goes to $\sum f(X_i)\,\mathbb{P}(A_i)$, and therefore to $\mathbb{M}$, whatever the partition of A (Prop. XII). The second summand, as we can partition A in subsets whose measure is as small as we want, converges to zero.

Therefore the limit of the averages of the observed values $f(x_k)$ of $f(X)$ exists, and it is the *mean value* or *probable value* of $f(X)$. □

*

* *

As we have seen in the previous chapter, the probability that the absolute value of the relative deviation, denoted $|\Lambda|$, is less than λ_1 is

$$\theta(\lambda_1) = \frac{2}{\sqrt{\pi}} \int_0^{\lambda_1} \mathrm{e}^{-\lambda^2} \, \mathrm{d}\lambda.$$

Therefore the probability law of the relative deviation Λ is

$$\mathbb{P}_\Lambda(\lambda) = \frac{1}{\sqrt{\pi}} \, \mathrm{e}^{-\lambda^2}$$

for all real λ.

O valor médio de λ, será

$$M(\lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e\lambda \, d\lambda = 0 \, ;$$

o valor médio de $|\lambda|$, será

$$M(|\lambda|) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\lambda^2} |\lambda| \, d\lambda$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} 2e^{-\lambda^2} \lambda \, d\lambda$$

$$= \frac{1}{\sqrt{\lambda}} \left[-e^{-\lambda^2} \right]_0^{\infty}$$

$$= \frac{1}{\sqrt{\pi}} \, .$$

o valor médio de (λ^2), será

$$M(\lambda^2) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-\lambda^2} \lambda^2 d\lambda$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \lambda \, . \, e^{-\lambda^2} 2\lambda \, d\lambda$$

$$= \frac{1}{\sqrt{\pi}} \left[-\lambda e^{-\lambda^2} + \int e^{-\lambda^2} d\lambda \right]_0^{\infty}$$

$$= \frac{1}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = \frac{1}{2} \, .$$

The mean value of Λ is

$$\mathbb{M}(\Lambda) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\lambda^2} \lambda \, \mathrm{d}\lambda = 0.$$

The mean value of $|\Lambda|$ is

$$\mathbb{M}(|\Lambda|) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\lambda^2} |\lambda| \, \mathrm{d}\lambda =$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} 2 \, \mathrm{e}^{-\lambda^2} \lambda \, \mathrm{d}\lambda =$$

$$= \frac{1}{\sqrt{\pi}} \left(- \, \mathrm{e}^{-\lambda^2} \right]_0^{\infty} =$$

$$= \frac{1}{\sqrt{\pi}} \, .$$

The mean value of Λ^2 is

$$\mathbb{M}\left[\Lambda^2\right] = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-\lambda^2} \lambda^2 \, \mathrm{d}\lambda =$$

$$= \frac{1}{\sqrt{\pi}} \int_{0}^{\infty} \lambda \cdot \, \mathrm{e}^{-\lambda^2} 2 \, \lambda \, \mathrm{d}\lambda =$$

$$= \frac{1}{\sqrt{\pi}} \left(-\lambda \, \mathrm{e}^{-\lambda^2} + \int \, \mathrm{e}^{-\lambda^2} \, \mathrm{d}\lambda \right]_0^{\infty} =$$

$$= \frac{1}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} = \frac{1}{2} \, .$$

Donde se conclue que

$$\frac{M(\lambda^2)}{[M(|\lambda|)]^2} = \frac{\frac{1}{2}}{\frac{1}{\pi}} = \frac{\pi}{2}.$$

A proposição XII do presente capítulo, dá a esta relação uma significação notável: a de podermos rectificar a circunferência, por meio de lançamentos à sorte.

Therefore

$$\frac{\mathbb{M}\left[\Lambda^2\right]}{\left[\mathbb{M}(|\Lambda|)\right]^2} = \frac{\frac{1}{2}}{\frac{1}{\pi}} = \frac{\pi}{2}\,.$$

Proposition XII from this chapter confers a remarkable interpretation to this result: it is possible, using random throws, to rectify the circumference.

CAPÍTULO VII

CHAPTER VII

CONCLUSION

CAPÍTULO VII

Nos capítulos precedentes vimos o que se entende por probabilidade dum fenómeno que pode identificar-se com uma tiragem feita, à sorte, numa classe finita de elementos, ou com um lançamento, feito à sorte, numa região, no caso de sermos nós mesmos os agentes dessa tiragem ou lançamento e a classe ou região serem por nós conhecidas qualitativa e quantitativamente. Vimos tambêm como reduzir qualquer conjunto de tiragens ou lançamentos, a uma só tiragem ou lançamento, feitos numa classe ou região.

Vejamos agora como, servindo-nos dos princípios neles estabelecidos, poderemos alargar o campo de acção desta sciência.

Mas, para mais fácil e claramente fazermos esse estudo, principiemos por fazer uma classificação dos factos que pretendemos estudar.

