



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

HISTÒRIA DE LA PROBABILITAT AL SEGLE XVIII: DE BERNOULLI A LAPLACE

Autora: Sara Antentas Oliveras

Director: Dr. Josep Vives Santa Eulàlia

Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 10 de juny de 2024

Abstract

Probability theory and mathematical statistics form the foundation of many areas that govern our day to day lives. It was during the XVIII century when the statistical framework would be assembled, and this event took place thanks to some of the greatest mathematics such as Jacob Bernoulli, De Moivre, Laplace, and Gauss, among many others. These intellectuals unified and shared knowledge in order to nurture themselves and develop the origins of probability and contemporary statistics.

Resum

La teoria de la probabilitat i l'estadística matemàtica conformen el fonament de molts àmbits que regeixen el nostre dia a dia. Fou durant el segle XVIII que es desenvoluparen les bases sobre les quals es forjaria l'estadística tal i com la coneixem actualment, i aquest esdeveniment tingué lloc gràcies a alguns il·lustres matemàtics com serien Jacob Bernoulli, De Moivre, Laplace i Gauss, entre molts d'altres. Aquests van unificar i compartir coneixement per tal de nodrir-se i desenvolupar els orígens de la probabilitat i l'estadística contemporània.

Agraïments

M'agradaria agrair el suport incondicional de la meva família i els meus amics i companys de classe al llarg d'aquests anys, ja que és gràcies a tots ells que s'ha fet possible que avui escrigui aquest treball de final de grau. Vull fer una menció especial al meu tutor de TFG, el Dr. Josep Vives Santa Eulàlia, ja que m'ha guiat i ajudat al llarg d'aquest camí, proporcionant-me les eines bibliogràfiques i tècniques necessàries per tal de poder fonamentar aquest treball. La confiança i el suport que m'ha proporcionat en tot moment ha sigut imprescindible per la realització del mateix. Als meus pares i al meu germà: sou el meu pilar fonamental, ja que sempre m'heu donat el suport necessari per animar-me a seguir lluitant. Als meus amics del doble grau de Matemàtiques i Física: sense vosaltres aquest camí hauria sigut molt més llarg i tediós. A la Clara, la Mima i la Marta: la vostra amistat és el millor que m'emporto de la meva etapa universitària. A la Universitat de Barcelona: per obrir les portes als joves a emprendre uns estudis universitaris i créixer com a persones en unes boniques facultats en aquesta pròspera i viva ciutat de Barcelona.

Índex

1	Introducció	1
1.1	El projecte	1
1.2	Estructura de la Memòria	2
2	La quantificació de la incertesa	3
2.1	Jacob Bernoulli	3
2.1.1	<i>Ars Conjectandi</i>	3
2.2	James Stirling	6
2.2.1	Fórmula de Stirling	7
2.3	Abraham De Moivre	8
2.3.1	<i>Doctrine of Chances</i>	9
2.3.2	Teorema de De Moivre-Laplace	15
2.4	Thomas Simpson	19
2.4.1	Mitjana d'un conjunt d'observacions a l'astronomia	20
3	La probabilitat inversa	21
3.1	Pierre-Simon Laplace	21
3.1.1	Probabilitat Inversa	21
3.1.2	Les memòries de Laplace de 1774	24
3.1.3	La corba d'error de 1772-1774	26
3.1.4	L'origen de la probabilitat inversa	28
3.1.5	Les memòries de Laplace de 1777-1781	31
3.2	Thomas Bayes	32
3.2.1	La “taula de billar” de Bayes	32
3.2.2	L'anàlisi de Laplace	36
3.2.3	El Teorema Central del Límit	38
4	La síntesi de Gauss-Laplace	39
4.1	Carl Friedrich Gauss	40
4.1.1	Gauss el 1809	40
4.1.2	Exemple del mètode de mínims quadrats	45
4.2	Laplace i les mareas de l'atmosfera	46
4.3	La situació el 1827	50
5	Conclusions	51

1 Introducció

1.1 El projecte

Per poder conèixer una ciència en la seva complexitat és essencial entendre la història de la mateixa. Per aquesta raó, amb aquest treball bibliogràfic s'ha volgut indagar en els orígens de l'estadística i la probabilitat, tot establint un fil conductor amb la llei dels grans nombres i el teorema central del límit, dos resultats que són de gran importància i que constitueixen els fonaments de l'estadística actual. També és interessant prendre consciència de les dificultats a les quals es van afrontar els seus autors en el moment de desenvolupar aquests brillants resultats, tot considerant-los en el seu context històric. Així doncs, per comprendre les matemàtiques contemporànies, és interessant remuntar-nos als inicis d'aquesta bonica ciència.

Al llarg del segle XVIII, la teoria de la probabilitat evoluciona de forma significativa, passant d'uns primers raonaments fonamentats en el joc i l'atzar fins a una disciplina matemàtica més formal. En la realització d'aquest treball s'han estudiat diverses figures que han sigut cabdals en aquest camí, com serien Jacob Bernoulli, Abraham De Moivre, Pierre-Simon Laplace i Carl Friedrich Gauss, entre d'altres. Aquests matemàtics destaquen per les seves contribucions innovadores, les quals estableixen les bases de la teoria de la probabilitat moderna. Al llarg d'aquest segle, la teoria de la probabilitat es comença a veure com una eina robusta per fer front a la incertesa i fer prediccions en diversos camps científics. Les aportacions d'aquests i d'altres matemàtics al segle XVIII estableixen les bases per a la formalització de la teoria de la probabilitat i l'estadística als segles XIX i XX, transformant-la en un component fonamental de les matemàtiques i la ciència moderna.

Aquest TFG té per objectiu seguir una línia temporal al llarg del segle XVIII, juntament amb alguns resultats i obres publicades a inicis del segle XIX. Destaquen alguns fets comuns d'entre tots els autors tractats, com seria el desig de trobar explicacions i fonaments teòrics per donar resposta a molts esdeveniments en l'àmbit de l'astronomia, la qual fou una ciència elemental per despertar en alguns matemàtics les ànsies d'indagar en l'estadística i la probabilitat. Al llarg del treball es mostren alguns exemples relacionats amb càlculs astronòmics i intents per donar resposta a algunes preguntes en aquest àmbit. També s'exemplifiquen altres resultats per fer més entenedora l'explicació dels raonaments presentats. El punt de partida del treball és l'extraordinària obra de Jacob Bernoulli, *Ars Conjectandi* [2], publicada el 1713, seguint amb altres llibres de gran transcendència, com serien *The Doctrine of Chances* [4], d'Abraham De Moivre, la primera publicació del qual fou el 1718, o bé el meravellós treball de Pierre-Simon Laplace, *Théorie Analytique des Probabilités* [17], del 1812. També és destacable l'obra de Carl Friedrich Gauss de 1809, *Theoria motus corporum coelestium* [9]. Cal remarcar que el TFG s'ha fonamentat amb un conjunt de llibres sobre probabilitat i estadística, els quals presenten diferents enfocaments i resultats matemàtics que han permès establir un fil conductor en el qual es presenten els resultats de forma cronològica i conceptualment coherent, tenint en compte que l'autora d'aquest treball és una estudiant de l'últim curs en matemàtiques i en física. S'ha intentat recopilar informació variada i contrastada de les diferents fonts, les quals estan mencionades a la bibliografia. A més, s'ha complementat la informació amb llibres d'autors de renom, com serien Chung [3] i Stigler [26], a part d'algunes memòries i documents publicades pels propis autors que apareixen al llarg del treball, com serien Thomas Bayes, Pierre-Simon Laplace, Joseph Louis Lagrange i Thomas Simpson, entre d'altres.

1.2 Estructura de la Memòria

La memòria s'ha estructurat en tres parts diferenciades i ordenades cronològicament. Primerament, s'ha tractat la quantificació de la incertesa, seguida de l'estudi de la probabilitat inversa, i finalitzant amb la síntesi de Gauss-Laplace. Dins de cada una d'aquestes seccions s'han presentat els autors més transcendents i els resultats matemàtics més destacables de cada àmbit. En tot moment s'ha intentat mantenir un caràcter formal i coherent pel que fa als desenvolupaments matemàtics, de manera que en algunes ocasions s'ha considerat oportú fer esment als resultats de l'època escrits en notació moderna, fet que facilita la comprensió dels passos i les idees matemàtiques més rellevants. Així doncs, la memòria segueix un fil conductor amb la llei dels grans nombres i el teorema central del límit, tot posant èmfasi en la trajectòria històrica dels autors i els seus resultats més transcendents al llarg del segle XVIII.

2 La quantificació de la incertesa

2.1 Jacob Bernoulli

L'origen del desenvolupament matemàtic en l'àmbit de la probabilitat es troba en els jocs d'atzar, a partir dels quals es forja la primera quantificació de la incertesa.

Jacob Bernoulli (Basilea, 1654 - 1705) és conegut com el pare de la quantificació de la incertesa. Fou professor a la Universitat de Basilea des de 1687, contemporani i ocasionalment rival de Isaac Newton. La rebel·lió contra els desitjos del seu pare el portaren a estudiar matemàtiques i astronomia. Va viatjar per Europa des de 1676 fins el 1683. Amb el seu retorn a Suïssa, va començar a treballar de professor de mecànica a la Universitat de Basilea, i el 1684 va presentar la seva tesi doctoral *Solutionem tergemini problematis*. El 1687 va ser nominat catedràtic de matemàtiques a la Universitat de Basilea.

Jacob Bernoulli i el seu germà John, reconeguts com a grans matemàtics, l'any 1684 van desenvolupar el càlcul diferencial a partir de diverses indicacions i solucions publicades per Leibniz. Posteriorment treballaren en el càlcul integral. A més, l'avenç dels seus estudis de corbes i alguns problemes de minimització acabarien evolucionant cap al càlcul de variacions per part d'Euler i Lagrange. Generalment, els germans Bernoulli eren rivals, fet que explica la demora en la publicació dels documents pòstums de Jacob Bernoulli, els quals es van endarrerir fins a vuit anys després de la seva mort, i que finalment foren editats i publicats pel seu nebot Nicholas.

Algunes de les moltes contribucions de Jacob Bernoulli van ser el descobriment de la constant matemàtica fonamental e (1683); el paral·lelisme entre la lògica i l'àlgebra (document publicat el 1685); un treball sobre probabilitat del 1685 i un de geometria del 1687. El 1689 va presentar la seva "lleï dels grans nombres" en teoria de probabilitats. Entre 1682 i 1704 publicà cinc tractats sobre sèries infinites. El 1690 va demostrar que el problema de determinar la isòcrona (corba de descens constant) era equivalent a resoldre una equació diferencial no lineal de primer ordre, la qual va resoldre amb el que ara anomenem "separació de variables". Aquest document de 1690 fou important ja que va aparèixer per primera vegada el terme "integral" amb el seu significat d'integració. El 1696, Bernoulli va resoldre l'actualment coneguda "equació diferencial de Bernoulli".

La mort del matemàtic va deixar un llegat de treballs sense publicar. Bernoulli va treballar durant vint anys amb problemes sobre la determinació de la probabilitat, i els fruits d'aquest treball es recopilaren en el tractat publicat per Nicholas l'any 1713, conegut com *Ars Conjectandi* [2] (del Llatí, "*L'art de conjecturar*"). Aquesta obra s'ha considerat com la precursora de la teoria de la probabilitat, remarcable en diversos aspectes, des dels seus importants avenços en combinatòria fins a la innovadora anàlisi de la interpretació de l'evidència.

2.1.1 *Ars Conjectandi*

De l'obra *Ars Conjectandi* [2], ens centrem en la quarta part, la qual emmarca la primera lleï dels grans nombres (1713). Aquesta consisteix en un teorema segons el qual, si el nombre d'observacions d'un fenomen aleatori és molt gran, aleshores la freqüència relativa d'un esdeveniment convergeix de forma quasi segura a la probabilitat de l'esdeveniment. Dit d'una altra manera, la probabilitat empírica d'un succés (la freqüència relativa de

vegades que s'ha realitzat l'esdeveniment en una sèrie llarga de repeticions) convergeix a la probabilitat teòrica.

Per tal d'arribar al desenvolupament d'aquest il·lustre teorema, Bernoulli nota que, a priori, es pot determinar directament la probabilitat en els jocs en què s'usen daus homogenis o bé urnes amb tiquets igualment accessibles de diferents colors, ja que únicament s'han d'enumerar els casos possibles o “fèrtils” i prendre la relació entre aquests i el nombre total de casos, ja siguin “fèrtils” o “estèrils”. No obstant, la dificultat rau en aquells problemes en els quals no és possible l'enumeració de casos igualment probables.

Per determinar la proporció de casos favorables o fèrtils es pot fer de forma empírica. No obstant, l'objectiu del matemàtic és intentar demostrar que, com major és el número d'evidències sobre la proporció desconeguda de casos, més a prop estem de conèixer aquesta última; és a dir, la incertesa decreix a mesura que augmenta el nombre d'observacions.

Un exemple seria considerar 3000 pedretes blanques i 2000 pedretes negres dins una urna. Per determinar-ne el número es treu una pedreta rere l'altra, retornant la pedreta triada abans d'escollir la següent, i observant la freqüència amb la qual se n'obté una de blanca o negra. L'objectiu és determinar si es pot realitzar aquest procediment prou vegades per tal que la probabilitat d'escollir pedretes blanques i negres augmenti suficientment per obtenir la mateixa proporció 3 : 2 que les pedretes de la urna.

A continuació, es presenta un replantejament del problema i una solució del mateix en notació moderna. Sigui X el nombre de casos favorables o fèrtils, N el número total d'observacions i p la proporció desconeguda. Aleshores, per qualsevol nombre positiu petit donat ϵ i qualsevol nombre positiu gran donat c (prenem $c = 10$, 100 o 1000), N ha de satisfer

$$P\left(\left|\frac{X}{N} - p\right| \leq \epsilon\right) > cP\left(\left|\frac{X}{N} - p\right| > \epsilon\right). \quad (2.1)$$

Aquesta afirmació es pot convertir en l'actual llei feble dels grans nombres de Bernoulli:

$$P\left(\left|\frac{X}{N} - p\right| > \epsilon\right) < \frac{1}{(c+1)}. \quad (2.2)$$

Com que c és arbitrària, donat qualsevol $\epsilon > 0$ es pot determinar una N suficientment gran tal que es compleix (2.2), i així queda demostrada la llei de Bernoulli. Aquest resultat és inacurat en alguns aspectes, així com també ho és la concessió a Bernoulli del primer exemple d'estimació d'un interval de probabilitat. En notació moderna, la prova de (2.1) consisteix en el desenvolupament que segueix. Es comença reescriuint l'equació,

$$P(|X - Np| \leq N\epsilon) > cP(|X - Np| > N\epsilon).$$

Bernoulli pren N com a múltiple de $t = r + s$ ($N = nt$), on $Np = nr$ i $N\epsilon = n$ són enters, i porcions d'igual longitud $N\epsilon$ de X . Centrant-se primer en els valors de X superiors a Np , defineix

$$\begin{aligned} A_0 &= P(Np < X \leq Np + N\epsilon) \\ A_1 &= P(Np + N\epsilon < X \leq Np + 2N\epsilon) \\ A_2 &= P(Np + 2N\epsilon < X \leq Np + 3N\epsilon) \\ &\text{etc.} \end{aligned}$$

L'objectiu és demostrar que, donat c , es pot trobar $N_0(c)$ tal que per $N \geq N_0(c)$, aleshores

$$A_0 > c(A_1 + A_2 + \dots)$$

De forma semblant, pels valors de X inferiors a Np es pot trobar una desigualtat equivalent.

Així doncs, sigui η dependent únicament de N i p , aleshores es compleix

$$A_{k+1} < \eta A_k \quad (2.3)$$

per $k = 0, 1, \dots$

Cal demostrar (2.3). Sigui

$$b_k = Np + kN\epsilon + 1 = n(r + k) + 1$$

aleshores

$$\begin{aligned} \frac{A_{k+1}}{A_k} &= \frac{P(X = b_{k+1}) + P(X = b_{k+1} + 1) + \dots + P(X = b_{k+1} + N\epsilon - 1)}{P(X = b_k) + P(X = b_k + 1) + \dots + P(X = b_k + N\epsilon - 1)} \\ &< \max \left\{ \frac{P(X = b_{k+1})}{P(X = b_k)}, \dots, \frac{P(X = b_{k+1} + N\epsilon - 1)}{P(X = b_k + N\epsilon - 1)} \right\}. \end{aligned}$$

Sabent que per $F, G, H, \dots, P, Q, R, \dots$, positius es compleix (amb igualtat únicament quan totes les fraccions són iguals)

$$\frac{F + G + H + \dots}{P + Q + R + \dots} \leq \max \left\{ \frac{F}{P}, \frac{G}{Q}, \frac{H}{R}, \dots \right\}.$$

Amb una anàlisi detallada de la probabilitat, Bernoulli demostra que:

$$\frac{P(X = a)}{P(X = b)} > \frac{P(X = a + l)}{P(X = b + l)}$$

sempre que $l > 0$ i $b < a$. Així,

$$\frac{A_{k+1}}{A_k} < \frac{P(X = b_{k+1})}{P(X = b_k)} \leq \frac{P(X = b_1)}{P(X = b_0)} = \eta.$$

Per tant, η tendeix a zero quan N tendeix a infinit.

Per tant, per una determinada p , $\eta \rightarrow 0$ quan $N \rightarrow \infty$, i s'arriba a

$$A_k < \eta^k A_0.$$

Conseqüentment,

$$A_1 + A_2 + \dots < A_0(\eta + \eta^2 + \eta^3 + \dots) = A_0 \frac{\eta}{1 - \eta} \Rightarrow A_0 > \frac{(1 - \eta)}{\eta} (A_1 + A_2 + \dots)$$

i per tant, per N suficientment gran,

$$\frac{(1 - \eta)}{\eta} \geq c.$$

Per trobar un valor N pel qual es pugui aconseguir amb certesa la proporció $\frac{r}{r+s}$, Bernoulli realitza una anàlisi exhaustiva de la relació de probabilitats

$$\eta = \frac{P(X = b_1)}{P(X = b_0)}$$

i demostra que es pot establir $\frac{(1-\eta)}{\eta} > c$, prenent N com l'enter més gran o igual al major dels termes de

$$mt + \frac{st(m-1)}{r+1} \quad (2.4)$$

on

$$m \geq \frac{\log[c(s-1)]}{\log(r+1) - \log r}$$

amb m enter, i

$$mt + \frac{rt(m-1)}{s+1} \quad (2.5)$$

on

$$m \geq \frac{\log[c(r-1)]}{\log(s+1) - \log s}$$

amb $t = r + s$.

Un exemple pot ser prendre $r = 30$ i $s = 20$. L'expressió (2.5) és la més gran, obtenint així $N = 25550$ per $c = 1000$, $N = 31258$ per $c = 10000$ i $N = 36966$ per $c = 1000000$.

Si bé és cert que el desenvolupament de Bernoulli es podria refinar, s'ha de considerar com un gran descobriment de l'època, i segurament com la primera aproximació al càlcul de la incertesa. L'autor demostra de forma quantitativa que, a mesura que augmenta el nombre d'observacions, disminueix la incertesa d'un resultat concret.

2.2 James Stirling

Abans de continuar amb els desenvolupaments de De Moivre, és interessant destacar l'aportació del matemàtic James Stirling, ja que la coneguda fórmula de Stirling serà imprescindible per l'èxit del teorema de De Moivre.

James Stirling (Garden, Stirlingshire, 1692 - Edimburg, 1770) va ser un matemàtic escocès que va fer grans aportacions en la teoria de les sèries infinites i del càlcul infinitesimal. No es coneix informació fidedigna sobre els primers anys de vida de Stirling. El 1711 es va matricular al *Balliol College*, Oxford, amb beques de postgrau per a les quals inicialment va ser exempt com a jacobita (partidari del rei exiliat, Jaume II). Després de la rebel·lió jacobita de 1715, se li va retirar l'exempció i va perdre les seves beques. Va romandre a Oxford fins el 1717, però no es va poder graduar.

El 1717, va publicar un suplement, titulat *Lineae Tertii Ordines Newtonianae*, a l'enumeració de 72 formes de la corba cúbica d'Isaac Newton, que va dedicar a l'ambaixador venecià a Londres. Va presentar *Methodus Differentialis Newtoniana Illustrata* (1719) a través de Newton a la *Royal Society* de Londres, on posteriorment seria escollit com a membre el 1726, gràcies al patrocini de Newton. El 1722 va tornar a Glasgow, i a finals de 1724 va anar a Londres, on va treballar com a professor d'escola. Durant la seva estada a Londres va publicar la seva obra més important, *Methodus Differentialis sive Tractatus de Summatione et Interpolatione Serierum Infinitarum* (1730), un tractat sobre la sèrie infinita, la suma, la interpolació i la quadratura. Conté també la coneguda fórmula de Stirling, tot i que Abraham De Moivre va trobar resultats equivalents simultàniament.

