
ut videam
ANO
1926 X 1936
DA REVOLUÇÃO
NACIONAL

INTRODUCÇÃO

Á

THEORIA DOS ERROS DAS OBSERVAÇÕES

POR

Sidonio Bernardino Cardoso da Silva Paes

LICENCIADO EM MATHEMATICA

COIMBRA

Imprensa da Universidade

—

1898

DISSERTAÇÃO INAUGURAL

PARA O

ACTO DE CONCLUSÕES MAGNAS

NA

FACULDADE DE MATHEMATICA

DA

UNIVERSIDADE DE COIMBRA

A

MINHA MÃE

E A

MINHA ESPOSA

PREFACIO

Com perto de cem annos de existencia a Theoria dos erros das observações é um assumpto ácerca do qual restam ainda muitas duvidas.

Reduzida ao principio a simples regras empiricas ou aconselhadas pelo bom senso, é um facto digno de notar-se que estas regras têem permanecido quasi sem modificação, porque se é verdade que a sua demonstração tem escapado ás investigações tenazes dos espiritos mais luminosos, tambem não é menos verdade que nunca se conseguiu substituil-as por outras melhores.

Desde Gauss, Lagrange e Laplace nos principios d'este seculo até Bertrand e Poincaré nos ultimos tempos, as questões de que esta Theoria se occupa e cuja alta importancia em diversas sciencias, especialmente na Astronomia, na Geodesia e na Physica, é de sobejo conhecida, têem sido dissecadas pela analyse mais variada, partindo dos pontos de vista mais differentes, em busca de deduzir de um pequeno numero de principios de caracter axiomatico, por um encadeamento logico de raciocinios, essas enigmaticas regras, cuja verdade é mais facil de presentir do que de basear solidamente.

Assumindo a hypothese de que é possivel assignar uma probabilidade numerica a cada grandeza dos erros das observações, a deducção da lei que segue essa probabilidade tem sido uma das maiores difficuldades.

Gauss chegou por uma via indirecta a assignar-lhe uma fórmula, mas elle proprio não depositava, ao que parece, confiança na sua demonstração.

Varios outros caminhos tem sido seguidos para a descoberta d'essa lei; acceitaveis ou não, têem todos conduzido á mesma fórmula — á lei de Gauss —.

Ora emquanto Bertrand no seu *Calcul des probabi-*

lités, apparecido em 1889, ataca esta lei, Poincaré, num livro com o mesmo titulo que veio á luz em 1896, defende-a e toma-a como base do methodo dos menores quadrados, notando todavia que aquella lei e portanto este methodo só são scientificamente legitimos em certas e determinadas circumstancias, que elle tracta de precisar.

«Il ne faudrait pas avoir une sorte de superstition pour la méthode des moindres carrés, à laquelle va nous conduire la loi de Gauss. Nous avons vu que l'on avait parfois des raisons de ne pas adopter cette loi» (*).

Mas se não é permittido crêr cegamente na lei de Gauss, parece exaggerado banil-a, porque admittindo

(*) Poincaré, *Calcul des probabilités,* pag. 196.

que ella não seja rigorosa, a verdadeira lei não será muito differente.

É o que mostra a analyse e a experiencia confirma.

*

Por outro lado as tentativas feitas para estabelecer a Theoria dos erros independentemente da lei da probabilidade do erro não têem talvez sido mais felizes.

A determinação dos valores mais convenientes das incognitas parecia accessivel no caso mais simples das observações directas de egual precisão. Apesar d'essa simplicidade ainda se não conseguiu uma solução que se imponha. A prova d'esta affirmação fornece-a o apparecimento em 1890 de uma memoria de Estienne em que este auctor, influenciado pela leitura da critica de

Bertrand, pretende substituir a media arithmetica pelo valor mediano das observações.

Se no caso simples a que nos referimos fosse possivel determinar satisfactoriamente o valor mais vantajoso da incognita, ter-se-ia ahi a chave de toda a Theoria, procurando reduzir os outros casos mais complicados áquelle de que se tracta.

A esta marcha do particular para o geral seria preferivel, porém, dar uma demonstração geral do principio dos menores quadrados. Fez-se essa demonstração; infelizmente ella implica a restricção de que as equações finaes sejam lineares.

*

Tudo isto prova o que dissemos no principio d'este prefacio. Não se disse a ultima palavra sobre a Theoria dos erros.

Não poderemos nós tambem proferil-a. O nosso fim é sómente fazer a exposição dos principaes trabalhos conhecidos sobre os principios em que se baseia a determinação de incognitas, ligadas a uma serie de observações.

Pretendemos pôr em evidencia o que ha de hypothetico e o que ha de rigoroso e de aproveitavel nestas investigações, de maneira que d'este estudo resalte claramente o estado da questão.

CAPITULO I

Noções preliminares

I. Natureza dos erros das observações. Objecto e fim da Theoria dos erros

1. Quando se observa uma grandeza desconhecida susceptivel de ser definida por um só numero, as differenças para este numero dos resultados das observações, chamam-se *erros verdadeiros* ou simplesmente *erros* das observações.

Estes erros serão numeros qualificados para distinguir os dois sentidos differentes em que podem commetter-se, e é possivel represental-os sobre uma linha recta, como é obvio; d'ahi lhes vem o nome de *lineares*. Erros lineares.

Casos ha, porém, em que a grandeza desconhecida é complexa. Para determinar sem ambiguidade a natureza do desvio, cada observação deverá fornecer mais do que um numero, e haverá em correspondencia erros não addicionaveis.

Isto succede por exemplo quando se procura a posição de um ponto no plano ou no espaço. Os desvios ou erros completos são então grandezas geometricas, para cuja representação não basta a segmentação de uma recta; é necessario fazer intervir, além da grandeza e sentido, um Erros no plano e no espaço.

terceiro elemento a direcção. Chamam-se *erros no plano* e *erros no espaço*.

Trataremos apenas da primeira especie de erros — os lineares —.

2. O observador nunca póde adquirir a certeza de que o resultado da sua medição seja verdadeiro.

A discordancia que geralmente se manifesta entre os resultados das observações dá-nos a prova da existencia de erros em todas as observações differentes menos uma, que poderá estar ou não inquinada; mas ainda que houvesse completo accordo não se poderia d'ahi concluir que o valor observado tivesse coincidido com o valor verdadeiro: esta circumstancia, como faz notar Chauvenet (*), indicaria apenas que se ficou no processo physico muito áquem do limite de precisão (limits of our measuring powers).

Numa tal incerteza reside a differença essencial entre a determinação de uma grandeza pelo calculo ou pela observação.

Erros elementares e erro total.

3. Uma primeira nota que deve fazer-se a respeito dos erros das observações é que elles são devidos quasi sempre a muitas causas differentes. Chamando *erro elementar* ao erro proveniente de uma só origem, o erro da observação será a somma algebrica dos erros elementares, e poderá sob este ponto de vista denominar-se *erro total*.

É costume, tendo em attenção o modo como actuam as diversas especies de causas, dividir os erros em duas classes, conforme são devidos a causas que obram de um modo re-

(*) *Spherical and practical Astronomy*, vol. II, pag. 470, nota.

gular durante a serie das observações, ou ao contrario provêm de origens cuja acção depende de circumstancias que variam de uma observação para outra de uma maneira irregular, podendo considerar-se como acontecimentos independentes.

Aos erros pertencentes á primeira classe dá-se o nome de *regulares* ou *systematicos,* e comprehendem como caso particular os erros *constantes.* **Erros regulares e irregulares.**

Os outros chamam-se *irregulares* e tambem *fortuitos* ou *accidentaes.*

4. Convem não interpretar estas denominações á lettra, porque o caracter verdadeiramente distinctivo d'estas duas especies de erros é o conhecimento ou ignorancia da lei.

É facil de ver por um exemplo classico o que ha de artificial naquella classificação. Os erros resultantes da irregularidade de divisão de um instrumento de medição de angulos deviam considerar-se systematicos e são até constantes, quando o instrumento se emprega para a medição no mesmo logar do limbo. Contam-se, porém, em regra, como fortuitos porque na medição de differentes angulos são empregados logares differentes do limbo.

Ainda um exemplo de outra natureza: na theoria da refracção suppõe-se um certo estado atmospherico que, em geral, será diverso do verdadeiro. Durante um tempo limitado póde esta causa de erro dar logar a erros systematicos, constantes mesmo, mas com a continuação das observações o estado real da atmosphera variará e os erros podem tornar-se irregulares.

Uma mesma origem de erros póde, pois, produzir erros systematicos e accidentaes conforme as circumstancias; portanto a classificação não corresponde a uma distincção precisa entre as causas de erros.

5. Mas considerados os erros em si, isto é, independentemente das causas que os produziram, é importante fixar que os erros irregulares possuem a propriedade de serem extremamente variaveis de observação para observação, podendo considerar-se como acontecimentos casuaes, o que lhes valeu a denominação de fortuitos.

Os erros systematicos desviam os resultados da observação do valor verdadeiro quasi sempre no mesmo sentido, seguindo em cada caso uma lei particular. Não têem em commum alguma coisa essencial que permitta um estudo de conjuncto, e necessitam porisso em cada genero de observações e para cada natureza de causas um tratamento differente.

Objecto da theoria dos erros.

É impossivel abrangel-os numa theoria geral, ao contrario do que succede com os erros accidentaes, que constituem o objecto da Theoria dos erros.

6. Accentuado o caracter dos erros accidentaes — a variabilidade —, convem desde já notar que, relativamente á variação da grandeza, se é certo poder o erro de uma observação theoricamente assumir todos os valores comprehendidos entre certos limites, na pratica os valores possiveis dos erros formam uma serie descontinua, em virtude do limite de observação nitida inherente a cada especie de observações.

Erro inevitavel.

Ha, pois, em cada serie de observações um erro independente das circumstancias em que cada observação foi feita a que se dá o nome de inevitavel. Elle será sempre menor do que uma unidade da ordem mais elevada, a que se attende nas observações.

7. É possivel ainda reconhecer nos erros algumas propriedades importantes.

Resulta, como dissemos, o erro de uma observação de um maior ou menor numero de origens de erros, que podem ser ou não independentes umas das outras, e cuja importancia relativa póde ser susceptivel de muitos graus.

O numero das origens está em proporção com a complexidade do processo empregado para a medição. Esse numero será quasi sempre grande, porque o aperfeiçoamento da arte de observar pelo que respeita á precisão tem trazido comsigo, em geral, a complicação do processo. **Grande numero das causas de erro.**

Por outro lado esse mesmo aperfeiçoamento fez com que as differentes causas de erro tenham uma importancia proximamente egual. É este um facto sabido por todos os que conhecem os instrumentos e methodos de observação modernos, o que torna desnecessario insistir sobre este ponto. **Proxima egualdade dos erros elementares.**

É de presumir assim que cada erro elementar tenha muito pequena influencia no resultado, propriedade muito importante de que para ao deante lançaremos mão.

Por fim notaremos que, se existem origens de erros dependentes umas das outras, é porém certo que são quasi sempre muito mais numerosas as que não são ligadas por nenhuma relação. **Independencia das origens.**

Em resumo, parece que se não deve estar muito longe da verdade considerando o erro total como formado de um numero muito grande de erros elementares, independentes uns dos outros, da mesma ordem de grandeza e tendo cada um pequena importancia em relação áquelle erro.

8. Na epocha em que se observa a theoria ou o methodo de cada especie de observações tem attingido um certo grau de progresso com que somos forçados a contentar-nos. Os meios auxiliares têem imperfeições que hão de tolerar-se, emquanto se não dispõe de outros melhores.

Finalmente o estudo de desaccordo inevitavel entre a maneira theorica de medir e o modo como se executou a medição, póde levar á descoberta de algumas correcções, que devem ser introduzidas nos resultados, mas ha em cada caso um limite, impossivel de ultrapassar.

Alguma cousa, porém, resta de arbitrario. Dentro dos limites mais ou menos estreitos que o tempo impõe sempre, o observador póde repetir as observações, dispõe, por assim dizer, do numero d'ellas.

Fim da Theoria dos erros.

A Theoria dos erros tem precisamente por fim aproveitar esta arbitrariedade de modo a achar os valores mais vantajosos das incognitas relacionadas com as observações.

Eliminação dos erros systematicos

9. Mas antes de applicar esta Theoria a uma serie de resultados de observação suppõe-se que tem sido feita com o maior cuidado a indagação das causas de erro, a determinação dos seus effeitos e a correcção correspondente das observações.

Este estudo previo não revela nunca todos os erros elementares. Não escaparão em geral a um exame demorado os erros constantes ou variando segundo uma lei simples de observação para observação. Ao contrario, é natural suppôr que os erros mais difficeis de determinar sejam aquelles que affectam as observações de uma maneira irregular.

Podemos, pois, admittir, com uma confiança tanto mais proxima da certeza quanto mais minuciosa tiver sido a investigação preliminar, que os erros subsistentes são os que nós chamamos fortuitos.

Este postulado que é necessario collocar logo no principio do nosso estudo não affecta de modo algum o rigor da Theoria, mas póde pôr em duvida a legitimidade da sua applicação, porque não ha meio seguro de verificar se elle tem logar em cada caso.

10. Não seria possivel sobre um desconhecimento completo do erro assentar qualquer estudo serio. Não é tambem essa a nossa situação. Acabamos de ver que se suppõe feita a eliminação do erro systematico e indicamos nos paragraphos anteriores qual é o caracter dos erros accidentaes. Vimos além disso algumas propriedades dos erros elementares e as suas relações com o erro total.

É sobre estas presumpções ou sobre outras mais precisas, mas talvez por isso mesmo menos provaveis, que se eleva a Theoria dos erros.

A natureza dos dados não permitte que se espere uma solução certa, mas apenas uma solução que mereça tanta confiança como as supposições de que tivermos partido. Não se poderá pedir que se determinem os verdadeiros valores das incognitas, mas sómente os valores mais convenientes. Mas será preciso que se diga primeiro o que se entende rigorosamente por tal expressão. É este um ponto delicado de que brevemente nos occuparemos.

II. A lei dos erros. Definição dos valores mais vantajosos das incognitas

11. A identificação de um erro fortuito com um acontecimento casual conduz naturalmente a considerar a Theoria dos erros como um ramo do calculo das probabilidades.

A cada erro corresponde uma certa probabilidade que deverá depender de muitas circumstancias; taes são a grandeza do erro, a individualidade do observador, a natureza do instrumento empregado, a grandeza observada.

Em geral faz-se a supposição de que ella depende só da grandeza do erro. A definição que costuma dar-se envolve ainda, como vae ver-se, outras faltas de rigor, que se permittem para tornar possivel ou, pelo menos, mais facil a analyse.

Definição da lei dos erros.

Consideremos o erro como uma grandeza continua e designemol-o por x. Já vimos atrás que os valores possiveis dos erros formavam na pratica uma serie discreta, separada, porém, por muito pequenos intervallos.

Seja agora $f(x)$ a probabilidade do erro inherente a uma observação estar comprehendido entre o e x. Supponhamos esta funcção analytica e façamos

$$\frac{df(x)}{dx} = \varphi(x).$$

Como $f(x + \Delta x) - f(x)$ é a probabilidade de que o erro de uma observação esteja comprehendido entre x e $x + \Delta x$, poderemos representar a probabilidade de um erro entre os limites x e $x + dx$ por $\varphi(x)dx$.

É esta funcção $\varphi(x)$ que se chama lei de probabilidade dos erros, ou mais simplesmente *lei dos erros.*

12. Fizemos arbitrariamente a supposição de que $f(x)$ fosse uma funcção analytica de x; $\varphi(x)$ será tambem. Ora esta supposição além de ser arbitraria, será, em geral, incompativel com uma outra condição a que em rigor deveria satisfazer a ultima funcção: a de se annullar para todos os valores de x não comprehendidos entre os limites dos erros. Esta condição resulta, com effeito, de que $\varphi(x)$ é a derivada de $f(x)$, e de que esta ultima funcção é evidentemente constante fóra dos limites dos erros. Teremos, pois, de nos contentar, para que desappareça a incompatibilidade com que os valores da lei dos erros além dos limites d'estes sejam muito pequenos.

Convem por outro lado para que a theoria comprehenda todas as especies de observações deixar indeterminados os limites dos erros. Extendem-se esses limites até $-\infty$ e $+\infty$. Ter-se-ha assim a egualdade rigorosa ou approximada

Extensão dos limites dos erros.

$$\int_{-\infty}^{+\infty} \varphi(x)dx = 1;$$

rigorosa se $\varphi(x)$ se annulla fóra dos limites dos erros, por que designando esses limites por a e b será

$$\int_{-\infty}^{+\infty} \varphi(x)dx = \int_{a}^{b} \varphi(x)dx$$

que, representando a probabilidade do erro cahir entre a e b, é a certeza, e portanto egual á unidade; approximada porque não se realisando o annullamento deve pelo menos $\varphi(x)$ ser de muito pequena importancia além dos limites.

Curva de probabilidade do erro.

13. A lei dos erros é susceptivel de uma representação geometrica. Basta para isso tomar para abscissas os erros e para ordenadas as probabilidades correspondentes. A curva determinada pelas extremidades das ordenadas tem por equação

$$y = \varphi(x).$$

Nesta representação a probabilidade de um erro cahir entre dois limites tem uma significação notavel. É a area limitada pelo eixo dos x, as duas ordenadas correspondentes aos limites dados e a curva.

A area total limitada pela curva e o eixo dos x deverá servir de unidade na avaliação das areas, visto representar a certeza.

Póde tambem representar-se a lei dos erros por uma

Representação da lei dos erros por uma recta material.

recta material. Os erros contam-se sobre a recta como anteriormente sobre o eixo dos x. A cada ponto x attribuir-se-ha uma massa egual a $\varphi(x)$. D'esta maneira a probabilidade de que um erro esteja comprehendido entre dois limites será egual á massa do segmento da recta que elles determinam.

14. Partindo da noção da lei dos erros procuraremos agora determinar a probabilidade de qualquer valor da grandeza observada em presença da serie de observações dada. É um problema de probabilidade de causas, que é resolvido pela conhecida fórmula de Bayes.

Consideremos o caso mais simples da Theoria dos erros — observações directas de egual precisão de uma grandeza desconhecida.

A expressão — observações de egual precisão — carece de ser definida porque, embora muito usada por todos os auctores, raras vezes se põe em evidencia o sentido em que ella deve ser tomada.

Diremos que as observações têem egual precisão quando a probabilidade do erro segue a mesma lei para todas as observações.

Seja

$$o_1, o_2, \ldots o_n$$

uma serie de observações nas condições referidas. Designemos a lei dos erros por $\varphi(x)$ ou para mais generalidade por $\varphi(o, v)$, sendo o uma das observações e v o valor verdadeiro da incognita.

Se accrescentarmos ás supposições já feitas a de que os resultados das observações sejam acontecimentos independentes uns dos outros, a probabilidade do concurso de todos esses acontecimentos, suppondo que o valor verdadeiro é v, será pelo theorema das probabilidades compostas, proporcional ao producto

$$\varphi(o_1, v)\varphi(o_2, v) \ldots \varphi(o_n, v) do_1 do_2 \ldots do_n.$$

Seja ainda $\psi(v)dv$ a probabilidade *a priori* para que v esteja comprehendido entre v e $v + dv$.

A probabilidade *a posteriori*, isto é, depois de feitas as observações do mesmo acontecimento póde então representar-se pela fórmula de Bayes, como dissemos. Ter-se-ha, designando por p esta ultima probabilidade

Probabilidade *a posteriori* dos valores da incognita.

$$p = \frac{\varphi(o_1, v)\varphi(o_2, v) \ldots \varphi(o_n, v)\psi(v)dv}{\int_{-\infty}^{+\infty} \varphi(o_1, v)\varphi(o_2, v) \ldots \varphi(o_n, v)\psi(v)dv} \ldots\ldots (1).$$

D'este modo, suppondo conhecidas as duas fórmas φ e ψ, teremos o meio de fazer corresponder a cada valor attribuido a v uma determinada probabilidade.

Poderá tambem a marcha d'esta probabilidade ser dada por uma curva, como para os erros das observações.

Tomando os valores de v para abscissas e os de y para ordenadas, a sua equação será

$$y = \frac{\varphi(o_1, v)\varphi(o_2, v) \ldots \varphi(o_n, v)\psi(v)}{\int_{-\infty}^{+\infty} \varphi(o_1, v)\varphi(o_2, v) \ldots \varphi(o_n, v)\psi(v)dv} \ldots\ldots (2).$$

15. Pergunta-se agora: que valor deve ter p para que o valor correspondente de v seja o mais conveniente? Ou antes: como se deve definir o valor mais conveniente da incognita?

Sobre este ponto não existe accordo.

Primeiro principio de Gauss.

O primeiro principio de Gauss para o valor mais vantajoso era que esse valor fosse aquelle a que correspondesse maior probabilidade.

p devia ser maximo, ou

$$\frac{d[\varphi(o_1, v)\varphi(o_2, v)\ldots\varphi(o_n, v)\psi(v)]}{dv} = o \ldots\ldots\ (3).$$

Principio de Laplace.

Laplace adoptou um principio differente. Para elle quando se procura determinar uma grandeza pela observação cada erro que se commette póde ser comparado a uma perda num jogo de azar. Sendo assim o melhor valor será aquelle que der um total de perdas menor. Cada perda é avaliada pelo valor absoluto do erro multiplicado pela probabilidade respectiva.

Chama *erro total provavel* á somma dos productos analogos obtidos dando ao erro todos os valores possiveis, e escolhe para valor da incognita, aquelle que torna esta somma minima.

Segundo principio de Gauss.

Gauss abandonando o seu primeiro principio adoptou um outro que differe só do de Laplace em substituir o valor absoluto do erro pelo seu quadrado. Esta substituição tem vantagem para a applicação da analyse, porque o quadrado do erro é uma funcção continua do mesmo, emquanto que o valor absoluto não é.

É claro tambem que qualquer potencia par do erro satisfaz, mas o quadrado offerece a vantagem de ser a mais simples funcção continua.

Vê-se que a comparação do erro a uma perda no jogo, fundamento d'estas differentes soluções, deixa subsistir um grande grau de arbitrariedade.

Qualquer que seja, em obediencia a este principio, a solução adoptada será ella preferivel á escolha do valor mais provavel que resulta do primeiro principio de Gauss?

Não poderiamos dar uma melhor resposta a esta pergunta do que transcrever os seguintes periodos do *Calcul des probabilités* de Bertrand:

«La valeur la plus probable est celle dont la probabilité est la plus grande. Peu importent les autres. Elles doivent tontes, cependant, diriger le choix à faire. S'il est utile d'accroître la probabilité des petites erreurs, il est désirable aussi de diminuer celle des grandes. S'attacher seulement à choisir la valeur la plus probable, c'est imiter le joueur qui, pouvant espérer un grand nombre de pertes, prendrait ses décisions de manière à accroître la chance de gagner le plus gros lot, sans aucunement se soucier des autres».

Eis ainda o que diz Poincaré no seu *Calcul des probabilités:*

«La quantité que l'on doit prendre pour z, ce n'est pas la valeur la plus probable, c'est la valeur probable. En effet, la valeur la plus probable est celle qui correspond à la plus grande valeur de p; elle peut être très diffèrente de tontes les autres, tandis que celles-ci peuvent se grouper très près l'une de l'autre, ce qui donne fort à croire qu'elles diffèrent très peu de la vèritable valeur. Elles n'interviement pas dans la valeur la plus probable, tandis qu'elles contribuent tontes à la valeur probable...» [1].

Este valor provavel tem a seguinte definição: Sejam $v_1, v_2, \ldots v_n$ os valores possiveis de uma grandeza; $p_1, p_2, \ldots p_n$ as suas respectivas probabilidades, elle será egual á somma

$$v_1 p_1 + v_2 p_2 + \ldots + v_n p_n.$$

---

[1] z designa o verdadeiro valor da grandeza observada e corresponde a v na nossa notação; p tem a mesma significação que aqui.