Para isso admitiremos que, seres a nós semilhantes e agentes doutra natureza, podem, *em certas circumstâncias,* fazer tiragens análogas às tiragens (ou lançamentos) feitas, à sorte, por nós mesmos. Depois justificaremos e esclareceremos esta hipótese.

Admitida ela, devidiremos os fenómenos que cahem debaixo da alçada desta sciência, em três grupos. No pri-

CHAPTER VII

CONCLUSION

In the preceding chapters, we have investigated the probability of events which can be thought of as random extractions from a finite set, or as random throws in a region, on the assumption that we are, ourselves, the agents of the random selection, and that the set or the region are qualitatively and quantitatively specified. We also described how a sequence of random extractions or of random throws can be reduced to a single extraction from a finite set or to a single throw in a region.

We now describe how the scope of Probability can be broadened, using the principles formerly established.

For clarity, we start with a classification of the facts we want to investigate.

To do so, we shall admit the possibility that someone like us, or even essentially diverse agents can, *in some circumstances*, perform random extractions (or random throws) with analogous outcomes to those performed by us.

Once this has been accepted as admissible, we shall consider three groups of phenomena whose probability can be investigated.

meiro grupo, ficarão os fenómenos que podem identificar-se com lançamentos ou tiragens, feitas à sorte, por nós mesmos; no segundo, os fenómenos análogos a tiragens ou lançamentos, feitos por um ser semelhante a nós; no terceiro, os fenómenos análogos a lançamentos ou tiragens feitas por agentes doutra natureza.

Cada um dêstes grupos, devidi-lo-hemos em dois sub-grupos. No 1.º desses sub-grupos, ficarão os fenómenos dos quais se considera um número finito de modalidades; no 2.º, os fenómenos dos quais se considera um conjunto de modalidades formando um *contínuo*.

Em cada um dêstes sub-grupos consideraremos, ainda, três casos. Assim, no 1.º sub-grupo de cada grupo, podemos conhecer o fenómeno qualitativa e quantitativamente (1.º caso); conhecê-lo qualitativamente e desconhecê-lo quantitativamente (2.º caso); ou desconhecê-lo qualitativa e quantitativamente (3.º caso).

No 2.º sub-grupo de cada grupo, podemos conhecer a lei de probabilidade do fenómeno e o seu campo de existência (1.º caso); conhecer o campo de existência, mas desconhecer a lei (2.º caso); desconhecer a lei e o seu campo de existência (3.º caso).

A primeira divisão tem por critério a natureza do agente da tiragem ou lançamento; a segunda, a natureza do fenómeno; a terceira o nosso grau de conhecimento.

Como vimos na Introdução a estes *Elementos,* nós supunhamos conhecido todo o fenómeno que podia indentificar-se com uma tiragem (ou lançamento) feita, à sorte, por nós mesmos, numa classe (ou região) conhecida qualitativa e quantitativamente. É esse o nosso *fenómeno padrão,* o nosso *facto elementar*. Todo o fenómeno, para

In the first group, we enclose the phenomena that can be assimilated either to random extractions or to random throws performed by us; in the second group, phenomena that can be viewed as random extractions or as random throws, but performed by someone similar to us; in the third group, phenomena that can be assimilated either to random extractions or to random throws, done by some essentially different agent.

Each of those three groups will be further divided into two subgroups. In the first one, we consider the phenomena that can have a finite number of possible outcomes. In the second one, the phenomena whose possible outcomes conceptually form a continuous region.

In each of those subgroups, we consider three possible situations. In the first subgroup of each of the three groups, the three possible situations are: the set of phenomenon on outcomes is qualitatively and quantitatively known (first case); it is qualitatively known, but quantitatively unknown (second case); or it is unknown, both qualitatively and quantitatively (third case).

In the second subgroup of each group, we may know the probability law of the phenomenon, and the corresponding support (first case); or we may know the support, but ignore the probability law (second case); or we may ignore both the probability law and its support (third case).

The criterion used in the first classification is the nature of the agent of random extractions or of random throws. The second classification is done on the nature of the phenomenon; the third classification is based on our degree of knowledge of the phenomenon.

As we have seen in the *Introduction* to these *Elements of Probability Calculus*, we consider a phenomenon which can be identified to a random extraction (or to a random throw) done by us, ourselves, from qualitatively and quantitatively known finite set (or in a qualitatively and quantitatively known bounded region) as known, in the sense that everything is well specified. This is the description of our *standard phenomenon*, *standard model*, or *elementary fact*.

que possa fazer parte do estudo desta sciência, deve poder reduzir-se a este.

Principiaremos essa redução pelo

1.º Grupo

que, como já vimos, é caracterizado por nele, as tiragens ou lançamentos serem feitos, à sorte, por nós mesmos.

Neste grupo, vimos, tambêm, haver dois sub-grupos, sendo o

1.º SUB-GRUPO

constituído pelos fenómenos nos quais consideravamos um número finito de modalidades, isto é, por fenómenos análogos a tiragens feitas em classes finitas e descontínuas, ou a lançamentos feitos em regiões, divididas num número finito de partes.