Des de 1734 va treballar per la companyia *Scotch Mines* a Escòcia.

2.2.1 Fórmula de Stirling

Comencem recordant la llei binomial, la qual compta el número d'èxits que s'obtenen quan es realitzen n observacions independents d'algun experiment aleatori amb dos resultats possibles, designats com a èxit i fracàs.

La funció de probabilitat de P_n , corresponent a la mesura de la probabilitat (a l'espai real) de la llei binomial de paràmetres $n \geq 1$ (corresponent al número d'assajos) i $p \in (0, 1)$ (relatiu a la probabilitat d'èxit) és

$$P_n(k) = \binom{n}{k} p^k q^{n-k}$$

amb $k \in \{0, 1, \dots, n\}$ i $q := 1 - p$.

Abans d'enunciar la fórmula de Stirling, és necessari conèixer el següent lema [3].

Lema 2.1. Si $|x| \leq \frac{2}{3}$, aleshores

$$\log(1+x) = x - \frac{x^2}{2} + \theta(x) \quad (2.6)$$

amb $|\theta(x)| \leq |x|^3$.

Prova: Usant Taylor per desenvolupar $\log(1+x)$, amb $|x| < 1$, s'obté

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \cdot \frac{x^n}{n}.$$

Considerant

$$\theta(x) := \sum_{n=3}^{\infty} (-1)^{n-1} \cdot \frac{x^n}{n},$$

usant l'afitació de la funció θ següent

$$|\theta(x)| \leq \sum_{n=3}^{\infty} \frac{|x|^n}{n} \leq \frac{1}{3} \sum_{n=3}^{\infty} |x|^n = \frac{|x|^3}{3(1-|x|)},$$

i prenent la suposició $|x| \leq \frac{2}{3}$, s'arriba a $|\theta(x)| \leq |x|^3$. Per tant, es pot reescriure $\log(1+x)$ com

$$\log(1+x) = x - \frac{x^2}{2} + \theta(x).$$

□

A continuació, es presenta la fórmula de Stirling [3]:

$$n! = K n^n e^{-n} \sqrt{n} e^{w(n)}$$

on $K > 0$ i $\lim_{n \rightarrow \infty} w(n) = 0$.

La demostració d'aquesta és com segueix: Prenent logaritmes a banda i banda, és suficient demostrar que

$$\log n! = \log K + \left(n + \frac{1}{2}\right) \log n - n + w(n).$$

Definint

$$d_n := \log n! - \left(n + \frac{1}{2}\right) \log n + n$$

per $n = 1$ s'obté $d_1 = 1$ i, altrament,

$$d_n - d_{n+1} = \left(n + \frac{1}{2}\right) \log \left(1 + \frac{1}{n}\right) - 1.$$

Usant el Lema 2.1,

$$d_n - d_{n+1} = \left(n + \frac{1}{2}\right) \left(\frac{1}{n} - \frac{1}{2n^2} + \theta\left(\frac{1}{n}\right)\right) - 1 = \left(n + \frac{1}{2}\right) \theta\left(\frac{1}{n}\right) - \frac{1}{4n^2}.$$

Alehores, usant altrament el Lema 2.1 pel cas $n \geq 2$, s'arriba a

$$|d_n - d_{n+1}| \leq \left(n + \frac{1}{2}\right) \frac{1}{n^3} + \frac{1}{4n^2} = \frac{1}{n^2} + \frac{1}{2n^3} + \frac{1}{4n^2} \leq \frac{3}{2n^2}$$

i per tant,

$$\sum_{n=1}^{\infty} (d_{n+1} - d_n) < \infty$$

Prenent N tendint a infinit, es compleix que

$$d_N = 1 + \sum_{n=1}^{N-1} (d_{n+1} - d_n) \rightarrow C$$

on $C \in \mathbb{R}$ és una certa constant. Llavors, definint $K := e^C > 0$, s'assoleixen els resultats

$$\lim_{n \rightarrow \infty} \frac{n! e^n}{n^{n+\frac{1}{2}}} = K \quad \text{i} \quad \lim_{n \rightarrow \infty} w(n) \rightarrow 0.$$

□

2.3 Abraham De Moivre

Jacob Bernoulli va semblar les llavors de la quantificació de la incertesa, però De Moivre protagonitzaria un pas fonamental en aquest àmbit.

Abraham De Moivre (Vitry, 1667 - Londres 1754) va néixer protestant i es mudà a Londres amb tan sols 21 anys, després de patir dos anys d'empresonament a França. A Anglaterra va continuar els seus estudis com a matemàtic, i el 1697 fou elegit membre de la *Royal Society*. La transcendència de l'autor recau en tres obres que tracten diferents aspectes matemàtics en l'àmbit de l'atzar, les quals són *The Doctrine of Chances* [4] de 1718, *Annuities upon Lives* [5] de 1725 i *Miscellanea Analytica* [6] de 1730.

De Moivre afirma que el 1711 només havia llegit el tractat de Huygens *De Ratiociniis in Ludo Aleae* (1657) i un estudi escrit per John Arbuthnot l'any 1692 basat en el treball de Huygens. El 1718 el matemàtic havia topat amb l'obra de Montmort *Essay d'analyse sur les jeux de hasard* (1713) i amb el llibre de Bernoulli *Ars Conjectandi* [2] del 1713. A partir de 1721, De Moivre va començar a avançar en l'aproximació dels termes d'una expansió binomial. La primera publicació dels seus resultats es recopilaren el 1730 a l'obra *Miscellanea Analytica* [6], i foren completats l'any 1733, amb la publicació del que avui en dia es coneix com a “aproximació normal a la distribució binomial”.

2.3.1 *Doctrine of Chances*

Al segle XVIII, es planteja la qüestió de veure si és possible trobar una llei P que approximi P_n per n gran, de manera equivalent a calcular $P_n((a, b])$ per n gran, on $(a, b]$ és un interval.

El llibre V de *Miscellanea Analytica* [6], de 1730, comença citant la part de l'obra *Ars Conjectandi* [2] en la qual Bernoulli aborda el problema de concretar el nombre d'experiments necessaris per determinar la proporció real de casos, dins una determinada aproximació. L'enfocament de De Moivre en aquesta qüestió s'allunya lleugerament de la visió de Bernoulli, ja que el primer busca obtenir una aproximació del coeficient

$$\binom{n}{k} p^k q^{n-k}$$

en comptes de fer-ho directament amb el terme

$$P(|X - Np| \leq N\epsilon) = \sum_{\substack{k \in \{0, \dots, n\} \\ |k - Np| \leq N\epsilon}} \binom{n}{k} p^k q^{n-k}.$$

És en aquesta obra on apareix la fórmula de Stirling en la seva forma definitiva, usada posteriorment per De Moivre el 1733 per descriure la llei normal com una aproximació a la llei binomial. A més, se li atribueix a De Moivre la primera aproximació de la llei normal o gaussiana, a part de la primera forma més prematura del teorema central del límit.

De Moivre inicia les seves investigacions, publicades a l'obra *Doctrine of Chances* [7], amb la distribució binomial simètrica, corresponent al cas $p = q = \frac{1}{2}$. Laplace, l'any 1812, estendria aquest resultat a la seva obra *Théorie analytique des probabilités* [17] pel cas $p \in (0, 1)$.

Comença estudiant el comportament aproximat del terme $(1+1)^n$ del binomi simètric per a n gran. L'any 1730 descobreix que, per n suficientment gran, la relació entre el terme màxim de $(1+1)^n$ (corresponent a $\binom{n}{n/2}$, en notació moderna) i la suma de tots els termes (és a dir, 2^n) pren la següent aproximació:

$$\frac{\binom{n}{n/2}}{2^n} \cong 2 \frac{21}{125} \cdot \frac{(1 - \frac{1}{n})^n}{\sqrt{n-1}}, \quad (2.7)$$

corresponent al terme del mig de la binomial.

Aquesta és la coneguda aproximació de la fórmula de Stirling. De Moivre determina la constant que apareix a (2.7),

$$A = 2 \frac{21}{125} = 2,168$$

a través de l'avaluació numèrica dels primers quatre termes de la sèrie

$$\ln \left(\frac{A}{2} \right) = \frac{1}{12} - \frac{1}{360} + \frac{1}{1260} - \frac{1}{1680} + \dots$$

Més endavant, el descobriment de Stirling

$$\ln \sqrt{2\pi} = 1 - \frac{1}{12} + \frac{1}{360} - \frac{1}{1260} + \dots$$

constata que el valor correcte de la constant és

$$A = \frac{2e}{\sqrt{2\pi}}$$

Amb aquesta A , i aproximant $(1 - \frac{1}{n})^n$ per e^{-1} i $\sqrt{n-1}$ per $\sqrt{n}$, De Moivre presenta la relació del terme màxim de $(1+1)^n$ i 2^n com

$$\frac{\binom{n}{n/2}}{2^n} \cong \frac{2}{\sqrt{nc}}$$

amb $c = 2\pi$ (valor que prendrem al llarg d'aquesta secció sempre que aparegui la constant c). Això no és més que la probabilitat de que la binomial X prengui el valor $\frac{n}{2}$ sota la condició de que n sigui un nombre enter parell.

Llavors, per trobar la relació entre els altres elements de la sèrie i aquest terme màxim, l'any 1730 escriu la relació del terme màxim M i un terme Q que es troba a una distància p del màxim com l'aproximació

$$\frac{M}{Q} = \frac{(m+p-1)^{m+p-\frac{1}{2}}(m-p+1)^{m-p+\frac{1}{2}}(m+p)/m}{m^{2m}}$$

on $m = \frac{n}{2}$. Posteriorment, el 1733 redefineix p per l , per determinar si n és gran en comparació a l . Així, si $\frac{n}{2}$ és una quantitat infinitament gran, el logaritme de la raó entre un terme que dista del central a una distància l i el terme central és $\frac{-2l^2}{n}$:

$$\ln\left(\frac{Q}{M}\right) = \ln\left[\frac{P(X = \frac{n}{2} + l)}{P(X = \frac{n}{2})}\right] \cong -\frac{2l^2}{n} \quad (2.8)$$

que, eliminant logaritmes, es pot reformular per obtenir l'expressió

$$P\left(X = \frac{n}{2} + l\right) \cong P\left(X = \frac{n}{2}\right) e^{-2l^2/n} \cong \frac{2}{\sqrt{nc}} e^{-2l^2/n},$$

i aquesta és la probabilitat d'obtenir el valor $\frac{n}{2} + l$ en una binomial simètrica X .

Per tal d'arribar a aquest resultat, el procediment a seguir seria semblant al que es presenta a continuació:

$$\begin{aligned} \ln\left(\frac{M}{Q}\right) &= \left(m+l-\frac{1}{2}\right) \ln(m+l-1) + \left(m-l+\frac{1}{2}\right) \ln(m-l+1) - 2m \ln m \\ &+ \ln\left(\frac{m+l}{m}\right) = \left(m+l-\frac{1}{2}\right) \ln\left(1+\frac{l-1}{m}\right) + \ln\left(1+\frac{l}{m}\right) \\ &\cong (m+l) \ln\left(1+\frac{l}{m}\right) + (m-l) \ln\left(1-\frac{l}{m}\right) \\ &\cong (m+l) \left(\frac{l}{m} - \frac{l^2}{2m^2}\right) + (m-l) \left(-\frac{l}{m} - \frac{l^2}{2m^2}\right) = \frac{l^2}{m} = \frac{2l^2}{n}. \end{aligned}$$

Per tant, efectivament s'arriba a

$$\ln\left(\frac{Q}{M}\right) \cong -\frac{2l^2}{n}$$

ja que es negligeixen els termes d'ordre menor a $\frac{l^2}{m}$.

De Moivre, a part d'abordar aquesta funció exponencial com una aproximació a la distribució binomial, entreveu que el seu comportament es pot associar al d'una “corba” amb dos punts d'inflexió, un a cada costat del terme màxim, corresponents a una distància de

$$\frac{(n+2)^{\frac{1}{2}}}{2}$$

d'aquest màxim (quan n és gran, es considera $\frac{n^{1/2}}{2}$ com a terme màxim).

El matemàtic usa aquesta corba per perfeccionar el primer intent de quantificació de la incerensa de Bernoulli. El límit superior que havia albirat Jacob Bernoulli fou realment decebedor, ja que trobà que es requereixen 25550 experiments per tal de saber la proporció de casos fèrtils i estèrils en una de cada cinquanta parts, fet que impossibilita el coneixement de quelcom de forma fiable, realitzant un nombre raonable d'experiments.

Bernoulli, frustrat amb els resultats obtinguts, decideix finalitzar de forma abrupta el seu treball, recopilat a l'obra *Ars Conjectandi* [2]. Tot i el fracàs, els resultats d'aquest il·lustre matemàtic obren les portes als futurs desenvolupaments i correccions a mans de De Moivre. Bernoulli estableix que la proporció empírica es troba en una de cada cinquanta parts de la proporció real, factor que desencadena el fracàs del matemàtic. De Moivre, estudiant la “corba binomial”, troba un límit més precís que dona lloc a una estimació més acurada de la incertesa.

Per arribar al resultat (2.8) l'any 1733, De Moivre fa ús de les expressions

$$\frac{M}{2^n} \cong \frac{2}{\sqrt{nc}}$$

i

$$\ln \left(\frac{Q}{M} \right) \cong -\frac{2l^2}{n},$$

que en notació actual equivalen a

$$\frac{M}{2^n} = P \left(X = \frac{n}{2} \right)$$

i

$$\frac{Q}{M} = \frac{P(X = \frac{n}{2} + l)}{P(X = \frac{n}{2})}.$$

Escrivint l'exponencial $e^{-2l^2/n}$ com la sèrie

$$1 - \frac{2l^2}{n} + \frac{4l^4}{2n^2} - \frac{8l^6}{6n^3} + \frac{16l^8}{24n^4} - \frac{32l^{10}}{120n^5} + \frac{64l^{12}}{720n^6} + \dots$$

i integrant-la terme a terme entre 0 i l ,

$$l - \frac{2l^3}{3n} + \frac{4l^5}{2 \cdot 5n^2} - \frac{8l^7}{6 \cdot 7n^3} + \frac{16l^9}{24 \cdot 9n^4} - \frac{32l^{11}}{120 \cdot 11n^5} + \dots$$

només queda multiplicar-la per $2\sqrt{nc}$ per obtenir la desitjada suma de termes des de 0 fins a l . L'autor nota que la sèrie només depèn de $\frac{l}{\sqrt{n}}$, i per tant l'avalua a $l = s\sqrt{n}$, aconseguint així la suma de termes des de $l = 0$ fins $l = s\sqrt{n}$ com

$$\frac{2}{\sqrt{c}} \left(s - \frac{2s^3}{3} + \frac{4s^5}{2 \cdot 5} - \frac{8s^7}{6 \cdot 7} + \frac{16s^9}{26 \cdot 9} - \frac{32s^{11}}{120 \cdot 11} + \dots \right). \quad (2.9)$$

En notació actual, els càlculs de De Moivre són tal i com es presenten a continuació. Es vol trobar la suma de $P(X = \frac{n}{2} + l)$ des de $l = 0$ fins a $l = s\sqrt{n}$. A partir de

$$P\left(X = \frac{n}{2} + l\right) = P\left(X = \frac{n}{2}\right) \cdot \frac{P\left(X = \frac{n}{2} + l\right)}{P\left(X = \frac{n}{2}\right)} \cong \frac{2}{\sqrt{nc}} e^{-2l^2/n} \quad (2.10)$$

s'aproxima la suma

$$\sum_{l=0}^{s\sqrt{n}} P\left(X = \frac{n}{2} + l\right) \quad (2.11)$$

per la integral

$$\frac{2}{\sqrt{nc}} \int_0^{s\sqrt{n}} e^{-2l^2/n} dl$$

i s'avalua aquesta última expandint l'integrand en una sèrie i integrant terme a terme. Aleshores, substituint $l = s\sqrt{n}$, s'arriba a (2.9) com una aproximació a (2.11).

De Moivre considera que aquesta sèrie convergeix prou ràpid per valors de $s \leq \frac{1}{2}$, però per valors de s superiors o iguals a 1 proposa una avaluació de la integral a través de fórmules de quadratura estàndards. Conseqüentment, troba la suma des de $l = 0$ fins a $l = \sqrt{n}$ mitjançant la suma des de $l = 0$ fins a $l = \frac{\sqrt{n}}{2}$ de (2.8) i afegint

$$\frac{2}{\sqrt{nc}} \cdot \frac{(A + D) + 3(B + C)}{8} \cdot \frac{1}{2} \sqrt{n}$$

on A, B, C, D són $e^{-2l^2/n}$ avaluades, respectivament, a $l = \frac{3}{6}\sqrt{n}$, $l = \frac{4}{6}\sqrt{n}$, $l = \frac{5}{6}\sqrt{n}$ i $l = \frac{6}{6}\sqrt{n}$.

Finalment, De Moivre compona un procediment per sumar els termes del binomi, ja sigui a través de la integració d'una sèrie o bé a partir d'una quadratura de quatre punts. El 1733 es fa accessible l'aproximació de $e^{-2l^2/n}$, a diferència de l'anterior expressió (2.10), estudiada l'any 1730 pel mateix autor. Demuestra aquesta accessibilitat en una sèrie de corollaris que conformen la primera taula de la distribució normal. Calcula el valor de (2.8), corresponent a 0,341344 per $s = \frac{1}{2}$, que en notació contemporània es correspon a

$$P\left(0 \leq X - \frac{n}{2} \leq \frac{\sqrt{n}}{2}\right) \cong 0,341344$$

i troba la probabilitat de que el número d'esdeveniments d'un experiment binomial simètric caigui entre $\frac{n}{2} - l$ i $\frac{n}{2} + l$, per $l = \frac{\sqrt{n}}{2}$, $\sqrt{n}$, $\frac{3\sqrt{n}}{2}$. També esbrina el valor aproximat de l tal que la probabilitat de que el nombre d'esdeveniments s'allunyi de $\frac{n}{2}$ per més de l és $\frac{1}{2}$:

$$P\left(\frac{n}{2} - l \leq X \leq \frac{n}{2} + l\right) = \frac{1}{2}.$$

És equivalent a solucionar

$$\int_{-a}^a \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \frac{1}{2}$$

per trobar $a = \frac{1}{\sqrt{2}} = 0,707$, un valor que no es distancia massa de la solució exacta $a = 0,674$, i que representa una primera apreciació del posterior “error probable” (la desviació mitjana).

Alguns dels valors que De Moivre computa el 1733 de la integral normal

$$\int_{-a}^a \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

com aproximacions a les probabilitats binomials

$$P\left(\frac{n}{2} - \frac{a\sqrt{n}}{2} \leq X \leq \frac{n}{2} + \frac{a\sqrt{n}}{2}\right)$$

són:

- Per $a = 1$, De Moivre troba el valor 0,682688 (valor exacte: 0,682689)
- Per $a = 2$, De Moivre troba el valor 0,95428 (valor exacte: 0,95450)
- Per $a = 3$, De Moivre troba el valor 0,99874 (valor exacte: 0,99730)

Com que, usant la fórmula de Stirling (demostrada abans per De Moivre per tractar el seu cas singular que no pas per Stirling) s'obté el resultat

$$\binom{2m}{m} \frac{1}{2^{2m}} \sim \frac{1}{\sqrt{\pi n}},$$

aleshores el quocient dels dos termes convergeix a 1 a mesura que n tendeix a infinit (recordem que $n = 2m$), equivalent a una convergència a 0 dels dos termes a igual velocitat.

De forma semblant, usant la mateixa aproximació es pot veure que, per a qualsevol $k \in \{-m, \dots, m\}$,

$$\binom{2m}{m+k} \frac{1}{2^{2m}}$$

també convergeix a 0 com $\frac{1}{\sqrt{\pi n}}$. Aquest fet prova que, per a qualsevol l fixat,

$$\sum_{k=-l}^l \binom{2m}{m+k} \frac{1}{2^{2m}} \rightarrow 0, \quad \text{si} \quad m \rightarrow \infty.$$

D'altra banda, si l'interval $[-l, l]$ creix com $\sqrt{n}$ a mesura que n tendeix a infinit, aleshores es pot compensar el comportament de tipus $\frac{1}{\sqrt{n}}$ de cada un dels termes, i per tant les sumes convergeixen cap a algun límit entre 0 i 1, resultat que apareix en el Teorema de De Moivre-Laplace.

