

Sala 5
Gab. —
Est. 55
Tab. 20
N.º 35

Sala 5
Gab. —
Est. 58
Tab. 20
N.º 35

ARNDT
ENIG
FERNANDES
L. N.º 100 Coimbra - COIMBRA



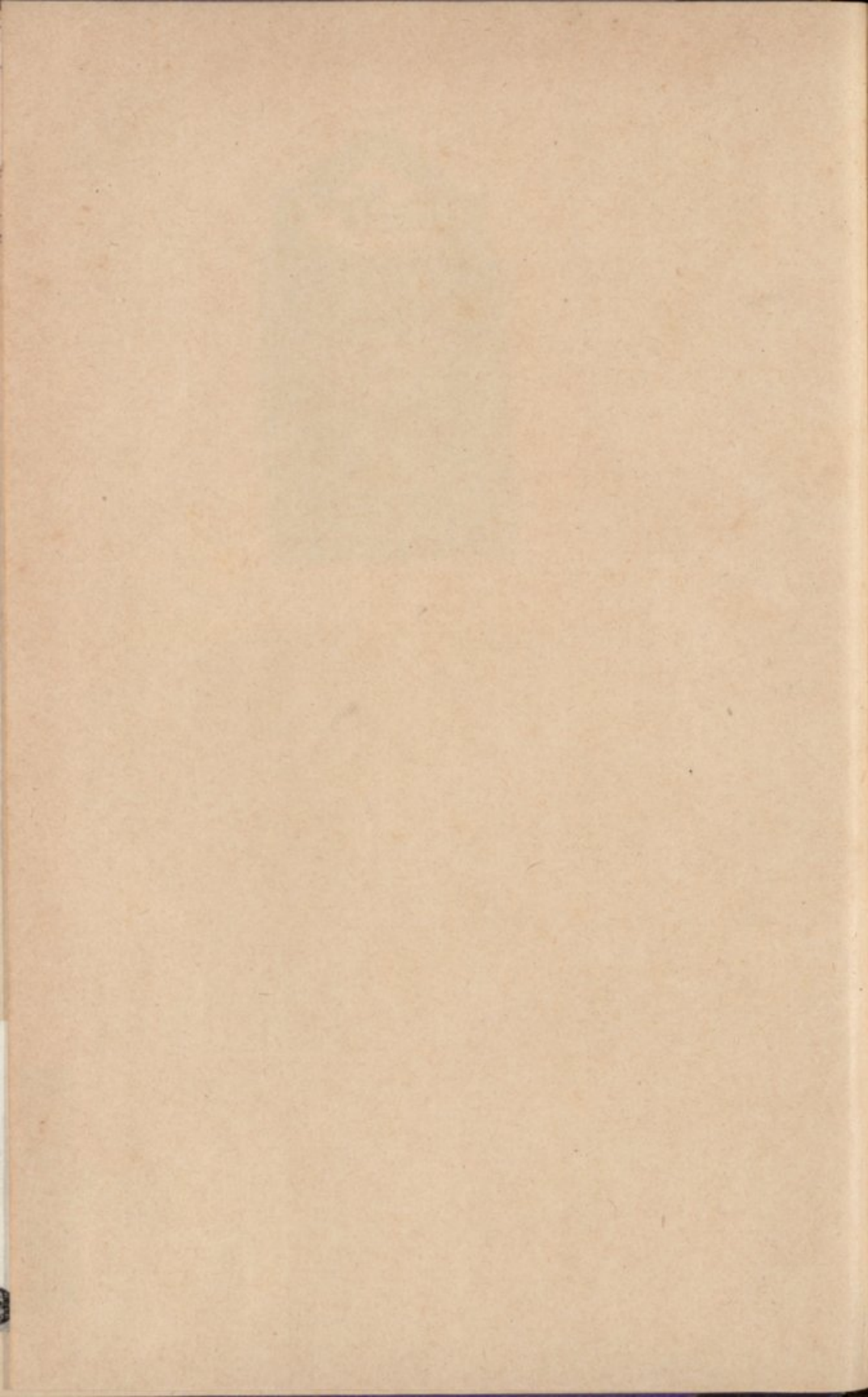
UNIVERSIDADE DE COIMBRA
Biblioteca Geral



1301500207



b12237280



MANUEL DOS REIS

SÔBRE

OS

PRINCÍPIOS FUNDAMENTAIS

DO

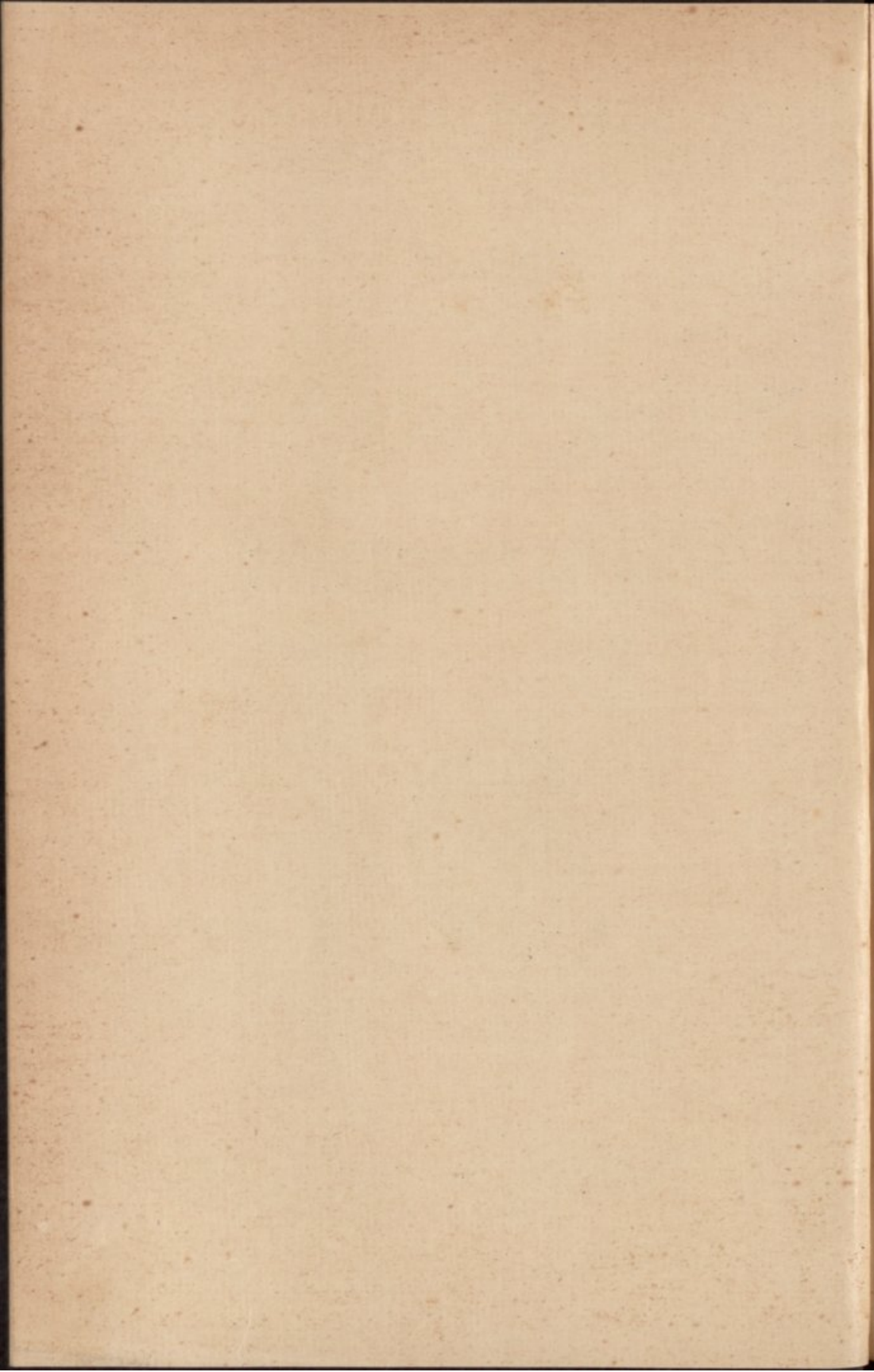
CÁLCULO DAS PROBABILIDADES



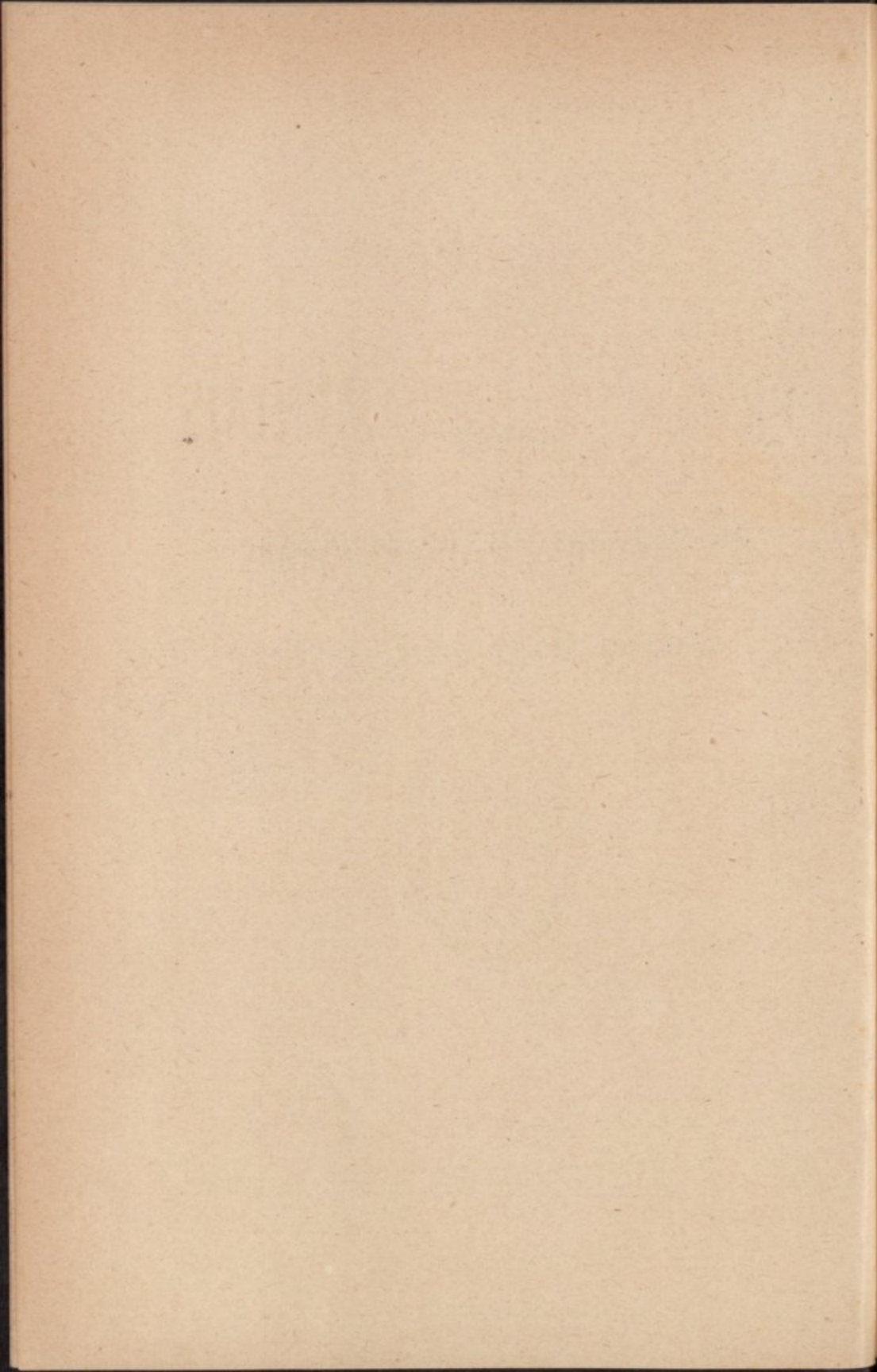
COIMBRA

IMPRESA DA UNIVERSIDADE

1929



SÔBRE
OS
PRINCÍPIOS FUNDAMENTAIS
DO
CÁLCULO DAS PROBABILIDADES



MANUEL DOS REIS

SÔBRE

OS

PRINCÍPIOS FUNDAMENTAIS

DO

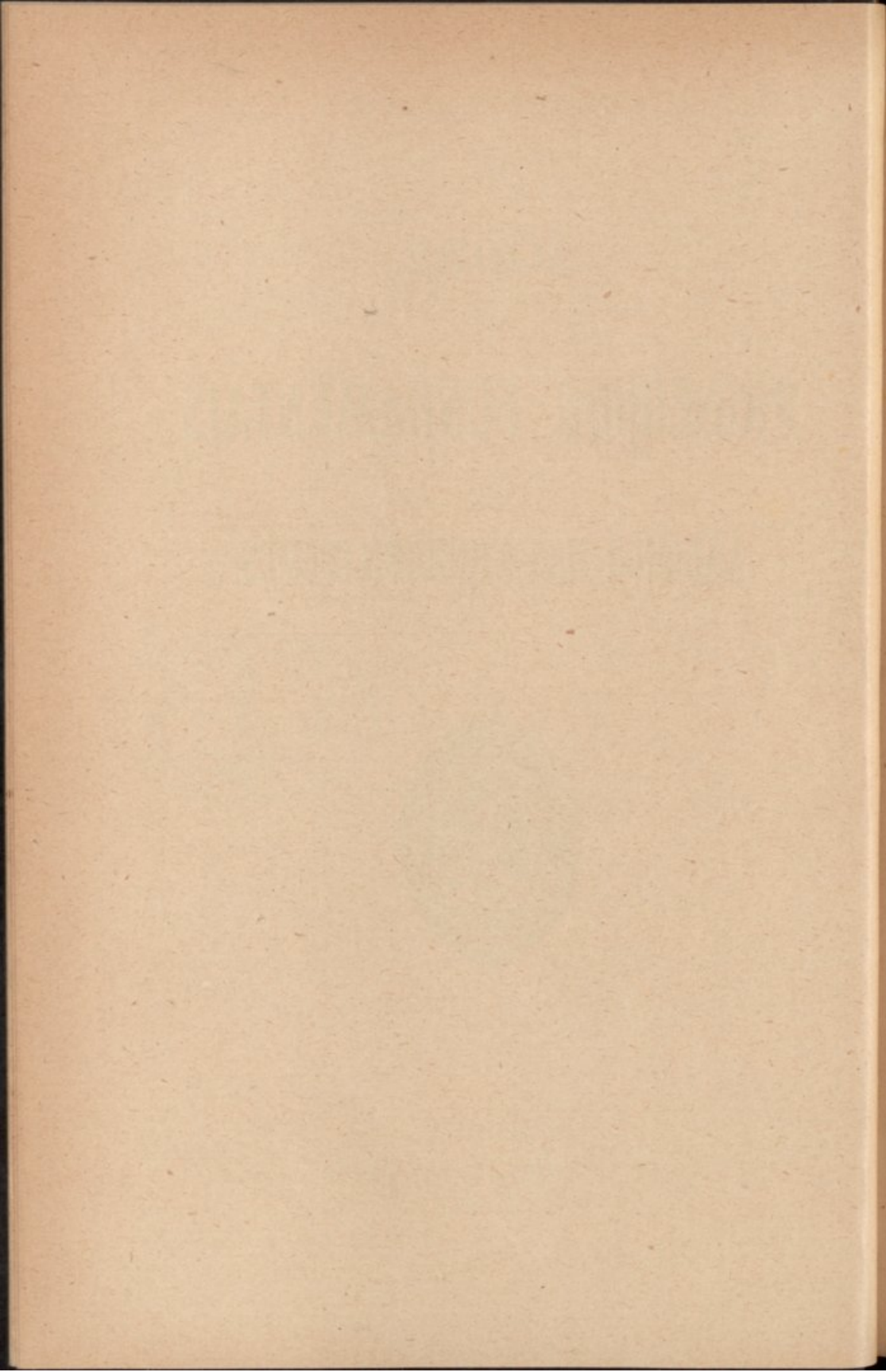
CÁLCULO DAS PROBABILIDADES



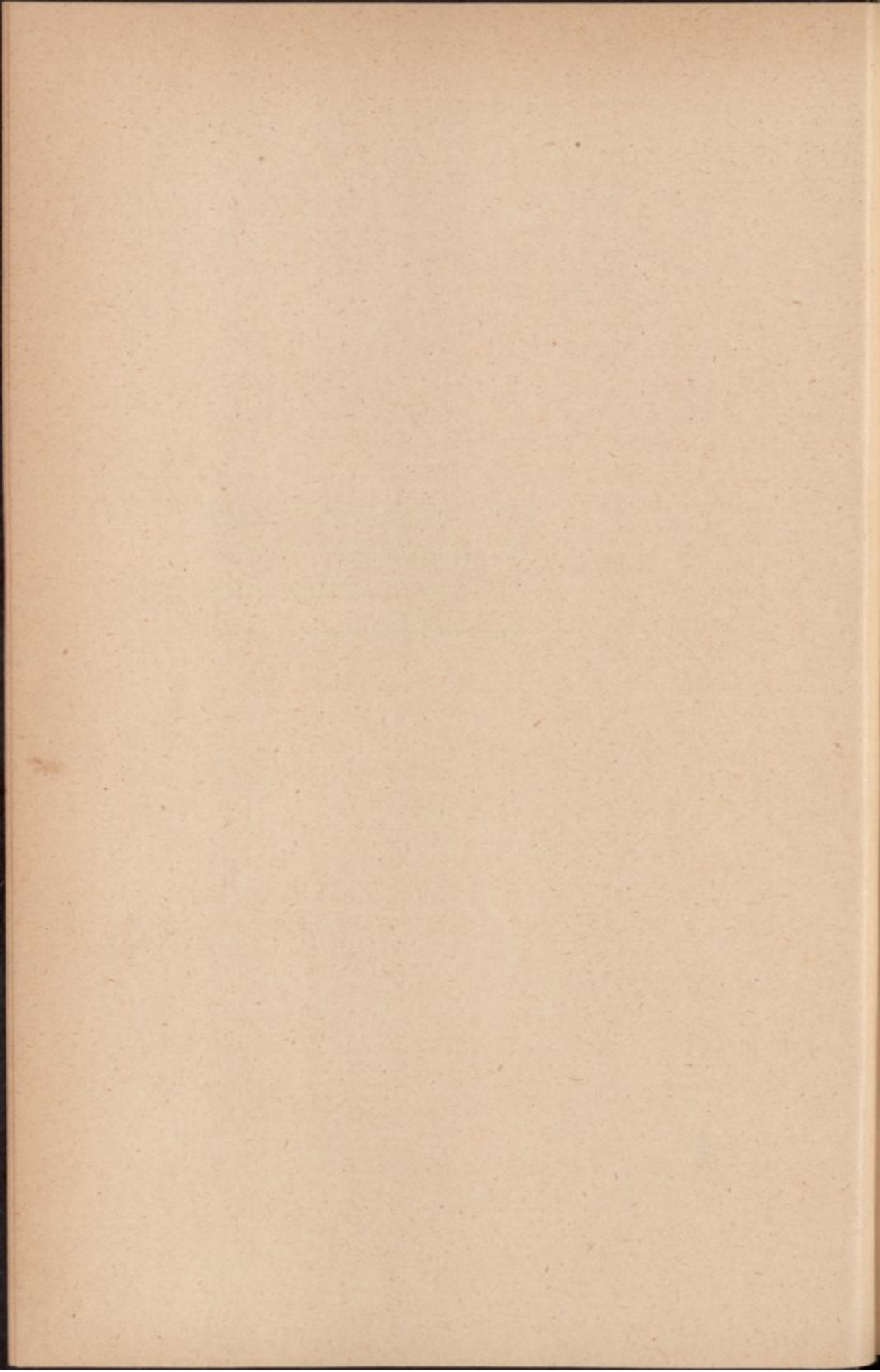
COIMBRA

IMPRESA DA UNIVERSIDADE

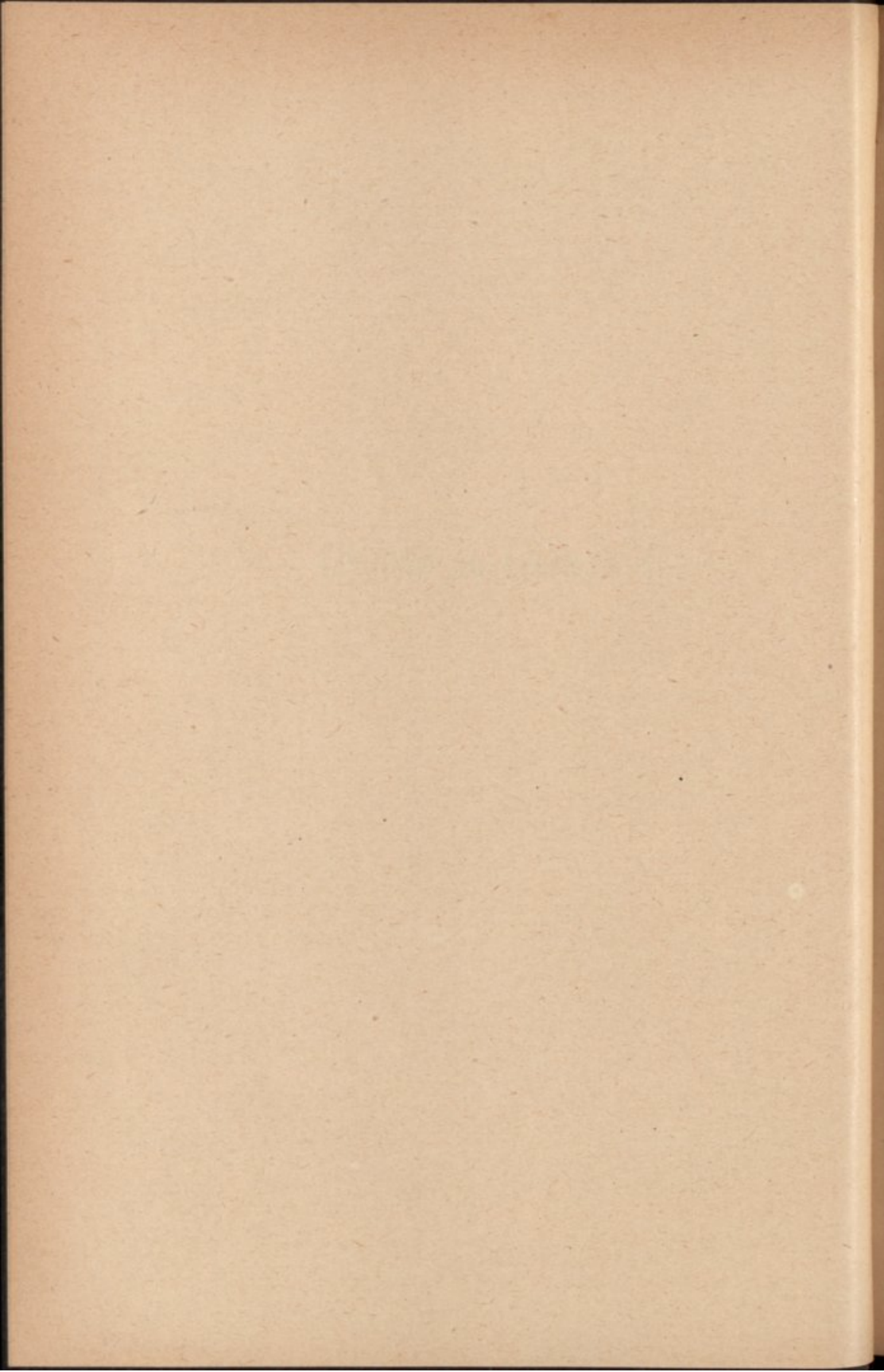
1929



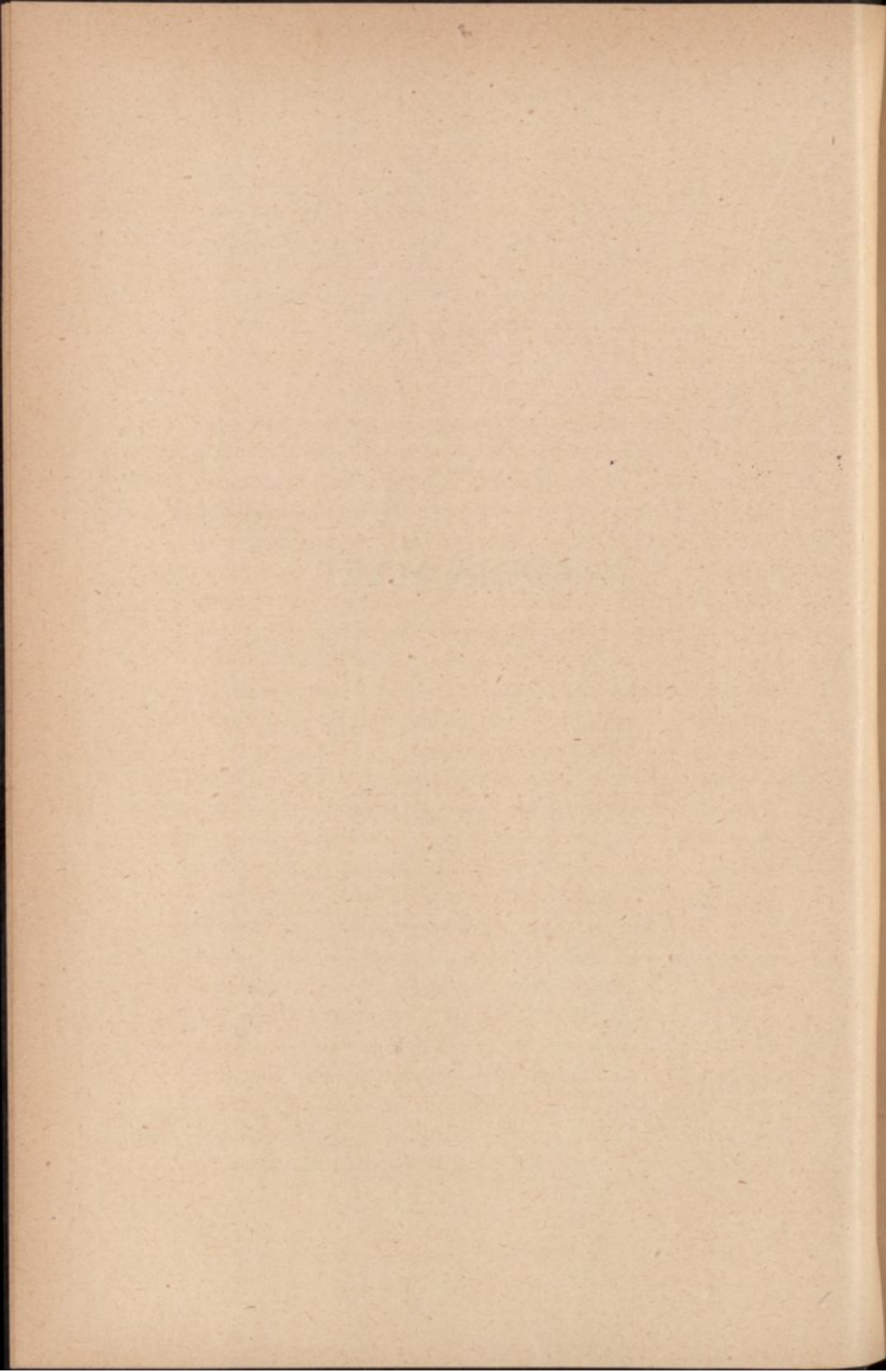
DISSERTAÇÃO PARA O EXAME
DE DOUTORAMENTO EM MA-
TEMÁTICA NA UNIVERSIDADE
DE COIMBRA : : : : : :



À MEMÓRIA DE MEU PAI



A MINHA MÃE



PREFÁCIO

A presente dissertação tem por origem um curso que fizemos na Faculdade de Ciências de Coimbra durante o segundo semestre de 1926-27. Todavia as nossas ideas sôbre os fundamentos do Cálculo das Probabilidades modificaram-se sensivelmente num intervalo de dois anos e os auditores dêsse curso, que porventura lerem as páginas que hoje damos à publicidade, difficilmente reconhecerão em certos pontos as nossas preleções. Acresce que nos propuzémos escrever uma obra essencialmente teórica, deixando de lado tudo quanto tem carácter de applicação, e que inserimos, entre outras, certas adições e generalizações inadequadas a um curso elementar.

A-pesar-do que se tem escrito sôbre a matéria, a construção dos fundamentos da Teoria das probabilidades deixa muito a desejar. Reconheceu-o entre nós o sr. dr. Pacheco de Amorim que, na sua Dissertação inaugural, apresentou uma tentativa interessante no sentido de pôr ordem lógica no caótico das exposições clássicas das bases da Teoria. Há porém nas ideas do sr. dr. Pacheco de Amorim, às quais, não obstante, devemos muito, uns senões que as tornam inaceitáveis em bloco. O mais importante é a separação radical estabelecida entre o princípio da divisão e o princípio da multiplicação, fazendo daquele uma definição da probabilidade mediante o artificio da distinção entre probabilidade e «possibilidade». Sem que pretendamos ter exposto os fundamentos do cálculo das proba-

bilidades in aeternum, cremos que do nosso trabalho resulta claramente a inutilidade desta distinção e a equivalência dos princípios da divisão e da multiplicação das probabilidades. Desta equivalência, aliás, fazem uso os autores clássicos; sòmente não dão dela uma justificação cabal.

Pecam as teorias gerais da probabilidade, que têm sido propostas, não só por ilogismo no estabelecimento dos princípios fundamentais, mas ainda por deficiência de matéria. A probabilidade numerável é posta geralmente de lado e quando isto não sucede só lhe é feita uma referência passageira. A Memória que Borel publicou nos Rendiconti de Palermo e depois inseriu na 2.^a edição das suas Leçons sur la théorie des fonctions não contém uma teoria desta probabilidade, mas sim o estudo de certos problemas particulares, alguns dos quais pouco têm que ver com a probabilidade numerável pròpriamente dita. A probabilidade contínua é estudada no caso particular das regiões e ordinariamente dum modo imperfeito. Comete-se por exemplo a inconseqüência de teorizar para volumes de espaços euclidianos e logo em seguida estudar problemas relativos a domínios de variedades destes espaços, que são em geral não-euclidianas. Pareceu-nos pois útil fazer o estudo da probabilidade contínua nos espaços riemannianos, e não só para regiões, mas para conjuntos mensuráveis quaisquer, utilizando os belos trabalhos de Lebesgue sòbre a medida e a integração, e completando-os para este efeito

onde se tornou necessário. Longe de nos confinarmos aos conjuntos homogêneos, só os utilizámos como ponto de partida, sendo o alvo dos nossos esforços o caso da heterogeneidade, que é o caso geral. Nos conjuntos finitos ou numeráveis, introduzimos as probabilidades irracionais; nos conjuntos da potência do contínuo, as densidades da probabilidade descontínuas. Esta orientação geral reflectiu-se naturalmente na teoria da esperança, que cremos ter apresentado dum modo rigoroso e completo, não obstante as abreviações que em todo o livro tivemos, por vezes, de fazer, ou porque não nos agradasse repetirmo-nos, ou porque a menor importância do assunto a isso aconselhava, ou ainda porque nos era materialmente vedado ultrapassar certos limites.