O

1.º caso

dêste sub-grupo foi já estudado nos capítulos I, II e III dos presentes *Elementos*. É nele que está incluído o fenómeno padrão e é, portanto, a ele que teremos de reduzir o

2.º caso

que passamos a estudar e, em seguida o 3.º. Este 2.º caso é caracterizado por as tiragens serem feitas em classes conhecidas qualitativamente e desconhecidas quantitativamente.

Reduzir o 2.º caso ao 1.º; será, pois, determinar quantitativamente a classe em que as tiragens são feitas.

Essa determinação poder-se-hà fazer com uma *probabi-*

Only the phenomena amenable to the standard model can be the object of Probability.

We shall start our program of standardization of phenomena with the

1st Group

which, as we have seen, is the one in which random extractions or random throws are performed by us. In this group, as we said before, we must consider two subgroups, of which the

1st SUBGROUP

contains those phenomena having a finite number of outcomes. Such phenomena may be conceptualized as random extractions from finite discrete sets, or to random throws in bounded regions divided into a finite number of parts.

Its

1st case

has been studied in Chapters I, II and III of these *Elements of Probability Calculus*. It includes the *standard model*; thus, the starting point of our standardization program must investigate how to reduce to it the

2nd case

which deals with phenomena amenable to random extractions from sets qualitatively known but quantitatively unknown.

In this situation, the aim of the reduction procedure is, therefore, the quantitative determination of the set from which the random extractions are done.

This can be done with

lidade [1] *e aproximação* tão grandes quanto se queira, logo que possamos fazer tantas tiragens quantas quizermos (3.º teorema de J. BERNOULLI).

A determinação será, pois, *aproximada e provável*. Mas a probabilidade de que a aproximação dê lugar a um êrro, em valôr absoluto, inferior a ε, por mais pequeno que ε seja, diferirá de *um* tão pouco quanto quizermos. Este 2.º caso, fica, pois, separado do 1.º, pelo mesmo hiato que separa a *probabilidade* da *certeza* [2].

---

[1] Com «uma probabilidade tão grande quanto se queira» deve entender-se — uma probabilidade tão próxima de *um* quanto se queira.

[2] A *certeza* é, para nós, a *probabilidade* de tirar uma bola branca duma urna que só contêm bolas dessa côr. Para LAPLACE havia uma *diferença essencial* entre *probabilidade* e *certeza:* «Quand tous les cas sont favorables à un évènement, sa probabilité se change en certitude, et son expression devient égale à l'unité. Sous ce rapport, la certitude et la probabilité sont comparables, quoiqu'il y ait une différence essentielle entre les deux états de l'esprit, lorsqu'une vérité lui est rigoureusement démontrée, ou lorsqu'il aperçoit encore une petite source d'erreur» (LAPLACE, *Essai Philosophique sur Les probabilités*). Para JACOB BERNOULLI, essa *diferença essencial* não existia: «Certitudo rerum, spectata in ordine ad nós, non omnium eadem est, sed multipliciter variat secundum majis et minus. Illa de quibus revelatione, ratione, sensu, experientia, ἀνιοψία aut aliter ita constat, ut de eorum existentia vel futuritione nullo modo dubitare possimus, summa et absoluta certitudine gaudent. Caetera omnia imperfectiorem ejus mensuram in mentibus nostris obtinent, majorem minoremve, prout plures vel pauciores sunt probabilitates, quae suadent rem aliquam esse, fore aut fuisse».

Probabilitas enim est gradus certitudinis, et ab hac differt ut pars à toto (J. BERNOULLI, *Ars Conjectandi,* Pars Quarta, C. I).

high probability and precision, i.e. with the accuracy we wish in the approximation, and with as probability as large [24] as we want, insofar as it is feasible to perform as many trials as needed (J. Bernoulli's 3rd theorem).

The determination is therefore *approximate* and *probabilistic*. But the probability that the approximation produces an error whose absolute value is lower than ε, however small ε is, will approach 1 as much as we want. Thus this 2nd case is separated from the 1st case by the same hiatus that separates *probability* from *certitude*. [25]

---

[24] I.e., $1-\varepsilon$, with ε as small as desired.

[25] Certitude is the probability of extracting one white ball from an urn containing only white balls. Laplace " Quand tous les cas sont favorables à un évènement, sa probabilité se change en certitude, et son expression devient égale à l'unité. Sous ce rapport, la certitude et la probabilité sont comparables, quoiqu'il y ait une différence essentielle entre les deux états de l'esprit, lorsqu'une vérité lui est rigoureusement démontrée, ou lorsqu'il aperçoit encore une petite source d'erreur." (LAPLACE, *Essai Philosophique sur les Probabilités*). In Jacob Bernoulli's view, there is no *essential difference* between probability and certitude: "Certitudo rerum, spectata in ordine ad nos, non omnium eadem est, sed multipliciter variat secundum majis et minus. Illa de quibus revelatione, ratione, sensu, experientia, *άνιοψία* aut aliter ita constat, ut de eorum existentia vel futuritione nullo modo dubitare possimus, summa et absoluta certitudine gaudent. Caetera omnia imperfectiorem ejus mensuram in mentibus nostris obtinent, majorem minoremve, prout plures vel pauciores sunt probabilitates, quae suadent rem aliquam esse, fore aut fuisse.