D'aquesta manera, De Moivre reconeix que l'escala sobre la qual s'han de jutjar les desviacions del centre és $\sqrt{n}$, és a dir, les distàncies des del centre s'han de mesurar com a múltiples de l'arrel quadrada del nombre d'esdeveniments (n). Per exemple, si $\sqrt{n} = 60$, una diferència de 50 és $\frac{5}{6}$ unitats del centre. A més, defineix $\sqrt{n}$ com el "mòdul".

És en el treball d'aquest reconegut matemàtic on es pot començar a constatar que, a l'afegir mesures independents, la precisió augmenta amb l'increment de l'arrel quadrada de la mida de la mostra, $\sqrt{n}$. Aquest desenvolupament quantitatiu és molt rellevant per poder donar pas a l'extensió d'aquests conceptes a mans d'altres grans matemàtics, principalment Laplace.

Per tant, a partir de (2.10) De Moivre conclou que

$$P\left(\left|X - \frac{n}{2}\right| \leq l\right) \approx \frac{2}{\sqrt{2\pi n}} \int_{-l}^l e^{-\frac{2s^2}{n}} ds,$$

que constitueix la primera aparició de la corba normal, amb el terme de la dreta conegut actualment com la distribució normal de mitjana 0 i variància $\frac{n}{4}$.

Recordem que la funció de densitat d'una funció normal estàndard és

$$\phi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$$

i la seva funció de distribució és

$$\Phi(x) = \int_{-\infty}^x \phi(t) dt.$$

La majoria de resultats i exemples proposats per De Moivre s'ocupen del cas simètric, amb probabilitats proporcionals als termes $(1+1)^n$. L'autor afegeix l'expansió de $(a+b)^n$, corresponent als resultats pel cas general, i determina les expressions

$$\frac{M}{(a+b)^n} \cong \frac{a+b}{\sqrt{abnc}}$$

i

$$\ln\left(\frac{Q}{M}\right) \cong -\frac{(a+b)^2}{2abn} \cdot l^2$$

on, seguint amb la notació anterior, M correspon al terme màxim de l'expansió i Q al terme a distància l de M . En notació actual, equival a

$$P(X = np) \cong \frac{1}{\sqrt{p(1-p)nc}}$$

i

$$\ln\left[\frac{P(X = np + l)}{P(X = np)}\right] \cong -\frac{l^2}{2np(1-p)}$$

on X té una distribució binomial (n, p) , $p = \frac{a}{(a+b)}$, i np és enter. Aquest resultat equival a la corba normal de mitjana 0 i variància npq :

$$P(|X - np| \leq l) \approx \frac{1}{\sqrt{2\pi npq}} \int_{-l}^l e^{-\frac{s^2}{2npq}} ds.$$

Aleshores, l'autor proposa el corollari que segueix, pel qual no presenta cap demostració del resultat.

Corollari 2.2. *Si les probabilitats d'èxit i fracàs es troben en una proporció qualsevol de desigualtat, els problemes relacionats amb la suma dels termes del binomi $(a+b)^n$ es resoldran amb la mateixa facilitat que aquells en què les probabilitats d'èxit i fracàs estiguin en proporció d'igualtat [[7], p. 250].*

Tot i que De Moivre obté uns resultats molt més precisos que Bernoulli, no es percep una evolució notòria en l'aplicació dels mateixos. El desenvolupament més transcendent de la quantificació de la incertesa i els primers grans èxits en capgirar la visió de De Moivre i crear una teoria matemàtica de la inferència es troben en les obres de Laplace. No obstant, altres autors previs també aborden el problema de la inferència, el qual s'allunya del treball de De Moivre en l'expansió binomial, però les tècniques usades pel seu tractament matemàtic es basen en els resultats de De Moivre.

2.3.2 Teorema de De Moivre-Laplace

En notació actual, podem presentar una proposició alternativa, amb la seva corresponent demostració, pel teorema de De Moivre-Laplace, que apareix per primera vegada a la segona edició de *Doctrine of Chances* [7], publicada el 1738. El teorema de De Moivre-Laplace és un cas particular del teorema central del límit, i enuncia que la distribució normal es pot usar com una aproximació de la distribució binomial, sota unes determinades condicions.

Abans, però, presentem un lema [3] que serà de molta utilitat.

Lema 2.3. *Considerem una mesura de probabilitat binomial, amb paràmetres n i p , amb esperança np i variància npq (n correspon al número d'assajos, p a la probabilitat d'èxit, i $q = 1 - p$ a la probabilitat de fracassar). Per a tot $k \in \{0, 1, \dots, n\}$, considerem els valors*

$$x(k, n) := \frac{k - np}{\sqrt{npq}}. \quad (2.12)$$

Sigui $A > 0$ fixat. Pels k que satisfan $|x(k, n)| \geq A$, s'obté

$$\psi(k, n) := \binom{n}{k} p^k q^{n-k} \sim \varphi(k, n) := \frac{1}{K\sqrt{npq}} \exp \left\{ -\frac{x(k, n)^2}{2} \right\}$$

uniformement en k (K és la constant que apareix a la fórmula de Stirling).

Prova: Com que $|x(k, n)| \geq A$, aleshores

$$k \in (np - A\sqrt{npq}, np + A\sqrt{npq}) \iff \frac{k}{np} \in \left(1 - A\sqrt{\frac{q}{np}}, 1 + A\sqrt{\frac{q}{np}} \right).$$

Per tant, per a tot k satisfent $|x(k, n)| \geq A$, com que

$$\left| \frac{k}{np} - 1 \right| \geq A\sqrt{\frac{q}{np}} \rightarrow 0,$$

s'obté que $k \sim np$ uniformement. Anàlogament, es demostra que $n - k \sim nq$, de manera que quan $n \rightarrow \infty$, també es compleix que $k \rightarrow \infty$ i que $n - k \rightarrow \infty$.

Aleshores, s'aplica la fórmula de Stirling:

$$\begin{aligned} \binom{n}{k} p^k q^{n-k} &= \frac{n!}{k!(n-k)!} p^k q^{n-k} \\ &= \frac{Kn^n e^{-n} \sqrt{n} e^{w(n)}}{[Kk^k e^{-k} \sqrt{k} e^{w(k)}] \cdot [K(n-k)^{n-k} e^{-(n-k)} \sqrt{n-k} e^{w(n-k)}]} p^k q^{n-k} \\ &= \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \frac{1}{K\sqrt{npq}} \sqrt{\frac{np}{k}} \sqrt{\frac{nq}{n-k}} e^{w(n)-w(k)-w(n-k)}. \end{aligned}$$

Ara només cal provar que

$$|\log \psi(k, n) - \log \varphi(k, n)| \rightarrow 0$$

uniformement en k . Comencem desenvolupant l'expressió

$$\begin{aligned} |\log \psi(k, n) - \log \varphi(k, n)| &= \left| k \log \left(\frac{np}{k} \right) + (n-k) \log \left(\frac{nq}{n-k} \right) \right. \\ &\quad \left. + \frac{1}{2} \log \left(\frac{np}{k} \right) + \frac{1}{2} \log \left(\frac{nq}{n-k} \right) + w(n) - w(k) - w(n-k) + \frac{x(k, n)^2}{2} \right|. \end{aligned}$$

Només s'ha de demostrar que la següent expressió convergeix a zero:

$$\left| k \log \left(\frac{np}{k} \right) + (n-k) \log \left(\frac{nq}{n-k} \right) + \frac{x(k,n)^2}{2} \right|. \quad (2.13)$$

Usant la definició de $x(k,n)$, s'obté

$$\frac{np}{k} = 1 - \frac{x(k,n)\sqrt{npq}}{k} \quad \text{i} \quad \frac{nq}{n-k} = 1 + \frac{x(k,n)\sqrt{npq}}{n-k}.$$

Substituint aquestes expressions a (2.13), resta demostrar la convergència a zero de

$$k \log \left(1 - \frac{x(k,n)\sqrt{npq}}{k} \right) + (n-k) \log \left(1 + \frac{x(k,n)\sqrt{npq}}{n-k} \right) + \frac{x(k,n)^2}{2}. \quad (2.14)$$

A partir de $|x(k,n)| \geq A$,

$$\left| \frac{x(k,n)\sqrt{npq}}{k} \right| \leq \frac{Anpq}{k\sqrt{npq}} = \frac{Aq}{\sqrt{npq}} \frac{np}{k}.$$

Notem que el terme $\frac{np}{k}$ convergeix a 1 (vist anteriorment), mentre que $\frac{Aq}{\sqrt{npq}}$ convergeix a 0, de manera que s'aconsegueix l'acotació

$$\left| \frac{x(k,n)\sqrt{npq}}{k} \right| \leq \frac{2}{3}.$$

Anàlogament es pot estudiar el cas

$$\left| \frac{x(k,n)\sqrt{npq}}{n-k} \right| \leq \frac{2}{3}$$

i s'arriba a la mateixa fitació.

A continuació, cal demostrar que la següent expressió convergeix a 0. Substituïm (2.6) a (2.14), i obtenim

$$\begin{aligned} & -k \frac{x(k,n)\sqrt{npq}}{k} - \frac{k}{2} \left(\frac{x(k,n)\sqrt{npq}}{k} \right)^2 + k \cdot \theta \left(-\frac{x(k,n)\sqrt{npq}}{k} \right) + (n-k) \frac{x(k,n)\sqrt{npq}}{n-k} \\ & - \frac{n-k}{2} \left(\frac{x(k,n)\sqrt{npq}}{n-k} \right)^2 + (n-k) \cdot \theta \left(\frac{x(k,n)\sqrt{npq}}{n-k} \right) + \frac{x(k,n)^2}{2} \\ & = \frac{x(k,n)^2}{2} \left(1 - \frac{npq}{k} - \frac{npq}{n-k} \right) + k \cdot \theta \left(-\frac{x(k,n)\sqrt{npq}}{k} \right) + (n-k) \cdot \theta \left(\frac{x(k,n)\sqrt{npq}}{n-k} \right). \end{aligned}$$

Prenent valor absolut al primer terme, obtenim

$$\begin{aligned} & \left| \frac{x(k,n)^2}{2} \left(1 - \frac{npq}{k} - \frac{npq}{n-k} \right) \right| \leq \frac{A^2}{2} \left| 1 - \frac{npq}{k} - \frac{npq}{n-k} \right| \\ & \leq \frac{A}{2} \left(q \left| 1 - \frac{np}{k} \right| + p \left| 1 - \frac{nq}{n-k} \right| \right). \end{aligned}$$

Com que els termes $\frac{npq}{k}$ i $\frac{npq}{n-k}$ convergeixen a 1, aleshores aquest primer terme convergeix a 0 uniformement en k . Els altres dos termes presenten les acotacions que segueixen:

$$\left| k \cdot \theta \left(-\frac{x(k, n)\sqrt{npq}}{k} \right) \right| \leq \frac{A^3(npq)^{3/2}}{k^2}$$

i

$$\left| (n-k) \cdot \theta \left(\frac{x(k, n)\sqrt{npq}}{n-k} \right) \right| \leq \frac{A^3(npq)^{3/2}}{(n-k)^2}.$$

Sumant aquestes dues fites s'arriba a

$$A^3(npq)^{3/2} \left(\frac{1}{k^2} + \frac{1}{(n-k)^2} \right) = \frac{A^3}{\sqrt{npq}} \left[\left(\frac{npq}{k} \right)^2 + \left(\frac{npq}{n-k} \right)^2 \right].$$

Aleshores, notem que $\left[\left(\frac{npq}{k} \right)^2 + \left(\frac{npq}{n-k} \right)^2 \right]$ convergeix a $q^2 + p^2$ uniformement en k , i $\frac{A^3}{\sqrt{npq}}$ convergeix a 0. $\square$

El Teorema de De Moivre-Laplace [7] diu el següent:

Teorema 2.4. *Segui P_n la mesura de probabilitat amb llei binomial de paràmetres n i p , i siguin $a, b \in \mathbb{R}$ amb $a < b$, aleshores*

$$\lim_n P_n((np + a\sqrt{npq}, np + b\sqrt{npq}]) = \int_a^b \phi_K(x) dx$$

on

$$\phi_K(x) = \frac{1}{K} e^{-x^2/2}.$$

Prova: Primer, usant la definició de funció de probabilitat de la llei binomial,

$$P_n((np + a\sqrt{npq}, np + b\sqrt{npq}]) = \sum_{a < x(k, n) \leq b} \binom{n}{k} p^k q^{n-k}.$$

Fixat l'interval (a, b) , es pot trobar A tal que $(a, b] \subseteq [-A, A]$. Si ara s'aplica el Lema 2.3, s'arriba a

$$\sum_{a < x(k, n) \leq b} \binom{n}{k} p^k q^{n-k} \sim \sum_{a < x(k, n) \leq b} \frac{1}{K\sqrt{npq}} \exp \left\{ -\frac{x(k, n)^2}{2} \right\}.$$

Es pot reescriure la suma de la dreta usant

$$x(k+1, n) - x(k, n) = \frac{1}{\sqrt{npq}}$$

i s'obté

$$\sum_{a < x(k, n) \leq b} \frac{1}{K} \exp \left\{ -\frac{x(k, n)^2}{2} \right\} \cdot \frac{1}{\sqrt{npq}} = \sum_{a < x(k, n) \leq b} \phi_K(x(k, n))(x(k+1, n) - x(k, n)).$$

Observem que els valors de $x(k, n)$ recorren l'interval

$$\left[-\sqrt{\frac{np}{q}}, \sqrt{\frac{nq}{p}} \right]$$

amb increments de valor $\frac{1}{\sqrt{npq}}$. Aleshores, a mesura que n tendeix a infinit, aquest interval acaba cobrint $(a, b]$, de manera que s'obté una partició de $(a, b]$ a partir dels valors $x(k, n)$, de pas $\frac{1}{\sqrt{npq}}$. Així doncs, existeixen j i l complint

$$x(j-1, n) \leq a < x(j, n) < \dots < x(l, n) \leq b < x(l+1, n).$$

Finalment, s'arriba al sumatori

$$\sum_{k=j}^l \phi_K(x(k, n))(x(k+1, n) - x(k, n)).$$

A mesura que n tendeix a infinit, el pas de la partició tendeix a 0 $\left(\frac{1}{\sqrt{npq}} \rightarrow 0\right)$, i la suma tendeix a

$$\int_a^b \phi_K(x) dx.$$

□

Per concretar el resultat trobat, és interessant veure la següent demostració [3]: Sigui S_n una variable aleatòria amb llei binomial $(Bin(n, p))$ i n suficientment gran, aleshores

$$P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) \rightarrow \int_a^b \phi_K(x) dx, \quad \text{si} \quad n \rightarrow \infty.$$

Recordem la desigualtat de Txebychev, la qual ens diu que si S_n és una variable aleatòria, i coneixem l'esperança $E(S_n)$ i la variància $Var(S_n)$ de la mateixa, aleshores, per a qualsevol $b > 0$,

$$P(|S_n - E(S_n)| > b) \leq \frac{Var(S_n)}{b^2}.$$

Per tant, aplicant aquesta desigualtat i considerant que $\frac{S_n - np}{\sqrt{npq}}$ és una variable centrada de variància 1, per a qualsevol b es compleix que

$$P\left(\left|\frac{S_n - np}{\sqrt{npq}}\right| > b\right) \leq \frac{1}{b^2} \quad \Rightarrow \quad P\left(\left|\frac{S_n - np}{\sqrt{npq}}\right| \leq b\right) \geq 1 - \frac{1}{b^2}.$$

Llavors,

$$1 - \frac{1}{b^2} \leq P\left(\left|\frac{S_n - np}{\sqrt{npq}}\right| \leq b\right) \leq 1.$$

Quan n tendeix a infinit,

$$1 - \frac{1}{b^2} \leq \int_{-b}^b \phi_K(x) dx \leq 1$$

i quan b tendeix a infinit,

$$\int_{\mathbb{R}} e^{-x^2/2} dx = K \quad \Rightarrow \quad K = \sqrt{2\pi}$$

Hem arribat a aquest resultat usant el canvi de variable $t = \frac{x}{\sqrt{2}}$, de manera que s'obté

$$\int_{\mathbb{R}} e^{-t^2} dt = \sqrt{\pi}.$$

Així, com que hem trobat ϕ amb $K = \sqrt{2\pi}$, corresponent a la llei normal estàndard,

$$\phi_K(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

s'arriba a un dels resultats més transcendents en Estadística:

$$P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) \approx P(a < Z \leq b)$$

on Z correspon a una variable aleatòria normal estàndard.

2.4 Thomas Simpson

Arribat a aquest punt, alguns matemàtics centren els seus esforços en el problema de combinar observacions discordants d'un cos astronòmic. Si cinc observadors registren cinc temps diferents pel pas d'una estrella davant el punt de mira d'un telescopi, es vol determinar la relació d'aquests resultats.

Simpson és el primer en publicar els seus resultats a aquest problema l'any 1755. Tot i no arribar a l'èxit, les aportacions d'aquest matemàtic van ser crucials pel posterior desenvolupament triomfant del mateix a mans de Laplace.

Thomas Simpson (Market Bosworth, 1710 - Amiens, 1761) va ser un matemàtic molt talentós i autodidacte, que va seguir De Moivre tant cronològicament com intel·lectualment. La seva primera feina fou de teixidor, professió a la qual es dedicava el seu pare. Va ensenyar matemàtiques a Nuneaton, Warwickshire, des del 1725 fins el 1733. El 1730, amb tan sols dinou anys, es va casar amb una vídua de cinquanta anys amb dos fills, i amb vint-i-cinc anys es va mudar a Londres, on treballava com a teixidor durant el dia i com a professor de matemàtiques al vespre.

El 1736, va ser un dels primers membres de la *Spitalfields Mathematical Society*. A més, va destacar d'entre un grup de professors itinerants que ensenyaven a les cafeteries de Londres. El 1743 va ser nomenat cap de matemàtiques de la *Royal Military Academy* de Woolwich, i dos anys més tard va ser escollit membre de la *Royal Society*. També va ser elegit membre de la Reial Acadèmia Sueca de Ciències el 1758.

Va publicar *A New Treatise of Fluxions* (1737), i estudià la interpolació i els mètodes numèrics d'integració. Cal destacar que la coneguda "regla de Simpson" té com a base els coneixements adquirits de Newton, tal i com el mateix autor va reconèixer. Va treballar en la teoria de la probabilitat, i l'any 1740 publicà el tractat *The Nature and Laws of Chance* [23] sobre les matemàtiques dels jocs d'atzar. El 1742, publicà el llibre *The Doctrine of Annuities and Reversions* [24]. Ambdues obres estan basades en el treball previ de De Moivre, presentades amb un format més breu i escrites a un nivell més accessible pel públic.

Treballà en teoria d'errors i va voler demostrar que la mitjana aritmètica era millor que una sola observació, justificació que aparegué en les seves memòries de 1757. Va escriure

Mathematical Dissertations (1743) i *The Doctrine and Application of Fluxions* (1750). A més, a *Miscellaneous Tracts* (1757) va discutir alguns problemes d'astronomia.

L'obra de Simpson que ens interessa [25] fou llegida a la Royal Society el 10 d'Abril de 1755, moment en el qual era professor a la *Royal Military Academy* a Woolwich i membre de la *Royal Society*.

2.4.1 Mitjana d'un conjunt d'observacions a l'astronomia

Simpson aborda el problema de prendre la mitjana d'una sèrie d'observacions en l'àmbit de l'astronomia. Tot i que el seu tractament és realment limitat, aquest pren importància per dos factors clau, un de conceptual i l'altre més tècnic. Per fer-ho, utilitza sis mesures i adopta una hipòtesi molt concreta sobre el caràcter probabilístic de les observacions, per concloure que la mitjana proporciona un resultat millor que en el cas d'una única observació.

Pel desenvolupament conceptual, es centra en els errors generats en les observacions, en les diferències entre les observacions registrades i la posició real del cos estudiat, en comptes de focalitzar-se en les observacions per elles mateixes o en el cos astronòmic albirat.