Os dois trabalhos que inserimos nos §§ 15 e 16 deviam fazer parte dum segundo capítulo que projectámos escrever sobre as propriedades gerais dos grandes números, e que seria o principal desta dissertação. As lacunas que possa haver nos treze primeiros §§ podem explicar-se até certo ponto por terem sido redigidos em vista das questões que no segundo capítulo desejávamos tratar. Porém motivos ponderosos impedem-nos de ser extenso, e assim tivemos de nos limitar a extrair das nossas notas o que nos pareceu ter mais originalidade, sem nos preocupar o ex-abrupto com que os dois §§ são introduzidos. No § 15 demonstramos dois teoremas novos sobre limites de probabilidade

para variáveis aleatórias independentes, nos quais se contém, como corolários particularíssimos, duas proposições devidas ao sr. dr. Pacheco de Amorim. No § 16 mostramos que o princípio em que Gauss baseou a sua teoria definitiva do método dos menores quadrados se justifica tanto melhor quanto maior é o número de observações. Os teoremas do § 15 valem também para vectores independentes; mas isto será objecto dum artigo que oportunamente publicaremos.

A precipitação com que redigimos este volume acarretou-lhe naturalmente defeitos, para os quais pedimos a indulgência do leitor.

Coimbra, Maio de 1929.

PRELIMINARES

No Cálculo das probabilidades consideram-se acontecimentos, passados, presentes ou futuros, possíveis mediante a realização de circunstâncias determinadas, mas geralmente incertos. A incerteza da realização dum acontecimento é meramente accidental se o acontecimento e as circunstâncias que lhe condicionam a possibilidade pertencem inteiramente à ordem física. Imperando nesta ordem de fenómenos um determinismo rigoroso, como admite a ciência moderna, a incerteza dum acontecimento só pode existir pela ignorância das relações de causalidade existentes entre êle e as circunstâncias que o tornam possível. Se acontecimento e circunstâncias não pertencem inteiramente à ordem física, aquela incerteza pode ser essencial. É o que se dá por exemplo nos jogos de azar, onde, nas circunstâncias que condicionam a possibilidade dum acontecimento não observado, intervém em condições especiais um acto humano que escapa, pelo menos em parte, ao determinismo físico. Esta distinção, importante para julgar do valor do Cálculo das probabilidades, não interessa porém à sua construção formal.

Poisson definiu a probabilidade dum acontecimento como sendo a nossa razão de crer que êle se realizou ou realizará. Se não queremos entrar em discussões de carácter filosófico, podemos aceitar esta definição. Duma urna que continha bolas brancas e pretas tirou-se à sorte uma bola, de cuja côr não fui informado; tenho motivo de

crer que ela é branca e tenho motivo de crer que ela é preta; a saída de bola branca e a saída de bola preta são dois acontecimentos que, para mim, têm probabilidade. Amanhã choverá em Coimbra, ou não choverá; tenho motivo de crer que sim e motivo de crer que não; chover amanhã em Coimbra, ou não chover, são ainda dois acontecimentos que, para mim, têm probabilidade.

Informo-me porém de que na urna do primeiro exemplo havia no momento da tiragem mais bolas brancas do que pretas; tenho agora mais razão de crer na saída duma bola branca que na saída duma bola preta; ao primeiro acontecimento attribuo maior probabilidade que ao segundo. Um meteorologista, que sabe que o bom tempo que hoje faz em Coimbra apresenta bons indícios de estabilidade, tem mais razão do que eu de crer que amanhã não choverá; para êle, é mais provável o bom tempo amanhã que para mim.

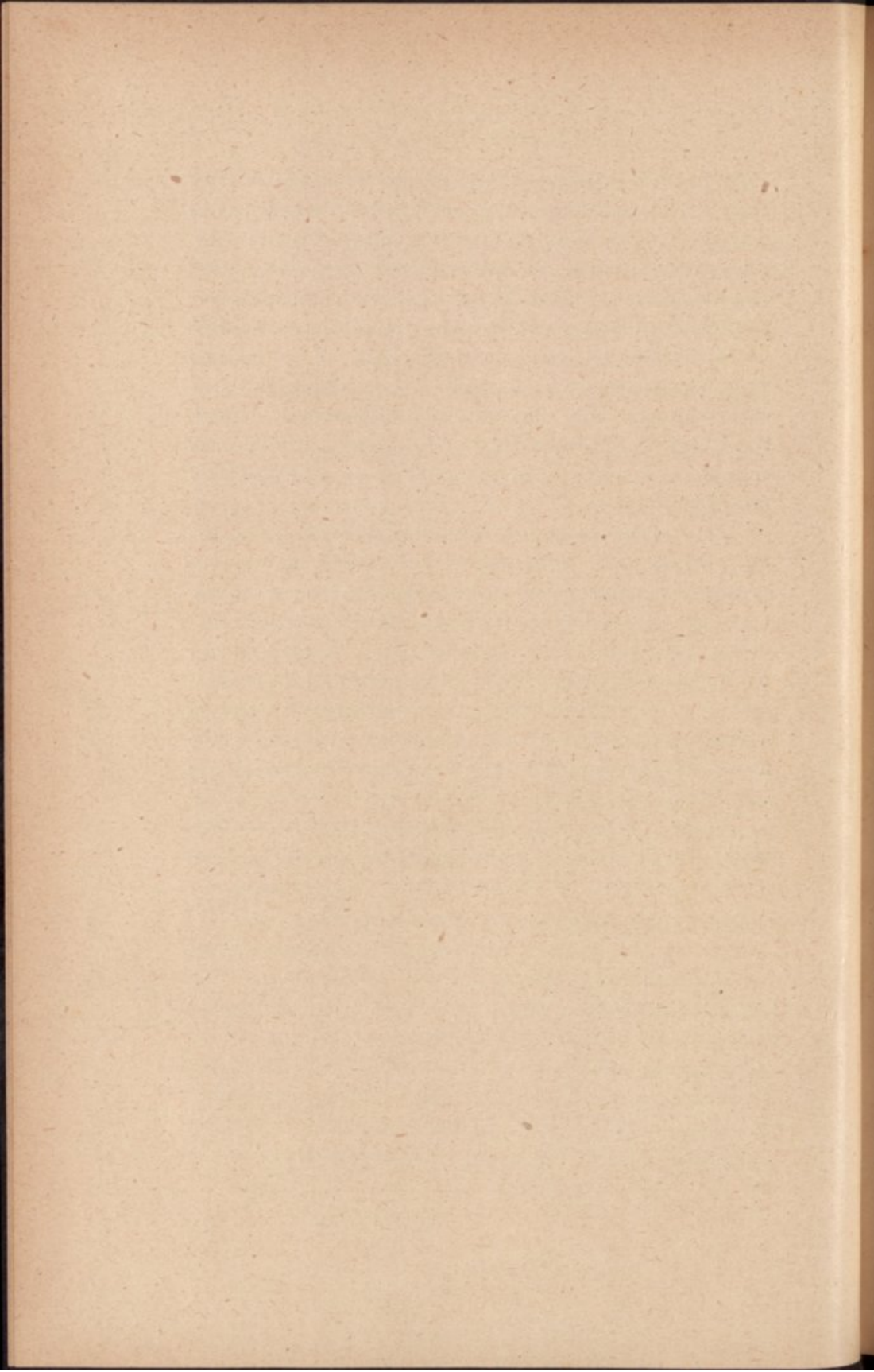
Êstes exemplos mostram que a probabilidade é uma grandeza e que ela varia com o grau do nosso conhecimento acerca das circunstâncias que aos acontecimentos dão possibilidade de realização. Há grandezas que não são quantidades matemáticas, mas a probabilidade não está neste caso. Veremos como esta grandeza é susceptível de medida e portanto como é possível um Cálculo das probabilidades.

O que produz a incerteza dum acontecimento, possível em circunstâncias determinadas, é que há pelo menos outro acontecimento possível nessas circunstâncias e incompatível com êle. Somos assim naturalmente conduzidos a agrupar um acontecimento possível com todos os acontecimentos possíveis nas mesmas circunstâncias e incompatíveis com êle e entre si. Um tal conjunto de acontecimentos é coerente, neste sentido: obtém-se sempre o mesmo conjunto partindo de qualquer dos seus elementos

e empregando o mesmo critério de agrupamento. A estes conjuntos chamaremos *totais*, por opposição aos conjuntos *parciais*, constituídos por acontecimentos possíveis e incompatíveis entre si, mas incompatíveis ainda com outros acontecimentos; é claro que um conjunto parcial é parte alíquota dum conjunto total. Dado um conjunto total de acontecimentos possíveis e incompatíveis, se se realizam as circunstâncias que condicionam a possibilidade dos seus elementos, produz-se um destes necessariamente; de contrário, haveria outro acontecimento possível nas mesmas circunstâncias e incompatível com êles, e o conjunto dado não seria total.

Um conjunto total pode ter um número de elementos finito ou infinito. Se êste número é infinito, há ainda a distinguir os dois casos mais importantes que podem apresentar-se: o dos conjuntos numeráveis e o dos conjuntos com a potência do contínuo. Isto dá lugar à divisão das questões de probabilidade em três categorias: as questões de probabilidade finita, de probabilidade numerável e de probabilidade contínua. As questões das duas primeiras categorias podem reunir-se sob a denominação de questões de probabilidade descontínua.

No que vai seguir-se pretendemos essencialmente formular em abstracto os princípios indispensáveis à resolução de questões das três categorias.



§ 1.º

Da probabilidade
nos conjuntos finitos homogéneos

1. Consideremos um conjunto finito total A de acontecimentos possíveis e incompatíveis e suponhamos que, para cada par de elementos de A , não há motivo para nos inclinarmos a admitir a realização dum elemento de preferência à do outro (quer êste motivo não exista absolutamente, como sucede com a tiragem à sorte feita numa urna, quer seja simplesmente ignorado por nós).

Um tal conjunto denominá-lo-hemos *homogéneo*. O nosso primeiro objecto é definir a *probabilidade* matemática de cada elemento de A , i. e. procurar como deve atribuir-se a cada elemento de A um número, que será a medida da probabilidade da sua realização.

Sendo o conjunto A homogéneo, é claro que devemos atribuir por probabilidade a todos os seus elementos um mesmo número p , função do número deles, n ; êsse número p tomá-lo-hemos positivo. Os elementos dum conjunto homogéneo são pois igualmente prováveis.

Uma vez fixado êste número p , a questão que se põe em seguida é a de calcular a probabilidade de que o acontecimento produzido pertença a um dado conjunto A' contido em A . Se fôr m o número de elementos de A' ($m \leq n$), esta probabilidade depende necessariamente de n e m ou, o que é o mesmo, de m e p ; suporemos que ela é o valor duma função

$$f_m(x_1, x_2, \dots, x_m)$$

de m variáveis independentes x_1, \dots, x_m , para

$$x_1 = x_2 = \dots = x_m = p.$$

É evidente que a função f_m tem de ser simétrica e deve satisfazer ainda às condições

$$f_m(x_1, \dots, x_m) > f_{m-1}(x_1, \dots, x_{m-1})$$

para

$$x_1 = \dots = x_m = p,$$

e

$$f_1(p) = p.$$

A fim de obter o máximo de simplicidade, adoptaremos a função

$$x_1 + x_2 + \dots + x_m,$$

que verifica tôdas as condições precedentes. A probabilidade de que o acontecimento produzido pertença indeterminadamente ao conjunto A' , ou, abreviadamente, a probabilidade do conjunto A' , será portanto igual à soma das probabilidades dos elementos dêste:

$$P_{A'} = p + p + \dots + p = mp.$$

Se applicamos esta regra, que é uma forma do *princípio da adição das probabilidades*, ao próprio conjunto total A , obtemos o número

$$np$$

para valor da probabilidade de que o elemento de A realzado pertença a A . Mas de tal facto há certeza, e vê-se

que a certeza nos aparece matematicamente como uma forma especial de probabilidade: à certeza vem também atribuído um número. Isto não pode significar que a certeza, estado psicológico uniforme, que não admite grau, seja em si susceptível de medida, mas sim que os diversos graus de probabilidade se lhe podem comparar sob o aspecto da quantidade; neste sentido se deve entender a frase de Jacob Bernoulli «*probabilitas enim est gradus certitudinis, et ab hac differt ut pars a toto*» (1). A certeza impõe-se-nos pois como unidade natural da probabilidade; e assim faremos

$$np = 1,$$

o que determina o número p que procurávamos:

$$p = \frac{1}{n}.$$

A probabilidade dum acontecimento dum conjunto finito total homogéneo é portanto o quociente da unidade pelo número de elementos do conjunto.

A probabilidade elementar é tanto menor quanto maior fôr o número de elementos do conjunto total que se considere. Quando êste número cresce ilimitadamente, a probabilidade elementar tende para zero; o estado psicológico que a acompanha tende visivelmente para a certeza da não-realização do elemento, e por isso zero simboliza a impossibilidade.

A probabilidade, tal como acabamos de defini-la, é susceptível de todos os valores racionais compreendidos entre zero e 1, e significa um estado de espírito variável,

(1) *Ars conjectandi, Basileae, MDCCXIII. Pars Quarta, pág. 211.*

intermediário a dois estados de certeza: certeza da realização do acontecimento e certeza da sua não-realização.

Mais tarde veremos como se introduzem as probabilidades irracionais.

2. A probabilidade do conjunto A' parcial de A é dada por

$$P_{A'} = m p = \frac{m}{n};$$

é o quociente do número de elementos de A' pelo número de elementos do conjunto total.

Como zero é a probabilidade dum acontecimento impossível, e é impossível que o acontecimento produzido em A pertença a um conjunto parcial de A que não contenha elemento algum, a fórmula anterior aplica-se ainda quando fôr $m = 0$.

A expressão $\frac{m}{n}$, encontrada para $P_{A'}$, exige que a probabilidade dum conjunto B' de m' elementos, contido num conjunto total B de n' elementos, seja igual a $P_{A'}$ se m' e n' forem proporcionais a m e n . Laplace ⁽¹⁾ e Poisson ⁽²⁾ mostraram que assim é. O raciocínio de Poisson, mais convincente que o de Laplace, consiste no seguinte:

Suponhamos por exemplo duas urnas A e B , a primeira com uma bola branca e duas pretas, a segunda com três bolas brancas e seis pretas, em cada uma das quais se efectua à sorte a tiragem duma bola, e comparemos as

(1) *Théorie analytique des probabilités*, 3.^e édition, Paris, 1920, in *Oeuvres complètes*, t. VII, Paris, 1886, pág. IX.

(2) *Recherches sur la probabilité des jugements*, Paris, 1837, pág. 31 e 32.

probabilidades da saída duma bola branca nas duas tiragens. A totalidade das bolas da primeira urna representa o conjunto total A , a das bolas da segunda o conjunto total B , e as bolas brancas da primeira e da segunda respectivamente os conjuntos parciais A' e B' . Consideremos agora três outras urnas α , β e γ , cada uma com uma bola branca e duas pretas; para efectuar a comparação das duas probabilidades mencionadas, comparemo-las à probabilidade duma bola branca na tiragem à sorte duma das urnas α , β , γ seguida da tiragem à sorte duma bola na urna saída. É claro que, tendo estas três urnas igual composição, é indiferente tirar à sorte uma delas para em seguida tirar desta uma bola, ou simplesmente tirar uma bola duma delas previamente escolhida: entre elementos idênticos, a escolha não dá resultado diferente da tiragem à sorte. A terceira probabilidade é pois igual à primeira, porque a composição das urnas α , β , γ é a mesma que a da urna A . Por outro lado, sendo igualmente prováveis as três bolas de cada uma das urnas α , β , γ , e sendo igualmente prováveis estas três urnas, as bolas de tôdas três, consideradas em conjunto, são também igualmente prováveis, evidentemente; assim êste conjunto, formado por três bolas brancas e seis pretas, tôdas igualmente prováveis, não difere do conjunto das bolas da urna B , e portanto a terceira probabilidade é igual à segunda. Resulta imediatamente a igualdade das duas probabilidades que queríamos comparar, e é visível que a demonstração comporta tôda a generalidade.

Todavia, regressando à consideração do conjunto genérico A , êste raciocínio não implica que $P_{A'}$ seja necessariamente da forma $\frac{m}{n}$. Uma função arbitrária de $\frac{m}{n}$, atribuída a $P_{A'}$, seria compatível com êle. Pode provar-se que $\left(\frac{m}{n}\right)^k$, onde é $k > 0$, é a única função de m e n

que convém atribuir a P_N se, em vez de

$$x_1 + x_2 + \dots + x_m,$$

se toma, mais geralmente, para

$$f_m(x_1, \dots, x_m)$$

uma qualquer função homogénea de grau 1, sob certas restrições, o que não discorda da condição

$$f_1(x_1) = x_1.$$

Com efeito, será então

$$\begin{aligned} P_N &= f_m(p, p, \dots, p) \\ &= p f_m(1, 1, \dots, 1), \end{aligned}$$

ou

$$P_N = p \varphi(m),$$

pondo

$$\varphi(m) = f_m(1, 1, \dots, 1);$$

a função positiva $\varphi(m)$ é crescente, em virtude da condição

$$f_m(1, 1, \dots, 1) > f_{m-1}(1, \dots, 1);$$

e por ser

$$1 = p \varphi(n),$$

teremos

$$P_N = \frac{\varphi(m)}{\varphi(n)}.$$

Vejamos qual é a função φ que satisfaz à equação fun-

cional

$$\frac{\varphi(m)}{\varphi(n)} = F\left(\frac{m}{n}\right).$$

Se φ é derivável, teremos

$$\frac{\varphi'(m)}{\varphi(n)} = F'\left(\frac{m}{n}\right) \cdot \frac{1}{n}$$

e

$$\frac{\varphi(m) \cdot \varphi'(n)}{[\varphi(n)]^2} = F'\left(\frac{m}{n}\right) \cdot \frac{m}{n^2},$$

onde $\varphi'(m)$ representa a derivada de $\varphi(m)$ relativamente a m , $\varphi'(n)$ a de $\varphi(n)$ relativamente a n e $F'\left(\frac{m}{n}\right)$ a de $F\left(\frac{m}{n}\right)$ relativamente a $\frac{m}{n}$.

As duas últimas fórmulas dão

$$\frac{\varphi'(m)}{\varphi(m)} : \frac{\varphi'(n)}{\varphi(n)} = \frac{n}{m},$$

e portanto a função φ satisfaz à equação diferencial

$$\xi \frac{\varphi'(\xi)}{\varphi(\xi)} = k,$$

onde k é uma constante positiva; o integral geral desta equação é

$$\varphi(\xi) = C \xi^k,$$

onde C representa uma constante arbitrária. Resulta

daqui para P_N o valor

$$\left(\frac{m}{n}\right)^k,$$

como tínhamos afirmado.

Se definíssemos assim P_N , ainda a probabilidade variaria entre 0 e 1. Apareceriam desde logo, como possíveis, certos valores irracionais da probabilidade, se k não fôsse inteiro. A função

$$f_m(x_1, \dots, x_m)$$

seria ainda susceptível duma grande indeterminação. O mais simples seria pôr

$$f_m = \left(x_1^{\frac{1}{k}} + \dots + x_m^{\frac{1}{k}}\right)^k,$$

que verifica tôdas as condições precedentes. Mas, mesmo entre as funções algébricas elementares, esta estaria longe de ser a única satisfatória; por exemplo, a função

$$\sqrt{\frac{2}{m+1}}(x_1^2 + \dots + x_m^2 + x_1 x_2 + x_1 x_3 + \dots + x_{m-1} x_m),$$

que verifica também essas condições, satisfaria ao caso $k = \frac{1}{2}$; e o produto desta função por \sqrt{m} satisfaria ao caso $k = 1$. Note-se porém que nestas duas funções figura m entre os coeficientes das variáveis, o que se não dá na primeira, que parece ser a única dêste género, entre as funções algébricas elementares que verificam as nossas condições.

Com efeito, de

$$P_N = \left(\frac{m}{n}\right)^k$$

e

$$p = \frac{1}{n^k},$$

vem

$$(P_{A'})^{\frac{1}{k}} = \frac{m}{n},$$

e

$$p^{\frac{1}{k}} = \frac{1}{n};$$

e a relação entre os m números

$$\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n},$$

com coeficientes independentes de m , que produz $\frac{m}{n}$, é claramente

$$\frac{1}{n} + \frac{1}{n} + \dots + \frac{1}{n};$$

daqui resulta

$$P_{A'} = \left(p^{\frac{1}{k}} + \dots + p^{\frac{1}{k}} \right)^k,$$

onde o segundo membro é exactamente o valor da função

$$\left(x_1^{\frac{1}{k}} + \dots + x_m^{\frac{1}{k}} \right)^k$$

para

$$x_1 = \dots = x_m = p.$$

Em conclusão, o princípio de adição e a definição corrente

$$P_{A'} = \frac{m}{n}$$

não são necessários no Cálculo das probabilidades. Mas $P_{A'}$ é necessariamente da forma $F\left(\frac{m}{n}\right)$, e particularmente da forma $\left(\frac{m}{n}\right)^k$ nas condições precisas que indicámos.

O princípio da multiplicação, de que adiante trataremos, também não é necessário em geral; mas é-o, não só para $P_{A'} = \frac{m}{n}$, mas ainda para $P_{A'} = \left(\frac{m}{n}\right)^k$, como facilmente se pode verificar. O princípio da multiplicação tem pois um grau maior de necessidade que o princípio de adição, e efectivamente os dois princípios são essencialmente diferentes.

As considerações que acabamos de fazer sobre a definição da probabilidade pode parecer que conduzem a uma infinidade de Cálculos das Probabilidades diferentes, análogamente ao que succedeu em Geometria com a descoberta dos espaços não euclidianos, sobretudo quando se atende à variedade de formas que pode revestir o princípio de adição. Mas o caso é muito diferente. À variedade de formas possíveis do princípio de adição contrapõe-se a relativa uniformidade da expressão final da probabilidade, suficiente para fazer dos infinitos Cálculos das Probabilidades, realmente possíveis, formas sem interesse, mesmo teórico, dum único Cálculo das Probabilidades, essência de todos: o Cálculo das Probabilidades tradicional. Efectivamente, mesmo no caso mais geral, em que a probabilidade $P_{A'}$ é $F\left(\frac{m}{n}\right)$ e F satisfaz a condições evidentes em virtude do que precede, os resultados do cálculo das diversas probabilidades obter-se-iam imediatamente a partir dos resultados do Cálculo clássico, pois que os primeiros são os valores duma função F da variável cujo campo de existência é o conjunto dos segundos. Também as proposições teóricas sobre limites de probabilidades resultariam imediatamente das proposições ordinárias, se se tomasse para F uma função contínua.

3. Sejam A_1', A_2', \dots, A_s' conjuntos parciais de A sem elementos comuns quando considerados dois a dois, e sejam m_1, m_2, \dots, m_s , respectivamente, os números de elementos destes conjuntos. O conjunto

$$A' = A_1' + \dots + A_s'$$

tem evidentemente a probabilidade

$$\frac{m_1 + \dots + m_s}{n},$$

e daqui resulta imediatamente

$$P_{A'} = P_{A_1'} + P_{A_2'} + \dots + P_{A_s'}.$$

A probabilidade de que o acontecimento produzido pertença indeterminadamente a um conjunto indeterminado dentre vários conjuntos mutuamente exclusivos é igual à soma das probabilidades de que o acontecimento pertença indeterminadamente a cada um destes conjuntos. Êste enunciado é a forma geral do *princípio da adição* nos conjuntos finitos homogêneos.

Se um conjunto parcial é a soma de conjuntos não mutuamente exclusivos, a sua probabilidade é evidentemente menor que a soma das probabilidades dos conjuntos parcelares.

4. Se se realizou um acontecimento dum conjunto total A primitivamente homogêneo, e se sabe que êsse acontecimento é um elemento indeterminado dum conjunto parcial A' , as probabilidades dos elementos de A alteram-se com êste conhecimento, e o conjunto A adquire uma heterogeneidade particular. É claro, com efeito,

que os elementos de A não pertencentes a A' se tornaram impossíveis, e portanto adquiriram probabilidades nulas; e que os elementos de A' , tornado verdadeiro conjunto total, não podem conservar as probabilidades que tinham antes, porque a soma destas é menor que 1. Mas estes elementos adquirem novas probabilidades iguais entre si, pois o conhecimento que se teve acerca da natureza do acontecimento realizado não introduz motivo para deixarmos de considerá-los igualmente prováveis. Se designarmos por p' o valor comum destas novas probabilidades e por m o número de elementos de A' , teremos

$$m p' = 1,$$

donde resulta

$$p' = \frac{1}{m} = \frac{1}{n} : \frac{m}{n} = \frac{p}{P_{A'}}.$$

A nova probabilidade é pois igual à primitiva dividida pela primitiva probabilidade do conjunto A' , agora tornado total.

Se A'' é um conjunto parcial de A' , será a sua nova probabilidade dada por

$$P'_{A''} = \sum_{A''} \frac{p}{P_{A'}} = \frac{P_{A''}}{P_{A'}};$$

esta probabilidade é ainda igual ao quociente da probabilidade primitiva pela probabilidade primitiva de A' . É o *princípio da divisão das probabilidades*, enunciado pela primeira vez pelo geômetra inglês Bayes num escrito póstumo⁽¹⁾ publicado pela Royal Society de Londres em 1763.

(1) *An Essay towards solving a Problem in the Doctrine of Chances*, in-*Philosophical Transactions*, Vol. LIII; Propos. 5, pág. 281.

Pode exprimir-se êste princípio dizendo que as novas probabilidades são proporcionais às antigas.

§ 2.º

**Da probabilidade
nos conjuntos finitos heterogêneos**

5. Consideremos um conjunto total homogêneo A , e sejam A'_1, A'_2, \dots, A'_s conjuntos parciais de A , sem elementos comuns quando considerados dois a dois, e tais que

$$A = A'_1 + A'_2 + \dots + A'_s.$$

Suponhamos um conjunto B de acontecimentos B_1, B_2, \dots, B_s , ligados por uma relação de causalidade, necessária ou convencional, aos acontecimentos do conjunto A , dêste modo: realiza-se o acontecimento B_i , quando se realiza um acontecimento qualquer do conjunto parcial A'_i ($i = 1, 2, \dots, s$). É claro que os acontecimentos B_i são todos possíveis (se nenhum dos conjuntos A'_i é nulo), incompatíveis e formam conjunto total. Mas o conjunto B é em geral heterogêneo. A probabilidade do elemento B_i é com efeito a probabilidade do conjunto A'_i , portanto variável com i se os conjuntos A'_i não têm todos o mesmo número de elementos.

Reciprocamente, todo o conjunto B com probabilidades elementares racionais pode considerar-se relacionado com um conjunto homogêneo A como acabamos de dizer.

Façamos

$$p_i = P_{A'_i} \quad i = 1, 2, \dots, s;$$

os números p_i são, como dissémos, as probabilidades dos

elementos B_i de B e, por ser (n.º 3)

$$1 = P_A = P_{A'_1} + P_{A'_2} + \dots + P_{A'_s},$$

tem-se

$$p_1 + p_2 + \dots + p_s = 1.$$

Como nos conjuntos homogéneos, a soma das probabilidades dos elementos dum conjunto total é a unidade.

Se B' é um conjunto parcial de B , de elementos B_1, B_2, \dots, B_h ($h < s$), a sua probabilidade é evidentemente a probabilidade do conjunto

$$A'_1 + A'_2 + \dots + A'_h;$$

esta probabilidade é igual à soma das probabilidades dos conjuntos parcelares, donde resulta

$$P_{B'} = p_1 + p_2 + \dots + p_h,$$

isto é, a probabilidade dum conjunto parcial é igual à soma das probabilidades dos seus elementos, como nos conjuntos homogéneos.

Se B'_1, B'_2, \dots, B'_k são conjuntos parciais de B , sem elementos comuns quando considerados dois a dois, e se fazemos

$$B' = B'_1 + B'_2 + \dots + B'_k,$$

será portanto

$$P_{B'} = P_{B'_1} + P_{B'_2} + \dots + P_{B'_k},$$

que é a forma geral do *princípio da adição* nos conjuntos finitos heterogéneos.

Se os conjuntos B'_i não são mutuamente exclusivos, a probabilidade da sua soma é menor que a soma das suas probabilidades.

6. Se se produziu um acontecimento do conjunto total B e se sabe que êle é um elemento indeterminado dum conjunto parcial B' , as probabilidades dos elementos de B alteram-se. Os elementos de $B - B'$ tornam-se impossíveis, e portanto adquirem probabilidades nulas. Os elementos de B' , que se tornou total, não podem conservar as probabilidades primitivas, porque a soma destas é menor que 1. É fácil calcular as novas probabilidades dos elementos de B' . Sejam como precedentemente B_1, B_2, \dots, B_h ($h < s$) êstes elementos, e A'_1, A'_2, \dots, A'_h os conjuntos parciais de A que lhes correspondem. O conhecimento que tivemos acerca da natureza do acontecimento produzido de B equivale ao conhecimento de que se produziu um elemento do conjunto homogêneo A e que êste elemento pertence ao conjunto parcial de A

$$A' = A'_1 + A'_2 + \dots + A'_h;$$

ora êste conhecimento muda a probabilidade de A'_i ($i = 1, 2, \dots, h$) em (n.º 4)

$$P'_{A'_i} = \frac{P_{A'_i}}{P_{A'}},$$

donde resulta imediatamente que a nova probabilidade de B_i é

$$p'_i = \frac{p_i}{P_{B'}} = \frac{p_i}{p_1 + p_2 + \dots + p_h}.$$

Em geral, se B'' é um conjunto parcial do novo conjunto total B' , a nova probabilidade de B'' é dada pela fórmula

$$P'_{B''} = \frac{P_{B''}}{P_{B'}},$$

que resulta da precedente por meio duma somação esten-

dida aos elementos de B'' . As novas probabilidades são pois proporcionais às antigas. Estas fórmulas encerram o *princípio da divisão* para os conjuntos finitos heterogéneos.