Probabilitas enim est gradus certitudinis, et ab hac differt ut pars a toto." (J. BERNOULLI, *Ars Conjectandi*, Pars Quarta, Chap. I)

(Laplace's text: "When all cases are favorable to an event, its probability becomes certitude, and its value is unity. In this perspective, probability and certitude are comparable, although there is an essential difference between the two states of mind, resulting from the rigorous proof of a true statement, or from an argument where a possible source of error is still perceived."

Bernoulli's text: "Our view on the certitude of things is not always the same, it varies, being high in what concerns some, low in respect to others. We have complete and absolute certitude on those things that we know by revelation, by the exercise of the intellect or of the senses, by experience, by direct observation, or otherwise constated, and in no way doubt that they will exist or occur in the future. Under other circumstances, our mind assigns to things some lower degree of belief, higher or lower according to whether we judge large or small the probability that they exist, existed or will exist.

Probability is, thus, a degree of certitude, and differs from it as a part differs from the whole.")

A identificação não poderá deixar de ser *provável,* mas à falta de melhor, aceitá-la-hemos, visto que o conhecimento duma probabilidade representa, para nós, alguma coisa de útil.

Na avaliação da probabilidade das coisas «reside toda a sciência dos filósofos e toda a prudência dos políticos» *(... in quo solo omnis Philosophi sapientia et Politici prudentia versatur)* (1).

A identificação do

3.º caso

com o 1.º, faz-se do mesmo modo; os elementos a determinar são, agora, dois, em vez de um. Mas o processo da sua determinação é ainda o mesmo do caso antecedente.

Consideremos, agora, o

2.º SUB-GRUPO

que, como vimos era constituído por fenómenos dos quais se considerava um número infinito de modalidades, formando um *contínuo* que nós suporemos de *segunda espécie,* na terminologia de H. POINCARÉ (2). Suporemos, pois, que a cada modalidade do fenómeno em questão, fazemos corresponder um ponto do espaço (a um número conveniente de dimensões). O

1.º caso

deste sub-grupo, que é caracterizado por nele se supôr conhecida a lei de probabilidade e o seu campo de exis-

(1) JACOB BERNOULLI, *Ars Conjectandi,* pars quarta, cap. II.
(2) H. POINCARÉ, *La Science et l'Hypothèse,* chap. II.

The identification cannot be but *probable*, but, having no better choice we accept it, since probabilistic knowledge is a useful degree of knowledge.

"All the science of philosophers and all the prudence of politicians deals with" the evaluation of the probability of events ([...] *in quo solo omnis Philosophi sapientia et Politici prudentia versatur*) [26]

The identification of the

3rd case

with the 1st one is done in a similar fashion, with the extra task of identifying qualitatively the set of possible outcomes. But the identification procedure is, in all steps, similar to the former one.

Let us now consider the

2nd SUBGROUP

which, as we have seen, contains the phenomena whose set of possible outcomes is infinite, building up a *continuous* that we shall assume to be of the *second species*, in the terminology of H. Poincaré [27]. We shall therefore assume that to each possible outcome of the phenomenon we associate one point in a space with the convenient number of dimensions. The

1st case

from this subgroup, characterized by the fact that its probability law and

(26) J. BERNOULLI, *Ars Conjectandi*, Pars Quarta, Chap. II.
(27) H. POINCARÉ, *La Science et l'Hypothése*, Chap. II.

tência, foi já estudado nos capítulos II, III e IV. Passemos, portanto, para o

2.º caso

no qual se supõe conhecido o campo de existência da lei de probabilidade, sendo, porêm, essa lei desconhecida. A redução deste grupo ao antecedente, consiste na determinação dessa lei.

Quando fazemos uma série de lançamentos, à sorte, numa região e observamos os pontos directamente lançados, a razão do número de pontos que cahem dentro duma parte dessa região para o número de pontos que cahem dentro doutra de igual extensão, tende para unidade, à medida que o número de pontos lançados aumenta (3.º teorema de BERNOULLI). Isto é, à medida que o número de lançamentos aumenta, a distribuição dos pontos lançados tende a tornar-se uniforme.

Já se não dá o mesmo se, em vez de observarmos o ponto lançado directamente, à sorte, observarmos uma sua imagem ou projecção. Neste caso, ainda em virtude do mesmo teorema de BERNOULLI, a distribuição far-se-hà de harmonia com a correspondente *lei de probabilidade.* Os pontos observados condensar-se-hão nas vizinhanças dos *máximos* dessa lei. Conhecendo a lei de probabilidade podemos, pois, *prevêr* a distribuição das imagens ou projecções dos pontos lançados à sorte, com uma probabilidade que aumenta com o número de pontos a abservar.