Os valores possiveis de v são aqui todos os comprehendidos entre $-\infty$ e $+\infty$, a probabilidade de cada um é dada pela fórmula (1). Logo o valor provavel de v será

$$v_p = \frac{\int_{-\infty}^{+\infty} v\varphi(o_1, v)\varphi(o_2, v)\ldots\varphi(o_n, v)\psi(v)dv}{\int_{-\infty}^{+\infty} \varphi(o_1, v)\varphi(o_2, v)\ldots\varphi(o_n, v)\psi(v)dv} \ldots\ldots (4).$$

É facil de ver que este principio é o mesmo que o segundo de Gauss. Com effeito de (4) tira-se

$$\int_{-\infty}^{+\infty} (v - v_p)\varphi(o_1, v)\varphi(o_2, v)\ldots\varphi(o_n, v)\psi(v)dv = o$$

ou

$$\int_{-\infty}^{+\infty} (v - v_p)ydv = o \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

em virtude de (2).

Ora a equação (5) representa a condição do minimo da expressão

$$\int_{-\infty}^{+\infty} (v - v_p)^2 ydv$$

em relação a v_p.

É precisamente o segundo principio de Gauss que tambem se póde exprimir, em vista d'isto, da seguinte fórma: o valor mais vantajoso é o valor provavel.

A condição analytica do principio de Laplace é differente. Exige-se que seja minima a somma dos valores dos erros multiplicados pelas probabilidades, fazendo abstracção dos signaes, isto é

$$\int_{-\infty}^{+\infty} |v-x|\, y dv = \min.$$

em relação a x, ou ainda

$$\int_{-\infty}^{x} (x-v) y dv + \int_{x}^{+\infty} (v-x) y dv = \min.$$

para o que deve ser

$$\int_{-\infty}^{x} y dv = \int_{x}^{+\infty} y dv \ldots\ldots\ldots\ldots\ldots\ldots (6).$$

16. As condições (3), (5) e (6) são susceptiveis de uma interpretação geometrica notavel.

Interpretação geometrica dos tres principios.

O primeiro principio de Gauss corresponde a adoptar para valor da incognita a abscissa do ponto que tem a ordenada maxima.

O principio de Laplace corresponde, como o mostra (6), a adoptar a abscissa cuja ordenada divide ao meio a superficie total limitada pela curva (2) e pelo eixo dos v.

Finalmente pelo segundo principio de Gauss o valor mais vantajoso será a abscissa do centro de gravidade da mesma superficie; é o que resulta immediatamente da equação (5) que dá

$$v_p = \frac{\int_{-\infty}^{+\infty} y v dv}{\int_{-\infty}^{+\infty} y dv}.$$

17. Pelo que temos dicto vê-se que não é facil, como tinhamos previsto, definir os valores mais vantajosos sob o ponto de vista do calculo das probabilidades.

O primeiro principio deve ser abandonado. Mas qual escolher dos outros dois e da infinidade de principios analogos?

Não temos actualmente nenhuma razão para nos decidirmos, a não ser as vantagens analyticas, que conferem a supremacia ao segundo principio de Gauss.

É elle, por isso, o que hoje se prefere.

Tem agora uma significação precisa o problema da determinação do valor mais vantajoso da incognita: equivale á determinação do valor provavel.

Este ha de ser uma funcção das observações dada pela formula (3), para cuja determinação o caminho naturalmente indicado é o da investigação das funcções φ e ψ. Esta ultima depende do que se sabe *a priori* da probabilidade de v; deve, pois, em rigor ter um valor particular para cada caso. A primeira é a lei dos erros, que é necessario investigar.

No resultado final não se encontra vestigio d'esta lei; é, pois, natural suppôr que pelo menos nalgum caso simples seja possivel chegar á determinação do valor preferivel sem conhecer a lei dos erros.

Se assim fôr, supposta tambem conhecida a funcção ψ, na equação que dá o valor mais vantajoso fica tudo conhecido á excepção da fórma da funcção φ, que poderá talvez determinar-se. Este foi o caminho primitivamente seguido por Gauss. Adoptou como valor mais vantajoso no caso das observações directas de egual precisão a media arithmetica, e identificando-a com o valor mais provavel (primeiro principio) conseguiu deduzir a fórma de φ.

Mas têem sido empregados outros processos para a deducção analytica da lei dos erros; entre elles são sobretudo importantes os que se fundam no estudo dos erros elementares.

Póde tambem procurar-se deduzir em todos os casos os valores mais vantajosos independentemente da lei de probabilidade do erro.

Como ponto de partida de todos estes trabalhos acham-se hypotheses mais ou menos plausiveis e cuja apreciação iremos fazendo *pari passo*.

18. Consideramos até aqui para o estabelecimento da probabilidade de cada valor attribuido á incognita o caso particular das observações directas de egual precisão. É facil de ver como deve modificar-se a fórmula (1) para que possa applicar-se ao caso mais geral da Theoria dos erros. Generalisação.

Sejam

$$v_1,\ v_2, \ldots, v_p$$

p incognitas, que não foram observadas directamente, tendo-se, em vez d'ellas, observado as funcções

$$u_1 = f_1(v_1,\ v_2 \ldots,\ v_p),$$
$$u_2 = f_2(v_1,\ v_2 \ldots,\ v_p),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$u_n = f_n(v_1,\ v_2 \ldots,\ v_p),$$

cujo numero n é maior do que p.

Sejam, além d'isso

$$o_1,\ o_2, \ldots, o_n$$

os resultados da observação das funcções

$$u_1,\ u_2, \ldots u_n.$$

Pede-se a probabilidade *a posteriori* de cada valor attribuido ás incognitas, ou mais rigorosamente as probabilidades *a posteriori* para que v_1 esteja comprehendido entre v_1 e $v_1 + dv_1$, v_2 entre v_2 e $v_2 + dv_2, \ldots$, v_p entre v_p e $v_p + dv_p$.

A priori esta probabilidade será dada por uma funcção que varia de caso para caso, e que nós representaremos por

$$\psi(v_1,\ v_2 \ldots,\ v_p) dv_1 dv_2 \ldots\ dv_p.$$

Suppondo que os verdadeiros valores das incognitas são os attribuidos, a probabilidade do concurso dos valores observados $o_1,\ o_2, \ldots,\ o_n$ poderá ser representada pelo producto

$$\varphi_1(e_1)\varphi_2(e_2) \ldots \varphi_n(e_n) de_1 de_2 \ldots de_n$$

onde

$$\varphi_i(e_i) de_i$$

representa a probabilidade do erro $e_i = u_i - o_i$ da observação o_i. Isto resulta da applicação do theorema das probabilidades compostas, e exige que as observações possam considerar-se acontecimentos independentes uns dos outros.

A fórmula de Bayes dá agora para a probabilidade pedida

$$P = \frac{\varphi_1(e_1)\varphi_2(e_2) \ldots \varphi_n(e_n)\psi(v_1,\ v_2, \ldots,\ v_p) dv_1 dv_2 \ldots dv_p}{\int_{-\infty}^{+\infty} \varphi_1(e_1)\varphi_2(e_2) \ldots \varphi_n(e_n)\psi(v_1,\ v_2, \ldots,\ v_p) dv_1 dv_2 \ldots dv_p} \ldots (7).$$

E para os valores mais vantajosos das incognitas, adoptando o segundo principio de Gauss, teremos a seguinte fórmula que substitue a fórmula (4)

$$\overline{v_i} = \frac{\int_{-\infty}^{+\infty} v_i \varphi_1 \varphi_2 \dots \varphi_n \psi dv_1 dv_2 \dots dv_p}{\int_{-\infty}^{+\infty} \varphi_1 \varphi_2 \dots \varphi_n \psi dv_1 dv_2 \dots dv_p} \dots\dots\dots (8)$$

e que dá o valor provavel da incognita v_i.

CAPITULO II

A lei dos erros fundada no postulado de Gauss

I. Exame do postulado de Gauss

19. Já dissemos no capitulo antecedente que Gauss tomou como base da deducção analytica da lei dos erros a seguinte proposição, que durante muito tempo passou por axiomatica:

Dada uma serie de observações directas egualmente precisas da mesma grandeza desconhecida, a media arithmetica das observações é o valor preferivel.

Isto não é claro e a prova é que são muito numerosas as tentativas feitas para o demonstrar, encontrando-se entre os seus auctores alguns dos maiores geometras. Veremos que a demonstração do postulado de Gauss tem escapado aos estudos mais porfiados, mas que taes trabalhos não têem sido completamente perdidos, servindo para esclarecer a verdadeira natureza da media arithmetica.

20. Seja

$$o_1, o_2, \ldots, o_n \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

uma serie de numeros resultante da observação repetida n

vezes da mesma grandeza desconhecida V, merecendo-nos todas as observações egual confiança.

Á serie (1) corresponde a serie

$$e_1, e_2, \ldots, e_n$$

dos erros de observação, em que

$$\left.\begin{array}{l} e_1 = \mathrm{V} - o_1 \\ e_2 = \mathrm{V} - o_2 \\ \cdots\cdots\cdots\cdots \\ e_n = \mathrm{V} - o_n \end{array}\right\} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2).$$

Se $n = 1$ a serie (1) reduz-se ao primeiro termo o_1 e não ha nesse caso outro partido a tomar senão o de attribuir a V o valor o_1.

Outro caso póde apresentar-se em que não é permitiida hesitação: é o de serem eguaes todos os termos da serie (1), qualquer que seja n devemos adoptar para V o valor commum d'esses termos.

Em todos os mais casos a solução não é intuitiva. Estienne, num trabalho (*) que adeante examinaremos, acha que é muito plausivel fazer esta pergunta:

«De todas as medidas feitas qual é a melhor»?

Mas não será menos plausivel perguntar em seguida se ha um valor melhor do que qualquer das medidas feitas. Reunindo as duas questões o problema a resolver é:

Qual é a combinação das observações que deve adoptar-se para valor da grandeza observada?

(*) *Étude sur les erreurs d'observation*, Paris, 1890.

21. Supponhamos que nada se sabe *a priori* ácerca do valor da incognita, e admittamos além d'isso que os erros positivos e negativos de egual valor absoluto são egualmente provaveis.

Caso particular de duas observações.

Sejam o_1 e o_2 as observações; v o valor da incognita; $\varphi(v-o)$ a lei dos erros, da qual sabemos apenas que deve ser uma funcção par do erro em virtude da hypothese antecedente.

Póde neste caso demonstrar-se que a media arithmetica é o valor mais provavel.

A condição a satisfazer é, segundo o que vimos no capitulo primeiro,

$$\varphi(v-o_1)\varphi(v-o_2)\psi(v)=\max.$$

Mas em vista da ignorancia supposta ácerca do valor de v, a probabilidade *a priori* de que v esteja comprehendido entre v e $v+dv$, que nós representamos por $\psi(v)dv$ reduz-se a dv, porque, devendo todos os valores de v considerar-se egualmente possiveis, aquella probabilidade será independente de v e proporcional ao intervallo dv. Logo

$$\psi(v)=1$$

e a condição póde escrever-se, tomando as derivadas logarithmicas

$$\frac{\varphi'(v-o_1)}{\varphi(v-o_1)}+\frac{\varphi'(v-o_2)}{\varphi(v-o_2)}=0$$

ou

$$\frac{\varphi'(v-o_1)}{\varphi(v-o_1)}-\frac{\varphi'(o_2-v)}{\varphi(o_2-v)}=0$$

por ser φ' uma funcção impar.

Ora a ultima equação é evidentemente satisfeita pela solução

$$v = \frac{o_1 + o_2}{2},$$

como se pretendia demonstrar.

Esta fórma de demonstração não é, porém, applicavel a maior numero de observações, e além d'isso o valor prefeferivel não é o valor mais provavel.

Para o caso considerado de duas observações parece, porém, que o valor preferivel coincide com a media arithmetica, porque, diz-se, o que está mais de harmonia com o facto de serem as observações da mesma precisão e com a hypothese da egual probabilidade de cada grandeza absoluta do erro é suppôr que os dois erros são eguaes e de signaes contrarios, e tomar, portanto, para valor verdadeiro a media arithmetica. Mas deve notar-se que só para n muito grande é que se póde esperar que os erros egualmente provaveis sejam egualmente numerosos, e ainda só approximadamente.

22. Para maior numero de observações é necessario introduzir mais hypotheses. Admittamos as seguintes:

1.ª Os erros positivos e negativos de egual valor absoluto são egualmente provaveis.

2.ª O valor preferivel deve ser uma funcção symetrica das observações.

3.ª A regra para achar este valor ha de ser tal que forneça o mesmo resultado numerico quer se applique ás observações dadas, quer se substituam algumas pelo valor, supposto conhecido, que ellas por si só fariam attribuir á incognita. Generalisação de Encke.

A primeira hypothese conduz, como se viu acima, á adopção da media arithmetica no caso de $n = 2$.

Seja agora $n = 3$ na serie (1) de observações.

Se tivessemos sómente as duas observações o_1 e o_2 o valor mais vantajoso seria $\frac{o_1+o_2}{2}$; se as observações dadas fossem o_1 e o_3, esse valor seria $\frac{o_1+o_3}{2}$; finalmente o valor adoptado seria $\frac{o_2+o_3}{2}$, se as observações fossem o_2 e o_3. Tomando agora cada um d'estes valores e a observação restante, deve, em virtude da terceira hypothese, haver uma combinação f que dê o mesmo resultado que se considerassemos as tres observações o_1, o_2, o_3 dadas.

Ter-se-ha para o valor procurado

$$x=f\left(\frac{o_1+o_2}{2},\ o_3\right)=f\left(\frac{o_1+o_3}{2},\ o_2\right)=f\left(\frac{o_2+o_3}{2},\ o_1\right).$$

Mas pela segunda hypothese x deve ser funcção symetrica de o_1, o_2 e o_3, para o que é preciso que a quantidade isolada esteja ligada com as outras duas como ellas estão ligadas entre si.

Para isso deve ter-se

$$x=\mathrm{F}(o_1+o_2+o_3)\ (*).$$

(*) Póde chegar-se a esta expressão mais claramente, como faz Chauvenet (*Sph. and pract. astr.*, vol. II, pag. 474). Pondo

$$s=o_1+o_2+o_3$$

será

$$\begin{aligned} x&=f\left(\frac{s-o_3}{2},\ o_3\right)=\mathrm{F}(s,\ o_3)\\ &=f\left(\frac{s-o_2}{2},\ o_2\right)=\mathrm{F}(s,\ o_2)\\ &=f\left(\frac{s-o_1}{2},\ o_1\right)=\mathrm{F}(s,\ o_1). \end{aligned}$$

Como s já é symetrico relativamente a o_1, o_2 e o_3 será preciso para que estas egualdades tenham logar que se tenha simplesmente $x=\mathrm{F}(o_1+o_2+o_3)$.

Mas se $o_1 = o_2 = o_3$ será

$$x = \mathrm{F}(3o_1) = o_1.$$

Logo a fórma F representa a divisão por 3 e portanto

$$x = \frac{o_1 + o_2 + o_3}{3}.$$

Empreguemos o processo ordinario de generalisação. Se para a serie (1) se tiver

$$x_n = \frac{o_1 + o_2 + \ldots + o_n}{n}$$

para a serie $o_1, o_2 \ldots, o_n, o_{n+1}$, ter-se-ha, designando por x_{n+1} o valor mais vantajoso,

$$x_{n+1} = \frac{o_1 + o_2 \ldots + o_n + o_{n+1}}{n+1}.$$

Com effeito, considerando o grupo

$$\frac{o_1 + o_2 + \ldots + o_n}{n},\ o_{n+1},$$

será

$$x_{n+1} = \psi\left(\frac{o_1 + o_2 + \ldots + o_n}{n},\ o_{n+1}\right)$$

e por um raciocinio analogo ao que se fez para tres observações

$$x_{n+1} = \theta(o_1 + o_2 + \ldots + o_n + o_{n+1})$$

$$= \frac{o_1 + o_2 + \ldots + o_{n+1}}{n+1}$$

c. s. p. d.

Esta demonstração, devida a Encke, é rigorosa, mas as hypotheses d'onde se partiu não podem acceitar-se todas.

Critica.

A primeira e a segunda são geralmente admittidas. Em observações de egual precisão é natural suppôr que os erros eguaes em valor absoluto sejam egualmente provaveis num sentido ou noutro, visto tractar-se de erros accidentaes. Por outro lado, não havendo razão para preferir uma a outra observação *a priori,* todas ellas devem concorrer egualmente para o resultado, o que equivale a suppôr que este deve ser uma funcção symetrica das observações.

A terceira hypothese, ao contrario, parece completamente arbitraria; não se póde apresentar nenhuma razão séria que obrigue a nossa escolha a regular-se por tal norma.

Demais, na primeira hypothese a mudança que se faz da egual probabilidade em egual frequencia não póde, como já dissemos, admittir-se só para duas observações.

Demonstração de Schiaparelli.

23. Admittam-se agora as tres hypotheses que seguem:

1.ª Qualquer que seja a unidade em que se exprimam as observações o valor preferivel é o mesmo.

2.ª Accrescentando a todas as observações a mesma grandeza arbitraria, o valor preferivel variará tambem d'essa grandeza.

3.ª Fazendo variar uma das observações, a variação correspondente do valor preferivel é independente da observação, que se fez variar.

A primeira hypothese equivale a suppôr que, multiplicando todas as observações pela mesma constante, o valor mais vantajoso venha tambem multiplicado por essa constante, ou que elle seja uma funcção homogenea e linear das observações. O theorema de Euler dá então designando por

$$x = f(o_1, o_2, \ldots, o_n)$$

o valor preferivel

$$\frac{\partial f}{\partial o_1} o_1 + \frac{\partial f}{\partial o_2} o_2 + \ldots + \frac{\partial f}{\partial o_n} o_n = x \ldots\ldots\ldots (3).$$

Addicionando agora a mesma quantidade a a cada uma das observações e attendendo á segunda hypothese, vem

$$x + a = f(o_1 + a, o_2 + a, \ldots o_n + a)$$

$$= f(o_1, o_2 \ldots, o_n) + \left(\frac{\partial f(o_1, o_2, \ldots o_n)}{\partial o_1} + \frac{\partial f(o_1, o_2, \ldots o_n)}{\partial o_2} + \ldots \right) a + \ldots$$

a segunda egualdade resultando do desenvolvimento segundo a fórmula de Taylor; ella deve ter logar qualquer que seja a, logo será

$$\frac{\partial f}{\partial o_1} + \frac{\partial f}{\partial o_2} + \ldots + \frac{\partial f}{\partial o_n} = 1 \ldots\ldots\ldots\ldots (4).$$

Finalmente para exprimir analyticamente a terceira hypothese designemos por $\partial \varepsilon$ a variação muito pequena que se faz experimentar á observação o_1; será, com desprezo das potencias de ε superiores á primeira,

$$\partial x = \frac{\partial f}{\partial o_1} \partial \varepsilon.$$

Soffra agora a mesma variação a observação o_2, ter-se-ha

$$\partial x = \frac{\partial f}{\partial o_2} \partial\varepsilon$$

e semelhantemente para as outras observações; d'onde

$$\frac{\partial f}{\partial o_1} = \frac{\partial f}{\partial o_2} = \ldots = \frac{\partial f}{\partial o_n} \ldots\ldots\ldots\ldots\ldots (5).$$

De (4) e (5) tira-se

$$\frac{\partial f}{\partial o_1} = \frac{\partial f}{\partial o_2} = \ldots = \frac{\partial f}{\partial o_n} = \frac{1}{n};$$

e substituindo estes valores em (3) vem finalmente

$$x = \frac{o_1 + o_2 + \ldots + o_n}{n}.$$

Critica. Deve-se esta elegante demonstração a Schiaparelli; ella tem o defeito de exigir muitas hypotheses, que não são intuitivas.

A primeira e a segunda são propriedades do valor verdadeiro, mas porque escolher estas e não outras? Além d'isso, devendo nós desistir de achar o valor verdadeiro, porque rasão attribuir ao valor mais vantajoso propriedades de que aquelle gosa? Porque assim nos approximamos do valor verdadeiro, não se póde concluir que d'outro modo não nos approximemos mais.

A terceira hypothese póde justificar-se pela rasão de que, sendo as observações de egual precisão, devem actuar da mesma maneira no resultado.

24. Considerando apenas como hypotheses fundamentaes esta ultima e a já admittida por Encke de que no caso de duas observações a media é o valor preferivel, póde tambem demonstrar-se o postulado de Gauss. Demonstração de Stone.

Na serie (1) de observações dê-se a cada uma das observações por sua vez um augmento $\partial\varepsilon$ qualquer. A variação correspondente do valor mais vantajoso $x = f(o_1, o_2 \ldots o_n)$ será a mesma qualquer que seja a observação que variou.

Ter-se-hão, pois, as seguintes egualdades:

$$\begin{aligned}
\partial x &= \frac{\partial f}{\partial o_1}\partial\varepsilon + \frac{1}{2}\frac{\partial^2 f}{\partial o_1^2}(\partial\varepsilon)^2 + \ldots \\
&= \frac{\partial f}{\partial o_2}\partial\varepsilon + \frac{1}{2}\frac{\partial^2 f}{\partial o_2^2}(\partial\varepsilon)^2 + \ldots \\
&= \ldots\ldots\ldots\ldots\ldots\ldots \\
&= \frac{\partial f}{\partial o_n}\partial\varepsilon + \frac{1}{2}\frac{\partial^2 f}{\partial o_n^2}(\partial\varepsilon)^2 + \ldots
\end{aligned}$$

para qualquer valor de $\partial\varepsilon$. Ellas separam-se, porisso, em

$$\frac{\partial f}{\partial o_1} = \frac{\partial f}{\partial o_2} = \ldots = \frac{\partial f}{\partial o_n},$$

$$\frac{\partial^2 f}{\partial o_1^2} = \frac{\partial^2 f}{\partial o_2^2} = \ldots = \frac{\partial^2 f}{\partial o_n^2},$$

d'onde

$$\frac{\partial^2 f}{\partial o_1 \partial o_2} = \frac{\partial^2 f}{\partial o_1 \partial o_3} = \ldots = \frac{\partial^2 f}{\partial o_1^2} = \ldots$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$\frac{\partial^{p+q} f}{\partial o_1^p \partial o_2^q} = \frac{\partial^{p+q} f}{\partial o_1^{p+q}} = \ldots$$

Façamos $o_1 = a_1 + b_1$, $o_2 = a_2 + b_2, \ldots o_n = a_n + b_n$; desenvolvamos em ordem as potencias dos b a funcção x das observações, virá, representando por θ uma quantidade comprehendida entre 0 e 1,

$$x = f(a_1, a_2, \ldots, a_n) + \left(\frac{\partial}{\partial o_1} b_1 + \frac{\partial}{\partial o_2} b_2 + \ldots + \frac{\partial}{\partial o_n} b_n\right) f +$$

$$+ \frac{1}{2}\left(\frac{\partial}{\partial o_1} b_1 + \frac{\partial}{\partial o_2} b_2 + \ldots + \frac{\partial}{\partial o_n} b_n\right)^2 f + \ldots +$$

$$+ \frac{1}{p!}\left(\frac{\partial}{\partial o_1} b_1 + \frac{\partial}{\partial o_2} b_2 + \ldots + \frac{\partial}{\partial o_n} b_n\right)^p f(a_1 + \theta b_1, a_2 + \theta b_2, \ldots, a_n + \theta b_n)$$

que, attendendo ás relações anteriores, se póde escrever

$$x = f(a_1, a_2, \ldots, a_n) + (b_1 + b_2 + \ldots + b_n) \frac{\partial f}{\partial o_1} +$$

$$+ \frac{1}{2}(b_1 + b_2 + \ldots + b_n)^2 \frac{\partial^2 f}{\partial o_1^2} + \ldots$$

$$+ \frac{1}{p!}(b_1 + b_2 + \ldots + b_n)^p \frac{\partial^p}{\partial o_1^p} f(a_1 + \theta b_1, a_2 + \theta b_2, \ldots, a_n + \theta b_n).$$

Das quantidades a_i e b_i uma é arbitraria á nossa escolha. Determinemos os a_i pela condição de serem todos eguaes entre si e á media arithmetica dos o_i será em consequencia

$$b_1 + b_2 + \ldots + b_n = 0$$

e todos os termos do desenvolvimento de x desapparecem á excepção do primeiro. Tem-se

$$x = f(a_1, a_2, \ldots, a_n) = \psi(m) \ldots\ldots\ldots\ldots \quad (6)$$

sendo

$$m=\frac{o_1+o_2+\ldots+o_n}{n}.$$

Temos assim expressa analyticamente a primeira hypothese; vê-se que ella equivale a suppôr o valor mais vantajoso uma funcção da media arithmetica das observações.