7. Quando só interessa um acontecimento B_1 dum conjunto total

$$B = (B_1, B_2, \dots, B_n)$$

de acontecimentos incompatíveis de probabilidades

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

usa-se, por motivo de simplicidade, reduzir a dois os elementos do conjunto B , dêste modo: um elemento do novo conjunto será o acontecimento B_1 ; o outro elemento será o acontecimento que consiste na realização dum acontecimento de B que não seja B_1 , i. e. na realização de B_2 , ou de B_3 , ... ou de B_n . Êste segundo elemento diz-se o acontecimento contraditório ⁽¹⁾ de B_1 , e tem a probabilidade

$$q_1 = p_2 + p_3 + \dots + p_n.$$

Por ser

$$p_1 + p_2 + \dots + p_n = 1,$$

resulta

$$p_1 + q_1 = 1,$$

i. e., a soma das probabilidades de dois acontecimentos contraditórios é a unidade.

⁽¹⁾ Diz-se ordinariamente *contrário*, o que é incorrecto, porque contrários a B_1 são todos os acontecimentos de B . Em Lógica a distinção entre os dois termos está feita há muito tempo e parece-nos conveniente, para evitar confusões, introduzi-la aqui.

§ 3.º

Da multiplicação das probabilidades

8. Sejam dois conjuntos totais

$$A = (A_1, A_2, \dots, A_m),$$

$$B = (B_1, B_2, \dots, B_n),$$

de acontecimentos possíveis e incompatíveis em circunstâncias determinadas. Produzir-se há nessas circunstâncias um acontecimento composto de dois, um de A e outro de B . Consideremos o conjunto de todos êstes acontecimentos compostos possíveis; êste conjunto é evidentemente formado de elementos incompatíveis, e é total; representá-lo-hemos por AB . Os seus elementos são da forma $A_i B_j$, mas não são em geral tôdas as combinações matemáticas dos elementos de A com os de B , porque algumas destas combinações podem ser de realização impossível. Deve notar-se porém que na expressão $A_i B_j$ dos elementos compostos possíveis os índices i e j tomam, pelo menos uma vez, cada um dos valores $1, 2, \dots, m$ e $1, 2, \dots, n$, respectivamente; de contrário, um pelo menos dos elementos de A , ou de B , seria impossível, contra a hipótese.

Designemos por r_{ij} as probabilidades dos elementos $A_i B_j$ de AB . Será

$$\sum_{i,j} r_{ij} = 1,$$

onde a dupla soma se estende a tôdas as combinações de realização possível dos índices i e j .

A probabilidade do elemento A_i de A é evidentemente

a probabilidade do conjunto parcial de AB formado por todos os elementos em cuja composição entra A_i ; é pois

$$P_{A_i} = \sum_j r_{ij},$$

onde a soma se estende a todos os valores de j de combinação possível com o número i . Verifica-se que é

$$\sum_i P_{A_i} = 1,$$

onde a soma se estende a todos os valores de i desde 1 até m ; como devia ser.

A probabilidade do elemento B_j de B é análogamente

$$P_{B_j} = \sum_i r_{ij},$$

onde a soma se estende a todos os valores de i compatíveis com o número j . É também

$$\sum_j P_{B_j} = 1,$$

onde a soma é feita de $j=1$ a $j=n$.

Se se sabe que A_i se produziu, a probabilidade de B_j , sendo j um número compatível com i , é, pelo princípio da divisão,

$$P_{B_j, A_i} = \frac{r_{ij}}{\sum_j r_{ij}},$$

onde o denominador é exactamente o valor de P_{A_i} ; resulta portanto que a probabilidade do acontecimento composto

$A_i B_j$ é dada por

$$(1) \quad P_{A_i B_j} = P_{A_i} \cdot P_{B_j, A_i}.$$

Verifica-se também a relação, de si evidente,

$$\sum_j P_{B_j, A_i} = 1,$$

onde a soma se estende a todos os valores de j compatíveis com o número i .

Semelhantemente, se se sabe que B_j se realizou, a probabilidade de A_i , sendo i compatível com j , é

$$P_{A_i, B_j} = \frac{r_{ij}}{\sum_i r_{ij}},$$

onde o denominador é exactamente P_{B_j} ; donde resulta

$$(2) \quad P_{A_i B_j} = P_{B_j} \cdot P_{A_i, B_j}.$$

Verifica-se, como precedentemente, a relação

$$\sum_i P_{A_i, B_j} = 1,$$

onde a soma se estende a todos os valores de i compatíveis com j .

As fórmulas (1) e (2) traduzem o *princípio da multiplicação das probabilidades*, que pode enunciar-se assim: a probabilidade dum acontecimento composto de dois outros acontecimentos, i. e. a probabilidade de que dois acontecimentos se realizem ambos, é igual ao produto da probabilidade dum qualquer destes pela probabilidade que adquire o outro na hipótese de o primeiro se ter realizado.

Este princípio não é essencialmente diferente do princípio da divisão, como se vê pelo que precede. Mas presta importantes serviços no Cálculo das Probabilidades, porque, ordinariamente, a probabilidade dum acontecimento, e a de outro na hipótese de se produzir o primeiro, são mais fáceis de calcular que a probabilidade do acontecimento composto de ambos.

9. O resultado precedente generaliza-se com facilidade. Sejam em certas circunstâncias

$$A = (A_1, A_2, \dots, A_l),$$

$$B = (B_1, B_2, \dots, B_m),$$

$$C = (C_1, C_2, \dots, C_n),$$

três conjuntos totais de acontecimentos possíveis e incompatíveis. Se essas circunstâncias se realizam, produzirse-há um acontecimento composto de três: um de A , outro de B e outro de C . Em geral, nem tôdas as combinações da forma $A_i B_j C_k$ são possíveis; mas no conjunto ABC das combinações possíveis, que é evidentemente um conjunto total de acontecimentos possíveis e incompatíveis, tomará o índice i cada um dos valores 1, 2, ..., l pelo menos uma vez, e facto análogo se dará com j e k ; de contrário, um pelo menos dos elementos de A , B ou C seria impossível. Pode dizer-se também que se produzirá um acontecimento composto de dois: um da forma AB e outro da forma C . Isto equivale a considerar com o conjunto C , em vez dos dois conjuntos A e B , o conjunto total AB dos acontecimentos possíveis da forma AB , evidentemente incompatíveis, e reduz o cálculo da probabilidade dum acontecimento composto de três ao cálculo da probabilidade dum acontecimento com-

posto de dois. Assim, designando por $P_{C,AB}$ a probabilidade do acontecimento C na hipótese da realização de ambos os acontecimentos A e B , teremos sucessivamente, pelo n.º anterior,

$$\begin{aligned} P_{ABC} &= P_{AB} \cdot P_{C,AB} \\ &= P_A \cdot P_{B,A} \cdot P_{C,AB}, \\ &= P_B \cdot P_{A,B} \cdot P_{C,BA}. \end{aligned}$$

Se, em vez de substituir aos dois conjuntos A e B o conjunto AB , substituíssemos BC a B e C , e depois CA a C e A , obteríamos ainda

$$\begin{aligned} P_{ABC} &= P_B \cdot P_{C,B} \cdot P_{A,BC}, \\ &= P_C \cdot P_{B,C} \cdot P_{A,CB}; \\ &= P_C \cdot P_{A,C} \cdot P_{B,CA}, \\ &= P_A \cdot P_{C,A} \cdot P_{B,AC}. \end{aligned}$$

Estas seis relações mostram que a probabilidade dum acontecimento composto de três outros é igual ao produto das probabilidades dos acontecimentos componentes, enumerados numa ordem qualquer, mas tomando cada probabilidade, a partir da segunda, na hipótese da realização dos acontecimentos precedentes.

Observemos que, pelo n.º anterior, é

$$\sum_i P_{A_i, B_j, C_k} = 1,$$

onde a soma se estende a todos os valores de i desde 1 a l ; e que têm lugar duas relações análogas para os acontecimentos B_j e C_k .

O teorema que acabámos de demonstrar é válido no

caso dum número finito qualquer de acontecimentos componentes; com efeito, se A, B, ..., H, K forem μ acontecimentos todos compatíveis, a relação

$$P_{AB\dots HK} = P_{AB\dots H} \cdot P_{K, AB\dots H}$$

e análogas mostram que êle vale para μ acontecimentos se valer para $\mu - 1$. É a forma geral do princípio da multiplicação das probabilidades nos conjuntos finitos.

10. Os μ acontecimentos todos compatíveis A, B, ..., H, K, L dizem-se *independentes*, quando a probabilidade de cada um deles nas diversas hipóteses possíveis sôbre a realização ou não-realização dos outros é a mesma. Representando respectivamente por A', B', ..., H', K', L' os contraditórios daqueles acontecimentos, ter-se-há então

$$(1) \quad \begin{aligned} P_{L, AB\dots HK} &= P_{L, AB\dots HK'} = \dots = P_{L, A'B\dots HK} = \\ &= P_{L, AB\dots H'K'} = \dots = P_{L, A'B'\dots H'K'}, \end{aligned}$$

e relações semelhantes para as probabilidades análogas de K, H, ..., B, A.

Ora, o princípio da multiplicação fornece a igualdade

$$(2) \quad P_{L, AB\dots H} \cdot P_{K, AB\dots HL} = P_{K, AB\dots H} \cdot P_{L, AB\dots HK},$$

visto que os produtos dos seus dois membros por $P_{AB\dots H}$ são ambos iguais a $P_{AB\dots HKL}$. Substituindo K pelo seu contraditório K', será também

$$P_{L, AB\dots H} \cdot P_{K', AB\dots HL} = P_{K', AB\dots H} \cdot P_{L, AB\dots HK'},$$

que pode escrever-se, atendendo à primeira igualdade (1)

e à observação do n.º precedente,

$$P_{L, AB \dots H} \cdot (1 - P_{K, AB \dots HL}) = (1 - P_{K, AB \dots H}) \cdot P_{L, AB \dots HK};$$

desta igualdade e de (2) resulta imediatamente

$$P_{L, AB \dots H} = P_{L, AB \dots HK}.$$

Tratando de modo análogo as outras relações (1) encontra-se

$$(3) \quad \begin{aligned} P_{L, AB \dots H} &= P_{L, AB \dots H'} = \dots = P_{L, A'B' \dots H} = \dots = \\ &= P_{L, A'B' \dots H'} = P_{L, AB \dots HK}, \end{aligned}$$

e relações semelhantes para as probabilidades de L nas diversas hipóteses sôbre a realização ou não-realização dos acontecimentos de cada uma das outras combinações dos $\mu - 1$ acontecimentos A, B, ..., H, K, $\mu - 2$ a $\mu - 2$. A aplicação reiterada dêste processo daria finalmente

$$P_L = P_{L, A} = P_{L, AB} = \dots = P_{L, AB \dots H} = P_{L, AB \dots HK},$$

e relações análogas resultantes da substituição de A por B, C, ..., K no segundo membro, ou de A B por A C, B C, ..., A K, ..., H K no terceiro membro, etc., havendo ainda para todos estes membros relações análogas a (1) e (3). Faz-se para A, B, ..., H, K o que acaba de fazer-se para L. Conclui-se que, se μ acontecimentos compatíveis são independentes, a probabilidade de qualquer deles é a mesma não só nas diversas hipóteses sôbre a realização ou não-realização dos outros, mas ainda nas diversas hipóteses sôbre a realização ou não-realização de alguns dos outros, quaisquer que sejam, e sempre igual à

probabilidade do acontecimento fora de tôdas estas hipóteses.

Resulta imediatamente que tôdas as fórmulas que exprimem o princípio da multiplicação para μ acontecimentos independentes se reduzem à fórmula única

$$P_{AB \dots KL} = P_A \cdot P_B \dots P_K \cdot P_L,$$

i. e., que a probabilidade dum acontecimento composto de acontecimentos independentes é igual ao produto das probabilidades dos acontecimentos componentes. É uma proposição muito importante.

11. Diremos que vários conjuntos totais de elementos incompatíveis

$$A = (A_1, A_2, \dots, A_\alpha),$$

$$B = (B_1, B_2, \dots, B_\beta),$$

$$\dots \dots \dots$$

$$L = (L_1, L_2, \dots, L_\lambda),$$

todos possíveis em circunstâncias determinadas, são independentes, quando tôdas as combinações $A B \dots L$ são possíveis e constituídas por acontecimentos independentes. Se designarmos respectivamente por

$$p_1, p_2, \dots, p_\alpha,$$

$$q_1, q_2, \dots, q_\beta,$$

$$\dots \dots \dots$$

$$u_1, u_2, \dots, u_\lambda$$

as probabilidades dos elementos de A, B, \dots, L , a proba-

bilidade duma combinação $AB\dots L$ é da forma

$$pq\dots u,$$

onde p terá o índice de A , q o de B , ... u o de L .

12. Os três princípios que demonstrámos são a chave da solução dos problemas de probabilidade. Trataremos com êles duas questões clássicas de grande importância nas aplicações.

A primeira é a da probabilidade das causas. Por *causa* entende-se aqui um acontecimento provável cuja realização dá a outro acontecimento uma certa probabilidade. A questão formula-se assim.

Seja C_1, C_2, \dots, C_n um conjunto total de causas possíveis e incompatíveis dum acontecimento A ; representemos por p_i a probabilidade de C_i e por q_i a probabilidade de A na hipótese da realização de C_i ($i=1, 2, \dots, n$). Se se sabe que A se realizou, qual é a probabilidade de ter sido acompanhado de C_i , ou, como costuma dizer-se, de a sua realização se ter dado por virtude da causa C_i ?

Estamos em presença de dois conjuntos totais de acontecimentos possíveis e incompatíveis: o conjunto das causas C_i e o conjunto (A, A') constituído pelo acontecimento A e pelo seu contraditório; ou, o que é o mesmo, dum conjunto total de acontecimentos compostos, da forma C_iA ou C_iA' . Pelo princípio da multiplicação, a probabilidade de C_iA é $p_i q_i$; pelo princípio da adição, a probabilidade de A é

$$p_1 q_1 + p_2 q_2 + \dots + p_n q_n;$$

finalmente, pelo princípio da divisão, a probabilidade de

$C_i A$ se A se realizou é

$$\frac{p_i q_i}{p_1 q_1 + \dots + p_n q_n}.$$

É a solução procurada.

A segunda questão é a denominada da probabilidade dos acontecimentos futuros deduzida dos acontecimentos observados. Nas condições da questão precedente, seja A por sua vez causa dum outro acontecimento B , e designemos por r_i a probabilidade que A , quando produzido pela causa C_i , dá a B ($i=1, 2, \dots, n$). Sabe-se que se realizou A e pede-se a probabilidade de que seja acompanhado de B .

Temos agora três conjuntos totais: os dois da questão anterior e o conjunto (B, B') formado por B e pelo seu contraditório; ou, o que é a mesma coisa, um conjunto total de acontecimentos compostos, da forma $C_i A B$, ou $C_i \bar{A} B'$, ou $C_i A' B$, ou $C_i A' B'$. A probabilidade de A é, como já vimos,

$$p_1 q_1 + \dots + p_n q_n;$$

a probabilidade de $C_i A B$ é $p_i q_i r_i$; a probabilidade de $A B$ é

$$p_1 q_1 r_1 + \dots + p_n q_n r_n;$$

emfim, a probabilidade de B se A se realizou é o quociente da probabilidade de $A B$ pela probabilidade de A , ou seja

$$\frac{p_1 q_1 r_1 + \dots + p_n q_n r_n}{p_1 q_1 + \dots + p_n q_n},$$

o que resolve a questão.

Se A não é causa de B , mas B tem as mesmas causas

que A , a expressão que acabamos de deduzir é ainda a probabilidade de que A seja acompanhado de B , sendo r_1, r_2, \dots, r_n as probabilidades que dão a B as diversas causas supostas.

§ 4.º

Da probabilidade nos conjuntos numeráveis

13. Estudaremos agora a probabilidade nos conjuntos numeráveis, considerados como limites de conjuntos finitos. Mais adiante daremos outra teoria desta probabilidade, na qual o conjunto numerável total sempre nos aparece como dado e os seus elementos podem ter probabilidades racionais ou irracionais.

Seja

$$M_n = (A_1, A_2, \dots, A_n, M_{n+1})$$

um conjunto finito total de acontecimentos possíveis e incompatíveis, de probabilidades

$$p_1, p_2, \dots, p_n, q_{n+1};$$

ter-se-á

$$p_1 + p_2 + \dots + p_n + q_{n+1} = 1.$$

Suponhamos M_n variável com n , de tal modo que, para cada inteiro positivo i , não variem A_i nem p_i , e que q_{n+1} tenda para zero quando n cresce infinitamente. O conjunto M_n tenderá para o conjunto numerável

$$A = (A_1, A_2, \dots, A_n, \dots).$$

Com efeito, o limite completo de M_n , quer dizer o con-

junto formado pelos elementos que figuram num número infinito de conjuntos da sucessão

$$M_1, M_2, \dots, M_n, \dots,$$

é evidentemente A , i. e.

$$\overline{\lim}_{n \rightarrow \infty} M_n = A;$$

o limite restrito de M_n , ou seja o conjunto dos elementos que figuram em todos os M_n com excepção dum número finito destes conjuntos, é ainda A , i. e.

$$\underline{\lim}_{n \rightarrow \infty} M_n = A;$$

esta igualdade e a precedente dão pois

$$\lim_{n \rightarrow \infty} M_n = A.$$

Análogamente, o conjunto das probabilidades dos elementos de M_n tenderá para o conjunto numeravel

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

constituído pelas probabilidades dos elementos de A .

É claro que o conjunto A é um conjunto total de acontecimentos possíveis e incompatíveis. E como, por ser

$$\lim_{n \rightarrow \infty} q_{n+1} = 0,$$

a série

$$p_1 + p_2 + \dots + p_n + \dots$$

converge para 1, vê-se que num conjunto numerável total a soma das probabilidades de todos os seus elementos é igual à unidade.

Reciprocamente, todo o conjunto tal como A pode ser considerado como limite duma sucessão de conjuntos tais como M_n .

14. É fácil dar um exemplo real dum tal conjunto de acontecimentos. No jôgo de «cara ou X», a «cara» pode aparecer pela primeira vez no primeiro lançamento, ou no segundo, ou no terceiro, e assim indefinidamente. Se consideramos como um acontecimento a saída de «cara» pela primeira vez no lançamento de ordem i ($i = 1, 2, 3, \dots$), estamos em presença dum conjunto numerável total de acontecimentos possíveis e incompatíveis. As probabilidades dos seus elementos são respectivamente

$$\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots$$

e a soma de todos êstes números é igual a 1. Vê-se imediatamente como êste conjunto é limite duma sucessão de conjuntos finitos, igualmente reais. O termo geral desta sucessão é o conjunto formado pelos $n+1$ acontecimentos

«cara» no 1.^o lançamento,
 «cara» no 2.^o lançamento,

 «cara» no n .^o lançamento,
 X no n .^o lançamento,

que constituem uma totalidade finita de acontecimentos possíveis e incompatíveis, de probabilidades respectiva-

mente iguais a

$$\frac{1}{2}, \frac{1}{2^2}, \dots, \frac{1}{2^n}, \frac{1}{2^n};$$

a soma destes números é 1, como tinha de ser, e o último tende para zero.

15. Definiremos a probabilidade dum conjunto contido em A como limite da probabilidade dum conjunto contido em M_n que tenda para o primeiro. Esta definição concorda com o facto de ser 1 a probabilidade do conjunto A . Definiremos também a probabilidade dum elemento de $A' < A$, se se realizou um elemento indeterminado de A' , como limite da probabilidade do mesmo elemento em M_n , se se realizou um elemento indeterminado dum conjunto parcial de M_n que tenda para A' .

Seja A' um conjunto finito contido em A . A partir dum certo valor de n , todos os elementos de A' estão contidos no conjunto finito M_n . Em M_n , a probabilidade do conjunto A' é igual à soma das probabilidades dos elementos que o constituem. Passando ao limite, conclui-se que a probabilidade dum conjunto finito parcial de A é igual à soma das probabilidades dos seus elementos. É o *princípio da adição finita das probabilidades* nos conjuntos totais numeráveis.

Seja

$$A' = (A_{i_1}, A_{i_2}, \dots, A_{i_m}, \dots),$$

$$i_1 < i_2 < \dots < i_m < \dots,$$

um conjunto numerável parcial de A . A partir dum certo valor de n todos os elementos do conjunto finito

$$A'_m = (A_{i_1}, A_{i_2}, \dots, A_{i_m}),$$

onde m é arbitrariamente grande, estão contidos no con-

junto finito M_n . Em M_n , a probabilidade de A'_m é

$$p_i + p_{i_1} + \dots + p_{i_m}.$$

Se agora cresce m indefinidamente, o mesmo sucede a i_m , e portanto a n ; M_n tende para A , A'_m tende evidentemente para A' , e a probabilidade de A'_m tende para a soma da série

$$p_i + p_{i_1} + \dots + p_{i_m} + \dots,$$

que converge para um número menor que 1. Logo, a probabilidade dum conjunto numerável A' contido em A é igual à soma das probabilidades dos seus elementos. É o *principio da adição infinita das probabilidades*.

16. Suponhamos que se realizou um acontecimento de A , e que se sabe que êle é um elemento indeterminado dum conjunto parcial A' de A . Os elementos de A estranhos a A' tornam-se impossíveis, adquirindo probabilidades nulas. O conjunto A' torna-se total; os seus elementos não podem pois conservar as probabilidades primitivas, porque a soma delas é menor que 1. Vejamos como se modificam estas probabilidades.

Se o conjunto A' é finito, a partir dum certo valor de n todos os seus elementos estão contidos no conjunto finito total M_n . A nova probabilidade dum elemento A_i de A' é, em M_n , igual a

$$\frac{p_i}{\sum_{A'} p_i},$$

i. e. independente de n ; passando ao limite, é ainda esta fracção o valor da nova probabilidade de A_i em A . A nova probabilidade de cada elemento do novo conjunto

total A' é pois o quociente da probabilidade primitiva d'êste elemento pela primitiva probabilidade de A' :

$$p'_i = \frac{p_i}{P_{A'}}.$$

Se A'' é um conjunto parcial de A' , a nova probabilidade de A'' , pelo princípio da adição nos conjuntos finitos, é

$$P'_{A''} = \sum_{A''} p'_i = \frac{P_{A''}}{P_{A'}},$$

i. e., igual ao quociente das antigas probabilidades de A'' e A' .

Se o conjunto A' é infinito,

$$A' = (A_{i_1}, A_{i_2}, \dots, A_{i_m}, \dots),$$

$$i_1 < i_2 < \dots < i_m < \dots,$$

a partir dum certo valor de m o elemento A_{i_h} , onde h é arbitrariamente dado, está contido no conjunto finito

$$A'_m = (A_{i_1}, A_{i_2}, \dots, A_{i_m});$$

e a êsse valor de m corresponde um valor de n tal que, a partir dêle, A'_m está contido no conjunto finito M_n . Em M_n , a probabilidade de A_{i_h} , se se realizou um elemento indeterminado de A'_m , é

$$\frac{p_{i_h}}{P_{A'_m}}.$$

Façamos crescer m infinitamente; o mesmo sucederá a i_m , e portanto a n ; M_n tenderá para A , A'_m para A' , $P_{A'_m}$ para $P_{A'}$; logo, a probabilidade de A_{i_h} se se realizou um ele-

mento indeterminado de A' é

$$p'_{i_h} = \frac{P_{i_h}}{P_{A'}}.$$

Verifica-se que é igual a 1 a soma de tôdas estas probabilidades.

Se A'' é um conjunto finito ou infinito contido no novo conjunto total A' , será, pelo princípio da adição nos conjuntos numeráveis,

$$P'_{A''} = \sum_{A''} p'_{i_h} = \frac{P_{A''}}{P_{A'}},$$

i. é, a probabilidade dum conjunto A'' contido em A' , se se realizou um elemento indeterminado de A' , é o quociente das primitivas probabilidades de A'' e A' .

Os resultados que acabamos de obter constituem o *princípio da divisão das probabilidades* nos conjuntos numeráveis.

17. As séries que dão as probabilidades dos conjuntos numeráveis são absolutamente convergentes, porque os seus termos são positivos. A probabilidade dum conjunto numerável é portanto independente da ordem dos seus elementos.

18. Sejam, em circunstâncias determinadas,

$$A = (A_1, A_2, \dots, A_m, \dots),$$

$$B = (B_1, B_2, \dots, B_n, \dots)$$

dois conjuntos totais numeráveis, ou um finito e outro numerável, de acontecimentos possíveis e incompatíveis. Se essas circunstâncias se realizam, produzir-se há um

acontecimento composto de um de A e outro de B . A expressão geral $A_i B_j$ deste acontecimento composto não conterá geralmente tôdas as combinações matemáticas dos elementos de A com os de B , porque algumas destas combinações podem ser impossíveis; mas i tomará nela cada um dos valores $1, 2, \dots, m, \dots$ pelo menos uma vez e coisa análoga sucederá com j , pois são possíveis todos os elementos de A e de B . Os acontecimentos compostos possíveis $A_i B_j$ formarão um conjunto total numerável AB ; designemos por r_{ij} as suas probabilidades. Teremos imediatamente, pelo princípio da adição,

$$P_{A_i} = \sum_j r_{ij}, \quad P_{B_j} = \sum_i r_{ij},$$

onde a primeira soma se estende a todos os valores de j de composição possível com o número i e a segunda a todos os valores de i de composição possível com o número j . Por ser

$$\sum_i \sum_j r_{ij} = 1,$$

verifica-se que é

$$\sum_i P_{A_i} = \sum_j P_{B_j} = 1.$$

Será também, pelo princípio da divisão,

$$P_{A_i, B_j} = \frac{r_{ij}}{P_{B_j}}, \quad P_{B_j, A_i} = \frac{r_{ij}}{P_{A_i}},$$

e verifica-se que é

$$\sum_i P_{A_i, B_j} = \sum_j P_{B_j, A_i} = 1;$$

as duas fórmulas precedentes dão

$$P_{A_i B_j} = P_{A_i} \cdot P_{B_j, A_i} = P_{B_j} \cdot P_{A_i, B_j},$$

o que exprime o *princípio da multiplicação das probabilidades* nos conjuntos numeráveis: a probabilidade dum acontecimento composto de dois é igual ao produto da probabilidade dum dêles pela probabilidade que adquire o outro se o primeiro se realizou.

19. Repetindo textualmente o que fizemos no n.º 9, estende-se o princípio da multiplicação a qualquer número finito de acontecimentos pertencentes a conjuntos totais, dos quais um pelo menos é numerável.

20. Num conjunto total numerável A pode considerar-se também, ao lado dum dos seus acontecimentos A , o acontecimento contraditório A' , que consiste na realização dum acontecimento indeterminado do conjunto parcial $A - A$. A soma das probabilidades de dois acontecimentos contraditórios é evidentemente 1.

21. Vários acontecimentos todos compatíveis A, B, \dots, K, L , elementos de conjuntos totais dos quais um pelo menos é numerável, dizem-se *independentes* se a probabilidade de cada um dêles nas diversas hipóteses possíveis sobre a realização ou não-realização dos outros é a mesma. Repetindo textualmente o que se contém no n.º 10, obtêm-se os mesmos resultados, a saber: a probabilidade de cada um dos acontecimentos A, B, \dots, L é a mesma não só nas diversas hipóteses sobre a realização ou não-realização dos outros, mas também nas diversas hipóteses sobre a realização ou não-realização de alguns dos outros, e sempre igual à probabilidade do mesmo acontecimento fora de

tôdas estas hipóteses; a probabilidade do acontecimento composto $AB \dots L$ é igual ao produto das probabilidades dos acontecimentos componentes:

$$P_{AB \dots L} = P_A \cdot P_B \dots P_L.$$

22. Vários conjuntos totais, finitos ou numeráveis,

$$A = (A_1, A_2, \dots),$$

$$B = (B_1, B_2, \dots),$$

$$\dots \dots \dots$$

$$L = (L_1, L_2, \dots),$$

dizem-se independentes quando tôdas as combinações da forma $AB \dots L$ são possíveis e constituídas por acontecimentos independentes. A probabilidade de cada uma destas combinações é o produto das probabilidades dos seus elementos.

23. Os conjuntos numeráveis cuja probabilidade acabamos de estudar são heterogéneos, e observámos que tais conjuntos existem efectivamente. Um conjunto total numerável só pode ser homogéneo se as probabilidades dos seus elementos forem tôdas nulas; mas então a probabilidade da realização dum elemento indeterminado dêsse conjunto, que corresponde a uma certeza e é portanto 1, não será igual à soma

$$0 + 0 + \dots$$

das probabilidades dos seus elementos, que é nula. O princípio da adição deixa de ser geralmente verdadeiro, o que mostra que tais conjuntos ocupam na teoria um lugar excepcional. Note-se ainda que a probabilidade nula deixa,

em geral, de significar impossibilidade, como a probabilidade 1 deixa, em geral, de significar certeza.

Faz-se o estudo da probabilidade nestes conjuntos considerando-os como limites de conjuntos homogêneos finitos. Seja

$$A_n = (A_1, A_2, \dots, A_n)$$

um conjunto finito homogêneo total, de elementos possíveis e incompatíveis. A probabilidade dum elemento determinado dêste conjunto é $\frac{1}{n}$. Suponhamos A_n variável, sem que a homogeneidade se destrua e de modo que A_{n+1} difira de A_n pela introdução dum elemento novo A_{n+1} . O conjunto A_n tenderá para o conjunto numerável

$$A = (A_1, A_2, \dots, A_n, \dots)$$

e as probabilidades dos elementos de A_n tenderão para zero. É claro que A é um conjunto total de elementos possíveis e incompatíveis, e que não há razão para lhe atribuir heterogeneidade. As probabilidades dos seus elementos são portanto tôdas nulas, e o princípio da adição não se verifica para o conjunto A , como já fizemos notar. Vê-se que a probabilidade dum elemento de A é o limite da sucessão das probabilidades do mesmo elemento nos conjuntos A_n que o contêm.

Definiremos a probabilidade dum conjunto A' contido em A como sendo o limite da probabilidade dum conjunto contido em A_n , que tenda para A' . Isto concorda com a atribuição do valor 1 à probabilidade do conjunto total, feita atrás.

Se A' é finito, tem uma probabilidade nula. Com efeito, a partir dum certo valor de n , A' estará contido

em A_n ; a sua probabilidade em A_n é

$$\frac{m}{n},$$

designando por m o número dos seus elementos; m é fixo; se n cresce ilimitadamente, esta fracção tende para zero.