Recíprocamente, nós poderemos, observando um grande número de lançamentos, determinar a lei de probabilidade do ponto observado, com uma probabilidade tão grande quanto se queira; ou antes, podemos determinar o valôr the corresponding support are known, has been dealt with in Chapters II, III and IV. We may therefore proceed to the

2nd case

in which the support is known, but the probability law itself is not known. The reduction of this case to the previous one consists, therefore, in the determination of the probability law.

When we execute a sequence of random throws in some region and we observe directly the points that result from each trial, the ratio of the number of points that lie in a given subregion to the number of points that lie in another region of the same size will converge to one as the number of trials increases (Bernoulli's 3rd theorem). In other words, as the number of trials increases, the distribution of the points approaches the uniform distribution.

But if, instead of directly observing the points randomly chosen, our goal is to study the law of the point's projections or of some other image point, the probability law is no longer uniform. But, according to the above mentioned Bernoulli's 3rd theorem, the distribution of these image points will be governed by the corresponding probability law. The observed points will be concentrate in the neighborhood of the maxima of such probability law. When the probability law is known, we can *forecast* the distribution of the points projections or of other image points of the randomly thrown points, and the probability that this forecast agrees with the reality is increasing with the number of points.

The other way round, observing a large number of points, we can determine the corresponding probability law, with a probability as large as we wish; more precisely, we may compute the value do integral dessa função desconhecida, estendido a qualquer segmento da região considerada (3.º teorema de Bernoulli).

Este facto fornece-nos dois métodos para a determinação da lei desconhecida.

1.º

Pode dar-se que, razões inerentes à natureza do fenómeno a estudar, nos justifiquem a adopção, *á priori,* duma lei de probabilidade, como, por exemplo, acontece com os êrros de observação. Neste caso, faremos um grande número de séries de lançamentos, sendo cada série formada por um número de lançamentos suficientemente grande, para que seja deminuta a probabilidade de que a sua distribuíção se não harmonize com a distribuíção prevista pela lei admitida *à priori,* se essa lei fôr a verdadeira.

A razão do número de séries que se harmonizarem com a lei admitida, para o número total de séries, dar-nos-hà um número a que (como já vimos no 2.º caso do 1.º grupo) poderemos chamar probabilidade dessa lei. Se julgarmos essa probabilidade suficiente, a lei será admitida; se não, será regeitada.

O

2.º

método consiste no seguinte: divide-se a região em que varia o ponto observado, região esta que se supõe conhecida, num número de partes suficientemente [1] grande.

Faz-se, em seguida, um número suficientemente grande de lançamentos à sorte. A razão do número de pontos

---

(1) O número de partes e o número de lançamentos a fazer, dependerá do rigôr de que precisarmos nos resultados.

of the integral, in any given interval from its support, of that unknown probability function (Bernoulli's 3rd theorem).

From this fact we get two methods to determine the unknown law:

1st method

It may happen that some reasons which are inherent to the nature of the phenomenon we are studying point towards the adoption, *a priori*, of some specified probability law — as it happens, for instance, when we are dealing with observation errors. In order to decide whether this is so, we perform a large number of sequences of trials, each sequence with a number large enough of trials, so that the probability that its distribution doesn't agree with the *a priori* law, provided this one is the true one, is negligible.

The ratio of the number of sequences whose empirical law matches with the hypothesized probability law to the total number of series approaches a number that (as we have already seen in the 2nd case of the 1st subgroup) we may call the probability of that law. If in our view that probability is large enough, the *a priori* law is maintained; otherwise, it is rejected.

The

2nd method

is the following: partition the region — which is assumed to be known — where the observed point varies into a large enough [28] number of subregions.

Once this has been done, we perform a large enough sequence of random throws in that region. The ratio of the number of points

[28] The number of subregions and the number of trials can be tuned so that we can achieve the accuracy we wish in the final results.

observados em cada uma dessas partes, para o número total de lançamentos, dár-nos-hà, com uma aproximação e probabilidade que podemos tornar tão grandes quanto queiramos, o integral da função desconhecida, ao longo de cada uma dessas regiões parciais.

Dividindo cada número achado, pela grandeza da região correspondente, obtemos outros tantos valôres da função procurada. Mas isto não basta para determinar a lei.

Hà uma infinidade de funções que, integradas ao longo das regiões consideradas, dão números iguais aos antecedentes. Como escolher uma, dentre tantas?

Todas as funções cujos integrais sejam iguais aos números achados, são igualmente bôas visto que se harmonizam igualmente com o 3.º teorema de Bernoulli. Escolheremos, à falta de razão mais forte, aquela que mais bem se preste às aplicações que lhe quizermos dar, isto é, a mais simples, para o fim que tenhamos em vista.

Como no sub-grupo antecedente, o

3.º caso

reduz-se ao 2.º. Os pontos observados distribuir-se-hão numa região de contôrno absolutamente arbitrário.

Podemos, mesmo, supôr que o campo de existência da lei é ilimitado em todos os sentidos; a determinação da lei nos diz depois quais as partes dessa região de probabilidade nula; isto é, a própria lei limitará o seu campo de existência.