No caso de duas observações tem-se

$$\psi(m)=m=\frac{o_1+o_2}{2}.$$

Vejamos o que succede para tres observações. Será

$$x=\psi\left(\frac{o_1+o_2+o_3}{3}\right).$$

Mas

$$\frac{o_1+o_2+o_3}{3}=\frac{2\,\frac{o_1+o_2}{2}+o_3}{3}=\frac{2m+o_3}{3}=m+\frac{o_3-m}{3}.$$

D'onde

$$x=\psi\left(m+\frac{o_3-m}{3}\right)$$

$$=\psi(m)+\frac{o_3-m}{3}\psi'(m)+\frac{1}{2}\left(\frac{o_3-m}{3}\right)^2\psi''(m)+\ldots$$

$$=m+\frac{o_3-m}{3}=\frac{o_1+o_2+o_3}{3},$$

por ser

$$\psi(m)=m,\ \psi'(m)=1,\ \psi''(m)=0,\ \psi'''(m)=0\ldots$$

Generalisando: se para n observações se tiver

$$\psi(m) = m$$

designando aqui m a media das n observações, sobrevindo mais uma observação ter-se-ha

$$\psi\left(\frac{o_1 + o_2 + \ldots + o_n + o_{n+1}}{n+1}\right) = \frac{o_1 + o_2 + \ldots + o_{n+1}}{n+1},$$

como se veria por um calculo perfeitamente analogo ao que se fez para $n = 3$.

Ora tendo logar esta propriedade por hypothese para $n = 2$, terá tambem logar para todo e qualquer valor de n.

É a demonstração de Stone.

Critica.

Ella põe bem evidencia a significação da hypothese terceira de Schiaparelli, a mais plausivel das tres em que se funda a demonstração d'este auctor. Adoptal-a é a mesma coisa que sujeitar o valor preferivel a ser, não a media, mas uma funcção da media.

Para reduzir o valor preferivel á media, Stone admittiu sómente que o postulado já estava demonstrado para duas observações. A demonstração é melhor do que as duas anteriores porque só reclama duas hypotheses em vez de tres e essas são talvez as mais acceitaveis de todas as que as outras demonstrações exigiam.

Outras investigações.

25. Mais simples de que todas estas demonstrações mas tambem menos convincente é o raciocinio feito por Ellis.

Sommemos membro a membro as egualdades (2) do principio do capitulo; vem:

$$\Sigma(V - o_i) = \Sigma e_i$$

d'onde

$$V=\frac{\Sigma o_i}{n}+\frac{\Sigma e_i}{n}\dots\dots\dots\dots\dots\dots (7).$$

Qualquer que seja a combinação das observações adoptada, designando-a por C, e por $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, os erros correspondentes, isto é, fazendo

$$\left.\begin{array}{l} C-o_1=\varepsilon_1 \\ C-o_2=\varepsilon_2 \\ \dots\dots\dots \\ C-o_n=\varepsilon_n \end{array}\right\}\dots\dots\dots\dots\dots\dots (8)$$

vem tambem

$$\Sigma(C-o_i)=\Sigma\varepsilon_i$$

$$C=\frac{\Sigma o_i}{n}+\frac{\Sigma\varepsilon_i}{n}\dots\dots\dots\dots\dots\dots (9).$$

Tomando o limite para $n=\infty$ do valor V dado por (7) obtem-se $\frac{\Sigma o_i}{n}$, que é a media arithmetica das observações, comtanto que não exista nenhuma causa permanente que torne a somma dos erros positivos differente da dos negativos, visto que Σe_i será cada vez provavelmente menor e com certeza sempre finito e n tenderá para ∞.

Mas a egualdade (9) mostra que o mesmo succede a qualquer outra combinação das observações, se comtudo ε fôr uma funcção impar de e, pois que no limite ter-se-ha

$$\frac{\Sigma\varepsilon_i}{n}=\frac{\Sigma f(e_i)}{n}=0.$$

Para a determinação do valor preferivel existem então uma infinidade de regras encerradas na equação

$$\Sigma f(V - o_i) = 0,$$

em que f é uma funcção impar.

Estas regras são todas inexactas quando n é finito e coincidem todas no limite com o valor verdadeiro.

Para escolher d'entre estas regras a da media arithmetica póde apresentar-se como razão a commodidade do calculo. Mas esta razão não colhe porque, como observou Tait, a escolha da media arithmetica entre outras regras é comparavel com o caso de se pretender demonstrar que a attracção varia proporcionalmente á distancia e não na razão inversa do quadrado com o unico fundamento de que o problema dos tres corpos se torna simples e póde resolver-se rigorosamente, ao contrario do que succede com esta ultima lei.

Tambem a nota de Glaisher não resolve a duvida de uma maneira completa. Não se póde affirmar que, sendo os erros muito pequenos, haja praticamente coincidencia entre as equações $\Sigma(V - o_i) = 0$ e $\Sigma f(V - o_i) = 0$. Desenvolvendo $f(e)$ segundo as potencias de e, ter-se-hia

$$f(e) = Ae^2 + Be^4 + Ce^5 + \ldots$$

e despresando os termos da terceira ordem em deante

$$\Sigma f(e_i) = A\Sigma e_i$$

o que tornaria indistinctas as duas equações.

Mas este raciocinio suppõe $A \lessgtr 0$.

Se, porém, $A = 0$ e $B \lessgtr 0$ teremos a equação

$$\Sigma(V - o_i)^3 = 0.$$

Se $A=0,\ B=0,\ C \lessgtr 0$

$$\Sigma(V-o_i)^5=0$$

e assim por deante.

Esta nota reduz em todo o caso a infinidade de fórmas possiveis da funcção f. f será uma potencia de grau impar da variavel.

+ **26.** Os trabalhos mais elucidativos sobre a media arithmetica são com certeza os devidos a De Morgan e Ferrero, que vamos agora expôr conjunctamente, porque elles completam-se um ao outro.

Entre as hypotheses feitas escolheram estes auctores para fundamento das suas investigações as duas seguintes:

1) O valor mais vantajoso da incognita deve coincidir com o valor commum das observações, quando estas são eguaes (vid. n.° 20).

2) O mesmo valor deve ser uma funcção symetrica das observações (vid. n.° 22).

Seja a serie (1) de observações, e $x=f(o_1, o_2, \ldots, o_n)$ o valor preferivel. Attribuam-se a esta funcção f as propriedades da uniformidade e continuidade, dentro dos limites das observações. A primeira equivale a suppôr que, para cada systema de valores $o_1, o_2, \ldots, o_n$ comprehendidos entre aquelles limites, existe um valor determinado que deve considerar-se como o valor mais vantajoso da incognita. Isto suppõe que se conhecem condições de preferencia taes que aquelle valor seja um só. A segunda condição — a continuidade — exige que o valor preferivel varie por graus insensiveis quando qualquer das observações variar tambem da mesma fórma, o que é natural admittir, visto que aquelle valor depende essencialmente dos valores das observações.

Supponhamos que as observações $o_1, o_2, \ldots o_n$ tenham variado de do_1, do_2, do_n, ter-se-ha

$$dx = df = \frac{\partial f}{\partial o_1} do_1 + \frac{\partial f}{\partial o_2} do_2 + \ldots + \frac{\partial f}{\partial o_n} do_n \ldots . \quad (10).$$

Para o caso particular de ser

$$o_1 = o_2 = \ldots = o_n = o_i$$

tem-se

$$f = o_i \text{ (hypothese primeira) e } do_1 = do_2 = \ldots = do_n = do_i.$$

A hypothese segunda dá então

$$\frac{\partial f}{\partial o_1} = \frac{\partial f}{\partial o_2} = \ldots = \frac{\partial f}{\partial o_n} = \mathrm{A},$$

que transforma (10) em

$$df = n\mathrm{A}do_i.$$

Mas $df = do_i$, logo as derivadas parciaes de F tem o valor commum

$$\mathrm{A} = \frac{1}{n}$$

no caso particular da coincidencia de todas as observações.

Fóra d'este caso podemos, pois, representaI-as pelas seguintes fórmulas.

$$\frac{df}{do_1}=\frac{1+\varepsilon_1}{n},\ \frac{df}{do_2}=\frac{1+\varepsilon_2}{n},\ldots,\ \frac{df}{do_n}=\frac{1+\varepsilon_n}{n}$$

suppondo que $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ tendem para zero quando o_1, $o_2, \ldots, o_n$ se approximam umas das outras. E é claro que para um grau sufficiente de concordancia das observações aquellas quantidades positivas ou negativas $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n$ tornar-se-hão menores do que a unidade e portanto as derivadas parciaes serão positivas a partir d'esse momento.

Esta permanencia do signal $+$ para as derivadas é a condição para que f seja uma funcção crescente com as observações. Approximando-se estas da menor, a funcção approxima-se egualmente decrescendo; ao contrario crescendo as observações até attingirem todas o valor da maior, f cresce tambem e tem por limite esse mesmo valor. Logo f acha-se comprehendida entre a menor e a maior das observações, se todavia estas fôrem sensivelmente concordantes.

Chamam-se, porisso, *valores medios* a todas as funcções uniformes e continuas das observações que satisfazem ás hypotheses primeira e segunda, **Valores medios.**

A media arithmetica é evidentemente um valor medio. E é notavel que mesmo quando as observações não são concordantes ella fica sempre comprehendida entre os valores extremos dos resultados de observação.

27. Procuraremos agora marcar o logar da media arithmetica em relação aos outros valores medios. Façamos **Logar da media entre os valores medios.**

$$o_1=a+\delta_1,\ o_2=a+\delta_2,\ \ldots,\ o_n=a+\delta_n.$$

Ter-se-ha

$$f(a+\delta_1,\ a+\delta_2,\dots,a+\delta_n)=f(a,\ a,\dots,a)$$

$$+\ \left(\frac{\partial}{\partial o_1}\delta_1+\frac{\partial}{\partial o_2}\delta_2+\dots+\frac{\partial}{\partial o_n}\delta_n\right)f$$

$$+\frac{1}{2}\left(\frac{\partial}{\partial o_1}\delta_1+\frac{\partial}{\partial o_2}\delta_2+\dots+\frac{\partial}{\partial o_n}\delta_n\right)^2 f+$$

$$+\frac{1}{3!}\left(\frac{\partial}{\partial o_1}\delta_1+\frac{\partial}{\partial o_2}\delta_2+\dots+\frac{\partial}{\partial o_n}\delta^n\right)^3 f+\dots\dots\dots(11).$$

Mas para $o_1=o_2=\dots=o_n=a$, é $f(a,\ a,\dots,a)=a$ e em virtude de f ser symetrica, tem-se

$$\frac{\partial f}{\partial o_1}=\frac{\partial f}{\partial o_2}=\dots=A$$

$$\frac{\partial^2 f}{\partial o_1^2}=\frac{\partial^2 f}{\partial o_2^2}=\dots=B.$$

$$\frac{\partial^2 f}{\partial o_1\partial o_2}=\frac{\partial^2 f}{\partial o_1\partial o_3}=\dots=C$$

$$\frac{\partial^3 f}{\partial o_1^3}=\frac{\partial^3 f}{\partial o_2^3}=\dots=D$$

$$\frac{\partial^3 f}{\partial o_1^2\partial o_2}=\frac{\partial^3 f}{\partial o_1^2\partial o_3}=\dots=E$$

$$\frac{\partial^3 f}{\partial o_1\partial o_2\partial o_3}=\frac{\partial^3 f}{\partial o_1\partial o_2\partial o_4}=\dots=F$$

....

o que transforma a primeira egualdade em

$$f(o_1, o_2, \ldots o_n) = a + A\Sigma\delta + \frac{1}{2}(B\Sigma\delta^2 + C\Sigma\delta_p\delta_q)$$
$$+ \frac{1}{3!}(D\Sigma\delta^3 + E\Sigma\delta_p^2\delta_q + F\Sigma\delta_p\delta_q\delta_r) + \ldots,$$

designando por $p, q, r\ldots$ numeros que devem ter todos os valores inteiros e differentes comprehendidos entre 1 e n.

Notemos agora que reduzindo-se o primeiro membro a $a + \delta$ para $\delta_1 = \delta_2 = \ldots = \delta_n = \delta$, o mesmo deve succeder ao segundo para o que deve ter-se

$$nA = 1,$$
$$B + (n-1)C = 0,$$
$$D + (n-1)E + (n-1)(n-2)F = 0,$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

em virtude de ser n o numero de termos de $\Sigma\delta$, de $\Sigma\delta^2$, $\Sigma\delta^3$, $\ldots$; $n(n-1)$ o de $\Sigma\delta_p\delta_q$, $\Sigma\delta_p^2\delta_q, \ldots$; $n(n-1)(n-2)$ o de $\Sigma\delta_p \delta_q\delta_r$ etc.

D'onde $A = \frac{1}{n}$, e

$$f(o_1, o_2, \ldots, o_n) = \frac{\Sigma o_i}{n} + \frac{1}{2}(B\Sigma\delta^2 + C\Sigma\delta_p\delta_q)$$
$$+ \frac{1}{3!}(D\Sigma\delta^3 + E\Sigma\delta_p^2\delta_q + F\Sigma\delta_p\delta_q\delta_r) +$$
$$+ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (12).$$

Os coefficientes B, C, D, E, F... dependem não só da natureza de f, mas tambem do valor de a e de n, e são desconhecidos.

Supponhamos agora que a é a media arithmetica; neste caso $\delta_1, \delta_2 \ldots \delta_n$ representam desvios entre as observações e a media, os quaes devem ser muito pequenos se as observações fôrem boas, e portanto sensivelmente concordantes.

Desprezemos as potencias dos δ superiores á segunda; (12) reduz-se então a

$$r(o_1, o_2 \ldots o_n) = \frac{\Sigma o}{n} + \frac{1}{2}(\mathrm{B}-\mathrm{C})\Sigma\delta^2 \ldots\ldots\ldots (13)$$

attendendo a que sendo a a media arithmetica

$$\Sigma\delta = o$$

e portanto

$$(\Sigma\delta)^2 = \Sigma\delta^2 + \Sigma\delta_p\delta_q = o.$$

Ora póde ver-se por differentes exemplos que $\frac{n}{2}(\mathrm{B}-\mathrm{C})$ não é crescente com o numero das observações. Assim, attribuindo a f varias fórmas e calculando os valores correspondentes de B e C, acha-se facilmente:

$$\frac{1}{2}(\mathrm{B}-\mathrm{C}) = -\frac{1}{2na} \qquad f = \sqrt[n]{o_1 o_2 \ldots o_n}$$

$$\frac{1}{2}(\mathrm{B}-\mathrm{C}) = \frac{1}{2na} \qquad f = \sqrt{\frac{o_1^2 + o_2^2 + \ldots + o_n^2}{n}}$$

para

$$\frac{1}{2}(\mathrm{B}-\mathrm{C}) = \frac{m-1}{2na} \qquad f = \sqrt[m]{\frac{o_1^m + o_2^m + \ldots + o_n^m}{n}}$$

$$\frac{1}{2}(\mathrm{B}-\mathrm{C}) = \frac{1}{n(n-1)a} \qquad f = \sqrt{\frac{o_1 o_2 + o_1 o_3 + \ldots + o_{n-1} o_n}{\frac{1}{2}n(n-1)}}$$

Admittindo, pois, que assim succeda sempre e escrevendo (13) sob a fórma

$$f(o_1,\ o_2 \ldots o_n) = \frac{\Sigma o_i}{n} + \alpha \frac{\Sigma \delta^2}{n}$$

vê-se que a differença entre a media e qualquer outro valor medio será tanto menor quanto menor fôr $\frac{\Sigma \delta^2}{n}$, isto é, quanto melhores forem as observações e quanto maior fôr o seu numero.

Todos os valores medios differem muito pouco uns dos outros, se as observações são sensivelmente concordantes, e a media arithmetica é o limite commum para que tendem os valores medios quando as observações tendem a concordar.

28. De Morgan foi mais longe nas suas conclusões. Diz elle que mostrando (12) que cada valor medio é egual á media arithmetica mais uma somma de termos, sobre cujo signal e valor absoluto se não póde saber nada *a priori*, não ha razão para suppôr que o valor mais vantajoso seja maior ou menor do que a media arithmetica e que portanto deve adoptar-se esta. Demonstração de De Morgan.

Mas este raciocinio é evidentemente falso. Do facto de que um ponto, por exemplo, deve estar sobre uma recta entre dois pontos dados a e b, não se póde concluir que o logar mais provavel seja o meio.

Se não se sabe mais nada, nada mais se póde prever.

29. Ferrero deu da media arithmetica uma demonstração notavel; mas é claro que não determinando as hypotheses feitas no n.º 26 de uma maneira completa o valor preferivel é necessario accrescentar agora novas hypotheses Demonstração de Ferrero.

para que elle se reduza á media. Ellas são a primeira e a segunda de Schiaparelli (vid. n.º 23).

A egualdade (12) póde ainda escrever-se, para maior simplicidade

$$f(o_1 o_2 \ldots o_n) = \frac{\Sigma o_i}{n} + A_2 + A_3 + \ldots$$

em que $A_n \ldots$ é uma funcção homogenea inteira de ∂_1, ∂_2, $\partial_3, \ldots, \partial_n$, cuja fórma geral é

$$A_n = \alpha \Sigma \partial^n + \beta \Sigma \partial_p^{n-1} \partial_q + \ldots + \xi \Sigma \partial_p^i \partial_q^j \partial_r^k \ldots + \ldots,$$

sendo

$$i + j + k + \ldots = n.$$

Os coefficientes α, $\beta \ldots \xi \ldots$ são funcções de $\frac{\Sigma o_i}{n} = a$; pondo $\xi = \varphi(a)$ e multiplicando todas as observações por uma constante c, a fórma geral de A_n tomar-se-ha em

$$c^n \varphi(ca) \Sigma \partial_p^i \partial_q^j \partial_r^k$$

e em virtude da primeira hypothese do n.º 23 será

$$c^n \varphi(ca) = c \varphi(a)$$

ou

$$c^{n-1} \varphi(ca) = \varphi(a).$$

Para $a = 1$ ξ toma um valor particular que designaremos por ξ_1 e a ultima egualdade dá

$$\varphi(c) = \frac{\xi_1}{c^{n-1}}$$

e mudando c em a

$$\xi = \frac{\xi_1}{a^{n-1}}.$$

Logo

$$\mathrm{A}_n = \frac{1}{a^{n-1}}(\alpha_1 \Sigma \partial_n + \beta_1 \Sigma \partial_p^{n-1} \partial_q + \ldots + \xi_1 \partial_p^i \partial_q^j \partial_r^k + \ldots) = \frac{\mathrm{A}'_n}{a^{n-1}}$$

em que A'_n é uma funcção sómente de $\partial_1, \partial_2, \ldots, \partial_n$. Attribuindo a n os valores 2, 3 ..., ter-se-ha

$$f(o_1, o_2 \ldots o_n) = a + \frac{\mathrm{A}'_2}{a} + \frac{\mathrm{A}'_3}{a^2} + \ldots \quad \ldots \ldots \text{(14)}$$

que exprime a primeira hypothese.

Para introduzir a segunda, junctemos a todas as observações a mesma constante k. f muda-se em $f+k$, a em $a+k$; os ∂_i, porém, permanecem os mesmos por ser $\partial_i = a - o_i = a + k - (o_i + k)$, teremos, pois, a egualdade

$$f(o_1, o_2 \ldots o_n) + k = a + k + \frac{\mathrm{A}'_2}{a+k} + \frac{\mathrm{A}'_3}{(a+k)^2} + \ldots \text{(15)}.$$

Para que sejam satisfeitas as duas egualdades (14) e (15) é necessario que

$$\mathrm{A}'_2 = \mathrm{A}'_3 = \ldots = 0,$$

d'onde

$$f(o_1, o_2 \ldots o_n) = a = \frac{\Sigma o_i}{n}.$$

As duas hypotheses de Schiaparelli aqui usadas de novo, são como já dissemos propriedades do verdadeiro valor e ellas conduzem á annullação das derivadas de ordem superior á primeira.

O valor verdadeiro é uma constante em relação ás observações. Vê-se, pois, que as supposições feitas approximam tambem por este lado o valor preferivel, d'aquelle valor.

30. As investigações que precedem tinham por fim demonstrar que a media arithmetica era o valor mais vantajoso da incognita. As demonstrações dadas não satisfazem, como temos visto; o defeito commum é que as hypotheses em que se baseiam não são mais simples em geral do que o postulado de Gauss.

Propriedades da media arithmetica.

D'estes trabalhos resultam, porém, varias consequencias notaveis, as principaes das quaes vamos aqui resumir:

1.º O valor preferivel deve ser um valor medio.

2.º A media arithmetica das observações é um dos valores medios.

3.º Representa uma hypothese media entre as infinitas hypotheses que se podem fazer sobre o valor preferivel.

4.º Distingue-se além d'isso dos outros valores medios por ser o limite commum de que elles se approximam, quando as observações tendem a concordar.

5.º Approxima-se, por virtude da sua linearidade, do valor verdadeiro, pois que as derivadas da segunda ordem em diante em relação ás observações são nullas.

Justificações da media arithmetica.

31. Têem ainda sido dadas varias justificações da media arithmetica, mas ellas não demonstram a sua superioridade sobre os outros valores medios. Entre ellas não podemos deixar de nomear as de Lagrange, Laplace e Gauss.

Lagrange, partindo da hypothese de que em cada observação se possam commetter com egual probabilidade os erros

$o, +1, -1$, demonstrou que a probabilidade de o erro da media ficar comprehendido dentro de numeros dados augmenta com o numero de observações, approximando-se rapidamente da unidade.

Por outro lado procurando o erro da media para o qual a probabilidade é maior, achou-o egual á somma dos productos dos differentes erros a que uma observação isolada está sujeita multiplicados pelas respectivas probabilidades, isto é, ao valor provavel do erro d'essa observação.

Admittindo que os erros positivos e negativos de egual valor são egualmente provaveis, a curva de probabilidade dos erros será symetrica em relação ao eixo das ordenadas, e o centro de gravidade da superficie da curva terá uma abscissa nulla. Essa abscissa sendo precisamente o valor provavel do erro, segue-se que o erro da media a que corresponde maior probabilidade é o erro zero.

Mas estas considerações não provam que não haja uma outra combinação differente da media para a qual o erro zero não tenha ainda uma maior probabilidade.

Por outro lado Laplace, firmado na mesma hypothese, demonstra que, para um numero *muito grande* de observações egualmente precisas, é mais provavel que a somma dos erros seja egual a zero do que seja differente, e que portanto é mais provavel que a incognita seja egual á media arithmetica do que diffira d'ella.

Isto não prova que não exista uma outra combinação que seja egual á incognita ainda com maior probabilidade; e além d'isso os desprezos feitos no decurso do calculo não permittem applicar esta conclusão senão a um numero muito consideravel de observações.

Finalmente Gauss conseguiu provar que o valor provavel do quadrado do erro da media arithmetica tende para zero á medida que augmenta o numero de observações.

Com effeito seja $\varphi(e)$ a lei dos erros, e supponhamos que se tem

$$\varphi(e)=\varphi(-e)$$

o que equivale a admittir a hypothese de Lagrange e de Laplace.

O quadrado do erro da media arithmetica das observações da serie (1) será $\left[V-\frac{o_1+o_2+\dots+o_n}{n}\right]^2=$

$$\varepsilon^2=\left(\frac{e_1+e_2+\dots+e_n}{n}\right)^2=\frac{\Sigma e_1^2}{n^2}+\frac{2\Sigma e_1 e_2}{n^2}$$

cujo valor provavel é, por definição

$$\overline{\varepsilon^2}=\int \varepsilon^2\varphi(e_1)\varphi(e_2)\dots\varphi(e_n)de_1de_2\dots de_n$$

$$=\frac{\overline{\Sigma e_1^2}}{n^2}+\frac{2\overline{\Sigma e_1 e_2}}{n^2}=\frac{\Sigma\overline{e_1^2}}{n^2}$$

porque $\overline{\Sigma e_i e_{i'}}=0$, em vista de e_i e $e_{i'}$ serem funcções impares e a lei dos erros ser par.

Notando agora que

$$\int e_i^2\varphi(e_1)\varphi(e_2)\dots\varphi(e_n)de_i=\int e_{i'}^2\varphi(e_1)\varphi(e_2)\dots\varphi(e_n)de_{i'}$$

vê-se que

$$\overline{\varepsilon^2}=\frac{n\overline{e_1^2}}{n^2}=\frac{\overline{e_1^2}}{n}$$

o que demonstra a proposição. Quanto maior fôr n, tanto mais provavel será a media

32. Estienne acha que o valor preferivel deve ser dado pela seguinte regra:

Regra de Estienne.