Como uma soma de zeros é zero, vale neste caso o princípio da adição.

Se A' é infinito, a sua probabilidade pode não existir, ou ser um número do intervalo fechado $(0,1)$. Esta probabilidade depende essencialmente da ordem atribuída aos elementos do conjunto total. Isto confirma o carácter excepcional dos conjuntos totais numeráveis homogéneos.

Por exemplo, para o conjunto total homogéneo dos números inteiros positivos na ordem natural, a probabilidade do conjunto das potências de 10 é nula, a do conjunto dos números pares é $\frac{1}{2}$, a do conjunto dos números que não são potências de 10 é 1; a do conjunto que é soma dos conjuntos dos números de 1 a 5, de 10 a 50, de 100 a 500, etc. não existe, porque a respectiva sucessão das probabilidades nos conjuntos A_n tem pelo menos os dois valores de acumulação $\frac{4}{9}$ e $\frac{8}{9}$, como é fácil de ver.

Sem dificuldade se demonstraria que, se vários conjuntos parciais finitos ou infinitos sem elementos comuns dois a dois têm probabilidade, é válido o princípio da adição relativamente a estes conjuntos e à sua soma se eles são em número finito; se são em infinidade numerável, pode o princípio ser válido ou não. Em particular, se se decompõe o conjunto total dado numa infinidade numerável de conjuntos parciais incompatíveis dotados de probabilidade, a série das probabilidades destes conjuntos é convergente, mas pode ter ou não ter por soma 1, mesmo

para uma ordem determinada dos elementos do conjunto e ainda que os conjuntos parciais sejam todos infinitos.

Por exemplo, o conjunto total dos números naturais pode decompor-se na seguinte infinidade numerável de conjuntos infinitos parciais incompatíveis: o dos números ímpares, o dos produtos dos números ímpares por 2, o dos produtos dos números ímpares por 2^2 , etc. Se se supõem os elementos do conjunto total ordenados naturalmente, e se êle é homogéneo, as probabilidades dêstes conjuntos parciais são respectivamente

$$\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots$$

e têm por soma 1. Decomponhamos agora o conjunto total nos conjuntos parciais seguintes: o das potências de 2, incluindo a unidade, o das potências de 3, o das potências de 5, o das potências de 6, etc., partindo sempre dos números que não são potências e não excluindo nenhum. Estes conjuntos são incompatíveis, infinitos e em infinidade numerável, e as suas probabilidades são tôdas nulas; portanto a sua soma é zero e não 1.

A-pesar-de tôdas estas irregularidades, digamos assim, a proposição que acabamos de indicar permitiria construir a teoria da probabilidade nos conjuntos numeráveis heterogéneos a partir dos homogéneos. Esta seria a marcha paralela à que adoptámos para os conjuntos finitos, e que adoptaremos para os conjuntos da potência do contínuo. Acresce que a questão ficaria tratada com tôda a generalidade, porque as probabilidades elementares, sendo limites de sucessões de números racionais, seriam racionais ou irracionais. Mas opõe-se a que procedamos assim o carácter de excepção dos conjuntos numeráveis homogéneos, que já evidenciámos, e o facto de estes conjuntos só se apre-

sentarem efectivamente em probabilidade contínua, e ainda aí anormalmente, como adiante observaremos. Êste caso, para mais, está longe de ter a importância que possui o dos conjuntos numeráveis heterogêneos; deixá-lo hemos de lado.

§ 5.º

**Da probabilidade
nos conjuntos homogêneos
de medida positiva**

24. Seja E_n um espaço riemanniano de n dimensões e A um conjunto, limitado ou não, de pontos A dêste espaço, dotado de medida lebesguiana n -dimensional positiva a . Suponhamos que, em circunstâncias determinadas, é possível para cada ponto de A que com êle fique coincidindo um ponto invariável duma figura Φ móvel em E_n , com deformação ou sem ela, e que esta possibilidade existe só para os pontos de A . O conjunto de tais coincidências é um conjunto total de acontecimentos possíveis e incompatíveis, com a potência do contínuo ⁽¹⁾. Para abreviar, em vez de coincidência possível ou realizada do ponto invariável da figura Φ com o ponto A de A diremos simplesmente ponto possível ou realizado A .

Se os pontos do conjunto A possuem todos a mesma

⁽¹⁾ Se um conjunto A tem medida exterior finita, a sua medida interior é o limite superior das medidas dos conjuntos fechados limitados contidos em A . Se esta medida interior é positiva, os conjuntos fechados limitados contidos em A não podem ser todos finitos ou numeráveis, porque então a medida interior de A seria nula. Há portanto em A um conjunto fechado limitado não numerável e sabe-se como Cantor demonstrou que um tal conjunto tem a potência do contínuo.

possibilidade, i. é, se para cada par de tais pontos não existe motivo que nos incline a admitir a realização dum dêles de preferência à do outro, denominaremos o conjunto A *homogêneo* e diremos que o número

$$\frac{1}{a}$$

é a sua *densidade de probabilidade*.

Consideremos um conjunto mensurável A' contido em A . Ao integral lebesguiano da densidade de probabilidade estendido a A' chamaremos probabilidade da realização dum ponto indeterminado de A' , ou simplesmente probabilidade do conjunto mensurável A' , e escreveremos

$$P_{A'} = \int_{A'} \frac{1}{a} dA.$$

Por ser

$$\int_{A'} \frac{1}{a} dA = \frac{a'}{a},$$

onde a' representa a medida n -dimensional de A' , resulta que a probabilidade dum conjunto mensurável parcial é igual ao quociente da sua medida pela medida do conjunto total.

Um conjunto parcial de medida nula tem uma probabilidade nula e reciprocamente. Em particular, um conjunto parcial cujos pontos estão todos situados numa região m -dimensional, sendo $m < n$, tem uma probabilidade igual a zero.

É claro que os pontos dum conjunto parcial de medida nula são possíveis como os outros. Em probabilidade contínua, a probabilidade nula deixa pois, em geral, de

significar uma certeza matemática, correspondendo-lhe apenas a certeza moral, ou prática, da não-realização do acontecimento que a possui.

Semelhantemente, é igual à unidade a probabilidade dum conjunto parcial da mesma medida que o conjunto total, e por isso a probabilidade unidade deixa de significar aqui, em geral, a certeza absoluta da realização do acontecimento. As probabilidades 0 e 1 só significam certeza matemática nos casos extremos do conjunto nulo e do conjunto total.

Sobre o modo de definir a densidade de probabilidade dum conjunto homogéneo e a probabilidade dum conjunto parcial poderiam fazer-se considerações análogas às que apresentámos, a propósito de questões semelhantes, em probabilidade descontínua.

A probabilidade relativa a um conjunto total de medida positiva aparece-nos susceptível de todos os valores do intervalo fechado (0,1), racionais ou irracionais, contrariamente ao que sucede nos conjuntos totais finitos ou numeráveis estudados precedentemente. Esta observação permitir-nos há introduzir para os conjuntos totais finitos ou numeráveis as probabilidades irracionais, como adiante veremos.

25. Seja $A_1', A_2', \dots, A_m', \dots$ uma infinidade numerável de conjuntos parciais de A , sem pontos comuns quando considerados dois a dois, de medidas $a_1', a_2', \dots, a_m', \dots$.
A soma

$$A' = A_1' + A_2' + \dots$$

dêstes conjuntos tem por medida a série convergente de termos não negativos

$$a_1' + a_2' + \dots$$

e portanto a probabilidade de A' é dada pela fórmula

$$P_{A'} = \frac{a_1' + a_2' + \dots}{a},$$

que se pode escrever

$$P_{A'} = P_{A_1'} + P_{A_2'} + \dots$$

Nesta igualdade consiste o *princípio da adição infinita das probabilidades*, que deixamos demonstrado para conjuntos parciais mutuamente exclusivos. Mas êle é válido também para conjuntos parciais não exclusivos, como facilmente se compreende, desde que, para cada par de conjuntos parciais, o conjunto dos pontos comuns tenha medida nula.

Se, em particular, os conjuntos A_i' são todos nulos a partir de certa ordem, tem-se o *princípio da adição finita das probabilidades*.

26. Se, tendo-se realizado um ponto de A , se sabe que êle pertence a um conjunto parcial A' de medida positiva a' menor que a , êste conhecimento altera a densidade de probabilidade do conjunto A , que adquire uma heterogeneidade particular. Os pontos do conjunto A' , complementar de A' , que tem a medida positiva $a - a'$, tornam-se impossíveis, e a densidade de probabilidade de A nos pontos de A' torna-se nula. Os pontos de A' tornam-se os únicos possíveis, i. é, A' torna-se verdadeiro conjunto total, e portanto a densidade de probabilidade de A em A' não pode conservar-se igual a $\frac{1}{a}$, porque $\frac{a'}{a}$ é menor do que 1. A nova densidade de probabilidade determina-se imediatamente se atendermos a que A' , tornado total, não perdeu a homogeneidade, pois

o conhecimento adquirido que lhe alterou a densidade de probabilidade não introduz motivo para deixarmos de considerar esta constante. Assim, a nova densidade de probabilidade é evidentemente $\frac{1}{a'}$; e se A'' é um conjunto parcial de A' , de medida a'' , a sua nova probabilidade será

$$P'_{A''} = \frac{a''}{a'} = \frac{a''}{a} : \frac{a'}{a} = \frac{P_{A''}}{P_{A'}}$$

portanto igual ao quociente da sua probabilidade primitiva pela primitiva probabilidade do novo conjunto total. Êste enunciado é o *princípio da divisão* das probabilidades.

Se o novo conjunto total A' tem a mesma medida que A , a densidade de probabilidade mantém-se invariável e as novas probabilidades não diferem das antigas.

Se o novo conjunto total tem medida nula, a nova densidade de probabilidade não existe. Da probabilidade nos conjuntos totais de medida nula trataremos mais tarde, em artigo especial.

§ 6.º

Da probabilidade nos conjuntos heterogêneos de medida positiva

27. Suponhamos que entre os pontos M do espaço E_n e os pontos N de outro espaço n -dimensional F_n se estabeleceu uma correspondência biunívoca e contínua definida por

$$N = \varphi(M), \quad M = \phi(N);$$

cada uma destas relações simboliza n equações que expri-

mem as coordenadas dum ponto genérico dum dos espaços em funções contínuas das coordenadas do ponto correspondente do outro. Admitamos que tôdas estas funções possuem derivadas parciais contínuas de primeira e de segunda ordem. Sendo igual a 1 o produto dos dois determinantes funcionais

$$\frac{\partial(\varphi)}{\partial(M)} \quad \text{e} \quad \frac{\partial(\phi)}{\partial(N)},$$

necessariamente finitos, nenhum dêles pode anular-se nem, por conseguinte, mudar de sinal.

Dêste modo, aos pontos A dum conjunto A de E_n corresponderão biunívocamente os pontos B dum conjunto B de F_n . Vamos demonstrar uma proposição fundamental relativa à mensurabilidade dêstes conjuntos.

É claro que a um conjunto aberto dum dos dois espaços corresponde um conjunto aberto do outro. Se a correspondência entre E_n e F_n fôr tal que a medida dum conjunto aberto de E_n seja finita quando fôr finita a medida do conjunto correspondente de F_n e reciprocamente, cada um dos conjuntos A e B será mensurável de medida finita quando o fôr o outro.

Com efeito, seja por exemplo A mensurável de medida finita, e consideremos em E_n um conjunto aberto R de medida finita que contenha A . Ao conjunto R corresponde em F_n um conjunto aberto S que contém B ; e como a medida de S é também finita, vê-se já que é finita a medida exterior de B . Sabe-se que um conjunto aberto é soma duma infinidade numerável de regiões fechadas mensuráveis que não têm pontos comuns senão nas fronteiras. Se concebermos o conjunto aberto R decomposto dêste modo em regiões fechadas α , ficará o conjunto aberto S decomposto numa infinidade numerável de regiões fechadas β , correspondentes àquelas, com pontos comuns

só nas fronteiras respectivas e também mensuráveis. Supondo, para maior simplicidade, que E_n e F_n estão ambos referidos a coordenadas para as quais é igual a 1 o discriminante da forma quadrática que exprime o quadrado do elemento linear, o que é sempre possível, tem-se

$$\beta' = \int_{\alpha} \left| \frac{\partial(\varphi)}{\partial(M)} \right| dM$$

e portanto

$$\sum \beta' = \int_{\Sigma \alpha} \left| \frac{\partial(\varphi)}{\partial(M)} \right| dM,$$

onde β' representa a medida da região β e as somas se estendem à infinidade numerável dos α e dos β' . A primeira fórmula é bem conhecida com integral de Riemann; é pois exacta com integral de Lebesgue, porque nela os dois integrais são idênticos, e disto resulta a exactidão da segunda, porque o integral de Lebesgue é completamente aditivo. A segunda fórmula pode escrever-se também, representando por s a medida de S , dêste modo:

$$s = \int_R \left| \frac{\partial(\varphi)}{\partial(M)} \right| dM.$$

A medida exterior de B é o limite inferior dos números tais como s . Por ser $R > A$, temos evidentemente

$$s \geq \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA.$$

Mas temos também

$$s - \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA = \int_{R-A} \left| \frac{\partial(\varphi)}{\partial(M)} \right| dM;$$

notando que a medida de A é o limite inferior das medidas dos conjuntos R e que o integral de Lebesgue é uma função de conjunto absolutamente contínua, vê-se que se pode escolher R de modo que o segundo membro desta igualdade seja inferior a um número positivo arbitrário ε ; a medida do conjunto S correspondente satisfará então à desigualdade

$$s < \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA + \varepsilon,$$

que, com a desigualdade precedente, prova que a medida exterior de B é dada pelo integral

$$m_e B = \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA.$$

A medida interior de B é

$$m_i B = s - m_e(S - B),$$

portanto igual a

$$s - \int_{R-A} \left| \frac{\partial(\varphi)}{\partial(M)} \right| dM = \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA = m_e B.$$

O conjunto B é pois mensurável de medida finita

$$m B = \int_A \left| \frac{\partial(\varphi)}{\partial(A)} \right| dA,$$

e a proposição fica demonstrada.

Em particular, se A é de medida nula, também B é de medida nula. Se A é de medida positiva, também B é de medida positiva. Se A' é um conjunto parcial de A

da mesma medida que A , o conjunto B' correspondente a A' tem a mesma medida que B .

28. Seja A um conjunto total homogêneo de E_n , de medida positiva a , F_n um espaço cujos pontos têm com os de E_n a correspondência definida no n.º precedente, e B o conjunto de F_n correspondente a A ; B terá também medida positiva. A correspondência biunívoca existente entre os pontos dos dois conjuntos A e B permite dizer que se realiza o ponto B de B quando se realiza o ponto A de A correspondente a B , e reciprocamente. Pode de resto supor-se uma figura Ψ móvel em F_n , com deformação ou sem ela, e relacionada com a figura Φ móvel em E_n de tal modo que um ponto invariável de Ψ fique coincidindo com B quando o ponto invariável de Φ ficar coincidindo com A . O conjunto das coincidências do ponto invariável de Ψ com os pontos de B será então um conjunto total de acontecimentos possíveis e incompatíveis, e a probabilidade da realização dum ponto indeterminado dum conjunto mensurável B' contido em B é igual à probabilidade do consurável A' de A correspondente a B' :

$$P_{B'} = P_{A'} = \frac{a'}{a},$$

onde a' designa a medida de A' . Pelo n.º precedente, esta medida é dada pela fórmula

$$a' = \int_{B'} \left| \frac{\partial(\phi)}{\partial(B)} \right| dB$$

e temos portanto

$$P_{B'} = \int_{B'} \frac{1}{a} \left| \frac{\partial(\phi)}{\partial(B)} \right| dB.$$

À função positiva

$$f(B) = \frac{1}{a} \left| \frac{\partial(\phi)}{\partial(B)} \right|$$

chamaremos *densidade de probabilidade* do conjunto B . Diremos que este conjunto é *heterogéneo* se esta função se não reduz a uma constante.

A probabilidade do conjunto mensurável parcial B' exprime-se pois pela fórmula

$$P_{B'} = \int_{B'} f(B) dB,$$

onde $f(B)$ é a densidade de probabilidade no conjunto heterogéneo total B de medida positiva. Esta função é positiva, contínua e integrável em B . Adiante veremos como se podem introduzir densidades de probabilidade descontínuas.

Se B' tem medida nula, a sua probabilidade é zero; se B' tem a medida de B a sua probabilidade é 1. Em particular, é zero a probabilidade dum conjunto nulo e 1 a do conjunto total; nestes dois casos a probabilidade é certeza.

29. Se B'_1, B'_2, \dots são conjuntos mensuráveis parciais de B , em infinidade numerável, sem pontos comuns quando considerados dois a dois, então, pondo

$$B' = B'_1 + B'_2 + \dots,$$

tem-se, em virtude da aditividade completa do integral de Lebesgue,

$$P_{B'} = P_{B'_1} + P_{B'_2} + \dots;$$

é o *princípio da adição infinita das probabilidades*,

Se, em particular, a partir dum valor de i , os conjuntos B_i são todos nulos, tem-se o *princípio da adição finita*.

30. Na hipótese de se ter realizado um ponto do conjunto total B e de se saber que êle é um ponto indeterminado dum conjunto B' de medida positiva contido em B , a densidade de probabilidade de B , em geral, altera-se. Os pontos do conjunto $B - B'$ tornam-se impossíveis; B' torna-se total. Ora o conhecimento havido acerca do ponto de B realizado equivale ao conhecimento de que se realizou um ponto do conjunto homogéneo A de E_n , e de que êste ponto é um ponto indeterminado do conjunto A' , também de medida positiva, contido em A e correspondente a B' . Se A'' é o conjunto de E_n correspondente a um conjunto mensurável B'' contido em B' , A'' será mensurável e contido em A' , e a sua probabilidade muda-se em

$$\frac{P_{A''}}{P_{A'}}$$

fracção cujos termos são as probabilidades primitivas de A'' e A' . O conjunto B'' adquiriu portanto a probabilidade

$$P_{B''} = \frac{P_{B''}}{P_{B'}} = \int_{B''} \frac{1}{P_{B'}} f(B) dB,$$

onde $P_{B''}$ e $P_{B'}$ representam as primitivas probabilidades de B'' e B' . Vê-se que a densidade de probabilidade em B' se alterou se é $P_{B'} < 1$, i. é, se a medida positiva de B' é menor que a de B . A primeira igualdade estabelece para o conjunto heterogéneo B o *princípio da divisão das probabilidades*.

§ 7.º

Das probabilidades irracionais
nos conjuntos finitos ou numeráveis

31. Precedentemente, estudámos os conjuntos finitos e os conjuntos numeráveis com probabilidades elementares racionais. Os conjuntos finitos ou numeráveis com probabilidades elementares racionais ou irracionais introduzem-se como segue.

Seja, num espaço E_n , um conjunto total A de pontos possíveis e incompatíveis, de medida n -dimensional positiva e de densidade de probabilidade contínua. Decomponhamos êste conjunto num número finito ou numa infinidade numerável de conjuntos de medida positiva

$$A'_1, A'_2, \dots, A'_n, \dots$$

sem pontos comuns quando considerados dois a dois, e suponhamos um conjunto de acontecimentos

$$B = (B_1, B_2, \dots, B_n, \dots)$$

em correspondência biunívoca com os A'_i definida dêste modo: quando se realiza um ponto qualquer de A'_i , realiza-se *ipso facto* o acontecimento B_i ($i=1, 2, \dots, n \dots$). Ê claro que B é um conjunto total de acontecimentos possíveis e incompatíveis, finito ou numerável; e tem-se, representando por p_i a probabilidade do elemento B_i ,

$$p_i = P_{A'_i}.$$

Os números p_i são todos positivos, e podem ser racionais ou irracionais. Como é

$$1 = P_A = P_{A'_1} + P_{A'_2} + \dots,$$

tem-se

$$p_1 + p_2 + \dots = 1.$$

Reciprocamente, todo o conjunto total finito ou numerável tal como B pode considerar-se relacionado desta maneira com um conjunto de pontos tal como A , decomposto convenientemente em conjuntos parciais incompatíveis.

Se fôr

$$B' = (B_{i_1}, B_{i_2}, \dots)$$

um conjunto parcial de B , finito ou numerável, a probabilidade da realização dum elemento indeterminado de tal conjunto é a probabilidade de que se realize um ponto indeterminado do conjunto de medida positiva

$$A' = A'_{i_1} + A'_{i_2} + \dots;$$

tem-se portanto

$$P_{B'} = p_{i_1} + p_{i_2} + \dots,$$

que é o *princípio da adição*, finita ou infinita, das probabilidades.

32. Na hipótese de se ter realizado um elemento de B e de se saber que êle é um elemento indeterminado do conjunto parcial B' , ou, o que é o mesmo, na hipótese de se ter realizado um ponto de A e de se saber que êle é um ponto indeterminado de A' , a probabilidade dum conjunto B'' contido em B' é manifestamente a probabilidade

do conjunto de medida positiva A'' soma dos conjuntos A' , correspondentes aos elementos de B'' , o qual está contido em A' ; é pois

$$P'_{B''} = P'_{A'} = \frac{P_{A''}}{P_{A'}} = \frac{P_{B''}}{P_{B'}} ,$$

resultado que traduz o *princípio da divisão* das probabilidades.

33. A repetição textual do que escrevemos nos n.ºs 18 e 19 conduz ainda no presente caso ao *princípio da multiplicação* sob a forma mais geral. Exactamente como no n.º 21 se formula no caso presente a noção de independência de vários acontecimentos compatíveis, e se demonstra a forma particular que assume o princípio da multiplicação para acontecimentos independentes. A noção de independência de vários conjuntos totais finitos ou numeráveis formula-se também como no n.º 22.

§ 8.º

Da composição dos conjuntos mensuráveis

34. Para o estudo da multiplicação das probabilidades nos conjuntos de medida positiva é indispensável que demonstremos previamente alguns teoremas relativos ao que denominamos *composição* ⁽¹⁾ dos conjuntos mensuráveis.

Sejam E_m e E_n dois espaços riemannianos de m e n

(1) Seria adequado o termo *multiplicação*, mas originaria confusões nos conjuntos de números.

dimensões respectivamente, e

$$ds^2 = \sum_{\alpha, \beta} g_{\alpha\beta} dx_{\alpha} dx_{\beta},$$

$$ds^2 = \sum_{\gamma, \delta} h_{\gamma\delta} dy_{\gamma} dy_{\delta}$$

as respectivas formas quadráticas fundamentais.

Ao espaço riemanniano E_{m+n} de $m+n$ dimensões, definido pela forma quadrática

$$ds^2 = \sum_{\alpha, \beta} g_{\alpha\beta} dx_{\alpha} dx_{\beta} + \sum_{\gamma, \delta} h_{\gamma\delta} dy_{\gamma} dy_{\delta},$$

chamaremos *espaço composto dos espaços* E_m e E_n , e escreveremos

$$E_{m+n} = E_m \cdot E_n.$$

É claro que o resultado da composição de dois espaços é independente das coordenadas a que estão referidos, pois uma mudança de coordenadas em E_m e outra em E_n dão uma mudança de coordenadas particular em E_{m+n} .

Por exemplo, a composição de dois espaços lineares

$$ds^2 = dx^2, \quad ds^2 = dy^2$$

dá o plano euclidiano

$$ds^2 = dx^2 + dy^2;$$

a composição do plano euclidiano com um espaço linear

$$ds^2 = dz^2$$

dá o espaço euclidiano tridimensional

$$ds^2 = dx^2 + dy^2 + dz^2;$$

a composição de dois espaços euclidianos, um de m e outro de n dimensões, dá o espaço euclidiano de $m+n$ dimensões.

Um ponto de E_{m+n} tem as coordenadas dum ponto M de E_m e dum ponto N de E_n . Diremos que o primeiro ponto é composto dos segundos, e representá-lo hemos por $M.N$.

Reciprocamente, a composição dum ponto de E_m com um ponto de E_n dá um ponto de E_{m+n} . Os pontos de E_{m+n} compostos de todos os pontos dum conjunto A de E_m com todos os pontos dum conjunto B de E_n formam um conjunto que diremos composto daqueles dois e que representaremos por $A.B$.

O resultado da composição de três espaços é o resultado da composição com o terceiro do espaço composto dos dois primeiros, etc.

A composição de espaços é uma operação unívoca, comutativa e associativa.

35. Designemos respectivamente por g e h os discriminantes das formas fundamentais de E_m e E_n . O elemento de volume em E_m será

$$dU = \sqrt{g} dx_1 dx_2 \dots dx_m,$$

e em E_n

$$dV = \sqrt{h} dy_1 dy_2 \dots dy_n.$$

É visível que o discriminante da forma fundamental de E_{m+n} é

$$gh;$$

o elemento de volume neste espaço será pois

$$dW = \sqrt{gh} dx_1 \dots dx_m dy_1 \dots dy_n,$$

e tem-se

$$dW = dU dV;$$

o elemento de volume no espaço composto é igual ao produto dos elementos de volume nos dois espaços componentes. Esta proposição estende-se imediatamente ao resultado da composição de qualquer número de espaços.

Consideremos uma região de medida U em E_m e uma região de medida V em E_n . Temos

$$U = \int_U dU, \quad V = \int_V dV.$$

Para uma região de medida W de E_{m+n} temos também

$$W = \int_W dW.$$

Se a segunda região é a região composta das duas primeiras, será

$$W = UV$$

Tem-se com efeito, notando que g é independente de y_1, \dots, y_n ,

$$W = \int \dots \int \sqrt{g} dx_1 \dots dx_m \int \dots \int \sqrt{h} dy_1 \dots dy_n,$$

onde o primeiro integral é estendido à região ρ_1 constituída pelos pontos duma variedade de m dimensões

$$\begin{cases} x_1, \dots, x_m & \text{variáveis,} \\ y_1, \dots, y_n & \text{constantes,} \end{cases}$$

correspondente a todos os sistemas de valores que x_1, \dots, x_m tomam em W , e o segundo integral é estendido à região variável ρ_2 de n dimensões constituída pelos pontos de W correspondentes a um sistema variável de valores de x_1, \dots, x_m . Esta proposição resulta imediatamente, se U e V são convexas, da redução do integral $\int_W dW$ a

integrais simples, e generaliza-se muito simplesmente para U e V quaisquer decompondo estas regiões em partes convexas e efectuando uma dupla soma. Ora, a variedade a que pertence a região ρ_1 é idêntica a E_m , e ρ_1 é evidentemente idêntica a U . Quanto às regiões ρ_2 , cada uma delas pertence a uma variedade idêntica a E_n e é manifestamente idêntica a V . Podemos pois escrever

$$W = \int_U \dots \int \sqrt{g} dx_1 \dots dx_m \int_V \dots \int \sqrt{h} dy_1 \dots dy_n,$$

ou, por V e h serem independentes de x_1, \dots, x_m ,

$$W = \int_U dU \cdot \int_V dV = UV,$$

como tínhamos afirmado.

Mais geralmente, se $\varphi(A)$ é uma função integravel-(R) em U e $\phi(B)$ uma função integravel-(R) em V , tem-se

$$(1) \quad \int_W \varphi(A) \phi(B) dW = \int_U \varphi(A) dU \cdot \int_V \phi(B) dV,$$

pelas mesmas razões.

Para uma função $f(A.B)$ integrável-(R) em W tem-se

$$(2) \quad \int_W f(A.B) dW = \int_U dU \int_V f(A.B) dV.$$

36. Consideremos agora um conjunto A mensurável de E_m e um conjunto B mensurável de E_n . O conjunto $A.B$ de E_{m+n} , resultante da composição de A e B , é também mensurável e tem por medida o produto da medida de A pela medida B .

Com efeito, suponhamos em E_m um conjunto aberto R de medida finita que contenha A , e em E_n um conjunto aberto S de medida finita que contenha B . Cada um dos conjuntos R e S é a soma duma infinidade numerável de regiões fechadas com pontos comuns só nas fronteiras. Compondo uma região α de R com uma região β de S , obtém-se em E_{m+n} uma região fechada; fazendo o mesmo para tôdas as regiões de R e S , obtém-se em E_{m+n} um conjunto aberto T , soma duma infinidade numerável de regiões fechadas com pontos comuns só nas fronteiras, o qual contém $A.B$. A medida do conjunto T é o produto das medidas dos conjuntos R e S , porque, em virtude do teorema precedente, é

$$m T = m \sum \alpha . \beta = \sum m (\alpha . \beta) = \sum m \alpha . m \beta = \sum m \alpha . \sum m \beta.$$

Resulta daqui que a medida exterior de $A.B$ não é superior ao produto das medidas de A e B :

$$(1) \quad m_e(A.B) \leq m A . m B.$$

Efectivamente, sendo a medida de A o limite inferior das medidas dos conjuntos R , para ϵ positivo arbitrário

pode tomar-se R tal que

$$mR < mA + \varepsilon;$$

análogamente, pode tomar-se S tal que

$$mS < mB + \varepsilon;$$

estas desigualdades dão

$$(2) \quad mT < mA \cdot mB + \varepsilon(mA + mB) + \varepsilon^2$$

Por outro lado, tem-se

$$(3) \quad m_e(A \cdot B) \leq mT.$$

Ora, se fôsse

$$m_e(A \cdot B) > mA \cdot mB,$$

ter-se-ia, designando por h a diferença positiva dêstes dois números,

$$m_e(A \cdot B) = mA \cdot mB + h,$$

donde, por (3),

$$mA \cdot mB + h \leq mT,$$

desigualdade incompatível com (2) para ε suficientemente pequeno. A fórmula (1) fica pois demonstrada.

Designemos agora por A_1 e B_1 os complementares de A relativamente a R e de B relativamente a S . Será evidentemente

$$T = R \cdot S = A \cdot B + A \cdot B_1 + A_1 \cdot B + A_1 \cdot B_1,$$

donde resulta

$$m_e(T - A \cdot B) \leq m_e(A \cdot B_1) + m_e(A_1 \cdot B) + m_e(A_1 \cdot B_1),$$

ou, por (1),

$$\begin{aligned} m_e(T - A \cdot B) &\leq m A \cdot m B_1 + m A_1 \cdot m B + m A_1 \cdot m B_1 = \\ &= (m A + m A_1)(m B + m B_1) - m A \cdot m B = \\ &= m R \cdot m S - m A \cdot m B = m T - m A \cdot m B; \end{aligned}$$

como a medida interior de $A \cdot B$ é

$$m T - m_e(T - A \cdot B),$$

tem-se portanto

$$m_i(A \cdot B) \geq m A \cdot m B,$$

desigualdade que, com (1), mostra que é necessariamente

$$m_i(A \cdot B) = m_e(A \cdot B) = m A \cdot m B,$$

em virtude da relação conhecida

$$m_i(A \cdot B) \leq m_e(A \cdot B).$$

O conjunto $A \cdot B$ é pois mensurável e tem-se

$$m(A \cdot B) = m A \cdot m B,$$

como desejávamos provar.