Simpson comença assumint una hipòtesi concreta per la distribució d'errors i estudia l'error mitjà, en lloc de l'observació mitjana. Sense saber la posició del cos observat, però coneixent la distribució d'errors, basa la seva anàlisi en aquesta última per poder fer front al problema. Escull un problema d'inferència (problema en el qual s'analitzen dades generades d'acord a una distribució de probabilitat desconeguda, de la qual s'indueix el seu comportament amb un risc d'error avaluable) i es limita a estudiar les distribucions d'error fixades i conegudes per hipòtesi. L'actual "invariància de localització" del problema garanteix que l'error de la mitjana és igual a la mitjana dels errors, i la incertesa de la inferència (l'error d'estimació usant la mitjana) es pot mesurar en termes d'una determinada distribució.

Aquest és el punt clau que distingeix el problema de Simpson del de De Moivre, i el fa tractable inclús sense disposar d'una teoria de la inferència. Les distribucions binomials d'error no estan fixades, i per tant no es poden prendre com a conegudes (ni per hipòtesi) a no ser que la proporció real es prengui com a notòria.

Simpson suposa que cada una de les n observacions independents és susceptible a errors de possibles mides $-v, -v + 1, \dots, -3, -2, -1, 0, 1, 2, 3, \dots, v$, amb probabilitats proporcionals a

$$r^{-v}, \dots, r^{-3}, r^{-2}, r^{-1}, r^{-0}, r^1, r^2, \dots, r^v$$

o bé a

$$r^{-v}, 2r^{-v+1}, 3r^{-v+2}, \dots, (v+1)r^0, \dots, 3r^{v-2}, 2r^{v-1}, r^v$$

Cada elecció de $r > 0$ dona una distribució d'errors. A primera vista, aquesta sembla una tria estranya per a famílies de distribucions d'errors. Simpson explica que aquesta tria fou dictada per conveniència matemàtica, i és aquí on rau el desenvolupament tècnic destacable de l'autor. Sembla ser que aquesta elecció de les distribucions d'error deriva de resultats previs de De Moivre.

Simpson només s'interessa en aquells casos on $r = 1$ (distribucions d'error simètriques), ja que l'únic càlcul que fa és pel cas en què la distribució d'error és proporcional a 1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1.

El desenvolupament matemàtic de Simpson és innovador però molt limitat. Determina el què els millors astrònoms ja havien anticipat anys abans (com més observacions i experiments es realitzen, la conclusió esdevé menys susceptible d'error), però la idea de basar-ho en càlculs matemàtics és el tret revolucionari. A més, intueix que el concepte de distribucions d'error pot donar lloc a la mesura de la incertesa.

3 La probabilitat inversa

3.1 Pierre-Simon Laplace

Pierre-Simon Laplace (Beaumont-en-Auge, Normandia, 1749 - París, 1827) va ser un astrònom, físic i matemàtic francès. Va fer aportacions rellevants en l'àmbit de la mecànica newtoniana i també en el posterior descobriment de la transformada de Laplace, revelant, a més, l'equació de Laplace. El seu treball com a estadístic es centra en la teoria analítica de la probabilitat.

Els primers desenvolupaments de Laplace relacionats amb l'aplicació de la probabilitat a la inferència es van produir entre el 1772 i el 1781. Es registren dues memòries importants de l'autor durant aquest període, les quals són *Memòria sobre la probabilitat de les causes dels esdeveniments* [11] del 1774, i *Memòria sobre probabilitats* [13] del 1781. Aquestes es troben entre els treballs més importants i complexos en la història de la probabilitat matemàtica, i en conjunt es pot considerar l'obra més influent del segle XVIII sobre l'ús de la probabilitat en inferència.

Va estudiar a la universitat de Caen, i el 1767 va entrar a treballar com a professor a l'Escola Militar de París. El 1785 va ser elegit membre de l'Acadèmia de Ciències, i el 1795 membre de la càtedra de matemàtiques del Nou Institut de les Ciències i les Arts, que va presidir el 1812. Va publicar el primer dels cinc volums de *Traité de mécanique céleste* [16]. El seu antic alumne, Napoleó Bonaparte, el 1805 li va conferir la Legió d'honor, i el 1806 el títol de comte de l'Imperi. Va publicar *Théorie analytique des probabilités* [17] el 1812 i *Assaig filosòfic de les probabilitats* [18] el 1814. El 1816 fou elegit membre de l'Acadèmia Francesa.

3.1.1 Probabilitat Inversa

Concebuda una distribució aleatòria d'errors, la inversió es pot seguir gairebé de forma involuntària. Sigui e l'error, O l'observació i P el punt observat, aleshores $O = P + e$ implica $P = O - e$. Si tenim una distribució aleatòria i simètrica de e , i suposant P fixat, s'obté una distribució per O ; en canvi, prenent O tal i com s'ha determinat, s'arriba a una distribució per P . Tractant només la diferència aleatòria $e = O - P$, s'obté una situació simètrica on la inversió esdevé més natural. Així, la distribució d'errors proporciona l'element aleatori tant per a les afirmacions de probabilitat "endavant" (en el temps) com "inverses", i també proporciona una idea de la inferència inversa, raonament probabilístic de l'efecte (O) a la causa (P), que donarà fruits en altres aplicacions.

Les memòries de Laplace de 1774 [11] i 1781 [13] comencen amb una discussió general de la probabilitat, aplicacions en la vida quotidiana i estudis de problemes que involucren jocs. Posteriorment, es focalitza en el problema d'escollir la mitjana d'un conjunt d'observacions astronòmiques.

A les cròniques de 1774 [11], l'autor comença enunciant un principi de probabilitat inversa i dona quatre exemples de l'aplicació del mateix. El quart exemple presenta una clara aplicació pràctica: Determinar la mitjana entre tres observacions d'un mateix fenomen. El principi és el següent:

“If an event can be produced by a number n of different causes, then the probabilities of these causes given the event are to each other as the probabilities of the event given the causes, and the probability of the existence of each of these is equal to the probability of the event given that cause, divided by the sum of all the probabilities of the event given each of these causes.” [[11], p. 634]

Aquesta és una afirmació de l'actual teorema de Bayes prenent, a priori, la igualtat en la probabilitat de les causes. Per facilitar la discussió, introduïm una notació que no apareix en els documents de Laplace o Bayes: Siguin n causes $A_1, A_2, \dots, A_n$ i sigui E l'esdeveniment mencionat en la discussió, aleshores el principi és:

$$\frac{P(A_i|E)}{P(A_j|E)} = \frac{P(E|A_i)}{P(E|A_j)} \quad (3.1)$$

o bé

$$P(A_i|E) = \frac{P(E|A_i)}{\sum_j P(E|A_j)}. \quad (3.2)$$

En notació moderna equival a

$$P(A_i|E) = \frac{P(E|A_i)P(A_i)}{\sum_j P(E|A_j)P(A_j)}.$$

Pot ser interessant fer un petit parèntesi per recordar la definició de probabilitat condicional $P(A|B)$, corresponent a la “probabilitat de A donada B ”, on A i B són dos successos. Donat un espai de probabilitat i dos successos A i B amb $P(B) > 0$, aleshores

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

Així doncs, el principi de Laplace és anàleg al teorema de Bayes suposant igualtat de probabilitats, és a dir, $P(A_i) = \frac{1}{n}$. L'enfoc de Laplace al càlcul de probabilitats, manllevat de Jacob Bernoulli i d'Abraham De Moivre, consisteix en reduir tots els problemes a una enumeració de casos igualment probables. En aquest context, assumir causes igualment probables significa que el problema està especificat (les “causes” estan enumerades i definides) de manera que són igualment possibles.

L'origen del principi de Laplace no sorgeix dels desenvolupaments de Bayes, ja que la visió del primer divergeix molt de la del darrer. Bayes realitza una progressió ordenada, començant amb una definició de la probabilitat i una derivació de la relació

$$P(A_i|E) = \frac{P(A_i \cap E)}{P(E)},$$

la qual es deriva explícitament i de forma rigorosa dels primers principis. A més, inclou un argument per justificar l'acceptació de causes igualment probables. En canvi, Laplace no presenta cap demostració dels seus resultats o l'origen del principi exposat anteriorment, la demostració del qual apareix més tard, a *Théorie analytique des probabilités* [17], i on

es fa explícita l'assumpció de causes igualment probables. No obstant, el 1781 presenta una primera derivació, la qual es mostra tot seguit (en notació moderna).

Es considera l'ocurrència de l'esdeveniment E dues vegades. Sigui E_1 el succés en la primera prova i E_2 en la segona, i sigui E l'esdeveniment genèric en qualsevol assaig. Llavors, $P(E_2|E_1)$ es pot calcular de dues formes diferents:

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)} \quad (3.3)$$

o bé

$$P(E_2|E_1) = \sum_i P(E|A_i)P(A_i|E). \quad (3.4)$$

L'equació (3.3) és la definició de probabilitat condicional, mentre que (3.4) segueix de l'actual "lei de probabilitat total", i la demostració és com segueix: Suposant que E_1 i E_2 són condicionalment independents (donada la causa A_i):

$$\begin{aligned} P(E_2|E_1) &= \sum_i P(E_2 \cap A_i|E_1) = \sum_i P(E_2|A_i \cap E_1)P(A_i|E_1) \\ &= \sum_i P(E_2|A_i)P(A_i|E_1) = \sum_i P(E|A_i)P(A_i|E). \end{aligned}$$

Laplace usa la suposició de causes igualment probables (A_i) per obtenir

$$P(E_1) = \frac{1}{n} \sum_i P(E_1|A_i) = \frac{1}{n} \sum_i P(E|A_i)$$

i també

$$\begin{aligned} P(E_1 \cap E_2) &= \frac{1}{n} \sum_i P(E_1 \cap E_2|A_i) = \frac{1}{n} \sum_i P(E_1|A_i)P(E_2|A_i) \\ &= \frac{1}{n} \sum_i P(E|A_i)P(E|A_i) = \frac{1}{n} \sum_i [P(E|A_i)]^2. \end{aligned}$$

Substituint aquest resultat a (3.3) i equiparant-lo a (3.4), s'arriba a

$$\sum_i P(E|A_i)P(A_i|E) = \frac{\sum_i [P(E|A_i)]^2}{\sum_i P(E|A_i)}.$$

De la mateixa manera, considerant la probabilitat que E succeeixi $k+1$ vegades, l'autor obté

$$\sum_i [P(E|A_i)]^k P(A_i|E) = \frac{\sum_i [P(E|A_i)]^{k+1}}{\sum_i P(E|A_i)}$$

per $k = 1, 2, \dots, n-1$. Així, aconseguim $n-1$ equacions lineals a $P(A_i|E)$, a les quals s'afegeix l'equació $\sum_i P(A_i|E) = 1$. Per tant, el sistema es pot resoldre i la solució és

$$P(A_i|E) = \frac{P(E|A_i)}{\sum_j P(E|A_j)}.$$

S'arriba així a la demostració del principi de Laplace exposat anteriorment.

3.1.2 Les memòries de Laplace de 1774

Laplace aborda el problema de l'elecció de les mitjanes per primera vegada el 1772 amb grans restriccions, ja que només s'estén al cas de tres observacions. Amb aquest, topa amb una equació de grau quinze, fet que el porta a considerar les limitacions dels seus resultats i conclou que no mereixen ser publicats. No obstant, el treball conté idees innovadores, com la introducció d'un criteri per escollir una “millor mitjana”, o bé la demostració de que la nova probabilitat inversa es pot aplicar a problemes pràctics no trivials.

La discussió de Laplace sobre probabilitat inversa que es troba a les seves memòries del 1774 [11] va acompanyada de les figures 1, 2 i 3 que apareixen a continuació.

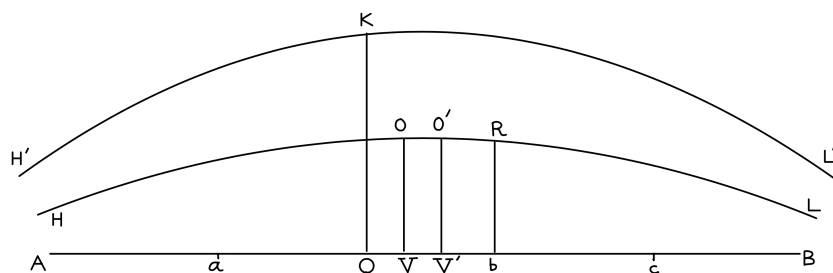


Figura 1: Diagrama de les distribucions posteriors [[11], p. 656].

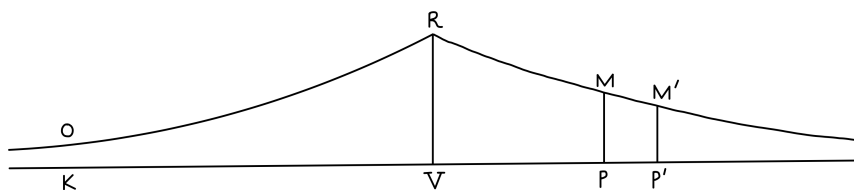


Figura 2: Diagrama de la doble densitat exponencial [[11], p. 656].

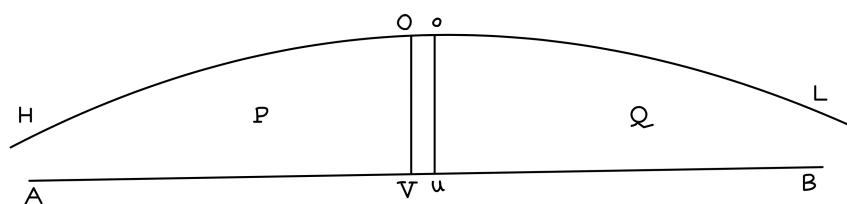


Figura 3: Diagrama que il·lustra que la mediana posterior minimitza l'error esperat posterior [[11], p. 656].

Laplace considera el següent problema: Donats tres temps observats d'un fenomen (a , b i c) al llarg de l'eix temporal AB (Figura 1), trobar el punt V que s'hauria de prendre com a temps real de l'esdeveniment. Aleshores, suposa que si V és l'instant real del fenomen, la probabilitat d'una observació que difereix de la real per una quantitat x ve donada per una corba $y = \phi(x)$ (corresponent a ORM a la Figura 2). Llavors, queda estudiar quina és la corba que s'ha de prendre com la corba d'error $\phi(x)$, i com s'ha de determinar la mitjana, donada $\phi(x)$.

L'autor determina tres condicions que $\phi(x)$ ha de satisfer. Primerament, la corba ha

de ser simètrica respecte a V , ja que és igual de probable que l'observació es desviï cap a la dreta o cap a l'esquerra de la real. Seguidament, ha de decreïxer cap a l'eix KP (Figura 2), ja que la probabilitat de que l'observació difereixi de la veritable per una distància infinita és zero. En tercer lloc, l'àrea sota la corba ha de ser 1, ja que l'observació caurà en un dels punts de la línia KP . Aquestes condicions no determinen x , però sí que adverteixen algunes reflexions generals sobre el problema.

En notació moderna, l'enfocament del problema és com segueix: Donat l'instant V real, la probabilitat de tres observacions, assumint valors a , b i c , és

$$\phi(a, b, c|V) = \phi(a - V)\phi(b - V)\phi(c - V)$$

on $\phi(a, b, c|V)$ correspon a una expressió genèrica de la “densitat de probabilitat” i les expressions de la dreta corresponen als errors de la corba de Laplace. Aleshores, el principi de Laplace (3.1), o teorema de Bayes amb distribució a priori uniforme en V , estipula la distribució de V (donades a, b, c) com

$$\phi(V|a, b, c) \propto \phi(a, b, c|V) \quad (3.5)$$

o bé, de (3.2),

$$\phi(V|a, b, c) = \frac{\phi(a - V)\phi(b - V)\phi(c - V)}{\int \phi(a - U)\phi(b - U)\phi(c - U)dU}. \quad (3.6)$$

El matemàtic procedeix tal i com s'ha exposat, però amb una subtil diferència que el portaria a desenvolupar el principi invers. Reformula el problema en termes de diferències: Siguin $p = b - a$ i $q = c - b$ les diferències entre les observacions, i sigui $x = V - a$ la correcció que s'hauria d'afegir a a per tal de trobar V . La probabilitat de les diferències x , p i q ve donada per

$$\phi(x)\phi(p - x)\phi(p + q - x).$$

Aleshores, invoca el seu principi. Si l'instant “real” es troba a V' enlloc de V , i $x' = V' - a$, aquesta probabilitat esdevé

$$\phi(x')\phi(p - x')\phi(p + q - x').$$

La relació de les probabilitats en què l'instant real sigui V o V' és

$$\phi(x)\phi(p - x)\phi(p + q - x) : \phi(x')\phi(p - x')\phi(p + q - x').$$

Conseqüentment, el seu principi confereix la distribució de probabilitat de la correcció x , donades p i q proporcionals a

$$y = \phi(x)\phi(p - x)\phi(p + q - x),$$

que correspon a la corba *HOL* de la Figura 1.

En termes de l'expressió moderna més genèrica per a la “densitat de probabilitat”, Laplace troba

$$\phi(x, p, q) = \phi(x)\phi(p - x)\phi(p + q - x).$$

El punt V només es considera en el primer error x (quantitat aleatòria no observable). Llavors, x (equivalent a la diferència entre una observació i la real) es converteix en l'objecte d'investigació, la “causa” a trobar dels “esdeveniments” observables p i q . Donada

la causa x , la distribució d'aquests esdeveniments, $\phi(p, q|x)$, és proporcional a $\phi(x, p, q)$, i pel principi de Laplace, la distribució de la causa és també

$$\phi(x|p, q) \propto \phi(x, p, q).$$

El problema es focalitza en l'ús d'aquesta corba de probabilitats de la correcció x per escollir un valor x profitós. En notació actual equival a preguntar-se pel millor valor singular, donada la distribució posterior de l'instant real V . Laplace descriu dos criteris per seleccionar-lo: Primerament, s'hauria de triar un valor de x tal que la mesura real caigui abans o després d'aquest amb la mateixa probabilitat, i ho denomina “mitjana de la probabilitat” (actualment “mediana posterior”); alternativament, s'escull la x que minimitza la suma d'errors, multiplicada per les probabilitats respectives, denominada “mitjana d'error” o “mitjana astronòmica”, ja que considera que aquesta és la que els astrònoms haurien de prioritzar. Avui en dia equival al valor que minimitza la pèrdua esperada posterior, mesurada com el valor absolut de la diferència entre la x triada i la real. A més, demostra que la mitjana de probabilitat i la mitjana astronòmica són idèntiques, i es troba bisecant l'àrea sota la corba *HOL*.

La demostració surt de la Figura 3: Sigui P el centre de gravetat de $AVOH$ i z' la distància de P a OV . Sigui Q el centre de gravetat de $BuoL$ i z la distància de uo a Q . Si la distància des de V fins a u és un infinitèsim dx (amb V el valor seleccionat), la pèrdua esperada posterior és

$$Mz + Nz' + \frac{1}{2}y(dx)^2$$

on M és l'àrea de $BuoL$, N l'àrea de $AVOH$ i y l'altura de OV . Similarment, la pèrdua esperada escollint u és

$$M(z - dx) + N(z' + dx) + \frac{1}{2}y(dx)^2.$$

Si V ha de ser el valor de minimització, aquestes han de ser iguals, i la seva diferència és $Ndx - Mdx$. Per tant, $N = M$.

3.1.3 La corba d'error de 1772-1774

A nivell teòric, Laplace soluciona el problema d'escollir la millor mitjana entre mesures discordants. Per minimitzar l'error esperat, biseca la distribució de probabilitats del valor real V (o equivalentment, de la correcció x de la primera mesura). Aquest tractament és vàlid per qualsevol nombre d'observacions, malgrat el matemàtic es restringeix a només tres observacions. Per precisar una mitjana a partir d'aquest criteri, és necessari concretar la forma d'una corba d'errors $\phi(x)$, fet que esdevé un gran repte. L'autor fa dues grans aportacions en aquest àmbit, cap de les quals la considera totalment exitosa. El primer intent es publica el 1772, en el qual es fa notòria la necessitat de restringir-se a tres observacions, i el 1777 publica el segon tempteig. En ambdós casos, Laplace assoleix uns avenços matemàtics remarcables per l'època.

Determinar una corba d'error és relativament fàcil. No obstant, el problema recau en la tria de la corba correcta per poder obtenir la mitjana. Les dues derivacions més primerenques de Laplace es desenvolupen en el paradigma del principi de casos igualment probables. En el seu primer treball considera i descarta dues possibles corbes abans de decidir-se per la que actualment porta el seu nom. Inicialment, nota que si els punts

P i P' de la Figura 2 són igualment probables, la corba $\phi(x)$ ha de ser una constant, fet que s'oposa a la suposició que els errors grans són menys probables que els petits. Posteriorment, considera (i també descarta) que $\phi(x)$ disminueix de forma constant a mesura que x augmenta.