Ordenem-se por ordem de grandeza os resultados das observações; se o seu numero é impar, o do meio é o valor melhor; se o seu numero é par podem adoptar-se indifferentemente os dois do meio ou qualquer valor intermediario. Para se ser inteiramente rigoroso deve accrescentar-se, como Estienne faz tambem, que a regra no caso de n ser impar está enunciada de uma maneira abreviada; poder-se-hia adoptar tambem qualquer valor comprehendido entre $\frac{o_{n-1}}{2}$ e $\frac{o_{n+3}}{2}$.

Para demonstrar a sua regra Estienne admitte que a caracteristica dos erros accidentaes é serem egualmente provaveis num sentido como noutro.

O principal vicio da demonstração está precisamente em partir d'este fundamento, pois que assim põe-se de parte a variação da probabilidade do erro com a sua grandeza, o que todos reconhecem.

Seja a serie (1) que suppomos ordenada segundo a grandeza crescente das observações.

Determinemos a probabilidade *a posteriori* de que o verdadeiro valor esteja comprehendido entre o_i e o_{i+1}. Formemos o numerador da fórmula de Bayes. A probabilidade *a priori* da causa, supposta uma absoluta ignorancia antes de serem feitas as observações, será (n.º 21) proporcional ao intervallo dv; a probabilidade que esta causa dá ao acontecimento referido, isto é, terem-se commettido i erros por defeito e $n-i$ por excesso, é proporcional a

$$\frac{n(n-1)\dots(n-i+1)}{1.2.3\dots i}\left(\frac{1}{2}\right)^n = C_n^i\left(\frac{1}{2}\right)^n$$

pois que os erros accidentaes têem a probabilidade $\frac{1}{2}$ de serem por excesso ou por defeito.

Ora vê-se que o coefficiente binomial C_n^i tem o seu maximo, sendo n par, se o valor v fica comprehendido entre $o_{\frac{n}{2}}$ e $o_{\frac{n}{2}+1}$, e sendo n impar se v fica entre $o_{\frac{n-1}{2}}$ e $o_{\frac{n+3}{2}}$.

Logo o valor mais provavel será o valor mediano dado pela regra enunciada.

Em rigor o raciocinio feito suppõe que o verdadeiro valor não coincide com qualquer dos resultados de observação, no qual caso o numero dos erros commettidos seria menor do que n. Estienne examina este caso e chega ainda á mesma conclusão.

33. Emquanto a media arithmetica é o valor que torna minima a somma dos quadrados dos erros das observações, como é evidente, o valor mediano torna minima a somma dos valores absolutos dos mesmos erros.

É o que Estienne demonstra servindo-se da representação sobre uma linha recta dos resultados das observações.

É o que tambem se póde ver analyticamente d'este modo:

Expressão analytica da regra.

Designando por C a combinação das observações, procuraremos o minimo de

$$|C-o_1|+|C-o_2|+\ldots+|C-o_n| \ldots\ldots\ldots (16)$$

Notemos primeiro que para $C \overline{\gtrless} o_n$ a expressão reduz-se a

$$nC-\Sigma o_i$$

cujo menor valor é

$$no_n-\Sigma o_i. \qquad [C=o_n]$$

E para $C \overline{\lessgtr} o_1$ transforma-se em

$$\Sigma o_i-nC,$$

cujo menor valor é

$$\Sigma o_i - n o_1. \qquad [C = o_1]$$

A partir d'estes valores menores a funcção cresce no primeiro caso até ao ∞ e no segundo até Σo_i.

Consideremos agora os valores de C comprehendidos entre o_1 e o_n

$$o_1 < C < o_n.$$

Designemos por p o numero das parcellas binomias para as quaes C prepondera; será $n-p$ o das restantes.

A expressão (16) poderá escrever-se sem o signal que indica o modulo da seguinte maneira:

$$C - o_1 + C - o_2 + \ldots + C - o_p - C + o_{p+1} - \ldots - C + o_n =$$

$$= C(2p-n) - \sum_{i}^{p} o_i + \sum_{p+1}^{n} o_i \ldots\ldots\ldots\ldots (17)$$

onde C deve satisfazer ás condições

$$o_p \overline{\gtrless} C \overline{\gtrless} o_{p+1}.$$

Supponhamos n par. Se $n = 2p$ (17) muda-se em

$$\sum_{\frac{n}{2}+1}^{n} o_i - \sum_{1}^{\frac{n}{2}} o_i = M.$$

Se $2p > n$ ou $2p = n + 2k$ (17) torna-se em

$$2kC - \sum_{1}^{\frac{n}{2}+k} o_i + \sum_{\frac{n}{2}+k+1}^{n} o_i = 2kC + M - 2\left(o_{\frac{n}{2}+1} + o_{\frac{n}{2}+2} + \dots + o_{\frac{n}{2}+k}\right)$$

$$\geqq 2ko_{\frac{n}{2}+k} + M - 2\left(o_{\frac{n}{2}+1} + \dots + o_{\frac{n}{2}+k}\right)$$

$$\geqq M.$$

Se $2p < n$ é facil de ver por uma analyse perfeitamente semelhante que ainda

$$(17) \geqq M$$

d'onde se conclue que M é o minimo valor da expressão (17), que é attingido dando a C qualquer valor comprehendido entre $o_{\frac{n}{2}}$ e $o_{\frac{n}{2}+1}$.

Analogamente se demonstraria que, se n impar, o minimo valor de (17) é dado ainda pela regra de Estienne.

34. Para estabelecer directamente a excellencia da regra do valor mediano, Estienne prova:

1.º Que a regra da media arithmetica não é plausivel;

2.º Que a regra do valor mediano é preferivel.

1.º Seja a media de n observações julgadas egualmente boas $o_1, o_2, \dots, o_n$

$$M = \frac{o_1 + o_2 + \dots + o_n}{n}.$$

Sobrevem uma outra observação o_{n+1} que não inspira mais duvida do que as precedentes. Applicando a mesma regra, o valor preferivel será agora

$$M' = \frac{nM + o_{n+1}}{n+1} = M + \frac{o_{n+1} - M}{n+1}.$$

A differença entre as duas medias $\frac{o_{n+1} - M}{n+1}$ será, pois, tanto maior quanto mais o_{n+1} differir de M, isto é, quanto peior parece essa observação.

Este raciocinio é vicioso; para adoptar a media M' suppozeram-se primeiro as observações egualmente boas e para a rejeitar diz-se depois que a observação o_{n+1} é peior do que as precedentes.

2.º Para demonstrar que o valor mediano é em geral mais approximado do verdadeiro valor de que a media arithmetica Estienne funda-se em que os erros pequenos são os mais provaveis, e raciocina d'este modo.

Sejam fig. 1 $o_1, o_2, \ldots, o_m, o_{m+1}, \ldots, o_{2m}$ os pontos representativos das observações feitas.

Seja G o centro de gravidade dos m primeiros pontos, G' o centro de gravidade dos m ultimos.

Supponhamos que o verdadeiro valor V cahe á esquerda da região mediana limitada por o_m e o_{m+1}.

Segundo o fundamento adoptado, os pontos devem condensar-se em torno de V, de modo que o centro de gravidade G estará provavelmente mais proximo da região mediana do que o ponto G'. Ora M, ponto representativo da media devendo ficar no meio de GG' ficará portanto á direita da região mediana e em consequencia mais distanciado de V.

Notemos porém que em virtude da hypothese fundamental se é certo que os m primeiros pontos estão em geral mais condensados que os restantes, a distribuição dos primeiros será menos desegual do que a dos segundos. O ponto G ficará proximamente ao meio de AO, emquanto que o ponto G′ deve ficar muito mais proximo de O do que do outro extremo.

Além d'isso dizer que os erros menores são mais provaveis do que os maiores é admittir que a probabilidade do erro depende da sua grandeza. Porque não admittir então que os erros de egual valor absoluto são egualmente provaveis?

A presumpção de que os erros sejam egualmente provaveis num sentido como noutro está ligada a esta de fórma que difficilmente se admittirá uma sem a outra.

Além d'estas objecções póde ainda notar-se que Estienne demonstra que o valor mediano é o valor mais provavel e não o valor provavel, como seria preciso para que se adoptasse.

Em resumo, se o postulado de Gauss não se acha demonstrado, tambem não parece que o valor mediano possa ser preferido á media arithmetica.

35. Bertrand fez ainda áquelle postulado a objecção seguinte:

Nota de Bertrand.

Se medirmos uma certa funcção $f(x)$ e obtivermos os resultados

$$f(o_1),\ f(o_2),\ \ldots f(o_n)$$

será

$$f(x) = \frac{f(o_1) + f(o_2) \ldots + f(o_n)}{n}.$$

Mas dos resultados obtidos tiram-se para x os valores

$$o_1,\ o_2,\ \ldots o_n,$$

que dão

$$x = \frac{o_1 + o_2 + \ldots + o_n}{n}.$$

Logo

$$f\left(\frac{o_1 + o_2 + \ldots + o_n}{n}\right) = \frac{f(o_1) + f(o_2) + \ldots + f(o_n)}{n},$$

o que é em geral falso.

Todavia se os erros são muito pequenos, ter-se-hão as seguintes egualdades approximadas.

$$f(o_1) = f(x) + (o_1 - x)f'(x),$$
$$f(o_2) = f(x) + (o_2 - x)f'(x),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$f(o_n) = f(x) + (o_n - x)f'(x);$$

d'onde, sommando e dividindo por n

$$f(x) = f(x) + \frac{\Sigma(o_1 - x)f'(x)}{n}$$

para o que é preciso que seja

$$\Sigma(o_1 - x) = 0 \quad \text{ou} \quad x = \frac{o_1 + o_2 + \ldots + o_n}{n}.$$

II. Deducção da lei dos erros

36. Gauss, partindo do seu postulado, deduziu a lei dos erros da maneira que segue.

Seja a mesma serie (1) de observações egualmente precisas.

Primitivo processo.

A condição a que deve satisfazer o valor preferivel é, conforme ao primeiro principio de Gauss,

$$\frac{\varphi'(v-o_1)}{\varphi(v-o_1)}+\frac{\varphi'(v-o_2)}{\varphi(v-o_2)}+\ldots+\frac{\varphi'(v-o_n)}{\varphi(v-o_n)}=0\ldots\ldots(1)$$

sendo $\varphi(v-o)$ a lei dos erros e suppondo $\psi(v)=1$.

A esta condição deve satisfazer, portanto, a solução

$$v=\frac{o_1+o_2+\ldots+o_n}{n}.$$

Temos assim duas equações simultaneas

$$\left.\begin{aligned} f(y_1)+f(y_2)+\ldots+f(y_n)&=0\\ y_1+y_2+\ldots+y_n&=0\end{aligned}\right\}\ldots\ldots\ldots\ldots(2)$$

e

na primeira das quaes é $f(y_i)=\dfrac{\varphi'(y_i)}{\varphi(y_i)}$ e em ambas $y_i=v-o_i$.

Do systema (2) deduz-se este outro

$$f'(y_1)dy_1+f'(y_2)dy_2+\ldots+f'(y_n)dy_n=0,$$

$$dy_1+dy_2+\ldots+dy_n=0;$$

d'onde, em geral,

$$f'(y) = \alpha$$

designando por α uma constante.

Integrando, vem

$$f(y) = \alpha y + \beta \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

e integrando de novo

$$l.\varphi(y) = \frac{\alpha}{2} y^2 + \beta y + \gamma.$$

Para determinar as constantes notemos primeiro que, segundo (2) e (3) será

$$f(y_1) + f(y_2) + \ldots + f(y_n) = \Sigma \alpha y + n\beta = 0,$$

$$\Sigma y = 0;$$

logo

$$\beta = 0$$

e

$$l.\varphi(y) = \frac{\alpha}{2} y^2 + \gamma, \ldots\ldots\ldots\ldots\ldots\ldots (4),$$

onde α deve ser negativa, porque $\varphi(y)$ annulla-se para $y = \infty$. Fazendo $\alpha = -h^2$ e $\gamma = l.c$ póde dar-se á equação (4) a fórma

$$\varphi(y) = c\,e^{-h^2 y^2} \ldots\ldots\ldots\ldots\ldots (5).$$

Mas como

$$\int_{-\infty}^{+\infty} e^{-h^2y^2} dy = \frac{\sqrt{\pi}}{h}$$

tem-se

$$\int_{-\infty}^{+\infty} \varphi(y) dy = c\frac{\sqrt{\pi}}{h} = 1$$

como se viu no capitulo anterior.

Da ultima equação tira-se

$$c = \frac{h}{\sqrt{\pi}}$$

o que transforma (5) em

$$\varphi(y) = \frac{h}{\sqrt{\pi}} e^{-h^2y^2}.$$

A constante h é uma caracteristica da precisão das observações. Se as observações não tiverem a mesma precisão, as leis dos erros differirão nos valores de h.

37. Não é rigoroso suppôr φ uma funcção só do erro, como atraz dissemos já. Tambem raras vezes a nossa ignorancia a respeito do valor procurado será tal que permitta considerar $\psi(v) = 1$.

Poincaré procurou libertar-se d'estas hypotheses e bem assim do primeiro principio de Gauss, substituindo-o pelo segundo.

Analyse de Poincaré.

A media arithmetica deve ser o valor provavel e a condição a verificar suppondo, que se obtiveram p resultados eguaes a o_1, p eguaes a $o_2, \ldots, p$ eguaes a o_n, será

$$\frac{\int_{-\infty}^{+\infty} v\psi(v)\Phi^p dv}{\int_{-\infty}^{+\infty} \psi(v)\Phi^p dv} = \frac{po_1 + po_2 + \ldots + po_n}{pn} = \frac{o_1 + o_2 + \ldots + o_n}{n} \ldots (6),$$

onde se fez para simplificar

$$\Phi = \varphi(o_1, v)\,\varphi(o_2, v)\ldots\varphi(o_n, v)\ldots\ldots\ldots(7).$$

Supponhamos que p cresce indefinidamente; o limite do primeiro membro de (6) será (*), designando por v_m o valor

---

(*) Com effeito, considerando a relação

$$\frac{\Sigma v_i\psi(v_i)\Phi_i^p}{\Sigma\psi(v_i)\Phi_i^p} = \mathrm{R},$$

em que Φ_i é o resultado da substituição em Φ de v por v_i, teremos

$$\mathrm{R} = \frac{\Sigma v_i\psi(v_i)\left(\frac{\Phi_i}{\Phi_m}\right)^p}{\Sigma\psi(v_i)\left(\frac{\Phi_i}{\Phi_m}\right)^p}.$$

Ora como por hypothese Φ_m é a maior das quantidades Φ, será

$$\lim_{p=\infty}.\, \mathrm{R} = \frac{v_m\psi(v_m)}{\psi(v_m)} = v_m,$$

porque todas as fracções $\frac{\Phi_i}{\Phi_m}$ são menores do que a unidade á excepção da que corresponde a $i = m$.

de v que torna Φ maximo,

$$\frac{v_m \psi(v_m)}{\psi(v_m)} = v_m$$

e a condição (6) reduz-se a

$$v_m = \frac{o_1 + o_2 + \ldots + o_n}{n},$$

isto é, a media arithmetica deve tornar maxima a funcção Φ dada por (7).

Tomando as derivadas logarithmicas a condição póde então substituir-se por

$$\frac{\varphi'_v(o_1, v)}{\varphi(o_1, v)} + \frac{\varphi'_v(o_2, v)}{\varphi(o_2, v)} + \ldots + \frac{\varphi'_v(o_n, v)}{\varphi(o_n, v)} = 0 \ldots\ldots (8)$$

e

$$o_1 + o_2 + \ldots + o_n = nv$$

ou ainda collocando $\frac{\varphi'_v}{\varphi} = f$, fazendo variar o_1 de do_1, o_2 de $do_2 \ldots$, e conservando v constante

$$\frac{df(o_1, v)}{do_1} do_1 + \frac{df(o_2, v)}{do_2} do_2 + \ldots + \frac{df(o_n, v)}{do_n} do_n = 0.$$

$$do_1 + do_2 + \ldots + do_n = 0.$$

Para o mesmo valor de v, do_1, $do_2, \ldots do_n$ podem ter uma infinidade de valores differentes, logo é preciso que se tenha

$$\frac{df(o_i, v)}{dv_i} = \chi(v) = A',$$

isto é, que cada derivada parcial seja uma funcção sómente de v.

Integrando em ordem a o_i, vem

$$f = A'o_i + B' \dots\dots\dots\dots\dots\dots (9)$$

sendo B′ uma funcção arbitraria de v.

Integrando de novo em relação a v acha-se, attendendo a que $f = \frac{\varphi'_v}{\varphi}$,

$$l.\varphi(o_i, v) = Ao_i + B + l.\theta(o_i),$$

designando por $l.\theta(o_i)$ uma funcção arbitraria de o_i, e sendo $A = \int A'dv$ e $B = \int B'dv$.

Passando de logarithmos para exponenciaes resulta

$$\varphi(o_i, v) = \theta(o_i)\, e^{Ao_i + B} \dots\dots\dots\dots (10).$$

38. Procuremos determinar as funcções arbitrarias. Dando na equação (9) ao indice i os valores $1, 2, 3, \dots, n$, sommando as n equações resultantes e tendo em attenção (8), virá

Determinação das funcções arbitrarias.

$$A'\Sigma o_i + nB' = n(A'v + B') = 0$$

ou

$$A'v + B' = 0 \dots\dots\dots\dots\dots\dots (11),$$

equação que determina uma das funcções arbitrarias A′ ou B′ em funcção da outra.

Vejamos se é possivel achar alguma relação entre A e $\psi(v)$, cuja significação nós conhecemos. Supponhamos p observações eguaes a o_i; como a media arithmetica é o_i, deve ter-se em virtude de (6)

$$\int_{-\infty}^{+\infty} v\psi(v)\,\Phi^p dv = o_i \int \psi(v)\,\Phi^p dv$$

onde Φ é agora egual a $\theta(o_i)e^{Ao_i+B}$; a equação transforma-se assim em

$$\theta^p(o_i)\int(v-o_i)\psi(v)\,e^{p(Ao_i+B)}dv = 0$$

e deve ter logar para qualquer valor de p e de o_i. Como $\theta(o_i)$ não póde ser egual a zero, aliás $\varphi(o_i, v)$ seria sempre nullo, como o mostra (10), é preciso que

$$\int(v-o_i)\,\psi(v)e^{p(Ao_i+B)}\,dv = 0 \ldots\ldots\ldots (12).$$

Temos assim uma equação que liga as duas funcções $A(v)$ e $\psi(v)$.

Introduzamos uma variavel auxiliar u, tal que

$$A(v)o_i + B(v) = u_m^2 - u^2 \ldots\ldots\ldots\ldots (13),$$

sendo u_m^2 o maximo de $Ao_i + B$, que (11) nos mostra ter logar para $v = o_i$. Será, pois,

$$A(o_i)o_i + B(o_i) = u_m^2$$

e u será real para que o seu quadrado seja uma quantidade positiva.

Supponde a equação (13) resolvida em ordem a v, ter-se-hia v expresso em u, e fazendo a mudança de variavel obter-se-hia um resultado da fórma

$$(v - o_i)\psi(v)dv = F(u)du.$$

A equação (12) transforma-se em consequencia na seguinte

$$\int F(u)e^{p(u_m^2 - u^2)}du = 0,$$

ou dividindo pela constante $e^{pu_m^2}$

$$\int F(u)e^{-pu^2}du = 0 \ldots\ldots\ldots\ldots\ldots (14).$$

Mudando u em $-u$, vem

$$\int F(-u)e^{-pu^2}du = 0;$$

e sommando estas duas equações membro a membro

$$\int [F(u) + F(-u)]e^{-pu^2}du = 0 \ldots\ldots\ldots (15).$$

Ha dois casos a considerar:

1.º $F(u)$ é uma funcção impar. Neste caso a condição é evidentemente satisfeita qualquer que seja p.

2.º $F(u)$ é par. Desenvolvendo a funcção $F(u) + F(-u)$ segundo as potencias crescentes de u, ter-se-ha uma expressão da fórma

$$F(u) + F(-u) = \alpha u^{2n} + \beta u^{2n+2} + \ldots$$

e (15) tornar-se-ha em

$$\int (\alpha u^{2n} + \beta u^{2n+2} + \ldots)\, e^{-pu^2}\, du = 0.$$

Fazendo agora a substituição

$$u = \frac{z}{\sqrt{p}}$$

a equação torna-se em

$$\int \left(\frac{\alpha z^{2n}}{p^{n+\frac{1}{2}}} + \frac{\beta z^{2n+2}}{p^{n+\frac{1}{2}}p} + \ldots \right) e^{-z^2}\, dz = 0.$$

Multiplicando por $p^{n+\frac{1}{2}}$, vem

$$\int \left(\alpha z^{2n} + \frac{\beta z^{2n+2}}{p} + \ldots \right) e^{-z^2}\, dz = 0.$$

Ora para p muito grande o primeiro membro toma approximadamente o valor

$$\alpha \int_{-\infty}^{+\infty} z^{2n}\, e^{-z^2}\, dz.$$

Sendo os elementos d'este integral todos positivos, esta quantidade não póde ser nulla; a egualdade não se verifica e por conseguinte a supposição inicial de ser F (u) par é absurda.

Por outro lado differenciando a equação (13) acha-se

$$(A'o_1 + B')\, dv = -2u du,$$

ou tendo em attenção (11) para eliminar B′

$$A'(v - o_i)\,dv = 2u\,du.$$

Dividindo membro a membro a equação

$$(v - o_i)\,\psi(v)\,dv = F(u)\,du$$

pela precedente acha-se a relação

$$\frac{\psi(v)}{A'} = \frac{F(u)}{2u}.$$

A propriedade descoberta para $F(u)$ da imparidade vae-nos servir agora para mostrar que esta relação é constante.

Com effeito, se assim fôr, deve ter-se

$$\frac{\psi(v_1)}{A'(v_1)} = \frac{\psi(v_2)}{A'(v_2)}$$

quaesquer que sejam v_1 e v_2. Para estes valores as funcções $A(v)$ e $B(v)$ tomam valores particulares que designaremos respectivamente por A_1 e B_1, A_2 e B_2. A equação

$$A_1 o_i + B_1 = A_2 o_i + B_2$$

resolvida em ordem a o_i determina esta quantidade de modo que a funcção $u_i^2 - u^2$ e portanto u^2 passe pelo mesmo valor para v_1 e v_2. O mesmo succederá a $\frac{F(u)}{2u}$ que, sendo uma funcção par em virtude da egual paridade dos dois termos da fracção, se póde considerar como funcção de u^2 sómente.

Na fórmula (10) póde agora substituir-se A' por $\psi(v)$; ter-se-ha, por ser $B = -\int A'vdv + \text{const.}$,

$$\varphi(o_i, v) = \theta(o_i)\, e^{A o_i + B}$$

$$= \theta(o_i)\, e^{-\int A'(v - o_i)\, dv}$$

$$= \theta(o_i)\, e^{-\int \psi(v)(v - o_i)\, dv} \quad \text{........(16).}$$

39. Restam as duas funcções arbitrarias $\psi(v)$ e $\theta(o_i)$. Gauss tomou, como vimos, $\psi(v) = 1$. Isto suppõe que nada se sabe *a priori* da probabilidade de v, o que não succederá sempre.

A fórmula (16) mostra que φ depende de ψ. Ora, nota Poincaré, não ha razão nenhuma para esta dependencia; porque emquanto ψ depende dos nossos conhecimentos *a priori* sobre a probabilidade de z, φ depende da habilidade do observador e da probabilidade *a priori* para que elle erre.

Ainda quando se descobrisse esta razão, ha uma circumstancia que limita muito a influencia de ψ e permitte suppol-a egual á unidade. Seja qual fôr a lei dos erros, ella deve satisfazer como vimos no capitulo primeiro á condição de decrescer muito rapidamente quando o erro augmenta, e de ter valores despreziveis fóra dos limites dos erros. Estes limites são muito estreitos, se as observações são boas.