37. Esta última fórmula pode escrever-se

$$\int_{A \cdot B} d(A \cdot B) = \int_A dA \cdot \int_B dB,$$

e vê-se que ela generaliza para conjuntos mensuráveis a

fórmula

$$\int_{U.V} dU \cdot dV = \int_U dU \cdot \int_V dV$$

estabelecida precedentemente para regiões.

Demonstraremos agora a fórmula

$$(1) \quad \int_{A.B} \varphi(A) \phi(B) d(A.B) = \int_A \varphi(A) dA \cdot \int_B \phi(B) dB,$$

extensão aos conjuntos mensuráveis da fórmula (1) do n.º 35.

Supondo primeiro os conjuntos A e B limitados, podemos encerrar numa região limitada M de E_m o conjunto A e numa região limitada N de E_n o conjunto B . O conjunto $A.B$ ficará encerrado na região $M.N$. Façamos

$$\Phi(M) = \begin{cases} \varphi(M) & \text{em } A, \\ 0 & \text{em } M-A, \end{cases}$$

$$\Psi(N) = \begin{cases} \phi(N) & \text{em } B, \\ 0 & \text{em } N-B; \end{cases}$$

será evidentemente

$$\Phi(M) \Psi(N) = \begin{cases} \varphi(M) \phi(N) & \text{em } A.B, \\ 0 & \text{em } M.N - A.B, \end{cases}$$

e portanto

$$(2) \quad \int_{M.N} \Phi(M) \Psi(N) d(M.N) = \int_{A.B} \varphi(A) \phi(B) d(A.B),$$

e têm-se igualdades análogas para o integral de $\Phi(M)$ em M e para o de $\Psi(N)$ em N .

Ora, sabe-se que o integral lebesguiano que figura no primeiro membro de (2) se reduz a integrais lebesguianos simples sucessivos, cujos limites se determinam do mesmo modo que na redução dos integrais múltiplos riemannianos. As razões que permitiram estabelecer a fórmula (1) do n.º 35 provam aqui, por conseguinte, que o integral que figura no primeiro membro de (2) é igual ao produto dos integrais de $\Phi(M)$ em M e de $\Psi(N)$ em N , donde resulta imediatamente a fórmula (1).

Se o conjunto A não é limitado, consideremos uma sucessão infinita de conjuntos parciais de A

$$A'_1, A'_2, \dots$$

limitados e mensuráveis, tais que cada um contenha o seu antecedente. Esta sucessão tem sempre limite; suponhamo-la tal que tenda para A . É sabido que nestas condições a sucessão das medidas dos conjuntos A'_i tende para a medida de A . A fórmula (1) aplica-se a A'_i e B . Mas é, simbolicamente,

$$\int_{A \cdot B} - \int_{A'_i \cdot B} = \int_{A \cdot B - A'_i \cdot B},$$

$$\int_A - \int_{A'_i} = \int_{A - A'_i};$$

como a medida de $A - A'_i$ tende para zero, e o integral lebesguiano é absolutamente contínuo, a segunda diferença tende para zero; o mesmo sucede à primeira, por ser

$$m(A \cdot B - A'_i \cdot B) = (m A - m A'_i) \cdot m B \rightarrow 0;$$

no limite, pois, tem-se ainda a fórmula (1).

Se B é também ilimitado, procede-se para B como para A , e encontra-se o mesmo resultado.

Note-se que a sucessão A'_1, A'_2, \dots existe sempre. Sabe-se com efeito que, se A é mensurável, existe sempre uma sucessão $\alpha_1, \alpha_2, \dots$ de conjuntos limitados mensuráveis, sem pontos comuns dois a dois, tal que

$$A = \alpha_1 + \alpha_2 + \dots;$$

fazendo

$$A'_1 = \alpha_1, \quad A'_2 = \alpha_1 + \alpha_2, \quad A'_3 = \alpha_1 + \alpha_2 + \alpha_3, \dots,$$

os A'_i serão também limitados e mensuráveis, e tem-se

$$A'_1 < A'_2 < A'_3 < \dots;$$

ora o limite desta sucessão é evidentemente A . Verifica-se que

$$m A = m\alpha_1 + m\alpha_2 + \dots = \lim_{i \rightarrow \infty} (m\alpha_1 + \dots + m\alpha_i) = \lim_{i \rightarrow \infty} m A'_i.$$

De modo análogo se demonstraria a fórmula mais geral, extensão da fórmula (2) do n.º 35,

$$\int_{A.B} f(A.B) d(A.B) = \int_A dA \int_B f(A.B) dB,$$

em cujo segundo membro pode ser invertida a ordem das integrações.

Todos estes resultados se estendem sem dificuldade a qualquer número de conjuntos mensuráveis componentes.

§ 9.º

Da multiplicação das probabilidades nos conjuntos mensuráveis

38. Seja A um conjunto de E_m e B um conjunto de E_n , ambos de medida positiva, e suponhamos que em certas circunstâncias A e B são conjuntos totais de pontos possíveis e incompatíveis. Se essas circunstâncias se realizam, produzir-se há um ponto A de A e um ponto B de B ou, o que é equivalente, o ponto $A.B$ de E_{m+n} composto de A e B . O conjunto de todos estes pontos compostos possíveis, que representaremos por AB , é evidentemente constituído por elementos incompatíveis e é total. Deve observar-se que o conjunto AB está contido no conjunto composto $A.B$ mas, em geral, não é idêntico a êle porque certos pontos de $A.B$ podem ser impossíveis.

Suporemos que para cada ponto de A o conjunto dos pontos de B compatíveis com êle é mensurável, e *vice-versa*. Esta suposição contém, quando um dos conjuntos A ou B é unidimensional, a da mensurabilidade da função na teoria geométrica do integral dada por Lebesgue na sua *Thèse*, e verifica-se em casos muito extensos; em especial, no caso importante de AB ser idêntico ao conjunto composto $A.B$, como é evidente.

Suporemos ainda que, se A' e B' forem dois conjuntos mensuráveis contidos respectivamente em A e B , o conjunto $A'B'$ dos pontos possíveis compostos dos pontos de A' e B' é sempre mensurável. Talvez esta proposição seja conseqüência da suposição precedente, mas não pudemos encontrar laço lógico entre ambas. Para o nosso objecto, basta que observemos que as duas suposições não são incompatíveis. No caso particular de Lebesgue, a que

já nos referimos, a segunda é de facto consequência da primeira, como este ilustre géometra demonstrou. No caso em que AB é idêntico a $A.B$, aquela relação lógica é evidente, porque será também $A'B'$ idêntico a $A'.B'$.

Pela segunda suposição, AB é mensurável. Rejeitaremos os casos em que a medida de AB é nula e admitiremos portanto que esta medida é positiva. É o que se dá ainda, quando, em particular, AB é idêntico a $A.B$.

Consideremos dois conjuntos compatíveis de medida positiva A' e B' contidos respectivamente em A e B . É claro que a probabilidade do conjunto parcial $A'B'$ de AB , constituído por todos os pontos possíveis compostos dos pontos de A' e B , é igual à probabilidade de A' , i. é

$$P_{A'B} = P_{A'};$$

análogamente é

$$P_{AB'} = P_{B'}.$$

Na hipótese de se ter realizado um ponto indeterminado de A' , a probabilidade dum ponto indeterminado de B' é evidentemente igual à probabilidade do conjunto $A'B'$ na hipótese da realização dum ponto indeterminado de $A'B$, i. é, pelo princípio da divisão,

$$P_{B', A'} = \frac{P_{A'B'}}{P_{A'B}},$$

donde resulta, pela primeira das duas igualdades precedentes,

$$(1) \quad P_{A'B'} = P_{A'} \cdot P_{B', A'};$$

e tem-se semelhantemente

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

Estas duas fórmulas traduzem o *princípio da multiplicação* nos conjuntos mensuráveis.

Do mesmo modo que para os conjuntos finitos ou numeráveis se estende este princípio a um número qualquer de conjuntos mensuráveis totais.

Se A' fôr de medida nula, será $A'B$ de medida nula; com efeito é $A'B < A'.B$, e $A'.B$ tem a medida

$$m A'.m B = 0$$

em virtude dum teorema precedente. O conjunto $A'B' < A'B$ é também de medida nula, e vê-se que a relação $P_{A'} = 0$ tem como consequência $P_{A'B'} = 0$.

Designemos por $f(A.B)$ a densidade de probabilidade de AB . Ter-se há

$$P_{A'B'} = \int_{A'B'} f(A.B) d(A.B).$$

Notando que é $A'B' < A'.B'$, poremos

$$\Phi(A.B) = \begin{cases} f(A.B) & \text{em } A'B', \\ 0 & \text{em } A'.B' - A'B'; \end{cases}$$

virá então

$$P_{A'B'} = \int_{A'.B'} \Phi(A.B) d(A.B),$$

ou (n.º 37 *in fine*)

$$P_{A'B'} = \int_{A'} dA \int_{B'} \Phi(A.B) dB,$$

donde resulta imediatamente

$$(2) \quad P_{A'B'} = \int_{A'} dA \int_{B'_A} f(A.B) dB,$$

designando por B'_A o conjunto dos pontos de B' compatíveis com o ponto A de A' . Tem-se uma fórmula análoga com as duas integrações na ordem inversa.

Observemos que é sempre, qualquer que seja $F(A.B)$ integrável em $A' B'$,

$$(3) \quad \int_{A' B'} F(A.B) d(A.B) = \int_{A'} dA \int_{B'_A} F(A.B) dB.$$

Notemos finalmente que, designando por $\varphi(A)$ a densidade de probabilidade de A , resulta da fórmula (1)

$$P_{A' B'} = \int_{A'} P_{B', A'} \varphi(A) dA;$$

a comparação com (2) mostra que pode escrever-se a relação

$$(4) \quad \int_{B'_A} f(A.B) dB = \varphi(A) \cdot P_{B', A'}$$

entre as densidades f e φ , e tem-se uma relação análoga entre f e a densidade $\phi(B)$ de B .

As fórmulas (3) e (4) são importantes na teoria da esperança, como veremos adiante, e generalizam-se facilmente.

39. Os dois conjuntos parciais A' e B' , de medidas positivas inferiores respectivamente às medidas de A e B , dizem-se independentes quando os pontos do conjunto composto $A' B'$ são todos possíveis e são satisfeitas as relações

$$P_{B', A'} = P_{B', \bar{A}'}, \quad P_{A', B'} = P_{A', \bar{B}'},$$

onde \bar{A} e \bar{B} representam os contraditórios de A e B , i. é, os seus complementares para A e B respectivamente.

Como nos conjuntos finitos ou numeráveis prova-se facilmente que os dois membros da primeira relação são iguais a $P_{B'}$ e os da segunda a $P_{A'}$. Tem-se pois, neste caso,

$$P_{A' B'} = P_{A'} \cdot P_{B'},$$

relação que exprime o princípio da multiplicação para dois conjuntos mensuráveis parciais independentes.

40. Diremos que os dois conjuntos totais A e B são independentes, quando os pontos do conjunto composto $A.B$ forem todos possíveis e além disso forem independentes todos os pares de conjuntos parciais A' e B' de medidas positivas inferiores respectivamente às medidas de A e B .

Já observámos que o conjunto $A.B$ satisfaz às condições de mensurabilidade formuladas no começo d'este §. Se representarmos por $f(A.B)$ a densidade de probabilidade de $A.B$ no ponto $A.B$ e por $\varphi(A)$ e $\psi(B)$ as densidades de probabilidade de A e B , a última igualdade do n.º precedente pode escrever-se

$$\int_{A.B'} f(A.B) d(A.B) = \int_{A'} \varphi(A) dA \cdot \int_{B'} \psi(B) dB.$$

Mas, pelo n.º 37, o segundo membro desta relação é igual a

$$\int_{A.B'} \varphi(A) \psi(B) d(A.B);$$

como A' e B' são arbitrários pode pôr-se, pois, em todos

os pontos de $A.B$,

$$f(A.B) = \varphi(A) \cdot \psi(B).$$

Expressaremos este facto dizendo que a densidade de probabilidade do conjunto composto de dois conjuntos independentes é igual ao produto das densidades de probabilidade dos conjuntos componentes.

A extensão da noção de independência e dos resultados precedentes ao caso dum número qualquer de conjuntos totais é manifesta.

§ 10.º

Das densidades de probabilidade descontínuas

41. O processo da passagem ao limite, que já empregámos na primeira teoria da probabilidade dos conjuntos numeráveis, permite introduzir nos conjuntos mensuráveis as densidades de probabilidade descontínuas. Para não nos repetirmos fastidiosamente, limitar-nos hemos a algumas breves indicações.

Seja A um conjunto de E_n , de medida positiva, que, em certas circunstâncias, possui a densidade de probabilidade contínua $f_1(A)$, noutras circunstâncias a densidade de probabilidade contínua $f_2(A)$, e assim indefinidamente. Se existir um grupo de circunstâncias que se possa considerar limite da sucessão dos grupos de circunstâncias dados, e se a sucessão

$$f_1(A), f_2(A), \dots$$

convergir para uma função positiva $f(A)$ integrável em A ,

diremos que $f(A)$ é a densidade de probabilidade de A no grupo de circunstâncias limite.

A nova densidade de probabilidade, se não fôr contínua, será uma função de Baire da classe I. Como A é mensurável, esta função é mensurável também, como é sabido. Para que ela seja integrável, basta supor que as funções integráveis f_1, f_2, \dots satisfazem à condição

$$f_i \leq \Psi, \quad i = 1, 2, \dots,$$

onde Ψ representa uma função integrável. Se A' fôr então um conjunto parcial de A , mensurável, será, por um conhecido teorema de Lebesgue,

$$\int_{A'} f(A) dA = \lim_{i \rightarrow \infty} \int_{A'} f_i(A) dA,$$

ou

$$(1) \quad P_{A'} = \lim_{i \rightarrow \infty} P_{A'}^{(i)};$$

a nova probabilidade do conjunto A' é pois o limite da sucessão das probabilidades do mesmo conjunto para a infinidade numerável das densidades contínuas dadas. Se A' é de medida nula, é $P_{A'} = 0$; se A' tem a medida de A , é $P_{A'} = 1$; se A' tem medida positiva menor que a de A , é $0 < P_{A'} < 1$.

A relação fundamental (1) permite demonstrar facilmente para a densidade descontínua $f(A)$ todos os resultados precedentemente obtidos para densidades contínuas. Não insistiremos nisso.

42. Em lugar duma sucessão de densidades contínuas, podemos supor agora uma sucessão de densidades descontínuas da classe I de Baire. O limite desta sucessão, se

existir e não pertencer às classes 0 ou I, será, quando integrável, uma densidade de probabilidade da classe II, acerca da qual se podem repetir as considerações do n.º precedente.

A aplicação reiterada dêste processo permite definir densidades de probabilidade de qualquer classe de Baire de ordem finita. Consideremos em seguida uma sucessão de densidades de probabilidade pertencentes a classes de ordem finita, formando as ordens destas classes uma sucessão ilimitada superiormente; se esta sucessão tem um limite positivo integrável que não é de classe de ordem finita, êsse limite será uma densidade de probabilidade pertencente à classe da primeira ordem transfinita ω . Continuando assim transfinitamente, introduzem-se densidades de probabilidade de tôdas as classes de Baire que, como se sabe, são numeráveis com os ordinais cantorianos inferiores a Ω .

Pode abandonar-se a restrição da positividade. Introduzem-se então densidades não negativas de tôdas as classes, incluindo a classe 0.

As funções de Baire, denominadas por Lebesgue «funções representáveis analiticamente» por provirem da aplicação reiterada da passagem ao limite a sucessões de polinómios, segundo o célebre teorema de Weierstrass sobre as funções contínuas, são tôdas mensuráveis num conjunto mensurável. Para que possam ser densidades de probabilidade, além de não negativas têm de ser integráveis. Uma função de Baire será necessariamente integrável quando a sucessão de funções de Baire integráveis de classes inferiores que a define fôr limitada superiormente por uma função integrável qualquer. Já mostrámos no n.º precedente como se utiliza esta proposição.

§ 11.º

Da probabilidade nos conjuntos de medida infinita

43. Ainda por passagem ao limite se pode fazer o estudo da probabilidade nos conjuntos de medida infinita. Êste caso não tem a importância do que temos considerado nos §§ precedentes e por isso nos ocuparemos dele muito brevemente.

Recordaremos o que se entende por conjunto de medida infinita. Seja primeiro A um conjunto de E_n , de medida finita e O um conjunto aberto qualquer de E_n , de medida finita também; a parte comum de A e O , que se representa por OA (nos §§ precedentes esta notação tem outro significado), é mensurável de medida finita, i. é

$$m_i(OA) = m_e(OA);$$

e como

$$m_i(OA) = mO - m_e(O - OA),$$

tem-se

$$(1) \quad mO = m_e(O - OA) + m_e(OA).$$

Se o conjunto aberto O tem medida infinita (i. é, medida exterior infinita), a relação (1) toma a forma $\infty = \infty$, e é ainda verificada.

Reciprocamente, se a relação (1) se verifica para um conjunto A , de medida exterior finita, qualquer que seja o conjunto aberto O , A é mensurável. Com efeito, tomemos um O de medida finita que contenha A ; será então $OA = A$ e a relação (1) dará

$$mO = m_e(O - A) + m_e A,$$

ou

$$m_i A = m_e A,$$

o que prova a afirmação.

Pôsto isto, se um conjunto A de medida exterior infinita satisfaz à relação (1) qualquer que seja o conjunto aberto O , diz-se que A é mensurável de medida infinitamente grande. Muitas propriedades dos conjuntos de medida finita estendem-se a conjuntos mensuráveis quaisquer, de medida finita ou infinita. Recordaremos a seguinte, de que vamos fazer uso: se A é de medida infinita, pode pôr-se sempre

$$A = \alpha_1 + \alpha_2 + \dots,$$

onde os α_i , em infinidade numerável, são limitados, de medida positiva e estranhos. Fazendo

$$\alpha_1 = A'_1, \alpha_1 + \alpha_2 = A'_2, \dots,$$

obtém-se uma sucessão

$$A'_1, A'_2, \dots$$

de conjuntos de medidas positivas crescentes contidos em A , que tende para A ; e é

$$\lim_{i \rightarrow \infty} m A'_i = \infty.$$

Suponhamos agora uma sucessão infinita de funções de Baire

$$f_1(A), f_2(A), \dots$$

definidas em A , não negativas, tais que

$$\int_{A'_i} f_i(A) dA = 1, \quad i = 1, 2, \dots,$$

e que

$$\lim_{i \rightarrow \infty} f_i(A) = f(A).$$

Se em certas circunstâncias A'_1 é total com a densidade de probabilidade f_1 , se noutras circunstâncias A'_2 é total com a densidade de probabilidade f_2 , e assim indefinidamente, diremos que A é total para o grupo de circunstâncias limite, se êste existir, com a densidade de probabilidade f .

Se as funções f_i verificam a condição de ser

$$f_i \leq \Psi,$$

onde Ψ é uma função integrável em qualquer conjunto limitado mensurável A' contido em A , será f integrável nêsse conjunto e ter-se à

$$(2) \quad \int_{A'} f(A) dA = \lim_{i \rightarrow \infty} \int_{A'} f_i(A) dA.$$

Com efeito, contendo cada conjunto A'_i todos os precedentes e tendendo para A a sucessão (A'_i) , a partir de certo valor de i estará A' contido em A'_i e em todos os conjuntos de índice superior; as funções

$$f_i, f_{i+1}, \dots$$

serão integráveis em A' , e a relação (2) fica demonstrada pelo teorema de Lebesgue citado anteriormente (n.º 41).

Essa relação pode escrever-se

$$(2^*) \quad P_{A'} = \lim_{i \rightarrow \infty} P_{A'_i}^{(i)},$$

onde o primeiro membro representa a probabilidade de A' para a densidade f e no segundo membro figura a probabilidade do mesmo conjunto para a densidade f_i .

Se o conjunto de medida finita A' é ilimitado, pode considerar-se como limite duma sucessão de conjuntos limitados mensuráveis β_i , cada um dos quais contém o precedente e cujas medidas tendem para a de A' . Como os integrais de f nos β_i formam uma sucessão não decrescente e não excedem 1, tendem para um limite; é sabido que nestas condições f é integrável em A' e tem-se

$$\int_{A'} f(A) dA = \lim_{i \rightarrow \infty} \int_{\beta_i} f(A) dA,$$

ou

$$(3) \quad P_{A'} = \lim_{i \rightarrow \infty} P_{\beta_i}.$$

Finalmente, se A' é de medida infinita, formemos a sucessão de termo geral $A'_i A'$ (parte comum de A'_i e A'); poremos

$$(4) \quad P_{A'} = \lim_{i \rightarrow \infty} P_{A'_i A'}^{(i)},$$

quando êste limite existir. Em particular, é $A'_i A = A'_i$, $P_{A'_i}^{(i)} = 1$, e tem-se portanto

$$P_A = 1.$$

Fácilmente se veria que a definição (4) concorda com a fórmula (3). A concordância com (2*) é evidente.

As fórmulas (2*), (3) e (4) permitem construir tóda a teoria da probabilidade nos conjuntos de medida infinita.

44. Se os conjuntos A'_i , como totais, são homogéneos, é

$$f_i(A) = \frac{1}{m A'_i},$$

portanto

$$f(A) = 0.$$

O conjunto total A é pois homogéneo, com densidade de probabilidade nula. A probabilidade dum conjunto parcial de medida finita é sempre nula. Se supomos feita a decomposição

$$A = \alpha_1 + \alpha_2 + \dots,$$

onde os α_i são estranhos e de medidas positivas, e consideramos a realização dum ponto indeterminado de α_i como um acontecimento ($i = 1, 2, \dots$), temos um exemplo dum conjunto numerável total de acontecimentos possíveis e incompatíveis, todos de probabilidades nulas, i. é, dum conjunto numerável total homogéneo. No n.º 23 estudámos a probabilidade de tais conjuntos. A observação que acabamos de fazer permitiria estudá-la dum novo ponto de vista. Aqui, porém, temos que limitar-nos a esta indicação.

§ 12.º

Da probabilidade nos conjuntos de medida nula

45. Para não omitir caso algum, resta-nos dizer duas palavras sôbre a probabilidade nos conjuntos totais de medida nula. É claro que em tais conjuntos não pode

haver densidade de probabilidade, porque um integral estendido a um conjunto de medida nula é sempre nulo.

Seja em E_n um conjunto A de medida nula. Vamos efectuar uma redução importante.

Duas suposições são possíveis, antes de mais: A está ou não está contido num E_{n-1} . Ora na primeira há três outras a considerar: A é em E_{n-1} mensurável de medida maior que zero, caso estudado nos §§ precedentes; ou não mensurável; ou de medida nula, caso que não difere do inicial e se submete à mesma redução.

Só nos restam, pois, duas suposições: A não está contido num E_{n-1} , ou é um conjunto não mensurável dum E_{n-1} . Na primeira, A pode ser finito, numerável, ou ter a potência do contínuo. Não há outra alternativa se admitimos com Cantor a «hipótese do contínuo», i. é, que a potência do contínuo é a segunda potência. Na segunda suposição, A tem sempre a potência do contínuo, de contrário teria medida nula em E_{n-1} . Como o caso de A ser finito ou numerável já está estudado, ficamos reduzidos ao caso em que A tem a potência do contínuo.

Consideremos portanto em E_n um conjunto A de medida nula, da potência do contínuo, que não está contido num E_{n-1} ou é um conjunto não mensurável dum E_{n-1} . Seja B , de medida positiva, um conjunto total de pontos possíveis e incompatíveis de outro espaço F_n , e suponhamos que entre os pontos de B e os de A se estabeleceu uma correspondência biunívoca qualquer, o que é possível por terem os dois conjuntos a mesma potência. Se consideramos realizado um ponto de A quando se realiza o ponto correspondente de B , A é um conjunto total de pontos possíveis e incompatíveis. A probabilidade dum conjunto parcial A' correspondente a um conjunto mensurável B' contido em B é a probabilidade de B' . Assim, cada proposição da teoria da probabilidade nos conjuntos

totais de medida positiva fornece imediatamente uma posição análoga para a probabilidade nos conjuntos totais de medida nula, e esta observação dispensa-nos de fazer mais considerações sôbre o assunto.

§ 13.º

Da esperança das variáveis aleatórias

46. Denominaremos *variável aleatória* um símbolo que representa um elemento indeterminado dum conjunto total de números possíveis e incompatíveis. Êstes números são os valores da variável aleatória.

Os valores duma variável aleatória podem ser todos diferentes e neste caso distinguem-se por si mesmos; ou não, e então distinguem-se dois valores iguais pelos acontecimentos concretos a que andam associados.

Por exemplo, os «pontos» possíveis no lançamento de dois dados, 2, 3, ..., 12, são os valores duma variável aleatória. Mas também o conjunto 2, 3, 3, 4, ..., 12, que difere do precedente em ter o elemento 3 repetido, pode ser considerado como conjunto dos valores duma variável aleatória, se um dos elementos 3 é o «ponto» obtido quando o primeiro dado fornece um 1 e o segundo um 2, e o outro elemento 3 é o «ponto» obtido quando se dá o inverso, conservando os restantes elementos o significado primitivo.

Duas ou mais variáveis aleatórias dizem-se independentes, se são independentes os conjuntos dos valores de cada uma delas.

Além da noção fundamental de probabilidade, anda ligada às variáveis aleatórias uma noção derivada extremamente importante, de que vamos ocupar-nos.

47. Seja x uma variável aleatória, de valores

$$x_1, x_2, \dots, x_n$$

em número finito, com as probabilidades

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

A soma

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n$$

é por definição a *esperança* matemática da variável aleatória x .

Para representar a esperança de x empregaremos a notação

$$\text{esp } x.$$

Esta noção foi primeiro formulada a propósito de problemas de jôgo, e aí reside a razão do nome. Suponhamos duas pessoas que jogam nas seguintes condições: existe um prémio a_1 , que será recebido pela primeira se se realizar um acontecimento A e pela segunda se se realizar o acontecimento contrário A' . Se as duas pessoas resolvem não jogar, é justo que dividam entre si o prémio proporcionalmente às probabilidades p e q dos dois acontecimentos ⁽¹⁾. As partes que então competem a cada uma são as *esperanças* que tinham as duas pessoas, antes daquela resolução, relativamente ao prémio a_1 e valem respectivamente

$$p a_1 \quad \text{e} \quad q a_1,$$

ou

$$p a_1 + q \cdot 0 \quad \text{e} \quad p \cdot 0 + q a_1.$$

⁽¹⁾ Supomos que as duas pessoas não contribuíram para a formação do prémio.

Esta noção de esperança concorda com a nossa definição; o ganho da primeira pessoa é com efeito uma variável aleatória com os valores a_1 e 0 de probabilidades p e q , e coisa análoga se dá com o ganho da segunda.

Supondo duas outras pessoas que jogam outro prémio a_2 nas mesmas condições, serão

$$p a_2 \quad \text{e} \quad q a_2$$

as esperanças respectivas. Imaginemos que a primeira pessoa do primeiro jôgo e a segunda do segundo jôgo cedem os seus direitos a uma quinta pessoa. Se nenhum dos dois jogos se efectua, esta pessoa receberá a soma das partes que competiam a ambas, i. é,

$$p a_1 + q a_2;$$

êste número é pois a esperança da quinta pessoa relativamente à variável aleatória de valores a_1 e a_2 com probabilidades p e q .

Importante nas questões de jôgo, a noção de esperança não importa menos à teoria geral das probabilidades. Esta importância assenta nas duas propriedades seguintes.

Sejam x e y duas variáveis aleatórias dependentes ou independentes, susceptíveis dos valores

$$x_1, x_2, \dots, x_m$$

e

$$y_1, y_2, \dots, y_n$$

respectivamente. A soma $x + y$ destas variáveis é a variável aleatória susceptível dos valores da expressão

$$x_i + y_j,$$

$$i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n,$$

correspondentes aos sistemas de números (x_i, y_j) de realização possível. Designemos por r_{ij} as probabilidades dêstes sistemas; será

$$\text{esp}(x + y) = \sum_{i,j} (x_i + y_j) r_{ij} = \sum_{i=1}^m (x_i \sum_j r_{ij}) + \sum_{j=1}^n (y_j \sum_i r_{ij}),$$

onde as somas interiores do terceiro membro se estendem respectivamente a todos os valores de j compatíveis com o número i e a todos os valores de i compatíveis com o número j . Ora (n.º 8) a primeira destas somas é a probabilidade de x_i e a segunda a probabilidade de y_j ; logo

$$\text{esp}(x + y) = \text{esp } x + \text{esp } y;$$

i. é, a esperança da soma de duas variáveis aleatórias, dependentes ou independentes, é igual à soma das esperanças das parcelas.

Esta propriedade vale para qualquer número de variáveis aleatórias. Tem-se com efeito

$$\text{esp}(x + y + z) = \text{esp}(x + y) + \text{esp } z = \text{esp } x + \text{esp } y + \text{esp } z,$$

e assim sucessivamente.

O produto xy é a variável aleatória susceptível dos valores da expressão

$$x_i y_j$$

correspondentes aos sistemas (x_i, y_j) de realização possível. É pois

$$\text{esp}(xy) = \sum_{i,j} x_i y_j r_{ij}.$$

Se as duas variáveis aleatórias x e y são independentes,

todos os sistemas (x_i, y_j) são de realização possível e tem-se

$$r_{ij} = p_i q_j,$$

onde p_i é a probabilidade de x_i e q_j a de y_j (n.º 11); logo

$$\text{esp}(xy) = \sum_{i=1}^m \sum_{j=1}^n x_i y_j p_i q_j = \sum_{i=1}^m p_i x_i \cdot \sum_{j=1}^n q_j y_j,$$

ou

$$\text{esp}(xy) = \text{esp } x \cdot \text{esp } y,$$

i. é, a esperança do produto de duas variáveis aleatórias *independentes* é igual ao produto das esperanças dos factores.

Vale a mesma propriedade para qualquer número de variáveis aleatórias independentes.