Finalment, troba una hipòtesi acceptable. Suposa que tant $\phi(x)$ com $\frac{d\phi(x)}{dx}$ decreixen amb x (per x positives), i vol determinar la relació entre els seus ritmes de disminució. Si la proporció entre dues diferències consecutives infinitament petites és igual a les ordenades corresponents (ja que ara es considera la mateixa llei per les ordenades que per les seves diferències), aleshores

$$\frac{d\phi(x+dx)}{d\phi(x)} = \frac{\phi(x+dx)}{\phi(x)}.$$

Llavors,

$$\frac{d\phi(x)}{dx} = -m\phi(x) \Rightarrow \phi(x) = Ce^{-mx}.$$

La constant C s'obté de prendre l'àrea de la corba ORM (Figura 2) igual a la unitat, i s'arriba a $C = \frac{m}{2}$. El resultat és

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

on e és el número amb logaritme hiperbòlic igual a la unitat. S'ha suposat que x es pren positiva, però en notació actual s'hauria d'especificar el valor absolut

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

Laplace desenvolupa una expressió elegant per una corba d'error, que actualment es coneix com a “distribució exponencial doble” o “distribució de Laplace”. Limita la seva discussió a únicament tres mesures (denominades a, b i c en aquest plantejament), degut a les dificultats que troba quan intenta avaluar la mitjana a partir de bisecar la corba HOL (Figura 3).

Sense pèrdua de generalitat, comença suposant $a < b < c$. Avalua la integral de $\phi(x)\phi(p-x)\phi(p+q-x)$, considerant que cada una de les quatre regions que determinen les observacions respectives presenten diferents formes. Aleshores, procedeix a buscar el millor valor de x , d'acord al seu criteri. Pel cas $p = b - a > q = c - b$, obté

$$x = p + \frac{1}{m} \ln \left(1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right). \quad (3.7)$$

A partir d'aquest resultat, es pot concloure que la mitjana aritmètica no és el millor mètode per combinar observacions, ja que aquesta correspondria a la correcció

$$x = \frac{2p+q}{3},$$

un resultat que només s'obté de (3.7) quan $m = 0$, o quan és infinitament petita, i s'aconsegueix de la següent manera:

$$\lim_{m \rightarrow 0} \frac{\ln \left(1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right)}{m} \approx \lim_{m \rightarrow 0} \frac{\frac{1}{3}(-p)e^{-mp} - \frac{1}{3}(-q)e^{-mq}}{1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq}} = \frac{-p+q}{3}$$

i per tant,

$$x = p + \frac{-p + q}{3} = \frac{2p + q}{3}.$$

Laplace rebutja aquest valor per m , ja que tots els punts de la línia KP igualment probables (Figura 2) es trobarien a una distància extremadament gran, cosa absurda. D'altra banda, aquest resultat (3.7) també presenta un problema, ja que el millor valor de x depèn del valor indeterminat de m . Laplace aborda aquest problema aplicant el principi de probabilitat inversa.

La distribució de probabilitat de x , p i q , donada m , és $\phi(x)\phi(p-x)\phi(p+q-x)$. Aleshores, la distribució de p i q donada m és la integral d'aquesta funció respecte a x , i pel seu principi la distribució de m , donades p i q , és proporcional a aquesta integral, ja que es compleix (3.5). Anteriorment, ja havia tobat aquesta integral (l'àrea sota *HOL* a la Figura 3):

$$\frac{m^2}{4} e^{-m(p+q)} \left(1 - \frac{1}{3} e^{-mp} - \frac{1}{3} e^{-mq} \right).$$

Per tant, la probabilitat de m és proporcional a

$$m^2 dm \cdot e^{-m(p+q)} \left(1 - \frac{1}{3} e^{-mp} - \frac{1}{3} e^{-mq} \right).$$

En notació moderna es pot descriure la seva cadena de raonaments tal i com segueix:

$$\phi(x, p, q|m) = \phi(x)\phi(p-x)\phi(p+q-x)$$

i per tant,

$$\phi(p, q|m) = \int_{-\infty}^{\infty} \phi(x)\phi(p-x)\phi(p+q-x)dx.$$

Pel principi de Laplace, com que

$$\phi(m|p, q) \propto \phi(p, q|m)$$

aleshores

$$\begin{aligned} \phi(m|p, q) &\propto \int_{-\infty}^{\infty} \phi(x)\phi(p-x)\phi(p+q-x)dx \\ &\propto m^2 e^{-m(p+q)} \left(1 - \frac{1}{3} e^{-mp} - \frac{1}{3} e^{-mq} \right). \end{aligned}$$

Llavors, la probabilitat que $m = 0$ o que sigui infinitament petita és immensament menor que la probabilitat que m sigui igual a qualsevol quantitat finita (diferent de zero).

3.1.4 L'origen de la probabilitat inversa

Ara, l'objectiu de Laplace és estimar la correcció x òptima, amb m desconeguda, i és aquí on el matemàtic comet un subtil error. Referent a la Figura 1, sigui y la probabilitat (corresponent a m) d'obtenir l'instant real del fenomen a una distància x del punt a , aleshores la probabilitat total és proporcional a

$$\int y \cdot m^2 dm \cdot e^{-m(p+q)} \left(1 - \frac{1}{3} e^{-mp} - \frac{1}{3} e^{-mq} \right) \quad (3.8)$$

on la integral va de $m = 0$ a $m = \infty$. A continuació, construint una nova corba $H'KL'$ (Figura 1) sobre l'eix AB (amb ordenades proporcionals a aquesta quantitat), l'ordenada KQ que divideix l'àrea d'aquesta corba en dues parts iguals talla l'eix en el punt que s'hauria de prendre com la mitjana entre les tres observacions. L'àrea d'aquesta nova corba és proporcional a la integral del producte de l'àrea de la corba HOL per

$$m^2 dm \cdot e^{-m(p+q)} \left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right).$$

Per determinar x , sota una suposició particular per m , es calcula

$$m^2 e^{-m(2p+q-x)} = m^2 e^{-m(p+q)} \left(1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right)$$

i per tant s'arriba a

$$\begin{aligned} & \int m^4 dm \cdot e^{-m(3p+2q-x)} \left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right) \\ &= \int m^4 dm \cdot e^{-m(2p+2q)} \left(1 + \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right) \left(1 - \frac{1}{3}e^{-mp} - \frac{1}{3}e^{-mq} \right) \end{aligned}$$

on les integrals van de $m = 0$ a $m = \infty$.

En notació moderna, aquests resultats s'interpreten tal i com segueix: y equival a $\phi(x|p, q, m)$, i la integral (3.8) correspon a

$$\int_0^\infty \phi(x|p, q, m) \phi(m|p, q) dm.$$

Aleshores, la mediana posterior (x_0) s'obté resolent l'equació

$$\int_{-\infty}^{x_0} \int_0^\infty \phi(x|p, q, m) \phi(m|p, q) dm dx = \frac{1}{2} \int_{-\infty}^\infty \int_0^\infty \phi(x|p, q, m) \phi(m|p, q) dm dx. \quad (3.9)$$

Com que

$$\int_0^\infty \phi(x|p, q, m) \phi(m|p, q) dm = \int_0^\infty \phi(x, m|p, q) dm = \phi(x|p, q)$$

llavors l'equació (3.9) és equivalent a

$$\int_{-\infty}^{x_0} \phi(x|p, q) dx = \frac{1}{2} \int_{-\infty}^\infty \phi(x|p, q) dx \quad (3.10)$$

i l'afirmació de Laplace és força correcta. Si el matemàtic hagués trobat i solucionat l'equació (3.9), hauria topat amb la solució del problema, ja que podem veure que (3.10) equival a la funció de distribució d'una funció normal estàndard. No obstant, aquest afirma, erròniament, que “evidentment” l'àrea d'aquesta nova corba,

$$\int_{-\infty}^\infty \int_0^\infty \phi(x|p, q, m) \phi(m|p, q) dm dx$$

és proporcional a la integral de l'àrea de la corba HOL ,

$$\int \phi(x, p, q|m) dx$$

multiplicada per $\phi(m|p, q)$, és a dir,

$$\int_{-\infty}^{\infty} \int_0^{\infty} \phi(x, p, q|m) \phi(m|p, q) dx dm. \quad (3.11)$$

Això seria cert si

$$\phi(x|p, q, m) \propto \phi(x, p, q|m) \quad (3.12)$$

que es compleix només si la constant de proporcionalitat de la relació (3.12) depèn tant de m com de p i q , ja que

$$\phi(x|p, q, m) = \frac{\phi(x, p, q|m)}{\phi(p, q|m)}.$$

No obstant, aquesta substitució no està permesa en una integral respecte a m (ja que, a l'integrar, desapareix la dependència respecte a aquesta variable), de manera que l'afirmació és falsa i, conseqüentment, també ho és l'afirmació que segueix. En notació actual, sigui m coneguda, x_0 ve determinada per

$$\int_{-\infty}^{x_0} \phi(x, p, q|m) dx = \frac{1}{2} \int_{-\infty}^{\infty} \phi(x, p, q|m) dx$$

que es troba a través de

$$\int_{-\infty}^{x_0} \int_0^{\infty} \phi(x, p, q|m) \phi(m|p, q) dm dx = \frac{1}{2} \int_{-\infty}^{\infty} \int_0^{\infty} \phi(x, p, q|m) \phi(m|p, q) dm dx.$$

El problema recau en la substitució de forma agosarada de la densitat conjunta per la densitat condicional, i viceversa. Usant de forma repetida les proporcionalitats

$$\phi(u, v) \propto \phi(u|v) \quad \text{i} \quad \phi(u, v) \propto \phi(v|u)$$

desapareix la dependència de les constants de proporcionalitat com a funcions d'una de les variables. En canvi, si Laplace hagués treballat amb les declaracions completes

$$\phi(u, v) = \phi(u|v) \phi(v) \quad \text{i} \quad \phi(u, v) = \phi(v|u) \phi(u)$$

la dependència s'hauria mantingut i hauria evitat aquest error. Així doncs, no és correcte acceptar com a certa l'expressió (3.11).

Laplace, el 1774, no és coneixedor de la notació escaient de probabilitat condicional. D'haver procedit de forma adequada, segurament hauria arribat a un problema analític molt més simple. Per a les seves distribucions prèvies,

$$\phi(x|p, q, m) \propto \frac{\phi(x, p, q|m)}{\phi(m|p, q)}$$

on la constant de proporcionalitat depèn només de p i q . El problema es redueix a resoldre

$$\int_{-\infty}^{x_0} \int_0^{\infty} \phi(x, p, q|m) dm dx = \frac{1}{2} \int_{-\infty}^{\infty} \int_0^{\infty} \phi(x, p, q|m) dm dx$$

o bé

$$\int_0^{\infty} m^2 e^{-m(2p+q-x_0)} dm = \int_0^{\infty} m^2 e^{-m(p+q)} \left(1 + \frac{1}{3} e^{-mp} - \frac{1}{3} e^{-mq} \right) dm,$$

un parell d'integrals més simples que les obtingudes per Laplace.

L'equació correcta es redueix a una cúbica en x_0 ,

$$(2p + q - x_0)^3 = \left[(p + q)^{-3} + \frac{(2p + q)^{-3}}{3} - \frac{(p + 2q)^{-3}}{3} \right]^{-1}.$$

3.1.5 Les memòries de Laplace de 1777-1781

En el context històric dels desenvolupaments de Laplace, una corba d'error teòrica només té sentit si es deriva del primer principi de la teoria de la probabilitat, el qual es redueix a l'estudi de casos igualment probables. Si e^{-mx} no funciona, Laplace ha de buscar una alternativa per aplicar el principi de raó insuficient, la qual desenvolupa el 1777.

Sembla ser que Laplace reconsidera el problema com a conseqüència de la publicació de les memòries de Lagrange [10] de 1776. Aquest últim usa el càlcul de probabilitats per estudiar els avantatges de prendre la mitjana d'entre els resultats de diferents observacions. Des del punt de vista estadístic, aquest és un treball força limitat, però destaca el fet que l'autor es centra en la mitjana aritmètica i la seva distribució de probabilitats, on destaca la tècnica que utilitza per determinar aquesta última.

L'estudi de Lagrange es basa en la mitjana aritmètica. Pren fixes la distribució d'error i la quantitat a determinar, i constata que el valor d'aquesta mitjana s'ha de jutjar en termes d'afirmacions probabilístiques sobre els possibles valors. També determina la correcció del biaix necessària quan el centre de gravetat de la distribució és diferent de zero, i tracta breument l'estimació d'una distribució d'errors discreta desconeguda.

Laplace, el 1777, estableix dos punts de vista diferents per abordar el problema, depenent de si es consideren les observacions abans o després d'haver-se produït. En el primer cas, per trobar la mitjana, s'ha de determinar a priori la millor funció dels resultats de les observacions. En el segon cas, la mitjana sorgeix estipulant una funció similar a posteriori, tenint en compte les distàncies entre les respectives observacions. Aquestes dues aproximacions porten a resultats diferents, i Laplace afirma que la segona opció és la correcta.

Les memòries de Laplace de 1777 [12] i les de 1781 [13] es centren en gran mesura a l'explicació d'un argument d'extremada complexitat per tal de justificar que la corba

$$y = \frac{1}{2a} \log \left(\frac{a}{|x|} \right), \quad |x| \leq a,$$

s'ha de prendre com una distribució d'errors, on a representa el límit superior dels errors possibles.

L'argument de l'autor pretén arribar a una densitat unimodal, simètrica i aleatòria. En terminologia actual, el procediment és com segueix: Sigui l'interval unitari dividit en $n + 1$ espaiats, ordenats en ordre decreixent:

$$d_1 > d_2 > \dots > d_{n+1} \quad \text{on} \quad d_1 + d_2 + \dots + d_{n+1} = 1$$

S'escullen n punts aleatoris de l'interval. Si n és molt gran, i s'eleva les ordenades d'una corba a les alçades $d_1, d_2, \dots, d_{n+1}$ (igualment espaiades al llarg de l'abscissa entre 0 i a), aleshores cal calcular els valors esperats d'aquestes ordenades (amb n tendint a infinit). El resultat és una corba proporcional a $\log \left(\frac{a}{x} \right)$ a l'interval de 0 a a . Per simetria, s'obté una corba proporcional a $\log \left(\frac{a}{|x|} \right)$ en $(-a, a)$. Amb la condició d'àrea total igual a 1, s'arriba a

$$\frac{1}{2a} \log \left(\frac{a}{|x|} \right). \quad (3.13)$$

Basant el seu argument en una distribució de probabilitat aleatòria, Laplace manté el principi de raó insuficient. A priori, no hi ha cap distribució preferida. A les memòries del

1781, admet que només en investigacions molt “delicades”, com seria el trànsit de Venus a través del Sol, es realitzaria una anàlisi amb aquesta corba (3.13).

3.2 Thomas Bayes

Thomas Bayes (1701 - Royal Tunbridge Wells, Anglaterra, 1761) va ser un estadístic, filòsof i ministre presbiterià anglès, conegut per la formulació d'un cas concret del teorema que porta el seu nom: “Teorema de Bayes”. El matemàtic no va arribar a publicar mai els seus desenvolupaments més transcendents, sinó que aquests es van editar pòstumament.

Era fill d'una família inconformista molt poderosa de Sheffield, el pare del qual (Joshua Bayes) era el ministre presbiterià de Londres. El 1719 es matriculà a la universitat d'Edimburg per estudiar lògica i teologia, i més tard va exercir com a ministre de la capella de *Mount Sion* fins el 1752. Va publicar dos treballs al llarg de la seva vida, un de teològic, titulat *Divine Benevolence, or an Attempt to Prove That the Principal End of the Divine Providence and Government is the Happiness of His Creatures* (1731); i l'altre matemàtic, *An Introduction to the Doctrine of Fluxions, and a Defence of the Mathematicians Against the Objections of the Author of The Analyst*, publicat de forma anònima el 1736, en el qual defensava el fonament lògic del càlcul d'Isaac Newton davant les crítiques de George Berkeley, autor de *The Analyst*. Bayes fou elegit membre de la *Royal Society* el 1742.

Durant els darrers anys de la seva vida s'interessà molt per la probabilitat, afecció que segurament va sorgir de la revisió dels treballs escrits per Thomas Simpson el 1755, o bé gràcies a l'aprenentatge de matemàtiques i probabilitat a través d'un llibre de De Moivre. Els seus treballs i conclusions sobre la teoria de la probabilitat van ser manuscrits i publicats pel seu amic, Richard Price, després de la seva mort.

3.2.1 La “taula de billar” de Bayes

Les observacions astronòmiques motiven Laplace a desenvolupar la probabilitat inversa, i segurament aquestes també porten Bayes a considerar la mateixa qüestió. Ambdós són referents pel què fa a la revolució copernicana en el marc de l'estadística, abastant el problema original de Jacob Bernoulli. L'objectiu d'aquest últim era mostrar com es podia aprendre de l'experimentació amb la natura. La solució al problema es faria notòria durant l'última meitat del segle XVIII.

Thomas Bayes aconsegueix el primer resultat encertat, amb algunes limitacions, després de l'article de Simpson de 1755 [25]. No obstant, en un inici, els seus resultats no tenen un impacte destacable, malgrat ser publicats a *Philosophical Transactions* [1] el 1764. El tractat de Bayes “*Cap a la resolució d'un problema en la doctrina de l'atzar*” [1], avui en dia és de lectura complexa, ja que usa la descripció geomètrica de Newton. Però tot i la difícil comprensió de la solució, el problema s'expressa de forma clara i concisa:

Given the number of times in which an unknown event has happened and failed: Required the chance that the probability of its happening in a single trial lies somewhere between any two degrees of probability that can be named. [[26], p. 123]

En notació moderna, el problema és: Sigui X el número de vegades que succeeix l'esdeveniment en n assajos, i θ la probabilitat que esdevingui en un únic intent. Aleshores, Bayes vol trobar $P(a < \theta < b|X)$, coneixent X .

Per a la solució del problema, comença demostrant les proposicions (3.14) i (3.15). Si E_1 i E_2 són esdeveniments ordenats en el temps (E_1 anterior a E_2), aleshores es compleix

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}. \quad (3.14)$$

Seguidament, presenta una demostració diferent per a

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}, \quad (3.15)$$

corresponent al teorema de Bayes en la seva primera aproximació. És interessant destacar el fet que l'autor considera que (3.14) i (3.15) són fonamentalment diferents, i per tant requereixen les seves respectives demostracions.

Bayes tracta el problema del binomi en termes d'un anàleg físic que difereix de l'urna de Bernoulli, distinció que fou crucial en l'èxit del primer. Planteja una taula quadrada ($ABCD$ a la Figura 4) i dues boles W i O . Aleshores, es llança la bola W al llarg de la taula, de manera que la probabilitat d'aturar-se a qualsevol punt de la taula sigui la mateixa (θ correspon a la posició de W). Seguidament, es fa rodar la bola O n vegades, i es compta el nombre de vegades que s'atura a la dreta de W (valor representat per X).

En el diagrama de Bayes, W determina una línia os paral·lela als costats BC i AD , i θ és la relació entre Ao i AB ($\theta = \frac{Ao}{AB}$); X és el número de vegades que O s'atura en el rectangle $osDA$ (esdeveniment M).

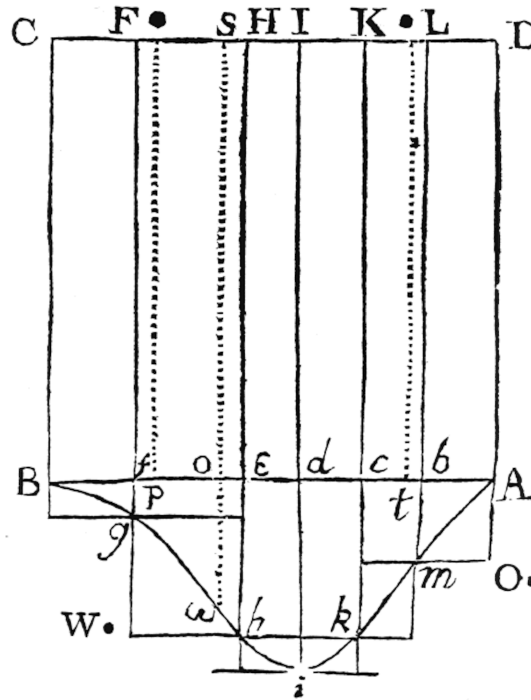


Figura 4: “Taula de billar” de Bayes: Taula quadrada $ABCD$ amb dues boles W i O . La bola W es fa rodar i s'atura a la línia os . Seguidament, la bola O es fa rodar n vegades i es compta el número de vegades que s'atura dins el rectangle $osDA$. La corba inferior mostra la densitat posterior de la fracció Ao/AB , donat el recompte registrat. També es mostren les aproximacions superiors i inferiors a aquesta densitat. [[1], p. 385]

Bayes puntualitza que les posicions finals de les boles estan uniformement distribuïdes sobre el pla, i per tant θ també ho està a l'interval unitat. Amb θ determinat, X presenta una distribució binomial, on θ és la probabilitat d'èxit en un únic intent. Així, la distribució de θ perd el caràcter discret que tenia a l'urna de Bernoulli, i el problema adopta un nou caràcter simètric.