Os valores possiveis da grandeza observada comprehendidos entre os resultados das observações serão, pois, em geral, *a priori* considerados como egualmente provaveis, por nada se saber sobre elles. Ora $\psi(v)$ entra no valor provavel e tambem no mais provavel multiplicada pelo producto $\varphi(o_1, v)\, \varphi(o_2, v) \ldots \varphi(o_n, v)$. Como este producto é fóra d'aquelles limites infinitamente pequeno, pouco será alterado pela introducção do factor finito $\psi(v)$, que é precisamente n'essas regiões que se poderá tomar differente da unidade.

Relativamente a $\theta(o_i)$, vê-se que será constante quando se suppozer com Gauss que a lei dos erros depende sómente do erro. Mas em rigor muitas vezes poderá prever-se que $\theta(o_i)$ seja variavel com os valores das observações, e mesmo descobrir o modo d'essa variação. Quando se commette por exemplo o *erro decimal*, $\theta(o_i)$ é uma funcção periodica que se torna maxima para certas decimaes.

Supppondo $\psi(v)$ e $\theta(o_i)$ constantes recáe-se na lei de Gauss.

40. A hypothese da media arithmetica conduz, quando se suppõe a lei dos erros funcção sómente do erro, á lei de Gauss. Reciprocamente a lei de Gauss conduz á media arithmetica como valor preferivel. A lei de Gauss e a media arithmetica.

Basta para o demonstrar formar partindo d'aquella lei o valor mais provavel ou o valor provavel. Em qualquer dos casos como vimos deve ser maximo o producto

$$\varphi(v-o_1)\,\varphi(v-o_2)\ldots\varphi(v-o_n)$$

$$=\left(\frac{h}{\sqrt{\pi}}\right)^n e^{-h^2\Sigma(v-o^i)^2}$$

em relação a v, o que dá

$$\Sigma(v-o_i)^2=\text{min.}$$

ou

$$v=\frac{\Sigma o_i}{n}.$$

Se a lei dos erros tem a fórma mais geral (16) tambem é facil verificar que a media arithmetica é o valor provavel.

Com effeito, se assim fôr deve ter-se, em virtude da definição de valor provavel, designando por M a media

$$\int_{-\infty}^{+\infty}(v-\mathrm{M})\,\psi(v)\,\varphi(o_1, v)\,\varphi(o_2, v)\ldots\varphi(o_n, v)\,dv=0$$

ou

$$\int_{-\infty}^{+\infty}(v-\mathrm{M})\psi(v)\theta(o_1)\theta(o_2)\ldots\theta(o_n)e^{-n\int(v-\mathrm{M})\psi(v)dv}\,dv=0\ldots(17).$$

Podemos fazer

$$\int(v-\mathrm{M})\,\psi(v)\,dv=u^2$$

visto que o expoente de e é negativo.

Da ultima egualdade tira-se

$$(v-\mathrm{M})\,\psi(v)\,dv=2udu$$

e portanto a equação (17) transforma-se em

$$2\theta(o_1)\,\theta(o_2)\ldots\theta(o_n)\int_{-\infty}^{+\infty}ue^{-nu^2}du=0$$

que é uma identidade por ser ue^{-nu^2} uma funcção impar de u.

41. Póde demonstrar-se que a lei de Gauss é a unica para a qual a media é sempre o valor mais provavel, comtanto que não haja erros systematicos.

Esta restricção equivale a suppôr que a lei dos erros é par, ou

$$\varphi(y) = \varphi(-y),$$

porque, devendo considerar-se egualmente provaveis n'um sentido ou n'outro os erros eguaes em valor absoluto, $\varphi(y)$ não deve alterar-se pela mudança de signal da variavel.

$\varphi(y)$ reduzida á fórma de exponencial deve poder representar-se por

$$e^{-\psi(y^2)}$$

em virtude da condição referida.

O valor mais provavel satisfaz á equação (3) do capitulo primeiro, que se transforma aqui em

$$\frac{d}{dv} e^{-[\psi(v-o_1)^2 + \psi(v-o_2)^2 + \ldots + \psi(v-o_n)^2]} = 0$$

ou

$$(v-o_1)\,\psi'(v-o_1)^2 + (v-o_2)\,\psi'(v-o_2)^2 + \ldots + (v-o_n)\psi'(v-o_n)^2 = 0.$$

Para que v coincida com a media deverá ser

$$v-o_1 + v-o_2 + \ldots + v-o_n = 0.$$

Das duas ultimas equações tira-se

$$\psi'(v-o_i)^2 = \psi'(y^2) = \text{const.}$$

Integrando:

$$\psi(y^2) = c_1 y^2 + c_2$$

c_1 e c_2 sendo constantes.

Logo

$$\varphi(y)=e^{-\psi(y^2)}=Ce^{-c_1y^2}$$

onde

$$l.\,C=-c_2.$$

Fica assim demonstrado que na hypothese feita a lei de Gauss é a unica que indica a media como valor mais provavel.

42. A deducção da lei de Gauss assenta n'uma proposição que se não sabe demonstrar em geral. Mas ha um caso em que, sendo o postulado de Gauss demonstravel rigorosamente, todavia a lei de Gauss apparece apenas como uma formula approximada.

Bertrand poz este facto em evidencia da seguinte maneira: Seja a grandeza que se pretende medir a proporção constante do numero de espheras brancas para o numero total de espheras contidas n'uma urna de composição desconhecida. N'uma primeira prova tiraram-se uma a uma μ espheras, tornando-as a metter na urna á medida que se foram tirando. Sejam d'estas b brancas; a relação $\frac{b_1}{\mu}$ dá uma medida da incognita. Effectuada a mesma operação n vezes, obtiveram-se os numeros

$$\frac{b_1}{\mu},\frac{b_2}{\mu},\ldots,\frac{b_n}{\mu},$$

resultados de observações egualmente precisas da incognita x. O valor mais provavel de x é o que torna maxima a probabilidade *a priori* do acontecimento observado. Ora x representa evidentemente a probabilidade assignada á sahida d'uma bola branca pela composição da urna. Logo a probabilidade *a priori* do resultado observado será

$$x^{b_1+b_2+\ldots+b_n}(1-x)^{n\mu-b_1-b_2-\ldots-b_n};$$

porque $1-x$ representa a probabilidade da tiragem d'uma bola não branca, e na primeira prova tiraram-se $\mu-b_1$, na segunda $\mu-b_2$, ... d'aquella especie.

O maximo é dado pela egualdade a zero da derivada em relação a x, ou

$$\frac{x}{b_1+b_2+\ldots+b_n}=\frac{1-x}{n\mu-b_1-b_2-\ldots-b_n},$$

d'onde

$$x=\frac{b_1+b_2+\ldots+b_n}{n\mu}$$

que é a media arithmetica das observações.

Procuremos agora a lei dos erros. É, como se sabe, o coefficiente differencial da probabilidade do erro $x-\frac{b}{\mu}$ caír entre y e $y+dy$.

A fórmula de Bayes dá

$$\varphi(y)dy=\frac{\left(\frac{b}{\mu}+y\right)^b\left(\frac{\mu-b}{\mu}-y\right)^{\mu-b}dy}{\int_{-\frac{b}{\mu}}^{\frac{\mu-b}{\mu}}\left(\frac{b}{\mu}+y\right)^b\left(\frac{\mu-b}{\mu}-y\right)^{\mu-b}dy}.$$

Fazendo para simplificar $\frac{b}{\mu}=p$, $\frac{\mu-b}{\mu}=q$, $\mu-b=c$, vem

$$\varphi(y)dy=\frac{\left(1+\frac{y}{p}\right)^b\left(1-\frac{y}{q}\right)^c dy}{\int_{-p}^{q}\left(1+\frac{y}{p}\right)^b\left(1-\frac{y}{p}\right)^c dy}.$$

Desenvolvamos cada um dos factores do numerador segundo a formula do binomio de Newton e effectuemos o producto:

$$\left(1+\frac{y}{p}\right)^b\left(1-\frac{y}{q}\right)^c=1+\left(\frac{b}{p}-\frac{c}{q}\right)y$$

$$+\left(\frac{b(b-1)}{p^2}-\frac{2bc}{pq}+\frac{c(c-1)}{q^2}\right)\frac{y^2}{2}+\ldots$$

$$=1-\frac{\mu}{2pq}y^2=e^{-\frac{\mu}{2pq}y^2}$$

desprezando os termos de ordem mais elevada em relação a y.

Teremos então

$$\int_{-p}^{q}\left(1+\frac{y}{p}\right)^b\left(1-\frac{y}{q}\right)^c dy=\int_{-p}^{q}e^{-\frac{\mu}{2pq}y^2}dy$$

que, fazendo a mudança de variavel

$$y=t\sqrt{\frac{2pq}{\mu}}$$

e collocando

$$\sqrt{\frac{\mu p}{2q}} = l,$$

se muda em

$$\sqrt{\frac{2pq}{\mu}} \int_{-l}^{+l} e^{-t^2}\, dt.$$

Para valores grandes de μ e de p, podemos substituir $-l$ e $+l$ por $-\infty$ e $+\infty$ e o integral reduz-se a

$$\sqrt{\frac{2\pi pq}{\mu}}.$$

Com estes valores approximados, ter-se-ha

$$\varphi(y)dy = \sqrt{\frac{\mu^3}{2\pi b(\mu - b)}}\, e^{\frac{-\mu^3 y^2}{2b(\mu - b)}} dy$$

$$= \frac{h}{\sqrt{\pi}} e^{-h^2 y^2}$$

fazendo $h = \sqrt{\frac{\mu^3}{2b(\mu - b)}}$.

Chegamos, pois, á lei de Gauss, mas como uma expressão approximada da lei de probabilidade do erro, como uma lei limite, e não como uma lei rigorosa dos erros.

Esta contradicção explica-se, notando que, se é verdade que na deducção da lei de Gauss do seu postulado se não fez desprezo algum, a verdade é tambem que attendendo á fórma da demonstração, a lei a que se chega não é propriamente a lei dos erros das observações, mas sim a lei dos desvios da media arithmetica.

43. Póde dar-se uma justificação da lei de Gauss, como lei approximada, provando que a probabilidade do erro da media segue esta lei, se não houver erros systematicos e se o numero n das observações fôr muito grande.

O erro da media é egual á media dos erros das observações ou

$$\varepsilon = \frac{e_1 + e_2 + \ldots + e_n}{n}.$$

O seu valor provavel $\bar{\varepsilon}$ é

$$\bar{\varepsilon} = \int_{-\infty}^{+\infty} \varphi(e_1)\,\varphi(e_2)\ldots\varphi(e_n)\,\varepsilon\, de_1\, de_2 \ldots de_n$$

que é nullo visto ser φ uma funcção par e ε impar.

O erro da media segue a lei de Gauss.

Qualquer potencia impar de ε terá egualmente um valor provavel nullo.

Consideremos o valor provavel de ε^{2p}; tem-se

$$\overline{\varepsilon^{2p}} = \frac{\overline{(e_1 + e_2 + \ldots + e_n)^{2p}}}{n^{2p}}.$$

Mas

$$(e_1 + e_2 + \ldots + e_n)^{2p} = \sum \frac{2p!}{\alpha_1!\,\alpha_2!\ldots\alpha_\mu!}\, e_1^{\alpha_1}\, e_2^{\alpha_2} \ldots e_\mu^{\alpha_\mu}$$

sendo

$$\alpha_1+\alpha_2+\ldots+\alpha_\mu=2p,$$

e não sendo preciso considerar para a formação do valor provavel senão os termos em que os expoentes $\alpha_1, \alpha_2, \ldots \alpha_\mu$, são todos pares, porque os restantes têem um valor provavel nullo.

O valor provavel de

$$e_1^{\alpha_1} e_2^{\alpha_2} \ldots e_\mu^{\alpha_\mu}$$

fica invariavel quando se permutam os indices dos e e se conservam os expoentes, porque o integral é o mesmo só com mudança de notações.

Seja N o numero de termos differentes assim obtidos, M_{α_i} o valor provavel de $e_1^{\alpha_i}$; o valor provavel do producto será $M_{\alpha_1} M_{\alpha_2} \ldots M_{\alpha_\mu}$ e ter-se-ha

$$\overline{(e_1+e_2+\ldots+e_n)^{2p}}=\sum \frac{2p!}{\alpha_1!\,\alpha_2!\ldots\alpha_\mu!} \mathrm{N} \mathrm{M}_{\alpha_1} \mathrm{M}_{\alpha_2} \ldots \mathrm{M}_{\alpha_\mu}.$$

Os termos d'esta somma obter-se-hão fazendo agora variar os α, da maneira dicta.

Para calcular N, supponhamos primeiro os α todos differentes; será o numero de arranjos de n lettras μ a μ, visto que permutando os indices μ obteremos sempre termos differentes e os μ podem ser quaesquer dos indices 1, 2...n. Ora esse numero é, como se sabe,

$$n(n-1)\ldots(n-\mu+1).$$

Supponhamos agora μ_1 expoentes eguaes a α_1, μ_2 eguaes a $\alpha_2, \ldots, \mu_k$ eguaes a α_k, sendo

$$\mu_1 + \mu_2 + \ldots + \mu_k = \mu$$

e $\alpha_1, \alpha_2, \ldots, \alpha_k$ differentes. A permutação das lettras que têem o mesmo expoente dá termos eguaes; o numero é precisamente o das permutações d'essas lettras em cada arranjo, isto é, $\mu_1!$ quando se permutam as lettras de expoente α_1, $\mu_2!$ as de expoeute α_2 e assim por deante, sendo pois necessario para evitar as repetições dividir o numero total de arranjos pelo producto d'aquelles numeros. Será

$$N = \frac{n(n-1)\ldots(n-\mu+1)}{\mu_1!\,\mu_2!\ldots\mu_k!}.$$

Em N haverá termos em $n^\mu, n^{\mu-1}, \ldots$; se n é muito grande como aqui supposemos, os termos de grau mais elevado serão muito superiores aos de grau menos elevado. Consideremos só os termos em n^μ; o maior valor de μ é p, porque esse valor correspondendo ao menor valor par dos α, será ao mesmo tempo

$$\alpha_1 = \alpha_2 = \ldots = 2,$$

e

$$\alpha_1 + \alpha_2 + \ldots + \alpha_\mu = 2p;$$

logo

$$\mu = p.$$

Mas por outro lado d'entre os k valores de μ, só subsiste um: logo será approximadamente

$$\mathrm{N}=\frac{n(n-1)\ldots(n-p+1)}{p!}=\frac{n^p}{p!}.$$

Substituindo na formula que dá o valor provavel e notando que é agora

$$\mathrm{M}_{\alpha_1}=\mathrm{M}_{\alpha_2}=\ldots=\mathrm{M}_{\alpha_\mu}=\mathrm{M}_2,$$

vem

$$\overline{(e_1+e_2+\ldots+e_n)^{2p}}=\frac{2p!}{2^p}\,\frac{n^p}{p!}\,\mathrm{M}_2^p$$

e

$$\overline{\varepsilon^{2p}}=\frac{2p!}{p!2^p}\left(\frac{\mathrm{M}_2}{n}\right)^p$$

$$=\frac{2p!}{p!2^p}\,\overline{\varepsilon^2}^{\,p}\ldots\ldots\ldots\ldots(18)$$

por ser (n.º 31)

$$\frac{\mathrm{M}_2}{n}=\overline{\varepsilon^2}\,.$$

Supponhamos que ε segue a lei de Gauss.

O valor provavel d'uma potencia impar de ε será ainda nullo por ser φ par. Procuremos o valor provavel de ε^{2p} com esta lei, isto é, sendo

$$\varphi(\varepsilon)=\frac{h}{\sqrt{\pi}}e^{-h^2\varepsilon^2}.$$

Será

$$\overline{\varepsilon^{2p}} = \int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2} \varepsilon^{2p} d\varepsilon.$$

Ora

$$\int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} e^{-h^2\varepsilon^2} d\varepsilon = 1,$$

d'onde

$$\int_{-\infty}^{+\infty} e^{-h^2\varepsilon^2} d\varepsilon = \sqrt{\pi}\,(h^2)^{-\frac{1}{2}}.$$

Differenciando p vezes em ordem a h^2, vem

$$\int_{-\infty}^{+\infty} e^{-h^2\varepsilon^2} (-\varepsilon^2)^p d\varepsilon = \sqrt{\pi}\,(h^2)^{-p-\frac{1}{2}} \left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\cdots\left(-\frac{2p-1}{2}\right);$$

ou

$$\int_{-\infty}^{+\infty} e^{-h^2\varepsilon^2} \varepsilon^{2p} d\varepsilon = \sqrt{\pi}\,(h^2)^{-p-\frac{1}{2}} \frac{1\,.\,3\,.\,5\ldots(2p-1)}{2^p}$$

$$= \sqrt{\pi}\,(h^2)^{-p-\frac{1}{2}} \frac{2p!}{2^p\,.\,2\,.\,4\,.\,6\ldots 2p}$$

$$= \sqrt{\pi}\,(h^2)^{-p-\frac{1}{2}} \frac{2p!}{2^{2p}\,.\,p!}.$$

D'onde, multiplicando por

$$\frac{(h^2)^{\frac{1}{2}}}{\sqrt{\pi}},$$

se deduz

$$\overline{\varepsilon^{2p}} = \frac{2p!}{p!2^{2p}} \frac{1}{h^{2p}} \ldots\ldots\ldots\ldots\ldots\ldots (19).$$

Valor provavel de uma potencia do erro segundo a lei de Gauss.

Para que (18) e (19) sejam compativeis é preciso que

$$\overline{\varepsilon^{2}}^{p} = \left(\frac{1}{2h^2}\right)^p,$$

ou

$$\overline{\varepsilon^2} = \frac{1}{2h^2};$$

o que tem logar para a lei de Gauss, como o mostra (19) fazendo $p = 1$, e não póde ter logar para outra lei, como agora vamos provar.

Só a lei de Gass conduz á expressão (19).

Com effeito, supponhamos que existam duas fórmas de φ que conduzam ao mesmo resultado (19), isto é, á mesma formula para o valor provavel.

Consideremos a expressão

$$e^{-n(y_0 - y)^2} = \Sigma A_p y^{2p},$$

representando o segundo membro o desenvolvimento em serie segundo as potencias de y, sempre convergente. O valor provavel d'esta expressão é com a lei de Gauss

$$\int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} e^{-h^2y^2} e^{-n(y_0 - y)^2} dy = \Sigma A_p \overline{y^{2p}};$$

por serem os A_p coefficientes numericos sempre os mesmos, não é necessario tomar o valor provavel d'elles.

Com qualquer outra lei que designaremos por φ será

$$\int_{-\infty}^{+\infty} \varphi(y)\, e^{-n(y_0-y)^2} dy = \Sigma A_p \overline{y^{2p}}.$$

Logo

$$\frac{\int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} e^{-h^2y^2} e^{-n(y_0-y)^2} dy}{\int_{-\infty}^{+\infty} \varphi(y)\, e^{-n(y_0-y)^2} dy} = 1.$$

O limite d'esta relação para $n = \infty$ é, em virtude de ser $e^{-n(y^0-y)^2}$ maxima para $y = y_0$, e pelo que vimos na nota do n.º 37

$$\frac{\frac{h}{\sqrt{\pi}} e^{-hy_0^2}}{\varphi(y_0)} = 1,$$

d'onde

$$\varphi(y_0) = \frac{h}{\sqrt{\pi}} e^{-hy^{02}}$$

e portanto as duas leis coincidiriam.

44. Vimos na primeira parte d'este capitulo as duvidas que póde haver sobre a admissão do postulado de Gauss. D'esta segunda parte resulta que a lei de Gauss não apparece como uma consequencia rigorosa d'esse postulado.

A analyse d'algumas das objecções de Bertrand feita por Poincaré, mostra que para pôr a lei de probabilidade ao abrigo d'ellas é necessario dar-lhe uma fórma mais geral que contém duas funcções $\psi(v)$ e $\theta(o_i)$, impossiveis geralmente de determinar com rigor, sobre as quaes mesmo frequentes vezes nada se sabe.

A lei de Gauss corresponde ás hypotheses mais simples que podem fazer-se sobre estas funcções.

Por outro lado não é natural este fundamento da lei de probabilidade dos erros. O modo como os erros são originados, a sua natureza intima é completamente posta de parte.

CAPITULO III

A lei dos erros fundada no estudo dos erros elementares

I. Lei do erro total em funcção das leis dos erros elementares

45. Supponhamos primeiro o caso de dois erros elementares independentes y_1 e y_2, cujas leis são $\varphi_1(y_1)$, $\varphi_2(y_2)$, os quaes compõem o erro total z; sejam além d'isso a_1 e b_1 os limites inferior e superior do primeiro erro e a_2 e b_2 os limites correspondentes do segundo.

Pretende-se determinar a lei $\varphi(z)$ do erro composto em funcção de φ_1 e φ_2.

Se o erro y_1 cáe entre y_1 e $y_1 + dy_1$ e o erro z entre z e $z + dz$, y_2 deve ficar comprehendido evidentemente entre $z - y_1$ e $z - y_1 + dz$.

A probabilidade de y_1 estar entre y_1 e $y_1 + dy_1$ é $\varphi(y_1)dy_1$; analogamente a probabilidade de y_2 estar entre $z - y_1$ e $z - y_1 + dz$ é $\varphi_2(z - y_1)dz$. O concurso d'estes dois acontecimentos tem a probabilidade composta

$$\varphi_1(y_1)\varphi_2(z - y_1)dy_1dz.$$

Dando a y_1 todos os valores compativeis com o valor z, a probabilidade de que o erro total caia entre z e $z+dz$, obter-se-ha integrando aquella expressão em ordem a y_1 e como esta probabilidade é $\varphi(z)dz$, ter-se-ha

Caso particular de dois erros elementares.

$$\varphi(z)=\int\varphi_1(y_1)\varphi_2(z-y_1)dy_1.$$

Procuremos determinar os limites do integral.

Seja, fig. 2, $\varphi_1(y_1)$ representada pela recta material $o_1 a_1 b_1$, $\varphi_2(y_2)$ pela recta material $o_2 a_2 b_2$. Os valores possiveis de y_1 cáem na primeira linha entre a_1 e b_1, e os de y_2 na segunda entre a_2 e b_2. Tomemos uma terceira recta parallela oz para representar o erro total.

Cada valor de z assigna dois limites $z-b_2$ e $z-a_2$ aos valores de y_1 compativeis com elle, aos quaes correspondem na linha oz os pontos do segmento $z-b_2$, $z-a_2$. Mas, como os valores possiveis de y_1 ficam comprehendidos entre a_1 e b_1, só devem considerar-se os valores de y_1 representados por pontos correspondentes nos dois segmentos a_1, b_1 e $z-b_2$, $z-a_2$.

Esses limites dependem da posição do ultimo segmento, isto é, do valor de z.

Supponhamos que o ponto $z-a_2$ cáia debaixo de a_1, será

$$z=a_1+a_2,$$

e são estes os unicos pontos que se correspondem nos dois segmentos.

Se é $z-b_2$ que cáe debaixo de a_1, será

$$z=a_1+b_2$$

e os limites de y_1 serão

$$a_1 \quad \text{e} \quad a_1 + b_2 - a_2.$$

Póde $z - a_2$ cair debaixo de b_1: ter-se-ha

$$z = b_1 + a_2$$

e os limites serão

$$b_1 + a_2 - b_2 \quad \text{e} \quad b_1.$$

Finalmente se $z - b_2$ cair debaixo de b_1 será

$$z = b_1 + b_2$$

e não haverá nenhum par de pontos em correspondencia nos dois segmentos.

Estes quatro valores de z são evidentemente pontos de descontinuidade de $\varphi(z)$, havendo, pois, em geral, tres funcções differentes correspondentes aos tres intervallos limitados pelos quatro valores de z.

46. Supponhamos $a_1 = -b_1$, $a_2 = -b_2$, esses valores são

$$z_1 = -b_1 - b_2\,,\ z_2 = -b_1 + b_2\,,\ z_3 = b_1 - b_2\,,\ z_4 = b_1 + b_2$$

e, se $b_1 > b_2$, acham-se dispostos pela sua ordem de grandeza.