Se os valores duma variável aleatória são todos iguais, tem-se uma constante aleatória. A esperança duma constante aleatória é a própria constante; com efeito, é

$$\text{esp } a = p_1 a + p_2 a + \dots + p_n a = a,$$

por ser

$$p_1 + p_2 + \dots + p_n = 1.$$

Somando a cada um dos valores duma variável aleatória x um número c , obtêm-se os valores duma nova variável aleatória $c + x$, e tem-se

$$\text{esp}(c + x) = p_1(c + x_1) + \dots + p_n(c + x_n) = c + \text{esp } x.$$

Multiplicando cada um dos valores de x por c , obtêm-se os valores da variável aleatória cx ; tem-se

$$\text{esp}(cx) = p_1 c x_1 + \dots + p_n c x_n = c \cdot \text{esp } x.$$

48. Seja agora x uma variável aleatória de valores

$$x_1, x_2, \dots, x_n, \dots$$

em infinidade numerável, com as probabilidades positivas

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

Se a série

$$p_1 x_1 + p_2 x_2 + \dots + p_n x_n + \dots$$

fôr absolutamente convergente, a sua soma define a esperança da variável aleatória x .

Quando a variável x é limitada, a esperança de x existe sempre. Mas esta condição suficiente não é necessária. Por exemplo, a variável susceptível dos valores

$$1, 2, 3, \dots$$

com as probabilidades

$$\frac{1}{2}, \frac{1}{2^2}, \frac{1}{2^3}, \dots$$

tem a esperança 2.

A esperança das variáveis aleatórias numeráveis possui as mesmas propriedades que a esperança das variáveis aleatórias finitas.

Sejam x e y duas variáveis aleatórias numeráveis, dependentes ou independentes, susceptíveis dos valores

$$x_1, x_2, \dots, x_n, \dots$$

e

$$y_1, y_2, \dots, y_n, \dots$$

respectivamente. Se existirem as esperanças de x e y ,

existirá também a esperança de $x + y$ e será igual à soma daquelas duas.

Com efeito, a soma das variáveis x e y é a variável aleatória susceptível dos valores da expressão

$$x_i + y_j, \\ i = 1, 2, \dots, n, \dots, \quad j = 1, 2, \dots, n, \dots$$

correspondentes aos sistemas (x_i, y_j) de realização possível. Se designamos por r_{ij} as probabilidades destes sistemas, temos

$$(1) \quad \text{esp}(x + y) = \sum_{i,j} (x_i + y_j) r_{ij}$$

se a série dupla do segundo membro fôr absolutamente convergente. Ora esta série resulta da adição ordenada das duas séries duplas

$$\sum_{i,j} x_i r_{ij} \quad \text{e} \quad \sum_{i,j} y_j r_{ij},$$

que são absolutamente convergentes por convergirem absolutamente as duas séries simples

$$(2) \quad \sum_{i=1}^{\infty} (x_i \sum_j r_{ij}) \quad \text{e} \quad \sum_{j=1}^{\infty} (y_j \sum_i r_{ij})$$

obtidas a partir delas somando-as respectivamente por linhas e por colunas; estas séries simples representam com efeito as esperanças de x e y , porque as séries

$$\sum_j r_{ij} \quad \text{e} \quad \sum_i r_{ij}$$

são respectivamente (n.º 18) as probabilidades de x_i e y_j .

A série que figura no segundo membro de (1) é portanto absolutamente convergente e igual à soma das séries (2), i. é,

$$\text{esp}(x + y) = \text{esp } x + \text{esp } y,$$

como tínhamos afirmado.

O produto das variáveis x e y é a variável aleatória xy susceptível dos valores da expressão

$$x_i y_j$$

correspondentes aos sistemas (x_i, y_j) de realização possível. É portanto

$$\text{esp}(xy) = \sum_i \sum_j x_i y_j r_{ij}$$

se a série do segundo membro fôr absolutamente convergente. Ora, se as duas variáveis aleatórias x e y forem independentes, todos os sistemas (x_i, y_j) são de realização possível e tem-se

$$r_{ij} = p_i q_j,$$

onde p_i e q_j são respectivamente as probabilidades de x_i e y_j (n.º 22). A série dupla

$$\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_i y_j p_i q_j$$

tem por termo geral o produto dos termos gerais das séries simples

$$\sum_{i=1}^{\infty} p_i x_i \quad \text{e} \quad \sum_{j=1}^{\infty} q_j y_j,$$

que convergem absolutamente para as esperanças de x e y ; é por conseguinte absolutamente convergente e tem por soma o produto destas esperanças, donde resulta imediatamente

$$\text{esp}(xy) = \text{esp } x \cdot \text{esp } y.$$

A esperança do produto de duas variáveis aleatórias numeráveis *independentes* existe quando existirem as esperanças dos factores, e é igual ao produto destas.

Os dois teoremas que acabamos de demonstrar têm lugar para um número qualquer de variáveis aleatórias.

Como no n.º precedente, demonstram-se imediatamente as igualdades:

$$\text{esp } a = a,$$

onde a é uma constante aleatória numerável;

$$\text{esp}(c + x) = c + \text{esp } x,$$

$$\text{esp}(cx) = c \cdot \text{esp } x,$$

onde x é uma variável aleatória numerável e c um número qualquer.

49. Consideremos finalmente o caso de variáveis aleatórias cujos valores formam conjuntos com a potência do contínuo.

Seja A um conjunto total, de medida positiva, de pontos possíveis e incompatíveis dum espaço E_n , e suponhamos que a cada um dos seus pontos A corresponde um número $x = \xi(A)$; x será uma variável aleatória, e a esperança de x é por definição o integral

$$\int_A \xi(A) \varphi(A) dA,$$

onde $\varphi(A)$ representa a densidade de probabilidade de A . Se $\xi(A)$ é limitada e mensurável, a esperança de x existe sempre. Esta condição suficiente não é porém necessária.

Para tais variáveis aleatórias subsistem ainda as proposições precedentes. Consideremos duas variáveis aleatórias

$$x = \xi(A), \quad y = \eta(B),$$

dependentes ou independentes, a primeira definida num conjunto A dum E_m e a segunda num conjunto B dum E_n , sendo A e B de medida positiva, com densidades de probabilidade $\varphi(A)$ e $\psi(B)$. Se as esperanças de x e y existirem, existirá também a de $x+y$ e será igual à soma daquelas.

Ter-se há com efeito

$$(1) \quad \text{esp}(x+y) = \int_{AB} (\xi(A) + \eta(B)) f(A \cdot B) d(A \cdot B),$$

onde $f(A \cdot B)$ é a densidade de probabilidade no conjunto AB de E_{m+n} constituído pelos pontos possíveis compostos dos pontos de A e de B , se o integral do segundo membro existir. Êste integral existe necessariamente quando existirem os dois integrais

$$(2) \quad \int_{AB} \xi(A) f(A \cdot B) d(A \cdot B), \quad \int_{AB} \eta(B) f(A \cdot B) d(A \cdot B)$$

e, pela fórmula (3) do n.º 38, êstes integrais existirão se existirem os integrais

$$(3) \quad \int_A \xi(A) dA \int_{B_A} f(A \cdot B) dB, \quad \int_B \eta(B) dB \int_{A_B} f(A \cdot B) dA,$$

onde B_A significa o conjunto dos pontos de B compatíveis com o ponto A de A , e A_B o conjunto dos pontos de A compatíveis com o ponto B de B . Ora, pela fórmula (4) do mesmo n.º, os integrais estendidos a B_A e a A_B são respectivamente iguais a $\varphi(A)$ e $\psi(B)$, porque é evidentemente

$$P_{B,A} = P_{A,B} = 1.$$

Logo os integrais (3) não diferem dos integrais

$$(4) \quad \int_A \xi(A) \varphi(A) dA, \quad \int_B \eta(B) \psi(B) dB,$$

que são as esperanças de x e y e existem por hipótese, e portanto existe a esperança de $x + y$. Pois que os integrais (3) são respectivamente iguais aos integrais (2) e a soma destes é o integral (1), tem-se

$$\text{esp}(x + y) = \text{esp } x + \text{esp } y,$$

como nos casos anteriores.

A esperança do produto xy é análogamente

$$(5) \quad \text{esp}(xy) = \int_{A \cdot B} \xi(A) \eta(B) f(A \cdot B) d(A \cdot B),$$

quando este integral existir. Se as variáveis aleatórias x e y são *independentes*, é

$$A \cdot B = A \cdot B$$

e (n.º 40)

$$f(A \cdot B) = \varphi(A) \psi(B);$$

pela fórmula (1) do n.º 37, o integral (5) existe porque existem os integrais (4) e é igual ao produto destes, i. é,

$$\text{esp}(xy) = \text{esp } x \cdot \text{esp } y,$$

como nos casos precedentes.

Estas proposições são válidas para qualquer número de variáveis aleatórias.

Se a é uma constante aleatória no conjunto A , tem-se

$$\text{esp } a = \int_A a \varphi(A) dA = a;$$

tem-se também, sendo c um número qualquer e x uma variável aleatória em A ,

$$\text{esp}(c+x) = \int_A (c + \xi(A)) \varphi(A) dA = c + \text{esp } x$$

e

$$\text{esp}(cx) = \int_A c \xi(A) \varphi(A) dA = c \cdot \text{esp } x,$$

como precedentemente.

50. Se as variáveis aleatórias são definidas em conjuntos de medida infinita, encontram-se os mesmos resultados mediante o processo da passagem ao limite, tal como o empregámos no § 11.

Se elas forem definidas em conjuntos de medida nula, com a potência do contínuo, chega-se ainda aos mesmos resultados, o que se torna evidente atendendo a que no § 12 reduzimos o estudo da probabilidade nos conjuntos

de medida nula ao estudo da probabilidade em conjuntos de medida positiva.

§ 14.º

Dos vectores aleatórios

51. Suponhamos que a cada elemento A dum conjunto total A de acontecimentos possíveis e incompatíveis corresponde, não um só número, mas um sistema de k números cuja ordem não é indiferente

$$x_A, y_A, \dots, w_A.$$

Se representamos por x um número indeterminado dentre os x_A e damos a $y \dots w$ significação análoga, serão

$$x, y, \dots, w$$

k -variáveis aleatórias definidas no conjunto A . Diz-se que estas variáveis são as componentes dum *vector aleatório* \mathbf{v} definido em A ; escreveremos

$$\mathbf{v} = (x, y, \dots, w).$$

Se \mathbf{v}' é outro vector aleatório de k componentes, definido num conjunto total B da mesma espécie que A ,

$$\mathbf{v}' = (x', y', \dots, w'),$$

ao vector

$$\mathbf{v} + \mathbf{v}' = (x + x', y + y', \dots, w + w')$$

chama-se *soma* de \mathbf{v} e \mathbf{v}' .

Analogamente, ao vector

$$\mathbf{v} \mathbf{v}' = (x x', y y', \dots, w w')$$

chamaremos *produto* ⁽¹⁾ de \mathbf{v} e \mathbf{v}' .

Se a, b, \dots, l designarem as esperanças das variáveis x, y, \dots, w , o vector

$$\mathbf{a} = (a, b, \dots, l)$$

é por definição a *esperança* de \mathbf{v} .

Os dois vectores \mathbf{v} e \mathbf{v}' dizem-se *independentes*, se são independentes os dois conjuntos A e B . Neste caso todos os grupos

$$x \text{ e } x', x \text{ e } y', \dots, x \text{ e } w', \dots, w \text{ e } w'$$

são pares de variáveis aleatórias independentes.

Pôsto isto, os resultados do § anterior permitem enunciar imediatamente as seguintes proposições:

a) a esperança da soma de dois vectores aleatórios é igual à soma das esperanças das parcelas;

b) a esperança do produto de dois vectores aleatórios *independentes* é igual ao produto das esperanças dos factores.

Além disso, a esperança dum vector aleatório constante é o próprio vector; a esperança da soma ou do produto dum vector constante e dum vector variável é igual à soma ou ao produto do primeiro vector e da esperança do segundo.

(1) Esta noção nada tem que ver com os produtos de vectores definidos em Geometria.

§ 15.º

Dois teoremas-limites
sobre variáveis aleatórias independentes

52. Em 1901, o geómetra russo Liapunoff⁽¹⁾ demonstrou a bela proposição seguinte.

Seja $x_1, x_2, \dots, x_n, \dots$ uma sucessão infinita de variáveis aleatórias independentes. Se, para todos os valores do índice i , existem as quantidades

$$a_i = \text{esp } x_i, \quad b_i = \text{esp } (x_i - a_i)^2, \quad b_i^{(2+\delta)} = \text{esp } |x_i - a_i|^{2+\delta},$$

sendo δ um número positivo, e se o quociente

$$\frac{b_1^{(2+\delta)} + b_2^{(2+\delta)} + \dots + b_n^{(2+\delta)}}{(\sqrt{b_1 + b_2 + \dots + b_n})^{2+\delta}}$$

converge para zero com $\frac{1}{n}$, a probabilidade das desigualdades

$$t_1 < \frac{x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)}{\sqrt{2(b_1 + b_2 + \dots + b_n)}} < t_2,$$

onde t_1 e t_2 são dois números quaisquer, tende para o limite

$$\frac{1}{\sqrt{\pi}} \int_{t_1}^{t_2} e^{-t^2} dt$$

quando n cresce infinitamente.

(1) *Comptes Rendus de l'Ac. des Sc. de Paris*, t. CXXXII, págs. 126 e 814. Demonstração elegante em Markoff, *Démonstration du second théorème-limite du Calcul des Probabilités*, S. Petersburgo, 1913.

Esta proposição tinha sido demonstrada anteriormente, sob condições muito mais restritivas, por Tchebychef⁽¹⁾, e também por Markoff⁽²⁾, que introduziu na demonstração de Tchebychef aperfeiçoamentos importantes. Foi descoberta por Laplace⁽³⁾ e generalizada por Poisson⁽⁴⁾, mas nenhum destes dois géometras a demonstrou com rigor. Há poucos anos, Lindeberg⁽⁵⁾ conseguiu tornar menos restritivas as condições de Liapunoff, mas com sacrifício da simplicidade do enunciado.

Vamos deduzir da proposição citada os seguintes teoremas.

I. Sejam $x_1, x_2, \dots, x_n, \dots$ as variáveis aleatórias de Liapunoff. Se a sucessão

$$b_1, b_2, \dots, b_n, \dots$$

é limitada superiormente e se ε e η são dois números positivos arbitrários, para valores de n suficientemente grandes a probabilidade da desigualdade

$$(1) \quad \frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} < \varepsilon,$$

onde k é um número positivo menor que 2, é superior a $1 - \eta$.

II. Supostas as mesmas variáveis aleatórias, se a sucessão

$$b_1, b_2, \dots, b_n, \dots$$

(1) *Sur deux théorèmes relatifs aux probabilités*, in-*Acta Mathematica*, t. 14.

(2) *Über die Wurzeln der Gleichung $ex^3 \cdot d^m e^{-x^2} : dx^m = 0$* , 1898, in-*Wahrscheinlichkeitsrechnung, Anhang I*, trad. alemã, Leipzig, 1912.

(3) *Op. cit.*, Livre II, Chap. IV, § 22; Chap. IX, § 39.

(4) *Op. cit.*, Chap. IV, §§ 97 a 101.

(5) *Comptes Rendus*, t. CLXXIV, 1922, pág. 1400.

é limitada inferiormente por um número positivo e se ε e η são dois números positivos arbitrários, para valores suficientemente grandes de n a probabilidade da desigualdade (1), onde k é um número maior que 2, é inferior a η .

Designemos ⁽¹⁾ por t um número positivo qualquer; pelo teorema de Liapunoff, a probabilidade da dupla desigualdade

$$-t\sqrt{2B_n} < x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n) < t\sqrt{2B_n},$$

onde é

$$B_n = b_1 + b_2 + \dots + b_n,$$

ou, o que é o mesmo para $k > 0$, a probabilidade da desigualdade

$$(2) \quad \frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} < \frac{(t\sqrt{2B_n})^k}{n},$$

tende para

$$\Theta(t) = \frac{1}{\sqrt{\pi}} \int_{-t}^{+t} e^{-t^2} dt$$

quando n diverge para $+\infty$. Lembremos que a função $\Theta(t)$ se anula para $t=0$, é contínua para todos os valores de t , cresce com t e é

$$\lim_{t \rightarrow \infty} \Theta(t) = 1;$$

e que, por conseguinte, existe um valor suficientemente grande de t para o qual a diferença $1 - \Theta(t)$ é igual a um número positivo arbitrariamente pequeno, e um valor sufi-

⁽¹⁾ A demonstração que segue inspira-se duma das demonstrações do teorema de Bernoulli dadas por Tehebychef; cf. Markoff, *Wahrscheinl.*, Kap. II, § 10.

cientemente pequeno de t para o qual $\theta(t)$ é igual a um número positivo tão pequeno quanto se queira. Notemos também que, se existir um número L tal que se tenha

$$b_i < L$$

para $i = 1, 2, 3, \dots$, será satisfeita a desigualdade

$$(3) \quad B_n < nL$$

qualquer que seja n ; e que, análogamente, existindo um número $l > 0$ tal que seja

$$l < b_i$$

para $i = 1, 2, 3, \dots$, ter-se há para todos os valores de n

$$(3') \quad nl < B_n.$$

Demonstremos o teorema I, em que é $0 < k < 2$. Façamos

$$\eta = \eta' + \eta'', \quad \eta' > 0, \quad \eta'' > 0$$

e disponhamos do número t de modo que seja

$$1 - \theta(t) = \eta'.$$

Representemos por Q a probabilidade da desigualdade (2); existe um número n_0 tal que é

$$\theta(t) - Q < \eta''$$

para $n > n_0$; portanto tem-se

$$1 - Q < \eta' + \eta'',$$

ou

$$Q > 1 - \eta$$

para $n > n_0$. Mas, por virtude de (3), é

$$\frac{(t\sqrt{2B_n})^k}{n} < (t\sqrt{2L})^k \cdot n^{\frac{k-2}{2}},$$

logo, a probabilidade da desigualdade

$$(4) \frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} < (t\sqrt{2L})^k \cdot n^{\frac{k-2}{2}}$$

é maior que a de (2). Se ainda sujeitamos n à condição

$$(t\sqrt{2L})^k \cdot n^{\frac{k-2}{2}} < \varepsilon,$$

ou

$$n > \left(\frac{(t\sqrt{2L})^k}{\varepsilon} \right)^{\frac{2}{2-k}} = n_1,$$

a probabilidade de (1) será maior que a de (4). Conclui-se que a probabilidade da desigualdade (1) é maior que $1 - \eta$ para todos os valores de n superiores ao maior dos números n_0 e n_1 , como pretendíamos provar.

Passemos ao teorema II, onde é $k > 2$. Façamos ainda

$$\eta = \eta' + \eta'', \quad \eta' > 0, \quad \eta'' > 0$$

e disponhamos agora do número t de modo que seja

$$\theta(t) = \eta'.$$

Designemos por Q' a probabilidade da desigualdade (2);

existe um número n'_0 tal que é

$$\theta(t) - Q' > -\eta''$$

para $n > n'_0$; portanto tem-se

$$\eta' - Q' > -\eta'',$$

ou

$$Q' < \eta$$

para $n > n'_0$. Mas é, por (3'),

$$(t\sqrt{2l})^k \cdot n^{\frac{k-2}{2}} < \frac{(t\sqrt{2}B_n)^k}{n},$$

logo, a probabilidade da desigualdade

$$(5) \frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} < (t\sqrt{2l})^k \cdot n^{\frac{k-2}{2}}$$

é menor que a de (2). Se ainda sujeitamos n à condição

$$\varepsilon < (t\sqrt{2l})^k \cdot n^{\frac{k-2}{2}},$$

ou

$$n > \left(\frac{\varepsilon}{(t\sqrt{2l})^k} \right)^{\frac{2}{k-2}} = n'_1,$$

a probabilidade de (1) será menor que a de (5). Conclui-se que a probabilidade da desigualdade (1) é menor que η para todos os valores de n superiores ao maior dos números n'_0 e n'_1 , o que demonstra o teorema.

53. O teorema I contém uma restrição de que podemos libertar-nos. Não é, com efeito, necessário que exis-

tam as esperanças $b_i^{(2+\delta)}$ para que êle tenha lugar. Basta que existam a_i e b_i , e que o conjunto $b_1, b_2, \dots, b_n, \dots$ seja limitado superiormente. Para demonstrar isto, em vez do teorema de Liapunoff, que já não podemos invocar, utilizaremos a conhecida desigualdade de Tchebychef⁽¹⁾.

Sejam x_1, x_2, \dots, x_n variáveis aleatórias independentes, a_1, a_2, \dots, a_n as suas esperanças, b_1, b_2, \dots, b_n as esperanças de

$$(x_1 - a_1)^2, (x_2 - a_2)^2, \dots, (x_n - a_n)^2$$

e t um número qualquer. A probabilidade de que a diferença

$$x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)$$

pertença ao intervalo de extremidades

$$\mp t\sqrt{b_1 + b_2 + \dots + b_n}$$

é maior que

$$1 - \frac{1}{t^2}.$$

A demonstração do teorema I, liberto da restrição que referimos, a partir desta proposição elementar é muito simples. Façamos com efeito

$$t^2 = \frac{1}{\eta};$$

a probabilidade das desigualdades

$$-\sqrt{\frac{B_n}{\eta}} \leq x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n) \leq +\sqrt{\frac{B_n}{\eta}},$$

(1) *Des valeurs moyennes*, in-*Journal de Liouville*, 11.º série, t. XII, 1867.

ou, por ser $k > 0$, da desigualdade

$$\frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} \leq \frac{1}{n} \left(\frac{B_n}{\eta} \right)^{\frac{k}{2}},$$

é maior que $1 - \eta$; ora, visto ser

$$B_n < nL$$

e, portanto,

$$\frac{1}{n} \left(\frac{B_n}{\eta} \right)^{\frac{k}{2}} < \left(\frac{L}{\eta} \right)^{\frac{k}{2}} \cdot n^{\frac{k-2}{2}},$$

será *a fortiori* maior que $1 - \eta$ a probabilidade da desigualdade

$$\frac{|x_1 + x_2 + \dots + x_n - (a_1 + a_2 + \dots + a_n)|^k}{n} < \varepsilon$$

desde que se dêem a n valores que satisfaçam à desigualdade

$$\left(\frac{L}{\eta} \right)^{\frac{k}{2}} \cdot n^{\frac{k-2}{2}} < \varepsilon,$$

ou

$$n > \left(\frac{1}{\varepsilon} \left(\frac{L}{\eta} \right)^{\frac{k}{2}} \right)^{\frac{2}{2-k}},$$

e o teorema fica demonstrado.

54. No caso particular de ser $k=1$, o teorema I, sob a forma mais extensa que acabamos de lhe dar, é idêntico ao teorema de Bernoulli generalizado, que se encontra já em Laplace e em Poisson, mas de que Tchebychef foi o primeiro a dar demonstração rigorosa, aliás muito sim-

ples, na Memória citada no número precedente. Para

$$(1) \quad x_i = \begin{cases} 1 & \text{com a probabilidade } p_i, \\ 0 & \text{com a probabilidade } q_i, \end{cases}$$

$$p_i + q_i = 1, \quad i = 1, 2, 3, \dots,$$

êste teorema reduz-se ao teorema de Poisson⁽¹⁾, no qual com efeito está contido o velho teorema da *Ars conjectandi*.

A aplicação do teorema I, sob a forma do número anterior, às variáveis (1), para as quais é

$$\text{esp } x_i = p_i, \quad \text{esp } (x_i - p_i)^2 = p_i q_i,$$

dá imediatamente o teorema seguinte.

a) Seja dada uma sucessão infinita de provas independentes e sejam

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

as probabilidades respectivas dum acontecimento A. Se ε e η são dois números positivos arbitrários, para valores de n suficientemente grandes a probabilidade da desigualdade

$$(2) \quad \frac{|m - (p_1 + p_2 + \dots + p_n)|^k}{n} < \varepsilon,$$

onde m designa a freqüência absoluta de A e k é um número positivo menor que 2, é superior a $1 - \eta$.

A aplicação do teorema II às mesmas variáveis dá o teorema que segue.

b) Sejam

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

(1) *Op. cit.*, cap., iv, §§ 94 a 96.



as probabilidades dum acontecimento A numa sucessão infinita de provas independentes e

$$q_1, q_2, \dots, q_n, \dots$$

as probabilidades do acontecimento contraditório. Se a sucessão numérica

$$p_1 q_1, p_2 q_2, \dots, p_n q_n, \dots$$

é limitada inferiormente por um número positivo e se ε e η designam dois números positivos arbitrários, para valores de n suficientemente grandes a probabilidade da desigualdade (2), onde m tem a mesma significação e k é um número maior que 2, é inferior a η .

Basta mostrar, com efeito, que as variáveis (1) satisfazem neste caso às condições de Liapounoff. Ora, temos aqui

$$b_i^{(2+\delta)} = \text{esp} |x_i - p_i|^{2+\delta} = p_i q_i (p_i^{1+\delta} + q_i^{1+\delta}) \leq p_i q_i$$

para $\delta > 0$, e

$$B_n = \sum_1^n p_i q_i;$$

donde resulta que o quociente

$$\frac{\sum_1^n b_i^{(2+\delta)}}{B_n^{1+\frac{\delta}{2}}} \leq \frac{1}{\left(\sum_1^n p_i q_i\right)^{\frac{\delta}{2}}}$$

tende para zero com $\frac{1}{n}$, por ser divergente a série

$$\sum_1^\infty p_i q_i.$$

Os corolários de *a*) e *b*) relativos ao caso especial de serem iguais a *p* tôdas as probabilidades $p_1, p_2, \dots, p_n, \dots$ contêm as duas proposições seguintes, demonstradas pelo sr. Dr. Pacheco de Amorim, por um método muito diferente, na sua Dissertação Inaugural (1).

1) A probabilidade de que $\alpha = |m - np|$ seja tal que

$$\frac{\alpha^{\rho+1}}{n^\rho} \geq \varepsilon \quad (\rho > 1),$$

onde ε é um número positivo arbitrário, tende para zero com $\frac{1}{n}$.

2) A probabilidade de que seja

$$\frac{\alpha}{\sqrt[\rho]{n}} < \varepsilon \quad (\rho > 2)$$

tende para zero com $\frac{1}{n}$.

A desigualdade da proposição 2) é com efeito equivalente a

$$\frac{|m - np|^\rho}{n} < \varepsilon^\rho,$$

i. é, a proposição 2) é idêntica ao corolário de *b*).

A proposição 1) afirma que a desigualdade contraditória de

$$\frac{|m - np|^{1 + \frac{1}{\rho}}}{n} \geq \varepsilon^{\frac{1}{\rho}}$$

(1) *Elementos de Cálculo das Probabilidades*, Coimbra, 1914; cap. v, propos. ix e x.

tem uma probabilidade que tende para 1; esta proposição está pois contida no corolário de a): corresponde ao caso $1 < k < 2$.

Os dois corolários referidos podem demonstrar-se directamente, pelo mesmo método que os teoremas I e II, a partir do teorema de Laplace, segundo o qual a probabilidade das desigualdades

$$np + t_1 \sqrt{2npq} < m < np + t_2 \sqrt{2npq}$$

tende para o limite

$$\frac{1}{\sqrt{\pi}} \int_{t_1}^{t_2} e^{-t^2} dt$$

quando n cresce infinitamente. Êste teorema é um caso particularíssimo do teorema de Liapunoff. Tchebychef deu dêle uma demonstração rigorosa e elementar, que Markoff expõe com muita elegância no seu Tratado (4).

§ 16.º

Justificação do principio de Gauss no problema da combinação das observações

55. Consideremos primeiro o caso das observações directas. Seja a uma grandeza desconhecida e sejam a_1, a_2, \dots, a_n as medidas desta grandeza em n observações.

(4) *Op. cit.*, Kap. II, § 9.

Poderemos escrever as equações aproximadas

$$(1) \quad a \neq a_1, a \neq a_2, \dots, a \neq a_n.$$

Designando por x_1, x_2, \dots, x_n os erros possíveis das observações efectuadas, que suporemos independentes, e admitindo que as observações são desprovidas de erro constante, teremos

$$\text{esp } x_i = 0, \quad i = 1, 2, \dots, n.$$

Propomo-nos procurar a combinação linear das equações (1) que os princípios da teoria das probabilidades mostrem ser a mais conveniente. Precisando o sentido destas palavras, diremos que se trata, em rigor, de determinar um sistema de factores m_1, m_2, \dots, m_n pelos quais devam ser multiplicadas respectivamente, e depois somadas, as equações (1), de modo que seja máxima a probabilidade de que o erro da equação resultante, depois de dividida pela soma $m_1 + m_2 + \dots + m_n$ suposta diferente de zero, fique compreendido num intervalo arbitrário $(-\varepsilon, +\varepsilon)$. Mas é claro que a questão assim posta é insolúvel, porque a expressão rigorosa desta probabilidade é tão inteiramente desconhecida como as leis de probabilidade dos erros das n observações, e que a razão desta insolubilidade não reside na imperfeição da sciência das probabilidades, mas na natureza das coisas. O que nos permite pô-la dum modo levemente diferente, que comporte solução, é o facto notável descoberto por Laplace, demonstrado rigorosamente e sucessivamente generalizado por Tchebychef e Liapunoff, de que, sob certas condições, aquela probabilidade tende para um limite de expressão perfeitamente determinada e muito simples quando n cresce infinitamente. Procuraremos portanto a combinação linear das

equações (1) que dá máximo para o limite da probabilidade de que o erro, que a afecta, pertença ao intervalo arbitrário $(-\varepsilon, +\varepsilon)$. Ver-se há que tal combinação é independente de ε . Esta solução aproxima-se da solução ideal impossível tanto mais quanto maior fôr n .

Uma combinação linear qualquer das equações (1) é da forma

$$(2) \quad a = q_1 a_1 + q_2 a_2 + \dots + q_n a_n,$$

com

$$q_i = \frac{m_i}{m_1 + m_2 + \dots + m_n}, \quad i = 1, 2, \dots, n;$$

os q_i estão pois ligados pela equação

$$(3) \quad q_1 + q_2 + \dots + q_n = 1.$$

O erro possível da equação (2) é a variável aleatória

$$q_1 x_1 + q_2 x_2 + \dots + q_n x_n.$$

Façamos

$$\text{esp } x_i^2 = b_i;$$

teremos

$$\text{esp } (q_i x_i) = 0, \quad \text{esp } (q_i x_i)^2 = q_i^2 b_i,$$

$$\text{esp } (q_1 x_1 + q_2 x_2 + \dots + q_n x_n) = 0.$$

A equação (2) não tem pois erro constante.