Do mesmo modo, no 3.º caso do 1.º sub-grupo, nós podiamos supôr que o fenómeno era qualitativamente indeterminado; a determinação quantitativa da classe nos diria depois quais as modalidades de probabilidade nula.

we observe in each subregion to the total number of random throws gives us, with the degree of probability and approximation we want, the integral of the unknown function in each of the partial subregions.

Dividing each number computed as described by the size of the corresponding subregion, we obtain an equal number of points of the function we wish. This, however, isn't sufficient to determine the probability law.

In fact, there is an infinite number of functions that, when integrated in the considered subregions, furnish the same results, namely the ones computed as described above. How to select one among this infinite number of possibilities?

All the functions whose integrals in the subregions match those numbers are equally plausible, since they have the same degree of agreement with Bernoulli's 3rd theorem. Among them, we choose the simpler one, the one which is more adequate for our goals, if no deeper reasons can guide our judgement.

As in the preceding subgroup, the

3rd case

reduces to the 2nd one. The observed points are distributed in a region of arbitrary boundary.

We can even assume that the support of the probability law is unbounded in all directions; the specification of the law will indicate, afterwards, which subregions do have null probability; in other words, the law itself will limit its domains of existence.

We might as well assume, as in the 3rd case of the 1st subgroup, that the phenomenon was qualitatively unknown; the quantitative determination of its probability law would then specify the events of null probability.

Debaixo deste ponto de vista, estes dois últimos casos não são distintos um do outro. Convêm, porêm, distingui-los, para facilitar a exposição.

Nota

No que acabamos de dizer, nós supômos, implicitamente que as classes (ou regiões) em que as tiragens são feitas, se manteem qualitativa e quantitativamente invariáveis. De contrário, nada podia fazer-se. A não ser que a variação se fizesse duma maneira lenta e regular e fôsse possível determinar a lei de variação, de modo a poder fazer as devidas *correcções.*

Passemos agora para o

2.º Grupo

de fenómenos e justifiquemos, em primeiro lugar, a hipótese em que apoiamos a sua construção, ou antes, expliquemos o que queremos dizer com ela.

A proposição *tirar, à sorte, um elemento duma classe,* tem para nós um sentido quando nós somos os agentes da tiragem.

Mas quando o agente da tiragem seja um ser semilhante a nós, essa proposição não tem, para nós, sentido, ou antes, não tem, para nós, um sentido diferente do da proposição *tirar um elemento duma classe.* Casos haverá, porêm, em que seja legítimo dar a esta proposição o mesmo sentido da primeira.

Com efeito, em que condições devemos nós de fazer uma tiragem, para que a possamos dizer *feita, à sorte?*

Em primeiro lugar, devemos desconhecer, por completo, In that perspective, these two cases are identical. We insist however in distinguishing the two cases, for the sake of clarity in the exposition.

Note

In all that has been said it was implicitly assumed that the sets and regions in which the random operations are performed remain invariable during the experiments, both qualitatively and quantitatively. Otherwise, no sound conclusions could be reached, unless the law of such variation was known, thus allowing the necessary *corrections*.

Let's now investigate the

2nd Group

of phenomena, starting with an explanation of the hypothesis assumed for its construction, and of its meaning.

The proposition *to extract, at random, one element from a set* has for us a precise meaning, when we are the agents of such random selection.

But when the agent of the selection is someone else, this proposition is ambiguous, in the sense that it has no different meaning from the proposition *to extract one element from a set.* In some situations, however, it is legitimate to retain the expression *random*, essential for our study.

What are the conditions needed to accept that the selection *has been done at random*?

First of all, the distribution of the elements in the class where the a distribuíção dos elementos na classe em que a tiragem é feita.

Em segundo lugar, devemos fazer a tiragem de modo que não possamos *prevêr* o elemento que vai sair, nem tão pouco a sua qualidade, não podendo essa *previsão* ser feita *por nós,* nem por nenhum *ser semelhante a nós.*

Ora, todas as vezes que as tiragens sejam feitas nestas circunstâncias, por um ser semelhante a nós, nada se opõe, *à priori,* a que façamos a hipótese de que as tiragens darão os mesmos resultados que dariam se elas fôssem feitas por nós mesmos.

Assim, dada uma urna, contendo nove décimos de bolas brancas para um décimo de bolas pretas, nós apostariamos que saíria uma bola branca, numa tiragem feita, à sorte, se nós mesmos fôssemos os agentes da tiragem. E continuariamos a apostar na mesma côr, se, o agente da tiragem sendo outrem, ela fôsse feita nas circunstàncias atrás descritas, isto é, se puzéssemos o agente da tiragem em condições de não poder *prevêr* o elemento que ia tirar, nem de poder ter o menor indício da distribuíção das bolas dentro da urna. É na admissão desta hipótese que se fundam todos os jogos de azar e é para a justificação dela que, por exemplo, se baralham as cartas dum baralho, antes de serem dadas e se dá ao verso das mesmas uma aparência identica.

*

Todas as vezes que esta hipótese não possa ser feita, por falta duma ou de ambas as condições, o fenómeno saírá fora do domínio desta sciência.

selection is done must be ignored.