La resolució al problema és com segueix: Sigui W el primer esdeveniment (E_1) i sigui n (número de rodaments d' O) el segon esdeveniment (E_2). De la distribució uniforme de W i el fet que, donada θ (posició de W), la distribució de X (número d'esdeveniments M) és binomial, Bayes usa (3.14) per derivar la probabilitat conjunta de que θ es trobi entre b i f , i M succeeixi p vegades (és a dir, $X = p$). En notació actual,

$$P(\{b < \theta < f\} \cap \{X = p\}) = \int_b^f \binom{n}{p} \theta^p (1 - \theta)^{n-p} d\theta. \quad (3.16)$$

De fet, descriu aquesta probabilitat com l'àrea $fghikmb$ a la Figura 4, on la corba dibuixada és $y = Ex^p r^q$ (x correspon a θ , r a $1 - \theta$, $E = \binom{n}{p}$, i $q = n - p$). Aleshores, prenent $b = 0$ i $f = 1$,

$$P(X = p) = \int_0^1 \binom{n}{p} \theta^p (1 - \theta)^{n-p} d\theta.$$

Més tard trobaria el resultat de la integral, que és $\frac{1}{n+1}$. Usant (3.15), arriba a

$$P(b < \theta < f | X = p) = \frac{P(\{b < \theta < f\} \cap \{X = p\})}{P(X = p)} = \frac{\int_b^f \binom{n}{p} \theta^p (1 - \theta)^{n-p} d\theta}{\int_0^1 \binom{n}{p} \theta^p (1 - \theta)^{n-p} d\theta}.$$

La demostració és remarcable pel fet que tracta les densitats de probabilitat com a àrees. Per trobar (3.16), s'aproxima la corba per sobre i per sota per rectangles (funcions lineals aproximades a trossos a la Figura 4) i s'obté que la probabilitat (3.16) es troba entre les àrees delimitades, respectivament.

Bayes justifica que, en situacions en les quals a priori es desconeix la probabilitat θ que succeeixi un esdeveniment, aleshores aquesta es pot prendre uniformement distribuïda, anàlogament a la bola sobre la taula de billar, i usar el resultat (3.16). Aquest resultat es coneix com a “principi de raó insuficient”: Com que no tenim cap raó per afavorir un valor θ sobre un altre, intervals d'igual valor es poden descriure com a igualment probables.

Originalment, pren la distribució uniforme de θ com un postulat axiomàtic, però després es decanta per un enfocament diferent. Per establir una analogia entre la taula de billar i altres aplicacions on la probabilitat θ sigui desconeguda, és necessari determinar el punt lògic en el qual es verifica aquest paral·lelisme. Bayes afirma que és X , i no θ , la que presenta una distribució uniforme.

Per a la taula de billar, és clar de la descripció anterior que θ està distribuïda uniformement, i una conseqüència matemàtica d'aquest fet és que la distribució marginal de X és uniforme

$$P(X = p) = P(\{0 < \theta < 1\} \cap \{X = p\}) = \int_0^1 \binom{n}{p} \theta^p (1 - \theta)^{n-p} d\theta = \frac{1}{n+1},$$

independentment de p .

Una altra aplicació de probabilitat inversa consisteix en considerar un partit de tennis de n sets, del qual es desconeix la probabilitat que el jugador 1 guanyi algun set. Per tant, d'entre els n intents, no preval un valor de X per sobre d'un altre. Així doncs, aplicant el resultat de Bayes, com que es desconeix el valor de θ , s'obté

$$P(X = p) = \frac{1}{n + 1},$$

per a tot p .

Un possible raonament de Bayes és: Si existeixen p^* i p tals que

$$P(X = p^*) > P(X = p),$$

aleshores $X = p^*$ és més probable que $X = p$. Llavors, esperàriem que $X = p^*$ fos més probable que $X = p$, i per tant la nostra “expectativa” per $\frac{X}{n}$ (estimació normal de θ) seria major per un valor que per un altre, cosa que contradiu el fet que suposem que desconeixem absolutament el valor de θ . Aquest és un argument semblant al principi de raó insuficient, però en aquest cas s'aplica a l'observable empíric X enlloc de θ .

“No saber res” equival a

$$P(X = p) = \frac{1}{n + 1}. \quad (3.17)$$

En un partit de tennis en el qual “no sabem res” a priori, tots els resultats $X = j$, amb $j = 0, 1, \dots, n$, (i per tant totes les proporcions $\frac{X}{n}$) són igualment probables. Bayes afirma que si podem reconèixer una situació que s'ajusta a aquesta descripció, aleshores és una situació anàloga a la taula de billar, i es pot aplicar el raonament derivat. Si se satisfà (3.17), aleshores

$$P[f(X) = f(p)] = \frac{1}{n + 1}$$

per qualsevol funció estrictament monòtona f . Desconèixer X equival a desconèixer $f(X)$, ja que X (a diferència de θ) és discreta.

Bayes ignorava que el raonament depèn d'una subtileza matemàtica. La condició (3.17), per a tot p , requereix d'una distribució uniforme per θ , però no és una condició suficient, a no ser que es compleixi per a tot p i per a tot n . En qualsevol esdeveniment, l'argument està estrictament limitat a la situació binomial.

Tot i el canvi de paradigma de Bayes per caracteritzar una distribució a priori de l'observable X en lloc d'un θ no observable, l'autor es mostra reticent a fer públics els seus resultats degut a un aspecte matemàtic, relacionat amb l'avaluació de la integral (3.16).

La integral

$$\int_0^f \theta^p (1 - \theta)^q d\theta \quad (3.18)$$

és actualment coneguda com a “funció beta incompleta”, i la seva dificultat recau en els valors de p i q , ja que, quan són petits, la resolució no presenta complicacions, però per valors grans es converteix en un problema formidable. Pel cas en què q és petit, s'expandeix $(1 - \theta)^q$ per la sèrie binomial, es multiplica per θ^p i s'integra terme a terme. No obstant, pels casos més importants en què p i q són ambdós grans, aquest procediment no és factible, ja que les sèries resultants convergeixen molt lentament. Així doncs, Bayes considera que el seu treball està incomplet i que, per tant, no confereix una solució útil a problemes pràctics.

3.2.2 L'anàlisi de Laplace

El tractament de Bayes sobre el binomi no despertà massa interès en el moment de la seva publicació, degut a ser matemàticament incomplet. Una dècada més tard, Laplace topa amb el mateix problema, però aquest disposa de millors habilitats analítiques.

L'estudi de Laplace es prolonga quatre dècades, des de la seva primera aproximació en els treballs de probabilitat inversa de 1774 [11] fins a la publicació més detallada de *Théorie analytique des probabilités* [17] de 1812. Després d'introduir el principi de probabilitat inversa, el matemàtic enuncia i demostra un bonic teorema. Suposa que si es treuen $p + q$ tiquets d'una urna que conté tiquets blancs i negres amb una proporció desconeguda x , aleshores s'obtenen p tiquets blancs i q tiquets negres. El problema recau en determinar la proporció desconeguda x . El teorema és el següent:

One can suppose that the numbers p and q are so large that it becomes as close to certainty as one wishes that the ratio of the number of white tickets to the total number of tickets contained in the urn is included between the two limits $\frac{p}{p+q} - w$ and $\frac{p}{p+q} + w$. w can be supposed less than any given quantity. [[11], p. 131]

En notació actual, Laplace afirma la consistència de la raó empírica $\frac{p}{p+q}$, ja que, augmentant el nombre d'assajos, es pot garantir que $\frac{p}{p+q}$ s'aproxima a x . És a dir,

$$P\left\{\left|x - \frac{p}{p+q}\right| \leq w | p, q\right\} \rightarrow 1, \quad \text{si } p + q \rightarrow \infty, \quad \text{per a qualsevol } w > 0.$$

La probabilitat en qüestió és respecte a la distribució (a posteriori) x , condicional a p i q .

Laplace comença aplicant el seu principi (3.6) a la situació binomial, i conclou que la distribució posterior de la proporció desconeguda x és

$$\frac{x^p(1-x)^q dx}{\int_0^1 x^p(1-x)^q dx} = \frac{(p+1)(p+2) \cdots (p+q+1)}{1 \cdot 2 \cdot 3 \cdots q} \cdot x^p(1-x)^q dx. \quad (3.19)$$

Aleshores, el problema esdevé el mateix que havia turmentat Bayes en el seu moment, és a dir, avaluar la integral de l'expressió (3.19) des de $\frac{p}{p+q} - w$ fins a $\frac{p}{p+q} + w$. Laplace desenvolupa el "mètode de Laplace per a l'avaluació asimptòtica d'integrals definides". L'integrand $x^p(1-x)^q$ és com a màxim $\frac{p}{p+q}$, de forma que expandeix $x^p(1-x)^q$ al voltant d'aquest punt. Sigui

$$x = \frac{p}{p+q} + z$$

i

$$\begin{aligned} \int_{\frac{p}{p+q}-w}^{\frac{p}{p+q}+w} x^p(1-x)^q dx &= \frac{p^p q^q}{(p+q)^{p+q}} \int \left(1 + \frac{p+q}{p} z\right)^p \left(1 - \frac{p+q}{q} z\right)^q dz \\ &\cong \frac{p^p q^q}{(p+q)^{p+q}} \int_0^w 2 \cdot \exp\left[-\frac{(p+q)^3}{2pq} z^2\right] dz \end{aligned}$$

on l'última igualtat s'obté a partir d'una anàlisi exhaustiva per part de l'autor. Aleshores, multiplicant als dos costats per

$$\frac{(p+1)(p+2) \cdots (p+q+1)}{1 \cdot 2 \cdot 3 \cdots q}$$

i aplicant la fórmula de Stirling, s'obté, en notació moderna,

$$P\left(\left|\frac{p}{p+q} - x\right| \leq w|p, q\right) \cong \frac{2}{\sqrt{2\pi\sigma^2}} \int_0^w \exp\left(-\frac{z^2}{2\sigma^2}\right) dz = \frac{2}{\sqrt{2\pi}} \int_0^{\frac{w}{\sigma}} \exp\left(-\frac{u^2}{2}\right) du$$

on

$$\sigma^2 = \frac{pq}{(p+q)^3}.$$

Aquest resultat correspon a l'aproximació normal a la distribució beta posterior, però aquest no és l'objectiu de Laplace el 1774. Per tant, avalua la integral

$$\int_0^\infty e^{-\frac{u^2}{2}} du = \sqrt{\frac{\pi}{2}}$$

i estipula que, sempre que $p+q \rightarrow \infty$ (i per tant $\sigma^2 \rightarrow 0$, doncs $\frac{w}{\sigma} \rightarrow \infty$), es compleix

$$P\left(\left|\frac{p}{p+q} - x\right| \leq w|p, q\right) \rightarrow 1.$$

El 1781, Laplace amplia la seva anàlisi del 1774 per obtenir aproximacions de sèries arbitràriament precises a la integral beta. Es centra en la funció beta incompleta, la integral (3.18) per p i q grans, i manipula les sèries per poder treballar amb elles. No obstant, l'esforç realitzat és immens i els resultats són poc elegants. El 1785 publica una memòria [15] que tracta més generalment amb integrals de productes de funcions elevades a potències grans, i es centra en aproximacions normals a aquestes integrals. Per facilitar l'avaluació de les mateixes, deriva una sèrie per a la integral normal. Per T petita,

$$\int_0^T e^{-t^2} dt = T - \frac{1}{3}T^3 + \frac{1}{1 \cdot 2} \cdot \frac{T^5}{5} - \frac{1}{1 \cdot 2 \cdot 3} \cdot \frac{T^7}{7} + \frac{1}{1 \cdot 2 \cdot 3 \cdot 4} \cdot \frac{T^9}{9} - \dots$$

Per T gran,

$$\int_T^\infty e^{-t^2} dt = \frac{e^{-T^2}}{2T} \left(1 - \frac{1}{2T^2} + \frac{1 \cdot 3}{2^2 T^4} - \frac{1 \cdot 3 \cdot 5}{2^3 T^6} + \dots\right).$$

Així, Laplace proporciona les eines necessàries per a la inferència sobre una propietat binomial. Usa aquests resultats per primera vegada en l'anàlisi de les estadístiques sobre la proporció de sexes en néixer. De les *Mémoires* de l'Acadèmia de 1771, troba que, des del 1745 fins al 1770, havien nascut a París 251.527 nens i 241.945 nenes, i estudia si això és indicatiu d'una major propensió al naixement masculí. Pren x com la probabilitat que un naixement donat sigui masculí i, d'una aplicació directa del seu anàlisi, troba la probabilitat posterior (en notació moderna):

$$P\left(x \leq \frac{1}{2} | p = 251.527, q = 241.945\right) = 1.1521 \times 10^{-42}.$$

Per tant, Laplace considera cert que $x > \frac{1}{2}$. A Londres, el predomini de mascles era encara major, ja que des del 1664 fins al 1757 havien nascut 737.629 nens i 698.958 nenes. Aleshores, es pregunta si la probabilitat d'un naixement masculí a París és la mateixa que a Londres, on les freqüències relatives eren de 0,50971 i 0,51346, respectivament.

Sigui u la probabilitat de naixements masculins a París i s la probabilitat de naixements masculins a Londres, Laplace estén la seva anàlisi al càlcul de la probabilitat posterior:

$$P(u > s | \text{dades}) = \frac{1}{410.458}.$$

Conclou que és bastant probable que a Londres, més que a París, hi hagués alguna causa que facilités el naixement de nens (com podria ser el clima, el menjar o els costums). Quan un càlcul similar, comparant la probabilitat de naixements masculins a París amb el del Regne de Nàpols (0,51162), dona una probabilitat posterior d'aproximadament $\frac{1}{100}$, Laplace rebut que aquesta probabilitat no és prou extrema per a una declaració irrevocable.

3.2.3 El Teorema Central del Límit

El resultat més important de Laplace en teoria de probabilitats és el teorema central del límit, publicat el 1810. Es pot considerar com una generalització del teorema del límit de De Moivre. Qualsevol suma o mitjana (no només el nombre total d'èxits amb n assajos) serà aproximadament una distribució normal, sempre que el número de termes sigui elevat. Les consideracions de regularitat i els casos particulars apareixerien més tard.

No es coneix exactament el moment en el qual Laplace desenvolupa aquest treball. L'anàlisi es basa en l'ús de l'actual “transformada de Fourier” o funcions característiques, conseqüència directa d'una tècnica que Lagrange usa el 1776 (l'avui coneguda “transformada de Laplace”), que és un desenvolupament de les funcions generadores de De Moivre i Simpson. No obstant, la clau en el triomf de Laplace és una fórmula d'inversió ideada pel propi autor. De Moivre, Simpson i Lagrange sabien que (en notació moderna) si $X_1, X_2, \dots, X_n$ són variables aleatòries independents i

$$P(X_i = k) = p_k$$

aleshores la probabilitat que la suma

$$S_n = X_1 + X_2 + \dots + X_n$$

prengui el valor k és el coeficient de s^k en l'expansió de

$$(p_0 + sp_1 + s^2p_2 + \dots)^n.$$

La pregunta recau en com recuperar aquest coeficient d'una forma que sigui susceptible d'anàlisi.

El 1782, Laplace escriu una memòria extensiva [14] de generació de funcions, i en fa ús en la resolució d'equacions diferencials. El 1785, considera la següent funció de generació, que en notació moderna pren la forma

$$(s^{-1}p_{-1} + sp_1)^n = s^{-n}a_{-n} + s^{-n+1}a_{-n+1} + \dots + a_0 + \dots + a_ns^n$$

on $p_{-1} = p_1 = 1$, i estima el problema d'extreure el coeficient a_k de s^k a l'expansió. Fixant $s = e^{t\sqrt{-1}}$,

$$(e^{-t\sqrt{-1}} + e^{t\sqrt{-1}})^n = e^{-nt\sqrt{-1}}a_{-n} + \dots + a_0 + \dots + a_ne^{nt\sqrt{-1}}$$

el problema de recuperar el terme del mig (suposant n parell) consisteix en trobar el terme constant (i per tant lliure de s (o t)). És interessant mencionar que l'aparició de $\sqrt{-1}$ actualment equival a i . Com que

$$e^{t\sqrt{-1}} = \cos t + \sqrt{-1} \sin t$$

els termes no constants es poden eliminar integrant l'expressió que segueix:

$$s^k = e^{kt\sqrt{-1}} = \cos(kt) + \sqrt{-1} \sin(kt)$$

i per tant

$$\int_0^{2\pi} \cos(kt) dt = \int_0^{2\pi} \sin(kt) dt = 0$$

tret de $k \neq 0$. Si $k = 0$, la primera integral és igual a 2π , i la segona a 0. Així, el terme del mig es pot recuperar integrant, i és

$$\frac{1}{2\pi} \int_0^{2\pi} \left(e^{-t\sqrt{-1}} + e^{t\sqrt{-1}} \right)^n dt = \frac{2^n}{2\pi} \int_0^{2\pi} \cos^n(t) dt.$$

Altres termes (per exemple el k -èssim) es poden recuperar multiplicant la funció de generació per

$$s^{-k} = e^{-kt\sqrt{-1}}$$

abans d'integrar.

Quan es publica *Théorie analytique des probabilités* [17], aquesta eina ja està completament desenvolupada, i combinada amb la capacitat inigualable de Laplace per derivar aproximacions asimptòtiques a integrals, aconseguix demostrar que sumes o mitjanes força generals presenten distribucions que s'aproximen correctament a la corba normal. Així doncs, el matemàtic desenvolupa un aparell molt extens per l'anàlisi de distribucions binomials, a part d'idear alguns principis per escollir la mitjana d'entre varies observacions, tot i que aquí la seva desafortunada tria de les corbes d'error limita l'ús pràctic dels mateixos.

El 1809 Gauss proporciona la clau. No obstant, Laplace encara no coneixia els resultats de Gauss quan publicà els seus desenvolupaments el 1810. El treball de Gauss en aquest àmbit és curt, però els resultats són extraordinaris, ja que amb només dos anys aconseguix una síntesi remarcable.

4 La síntesi de Gauss-Laplace

Després de la publicació del mètode dels mínims quadrats de Legendre [22] de 1805, es podria considerar que s'inicia el desenvolupament de l'Estadística matemàtica. Aquest proposa un mètode fàcil i intuïtiu per combinar observacions, amb una aplicació relativament simple i directa a varis problemes astronòmics i geodèsics. No obstant, vist des del punt de vista modern, manca una consideració formal de la probabilitat i la corresponent relació amb els mínims quadrats, imprescindible per poder estimar la precisió dels resultats a l'usar aquest procediment.

El mètode dels mínims quadrats dona resultats que es podien anomenar “millors”, ja que minimitzen la suma dels errors al quadrat; però sense conèixer la natura estocàstica de les observacions, la quantificació de la incertesa no és possible, de manera que no es pot donar resposta a la qüestió: Què tan bo és “millor”?

4.1 Carl Friedrich Gauss

Carl Friedrich Gauss (Braunschweig, 1777 - Göttingen, 1855) va ser un gran matemàtic, astrònom, geodèsic i físic alemany. És considerat un dels matemàtics més influents de la història, ja que va fer grans contribucions en el camp de les matemàtiques i la ciència. Se'l coneix amb el sobrenom de *Princeps mathematicorum* (del llatí “el príncep dels matemàtics”). Durant gairebé cinquanta anys (des de 1807 fins a la seva mort), va ser professor d'astronomia i director de l'observatori de Göttingen.