As fórmas de $\varphi(z)$ serão

$$\varphi_\alpha(z) = \int_{-b_1}^{z+b_2} \varphi_1(y_1)\varphi_2(z-y_1)\,dy_1$$

$$\varphi_\beta(z) = \int_{z-b_2}^{z+b_2} \varphi_1(y_1)\varphi_2(z-y_1)dy_1$$

$$\varphi_\gamma(z) = \int_{z-b_2}^{b_1} \varphi_1(y_1)\varphi_2(z-y_1)dy_1$$

sendo $\varphi_\alpha(z)$ a lei dos erros comprehendidos entre z_1 e z_2, porque, como se viu acima, os valores limites de y_1, são $a_1 = -b_1$ e $a_1 + b_2 - a_2 = -b_1 + b_2 + b_2 = z_2 + b_2$; $\varphi_\beta(z)$ a lei dos erros comprehendidos entre z e z_3, e $\varphi_\gamma(z)$ entre z_3 e z_4, como facilmente se vê d'um modo analogo.

Se suppozermos por exemplo que cada um dos erros elementares tem uma probabilidade constante, ficando y_1 entre $-b_1$ e $+b_1$ e y_2 entre $-b_2$ e $+b_2$, será

$$\varphi_1(x_1) = \frac{1}{2b_1},\ \varphi_2(x) = \frac{1}{2b_2}$$

e

$$\varphi_\alpha(z) = \frac{z + b_1 + b_2}{4b_1b_2}$$

$$\varphi_\beta(z) = \frac{2b_2}{4b_1b_2}$$

$$\varphi_\gamma(z) = \frac{-z + b_1 + b_2}{4b_1b_2}.$$

Ora emquanto as curvas de probabilidade de cada um dos erros elementares são rectas parallelas ao eixo onde se contam os erros, a do erro composto é uma linha quebrada com a fórma ABCD da fig. 3. Com effeito, a parte $\varphi_\beta(z)$ é constante, portanto a curva de probabilidade de z entre z_2 e z_3 será uma recta parallela ao eixo dos erros; φ_α e φ_γ são do primeiro grau em z, as curvas de probabilidade serão duas rectas, cujas equações são, designando por y a ordenada,

$$y = \frac{1}{4b_1b_2} z + \frac{b_1 + b_2}{4b_1b_2}$$

e

$$y = -\frac{1}{4b_1b_2} z + \frac{b_1 + b_2}{4b_1b_2}.$$

Como os coefficientes angulares são eguaes e de signaes contrarios, estas rectas formam angulos supplementares com o eixo dos z. Notemos além d'isso que a primeira encontra este eixo no ponto $z_1 = -b_1 - b_2$ e a segunda no ponto $z_4 = b_1 + b_2$, e que para os valores $z = z_2 = -b_1 + b_2$ e $z = z_3 = b_1 - b_2$, y toma sempre o mesmo valor $\frac{1}{2b_1}$. D'ahi a fórma attribuida á curva da probabilidade na figura citada.

47. Supponhamos agora que um dos erros elementares y_2 por exemplo, é susceptivel de todos os valores entre $-\infty$ e $+\infty$. Então $\varphi(z)$ terá só uma fórma

$$\varphi(z) = \int_{a_1}^{b_1} \varphi_1(y_1)\, \varphi_2(z - y_1)\, dy_1,$$

porque os valores de y_1 compativeis com os valores de z são agora todos os do segmento $a_1 b_1$.

Desenvolva-se $\varphi_2(z-y_1)$ segundo as potencias crescentes de y_1, será

$$\varphi(z)=\varphi_2(z)\int_{a_1}^{b_1}\varphi_1(y_1)dy_1-\varphi'_2(z)\int_{a_1}^{b_1}y_1\varphi_1(y_1)dy_1$$

$$+\frac{\varphi''_2(z)}{1\,.\,2}\int_{a_1}^{b_1}y_1^2\varphi_1(y_1)dy_1-\ldots$$

$$=\varphi_2(z)-k'_1\varphi'_2(z)+\frac{1}{1\,.\,2}k''_1\varphi''_2(z)-\ldots(1),$$

notando que é

$$\int_{a_1}^{b_1}\varphi_1(y_1)dy_1=1,$$

e fazendo para simplificar

$$\int_{a_1}^{b_1}y_1\varphi_1(y_1)dy_1=k'_1\,,$$

$$\int_{a_1}^{b_1}y_1^2\varphi_1(y_1)dy_1=k''_1\,,$$

....

48. Considerando tres erros elementares y_1, y_2, y_3, poderia primeiro compôr-se y_1 com y_2 e depois o resultado com y_3. A lei seria

$$\varphi(z) = \int\int \varphi_1(y_1)\,\varphi_2(y_2)\,\varphi_3(z - y_1 - y_2)\,dy_1\,dy_2.$$

Haveria 8 pontos de descontinuidade para z correspondentes ás combinações dos 4 valores anteriormente achados com os limites do novo erro y_3. Aos 7 intervallos que estes 8 valores delimitam corresponderiam outras tantas fórmas de $\varphi(z)$.

Em geral para n erros elementares haverá 2^n pontos de descontinuidade para $\varphi(z)$, que terá assim $2^n - 1$ fórmas differentes.

Caso geral. Consideremos este caso geral de ser o erro total formado de n erros elementares $y_1, y_2, \ldots y_n$, cujas leis são respectivamente $\varphi_1(y_1), \varphi_2(y_2), \ldots, \varphi_n(y_n)$ e que se acham comprehendidos entre os limites $-a_1$ e $+a_1, -a_2$ e $+a_2, \ldots, -a_n$ e $+a_n$.

A probabilidade de que o erro total cáia entre z e $z + dz$, será evidentemente representada pelo integral multiplo

$$\varphi(z)dz = \int\int \ldots \int \varphi_1(y_1)\,\varphi_2(y_2) \ldots \varphi_n(y_n)dy_1 dy_2 \ldots dy_n \ldots \quad (2).$$

Como o erro total deve estar comprehendido entre z e $z + dz$, e como $z = y_1 + y_2 + \ldots + y_n$ é preciso escolher dos valores possiveis das n variaveis aquelles cuja somma estiver comprehendida entre z e $z + dz$.

Introduzamos debaixo do signal integral um factor de descontinuidade gozando da propriedade de se annullar logo que a mesma somma está fóra do referido intervallo e de ter o valor 1 para todo esse intervallo. Satisfaz a estas condições o factor

$$\frac{dz}{2\pi}\int_{-\infty}^{+\infty}\cos(y_1+y_2+\ldots+y_n-z)\theta d\theta \ldots\ldots (3).$$

Com effeito considerando o integral

$$\frac{1}{\pi}\int_0^{\infty}\frac{\operatorname{sen}(s-z)\,\theta d\theta}{\theta}=\mp\frac{1}{2}$$

segundo é $s \lessgtr z$, será

$$\frac{1}{\pi}\int_0^{\infty}\frac{\operatorname{sen}(s-z-dz)\,\theta d\theta}{\theta}-\frac{1}{\pi}\int_0^{\infty}\frac{\operatorname{sen}(s-z)\,\theta d\theta}{\theta}$$

egual a zero para $s>z+dz$ ou para $s<z$, porque para estes valores ambos os termos tomam o mesmo valor $+\frac{1}{2}$ ou $-\frac{1}{2}$; para os valores de s comprehendidos entre z e $z+dz$ tornar-se-ha egual a -1 e finalmente para $s=z$ e $s=z+dz$, egual a $-\frac{1}{2}$.

Mas

$$\operatorname{sen}(s-z-dz)\,\theta=\operatorname{sen}(s-z)\,\theta-\theta dz\cos(s-z)\,\theta+\ldots.$$

Desprezando os termos de ordem superior á primeira a differença dos dois integraes considerados tornar-se-ha em

$$-\frac{dz}{\pi}\int_0^{\infty}\cos(s-z)\,\theta\,d\theta=-\frac{dz}{2\pi}\int_{-\infty}^{+\infty}\cos(s-z)\,\theta\,d\theta$$

que para $s=y_1+y_2+\ldots+y_n$ é o factor (3) com o signal contrario. Em vez do valor -1, tomará esse factor o valor $+1$

quando $z < s < z + dz$ e o mesmo valor zero fóra d'este dominio.

Passando de funcções trigonometricas para exponenciaes, póde em vez d'esse factor tomar-se

$$\frac{dz}{2\pi}\int_{-\infty}^{+\infty} e^{(y_1+y_2+\ldots+y_n-z)\,\theta\sqrt{-1}}\,d\theta$$

que introduzido em (2) dá, dividindo ambos os membros por dz

Lei do erro total.

$$\varphi(z)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\int_{-a_1}^{+a_1}\varphi_1(y_1)e^{\theta y_1\sqrt{-1}}dy_1\int_{-a_2}^{+a_2}\varphi_2(y_2)e^{\theta y_2\sqrt{-1}}dy_2\ldots\right.$$

$$\left.\ldots\int_{-a_n}^{+a_n}\varphi_n(y_n)\,e^{\theta y_n\sqrt{-1}}dy_n\right]e^{-z\theta\sqrt{-1}}d\theta\ldots(4).$$

É a pretendida formula da *lei do erro total* ou *composto* em funcção das leis dos erros elementares.

49. Se conhecessemos essas leis e os limites dos erros teriamos assim, pelo menos theoricamente, a maneira de obter a lei do erro total.

A investigação das leis dos erros elementares seria, pois, a primeira tarefa.

Leis dos erros elementares.

Supponhamos o erro y_1 proveniente de uma só origem ou causa d'erros, e que se descobriu a relação de dependencia entre esta causa e uma grandeza x cuja lei de probabilidade é conhecida. Seja $\psi(x)$ esta lei e $y_1 = \theta(x)$ a relação entre y_1 e x.

A lei $\varphi(y_1)$ do erro y_1 deduz-se então facilmente. Com effeito, correspondendo a cada valor de x um certo valor de y_1, a probabilidade de que x esteja comprehendido entre dois limites x e $x+dx$ é egual á probabilidade de que y_1 esteja entre os limites correspondentes y_1 e y_1+dy_1. Portanto

$$\varphi(y_1)\, dy_1 = \psi(x)\, dx.$$

Mas $dy_1 = \theta'(x)\, dx$, logo

$$\varphi(y_1) = \frac{\psi(x)}{\theta'(x)}.$$

Esta formula póde facilitar em alguns casos a indagação das leis dos erros elementares.

Em geral, porém, não só será extremamente difficil descobrir estas leis, mas nem mesmo se conseguirão achar todas as origens de erros.

Para supprir esta deficiencia é necessario estabelecer hypotheses mais ou menos geraes sobre os erros elementares, das quaes se deduza a lei do erro total.

A formula a que chegamos no numero anterior póde utilisar-se nessa deducção.

II. Deducção da lei do erro total partindo de hypotheses sobre os erros elementares

50. No capitulo I, n.º 7, enunciamos algumas propriedades que é natural attribuir aos erros elementares, porque ellas estão a maior parte das vezes d'accordo com os factos. Taes são: o grande numero de erros elementares, a sua independencia relativa, a pequena importancia de cada um d'elles em relação ao total e a importancia quasi egual de uns em relação aos outros.

Essas propriedades entram em todas ou quasi todas as hypotheses que se tem feito sobre a composição do erro total.

Convém aqui precisar melhor como se deve avaliar a importancia de uma causa d'erro.

É claro que é necessario attender não só á grandeza absoluta de cada erro que essa causa póde produzir mas tambem á sua probabilidade respectiva. As considerações que fizemos no capitulo I, n.º 15, a proposito da escolha dos valores mais convenientes das incognitas, applicam-se aqui tambem. O erro mais provavel não daria uma apreciação justa da influencia da causa de erro, e parece que agora como então é o valor provavel das differentes potencias do erro que resolve o problema.

Assim para o erro y cuja lei é $\varphi(y)$ deveremos considerar para avaliação da sua importancia as quantidades

$$\begin{aligned} k' &= \int y\varphi(y)\,dz \\ k'' &= \int y^2\varphi(y)\,dz \\ k''' &= \int y^3\varphi(y)\,dz \\ &\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

Mas as potencias pares são as mais proprias; querendo empregar as impares devem tomar-se os erros com o mesmo signal.

51. Admittamos as seguintes hypotheses:

Theoria de Hagen.

1.ª Os erros elementares são em grande numero, eguaes, muito pequenos e independentes uns dos outros.

2.ª Cada um d'elles póde actuar no resultado da observação com a mesma probabilidade no sentido positivo ou negativo.

Seja $2n$ o numero dos erros elementares; ε a grandeza absoluta de qualquer d'elles. As supposições feitas equivalem a comparar a formação do erro total a um lance de $2n$ dados, taes que só podem apparecer duas faces marcadas uma com

$+\varepsilon$ e outra com $-\varepsilon$ e qualquer d'ellas com a probabilidade $\frac{1}{2}$. O erro total será egual á somma algebrica de todos os numeros que sahem no lance.

A seguinte tabella, cuja construcção é obvia, dá os valores possiveis do erro total, o numero de casos favoraveis á sua producção e o numero total de casos possiveis, para $2n$ erros elementares.

Numero de erros elementares eguaes a $-\varepsilon$	Numero de erros elementares eguaes a $+\varepsilon$	Valores do erro total	Numero de casos favoraveis	Numero total de casos possiveis
0	$2n$	$+2n\varepsilon$	1	
1	$2n-1$	$+2(n-1)\varepsilon$	$2n$	
2	$2n-2$	$+2(n-2)\varepsilon$	$\binom{2n}{2}$	
.	.	.	.	
.	.	.	.	
.	.	.	.	
n	n	0	$\binom{2n}{n}$	2^{2n}
.	.	.	.	
.	.	.	.	
.	.	.	.	
$2n-2$	2	$-2(n-2)\varepsilon$	$\binom{2n}{2}$	
$2n-1$	1	$-2(n-1)\varepsilon$	$2n$	
$2n$	0	$-2n\varepsilon$	1	

Tendo o numero de casos favoraveis e o numero de casos possiveis, basta, para calcular a probabilidade de que o erro total tenha um valor determinado $2m\varepsilon$, tomar a relação entre esses dois numeros, isto é

$$P_{2m\varepsilon}=\binom{2n}{n-m}2^{-2n}.$$

A maior d'estas probabilidades é a do erro total nullo ou

$$P_0 = \binom{2n}{n} 2^{-2n}.$$

Consideremos a relação

$$\frac{P_{2m\varepsilon}}{P_0} = \frac{\binom{2n}{n-m}}{\binom{2n}{n}}$$

$$= \frac{n!\,n!}{(n-m)!\,(n+m)!}.$$

Conhece-se do *Calculo das probabilidades* a formula de Stirling para calcular a factorial de um numero:

$$x! = \sqrt{2\pi}\, x^{x+\frac{1}{2}} e^{-x+\frac{1}{12x}-\frac{1}{360x^3}+\frac{1}{1260x^5}-\cdots}.$$

Appliquemol-a á formula anterior:

$$\frac{P_{2m\varepsilon}}{P_o} = \frac{2\pi n^{2\left(n+\frac{1}{2}\right)}}{2\pi (n-m)^{n-m+\frac{1}{2}} (n+m)^{n+m+\frac{1}{2}}}$$

$$\times \frac{e^{-2n+\frac{1}{6n}-\frac{1}{180n^3}+\cdots}}{e^{-2n+\frac{n}{6(n^2-m^2)}-\frac{n^3+3nm^2}{180(n^2-m^2)^3}+\cdots}}$$

$$= \left(\frac{n}{n-m}\right)^{n-m+\frac{1}{2}} \left(\frac{n}{n+m}\right)^{n+m+\frac{1}{2}}$$

$$\times e^{+\frac{1}{6}\left(\frac{1}{n}-\frac{n}{n^2-m^2}\right)-\frac{1}{180}\left(\frac{1}{n^3}-\frac{n^3+3nm^2}{(n^2-m^2)^3}\right)+\cdots} \quad \ldots\ldots (1).$$

E como é

$$l(1+x)=\frac{x}{1}-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots,$$

vem

$$l\left(\frac{n}{n-m}\right)=-l\left(1-\frac{m}{n}\right)=+\frac{m}{n}+\frac{1}{2}\left(\frac{m}{n}\right)^2+\frac{1}{3}\left(\frac{m}{n}\right)^3+\frac{1}{4}\left(\frac{m}{n}\right)^4+\ldots,$$

$$l\left(\frac{n}{n+m}\right)=-l\left(1+\frac{m}{n}\right)=-\frac{m}{n}+\frac{1}{2}\left(\frac{m}{n}\right)^2-\frac{1}{3}\left(\frac{m}{n}\right)^3+\frac{1}{4}\left(\frac{m}{n}\right)^4-\ldots,$$

d'onde

$$\left(n-m+\frac{1}{2}\right)l\left(\frac{n}{n-m}\right)+\left(n+m+\frac{1}{2}\right)l\left(\frac{n}{n+m}\right)$$

$$=-2m\frac{m}{n}+(2n+1)\frac{1}{2}\left(\frac{m}{n}\right)^2-2m\frac{1}{3}\left(\frac{m}{n}\right)^3+(2n+1)\frac{1}{4}\left(\frac{m}{n}\right)^4-\ldots$$

$$=-\left(n-\frac{1}{2}\right)\left(\frac{m}{n}\right)^2-\left(\frac{1}{6}n-\frac{1}{4}\right)\left(\frac{m}{n}\right)^4-\left(\frac{1}{15}n-\frac{1}{6}\right)\left(\frac{m}{n}\right)^6-$$

$$-\left(\frac{1}{28}n-\frac{1}{8}\right)\left(\frac{m}{n}\right)^8-\ldots$$

$$=-\frac{m^2}{n}-\frac{1}{6}\frac{m^4}{n^3}-\frac{1}{15}\frac{m^6}{n^5}-\frac{1}{28}\frac{m^8}{n^7}\ldots$$

desprezando $\frac{1}{2},\frac{1}{4},\frac{1}{6},\frac{1}{8}\ldots$ em relação a $n,\frac{n}{6},\frac{n}{15},\frac{n}{28},\ldots$

Portanto, passando dos logarithmos para as exponenciaes a ultima egualdade torna-se

$$\left(\frac{n}{n-m}\right)^{n-m+\frac{1}{2}}\left(\frac{n}{n+m}\right)^{n+m+\frac{1}{2}}=e^{-\frac{m^2}{n}-\frac{1}{6}\frac{m^4}{n^3}-\frac{1}{28}\frac{m^8}{n^7}-\cdots} \quad (2).$$

Mas é além d'isso

$$+\frac{1}{6}\left(\frac{1}{n}-\frac{n}{n^2-m^2}\right)=-\frac{1}{6n^2}\frac{m^2}{n}-\frac{1}{6n^2}\frac{m^4}{n^3}-\frac{1}{6n^2}\frac{m^6}{n^5}-$$
$$-\frac{1}{6n^2}\frac{m^8}{n^7}-\cdots\cdots \quad (3)$$

$$-\frac{1}{180}\left(\frac{1}{n^3}-\frac{n^3+3nm^2}{(n^2-m^2)^3}\right)=+\frac{1}{30n^4}\frac{m^2}{n}+\frac{1}{12n^4}\frac{m^4}{n^3}$$
$$+\frac{7}{45n^4}\frac{m^6}{n^5}+\frac{1}{4n^4}\frac{m^8}{n^7}-\cdots\cdots \quad (4).$$

Tendo em attenção as formulas (2), (3) e (4), transforma-se (1) em

$$\frac{P_{2m\varepsilon}}{P_0}=e^{-\frac{m^2}{n}\left(1+\frac{1}{6n^2}-\cdots\right)-\frac{1}{6}\frac{m^4}{n^3}\left(1+\frac{1}{n^2}-\cdots\right)}$$
$$\times e^{-\frac{1}{15}\frac{m^6}{n^5}\left(1+\frac{5}{2n^2}-\cdots\right)-\frac{1}{28}\frac{m^8}{n^7}\left(1+\frac{14}{3n^2}-\cdots\right)-\cdots}$$
$$=e^{-\frac{m^2}{n}-\frac{1}{6}\frac{m^4}{n^3}-\frac{1}{15}\frac{m^6}{n^5}-\frac{1}{28}\frac{m^8}{n^7}-\cdots}$$

desprezando em relação á unidade $\frac{1}{n^2}, \frac{1}{n^4}\cdots$

Por outro lado, applicando a mesma formula de Stirling a P_0 sómente, vem

$$P_0 = \binom{2n}{n} 2^{-2n} = \frac{2n\,!}{n\,!\,n\,!} 2^{-2n}$$

$$= \frac{\sqrt{2\pi}.(2n)^{2n+\frac{1}{2}}.e^{-2n+\frac{1}{12.2n}-\frac{1}{360(2n)^3}+\cdots}}{2\pi.n^{2\left(n+\frac{1}{2}\right)}.e^{-2n+\frac{1}{6n}-\frac{1}{180n^3}+\cdots}} 2^{-2n}$$

$$= \frac{1}{\sqrt{n\pi}} e^{\cdot -\frac{1}{8n}+\frac{1}{192n^3}+\cdots}$$

$$= \frac{1}{\sqrt{n\pi}} \cdot \left(1 - \frac{1}{8n} + \frac{1}{128n^2} + \frac{5}{1024n^3} + \cdots\right)$$

$$= \frac{1}{\sqrt{n\pi}}$$

approximadamente.

Logo, fazendo ainda despresos da mesma ordem,

$$P_{2m\varepsilon} = \frac{1}{\sqrt{n\pi}}\left(1 - \frac{m^2}{n} + \frac{1}{2!}\left(\frac{m^2}{n}\right)^2 - \frac{1}{3!}\left(\frac{m^2}{n}\right)^4 + \frac{1}{4!}\left(\frac{m^2}{n}\right)^4 - \cdots\right)$$

e, porque a expressão encerrada entre parenthesis é o desenvolvimento de $e^{-\frac{m^2}{n}}$,

$$P_{2m\varepsilon} = \frac{1}{\sqrt{n\pi}} e^{-\frac{m^2}{n}}$$

Pondo agora $2m\varepsilon = y$, vem

$$P_{2m\varepsilon} = \frac{1}{\sqrt{n\pi}} e^{-\frac{y^2}{4n\varepsilon^2}};$$

ou fazendo $\frac{1}{4n\varepsilon^2} = h^2$

$$P_{2m\varepsilon} = \frac{2\varepsilon h}{\sqrt{\pi}} e^{-h^2y^2}.$$

Vê-se que $P_{2m\varepsilon}$ é proporcional a $\frac{h}{\sqrt{\pi}} e^{-h^2y^2}$ e póde portanto escrever-se

$$\varphi(y) = \frac{h}{\sqrt{\pi}} e^{-h^2y^2}.$$

Chegamos á lei de Gauss, com uma lei approximada, tanto mais quanto maior fôr o numero de origens d'erros.

O merecimento d'esta theoria, devida a Hagen, está no caracter elementar que ella tem. As hypotheses d'onde se partiu é que infelizmente não são plausiveis. Por virtude d'ellas os erros elementares seguiriam todos a mesma lei, o que não é admissivel.

52. O mesmo defeito tem a hypothese mais geral, sobre que Laplace assentou a sua theoria, que consiste em suppôr os erros elementares em grande numero e sujeitos a uma lei commum que attribue probabilidade egual a erros numericamente eguaes.

Theorias de Laplace e Tait.

A theoria de Hagen suppõe além d'isso a egualdade dos erros elementares, o que a torna menos verdadeira; mas é em compensação mais simples, e dá uma concepção muito lucida do modo como os erros elementares concorrem para a formação do erro total.

Uma outra theoria, que se approxima da de Hagen, é a de Tait. Este auctor compara o erro resultante de qualquer origem ao desvio que existe entre o resultado de um grande numero de tiragens de uma urna contendo espheras brancas e pretas e o resultado mais provavel da tiragem.

O erro total é aquelle que tem maior probabilidade de coexistir com os desvios correspondentes ás differentes origens, suppondo para cada origem uma urna, com um conteúdo desconhecido de espheras brancas e pretas.

Para cada erro elementar acha uma lei exponencial e tambem para o erro total da fórma

$$\varphi(z) = Ce^{-h^2z^2}$$

53. Mais geral é a hypothese de Bessel: Os erros elementares actuam segundo leis quaesquer, satisfazendo unicamente á condição de serem funcções pares e de os valores provaveis dos quadrados dos erros elementares serem grandezas da mesma ordem. Theoria de Bessel.

Se na formula (4) do n.º 48 substituirmos as exponenciaes pelas funcções trignometricas e attendermos a que por hypothese se tem $\varphi_i(y_i) = \varphi_i(-y_i)$ os senos desapparecerão e teremos simplesmente

$$\varphi(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \mathrm{P} \cos z\theta \, d\theta$$

onde P é um producto de n factores da fórma

$$\int_{-a_i}^{+a_i} \varphi_i(y_i) \cos \theta y_i \, dz_i$$

que se obtêm dando a i os valores $1, 2, \ldots, n$.