Further, the extraction has to be done in such a way that it is impossible to *forecast* either which element will be selected, or its quality; this *impossibility of forecasting* must hold in what concerns not only *us, ourselves*, as *anyone similar to us.*

Hence, if the extractions are done by someone similar to us, according to the above requirements, nothing opposes, *a priori*, the hypothesis that the outcome of such extractions has the some informative value that would have the outcome of a random selection done by us.

For instance, assuming that 90% of the balls in one urn are white and the remaining 10% are black, we would favor a bet that the outcome of a random extraction would be white ball, in case we would be the agent of the random extraction. But we would surely stick to this bet, the extraction being performed by someone else, in case we would be satisfied that the circumstances of the extraction were as described: namely, the agent of the extraction was unable to *predict* the element he would extract, and ignorant of how the balls are mixed in the urn. All games of chance take those assumptions for granted; shuffling the deck before dealing the cards, and the fact that the back of all cards in a deck is the same, fulfills the above requirements.

*

Whenever the above requirements are not fulfilled, the phenomenon is out of the scope of the Science of Probability.

*

A hipótese de que esse agente está a fazer tiragens à sorte, deve ser posta de parte, todas as vezes que os resultados se não harmonisem com as leis de BERNOULLI e análogas, porque aqui, como aliás em qualquer sciência, as condições exteriores podem enganar-nos.

Assim o pensava o padre GALIANI, filósofo e economista do século XVIII, como se vê por esta anedota que dele conta BERTRAND:

«Un jour, à Naples, un home de Basilicate, en présence de l'abbé GALIANI, agita trois dés dans un cornet e paria d'amener rafle de 6; il l'amena sur-le-champ. Cette chance est possible, dit-on; l'home réunit une seconde fois, et l'on répéta la même chose; il réunit les dés dans le cornet trois, quatre, cinq fois, et toujours rafle de 6. *Sangue di Bacco,* s'ecria l'abbé, *les dés sont pipés*» (1).

*

Pelo que acabamos de vêr, os fenómenos do segundo grupo, ou bem se identificam com os do primeiro e nesse caso já estão estudados; ou não se identificam e nesse caso não fazem parte do objecto desta sciência.

*

No que acabamos de dizer àcêrca dos fenómenos do

(1) J. BERTRAND *Calcul des Probabilités,* Préface.

*

We must reject the hypothesis that another agent is doing random extractions whenever the outcome doesn't agree with Bernoulli's and similar laws. In fact, as in any other Science, external circumstances can cheat us.

This is clearly demonstrated in an anecdote that Bertrand reports about Galiani, philosopher and economist of the XVIII century:

"Un jour, à Naples, un homme de Basilicate, en présence de l'abbé Galiani, agita trois dés dans un cornet et paria d'amener rafle de 6; il l'amena sur-le-champ. Cette chance est possible, dit-on; l'homme réunit une seconde fois, et l'on répéta la même chose; il réunit les dés dans le cornet trois, quatre, cinq fois, et toujours rafle de 6. *Sangue di Bacco*, s'écria l'abbé, *les dés sont pipés*". [29]

*

As we have just discussed, the phenomena in the second group can be reduced to phenomena in the first group, and in that case we can view them as studied, or they are not amenable to phenomena in the first group, and they do not fall in the scope of Probability.

*

In what we said above about phenomena in

[29] J. BERTRAND, *Calcul des Probabilités*, Préface.
"One day, at Naples, a man from Basilicate, the abbot Galiani being present, shacked a cornet with three dice and bet he would throw a 6; he did, in fact, throw a 6. This is possible, no doubt; but he made a second throw, obtaining once again 6, and the same happened in a third throw, a fourth throw, a fifth throw — always he got a 6. *Sangue di Bacco*, said the abbot, *these dice are loaded*."

segundo grupo, não nos referimos explicitamente aos fenómenos do segundo sub-grupo, porque a analogia que existe entre a *composição qualitativa* duma classe e o campo de existência da lei de probabilidade, *composição quantitativa* e *forma* da mesma lei, nos dispensa dessa referência.

Consideremos, agora, o

3.º Grupo

de fenómenos.

Quando fazemos uma série de tiragens (ou lançamentos), à sorte, os elementos obtidos sucedem-se desordenadamente, isto é, sem obedecerem a nenhuma lei. De contrário, podê-los-hiamos *prever*, o que é incompatível com o que a intuição nos diz das tiragens feitas à sorte, consideradas nos casos já estudados. Ora, quando as modalidades dum qualquer fenómeno se sucedem dum modo desordenado, esse facto desperta, também, em nós, uma vaga noção de *sorte* ou *acaso*.

Poderemos nós reduzir, identificar, esta vaga noção de *acaso*, agora considerada, com a de hà pouco? Isto é, poderemos nós determinar *quantitativamente* uma classe, formada *qualitativamente* pelas modalidades do fenómeno em questão, de modo a podermos *supôr* que os fenómenos, produzidos pelas suas causas naturais, se sucedem como se fôssem tirados, à sorte, por nós mesmos, nessa classe? Ou, no caso de conhecermos apenas em parte o número de modalidades do fenómeno (o que podemos *sempre* supôr), poderemos nós determinar qualitativa e quantitativamente a classe correspondente?