Durant la seva estada a la Universitat de Göttingen va proposar diversos teoremes matemàtics. Va escriure dues obres mestres *Disquisitiones Arithmeticae* el 1798 i *Theoria motus corporum coelestium* [9] el 1809. Va proporcionar la segona i la tercera demostracions completes del teorema fonamental de l'àlgebra, i va contribuir en la teoria de nombres, a més de desenvolupar les teories de les formes quadràtiques binàries i ternàries. Se'l considera un dels pioners en la geometria no euclidiana. En l'àmbit de l'astronomia, va desenvolupar la constant gravitatòria gaussiana i el mètode dels mínims quadrats.

Va fer una gran part del treball de camp i de l'avaluació científica del projecte de mesura de l'arc des de 1820 fins el 1844. A més, va ser un dels fundadors de la geofísica i va formular els principis fonamentals del magnetisme, camp en el qual va realitzar bastantes investigacions pràctiques. De les seves investigacions sorgiren les invencions de l'heliotrop (1821), un magnetòmetre (1833) i el primer telègraf electromagnètic (1833), aquest últim juntament amb Wilhelm Eduard Weber.

Moltes de les obres de Gauss van ser editades i publicades després de la seva mort.

4.1.1 Gauss el 1809

Actualment, el mètode dels mínims quadrats està associat a una tècnica consistent en obtenir l'equació d'una corba que, sota algun criteri determinat, s'“apropi” o s'“adapti” el millor possible als punts observats d'una distribució bidimensional. No obstant, en els seus inicis, el mètode dels mínims quadrats consistia en una tècnica geodèsica, creada per resoldre sistemes d'equacions en els quals el número d'equacions fos major que el d'incògnites.

El paper de l'astronomia ha sigut fonamental en el desenvolupament del mètode dels mínims quadrats. Legendre publica aquest mètode, consistent en minimitzar la suma d'errors al quadrat, l'any 1805, en un apèndix del seu llibre sobre l'òrbita dels cometes *Nouvelles méthodes pour la détermination des orbites des comètes* [22].

Els astrònoms s'enfronten al problema de calcular les posicions dels cossos celestes a partir de poques observacions, i és en aquest moment quan Gauss desenvolupa un sistema per calcular les òrbites, el qual és un mètode basat en el dels mínims quadrats, que s'usa actualment.

Aquesta pròxima etapa s'inicia amb la publicació de *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium* [9] (“Teoria del moviment dels cossos celestes que es mouen al voltant del sol en seccions còniques”) l'any 1809, a mans del matemàtic Carl Friedrich Gauss. El tema principal del llibre és una investigació sobre les matemàtiques amagades darrere les òrbites planetàries, on desputen conceptes com la distribució normal o la “lleï dels errors”. Apareix també un apartat que tracta la combinació d'observacions, que aborda essencialment el mateix problema que Laplace i Legendre havien estudiat, però amb una diferència significativa, ja que Gauss planteja la

qüestió en termes explícitament probabilístics.

Suposa que existeixen μ funcions lineals desconegudes ($V, V', V'', \text{etc.}$),

$$\begin{aligned} V &= ap + bq + cr + ds + \dots \\ V' &= a'p + b'q + c'r + d's + \dots \\ V'' &= a''p + b''q + c''r + d''s + \dots \\ &\text{etc.} \end{aligned}$$

on $a, b, c, \dots$ són observables, i $p, q, r, \dots$ són v coeficients desconeguts que es troben a partir de l'observació directa, i $M, M', M'', \text{etc.}$ són els valors observats d'aquestes funcions. Dedueix que aquests valors estan afectats dels errors ($\Delta, \Delta', \Delta'', \text{etc.}$):

$$\begin{aligned} \Delta &= V - M \\ \Delta' &= V' - M' \\ \Delta'' &= V'' - M'' \\ &\text{etc.} \end{aligned}$$

les probabilitats dels quals venen donades per la corba $\varphi(\Delta), \varphi(\Delta'), \varphi(\Delta''), \text{etc.}$

Gauss usa resultats d'obres anteriors, en particular de Laplace, per formular els seus desenvolupaments. Escull un paradima amb més equacions lineals que incògnites, i usa l'adaptació de Laplace del teorema de Bayes. Inicialment, suposa els valors de les incògnites igualment probables, fet que el porta a triar aquells valors per $p, q, \text{etc.}$ que maximitzen

$$\Omega = \varphi(\Delta)\varphi(\Delta')\varphi(\Delta'')\dots$$

com a sistema de valors més probable. Aquests valors es poden trobar derivant Ω respecte a $p, q, r, \text{etc.}$, igualant els resultats a zero i resolent les equacions simultànies. No obstant, aquestes equacions involucren la corba d'error $\varphi(\Delta)$, de manera que primerament s'ha de determinar una fórmula per la mateixa.

Fins aquí, els resultats de Gauss no discerneixen massa dels que Laplace havia fet per a mesures simples. Aleshores, el primer inverteix el procés que va seguir el darrer per aproximar la corba d'error, el qual va començar amb el principi de raó insuficient, raonant a una corba, i després desenvolupa un mètode per combinar observacions.

Gauss nota que, a priori, només pot fer afirmacions generals sobre $\varphi(\Delta)$. Considera que és simètrica; com a màxim és $\Delta = 0$; i fora del rang de valors possibles és zero. Enlloc d'imposar condicions addicionals, assumeix la conclusió. Pren com a axioma que el valor més probable d'una única incògnita observada diverses vegades sota les mateixes circumstàncies és la mitjana aritmètica de les observacions. Llavors, demostra que en aquest cas, on $V = V' = V'' = \dots = p$, prenent

$$p = \frac{1}{\mu}(M + M' + M'' + \dots)$$

es maximitza Ω només si

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

per alguna constant positiva h , on h correspon a una mesura de la precisió de l'observació. Aleshores, mostra que una situació més general d'aquesta distribució d'error porta al mètode dels mínims quadrats, proporcionant valors per $p, q, r, \text{etc.}$ que maximitzen Ω .

Així doncs, a partir de la llei de probabilitat d'ocurrència de l'error, que és

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

obté que la probabilitat de que succeeixin els μ errors alhora és

$$P = \left(\frac{h}{\sqrt{\pi}} \right)^p e^{-h^2(\Delta^2 + \Delta'^2 + \Delta''^2 + \dots)}.$$

Com que Gauss ignora on es produeixen aquests errors, pensa en substituir-los allà on la probabilitat anterior proporcioni el seu valor màxim, és a dir, allà on la suma d'errors $(\Delta^2 + \Delta'^2 + \Delta''^2 + \dots)$ sigui mínima.

Vist en perspectiva, es poden veure algunes incongruències en el raonament de Gauss, ja que és un argument circular i confús. Esquemàticament, el matemàtic segueix els següents passos: Primer, la mitjana aritmètica (un cas singular del mètode dels mínims quadrats) és la més probable només en el cas on els errors estiguin distribuïts normalment; seguidament, la mitjana aritmètica és una forma excel·lent de combinar observacions tal que els errors es prenguin distribuïts normalment; finalment, la suposició que els errors estan distribuïts normalment retorna als mínims quadrats. El mateix Gauss, revisant aquest argument anys més tard, considera que aquest és difícil d'acceptar.

L'obra de Gauss arriba a París el maig de 1809. A l'abril de 1810, Laplace presenta la seva memòria a l'Acadèmia, la qual conté el teorema central del límit. Sembla ser que Laplace no hauria llegit el treball de Gauss quan es va conjecturar, ja que posa molt èmfasi en l'anàlisi, i no hi ha prova de que Laplace hagués pogut deduir que e^{-t^2} pogués ser una possible corba d'error. S'aconsegueix un resultat important, però les úniques aplicacions d'aquest s'atribueixen a mitjanes simples, com la inclinació mitjana de les òrbites dels cometes. A finals del 1810 tot això canvia.

Laplace troba el treball de Gauss poc després de l'abril de 1810, i adverteix que la derivació d'aquest no té sentit. És en aquest moment quan Laplace troba la connexió entre el teorema del límit i l'estimació lineal, de manera que usa el teorema central del límit com a alternativa a la derivació de Gauss. Reformula el teorema de pas al límit i mostra com podria proporcionar una justificació més coherent per l'elecció de Gauss de $\varphi(\Delta)$ com a corba d'error: Si els errors que considera Gauss en la seva formulació fossin agregats, aleshores el teorema de pas al límit implica que han d'estar distribuïts aproximadament com una corba normal, o Gaussiana. Un cop determinada la corba d'error, tot el desenvolupament de mínims quadrats encaixa de forma correcta, tal i com Gauss havia demostrat.

En retrospectiva, la tria d'una corba d'error pot semblar una qüestió minoritària del problema, però la lentitud de progrés mostra que de fet va ser central. El 1809 Gauss va més enllà de considerar que aquesta corba d'error condueix als mínims quadrats, explora conseqüències remarcables i desenvolupa formes de simplificar l'ús del mètode. Nota que, si els errors es distribueixen en funció de la corba $\varphi(\Delta)$, aleshores les incògnites p, q, r , etc., segueixen una distribució proporcional a

$$e^{-h^2 W},$$

on

$$W = \Delta^2 + \Delta'^2 + \Delta''^2 + \dots.$$

Procedeix a mostrar com aquest fet es pot usar per avaluar la precisió del mètode. Determina les distribucions marginals a posteriori de les incògnites (per separat) i la seva precisió corresponent (en notació actual correspon a $1/(\sqrt{2} \cdot \text{desviació estàndard})$). També mostra com els procediments descrits es poden generalitzar a mesures $M, M', M'', \text{etc.}$ fetes amb precisions desiguals (però conegudes) i a funcions desconegudes $V, V', V'', \text{etc.}$ que no són lineals en $p, q, r, \text{etc.}$ (només localment lineals). Finalment, discuteix un algorisme d'eliminació per accelerar el càlcul i resoldre simultàniament les incògnites $p, q, r, \text{etc.}$

El desenvolupament de Gauss té altres limitacions, a part de la seva tria enganyosa per la corba d'error. Per estimar les incògnites, l'autor només esmenta precisió relativa. El coeficient desconegut p s'estima amb una precisió de 4,96 vegades la d'una única mesura (h igual a la unitat). Manca conèixer la incertesa de l'estimació de p en la seva pròpia escala de mesura, i tampoc s'explica com determinar la precisió desconeguda h d'una única mesura.

El principi de Legendre dels mínims quadrats era considerada una eina essencial de l'estadística matemàtica, i Gauss es referia al mateix com el “nostre principi”, reivindicant que ell ja usava el mètode des de 1795. Això va originar que Legendre i Gauss entressin en una disputa important.

En el suplement de Laplace de 1810, limita el seu estudi a la probabilitat inversa, mostra com el seu teorema del límit pot millorar l'argument de Gauss per prendre la corba normal com a distribució d'error, i utilitza el seu propi teorema de 1774 per reiterar que els mínims quadrats proporcionen les estimacions més precises (en el sentit de l'error esperat posterior més petit).

Posteriorment, el 1811, Laplace publica una llarga memòria en la qual usa el teorema central del límit d'una forma relativament diferent, una obra que es repeteix en gran part en el llibre publicat el 1812, *Théorie analytique des probabilités* [17]. L'autor comença amb una situació simple, prenent una aproximació propera en la qual els errors es relacionen amb les observacions per

$$\epsilon^{(i)} = p^{(i)}z - \alpha^{(i)}.$$

En aquesta notació, $p^{(i)}$ i $\alpha^{(i)}$ venen donades per l'observació, z és el coeficient desconegut a determinar, i els $\epsilon^{(i)}$ són errors. Aquí, Laplace s'allunyava de Gauss evitant qualsevol suposició de les distribucions d'errors, i prenent els errors distribuïts de forma simètrica respecte a zero. Nota que si una funció lineal d'errors $\sum m^{(i)}\epsilon^{(i)}$ és igual a zero, aleshores com que se satisfà

$$\sum m^{(i)}\epsilon^{(i)} = z \sum m^{(i)}p^{(i)} - \sum m^{(i)}\alpha^{(i)} \quad (4.1)$$

un cop especificats els pesos $m^{(i)}$, es pot solucionar l'equació (4.1) per trobar

$$z = \frac{\sum m^{(i)}\alpha^{(i)}}{\sum m^{(i)}p^{(i)}}.$$

Suposa que els pesos són enters positius o negatius. Aquesta aproximació li dona una estimació per a cada concreció dels pesos. Aquesta és essencialment una formulació algebraica de la forma en què Laplace generalitza el mètode de Mayer: Donat un sistema amb més equacions inconsistentes que incògnites, prendre tantes combinacions lineals de l'equació com incògnites hi hagi i resoldre per les incògnites. En aquest cas, procedeix amb una sola incògnita, i a més assumeix que els pesos $m^{(i)}$ són tots nombres sencers;

per tant, la suma és només un simple agregat de les equacions originals d'observació (com en el cas del mètode de Mayer), però amb l'equació i -èssima comptada $m^{(i)}$ vegades en l'agregació (semblant al mètode de ponderació que usà pel tractament de les longituds d'arc a l'obra *Mécanique Céleste* [21] del mateix Laplace).

Fins aquí, Laplace pensa de forma probabilística; els errors s'han de prendre com a aleatoris, distribuïts simètricament al voltant de zero. Com que no s'obté el valor exacte

$$\sum m^{(i)}\epsilon^{(i)} = 0$$

ja que difereix per una quantitat aleatòria, la relació anterior no és exactament igual a z , sinó que l'error equival a una quantitat u :

$$z = \frac{\sum m^{(i)}\alpha^{(i)}}{\sum m^{(i)}p^{(i)}} + u$$

on

$$u = \frac{\sum m^{(i)}\epsilon^{(i)}}{\sum m^{(i)}p^{(i)}}.$$

Ara, per especificar els pesos, Laplace usa una petita extensió del seu teorema central del límit per mostrar que l'error u és (si el número d'equacions és gran) aproximadament normal i, per aquesta distribució limitant, l'error esperat és proporcional a

$$\frac{\sqrt{\sum m^{(i)2}}}{\sum m^{(i)}p^{(i)}}.$$

A més, la probabilitat de qualsevol interval simètric al voltant de zero varia inversament amb aquesta quantitat.

Aleshores, resol que aquesta quantitat és la més petita si els $m^{(i)}$ són proporcionals als $p^{(i)}$, però aquesta només correspon a l'aproximació dels mínims quadrats,

$$\frac{\sum p^{(i)}\alpha^{(i)}}{\sum p^{(i)2}}.$$

En notació moderna, la conclusió de Laplace es resumeix de la següent manera: Totes les estimacions de z que són funcions lineals de l'observable $\alpha^{(i)}$ (on els coeficients $\alpha^{(i)}$ poden dependre de $p^{(i)}$) presenten, aproximadament, una distribució normal, i dins d'aquesta classe l'estimació de mínims quadrats té l'error esperat més petit.

Generalitza aquesta conclusió a problemes d'estimació lineal amb diverses incògnites, amb un argument que deriva la distribució normal multivariant de dos o més estimacions de mínims quadrats. Aquestes deduccions, tot i que es basen en arguments asimptòtics, no suposen hipòtesis estrictes sobre la corba d'error. Laplace fins i tot suggereix de nou que la corba

$$\frac{1}{2a} \log \left(\frac{a}{|x|} \right),$$

derivada el 1777, es podria considerar com una possible corba d'error.

Per treballar amb un gran nombre d'equacions, el mètode dels mínims quadrats proporciona una major precisió, ja que s'aconsegueix l'error esperat més petit i el més probable

d'estar a prop de la quantitat estimada. Més tard, s'hi referiria com “el mètode més avantatjós”.

El desenvolupament, lluny d'estar acabat, conforma una síntesi remarcable gràcies a la complexitat tècnica que presenta. No obstant, encara s'havien d'idear mètodes per determinar la precisió de les estimacions i algoritmes per simplificar el càlculs. A la secció 4.2, es presenta una aplicació dels resultats de l'últim treball de Laplace, publicat el 1827, després de la mort del matemàtic.

4.1.2 Exemple del mètode de mínims quadrats

Per fer el raonament anterior més entenedor, i sense perdre generalitat, considerem un sistema de $\mu = 3$ equacions amb $\nu = 2$ incògnites. Aleshores,

$$\begin{aligned} V &= ap + bq + c \\ V' &= a'p + b'q + c' \\ V'' &= a''p + b''q + c'' \end{aligned}$$

on les “equacions de condició” equivalen a

$$\begin{aligned} \Delta &= V - M = ap + bq + (c - M) = ap + bq + d \\ \Delta &= V' - M' = a'p + b'q + (c' - M') = a'p + b'q + d' \\ \Delta &= V'' - M'' = a''p + b''q + (c'' - M'') = a''p + b''q + d'' \end{aligned}$$

La suma d'errors al quadrat,

$$W = \Delta^2 + \Delta'^2 + \Delta''^2 = (ap + bq + d)^2 + (a'p + b'q + d')^2 + (a''p + b''q + d'')^2$$

ha de ser mínima, de manera que s'han de satisfer les següents condicions:

$$\begin{aligned} \frac{\partial W}{\partial p} &= 2 [(ap + bq + d)a + (a'p + b'q + d')a' + (a''p + b''q + d'')a''] = 0 \\ \frac{\partial W}{\partial q} &= 2 [(ap + bq + d)b + (a'p + b'q + d')b' + (a''p + b''q + d'')b''] = 0 \end{aligned}$$

Simplificant les expressions anteriors s'obté el sistema d'equacions normals que es mostra tot seguit

$$\begin{aligned} \left(\sum a^i{}^2 \right) p + \left(\sum a^i b^i \right) q + \sum a^i d^i &= 0 \\ \left(\sum a^i b^i \right) p + \left(\sum b^i{}^2 \right) q + \sum b^i d^i &= 0 \end{aligned}$$

Aquest sistema ja té el mateix número d'incògnites que d'equacions, i els errors han desaparegut. A més, es compleix que

$$\begin{aligned} \sum a^i b^i &= \frac{1}{2} \left[\sum (a^i + b^i)^2 - \sum a^i{}^2 - \sum b^i{}^2 \right] \\ \sum a^i d^i &= \frac{1}{2} \left[\sum (a^i + d^i)^2 - \sum a^i{}^2 - \sum d^i{}^2 \right] \\ \sum b^i d^i &= \frac{1}{2} \left[\sum (b^i + d^i)^2 - \sum b^i{}^2 - \sum d^i{}^2 \right] \end{aligned}$$

de manera que tots els coeficients del sistema d'equacions normals o bé són sumes de quadrats o bé es redueixen a sumar i restar aquestes sumes de quadrats (les quals es troben en les taules dels llibres de text de finals del segle XIX).

4.2 Laplace i les marees de l'atmosfera

A l'obra *Mécanique Céleste* [21], Laplace aplica la teoria gravitatòria per explicar l'efecte de la lluna sobre les marees del mar de forma exitosa. El 1823, intenta aplicar la mateixa teoria a les marees atmosfèriques, amb l'objectiu de mesurar l'efecte de la lluna en la pressió baromètrica a París, comparant les variacions diàries de pressió en cada fase lunar.

La lluna exerceix una gran influència sobre el mar, de manera que Newton i científics posteriors creien que havia d'existir un efecte similar sobre l'atmosfera. Laplace vol demostrar l'existència d'aquest efecte, que pot ser degut directament per l'atracció gravitatòria o indirectament per l'elevació i disminució del nivell del mar. La seva teoria de les marees afirma que la magnitud de l'efecte es pot mesurar comparant els canvis en les pressions baromètriques entre les 9:00 i les 15:00 durant els quatre dies que envolten les sizígies (els dos dies de cada mes quan la lluna, la terra i el sol estan alineats) amb els canvis diaris de pressió durant els quatre dies que envolten les quadratures (els dos dies de cada mes quan la lluna, terra i sol formen un angle recte).

Usa les dades de l'observatori de París, corresponents a vuit anys de mesures baromètriques preses tres vegades al dia (a les 9:00, a les 12:00 i a les 15:00). Laplace redueix la seva teoria a un sistema d'equacions lineals:

$$x \cos(2iq) + y \sin(2iq) = E_i$$

$$y \cos(2iq) - x \sin(2iq) = F_i$$

per $i = -1, 0, 1, 2$. x i y són quantitats desconegudes, a estimar a partir de les dades; q és una quantitat coneguda, corresponent al moviment sinòdic de la lluna; i E_i i F_i (l'índex i representa el dia de la fase lunar) es calculen a través de les dades com segueix:

$$E_i = A_i'' - A_i + B_i - B_i'' \quad (4.2)$$

$$F_i = (2A_i' - (A_i + A_i'') - 2B_i' + (B_i + B_i'')) \left(1 + \frac{1}{19}\right) \quad (4.3)$$

A_i és la mitjana de les mesures de la pressió baromètrica de les 9:00 per l' i -èssim dia després d'una sizígia, presa al llarg de vuit anys; A_i' és la mateixa mitjana per les mesures de les 12:00, i A_i'' per les mesures de les 15:00. B_i , B_i' i B_i'' representen les mesures mitjanes corresponents a l' i -èssim dia després de la quadratura. Per tant, E_i és la variació baromètrica diària mitjana (de les 9:00 a les 15:00) per l' i -èssim dia després de la sizígia menys el canvi mitjà de l' i -èssim dia després de la quadratura. F_i és proporcional a la diferència entre les taxes de variació mitjanes, o diferències segones, pels mateixos dos dies.