Seja $(-\varepsilon, +\varepsilon)$ um intervalo arbitrariamente dado, e consideremos a equação

$$\varepsilon = t \sqrt{2(q_1^2 b_1 + q_2^2 b_2 + \dots + q_n^2 b_n)}$$

com as variáveis t, q_1, q_2, \dots, q_n . Vejamos que sistema

de valores destas variáveis torna máximo o integral

$$(4) \quad \frac{1}{\sqrt{\pi}} \int_{-t}^{+t} e^{-z^2} dz,$$

considerado por ora como simples valor aproximado da probabilidade da desigualdade

$$(5) \quad |q_1 x_1 + q_2 x_2 + \dots + q_n x_n| < \varepsilon.$$

É claro que o integral é máximo para t máximo, e que t é máximo quando a soma

$$(6) \quad q_1^2 b_1 + q_2^2 b_2 + \dots + q_n^2 b_n$$

fôr mínima. As condições de mínimo desta função, em virtude de (3), obtêm-se igualando a zero as derivadas parciais de

$$f = q_1^2 b_1 + \dots + q_n^2 b_n - 2\mu(q_1 + \dots + q_n - 1),$$

considerando μ como constante; são portanto

$$q_i b_i - \mu = 0, \quad i = 1, 2, \dots, n,$$

ou, pondo

$$\frac{1}{b_i} = g_i$$

e atendendo a (3),

$$(7) \quad q_i = \frac{g_i}{g_1 + g_2 + \dots + g_n}$$

Há realmente um mínimo de (6) para estes valores dos q_i ,

por ser

$$\frac{\partial^2 f}{\partial q_i \partial q_j} = \begin{cases} 2b_i > 0 & \text{se } j = i, \\ 0 & \text{se } j \neq i. \end{cases}$$

Para os valores (7) dos q_i , que tornam máximo o integral (4), a desigualdade (5) pode escrever-se, como é fácil de ver,

$$(8) \quad |g_1 x_1 + g_2 x_2 + \dots + g_n x_n| < t \sqrt{2(g_1 + g_2 + \dots + g_n)}.$$

Por ser

$$\text{esp}(g_i x_i) = 0, \quad \text{esp}(g_i x_i)^2 = g_i^2 b_i = g_i,$$

a expressão $g_1 + g_2 + \dots + g_n$ é a soma das esperanças dos quadrados das variáveis aleatórias $g_i x_i$.

No § precedente demos o enunciado do teorema de Liapunoff. Mencionaremos aqui um corolário d'este teorema que nos vai ser útil.

Com as notações daquele §, suponhamos que as variáveis aleatórias independentes $x_i - a_i$ são limitadas superiormente em valor absoluto por um mesmo número M , e suponhamos além disso que a soma $b_1 + b_2 + \dots + b_n$ cresce infinitamente com n . Será

$$b_i^{(2+\delta)} = \text{esp} |x_i - a_i|^{2+\delta} < M^\delta b_i, \quad i = 1, 2, \dots, n, \dots$$

e também

$$\frac{b_1^{(2+\delta)} + \dots + b_n^{(2+\delta)}}{(b_1 + \dots + b_n)^{1 + \frac{\delta}{2}}} < \frac{M^\delta}{(b_1 + \dots + b_n)^{\frac{\delta}{2}}} \rightarrow 0$$

para $n \rightarrow \infty$. As condições de Liapunoff são verificadas.

Este corolário permite afirmar, voltando à questão que nos ocupa, que a probabilidade da desigualdade (8) tem por limite o integral (4) para $n \rightarrow \infty$, se fôr

$$|g_i x_i| < M, \quad i = 1, 2, \dots, n, \dots,$$

e

$$g_1 + g_2 + \dots + g_n \rightarrow \infty.$$

Ora estas condições verificam-se desde que seja

$$|x_i| < N,$$

$$0 < l < g_i < L,$$

$$i = 1, 2, \dots, n, \dots,$$

i. é, desde que os valores absolutos dos êrros possíveis não possam exceder um certo número e que os números g_i , proporcionais aos pesos das observações, não possam tornar-se inferiores a certo limite positivo nem superiores a outro. É o que se dá efectivamente com as boas observações.

Resulta daqui que, para os valores (7) dos q_i , o integral (4) não só é um valor aproximado da probabilidade da desigualdade (5), mas é também o limite para que tende esta probabilidade quando o número de observações cresce infinitamente; e portanto a combinação linear (2) correspondente a êsses valores dos q_i é a combinação mais conveniente que procurávamos.

Observando que se tem

$$\begin{aligned} \text{esp}(q_1 x_1 + \dots + q_n x_n)^2 &= \text{esp}(q_1^2 x_1^2 + \dots + q_n^2 x_n^2 + 2q_1 q_2 x_1 x_2 + \dots) = \\ &= q_1^2 b_1 + \dots + q_n^2 b_n \end{aligned}$$

por serem independentes x_1, x_2, \dots, x_n e ser consequen-

temente

$$\text{esp}(x_i x_j) = \text{esp } x_i \cdot \text{esp } x_j = 0, \quad i \leq j,$$

conclui-se que a combinação linear mais conveniente das observações dadas é a que torna mínima a esperança do quadrado do erro possível da incógnita. É o princípio que Gauss postulou na sua segunda teoria da combinação das observações (1) e que conduz, como é sabido, ao método dos menores quadrados.

A combinação mais conveniente é independente de ε , como tínhamos afirmado. Além disso, por ser

$$q_1^2 b_1 + \dots + q_n^2 b_n = \frac{1}{g_1 + \dots + g_n},$$

pondo

$$g_1 + g_2 + \dots + g_n = 2h^2$$

tem-se a igualdade

$$t = h\varepsilon,$$

e portanto a probabilidade da desigualdade (5), quando n é suficientemente grande, é praticamente igual a

$$\frac{1}{\sqrt{\pi}} \int_{-h\varepsilon}^{+h\varepsilon} e^{-z^2} dz.$$

Fazendo a substituição

$$z = h\xi,$$

vê-se que a probabilidade de que o erro possível da in-

(1) *Theoria combinationis observationum erroribus minimis obnoxiae: Pars prior*, 1821; *Pars posterior*, 1823; *Supplementum*, 1826. In-*Werke*, Bd. IV, 1873.

cógnita, na combinação mais conveniente dum grande número de observações, seja um número do intervalo $(-\varepsilon, +\varepsilon)$ é dada praticamente pelo integral

$$\frac{h}{\sqrt{\pi}} \int_{-\varepsilon}^{+\varepsilon} e^{-h^2 \xi^2} d\xi,$$

e satisfaz por conseguinte à lei exponencial que Gauss, na sua primeira teoria do método dos menores quadrados, pretendia impôr ao êrro de cada observação.

56. Consideremos agora o caso das observações indirectas. Para simplificar a escrita, suporemos o número de incógnitas igual a três. Mas o método é geral.

Como no n.º anterior, admitimos que as observações são independentes e desprovidas de êrros constantes.

Suponhamos que n observações nos forneceram as equações aproximadas

$$(1) \quad \left\{ \begin{array}{l} H_1 \alpha + K_1 \beta + L_1 \gamma \cong \delta_1, \\ H_2 \alpha + K_2 \beta + L_2 \gamma \cong \delta_2, \\ \dots \dots \dots \\ H_n \alpha + K_n \beta + L_n \gamma \cong \delta_n, \end{array} \right.$$

onde $\delta_1, \delta_2, \dots, \delta_n$ são os resultados directos das observações, α, β, γ as incógnitas, $H_i, K_i, L_i, (i=1, 2, \dots, n)$ números conhecidos e $n > 3$. Para obter a combinação mais conveniente das n observações, procuraremos três sistemas de factores

$$\begin{array}{l} h_1, h_2, \dots, h_n \\ k_1, k_2, \dots, k_n \\ l_1, l_2, \dots, l_n \end{array}$$

pelos quais se devam multiplicar sucessivamente as equações (1) afim de que as três equações resultantes da adição dos produtos das equações dadas pelos factores de cada sistema dêem para as incógnitas valores aproximados tais, que seja máximo o limite para $n \rightarrow \infty$ da probabilidade de que o vector aleatório, cujas componentes são os erros possíveis destes valores aproximados, tenha a extremidade dentro dum certo elipsoide de volume arbitrário ε , concêntrico à origem do vector. Ver-se há que os sistemas de factores encontrados são independentes de ε .

Sejam

$$x_1, x_2, \dots, x_n$$

os erros possíveis das equações (1). Tem-se

$$\text{esp } x_i = 0, \quad i = 1, 2, \dots, n,$$

por não existirem erros constantes.

As três equações formadas como indicámos a partir de (1) dão para α, β, γ valores aproximados da forma

$$(2) \quad \begin{cases} \alpha \approx q_1 \delta_1 + \dots + q_n \delta_n, \\ \beta \approx r_1 \delta_1 + \dots + r_n \delta_n, \\ \gamma \approx s_1 \delta_1 + \dots + s_n \delta_n, \end{cases}$$

com

$$q_i = \frac{D'_i}{D}, \quad r_i = \frac{D''_i}{D}, \quad s_i = \frac{D'''_i}{D},$$

$$i = 1, 2, \dots, n,$$

onde D representa o determinante, suposto diferente de zero,

$$\begin{vmatrix} \Sigma h_i H_i & \Sigma h_i K_i & \Sigma h_i L_i \\ \Sigma k_i H_i & \Sigma k_i K_i & \Sigma k_i L_i \\ \Sigma l_i H_i & \Sigma l_i K_i & \Sigma l_i L_i \end{vmatrix}$$

e D'_i, D''_i, D'''_i são os determinantes que se obtêm substituindo em D , respectivamente, a primeira, a segunda, a terceira colunas por uma coluna formada pelos números h_i, k_i, l_i .

Entre os coeficientes q_i, r_i, s_i existem as relações evidentes

$$(3) \quad \begin{cases} \Sigma q_i H_i = 1, & \Sigma q_i K_i = 0, & \Sigma q_i L_i = 0, \\ \Sigma r_i H_i = 0, & \Sigma r_i K_i = 1, & \Sigma r_i L_i = 0, \\ \Sigma s_i H_i = 0, & \Sigma s_i K_i = 0, & \Sigma s_i L_i = 1. \end{cases}$$

Os erros possíveis das equações aproximadas (2) são respectivamente

$$(4) \quad \begin{cases} q_1 x_1 + \dots + q_n x_n, \\ r_1 x_1 + \dots + r_n x_n, \\ s_1 x_1 + \dots + s_n x_n. \end{cases}$$

Façamos

$$b_i = \text{esp } x_i^2, \quad i = 1, 2, \dots, n,$$

e consideremos os n vectores independentes

$$(q_1 x_1, r_1 x_1, s_1 x_1), \dots, (q_n x_n, r_n x_n, s_n x_n);$$

teremos

$$\begin{aligned} \text{esp } (q_i x_i) &= 0, \dots, \quad \text{esp } (q_i x_i)^2 = q_i^2 b_i, \dots, \\ \text{esp } (q_1 x_1 + \dots + q_n x_n) &= 0, \dots \end{aligned}$$

Sabe-se⁽¹⁾ que a probabilidade de que a soma destes vectores, aplicada à origem das coordenadas cartesianas ortogonais t, t', t'' , tenha a extremidade dentro dum vo-

(1) Cf. p. ex. Markoff, *op. cit.*, Kap. v, § 33. Markoff só considera vectores de duas componentes.

lume V , é aproximadamente igual a

$$(5) \quad \frac{1}{(\sqrt{2\pi})^3 \sqrt{\Delta}} \iiint_V e^{-\frac{R(t, t', t'')}{2\Delta}} dt dt' dt'',$$

onde $R(t, t', t'')$ é a forma quadrática definida positiva

$$Mt^2 + M't'^2 + M''t''^2 + 2Nt't' + 2N't''t + 2N''t't'',$$

cujos coeficientes são os complementos algébricos dos elementos A, A', A'', B, B', B'' do determinante

$$\Delta = \begin{vmatrix} A & B'' & B' \\ B'' & A' & B \\ B' & B & A'' \end{vmatrix}$$

com

$$\begin{aligned} A &= q_1^2 b_1 + \dots + q_n^2 b_n, & B &= r_1 s_1 b_1 + \dots + r_n s_n b_n, \\ A' &= r_1^2 b_1 + \dots + r_n^2 b_n, & B' &= s_1 q_1 b_1 + \dots + s_n q_n b_n, \\ A'' &= s_1^2 b_1 + \dots + s_n^2 b_n, & B'' &= q_1 r_1 b_1 + \dots + q_n r_n b_n. \end{aligned}$$

A equação

$$(6) \quad \frac{R(t, t', t'')}{2\Delta} = u,$$

onde u é um parâmetro variável nulo ou positivo, representa uma família de elipsóides reais com o centro na origem das coordenadas; o volume dum destes elipsóides é, como se sabe, expresso por

$$\frac{4}{3} \pi \sqrt{\frac{(2\Delta u)^3}{\Delta_1}},$$

onde Δ_1 é o determinante adjunto de Δ :

$$\Delta_1 = \begin{vmatrix} M & N'' & N' \\ N'' & M' & N \\ N' & N & M' \end{vmatrix}.$$

A relação conhecida

$$\Delta_1 = \Delta^2$$

dá à expressão do volume do elipsóide a forma

$$(7) \quad \frac{8}{3} \pi \sqrt{2\Delta} u^{\frac{3}{2}}.$$

Suponhamos que o volume V é a camada compreendida entre as superfícies infinitamente vizinhas dos dois elipsóides correspondentes aos valores u e $u + du$ do parâmetro. Atendendo a que a função integranda da expressão (5) é igual a

$$e^{-u}$$

nesta camada infinitesimal, e a que o volume da mesma camada é

$$4\pi \sqrt{2\Delta} \sqrt{u} du,$$

como se reconhece diferenciando (7), vê-se facilmente que a probabilidade de que o vector êrro das equações (2), cujas componentes são as somas (4), tenha a extremidade na camada considerada é dada aproximadamente pela expressão

$$\frac{2}{\sqrt{\pi}} e^{-u} \sqrt{u} du.$$

Portanto, a probabilidade de que a extremidade do vector

erro esteja dentro do elipsóide correspondente ao valor u do parâmetro tem por valor aproximado

$$(8) \quad \frac{2}{\sqrt{\pi}} \int_0^u e^{-u} \sqrt{u} du.$$

Seja ε um número positivo arbitrariamente dado e consideremos a equação

$$\frac{8}{3} \pi \sqrt{2 \Delta} u^{\frac{3}{2}} = \varepsilon,$$

com as variáveis u, q_i, r_i, s_i ($i=1, 2, \dots, n$). A expressão (8) dá a probabilidade aproximada de que o vector erro tenha a extremidade dentro do elipsóide da família (6) de volume ε . Esta probabilidade é evidentemente máxima quando u fôr máximo; mas u será máximo quando Δ fôr mínimo, e portanto a combinação linear das equações (1) que dá máximo à probabilidade aproximada (8) é aquela em que os q_i, r_i, s_i tornam mínimo o determinante

$$\begin{vmatrix} A & B'' & B' \\ B'' & A' & B \\ B' & B & A'' \end{vmatrix}$$

Ora, não é difícil mostrar que a combinação, que torna mínimo este determinante, torna também mínimas as funções A, A', A'' .

As condições de mínimo de Δ obtêm-se com efeito, por virtude de (3), igualando a zero as derivadas parciais da

função

$$\begin{aligned} \Delta - 2\lambda (\Sigma q_i H_i - 1) - 2\lambda' \Sigma q_i K_i - 2\lambda'' \Sigma q_i L_i \\ - 2\mu \Sigma r_i H_i - 2\mu' (\Sigma r_i K_i - 1) - 2\mu'' \Sigma r_i L_i \\ - 2\nu \Sigma s_i H_i - 2\nu' \Sigma s_i K_i - 2\nu'' (\Sigma s_i L_i - 1), \end{aligned}$$

onde $\lambda, \lambda', \lambda'', \mu, \mu', \mu'', \nu, \nu', \nu''$ devem considerar-se como constantes, o que dá as equações

$$\frac{\partial \Delta}{\partial q_i} - 2\lambda H_i - 2\lambda' K_i - 2\lambda'' L_i = 0,$$

$$\frac{\partial \Delta}{\partial r_i} - 2\mu H_i - 2\mu' K_i - 2\mu'' L_i = 0,$$

$$\frac{\partial \Delta}{\partial s_i} - 2\nu H_i - 2\nu' K_i - 2\nu'' L_i = 0,$$

$$i = 1, 2, \dots, n.$$

Encontra-se sem dificuldade que as derivadas $\frac{\partial \Delta}{\partial q_i}, \frac{\partial \Delta}{\partial r_i}, \frac{\partial \Delta}{\partial s_i}$ valem respectivamente

$$2b_i (M q_i + N'' r_i + N' s_i),$$

$$2b_i (N'' q_i + M' r_i + N s_i),$$

$$2b_i (N' q_i + N r_i + M'' s_i),$$

e assim as condições de mínimo encontradas podem escrever-se, pondo $\frac{1}{b_i} = g_i$,

$$M q_i + N'' r_i + N' s_i = g_i (\lambda H_i + \lambda' K_i + \lambda'' L_i),$$

$$N'' q_i + M' r_i + N s_i = g_i (\mu H_i + \mu' K_i + \mu'' L_i),$$

$$N' q_i + N r_i + M'' s_i = g_i (\nu H_i + \nu' K_i + \nu'' L_i),$$

$$i = 1, 2, \dots, n.$$

Destas equações resultam imediatamente as equações equivalentes

$$(9) \quad \begin{cases} q_i = \frac{g_i}{\Delta^2} (\zeta H_i + \zeta' K_i + \zeta'' L_i), \\ r_i = \frac{g_i}{\Delta^2} (\eta H_i + \eta' K_i + \eta'' L_i), \\ s_i = \frac{g_i}{\Delta^2} (\theta H_i + \theta' K_i + \theta'' L_i), \end{cases}$$

onde ζ, η, θ se obtêm substituindo respectivamente em Δ_1 a primeira, a segunda, a terceira colunas por uma coluna formada pelos elementos λ, μ, ν ; e $\zeta', \eta', \theta', \zeta'', \eta'', \theta''$ têm expressões semelhantes com as letras λ, μ, ν acentuadas uma e duas vezes.

Em virtude das três primeiras equações (3), a primeira (9) fornece as três seguintes

$$(10) \quad \begin{cases} \frac{\zeta}{\Delta^2} \sum g_i H_i^2 + \frac{\zeta'}{\Delta^2} \sum g_i H_i K_i + \frac{\zeta''}{\Delta^2} \sum g_i H_i L_i = 1, \\ \frac{\zeta}{\Delta^2} \sum g_i H_i K_i + \frac{\zeta'}{\Delta^2} \sum g_i K_i^2 + \frac{\zeta''}{\Delta^2} \sum g_i K_i L_i = 0, \\ \frac{\zeta}{\Delta^2} \sum g_i H_i L_i + \frac{\zeta'}{\Delta^2} \sum g_i K_i L_i + \frac{\zeta''}{\Delta^2} \sum g_i L_i^2 = 0, \end{cases}$$

que dão os valores que as fracções

$$\frac{\zeta}{\Delta^2}, \frac{\zeta'}{\Delta^2}, \frac{\zeta''}{\Delta^2}$$

possuem na combinação que torna Δ mínimo, os quais, substituídos na primeira (9), fornecem o valor procurado de q_i ; verifica-se facilmente, com efeito, que o determinante do sistema (10) é positivo.

Procuremos agora as condições de mínimo de A . Devemos igualar a zero as derivadas parciais da função

$$A - 2\rho(\Sigma q_i H_i - 1) - 2\rho' \Sigma q_i K_i - 2\rho'' \Sigma q_i L_i,$$

considerando ρ , ρ' , ρ'' como constantes, o que dá as equações

$$\frac{\partial A}{\partial q_i} - 2\rho H_i - 2\rho' K_i - 2\rho'' L_i = 0,$$

ou

$$(11) \quad q_i = g_i(\rho H_i + \rho' K_i + \rho'' L_i), \\ i = 1, 2, \dots, n.$$

Verifica-se como no número anterior que estas equações dão efectivamente um mínimo para A . Quanto ao mínimo de Δ , é evidente que êle existe em virtude das considerações com que o introduzimos.

Das equações (11) resultam, pelas três primeiras equações (3), as seguintes

$$\begin{aligned} \rho \Sigma g_i H_i^2 + \rho' \Sigma g_i H_i K_i + \rho'' \Sigma g_i H_i L_i &= 1, \\ \rho \Sigma g_i H_i K_i + \rho' \Sigma g_i K_i^2 + \rho'' \Sigma g_i K_i L_i &= 0, \\ \rho \Sigma g_i H_i L_i + \rho' \Sigma g_i K_i L_i + \rho'' \Sigma g_i L_i^2 &= 0, \end{aligned}$$

que, confrontadas com (10), provam que, para Δ mínimo, os valores de

$$\frac{\xi}{\Delta^2}, \quad \frac{\xi'}{\Delta^2}, \quad \frac{\xi''}{\Delta^2}$$

são iguais a ρ , ρ' , ρ'' respectivamente; e portanto, pelo confronto de (11) com a primeira (9), que os valores dos q_i relativos ao mínimo de Δ tornam mínimo A .

Procedendo do mesmo modo com A' e A'' chegaríamos a resultado análogo. A combinação das equações (1) que torna máxima a probabilidade aproximada (8) é portanto a que torna mínimos A , A' , A'' , como tínhamos afirmado; verifica-se que ela é independente de ε .

Ora A , A' , A'' são as esperanças dos quadrados dos erros (4). Por outro lado, Casteljano demonstrou recentemente (4) que, sob certas condições análogas às do corolário de Liapunoff e que podem admitir-se como realizadas nas boas observações, o integral (5), e portanto o integral (8) que resulta de (5) rigorosamente, não só é um valor aproximado da probabilidade mas é mesmo o limite para que tende a probabilidade quando n cresce infinitamente. Conclui-se que a combinação que acabamos de encontrar é a combinação mais conveniente que procurávamos, e que ela torna mínimas as esperanças dos quadrados dos erros possíveis das incógnitas. É o princípio de Gauss, do qual este géometra deduziu o método dos menores quadrados.

Como observámos no começo deste número, a análise precedente é geral. Partindo do integral

$$\frac{1}{(\sqrt{2\pi})^m \sqrt{\Delta}} \int \int \dots \int_V e^{-\frac{R(t_1, t_2, \dots, t_m)}{2\Delta}} dt_1 dt_2 \dots dt_m,$$

limite da probabilidade de que a soma de n vectores independentes de m componentes, com esperanças nulas, tenha a extremidade dentro dum supervolume V do espaço eu-

(4) *Calcolo delle Probabilità*, vol. II, 2.^a ed., Bolonha, 1928; Apêndice, art. IV, pág. 218. O autor considera vectores de duas componentes, mas o seu método generaliza-se sem dificuldade.

clidiano m -dimensional, e utilizando a expressão

$$\frac{(\pi u)^{\frac{m}{2}}}{\Gamma\left(1 + \frac{m}{2}\right) \sqrt{|a_{ij}|}}$$

do volume do hiperelipsóide

$$a_{11} t_1^2 + \dots + a_{mm} t_m^2 + 2 a_{12} t_1 t_2 + \dots + 2 a_{m-1, m} t_{m-1} t_m = u,$$

na qual $|a_{ij}|$ é o discriminante da forma quadrática definida positiva do primeiro membro desta equação e Γ representa a função euleriana de segunda espécie, acha-se como precedentemente que o limite da probabilidade, de que o vector êrro dum combinação linear de n equações com m incognitas ($n > m$) tenha a extremidade dentro do hiperelipsóide

$$\frac{R(t_1, t_2, \dots, t_m)}{2 \Delta} = u,$$

é igual ao integral

$$(12) \quad \frac{m}{2 \Gamma\left(1 + \frac{m}{2}\right)} \int_0^u e^{-z} z^{\frac{m}{2}-1} dz.$$

Fazendo o volume dêste hiperelipsóide igual a um número positivo arbitrário ε , i. é, pondo

$$(13) \quad \frac{(\sqrt{2} \pi)^m \sqrt{\Delta}}{\Gamma\left(1 + \frac{m}{2}\right)} u^{\frac{m}{2}} = \varepsilon,$$

conclui-se que o limite da probabilidade de que a extre-

midade do vector êrro esteja dentro do hiperelipsóide de volume ε é máxima quando u fôr máximo, e isto dá-se quando Δ fôr mínimo. Sem dificuldade se prova depois que os sistemas de factores, que tornam mínimo o determinante Δ e que dão portanto a combinação mais conveniente, tornam também mínimos os elementos da sua diagonal principal, i. é, as esperanças dos quadrados dos êrros possíveis das incógnitas, o que é o princípio de Gauss.

A fórmula (12) permite calcular aproximadamente a probabilidade de que o vector êrro da combinação mais conveniente tenha a extremidade dentro do hiperelipsóide considerado de volume ε , quando o número de observações é muito grande. Se, em (13), fazemos igual a $\frac{1}{k}$ o coeficiente de $u^{\frac{m}{2}}$, obtemos

$$u = (k\varepsilon)^{\frac{2}{m}}$$

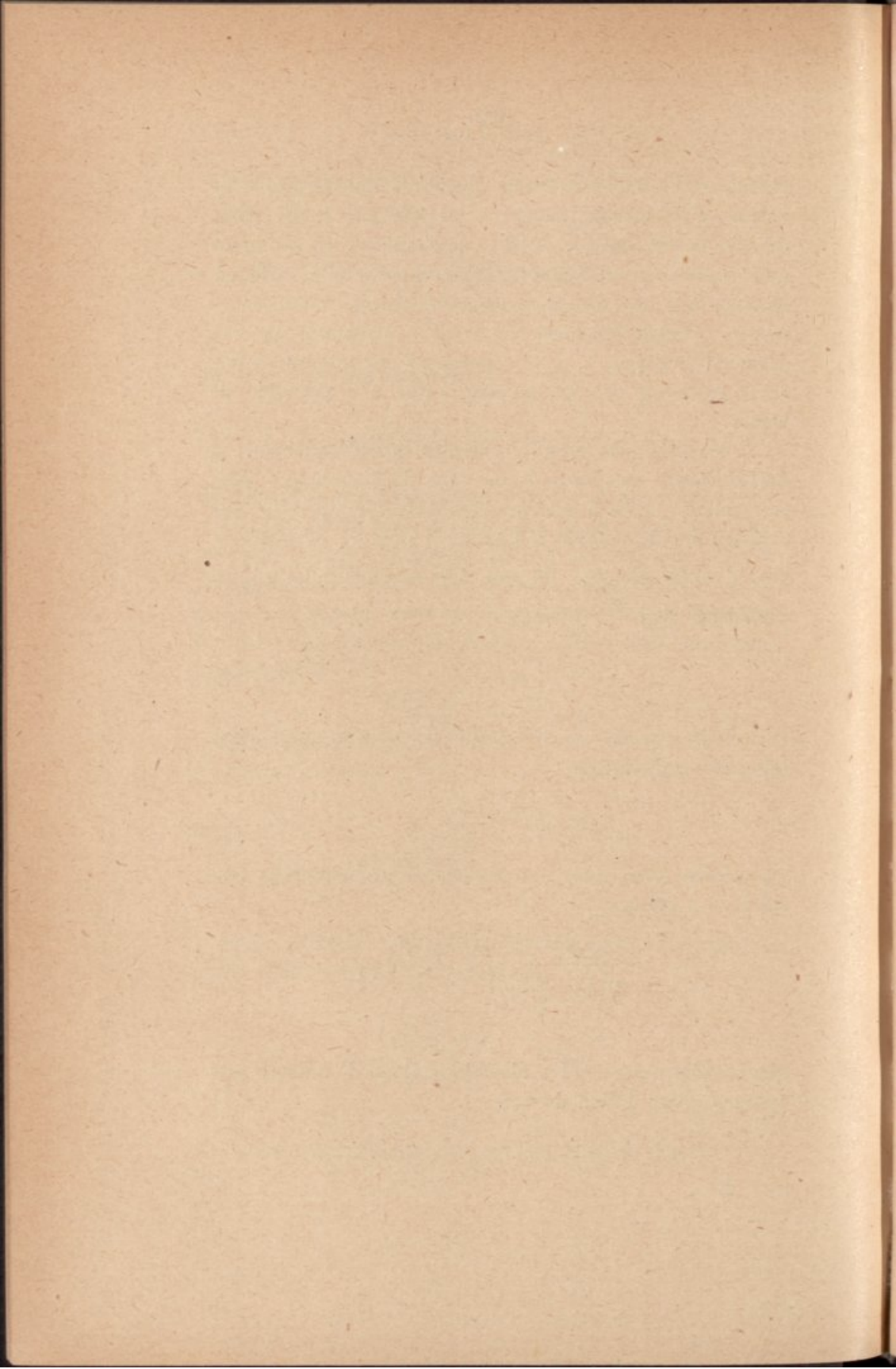
para limite superior do integral (12); fazendo ainda neste integral a substituição

$$z^{\frac{m}{2}} = k\varphi,$$

encontra-se para valor aproximado da probabilidade considerada o integral

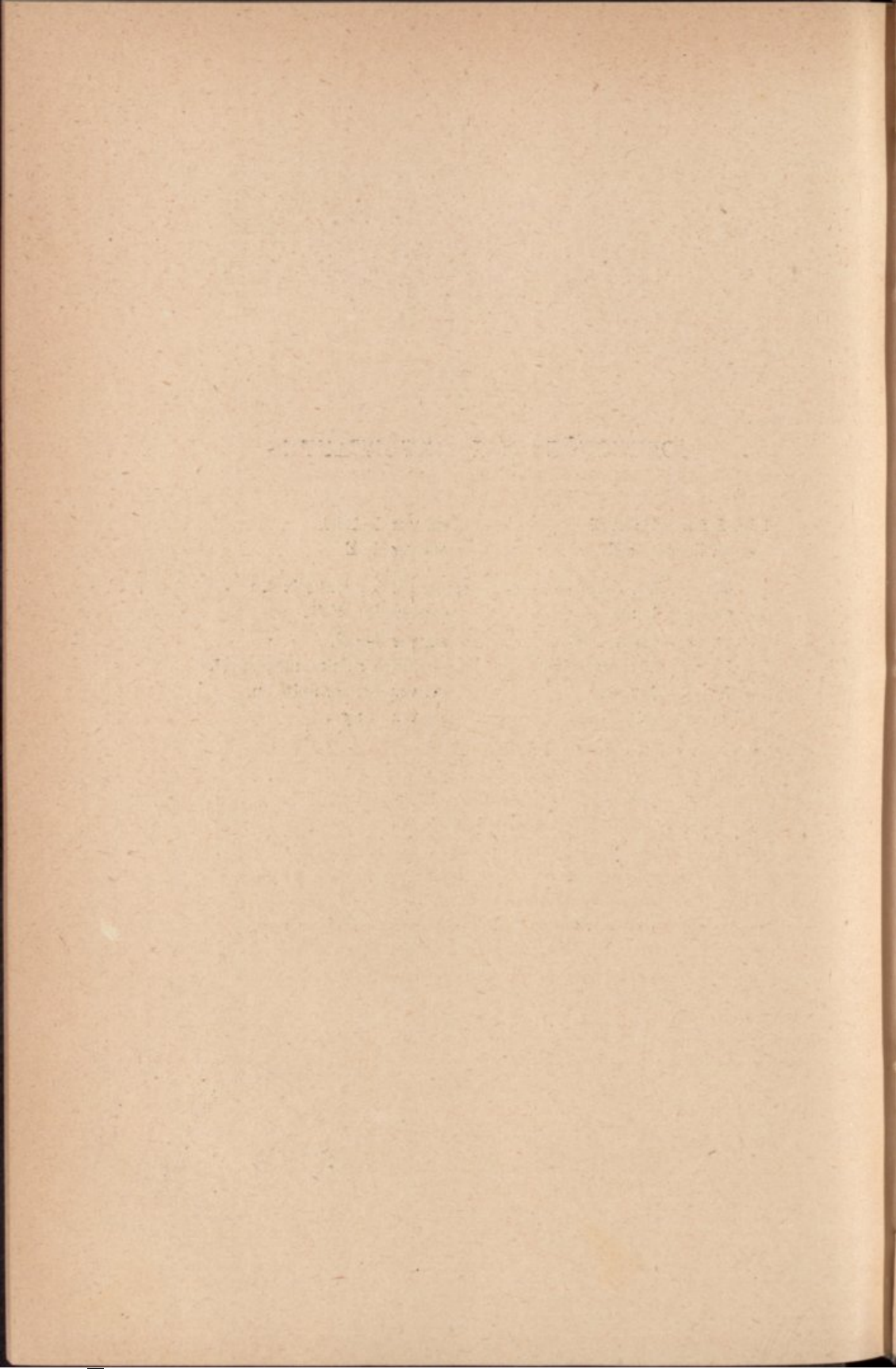
$$\frac{k}{\Gamma\left(1 + \frac{m}{2}\right)} \int_0^{\varepsilon} e^{-\sqrt{l^2}\varphi^2} d\varphi,$$

que contém para $m=1$ a fórmula análoga do número precedente, como é fácil de ver.