Mas desenvolvendo em serie $\cos\theta y_i$ vem

$$\int_{-a_i}^{+a_i}\varphi_i(y_i)\cos\theta y_i\,dz_i = 1 - \frac{k_i''}{2}\theta^2 + \frac{k_i^{\text{iv}}}{24}\theta^4 - \frac{k_i^{\text{vi}}}{720}\theta^6 + \ldots$$

designando por k_i, k_i^{iv}, ... os valores provaveis do quadrado, da 4.ª potencia ... de y_i. Tomando os logarithmos neperianos de ambos os membros e representando o primeiro membro por S_i vem

$$\begin{aligned} \text{l. } S_i = & -\frac{k_i''}{2}\theta^2 + \frac{k_i^{\text{iv}}}{24}\theta^4 - \frac{k_i^{\text{vi}}}{720}\theta^6 + \ldots \\ & -\frac{k_i''^2}{8}\theta^4 + \frac{k_i''\,k_i^{\text{iv}}}{48}\theta^6 + \ldots \\ & -\frac{k_i''^3}{24}\theta^6 + \ldots \\ & + \ldots\ldots\ldots\ldots\ldots\ldots \end{aligned}$$

ou

$$\text{l. } S_i = -\frac{k_i''}{2}\theta^2 - \frac{3k_i''^2 - k_i^{\text{iv}}}{24}\theta^2 - \frac{30k_i''^3 - 15k_i''\,k_i^{\text{iv}} + k_i^{\text{vi}}}{720}\theta^6 + \ldots.$$

Logo

$$\text{l. } P = -\frac{\Sigma k''}{2}\theta^2 - \frac{3\Sigma k''^2 - \Sigma k^{\text{iv}}}{24}\theta^4 - \frac{30\Sigma k''^3 - 15\Sigma k'' k^{\text{iv}} + \Sigma k^{\text{v}}}{720}\theta^6 + \ldots$$

e

$$\varphi(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-\frac{\Sigma k''}{2}\theta^2}\cos z\theta\,d\theta\left[1 - \frac{3\Sigma k''^2 - \Sigma k^{\text{iv}}}{24}\theta^4 \right.$$

$$\left. - \frac{30\Sigma k''^3 - 15\Sigma k'' k^{\text{iv}} + \Sigma k^{\text{vi}}}{720}\theta^6 + \ldots\right].$$

Attendendo agora a que é

$$\int_{-\infty}^{+\infty} \theta^{2n} d\theta \cos z\theta \, e^{-a^2\theta^2} = \frac{\sqrt{\pi}}{a} \frac{d^{2n}}{dz^{2n}} e^{-\frac{z^2}{4a^2}}$$

como o mostrou Laplace (*Theorie analytique des Probabilités*, liv. I, n.º 25), e fazendo $a = \frac{\Sigma k''}{2}$ e $n = 1, 2, 3 \ldots$, vem

$$\varphi(z) = \frac{e^{-\frac{z^2}{2\Sigma k''}}}{\sqrt{2\pi\Sigma k''}} (1 - s)$$

onde

$$s = \frac{3\Sigma k''^2 - \Sigma k^{iv}}{24(\Sigma k'')^2}\left(3 - \frac{6z^2}{\Sigma k''} + \frac{z^4}{\Sigma k''^2}\right)$$
$$+ \frac{30\Sigma k''^3 - 15\Sigma k'' k^{iv} + \Sigma k^{vi}}{720(\Sigma k'')^3}\left(15 - \frac{45z^2}{\Sigma k''} + \frac{15z^4}{(\Sigma k'')^2} - \frac{z^6}{(\Sigma k'')^3}\right)$$
$$+ \ldots\ldots\ldots\ldots\ldots\ldots$$

É facil agora de vêr que, por ser n grande, póde desprezar-se em $\varphi(z)$ a parte multiplicada por s.

Com effeito, tem-se

$$\frac{e^{-\frac{z^2}{2\Sigma k''}}}{\sqrt{2\pi\Sigma k''}} s < \frac{1}{\sqrt{2\pi\Sigma k''}}\left(\frac{3\Sigma k''^2 - \Sigma k^{iv}}{8(\Sigma k'')^2} + \frac{30\Sigma k''^3 - 15\Sigma k'' k^{iv} + \Sigma k^{vi}}{48(\Sigma k'')^3}\right)$$

correspondendo o segundo membro a $z = 0$.

Procuremos avaliar a ordem d'esta expressão. Sendo k'' o valor provavel do quadrado de um erro elementar e sendo este valor da mesma ordem de grandeza, qualquer que seja a origem d'erro considerada, como se suppoz inicialmente, poderemos considerar $\Sigma k''$ da ordem $n\varepsilon^2$, $\Sigma k''^2$ e Σk^{iv} da ordem $n\varepsilon^4$, $\Sigma k''^3$, $\Sigma k''k^{\text{iv}}$ e Σk^{vi} da ordem $n\varepsilon^6$, ... E portanto será o 1.º termo de s da ordem $n^{-\frac{3}{2}}$ e o segundo da ordem $n^{-\frac{5}{2}}$, isto é quantidades muito pequena.

Logo, desprezando-os, terá $\varphi(z)$ a fórma da lei de Gauss.

Theoria de Crofton.

54. Resta-nos expôr a theoria de Crofton, que é sem duvida a que satisfaz melhor, pela grande generalidade que tem a hypothese sobre que assenta.

Suppõe-se sómente que ha um grande numero de origens de erros, independentes, e taes que cada uma d'ellas actuando isoladamente produz erros de muito pequena importancia em relação aos que resultam de todas as restantes origens actuando conjunctamente.

Esta importancia avalia-se, como já dissémos, pelos valores provaveis das potencias dos erros. Diz-se que um erro tem uma importancia ou uma influencia i vezes menor do que um outro quando os valores provaveis das potencias successivas do primeiro são diminuidos na proporção das potencias de egual grau do numero i, e o primeiro erro chama-se um *diminutivo* do segundo.

Assim, sendo k', k'', k'''... os valores provaveis de um erro y_2, resultante da combinação de $n-1$ origens de erros, o erro y_1 que resulta da origem restante terá os valores provaveis correspondentes: $\frac{k'}{i}$, $\frac{k''}{i^2}$, $\frac{k'''}{i^3}$, ...

i deve ser por hypothese muito grande e em consequencia poderão desprezar-se $\frac{k'''}{i^3}$, $\frac{k^{\text{iv}}}{i^4}$, ... em relação a $\frac{k'}{i}$, $\frac{k''}{i^2}$. O mesmo se não póde dizer de $\frac{k''}{i^2}$ em relação a $\frac{k'}{i}$, porque k' póde ter um valor muito pequeno ou nullo, por ser formado de termos positivos e negativos.

O erro total z póde imaginar-se composto dos dois erros y_1 e y_2 e a sua lei será dada (n.º 47) por

$$\varphi(z)=\varphi_2(z)-k_1'\varphi_2'(z)+\frac{k_1''}{2}\varphi_2''(z)$$

sendo k_1' e k_1'' os valores provaveis de y_1 e y_1^2, e desprezando os termos do grau superior ao 2.º em virtude do que se disse.

Esta egualdade mostrou a Crofton que a lei do erro total, não dependendo senão de k_1' e k_1'', deve ficar a mesma quando em vez da verdadeira lei do erro elementar se tomar outra que permitta desprezar os valores provaveis k_1''', $k_1^{\text{iv}}\ldots$ em relação aos dois primeiros; e levou-o á descoberta da lei

$$\varphi_1(y_1)=\frac{1}{\sqrt{2\pi(k_1''-k_1'^2)}}\,e^{-\frac{(x-k_1')^2}{2(k_1''-k_1'^2)}}=\frac{1}{\sqrt{\pi}}$$

que satisfaz áquellas condições.

Com effeito

$$\int_{-\infty}^{+\infty}y_1\varphi(y_1)dy_1=\frac{1}{\sqrt{2\pi(k_1''-k_1'^2)}}\left[\int_{-\infty}^{+\infty}(y_1-k_1')\,e^{-\frac{(y_1-k_1')^2}{2(k_1''-k_1'^2)}}dy_1\right.$$
$$\left.+k_1'\int_{-\infty}^{+\infty}e^{-\frac{(y_1-k_1')^2}{2(k_1''-k_1'^2)}}dy_1\right]=k_1'$$

e analogamente:

$$\int_{-\infty}^{+\infty}y_1^2\varphi(y_1)dy_1=C\left[\int_{-\infty}^{+\infty}(y_1-k_1')^2\,e^{-u^2}dy_1+2k_1'\int_{-\infty}^{+\infty}e^{-u^2}dy_1\right.$$
$$\left.-k_1'^2\int_{-\infty}^{+\infty}e^{-u^2}dy_1\right]=k_1''$$

representando nesta fórmula, para a tornar mais breve, $\frac{1}{\sqrt{2\pi(k_1''-k_1'^2)}}$ por C e o expoente de e por u^2.

Tem-se tambem

$$\int_{-\infty}^{+\infty} y_1^3\varphi(y_1)dy_1 = k_1'(3k_1''-2k_1'^2)$$

$$\int_{-\infty}^{+\infty} y_1^4\varphi(y_1)dy_1 = 3k_1''^2 - 2k_1'^4$$

. . . .

Por estas fórmulas vê-se que os valores provaveis da primeira e segunda potencia de y se conservaram os mesmos: k_1' e k_1''. Emquanto ao da terceira, quarta, etc., temos as suas expressões em funcção de k_1 e k_1'' e por ellas se vê que a sua ordem é $\frac{1}{i^3}$ e $\frac{1}{i^4}$, attendendo a que é k_1' da ordem $\frac{1}{i}$ e k_1'' da ordem $\frac{1}{i^2}$. Podem portanto desprezar-se, e a lei proposta satisfaz.

Para cada um dos erros elementares $y_1, y_2, \ldots, y_n$ adoptamos uma lei da mesma fórma, bastando para a sua expressão analytica affectar as quantidades y, k' e k'', respectivamente dos indices 1, 2, . . ., n.

Consideremos primeiro sómente os erros y_1 e y_2, seja z_1 o erro composto e $f_1(z_1)$ a sua lei, será em virtude do que se viu no n.º 45

$$f_1(z_1) = \int \varphi_1(y_1)\varphi_2(z_1-y_1)dy_1$$

$$= \frac{1}{\pi\alpha_1\alpha_2}\int_{-\infty}^{+\infty} e^{-\frac{(y_1-k_1')^2}{\alpha_1^2}-\frac{(z_1-y_1-k_2')^2}{\alpha_2^2}}\, dy_1$$

onde se fez para abreviar $\alpha_1 = \sqrt{2(k_1''-k_1'^2)}$ e $\alpha_2 = \sqrt{2(k_2''-k_2'^2)}$.

Para effectuar a integração notemos que $f_1(z_1)$ se póde pôr sob a fórma

$$f_1(z_1)=\frac{1}{\pi\sqrt{\alpha_1^2+\alpha_2^2}}\,e^{-\frac{(z_1-k_1'-k_2')^2}{\alpha_1^2+\alpha_2^2}}\int e^{-t^2}\,dt$$

fazendo a substituição

$$\frac{\sqrt{\alpha_1^2+\alpha_2^2}}{\alpha_1\alpha_2}\,y_1-\frac{k_1'\alpha_2^2-k_2'\alpha_1^2+\alpha_1^2 y_1}{\alpha_1\alpha_2\sqrt{\alpha_1^2+\alpha_2^2}}=t,$$

que lembra, logo que se tenha ordenado o expoente de e segundo as potencias de y_1. Ora $\int_{-\infty}^{+\infty} e^{-t^2}\,dt=\sqrt{\pi}$, portanto

$$f_1(z_1)=\frac{1}{\sqrt{\pi(\alpha_1^2+\alpha_2^2)}}\,e^{-\frac{(z_1-k_1'-k_2')^2}{\alpha_1^2+\alpha_2^2}}.$$

Considere-se agora o terceiro erro y_3, e designe-se por z_2 o erro composto; vê-se mesmo sem effectuar o calculo que se obterá para a lei de z_2

$$f_2(z_2)=\frac{1}{\sqrt{\pi(\alpha_1^2+\alpha_2^2+\alpha_3^2)}}\,e^{-\frac{(z_2-k_1'-k_2'-k_3')^2}{\alpha_1^2+\alpha_2^2+\alpha_3^2}}.$$

E finalmente para $f_{n-1}(z_{n-1})=\varphi(z)$

$$\varphi(z)=\frac{1}{\sqrt{\pi(\alpha_1^2+\alpha_2^2+\ldots+\alpha_n^2)}}\,e^{-\frac{(z-k_1'-k_2'-\ldots-k_n')}{\alpha_1^2+\alpha_2^2+\ldots+\alpha_n^2}}$$

$$=\frac{1}{\sqrt{2\pi(\Sigma k''-\Sigma k'^2)}}\,e^{-\frac{(z-\Sigma k')^2}{2(\Sigma k''-\Sigma k'^2)}}\ \ldots\ldots\ldots\ldots\ldots(5).$$ Lei de Crofton.

É a lei do erro composto a que se chega admittindo apenas as hypotheses feitas no principio d'esta deducção.

A funcção $\varphi(z)$ é analoga á lei supposta para cada erro elementar. É facil de vêr, porisso, que semelhantemente ao que succedia com esta ultima se terá, designando por K' o valor provavel da primeira potencia de z

$$K' = \int_{-\infty}^{+\infty} z\varphi(z)dz = \Sigma k'.$$

Se agora ás hypotheses estabelecidas se accrescenta a de que a lei dos erros seja par, ou que os erros positivos e negativos de egual valor absoluto sejam egualmente provaveis, os valores provaveis das potencias impares são nullos e portanto

$$K' = \Sigma k' = 0$$

o que reduz $\varphi(z)$ á lei de Gauss.

Critica.

55. A theoria de Crofton que acabamos de expôr apresenta sobre as anteriores vantagens incontestaveis.

A hypothese de Laplace e a de Hagen, que é um caso particular da primeira, são muito restrictas pois que obrigam todos os erros elementares a seguir uma mesma lei, o que está em desaccordo manifesto com os factos.

As causas ou origens de erros actuam de maneiras muito variadas, pelo menos na apparencia, e *a priori* a unica hypothese razoavel é deixar indeterminada a acção d'essas causas.

A theoria de Tait está livre quasi inteiramente d'esta critica, mas em compensação parte de uma supposição perfeitamente arbitraria. Demais essa hypothese traz tambem a consequencia de que os erros elementares sigam todos a lei exponencial, o que não póde ser acceite, e mostra que implicitamente se acha contida na hypothese fundamental uma restricção sobre as fórmas das leis dos erros elementares.

A analyse de Bessel apresenta uma enorme vantagem sobre as tres anteriores. Aqui os erros elementares seguem leis quaesquer, comtanto que sejam funcções pares. Exige-se além d'isso que as differentes origens tenham uma importancia relativa proximamente egual.

Isto succederá a maior parte das vezes nas observações modernas.

Mas a hypothese de Crofton é ainda mais geral pois que apenas se exige que cada origem actuando só produza erros de muito pequena importancia em relação aos que produziriam as outras todas actuando simultaneamente.

Como o numero de origens é supposto muito grande, e isso tanto nesta theoria como nas precedentes, é claro que esta hypothese equivale a suppôr que os erros elementares sejam diminutivos do erro total. A amplitude d'esses erros póde ser finita e até mesmo infinita. Além d'isso a importancia relativa das origens póde ser bastante diversa, comtanto que a de cada uma fique pequena em relação á combinação das outras.

Póde haver erros systematicos e a theoria fornece a fórmula geral (5) a que chamamos lei de Crofton.

Não se chega porém á lei de Gauss, sem suppôr eliminados os erros systematicos.

Todas as outras hypotheses podem ser consideradas como casos particulares da de Crofton e porisso a sua theoria abrange não só os erros compativeis com estas hypotheses mas ainda muitos outros.

Ella fixa além d'isso as condições em que é licito esperar que os erros de observação sigam approximadamente a lei de Gauss. Essas condições são:

1.ª Não existencia de erros systematicos.

2.ª Grande numero de causas de erros.

3.ª Erros elementares de uma importancia muito pequena em relação ao erro total.

Justificação da lei de Gauss.

56. Eis aqui ainda como se póde fazer uma verificação da lei de Gauss, admittindo que os erros satisfazem ás tres condições acima enunciadas.

Sejam $y_1, y_2, \ldots, y_n$ os erros elementares, $\varphi_1, \varphi_2, \ldots, \varphi_n$ as suas leis, $z = y_1 + y_2 + \ldots + y_n$ o erro total. Se z seguir a lei de Gauss, o valor provavel de z^{2p} será (capitulo II, n.º 43)

$$\overline{z^{2p}} = \frac{2p!}{p!}\left(\frac{1}{2^2 . h}\right)^p = \frac{2p!}{p!}\left(\frac{M}{2}\right)^p$$

onde $M = \frac{1}{2h} = \overline{z^2}$.

Para $p = 2, 3, \ldots$ vem

$$\overline{z^4} = 3M^2, \ \overline{z^6} = 15M^3, \ldots$$

Calculemos directamente estes valores provaveis.

Como é

$$z^2 = (y_1 + y_2 + \ldots + y_n)^2 = \Sigma y_1^2 + 2\Sigma y_1 y_2$$

será

$$\overline{z^2} = \Sigma \overline{y_1^2}$$

porque

$$\Sigma\overline{y_1}\,\overline{y_2} = \int\int \varphi_1(y_1)\varphi_2(y_2)y_1y_2dy_1dy_2$$

$$= \int \varphi_1(y_1)y_1dy_1 \int \varphi_2(y_2)y_2dy_2 = 0,$$

por serem φ_1, φ_2 funcções pares e y_1, y_2 impares.

Colloquemos

$$M = \Sigma\overline{y_1^2}.$$

Não escrevendo os termos em que entram potencias de grau impar dos erros, porque os seus valores provaveis são pelo menos approximadamente nullos, acha-se para o valor provavel de z^4

$$\overline{z^4} = \Sigma\overline{y_1^4} + 6\Sigma\overline{y_1^2}\,\overline{y_2^2}$$

ou ainda

$$\overline{z^4} = 6\Sigma\overline{y_1^2}\,\overline{y_2^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

desprezando $\Sigma\overline{y_1^4}$ em relação ao outro Σ. Para justificar este desprezo basta notar que a primeira somma contem n termos e a segunda $\frac{n(n-1)}{2}$.

Ora por outro lado é

$$M^2 = (\overline{z^2})^2 = (\Sigma\overline{y_1^2})^2 = \Sigma(\overline{y_1^2})^2 + 2\Sigma\overline{y_1^2}\,\overline{y_2^2} = 2\Sigma\overline{y_1^2}\,\overline{y_2^2}$$

o desprezo obedecendo ás mesmas razões que atráz. Pode então escrever-se

$$3M^2 = 6\Sigma\overline{y_1^2}\,\overline{y_2^2} \ldots\ldots\ldots\ldots\ldots\ldots (7).$$

Comparando (6) e (7) vem

$$\overline{z^4} = 3M^2$$

conforme á lei de Gauss.

Passemos ao calculo de $\overline{z^6}$:

$$\overline{z^6} = \Sigma\overline{y_1^6} + 15\Sigma\overline{y_1^4}\,\overline{y_2^2} + 90\Sigma\overline{y_1^2}\,\overline{y_2^2}\,\overline{y_3^2} = 90\Sigma\overline{y_1^2}\,\overline{y_2^2}\,\overline{y_3^2}\dots \quad (8)$$

approximadamente, attendendo a que o numero de termos do primeiro Σ é n, do segundo $n(n-1)$ e do terceiro $\dfrac{n(n-1)(n-2)}{2.3}$, como é facil de vêr.

Mas é

$$M^3 = (\Sigma\overline{y_1^2})^3 = \Sigma(\overline{y_1^2})^3 + 3\Sigma(\overline{y_1^2})^2\,\overline{y_2^2} + 6\Sigma\overline{y_1^2}\,\overline{y_2^2}\,\overline{y_3^2}$$

ou sensivelmente

$$M^3 = 6\Sigma\overline{y_1^2}\,\overline{y_2^2}\,\overline{y_3^2}$$

Logo

$$15\,M^3 = 90\,\Sigma\overline{y_1^2}\,\overline{y_2^2}\,\overline{y_3^2}\dots\dots\dots\dots \quad (9)$$

e pelo confronto de (8) com (9)

$$\overline{z^6} = 15\,M^3$$

conforme á lei de Gauss.

Em geral $\overline{z^{2p}}$ será sensivelmente

$$\overline{z^{2p}} = \frac{2p!}{2^p}\,\Sigma\overline{y_1^2}\,\overline{y_2^2}\dots\overline{y_n^2}\dots\dots\dots\dots \quad (10),$$

considerando apenas o Σ que tem maior numero de termos.

Ora por outro lado é

$$M^p = (\Sigma \overline{y_1^2})^p,$$

$$= p! \Sigma \overline{y_1^2}\,\overline{y_2^2} \ldots \overline{y_n^2}$$

e

$$\frac{2p!}{p!\,2^p} M^p = \frac{2p!}{2^p}\, y\overline{\Sigma_2^2}\,\overline{y_3^2} \ldots \overline{y_n^2} \ldots\ldots\ldots\ldots (11).$$

Comparando (10) e (11) vem

$$z^{2p} = \frac{2p!}{p!}\left(\frac{M}{2}\right)^p$$

formula que só a lei de Gauss dá para o valor provavel de uma potencia par do erro.

D'esta analyse conclue-se que, nas condições suppostas, o erro total não segue rigorosamente a lei de Gauss, mas não se afasta muito d'ella. É portanto uma confirmação da theoria de Crofton.

CAPITULO IV

O principio dos menores quadrados

I. O principio dos menores quadrados como consequencia da lei de Gauss

57. No Capitulo I, II, vimos como a lei dos erros entra na formação dos valores mais vantajosos em todos os casos que podem considerar-se na Theoria dos erros.

A probabilidade de cada systema de valores attribuidos ás incognitas é dada pela fórmula (7) do n.º 18, no caso mais geral definido nesse mesmo numero.

Suppondo ψ constante, a probabilidade maxima corresponde á condição

$$\Phi = \varphi_1(e_1)\varphi_2(e_2)\ldots\varphi_n(e_n) = \max.$$

Admittamos a lei de Crofton dada pela fórmula (5) do n.º 54, que póde escrever-se, designando por K_i' e K_i'' os valores provaveis da primeira e da segunda potencia do erro e_i da seguinte maneira:

$$\varphi(e_i) = \frac{1}{\sqrt{2\pi(K_i'' - K_i'^2)}} e^{-\frac{(e_i - K_i')^2}{2(K_i'' - K_i'^2)}}$$

por ser

$$K_i' = \Sigma k_i'$$

e

$$K_i'' = \Sigma k_i'' + (\Sigma k_i')^2 - \Sigma k_i'^2.$$

Com esta lei a condição referida torna-se evidentemente em

$$\sum \frac{(e_i - K_i')^2}{K_i'' - K_i'^2} = \text{min.} \ldots\ldots\ldots\ldots\ldots\ldots (1).$$

Se, porém, se suppõe que não existem erros systematicos, a lei de Crofton transforma-se na de Gauss, e a condição a que têm de satisfazer os valores mais provaveis das incognitas é, por ser neste caso $K_i' = 0$,

$$\sum \frac{e_i^2}{K_i''} = \text{min.} \ldots\ldots\ldots\ldots\ldots\ldots (2).$$

Finalmente, se os erros das observações seguirem todos a mesma lei, os denominadores das fracções que compõem o Σ são eguaes e a condição reduz-se então a

$$\Sigma e_i^2 = \text{min.} \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Os valores mais provaveis das incognitas satisfazem ao principio dos menores quadrados.

que exprime que deve ser minima a somma dos quadrados dos erros das observações. D'aqui vem o nome abreviado de *principio dos menores quadrados* ao principio expresso por (3) e ainda aos principios mais geraes dados por (2) e (1).