Admitiremos que sim.

the second group, we skipped the phenomena in the second subgroup. In fact, this is a reasonable option, in view of the analogy between the qualitative composition of a finite set and the support of the probability law, and between the quantitative composition in the former case and the form of the probability law.

Let us now consider the

3rd Group

of phenomena.

When we execute a sequence of random extractions (or of random throws), the elements we get are unordered, with runs that apparently do not follow any law. Otherwise, we could *predict* them, and this is incompatible with what intuition tells us about random extractions. Henceforth, when the outcomes of some phenomenon do occur in an unordered fashion, this hazardous character gives us a vague feeling that it is governed by *chance*.

Can we identify that vague feeling of chance intervention with the random character of sequences of random extrations we have studied? In other words, can we quantitatively determine the set of the qualitative outcomes of the said phenomenon, so that we may *assume* that the outcomes produced by natural causes have the same random character possessed by random extractions done by us in that set?

Or, if only part of the possible outcomes of the phenomenon is known (and this may always be *assumed*), can we in all cases determine qualitatively and quantitatively the corresponding set?

We shall hypothesize that this is so.

Como justificar esta hipótese?

Duma só maneira: verificando que as conclusões que dela se tiram, se harmonizam com os factos.

Essas conclusões teem todas, por remate, os teoremas de Bernoulli e proposições análogas.

A primeira conclusão a verificar, seria a correspondente ao 3.° teorema de Bernoulli, porque ela nos servia para determinar qualitativa e quantitativamente a classe.

Ora, o 3.° teorema de Bernoulli diz-nos que: hà uma probabilidade *sempre crescente* de que a razão do número de fenómenos observados duma modalidade para o número total de experiências, difira da probabilidade dessa modalidade tão pouco quante se queira. Por outras palavras: se, após cada experiência, dividirmos o número de vezes que obtivemos cada uma das modalidades, pelo número total de experiências, obtemos números que tendem para um limite que é a probabilidade da modalidade correspondente; a probabilidade de que esses números se avizinhem dos respectivos limites, aumenta com o número de experiências.

De modo que, enquanto o número de experiências fôr deminuto, os números obtidos variarão muito irregularmente (visto que a probabilidade de que se avizinhem do limite, é deminuta); mas essa irregularidade irá deminuindo à medida que o número de experiências aumentar.

Se os factos observados se não harmonizarem com o que acabamos de dizer, é porque a hipótese de que partimos não é ligítima, quer porque os fenómenos não sejam identicos a tiragens feitas à sorte, quer porque a classe em que as tiragens são feitas, varie.

Se a verificação se fizer, não só para o 3.° teorema de

How can we justify such a hypothesis?

There is one way out: we must verify whether the conclusions that we can deduce from it conform to the observed facts.

Bernoulli's theorems, and those analogous to them, are the appropriate tools to judge whether this is so.

The first step is to analyze what we get in the light of Bernoulli's 3rd theorem, since it is useful in the qualitative and quantitative determination of the law.

In fact, Bernoulli's 3rd theorem tells us that the probability that the relative frequency of each possible outcome approaches, as much as desired, the probability of that outcome is *always increasing*. In other words: if after experiment we divide the number of times we got each of the outcomes by the total number of experiments, we obtain relative frequencies that converge to the probabilities of the outcomes. The probability that those numbers get closer and closer to the corresponding limits increases with the number of trials.

So, while the number of trials is moderate, those numbers will fluctuate showing some irregularity (since the probability that they are close to their limits is small); but their fluctuation will be smoother and smoother when the number of trials increases.

If the observed facts do not agree with this pattern of behavior, this must be interpreted as an indication that the hypothesis we assumed is wrong, either these phenomena have a pattern that is not identifiable with random extractions, or the phenomenon is varying in time.

If the facts are in agreement with Bernoulli's 3rd theorem and

BERNOULLI, mas também para as proposições análogas, é porque a hipótese era legítima durante o intervalo de tempo gasto nessa verificação.

E, enquanto que o meio em que o fenómeno se produz, não variar sensivelmente, nada se opõe a que continuemos a ter a hipótese por aceitável. Se o meio variar, a hipótese ainda pode ser aceitável; mas pode a variação do meio alterar a composição da classe e nesse caso será preciso determiná-la de novo e tantas vezes quantas as precisas para determinar a lei de variação, se essa lei existir.

with the analogous theorems we discussed in Chapter V, we may maintain the hypothesis that the observed outcomes of the phenomenon behaved like random extractions, during the period taken in the verification.

So, while there is no reason to question the stability of the process producing this phenomenon, there is no substantial reason to doubt that hypothesis. Even if the process changes, the hypothesis may still be acceptable, but further investigation has to be carried out, to make the necessary adaptations and amendments, and as often as required to determine the eventual pattern of variation.

INDICE

CONTENTS

INDICE

CONTENTS