Aleshores, Laplace assumeix que les desviacions, degudes a “causes irregulars”, de les mesures diàries segueixen una mateixa llei, no necessàriament normal. Tractant les diferents mesures com si s'haguessin fet independentment, aplica un mètode d'anàlisi (actualment conegut com a mínims quadrats ponderats) presentat en el tercer apèndix del 1820 de *Théorie analytique des probabilités* [17], que denomina com “el mètode més avançatjós”. Difereix del mètode dels mínims quadrats de Legendre, considerat per Laplace com un cas especial d'aquest, ja que en aquest les observacions (E_i i F_i) s'han de calcular com a combinacions lineals d'altres observacions independents, cada una amb una distribució singular. Aquesta estructura llavors s'usa per trobar el millor mètode de combinar les observacions, mitjançant una apel·lació al seu teorema del límit, tal i com havia fet a la seva memòria de 1811.

Un cop aplicat aquest mètode per calcular x i y , Laplace avalua la probabilitat que aquestes observacions indiquin una marea lunar. Primer, nota que no n'hi ha prou amb combinar un gran nombre d'observacions de la forma més avantatjosa, sinó que s'ha d'aconseguir un mètode per determinar la probabilitat en què l'error dels resultats obtinguts estigui contingut en límits estrets.

Per comprovar la hipòtesi de que els canvis baromètrics no estan influenciats per la fase lunar, compara la variació mitjana durant 792 dies a prop de les sizígies amb el canvi mitjà durant els mateixos dies a prop de les quadratures. En base a l'aproximació donada pel teorema central del límit, quan no hi ha una diferència regular en els canvis baromètrics, l'atzar és l'únic capaç de produir una dissemblança en les mitjanes, que no pot ser superior a la realment observada amb probabilitat 0,843. Laplace afirma que aquest nombre no és prou gran per confirmar l'existència d'una diferència real, i calcula que si l'efecte lunar és de la mida estimada, aleshores calen nou vegades més dades per confirmar la seva existència. És a dir, Laplace afirma que la diferència observada entre mitjanes és significativament al nivell 0,01 només si està basada en 72 anys de dades.

Aquest document és important com a exemple de l'aplicació de l'estadística a un problema científic. No obstant, és rellevant destacar que Laplace comet tres errors subtils però importants.

Primer, Laplace suposa que les diferents lectures baromètriques són independents, sense considerar que les lectures d'un mateix dia estan força correlacionades. Tot i això, l'autor observa que les variacions de dies diferents són molt majors que les variacions d'un mateix dia, fet que el porta a formular la seva teoria en termes de canvis diaris més que en termes de lectures aïllades. Tot i ser conscient d'aquesta correlació, la ignora en l'aplicació del seu "mètode més avantatjós", una omisió que el porta a ponderar incorrectament les observacions.

Si Laplace hagués tingut en compte aquesta correlació, potser hauria tractat les variàncies de E_i i F_i amb la proporció 3 : 1 enlloc de 1 : 3, que és la que realment va emprar. L'ús de 1 : 3 per la proporció de la variància de E_i amb F_i negligeix el factor $1 + \frac{1}{19}$, i es basa en les definicions (equacions (4.2) i (4.3)) i les suposicions que totes les A i les B són independents, amb la mateixa variància. És a dir, assumint que les mesures preses en diferents instants del dia són independents i igualment variables, les variàncies són:

$$\begin{aligned}\text{Var}(E_i) &= \text{Var}(A_i'' - A_i + B_i - B_i'') \\ &= \text{Var}(A_i'') + \text{Var}(A_i) + \text{Var}(B_i) + \text{Var}(B_i'') \\ &= 4\sigma^2\end{aligned}$$

i també

$$\begin{aligned}\text{Var}\left(\left(1 + \frac{1}{19}\right)(2A_i' - (A_i + A_i'') - 2B_i' + (B_i + B_i''))\right) \\ &= \left(1 + \frac{1}{19}\right)^2 (4\text{Var}(A_i') + \text{Var}(A_i) + \text{Var}(A_i'') + 4\text{Var}(B_i') + \text{Var}(B_i) + \text{Var}(B_i'')) \\ &= \left(1 + \frac{1}{19}\right)^2 12\sigma^2.\end{aligned}$$

Exceptuant el factor $(1 + \frac{1}{19})^2$, aquests valors es troben en la proporció 1 : 3. D'altra banda, si es té en compte la correlació mencionada,

$$U_i = A_i'' - A_i' \quad \text{i} \quad V_i = A_i' - A_i$$

aleshores

$$U_i + V_i = A_i'' - A_i \quad \text{i} \quad V_i - U_i = 2A_i' - (A_i + A_i'').$$

Una anàlisi de dades meteorològiques de París del 1823 suggereix que la correlació de U_i i V_i no s'allunya massa de $1/2$, un valor que dona la proporció $3 : 1$, negligint de nou el factor $1 + \frac{1}{19}$. Per tant, sigui

$$\text{Var}(U_i) = \text{Var}(V_i) = \tau^2$$

aleshores

$$\begin{aligned} \text{Var}(U_i + V_i) &= \text{Var}(U_i) + \text{Var}(V_i) + 2\text{corr}(U_i, V_i)\sqrt{\text{Var}(U_i)\text{Var}(V_i)} \\ &= \tau^2 + \tau^2 + 2 \cdot \frac{1}{2} \cdot \tau^2 = 3\tau^2 \end{aligned}$$

i similarment,

$$\text{Var}(U_i - V_i) = \tau^2 + \tau^2 - 2 \cdot \frac{1}{2} \cdot \tau^2 = \tau^2$$

donant la proporció $3 : 1$.

El segon error de Laplace es troba en el seu test de significació, en què compara els dos canvis de mitjana. Si un assumeix que les lectures fetes en dies diferents són independents amb variàncies iguals, l'ús del teorema central del límit és incorrecte. L'autor estima la variància agrupant les dues mostres com una de sola, en comptes d'agrupar dues estimacions de la mostra, una de les quals basada en els canvis propers a les sizígies i l'altra en els canvis propers a les quadratures. Així, Laplace calcula la variància suposant que la hipòtesi nul·la (la qual s'ha construït per refutar, amb l'objectiu de donar suport a una hipòtesi alternativa) és correcta. No obstant, si aquesta fos falsa, usant la suma total dels quadrats, l'estimació de la variància tendiria a créixer més del compte.

El tercer error de Laplace consisteix en suposar que les vuit equacions són independents, cosa qüestionable pel fet que tant E_i com F_i involucren les mateixes mesures. No obstant, el veritable efecte d'aquest error és petit, ja que E_i inclou diferències de canvis i F_i segones diferències, de manera que E_i i F_i no estan correlacionades si la variància de l'alteració des de les 9:00 fins a les 12:00 és igual a la variació de les 12:00 a les 15:00, cosa que no està lluny de la realitat. Laplace percep aquest possible entrebanc més tard, i apareix en el seu últim treball, publicat després de la seva mort, en el qual recupera aquest problema i repeteix l'anàlisi de 1823, però amb una diferència important.

En el tercer additament de *Théorie analytique des probabilités* [17], Laplace presenta l'equació

$$l^{(s)}x + p^{(s)}y + q^{(s)}z + \dots = a^{(s)} + m^{(s)}\gamma^{(s)} + n^{(s)}\lambda^{(s)} + r^{(s)}\delta^{(s)} + \dots \quad (4.4)$$

com a base per una de les “equacions de condició” en situacions en les quals el seu mètode més avantatjós és aplicable. l, p, q, a, m, n , i r venen representades amb els coeficients donats; x, y i z són quantitats a estimar; i γ, λ i δ són observacions independents amb diferents distribucions possibles.

Laplace considera la part dreta de l'equació (4.4) com una observació derivada, calculada a partir d'altres a través de normes conegudes. Tot i la manca d'independència mencionada anteriorment, les equacions de condició ((4.2) i (4.3)) per les marees lunars són d'aquesta forma. A l'obra de Laplace *Mémoire sur le flux et reflux lunaire atmosphérique* [20] de 1827, retorna a aquest problema amb dades addicionals registrades

durant tres anys, i considera un tractament diferent pel problema. Determina els factors que s'han de multiplicar a les diferents equacions de condició per obtenir els resultats “més avantatjats”.

Comença resolent les equacions de condició amb parelles per x i y , agrupa els termes que involucraven les mateixes mesures i aplica el seu mètode més avantatjós, usant només el cas relatiu a una única incògnita. Com que E_i i F_i estan calculades a partir de les mateixes mesures, no poden ser independents. Considerant aquest canvi en la seva anàlisi, Laplace aconsegueix abordar correctament múltiples problemes de regressions lineals amb errors correlacionats i una covariància coneguda.

Laplace repeteix els altres errors previs. Assumeix la independència de mesures diferents i avalua si la seva estimació de x difereix significativament de zero usant una mostra agrupada per valorar la variància dels canvis de pressió diaris. Determina x igual a 0,031758 i calcula la probabilitat que l'atzar produeixi una estimació de 0,3617 entre els límits $\pm 0,031758$. Declara que si aquesta probabilitat s'acosta a la unitat, aleshores és molt probable que el valor de x no es degui únicament a irregularitats de la probabilitat, sinó que existeix una causa constant, la qual només pot ser l'efecte de la lluna sobre l'atmosfera. Però la diferència considerable entre aquesta probabilitat i la certa, representada per la unitat, mostra que, tot i el gran nombre d'observacions usades, l'acció d'aquesta només s'indica amb una probabilitat molt petita; de manera que és incert que la seva existència sigui perceptible a París.

Tot i que Laplace reconeix i té en compte la possible dependència entre observacions, aquest document de 1827 mostra com la comprensió del matemàtic es veu molt limitada i potser lligada al cas en què la dependència és explícita notacionalment.

A l'última secció d'aquesta publicació, investiga l'homogeneïtat de les dades al llarg de l'any i intenta determinar si una diferència notable en els canvis de la mitjana de la pressió baromètrica durant quatre trimestres és realment significativa.

Laplace tracta aquest problema (actualment conegut com a “variància del problema”) mitjançant quatre tests diferents. Primer considera la diferència entre el canvi mitjà de febrer a abril i el canvi global mitjà, corresponent a $0,940 - 0,763 = 0,177$, i avalua la probabilitat que una diferència igual o major a aquesta sigui deguda únicament per l'atzar. Troba el valor 0,0000015815 per aquesta probabilitat i considera molt probable que la discrepància sigui indicativa d'alguna causa “constant”. Repeteix aquest anàlisi per tres trimestres més, i troba que de novembre a gener el canvi mitjà també difereix significativament respecte la mitjana global, però que les discrepàncies entre les mitjanes dels altres trimestres i la mitjana global es pot atribuir a les irregularitats de l'atzar.

Tot i que l'anàlisi de Laplace era correcta en molts aspectes, la derivació de la distribució de la diferència entre aquestes mitjanes es basa en la suposició que són independents. Aquí, com que només presenta les mitjanes numèricament i no usa cap notació per les mateixes, Laplace no reconeix la correlació entre les mitjanes trimestrals i la mitjana global (ja que les primeres determinen la darrera).

Curiosament, les conclusions generals de Laplace en aquestes investigacions semblen ser correctes, tot i els errors en la seva anàlisi. Cal remarcar que París no era el lloc més adient per mesurar la marea atmosfèrica lunar, ja que el seu efecte és extremadament petit (fins el 1945 no es va determinar de forma exitosa la seva magnitud).

4.3 La situació el 1827

Laplace mor el 5 de març de 1827, establint així la fi d'una era. L'obra *Théorie analytique des probabilités* [17] de Laplace passa per tres edicions i quatre suplementes. Gauss publica la seva extensió de la teoria de Laplace el 1823, abandonant el seu argument circular de 1809 per la corba d'error normal. Gràcies a la seva extremada simplicitat, la derivació de Gauss de 1809 dels mínims quadrats es converteix en un element bàsic dels llibres de text fins i tot al segle XX i, per tant, és el treball estadístic de Gauss més difós i influent.

La síntesi de Gauss-Laplace reuneix la combinació d'observacions mitjançant l'agregació d'equacions de condició lineals amb l'ús de la probabilitat matemàtica per avaluar la incertesa i fer inferències. En molts aspectes és un dels èxits més importants en la història de la ciència.

Tanmateix, les aplicacions d'aquesta meravellosa eina estadística es van mantenir confinades a les disciplines que la van generar, convertint-se en habituals en el treball astronòmic i geodèsic, i mantenint-se gairebé desconegudes en les ciències socials, tot i que la seva utilitat en aquestes últimes era coneguda des de l'època de Bernoulli. Aquest fet suggereix que la situació no era tan senzilla com pot semblar a primera vista, ja que restaven grans barreres conceptuals per superar abans que la nova tecnologia pogués estendre's a les ciències socials.

5 Conclusions

Un cop realitzat aquest TFG, es pot considerar que s'han complert amb èxit els objectius inicials, ja que s'ha pogut establir un eix temporal en el qual s'han estudiat els matemàtics més importants del segle XVIII que consolidaren la base de l'estadística i la probabilitat actual, i s'ha fet un desenvolupament exhaustiu sobre els resultats més transcendents de les seves obres. D'aquesta manera, la llei dels grans nombres i el teorema central del límit s'han emprat com a fil conductor del treball, els quals són dos resultats essencials en probabilitat i estadística.

L'obra *Ars Conjectandi* [2] de Jacob Bernoulli, publicada el 1713, introdueix la llei dels grans nombres, demostrant que la mitjana dels resultats obtinguts a partir d'un gran nombre d'assaigs ha de ser propera al valor esperat. Amb aquest treball, s'emfatitza el concepte de convergència i, a més, proporciona una base rigorosa per a l'ús de la probabilitat en la predicció de resultats a llarg termini.

Seguint amb l'obra d'Abraham De Moivre, *The Doctrine of Chances* [4], publicada per primera vegada el 1718, és interessant destacar l'aportació de la mateixa en els desenvolupaments matemàtics en l'àmbit de la probabilitat i l'estadística. De Moivre introdueix la distribució normal, tot i que es refereix a ella com a "aproximació a la distribució binomial". El seu treball sobre el teorema central del límit, que formula d'una forma força rudimentària, demostra que la suma d'un gran nombre de variables aleatòries independents i idènticament distribuïdes tendeix a una distribució normal.

Les observacions astronòmiques motiven a Laplace a indagar en els desenvolupaments de probabilitat inversa, i Thomas Bayes adopta un paper molt important en aquesta revolució copernicana en el marc de l'estadística, amb l'objectiu d'aprendre de l'experimentació amb la natura. Tot i que els seus resultats són incomplets, la "taula de billar" de Bayes pren un paper fonamental en els futurs desenvolupaments a mans d'altres matemàtics.

Posteriorment, el 1812, Pierre-Simon Laplace publica la seva gran obra mestra *Théorie analytique des probabilités* [17]. Aquesta consolida i amplia els treballs anteriors, proporcionant un tractament exhaustiu de la teoria de la probabilitat. Laplace introdueix funcions generadores i formalitza el mètode dels mínims quadrats. També aplica la seva teoria de la probabilitat a problemes d'astronomia, física i demografia, demostrant així la seva àmplia aplicabilitat. És remarcable destacar les dificultats i els errors amb els quals s'enfronta Laplace, com serien els deguts a la manca de coneixement de la notació escaient de probabilitat condicional. Finalment, amb l'obra de Carl Friedrich Gauss de 1809, *Theoria motus corporum coelestium* [9], es desenvolupa un mètode per a calcular òrbites, el qual es basa en el dels mínims quadrats. Juntament amb els resultats de Laplace, s'acaben de consolidar els resultats anteriors i es desenvolupa la base de l'estadística actual.

En conclusió, cada un dels autors tractats en aquest treball ha sigut un pilar fonamental per tal de poder construir les bases de l'estadística moderna, destacant la importància de cada un d'ells en tant que els resultats i publicacions de cada autor anterior va ser cabdal per poder arribar a procediments i descobriments més exactes per part dels matemàtics posteriors. A més, al llarg del treball es fa evident que alguns progressos es van veure alentits degut a les dificultats amb les quals es van trobar els matemàtics per trobar una corba d'error adient. Així doncs, fou gràcies a aquesta cadena d'autors i l'evolució del seus coneixements que va ser possible arribar als magnífics resultats en el marc de l'estadística i la probabilitat, com serien la llei dels grans nombres i el teorema central del límit.

Referències

- [1] Bayes, T.: An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* (1763, 53: 370-418), 1764.
- [2] Bernoulli, J.: *Ars Conjectandi*, Basilea, 1713.
- [3] Chung, K. L.: *Teoría Elemental de la Probabilidad y de los Procesos Estocásticos*, 3a Edició, Editorial Reverté, S. A., 1983.
- [4] De Moivre, A.: *The Doctrine of Chances*, 1a Edició, Londres, 1718.
- [5] De Moivre, A.: *Annuities upon Lives*, Londres, 1725.
- [6] De Moivre, A.: *Miscellanea Analytica de Seriebus et Quadraturis*, Londres, 1730.
- [7] De Moivre, A.: *The Doctrine of Chances*, 2a Edició, Londres, 1738.
- [8] De Moivre, A.: *The Doctrine of Chances*, 2a Edició, Londres, 1756.
- [9] Gauss, C. F.: *Theoria motus corporum celestium*, Hamburg, 1809.
- [10] Lagrange, J. L.: Mémoire sur l'utilité de la méthode de prendre le milieu entre les résultats de plusieurs observations; dans lequel on examine les avantages de cette méthode par le calcul des probabilités; & ou l'on resoud diferens problèmes relatifs à cette matière. *Miscellanea Taurinensia* 5: 167-232, 1776.
- [11] Laplace, P.S.: Mémoire sur la probabilité des causes par les évènements. *Mémoires de l'Académie Royale des Sciences présentés par divers savans* 6: 621-56, 1774.
- [12] Laplace, P.S.: Recherches sur le milieu qu'il faut choisir entre les résultats de plusieurs observations, pp. 228-256, 1777.
- [13] Laplace, P.S.: Mémoire sur les probabilités. *Mémoires de l'Académie Royale des Sciences de Paris* (1778, pp. 227-332), 1781.
- [14] Laplace, P.S.: Mémoire sur les suites. *Mémoires de l'Académie Royale des Sciences de Paris* (1779, pp. 207-309), 1782.
- [15] Laplace, P.S.: Mémoire sur les approximations des formules qui sont fonctions de très grands nombres. *Mémoires de l'Académie Royale des Sciences de Paris* (1782), 1785.
- [16] Laplace, P.S.: *Traité de mécanique céleste*, Vols. 1-4, París, 1799-1805.
- [17] Laplace, P.S.: *Théorie analytique des probabilités*, París, 1812.
- [18] Laplace, P.S.: *Essai philosophique sur les probabilités*, París, 1814.
- [19] Laplace, P.S.: *Traité de mécanique céleste*, Vol. 5, París, 1825.
- [20] Laplace, P.S.: Mémoire sur le flux et reflux lunaire atmospherique, 1827.
- [21] Laplace, P.S.: *Mécanique Céleste*, Vols. 1-4, Boston, 1829-1839.
- [22] Legendre, A. M.: *Nouvelles méthodes pour la détermination des orbites des comètes*, París, 1805.

- [23] Simpson, T.: *The Nature of Laws of Chance*, London, 1740.
- [24] Simpson, T.: *The Doctrine of Annuities and Reversions*, London, 1742.
- [25] Simpson, T.: A letter to the Right Honorable George Earl of Macclesfield, President of the Royal Society, on the advantage of taking the mean of a number of observations, in practical astronomy. *Philosophical Transactions of the Royal Society of London* 49: 82-93, 1755.
- [26] Stigler, S. M.: *The History of Statistics: The Measurement of Uncertainty before 1900*, Balknap Press, 1986.