Qualquer d'estas condições desdobra-se em tantas equações quantas as incognitas, visto ser (n.º 18)

$$e_i = u_i - o_i$$

e

$$u_i = f_i(v_1, v_2, \ldots, v_p)$$

e termos de egualar a zero separadamente as derivadas parciaes dos primeiros membros de (1), (2) ou (3) em relação a cada uma das incognitas $v_1, v_2, \ldots, v_p$.

Entrevê-se assim a possibilidade de resolver o problema da determinação dos valores mais provaveis das incognitas. O methodo que se emprega em cuja pratica não entraremos, é conhecido pelo nome de *methodo dos menores quadrados*. Elle applica-se sómente ao caso em que as funcções u_i são lineares em ordem ás incognitas. Os outros casos reduzem-se a este commettendo desprezos.

58. Vê-se que o principio dos menores quadrados dá os valores mais provaveis das incognitas, e não os valores provaveis, e são estes os que devem considerar-se mais vantajosos.

Condições em que os valores provaveis coincidem com os mais provaveis.

Admittindo, porém, a lei de Gauss e suppondo os u, funcções lineares dos v, demonstra-se a coincidencia d'estas duas especies de valores.

Com effeito, designando por $v_1^0, v_2^0, \ldots, v_p^0$ os valores das incognitas que satisfazem á condição (2), se estes são os valores provaveis, ter-se-ha em virtude da fórmula (8) do n.° 18

$$\int_{-\infty}^{+\infty} (v_i - v_i^\circ)\Phi dv_1\, dv_2 \ldots dv_p = 0.$$

Esta egualdade verificar-se-ha se $(v_i - v_i^\circ)\Phi$ fôr uma funcção impar de $(v_i - v_i^\circ)$, para o que basta demonstrar que Φ é uma funcção par.

Ora Φ é, áparte o coefficiente constante, egual a

$$e^{-\Sigma \frac{e_i^2}{K_i''}}.$$

E como $e_i = u_i - o_i$ e u_i do primeiro grau em relação aos v vê-se que o expoente é do segundo grau em relação a estas ultimas quantidades, e que, segundo (2), deve attingir o minimo para os valores v° dos v.

Designando esse expoente por E, poderemos fazer

$$E = E_\circ + E_2$$

representando por $E_\circ$ o minimo de E, constante, e por E_2 uma funcção de

$$(v_1 - v_1^\circ), (v_2 - v_2^\circ), \ldots, (v_p - v_p^\circ)$$

que deve ser essencialmente positiva, do segundo grau e annullar-se para

$$v_1 = v_1^\circ,\; v_2 = v_2^\circ, \ldots, v_p = v_p^\circ;$$

logo será E uma funcção homogenea e do segundo grau em relação aos $(v - v^\circ)$, como se pretendia demonstrar.

O mesmo succederá approximadamente se, não sendo os u funcções lineares das incognitas, todavia estas têm um campo de variação muito restricto, como nota Poincaré, a quem se deve a demonstração precedente.

II. Demonstração do principio dos menores quadrados, prescindindo da lei dos erros

Analyse de Yarochenko.

59. Supponhamos agora que, prescindindo da lei dos erros, se pretende determinar os valores mais vantajosos das incognitas.

Sejam os u funccões lineares dos v, de sorte que se tenham as n equações.

$$a_i v_1 + b_i v_2 + c_i v_3 + \ldots - o_i = e_i \qquad (i = 1, 2, 3, \ldots n).$$

Admittamos que *os valores de e_i eguaes e de signaes contraríos sejam egualmente provaveis,* isto é, que cada erro e_i siga uma lei par.

Com esta unica hypothese vamos demonstrar o principio dos menores quadrados, seguindo nos seus traços geraes uma elegante analyse de Yarochencko, professor da Universidade de Odessa, fundada no theorema seguinte de Tchébychef:

Designando por $a, b, c, \ldots$ os valores provaveis das quantidades $x, y, z, \ldots$ e por $a_1, b_1, c_1, \ldots$ os valores provaveis dos seus quadrados $x^2, y^2, z^2, \ldots$, a probabilidade de que a somma $x + y + z + \ldots$ esteja comprehendida entre os limites

$$a + b + c + \ldots + \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots}$$

e

$$a + b + c + \ldots - \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots}$$

será sempre maior do que $1-\frac{1}{\alpha^2}$, qualquer que seja α (*).

Multiplique-se cada uma das equações dadas por um certo multiplicador λ_i e sommem-se os resultados; effectuem-se as mesmas operações com um segundo systema de multiplicadores λ_i; e assim por diante.

Ter-se-ha

$$\left.\begin{array}{l} v_1=\Sigma\lambda\, o+\Sigma\lambda\, e \\ v_2=\Sigma\lambda' o+\Sigma\lambda' e \\ v_3=\Sigma\lambda'' o+\Sigma\lambda'' e \\ \ldots\ldots\ldots\ldots\ldots \end{array}\right\}\ldots\ldots\ldots\ldots\ldots\ldots \quad (4),$$

se os systemas de multiplicadores são escolhidos de modo a satisfazer ás equações

$$\left.\begin{array}{l} \Sigma\lambda\, a=1,\ \ \Sigma\lambda\, b=0,\ \ \Sigma\lambda\, c=0,\ldots \\ \Sigma\lambda' a=1,\ \ \Sigma\lambda' b=0,\ \ \Sigma\lambda' c=0,\ldots \\ \Sigma\lambda'' a=1,\ \ \Sigma\lambda'' b=0,\ \ \Sigma\lambda'' c=0,\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\}\ldots\ldots\ldots \quad (5).$$

Como cada systema de multiplicadores contém n e o numero das equações (5) em cada linha é egual ao numero das incognitas v e portanto menor do que n, os λ não ficam determinados e por isso podemos dispôr d'essa indeterminação de maneira que os valores mais vantajosos sejam

$$v_1=\Sigma\lambda o,\ \ v_2=\Sigma\lambda' o,\ \ v_3=\Sigma\lambda'' o,\ldots$$

Para isso procuremos os limites dos erros $\Sigma\lambda e$, $\Sigma\lambda' e$, $\Sigma\lambda'' e,\ldots$ d'estes valores.

(*) Para a demonstração d'este theorema vid. *Journ. de Liouville,* 2^e^ série, t. XII, p. 177, 1867.

Sejam P_i', P_i'' os valores provaveis de $\lambda_i e_i$ e $\lambda_i^2 e_i^2$.

O Theorema citado de Tchébychef diz-nos que a probabilidade de $\Sigma \lambda e$ estar encerrado entre os limites

$$\Sigma P' + \frac{\sqrt{n}}{t} \sqrt{\Sigma P'' - \Sigma P'^2} \quad \text{e} \quad \Sigma P' - \frac{\sqrt{n}}{t} \sqrt{\Sigma P'' - \Sigma P'^2}$$

será sempre maior do que $1 - \frac{t^2}{n}$.

Mas em virtude da hypothese d'onde partimos é evidente que

$$P_i' = 0.$$

Além d'isso, designando como atraz por K_i'' o valor provavel de e_i^2, será

$$P_i'' = \lambda_i^2 K_i''.$$

Logo os limites procurados são

$$\frac{1}{t}\sqrt{n \Sigma K'' \lambda^2} \quad \text{e} \quad \frac{1}{t}\sqrt{n \Sigma K'' \lambda^2},$$

e analogamente para as outras incognitas.

Estes limites serão os menores possiveis se escolhermos os λ de maneira que

$$\Sigma K'' \lambda^2 = \text{min.}, \quad \Sigma K'' \lambda'^2 = \text{min.}, \quad \Sigma K'' \lambda''^2 = \text{min.}, \ldots \quad (6)$$

e os valores das incognitas correspondentes a estes valores dos λ serão os mais vantajosos. Para cada probabilidade $1 - \frac{t^2}{n}$ dada commettem-se com estes valores os menores erros possiveis.

Vamos agora vêr que as condições (6), junctas ás (5), equivalem a (2) ou ao principio dos menores quadrados.

Com effeito, consideremos o primeiro systema dos λ, que deve satisfazer ás seguintes condições:

$$\Sigma K''\lambda^2 = \text{min.}, \quad \Sigma\lambda a - 1 = 0, \quad \Sigma\lambda b = 0, \quad \Sigma\lambda c = 0, \ldots. \quad (7).$$

Multipliquem-se estas equações á excepção da primeira por factores indeterminados $2q_1$, $2q_2$, $2q_3$, ...

Podemos substituil-as todas pela condição unica:

$$\Sigma K''\lambda^2 - 2q_1(\Sigma\lambda a - 1) - 2q_2\,\Sigma\lambda b - 2q_3\,\Sigma\lambda c - \ldots = \text{min.};$$

que se desdobra nas seguintes, por derivação:

$$\left.\begin{array}{l} K_1''\lambda_1 = q_1 a_1 + q_2 b_1 + q_3 c_1 + \ldots \\ K_2''\lambda_2 = q_1 a_2 + q_2 b_2 + q_3 c_2 + \ldots \\ K_3''\lambda_3 = q_1 a_3 + q_2 b_3 + q_3 c_3 + \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \ldots\ldots \quad (8).$$

Multipliquemos agora estas equações por $\frac{a_1}{K_1''}, \frac{a_2}{K_2''}, \frac{a_3}{K_3''}, \ldots$ e sommemos; depois por $\frac{b_1}{K''_2}, \frac{b_2}{K''_2}, \ldots$ e sommemos de novo; e assim por diante. Tendo em attenção (7), virá

$$\left.\begin{array}{l} 1 = q_1\Sigma\frac{aa}{K''} + q_2\Sigma\frac{ab}{K''} + q_3\Sigma\frac{ac}{K''} + \ldots \\ 0 = q_1\Sigma\frac{ba}{K''} + q_2\Sigma\frac{bb}{K''} + q_3\Sigma\frac{bc}{K''} + \ldots \\ 0 = q_1\Sigma\frac{ca}{K''} + q_2\Sigma\frac{cb}{K''} + q_3\Sigma\frac{cc}{K''} + \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \ldots\ldots \quad (9).$$

Resolvendo este systema, vem

$$q_1 = \frac{A_1}{R},\ q_2 = \frac{A_2}{R},\ q_3 = \frac{A_3}{R}, \ldots\ldots\ldots\ldots,$$

designando por R o determinante dos coefficientes das incognitas e por A_1, A_2, $A_3\ldots$ os menores d'este determinante relativos aos elementos da 1.ª linha.

Substituindo em (8) estes valores, vem

$$\left.\begin{array}{l} R\lambda_1 = \dfrac{a_1}{K_1''} A_1 + \dfrac{b_1}{K_1''} B_1 + \dfrac{c_1}{K_1''} C_1 + \ldots \\[2ex] R\lambda_2 = \dfrac{a_2}{K_2''} A_2 + \dfrac{b_2}{K_2''} B_2 + \dfrac{c_2}{K_2''} C_2 + \ldots \\[2ex] R\lambda_3 = \dfrac{a_3}{K_3''} A_3 + \dfrac{b_3}{K_3''} B_3 + \dfrac{c_3}{K_3''} C_3 + \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \ldots\ldots (10).$$

Multiplicando agora estas equações por o_1, o_2, $o_3\ldots$ e sommando, vem a 1.ª das seguintes

$$\left.\begin{array}{l} Rv_1 = A_1 \Sigma \dfrac{ao}{K''} + A_2 \Sigma \dfrac{bo}{K''} + A_3 \Sigma \dfrac{co}{K''} + \ldots \\[2ex] Rv_2 = B_1 \Sigma \dfrac{ao}{K''} + B_2 \Sigma \dfrac{bo}{K''} + B_3 \Sigma \dfrac{co}{K''} + \ldots \\[2ex] Rv_3 = C_1 \Sigma \dfrac{ao}{K''} + C_2 \Sigma \dfrac{bo}{K''} + C_3 \Sigma \dfrac{co}{K''} + \ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{array}\right\} \ldots\ldots (11),$$

obtendo-se as restantes por uma analyse semelhante applicada ás incognitas v_2, v_3, $\ldots$, e em que B_1, B_2, $B_3\ldots$ são os menores de R em relação aos elementos da 2.ª linha; C_1, C_2, C_3,..., em relação aos elementos da 3.ª linha,... etc.

As equações (11) mostram que podem considerar-se as incognitas como sendo soluções do systema seguinte:

$$\left.\begin{array}{l}
v_1 \Sigma \dfrac{aa}{K''} + v_2 \Sigma \dfrac{ab}{K''} + v_3 \Sigma \dfrac{ac}{K''} + \ldots = \Sigma \dfrac{ao}{K''} \\[2ex]
v_1 \Sigma \dfrac{ba}{K''} + v_2 \Sigma \dfrac{bb}{K''} + v_3 \Sigma \dfrac{bc}{K''} + \ldots = \Sigma \dfrac{bo}{K''} \\[2ex]
v_1 \Sigma \dfrac{ca}{K''} + v_2 \Sigma \dfrac{cb}{K''} + v_3 \Sigma \dfrac{cc}{K''} + \ldots = \Sigma \dfrac{co}{K''} \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{array}\right\} \ldots\ (12),$$

attendendo a que os 2.ºˢ membros de (11) são os resultados da substituição no determinante symetrico R dos elementos da 1.ª, 2.ª, 3.ª,... columnas ou da 1.ª, 2.ª, 3.ª,... linhas pelos 2.ºˢ membros de (12).

Ora o systhema (12) é equivalente áquelle em que se desdobra a condição (2) pela derivação parcial em ordem a cada uma das incognitas. Com effeito, obtem-se d'este modo as chamadas equações *normaes:*

$$\left.\begin{array}{l}
a_1 \dfrac{e_1}{K_1} + a_2 \dfrac{e_2}{K_2} + a_3 \dfrac{e_3}{K_3} + \ldots = 0 \\[2ex]
b_1 \dfrac{e_1}{K_1} + b_2 \dfrac{e_2}{K_2} + b_3 \dfrac{e_3}{K_3} + \ldots = 0 \\[2ex]
c_1 \dfrac{e_2}{K_1} + c_2 \dfrac{e_2}{K_2} + c_3 \dfrac{e_3}{K_3} + \ldots = 0 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots
\end{array}\right\} \ldots\ldots\ldots\ (13),$$

que se transformam em (12) substituindo e_1, e_2, e_3... pelas suas expressões tiradas das equações dadas.

Observações de egual precisão.

60. Notemos que esta condição simplifica-se quando se suppõe que os erros seguem todos a mesma lei e estão comprehendidos entre os mesmos limites. Porque n'esse caso será

$$K_1'' = K_2'' = K_3'' = \ldots$$

e as condições (6) reduzem-se a

$$\Sigma \lambda^2 = \text{min.}, \ \Sigma \lambda'^2 = \text{min.}, \ \Sigma \lambda''^2 = \text{min.}, \ldots,$$

que, como o mostra uma analyse semelhante á do n.º antecedente, são equivalentes a

$$\Sigma e^2 = \text{min.}$$

É o que succede no caso das observações de egual precisão, em que deve attribuir-se egual probabilidade á supposição de qualquer dos erros $e_1, e_2, e_3 \ldots$ ter um valor determinado.

Finalmente no caso das observações directas egualmente precisas as equações dadas serão de fórma

$$v - o_i = e_i$$

e a ultima condição dará

$$v = \frac{\Sigma o_i}{n},$$

isto é, a media arithmetica das observações.

Existencia de erros systematicos.

61. Os raciocinios feitos nos dois numeros precedentes partem da supposição essencial de que os erros de cada observação com o mesmo valor absoluto sejam egualmente provaveis, isto é, que não existam erros systematicos.

Notamos já no capitulo I, n.º 6, que a imperfeição dos instrumentos de que o observador usa não lhe permitte ir alem de uma certa ordem de unidades na avaliação de uma grandeza pela observação.

O erro verdadeiro e_i da observação o_i compôr-se-ha de um numero inteiro N_i de unidades u d'esta ordem e mais uma fracção que representaremos por I_i. Será

$$e_i = N_i u + I_i, \qquad [I_i < u].$$

A parcella I_i póde considerar-se sempre positiva; a outra parcella $N_i u$ será positiva ou negativa conforme o erro verdadeiro é maior ou menor do que I_i. Esta ultima quantidade é, pois, uma constante; é um erro que ha de commetter-se sempre, cuja probabilidade é egual á unidade, e por isso se chama *inevitavel*.

Suppondo desviados os erros systematicos, a parte $E_i = N_i u$ do erro será accidental e portanto poderá sujeitar-se á hypothese fundamental do n.º 59.

Seria agora facil de refazer a analyse d'esse mesmo numero e de chegar á condição

$$\sum \frac{(e-I)^2}{\chi''} = \sum \frac{E^2}{\chi''} = \text{min.} \ldots\ldots\ldots\ldots (14)$$

onde χ_i'' é o valor provavel do quadrado do erro E_i. Basta notar que o valor de v_1, por exemplo, terá agora a fórma

$$v_1 = \Sigma\lambda o + \Sigma\lambda I + \Sigma\lambda E.$$

Os limites provaveis de $\Sigma\lambda E$ serão em virtude do theorema de Tchébychef e de considerações perfeitamente semelhantes ás que se fizeram para $\Sigma\lambda e$ no citado numero,

$$\frac{1}{t}\sqrt{n\Sigma\chi''\lambda^2} \quad \text{e} \quad -\frac{1}{t}\sqrt{n\Sigma\chi''\lambda^2}$$

com uma probabilidade maior do que $1-\frac{t^2}{n}$.

Esses limites serão os mais estreitos possivel se

$$\Sigma\chi''\lambda^2 = \text{min.},$$

condição que equivale a (14).

O valor mais vantajoso de v_1 é, pois,

$$v_1 = \Sigma\lambda o + \Sigma\lambda \mathrm{I}.$$

Resta-nos mostrar a identidade entre as condições (14) e (1).

Para isso vamos primeiro demonstrar, com Yarochencko, que o erro inevitavel de uma observação é egual ao valor provavel do erro verdadeiro d'esta observação.

Seja a observação v_i, $e_i = \mathrm{E}_i + \mathrm{I}_i$ o erro verdadeiro, o valor K_i' provavel d'este erro será, designando por $\mathrm{P}_{i,0}, \mathrm{P}_{i,1}, \ldots, \mathrm{P}_{i,n}$ as probabilidades dos differentes valores possiveis de E_i,

$$\begin{aligned}\mathrm{K}_i' &= \mathrm{I}_i(\mathrm{P}_{i,0} + 2\,\mathrm{P}_{i,1} + \ldots + 2\,\mathrm{P}_{i,n}) + \chi'_i \\ &= \mathrm{I}_i.\end{aligned}$$

Com effeito, é, por ser E_i um erro accidental, $\chi_i' = 0$; e por outro lado

$$\mathrm{P}_{i,0} + 2\,\mathrm{P}_{i,}1 + \ldots + 2\,\mathrm{P}_{i,n} = 1,$$

porque a somma das probabilidades de todas as supposições possiveis sobre o valor de uma grandeza equivale á certeza.

Procuremos agora K_i'', valor provavel de e_i^2. É

$$K_i'' = \overline{(E_i + I_i)^2}$$
$$= \overline{E_i^2} + \overline{2\,E_i I_i} + \overline{I_i^2}.$$

$\overline{E_i^2}$ é a quantidade que nós designamos acima por χ_i''. O segundo termo $\overline{2\,E_i I_i}$ é nullo, porque

$$\overline{2\,E_i I_i} = 2\,\overline{I_i}\,.\,\overline{E_i} = 2K_i'\chi_i' = 0\,.$$

Finalmente

$$\overline{I_i^2} = I_i^2 = K_i'^2\,.$$

Logo

$$K_i'' = \chi_i'' + K_i'^2\,;$$

e a condição (14) transforma-se em

$$\sum \frac{(e - K')^2}{K'' - K'^2} = \text{min.},$$

como se pretendia demonstrar.

O raciocinio feito applica-se ao caso mais geral de existirem outros erros systematicos alem do erro inevitavel; K_i' seria então egual á parte systematica do erro.

62. Em resumo: Se o systema de equações dado fôr linear, quaesquer que sejam as leis dos erros os valores mais vantajosos das incognitas são os que satisfazem á condição (1), correspondente á lei de Crofton. Não existindo erros systematicos, a condição simplifica-se reduzindo-se a (2), que é uma consequencia da lei de Gauss. Nota final.

Como, porém, existe sempre, pelo menos, o erro inevitavel a lei de Gauss não será applicavel com rigor, mas será tanto mais approximada quanto menor valor tiver a quantidade K_i'.

A demonstração de Yarochencko, que é uma simplicação da analyse de Laplace, que aqui não podemos expôr, tem, como esta, o defeito de exigir que as equações finaes que dão os valores mais vantajosos sejam uma combinação linear das equações dadas. Alem d'isso estas são tambem, por hypothese, lineares.

Em rigor o que fica demonstrado é que, no caso das equações propostas serem lineares, d'entre os systhemas de equações tambem lineares que se podem substituir ao dado, é mais vantajoso aquelle que resulta da condição (1), havendo erros systematicos ou da condição (2), não os havendo.

Esta solução, que deixa muito a desejar sob o ponto de vista theorico, satisfaz praticamente, porque os calculos tornar-se-hiam extremamente complicados se fosse preciso recorrer a combinações não lineares; e basta para isso lembrar que a applicação do methodo dos menores quadrados é já muito trabalhosa.

No caso das observações de egual precisão os calculos são muito menos laboriosos e mal se póde dar a razão de preferencia que acabamos de apontar. Fazer a restricção de que as equações finaes sejam lineares equivale a (vid. n.º 25) adoptar o raciocinio de Ellis. Mas se as observações são sensivelmente concordantes nós vimos no capitulo II, I que a media arithmetica pouco póde differir dos outros valores que é admissivel attribuir á incognita; vimos alem d'isso que não se conseguiu ainda provar que qualquer d'estes valores seja preferivel á media arithmetica, embora não tenha tambem sido possivel demonstrar a preferencia d'esta sobre todos aquelles. Emquanto, pois, esta ignorancia persistir, a media offerece sobre as restantes soluções a vantagem de simplicidade e por isso é geralmente adoptada.

Mas devemos observar, e neste ponto estamos de accordo com Estienne, que a media arithmetica não póde nunca eliminar os erros systematicos e em particular o erro inevitavel, que, affectando todos os resulfados das observações de egual modo, se conserva integralmente na media.

Pelo que respeita aos outros casos mais complexos a analyse de Yarochencko fornece uma justificação muito notavel do methodo dos menores quadrados.

Sob o ponto de vista da constituição da Theoria dos erros como sciencia, é todavia preferivel a primeira demonstração que demos do principio dos menores quadrados.

D'este modo a Theoria dos erros apparece como um ramo do *Calculo das probabilidades,* definem-se os valores mais vantajosos das incognitas, determina-se a lei dos erros de que elles dependem, e o principio dos menores quadrados resulta como uma nova maneira de definir aquelles valores, que conduz directamente ao methodo dos menores quadrados.

Ao contrario no processo de Yarochencko não se dá uma definição prévia dos valores mais vantajosos das incognitas; impõe-se uma determinada combinação das equações propostas e d'entre as soluções possiveis a que essa marcha conduz escolhe-se então a mais conveniente.

É preciso, porém, accrescentar que, não sendo rigorosa a deducção analytica da lei dos erros, só a verificação experimental poderá decidir definitivamente se a approximação é sufficiente.

FIM.

Fig. 1

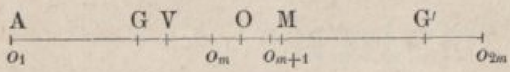


Fig. 2

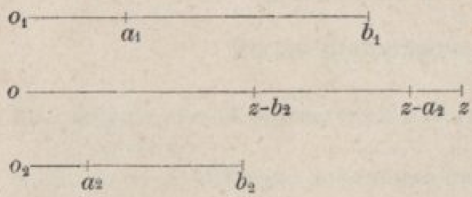
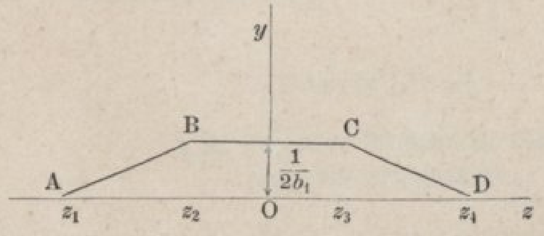


Fig. 3



INDICE

CAPITULO IV

O principio dos menores quadrados

1898

S. PAES. DISSERTAÇÃO INAUGURAL

MATHEMATICA