CORRECÇÕES MAIS IMPORTANTES

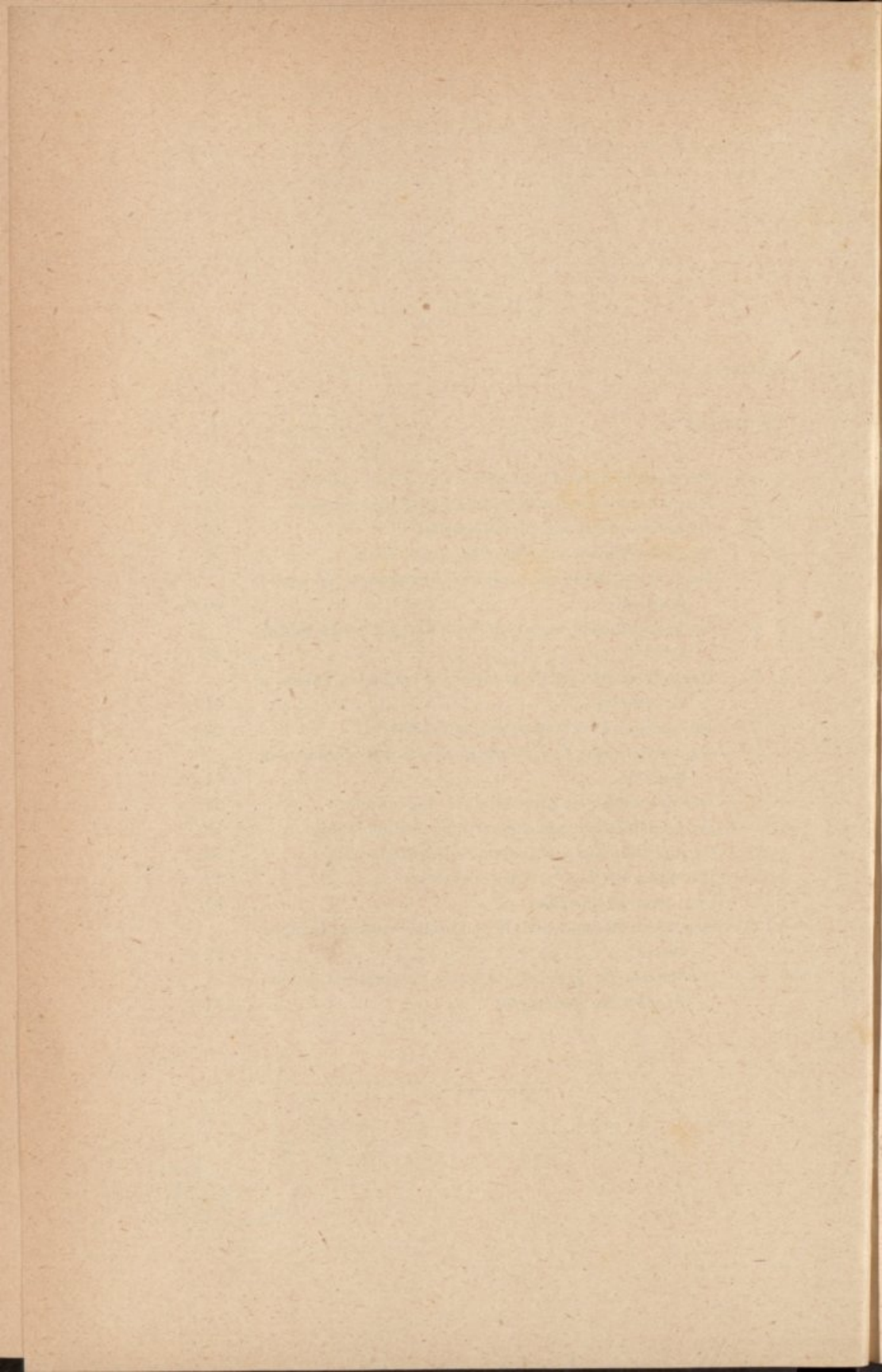
Pág. 4, nota (1): 1820	em vez de 1920.
» 34, linha 12: Σ i	em vez de Σ
» 48, » 20: conjunto mensurável	em vez de: consurável.
» 65, » 11: $A' B$	em vez de: $A' B'$.
» 67, » última: $\overline{B'}$	em vez de: \overline{B} .
» 77, » 13: numerável	em vez de: finito, numerável.
» 77, » 19: ser	em vez de: ser finito ou.
» 108, » 6: $+ g_n$	em vez de: g_n .

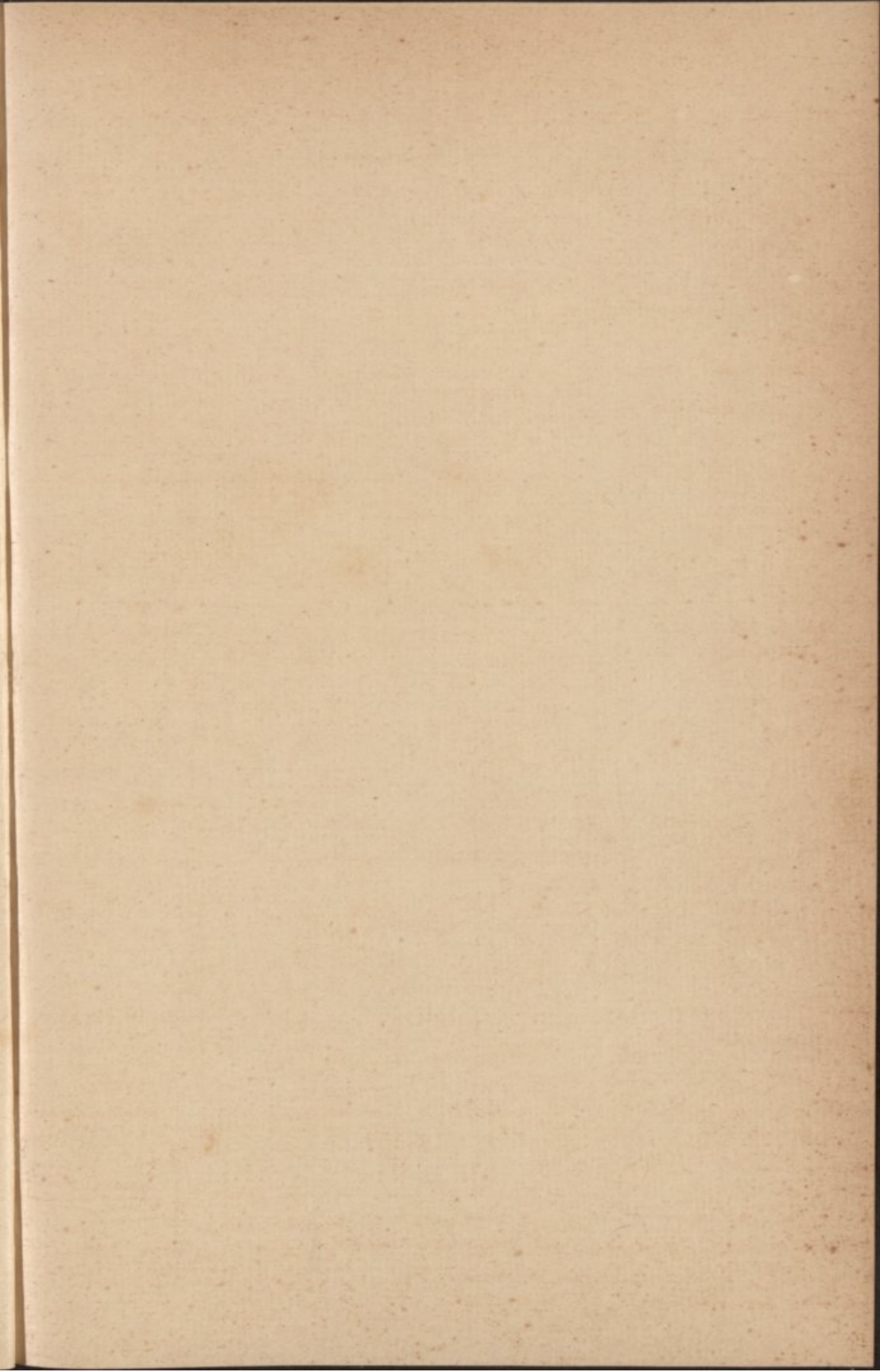


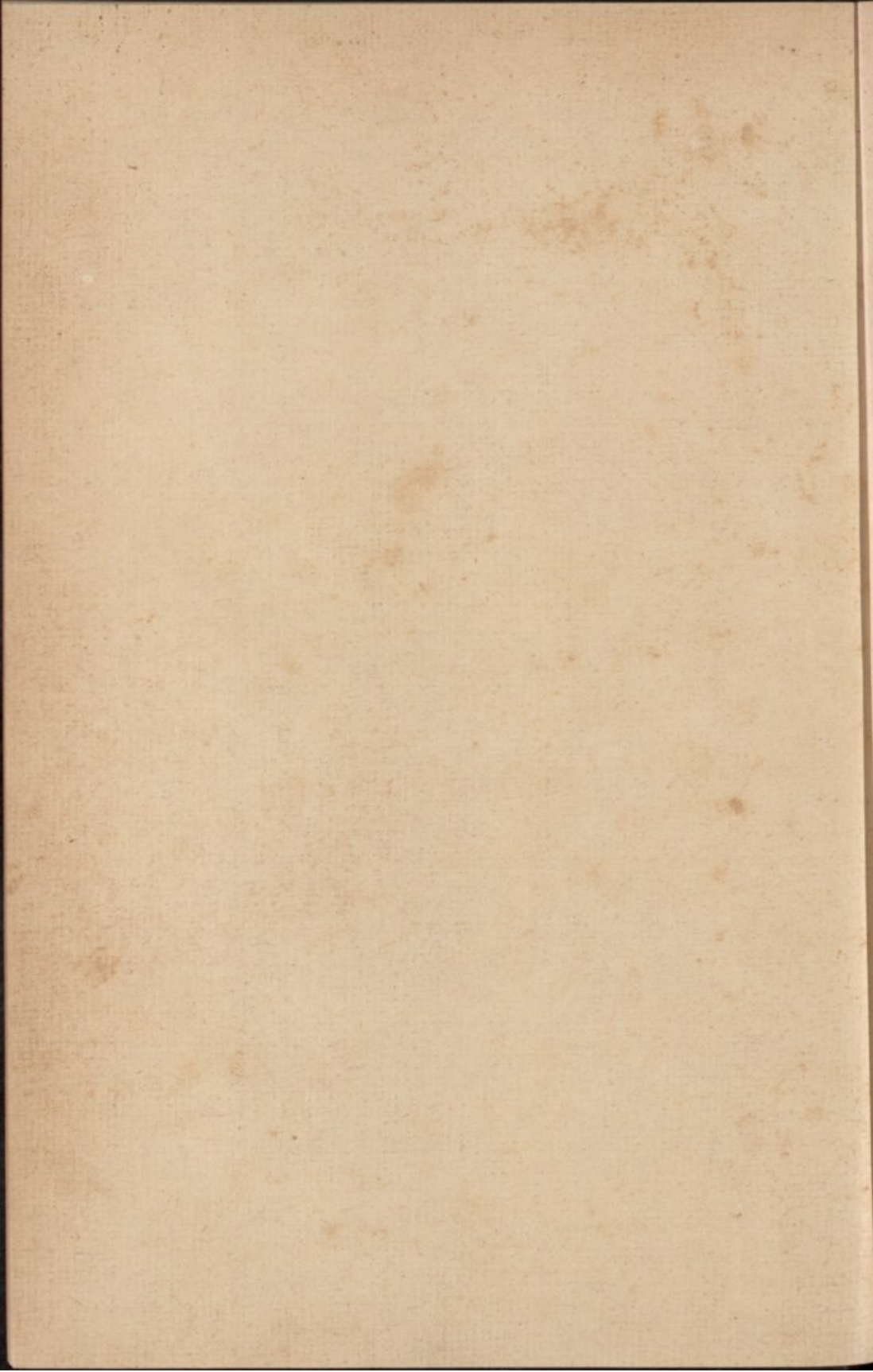
ÍNDICE

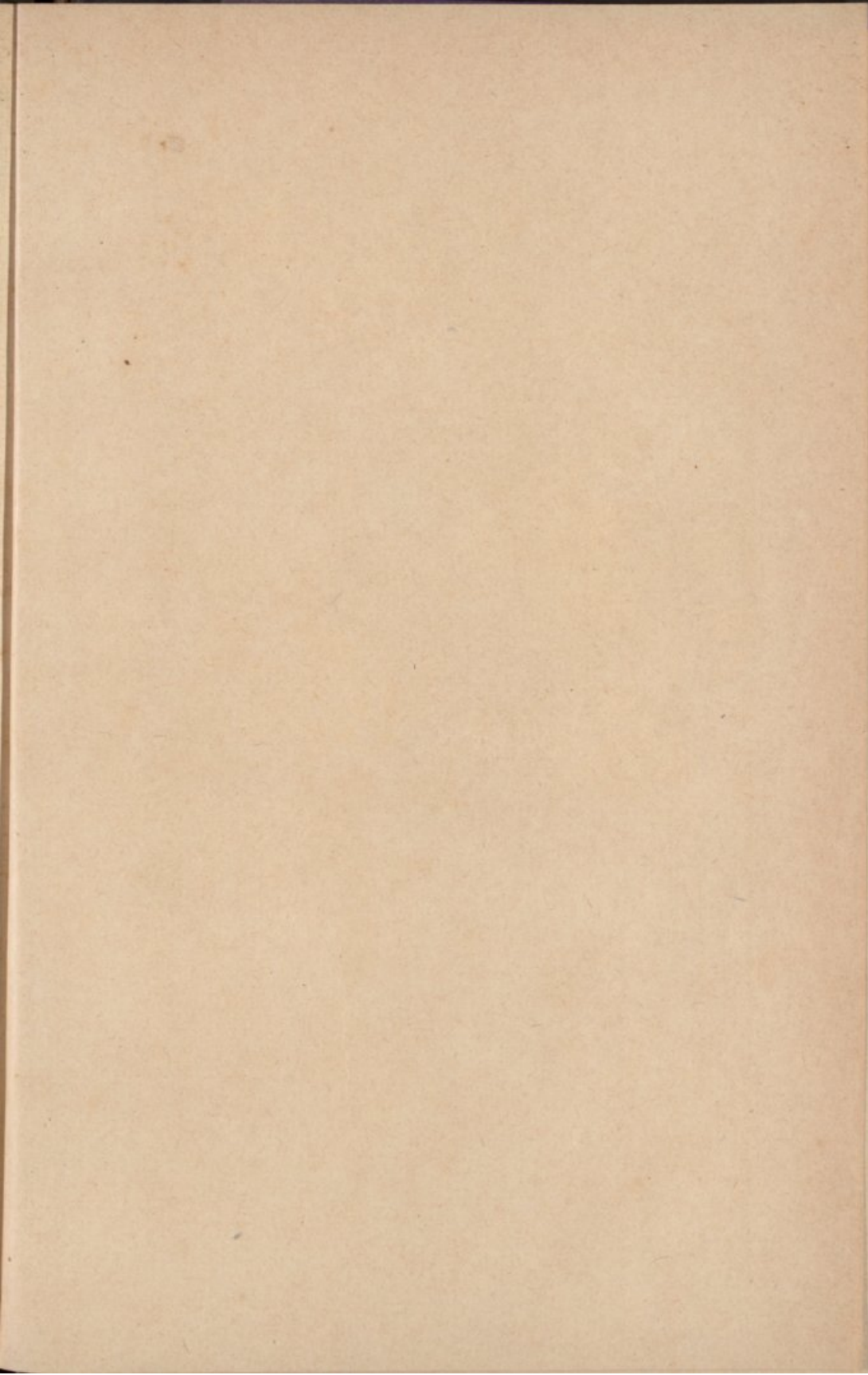
	Págs.
PREFÁCIO	xi
PRELIMINARES	[1]
§ 1.º — Da probabilidade nos conjuntos finitos homogéneos . . .	1
§ 2.º — Da probabilidade nos conjuntos finitos heterogéneos. . .	13
§ 3.º — Da multiplicação das probabilidades	17
§ 4.º — Da probabilidade nos conjuntos numeráveis	27
§ 5.º — Da probabilidade nos conjuntos homogéneos de medida positiva	40
§ 6.º — Da probabilidade nos conjuntos heterogéneos de medida positiva	44
§ 7.º — Das probabilidades irracionais nos conjuntos finitos ou numeráveis	51
§ 8.º — Da composição dos conjuntos mensuráveis	53
§ 9.º — Da multiplicação das probabilidades nos conjuntos men- suráveis.	64
§ 10.º — Das densidades de probabilidade descontínuas	69
§ 11.º — Da probabilidade nos conjuntos de medida infinita. . . .	72
§ 12.º — Da probabilidade nos conjuntos de medida nula	76
§ 13.º — Da esperança das variáveis aleatórias	78
§ 14.º — Dos vectores aleatórios	90
§ 15.º — Dois teoremas-limites sôbre variáveis aleatórias indepen- dentes	92
§ 16.º — Justificação do principio de Gauss no problema da com- binação das observações	103

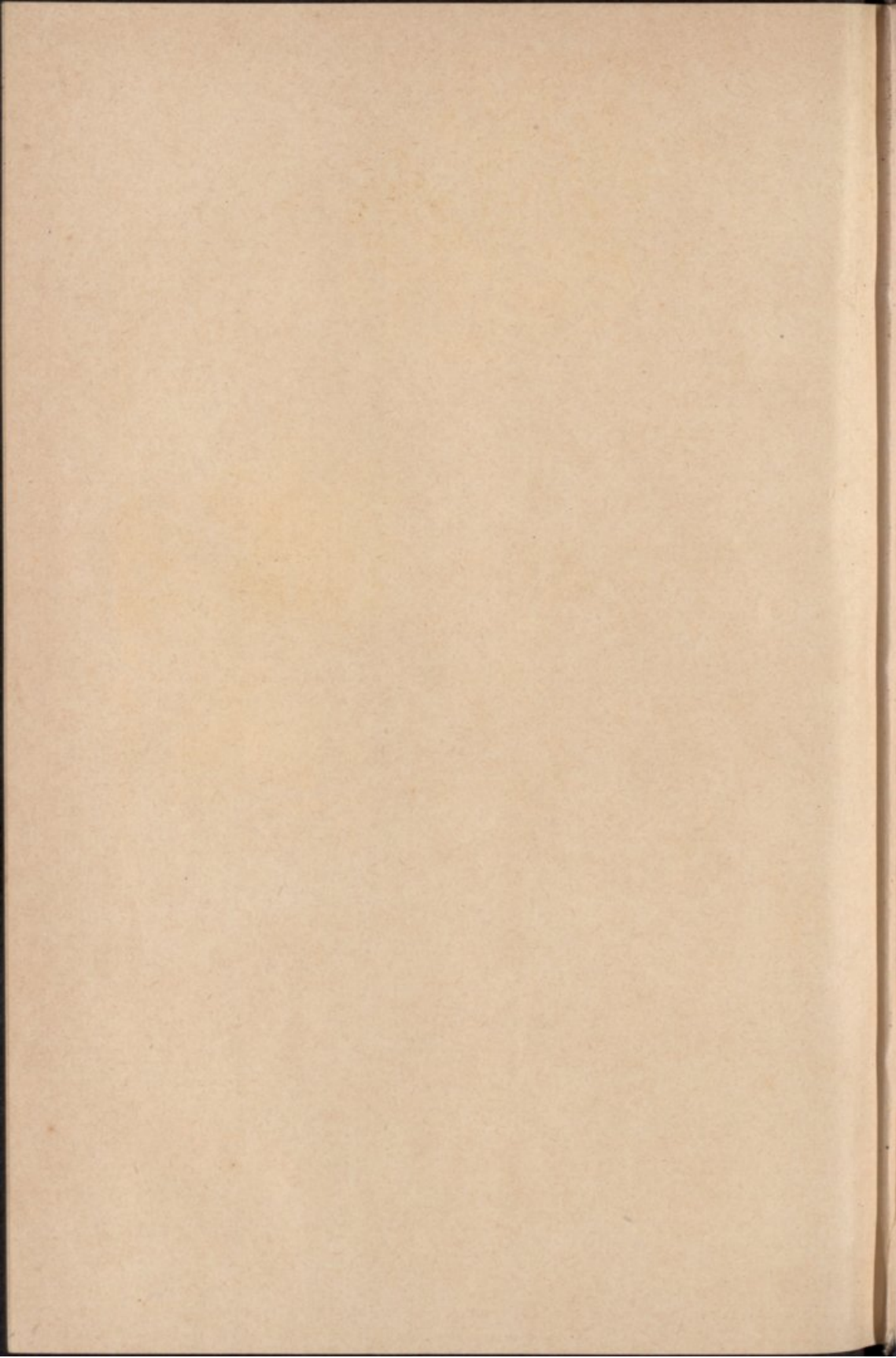


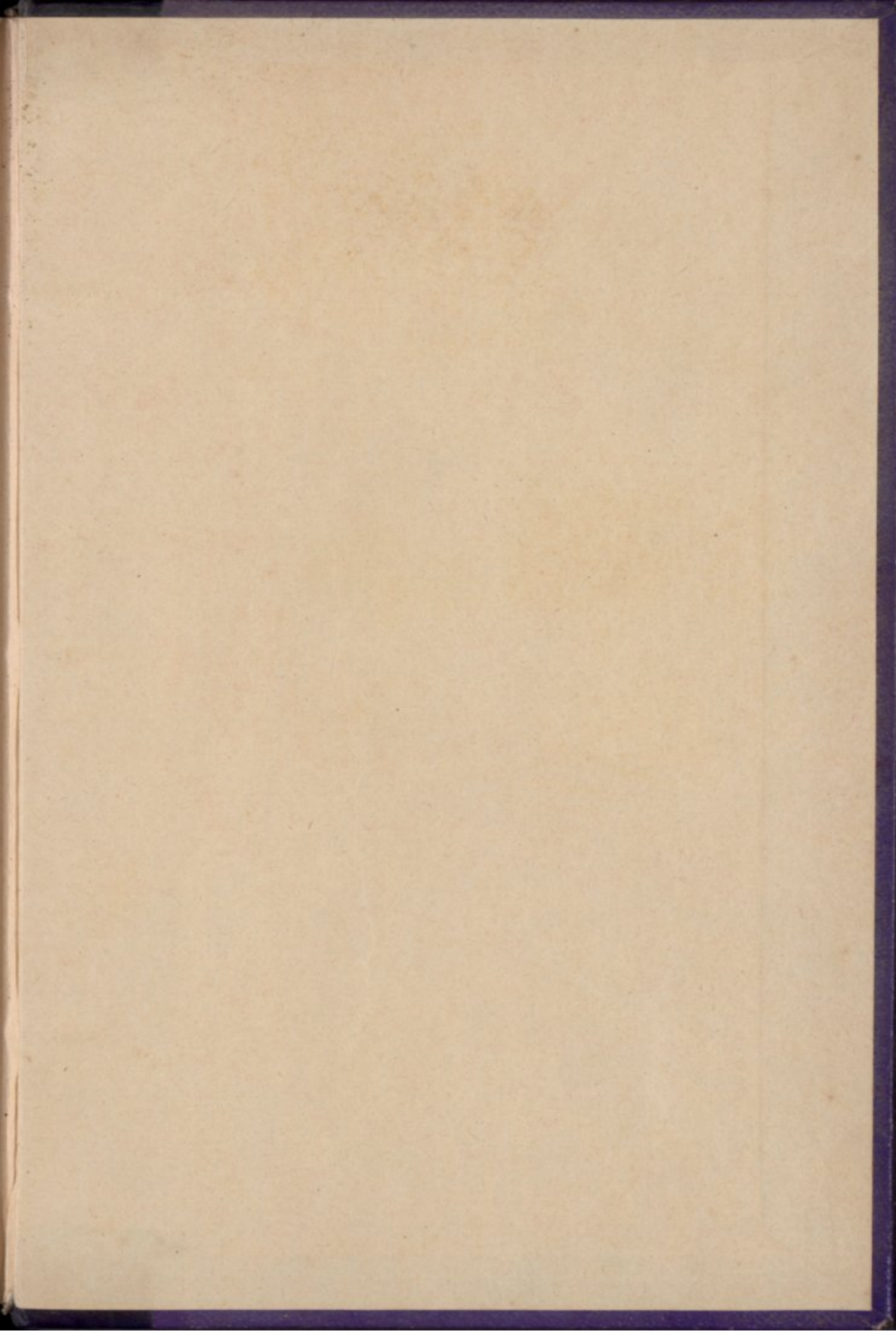


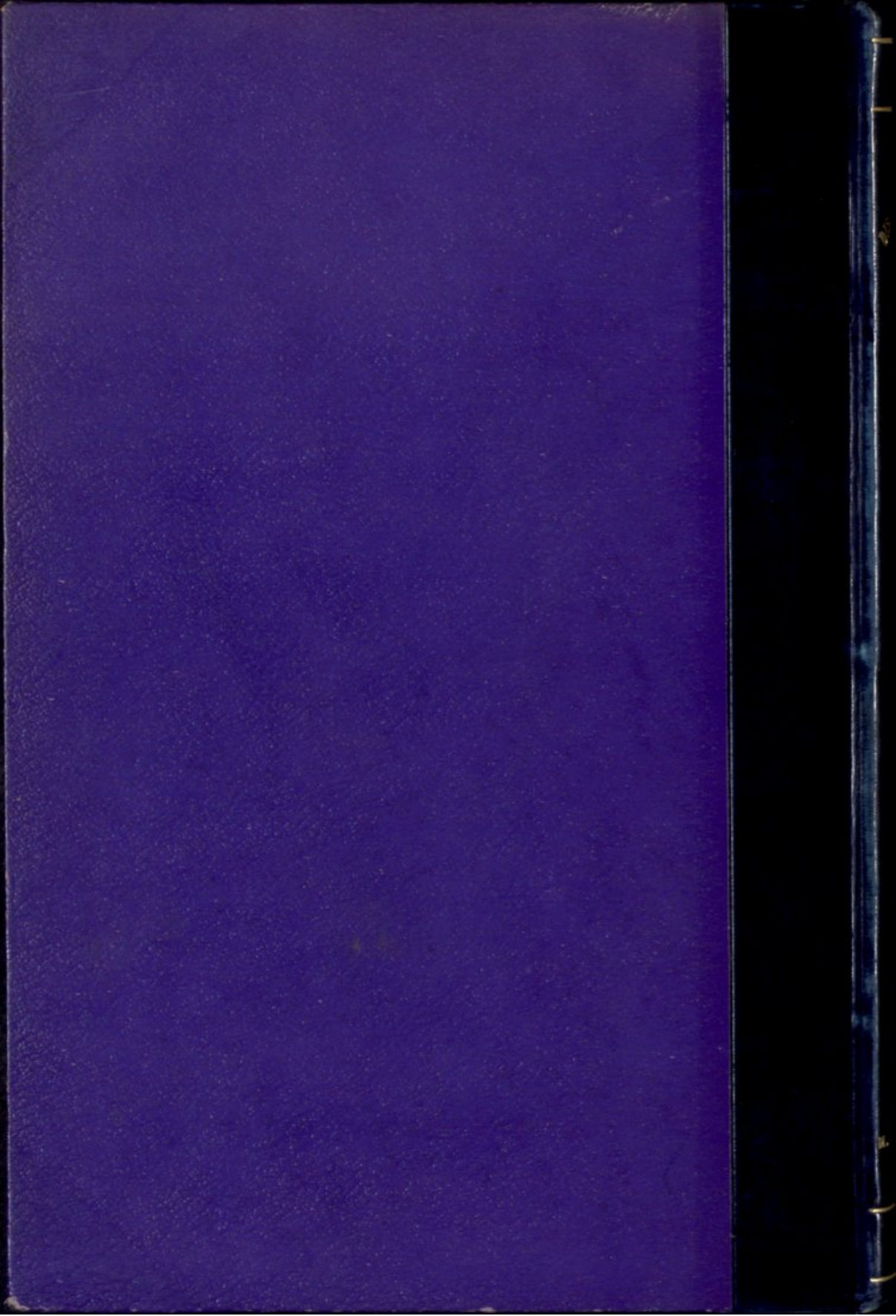












DISPERPAÇÃO

CÁLCULO DAS PROBABILIDADES

M. DOS REIS